

The stochastic counterpart of conservation laws with heterogeneous conductivity fields: Application to deterministic problems and uncertainty quantification



Amir H. Delgosaie^a, Peter W. Glynn^b, Patrick Jenny^c, Hamdi A. Tchelepi^{a,*}

^a Department of Energy Resources Engineering, Stanford University, 367 Panama Street, Stanford, CA, 94305, USA

^b Department of Management Science and Engineering, Stanford University, 475 Via Ortega, Stanford, CA, 94305, USA

^c Institute of Fluid Dynamics, ETH Zürich, Sonneggstrasse 3, CH-8092 Zürich, Switzerland

ARTICLE INFO

Article history:

Received 8 June 2018

Received in revised form 15 November 2018

Accepted 20 December 2018

Available online 27 February 2019

Keywords:

Heterogeneous diffusion

Stochastic modeling

Backward equations

Stochastic PDE

ABSTRACT

Conservation laws in the form of elliptic and parabolic partial differential equations (PDEs) are fundamental to the modeling of many problems such as heat transfer and flow in porous media. Many of such PDEs are stochastic due to the presence of uncertainty in the conductivity field. Based on the relation between stochastic diffusion processes and PDEs, Monte Carlo (MC) methods are available to solve these PDEs. These methods are especially relevant for cases where we are interested in the solution in a small subset of the domain. The existing MC methods based on the stochastic formulation require restrictively small time steps for high-variance conductivity fields. Moreover, in many applications the conductivity is piecewise constant and the existing methods are not readily applicable in these cases. Here we provide an algorithm to solve one-dimensional elliptic problems that bypasses these two limitations. The methodology is demonstrated using problems governed by deterministic and stochastic PDEs. It is shown that the method provides an efficient alternative to compute the statistical moments of the solution to a stochastic PDE at any point in the domain. A variance reduction scheme is proposed for applying the method for efficient mean calculations.

© 2019 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Elliptic and parabolic partial differential equations (PDEs) arise in many applications in describing conservation laws. The mass conservation equation resulting from Fick's law, the heat equation, and the pressure equation in the context of flow in porous media are some of the prominent examples of such applications. These equations have the following general form

$$\nabla \cdot (K \nabla p) = c \frac{\partial p}{\partial t}, \quad (1)$$

when the problem is time dependent and

$$\nabla \cdot (K \nabla p) = 0, \quad (2)$$

* Corresponding author.

E-mail addresses: amirdel@stanford.edu (A.H. Delgosaie), glynn@stanford.edu (P.W. Glynn), jenny@ifd.mavt.ethz.ch (P. Jenny), tchelepi@stanford.edu (H.A. Tchelepi).

when the problem is at steady state. Here, p is the unknown (concentration, temperature, or pressure) and K is the conductivity tensor. In these problems, the flux (of mass or heat) is governed by a Fick's type law, i.e.,

$$q = -K\nabla p. \quad (3)$$

In many practical settings, closed-form solutions of PDEs (1) and (2) do not exist, and numerical methods such as finite-volume (FV) and finite-element methods are used to compute numerical solutions of these PDEs [1]. These numerical methods rely on discretization of the PDE for the domain of interest and deriving a set of linear equations for the solution of the PDE on the discretized grid. Once the system is reduced to a linear system, efficient numerical methods can be employed to solve it [2]. These linear solvers compute the solution for all grid points simultaneously.

In stochastic modeling, elliptic and parabolic PDEs that are similar to equations (1) and (2) arise when calculating various expected values for a stochastic diffusion process [3]. In the stochastic modeling nomenclature, these PDEs are referred to as backward equations, and the differential operator describing the left-hand-side of equations (1) and (2) is referred to as \mathcal{L} . In the view of the connection between diffusion processes and PDEs (the Feynman-Kac formulation), the stochastic counterpart of \mathcal{L} has been used in methods such as backward walks and random walks on boundary [4,5] to solve elliptic and parabolic PDEs. Moreover, methods based on the Feynman-Kac formulation are available that can handle Dirichlet, Neumann, and Robin boundary conditions [6–8].

Recognizing that equations (1) and (2) correspond to backward equations of specific stochastic processes has multiple advantages. First, unlike solving linear systems, by using the stochastic representation of the problem, the solution for any subset of points in the domain can be found independently of the solution at points outside the subset of interest. Second, the numerical solution can be computed at any point in the domain without the need for a mesh. Third, the stochastic solution strategy is 'embarrassingly parallel', which allows for efficient implementations that can take full advantage of GPUs and massively parallel CPUs.

In many applications, such as flow in natural porous formations, the conductivity field, K , is highly heterogeneous. This heterogeneity poses additional challenges for using the stochastic counterpart of \mathcal{L} in Monte Carlo algorithms. At the same time, there is often uncertainty associated with K . The advantages of stochastic formulations become more relevant in the case where K is modeled as a random field. The most common way to find the solution under uncertainty in the conductivity field is to solve the equations numerically for an ensemble of conductivity realizations to compute the statistical moments, or the distribution, of the unknown variable. This can be a very expensive computational procedure. Using the stochastic formulation, the moments of the solution or the one-point probability density function (PDF) at any location in the domain is obtained by computing the solution only at the point of interest - independently of other points - for the different realizations of the conductivity field.

