BIT

Accurate discretization of poroelasticity without Darcy stability

Stokes–Biot stability revisited

Kent-Andre Mardal[1] · Marie E. Rognes[2] · Travis B. Thompson[2,3]

Abstract

In this manuscript we focus on the question: what is the correct notion of Stokes–Biot stability? Stokes–Biot stable discretizations have been introduced, independently by several authors, as a means of discretizing Biot's equations of poroelasticity; such schemes retain their stability and convergence properties, with respect to appropriately defined norms, in the context of a vanishing storage coefficient and a vanishing hydraulic conductivity. The basic premise of a Stokes–Biot stable discretization is: one part Stokes stability and one part mixed Darcy stability. In this manuscript we remark on the observation that the latter condition can be generalized to a wider class of discrete spaces. In particular: a parameter-uniform inf-sup condition for a mixed Darcy sub-problem is not strictly necessary to retain the practical advantages currently enjoyed by the class of Stokes–Biot stable Euler–Galerkin discretization schemes.

Communicated by Rolf Stenberg.

✉ Travis B. Thompson
thompsont@maths.ox.ac.uk

[1] Department of Mathematics, University of Oslo and Simula Research Laboratory, Fornebu, Norway

[2] Department of Scientific Computing and Numerical Analysis, Simula Research Laboratory, Fornebu, Norway

[3] Mathematical Institute, Oxford University, Oxford, UK

1 Introduction

In this note, we consider a three-field formulation of the time-dependent Biot equations describing flow through an isotropic, porous and linearly elastic medium, reading as: find the elastic displacement u, the Darcy flux z and the (negative) fluid pressure p such that

$$- \operatorname{div} \sigma(u) - \alpha \nabla p = f, \tag{1.1a}$$

$$\frac{1}{\kappa} z - \nabla p = g, \tag{1.1b}$$

$$\alpha \operatorname{div} \partial_t u + \operatorname{div} z - c_0 \, \partial_t p = s, \tag{1.1c}$$

for a given body force f, source s, and given g (typically $g = 0$) over a domain $\Omega \subset \mathbb{R}^d$ ($d = 1, 2, 3$). The expression $\sigma(u)$ denotes the isotropic elastic stress tensor, $\sigma(u) = \mu \varepsilon(u) + \lambda \operatorname{tr} \varepsilon(u)$, where tr is the matrix trace. The material parameters are the elastic Lamé parameters μ and λ, the Biot-Willis coefficient α, the storage coefficient $c_0 \geq 0$ and the hydraulic conductivity $\kappa = K/\mu_f > 0$, in which K is the material permeability, and μ_f is the fluid viscosity. Moreover, ε denotes the (row-wise) symmetric gradient, div is the divergence, ∇ is the gradient, and ∂_t denotes the (continuous) time-derivative.

The three field formulation (1.1a–1.1c) combines one scalar, time-dependent partial differential equation and two, stationary, vector partial differential equations. This combination of time-dependent and time-independent equations can lead to non-trivial issues when considering discretizations of the time derivative; as a result: several splitting scheme approaches have been proposed [7,11,18,21,32]. In this manuscript we will focus on a monolithic approach, namely a straightforward backward Euler scheme, where all unknowns are solved for simultaneously. In the case of monolithic time discretization schemes: robustness with respect to material parameters in spatial discretizations of (1.1) is a central concern and has been the topic of several recent investigations; c.f. e.g. [17,19,20,30]. A notable difficulty, both practically and theoretically, is that the parameter λ may be very large, while κ may be very small. The former corresponds to the (nearly) incompressible regime, while the latter corresponds to the (nearly) impermeable regime. Special care is required in the formulation and analysis of discretizations of (1.1) to retain stability and convergence within these parameter ranges.

Thus far, authors have analyzed mixed discretizations of (1.1) in the nearly-incompressible, and nearly-impermeable parameter regimes separately. For instance, a mixed discretization based on a total-pressure formulation [21–24,28] has been well-studied and addresses the case of $\lambda \to \infty$. In the context of vanishingly small hydraulic conductivity, the concept of a *Stokes–Biot stable* discretization has emerged [17,22,27,30] as a guide for the design of discrete schemes that retain their convergence properties as $\kappa \to 0$.

Remark 1.1 It is worth noting that the term *Stokes–Biot stability* refers, in the contemporary literature, to a particular type of dual inf-sup condition. It does not allude to an interface problem; that is, one should not confuse this term with that of a '*Stokes-Darcy problem*', which refers to coupled Stokes and Darcy flow at an interface.

1.1 An intuition for the Stokes–Biot stability condition

To motivate an intuitive view on the current notion of Stokes–Biot stability, we begin by considering a three-field variational formulation of a related system of (time-independent) equations: find $u \in U$, $z \in W$, and $p \in Q$ such that

$$(\sigma(u), \varepsilon(v)) + (\operatorname{div} v, p) = (f, v) \qquad \forall v \in U, \quad (1.2a)$$

$$\tau \kappa^{-1}(z, w) + \tau(\operatorname{div} w, p) = \tau(g, w) \qquad \forall w \in W, \quad (1.2b)$$

$$(\operatorname{div} u, q) + (\tau \operatorname{div} z, q) - (c_0 p, q) = (\tau s + \operatorname{div} \bar{u} - c_0 \bar{p}, q), \qquad \forall q \in Q, \quad (1.2c)$$

for given $f, g, s, \bar{u}, \bar{p}$ and with $(\cdot, \cdot)$ denoting the standard L^2-inner product over the domain Ω. The continuous formulation (1.2) is representative of the equations resulting from an implicit Euler time discretization of (1.1) with time step $\tau > 0$ and a prescribed set of homogeneous boundary conditions. To continue, set $\tau = 1$; in this case we refer to (1.2) as (a mixed variational formulation of) a steady equation of Biot type; that is, the left-hand side is free of any time derivatives. The system (1.2) forms a generalized saddle-point system which can be informally related [27] to a stand-alone Stokes-like, and stand-alone mixed-Darcy system. For the former, multiply (1.2b) by κ, take $\kappa = c_0 = 0$ and assume $s = \operatorname{div} \bar{u} = 0$. If (u, z, p) solves (1.2) under these conditions, then $z = 0$ almost everywhere and and (1.2) reduces to: find $u \in U$ and $q \in Q$ such that:

$$(\nu \varepsilon(u), \varepsilon(v)) + (\operatorname{div} v, p) = (f, v), \qquad (1.3a)$$

$$(\operatorname{div} u, q) = (\operatorname{div} \bar{u}, q) = 0, \qquad (1.3b)$$

for all $v \in U$ and $q \in Q$, with $\nu = 2\mu$. On the other hand, if $c_0 = 0$ and the solution (u, z, p) to (1.2) satisfies $\operatorname{div} u = 0$ then (z, p) solve the mixed Darcy problem: find $z \in W$ and $p \in Q$ such that

$$\left(\kappa^{-1} z, w\right) + (\operatorname{div} w, p) = (g, w), \qquad (1.4a)$$

$$(\operatorname{div} z, q) = (\tilde{s}, q), \qquad (1.4b)$$

for all $w \in W$ and $q \in Q$ for given $g, \tilde{s}$. These observations hint at a close relationship between the Stokes equations, Darcy equations and the (steady) Biot-like system (1.2). With this background: the Stokes–Biot stability concept [17,22,27,30] introduces two conditions for finite element discretizations $U_h \times W_h \times Q_h$ of (1.1) or (1.2):

(i) the displacement-pressure pairing $U_h \times Q_h$ is a stable pair, in the sense of Babuška-Brezzi [6]), for the incompressible Stokes equations (1.3),

(ii) the flux-pressure pairing $W_h \times Q_h$ is a stable pair for the mixed Darcy problem (1.4).

1.2 The Darcy assumption of Stokes–Biot stability

Stokes–Biot stable discrete schemes should retain their convergence properties even when $\kappa \to 0$. Indeed, a-priori error estimates, in appropriate parameter-dependent

Table 1 Relative approximation errors for the displacement $\|\tilde{u} - u_h\|_1/\|\tilde{u}\|_1$ (top three rows in each table), pressure $\|\tilde{p} - p_h\|/\|\tilde{p}\|$ (middle three rows) and flux $\|\tilde{z} - z_h\|_{\mathrm{div}}/\|\tilde{z}\|_{\mathrm{div}}$ (bottom three rows) for varying κ on a series of uniform meshes $\mathcal{T}_h$ with mesh size h

κ	h					
	1/8	1/16	1/32	1/64	1/128	Rate
(a) $P_2^d \times RT_0 \times DG_0$						
10^0	1.64×10^{-1}	4.45×10^{-2}	1.13×10^{-2}	2.84×10^{-3}	7.11×10^{-4}	2.0
10^{-4}	1.64×10^{-1}	4.45×10^{-2}	1.13×10^{-2}	2.84×10^{-3}	7.11×10^{-4}	2.0
10^{-8}	1.64×10^{-1}	4.45×10^{-2}	1.13×10^{-2}	2.84×10^{-3}	7.11×10^{-4}	2.0
10^0	4.00×10^{-1}	1.03×10^{-1}	5.05×10^{-2}	2.53×10^{-2}	1.26×10^{-2}	1.0
10^{-4}	2.08×10^{2}	1.62×10^{1}	1.13	7.61×10^{-2}	1.34×10^{-2}	2.5
10^{-8}	2.50×10^{2}	2.41×10^{1}	2.52	2.81×10^{-1}	3.51×10^{-2}	3.0
10^0	1.30	1.76×10^{-1}	6.51×10^{-2}	3.18×10^{-2}	1.59×10^{-2}	1.0
10^{-4}	3.06×10^{3}	6.76×10^{2}	1.19×10^{2}	1.65×10^{1}	2.11	3.0
10^{-8}	4.51×10^{3}	2.27×10^{3}	1.31×10^{3}	7.09×10^{2}	3.61×10^{2}	1.0
(b) $P_2^d \times P_1^d \times DG_0$						
10^0	1.64×10^{-1}	4.45×10^{-2}	1.13×10^{-2}	2.84×10^{-3}	7.12×10^{-4}	2.0
10^{-4}	1.64×10^{-1}	4.45×10^{-2}	1.13×10^{-2}	2.84×10^{-3}	7.11×10^{-4}	2.0
10^{-8}	1.64×10^{-1}	4.45×10^{-2}	1.13×10^{-2}	2.84×10^{-3}	7.11×10^{-4}	2.0
10^0	1.51×10^{2}	1.95×10^{1}	5.50	2.65	1.34	1.0
10^{-4}	2.44×10^{2}	2.33×10^{1}	2.42	2.75×10^{-1}	3.52×10^{-2}	3.0
10^{-8}	2.50×10^{2}	2.41×10^{1}	2.52	2.82×10^{-1}	3.56×10^{-2}	3.5
10^0	1.12	1.71×10^{-1}	7.28×10^{-2}	3.62×10^{-2}	1.81×10^{-2}	1.0
10^{-4}	4.56×10^{2}	8.98×10^{1}	1.25×10^{1}	1.31	1.17×10^{-1}	3.5
10^{-8}	4.87×10^{2}	1.24×10^{2}	2.85×10^{1}	6.59	1.57	2.1

The last column rate denotes the order of convergence using for the last two values in each row

The exact solutions $\tilde{u}$, $\tilde{p}$, $\tilde{z}$, defined in Sect. 6, were represented by continuous piecewise cubic interpolants in the error computations

Similar results were obtained for $\kappa = 10^{-2}, 10^{-6}, 10^{-10}$ (data not shown)

$c_0 = 0$. (A): $U_h \times W_h \times Q_h = P_2^d(\mathcal{T}_h) \times RT_0(\mathcal{T}_h) \times DG_0(\mathcal{T}_h)$. (B): $U_h \times W_h \times Q_h = P_2^d(\mathcal{T}_h) \times P_1^d(\mathcal{T}_h) \times DG_0(\mathcal{T}_h)$

norms, have been advanced for both non-conforming [17,20,22] and conforming [30] discretizations of (1.1) or (1.2) satisfying the Stokes–Biot conditions (i) and (ii). Consider a numerical test with two closely-related choices of discrete spaces; the finite element pairings $P_2^d \times RT_0 \times DG_0$ (product space of continuous piecewise quadratic vector fields, lowest order Raviart-Thomas elements and piecewise constants) and $P_1^d \times RT_0 \times DG_0$. The former pairing satisfies conditions (i) and (ii) above (for given $\kappa > 0$), and is observed to converge even for $\kappa \ll 1$, see e.g. [30] or Table 1a. The latter pairing, which violates condition (i), can easily fail to converge when κ is sufficiently small (c.f. [30, Table 2.1] or [27, Section 6]). This numerical observation demonstrates that condition (i), Stokes stability, is indeed an integral player in discretizations of (1.1) that retain their convergence behaviour as $\kappa \to 0$.

The importance of the Stokes stability condition is not surprising from a theoretical perspective. In the early Stokes–Biot literature, condition (i) plays a formative role [30] in showing that Euler–Galerkin discretizations of (1.1) remain inf-sup stable as $\kappa \to 0$. Conversely, the Darcy stability condition (ii) is used to construct a projection that facilitates an a-priori analysis; the condition is not used in the stability argument. This raises the question: is condition (ii) necessary to guarantee convergence as $\kappa \to 0$? This question is important; the Darcy stability condition can easily fail to hold uniformly in $\kappa \ll 1$, thereby placing the previous analytic projection technique on questionable grounds. This observation was implicitly noted by other authors; c.f. for instance [17, Rmk. 5]. More precisely, the continuous mixed Darcy problem (1.4) does not satisfy the Babuska-Brezzi conditions [6] with bounds independent of $0 < \kappa \ll 1$ in the standard $H(\text{div}) \times L^2$ norm. To compensate, permeability-weighted flux and pressure norms, such as e.g. $\kappa^{-1/2} H(\text{div}) \times \kappa^{1/2} L^2$, have been suggested as viable alternatives [17]. However, resorting to a permeability-weighted pressure space is not entirely satisfactory; the relation between (1.2) and the Stokes Eq. (1.3), resulting from $\kappa \to 0$, points at $p \in L^2$ rather than $p \in \kappa^{1/2} L^2$.

Moreover, numerical experiments demonstrate convergence of the pressure in the L^2-norm even for diminishing κ, see e.g. Table 1a for the pairing $P_2^d \times RT_0 \times DG_0$. Conversely, consider the pairing $P_2^d \times P_1^d \times DG_0$ which violates the Darcy condition (ii), for any $\kappa > 0$, and thus does not satisfy the Stokes–Biot stability conditions. However, numerical experiments with this pairing, see Table 1b, show the hallmark of Stokes–Biot stable schemes. That is, they appear stable, with the displacement and pressure errors converging at comparable rates as for $P_2^d \times RT_0 \times DG_0$, for small κ; this behaviour even holds when $c_0 = 0$. These observations call into question the precise role of the Darcy stability assumption in conforming mixed finite element discretizations of (1.1) or (1.2).

