

Conditioning by Projection for the Sampling from Prior Gaussian Distributions

Alsadig Ali[1], Abdullah Al-Mamun[2], Felipe Pereira[1], and Arunasalam Rahunanthan[3(✉)]

[1] Department of Mathematical Sciences, The University of Texas at Dallas, Richardson, TX 75080, USA
[2] Institute of Natural Sciences, United International University, Dhaka, Bangladesh
[3] Department of Mathematics and Computer Science, Central State University, Wilberforce, OH 45384, USA
aRahunanthan@centralstate.edu

Abstract. In this work we are interested in the (ill-posed) inverse problem for absolute permeability characterization that arises in predictive modeling of porous media flows. We consider a Bayesian statistical framework with a preconditioned Markov Chain Monte Carlo (MCMC) algorithm for the solution of the inverse problem. Reduction of uncertainty can be accomplished by incorporating measurements at sparse locations (static data) in the prior distribution. We present a new method to condition Gaussian fields (the log of permeability fields) to available sparse measurements. A truncated Karhunen-Loève expansion (KLE) is used for dimension reduction. In the proposed method the imposition of static data is made through the projection of a sample (expressed as a vector of independent, identically distributed normal random variables) onto the nullspace of a data matrix, that is defined in terms of the KLE. The numerical implementation of the proposed method is straightforward. Through numerical experiments for a model of second-order elliptic equation, we show that the proposed method in multi-chain MCMC studies converges much faster than the MCMC method without conditioning. These studies indicate the importance of conditioning in accelerating the MCMC convergence.

Keywords: Conditioning · Predictive simulations · Bayesian framework · MCMC

1 Introduction

Subsurface formations have spatial variability in their hydraulic properties in multiple length scales. It has been established that such variability plays an important role in determining fluid flow patterns in heterogeneous formations [4,18,24,25]. The characterization of such formations is essential for predictive

© Springer Nature Switzerland AG 2021
O. Gervasi et al. (Eds.): ICCSA 2021, LNCS 12952, pp. 653–669, 2021.
https://doi.org/10.1007/978-3-030-86973-1_46

simulations of multiphase flows in problems such as oil recovery, CO_2 sequestration, or contaminant transport. Procedures for characterization, such as a Bayesian framework using the Markov chain Monte Carlo (MCMC) algorithm, entails the solution of an ill-posed inverse problem. From static and dynamic data, the unknown coefficients of a governing system of partial differential equations that model fluid flow have to be recovered. In the inversion process, the available static data (for instance, measured absolute permeability values at sparse locations) have to be incorporated in the sampling of the (unknown) permeability field so that uncertainty is reduced as much as possible. This way, a conditioning procedure has to be applied such that all samples considered by the inversion method honor available measurements.

MCMC methods have been applied in several areas of science and engineering, and have attracted the attention of many research groups [19,33]. There are many recent and important developments related to gradient-based MCMC procedures. We mention, among many others, the Hamiltonian Monte Carlo [8], and the active subspace methods [10]. There are also Hessian-based procedures [34]. However, for subsurface flow problems gradient and Hessian calculations are computationally very expensive. Thus, for the problems we have in mind we consider the preconditioned (or two-stage) MCMC introduced in [9,16]. Some developments of this method include the multi-physics version [23], the parallelization in multi-core devices [21] and its extension to deformable subsurface formations [5]. This procedure has also been successfully applied in geophysics [44]. The Bayesian framework for inversion that we will use in this work has been carefully discussed in [1].

Our focus in this work is the conditioning of samples to sparse measured data. The importance of conditioning in reservoir simulation has been recognized for many years (see, for instance, [26]). The sequential Gaussian simulation (SGS) [13] has been widely used for this purpose. More recently a truncated Karhunen-Loève expansion (KLE) [16,22,29] has been applied both for dimensional reduction of the (computationally infeasible) large stochastic dimensions of fine grid discretizations of subsurface flow problems as well as for conditioning. In [37] one can find a discussion of the advantages of the KLE over the SGS. There are two lines of work for the conditioning of samples using the KLE. In [30] the basic KLE construction is altered to incorporate static data. In another line of work the vector of independent, identically distributed (i.i.d.) $\mathcal{N}(0,1)$ Gaussian variables that enter the KLE construction is modified to take data into account while keeping the KLE eigenvectors unaltered. This approach can be seen in [36] where a search through a set of matrices of size determined by the number of measurements is performed, and they take the one with the best condition number. In [37] the authors use the expressions for conditional means and covariances of Gaussian random variables [46] that require the inversion of a data covariance matrix.

The method that we introduce here does not alter the basic KLE construction and is computationally very inexpensive. By using the measured data in the KLE we define a data matrix that models Gaussian perturbations on top of a surface

defined by kriging [12] of the available data. The proposed procedure consists of replacing the vectors of i.i.d., $\mathcal{N}(0,1)$ Gaussian variables by the closest vector to them (in the least-squares sense) in the nullspace of the data matrix. This way, the Gaussian fields produced by the KLE honor exactly all the data available. We remark that all one needs to implement the proposed method is a basis for the nullspace of the data matrix. Such a basis is used to construct the projection onto the nullspace of the data matrix. Besides introducing this new method, we also present multi-chain MCMC studies for a model elliptic equation. We show convergence of this method for large dimensional problems. Moreover, these studies are used to illustrate the importance of conditioning for the convergence of MCMC methods. Although outside the scope of this work we believe the method presented here can also be applied to other sampling strategies such as the hierarchical method discussed in [6]. This topic is currently being considered by the authors.

This work is organized as follows. In Sect. 2 we discuss the model problem, kriging and dimensional reduction. Section 3 introduces the new method for conditioning by projection. In Sect. 4 we explain the Bayesian framework for inversion. We describe the preconditioned MCMC method with and without conditioning for the construction of permeability fields. Also, the method we use for convergence assessment of MCMC methods is presented. Section 5 is dedicated to a discussion of numerical results from our experiments. Conclusions and some remarks appear in Sect. 6.

2 Model Problem and Dimensional Reduction

2.1 The Model Problem

In the modeling of incompressible subsurface flow problems, such as in two-phase flows in oil reservoirs [14,38], the system of governing partial differential equations consists of a second order elliptic equation coupled to a hyperbolic dominated equation that describes fluid transport. In this work, we illustrate the proposed method for conditioning Gaussian fields in terms of the elliptic equation that enters in this system.