Anker et al. [9] recently used the Feynman-Kac formulation to solve elliptic problems in heterogeneous conductivity fields. They also proposed a method to efficiently find the mean solution in random heterogeneous domains. In the stochastic counterpart of \mathcal{L} for heterogeneous conductivity fields, the gradient of conductivity appears as the drift function, so the method is valid only for smooth conductivity fields. This is a good assumption in many scenarios where K is indeed sufficiently smooth or it can be well approximated by smooth functions. This is true both for deterministic and stochastic K , where the conductivity can be represented by truncated Fourier series or Karhunen-Loeve (KL) expansions [10].

There are additional challenges remaining for using the stochastic counterpart of equations (1) and (2). In many practical applications the conductivity field is piecewise constant and the method proposed in [9] would not be applicable in those settings. Moreover, even when solving elliptic problems with the stochastic formulation, the location of the computational particles are incremented in pseudo time steps. For highly heterogeneous conductivity fields which are common in porous media applications, the method would be inaccurate unless for very small time steps. In general, the correlation length and the variance of the conductivity field determine the right time step size. For some cases this would render the stochastic approach computationally too expensive.

In this work, we first review the correspondence between \mathcal{L} , the operator in the backward equations of diffusion processes, and the differential operator in the steady-state and transient conservation equations with heterogeneous coefficients. Next, a stochastic algorithm is proposed to solve the elliptic problem in one dimension that would be valid for piecewise constant conductivity fields. We provide a heuristic proof of the method based on the exit probability of diffusion processes. A rigorous proof of the proposed method can be provided based on skew Brownian motion, similar to [11,12]. In one dimension, the proposed method is exact and is not sensitive to the variance or the correlation length of the conductivity field.

The paper is organized as follows. In section 2 the Feynman-Kac formulation for heterogeneous media is reviewed. Potential issues for using the stochastic formulation in high-variance conductivity fields are discussed in section 3. The stochastic algorithm for solving elliptic PDEs with piecewise constant coefficients is presented in section 4 and a deterministic example is provided in section 5. Various examples are provided for using the proposed stochastic algorithm for uncertainty quantification in section 6. A variance reduction scheme for calculating the mean solution to a stochastic PDE is discussed. Finally, conclusions and possible extensions of the algorithm are discussed in section 7.

2. Conservation laws versus backward equations

Consider $X(t)$ that satisfies the following stochastic differential equation (SDE)

$$dX(t) = \mu(X(t))dt + \sigma(X(t))dB(t). \tag{4}$$

We are interested in computing various expected values of functions of $X(t)$ conditional on starting the process at $X(0) = x$. It is standard to adapt the following notation

$$E_x[\cdot] = E[\cdot | X(0) = x]. \tag{5}$$

As it is shown in [3,13],

$$u^*(t, x) = E[r(X(t)) | X(0) = x] = E_x[r(X(t))], \tag{6}$$

which is the expected value of $r(X(t))$ if we start from $X(0) = x$, can be found by solving the following parabolic PDE

$$\begin{aligned} \mathcal{L}u^*(x, t) &= \frac{\partial u^*(x, t)}{\partial t} \\ u^*(x, 0) &= r(x). \end{aligned} \tag{7}$$

In equation (7) the differential operator \mathcal{L} is defined as follows

$$\mathcal{L} = \sum_{i=1}^d \mu_i(x) \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^d b_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j}. \tag{8}$$

Here $b_{ij}(x)$ refers to the elements of $\sigma(x)\sigma(x)^T$. In these equations, $X(t), \mu(X(t)) \in \mathbb{R}^d$ and $\sigma(X(t)) \in \mathbb{R}^{d \times d}$. $B(t) \in \mathbb{R}^d$ is a d -dimensional Brownian motion. To be concrete, in two dimensions

$$\begin{bmatrix} b_{11}(x) & b_{12}(x) \\ b_{21}(x) & b_{22}(x) \end{bmatrix} = \sigma(x)\sigma^T(x) = \begin{bmatrix} \sigma_{11}^2(x) + \sigma_{12}^2(x) & \sigma_{11}(x)\sigma_{21}(x) + \sigma_{12}(x)\sigma_{22}(x) \\ \sigma_{11}(x)\sigma_{21}(x) + \sigma_{12}(x)\sigma_{22}(x) & \sigma_{22}^2(x) + \sigma_{21}^2(x) \end{bmatrix}. \tag{9}$$

As shown in [13], the rigorous connection between SDEs and PDEs involves the use of Ito's formula. In order to apply Ito's formula to show the connection between SDEs and PDEs, the solution to the PDE needs to be continuously differentiable in time, and piecewise twice continuously differentiable in the spatial variable. In this section we assume that the solutions to the considered PDEs satisfy the necessary conditions for applying Ito's formula.

The connection between equations (4) and (7) makes it possible to solve a PDE similar to equation (7) using MC simulation of its SDE counterpart. The algorithm is as follows: to find $u^*(x, t)$, launch random walks from x at time t ; evolve them according to equation 4 for t time units backward in time (to $t = 0$), and store $r(X(0))$. The solution is then the average of the stored values $r(X(0))$. For homogeneous isotropic conductivity fields ($K_{ij}(x) = K\delta_{ij}$), this method is referred to as 'backward walks' [14].

Similarly, the expected value

$$u^*(x) = E_x[r(X(T))], \tag{10}$$

where $T = \inf\{t \geq 0 : X(t) \in C^c\}$ is the first hitting time of C^c (complement of set C), can be found by solving the following elliptic PDE

$$\begin{aligned} \mathcal{L}u^*(x) &= 0, \quad x \in C \\ u^*(x) &= r(x), \quad x \in C^c. \end{aligned} \tag{11}$$

Hence, to solve a PDE similar to equation (11), one can use MC simulations of equation (4). The algorithm is as follows: to find $u^*(x)$, start many random walks from x , and evolve them according to equation (4) until they hit the boundary for the first time, then store the boundary value at the hitting times. The solution can be obtained by averaging these boundary values. For cases where the conductivity field is homogeneous, this method has been developed and is referred to as random walk on boundary [5].

