

## FINITE ELEMENT DISCRETIZATION OF DARCY'S EQUATIONS WITH PRESSURE DEPENDENT POROSITY

VIVETTE GIRAULT<sup>1</sup>, FRANÇOIS MURAT<sup>2</sup> AND ABNER SALGADO<sup>3</sup>

**Abstract.** We consider the flow of a viscous incompressible fluid through a rigid homogeneous porous medium. The permeability of the medium depends on the pressure, so that the model is nonlinear. We propose a finite element discretization of this problem and, in the case where the dependence on the pressure is bounded from above and below, we prove its convergence to the solution and propose an algorithm to solve the discrete system. In the case where the dependence on the pressure is exponential, we propose a splitting scheme which involves solving two linear systems, but parts of the analysis of this method are still heuristic. Numerical tests are presented, which illustrate the introduced methods.

**Mathematics Subject Classification.** 76S05, 65N30.

Received 15 January 2009.

Published online February 23, 2010.

### 1. INTRODUCTION

#### 1.1. Position of the paper

The system of equations commonly referred to as Darcy's law was obtained on an experimental basis by Darcy [14], more than 150 years ago. It approximates the balance of linear momentum of a fluid flowing through a porous rigid body and is the simplest model of flow of a viscous incompressible fluid through a porous medium. Darcy's equations were obtained rigorously by Homogenization; without being exhaustive, we refer the reader to the works of Ene and Sanchez-Palencia [16], Allaire [2], Cioranescu *et al.* [13], Pastukhova [30], and Skjetne and Auriault [34].

Recently, in [31], Rajagopal developed systematically a family of models within the framework of Mixture Theory, deriving first Darcy's system, and next relaxing one or more restrictions that were used in deriving this law. The steady nonlinear model studied in the present work is one of the numerous models obtained through this approach, see Section 3.5 of [31]. It is a much simplified version of a model of enhanced oil recovery, where oil is forced to flow through rocks by injecting steam at high pressure. This model is simplified because only one fluid is considered and the viscous and inertial effects are neglected, thus resulting in a steady system. On the other hand, it is nonlinear because the porosity of the solid medium is allowed to depend exponentially

---

*Keywords and phrases.* Porous media flows, Darcy equations, finite elements.

<sup>1</sup> UPMC Univ. Paris 06, UMR 7598, Laboratoire Jacques-Louis Lions, 75005 Paris, France. [girault@ann.jussieu.fr](mailto:girault@ann.jussieu.fr)

<sup>2</sup> CNRS, UMR 7598, Laboratoire Jacques-Louis Lions, 75005 Paris, France. [murat@ann.jussieu.fr](mailto:murat@ann.jussieu.fr)

<sup>3</sup> Department of Mathematics, Texas A&M University, College Station, TX 77843-3368, USA. [abnersg@math.tamu.edu](mailto:abnersg@math.tamu.edu)

on the pressure. Indeed, it has been observed experimentally that high variations on the pressure induce an exponential variation on the porosity of the medium.

Let  $\Omega$  be a bounded domain in  $\mathbb{R}^d$ , with  $d = 2, 3$ . The boundary,  $\partial\Omega$ , of this domain is divided into two parts  $\Gamma_w$  and  $\Gamma$ . We are interested in the following model, which as we have stated above was derived by Rajagopal [31],

$$\begin{cases} \alpha(p)\mathbf{u} + \nabla p = \mathbf{f}, & \text{in } \Omega, \\ \nabla \cdot \mathbf{u} = 0, & \text{in } \Omega, \\ p = 0, & \text{on } \Gamma_w, \\ \mathbf{u} \cdot \mathbf{n} = g & \text{on } \Gamma, \end{cases} \quad (1.1)$$

where the unknowns are the velocity  $\mathbf{u}$  and the pressure  $p$  of the fluid. The function  $\alpha$  is the permeability of the medium; for simplicity, it is assumed homogeneous, but it depends exponentially on the pressure:

$$\alpha(\xi) = \alpha_0 e^{\gamma \xi}, \quad (1.2)$$

for positive parameters  $\alpha_0$  and  $\gamma$ . The homogeneous boundary condition in the third row of (1.1) is just introduced to simplify the discussion. More generally, a non homogeneous boundary condition can be prescribed on  $p$ :  $p = p_w$  on  $\Gamma_w$ . Owing to the nature of  $\alpha(p)$ , the subsequent analysis readily carries over to this case for adequately smooth boundary data  $p_w$ ; see Remark 2.2.

For the sake of brevity, in what follows we shall refer to equations (1.1) simply as the nonlinear Darcy equations. Of course, there are other nonlinear Darcy's model, such as the well-known Forchheimer model introduced by Forchheimer in [19]. The reader can refer to a discrete scheme, closely related to the one studied here, for a steady Forchheimer model studied by Girault and Wheeler in [21].

The analysis of the nonlinear Darcy equations is difficult because of the exponential nonlinearity. Following the work of Azaïez *et al.* in [3], we propose in a first part to discretize (1.1) when the function  $\alpha$  is truncated above and below. We introduce a straightforward finite element scheme, such as  $\mathbb{P}_{k-1}$  for each component of the velocity and  $\mathbb{P}_k$  for the pressure, similar to the scheme studied by Roberts and Thomas in [32] and by Kim and Park in [27]. When the exact solution is sufficiently small so that it satisfies a sufficient condition for uniqueness, we establish optimal *a priori* error estimates, and geometric convergence of a successive approximation algorithm for computing the discrete solution. We also study the case when the exact solution is nonsingular in the sense of Brezzi *et al.* [11], but is not necessarily unique. We give sufficient conditions for the finite element scheme to have a nonsingular solution, establish convergence and *a priori* error estimates, and study the convergence of Newton's algorithm for computing this solution. In particular, we prove that Newton's method converges quadratically, but not uniformly. This confirms the convergence analysis for nonlinear second order elliptic problems studied by Douglas and Dupont in [15] and by Park in [29].

The problem with fully exponential porosity is studied in a second part. To begin with, the velocity is eliminated by

- (1) dividing the equation by the exponential,
- (2) taking the divergence of the equation,
- (3) and making a change in variable.

This splits the problem into exactly two consecutive *linear* equations: first a diffusion-convection-reaction equation and next a linear Darcy system. These are discretized by an easy variant of the finite element scheme used in the first approach. The analysis of each discrete linear system is straightforward, but the global analysis of the complete algorithm is still an open problem.

We present numerical experiments for testing each method. As expected, the split algorithm (that involves no iteration) performs better than the successive approximation algorithm.

The rest of this work is organized as follows. In Section 1.2 we set up the notation and conventions that will be used in the sequel. The next three sections are devoted to the case where the function  $\alpha$  is bounded above and below and is uniformly Lipschitz-continuous. In Section 2, we present some results established in [3] on the existence of the solution to the nonlinear Darcy equation, as well as some sufficient conditions for uniqueness

and regularity. Section 2.2 gives sufficient conditions for the existence of a nonsingular solution. Section 3 is devoted to the analysis of the discrete problem. In the case where the solution is unique, we prove in Section 3.1 optimal error estimates and convergence of the successive approximation algorithm. In Section 3.2, we approximate nonsingular solutions, and we analyze Newton's method when used to find the discrete nonsingular solution. The splitting method is developed and studied in Section 4. Finally, Section 5 gives some numerical experiments which illustrate the theory and methods developed in the previous sections. We consider the method developed in Section 3.1 and show its performance on a series of model problems for different types of finite element spaces. We also test the algorithm of Section 4 for various types of finite element spaces on each of the sub-problems involved.

### 1.2. Notation and conventions

Henceforth we denote by  $\Omega$  a bounded connected domain in  $\mathbb{R}^d$ , with  $d = 2$  or  $3$ . As usual, we denote by  $L^p(\Omega)$  the space of Lebesgue integrable functions with exponent  $p \in [1, \infty]$  defined on  $\Omega$ , normed, for  $1 \leq p < \infty$ , by

$$\forall v \in L^p(\Omega), \|v\|_{L^p(\Omega)} = \left( \int_{\Omega} |v|^p \right)^{1/p},$$

with the usual extension when  $p = \infty$ . By  $W_p^s(\Omega)$ , for an integer  $s$ , we denote the Sobolev space of functions in  $L^p(\Omega)$  with partial derivatives of order up to  $s$  in  $L^p(\Omega)$ , namely

$$W_p^s(\Omega) = \{v \in L^p(\Omega); \partial^k v \in L^p(\Omega) \forall |k| \leq s\},$$

equipped with the seminorm

$$|v|_{W_p^s(\Omega)} = \left[ \sum_{|k|=s} \int_{\Omega} |\partial^k v|^p \right]^{1/p},$$

and norm (for which it is a Banach space)

$$\|v\|_{W_p^s(\Omega)} = \left[ \sum_{0 \leq |k| \leq s} |v|_{W_p^k(\Omega)}^p \right]^{1/p}.$$

When  $s$  is not an integer,  $W_p^s(\Omega)$  is defined by interpolation (*cf.* Lions and Magenes [28], or Berg and Löfström [6]). In this case, there are several definitions with equivalent norms. Here, we choose the following seminorm and norm: let  $s = m + s'$  for an integer  $m \geq 0$  and  $0 < s' < 1$ , then we set

$$|v|_{W_p^s(\Omega)} = \left( \sum_{|k|=m} \int_{\Omega} \int_{\Omega} \frac{|\partial^k v(\mathbf{x}) - \partial^k v(\mathbf{y})|^p}{|\mathbf{x} - \mathbf{y}|^{d+ps'}} \, d\mathbf{x} \, d\mathbf{y} \right)^{1/p},$$

$$\|v\|_{W_p^s(\Omega)} = \left( \|v\|_{W_p^m(\Omega)}^p + |v|_{W_p^s(\Omega)}^p \right)^{1/p}.$$

When  $p = 2$ , we set  $H^s(\Omega) := W_2^s(\Omega)$ , for any  $s$ . In particular, we have the following trace property on a domain  $\Omega$  with Lipschitz-continuous boundary  $\partial\Omega$ : if  $v$  belongs to  $H^s(\Omega)$  for some real number  $s \in ]1/2, 1]$ , then its trace on  $\partial\Omega$  belongs to  $H^{s-1/2}(\partial\Omega)$  (*cf.* for instance Grisvard [23], Thm. 1.5.1.2), and there exists a constant  $C$  such that

$$\forall v \in H^s(\Omega), \|v\|_{H^{s-1/2}(\partial\Omega)} \leq C \|v\|_{H^s(\Omega)}.$$

Finally, if  $\Gamma$  is a subset of  $\partial\Omega$  with positive measure,  $|\Gamma| > 0$ , we say that a function  $g$  in  $H^{1/2}(\Gamma)$  belongs to  $H_{00}^{1/2}(\Gamma)$  if its extension by zero to  $\partial\Omega$  belongs to  $H^{1/2}(\partial\Omega)$ . For a discussion on this space see Tartar [35] for instance.

For vector-valued functions we use boldface and the spaces of these functions are denoted, for instance, by  $L^p(\Omega)^d$ .

Whenever  $E$  is a normed space,  $\|\cdot\|_E$  denotes its norm and  $E'$  its dual. We use the convention that when taking the supremum of any quantity  $A$  over the elements of a space  $E$ ,

$$\sup_{x \in E} A$$

actually means the supremum over the elements of  $E$  which are nonzero.

By  $C$  we denote a constant, the value of which might change at each occurrence. This constant may depend on the problem data. When discussing discretization, this constant can also depend on the exact solution of the problem, but it does not depend on the discretization parameter  $h$ .

The constant in the Sobolev embedding  $H^1(\Omega) \hookrightarrow L^6(\Omega)$  (see Adams [1] or Tartar [35]) shall appear repeatedly and, therefore, we assign it the symbol  $C(\Omega)$ . More precisely,  $C(\Omega)$  is the smallest constant such that

$$\|q\|_{L^6(\Omega)} \leq C(\Omega) |q|_{H^1(\Omega)}, \quad \forall q \in H^1(\Omega).$$

Finally, we must say that all the reasoning carried on in the following sections assume that the space dimension  $d$  equals three. This is done only for the sake of definiteness. The reader can easily verify that similar arguments and less restrictive assumptions can be used to obtain the same results in the case when  $d = 2$ .

## 2. ANALYSIS OF THE PROBLEM

Before considering the discretization of problem (1.1) we will discuss some properties of its exact solution, namely, its existence and sufficient conditions for this solution to be globally unique and possess certain smoothness properties. When the nonlinear Darcy equations have more than one solution we shall discuss so-called *nonsingular solutions*, in the sense of [11]. This shall prove useful for the development and analysis of the discretization.

### 2.1. Variational formulation

We intend to study problem (1.1) under the following assumptions:

- The domain  $\Omega$  has a Lipschitz-continuous boundary  $\partial\Omega$  divided into two parts  $\Gamma_w$  and  $\Gamma$ , also with Lipschitz continuous boundaries.
- The part of the boundary  $\Gamma_w$  has positive surface measure.
- The function  $\alpha$  is continuous from  $\mathbb{R}$  to  $\mathbb{R}$  and there are two positive constants  $\alpha_{\min}$  and  $\alpha_{\max}$  such that

$$\alpha_{\min} \leq \alpha(\xi) \leq \alpha_{\max}, \quad \forall \xi \in \mathbb{R}. \tag{2.1}$$

- The function  $\alpha$  is uniformly Lipschitz-continuous on  $\mathbb{R}$ . That is, there is a constant  $L_\alpha > 0$  such that for all  $\xi_1, \xi_2 \in \mathbb{R}$

$$|\alpha(\xi_1) - \alpha(\xi_2)| \leq L_\alpha |\xi_1 - \xi_2|. \tag{2.2}$$

**Remark 2.1.** Assumptions (2.1) and (2.2) are not true when the function  $\alpha$  is unbounded, as it is the case when it is exponential. However, these assumptions can be easily recovered by truncating the original function  $\alpha$ . Obviously, the solution of the truncated problem will not in general solve the original one. The analysis of how these two problems are related is beyond the scope of this paper.

It is well known that Darcy's equations have several variational formulations. We have chosen here the formulation that treats the boundary condition on  $p$  as an essential one and leads, roughly speaking, to taking  $\mathbf{u}$  in  $L^2(\Omega)^3$  and  $p$  in  $H^1(\Omega)$ . This choice is motivated by the fact that the forthcoming analysis of the nonlinear term  $\alpha(p)\mathbf{u}$  uses intensively the fact that  $p$  is in  $H^1(\Omega)$ . Moreover, a velocity  $\mathbf{u}$  in  $L^2(\Omega)^3$  is easily discretized. Another option consists in taking  $\mathbf{u}$  in  $H(\text{div}; \Omega)$  and  $p$  in  $L^2(\Omega)$ . Then  $\mathbf{u}$  must be discretized with mixed finite elements, with the advantage that this leads to a locally conservative scheme. But the drawback is that the analysis of the nonlinear term is not so clear.

Let us define the space

$$H_w^1(\Omega) := \{q \in H^1(\Omega) : q|_{\Gamma_w} = 0\},$$

and assume, for the sake of simplicity, that  $p_w = 0$ . Then the variational formulation is the following

Given  $\mathbf{f} \in L^2(\Omega)^3$  and  $g \in H_{00}^{1/2}(\Gamma)'$ , find a pair  $(\mathbf{u}, p) \in L^2(\Omega)^3 \times H_w^1(\Omega)$  such that

$$\begin{cases} a_p(\mathbf{u}, \mathbf{v}) + b(\mathbf{v}, p) = \int_{\Omega} \mathbf{f} \cdot \mathbf{v}, & \forall \mathbf{v} \in L^2(\Omega)^3, \\ b(\mathbf{u}, q) = \langle g, q \rangle_{\Gamma}, & \forall q \in H_w^1(\Omega). \end{cases} \tag{2.3}$$

The bilinear forms  $a_{\xi}(\cdot, \cdot)$  for any measurable function  $\xi$  on  $\Omega$  and  $b(\cdot, \cdot)$  are defined by

$$a_{\xi}(\mathbf{v}, \mathbf{w}) := \int_{\Omega} \alpha(\xi) \mathbf{v} \cdot \mathbf{w}, \tag{2.4}$$

$$b(\mathbf{v}, q) := \int_{\Omega} \mathbf{v} \cdot \nabla q, \tag{2.5}$$

and  $\langle \cdot, \cdot \rangle_{\Gamma}$  denotes the duality pairing between  $H_{00}^{1/2}(\Gamma)$  and its dual space  $H_{00}^{1/2}(\Gamma)'$ .

It is readily checked that under assumption (2.1) the forms  $a_{\xi}(\cdot, \cdot)$  and  $b(\cdot, \cdot)$  are continuous on  $L^2(\Omega)^3 \times L^2(\Omega)^3$  and  $L^2(\Omega)^3 \times H^1(\Omega)$  respectively. Thus, standard arguments yield the equivalence of problem (2.3) with the system (1.1) in the distribution sense.

**Remark 2.2.** The above variational formulation is defined for homogeneous boundary conditions:  $p_w = 0$ . Standard techniques (*i.e.* lifting arguments) allow us to reduce the case of nonhomogeneous Dirichlet boundary conditions on the pressure  $p$  to the present one. For this, it is sufficient to assume that  $p_w \in H^{1/2}(\Gamma_w)$  and notice that the function  $\xi \mapsto \alpha(\xi - \bar{p}_w)$ , where  $\bar{p}_w$  is a proper lifting of  $p_w$ , has the same properties as  $\xi \mapsto \alpha(\xi)$ . Hence, there is no loss of generality in considering only homogeneous Dirichlet boundary conditions.

The existence of a solution to problem (2.3) is studied in [3]. For the sake of completeness we list here the results that later prove useful for our purposes. Regarding existence we have the following theorem.

**Theorem 2.3.** Assume that the function  $\alpha$  satisfies assumption (2.1). Then, for any data  $(\mathbf{f}, g) \in L^2(\Omega)^3 \times H_{00}^{1/2}(\Gamma)'$  problem (2.3) has a solution  $(\mathbf{u}, p) \in L^2(\Omega)^3 \times H_w^1(\Omega)$ . Moreover, this solution satisfies

$$\|\mathbf{u}\|_{L^2(\Omega)^3} + \|p\|_{H^1(\Omega)} \leq C \left( \|\mathbf{f}\|_{L^2(\Omega)^3} + \|g\|_{H_{00}^{1/2}(\Gamma)'} \right). \tag{2.6}$$

A sufficient condition for the global uniqueness of the solution is:

**Proposition 2.4.** Assume that the function  $\alpha$  satisfies assumptions (2.1) and (2.2). If problem (2.3) has a solution  $(\mathbf{u}, p)$  such that  $\mathbf{u} \in L^3(\Omega)^3$  and satisfies

$$\frac{\alpha_{\max} + \alpha_{\min}}{\alpha_{\min}} C(\Omega) L_{\alpha} \|\mathbf{u}\|_{L^3(\Omega)^3} < 1, \tag{2.7}$$

then, there is no other solution of problem (2.3).

This uniqueness result is stated in [3] under the condition that  $\mathbf{u}$  belongs to  $L^r(\Omega)^d$  with  $r > d$ , where  $d$  is the dimension. This is due to the Sobolev imbedding when  $d = 2$  (see Rem. 2.8). However, when  $d = 3$ , the proof in [3] is also valid with  $r = 3$ .

Finally, concerning the regularity of the solution the following result holds.

**Proposition 2.5.** *There exists a real number  $\rho_0 > 2$  only depending on the geometry of  $\Omega$  such that, for all  $\rho$ ,  $2 < \rho \leq \rho_0$ , and for all data  $(\mathbf{f}, g) \in L^\rho(\Omega)^3 \times W_\rho^{-1/\rho}(\Gamma)$ , any solution  $(\mathbf{u}, p)$  of problem (2.3) belongs to  $L^\rho(\Omega)^3 \times W_\rho^1(\Omega)$ .*

**Remark 2.6.** The existence of  $\rho_0$  is obtained in [3] by a perturbation argument, but in dimension  $d = 3$ , there is no guarantee that  $\rho_0 \geq 3$ . Therefore, in general, condition (2.7) for global uniqueness cannot be checked from the data.