1.3 Stokes–Biot stability revisited

In this manuscript, we advance a theoretical point. Namely, that a full Darcy inf-sup assumption is not necessary and can be relaxed; at least in the case of conformal Euler–Galerkin discretizations of (1.1) or (1.2). Instead, we will see is that the following two assumptions are key:

(I) the displacement-pressure pairing $U_h \times Q_h$ is a stable pair for the incompressible Stokes Eq. (1.3); and that
(II) the inclusion div $W_h \subseteq Q_h$ holds.

We return to, and formalize, these *minimal Stokes–Biot stability conditions* in Sect. 3. In practice, the class of minimally Stokes–Biot stable discretizations are a superset of Stokes–Biot stable discretizations; one could then naturally consider dropping the distinction and, instead, viewing Stokes–Biot stability from this alternative point of view. We will show that the relaxed conditions produce schemes that retain their stability and convergence properties, in appropriate norms, as $\kappa \to 0$; motivated by the literature in applied porous-media modeling, we also note that this holds true for applications where $0 \leq c_0 < 1$ is chosen independently of other parameters.

Our primary purpose in this manuscript is theoretical in nature. After introducing the relaxed conditions, we comment on the inf-sup stability and advance an a-priori analysis that does not employ a Galerkin projection technique; thus avoiding either an implicit dependence on κ^{-1} in any projection estimates or the problematic question of uniform inf-sup Darcy stability as $\kappa \to 0$. Unlike some previous endeavors of convergence estimates, we will conduct our estimates in the full norm used [17,30] for the inf-sup stability. In particular, we introduce the norms in which Euler–Galerkin schemes, satisfying the relaxed conditions, are well posed and we show that the corresponding a priori error convergence rates: hold in the limit as $\kappa \to 0$; and coincide with canonically expected rates for well known mixed three-field finite element paradigms (e.g. first order for discretizations using linear or Raviart–Thomas type flux approximations, etc). Our objective, in clarifying these nuanced issues, is to establish a more consistent theory of Stokes–Biot stable schemes and to demonstrate an alternative, but standard, approach for their convergence analysis; such a view may also lead to downstream advances in the design of more efficient numerical schemes. The remainder of this manuscript is organized as follows: Sect. 2 describes basic spaces and notation that will be used throughout; Sect. 3 overviews the current view of Stokes–Biot stability [17,22,27,30]; Sect. 4 introduces a slight relaxation on the Stokes–Biot stable conditions and recalls a well-posedness argument for Euler–Galerkin discrete schemes; Sect. 5 is aimed at a priori estimates for discretizations satisfying the relaxed conditions; finally, Sect. 6 is a numerical example demonstrating the retention of convergence behaviour as $\kappa \to 0$.

Remark 1.2 In this manuscript, we are concerned with discretizations that retain their stability and convergence as $\kappa \to 0$. The case of $\lambda \to \infty$ has been investigated separately [21–24,28] by introducing a total pressure, $\hat{p} = \lambda \nabla \cdot u - p$, to achieve robustness with respect to λ when $\kappa \approx 1$ is assumed. This view is similar to Herrmann's method [16], where a 'solid pressure' term $p_s = \lambda \nabla \cdot u$, for elasticity systems in primal form with $\mu \ll \lambda$. One may wonder if these methods can be brought together in a conformal setting. This has not yet been investigated in the literature, and this is not the question we investigate in this manuscript; our current focus is to further the understanding of Stokes–Biot stable discretizations.

2 Notation and preliminaries

2.1 Sobolev spaces and norms

Let $\Omega \subset \mathbb{R}^d$ for $d = 1, 2, 3$ be an open and bounded domain with piecewise C^2 boundary [26,31,35]. We will consider discretizations of Ω by simplicial complexes of order d. All triangulations, $\mathcal{T}_h$ of Ω, will be assumed to be shape regular with the maximal element diameter, also referred to as the mesh resolution or mesh size, of $\mathcal{T}_h$ denoted by h.

We let $L^2(\Omega; \mathbb{R}^d)$, $H(\text{div}, \Omega)$ and $H^1(\Omega; \mathbb{R}^d)$ denote the standard Sobolev spaces of square-integrable fields over Ω, fields with square-integrable divergence, and fields with square-integrable gradient, respectively, and define the associated standard norms

$$\|f\|^2 = (f, f),$$
$$\|f\|_1^2 = (f, f)_1 = (f, f) + (\nabla f, \nabla f),$$
$$\|f\|_{\mathrm{div}}^2 = (f, f)_{\mathrm{div}} = (f, f) + (\mathrm{div}\, f, \mathrm{div}\, f).$$

with $(\cdot, \cdot)_\Omega$ denoting the standard $L^2(\Omega)$-inner product. We will frequently drop the arguments Ω and $\mathbb{R}^d$ from the notation when the meaning is clear from the context. The notation $H^1_\Gamma(\Omega)$ represents those functions in H^1 with zero trace on $\Gamma \subseteq \partial\Omega$. Similarly, $H_\Gamma(\mathrm{div}, \Omega)$ denotes fields in $H(\mathrm{div}, \Omega)$ with zero (normal) trace on $\Gamma \subseteq \partial\Omega$ in the appropriate sense [4]. We also define the standard space of square-integrable functions with zero average:

$$L^2_0(\Omega) = \left\{ p \in L^2(\Omega) \mid \int_\Omega p \, \mathrm{d}x = 0 \right\}.$$

We will also use parameter-weighted norms. For a Banach space X and real parameter $\alpha > 0$, the space αX signifies X equipped with the α-weighted norm $\|f\|_{\alpha X} = \alpha \|f\|_X$. Finally, for a coercive and continuous bilinear form $a : V \times V \to \mathbb{R}$, we will also write

$$\|v\|_a^2 = a(v, v).$$

2.2 Intersections and sums of Hilbert spaces

Let $X \subset Z$ and $Y \subset Z$ be two Hilbert spaces with a common ambient Hilbert space Z. The intersection space, denoted $X \cap Y$, is a Hilbert space with norm

$$\|x\|_{X \cap Y}^2 = \|x\|_X^2 + \|x\|_Y^2.$$

For instance, to illustrate our notation, the norm on the intersection space $\kappa^{-1/2}L^2 \cap H(\mathrm{div})$ is given by

$$\|v\|_{\kappa^{-1/2}L^2 \cap H(\mathrm{div})}^2 = \|v\|_{\kappa^{-1/2}L^2}^2 + \|v\|_{H(\mathrm{div})}^2 = \kappa^{-1}\|v\|_{L^2}^2 + \|v\|_{H(\mathrm{div})}^2.$$

The sum space $X + Y$ is the set $\{z = x + y \mid x \in X, y \in Y\}$ equipped with the norm

$$\|z\|_{X+Y}^2 = \inf_{\substack{z=x+y \\ x \in X, y \in Y}} \|x\|_X^2 + \|y\|_Y^2,$$

and is also a Hilbert space. See e.g. [3, Ch. 2] for a further discussion of sum and intersection spaces.

2.3 Operators

For a given time step size τ, times t^{m-1} and t^m and fields $u^m \approx u(t^m)$ and $u^{m-1} \approx u(t^{m-1})$, we will make use of a discrete derivative notation

$$\partial_\tau u^m = \frac{u^m - u^{m-1}}{\tau}. \tag{2.1}$$

2.4 Finite element spaces

Now, suppose that $\Omega \subset \mathbb{R}^d$ is a polygonal and let $C^k(\Omega)$ denote the space of k-continuously differentiable functions defined on Ω. Let $D \subseteq \Omega$ and let $P^k(D) \subset C^\infty(D)$ denote the set of polynomials of total degree k defined on D. Let $\mathcal{T}_h$ be a simplicial triangulation of Ω and let $T \in \mathcal{T}_h$ be any simplex; we denote the restriction of a function f to $T \in \mathcal{T}_h$ by f_T. The notation for the Lagrange elements of order k used here is then

$$P_k(\mathcal{T}_h) = \left\{ f \in C^0(\Omega) \mid f_T \in P^k(T), \quad \forall T \in \mathcal{T}_h \right\}. \tag{2.2}$$

The notation $P_k^d(\mathcal{T}_h)$ will be used to represent the d-dimensional (vector) Lagrange spaces in $\mathbb{R}^d$. The discontinuous Galerkin spaces of order k relax the overall continuity requirement of the Lagrange finite element spaces; they are defined by

$$DG_k(\mathcal{T}_h) = \left\{ f \in L^2(\Omega) \mid f_T \in P^k(T) \quad \forall T \in \mathcal{T}_h \right\}. \tag{2.3}$$

A comprehensive discussion on Lagrange and discontinuous Galerkin elements and their interpolation properties can be found in e.g. [8] and [29] respectively. We will also make use of the Brezzi-Douglas-Marini and Raviart–Thomas finite element spaces [4, Sec. 2.3]. Throughout the rest of the manuscript we use the notation P_k, P_k^d, DG_k, BDM_k and RT_k in reference to the spaces defined above; that is, we drop the additional mesh domain specification.

2.5 Boundary and initial conditions

General boundary conditions for (1.1) start by considering two distinct, non-overlapping partitions of the $d-1$ dimensional boundary $\partial\Omega$. The first, corresponding to the displacement, is $\partial\Omega = \overline{\Gamma_c} \cup \overline{\Gamma_t}$ and the second, corresponding to the pressure, is denoted $\partial\Omega = \overline{\Gamma_p} \cup \overline{\Gamma_f}$; the non-overlapping condition means $\Gamma_c \cap \Gamma_t = \emptyset$ and $\Gamma_p \cap \Gamma_f = \emptyset$. The general form of the typical boundary conditions are then expressed as

$$\begin{aligned} u &= 0, \text{ on } \Gamma_c, \text{ and} & z \cdot n &= 0, \text{ on } \Gamma_f, \\ p &= 0, \text{ on } \Gamma_p, \text{ and} & \hat{\sigma}(u, p) \cdot n &= 0 \text{ on } \Gamma_t, \end{aligned} \tag{2.4}$$

where $\hat{\sigma}(u, p) = \sigma(u) + p\, I_d$ and I_d is the $d \times d$ identity matrix. We will consider a simplification of the boundary conditions, above. The simplification that we will

consider is that which was studied in the defining work on Stokes–Biot stable discretizations [17,19,26,30]. These conditions take $\Gamma_f = \Gamma_c$ and $\Gamma_p = \Gamma_t$ with the $d-1$ dimensional Lebesgue measure $|\Gamma_c| > 0$. Thus we have

$$
\begin{aligned}
u = 0, \quad &\text{on } \Gamma_c, \text{ and} \quad z \cdot n = 0, \text{ on } \Gamma_c, \\
p = 0, \quad &\text{on } \Gamma_t, \text{ and } \sigma(u) \cdot n = 0 \text{ on } \Gamma_t,
\end{aligned}
\tag{2.5}
$$

Let $\eta(x, t)$ denote the fluid content with equation

$$
\eta(x, t) = c_0 p(x, t) + \operatorname{div} u(x, t).
$$

We follow [31] and remark: that under appropriate regularity assumptions on the sources and initial data, (i.e. source data in $C^\alpha(0, T; (\Omega))$ where α is the Biot-Willis coefficient, boundary data in $C^\alpha(0, T; L^2(\partial\Omega))$, initial fluid content $\eta(x, 0) \in L^2(\Omega)$, etc), then there exists a unique solution to (1.1) satisfying the boundary conditions [31,35].

Remark 2.1 A full discussion on regularity details for the source, initial and boundary data can be found in [35, Theorem 1], and [31, Sect. 3 and 4]. We also note that the boundary conditions (2.5) reflect a restriction that may not be practical for many applications. These boundary conditions coincide with those initially considered in the Stokes–Biot stability literature (e.g. [30]) and allow the key ideas behind the Stokes–Biot (respectively, minimal Stokes–Biot) conditions, discussed in Sect. 3 (respectively, Sect. 4), to be discussed simply. A discussion of more general conditions can be found in Sect. 7, and in e.g. [17].

2.6 Material parameters

To facilitate the analysis here, we will assume that the material parameters of (1.1a)-(1.1c), i.e. μ, λ, α, κ, and c_0, are constant in space and time. For simplicity and without loss of generality we set $\alpha = 1$. This view can either be interpreted literally or as having divided (1.1a–1.1c) through by α to obtain rescaled material parameters. Moreover, one need not look far [12,13,17,25,27,33,34] to find applications where κ is small, and the storage coefficient c_0 varies over a wide range of values in the presence of only modest choices of λ. For instance, the literature contains examples of low hydraulic conductivities where both λ and c_0 are approximately unity [17]; in various soft-tissues, $\lambda \approx 10^2$ and $c_0 \approx 10^{-5}$ have been used [12,13], in addition to $\lambda \approx 10^1$ or $\lambda \approx 10^3$ with $c_0 \approx 10^{-10}$ [25,33], and even $c_0 = 0$ [27,34]. This wide variation in c_0, while λ remains modest, can be due to several reasons: an ad-hoc modeling assumption; to simplify numerical methods when storage coefficients are near the limits of computing precision; or due to the fact that, especially in biological applications, measurements for certain parameters may be unavailable and values are often estimated, chosen, or substituted from those, of similar biological regime, for which reasonable parameter estimates are available.

This manuscript is only concerned with Stokes–Biot stable discretizations; these discretizations are designed to retain their stability and convergence properties in

the presence of diminished hydraulic conductivity. Given the wide variety of storage coefficients which appear, in the applied literature, in the presence of values for $\lambda \in [10^1, 10^3]$ we take the view here that c_0 and λ are independent parameters; this is not to assert that the linear poroelasticity theory does not imply that $\lambda \to \infty$ as $c_0 \to 0$. Rather, we do this to make a secondary, strictly-numerical observation: that Stokes–Biot stable schemes, and the relaxation we propose herein, also retain their stability and convergence properties as $\kappa \to 0$ for every fixed choice of $0 \leq c_0 < 1$. We will therefore assume that $0 < \kappa \leq 1$ and $0 \leq c_0 < 1$ are fixed, but otherwise arbitrary, constants.

