

A multigrid multilevel Monte Carlo method for transport in the Darcy–Stokes system



Prashant Kumar^{a,b,*}, Peiyao Luo^c, Francisco J. Gaspar^a, Cornelis W. Oosterlee^{a,c}

^a CWI – Centrum Wiskunde & Informatica, Amsterdam, the Netherlands

^b Faculty of Aerospace Engineering, Delft University of Technology, Delft, the Netherlands

^c DIAM, Delft University of Technology, Delft, the Netherlands

ARTICLE INFO

Article history:

Received 4 September 2017

Received in revised form 16 May 2018

Accepted 28 May 2018

Available online 31 May 2018

Keywords:

UQ

Darcy–Stokes flow

Contaminant transport

MLMC

Multigrid

Uzawa smoother

ABSTRACT

A multilevel Monte Carlo (MLMC) method for Uncertainty Quantification (UQ) of advection-dominated contaminant transport in a coupled Darcy–Stokes flow system is described. In particular, we focus on high-dimensional epistemic uncertainty due to an unknown permeability field in the Darcy domain that is modelled as a lognormal random field. This paper explores different numerical strategies for the subproblems and suggests an optimal combination for the MLMC estimator. We propose a specific monolithic multigrid algorithm to efficiently solve the steady-state Darcy–Stokes flow with a highly heterogeneous diffusion coefficient. Furthermore, we describe an Alternating Direction Implicit (ADI) based time-stepping for the flux-limited quadratic upwinding discretization for the transport problem. Numerical experiments illustrating the multigrid convergence and cost of the MLMC estimator with respect to the smoothness of permeability field are presented.

© 2018 Elsevier Inc. All rights reserved.

1. Introduction

Transport in a Darcy–Stokes flow system can be used to analyze a large number of dynamical processes. This model is, for example, of importance in the study of accidental discharge of radioactive contaminants or chemical spillage in surface water bodies and the subsequent transport to the connected aquifers. A coupled Darcy–Stokes system is then used to simulate the interaction between the surface water and the groundwater flow. The coupling is achieved by imposing interface conditions based on mass conservation, balancing the normal stress and a special condition, called the Beavers–Joseph–Saffman (BJS) interface condition [1,2] that relates the shear stress and tangential velocity along the interface. The steady-state velocity field derived from this model is then utilized for the advection of chemical components in a transport model. Many relevant physical phenomena concerning the mass transport such as molecular diffusion, mechanical dispersion, adsorption, can be conveniently incorporated in this model.

The mathematical theory and analysis of transport in a coupled Darcy–Stokes flow is well developed and a number of stable, convergent numerical methods have been proposed. These include finite volume (FV), mixed finite element (MFE) or more advanced, locally conservative discontinuous Galerkin (DG) schemes, see [3–10] and references therein. These schemes require the knowledge of physical quantities like fluid viscosity, permeabilities and experimentally measured BJS interface parameters for approximating the Darcy–Stokes flow. Furthermore, the chemical transport equation also requires input such

* Corresponding author.

E-mail addresses: pkumar@cwil.nl (P. Kumar), p.luo@tudelft.nl (P. Luo), gaspar@cwil.nl (F.J. Gaspar), c.w.oosterlee@cwil.nl (C.W. Oosterlee).

as initial and inflow boundary conditions. In many cases, complete information of these physical quantities is not available and they may be modelled in a probabilistic framework. For instance, it is well known that the permeability field can be modelled as a lognormal random field [11–13]. Once these uncertainties are incorporated in the mathematical models, the goal is to obtain the statistics of certain quantities of interest, for example, the mean spatial concentration of contaminants in an aquifer after a certain interval from the time of discharge.

Many Uncertainty Quantification (UQ) techniques exist in the literature but they are not universally applicable to all engineering problems. These techniques are broadly categorized into intrusive and non-intrusive approaches. Notably one of the most popular intrusive techniques is the stochastic Galerkin method based on a generalized Polynomial Chaos (gPC) expansion [14,15]. The stochastic Galerkin formulation for a coupled set of PDEs is however not straightforward, as existing iterative solvers cannot be directly utilized and may need severe modifications. Another class of well-established UQ techniques includes stochastic collocation methods [15,16] that are based on a deterministic sampling approach where the nodes in the random space are computed using *cubature rules*. The main advantage of the stochastic collocation method is its non-intrusiveness, hence, existing deterministic solvers can be utilized. The computational cost is governed by the number of nodes that however increases rapidly with an increase in the stochastic dimensions.

For our problem to have any practical relevance, the stochastic model should consider random fields with low spatial regularity and a small correlation length. This usually translates to a high-dimensional UQ problem [17] and the aforementioned deterministic approaches may not be computationally tractable. Historically, Monte Carlo (MC) type methods have been proven to be effective UQ tools for such problems as they do not suffer from the *curse of dimensionality*, are easy to implement and have a high parallelization potential. For the standard MC method, the statistical error converges as $(\text{Var}[Q]/N)^{\frac{1}{2}}$, where N is the number of MC samples and $\text{Var}[Q]$ is the variance of the Quantity of Interest (QoI). This slow convergence with respect to N is the main drawback of the method. Problems involving a spatio-temporal grid, ensuring a low root-mean-square error (RMSE) will require a large number of samples on a very fine mesh making the estimator expensive.

Over the past few years, a large number of sampling and variance reduction techniques have been proposed to overcome this limitation. For instance, the authors in [18] have applied a Quasi-Monte Carlo (QMC) sampling technique to improve the convergence rate for the stochastic Darcy flow subproblem. For the same problem, to reduce the grid size requirements, high-order numerical schemes were applied by [19] that resulted in lower asymptotic costs. More recently, the Monte Carlo method has been generalized to multiple grid levels, exhibiting an exceptional improvement over the standard MC [20,21]. The improved efficiency of these multilevel Monte Carlo (MLMC) methods comes from building the estimate for the QoI, on a hierarchy of grids, by exploiting the linearity of the expectation operator, i.e. $\mathbb{E}[Q_\ell] = \mathbb{E}[Q_0] + \sum_{\ell=1}^{\ell} \mathbb{E}[Q_\ell - Q_{\ell-1}]$. On the coarsest grid expectations are inexpensive to compute accurately and for large values of ℓ , where the numerical solution is comparatively expensive, fewer samples are required as the variance of the correction term $\text{Var}[Q_\ell - Q_{\ell-1}]$ is significantly smaller compared to the pure sampling variance, $\text{Var}[Q_\ell]$. While offering large savings over the standard MC method, MLMC retains all the important properties of MC methods like parallelization potential and combination with other complementary variance reduction techniques [18,22].

The purpose of this paper is to describe in a systematic manner a numerical strategy to design an efficient MLMC estimator for UQ of stochastic transport in the Darcy–Stokes system. An efficient MLMC estimator requires careful consideration of the numerical techniques for the approximation of the QoI. The paper is organized in the following way:

- In Section 2, we describe the stochastic transport in the Darcy–Stokes flow. The mixed formulation of the Darcy equation is used that is coupled with the Stokes flow using three interface conditions. The stochastic extension of the problem is obtained by modelling the permeability as a lognormal random field. The spatial covariance of the random field is derived from the parameterized Matérn function [23] that allows simulation of fields with different degrees of smoothness.
- In Section 3, the Finite Volume (FV) discretization of the Darcy–Stokes flow on a *staggered grid* is described. A special discretization along the interface is proposed taking into account the Beaver–Joseph–Saffmann interface condition that depends on the random permeability along the interface and a specific parameter. The optimal choice for the spatio-temporal discretization of the transport equation is based on the regularity of the Darcy–Stokes solution. Typically, the error in the FV approximation of the velocity field depends on the spatial regularity of the permeability field and also on the smoothness of boundary conditions along the two domains. Thus, the discretization scheme for the transport equation should be related to the accuracy of the velocity approximation. Using higher-order schemes for low regularity problems may lead to an expensive MLMC estimator without any improvement in the numerical accuracy. We also show that, for advection dominated transport with sharp gradients and discontinuities, low-order schemes are very diffusive and are less suited for MLMC applications.
- A monolithic multigrid solver for the Darcy–Stokes problem based on an Uzawa smoother [24] is proposed in Section 4. This smoother employs an equation-wise decoupled relaxation for the pressure and velocity unknowns. For the velocity a symmetric Gauss–Seidel iteration is employed whereas for the pressure a Richardson iteration is applied. The Richardson iteration takes into account local fluctuations in the permeability field to derive the optimal relaxation parameter. The multigrid algorithm is based on cell-centred permeability coefficients and direct coarse-grid discretization based on cell-centred averaging. The proposed solver is robust and also works well on very coarse grids. We generalize this multigrid method to multi-block problems by the *grid partitioning technique* [25]. To incorporate the coarse time-step in the MLMC hierarchy, an implicit time stepping is desired for the contaminant transport equation. We consider an

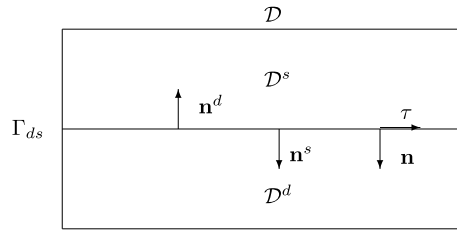


Fig. 1. Geometry of the Darcy–Stokes problem coupled with the transport equation. Subdivision of the domain \mathcal{D} into a free-flow subregion \mathcal{D}^s and a porous medium subdomain \mathcal{D}^d , by an internal interface Γ_{ds} .

Alternating Direction Implicit (ADI) based solver for the discrete transport equation which breaks the problem into two 1D problems, greatly reducing the cost of time-stepping.

- Single and multilevel Monte Carlo estimators are defined in Sections 5 and 6, respectively. Section 7 is devoted to numerical experiments performed on two test problems: a 2-block problem with no-slip interface condition and a more realistic 4-block case with the BJS interface condition. We thoroughly test the multigrid method with respect to different Matérn parameters to identify the optimal cycling strategy. Finally, we combine these components to obtain the solution of the stochastic transport problem using MLMC estimator and compare the asymptotic cost with the standard Monte Carlo method.

2. Stochastic transport in Darcy–Stokes system

We consider the transport equation defined in the bounded domain $\mathcal{D} \subset \mathbb{R}^2$, with boundary $\partial\mathcal{D}$ and in a finite time interval $\mathcal{T} = (0, T]$, for $T < \infty$. The transport equation coupled with the steady-state Darcy–Stokes flow on \mathcal{D} , is subdivided into a porous medium $\mathcal{D}^d \subset \mathbb{R}^2$, where the flow is described by Darcy’s law and the free-flow region $\mathcal{D}^s \subset \mathbb{R}^2$, governed by the Stokes equations with boundaries $\partial\mathcal{D}^d$ and $\partial\mathcal{D}^s$, respectively (Fig. 1). The internal interface is defined as $\Gamma_{ds} = \partial\mathcal{D}^d \cap \partial\mathcal{D}^s$. Furthermore, we denote by ω an event in the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω is the sample space with σ -field \mathcal{F} and probability measure \mathbb{P} . Our description of the stochastic flow model follows from [26–28] where the deterministic counterpart of this problem is considered.

Porous medium description. The steady-state single-phase flow in a porous medium can be modelled by Darcy’s law and the incompressibility condition

$$\begin{cases} \eta K^{-1} \mathbf{u}^d + \nabla p^d = \mathbf{0} & \text{in } \mathcal{D}^d \times \Omega, \quad (\text{a}) \\ \nabla \cdot \mathbf{u}^d = f^d & \text{in } \mathcal{D}^d \times \Omega. \quad (\text{b}) \end{cases} \tag{2.1}$$

The fluid pressure is represented by p^d and the velocity vector $\mathbf{u}^d = (u^d, v^d)$ denotes the horizontal and vertical components. The fluid viscosity is denoted by the positive constant η and K is the spatially variable permeability field. The known source (sink) term is indicated by $f^d \in L^2(\mathcal{D}^d)$, where $L^2(\mathcal{D}^d)$ is the space of square-integrable functions in \mathcal{D}^d .

We divide the boundary $\partial\mathcal{D}^d \setminus \Gamma_{ds}$ into two disjoint sets, $\partial\mathcal{D}_D^d$ and $\partial\mathcal{D}_N^d$, where deterministic Dirichlet and Neumann boundary conditions are prescribed,

$$\begin{cases} p^d = g_D^d & \text{on } \partial\mathcal{D}_D^d, \quad (\text{a}) \\ \mathbf{u}^d \cdot \mathbf{n} = g_N^d & \text{on } \partial\mathcal{D}_N^d, \quad (\text{b}) \end{cases} \tag{2.2}$$

respectively, with \mathbf{n} the outward normal to the boundary $\partial\mathcal{D}_N^d$ and $g_D^d \in L^2(\partial\mathcal{D}_D^d)$, $g_N^d \in L^2(\partial\mathcal{D}_N^d)$.

We will operate under the assumption that the solid framework is rigid and there is no interaction between the fluid and the solid matrix of the porous medium.

We model the permeability field by means of a lognormal field, i.e. $\log K(\mathbf{x}, \omega) := Z(\mathbf{x}, \omega)$ is a zero-mean Gaussian random field for $\mathbf{x} \in \overline{\mathcal{D}^d} = \mathcal{D}^d \cup \partial\mathcal{D}^d$ and $\omega \in \Omega$. Therefore

$$\begin{cases} \mathbb{E}[Z(\mathbf{x}, \cdot)] = 0, \\ \mathbf{cov}(Z(\mathbf{x}_1, \cdot), Z(\mathbf{x}_2, \cdot)) = \mathbb{E}[Z(\mathbf{x}_1, \cdot)Z(\mathbf{x}_2, \cdot)], \quad \mathbf{x}_1, \mathbf{x}_2 \in \overline{\mathcal{D}^d}. \end{cases} \tag{2.3}$$

The lognormal property ensures a positive permeability throughout the domain. For further simplification, we consider an isotropic and stationary Gaussian process, which can be obtained from a homogeneous covariance function $C_\phi : \mathbb{R} \rightarrow \mathbb{R}$ such that

$$\mathbf{cov}(Z(\mathbf{x}_1, \cdot), Z(\mathbf{x}_2, \cdot)) = C_\phi(r) \quad \text{with } r = \|\mathbf{x}_1 - \mathbf{x}_2\|_2. \tag{2.4}$$

For the problem to be well-posed, we assume

$$K_{\min} := \min K(\mathbf{x}, \omega) > 0 \quad \text{and} \quad K_{\max} := \max K(\mathbf{x}, \omega) < \infty. \tag{2.5}$$

To generate samples of the Gaussian random field Z , the covariance matrix is derived from the family of *Matérn functions* [29] characterized by the parameter set $\Phi = (\nu_c, \lambda_c, \sigma_c^2)$, and has the standard form

$$C_\Phi(r) = \sigma_c^2 \frac{2^{1-\nu_c}}{\Gamma(\nu_c)} \left(2\sqrt{\nu_c} \frac{r}{\lambda_c}\right)^{\nu_c} K_{\nu_c} \left(2\sqrt{\nu_c} \frac{r}{\lambda_c}\right). \tag{2.6}$$

Here, Γ is the gamma function and K_{ν_c} is the modified Bessel function of the second kind. The parameter $\nu_c \geq 1/2$ defines the smoothness, $\sigma_c^2 > 0$ is the variance and $\lambda_c > 0$ is the correlation length of the Gaussian process. The Matérn model has great flexibility in modelling spatial processes because of parameter ν_c , which governs the differentiability of the random field. For $\nu_c = 1/2$, the Matérn function corresponds to an exponential model and for $\nu_c \rightarrow \infty$ to a Gaussian model. Furthermore, the other two parameters λ_c and σ_c^2 dictate the number of peaks and the amplitude of the random field, respectively. The realizations of Gaussian random fields are Hölder continuous i.e., $Z(\mathbf{x}, \cdot) \in C^\kappa(\overline{\mathcal{D}}^d)$ almost surely with the exponent $0 < \kappa < \nu_c$ [30]. For $\nu_c > 1$, the realizations are continuously differentiable.

Free-flow description. We consider a creeping, steady-state flow connected to the porous flow regime. We assume a viscous, incompressible Newtonian fluid flow which can be modelled by the Stokes equations, consisting of the momentum and continuity equations

$$\begin{cases} -\nabla \cdot \boldsymbol{\sigma}^s = \mathbf{f}^s & \text{in } \mathcal{D}^s \times \Omega, & \text{(a)} \\ \nabla \cdot \mathbf{u}^s = \mathbf{0} & \text{in } \mathcal{D}^s \times \Omega, & \text{(b)} \end{cases} \tag{2.7}$$

respectively, with $\mathbf{u}^s = (u^s, v^s)$ as the fluid velocity components and $\mathbf{f}^s = (f_1^s, f_2^s) \in L^2(\mathcal{D}^s)^2$ is the force. The fluid stress tensor $\boldsymbol{\sigma}^s$ is defined by

$$\begin{cases} \boldsymbol{\sigma}^s = -p^s \mathbf{I} + 2\eta \mathbf{R}(\mathbf{u}^s), & \text{(a)} \\ \mathbf{R} := \frac{1}{2}(\nabla \mathbf{u}^s + (\nabla \mathbf{u}^s)^T), & \text{(b)} \end{cases} \tag{2.8}$$

where p^s denotes the fluid pressure and \mathbf{R} is the strain tensor. In 2D, the components of the stress tensor (2.8)(a) are given by

$$\boldsymbol{\sigma}^s := \begin{bmatrix} \sigma_{xx}^s & \sigma_{xy}^s \\ \sigma_{yx}^s & \sigma_{yy}^s \end{bmatrix}, \tag{2.9}$$

and is related to the primitive variables via

$$\begin{cases} \sigma_{xx}^s = -p^s + 2\eta \frac{\partial u^s}{\partial x}, \\ \sigma_{xy}^s = \sigma_{yx}^s = \eta \left(\frac{\partial u^s}{\partial y} + \frac{\partial v^s}{\partial x} \right), \\ \sigma_{yy}^s = -p^s + 2\eta \frac{\partial v^s}{\partial y}. \end{cases} \tag{2.10}$$

The exterior boundary $\partial \mathcal{D}^s \setminus \Gamma_{ds}$ is partitioned into two disjoint sets, $\partial \mathcal{D}_D^s$ and $\partial \mathcal{D}_N^s$, and the free-flow model is completed by imposing the following boundary conditions

$$\begin{cases} \mathbf{u}^s = \mathbf{g}_D^s & \text{on } \partial \mathcal{D}_D^s, & \text{(a)} \\ \boldsymbol{\sigma}^s \cdot \mathbf{n} = \mathbf{g}_N^s & \text{on } \partial \mathcal{D}_N^s, & \text{(b)} \end{cases} \tag{2.11}$$

where \mathbf{n} is the outward normal to the boundary $\partial \mathcal{D}_N^s$ and $\mathbf{g}_D^s \in L^2(\partial \mathcal{D}_D^s)^2$, $\mathbf{g}_N^s \in L^2(\partial \mathcal{D}_N^s)^2$.