### 2.2. Nonsingular solutions

Let us now consider the case when the solution is only locally unique. In this case, although problem (2.3) may have more than one solution, we assume that there exists an *isolated* solution. That is, there exists a neighborhood of this solution where no other solution exists. A sufficient condition for this to hold is that the solution is *nonsingular* (see [11] or Girault and Raviart [20]). In this paragraph we analyze the properties of nonsingular solutions, and give sufficient conditions for such a solution to exist.

First we cast problem (2.3) in a more convenient, but nevertheless equivalent, functional setting. With this purpose let us define the data space

$$\mathfrak{Y} := L^2(\Omega)^3 \times H_{00}^{1/2}(\Gamma)',$$

with norm

$$\|(\mathbf{f}, g)\|_{\mathfrak{Y}} := \|\mathbf{f}\|_{L^2(\Omega)^3} + \|g\|_{H_{00}^{1/2}(\Gamma)'},$$

and the solution space

$$\mathfrak{X} := L^2(\Omega)^3 \times H_w^1(\Omega),$$

with norm

$$\|(\mathbf{u}, p)\|_{\mathfrak{X}} := \|\mathbf{u}\|_{L^2(\Omega)^3} + \|p\|_{H^1(\Omega)}.$$

We also define  $T$  as the solution operator to the linear Darcy problem, *i.e.*  $T : \mathfrak{Y} \rightarrow \mathfrak{X}$  is such that, for every  $\eta = (\mathbf{f}, g) \in \mathfrak{Y}$ ,  $\mathfrak{X} \ni x = (\mathbf{u}, p) = T\eta = T(\mathbf{f}, g)$  solves

$$\begin{cases} \bar{\alpha}\mathbf{u} + \nabla p = \mathbf{f}, & \text{in } \Omega, \\ \nabla \cdot \mathbf{u} = 0, & \text{in } \Omega, \\ p = 0, & \text{on } \Gamma_w, \\ \mathbf{u} \cdot \mathbf{n} = g, & \text{on } \Gamma, \end{cases} \tag{2.8}$$

for a fixed  $\bar{\alpha} > 0$ .

It is classical that problem (2.8) is well-posed. This implies that  $T \in \mathcal{L}(\mathfrak{Y}, \mathfrak{X})$ . In other words, there is a constant  $C > 0$  such that for every  $(\mathbf{f}, g) \in \mathfrak{Y}$

$$\|T(\mathbf{f}, g)\|_{\mathfrak{X}} \leq C \|(\mathbf{f}, g)\|_{\mathfrak{Y}}. \tag{2.9}$$

By assumption (2.1) we get that  $\alpha \in L^\infty(\mathbb{R})$ . Then, for any  $(\mathbf{u}, p) \in \mathfrak{X}$ ,  $\alpha(p)\mathbf{u} \in L^2(\Omega)^3$  and we can define the map  $G : \mathfrak{X} \rightarrow \mathfrak{Y}$  as follows. If  $x = (\mathbf{u}, p) \in \mathfrak{X}$ , then

$$G(x) := \begin{pmatrix} (\alpha(p) - \bar{\alpha})\mathbf{u} - \mathbf{f} \\ -g \end{pmatrix} \in \mathfrak{Y}.$$

Finally, let us define  $F : \mathfrak{X} \rightarrow \mathfrak{X}$  as

$$F(x) := x + TG(x).$$

With this notation, problem (2.3) can be equivalently restated as:

Find  $x = (\mathbf{u}, p) \in \mathfrak{X}$  such that

$$F(x) = 0. \tag{2.10}$$

We are now in a position to define the notion of nonsingular solutions.

**Definition 2.7** ([11]). Let  $x \in \mathfrak{X}$  solve problem (2.10). This solution is called *nonsingular* if the linear operator

$$F'(x) = I + TG'(x)$$

is an isomorphism of  $\mathfrak{X}$ . Here  $F'(x)$  and  $G'(x)$  denote the Fréchet derivative of the maps  $F$  and  $G$  at point  $x$ , respectively.

Our main interest in this paragraph is to provide sufficient conditions for a solution to be nonsingular in this sense. First of all, by assumption (2.2) we know that the derivative of  $\alpha$  exists a.e. on  $\mathbb{R}$  (cf. Folland [18]). Denoting this derivative by  $\dot{\alpha}$  we can, formally, obtain the derivative of the map  $G$ . Let  $x = (\mathbf{u}, p), y = (\mathbf{v}, q) \in \mathfrak{X}$ , then

$$G'(x)y = \begin{pmatrix} (\alpha(p) - \bar{\alpha}) \mathbf{v} + \dot{\alpha}(p)q\mathbf{u} \\ 0 \end{pmatrix}. \tag{2.11}$$

From this we can conclude that if  $x = (\mathbf{u}, p) \in L^3(\Omega)^3 \times H_w^1(\Omega) \subset \mathfrak{X}$ , the Fréchet derivative of the map  $G$  is well-defined, given by equation (2.11), and  $G'(x) \in \mathcal{L}(\mathfrak{X}, \mathfrak{Y})$ .

**Remark 2.8.** We need  $\mathbf{u} \in L^3(\Omega)^3$  because of the term  $\dot{\alpha}(p)q\mathbf{u}$ . Indeed, by assumption (2.2), Hölder's inequality and the Sobolev embedding  $H^1 \hookrightarrow L^6$ , we have

$$\begin{aligned} \int_{\Omega} |\dot{\alpha}(p)q\mathbf{u}|^2 &\leq L_{\alpha}^2 \left( \int_{\Omega} q^6 \right)^{1/3} \left( \int_{\Omega} |\mathbf{u}|^3 \right)^{2/3} \\ &\leq C(\Omega)^2 L_{\alpha}^2 \|q\|_{H^1(\Omega)}^2 \|\mathbf{u}\|_{L^3(\Omega)}^2, \end{aligned}$$

where all inequalities are sharp. Clearly, if  $d = 2$  we should require  $\mathbf{u} \in L^{2+\epsilon}(\Omega)^2$  for some  $\epsilon > 0$ . In both cases, we must assume that the velocity  $\mathbf{u}$  lies in a smaller space than  $L^2(\Omega)^d$  for the derivative to make sense. This is in contrast to the common feature of many nonlinear operators arising in the analysis of partial differential equations that describe physical phenomena. For such an operator, its derivative is everywhere defined and the range of the derivative is a smaller space (i.e. more smooth or regular) than the data space. For this reason, we say that the operator  $G$  does not have *regularizing properties*. The fact that for problem (1.1) the nonlinearity  $G$  does not have regularizing properties lies at the heart of all the difficulties that its theoretical and numerical analysis present.

We now give sufficient conditions for a solution of problem (2.10) to be nonsingular in the sense of Definition 2.7.

**Proposition 2.9.** Assume that for problem (2.10) the function  $\alpha$  is such that conditions (2.1) and (2.2) hold. Let  $x = (\mathbf{u}, p) \in \mathfrak{X}$  be a solution to problem (2.10). If  $\mathbf{u} \in L^3(\Omega)^3$  and

$$\frac{\alpha_{\max} + \alpha_{\min}}{\alpha_{\min}} C(\Omega) L_{\alpha} \|\mathbf{u}\|_{L^3(\Omega)^3} < 1, \tag{2.12}$$

then this solution is nonsingular.

*Proof.* We need to show that the map  $I + TG'(x)$  is an isomorphism of  $\mathfrak{X}$ . Since the operator is continuous, by the Open Mapping Theorem (see Helemskii [25]) it is sufficient to show that the operator is bijective. That is, given any  $z = (\mathbf{w}, r) \in \mathfrak{X}$  there exists a unique  $y = (\mathbf{v}, q) \in \mathfrak{X}$  such that

$$y + TG'(x)y = z,$$

or

$$(y - z) = T(-G'(x))y.$$

In other words, we must prove that the problem: Find  $(\mathbf{v}, q) \in \mathfrak{X}$  such that

$$\begin{cases} \bar{\alpha}(\mathbf{v} - \mathbf{w}) + \nabla(q - r) = (\bar{\alpha} - \alpha(p))\mathbf{v} - \dot{\alpha}(p)q\mathbf{u}, & \text{in } \Omega, \\ \nabla \cdot (\mathbf{v} - \mathbf{w}) = 0, & \text{in } \Omega, \\ (\mathbf{v} - \mathbf{w}) \cdot \mathbf{n} = 0, & \text{on } \Gamma, \\ q - r = 0, & \text{on } \Gamma_w, \end{cases}$$

always has a unique solution. Doing the elementary change of variables  $(\mathbf{V}, Q) = (\mathbf{v} - \mathbf{w}, q - r) \in \mathfrak{X}$  this problem can be equivalently restated as: Find  $(\mathbf{V}, Q) \in \mathfrak{X}$  such that

$$\begin{cases} \alpha(p)\mathbf{V} + \nabla Q = \mathbf{F}(Q), & \text{in } \Omega, \\ \nabla \cdot \mathbf{V} = 0, & \text{in } \Omega, \\ \mathbf{V} \cdot \mathbf{n} = 0, & \text{on } \Gamma, \\ Q = 0, & \text{on } \Gamma_w, \end{cases}$$

where

$$\mathbf{F}(Q) := (\bar{\alpha} - \alpha(p))\mathbf{w} - \dot{\alpha}(p)r\mathbf{u} - \dot{\alpha}(p)Q\mathbf{u} = \mathbf{F} + \bar{\mathbf{F}}(Q),$$

with

$$\mathbf{F} = (\bar{\alpha} - \alpha(p))\mathbf{w} - \dot{\alpha}(p)r\mathbf{u}, \quad \bar{\mathbf{F}}(Q) = \dot{\alpha}(p)Q\mathbf{u}.$$

Notice that, since  $\mathbf{u} \in L^3(\Omega)^3$  then  $\mathbf{F}(Q) \in L^2(\Omega)^3$ . This problem can be written in variational form as: Find  $(\mathbf{V}, Q) \in \mathfrak{X}$  such that

$$\begin{cases} \int_{\Omega} \alpha(p)\mathbf{V} \cdot \mathbf{W} + \int_{\Omega} \mathbf{W} \cdot \nabla Q = \int_{\Omega} \mathbf{F}(Q) \cdot \mathbf{W}, & \forall \mathbf{W} \in L^2(\Omega)^3, \\ \int_{\Omega} \mathbf{V} \cdot \nabla R = 0, & \forall R \in H_w^1(\Omega). \end{cases} \tag{2.13}$$

We observe that (2.13) is a linear Darcy's system with an affine perturbation  $\mathbf{F}(Q)$ . If we define the bilinear form  $\mathcal{A} : \mathfrak{X} \times \mathfrak{X} \rightarrow \mathbb{R}$  by

$$\mathcal{A}((\mathbf{V}, Q), (\mathbf{W}, R)) := \int_{\Omega} \alpha(p)\mathbf{V} \cdot \mathbf{W} + \int_{\Omega} \mathbf{W} \cdot \nabla Q + \int_{\Omega} \mathbf{V} \cdot \nabla R,$$

and assume for the moment that  $\bar{\mathbf{F}}(Q) = \mathbf{0}$ , *i.e.*  $\mathbf{F}(Q)$  does not depend on  $Q$ , then, problem (2.13) has a unique solution if and only if:

- (1) There exists a constant  $\beta_{\mathcal{A}} > 0$  such that

$$\inf_{(\mathbf{V}, Q) \in \mathfrak{X}} \sup_{(\mathbf{W}, R) \in \mathfrak{X}} \frac{\mathcal{A}((\mathbf{V}, Q), (\mathbf{W}, R))}{\|(\mathbf{V}, Q)\|_{\mathfrak{X}} \|(\mathbf{W}, R)\|_{\mathfrak{X}}} \geq \beta_{\mathcal{A}}. \tag{2.14}$$

- (2) The form  $\mathcal{A}$  has the following property:

$$(\mathcal{A}((\mathbf{V}, Q), (\mathbf{W}, R)) = 0 \quad \forall (\mathbf{V}, Q) \in \mathfrak{X}) \Rightarrow (\mathbf{W}, R) = (\mathbf{0}, 0). \tag{2.15}$$



These two properties are equivalent to the fact that the linear Darcy problem defined by the form  $\mathcal{A}$  is well-posed, which is a classical result. This also implies the *a priori* estimate

$$\|\mathbf{V}\|_{L^2(\Omega)^3} + |Q|_{H^1(\Omega)} \leq C \|\mathbf{F}\|_{L^2(\Omega)^3}, \tag{2.16}$$

for some  $C > 0$  that does not depend on  $\mathbf{F}$ ,  $\mathbf{V}$  or  $Q$ . Now, the well-posedness of (2.13) follows immediately by proving that the affine mapping  $S \mapsto Q$ , where  $Q$  is the second component of the solution pair  $(\mathbf{V}, Q)$  of (2.13) with data  $\mathbf{F}(S)$  is a contraction: There exists  $k \in ]0, 1[$  such that

$$\forall S \in H^1(\Omega), |Q|_{H^1(\Omega)} \leq k |S|_{H^1(\Omega)}.$$

To do this, let  $S$  be given in  $H^1(\Omega)$ , set  $\mathbf{F} = \mathbf{0}$ , and take  $\mathbf{W} = \mathbf{V}$  in the first equation of problem (2.13). The second equation, together with condition (2.1) imply

$$\begin{aligned} \alpha_{\min} \|\mathbf{V}\|_{L^2(\Omega)^3}^2 &\leq \int_{\Omega} \alpha(p) \mathbf{V} \cdot \mathbf{V} \\ &= \int_{\Omega} \bar{\mathbf{F}}(S) \cdot \mathbf{V} \\ &\leq \|\bar{\mathbf{F}}(S)\|_{L^2(\Omega)^3} \|\mathbf{V}\|_{L^2(\Omega)^3}, \end{aligned}$$

or

$$\|\mathbf{V}\|_{L^2(\Omega)^3} \leq \frac{1}{\alpha_{\min}} \|\bar{\mathbf{F}}(S)\|_{L^2(\Omega)^3}.$$

By taking  $\mathbf{W} = \nabla Q$  we obtain

$$\begin{aligned} |Q|_{H^1(\Omega)}^2 &= \int_{\Omega} \nabla Q \cdot \nabla Q \\ &= \int_{\Omega} \bar{\mathbf{F}}(S) \cdot \nabla Q - \int_{\Omega} \alpha(p) \mathbf{V} \cdot \nabla Q \\ &\leq \|\bar{\mathbf{F}}(S)\|_{L^2(\Omega)^3} |Q|_{H^1(\Omega)} + \alpha_{\max} \|\mathbf{V}\|_{L^2(\Omega)^3} |Q|_{H^1(\Omega)} \\ &\leq \left(1 + \frac{\alpha_{\max}}{\alpha_{\min}}\right) \|\bar{\mathbf{F}}(S)\|_{L^2(\Omega)^3} |Q|_{H^1(\Omega)} \\ &\leq \left(1 + \frac{\alpha_{\max}}{\alpha_{\min}}\right) \|\dot{\alpha}(p) S \mathbf{u}\|_{L^2(\Omega)^3} |Q|_{H^1(\Omega)}. \end{aligned}$$

Since

$$\|\dot{\alpha}(p) S \mathbf{u}\|_{L^2(\Omega)^3} \leq C(\Omega) L_{\alpha} \|\mathbf{u}\|_{L^3(\Omega)^3} |S|_{H^1(\Omega)},$$

we derive

$$|Q|_{H^1(\Omega)} \leq \left(1 + \frac{\alpha_{\max}}{\alpha_{\min}}\right) C(\Omega) L_{\alpha} \|\mathbf{u}\|_{L^3(\Omega)^3} |S|_{H^1(\Omega)}.$$

Therefore the mapping  $S \mapsto Q$  is a contraction if

$$\left(1 + \frac{\alpha_{\max}}{\alpha_{\min}}\right) C(\Omega) L_{\alpha} \|\mathbf{u}\|_{L^3(\Omega)^3} < 1,$$

which is condition (2.12). □

**Remark 2.10.** We see that (2.12) coincides with the condition for global uniqueness (2.7). This reflects that the nonlinearity  $G$  does not have regularizing properties. Nevertheless, these are only sufficient conditions, and it is plausible that problem (1.1) has a nonsingular solution without satisfying condition (2.12).

### 3. THE DISCRETE PROBLEM

Having analyzed the mathematical properties of problem (1.1) we now proceed to propose several methods for its approximate solution. With this purpose, let  $h$  be a discretization parameter (that will tend to zero). For every  $h > 0$  we introduce two finite dimensional spaces  $X_h \subset L^2(\Omega)^3$  and  $M_h \subset H^1_w(\Omega)$  such that:

- (1) The pair of spaces  $(X_h, M_h)$  is stable, in the sense that they satisfy a uniform inf-sup condition ([10,20], Ern and Guermond [17] or Boffi *et al.* [7]). That is, there exists a constant  $\beta > 0$  independent of  $h$  such that

$$\sup_{\mathbf{w}_h \in X_h} \frac{b(\mathbf{w}_h, q_h)}{\|\mathbf{w}_h\|_{L^2(\Omega)^3}} \geq \beta |q_h|_{H^1(\Omega)}, \quad \forall q_h \in M_h, \tag{3.1}$$

where the form  $b$  is defined in (2.5).

- (2) There exist continuous interpolation operators  $\pi_h : L^2(\Omega)^3 \rightarrow X_h$ ,  $\mathcal{I}_h : H^1(\Omega) \rightarrow M_h$  and an integer  $k \geq 1$ , such that for all  $(\mathbf{v}, q) \in H^k(\Omega) \times H^{k+1}(\Omega)$

$$\|\mathbf{v} - \pi_h \mathbf{v}\|_{L^2(\Omega)^3} \leq Ch^k \|\mathbf{v}\|_{H^k(\Omega)^3}, \tag{3.2}$$

and

$$|q - \mathcal{I}_h q|_{H^1(\Omega)} \leq Ch^k |q|_{H^{k+1}(\Omega)}. \tag{3.3}$$

In order to find examples of such discrete spaces, assume to simplify that  $\Omega$  is a polyhedron, and let  $\mathcal{T}_h$  be a family of triangulations of  $\bar{\Omega}$ , made of tetrahedra with diameter bounded by  $h$ . We suppose that  $\mathcal{T}_h$  is regular in the following sense (*cf.* Ciarlet [12]): There exists a constant  $\sigma > 0$ , independent of  $h$ , such that

$$\forall T \in \mathcal{T}_h, \frac{h_T}{\rho_T} \leq \sigma, \tag{3.4}$$

where  $h_T$  is the diameter of  $T$  and  $\rho_T$  is the diameter of the ball inscribed in  $T$ . Then, for any integer  $k \geq 1$ , the following pair of spaces satisfy conditions (3.1)–(3.3):

$$X_h := \{ \mathbf{v}_h \in L^2(\Omega)^3 : \mathbf{v}_h|_T \in \mathbb{P}_{k-1}^3, \forall T \in \mathcal{T}_h \}, \tag{3.5}$$

and

$$M_h := \{ q_h \in C^0(\bar{\Omega}) : q_h|_T \in \mathbb{P}_k, \forall T \in \mathcal{T}_h \}. \tag{3.6}$$

For a proof the reader can consult standard references, for instance [10,17,20].

