Uncertainty quantification for flow in highly heterogeneous porous media

D. Xiu and D.M. Tartakovsky[a]*

[a]Theoretical Division, Los Alamos National Laboratory,
Mathematical Modeling and Analysis Group (T-7), MS B284, Los Alamos, NM 87545

Natural porous media are highly heterogeneous and characterized by parameters that are often uncertain due to the lack of sufficient data. This uncertainty (randomness) occurs on a multiplicity of scales. We focus on geologic formations with the two dominant scales of uncertainty: a large-scale uncertainty in the spatial arrangement of geologic facies and a small-scale uncertainty in the parameters within each facies. We propose an approach that combines random domain decompositions (RDD) and polynomial chaos expansions (PCE) to account for the large- and small-scales of uncertainty, respectively. We present a general framework and use a one-dimensional flow example to demonstrate that our combined approach provides robust, non-perturbative approximations for the statistics of the system states.

1. INTRODUCTION

Modeling of flow and transport in natural porous media is hampered by the insufficiency of available data. To make predictions under such conditions, one needs to assign the values of parameters to the points (cells) on a computational grid, where parameter data are absent. This is commonly done by treating such parameters as random fields, whose statistics are inferred from available data. This renders governing flow and transport equations stochastic even though the underlying physical phenomena are deterministic. While the parameter statistics are often highly non-Gaussian, and exhibit non-trivial correlation structures, most stochastic approaches assume the opposite.

Consider, for instance, the moment equations approach [1–4] that derives a set of deterministic equations for the statistical moments, usually the ensemble mean and (co)variance, of system states (hydraulic head, saturation, concentration, etc.). This approach requires closure approximations, such as perturbation expansions in the variances of system parameters (e.g., log hydraulic conductivity). This formally limits the applicability of these methods to mildly heterogeneous media, i.e., to media whose parameter variances are small. While this approach might work remarkably well even for some nonlinear problems [5], it often fails for others [6].

*This research was performed under the auspices of the U.S. Department of Energy, under contract W-7405-ENG-36. This work was supported in part by the U.S. Department of Energy under the DOE/BES Program in the Applied Mathematical Sciences, Contract KC-07-01-01, and in part by the LDRD Program at Los Alamos National Laboratory. This work made use of shared facilities supported by SAHRA (Sustainability of semi-Arid Hydrology and Riparian Areas) under the STC Program of the National Science Foundation under agreement EAR-9876800.

A non-perturbative alternative relies on polynomial chaos expansions. The classical Wiener polynomial chaos [7] defines the span of Hermite polynomial functionals of a Gaussian process and converges to any L^2 functional in the L^2 sense [8]. While the Wiener-Hermite polynomial chaos expansions have been applied successfully to analyze the propagation of uncertainty in some porous media [9], their theoretical limitations are well established [10,11]. Some of these limitations can be overcome by the use of generalized polynomial chaos expansions, which employ a wide range of orthogonal polynomials mostly from the Askey scheme [12]. The main advantage of the generalized polynomial chaos is its ability to represent accurately and efficiently many non-Gaussian stochastic processes [13–16]. However, as we demonstrate in this study, the generalized polynomial chaos might become computationally inefficient when applied to multi-modal processes that arise routinely in modeling flow and transport in composite porous media. The present study is devoted to overcoming this shortcoming.

A computational framework that we adopt here combines the generalized polynomial chaos with the random domain decomposition (RDD) approach [17–19]. The key advantage of RDD is that it provides robust closures (accurate approximations) for moment equations even when environments are highly heterogeneous and the statistical distributions and correlation structures of parameters are complex. RDD relies on the fact that a high degree of heterogeneity usually arises from the presence of different materials (populations) in the environment. Specifically, RDD replaces a non-Gaussian, multi-modal parameter field $Y(\boldsymbol{x})$ with a two-scale random process. The large scale randomness arises due to uncertainty in internal boundaries of materials (populations). The small scale randomness corresponds to uncertainty in parameters within each material. In other words, a non-Gaussian, multi-modal probability density function $p_Y(y)$ is replaced with a joint probability density function $p_Y(y,\gamma) = p_Y(y|\gamma)p_\Gamma(\gamma)$. The conditional probability density function $p_Y(y|\gamma)$ describes the distribution of Y within each material conditioned on the boundary location Γ, whose probability density function is $p_\Gamma(\gamma)$. Hence it has convenient properties, such as unimodality and convenient correlations.