Interface conditions. The coupling between the two problems is attained by imposing three interface conditions for mass conservation, normal stress balance and the third condition relating the slip velocity to the shear stress along Γ_{ds}

$$\begin{cases} \mathbf{u}^s \cdot \mathbf{n}^s = \mathbf{u}^d \cdot \mathbf{n}^s & \text{on } \Gamma_{ds} \times \Omega, & \text{(a)} \\ -\mathbf{n}^s \cdot \boldsymbol{\sigma}^s \cdot \mathbf{n}^s = p^d & \text{on } \Gamma_{ds} \times \Omega, & \text{(b)} \\ \mathbf{u}^s \cdot \boldsymbol{\tau} + \left(\frac{\sqrt{K}}{\alpha_{BJ}}\right) \boldsymbol{\tau} \cdot \boldsymbol{\sigma}^s \cdot \mathbf{n}^s = \mathbf{0} & \text{on } \Gamma_{ds} \times \Omega, & \text{(c)} \end{cases} \tag{2.12}$$

respectively. We denote by $\boldsymbol{\tau}$ and \mathbf{n}^s the unit tangential and normal vectors to the interface Γ_{ds} , respectively. The third interface condition (2.12)(c) was originally derived by Beaver and Joseph [1] on the basis of experimentation and dimensional analysis, and further has been mathematically proven by Saffmann [2]. It is often referred to as the Beaver–Joseph–Saffmann (BJS) condition. The literature on BJS is growing rapidly as it is well able to represent the physics along the interface of viscous fluid flow and a porous media, for more details see [8,31,32]. Note that friction constant $\frac{\sqrt{K}}{\alpha_{BJ}}$ depends on the permeability along the interface and is thus here a random variable. Parameter $\alpha_{BJ} > 0$ is a dimensionless quantity measured experimentally and can also be prone to uncertainty.

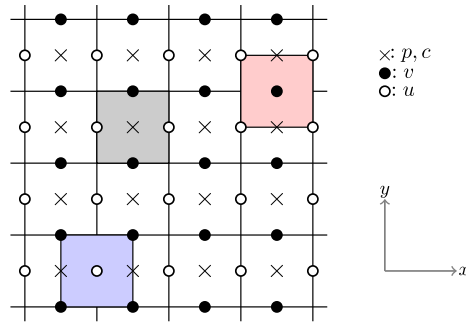


Fig. 2. Staggered grid location of unknown and corresponding control volumes.

Alternatively, one obtains a *no-slip interface* condition by neglecting the second term from (2.12)(c)

$$\mathbf{u}^s \cdot \boldsymbol{\tau} = \mathbf{0} \quad \text{on } \Gamma_{ds} \times \Omega. \tag{2.13}$$

Transport model. The generic single-component transport equation with random initial data gives us the following equation:

$$\begin{cases} \phi \frac{\partial c}{\partial t} + \nabla \cdot (\mathbf{c}\mathbf{u} - \mathbb{D}\nabla c) = \phi f^t & \text{in } \mathcal{D} \times \mathcal{T} \times \Omega, & \text{(a)} \\ c = c_0 & \text{in } \mathcal{D} \times \Omega, \quad t = 0, & \text{(b)} \end{cases} \tag{2.14}$$

where c denotes the concentration of the chemical component, typically expressed in terms of moles per unit volume; $\phi \in (0, 1]$ is the known porosity of the medium while $f^t \in L^2(\mathcal{D})$ is a net volumetric source for c . The initial concentration $c_0 := c_0(\mathbf{x}, \omega) \in L^2(\mathcal{D})$ corresponds to a random field. The steady-state velocity field $\mathbf{u} = \mathbf{u}^s \cup \mathbf{u}^d$ is derived from the solution of the coupled Darcy–Stokes problem. In the Darcy domain, the hydrodynamical dispersion is represented by the tensor field \mathbb{D} accounting for the molecular diffusion and mechanical dispersion. Molecular diffusion which takes place due to a concentration gradient that is more significant compared to the mechanical dispersion resulting from micro-scale variations in the velocity field compared to the average flow. In this paper, we consider a diffusion tensor defined in [33,34] for an isotropic porous media with tensor components:

$$\begin{cases} D_{xx} = D_L \frac{u^2}{|\mathbf{u}|} + D_T \frac{v^2}{|\mathbf{u}|} + D^*, \\ D_{yy} = D_L \frac{v^2}{|\mathbf{u}|} + D_T \frac{u^2}{|\mathbf{u}|} + D^*, \\ D_{xy} = D_{yx} = (D_L - D_T) \frac{uv}{|\mathbf{u}|}, \end{cases} \tag{2.15}$$

where $D_L, D_T > 0$ are longitudinal and transverse dispersivity, respectively, and $D^* > 0$ is the effective molecular diffusion. In the Stokes domain, the dispersion tensor is usually taken to be isotropic, i.e. $\mathbb{D} = D\mathbf{I}$.

The transport model is completed by applying a Cauchy boundary condition at the inflow boundary and a non-dispersive mass flux condition at outflow boundaries, formally expressed as

$$\begin{cases} (\mathbf{c}\mathbf{u} - \mathbb{D}\nabla c) \cdot \mathbf{n} = (c_{in}\mathbf{u}) \cdot \mathbf{n} & \partial\mathcal{D}_{in} \times \mathcal{T}, & \text{(a)} \\ \mathbb{D}\nabla c \cdot \mathbf{n} = 0 & \partial\mathcal{D}_{out} \times \mathcal{T}, & \text{(b)} \end{cases} \tag{2.16}$$

respectively. The inflow boundary is defined as $\partial\mathcal{D}_{in} := \{\mathbf{x} \in \partial\mathcal{D} : \mathbf{u} \cdot \mathbf{n} < 0\}$ and the outflow boundary as $\partial\mathcal{D}_{out} := \partial\mathcal{D} \setminus \partial\mathcal{D}_{in}$.

For ease of presentation, we consider a simplified chemical transport model ignoring the adsorption phenomena and also assume that the injection of the chemical component does not have any effect on the steady-state Darcy–Stokes flow field.

3. Finite volume discretization

For the spatial discretizations, we employ the finite volume scheme on a staggered grid for the Darcy–Stokes system. The domain is subdivided into square blocks of size $h \times h$ conforming with $\partial\mathcal{D}$ and Γ_{ds} . The locations and indexing of unknowns u, v and p, c along with their respective control volumes $\mathcal{D}_h^1, \mathcal{D}_h^2$ and \mathcal{D}_h^3 are shown in Figs. 2–3.

As the same variables describe different physics in the two subdomains, discretization along the interface becomes involved. A large numerical error or even a reduction in the order of grid convergence may be encountered if interface conditions are not handled properly. A staggered arrangement of the unknowns greatly simplifies the discretization along the interface and has been proven to be effective in reducing numerical error along the interface [28]. Moreover, a staggered grid is also a convenient way of avoiding spurious oscillations in the numerical solution [35] and obtaining conservation of mass throughout the system, also on a relatively coarse grid.

To make this paper self-contained, we briefly discuss the spatio-temporal discretization of the coupled problem. For the Darcy–Stokes approximation, we will closely follow the description from [28].

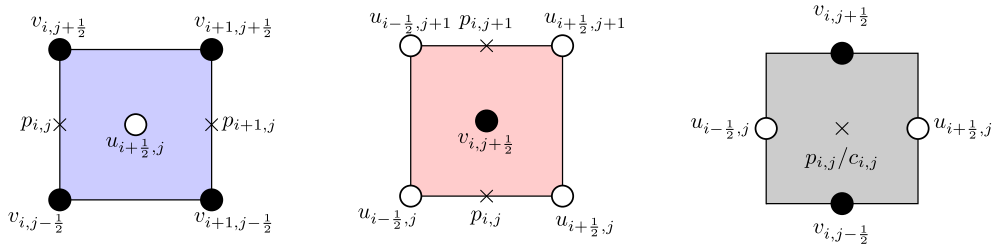


Fig. 3. Control volumes $\mathcal{D}_h^1(i + \frac{1}{2}, j)$ (left), $\mathcal{D}_h^2(i, j + \frac{1}{2})$ (middle) and $\mathcal{D}_h^3(i, j)$ (right) for the primary unknowns u , v and p , c respectively, together with the corresponding indexing for each variable.

3.1. Discretization of Darcy–Stokes flow

3.1.1. Discretization of Darcy equation

The discrete equations for velocities and pressure in the mixed formulation are easy to determine. The discretization for the horizontal velocity is obtained by integrating the Darcy equation (2.1)(a) in the control volume $\mathcal{D}_h^1(i + \frac{1}{2}, j)$

$$\eta K_{i+\frac{1}{2},j}^{-1} u_{i+\frac{1}{2},j}^d + \frac{p_{i+1,j}^d - p_{i,j}^d}{h} = 0. \tag{3.1}$$

Similarly, the discrete equations for the vertical velocities are obtained by integrating the Darcy equation in the control volume $\mathcal{D}_h^2(i, j + \frac{1}{2})$. For the pressure, we integrate the continuity equation (2.1)(b) in the volume $\mathcal{D}_h^3(i, j)$

$$\frac{u_{i+\frac{1}{2},j}^d - u_{i-\frac{1}{2},j}^d}{h} + \frac{v_{i,j+\frac{1}{2}}^d - v_{i,j-\frac{1}{2}}^d}{h} = f_{i,j}^d. \tag{3.2}$$

For boundaries where the pressure is prescribed (2.2)(a), we integrate the Darcy equation over half volumes. At boundaries for which the velocities are known, (2.2)(b) is directly applied.

3.1.2. Discretization of Stokes equation

Integrating the first component of the momentum equation (2.7)(a) in the control volumes $\mathcal{D}_h^2(i, j + \frac{1}{2})$,

$$-\left(\frac{(\sigma_{xx})_{i+1,j} - (\sigma_{xx})_{i,j}}{h} + \frac{(\sigma_{xy})_{i+\frac{1}{2},j+\frac{1}{2}} - (\sigma_{xy})_{i+\frac{1}{2},j-\frac{1}{2}}}{h} \right) = (f_1^s)_{i+\frac{1}{2},j}. \tag{3.3}$$

The components in the above equation are approximated using (2.10) as

$$\left\{ \begin{aligned} (\sigma_{xx})_{i+1,j} &= -p_{i+1,j}^s + 2\eta \frac{u_{i+\frac{3}{2},j}^s - u_{i+\frac{1}{2},j}^s}{h}, \\ (\sigma_{xx})_{i,j} &= -p_{i,j}^s + 2\eta \frac{u_{i+\frac{1}{2},j}^s - u_{i-\frac{1}{2},j}^s}{h}, \\ (\sigma_{xy})_{i+\frac{1}{2},j+\frac{1}{2}} &= \eta \left(\frac{u_{i+\frac{1}{2},j+1}^s - u_{i+\frac{1}{2},j}^s}{h} + \frac{v_{i+1,j+\frac{1}{2}}^s - v_{i,j+\frac{1}{2}}^s}{h} \right), \\ (\sigma_{xy})_{i+\frac{1}{2},j-\frac{1}{2}} &= \eta \left(\frac{u_{i+\frac{1}{2},j}^s - u_{i+\frac{1}{2},j-1}^s}{h} + \frac{v_{i+1,j-\frac{1}{2}}^s - v_{i,j-\frac{1}{2}}^s}{h} \right). \end{aligned} \right. \tag{3.4}$$

Using these approximations in (3.3), we finally get

$$-\frac{2\eta}{h^2} (u_{i+\frac{3}{2},j}^s - 2u_{i+\frac{1}{2},j}^s + u_{i-\frac{1}{2},j}^s) - \frac{\eta}{h^2} (u_{i+\frac{1}{2},j+1}^s - 2u_{i+\frac{1}{2},j}^s + u_{i+\frac{1}{2},j-1}^s) - \frac{\eta}{h^2} (v_{i+1,j+\frac{1}{2}}^s - v_{i,j+\frac{1}{2}}^s - v_{i+1,j-\frac{1}{2}}^s + v_{i,j-\frac{1}{2}}^s) + \frac{1}{h} (p_{i+1,j}^s - p_{i,j}^s) = (f_1^s)_{i+\frac{1}{2},j}. \tag{3.5}$$

Similarly, the second component of the momentum equation is integrated in the volume $\mathcal{D}_h^2(i, j + \frac{1}{2})$. The continuity equation (2.7)(b) is as given in (3.2) over the volume $\mathcal{D}_h^3(i, j)$.

The Dirichlet boundary condition (2.11)(a) can be directly utilized for the approximation of the stress component in (3.4). If the stress components are prescribed along the boundary (2.11)(b) then the Stokes equations are integrated over half volumes along that boundary.

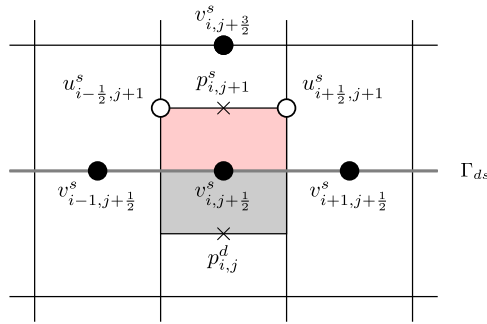


Fig. 4. Locations of unknowns required for discretization along the interface. (For interpretation of the colours in the figure(s), the reader is referred to the web version of this article.)

3.1.3. Discretization of interface equations

An appropriate interface discretization is crucial for achieving a strong numerical coupling between two sub-solutions. As the vertical velocities lie along the interface, we integrate the Stokes equation for v^s over the half volume indicated in red in Fig. 4,

$$-\left(\frac{(\sigma_{xy})_{i+\frac{1}{2},j+\frac{1}{2}} - (\sigma_{xy})_{i-\frac{1}{2},j+\frac{1}{2}}}{h} + \frac{(\sigma_{yy})_{i,j+1} - (\sigma_{yy})_{i,j+\frac{1}{2}}}{h/2}\right) = (f_2^s)_{i,j+\frac{1}{2}}. \tag{3.6}$$

Now, we describe the computation of these stress components in the above equation. To approximate these terms, we will utilize the interface conditions (2.12)(a)–(c).

For the normal stress component $(\sigma_{yy})_{i,j+1}$, we use

$$(\sigma_{yy})_{i,j+1} = -p_{i,j+1}^s + 2\eta \frac{v_{i,j+\frac{3}{2}}^s - v_{i,j+\frac{1}{2}}^s}{h}. \tag{3.7}$$

Next, the normal stress $(\sigma_{yy})_{i,j+\frac{1}{2}}$ at the interface is derived by equilibrating the normal stress with the pressure at the interface

$$(\sigma_{yy})_{i,j+\frac{1}{2}} = p_{i,j+\frac{1}{2}}^d. \tag{3.8}$$

Pressure $p_{i,j+\frac{1}{2}}^d$ at the interface is not known and is obtained by integrating the Darcy equation over the half volume indicated by the gray box in Fig. 4, as

$$\eta K_{i,j+\frac{1}{2}}^{-1} v_{i,j+\frac{1}{2}}^d + \frac{p_{i,j+\frac{1}{2}}^d - p_{i,j}^d}{h/2} = 0. \tag{3.9}$$

Equation (3.8) can then be rewritten as

$$(\sigma_{yy})_{i,j+\frac{1}{2}} = -p_{i,j}^d + \frac{\eta h}{2K_{i,j+\frac{1}{2}}} v_{i,j+\frac{1}{2}}^d. \tag{3.10}$$

The stress component $(\sigma_{xy})_{i+\frac{1}{2},j+\frac{1}{2}}$ is approximated as

$$(\sigma_{xy})_{i+\frac{1}{2},j+\frac{1}{2}} = \eta \left(\frac{u_{i+\frac{1}{2},j+1}^s - u_{i+\frac{1}{2},j+\frac{1}{2}}^s}{h/2} + \frac{v_{i+1,j+\frac{1}{2}}^s - v_{i,j+\frac{1}{2}}^s}{h} \right). \tag{3.11}$$

Here, the horizontal component of the velocity at the interface $u_{i+\frac{1}{2},j+\frac{1}{2}}^s$ is derived using the BJS condition (2.12)(c). Therefore,

$$\left(\frac{\sqrt{K_{i,j+\frac{1}{2}}}}{\alpha_{BJ}}\right)^{-1} u_{i+\frac{1}{2},j+\frac{1}{2}}^s - \eta \left(\frac{u_{i+\frac{1}{2},j+1}^s - u_{i+\frac{1}{2},j+\frac{1}{2}}^s}{h/2} + \frac{v_{i+1,j+\frac{1}{2}}^s - v_{i,j+\frac{1}{2}}^s}{h} \right) = 0. \tag{3.12}$$

Using (3.11) and (3.12), one obtains:

$$(\sigma_{xy})_{i+\frac{1}{2},j+\frac{1}{2}} = \frac{2\eta \tilde{m}_{i,j+\frac{1}{2}}}{h} u_{i+\frac{1}{2},j+1}^s + \eta \tilde{m}_{i,j+\frac{1}{2}} \frac{v_{i+1,j+\frac{1}{2}}^s - v_{i,j+\frac{1}{2}}^s}{h}, \tag{3.13}$$

where $\tilde{m}_{i,j+\frac{1}{2}} = \left(1 - \frac{2\eta\sqrt{K_{i,j+\frac{1}{2}}}}{h\alpha_B + 2\eta\sqrt{K_{i,j+\frac{1}{2}}}}\right)$. The final component $(\sigma_{xy})_{i-\frac{1}{2},j+\frac{1}{2}}$ is computed in a similar manner. Using all stress components derived above, we rewrite (3.6) as

$$\begin{aligned}
 & -\frac{2\eta\tilde{m}_{i,j+\frac{1}{2}}}{h^2}(u_{i+\frac{1}{2},j+1}^s - u_{i-\frac{1}{2},j+1}^s) - \frac{\eta\tilde{m}_{i,j+\frac{1}{2}}}{h^2}(v_{i+1,j+\frac{1}{2}}^s - 2v_{i,j+\frac{1}{2}}^s + v_{i-1,j+\frac{1}{2}}^s) \\
 & + \frac{2}{h}(p_{i,j+1}^s - p_{i,j}^d) - \frac{4\eta}{h^2}(v_{i,j+\frac{3}{2}}^s - v_{i,j+\frac{1}{2}}^s) + \frac{v_{i,j+\frac{1}{2}}^s}{K_{i,j+\frac{1}{2}}} = (f_2^s)_{i,j+\frac{1}{2}}.
 \end{aligned} \tag{3.14}$$

3.2. Discretization of transport equation

The numerical scheme to approximate the advection–diffusion part of the transport equation depends on the type of problem. In the case of contaminant transport in the Darcy–Stokes system, convection is the dominant cause for the flow movement [33,34]. An important requirement for designing an efficient MLMC estimator is using numerical schemes that are stable on coarse meshes. For example, for time-stepping an implicit method is favourable due to unconditional stability with respect to time-step size. Similarly, an upwind based discretization can give stable solutions on relatively coarse grids compared to the central differencing based methods. We briefly discuss the implementation details of the spatio-temporal discretization of the transport equation.

