

Newton's Method and Its Hybrid with Machine Learning for Navier-Stokes Darcy Models Discretized by Mixed Element Methods

Jianguo Huang¹, Hui Peng^{1,*} and Haohao Wu¹

¹ School of Mathematical Sciences, and MOE-LSC, Shanghai Jiao Tong University, Shanghai 200240, China.

Received 22 March 2024; Accepted (in revised version) 4 September 2024

Abstract. This paper focuses on discussing Newton's method and its hybrid with machine learning for the steady state Navier-Stokes Darcy model discretized by mixed element methods. First, a Newton iterative method is introduced for solving the relative discretized problem. It is proved technically that this method converges quadratically with the convergence rate independent of the mixed element mesh size, under certain standard conditions. Later on, a deep learning algorithm is proposed for solving this nonlinear coupled problem. Following the ideas of an earlier work by Huang, Wang and Yang (2020), an Int-Deep algorithm is constructed by combining the previous two methods so as to further improve the computational efficiency and robustness. A series of numerical examples are reported to show the numerical performance of the proposed methods.

AMS subject classifications: 65N12, 65N15, 65N22, 65N30

Key words: Navier-Stokes Darcy model, deep learning, Newton iterative method, Int-Deep method, convergence analysis.

1 Introduction

The Navier-Stokes Darcy model is frequently encountered in various industrial engineering scenarios, for instance, in groundwater [10,14,27], flow in porous media [2,3], industrial filtrations [18] and so on. Due to its importance, many numerical methods have been developed and analyzed for the nonlinear coupled model, which can be roughly divided into two categories. The first one focuses on numerically solving the coupled

*Corresponding author. Email addresses: jghuang@sjtu.edu.cn (J. Huang), penghui23@sjtu.edu.cn (H. Peng), wu1150132305@sjtu.edu.cn (H. Wu)

system directly. Historically, Girault and Rivi  re [17] propose and analyze a discontinuous Galerkin method for the Navier-Stokes Darcy model. Following [17], Chidyagwai and Rivi  re [12] develop a hybrid coupled method, where the continuous finite elements are employed in the free flow region and the discontinuous Galerkin finite elements in the porous medium region, respectively. Wu and Mei [32] discretize the model with a non-conforming finite volume element method. Cesmelioglu and Rhebergen present and analyze in [11] a strongly conservative hybridizable discontinuous Galerkin scheme. The other strategy is to decouple the model first and then numerically solve two subproblems separately. In [9], Cai, Mu and Xu propose a decoupled and linearized two-grid algorithm. The two-grid decoupled method is further studied in [25, 36]. In addition, Cai, Huang and Mu also develop a multi-grid algorithm in [8]. He et al. [20] design a domain decomposition method for the Navier-Stokes Darcy model with the BJ interface condition. In order to further improve computational efficiency, Du et al. develop in [16, 31] a series of parallel algorithms based on decoupled model. We mention that all the previous methods are used to solve steady-state Navier-Stokes Darcy models. In fact, there are deep and through works on numerical solution for unsteady state models, and we skip the details due to the limit of space.

Both the direct and decoupling methods require solving nonlinear system of equations globally or locally through iterative methods, e.g. Newton iterative method. However, there are few works on investigating the global convergence analysis of these iterative methods. As far as we know, only Badea, Discacciati, and Quarteroni [4] study Newton iterative method for the Navier-Stokes Darcy model in the case of infinite dimensions. The analysis is rather involved; the key techniques are first reformulating the coupled nonlinear problem as an interface equation and then developing the underlying convergence analysis of Newton iterative method by means of the Kantorovich theorem. However, when applying the previous arguments to the nonlinear system arising from discretization of the Navier-Stokes Darcy model by mixed element methods, it is very hard to show the convergence rate is uniformly bounded with respect to the finite element mesh size h . This study is important for real applications. In fact, if the convergence rate goes to 1 when h goes to zero, then we even cannot observe convergence of the underlying iterative method due to the rounding-off error. The other point to be emphasized is that Newton's method is locally convergent, so it is very challenging to develop a strategy on the choice of initial guesses to ensure convergence. For our problem under discussion, a common way is to set the initial guess as the numerical solution of the corresponding linear Stokes Darcy problem. However, such a choice is not robust with respect to the model coefficients; we refer to numerical results in Subsection 5.1.2 for details. Therefore, it's necessary to develop an effective approach to choosing initial guesses in order to improve the computational performance and robustness of Newton iterative method.

With the above discussion in mind, our study in this paper is twofold. First of all, we analyze in detail the convergence of Newton iterative method for the Navier-Stokes Darcy model discretized by mixed finite element methods. Different from the approach

in [4], we derive the required result (cf. Theorem 3.1) in another way. Our analysis relies on a key observation which indicates that a critical interface term in the Navier-Stokes equations, which is connected with Darcy equations, can be decomposed into a positive term plus an easily estimated term. Combining this finding with the analysis of Newton's iteration for Navier-Stokes model in [22], we prove that the iterative method (cf. (3.1)-(3.3)) is convergent quadratically under certain standard conditions, with the convergence rate independent of the finite element mesh size h .

On the other hand, it is observed that a deep learning initialized iterative (Int-Deep) method is introduced in [23] for nonlinear variational problems. It is a hybrid iterative method, consisting of two phases. In the first phase, an expectation minimization problem formulated for a given nonlinear partial differential equations (PDE) is solved by deep learning (DL) methods. In the second phase, the numerical solution from the first phase is interpolated as the initial guess, and some Newton-type iterative methods are used to solve the finite-dimensional problem discretized by mixed finite element methods, which converges rapidly to the mixed element solution. This method has been used successfully to solve semi-linear elliptic problems, linear and nonlinear eigenvalue problems (cf. [23]). The success of the method may rely partly on the so-called frequency principle mentioned in [33,34], which asserts that a deep neural network (DNN) tends to fit training data by a low-frequency function. This result implies that the DL solution with few iteration steps may capture the low frequency components of the exact solution, so it is reasonable to act as an initial guess of an iterative method for a variational problem.

Thus, the second goal of this paper is to use the ideas of the above Int-Deep method to solve the Navier-Stokes Darcy model combined with Newton iterative method just mentioned. To this end, we first design a physics informed neural networks (PINN)-type [28,29,35] DL algorithm for the Navier-Stokes Darcy model, through expressing the unknowns with ResNet functions [21] and constructing a log-loss function (see (4.5)) to improve computational efficiency. Furthermore, following the ideas in [23], we introduce the Int-Deep algorithm for the discrete Navier-Stokes Darcy model (cf. (2.6)-(2.7)). In other words, we solve the discrete problem by means of Newton iterative method, with the interpolant of the DNN solution as an initial guess. Finally, we perform a series of numerical examples to confirm that this Int-Deep algorithm is able to reach the accuracy of the mixed element method with few iteration steps. In particular, our numerical experiments in this paper indicate that the algorithm is robust with respect to the underlying physical parameters, e.g., the viscosity coefficient of fluid and the hydraulic conductivity tensor of the porous medium.

The rest of the paper is organized as follows. In Section 2, we introduce the model problem and recall some existing results of the mixed element method for the nonlinear coupled system. In Section 3, we discuss the stability and convergence of Newton iterative method under some conditions. In Section 4, a DL algorithm for the Navier-Stokes Darcy model is designed, and then the Int-Deep algorithm is developed for the Navier-Stokes Darcy model. In Section 5, we present some numerical examples to demonstrate

the accuracy and efficiency of the Int-Deep algorithm. Finally, Section 6 offers some concluding remarks.

2 The Navier-Stokes Darcy problem and its discretization

In this section, we aim to introduce the Navier-Stokes Darcy model and its mixed element discretization. We also recall some important existing results which will be frequently used later on.

2.1 Navier-Stokes Darcy model

Let $\Omega \subset R^d$ ($d = 2, 3$) be a bounded polytopal domain, which is subdivided into a free fluid region Ω_f and a porous media region Ω_p by an interface Γ , as shown in Fig. 1. Let $\Gamma_f = \partial\Omega_f \setminus \Gamma$ and $\Gamma_p = \partial\Omega_p \setminus \Gamma$ be the outer boundaries of the domains Ω_f and Ω_p , respectively. Denote by $\mathbf{n}_f$ and $\mathbf{n}_p$ the unit outward normal vectors on $\partial\Omega_f$ and $\partial\Omega_p$, respectively. The unit tangential vector on Γ is denoted by $\boldsymbol{\tau}$.

The free flow in Ω_f is governed by the steady Navier-Stokes equations:

$$\begin{cases} -\operatorname{div}(\mathbf{T}(\mathbf{u}, p)) + \rho(\mathbf{u} \cdot \nabla)\mathbf{u} = \mathbf{f}_f & \text{in } \Omega_f, \\ \operatorname{div} \mathbf{u} = 0 & \text{in } \Omega_f, \end{cases} \quad (2.1)$$

where $\mathbf{u}$ and p indicate the velocity field and the pressure field of the flow, respectively; $\nu > 0$ stands for the viscosity coefficient, $\mathbf{T}(\mathbf{u}, p) = 2\nu\mathbf{D}(\mathbf{u}) - p\mathbf{I}$ is the stress tensor, and $\mathbf{D}(\mathbf{u}) = \frac{1}{2}(\nabla\mathbf{u} + \nabla^T\mathbf{u})$ is the deformation rate tensor; ρ denotes the density of the flow and $\mathbf{f}_f$ is the external force.

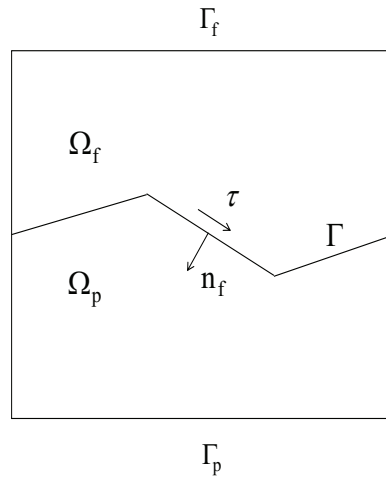


Figure 1: Domain schematic for Navier-Stokes Darcy coupled flow.

The flow motion in the porous media region Ω_p is described by Darcy's law:

$$-\operatorname{div}(\mathbf{K}\nabla\varphi) = f_p \quad \text{in } \Omega_p, \quad (2.2)$$

where φ is the hydraulic head, f_p is a source term and $\mathbf{K}$ represents the hydraulic conductivity tensor. In this paper, we assume that $\mathbf{K}$ is symmetric $K_{ij} = K_{ji}$, and positive definite, i.e.

$$\alpha_1(\mathbf{x}, \mathbf{x}) \leq (\mathbf{K}\mathbf{x}, \mathbf{x}) \leq \alpha_2(\mathbf{x}, \mathbf{x}) \quad \forall \mathbf{x} \in \Omega_p,$$

for two positive constants α_1 and α_2 .

The physical quantities are coupled on the interface Γ through three interface conditions:

$$\begin{cases} \mathbf{u} \cdot \mathbf{n}_f = -\mathbf{K}\nabla\varphi \cdot \mathbf{n}_f, \\ [-\mathbf{T}(\mathbf{u}, p)\mathbf{n}_f] \cdot \mathbf{n}_f = \rho g \varphi, \\ [-\mathbf{T}(\mathbf{u}, p)\mathbf{n}_f] \cdot \boldsymbol{\tau}_j = \frac{\nu\alpha}{\sqrt{\boldsymbol{\tau}_j \cdot \nu \mathbf{K} \cdot \boldsymbol{\tau}_j}} \mathbf{u} \cdot \boldsymbol{\tau}_j, \quad j=1, \dots, d-1, \end{cases} \quad (2.3)$$

where g is the gravitational acceleration, $\{\boldsymbol{\tau}_j\}_{j=1}^{d-1}$ are linearly independent unit tangential vectors on Γ , and α is an experimentally determined positive parameter.

For simplicity, we consider homogeneous Dirichlet boundary conditions on the outer boundaries:

$$\begin{cases} \mathbf{u} = \mathbf{0} & \text{on } \Gamma_f, \\ \varphi = 0 & \text{on } \Gamma_p. \end{cases} \quad (2.4)$$

In order to describe the variational formulation of this model, we have to introduce some function spaces in advance. Denote the admissible spaces for velocity, pressure and hydraulic head, respectively by

$$\begin{aligned} \mathbf{X}_f &= \left\{ \mathbf{v} \in \mathbf{H}^1(\Omega_f) = [H^1(\Omega_f)]^d \mid \mathbf{v} = \mathbf{0} \text{ on } \Gamma_f \right\}, \quad Q = L^2(\Omega_f), \\ X_p &= \left\{ \psi \in H^1(\Omega_p) \mid \psi = 0 \text{ on } \Gamma_p \right\}. \end{aligned}$$

Then, the variational formulation of the coupled Navier-Stokes Darcy model (2.1) – (2.4) is given as follows.

Find $\mathbf{u} \in \mathbf{X}_f, \varphi \in X_p$ and $q \in Q$ such that

$$\begin{cases} a_f(\mathbf{u}, \mathbf{v}) + a_p(\varphi, \psi) + a_\Gamma(\mathbf{u}, \psi; \mathbf{v}, \varphi) + c(\mathbf{u}; \mathbf{u}, \mathbf{v}) + b(\mathbf{v}, p) \\ \quad = (\mathbf{f}_f, \mathbf{v}) + \rho g(f_p, \psi) \quad \forall \mathbf{v} \in \mathbf{X}_f, \psi \in X_p, \\ b(\mathbf{u}, q) = 0 \quad \forall q \in Q, \end{cases} \quad (2.5)$$

where

$$\begin{aligned} a_f(\mathbf{u}, \mathbf{v}) &= \int_{\Omega_f} 2\nu \mathbf{D}(\mathbf{u}) : \mathbf{D}(\mathbf{v}) + \sum_{j=1}^{d-1} \frac{\nu \alpha}{\sqrt{\boldsymbol{\tau}_j \cdot \nu \mathbf{K} \boldsymbol{\tau}_j}} \int_{\Gamma} (\mathbf{u} \cdot \boldsymbol{\tau}_j) (\mathbf{v} \cdot \boldsymbol{\tau}_j), \\ a_p(\varphi, \psi) &= \rho g \int_{\Omega_p} \mathbf{K} \nabla \varphi \cdot \nabla \psi, \quad a_{\Gamma}(\mathbf{u}, \psi; \mathbf{v}, \varphi) = \rho g \int_{\Gamma} (\varphi \mathbf{v} - \psi \mathbf{u}) \cdot \mathbf{n}_f, \\ c(\mathbf{u}; \mathbf{v}, \mathbf{w}) &= \rho \int_{\Omega_f} (\mathbf{u} \cdot \nabla) \mathbf{v} \cdot \mathbf{w}, \quad b(\mathbf{v}, p) = - \int_{\Omega_f} p \operatorname{div} \mathbf{v}, \\ (\mathbf{f}_f, \mathbf{v}) &= \int_{\Omega_f} \mathbf{f}_f \cdot \mathbf{v}, \quad (f_p, \psi) = \int_{\Omega_p} f_p \psi. \end{aligned}$$

Here and hereafter, we omit the infinitesimal element over an integration symbol when there is no confusion caused.