Remark 2.2 The bilinear forms defined in Sect. 3 are parameter-dependent. Thus, the arguments advanced in this manuscript may potentially be extended to parameters that vary in space or time, provided they satisfy suitable regularity requirements to justify the requiste manipulations. As parameter, or data, regularity is not the focus on the current work, we do not take up this issue herein; we belay the topic and consider constant (i.e. constant μ, λ and arbitrary but fixed $0 < \kappa \leq 1$, and $0 \leq c_0 \leq 1$) parameters.

3 The Stokes–Biot stability conditions for conforming Euler–Galerkin schemes

Combining the nature of (1.1) with the boundary conditions (2.5), we define the spaces

$$U = H^1_{\Gamma_c}(\Omega), \quad W = H_{\Gamma_c}(\text{div}, \Omega), \quad Q = L^2(\Omega). \tag{3.1}$$

We consider the following variational formulation of (1.1) over the time interval $(0, T]$: for a.e. $t \in (0, T]$, find the displacement u, flux z and pressure p such that $u(t) \in U$, $z(t) \in Z$ and $p(t) \in Q$ satisfy

$$a(u, v) + b(v, p) = (f, v) \qquad v \in V, \tag{3.2a}$$
$$c(z, w) + b(w, p) = (g, w) \qquad w \in W, \tag{3.2b}$$
$$b(\partial_t u, q) + b(z, q) - d(\partial_t p, q) = (s, q) \qquad q \in Q. \tag{3.2c}$$

The bilinear forms in (3.2) are given by:

$$a(u, v) = (\sigma(u), \varepsilon(v)), \quad b(u, q) = (\text{div } u, q),$$
$$c(z, w) = \left(\kappa^{-1} z, w\right), \quad d(p, q) = (c_0 p, q). \tag{3.3}$$

As noted in [30]: the existence and uniqueness of a solution (u, z, p) to (3.2), with continuous dependence on f, g and s, has been established by previous authors [26, 31,35].

Remark 3.1 If Dirichlet conditions are imposed for the displacement on the entire boundary and thus the pressure is only determined up to a constant (i.e. if $\Gamma_c = \partial\Omega$) we instead let $Q = L^2_0$.

3.1 An Euler–Galerkin discrete scheme

Following [30] we consider Euler–Galerkin discretizations; i.e., conforming finite element spaces in space and an implicit Euler in time, of (3.2). Let $0 = t_0 < t_1 < \cdots < t_N = T$ be a uniform partition of the time interval $[0, T]$. The constant time step is then $\tau = \tau_m = t^m - t^{m-1}$. For the function $f(t, x)$, evaluation at t^m is denoted by $f^m = f(t^m, x)$, and similarly for g and s. We define conforming discrete spaces

$$U_h \subset U, \quad W_h \subset W, \quad Q_h \subset Q. \tag{3.4}$$

The Euler–Galerkin discrete scheme of Biot's equations then reads as follows: for each time iterate $m \in \{1, 2, \ldots, N\}$, given f^m, g^m, s^m, $\operatorname{div} u_h^{m-1}$, and, iff $c_0 > 0$, p_h^{m-1}, we seek $(u_h^m, z_h^m, p_h^m) \in U_h \times W_h \times Q_h$ such that

$$a(u_h^m, v) + b(v, p_h^m) = \left(f^m, v\right), \tag{3.5a}$$

$$\tau c(z_h^m, w) + \tau b(w, p_h^m) = \tau \left(g^m, w\right), \tag{3.5b}$$

$$b(\partial_\tau u_h^m, q) + b(z_h^m, q) - d(\partial_\tau p_h^m, q) = \left(s^m, q\right), \tag{3.5c}$$

for all $v \in U_h$, $w \in W_h$ and $q \in Q_h$, and where we have made use of the discrete derivative notation (2.1).

3.2 The Stokes–Biot stability conditions

The Stokes–Biot stability conditions were introduced independently, in slightly different contexts, by several authors [17,22,27,30] and guide the selection of discrete spaces, $U_h \times W_h \times Q_h$, for (3.5). We recall a succinct statement of the (conforming) Stokes–Biot stability conditions, used in analogous forms by all original authors [17,22,30], here for posterity:

Definition 3.1 *(c.f.* [30, Defn. 3.1]*)* The discrete spaces $U_h \subset U$, $W_h \subset W$ and $Q_h \subset Q$ are called a Stokes–Biot stable discretization if and only if the following conditions are satisfied:

(i) The bilinear form a, as defined by (3.3), is bounded and coercive on U_h;
(ii) The pairing (U_h, Q_h) is Stokes stable;
(iii) The pairing (W_h, Q_h) is Darcy (Poisson) stable.

We remark that [17,22] were not conforming. More precisely, the Stokes and Darcy stability assumptions of Definition 3.1 entail that the relevant discrete spaces are stable in the (discrete) Babuška-Brezzi sense [4,6] for the discrete Stokes and Darcy problems, respectively. We will now examine the Darcy stability condition more closely.

3.3 The Darcy stability condition

The discrete Darcy problem reads as: find $(z_h, p_h) \in W_h \times Q_h$ such that (1.4) holds for all $w \in W_h$ and $q \in Q_h$. Assume that $W \subseteq W$ and $Q_h \subset Q$ are equipped

with norms $\| \cdot \|_W$ and $\| \cdot \|_Q$, respectively. The space $W_h \times Q_h$ is Darcy stable in the (discrete) Babuška-Brezzi sense if the discrete Babuška-Brezzi conditions are satisfied, in particular, if there exists constants $\alpha > 0$ and $\beta > 0$, independent of h, such that

$$c(w, w) \geq \alpha \|w\|_W^2 \quad \forall \, w \in \ker b = \{w \in W_h \,|\, b(w, q) = 0 \,\forall q \in Q_h\}, \qquad (3.6)$$

$$\inf_{q \in Q_h} \sup_{w \in W_h} \frac{b(w, q)}{\|w\|_W \|q\|_Q} \geq \beta > 0, \qquad (3.7)$$

with b and c as defined by (3.3). It is also assumed that b and c are continuous over $W \times Q$ and $W \times W$ with respect to the relevant norms; i.e. there exist constants $C_b > 0$ and $C_c > 0$, independent of h, such that

$$b(v, q) \leq C_b \|v\|_W \|q\|_Q, \quad c(v, w) \leq C_c \|v\|_W \|w\|_W, \qquad (3.8)$$

for all $v, w \in W, q \in Q$.

The assumption of discrete Darcy stability, and thus the existence of solutions to the discrete Darcy problem, has been used to define Galerkin projectors for use in the a-priori analysis of the Biot Eq. (3.5) (c.f. for instance [30, Sect. 4.2]). Given $z(t) \in W$ and $p(t) \in Q$ solving the continuous Biot Eq. (3.2), these projectors $\Pi_{W_h} z(t)$ and $\Pi_{Q_h} p(t)$ solve the discrete Darcy problem (1.4) for all $w \in W_h, q \in Q_h$ with right-hand sides given by $(g, w) = c(z(t), w) + b(w, p(t))$ and $(s, q) = b(z(t), q)$. For an a-priori analysis based on such a Galerkin-projection approach to be optimal, including in the limit as $\kappa \to 0$, the continuity constants C_b, C_c and the Babuška-Brezzi stability constants α, β must be independent of $0 < \kappa \leq 1$.

Attaining κ-independent continuity and stability constants is non-trivial for the Darcy problem, and the norms that are selected for W and Q play a vital role. For instance, the standard pairing $H(\text{div}) \times L^2$ with the natural norms is not appropriate as e.g. c is not continuous with respect to the $H(\text{div})$ norm: the continuity bound C_c depends on κ. However, the following pairings for $W \times Q$ are all meaningful for (1.4) or its dual $L^2 \times H^1$ formulation:

(A) $\left(\kappa^{-1/2} L^2 \cap H(\text{div})\right) \times \left(L^2 + \kappa^{1/2} H^1\right)$
(B) $\kappa^{-1/2} H(\text{div}) \times \kappa^{1/2} L^2$
(C) $\kappa^{-1/2} L^2 \times \kappa^{1/2} H^1$

In particular, the inf-sup condition (3.7) holds with inf-sup constant β independent of κ for each of these pairings. We remark that $\|p\|_{L^2 + \kappa^{1/2} H^1} \leq \|p\|$ and $\|p\|_{\kappa^{1/2} L^2} \leq \|p\|$ for $\kappa \leq 1$. The κ-independent inf-sup condition for (A) was recently shown in [2], the inf-sup condition for (B) follows directly by a scaling of the flux by $\kappa^{-1/2}$ and the pressure by $\kappa^{1/2}$. Finally, the inf-sup condition of (C) follows directly from Poincare's inequality with a similar scaling as in (B). The boundedness of $b(z, p)$ can be established for each of the pairings above. The pairing of (C) corresponds to the case of the $L^2 \times H^1$ formulation of the mixed Darcy problem, i.e. $b(z, p) = (z, \nabla p)$ with $z \in W = L^2$ and $p \in Q = H^1$, but boundedness is proved in the same manner as for (B). In the case of (B): applying Cauchy-Schwarz and the weighted norm definitions immediately gives

$$|b(z, p)| \leq \| \operatorname{div} z\| \|p\| \leq \|z\|_{H(\operatorname{div})} \|p\| = \left(\kappa^{-1/2}\|z\|_{H(\operatorname{div})}\right) \left(\kappa^{1/2}\|p\|\right).$$

The case of (A) is complicated by the definition of the sum norm on the pressure space Q, and a one-line argument is not possible without additional context; see [2] for details.

Options (A) and (B) above fit naturally with the variational formulation of (3.5) and spaces (3.1). In the following, we suggest that a natural norm for the Darcy flux is

$$\|\|z\|\|^2 = \frac{\tau}{\kappa} (z, z) + \tau^2 (\operatorname{div} z, \operatorname{div} z), \tag{3.9}$$

which is equivalent to the norm of the flux in (A) above for the relevant range of κ when $\tau > 0$. However, both options (A) and (B) have disadvantages. For (B), the pressure norm (on Q) becomes progressively weaker as κ nears 0 while the norm of the flux divergence (on W) is unnecessarily large compared with e.g. (3.9). The primary drawback to using (A) is that the pressure norm is implicitly defined. This fact means that an a-priori analysis based on the method of projections is more complex to carry out in practice; it is not clear that standard analytic techniques, e.g. in [17,19,22,30] among others, could be used directly when the norm of $L^2 + \kappa^{1/2} H^1$ is chosen for the pressure space.

We will argue instead that an a-priori analysis of (3.5) based on the use of a Galerkin projection of the form (1.4) is not necessary; thus alleviating the need for an explicit uniform-in-κ Darcy stability condition on (W_h, Q_h). Neither (3.7) nor the saddle-point stability of (1.4) in general play a role in the well-posedness of (3.5). Condition (iii) of Definition 3.1 will thus be replaced by a less restrictive condition. An important consequence of relaxing the uniform-in-κ Darcy stability hypothesis is that the standard L^2-norm on Q can, and will, be used.

4 Minimal Stokes–Biot stability

In this section we state the definition of minimal Stokes–Biot stability and recall a previous inf-sup condition in the spirit of the Banach-Nečas-Babuška theorem. In particular, the minimal Stokes–Biot stability conditions (c.f. Definition 4.1) relinquish the Darcy stability assumption in favor of a containment condition. In practice, this containment condition is satisfied for discrete flux-pressure pairings that are Darcy stable, though other discrete spaces satisfy this condition which are not stable pairings for the mixed Darcy problem. Throughout this section we assume that U, W and Q are defined by (3.1). The norm on U is taken to be the usual $H^1(\Omega)$-norm $\| \cdot \|_1$, the norm on Q is the standard L^2-norm $\| \cdot \|$, while the norm $\|\| \cdot \|\|$ on W is the weighted norm defined by (3.9). The norm (3.9) was first introduced in [19, Sect. 3.1]. The bilinear forms a, b, c, d are as defined by (3.3).

4.1 Minimal Stokes–Biot conditions

We now introduce our set of minimal Stokes–Biot stability conditions. For clarity and completeness (rather than e.g. brevity), we include the precise stability conditions in the definition here. In essence, between Definitions 3.1 and 4.1, only condition (iii) changes.

Definition 4.1 A family of conforming discrete spaces $\{U_h \times W_h \times Q_h\}_h$ with $U_h \subset U$, $W_h \subset W$ and $Q_h \subset Q$ is called minimally Stokes–Biot stable if and only if

(i) The bilinear form a is continuous and coercive on $U_h \times U_h$; i.e. there exists constants $C_a > 0$ and $\gamma_a > 0$ independent of h such that

$$a(u, u) \geq \gamma_a \|u\|_1^2, \quad a(u, v) \leq C_a \|u\|_1 \|v\|_1, \quad \forall u, v \in U_h. \qquad (4.1)$$

(ii) The pairings $\{U_h \times Q_h\}_h$ are Stokes stable in the discrete Babuška-Brezzi sense [5, 6]; i.e. in particular there exists an inf-sup constant $\beta_S > 0$ independent of h such that

$$\inf_{q \in Q_h} \sup_{v \in U_h} \frac{b(v, q)}{\|v\|_1 \|q\|} \geq \beta_S > 0. \qquad (4.2)$$

(iii) div $W_h \subseteq Q_h$ for each h.

The classical flux-pressure pairings, e.g. $RT_k \times DG_k$ or $BDM_{k+1} \times DG_k$ for $k = 0, 1, 2, \ldots$, satisfying Definition 3.1(iii) also satisfy the conditions of minimal Stokes–Biot stability; in particular Definition 4.1(iii). However, the minimal Stokes–Biot condition also includes discretizations which are not encompassed by Definition 3.1. For instance: flux-pressure pairings where the flux is taken from the space of continuous Lagrange polynomials can satisfy Definition 4.1 while not satisfying Definition 3.1. An illustration of this can be found in the family of discretizations where the displacement-pressure pairing are of Scott-Vogelius type; these either have the form $P_k^d \times RT_m \times DG_{k-1}$ or $P_k^d \times P_m^d \times DG_{k-1}$ where $k \geq 4$ and $0 \leq m \leq k-1$. The flux-pressure pairings $RT_m \times DG_{k-1}$, for $m < k - 1$, and $P_m^d \times DG_{k-1}$, for $m \leq k - 1$, are *not* Darcy stable but *do satisfy* the minimal Stokes–Biot stability containment condition of Definition 4.1(iii).

