Journal of Computational Physics 228 (2009) 591–618

Contents lists available at ScienceDirect

Journal of Computational Physics

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

A stochastic multiscale framework for modeling flow through random heterogeneous porous media

B. Ganapathysubramanian [1], N. Zabaras *

Materials Process Design and Control Laboratory, Sibley School of Mechanical and Aerospace Engineering, 101 Frank H.T. Rhodes Hall, Cornell University, Ithaca, NY 14853-3801, USA

ARTICLE INFO

Article history:
Received 4 May 2008
Received in revised form 2 October 2008
Accepted 7 October 2008
Available online 17 October 2008

Keywords:
Stochastic partial differential equations
Collocation methods
Sparse grids
Scalable algorithms
Non-linear model reduction
Manifold learning
Variational multiscale methods
Mixed finite elements
Data-driven modeling

ABSTRACT

Flow through porous media is ubiquitous, occurring from large geological scales down to the microscopic scales. Several critical engineering phenomena like contaminant spread, nuclear waste disposal and oil recovery rely on accurate analysis and prediction of these multiscale phenomena. Such analysis is complicated by inherent uncertainties as well as the limited information available to characterize the system. Any realistic modeling of these transport phenomena has to resolve two key issues: (i) the multi-length scale variations in permeability that these systems exhibit, and (ii) the inherently limited information available to quantify these property variations that necessitates posing these phenomena as stochastic processes.

A stochastic variational multiscale formulation is developed to incorporate uncertain multiscale features. A stochastic analogue to a mixed multiscale finite element framework is used to formulate the physical stochastic multiscale process. Recent developments in linear and non-linear model reduction techniques are used to convert the limited information available about the permeability variation into a viable stochastic input model. An adaptive sparse grid collocation strategy is used to efficiently solve the resulting stochastic partial differential equations (SPDEs). The framework is applied to analyze flow through random heterogeneous media when only limited statistics about the permeability variation are given.

1. Introduction

Thermal and hydrodynamic transport in random heterogeneous media are ubiquitous processes occurring in various scales ranging from the large scale (e.g. geothermal energy systems, oil recovery, geological heating of the earth's crust) to smaller scales (e.g. heat transfer through composites, polycrystals, flow through pores, inter-dendritic flow in solidification, heat transfer through fluidized beds). There has been increasing interest in reliably modeling and predicting the thermal and hydrodynamic behavior of such media. One of the challenging mathematical issues in the analysis of transport through heterogeneous random media is the multiscale nature of the property variations. Complete response evaluation involving full-scale spatial and temporal resolution simulations of multiscale systems is extremely expensive. Computational techniques have been developed that solve for a coarse-scale solution by defining an appropriate coarse-scale problem that captures the effect of the subgrid-scales [1]. The more popular techniques developed for such upscaling fall under the category of multiscale methods viz. the variational multiscale (VMS) method (operator upscaling) [2–7], the heterogenous

* Corresponding author. Fax: +1 607 255 1222.
E-mail address: zabaras@cornell.edu (N. Zabaras).
URL: http://mpdc.mae.cornell.edu/ (N. Zabaras).
[1] Currently with the Department of Mechanical Engineering, Iowa State University, Ames, Iowa.

0021-9991/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jcp.2008.10.006

multiscale method [8,9] and the multiscale finite element method [10–14]. Further related techniques are presented in [15,16]. These computationally scalable techniques have resulted in the development of black box simulators that have been used with significant success in the solution of large scale transport problems in complex geometries.[2]

The multiscale analysis of such systems inherently assumes that the complete multiscale variation of the permeability is known. This assumption limits the applicability of these frameworks since it is usually not possible to experimentally determine the complete structure of the media at smaller scales [17]. In most cases, only a few statistical descriptors of the property variation or the property variation in small test regions can be experimentally determined. This limited information about the permeability necessitates viewing the permeability variation as a random field that satisfies certain statistical properties/correlations. This naturally results in describing the physical phenomena as stochastic partial differential equations (SPDEs) instead of partial differential equations (PDEs).

In the past decade, there has been tremendous progress in posing and solving SPDEs. Several techniques like Generalized Polynomial Chaos expansion [18–24], wavelet expansion and collocation based strategies [25–28] have been developed to solve SPDEs. These techniques have been applied with significant success to solve single scale stochastic problems. In recent years, there has been significant interest in coupling deterministic multiscale methods with techniques in stochastic analysis to investigate critical multiscale systems in the presence of uncertainties.

There are two key questions that have to be sequentially answered to construct a stochastic multiscale framework that models transport phenomena only given limited information: (1) techniques to convert/encode limited statistical information and/or sampled property values into a viable (stochastic) input model for the (multiscale) permeability variation and (2) given this input multiscale stochastic model, techniques to solve the stochastic multiscale equations.

The construction of viable stochastic input models based on limited data is a very interesting and challenging mathematical problem. The recent work in [29] looks at developing probabilistic models of random coefficients in SPDEs using a maximum likelihood framework. The random domain decomposition (RDD) [30,31] method was used to construct probabilistic models for highly variable permeability distributions [32]. All these techniques are very specialized to particular applications and require some amount of expert knowledge in allocating probability distributions. In [33,34], we have recently developed techniques to utilize statistical information about the variability in the property of random media and produce viable low-dimensional descriptors as inputs to the SPDE describing the evolution of the dependent variable.

The basic idea to solve the stochastic multiscale set of equations is to extend deterministic multiscale methods to their stochastic analogues. Spectral strategies to pose and solve stochastic multiscale problems have been investigated by Xu [35] and Asokan and Zabaras [36]. Collocation based strategies to pose and solve stochastic multiscale problems have recently been developed in [33]. The key is to define appropriate ways to link the subgrid-scale stochastic variation with the coarse-scale stochastic variation of the dependent variables.

In the present work we are interested in analyzing flow through random heterogeneous media given limited statistical information about the multiscale permeability variation. We link stochastic analysis and multiscale methods to investigate this problem. The key contributions of the current work are: (1) *formulation of a stochastic variational multiscale method* to incorporate the effects of stochastic multiscale permeability variations. (2) *Utilizing various data-driven strategies* to encode the limited information and subsequently construct a finite-dimensional representation of the multiscale permeability variation. (3) *Utilizing an adaptive sparse grid collocation strategy* to effectively solve the multiscale SPDEs. Even though (2) and (3) have been developed elsewhere, this is the first time (to our best knowledge) that data-driven model generation and fast adaptive collocation strategies have been coupled with a consistent stochastic variational multiscale framework. This sequential development – from limited data to plausible stochastic models to a computationally efficient solution strategy – also serves as a roadmap for seamlessly constructing stochastic multiscale frameworks using deterministic solvers.

2. Problem definition

Fig. 1 shows a schematic of the problem of interest. Denote the domain as $\mathcal{D} \subset \mathbb{R}^{n_{sd}}$, where n_{sd} is the number of spatial dimensions. This paper deals with developing (stochastic multiscale) methods for the simulation of pressure and (phase) velocities in porous media flow across this domain. Stochastic multiscale structures in the domain are reflected in the coefficients of the governing partial differential equations. Assuming incompressibility, the problem reduces to a variable-coefficient elliptic equation. Without loss of generality, we will study the following elliptic problem [37] describing steady-state flow through heterogeneous porous media:

$$\nabla \cdot \boldsymbol{u} = f, \tag{1}$$

$$\boldsymbol{u} = -k\nabla p, \tag{2}$$

where p is the pressure, $\boldsymbol{u}$ is the flow velocity and k is the permeability.[3] The characteristic length scale of $\mathcal{D}$ is L. Denote the length scale of permeability fluctuation as l. In the problems that we are interested in solving, the characteristic length of the domain is a couple of orders of magnitude larger than the characteristic length scale of the permeability variations, $l \ll L$.

[2] In the context of flow through porous media, these strategies are usually applied to geological scale domains and petro-physical property models. In such problems, the property variations are homogenized over individual cell blocks. Each individual cell is large enough for the concept of the RVE to be valid [17]. The number of such cells in the geological domain necessitates multiscaling.

[3] More precisely, k is the permeability divided by the fluid viscosity. Since we assume the fluid viscosity is uniform, k refers to a scaled permeability, without loss of generality. $\boldsymbol{u}$ is the volumetric flow density.

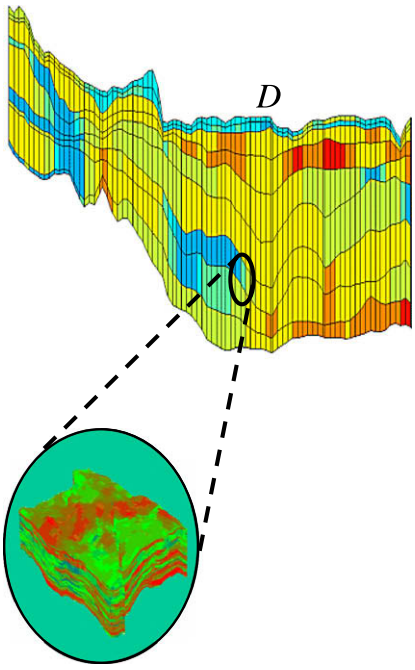


Fig. 1. Schematic of the problem of interest.

The variables $(\boldsymbol{u}(\boldsymbol{x}), p(\boldsymbol{x}))$ depend on the (multiscale) permeability distribution, $k(\boldsymbol{x})$ in the domain. However, the complete permeability distribution in unknown. Only some limited statistics and/or snapshots of the permeability are given. This limited information available to characterize the permeability necessitates assuming that the permeability is a realization of a random field. This is mathematically stated as follows: Let Ω be the space of all allowable permeability variations. This is our event space. Every point $k = \{k(\boldsymbol{x}, \omega), \forall \boldsymbol{x} \in \mathcal{D}, \omega \in \Omega\}$ in this space is equiprobable. Consequently, we can define a σ-algebra $\mathcal{F}$ and a corresponding probability measure $\mathcal{P} : \mathcal{F} \to [0, 1]$ to construct a complete probability space $(\Omega, \mathcal{F}, \mathcal{P})$ of allowable permeability. To make this abstract description amenable to numerical simulation, a finite-dimensional approximation/representation [38] of this abstract set is necessary. Various data-driven strategies to represent the set Ω as a finite-dimensional function are discussed in Section 4. The stochastic permeability is represented as

$$k(\boldsymbol{x}, \omega) \equiv k(\boldsymbol{x}, Y_1, \ldots, Y_N) \equiv k(\boldsymbol{x}, \boldsymbol{Y}), \tag{3}$$

where $Y_1, \ldots, Y_N$ are uncorrelated random variables. Usually, the permeability field is spatially correlated [17] and is assumed to be stationary. The pressure and velocity are then characterized by:

$$\nabla \cdot \boldsymbol{u}(\boldsymbol{x}, \boldsymbol{Y}) = f(\boldsymbol{x}), \tag{4}$$

$$\boldsymbol{u}(\boldsymbol{x}, \boldsymbol{Y}) = -k(\boldsymbol{x}, \boldsymbol{Y}) \nabla p(\boldsymbol{x}, \boldsymbol{Y}), \tag{5}$$

where the source/sink term $f(\boldsymbol{x})$ is taken to be deterministic. Furthermore, f is also assumed to not have a multiscale character. These assumptions are simply to make the subsequent developments clear. It is fairly straightforward to account for uncertainties in f as well as to analyze the effects of a multiscale f [39].

For the problem to be physically relevant, we assume that the stochastic permeability satisfies some conditions, particularly:

Assumption 1. k is positive and uniformly coercive:

$$\begin{aligned} &\exists k_{min}, k_{max} \in (0, \infty) : \\ &\mathcal{P}(\omega \in \Omega : k(\boldsymbol{x}, \omega) \in [k_{min}, k_{max}], \quad \forall \boldsymbol{x} \in \mathcal{D}) = 1. \end{aligned} \tag{6}$$

As stated earlier, the abstract representation of $k(., \omega)$ in Ω is replaced by a more tractable finite-dimensional representation $k(., \boldsymbol{Y})$, with $\boldsymbol{Y} \in \Gamma \subset \mathbb{R}^N$. Corresponding to the probability measure $\mathcal{P} : \mathcal{F} \to [0, 1]$, we denote the equivalent probability measure $\rho : \Gamma \to [0, 1]$.

The governing equations for the velocity and pressure are represented in mixed form as follows:

$$k^{-1}\boldsymbol{u} + \nabla p = 0, \tag{7}$$

$$\nabla \cdot \boldsymbol{u} = f. \tag{8}$$

The above constitutive and conservation equations are supplemented with the following boundary conditions

$$p = p_o \quad \text{on} \quad \partial\mathcal{D}_p, \tag{9}$$

$$\boldsymbol{u}\cdot\boldsymbol{n} = u_o \quad \text{on} \quad \partial\mathcal{D}_u. \tag{10}$$

Without loss of generality, we assume that the boundary conditions are deterministic and that the Neumann condition is homogeneous [39], $u_o = 0$ on $\partial\mathcal{D}_u$.

The basic idea is to solve the problem on a coarse spatial discretization, $\mathcal{D}^c$, while taking into account the subgrid-scale variation in the stochastic permeability. In the next section, we detail a stochastic extension to the variational multiscale method to solve this problem. The stochastic multiscale formulation is based on the multiscale formulation detailed in the papers by Juanes and Dub [39] and Arbogast and Boyd [40]. The augmentation of the deterministic multiscale formulation to its stochastic counterpart mainly depends on two straightforward developments: (a) defining the appropriate stochastic function spaces (both finite- and infinite-dimensional), and (b) defining the appropriate localization assumptions. For brevity of presentation, we will concentrate on these two aspects. The interested reader in referred to [39,40] for other details.

3. Stochastic variational multiscale formulation

We now introduce the appropriate function spaces in which the velocity and pressure lie. In contrast to their deterministic counterparts, the velocity and pressure defined here are stochastic processes. Most dependent variables one encounters in stochastic analysis are usually random processes [41–44] that are defined in product spaces. These variables (for instance $p(\boldsymbol{x},\boldsymbol{Y})$) usually have one structure in the stochastic space $p(.,\boldsymbol{Y})$ and another in the physical space $p(\boldsymbol{x},.)$. The numerical analysis/approximation [41,42] of such functions can be performed by defining the appropriate *tensor product spaces*. The interested reader is referred to [41–44] for insightful discussions on the definitions of these product spaces. Following [41,42], we define appropriate function spaces that encode variations of the function in the physical domain $\mathcal{D}$ and in the stochastic space Γ.

Remark 3.1. We will consider stochastic functions that belong to $\mathcal{S}$, the space of all square integrable functions, with measure $\rho(\boldsymbol{Y})$. Given a deterministic function space A, its stochastic counterpart is the tensor product space $\mathcal{A} \equiv \mathcal{S}\otimes A$.

We introduce the following tensor product function spaces:

$$\mathcal{W} \equiv \mathcal{S}\otimes W \equiv L^2(\Gamma)\otimes L^2(\mathcal{D}), \quad \text{with}, \quad (p,p) \equiv \|p\|^2_{\mathcal{W}} := \int_\Gamma d\rho(\boldsymbol{Y}) \int_{\mathcal{D}} p^2 dx, \tag{11}$$

$$\mathcal{H} \equiv \mathcal{S}\otimes H \equiv L^2(\Gamma)\otimes H(div,\mathcal{D}), \quad \text{with}, \quad (\boldsymbol{u},\boldsymbol{u}) \equiv \|\boldsymbol{u}\|^2_{\mathcal{H}} := \int_\Gamma d\rho(\boldsymbol{Y}) \int_{\mathcal{D}} \boldsymbol{u}\cdot\boldsymbol{u} dx, \tag{12}$$

$$\mathcal{V} = \{\boldsymbol{u} : \boldsymbol{u}\in\mathcal{H}, \boldsymbol{u}(.,\boldsymbol{Y})\cdot\boldsymbol{n} = 0 \quad \text{on} \quad \partial\mathcal{D}_u, \forall \boldsymbol{Y}\in\Gamma\}. \tag{13}$$

The problem defined by Eqs. (7) and (8) along with the boundary conditions Eqs. (9) and (10) can be written in mixed variational form: Find $(\boldsymbol{u},p)\in\mathcal{V}\times\mathcal{W}$ such that

$$(\boldsymbol{v}, k^{-1}\boldsymbol{u}) - (\nabla\cdot\boldsymbol{v}, p) = -\langle \boldsymbol{v}\cdot\boldsymbol{n}, p_o\rangle, \quad \forall \boldsymbol{v}\in\mathcal{V}, \tag{14}$$

$$(w, \nabla\cdot\boldsymbol{u}) = (w,f), \quad \forall w\in\mathcal{W}, \tag{15}$$

where $\langle f,g\rangle$ is defined as $\int_\Gamma d\rho(\boldsymbol{Y}) \int_{\partial\mathcal{D}_p} fg d\boldsymbol{x}$ and $(f,g) = \int_\Gamma d\rho(\boldsymbol{Y}) \int_{\mathcal{D}} fg d\boldsymbol{x}$.