The existence and uniqueness of solution of (2.5) can be found in [17].

2.2 Finite element approximation

In this subsection, we shall give the mixed element discretization of the Navier-Stokes Darcy model and review some existing results for subsequent analysis.

Let $\mathcal{T}_{f,h}$ (resp. $\mathcal{T}_{p,h}$) be a nondegenerate quasi-uniform triangulation of Ω_f (resp. Ω_p). The partition $\mathcal{T}_{f,h}$ matches with the partition $\mathcal{T}_{p,h}$ at the interface Γ . Let $\mathcal{T}_h = \mathcal{T}_{f,h} \cup \mathcal{T}_{p,h}$ be the partition of the whole domain Ω . For each element $T \in \mathcal{T}_h$, we denote by h_T the diameter of T ; the mesh size of $\mathcal{T}_h$ is defined by $h = \max_{T \in \mathcal{T}_h} h_T$. Let k be a non-negative integer, and denote by $\mathbb{P}_k(T)$ the set of polynomials with the degree no more than k on element $T \in \mathcal{T}_h$. Moreover, we use C (with or without subscripts) to denote a generic positive constant independent of the mesh size h , which may take different values at different occurrences.

There are various combinations of the stable Stokes finite element pairs [6, 24] (such as MINI element, Taylor-Hood element, conforming Crouzeix-Raviart element) with continuous finite elements for Darcy equations. For convenience, we focus on the traditional Taylor-Hood element as an typical example in this paper. It is not difficult to extend the corresponding theory to other elements.

Define some finite element spaces by

$$\begin{aligned} \mathbf{X}_{f,h} &= \{ \mathbf{v}_h \in \mathbf{X}_f, \mathbf{v}_h|_T \in [\mathbb{P}_2(T)]^d \ \forall T \in \mathcal{T}_{f,h} \}, \\ Q_h &= \{ q_h \in Q \cap C^0(\Omega), q_h|_T \in \mathbb{P}_1(T) \ \forall T \in \mathcal{T}_{f,h} \}, \\ X_{p,h} &= \{ \psi_h \in X_p, \psi_h|_T \in \mathbb{P}_2(T) \ \forall T \in \mathcal{T}_{p,h} \}. \end{aligned}$$

With these discrete function spaces, we give the mixed element method for the coupled

Navier-Stokes Darcy model (2.1)-(2.4): Find $\mathbf{u}_h \in \mathbf{X}_{f,h}$, $p_h \in Q_h$, and $\varphi_h \in X_{p,h}$ such that

$$\begin{aligned} a_f(\mathbf{u}_h, \mathbf{v}_h) + a_p(\varphi_h, \psi_h) + a_\Gamma(\mathbf{u}_h, \psi_h; \mathbf{v}_h, \varphi_h) + c(\mathbf{u}_h; \mathbf{u}_h, \mathbf{v}_h) + b(\mathbf{v}_h, p_h) \\ = (f_f, \mathbf{v}_h) + \rho g(f_p, \psi_h) \quad \forall \mathbf{v}_h \in \mathbf{X}_{f,h}, \quad \psi_h \in X_{p,h}, \end{aligned} \quad (2.6)$$

$$b(\mathbf{u}_h, q_h) = 0 \quad \forall q_h \in Q_h. \quad (2.7)$$

To proceed with the forthcoming discussion about the well-posedness and convergence of the mixed element method, we first recall some basic inequalities in Sobolev spaces as follows [1]. There exist constants C_p , C_t , C_k and C_s only depending on the domain Ω_f , and $\tilde{C}_p$ and $\tilde{C}_t$ only depending on the domain Ω_p , such that for all $\mathbf{v} \in \mathbf{X}_f$ and $\psi \in X_p$,

$$\begin{aligned} (\text{the Poincaré inequality}) \quad & \|\mathbf{v}\|_{0,\Omega_f} \leq C_p |\mathbf{v}|_{1,\Omega_f}, & \|\psi\|_{0,\Omega_p} \leq \tilde{C}_p |\psi|_{1,\Omega_p}; \\ (\text{the trace inequality}) \quad & \|\mathbf{v}\|_{0,\Gamma} \leq C_t |\mathbf{v}|_{1,\Omega_f}, & \|\psi\|_{0,\Gamma} \leq \tilde{C}_t |\psi|_{1,\Omega_p}; \\ (\text{Korn's inequality}) \quad & |\mathbf{v}|_{1,\Omega_f} \leq C_k \|\mathbf{D}(\mathbf{v})\|_{0,\Omega_f}; \\ (\text{the Sobolev inequality}) \quad & \|\mathbf{v}\|_{L^6(\Omega_f)} \leq C_s |\nabla \mathbf{v}|_{1,\Omega_f}. \end{aligned}$$

Using the Hölder inequality, the Sobolev inequality and Korn's inequality, we immediately have (cf. [30])

$$c(\mathbf{u}; \mathbf{v}, \mathbf{w}) \leq \mathcal{N}_d \|\mathbf{D}(\mathbf{u})\|_{0,\Omega_f} \|\mathbf{D}(\mathbf{v})\|_{0,\Omega_f} \|\mathbf{D}(\mathbf{w})\|_{0,\Omega_f}, \quad (2.8)$$

with $\mathcal{N}_d$ depending only on C_s , C_k and d .

In order to describe some important results, we define

$$\|f_f\|_* = \frac{(f_f, \mathbf{v})}{\|\mathbf{D}(\mathbf{v})\|_{0,\Omega_f}}, \quad \|f_p\|_* = \frac{(f_p, \psi)}{|\psi|_{1,\Omega_p}}.$$

The theoretical results given below indicate the well-posedness and convergence of the mixed element method (2.6)-(2.7). We refer the interesting readers to [9] for details along this line.

Theorem 2.1. Assume that $f_f \in [L^2(\Omega_f)]^d$, $f_p \in L^2(\Omega_p)$ and the viscosity coefficient ν satisfies

$$\nu^{\frac{3}{2}} \geq \sqrt{2} \mathcal{N}_d \mathcal{R}, \quad (2.9)$$

with

$$\mathcal{R} = \left(\frac{1}{\nu} \|f_f\|_*^2 + \frac{\rho g}{\alpha_1} \|f_p\|_*^2 \right)^{\frac{1}{2}}. \quad (2.10)$$

Then the finite element scheme (2.6)-(2.7) has a unique solution satisfying the following estimate

$$2\nu \|\mathbf{D}(\mathbf{u}_h)\|_{0,\Omega_f}^2 + \rho g \|\mathbf{K} \nabla \varphi_h\|_{0,\Omega_p}^2 \leq \mathcal{R}^2. \quad (2.11)$$

Remark 2.1. Under the assumption (2.9), it follows from (2.10) that

$$\|D(\mathbf{u}_h)\|_{0,\Omega_f} \leq \frac{\nu}{2\mathcal{N}_d}. \quad (2.12)$$

Lemma 2.1. There is a constant $\beta > 0$ independent of h such that

$$\inf_{q_h \in Q_h} \sup_{\mathbf{v}_h \in \mathbf{X}_{f,h}} \frac{b(\mathbf{v}_h, q_h)}{\|D(\mathbf{v}_h)\|_{0,\Omega_f} \|q_h\|_{0,\Omega_f}} \geq \beta. \quad (2.13)$$

Theorem 2.2. Let $(\mathbf{u}, p, \varphi)$ be the solution of problem (2.1)-(2.4) and $(\mathbf{u}_h, p_h, \varphi_h)$ be the solution of finite element scheme (2.6)-(2.7). Assume that $\mathbf{u} \in \mathbf{H}^3(\Omega_f)$, $p \in H^2(\Omega_f)$ and $\varphi \in H^3(\Omega_p)$. Then the following estimates hold:

$$\begin{aligned} & 2\nu \|D(\mathbf{u} - \mathbf{u}_h)\|_{0,\Omega_f}^2 + \sum_{j=1}^{d-1} \|(\mathbf{u} - \mathbf{u}_h) \cdot \boldsymbol{\tau}_j\|_{0,\Gamma}^2 + \rho g \|K^{\frac{1}{2}} \nabla(\varphi - \varphi_h)\|_{0,\Omega_p}^2 \\ & \leq Ch^4 (\|\mathbf{u}\|_{3,\Omega_f}^2 + \|p\|_{2,\Omega_f}^2 + \|\varphi\|_{3,\Omega_p}^2), \end{aligned} \quad (2.14)$$

$$\|p - p_h\|_{0,\Omega_f}^2 \leq Ch^4 (\|\mathbf{u}\|_{3,\Omega_f}^2 + \|p\|_{2,\Omega_f}^2 + \|\varphi\|_{3,\Omega_p}^2). \quad (2.15)$$

3 Newton's method for the Navier-Stokes Darcy model

In this section, we give Newton iterative method for solving the Navier-Stokes Darcy model under the mixed element discretization and provide stability and convergence analysis of the iterative method.

According to Newton iterative method for Navier-Stokes Darcy problem in infinite dimension (cf. [4]), its finite element analogue can be described as follows.

Given $\mathbf{u}_0^h \in \mathbf{X}_{f,h}$, for $n \geq 1$, find $\mathbf{u}_n^h \in \mathbf{X}_{f,h}$, $p_n^h \in Q_h$ and $\varphi_n^h \in X_{p,h}$ such that

$$\begin{aligned} & a_f(\mathbf{u}_n^h, \mathbf{v}_h) + b(\mathbf{v}_h, p_n^h) + \rho g \langle \varphi_n^h, \mathbf{v}_h \cdot \mathbf{n}_f \rangle_\Gamma + c(\mathbf{u}_n^h; \mathbf{u}_{n-1}^h, \mathbf{v}_h) + c(\mathbf{u}_{n-1}^h; \mathbf{u}_n^h, \mathbf{v}_h) \\ & = (\mathbf{f}_f, \mathbf{v}_h) + c(\mathbf{u}_{n-1}^h; \mathbf{u}_{n-1}^h, \mathbf{v}_h), \end{aligned} \quad (3.1)$$

$$b(\mathbf{u}_n^h, q_h) = 0, \quad (3.2)$$

$$a_p(\varphi_n^h, \psi_h) = \rho g \langle \psi_h, \mathbf{u}_n^h \cdot \mathbf{n}_f \rangle_\Gamma + \rho g (f_p, \psi_h), \quad (3.3)$$

for any $\mathbf{v}_h \in \mathbf{X}_{f,h}$, $q_h \in Q_h$, $\psi_h \in X_{p,h}$.

To analyze Newton iterative method, we next introduce two auxiliary problems given below.

For any $\zeta \in L^2(\Gamma)$ or $\eta \in L^2(\Omega_p)$, find $\phi_h \in X_{p,h}$ such that

$$a_p(\phi_h, \psi_h) = \rho g \langle \zeta, \psi_h \rangle_\Gamma \quad \forall \psi_h \in X_{p,h}, \quad (3.4)$$

or

$$a_p(\phi_h, \psi_h) = \rho g (\eta, \psi_h) \quad \forall \psi_h \in X_{p,h}. \quad (3.5)$$

It is easy to prove by the Lax-Milgram lemma that problem (3.4) or problem (3.5) has a unique solution, which induces a linear operator $\phi_h = T_h(\zeta)$ or $\phi_h = T_h(\eta)$. Furthermore, we have the following estimate which is easy to derive but important in our forthcoming analysis

$$\|T_h(\eta)\|_{1,\Omega_p} \leq C_r \|\eta\|_*, \quad (3.6)$$

with C_r a positive constant independent of h .

The discussion on stability and convergence will be based on the following assumptions about the initial function $\mathbf{u}_0^h \in \mathbf{X}_{f,h}$ in Newton iterative method (3.1)-(3.3):

$$2\nu \|D(\mathbf{u}_0^h)\|_{0,\Omega_f} \leq C_1 (\|\mathbf{f}_f\|_* + \|\mathbf{f}_p\|_*), \quad (3.7)$$

$$\|D(\mathbf{u}_h - \mathbf{u}_0^h)\|_{0,\Omega_f} \leq \frac{C_1}{\nu} (\|\mathbf{f}_f\|_* + \|\mathbf{f}_p\|_*), \quad (3.8)$$

where

$$C_1 = 1 + \rho g C_r \tilde{C}_t C_k.$$

Lemma 3.1. Assume the initial conditions (3.7)-(3.8) and the following stability condition hold:

$$\|\mathbf{f}_f\|_* + \|\mathbf{f}_p\|_* \leq \frac{\nu^2}{9C_1 N_d}. \quad (3.9)$$

Then $\mathbf{u}_m^h$ defined by Newton iterative method (3.1)-(3.3) satisfies

$$\|D(\mathbf{u}_m^h)\|_{0,\Omega_f} \leq \frac{3C_1}{2\nu} (\|\mathbf{f}_f\|_* + \|\mathbf{f}_p\|_*), \quad (3.10)$$

for all $m \geq 0$.