Finally, we define the discrete solution space

$$\mathfrak{X}_h := X_h \times M_h,$$

normed by  $\|\cdot\|_{\mathfrak{X}}$ . Clearly,  $\mathfrak{X}_h \subset \mathfrak{X}$ . For the sequel, it is also useful to introduce the space

$$V_h = \{ \mathbf{v}_h \in X_h : \forall q_h \in M_h, b(\mathbf{v}_h, q_h) = 0 \}, \tag{3.7}$$

and its orthogonal in  $X_h$

$$V_h^\perp = \{ \mathbf{v}_h \in X_h : \forall \mathbf{w}_h \in V_h, \int_\Omega \mathbf{v}_h \cdot \mathbf{w}_h = 0 \}. \tag{3.8}$$

For each such pair of discrete spaces we define the Galerkin solution to problem (2.3) as the pair  $x_h = (\mathbf{u}_h, p_h) \in \mathfrak{X}_h$  such that

$$\begin{cases} a_{p_h}(\mathbf{u}_h, \mathbf{v}_h) + b(\mathbf{v}_h, p_h) = \int_\Omega \mathbf{f} \cdot \mathbf{v}_h, & \forall \mathbf{v}_h \in X_h, \\ b(\mathbf{u}_h, q_h) = \langle g, q_h \rangle_\Gamma, & \forall q_h \in M_h. \end{cases} \tag{3.9}$$

Under assumptions (2.1) and (3.1), the existence of a solution for this problem can be established by the same techniques used in Theorem 2.3 (cf. [3]). It is even simpler, since problem (3.9) is already set in finite dimension. All solutions of problem (3.9) satisfy uniform *a priori* estimates and (3.2) and (3.3) suffice to establish weak convergence (up to subsequences) of any solution of (3.9) to some solution of (2.3).

In the remainder of this section we analyze this discrete problem. For the case when the solution is unique we prove optimal error estimates and propose an algorithm to find such an approximate solution. The algorithm is proved to converge independently of the discretization parameter. For the nonuniqueness case, in the spirit of [11,20], we show that for  $h$  small enough there exists a nonsingular solution to (3.9) in a neighborhood of the nonsingular solution to the exact problem. We analyze some properties of the application of Newton's method to this problem, and we obtain estimates on its speed of convergence and conditions on the initial approximation. The main difficulty in this analysis is that there exist  $x$  in  $\mathfrak{X}$  for which the operator  $G'(x)$  is not bounded in  $\mathcal{L}(\mathfrak{X}, \mathfrak{Y})$ . More precisely, we require that the first component of  $x$  belong to  $L^3(\Omega)^3$ , a smaller space than  $L^2(\Omega)^3$ . This again is related to the fact that the nonlinearity  $G$  does not have regularizing properties.

### 3.1. The uniqueness case

Recall that condition (2.7) is sufficient for the solution to problem (2.3) to be unique. In the setting that we have described, and under a similar assumption, we have the following *a priori* estimate.

**Theorem 3.1.** *Let the pair of finite dimensional spaces  $\mathfrak{X}_h$  satisfy condition (3.1). Assume that the solution  $x = (\mathbf{u}, p) \in \mathfrak{X}$  to (2.3) is such that  $\mathbf{u} \in L^3(\Omega)^3$  and is small enough, in the sense that*

$$\frac{1}{\beta} \frac{\alpha_{\max} + \alpha_{\min}}{\alpha_{\min}} C(\Omega) L_\alpha \|\mathbf{u}\|_{L^3(\Omega)^3} \leq \theta < 1. \tag{3.10}$$

Then both (2.3) and (3.9) have a unique solution and there exists a constant  $C > 0$  independent of  $h$  such that the solution  $x_h = (\mathbf{u}_h, p_h) \in \mathfrak{X}_h$  of problem (3.9) satisfies

$$\|\mathbf{u} - \mathbf{u}_h\|_{L^2(\Omega)^3} + |p - p_h|_{H^1(\Omega)} \leq C \left( \inf_{\mathbf{v}_h \in X_h} \|\mathbf{u} - \mathbf{v}_h\|_{L^2(\Omega)^3} + \inf_{q_h \in M_h} |p - q_h|_{H^1(\Omega)} \right). \tag{3.11}$$

*Proof.* The proof proceeds in three steps.

1) The second equation in (3.9) can be viewed as a non-homogeneous constraint; let us show that we can approximate  $\mathbf{u}$  with functions of  $X_h$  that satisfy this constraint. For this, let  $\mathbf{v}_h$  be an arbitrary function of  $X_h$ , define  $\mathbf{r}_h$  in  $X_h$  by

$$\forall q_h \in M_h, b(\mathbf{r}_h, q_h) = b(\mathbf{u} - \mathbf{v}_h, q_h),$$

and set  $\mathbf{w}_h := \mathbf{r}_h + \mathbf{v}_h$ . It follows from (3.1) and the Babuška–Brezzi's theory (cf. [4,9] or [10,17,20]) that this equation has a solution  $\mathbf{r}_h \in X_h$ , unique in  $V_h^\perp$ , and such that

$$\beta \|\mathbf{r}_h\|_{L^2(\Omega)^3} \leq \|\mathbf{u} - \mathbf{v}_h\|_{L^2(\Omega)^3}. \tag{3.12}$$

Thus

$$b(\mathbf{w}_h, q_h) = b(\mathbf{u}, q_h) = \langle g, q_h \rangle = b(\mathbf{u}_h, q_h), \quad \forall q_h \in M_h,$$

and  $\mathbf{u}_h - \mathbf{w}_h \in V_h$ . This implies

$$\begin{aligned} \alpha_{\min} \|\mathbf{u}_h - \mathbf{w}_h\|_{L^2(\Omega)^3} &\leq \sup_{\mathbf{y}_h \in V_h} \frac{a_{p_h}(\mathbf{u}_h - \mathbf{w}_h, \mathbf{y}_h)}{\|\mathbf{y}_h\|_{L^2(\Omega)^3}} \\ &\leq \sup_{\mathbf{y}_h \in V_h} \frac{a_{p_h}(\mathbf{u}_h - \mathbf{u}, \mathbf{y}_h)}{\|\mathbf{y}_h\|_{L^2(\Omega)^3}} + \sup_{\mathbf{y}_h \in V_h} \frac{a_{p_h}(\mathbf{u} - \mathbf{w}_h, \mathbf{y}_h)}{\|\mathbf{y}_h\|_{L^2(\Omega)^3}} \\ &\leq \sup_{\mathbf{y}_h \in V_h} \frac{a_{p_h}(\mathbf{u}_h - \mathbf{u}, \mathbf{y}_h)}{\|\mathbf{y}_h\|_{L^2(\Omega)^3}} + \alpha_{\max} \|\mathbf{u} - \mathbf{w}_h\|_{L^2(\Omega)^3}. \end{aligned}$$

2) Subtract the first equation of (2.3) from the first equation in (3.9) with test function  $\mathbf{y}_h \in V_h$ . Since  $X_h \subset L^2(\Omega)^3$ ,

$$\begin{aligned} a_{p_h}(\mathbf{u}_h - \mathbf{u}, \mathbf{y}_h) &= \int_{\Omega} (\alpha(p) - \alpha(p_h)) \mathbf{u} \cdot \mathbf{y}_h + \int_{\Omega} \mathbf{y}_h \cdot \nabla(p - p_h) \\ &\leq L_{\alpha} \|p - p_h\|_{L^6(\Omega)} \|\mathbf{u}\|_{L^3(\Omega)^3} \|\mathbf{y}_h\|_{L^2(\Omega)^3} + b(\mathbf{y}_h, p - p_h) \\ &\leq C(\Omega) L_{\alpha} |p - p_h|_{H^1(\Omega)} \|\mathbf{u}\|_{L^3(\Omega)^3} \|\mathbf{y}_h\|_{L^2(\Omega)^3} + b(\mathbf{y}_h, p - q_h) + b(\mathbf{y}_h, q_h - p_h). \end{aligned}$$

This yields

$$\alpha_{\min} \|\mathbf{u}_h - \mathbf{w}_h\|_{L^2(\Omega)^3} \leq C(\Omega) L_{\alpha} |p - p_h|_{H^1(\Omega)} \|\mathbf{u}\|_{L^3(\Omega)^3} + |p - q_h|_{H^1(\Omega)} + \alpha_{\max} \|\mathbf{u} - \mathbf{w}_h\|_{L^2(\Omega)^3},$$

where the last inequality holds since  $\mathbf{y}_h \in V_h$ . Finally, by the triangle inequality and (3.12)

$$\begin{aligned} \alpha_{\min} \|\mathbf{u} - \mathbf{u}_h\|_{L^2(\Omega)^3} &\leq (\alpha_{\min} + \alpha_{\max}) \left(1 + \frac{1}{\beta}\right) \|\mathbf{u} - \mathbf{v}_h\|_{L^2(\Omega)^3} \\ &\quad + C(\Omega) L_{\alpha} |p - p_h|_{H^1(\Omega)} \|\mathbf{u}\|_{L^3(\Omega)^3} + |p - q_h|_{H^1(\Omega)}. \end{aligned} \quad (3.13)$$

3) Let  $q_h \in M_h$  be arbitrary. By the inf-sup condition (3.1),

$$\begin{aligned} \beta |p_h - q_h|_{H^1(\Omega)} &\leq \sup_{\mathbf{y}_h \in X_h} \frac{b(\mathbf{y}_h, p_h - q_h)}{\|\mathbf{y}_h\|_{L^2(\Omega)^3}} \\ &\leq \sup_{\mathbf{y}_h \in X_h} \frac{b(\mathbf{y}_h, p_h - p)}{\|\mathbf{y}_h\|_{L^2(\Omega)^3}} + \sup_{\mathbf{y}_h \in X_h} \frac{b(\mathbf{y}_h, p - q_h)}{\|\mathbf{y}_h\|_{L^2(\Omega)^3}} \\ &\leq \sup_{\mathbf{y}_h \in X_h} \frac{b(\mathbf{y}_h, p_h - p)}{\|\mathbf{y}_h\|_{L^2(\Omega)^3}} + |p - q_h|_{H^1(\Omega)}. \end{aligned}$$

Subtracting the first equation of (2.3) from the first equation of (3.9), since  $X_h \subset L^2(\Omega)^3$  we obtain

$$\begin{aligned} b(\mathbf{y}_h, p_h - p) &= \int_{\Omega} (\alpha(p) - \alpha(p_h)) \mathbf{u} \cdot \mathbf{y}_h + \int_{\Omega} \alpha(p_h) (\mathbf{u} - \mathbf{u}_h) \cdot \mathbf{y}_h \\ &\leq C(\Omega) L_{\alpha} |p - p_h|_{H^1(\Omega)} \|\mathbf{u}\|_{L^3(\Omega)^3} \|\mathbf{y}_h\|_{L^2(\Omega)^3} + \alpha_{\max} \|\mathbf{u} - \mathbf{u}_h\|_{L^2(\Omega)^3} \|\mathbf{y}_h\|_{L^2(\Omega)^3}, \end{aligned}$$

which implies

$$|p_h - q_h|_{H^1(\Omega)} \leq \frac{1}{\beta} |p - q_h|_{H^1(\Omega)} + \frac{C(\Omega) L_{\alpha}}{\beta} \|\mathbf{u}\|_{L^3(\Omega)^3} |p - p_h|_{H^1(\Omega)} + \frac{\alpha_{\max}}{\beta} \|\mathbf{u} - \mathbf{u}_h\|_{L^2(\Omega)^3}.$$

By the triangle inequality

$$|p - p_h|_{H^1(\Omega)} \leq \left(1 + \frac{1}{\beta}\right) |p - q_h|_{H^1(\Omega)} + \frac{C(\Omega) L_{\alpha}}{\beta} \|\mathbf{u}\|_{L^3(\Omega)^3} |p - p_h|_{H^1(\Omega)} + \frac{\alpha_{\max}}{\beta} \|\mathbf{u} - \mathbf{u}_h\|_{L^2(\Omega)^3}.$$

Assumption (3.10) implies

$$\frac{\alpha_{\max} + \alpha_{\min}(1 - \theta)}{\alpha_{\max} + \alpha_{\min}} |p - p_h|_{H^1(\Omega)} \leq \left(1 + \frac{1}{\beta}\right) |p - q_h|_{H^1(\Omega)} + \frac{\alpha_{\max}}{\beta} \|\mathbf{u} - \mathbf{u}_h\|_{L^2(\Omega)^3}.$$

Combining this last inequality, assumption (3.10), and (3.13) we obtain

$$\|\mathbf{u} - \mathbf{u}_h\|_{L^2(\Omega)^3} \leq C \left( \|\mathbf{u} - \mathbf{v}_h\|_{L^2(\Omega)^3} + |p - q_h|_{H^1(\Omega)} \right) + \frac{\alpha_{\max}\theta}{\alpha_{\max} + \alpha_{\min}(1 - \theta)} \|\mathbf{u} - \mathbf{u}_h\|_{L^2(\Omega)^3}.$$

Since

$$1 - \frac{\alpha_{\max}\theta}{\alpha_{\max} + \alpha_{\min}(1 - \theta)} = \frac{\alpha_{\max} + \alpha_{\min}}{\alpha_{\max} + \alpha_{\min}(1 - \theta)}(1 - \theta) > 0$$

and the pair  $(\mathbf{v}_h, q_h) \in \mathfrak{X}_h$  is arbitrary we obtain the desired result. □

**Remark 3.2.** For the pair of finite element spaces (3.5), (3.6) condition (3.1) holds with  $\beta = 1$ . Hence, in this case, assumption (3.10) is the same as (2.7).

The next corollary follows readily from this theorem.

**Corollary 3.3.** *Under the setting of Theorem 3.1, if the spaces  $X_h$  and  $M_h$  satisfy assumptions (3.2) and (3.3), then*

$$\lim_{h \rightarrow 0} \left( \|\mathbf{u} - \mathbf{u}_h\|_{L^2(\Omega)^3} + |p - p_h|_{H^1(\Omega)} \right) = 0.$$

Moreover, if the exact solution  $(\mathbf{u}, p) \in H^s(\Omega)^3 \times H^{s+1}(\Omega)$  for some real number  $s \in [0, k]$ , then there is a constant  $C > 0$  independent of  $h$  such that

$$\|\mathbf{u} - \mathbf{u}_h\|_{L^2(\Omega)^3} + |p - p_h|_{H^1(\Omega)} \leq Ch^s \left( \|\mathbf{u}\|_{H^s(\Omega)^3} + |q|_{H^{s+1}(\Omega)} \right).$$

*Proof.* The conclusion of Theorem 3.1, an elementary density argument and assumptions (3.2) and (3.3) give that the Galerkin solution converges to the exact solution as  $h \rightarrow 0$ . If the exact solution is more regular, assumptions (3.2) and (3.3) give the claimed error estimates. □

We now propose an iterative scheme to solve the discrete nonlinear system (3.9). Although the scheme requires assembling a new matrix at each iterative step, we show that, under an assumption similar to (2.7), the speed of convergence to the Galerkin solution is independent of the discretization parameter  $h$ .

The proposed scheme is the following:

Given an arbitrary initial approximation  $p_h^{(0)} \in M_h$ , for  $n = 0, 1, 2, \dots$  find  $(\mathbf{u}_h^{(n+1)}, p_h^{(n+1)}) \in \mathfrak{X}_h$  that solve

$$\begin{cases} a_{p_h^{(n)}}(\mathbf{u}_h^{(n+1)}, \mathbf{v}_h) + b(\mathbf{v}_h, p_h^{(n+1)}) = \int_{\Omega} \mathbf{f} \cdot \mathbf{v}_h, & \forall \mathbf{v}_h \in X_h, \\ b(\mathbf{u}_h^{(n+1)}, q_h) = \langle g, q_h \rangle_{\Gamma}, & \forall q_h \in M_h. \end{cases} \tag{3.14}$$

Now we prove that this scheme converges independently of the discretization parameter.

**Proposition 3.4.** *Assume that the pair of spaces  $(X_h, M_h)$  satisfies condition (3.1). Let the solution to (3.9) be small enough, in the sense that there are two constants  $\theta < 1$  and  $h_0 > 0$  such that for every  $h \leq h_0$*

$$\frac{\alpha_{\max} + \alpha_{\min}}{\alpha_{\min}} C(\Omega) L_{\alpha} \|\mathbf{u}_h\|_{L^3(\Omega)^3} \leq \theta. \tag{3.15}$$

Then for the iterative scheme (3.14) the following error estimates hold

$$\left\| \mathbf{u}_h - \mathbf{u}_h^{(n+1)} \right\|_{L^2(\Omega)^3} \leq \frac{1}{\alpha_{\max} + \alpha_{\min}} \frac{\theta^{n+1}}{\beta^n} \left| p_h - p_h^{(0)} \right|_{H^1(\Omega)},$$

and

$$\left| p_h - p_h^{(n+1)} \right|_{H^1(\Omega)} \leq \left( \frac{\theta}{\beta} \right)^{n+1} \left| p_h - p_h^{(0)} \right|_{H^1(\Omega)}.$$

*Proof.* Take the difference of equations (3.9) and (3.14), we obtain

$$\begin{cases} \int_{\Omega} \left( \alpha(p_h) \mathbf{u}_h - \alpha(p_h^{(n)}) \mathbf{u}_h^{(n+1)} \right) \cdot \mathbf{v}_h + b \left( \mathbf{v}_h, p_h - p_h^{(n+1)} \right) = 0, & \forall \mathbf{v}_h \in X_h, \\ b \left( \mathbf{u}_h - \mathbf{u}_h^{(n+1)}, q_h \right) = 0, & \forall q_h \in M_h. \end{cases}$$

Set  $\mathbf{v}_h = \mathbf{u}_h - \mathbf{u}_h^{(n+1)}$ , then

$$\begin{aligned} \alpha_{\min} \left\| \mathbf{u}_h - \mathbf{u}_h^{(n+1)} \right\|_{L^2(\Omega)^3}^2 &\leq \left| \int_{\Omega} \left( \alpha(p_h^{(n)}) - \alpha(p_h) \right) \mathbf{u}_h \cdot \left( \mathbf{u}_h - \mathbf{u}_h^{(n+1)} \right) \right| \\ &\leq C(\Omega) L_{\alpha} \left| p_h - p_h^{(n)} \right|_{H^1(\Omega)} \left\| \mathbf{u}_h \right\|_{L^3(\Omega)^3} \left\| \mathbf{u}_h - \mathbf{u}_h^{(n+1)} \right\|_{L^2(\Omega)^3}, \end{aligned}$$

which by (3.15) implies

$$\left\| \mathbf{u}_h - \mathbf{u}_h^{(n+1)} \right\|_{L^2(\Omega)^3} \leq \frac{\theta}{\alpha_{\max} + \alpha_{\min}} \left| p_h - p_h^{(n)} \right|_{H^1(\Omega)}. \tag{3.16}$$

By the inf-sup condition (3.1),

$$\begin{aligned} \beta \left| p_h - p_h^{(n+1)} \right|_{H^1(\Omega)} &\leq \sup_{\mathbf{v}_h \in X_h} \frac{b \left( \mathbf{v}_h, p_h - p_h^{(n+1)} \right)}{\left\| \mathbf{v}_h \right\|_{L^2(\Omega)^3}} \\ &= \sup_{\mathbf{v}_h \in X_h} \frac{\int_{\Omega} \left( \alpha(p_h) \mathbf{u}_h - \alpha(p_h^{(n)}) \mathbf{u}_h^{(n+1)} \right) \cdot \mathbf{v}_h}{\left\| \mathbf{v}_h \right\|_{L^2(\Omega)^3}} \\ &\leq \sup_{\mathbf{v}_h \in X_h} \frac{\int_{\Omega} \left( \alpha(p_h) - \alpha(p_h^{(n)}) \right) \mathbf{u}_h \cdot \mathbf{v}_h}{\left\| \mathbf{v}_h \right\|_{L^2(\Omega)^3}} + \sup_{\mathbf{v}_h \in X_h} \frac{\int_{\Omega} \alpha(p_h^{(n)}) \left( \mathbf{u}_h - \mathbf{u}_h^{(n+1)} \right) \cdot \mathbf{v}_h}{\left\| \mathbf{v}_h \right\|_{L^2(\Omega)^3}} \\ &\leq C(\Omega) L_{\alpha} \left| p_h - p_h^{(n)} \right|_{H^1(\Omega)} \left\| \mathbf{u}_h \right\|_{L^3(\Omega)^3} + \alpha_{\max} \left\| \mathbf{u}_h - \mathbf{u}_h^{(n+1)} \right\|_{L^2(\Omega)^3}. \end{aligned}$$

By condition (3.15) and inequality (3.16)

$$\begin{aligned} \beta \left| p_h - p_h^{(n+1)} \right|_{H^1(\Omega)} &\leq \frac{\theta}{\alpha_{\max} + \alpha_{\min}} (\alpha_{\max} + \alpha_{\min}) \left| p_h - p_h^{(n)} \right|_{H^1(\Omega)} \\ &= \theta \left| p_h - p_h^{(n)} \right|_{H^1(\Omega)}. \end{aligned}$$

From this inequality and (3.16) the claimed error bounds follow. □

**Remark 3.5.** One might argue that the previous error bounds do not guarantee convergence of the algorithm, since the value of  $\beta$  is not known and, hence, the ratio  $\theta/\beta$  could be greater than one. Using a similar assumption as (3.10), namely

$$\frac{1}{\beta} \frac{\alpha_{\max} + \alpha_{\min}}{\alpha_{\min}} C(\Omega) L_{\alpha} \left\| \mathbf{u}_h \right\|_{L^3(\Omega)^3} \leq \theta,$$

we can bypass this constraint. Moreover, as we have mentioned before, for the concrete examples of spaces (3.5)–(3.6) we have  $\beta = 1$ .

