ESAIM: M2AN 43 (2009) 353–375
DOI: 10.1051/m2an:2008048

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

A POSTERIORI ERROR ANALYSIS OF EULER-GALERKIN APPROXIMATIONS TO COUPLED ELLIPTIC-PARABOLIC PROBLEMS *

Alexandre Ern[1] and Sébastien Meunier[2]

Abstract. We analyze Euler-Galerkin approximations (conforming finite elements in space and implicit Euler in time) to coupled PDE systems in which one dependent variable, say u, is governed by an elliptic equation and the other, say p, by a parabolic-like equation. The underlying application is the poroelasticity system within the quasi-static assumption. Different polynomial orders are used for the u- and p-components to obtain optimally convergent *a priori* bounds for all the terms in the error energy norm. Then, a residual-type *a posteriori* error analysis is performed. Upon extending the ideas of Verfürth for the heat equation [*Calcolo* **40** (2003) 195–212], an optimally convergent bound is derived for the dominant term in the error energy norm and an equivalence result between residual and error is proven. The error bound can be classically split into time error, space error and data oscillation terms. Moreover, upon extending the elliptic reconstruction technique introduced by Makridakis and Nochetto [*SIAM J. Numer. Anal.* **41** (2003) 1585–1594], an optimally convergent bound is derived for the remaining terms in the error energy norm. Numerical results are presented to illustrate the theoretical analysis.

Mathematics Subject Classification. 65M60, 65M15, 74F10.

Received July 23, 2007.
Published online December 5, 2008.

1. Introduction

The main motivation for this work is poroelasticity problems with hydro-mechanical couplings. We consider a linearly elastic porous medium Ω saturated by a slightly compressible and viscous fluid within the so-called quasi-static assumption in which inertia effects in the elastic structure are negligible. Given a simulation time $T > 0$, the problem consists in finding a displacement field $u : [0, T] \times \Omega \to \mathbb{R}^3$ and a pressure field $p : [0, T] \times \Omega \to \mathbb{R}$ such that

$$-\nabla\cdot\sigma(u) + b\nabla p = f, \qquad \text{in } [0, T] \times \Omega, \tag{1.1}$$

$$\partial_t(\tfrac{1}{M}p + b\nabla\cdot u) - \nabla\cdot(\kappa\nabla p) = g, \qquad \text{in } [0, T] \times \Omega. \tag{1.2}$$

Keywords and phrases. Finite element method, energy norm, *a posteriori* error analysis, hydro-mechanical coupling, poroelasticity.

* *This work was partially supported by EdF R&D and the Groupement MOMAS (PACEN/CNRS, ANDRA, BRGM, CEA, EdF, IRSN). The second author was supported by a CIFRE Ph.D. fellowship.*

[1] Université Paris-Est, CERMICS, École des Ponts, 77455 Marne-la-Vallée Cedex 2, France. ern@cermics.enpc.fr

[2] EDF R&D, 1 avenue du Général de Gaulle, 92141 Clamart Cedex, France. sebastien.meunier@edf.fr

Article published by EDP Sciences

Here, $\sigma(u) = 2\lambda_1\varepsilon(u) + \lambda_2(\nabla\cdot u)I$ is the so-called effective stress tensor, $\varepsilon(u) = \frac{1}{2}(\nabla u + (\nabla u)^t)$ the linearized strain tensor, λ_1 and λ_2 the Lamé coefficients assumed to be positive constants, I the identity matrix in $\mathbb{R}^3$, b the Biot-Willis coefficient, M the Biot modulus, κ the permeability of the medium, while f and g are given data, typically the volumetric body forces and their divergence, respectively. The system (1.1)–(1.2) is supplemented with initial and boundary conditions discussed below. The poroelasticity system can be traced back to the pioneering work of von Terzaghi [25] and Biot [4]. Equations (1.1)–(1.2) respectively express the balance of momentum and the conservation of mass. The quasi-static assumption means that the term $\rho\partial_{tt}u$ (where ρ denotes the density of the elastic structure) has been neglected in the momentum balance. The Biot modulus combines compressibility and porosity effects; it is often assumed to be very large when dealing with the so-called Biot's consolidation problem, but this assumption will not be made here. For the sake of simplicity, we will assume that the coefficients b and M are given constants. A mathematical analysis of the system (1.1)–(1.2), including existence and uniqueness of strong and weak solutions based on the theory of linear degenerate evolution equations in Hilbert spaces, has been carried out by Showalter [19,20]. Boundary conditions can be prescribed by considering two partitions of the boundary. The first partition is used for the displacement field (either the displacement itself or a traction force is prescribed), while the other partition is used for the pressure field (either the pressure itself or a flux is prescribed). For the sake of simplicity, we assume here that any portion of the boundary is clamped or drained, *i.e.* at least a Dirichlet condition is enforced on the displacement or on the pressure everywhere. Furthermore, an initial condition must be enforced on the quantity $\frac{1}{M}p + b\nabla\cdot u$. Although the evolution problem related to (1.1)–(1.2) is essentially of parabolic type under minimal smoothness requirements on the data, we refer to it as a coupled elliptic-parabolic problem to stress the fact that equation (1.1) is of elliptic type for the displacement and equation (1.2) is of parabolic type for the pressure.

In the present work, we assume that the data (including boundary and initial conditions) are smooth enough for a strong solution to exist up to initial time, and we are concerned with the analysis of Euler-Galerkin approximations using a backward Euler scheme in time and conforming finite elements in space. The *a priori* analysis of Euler-Galerkin approximations for Biot's consolidation problem is covered in the work of Murad, Loula, and coworkers [14–16], including the semi-discrete and fully discrete cases and long-time behavior. The problem under scrutiny here is somewhat different since we do not assume that the Biot modulus takes very large values, *i.e.* we do not discard the pressure time-derivative in (1.2). As a result, we shall briefly address below the *a priori* error analysis of the Euler-Galerkin approximation to the evolution problem (1.1)–(1.2). The energy norm associated with the present problem controls the $L^\infty(0,t;H^1_x)$-norm (L^∞ in time and H^1 in space) of the displacement and the $L^\infty(0,t;L^2_x)$- and $L^2(0,t;H^1_x)$-norms of the pressure. This implies that error bounds with optimal convergence orders in space require the use of different polynomial degrees in the finite element spaces for the displacement and for the pressure, namely one degree higher for the displacement than for the pressure. The technique of Wheeler [26] originally designed to obtain optimally convergent $L^\infty(0,t;L^2_x)$ *a priori* error bounds for the heat equation can be adapted to the present framework. The same technique has already been used in [14–16] for Biot's consolidation problem. Note that the use of different polynomial degrees for the displacement and the pressure stems here solely from the derivation of optimally convergent error bounds, and that using equal-order polynomials still yields a stable discrete problem even though the underlying Stokes problem associated with the elimination of the displacement is not stable.

The *a posteriori* error analysis of evolution problems related to poroelasticity is a much less explored field. In the present work, we derive *a posteriori* energy-norm error bounds of residual type. Two approaches are undertaken. The first one uses standard energy techniques and yields an error bound which is similar to those previously derived for the heat equation by Picasso [17], Verfürth [24], Chen and Feng [6], and Bergam *et al.* [3]. The error bound comprises a term associated with time errors (evaluated from the pressure differences at two consecutive time steps), one associated with space errors (evaluated from the finite element residuals for the displacement and for the pressure) and a data oscillation term. Furthermore, taking inspiration from the work of Verfürth [24], an equivalence result is established between the residual measured in a suitable dual norm and the error measured in the energy norm supplemented with some time-derivatives measured in weaker norms. A nontrivial novelty with respect to the heat equation is the use of a $L^1(0,t;H^{-1}_x)$-norm for the time-derivative of the displacement equation. A further important difference is that owing to the use of different polynomial orders for the displacement and for the pressure, the convergence rate with respect to mesh size of the derived bound is not optimally convergent with respect to all of the terms in the energy norm, in particular those concerning the displacement. To tackle this difficulty, we make use of the elliptic reconstruction technique introduced for linear parabolic problems by Makridakis and Nochetto [12] and further analyzed by Lakkis and Makridakis [11]. This technique, which can be regarded as the counterpart of the elliptic projection method introduced by Wheeler for the *a priori* error analysis, was designed to obtain optimally convergent *a posteriori* error bounds in the $L^\infty(0,t;L^2_x)$-norm (and other higher-order norms) for linear parabolic problems. Another novelty of the present work is to extend this technique to coupled elliptic-parabolic problems such as the poroelasticity equations to derive error bounds for the displacement exhibiting optimal convergence behavior with respect to mesh size. This extension is not straightforward since we shall see that the reconstructed fields at the continuous level must still depend on some time-derivative of the displacement to account for the fact that the divergence of the discrete displacement does not belong to the pressure finite element space. Other approaches to derive $L^\infty(0,t;L^2_x)$-norm *a posteriori* error bounds for parabolic equations can be found, among others, in the work of Eriksson and Johnson [8,9] and of Thomée [21] based on duality techniques and in the work of Babuška *et al.* [1] using a double integration in time. Finally, we observe that an alternative approach to derive *a posteriori* error bounds is the use of goal-oriented, duality-weighted techniques; see, *e.g.*, Becker and Rannacher [2] for a thorough description and various applications.

This paper is organized as follows. Section 2 presents the setting under scrutiny in an abstract framework. The advantage of working with an abstract setting rather than with the poroelasticity system is that it allows more generality on the problem to be treated (easily incorporating for instance various boundary conditions) and that it allows to identify more clearly the main arguments in the proofs. Furthermore, although the displacement can be eliminated to yield a parabolic-type evolution equation for the sole pressure (see Sect. 2.4), we have chosen to work with the mixed pressure-displacement formulation since the displacement is often an important variable in poroelasticity applications. Section 3 is devoted to the *a priori* error analysis. The main result is Theorem 3.1. Section 4 deals with the *a posteriori* error analysis. The main results are Theorems 4.1 and 4.2 for the direct approach and Theorem 4.3 for the approach using elliptic reconstruction. Section 5 contains numerical results illustrating the error analysis. Finally, Section 6 draws some conclusions.

2. The setting

2.1. The continuous problem

Let V_a and V_d be two Hilbert spaces respectively equipped with symmetric, continuous and coercive bilinear forms a and d. The norms induced by these forms are denoted by $\|\cdot\|_a$ and $\|\cdot\|_d$ respectively. Let V_a' (resp., V_d') be the dual space of V_a (resp., V_d) with duality product denoted by $\langle\cdot,\cdot\rangle_a$ (resp., $\langle\cdot,\cdot\rangle_d$) and norm $\|\cdot\|_{a'} = \sup_{0\neq v\in V_a} |\langle\cdot,v\rangle_a|/\|v\|_a$ (resp., $\|\cdot\|_{d'} = \sup_{0\neq q\in V_d} |\langle\cdot,q\rangle_d|/\|q\|_d$). Let L_a (resp., L_d) be a Hilbert space equipped with a scalar product $(\cdot,\cdot)_{L_a}$ (resp., $(\cdot,\cdot)_{L_d}$) with dense and continuous injection $V_a \hookrightarrow L_a$ (resp., $V_d \hookrightarrow L_d$); in particular, let $\tilde\gamma$ be such that for all $q \in V_d$, $\|q\|_{L_d} \le \tilde\gamma\|q\|_d$. Identifying L_a (resp., L_d) with its dual space, it is inferred that $V_a \hookrightarrow L_a \equiv L_a' \hookrightarrow V_a'$ (resp., $V_d \hookrightarrow L_d \equiv L_d' \hookrightarrow V_d'$). Moreover, let c be a symmetric, continuous and coercive bilinear form defined over $L_d \times L_d$ inducing a norm $\|\cdot\|_c$; in particular, let $\hat\gamma$ be such that for all $q \in L_d$, $\|q\|_c \le \hat\gamma\|q\|_{L_d}$, so that for all $q \in V_d$, $\|q\|_c \le \gamma\|q\|_d$ with $\gamma := \hat\gamma\tilde\gamma$. Finally, let b be a continuous bilinear form defined over $V_a \times L_d$ with continuity constant β, *i.e.*, for all $(v,q) \in V_a \times L_d$, $|b(v,q)| \le \beta\|v\|_a\|q\|_c$.

The elements of the spaces defined above are functions of the space variable x. In the sequel, we shall deal with functions of time and space. The time variable varies over the interval $[0,T]$ for a fixed $T > 0$. Henceforth, $L^p(0,T;Z)$, $p \in [1,+\infty]$, denotes the vector space of functions f in space and time such that for a.e. $t \in [0,T]$, $f(t) := f(t,\cdot)$ is in Z (where Z denotes any of the spaces defined above) and $\int_0^T \|f(s)\|_Z^p \mathrm{d}s < +\infty$ if $p \neq +\infty$ or $\sup_{s\in[0,T]} \|f(s)\|_Z < +\infty$ if $p = +\infty$. Similarly, $H^1(0,T;Z)$ denotes the subspace of $L^2(0,T;Z)$ consisting in functions f with square integrable distributional time-derivative $\partial_t f$ over $[0,T]$. Functions in $H^1(0,T;Z)$ admit pointwise values in Z for all $t \in [0,T]$. Whenever $Z \equiv L_d$, the space Z is equipped with the norm $\|\cdot\|_c$.

Given data $f \in H^1(0,T;L_a)$, $g \in H^1(0,T;L_d)$, and $p_0 \in V_d$, we seek for the strong solution $(u,p) \in H^1(0,T;V_a) \times H^1(0,T;V_d)$ such that for a.e. $t \in [0,T]$,

$$a(u,v) - b(v,p) = \langle f, v\rangle_a, \qquad \forall v \in V_a, \tag{2.1}$$
$$c(\partial_t p, q) + b(\partial_t u, q) + d(p,q) = \langle g, q\rangle_d, \qquad \forall q \in V_d, \tag{2.2}$$

completed with the initial condition $p(0) := p_0$. Note that in the present setting, equation (2.1) holds up to $t = 0$, thus uniquely determining the initial value of u in terms of p_0 and $f(0)$. Letting $u_0 := u(0) \in V_a$, the *a priori* bound $\|u_0\|_a \le \beta\|p_0\|_c + \|f(0)\|_{a'}$ is readily inferred by taking $v := u_0$ in (2.1).