Proof. We prove (3.10) by mathematical induction. The initial condition (3.7) implies that (3.10) holds for $m = 0$. Assuming that (3.10) holds for $m = J$ and we aim to prove the estimate holds for $m + 1$. Taking $\zeta = \mathbf{u}_{J+1}^h \cdot \mathbf{n}_f$ in (3.4) gives $\phi_{J+1}^b = T_h(\mathbf{u}_{J+1}^h \cdot \mathbf{n}_f)$. Taking $\eta = \mathbf{f}_p$ in (3.5) gives ϕ^i , which satisfies $\|\phi^i\|_{1,\Omega_p} \leq C_r \|\mathbf{f}_p\|_*$ from (3.6). According to the superposition principle, $\phi_{J+1}^h = \phi_{J+1}^b + \phi^i$ solves (3.3), which implies that

$$\begin{aligned} \rho g \langle \phi_{J+1}^h, \mathbf{u}_{J+1}^h \cdot \mathbf{n}_f \rangle_\Gamma &= \rho g \langle \phi_{J+1}^b, \mathbf{u}_{J+1}^h \cdot \mathbf{n}_f \rangle_\Gamma + \rho g \langle \phi^i, \mathbf{u}_{J+1}^h \cdot \mathbf{n}_f \rangle_\Gamma \\ &= a_p(T_h(\mathbf{u}_{J+1}^h \cdot \mathbf{n}_f), \phi_{J+1}^b) + \rho g \langle \phi^i, \mathbf{u}_{J+1}^h \cdot \mathbf{n}_f \rangle_\Gamma \\ &= a_p(\phi_{J+1}^b, \phi_{J+1}^b) + \rho g \langle \phi^i, \mathbf{u}_{J+1}^h \cdot \mathbf{n}_f \rangle_\Gamma \\ &\geq \rho g \langle \phi^i, \mathbf{u}_{J+1}^h \cdot \mathbf{n}_f \rangle_\Gamma. \end{aligned} \quad (3.11)$$

Taking $(v_h, q_h) = (u_h^{J+1}, p_h^{J+1}) \in \mathbf{X}_{f,h} \times Q_h$ in (3.1) with $n = J+1$ and using Korn's inequality, the estimates (2.8), (3.11) and the induction assumption, we have

$$\begin{aligned}
& 2\nu \|D(u_{J+1}^h)\|_{0,\Omega_f}^2 \\
& \leq (f_f, u_{J+1}^h) - \rho g \langle \phi_{J+1}^h, u_{J+1}^h \cdot \mathbf{n}_f \rangle_\Gamma \\
& \quad + c(u_J^h; u_J^h, u_{J+1}^h) - c(u_{J+1}^h; u_J^h, u_{J+1}^h) - c(u_J^h; u_{J+1}^h, u_{J+1}^h) \\
& \leq \|f_f\|_* \|D(u_{J+1}^h)\|_{0,\Omega_f} - \rho g \langle \phi_{J+1}^h, u_{J+1}^h \cdot \mathbf{n}_f \rangle_\Gamma \\
& \quad + \mathcal{N}_d \|D(u_J^h)\|_{0,\Omega_f}^2 \|D(u_{J+1}^h)\|_{0,\Omega_f} + 2\mathcal{N}_d \|D(u_J^h)\|_{0,\Omega_f} \|D(u_{J+1}^h)\|_{0,\Omega_f}^2 \\
& \leq C_1 (\|f_f\|_* + \|f_p\|_*) \|D(u_{J+1}^h)\|_{0,\Omega_f} + \frac{9\mathcal{N}_d}{4\nu^2} C_1^2 (\|f_f\|_* + \|f_p\|_*)^2 \|D(u_{J+1}^h)\|_{0,\Omega_f} \\
& \quad + \frac{3\mathcal{N}_d}{\nu} C_1 (\|f_f\|_* + \|f_p\|_*) \|D(u_{J+1}^h)\|_{0,\Omega_f}^2.
\end{aligned}$$

This, in conjunction with (3.9), gives rise to

$$\left(\nu - \frac{\nu}{6}\right) \|D(u_{J+1}^h)\|_{0,\Omega_f} \leq \frac{5}{4} C_1 (\|f_f\|_* + \|f_p\|_*),$$

which implies (3.10). $\square$

Lemma 3.2. *The function u_h defined by the mixed element method (2.6)-(2.7) satisfies*

$$\nu \|D(u_h)\|_{0,\Omega_f} \leq C_1 (\|f_f\|_* + \|f_p\|_*). \quad (3.12)$$

Proof. Taking $v_h = u_h$, $\psi_h = 0$ in (2.6) and noting that

$$b(u_h, p_h) = 0,$$

we get

$$2\nu \|D(u_h)\|_{0,\Omega_f}^2 = (f_f, u_h) - c(u_h; u_h, u_h) - \rho g \langle \phi_h, u_h \cdot \mathbf{n}_f \rangle_\Gamma. \quad (3.13)$$

In view of the estimates (2.8) and (2.12), the trilinear form in the above equation can be bounded as

$$c(u_h; u_h, u_h) \leq \mathcal{N}_d \|D(u_h)\|_{0,\Omega_f}^3 \leq \frac{\nu}{2} \|D(u_h)\|_{0,\Omega_f}^2. \quad (3.14)$$

In order to handle the interface term $\rho g \langle \phi_h, u_h \cdot \mathbf{n}_f \rangle_\Gamma$, taking $v_h = \mathbf{0}$ in (2.6), we get

$$a_p(\phi_h, \psi_h) = \rho g (f_p, \psi_h) + \rho g \langle u_h \cdot \mathbf{n}_f, \psi_h \rangle_\Gamma. \quad (3.15)$$

Applying a similar argument used for proving Lemma 3.1, we decompose φ_h as $\varphi_h = \phi_h^b + \phi_h^i$, where ϕ_h^b is the solution of (3.4) with $\zeta = \mathbf{u}_h \cdot \mathbf{n}_f$, and ϕ_h^i is the solution of (3.5) with $\eta = f_p$. Then, for the interface term, we have

$$\begin{aligned} \rho g \langle \varphi_h, \mathbf{u}_h \cdot \mathbf{n}_f \rangle_\Gamma &= \rho g \langle \phi_h^b, \mathbf{u}_h \cdot \mathbf{n}_f \rangle_\Gamma + \rho g \langle \phi_h^i, \mathbf{u}_h \cdot \mathbf{n}_f \rangle_\Gamma \\ &= a_p(\phi_h^b, \phi_h^b) + \rho g \langle \phi_h^i, \mathbf{u}_h \cdot \mathbf{n}_f \rangle_\Gamma \\ &\geq \rho g \langle \phi_h^i, \mathbf{u}_h \cdot \mathbf{n}_f \rangle_\Gamma. \end{aligned} \quad (3.16)$$

The combination of (3.14)-(3.16) implies the estimate (3.12). $\square$

Finally, we derive error estimates of Newton iterative method. Let $(\mathbf{u}_h, \varphi_h, p_h)$ be the solution of the scheme (2.6)-(2.7) and $(\mathbf{u}_n^h, p_n^h, \varphi_n^h)$ be the solution of Newton iterative method (3.1)-(3.3). Set $\mathbf{e}_n^u = \mathbf{u}_h - \mathbf{u}_n^h$, $e_n^p = p_h - p_n^h$ and $e_n^\varphi = \varphi_h - \varphi_n^h$.

Theorem 3.1. *Under the assumptions of Lemma 3.1, the following error estimates hold true.*

$$\|\mathbf{D}(\mathbf{e}_n^u)\|_{0,\Omega_f} \leq \frac{C_1}{\nu} \left(\frac{3C_1\mathcal{N}_d}{2\nu^2} (\|\mathbf{f}_f\|_* + \|f_p\|_*) \right)^{2^n-1} (\|\mathbf{f}_f\|_* + \|f_p\|_*), \quad (3.17)$$

$$|\nabla e_n^\varphi|_{1,\Omega_p} \leq \frac{\tilde{C}_t C_t C_k C_1}{\alpha_1 \nu} \left(\frac{3C_1\mathcal{N}_d}{2\nu^2} (\|\mathbf{f}_f\|_* + \|f_p\|_*) \right)^{2^n-1} (\|\mathbf{f}_f\|_* + \|f_p\|_*), \quad (3.18)$$

$$\begin{aligned} \|e_n^p\|_{0,\Omega_f} &\leq \frac{C_1}{\beta} \left(\frac{10}{3} + \frac{\tilde{C}_t^2 C_t^2 C_k^2}{\alpha_1 \nu} \right) \\ &\quad \left(\frac{3C_1\mathcal{N}_d}{2\nu^2} (\|\mathbf{f}_f\|_* + \|f_p\|_*) \right)^{2^n-1} (\|\mathbf{f}_f\|_* + \|f_p\|_*), \end{aligned} \quad (3.19)$$

for all $n \geq 1$.

Proof. Subtracting (3.1)-(3.3) from the mixed element scheme (2.6)-(2.7), we find the error equations for Newton iterative method are determined by

$$\begin{aligned} a_f(\mathbf{e}_n^u, \mathbf{v}_h) + a_p(e_n^\varphi, \psi_h) + b(\mathbf{v}_h, e_n^p) + a_\Gamma(\mathbf{e}_n^u, \psi_h; \mathbf{v}_h, e_n^\varphi) + c(\mathbf{e}_n^u; \mathbf{u}_{n-1}^h, \mathbf{v}_h) \\ + c(\mathbf{u}_{n-1}^h; \mathbf{e}_n^u, \mathbf{v}_h) + c(\mathbf{e}_{n-1}^u; \mathbf{e}_{n-1}^u, \mathbf{v}_h) = 0 \quad \forall \mathbf{v}_h \in \mathbf{X}_{f,h}, \psi_h \in X_{p,h}, \end{aligned} \quad (3.20)$$

$$b(\mathbf{e}_n^u, q_h) = 0 \quad \forall q_h \in Q_h. \quad (3.21)$$

Taking $\mathbf{v}_h = \mathbf{e}_n^u$, $\psi_h = 0$ in (3.20) and noting that $b(\mathbf{e}_n^u, e_n^p) = 0$, we achieve

$$\begin{aligned} 2\nu \|\mathbf{D}(\mathbf{e}_n^u)\|_{0,\Omega_f}^2 \\ = -\rho g \langle e_n^\varphi, \mathbf{e}_n^u \cdot \mathbf{n}_f \rangle_\Gamma - c(\mathbf{e}_{n-1}^u; \mathbf{e}_{n-1}^u, \mathbf{e}_n^u) - c(\mathbf{e}_n^u; \mathbf{u}_{n-1}^h, \mathbf{e}_n^u) - c(\mathbf{u}_{n-1}^h; \mathbf{e}_n^u, \mathbf{e}_n^u) \\ \leq \mathcal{N}_d \|\mathbf{D}(\mathbf{e}_{n-1}^u)\|_{0,\Omega_f}^2 \|\mathbf{D}(\mathbf{e}_n^u)\|_{0,\Omega_f} + 2\mathcal{N}_d \|\mathbf{D}(\mathbf{u}_{n-1}^h)\|_{0,\Omega_f} \|\mathbf{D}(\mathbf{e}_n^u)\|_{0,\Omega_f}^2, \end{aligned}$$

and with (3.10) and (3.9), we further have

$$\|D(e_n^u)\|_{0,\Omega_f} \leq \frac{3\mathcal{N}_d}{2\nu} \|D(e_{n-1}^u)\|_{0,\Omega_f}^2, \quad (3.22)$$

for $n \geq 1$.

Now, we prove (3.17) by mathematical induction. It follows from the initial condition (3.7) that (3.17) holds for $n=0$. Assuming that (3.17) holds for $n=J$, we are going to prove that the estimate still holds for $n=J+1$. Applying the estimate (3.22) and the induction assumption gives

$$\begin{aligned} \|D(e_{J+1}^u)\|_{0,\Omega_f} &\leq \frac{3\mathcal{N}_d}{2\nu} \|D(e_J^u)\|_{0,\Omega_f}^2 \\ &\leq \frac{3\mathcal{N}_d}{2\nu} \cdot \frac{C_1^2}{\nu^2} \left(\frac{3C_1\mathcal{N}_d}{2\nu^2} (\|f_f\|_* + \|f_p\|_*) \right)^{2^{J+1}-2} (\|f_f\|_* + \|f_p\|_*)^2 \\ &= \frac{C_1}{\nu} \left(\frac{3C_1\mathcal{N}_d}{2\nu^2} (\|f_f\|_* + \|f_p\|_*) \right)^{2^{J+1}-1} (\|f_f\|_* + \|f_p\|_*), \end{aligned}$$

which is (3.17) with $n=J+1$. So we prove (3.17) holds.