3.1. Variational multiscale approach

The exact solution $\boldsymbol{u}$ is assumed to be made up of contributions [5] from two different (spatial) scales namely, the coarse-scale solution $\boldsymbol{u}_c(\boldsymbol{x},.)$ that can be resolved using a coarse (spatial) mesh and a subgrid solution $\boldsymbol{u}_f(\boldsymbol{x},.)$:

$$\boldsymbol{u} = \boldsymbol{u}_c + \boldsymbol{u}_f, \quad p = p_c + p_f. \tag{16}$$

This additive sum decomposition induces a similar decomposition for the spatial part of full-scale tensor product function spaces into a direct sum of a coarse-scale and a subgrid tensor product function spaces, i.e.

$$\mathcal{W} = \mathcal{W}_c \oplus \mathcal{W}_f, \quad \mathcal{H} = \mathcal{H}_c \oplus \mathcal{H}_f, \quad \mathcal{V} = \mathcal{V}_c \oplus \mathcal{V}_f. \tag{17}$$

The main idea is to develop models for characterizing the effect of the subgrid solution $\boldsymbol{u}_f(\boldsymbol{x},.)$ on the coarse-scale solution and to subsequently derive a modified coarse-scale formulation that only involves $\boldsymbol{u}_c(\boldsymbol{x},.)$. The additive decomposition provides a way of splitting the full-scale problem given by Eqs. (14) and (15) into a coarse-scale problem and a subgrid problem. Testing against the coarse-scale test functions results in the coarse-scale variational problem: Find $(\boldsymbol{u}_c,p_c)\in\mathcal{V}_c\times\mathcal{W}_c$ such that

$$(\boldsymbol{v}_c, k^{-1}(\boldsymbol{u}_c+\boldsymbol{u}_f)) - (\nabla\cdot\boldsymbol{v}_c, (p_c+p_f)) = -\langle \boldsymbol{v}_c\cdot\boldsymbol{n}, p_o\rangle, \quad \forall \boldsymbol{v}_c\in\mathcal{V}_c, \tag{18}$$

$$(w_c, \nabla\cdot(\boldsymbol{u}_c+\boldsymbol{u}_f)) = (w_c, f), \quad \forall w_c\in\mathcal{W}_c. \tag{19}$$

Similarly testing against the subgrid test functions results in the subgrid variational problem: Find $(\boldsymbol{u}_f, p_f) \in \mathcal{V}_f \times \mathcal{W}_f$ such that

$$(\boldsymbol{v}_f, k^{-1}(\boldsymbol{u}_c + \boldsymbol{u}_f)) - (\nabla \cdot \boldsymbol{v}_f, (p_c + p_f)) = -\langle \boldsymbol{v}_f \cdot \boldsymbol{n}, p_o \rangle, \quad \forall \boldsymbol{v}_f \in \mathcal{V}_f, \tag{20}$$

$$(w_f, \nabla \cdot (\boldsymbol{u}_c + \boldsymbol{u}_f)) = (w_f, f), \quad \forall w_f \in \mathcal{W}_f. \tag{21}$$

Note that both the coarse-scale variational equations Eqs. (18) and (19) as well as the subgrid variational equations Eqs. (20) and (21) contain coarse and subgrid variations in the dependent variable $(\boldsymbol{u}_c, p_c; \boldsymbol{u}_f, p_f)$. The key is to solve Eqs. (20) and (21) for $\boldsymbol{u}_f$ and construct a functional representation of the subgrid variation, $\boldsymbol{u}_f$ and p_f *in terms of the coarse-scale variation*, $\boldsymbol{u}_c$:

$$\boldsymbol{u}_f = \Phi(\boldsymbol{u}_c), \quad p_f = \Psi(\boldsymbol{u}_c). \tag{22}$$

This representation can be subsequently used to remove explicit dependence of $\boldsymbol{u}_f$ and p_f in Eqs. (18) and (19). The key problem is now to solve Eqs. (20) and (21). But these equations are defined over the complete global domain, $\mathcal{D}$. Solving the subgrid problem on the global domain is equivalent to solving the full-scale problem itself. To make the problem computationally tractable, some localization assumptions have to be made to convert this global subgrid problem into a set of local subgrid problems [39].

3.2. Finite element approximation spaces

Before defining the localization assumptions, the finite element approximation spaces at the full-scale and coarse-scale are first defined. As stated before (Remark 3.1), the stochastic function spaces are simply tensor product spaces of $\mathcal{S}$ with the corresponding deterministic spaces.

3.2.1. Full-scale approximation spaces

Let $\mathcal{V}_h$ and $\mathcal{W}_h$ be finite-dimensional subspaces of the spatial part of the corresponding continuum spaces $\mathcal{V}$ and $\mathcal{W}$. That is

$$\mathcal{W}_h = \mathcal{S} \otimes W_h, \quad \mathcal{V}_h = \mathcal{S} \otimes V_h. \tag{23}$$

Note that W_h and V_h should satisfy the discrete inf–sup condition [45]. Consider a partition, $\mathcal{T}_h$ of the domain $\mathcal{D}$ into non-overlapping elements e_i, $\mathcal{T}_h = \bigcup_{i=1}^{N_h} e_i$, where N_h is the number of elements of the grid. Following [39], define also the skeleton of the partition, $\mathcal{SP}_h = \bigcup_{a=1}^{M_h} \gamma_a$, where M_h is the number of element faces denoted by γ_a. The partition $\mathcal{T}_h$ is denoted as the full-scale grid, on which the multiscale permeability is defined. The deterministic finite element approximation space for the velocity, V_h is taken to be the lowest-order Raviart-Thomas [46,39] space, $RT_0(\mathcal{T}_h)$, and the deterministic finite element approximation space for the full-scale pressure, W_h is taken to be the space of piece-wise constants on the full-scale mesh, $P_0(\mathcal{T}_h)$.

3.2.2. Coarse-scale approximation spaces

Consider a coarse-scale partition of the domain, $\mathcal{D}$. Denote this partition as $\mathcal{T}_c = \bigcup_{i=1}^{N_c} E_i$. Denote by $\mathcal{SP}_c = \bigcup_{a=1}^{M_c} \Lambda_a$ the associated skeleton of the coarse-scale discretization. Here, N_c is the number of coarse elements and M_c is the number of coarse element faces denoted by Λ_a. Following [39], we also assume for simplicity that the partitions $\mathcal{T}_h$ and $\mathcal{T}_c$ – the full- and coarse-grid, respectively – are nested, conforming, and consist of rectangular elements. Since $\mathcal{W}_h = \mathcal{S} \otimes P_o(\mathcal{T}_h)$, choose the coarse-scale pressure as belonging to the space of piecewise constant functions, $\mathcal{W}_c^h = \mathcal{S} \otimes \mathcal{P}_o(\mathcal{T}_c)$. The choice of the coarse-scale velocity approximation spaces now has to be compatible with this choice of the pressure approximation. In the analogous deterministic developments, the two choices for such spaces were the lowest-order Raviart-Thomas space (used in [39]) and the Brezzi-Douglas-Marini space of order 1 (used in [3]). We choose to utilize the stochastic analogue to the lowest-order Raviart-Thomas space in the current work. We can now associate the coarse-scale velocity with the lowest-order Raviart-Thomas space, $RT_0(\mathcal{T}_c)$ as:

$$\mathcal{V}_c^h = \mathcal{S} \otimes V_c^h, \quad V_c^h = \left\{ \boldsymbol{u}_c : \boldsymbol{u}_c = \sum_{a=1}^{M_c} \boldsymbol{N}_a^c u_a^c, u_a^c = 0 \quad \forall \Lambda_a \in \partial \mathcal{D}_u \right\}, \tag{24}$$

where $\mathcal{V}_c^h$ is the finite-dimensional approximation to the coarse-scale continuum space $\mathcal{V}_c$. Here, $\boldsymbol{N}_a^c$ is the RT_0 basis function associated with face Λ_a, and u_a^c is the corresponding degree of freedom (the integrated coarse-flux through Λ_a [39]). The coarse-scale pressure approximation is piecewise constant on the coarse-mesh, $P_0(\mathcal{T}_c)$:

$$\mathcal{W}_c^h = \mathcal{S} \otimes W_c^h, \quad W_c^h = \left\{ w_c : w_c = \sum_{i=1}^{N_c} \phi_i^c w_i^c \right\}, \tag{25}$$

where $\mathcal{W}_c^h$ is the finite-dimensional approximation to the coarse-scale continuum space $\mathcal{W}_c$. Here, ϕ_i^c is the pressure basis function for the coarse element i, which is equal to one in element E_i and zero in all other elements. w_i^c is the corresponding pressure degree of freedom (the average pressure in coarse element E_i).

3.2.3. Subgrid-scale approximation spaces

The subgrid-scale velocities have to accurately mimic the full-scale velocities. To ensure this, they are restricted to belong to the lowest-order Raviart-Thomas space on the full-scale grid *within each coarse element* [39]. Denote by $E_{i,h} = \mathcal{T}_h(E_i)$ the full-scale grid defined over the coarse element E_i. The subgrid-scale velocity field defined on each coarse element satisfies the condition

$$V^h_{f,i} \subset RT_0(E_{i,h}), \tag{26}$$

$$\mathcal{V}^h_{f,i} = \mathcal{S} \otimes V^h_{f,i}, \tag{27}$$

where $\mathcal{V}^h_{f,i}$ is the finite-dimensional approximation to the subgrid-scale continuum space $\mathcal{V}_f$, constrained to a coarse element E_i. The elements of $\mathcal{V}^h_{f,i}$ can naturally be extended to all of $\mathcal{D}$ by zero. The global subgrid velocity space can then be defined as the direct sum of the subgrid spaces over the coarse elements

$$\mathcal{V}^h_f = \bigoplus_{i=1}^{N_c} \mathcal{V}^h_{f,i}. \tag{28}$$

Similarly, the subgrid pressure space is restricted to belong to the space of piecewise constant functions on each coarse element

$$W^h_{f,i} \subset P_0(E_{i,h}), \tag{29}$$

$$\mathcal{W}^h_{f,i} = \mathcal{S} \otimes W^h_{f,i}, \tag{30}$$

where $\mathcal{W}^h_{f,i}$ is the finite-dimensional approximation to the subgrid-scale continuum space $\mathcal{W}_f$, constrained to a coarse element E_i. The elements of these spaces are extended to zero functions to the entire spatial domain [39], $\mathcal{D}$, and the subgrid pressure space is defined as

$$\mathcal{W}^h_f = \bigoplus_{i=1}^{N_c} \mathcal{W}^h_{f,i}. \tag{31}$$

3.3. Localization assumptions

Having defined the appropriate finite element approximation spaces, we now move to the problem of reducing the global subgrid problem (defined by Eqs. (20) and (21)) into a set of local subgrid problems defined over the coarse elements E_i.

The key localization assumption in the construction of the multiscale framework is to ensure that the approximation is locally conservative at both scales [39]. This implies that the discrete version of the conservation equation is satisfied on each element in the coarse and full-scale grids. Furthermore, we also assume that this condition is satisfied *not in a distributed sense, but for each stochastic realization (i.e. point-wise in stochastic space)*. This leads to the following condition

$$(w_c, \nabla \cdot \boldsymbol{u}_c)_{E_i} = (w_c, f)_{E_i}, \tag{32}$$

where $(.,.)_{E_i}$ is the corresponding inner product defined over the coarse element E_i. This condition (in its strong, point-wise form) is equivalent to [39]

$$\nabla \cdot \boldsymbol{u}_c = \Pi_c f, \tag{33}$$

where $\Pi_c f$ is the projection of the source/sink term onto the space W_c of piecewise constants on the coarse-grid. As stated in the beginning of this section, we assume that the sourse/sink function does not display a multiscale nature, i.e. it is equal to its projection on the space of coarse-scale pressures ($f = \Pi_c f$). Substituting Eq. (32) into the coarse-scale conservation equation Eq. (19) results in

$$(w_c, \nabla \cdot \boldsymbol{u}_f)_{E_i} = 0, \quad \forall E_i \in \mathcal{T}_c. \tag{34}$$

Since w_c is constant in each coarse element, using the divergence theorem and the assumption that this condition is valid point-wise in stochastic space, results in the following condition on the subgrid velocities in each coarse element

$$\int_{\partial E_i} \boldsymbol{u}_f \cdot \boldsymbol{n} d\boldsymbol{x} = 0, \quad \forall E_i \in \mathcal{T}_c. \tag{35}$$

Eq. (35) is the essential condition [39] that guarantees mass conservation at both scales and allows for the localization of the subgrid problem. This localization assumption converts the global subgrid problem into a set of local Neumann problems with boundary conditions defined by Eq. (35). Furthermore, the subgrid test function $\boldsymbol{v}_f$ must satisfy $\boldsymbol{v}_f \cdot \boldsymbol{n} = 0$ on ∂E_i which results in the following constraint

$$(\nabla \cdot \boldsymbol{v}_f, p_c)_{E_i} = 0, \quad \forall E_i \in \mathcal{T}_c. \tag{36}$$

Finally, we define the space of subgrid pressure $\mathcal{W}^h_f$ as the orthogonal complement of $\mathcal{W}^h_c$ in $\mathcal{W}_h$ (i.e. $(w_f, w_c) = 0$). Now, since $\mathcal{W}^h_c = \mathrm{div}\mathcal{V}^h_c$, this results in an additional constraint

$$(\nabla \cdot \boldsymbol{v}_c, p_f)_{E_i} = 0, \quad \forall E_i \in \mathcal{T}_c. \tag{37}$$

3.4. The subgrid problem

Using the localization assumption Eqs. (35)–(37) the localized subgrid problem can be defined as follows: For each coarse element, $E_i = 1, \ldots, N_c$, find $(\boldsymbol{u}_f, p_f) \in \mathcal{V}^h_{f,i} \times \mathcal{W}^h_{f,i}$, such that

$$(\boldsymbol{v}_f, k^{-1}\boldsymbol{u}_f)_{E_i} - (\nabla \cdot \boldsymbol{v}_f, p_f)_{E_i} = -(\boldsymbol{v}_f, k^{-1}\boldsymbol{u}_c)_{E_i}, \quad \forall \boldsymbol{v}_f \in \mathcal{V}^h_{f,i}, \tag{38}$$

$$(w_f, \nabla \cdot \boldsymbol{u}_f)_{E_i} = (w_f, f - \nabla \cdot \boldsymbol{u}_c)_{E_i}, \quad \forall w_f \in \mathcal{W}^h_{f,i}. \tag{39}$$

Since $f = \nabla \cdot \boldsymbol{u}_c$ (assumption that the sourse/sink exhibits no multiscale character), the RHS of Eq. (39) is identically zero. The interested reader is referred to [39] for detailed discussion of the case when the source/sink function displays multiscale behavior where deterministic analysis is provided. It is straightforward to extend those arguments to the stochastic case. Note that given $\boldsymbol{u}_c$ and the appropriate boundary conditions, the above problem has a unique solution for the subgrid velocity and pressure. The global subgrid-scale solution $(\boldsymbol{u}_f, p_f)$ is obtained by patching together the solutions on each coarse element.

3.5. Multiscale basis functions

Eqs. (38) and (39) represent the subgrid-scale velocity and pressure in terms of the coarse-scale velocity. Recollect that the coarse-scale velocity is represented in terms of the RT_0 basis functions (Eq. (24)) and the M_c degrees of freedom as $\boldsymbol{u}_c = \sum_{a=1}^{M_c} \boldsymbol{N}^c_a u^c_a$. We therefore represent the subgrid-scale variation in terms of the finite number of coarse-scale degrees of freedom [39] as

$$\boldsymbol{u}_f = \sum_{a=1}^{M_c} \boldsymbol{N}^f_a u^c_a, \quad p_f = \sum_{a=1}^{M_c} \phi^f_a u^c_a. \tag{40}$$

The above equations transform the problem of finding the function $\Phi(\boldsymbol{u}_c)$ and $\Psi(\boldsymbol{u}_c)$ (defined in Eq. (22)) into finding the subgrid scale basis functions $(\boldsymbol{N}^f_a, \phi^f_a)$ associated with each coarse-scale degree of freedom. For computational efficiency, the sum of the coarse-scale and subgrid-scale components is calculated instead (defined as the multiscale basis function which represents the multiscale velocity)

$$\boldsymbol{u} = \sum_{a=1}^{M_c} (\boldsymbol{N}^c_a + \boldsymbol{N}^f_a) u^c_a = \sum_{a=1}^{M_c} \boldsymbol{N}^{ms}_a u^c_a. \tag{41}$$

The multiscale basis functions are associated with each coarse-scale interface Λ_a. The multiscale basis functions are constructed based on defining the Green's function to the subgrid Eqs. (38) and (39) [5,6]. In the context of deterministic mixed multiscale methods, various authors have constructed these multiscale basis functions [3,13,14]. We utilize the deterministic definition of the multiscale basis functions used in [37,39] to define the stochastic multiscale basis functions. Each multiscale basis function is the solution to a flow problem restricted to a pair of adjacent coarse elements that share a common coarse interface. The course terms are specified in such a way that the flow through the interface is identically one (based on the Green's function idea). It is straightforward to extend this definition of the multiscale basis function to the case when the permeability is stochastic.