In each of the control volumes $\mathcal{D}_h^3(i, j)$ (see Fig. 3), the integral formulation of (2.14)(a) takes the form

$$\frac{\partial}{\partial t} \int_{\mathcal{D}_h^3(i,j)} \phi c \, d\mathbf{x} + \int_{\mathcal{D}_h^3(i,j)} \nabla \cdot (\mathbf{c}\mathbf{u} - \mathbb{D}\nabla c) \, d\mathbf{x} = \int_{\mathcal{D}_h^3(i,j)} \phi f^t \, d\mathbf{x}. \tag{3.15}$$

The velocity field \mathbf{u} is assumed to be exact but for the current problem this is derived from the solution of coupled Darcy–Stokes system. Also, for simplicity, we will assume $D_{xy} = D_{yx} = 0$ resulting in a diagonal dispersion tensor \mathbb{D} .

3.3. Spatial discretization

Using the Gauss divergence theorem, the second integral in (3.15) is reformulated as a boundary integral for the boundary $\partial\mathcal{D}_h^3(i, j) = \bigcup_{k=1}^4 \partial\mathcal{D}_h^3(i, j, k)$, such that

$$\int_{\mathcal{D}_h^3(i,j)} \nabla \cdot (\mathbf{c}\mathbf{u} - \mathbb{D}\nabla c) \, d\mathbf{x} = \sum_{k=1}^4 \int_{\partial\mathcal{D}_h^3(i,j,k)} (\mathbf{c}\mathbf{u} - \mathbb{D}\nabla c) \cdot \mathbf{n}_k \, dS_k, \tag{3.16}$$

where \mathbf{n}_k denotes the unit normal vector to the corresponding face. Flow in the Stokes domain is well-behaved and the central differencing scheme to evaluate the fluxes in (3.16) will result in a second-order accurate solution locally on very fine grids. However, in the Darcy domain, a permeability field with very small correlation lengths may exhibit a highly fluctuating flow and may give rise to non-physical oscillations in the solution even on finer meshes. The quality of solution will depend on the cell Péclet number is defined as

$$Pe_h(i, j) := \min \left\{ \left| \frac{u}{D_{xx}} \right|_{i-\frac{1}{2},j}, \left| \frac{u}{D_{xx}} \right|_{i+\frac{1}{2},j}, \left| \frac{v}{D_{yy}} \right|_{i,j-\frac{1}{2}}, \left| \frac{v}{D_{yy}} \right|_{i,j+\frac{1}{2}} \right\} h. \tag{3.17}$$

To avoid oscillations it is desired to have $Pe_h(i, j) \leq 2$ in each cell. For the current problem, the velocities u_h, v_h are random variables and it is difficult to bound the cell Péclet number, especially on coarser grids. A first-order Upwind Differencing Scheme (UDS) for the approximation of the convective fluxes seems a convenient choice. The main disadvantage of the first-order upwind scheme is however excessive numerical diffusion that smears out sharp features in the solution. Therefore, to prevent this we use a Quadratic Upwind Interpolation for Convective Kinematics (QUICK) [36] scheme for the approximation of the convective fluxes. For the diffusion part we will use the central differencing scheme. This scheme is less diffusive compared to the first-order solves and is also highly stable. Furthermore, the QUICK scheme is formally third-order accurate in space thus can capture the additional smoothness in the Stokes region. The QUICK scheme is however, not monotone, which means that flux limiters will be employed for problems with sharp gradients. Many versions of the flux-limited QUICK method exist in the literature [37,38] mostly based on the deferred correction (or sometimes referred to as defect correction) framework where the first-order scheme is improved using a higher-order correction. In this article, we consider the implicit version for the same. The flux for the face “BC” (see Fig. 5) is defined as follows

$$\langle F_{i+\frac{1}{2},j} \rangle := \langle F_{i+\frac{1}{2},j} \rangle^{UDS} + \Psi(r_{i+\frac{1}{2},j}) \left(\langle F_{i+\frac{1}{2},j} \rangle^{QUICK} - \langle F_{i+\frac{1}{2},j} \rangle^{UDS} \right), \tag{3.18}$$

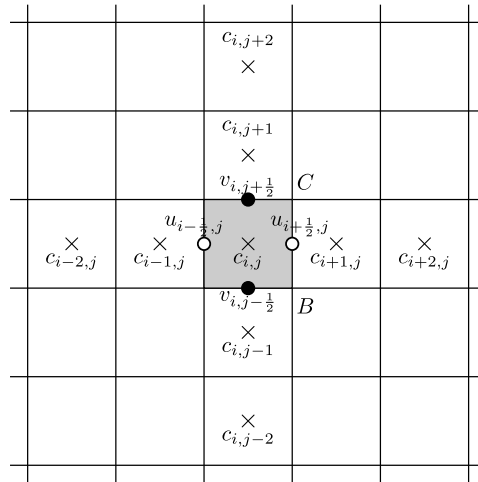


Fig. 5. Stencil for the QUICK scheme for purely convection problem.

where $\langle F_{i+\frac{1}{2},j} \rangle^{UDS}$ and $\langle F_{i+\frac{1}{2},j} \rangle^{QUICK}$ are face-averaged convective fluxes computed using the UDS and QUICK scheme, respectively, and are computed as

$$\langle F_{i+\frac{1}{2},j} \rangle^{UDS} = \begin{cases} u_{i+\frac{1}{2},j}(c_{i,j}), & \text{if } u_{i+\frac{1}{2},j} \geq 0, \\ u_{i+\frac{1}{2},j}(c_{i+1,j}), & \text{if } u_{i+\frac{1}{2},j} \leq 0, \end{cases} \tag{3.19}$$

and

$$\langle F_{i+\frac{1}{2},j} \rangle^{QUICK} = \begin{cases} u_{i+\frac{1}{2},j} \left(\frac{3}{8}c_{i+1,j} + \frac{3}{4}c_{i,j} - \frac{1}{8}c_{i-1,j} \right), & \text{if } u_{i+\frac{1}{2},j} \geq 0, \\ u_{i+\frac{1}{2},j} \left(\frac{3}{8}c_{i,j} + \frac{3}{4}c_{i+1,j} - \frac{1}{8}c_{i+2,j} \right), & \text{if } u_{i+\frac{1}{2},j} \leq 0. \end{cases} \tag{3.20}$$

The limiter function is $\Psi(r_{i+\frac{1}{2},j})$, where $r_{i+\frac{1}{2},j}$ is the measure of the local smoothness. This is computed by taking the ratio of successive gradients along the stream-wise direction to decide the weights for the first-order and QUICK schemes, where

$$\tilde{r}_{i+\frac{1}{2},j} = \begin{cases} \frac{c_{i+1,j} - c_{i,j}}{c_{i,j} - c_{i-1,j}}, & \text{if } u_{i+\frac{1}{2},j} \geq 0, \\ \frac{c_{i+2,j} - c_{i+1,j}}{c_{i+1,j} - c_{i,j}}, & \text{if } u_{i+\frac{1}{2},j} \leq 0, \end{cases} \tag{3.21}$$

and $r_{i+\frac{1}{2},j} := \min(\tilde{r}_{i+\frac{1}{2},j}, \tilde{r}_{i+\frac{1}{2},j}^{-1})$. The inverse of \tilde{r} is used to find large denominators in (3.21). The choice of the limiter function Ψ is somewhat arbitrary. We have numerically tested a number of classical TVD-limiter functions (for e.g., see [39]) and they performed reasonably well. In the current application, where we will encounter discontinuous boundary and initial conditions, the Koren limiter [40] with $\Psi(r) := \max[0, \min(2r, (2+r)/3, 2)]$ was able to reconstruct the propagating sharp fronts. The gradients can be computed using the solution from the previous time-step and then, modify the fluxes using (3.18) to obtain the improved solution. Note that the staggered arrangement of variables for the Darcy–Stokes flow is ideal here as the velocities at the cell faces are directly obtained from the solution of discrete Darcy–Stokes flow. The inflow and outflow boundary conditions (2.16)(a)–(b) are utilized to compute fluxes for control volumes that lie along the domain boundary.

3.4. Temporal discretization

For the temporal discretizations, many well established methods are available in the literature. In our application, flexibility in the time-step size is needed and therefore we prefer an implicit time-stepping scheme. We can formally rewrite the transport equation (2.14)(a) in a semi-discrete form as

$$\frac{\partial c_h}{\partial t} + \mathcal{L}_h c_h = f_h \quad \text{in } \mathcal{D}_h \times \mathcal{T} \times \Omega, \tag{3.22}$$

with \mathcal{L}_h denoting the spatial discretization matrix. For ease of presentation, we have omitted the porosity $\phi(\mathbf{x})$ term in (3.22) by scaling \mathcal{L}_h by $\phi_{i,j}$. With the time-step size $\Delta t > 0$ and for any given integer $m \geq 1$, we define the temporal

grid points as $t^m := m\Delta t$ and will denote by c_h^m the approximation of $c(\cdot, t^m)$. To solve (3.22), we propose to use an Alternating Direction Implicit based solution method that decomposes the semi-discrete form (3.22) into two linear systems each requiring inversion of a tridiagonal matrix for UDS or a pentadiagonal matrix for QUICK scheme. The ADI solver uses two half time-steps. First, an intermediate quantity $c^{m-1/2}$ is generated by performing an implicit Euler time-stepping along the x -coordinate and an explicit Euler along y -direction. Using this intermediate quantity, the reverse is performed for the next half step. Motivated by the classical Peaceman–Rachford approach [41], one can express ADI method as

$$\begin{cases} (I_h + \frac{\Delta t}{2} \mathcal{L}_h^x) c_h^{m-1/2} = \frac{\Delta t}{2} f_h^{m-1/2} + (I_h - \frac{\Delta t}{2} \mathcal{L}_h^y) c_h^{m-1}, & \text{(a)} \\ (I_h + \frac{\Delta t}{2} \mathcal{L}_h^y) c_h^m = \frac{\Delta t}{2} f_h^m + (I_h - \frac{\Delta t}{2} \mathcal{L}_h^x) c_h^{m-1/2}, & \text{(b)} \end{cases} \quad (3.23)$$

where I_h is an identity matrix, \mathcal{L}_h^y and \mathcal{L}_h^x are tridiagonal (pentadiagonal) matrices derived from discretizations along x -coordinate and y -coordinate, respectively. There are many standard algorithms available for inverting tridiagonal or pentadiagonal matrices, for example, the Thomas algorithm can efficiently invert a tridiagonal matrix with only 8 floating point operations per unknown. Furthermore, this ADI scheme is second-order accurate in time which can be easily verified using Taylor’s expansion.

4. Multigrid solver for Darcy–Stokes problem with lognormal diffusion

Multigrid methods are generally recognized as fast efficient solution methods for a large class of linear and non-linear problems. We propose a *monolithic multigrid algorithm* to solve the coupled Darcy–Stokes system with a heterogeneous stochastic permeability field. This multigrid algorithm stems from the work done in [28], where a fixed permeability value was used throughout the Darcy domain. We extend the method to the cases with highly heterogeneous permeability fields. Solving the Darcy flow with highly fluctuating and discontinuous diffusion parameter is quite challenging. In the context of geometric multigrid methods, basically two approaches exist, based on either the cell-centred or the vertex-centred location of unknowns, see [19,25,42] and references therein. The vertex-centred approach requires transfer operators that are dependent on the diffusion parameter. Such multigrid implementations may be expensive as these transfer operators need to be modified on all grid levels. Due to this, the cell-centred version is somewhat more beneficial as it is possible to achieve a decent multigrid convergence using constant transfer operators [19,42]. We will demonstrate that the cell-centred approach can be extended to the coupled Darcy–Stokes system.

We would also like to point out that another common alternative for solving the linear systems derived from a coupled model is the Domain Decomposition Method (DDM). In a DDM, one splits the main boundary value problem into smaller boundary value subproblems. Typically, the boundary information between subproblems is exchanged during every iteration until the converged solution is reached. Contrary to this, monolithic solution approaches treat the coupled system as a single problem. An important aspect of this approach is that the coupling variables between the subproblems are treated simultaneously, thus after every iteration, the three fields $(u, v, p)^T$ are updated throughout the domain. The monolithic approach has proven to be very efficient when the subproblems are strongly coupled [43,44].

4.1. Uzawa smoother for saddle-point system

Discretization of each subproblem, the mixed form of Darcy flow and the Stokes equations, yields a saddle-point system. This saddle-point structure can be maintained for the coupled Darcy–Stokes system by ordering the velocity unknowns together for both the subproblems followed by the pressure unknowns. This results in the following linear system

$$\begin{bmatrix} \mathcal{A}_h & \mathcal{B}_h^T \\ \mathcal{B}_h & 0 \end{bmatrix} \begin{bmatrix} \mathbf{u}_h \\ p_h \end{bmatrix} = \begin{bmatrix} \mathbf{g}_h \\ f_h \end{bmatrix}, \quad (4.1)$$

where $\mathbf{u}_h = (\mathbf{u}_h^d, \mathbf{u}_h^s)^T$, $p_h = (p_h^d, p_h^s)^T$, $\mathbf{g}_h = (\mathbf{0}, \mathbf{f}_h^s)^T$ and $f_h = (f^d, 0)^T$. For both subproblems, \mathcal{B}_h^T and \mathcal{B}_h represent the discrete gradient and minus discrete divergence operators, respectively, and \mathcal{A}_h is the discrete representation of either the Laplacian operator $-\eta\Delta$ for the Stokes equations, or $\eta K_h^{-1} I_h$ for the Darcy equation.

The multigrid method proposed here uses a special class of relaxation methods called the Uzawa smoothers [24]. The Uzawa smoother is basically an equation-wise, decoupled smoother where the velocity components in the Darcy and Stokes domains are first updated, after which the pressure field is updated. In the following, we provide the details of this smoother.

Consider a splitting of the matrix coefficients of the saddle point system (4.1) as

$$\begin{bmatrix} \mathcal{A}_h & \mathcal{B}_h^T \\ \mathcal{B}_h & 0 \end{bmatrix} = \begin{bmatrix} \mathcal{M}_h & 0 \\ \mathcal{B}_h & -\zeta_h^{-1} I_h \end{bmatrix} - \begin{bmatrix} \mathcal{M}_h - \mathcal{A}_h & -\mathcal{B}_h^T \\ 0 & -\zeta_h^{-1} I_h \end{bmatrix}, \quad (4.2)$$

where \mathcal{M}_h is a smoother for the operator \mathcal{A}_h and ζ_h is a positive parameter. For a given approximation of the solution $(\mathbf{u}_h^{k-1}, p_h^{k-1})^T$, the relaxed approximation $(\mathbf{u}_h^k, p_h^k)^T$, can be defined in the following way:

$$\begin{bmatrix} \mathcal{M}_h & 0 \\ \mathcal{B}_h & -\zeta_h^{-1} I_h \end{bmatrix} \begin{bmatrix} \mathbf{u}_h^k \\ p_h^k \end{bmatrix} = \begin{bmatrix} \mathcal{M}_h - \mathcal{A}_h & -\mathcal{B}_h^T \\ 0 & -\zeta_h^{-1} I_h \end{bmatrix} \begin{bmatrix} \mathbf{u}_h^{k-1} \\ p_h^{k-1} \end{bmatrix} + \begin{bmatrix} \mathbf{g}_h \\ f_h \end{bmatrix}. \quad (4.3)$$

In terms of velocity and pressure variables, the relaxation step can be written down as:

$$\begin{cases} \mathbf{u}_h^k = \mathbf{u}_h^{k-1} + \mathcal{M}_h^{-1} (\mathbf{g}_h - \mathcal{A}_h \mathbf{u}_h^{k-1} - \mathcal{B}_h^T p_h^{k-1}), & \text{(a)} \\ p_h^k = p_h^{k-1} + \zeta_h (\mathcal{B}_h \mathbf{u}_h^k - f_h), & \text{(b)} \end{cases} \tag{4.4}$$

respectively. In [27,28], the choice of \mathcal{M}_h was based on a symmetric Gauss–Seidel iteration for (4.4)(a) that gives

$$\mathcal{M}_h = (\Lambda_h + L_h) \Lambda_h^{-1} (\Lambda_h + U_h), \tag{4.5}$$

where Λ_h , L_h and U_h are the diagonal, strictly lower and strictly upper parts of \mathcal{A}_h , respectively. A symmetric Gauss–Seidel iteration comprises a forward and a backward sweep for velocity in the entire domain.