A more pragmatic example is the discretization $P_2^d \times RT_0 \times DG_0$. This discretization is both Stokes–Biot stable and minimally Stokes–Biot stable; of note is that $P_2^d \times P_1^d \times DG_0$ is not Stokes–Biot stable but is minimally Stokes–Biot stable. The $P_2^d \times RT_0 \times DG_0$ discretization is a prototype for the minimal-dof displacement enrichment of a $P_1^d \times RT_0 \times DG_0$ approach studied in [30]. The comparison between $P_2^d \times RT_0 \times DG_0$ and $P_2^d \times P_1^d \times DG_0$ serves as a motivation for Definition 4.1, and will be studied in Sect. 6. A further discussion of spaces that satisfy the minimal Stokes–Biot stability condition is given in Sect. 7.

4.2 An inf-sup condition for minimal Stokes–Biot stable Euler–Galerkin schemes

The variational problem (3.5) can be shown to satisfy a requirement of the Banach-Nečas-Babuška theorem with respect to the weighted norm (3.9) and Definition 4.1. In fact, this result was proved in [19].

Proposition 4.1 (Theorem 1, [19]) *Let* $\|(u_h, w_h, q_h)\|_{UWQ}$ *be defined by*

$$\|(u_h, w_h, q_h)\|_{UWQ} = \left(\|u_h\|_1^2 + \|\|w_h\|\|^2 + \|q_h\|^2 \right)^{1/2}$$

where $\|\|w_h\|\|$ *is defined by (3.9). Define a composite bilinear form, on* $U_h \times W_h \times Q_h$ *and corresponding to (3.5), by the formula*

$$\begin{aligned}
\mathcal{B}(u_h, z_h, p_h; v_h, r_h, q_h) &= a(u_h, v_h) + b(v_h, p_h) + \tau\, c(z_h, r_h) \\
&\quad + \tau\, b(r_h, p_h) + b(u_h, q_h) + \tau\, b(z_h, q_h) - d(p_h, q_h)
\end{aligned}$$

Suppose $U_h \times W_h \times Q_h$ *satisfy the assumptions of Definition 4.1. Then* $\mathcal{B}$ *is continuous and there exists a constant* $\gamma > 0$, *independent of* κ *and* c_0, *such that*

$$\sup_{(v_h, r_h, q_h) \in U_h \times W_h \times Q_h} \frac{\mathcal{B}(u_h, z_h, p_h; v_h, r_h, q_h)}{\|(v_h, r_h, q_h)\|_{UWQ}} \geq \gamma \|(u_h, z_h, p_h)\|_{UWQ}$$

Proof The proof follows from the arguments in [19, Theorem 1]. □

Remark 4.1 Work by previous authors [17,19] shows that the assumptions of Definition 4.1 were nascent in the literature. The proof [19] of Proposition 4.1 is independent of $0 \leq c_0$, and does not invoke Darcy stability, but does, in fact, use condition (iii) of Definition 4.1. In fact, another version of Proposition 4.1 was also proved, independently, in [17, Theorem 3.2, Case I]; the proof, once more, is independent of c_0 and does not assume that the divergence maps the flux space surjectively onto the pressure space (i.e. Darcy stability). A nice mention of the case $U = H_0^1$ and $Q = L_0^2$ can also be found therein. The arguments of [17, Theorem 3.2, Case I] follow similarly to those of [19, Theorem 2].

Corollary 4.1 *Assume that the assumptions of Definition 4.1 hold; then (3.5) is well posed.*

Proof The Banach-Nečas-Babuška theorem [8, Theorem 2.6], applied to (3.5), requires that two conditions are satisfied. The first condition is that of Proposition 4.1, which has been proved, independently, by several authors. The second condition, which remains to be verified, is that if an element $(v_h, r_h, q_h) \in U_h \times W_h \times Q_h$ is such that

$$\mathcal{B}(u_h, z_h, p_h; v_h, r_h, q_h) = 0, \quad \forall (u_h, z_h, p_h) \in U_h \times W_h \times Q_h,$$

then $v_h = r_h = q_h = 0$ must follow. To show that this condition also holds true, fix $(v_h, r_h, q_h) \in U_h \times W_h \times Q_h$ and suppose that the above implication holds; we need

to show that, in this case, it must be that $v_h = r_h = q_h = 0$. Towards this end we consider two cases: the first case is if $c_0 > 0$, and the second case is if $c_0 = 0$. For the first case, select $u_h = v_h$, $z_h = r_h$ and $p_h = -q_h$, along with (3.9), the hypothesis above and (3.3), to get

$$\mathcal{B}(v_h, r_h, -q_h; v_h, r_h, q_h) = a(v_h, v_h) + \frac{\tau}{\kappa}\|r_h\|^2 + c_0\|q_h\|^2 = 0.$$

Coercivity (c.f. (4.1)) gives $\gamma_a\|v_h\|_1^2 \leq a(v_h, v_h)$ and $v_h = r_h = q_h = 0$ follows. For the second case, assume that $c_0 = 0$. The Stokes stability assumption (Definition 4.1(ii)) implies that (e.g. [5, p. 136]) there exists $y_h \in U_h$ such that

$$(\text{div } y_h, q_h) = \|q_h\|^2, \tag{4.3}$$
$$\beta_S\|y_h\|_1 \leq \|q_h\|, \tag{4.4}$$

where β_S is the Stokes inf-sup constant of (4.2). Let $\delta \geq 0$ be a yet-undetermined, but fixed, constant and choose $u_h = v_h + \delta y_h$, $z_h = r_h$, and $p_h = -q_h$. With these choices, and (3.3), we have

$$\mathcal{B}(v_h + \delta y_h, r_h, -q_h; v_h, r_h, q_h) = a(v_h, v_h) + \delta a(y_h, v_h) + \frac{\tau}{\kappa}\|r_h\|^2 + \delta\|q_h\|^2 = 0.$$

The coercivity and continuity assumptions (c.f. (4.1)), together with Cauchy's inequality with epsilon, (4.4) and gathering of like terms gives

$$(\gamma_a - \delta C_a \epsilon)\|v_h\|_1^2 + \frac{\tau}{\kappa}\|r_h\|^2 + \delta\left(1 - \frac{C_a}{4\beta_S^2\epsilon}\right)\|q_h\|^2 \leq 0. \tag{4.5}$$

We can now select the appropriate constants δ and ϵ, as e.g.

$$\epsilon = 2\frac{C_a}{4\beta_S^2} > 0, \quad \delta = \frac{\gamma_a \beta_S^2}{C_a^2} > 0, \tag{4.6}$$

from which it follows that

$$\gamma_a\|v_h\|_1^2 + \frac{\tau}{\kappa}\|r_h\|^2 + \frac{1}{2}\frac{\gamma_a \beta_S^2}{C_a^2}\|q_h^m\|^2 \leq 0. \tag{4.7}$$

and thus $v_h = r_h = q_h = 0$. Thus, the second condition of the Banach-Nečas-Babuška theorem [8, Theorem 2.6] holds, irregardless of c_0, and the result follows. □

5 A priori error estimates for minimally Stokes–Biot stable schemes

In this section, we derive a-priori error estimates for the Euler–Galerkin discrete Biot equations (3.5) using the assumptions of Definition 4.1. The final result is summarized

in Proposition 5.2 of Sect. 5.4. We will assume the point of view of minimal Stokes–Biot stability as defined by Definition 4.1 and that U_h contains the continuous nodal Lagrange elements P_r for some $r \geq 1$. We begin by establishing basic assumptions on the spaces U_h, W_h and Q_h, and define projection operators in Sect. 5.1.

5.1 Projections and approximability

As in the previous section, let U, W, Q be given by (3.1) with norms $\| \cdot \|_1$, $\|\|\cdot\|\|$ cf. (3.9), and $\| \cdot \|$, respectively. Assume that the discrete spaces $U_h \times W_h \times Q_h$ satisfy the assumptions of Definition 4.1. We denote the (continuous) solutions to (3.2) at time t^m by (u^m, z^m, p^m) for $m = 1, 2, \ldots, N$ while (u_h^m, z_h^m, p_h^m) represent the solutions of the discrete problem (3.5). For use in the subsequent error analysis, we make basic assumptions on the spaces U_h, W_h and Q_h, and define projection operators $\Pi_{U_h} : U \to U_h$, $\Pi_{W_h} : W \to W_h$ and $\Pi_{Q_h} : Q \to Q_h$ as follows.

Q_h: Define Π_{Q_h} to be the standard L^2-projection into Q_h. Then

$$\|q - \Pi_{Q_h} q\| \lesssim \inf_{q_h \in Q_h} \|q - q_h\|,$$

for all $q \in Q$. If Q_h contains piecewise polynomials of order $k = k_Q \geq 0$, then in particular

$$\|q - \Pi_{Q_h} q\| \lesssim h^{k_Q+1} \|q\|_{k_Q+1}, \quad \forall w \in H^k. \tag{5.1}$$

W_h: Assume that W_h contains (at least) piecewise polynomial (vector) fields of order $k = k_W \geq 0$. We assume the existence of a generic discrete interpolant $\Pi_{W_h} : W \to W_h$ satisfying either

$$\|w - \Pi_{W_h} w\| \lesssim h^{k_W+1} \|w\|_{k_W+1} \text{ and } \| \operatorname{div}(w - \Pi_{W_h} w)\|$$
$$\lesssim h^{k_W+1} \| \operatorname{div} w\|_{k_W+1}, \tag{5.2}$$

for $w \in H^{k_W+2}$, or

$$\|w - \Pi_{W_h} w\| \lesssim h^{k_W+1} \|w\|_{k_W+1} \text{ and } \|w - \Pi_{W_h} w\|_1 \lesssim h^{k_W} \|w\|_{k_W+1}. \tag{5.3}$$

for $w \in H^{k+1}(\Omega)$. The estimates (5.2) are characteristic of a Raviart-Thomas type, RT_k ($k = 0, 1, 2, \ldots$), interpolant whereas (5.3) could correspond to a continuous Lagrange interpolant of order $k \geq 1$ [8].

U_h: Following [30], we define $\Pi_{U_h} : U \to U_h$ as a modified elliptic projection satisfying for $u \in U$:

$$a(\Pi_{U_h} u, v) = a(u, v) + b(v, q - \Pi_{Q_h} q) \quad \forall v \in U_h, \tag{5.4}$$

where $q \in Q$ is given and will, in practice, be selected as the exact pressure solution to (3.2) at given times.

Assume that U_h contains (at least) continuous piecewise polynomial (vector) fields of order $k_U \geq 1$. There then exists an interpolant, $I^{k_U} : U \to U_h$, such that

$$\|u - I^{k_U}u\|_1 \lesssim h^{k_U}\|u\|_{k_U+1}$$

for all $u \in H^{k_U+1}$, c.f. e.g [8]. Then for $u \in U$ we have

$$\|u - \Pi_{U_h}u\|_1 \leq \|u - I^{k_U}u\|_1 + \|I^{k_U}u - \Pi_{U_h}u\|_1 \lesssim h^{k_U}\|u\|_{k_U+1}$$
$$+ \|I^{k_U}u - \Pi_{U_h}u\|_1.$$

Using assumption (i) of Definition 4.1 and (5.4) with $v = \Pi_{U_h}u - I^{k_U}u$ imply that

$$\gamma_a\|\Pi_{U_h}u - I^{k_U}u\|_1^2 \leq a(\Pi_{U_h}u - I^{k_U}u, \Pi_{U_h}u - I^{k_U}u)$$
$$= a(u - I^{k_U}u, \Pi_{U_h}u - I^{k_U}u)$$
$$+ b(\Pi_{U_h}u - I^{k_U}u, q - \Pi_{Q_h}q)$$
$$\leq \|\Pi_{U_h}u - I^{k_U}u\|_1 \left(C_a\|u - I^{k_U}u\|_1 + \|q - \Pi_{Q_h}q\|\right).$$

Combining the above with assumption (5.1) gives

$$\|u - \Pi_{U_h}u\|_1 \lesssim h^{k_U}\|v\|_{k_U+1} + h^{k_Q+1}\|q\|_{k_Q+1}, \qquad (5.5)$$

where $q \in Q$ is the fixed function defining the elliptic projection (5.4).

5.2 Interpolation notation and identities

Following standard notation [17,22,30], the error at time $t^m > 0$ can be decomposed into interpolation errors ρ and approximation errors e:

$$u^m - u_h^m = \left(u^m - \Pi_{U_h}u^m\right) - \left(u_h^m - \Pi_{U_h}u^m\right) = \rho_u^m - e_u^m$$
$$z^m - z_h^m = \left(z^m - \Pi_{W_h}z^m\right) - \left(z_h^m - \Pi_{W_h}z^m\right) = \rho_z^m - e_z^m \qquad (5.6)$$
$$p^m - p_h^m = \left(p^m - \Pi_{Q_h}p^m\right) - \left(p_h^m - \Pi_{Q_h}p^m\right) = \rho_p^m - e_p^m.$$

The interpolation errors satisfy the following identities. Since div $W_h \subseteq Q_h$ and by the definition of the L^2-projection Π_{Q_h}, we have that

$$b(w, \rho_p^m) = \left(\text{div } w, p^m - \Pi_{Q_h}p^m\right) = 0 \quad \forall w \in W_h. \qquad (5.7)$$

Similarly, by the definition of Π_{Q_h},

$$d(\partial_\tau \rho_p^m, q) = c_0 \left(\partial_\tau \rho_p^m, q\right) = 0 \quad \forall q \in Q_h, \qquad (5.8)$$

where we recall the discrete derivative notation (2.1). Finally, (5.4) directly gives

$$a(\rho_u^m, v) + b(v, \rho_p^m) = 0, \quad \forall v \in U_h. \tag{5.9}$$

Taking the difference between the continuous Eq. (3.2) and discrete scheme (3.5), after multiplying (3.2b) by τ, combined with the cancellations (5.7)–(5.9), yield the following error equations at t^m: (e_u^m, e_z^m, e_p^m) satisfies

$$a(e_u^m, v) + b(v, e_p^m) = 0 \qquad \forall v_h \in U_h, \tag{5.10a}$$

$$\tau c(e_z^m, w) + \tau b(w, e_p^m) = \tau c(\rho_z^m, w), \qquad \forall w \in W_h, \tag{5.10b}$$

$$b(\partial_\tau e_u^m, q) + b(e_z^m, q) - d(\partial_\tau e_p^m, q) = \left(R^m, q\right) \qquad \forall q \in Q_h, \tag{5.10c}$$

where

$$R^m = \operatorname{div}(\partial_t u^m - \partial_\tau u^m) + \operatorname{div}(\partial_\tau \rho_u^m) + \operatorname{div} \rho_z^m + c_0(\partial_t p^m - \partial_\tau p^m), \tag{5.11}$$

by way of the general identity

$$\partial_t u^m - \partial_\tau u_h^m = \partial_t u^m - \partial_\tau u^m + \partial_\tau \rho_u^m - \partial_\tau e_u^m, \tag{5.12}$$

and similarly for p.

