Appl. Numer. Math. 207 (2025) 303–322

Contents lists available at ScienceDirect

Applied Numerical Mathematics

journal homepage: www.elsevier.com/locate/apnum

Research Paper

Optimal control for Darcy's equation in a heterogeneous porous media

SeongHee Jeong *, Sanghyun Lee

Department of Mathematics, Florida State University, Tallahassee, 32306, FL, USA

ARTICLE INFO

Keywords:
Optimal control
Darcy's flow
Heterogeneity
C^0 interior penalty

ABSTRACT

In this paper, we investigate optimal control problems in heterogeneous porous media. The optimal control problem is governed by the Darcy's flow equation; where the pressure is the state variable and the source/sink is the control variable. Then we introduce the reduced optimal control problem which contains only the state variable by replacing the control variable with a dependent quantity of the state variable based on the Darcy's equation. Here we employ C^0 interior penalty finite element methods for the spatial discretization to solve the reduced optimal control problem resulting in a fourth-order variational inequality. We use $\mathbb{P}_2$ Lagrange finite elements for C^0 interior penalty methods, which require fewer degrees of freedom than C^1 finite element methods. We provide a priori error estimates and stability analyses by considering a heterogeneous permeability coefficient. Several numerical examples validate the given theories and illustrate the capabilities of the proposed algorithm.

1. Introduction

Subsurface flow systems play a pivotal role in various critical applications, including groundwater management, oil reservoir engineering, and environmental remediation. Accurate modeling and prediction of these systems are paramount for informed decision-making, but they are often challenged by uncertainties in material properties data such as permeabilities [1].

Uncertainties in permeabilities and material properties arise due to the complex, heterogeneous nature of subsurface formations [2]. Traditional forward modeling approaches, which rely on fixed parameter values, often fall short of capturing the inherent variability of these systems. Optimal control problems, on the other hand, provide a means to calibrate models using observed data and desired constraints, thereby improving the accuracy of predictions and understanding of the phenomena.

Optimal control [3–5] allows for the determination of not only the best estimate of parameter values but also the identification of control strategies to optimize system performance. Whether it involves managing groundwater resource [6], enhancing subsurface energy system [7], or mitigating groundwater contamination [8], optimal control techniques help in achieving desired objectives while considering the uncertainties in data.

This paper emphasizes the broader perspective of achieving optimal system behavior under conditions of uncertainty in heterogeneous porous media by considering Darcy's flow equation, $-\nabla \cdot (\mathbf{K}\nabla p) = \eta$. Here we consider the pressure p as a state variable, and the source/sink η as a control variable with the heterogeneous permeability $\mathbf{K}$. Thus, we aim to optimally control the source/sink term without any prior information, to obtain the targeted pressure value.

* Corresponding author.
E-mail address: sjeong@fsu.edu (S. Jeong).

https://doi.org/10.1016/j.apnum.2024.08.027
Received 9 April 2024; Received in revised form 13 July 2024; Accepted 30 August 2024

Available online 4 September 2024
0168-9274/Published by Elsevier B.V. on behalf of IMACS.

In this work, we introduce the reduced optimal control problem, which contains only the state variable, by replacing the control variable using Darcy's equation. The advantage of reduced optimal control [9–12] includes the reduction in the number of variables that leads to the simplified optimization problem. Moreover, eliminating the control variable and replacing it with the state variable helps handle the pointwise state constraints [13,14]. However, this results in having the fourth order term in the cost functional, and hence we need to choose an appropriate numerical method to discretize the fourth order problem in order to get a good approximation to the solution of the continuous problem.

Various finite element methods have been employed to approximate a solution for fourth-order problems. As continuous problems within the domain Ω are formulated in the Sobolev space $H^2(\Omega)$ [15–17], finite element spaces [18] in conforming methods must be subspaces of $H^2(\Omega)$—specifically, they are C^1 finite element spaces. However, the drawback of these conforming methods lies in their complexity, as C^1 continuity imposes numerous conditions on the vertices and edges/faces of an element. Meeting these conditions requires a substantial number of degrees of freedom. For instance, the use of Argyris triangular elements [19], which belong to $\mathbb{P}_5$ with 21 degrees of freedom, or Macro elements [20], which are C^1 with piecewise cubic polynomials and 12 degrees of freedom, becomes necessary. The reduction of degrees of freedom is possible through the utilization of non-conforming finite element methods, as only weak continuity conditions need to be satisfied for the finite element functions and their derivatives. Yet, constructing effective nonconforming finite elements, particularly for more intricate fourth order problems, requires a considerable amount of ingenuity. On the other hand, one can explore mixed finite element methods by decomposing the biharmonic problem into two second order problems. Challenges in this approach arise in selecting a compatible pair to meet the inf-sup conditions and in solving the associated saddle point problem.

In this paper, we employ the C^0 interior penalty finite element method (C^0IP-FEM) [21–23], which belongs to the class of discontinuous Galerkin methods [24], where the discontinuity involves the first order or higher order derivatives. The lowest order elements in this family are as straightforward as classical nonconforming finite elements. The advantages of C^0IP-FEM include that the method allows to capture the smoothness of solutions using higher order elements that are as efficient as higher order C^1 elements, and considerably simpler. Unlike mixed methods, this approach can be readily extended to tackle more complex fourth order problems, such as those encountered in strain gradient elasticity problems.

Thus, we focus on utilizing the C^0IP-FEM for solving the optimal control problem [12,25,10,9,11] governed by Darcy's flow equation in heterogeneous porous media. Unlike most existing work, which considers the state to be close to the desired state in the whole domain [26–28,10,11], we employ the general cost functional which tracks points [29–34], curves, and regions in the domain [12] so that the state is close to the desired state in the specific parts of the domain with pointwise state constraints. We derive error estimates and convergence analyses by considering diffusive heterogeneity in the permeability coefficient. To illustrate the capabilities of our proposed algorithm, we provide several numerical examples testing various scenarios. For example, different states including points, lines, and subdomains in the heterogeneous porous media are considered. Although the discrete problem is fourth-order, by using C^0 interior penalty methods with $\mathbb{P}_2$ Lagrange finite element, the computational cost is considerably low. The discrete problem is solved by the primal-dual active set algorithm [35,36].

This paper is organized as follows. In Section 2, the governing system is introduced, including the optimal control problem with Darcy's equation and pointwise state constraints. The well-posedness of the optimal control problem is considered in Section 3 as well as the regularity results for the optimal state and optimal control. In Section 4, we utilize the C^0IP-FEM to the reduced optimal control problem, and the modified bilinear form with the (non-constant) permeability coefficient is obtained. Moreover, convergence error analyses are provided. We demonstrate the numerical algorithm and exhibit various numerical examples in Section 5. The concluding remarks are presented in Section 6.

2. Governing system

Let Ω be a bounded convex domain in $\mathbb{R}^2$, with the boundary $\partial\Omega$. The optimal control problem that we focus on is to find

$$(\bar{p},\bar{\eta}) = \underset{(p,\eta)\in\mathbb{U}}{\operatorname{argmin}} \frac{1}{2}\left[\int_\Omega |p-p_d|^2\,d\nu + \beta\int_\Omega |\eta|^2\,d\mathbf{x}\right], \tag{2.1}$$

where $\mathbb{U}\subset H_0^1(\Omega)\times L_2(\Omega)$ and $\beta>0$, subject to the Darcy's flow equation

$$-\nabla\cdot(\mathbf{K}\nabla p) = \eta \quad \text{in } \Omega, \tag{2.2a}$$

$$p = 0 \quad \text{on } \partial\Omega, \tag{2.2b}$$

and the pointwise state constraints

$$p_-(\mathbf{x}) \le p(\mathbf{x}) \le p_+(\mathbf{x}) \quad \text{a.e. } \mathbf{x}=(x,y)\in\Omega. \tag{2.3}$$

Darcy's flow equation (2.2) models the flow in porous media, characterized by the pressure scalar function $p:\Omega\to\mathbb{R}$, the source/sink term η which is often referred as the flow rate, and the permeability tensor $\mathbf{K}\in\mathbb{R}^{2\times2}$. Here we assume $\mathbf{K}\in[H^2(\Omega)]^{2\times2}$. The permeability tensor is symmetric and uniformly positive definite. Positive constants k_0 and k_1 exist such that for any $\mathbf{x}\in\Omega$, the following inequalities hold for all $\xi\in\mathbb{R}^2$:

$$k_0\xi^T\xi \le \xi^T\mathbf{K}(\mathbf{x})\xi \le k_1\xi^T\xi, \quad \forall\xi\in\mathbb{R}^2.$$

The pressure p is considered to be the state variable, η is the control variable, and p_d is the desired state. Through the optimal control problem (2.1), we aim p to be very close to p_d in the tracking points, lines, and subdomains of the domain Ω while p and η satisfy the Darcy's equation (2.2).

In (2.1), we recall the definition of the Radon measure ν on $\bar{\Omega}$ in [12]:

$$\int_{\Omega} f \, d\nu = \sum_{j=1}^{J} f(\mathscr{P}_j) w^j_{\mathscr{P}} + \sum_{l=1}^{L} \int_{\mathscr{C}_l} f w^l_{\mathscr{C}} \, ds + \sum_{m=1}^{M} \int_{\mathscr{E}_m} f w^m_{\mathscr{E}} \, d\mathbf{x}, \tag{2.4}$$

where $\mathscr{P} = \{\mathscr{P}_1, \dots, \mathscr{P}_J\}$ is a finite set of points in Ω, $\mathscr{C} = \{\mathscr{C}_1, \dots, \mathscr{C}_L\}$ is the union of the curves where $\mathscr{C}_l \subset \Omega$, and $\mathscr{E} = \{\mathscr{E}_1, \dots, \mathscr{E}_M\}$ is a union of the subdomains where each $\mathscr{E}_m \subset \Omega$. The weight functions $w^j_{\mathscr{P}}$, $w^l_{\mathscr{C}}$ and $w^m_{\mathscr{E}}$ are bounded nonnegative Borel measurable functions defined on $\mathscr{P}$, $\mathscr{C}$ and $\mathscr{E}$, respectively. Then, the desired state p_d is a target function of the pressure function p which is given as

$$p_d := \begin{cases} p_{\mathscr{P}} & \text{on} \quad \mathscr{P} \\ p_{\mathscr{C}} & \text{on} \quad \mathscr{C} \setminus \mathscr{P} \\ p_{\mathscr{E}} & \text{on} \quad \mathscr{E} \setminus (\mathscr{C} \cup \mathscr{P}) \end{cases} \tag{2.5}$$

such that

$$\|p_d\|^2_{L_2(\Omega;\nu)} := \int_{\Omega} |p_d|^2 \, d\nu < \infty.$$

We observe that the considered optimal control problem enables the generalization of the optimal control region. For instance, the points denoted as $\mathscr{P}$ and the curves represented by $\mathscr{C}$ are regarded as injection/production wells and fractures in the subsurface engineering problems, respectively.

Finally, we assume that the given functions $p_{\pm}$ in (2.3) satisfy [12]

$$p_{\pm} \in W^{3,q}(\Omega) \quad \text{for} \quad q > 2, \tag{2.6a}$$

$$p_- < p_+ \quad \text{on} \quad \bar{\Omega}, \tag{2.6b}$$

$$p_- < 0 < p_+ \quad \text{on} \quad \partial\Omega. \tag{2.6c}$$

In this work, instead of seeking for both the state p and the control η simultaneously, we reformulate the optimal control problem (2.1) by utilizing the linear partial differential operator $\mathscr{L}$, where

$$\mathscr{L} p := -\nabla \cdot (\mathbf{K} \nabla p), \tag{2.7}$$

so that η can be written as $\eta = \mathscr{L} p$ by (2.2a). Thus, the reduced optimal control problem is obtained as following: Find $\bar{p} \in U$ such that

$$\bar{p} = \underset{p \in U}{\operatorname{argmin}} \frac{1}{2} \left[\int_{\Omega} |p - p_d|^2 \, d\nu + \beta \int_{\Omega} |\mathscr{L} p|^2 \, d\mathbf{x} \right], \tag{2.8}$$

where the admissible set is

$$U = \{ p \in H^2(\Omega) \cap H^1_0(\Omega) : \mathscr{L} p \in L_2(\Omega) \text{ and } p_- \le p \le p_+ \text{ in } \Omega \}. \tag{2.9}$$

Note that according to (2.6b)-(2.6c), the admissible set U is nonempty.