The stochastic multiscale basis functions $(\boldsymbol{N}^{ms}_a, \phi^{ms}_a)$ for the interface Λ_a (which is shared by the coarse elements E_i and E_j) are the solution to the following stochastic local problem:

$$k^{-1}\boldsymbol{N}^{ms}_a + \nabla \phi^{ms}_a = 0, \quad \text{in} \quad E_i \cup E_j, \tag{42}$$

$$\boldsymbol{N}^{ms}_a \cdot \boldsymbol{n} = 0, \quad \text{on} \quad \partial(E_i \cup E_j), \tag{43}$$

$$\nabla \cdot \boldsymbol{N}^{ms}_a = \begin{cases} +k(\boldsymbol{x}, .)/\int_{E_i} k(\boldsymbol{x}, .) d\boldsymbol{x}, & \text{if} \quad \boldsymbol{x} \in E_i, \\ -k(\boldsymbol{x}, .)/\int_{E_j} k(\boldsymbol{x}, .) d\boldsymbol{x}, & \text{if} \quad \boldsymbol{x} \in E_j. \end{cases} \tag{44}$$

3.6. The coarse-scale problem

We next turn to the coarse-scale variational equations. Based on the localization assumptions Eqs. (35)–(37), the coarse-scale variational equations Eqs. (18) and (19) reduce to the following: Find $(\boldsymbol{u}_c, p_c) \in \mathcal{V}^h_c \times \mathcal{W}^h_c$, such that

$$(\boldsymbol{v}_c, k^{-1}\boldsymbol{u}_c) + (\boldsymbol{v}_c, k^{-1}\boldsymbol{u}_f) - (\nabla \cdot \boldsymbol{v}_c, p_c) = -\langle \boldsymbol{v}_c \cdot \boldsymbol{n}, p_o \rangle, \quad \forall \boldsymbol{v}_c \in \mathcal{V}^h_c, \tag{45}$$

$$(w_c, \nabla \cdot \boldsymbol{u}_c) = (w_c, f), \quad \forall w_c \in \mathcal{W}^h_c. \tag{46}$$

Following [39], the problem is expressed in its equivalent symmetric form by defining the stochastic multiscale velocity space, $\mathcal{V}^h_{ms}$ and stochastic multiscale pressure spaces $\mathcal{W}^h_{ms}$ as follows:

$$\mathcal{V}^h_{ms} = \mathcal{S} \otimes V^h_{ms}, \quad V^h_{ms} = \left\{ \boldsymbol{u}_{ms} : \boldsymbol{u}_{ms} = \sum_{a=1}^{M_c} \boldsymbol{N}^{ms}_a u^c_a, u^c_a = 0 \quad \forall \Lambda_a \in \partial\mathcal{D}_u \right\}, \tag{47}$$

$$\mathcal{W}^h_{ms} = \mathcal{S} \otimes W^h_{ms}, \quad W^h_{ms} = \left\{ w_{ms} : w_{ms} = \sum_{i=1}^{N_c} \phi^c_i \left(w^c_i + \sum_a \tilde{\phi}^{ms}_a u^c_a \right) \right\}, \tag{48}$$

where $\tilde{\phi}^{ms}_a = \phi^{ms}_a - \frac{1}{|E_i|} \int_{E_i} \phi^{ms}_a d\boldsymbol{x}$.

With this definition of the finite-dimensional function spaces, the stochastic coarse-scale problem can be written as: Find $(\boldsymbol{u}_{ms}, p_{ms}) \in \mathcal{V}^h_{ms} \times \mathcal{W}^h_{ms}$, such that

$$(\boldsymbol{v}_{ms}, k^{-1}\boldsymbol{u}_{ms}) - (\nabla \cdot \boldsymbol{v}_{ms}, p_{ms}) = -\langle \boldsymbol{v}_{ms} \cdot \boldsymbol{n}, p_o \rangle, \quad \forall \boldsymbol{v}_{ms} \in \mathcal{V}^h_{ms}, \tag{49}$$

$$(w_{ms}, \nabla \cdot \boldsymbol{u}_{ms}) = (w_{ms}, f), \quad \forall w_c \in \mathcal{W}^h_c. \tag{50}$$

Once the coarse-scale variables are solved, the multiscale velocity and pressure can be reconstructed (if necessary) by resorting to the additive decomposition Eq. (16). The subgrid part is obtained by a linear combination of the coarse-scale fluxes and the subgrid-scale basis functions.

3.7. The stochastic multiscale framework

The abstract framework to solve the stochastic multiscale problem defined by Eqs. (4) and (5) is given in Fig. 2. As seen in Fig. 2, the key steps involved are

- From limited data determine the multiscale stochastic permeability. This requires data-driven strategies to construct viable finite-dimensional representation of the stochastic permeability. Different techniques are discussed in Section 4.
- Find the stochastic multiscale basis functions for each coarse element interface by solving Eqs. (42)–(44).
- Use these multiscale basis functions to solve for the stochastic coarse-scale velocities and pressure given by Eqs. (49) and (50). Adaptive techniques to solve the above set of SPDEs are discussed in Section 5.

4. Representing the multiscale stochastic permeability: Constructing low-dimensional models

The first step towards solving the SPDEs is to obtain some form of numerical representation of the input random processes – i.e. the stochastic permeability, $k(\boldsymbol{x}, \omega)$. A finite-dimensional representation of the abstract probability space is a necessary prerequisite for developing any viable framework to solve Eqs. (4) and (5).

4.1. The Karhunen-Loève expansion

In many situations, extensive experimental studies are available and semi-variograms of permeability distribution are constructed [17,47]. Furthermore, some notion of the mean permeability variation is also available. The existence of some correlation structure of the permeability field provides an elegant way of representing it as a finite set of uncorrelated random variables. The Karhunen-Loève expansion can be used to convert these experimental statistics into a viable low-dimensional stochastic model of the permeability variation.

The KL expansion can be written as

$$k(\boldsymbol{x}, \omega) = E(k(\boldsymbol{x}, \omega)) + \sum_{i=0}^{\infty} \sqrt{\lambda_i} f_i(\boldsymbol{x}) Y_i(\omega), \tag{51}$$

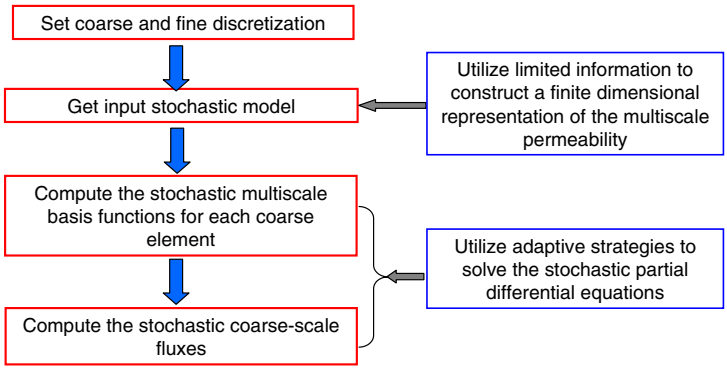


Fig. 2. The stochastic multiscale framework.

where $E(k(\boldsymbol{x},\omega))$ denotes the mean of the process and $\{Y_i(\omega)\}_{i=0}^{\infty}$ forms a set of uncorrelated random variables whose distribution has to be determined [48]. $(\lambda_i, f_i(\boldsymbol{x}))_{i=0}^{\infty}$ form the eigenpairs of the covariance function:

$$\int_{\mathcal{D}} R_{hh}(\boldsymbol{x}_1,\boldsymbol{x}_2) f_i(\boldsymbol{x}_2) dy_2 = \lambda f_1(\boldsymbol{x}_1). \tag{52}$$

The chief characteristic of the KLE is that the spatial randomness has been decomposed into a set of deterministic functions multiplying random variables. These deterministic functions can also be thought of as representing the scales of fluctuations of the process. The KLE is mean-square convergent to the original process $k(\boldsymbol{x},\omega)$. More interestingly, the first few terms of this expansion represent most of the process with arbitrary accuracy. The expansion in Eq. (51) is typically truncated to a finite number of summation terms, N.

4.2. Non-KLE based data-driven methods

As long as variograms and other correlation functions (which can be expanded via KLE) are available, using the KLE provides the simplest way of constructing a finite-dimensional descriptor of the stochastic permeability. In cases where correlation statistics vary spatially across the domain (or when the KLE cannot be used to generate a descriptor), one can utilize geostatistical methods [17,47] to construct plausible realizations of the permeability. Each plausible permeability distribution is given at the nodal points of the full-scale discretization of the domain. Denote such permeability realizations as $\boldsymbol{k}^{\mathcal{D}} = \{k(x_1),\ldots,k(x_{N_h})\}$, where N_h is the number of nodal points in the full-scale discretization of the domain. $\boldsymbol{k}^{\mathcal{D}}$ belongs to Ω, the set of plausible permeability distributions:

$$\Omega = \{\boldsymbol{k}^{\mathcal{D}} \in \mathbb{R}^{N_h} | \boldsymbol{k}^{\mathcal{D}} \text{satisfies the experimental statistics } S = \{S_1,\ldots,S_p\}\}. \tag{53}$$

Data-driven strategies can be utilized to incorporate these plausible realizations into a global descriptor of the stochastic permeability.[4] Depending on the number of such realizations (which is determined by accuracy and computational complexity considerations), two data-driven techniques for representing the stochastic permeability can be utilized. If the number of plausible data sets is limited ($\sim$10–100), it is possible to use linear model reduction ideas (i.e. Principal Component Analysis (PCA), Proper Orthogonal Decomposition (POD)) to construct a low-dimensional stochastic input model for the permeability variation. On the other hand, when the number of realizations available is fairly large, one can utilize non-linear model reduction strategies (manifold learning algorithms) to construct low-dimensional representations.

4.2.1. Linear model reduction strategies

Using geostatistical methods [17,47], it is possible to construct a finite set, $\boldsymbol{k}_1^{\mathcal{D}},\ldots,\boldsymbol{k}_M^{\mathcal{D}}$, of plausible permeability distributions in the domain that satisfy the experimental data. Note that we denote each realization of the permeability in the domain $\boldsymbol{k}_i^{\mathcal{D}}$ as a vector of length N_h. If the number, M, of such reconstructed vectors is not large ($M <\sim 100$), linear model reduction strategies like PCA and POD can be utilized to encode the information contained in the M plausible snapshots of the permeability distribution.

The average permeability vector is computed as $\langle \boldsymbol{k}^{\mathcal{D}} \rangle = \frac{1}{M}\sum_{i=1}^{M} \boldsymbol{k}_i^{\mathcal{D}}$. The average permeability vector, $\langle \boldsymbol{k}^{\mathcal{D}} \rangle$ is then subtracted from all the permeability vectors as $\boldsymbol{k}_i^{\mathcal{D}} \leftarrow \boldsymbol{k}_i^{\mathcal{D}} - \langle \boldsymbol{k}^{\mathcal{D}} \rangle$ for $i = 1,\ldots,M$. The eigenvectors $U^{(k)}$ of the $N_h \times N_h$ covariance matrix $C = \frac{1}{M}\sum_{i=1}^{M} \boldsymbol{k}_i^{\mathcal{D}} \boldsymbol{k}_i^{\mathcal{D}}$ satisfying the equation

$$CU^{(k)} = \lambda_k U^{(k)}, \quad k = 1,\ldots,N_h, \tag{54}$$

along with the eigenvalues λ_k are computed. This set of eigen-images and eigenvalues form the best basis to represent the permeability vector [49,50]. The first N eigen-images (usually $N \ll M$) representing most of the energy spectrum of the decomposition is chosen. Any random (plausible) permeability vector (I) belonging to the space of allowable permeability distributions Ω can then be represented as a unique linear combination of the N eigen-images:

$$I = \langle \boldsymbol{k}^{\mathcal{D}} \rangle + \sum_{i=1}^{N} Y_i U^{(i)}, \quad I \in \Omega, \quad \boldsymbol{Y} \equiv (Y_1,\ldots,Y_N) \in \mathbb{R}^N. \tag{55}$$

This transformation serves as the low-dimensional representation that can be used as a stochastic input model for the permeability. Note, however, that not every $\boldsymbol{Y} \in \mathbb{R}^N$ results in a plausible permeability distribution [33] (i.e. the transformation is not surjective). This can be accounted for by reducing the domain of variability of $\boldsymbol{Y}$ to make the transformation surjective. Such 'subspace reducing' methods are detailed in [33].

4.2.2. Non-linear model reduction strategies

The PCA based model reduction scheme constructs the *closest linear subspace* of the high-dimensional input space. This is a fairly good approximation when the number of data sets is limited. But as the amount of data available increases, PCA based techniques tend to consistently over-estimate the actual dimensionality of the space [34]. This is primarily due to

[4] Another way is to first build a facies model [32] and then perform a KLE within each facie. This could however enormously increase the total (stochastic) dimensionality of the problem.

the fact that the space of all plausible permeability distributions (Ω) is a non-linear space. Non-linear transformation strategies offer the possibility of constructing optimal low-dimensional representations of this space. As before, assume that M plausible permeability vectors are available. Using the unordered data $\{\boldsymbol{k}_i^{\mathcal{D}}\}_{i=1}^{M}$, the problem of interest is to find a low-dimensional parametrization of Ω, i.e. a set $\Gamma \in \mathbb{R}^N$, $N \ll N_h$, such that there is a one-to-one correspondence between Ω and Γ.

The solution strategy is based on the so-called principle of 'manifold learning' [34]. The basic strategy is to show that this set of unordered points lies on a manifold embedded in a high-dimensional space. That is, Ω is a manifold embedded in a high-dimensional space. The mathematical framework is then to 'unravel and smoothen' this manifold and represent it as a smooth low-dimensional curve, Γ. *This 'unravelling and smoothing' corresponds to a topological transformation that preserves some notion of the geometry of the manifold*. By keeping specific geometrical features of the manifold invariant one can construct a low-dimensional representation that is equivalent to the manifold. A natural choice of a geometric feature is the distance metric. This results in an isometric mapping to transform Ω into Γ. The important idea is that the distance that encodes the geometric information about the non-linear manifold in the *geodesic distance*. The geodesic distance reflects the true geometry of the manifold embedded in the high-dimensional space.

Construction of Γ reduces to finding a low-dimensional representation, $\{\boldsymbol{Y}_i\}$ of the given data points $\boldsymbol{k}_1^{\mathcal{D}}, \ldots, \boldsymbol{k}_M^{\mathcal{D}}$ such that $\{\boldsymbol{Y}_i\}$ is isometric to $\boldsymbol{k}_1^{\mathcal{D}}, \ldots, \boldsymbol{k}_M^{\mathcal{D}}$ based on the geodesic distances between the points. The geodesic distance is approximately computed via the concept of graph distance. The unknown geodesic distances in Ω between the data points are computed in terms of a *graph distance* with respect to a neighborhood graph G constructed on the data points [34,51]. This data is compactly stored as a matrix (denoted as $\mathbf{M}$). The estimation of the low-dimensional representation of $\{\boldsymbol{k}_1^{\mathcal{D}}, \ldots, \boldsymbol{k}_M^{\mathcal{D}}\}$ can now be posed as:

Find a configuration of points $\{\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_M\}$, $\boldsymbol{Y}_i \in \mathbb{R}^N$ *such that these points yield a Euclidean distance matrix whose elements are identical to the elements of the geodesic distance matrix* $\mathbf{M}$ (with $N \ll N_h$). That is, find $\{\boldsymbol{Y}_i\}_{i=1}^{M}$ such that $\|\boldsymbol{Y}_i - \boldsymbol{Y}_j\| \approx M_{ij}$. The principle of Multi-dimensional scaling (MDS) can subsequently be used to compute the set of low-dimensional points that best represent the high-dimensional points [52,53]. The MDS procedure essentially computes the eigen-decomposition of the geodesic matrix and sets the low-dimensional points as linear combinations of the largest N eigenvectors of the geodesic matrix. This provides a natural of constructing Γ from the M low-dimensional points $\{\boldsymbol{Y}_i\}_{i=1}^{M}$ [34]:

$$\Gamma \equiv \{\boldsymbol{Y} \in \mathbb{R}^N | \boldsymbol{Y} \in \text{Convex hull } (\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_M)\}. \tag{56}$$

The intrinsic dimensionality N of the low-dimensional representation can be estimated by using a variant of the Breadwood–Halton–Hammersley [54] theorem where N is linked to the rate of convergence of the length functional of the minimal spanning tree of the geodesic distance matrix of the unordered data points in the high-dimensional space [55–57]. For an arbitrary point $\boldsymbol{Y} \in \Gamma \subset \mathbb{R}^N$, the corresponding permeability distribution $k(\boldsymbol{Y})$ is constructed as a linear combination of the nearest N input data points, i.e.

$$k(\boldsymbol{Y}) = \sum_{i=1}^{N} w_i k(\boldsymbol{Y}_i), \tag{57}$$

where $\{\boldsymbol{Y}_i\}_{i=1}^{N}$ are the N closest points to $\boldsymbol{Y}$ from the set of M points $\{\boldsymbol{Y}_i\}_{i=1}^{M}$, and the weights w_i are inversely proportional to the distance of $\boldsymbol{Y}$ from each point $\boldsymbol{Y}_i$ as $w_i = w/||Y - Y_i||$ with $1/w = \sum_{i=1}^{N} 1/||\boldsymbol{Y} - \boldsymbol{Y}_i||$.

5. Solving stochastic partial differential equations

The above generated N-dimensional representation of the stochastic multiscale permeability is utilized as an input stochastic model for the solution of SPDEs Eqs. (7) and (8). The stochastic input model is utilized in the solution of the stochastic multiscale basis functions (defined by Eqs. (42)–(44)) which are the inputs to solve the stochastic coarse-scale variables Eqs. (49) and (50).

We utilize an adaptive sparse grid collocation strategy for constructing the stochastic solution [25]. We briefly describe the development of the adaptive sparse grid collocation strategy here. The interested reader is referred to our recent work in [28].

The basic idea of the stochastic collocation method is to approximate the stochastic space using multi-dimensional interpolating functions. The method uses *realizations* of the function (i.e. the solution of the SPDE, $(\boldsymbol{u}(.,\boldsymbol{Y}_i), p(.,\boldsymbol{Y}_i))$) at a finite set of collocation points $\{\boldsymbol{Y}_i\}_{i=1}^{M} \in \Gamma$. These finite number of deterministic solutions are used in constructing an interpolant of the dependent stochastic variables $(\boldsymbol{u}(.,\boldsymbol{Y}), p(.,\boldsymbol{Y}))$ by using linear combinations of the solutions $u(\cdot,\boldsymbol{Y}_i)$. The two key question issues to be resolved are (a) Find the optimal points to sample this multi-dimensional space, and (b) The mathematical framework to construct the adaptive multi-dimensional interpolation once the sampling is performed.