Section 2 formulates the problem of diffusion in random composite media. We outline the generalized polynomial chaos expansion approach in Section 3 and demonstrate its limitations for multi-modal distributions of system parameters. In Section 4, we employ a random domain decomposition to extend the range of applicability of the polynomial chaos expansions to such parameters. Section 5 provides two computational examples, and analyzes the accuracy of the proposed approach.

2. PROBLEM FORMULATION

Consider steady-state saturated flow in a domain Ω,

$$\nabla \cdot K \nabla h(\boldsymbol{x}) + f = 0, \qquad \boldsymbol{x} \in \Omega. \tag{1}$$

The hydraulic conductivity $K(\boldsymbol{x})$ of a porous medium is sampled at selected locations $\boldsymbol{x}_i$, $i = 1, \ldots, N$, as shown in Figure 1a. To simplify presentation, we assume that the source function $f(\boldsymbol{x})$ and boundary conditions are deterministic. Randomness in these quantities is additive and can be easily incorporated in subsequent analysis [20].

The flow equation (1) is under-determined, since the values of K at points other than $\{\boldsymbol{x}_i\}$, are unknown. To quantify the uncertainty in K, it is common [1–4] to model it

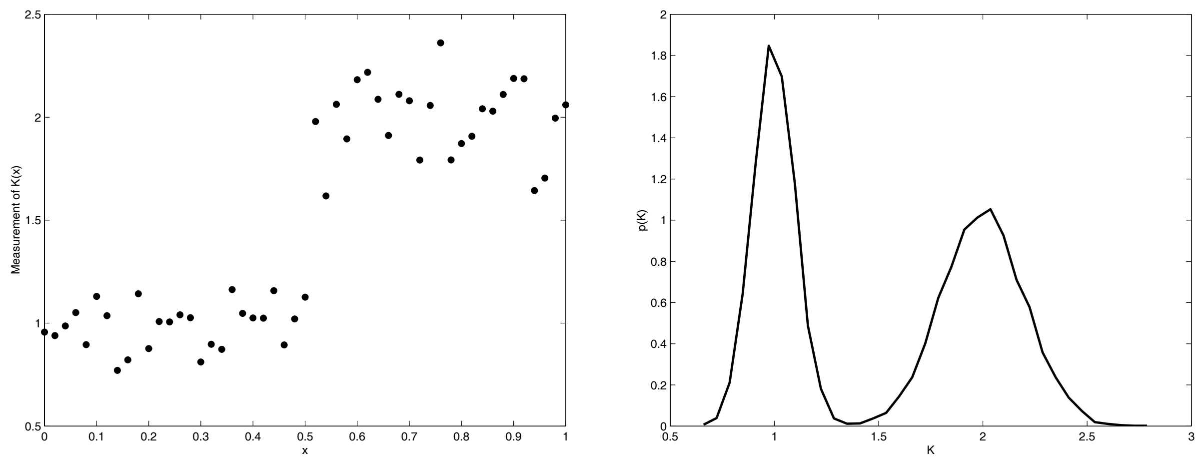

Figure 1. (a) Measurements of hydraulic conductivity K at selected locations in a one-dimensional porous medium, and (b) the corresponding statistically homogeneous probability density function $p_K(k)$.

as a random field, whose sample statistics are inferred from data. Figure 1b depicts a stationary (statistically homogeneous) probability density function $p_K(k)$ constructed from the K data in Figure 1a.

The randomness of K renders the flow equation (1) stochastic, so that its solution, a distribution of hydraulic head $h(\boldsymbol{x})$, is a probability density function or, equivalently, the corresponding ensemble moments.

We use the Reynolds decomposition to represent random fields $\mathcal{R} = \langle \mathcal{R} \rangle + \mathcal{R}'$ as the sum of their ensemble means $\langle \mathcal{R} \rangle$ and zero-mean fluctuations $\mathcal{R}'$. Taking the ensemble mean of (1) yields the mean flow equation

$$\nabla \cdot [\langle K \rangle \nabla \langle h \rangle + \langle K' \nabla h' \rangle] + f = 0, \qquad \boldsymbol{x} \in \Omega, \tag{2}$$

that contains the second mixed moment, $\langle K' \nabla h' \rangle$, an expression for which is not known. The need to approximate this term is often referred to as a closure problem. One of the most widely used closures relies on perturbation expansions in σ_K^2 or σ_Y^2, the variances of K or $Y = \ln K$ [1–4]. This requires the perturbation parameter σ_K^2 (or σ_Y^2) to be small, which is not the case for most multi-modal distributions, such as the one shown in Figure 1b.