Remark 2.1. Due to the convexity of Ω, the constraints (2.2) and $\mathbf{K} \in [H^2(\Omega)]^{2\times 2}$ imply that $p \in H^2(\Omega)$ if $(p, \eta) \in \mathbb{U}$ by elliptic regularity. Therefore the cost functional

$$\int_{\Omega} |p - p_d|^2 \, d\nu + \beta \int_{\Omega} |\eta|^2 \, d\mathbf{x}$$

is well-defined by the Sobolev embedding theorem $H^2(\Omega) \subset C(\bar{\Omega})$ [15].

3. Well-posedness and regularity results

In this section, we briefly recapitulate the well-posedness and regularity for the given problem based on [12]. Note that the main difference compared to [12] is that we consider the coefficient $\mathbf{K}$ in the elliptic partial differential equation constraint (2.2).

Theorem 3.1. *The reduced optimal control problem (2.8)-(2.9) has a unique solution.*

Proof. First one can rewrite (2.8) as following:

$$\bar{p} = \underset{p\in U}{\operatorname{argmin}} \frac{1}{2}\left[\int_\Omega |p - p_d|^2\, d\nu + \beta \int_\Omega |\mathscr{L}p|^2\, d\mathbf{x}\right]$$
$$= \underset{p\in U}{\operatorname{argmin}} \frac{1}{2}\left[\int_\Omega |p|^2\, d\nu + \beta \int_\Omega |\mathscr{L}p|^2\, d\mathbf{x} - 2\int_\Omega p\, p_d\, d\nu\right].$$

Define the bilinear form $\mathscr{A}(\cdot,\cdot)$ as

$$\mathscr{A}(v,w) := \int_\Omega vw\, d\nu + \beta \int_\Omega (\mathscr{L}v)(\mathscr{L}w)\, d\mathbf{x}, \tag{3.1}$$

and the linear functional $F(\cdot)$ as

$$F(v) := \int_\Omega v\, p_d\, d\nu.$$

Then the reduced minimization problem (2.8) is written as

$$\bar{p} = \underset{p\in U}{\operatorname{argmin}} \left[\frac{1}{2}\mathscr{A}(p,p) - F(p)\right]. \tag{3.2}$$

By (2.4), the Cauchy-Schwarz inequality, and the Sobolev embedding theorem,

$$|\mathscr{A}(v,w)| \le C\left(\|v\|_{H^2(\Omega)\cap H_0^1(\Omega)}\|w\|_{H^2(\Omega)\cap H_0^1(\Omega)}\right),$$

and

$$|F(v)| < \infty,$$

for any $v, w \in H^2(\Omega)\cap H_0^1(\Omega)$. Moreover, according to (2.4) and the trace theorem,

$$\mathscr{A}(v,v) \ge C\|v\|^2_{H^2(\Omega)\cap H_0^1(\Omega)},$$

for any $v \in H^2(\Omega)\cap H_0^1(\Omega)$. This tells us that the bilinear form $\mathscr{A}(\cdot,\cdot)$ is continuous and coercive, and the linear functional $F(\cdot)$ is bounded. Since U is a nonempty, closed, and convex subset of $H^2(\Omega)\cap H_0^1(\Omega)$, by the Riesz Representation theorem and the Projection theorem [16], the reduced optimal control problem (3.2) has a unique solution. □

Now we turn to the first order (necessary) optimality condition by using Theorem 3.1 which tells us that $\bar{p}$ is the unique optimal solution to the minimization problem (2.8)-(2.9). This then leads us to the generalized Karush-Kuhn-Tucker (KKT) conditions. In order to deduce the conditions, we derive the variational inequality corresponding to the optimal control problem (2.8)-(2.9).

Lemma 3.1. *The optimal control problem (2.8)-(2.9) has a unique solution $\bar{p} \in U$ characterized by the fourth order variational inequality [37]*

$$\beta \int_\Omega (\mathscr{L}\bar{p})(\mathscr{L}(p-\bar{p}))\, d\mathbf{x} + \int_\Omega (\bar{p} - p_d)(p-\bar{p})\, d\nu \ge 0 \quad \forall p \in U. \tag{3.3}$$

Proof. Let $p \in U$ be arbitrary. We define the function $\Theta : [0,1] \to \mathbb{R}$ such that

$$\Theta(t) := J\left((1-t)\bar{p} + tp\right),$$

where

$$J(p) := \frac{1}{2}\left[\int_\Omega |p - p_d|^2\, d\nu + \beta \int_\Omega |\mathscr{L}p|^2\, d\mathbf{x}\right].$$

Since $\bar{p} \in U$ is the solution to the optimization problem (2.8)-(2.9), it implies that $\bar{p} \in U$ is the minimizer of $J(p)$, and thus

$$\Theta(0) = J(\bar{p}) \le J\left((1-t)\bar{p} + tp\right) = \Theta(t) \quad \text{for} \quad 0 \le t \le 1.$$

This leads us that $\Theta'(0) \ge 0$. Therefore,

$$0 \le \Theta'(0) = \left[\frac{d}{dt} J\left((1-t)\bar{p} + tp\right)\right]_{t=0}$$
$$= \beta \int_\Omega (\mathscr{L}\bar{p})(\mathscr{L}(p-\bar{p}))\, d\mathbf{x} + \int_\Omega (\bar{p} - p_d)(p-\bar{p})\, d\nu,$$

for any $p \in U$, and hence the result follows. □

The fourth order variational inequality (3.3) is equivalent to the following generalized KKT conditions [38,39]:

$$\beta \int_\Omega (\mathscr{L}\bar{p})(\mathscr{L}g)\, d\mathbf{x} + \int_\Omega (\bar{p} - p_d) g\, d\nu = \int_\Omega g\, d\lambda \quad \forall g \in H^2(\Omega) \cap H_0^1(\Omega), \tag{3.4}$$

where λ is a bounded regular Borel measure satisfying

$$\lambda \ge 0 \quad \text{if} \quad \bar{p} = p_+, \tag{3.5a}$$

$$\lambda \le 0 \quad \text{if} \quad \bar{p} = p_-, \tag{3.5b}$$

$$\lambda = 0 \quad \text{otherwise.} \tag{3.5c}$$

See [40] for the proof of the KKT conditions.

Finally, we can find the local and global regularity results for the optimal state of the optimal control problem (2.8)-(2.9). The solution $\bar{p}$ belongs to [41,42,12]

$$W^{3,s}_{\text{loc}}(\Omega)$$

for any $s \in (1,2)$. Furthermore, globally [43,44,12]

$$\bar{p} \in H^{2+\alpha}(\Omega) \tag{3.6}$$

for some $\alpha \in (0,1)$, where α is the index of elliptic regularity determined by Ω. Note that for a rectangular domain [44,34], it is known that $\alpha = 1 - \varepsilon$ where $\varepsilon > 0$.

4. C^0 interior penalty methods and error analysis

In this section, we introduce the C^0IP-FEM for spatial discretization of the reduced optimal control problem and provide error convergence analyses. One of the main advantages of C^0IP-FEM for the fourth order variational problem (3.3) is its lower cost compared to that of the conforming continuous Galerkin methods.

4.1. C^0 interior penalty finite element methods

Let $\mathscr{T}_h = \{T\}$ be a shape-regular triangulation of the domain Ω into triangular elements with a mesh size $h = \max_T h_T$, where h_T is the diameter of $T \in \mathscr{T}_h$. Let $\mathbb{P}_k(T)$ be the space of all polynomials of degree at most $k \ge 0$ on a set T, and the nodal interpolation operator for the $\mathbb{P}_1$ finite element space associated with $\mathscr{T}_h$ is denoted by I_h.

In this paper, we employ the $\mathbb{P}_2$ Lagrange triangular elements [21]. Thus, we define the finite element space as

$$V_h := \left\{\psi \in H_0^1(\Omega) \mid \psi|_T \in \mathbb{P}_2(T) \quad \forall T \in \mathscr{T}_h\right\}.$$

The C^0IP-FEM is a class of discontinuous Galerkin methods, where the discontinuity is in the normal derivative across element boundaries. Since V_h is not a subspace of $H^2(\Omega)$ ($V_h \subseteq C(\bar{\Omega})$ but $V_h \nsubseteq C^1(\bar{\Omega})$), this is a H^2-nonconforming method. The discontinuity in normal derivatives yields the jump and the average across edges in the discrete bilinear form. We note that this method relaxes the C^1 continuity requirement by replacing it with interior penalty techniques.

To construct C^0IP-FEM, we first define the piecewise Sobolev space

$$H^2(\mathscr{T}_h) = \{v \in L_2(\Omega) : v_T = v|_T \in H^2(T) \quad \forall T \in \mathscr{T}_h\}.$$

We denote by $\mathscr{E}_h$ the set of all edges in the mesh and by $\mathscr{E}_h^I$ the set of all the interior edges. For each $T \in \mathscr{T}_h$, denote the boundary of T by ∂T. If $e \in \mathscr{E}_h^I$, we assign to e a fixed unit normal vector $\mathbf{n}_e$ pointing from T_- to T_+. Let v be any piecewise $H^2(\mathscr{T}_h)$ function, then the jump value of the normal derivative of v on $e \in \mathscr{E}_h^I$ (shared by the two neighboring elements $T_\pm \in \mathscr{T}_h$) is defined as

$$[\![\partial v/\partial \mathbf{n}_e]\!] := \partial v_+/\partial \mathbf{n}_e - \partial v_-/\partial \mathbf{n}_e,$$

where $v_\pm = (v|_{T_\pm})|_e$. If $e \in \partial\Omega$, then e belongs to only one triangle T, and we define

$$[\![\partial v/\partial \mathbf{n}_e]\!] := -\partial v|_T/\partial \mathbf{n}_e,$$

where we take $\mathbf{n}_e$ to be the unit normal pointing outside Ω. Next, let v be any piecewise $H^s(\mathscr{T}_h)$ function for $s > 5/2$, then average of the second order normal derivative of v on $e \in \mathscr{E}_h^I$ is defined as

$$\left\{\left\{\partial^2 v/\partial \mathbf{n}_e^2\right\}\right\} := \frac{1}{2}\left(\left(\partial^2 v_+/\partial \mathbf{n}_e^2\right) + \left(\partial^2 v_-/\partial \mathbf{n}_e^2\right)\right).$$

If $e \in \partial\Omega$, then

$$\left\{\left\{\partial^2 v/\partial \mathbf{n}_e^2\right\}\right\} := \partial^2 v|_T/\partial \mathbf{n}_e^2,$$

where we take $\mathbf{n}_e$ to be the unit normal pointing outside Ω. Furthermore, with the permeability coefficient $\mathbf{K}$, we have

$$\left[\!\left[(\mathbf{K}\nabla v)\cdot \mathbf{n}_e\right]\!\right] := (\mathbf{K}_+\nabla v_+)\cdot \mathbf{n}_e - (\mathbf{K}_-\nabla v_-)\cdot \mathbf{n}_e$$

and

$$\left\{\left\{\nabla((\mathbf{K}\nabla v)\cdot \mathbf{n}_e)\cdot \mathbf{n}_e\right\}\right\} := \frac{1}{2}\left[\nabla\left((\mathbf{K}_+\nabla v_+)\cdot \mathbf{n}_e\right) + \nabla\left((\mathbf{K}_-\nabla v_-)\cdot \mathbf{n}_e\right)\right],$$

where $\mathbf{K}_\pm = (\mathbf{K}|_{T_\pm})|_e$.

