

Article

# Analysis of the Stochastic Quarter-Five Spot Problem Using Polynomial Chaos

Hesham AbdelFattah <sup>1</sup>, Amnah Al-Johani <sup>2</sup> and Mohamed El-Beltagy <sup>1,\*</sup>

<sup>1</sup> Engineering Mathematics and Physics Department, Engineering Faculty, Cairo University, Giza 12613, Egypt; heshammustafa1992@gmail.com

<sup>2</sup> Mathematics Department, Science Faculty, Tabuk University, Tabuk 71491, Saudi Arabia; aalgohani@ut.edu.sa

\* Correspondence: zbeltagy@eng.cu.edu.eg

Academic Editor: David Faux

Received: 26 May 2020; Accepted: 13 July 2020; Published: 24 July 2020



**Abstract:** Analysis of fluids in porous media is of great importance in many applications. There are many mathematical models that can be used in the analysis. More realistic models should account for the stochastic variations of the model parameters due to the nature of the porous material and/or the properties of the fluid. In this paper, the standard porous media problem with random permeability is considered. Both the deterministic and stochastic problems are analyzed using the finite volume technique. The solution statistics of the stochastic problem are computed using both Polynomial Chaos Expansion (PCE) and the Karhunen-Loeve (KL) decomposition with an exponential correlation function. The results of both techniques are compared with the Monte Carlo sampling to verify the efficiency. Results have shown that PCE with first order polynomials provides higher accuracy for lower (less than 20%) permeability variance. For higher permeability variance, using higher-order PCE considerably improves the accuracy of the solution. The PCE is also combined with KL decomposition and faster convergence is achieved. The KL-PCE combination should carefully choose the number of KL decomposition terms based on the correlation length of the random permeability. The suggested techniques are successfully applied to the quarter-five spot problem.

**Keywords:** porous media; quarter-five spot problem; random variations; polynomial chaos; Karhunen-Loeve decomposition

## 1. Introduction

Fluid flow in porous media is a subject of great interest due to its wide range of applications such as water/oil flow in underground reservoirs, modeling the underground spreading of toxic, and biological applications such as the transport in lungs and arteries [1]. When dealing with porous media, such as geological formations, we do not have complete information about their properties, such as permeability and porosity. This leads to uncertainty about the values of the formation properties and, in turn, in calculating flow in such formations [2]. For evaluating such uncertainties, a probabilistic framework is to be used. The input data are considered as random quantities, which are described in terms of stochastic processes.

Uncertainty quantification (UQ) methods are used for analyzing the system using a model by the parameterization of the uncertain input data (such as permeability) using a set of independent random variables. Then, the model is solved to determine the response of the system to the given random input data [3].

Most of the numerical models used for uncertainty quantification are based on the Monte Carlo (MC) sampling method [4–9]. This method is known for the simplicity of its implementation; however, it requires a large number of iterations of the numerical model in order to provide reliable results and

therefore may be not feasible for complex simulators. In order to overcome these limitations, other variants to the standard MC method were developed such as the multilevel Monte Carlo method (MLMC) [4], Sequential Monte Carlo (SMC) Methods [5], and quasi-Monte Carlo methods [6]. In addition to the MC method and its variations, other approaches including analytical models [7] and reduced-order models [8,9] were introduced. The Karhunen-Loève (KL) expansion is an order reduction method that can be used to provide a low-dimensional representation for a Gaussian random field by performing a projection of the original system with respect to a set of spatially dependent basis and a set of independent standard normal random coefficients, where the eigenvalues and eigenvectors of the autocorrelation function of the Gaussian field are utilized in the expansion [8–10].

Another method is the polynomial chaos expansion (PCE) [11], where the random variables are expanded using polynomial chaos basis, then a set of equations are solved using the Galerkin technique to determine the expansion coefficients [12]. For the classical polynomial chaos expansion, the polynomial chaos basis is the Hermite polynomials, and the random variables are Gaussian. Theoretically, convergence is achieved for any  $L_2$  functionals in the random space [13]; however, only Gaussian and near-Gaussian random fields achieve optimal convergence rate. A more general approach, the generalized polynomial chaos (gPC), was proposed in [14], where the random variables are not restricted to Gaussian random variables and the expansion is in terms of polynomials chosen from the hypergeometric polynomials of the Askey scheme [15]. However, gPC can be used only when the input pdf is known in order to choose the proper basis functions, which is not the case in many applications [16]. To overcome this issue, the arbitrary polynomial chaos (aPC) technique was proposed [16,17], where in aPC, the polynomial chaos expansion is based on the moments of the random data instead of the PDF.