A two-scale non-perturbative closure that we pursue here is based on a combination of random domain decompositions (RDD) and polynomial chaos expansions (PCE). RDD is used to decompose the computational domain into sub-domains that have convenient statistical properties, such as unimodality. PCE takes advantage of these properties to compute efficiently and accurately the statistics of the system states. The main features of this approach are described in the following sections.

3. GENERALIZED POLYNOMIAL CHAOS

The generalized polynomial chaos represents a second-order stochastic process $X(\omega)$, viewed as a function of a random event ω, as

$$X(\omega) = \sum_{j=0}^{\infty} a_j \Phi_j[\boldsymbol{\xi}(\omega)]. \tag{3}$$

Here $\{\Phi_j(\boldsymbol{\xi})\}$ are (multi-dimensional) orthogonal polynomials of the multi-dimensional random vector $\boldsymbol{\xi}$, which satisfies the orthogonality relation

$$\langle \Phi_i \Phi_j \rangle = \langle \Phi_i^2 \rangle \delta_{ij}, \tag{4}$$

where δ_{ij} is the Kronecker delta. The ensemble average of the product $\Phi_i \Phi_j$ is an inner product in the Hilbert space determined by the support of the random variables,

$$\langle p(\boldsymbol{\xi}) q(\boldsymbol{\xi}) \rangle = \int p(\boldsymbol{\xi}) q(\boldsymbol{\xi}) w(\boldsymbol{\xi}) d\boldsymbol{\xi}, \tag{5}$$

with $w(\boldsymbol{\xi})$ denoting the weighting function. In the discrete case, the above orthogonal relation takes the form

$$\langle p(\boldsymbol{\xi}) q(\boldsymbol{\xi}) \rangle = \sum_{\boldsymbol{\xi}} p(\boldsymbol{\xi}) q(\boldsymbol{\xi}) w(\boldsymbol{\xi}). \tag{6}$$

Equation (3) defines a one-to-one correspondence between the type of the orthogonal polynomials $\{\Phi\}$ and the type of the random variables $\boldsymbol{\xi}$. It is determined by choosing the type of orthogonal polynomials $\{\Phi\}$ in such a way that their weighting function $w(\boldsymbol{\xi})$ in the orthogonality relation (5) has the same form as the probability distribution function of the underlying random variables $\boldsymbol{\xi}$. For example, the weighting function of Hermite orthogonal polynomials is $\exp(-\boldsymbol{\xi}^T \boldsymbol{\xi}/2)/\sqrt{(2\pi)^n}$, and is the same as the probability density function of the $n-$dimensional Gaussian random variables $\boldsymbol{\xi}$. Hence, the classical Wiener polynomial chaos is an expansion of Hermite polynomials in terms of Gaussian random variables.

In practice, one has to truncate the infinite summation in (3), so that

$$X(\omega) = \sum_{j=0}^{M} a_j \Phi_j(\boldsymbol{\xi}), \tag{7}$$

where $\boldsymbol{\xi}$ is an $n-$dimensional random vector. If the highest order of a polynomial $\{\Phi\}$ is m, then the total number of expansion terms is $(M+1) = (n+m)!/(n!m!)$. Cameron and Martin [8] proved the convergence of Hermite-chaos expansion. The convergence of general non-Hermite expansions has been demonstrated both numerically [13–15] and analytically [21] for linear elliptic equations.