5.3 Discrete approximation error estimates

In this section we estimate the discrete errors described by (5.10) in their respective norms; that is, $\|e_u^m\|_1$, $\|\|e_z^m\|\|$ and $\|e_p^m\|$. In contrast to e.g. [30], we do not make use of the restrictive uniform-in-κ Darcy stability assumption. In turn, the error equations require a more technical analysis and we have adapted related methods originally used to study κ fixed [22] and vanishing (c_0) storage coefficient. Despite the more technical approach, the resulting estimates presented in Proposition 5.2 is directly comparable to related results in the literature; c.f. [17, Lem. 3], [22, Thm. 4.1] and [30, Thm 4.6]. We conclude that the concept of minimal Stokes–Biot stability provides analogous error estimates for a more general set of conforming discrete spaces than the original Stokes–Biot stability concept.

During the course of the analysis will make use of the following useful inequality

Lemma 5.1 [22, Lemma 3.2] *Suppose that A, B, C > 0 and D ≥ 0 satisfy*

$$A^2 + B^2 \leq CA + D.$$

Then either $A + B \leq 4C$ or $A + B \leq 2\sqrt{D}$ holds.

Proposition 5.1 *Suppose that $U_h \times W_h \times Q_h$ is minimally Stokes–Biot stable (by satisfying the assumptions of Definition 4.1). Then, the discrete approximation errors (e_u^m, e_z^m, e_p^m) described by (5.10) satisfy the inequality:*

$$\|e_u^m\|_1 + \|e_p^m\| + \||e_z^m|\| \lesssim \|e_u^0\|_1 + \|e_p^0\|_d + \tau^{1/2}\|e_z^0\|_c$$
$$+ \left(\int_0^T \|\rho_z\|_c^2 \, ds\right)^{1/2} + \tau \int_0^T \|\rho_{\partial_t z}\|_c^2 \, ds + C_\tau^T, \qquad (5.13)$$

with inequality constant depending on C_a, γ_a^{-1} and where

$$C_\tau^m \equiv \int_0^{t_m} \|\operatorname{div} \rho_z\| + \|\rho_{\partial_t u}\|_1 + \tau \left(c_0\|\partial_{tt} p\| + \|\partial_{tt} u\|_1\right) \, ds.$$

Proof In an analogous fashion as for Proposition 4.1, multiplying (5.10c) by τ, selecting $v = e_u^m - e_u^{m-1}$, $w = e_z^m$, and $q = -e_p^m$ in (5.10) and summing gives

$$a(e_u^m - e_u^{m-1}, e_u^m) + \tau c(e_z^m, e_z^m) + d(e_p^m - e_p^{m-1}, e_p^m) = \tau c(\rho_z^m, e_z^m) - \tau\left(R^m, e_p^m\right). \qquad (5.14)$$

For any (continuous) symmetric bilinear form a with induced norm $\|\cdot\|_a$ we have the inequality [9]

$$\frac{1}{2}\left(\|\chi\|_a^2 - \|\chi - \xi\|_a^2\right) \leq a(\xi, \chi). \qquad (5.15)$$

Using the above, and the symmetry of both $a(\cdot, \cdot)$ and $d(\cdot, \cdot)$, it follows that the left-hand side of (5.14) is bounded below by

$$\frac{1}{2}\|e_u^m\|_a^2 - \frac{1}{2}\|e_u^{m-1}\|_a^2 + \tau\|e_z^m\|_c^2 + \frac{1}{2}\|e_p^m\|_d^2 - \frac{1}{2}\|e_p^{m-1}\|_d^2. \qquad (5.16)$$

On the other hand, Cauchy-Schwarz and Young's inequality give

$$|\tau c(\rho_z^m, e_z^m)| \leq \frac{\tau}{2}\|\rho_z^m\|_c^2 + \frac{\tau}{2}\|e_z^m\|_c^2. \qquad (5.17)$$

From the Stokes stability assumption (4.2) and (5.10a) we have the estimate

$$\beta_S\|e_p^m\| \leq \sup_{v \in U_h} \frac{b(v, e_p^m)}{\|v\|_1} = \sup_{v \in U_h} \frac{-a(e_u^m, v)}{\|v\|_1} \leq C_a\|e_u^m\|_1. \qquad (5.18)$$

Then, Cauchy-Schwarz, (5.18) and the coercivity of a gives

$$\tau\left|\left(R^m, e_p^m\right)\right| \leq C_a\beta_S^{-1}\gamma_a^{-1/2}\tau\|R^m\|\|e_u^m\|_a. \qquad (5.19)$$

Combining (5.16–5.19) yields

$$\|e_u^m\|_a^2 - \|e_u^{m-1}\|_a^2 + \tau\|e_z^m\|_c^2 + \|e_p^m\|_d^2 - \|e_p^{m-1}\|_d^2 \lesssim \tau\left(\|\rho_z^m\|_c^2 + \|R^m\|\|e_u^m\|_a\right). \qquad (5.20)$$

with inequality constant depending on $C_a\beta_S^{-1}\gamma_a^{-1/2}$.

Estimate of $\|e_u^m\|_a$: Following a technique from [22], let J be the integer index where $\|e_u^m\|_a$ (for $m = 1, \ldots, N$) obtains its maximal value. Summing (5.20) from $m = 1$ to $m = J$, using the maximality assumption, and re-arranging terms yields

$$\|e_u^J\|_a^2 + \tau \sum_{m=1}^{J} \|e_z^m\|_c^2 + \|e_p^J\|_d^2 \lesssim \|e_u^0\|_a^2 + \|e_p^0\|_c^2 + \sum_{m=1}^{J} \tau \|\rho_z^m\|_c^2 + \sum_{m=1}^{J} \tau \|R^m\| \|e_u^J\|_a.$$

(5.21)

We can apply Lemma 5.1–5.21 by taking $A = \|e_u^J\|_a$, $B = \|e_p^J\|_d$ and dropping the additional left-hand side term; then we choose

$$C = \sum_{m=1}^{J} \tau \|R^m\|, \quad D = \|e_u^0\|_a^2 + \|e_p^0\|_c^2 + \sum_{m=1}^{J} \tau \|\rho_z^m\|_c^2,$$

and, provided appropriate temporal regularity of the exact solution, have

$$\sum_{m=1}^{J} \tau \|\rho_z^m\|_c^2 \lesssim \int_0^{t^J} \|\rho_z\|_c^2 \, ds.$$

Lemma 5.1, with the above and the triangle inequality, implies

$$\|e_u^J\|_a + \|e_p^J\|_d \lesssim \|e_u^0\|_a + \|e_p^0\|_d + \sum_{m=1}^{J} \tau \|R^m\| + \left(\int_0^{t^J} \|\rho_z\|_c^2 \right)^{1/2}.$$

(5.22)

Bound of $\tau \|R^m\|$: We now develop a bound for the terms $\tau \|R^m\|$; c.f. (5.11). From the fundamental theorem of calculus and integration by parts we have the general result

$$\partial_t f^m - \partial_\tau f^m = \frac{1}{\tau} \int_{t^{m-1}}^{t^m} (s - t^{m-1}) \partial_{tt} f(s) \, ds$$

for any $m = 1, \ldots, N$, assuming sufficient temporal regularity of the field f. We therefore, again under the assumption of sufficient spatial and temporal regularity, have the inequalities

$$c_0 \|\partial_t p^m - \partial_\tau p^m\| \leq \int_{t^{m-1}}^{t^m} c_0 \|\partial_{tt} p\| \, ds$$

(5.23)

$$\|\operatorname{div} (\partial_t u^m - \partial_\tau u^m)\| \leq \int_{t^{m-1}}^{t^m} \|\partial_{tt} u\|_1 \, ds,$$

(5.24)

which control the first and fourth terms of $\|R^m\|$.

For the second term of R^m we have $\|\operatorname{div} \partial_\tau \rho_u^m\| \leq \|\partial_\tau \rho_u^m\|_1$. Rearranging the terms of $\partial_\tau \rho_u^m$, applying the fundamental theorem of calculus and using the commutation of the time derivative with the elliptic projection (5.4) yields

$$\|\partial_\tau \rho_u^m\|_1 = \|\frac{u^m - u^{m-1}}{\tau} - \frac{\Pi_{U_h} u^m - \Pi_{U_h} u^{m-1}}{\tau}\|_1 \leq \frac{1}{\tau} \int_{t^{j-1}}^{t^j} \|\rho_{\partial_t u}\|_1 \, ds. \quad (5.25)$$

For the third term of R^m we have, again up to sufficient temporal regularity of the exact solution, that

$$\sum_{m=1}^{J} \tau \|\operatorname{div} \rho_z^m\| \lesssim \int_0^{t^J} \|\operatorname{div} \rho_z\| \, ds. \quad (5.26)$$

Summarizing, (5.23–5.26) thus yield

$$\sum_{m=1}^{J} \tau \|R^m\| \lesssim \int_0^{t^J} \|\operatorname{div} \rho_z\| + \|\rho_{\partial_t u}\|_1 + \tau \left(c_0\|\partial_{tt} p\| + \|\partial_{tt} u\|_1\right) \, ds \equiv C_\tau^J. \quad (5.27)$$

And so, the estimate (5.22) becomes

$$\|e_u^J\|_a + \|e_p^J\|_d \lesssim \|e_u^0\|_a + \|e_p^0\|_d + \left(\int_0^{t^J} \|\rho_z\|_c^2 \, ds\right)^{1/2} + C_\tau^J \quad (5.28)$$

Clearly, by Definition 4.1(i), this also gives a bound for $\|e_u^m\|_1$ (depending on γ_a^{-1}) for $m = 1, \ldots, N$.

Estimate of $\|e_p^m\|$: The norm $\|e_p^J\|_d$ in e.g. (5.28) vanishes in the limit as $c_0 \to 0$. An alternative bound for $\|e_p^m\|$ can be derived from the Stokes stability assumption, Definition 4.1(ii). In particular, using (5.18) and (5.28) it follows that for each $1 \leq m \leq N$:

$$\|e_p^m\| \lesssim \|e_u^m\|_1 \lesssim \|e_u^J\|_a, \quad (5.29)$$

with inequality constant $C_a \beta_S^{-1} \gamma_a^{-1/2}$ and where J is the index where $\|e_u^J\|_1$ is maximal. Thus $\|e_p^m\|$ can be bounded by the right hand side of (5.28), independently of c_0.

Estimate of $\tau \|e_z^m\|_c^2$: In order to estimate the flux error in the norm defined by (3.9), i.e. $\|\|e_z^m\|\|$, it will be advantageous to consider the constituents separately; e.g. $\tau \|e_z^m\|_c^2$ and $\tau^2 \|\operatorname{div} e_z^m\|^2$.

We begin by considering the first component and again argue based on maximality. Take the difference of the error equation (5.10a) at time levels $m, m-1$ and dividing by τ to get

$$a(\partial_\tau e_u^m, v) + b(v, \partial_\tau e_p^m) = 0 \quad \text{for } v \in U_h. \quad (5.30)$$

Similarly taking the difference of (5.10b) at time levels m and $m-1$, and divide by τ^2 to get

$$c(\partial_\tau e_z^m, w) + b(w, \partial_\tau e_p^m) = c(\partial_\tau \rho_z^m, w) \quad \text{for } w \in W_h.$$

Choose $v = \partial_\tau e_u^m$, $w = e_z^m$ in the above as well as $q = -\partial_\tau e_p^m$ in (5.10c); summing these three equations, using Cauchy-Schwarz on the right-hand side, and coercivity

on the left-hand side gives

$$\gamma_a \|\partial_\tau e_u^m\|_1^2 + \|\partial_\tau e_p^m\|_d^2 + c(\partial_\tau e_z^m, e_z^m) \leq \|\partial_\tau \rho_z^m\|_c \|e_z^m\|_c + \|R^m\| \|\partial_\tau e_p^m\|.$$

From Definition 4.1(ii) and (5.30) we have that $\|\partial_\tau e_p^m\| \leq C_a \beta_S^{-1} \|\partial_\tau e_u^m\|_1$ by the analogue of (5.18). Using this on the right-most term of the above, alongside Cauchy's inequality with epsilon and choosing epsilon appropriately, gives

$$\|\partial_\tau e_u^m\|_1^2 + \|\partial_\tau e_p^m\|_d^2 + c(\partial_\tau e_z^m, e_z^m) \lesssim \|\partial_\tau \rho_z^m\|_c \|e_z^m\|_c + \|R^m\|^2,$$

with inequality constant depending on $C_a \beta_S \gamma_a^{-1}$. Dropping the positive displacement and pressure left-hand side terms, multiplying both sides by τ, and using the symmetry of c together with the inequality (5.15) give

$$\|e_z^m\|_c^2 - \|e_z^{m-1}\|_c^2 \lesssim \tau \|\partial_\tau \rho_z^m\|_c \|e_z^m\|_c + \tau \|R^m\|^2.$$

Let M be the index where $\|e_z^m\|_c^2$ achieves its maximum for $1 \leq m \leq N$. Summing the above from $m = 1$ to $m = M$, using the maximality of $\|e_z^M\|_c$, multiplying both sides by τ and re-arranging yields

$$\tau \|e_z^M\|_c^2 \lesssim \tau \|e_z^0\|_c^2 + \tau \left(\sum_{m=1}^M \tau \|\partial_\tau \rho_z^m\|_c \right) \|e_z^M\|_c + \sum_{m=1}^M \left(\tau \|R^m\| \right)^2. \tag{5.31}$$

By the fundamental theorem of calculus, we have

$$\tau \left(\|\partial_\tau \rho_z^m\|_c \right) = \|\rho_z^m - \rho_z^{m-1}\|_c \leq \int_{t^{m-1}}^{t^m} \|\rho_{\partial_t z}\|_c \, ds.$$

Applying Hölder's inequality on the right-most term, above, gives

$$\int_{t^{m-1}}^{t^m} \|\rho_{\partial_t z}\|_c \leq \left(\int_{t^{m-1}}^{t^m} 1 \, dt \right)^{1/2} \left(\int_{t^{m-1}}^{t^m} \|\rho_{\partial_t z}\|_c^2 \right)^{1/2}$$

so that

$$\tau \left(\|\partial_\tau \rho_z^m\|_c \right) \leq \tau^{1/2} \left(\int_{t^{m-1}}^{t^m} \|\rho_{\partial_t z}\|_c^2 \right)^{1/2}. \tag{5.32}$$