Next, for smoothing of the pressure variable, a Richardson iteration (4.4)(b) with appropriate relaxation parameters is applied. For any control volume $\mathcal{D}_h^3(i, j)$, we use the following relaxation parameters

$$\zeta_h(i, j) = \begin{cases} \eta & \text{in } \mathcal{D}^s, \\ \frac{\eta h^2}{5\bar{K}_h(i, j)} & \text{in } \mathcal{D}^d, \end{cases} \tag{4.6}$$

where η is the fluid viscosity and $\bar{K}_h(i, j)$ is derived by applying a *half-weighting* (HW) operator

$$\bar{K}_h(i, j) = \frac{1}{8} [4K_{i,j} + K_{i-1,j} + K_{i+1,j} + K_{i,j-1} + K_{i,j+1}] \tag{4.7}$$

This is a generalization of the optimal relaxation parameters that were derived using a Local Fourier Analysis (LFA) in [27, 28]. Usually, an LFA is performed on the multigrid components obtained by freezing the coefficient field. In the case of a variable coefficient field, ζ_h needs to be modified locally. In our experience, using the weighted average (4.7) resulted in a robust convergence rate that we will later demonstrate numerically.

Remark 4.1. We would like to mention that obtaining an analytic bound on the smoothing factor of the Uzawa smoother for the stochastic Darcy–Stokes flow is rather involved. We leave this for our future work.

4.2. Multigrid algorithm with grid partitioning

To cover realistic problems, we propose a multi-block multigrid method which is based on the grid partitioning technique [25]. A multi-block algorithm requires communication between the Stokes and Darcy domains during the multigrid iteration. In our case, the stencils do not use variables located more than one cell away from the current control volume, therefore padding the Stokes block with one extra row to store the variables from the Darcy block is sufficient to achieve the data exchange.

The multigrid hierarchy is based on uniform coarsening, i.e. the cell-width is doubled in each coarsening step. A two-block two-grid method with variable permeability field can be described by the following steps:

1. Fine grid pre-smoothing. Relax the velocities in two blocks using the symmetric Gauss–Seidel smoother. The vertical velocity from the Stokes block along the interface is transferred to the Darcy block after which the pressure is updated by performing Richardson iterations with optimal relaxation parameters given by (4.6). The updated Darcy pressure unknowns are then transferred to the Stokes overlap control volumes.

2. Defect computation and restriction. The defect (residual) is computed for each variable. Next, the residuals from the Darcy block are transferred to the Stokes overlap region. We use fixed restriction operators to transfer residuals to the coarse grid. For the velocities, we use a six-point stencil whereas for pressures a four-point restriction is applied. These stencils are denoted by

$$[I_h^{2h}]^u = \frac{1}{8} \begin{bmatrix} 1 & 2 & 1 \\ & * & \\ 1 & 2 & 1 \end{bmatrix}_h^{2h}, \quad [I_h^{2h}]^v = \frac{1}{8} \begin{bmatrix} 1 & & 1 \\ 2 & * & 2 \\ 1 & & 1 \end{bmatrix}_h^{2h}, \quad [I_h^{2h}]^p = \frac{1}{4} \begin{bmatrix} 1 & & 1 \\ & * & \\ 1 & & 1 \end{bmatrix}_h^{2h}, \tag{4.8}$$

respectively.

3. Coarse grid relaxation. The coarse grid defect equation is solved exactly by a direct method, e.g., Gaussian elimination. Alternately, one can perform multiple smoothing iterations as described in step 1 to get an approximate solution on the coarse grid. We use direct discretization to obtain the discrete system on the coarse grids and the coarse grid representation of permeability K_{2h} is also obtained by using the restriction operator $[I_h^{2h}]^p$ given in (4.8). Finally, vertical Stokes velocity unknowns at the interface are transferred to the Darcy block.

4. Coarse grid correction. The error is interpolated to the fine grid, where it is used correct the fine grid solution. The prolongation operators are chosen as the adjoint of the restriction operators.

5. Fine grid post-smoothing. The interpolation process introduces new errors in the solution that need to be eliminated. For this, several post-smoothing steps are performed with the same procedure as used in pre-smoothing given in step 1.

Although the above multi-block algorithm resembles the domain decomposition method, the main difference is that the communication between the two domains is performed on *all multigrid levels* and not just on the finest and coarsest grids. This makes multi-block method typically more efficient than the traditional DDM.

The two-block method can be easily extended to more than two blocks that may have more interfaces without any deterioration of the multigrid convergence rate. Furthermore, this two-grid routine is easily extended to V- and W-cycling strategies. Numerical tests showed that for this problem, the W(2, 2)-cycle exhibits a fast convergence rate regardless of the roughness of the coefficient fields.

5. Single-level Monte Carlo method

In this section, we describe steps to formulate a single-level MC estimator along with the bounds for the errors incurred from the finite volume and statistical approximations of the coupled problem. These error estimates will be needed for the development and analysis of the MLMC estimator.

5.1. MC estimator

We are interested in estimating the statistical moments of the random concentration field at some final time T . We denote by $c(\cdot, t^m)$ the solution at a discrete time $t^m = T$ where $m = T/\Delta t$. For simplicity, we consider a spatio-temporal grid with $h = \Delta t$ and denote by $c_h^m(\omega_i)$ the approximation of the concentration field with the random input $\omega_i \in \Omega$. The *single-level MC estimator* \mathcal{E}_h^{MC} employs N independent, identically distributed (i.i.d.) realizations of the solution field $\{c_h^m(\omega_i)\}_{i=1}^N$ and is defined as

$$\mathbb{E}[c(\cdot, t^m)] \approx \mathcal{E}_N^{MC}[c_h^m] := \frac{1}{N} \sum_{i=1}^N c_h^m(\omega_i). \tag{5.1}$$

To compute a sample, $c_h^m(\omega_i)$, we proceed as follows: First, a random permeability field and other associated stochastic parameters are generated. Then, using the multiblock multigrid method, we solve the Darcy–Stokes flow for the velocity field. With these velocities, the linear system for the transport model is derived. Next, the ADI solver is employed for time-stepping and after m time steps, we obtain a sample of c_h^m . These samples are then averaged. The higher moments of the QoI can be computed in a similar fashion from the sample ensemble.

The proposed MC estimator contains two additive errors due to the finite sampling and the FV approximation and one may separate these two errors as:

$$\left\| \mathbb{E}[c(\cdot, t^m)] - \mathcal{E}_N^{MC}[c_h^m] \right\|_{L^2(\Omega; \mathcal{D})} \leq \left\| \mathbb{E}[c(\cdot, t^m)] - \mathcal{E}_N^{MC}[c(\cdot, t^m)] \right\|_{L^2(\Omega; \mathcal{D})} + \left\| \mathcal{E}_N^{MC}[c(\cdot, t^m)] - \mathcal{E}_N^{MC}[c_h^m] \right\|_{L^2(\Omega; \mathcal{D})} \tag{5.2}$$

respectively. Here, the function space $L^2(\Omega; \mathcal{D})$ corresponds to the space of square integrable measurable functions $f : \Omega \rightarrow L^2(\mathcal{D})$ for the given complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ defined earlier. These spaces are equipped with the norm

$$\|f(\mathbf{x}, t, \omega)\|_{L^2(\Omega; \mathcal{D})} := \mathbb{E} \left[\|f(\mathbf{x}, t, \omega)\|_{L^2(\mathcal{D})}^2 \right]^{\frac{1}{2}} = \left(\int_{\Omega} \|f(\mathbf{x}, t, \omega)\|_{L^2(\mathcal{D})}^2 d\mathbb{P}(\omega) \right)^{\frac{1}{2}}. \tag{5.3}$$

The above error norm is regarded to be the root mean squared error (RMSE) and is widely used in the literature to quantify the error associated with Monte Carlo estimators [45,46].

5.1.1. Bound for discretization error

The spatio-temporal discretization introduces a bias in the MC estimator. Therefore, it is imperative to quantify this bias as a function of the mesh parameters. The approximations of the subproblems (2.1)(a)–(2.16)(b) may exhibit different grid convergence rates depending on the regularity of the solution variables and the discretization order of the numerical scheme employed. The error in the transport equations depends on the error in the approximation of the velocity in \mathcal{D} . Thus, the second term in the right-hand side of (5.2) can be formally expressed as

$$\left\| \mathcal{E}_N^{MC}[c(\cdot, t^m) - c_h^m] \right\|_{L^2(\Omega; \mathcal{D})} \leq C_0 (h^{\alpha_d} + h^{\alpha_s} + h^{\alpha_{t1}} + h^{\alpha_{t2}}). \tag{5.4}$$

Here, the constant $C_0 > 0$ is independent of h and convergence rates $\alpha_d, \alpha_s, \alpha_{t1}, \alpha_{t2}$ are also positive. The rates α_s and α_d are used to express the errors from the approximation of the Stokes and Darcy equations, respectively. The errors due to the spatial and temporal discretizations of the transport equations are quantified using the rates α_{t1} and α_{t2} , respectively. Clearly, the asymptotic convergence will be dictated by the term which decays at the slowest rate, i.e.,

$$\left\| \mathcal{E}_N^{MC}[c(\cdot, t^m) - c_h^m] \right\|_{L^2(\Omega; \mathcal{D})} \leq C_1 h^\alpha, \tag{5.5}$$

where $\alpha = \min\{\alpha_d, \alpha_s, \alpha_{t1}, \alpha_{t2}\}$ and $C_1 > 0$. For the current problem, the dominant error comes from the approximation of the Darcy subproblem that further depends on the spatial regularity of the random permeability field.

One can use the relative L_2 -error measure $\|\mathbb{E}[c_h^m - c_{2h}^m]\|_{L^2(\Omega; \mathcal{D})}$ as an indicator for the FV error convergence. The relative error can simply be associated with the absolute error using the reverse triangle inequality

$$\begin{aligned} \left| \|\mathbb{E}[c_h^m - c(\cdot, t^m)]\|_{L^2(\Omega; \mathcal{D})} - \|\mathbb{E}[c_{2h}^m - c(\cdot, t^m)]\|_{L^2(\Omega; \mathcal{D})} \right| &\leq \|\mathbb{E}[c_h^m - c(\cdot, t^m) + c(\cdot, t^m) - c_{2h}^m]\|_{L^2(\Omega; \mathcal{D})}, \\ &= \|\mathbb{E}[c_h^m - c_{2h}^m]\|_{L^2(\Omega; \mathcal{D})}. \end{aligned} \tag{5.6}$$

Assuming $\|\mathbb{E}[c_h^m - c(\cdot, t^m)]\|_{L^2(\Omega; \mathcal{D})} = \mathcal{O}(h^\alpha)$ and using (5.6), we have the bound

$$\|\mathbb{E}[c_h^m - c(\cdot, t^m)]\|_{L^2(\Omega; \mathcal{D})} \leq \|\mathbb{E}[c_h^m - c_{2h}^m]\|_{L^2(\Omega; \mathcal{D})} / (2^\alpha - 1). \tag{5.7}$$

Relative errors can be numerically computed on a hierarchy of grids and the value of α can be empirically determined via regression.

5.1.2. Bound for sampling error

Once the FV error corresponding to the grid parameter h is determined, the next task is to reduce the sampling error up to the order of the discretization error. Using (5.3), we can write the mean square error (MSE) as

$$\begin{aligned} &\left\| \mathbb{E}[c(\cdot, t^m)] - \mathcal{E}_N^{MC}[c(\cdot, t^m)] \right\|_{L^2(\Omega; \mathcal{D})}^2 \\ &= \mathbb{E} \left[\left\| \mathbb{E}[c(\cdot, t^m)] - \frac{1}{N} \sum_{i=1}^N c(\cdot, t^m, \omega_i) \right\|_{L^2(\mathcal{D})}^2 \right] \\ &= \mathbb{E} \left[\frac{1}{N^2} \left\| \sum_{i=1}^N \mathbb{E}[c(\cdot, t^m)] - c(\cdot, t^m, \omega_i) \right\|_{L^2(\mathcal{D})}^2 \right] \\ &= \frac{1}{N^2} \mathbb{E} \left[\int_{\mathcal{D}} \left(\sum_{i=1}^N \mathbb{E}[c(\mathbf{x}, t^m)] - c(\mathbf{x}, t^m, \omega_i) \right) \left(\sum_{j=1}^N \mathbb{E}[c(\mathbf{x}, t^m)] - c(\mathbf{x}, t^m, \omega_j) \right) d\mathbf{x} \right]. \end{aligned} \tag{5.8}$$

Due to the fact that the samples are mutually independent, the terms with $i \neq j$ vanish and by linearity of the expectation, we may write

$$\begin{aligned} \left\| \mathbb{E}[c(\cdot, t^m)] - \mathcal{E}_N^{MC}[c(\cdot, t^m)] \right\|_{L^2(\Omega; \mathcal{D})}^2 &= \frac{1}{N^2} \sum_{i=1}^N \mathbb{E} \left[\int_{\mathcal{D}} (\mathbb{E}[c(\mathbf{x}, t^m)] - c(\mathbf{x}, t^m, \omega_i))^2 d\mathbf{x} \right] \\ &= \frac{1}{N^2} \sum_{i=1}^N \|\mathbb{E}[c(\cdot, t^m)] - c(\cdot, t^m)\|_{L^2(\Omega; \mathcal{D})}^2 \\ &= \frac{\|\mathcal{V}[c(\cdot, t^m)]\|_{L^2(\Omega; \mathcal{D})}}{N}, \end{aligned} \tag{5.9}$$

where the term $\|\mathcal{V}[c(\cdot, t^m)]\|_{L^2(\Omega; \mathcal{D})} := \|\mathbb{E}[c(\cdot, t^m)] - c(\cdot, t^m)\|_{L^2(\Omega; \mathcal{D})}^2$ denotes the sample variance and can be numerically estimated. Assuming the square root of the sample variance to be bounded by some positive constant, i.e., $\|\mathcal{V}[c(\cdot, t^m)]\|_{L^2(\Omega; \mathcal{D})}^{1/2} \leq C_2$, the error incurred by the MC estimator (5.2) can be bounded using (5.5) and (5.9) as

$$\left\| \mathbb{E}[c(\cdot, t^m)] - \mathcal{E}_N^{MC}[c_h^m] \right\|_{L^2(\Omega; \mathcal{D})} \lesssim C_1 h^\alpha + N^{-\frac{1}{2}} C_2. \tag{5.10}$$

It is pointed out that for a given h , the proposed MC estimator can at most achieve an accuracy of the order of the discretization bias, i.e., $\mathcal{O}(h^\alpha)$, as $N \rightarrow \infty$. Therefore, the optimal number of MC samples should be obtained by balancing the sampling error with the discretization error.

For a given mesh size h , number of MC samples N and final time T , it is straightforward to compute the cost of the MC estimator. As the proposed multigrid solver for the Darcy–Stokes equation and the ADI solver for the transport equation are both asymptotically optimal (i.e. computational cost is proportional to the number of unknowns), their combined cost can be expressed as $\mathcal{O}(h^{-2})$. Therefore, the cost of time-stepping with steps $m = T/h$ ($h = \Delta t$), will be

$$\mathcal{W}_h^m = \mathcal{O}(mh^{-2}) = \mathcal{O}(h^{-3}). \tag{5.11}$$

For the sake of generality, we assume that the computational cost per sample is $\mathcal{O}(h^{-\gamma})$, where $\gamma \geq 3$. To balance the sampling error and the spatio-temporal discretization errors in (5.10), we need $N = \mathcal{O}(h^{-2\alpha})$. Therefore, the work done to compute N samples for the estimator $\mathcal{E}_N^{MC}[c_h^m]$ can be expressed as

$$\mathcal{W}_{h,N}^{MC} = \mathcal{O}(Nh^{-\gamma}) = \mathcal{O}(h^{-(2\alpha+\gamma)}). \tag{5.12}$$

The accuracy versus asymptotic work can be expressed as

$$\left\| \mathbb{E}[c(\cdot, t^m)] - \mathcal{E}_N^{MC}[c_h^m] \right\|_{L^2(\Omega; \mathcal{D})} \lesssim \left(\mathcal{W}_{h,N}^{MC} \right)^{-\alpha/(2\alpha+\gamma)}. \tag{5.13}$$

Note that the cost of the MC estimator depends on the efficiency of the solver and the mean grid convergence rate of the QoI, represented by the parameters γ and α , respectively. Improvements in the asymptotic cost can be obtained by using an optimal solver (for e.g. a full multigrid solver) or by employing a high-order discretization scheme, see e.g., [19]. For subsurface problems, the convergence rate of the discretization error is usually slow due to the low spatial regularity. As a consequence, samples should be computed on a relatively fine mesh to obtain a reliable estimate of the statistics of the QoI making the plain MC estimator very expensive.