Our first result is an *a priori* estimate for the strong solution. The energy norm for problem (2.1)–(2.2) is defined for all $t \in [0,T]$ as

$$\|(u,p)\|^2_{X(0,t)} = \|u\|^2_{L^\infty(0,t;V_a)} + \|p\|^2_{L^\infty(0,t;L_d)} + \|p\|^2_{L^2(0,t;V_d)}. \tag{2.3}$$

Proposition 2.1. *The following holds for a.e.* $t \in [0,T]$,

$$\tfrac{1}{2}\|(u,p)\|^2_{X(0,t)} \le 2\left(2\|f\|_{L^\infty(0,T;V_a')} + \|\partial_t f\|_{L^1(0,T;V_a')}\right)^2 + \|g\|^2_{L^2(0,T;V_d')} + \|u_0\|^2_a + \|p_0\|^2_c. \tag{2.4}$$

Proof. Since (u,p) is a strong solution, for a.e. $t \in [0,T]$, $v = \partial_t u$ is in V_a and $q = p$ is in V_d. Using these test functions in (2.1)–(2.2) yields

$$\tfrac{1}{2}d_t\|u\|^2_a + \tfrac{1}{2}d_t\|p\|^2_c + \|p\|^2_d = \langle f, \partial_t u\rangle_a + \langle g, p\rangle_d.$$

Hence,

$$\tfrac{1}{2}d_t\|u\|^2_a + \tfrac{1}{2}d_t\|p\|^2_c + \tfrac{1}{2}\|p\|^2_d \le \langle f, \partial_t u\rangle_a + \tfrac{1}{2}\|g\|^2_{d'}.$$

Let $t \in (0,T)$. Integrating the above inequality over $(0,t)$ and integrating by parts in time the term $\langle f, \partial_t u\rangle_a$ yields

$$\tfrac{1}{2}\|(u,p)\|^2_{X(0,t)} \le \langle f(t), u(t)\rangle_a - \langle f(0), u(0)\rangle_a - \int_0^t \langle \partial_t f(s), u(s)\rangle_a \mathrm{d}s + \tfrac{1}{2}\|g\|^2_{L^2(0,T;V_d')} + \tfrac{1}{2}\|u_0\|^2_a + \tfrac{1}{2}\|p_0\|^2_c.$$

As a result, it is inferred that for a.e. $t \in [0,T]$,

$$\tfrac{1}{2}\|(u,p)\|^2_{X(0,t)} \le C_1\|u\|_{L^\infty(0,T;V_a)} + C_2,$$

with $C_1 = 2\|f\|_{L^\infty(0,T;V_a')} + \|\partial_t f\|_{L^1(0,T;V_a')}$ and $C_2 = \frac{1}{2}\|g\|^2_{L^2(0,T;V_d')} + \frac{1}{2}\|u_0\|^2_a + \frac{1}{2}\|p_0\|^2_c$. Hence,

$$\tfrac{1}{2}\|u\|^2_{L^\infty(0,T;V_a)} \le C_1\|u\|_{L^\infty(0,T;V_a)} + C_2,$$

so that $\|u\|^2_{L^\infty(0,T;V_a)} \le 4(C_1^2 + C_2)$. This yields $\frac{1}{2}\|(u,p)\|^2_{X(0,t)} \le 2(C_1^2 + C_2)$, *i.e.*, (2.4). □

An important consequence of Proposition 2.1 is the uniqueness of the strong solution of (2.1)–(2.2). In the sequel, we assume the existence of the strong solution.

Remark 2.1. Proposition 2.1 holds in the slightly more general setting where $f \in H^1(0,T;V_a')$, $g \in L^2(0,T;V_d')$, $p_0 \in L_d$, and $p \in L^2(0,T;V_d) \cap H^1(0,T;L_d)$.

Application to poroelasticity. The evolution problem (1.1)–(1.2) fits the present setting. For the sake of simplicity, we consider homogeneous Dirichlet boundary conditions both for the displacement and the pressure. Then, letting

$$V_a = [H^1_0(\Omega)]^3, \qquad L_a = [L^2(\Omega)]^3, \qquad V_d = H^1_0(\Omega), \qquad L_d = L^2(\Omega), \tag{2.5}$$

we define the bilinear forms

$$a(u,v) = \int_\Omega \sigma(u){:}\varepsilon(v), \qquad b(v,p) = \int_\Omega bp\nabla{\cdot}v, \tag{2.6}$$

$$c(p,q) = \int_\Omega \tfrac{1}{M}pq, \qquad d(p,q) = \int_\Omega \kappa\nabla p{\cdot}\nabla q. \tag{2.7}$$

The coercivity of a on $V_a \times V_a$ (resp., d on $V_d \times V_d$) results from Korn's First Inequality (resp., Poincaré's Inequality). The bilinear form b is clearly continuous on $V_a \times L_d$ with $\beta = b(M/\lambda_2)^{1/2}$. Proposition 2.1 means that the displacement is controlled in the $L^\infty(0,t;H^1_x)$-norm and that the pressure is controlled in the $L^\infty(0,t;L^2_x)$- and $L^2(0,t;H^1_x)$-norms.

2.2. The discrete problem

Problem (2.1)–(2.2) is approximated by an Euler-Galerkin scheme, namely conforming finite elements in space and an implicit Euler scheme in time. Let $\{V_{ah}\}_{h>0}$ and $\{V_{dh}\}_{h>0}$ be two families of finite-dimensional subspaces of V_a and V_d respectively. The parameter h refers to the size of an underlying mesh family denoted by $\{\mathcal{T}_h\}_{h>0}$. Let $0 = t_0 < t_1 < \ldots < t_N = T$ be a sequence of discrete times and for all $n \in \{1,\ldots,N\}$, set $\tau_n = t_n - t_{n-1}$ and $I_n = (t_{n-1}, t_n)$. Henceforth, a superscript n indicates the value taken by any function of space and time at the discrete time t_n. For instance, $u^n := u(t_n) \in V_a$. For the sake of simplicity, we have chosen to restrict ourselves to fixed meshes and to postpone the study of discretizations with time-dependent meshes to future work.

The discrete problem consists in seeking $\{u_h^n\}_{n=1}^N \in [V_{ah}]^N$ and $\{p_h^n\}_{n=1}^N \in [V_{dh}]^N$ such that for all $n \in \{1,\ldots,N\}$,

$$a(u_h^n, v_h) - b(v_h, p_h^n) = (f_h^n, v_h)_{L_a}, \qquad \forall v_h \in V_{ah}, \tag{2.8}$$

$$c(\delta_t p_h^n, q_h) + b(\delta_t u_h^n, q_h) + d(p_h^n, q_h) = (g_h^n, q_h)_{L_d}, \qquad \forall q_h \in V_{dh}, \tag{2.9}$$

where $\delta_t p_h^n = \tau_n^{-1}(p_h^n - p_h^{n-1})$ and $\delta_t u_h^n = \tau_n^{-1}(u_h^n - u_h^{n-1})$. Given a pair $(u_{0h}, p_{0h}) \in V_{ah} \times V_{dh}$, the initial condition is $(u_h^0, p_h^0) := (u_{0h}, p_{0h})$. The data $\{f_h^n\}_{n=1}^N \in [L_a]^N$ and $\{g_h^n\}_{n=1}^N \in [L_d]^N$ are approximations of $\{f^n\}_{n=1}^N$ and $\{g^n\}_{n=1}^N$ respectively.

Lemma 2.1. *The discrete problem is well-posed.*

Proof. For all $n \in \{1,\ldots,N\}$, equations (2.8)–(2.9) yield a square linear system for the components of (u_h^n, p_h^n) once bases of V_{ah} and V_{dh} are chosen, so it suffices to prove the uniqueness of the discrete solution. Testing with $v_h = u_h^n - u_h^{n-1}$ and $q_h = \tau_n p_h^n$ and using the fact that $a(x,x-y) = \frac{1}{2}a(x,x) + \frac{1}{2}a(x-y,x-y) - \frac{1}{2}a(y,y)$ owing to the symmetry of the bilinear form a (along with a similar property for the bilinear form c) yields

$$\tfrac{1}{2}\|u_h^n\|_a^2 + \tfrac{1}{2}\|u_h^n - u_h^{n-1}\|_a^2 + \tfrac{1}{2}\|p_h^n\|_c^2 + \tfrac{1}{2}\|p_h^n - p_h^{n-1}\|_c^2 + \tau_n\|p_h^n\|_d^2 = \tfrac{1}{2}\|u_h^{n-1}\|_a^2 + (f_h^n, u_h^n - u_h^{n-1})_{L_a} + \tfrac{1}{2}\|p_h^{n-1}\|_c^2 + \tau_n(g_h^n, p_h^n)_{L_d}.$$

Hence,

$$\tfrac{1}{2}\|u_h^n\|_a^2 + \tfrac{1}{2}\|p_h^n\|_c^2 + \tfrac{1}{2}\tau_n\|p_h^n\|_d^2 \le \tfrac{1}{2}\|u_h^{n-1}\|_a^2 + \tfrac{1}{2}\|f_h^n\|_{a'}^2 + \tfrac{1}{2}\|p_h^{n-1}\|_c^2 + \tfrac{1}{2}\tau_n\|g_h^n\|_{d'}^2.$$

This shows the uniqueness of the solution of the square linear system. □

The proof of Lemma 2.1 hints at how the discrete scheme (2.8)–(2.9) could be modified if time-dependent meshes were used. In this case, we work with two families $\{V_{ah}^n\}_{n=0}^N$ and $\{V_{dh}^n\}_{n=0}^N$ of finite-dimensional subspaces such that for all $n \in \{0,\ldots,N\}$, $V_{ah}^n \subset V_a$ and $V_{dh}^n \subset V_d$. The discrete scheme takes the general form (2.8)–(2.9) with test functions $v_h \in V_{ah}^n$ and $q_h \in V_{dh}^n$. However, if the expression for the time-derivative of the displacement is kept unchanged, the argument deployed in the above proof breaks down because it is no longer possible to use

$v_h = u_h^n - u_h^{n-1}$ as a test function since in general $v_h \notin V_{ah}^n$ (unless the restrictive assumption $V_{ah}^{n-1} \subset V_{ah}^n$ is made for all $n \in \{1, \dots, N\}$). To circumvent this difficulty, let $\mathfrak{R}_{ah}^{n*} : V_a \to V_{ah}^n$ be the Riesz projection operator defined such that for all $v \in V_a$,

$$a(v - \mathfrak{R}_{ah}^{n*}(v), v_h) = 0, \qquad \forall v_h \in V_{ah}^n, \tag{2.10}$$

and use $\delta_t u_h^n = \tau_n^{-1}(u_h^n - \mathfrak{R}_{ah}^{n*}(u_h^{n-1}))$ in (2.9). Then, proceeding as in the proof of Lemma 2.1 with the test functions $v_h = u_h^n - \mathfrak{R}_{ah}^{n*}(u_h^{n-1}) \in V_{ah}^n$ and $q_h = \tau_n p_h^n \in V_{dh}^n$ and observing that $a(u_h^n, v_h) = a(u_h^n, u_h^n - u_h^{n-1})$ since $u_h^n \in V_{ah}^n$, the same stability result is recovered, and thus well-posedness.

2.3. **Continuous and discrete differential operators**

To formulate the *a posteriori* error bounds in the usual form, it is convenient to associate differential operators (in space) with the bilinear forms a, b, c, and d. To this purpose, we define the following continuous operators: $A \in \mathcal{L}(V_a; V_a')$ s.t. $\langle Av, w\rangle_a = -a(v, w)$, $B \in \mathcal{L}(V_a; L_d)$ s.t. $(Bv, q)_{L_d} = b(v, q)$ with (formal) adjoint $B^* \in \mathcal{L}(V_d; L_a)$ s.t. $(B^*q, v)_{L_a} = b(v, q)$, $C \in \mathcal{L}(L_d; L_d)$ s.t. $(Cq, r)_{L_d} = c(q, r)$, and $D \in \mathcal{L}(V_d; V_d')$ s.t. $\langle Dq, r\rangle_d = -d(q, r)$. We use the same notation for the extension $B^* \in \mathcal{L}(L_d; V_a')$ s.t. $\langle B^*q, v\rangle_a = b(v, q)$. Problem (2.1)–(2.2) can be rewritten in the form

$$-Au - B^*p = f, \tag{2.11}$$

$$C\partial_t p + B\partial_t u - Dp = g, \tag{2.12}$$

these equalities holding for a.e. $t \in [0, T]$ in V_a' and V_d' respectively. For the poroelasticity system, $Av = \nabla\cdot\sigma(v)$, $Bv = b\nabla\cdot v$, $B^*q = -b\nabla q$, $Cq = \frac{1}{M}q$, and $Dq = \nabla\cdot(\kappa\nabla p)$. In the simplified setting where all the physical parameters are equal to unity, we obtain

$$A = \Delta, \qquad B = \nabla\cdot, \qquad B^* = -\nabla, \qquad C = I_{L^2(\Omega)}, \qquad D = \Delta. \tag{2.13}$$

In the sequel, we shall also consider the operators A_{loc} and D_{loc} which are localized versions to mesh cells of the corresponding global differential operators, that is, those operators act locally on each mesh cell as their global counterpart without taking into account possible discontinuities across mesh interfaces.

At the discrete level, we also consider the operators $A_h \in \mathcal{L}(V_{ah}; V_{ah})$ s.t. $(A_h v_h, w_h)_{L_a} = -a(v_h, w_h)$, $B_h \in \mathcal{L}(V_{ah}; V_{dh})$ s.t. $(B_h v_h, q_h)_{L_d} = b(v_h, q_h)$ with adjoint $B_h^* \in \mathcal{L}(V_{dh}; V_{ah})$ s.t. $(B_h^* q_h, v_h)_{L_a} = b(v_h, q_h)$, and $D_h \in \mathcal{L}(V_{dh}; V_{dh})$ s.t. $(D_h q_h, r_h)_{L_d} = -d(q_h, r_h)$. Observe that duality products have been replaced by L_a- and L_d-scalar products. The discrete problem (2.8)–(2.9) can be rewritten in the form

$$-A_h u_h^n - B_h^* p_h^n = f_h^n, \tag{2.14}$$

$$C\delta_t p_h^n + B_h \delta_t u_h^n - D_h p_h^n = g_h^n, \tag{2.15}$$

these equalities holding for all $n \in \{1, \dots, N\}$ in V_{ah} and V_{dh} respectively. For later use, we let $f_h^0 := -A_h u_{0h} - B_h^* p_{0h}$, so that (2.14) also holds for $n = 0$.