Let us consider $\Omega \subset \mathbb{R}^2$, a bounded domain with Lipschitz boundary, let $\boldsymbol{v}(\boldsymbol{x})$ and $p(\boldsymbol{x})$ denote the Darcy velocity and the pressure of the fluid, respectively, where $\boldsymbol{x} \in \Omega$, and $k(\boldsymbol{x})$ is the absolute permeability. Using Darcy's law, we have the following model elliptic problem:

$$\begin{aligned} \boldsymbol{v}(\boldsymbol{x}) &= -k(\boldsymbol{x})\nabla p(\boldsymbol{x}), \\ \nabla \cdot \boldsymbol{v}(\boldsymbol{x}) &= f, \quad \boldsymbol{x} \in \Omega, \end{aligned} \tag{1}$$

where $f \in L^2(\Omega)$ is a given source term, together with appropriate Dirichlet and Neumann type boundary conditions $p = g_p \in H^{\frac{1}{2}}(\partial\Omega_p)$ and $\boldsymbol{v} \cdot \hat{\boldsymbol{n}} = g_{\boldsymbol{v}} \in H^{-\frac{1}{2}}(\partial\Omega_{\boldsymbol{v}})$, the pressure boundary data and normal velocity data, respectively, with $\partial\Omega = \overline{\partial\Omega_p} \cup \overline{\partial\Omega_{\boldsymbol{v}}}$ and $\overline{\partial\Omega_p} \cap \overline{\partial\Omega_{\boldsymbol{v}}} = \varnothing$, $\hat{\boldsymbol{n}}$ is the exterior unit normal vector.

The numerical approximation of the above elliptic problem is derived from the weak formulation of the velocity-pressure system. The weak formulation can be found in [17].

2.2 Kriging

Kriging is a type of interpolation method, which uses a local estimation technique that gives the best linear unbiased estimator of unknown values [12,27]. The main idea of kriging is to estimate the values of a function (or random field) in a domain by calculating a weighted average of a given point in some neighborhood. The kriging takes into account (i) the distances between the estimated points in the domain and the given data points, and (ii) the distances between the given data points themselves. Thus, the kriging gives more weights to the nearest data points. When a point $\boldsymbol{x}_0$ coincides with a given data point $\boldsymbol{x}_i$, the kriging gives the exact value and the variance is zero. In the following subsection, we discuss the KLE, and in Sect. 3, we combine both kriging and KLE.

2.3 Dimensional Reduction

We use the KLE to reduce the large dimensional uncertainty space describing the permeability field $k(\boldsymbol{x})$. In this subsection, we present a brief description of the KLE [20,29]. Let $\boldsymbol{x} \in \Omega$, and suppose $\log[k(\boldsymbol{x})] = Y^k(\boldsymbol{x})$ is a Gaussian field. Also, let us assume $Y^k(\boldsymbol{x})$ is a second-order stochastic process, that is $Y^k(\boldsymbol{x}) \in L^2(\Omega)$ with a probability one. Set $E[(Y^k)^2] = 0$. Thus, for a given orthonormal basis $\{\varphi_i\}$ in $L^2(\Omega)$, $Y^k(\boldsymbol{x})$ can be written as the following:

$$Y^k(\boldsymbol{x}) = \sum_{i=1}^{\infty} Y_i^k \varphi_i(\boldsymbol{x}), \quad \text{with random coefficients} \quad Y_i^k = \int_{\Omega} Y^k(\boldsymbol{x})\varphi_i(\boldsymbol{x})d\boldsymbol{x}, \tag{2}$$

where $\varphi_i(\boldsymbol{x})$ is an eigenfunction satisfying

$$\int_{\Omega} R(\boldsymbol{x}_1, \boldsymbol{x}_2)\varphi_i(\boldsymbol{x}_2)d\boldsymbol{x}_2 = \lambda_i \varphi_i(\boldsymbol{x}_1), \quad i = 1, 2, ..., \tag{3}$$

with $\lambda_i = E[(Y_i^k)^2] > 0$ and $R(\boldsymbol{x}_1, \boldsymbol{x}_2)$ is the covariance function, $\text{Cov}(Y^k(\boldsymbol{x}_1), Y^k(\boldsymbol{x}_2))$. Letting $\theta_i^k = Y_i^k/\sqrt{\lambda_i}$, the KLE in Eq. (2) can be written as

$$Y^k(\boldsymbol{x}) = \sum_{i=1}^{\infty} \sqrt{\lambda_i}\theta_i^k \varphi_i(\boldsymbol{x}), \tag{4}$$

where λ_i and φ_i satisfy Eq. (3). We arrange the eigenvalues in the descending order. The infinite series in Eq. (4) is called the Karhunen-Loève expansion. In general, we only need first n dominating terms such that the energy E is more than 95% [28], i.e.

$$E = \frac{\sum_{i=1}^{n} \lambda_i}{\sum_{i=1}^{\infty} \lambda_i} \geq 95\%. \tag{5}$$

Then, a truncated KLE is defined by

$$Y_n^k(\boldsymbol{x}) = \sum_{i=1}^{n} \sqrt{\lambda_i}\theta_i^k \varphi_i(\boldsymbol{x}). \tag{6}$$

From now on we will use $Y(\boldsymbol{x})$ instead of $Y_n^k(\boldsymbol{x})$ for samples generated by the truncated KLE.

3 Conditioning by Projection

Suppose the Gaussian field $Y(\boldsymbol{x})$ defined in Eq. (6) is a Gaussian perturbation on top of a kriged field $\hat{Y}(\boldsymbol{x})$ discussed in Sect. 2.2. That is,

$$Y(\boldsymbol{x}) = \hat{Y}(\boldsymbol{x}) + \sum_{i=1}^{n} \sqrt{\lambda_i}\varphi_i(\boldsymbol{x})\theta_i \ . \tag{7}$$

This can also be written as

$$\begin{aligned} Y(\boldsymbol{x}) - \hat{Y}(\boldsymbol{x}) &= \sum_{i=1}^{n} \sqrt{\lambda_i}\varphi_i(\boldsymbol{x})\theta_i \\ &= \boldsymbol{\phi}^T(\boldsymbol{x})\sqrt{D}\boldsymbol{\theta}, \end{aligned} \tag{8}$$

where D is a diagonal matrix of the dominating n eigenvalues and $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_n)$. For each $\boldsymbol{x}$, $\boldsymbol{\phi}(\boldsymbol{x}) \in \mathbb{R}^n$. Furthermore, assume that m measured data values $Y^1, Y^2, \ldots, Y^m$ of $Y(\boldsymbol{x})$ are given at sparse locations $\hat{\boldsymbol{x}} = (\hat{\boldsymbol{x}}_1, \hat{\boldsymbol{x}}_2, \ldots, \hat{\boldsymbol{x}}_m)$ in the domain Ω such that

$$Y(\hat{\boldsymbol{x}}_i) = Y^i, \quad i = 1, 2, \ldots, m \tag{9}$$

are known. Thus, by substituting (9) in Eq. (8) we have

$$\begin{aligned} \boldsymbol{\phi}^T(\hat{\boldsymbol{x}})\sqrt{D}\boldsymbol{\theta} &= Y(\hat{\boldsymbol{x}}) - \hat{Y}(\hat{\boldsymbol{x}}) \\ &= \mathbf{0}. \end{aligned} \tag{10}$$