6. Multilevel Monte Carlo method

6.1. MLMC estimator

In the following, we discuss the key elements of the MLMC scheme as explained in [20,21,45,47]. First, we construct a hierarchy of spatial grids \mathcal{D}_{h_ℓ} for $\ell = 0, 1, \dots, L$, with uniform refinement such that cell-width is related as

$$h_\ell = 2^{-\ell}h_0, \quad \text{for } h_0 > 0, \tag{6.1}$$

and the temporal grid hierarchy $\{\mathcal{T}_\ell\}_{\ell=0}^L$ with time-step size Δt_ℓ . For simplicity, we set $\Delta t_\ell = h_\ell$. Now, using the linearity of the expectation operator, we can build the estimate on the hierarchy of spatio-temporal meshes as

$$\mathbb{E}[c_{h_L}^m] = \mathbb{E}[c_{h_0}^m] + \sum_{\ell=1}^L \mathbb{E}[c_{h_\ell}^m - c_{h_{\ell-1}}^m]. \tag{6.2}$$

Each of the expectations in (6.2) can be approximated using the standard MC estimator defined in (5.1) as

$$\begin{aligned} \mathcal{E}_L^{ML}[c_{h_L}^m] &:= \sum_{\ell=0}^L \mathcal{E}_{N_\ell}^{MC}[c_{h_\ell}^m - c_{h_{\ell-1}}^m] \\ &= \sum_{\ell=0}^L \left(\frac{1}{N_\ell} \sum_{i=1}^{N_\ell} c_{h_\ell}^m(\omega_i) - c_{h_{\ell-1}}^m(\omega_i) \right), \end{aligned} \tag{6.3}$$

with $N_\ell \in \mathbb{N}$ denoting the level-dependent number of MC samples. For notational convenience, we set $c_{h_{-1}}^m = 0$. The MC estimators at different levels are mutually independent. The samples $c_{h_\ell}^m(\omega_i) - c_{h_{\ell-1}}^m(\omega_i)$ at any level ℓ correspond to the difference between the concentration fields at time t^m on two discretization levels $h_\ell, \Delta t_\ell$ and $h_{\ell-1}, \Delta t_{\ell-1}$ both generated using the same random input ω_i .

Similar to the MC estimator, the MLMC estimator $\mathcal{E}_L^{ML}[c_{h_L}^m]$ is obtained by two approximations

$$\mathbb{E}[c(\cdot, t^m)] \approx \mathbb{E}[c_{h_L}^m] \approx \mathcal{E}_L^{ML}[c_{h_L}^m]. \tag{6.4}$$

Therefore, the mean squared error (MSE) in $\mathcal{E}_L^{ML}[c_{h_L}^m]$ can be represented as

$$\left\| \mathbb{E}[c(\cdot, t^m)] - \mathcal{E}_L^{ML}[c_{h_L}^m] \right\|_{L^2(\Omega; \mathcal{D})}^2 \leq \left\| \mathbb{E}[c(\cdot, t^m)] - \mathbb{E}[c_{h_L}^m] \right\|_{L^2(\Omega; \mathcal{D})}^2 + \left\| \mathbb{E}[c_{h_L}^m] - \mathcal{E}_L^{ML}[c_{h_L}^m] \right\|_{L^2(\Omega; \mathcal{D})}^2. \tag{6.5}$$

The first term corresponds to the discretization error and is bounded in a similar manner as in (5.5), that is,

$$\left\| \mathbb{E}[c(\cdot, t^m)] - \mathbb{E}[c_{h_L}^m] \right\|_{L^2(\Omega; \mathcal{D})} \leq C_1 h_L^\alpha = \mathcal{O}(2^{-\alpha L}). \tag{6.6}$$

Using (6.2) and (6.3), we can bound the second error term in (6.5) as

$$\begin{aligned}
 \left\| \mathbb{E}[c_{h_L}^m] - \mathcal{E}_L^{ML}[c_{h_L}^m] \right\|_{L^2(\Omega; \mathcal{D})}^2 &= \mathbb{E} \left[\left\| \sum_{\ell=0}^L \mathbb{E}[c_{h_\ell}^m - c_{h_{\ell-1}}^m] - \mathcal{E}_{N_\ell}^{MC}[c_{h_\ell}^m - c_{h_{\ell-1}}^m] \right\|_{L^2(\mathcal{D})}^2 \right] \\
 &= \mathbb{E} \sum_{\ell=0}^L \left[\left\| \mathbb{E}[c_{h_\ell}^m - c_{h_{\ell-1}}^m] - \mathcal{E}_{N_\ell}^{MC}[c_{h_\ell}^m - c_{h_{\ell-1}}^m] \right\|_{L^2(\mathcal{D})}^2 \right] \\
 &= \sum_{\ell=0}^L \mathbb{E} \left[\left\| \mathbb{E}[c_{h_\ell}^m - c_{h_{\ell-1}}^m] - \mathcal{E}_{N_\ell}^{MC}[c_{h_\ell}^m - c_{h_{\ell-1}}^m] \right\|_{L^2(\mathcal{D})}^2 \right] \tag{6.7} \\
 &= \sum_{\ell=0}^L \left[\frac{1}{N_\ell^2} \sum_{i=1}^{N_\ell} \mathbb{E} \left[\left\| \mathbb{E}[c_{h_\ell}^m - c_{h_{\ell-1}}^m] - [c_{h_\ell}^m(\omega_i) - c_{h_{\ell-1}}^m(\omega_i)] \right\|_{L^2(\mathcal{D})}^2 \right] \right] \\
 &= \sum_{\ell=0}^L \frac{\mathcal{V}_\ell}{N_\ell},
 \end{aligned}$$

where $\mathcal{V}_\ell := \left\| \mathcal{V}[c_{h_\ell}^m - c_{h_{\ell-1}}^m] \right\|_{L^2(\Omega; \mathcal{D})} = \left\| \mathbb{E}[c_{h_\ell}^m - c_{h_{\ell-1}}^m] - [c_{h_\ell}^m - c_{h_{\ell-1}}^m] \right\|_{L^2(\Omega; \mathcal{D})}^2$ corresponds to the level-dependent variance. Finally, the error in the MLMC estimator can be written in a compact form as

$$\left\| \mathbb{E}[c(\cdot, t^m)] - \mathcal{E}_L^{ML}[c_{h_L}^m] \right\|_{L^2(\Omega; \mathcal{D})}^2 \lesssim (C_1 h_L^\alpha)^2 + \sum_{\ell=0}^L \frac{\mathcal{V}_\ell}{N_\ell}. \tag{6.8}$$

The above error represents the error in the MLMC estimator in mean square sense. We further assume that the level-dependent variance decays with a rate β as

$$\mathcal{V}_\ell = \mathcal{O}(h_\ell^\beta) = \mathcal{O}(2^{-\beta\ell}), \quad \ell = 1, 2, \dots, L, \quad \beta > 0. \tag{6.9}$$

The value of β can also be derived a-priori in a similar fashion as α . For any level ℓ , the optimal number of samples is obtained by balancing the level-dependent sampling error \mathcal{V}_ℓ/N_ℓ with the discretization bias term $\mathcal{O}(h_\ell^{2\alpha})$. A convenient way of achieving this is by fixing the number of samples N_L as a free parameter, see e.g. [45–47]. Using this, we obtain the sequence of MLMC samples

$$N_\ell = \lceil N_L 2^{\beta(L-\ell)} \rceil. \tag{6.10}$$

Ideally, N_L should be chosen such that the sampling error on the finest level also satisfies $\mathcal{V}_L/N_L = \mathcal{O}(h_L^{2\alpha})$. In practice, the value of N_L should be very small, i.e. $\mathcal{O}(10)$ and can be set heuristically. Also, note that the sample variance \mathcal{V}_0 on the coarsest level does not depend on the parameter β , therefore \mathcal{V}_0/N_0 will reduce to discretization error only asymptotically (for large L). Since the cost per sample on level ℓ is $\mathcal{W}_{h_\ell}^m = \mathcal{O}(h_\ell^{-\gamma}) = \mathcal{O}(2^{\gamma\ell})$, the total cost of the MLMC estimator can be expressed as

$$\mathcal{W}_L^{ML} = \sum_{\ell=0}^L N_\ell \mathcal{W}_{h_\ell}^m = \mathcal{O} \left(\sum_{\ell=0}^L 2^{(\gamma-\beta)\ell} \right). \tag{6.11}$$

This leads to the following three scenarios: 1) For $\beta > \gamma$, the dominant cost comes from the coarsest level; 2) for $\beta = \gamma$ all levels equally contribute to the total cost and 3) for $\beta < \gamma$, the dominant cost comes from the finest level. The authors in [20,21,45,47] have estimated the asymptotic work versus accuracy of the MLMC estimator that can be stated as follows

$$\left\| \mathbb{E}[c(\cdot, t^m)] - \mathcal{E}_L^{ML}[c_{h_L}^m] \right\|_{L^2(\Omega; \mathcal{D})} \lesssim \begin{cases} (\mathcal{W}_L^{ML})^{-\frac{1}{2}} & \text{if } \beta > \gamma, \\ (\mathcal{W}_L^{ML})^{-\frac{1}{2}} \log(\mathcal{W}_L^{ML})^{\frac{1}{2}} & \text{if } \beta = \gamma, \\ (\mathcal{W}_L^{ML})^{\frac{-\alpha}{2\alpha+\gamma-\beta}} & \text{if } \beta < \gamma. \end{cases} \tag{6.12}$$

Note that for all these cases, the MLMC estimator has a better asymptotic cost than the standard Monte Carlo method $(\mathcal{W}_{h,N}^{MC})^{-\alpha/(2\alpha+\gamma)}$. Further, if $\beta < \gamma$ and $\beta = 2\alpha$, we obtain an optimal MLMC estimator in the sense that its accuracy versus cost has the same order as the deterministic version of the problem. Here, parameters α and β depend on the numerical schemes used to approximate the subproblems as well as the smoothness of the quantity of interest. Also, these rates can sometimes be improved by using high-order schemes and may help in reducing the overall asymptotic cost of the MLMC estimator. The rate γ however increases with the dimension of the problem and when using an optimal solver it is roughly equal to the spatio-temporal dimension.

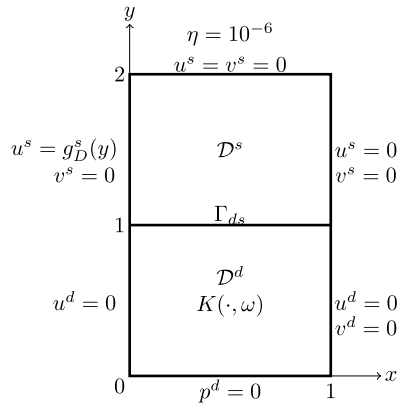


Fig. 6. Geometry of the 2-block problem with no-slip interface condition. Subdivision of the domain into a free-flow domain \mathcal{D}^s and a porous flow domain \mathcal{D}^d , by the interface Γ_{ds} .

Table 1
Different combinations of the Matérn parameters $\Phi = (\nu_c, \lambda_c, \sigma_c^2)$ with increasing complexity from left to right.

Φ_1	Φ_2	Φ_3	Φ_4
(1.5, 0.3, 1)	(0.5, 0.3, 1)	(1.5, 0.1, 3)	(0.5, 0.1, 3)

7. Numerical experiments

We present numerical experiments for the multigrid solver and the MLMC method separately. First, we study the accuracy of the FV scheme and the convergence and robustness of the solver with respect to the Matérn parameters using two test problems: 1) A 2-block case with the no-slip interface condition and simple boundary conditions for the Stokes and the Darcy subproblems, and 2) A realistic 4-block problem with the BJS interface condition and discontinuous boundary conditions. Here, we will also consider a non-isotropic covariance model to generate randomly layered permeability fields. Secondly, we will compute the statistical solution of the transport in Darcy–Stokes flow using the MLMC method. Furthermore, we compare the cost and accuracy of the standard MC method with the MLMC method.

7.1. Multigrid convergence

7.1.1. No-slip interface condition on 2-blocks

For the first test case, the computational domain is taken to be $\mathcal{D} \cup \partial\mathcal{D} = [0, 1] \times [0, 2]$, where $\mathcal{D}^d = (0, 1) \times (0, 1)$ and $\mathcal{D}^s = (0, 1) \times (1, 2)$ and the interface $\Gamma_{ds} = [0, 1] \times \{1\}$. For simplicity, we consider the source terms in both domains to be zero. Fig. 6 explains the geometry of this problem along with the boundary conditions. The inflow in the Stokes domain is described by a parabolic function $g_D^s = (y - 1)(2 - y)$ and tangential flow at the top and the right boundary is applied. In the Darcy region, we set tangential flow at the left and right boundaries and set $p = 0$ at the bottom boundary to mimic the effect of a gravity force. With this configuration of boundary conditions, we will have outflow only through the bottom Darcy boundary, i.e., $\Gamma_D^d \in [0, 1] \times \{0\}$.

The lognormal random field K is generated on a cell-centred grid and the permeability values at the face-centres are obtained using the harmonic mean of the two adjacent cell-centres, see [48] for details. Factorization of the covariance matrix and sampling of the Gaussian random field is done using the *circulant embedding technique*, outlined in [49].

In Table 1, four different Matérn parameters of increasing order of complexity in terms of solvability of the linear system are presented. Also, we show one random realization of the Gaussian random field generated using each of the four Matérn parameters in Fig. 7. These examples clearly show that for the considered Matérn parameters, the simulated permeability fields can exhibit large fluctuations and a variation of more than 10 orders of magnitude can be encountered for the cases with $\sigma_c^2 = 3$.

As the performance of multigrid will depend on the particular realization of the random permeability field, we take a statistical approach to study the convergence and robustness of the multigrid algorithm. For a fixed Matérn parameter set Φ and h , we generate N_{MG} samples of the random field and record the multigrid convergence factors $\{q_i\}_{i=1}^{N_{MG}}$ for each solve. The samples of convergence factor q_i for the i -th realization of the random field are defined as

$$q_i := \left\{ \frac{\|res^{k_i}\|_\infty}{\|res^0\|_\infty} \right\}^{\frac{1}{k_i}}, \quad \text{for } i = 1, 2, \dots, N_{MG}, \tag{7.1}$$

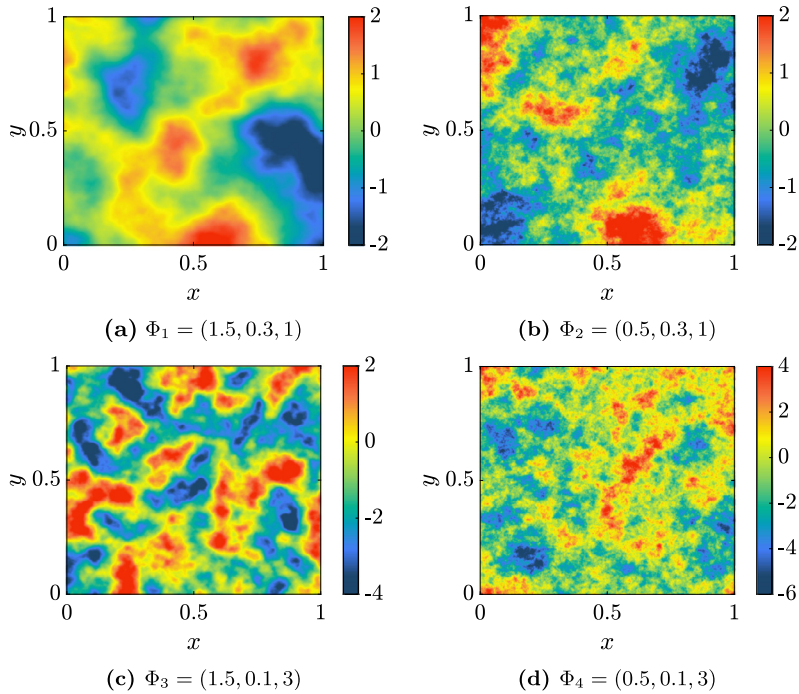


Fig. 7. Examples of Gaussian random field ($\log K$) generated using parameter sets Φ_1 – Φ_4 . The colour bars indicate the variation in the order of magnitude.

where $\|res^0\|_\infty$ is the infinity-norm of the residual from an initial solution and $\|res^{k_i}\|_\infty$ is the residual after k_i iterations of the multigrid cycle. We use these to compute the mean multigrid convergence as

$$\langle q \rangle = \frac{1}{N_{MG}} \sum_{i=1}^{N_{MG}} q_i. \tag{7.2}$$

We compute the mean and variance of the convergence factors for the V- and W-cycles with different combinations of the number of pre- and post-smoothing steps. Similarly, the average number of multigrid iterations $\langle k \rangle$ is computed. We choose random numbers as the initial solution, and the stopping criterion is reached when the k_i -th iteration satisfies

$$\frac{\|res^{k_i}\|_\infty}{\|res^0\|_\infty} \leq \epsilon_{MG}, \tag{7.3}$$

with the tolerance ϵ_{MG} set to 10^{-10} . The relaxation parameters for the Richardson iteration ζ are taken as described in (4.6). In the Stokes domain $\zeta_h = \eta = 10^{-6}$ and in the Darcy domain, we take spatially varying $\zeta_h(i, j) = \eta h^2 / 5 \bar{K}_h(i, j)$, where i, j are cell indices.

In Tables 2–5, we list the average number of multigrid iterations $\langle k \rangle$ (rounded-up to the nearest integer) for different combinations of W-cycles required to achieve the stopping criterion on different grids for $N_{MG} = 100$. The V-cycle variants performed poorly for all Matérn parameters and also divergence was observed in case of Φ_4 , hence they are not included in the tables. Also, we omit results with other values of the fluid viscosity as it did not influence the convergence factors of these cycles.