Let us consider the boundedness of $|e_n^\varphi|_1$. Taking $v_h = \mathbf{0}$, $\psi_h = e_n^\varphi$ in (3.21) and applying (3.17) gives

$$\begin{aligned} \alpha_1 |\nabla e_n^\varphi|_{1,\Omega_p}^2 &\leq \|e_n^\varphi\|_{0,\Gamma} \|e_n^u\|_{0,\Gamma} \\ &\leq \tilde{C}_t |\nabla e_n^\varphi|_{1,\Omega_p} C_t C_k \|D(e_n^u)\|_{0,\Omega_f} \\ &\leq \frac{\tilde{C}_t C_t C_k C_1}{\nu} |\nabla e_n^\varphi|_{1,\Omega_p} \left(\frac{3C_1\mathcal{N}_d}{2\nu^2} (\|f_f\|_* + \|f_p\|_*) \right)^{2^n-1} (\|f_f\|_* + \|f_p\|_*). \end{aligned}$$

Finally, we estimate e_n^p . By using the discrete inf-sup condition (2.13), and the estimates (3.10), (3.9), (3.17) and (3.18), we obtain

$$\begin{aligned} \beta \|e_n^p\|_{0,\Omega_f} &\leq 2\nu \|D(e_n^u)\|_{0,\Omega_f} + \rho g C_t \tilde{C}_t C_k |e_n^\varphi|_{1,\Omega_p} \\ &\quad + 2\mathcal{N}_d \|D(u_{n-1}^h)\|_{0,\Omega_f} \|D(e_n^u)\|_{0,\Omega_f} + \mathcal{N}_d \|D(e_{n-1}^u)\|_{0,\Omega_f}^2 \\ &\leq \left(2\nu + \frac{\nu}{3} \right) \|D(e_n^u)\|_{0,\Omega_f} + \rho g C_t \tilde{C}_t C_k |e_n^\varphi|_{1,\Omega_p} + \frac{3}{2} \mathcal{N}_d \|D(e_{n-1}^u)\|_{0,\Omega_f}^2 \\ &\leq \frac{7C_1}{3} \left(\frac{3C_1\mathcal{N}_d}{2\nu^2} (\|f_f\|_* + \|f_p\|_*) \right)^{2^n-1} (\|f_f\|_* + \|f_p\|_*) \\ &\quad + \frac{\tilde{C}_t^2 C_t^2 C_k^2 C_1}{\alpha_1 \nu} \left(\frac{3C_1\mathcal{N}_d}{2\nu^2} (\|f_f\|_* + \|f_p\|_*) \right)^{2^n-1} (\|f_f\|_* + \|f_p\|_*) \\ &\quad + C_1 \left(\frac{3C_1\mathcal{N}_d}{2\nu^2} (\|f_f\|_* + \|f_p\|_*) \right)^{2^n-1} (\|f_f\|_* + \|f_p\|_*) \\ &= \left(\frac{10}{3} + \frac{\tilde{C}_t^2 C_t^2 C_k^2}{\alpha_1 \nu} \right) C_1 \left(\frac{3C_1\mathcal{N}_d}{2\nu^2} (\|f_f\|_* + \|f_p\|_*) \right)^{2^n-1} (\|f_f\|_* + \|f_p\|_*). \end{aligned}$$

The proof is complete. $\square$

Corollary 3.1. *The initial function $\mathbf{u}_0^h \in \mathbf{X}_{f,h}$ defined by the following mixed element method for Stokes Darcy problem:*

$$a_f(\mathbf{u}_0^h, \mathbf{v}_h) + b(\mathbf{v}_h, p_0^h) + \rho g \langle \phi_0^h, \mathbf{v}_h \cdot \mathbf{n}_f \rangle_\Gamma = (\mathbf{f}_f, \mathbf{v}_h) \quad \forall \mathbf{v}_h \in \mathbf{X}_{f,h}, \quad (3.23)$$

$$b(\mathbf{u}_0^h, q_h) = 0 \quad \forall q_h \in Q_h, \quad (3.24)$$

$$a_p(\phi_0^h, \psi_h) = \rho g \langle \psi_h, \mathbf{u}_0^h \cdot \mathbf{n}_f \rangle_\Gamma + \rho g (f_p, \psi_h) \quad \forall \psi_h \in X_{p,h}, \quad (3.25)$$

satisfies the initial conditions (3.7)-(3.8).

Proof. Taking $\zeta = \mathbf{u}_0^h \cdot \mathbf{n}_f$ in auxiliary problem (3.4) and $\eta = f_p$ in (3.5), the corresponding solutions are denoted by ϕ_0^b and ϕ^i , respectively. Then, we have $\phi_0^b = T_h(\mathbf{u}_0^h \cdot \mathbf{n}_f)$ and $\|\phi^i\|_{1,\Omega_p} \leq C_r \|f_p\|_*$. According to the superposition principle, we know that $\phi_0^h = \phi_0^b + \phi^i$ solves (3.25), which implies that

$$\begin{aligned} \rho g \langle \phi_0^h, \mathbf{v}_0^h \cdot \mathbf{n}_f \rangle_\Gamma &= \rho g \langle \phi_0^b, \mathbf{v}_0^h \cdot \mathbf{n}_f \rangle_\Gamma + \rho g \langle \phi^i, \mathbf{v}_0^h \cdot \mathbf{n}_f \rangle_\Gamma \\ &= a_p(T_h(\mathbf{v}_0^h \cdot \mathbf{n}_f), \phi_0^b) + \rho g \langle \phi^i, \mathbf{v}_0^h \cdot \mathbf{n}_f \rangle_\Gamma. \end{aligned} \quad (3.26)$$

Taking $\mathbf{v}_h = \mathbf{u}_0^h$ in (3.23), and using (3.26) and (3.6), we have

$$\begin{aligned} 2\nu \|\mathbf{D}(\mathbf{u}_0^h)\|_{0,\Omega_f}^2 &\leq (\mathbf{f}_f, \mathbf{u}_0^h) - \rho g \langle \phi_0^h, \mathbf{u}_0^h \cdot \mathbf{n}_f \rangle_\Gamma \\ &\leq (\mathbf{f}_f, \mathbf{u}_0^h) - a_p(\phi_0^b, \phi_0^b) - \rho g \langle \phi^i, \mathbf{u}_0^h \cdot \mathbf{n}_f \rangle_\Gamma \\ &\leq (\mathbf{f}_f, \mathbf{u}_0^h) - \rho g \langle \phi^i, \mathbf{u}_0^h \cdot \mathbf{n}_f \rangle_\Gamma \\ &\leq \|\mathbf{f}_f\|_* \|\mathbf{D}(\mathbf{u}_0^h)\|_{0,\Omega_f} + \rho g C_r \tilde{C}_t C_k \|f_p\|_* \|\mathbf{D}(\mathbf{u}_0^h)\|_{0,\Omega_f} \\ &\leq C_1 (\|\mathbf{f}_f\|_* + \|f_p\|_*) \|\mathbf{D}(\mathbf{u}_0^h)\|_{0,\Omega_f}, \end{aligned} \quad (3.27)$$

which yields (3.7).

Next, we verify the initial condition (3.8). Let $\mathbf{u}_h, p_h, \phi_h$ be the solution of the finite element scheme (2.6)-(2.7). For convenience, set $\mathbf{e}_0^u = \mathbf{u}_h - \mathbf{u}_h^0, e_0^p = p_h - p_h^0, e_0^\phi = \phi_h - \phi_h^0$.

Subtracting (3.23)-(3.25) from (2.6)-(2.7) gives the error equations:

$$a_f(\mathbf{e}_0^u, \mathbf{v}_h) + a_p(e_0^\phi, \psi_h) + a_\Gamma(\mathbf{e}_0^u, \psi_h; \mathbf{v}_h, e_0^\phi) + b(\mathbf{v}_h, e_0^p) + c(\mathbf{u}_h; \mathbf{u}_h, \mathbf{v}_h) = 0, \quad (3.28)$$

$$b(\mathbf{e}_0^u, q_h) = 0, \quad (3.29)$$

for any $\mathbf{v}_h \in \mathbf{X}_{f,h}, \psi_h \in X_{p,h}$ and $q_h \in Q_h$.

Taking $\mathbf{v}_h = \mathbf{e}_0^u, \psi_h = 0$ in (3.28) and noting that $b(\mathbf{e}_0^u, e_0^p) = 0$, we have

$$\nu \|\mathbf{D}(\mathbf{e}_0^u)\|_{0,\Omega_f}^2 = -\rho g \langle e_0^\phi, \mathbf{e}_0^u \cdot \mathbf{n}_f \rangle_\Gamma - c(\mathbf{u}_h; \mathbf{u}_h, \mathbf{e}_0^u). \quad (3.30)$$

Now, it suffices to estimate the terms on the right side of (3.30). Taking $v_h = \mathbf{0}$, $\psi_h = e_0^\varphi$ in (3.28) yields that

$$\rho g \langle e_0^u \cdot \mathbf{n}_f, e_0^\varphi \rangle_\Gamma = a_p(e_0^\varphi, e_0^\varphi) \geq 0. \quad (3.31)$$

Using the estimates (2.8) and (3.12) and the assumption (3.9), we have

$$\begin{aligned} |c(\mathbf{u}_h; \mathbf{u}_h, e_0^u)| &\leq \mathcal{N}_d \|\mathbf{D}(\mathbf{u}_h)\|_{0,\Omega_f}^2 \|\mathbf{D}(e_0^u)\|_{0,\Omega_f} \\ &\leq \frac{9\mathcal{N}_d C_1^2}{4v^2} (\|\mathbf{f}_f\|_* + \|\mathbf{f}_p\|_*)^2 \|\mathbf{D}(e_0^u)\|_{0,\Omega_f} \\ &\leq \frac{9\mathcal{N}_d C_1^2}{4v^2} \cdot \frac{v^2}{9C_1 \mathcal{N}_d} (\|\mathbf{f}_f\|_* + \|\mathbf{f}_p\|_*) \|\mathbf{D}(e_0^u)\|_{0,\Omega_f} \\ &\leq \frac{1}{4} C_1 (\|\mathbf{f}_f\|_* + \|\mathbf{f}_p\|_*) \|\mathbf{D}(e_0^u)\|_{0,\Omega_f}. \end{aligned} \quad (3.32)$$

By the estimates (3.31) and (3.32),

$$\|\mathbf{D}(e_0^u)\|_{0,\Omega_f} \leq \frac{C_1}{4v} (\|\mathbf{f}_f\|_* + \|\mathbf{f}_p\|_*),$$

which implies the initial condition (3.8). $\square$

4 The Int-Deep method

In this section, we first design a PINN-type (cf. [28, 29]) deep learning method for the Navier-Stokes Darcy model and then propose the Int-Deep algorithm, where the solution obtained from the DL method serves as an initial function for Newton iterative method.

4.1 Deep learning method

In a DL method, the solution of a PDE is approximated by deep neural networks functions. In this work, we employ the classical residual neural network (ResNet) [21] to approximate the solutions of the Navier-Stokes Darcy model. For this purpose, we first recall its structure. Given an input $\mathbf{x} \in \mathbb{R}^d$, consider the following residual network architecture with skip connection in each layer.

$$\mathbf{h}_0 = \mathbf{V}\mathbf{x}, \quad (4.1)$$

$$\mathbf{g}_l = \sigma(\mathbf{W}_l \mathbf{h}_{l-1} + \mathbf{b}_l), \quad l = 1, \dots, L, \quad (4.2)$$

$$\mathbf{h}_l = \mathbf{h}_{l-1} + \mathbf{g}_l, \quad l = 1, \dots, L, \quad (4.3)$$

$$\boldsymbol{\phi}(\mathbf{x}; \tilde{\boldsymbol{\theta}}) = \mathbf{a}^T \mathbf{h}_L, \quad (4.4)$$

where $\tilde{\boldsymbol{\theta}} = \{\mathbf{W}_l, \mathbf{b}_l, \mathbf{a}\}$, $\mathbf{V} \in \mathbb{R}^{m \times d}$,

$$V(i, j) = \begin{cases} 1 & i = j, \quad j \leq d, \\ 0 & \text{otherwise,} \end{cases} \quad \mathbf{W}_l \in \mathbb{R}^{m \times m}, \quad \mathbf{b}_l \in \mathbb{R}^m,$$

$a \in \mathbb{R}^{m \times n}$, L is the number of layers, n is the number of neurons of output layer, m is the width of the residual blocks and σ is an activation function.

Next, we present the DL method in detail. Following the approach in [5], for the given essential boundary conditions $u = g_u(x)$, $x \in \Gamma_f$, $\varphi(x) = g_\varphi(x)$, $x \in \Gamma_p$, we construct three neural network functions $\phi_u(x; \theta_u)$, $\phi_p(x; \theta_p)$ and $\phi_\varphi(x; \theta_\varphi)$ as follows.

$$\begin{aligned}\phi_u(x; \theta_u) &= B_u(x)N_u(x; \theta_u) + \bar{g}_u(x), \\ \phi_p(x; \theta_p) &= N_p(x; \theta_p), \\ \phi_\varphi(x; \theta_\varphi) &= B_\varphi(x)N_\varphi(x; \theta_\varphi) + \bar{g}_\varphi(x),\end{aligned}$$

where $B_u(x)$ and $B_\varphi(x)$ are two smooth scalar functions which satisfy the conditions $B_u(x)|_{\Gamma_f} = 0$ and $B_\varphi(x)|_{\Gamma_p} = 0$; $N_u(x; \theta_u)$, $N_p(x; \theta_p)$ and $N_\varphi(x; \theta_\varphi)$ are three distinct residual neural network functions; $\bar{g}_u$ is the extension of g_u to Ω_f , and $\bar{g}_\varphi$ is the extension of g_φ to Ω_p . Obviously, all the above neural network functions automatically satisfy the prescribed boundary conditions.

Setting $\theta := \{\theta_u, \theta_p, \theta_\varphi\}$, we respectively approximate u , p and φ by

$$u(x) \approx \phi_u(x; \theta_u), \quad p(x) \approx \phi_p(x; \theta_p), \quad \varphi(x) \approx \phi_\varphi(x; \theta_\varphi),$$

with the network parameters determined by minimizing the log square loss

$$\theta^* = \underset{\theta}{\operatorname{argmin}} \log_2(I(\theta_u, \theta_p, \theta_\varphi) + 1), \quad (4.5)$$

where

$$\begin{aligned}I(\theta_u, \theta_p, \theta_\varphi) &= \gamma_1 \left(|\Omega_f| \mathbb{E}_\xi \left(\left\| -\operatorname{div} T(\phi_u(\xi; \theta_u), \phi_p(\xi; \theta_p)) + \rho(\phi_u(\xi; \theta_u) \cdot \nabla) \phi_u(\xi; \theta_u) - f_f(\xi) \right\|^2 \right. \right. \\ &\quad \left. \left. + |\operatorname{div} \phi_u(\xi; \theta_u)|^2 \right) + |\Omega_p| \mathbb{E}_\eta \left(|f_p(\eta) + \operatorname{div}(K \nabla \phi_\varphi(\eta; \theta_\varphi))|^2 \right) \right) \\ &\quad + \gamma_2 \left(|\Gamma| \mathbb{E}_\zeta \left(|\phi_u(\zeta; \theta_u) \cdot n_f + K \nabla \phi_\varphi(\zeta; \theta_\varphi) \cdot n_f|^2 \right. \right. \\ &\quad \left. \left. + \|K\|_2 \left| (\rho g \phi_\varphi(\zeta; \theta_\varphi) + (T(\phi_u(\zeta; \theta_u), \phi_p(\zeta; \theta_p)) n_f) \cdot n_f) \right|^2 \right. \right. \\ &\quad \left. \left. + \sum_{j=1}^{d-1} \left(|(T(\phi_u(\zeta; \theta_u), \phi_p(\zeta; \theta_p)) n_f) \cdot \tau_j + \frac{\nu \alpha_{BJ}}{\sqrt{\tau_j \cdot \nu K \cdot \tau_j}} \phi_u(\zeta; \theta_u) \cdot \tau_j|^2 \right) \right) \right).\end{aligned}$$

Here γ_1, γ_2 are two positive parameters, $\|\cdot\|$ is the Euclidean norm of a vector, $\|\cdot\|_2$ is the 2-norm of a matrix, and ξ, η, ζ are random variables that are uniformly distributed on Ω_f, Ω_p and Γ , respectively. Furthermore, we use the Monte-Carlo method to approximate the related expectation to derive the discretization of (4.5), which is solved by the stochastic gradient descent (SGD) method or its variants (e.g. Adam [26]).