**Remark 3.6.** In addition to (3.1)–(3.3), assume that the following *inverse inequality* holds

$$\left\| \mathbf{v}_h \right\|_{L^3(\Omega)^3} \leq Ch^{-1/2} \left\| \mathbf{v}_h \right\|_{L^2(\Omega)^3}, \quad \forall \mathbf{v}_h \in X_h. \tag{3.17}$$

If the exact solution  $(\mathbf{u}, p)$  belongs to  $H^s(\Omega)^3 \times H^{s+1}(\Omega)$  for some real number  $s$  with  $\frac{1}{2} < s \leq 1$ , then the uniqueness condition (2.7) implies (3.15). Indeed, under these assumptions we have

$$\|\mathbf{u} - \mathbf{u}_h\|_{L^3(\Omega)^3} = \mathcal{O}(h^{s-\frac{1}{2}}),$$

hence, if

$$\frac{\alpha_{\max} + \alpha_{\min}}{\alpha_{\min}} C(\Omega) L_\alpha \|\mathbf{u}\|_{L^3(\Omega)^3} \leq \Theta < 1,$$

then,

$$\frac{\alpha_{\max} + \alpha_{\min}}{\alpha_{\min}} C(\Omega) L_\alpha \|\mathbf{u}_h\|_{L^3(\Omega)^3} \leq (1 + \mathcal{O}(h^{s-\frac{1}{2}}))\Theta.$$

If  $h$  is small enough, we obtain condition (3.15).

### 3.2. The nonuniqueness case. Approximation of nonsingular solutions

First, we introduce a final assumption on the function  $\alpha$ , namely

$$\alpha \in W_\infty^2(\mathbb{R}). \tag{3.18}$$

As we have mentioned before, in the truncated case this is not restrictive for the problem we are treating.

Next, we complement (3.1)–(3.3) and (3.17) with an additional *inverse inequality*:

$$\|q_h\|_{L^\infty(\Omega)} \leq Ch^{-1/2} |q_h|_{H^1(\Omega)}, \quad \forall q_h \in M_h. \tag{3.19}$$

Both inverse inequalities (3.17) and (3.19) hold when the family of triangulations  $\mathcal{T}_h$  is quasi-uniform (or uniformly regular) in the following sense (cf. [12]): In addition to (3.4), there exists a constant  $\tau > 0$ , independent of  $h$ , such that

$$\forall T \in \mathcal{T}_h, h_T \geq \tau h. \tag{3.20}$$

We are now concerned with the approximation of nonsingular solutions to (2.10) under the hypotheses (3.1)–(3.3), (3.17), and (3.19). In order to do that, let us define the discrete solution operator to the linear Darcy equations  $T_h : \mathfrak{Y} \rightarrow \mathfrak{X}_h$ . That is, for any  $\eta = (\mathbf{f}, g) \in \mathfrak{Y}$ ,  $\mathfrak{X}_h \ni x_h = (\mathbf{u}_h, p_h) = T_h \eta = T_h(\mathbf{f}, g)$  solves

$$\begin{cases} a(\mathbf{u}_h, \mathbf{v}_h) + b(\mathbf{v}_h, p_h) = \int_\Omega \mathbf{f} \cdot \mathbf{v}_h, & \forall \mathbf{v}_h \in X_h, \\ b(\mathbf{u}_h, q_h) = \langle g, q_h \rangle_\Gamma, & \forall q_h \in M_h, \end{cases}$$

where the bilinear form  $a : L^2(\Omega)^3 \times L^2(\Omega)^3$  is defined by

$$a(\mathbf{u}, \mathbf{v}) := \bar{\alpha} \int_\Omega \mathbf{u} \cdot \mathbf{v}.$$

It is a classical matter [10,17] to show that, under assumption (3.1), this operator is well-defined, injective,  $T_h \in \mathcal{L}(\mathfrak{Y}, \mathfrak{X}_h)$ , and there is a constant  $C$  independent of  $h$  such that

$$\|T_h(\mathbf{f}, g)\|_{\mathfrak{X}} \leq C \|(\mathbf{f}, g)\|_{\mathfrak{Y}}, \quad \forall (\mathbf{f}, g) \in \mathfrak{Y}. \tag{3.21}$$

We can also define the discrete nonlinearity. This is an operator  $G_h : \mathfrak{X}_h \rightarrow X_h \times H_{00}^{1/2}(\Gamma)' \subset \mathfrak{Y}$ , such that if  $x_h = (\mathbf{u}_h, p_h) \in \mathfrak{X}_h$ , then  $G_h(x_h) := (\mathbf{F}_h, -g)$ , where  $\mathbf{F}_h \in X_h$  is the unique solution to

$$\int_\Omega \mathbf{F}_h \cdot \mathbf{v}_h = \int_\Omega [(\alpha(p_h) - \bar{\alpha}) \mathbf{u}_h - \mathbf{f}] \cdot \mathbf{v}_h, \quad \forall \mathbf{v}_h \in X_h.$$

Finally, define the operator  $F_h : \mathfrak{X}_h \rightarrow \mathfrak{X}_h$  by

$$F_h(x_h) := x_h + T_h G_h(x_h).$$

With this notation, problem (3.9) can be equivalently rewritten as:

Find  $x_h \in \mathfrak{X}_h$  such that

$$F_h(x_h) = 0. \tag{3.22}$$

The approximation properties of the operator  $T_h$  are the following.

**Proposition 3.7.** *Assume that (3.1)–(3.3) hold. Let  $(\mathbf{f}, g) \in \mathfrak{Y}$  be such that  $T(\mathbf{f}, g) \in H^s(\Omega)^3 \times H^{1+s}(\Omega) \subset \mathfrak{X}$ , for some  $0 < s \leq k$ . Then, there is a constant  $C > 0$ , independent of  $h$  such that*

$$\|(T - T_h)(\mathbf{f}, g)\|_{\mathfrak{X}} \leq Ch^s \|T(\mathbf{f}, g)\|_{H^s(\Omega)^3 \times H^{1+s}(\Omega)}. \tag{3.23}$$

*Proof.* It is a direct consequence of assumptions (3.1)–(3.3), together with a basic interpolation argument [6].  $\square$

**Corollary 3.8.** *Under the hypotheses of Proposition 3.7, the operator  $T_h$  satisfies*

$$\lim_{h \rightarrow 0} \|(T - T_h)\|_{\mathcal{L}(\mathfrak{Y}, \mathfrak{X})} = 0. \tag{3.24}$$

*Proof.* Standard regularity results for the linear Darcy problem (2.8) imply that, for sufficiently small  $s > 0$ ,  $T(\mathbf{f}, g) \in H^s(\Omega)^3 \times H^{1+s}(\Omega)$  if  $(\mathbf{f}, g)$  belongs to  $\mathfrak{Y} := H^s(\Omega)^3 \times H^{s-1/2}(\partial\Omega)$ , which is a dense subset of  $\mathfrak{Y}$ . The boundedness of operator  $T$  (see (2.9)), together with inequality (3.23) imply

$$\begin{aligned} \sup_{(\mathbf{f}, g) \in \mathfrak{Y}} \frac{\|(T - T_h)(\mathbf{f}, g)\|_{\mathfrak{X}}}{\|(\mathbf{f}, g)\|_{\mathfrak{Y}}} &= \sup_{(\mathbf{f}, g) \in \mathfrak{Y}} \frac{\|(T - T_h)(\mathbf{f}, g)\|_{\mathfrak{X}}}{\|(\mathbf{f}, g)\|_{\mathfrak{Y}}} \\ &\leq Ch^s \frac{\|T(\mathbf{f}, g)\|_{\mathfrak{X}}}{\|(\mathbf{f}, g)\|_{\mathfrak{Y}}} \\ &\leq Ch^s, \end{aligned}$$

from which (3.24) clearly follows.  $\square$

We are interested in approximating a nonsingular solution  $x = (\mathbf{u}, p) \in \mathfrak{X}$  to (2.10). For this, we must assume that there is a real number  $s > 1/2$  such that

$$(\mathbf{u}, p) \in H^s(\Omega)^3 \times H^{1+s}(\Omega). \tag{3.25}$$

**Remark 3.9.** Since  $s > 1/2$ , (3.25) implies that  $(\mathbf{u}, p) \in L^3(\Omega)^3 \times C^0(\bar{\Omega})$ , see [1].

To alleviate the notation, define

$$x_h^0 := (\mathbf{u}_h^0, p_h^0) = (\pi_h \mathbf{u}, \mathcal{I}_h p) \in \mathfrak{X}_h, \tag{3.26}$$

where  $\pi_h$  and  $\mathcal{I}_h$  are the interpolation operators of (3.2) and (3.3) respectively. Important properties of the interpolant  $x_h^0$  and the operator  $F_h'(x_h^0)$  are established below.

**Lemma 3.10.** *Let the function  $\alpha$  satisfy conditions (2.1), (2.2) and (3.18). Let the solution  $(\mathbf{u}, p) \in \mathfrak{X}$  to problem (2.10) be nonsingular and satisfy the smoothness condition (3.25). If the pair of spaces  $(X_h, M_h)$  satisfies assumptions (3.2), (3.3), then there exists a constant  $C > 0$  independent of  $h$ , such that*

$$\|\mathbf{u} - \mathbf{u}_h^0\|_{L^2(\Omega)^3} \leq Ch^s |\mathbf{u}|_{H^s(\Omega)^3}, \tag{3.27}$$



and

$$|p - p_h^0|_{H^1(\Omega)} \leq Ch^s |p|_{H^{1+s}(\Omega)}. \tag{3.28}$$

Moreover, if the pair  $(X_h, M_h)$  also satisfies conditions (3.1), (3.17) and (3.19), then there exists a  $h_0 > 0$  such that for every  $h \leq h_0$  the operator  $F'_h(x_h^0)$  is an isomorphism of  $\mathfrak{X}_h$  and the norm of its inverse is bounded independently of  $h$ .

*Proof.* Inequalities (3.27) and (3.28) are a simple consequence of (3.2), (3.3) and assumption (3.25) via interpolation [6].

To show that  $F'_h(x_h^0)$  is an isomorphism of  $\mathfrak{X}_h$ , notice that

$$I + T_h G'_h(x_h^0) = I + T_h G'(x) + T_h (G'(x_h^0) - G'(x)) + T_h (G'_h(x_h^0) - G'(x_h^0)).$$

Let us consider each term separately.

(1)  $I + T_h G'(x)$ . Notice, first of all, that if  $y_h \in \mathfrak{X}_h$ , then  $(I + T_h G'(x)) y_h \in \mathfrak{X}_h$ . Moreover,

$$I + T_h G'(x) - F'(x) = (T_h - T) G'(x).$$

Since  $x$  is a nonsingular solution,  $F'(x)$  is an isomorphism of  $\mathfrak{X}$ . Corollary 3.8 and an application of the Theorem about the Perturbation of an Invertible Operator (see [26], Thm. 4, p. 207 for instance) imply that there is  $h_0^{(1)} > 0$  such that for all  $h \leq h_0^{(1)}$  the operator  $I + T_h G'(x)$  is an isomorphism of  $\mathfrak{X}$ . Hence it is an isomorphism of  $\mathfrak{X}_h$ . Thus, the result of the lemma will be proved if we show that the remaining two terms tend to zero (in the  $\|\cdot\|_{\mathcal{L}(\mathfrak{X}_h)}$ -norm) as  $h \rightarrow 0$ .

(2)  $T_h(G'(x_h^0) - G'(x))$ . Let  $y_h = (\mathbf{v}_h, q_h)$ ; using the definition of the derivatives, for any  $\mathbf{w} \in L^2(\Omega)^3$

$$\begin{aligned} \langle (G'(x_h^0) - G'(x))y_h, (\mathbf{w}, 0) \rangle &= \int_{\Omega} (\alpha(p_h^0) - \alpha(p)) \mathbf{v}_h \cdot \mathbf{w} + \int_{\Omega} (\dot{\alpha}(p_h^0) \mathbf{u}_h^0 - \dot{\alpha}(p) \mathbf{u}) q_h \cdot \mathbf{w} \\ &= \int_{\Omega} (\alpha(p_h^0) - \alpha(p)) \mathbf{v}_h \cdot \mathbf{w} + \int_{\Omega} (\dot{\alpha}(p_h^0) - \dot{\alpha}(p)) q_h \mathbf{u} \cdot \mathbf{w} \\ &\quad + \int_{\Omega} \dot{\alpha}(p_h^0) (\mathbf{u}_h^0 - \mathbf{u}) q_h \cdot \mathbf{w}. \end{aligned}$$

Consider each term separately. By (2.2) and the inverse inequality (3.17)

$$\begin{aligned} \int_{\Omega} (\alpha(p_h^0) - \alpha(p)) \mathbf{v}_h \cdot \mathbf{w} &\leq C(\Omega) L_{\alpha} |p - p_h^0|_{H^1(\Omega)} \|\mathbf{v}_h\|_{L^3(\Omega)^3} \|\mathbf{w}\|_{L^2(\Omega)^3} \\ &\leq Ch^{-1/2} |p - p_h^0|_{H^1(\Omega)} \|\mathbf{v}_h\|_{L^2(\Omega)^3} \|\mathbf{w}\|_{L^2(\Omega)^3}. \end{aligned}$$

By (3.18) and the inverse inequality (3.19)

$$\begin{aligned} \int_{\Omega} (\dot{\alpha}(p_h^0) - \dot{\alpha}(p)) q_h \mathbf{u} \cdot \mathbf{w} &\leq C \|q_h\|_{L^{\infty}(\Omega)} \int_{\Omega} |p - p_h^0| |\mathbf{u}| |\mathbf{w}| \\ &\leq Ch^{-1/2} |p - p_h^0|_{H^1(\Omega)} \|q_h\|_{H^1(\Omega)} \|\mathbf{u}\|_{L^3(\Omega)^3} \|\mathbf{w}\|_{L^2(\Omega)^3}. \end{aligned}$$

Finally, by (2.2) and the inverse inequality (3.19)

$$\begin{aligned} \int_{\Omega} \dot{\alpha}(p_h^0) (\mathbf{u}_h^0 - \mathbf{u}) q_h \cdot \mathbf{w} &\leq L_{\alpha} \|\mathbf{u} - \mathbf{u}_h^0\|_{L^2(\Omega)^3} \|q_h\|_{L^{\infty}(\Omega)} \|\mathbf{w}\|_{L^2(\Omega)^3} \\ &\leq Ch^{-1/2} \|\mathbf{u} - \mathbf{u}_h^0\|_{L^2(\Omega)^3} \|q_h\|_{H^1(\Omega)} \|\mathbf{w}\|_{L^2(\Omega)^3}. \end{aligned}$$

Thus, by the stability property (3.21) of  $T_h$ ,

$$\begin{aligned} \|T_h(G'(x_h^0) - G'(x))\|_{\mathcal{L}(\mathfrak{X}_h)} &\leq C \|G'(x_h^0) - G'(x)\|_{\mathcal{L}(\mathfrak{X}_h, \mathfrak{Y})} \\ &= C \sup_{y_h \in \mathfrak{X}_h} \frac{\|(G'(x_h^0) - G'(x))y_h\|_{\mathfrak{Y}}}{\|y_h\|_{\mathfrak{X}}} \\ &= C \sup_{y_h \in \mathfrak{X}_h} \sup_{\mathbf{w} \in L^2(\Omega)^3} \frac{\langle (G'(x_h^0) - G'(x))y_h, \mathbf{w} \rangle}{\|y_h\|_{\mathfrak{X}} \|\mathbf{w}\|_{L^2(\Omega)^3}} \\ &\leq Ch^{-1/2} \left( |p - p_h^0|_{H^1(\Omega)} + \|\mathbf{u} - \mathbf{u}_h^0\|_{L^2(\Omega)^3} \right), \end{aligned}$$

which by the approximation properties (3.27) and (3.28) of  $x_h^0$  and the fact that  $s > 1/2$  implies that this last quantity tends to zero as  $h \rightarrow 0$ .

(3)  $T_h(G'_h(x_h^0) - G'(x_h^0))$ . It is sufficient to notice that for any  $\mathbf{w}_h \in X_h$

$$\langle (G'_h(x_h^0) - G'(x_h^0))y_h, (\mathbf{w}_h, 0) \rangle = 0. \quad \square$$

**Remark 3.11.** In the example (3.5), (3.6), as in most finite element spaces, inverse estimates such as (3.17) and (3.19) hold locally. Therefore they may be applied locally when used in proving the interpolation Lemma 3.10, because interpolation properties are also local. In this case, the statement of Lemma 3.10 is valid even if the triangulation is not quasi-uniform. But of course intermediate results would have to be stated differently. For instance the bound for

$$\int_{\Omega} (\alpha(p_h^0) - \alpha(p)) \mathbf{v}_h \cdot \mathbf{w}$$

would read, for  $s > \frac{1}{2}$ :

$$\int_{\Omega} (\alpha(p_h^0) - \alpha(p)) \mathbf{v}_h \cdot \mathbf{w} \leq Ch^{s-1/2} |p|_{H^{1+s}(\Omega)} \|\mathbf{v}_h\|_{L^2(\Omega)^3} \|\mathbf{w}\|_{L^2(\Omega)^3}.$$

However, this does not apply to inverse inequalities that are used in conjunction with global error estimates, such as in Remark 3.6 or in Lemma 3.12 below, in which case some restriction on the mesh cannot be avoided.

Once we know the main properties of the operator  $F'_h(x_h^0)$ , it is possible to study  $F'_h(y_h)$  for  $y_h$  close to  $x_h^0$ .