Inserting (5.32) and (5.27) into (5.31), using Young's inequality on the second term on the right-hand side and rearranging yields

$$\tau \|e_z^M\|_c^2 \lesssim \tau \|e_z^0\|_c^2 + \tau \left(\tau \sum_{m=1}^M \|\partial_\tau \rho_z^m\|_c \right)^2 + \sum_{m=1}^M \left(\tau \|R^m\| \right)^2,$$

$$\lesssim \tau \|e_z^0\|_c^2 + \tau^2 \int_0^{t_M} \|\rho_{\partial_t z}\|_c^2 \, ds + \left(C_\tau^M \right)^2 \tag{5.33}$$

Estimate of $\tau^2 \|\operatorname{div} e_z^m\|$:

Now we estimate the second, and final, term in the flux norm (3.9). Let K denote the index where $\|\operatorname{div} e_z^K\|$ is maximal. Using Definition 4.1(iii), and selecting $q = \tau \operatorname{div} e_z^K$ in the error equation (5.10c) for $m = K$ yields

$$\left(\operatorname{div}(e_u^K - e_u^{K-1}), \operatorname{div} e_z^K\right) + \tau \left(\operatorname{div} e_z^K, \operatorname{div} e_z^K\right) - \left(c_0(e_p^K - e_p^{K-1}), \operatorname{div} e_z^K\right)$$
$$= \tau \left(R^K, \operatorname{div} e_z^K\right).$$

Thus, re-arranging terms, using Cauchy-Schwarz and the triangle inequalities, and dividing by $\|\operatorname{div} e_z^K\|$ gives

$$\tau \|\operatorname{div} e_z^K\| \lesssim \|e_u^K\|_1 + \|e_u^{K-1}\|_1 + c_0\|e_p^K\| + c_0\|e_p^{K-1}\| + \tau\|R^K\|$$
$$\lesssim \|e_u^J\|_1 + \tau\|R^K\|,$$

where the last inequality follows from the majorization of the terms e_u^K, e_p^K, e_u^{K-1}, e_p^{K-1} by the maximum $e_u^J = \max_{j=1,2,\ldots,N} e_u^j$ and inequality (5.29). Noting that

$$\left(\|e_u^J\|_1 + \tau\|R^K\|\right)^2 \lesssim \|e_u^J\|_1^2 + \tau^2\|R^K\|^2 \lesssim \|e_u^J\|_1^2 + \sum_{m=1}^K \tau^2\|R^m\|^2,$$

and employing (5.28), (5.27) and taking $I = \max\{J, K\}$ then gives

$$\tau^2 \|\operatorname{div} e_z^K\|^2 \lesssim \|e_u^0\|_1^2 + \|e_p^0\|_d^2 + \int_0^{t_I} \|\rho_z\|_c^2 \, ds + \left(C_\tau^I\right)^2. \tag{5.34}$$

Finally, to establish (5.13), combine the definition of the weighted flux norm (3.9), (5.28), (5.29) (5.33), and (5.34) and use the fact that the integral from 0 to T majorizes all of the time-integral right-hand sides of the summed expressions. □

5.4 Convergence estimates

To specialize the general results of Proposition 5.1 we will first suppose the exact solutions to (3.2a)- suitable regularity assumptions. Moreover, we assume the interpolants, discussed in 5.1, satisfy approximation inequalities of a certain order. Towards that end let $U_h \times W_h \times Q_h$ satisfy the assumptions of Definition 4.1. For a reflexive Banach space X, a time interval $(a, b) \subseteq \mathbb{R}$ and a measurable $f : (a, b) \to X$ we define the canonical space-time norm [10]

$$\|f\|_{L^p(a,b;X)} = \left(\int_a^b \|f(s)\|_X \, ds\right)^{1/p}. \tag{5.35}$$

As in the case of spatial derivatives, the usual Sobolev notation $f \in H^r(a, b; X)$ means that $f \in L^2(a, b; X)$ and that $\partial_t f$, $\partial_t^2 f$, ..., $\partial_t^r f$ are also in $L^2(a, b; X)$. In the sections that follow we will sometimes use the abbreviations $\|f\|_{L^2 X}$ or $\|f\|_{H^r X}$ to signify (5.35).

Proposition 5.2 *Suppose the assumptions of Proposition 5.1 hold. Let $k \geq 0$ be the greatest integer such that the orthogonal projection, $\Pi_{Q_h} : Q \to Q_h$, satisfies (5.1). Suppose $r \geq 1$ is the maximal integer such that P_r, the space of continuous Lagrange polynomials of order r, is contained in U_h; suppose an interpolation, from W to W_h, satisfying either (5.2) or (5.3) exists and let $s > 0$ be the maximal integer satisfying the respective inequality. Suppose that the exact solutions to (3.2a)-(3.2c) satisfy the regularity assumptions*

$$u(t) \in L^\infty((0, T]; H^{r+1} \cap U)\partial_t u \in L^1((0, T]; H^{r+1} \cap U) \ \partial_{tt} u \in L^1((0, T]; H^1))$$
$$z(t) \in L^\infty((0, T]; H^{s+1} \cap W) \cap L^\infty((0, T]; H^{s+1}_{\kappa^{-1}} \cap W) \quad \partial_t z \in L^2((0, T]; H^{s+1}_{\kappa^{-1}} \cap W)$$
$$p(t) \in L^\infty((0, T]; H^{k+1} \cap Q) \quad\quad\quad\quad\quad\quad\quad \partial_{tt} p \in L^1((0, T]; L^1),$$

and that the initial iterates, (u_h^0, z_h^0, p_h^0), satisfy the estimates

$$\|u(0) - u_h^0\|_1 + \tau^{1/2}\|z(0) - z_h^0\|_c + \|p(0) - p_h^0\|_d$$
$$\lesssim \ h^r \|u(0)\|_{H^{r+1}} + \tau^{1/2}h^{s+1}\|z(0)\|_{\kappa^{-1}H^{s+1}} + h^{k+1}\|p(0)\|_{H^{k+1}}, \quad (5.36)$$

consistent with the projections of sect. 5.1. Then for $c = \min\{k, r, s\}$ we have

$$\|u^m - u_h^m\|_1 + \||z^m - z_h^m|\| + \|p^m - p_h^m\| \lesssim h^c M_1 + \tau M_2 \quad (5.37)$$

where M_1 and M_2 are given by

$$M_1 = h^{r-c} \left(\|\partial_t u\|_{L^1 H^{r+1}} + \|u\|_{L^\infty H^{r+1}} \right) + h^{s-c} \left(\|z\|_{L^1 H^{s+1}} \right.$$
$$\left. + (h + \tau^{1/2})\|z\|_{L^\infty H^{s+1}_{\kappa^{-1}}} \right) + h^{k-c}\|p\|_{L^\infty H^{k+1}}$$
$$M_2 = c_0 \|\partial_{tt} p\|_{L^1 L^1} + \|\partial_{tt} u\|_{L^1 H^1} + h^{s+1}\|\partial_t z\|_{L^2 H^{s+1}_{\kappa^{-1}}} + h^s\|z\|_{L^\infty H^{s+1}}$$

Proof First, note that since Π_{Q_h} satisfies (5.1) and since $P_r \subset U_h$ then, according to the argument directly preceding (5.5), the inequality (5.5) holds. Using the triangle inequality,

$$\|e_u^0\|_1 + \tau^{1/2}\|e_z^0\|_c + \|e_p^0\|_d \leq \|u(0) - u_h^0\|_1 + \tau^{1/2}\|z(0) - z_h^0\|_c$$
$$+ \|p(0) - p_h^0\|_d + \|\rho_u^0\|_1 + \|\rho_z^0\|_c + \|\rho_p^0\|_d,$$

along with (5.36) and the projection estimates of sect. 5.1, applied to the last three terms above, gives

$$\|e_u^0\|_1 + \tau^{1/2}\|e_z^0\|_c + \|e_p^0\|_d \lesssim \ h^r \|u(0)\|_{H^{r+1}} + \tau^{1/2}h^{s+1}\|z(0)\|_{\kappa^{-1}H^{s+1}}$$
$$+ h^{k+1}\|p(0)\|_{H^{k+1}}. \quad (5.38)$$

Then (5.37) follows from the triangle inequality, with respect to the error decomposition (5.6), along with: the discrete error estimates (5.13); discrete initial iterate error estimates (5.38); and interpolation estimates (5.1–5.4). □

Remark 5.1 Further assumptions on the discrete spaces, beyond the minimal Stokes–Biot stability of Defn. 4.1, can lead to slightly different versions of Proposition 5.2. For instance, if (W_h, Q_h) are such that the usual Raviart-Thomas type projection commutation relation

$$\left(\text{div } z - \text{div } \Pi_{W_h} z, q_h = 0\right), \quad \text{for all } q_h \in Q_h,$$

holds for each $z \in W$ then $(\text{div } \rho_z, q_h) = 0$ so that, for instance, the contribution $\|z\|_{L^1 H^{s+1}}$ vanishes from M_1; this term arises from $\|\text{div } \rho_z\|$ in (5.13). This observation is used in [30] where $W_h = RT_0$ is fixed.

6 Numerical experiments

Turning to numerical evaluation of our theoretical findings, we investigate the stability and numerical convergence properties, for $0 < \kappa \ll 1$ and $0 \le c_0 \le 1$, of two mixed finite element pairings:

(i) $U_h \times W_h \times Q_h = P_2^2 \times RT_0 \times DG_0$, and
(ii) $U_h \times W_h \times Q_h = P_2^2 \times P_1^2 \times DG_0$

The first discretization is a canonical choice from the original view of conforming Stokes–Biot [19,30] stability whereas the second choice is only minimally Stokes–Biot stable. We define a manufactured, smooth exact solution set over the unit square $\Omega = [0, 1] \times [0, 1]$, with coordinates $x = (x_1, x_2) \in \Omega$, given by

$$u(t, x) = \begin{pmatrix} t \sin(\pi x_1) \sin(\pi x_2) \\ 2t \sin(3\pi x_1) \sin(4\pi x_2) \end{pmatrix},$$

$$p(t, x) = (t + 1) \left(((x_1 - 1)x_1(x_2 - 1)x_2)^2 - \frac{1}{900} \right).$$

for $t \in (0, T), T = 1.0$. These solutions satisfy the homogeneous boundary conditions $u_{|\partial\Omega} = 0$ and $z_{|\partial\Omega} \cdot n = 0$ where n is the outward boundary normal to the unit square. By construction $p(t, \cdot) \in L_0^2(\Omega)$ for each t. For each discretization we consider three parameter scenarios: vanishing storage ($c_0 = 0$), fixed storage ($c_0 = 1$) and diminishing hydraulic conductivity ($\kappa \to 0$), and fixed hydraulic conductivity ($\kappa = 1.0$) and vanishing storage ($c_0 \to 0$). For simplicity, we here consider unit Lamé parameters: $\mu = \lambda = 1.0$. We let the time step size $\Delta t = T = 1.0$ as the test case is linear in time. For solving (3.2) numerically, we used the FEniCS finite element software suite [1]. The zero average-value condition on the pressure is enforced via a single real Lagrange multiplier. Linear systems were solved using MUMPS.

In Sects. 6.1 and 6.2 we examine the numerical errors for a fixed, minimally Stokes–Biot stable discretization on a series of uniform meshes, $\mathcal{T}_h$, with mesh size

h. Each of the error tables in Sects. 6.1 and 6.2 follow the same general format. In general, the relative displacement errors $\|\tilde{u}(T) - u_h(T)\|_1 / \|\tilde{u}(T)\|_1$ are reported in the first set of table rows; the relative pressure errors $\|\tilde{p}(T) - p_h(T)\| / \|\tilde{p}(T)\|$ appear in the second set of table rows; and the relative flux errors $\|\|\tilde{z}(T) - z_h(T)\|\| / \|\|\tilde{z}(T)\|\|$ appear in the final set of table rows. The last column of each table ('Rate') denotes the order of convergence using for the last two values in each row. For each investigation, either c_0 or κ varies while the other is fixed; the result of the variable parameter is reported for the values 10^{-r} for $r = 0, 4, 8$ and 12 but intermediate results identical to the previous case are suppressed. For instance, if $\kappa = 10^0$, $\kappa = 10^{-4}$, and $\kappa = 10^{-8}$ all yield the same errors for a given quantity, then only the errors for $\kappa = 10^0$ and $\kappa = 10^{-12}$ are reported in the corresponding row. In many instances, the displacement errors correspond directly to those of a previous case; in this event, we refer to the appropriate table.

6.1 Convergence of a Stokes–Biot stable pairing

We first consider the convergence properties for the pairing $U_h \times W_h \times Q_h = P_2^2(\mathcal{T}_h) \times RT_0(\mathcal{T}_h) \times DG_0(\mathcal{T}_h)$. This discretization satisfies the minimal Stokes–Biot stability conditions and also the Darcy stability condition when κ is uniformly bounded below. We note, however, that the Darcy stability condition fails to hold [2] uniformly as κ tends to zero but that Definition 4.1(iii) does indeed hold regardless of κ. As discussed in the previous section, we report on the relative approximation errors for the displacement, pressure, and flux for a series of uniform meshes $\mathcal{T}_h$ with mesh size h. The exact solutions $\tilde{u}, \tilde{p}, \tilde{z}$ were represented by continuous piecewise cubic interpolants in the error computations.

6.1.1 Vanishing storage $c_0 = 0$, varying conductivity $0 < \kappa \leq 1$

(see: Table 2) We observe that the displacement error converges at the expected and optimal rate (2) for κ ranging from 1 down to 10^{-12}. Overall, the displacement errors remain essentially unchanged as c_0 and κ vary. (We therefore do not report or discuss these further here.) The behaviour for the flux and pressure errors is less regular. The flux and pressure approximation errors increase as κ decreases, but seem to stabilize i.e. not increase substantially further from $\kappa = 10^{-4}$ to 10^{-8} and to 10^{-12}. Moreover, for each κ, the pressure and flux errors decrease with decreasing mesh size. Indeed, for $h = 1/128$, the pressure errors are of similar magnitude for the range of hydraulic conductivities (κ) tested. For a comparison to a minimally Stokes–Biot stable analogue, see Sect. 6.2.1 and Table 5.