With these preparations, the DL algorithm for Navier-Stokes Darcy model (2.1)-(2.4) is described as Algorithm 1.

Algorithm 1 Deep learning method for Navier-Stokes Darcy problem.**Input:** the maximum number of training $Epoch$, the learning rate η , sample size N_1 , N_2 and N_3 .**Output:**

$$\mathbf{u}^{DL} = \boldsymbol{\phi}_u(\mathbf{x}; \boldsymbol{\theta}_u^{Epoch}), \quad p^{DL} = \phi_p(\mathbf{x}; \boldsymbol{\theta}_p^{Epoch}), \quad \varphi^{DL} = \phi_\varphi(\mathbf{x}; \boldsymbol{\theta}_\varphi^{Epoch}).$$

Initialization: Let $l=0$, $\eta^0 = \eta$, the default initialization method in PyTorch is used to initialize the parameters of neural networks $\boldsymbol{\theta}_u^0, \boldsymbol{\theta}_p^0, \boldsymbol{\theta}_\varphi^0$.

while $l < Epoch$ **do**

Generate uniformly distributed points $\{\boldsymbol{\xi}_i\}_{i=1}^{N_1}$ in Ω_f , $\{\boldsymbol{\eta}_i\}_{i=1}^{N_2}$ in Ω_p , and $\{\boldsymbol{\zeta}_i\}_{i=1}^{N_3}$ on γ , respectively.

Computing loss:

$$L = \log_2(I_1(\boldsymbol{\theta}_u^l, \boldsymbol{\theta}_p^l, \boldsymbol{\theta}_\varphi^l) + 1), \quad (4.6)$$

where

$$\begin{aligned} & I_1(\boldsymbol{\theta}_u^l, \boldsymbol{\theta}_p^l, \boldsymbol{\theta}_\varphi^l) \\ &= \gamma_1 \left(\frac{|\Omega_f|}{N_1} \sum_{i=1}^{N_1} \left(\left\| -\operatorname{div} \mathbf{T}(\boldsymbol{\phi}_u(\boldsymbol{\xi}_i; \boldsymbol{\theta}_u^l), \phi_p(\boldsymbol{\xi}_i; \boldsymbol{\theta}_p^l)) + \rho(\boldsymbol{\phi}_u(\boldsymbol{\xi}_i; \boldsymbol{\theta}_u^l) \cdot \nabla) \boldsymbol{\phi}_u(\boldsymbol{\xi}_i; \boldsymbol{\theta}_u^l) - \mathbf{f}_f(\boldsymbol{\xi}_i) \right\|^2 \right. \right. \\ & \quad \left. \left. + \left| \operatorname{div} \boldsymbol{\phi}_u(\boldsymbol{\xi}_i; \boldsymbol{\theta}_u^l) \right|^2 \right) + \frac{|\Omega_p|}{N_2} \sum_{i=1}^{N_2} \left(\left| f_p(\boldsymbol{\eta}_i) + \operatorname{div}(\mathbf{K} \nabla \phi_\varphi(\boldsymbol{\eta}_i; \boldsymbol{\theta}_\varphi^l)) \right|^2 \right) \right) \\ & \quad + \gamma_2 \left(\frac{|\Gamma|}{N_3} \sum_{i=1}^{N_3} \left(\left| \boldsymbol{\phi}_u(\boldsymbol{\zeta}_i; \boldsymbol{\theta}_u^l) \cdot \mathbf{n}_f + \mathbf{K} \nabla \phi_\varphi(\boldsymbol{\zeta}_i; \boldsymbol{\theta}_\varphi^l) \cdot \mathbf{n}_f \right|^2 \right. \right. \\ & \quad \left. \left. + \left\| \mathbf{K} \right\|_2 \left(\rho g \phi_\varphi(\boldsymbol{\zeta}_i; \boldsymbol{\theta}_\varphi^l) + (\mathbf{T}(\boldsymbol{\phi}_u(\boldsymbol{\zeta}_i; \boldsymbol{\theta}_u^l), \phi_p(\boldsymbol{\zeta}_i; \boldsymbol{\theta}_p^l)) \mathbf{n}_f) \cdot \mathbf{n}_f \right) \right)^2 \right. \\ & \quad \left. + \sum_{j=1}^{d-1} \left(\left| (\mathbf{T}(\boldsymbol{\phi}_u(\boldsymbol{\zeta}_i; \boldsymbol{\theta}_u^l), \phi_p(\boldsymbol{\zeta}_i; \boldsymbol{\theta}_p^l)) \mathbf{n}_f) \cdot \boldsymbol{\tau}_j + \frac{\nu \alpha_{BJ}}{\sqrt{\boldsymbol{\tau}_j \cdot \nu \mathbf{K} \cdot \boldsymbol{\tau}_j}} \boldsymbol{\phi}_u(\boldsymbol{\zeta}_i; \boldsymbol{\theta}_u^l) \cdot \boldsymbol{\tau}_j \right|^2 \right) \right). \end{aligned}$$

Updating parameters $\boldsymbol{\theta}_u^{l+1} = \boldsymbol{\theta}_u^l - \eta^l \nabla_{\boldsymbol{\theta}_u} L$, $\boldsymbol{\theta}_p^{l+1} = \boldsymbol{\theta}_p^l - \eta^l \nabla_{\boldsymbol{\theta}_p} L$, $\boldsymbol{\theta}_\varphi^{l+1} = \boldsymbol{\theta}_\varphi^l - \eta^l \nabla_{\boldsymbol{\theta}_\varphi} L$.

Let $l = l + 1$, $\eta^l = \eta^{l-1} \times (0.01^{\frac{1}{100000}})$.

end while

4.2 The Int-Deep method for solving the Navier-Stokes Darcy model

Since Newton iterative method is only locally convergent, it is very challenging to choose reasonably an initial guess. One natural choice is taking the initial guess as the solutions to the Stokes Darcy model (cf. (3.23)-(3.25)). However, the convergence is very sensitive to the physical parameters in the model. Motivated by the ideas in [23], we are tempted to choose $I_h \mathbf{u}^{DL}$ as an initial guess, where $\mathbf{u}^{DL}$ denotes the deep learning solution obtained from Algorithm 1 with few iteration steps, and I_h is the usual nodal interpolation operator (cf. [7, 13]). Similar to [23], this method is referred to as the Int-Deep method for the Navier-Stokes Darcy problem (2.1)-(2.4), and is described as Algorithm 2. It is

worthy to emphasize that the later choice may rely on the so-called frequency principle mentioned in [33, 34], as also emphasized in the introduction part. In particular, in the appendix we present an one-dimensional numerical experiment to demonstrate that the numerical solution of the PINN-type method is tempted to capture the low-frequency components of the exact solution in the first few training steps, which can be used as a good initial guess of the Newton's method, so that the later method can further capture high-frequency components of the exact solution and converges rapidly. Furthermore, our numerical results in this paper also show the superiority of such a choice for solving the Navier-Stokes Darcy model.

Algorithm 2 Int-Deep method for Navier-Stokes Darcy problem.

Input: the target accuracy ϵ , the maximum number of iterations $N_{\max}$, the approximate solution in a form of a DNN $\mathbf{u}^{DL}$.

Output: $\mathbf{u}_{ID}^h = \mathbf{u}_n^h$, $p_{ID}^h = p_n^h$, $\varphi_{ID}^h = \varphi_n^h$.

Initialization: Let $\mathbf{u}_0^h = I_h \mathbf{u}^{DL}$, $n = 1$, and $e_0 = 1$.

while $e_{n-1} > \epsilon$ and $n \leq N_{\max}$ **do**

find $\mathbf{u}_n^h \in \mathbf{X}_{f,h}$, $p_n^h \in Q_h$, and $\varphi_n^h \in X_p$ such that

$$\begin{aligned} & a_f(\mathbf{u}_n^h, \mathbf{v}_h) + b(\mathbf{v}_h, p_n^h) + \rho g \langle \phi_n^h, \mathbf{v}_h \cdot \mathbf{n}_f \rangle_\Gamma + c(\mathbf{u}_n^h; \mathbf{u}_{n-1}^h, \mathbf{v}_h) + c(\mathbf{u}_{n-1}^h; \mathbf{u}_n^h, \mathbf{v}_h) \\ & = (f_f, \mathbf{v}_h) + c(\mathbf{u}_{n-1}^h; \mathbf{u}_{n-1}^h, \mathbf{v}_h), \\ & b(\mathbf{u}_n^h, q_h) = 0, \\ & a_p(\varphi_n^h, \psi_h) = \rho g \langle \psi_h, \mathbf{u}_n^h \cdot \mathbf{n}_f \rangle_\Gamma + \rho g(f_p, \psi_h). \end{aligned}$$

$$e_n = \max \left(\frac{\|\mathbf{u}_n^h - \mathbf{u}_{n-1}^h\|_0, \Omega_f}{\|\mathbf{u}_{n-1}^h\|_0, \Omega_f}, \frac{\|\varphi_n^h - \varphi_{n-1}^h\|_0, \Omega_p}{\|\varphi_{n-1}^h\|_0, \Omega_p}, \frac{\|p_n^h - p_{n-1}^h\|_0, \Omega_f}{\|p_{n-1}^h\|_0, \Omega_f} \right), n = n + 1.$$

end while

5 Numerical experiments

In this section, we shall present some numerical experiments to illustrate the performance of Algorithm 1 and Algorithm 2 for Navier-Stokes Darcy problem.

The specific details about the implementation of algorithms are provided in advance. The activation function in all neural networks is taken as $\sigma = \max\{x^3, 0\}$. During the training process, we train the neural networks using the Adam optimizer [26] with a learning rate of $\eta = 5e-03$. In Algorithm 1, we use ResNets of width 50, depth 5 for the Navier-Stokes problem and width 50, depth 6 for the Darcy problem. The batch size is taken as: $N_1 = N_2 = 1024$, $N_3 = 256$ for all 2D examples; $N_1 = N_2 = 3000$, $N_3 = 512$ for the

3D example. In Algorithm 2, we set $N_{\max} = 20$, $\epsilon = 1e-07$ for all examples.

Some notations in the following part are summarized here. In numerical experiments, different strategies are employed to generate initial guesses for Newton iterative method, and corresponding results will be compared with the Int-Deep algorithm. We will refer to $\mathbf{u}_0^{DL}$ as the initial guess $I_h \mathbf{u}^{DL}$ where $\mathbf{u}^{DL}$ is provided by Algorithm 1 with 200 epochs in the Adam, $\mathbf{u}_0^{SD}$ as the initial guess obtained by solving the discretized Stokes Darcy problem (3.23)-(3.25), and C means a finite element function whose vector components under the nodal basis functions are taken as a constant C . The number of epochs in the Adam for DL method is denoted by #Epoch and the number of iterations in Newton iterative method is denoted by #K. The accuracy of the solution of Algorithm 1 is measured by the discrete maximum norm: for any $v \in C(\bar{\Omega})$,

$$\|v\|_{0, \infty, h} = \max_{x \in \Omega_h} |v(x)|,$$

where Ω_h is the set of all vertices in $\mathcal{T}_h$. We measure the accuracy of solution of Algorithm 2 by the relative L^2 -error and H^1 -error:

$$\begin{aligned} \|e_u^h\|_0 &= \frac{\|u - u_{ID}^h\|_{0, \Omega_f}}{\|u\|_{0, \Omega_f}}, \quad \|e_p^h\|_0 = \frac{\|p - p_{ID}^h\|_{0, \Omega_f}}{\|p\|_{0, \Omega_f}}, \quad \|e_h^{\varphi}\|_0 = \frac{\|\varphi_h - \varphi_{ID}^h\|_{0, \Omega_p}}{\|\varphi\|_{0, \Omega_p}}, \\ \|e_u^h\|_1 &= \frac{|u - u_{ID}^h|_{1, \Omega_f}}{|u|_{1, \Omega_f}}, \quad \|e_h^{\varphi}\|_1 = \frac{|\varphi - \varphi_{ID}^h|_{1, \Omega_p}}{|\varphi|_{1, \Omega_p}}. \end{aligned}$$

5.1 A 2D-example with a closed-form solution

In this example, we will verify the convergence and robustness of our algorithms. This example is adapted from an example in Subsection 9.2 of [15] so that the required interface conditions are fulfilled. Moreover, the external force $\mathbf{f}_f$ and the source term f_p are independent of the parameters ν and K . The details are given as follows.

We consider the coupled problem (2.1)-(2.4) in $\Omega \in \mathbb{R}^2$ with $\Omega_f = (0, 1) \times (1, 2)$, $\Omega_p = (0, 1) \times (0, 1)$ and the interface $\Gamma = (0, 1) \times \{1\}$. The exact solution is given by

$$\begin{cases} u = y^2 - 2y + 1 + (y-1)(1-2x) & \text{in } \Omega_f, \\ v = x^2 - x + (y-1)^2 & \text{in } \Omega_f, \\ p = 2\nu(x + 2(y-1)) + \frac{1}{3\kappa}, & \text{in } \Omega_f, \\ \varphi = \frac{1}{\kappa}(x(1-x)(y-1) + \frac{y^3}{3} - y^2 + y) + \frac{2\nu x}{g} & \text{in } \Omega_p, \end{cases}$$

where the components of $\mathbf{u}$ are denoted by (u, v) for convenience. For simplicity, all the parameters except ν and κ ($K = \kappa I$) in the coupled model are set to be 1. In (4.6), we set $\gamma_1 = 100$, $\gamma_2 = 1$; $B_u(\mathbf{x}) = x(x-1)(y-2)$ and $B_\varphi(\mathbf{x}) = xy(x-1)$. To evaluate the test error of DL method, we adopt a mesh size $h = \frac{1}{1024}$ to generate the uniform triangulation $\mathcal{T}_h$ as the test locations.