To derive the discrete optimal control problem, we define the piecewise operator $\mathscr{L}_T$ with respect to $\mathscr{T}_h$ as

$$\mathscr{L}_T v_h = -\nabla\cdot(\mathbf{K}\nabla v_h)\Big|_T, \tag{4.1}$$

and we assume $\mathbf{K}$ is piecewise constant, and thus $\nabla\mathbf{K} = 0$ piecewise accordingly. Then the discrete problem of (2.8) is to find

$$\bar{p}_h = \underset{p_h\in U_h}{\operatorname{argmin}} \frac{1}{2}\left[\int_\Omega |p_h - p_d|^2\, dv + \beta b_h(p_h,p_h)\right], \tag{4.2}$$

where

$$U_h = \{p_h \in V_h :\ I_h p_- \le I_h p_h \le I_h p_+\}. \tag{4.3}$$

Here, $b_h(\cdot,\cdot)$ is the discrete bilinear form obtained by using C^0IP-FEM:

$$\begin{aligned} b_h(v_h,w_h) = &\sum_{T\in\mathscr{T}_h}\int_T (\mathscr{L}_T v_h)(\mathscr{L}_T w_h)\, d\mathbf{x} \\ &+ \sum_{e\in\mathscr{E}_h^I}\sigma h_e^{-1}\int_e \left[\!\left[(\mathbf{K}\nabla v_h)\cdot\mathbf{n}_e\right]\!\right]\left[\!\left[(\mathbf{K}\nabla w_h)\cdot\mathbf{n}_e\right]\!\right] ds \\ &+ \sum_{e\in\mathscr{E}_h^I}\int_e \left\{\left\{\nabla((\mathbf{K}\nabla v_h)\cdot\mathbf{n}_e)\cdot\mathbf{n}_e\right\}\right\}\left[\!\left[(\mathbf{K}\nabla w_h)\cdot\mathbf{n}_e\right]\!\right] ds \\ &+ \sum_{e\in\mathscr{E}_h^I}\int_e \left\{\left\{\nabla((\mathbf{K}\nabla w_h)\cdot\mathbf{n}_e)\cdot\mathbf{n}_e\right\}\right\}\left[\!\left[(\mathbf{K}\nabla v_h)\cdot\mathbf{n}_e\right]\!\right], ds \end{aligned} \tag{4.4}$$

where h_e is the diameter of the edge e, and $\sigma > 0$ is a penalty parameter.

Similar to the continuous problem, U_h is nonempty and hence the minimization problem (4.2) has a unique solution $\bar{p}_h \in U_h$ characterized by the discrete variational inequality

$$\int_\Omega (\bar{p}_h - p_d)(p_h - \bar{p}_h)\, dv + \beta b_h(\bar{p}_h, p_h - \bar{p}_h) \ge 0 \qquad \forall\, p_h \in U_h. \tag{4.5}$$

Then, we can rewrite (4.5) as follows:

$$\mathscr{A}_h(\bar{p}_h, p_h - \bar{p}_h) - \int_\Omega p_d(p_h - \bar{p}_h)\, dv \ge 0 \qquad \forall\, p_h \in U_h, \tag{4.6}$$

where the bilinear form is

$$\mathscr{A}_h(v_h,w_h) := \beta b_h(v_h,w_h) + \int_\Omega v_h w_h\, dv, \tag{4.7}$$

which approximates the bilinear form (3.1).

Moreover, through symmetrization interior penalty term

$$\int_e \left\{\left\{\nabla((\mathbf{K}\nabla w_h)\cdot\mathbf{n}_e)\cdot\mathbf{n}_e\right\}\right\}\left[\!\left[(\mathbf{K}\nabla v_h)\cdot\mathbf{n}_e\right]\!\right] ds,$$

and the penalty stabilization term

$$\sigma h_e^{-1} \int_e [\![(\mathbf{K}\nabla v_h)\cdot \mathbf{n}_e]\!]\,[\![(\mathbf{K}\nabla w_h)\cdot \mathbf{n}_e)]\!]\,ds,$$

we can get the symmetric positive-definiteness (SPD) of the bilinear form $b_h(\cdot,\cdot)$ with sufficiently large penalty parameter σ, and thus the discrete problem preserves the SPD property of the continuous problem.

4.2. Convergence analysis

Next, we will derive the error estimates based on the convergence analysis from [12]. We define the mesh-dependent H^2-norm as follows:

$$|v|^2_{H^2(\Omega;\mathscr{T}_h)} := \sum_{T\in\mathscr{T}_h} |\mathscr{L}_T v|^2_{L_2(T)} + \sum_{e\in\mathscr{E}_h^I} h_e^{-1} \|\,[\![(\mathbf{K}\nabla v)\cdot \mathbf{n}_e]\!]\,\|^2_{L_2(e)}. \tag{4.8}$$

Thus, there exist positive constants C_1 and C_2 such that [22,21]

$$|b_h(v_h,w_h)| \le C_1 |v_h|_{H^2(\Omega;\mathscr{T}_h)} |w_h|_{H^2(\Omega;\mathscr{T}_h)} \qquad \forall v_h, w_h \in V_h, \tag{4.9}$$

$$b_h(v_h,v_h) \ge C_2 |v_h|^2_{H^2(\Omega;\mathscr{T}_h)} \qquad \forall v_h \in V_h, \tag{4.10}$$

provided the penalty parameter σ is sufficiently large.

We also define the mesh-dependent energy norm $\|\cdot\|_h$ by

$$\|v\|_h^2 := \beta |v|^2_{H^2(\Omega,\mathscr{T}_h)} + \|v\|^2_{L_2(\Omega;\nu)}. \tag{4.11}$$

Then, according to (4.7), (4.8)-(4.10), and (4.11), there are positive constants C_1 and C_2 such that

$$\mathscr{A}_h(v_h,w_h) \le C_1 \|v_h\|_h \|w_h\|_h \qquad \forall\, v_h, w_h \in V_h, \tag{4.12}$$

$$\mathscr{A}_h(v_h,v_h) \ge C_2 \|v_h\|_h^2 \qquad \forall\, v_h \in V_h. \tag{4.13}$$

Finally, we can derive the following theorem.

Theorem 4.1. *There is a constant $C>0$ independent of h such that*

$$\|\bar{p} - \bar{p}_h\|_h \le C h^{\alpha}, \tag{4.14}$$

where $\alpha \in (0,1)$ is the index of elliptic regularity.

Proof. The detailed proof can be found in [12].

The following error estimates for the Lagrange interpolation operator, $\Pi_h : H^2(\Omega)\cap H_0^1(\Omega) \to V_h$, are based on the Bramble-Hilbert lemma [45,46], and the local and global regularity estimates, can be found in [21,47]. We have, by (4.1),

$$\|\bar{p} - \Pi_h\bar{p}\|_{L_2(\Omega)} \le C h^{2+\alpha}, \tag{4.15}$$

$$|\bar{p} - \Pi_h\bar{p}|_{H^1(\Omega)} \le C h^{1+\alpha}, \tag{4.16}$$

$$\|\bar{p} - \Pi_h\bar{p}\|_{L_\infty(\Omega)} \le C h^{1+\alpha}, \tag{4.17}$$

$$\left(\sum_{T\in\mathscr{T}_h} |\mathscr{L}_T(\bar{p} - \Pi_h\bar{p})|^2_{L_2(T)}\right)^{\frac{1}{2}} \le C h^{\alpha}. \tag{4.18}$$

By (4.16), (4.18), and the trace inequality with scaling, one can obtain

$$\left(\sum_{e\in\mathscr{E}_h^I} h_e^{-1} \|(\mathbf{K}\nabla(\bar{p} - \Pi_h\bar{p}))\cdot \mathbf{n}_e\|^2_{L_2(e)}\right)^{1/2} \le C h^{\alpha}. \tag{4.19}$$

Furthermore, it follows from (2.4), and (4.15)-(4.19) that

$$\|\bar{p} - \Pi_h\bar{p}\|_{L_2(\Omega;\nu)} \le C h^{1+\alpha},$$

and thus

$$\|\bar{p} - \Pi_h\bar{p}\|_h \le C h^{\alpha}. \tag{4.20}$$

Combining with (4.13), (4.20) and (4.6), we obtain

$$\begin{aligned}\|\bar{p}-\bar{p}_h\|_h^2 &\le 2\|\bar{p}-\Pi_h\bar{p}\|_h^2 + 2\|\Pi_h\bar{p}-\bar{p}_h\|_h^2\\ &\le C\left[h^{2\alpha} + \mathscr{A}_h(\Pi_h\bar{p}-\bar{p}_h, \Pi_h\bar{p}-\bar{p}_h)\right]\\ &\le C\left(h^{2\alpha} + h^{\alpha}\|\Pi_h\bar{p}-\bar{p}_h\|_h\right)\\ &\le C\left[h^{2\alpha} + h^{\alpha}(\|\Pi_h\bar{p}-\bar{p}\|_h + \|\bar{p}-\bar{p}_h\|_h)\right]\\ &\le C\left(h^{2\alpha} + h^{\alpha}\|\bar{p}-\bar{p}_h\|_h\right),\end{aligned}$$

and hence the inequality of arithmetic and geometric means implies (4.14). □

5. Numerical examples

In this section, we present a set of numerical examples validating the theoretical results presented in Section 4 and demonstrating the capabilities of the proposed algorithm in different scenarios. The computations are performed with in-house developed code based on Matlab [40].

5.1. Global algorithm

First, we briefly summarize the global algorithm to solve the given problem. The variational inequality (4.6) can be written in the algebraic form as follows:

$$\begin{cases} \mathtt{p}^T\mathtt{A}(\mathtt{q}-\mathtt{p}) - \mathtt{f}^T(\mathtt{q}-\mathtt{p}) \ge 0 \quad \forall \mathtt{q}\in\mathbb{R}^N \\ I_h p_- \le \mathtt{p} \le I_h p_+, \end{cases} \tag{5.1}$$

where $\mathtt{p}\in\mathbb{R}^N$ is the solution vector, $\mathtt{A}\in\mathbb{R}^{N\times N}$ is the matrix obtained from (4.7) and (4.4), and $\mathtt{f}\in\mathbb{R}^N$ is the vector obtained from the second term of the left-hand side of (4.6). Here, N is the total number of nodal points. As we introduce the dual unknown $\mu\in\mathbb{R}^N$, the discrete problem (5.1) becomes the following system

$$\begin{cases} \mathtt{A}\mathtt{p} - \mathtt{f} + \mu = 0 \\ \mu = \max\{0, \mu + C(\mathtt{p} - I_h p_+)\} + \min\{0, \mu + C(\mathtt{p} - I_h p_-)\}, \end{cases} \tag{5.2}$$

where $C>0$. Then, the equivalent KKT conditions are obtained as follows:

$$\begin{cases} \mathtt{A}\mathtt{p} + \mu = \mathtt{f} \\ I_h p_- \le \mathtt{p} \le I_h p_+ \\ \mu \ge 0 \quad \text{if} \quad \mathtt{p} = I_h p_+ \\ \mu \le 0 \quad \text{if} \quad \mathtt{p} = I_h p_- \\ \mu = 0 \quad \text{if} \quad I_h p_- \le \mathtt{p} \le I_h p_+. \end{cases}$$

Remark 5.1. Suppose $\mathtt{p}$ and μ are the unique solution to (5.2). The primal dual active set algorithm [36] converges superlinearly provided that the initial state $\mathtt{p}^0$ and the initial dual μ^0 are sufficiently close to $\mathtt{p}$ and μ, respectively.

Algorithm 1 Primal dual active set algorithm.

set initials: $\mathtt{p}^0$, μ^0.