This honors m data values at the corresponding sparse locations $\hat{\boldsymbol{x}}$. We can rewrite (10) as follows

$$A\boldsymbol{\theta} = \mathbf{0} \tag{11}$$

where $A = \boldsymbol{\phi}^T(\hat{\boldsymbol{x}})\sqrt{D} \in \mathbb{R}^{m \times n}$ is our data matrix. The solution of this homogeneous system of Eqs. (11) is the nullspace $N(A)$, a subspace of $\mathbb{R}^n$. Given an arbitrary vector $\boldsymbol{\theta} \in \mathbb{R}^n$ of i.i.d., $\mathcal{N}(0, 1)$ Gaussian variables that would be used in the KLE (7), in order to honor the given m data we select the closest vector to $\boldsymbol{\theta}$ in the subspace $N(A)$. This vector is obtained by projecting $\boldsymbol{\theta}$ onto the nullspace of the data matrix A.

Assume that A has rank r, we can easily find an orthonormal basis $\beta = \{\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_{n-r}\}$ for the nullspace of A. We assume that $r < n$, so that A has

a non-trivial nullspace. The projection matrix P onto the nullspace of A is given by [43]

$$P = QQ^T, \tag{12}$$

where Q is a matrix that has $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_{n-r}$ as columns. The projection vector of $\boldsymbol{\theta}$ onto the nullspace of the data matrix A is given by

$$\hat{\boldsymbol{\theta}} = P\boldsymbol{\theta}. \tag{13}$$

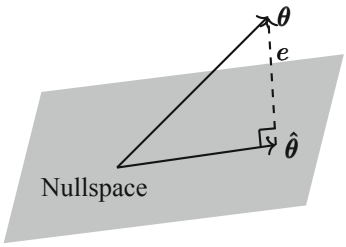

Fig. 1. Projection of the i.i.d., $\mathcal{N}(0,1)\,\boldsymbol{\theta}$ showing the closet vector $\hat{\boldsymbol{\theta}}$ in the nullspace.

Figure 1 illustrates the projection of the i.i.d., $\mathcal{N}(0,1)$ Gaussian variables $\boldsymbol{\theta}$ vector onto the nullspace of the data matrix A. The error vector defined as $\boldsymbol{e} = \boldsymbol{\theta} - \hat{\boldsymbol{\theta}}$ is minimized in the least squares sense. It is perpendicular to the projection vector $\hat{\boldsymbol{\theta}}$. Finally, we can rewrite Eq. (7) as

$$Y(\boldsymbol{x}) = \hat{Y}(\boldsymbol{x}) + \sum_{i=1}^{n} \sqrt{\lambda_i}\varphi_i(\boldsymbol{x})\hat{\theta}_i, \tag{14}$$

where $\hat{\theta}_i, i = 1, \ldots, n$ are the components of $\hat{\boldsymbol{\theta}}$ in Eq. (13). If $\boldsymbol{x} = \hat{\boldsymbol{x}}_i, i = 1, 2, \ldots, m$, in Eq. (14), the second term on the right-hand side is zero. This leads to $Y(\boldsymbol{x}_i) = \hat{Y}(\hat{\boldsymbol{x}}_i)$, and thus, we honor the exact values of the permeability field at the known locations $\hat{\boldsymbol{x}}_1, \ldots, \hat{\boldsymbol{x}}_m$.

4 The Bayesian Framework

4.1 Posterior Exploration

Our primary interest in this article is to characterize the permeability field of our domain of study. Here, we discuss the characterization of the permeability field conditioned on available pressure data. We use half (in a chessboard pattern) of the total pressure data in the discretized domain [45]. We denote the (log of the) permeability field and the reference pressure data by $\boldsymbol{\eta}$ and R_p, respectively. A Bayesian statistical approach along with a preconditioned MCMC method with

and without conditioning is used to solve the inverse problem. The posterior probability is calculated using the Bayes' theorem with respect to R_p as follows:

$$P(\boldsymbol{\eta}|R_p) \propto P(R_p|\boldsymbol{\eta})P(\boldsymbol{\eta}), \tag{15}$$

where $P(\boldsymbol{\eta})$ represents the Gaussian prior distribution. The normalizing constant is disregarded in this investigation as we do an iterative search in our MCMC method. We construct the permeability field $\boldsymbol{\eta}(\boldsymbol{\theta})$ using the KLE, where $\boldsymbol{\theta}$ is generated by the MCMC method. The likelihood is considered to be a Gaussian function as in [16], and as a result of that we have our likelihood function as

$$P(R_p|\boldsymbol{\eta}) \propto \exp\Big(-(R_p - R_{\boldsymbol{\eta}})^\top \Sigma (R_p - R_{\boldsymbol{\eta}})\Big), \tag{16}$$

where $R_{\boldsymbol{\eta}}$ denotes the simulated pressure data. We set the covariance matrix Σ as $\Sigma = \boldsymbol{I}/2\sigma_R^2$, where $\boldsymbol{I}$ and σ_R^2 represents the identity matrix and the precision parameter, respectively.

We use a MCMC algorithm to sample from the posterior distribution (15). In the MCMC algorithm, we have an instrumental distribution $I(\boldsymbol{\eta}_p|\boldsymbol{\eta})$ by which we propose $\boldsymbol{\eta}_p = \boldsymbol{\eta}(\boldsymbol{\theta}_p)$ at every iteration, where $\boldsymbol{\eta}$ represents the previously accepted state. We solve the forward problem for a given permeability field on the numerical simulator and compute the acceptance probability of a proposed proposal using

$$\alpha(\boldsymbol{\eta}, \boldsymbol{\eta}_p) = \min\left(1, \frac{I(\boldsymbol{\eta}|\boldsymbol{\eta}_p)P(\boldsymbol{\eta}_p|R_p)}{I(\boldsymbol{\eta}_p|\boldsymbol{\eta})P(\boldsymbol{\eta}|R_p)}\right). \tag{17}$$

Below we describe the preconditioned MCMC method with and without conditioning.