In order to use "backward walks" and "random walks on boundary" for problems where the conductivity field is heterogeneous, we need to find the stochastic counterpart of the differential operator in equations (1) and (2). In the following section, we illustrate this stochastic representation by expanding the conservation equations and comparing the expanded operator with \mathcal{L} . The comparison is performed for a two-dimensional system; however, the argument can readily be extended to n dimensions.

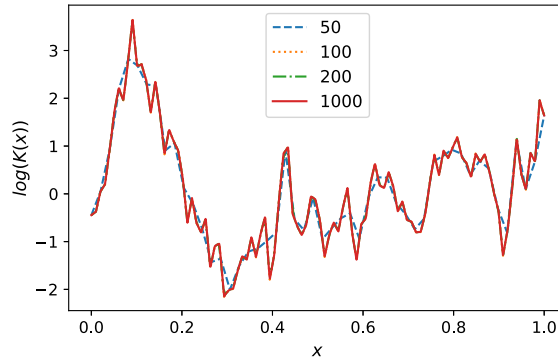


Fig. 1. One realization of $\log(K(x))$ described in section 3, evaluated with four different resolutions.

2.1. Comparison of the differential operators

Since the linear operator for both elliptic and parabolic problems is the same, we focus on the elliptic problem. Assuming K is differentiable, expansion of equation (2) leads to

$$\frac{\partial K_{11}}{\partial x_1} \frac{\partial p}{\partial x_1} + \frac{\partial K_{22}}{\partial x_2} \frac{\partial p}{\partial x_2} + \frac{\partial K_{12}}{\partial x_1} \frac{\partial p}{\partial x_2} + \frac{\partial K_{21}}{\partial x_2} \frac{\partial p}{\partial x_1} + K_{11} \frac{\partial^2 p}{\partial x_1^2} + K_{22} \frac{\partial^2 p}{\partial x_2^2} + K_{12} \frac{\partial^2 p}{\partial x_1 \partial x_2} + K_{21} \frac{\partial^2 p}{\partial x_2 \partial x_1} = 0. \quad (12)$$

Equation (12) can be written as $\mathcal{L}_1 p = 0$, where

$$\mathcal{L}_1 = \left(\left(\frac{\partial K_{11}}{\partial x_1} + \frac{\partial K_{21}}{\partial x_2} \right) \frac{\partial}{\partial x_1} + \left(\frac{\partial K_{22}}{\partial x_2} + \frac{\partial K_{12}}{\partial x_1} \right) \frac{\partial}{\partial x_2} \right) + \left(K_{11} \frac{\partial^2}{\partial x_1^2} + K_{22} \frac{\partial^2}{\partial x_2^2} + K_{12} \frac{\partial^2}{\partial x_1 \partial x_2} + K_{21} \frac{\partial^2}{\partial x_2 \partial x_1} \right). \quad (13)$$

The first part of \mathcal{L}_1 matches the drift term of the backward operator. Note that the derivatives of the conductivity field constitute the drift term, or the preferential direction for the random walks. In two dimensions, comparison of b_{ij} and the second-order term in \mathcal{L}_1 yields

$$\begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \sigma_{11}^2 + \sigma_{12}^2 & \sigma_{11}\sigma_{21} + \sigma_{12}\sigma_{22} \\ \sigma_{11}\sigma_{21} + \sigma_{12}\sigma_{22} & \sigma_{22}^2 + \sigma_{21}^2 \end{bmatrix}. \quad (14)$$

System (14) consists of four equations and four unknowns. For symmetric K , this system can easily be solved to find σ_{ij} corresponding to \mathcal{L}_1 .

3. Challenges in using the stochastic formulation for high variance conductivity fields

Here, we consider a one-dimensional case, where $K(x)$ is one realization of a conductivity field with a log-normal distribution with mean zero and a variance of four ($\log(K) \sim \mathcal{N}(0, 4)$), with an exponential covariance function. Such conductivity fields are very common in porous media applications. The domain is $D = [0, 1]$ and the boundary conditions are $p(0) = 1$ and $p(1) = 0$. The dimensionless correlation length (l_γ) of $K(x)$ is equal to 0.05. We draw this realization using the Karhunen-Loeve (KL) expansion of the conductivity field. By truncating the KL expansion after 91 terms, to capture 99% of the energy of the field, we ensure that the conductivity is differentiable and the formulation discussed in the previous section is applicable. Fig. 1 shows one realization of $\log(K)$ generated from this KL expansion with multiple different resolutions. The code available at [15] was used for generating the KL expansion.

Fig. 2 shows the realization of the permeability field used in this example and its derivative. To ensure that we have resolved the permeability field sufficiently, $K(x)$ was evaluated at $n = 1000$ equidistant points. This example illustrates a potential challenge for solving the elliptic problem with the Feynman-Kac formulation even when the conductivity is differentiable. The derivative of $K(x)$ is a drift velocity. Starting at a point with a very high value of the drift function, the time step should be selected such that the particle can still see the variations in K . More specifically, one should resolve the smallest wavelength in the truncated KL expansion, which typically requires resolving a length scale $\ll l_\gamma$. The distribution of the time step that would satisfy $\frac{dK}{dx} dt < 0.1l_\gamma$ is shown in Fig. 3. This figure illustrates that adaptive time stepping or a very small constant time step is required for ensuring that a particle sees the variations in K as it travels through the domain. The time step restrictions would become more strict with higher variance and smaller correlation length of $\log(K)$. Choosing the right time step would also require the characterization on dK/dx .

