A stochastic method for flow in random porous media: Karhunen-Loeve expansion, polynomial chaos, and probabilistic collocation

Darin Comeau

November 26, 2010

Abstract

This paper presents a method put forth by Li & Zhang in their 2007 paper *Probabilistic collocation method for flow in porous media: Comparisons with other stochastic methods* [LZ07]. A stochastic method for flow in random porous media is detailed using Karhunen-Loeve expansion to approximate the log conductivity field Y, a polynomial chaos expansion to approximate the hydraulic head field h, and a probabilistic collocation method to determine the coefficients of the polynomial chaos expansion of h. The method is then implemented in a 1D domain, and compared against a benchmark of Monte Carlo simulation with 10,000 realizations.

1 Introduction

1.1 Background

The ability to manage and protect groundwater supplies in a responsible and scientifically defensible manner rests on one's ability to model complex subsurface processes. Due to the heterogeneity of a typical subsurface environment, together with a lack of sufficient site characterizations, accurate and verifiable predictions of subsurface flow are notoriously difficult in a deterministic setting. As population growth in Arizona and global climate change place an ever increasing strain on potable water, groundwater plays a key role in meeting demand for fresh water, especially in the semi-arid climate of Arizona. Through developing the tools to better understand and model subsurface flow and transport on large domains, we will enhance the ability to properly manage and protect Arizona's increasingly strained, and valuable, groundwater supply.

The subsurface environment is a great source of uncertainty, naturally lending itself to stochastic methods. Straight-forward approaches to modeling flow in a random porous media, such as Monte Carlo simulation, are not feasible for large-scale problems. Considerable work has been done, theoretical and computational, in developing alternative methods that are computationally efficient, and can account for heterogeneity in the subsurface environment. In this paper we focus on a method recently put forward by Li & Zhang in their 2007 paper *Probabilistic collocation method for flow in porous media: Comparisons with other stochastic methods* [LZ07].

1.2 Problem Statement

The governing equations for steady-state subsurface flow in saturated conditions in a domain $D \subset \mathbb{R}^3$ are the continuity equation

$$\nabla \cdot q(\mathbf{x}) = g(\mathbf{x}) \qquad \mathbf{x} \in D \tag{1}$$

and Darcy's law

$$q(\mathbf{x}) = -K(\mathbf{x})\nabla h(\mathbf{x}) \qquad \mathbf{x} \in D \tag{2}$$

with prescribed boundary conditions.

$$h(\mathbf{x}) = H(\mathbf{x}) \qquad \mathbf{x} \in \Gamma_D \tag{3}$$

$$q(\mathbf{x}) \cdot n(\mathbf{x}) = Q(\mathbf{x}) \qquad \mathbf{x} \in \Gamma_N \tag{4}$$

where Γ_D, Γ_N are the portions of the boundary with prescribed Dirichlet and Neumann boundary conditions, respectively. Here q is the flux, h is the hydraulic head, K is the hydraulic conductivity, and g is the forcing term of water sources and/or sinks. Combining 1 and 2 we have the single flow equation

$$\nabla \cdot [K(\mathbf{x})\nabla h(\mathbf{x})] = g(\mathbf{x}) \tag{5}$$

This equation can be solved provided we have a good characterization of K, which in realistic settings is not the case. Sparse measurements and uncertainty in those measurements present one set of challenges, and the high degree of variability of the hydraulic conductivity field K make solving the problem in a deterministic setting unrealistic. Instead we view the problem in a stochastic setting, and allow K to be a random field, and rendering the dependent flux and hydraulic head fields also random, resulting in a set of stochastic partial differential equations. Since the solution h to (5) is now a random process, our aim is to solve for the statistical moments of h, notably its mean and variance. K is commonly taken to be log-normally distributed, so we will instead deal with the normally distributed log conductivity $Y = \log K$, with mean $\langle Y \rangle$, constant variance σ_Y^2, and the corresponding flow equation:

$$\nabla \cdot [\exp(Y(\mathbf{x}))\nabla h(\mathbf{x})] = g(\mathbf{x}) \tag{6}$$

Following [LZ07], our approach will be:

1. Approximate the log conductivity field Y with a truncated Karhunen-Loeve expansion.

2. Approximate the random hydraulic head field h with a truncated polynomial chaos expansion.

3. Evaluate the coefficients $c_n(\mathbf{x})$ of the polynomial chaos expansion of h using the probabilistic collocation method.