The choice of the optimal interpolating sampling is a well studied problem [58]. For a one-dimensional function, Clenshaw-Curtis points at the non-equidistant extrema of the Chebyshev polynomials [26,59] as well as the Newton-Cotes formulae using equidistant support nodes have been shown to be optimal [28]. It is usually advantageous to choose the collocation points in a nested fashion to obtain many recurring points with increasing order of interpolation [28]. Having chosen the optimal set of points in one dimension, one can construct the interpolant approximation to a one-dimensional function f as:

$$\mathcal{U}^i(f) = \sum_{j=1}^{m_i} f(Y_j^i) \cdot a_j^i, \tag{58}$$

with the set of support nodes $X^i = \{Y_j^i | Y_j^i \in [0,1] \text{ for } j = 1,2,\ldots,m_i\}$, where $i \in \mathbb{N}$, $a_j^i \equiv a_j(Y_j^i) \in C([0,1])$ are the interpolation nodal basis functions, and m_i is the number of elements of the set X^i. The nodal basis functions are usually Lagrange polynomials [27,25]. The multi-dimensional interpolation function can then be constructed by using full tensor product of the corresponding 1D interpolation rule.

$$(\mathcal{U}^{i_1} \otimes \cdots \otimes \mathcal{U}^{i_N})(f) = \sum_{j_1=1}^{m_1} \cdots \sum_{j_N=1}^{m_N} f(Y_{j_1}^{i_1}, \ldots, Y_{j_N}^{i_N}) \cdot (a_{j_1}^{i_1} \otimes \cdots \otimes a_{j_N}^{i_N}). \tag{59}$$

Obviously, the number of support points grows very quickly as the number of stochastic dimensions increases in the full tensor product case. This resulted in the development of the sparse grid interpolation method based on the Smolyak algorithm [60].

5.1. Sparse grid interpolation

Using the Smolyak algorithm [60], univariate interpolation formulae are extended to the multivariate case by using tensor products in a special way – providing an interpolation strategy with potentially orders of magnitude reduction in the number of support nodes required. The Smolyak algorithm constructs the sparse interpolant, $A_{q,N}$, using products of 1D functions.

Consider the incremental interpolant, Δ^i given by [28,25]

$$\mathcal{U}^0 = 0, \quad \Delta^i = \mathcal{U}^i - \mathcal{U}^{i-1}. \tag{60}$$

The Smolyak interpolation $\mathcal{A}_{q,N}$ is then given by

$$\mathcal{A}_{q,N}(f) = \sum_{|\mathbf{i}| \leqslant q} (\Delta^{i_1} \otimes \cdots \otimes \Delta^{i_N})(f) = \mathcal{A}_{q-1,N}(f) + \sum_{|\mathbf{i}|=q} (\Delta^{i_1} \otimes \cdots \otimes \Delta^{i_N})(f) = \mathcal{A}_{q-1,N}(f) + \Delta\mathcal{A}_{q,N}(f), \tag{61}$$

with $q \geqslant N$, $\mathcal{A}_{N-1,N} = 0$ and where the multi-index $\mathbf{i} = (i_1,\ldots,i_N) \in \mathbb{N}^N$ and $|\mathbf{i}| = i_1 + \cdots + i_N$. Here i_k, $k = 1,\ldots,N$, can be thought of as the level of interpolation along the kth direction. *The Smolyak algorithm essentially builds the interpolation function by adding a combination of 1D functions of order i_k with the constraint that the sum total ($|\mathbf{i}| = i_1 + \ldots + i_N$) across all dimensions is less than q*. The construction of the algorithm allows one to utilize all the previous results generated to improve the interpolation (this is immediately obvious from Eq. (61)). By choosing appropriate points for interpolating the 1D function, one can ensure that the sets of points X^i are nested ($X^i \subset X^{i+1}$).

In the context of incorporating adaptivity, we have utilized the Newton-Cotes grid using equidistant support nodes [28]. By using equidistant nodes, it is easy to refine the grid locally. Furthermore, by using the linear hat function as the univariate nodal basis function [61] one ensures a local support in contrast to the global support of using Lagrange polynomial (Eq. (58)). This ensures that discontinuities in the stochastic space can be resolved.

5.2. From nodal basis to hierarchical basis

The important step to naturally incorporate adaptivity in the sparse grid collocation framework is to move from a nodal basis definition of the interpolation formulae to a hierarchical basis definition of the interpolation functions [28]. We start from the 1D interpolating formula Eq. (58) using nodal basis as discussed in the previous section. By the definition of Eq. (60), we have $\Delta^i(f) = \mathcal{U}^i(f) - \mathcal{U}^{i-1}(f)$, with $\mathcal{U}^i(f) = \sum_{Y_j^i \in X^i} a_j^i \cdot f(Y_j^i)$, and $\mathcal{U}^{i-1}(\mathrm{f}) = \mathcal{U}^i(\mathcal{U}^{i-1}(\mathrm{f}))$, we obtain [28]

$$\Delta^i(f) = \sum_{Y_j^i \in X^i} a_j^i \cdot f(Y_j^i) - \sum_{Y_j^i \in X^i} a_j^i \cdot \mathcal{U}^{i-1}(f)(Y_j^i) = \sum_{Y_j^i \in X^i} a_j^i \cdot (f(Y_j^i) - \mathcal{U}^{i-1}(f)(Y_j^i)), \tag{62}$$

and, since $f(Y_j^i) - \mathcal{U}^{i-1}(f)(Y_j^i) = 0, \forall Y_j^i \in X^{i-1}$, we obtain

$$\Delta^i(f) = \sum_{Y_j^i \in X_\Delta^i} a_j^i \cdot (f(Y_j^i) - \mathcal{U}^{i-1}(f)(Y_j^i)), \tag{63}$$

where $X_\Delta^i = X^i \setminus X^{i-1}$. Clearly, X_Δ^i has $m_\Delta^i = m_i - m_{i-1}$ points, since $X_{i-1} \subset X_i$. The above equation is rewritten as [28]

$$\Delta^i(f) = \sum_{j=1}^{m_\Delta^i} a_j^i \cdot \underbrace{(f(Y_j^i) - \mathcal{U}^{i-1}(f)(Y_j^i))}_{w_j^i}. \tag{64}$$

Here, we define w_j^i as the 1D hierarchical surplus, *which is just the difference between the function value at the current level and the previous level*. We also define the set of functions a_j^i as the hierarchical basis functions. Now we apply the 1D Eq. (64) to obtain the sparse grid interpolation formula for the multivariate case in a hierarchical form.

$$\Delta\mathcal{A}_{q,N}(f) = \sum_{|\mathbf{i}|=q}\sum_{\mathbf{j}\in B_{\mathbf{i}}}(a_{j_1}^{i_1}\otimes\cdots\otimes a_{j_N}^{i_N})\cdot(f(Y_{j_1}^{i_1},\ldots,Y_{j_N}^{i_N}) - \mathcal{A}_{q-1,N}(f)(Y_{j_1}^{i_1},\ldots,Y_{j_N}^{i_N})). \tag{65}$$

where the multi-index set $B_{\mathbf{i}} := \{\mathbf{j}\in\mathbb{N}^N : Y_{j_k}^{i_k}\in X_\Delta^{i_k} \text{ for } j_k = 1,\ldots,m_\Delta^{i_k}, k = 1,\ldots,N\}$ and we define

$$w_{\mathbf{j}}^{\mathbf{i}} = f(Y_{j_1}^{i_1},\ldots,Y_{j_N}^{i_N}) - \mathcal{A}_{|\mathbf{i}|-1,N}(f)(Y_{j_1}^{i_1},\ldots,Y_{j_N}^{i_N}), \tag{66}$$

as the hierarchical surplus, which is just the difference between the function value at the current point and interpolation value from the coarser grid. As described in [28], we can work either in the nodal basis functional space or the hierarchical basis space. For smooth functions, the hierarchical surpluses tend to zero as the interpolation level tends to infinity. On the other hand, for non-smooth functions, steep gradients/finite discontinuities are indicated by the magnitude of the hierarchical surplus. The bigger the magnitude is, the stronger the underlying discontinuity is. Therefore, the hierarchical surplus is a natural candidate for error control and implementation of adaptivity.

5.3. Adaptive sparse grid interpolation

Following [28], we set some notation first. The $1D$ equidistant points of the sparse grid can be considered as a tree-like data structure as shown in Fig. 3. We can consider the interpolation level of a grid point Y as the depth of the tree $D(Y)$. Denote the father of a grid point as $F(Y)$, where the father of the root 0.5 is itself, i.e., $F(0.5) = 0.5$.

The conventional sparse grid in the N-dimensional space can be reconsidered as

$$\mathcal{H}_{q,N} = \left\{\mathbf{Y} = \{Y_1,\ldots,Y_N\} | \sum_{i=1}^{N} D(Y_i) \leqslant q\right\}. \tag{67}$$

We denote the sons of a grid point $\mathbf{Y} = (Y_1,\ldots,Y_N)$ by

$$\text{Sons}(\mathbf{Y}) = \{\mathbf{S} = (S_1,S_2,\ldots,S_N) | (F(S_1),S_2,\ldots,S_N) = \mathbf{Y}, \quad \text{or} \quad (S_1,F(S_2),\ldots,S_N) = \mathbf{Y},\ldots,(S_1,S_2,\ldots,F(S_N)) = \mathbf{Y}\}. \tag{68}$$

From this definition, it is noted that, in general, for each grid point there are two sons in each dimension, therefore, for a grid point in a N-dimensional stochastic space, there are $2N$ sons. It is also noted that, the sons are also the *neighbor points* of the father. The neighbor points are just the support nodes of the hierarchical basis functions in the next interpolation level [28]. By adding the neighbor points, we actually add the support nodes from the next interpolation level, i.e., we perform interpolation from level $|\mathbf{i}|$ to level $|\mathbf{i}|+1$. Therefore, in this way, we refine the grid locally while not violating the developments of the Smolyak algorithm Eq. (65).

The basic idea here is to use hierarchical surpluses as an error indicator to detect the smoothness of the solution and refine the hierarchical basis functions $a_{\mathbf{j}}^{\mathbf{i}}$ whose magnitude of the hierarchical surplus satisfies $|w_{\mathbf{j}}^{\mathbf{i}}| \geqslant \varepsilon$. If this criterion is satisfied, we simply add the $2N$ neighbor points of the current point from Eq. (68) to the sparse grid. It is noted that the growth of the points scales linearly with increasing dimensionality rather than the $O(2^N)$ tree-like scaling of the standard h-type adaptive refinement as in a random element-based framework, e.g. ME-gPC. Let $\varepsilon > 0$ be the parameter for the adaptive refinement threshold. We propose the following iterative refinement algorithm beginning with a coarsest adaptive sparse grid $\mathcal{G}_{N,N}$, i.e., we begin with the N-dimensional multi-index $\mathbf{i} = (1,\ldots,1)$, which is just a point $(0.5,\ldots,0.5)$.

(1) Set level of Smolyak construction $k = 0$.
(2) Construct the first level adaptive sparse grid $\mathcal{G}_{N,N}$.
- Calculate the function value at the point $(0.5,\ldots,0.5)$;
- Generate the $2N$ neighbor points and add them to the active index set;
- Set $k = k+1$.

(3) While $k \leqslant k_{max}$ and the active index set is not empty:
- Copy the points in the active index set to an old index set and clear the active index set.
- Calculate in parallel the hierarchical surplus of each point in the old index set according to

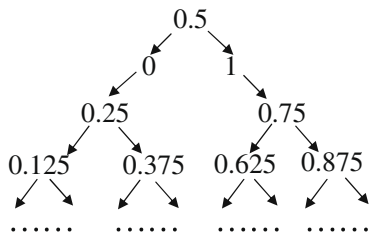


Fig. 3. $1D$ tree-like structure of the sparse grid.

$$w_{\mathbf{j}}^{\mathbf{i}} = f(Y_{j_1}^{i_1}, \ldots, Y_{j_N}^{i_N}) - \mathcal{G}_{N+k-1,N}(f)(Y_{j_1}^{i_1}, \ldots, Y_{j_N}^{i_N}). \tag{69}$$

Here, we use all of the existing collocation points in the current adaptive sparse grid $\mathcal{G}_{N+k-1,N}$. This allows us to evaluate the surplus for each point from the old index set in parallel.

- For each point in the old index set, if $|w_{\mathbf{j}}^{\mathbf{i}}| \geqslant \varepsilon$.
  - Generate $2N$ neighbor points of the current active point according to Eq. (68);
  - Add them to the active index set.

- Add the points in the old index set to the existing adaptive sparse grid $\mathcal{G}_{N+k-1,N}$. Now the adaptive sparse grid becomes $\mathcal{G}_{N+k,N}$.
- $k = k + 1$.

(4) Calculate the mean and the variance, the PDF and if needed realizations of the solution (see Section 5.4).

5.4. Post-processing operations

Any function $u(.,\mathbf{Y}) \in \Gamma$ can now be approximated by the following reduced form from Eq. (65):

$$u(\boldsymbol{x},\mathbf{Y}) = \sum_{|\mathbf{i}| \leqslant q} \sum_{\mathbf{j} \in B_{\mathbf{i}}} w_{\mathbf{j}}^{\mathbf{i}}(\boldsymbol{x}) \cdot a_{\mathbf{j}}^{\mathbf{i}}(\mathbf{Y}). \tag{70}$$

This is just a simple weighted sum of the value of the basis functions for all collocation points in the current sparse grid [28]. After obtaining the expression in Eq. (70), it is also easy to extract statistics [28]. The mean of the random solution can be evaluated as follows:

$$\mathbb{E}[u(\boldsymbol{x})] = \sum_{|\mathbf{i}| \leqslant q} \sum_{\mathbf{j} \in B_{\mathbf{i}}} w_{\mathbf{j}}^{\mathbf{i}}(\boldsymbol{x}) \cdot \int_{\Gamma} a_{\mathbf{j}}^{\mathbf{i}}(\mathbf{Y}) d\mathbf{Y}, \tag{71}$$

where the probability density function $\rho(\mathbf{Y})$ is 1 since the stochastic space is a unit hypercube $[0,1]^N$. As shown in [28], the multi-dimensional integral is simply the product of the 1D integrals which can be computed analytically. Denoting $\int_{\Gamma} a_{\mathbf{j}}^{\mathbf{i}}(\mathbf{Y}) d\mathbf{Y} = I_{\mathbf{j}}^{\mathbf{i}}$, we can rewrite Eq. (71) as

$$\mathbb{E}[u(\boldsymbol{x})] = \sum_{|\mathbf{i}| \leqslant q} \sum_{\mathbf{j} \in B_{\mathbf{i}}} w_{\mathbf{j}}^{\mathbf{i}}(\boldsymbol{x}) \cdot I_{\mathbf{j}}^{\mathbf{i}}. \tag{72}$$

Similarly, higher-order statistics are simple weighted sums of powers of the hierarchical surpluses.

6. The complete algorithm

The complete schematic of the stochastic multiscale procedure is illustrated in Fig. 4. The data-driven strategies (discussed in Section 4) convert the limited information available into a plausible, realistic stochastic representation of the multiscale permeability variation. The adaptive sparse grid collocation strategy is used to construct the stochastic multiscale basis functions over all the coarse elements. Following this, the stochastic coarse-scale equations are solved for the coarse-scale stochastic pressure and velocity.

7. Numerical examples

In this section, we apply the complete stochastic analysis framework- from data-driven model generation, the stochastic variational mixed multiscale formulation to the adaptive solution of the resulting SPDEs. In the first example, we compare the effect of uncertainties in the small scale permeability variations and the effect of uncertainties in large scale geological features on the flow characteristics in a domain. In the second example, we look at the effect of localized uncertainties in permeability and how they propagate into the complete domain.

7.1. Effect of uncertainties at different scales

The schematic of the domain of interest is shown in Fig. 5. The region is part of an outcrop (where extensive measurements of permeability have been performed [62]) in Lawyer Canyon in Texas. The domain of interest is a square of length 200 ft. Flow is driven by an injection well at the left bottom corner of the domain and a production well at the top right corner. There is a low-permeability fault running across the domain. Limited information about the spatial variation in the multiscale permeability is available in the form of semi-variograms and mean permeability. Furthermore, the exact location and characteristics of the fault are unknown. Hence both the multiscale permeability variation as well as the large scale fault features have to be considered to be stochastic.

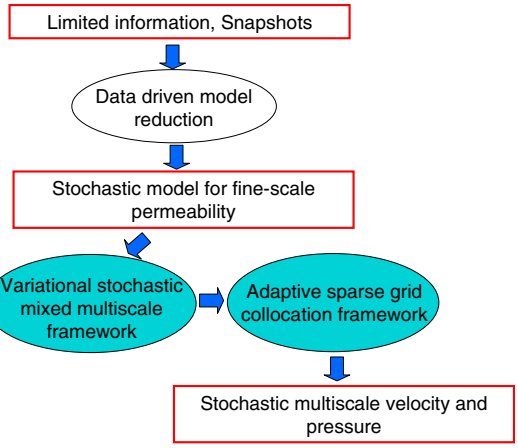


Fig. 4. Schematic of the developed stochastic multiscale framework.

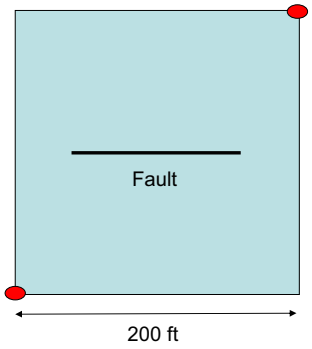


Fig. 5. Schematic of the domain. A fault runs across the middle of the domain.