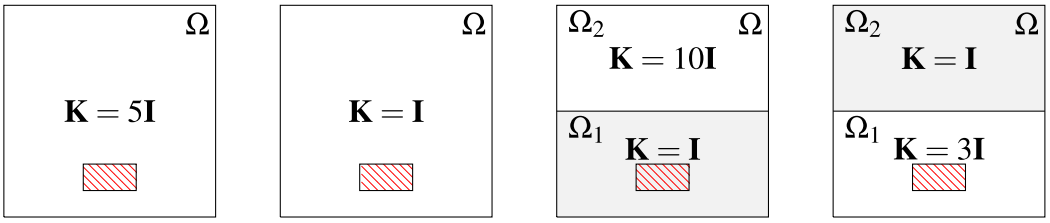
find active sets: $\mathscr{A}_-^k = \{i\in[0,N] : \mu^k(i) + C(\mathtt{p}^k(i) - I_h p_-(i)) < 0\}$,
$\mathscr{A}_+^k = \{i\in[0,N] : \mu^k(i) + C(\mathtt{p}^k(i) - I_h p_+(i)) > 0\}$;
inactive set: $\mathscr{I}^k = (\mathscr{A}_-^k \cup \mathscr{A}_+^k)^c$

while $\mu^k \ne \mu^{k+1}$ **do**

solve
$$\begin{bmatrix} \mathtt{A}(:,:) & \mathtt{I}^T\left(:,\mathscr{A}_-^k\right) & \mathtt{I}^T\left(:,\mathscr{A}_+^k\right) \\ \mathtt{I}\left(\mathscr{A}_-^k,:\right) & \mathtt{0}\left(\mathscr{A}_-^k,\mathscr{A}_-^k\right) & \mathtt{0}\left(\mathscr{A}_-^k,\mathscr{A}_+^k\right) \\ \mathtt{I}\left(\mathscr{A}_+^k,:\right) & \mathtt{0}\left(\mathscr{A}_+^k,\mathscr{A}_-^k\right) & \mathtt{0}\left(\mathscr{A}_+^k,\mathscr{A}_+^k\right) \end{bmatrix} \begin{bmatrix} \mathtt{p}^{k+1} \\ \mu^{k+1}\left(\mathscr{A}_-^k\right) \\ \mu^{k+1}\left(\mathscr{A}_+^k\right) \end{bmatrix} = \begin{bmatrix} \mathtt{f} \\ I_h p_-\left(\mathscr{A}_-^k\right) \\ I_h p_+\left(\mathscr{A}_+^k\right) \end{bmatrix}$$

update $\mathscr{A}_-^{k+1} = \{i\in[0,N] : \mu^k(i) + C(\mathtt{p}^{k+1}(i) - I_h p_-(i)) < 0\}$,
$\mathscr{A}_+^{k+1} = \{i\in[0,N] : \mu^k(i) + C(\mathtt{p}^{k+1}(i) - I_h p_+(i)) > 0\}$;
$\mathscr{I}^{k+1} = (\mathscr{A}_-^{k+1} \cup \mathscr{A}_+^{k+1})^c$

end while



(a) Case i) (b) Case ii) (c) Case iii) (d) Case iv)

Fig. 1. Example 1. Four different cases.

5.2. Primal dual active set

The discrete problem (5.2) is a quadratic program with simple box constraints, and here we employ the primal dual active set (PDAS) algorithm [48,36] to solve the problem. The main idea of the PDAS is to solve the system while the active sets

$$\mathscr{A}_-^k = \{i \in [0,N] : \mu^k(i) + C(\mathfrak{p}^k(i) - I_h p_-(i)) < 0\},$$
$$\mathscr{A}_+^k = \{i \in [0,N] : \mu^k(i) + C(\mathfrak{p}^k(i) - I_h p_+(i)) > 0\},$$

are updated by finding the nodal indices where the state does not satisfy the constraints

$$I_h p_-(i) \le \mathfrak{p}^k(i) \le I_h p_+(i).$$

Here, k is the iteration number, and i is the index of nodal points. See Algorithm 1 for more details. This algorithm terminates when the previous active sets are the same as the updated ones; equivalently, when the dual unknown μ remains the same. A simple direct solver is used to solve the linear system in the algorithm.

We note that the Lagrange multiplier λ in (3.5) associated with the pointwise state constraints (2.3) is a Borel measure. As a result, PDAS algorithm for the state-constrained problem exhibits mesh-dependent behavior and has no function space analysis as discussed in [49–51]. To overcome these shortcomings, mixed state-control constraints technique [51–53] or Moreau-Yosida regularization [54–57] techniques were developed. However, it is not trivial to utilize these approaches in our given problem.

In our case, we consider the reduced optimal control problem, which is obtained by replacing the control η with $\eta = \mathscr{L}p = -\nabla\cdot(\mathbf{K}\nabla p)$ according to the partial differential equation constraint (2.2a). This resulted in the C^0 interior penalty methods with only state variable, and it requires at least a quadratic Lagrange finite element. Thus, the mixed state-control constraints (or Lavrentiev-type regularization) approach, which reformulated the state variable p to the control variable η, is not applicable to our approach. On the other hand, considering the Moreau-Yosida regularization (or augmented Lagrangian-type penalization) approach is challenging due to our use of the partial differential equation to replace the control variable in the cost functional. Since this work focuses on the analyses and the application of C^0IP-FEM methods, considering mesh-independence PDAS algorithm for our reduced problem is out of the scope of this work but will be studied further. However, we have still illustrated the iteration numbers in the PDAS algorithm in one of our numerical examples.

5.3. Example 1. Heterogeneous domain with layers

In this section, we test four different cases in the computational domain $\Omega = (0,1)^2$. We consider the permeability tensor $\mathbf{K} = K\mathbf{I}$, where K is a piecewise constant, and $\mathbf{I}$ is the 2×2 identity matrix, in all the numerical experiments for the simplicity.

First two cases consider homogeneous $\mathbf{K}$ values;

Case i) $\mathbf{K} = 5\mathbf{I}$ in Ω, and

Case ii) $\mathbf{K} = \mathbf{I}$ in Ω.

Then, we consider heterogeneous domains by setting two different layers such as

Case iii) $\mathbf{K} = \mathbf{I}$ in Ω_1 and $\mathbf{K} = 10\mathbf{I}$ in Ω_2, and

Case iv) $\mathbf{K} = 3\mathbf{I}$ in Ω_1 and $\mathbf{K} = \mathbf{I}$ in Ω_2.

Here, $\Omega = \Omega_1 \cup \Omega_2$, where $\Omega_1 = \{(x,y) \mid y \le 0.5\}$ and $\Omega_2 = \{(x,y) \mid y > 0.5\}$.

For these cases, the desired state is given in a tracking subdomain region (i.e., $M = 1$ while $J = 0$ and $L = 0$). We recall that J is the number of tracking points, L is the number of tracking line segments, and M is the number of tracking subdomains; see (2.4). Here we set $\mathscr{E}_1 = [0.375, 0.125] \times [0.625, 0.25]$. Details are shown in Fig. 1. The weight is set to be $w^1_{\mathscr{E}} = 10^6$, and the penalty parameter is chosen to be $\beta = 1$ and $\sigma = 10$. The state constraints $p_\pm$ are given by

$$p_+(x,y) = 5 - ((x-0.5)^2 + (y-0.5)^2) \quad \text{and} \quad p_-(x,y) = -\infty,$$

and $p_d = p_{\mathscr{E}} = 4.84$.

Table 1
Example 1. Convergence results for each different cases.

h	Case i) E_h	order	Case ii) E_h	order	Case iii) E_h	order	Case iv) E_h	order
2^{-3}	1.88e-02	-	2.63e-02	-	2.79e-02	-	2.80e-02	-
2^{-4}	8.07e-03	1.22	1.19e-02	0.98	1.26e-02	1.15	1.14e-02	1.40
2^{-5}	1.93e-03	2.06	4.59e-03	0.95	5.21e-03	1.27	5.23e-03	1.12
2^{-6}	6.44e-04	1.59	1.44e-03	1.31	2.00e-03	1.38	3.34e-03	0.65
2^{-7}	1.34e-04	2.27	4.16e-04	1.82	8.64e-04	1.21	1.70e-03	0.97

Table 2
Example 1. Energy norm of the finest result $\tilde{p}$ when $h = 2^{-8}$ for each different cases.

$\|\tilde{p}\|_h$	Case i)	Case ii)	Case iii)	Case iv)
2^{-8}	7.30e+02	8.46e+02	8.40e+02	7.85e+02

Table 3
Example 1. The number of iterations in the PDAS algorithm for each case.

h	2^{-3}	2^{-4}	2^{-5}	2^{-6}	2^{-7}	2^{-8}
Case i)	1	1	3	6	12	24
Case ii)	1	2	3	5	16	32
Case iii)	1	2	3	5	7	20
Case iv)	4	9	16	31	62	130

5.3.1. Convergence

First, to verify the optimal convergence rate of the proposed method as proven in Theorem 4.1, we measure errors on five uniform meshes, starting with the initial mesh of $h = 2^{-3}$ halving them in each refinement cycle. The errors are measured in the relative norm $E_h := \|\tilde{p} - \bar{p}_h\|_h / \|\tilde{p}\|_h$, where $\tilde{p}$ is the numerical solution obtained with the finest grid $h = 2^{-8}$. The convergence results are summarized in Table 1, illustrating the expected optimal convergence rates. Moreover, Table 2 presents the magnitude of $\|\tilde{p}\|_h$.

5.3.2. Discussions

We emphasize the importance of considering heterogeneous media to explore realistic subsurface scenarios. In Fig. 2, we compare differences in the state variable $\bar{p}_h$ among four different cases, Case i) - iv). The left column depicts the values of $\bar{p}_h$, while the right column showcases the $\bar{p}_h$ values from a top-view perspective within the control domain $\mathscr{E}_1$. Although Case i) and Case ii) have similar results, the comparison between the homogeneous Case ii) (Fig. 2 c) - d)) and heterogeneous Case iii) (Fig. 2 e) - f)) reveals variations in the pressure values $\bar{p}_h$ due to the presence of a heterogeneous layer. Specifically, we observe that the Case ii) has more diffusive and higher pressure values on the upper domain compare to the Case iii). Furthermore, Case iv) illustrates significant different pressure values compared to that of Case iii) as the permeability $\mathbf{K}$ has opposite aspects in Case iii) and Case iv). Therefore, targeting on the pressure value $p_{\mathscr{E}}$ within the region $\mathscr{E}_1$ leads to distinct optimal control solution values depending on the permeability.

Next, Fig. 3 illustrates the pressure values $\bar{p}_h$ plotted over the line $x = 0.5$ in the domain Ω for each case. On the other hand, Fig. 4 illustrates the pressure values $\bar{p}_h$ plotted over the line $y = 0.25$. These figures provide clear differences in the optimal pressure values, as discussed in the above paragraph. We note that the target pressure value $p_d = p_{\mathscr{E}} = 4.84$ is almost matched.

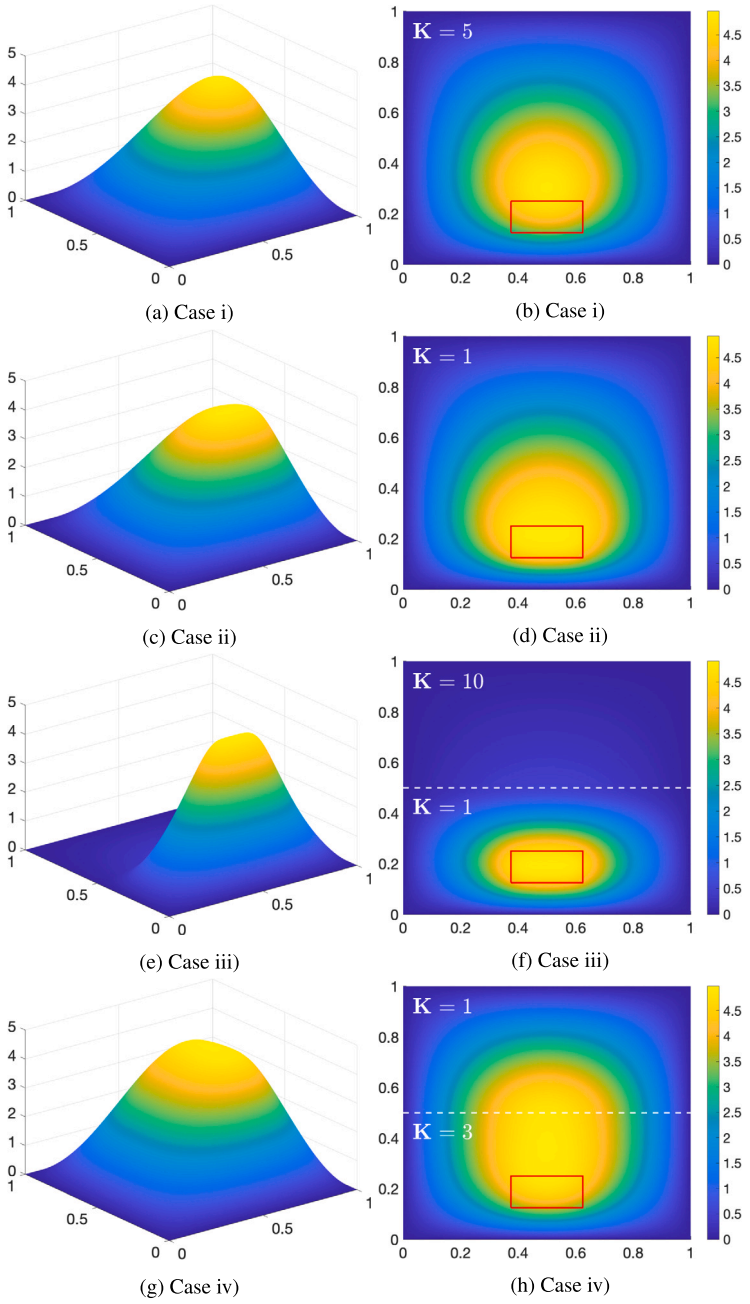
Fig. 5 illustrates the differences in the control variable $\bar{\eta}_h$ across the four distinct cases, Case i) - iv). These $\bar{\eta}_h$ values define the configuration of the source function (either injection or production) in the Darcy's flow equation to attain the targeted pressure value p_d within the specified subdomain $\mathscr{E}_1$.