4. Calculate statistical information about h using the coefficients $c_n(\mathbf{x})$.

2 Karhunen-Loeve Expansion

Let us write the log conductivity as $Y(\mathbf{x}, \theta)$ $\mathbf{x} \in D, \theta$ in some probability space Θ to emphasize the random nature of Y. We can decompose Y into its deterministic and random components:

$$Y(\mathbf{x}, \theta) = \overline{Y}(\mathbf{x}) + Y'(\mathbf{x}, \theta) \tag{7}$$

where $\overline{Y}(\mathbf{x}) = \langle Y(\mathbf{x}, \theta) \rangle$ is the mean, and $Y'(\mathbf{x}, \theta)$ is the random, zero-mean fluctuation. The covariance of Y is then given by $C_Y(\mathbf{x}, \mathbf{y}) = \langle Y'(\mathbf{x}, \theta) Y'(\mathbf{y}, \theta) \rangle$, and since this is necessarily a bounded, symmetric, positive definite operator, C_Y admits a spectral decomposition [GS91]:

$$C_Y(\mathbf{x}, \mathbf{y}) = \sum_{n=1}^{\infty} \lambda_n f_n(\mathbf{x}) f_n(\mathbf{y}) \tag{8}$$

where the (deterministic) eigenvalues and eigenfunctions $\lambda_n, f_n(\mathbf{x})$ are found as solutions to the Fredholm integral equation

$$\int_D C_Y(\mathbf{x}, \mathbf{y}) f_n(\mathbf{x}) d\mathbf{x} = \lambda_n f_n(\mathbf{y}) \tag{9}$$

Analytical expressions for the eigenvalues and eigenfunctions for the integral equation (9) can be found in certain cases, but in general the problem needs to be solved numerically, for example when the covariance kernel is not separable, or when the domain has complex geometry [LZ07]. The spectral expansion has the following properties [GS91]:

- The set of eigenfunctions $\{f_n(\mathbf{x})\}_{n=1}^{\infty}$ forms a complete orthonormal set for $L^2(D)$.

- For each eigenvalue λ_n, there corresponds at most a finite number of linearly independent eigenfunctions.

- There are at most a countably infinite set of eigenvalues, all of which are positive real numbers.

- The expansion converges uniformly.

We can then expand the random process $Y'(\mathbf{x},\theta)$ in terms of this orthogonal basis, with the coefficients $\xi_n(\theta)$ that are mean zero random variables:

$$Y'(\mathbf{x},\theta) = \sum_{n=1}^{\infty} \xi_n(\theta)\sqrt{\lambda_n} f_n(\mathbf{x}) \tag{10}$$

which is called the Karhunen-Loeve expansion [Kar47], [Loe46]. To determine properties of these random variables ξ_n, we note that

$$\begin{aligned} C(\mathbf{x},\mathbf{y}) &= \left\langle Y'(\mathbf{x},\theta)Y'(\mathbf{y},\theta)\right\rangle & (11)\\ &= \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \langle \xi_n(\theta)\xi_m(\theta)\rangle \sqrt{\lambda_n\lambda_m} f_n(\mathbf{x})f_m(\mathbf{y}) & (12)\end{aligned}$$

Multiplying by f_k, integrating over the domain and making use of orthogonality, we have

$$\begin{aligned} \int_D C(\mathbf{x},\mathbf{y})f_k(\mathbf{x}) &= \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \int_D \langle \xi_n(\theta)\xi_m(\theta)\rangle \sqrt{\lambda_n\lambda_m} f_n(\mathbf{x})f_m(\mathbf{y})f_k(\mathbf{x})d\mathbf{x} & (13)\\ \lambda_k f_k(\mathbf{y}) &= \sum_{m=1}^{\infty} \langle \xi_m(\theta)\xi_k(\theta)\rangle \sqrt{\lambda_m\lambda_k} f_m(\mathbf{y}) & (14)\end{aligned}$$

Multiplying again by $f_l(\mathbf{y})$ and integrating over the domain, we have

$$\begin{aligned} \int_D \lambda_k f_k(\mathbf{y})f_l(\mathbf{y})d\mathbf{y} &= \sum_{m=1}^{\infty} \int_D \langle \xi_m(\theta)\xi_k(\theta)\rangle \sqrt{\lambda_m\lambda_k} f_m(\mathbf{y})f_l(\mathbf{y})d\mathbf{y} & (15)\\ \lambda_k\delta_{kl} &= \langle \xi_l(\theta)\xi_k(\theta)\rangle \sqrt{\lambda_l\lambda_k} & (16)\end{aligned}$$