When dealing with Gaussian processes, the convergence of the KL expansion is optimal [18], however, the convergence rate may be much slower for other types of processes where optimal convergence is achieved based on the particular choice, made a priori, of the basis functions used in the expansion. Therefore, PCE and KL become less accurate for nonlinear dynamics and long-time integrations due to the fixed basis and the increasing complexity with the time. To overcome that issue, the dynamically-orthogonal (DO) [19] and the bi-orthogonal (BO) [20] techniques were introduced to allow for time-dependent basis, and therefore, more accurate for long time intervals. In both techniques, the stochastic coefficients and the spatial basis evolve according to the stochastic partial differential equation (SPDE) while preserving an orthonormality conditions for the basis functions.

The purpose of this paper is to solve one of the standard porous media problems, namely, the quarter-five spot problem after imposing uncertainty to the problem. The quarter five spot problem is used to model the case when a regular set of injection and production wells is present in the reservoir; as this problem already illustrates the features and difficulties of the general problem [21]. The problem is solved using both PCE and KL techniques and the results are compared with the MC method to verify the efficiency of each technique.

The remainder of this paper is organized as follows. First, in Section 2, the mathematical model is introduced, then the details of the problem are given in Section 3. Finally, in Section 4, results and conclusions are discussed.

## 2. Mathematical Statement of the Quarter-Five Spot Problem

### 2.1. The Deterministic Problem

Consider the problem of fluid flow in porous medium for a single-phase flow with an incompressible rock and fluid under isothermal conditions [22].

The problem is presented by the following partial differential equation (PDE):

$$\nabla \cdot \left[ \frac{-K}{\mu} (\nabla p + \rho g \nabla z) \right] = \frac{q}{\rho}, \quad (1)$$

where  $\rho$  denotes the density of the fluid,  $q$  denotes sources and sinks (out flow and in flow of fluid per unit volume),  $K$  denotes the permeability,  $z$  is the vertical coordinate, and  $\mu$  is the fluid viscosity.

Let  $G = -g\nabla z$  denote the downward force for gravity. Equation (1) can be expressed as

$$\nabla \cdot \left( \left( \frac{-K}{\mu} \frac{\partial P}{\partial x}, \frac{-K}{\mu} \frac{\partial P}{\partial y} \right) + \left( -\rho \frac{\partial G}{\partial x}, -\rho \frac{\partial G}{\partial y} \right) \right) = \frac{q}{\rho}, \quad (2)$$

As gravity forces are approximately constant inside a reservoir domain, we get

$$\frac{\partial G}{\partial x} = \frac{\partial G}{\partial y} = 0, \quad (3)$$

Therefore,

$$\nabla \cdot \left( -K \frac{\partial P}{\partial x}, -K \frac{\partial P}{\partial y} \right) = \frac{q}{\rho}, \quad (4)$$

where we assumed constant viscosity  $\mu$ .

If we assume water phase with  $\rho = 1$ , we get Equation (1) in the form

$$-K \frac{\partial^2 P}{\partial x^2} - K \frac{\partial^2 P}{\partial y^2} - \frac{\partial K}{\partial x} \frac{\partial P}{\partial x} - \frac{\partial K}{\partial y} \frac{\partial P}{\partial y} = q, \quad (5)$$

or in compact form,

$$\nabla \cdot \vec{v} = q, \quad \vec{v} = -(K\nabla P), \quad (6)$$

where  $\vec{v}$  is the macroscopic Darcy velocity.

Equation (6) is a second order linear PDE. Additionally, Equation (6) is elliptic as the eigenvalues of its coefficient matrix are all negative.

## 2.2. The Stochastic Problem

Due to our incomplete knowledge about the medium properties, subsurface flow involves some degree of uncertainty, and therefore we need a stochastic description of the medium properties [23]. We can impose uncertainty to the quarter-five spot problem by considering the hydraulic conductivity (permeability) as a random field  $K = K(x, \omega)$ , where  $x$  is the spatial coordinate and  $\omega$  is the random outcome on a given probability space  $\Omega$ . This, in turn, makes the solution  $P$  a random field. In our model, the boundary conditions and the source term are kept deterministic.

Thus, Equation (6) becomes an SPDE stated as

$$\nabla \cdot (-K(x; \omega) \nabla P(x; \omega)) = q, \quad (7)$$

The PCE method can be used to represent a second order process  $f \in L_2(\Omega)$ , where  $L_2(\Omega)$  is the space of square integrable  $L_2$  functions, such that a random field  $u(x; \omega)$  can be expressed as [24]

$$\begin{aligned} u(x; \omega) = & u_0(x) + \sum_{i_1=1}^{\infty} u_{i_1}(x) \varphi_1(\xi_{i_1}(\omega)) \\ & + \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{i_1} u_{i_1 i_2}(x) \varphi_2(\xi_{i_1}(\omega), \xi_{i_2}(\omega)) \\ & + \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{i_1} \sum_{i_3=1}^{i_2} u_{i_1 i_2 i_3}(x) \varphi_3(\xi_{i_1}(\omega), \xi_{i_2}(\omega), \xi_{i_3}(\omega)) + \dots, \end{aligned} \quad (8)$$

To simplify writing the equations, we can eliminate the arguments when there is no ambiguity.  $\varphi_n(\xi_{i_1}, \xi_{i_2}, \dots, \xi_{i_n})$  denotes the orthogonal polynomial of degree  $n$  with respect to the multidimensional random variables  $\xi = (\xi_{i_1}, \dots, \xi_{i_n}, \dots)$ .