Note that Table 3 shows the number of iterations in the PDAS algorithm. It is observed that the number of iterations increases as the mesh size decreases, and hence it exhibits that the PDAS is mesh-dependent for the state-constrained optimal control problem.

5.4. Example 2. Various tracking states

In this example, the domain setup is the same as in Example 1, Case iii). However, we test various tracking states, including lines that can be considered fractures, a point that can be regarded as an injection/production source, and combinations of these elements. Fig. 6 illustrates the details for each case.

Thus, $J = 1$, $L = 0$, and $M = 0$ for Case i) with $\mathscr{P}_1 = (0.5, 0.25)$. For Case ii), we set $J = 0$, $L = 2$, and $M = 0$, where $\mathscr{C}_1 = (0.25, 0.125) - (0.75, 0.125)$ and $\mathscr{C}_2 = (0.25, 0) - (0.625, 0.375)$. We mix different tracking states in the Case iii), such as $J = 1$, $L = 2$, and $M = 0$, where $\mathscr{P}_1 = (0.125, 0.125)$, and $\mathscr{C}_1 = (0.375, 0.25) - (0.75, 0.25)$ and $\mathscr{C}_2 = (0.375, 0.125) - (0.625, 0.375)$.



(a) Case i) (b) Case i)

(c) Case ii) (d) Case ii)

(e) Case iii) (f) Case iii)

(g) Case iv) (h) Case iv)

Fig. 2. Example 1. Illustrates the state variable $\bar{p}_h$ for each different cases. The right figure shows the tracking domain.

In addition, the weight functions and target functions are set as $w^1_{\mathscr{P}} = 10^5$ and $p_d = p_{\mathscr{P}} = 4.93$ for the Case i), $w^1_{\mathscr{C}} = w^2_{\mathscr{C}} = 10^5$ and $p_d = p_{\mathscr{C}} = 4.68$ for the Case ii), and $w^1_{\mathscr{P}} = w^1_{\mathscr{C}} = w^2_{\mathscr{C}} = 10^5$ and $p_{\mathscr{P}} = 4.73$ $p_{\mathscr{C}} = 4.84$ for the Case iii).

5.4.1. *Convergence*

First, to confirm the optimal convergence rate of the proposed method, as demonstrated in Example 1, we calculate errors on five uniform meshes, starting with the initial mesh of $h = 2^{-3}$ halving them in each refinement cycle. The errors are computed in the relative norm as in Example 1. The convergence results are presented in Tables 4, 5, showing the expected optimal convergence rates in Theorem 4.1.

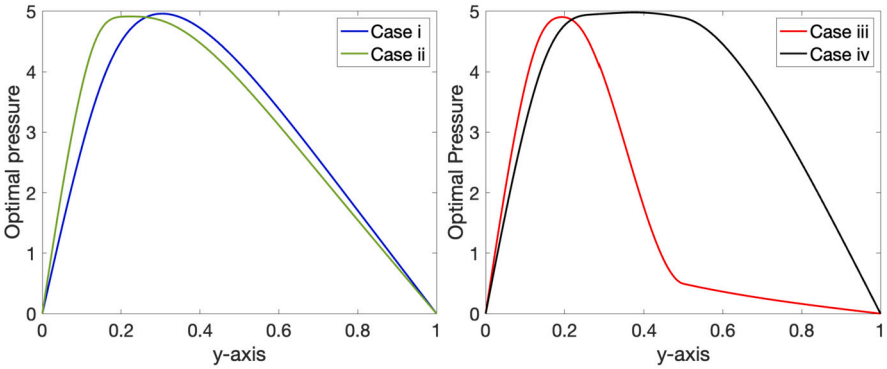


Fig. 3. Example 1. Optimal state $\bar{p}_h$ values over the line $x = 0.5$ for each case. We note that $p_d = p_{\mathscr{E}} = 4.84$.

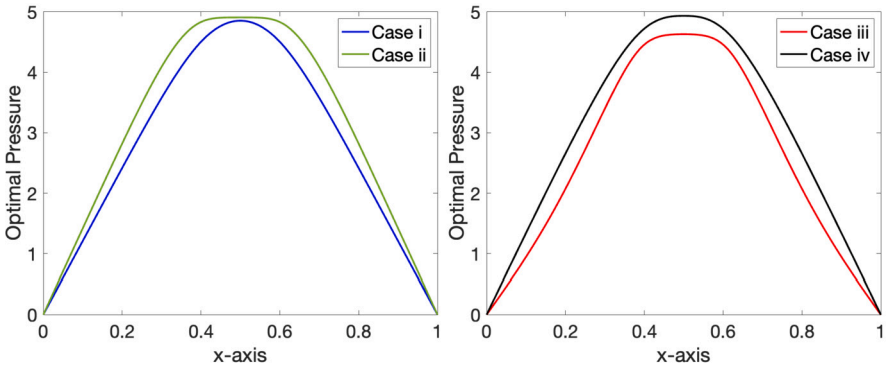


Fig. 4. Example 1. Optimal state $\bar{p}_h$ values over the line $y = 0.25$ for each case. We note that $p_d = p_{\mathscr{E}} = 4.84$.

Table 4
Example 2. Convergence results for each different cases.

	Case i)		Case ii)		Case iii)	
h	E_h	order	E_h	order	E_h	order
2^{-3}	1.17e-02	-	7.04e-02	-	2.52e-02	-
2^{-4}	6.40e-03	0.87	4.60e-02	0.61	2.45e-02	0.04
2^{-5}	3.31e-03	0.95	2.67e-02	0.78	1.57e-02	0.64
2^{-6}	1.63e-03	1.02	1.20e-02	1.15	7.70e-03	1.03
2^{-7}	6.33e-04	1.36	4.04e-03	1.57	2.91e-03	1.40

Table 5
Example 2. Energy norm of the finest result for each different cases.

$\|\bar{p}\|_h$	Case i)	Case ii)	Case iii)
2^{-8}	1.56e+03	1.47e+03	2.10e+03

5.4.2. Discussions

Fig. 7 illustrates the optimal solutions for the state variable $\bar{p}_h$, and the control variable $\bar{\eta}_h$ computed from the $\bar{p}_h$ for Case i). In Fig. 7 (a) and (b), it is evident that the point tracking on $\mathscr{P}_1$ is active, and we observe that the pressure value reached the desired state pressure values $p_d = p_{\mathscr{P}}$ for the point. The control variable $\bar{\eta}_h$ from Fig. 7 (c) shows the optimal source function values to obtain the desired state values.

Case ii) results are shown in Fig. 8. In this case, two lines ($\mathscr{C}_1$ and $\mathscr{C}_2$) are given as the targeted region, which could be considered as fractures in porous media. In Fig. 8 (a)-(b), we observe that the state pressure profile along the fracture lines are reaching the $p_{\mathscr{C}}$ value. Next, Fig. 8 (c)-(d) present the control values $\bar{\eta}_h$ computed from $\bar{p}_h$. The $\bar{\eta}_h$ corresponds to the optimally controlled source function value required to attain the targeted pressure values. Notably, $\bar{\eta}_h$ values are slightly higher at the end of the lines (or near the fracture tips).

Case iii) presented in Fig. 9 demonstrates the capabilities of our algorithm to combine multiple states. We observe that the pressure values $\bar{p}_h$ are reaching values p_d. Specifically, $p_{\mathscr{P}}$ for the point and $p_{\mathscr{C}}$ for the line. The value of $\bar{\eta}_h$ in Fig. 9 (c) - (d) represents the

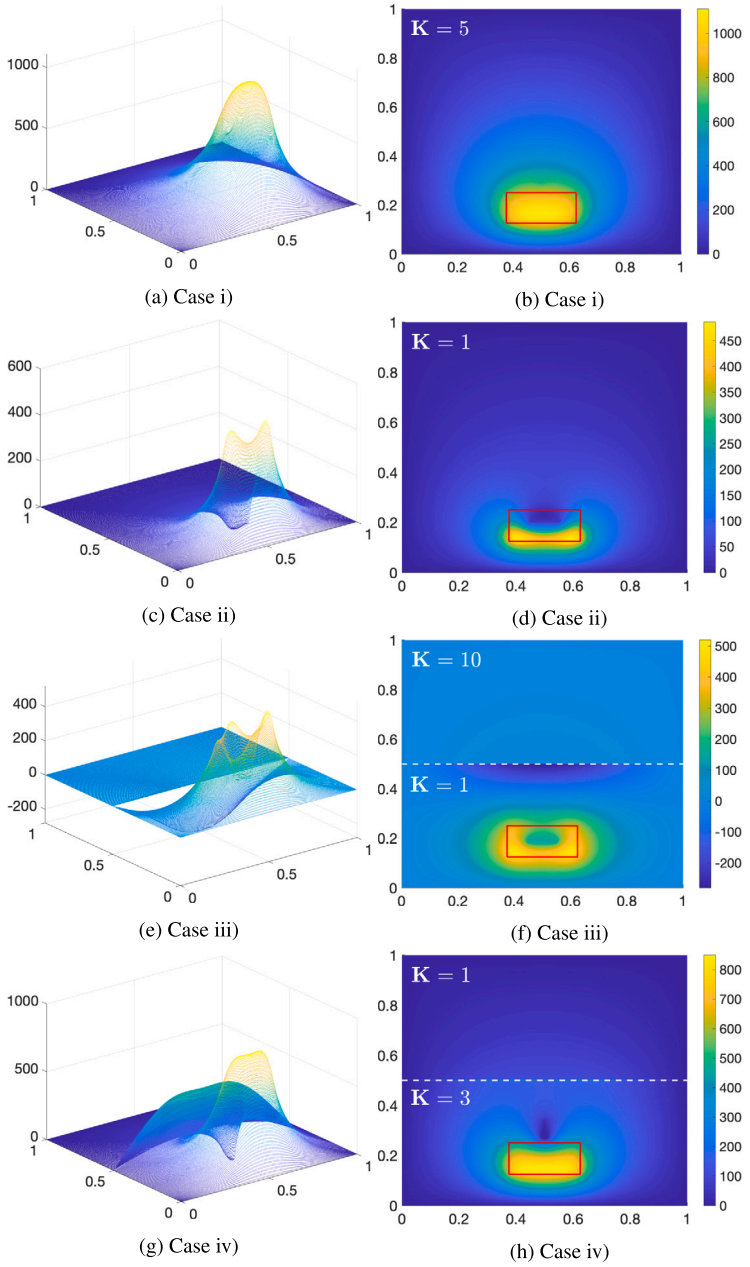


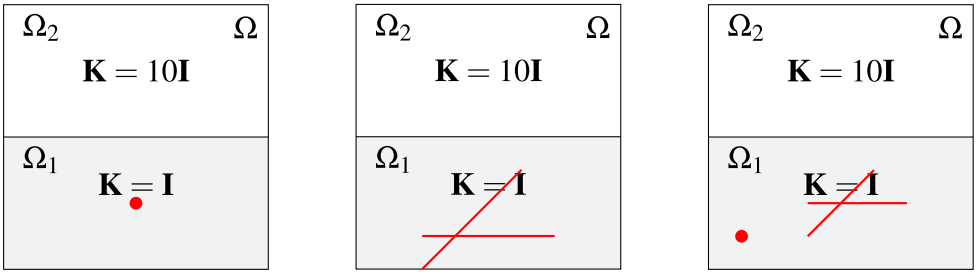
Fig. 5. Example 1. Illustrates the control variable $\bar{\eta}_h$ for each different case. We observe discontinuous values on the left column figures due to the choice of our C^0 interior penalty finite element space. The figures on the right column show the top view including the tracking domain.

optimal control variable $\bar{\eta}_h$ needed to achieve the target pressure values in Fig. 9 (a) - (b). Clearly, our algorithm can seek optimal state values with multiple tracking, as demonstrated in Fig. 9. To achieve the desired targeted pressure on both the point and the fracture, as shown in Fig. 9 (c), we observe that having a larger $\bar{\eta}_h$ value at the point is crucial, given that the area is much smaller compared to the lines.