Thus we see that the random variable coefficients ξ_n satisfy $\langle \xi_n(\theta)\xi_m(\theta)\rangle = \delta_{nm}$. Moreover, since we are assuming that our original process Y is Gaussian, the ξ_n are also Gaussian, and thus are orthogonal standard normal variables [GS91]. We can see how the variance σ_Y^2 is related to the expansion from (12):

$$\begin{aligned} C(\mathbf{x},\mathbf{x}) &= \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \delta_{nm}\sqrt{\lambda_n\lambda_m} f_n(\mathbf{x})f_m(\mathbf{x}) & (17)\\ \sigma_Y^2 &= \sum_{n=1}^{\infty} \lambda_n f_n^2(\mathbf{x}) & (18)\\ \int_D \sigma_Y^2 d\mathbf{x} &= \sum_{n=1}^{\infty} \int_D \lambda_n f_n^2(\mathbf{x})d\mathbf{x} & (19)\\ \sigma_Y^2 \mathrm{Vol}(D) &= \sum_{n=1}^{\infty} \lambda_n & (20)\end{aligned}$$

Thus if the (necessarily positive) λ_n are in decreasing order, we can consider only the first few terms of the Karhunen-Loeve expansion to approximate Y, and the number of terms needed will be determined by the rate of decay of λ_n. The benefit of using the Karhunen-Loeve expansion for Y rather than some other orthogonal series expansion is that the mean square error from the result of truncation of the decomposition is minimal. To see this, we follow [GS91], and let Y be expanded by any orthogonal series:

$$Y'(\mathbf{x},\theta) = \sum_{n=1}^{\infty} \zeta_n(\theta)\sqrt{\alpha_n} g_n(\mathbf{x}) \tag{21}$$

The truncation error ϵ_N in keeping only the first N terms is given by

$$\epsilon_N(\mathbf{x},\theta) = \sum_{n=N+1}^{\infty} \zeta_n(\theta)\sqrt{\alpha_n} g_n(\mathbf{x}) \tag{22}$$

$$\tag{23}$$

From (21), if we multiply both sides by $g_m(\mathbf{x})$ and integrate over D, we get an expression for $\zeta_n(\theta)$:

$$\xi_n(\theta) = \frac{1}{\sqrt{\alpha_n}} \int_D Y'(\mathbf{x},\theta) g_m(\mathbf{x}) d\mathbf{x} \tag{24}$$

Plugging this into (23), we get

$$\epsilon_N(\mathbf{x},\theta) = \sum_{n=N+1}^{\infty} g_n(\mathbf{x}) \int_D Y'(\mathbf{y},\theta) g_n(\mathbf{y}) d\mathbf{y} \tag{25}$$

$$\epsilon_N^2(\mathbf{x},\theta) = \sum_{n=N+1}^{\infty} \sum_{m=N+1}^{\infty} \beta_{nm} g_n(\mathbf{x}) g_m(\mathbf{x}) \int_D \int_D Y'(\mathbf{y},\theta) Y'(\mathbf{z},\theta) g_n(\mathbf{y}) g_m(\mathbf{z}) d\mathbf{y} d\mathbf{z} \tag{26}$$

where the β_{nm} are some coefficients from squaring the infinite series, but the only ones we will need are $\beta_{nn} = 1$. Taking expectation, and then integrating over D, we have

$$\epsilon_N^2(\mathbf{x}) = \sum_{n=N}^{\infty} \sum_{m=N}^{\infty} \beta_{nm} g_n(\mathbf{x}) g_m(\mathbf{x}) \int_D \int_D \left\langle Y'(\mathbf{y},\theta) Y'(\mathbf{z},\theta) \right\rangle g_n(\mathbf{y}) g_m(\mathbf{z}) d\mathbf{y} d\mathbf{z} \tag{27}$$

$$\epsilon_N^2 = \sum_{n=N}^{\infty} \sum_{m=N}^{\infty} \int_D \beta_{nm} g_n(\mathbf{x}) g_m(\mathbf{x}) d\mathbf{x} \int_D \int_D C_Y(\mathbf{y},\mathbf{z}) g_n(\mathbf{y}) g_m(\mathbf{z}) d\mathbf{y} d\mathbf{z} \tag{28}$$

$$= \sum_{n=N}^{\infty} \int_D \int_D C_Y(\mathbf{y},\mathbf{z}) g_n(\mathbf{y}) g_n(\mathbf{z}) d\mathbf{y} d\mathbf{z} \tag{29}$$