Moreover, in many practical applications a piecewise constant conductivity is available on a grid. In these scenarios K is no longer differentiable and the methods proposed in [9] are not readily applicable. In the next section we propose an algorithm for solving the one-dimensional elliptic problem that bypasses the restrictions on dt and extends the application of the stochastic formulation to problems with piecewise constant K .

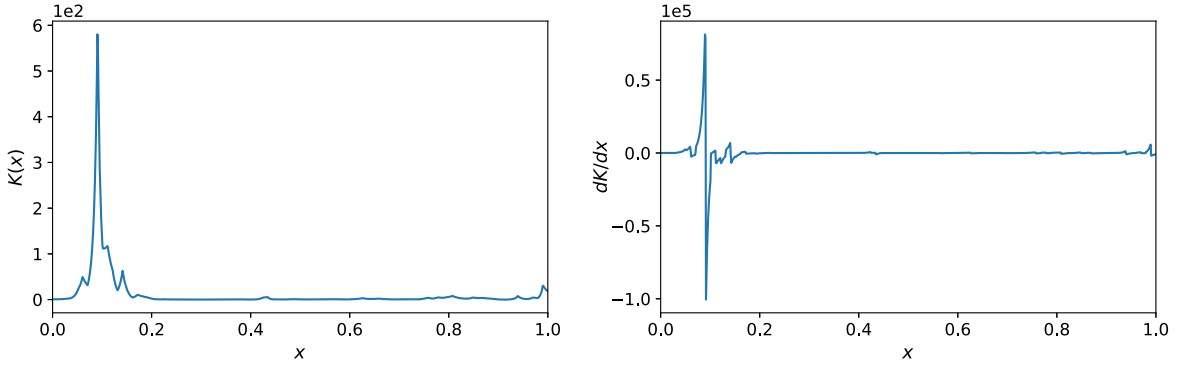


Fig. 2. One realization of the permeability field described in section 3 along with its derivative. The derivative is used as the drift term in the stochastic counterpart of the elliptic problem.

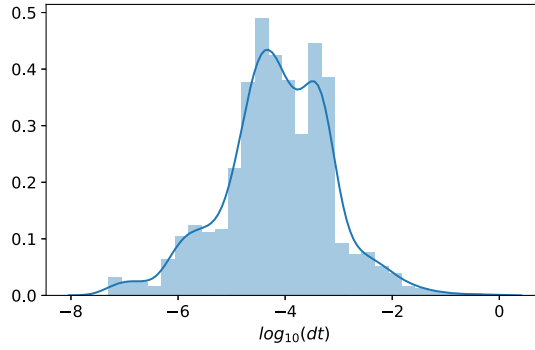


Fig. 3. Distribution of dt that would satisfy $\frac{dK}{dx} dt < 0.11y$.

4. An algorithm for solving elliptic problems with piecewise constant K in one dimension

Consider a 1D diffusion of the form

$$dX(t) = \mu(X(t))dt + \sigma(X(t))dB(t). \tag{15}$$

As shown in [13], starting at point x , the probability of exiting an interval through its right boundary before the left boundary is

$$P(T_r < T_l) = \frac{u(x) - u(l)}{u(r) - u(l)}, \tag{16}$$

where

$$u(x) = \int_0^x \exp\left(-\int_0^y \frac{2\mu(z)}{\sigma^2(z)} dz\right) dy. \tag{17}$$

From section 2, when K is differentiable, we know that $\mu(z) = dK(z)/dz$ and $\sigma(z) = \sqrt{2K(z)}$. Substituting into equation (17) we obtain

$$\begin{aligned} u(x) &= \int_0^x \exp\left(-\int_0^y \left\{\frac{d}{dz} \log(k(z))\right\} dz\right) dy = \int_0^x \exp\left(-\log(K(y)) + \log(K(0))\right) dy = \\ &= K(0) \int_0^x \exp\left(\log(1/K(y))\right) dy = K(0) \int_0^x \frac{1}{K(y)} dy. \end{aligned} \tag{18}$$

This shows that knowing K is sufficient for calculating the exit probability and dK/dx does not appear in the final result.

Using equation (18) we can find the exit probability for a two cell problem, where the permeability of the left cell and the right cell are given by K_l and K_r . We can repeat the calculations in (18) with a smooth approximation of the

conductivity, such that the drift function is defined everywhere in the two cell problem. Starting from the interface of the two cells, the probability of exiting from the right boundary is equal to

$$P(T_r < T_l) = \frac{1/K_l}{1/K_l + 1/K_r}. \quad (19)$$

For a two cell problem with piecewise constant conductivity the presented proof is not rigorous. This is the case since the conditions necessary for applying Ito's formula and establishing the relation between the SDE and the PDE are not satisfied (see section 2). A rigorous proof in that case can be provided by using skew Brownian motion and following an argument similar to [11,12].