To demonstrate the robustness of the generalized polynomial chaos expansions, let us consider a highly non-Gaussian unimodal random variable $X = 1 + B(1,6) + N(0,4) + 5U(0,1) + E(3)$, where $B(1,6)$ is a β random variable with parameters $\alpha = 1$ and $\beta = 6$,

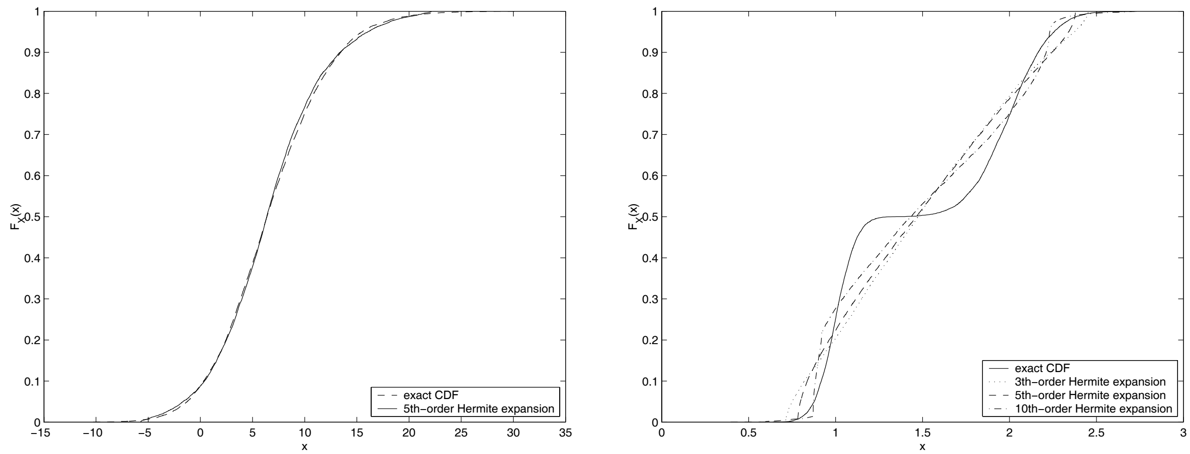

Figure 2. The Hermite expansions for the cumulative distribution functions of (a) a highly non-Gaussian, but unimodal random process and (b) a bi-modal process corresponding to the data in Figure 1.

$N(0,4)$ is a Gaussian random variable with zero mean and standard deviation 4, $U(0,1)$ is a uniform random variable in $(0,1)$, and $E(3)$ is an exponential random variable with mean 3. The total variance is $\sigma_X^2 = 26.8$, which clearly makes it unsuitable for small perturbation expansions. The fifth degree Hermite polynomial is sufficient to accurately approximate this distribution, as demonstrated in Figure 2a.

The situation is radically different, when one deals with system parameters that are neither Gaussian, nor unimodal. Figure 2b demonstrates that the Hermite polynomials are not adequate to represent the bi-modal cumulative density function corresponding to $p_K(k)$ shown in Figure 1b.

4. RANDOM DOMAIN DECOMPOSITION

To apply the generalized polynomial chaos expansions to systems whose parameters are multi-modal, we reformulate the problem (1) in terms of the random domain decomposition [17,18]. Within this framework, the randomness of $K(\boldsymbol{x})$ stems from two factors: large-scale uncertainty in the spatial arrangement of N sub-domains $\{\Omega_i\}_{i=1}^N$ (or, equivalently, the boundaries $\{\Gamma_{ij}\}$ between sub-domains Ω_i and Ω_j for $i \neq j$), and small-scale uncertainty of K within each sub-domain. Then $p_K(k)$, the probability density function of K, is replaced with a joint probability density function $p_K(k,\gamma) = p_K(k|\gamma)p_\Gamma(\gamma)$.

The reconstruction of $p_\Gamma(\gamma)$, the probability density function of the random boundary Γ, from measurements of $K(\boldsymbol{x})$ is a subject of an ongoing research [22], which we do not pursue here. The reverse relationship, however, is more straightforward. Indeed, if p_{K_i} $(i = 1,\ldots,N)$ are the probability density functions of $K = K_i$ within each random sub-domain Ω_i, then the probability density function of the mixture is

$$p_K(k;\boldsymbol{x}) = \sum_i P_i(\boldsymbol{x}) p_{K_i}(k). \tag{8}$$

Here $P_i(\boldsymbol{x})$ is the probability that a point $\boldsymbol{x} \in \Omega_i$, which is uniquely defined by $p_\Gamma(\gamma)$.