We want to minimize this quantity, subject to the orthogonality of $g_n(\mathbf{x})$, i.e. minimize the functional:

$$F\left[g_n(\mathbf{x})\right] = \sum_{n=N}^{\infty} \int_D \int_D C_Y(\mathbf{y},\mathbf{z}) g_n(\mathbf{y}) g_n(\mathbf{z}) d\mathbf{y} d\mathbf{z} - \lambda_n \left(\int_D g_n(\mathbf{z}) g_n(\mathbf{z}) d\mathbf{z} - 1 \right) \quad (30)$$

For each k, we take the derivative of (30) with respect to $g_k(\mathbf{x})$, and set equal to zero to get

$$\frac{\partial F\left[g_m(\mathbf{x})\right]}{\partial g_k(\mathbf{x})} = 2 \int_D \int_D C_Y(\mathbf{y},\mathbf{z}) g_k(\mathbf{y}) d\mathbf{y} d\mathbf{z} - 2\lambda_k \int_D g_k(\mathbf{z}) d\mathbf{z} \quad (31)$$

$$0 = \int_D \left[\int_D C_Y(\mathbf{y},\mathbf{z}) g_k(\mathbf{y}) d\mathbf{y} - \lambda_k g_k(\mathbf{z}) \right] d\mathbf{z} \quad (32)$$

This true when the eigenfunctions satisfy

$$\int_D C_Y(\mathbf{y},\mathbf{z}) g_k(\mathbf{y}) d\mathbf{y} = \lambda_k g_k(\mathbf{z}) d\mathbf{z} \quad (33)$$

which is exactly the Karhunen-Loeve expansion. Thus truncating all but the first N terms, we approximate our log conductivity field as

$$Y(\mathbf{x},\theta) = \overline{Y}(\mathbf{x}) + \sum_{n=1}^{N} \xi_n(\theta) \sqrt{\lambda_n} f_n(\mathbf{x}) \quad (34)$$

and our governing equation becomes

$$\nabla \cdot \left\{ \exp\left[\overline{Y}(\mathbf{x}) + \sum_{n=1}^{N} \xi_n(\theta) \sqrt{\lambda_n} f_n(\mathbf{x}) \right] \nabla h(\mathbf{x}) \right\} = g(\mathbf{x}) \quad (35)$$

3 Polynomial Chaos

Since we do not a priori know the covariance for the h, we cannot similarly decompose the hydraulic head field by a Karhunen-Loeve expansion, so an alternative expansion for h is necessary. Such an expansion would desirably also be orthogonal, and dependent on the random variables ξ_n that appear in the expansion of Y. For this we turn to the polynomial chaos expansion where h will be written as a linear combination of a basis of random functions with deterministic coefficients. This was first introduced by Wiener in [Wie38], and later gaining popularity in various disciplines [GS90].

For the formal development of polynomial chaos, we proceed as in [GS91]. To begin, let $\{\xi_n(\theta)\}_{n=1}^{\infty}$ be a set of orthogonal Gaussian random variables on a probability space Θ, i.e. the covariance between two distinct ξ_ns is zero. We recursively define subsets of the space of square-integrable random variables $L^2(\Theta)$. Let $\hat{\Gamma}_d$ be the set of all polynomials on $\{\xi_n(\theta)\}_{n=1}^{\infty}$ of degree $\leq d$, and let $\Gamma_d \subset \hat{\Gamma}_d$ be the set of all polynomials orthogonal to $\hat{\Gamma}_{d-1}$. We call Γ_d the *polynomial chaos of order* d. Elements of the dth polynomial chaos are polynomials of finite degree but infinite dimensional, since there are an infinite number of indeterminates $\{\xi_n(\theta)\}_{n=1}^{\infty}$ to select d from. The space spanned by the union of sets Γ_d will be a subspace of $L^2(\Theta)$, and under general conditions this subspace will be dense [GS91]. Then we can represent a general square-integrable random variable as