If  $\xi_{i_1}, \xi_{i_2}, \dots, \xi_{i_n}$  are independent standard Gaussian random variables, then  $\varphi_n(\xi_{i_1}, \xi_{i_2}, \dots, \xi_{i_n})$  represent the multidimensional Hermite Polynomials of degree  $n$  expressed as

$$\varphi_n(\xi_{i_1}, \dots, \xi_{i_n}) = (-1)^n e^{\frac{1}{2}\xi^T \xi} \frac{\partial^n}{\partial \xi_{i_1} \dots \partial \xi_{i_n}} \left[ e^{-\frac{1}{2}\xi^T \xi} \right] \quad (9)$$

For Gaussian random variables, Hermite polynomials are the most suitable orthogonal basis [11]. For other distributions, the basis polynomials are selected to optimize the convergence rate. This generalization is commonly known as gPC [25].

For computational purposes, if we have a random vector  $\xi$  of dimension  $N$ , then Equation (8) can be written in terms of a finite number of Hermite polynomials as

$$\mathbf{u}(x; \omega) = \sum_{i=0}^P u_i(x) \psi_i(\xi), \quad (10)$$

where the total number of expansion terms  $P + 1$  is determined by the random vector dimension  $N$  and the order  $d$  of the PCE [23].

$$P + 1 = \frac{(N + d)!}{N!d!}, \quad (11)$$

The functions  $\psi_i$  construct orthogonal basis in  $L_2(\Omega)$  with the orthogonality relation [24]

$$\langle \psi_i, \psi_j \rangle = \langle \psi_i^2 \rangle \delta_{ij}, \quad (12)$$

where  $\delta_{ij}$  is the Kronecker delta and  $\langle \cdot, \cdot \rangle$  is the ensemble average, that is, the inner product in the Hilbert space of the variables  $\xi$  defined by

$$\langle f(\xi), g(\xi) \rangle = \int_{\Omega} f(\xi) g(\xi) W(\xi) d\xi, \quad (13)$$

where  $W(\xi)$  is the weighting function corresponding to the polynomial chaos basis  $\psi_i$ .

In case that the polynomial chaos basis are the Hermite polynomials,  $W(\xi)$  is given by

$$W(\xi) = \frac{1}{\sqrt{(2\pi)^N}} e^{-\frac{1}{2}\xi^T \xi}, \quad (14)$$

where  $N$  is the number of independent Gaussian random variables.

The mean square error of the truncated PCE expansion,  $u(x; \omega)$ , is expressed as

$$e(x) = \sqrt{\mathbb{E}[u(x; \omega) - u_e(x; \omega)]^2}, \quad (15)$$

where  $u_e(x, \omega)$  is the exact solution and  $\mathbb{E}$  is the expectation operator.

For the case of univariate Gaussian stochastic germ, the following theorem provides an error bound. Theorem 1 [25]. Given  $u_e(x, \omega) \in L_2$ , being  $k$  times continuously differentiable, the convergence rate

$$\|u_e(x, \omega) - \sum_{i=0}^P u_i(x) W_i(\xi)\| \leq \frac{1}{\prod_{i=0}^{k-1} \sqrt{(n-i+1)}} \|u^{(k)}\|, \quad (16)$$

is obtained for an approximation of  $u_e(x, \omega)$  by a finite number of Hermite polynomials  $\{W_i\}_{i=0}^P$ , where the norm  $\|\cdot\|$  is induced by the inner product (14).

The KL expansion can be viewed as a special case of the PCE where only first-order polynomials have non-zero coefficients [26]. The KL expansion involves the spectral expansion of the correlation function of the process under consideration and is used to reduce the dimensionality of the random space [24]. Consider a second-order process  $Z(x; \omega)$  with a correlation function  $C(x, y)$ . The correlation function should be real, symmetric, and positive definite [14].

We can represent  $Z(x; \omega)$  as

$$Z(x; \omega) = \mathbb{E}[Z(x, \cdot)] + \sum_{i=1}^{\infty} \sqrt{\lambda_i} \phi_i(x) \xi_i(\omega), \quad (17)$$

where  $Z(x, \cdot) : \Omega \rightarrow \mathbb{R}$  is a random variable,  $\mathbb{E}[Z(x, \cdot)]$  is mean of the process,  $\xi_i(\omega)$  a set of uncorrelated random variables,  $\lambda_i$  and  $\phi_i(x)$  are the eigenvalues and eigenfunctions of the correlation function, respectively, i.e.,

$$\int_{\Omega} C(x, y) \phi_i(y) dy = \lambda_i \phi_i(x), \quad (18)$$

The KL expansion is effective when the correlation function is known, as it is characterized by an optimal convergence in the mean square sense [25]. However, the correlation function might not be known in some cases, which is a limitation of the KL expansion method.