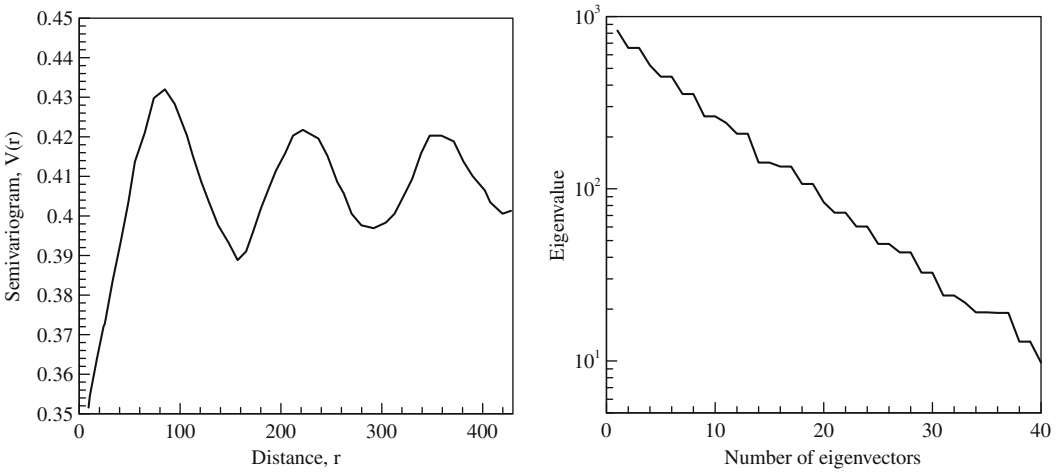


Fig. 6. (Left) Experimental semi-variogram of permeability (taken from [62]). (Right) The KL expansion of the log-permeability using the semi-variogram data.

7.1.1. Effect of stochastic multiscale permeability

We first investigate the effect of the stochastic multiscale permeability keeping the location of the fault fixed. The fault is assumed to originate at (40, 100) and have a length of 100 ft and a width of 20 ft. A reduced model for the permeability is constructed from the limited statistics available. Since in this case, the correlation structure of the permeability variation is

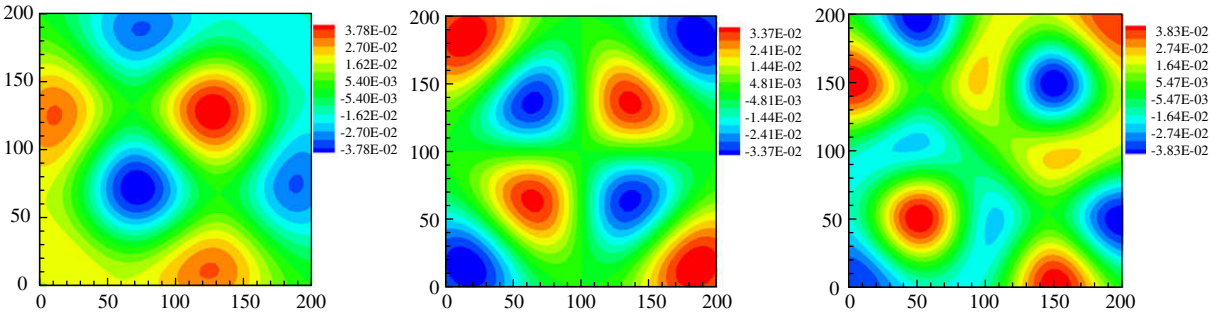


Fig. 7. Three eigenvectors constructed from the KL expansion.

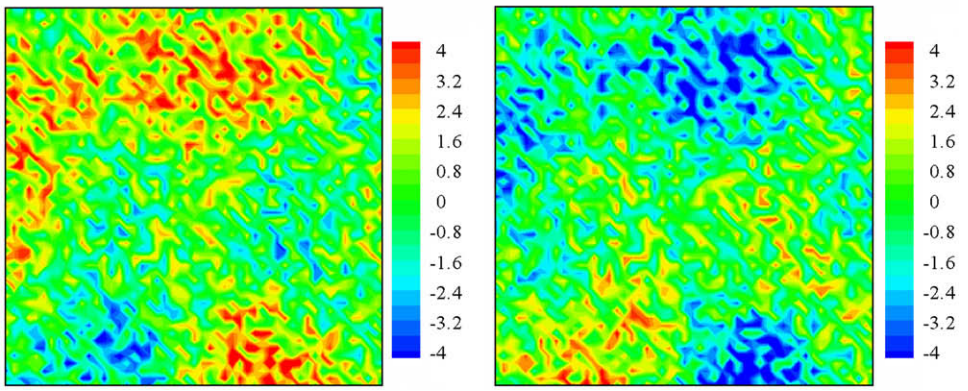


Fig. 8. Two extreme realizations of the log-permeability generated from the KL expansion.

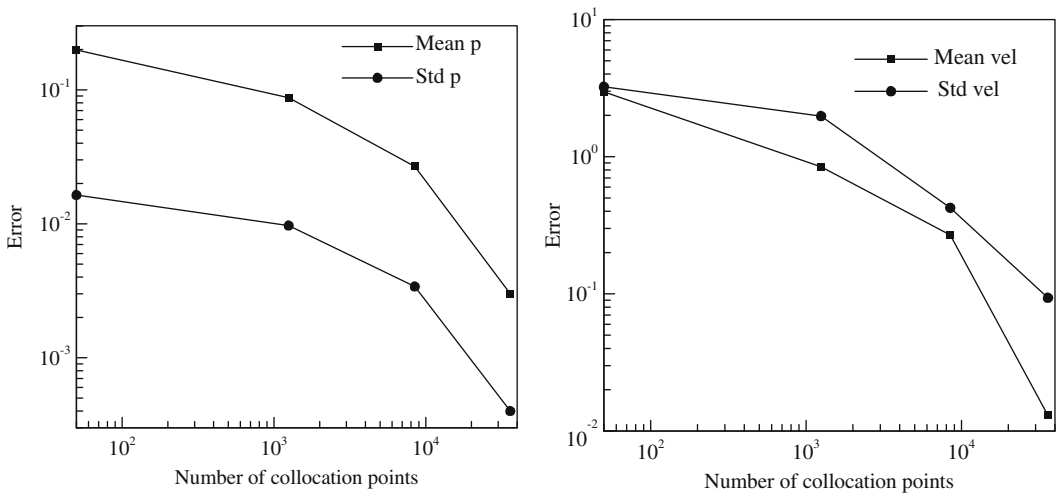


Fig. 9. Convergence of the stochastic solution (mean and standard deviation) with increasing number of collocation points. (Left) Stochastic coarse-scale pressure, (Right) stochastic coarse-scale *x*-direction flux.

available, we utilize the KL expansion to construct a viable stochastic model. The two-point correlation is extracted (Fig. 6 (left)) and the KL expansion is performed. Fig. 6 (right) plots the first 40 eigenvalues constructed from the KL expansion. Note the logarithmic scale used. The first 25 eigenvalues represent about 95% of the total information (the large number of stochastic dimensions required indicate that the correlation length is relatively small). The KL expansion is truncated to 25 terms and this generates a 25-dimensional input stochastic model for the permeability. The iid random variables in the

KL expansion (note that the KL expansion is for the log-permeability field) are standard normal appropriately scaled by the eigenvalues ($\sqrt{\lambda_i}\boldsymbol{y}_i$). The total standard deviation is 37.607. The spatial mean value of the log-permeability is -0.115, which gives a coefficient of variation of 1og k as $37.607/0.115 \sim 327$.

Three eigenvectors utilized in the KL expansion are shown in Fig. 7. Note that the eigenvectors are in fact eigenvectors of the log-permeability field.

Fig. 8 shows two extreme realizations of the log-permeability field constructed utilizing the KL model. Note the close to 7-orders of magnitude variation in the permeability variation in these realizations.

The multicale permeability is given as 100×100 fineblocks in the domain. Aggressive upgridding is utilized and a coarse-scale discretization of 10×10 is utilized in the stochastic variational mixed multiscale framework. The 100-fold coarse-graining results in the representation of the system by only 320 upscaled coarse variables.

Thirty nodes of our local Linux cluster (corresponding to 120 processors) was utilized. The total computational time was about 96 h to solve this $25 + 2$ (stochastic + spatial domain) dimensional problem. Fig. 9 (left) plots the convergence of the stochastic pressure (mean and standard deviation) with increasing number of collocation points, while Fig. 9 (right) plots the convergence of the stochastic x-direction flux with increasing number of collocation points. For the same depth of interpolation (depth of interpolation 8) about 1.3×10^9 collocation points would have been required using conventional sparse grid collocation.

Fig. 10 plots the mean values of the coarse-scale pressure and x-direction flux. Note that the pressure is stratified around the region where the permeability fault is located.

Fig. 11 plots the standard deviation of the stochastic coarse-scale pressure and x-direction flux. The low-permeability fault leads in preferential flow around the fault. This leads to stratification of pressure around the fault as well as negligible

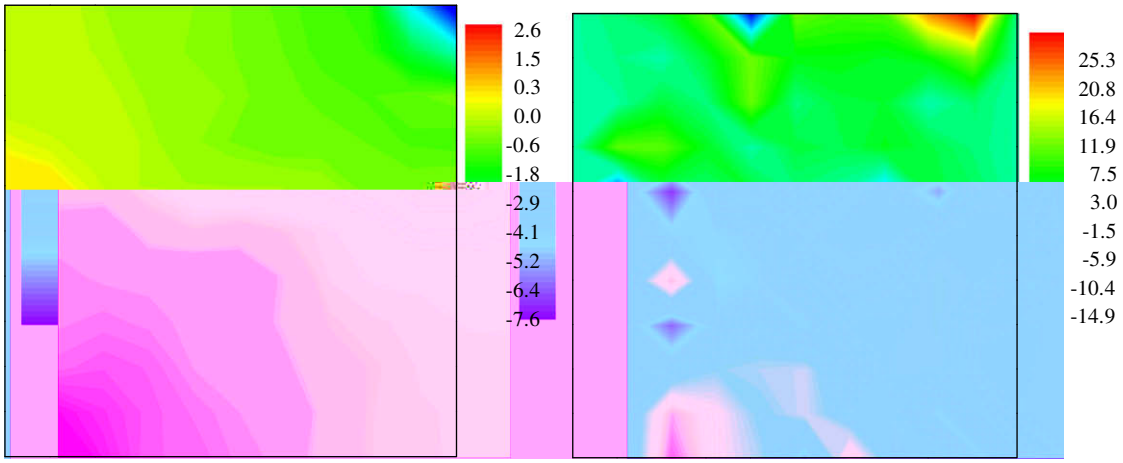


Fig. 10. Mean stochastic coarse-scale solution. (Left) Coarse-scale pressure, (Right) coarse-scale x-direction flux.

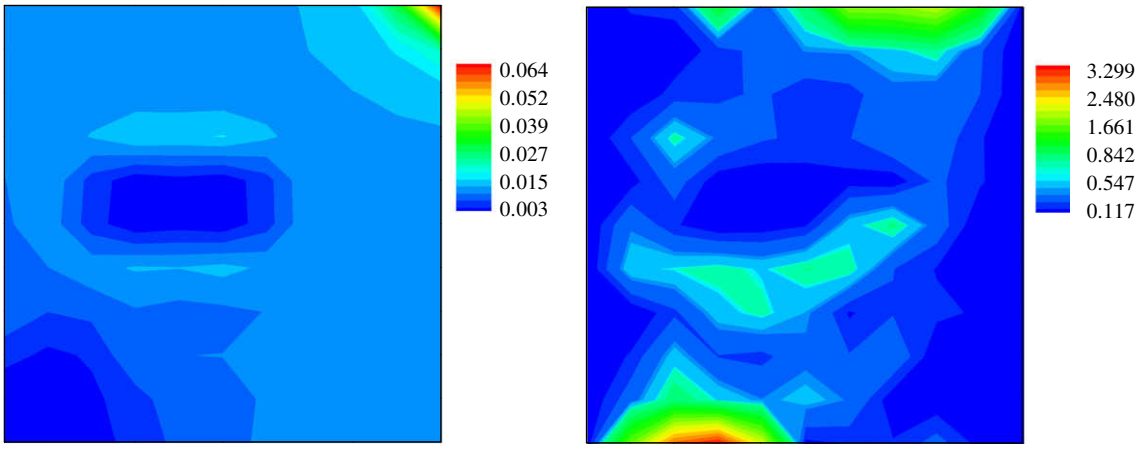


Fig. 11. Standard deviation of the stochastic coarse-scale solution. (Left) Coarse-scale pressure, (Right) coarse-scale x-direction flux.

flow. This is clearly seen as the very small standard deviation in pressure in the middle of the domain where the fault is located. Furthermore, this results in a large variation in flux around the permeability fault as can be seen in Fig. 11 (right).

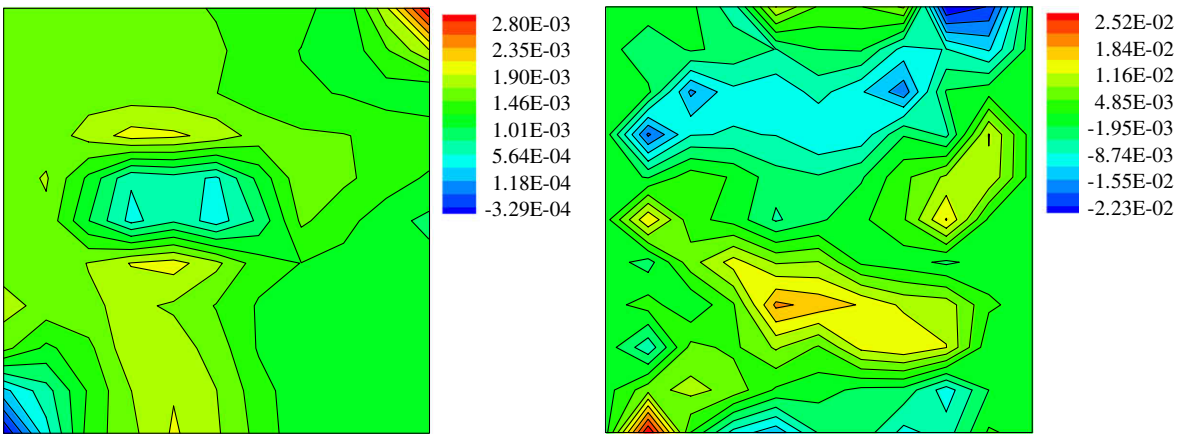


Fig. 12. Difference in the mean value of the stochastic solutions using the adaptive sparse grid strategy and Monte Carlo sampling. (Left) Mean pressure, (Right) mean x-direction flux.

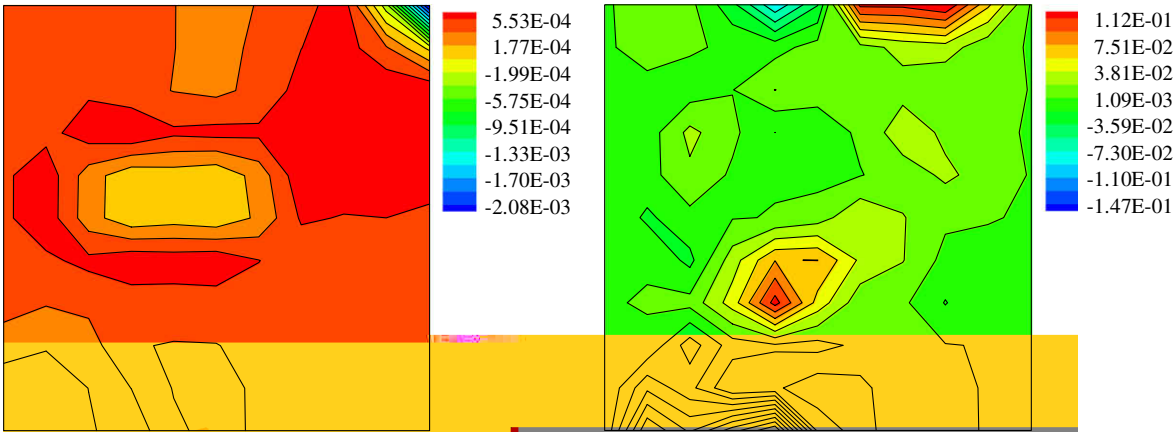
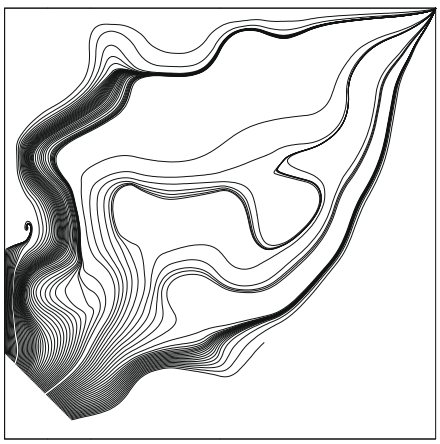


Fig. 13. Difference in the standard deviation of the stochastic solutions using the adaptive sparse grid strategy and Monte Carlo sampling. (Left) standard deviation in pressure, (Right) standard deviation in the x-direction flux.



To check the accuracy of the stochastic solution, we perform comparison with the stochastic solution constructed using Monte Carlo sampling strategies. Twenty-five-dimensional iid Gaussian random variables were generated. Each such 25-tuple was used in the KLE expansion representing the stochastic permeability (Fig. 6) to generate a realization of the permeability. 10^6 samples were utilized to compute the statistics using the Monte Carlo method. This took about 900 h using 30 nodes of our in-house Linux cluster. Fig. 12 plots the difference between the mean values of the coarse-scale pressure and x-direction flux computed from the two methods.

Notice the difference in the mean pressure is very close to zero close to the injection well as well as on the permeability fault. On the other hand, since the flow will preferentially try to avoid the fault (resulting in the largest deviation around this region), there will be high variability in the x-direction flux immediately above and below the permeability fault. This leads to the largest error occurring here as seen clearly in Fig. 12. This is also clearly illustrated in Fig. 13 which plots the difference between the standard deviation of the coarse-scale pressure and x-direction flux computed from the two methods.

Fig. 14 plots streamlines based on the mean coarse-scale velocity distribution in the domain. The streamlines preferentially circumvent the permeability barrier.

From the stochastic solution, the probability distribution function of the coarse-scale x-direction flux is plotted at two spatial locations (120, 40) and (180, 160). The PDFs are plotted in Fig. 15. PDFs of coarse-scale pressure (not shown) also exhibited similar trends. The stochastic permeability essentially caused a 'smearing' of the coarse-scale pressure and velocity fluxes.