2.4. **Elimination of the displacement**

An alternative viewpoint to the PDE system (2.11)–(2.12) consists in eliminating the displacement to infer the following parabolic-like evolution equation for the pressure

$$\partial_t(Cp + Lp) - Dp = \tilde{g}, \tag{2.16}$$

where $L = -BA^{-1}B^* \in \mathcal{L}(L_d; L_d)$ is self-adjoint and monotone ($(Lq, q)_{L_d} \geq 0$ for all $q \in L_d$) and where $\tilde{g} = g + BA^{-1}\partial_t f \in L^2(0, T; L_d)$. In addition, the operator L is coercive provided the operator B is surjective;

this assumption actually holds for the poroelasticity system. The same elimination can be performed at the discrete level yielding

$$\delta_t(Cp_h^n + L_h p_h^n) - D_h p_h^n = \tilde{g}_h^n, \tag{2.17}$$

where $L_h = -B_h A_h^{-1} B_h^*$ is again self-adjoint and monotone and g_h^n is defined accordingly. The monotonicity property of L_h is key to the stability of the discrete system since it ensures that the pressure-displacement coupling only introduces additional dissipation into the pressure evolution equation. Interestingly, the operator L_h needs not be coercive for the discrete problem (2.17) to be stable; this is actually not the case for the poroelasticity problem if equal-order polynomials are used for the displacement and the pressure. In the following section, we will see that the use of different polynomial orders is solely geared to obtain optimal convergence rates for all the terms in the energy norm. The situation is different in the singular limit where $C \to 0$ (that is, when the Biot modulus M goes to infinity) since in this case, the coercivity of the operator L_h is needed to control the L_d-norm of the pressure.

2.5. The steady problem

Both the *a priori* and *a posteriori* error analysis of the approximation of the steady version of (2.1)–(2.2) using the subspaces $\{V_{ah}\}_{h>0}$ and $\{V_{dh}\}_{h>0}$ will play a role in the error analysis for the time-dependent case. The steady version of (2.1)–(2.2) consists in seeking $\overline{u} \in V_a$ and $\overline{p} \in V_d$ such that

$$a(\overline{u}, v) - b(v, \overline{p}) = \langle \overline{f}, v\rangle_a, \qquad \forall v \in V_a, \tag{2.18}$$

$$d(\overline{p}, q) = \langle \overline{g}, q\rangle_d, \qquad \forall q \in V_d, \tag{2.19}$$

with data $\overline{f} \in V_a'$ and $\overline{g} \in V_d'$. It is straightforward to verify that the problem (2.18)–(2.19) is well-posed owing to its upper triangular structure and the coercivity of the bilinear forms a and d.

The discrete problem consists in seeking $\overline{u}_h \in V_{ah}$ and $\overline{p}_h \in V_{dh}$ such that

$$a(\overline{u}_h, v_h) - b(v_h, \overline{p}_h) = \langle \overline{f}, v_h\rangle_a, \qquad \forall v_h \in V_{ah}, \tag{2.20}$$

$$d(\overline{p}_h, q_h) = \langle \overline{g}, q_h\rangle_d, \qquad \forall q_h \in V_{dh}. \tag{2.21}$$

Here, we do not consider an approximation to the data $\overline{f}$ and $\overline{g}$. The discrete problem is conveniently reformulated using a Riesz projection operator $\mathfrak{R}_h : V_a \times V_d \to V_{ah} \times V_{dh}$ such that for all $(v,q) \in V_a \times V_d$, $\mathfrak{R}_h(v,q) := (\mathfrak{R}_{ah}(v,q), \mathfrak{R}_{dh}(q))$ is defined by

$$a(v - \mathfrak{R}_{ah}(v,q), v_h) - b(v_h, q - \mathfrak{R}_{dh}(q)) = 0, \qquad \forall v_h \in V_{ah}, \tag{2.22}$$

$$d(q - \mathfrak{R}_{dh}(q), q_h) = 0, \qquad \forall q_h \in V_{dh}. \tag{2.23}$$

It is clear that $(\overline{u}_h, \overline{p}_h)$ solves (2.20)–(2.21) if and only if $\overline{u}_h = \mathfrak{R}_{ah}(\overline{u}, \overline{p})$ and $\overline{p}_h = \mathfrak{R}_{dh}(\overline{p})$. The approximation properties of the operator $\mathfrak{R}_h$ can be found in [14]. The result is restated here for completeness.

Lemma 2.2. *The following holds for all* $(v,q) \in V_a \times V_d$,

$$\|v - \mathfrak{R}_{ah}(v,q)\|_a \le \inf_{v_h \in V_{ah}} \|v - v_h\|_a + \beta \|q - \mathfrak{R}_{dh}(q)\|_c, \tag{2.24}$$

$$\|q - \mathfrak{R}_{dh}(q)\|_d = \inf_{q_h \in V_{dh}} \|q - q_h\|_d. \tag{2.25}$$

Proof. Property (2.25) is classical. To establish (2.24), consider the operator $\mathfrak{R}_{ah}^*$ defined by (2.10) (the upper index n is dropped since meshes are kept fixed in time). Then, observe that since both $\mathfrak{R}_{ah}^*(v)$ and $\mathfrak{R}_{ah}(v,q)$

are in V_{ah},

$$\begin{aligned}\|\mathfrak{R}_{ah}(v,q)-\mathfrak{R}^*_{ah}(v)\|_a^2 &= a(\mathfrak{R}_{ah}(v,q)-\mathfrak{R}^*_{ah}(v),\mathfrak{R}_{ah}(v,q)-\mathfrak{R}^*_{ah}(v))\\ &= a(\mathfrak{R}_{ah}(v,q)-v,\mathfrak{R}_{ah}(v,q)-\mathfrak{R}^*_{ah}(v))+a(v-\mathfrak{R}^*_{ah}(v),\mathfrak{R}_{ah}(v,q)-\mathfrak{R}^*_{ah}(v))\\ &= -b(\mathfrak{R}_{ah}(v,q)-\mathfrak{R}^*_{ah}(v),q-\mathfrak{R}_{dh}(q))\\ &\le \beta\|\mathfrak{R}_{ah}(v,q)-\mathfrak{R}^*_{ah}(v)\|_a\|q-\mathfrak{R}_{dh}(q)\|_c.\end{aligned}$$

Hence, $\|\mathfrak{R}_{ah}(v,q)-\mathfrak{R}^*_{ah}(v)\|_a \le \beta\|q-\mathfrak{R}_{dh}(q)\|_c$ whence it follows by the triangle inequality that

$$\|v-\mathfrak{R}_{ah}(v,q)\|_a \le \|v-\mathfrak{R}^*_{ah}(v)\|_a+\|\mathfrak{R}_{ah}(v,q)-\mathfrak{R}^*_{ah}(v)\|_a \le \|v-\mathfrak{R}^*_{ah}(v)\|_a+\beta\|q-\mathfrak{R}_{dh}(q)\|_c,$$

readily yielding (2.24). □

To deduce from Lemma 2.2 asymptotic rates of convergence for the approximation error in terms of the mesh size when the exact solution is smooth enough, we introduce the following assumptions.

Hypothesis 2.1. *There exist constants c_1 and c_2, positive real numbers s_a and s_d, and subspaces $W_a \subset V_a$ and $W_d \subset V_d$ respectively equipped with norms $\|\cdot\|_{W_a}$ and $\|\cdot\|_{W_d}$, such that independently of h,*

$$\forall v \in W_a, \qquad \inf_{v_h\in V_{ah}} \|v-v_h\|_a \le c_1 h^{s_a}\|v\|_{W_a}, \tag{2.26}$$

$$\forall q \in W_d, \qquad \inf_{q_h\in V_{dh}} \|q-q_h\|_d \le c_2 h^{s_d}\|q\|_{W_d}. \tag{2.27}$$

Hypothesis 2.2. *There exist a constant c_3 and a positive real number δ such that for all $r\in L_d$, the unique solution $\phi\in V_d$ of the dual problem $d(q,\phi)=c(r,q)$ for all $q\in V_d$, is such that there is $\phi_h\in V_{dh}$ satisfying*

$$\|\phi-\phi_h\|_d \le c_3 h^{\delta}\|r\|_c. \tag{2.28}$$

Hypothesis 2.1 is classical in the context of finite element approximations. It will be used in the *a priori* error analysis. To keep technicalities at a minimum, a version of Hypothesis 2.1 localized to mesh cells is not considered. Hypothesis 2.2 is an elliptic regularity property associated with the bilinear form d on V_d. It is stated here in compact form, the usual statement consisting in assuming that the dual solution ϕ is in a subspace Y_d of V_d where the interpolation property (2.28) holds in the form $\|\phi-\phi_h\|_d \le c_3 h^{\delta}\|\phi\|_{Y_d}$. Hypothesis 2.2 will serve both in the *a priori* and the *a posteriori* error analysis. In the latter case, a sharper statement localized to mesh cells will be introduced in Section 4.2. For the time being, we will only use the following important consequence of Hypothesis 2.2:

$$\|q-\mathfrak{R}_{dh}(q)\|_c \le c_3 h^{\delta}\|q-\mathfrak{R}_{dh}(q)\|_d. \tag{2.29}$$

Indeed, letting ϕ be the dual solution associated with $r := q-\mathfrak{R}_{dh}(q)$ and observing that $d(r,\phi_h)=0$ for $\phi_h\in V_{dh}$ yields

$$\|q-\mathfrak{R}_{dh}(q)\|_c^2 = c(r,r) = d(r,\phi) = d(r,\phi-\phi_h) \le \|r\|_d\|\phi-\phi_h\|_d, \tag{2.30}$$

whence (2.29) readily follows. An important consequence of (2.24), (2.26), and (2.29) is that for all $(v,q)\in W_a\times W_d$,

$$\|v-\mathfrak{R}_{ah}(v,q)\|_a \le c_1 h^{s_a}\|v\|_{W_a}+\beta c_2 c_3 h^{s_d+\delta}\|q\|_{W_d}. \tag{2.31}$$

For the purpose of computational efficiency, it is reasonable to balance both sources of error in $\|v-\mathfrak{R}_{ah}(v,q)\|_a$. This motivates the following hypothesis.

Hypothesis 2.3. $s_a = s_d + \delta =: s$.

In the framework of Hypotheses 2.1–2.3, Lemma 2.2 yields for all $(v, q) \in W_a \times W_d$,

$$\|v - \mathfrak{R}_{ah}(v, q)\|_a \leq h^s(c_1\|v\|_{W_a} + \beta c_2 c_3 \|q\|_{W_d}), \tag{2.32}$$

$$\|q - \mathfrak{R}_{dh}(q)\|_c \leq h^s c_2 c_3 \|q\|_{W_d}, \tag{2.33}$$

$$\|q - \mathfrak{R}_{dh}(q)\|_d \leq h^{s-\delta} c_2 \|q\|_{W_d}. \tag{2.34}$$

As a result, whenever the exact solution of the steady problem (2.18)–(2.19) is smooth enough, namely $(\overline{u}, \overline{p}) \in W_a \times W_d$, the error $\|\overline{p} - \overline{p}_h\|_d$ converges asymptotically as $h^{s-\delta}$ while the error $\|\overline{u} - \overline{u}_h\|_a + \|\overline{p} - \overline{p}_h\|_c$ converges asymptotically as h^s. Since δ is positive, this means that the error $\|\overline{u} - \overline{u}_h\|_a + \|\overline{p} - \overline{p}_h\|_c$ converges at a faster rate than $\|\overline{p} - \overline{p}_h\|_d$. This difference in the convergence rates will be accounted for in the subsequent analysis of the time-dependent problem, the goal being to derive *a priori* and *a posteriori* error bounds that are optimally convergent for $\|p^n - p_h^n\|_d$ on the one hand and for $\|u^n - u_h^n\|_a + \|p^n - p_h^n\|_c$ on the other hand.

Application to poroelasticity. Consider the model problem (1.1)–(1.2) with the displacement (resp., the pressure) approximated in space by continuous Lagrange finite elements of degree $k \geq 1$ (resp., $l \geq 1$). Then, Hypothesis 2.1 holds with $s_a := k$, $W_a := [H_0^1(\Omega) \cap H^{k+1}(\mathcal{T}_h)]^3$, $s_d := l$ and $W_d := H_0^1(\Omega) \cap H^{l+1}(\mathcal{T}_h)$, where for $m \geq 0$, $H^m(\mathcal{T}_h)$ denotes the usual broken Sobolev space of order m. Hypothesis 2.2 means that the steady-state version of the pressure equation yields elliptic regularity, namely for all $r \in L^2(\Omega)$, the unique solution $\phi \in H_0^1(\Omega)$ to the dual problem $\int_\Omega \kappa \nabla\phi \cdot \nabla q = \int_\Omega rq$ for all $q \in H_0^1(\Omega)$ is in $H^2(\Omega)$. Then, (2.28) holds with $\delta := 1$. As a result, Hypothesis 2.3 implies

$$k = l + 1, \tag{2.35}$$

i.e. the polynomial interpolation for the displacement is one degree higher than that for the pressure. The most common choice in practice is $k = 2$ and $l = 1$, *i.e.* continuous piecewise quadratics are used to approximate the displacement and continuous piecewise linears are used to approximate the pressure.

3. *A priori* error analysis

The *a priori* error analysis is performed under the assumption that the exact solution is smooth, namely

$$u \in C_t^1(W_a) \cap C_t^2(V_a), \qquad p \in C_t^1(W_d) \cap C_t^2(L_d). \tag{3.1}$$

For all $n \in \{1, \ldots, N\}$, define

$$C_1^n(u,p) = 2\gamma^2 c_2^2 c_3^2 \|\partial_t p(s)\|^2_{L^\infty(I_n;W_d)} + 2\beta^2\gamma^2(c_1\|\partial_t u(s)\|_{L^\infty(I_n;W_a)} + \beta c_2 c_3 \|\partial_t p(s)\|_{L^\infty(I_n;W_d)})^2, \tag{3.2}$$

$$C_2^n(u,p) = \tfrac{1}{2}\gamma^2\hat{\gamma}^2 \|\partial_{tt}^2 p(s)\|^2_{L^\infty(I_n;L_d)} + \tfrac{1}{2}\beta^2\gamma^2 \|\partial_{tt}^2 u(s)\|^2_{L^\infty(I_n;V_a)}, \tag{3.3}$$

$$C^n(f,g) = \tfrac{1}{2}\|f^n - f_h^n\|^2_{\prime_a} + \tau_n \|g^n - g_h^n\|^2_{\prime_d}. \tag{3.4}$$

Moreover, it is assumed that the initial data u_{0h} and p_{0h} are chosen such that

$$\|u_0 - u_{0h}\|_a \leq c_4 h^s \|u_0\|_{W_a} \qquad \text{and} \qquad \|p_0 - p_{0h}\|_c \leq c_5 h^s \|p_0\|_{W_d}, \tag{3.5}$$

and we define

$$C(u_0,p_0) = (c_1 + c_4)^2 \|u_0\|^2_{W_a} + (\beta^2 c_2^2 c_3^2 + \tfrac{1}{2}(c_2 c_3 + c_5)^2)\|p_0\|^2_{W_d}. \tag{3.6}$$

One possible choice is $u_{0h} = \mathfrak{R}_{ah}(u_0, p_0)$ and $p_{0h} = \mathfrak{R}_{dh}(p_0)$, in which case we can take $C(u_0, p_0) = 0$.