4.2 MCMC Algorithms with and Without Conditioning

In our study, we use the MCMC method with and without conditioning. The following random walk sampler (RWS) [11] is used for both MCMC cases:

$$\boldsymbol{\theta}_p = \sqrt{1-\beta^2}\,\boldsymbol{\theta} + \beta\,\boldsymbol{\epsilon}, \tag{18}$$

where $\boldsymbol{\theta}_p$ denotes the current proposal and $\boldsymbol{\theta}$ represents the previously accepted proposal. The symbol β denotes the tuning parameter and $\boldsymbol{\epsilon}$ represents $\mathcal{N}(0,1)$-random variable.

We first discuss the algorithm of the MCMC method without conditioning [9,16]. The filtering step of this method is based on a coarse-scale model approximation of the governing Eq. (1). The coarse-scale discretization is similar to the fine-scale discretization and the permeability field $\boldsymbol{\eta}(\boldsymbol{\theta})$ is projected on the coarse-scale. An upscaling procedure [15] is used to set an effective permeability field on a coarse-grid that provides a similar average response as that of the underlying fine-scale problem. The numerical simulator is run on the coarse-scale model and produces the coarse-grid pressure field R_c. The coarse-scale and fine-scale acceptance probabilities are estimated as

Algorithm 1. MCMC with conditioning

For a given covariance function R generate KLE.
for $i = 1$ to M_{mcmc} **do**
  At $\boldsymbol{\eta}(\boldsymbol{\theta})$generate $\boldsymbol{\theta}_p$ using equation (18).
  Project $\boldsymbol{\theta}_p$ to the nullspace using (13) to get $\hat{\boldsymbol{\theta}}_p$.
  At $\hat{\boldsymbol{\theta}}_p$ construct $\boldsymbol{\eta}_p$ using equation (14).
  Compute the upscaled permeability on the coarse-scale using $\boldsymbol{\eta}_p$.
  Solve the forward problem on the coarse-scale to get R_c.
  Compute the coarse-scale acceptance probability $\alpha_c(\boldsymbol{\eta}, \boldsymbol{\eta}_p)$.
  if $\boldsymbol{\eta}_p$ is accepted **then**
    Use $\boldsymbol{\eta}_p$ in the fine-scale simulation to get R_f.
    Compute the fine-scale acceptance probability $\alpha_f(\boldsymbol{\eta}, \boldsymbol{\eta}_p)$.
    if $\boldsymbol{\eta}_p$ is accepted **then** $\boldsymbol{\eta} = \boldsymbol{\eta}_p$.
    end if
  end if
end for

$$\begin{aligned} \alpha_c(\boldsymbol{\eta}, \boldsymbol{\eta}_p) &= \min\left(1, \frac{I(\boldsymbol{\eta}|\boldsymbol{\eta}_p) P_c(\boldsymbol{\eta}_p|R_p)}{I(\boldsymbol{\eta}_p|\boldsymbol{\eta}) P_c(\boldsymbol{\eta}|R_p)}\right), \text{ and} \\ \alpha_f(\boldsymbol{\eta}, \boldsymbol{\eta}_p) &= \min\left(1, \frac{P_f(\boldsymbol{\eta}_p|R_p) P_c(\boldsymbol{\eta}|R_p)}{P_f(\boldsymbol{\eta}|R_p) P_c(\boldsymbol{\eta}_p|R_p)}\right), \end{aligned} \tag{19}$$

where P_c and P_f represent the posterior probabilities computed at coarse- and fine-scale, respectively. In the MCMC method with conditioning, we construct the permeability field $\boldsymbol{\eta}(\boldsymbol{\theta})$ using Eq. (14). The algorithm of the MCMC method without conditioning is described in [2] and the algorithm of the MCMC method with conditioning is presented in Algorithm 1.

4.3 Convergence Assessment

In this subsection, we discuss the convergence of the MCMC methods for a reliable characterization of the permeability field. Determining the stopping criterion in the MCMC simulations is a challenging task. For this reason, researchers use several MCMC convergence diagnostics to decide when it is safe to stop the simulations. Different diagnostic measures are discussed in [31,35,40,41]. A common approach is to start more than one MCMC chain using distinct initial conditions. In this paper, we study the convergence of the proposed MCMC methods using the Potential Scale Reduction Factor (PSRF) and its multivariate extension MPSRF [7]. Note that the PSRF does not take all the parameters into account and thus may not provide a very reliable indicator of the convergence of the parallel chains. However, the MPSRF takes all the parameters and their interactions into consideration. Therefore, the MPSRF is a better measure to determine the convergence of several chains that explore a high dimensional parameter space.

Let us consider l posterior draws of $\boldsymbol{\theta}$ (dimension n) for each of the $k > 1$ parallel chains. Let $\boldsymbol{\theta}_j^c$ be the generated samples at iteration c in jth chain of k parallel MCMCs. Then, the posterior variance-covariance matrix is computed by

$$\widehat{\mathbf{V}} = \frac{l-1}{l}\mathbf{W} + \left(1 + \frac{1}{k}\right)\frac{\mathbf{B}}{l}. \tag{20}$$

The within-chain covariance matrix $\mathbf{W}$ and between-chain covariance matrix $\mathbf{B}$ are defined by

$$\mathbf{W} = \frac{1}{k(l-1)} \sum_{j=1}^{k} \sum_{c=1}^{l} \left(\boldsymbol{\theta}_j^c - \bar{\boldsymbol{\theta}}_{j.}\right)\left(\boldsymbol{\theta}_j^c - \bar{\boldsymbol{\theta}}_{j.}\right)^T, \tag{21}$$

and

$$\mathbf{B} = \frac{l}{k-1} \sum_{j=1}^{k} \left(\bar{\boldsymbol{\theta}}_{j.} - \bar{\boldsymbol{\theta}}_{..}\right)\left(\bar{\boldsymbol{\theta}}_{j.} - \bar{\boldsymbol{\theta}}_{..}\right)^T, \tag{22}$$

respectively. $\bar{\boldsymbol{\theta}}_{j.}$ and $\bar{\boldsymbol{\theta}}_{..}$ represent the mean within the chain and the mean of the k combined chains, respectively, and T stands for transpose. The PSRFs are computed by

$$\mathrm{PSRF_i} = \sqrt{\frac{\mathrm{diag}(\widehat{\mathbf{V}})_i}{\mathrm{diag}(\mathbf{W})_i}}, \quad \text{where } i = 1, 2, ..., n. \tag{23}$$

The maximum of PSRF values, which is closer to 1, indicates that MCMC proposals sample from approximately the same posterior distribution. The MPSRF is defined as [32]

$$\mathrm{MPSRF} = \sqrt{\left(\frac{l-1}{l} + \left(\frac{k+1}{k}\right)\lambda\right)}, \tag{24}$$

where λ is the largest eigenvalue of the positive definite matrix $\mathbf{W}^{-1}\mathbf{B}/l$. If the proposals in k chains are sampled from almost the same distribution, the between chain covariance matrix becomes zero. Consequently, λ and the MPSRF approach to 0, and 1, respectively, for sufficiently large l. Thus, we achieve the required convergence of the chains.