$$
\begin{aligned}
\phi(\theta) \quad = \quad & a_0\Gamma_0 && (36)\\
+ \quad & \sum_{i_1=1}^{\infty} a_{i_1}\Gamma_1\left(\xi_{i_1}(\theta)\right) && (37)\\
+ \quad & \sum_{i_1=1}^{\infty}\sum_{i_2=1}^{i_1} a_{i_1 i_2}\Gamma_2\left(\xi_{i_1}(\theta),\xi_{i_2}(\theta)\right) && (38)\\
+ \quad & \sum_{i_1=1}^{\infty}\sum_{i_2=1}^{i_1}\sum_{i_3=1}^{i_2} a_{i_1 i_2 i_3}\Gamma_3\left(\xi_{i_1}(\theta),\xi_{i_2}(\theta),\xi_{i_3}(\theta)\right) && (39)\\
+ \quad & \cdots && (40)
\end{aligned}
$$

To clarify, $\Gamma_d(\xi_{i_1},\ldots,\xi_{i_d})$ is an dth degree polynomial in the variables $\xi_{i_1},\ldots,\xi_{i_d}$, chosen from (with possible repetition) the set $\{\xi_n\}_{n=1}^{\infty}$, and there are infinitely many such polynomials for each d, together forming an element from the dth order polynomial chaos. We can reindex and write this concisely as

$$\phi(\theta) \quad = \quad \sum_{j=0}^{\infty} c_j\Psi_j(\xi(\theta)) \qquad \xi = \{\xi_n(\theta)_{n=1}^{\infty}\} \tag{41}$$

To truncate the polynomial chaos expansion, there are two controlling factors - the highest order of the polynomial chaos d, and the number of random variables $\{\xi_n(\theta)\}_{n=1}^{N}$ used, which we call the dimension of polynomial chaos N. The number of terms p in a truncated expansion is then

$$p \quad = \quad \binom{N+d}{d} = \frac{(N+d)!}{N!d!} \tag{42}$$

Up to now there has been no assumption made on the family of random variables $\{\xi_n\}_{n=1}^{\infty}$ other have finite variance and are orthogonal with respect to some probability measure. If we now consider $\{\xi_n\}_{n=1}^{\infty}$ to be orthogonal standard normal variables, as those that appear in the Karhunen-Loeve expansion are, then the polynomial chaoses are orthogonal with respect to Gaussian measure. These are exactly the multidimensional Hermite polynomials, and we can use its generating function to evaluate the terms in the dth order polynomial chaos [GS91]:

$$\Gamma_d(\xi_{i_1},\ldots,\xi_{i_d}) \quad = \quad (-1)^d \exp\left(\frac{1}{2}\xi^T\xi\right)\frac{\partial^d}{\partial\xi_{i_1}\cdots\partial\xi_{i_d}}\exp\left(-\frac{1}{2}\xi^T\xi\right) \tag{43}$$

where $\xi = (\xi_{i_1},\ldots,\xi_{i_d})^T$ and we have suppressed the dependence on θ. So if we truncate the Karhunen-Loeve expansion at N terms and consider polynomial chaos up to order d, we can now approximate the hydraulic head field $h(\mathbf{x})$ with a truncated polynomial chaos expansion with $p = (N+d)!/(N!d!)$ as

$$\hat{h}(\mathbf{x}) \quad = \quad \sum_{j=1}^{p} c_j(\mathbf{x})\Psi_j(\xi) \tag{44}$$

Our task now is to evaluate the deterministic coefficients $c_j(\mathbf{x})$.

4 Probabilistic Collocation Method

To evaluate the coefficients $c_j(\mathbf{x})$, we use the weighted residual method. In general, for a differential operator $\mathcal{L}$ and stochastic differential equation of the form $\mathcal{L}y(\mathbf{x},\theta) = f(\mathbf{x})$, with approximate solution $\hat{y}(\mathbf{x},\theta) = \sum_{n=1}^{p} c_j(\mathbf{x})\Psi_j(\xi(\theta))$, we define the residual for the differential equation as

$$R\left(\xi(\theta)\right) \quad = \quad \mathcal{L}\hat{y} - f \tag{45}$$

We want to minimize this with respect to the random vector $\xi(\theta)$ in some sense, so we form a weak formulation and integrate against some chosen weight functions $\{w_j(\xi(\theta))\}$:

$$\langle R\left(\xi(\theta)\right) w_j(\xi(\theta))\rangle \quad = \quad 0 \tag{46}$$