As expected, we see an increase in the number of iterations with respect to the complexity of the random field. For all Φ , we see an improvement in the convergence rate with respect to size of the grid, stabilizing after a certain h . The stabilization is encountered on a relatively coarse grid ($1/h = 32$) for Φ_1 and on a relatively fine grid ($1/h = 128$) for Φ_4 . This happens because finer grids are better able to resolve the small-scale correlation structures in the permeability field.

We use the four-point restriction operator $[I_h^{2h}]^p$ from (4.8) to get the coarse grid representation of the permeability field by which we derive the discretization operator. For a random field with small correlation length λ_c and large variance σ_c^2 that produces highly fluctuating fields, the coarse grid representation may be very different from the finest grid version. This discrepancy further increases as we go to coarser levels. Therefore, it makes more sense to use a W-cycle which by definition performs more smoothing steps on coarse and intermediate grids where multigrid convergence may not be optimal.

It is worthwhile to mention that for Φ_1 and Φ_2 , we are able to achieve the same efficiency as was observed for constant K in [28], which required around 14 iterations to reduce residual by 10 orders of magnitude independent of the mesh size. This is typically regarded as h -independent convergence of a multigrid algorithm. For Φ_3 and Φ_4 , this efficiency is achieved on relatively finer grids.

Table 2
Average multigrid iterations $\langle k \rangle$ for Φ_1 .

h	W(1, 1)	W(1, 2)	W(2, 2)
1/8	36	21	16
1/16	34	20	15
1/32	35	20	15
1/64	34	19	14
1/128	33	19	14
1/256	33	19	14

Table 3
Average multigrid iterations $\langle k \rangle$ for Φ_2 .

h	W(1, 1)	W(1, 2)	W(2, 2)
1/8	35	23	17
1/16	34	21	16
1/32	35	21	16
1/64	35	21	15
1/128	35	20	15
1/256	35	20	15

Table 4
Average multigrid iterations $\langle k \rangle$ for Φ_3 .

h	W(1, 1)	W(1, 2)	W(2, 2)
1/8	60	41	32
1/16	66	43	34
1/32	53	34	27
1/64	42	25	20
1/128	40	22	16
1/256	37	21	16

Table 5
Average multigrid iterations $\langle k \rangle$ for Φ_4 .

h	W(1, 1)	W(1, 2)	W(2, 2)
1/8	59	41	33
1/16	66	42	32
1/32	57	37	28
1/64	46	30	23
1/128	40	25	18
1/256	39	23	17

We wish to emphasize that the choice of sampling mesh for K has a significant influence on the multigrid convergence. To study this, we also generate a permeability field on a staggered mesh such that the cell faces have different values than the cell-centres. This introduces more variability within a control volume and slows down the multigrid convergence. The mean convergence rate $\langle q \rangle$ along with the standard deviation for W(2, 2)-cycles is presented in Fig. 8 for cell-centred (left) and staggered (right) grid based sampling, respectively. Multigrid convergence for a cell-centred based sampling performs well on very coarse grids and the average convergence factors on finer grids are in the range 0.1–0.2. Whereas for the staggered grid based sampling, we see a deterioration in the multigrid convergence, especially on coarser grids for parameters Φ_3 and Φ_4 . As the MLMC estimator requires a relatively large number of samples on coarser grids, a cell-centred based sampling grid for K may give rise to a significant gain over the staggered sampling approach. We also remark that the FV error convergence rate does not deteriorate when a cell-centred approach is adapted which will be demonstrated later.

The most efficient variant of the W-cycle is determined on the basis of the average CPU times required to reach the stopping criterion. In Fig. 9 we plot the CPU times for different W-cycles versus the grid size for the “hardest” parameter set Φ_4 . The W(2, 2)-cycle was the fastest to achieve the stopping criterion. A similar trend is also observed for the other Matérn parameters. We would like to point out that the average CPU time for one W(2, 2)-cycle is about 0.65 seconds for a 256×512 grid. For many engineering applications, however, residual reduction of about 3 orders of magnitude may be sufficient and can be achieved in less than 8 iterations using the W(2, 2)-cycle. Also, we observe the $\mathcal{O}(h^{-2})$ scaling of the computational cost from $h = 1/128$ onwards for all the three variants of W-cycles. As the cost is proportional to the number of unknowns, we can conclude that we have an optimal multigrid solver for the coupled Darcy–Stokes problem, on staggered meshes. We will use the W(2, 2)-cycle to solve the Darcy–Stokes system for the MC samples.

Next, we numerically demonstrate the FV error convergence in the mean for the coupled Darcy–Stokes approximation. In Figs. 10–12, the relative errors in the horizontal and vertical components of the velocity for \mathcal{D}^s , \mathcal{D}^d and \mathcal{D} , respectively,

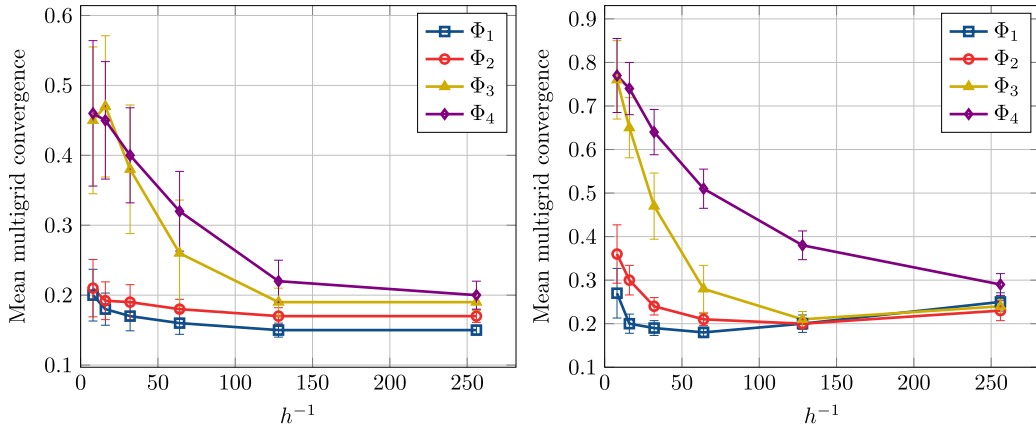


Fig. 8. Mean convergence rate (q) along with the standard deviation for W(2,2)-cycle when K is sampled on a cell-centred grid (left) and staggered grid (right).

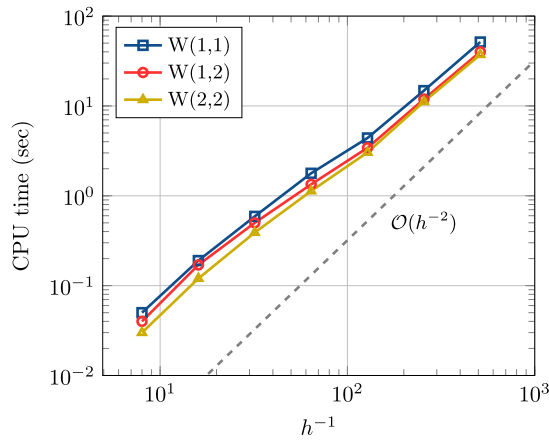


Fig. 9. Comparison of average CPU time for W-cycle variants for Φ_4 . Similar trend is observed for other Matérn parameters.

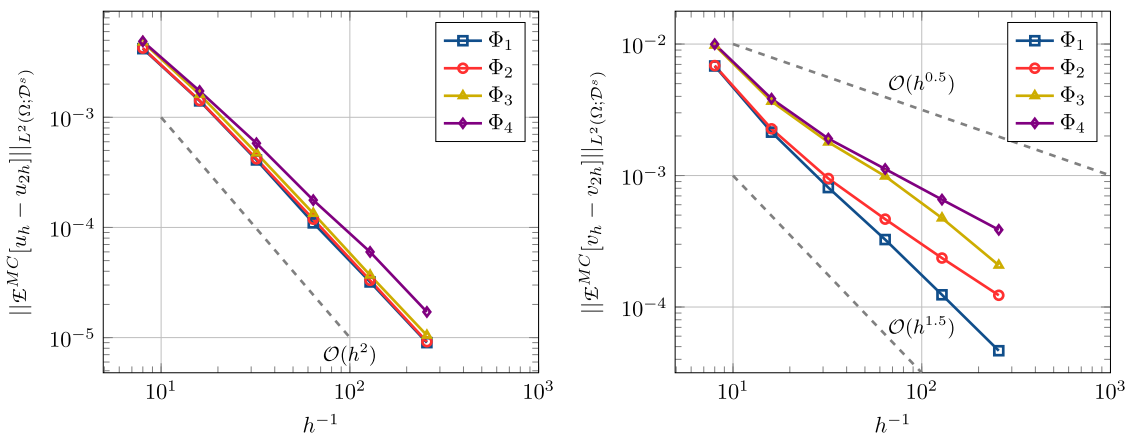


Fig. 10. Mean error convergence of horizontal (left) and vertical (right) velocity components for parameter sets Φ_1 – Φ_4 in the Stokes region \mathcal{D}^s .

are listed. These relative errors are computed using the Monte Carlo method with a sufficient number of samples such that the sampling error is less than the FV bias. In the Stokes region, a convergence close to second-order is observed for the horizontal velocity component, however for the vertical velocity component we see a dependence on the smoothness parameter of the Matérn function ν_c . This is due to the fact that for the no-slip boundary condition $u^s = 0$ is imposed along the interface without any influence of the permeability field. Whereas the vertical velocity component at the interface

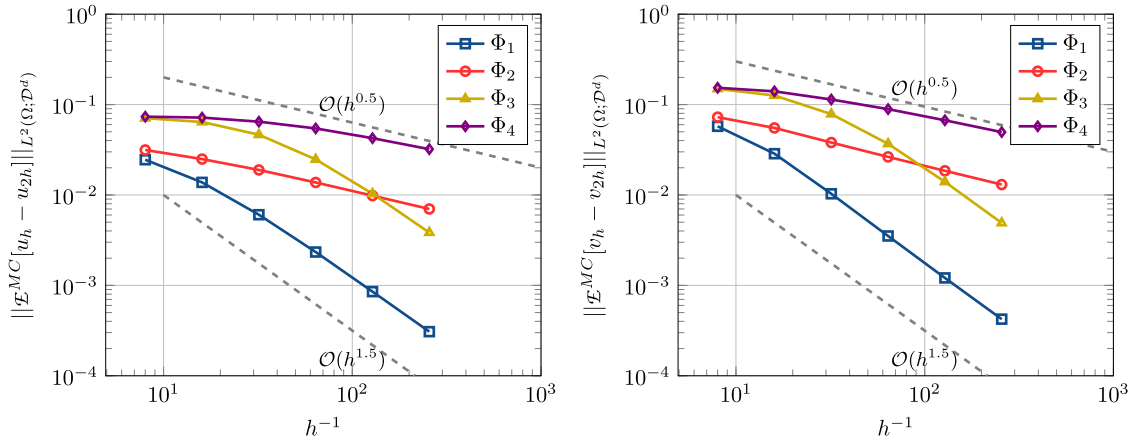


Fig. 11. Mean error convergence of horizontal (left) and vertical (right) velocity components for parameter sets Φ_1 – Φ_4 in the Darcy region \mathcal{D}^d .

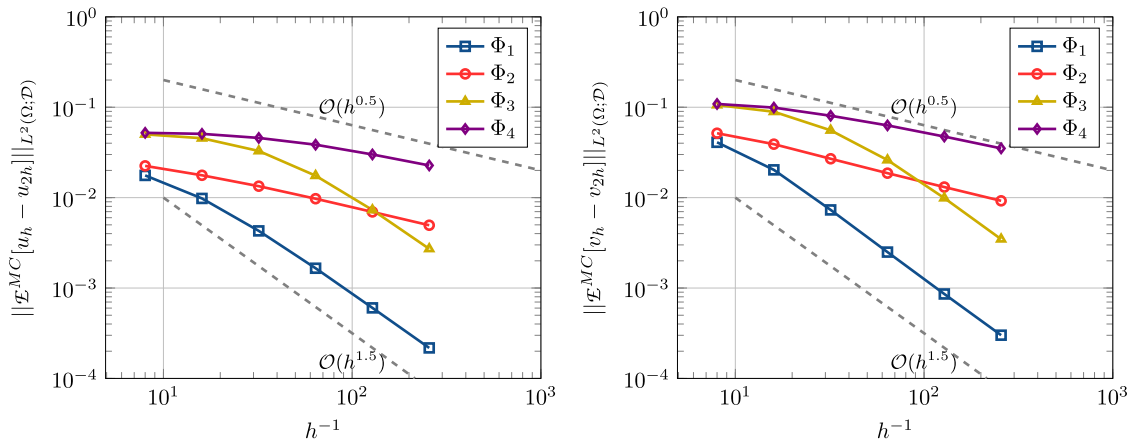


Fig. 12. Mean error convergence of horizontal (left) and vertical (right) velocity components for parameter sets Φ_1 – Φ_4 in the whole domain \mathcal{D} .

is based on the discretization of the Darcy flow (3.9) related to the random permeability field. Also note that for the BJS interface condition, the convergence of u^s will again depend on the permeability field, due to (3.12).

For the Darcy domain, the FV error decays with a rate equal to ν_c for both velocity components. The global error also shows a similar trend as the dominating error is due to the approximation of the Darcy subproblem. In general, the FV error increases with an increase in the variance parameter σ_c^2 .

7.1.2. Beaver–Joseph–Saffman interface condition on 4-blocks

Next, we test our method for a 4-block geometry with the BJS boundary condition at the Darcy–Stokes interface. This type of geometry is sometimes utilized to model a cross-flow filtration process [50]. The geometry and boundary constraints are depicted in Fig. 13. Here, \mathcal{D}^s represents a channel divided into 3-blocks and \mathcal{D}^d represents a porous media forming an interface with the second Stokes block \mathcal{D}_2^s . We impose a constant inflow velocity $u^s = 0.1$ at the left inlet boundary of the channel. On the right, there is an exit of length $1/6$ with zero normal stress $\sigma_{xx} = \sigma_{xy} = 0$.

We also analyze the performance of the multigrid algorithm with grid partitioning (see Section 4.2) when extended to more than two blocks. For the 4-block test case, the algorithm can be described in the following way. Given an initial approximation, we first update all variables on the first subgrid \mathcal{D}_1^s . Next, the updated boundary data from \mathcal{D}_1^s is communicated to the second block \mathcal{D}_2^s and then all the interior points are updated. Similarly, the variables in blocks \mathcal{D}_3^s and \mathcal{D}^d are updated. For each block, it is common to extend the computational mesh by one cell length along the boundary between the two neighbouring blocks in order to store the updated data. Note that this grid partitioning based multigrid is also highly parallelizable as all the blocks can be simultaneously relaxed on separate processors.

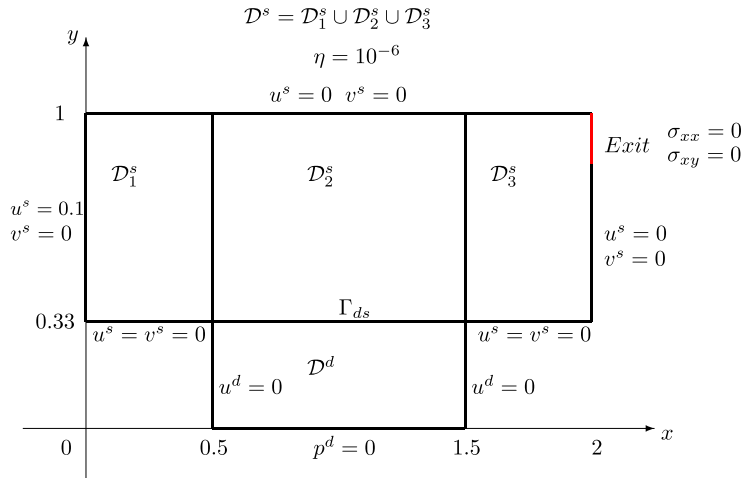


Fig. 13. Geometry of the 4-block problem with BJS interface condition. Subdivision of the domain \mathcal{D} into a free-flow subregion \mathcal{D}^s and a porous subdomain \mathcal{D}^d by an internal interface Γ_{ds} .

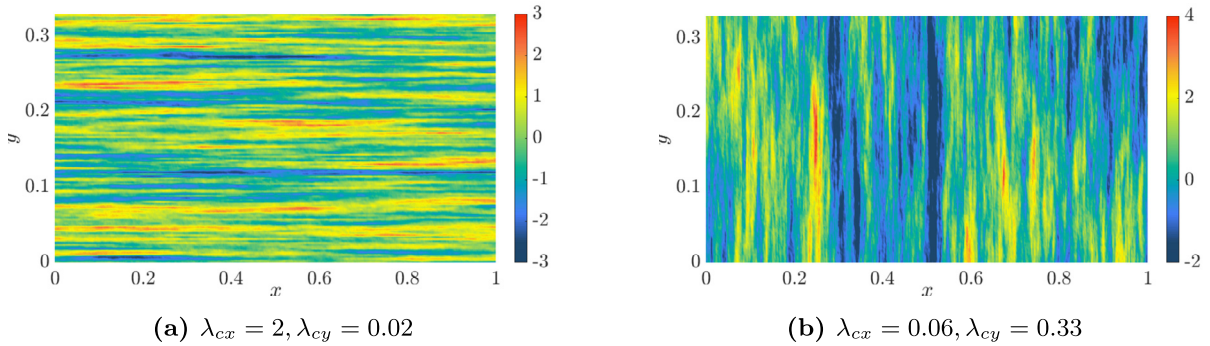


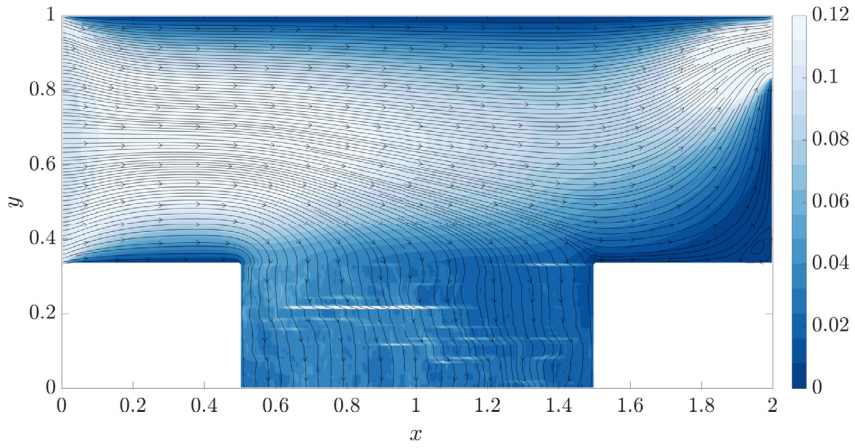
Fig. 14. Examples of Gaussian random field ($\log K$) with (a) horizontal layers and (b) vertical layers. For both the cases $\nu_c = 0.5$ and variance $\sigma_c^2 = 1$.