**Lemma 3.12.** *Under the assumptions of Lemma 3.10, there is a constant  $C_0 > 0$  independent of  $h$  such that*

$$\|G'_h(y_h) - G'_h(x_h^0)\|_{\mathcal{L}(\mathfrak{X}_h, \mathfrak{Y})} \leq C_0 h^{-1/2} \|y_h - x_h^0\|_{\mathfrak{X}}, \quad \forall y_h \in \mathfrak{X}_h. \quad (3.29)$$

*Proof.* Let  $y_h = (\mathbf{v}_h, q_h), z_h = (\mathbf{w}_h, r_h) \in \mathfrak{X}_h$ . For an arbitrary  $\mathbf{t}_h \in X_h$

$$\begin{aligned} \langle (G'_h(y_h) - G'_h(x_h^0))z_h, (\mathbf{t}_h, 0) \rangle &= \int_{\Omega} (\alpha(q_h) - \alpha(p_h^0)) \mathbf{w}_h \cdot \mathbf{t}_h + \int_{\Omega} \dot{\alpha}(q_h) (\mathbf{v}_h - \mathbf{u}_h^0) r_h \cdot \mathbf{t}_h \\ &\quad + \int_{\Omega} (\dot{\alpha}(q_h) - \dot{\alpha}(p_h^0)) \mathbf{u}_h^0 r_h \cdot \mathbf{t}_h \\ &\leq C \left( \|p_h^0 - q_h\|_{L^\infty(\Omega)} \|\mathbf{w}_h\|_{L^2(\Omega)^3} \|\mathbf{t}_h\|_{L^2(\Omega)^3} \right. \\ &\quad + \|\mathbf{u}_h^0 - \mathbf{v}_h\|_{L^3(\Omega)^3} |r_h|_{H^1(\Omega)} \|\mathbf{t}_h\|_{L^2(\Omega)^3} \\ &\quad \left. + \|p_h^0 - q_h\|_{L^\infty(\Omega)} \|\mathbf{u}_h^0\|_{L^3(\Omega)^3} |r_h|_{H^1(\Omega)} \|\mathbf{t}_h\|_{L^2(\Omega)^3} \right), \end{aligned}$$

hence

$$\|G'_h(y_h) - G'_h(x_h^0)\|_{\mathcal{L}(\mathfrak{X}_h, \mathfrak{Y})} \leq C \left( \|p_h^0 - q_h\|_{L^\infty(\Omega)} + \|\mathbf{u}_h^0 - \mathbf{v}_h\|_{L^3(\Omega)^3} \right).$$

This estimate and the inverse inequalities (3.17), (3.19) imply (3.29). □

**Remark 3.13.** Lemma 3.12 states that  $G'_h$  is Lipschitz-continuous in a neighborhood of  $x_h^0$ , but this continuity is not uniform with respect to  $h$ . One more time, the absence of regularizing properties for the nonlinearity  $G$  does not allow us to obtain uniform in  $h$  bounds.

It is important to know whether the consistency error  $F_h(x_h^0)$  tends to zero as  $h \rightarrow 0$ , and if this is the case at which rate. The following lemma shows that the convergence is optimal given the regularity of the exact nonsingular solution  $x$ .

**Lemma 3.14.** *Under the assumptions of the first part of Lemma 3.10, there is a constant  $C > 0$ , independent of  $h$  such that*

$$\|F_h(x_h^0)\|_{\mathfrak{X}} \leq Ch^s \left( |\mathbf{u}|_{H^s(\Omega)^3} + |p|_{H^{1+s}(\Omega)} \right). \tag{3.30}$$

*Proof.* Since  $F(x) = 0$ ,

$$F_h(x_h^0) = x_h^0 - x + T_h(G_h(x_h^0) - G(x)) + (T_h - T)G(x),$$

which implies

$$\|F_h(x_h^0)\|_{\mathfrak{X}} \leq \|x - x_h^0\|_{\mathfrak{X}} + \|(T - T_h)G(x)\|_{\mathfrak{X}} + \|T_h(G(x) - G_h(x_h^0))\|_{\mathfrak{X}}.$$

From (3.27) and (3.28),

$$\|x - x_h^0\|_{\mathfrak{X}} \leq Ch^s \left( |\mathbf{u}|_{H^s(\Omega)^3} + |p|_{H^{1+s}(\Omega)} \right).$$

Estimate (3.23) implies

$$\|(T - T_h)G(x)\|_{\mathfrak{X}} \leq Ch^s \|TG(x)\|_{H^s(\Omega)^3 \times H^{1+s}(\Omega)} = Ch^s \left( |\mathbf{u}|_{H^s(\Omega)^3} + |p|_{H^{1+s}(\Omega)} \right).$$

Finally, since  $T_h(G_h(x_h^0) - G(x))$  belongs to  $\mathfrak{X}_h$ , by the stability property (3.21) of  $T_h$  we see that it is sufficient to control the difference of the first coordinate of  $G(x) - G_h(x_h^0)$  when tested against an element of  $X_h$ . Let  $\mathbf{v}_h \in X_h$ , then using (3.27) and (3.28)

$$\begin{aligned} \int_{\Omega} [G(x) - G_h(x_h^0)]_1 \cdot \mathbf{v}_h &\leq (\bar{\alpha} + \alpha_{\max}) \|\mathbf{u} - \mathbf{u}_h^0\|_{L^2(\Omega)^3} \|\mathbf{v}_h\|_{L^2(\Omega)^3} + C(\Omega)L_\alpha |p - p_h^0|_{H^1(\Omega)} \|\mathbf{u}\|_{L^3(\Omega)^3} \|\mathbf{v}_h\|_{L^2(\Omega)^3} \\ &\leq Ch^s \left( |\mathbf{u}|_{H^s(\Omega)^3} + |p|_{H^{1+s}(\Omega)} \right) \|\mathbf{v}_h\|_{L^2(\Omega)^3}. \end{aligned} \tag{3.30}$$

According to the theory in [11,20], Lemmas 3.10, 3.12, and 3.14 allow us to prove our main result, namely, the existence of a nonsingular solution for the discrete problem and optimal error estimates for it.

**Theorem 3.15.** *Let  $\alpha$  satisfy (2.1), (2.2) and (3.18). Assume that problem (2.10) has a nonsingular solution  $x = (\mathbf{u}, p) \in H^s(\Omega)^3 \times H^{1+s}(\Omega) \subset \mathfrak{X}$ , for some  $s > 1/2$ . If the pair of spaces  $(X_h, M_h)$  satisfies (3.1), (3.2), (3.3), (3.17), and (3.19), then there is a  $h_0 > 0$  such that for all  $h \leq h_0$  the discrete problem (3.22) has a unique nonsingular solution  $x_h = (\mathbf{u}_h, p_h)$  in a neighborhood of the interpolant  $x_h^0 = (\mathbf{u}_h^0, p_h^0)$  of the exact nonsingular solution. Moreover, this solution satisfies the following error estimate*

$$\|x - x_h\|_{\mathfrak{X}} \leq Ch^s \left( |\mathbf{u}|_{H^s(\Omega)^3} + |p|_{H^{1+s}(\Omega)} \right), \tag{3.31}$$

where the constant  $C > 0$  does not depend on  $h$ .

*Proof.* Let us define

$$\epsilon_h := \|F_h(x_h^0)\|_{\mathfrak{X}},$$

and

$$M_h(\delta) := \sup_{y_h \in \mathfrak{X}_h, \|y_h - x_h^0\|_{\mathfrak{X}} < \delta} \|F'_h(y_h) - F'_h(x_h^0)\|_{\mathcal{L}(\mathfrak{X})}.$$

Lemma 3.10 implies that there is a  $h_0^{(1)} > 0$  such that for all  $h \leq h_0^{(1)}$  the operator  $F'_h(x_h^0)$  is an isomorphism of  $\mathfrak{X}_h$  with inverse bounded independently of  $h$ . Denote this bound by  $\Delta$ . Inequalities (3.29) and (3.30) imply that

$$2\Delta M_h(2\Delta\epsilon_h) \leq Ch^{s-1/2},$$

hence there is a  $h_0^{(2)} > 0$  such that for all  $h \leq h_0^{(2)}$

$$2\Delta M_h(2\Delta\epsilon_h) < 1.$$

Set  $h_0 = \min\{h_0^{(1)}, h_0^{(2)}\}$  and consider  $h \leq h_0$ .

Since the operator  $F'_h(x_h^0)$  is an isomorphism, solving problem (3.22) is equivalent to finding a fixed point of the map  $\Phi_h : \mathfrak{X}_h \rightarrow \mathfrak{X}_h$  defined by

$$\Phi_h(y_h) := y_h - [F'_h(x_h^0)]^{-1} F_h(y_h).$$

Denote

$$S := \{y_h \in \mathfrak{X}_h : \|y_h - x_h^0\|_{\mathfrak{X}} \leq 2\Delta\epsilon_h\}.$$

We shall show that  $\Phi_h$  is a contraction from  $S$  to  $S$ .

If  $y_h \in S$ ,

$$\Phi_h(y_h) - x_h^0 = [F'_h(x_h^0)]^{-1} (F'_h(x_h^0)(y_h - x_h^0) - (F_h(y_h) - F_h(x_h^0)) - F_h(x_h^0)).$$

By the Mean Value Theorem

$$F_h(y_h) - F_h(x_h^0) = \int_0^1 F'_h(x_h^0 + \theta(y_h - x_h^0))(y_h - x_h^0) d\theta,$$

from which follows

$$\begin{aligned} \|F'_h(x_h^0)(y_h - x_h^0) - (F_h(y_h) - F_h(x_h^0))\|_{\mathfrak{X}} &\leq \int_0^1 \|F'_h(x_h^0) - F'_h(x_h^0 + \theta(y_h - x_h^0))\|_{\mathcal{L}(\mathfrak{X}_h)} \|y_h - x_h^0\|_{\mathfrak{X}} d\theta \\ &\leq 2\Delta\epsilon_h M_h(2\Delta\epsilon_h). \end{aligned}$$

And, by the choice of  $h$

$$\|\Phi_h(y_h) - x_h^0\|_{\mathfrak{X}} \leq \Delta(2\Delta\epsilon_h M_h(2\Delta\epsilon_h) + \epsilon_h) = \Delta\epsilon_h(2\Delta M_h(2\Delta\epsilon_h) + 1) < 2\Delta\epsilon_h,$$

which means that  $\Phi_h(y_h) \in S$ .

Let  $y_h, z_h \in S$ , then a similar computation shows that

$$\|\Phi_h(y_h) - \Phi_h(z_h)\|_{\mathfrak{X}} \leq \Delta M_h(2\Delta\epsilon_h) \|y_h - z_h\|_{\mathfrak{X}} < \frac{1}{2} \|y_h - z_h\|_{\mathfrak{X}},$$

which implies that  $\Phi_h$  is a contraction and we can conclude that there is a unique  $x_h \in S$  such that  $x_h = \Phi_h(x_h)$ .

To realize that this solution is nonsingular, notice that

$$\|F'_h(x_h^0) - F'_h(x_h)\|_{\mathcal{L}(\mathfrak{X}_h)} \leq M_h(2\Delta\epsilon_h) < \frac{1}{2\Delta},$$

and apply the Theorem about the Perturbation of an Invertible Operator (see Kantorovich and Akilov [26], Thm. 4, p. 207 for instance).

Finally, to get the error estimate (3.31) it is sufficient to use (3.30), the triangle inequality; and properties (3.27) and (3.28) of  $x_h^0$ ,

$$\begin{aligned} \|x_h - x\|_{\mathfrak{X}} &\leq \|x_h - x_h^0\|_{\mathfrak{X}} + \|x_h^0 - x\|_{\mathfrak{X}} \\ &\leq 2\Delta\epsilon_h + Ch^s \left( |\mathbf{u}|_{H^s(\Omega)^3} + |p|_{H^{1+s}(\Omega)} \right) \\ &\leq Ch^s \left( |\mathbf{u}|_{H^s(\Omega)^3} + |p|_{H^{1+s}(\Omega)} \right). \end{aligned} \quad \square$$

This concludes the proof.

**Remark 3.16.** From the proof of this theorem we see that the discrete nonsingular solution  $x_h$  is unique in a ball larger than  $S$ . Namely, it is unique in the ball

$$S(\bar{\delta}) := \{y_h \in \mathfrak{X}_h : \|y_h - x_h^0\|_{\mathfrak{X}} < \bar{\delta}\},$$

where  $\bar{\delta}$  is such that  $\Delta M_h(\bar{\delta}) < 1$ . Both radii tend to zero as  $h \rightarrow 0$ . But, according to (3.30), the radius of  $S$  is  $\mathcal{O}(h^s)$ ,  $s > 1/2$ , whereas  $\bar{\delta} = \mathcal{O}(h^{1/2})$ .

We have obtained that the discrete problem (3.22) has a unique nonsingular solution in a neighborhood of the exact nonsingular solution. We now analyze the application of Newton's method to the solution of this discrete problem. The algorithm is the following:

Given  $x_h^{(0)} \in \mathfrak{X}_h$ , for  $n = 0, 1, \dots$  define  $x_h^{(n+1)}$  by

$$x_h^{(n+1)} = x_h^{(n)} - \left[ F'_h \left( x_h^{(n)} \right) \right]^{-1} F_h \left( x_h^{(n)} \right).$$

For this method to make sense  $F'_h \left( x_h^{(n)} \right)$  must be an isomorphism of  $\mathfrak{X}_h$  for all  $n$ . Let us introduce the following notation

$$S(x_h, \delta) := \{y_h \in \mathfrak{X}_h : \|y_h - x_h\|_{\mathfrak{X}} < \delta\},$$

and,

$$K := \frac{1}{4 \|T_h\|_{\mathcal{L}(\mathfrak{Y}, \mathfrak{X}_h)} C_0 \Delta},$$

where the constant  $C_0$  is the constant in inequality (3.29),  $\Delta$  is such that for  $h$  small enough

$$\left\| \left[ F'_h(x_h^0) \right]^{-1} \right\|_{\mathcal{L}(\mathfrak{X}_h)} \leq \Delta,$$

and  $x_h^0$  is the interpolant of  $x$  defined in (3.26).

**Lemma 3.17.** *There exists a real number  $h_0 > 0$  such that for all  $h \leq h_0$ , if  $\delta = \mathcal{O}(h^{1/2})$  and  $y_h \in S(x_h, \delta)$ , then the linear operator  $F'_h(y_h)$  is an isomorphism of  $\mathfrak{X}_h$ . Moreover, the norm of the inverse of this operator is bounded independently of  $h$ .*

*Proof.* Since

$$F'_h(y_h) = F'_h(x_h) + (F'_h(y_h) - F'_h(x_h)),$$

and, by Theorem 3.15, there exists  $h_0 > 0$  such that for all  $h \leq h_0$ ,  $F'_h(x_h)$  is an isomorphism of  $\mathfrak{X}_h$ , the result is obtained if we show that  $F'_h(y_h) - F'_h(x_h)$  is small enough. We know that,

$$\left\| [F'_h(x_h)]^{-1} \right\|_{\mathcal{L}(\mathfrak{X}_h)} \leq 2\Delta.$$

A similar argument as in the proof of Lemma 3.12 gives us that

$$\|F'_h(y_h) - F'_h(x_h)\|_{\mathcal{L}(\mathfrak{X}_h)} \leq C_0 h^{-1/2} \|T_h\|_{\mathcal{L}(\mathfrak{Y}, \mathfrak{X}_h)} \|y_h - x_h\|_{\mathfrak{X}}.$$

Hence, if

$$2C_0 \|T_h\|_{\mathcal{L}(\mathfrak{Y}, \mathfrak{X}_h)} \|y_h - x_h\|_{\mathfrak{X}} \Delta h^{-1/2} < 1,$$

then the Theorem about the Perturbation of an Invertible Operator implies that  $F'_h(y_h)$  is an isomorphism of  $\mathfrak{X}_h$ . Moreover, from this inequality we see that it is sufficient to set

$$\delta \leq Kh^{1/2},$$

where  $K$  is a constant independent of  $h$ . □

**Theorem 3.18.** *There exists a real number  $h_0 > 0$  such that for all  $h \leq h_0$ , if*

$$\delta \leq \epsilon Kh^{1/2},$$

for some real number  $\epsilon$  with  $0 < \epsilon < 1$ , and if the initial approximation of Newton's method  $x_h^{(0)}$  belongs to  $S(x_h, \delta)$ , then Newton's method converges to the discrete nonsingular solution  $x_h$  and the following error estimate holds

$$\|x_h^{(n+1)} - x_h\|_{\mathfrak{X}} \leq \frac{1}{K} h^{-1/2} \|x_h^{(n)} - x_h\|_{\mathfrak{X}}^2.$$

*Proof.* Assume  $h$  is small enough. Let us show by induction that if  $x_h^{(0)} \in S(x_h, \delta)$ , then  $x_h^{(n)} \in S(x_h, \delta)$  for all  $n > 0$ . If  $x_h^{(n)}$  is in  $S(x_h, \delta)$  and  $\delta$  is chosen as indicated, then by the previous lemma,  $K$  can be chosen independently of  $h$ , so that  $F'_h(x_h^{(n)})$  is an isomorphism of  $\mathfrak{X}_h$ , with

$$\left\| [F'_h(x_h^{(n)})]^{-1} \right\|_{\mathcal{L}(\mathfrak{X}_h)} \leq 4\Delta.$$

Furthermore with a similar argument as in the proof of Theorem 3.15 we obtain

$$\begin{aligned} x_h^{(n+1)} - x_h &= [F'_h(x_h^{(n)})]^{-1} \left( F'_h(x_h^{(n)})(x_h^{(n)} - x_h) - (F_h(x_h^{(n)}) - F_h(x_h)) \right) \\ &= [F'_h(x_h^{(n)})]^{-1} \int_0^1 [F'_h(x_h^{(n)}) - F'_h(x_h^{(n)} - \theta(x_h^{(n)} - x_h))] (x_h^{(n)} - x_h) d\theta. \end{aligned}$$

Then, by the induction hypothesis, a similar argument as in Lemma 3.12 and the choice of  $\delta$  and  $K$  imply

$$\begin{aligned} \|x_h^{(n+1)} - x_h\|_{\mathfrak{X}} &\leq \left\| \left[ F'_h \left( x_h^{(n)} \right) \right]^{-1} \right\|_{\mathcal{L}(\mathfrak{X}_h)} \|T_h\|_{\mathcal{L}(\mathfrak{Y}, \mathfrak{X}_h)} \\ &\quad \times \int_0^1 \left\| G'_h \left( x_h^{(n)} \right) - G'_h \left( x_h^{(n)} - \theta \left( x_h^{(n)} - x_h \right) \right) \right\|_{\mathcal{L}(\mathfrak{Y}, \mathfrak{X}_h)} d\theta \|x_h^{(n)} - x_h\|_{\mathfrak{X}} \\ &\leq 4\Delta \|T_h\|_{\mathcal{L}(\mathfrak{Y}, \mathfrak{X}_h)} C_0 h^{-1/2} \|x_h^{(n)} - x_h\|_{\mathfrak{X}}^2 \\ &\leq \epsilon \|x_h^{(n)} - x_h\|_{\mathfrak{X}}. \end{aligned}$$

On one hand, this shows that  $x_h^{(n+1)} \in S(x_h, \delta)$  and hence, by Lemma 3.17, that  $F'_h(x_h^{(n+1)})$  is an isomorphism of  $\mathfrak{X}_h$  for all  $n \geq 1$ , on the other hand this shows the claimed error estimate.  $\square$

**Remark 3.19.** As we can see, the initial guess in Newton’s method must be very close to the discrete solution. Moreover, the convergence of the method deteriorates as the discretization parameter  $h$  tends to zero. This is again related to the lack of regularizing properties for the nonlinearity  $G$ , as is reflected by Lemma 3.12.

#### 4. A SPLITTING ALGORITHM FOR EXPONENTIAL POROSITY

The preceding analysis does not apply to an exponential porosity  $\alpha$ , since assumptions (2.1) and (2.2) are not satisfied. So far, a rigorous analysis of this problem is beyond our reach. Nevertheless, for the exponential case, we propose a split formulation derived heuristically by taking the divergence of the first equation of (1.1) and making a change of variable. Thus, by precisely exploiting the exponential character of the porosity (1.2), we are able to decompose the nonlinear Darcy problem into a linear elliptic equation and a linear Darcy system. But this process is heuristic since we develop this method without even knowing whether in general problem (1.1), with the porosity defined as (1.2), does have a solution.

This section is organized as follows. First, we present the motivation behind the split formulation, next we study the properties of the solution to the auxiliary problem, *i.e.* the linear elliptic equation. Finally, we discretize the split formulation and we study the convergence of the resulting algorithm.