Let $K(\boldsymbol{x}) = K_i(\boldsymbol{x})\mathbb{I}_{\Omega_i}(\boldsymbol{x})$, where $\mathbb{I}_{\Omega_i}(\boldsymbol{x})$ is the indicator function. Then (1) can be rewritten as

$$\nabla \cdot K_i \nabla h(\boldsymbol{x}) + f = 0, \qquad \boldsymbol{x} \in \Omega_i, \qquad i = 1, \ldots, N. \tag{9}$$

Boundary conditions for (1) are supplemented by the conditions of the continuity of the state variable, h, and the normal component of $\boldsymbol{q}_{ij} = -K\nabla h$, the flux across the random boundaries $\Gamma_{ij}, \forall i \neq j$. Generalized polynomial chaos expansions are then applied to the relevant random fields within each sub-domain, and the resulting probability density function of h is matched along the random boundaries $\Gamma_{ij}, \forall i \neq j$.

To simplify presentation, we consider a two-subdomain problem $\Omega = \Omega_1 \cup \Omega_2$ and denote the boundary between these two sub-domains by $\Gamma = \Gamma_{12}$.

4.1. Conditional statistics

In the first step of the averaging procedure, we apply the polynomial chaos expansion within each sub-domain Ω_i $(i = 1, 2)$,

$$h_i(x) = \sum_{m=0}^{M} \hat{h}_{i,m}\Phi_m(\boldsymbol{\xi}), \quad K_i(x) = \sum_{m=0}^{M} \hat{K}_{i,m}\Phi_m(\boldsymbol{\xi}), \quad \text{for } x \in \Omega_i. \tag{10}$$

By applying these expansions to (9) and using a Galerkin projection onto each basis of $\{\Phi_m\}_{m=0}^M$, we obtain

$$\sum_{m=0}^{M}\sum_{n=0}^{M} \nabla \cdot \hat{K}_{i,m} \nabla \hat{h}_{i,n} \langle \Phi_m \Phi_n \Phi_l \rangle + f_i \delta_{l0} = 0, \qquad i = 1, 2. \tag{11}$$

Denoting $H_i = [\hat{h}_{i,0}, \ldots, \hat{h}_{i,M}]^T$ and $F_i = [f_i, 0, \ldots]^T$ allows us to rewrite (11) in a matrix form

$$\nabla \cdot B_i^T \nabla H_i + F_i = 0, \qquad i = 1, 2, \tag{12}$$

where $B_i(x) = [b_{i,nl}]_{n,l=0}^M$ is a symmetric matrix of size $(M+1) \times (M+1)$ whose entries are

$$b_{i,nl} = \sum_{m=0}^{M} \hat{K}_{i,m} \langle \Phi_m \Phi_n \Phi_l \rangle, \qquad i = 1, 2. \tag{13}$$

The continuity of head and flux is enforced at the inner boundary, i.e.,

$$h|_{x=\Gamma^-} = h|_{x=\Gamma^+}, \qquad K(x)\frac{dh}{dx}\bigg|_{x=\Gamma^-} = K(x)\frac{dh}{dx}\bigg|_{x=\Gamma^+},$$

which leads to

$$H_1|_{x=\Gamma^-} = H_2|_{x=\Gamma^+}, \qquad B_1^T \frac{dH_1}{dx}\bigg|_{x=\Gamma^-} = B_2^T \frac{dH_2}{dx}\bigg|_{x=\Gamma^+}. \tag{14}$$

Equations (12), (14), and external boundary conditions define a complete system for the stochastic head h conditioned on the random interface Γ. Solutions of this system provide the conditional statistics of hydraulic head, such as its conditional mean $\langle h(x)|\Gamma\rangle = \hat{h}_0(x)$ conditional variance $\sigma_h^2|\Gamma = \sum_{m=1}^M \hat{h}_m^2(x)\langle \Phi_m^2 \rangle$.

4.2. Averaging over geometries

In the second step, the statistics of h is obtained by averaging over the random geometry Γ, e.g.,

$$\begin{aligned}\langle h(x)\rangle &= \iint h(x;k,\gamma)p_K(k,\gamma)dkd\gamma = \iint h(x;k,\gamma)p_K(k|\gamma)p_\Gamma(\gamma)dkd\gamma \\ &= \int \langle h(x)|\gamma\rangle p_\Gamma(\gamma)d\gamma.\end{aligned} \tag{15}$$

To evaluate the above integral, we employ a quadrature rule,

$$\langle h(x)\rangle = \sum_{q=1}^{Q} \langle h(x)|\alpha_q\rangle w_q. \tag{16}$$

Here $\{\alpha_q, w_q\}_{q=1}^{Q}$ are the quadrature points and corresponding weights of the orthogonal polynomials $g_Q(\gamma)$, which satisfy

$$\int g_m(\gamma)g_n(\gamma)p_\Gamma(\gamma)d\gamma = d_m^2\delta_{mn}.$$

For example, if Γ can be parameterized by Gaussian random variables, $\{g_m\}$ take form of the Hermite polynomials. Expressions for the head variance and other statistics can be obtained in a similar manner.