5 Numerical Results

5.1 Setup of the Problem

In this subsection, we present the details of the simulation of our problem of interest. We consider a unit square-shaped physical domain and solve the elliptic Eq. (1) in the domain. The simulated elliptic solutions are used in the MCMC algorithms. We then present a comparative study between the MCMC methods with and without conditioning. In the MCMC method with conditioning, we

include the prior measurement of the permeability field at nine specific sparse locations. This permeability field is constructed using the KLE in which we consider the covariance function as

$$R(\boldsymbol{x}_1, \boldsymbol{x}_2) = \sigma_Y^2 \exp\left(-\frac{|x_1 - x_2|^2}{2L_x^2} - \frac{|y_1 - y_2|^2}{2L_y^2}\right), \tag{25}$$

where $L_x = 0.4$ and $L_y = 0.8$ are the correlation lengths and $\sigma_Y^2 = \mathrm{Var}[(Y^k)^2] = 1$. Figure 3 illustrates the decay of the eigenvalues for these values. Considering the fast decay of the eigenvalues, in this paper, we take the first twenty dominating eigenvalues in the KLE, which preserves almost 100% (using the Eq. (5)) of the total variance or energy of Y in the Eq. (6). We set the source term $f = 0$ and impose Dirichlet boundary conditions, $p = 1$ and $p = 0$, on the left and right boundaries, respectively. We set a Neumann-type boundary condition to zero (i.e., no-flow) everywhere else on the domain boundaries. In our study, we use a synthetic reference permeability field that is generated on a fine-grid of size 16×16. We then run the forward problem using the numerical simulator and generate the corresponding reference pressure field. See Fig. 2 for both fields.

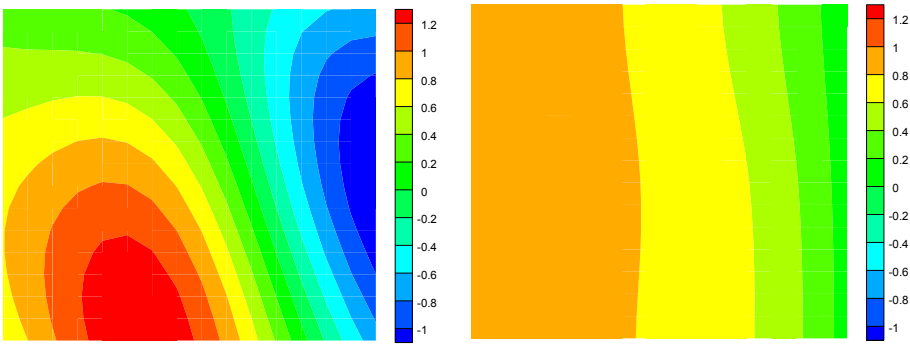

Fig. 2. Reference log permeability field (left) and the corresponding reference pressure field (right).

We take the tuning parameter $\beta = 0.85$ in (18) for both MCMC methods. As a screening step in the MCMC algorithm, we introduce a coarse-scale filter of size 8×8. The coarse-scale simulation runs almost two times faster than the corresponding fine-scale simulation while maintaining the general trend of the fine-scale simulation. We change one stochastic parameter at every iteration to construct a new permeability field for both MCMC studies.

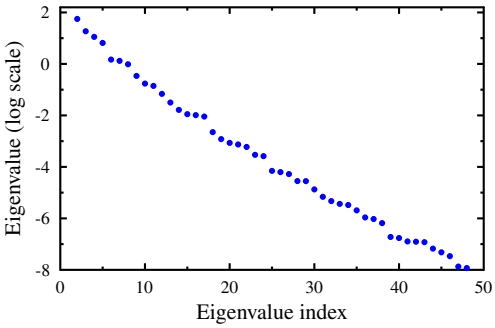

Fig. 3. Decay of eigenvalues in KLE.

5.2 Convergence Analysis of the MCMC Method

We analyze here the convergence of the MCMC algorithm for both studies in terms of the maximum of PSRFs and MPSRF. We run four MCMCs starting from distinct initial seeds for the MCMC method without conditioning. We also run another four chains, but using the same corresponding initial seeds in the former runs, for the MCMC method with conditioning. In Table 1 we show the acceptance rates for these two MCMC studies. We take $\sigma_F^2 = 10^{-4}$ and $\sigma_C^2 = 5 \times 10^{-3}$ for coarse- and fine-scale precisions, respectively, in Eq. (16).

Table 1. A comparison of acceptance rates for the MCMC with and without conditioning.

	MCMC without conditioning	MCMC with conditioning
σ_F^2	10^{-4}	10^{-4}
σ_C^2	5×10^{-3}	5×10^{-3}
Acceptance rate	53%	60%

We observe from the Table 1 that the MCMC method with conditioning has a better acceptance rate than the MCMC method without conditioning. In the calculation of the PSRFs and MPSRF, we take into account the $\boldsymbol{\theta}$ values that are accepted on the fine-scale. If there is no change in the theta values from one state to another in the Markov chain, the repetition of the accepted theta values on the fine-scale is considered in the calculation. Note that a MCMC method converges to the target distribution if both PSRFs and MPSRF values are below 1.2 [42]. Figure 4 depicts the maximum of the PSRFs and MPSRF for both cases with the same number of proposals, 26000, after burn-in. In each chain, 6500 proposals are considered. We observe that the MCMC with conditioning achieves a full

convergence. Towards the end of the curves, the maximum PSRF and MPSRF values are 1.14 and 1.16, respectively. Both PSRF and MPSRF curves for the MCMC method without conditioning are considerably away from the numerical value of 1.2. Thus, we conclude that the MCMC method with conditioning explores the posterior distribution much faster than the MCMC method without conditioning. Moreover, Fig. 4 demonstrates that the maximum of PSRFs is bounded above by the MPSRF as shown in [7]. In the next paragraph, we discuss the importance of this convergence in order to recover the underlying permeability field.