### 3. Solution Implementation

#### 3.1. The Deterministic Problem

Recall Equation (6):

$$\nabla \cdot \vec{v} = q, \quad \vec{v} = -(K \nabla p), \quad (19)$$

To have a complete solution, we need to specify the boundary conditions. Closed flow systems where there is no fluid flow across the external boundaries are modeled by specifying homogeneous Neumann boundary conditions:

$$\vec{v} \cdot \vec{n} = 0, \quad \text{for } \vec{x} \in \partial\Omega, \quad (20)$$

where  $\vec{n}$  is the normal at the boundary  $\partial\Omega$  of the computational domain and  $\vec{x}$  is any spatial point that exists on the domain's boundary  $\partial\Omega$ .

There are several numerical methods used to solve (4), three of these methods were discussed in [22]. The first method is a cell-centered finite-volume method called the two-point flux-approximation (TPFA) scheme that is widely used in the oil industry. The TPFA scheme uses two cells pressure averages to approximate the flux  $v_{i,k}$  across each face of the cell. In some cases, the orientation of the computational grid will affect the results of the TPFA scheme; thus, another finite-volume method called the multi-point flux-approximation (MPFA) scheme is used to overcome the inconsistencies of the TPFA scheme. Instead of using only two points such as the TPFA, the MPFA scheme utilizes each of the neighboring cell averages to approximate the flux. Another alternative to the MPFA scheme is the mixed finite-element method where the fluxes over cell edges are considered as unknowns in addition to the pressure values. However, due to the simplicity and computational speed of the TPFA, we will restrict our attention to this method. First, we write Equation (17) in the integral form inside a cell  $\Omega_i$  as the control volume.

$$\int_{\partial\Omega_i} \vec{v} \cdot \vec{n} ds = \int_{\Omega_i} q dV, \quad (21)$$

To compute the flux  $v_{i,k}$  across each face of the cell we use Darcy's law,

$$v_{i,k} = \int_{\Gamma_{i,k}} \vec{v} \cdot \vec{n} ds, \quad \Gamma_{i,k} = \partial\Omega_i \cap \partial\Omega_k, \quad (22)$$

where  $\Gamma_{i,k}$  is the interface between the two cells' boundaries  $\partial\Omega_i$  and  $\partial\Omega_k$  with an area  $A_{i,k}$  and a normal vector  $\vec{n}_{i,k}$ ,  $\Gamma_{i,k}$  is the opposite interface and has identical area  $A_{i,k} = A_{k,i}$  but opposite normal vector  $\vec{n}_{i,k} = -\vec{n}_{k,i}$ .

The integral can be approximated as

$$v_{i,k} \approx -A_{i,k} (K_i \nabla p_{i,k}) \cdot \vec{n}_{i,k}, \quad (23)$$

where  $\nabla p_{i,k}$  is the pressure gradient at the centroid of the interface  $\Gamma_{i,k}$ . Finite difference is used to approximate  $\nabla p_{i,k}$  using the pressure  $\widetilde{p}_{i,k}$  at the interface's centroid and the average pressure  $p_i$  inside the cell  $\Omega_i$ ,

$$v_{i,k} \approx -A_{i,k} K_i \frac{(\mathbf{p}_i - \widetilde{p}_{i,k}) \vec{x}_{i,k}}{\|\vec{x}_{i,k}\|} \cdot \vec{n}_{i,k} = T_{i,k} (p_i - \widetilde{p}_{i,k}), \quad (24)$$

where  $\vec{x}_{i,k}$  is the position vector between the center of the cell  $\Omega_i$  and the centroid of the interface  $\Gamma_{i,k}$ , and  $T_{i,k}$  is the one-sided transmissibility associated with the cell  $\Omega_i$ . Therefore, we have for cell  $\Omega_i$  the relation

$$T_{i,k}^{-1} v_{i,k} = (p_i - \widetilde{p}_{i,k}), \quad (25)$$

Similarly, we have for cell  $\Omega_k$  the relation

$$T_{k,i}^{-1} v_{k,i} = (p_k - \widetilde{p}_{k,i}), \quad (26)$$

By imposing the continuity of fluxes, we have  $v_{i,k} = -v_{k,i} = v_{ik}$  and  $\widetilde{p}_{i,k} = \widetilde{p}_{k,i} = \widetilde{p}_{ik}$ , and after eliminating  $\widetilde{p}_{ik}$  we get

$$v_{ik} = [T_{i,k}^{-1} + T_{k,i}^{-1}]^{-1} (p_i - p_k) = T_{ik} (p_i - p_k), \quad (27)$$

where  $T_{ik}$  is the transmissibility associated with the connection between the two cells  $\Omega_i$  and  $\Omega_k$ . Therefore, TPFA uses the cell averages  $p_i$  and  $p_k$ , to approximate the flux across the interface  $\Gamma_{i,k}$  between the cells  $\Omega_i$  and  $\Omega_k$ .