Theorem 3.1. *In the above framework, the following holds for all* $n \in \{1, \dots, N\}$,

$$\begin{aligned}\tfrac{1}{4}\|u^n - u_h^n\|_a^2 + \tfrac{1}{4}\|p^n - p_h^n\|_c^2 \leq h^{2s}C(u_0, p_0) + \sum_{m=1}^{n} C^m(f, g) + \sum_{m=1}^{n} [\tau_m h^{2s} C_1^m(u, p) + \tau_m^3 C_2^m(u, p)] \\ + h^{2s}(c_1^2\|u^n\|_{W_a}^2 + (\beta^2 + \tfrac{1}{2})c_2^2 c_3^2 \|p^n\|_{W_d}^2),\end{aligned} \tag{3.7}$$

and

$$\begin{aligned}\sum_{m=1}^{n} \tfrac{1}{8}\tau_m \|p^m - p_h^m\|_d^2 \leq h^{2s}C(u_0, p_0) + \sum_{m=1}^{n} C^m(f, g) + \sum_{m=1}^{n} [\tau_m h^{2s} C_1^m(u, p) + \tau_m^3 C_2^m(u, p)] \\ + \sum_{m=1}^{n} \tfrac{1}{4}\tau_m h^{2(s-\delta)} c_2^2 \|p^m\|_{W_d}^2.\end{aligned} \tag{3.8}$$

Proof. (i) For all $n \in \{1, \dots, N\}$, let us first bound the quantities

$$\eta_{ah}^n = \mathfrak{R}_{ah}(u^n, p^n) - u_h^n \qquad \text{and} \qquad \eta_{dh}^n = \mathfrak{R}_{dh}(p^n) - p_h^n.$$

Observe that

$$\begin{aligned} a(\eta_{ah}^n, v_h) - b(v_h, \eta_{dh}^n) &= \langle f^n - f_h^n, v_h \rangle_a, && \forall v_h \in V_{ah}, \\ c(\eta_{dh}^n - \eta_{dh}^{n-1}, q_h) + b(\eta_{ah}^n - \eta_{ah}^{n-1}, q_h) + \tau_n d(\eta_{dh}^n, q_h) &= \tau_n \langle g^n - g_h^n, q_h \rangle_d + c(\theta_{dh}^n, q_h) + b(\theta_{ah}^n, q_h), && \forall q_h \in V_{dh}, \end{aligned}$$

where $\theta_{dh}^n = \mathfrak{R}_{dh}(p^n) - \mathfrak{R}_{dh}(p^{n-1}) - \tau_n \partial_t p^n$ and $\theta_{ah}^n = \mathfrak{R}_{ah}(u^n, p^n) - \mathfrak{R}_{ah}(u^{n-1}, p^{n-1}) - \tau_n \partial_t u^n$. Testing with $v_h := \eta_{ah}^n - \eta_{ah}^{n-1} \in V_{ah}$ and $q_h := \eta_{dh}^n \in V_{dh}$ yields after some straightforward algebra

$$\begin{aligned}\tfrac{1}{2}\|\eta_{ah}^n\|_a^2 + \tfrac{1}{2}\|\eta_{ah}^n - \eta_{ah}^{n-1}\|_a^2 + \tfrac{1}{2}\|\eta_{dh}^n\|_c^2 + \tfrac{1}{2}\|\eta_{dh}^n - \eta_{dh}^{n-1}\|_c^2 + \tau_n\|\eta_{dh}^n\|_d^2 = \tfrac{1}{2}\|\eta_{ah}^{n-1}\|_a^2 + \tfrac{1}{2}\|\eta_{dh}^{n-1}\|_c^2 \\ + \langle f^n - f_h^n, \eta_{ah}^n - \eta_{ah}^{n-1}\rangle_a + \tau_n\langle g^n - g_h^n, \eta_{dh}^n\rangle_d + c(\theta_{dh}^n, \eta_{dh}^n) + b(\theta_{ah}^n, \eta_{dh}^n).\end{aligned}$$

Hence,

$$\tfrac{1}{2}\|\eta_{ah}^n\|_a^2 + \tfrac{1}{2}\|\eta_{dh}^n\|_c^2 + \tfrac{1}{4}\tau_n\|\eta_{dh}^n\|_d^2 \leq \tfrac{1}{2}\|\eta_{ah}^{n-1}\|_a^2 + \tfrac{1}{2}\|\eta_{dh}^{n-1}\|_c^2 + C^n(f, g) + \tau_n^{-1}\gamma^2\|\theta_{dh}^n\|_c^2 + \tau_n^{-1}\beta^2\gamma^2\|\theta_{ah}^n\|_a^2.$$

(ii) Let us now bound the quantities θ_{dh}^n and θ_{ah}^n. Observe that

$$\theta_{dh}^n = -\int_{I_n} [\partial_t p(s) - \mathfrak{R}_{dh}(\partial_t p(s))]\mathrm{d}s - \int_{I_n} (s - t_{n-1})\partial_{tt}^2 p(s)\mathrm{d}s.$$

Hence, owing to the regularity assumptions on the exact solution and equation (2.33),

$$\|\theta_{dh}^n\|_c \leq \tau_n h^s c_2 c_3 \|\partial_t p(s)\|_{L^\infty(I_n;W_d)} + \tfrac{1}{2}\tau_n^2 \hat{\gamma}\|\partial_{tt}^2 p(s)\|_{L^\infty(I_n;L_d)}.$$

Similarly, using (2.32),

$$\|\theta_{ah}^n\|_a \leq \tau_n h^s (c_1\|\partial_t u(s)\|_{L^\infty(I_n;W_a)} + \beta c_2 c_3 \|\partial_t p(s)\|_{L^\infty(I_n;W_d)}) + \tfrac{1}{2}\tau_n^2\|\partial_{tt}^2 u(s)\|_{L^\infty(I_n;V_a)}.$$

Therefore,

$$\tau_n^{-1}\gamma^2\|\theta_{dh}^n\|_c^2 + \tau_n^{-1}\beta^2\gamma^2\|\theta_{ah}^n\|_a^2 \leq \tau_n h^{2s} C_1^n(u, p) + \tau_n^3 C_2^n(u, p).$$

Summing up the above inequalities leads to

$$\tfrac{1}{2}\|\eta_{ah}^n\|_a^2+\tfrac{1}{2}\|\eta_{dh}^n\|_c^2+\sum_{m=1}^n \tfrac{1}{4}\tau_m\|\eta_{dh}^m\|_d^2 \le \tfrac{1}{2}\|\eta_{ah}^0\|_a^2+\tfrac{1}{2}\|\eta_{dh}^0\|_c^2+\sum_{m=1}^n [C^m(f,g)+\tau_m h^{2s}C_1^m(u,p)+\tau_m^3 C_2^m(u,p)].$$

(iii) We now bound the initial errors η_{ah}^0 and η_{dh}^0. In the case where $u_{0h}=\mathfrak{R}_{ah}(u_0,p_0)$ and $p_{0h}=\mathfrak{R}_{dh}(p_0)$, it is clear that $\eta_{ah}^0=0$ and $\eta_{dh}^0=0$. In the general case, use the triangle inequality to infer

$$\|\eta_{dh}^0\|_c=\|\mathfrak{R}_{dh}(p_0)-p_{0h}\|_c\le\|\mathfrak{R}_{dh}(p_0)-p_0\|_c+\|p_0-p_{0h}\|_c\le h^s(c_2c_3+c_5)\|p_0\|_{W_d}.$$

Similarly,

$$\|\eta_{ah}^0\|_a\le\|\mathfrak{R}_{ah}(u_0,p_0)-u_0\|_a+\|u_0-u_{0h}\|_a\le h^s((c_1+c_4)\|u_0\|_{W_a}+\beta c_2c_3\|p_0\|_{W_d}).$$

Hence,

$$\tfrac{1}{2}\|\eta_{ah}^n\|_a^2+\tfrac{1}{2}\|\eta_{dh}^n\|_c^2+\sum_{m=1}^n \tfrac{1}{4}\tau_m\|\eta_{dh}^m\|_d^2\le h^{2s}C(u_0,p_0)+\sum_{m=1}^n[C^m(f,g)+\tau_m h^{2s}C_1^m(u,p)+\tau_m^3C_2^m(u,p)].$$

(iv) The conclusion readily results from the triangle inequality and equations (2.32)–(2.33)–(2.34). □

Theorem 3.1 shows that whenever the exact solution is smooth enough and up to data approximation errors that can be made small enough, $\|u^n-u_h^n\|_a+\|p^n-p_h^n\|_c$ converges to order s in space and first-order in time, while $(\sum_{m=1}^n\tau_m\|p^m-p_h^m\|_d^2)^{1/2}$ converges to order $(s-\delta)$ in space and first-order in time.

Application to poroelasticity. When continuous piecewise quadratics (resp., linears) are used to approximate the displacement (resp., the pressure), $\|u^n-u_h^n\|_{H^1}+\|p^n-p_h^n\|_{L^2}$ converges to second-order in space and first-order in time, while $(\sum_{m=1}^n\tau_m\|p^m-p_h^m\|_{H^1}^2)^{1/2}$ converges to first-order in space and in time.

4. *A POSTERIORI* ERROR ANALYSIS

This section is devoted to the *a posteriori* error analysis for the discrete scheme (2.8)–(2.9). The main results are Theorems 4.1 and 4.2 for the direct approach and Theorem 4.3 for the approach using elliptic reconstruction.

4.1. **The direct approach**

The *a posteriori* error analysis relies on the stability of the continuous problem. Therefore, it is convenient to rewrite the discrete scheme as equations holding a.e. in $(0,T)$ rather than at the discrete times $\{t_n\}_{n=1}^N$. To this purpose, let $u_{h\tau}$ (resp., $p_{h\tau}$) be the continuous and piecewise affine function in time such that for all $n\in\{0,\dots,N\}$, $u_{h\tau}(t_n)=u_h^n$ (resp., $p_{h\tau}(t_n)=p_h^n$). Observe that $\partial_t u_{h\tau}$ and $\partial_t p_{h\tau}$ are defined a.e. in $(0,T)$. Similarly, let $f_{h\tau}$ be the continuous and piecewise affine function in time such that for all $n\in\{0,\dots,N\}$, $f_{h\tau}(t_n)=f_h^n$. We will also need to consider piecewise constant functions in time, namely $\pi^0 p_{h\tau}$ (resp., $\pi^0 g_{h\tau}$) equal to p_h^n (resp., g_h^n) on I_n for all $n\in\{1,\dots,N\}$. With the above notation, the discrete scheme (2.8)–(2.9) yields a.e. in $(0,T)$,

$$a(u_{h\tau},v_h)-b(v_h,p_{h\tau})=(f_{h\tau},v_h)_{L_a},\qquad\forall v_h\in V_{ah},\tag{4.1}$$
$$c(\partial_t p_{h\tau},q_h)+b(\partial_t u_{h\tau},q_h)+d(\pi^0p_{h\tau},q_h)=(\pi^0g_{h\tau},q_h)_{L_d},\qquad\forall q_h\in V_{dh}.\tag{4.2}$$

For the *a posteriori* error analysis, it is convenient to introduce the Galerkin residual $\mathfrak{G}_a$ (resp., $\mathfrak{G}_d$) which is a continuous and piecewise affine function in time with values in V_a' (resp., piecewise constant function in time with values in V_d') such that a.e. in $(0,T)$,

$$\langle \mathfrak{G}_a, v\rangle_a = (f_{h\tau}, v)_{L_a} - a(u_{h\tau}, v) + b(v, p_{h\tau}), \qquad \forall v \in V_a, \tag{4.3}$$

$$\langle \mathfrak{G}_d, q\rangle_d = (\pi^0 g_{h\tau}, q)_{L_d} - c(\partial_t p_{h\tau}, q) - b(\partial_t u_{h\tau}, q) - d(\pi^0 p_{h\tau}, q), \qquad \forall q \in V_d. \tag{4.4}$$

Note that $\mathfrak{G}_a \in H^1(0,T;V_a') \cap L^\infty(0,T;V_a')$ and $\mathfrak{G}_d \in L^2(0,T;V_d')$. Let

$$\mathcal{E}(f,g) = \|g - \pi^0 g_{h\tau}\|^2_{L^2(0,T;V_d')} + \left(2\|f - f_{h\tau}\|_{L^\infty(0,T;V_a')} + \|\partial_t(f - f_{h\tau})\|_{L^1(0,T;V_a')}\right)^2, \tag{4.5}$$

$$\mathcal{E}_{\rm dat} = \|u_0 - u_{0h}\|_a^2 + \|p_0 - p_{0h}\|_c^2 + 4\mathcal{E}(f,g), \tag{4.6}$$

$$\mathcal{E}_{\rm spc} = 4\|\mathfrak{G}_d\|^2_{L^2(0,T;V_d')} + 4\left(2\|\mathfrak{G}_a\|_{L^\infty(0,T;V_a')} + \|\partial_t \mathfrak{G}_a\|_{L^1(0,T;V_a')}\right)^2, \tag{4.7}$$

$$\mathcal{E}_{\rm tim} = \|p_{h\tau} - \pi^0 p_{h\tau}\|^2_{L^2(0,T;V_d)}. \tag{4.8}$$

$\mathcal{E}_{\rm dat}$ is referred to as a data oscillation term.