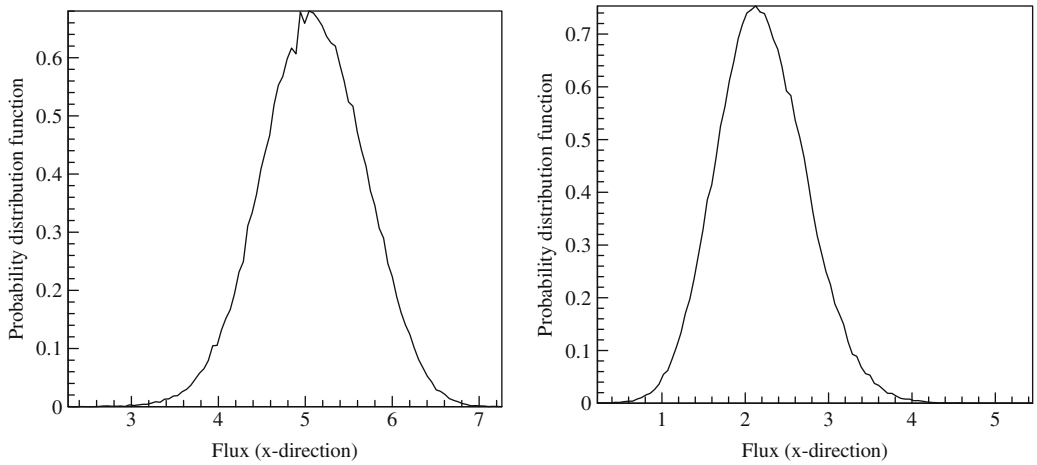


Fig. 15. Probability distribution functions of the coarse-scale flux (x-direction) for two points in the domain.

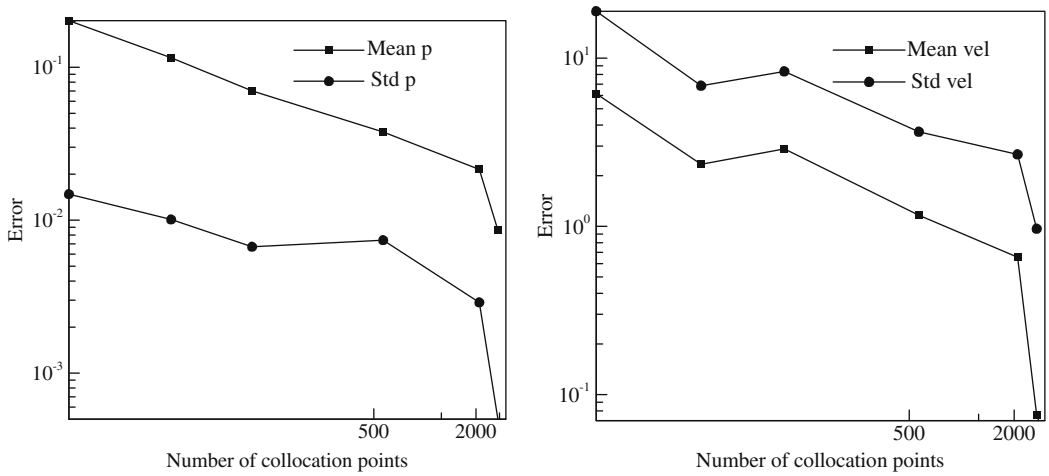


Fig. 16. Convergence of the stochastic solution (mean and standard deviation) with increasing number of collocation points. (Left) Stochastic coarse-scale pressure, (Right) stochastic coarse-scale x-direction flux.

7.1.2. Effect of uncertainty in large scale permeability barrier

Next we focus on the effects of uncertainty in the location of the permeability barrier on the coarse-scale stochastic pressure and velocity flux. We set the permeability variation to its mean value and utilize the same 100-fold coarse-grained discretization as in the previous problem. We assume that there is 10% uncertainty in the location, length and width of the permeability barrier. This is quantified by 4 independent uniform random variables, (Y_1, Y_2, Y_3, Y_4), on the domain $[-1, 1]$ as

$$
\begin{aligned}
x_l &= (0.2 + 0.05Y_1) * 200, \\
y_l &= (0.5 + 0.05Y_2) * 200, \\
L_{barrier} &= (0.5 + 0.05Y_3) * 200, \\
W_{barrier} &= (0.1 + 0.05Y_4) * 200,
\end{aligned}
$$

where (x_l, y_l) is the start position of the barrier, and $L_{barrier}$, $W_{barrier}$ are the length and width of the barrier, respectively. The variational stochastic mixed multiscale framework is utilized to solve this problem. 40 nodes of our local Linux cluster (corresponding to 160 processors) is utilized to solve the problem. The total computational time was about 240 min.

Fig. 16 (left) plots the convergence of the stochastic pressure with increasing number of collocation points, while Fig. 16 (right) plots the convergence of the stochastic x-direction flux with increasing number of collocation points. For the same

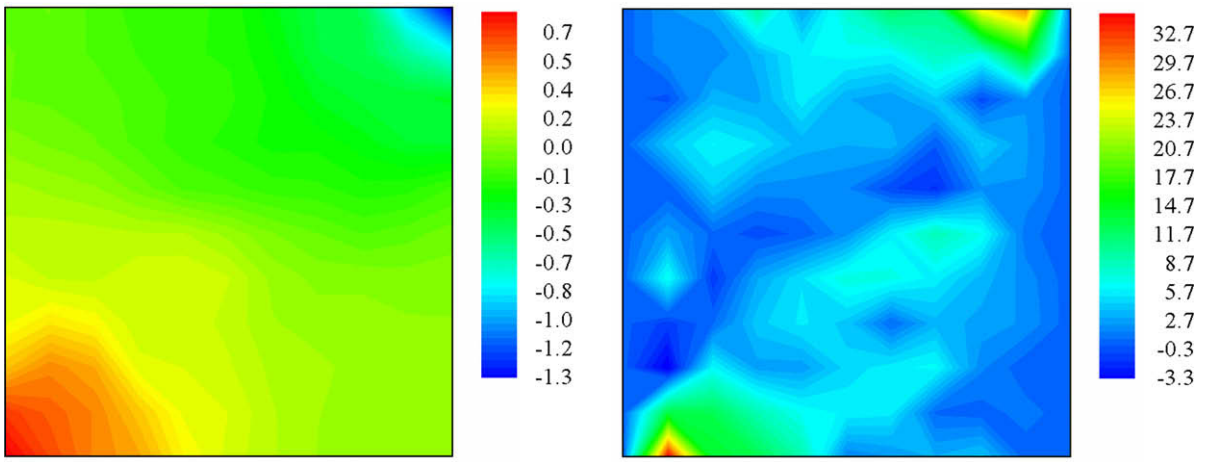


Fig. 17. Mean contours of the stochastic coarse-scale solution. (Left) Coarse-scale pressure, (Right) coarse-scale x-direction flux.

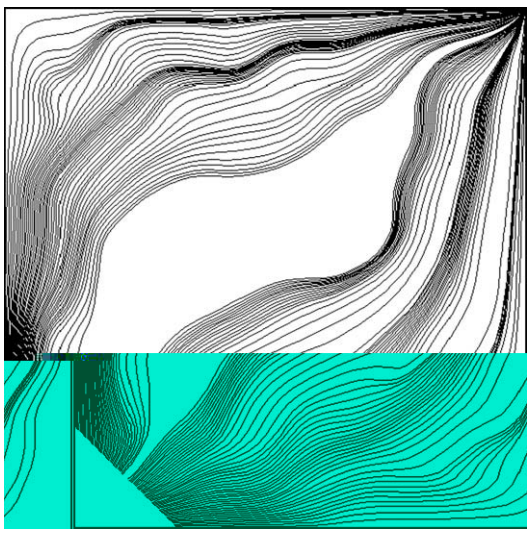


Fig. 18. Streamlines drawn on the mean stochastic coarse-scale velocity.

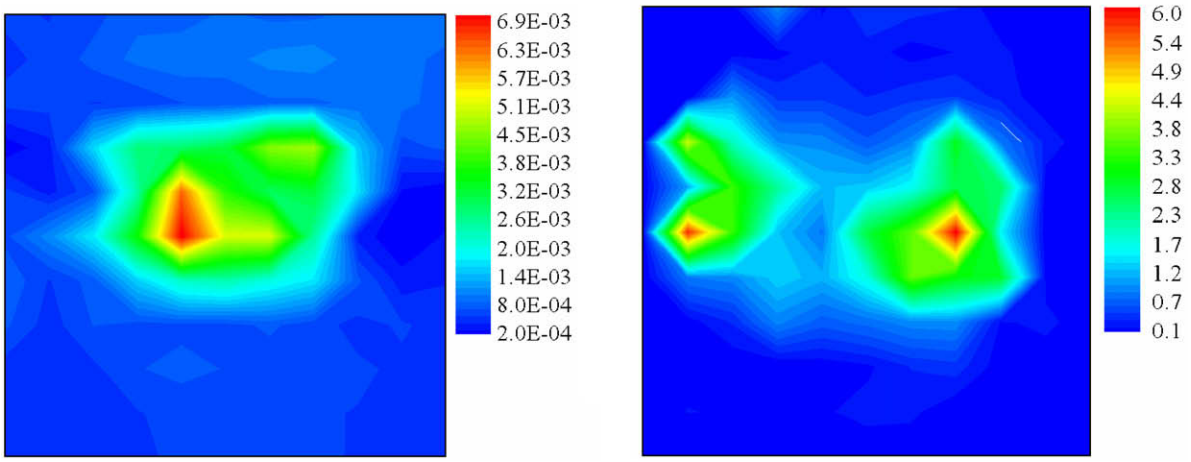


Fig. 19. Standard deviation of the stochastic coarse-scale solution. (Left) Coarse-scale pressure, (Right) coarse-scale *x*-direction flux.

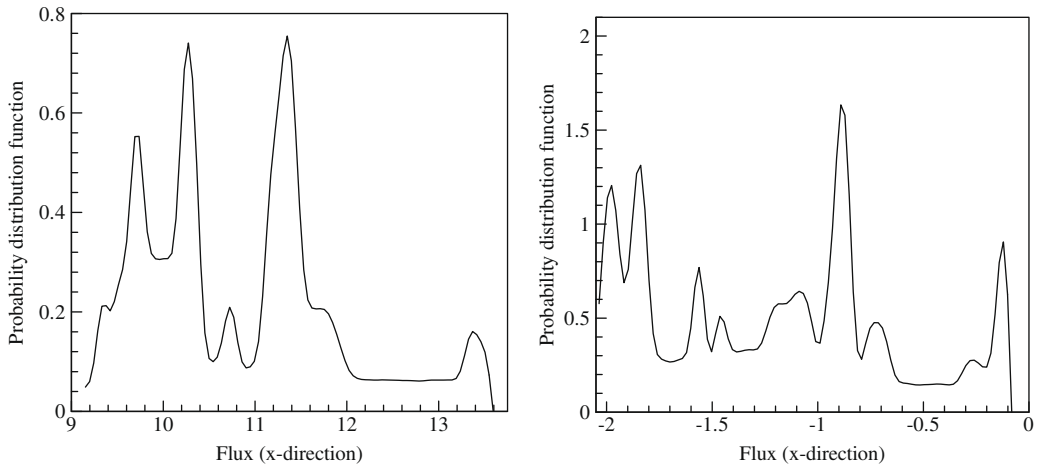


Fig. 20. Probability distribution functions of the coarse-scale flux (*x*-direction) for two points in the domain.

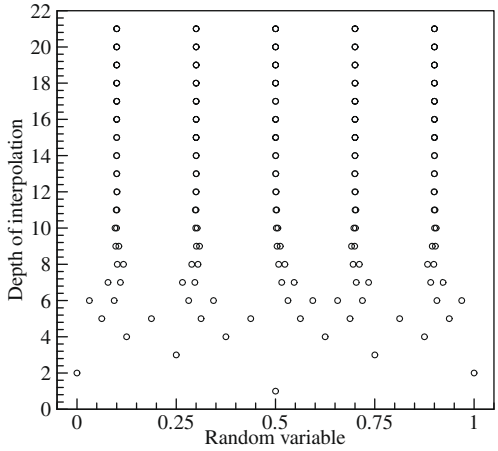


Fig. 21. The refinement of the adaptive sparse grid. Each row of points represents the refinement at a particular level of interpolation. The interpolation starts at the bottom.

depth of interpolation (depth of interpolation 7) about 8000 collocation points would have been required using conventional sparse grid collocation. The adaptive sparse grid strategy sampled the 4D stochastic space anisotropically with the random dimension corresponding to the x-location of the permeability barrier being the most sensitive dimension.

Fig. 17 plots the mean contours of the stochastic coarse-scale pressure and flux in the x-direction. The effects of the multiscale variation in the structure of permeability is visible in the uneven pressure contours near the injection and production wells. Furthermore, the pressure is essentially uniform close to the middle of the domain where the permeability barrier lies. This is also seen in Fig. 17 (right), where most of the flow circumvents the location of the fault. This is even more clearly seen in the streamlines drawn on the mean coarse-scale velocity field as shown in Fig. 18.

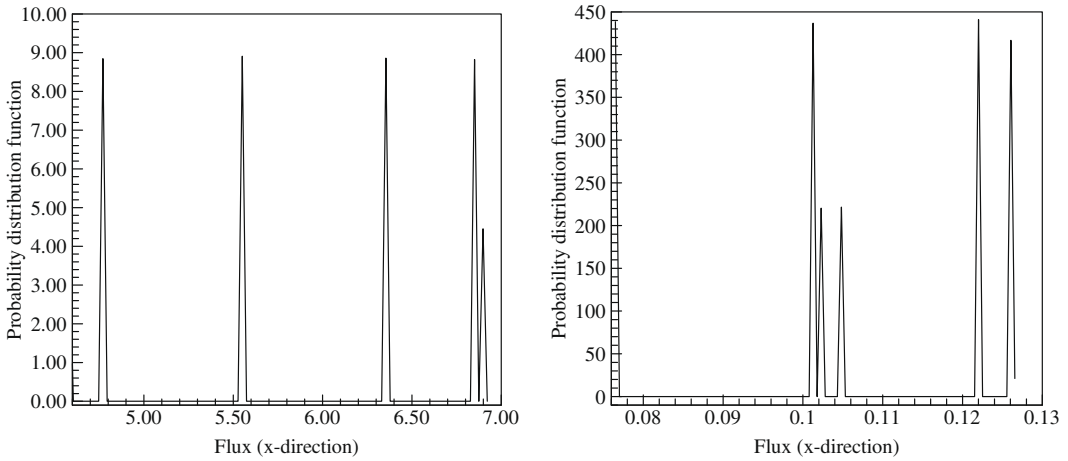


Fig. 22. Probability distribution functions of the coarse-scale flux (x-direction) for two points in the domain.

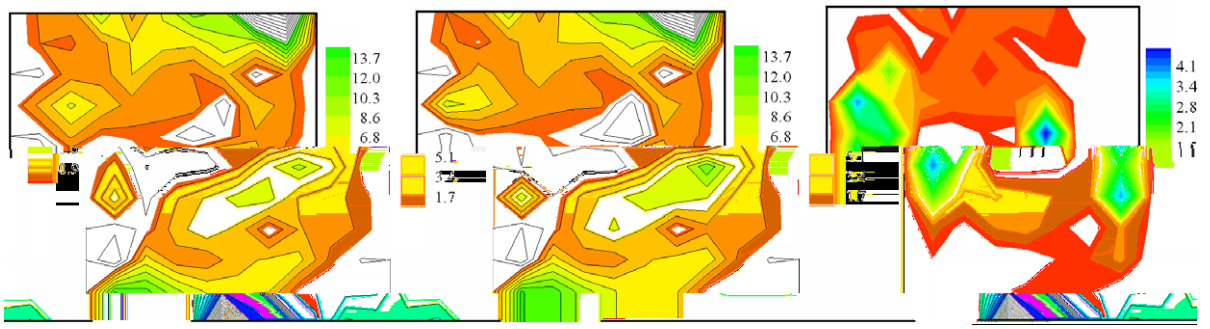


Fig. 23. (Left and center) Two realizations of the coarse-scale flux (x-direction) for a 1% perturbation in the x-location of the permeability barrier, (Right) difference between the two solutions.

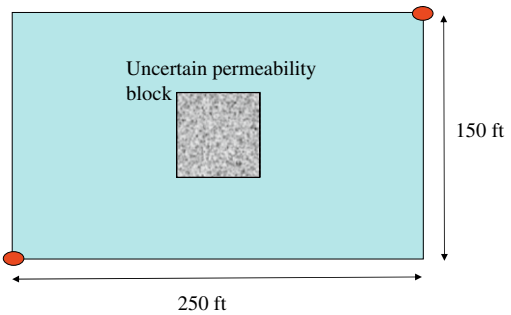


Fig. 24. Schematic of the domain. The permeability is uncertain in part of the domain.

Fig. 19 plots the standard deviation contours of the coarse-scale stochastic pressure and x-direction flux. As expected, most of the deviation in the pressure occurs close to the region of the permeability fault. Interestingly, most of the large deviation region is 'downstream' of the location of the uncertainty – along the direction of flow. This is also seen in Fig. 19 (right) where the region of large standard deviation is on either side of the permeability barrier and extends along the flow direction downstream towards the production well on the top right.

From the stochastic solution, the probability distribution function of the coarse-scale x-direction flux is plotted at two spatial locations. The first point (120, 40) has a relatively large standard deviation in the flux, while the second point (100, 60) has a smaller standard deviation. The PDFs are plotted in Fig. 20. Notice that the PDFs at both points have several peaks (in contrast with the previous case) pointing to the possibility of mode shifts existing. In particular, Fig. 20 (right) is the PDF of a point that lies very close to the permeability barrier and it clearly shows up to 10 distinct peaks. It appears that small variations in the location of the peak results in significant changes in the flow patterns leading to the multi-modal structure of the PDF.