To simulate layered porous media, we use a stationary non-isotropic Matérn model given by

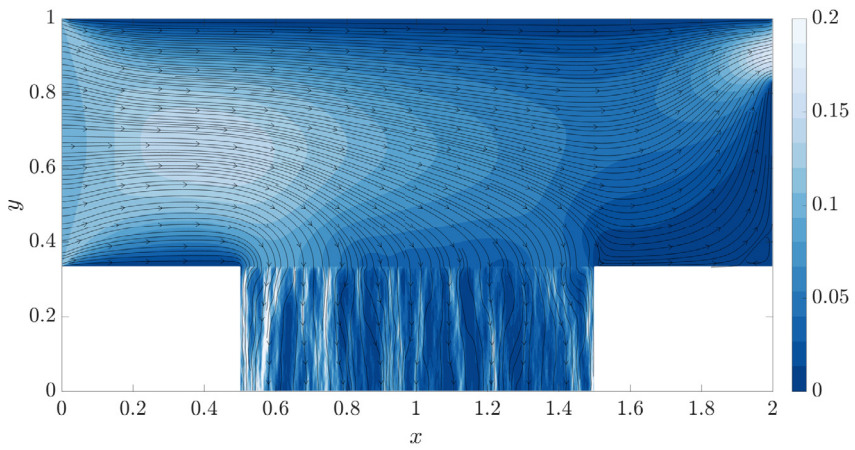
$$\begin{cases} C_\Phi(\mathbf{x}_1, \mathbf{x}_2) = \sigma_c^2 \frac{2^{1-\nu_c}}{\Gamma(\nu_c)} (2\sqrt{\nu_c \tilde{r}})^{\nu_c} K_{\nu_c}(2\sqrt{\nu_c \tilde{r}}), \\ \tilde{r} = \sqrt{\frac{(x_1 - x_2)^2}{\lambda_{cx}^2} + \frac{(y_1 - y_2)^2}{\lambda_{cy}^2}} \quad \text{with } \mathbf{x}_1 = (x_1, y_1), \mathbf{x}_2 = (x_2, y_2), \end{cases} \quad (7.4)$$

where λ_{cx} and λ_{cy} are correlation lengths along x - and y -coordinates, respectively. An example of a horizontally and vertically layered Gaussian field is depicted in Fig. 14. Also in Fig. 15, we present the streamlines and magnitude ($\sqrt{u^2 + v^2}$) of the velocity fields generated from the horizontally (Fig. 15a) and vertically (Fig. 15b) layered permeabilities. These solutions are based on a staggered mesh with $h = 1/384$ such that the number of cells on different blocks are $\mathcal{D}_1^s = 192 \times 256$, $\mathcal{D}_2^s = 384 \times 256$, $\mathcal{D}_3^s = 192 \times 256$ and $\mathcal{D}^d = 384 \times 128$. The effect of layered permeability is clearly visible in both figures resulting in randomly layered high and low mobility zones. The flow intensity through the right exit depends on the particular realization of the random fields.

In Fig. 16, we show the mean \pm standard deviation multigrid convergence for three parameter sets: isotropic, small vertical and small horizontal correlations. For all these cases, the smoothness is taken as $\nu_c = 0.5$ and variance $\sigma_c^2 = 1$. For the isotropic case with $\lambda_{cx} = 0.1$, $\lambda_{cy} = 0.1$, a similar convergence behaviour as observed for the no-slip interface condition test case is seen. For the vertically layered case with $\lambda_{cx} = 0.06$, $\lambda_{cy} = 0.33$, fast convergence rates are achieved. A deterioration in the multigrid convergence for horizontally layered permeability for $\lambda_{cx} = 2$, $\lambda_{cy} = 0.02$ is also observed. This is due to very high contrast between $x - y$ correlation lengths ($\lambda_{cx}/\lambda_{cy} = 100$). The mean convergence rates improves with grid refinement for all these hyper-parameters. Multigrid with semi-coarsening or line-wise smoothers may further improve the convergence [25].



(a) $\lambda_{cx} = 2, \lambda_{cy} = 0.02$



(b) $\lambda_{cx} = 0.06, \lambda_{cy} = 0.33$

Fig. 15. Flow field generated for (a) horizontally and (b) vertically layered porous media. These flows are computed on a staggered mesh with $h = 1/384$.

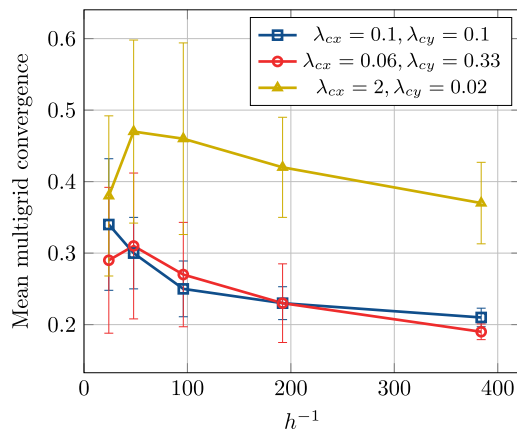


Fig. 16. Mean multigrid convergence (q) along with the standard deviation for W(2, 2)-cycle.

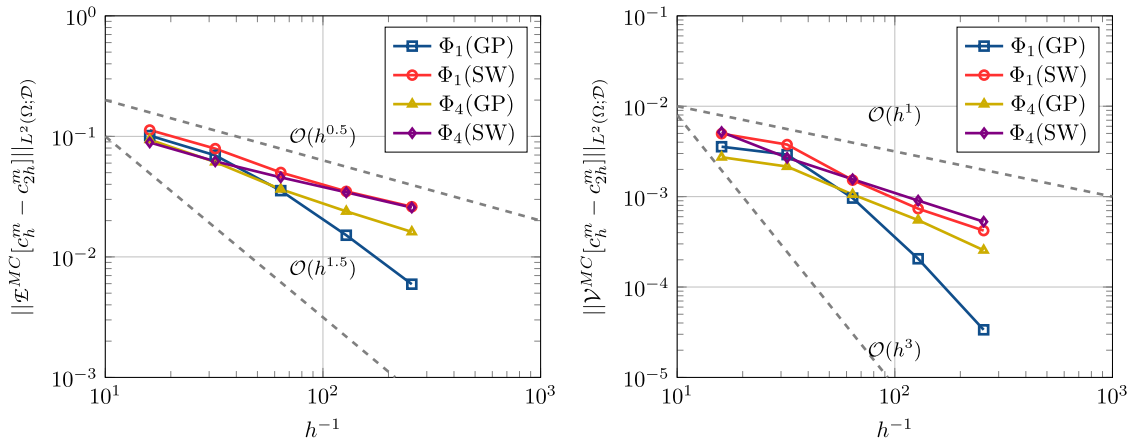


Fig. 17. The convergence of the FV bias and the level-dependent variance for different combinations of Matérn parameter and inflow boundary conditions.

7.2. MLMC convergence

So far, we have benchmarked the performance of the multigrid solver and accuracy of the FV scheme for the random Darcy–Stokes flow. Now, we use the MLMC estimator to compute moments of the stochastic transport in the Darcy–Stokes system. Note that one can also apply the MLMC estimator together with the multigrid solver to efficiently compute the statistical moments of the random flow field.

We consider the transport in the same 2-block problem with the no-slip condition (Section 7.1.1). The source term is taken to be zero, i.e., $f^t \equiv 0$. To investigate the convergence of the proposed spatio-temporal discretization for the fully coupled system, we use two deterministic inflow boundary conditions, based on a Gaussian plume (GP) and a discontinuous square-wave (SW), defined as

$$c_{in}(0, y, t) = \exp\left(-\frac{|y-1.5|^2}{\lambda^2}\right), \quad \text{for } \lambda = 0.1, y \in [1, 2], t \geq 0, \tag{7.5}$$

$$c_{in}(0, y, t) = 1, \quad \text{for } |y - 1.5| \leq 1/8, t \geq 0, \tag{7.6}$$

respectively. The initial concentration c_0 is obtained from c_{in} at $t = 0$. Uncertainty in the inflow condition or initial concentration can be easily incorporated within this test case. The dispersion parameter in the Stokes region is taken as $D = 10^{-6}$ and for the Darcy region parameters in (2.15) are $D_L = D_T = 10^{-5}$ with $D^* = 0$ resulting in a diagonal dispersion tensor. We consider the porosity $\phi = 1$ for the Stokes region and $\phi = 0.4$ for the porous medium.

In Fig. 17, we show the convergence of the FV bias (left) along with the decay of the level-dependent variance (right) for four combinations of Matérn parameters and inflow conditions. For the Gaussian plume model (7.5) the bias decays with the rate of $\mathcal{O}(h^{\nu_c})$ depending on the smoothness parameter of the permeability field as the dominant error comes from the discretization of Darcy–Stokes velocity. As the proposed scheme is second-order accurate, thus we get $\mathcal{O}(h^2)$ for $\nu_c \geq 2$. For the discontinuous square-wave function, we observe a rate of $\mathcal{O}(h^{0.5})$ regardless of the smoothness parameter. Further, we observe that the variance decays at twice the rate of the FV bias. The empirical values of the MLMC parameters α and β can be derived from these plots. It is pointed out that for the easiest parameter set Φ_1 (GP), we can get a cubic decay of samples with increasing level and for the hardest case Φ_4 (SW), we can achieve a linear decay only.

To illustrate the behaviour of the contaminant transport problem in the random flow field and compare the first-order UDS and the flux-limited QUICK scheme, we present two examples with the concentration distribution after $t = 6.00$ in Fig. 18 where the velocity fields are generated using the Matérn parameter set Φ_3 (top) and Φ_4 (bottom), respectively with the discontinuous inflow (7.6). The contaminant plume remains compact as long as it remains in the free-flow region but as soon as it reaches the interface, it starts to spread out due to the heterogeneity of the porous medium. The QUICK method is clearly less diffusive than the UDS and the flux limiter is able to remove oscillations near sharp discontinuities. Also, the velocity profile and concentration distribution from Φ_3 are visibly more smooth compared to Φ_4 .

To demonstrate the convergence of the MLMC method we will consider the most difficult case Φ_4 (SW). To measure the accuracy of the MLMC and MC estimators, we use the following L^2 -based relative error measure

$$\varepsilon_{rel} := \frac{\|\mathcal{E}_{ref}[c(\cdot, t^m)] - \mathcal{E}[c_h^m]\|_{L^2(\mathcal{D}_h)}}{\|\mathcal{E}_{ref}[c(\cdot, t^m)]\|_{L^2(\mathcal{D}_h)}}, \tag{7.7}$$

where \mathcal{E} can be replaced with either \mathcal{E}_h^{ML} (MLMC) or \mathcal{E}_N^{MC} (MC). The reference solution is computed on a relatively fine grid. For the current test problem, the reference solution is presented in Fig. 19, showing the mean and variance of the spatial concentration distribution at $t = 6.00$. The samples required for the MLMC estimator were computed using the multigrid

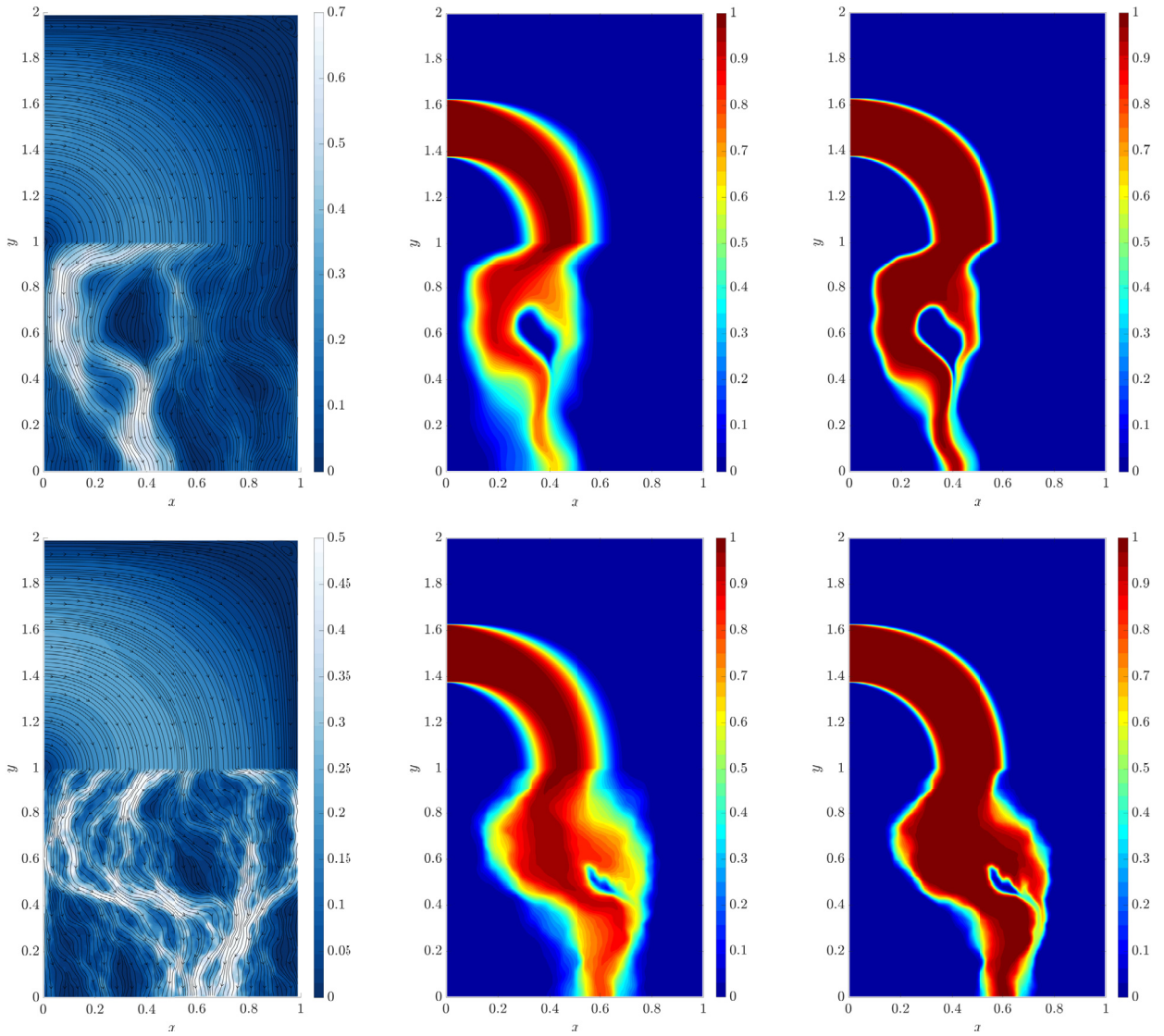


Fig. 18. Transport solution at $t = 6.00$ from UDS (middle) and flux-limited QUICK (right) scheme for same velocity field (left) with $h = \Delta t = 1/128$. The permeability field is generated using Φ_3 (top) and Φ_4 (bottom).

solver with $W(2, 2)$ -cycle to obtain the velocity field and the ADI time-stepping was applied to solve the transport equation. We observe that the plume remains compact in the free-flow region with very insignificant variance. In the Darcy region, the spatial mean and variance are symmetrically distributed. The reference solution is computed using a six level MLMC estimator ($L = 6$) with the coarsest level $h_0 = \Delta t_0 = 1/16$ and the finest level $h_L = \Delta t_L = 1/512$. We use $N_L = 8$ and the number of samples on coarser levels is derived using relation (6.10). For case Φ_4 (SW), we get $\alpha \approx 0.5$ and $\beta \approx 2\alpha \approx 1$ (see Fig. 17) resulting in a sample sequence $N_\ell = N_L 2^{(L-\ell)}$. Recall that with these rates, we end up in the third scenario from (6.12) with $\gamma > \beta$ resulting in an optimal MLMC estimator. For comparison, we also implement the standard MC estimator based on the sampling strategy discussed in Section 5.1.2. For this method, the number of samples are given by $N = \mathcal{O}(h^{-2\alpha}) = \mathcal{O}(h^{-1})$. We chose $N = 8$ for grid $h = 1/16$ and double the number of samples with subsequent refinements, when using the plain MC method.