Theorem 4.1. *For all* $n \in \{1,\dots,N\}$,

$$\tfrac14\|(u - u_{h\tau}, p - p_{h\tau})\|^2_{X(0,t_n)} + \tfrac12\|p - \pi^0 p_{h\tau}\|^2_{L^2(0,t_n;V_d)} \le \mathcal{E}_{\rm dat} + \mathcal{E}_{\rm spc} + \mathcal{E}_{\rm tim}. \tag{4.9}$$

Proof. Let $\xi = u - u_{h\tau}$, $\zeta = p - p_{h\tau}$, and $\zeta^* = p - \pi^0 p_{h\tau}$. Observe that a.e. in $(0,T)$,

$$a(\xi, v) - b(v, \zeta) = \langle f - f_{h\tau} + \mathfrak{G}_a, v\rangle_a, \qquad \forall v \in V_a,$$

$$c(\partial_t\zeta, q) + b(\partial_t \xi, q) + d(\zeta^*, q) = \langle g - \pi^0 g_{h\tau} + \mathfrak{G}_d, q\rangle_d, \qquad \forall q \in V_d.$$

Testing for a.e. $t \in (0,T)$ with $v := \partial_t \xi$ and $q := \zeta$ yields

$$\tfrac12 d_t\|\xi\|_a^2 + \tfrac12 d_t\|\zeta\|_c^2 + \tfrac12\|\zeta\|_d^2 + \tfrac12\|\zeta^*\|_d^2 = \langle f - f_{h\tau} + \mathfrak{G}_a, \partial_t\xi\rangle_a + \langle g - \pi^0 g_{h\tau} + \mathfrak{G}_d, \zeta\rangle_d + \tfrac12\|p_{h\tau} - \pi^0 p_{h\tau}\|_d^2,$$

where we have used the fact that owing to the symmetry of d,

$$d(\zeta, \zeta^*) = \tfrac12 d(\zeta,\zeta) + \tfrac12 d(\zeta^*,\zeta^*) - \tfrac12 d(\zeta - \zeta^*, \zeta - \zeta^*).$$

Since $f - f_{h\tau} + \mathfrak{G}_a \in H^1(0,T;V_a')$ and $g - \pi^0 g_{h\tau} + \mathfrak{G}_d \in L^2(0,T;V_d')$, we can conclude by proceeding as in the proof of Proposition 2.1. □

Since for all $m \in \{1,\dots,N\}$ and for all $s \in I_m$, $(p_{h\tau} - \pi^0 p_{h\tau})(s) = \tau_m^{-1}(s - t_m)(p_h^m - p_h^{m-1})$, it is straightforward to formulate $\mathcal{E}_{\rm tim}$ in terms of so-called time error indicators as follows:

$$\mathcal{E}_{\rm tim} = \sum_{m=1}^N \mathcal{E}_{\rm tim}^m \qquad \text{with} \qquad \mathcal{E}_{\rm tim}^m = \tfrac13 \tau_m \|p_h^m - p_h^{m-1}\|_d^2. \tag{4.10}$$

We now proceed to bound more explicitly $\mathcal{E}_{\rm spc}$. We assume that the various bilinear forms in the model problem (2.1)–(2.2) can be localized as follows: for all $(v,q) \in V_a \times V_d$ and for all $(\xi,\zeta) \in V_a \times V_d$ such that $A_{\rm loc}\xi \in L_a$ and $D_{\rm loc}\zeta \in L_d$,

$$a(\xi, v) = \sum_{T \in \mathcal{T}_h} [-(A_{\rm loc}\xi, v)_{L_a(T)} + (J_a\xi, v)_{L_a(\partial T)}], \tag{4.11}$$

$$d(\zeta, q) = \sum_{T \in \mathcal{T}_h} [-(D_{\rm loc}\zeta, q)_{L_d(T)} + (J_d\zeta, q)_{L_d(\partial T)}]. \tag{4.12}$$

Here, for a mesh cell $T \in \mathcal{T}_h$, $L_a(T)$ and $L_d(T)$ are local versions of L_a and L_d respectively, $(\cdot,\cdot)_{L_a(\partial T)}$ and $(\cdot,\cdot)_{L_d(\partial T)}$ are scalar products for functions defined on the boundary ∂T of T and J_a and J_d are suitable (jump) operators such that $J_a\xi = 0$ if $A\xi \in L_a$ and $J_d\zeta = 0$ if $D\zeta \in L_d$. In addition, for all $(v,\zeta) \in V_a \times V_d$ and for all $(q,r) \in L_d \times L_d$,

$$b(v,\zeta) = \sum_{T\in\mathcal{T}_h} (v, B^*\zeta)_{L_a(T)} = \sum_{T\in\mathcal{T}_h} (Bv,\zeta)_{L_d(T)}, \qquad \text{and} \qquad c(q,r) = \sum_{T\in\mathcal{T}_h} (Cq,r)_{L_d(T)}. \tag{4.13}$$

Finally, following [10,23], we assume that there exist two (Clément-type) interpolation operators $i_{ah}: V_a \to V_{ah}$ and $i_{dh}: V_d \to V_{dh}$ such that for all $(v,q) \in V_a \times V_d$,

$$\sum_{T\in\mathcal{T}_h} [h_T^{-2}\|v - i_{ah}(v)\|^2_{L_a(T)} + h_T^{-1}\|v - i_{ah}(v)\|^2_{L_a(\partial T)}] \le c_6\|v\|_a^2, \tag{4.14}$$

$$\sum_{T\in\mathcal{T}_h} [h_T^{-2}\|q - i_{dh}(q)\|^2_{L_d(T)} + h_T^{-1}\|q - i_{dh}(q)\|^2_{L_d(\partial T)}] \le c_7\|q\|_d^2, \tag{4.15}$$

where for all $T \in \mathcal{T}_h$, h_T denotes the diameter of T. In the context of poroelasticity where $V_a = [H_0^1(\Omega)]^3$ and $V_d = H_0^1(\Omega)$, the usual Clément [7] interpolation operator (modified to account for homogeneous Dirichlet boundary conditions) or the Scott-Zhang [18] interpolation operator can be used.

We define the following elementwise and jump residuals for all $m \in \{1,\dots,N\}$,

$$R_{uh}^m = f_h^m + A_{\mathrm{loc}}u_h^m + B^*p_h^m, \qquad J_{uh}^m = J_a u_h^m, \tag{4.16}$$

$$R_{ph}^m = g_h^m - C\delta_t p_h^m - B\delta_t u_h^m + D_{\mathrm{loc}}p_h^m, \qquad J_{ph}^m = J_d p_h^m. \tag{4.17}$$

For $m = 0$, R_{uh}^0, J_{uh}^0, and J_{ph}^0 are defined similarly, while we set $R_{ph}^0 = (D_{\mathrm{loc}} - D_h)p_{0h}$ for later use in Section 4.2. We also define for all $m \in \{0,\dots,N\}$, the so-called space error indicators

$$\widehat{\mathcal{E}}_u^m = \sum_{T\in\mathcal{T}_h} \widehat{\mathcal{E}}_{u,T}^m, \qquad \widehat{\mathcal{E}}_{u,T}^m = h_T^2\|R_{uh}^m\|^2_{L_a(T)} + h_T\|J_{uh}^m\|^2_{L_a(\partial T)}, \tag{4.18}$$

$$\widehat{\mathcal{E}}_{p,r}^m = \sum_{T\in\mathcal{T}_h} \widehat{\mathcal{E}}_{p,r,T}^m, \qquad \widehat{\mathcal{E}}_{p,r,T}^m = h_T^{2r}[h_T^2\|R_{ph}^m\|^2_{L_d(T)} + h_T\|J_{ph}^m\|^2_{L_d(\partial T)}], \tag{4.19}$$

for a real parameter $r \ge 0$. Here and in the sequel, integrals over element boundaries are restricted to those faces of the element that do not lie on the boundary of Ω where homogeneous Dirichlet boundary conditions are enforced. We will also use the transient version of the above quantities, namely for $m \in \{1,\dots,N\}$,

$$\widehat{\mathcal{E}}_u^m(\delta_t) = \sum_{T\in\mathcal{T}_h} \widehat{\mathcal{E}}_{u,T}^m(\delta_t), \qquad \widehat{\mathcal{E}}_{u,T}^m(\delta_t) = h_T^2\|\delta_t R_{uh}^m\|^2_{L_a(T)} + h_T\|\delta_t J_{uh}^m\|^2_{L_a(\partial T)}, \tag{4.20}$$

$$\widehat{\mathcal{E}}_{p,r}^m(\delta_t) = \sum_{T\in\mathcal{T}_h} \widehat{\mathcal{E}}_{p,r,T}^m(\delta_t), \qquad \widehat{\mathcal{E}}_{p,r,T}^m(\delta_t) = h_T^{2r}[h_T^2\|\delta_t R_{ph}^m\|^2_{L_d(T)} + h_T\|\delta_t J_{ph}^m\|^2_{L_d(\partial T)}], \tag{4.21}$$

where $\delta_t R_{uh}^m = \tau_m^{-1}(R_{uh}^m - R_{uh}^{m-1})$, $\delta_t J_{uh}^m = \tau_m^{-1}(J_{uh}^m - J_{uh}^{m-1})$, and so on.

Proposition 4.1. *In the above framework,*

$$\|\mathfrak{G}_a\|^2_{L^\infty(0,T;V_a')} \le c_6 \sup_{0\le m\le N} \widehat{\mathcal{E}}^m_u, \tag{4.22}$$

$$\|\partial_t\mathfrak{G}_a\|^2_{L^1(0,T;V_a')} \le c_6 \left(\sum_{m=1}^N \tau_m(\widehat{\mathcal{E}}^m_u(\delta_t))^{\frac{1}{2}}\right)^2, \tag{4.23}$$

$$\|\mathfrak{G}_d\|^2_{L^2(0,T;V_d')} \le c_7 \sum_{m=1}^N \tau_m\widehat{\mathcal{E}}^m_{p,0}. \tag{4.24}$$

Hence,

$$\mathcal{E}_{\rm spc} \le \mathcal{E}^\dagger_{\rm spc} := 4c_7 \sum_{m=1}^N \tau_m\widehat{\mathcal{E}}^m_{p,0} + 32c_6 \sup_{0\le m\le N} \widehat{\mathcal{E}}^m_u + 8c_6\left(\sum_{m=1}^N \tau_m(\widehat{\mathcal{E}}^m_u(\delta_t))^{\frac{1}{2}}\right)^2. \tag{4.25}$$

Proof. The proof is only sketched since it uses classical techniques of *a posteriori* error analysis. Observe that

$$\|\mathfrak{G}^m_a\|_{\prime a} = \sup_{0\ne v\in V_a} \frac{\langle \mathfrak{G}^m_a, v\rangle_a}{\|v\|_a} = \sup_{0\ne v\in V_a} \frac{\langle \mathfrak{G}^m_a, v - i_{ah}(v)\rangle_a}{\|v\|_a},$$

since (4.1) and (4.3) imply that $\langle \mathfrak{G}^m_a, v_h\rangle_a = 0$ for all $v_h \in V_{ah}$. Then, use definitions (4.11) and (4.13), assumption (4.14) and a Cauchy-Schwarz inequality to infer (4.22). The proof of (4.23) and (4.24) is similar. ☐

To establish an equivalence result between the residuals and the error, we extend the approach derived by Verfürth [24] for the heat equation. For all $t \in [0,T]$, we introduce the norm

$$\|(v,q)\|^2_{Y(0,t)} = \|(v,q)\|^2_{X(0,t)} + \|A\partial_t v + B^*\partial_t q\|^2_{L^1(0,t;V_a')} + \|C\partial_t q + B\partial_t v\|^2_{L^2(0,t;V_d')}, \tag{4.26}$$

which corresponds to the energy norm $\|\cdot\|_{X(0,t)}$ supplemented with some time-derivatives measured in weaker norms. We observe that a nontrivial novelty with respect to the heat equation is the use of a $L^1(0,t;V_a')$-norm for the time-derivative of the displacement equation.

Theorem 4.2. *For all $n \in \{1,\dots,N\}$,*

$$\tfrac{1}{4}\|(u-u_{h\tau}, p-p_{h\tau})\|^2_{Y(0,t_n)} + \tfrac{1}{2}\|p-\pi^0 p_{h\tau}\|^2_{L^2(0,t_n;V_d)} \le 3\mathcal{E}_{\rm dat} + 3\mathcal{E}^\dagger_{\rm spc} + \tfrac{5}{2}\mathcal{E}_{\rm tim}. \tag{4.27}$$

Furthermore,

$$\mathcal{E}_{\rm tim} \le 2\|(u-u_{h\tau}, p-p_{h\tau})\|^2_{X(0,t_n)} + 2\|p-\pi^0 p_{h\tau}\|^2_{L^2(0,t_n;V_d)}, \tag{4.28}$$

and

$$\mathcal{E}^\dagger_{\rm spc} \le c_{\dagger\dagger}(\|(u-u_{h\tau}, p-p_{h\tau})\|^2_{Y(0,T)} + \|p-\pi^0 p_{h\tau}\|^2_{L^2(0,T;V_d)}) + \mathcal{E}'_{\rm dat}, \tag{4.29}$$

where $c_{\dagger\dagger} = 12\max(8\tilde{\beta}c_6, c_7)c_\dagger^2$, $\tilde{\beta} = \max(1,\beta^2)$, the constant $c_\dagger$ only depends on the shape-regularity of the mesh family and on the maximum polynomial degrees used in the finite element spaces V_{ah} and V_{dh}, and where

$$\mathcal{E}'_{\rm dat} = 16c_6c_\dagger^2(6\|f-f_{h\tau}\|^2_{L^\infty(0,T;V_a')} + \|\partial_t(f-f_{h\tau})\|^2_{L^1(0,T;V_a')}) + 12c_7c_\dagger^2\|g-\pi^0 g_{h\tau}\|^2_{L^2(0,T;V_d')}. \tag{4.30}$$

Proof. (i) Proof of (4.27). Let $\xi = u - u_{h\tau}$, $\zeta = p - p_{h\tau}$, and $\zeta^* = p - \pi^0 p_{h\tau}$. Observe that a.e. in $(0,T)$,

$$\begin{aligned} -A\xi - B^*\zeta &= f - f_{h\tau} + \mathfrak{G}_a, && \text{in } V_a', \\ C\partial_t\zeta + B\partial_t\xi - D\zeta^* &= g - \pi^0 g_{h\tau} + \mathfrak{G}_d, && \text{in } V_d'. \end{aligned}$$

Furthermore, taking the time-derivative of the first equation yields a.e. in $(0,T)$,

$$-A\partial_t\xi - B^*\partial_t\zeta = \partial_t(f - f_{h\tau}) + \partial_t\mathfrak{G}_a, \qquad \text{in } V_a'.$$

Hence, integrating over $(0,t_n)$, using the triangle inequality and taking the square,

$$\begin{aligned}\|A\partial_t\xi + B^*\partial_t\zeta\|^2_{L^1(0,t_n;V_a')} &\le 2\|\partial_t(f-f_{h\tau})\|^2_{L^1(0,t_n;V_a')} + 2\|\partial_t\mathfrak{G}_a\|^2_{L^1(0,t_n;V_a')}\\ &\le \tfrac{1}{2}\mathcal{E}_{\text{dat}} + \tfrac{1}{4}\mathcal{E}^{\dagger}_{\text{spc}}.\end{aligned}$$