In the probabilistic collocation method, the weighting functions are chosen to be delta functions at prescribed collocation points ξ^j (which are really N dimensional vectors) [WTM96][TPPM97]:

$$w_j(\xi(\theta)) \quad = \quad \delta(\xi(\theta) - \xi^j) \tag{47}$$

Then (46) simply becomes

$$R\left(\xi^j\right) \quad = \quad 0 \tag{48}$$

for the (as yet to be determined) collocation points ξ^j. We use the superscript, as well as omit the dependence on θ, to distinguish that these are deterministic quantities. Since there are p coefficients to solve for, we need to choose p collocation points. We are insisting that our solution be exact when the random vector $\xi(\theta) = (\xi_1(\theta),\cdots,\xi_d(\theta))$ takes on a specific (deterministic) value $\xi^j = (\xi_1^1, \ldots, \xi_d^j)$, so we'll naturally want to choose values of ξ^j that have the highest probability.

For the flow equation (34), for each $j = 1, \ldots, p$, the residual equation has solution $\hat{h}^j(\mathbf{x})$, which satisfies

$$\nabla \cdot \left\{ \exp\left[\overline{Y}(\mathbf{x}) + \sum_{n=1}^{N} \xi_n^j \sqrt{\lambda_n} f_n(\mathbf{x}) \right] \nabla \hat{h}^j(\mathbf{x}) \right\} - g(\mathbf{x}) \quad = \quad 0 \tag{49}$$

where ξ^j is the jth collocation point. This is now a set of p deterministic equations, independent of each other, for which we can numerically solve for each $\hat{h}^j$. To compute the statistics of h, however, we simply cannot compute sample moments of the $\hat{h}^j$s, as is done in a Monte Carlo simulation. The $\hat{h}^j$s, which we refer to as representations rather than realizations, are not all equally probably, since the collocation points ξ^j are not equally probable. Instead, we need to return to the coefficients $C(\mathbf{x}) = [c_1(\mathbf{x}), \ldots, c_p(\mathbf{x})]^T$ of the polynomial chaos expansion in (44), which will satisfy the equation

$$ZC(\mathbf{x}) = \tilde{h}(\mathbf{x}) \qquad Z_{ji} = \Psi_i(\xi^j), \quad \tilde{h}(\mathbf{x}) = [\hat{h}_1(\mathbf{x}), \ldots, \hat{h}_p(\mathbf{x})]^T \tag{50}$$

Z is a $P \times P$ matrix, whose jth row is the elements of the polynomial chaos Ψ_i evaluated at the collocation points ξ^j. Thus to solve for $C(\mathbf{x})$, we need to ensure that Z has full rank, and this affects our choice of collocation points. To select the collocation points, since $\hat{h}$ is a linear combination of multidimensional Hermite polynomials, we should choose our points from the roots of the next degree Hermite polynomial, as is done in Gaussian quadrature [LZ07]. For polynomial chaos up to degree d, we would then choose the roots of the $d+1$st Hermite polynomial, and use these $d+1$ values as entries in $\xi^j = (\xi_1^j, \ldots, \xi_N^j)$, permitting repetition. We only need p collocation points, and we have $(d+1)^N$ possible choices for ξ^j, so we want to choose those vectors with the highest probability, recognizing that $\xi(\theta)$ is an N dimensional multivariate standard normal. After ranking the possible choices according to probability, for each point ξ^j chosen, we evaluate the terms in the polynomial $\Psi_i(\xi^j)$, add those values as a row in the Z matrix, and only accept the collocation point if Z still has full rank.

Once the collocation points ξ^j have been chosen and the $\hat{h}^j(\mathbf{x})$ found, we calculate the coefficients $c_j(\mathbf{x})$ with (50). We can then calculate the (approximate) moments of h by taking moments of (44). The Ψ_is are all mean zero except for the constant term $\Psi_1 = 1$, and are mutually orthogonal, so for example the first and second moments of h are given by

$$\langle h(\mathbf{x}) \rangle = c_1(\mathbf{x}) \qquad \sigma_h^2 = \sum_{j=2}^{p} c_j^2(\mathbf{x}) \sigma_{\Psi_j}^2 \tag{51}$$