##### 4.1. Motivation

Let  $(\mathbf{u}, p)$  be a solution of problem (1.1) with the porosity given by (1.2) and assume that  $p$  belongs to  $L^\infty(\Omega)$ . Since  $\alpha(p) > 0$ , we can divide the first equation in (1.1) by  $\alpha(p)$ , take the divergence of the result, and make a suitable change in variable. Using the second equation of (1.1), we obtain

$$0 = \nabla \cdot \mathbf{u} = \nabla \cdot \left( \frac{1}{\alpha(p)} \mathbf{f} - \frac{1}{\alpha(p)} \nabla p \right).$$

Since  $1/\alpha(p) = 1/\alpha_0 e^{-\gamma p}$ , then

$$\frac{1}{\alpha(p)} \nabla p = \frac{1}{\alpha_0} e^{-\gamma p} \nabla p = -\frac{1}{\alpha_0 \gamma} \nabla e^{-\gamma p},$$

and the above equation can be rewritten as

$$-\Delta e^{-\gamma p} = \gamma \nabla \cdot (e^{-\gamma p} \mathbf{f}). \tag{4.1}$$

Let us introduce the new variable

$$q = e^{-\gamma p} - 1. \tag{4.2}$$

Since  $p = 0$  on  $\Gamma_w$ ,

$$q = e^{-\gamma p|_{\Gamma_w}} - 1 = 0 \text{ on } \Gamma_w.$$

From (4.1) and (4.2), this new variable satisfies a.e. in  $\Omega$

$$-\Delta q - \gamma \nabla \cdot (q\mathbf{f}) = \gamma \nabla \cdot \mathbf{f},$$

$$\alpha(p) = \frac{\alpha_0}{q + 1}. \tag{4.3}$$

Assume that the right-hand side  $\mathbf{f}$  is smooth enough so that it has a normal trace on  $\Gamma$ . Then it is legitimate to multiply the first equation of (1.1) by  $\mathbf{n}$  on  $\Gamma$  and obtain

$$\alpha(p)g + \partial_n p = \mathbf{f} \cdot \mathbf{n}.$$

Denote  $\tilde{F} := \mathbf{f} \cdot \mathbf{n}$ . By (4.2),

$$\partial_n q + \gamma \tilde{F} q = \alpha_0 \gamma g - \gamma \tilde{F}.$$

Thus, for the variable  $q$ , we have obtained the following boundary value problem

$$\begin{cases} -\Delta q - \gamma \nabla \cdot (q\mathbf{f}) = \gamma \nabla \cdot \mathbf{f}, & \text{in } \Omega, \\ q = 0, & \text{on } \Gamma_w, \\ \partial_n q + \gamma \tilde{F} q = \alpha_0 \gamma g - \gamma \tilde{F}, & \text{on } \Gamma. \end{cases} \tag{4.4}$$

This motivates the following split formulation for problem (1.1):

- (1) Find  $q$  that solves (4.4).
- (2) In view of (4.3), define

$$\tilde{\alpha}(\mathbf{x}) = \frac{\alpha_0}{q(\mathbf{x}) + 1}, \quad \mathbf{x} \in \Omega. \tag{4.5}$$

- (3) Find  $(\mathbf{U}, P)$  that solve

$$\begin{cases} \tilde{\alpha} \mathbf{U} + \nabla P = \mathbf{f}, & \text{in } \Omega, \\ \nabla \cdot \mathbf{U} = 0, & \text{in } \Omega, \\ P = p_w, & \text{on } \Gamma_w, \\ \mathbf{U} \cdot \mathbf{n} = g, & \text{on } \Gamma. \end{cases} \tag{4.6}$$

Summing up, if  $(\mathbf{u}, p)$  is a solution of problem (1.1) and  $p$  belongs to  $L^\infty(\Omega)$ , then  $(q, \mathbf{U}, p)$  solves (4.4)–(4.6). The converse is partially established in the next subsection.

**Remark 4.1.** This formulation requires only the solution of two linear problems.

### 4.2. Analysis of the auxiliary problem

Let us first examine the well-posedness of the boundary value problem (4.4). For this, we write it in a variational form. Multiply the first equation of (4.4) by a sufficiently smooth function  $r$  that vanishes on  $\Gamma_w$ , apply Green’s formula and use the last equation of (4.4). We obtain

$$\int_{\Omega} \nabla q \cdot \nabla r + \gamma \int_{\Omega} q\mathbf{f} \cdot \nabla r = \alpha_0 \gamma \int_{\Gamma} gr - \gamma \int_{\Omega} \mathbf{f} \cdot \nabla r.$$



In the case  $d = 3$ , the minimal smoothness requirements for these integrals to be meaningful are  $q, r \in H^1(\Omega)$ ,  $\mathbf{f} \in L^3(\Omega)^3$ , and  $g \in H_{00}^{1/2}(\Gamma)'$ . Hence, the weak formulation of problem (4.4) that we will consider is the following:

Given  $\mathbf{f} \in L^3(\Omega)^3$  and  $g \in H_{00}^{1/2}(\Gamma)'$ , find  $q \in H_w^1(\Omega)$  such that

$$\int_{\Omega} \nabla q \cdot \nabla r + \gamma \int_{\Omega} q \mathbf{f} \cdot \nabla r = \alpha_0 \gamma \langle g, r \rangle_{\Gamma} - \gamma \int_{\Omega} \mathbf{f} \cdot \nabla r, \quad \forall r \in H_w^1(\Omega). \tag{4.7}$$

A sufficient condition for this problem to be well posed is the following.

**Proposition 4.2.** Assume there exists a constant  $\chi < 1$  such that

$$\gamma C(\Omega) \|\mathbf{f}\|_{L^3(\Omega)^3} \leq \chi < 1. \tag{4.8}$$

Then, problem (4.7) has a unique solution  $q \in H_w^1(\Omega)$ .

*Proof.* Let  $q = r$  in (4.7); Hölder's inequality and (4.8) give

$$\begin{aligned} \left| \gamma \int_{\Omega} q \mathbf{f} \cdot \nabla q \right| &\leq \gamma \|q\|_{L^6(\Omega)} \|\mathbf{f}\|_{L^3(\Omega)^3} \|\nabla q\|_{L^2(\Omega)^3} \\ &\leq \gamma C(\Omega) \|\mathbf{f}\|_{L^3(\Omega)^3} |q|_{H^1(\Omega)}^2 \\ &\leq \chi |q|_{H^1(\Omega)}^2. \end{aligned}$$

Then Lax–Milgram's Lemma implies that problem (4.7) is well-posed. □

**Remark 4.3.** Condition (4.8) is only sufficient for problem (4.4) to be well-posed. We do not want to provide a thorough analysis of this problem, but only to show that there are cases when the algorithm that we are developing is meaningful.

Next, we turn to problem (4.6). This problem is well-posed if  $\tilde{\alpha}$  defined by (4.5) belongs to  $L^\infty(\Omega)$  and is bounded away from zero. For this, it suffices that there exists a constant  $q_0 > 0$  such that

$$q + 1 \geq q_0 > 0, \text{ a.e. in } \Omega \tag{4.9}$$

and

$$q \in L^\infty(\Omega). \tag{4.10}$$

Condition (4.10) can be regarded as a restriction on the smoothness of the data and the domain. Sufficient conditions for assumption (4.9) to hold elude us at the moment, but we have the following partial result, in the simpler case when  $\Gamma_w = \partial\Omega$ .

**Proposition 4.4.** Assume that  $\Gamma_w = \partial\Omega$  and condition (4.8) holds. Then  $q$  satisfies

$$q + 1 \geq 0, \text{ a.e. in } \Omega.$$

*Proof.* Let us define the set

$$\Omega^- = \{\mathbf{x} \in \Omega : q(\mathbf{x}) + 1 \leq 0\},$$

and the function

$$r_0(\mathbf{x}) = \begin{cases} 0, & \mathbf{x} \notin \Omega^-, \\ -(q(\mathbf{x}) + 1), & \mathbf{x} \in \Omega^-. \end{cases}$$

Clearly,  $r_0 \in H^1(\Omega)$  and by definition  $r_0 \geq 0$  almost everywhere in  $\Omega$ . Moreover, since  $q + 1|_{\partial\Omega} = 1 > 0$  then  $r_0 \in H_0^1(\Omega)$ . By setting  $r = r_0$  in (4.7) and changing signs we obtain that

$$\int_{\Omega^-} |\nabla r_0|^2 + \gamma \int_{\Omega^-} r_0 \mathbf{f} \cdot \nabla r_0 = 0. \tag{4.11}$$

Owing to condition (4.8), equality (4.11) implies that

$$(1 - \chi) \int_{\Omega} |\nabla r_0|^2 \leq 0.$$

In other words  $\nabla r_0 = 0$ , a.e. in  $\Omega$ . Since  $r_0 \in H_0^1(\Omega)$ , we have  $r_0 = 0$ , a.e. in  $\Omega$  thus implying the result.  $\square$

Under restrictions (4.9), (4.10) and (4.8), we are able to show that the solution  $(\mathbf{U}, P)$  to (4.6) solves (1.1).

**Proposition 4.5.** *In addition to (4.8), assume that the solution  $q$  to problem (4.4) is in  $L^\infty(\Omega)$  and satisfies (4.9). Then problem (4.6) has a unique solution  $(\mathbf{U}, P)$  and this solution solves (1.1).*

*Proof.* By (4.9), there is a unique  $\tilde{P}$  such that a.e. in  $\Omega$ ,

$$e^{-\gamma\tilde{P}} = q + 1.$$

The assumption that  $q \in L^\infty(\Omega)$  together with (4.9) imply that  $\tilde{P} \in H^1(\Omega)$ . Moreover, since  $q = 0$  on  $\Gamma_w$ , we obtain  $\tilde{P} \in H_w^1(\Omega)$ .

Define  $\tilde{\mathbf{U}} \in L^2(\Omega)^3$  by

$$\alpha_0 \gamma \tilde{\mathbf{U}} := \nabla q + \gamma(q + 1)\mathbf{f};$$

by (4.4), this implies that

$$\nabla \cdot \tilde{\mathbf{U}} = 0.$$

Moreover, by the definition of  $\tilde{P}$ ,

$$\begin{aligned} \alpha_0 \gamma \tilde{\mathbf{U}} &= \nabla(e^{-\gamma\tilde{P}} - 1) + \gamma e^{-\gamma\tilde{P}} \mathbf{f} \\ &= -\gamma e^{-\gamma\tilde{P}} \nabla \tilde{P} + \gamma e^{-\gamma\tilde{P}} \mathbf{f}; \end{aligned}$$

hence

$$\alpha(\tilde{P})\tilde{\mathbf{U}} + \nabla \tilde{P} = \mathbf{f}.$$

The boundary condition on  $\tilde{\mathbf{U}}$  can be obtained in a similar way. This implies not only that the pair  $(\tilde{\mathbf{U}}, \tilde{P})$  solves (1.1), but also that

$$\frac{\alpha_0}{q + 1} \tilde{\mathbf{U}} + \nabla \tilde{P} = \mathbf{f}.$$

Since the solution to (4.6) is unique  $(\tilde{\mathbf{U}}, \tilde{P}) = (\mathbf{U}, P)$ .  $\square$

**Remark 4.6.** In the case of Dirichlet boundary conditions on the whole boundary:  $\Gamma_w = \partial\Omega$ , if we slightly restrict the angles of the domain and assume that  $\mathbf{f}$  is smoother, for instance  $\mathbf{f} \in L^6(\Omega)^3$  and  $\nabla \cdot \mathbf{f} \in L^2(\Omega)$ , then a bootstrap argument, and regularity results for the Laplace equation, show that  $q \in W_r^1(\Omega)$  for some  $r > 3$  and hence  $q$  is continuous. Therefore (4.10) is satisfied.

### 4.3. Discretization

Let us discretize (4.4)–(4.6). In order to approximate the linear Darcy system (4.6) we use the spaces  $X_h$  and  $M_h$  introduced in Section 3 and assume that they satisfy (3.1). We also introduce another finite dimensional space  $W_h \subset H^1_w(\Omega)$  to discretize (4.4). Then, the discrete algorithm is the following:

- (1) Find  $q_h \in W_h$  such that

$$\int_{\Omega} \nabla q_h \cdot \nabla s_h + \gamma \int_{\Omega} q_h \mathbf{f} \cdot \nabla s_h = \alpha_0 \gamma \int_{\Gamma} g s_h - \gamma \int_{\Omega} \mathbf{f} \cdot \nabla s_h, \quad \forall s_h \in W_h. \tag{4.12}$$

- (2) Compute the function

$$\tilde{\alpha}_h(\mathbf{x}) = \frac{\alpha_0}{q_h(\mathbf{x}) + 1}, \quad \mathbf{x} \in \Omega. \tag{4.13}$$

- (3) Find  $(\tilde{\mathbf{u}}_h, \tilde{p}_h) \in X_h \times M_h$  that solve the discrete linear Darcy system

$$\begin{cases} \int_{\Omega} \tilde{\alpha}_h \tilde{\mathbf{u}}_h \cdot \mathbf{v}_h + \int_{\Omega} \mathbf{v}_h \cdot \nabla \tilde{p}_h = \int_{\Omega} \mathbf{f} \cdot \mathbf{v}_h, & \forall \mathbf{v}_h \in X_h, \\ \int_{\Omega} \tilde{\mathbf{u}}_h \cdot \nabla r_h = \langle g, r_h \rangle_{\Gamma}, & \forall r_h \in M_h. \end{cases} \tag{4.14}$$

**Remark 4.7.** Note that finding this approximate solution involves solving only two consecutive linear problems.

**Remark 4.8.** Clearly, under assumption (4.8), problem (4.12) has a unique solution. Then, for the discrete version of the splitting method to make sense we need assumptions analogous to (4.9) and (4.10). When  $W_h$  has the same structure as in (3.6), (4.10) is always satisfied, although the upper bound may not be uniform with respect to  $h$ . Furthermore, if  $q_h(\mathbf{x}) + 1 > 0$  for all  $\mathbf{x}$  in  $\bar{\Omega}$ , then since problem (4.14) is set into finite dimension, it also has a unique solution. But of course, (4.9) is not guaranteed, although in the numerical experiments of Section 5.2, we observe indeed that the discrete solution satisfies  $q_h + 1 > 0$ .

### 4.4. Heuristic error analysis

Now, we present an error analysis of the algorithm (4.12)–(4.14), but this analysis is still heuristic because we must assume that the function  $q_h$  satisfies uniformly assumptions similar to (4.9) and (4.10). More precisely, we suppose that there are constants  $q_{\min}, q_{\max} > 0$  such that for every  $h > 0$ ,

$$0 < q_{\min} \leq q_h(\mathbf{x}) + 1 \leq q_{\max}, \quad \forall \mathbf{x} \in \bar{\Omega}. \tag{4.15}$$

With this, we can proceed in two directions: a straightforward analysis of (4.12)–(4.14), or a comparison with (3.9). In both cases, we suppose that (4.8) holds, so that (4.12) has a unique solution.

Let us proceed first with the second option, namely comparison with (3.9). We do not know whether the nonlinear Darcy problem with exponential porosity has a solution or not; and if so, which are its properties. For this reason, we shall carry this error analysis under the assumption that problem (1.1) with the function  $\alpha$  defined by (1.2) does have a solution. Moreover, we shall assume that the discrete problem defined by (3.9), with  $\alpha$  as in (1.2) has a unique solution for all  $h > 0$ .

**Proposition 4.9.** *In addition to (3.1) and (4.8), assume that the solution  $q_h$  to problem (4.12) satisfies (4.15). If the pair  $(\tilde{\mathbf{u}}_h, \tilde{p}_h) \in X_h \times M_h$  solves (4.14), then there exists a constant  $C > 0$  independent of  $h$  such that*

$$\|\mathbf{u}_h - \tilde{\mathbf{u}}_h\|_{L^2(\Omega)^3} + |p_h - \tilde{p}_h|_{H^1(\Omega)} \leq C \sup_{\mathbf{x} \in \bar{\Omega}} |\alpha(p_h(\mathbf{x})) - \tilde{\alpha}_h(\mathbf{x})| \|\mathbf{u}_h\|_{L^2(\Omega)^3}, \tag{4.16}$$

where  $(\mathbf{u}_h, p_h) \in X_h \times M_h$  solves (3.9).

*Proof.* Let us take the difference of equations (3.9) and (4.14). We obtain

$$\begin{cases} \int_{\Omega} (\alpha(p_h)\mathbf{u}_h - \tilde{\alpha}_h\tilde{\mathbf{u}}_h) \cdot \mathbf{v}_h + \int_{\Omega} \mathbf{v}_h \cdot \nabla(p_h - \tilde{p}_h) = 0, & \forall \mathbf{v}_h \in X_h, \\ \int_{\Omega} (\mathbf{u}_h - \tilde{\mathbf{u}}_h) \cdot \nabla r_h = 0, & \forall r_h \in M_h. \end{cases}$$

Let  $\mathbf{v}_h = \mathbf{u}_h - \tilde{\mathbf{u}}_h$ ; assumption (4.15) implies

$$\begin{aligned} \frac{\alpha_0}{q_{\max}} \|\mathbf{u}_h - \tilde{\mathbf{u}}_h\|_{L^2(\Omega)^3}^2 &\leq \int_{\Omega} \tilde{\alpha}_h |\mathbf{u}_h - \tilde{\mathbf{u}}_h|^2 \\ &= \left| \int_{\Omega} (\alpha(p_h) - \tilde{\alpha}_h) \mathbf{u}_h \cdot (\mathbf{u}_h - \tilde{\mathbf{u}}_h) \right|, \end{aligned}$$

whence

$$\|\mathbf{u}_h - \tilde{\mathbf{u}}_h\|_{L^2(\Omega)^3} \leq C \sup_{\mathbf{x} \in \bar{\Omega}} |\alpha(p_h(\mathbf{x})) - \tilde{\alpha}_h(\mathbf{x})| \|\mathbf{u}_h\|_{L^2(\Omega)^3}.$$

By the inf-sup condition (3.1)

$$\begin{aligned} \beta |p_h - \tilde{p}_h|_{H^1(\Omega)} &\leq \sup_{\mathbf{v}_h \in X_h} \frac{\int_{\Omega} (\alpha(p_h)\mathbf{u}_h - \tilde{\alpha}_h\tilde{\mathbf{u}}_h) \cdot \mathbf{v}_h}{\|\mathbf{v}_h\|_{L^2(\Omega)^3}} \\ &= \sup_{\mathbf{v}_h \in X_h} \frac{\int_{\Omega} \tilde{\alpha}_h (\mathbf{u}_h - \tilde{\mathbf{u}}_h) \cdot \mathbf{v}_h + \int_{\Omega} (\alpha(p_h) - \tilde{\alpha}_h) \mathbf{u}_h \cdot \mathbf{v}_h}{\|\mathbf{v}_h\|_{L^2(\Omega)^3}} \\ &\leq C \|\mathbf{u}_h - \tilde{\mathbf{u}}_h\|_{L^2(\Omega)^3} + \sup_{\mathbf{x} \in \bar{\Omega}} |\alpha(p_h(\mathbf{x})) - \tilde{\alpha}_h(\mathbf{x})| \|\mathbf{u}_h\|_{L^2(\Omega)^3} \\ &\leq C \sup_{\mathbf{x} \in \bar{\Omega}} |\alpha(p_h(\mathbf{x})) - \tilde{\alpha}_h(\mathbf{x})| \|\mathbf{u}_h\|_{L^2(\Omega)^3}. \end{aligned} \quad \square$$

This estimate should be regarded as the basic one. If the exact solution is smooth enough, it can easily be reduced, for instance, to max-norm error estimates for the pressure  $p$  and the auxiliary variable  $q$ .