Similarly,

$$\begin{aligned}\|C\partial_t\zeta + B\partial_t\xi\|^2_{L^2(0,t_n;V_d')} &\le 3\|g - \pi^0 g_{h\tau}\|^2_{L^2(0,t_n;V_d')} + 3\|\mathfrak{G}_d\|^2_{L^2(0,t_n;V_d')} + 3\|D\zeta^*\|^2_{L^2(0,t_n;V_d')}\\ &\le 7\mathcal{E}_{\text{dat}} + 7\mathcal{E}^{\dagger}_{\text{spc}} + 6\mathcal{E}_{\text{tim}},\end{aligned}$$

where we have used Theorem 4.1 and Proposition 4.1 to infer $\|D\zeta^*\|^2_{L^2(0,t_n;V_d')} = \|\zeta^*\|^2_{L^2(0,t_n;V_d)} \le 2(\mathcal{E}_{\text{dat}} + \mathcal{E}^{\dagger}_{\text{spc}} + \mathcal{E}_{\text{tim}})$ and the fact that $\frac{3}{4} \le 1$ to simplify the expression. Finally, still owing to Theorem 4.1 and Proposition 4.1,

$$\tfrac{1}{4}\|(\xi,\zeta)\|^2_{X(0,t_n)} + \tfrac{1}{2}\|p - \pi^0 p_{h\tau}\|^2_{L^2(0,t_n;V_d)} \le \mathcal{E}_{\text{dat}} + \mathcal{E}^{\dagger}_{\text{spc}} + \mathcal{E}_{\text{tim}}.$$

Summing up all the above contributions yields (4.27) since $\frac{23}{8} \le 3$.
(ii) Proof of (4.28). Use a triangle inequality.
(iii) Proof of (4.29). We bound the three terms in (4.25) starting from the second one. Using the technique of bubble functions introduced by Verfürth [22–24], for all $m \in \{0,\dots,N\}$, there is $\nu_a^m \in V_a$ such that

$$\widehat{\mathcal{E}}_u^m \le \langle \mathfrak{G}_a^m, \nu_a^m\rangle_a, \qquad \text{with} \qquad \|\nu_a^m\|_a \le c_{\dagger}(\widehat{\mathcal{E}}_u^m)^{1/2},$$

where the constant $c_{\dagger}$ depends on the shape-regularity of the mesh family and on the maximum polynomial degrees used in the finite element spaces V_{ah} and V_{dh}. Hence,

$$\widehat{\mathcal{E}}_u^m \le -\langle f^m - f_h^m + A\xi^m + B^*\zeta^m, \nu_a^m\rangle_a \le c_{\dagger}(\|f^m - f_h^m\|_{a'} + \|\xi^m\|_a + \beta\|\zeta^m\|_c)(\widehat{\mathcal{E}}_u^m)^{1/2}.$$

Hence, letting $\tilde\beta = \max(1,\beta^2)$,

$$\sup_{0\le m\le N} \widehat{\mathcal{E}}_u^m \le 3c_{\dagger}^2(\|f - f_{h\tau}\|^2_{L^\infty(0,T;V_a')} + \tilde\beta\|(\xi,\zeta)\|^2_{X(0,T)}).$$

Consider now the third term in (4.25). Using again bubble functions, for all $m \in \{1,\dots,N\}$, there is $\mu_a^m \in V_a$ such that

$$\tau_m^2\widehat{\mathcal{E}}_u^m(\delta_t) \le \langle \mathfrak{G}_a^m - \mathfrak{G}_a^{m-1}, \mu_a^m\rangle_a, \qquad \text{with} \qquad \|\mu_a^m\|_a \le c_{\dagger}\tau_m(\widehat{\mathcal{E}}_u^m(\delta_t))^{1/2}.$$

Therefore,

$$\begin{aligned}\tau_m^2\widehat{\mathcal{E}}_u^m(\delta_t) &\le \int_{I_m}\langle\partial_t\mathfrak{G}_a, \mu_a^m\rangle_a\mathrm{d}s\\ &= -\int_{I_m}\langle\partial_t(f - f_{h\tau})(s), \mu_a^m\rangle_a\mathrm{d}s - \int_{I_m}\langle A\partial_t\xi(s) + B^*\partial_t\zeta(s), \mu_a^m\rangle_a\mathrm{d}s\\ &\le c_{\dagger}\left(\|\partial_t(f - f_{h\tau})\|_{L^1(I_m;V_a')} + \|A\partial_t\xi + B^*\partial_t\zeta\|_{L^1(I_m;V_a')}\right)\tau_m(\widehat{\mathcal{E}}_u^m(\delta_t))^{1/2},\end{aligned}$$

yielding

$$\left(\sum_{m=1}^N \tau_m(\widehat{\mathcal{E}}_u^m(\delta_t))^{\frac{1}{2}}\right)^2 \le 2c_{\dagger}^2\left(\|\partial_t(f - f_{h\tau})\|^2_{L^1(0,T;V_a')} + \|A\partial_t\xi + B^*\partial_t\zeta\|^2_{L^1(0,T;V_a')}\right).$$

Finally, consider the first term in (4.25). Still using bubble functions, for all $m \in \{1, \dots, N\}$, there is $\nu_d^m \in V_d$ such that

$$\widehat{\mathcal{E}}_{p,0}^m \leq \langle \mathfrak{G}_d^m, \nu_d^m \rangle_d \qquad \text{with} \qquad \|\nu_d^m\|_d \leq c_\dagger (\widehat{\mathcal{E}}_{p,0}^m)^{1/2}.$$

Thus,

$$\begin{aligned}\tau_m \widehat{\mathcal{E}}_{p,0}^m &\leq -\int_{I_m} \langle (g - \pi^0 g_{h\tau})(s), \nu_d^m \rangle_d \,\mathrm{d}s + \int_{I_m} \langle C\partial_t \zeta(s) + B\partial_t \xi(s), \nu_d^m \rangle_d \,\mathrm{d}s - \int_{I_m} \langle D\zeta^*, \nu_d^m \rangle_d \,\mathrm{d}s \\ &\leq c_\dagger \left(3\|g - \pi^0 g_{h\tau}\|_{L^2(I_m;V_d')}^2 + 3\|C\partial_t\zeta + B\partial_t\xi\|_{L^2(I_m;V_d')}^2 + 3\|\zeta^*\|_{L^2(I_m;V_d)}^2 \right)^{1/2} \tau_m^{1/2} (\widehat{\mathcal{E}}_{p,0}^m)^{1/2},\end{aligned}$$

using this time a Cauchy-Schwarz inequality. Hence,

$$\sum_{m=1}^N \tau_m \widehat{\mathcal{E}}_{p,0}^m \leq 3c_\dagger^2 \left(\|g - \pi^0 g_{h\tau}\|_{L^2(0,T;V_d')}^2 + \|C\partial_t\zeta + B\partial_t\xi\|_{L^2(0,T;V_d')}^2 + \|\zeta^*\|_{L^2(0,T;V_d)}^2 \right).$$

Collecting the above inequalities yields the desired result. □

4.2. **Elliptic reconstruction**

Because of the use of different polynomial degrees for the displacement and for the pressure, the three terms in the energy norm $\|\cdot\|_{X(0,T)}$ do not have the same convergence rate in space. Indeed, the $L^2(0,T;V_d)$-norm for the pressure is dominant and converges at the same rate as the bound derived in the previous section. The purpose of this section is to derive a sharper bound for the $L^\infty(0,T;V_a)$-norm of the displacement and the $L^\infty(0,T;L_d)$-norm of the pressure which converges at the optimal rate.

For all $n \in \{0, \dots, N\}$, we define the elliptic reconstruction of $(u_h^n, p_h^n) \in V_{ah} \times V_{dh}$ as the functions $(U^n, P^n) \in V_a \times V_d$ such that

$$a(U^n, v) - b(v, P^n) = a(u_h^n, P_{ah}v) - b(P_{ah}v, p_h^n), \qquad \forall v \in V_a, \tag{4.31}$$
$$d(P^n, q) = d(p_h^n, P_{dh}q) - b(\delta_t u_h^n, q - P_{dh}q), \qquad \forall q \in V_d, \tag{4.32}$$

where P_{ah} (resp., P_{dh}) denotes the L_a-orthogonal projection from V_a onto V_{ah} (resp., the L_d-orthogonal projection from V_d onto V_{dh}). Henceforth, we use the convention that $\delta_t u_h^0 = 0$ and $\delta_t p_h^0 = 0$. Observe that

$$AU^n + B^* P^n = A_h u_h^n + B_h^* p_h^n, \tag{4.33}$$
$$DP^n = D_h p_h^n + (B - B_h)\delta_t u_h^n. \tag{4.34}$$

Indeed, for all $q \in V_d$,

$$\begin{aligned}\langle DP^n, q\rangle_d = -d(P^n, q) &= -d(p_h^n, P_{dh}q) + b(\delta_t u_h^n, q - P_{dh}q) \\ &= (D_h p_h^n, P_{dh}q)_{L_d} + ((B - B_h)\delta_t u_h^n, q)_{L_d} = (D_h p_h^n + (B - B_h)\delta_t u_h^n, q)_{L_d},\end{aligned} \tag{4.35}$$

since $D_h p_h^n \in V_{dh}$. Equation (4.33) is proved similarly. It is worthwhile to point out at this stage a nontrivial difference with the elliptic reconstruction technique for the heat equation, namely that the pressure reconstruction must also account for the fact that the divergence of the displacement is not in the pressure finite element space, thus the presence of the time-derivative in the right-hand side of (4.32) and (4.34).

The key idea to bound the errors $\|u^n - u_h^n\|_a$ and $\|p^n - p_h^n\|_c$ is to consider the decompositions

$$u^n - u_h^n = \omega_u^n - \rho_u^n, \qquad \omega_u^n = U^n - u_h^n, \qquad \rho_u^n = U^n - u^n, \tag{4.36}$$
$$p^n - p_h^n = \omega_p^n - \rho_p^n, \qquad \omega_p^n = P^n - p_h^n, \qquad \rho_p^n = P^n - p^n. \tag{4.37}$$

The quantities ω_u^n and ω_p^n can be handled following the *a posteriori* error analysis for the steady problem, while the quantities ρ_u^n and ρ_p^n can be bounded in terms of ω_u^n and ω_p^n and other computable quantities. The analysis requires a refinement of Hypothesis 2.2 by localizing the approximation property to mesh cells. In the context of poroelasticity, this assumption means that the steady-state version of the pressure equation yields elliptic regularity and that the finite element space V_{dh} satisfies the usual approximation properties.

Hypothesis 4.1. *There exist a constant c_8 and a positive real number δ such that for all $r \in L_d$, the unique solution $\phi \in V_d$ of the dual problem $d(q,\phi) = c(r,q)$ for all $q \in V_d$, is such that there is $\phi_h \in V_{dh}$ satisfying*

$$\sum_{T\in\mathcal{T}_h} h_T^{-2\delta}[h_T^{-2}\|\phi-\phi_h\|^2_{L_d(T)} + h_T^{-1}\|\phi-\phi_h\|^2_{L_d(\partial T)}] \le c_8\|r\|_c^2. \tag{4.38}$$

We first consider the quantities ω_u^n and ω_p^n.

Lemma 4.1. *In the above framework, the following holds for all $n \in \{0,\dots,N\}$,*

$$\|\omega_u^n\|_a^2 \le 2c_6\widehat{\mathcal{E}}_u^n + 2\beta^2 c_8\widehat{\mathcal{E}}_{p,\delta}^n, \tag{4.39}$$

$$\|\omega_p^n\|_c^2 \le c_8\widehat{\mathcal{E}}_{p,\delta}^n. \tag{4.40}$$

Proof. (i) Bound on $\|\omega_p^n\|_c$. Let ϕ be the dual solution associated with the data $r := \omega_p^n$ in Hypothesis 4.1. Then,

$$\|\omega_p^n\|_c^2 = c(r,r) = d(r,\phi) = d(r,\phi-\phi_h),$$

since owing to (4.32), $d(r,\phi_h) = d(P^n - p_h^n, \phi_h) = 0$ for $\phi_h \in V_{dh}$. Using (4.12) and (4.38) leads to

$$\|\omega_p^n\|_c^2 \le c_8 \sum_{T\in\mathcal{T}_h} h_T^{2\delta}[h_T^2\|D_{\mathrm{loc}}\omega_p^n\|^2_{L_d(T)} + h_T\|J_d\omega_p^n\|^2_{L_d(\partial T)}].$$

Using (2.15) and (4.34) yields for $n \ge 1$,

$$D_{\mathrm{loc}}\omega_p^n = DP^n - D_{\mathrm{loc}}p_h^n = D_h p_h^n + (B - B_h)\delta_t u_h^n - D_{\mathrm{loc}}p_h^n = -R_{ph}^n,$$

and this relation also holds for $n = 0$ by definition of R_{ph}^0. In addition, for all $n \ge 0$, $J_d P^n = 0$ since $DP^n \in V_{dh} \subset L_d$. As a result,

$$\|\omega_p^n\|_c^2 \le c_8\widehat{\mathcal{E}}_{p,\delta}^n.$$

(ii) Bound on $\|\omega_u^n\|_a$. Observe that

$$\begin{aligned}\|\omega_u^n\|_a = \sup_{0\ne v\in V_a}\frac{a(\omega_u^n,v)}{\|v\|_a} &\le \sup_{0\ne v\in V_a}\left(\frac{a(\omega_u^n,v)-b(v,\omega_p^n)}{\|v\|_a}\right) + \beta\|\omega_p^n\|_c\\ &= \sup_{0\ne v\in V_a}\left(\frac{a(\omega_u^n, v-i_{ah}(v)) - b(v-i_{ah}(v),\omega_p^n)}{\|v\|_a}\right) + \beta\|\omega_p^n\|_c,\end{aligned}$$

owing to (4.31) since $i_{ah}(v) \in V_{ah}$. Using (4.11), (4.13), and (4.14) leads to

$$\|\omega_u^n\|_a^2 \le 2c_6 \sum_{T\in\mathcal{T}_h}[h_T^2\|A_{\mathrm{loc}}\omega_u^n + B^*\omega_p^n\|^2_{L_a(T)} + h_T\|J_a\omega_u^n\|^2_{L_a(\partial T)}] + 2\beta^2\|\omega_p^n\|_c^2.$$

Owing to (2.14) and (4.33), $AU^n + B^*P^n = A_h u_h^n + B_h^* p_h^n = -f_h^n$ so that $A_{\mathrm{loc}}\omega_u^n + B^*\omega_p^n = AU^n + B^*P^n - A_{\mathrm{loc}}u_h^n - B^*p_h^n = -R_{uh}^n$. Moreover, $J_a\omega_u^n = -J_{uh}^n$ since $AU^n \in L_a$. This yields (4.39). □