5.1.1 The performance of Algorithm 1

In this part, we focus on the performance of Algorithm 1 and the results are presented in Tables 1-4. We denote the components of $\mathbf{u}^{DL}$ obtained by Algorithm 1 as (u^{DL}, v^{DL}) . Here are some observations.

1. As the viscosity coefficient ν decreases, the absolute discrete maximum errors of all unknowns can reach $\mathcal{O}(10^{-3})$ after 10000 epochs but it is difficult to further improve the accuracy with the increase of the iterations. When both the viscosity coefficient ν and hydraulic conductivity κ are set to be extremely small, the errors of pressure p and hydraulic head φ can only reach $\mathcal{O}(10^{-1})$ or $\mathcal{O}(10^{-2})$, and the error of velocity $\mathbf{u}$ can still reach $\mathcal{O}(10^{-3})$ after 10000 epochs.
2. For different values of viscosity coefficient ν and hydraulic conductivity κ , the absolute discrete maximum errors of velocity $\mathbf{u}$ can reach $\mathcal{O}(10^{-1})$ or $\mathcal{O}(10^{-2})$ after 200 epochs. Therefore, the approximate velocity $\mathbf{u}^{DL}$ generate by Algorithm 1 with $\mathcal{O}(100)$ iterations can serve as a good initial guess for Algorithm 2.

Table 1: The absolute discrete maximum errors of Algorithm 1 with $\nu=1$, $\kappa=1$.

#Epoch	$\ u - u^{DL}\ _{0, \infty, h}$	$\ v - v^{DL}\ _{0, \infty, h}$	$\ \kappa*(p - p^{DL})\ _{0, \infty, h}$	$\ \kappa*(\varphi - \varphi^{DL})\ _{0, \infty, h}$
200	2.5287e-01	2.6556e-01	6.2147e-01	4.4382e-01
500	9.1186e-02	1.1459e-01	5.3835e-01	8.2253e-02
1000	7.1441e-02	8.6823e-02	5.9128e-01	1.8872e-02
1500	3.5708e-02	4.9980e-02	6.2699e-01	5.3279e-03
2000	3.5881e-02	4.4790e-02	6.2020e-01	6.7503e-03
2500	2.1491e-02	3.2071e-02	5.9471e-01	1.0134e-02
3000	1.2654e-02	2.3105e-02	5.2633e-01	4.2378e-03
3500	5.4843e-03	1.4041e-02	2.6764e-01	5.1807e-03
5000	1.0809e-03	3.3421e-03	6.4445e-02	2.1010e-03
7500	1.2335e-03	1.0001e-03	3.8443e-02	7.7657e-04
10000	1.0256e-03	6.9977e-04	3.2053e-02	6.1445e-04
15000	9.9100e-04	4.9650e-04	2.8792e-02	5.7223e-04
20000	1.0007e-03	4.5690e-04	2.8224e-02	5.7376e-04
25000	1.0103e-03	4.5177e-04	2.8161e-02	5.7422e-04
30000	1.0018e-03	4.5314e-04	2.8140e-02	5.7152e-04

5.1.2 The performance of Algorithm 2

In this part, we pay attention to the stability and effectiveness of Algorithm 2. To this end, we first compare the iteration steps #K of Newton iterative method employing different initial guesses. Tables 5-8 show that for various values of ν and κ , the iterative steps of

Table 2: The absolute discrete maximum errors of Algorithm 1 with $\nu=1e-03$, $\kappa=1$.

#Epoch	$\ u - u^{DL}\ _{0, \infty, h}$	$\ v - v^{DL}\ _{0, \infty, h}$	$\ \kappa*(p - p^{DL})\ _{0, \infty, h}$	$\ \kappa*(\varphi - \varphi^{DL})\ _{0, \infty, h}$
200	1.5228e-01	7.8899e-02	1.0821e-01	1.3050e-01
500	3.2690e-02	6.7493e-02	6.9836e-02	3.3034e-02
1000	4.8649e-02	3.3678e-02	1.6692e-02	1.7064e-02
1500	1.5276e-02	1.5477e-02	4.3203e-03	4.0101e-03
2000	1.1154e-02	1.1725e-02	2.5474e-03	1.5258e-03
2500	8.1252e-03	1.0360e-02	1.5529e-03	1.9448e-03
3000	7.5076e-03	9.1052e-03	1.7594e-03	1.0383e-03
3500	6.5818e-03	8.9840e-03	1.8445e-03	8.2254e-04
5000	6.4812e-03	8.1881e-03	1.4697e-03	1.2283e-03
7500	5.9790e-03	7.7442e-03	1.4521e-03	1.1481e-03
10000	5.7681e-03	7.2189e-03	1.3932e-03	1.1389e-03
15000	5.5520e-03	6.7599e-03	1.3696e-03	1.0710e-03
20000	5.4419e-03	6.6387e-03	1.3253e-03	1.0894e-03
25000	5.4207e-03	6.6333e-03	1.2974e-03	1.0759e-03
30000	5.4197e-03	6.6308e-03	1.3030e-03	1.0741e-03

Table 3: The absolute discrete maximum errors of Algorithm 1 with $\nu=1e-05$, $\kappa=1$.

#Epoch	$\ u - u^{DL}\ _{0, \infty, h}$	$\ v - v^{DL}\ _{0, \infty, h}$	$\ \kappa*(p - p^{DL})\ _{0, \infty, h}$	$\ \kappa*(\varphi - \varphi^{DL})\ _{0, \infty, h}$
200	1.2267e-01	2.6950e-01	5.2884e-01	2.7065e-01
500	7.2845e-02	3.1214e-02	1.0978e-01	4.3782e-02
1000	8.4428e-02	6.1127e-02	1.0669e-02	2.3048e-02
1500	4.5015e-02	4.0511e-02	1.2635e-02	5.0194e-03
2000	2.6260e-02	2.7103e-02	6.4048e-03	4.2232e-03
2500	2.0447e-02	2.2834e-02	4.5544e-03	3.4207e-03
3000	1.8230e-02	2.1036e-02	4.3835e-03	3.2771e-03
3500	1.7129e-02	1.9659e-02	4.3157e-03	3.4534e-03
5000	1.4332e-02	1.7132e-02	3.4013e-03	2.5777e-03
7500	1.1559e-02	1.4041e-02	2.7369e-03	2.1010e-03
10000	1.0130e-02	1.2115e-02	2.3879e-03	1.7258e-03
15000	8.8846e-03	1.0759e-02	2.1902e-03	1.4898e-03
20000	8.6222e-03	1.0508e-02	2.1309e-03	1.4656e-03
25000	8.6160e-03	1.0461e-02	2.1178e-03	1.4720e-03
30000	8.6120e-03	1.0458e-02	2.1204e-03	1.4664e-03

Table 4: The absolute discrete maximum errors of Algorithm 1 with $\nu=1e-05$, $\kappa=1e-08$.

#Epoch	$\ u-u^{DL}\ _{0,\infty,h}$	$\ v-v^{DL}\ _{0,\infty,h}$	$\ \kappa*(p-p^{DL})\ _{0,\infty,h}$	$\ \kappa*(\varphi-\varphi^{DL})\ _{0,\infty,h}$
200	4.3039e-02	3.2012e-02	3.3333e-01	1.0108e-08
500	1.4437e-02	1.2377e-02	3.3333e-01	9.6933e-09
1000	3.0128e-03	3.5399e-03	3.3333e-01	9.7415e-09
1500	4.0819e-03	4.4276e-03	3.3333e-01	3.4149e-07
2000	1.9469e-03	1.9256e-03	3.3333e-01	2.0512e-04
2500	2.0033e-03	1.5658e-03	3.3333e-01	1.3423e-04
3000	1.7536e-03	3.2420e-03	3.3333e-01	3.8973e-02
3500	2.1632e-03	3.0824e-03	3.3333e-01	3.9326e-02
5000	1.2825e-03	2.8897e-03	3.3333e-01	3.9214e-02
7500	1.2791e-03	3.1104e-03	3.3333e-01	3.8878e-02
10000	1.4466e-03	3.1671e-03	3.3333e-01	3.8691e-02
15000	1.5980e-03	3.3951e-03	3.3333e-01	3.8457e-02
20000	1.6034e-03	3.3725e-03	3.3333e-01	3.8464e-02
25000	1.6084e-03	3.4061e-03	3.3333e-01	3.8458e-02
30000	1.6141e-03	3.4001e-03	3.3333e-01	3.8457e-02

Table 5: The number of iterations #K under different initial guesses with $\nu=1$, $\kappa=1$.

Initial value \ Mesh size	$h=\frac{1}{70}$	$h=\frac{1}{80}$	$h=\frac{1}{90}$	$h=\frac{1}{100}$
u_0^{SD}	4	4	4	4
0	5	5	5	5
1	5	5	5	5
u_0^{DL}	4	4	4	4

Table 6: The number of iterations #K under different initial guesses with $\nu=1e-03$, $\kappa=1$.

Initial value \ Mesh size	$h=\frac{1}{70}$	$h=\frac{1}{80}$	$h=\frac{1}{90}$	$h=\frac{1}{100}$
u_0^{SD}	×	×	×	×
0	×	×	×	×
1	7	7	7	7
u_0^{DL}	6	6	6	6

Table 7: The number of iterations #K under different initial guesses with $\nu = 1e-05$, $\kappa = 1$.

Initial value \ Mesh size	Mesh size			
	$h = \frac{1}{70}$	$h = \frac{1}{80}$	$h = \frac{1}{90}$	$h = \frac{1}{100}$
u_0^{SD}	×	×	×	×
0	×	×	×	×
1	×	×	×	×
u_0^{DL}	7	7	7	7

Table 8: The number of iterations #K under different initial guesses with $\nu = 1e-05$, $\kappa = 1e-08$.

Initial value \ Mesh size	Mesh size			
	$h = \frac{1}{70}$	$h = \frac{1}{80}$	$h = \frac{1}{90}$	$h = \frac{1}{100}$
u_0^{SD}	×	×	×	×
0	×	×	×	×
1	×	×	×	×
u_0^{DL}	6	6	6	5

Algorithm 2 are less than or equal to that of Newton iterative method with other initial guesses, and the Newton iterative method with other initial guesses even fails in convergence. Furthermore, when the parameter ν or κ is extremely small, only Algorithm 2 is convergent.

Remark 5.1. "×" indicates that the iterative method does not converge within $N_{\max} = 20$ steps.

Next, we investigate the accuracy of Algorithm 2 for different values of viscosity coefficient ν and hydraulic conductivity κ . The relative errors for u , p and φ in L^2 and H^1 norms are displayed in Tables 9-12. Here are our observations.

1. Regarding the Navier-Stokes part, with different values of ν and κ , the relative error in L^2 norm of the velocity u is of the order $\mathcal{O}(h^3)$ or higher, and in H^1 norm, the velocity u has an order of $\mathcal{O}(h^2)$ or higher. Besides, The relative error in L^2 norm of the pressure p is of the order $\mathcal{O}(h^3)$ or higher for different values of ν and κ .
2. For the Darcy part, we observe that the relative errors for φ in L^2 norm are of order $\mathcal{O}(h^3)$ and the relative errors of φ in H^1 norm are of order $\mathcal{O}(h^2)$, which is in line with our expectations. These results hold for different values of ν and κ .

5.2 A 2D-example without closed form solution

In this subsection, we test the Algorithm 2 by an example which is adapted from an example in [19] so that it satisfies our boundary conditions. In addition, we also consider

Table 9: The errors of Algorithm 2 with $\nu=1$, $\kappa=1$.

	$\ e_u^h\ _0$	order	$\ e_p^h\ _0$	order	$\ e_\phi^h\ _0$	order	$\ e_u^h\ _1$	order	$\ e_\phi^h\ _1$	order
$h=\frac{1}{70}$	4.2609e-11	-	3.1352e-10	-	4.1530e-08	-	5.3915e-10	-	1.1848e-05	-
$h=\frac{1}{80}$	2.5209e-11	3.9306	1.8491e-10	3.9539	2.7827e-08	2.9986	3.3530e-10	3.5569	9.0728e-06	1.9989
$h=\frac{1}{90}$	1.5872e-11	3.9280	1.1583e-10	3.9715	1.9547e-08	2.9988	2.2062e-10	3.5539	7.1694e-06	1.9991
$h=\frac{1}{100}$	1.0458e-11	3.9599	7.6790e-11	3.9012	1.4251e-08	2.9989	1.5172e-10	3.5535	5.8077e-06	1.9992

Table 10: The errors of Algorithm 2 with $\nu=1e-03$, $\kappa=1$.

	$\ e_u^h\ _0$	order	$\ e_p^h\ _0$	order	$\ e_\phi^h\ _0$	order	$\ e_u^h\ _1$	order	$\ e_\phi^h\ _1$	order
$h=\frac{1}{70}$	9.7483e-09	-	1.8392e-08	-	2.4446e-07	-	4.7831e-07	-	7.2070e-05	-
$h=\frac{1}{80}$	5.7736e-09	3.9227	1.0901e-08	3.9170	1.6379e-07	2.9989	3.0118e-07	3.4639	5.5187e-05	1.9989
$h=\frac{1}{90}$	3.6378e-09	3.9217	6.8692e-09	3.9208	1.1505e-07	2.9990	2.0013e-07	3.4701	4.3609e-05	1.9991
$h=\frac{1}{100}$	2.4060e-09	3.9237	4.5447e-09	3.9206	8.3879e-08	2.9991	1.3878e-07	3.4749	3.5326e-05	1.9992

Table 11: The errors of Algorithm 2 with $\nu=1e-05$, $\kappa=1$.

	$\ e_u^h\ _0$	order	$\ e_p^h\ _0$	order	$\ e_\phi^h\ _0$	order	$\ e_u^h\ _1$	order	$\ e_\phi^h\ _1$	order
$h=\frac{1}{70}$	8.7243e-08	-	7.2017e-08	-	2.4602e-07	-	1.1941e-05	-	7.2072e-05	-
$h=\frac{1}{80}$	5.3363e-08	3.6814	4.3147e-08	3.8365	1.6474e-07	3.0033	8.1309e-06	2.8779	5.5188e-05	1.9989
$h=\frac{1}{90}$	3.4379e-08	3.7328	2.7386e-08	3.8594	1.1566e-07	3.0027	5.7670e-06	2.9165	4.3610e-05	1.9991
$h=\frac{1}{100}$	2.3102e-08	3.7732	1.8225e-08	3.8650	8.4300e-08	3.0022	4.2260e-06	2.9509	3.5327e-05	1.9992

Table 12: The errors of Algorithm 2 with $\nu=1e-05$, $\kappa=1e-08$.