We investigate this possibility by reducing the problem to its most sensitive stochastic dimension. As stated earlier, the adaptive sparse grid collocation algorithm preferentially refines the stochastic solution along regions of steep variations and discontinuities. Most of the refinement proceeded along the dimension that characterized the uncertainty in the x-location of the permeability barrier (i.e. Y_1). We solved a problem with this as the only input stochastic variable. Fig. 21 plots the evolution of the adaptive grid. Notice that after the 8th level of interpolation, most of the refinement is concentrated at five points. This implies the definite existence of steep gradients or discontinuities around these points in stochastic space. This is clearly seen in Fig. 22, where the PDF of the coarse-scale flux in the x-direction is plotted at two spatial points. The PDFs in both cases consist of five sharp peaks.

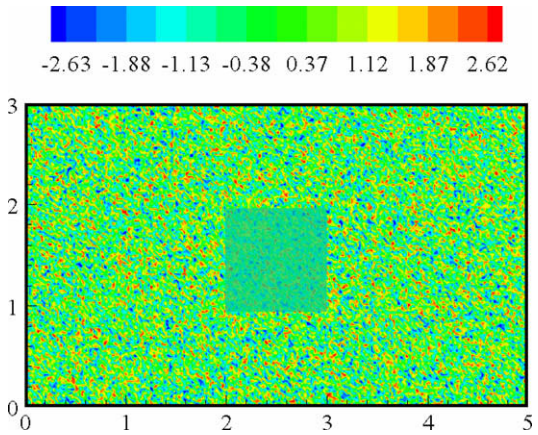


Fig. 25. The multiscale log-permeability distribution in the domain. The axes are scaled (1 unit = 50 ft).

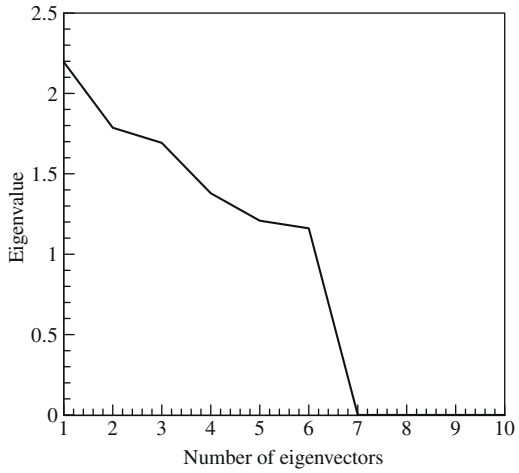


Fig. 26. The eigen-spectrum of the data-set. The first six eigen-modes represent most of the variation that the input data-set exhibits.

In Fig. 23, we illustrate this mode shift behavior by plotting two realizations of the coarse-scale stochastic flux (x-direction). These two realizations correspond to two points in stochastic space that are separated by 1% ($Y_1 = 0.295$ and 0.305). This corresponds to a mode shift located at $Y_1 = 0.30$ as seen in Fig. 21. The flow structure close to the permeability barrier is significantly different for the two cases.

In comparison with the uncertainty due to multiscale permeability variation, even a small uncertainty in the location of large scale structures results in significant variabilities in the flow patterns.

7.2. Effect of localized uncertainties

In this second example, we look at the effect of localized uncertainties in permeability and how they propagate into the complete domain. A schematic of the domain is shown in Fig. 24. The region consists of a 250×150 ft^2 domain where the

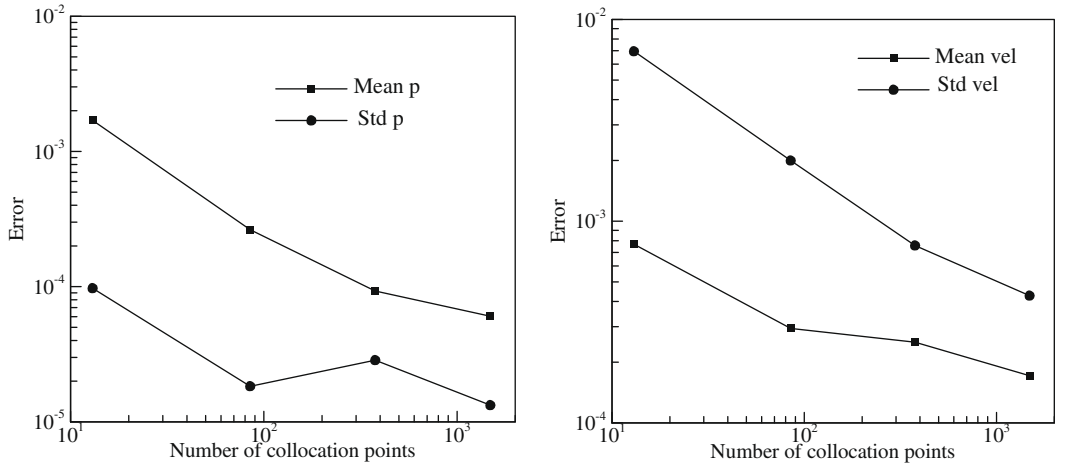


Fig. 27. Convergence of the stochastic solution (mean and standard deviation) with increasing number of collocation points. (Left) Stochastic coarse-scale pressure, (Right) stochastic coarse-scale x-direction flux.

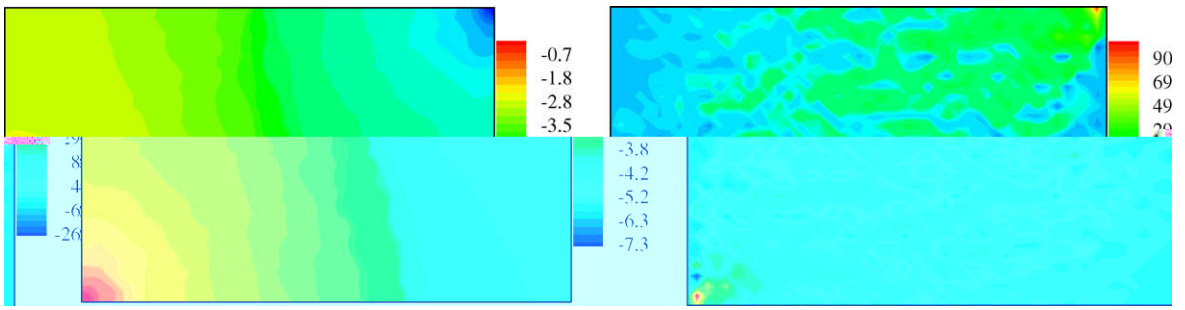


Fig. 28. Mean contours of the stochastic coarse-scale solution. (Left) Coarse-scale pressure, (Right) coarse-scale x-direction flux.

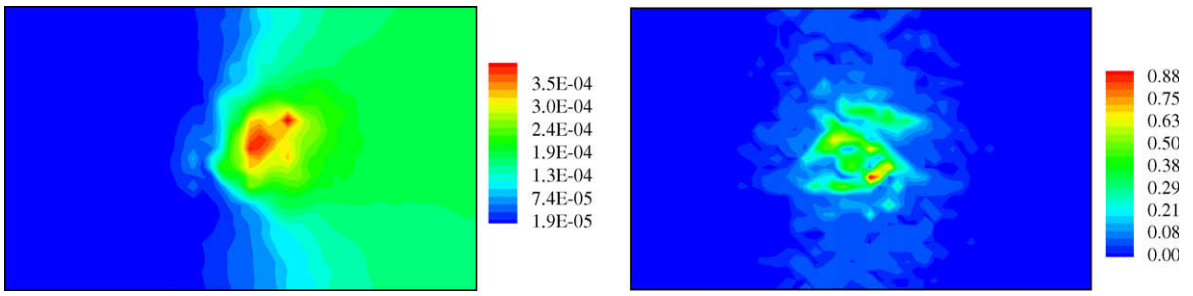


Fig. 29. Standard deviation of the stochastic coarse-scale solution. (Left) Coarse-scale pressure, (Right) coarse-scale x-direction flux.

multiscale permeability is given in 1 ft^2 blocks. An injection well is located at the bottom left corner of the domain while a production well is placed at the top right corner of the domain. No flow boundary conditions are enforced on all the four boundaries. The full-scale discretization of the domain corresponds to a 250×150 element triangulation.

The permeability in the domain is shown in Fig. 25. The deterministic permeability is taken from the 10th SPE comparative project [63]. The permeability shows variation of five orders of magnitude. Furthermore, there is a localized region of uncertain permeability in the central portion of the domain, corresponding to a 50×50 ft^2 domain. Aggressive coarse-scaling is performed at the complete domain which is characterized using a coarse 50×30 element discretization.

We are given a finite number of snapshots of plausible permeability distributions in the uncertain permeability block. Based on the number of available snapshots, we consider two cases (a) for small number of snapshots, a linear model reduction strategy is utilized to construct a data-driven model of the multiscale permeability in the domain, and (b) if a large number of snapshots are available, a non-linear model reduction strategy is utilized.

7.2.1. Small number of snapshots

In this case, a data-set containing 85 plausible permeability distributions in the uncertain part of the domain is available. Based on the discussion in Section 4.2.1, the covariance matrix C (see Eq. (54)) is constructed using these 85 permeability distributions and the eigenvalue problem solved. Fig. 26 plots the eigen-spectrum of the covariance matrix. The first six eigenmodes accurately represent the variability in the given data-set. These eigen-modes are utilized to construct the corresponding low-dimensional representation of the stochastic permeability distribution in the central block of the domain. The dimensionality of the representation is 6.

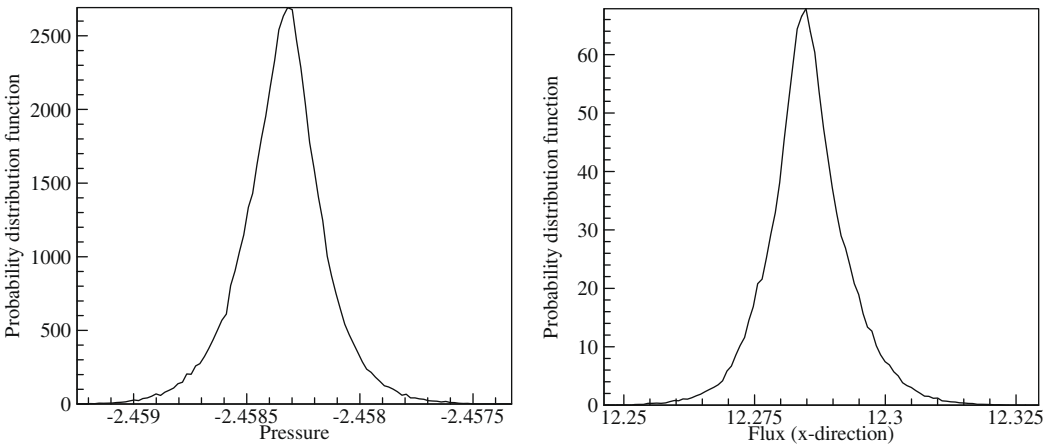


Fig. 30. Probability distribution functions of the coarse-scale pressure (Left) and flux (x-direction) (Right) for a point upstream of the uncertainty.

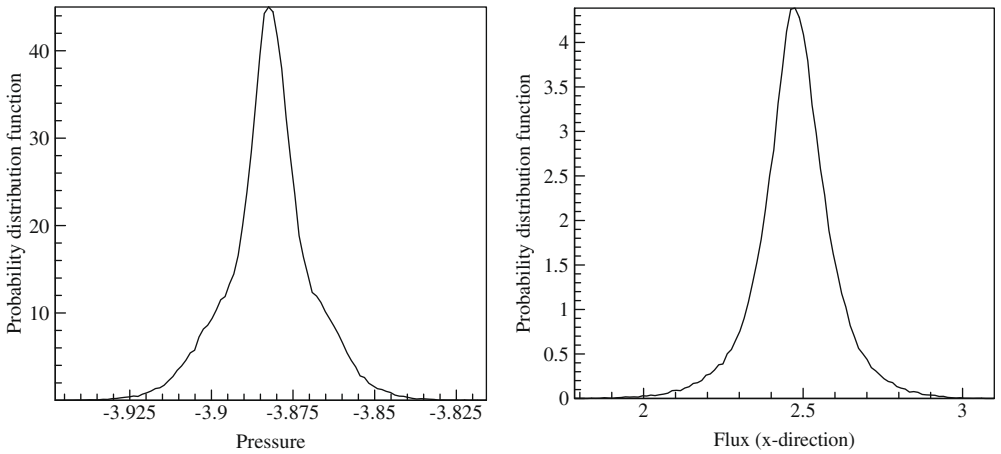


Fig. 31. Probability distribution functions of the coarse-scale pressure (Left) and flux (x-direction) (Right) for a point in the uncertain domain.

Twelve nodes of our local Linux cluster (corresponding to 48 processors) was utilized. The total computational time was about 24 h to solve this $6+2$ (stochastic + spatial domain) dimensional problem. Fig. 27 (left) plots the convergence of the stochastic pressure (mean and standard deviation) with increasing number of collocation points, while Fig. 27 (right) plots the convergence of the stochastic x-direction flux with increasing number of collocation points.

Fig. 28 plots the mean contours of the stochastic coarse-scale pressure and flux in the x-direction. Note the highly dendrite-like structure of the x-direction flux. Due to the large variation in the permeability, the flow tries to find the path with the largest permeability resulting in this dendrite-like structure.

Fig. 29 plots the standard deviation contours of the coarse-scale stochastic pressure and x-direction flux. Upstream of the uncertain permeability block, the standard deviation of the coarse-scale pressure and velocity are both negligible. Most of the deviation in the pressure occurs in the uncertain domain with its effects felt downstream of the uncertain domain.

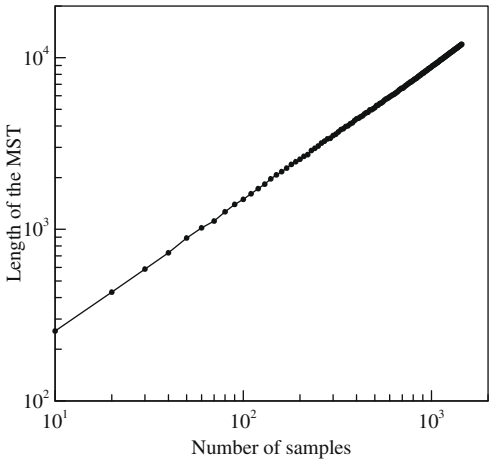


Fig. 32. Plot of the length functional of the MST of the graph G for various sample sizes. The intrinsic dimension is related to the slope of the graph.

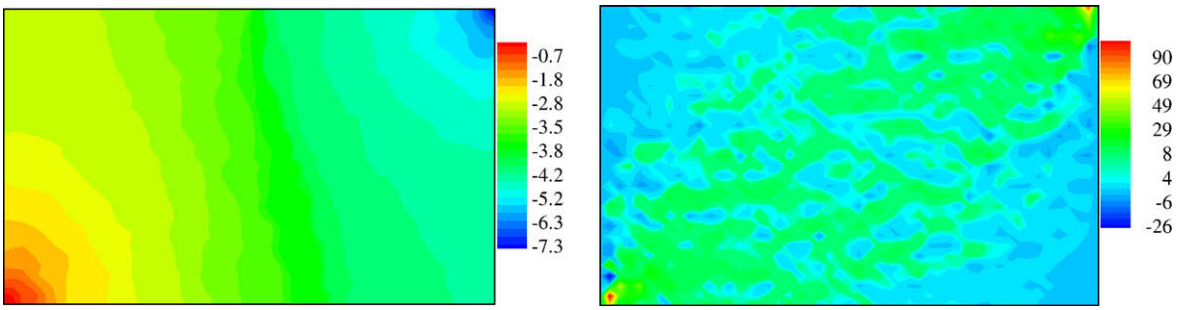


Fig. 33. Mean contours of the stochastic coarse-scale solution. (Left) Coarse-scale pressure, (Right) coarse-scale x-direction flux.

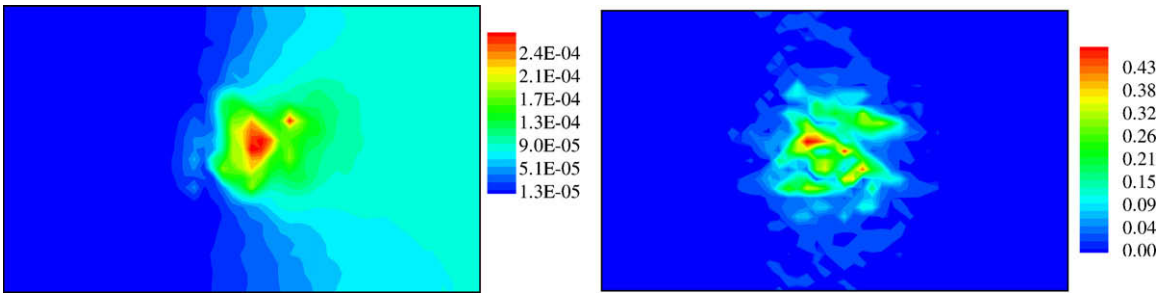


Fig. 34. Standard deviation of the stochastic coarse-scale solution. (Left) Coarse-scale pressure, (Right) coarse-scale x-direction flux.

The velocity flux has maximum deviation in the region of uncertainty and the deviation rapidly decays in the rest of the domain.

From the stochastic solution, the probability distribution function of the coarse-scale pressure and x-direction flux is plotted at two spatial locations. The first point (50, 50) is upstream of the localized uncertain region, while the second point (125, 75) is in the region of uncertain permeability. The PDFs are plotted in Figs. 30 and 31. For the point upstream, the probability distributions are peaked and have a very small range of variation. In comparison, the point in the region of uncertainty has a significantly larger range of variability.