5. COMPUTATIONAL EXAMPLE

To simplify presentation, we consider the one-dimensional version of (1) with $f = 0$ defined on $\Omega = (0,1)$. This equation is subject to the following boundary conditions, $q(0) = q_0$ and $h(1) = 0$.

Two random materials $[0,\alpha)$ and $(\alpha,1]$ are joined at the random location α. Then $p_K(k)$, the probability density function of K shown in Figure 1b, is replaced with a joint probability density function $p_K(k,\alpha) = p_K(k|\alpha)p_\alpha(\alpha)$. We assume that both $Y_i = \ln K_i(x)$ are Gaussian random fields with the exponential correlation functions. We further assume that the fields $K_1(x)$ and $K_2(x)$ are mutually uncorrelated, and that α is a Gaussian random variable.

This problem admits an analytical solution for random h,

$$h(x) = q_0\mathcal{H}(\alpha - x)\left[\int_x^\alpha \frac{ds}{K_1(s)} + \int_\alpha^1 \frac{ds}{K_2(s)}\right] + q_0\mathcal{H}(x-\alpha)\int_x^1 \frac{ds}{K_2(s)}, \tag{17}$$

where $\mathcal{H}(z)$ is the Heaviside function,

$$\mathcal{H}(z) = \begin{cases} 1 & z \geq 0 \\ 0 & z < 0. \end{cases} \tag{18}$$

Equation (17) leads to exact expressions for the mean and variance of hydraulic head.

Figure 3 demonstrates that the ensemble mean, $\langle h\rangle$, and standard deviation, σ_h, of h computed by the RDD-PCE approach outlined in Section 4 practically coincide with their exact counterparts. The first twenty terms are retained in the polynomial expansion. In these calculations, we set $\alpha = N(0.5, 0.05)$, $Y_1(x) = N(0, 0.1)$, $Y_2(x) = N(2, 0.2)$, $l_{Y_1} = 5$, $l_{Y_2} = 1$, and $q_0 = 1$.

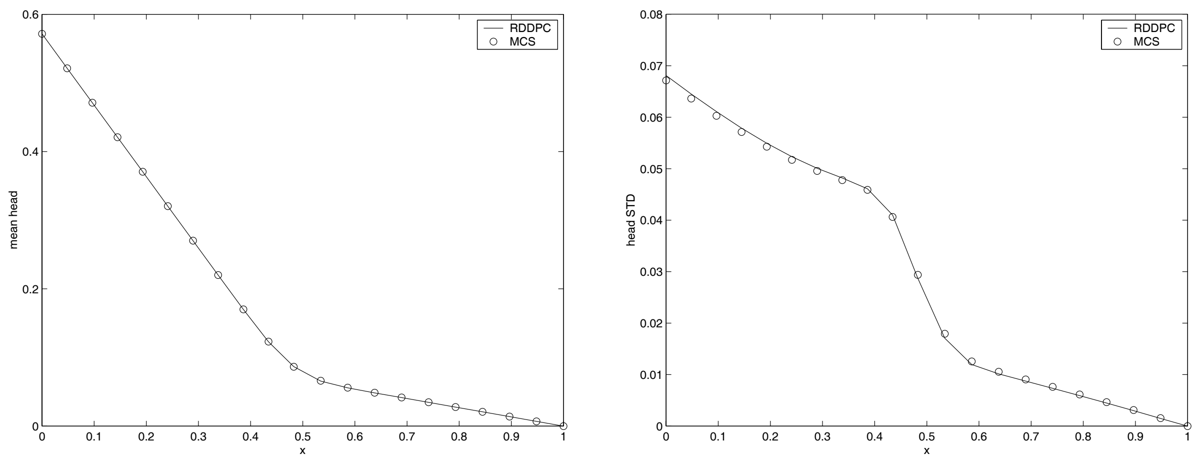

Figure 3. The ensemble mean (a) and standard deviation (b) of h given by the exact analytical solutions (circles) and by the polynomial chaos expansion (solid lines).