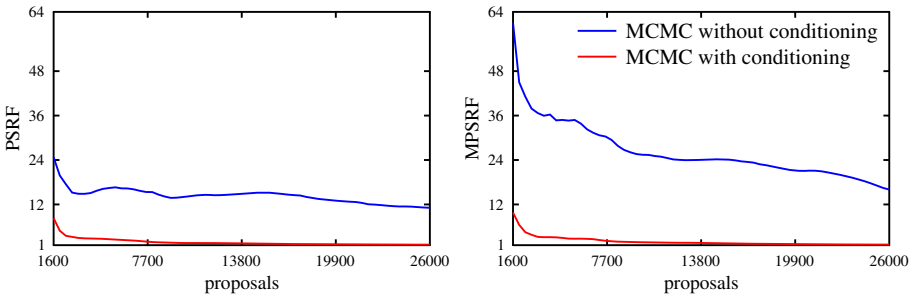

Fig. 4. The maximum of PSRFs and MPSRF for the MCMC method with and without conditioning on measured data.

Now, we discuss the permeability fields that are accepted in the MCMC methods. In Fig. 5 we show twelve permeability fields (three per chain) at 40, 5000, and 10000 iterations, respectively, in four chains for the MCMC method without conditioning. In Fig. 6 we present the corresponding permeability fields in four chains obtained from the MCMC method with conditioning. The process of constructing the permeability fields for this MCMC method with conditioning has two steps: In the first step, we create a permeability field using the KLE in which we provide the projected proposal as an input. Then we add the kriging data obtained from the measurements at nine locations (See Sect. 3). We observe that the MCMC method with conditioning recovers the reference permeability field in Fig. 2 much better than the MCMC method without conditioning. This result is consistent with the MCMC convergence discussed in the above paragraph. Thus, we conclude that the MCMC method with conditioning shows a better performance for the characterization of the permeability field in our problem.

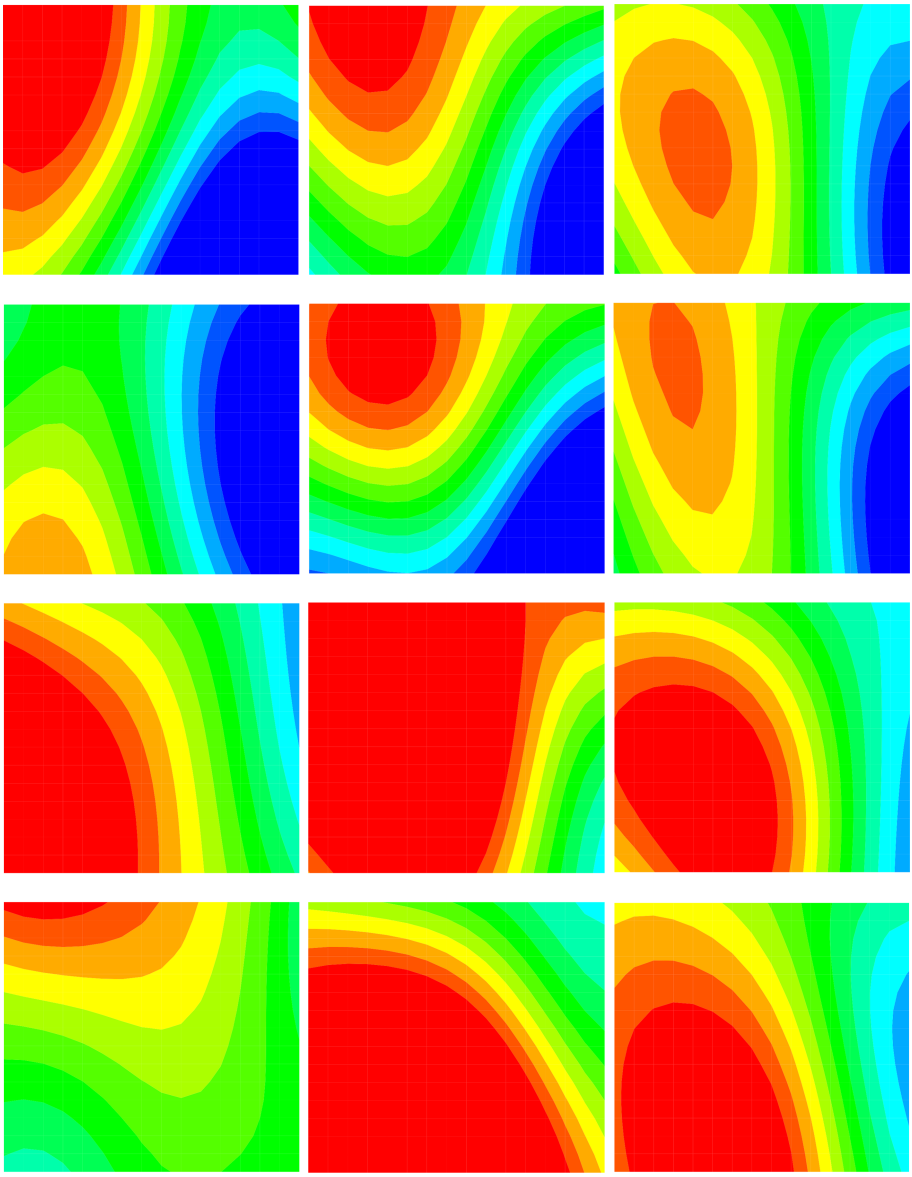

Fig. 5. Log permeability fields recovered in the MCMC without conditioning. Rows refer to four consecutive chains and columns refer to iterations 40, 5000 and 10000, respectively.