7.2.2. Large number of snapshots

In this case, a data-set containing 1500 plausible permeability distributions in the uncertain part of the domain is available. Based on the discussion in Section 4.2.2, the neighborhood graph G is constructed using these 1500 permeability distributions. Following this, the geodesic distance matrix is computed. The BHH theorem [34] is used to compute the intrinsic dimensionality of the dataset. The dimensionality is related to the slope of the length functional of the minimal spanning graph (MST) of the neighborhood graph G of the dataset (see [34]). Fig. 32 plots the variation in the length functional of the MST. The optimal dimensionality was estimated to be $d = 5$. The corresponding low-dimensional representation of the stochastic permeability distribution in the central block of the domain (Eq. (56)) is constructed.

Fifty nodes of our local Linux cluster (corresponding to 200 processors) was utilized. The total computational time was about 12 h to solve this $5 + 2$ (stochastic + spatial domain) dimensional problem.

Fig. 33 plots the mean contours of the stochastic coarse-scale pressure and flux in the x-direction. Comparing this with Fig. 28, it is apparent that the results are quite similar to those obtained by generating a data-driven model utilizing a much

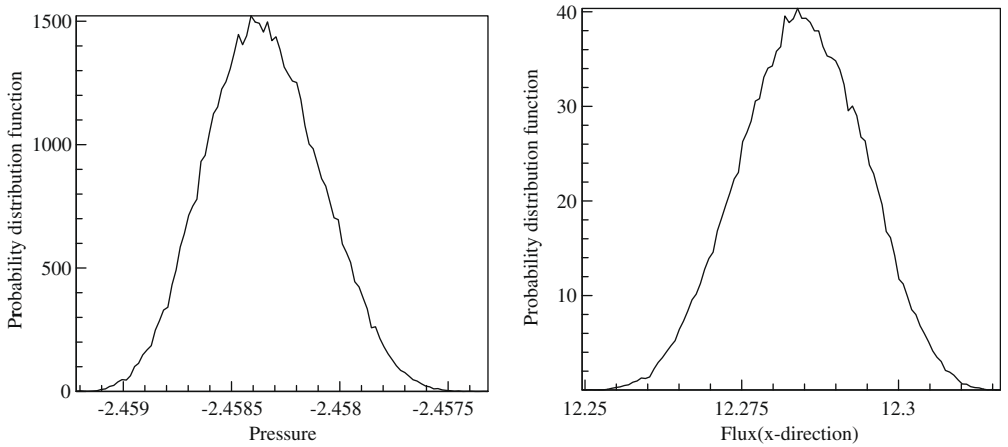


Fig. 35. Probability distribution functions of the coarse-scale pressure (left) and flux (x-direction) (right) for a point upstream of the uncertainty.

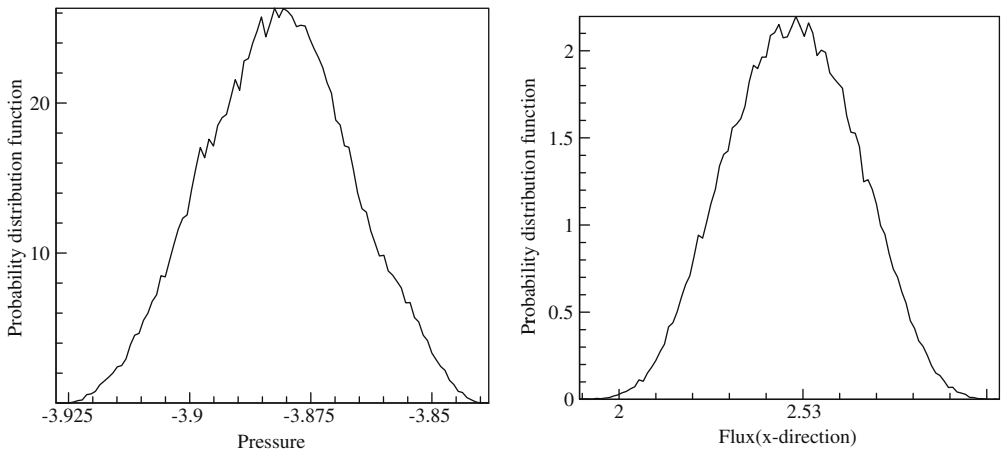


Fig. 36. Probability distribution functions of the coarse-scale pressure (left) and flux (x-direction) (right) for a point in the uncertain domain.

lower amount of initial information (using the linear model generation strategy). Limited amount of data is enough to construct a stochastic model for the permeability that results in fairly accurate modeling of the mean behavior. However, additional data results in substantial improvements in higher-order statistics.

Fig. 34 plots the standard deviation contours of the coarse-scale stochastic pressure and x-direction flux. The utilization in more information (larger number of snapshots to build the input model) results in an overall reduction in the standard deviation in the domain, though the general trends are similar (compare Figs. 29 and 34).

From the stochastic solution, the probability distribution function of the coarse-scale pressure and x-direction flux is plotted at the same two spatial locations. The PDFs are plotted in Figs. 35 and 36. For the point upstream, the probability distributions are peaked and have a very small range of variation. In comparison, the point in the region of uncertainty has a significantly larger range of variability. Comparing Fig. 36 with Fig. 31 illustrates the effect of the incorporation of more data into the construction of the stochastic input model.

8. Conclusions

We have extended the state-of-art in deterministic multiscale modeling of flow through heterogeneous media to a stochastic multiscale framework. We accomplish this in three steps. In the first step, a variational stochastic mixed multiscale formulation is developed to incorporate the effects of stochastic subgrid-scale permeability. The input to this is a finite-dimensional representation of the stochastic permeability that is constructed from limited permeability information in the form of snapshots or statistics. We utilize various data-driven model reduction strategies to embed this limited information into viable stochastic input models of the permeability. The resulting multiscale SPDEs are solved utilizing an adaptive sparse grid collocation strategy. We showcase the complete formulation through realistic large scale applications of flow through heterogeneous random media. This is (to the best knowledge of the authors) the first instance of a stochastic variational multiscale treatment of flow through random heterogeneous media. We are currently utilizing this framework in multiple areas: The first involves performing estimation of multiscale permeability distributions as well as contaminant detection problems. This multiscale strategy towards estimation and inverse problems offers two advantages (1) it provides a computationally attractive framework to hierarchically estimate spatial property variations and (2) it also naturally links coarse-scale integrated responses like flow, contaminant with the lower scale features via the stochastic multiscale basis functions. The robustness of this technique is yet to be investigated and will be the focus of a forthcoming publication. The second application of this development is to incorporate stochastic model reduction strategies to construct low-complexity surrogate representations of the stochastic multiscale equations for accelerated analysis.

Acknowledgments

This research was supported with grants to Cornell by the Computational Mathematics program of AFOSR (Grant F49,620-00-1-0373) and the Computational Mathematics program of the NSF (award DMS-0809062).

References

[1] T. Arbogast, S.L. Bryant, A two-scale numerical subgrid technique for waterflood simulations, SPE J. (2002) 446–457.
[2] T. Arbogast, S.E. Minkoff, P.T. Keenan, An operator-based approach to upscaling the pressure equation, Computational Methods in Water Resources XII, Computational Methods in Contamination and Remediation of Water Resources, vol. 1, Computational Mechanics Publications, Southampton, UK, 1998.
[3] T. Arbogast, Implementation of a locally conservative numerical subgrid upscaling scheme for two-phase darcy flow, Comput. Geosci. 6 (2002) 453–480.
[4] T. Arbogast, Numerical subgrid upscaling of two-phase flow in porous media, Lecture Notes in Physics, vol. 552, Springer, Berlin, 2000.
[5] T.J.R. Hughes, Multiscale phenomena: Green's functions, the dirichlet-to-neumann formulation, subgrid scale models, bubbles and the origin of stabilized methods, Comp. Meth. Appl. Mech. Eng. 127 (1995) 387–401.
[6] T.J.R. Hughes, G.R. Feijo, L. Mazzei, J.B. Quincy, The variational multiscale method a paradigm for computational mechanics, Comp. Meth. Appl. Mech. Engrg. 166 (1998) 3–24.
[7] R. Juanes, T.W. Patzek, A variational multiscale finite element method for multiphase flow in porous media, Finite Elem. Anal. Des. 41 (2005) 763–777.
[8] W.E.B. Engquist, The heterogeneous multiscale methods, Comm. Math. Sci. 1 (2003) 87–133.
[9] W.E.P. Ming, P. Zhang, Analysis of the heterogeneous multiscale method for elliptic homogenization problems, J. Am. Mathe. Soc. 18 (2005) 21–156.
[10] T.Y. Hou, X.H. Wu, A multiscale finite element method for elliptic problems in composite materials and porous media, J. Comput. Phys. 134 (1997) 169–189.
[11] T.Y. Hou, X.H. Wu, Z.Q. Cai, Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients, Math. Comput. 68 (1999) 913–943.
[12] P. Park, T.Y. Hou, Multiscale numerical methods for singularly-perturbed convection diffusion equations, Int. J. Comput. Meth. 1 (2004) 17–65.
[13] Y. Efendiev, V. Ginting, T.Y. Hou, R. Ewing, Accurate multiscale finite element methods for two-phase flow simulations, J. Comput. Phys. 220 (2006) 155–174.
[14] J.E. Aarnes, V. Kippe, K.A. Lie, Mixed multiscale finite elements and streamline methods for reservior simulation of large geomodels, Adv. Water Resources 28 (2005) 257–271.
[15] I. Babuska, U. Banerjee, J. Osborn, Generalized finite element methods: main ideas, results and perspective, Int. J. Comput. Meth. 1 (2004) 67–103.
[16] G. Sangalli, Capturing small scales in elliptic problems using a residual-free bubbles finite element method, SIAM MMS 1 (2003) 485–503.
[17] J.E. Aarnes, V. Kippe, K.-A. Lie, A.B. Rustad, Modelling of multiscale structures in flow simulations for petroleum reservoirs, in: Geometric Modelling, Numerical Simulation, and Optimization, Springer, Berlin, 2007.
[18] R.G. Ghanem, P.D. Spanos, Stochastic Finite Elements: A Spectral Approach, Dover publications, 1991.
[19] R. Ghanem, Probabilistic characterization of transport in heterogeneous porous media, Comput. Meth. Appl. Mech. Engrg. 158 (1998) 199–220.

[20] D. Xiu, G.E. Karniadakis, Modeling uncertainty in steady state diffusion problems via generalized polynomial chaos, Comput. Meth. Appl. Mech. Engrg. 191 (2002) 4927–4948.
[21] D. Xiu, G.E. Karniadakis, Modeling uncertainty in flow simulations via generalized polynomial chaos, J. Comput. Phys. 187 (2003) 137–167.
[22] M. Jardak, C.-H. Su, G.E. Karniadakis, Spectral polynomial chaos solutions of the stochastic advection equation, J. Sci. Comput. 17 (2002) 319–338.
[23] X. Wan, G.E. Karniadakis, An adaptive multi-element generalized polynomial chaos method for stochastic differential equations, J. Comput. Phys. 209 (2005) 617–642.
[24] D. Xiu, G.E. Karniadakis, The Wiener-Askey polynomial chaos for stochastic differential equations, SIAM J. Sci. Comput. 24 (2002) 619–644.
[25] B. Ganapathysubramanian, N. Zabaras, Sparse grid collocation schemes for stochastic natural convection problems, J. Comput. Phys. (2007) 652–685.
[26] D. Xiu, J.S. Hesthaven, High-order collocation methods for the differential equation with random inputs, SIAM J. Sci. Comput. 27 (2005) 1118–1139.
[27] D. Xiu, Efficient collocational approach for parametric uncertainty analysis, Comm. Comput. Phys. 2 (2007) 293–309.
[28] X. Ma, N. Zabaras, An adaptive hierarchical sparse grid collocation algorithm for the solution of stochastic differential equations, J. Comput. Phys., submitted for publication.
[29] C. Desceliers, R. Ghanem, C. Soize, Maximum likelihood estimation of stochastic chaos representations from experimental data, Int. J. Numer. Meth. Engng. 66 (2006) 978–1001.
[30] C.L. Winter, D.M. Tartakovsky, Mean flow in composite porous media, Geophys. Res. Lett. 27 (2000) 1759–1762.
[31] C.L. Winter, D.M. Tartakovsky, Groundwater flow in heterogeneous composite aquifers, Water Resour. Res. 38 (2002) 23.1–23.11.
[32] L. Guadagnini, A. Guadagnini, D.M. Tartakovsky, Probabilistic reconstruction of geologic facies, J. Hydrol. 294 (2004) 57–67.
[33] B. Ganapathysubramanian, N. Zabaras, Modelling diffusion in random heterogeneous media: data-driven models, stochastic collocation and the variational multi-scale method, J. Comput. Phys. 226 (2007) 326–353.
[34] B. Ganapathysubramanian, N. Zabaras, A non-linear dimension reduction methodology for generating data-driven stochastic input models, J. Comput. Phys. 227 (2008) 6612–6637.
[35] X.F. Xu, A multiscale stochastic finite element method on elliptic problems involving uncertainties, Comput. Meth. Appl. Mech. Engrg. 196 (2007) 2723–2736.
[36] B. Velamur Asokan, N. Zabaras, Variational multiscale stabilized FEM formulations for transport equations: stochastic advection-diffusion and incompressible stochastic Navier-Stokes equations, J. Comput. Phys. 202 (2005) 94–133.
[37] V. Kippe, J.E. Aarnes, K.A. Lie, A comparison of multiscale methods for elliptic problems in porous media flow, Comput. Geosci. 12 (2008) 377–398.
[38] I. Babuska, R. Tempone, G.E. Zouraris, Solving elliptic boundary value problems with uncertain coefficients by the finite element method: the stochastic formulation, Comput. Meth. Appl. Mech. Engrg. 194 (2005) 1251–1294.
[39] R. Juanes, F.-X. Dub, A locally-conservative variational multiscale method for the simulation of flow in porous media with multiscale source terms, Comput. Geosci. 12 (2008) 273–295.
[40] T. Arbogast, K.J. Boyd, Subgrid upscaling and mixed multiscale finite elements, SIAM J. Num. Anal. 44 (2006) 1150–1171.
[41] I. Babuska, R. Tempone, G.E. Zouraris, Galerkin finite elements approximation of stochastic finite elements, SIAM J. Numer. Anal. 42 (2004) 800–825.
[42] Ch. Schwab, R.A. Todur, Sparse finite elements for stochastic elliptic problems – higher moments, Computing 71 (1) (2003) 43–63.
[43] Ch. Schwab, R.A. Todur, Sparse finite elements for elliptic problems with stochastic loading, Numerische Mathematik 95 (2001) 707–734.
[44] T. Petersdorff, Ch. Schwab, Sparse finite element methods for operator equations with stochastic data, Appl. Mathe. 51 (2006) 145–180.
[45] F. Brezzi, M. Fortin, Mixed and Hybrid Finite Element Methods, Springer-Verlag, 1991.
[46] P.A. Raviart, J.M. Thomas, A mixed finite element method for second order elliptic problems, in: Mathematical Aspects of the Finite Element Method, Lecture Notes in Mathematics, vol. 606, Springer-Verlag, Berlin, 1977.
[47] GSLIB: Geostatistical Software Library, available at <http://www.pe.utexas.edu/Geosci/Software/GSLIB/gslib.html>.
[48] M. Loéve, Probability Theory, fourth ed., Springer-Verlag, Berlin, 1977.
[49] P. Holmes, J.L. Lumley, G. Berkooz, Turbulence, Coherent Structures, Dynamical Systems and Symmetry, Cambridge University Press, 1998.
[50] V. Sundararaghavan, N. Zabaras, Classification of three-dimensional microstructures using support vector machines, Comp. Mat. Sci. 32 (2005) 223–239.
[51] J. Tenenbaum, V. DeSilva, J. Langford, A global geometric framework for nonlinear dimension reduction, Science 290 (2000) 2319–2323.
[52] J.T. Kent, J.M. Bibby, K.V. Mardia, Multivariate Analysis (Probability and Mathematical Statistics), Elsevier, 2006.
[53] W. Hardle, L. Simar, Applied Multivariate Statistical Analysis, 2004, available online at <http://www.quantlet.com/mdstat/scripts/mva/htmlbook/>.
[54] J. Breadwood, J.H. Halton, J.M. Hamersley, The shortest path through many points, Proc. Cambridge Philos. Soc. 55 (1959) 299–327.
[55] J.A. Costa, A.O. Hero, Entropic graphs for manifold learning, in: Proceedings of the IEEE Asilomar Conference Signals, Systems, and Computers, Pacific Groove, CA, November, 2003.
[56] J.A. Costa, A.O. Hero, Geodesic entropic graphs for dimension and entropy estimation in manifold learning, IEEE Trans. Signal Process. 52 (2004) 2210–2221.
[57] J.A. Costa, A.O. Hero, Manifold Learning with Geodesic Minimal Spanning Trees, 2003 <arXiv:cs/0307038v1>.
[58] C. Canuto, M.Y. Hussaini, A. Quarteroni, T.A. Zang, Spectral Methods: Fundamentals in Single Domains Series, Springer, 2006.
[59] F. Nobile, R. Tempone, C. Webster, An anisotropic sparse grid collocation method for elliptic partial differential equations with random input data, MOX, Dipartimento di Matematica, Politecnico di Milano, Italy, 2007.
[60] S. Smolyak, Quadrature and interpolation formulas for tensor product of certain classes of functions, Soviet Math. Dokl. 4 (1963) 240–243.
[61] T. Gerstner, M. Griebel, Numerical integration using sparse grids, Numer. Algor. 18 (1998) 209–232.
[62] J.W. Jennings, S.l. Bryant, Outcrop permeability statistics and implication for fluid flow and reactive transport modeling, Bureau of Economic Geology, U. Texas Austin, Report, 2000.
[63] M.A. Christie, M.J. Blunt, Tenth SPE comparative solution project: a comparison of upscaling techniques, SPE Reserv. Eval. Eng. 4 (2001) 308–317.