**Corollary 4.10.** *In addition to (3.1) and (4.8), assume that the solution  $q$  to (4.7) belongs to  $L^\infty(\Omega)$  and satisfies (4.9). Assume, also, that the pair  $(\mathbf{u}, p)$  that solves (1.1) is such that  $p \in L^\infty(\Omega)$ . If  $q_h$  satisfies (4.15) then there is a constant  $C > 0$  independent of  $h$  such that*

$$\|\mathbf{u}_h - \tilde{\mathbf{u}}_h\|_{L^2(\Omega)^3} + |p_h - \tilde{p}_h|_{H^1(\Omega)} \leq C \left( \|p - p_h\|_{L^\infty(\Omega)} + \|q - q_h\|_{L^\infty(\Omega)} \right) \|\mathbf{u}_h\|_{L^2(\Omega)^3}. \quad (4.17)$$

*Proof.* Using (4.16) it is sufficient to bound the  $L^\infty$  norm of the difference  $\alpha(p_h) - \tilde{\alpha}_h$ . Then

$$\begin{aligned} \|\alpha(p_h) - \tilde{\alpha}_h\|_{L^\infty(\Omega)} &\leq \|\alpha(p) - \alpha(p_h)\|_{L^\infty(\Omega)} + \|\alpha(p) - \tilde{\alpha}_h\|_{L^\infty(\Omega)} \\ &\leq D \|p - p_h\|_{L^\infty(\Omega)} + \|\alpha(p) - \tilde{\alpha}_h\|_{L^\infty(\Omega)}, \end{aligned}$$

where the constant  $D$  satisfies

$$D \leq \alpha_0 \gamma \exp \left( \gamma \max \left\{ \|p\|_{L^\infty(\Omega)}, \|p_h\|_{L^\infty(\Omega)} \right\} \right).$$

Comparing (4.3) and (4.13), we obtain for a.e.  $\mathbf{x}$  in  $\Omega$

$$\begin{aligned} |\alpha(p(\mathbf{x})) - \tilde{\alpha}_h(\mathbf{x})| &\leq \alpha_0 \frac{|q_h(\mathbf{x}) - q(\mathbf{x})|}{|(q(\mathbf{x}) + 1)(q_h(\mathbf{x}) + 1)|} \\ &\leq \frac{\alpha_0}{|(q(\mathbf{x}) + 1)(q_h(\mathbf{x}) + 1)|} \|q_h - q\|_{L^\infty(\Omega)}. \end{aligned}$$

Assumptions (4.9) and (4.15) imply that there is a constant  $C > 0$  independent of  $h$  such that

$$|(q(\mathbf{x}) + 1)(q_h(\mathbf{x}) + 1)| > C \text{ for a.e. } \mathbf{x} \in \Omega,$$

whence (4.17). □

Finally, to be able to provide an order of convergence, we must assume one additional approximation property of the space  $M_h$ , and we must assume that the space  $W_h$  has adequate approximation properties. More precisely,

- (1) There is a constant  $C > 0$ , independent of  $h$ , such that for every  $r \in W_\infty^k(\Omega)$  the interpolation operator  $\mathcal{I}_h$  defined in (3.3) satisfies

$$\|r - \mathcal{I}_h r\|_{L^\infty(\Omega)} \leq Ch^k |r|_{W_\infty^k(\Omega)}. \tag{4.18}$$

- (2) There exists an interpolation operator  $\rho_h : H^1(\Omega) \rightarrow W_h$ , such that for all  $1 \leq s \leq \infty$ , if  $r \in W_s^{k+1}(\Omega)$

$$\|r - \rho_h r\|_{L^s(\Omega)} + h |r - \rho_h r|_{W_s^1(\Omega)} \leq Ch^{k+1} |r|_{W_s^{k+1}(\Omega)}, \tag{4.19}$$

where the constant  $C > 0$  does not depend on  $r$  or  $h$ .

- (3) There is a constant  $C > 0$  independent of  $h$ , such that for every  $r_h \in W_h$  the following *inverse inequality* holds

$$\|r_h\|_{L^\infty(\Omega)} \leq Ch^{-1/2} |r_h|_{H^1(\Omega)}. \tag{4.20}$$

**Remark 4.11.** The space  $M_h$  defined in (3.6) has properties (4.18) and (4.19) with the same interpolation operator  $\mathcal{I}_h$ . Hence, the triple  $(X_h, M_h, M_h)$  with  $X_h$  defined in (3.5) and  $M_h$  defined in (3.6) has all the desired properties for all  $k \geq 1$ .

Under these assumptions, we first bound the error of the auxiliary problem.

**Proposition 4.12.** *If (4.8) holds, the solution  $q_h$  of (4.12) satisfies*

$$|q - q_h|_{H^1(\Omega)} \leq 2 \left( 1 + \frac{\gamma C(\Omega)}{1 - \chi} \|\mathbf{f}\|_{L^3(\Omega)^3} \right) \inf_{r_h \in W_h} |q - r_h|_{H^1(\Omega)}. \tag{4.21}$$

*Proof.* By taking the difference between (4.12) and (4.7), inserting any function  $r_h$  in  $W_h$  and testing with  $s_h = q_h - r_h$ , we obtain

$$|q_h - r_h|_{H^1(\Omega)} (1 - \gamma C(\Omega) \|\mathbf{f}\|_{L^3(\Omega)^3}) \leq |q - r_h|_{H^1(\Omega)} (1 + \gamma C(\Omega) \|\mathbf{f}\|_{L^3(\Omega)^3}).$$

By virtue of (4.8), this implies that

$$|q_h - r_h|_{H^1(\Omega)} \leq \left( 1 + 2 \frac{\gamma C(\Omega) \|\mathbf{f}\|_{L^3(\Omega)^3}}{1 - \gamma C(\Omega) \|\mathbf{f}\|_{L^3(\Omega)^3}} \right) |q - r_h|_{H^1(\Omega)}.$$

Then (4.21) follows from (4.8) and the triangle inequality. □

Now we are able to prove a convergence result.

**Corollary 4.13.** *In addition to (4.8), assume that the solution  $q$  to problem (4.4) belongs to  $H^{k+1}(\Omega) \cap W_\infty^k(\Omega)$  and satisfies (4.9). Moreover, assume that the solution  $(\mathbf{u}, p)$  to (1.1) is such that  $p \in H^{k+1}(\Omega) \cap W_\infty^k(\Omega)$ . Then, if the space  $M_h$  satisfies (3.3), (3.19) and (4.18), and the space  $W_h$  satisfies (4.19) and (4.20), and if  $q_h$  satisfies (4.15), there exists a constant  $C > 0$  that does not depend on  $h$ , such that*

$$\|\mathbf{u}_h - \tilde{\mathbf{u}}_h\|_{L^2(\Omega)^3} + |p_h - \tilde{p}_h|_{H^1(\Omega)} \leq Ch^{k-1/2} \left( |p|_{W_\infty^k(\Omega)} + |p|_{H^{k+1}(\Omega)} + |q|_{W_\infty^k(\Omega)} + |q|_{H^{k+1}(\Omega)} \right) \|\mathbf{u}_h\|_{L^2(\Omega)^3}.$$

*Proof.* By property (4.18),

$$\begin{aligned} \|p - p_h\|_{L^\infty(\Omega)} &\leq \|p - \mathcal{I}_h p\|_{L^\infty(\Omega)} + \|\mathcal{I}_h p - p_h\|_{L^\infty(\Omega)} \\ &\leq Ch^k |p|_{W_\infty^k(\Omega)} + \|\mathcal{I}_h p - p_h\|_{L^\infty(\Omega)}. \end{aligned}$$

By the inverse inequality (3.19) and by (3.3)

$$\begin{aligned} \|\mathcal{I}_h p - p_h\|_{L^\infty(\Omega)} &\leq Ch^{-1/2} |\mathcal{I}_h p - p_h|_{H^1(\Omega)} \\ &\leq Ch^{-1/2} \left( |p - \mathcal{I}_h p|_{H^1(\Omega)} + |p - p_h|_{H^1(\Omega)} \right) \\ &\leq Ch^{-1/2} \left( h^k |p|_{H^{k+1}(\Omega)} + |p - p_h|_{H^1(\Omega)} \right). \end{aligned}$$

To estimate the term  $|p - p_h|_{H^1(\Omega)}$  it is sufficient to recall Corollary 3.3 in the uniqueness case, or (3.31) for nonsingular solutions (with  $s = k + 1$ ). We obtain

$$\|p - p_h\|_{L^\infty(\Omega)} \leq Ch^k |p|_{W_\infty^k(\Omega)} + Ch^{k-1/2} |p|_{H^{k+1}(\Omega)}.$$

Then we conclude the proof by applying (4.21) and the inverse inequality (4.20). □

**Remark 4.14.** The above estimates are suboptimal, but they show heuristically that the splitting algorithm does indeed converge. By using a more refined analysis, for instance the method of weighted norms of Nitsche (see [12], Brenner and Scott [8], Chap. 8, or Girault *et al.* [22], for more details) we may derive (again heuristically) optimal error estimates. The results of Section 5.2 give examples where the errors have indeed optimal order.

**Remark 4.15.** If  $q$  belongs to  $H^2(\Omega) \cap W_\infty^1(\Omega)$  and satisfies (4.9), then for all sufficiently small  $h$ ,  $q_h$  also satisfies (4.15).

Now, let us estimate the error of (4.12)–(4.14) without reverting to (3.9).

The estimate (4.21) is rigorous because it is derived solely under assumptions on the data. However, the remaining estimates are heuristic because we do not know how to estimate the error on  $\tilde{\mathbf{u}}_h$  without assuming that  $q_h$  satisfies (4.15) and  $q$  satisfies (4.9) and (4.10). Then we have the following result.

**Theorem 4.16.** *In addition to (3.1) and (4.8), suppose that the solution  $q$  to (4.7) satisfies (4.9) and (4.10), the solution  $\mathbf{U}$  of (4.6) belongs to  $L^3(\Omega)^3$ , and the solution  $q_h$  of (4.12) satisfies (4.15). Then*

$$\begin{aligned} \|\mathbf{U} - \tilde{\mathbf{u}}_h\|_{L^2(\Omega)^3} &\leq \left(1 + \frac{q_{\max}}{q_{\min}}\right) \left(1 + \frac{1}{\beta}\right) \inf_{\mathbf{v}_h \in X_h} \|\mathbf{U} - \mathbf{v}_h\|_{L^2(\Omega)^3} + \frac{q_{\max}}{q_{\min}} \frac{1}{q_0} C(\Omega) \|\mathbf{U}\|_{L^3(\Omega)^3} |q - q_h|_{H^1(\Omega)} \\ &\quad + \frac{q_{\max}}{\alpha_0} \inf_{r_h \in M_h} |P - r_h|_{H^1(\Omega)}, \end{aligned} \tag{4.22}$$

and

$$|P - \tilde{p}_h|_{H^1(\Omega)} \leq \left(1 + \frac{1}{\beta}\right) \inf_{r_h \in M_h} |P - r_h|_{H^1(\Omega)} + \frac{1}{\beta} \frac{\alpha_0}{q_{\min}} \left( \|\mathbf{U} - \tilde{\mathbf{u}}_h\|_{L^2(\Omega)^3} + \frac{C(\Omega)}{q_0} \|\mathbf{U}\|_{L^3(\Omega)^3} |q - q_h|_{H^1(\Omega)} \right). \tag{4.23}$$

*Proof.* First, the assumptions on  $q$  and  $q_h$  imply that  $\tilde{\alpha}$  and  $\tilde{\alpha}_h$  are well-defined and strictly positive. Next, by taking the difference between the first row of (4.14) and (4.6) in weak form, and inserting any element  $\mathbf{v}_h$  of  $X_h$  and  $r_h$  of  $M_h$ , we obtain for any  $\mathbf{w}_h$  in  $X_h$ ,

$$\int_\Omega \tilde{\alpha}_h(\tilde{\mathbf{u}}_h - \mathbf{v}_h) \cdot \mathbf{w}_h + \int_\Omega (\tilde{\alpha}_h - \tilde{\alpha}) \mathbf{U} \cdot \mathbf{w}_h + \int_\Omega \nabla(\tilde{p}_h - r_h) \cdot \mathbf{w}_h = \int_\Omega \tilde{\alpha}_h(\mathbf{U} - \mathbf{v}_h) \cdot \mathbf{w}_h + \int_\Omega \nabla(P - r_h) \cdot \mathbf{w}_h.$$

In order to eliminate  $\tilde{p}_h$ , we proceed as in Theorem 3.1: owing to (3.1), there exists  $\mathbf{v}_h$  in  $X_h$  such that  $\mathbf{w}_h := \tilde{\mathbf{u}}_h - \mathbf{v}_h$  belongs to  $V_h$  (see (3.7)), and

$$\|\mathbf{U} - \mathbf{v}_h\|_{L^2(\Omega)^3} \leq \left(1 + \frac{1}{\beta}\right) \inf_{\mathbf{v}_h \in X_h} \|\mathbf{U} - \mathbf{v}_h\|_{L^2(\Omega)^3}. \tag{4.24}$$

This choice of test function eliminates the last term in the left-hand side of the above difference. Then by applying (4.15), we derive

$$\|\tilde{\mathbf{u}}_h - \mathbf{v}_h\|_{L^2(\Omega)^3} \leq \frac{q_{\max}}{q_{\min}} \|\mathbf{U} - \mathbf{v}_h\|_{L^2(\Omega)^3} + \frac{q_{\max}}{\alpha_0} \|\tilde{\alpha}_h - \tilde{\alpha}\|_{L^6(\Omega)} \|\mathbf{U}\|_{L^3(\Omega)^3} + \frac{q_{\max}}{\alpha_0} |P - r_h|_{H^1(\Omega)}. \tag{4.25}$$

There remains to estimate  $\tilde{\alpha}_h - \tilde{\alpha}$ :

$$\|\tilde{\alpha}_h - \tilde{\alpha}\|_{L^6(\Omega)} \leq \frac{\alpha_0}{q_0 q_{\min}} C(\Omega) |q - q_h|_{H^1(\Omega)}. \tag{4.26}$$

Then (4.22) follows by substituting this bound into (4.25) and using (4.24) and the triangle inequality.

To obtain (4.23) notice that, by the discrete inf-sup condition (3.1), for any  $r_h \in M_h$

$$\beta |\tilde{p}_h - r_h|_{H^1(\Omega)} \leq \sup_{\mathbf{y}_h \in X_h} \frac{b(\mathbf{y}_h, \tilde{p}_h - r_h)}{\|\mathbf{y}_h\|_{L^2(\Omega)^3}} \leq |P - r_h|_{H^1(\Omega)} + \sup_{\mathbf{y}_h \in X_h} \frac{b(\mathbf{y}_h, P - \tilde{p}_h)}{\|\mathbf{y}_h\|_{L^2(\Omega)^3}},$$

which shows that it is sufficient to estimate  $b(\mathbf{y}_h, P - \tilde{p}_h)$ . By taking the difference of the first equation in (4.6) in weak form and the first equation of (4.14) we obtain

$$\begin{aligned} b(\mathbf{y}_h, P - \tilde{p}_h) &= \int_{\Omega} (\tilde{\alpha}_h \tilde{\mathbf{u}}_h - \tilde{\alpha} \mathbf{U}) \cdot \mathbf{y}_h = \int_{\Omega} \tilde{\alpha}_h (\tilde{\mathbf{u}}_h - \mathbf{U}) \cdot \mathbf{y}_h - \int_{\Omega} (\tilde{\alpha} - \tilde{\alpha}_h) \mathbf{U} \cdot \mathbf{y}_h \\ &\leq \|\tilde{\alpha}_h\|_{L^\infty(\Omega)} \|\mathbf{U} - \tilde{\mathbf{u}}_h\|_{L^2(\Omega)^3} \|\mathbf{y}_h\|_{L^2(\Omega)^3} + \|\tilde{\alpha} - \tilde{\alpha}_h\|_{L^6(\Omega)} \|\mathbf{U}\|_{L^3(\Omega)^3} \|\mathbf{y}_h\|_{L^2(\Omega)^3}, \end{aligned}$$

which, by (4.26) and (3.1) implies

$$|\tilde{p}_h - r_h|_{H^1(\Omega)} \leq \frac{1}{\beta} \left( |P - r_h|_{H^1(\Omega)} + \frac{\alpha_0}{q_{\min}} \left( \|\tilde{\mathbf{u}}_h - \mathbf{U}\|_{L^2(\Omega)^3} + \frac{C(\Omega)}{q_0} \|\mathbf{U}\|_{L^3(\Omega)^3} |q - q_h|_{H^1(\Omega)} \right) \right). \tag{4.27}$$

The error estimate (4.23) follows from (4.27) and the triangle inequality. □

**Remark 4.17.** Proposition 4.12 and Theorem 4.16 immediately yield straightforward orders of convergence for  $(\tilde{\mathbf{u}}_h, \tilde{p}_h)$ . We skip them for the sake of brevity.

### 5. NUMERICAL EXPERIMENTS

To illustrate the theory of the previous sections, we present a series of numerical experiments, in two and three dimensions, which show the performance of the developed methods in a series of test cases.

The numerical experiments in two dimensions were conducted using the package `FreeFem++` (see [24]). In this case, unless otherwise stated, the domain is  $\Omega = ]0, 1]^2$ , where the top and right sides are  $\Gamma_w$  and the other two sides are  $\Gamma$ .

The numerical experiments in three dimensions were carried out with the help of the `deal.II` library (see [5]). For the experiments in this dimension, the domain is  $\Omega = ]0, 1]^3$ , with  $\Gamma_w = \{(x, y, z) \in \partial\Omega : x = 1\} \cup \{(x, y, z) \in \partial\Omega : y = 1\}$  and  $\Gamma = \partial\Omega \setminus \bar{\Gamma}_w$ .

TABLE 1. 3-D. Iterative algorithm. Small porosity.  $\mathbb{Q}_1dc$ -velocity,  $\mathbb{Q}_1$ -pressure.

Level	$h$	$\ \mathbf{u} - \mathbf{u}_h\ _{L^2(\Omega)}$	Order	$ p - p_h _{H^1(\Omega)}$	Order	Iterations
1	0.5000	1.63 E+000	—	3.25 E+000	—	5
2	0.2500	9.35 E-001	0.80	1.72 E+000	0.92	9
3	0.1250	4.97 E-001	0.91	8.66 E-001	0.99	8
4	0.0625	2.53 E-001	0.97	4.35 E-001	0.99	8
5	0.0313	1.27 E-001	0.99	2.18 E-001	1.00	8

### 5.1. The uniqueness case

To test the algorithm developed in Section 3.1 we have conducted a series of numerical experiments, the results of which we present below. We always initiate the iterative process (3.14) with  $p_h^0 = 0$  and use the stopping criterion

$$\frac{\sqrt{\left\| \mathbf{u}_h^{(n+1)} - \mathbf{u}_h^{(n)} \right\|_{L^2(\Omega)^d}^2 + \left| p_h^{(n+1)} - p_h^{(n)} \right|_{H^1(\Omega)}^2}}{\sqrt{\left\| \mathbf{u}_h^{(n+1)} \right\|_{L^2(\Omega)^d}^2 + \left| p_h^{(n+1)} \right|_{H^1(\Omega)}^2}} < 10^{-10}.$$

#### 5.1.1. Small data

To test the algorithm in the case when the porosity does not have high variations, we define the porosity as

$$\alpha(\xi) = 1 + \frac{1}{1 + \xi^2}, \quad \xi \in \mathbb{R}.$$

Notice that  $1 \leq \alpha(\xi) \leq 2$ . We define the exact solution as

$$\mathbf{u}(x, y) = (-y^2, z^2, x^2)^\top, \quad p(x, y) = \sin(2\pi x) \sin(2\pi y) \sin(2\pi z).$$

These functions determine the right-hand side and boundary data.

The results of the algorithm obtained using a discontinuous- $\mathbb{Q}_1$  approximation of the velocity and a  $\mathbb{Q}_1$  approximation of the pressure are reported in Table 1. We see that the number of iterations does not depend on the discretization parameter, and the errors on the velocity and pressure have optimal order. We obtained similar results in two dimensions, using spaces  $\mathbb{P}_0$ - $\mathbb{P}_1$  and  $\mathbb{P}_1dc$ - $\mathbb{P}_2$ . For the sake of brevity, we do not present them here.