Finally, when substituting  $v_{ik}$  in Equation (19) we get the following system of equations,

$$\sum_k T_{ik} (p_i - p_k) = q_i, \quad (28)$$

where

$$T_{ik} = 2\Delta y \left[ \frac{\Delta x_i}{K_i} + \frac{\Delta x_k}{K_k} \right]^{-1}, \quad (29)$$

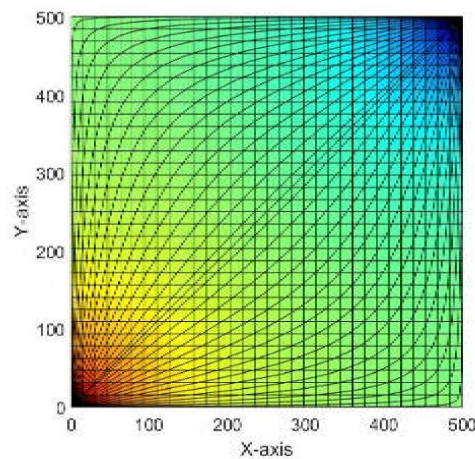
Therefore we get the following system of equations,

$$Au = f, \quad (30)$$

where  $A$  is a sparse, penta-diagonal symmetric matrix. Moreover, when dealing with a homogenous permeability field,  $A$  is a Toeplitz matrix. The system is made positive definite, which can be effectively solved using the Cholesky decomposition algorithm [27].

## Numerical Example

Equation (4) is solved on a  $32 \times 32$  uniform Cartesian grid covering a  $500 \times 500 \text{ m}^2$  area with a homogeneous isotropic permeability  $K = 100 \text{ md}$  with porosity of 0.2 for all points in the grid. There is an injection well at the origin with a flow rate of  $1 \text{ m}^3/\text{day}$  and 4 production wells with flow rate of  $-1 \text{ m}^3/\text{day}$  at the points  $(\pm 1, \pm 1)$ . We use the homogeneous Neumann boundary conditions. The flow in the five-spot is symmetric about both the coordinate axes. Therefore, the domain can be reduced to a quarter of its size. This problem is called the quarter-five spot problem. Figure 1 shows the pressure distribution on the grid while illustrating the position of the injection and production wells on the bottom left and top right corners of the grid, respectively. The area of largest pressure's value is the immediate zone surrounding the injection well (shown in red), then the pressure decreases gradually till it reaches its lowest value at the production well. In order to obtain a better view of the flow field, Pollock's method [28] for semi-analytical tracing of streamlines is used to track the path of flow of the particles as they move from cell to cell through the cells' centers. Additionally, traces are used, which are particles that flow with the fluid without changing the characteristics of the fluid.



**Figure 1.** Solution of the quarter-five spot problem on a  $32 \times 32$  uniform grid. The plot shows the pressure distribution on each point in the grid, where the highest pressure values are in the bottom left corner and the pressure decreases gradually until it reaches the lowest value at the top right corner.

The concentration of a tracer is given by a continuity equation

$$\frac{\partial(\phi C)}{\partial t} + \nabla \cdot (\vec{v} C) = q_C, \quad (31)$$

where  $C$  is the concentration,  $q_C$  is the inflow rate per unit volume for the tracer, and  $\phi$  is the rock porosity. In the case of steady state flow, with tracers whose concentration do not change with fluid compression or expansion and with concentration equal to one at the inflow boundary, we get the simplified equation

$$\nabla \cdot (\vec{v} C) = q_C \quad C_{inflow} = 1, \quad (32)$$

We define the time-of-flight  $\tau$  as the time required by a fluid particle to travel a given distance along a streamline.  $\tau$  is given by the differential equation

$$\vec{v} \cdot \nabla \tau = \phi. \quad (33)$$

Applying the time flight to a tracer, the communication patterns of the reservoir can be studied effectively. Therefore, in order to estimate the speed of flow inside the reservoir, we solve equation (6) subject to  $\tau = 0$  at the inflow to get the forward time-of-flight from inflow point and into the reservoir,

and backward time-of-flight from outflow points and backwards into the reservoir. The sum of forward and backward times gives the total time required by a particle to enter the reservoir from inflow and travel through until it exits through the outflow. Figure 2 shows total travel time to illustrate the areas of high speed flow and those of stagnant flow.

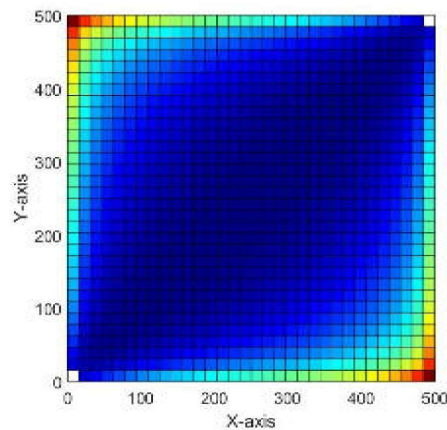


Figure 2. Total travel time to illustrate the areas of high speed flow and those of stagnant flow.