6.1.2 Fixed storage $c_0 = 1$, varying conductivity $0 < \kappa \leq 1$

(see: Table 3) For this case, we again observe that the flux and pressure approximation errors increase as κ decrease, but seem to stabilize and not increase substantially further from $\kappa = 10^{-4}$ to 10^{-8} and 10^{-12}. Again, for each κ, the pressure and flux errors decrease with decreasing mesh size and for $h = 1/128$, the pressure errors are

Table 2 Vanishing storage coefficient $c_0 = 0$, varying conductivity $0 < \kappa \leq 1$ for the (minimally) Stokes–Biot stable pairing $P_2^2(\mathcal{T}_h) \times RT_0(\mathcal{T}_h) \times DG_0(\mathcal{T}_h)$

| κ | h | | | | | |
	1/8	1/16	1/32	1/64	1/128	Rate
Displacement						
10^0	1.64×10^{-1}	4.45×10^{-2}	1.13×10^{-2}	2.84×10^{-3}	7.11×10^{-4}	2.0
10^{-12}	1.64×10^{-1}	4.45×10^{-2}	1.13×10^{-2}	2.84×10^{-3}	7.11×10^{-4}	2.0
Pressure						
10^0	2.63×10^{-1}	1.02×10^{-1}	5.05×10^{-2}	2.53×10^{-2}	1.26×10^{-2}	1.0
10^{-4}	1.04×10^2	8.12	5.69×10^{-1}	4.39×10^{-2}	1.28×10^{-2}	1.8
10^{-8}	1.25×10^2	1.21×10^1	1.26	1.42×10^{-1}	2.07×10^{-2}	2.8
10^{-12}	1.25×10^2	1.21×10^1	1.26	1.43×10^{-1}	2.09×10^{-2}	2.8
Flux						
10^0	6.88×10^{-1}	1.41×10^{-1}	6.39×10^{-2}	3.18×10^{-2}	1.59×10^{-2}	1.0
10^{-4}	3.62×10^2	4.38×10^1	5.35	6.46×10^{-1}	8.05×10^{-2}	3.0
10^{-8}	4.72×10^2	9.91×10^1	2.67×10^1	7.01	1.76	2.0
10^{-12}	4.72×10^2	9.91×10^1	2.67×10^1	7.04	1.79	2.0

Relative approximation erros for the time-dependent test case given in Sect. 6. Listed are the relative displacement (top), relative flux (middle) and relative pressure (bottom) errors for varying κ on a series of uniform meshes $\mathcal{T}_h$ with mesh size h. The displacement errors for $\kappa = 10^{-4}, 10^{-8}$ were identical to the data presented ($\kappa = 1$, $\kappa = 10^{-12}$) and are suppressed. The last column 'Rate' denotes the order of convergence using for the last two values in each row. Compare with Table 5

nearly identical for the range of hydraulic conductivites (κ) tested. For comparison, see Sect. 6.2.2 and Table 6.

6.1.3 Fixed conductivity $\kappa = 1$, varying storage $0 \leq c_0 \leq 1$

(see: Table 4) For this case, we observe nearly uniform behaviour as c_0 decreases. The pressure and flux errors are similar for the range of storage coefficients (c_0) considered, and converge at the optimal and expected rate (1). For comparison, see Sect. 6.2.3 and Table 7

6.2 Convergence of a minimally Stokes–Biot stable pairing

We now turn to consider the convergence properties for the pairing $U_h \times W_h \times Q_h = P_2^2(\mathcal{T}_h) \times P_1^2(\mathcal{T}_h) \times DG_0(\mathcal{T}_h)$ and again report on the relative approximation errors for the displacement, pressure and flux. This pairing does not satisfy a Darcy stability condition, for any value of κ, as advanced in the original Stokes–Biot stability criteria; it does satisfy the minimally Stokes–Biot criterion of Definition 4.1. Numerical results for this minimally Stokes–Biot stable discretization, for the three paradigms considered in Sect. 6.1, are presented in 6.2.1– 6.2.3 alongside specific comparisons to the standard Stokes–Biot stable case.

Table 3 Fixed storage coefficient $c_0 = 1$, varying conductivity $0 < \kappa \leq 1$ for the (minimally) Stokes–Biot stable pairing $P_2^2(\mathcal{T}_h) \times RT_0(\mathcal{T}_h) \times DG_0(\mathcal{T}_h)$

Displacement, c.f. Table 2

κ	h					
	1/8	1/16	1/32	1/64	1/128	Rate
Pressure						
10^0	2.61×10^{-1}	1.02×10^{-1}	5.05×10^{-2}	2.53×10^{-2}	1.26×10^{-2}	1.0
10^{-4}	2.41×10^1	1.95	1.43×10^{-1}	2.64×10^{-2}	1.26×10^{-2}	1.1
10^{-8}	2.56×10^1	2.43	2.80×10^{-1}	4.17×10^{-2}	1.33×10^{-2}	1.7
10^{-12}	2.56×10^1	2.43	2.80×10^{-1}	4.17×10^{-2}	1.33×10^{-2}	1.7
Flux						
10^0	6.86×10^{-1}	1.41×10^{-1}	6.39×10^{-2}	3.18×10^{-2}	1.59×10^{-2}	1.0
10^{-4}	1.03×10^2	1.98×10^1	3.91	5.90×10^{-1}	7.87×10^{-2}	2.9
10^{-8}	1.07×10^2	2.35×10^1	6.56	1.75	4.46×10^{-1}	2.0
10^{-12}	1.07×10^2	2.35×10^1	6.56	1.75	4.47×10^{-1}	2.0

The format follows that of Table 2 and the relative displacement errors are identical
Compare with Table 6

Table 4 Fixed hydraulic conductivity $\kappa = 1$, varying storage $0 < c_0 \leq 1$ for the (minimally) Stokes–Biot stable pairing $P_2^2(\mathcal{T}_h) \times RT_0(\mathcal{T}_h) \times DG_0(\mathcal{T}_h)$

c_0	h					
	1/8	1/16	1/32	1/64	1/128	Rate
Displacement, c.f. Table 2						
Pressure						
10^0	2.61×10^{-1}	1.02×10^{-1}	5.05×10^{-2}	2.53×10^{-2}	1.26×10^{-2}	1.0
10^{-12}	2.45×10^{-1}	1.01×10^{-1}	5.05×10^{-2}	2.53×10^{-2}	1.26×10^{-2}	1.0
Flux						
10^0	6.86×10^{-1}	1.41×10^{-1}	6.39×10^{-2}	3.18×10^{-2}	1.59×10^{-2}	1.0
10^{-12}	6.88×10^{-1}	1.41×10^{-1}	6.39×10^{-2}	3.18×10^{-2}	1.59×10^{-2}	1.0

The format follows that of Table 2 and the relative displacement errors are identical. Compare with Table 7

The results of this comparison supply computational evidence that Definition 3.1(iii) can be replaced by Definition 4.1(iii) while retaining the convergence properties first observed in [17,30]. Since the Darcy stability of Definition 3.1(iii) is not satisfied [2] uniformly in κ, our observations strongly suggest that the minimal Stokes–Biot stability assumptions, specifically Definition 4.1(iii), are in fact, the key component for discretizations that retain their convergence properties as κ tends to zero.

6.2.1 Vanishing storage $c_0 = 0$, varying conductivity $0 < \kappa \leq 1$

(see: Table 5) Comparing Table 5 with Table 2, we observe that the performance of the two element pairings is almost surprisingly similar. Again, the displacement converges at the optimal and expected rate (2), the pressure and flux errors increase with decreasing κ, but stabilize, and converge with decreasing mesh size. We further observe that the relative errors for the flux for this element pairing is smaller than for the $P_2^2 \times RT_0 \times DG_0$ case (bottom rows). For a comparison to a discretization satisfying Darcy stability (though not uniformly in κ) see Sect. 6.1.1 and Table 2.

6.2.2 Fixed storage $c_0 = 1$, varying conductivity $0 < \kappa \leq 1$

(see: Table 6) Comparing Table 6 with Table 3, we again observe highly comparable performance. The observations made for the $P_2^2 \times RT_0 \times DG_0$ case thus also apply for $P_2^2 \times P_1^2 \times DG_0$. For comparison, see Sect. 6.1.2 and Table 3.

6.2.3 Fixed conductivity $\kappa = 1$, varying storage $0 \leq c_0 \leq 1$

(see: Table 7) For this case, we observe similar convergence rates as e.g. in Table 4). The pressure error increases very moderately with decreasing c_0 (it doubles as c_0 is reduced by 12 orders of magnitude), but both the pressure and flux converges at the optimal and expected rate (1). For comparison, see Sect. 6.1.3 and Table 4.

7 Conclusion

The important concept of Stokes–Biot stability, introduced independently by [19,22, 27,30], has proven a practical key to the selection of conforming Euler–Galerkin discretizations of Biot's Eq. (1.1) that retain their convergence properties as the hydraulic conductivity ($0 < \kappa$) becomes arbitrarily small. The novel contributions of this manuscript are primarily theoretical in nature; we have shown that the Stokes–Biot stability perspective can be, formally, relaxed and we have introduced the notion of a minimally Stokes–Biot stable Euler–Galerkin discretization. The stability of minimally Stokes–Biot stable schemes is independent of both c_0 and κ (c.f. Sect. 4.2, [19, Theorem 1] and [17, Theorem 3.2, Case I]), and we have presented a convergence analysis in this context.

In particular, we differ from previous authors [30] by carrying out our convergence analysis without the use of a Galerkin projection based on the Darcy problem. In doing so, we are able to depart from both the Darcy stability assumption, in general, and any questions regarding the appropriate norms for uniform-in-κ Darcy stability. In fact, an analysis based on a uniform-in-κ Darcy stability assumption should take into account a pressure-space norm exhibiting one of the forms discussed in Section 3.3; namely $L^2 + \kappa^{1/2} H^1$, $\kappa^{1/2} L^2$ or $\kappa^{1/2} H^1$. Each of these pressure norms have related difficulties over the usual pressure L^2 norm used here. First, it is not entirely clear to the authors that the $L^2 + \kappa^{1/2} H^1$ norm can be treated with the otherwise-standard arguments presented here and in related [17,19,22,30] work. Second, the $\kappa^{1/2} L^2$ and

Table 5 Vanishing storage coefficient $c_0 = 0$, varying conductivity $0 < \kappa \leq 1$ for the minimally Stokes–Biot stable pairing $P_2^2(\mathcal{T}_h) \times P_1^2(\mathcal{T}_h) \times DG_0(\mathcal{T}_h)$

κ	h					
	1/8	1/16	1/32	1/64	1/128	Rate
Displacement						
10^0	1.64×10^{-1}	4.45×10^{-2}	1.13×10^{-2}	2.84×10^{-3}	7.14×10^{-4}	2.0
10^{-12}	1.64×10^{-1}	4.45×10^{-2}	1.13×10^{-2}	2.84×10^{-3}	7.11×10^{-4}	2.0
Pressure						
10^0	7.83×10^1	1.35×10^1	5.30	2.64	1.34	1.0
10^{-4}	1.22×10^2	1.16×10^1	1.21	1.39×10^{-1}	2.07×10^{-2}	2.7
10^{-8}	1.25×10^2	1.21×10^1	1.26	1.43×10^{-1}	2.09×10^{-2}	2.8
10^{-12}	1.25×10^2	1.21×10^1	1.26	1.43×10^{-1}	2.09×10^{-2}	2.8
Flux						
10^0	6.11×10^{-1}	1.51×10^{-1}	7.23×10^{-2}	3.62×10^{-2}	1.81×10^{-2}	1.0
10^{-4}	1.39×10^2	1.52×10^1	1.54	1.27×10^{-1}	9.42×10^{-3}	3.7
10^{-8}	1.45×10^2	1.76×10^1	2.14	2.43×10^{-1}	2.87×10^{-2}	3.1
10^{-12}	1.45×10^2	1.76×10^1	2.14	2.43×10^{-1}	2.87×10^{-2}	3.1

Listed are the relative displacement (top), relative flux (middle) and relative pressure (bottom) errors for varying κ on a series of uniform meshes $\mathcal{T}_h$ with mesh size h

The displacement errors for $\kappa = 10^{-4}, 10^{-8}$ were identical to the data presented ($\kappa = 1, \kappa = 10^{-12}$) and are suppressed

The last column 'Rate' denotes the order of convergence using for the last two values in each row. Compare with Table 2

Table 6 Fixed storage coefficient $c_0 = 1$, varying conductivity $0 < \kappa \leq 1$ for the minimally Stokes–Biot stable pairing $P_1^2(\mathcal{T}_h) \times DG_0(\mathcal{T}_h)$

Displacement, c.f. Table 5

κ	h					
	1/8	1/16	1/32	1/64	1/128	Rate
Pressure						
10^0	1.55×10^1	3.15	1.33	6.67×10^{-1}	3.36×10^{-1}	1.0
10^{-4}	2.54×10^1	2.41	2.78×10^{-1}	4.16×10^{-2}	1.33×10^{-2}	1.6
10^{-8}	2.56×10^1	2.43	2.80×10^{-1}	4.17×10^{-2}	1.33×10^{-2}	1.7
10^{-12}	2.56×10^1	2.43	2.80×10^{-1}	4.17×10^{-2}	1.33×10^{-2}	1.7
Flux						
10^0	5.97×10^{-1}	1.49×10^{-1}	7.21×10^{-2}	3.61×10^{-2}	1.81×10^{-2}	1.0
10^{-4}	3.14×10^1	3.42	3.69×10^{-1}	3.40×10^{-2}	3.26×10^{-3}	3.4
10^{-8}	3.16×10^1	3.49	3.90×10^{-1}	3.98×10^{-2}	4.33×10^{-3}	3.2
10^{-12}	3.16×10^1	3.49	3.90×10^{-1}	3.98×10^{-2}	4.33×10^{-3}	3.2

The format follows that of Table 5 and the relative displacement errors are identical. Compare with Table 3

Table 7 Fixed hydraulic conductivity $\kappa = 1$, varying storage $0 < c_0 \leq 1$ for the minimally Stokes–Biot stable pairing $P_2^2(\mathcal{T}_h) \times P_1^2(\mathcal{T}_h) \times DG_0(\mathcal{T}_h)$

Displacement, c.f. Table 5

c_0	h					
	1/8	1/16	1/32	1/64	1/128	Rate
Pressure						
10^0	1.55×10^1	3.15	1.33	6.67×10^{-1}	3.36×10^{-1}	1.0
10^{-12}	7.83×10^1	1.35×10^1	5.30	2.64	1.34	1.0
Flux						
10^0	5.97×10^{-1}	1.49×10^{-1}	7.21×10^{-2}	3.61×10^{-2}	1.81×10^{-2}	1.0
10^{-12}	6.11×10^{-1}	1.51×10^{-1}	7.23×10^{-2}	3.62×10^{-2}	1.81×10^{-2}	1.0

Compare with Table 4

$\kappa^{1/2} H^1$ weightings both degenerate as κ becomes small and must be balanced by an appropriate displacement norm so that Stokes stability (Definition 3.1(i), (ii) and Definition 4.1(i), (ii)) holds uniformly in κ as well. Conversely, our arguments bring together many standard techniques and makes definitively clear, by abdicating Darcy stability, that one need not consider any κ-weighted norms for the pressure space in order to ensure stability and approximation as κ diminishes. Moreover, the convergence analysis presented here is the first instance, of which we are aware, of an analysis carried out in the context of the full norm used to prove the Banach-Nečas-Babuška stability (c.f. Sect. 4.2) of the Euler–Galerkin discretization (3.5). Thus, as neither the current convergence analysis, nor the previously-established BNB stability result, rely on a Darcy stability assumption, Proposition 5.2 solidifies, and generalizes, previous convergence estimates [30]. The concept of minimal Stokes–Biot stability therefore broadens the original view of Stokes–Biot stability to include alternative spaces that may not be Darcy stable; even for a fixed choice of κ.