Notice that for the last level the number of cells equals 32 768 and

$$\dim X_h = 786\,432 \quad \dim M_h = 35\,937.$$

#### 5.1.2. Large data

To illustrate the case when the porosity has high variations, but is still bounded we consider

$$\alpha(\xi) = 1 + \frac{10}{1 + \xi^2}.$$

Notice that  $1 \leq \alpha(\xi) \leq 11$ . We define the exact solution to be

$$\mathbf{u}(x, y) = (-y^2, x^2)^\top, \quad p(x, y) = 10 \sin(2\pi x) \sin(2\pi y).$$

These functions determine the right-hand side and boundary data.



TABLE 2. 2-D. Iterative algorithm. Big porosity.  $\mathbb{P}_1dc$ -velocity,  $\mathbb{P}_2$ -pressure.

$h$	$\ \mathbf{u} - \mathbf{u}_h\ _{L^2(\Omega)}$	Order	$ p - p_h _{H^1(\Omega)}$	Order	Iterations
0.250000	2.07 E+000	—	9.27 E+000	—	14
0.125000	8.57 E-001	1.33	2.64 E+000	1.43	10
0.062500	2.66 E-001	1.27	6.76 E-001	1.81	9
0.031250	7.11 E-002	1.69	1.69 E-001	1.96	9
0.015625	1.81 E-002	1.90	4.22 E-002	2.00	10

TABLE 3. 3-D. Iterative algorithm. Exponential porosity.  $\mathbb{Q}_1dc$ -velocity,  $\mathbb{Q}_1$ -pressure.

Level	$h$	$\ \mathbf{u} - \mathbf{u}_h\ _{L^2(\Omega)}$	Order	$ p - p_h _{H^1(\Omega)}$	Order	Iterations
1	0.5000	3.26 E+000	—	3.25 E+000	—	8
2	0.2500	1.73 E+000	0.91	1.72 E+000	0.92	8
3	0.1250	8.93 E-001	0.96	8.68 E-001	0.98	7
4	0.0625	4.61 E-001	0.95	4.39 E-001	0.98	7
5	0.0313	2.50 E-001	0.88	2.25 E-001	0.96	7

The results of the algorithm obtained with a discontinuous- $\mathbb{P}_1$  approximation of the velocity and a  $\mathbb{P}_2$  approximation of the pressure are reported in Table 2. We see that the number of iterations does not depend on the discretization parameter, and the errors on the velocity and pressure have optimal order. Using lower order elements, *i.e.* a  $\mathbb{P}_0$ - $\mathbb{P}_1$  approximation, we obtain the same results.

5.1.3. Exponential porosity

Finally, although the theory developed for algorithm (3.14) does not cover the case of an unbounded (*i.e.* exponential) porosity, we nevertheless test this case. We set the porosity to be defined as in (1.2) with

$$\alpha_0 = 1, \quad \gamma = 1/4,$$

and the exact solution

$$\mathbf{u}(x, y) = \frac{1}{2}(-y^2, z^2, x^2)^\top, \quad p(x, y) = 2 + \sin(2\pi x) \sin(2\pi y) \sin(2\pi z).$$

These functions determine the right-hand side and boundary data.

The results of the algorithm obtained using a discontinuous- $\mathbb{Q}_1$  approximation of the velocity and a  $\mathbb{Q}_1$  approximation of the pressure are reported in Table 3. We see that the number of iterations does not depend on the discretization parameter, and the errors on the velocity and pressure have optimal order. In two dimensions, and on a similar problem, we obtain similar results using  $\mathbb{P}_0$ - $\mathbb{P}_1$  and  $\mathbb{P}_1dc$ - $\mathbb{P}_2$  approximations.

5.2. Splitting method

To test the algorithm developed in Section 4, let

$$\alpha_0 = 1, \quad \gamma = 1/4.$$

We define the exact solution to be

$$\mathbf{u}(x, y) = \frac{1}{2}(-y^2, z^2, x^2)^\top, \quad p(x, y) = 2 + \sin(2\pi x) \sin(2\pi y) \sin(2\pi z).$$

Notice that this is the same problem we solved in Section 5.1.3 using the iterative algorithm. The following triple of finite element spaces was used:  $X_h$ -discontinuous- $\mathbb{Q}_1$ ,  $M_h$ - $\mathbb{Q}_1$  and  $W_h$ - $\mathbb{Q}_1$ . The obtained results can be

TABLE 4. 3-D. Splitting algorithm.  $\mathbb{Q}_1dc$ -velocity space,  $\mathbb{Q}_1$ -pressure space,  $\mathbb{Q}_1$ -auxiliary variable.

Level	$h$	$\ \mathbf{u} - \mathbf{u}_h\ _{L^2(\Omega)}$	Order	$ p - p_h _{H^1(\Omega)}$	Order
1	0.5000	5.25 E+000	—	3.25 E+000	—
2	0.2500	2.80 E+000	0.91	1.72 E+000	0.92
3	0.1250	1.45 E+000	0.95	8.70 E-001	0.98
4	0.0625	7.73 E-001	0.91	4.44 E-001	0.97
5	0.0313	3.95 E-001	0.97	2.35 E-001	0.92

TABLE 5. 2-D. Computational time (s). Exponential porosity.

$h$	Iterative		Splitting			
	$(\mathbb{P}_0, \mathbb{P}_1)$	$(\mathbb{P}_1dc, \mathbb{P}_2)$	$(\mathbb{P}_0, \mathbb{P}_1, \mathbb{P}_1)$	$(\mathbb{P}_0, \mathbb{P}_1, \mathbb{P}_2)$	$(\mathbb{P}_1dc, \mathbb{P}_2, \mathbb{P}_1)$	$(\mathbb{P}_1dc, \mathbb{P}_2, \mathbb{P}_2)$
0.500000	0.21	0.74	0.02	0.04	0.04	0.06
0.250000	0.40	1.13	0.08	0.09	0.10	0.13
0.125000	1.20	3.35	0.23	0.27	0.53	0.59
0.062500	4.71	23.16	0.95	1.08	5.15	5.25
0.031250	23.69	248.62	5.81	7.00	69.87	82.07
0.015625	167.36	3341.34	50.64	65.48	1366.66	1702.59
0.007813	1711.00	—	713.58	894.86	—	—

seen in Table 4. The errors  $\|\mathbf{u} - \mathbf{u}_h\|_{L^2(\Omega)^3}$  and  $|p - p_h|_{H^1(\Omega)}$  asymptotically have optimal order. Testing the method on a similar two-dimensional problem, we can draw the same conclusions for the triples  $(\mathbb{P}_0, \mathbb{P}_1, \mathbb{P}_1)$ ,  $(\mathbb{P}_0, \mathbb{P}_1, \mathbb{P}_2)$ ,  $(\mathbb{P}_1dc, \mathbb{P}_2, \mathbb{P}_1)$  and  $(\mathbb{P}_1dc, \mathbb{P}_2, \mathbb{P}_2)$ .

### 5.3. Computational time

In order to estimate the computational complexity of the proposed algorithms, we compare the computational time involved in solving the following two dimensional problem:

$$\alpha(\xi) = e^{\xi/2},$$

$$\mathbf{u} = (-y^3, x^3)^\top, \quad p(x, y) = 2 + \sin(2\pi x) \sin(2\pi y).$$

We compare the iterative algorithm (3.14) and the splitting method of Section 4. The obtained results are shown in Table 5.

From the results shown in this Table we can clearly see that the splitting algorithm of Section 4 outperforms the iterative algorithm of Section 3.1. This is expected to be the case, since the splitting algorithm requires solving only two linear problems as opposed to the iterative algorithm; which although converges independently of the discretization parameter, requires the assembly and solution of a linear problem at each iterative step.

Finally, when comparing the computational times for the splitting algorithm using a fixed velocity-pressure pair but different approximation spaces for the auxiliary problem, we see that the computational times differ very little, their relative difference is never greater than 20%. This suggests that the most time consuming procedure is solving the linear Darcy problem (4.14). This is in agreement with the theory, as this problem has more unknowns and its matrix is indefinite. A better approach for the solution of this problem may reduce the time involved in solving this problem (see the work of Schöberl and Zulehner [33] and Zulehner [36] for instance).

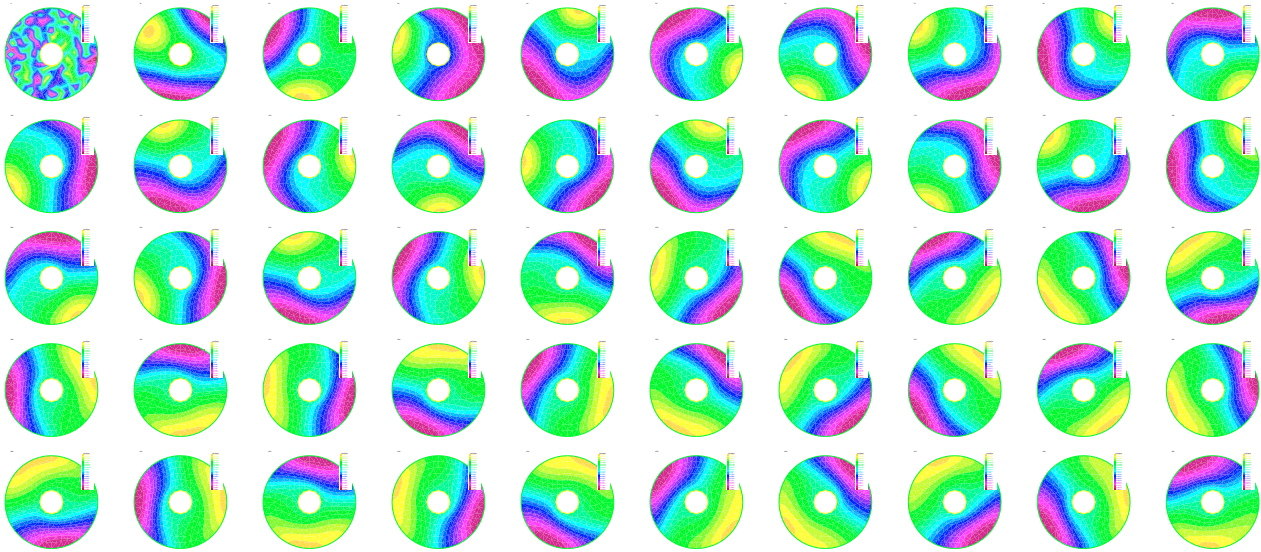


FIGURE 1. Approximate pressure for the iterative algorithm. Shown every ten (10) iterations.

**5.4. Numerical investigation of the convergence condition for the iterative algorithm**

In order to further investigate the properties of the iterative algorithm (3.14) and, more precisely, the role of condition (3.15) we solve the following particular problem in the domain

$$\Omega = \left\{ (x, y) \in \mathbb{R}^2 : 1 < \sqrt{x^2 + y^2} < 4 \right\},$$

with

$$\Gamma_w = \left\{ (x, y) \in \mathbb{R}^2 : \sqrt{x^2 + y^2} = 1 \right\},$$

and  $\Gamma = \partial\Omega \setminus \Gamma_w$ . In this domain we solve the nonlinear Darcy equations with exponential porosity. We set the right-hand side that corresponds to the exact solution

$$\begin{aligned} \mathbf{u}(x, y) &= (xr, -yr)^\top, \\ p(x, y) &= r, \end{aligned}$$

where  $r = \sqrt{x^2 + y^2}$ . In the numerical experiments that follow we use a  $(\mathbb{P}_0, \mathbb{P}_1, \mathbb{P}_1)$  approximation of the velocity-pressure-auxiliary variable. We set  $\alpha_0 = 2$  and vary the parameter  $\gamma$ . Experimentally we have obtained that if  $\gamma < 0.038$  the iterative algorithm converges independently of the initial guess, and it behaves the same way as the cases covered in Section 5.1.

For bigger values of the parameter  $\gamma$ , the splitting algorithm of Section 4 performs as before. However, the iterative algorithm does not converge anymore. Moreover, if we truncate the porosity function  $\alpha$  setting, for instance,

$$\alpha(\xi) = \begin{cases} \alpha_0, & \xi < 0, \\ \alpha_0 e^{\gamma \xi}, & 0 \leq \xi \leq 4.5, \\ \alpha_0 e^{4.5\gamma}, & \xi > 4.5, \end{cases}$$

where the choice of truncation is dictated by  $1 \leq p(x, y) \leq 4 \forall (x, y) \in \bar{\Omega}$ , the method still diverges. For  $\gamma = 0.2$ , a history of the behaviour of the approximate pressure is shown in Figure 1.

From Figure 1 we can see that although the approximate solution diverges, it does remain bounded, and it seems to be oscillating around more than one fixed functions. A detailed analysis of the reasons behind these phenomena is a topic for future research.

## 5.5. Perspectives

These numerical experiments confirm that the splitting formulation is very promising and is worth further investigation. Although it is doubtful that the exact problem (4.7) with arbitrary mixed boundary conditions satisfies the maximum principle, the maximum principle may be valid in particular geometrical configurations such as the one considered in the example of Section 5.4.

Proving that the discrete problem (4.12) satisfies the maximum principle is much more delicate. In the case of the Laplace equation, it is well-known that the maximum principle holds for finite elements of degree one on tetrahedra with acute angles. Extension of this result to (4.12) will be the object of future work.

*Acknowledgements.* The third author is partially supported by Award No. KUS-C1-016-04, made by King Abdullah University of Science and Technology (KAUST). Part of this work was done while the third author was visiting the Laboratoire Jacques-Louis Lions under the *Study Abroad Non-Degree Reciprocal Educational Exchange Program* between TAMU and UPMC. His stay was financed by the Master of the Mathematics Department of the Université Pierre et Marie Curie (Paris VI). The authors would like to thank Prof. K.R. Rajagopal for proposing this model and suggesting to work on it.

## REFERENCES

- [1] R.A. Adams, *Sobolev spaces*. Academic Press (1975).
- [2] G. Allaire, Homogenization of the Navier-Stokes equations with slip boundary conditions. *Comm. Pure Appl. Math.* **44** (1991) 605–641.
- [3] M. Azaïez, F. Ben Belgacem, C. Bernardi and N. Chorfi, *Spectral discretization of Darcy's equations with pressure dependent porosity*. Report 2009-10, Laboratoire Jacques-Louis Lions, France (2009).
- [4] I. Babuška, The finite element method with Lagrangian multipliers. *Numer. Math.* **20** (1973) 179–192.
- [5] W. Bangerth, R. Hartman and G. Kanschat, deal.II – a general-purpose object-oriented finite element library. *ACM Trans. Math. Softw.* **33** (2007) 24.
- [6] J. Berg and J. Löfström, *Interpolation spaces: An introduction, Comprehensive Studies in Mathematics* **223**. Springer-Verlag (1976).
- [7] D. Boffi, F. Brezzi, L. Demkowicz, R. Durán, R. Falk and M. Fortin, *Mixed finite elements, compatibility conditions, and applications, Lecture Notes in Mathematics* **939**. Springer-Verlag, Berlin, Germany (2008).
- [8] S.C. Brenner and L.R. Scott, *The mathematical theory of finite element methods, Texts in applied mathematics* **15**. Third edition, Springer-Verlag (2008).
- [9] F. Brezzi, On the existence, uniqueness and approximation of saddle-point problems arising from Lagrange multipliers. *RAIRO Anal. Numér.* **R2** (1974) 129–151.
- [10] F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods, Springer Series in Computational Mathematics*. Springer-Verlag, New York (1991).
- [11] F. Brezzi, J. Rappaz and P.-A. Raviart, Finite dimensional approximation of nonlinear problems. Part I: Branches of nonsingular solutions. *Numer. Math.* **36** (1980) 1–25.
- [12] P.-G. Ciarlet, Basic error estimates for elliptic problems, in *Handbook of Numerical Analysis II, Finite Element Methods (Part 1)*, P.G. Ciarlet and J.L. Lions Eds., Amsterdam, North-Holland (1991) 17–351.
- [13] D. Cioranescu, P. Donato and H.I. Ene, Homogenization of the Stokes problem with non-homogeneous boundary conditions. *Math. Appl. Sci.* **19** (1996) 857–881.
- [14] H. Darcy, *Les fontaines publiques de la ville de Dijon*. Victor Dalmont, Paris, France (1856).
- [15] J. Douglas and T. Dupont, A Galerkin method for a nonlinear Dirichlet problem. *Math. Comp.* **29** (1975) 689–696.
- [16] H.I. Ene and E. Sanchez-Palencia, Équations et phénomènes de surface pour l'écoulement dans un modèle de milieu poreux. *J. Mécanique* **14** (1975) 73–108.
- [17] A. Ern and J.-L. Guermond, *Theory and practice of finite elements, Applied Mathematical Sciences* **159**. Springer-Verlag, New York, USA (2004).
- [18] G.B. Folland, *Real analysis, modern techniques and their applications*. Second edition, Wiley Interscience (1999).
- [19] P. Forchheimer, Wasserbewegung durch Boden. *Z. Ver. Deutsh. Ing.* **45** (1901) 1782–1788.

- [20] V. Girault and P.-A. Raviart, *Finite element methods for Navier-Stokes equations – Theory and algorithms*, Springer Series in Computational Mathematics **5**. Springer-Verlag, Berlin, Germany (1986).
- [21] V. Girault and M.F. Wheeler, Numerical discretization of a Darcy-Forchheimer model. *Numer. Math.* **110** (2008) 161–198.
- [22] V. Girault, R. Nochetto and L.R. Scott, Maximum-norm stability of the finite-element Stokes projection. *J. Math. Pure. Appl.* **84** (2005) 279–330.
- [23] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, Monographs and Studies in Mathematics **24**. Pitman, Boston, USA (1985).
- [24] F. Hecht, A. Le Hyaric, O. Pironneau and K. Ohtsuka, *Freefem++*. Second Edition, Version 2.24-2-2. Laboratoire J.-L. Lions, UPMC, Paris, France (2008).
- [25] A.Ya. Helemskii, *Lectures and exercises on functional analysis*, Translations of Mathematical Monographs **233**. American Mathematical Society, USA (2006).
- [26] L.V. Kantorovich and G.P. Akilov, *Functional analysis*. Third edition, Nauka (1984) [in Russian].
- [27] D. Kim and E.J. Park, Primal mixed finite-element approximation of elliptic equations with gradient nonlinearities. *Comput. Math. Appl.* **51** (2006) 793–804.
- [28] J.L. Lions and E. Magenes, *Problèmes aux Limites non Homogènes et Applications*, I. Dunod, Paris, France (1968).
- [29] E.J. Park, Mixed finite element methods for nonlinear second order elliptic problems. *SIAM J. Numer. Anal.* **32** (1995) 865–885.
- [30] S.E. Pastukhova, Substantiation of the Darcy Law for a porous medium with condition of partial adhesion. *Sbornik Math.* **189** (1998) 1871–1888.
- [31] K.R. Rajagopal, On a hierarchy of approximate models for flows of incompressible fluids through porous solids. *M3AS* **17** (2007) 215–252.
- [32] J.E. Roberts and J.-M. Thomas, Mixed and Hybrid methods in *Handbook of Numerical Analysis II: Finite Element Methods (Part 1)*, P.G. Ciarlet and J.L. Lions Eds., Amsterdam, North-Holland (1991) 523–639.
- [33] J. Schöberl and W. Zulehner, Symmetric indefinite preconditioners for saddle point problems with applications to pde-constrained optimization problems. *SIAM J. Matrix Anal. Appl.* **29** (2007) 752–773.
- [34] E. Skjetne and J.L. Auriault, Homogenization of wall-slip gas flow through porous media. *Transp. Porous Media* **36** (1999) 293–306.
- [35] L. Tartar, *An introduction to Sobolev spaces and interpolation spaces*, Lecture Notes of the Unione Matematica Italiana **3**. Springer-Verlag, Berlin-Heidelberg (2007).
- [36] W. Zulehner, Analysis of iterative methods for saddle point problems: a unified approach. *Math. Comp.* **71** (2001) 479–505.