	$\ e_u^h\ _0$	order	$\ e_p^h\ _0$	order	$\ e_\phi^h\ _0$	order	$\ e_u^h\ _1$	order	$\ e_\phi^h\ _1$	order
$h=\frac{1}{70}$	1.1704e-05	-	1.6194e-07	-	2.6092e-07	-	2.6384e-03	-	7.2153e-05	-
$h=\frac{1}{80}$	7.9024e-06	2.9415	1.0636e-07	3.1485	1.7443e-07	3.0159	2.0308e-03	1.9603	5.5242e-05	2.0000
$h=\frac{1}{90}$	5.5862e-06	2.9450	7.3426e-08	3.1458	1.2229e-07	3.0152	1.6111e-03	1.9653	4.3648e-05	2.0000
$h=\frac{1}{100}$	4.1012e-06	2.9330	5.2725e-08	3.1434	8.9011e-08	3.0145	1.3101e-03	1.9630	3.5355e-05	2.0000

the situation where the parameter ν takes small values.

In the free fluid region $\Omega_f = (0, 2) \times (0, 1)$, the lid-driven problem is considered. The fluid is mainly driven by a rightward velocity ($\mathbf{u} = [1, 0]^T$) on the top side ($y = 1$), and no-slip boundary condition ($\mathbf{u} = \mathbf{0}$) is imposed on the left- and right-sides. There is no body force ($\mathbf{f}_f = \mathbf{0}$). We take the parameter $\nu = 1, 0.1, 0.01$.

In the porous media region $\Omega_p = (0, 2) \times (-1, 0)$. A heterogeneous permeability $\mathbf{K} = \kappa \mathbf{I}$ is taken. In three blocks: $(0.2, 0.6) \times (-0.8, -0.6)$, $(0.8, 1.2) \times (-0.7, -0.5)$ and $(1.4, 1.8) \times (-0.6, -0.4)$, the permeability κ is set to be $1e-06$. For the remaining region, $\kappa = 1$. There is no source ($f_p = 0$). And $\phi = 0$ is imposed on the left-, right-, and bottom-sides of the region.

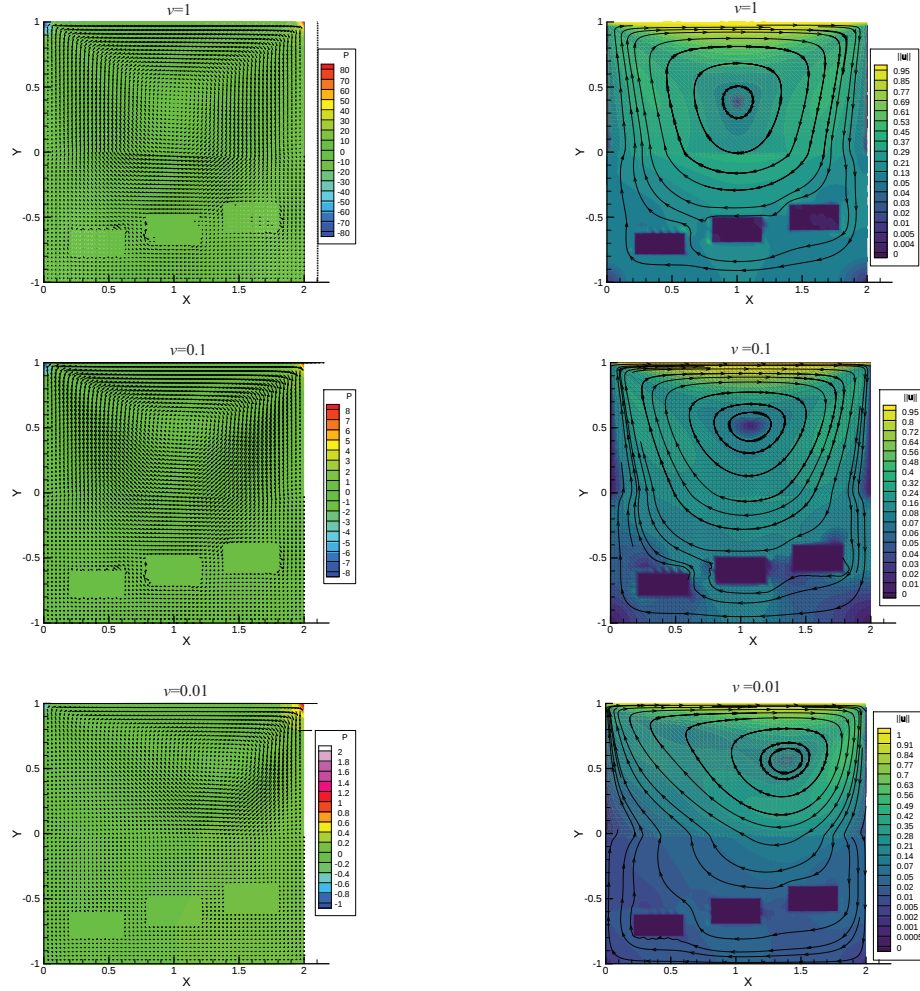


Figure 2: (Left) Numerical velocity and pressure under different viscosity coefficients ν on a triangular mesh with $h = \frac{1}{16}$; (Right) The streamlines of velocity.

For simplicity, all the parameters except ν and κ in the coupled model are set to 1. Besides, in (4.6), we take $\gamma_1 = 200$, $\gamma_2 = 1$, $B_u(x) = x(x-2)(y-1)$ and $B_\phi(x) = x(x-2)(y+1)$.

Since there is no exact solution for comparison, we focus on corresponding physical phenomena. From Fig. 2, we can clearly observe that there is a smooth exchange of flow between the free flow (Navier-Stokes) and porous-medium flow (Darcy) across the interface ($y = 0$). In Navier-Stokes region, singular pressure occurs at the two corners $(0, 1)$ and $(2, 1)$ and in Darcy region, the flow path is diverted around three low-permeability blocks. With the vary of ν , the fluid exhibits distinct flow patterns. For $\nu = 1$ and $\nu = 0.1$, fluid travels from the Navier-Stokes region to the Darcy region when $x > 1$, and from the Darcy region to the Navier-Stokes region when $x < 1$. For $\nu = 0.01$, fluid travels the

Navier-Stokes region to the Darcy region when $x > 1.3$, and from the Darcy region to the Navier-Stokes region when $x < 1.3$.

5.3 A 3D-example with closed form solution

Finally, we test our algorithms for solving a three-dimensional problem with closed form solution. Let $\Omega \subset \mathbb{R}^3$ with $\Omega_f = (0, 1) \times (0, 1) \times (0, 1)$, $\Omega_p = (0, 1) \times (0, 1) \times (1, 2)$ and the interface $\Gamma = (0, 1) \times \{1\}$. The exact solution (u, φ, p) is given by

$$\begin{cases} u = -(1-y)(1-z) & \text{in } \Omega_f, \\ v = -(1-x)(1-z) & \text{in } \Omega_f, \\ w = (1-x)(1-y) & \text{in } \Omega_f, \\ p = (1-z)(1-x-y-z+4xyz) & \text{in } \Omega_f, \\ \varphi = (1-x)(1-y)(1-z) & \text{in } \Omega_p, \end{cases}$$

where the components of $\mathbf{u}$ is denoted by (u, v, w) . All the parameters except ν in the coupled model are set to be 1. In (4.6), we set $\gamma_1 = 1$, $\gamma_2 = 1$, $B_u(x) = xyz(x-1)(y-1)$ and $B_\varphi(x) = xy(x-1)(y-1)(z-2)$. The uniform triangulation $\mathcal{T}_h$ with mesh size $h = \frac{1}{32}$ is adopted to be the test locations.

Similar to the 2D situation, we denote the components of $\mathbf{u}^{DL}$ obtained by Algorithm 1 as (u^{DL}, v^{DL}, w^{DL}) and test the computational performance of Algorithm 1 and Algorithm 2. From Tables 13-14, we observe that the absolute discrete maximum errors of u, v, w, p and φ can reach $\mathcal{O}(10^{-3})$ or $\mathcal{O}(10^{-4})$. However, decreasing the iteration

Table 13: The absolute discrete maximum errors of Algorithm 1 with $\nu=1$, $\kappa=1$.

#Epoch	$\ u - u^{DL}\ _{0, \infty, h}$	$\ v - v^{DL}\ _{0, \infty, h}$	$\ w - w^{DL}\ _{0, \infty, h}$	$\ p - p^{DL}\ _{0, \infty, h}$	$\ \varphi - \varphi^{DL}\ _{0, \infty, h}$
200	1.1280e-01	1.1474e-01	1.2573e-01	9.8521e-01	1.9437e-01
500	5.8955e-03	6.0989e-03	4.6234e-03	1.8171e-01	7.3369e-03
1000	3.9364e-03	4.6433e-03	3.1881e-03	6.4626e-02	4.4960e-03
1500	8.2522e-03	8.5444e-03	8.1683e-03	2.9225e-02	2.2214e-03
2000	1.0683e-03	1.0316e-03	7.5461e-04	1.5038e-02	9.4810e-04
2500	2.4907e-03	2.4303e-03	2.6002e-03	1.2858e-02	8.1199e-04
2500	2.4907e-03	2.4303e-03	2.6002e-03	1.2858e-02	8.1199e-04
3000	7.7432e-04	1.1575e-03	1.1627e-03	8.1957e-03	1.2877e-03
3500	4.7258e-04	6.2618e-04	4.2063e-04	6.4445e-03	8.3823e-04
5000	3.6597e-04	6.0589e-04	3.5250e-04	7.9574e-03	6.8734e-04
7500	2.8277e-04	4.7100e-04	3.0724e-04	3.4871e-03	4.7562e-04
10000	1.3952e-04	3.0117e-04	2.9271e-04	3.3414e-03	3.3065e-04
15000	1.6717e-04	2.4263e-04	2.3134e-04	2.9260e-03	2.9287e-04
20000	1.3941e-04	2.2699e-04	2.1378e-04	2.5956e-03	2.9260e-04

Table 14: The absolute discrete maximum errors of Algorithm 1 with $\nu=0.01$, $\kappa=1$.

#Epoch	$\ u - u^{DL}\ _{0, \infty, h}$	$\ v - v^{DL}\ _{0, \infty, h}$	$\ w - w^{DL}\ _{0, \infty, h}$	$\ p - p^{DL}\ _{0, \infty, h}$	$\ \varphi - \varphi^{DL}\ _{0, \infty, h}$
200	4.9421e-02	3.9340e-02	1.1578e-01	3.6239e-02	4.6330e-02
500	2.6120e-02	1.6990e-02	3.8038e-02	8.8871e-03	1.4511e-02
1000	1.7341e-02	1.1687e-02	2.1754e-02	8.6221e-03	8.6668e-03
1500	1.3706e-02	7.0937e-03	1.6307e-02	2.8401e-03	4.4740e-03
2000	1.2608e-02	5.5066e-03	1.3255e-02	2.1570e-03	3.2142e-03
2500	9.4278e-03	5.1800e-03	1.1713e-02	2.0543e-03	2.5913e-03
2500	9.4278e-03	5.1800e-03	1.1713e-02	2.0543e-03	2.5913e-03
3000	8.6336e-03	3.9181e-03	9.9709e-03	2.0516e-03	2.3925e-03
3500	7.9073e-03	3.6801e-03	9.5410e-03	2.6332e-03	1.8229e-03
5000	5.8340e-03	3.3927e-03	7.2748e-03	7.1876e-04	1.3675e-03
7500	3.9786e-03	3.3174e-03	5.5008e-03	3.2889e-04	1.0128e-03
10000	3.3260e-03	2.7889e-03	4.4905e-03	2.9798e-04	8.3913e-04
15000	2.7782e-03	2.3925e-03	3.9075e-03	1.8952e-04	6.8426e-04
20000	2.6754e-03	2.3359e-03	3.8142e-03	2.1757e-04	7.0270e-04

Table 15: The number of iterations #K under different initial values with $\nu=1$, $\kappa=1$.

Initial value \ Mesh size			
	$h = \frac{1}{4}$	$h = \frac{1}{8}$	$h = \frac{1}{16}$
u_0^{SD}	4	4	4
0	5	5	5
1	5	5	5
u_0^{DL}	4	4	4

Table 16: The number of iterations #K under different initial values with $\nu=0.01$, $\kappa=1$.

Initial value \ Mesh size			
	$h = \frac{1}{4}$	$h = \frac{1}{8}$	$h = \frac{1}{16}$
u_0^{SD}	7	7	7
0	8	8	8
1	8	8	8
u_0^{DL}	5	5	5

steps cannot further improve accuracy. In order to illustrate the accuracy of the Int-Deep method, we compare the iteration steps #K of Newton iterative method employing different initial guesses with $h = \frac{1}{2^k}$, $k=2, 3, 4$. The corresponding results are presented in Tables 15-16. As we can see, when the viscosity coefficient ν decreases, only the Int-Deep method can converge with few iteration steps.

Table 17: The errors of Algorithm 2 with $\nu=1$, $\kappa=1$.

	$\ e_u^h\ _0$	order	$\ e_p^h\ _0$	order	$\ e_\varphi^h\ _0$	order	$\ e_u^h\ _1$	order	$\ e_\varphi^h\ _1$	order
$h=\frac{1}{4}$	1.2611e-04	-	1.4169e-03	-	7.8653e-02	-	7.5859e-04	-	1.3749e-02	-
$h=\frac{1}{8}$	9.6019e-06	3.7152	1.9505e-02	2.0117	9.3657e-05	3.0179	2.2331e-04	2.6656	3.4647e-03	1.9886
$h=\frac{1}{16}$	7.2685e-07	3.7236	4.8542e-03	2.0065	1.1678e-05	3.0036	3.4100e-05	2.7112	8.6976e-04	1.9940

Table 18: The errors of Algorithm 2 with $\nu=0.01$, $\kappa=1$.