6. CONCLUSIONS

Polynomial chaos expansions (PCE) provide a valuable tool for quantifying uncertainty in physical systems with uncertain (random) system parameters. However they might become less efficient if these parameter have highly non-Gaussian, multi-modal distributions and/or short correlation lengths. To extend the range of applicability of PCE approaches, we combined them with a random domain decomposition (RDD). We used one-dimensional flow to demonstrate that this combined approach provides robust, perturbation-free approximations for the statistics of hydraulic head.

REFERENCES

1. G. Dagan, Flow and transport in porous formations, Springer-Verlag, New York, 1989.
2. J. H. Cushman, The physics of fluids in hierarchical porous media: Angstroms to miles, Kluwer Academic Pub., New York, 1997.
3. G. Dagan, S. P. Neuman (Eds.), Subsurface flow and transport: a stochastic approach, Cambridge, New York, 1997.
4. Y. Rubin, Applied stochastic hydrogeology, Oxford University Press, New York, 2003.
5. D. M. Tartakovsky, A. Guadagnini, M. Riva, Stochastic averaging of nonlinear flows in heterogeneous porous media, J. Fluid Mech. 492 (2003) 47 – 62.
6. K. D. Jarman, T. F. Russell, Moment equations for stochastic immiscible flow, SIAM J. on Multiscale Modeling and Simulation .
7. N. Wiener, The homogeneous chaos, Amer. J. Math. 60 (1938) 897–936.
8. R. Cameron, W. Martin, The orthogonal development of nonlinear functionals in series of Fourier-Hermite functionals, Ann. Math. 48 (1947) 385–392.
9. R. Ghanem, Scales of fluctuation and the propagation of uncertainty in random porous media, Water Resources Research 34 (1998) 2123.
10. A. Chorin, Hermite expansion in Monte-Carlo simulations, J. Comput. Phys. 8 (1971)

472–482.
11. H. Matthies, C. Bucher, Finite elements for stochastic media problems, Comput. Methods Appl. Mech. Engrg. 168 (1999) 3–17.
12. R. Askey, J. Wilson, Some basic hypergeometric polynomials that generalize Jacobi polynomials, Memoirs Amer. Math. Soc., AMS, Providence RI 319.
13. D. Xiu, G. Karniadakis, Modeling uncertainty in steady state diffusion problems via generalized polynomial chaos, Comput. Methods Appl. Math. Engrg. 191 (2002) 4927–4948.
14. D. Xiu, G. Karniadakis, The Wiener-Askey polynomial chaos for stochastic differential equations, SIAM J. Sci. Comput. 24 (2) (2002) 619–644.
15. D. Xiu, G. Karniadakis, Modeling uncertainty in flow simulations via generalized polynomial chaos, J. Comput. Phys. 187 (2003) 137–167.
16. D. Xiu, G. Karniadakis, A new stochastic approach to transient heat conduction modeling with uncertainty, Inter. J. Heat Mass Trans. 46 (2003) 4681–4693.
17. C. L. Winter, D. M. Tartakovsky, Mean flow in composite porous media, Geophys. Res. Lett. 27 (12) (2000) 1759 – 1762.
18. C. L. Winter, D. M. Tartakovsky, Groundwater flow in heterogeneous composite aquifers, Water Resour. Res. 38 (8) (2002) 10.1029/2001WR000450.
19. C. L. Winter, D. M. Tartakovsky, A. Guadagnini, Numerical solutions of moment equations for flow in heterogeneous composite aquifers, Water Resour. Res. 38 (5) (2002) 10.1029/2001WR000222.
20. D. M. Tartakovsky, S. P. Neuman, Extension of "Transient flow in bounded randomly heterogeneous domains 1. Exact conditional moment equations and recursive approximations", Water Resour. Res. 35 (6) (1999) 1921 – 1925.
21. I. Babuška, R. Tempone, G. Zouraris, Galerkin finite element approximations of stochastic elliptic differential equations, Tech. Rep. 02-38, TICAM (2002).
22. L. Guadagnini, A. Guadagnini, D. M. Tartakovsky, Probabilistic reconstruction of geologic facies, J. Hydrol. (to appear).