From the stochastic formulation of the flow problem we know the pressure solution at the interface, $p(x_l)$, is the expected value of the boundary condition at the hitting time of the boundary (see equation 10). For a two cell problem with boundary conditions p_l and p_r , the interface pressure would be

$$p(x_l) = E_{x_l}[p_T] = p_r \frac{1/K_l}{1/K_l + 1/K_r} + p_l \frac{1/K_r}{1/K_l + 1/K_r} = \frac{K_l p_l + K_r p_r}{K_l + K_r}. \quad (20)$$

Using a finite-volume method we arrive at the same solution for $p(x_l)$. In short, one can calculate the flux going through the interface by using the conductivity of the left cell and the right cell. The flux going through the interface is equal to

$$q = K_l \frac{p_l - p(x_l)}{\Delta x} = K_r \frac{p(x_l) - p_r}{\Delta x} \quad (21)$$

from which we can find the interface pressure

$$p(x_l) = \frac{K_l p_l + K_r p_r}{K_l + K_r}. \quad (22)$$

Following these observations we propose Algorithm 1 for finding the solution to a one-dimensional elliptic problem with piecewise constant coefficients. Since the method is exact, the only source of error in Algorithm 1 is the Monte Carlo error related to the number of trajectories starting from the point where we seek the solution.

```

s = 0
for i ← 1 to M do
  x = x0
  // if x is not a discontinuity point
  if x ∉ Xdiscontinuity then
    // Move the particle to one of the two closest discontinuity points with the corresponding
    // probability. xl and xr are the closest discontinuity points to the left and right of x.
    Pl = (xr - x)/(xr - xl)
    // draw a uniform random variable α ∈ (0,1)
    if α < Pl then x = xl
    else x = xr
  end
  Exit = checkExit(x)
  while not Exit do
    // advance the particle to the next discontinuity point
    al = (x - xl)/Kl
    ar = (xr - x)/Kr
    Pr = al/(ar + al)
    // draw a uniform random variable α ∈ (0,1)
    if α < Pr then x = xr
    else x = xl
    Exit = checkExit(x)
  end
  // add the value of the boundary condition to s
  s = s + BC(x)
end
return s/M

```

Algorithm 1: Algorithm for solving the elliptic problem with piecewise constant K .

5. An illustrative deterministic example

In this section, an example of using the proposed stochastic method for solving deterministic problems is provided. We consider one realization of a Gaussian conductivity field where $\log(K) \sim \mathcal{N}(0, 4)$ with an exponential correlation function with $l_y = 0.25$. Unlike section 3, the conductivity realization used here is a piecewise constant field generated by the

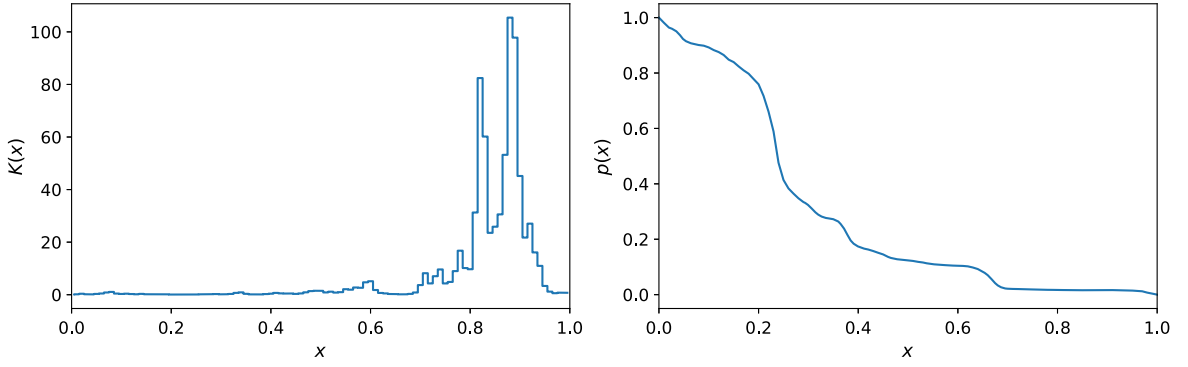


Fig. 4. Conductivity field and the corresponding pressure solution for the example in section 5.

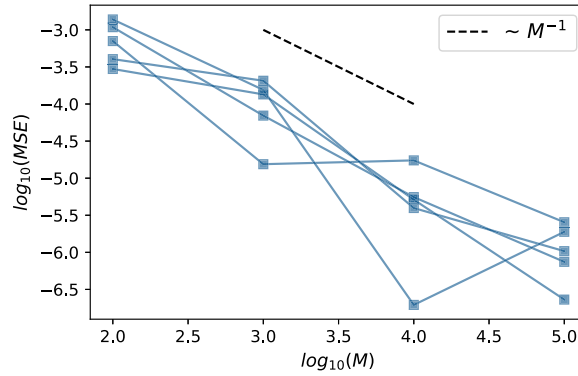


Fig. 5. Convergence of the solution obtained by Algorithm 1 to the analytical solution for the example in section 5. Different lines correspond to different experiments.

Fourier integral method [16]. The reference solution, $p^*(x)$, for a one-dimensional elliptic problem can be found analytically by integrating the one-dimensional version of equation (2):

$$\begin{aligned} \frac{d}{dx} \left(K(x) \frac{dp}{dx} \right) &= 0, \quad x \in (0, 1) \\ \text{s.t. } p(0) &= 1, \quad p(1) = 0 \end{aligned} \quad (23)$$

$$p^*(x) = 1 - \left(\int_0^1 \frac{dx}{K(x)} \right)^{-1} \int_0^x \frac{dx}{K(x)}.$$

The conductivity realization and the corresponding pressure solution are illustrated in Fig. 4.