### 3.2. The Stochastic Quarter-Five Spot Problem

We now proceed to solve Equation (7) while considering the uncertainty of the permeability field  $K(x; \omega)$ . The problem is first solved using only PCE, and then we will use KL expansion to reduce the dimension of the input space then apply PCE. For both scenarios, the pressure solution statistics are presented and the results are compared with the MC solution with a sample size of 10,000.

- First: Using PCE

Let permeability tensor  $K$  be a random field with a Gaussian distribution of mean  $100 \text{ md}$  and variance  $0.5 \text{ md}$ . Consequently, the pressure vector  $p$  and velocity vector  $\vec{v}$  become stochastic processes. In order to apply PCE, we express  $K$  as a series of orthogonal basis functions  $\psi_i$  (Hermite polynomials) and expansion coefficients  $k_i$

$$K(x; \omega) = \sum_{i=0}^P k_i(x) \psi_i(\xi), \tag{34}$$

Accordingly,

$$p(x; \omega) = \sum_{i=0}^P p_i(x) \psi_i(\xi), \tag{35}$$

and

$$q(x; \omega) = \sum_{i=0}^P q_i(x) \psi_i(\xi), \tag{36}$$

substituting into Equation (7) and eliminating the arguments for brevity, we get

$$\nabla \cdot \left[ - \sum_{i=0}^P k_i \psi_i \nabla \left( \sum_{j=0}^P p_j \psi_j \right) \right] = \sum_{i=0}^P q_i \psi_i, \tag{37}$$

Which can be simplified to become

$$\sum_{i=0}^P \sum_{j=0}^P -[k_i \nabla^2 p_j + \nabla k_i \cdot \nabla p_j] \psi_i \psi_j = \sum_{i=0}^P q_i \psi_j. \tag{38}$$



Applying inner product  $\langle \cdot, \cdot \rangle$  with each polynomial  $\psi_k$ ,  $k = 0, \dots, P$ , and utilizing the orthogonality relation to get a set of  $P + 1$  deterministic equations of the form

$$\sum_{i=0}^P \sum_{j=0}^P -[k_i \nabla^2 p_j + \nabla k_i \cdot \nabla p_j] e_{ijk} = q_K \langle \psi_k^2 \rangle, \tag{39}$$

where  $e_{ijk} = \langle \psi_i \psi_j \psi_k \rangle$ , which can be computed and stored priori [29]. We can apply the TPFA numerical scheme for this set of deterministic equations and obtain the pressure components  $p_i(x)$ .

- Second: Using KL expansion

As the realizations of a Gaussian distribution may take negative values, it is more practical to assume a log normal distribution for  $K$ . For that purpose, first we define a Gaussian process  $Z(x; \omega)$  with the KL expansion,

$$Z(x; \omega) = \mathbb{E}[Z(x, \cdot)] + \sum_{i=1}^M \sqrt{\lambda_i} \phi_i \xi_i, \tag{40}$$

where  $\mathbb{E}[Z(x, \cdot)]$  is the mean, and  $\lambda_i$  and  $\phi_i(x)$  are the eigenvalues and eigenfunctions of the exponential correlation function  $C(x, y)$ , respectively. The correlation  $C(x, y)$  is of the form [30,31]

$$C(x, y) = \sigma_z^2 \exp\left(\frac{-\|x - y\|}{l}\right), x, y \in D, \tag{41}$$

where  $l$  denotes the spatial correlation length and  $\sigma_z^2$  is the process variance.

The permeability  $K$  and  $Z$  are related such that  $K$  is the exponential of  $Z$  and the PCE of  $K$  would be

$$K(x; \omega) = e^{Z(x; \omega)} = \sum_{i=0}^P k_i(x) \psi_i(\xi) \tag{42}$$

To obtain the coefficients  $k_i(x)$ , we use the formula

$$k_i(x) = \frac{\langle \psi_i(\eta) \rangle}{\langle \psi_i^2 \rangle} \exp\left[\mathbb{E}[Z(x, \cdot)] + \frac{\sigma_z^2}{2}\right], \eta_j = \xi_j - z_j, \tag{43}$$

with  $\psi_i(\eta)$  given in Table 1.