	$\ e_u^h\ _0$	order	$\ e_p^h\ _0$	order	$\ e_\varphi^h\ _0$	order	$\ e_u^h\ _1$	order	$\ e_\varphi^h\ _1$	order
$h=\frac{1}{4}$	1.1766e-02	-	7.8758e-02	-	7.9584e-04	-	1.3575e-01	-	1.3819e-02	-
$h=\frac{1}{8}$	8.8893e-04	3.7264	1.9508e-02	2.0134	9.4602e-05	3.0725	2.1471e-02	2.6605	3.4671e-03	1.9948
$h=\frac{1}{16}$	6.6064e-05	3.7501	4.8543e-03	2.0067	1.1691e-05	3.0165	3.2525e-03	2.7227	8.6985e-04	1.9949

Table 17 and Table 18 show that the orders of relative errors in L^2 norm for u, p and φ are $\mathcal{O}(h^{3.7})$, $\mathcal{O}(h^2)$ and $\mathcal{O}(h^3)$, respectively, and the orders of relative errors in H^1 norm for u and φ are $\mathcal{O}(h^{2.7})$ and $\mathcal{O}(h^2)$, respectively, which is consistent with theoretical results.

6 Conclusion

In this paper, we are concerned with Newton's method and its hybrid with machine learning for solving Navier-Stokes Darcy model. We give the convergence analysis of Newton's method for the coupled nonlinear problem with mixed element discretization. Then, we develop a deep learning initialized iterative (Int-Deep) method for Navier-Stokes Darcy model. In this method, we employ the neural network solution generated by a few training steps as the initial guess for Newton iterative method with the FEM discretized problem. This hybrid method can converge to the true solution with the accuracy of FEM method rapidly. Our numerical experiments also indicate the method is robust with respect to the underlying physical parameters. It is our forthcoming topic to study this property theoretically.

Acknowledgments

J. Huang was partially supported by the China National KeyR&D Project (Grant 2020YFA0709800), the National Natural Science Foundation of China (Grant No. 12071289) and the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDA25010402). H. Peng was supported by the National Natural Science Foundation of China (Grant No. 12301519).

Appendix

In this appendix, we consider frequency analysis of the Int-Deep method for an one-dimensional nonlinear elliptic problem, described as follows.

$$\begin{cases} -u'' + u^2 = f, & 0 < x < 1, \\ u(0) = u(1) = 0. \end{cases}$$

The exact solution is given by

$$u(x) = \sin(\pi x) + \sin(4\pi x)/4 - \sin(8\pi x)/8 + \sin(16\pi x)/16 + \sin(24\pi x)/24,$$

which has low and high-frequency components. The deep learning method is shown in Algorithm 3.

Algorithm 3 Deep learning method.

Input: the maximum number of training *Epoch*, the learning rate η , sample size M .

Output: $u^{DL} = \phi(x, \theta^{Epoch})$.

Initialization: Let $l = 0$, $\eta^0 = \eta$, the default initialization method in PyTorch is used to initialize the parameters of neural networks θ .

while $l < Epoch$ **do**

Generate independent random variables $\{\xi_i\}_{i=1}^M$ that are uniformly distributed on $[0, 1]$.

Computing loss:

$$L = \log_2(l(\theta^l) + 1),$$

where

$$l(\theta^l) = \frac{1}{M} \sum_{i=1}^M (|-\phi''(\xi_i; \theta^l) + \phi^2(\xi_i; \theta^l) - f(\xi_i)|^2).$$

Updating parameters $\theta^{l+1} = \theta^l - \eta^l$.

Let $l = l + 1$, $\eta^l = \eta^{l-1} \times (0.01^{\frac{1}{100000}})$.

end while

In Algorithm 3, $\phi(x; \theta) = x(x-1)N(x, \theta)$, where $N(x, \theta)$ are ResNets of width 50, depth 5. Besides, the activation function is taken as $\sigma = \max\{x^3, 0\}$ and the batch size M is 512.

The details of Newton method are shown as follows.

Divide the interval $[0, 1]$ into N equal parts, with nodes x_i , $i = 0, 1, \dots, N$:

$$0 = x_0 < x_1 < \dots < x_N = 1.$$

Let

$$h = h_j = \frac{1}{N}, \quad e_j = [x_{j-1}, x_j], \quad j = 1, \dots, N.$$

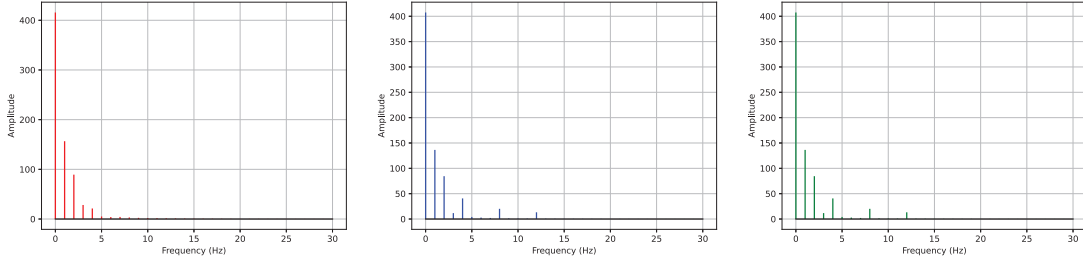


Figure 3: Amplitude spectra of u^{DL} (left), u (middle), and u^{Newton} (right).

Table 19: L^2 errors of u^{Newton} for different mesh sizes.

	Err	order
$h=\frac{1}{80}$	1.4559e-03	-
$h=\frac{1}{160}$	3.4788e-04	2.0161
$h=\frac{1}{320}$	8.6002e-05	2.0652
$h=\frac{1}{640}$	2.1440e-05	2.2716

Construct the finite element space

$$V_h = \{v \in C[0, 1] : v|_{e_i} \in P_1(e_i), v(0) = v(1) = 0\}.$$

The finite element solution $u_h \in V_h$ satisfies

$$a(u_h, v_h) + b(u_h, v_h) = F(v_h) \quad \forall v_h \in V_h,$$

where $a(u, v) = \int_0^1 u'v' dx$, $b(u, v) = \int_0^1 u^2v dx$ and $F(v) = \int_0^1 f v dx$.

At each Newton step n , we need to solve the following linear problem

$$a(u_n^h, v_h) + \tilde{b}(u_n^h, v_h; u_{n-1}^h) = F(v_h) + b(u_{n-1}^h, v_h) \quad \forall v_h \in V_h,$$

where $\tilde{b}(u_n^h, v_h; u_{n-1}^h) = \int_0^1 (2u_{n-1}^h u_n^h v_h) dx$.

Denote by u^{DL} the approximate solution obtained from Algorithm 3 with 200 epochs in the Adam algorithm, and u^{Newton} the solution of Newton's method with $I_h u^{DL}$ as the initial guess. On a uniform grid with $h = \frac{1}{640}$ over the interval $[0, 1]$, we perform a discrete Fourier transform on u^{DL} , u^{Newton} and the exact solution u , and then display their amplitude spectra of the first 31 frequencies in Fig. 3. Additionally, we display the L^2 errors of u^{Newton} for different mesh sizes h in Table 19.

Fig. 3 shows that u^{DL} captures the low-frequency components of the exact solution while u^{Newton} captures not only the previous components but also the high-frequency ones that u^{DL} fails to catch. Table 19 shows that u^{Newton} converges to the exact solution u .

References

- [1] R.A. Adams, Sobolev Spaces, Academic Press, New York, 1975.
- [2] T. Arbogast and D.S. Brunson, A computational method for approximating a Darcy-Stokes system governing a vuggy porous medium, *Comput. Geosci.*, 11 (2007), 207-218.
- [3] T. Arbogast and H.L. Lehr, Homogenization of a Darcy-Stokes system modeling vuggy porous media, *Comput. Geosci.*, 10 (2006), 291-302.
- [4] L. Badea, M. Discacciati, and A. Quarteroni, Numerical analysis of the Navier-Stokes Darcy coupling, *Numer. Math.*, 115 (2010), 195-227.
- [5] J. Berg and K. Nyström, A unified deep artificial neural network approach to partial differential equations in complex geometries, *Neurocomputing*, 317 (2018), 28-41.
- [6] D. Boffi, F. Brezzi, and M. Fortin, *Mixed Finite Element Methods and Applications*, Springer, Berlin, 2013.
- [7] S.C. Brenner and L.R. Scott, *The Mathematical Theory of Finite Element Methods* (Third Edition), Springer, New York, 2008.
- [8] M. Cai, P. Huang, and M. Mu, Some multilevel decoupled algorithms for a mixed Navier-Stokes/Darcy model, *Adv. Comput. Math.*, 44 (2018), 115-145.
- [9] M. Cai, M. Mu, and J. Xu, Numerical solution to a mixed Navier-Stokes Darcy model by the two-grid approach, *SIAM J. Numer. Anal.*, 47 (2009), 3325-3338.
- [10] Y. Cao, M. Gunzburger, X. Hu, F. Hua, X. Wang, and W. Zhao, Finite element approximations for Stokes-Darcy flow with Beavers-Joseph interface conditions, *SIAM J. Numer. Anal.*, 47 (2010), 4239-4256.
- [11] A. Cesmelioglu and S. Rhebergen, A hybridizable discontinuous Galerkin method for the coupled Navier-Stokes and Darcy problem, *J. Comput. Appl. Math.*, 422(2023), Paper No. 114923, 17 pp.
- [12] P. Chidyagwai and B. Rivière, On the solution of the coupled Navier-Stokes and Darcy equations, *Comput. Methods Appl. Mech. Engrg.*, 198 (2009), 3806-3820.
- [13] P.G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.
- [14] M. Discacciati, E. Miglio, and A. Quarteroni, Mathematical and numerical models for coupling surface and groundwater flows, *Appl. Numer. Math.*, 43 (2002), 57-74.
- [15] M. Discacciati and A. Quarteroni, Navier-Stokes/Darcy coupling: modeling, analysis, and numerical, *Rev. Mat. Complut.*, 22 (2009), 315-426.
- [16] G. Du and L. Zuo, Local and parallel partition of unity scheme for the mixed Navier-Stokes-Darcy problem, *Numer. Algorithms*, 91 (2022), 635-650.
- [17] V. Girault and B. Rivière, DG approximation of coupled Navier-Stokes and Darcy equations by Beaver-Joseph-Saffman interface condition, *SIAM J. Numer. Anal.*, 47 (2009), 2052-2089.
- [18] N. Hanspal, A. Waghode, V. Nassehi, and R. Wakeman, Numerical analysis of coupled Stokes/Darcy flow in industrial filtrations, *Transp. Porous Media.*, 64 (2006), 73-101.
- [19] G. Harper, J. Liu, S. Tavener, T. Wildey, Coupling Arbogast-Correa and Bernardi-Raugel elements to resolve coupled Stokes-Darcy flow problems, *Comput. Methods Appl. Mech. Engrg.*, 373 (2021), 113469.
- [20] X. He, J. Li, Y. Lin, and J. Ming, A domain decomposition method for the steady-state Navier-Stokes-Darcy model with Beavers-Joseph interface condition, *SIAM J. Sci. Comput.*, 37 (2015), 264-290.
- [21] K. He, X. Zhang, S. Ren, and J. Sun, Deep Residual Learning for Image Recognition, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, 770-778.

- [22] Y. He and J. Li, Convergence of three iterative methods based on the finite element discretization for the stationary Navier-Stokes equations, *Comput. Methods Appl. Mech. Engrg.*, 198 (2009), 1351-1359.
- [23] J. Huang, H. Wang, and H. Yang, Int-Deep: a deep learning initialized iterative method for nonlinear problems, *J. Comput. Phys.*, 419 (2020), 109675.
- [24] T.J.R. Hughes, *The Finite Element Method, Linear Static and Dynamic Finite Element Analysis*, Prentice-Hall, Englewood Cliffs, N.J, 1987.
- [25] H. Jia, H. Jia, and Y. Huang, A modified two-grid decoupling method for the mixed Navier-Stokes/Darcy model, *Comput. Math. Appl.*, 72 (2016), 1142-1152.
- [26] D.P. Kingma and J. Ba, Adam: A method for stochastic optimization, 2014, arXiv preprint arXiv:1412.6980.
- [27] W. Layton, H. Tran, and C. Trenchea, Analysis of long time stability and errors of two partitioned methods for uncoupling evolutionary groundwater-surface water flows, *SIAM J. Numer. Anal.*, 51 (2013), 248-272.
- [28] L. Lu, X. Meng, Z. Mao, and G. Karniadakis, DeepXDE: A Deep Learning Library for Solving Differential Equations, *SIAM Rev.*, 63 (2021), 208-228.
- [29] M. Raissi, P. Perdikaris, and G. Karniadakis, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, *J. Comput. Phys.*, 378 (2019), 686-707.
- [30] R. Temam, *Navier-Stokes Equations*, North-Holland, Amsterdam, 1979.
- [31] X. Wang, G. Du, L. Zuo, A novel local and parallel finite element method for the mixed Navier-Stokes-Darcy problem, *Comput. Math. Appl.*, 90(2021), 73-79.
- [32] Y. Wu and L. Mei, A non-conforming finite volume element method for the two-dimensional Navier-Stokes/Darcy system, *Comput. Appl. Math.*, 37(2018), 457-474.
- [33] Z. Xu, Y. Zhang, T. Luo, Y. Xiao, and Z. Ma, Frequency principle: Fourier analysis sheds light on deep neural networks, *Commun. Comput. Phys.*, 28 (2020), 1746-1767.
- [34] Z. Xu, Y. Zhang, and Y. Xiao, Training behavior of deep neural network in frequency domain, in *International Conference on Neural Information Processing*, Springer, 2019, 264-274.
- [35] Q. Yang, Y. Yang, T. Cui, Q. He, FDM-PINN: Physics-informed neural network based on fictitious domain method, *Int. J. Comput. Math.*, 100(2023), 511-524.
- [36] L. Zuo and Y. Hou, Numerical analysis for the mixed Navier-Stokes and Darcy problem with the Beavers-Joseph interface condition, *Numer. Methods Partial Differential Equations*, 31 (2015), 1009-1030.