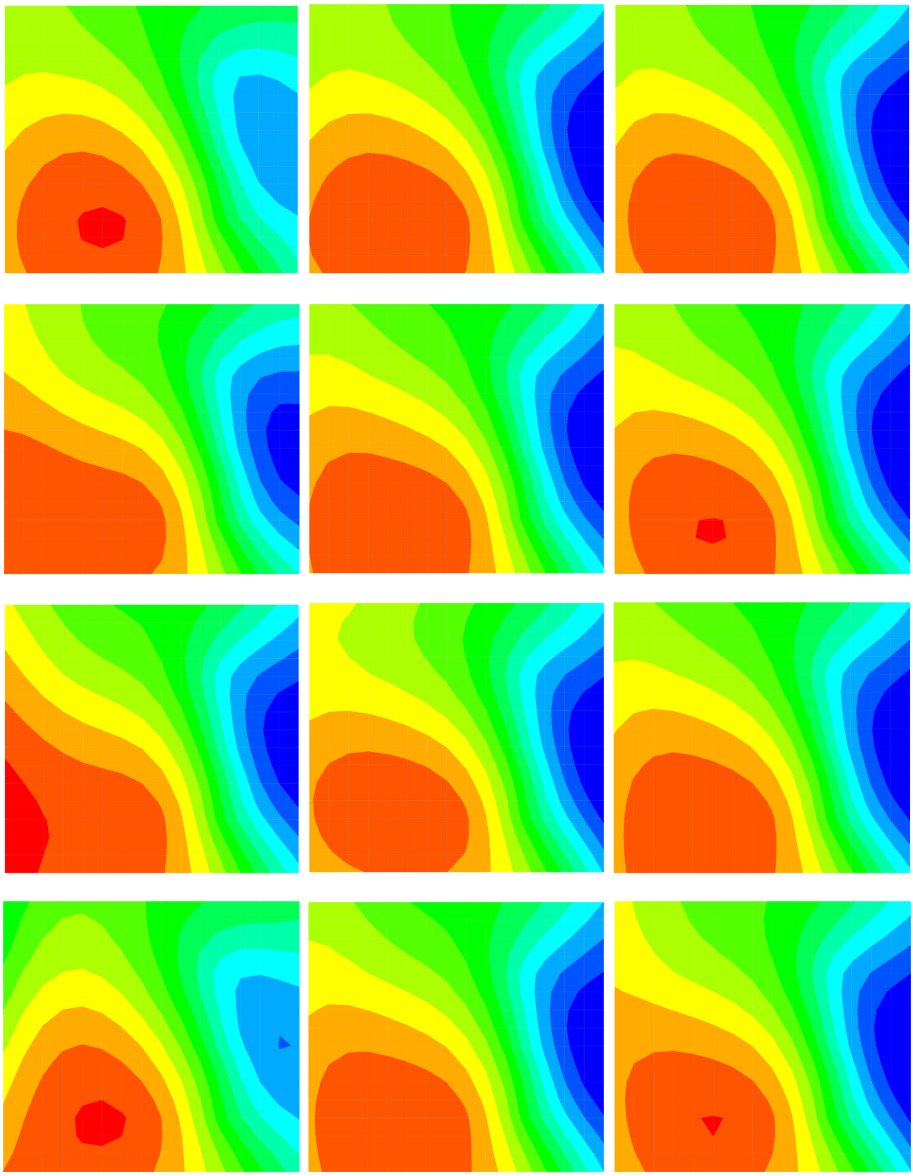

Fig. 6. Log permeability fields recovered in the MCMC with conditioning. Rows refer to four consecutive chains and columns refer to iterations 40, 5000 and 10000, respectively.

6 Conclusions

In this work, we considered the use of MCMC methods for the characterizations of subsurface formations. Usually, sparse measurements of quantities of interest

(such as the permeability field) are available in several field locations and they have to be incorporated in the characterization procedure to reduce uncertainty as much as possible.

We introduced a computationally efficient method for conditioning (the log of) permeability fields. The method is based on a projection onto the nullspace of a data matrix defined in terms of a KL expansion. We also presented multi-chain studies that illustrate the importance of conditioning in accelerating the MCMC convergence.

The authors and their collaborators are currently applying the Multiscale Perturbation Method [3,39] to further speed up MCMC studies.

Acknowledgment. A. Rahunanthan was supported by grants from the National Science Foundation (Grant No. HRD–1600818), and NIFA/USDA through Central State University's Evans-Allen Research Program (Fund No. NI201445XXXXG018-0001).

All the numerical simulations presented in this paper were performed on the GPU Computing cluster housed in the Department of Mathematics and Computer Science at Central State University.

References

1. Al-Mamun, A., Barber, J., Ginting, V., Pereira, F., Rahunanthan, A.: Contaminant transport forecasting in the subsurface using a bayesian framework. Appl. Math. Comput. **387**, 124980 (2020)
2. Ali, A., Al-Mamun, A., Pereira, F., Rahunanthan, A.: Markov chain monte carlo methods for fluid flow forecasting in the subsurface. In: Krzhizhanovskaya, V.V., et al. (eds.) ICCS 2020. LNCS, vol. 12143, pp. 757–771. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-50436-6_56
3. Ali, A., Mankad, H., Pereira, F., Sousa, F.S.: The multiscale perturbation method for second order elliptic equations. Appl. Math. Comput. **387**, 125023 (2020)
4. Borges, M.R., Furtado, F., Pereira, F., Amaral Souto, H.P.: Scaling analysis for the tracer flow problem in self-similar permeability fields. Multiscale Model. Simul. **7**(3), 1130–1147 (2008)
5. Borges, M.R., Pereira, F.: A novel approach for subsurface characterization of coupled fluid flow and geomechanical deformation: the case of slightly compressible flows. Comput. Geosci. **24**(4), 1693–1706 (2020)
6. Borges, M.R., Pereira, F., Amaral Souto, H.P.: Efficient generation of multi-scale random fields: a hierarchical approach. Int. J. Numer. Methods Biomed. Eng. **26**(2), 176–189 (2010)
7. Brooks, S., Gelman, A.: General methods for monitoring convergence of iterative simulations. J. Comput. Graph. Stat. **7**, 434–455 (1998)
8. Brooks, S., Gelman, A., Jones, G., Meng, X.L.: Handbook of Markov chain Monte Carlo. CRC Press, Boca Raton (2011)
9. Christen, J.A., Fox, C.: Markov chain Monte Carlo using an approximation. J. Comput. Graph. Stat. **14**(4), 795–810 (2005)
10. Constantine, P.G., Kent, C., Bui-Thanh, T.: Accelerating Markov chain Monte Carlo with active subspaces. SIAM J. Sci. Comput. **38**(5), A2779–A2805 (2016)
11. Cotter, S.L., Roberts, G.O., Stuart, A.M., White, D.: MCMC methods for functions: modifying old algorithms to make them faster. Stat. Sci. **28**(3), 424–446 (2013)