We now turn our attention to the quantities ρ_u^n and ρ_p^n. Let

$$\widehat{\mathcal{E}}_{\mathrm{dat}} = \|U^0 - u_0\|_a^2 + \|P^0 - p_0\|_c^2 + 2\mathcal{E}(f,g), \tag{4.41}$$

$$\widehat{\mathcal{E}}_{\mathrm{spc}} = \sum_{m=1}^{N} \tau_m [2\gamma^2 c_8 (1+2\beta^4)\widehat{\mathcal{E}}_{p,\delta}^m(\delta_t) + 4\beta^2\gamma^2 c_6 \widehat{\mathcal{E}}_u^m(\delta_t)], \tag{4.42}$$

$$\widehat{\mathcal{E}}_{\mathrm{tim}} = \tfrac{2}{3}\tilde{\gamma}^2\tau_1 \|C\delta_t p_h^1 + B\delta_t u_h^1 - D_h p_{0h} - g_h^1\|_{L_d}^2 + \sum_{m=2}^{N} \tfrac{2}{3}\tilde{\gamma}^2\tau_m^3 \|\delta_t(C\delta_t p_h^m + B\delta_t u_h^m - g_h^m)\|_{L_d}^2. \tag{4.43}$$

Lemma 4.2. *The following holds for all $n \in \{1,\dots,N\}$,*

$$\tfrac{1}{2}\|\rho_u^n\|_a^2 + \tfrac{1}{2}\|\rho_p^n\|_c^2 \le \widehat{\mathcal{E}}_{\mathrm{dat}} + \widehat{\mathcal{E}}_{\mathrm{spc}} + \widehat{\mathcal{E}}_{\mathrm{tim}}. \tag{4.44}$$

Proof. The bounds on ρ_u^n and ρ_p^n rely on the stability properties of the continuous problem. Thus, it is again convenient to handle equations holding a.e. in $[0,T]$ rather than at the discrete times $\{t_n\}_{n=0}^N$. Let U_τ (resp., P_τ) be the continuous and piecewise affine function in time such that for all $n \in \{0,\dots,N\}$, $U_\tau^n = U^n$ (resp., $P_\tau^n = P^n$). Let $\omega_{u\tau}$ and $\omega_{p\tau}$ be constructed in a similar way from $\{\omega_u^n\}_{n=0}^N$ and $\{\omega_p^n\}_{n=0}^N$. Define $\rho_{u\tau} = U_\tau - u$ and $\rho_{p\tau} = P_\tau - p$. Observe that for a.e. $t \in [0,T]$ and for all $v \in V_a$,

$$\begin{aligned} a(\rho_{u\tau}, v) - b(v, \rho_{p\tau}) &= a(U_\tau, v) - b(v, P_\tau) - (f,v)_{L_a} \\ &= a(u_{h\tau}, P_{ah}v) - b(P_{ah}v, p_{h\tau}) - (f,v)_{L_a} \\ &= (f_{h\tau}, P_{ah}v)_{L_a} - (f,v)_{L_a} = (f_{h\tau} - f, v)_{L_a}, \end{aligned}$$

while for all $q \in V_d$,

$$\begin{aligned} c(\partial_t\rho_{p\tau}, q) + b(\partial_t\rho_{u\tau}, q) + d(\rho_{p\tau}, q) &= c(\partial_t P_\tau, q) + b(\partial_t U_\tau, q) + d(P_\tau, q) - (g,q)_{L_d} \\ &= c(\partial_t\omega_{p\tau}, q) + b(\partial_t\omega_{u\tau}, q) + (D\pi^0 P_\tau, q)_{L_d} + d(P_\tau, q) + (\pi^0 g_{h\tau} - g, q)_{L_d} \\ &= c(\partial_t\omega_{p\tau}, q) + b(\partial_t\omega_{u\tau}, q) + (D(\pi^0 P_\tau - P_\tau), q)_{L_d} + (\pi^0 g_{h\tau} - g, q)_{L_d}, \end{aligned}$$

where $\pi^0 P_\tau$ is the piecewise constant function in time equal to P^n on I_n for all $n \in \{1,\dots,N\}$. Indeed, on each time interval I_n,

$$\begin{aligned} c(\partial_t p_{h\tau}, q) + b(\partial_t u_{h\tau}, q) &= (C\delta_t p_h^n + B\delta_t u_h^n, q)_{L_d} \\ &= (g_h^n + D_h p_h^n - B_h\delta_t u_h^n + B\delta_t u_h^n, q)_{L_d} = (g_h^n + DP^n, q)_{L_d}. \end{aligned}$$

Testing the above equations with $v := \partial_t\rho_{u\tau}$ and $q := \rho_{p\tau}$ yields

$$\begin{aligned} \tfrac{1}{2}d_t\|\rho_{u\tau}\|_a^2 + \tfrac{1}{2}d_t\|\rho_{p\tau}\|_c^2 + \|\rho_{p\tau}\|_d^2 = (f_{h\tau} - f, \partial_t\rho_{u\tau})_{L_a} + c(\partial_t\omega_{p\tau}, \rho_{p\tau}) + b(\partial_t\omega_{u\tau}, \rho_{p\tau}) \\ + (D(\pi^0 P_\tau - P_\tau), \rho_{p\tau})_{L_d} + (\pi^0 g_{h\tau} - g, \rho_{p\tau})_{L_d}, \end{aligned}$$

so that

$$\begin{aligned} \tfrac{1}{2}d_t\|\rho_{u\tau}\|_a^2 + \tfrac{1}{2}d_t\|\rho_{p\tau}\|_c^2 \le (f_{h\tau} - f, \partial_t\rho_{u\tau})_{L_a} + \gamma^2\|\partial_t\omega_{p\tau}\|_c^2 + \beta^2\gamma^2\|\partial_t\omega_{u\tau}\|_a^2 \\ + \tilde{\gamma}^2\|D(\pi^0 P_\tau - P_\tau)\|_{L_d}^2 + \|\pi^0 g_{h\tau} - g\|_{d'}^2. \end{aligned}$$

Proceeding as usual (details are skipped for brevity) yields for all $n \in \{1,\dots,N\}$,

$$\tfrac{1}{2}\|\rho_u^n\|_a^2 + \tfrac{1}{2}\|\rho_p^n\|_c^2 \le \widehat{\mathcal{E}}_{\mathrm{dat}} + \int_0^T 2\gamma^2\|\partial_t\omega_{p\tau}(s)\|_c^2\mathrm{d}s + \int_0^T 2\beta^2\gamma^2\|\partial_t\omega_{u\tau}(s)\|_a^2\mathrm{d}s + \int_0^T 2\tilde{\gamma}^2\|D(\pi^0 P_\tau - P_\tau)(s)\|_{L_d}^2\mathrm{d}s.$$

The second and third terms in the above right-hand side lead to the term $\widehat{\mathcal{E}}_{\rm spc}$ in (4.44) owing to the fact that $\omega_{p\tau}$ and $\omega_{u\tau}$ are piecewise affine in time so that for instance,

$$\int_0^T \|\partial_t \omega_{p\tau}(s)\|_c^2 {\rm d}s = \sum_{m=1}^N \tau_m^{-1} \|\omega_p^m - \omega_p^{m-1}\|_c^2.$$

By linearity and proceeding as in the proof of Lemma 4.1 yields

$$\tau_m^{-2}\|\omega_u^m - \omega_u^{m-1}\|_a^2 \leq 2c_6 \widehat{\mathcal{E}}_u^m(\delta_t) + 2\beta^2 c_8 \widehat{\mathcal{E}}_{p,\delta}^m(\delta_t) \qquad \text{and} \qquad \tau_m^{-2}\|\omega_p^m - \omega_p^{m-1}\|_c^2 \leq c_8 \widehat{\mathcal{E}}_{p,\delta}^m(\delta_t).$$

Finally, the last term in the above bound on ρ_u^n and ρ_p^n leads to the term $\widehat{\mathcal{E}}_{\rm tim}$ in (4.44) since

$$\int_0^T \|D(\pi^0 P_\tau - P_\tau)(s)\|_{L_d}^2 {\rm d}s = \sum_{m=1}^N \tfrac{1}{3}\tau_m \|D(P^m - P^{m-1})\|_{L_d}^2,$$

and for all $m \geq 0$, $DP^m = D_{\rm loc}\, p_h^m - R_{ph}^m$. □

Theorem 4.3. *For all $n \in \{1, \ldots, N\}$,*

$$\tfrac{1}{4}\|u^n - u_h^n\|_a^2 + \tfrac{1}{4}\|p^n - p_h^n\|_c^2 \leq \widehat{\mathcal{E}}_{\rm dat} + \widehat{\mathcal{E}}_{\rm spc} + \widehat{\mathcal{E}}_{\rm tim} + c_6 \widehat{\mathcal{E}}_u^n + c_8(\tfrac{1}{2} + \beta^2)\widehat{\mathcal{E}}_{p,\delta}^n. \tag{4.45}$$

Proof. Use Lemmas 4.1 and 4.2, and the triangle inequality. □

We will not attempt here to prove lower error bounds; this goes beyond the present scope. We will verify numerically in the following section that $\widehat{\mathcal{E}}_{\rm spc}$ and $\widehat{\mathcal{E}}_{\rm tim}$ yield the expected, optimal, order of convergence with respect to mesh size.

5. Numerical results

Two test cases are presented in this section. In the first one, an analytic solution is available; the aim of the test case is to verify the convergence rate of the derived error bounds and to evaluate the corresponding effectivity indices. The second test case is drawn from an excavation damage benchmark problem; its aim is to illustrate the generation of adaptive meshes in this context.

5.1. Test case with analytical solution

We consider the following analytical solution of (1.1)–(1.2) on the domain $\Omega = (0,1) \times (0,1)$,

$$u(t,x,y) = -\frac{\exp(-At)}{2\pi} \begin{bmatrix} \cos(\pi x)\sin(\pi y) \\ \sin(\pi x)\cos(\pi y) \end{bmatrix}, \quad p(t,x,y) = \exp(-At)\sin(\pi x)\sin(\pi y),$$

with $A = \frac{2\pi^2\kappa}{b+\frac{1}{M}}$, $\kappa = 0.05$, $b = 0.75$, $\frac{1}{M} = \frac{3}{28}$. The Lamé coefficients are $\lambda_1 = \frac{1}{2}$ and $\lambda_2 = \frac{1}{8}$, yielding a Poisson ratio $\nu = 0.4$ and a Young modulus $E = \frac{7}{20}$. Convergence rates in space are evaluated on a series of uniformly refined structured triangulations based on a boundary mesh step h_0. Finite elements are based on quadratics for the displacement and linears for the pressure.

Tables 1 and 2 present the convergence results (under space and time refinement respectively) for the approximation errors measured in various norms. All the convergence rates match those predicted by the *a priori* error analysis. In both cases, the total error is dominated by the $L^2(0,t;H_x^1)$-error on the p-component.

TABLE 1. Errors at final time and convergence rates under space refinement; $T = 0.1$, $\tau = 2.50\,\text{e–}4$.

h_0^{-1}	$\|u - u_{h\tau}\|_a$		$\|p - p_{h\tau}\|_c$		$\|p - p_{h\tau}\|_{L^2(0,T;V_d)}$		$\|p - \pi^0 p_{h\tau}\|_{L^2(0,T;V_d)}$	
4	8.12 e–3	–	5.66 e–3	–	2.75 e–2	–	2.75 e–2	–
8	2.15 e–3	1.92	1.49 e–3	1.92	1.45 e–2	0.92	1.45 e–2	0.92
16	5.34 e–4	2.01	3.73 e–4	2.00	7.33 e–3	0.98	7.33 e–3	0.98
32	1.32 e–4	2.02	9.21 e–5	2.01	3.68 e–3	0.99	3.68 e–3	0.99

TABLE 2. Errors at final time and convergence rates under time refinement; $T = 1$, $h_0 = 1/128$.

τ	$\|u - u_{h\tau}\|_a$		$\|p - p_{h\tau}\|_c$		$\|p - p_{h\tau}\|_{L^2(0,T;V_d)}$		$\|p - \pi^0 p_{h\tau}\|_{L^2(0,T;V_d)}$	
0.25	5.10 e–3	–	5.64 e–3	–	1.68 e–2	–	1.81 e–2	–
0.2	4.13 e–3	0.94	4.55 e–3	0.96	1.39 e–2	0.84	1.48 e–2	0.93
0.1	2.12 e–3	0.96	2.31 e–3	0.98	7.58 e–3	0.87	7.74 e–3	0.93
0.05	1.07 e–3	0.99	1.16 e–3	0.99	4.24 e–3	0.83	4.26 e–3	0.86

TABLE 3. *A posteriori* error bounds using the direct approach and convergence rates under space refinement; $T = 0.1$, $\tau = 2.50\,\text{e–}4$.

h_0^{-1}	η_1		η_2		η_3		η_4		$\mathcal{I}_{\text{eff}}$	$\mathcal{I}^*_{\text{eff}}$
4	6.34 e–2	–	9.53 e–2	–	8.17 e–3	–	1.45 e–3	–	3.13	11.43
8	3.33 e–2	0.93	2.57 e–2	1.89	2.75 e–3	1.57	7.67 e–4	0.92	2.18	15.33
16	1.71 e–2	0.96	6.63 e–3	1.96	7.13 e–4	1.94	3.89 e–4	0.98	1.70	23.81
32	8.62 e–3	0.98	1.68 e–3	1.98	1.80 e–4	1.99	1.96 e–4	0.99	1.45	41.19

To assess the *a posteriori* error bounds obtained with the direct approach, we evaluate the quantities

$$\eta_1 = \left(\sum_{m=1}^{N} \tau_m \widehat{\mathcal{E}}_{p,0}^m \right)^{\frac{1}{2}}, \quad \eta_2 = \sup_{0 \le m \le N} (\widehat{\mathcal{E}}_u^m)^{\frac{1}{2}}, \quad \eta_3 = \sum_{m=1}^{N} \tau_m (\widehat{\mathcal{E}}_u^m(\delta_t))^{\frac{1}{2}}, \quad \eta_4 = \left(\sum_{m=1}^{N} \tau_m \|p_h^m - p_h^{m-1}\|_d^2 \right)^{\frac{1}{2}}, \tag{5.1}$$

as well as the effectivity indices

$$\mathcal{I}_{\text{eff}} = \frac{\eta_1 + \eta_2 + \eta_3 + \eta_4}{(\|p - p_{h\tau}\|^2_{L^2(0,T;V_d)} + \|p - \pi^0 p_{h\tau}\|^2_{L^2(0,T;V_d)})^{1/2}}, \qquad \mathcal{I}^*_{\text{eff}} = \frac{\eta_1 + \eta_2 + \eta_3 + \eta_4}{\|u^N - u_h^N\|_a + \|p^N - p_h^N\|_c} \cdot \tag{5.2}$$

Recall that η_1, η_2, and η_3 are associated with the space error indicators, see (4.25), and that η_4 is associated with the time error indicators, see (4.10). For brevity, we concentrate here on these quantities. Tables 3 and 4 present the results obtained under space and time refinement, respectively. All the observed convergence rates match the theoretical predictions. Moreover, the effectivity index $\mathcal{I}_{\text{eff}}$ takes values between 2 and 3, indicating that the present *a posteriori* error bounds behave quite satisfactorily to control the pressure error in the $L^2(0,t;H_x^1)$-norm. As expected, the situation is quite different if one attempts to control the displacement error in the $L^\infty(0,t;H_x^1)$-norm. As reflected by the effectivity index $\mathcal{I}^*_{\text{eff}}$ which increases as the mesh is refined, this latter error converges faster to zero than the bound derived with the direct approach.