Further observations and practical considerations

The primary contribution of the current work is theoretical in nature; we have, in practice, removed the Darcy restriction for Stokes–Biot stability and demonstrated an alternative convergence analysis in this context. Nevertheless, practical questions regarding the suitability of both Stokes–Biot and minimally Stokes–Biot stable approaches, solving (3.2), can be asked. In particular, we now briefly discuss: the drawbacks of Stokes–Biot and minimally Stokes–Biot stable discretizations; what computational advantages, if any, are granted by the minimal Stokes–Biot perspective; and alternatives to the boundary conditions (2.5).

Both Stokes–Biot and minimally Stokes–Biot stables are not without their drawbacks. The requirements of both definition 3.1 and definition 4.1 are general; however, in practice, both approaches typically make use of discontinuous pressures. This theme is present in the literature for both conformal and non-conformal discretizations. In practice, the need of a discontinuous pressure space imposes restrictions on the choice

of elements. In this manuscript we have used $P_2^d \times RT_0 \times DG_0$ and $P_2^d \times P_1^d \times DG_0$ discretizations to illustrate a simple comparison (c.f. Sect. 6) via numerical experiments in 2D. In two dimensions, as we discussed in Sect. 4, one could also consider pairings of Scott-Vogelius type, i.e. $P_k^d \times RT_m \times DG_{k-1}$ or $P_k^d \times P_m^d \times DG_{k-1}$ where $k \geq 4$. In the context of minimally Stokes–Biot stable triples, we have that the flux space degree can be chosen as $0 < m \leq k-1$ in both cases; from the original Stokes–Biot point of view, one would require $m = k-1$ for the case or Raviart-Thomas elements and polynomial fluxes would not be admissible at all. In 3D, one could also consider extensions of the Scott-Vogelius elements [15], the enriched cubic displacement element and piecewise constant pressures introduced by Guzmán and Neilan [14], the bubble-enriched continuous linear displacement element and piecewise constant pressures as in [30], or the related bubble-enriched continuous quadratic elements with discontinuous linear pressures [8]; these spaces could be considered alongside fluxes of Raviart-Thomas, Brezzi-Douglas-Marini, and Lagrange type. One may also consider quadrilateral meshes by using the Stokes pairing [8] given by $Q_2^2 \times DG_1$ along with, for instance, RT_m ($m = 0$ or 1), BDM_k or P_k^d ($k = 1$ or 2) fluxes.

It is practical to note that minimally Stokes–Biot stable discretizations do not necessarily confer a computational advantage over those with Darcy stable (for fixed κ) flux-pressuring pairings when equal-order fluxes are selected. That is to say, for instance, that the RT_k fluxes will have fewer DOFs than the alternative P_k^d fluxes discussed in this manuscript. However, minimal Stokes–Biot stability makes it clear that one can lower the order of the flux space without adversely impacting the stability and convergence of the method. One could interpret this as a form of 'computational advantage' of minimal Stokes–Biot. Overall, however, this is not the important point of minimal Stokes–Biot stability. The important points are that: Darcy stability is not necessary; typical 'Darcy stable pairings' satisfy the minimal Stokes–Biot criteria; and that approximation in both contexts yield strikingly similar results. Indeed, numerical experiments (Sect. 6) show similar errors both with (Table 2–Table 4) and without (Table 5–Table 7) a Darcy stability assumption; even as κ becomes very small. Moreover, we would not expect an improvement in results if the norms of the flux and pressure were altered to provide for uniform-in-κ Darcy stability condition (c.f. (A) and (B) of Sect. 3.3)); this is due to the fact that approximation of the pressure in the $L^2 + \kappa^{1/2} H^1$ norm is similar to that of the L^2 norm while approximation in the other option, $\kappa^{1/2} L^2$, degrades as κ becomes small. Thus, the tenets of minimal Stokes–Biot stability (Definition 4.1) provide an approximation of the pressure in the most sensible norm; that is, the L^2 norm is a fortuitous choice for convergence analysis, assures the proper context for Stokes stability, and does not degrade as κ becomes small. Our conclusion is that one can think, instead, in terms of Definition 4.1 (iii) when designing, or analyzing, approaches for Biot when $\kappa \to 0$. This important point could certainly impact the design, or choice, of discretizations that do in fact confer a computational advantage of those where Darcy stability is a requirement.

Finally, we close with a brief revisitation of the boundary conditions discussed in Sect. 2.5. Extending (2.5) to inhomogeneous data is not a concern; essential boundary conditions conditions can be lifted by selecting a particular solution and natural boundary conditions yield right-hand side terms that vanish in the error Eq. (5.10), and do not

alter the stability arguments (Sect. 4). However, it is valid to note that the assumption that $\Gamma_f = \Gamma_c$ and $\Gamma_t = \Gamma_p$, in (2.4), may not be practical for problems of interest. The conditions (2.5) were considered in the original Stokes–Biot, or motivating, literature [19,30] which lead to their adoption here. The advantage of the boundary conditions (2.5) is that they provide for an overall discussion that scopes naturally between the case where $\Gamma_c \neq \partial\Omega$ and $\Gamma_c = \partial\Omega$ provided that $|\Gamma_c| > 0$ is assumed. In particular, as mentioned in Remark 3.1, if $\Gamma_c = \partial\Omega$ then conditions (2.5) imply $\Gamma_t = \emptyset$, the variational forms (3.2) and (3.5) are unchanged, and all results discussed hold when $Q = L_0^2(\Omega)$ is selected, instead, in (3.1).

It is reasonable to ask what other boundary condition configurations can be considered, and under what conditions. First, we note that the requirement that $\Gamma_c \cap \Gamma_t = \emptyset$ and $\Gamma_f \cap \Gamma_p = \emptyset$ arise from the early work in well posedness for Biot [31]. Moreover, the requirement that the positive measure of the clamped displacement boundary is non-zero, i.e. that $|\Gamma_c| > 0$, provides the coercive property $a(u, u) \geq \gamma_a \|u\|_1^2$ needed by both Definition 3.1 (Stokes–Biot stability) and Definition 4.1 (minimal Stokes–Biot stability); this is therefore a strict requirement of the proposed method. However, if both $|\Gamma_c| > 0$ and $|\Gamma_t| > 0$ then, as noted in [18] and used in [22], the conditions (2.5) can be relaxed to those of (2.4). In this case, the requirements of both Definition 3.1 and Definition 4.1 can be satisfied with $Q = L_2(\Omega)$ in (3.1). In this case, the variational formulations (3.2) and (3.5) are, again, unaltered and the results of the manuscript follow analagously. It is true, as discussed above, that restrictive boundary conditions, i.e. such as (2.5), are needed when $\Gamma_c = \partial\Omega$ in order to ensure the tenets of both Definition 3.1 and Definition 4.1; this is not an additional imposition of minimal Stokes–Biot stability (Definition 4.1) but rather of the Stokes–Biot perspective in general.

References

1. Alnæs, M., Blechta, J., Hake, J., Johansson, A., Kehlet, B., Logg, A., Richardson, C., Ring, J., Rognes, M., Wells, G.: The FEniCS Project Version 1.5. Archive of Num. Soft. **3** (2015)
2. Bærland, T., Kuchta, M., Mardal, K.A., Thompson, T.: An observation on the uniform preconditioners for the mixed Darcy problem. Numer. Methods Partial Differ. Equ. **36**(6), 1718–1734 (2020). https://doi.org/10.1002/num.22500
3. Bergh, J., Löfström, J.: Interpolation Spaces: A Series of Comprehensive Studies in Mathematics. Springer, New York (1976)

4. Boffi, D., Brezzi, F., Fortin, M.: Mixed Finite Element Methods and Applications, 1st edn. Springer, Berlin (2013)
5. Braess, D.: Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics, 2nd edn. Cambridge University Press, Cambridge (2002)
6. Brezzi, F.: On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers. Publications mathématiques et informatique de Rennes **S4**, 1–26 (1974)
7. Brun, M.K., Ahmed, E., Berre, I., Nordbotten, J.M., Radu, F.A.: Monolithic and splitting based solution schemes for fully coupled quasi-static thermo-poroelasticity with nonlinear convective transport. arXiv preprint arXiv:1902.05783 (2019)
8. Ern, A., Guermond, J.L.: Theory and Practice of Finite Elements. Springer, Berlin (2004)
9. Ern, A., Meunier, S.: A posteriori error analysis of Euler–Galerkin approximations to coupled elliptic-parabolic problems. ESAIM: M2AN **43**(2), 353–375 (2009). https://doi.org/10.1051/m2an:2008048
10. Evans, L.: Partial Differential Equations. American Mathematical Society, Providence, R.I. (2010)
11. Girault, V., Wheeler, M.F., Almani, T., Dana, S.: A priori error estimates for a discretized poro-elastic-elastic system solved by a fixed-stress algorithm. Oil Gas Sci. Technol. Revue d'IFP Energies nouvelles **74**, 24 (2019)
12. Guo, L., Li, Z.: Ventikos, Yea: On the validation of a multiple-network poroelastic model using arterial spin labeling MRI data. Front. Comput. Neurosci. **13**, 60 (2019)
13. Guo, L., Vardakis, J., Ventikos, Yea: Subject-specific multi-poroelastic model for exploring the risk factors associated with the early stages of Alzheimer's disease. Interface Focus **8**(1), 20170,019 (2018)
14. Guzman, J., Neilan, M.: Conforming and divergence-free stokes elements in three dimensions. IMA J. Numer. Anal. **34**(4), 1489–1508 (2019). https://doi.org/10.1090/mcom/3346
15. Guzman, J., Scott, L.: The scott-vogelius finite elements revisited. Math. Comput. **88**, 515–529 (2019). https://doi.org/10.1090/mcom/3346
16. Herrmann, L.R.: Elasticity equations for incompressible and nearly incompressible materials by a variational theorem. AIAA J. **3**(10), 1896–1900 (1965)
17. Hong, Q., Kraus, J.: Parameter-robust stability of classical three-field formulation of Biot's consolidation model. Electron. T. Numer. Anal. **48**, 202–226 (2018)
18. Hong, Q., Kraus, J., Lymbery, M., Wheeler, M.F.: Parameter-robust convergence analysis of fixed-stress split iterative method for multiple-permeability poroelasticity systems. Multiscale Model. Simul. **18**(2), 916–941 (2020)
19. Hu, X., Rodrigo, C., Gaspar, F.J., Zikatanov, L.: A nonconforming finite element method for the Biot's consolidation model in poroelasticity. J. Comput. Appl. Math. **310**, 143–154 (2017)
20. Kraus, J., Lederer, P., Lymbery, M., Schoberl, J.: Uniformly well-posed hybridized discontinuous Galerkin/hybrid mixed discretizations for Biot's consolidation model. Cold Spring Harbor Lab. (preprint) arXiv:2012.08584 (2020)
21. Kumar, S., Oyarzúa, R., Ruiz-Baier, R., Sandilya, R.: Conservative discontinuous finite volume and mixed schemes for a new four-field formulation in poroelasticity. ESAIM Math. Model. Numer. Anal. **54**(1), 273–299 (2020)
22. Lee, J.: Robust three-field finite element methods for Biot's consolidation model in poroelasticity. BIT Numer. Math. **58**(2), 347–372 (2018)
23. Lee, J., Mardal, K.A., Winther, R.: Parameter-robust discretization and preconditioning of Biot's consolidation model. SIAM J. Scie. Comput. **39**(1), A1–A24 (2017)
24. Lee, J., Piersanti, E., Mardal, K.A., Rognes, M.: A mixed finite element method for nearly incompressible multiple-network poroelasticity. SIAM J. Sci. Comput. **41**(2), A722–A747 (2019)
25. Li, X., Holst, H., Kleiven, S.: Influences of brain tissue poroelastic constants on intracranial pressure (ICP) during constant-rate infusion. Comput. Methods Biomech. Biomed. Eng. **16**(12), 1330–1343 (2013)
26. Lipnikov, K.: Numerical methods for the Biot model in poroelasticity. Ph.D. thesis, University of Houston (2002)
27. Lotfian, Z., Sivaselvan, M.: Mixed finite element formulation for dynamics of porous media. Int. J. Numer. Methods Eng. **115**, 141–171 (2018)
28. Oyarzúa, R., Ruiz-Baier, R.: Locking-free finite element methods for poroelasticity. SIAM J. Numer. Anal. **54**(5), 2951–2973 (2016)
29. Riviere, B.: Discontinuous Galerkin methods for Solving Elliptic and Parabolic Equations: Theory and Implementation. Society for Industrial and Applied Mathematics (2008)

30. Rodrigo, C., Hu, X., Ohm, P., Adler, J.H., Gaspar, F.J., Zikatanov, L.: New stabilized discretizations for poroelasticity and the Stokes' equations. Comput. Methods Appl. Mech. Eng. **341**, 467–484 (2018)
31. Showalter, R.: Diffusion in poro-elastic media. J. Math. Anal. Appl. **24**(251), 310–340 (2000)
32. Storvik, E., Both, J.W., Kumar, K., Nordbotten, J.M., Radu, F.A.: On the optimization of the fixed-stress splitting for biot's equations. Int. J. Numer. Methods Eng. **120**(2), 179–194 (2019)
33. Thompson, T., Riviere, B., Knepley, M.: An implicit discontinuous galerkin method for modeling acute edema and resuscitation in the small intestine. Math Med. Biol. **36**(4), 513–548 (2019)
34. Young, J., Riviere, B.: A mathematial model of intestinal oedema formation. Math. Med. Biol. **31**(1), 1–15 (2014)
35. Zenisek, A.: The existence and uniqueness theorem in Biot's consolidation theory. Aplikace Matematiky **29**(3), 194–211 (1984)