12. Delhomme, J.P.: Kriging in the hydrosciences. Adv. Water Resour. **1**(5), 251–266 (1978)
13. Deutsch, C.V., Journel, A.G., et al.: Geostatistical software library and user's guide. New York **119**(147) (1992)
14. Douglas, J., Furtado, F., Pereira, F.: On the numerical simulation of waterflooding of heterogeneous petroleum reservoirs. Comput. Geosci. **1**(2), 155–190 (1997)
15. Durlofsky, L.: Numerical calculation of equivalent grid block permeability tensors for heterogeneous porous media. Water Resour. Res. **27**(5), 699–708 (1991)
16. Efendiev, Y., Hou, T., Luo, W.: Preconditioning Markov Chain Monte Carlo simulations using coarse-scale models. SIAM J. Sci. Comput. **28**, 776–803 (2006)
17. Francisco, A., Ginting, V., Pereira, F., Rigelo, J.: Design and implementation of a multiscale mixed method based on a nonoverlapping domain decomposition procedure. Math. Comput. Simul. **99**, 125–138 (2014)
18. Furtado, F., Pereira, F.: Crossover from nonlinearity controlled to heterogeneity controlled mixing in two-phase porous media flows. Comput. Geosci. **7**(2), 115–135 (2003)
19. Ginting, V., Pereira, F., Rahunanthan, A.: Multiple Markov chains Monte Carlo approach for flow forecasting in porous media. Procedia Comput. Sci. **9**, 707–716 (2012)
20. Ginting, V., Pereira, F., Rahunanthan, A.: A multi-stage Bayesian prediction framework for subsurface flows. Int. J. Uncertain. Quantif. **3**(6), 499–522 (2013)
21. Ginting, V., Pereira, F., Rahunanthan, A.: A prefetching technique for prediction of porous media flows. Comput. Geosci. **18**(5), 661–675 (2014)
22. Ginting, V., Pereira, F., Rahunanthan, A.: Rapid quantification of uncertainty in permeability and porosity of oil reservoirs for enabling predictive simulation. Math. Comput. Simul. **99**, 139–152 (2014)
23. Ginting, V., Pereira, F., Rahunanthan, A.: Multi-physics Markov chain Monte Carlo methods for subsurface flows. Math. Comput. Simul. **118**, 224–238 (2015)
24. Glimm, J., Lindquist, W.B., Pereira, F., Peierls, R.: The multi-fractal hypothesis and anomalous diffusion. Mat. Aplic. Comp. **11**(2), 189–207 (1992)
25. Glimm, J., Lindquist, W.B., Pereira, F., Zhang, Q.: A theory of macrodispersion for the scale-up problem. Comput. Geosci. **13**(1), 97–122 (1993)
26. Hewett, T.A., Behrens, R.A.: Conditional simulation of reservoir heterogeneity with fractals. SPE Form. Eval. **5**(03), 217–225 (1990)
27. Journel, A.G.: Huijbregts: Mining eostatistics. Academic Press, Technical report (1978)
28. Laloy, E., Rogiers, B., Vrugt, J., Mallants, D., Jacques, D.: Efficient posterior exploration of a high-dimensional groundwater model from two-stage Markov chain Monte Marlo simulation and polynomial chaos expansion. Water Resour. **49**(5), 2664–2682 (2013)
29. Loève, M.: Probability Theory. Springer, Berlin (1997). https://doi.org/10.1007/978-1-4684-9464-8
30. Lu, Z., Zhang, D.: Conditional simulations of flow in randomly heterogeneous porous media using a KL-based moment-equation approach. Adv. Water Resour. **27**(9), 859–874 (2004)
31. Cowles, M.K., Carlin, B.P.: Markov chain Monte Carlo convergence diagnostics: a comparative review. J. Am. Stat. Assoc. **91**, 883–904 (1996)
32. Malyshkina, N.: Markov Switching Models: An Application to Roadway Safety. Ph.D. thesis, Purdue University (2008)

33. Mamun, A., Pereira, F., Rahunanthan, A.: Convergence analysis of MCMC methods for subsurface flow problems. In: Gervasi, O., Murgante, B., Misra, S., Stankova, E., Torre, C.M., Rocha, A.M.A.C., Taniar, D., Apduhan, B.O., Tarantino, E., Ryu, Y. (eds.) ICCSA 2018. LNCS, vol. 10961, pp. 305–317. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-95165-2_22
34. Martin, J., Wilcox, L.C., Burstedde, C., Ghattas, O.: A stochastic Newton MCMC method for large-scale statistical inverse problems with application to seismic inversion. SIAM J. Sci. Comput. **34**(3), A1460–A1487 (2012)
35. Mengersen, K.L., Robert, C.P., Guihenneuc-Jouyaux, C.: MCMC convergence diagnostics: a review. In: Bernardo, M., Berger, J.O., Dawid, A.P., Smtith, A.F.M. (eds.) Bayesian Statistics, vol. 6, pp. 415–440. Oxford University Press, Oxford (1999)
36. Mondal, A., Efendiev, Y., Mallick, B., Datta-Gupta, A.: Bayesian uncertainty quantification for flows in heterogeneous porous media using reversible jump Markov chain Monte Carlo methods. Adv. Water Resour. **33**(3), 241–256 (2010)
37. Ossiander, M., M.Peszynska, Vasylkivska, V.: Conditional stochastic simulations of flow and transport with Karhunen-Loève expansions, stochastic collocation, and sequential Gaussian simulation. J. Appl. Math. **2014** (2014)
38. Pereira, F., Rahunanthan, A.: A semi-discrete central scheme for the approximation of two-phase flows in three space dimensions. Math. Comput. Simul. **81**(10), 2296–2306 (2011)
39. Rocha, F.F., Mankad, H., Sousa, F.S., Pereira, F.: The multiscale perturbation method for two-phase reservoir flow problems. arXiv preprint arXiv:2103.11050 (2021)
40. Roy, V.: Convergence diagnostics for Markov chain Monte Carlo. Ann. Rev. Stat. Appl. **7**, 387–412 (2019)
41. Brooks, S.P., Roberts, G.O.: Convergence assessments of Markov chain Monte Carlo algorithms. Stat. Comput. **8**, 319–335 (1998)
42. Smith, B.J.: boa: An R package for MCMC output convergence assessment and posterior inference. J. Stat. Softw. **21**, 1–37 (2007)
43. Strang, G.: Linear Algebra and Learning from Data. Wellesley-Cambridge Press, Cambridge (2019)
44. Stuart, G.K., Minkoff, S.E., Pereira, F.: A two-stage Markov chain Monte Carlo method for seismic inversion and uncertainty quantification. Geophysics **84**(6), R1003–R1020 (2019)
45. Tong, X.T., Morzfeld, M., Marzouk, Y.M.: MALA-within-Gibbs samplers for high-dimensional distributions with sparse conditional structure. SIAM J. Sci. Comput. **42**(3), A1765–A1788 (2020)
46. Tong, Y.L.: The Multivariate Normal Distribution. Springer Series in Statistics, Springer, New York (1990). https://doi.org/10.1007/978-1-4613-9655-0