The convergence of the MC solution obtained by using Algorithm 1 to the analytic solution is shown in Fig. 5 for different numbers of trajectories released per point. The MC solution is calculated at all points $x \in S = \{0.25, 0.5, 0.75\}$. The mean square difference is defined as

$$MSE = \frac{1}{|S|} \sum_{x \in S} (p(x) - p^*(x))^2. \quad (24)$$

Since the method is exact the only source of error is the stochastic error due to the number of trajectories followed per solution point. This is consistent with the M^{-1} scaling of the error, where M is the number of random walks per point in S .

6. Illustrative examples for uncertainty quantification

Building on Algorithm 1, in this section we use “backward walks on boundary” to quantify the uncertainty in the solution of elliptic PDEs with a random heterogeneous piecewise constant conductivity. Algorithm 2 shows the modifications for uncertainty quantification.

In the examples provided in this section the conductivity field has an exponential correlation structure with $l_Y = 0.25$. The domain is $D = [0, 1]$ and the boundary conditions are $p(0) = 1$ and $p(1) = 0$. Different realizations of the described

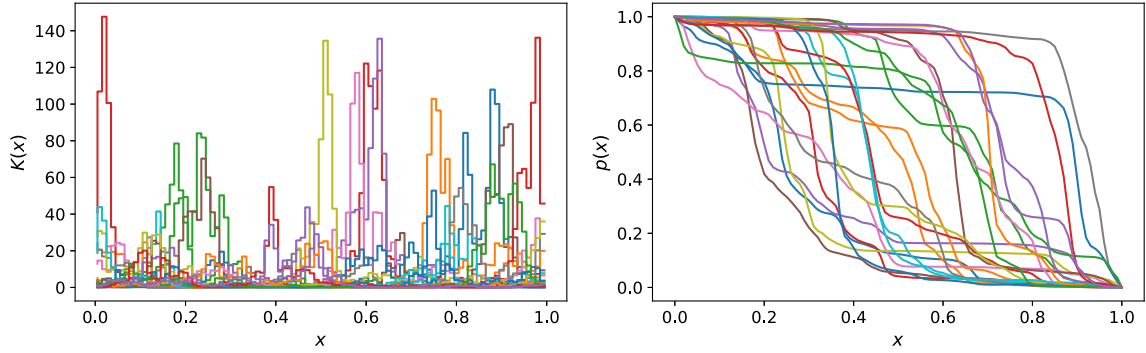


Fig. 6. Conductivity realizations sampled from the distribution described in section 6 and their corresponding analytical pressure solution.

conductivity field were generated and the flow equation was solved using analytic integration (equation (23)) for all realizations. Fig. 6 illustrates a number of these realizations and their corresponding pressure solution. The analytical solution for an ensemble of 100,000 realizations is used as the reference solution in the following examples.

```

// Initialization
S=[]
// Looping over different realizations of K
for i ← 1 to N do
  K = generateField()
  // Sample from solution distribution using algorithm 1
  p = solveMC(x0, K, M)
  S.store(p)
end
// return desired statistics (e.g. mean, histogram) using the estimated distribution of the solution
return desiredStatistic(S)

```

Algorithm 2: Uncertainty quantification using Algorithm 1.

6.1. Estimating the one-point distribution

In Fig. 7 the one-point histograms of $p(x)$ at $x = 0.25$ generated by Algorithm 2 are compared with the reference histogram. For generating the histograms in Fig. 7, $N = 100,000$ realizations were used in Algorithm 2, and M (the number of random walks followed in each realization) was varied between 10 and 1000. The histogram obtained from the analytical solution of 100,000 realizations was used as the reference. It can be observed that even for $M = 100$ random walks per realization the obtained histogram is very close to the reference. Fifty equal width bins were used for all histograms in Fig. 7. We define mean square error for a histogram as

$$MSE_{hist} = \frac{1}{n_{bins}} \sum_{i=1}^{n_{bins}} (p(x_i) - p^*(x_i))^2. \quad (25)$$

The convergence of the mean square error of the histograms obtained with Algorithm 2 to the reference histogram is compared with the convergence of the histogram calculated by analytic integration for different number of realizations (N) in the right portion of Fig. 7.

6.2. Estimating the mean solution

Since the proposed method can be used for estimating the one-point distribution of the solution, it can also be used for estimating the moments of the solution at any given point. Estimating the mean is specifically efficient using Algorithm 2. As it was shown in [9], by tracking one trajectory per realization ($M = 1$), the mean solution can be calculated very efficiently. This is the case, since the stochastic MC method provides an unbiased estimate of the solution in each realization, and calculating the mean involves averaging the solution of different realizations. Fig. 8 illustrates the convergence of the mean solution calculated with Algorithm 2 at all points $x \in S = \{0.25, 0.5, 0.75\}$ to the reference solution. Here the mean square error is define as