TABLE 4. *A posteriori* error bounds using the direct approach and convergence rates under time refinement; $T = 1$, $h_0 = 1/128$.

τ	η_1	η_2	η_3	η_4		$\mathcal{I}_{\text{eff}}$	$\mathcal{I}^*_{\text{eff}}$
0.25	4.33 e–3	1.07 e–4	1.15 e–4	4.70 e–2	–	1.48	4.80
0.2	4.37 e–3	1.02 e–4	1.04 e–4	3.85 e–2	0.90	1.51	4.95
0.1	4.45 e–3	9.93 e–5	8.25 e–5	2.01 e–2	0.93	1.62	5.58
0.05	4.49 e–3	1.02 e–4	7.38 e–5	1.03 e–2	0.96	1.77	6.69

TABLE 5. *A posteriori* error bounds using elliptic reconstruction and convergence rates under space refinement; $T = 1$, $\tau = 0.1$.

h_0^{-1}	η_5		η_6	η_7		η_8		$\mathcal{J}_{\text{eff}}$
4	1.34 e–0	–	1.14 e–2	1.55 e–2	–	2.18 e–2	–	169.43
8	2.01 e–1	2.74	7.82 e–3	5.70 e–3	1.44	5.83 e–3	1.90	120.26
16	4.36 e–2	2.21	7.58 e–3	1.79 e–3	1.67	1.50 e–3	1.96	97.49
32	1.04 e–2	2.07	7.55 e–3	4.93 e–4	1.86	3.77 e–4	1.98	19.59

To assess the *a posteriori* error bound using elliptic reconstruction, we evaluate the quantities

$$\eta_5 = \left(\sum_{m=1}^{N} \tau_m (\widehat{\mathcal{E}}^m_{p,1}(\delta_t) + \widehat{\mathcal{E}}^m_u(\delta_t)) \right)^{\frac{1}{2}}, \qquad \eta_7 = (\widehat{\mathcal{E}}^N_u)^{\frac{1}{2}}, \qquad \eta_8 = (\widehat{\mathcal{E}}^N_{p,1})^{1/2}, \tag{5.3}$$

$$\eta_6 = \left(\tau_1 \| C\delta_t p^1_h + B\delta_t u^1_h - D_h p_{0h} - g^1_h \|^2_c + \sum_{m=2}^{N} \tau^3_m \| \delta_t (C\delta_t p^m_h + B\delta_t u^m_h - g^m_h) \|^2_c \right)^{\frac{1}{2}}, \tag{5.4}$$

as well as the effectivity index

$$\mathcal{J}_{\text{eff}} = \frac{\eta_5 + \eta_6 + \eta_7 + \eta_8}{\|u^N - u^N_h\|_a + \|p^N - p^N_h\|_c}. \tag{5.5}$$

Recall that η_5 is associated with the space error indicators, see (4.42), and that η_6 is associated with the time error indicators, see (4.43). Moreover, η_7 and η_8 stem from the difference between the discrete solution and its elliptic reconstruction, see Lemma 4.1. Table 5 presents the results obtained under space refinement. The observed orders of convergence match theoretical predictions, with a slight super-convergence for η_5 and a slight sub-convergence for η_7 on the (very) coarse meshes. The quantity η_6, which is related to the time error, remains at a fairly constant value under space refinement. Additional tests (not reported here for brevity; see [13]) indicate that η_6 converges with order close to 1 under time refinement. Finally, we observe that the effectivity index with elliptic reconstruction is much larger than that with the direct approach, especially on coarse meshes. This can be expected since the elliptic reconstruction technique uses the stability constant of an adjoint problem, and this latter constant is usually large.

5.2. **Excavation damage test case**

This test case is drawn from [5]. A two-dimensional setting is considered. The computational domain is a square of length 60 m minus a circular sector centered at the lower left corner with angle $\frac{\pi}{2}$ representing the excavated part. The model parameters are a Young modulus $E = 5800$ MPa, a Poisson ratio $\nu = 0.3$, an hydraulic conductivity $\kappa = 1.02 \times 10^{-16}$ $\text{m}^2{\cdot}\text{Pa}^{-1}{\cdot}\text{s}^{-1}$, a Biot modulus $\frac{1}{M} = 2.69 \times 10^{-11}$ Pa^{-1}, and a Biot-Willis coefficient $b = 0.8$. Body forces result from gravity with a density $\rho = 10^3$ $\text{kg}{\cdot}\text{m}^{-3}$. The simulation time is 1.5×10^6 s (roughly 17 days). Mixed Dirichlet-Neumann boundary conditions are enforced. The order of magnitude for the pressure is about 5 MPa and that for the displacement is about 5 cm. The problem is solved

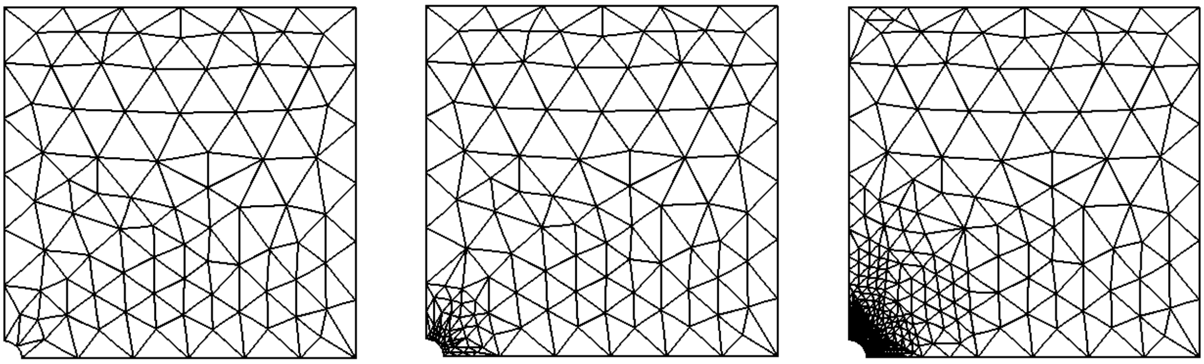

Figure 1. Initial mesh (left) and adaptive meshes generated after two (center) and four (right) steps of the refinement procedure.

Table 6. *A posteriori* error bounds for the excavation damage test case on adaptive meshes; the second column reports the number of mesh cells on each adaptive mesh.

Step	# mesh cells	η_1	η_2	η_3	η_4
1	196	1.25 e–1	4.24 e–1	4.24 e–1	9.19 e–7
2	278	2.65 e–2	7.63 e–2	7.64 e–2	1.24 e–6
3	392	1.76 e–2	3.46 e–2	3.46 e–2	1.31 e–6
4	564	1.50 e–2	2.00 e–2	2.00 e–2	1.32 e–6
5	890	1.13 e–2	1.18 e–2	1.18 e–2	1.34 e–6
6	1368	9.09 e–3	7.72 e–3	7.72 e–3	1.34 e–6
7	2268	7.04 e–3	4.69 e–3	4.69 e–3	1.36 e–6

in non-dimensional form using quadratics for the displacement and linears for the pressure. The time step is kept fixed at $\tau = 10^5$ s; this value is small enough to keep time discretization errors negligible (see Tab. 6).

Adaptive meshes are constructed using the space error indicators derived using the direct approach. On a given mesh, the approximate solution is constructed until final time, then mesh cells are marked (the 5% mesh cells yielding the largest contributions to the global error bound are marked), and a new mesh is generated. This algorithm is considered here solely for illustration purposes, being understood that further work is needed to optimize the computation using dynamically adapted meshes. The initial mesh and those generated after two and four steps of the above procedure are presented in Figure 1. The values taken by η_1, η_2, η_3 and η_4, see (5.1), are reported in Table 6. The quantities η_1, η_2 and η_3, which are associated with spatial errors, decrease faster than the number of adaptive mesh cells, while η_4, which is associated with time errors, remains negligible. The quantities η_2 and η_3 take similar values since for the present problem, both are dominated by the error at initial time. Further results can be found in [13].

6. Conclusions

We have analyzed Euler-Galerkin approximations to coupled elliptic-parabolic problems with application to poroelasticity. The *a priori* error analysis shows that equal-order polynomial interpolation can be used for the displacement and for the pressure, but that using polynomials of one degree higher for the displacement than for the pressure is preferable since this equilibrates the convergence rate of all the terms in the energy norm. Furthermore, we have obtained residual-type, energy-norm *a posteriori* error bounds for the displacement and for the pressure. Elaborating on the ideas of Verfürth, we have established an equivalence result between the residual measured in a dual norm and the error measured in the energy norm plus some time-derivatives measured in weaker norms. The *a posteriori* error bound converges optimally with respect to mesh size when compared with the dominant term in the error energy norm, namely that related to the $L^2(0,t;H^1_x)$-norm for the pressure. To obtain an optimally convergent bound for the two other terms in the error energy norm, namely those related to the $L^\infty(0,t;H^1_x)$-norm for the displacement and the $L^\infty(0,t;L^2_x)$-norm for the pressure, we have extended the elliptic reconstruction technique introduced by Makridakis and Nochetto for linear parabolic problems. One important difference with the heat equation is that the pressure reconstruction must also account for the fact that the divergence of the displacement is not in the pressure finite element space.

Acknowledgements. The authors are grateful to O. Boiteau, C. Chavant and G. Nicolas (EdF R&D) for fruitful discussions and instrumental help for the implementation in the software *Code_Aster*. The authors are also grateful to S. Nicaise (University of Valenciennes) and A. Veeser (University of Milan) for several comments on this work and to one of the referees for his constructive remarks.

References

[1] I. Babuška, M. Feistauer and P. Šolín, On one approach to a posteriori error estimates for evolution problems solved by the method-of-lines. *Numer. Math.* **89** (2001) 225–256.
[2] R. Becker and R. Rannacher, An optimal control approach to a posteriori error estimation in finite element methods. *Acta Numer.* **10** (2001) 1–102.
[3] A. Bergam, C. Bernardi and Z. Mghazli, *A posteriori* analysis of the finite element discretization of some parabolic equations. *Math. Comp.* **74** (2005) 1117–1138.
[4] M.A. Biot, General theory of three-dimensional consolidation. *J. Appl. Phys.* **12** (1941) 155–169.
[5] C. Chavant and A. Millard, *Simulation d'excavation en comportement hydro-mécanique fragile*. Technical report, EDF R&D/AMA and CEA/DEN/SEMT (2007) http://www.gdrmomas.org/ex_qualifications.html.
[6] Z. Chen and J. Feng, An adaptive finite element algorithm with reliable and efficient error control for linear parabolic problems. *Math. Comp.* **73** (2004) 1167–1193.
[7] P. Clément, Approximation by finite element functions using local regularization. *RAIRO Anal. Numér.* **9** (1975) 77–84.
[8] K. Eriksson and C. Johnson, Adaptive finite element methods for parabolic problems. I. A linear model problem. *SIAM J. Numer. Anal.* **28** (1991) 43–77.
[9] K. Eriksson and C. Johnson, Adaptive finite element methods for parabolic problems. II. Optimal error estimates in $l_\infty l_2$ and $l_\infty l_\infty$. *SIAM J. Numer. Anal.* **32** (1995) 706–740.
[10] A. Ern and J.-L. Guermond, *Theory and Practice of Finite Elements*, *Applied Mathematical Sciences* **159**. Springer-Verlag, New York (2004).
[11] O. Lakkis and Ch. Makridakis, Elliptic reconstruction and *a posteriori* error estimates for fully discrete linear parabolic problems. *Math. Comp.* **75** (2006) 1627–1658.
[12] Ch. Makridakis and R.H. Nochetto, Ellitpic reconstruction and *a posteriori* error estimates for elliptic problems. *SIAM J. Numer. Anal.* **41** (2003) 1585–1594.
[13] S. Meunier, *Analyse d'erreur a posteriori pour les couplages hydro-mécaniques et mise en œuvre dans Code_Aster*. Ph.D. Thesis, École nationale des ponts et chaussées, France (2007).
[14] M.A. Murad and A.F.D. Loula, Improved accuracy in finite element analysis of Biot's consolidation problem. *Comput. Meth. Appl. Mech. Engrg.* **95** (1992) 359–382.
[15] M.A. Murad and A.F.D. Loula, On stability and convergence of finite element approximations of Biot's consolidation problem. *Internat. J. Numer. Methods Engrg.* **37** (1994) 645–667.
[16] M.A. Murad, V. Thomée and A.F.D. Loula, Asymptotic behavior of semidiscrete finite-element approximations of Biot's consolidation problem. *SIAM J. Numer. Anal.* **33** (1996) 1065–1083.
[17] M. Picasso, Adaptive finite elements for a linear parabolic problem. *Comput. Methods Appl. Mech. Engrg.* **167** (1998) 223–237.
[18] R.L. Scott and S. Zhang, Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comp.* **54** (1990) 483–493.
[19] R.E. Showalter, Diffusion in deformable media. *IMA Volumes in Mathematics and its Applications* **131** (2000) 115–130.
[20] R.E. Showalter, Diffusion in poro-elastic media. *J. Math. Anal. Appl.* **251** (2000) 310–340.
[21] V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*. Springer-Verlag, Berlin (1997).
[22] R. Verfürth, *A posteriori* error estimations and adaptative mesh-refinement techniques. *J. Comput. Appl. Math.* **50** (1994) 67–83.
[23] R. Verfürth, *A Review of a Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*. Wiley, Chichester, UK (1996).
[24] R. Verfürth, *A posteriori* error estimates for finite element discretizations of the heat equation. *Calcolo* **40** (2003) 195–212.
[25] K. von Terzaghi, *Theoretical Soil Mechanics*. Wiley, New York (1936).
[26] M. Wheeler, *A priori* L_2 error estimates for Galerkin approximations to parabolic partial differential equations. *SIAM J. Numer. Anal.* **10** (1973) 723–759.