5 Example

Following Li & Zhang in [LZ07], we implemented this technique on a 1D example of steady state flow under saturated conditions with covariance $C_Y(x,y) = \sigma_Y^2 \exp(-|x-y|/\eta)$, where σ_Y^2 and η are the variance and correlation length, respectively. The eigenvalues and eigenfunctions for this covariance kernel can be computed analytically [ZL04]:

$$\lambda_n = \frac{2\eta\sigma_Y^2}{\eta^2\omega_n^2+1} \tag{52}$$

$$f_n(x) = \frac{1}{\sqrt{(\eta^2\omega_n^2+1)\frac{L}{2}+\eta}}\left[\eta\omega_n\cos(\omega_n x)+\sin(\omega_n x)\right] \tag{53}$$

where ω_n are the positive roots, in ascending order, of

$$(\eta^2\omega^2-1)\sin(\omega L) = 2\eta\omega\cos(\omega L) \tag{54}$$

We kept only the first $N=3$ terms of the Karhunen-Loeve expansion, representing about 80% of the energy in the process [LZ07], and polynomial chaos up to order $d=2$, for a total of $p=10$ collocation points and corresponding hydraulic head representations. Using (43), the polynomial chaos expansion of h is then of the form

$$\hat{h}(\mathbf{x}) = a_0(\mathbf{x}) + \sum_{i=1}^{3} a_i(\mathbf{x})\xi_i + \sum_{1=i\le j=3} a_{i,j}(\mathbf{x})\left(\xi_i\xi_j-\delta_{ij}\right) \tag{55}$$

The collocation points ξ^j, each a vector with three entries, were then chosen to have entries from the roots of the third Hermite polynomial $H_3(\xi)=\xi^3-3\xi$, which are $\xi = 0 \pm \sqrt{3}$, with 0 having the highest probability. For example, the point $\xi^1=(0,0,0)$ was chosen as the first collocation point.

The flow equation was solved on a domain of length 10 [L] units, with Dirichlet boundary conditions $H_0 = 7$ [L], $H_L = 5$ [L], correlation length $\eta = 4$ and variance $\sigma_Y^2 = 1$. Results were compared against a benchmark of 10,000 Monte Carlo simulations. To account for the covariance structure in generating log conductivity realizations, we first decomposed the covariance matrix into a Cholesky decomposition $C_Y = LL^T$ [Zha02]. We then generated a random vector α of independent, standard normal random variables, and generated a realization of Y by

$$Y = L\alpha \tag{56}$$

Then the covariance of the random generated vector Y is

$$\left\langle YY^T\right\rangle = L\left\langle \alpha\alpha^T\right\rangle L^T = LIL^T = C \tag{57}$$

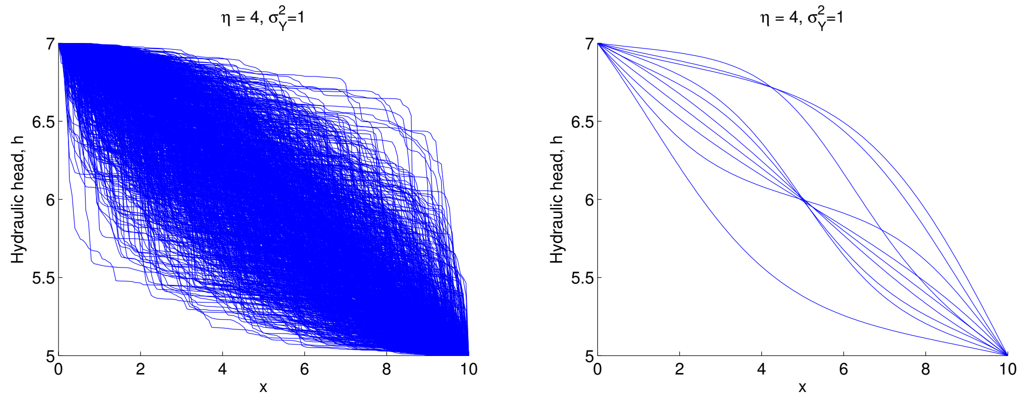

Figure 1: The first 1000 realizations of hydraulic head from 10,000 MC simulations (left). 10 representations of hydraulic head calculated using second-order PCM (right).

In Figure (1) we show the first 1000 realizations of hydraulic head from the Monte Carlo simulation, and the ten hydraulic head representations obtained using the outlined method. For reference, these are comparable to Figure 3 in [LZ07], where we used random dimension $N = 3$, and thus only $p = 10$ representations, vs. their $p = 28$. In Figure (2), we compare the variance of the hydraulic head field from the MC simulation versus that derived from the (51), and we see quite good agreement from such a few number of representations. We note Figure (2) is again comparable to Figure 2a in [LZ07].