**Table 1.**  $\psi_i(\eta)$  used to evaluate the PC coefficients  $k_i(x)$  [11].

$\psi_i(\xi)$	$\psi_i(\eta)$	$\psi_i(\eta)$
$\xi_i$	$\eta_i + z_i$	$z_i$
$\xi_i \xi_j - \delta_{ij}$	$(\eta_i + z_i)(\eta_j + z_j) - \delta_{ij}$	$z_i z_j$
$\xi_i \xi_j \xi_k - \xi_i \delta_{jk} - \xi_j \delta_{ik} - \xi_k \delta_{ij}$	$(\eta_i + z_i)(\eta_j + z_j)(\eta_k + z_k) - z_i \delta_{jk} - z_j \delta_{ik} - z_k \delta_{ij}$	$z_i z_j z_k$

After computing the coefficients  $k_i(x)$ , we apply PCE to we get the following set of deterministic equations to be solved to get the pressure components using the TPFA scheme,

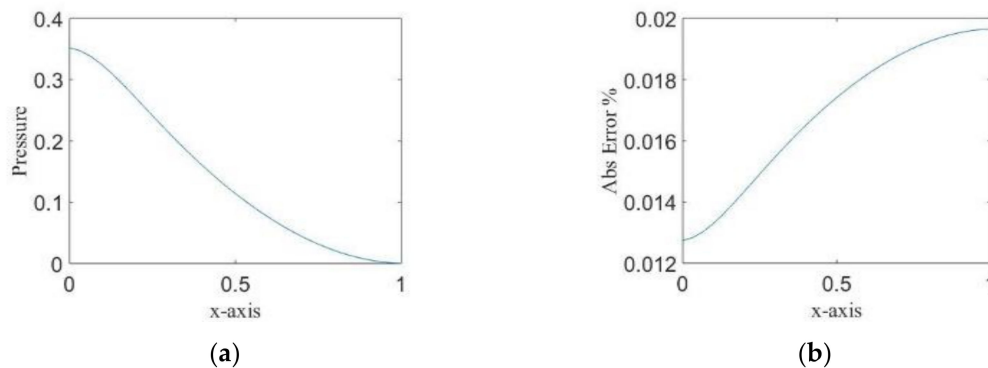
$$\sum_{i=0}^P \sum_{j=0}^P -[k_i \nabla^2 p_j + \nabla k_i \cdot \nabla p_j] e_{ijk} = q_K \psi_k^2, \tag{44}$$

where  $e_{ijk} = \langle \psi_i \psi_j \psi_k \rangle$ .

## 4. Results

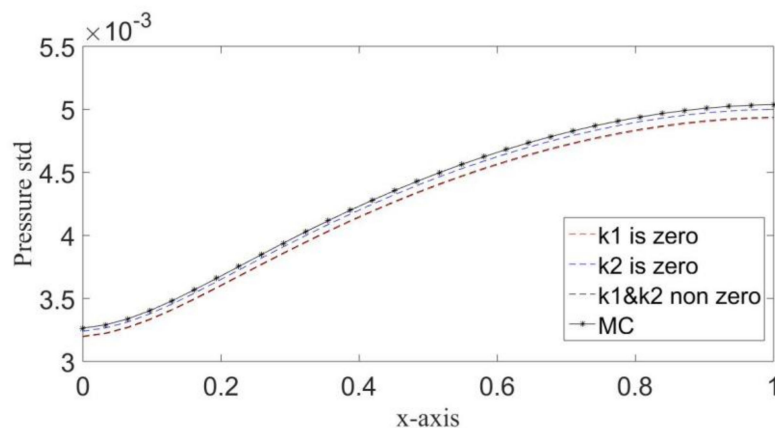
### 4.1. Case of PCE Only

We compare the differences in solutions obtained using MC and PCE methods. Figure 3 represents the normalized pressure mean solution produced by MC method (left) and the absolute error between the MC mean solution and the PCE mean solution (right). There was no considerable error between the two solutions (less than 0.02%) when assuming PCE with order  $d = 2$ , which proves the efficiency of the method for linear problems even when taking a low order expansion.



**Figure 3.** (a) The pressure mean solution using MC method. (b) The percentage of difference of solutions between MC and PCE methods.

Figure 4 shows variances in pressure solutions of MC and PCE methods when eliminating one of the PCE coefficients of the input random field  $K$ . It has been observed that the elimination has an insignificant impact on the accuracy of the solution. However, using more PCE coefficients tends to minimize the error between MC and PCE solutions. These results illustrate the fast convergence of the PCE solution for PDE (8), as assuming a higher-order PCE had a negligible effect on the accuracy of the solution.



**Figure 4.** Comparing the pressure standard deviation resulting from MC and PCE methods.

Table 2 illustrates the percentage of the difference between MC and PCE solutions when eliminating the first PCE coefficient, i.e.,  $k_1 = 0$ , then the second PCE coefficient  $k_2 = 0$ , then without elimination. The minimum percentage of error results when using all PCE coefficients, while using only the second PCE coefficient.  $k_1 = 0$  resulted in more accurate results than when  $k_2 = 0$ .