We have also calculated the cost functional values at each PDAS iteration for Case iii) with $h = 2^{-5}$ in Table 6. The cost functional value is computed by the following formula

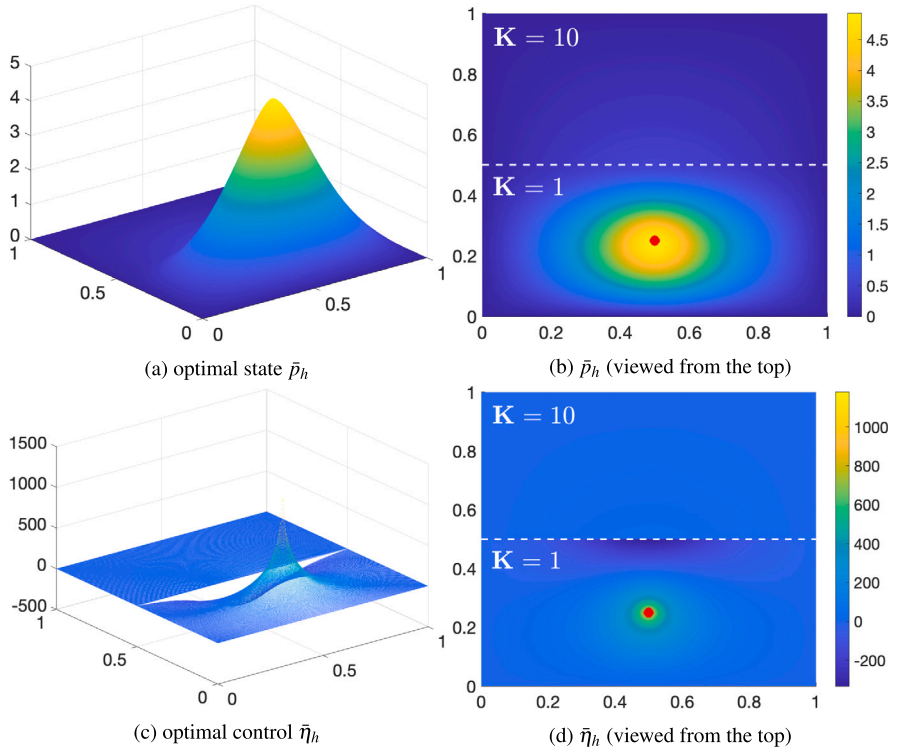
$$J := \frac{1}{2}\left[\int_{\Omega} |p_h - p_d|^2 \, dv + \beta b_h(p_h, p_h)\right]. \tag{5.3}$$

As the iteration runs, the value of the cost functional, J, converges to 4.07e+06 while the optimal solution satisfies the given constraints.



(a) Case i) a source point (b) Case ii) lines (fractures) (c) Case iii) a mixture

Fig. 6. Example 2. Three different cases.



(a) optimal state $\bar{p}_h$ (b) $\bar{p}_h$ (viewed from the top)

(c) optimal control $\bar{\eta}_h$ (d) $\bar{\eta}_h$ (viewed from the top)

Fig. 7. Example 2. Case i) Illustrates the optimal solutions for the state variable $\bar{p}_h$ and computed control variable $\bar{\eta}_h$. The right column figures present the top view of each value with the tracking domain (a point $\mathscr{P}_1$).

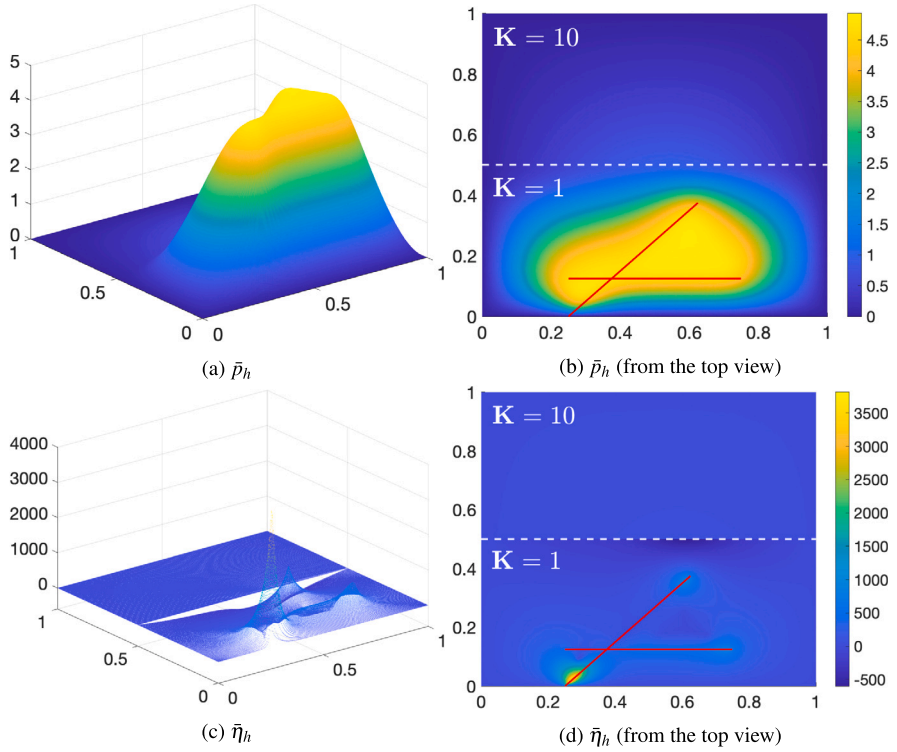
Table 6
Example 2. Case iii) Cost functional value at each PDAS iteration when $h = 2^{-5}$.

PDAS iteration	1	2	3	4	5	6	7	8	9
$J/(1.0e{+}06)$	4.10	4.08	4.07	4.07	4.07	4.07	4.07	4.07	4.07

Finally, Fig. 10 shows the evolution of the control $\bar{\eta}_h$ as the PDAS iteration runs. The last figure in Fig. 10, control evolution when PDAS iteration 28, presents the converged $\bar{\eta}_h$.

5.5. Example 3. General heterogeneity

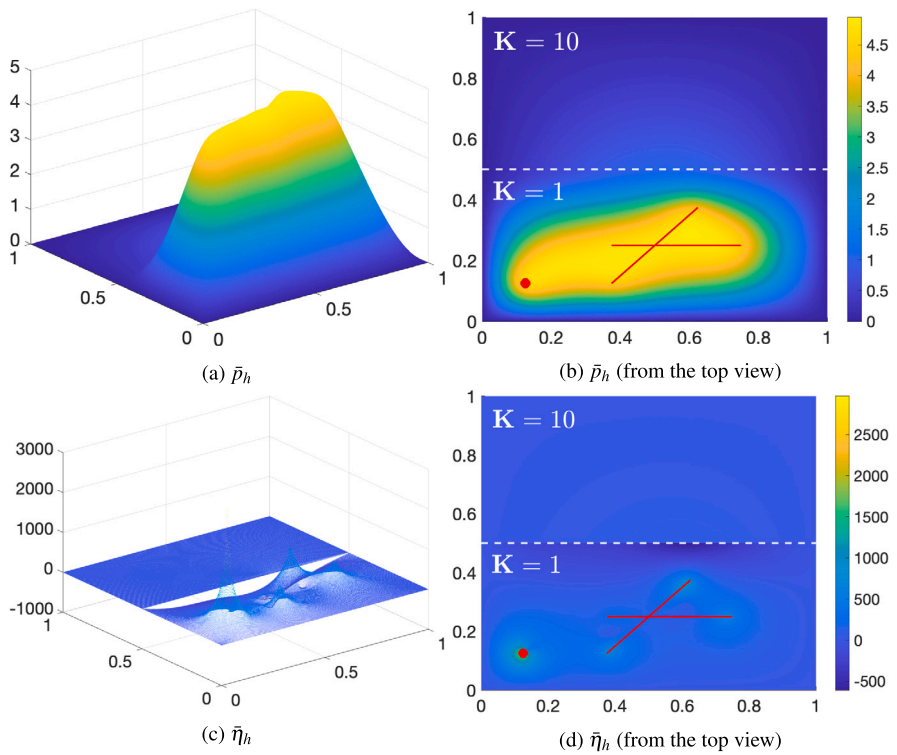
In this example, the domain is given as $\Omega = (0,1)^2$ and we test two different heterogeneous permeabilities with two line tracking that can be considered as an injection (stream) in the geothermal energy power plant. Thus, $J = 0$, $L = 2$, and $M = 0$, and $\mathscr{C}_1 = (0.25, 0.25) - (0.25, 1)$ and $\mathscr{C}_2 = (0.75, 0.25) - (0.75, 1)$. The target functions are $p^1_{\mathscr{C}} = 4.17$ and $p^2_{\mathscr{C}} = 1.82$, and the weights are $w^1_{\mathscr{C}} = w^2_{\mathscr{C}} = 10^5$. The penalty parameters are chosen as $\beta = 1$ and $\sigma = 10$. In this example, we set the state constraints $p_{\pm}$ given by



(a) $\bar{p}_h$ (b) $\bar{p}_h$ (from the top view)

(c) $\bar{\eta}_h$ (d) $\bar{\eta}_h$ (from the top view)

Fig. 8. Example 2. Case ii) Illustrates the optimal solutions for the state variable $\bar{p}_h$ and computed control variable $\bar{\eta}_h$. The right column figures present the top view of each value with the tracking domain (lines $\mathscr{C}_1$ and $\mathscr{C}_2$).



(a) $\bar{p}_h$ (b) $\bar{p}_h$ (from the top view)

(c) $\bar{\eta}_h$ (d) $\bar{\eta}_h$ (from the top view)

Fig. 9. Example 2. Case iii) Illustrates the optimal solutions for the state variable $\bar{p}_h$ and computed control variable $\bar{\eta}_h$. The right column figures present the top view of each value with the tracking domains (a point $\mathscr{P}_1$, and lines $\mathscr{C}_1$, and $\mathscr{C}_2$).

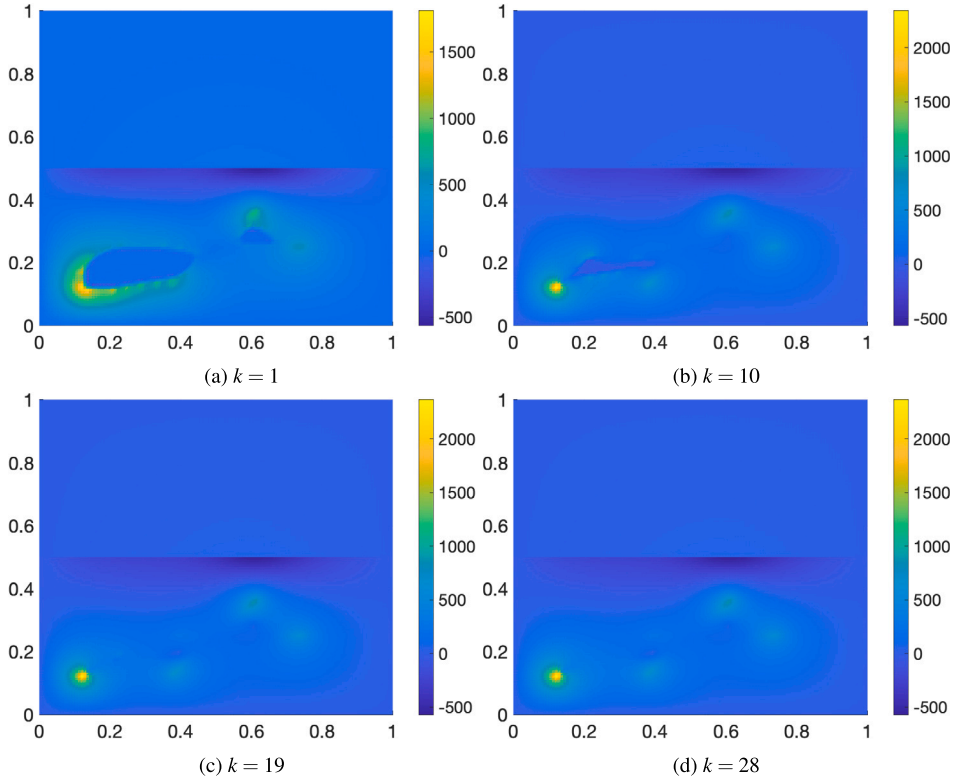


Fig. 10. Example 2. Case iii) Control values for each PDAS iteration k with h^{-7}.