In Fig. 20 (left), we compare the accuracy of the MLMC and the standard MC estimators. For each h , we repeated the experiment eight times and the average relative error denoted by $\bar{\epsilon}_{rel}$ is plotted. As expected, both the MLMC and MC estimators were able to achieve a similar accuracy converging as $\mathcal{O}(h^{0.5})$. Also, in Fig. 20 (right), we show the cost scaling for both estimators. The CPU time for the MLMC estimator grows with a slower rate than the standard MC estimator and we observe a speed-up of more than an order of magnitude for $h = 1/256$. For reference, we also indicate the theoretical cost (dotted grey lines) for the MC estimator ($\mathcal{W}^{-1/8}$) and the MLMC estimator ($\mathcal{W}^{-1/6}$), when using an optimal solver

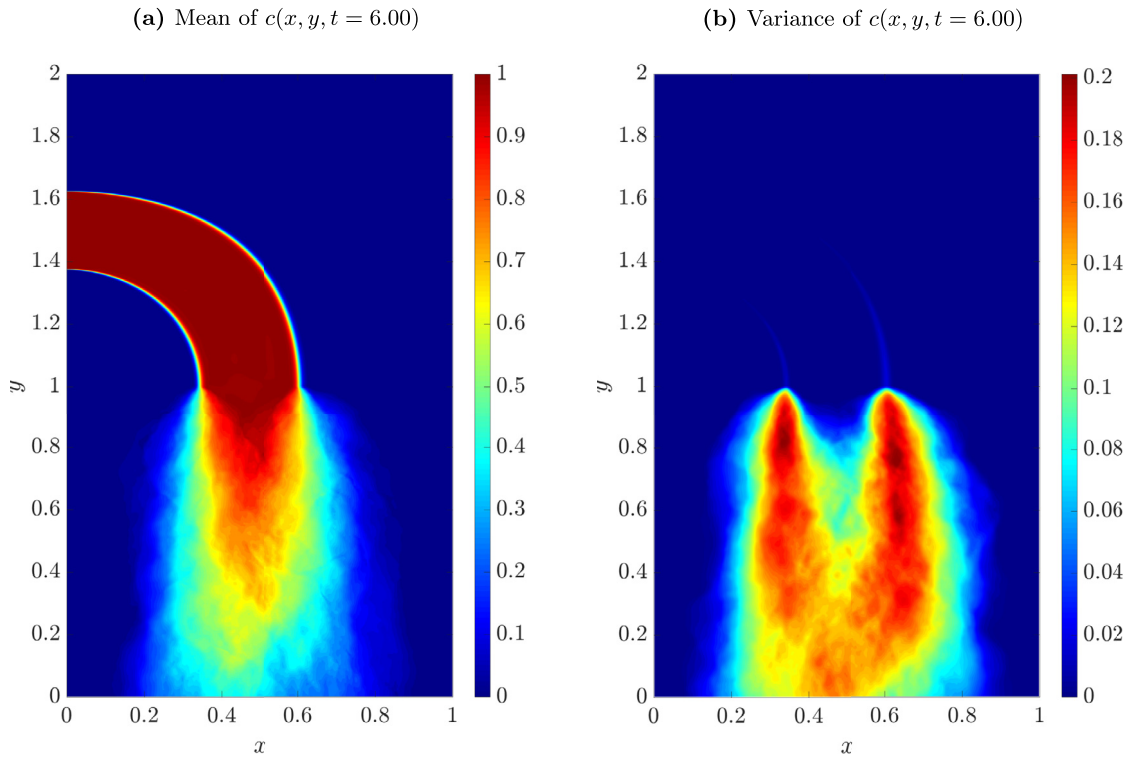


Fig. 19. Statistical solution of uncertain transport problem with parameter Φ_4 at $t = 6.00$ computed using the MLMC estimator. High variance is observed in the Darcy region. Sharp gradients are preserved in the Stokes domain with very insignificant variance.

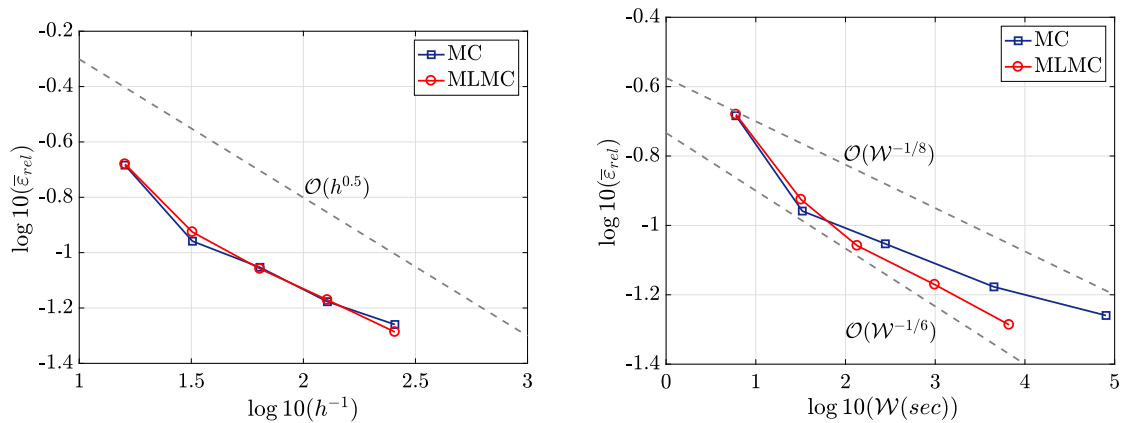


Fig. 20. Comparison of relative error convergence $\bar{\epsilon}_{rel}$ with grid size (left) and computational work \mathcal{W} versus accuracy (right) for MLMC and MC estimators.

(or $\gamma = 3$). Here, we have only presented the results corresponding to the most difficult test case Φ_4 (SW). In general, the asymptotic cost of the MLMC method further improves with the smoothness of the random field as well as with smooth inflow conditions. For example, it is possible to obtain a complexity of about $\mathcal{W}^{-1/2}$ (ignoring the logarithmic term) for case Φ_1 (GP) as $\beta \approx \gamma \approx 3$.

We also remark that the proposed multigrid MLMC method can be extended to parallel architectures. Specifically, one can achieve three degrees of parallelization over samples, levels and the multigrid solver, see [46,51,52] for recent advances. Out of these, the most challenging part is the parallelization over the solver that should be considered only on finer levels when the communication overhead becomes negligible. Moreover, the grid partitioning method explored in this paper can be very effective as each partitioned block can be relaxed on separate processors. We have already demonstrated that it is possible to obtain a satisfactory multigrid convergence using the 4-block example earlier.

8. Summary and conclusion

In this paper, we described an MLMC algorithm for Uncertainty Quantification of advection-dominated transport in a coupled Darcy–Stokes system with uncertain permeability. An important contribution of this paper is the robust multigrid solver that can solve the coupled Darcy–Stokes flow with highly heterogeneous permeability field very efficiently. The new monolithic multigrid method achieves *textbook multigrid convergence* for a wide range of Matérn parameters, thus making it a highly suitable solver for MLMC applications. We combined a flux-limited QUICK scheme with an ADI time-stepping resulting in a second-order accurate spatio-temporal discretization of the stochastic transport equation that can be solved optimally. This implicit version also resulted in a scheme that is stable on very coarse grids and greatly helped in reducing the cost of the MLMC method. Lastly, we showed that for very rough problems, we can attain an asymptotically optimal MLMC estimator that has same computational complexity as its deterministic counterpart.

Acknowledgements

This research is funded by the Shell–NWO/FOM program “Computational Sciences for Energy Research” (CSER) under the research grant No. 14CSER004. Francisco J. Gaspar and C.W. Oosterlee have received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 705402, POROSOS.

References

- [1] G.S. Beavers, D.D. Joseph, Boundary conditions at a naturally permeable wall, *J. Fluid Mech.* 30 (01) (1967) 197–207.
- [2] P.G. Saffman, On the boundary condition at the surface of a porous medium, *Stud. Appl. Math.* 50 (2) (1971) 93–101.
- [3] P. Angot, On the well-posed coupling between free fluid and porous viscous flows, *Appl. Math. Lett.* 24 (6) (2011) 803–810.
- [4] Y. Cao, M. Gunzburger, F. Hua, X. Wang, Coupled Stokes–Darcy model with Beavers–Joseph interface boundary condition, *Commun. Math. Sci.* 8 (1) (2010) 1–25.
- [5] M. Discacciati, E. Miglio, A. Quarteroni, Mathematical and numerical models for coupling surface and groundwater flows, *Appl. Numer. Math.* 43 (1–2) (2002) 57–74.
- [6] H.P. Langtangen, K.-A. Mardal, R. Winther, Numerical methods for incompressible viscous flow, *Adv. Water Resour.* 25 (8) (2002) 1125–1146.
- [7] W.J. Layton, F. Schieweck, I. Yotov, Coupling fluid flow with porous media flow, *SIAM J. Numer. Anal.* 40 (6) (2002) 2195–2218.
- [8] A. Mikelić, W. Jäger, On the interface boundary condition of Beavers, Joseph, and Saffman, *SIAM J. Appl. Math.* 60 (4) (2000) 1111–1127.
- [9] B. Rivière, I. Yotov, Locally conservative coupling of Stokes and Darcy flows, *SIAM J. Numer. Anal.* 42 (5) (2005) 1959–1977.
- [10] D. Vassilev, I. Yotov, Coupling Stokes–Darcy flow with transport, *SIAM J. Sci. Comput.* 31 (5) (2009) 3661–3684, <https://doi.org/10.1137/080732146>.
- [11] J.P. Delhomme, Spatial variability and uncertainty in groundwater flow parameters: a geostatistical approach, *Water Resour. Res.* 15 (2) (1979) 269–280, <https://doi.org/10.1029/WR015i002p00269>.
- [12] R.A. Freeze, A stochastic–conceptual analysis of one-dimensional groundwater flow in nonuniform homogeneous media, *Water Resour. Res.* 11 (5) (1975) 725–741, <https://doi.org/10.1029/WR011i005p00725>.
- [13] R.J. Hoeksema, P.K. Kitanidis, Analysis of the spatial structure of properties of selected aquifers, *Water Resour. Res.* 21 (4) (1985) 563–572, <https://doi.org/10.1029/WR021i004p00563>.
- [14] I. Babuska, R. Tempone, G.E. Zouraris, Galerkin finite element approximations of stochastic elliptic partial differential equations, *SIAM J. Numer. Anal.* 42 (2) (2004) 800–825.
- [15] D. Xiu, *Numerical Methods for Stochastic Computations: A Spectral Method Approach*, Princeton University Press, Princeton, 2010.
- [16] I. Babuška, F. Nobile, R. Tempone, A stochastic collocation method for elliptic partial differential equations with random input data, *SIAM J. Numer. Anal.* 45 (3) (2007) 1005–1034.
- [17] M. Ghanem, P. Spanos, *Stochastic Finite Elements: A Spectral Approach*, Springer, New York, 1991.
- [18] I. Graham, F. Kuo, D. Nuyens, R. Scheichl, I. Sloan, Quasi-Monte Carlo methods for elliptic PDEs with random coefficients and applications, *J. Comput. Phys.* 230 (10) (2011) 3668–3694.
- [19] P. Kumar, C.W. Oosterlee, R.P. Dwight, A multigrid multilevel Monte Carlo method using high-order finite-volume scheme for lognormal diffusion problem, *Int. J. Uncertain. Quantificat.* 7 (1) (2017) 57–81.
- [20] K. Cliffe, M.B. Giles, R. Scheichl, A.L. Teckentrup, Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients, *Comput. Vis. Sci.* 14 (2011) 3–15.
- [21] M.B. Giles, Multilevel Monte Carlo path simulation, *Oper. Res.* 256 (2008) 981–986.
- [22] F.Y. Kuo, C. Schwab, I.H. Sloan, Multilevel quasi-Monte Carlo finite element methods for a class of elliptic PDEs with random coefficients, *Found. Comput. Math.* 15 (2) (2015) 411–449, <https://doi.org/10.1007/s10208-014-9237-5>.
- [23] C.E. Rasmussen, *Gaussian Processes for Machine Learning*, MIT Press, 2006.
- [24] J. Maitre, F. Musy, P. Nigon, A fast solver for the Stokes equations using multigrid with a Uzawa smoother, in: *Advances in Multi-Grid Methods*, Springer, 1985, pp. 77–83.
- [25] U. Trottenberg, C.W. Oosterlee, A. Schuller, *Multigrid*, Elsevier Academic Press, San Diego, CA, 2000.
- [26] R. Eymard, T. Gallouët, R. Herbin, Finite volume methods, *Handb. Numer. Anal.* 7 (2000) 713–1018.
- [27] F.J. Gaspar, Y. Notay, C.W. Oosterlee, C. Rodrigo, A simple and efficient segregated smoother for the discrete Stokes equations, *SIAM J. Sci. Comput.* 36 (3) (2014) A1187–A1206, <https://doi.org/10.1137/130920630>.
- [28] P. Luo, C. Rodrigo, F.J. Gaspar, C.W. Oosterlee, Uzawa smoother in multigrid for the coupled porous medium and Stokes flow system, *SIAM J. Sci. Comput.* 39 (5) (2017) S633–S661, <https://doi.org/10.1137/16M1076514>.
- [29] M. Handcock, J. Wallis, An approach to statistical spatial-temporal modeling of meteorological fields (with discussion), *J. Am. Stat. Assoc.* 89 (1994) 368–390.
- [30] R. Adler, *The Geometry of Random Fields*, Society for Industrial and Applied Mathematics, 2010.
- [31] F. dell’Isola, A. Madeo, P. Seppacher, Boundary conditions at fluid-permeable interfaces in porous media: a variational approach, *Int. J. Solids Struct.* 46 (17) (2009) 3150–3164, <https://doi.org/10.1016/j.ijsolstr.2009.04.008>.
- [32] L. Payne, B. Straughan, Analysis of the boundary condition at the interface between a viscous fluid and a porous medium and related modelling questions, *J. Math. Pures Appl.* 77 (4) (1998) 317–354, [https://doi.org/10.1016/S0021-7824\(98\)80102-5](https://doi.org/10.1016/S0021-7824(98)80102-5).
- [33] J. Bear, *Dynamics of Fluids in Porous Media*, Courier Corporation, 2013.

- [34] C. Zheng, P.P. Wang, MT3DMS: a modular three-dimensional multi-species transport model for simulation of advection, dispersion, and chemical reactions of contaminants in ground-water systems. Documentation and user's guide.
- [35] F.H. Harlow, J.E. Welch, Numerical calculation of time-dependent viscous incompressible flow of fluid with free surface, *Phys. Fluids* 8 (12) (1965) 2182–2189.
- [36] B.P. Leonard, A stable and accurate convective modelling procedure based on quadratic upstream interpolation, *Comput. Methods Appl. Mech. Eng.* 19 (1) (1979) 59–98.
- [37] B. Leonard, S. Mokhtari, Beyond first-order upwinding: the ultra-sharp alternative for non-oscillatory steady-state simulation of convection, *Int. J. Numer. Methods Eng.* 30 (4) (1990) 729–766.
- [38] B. Leonard, The ULTIMATE conservative difference scheme applied to unsteady one-dimensional advection, *Comput. Methods Appl. Mech. Eng.* 88 (1) (1991) 17–74.
- [39] P.K. Sweby, High resolution schemes using flux limiters for hyperbolic conservation laws, *SIAM J. Numer. Anal.* 21 (5) (1984) 995–1011.
- [40] B. Koren, A robust upwind discretization method for advection, diffusion and source terms, *Notes Numer. Fluid Mech.* 45 (1993) 117–138.
- [41] D.W. Peaceman, J.H.H. Rachford, The numerical solution of parabolic and elliptic differential equations, *J. Korea Soc. Ind. Appl. Math.* 3 (1) (1955) 28–41, <https://doi.org/10.1137/0103003>.
- [42] M. Khalil, P. Wesseling, Vertex-centered and cell-centered multigrid for interface problems, *J. Comput. Phys.* 98 (1992) 1–10.
- [43] P. Chidyagwai, B. Rivière, Numerical modelling of coupled surface and subsurface flow systems, *Adv. Water Resour.* 33 (1) (2010) 92–105.
- [44] S. Turek, J. Hron, M. Mádlík, M. Razzaq, H. Wobker, J.F. Acker, Numerical Simulation and Benchmarking of a Monolithic Multigrid Solver for Fluid–Structure Interaction Problems with Application to Hemodynamics, Springer, Berlin, Heidelberg, 2010, pp. 193–220.
- [45] S. Mishra, C. Schwab, Sparse tensor Multi-level Monte Carlo finite volume methods for hyperbolic conservation laws with random initial data, *Math. Comput.* 81 (280) (2012) 1979–2018.
- [46] S. Mishra, C. Schwab, J. Šukys, Multi-level Monte Carlo finite volume methods for nonlinear systems of conservation laws in multi-dimensions, *J. Comput. Phys.* 231 (8) (2012) 3365–3388.
- [47] S. Mishra, C. Schwab, J. Šukys, Multi-level Monte Carlo finite volume methods for uncertainty quantification of acoustic wave propagation in random heterogeneous layered medium, *J. Comput. Phys.* 312 (2016) 192–217.
- [48] P. Kumar, C. Rodrigo, F.J. Gaspar, C.W. Oosterlee, On cell-centered multigrid methods and Local Fourier Analysis for PDEs with random coefficients, preprint, arXiv:1803.08864.
- [49] C. Dietrich, G. Newsam, Fast and exact simulation of stationary Gaussian processes through circulant embedding of the covariance matrix, *SIAM J. Sci. Comput.* 18 (1997) 1088–1107.
- [50] N. Hanspal, A. Waghode, V. Nassehi, R. Wakeman, Development of a predictive mathematical model for coupled Stokes/Darcy flows in cross-flow membrane filtration, *Chem. Eng. J.* 149 (1–3) (2009) 132–142, <https://doi.org/10.1016/j.cej.2008.10.012>.
- [51] J.S. Šukys, C. Schwab, Static load balancing for multi-level Monte Carlo finite volume solvers, in: *Parallel Processing and Applied Mathematics*, Springer, Berlin, Heidelberg, 2012, pp. 245–254.
- [52] D. Drzisga, B. Gmeiner, U. Råde, R. Scheichl, B. Wohlmuth, Scheduling massively parallel multigrid for multilevel Monte Carlo methods, *SIAM J. Sci. Comput.* 39 (5) (2017) S873–S897, <https://doi.org/10.1137/16M1083591>.