$$MSE = \frac{1}{|S|} \sum_{x \in S} (\bar{p}(x) - \bar{p}^*(x))^2, \quad (26)$$

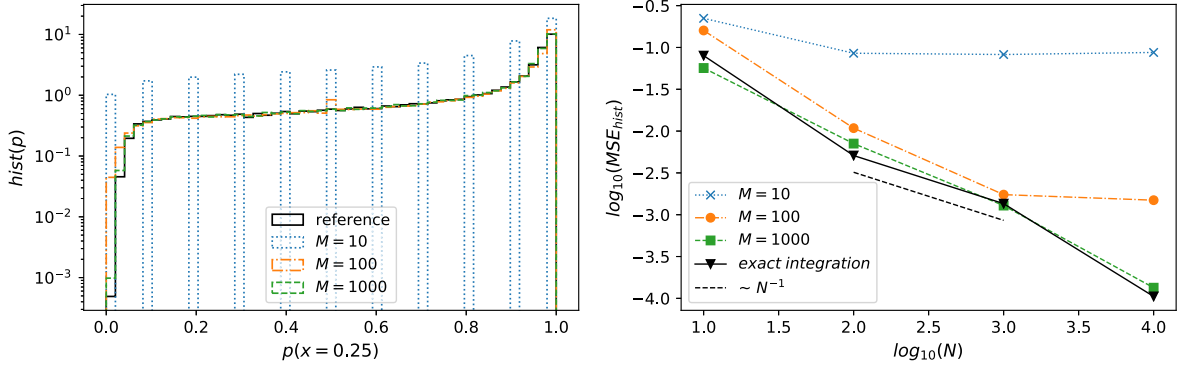


Fig. 7. Left: comparison of the one point histogram of pressure generated by Algorithm 2 and the reference histogram at $x = 0.25$. Right: Convergence of the histograms obtained from Algorithm 2 to the reference histogram.

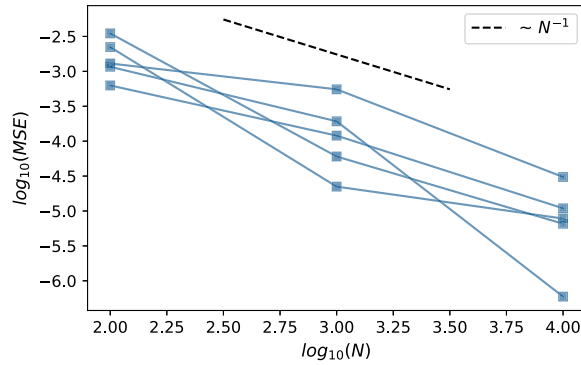


Fig. 8. Convergence of the mean solution calculated with Algorithm 2 with $M = 1$ to the reference solution for the example described in section 6.2. Different lines correspond to different experiments.

where $\bar{p}(x)$ and $\bar{p}^*(x)$ are the mean solutions calculated using Algorithm 2 and analytic integration, respectively. These results illustrate that the proposed method can be used to efficiently find the mean solution in highly heterogeneous conductivity fields.

6.3. Variance reduction for mean calculation

In applications such as flow in porous media, it is common to have a trend in the log-conductivity field. Based on the work in [17], the algorithm proposed for mean calculation can be modified to use the trend in the conductivity field to reduce the variance of the estimated mean. In short, for every realization of K , one could track a particle in that realization along with a shadow particle in the trend conductivity field using the same random numbers and store the boundary conditions at the hitting points of the boundary for both particles. The contribution of that realization to the mean would then be the sum of the difference between the boundary conditions and the solution of the trend conductivity field, which can be calculated once. This idea is outlined in Algorithm 3.

Here we present an example of variance reduction for calculating the mean in such a setting. The mean trend in $\log(K)$ is defined as

$$\log(K)_{trend} = 2 - x. \quad (27)$$

The log conductivity realizations are generated by adding a Gaussian noise process with mean zero, standard deviation $\sigma = 0.25$ and an exponential correlation structure with $l_y = 0.25$ to this trend. A number of conductivity realizations generated using this procedure and their corresponding pressure solutions are illustrated in Fig. 9. Figure 10 shows the variance reduction for calculating the mean using shadow particles in the trend conductivity field. The proposed variance reduction technique works best for cases where the solution to the trend conductivity field is highly correlated with the solution for different realizations of K . In our example, this is the case for relatively low variance of the added noise to the log conductivity.

```

// Initialization
s=0
// Looping over different realizations of K
for i ← 1 to N do
  K = generateField()
  // Generate a large array of uniform random variables to use for the current realization and the
  // shadow process
  U = uniformRandArray()
  pi = solveMC(x0, K, M = 1, U)
  // Track a shadow particle in the trend conductivity field
  pshadow = solveMC(x0, Ktrend, M = 1, U)
  s += pi - pshadow
end
// return MC estimate of the mean solution. Here we assume the solution for the trend field, ptrend, is
// pre-computed.
return ptrend(x0) + s/N

```

Algorithm 3: Using a shadow process for variance reduction in mean calculation.

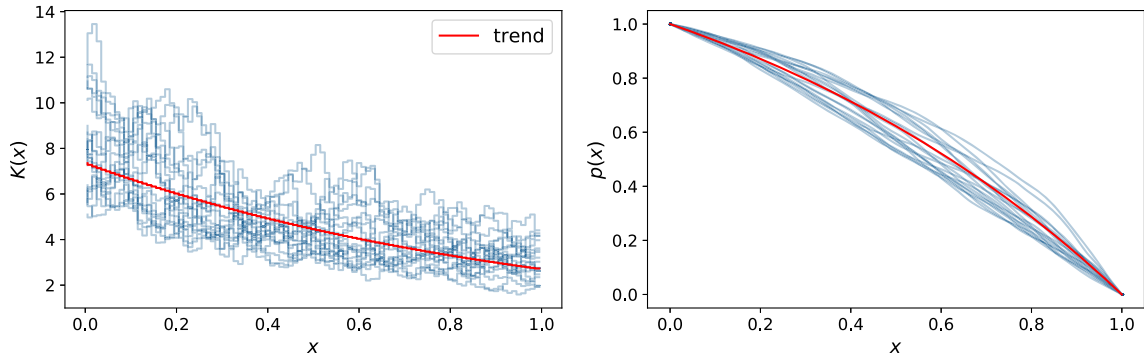


Fig. 9. Multiple realizations of the conductivity field along with their corresponding pressure solution for the example in section 6.3.