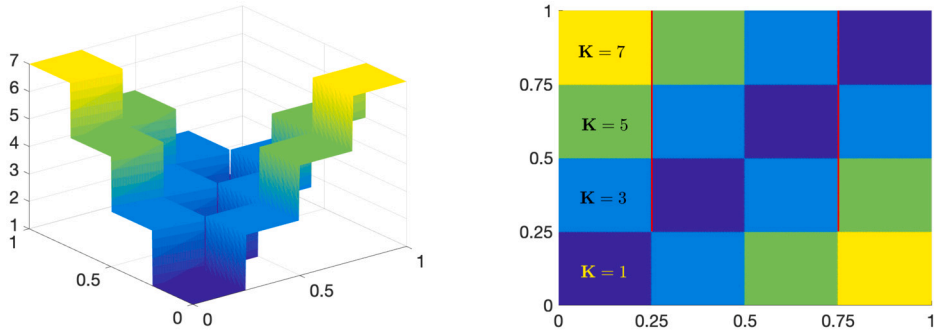


Fig. 11. Example 3. Case i) Illustrations of heterogeneous $\mathbf{K}(x, y)$ with the tracking lines.

$$p_+(x, y) = 2\cos(2\pi(x - 0.1)) + 3 \quad \text{and} \quad p_-(x, y) = -\infty,$$

First, in Case i), we set the permeability as $\mathbf{K}(x, y) = K(x, y)\mathbf{I}$ as shown in Fig. 11. Next, in Case ii), we set the permeability as (see Fig. 12) $\mathbf{K}(x, y) = K(x, y)\mathbf{I}$ where

$$K(x, y) = \begin{cases} 1 & \text{if} \quad y < -x + 0.5 \\ 3 & \text{if} \quad -x + 0.5 \le y < -x + 1 \\ 5 & \text{if} \quad -x + 1 \le y < -x + 1.5 \\ 7 & \text{if} \quad -x + 1.5 \le y \end{cases}.$$

5.5.1. Convergence

First, to validate the optimal convergence rate of the proposed method, we compute errors as in previous Examples. The convergence results are shown in Tables 7, 8, presenting the expected optimal convergence rates, which is $1 - \varepsilon$, $\varepsilon > 0$.

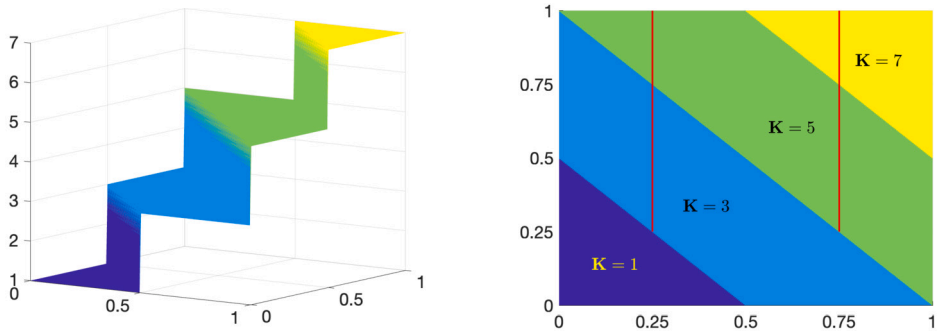


Fig. 12. Example 3. Case ii) Illustrations of heterogeneous $\mathbf{K}(x, y)$ with two tracking lines.

Table 7
Example 3. Convergence results for each different cases.

	Case i)		Case ii)	
h	E_h	order	E_h	order
2^{-3}	3.92e-02	-	3.11e-02	-
2^{-4}	2.18e-02	0.85	2.07e-02	0.59
2^{-5}	1.27e-02	0.78	1.43e-02	0.53
2^{-6}	9.20e-03	0.47	8.93e-03	0.68
2^{-7}	5.34e-03	0.78	4.42e-03	1.01

Table 8
Example 3. Energy norm of the finest result for each different cases.

$\|\bar{p}\|_h$	Case i)	Case ii)
2^{-8}	1.13e+03	1.19e+03

Table 9
Example 3. Case ii) Cost functional value at each PDAS iteration when $h = 2^{-6}$.

PDAS iteration	1	4	6	8	11	16	21	24	29
cost/(1.0e+06)	1.66	1.40	1.32	1.27	1.23	1.21	1.20	1.20	1.20

5.5.2. Discussions

Fig. 13 shows the optimal solutions for the state $\bar{p}_h$, and the control $\bar{\eta}_h$ for Case i). In Fig. 13 (a) and (b), it is clear that the state $\bar{p}_h$ reached two different desired states on each line tracking. The control $\bar{\eta}_h$ from Fig. 13 (c) and (d), displays the optimal source values required to achieve the desired state values on two line tracking. It has shown that to get the greater desired state value (the desired state for $\mathscr{C}_1$ is greater than for $\mathscr{C}_2$), the source should also be more powerful in the stream.

Fig. 14 depicts the optimal solutions for the state $\bar{p}_h$, and the control $\bar{\eta}_h$ for Case ii). In Fig. 14 (a) and (b), it is demonstrated that the state $\bar{p}_h$ reaches the desired states on $\mathscr{C}_1$ and $\mathscr{C}_2$. The control $\bar{\eta}_h$ from Fig. 14 (c) and (d), displays the optimal source function values to attain the desired states. As shown in Fig. 13 (c) and 14 (c), the profile of the control $\bar{\eta}_h$ follows the behavior of the given permeability $\mathbf{K}$.

We have also computed the cost functional values at each PDAS iteration for Example 3, Case ii) with $h = 2^{-6}$ in Table 9. As the iteration runs, the value of the cost functional converges to 1.20e+06 while the optimal solution satisfies the constraints.

Finally, in Fig. 15, we provide the active sets for Case i) and Case ii). These active sets are obtained at the last loop in the primal-dual active set algorithm when the algorithm terminates. Since we have only an upper state constraint for both cases, it shows the region in blue where the state reaches the upper pointwise constraint p_+. In other words, the optimal state approximation $\bar{p}_h$ contacts the pointwise constraint in blue nodes while the rest of the part of $\bar{p}_h$ is less than the constraint (blank part in the Fig. 15.)

6. Conclusion

In this work, we developed a computational framework for the optimal control problem for solving Darcy's flow equation in a heterogeneous porous media. Different layers for the permeability and various setups for the tracking domain (e.g. line, point, region) were utilized to simulate realistic scenarios. Convergence analysis is derived and validated with several test cases, and numerical examples are presented to illustrate the capabilities of the proposed method.

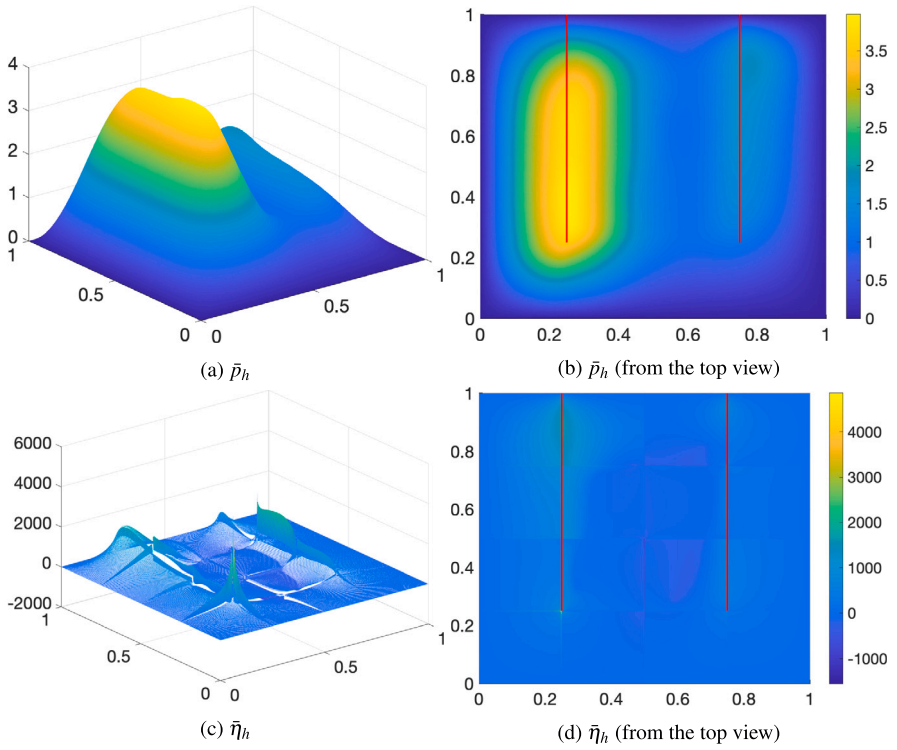


Fig. 13. Example 3. Case i) Illustrates the optimal solutions for the state variable $\bar{p}_h$ and computed control variable $\bar{\eta}_h$. The right column figures present the top view of each value.

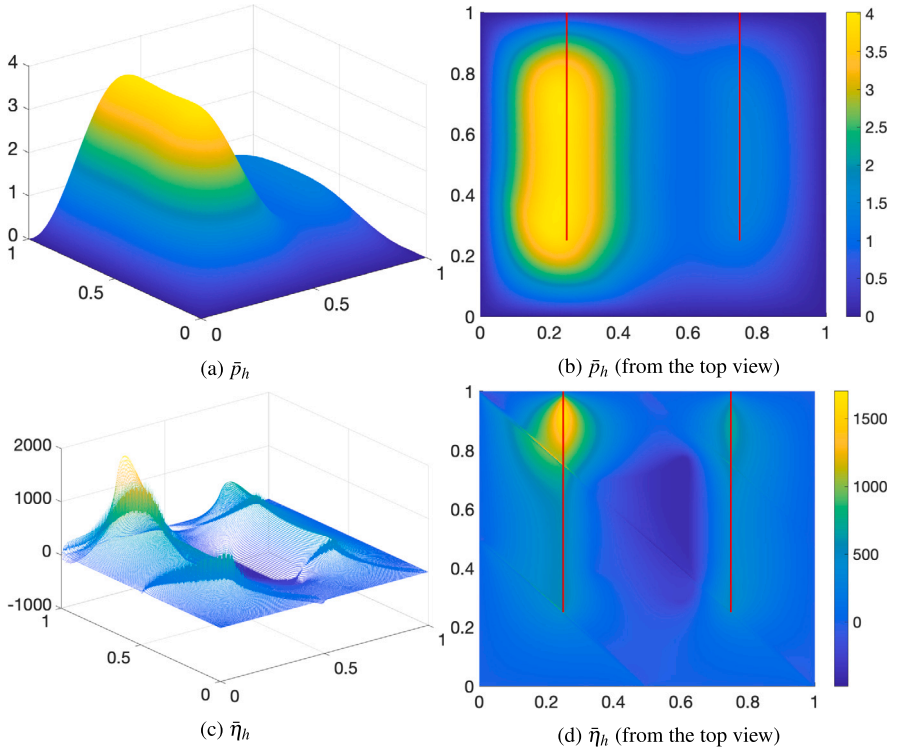


Fig. 14. Example 3. Case ii) Illustrates the optimal solutions for the state variable $\bar{p}_h$ and computed control variable $\bar{\eta}_h$. The right column figures present the top view of each value.