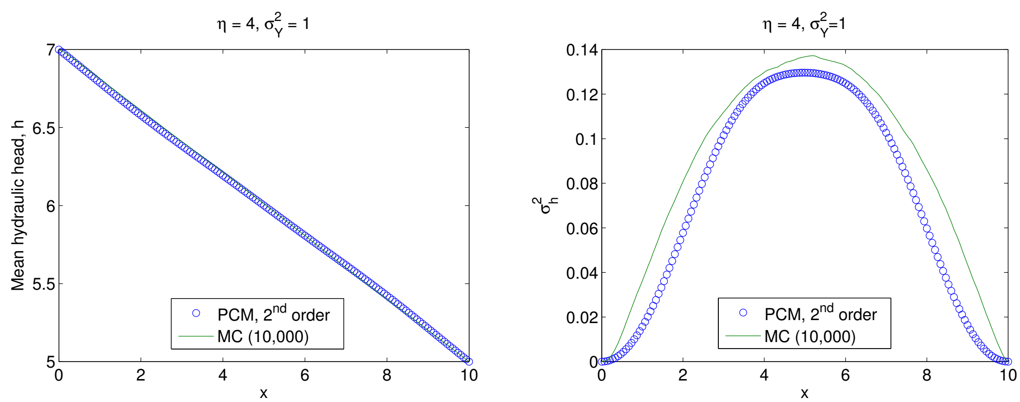

Figure 2: The mean (left) and variance (right) of the hydraulic head derived from the second order PCM as in (51), and MC.

6 Conclusion

We implemented a method for modeling flow in a random porous media put forth by Li & Zhang in [LZ07], reproducing their results in a lower random dimension. Using a Karhunen-Loeve expansion of the log conductivity parameter field, a polynomial chaos expansion of the dependent hydraulic head field, and the probabilistic collocation method to determine the coefficients of the polynomial chaos expansion, we were able to generate the statistics of the hydraulic head field by solving the flow equation only 10 times. Compared to a benchmark of a Monte Carlo simulation with 10,000 realizations, the first and second moments of the hydraulic head field were in close agreement, considering the 3 orders of magnitude difference in computation.

I would like to thank and acknowledge the NSF VIGRE fellowship for funding support over the summer and fall 2010 semesters, and my advisor Dr. Juan Restrepo, as well as my committee members Dr. Shlomo Neuman, Dr. Larry Winter, and Dr. Rabindra Bhattacharya for academic support.

References

[GS90] R. Ghanem and PD Spanos. Polynomial chaos in stochastic finite elements. *Journal of Applied Mechanics*, 57(1):197–202, 1990.

[GS91] RG Ghanem and PD Spanos. *Stochastic finite elements: a spectral approach*. Springer-Verlag, New York, 1991.

[Kar47] K. Karhunen. *Uber lineare Methoden in der Wahrscheinlichkeitsrechnung*. PhD thesis, Helsinki., 1947.

[Loe46] M. Loeve. Fonctions aléatoires du second ordre. *Rev. Sci*, 84(4):195–206, 1946.

[LZ07] H. Li and D. Zhang. Probabilistic collocation method for flow in porous media: Comparisons with other stochastic methods. *Water Resour. Res*, 43:44–48, 2007.

[TPPM97] M.A. Tatang, W. Pan, R.G. Prinn, and G.J. McRae. An efficient method for parametric uncertainty analysis of numerical geophysical models. *Journal of Geophysical Research*, 102(D18):21925, 1997.

[Wie38] N. Wiener. The homogeneous chaos. *American Journal of Mathematics*, 60(4):897–936, 1938.

[WTM96] M.D. Webster, M.A. Tatang, and G.J. McRae. Application of the probabilistic collocation method for an uncertainty analysis of a simple ocean model. *Report no. 4*, 1996.

[Zha02] Y.K. Zhang. *Stochastic methods for flow in porous media: coping with uncertainties*. Academic Press, 2002.

[ZL04] D. Zhang and Z. Lu. An efficient, high-order perturbation approach for flow in random porous media via Karhunen-Loeve and polynomial expansions. *Journal of Computational Physics*, 194(2):773–794, 2004.