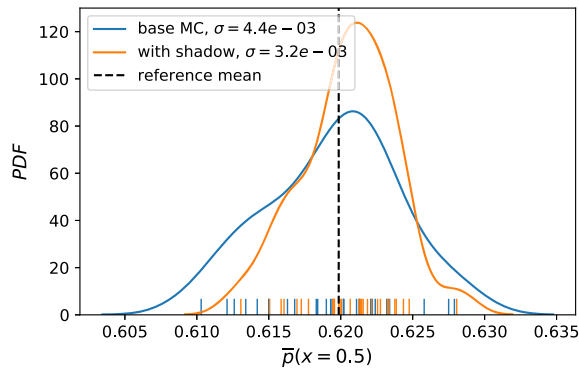


Fig. 10. The distribution of the mean solution at $x=0.5$ calculated by the base MC (Algorithm 2 with $M=1$) and using shadow particles (Algorithm 3) for thirty different experiments. By using shadow particles, the variance of the calculated mean is reduced by 27%.

7. Conclusions and future work

In this work, we reviewed the stochastic counterpart of the differential operator in elliptic and parabolic conservation equations with heterogeneous conductivity fields. Numerical challenges due to time step restrictions for using this formulation were discussed. A Monte Carlo algorithm is proposed to solve the elliptic problem for one-dimensional domains with piecewise constant conductivity. An example was provided to illustrate that this method is capable of accurately obtaining the solution of a deterministic PDE. Moreover, the proposed method was used to calculate accurate estimates of the one-point distribution of the solution. It was shown that the proposed stochastic method can provide a very efficient alternative for estimating the mean solution of a random PDE at specific points of interest in the domain. Finally, a variance reduction scheme was proposed for applying the method for efficient mean calculation.

The proposed stochastic simulations can be accelerated using numerical methods designed for the simulation of stochastic processes. Variance reduction strategies such as control variate schemes can be used to decrease the required number of MC trials for a given precision and will be the subject of future investigations. Moreover, the known statistics of the conduc-

tivity field can be used to accelerate path generation for random walk simulations. In two and three-dimensional examples that will be the subject of future work, these accelerations can play a key role. In these higher dimensional domains, extra attention should be given to the simulation of the random paths close to the boundary (e.g. calculating accurate estimates of the exit locations). Furthermore, since the random walks only experience the random field locally, generation of complete realizations of the random field can be avoided. In subsurface flow simulations, this could lead to significant computational cost savings in generating geostatistical models for uncertainty quantification. Finally, an effective implementation strategy to increase the efficiency of the algorithm is to partition the particle paths and have dedicated cores that simulate paths in each partition. This implementation strategy will be explored in future investigations.

Conflict of interest statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Amir H. Delgoushaie is grateful to Daniel M. Tartakovsky and Joseph Bakarji from the Energy Resources Engineering department at Stanford University for several helpful discussions. Funding for this project was provided by the Stanford University Reservoir Simulation Industrial Affiliates (SUPRI-B) program.

References

- [1] H. Versteeg, W. Malalasekera, *The Finite Volume Method*, 1995.
- [2] L.N. Trefethen, D. Bau III, *Numerical Linear Algebra*, vol. 50, SIAM, 1997.
- [3] B. Øksendal, in: *Stochastic Differential Equations*, Springer, 2003, pp. 65–84.
- [4] É. Pardoux, in: *Stochastic Analysis and Related Topics VI*, Springer, 1998, pp. 79–127.
- [5] K.K. Sabelfeld, N.A. Simonov, *Random Walks on Boundary for Solving PDEs*, Walter de Gruyter, 1994.
- [6] F. Buchmann, W. Petersen, *BIT Numer. Math.* 43 (2003) 519–540.
- [7] Y. Hu, *Stoch. Process. Appl.* 48 (1993) 107–121.
- [8] Y. Zhou, W. Cai, *J. Sci. Comput.* 69 (2016) 107–121.
- [9] F. Anker, C. Bayer, M. Eigel, M. Ladhak, J. Neumann, J. Schoenmakers, *SIAM J. Sci. Comput.* 39 (2017) A1168–A1200.
- [10] J.W. Woods, H. Stark, *Probability, Random Processes, and Estimation Theory for Engineers*, Prentice-hall, 2000.
- [11] J.-M. Harrison, L.-A. Shepp, *Ann. Probab.* (1981) 309–313.
- [12] A. Lejay, M. Martinez, et al., *Ann. Appl. Probab.* 16 (2006) 107–139.
- [13] C. Gardiner, *Stochastic Methods*, vol. 4, Springer, Berlin, 2009.
- [14] A.J. Chorin, O.H. Hald, *Stochastic Tools in Mathematics and Science*, vol. 3, Springer, 2009.
- [15] I. Bilionis, *Introduction to uncertainty quantification*, <https://github.com/PredictiveScienceLab/uq-course>. (Accessed 20 March 2018).
- [16] E. Pardo-Iguzquiza, *Math. Geol.* 25 (1993) 177.
- [17] M.H. Gorji, N. Andric, P. Jenny, *J. Comput. Phys.* 295 (2015) 644–664.