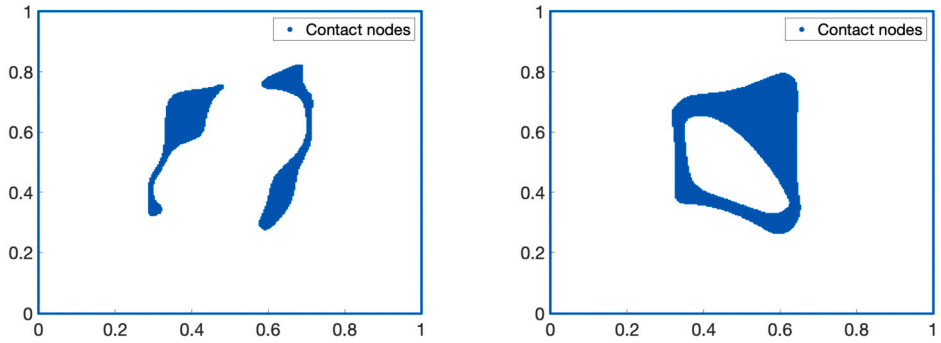


Fig. 15. Example 3. Illustrations of active sets for Case i) and Case ii). (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

CRediT authorship contribution statement

SeongHee Jeong: Writing – review & editing, Writing – original draft, Visualization, Software, Investigation, Formal analysis. **Sanghyun Lee:** Writing – review & editing, Writing – original draft, Visualization, Investigation.

Acknowledgements

The work by S. Lee is based upon the work supported by the National Science Foundation under Grant DMS-2208402.

References

[1] R.A. Freeze, P. Witherspoon, Theoretical analysis of regional groundwater flow: 2. effect of water-table configuration and subsurface permeability variation, Water Resour. Res. 3 (2) (1967) 623–634.
[2] C. Scheidt, L. Li, J. Caers, Quantifying Uncertainty in Subsurface Systems, vol. 236, John Wiley & Sons, 2018.
[3] M. Hinze, R. Pinnau, M. Ulbrich, S. Ulbrich, Optimization with PDE Constraints, vol. 23, Springer Science & Business Media, 2008.
[4] J.L. Lions, Optimal Control of Systems Governed by Partial Differential Equations, vol. 170, Springer, 1971.
[5] F. Tröltzsch, Optimal Control of Partial Differential Equations: Theory, Methods, and Applications, vol. 112, American Mathematical Soc., 2010.
[6] A.S. Mayer, C.T. Kelley, C.T. Miller, Optimal design for problems involving flow and transport phenomena in saturated subsurface systems, Adv. Water Resour. 25 (8–12) (2002) 1233–1256.
[7] M.D. Aliyu, H.-P. Chen, Optimum control parameters and long-term productivity of geothermal reservoirs using coupled thermo-hydraulic process modelling, Renew. Energy 112 (2017) 151–165.
[8] M. Gisser, D.A. Sanchez, Competition versus optimal control in groundwater pumping, Water Resour. Res. 16 (4) (1980) 638–642.
[9] S.C. Brenner, L.-y. Sung, Y. Zhang, A quadratic C^0 interior penalty method for an elliptic optimal control problem with state constraints, in: Recent Developments in Discontinuous Galerkin Finite Element Methods for Partial Differential Equations: 2012 John H Barrett Memorial Lectures, 2014, pp. 97–132.
[10] S.C. Brenner, L.-y. Sung, Z. Tan, A cubic C^0 interior penalty method for elliptic distributed optimal control problems with pointwise state and control constraints, Results Appl. Math. 7 (2020) 100119.
[11] S.C. Brenner, L.-y. Sung, Y. Zhang, C^0 interior penalty methods for an elliptic state-constrained optimal control problem with Neumann boundary condition, J. Comput. Appl. Math. 350 (2019) 212–232.
[12] S.C. Brenner, S. Jeong, L.-y. Sung, Z. Tan, C^0 interior penalty methods for an elliptic distributed optimal control problem with general tracking and pointwise state constraints, Comput. Math. Appl. 155 (2024) 80–90.
[13] W. Liu, W. Gong, N. Yan, A new finite element approximation of a state-constrained optimal control problem, J. Comput. Math. (2009) 97–114.
[14] J. Frehse, On the regularity of the solution of the biharmonic variational inequality, Manuscr. Math. 9 (1) (1973).
[15] R.A. Adams, J.J. Fournier, Sobolev Spaces, Elsevier, 2003.
[16] L.C. Evans, Partial Differential Equations, vol. 19, American Mathematical Society, 2022.
[17] V. Maz'ya, Sobolev Spaces, Springer, 2013.
[18] S.C. Brenner, L.R. Scott, The Mathematical Theory of Finite Element Methods, Springer, 2008.
[19] J.H. Argyris, I. Fried, D.W. Scharpf, The tuba family of plate elements for the matrix displacement method, Aeronaut. J. 72 (692) (1968) 701–709.
[20] R.W. Clough, J.L. Tocher, Finite element stiffness matrices for analysis of plate bending, in: Proceedings of the Conference on Matrix Methods in Structural Mechanics, Wright Patterson A.F.B., Ohio, 1965.
[21] S.C. Brenner, L.-y. Sung, C^0 interior penalty methods for fourth order elliptic boundary value problems on polygonal domains, J. Sci. Comput. 22 (1–3) (2005) 83–118.
[22] S.C. Brenner, C^0 interior penalty methods, in: Frontiers in Numerical Analysis-Durham 2010, Springer, 2011, pp. 79–147.
[23] S.C. Brenner, T. Gudi, L.-y. Sung, An a posteriori error estimator for a quadratic C^0 interior penalty method for the biharmonic problem, IMA J. Numer. Anal. 30 (3) (2010) 777–798.
[24] G. Engel, K. Garikipati, T. Hughes, M. Larson, L. Mazzei, R.L. Taylor, Continuous/discontinuous finite element approximations of fourth-order elliptic problems in structural and continuum mechanics with applications to thin beams and plates, and strain gradient elasticity, Comput. Methods Appl. Mech. Eng. 191 (34) (2002) 3669–3750.
[25] S.C. Brenner, L.-y. Sung, A new convergence analysis of finite element methods for elliptic distributed optimal control problems with pointwise state constraints, SIAM J. Control Optim. 55 (4) (2017) 2289–2304.
[26] S.C. Brenner, T. Gudi, K. Porwal, L.-y. Sung, A Morley finite element method for an elliptic distributed optimal control problem with pointwise state and control constraints, ESAIM Control Optim. Calc. Var. 24 (3) (2018) 1181–1206.
[27] S.C. Brenner, L.y. Sung, J. Gedicke, P_1 finite element methods for an elliptic optimal control problem with pointwise state constraints, IMA J. Numer. Anal. 40 (1) (2020) 1.

[28] S. Brenner, M. Oh, L.-y. Sung, P_1 finite element methods for an elliptic state-constrained distributed optimal control problem with Neumann boundary conditions, Results Appl. Math. 8 (2020) 100090.
[29] A. Allendes, F. Fuica, E. Otarola, Error estimates for a pointwise tracking optimal control problem of a semilinear elliptic equation, SIAM J. Control Optim. 60 (3) (2022) 1763–1790.
[30] A. Allendes, E. Otarola, R. Rankin, A.J. Salgado, Adaptive finite element methods for an optimal control problem involving Dirac measures, Numer. Math. 137 (1) (2017) 159–197.
[31] N. Behringer, D. Meidner, B. Vexler, Finite element error estimates for optimal control problems with pointwise tracking, arXiv preprint, arXiv:1802.02918, 2018.
[32] C. Brett, A. Dedner, C. Elliott, Optimal control of elliptic pdes at points, IMA J. Numer. Anal. 36 (3) (2016) 1015–1050.
[33] C. Brett, C.M. Elliott, M. Hintermüller, C. Löbhard, Mesh adaptivity in optimal control of elliptic variational inequalities with point-tracking of the state, Interfaces Free Bound. 17 (1) (2015) 21–53.
[34] E. Casas, L_2 estimates for the finite element method for the Dirichlet problem with singular data, Numer. Math. 47 (1985) 627–632.
[35] M. Bergounioux, K. Kunisch, Primal-dual strategy for state-constrained optimal control problems, Comput. Optim. Appl. 22 (2002) 193–224.
[36] M. Hintermüller, K. Ito, K. Kunisch, The primal-dual active set strategy as a semismooth Newton method, SIAM J. Optim. 13 (3) (2002) 865–888.
[37] D. Kinderlehrer, G. Stampacchia, An Introduction to Variational Inequalities and Their Applications, SIAM, 2000.
[38] K. Ito, K. Kunisch, Lagrange Multiplier Approach to Variational Problems and Applications, SIAM, 2008.
[39] W. Rudin, Real and Complex Analysis, Mathematics Series, McGraw-Hill, 1987.
[40] S. Jeong, Finite element methods for elliptic optimal control problems with general tracking, Ph.D. thesis, Louisiana State University, 2023.
[41] S. Agmon, A. Douglis, L. Nirenberg, Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions. I, Commun. Pure Appl. Math. 12 (4) (1959) 623–727.
[42] L. Hörmander, The Analysis of Linear Partial Differential Operators. III, Springer-Verlag, Berlin, 1985.
[43] P. Grisvard, Elliptic Problems in Nonsmooth Domains, SIAM, 2011.
[44] M. Dauge, Elliptic Boundary Value Problems on Corner Domains, Lecture Notes in Mathematics, vol. 1341, 1988.
[45] J.H. Bramble, S. Hilbert, Estimation of linear functionals on Sobolev spaces with application to Fourier transforms and spline interpolation, SIAM J. Numer. Anal. 7 (1) (1970) 112–124.
[46] T. Dupont, R. Scott, Polynomial approximation of functions in Sobolev spaces, Math. Comput. 34 (150) (1980) 441–463.
[47] S.C. Brenner, L.-y. Sung, H. Zhang, Y. Zhang, A quadratic C^0 interior penalty method for the displacement obstacle problem of clamped Kirchhoff plates, SIAM J. Numer. Anal. 50 (6) (2012) 3329–3350.
[48] M. Bergounioux, K. Ito, K. Kunisch, Primal-dual strategy for constrained optimal control problems, SIAM J. Control Optim. 37 (4) (1999) 1176–1194.
[49] M. Hintermüller, K. Kunisch, Feasible and noninterior path-following in constrained minimization with low multiplier regularity, SIAM J. Control Optim. 45 (4) (2006) 1198–1221.
[50] M. Bergounioux, M. Haddou, M. Hintermüller, K. Kunisch, A comparison of a Moreau–Yosida-based active set strategy and interior point methods for constrained optimal control problems, SIAM J. Optim. 11 (2) (2000) 495–521.
[51] M. Hintermüller, Mesh independence and fast local convergence of a primal-dual active-set method for mixed control-state constrained elliptic control problems, ANZIAM J. 49 (1) (2007) 1–38.
[52] M. Hintermüller, I. Yousept, A sensitivity-based extrapolation technique for the numerical solution of state-constrained optimal control problems, ESAIM Control Optim. Calc. Var. 16 (3) (2010) 503–522.
[53] M. Hintermüller, F. Tröltzsch, I. Yousept, Mesh-independence of semismooth Newton methods for Lavrentiev-regularized state constrained nonlinear optimal control problems, Numer. Math. 108 (4) (2008) 571–603.
[54] M. Bergounioux, K. Kunisch, Augmented Lagrangian techniques for elliptic state constrained optimal control problems, SIAM J. Control Optim. 35 (5) (1997) 1524–1543.
[55] M. Hintermüller, M. Hinze, Moreau–Yosida regularization in state constrained elliptic control problems: error estimates and parameter adjustment, SIAM J. Numer. Anal. 47 (3) (2009) 1666–1683.
[56] M. Hintermüller, K. Kunisch, PDE-constrained optimization subject to pointwise constraints on the control, the state, and its derivative, SIAM J. Optim. 20 (3) (2010) 1133–1156.
[57] M. Hintermuüller, A. Schiela, W. Wollner, The length of the primal-dual path in Moreau–Yosida-based path-following methods for state constrained optimal control, SIAM J. Optim. 24 (1) (2014) 108–126.