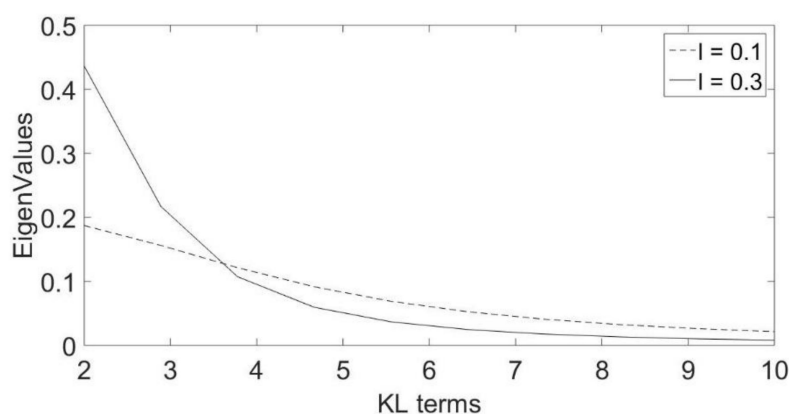
**Table 2.** PCE vs. MC solution statistics.

	$k_1=0$	$k_2=0$	$k_1&k_2$ non zero
Mean difference percentage	0.0035	0.0039	0.0031
Variance difference percentage	2.56	2.61	2.52

#### 4.2. Case of KL with PCE

In the second case, we apply the KL expansion to the input, in order to reduce the dimension of the input space and then use PCE. The permeability field  $K$  is of log-normal distribution in this case. The MC method used for validation is modified by utilizing the KL expansion on a Gaussian field  $Z$  and constructing  $K$  by taking the exponential of  $Z$  in every iteration.

First, we compare the convergence of the eigenvalues of the correlation function  $C(x, y)$  for different spatial correlation lengths. Figure 5 shows that as the correlation length decreases, the convergence would require more expansion terms. This can be further illustrated in Table 3, where the minimum number of KL terms required for convergence is provided for four different correlation lengths. We compared the solutions resulting from using PCE only and those resulting from using KL then PCE (KL-PCE) using different correlation lengths, where the results are compared against the MC solution. It has been noticed that there are no considerable differences in the mean solutions for all methods. However, the variance solution differs considerably based on the used method and the parameters  $\sigma^2$  and  $l$  of the correlation function  $C(x, y)$ .

**Figure 5.** Comparing Eigenvalues convergence rate for  $l = 0.1$  and  $l = 0.3$ .**Table 3.** PCE vs. MC solution statistics.

Correlation Length $l$	$l=0.1$	$l=0.3$	$l=0.5$	$l=1$
Number of terms for convergence	> 12	12	8	5

The problem is solved using PCE of order  $d = 1$ , and PCE of order  $d = 2$  after applying KL. The results show that PCE of order  $d = 1$  provides accurate results for low variance  $\sigma^2$  values (less than 0.5). Even for highly heterogeneous mediums where we can set

[32], PCE with first-order polynomials provides acceptable accuracy (less than 1% error when compared to MC) with no much need for higher-order PCE. However, for  $\sigma^2 > 1$ , using higher-order PCE is recommended for more accurate results. Therefore, we can apply KL in our problem before applying PCE as this leads to a faster convergence. It should be noted that when using KL the correlation length must be taken into consideration, as using higher correlation length would significantly reduce the accuracy of the solution if the minimum number of KL terms required for the eigenvalues convergence is not used. This can be observed in Table 4, where we compare the mean square error (MSE) percentages between MC and KL/PCE variance solutions.

**Table 4.** PCE vs MC variances MSE.

Method	MSE% at $\sigma=0.1$	MSE% at $\sigma=0.5$	MSE% at $\sigma=1$
PCE	0.02	0.4	1
KL-PCE $l = 0.1$	0.01	0.5	0.1
KL-PCE $l = 0.3$	0.02	0.3	0.9
KL-PCE $l = 0.5$	0.1	0.8	4

## 5. Conclusions

In this paper, we studied the quarter-five spot problem. We solved both the deterministic problem and the stochastic problem after imposing uncertainty to the problem. We considered two cases for the stochastic problem: The first case is solved using both MC and PCE methods and the results are compared. It was shown that for the given problem, assuming a PCE with order  $d = 2$  produces an error in the mean solution less than 0.02%, while reducing the computational time to at least 30% of the time needed by the MC method. In the second case, we compared the effect of applying the KL expansion with different correlation function parameters before applying the PCE. It has been shown that for systems with low heterogeneity (low values of process variance  $\sigma^2$ ), PCE with first-order polynomials provides accurate results (less than 1% error). Even for a highly heterogeneous medium with  $\sigma^2 = 1$ , PCE with first-order polynomials provides acceptable accuracy in most cases. However, if the value of  $\sigma^2$  increases, higher-order PCE may be required, and the KL expansion can be applied for better convergence of the PCE. When using the KL expansion, the correlation length of the correlation function should be taken into consideration before choosing the number of KL expansion terms. It was observed that a correlation function with a higher correlation length may require using the full number of KL expansion terms needed for the convergence of the eigenvalues of the correlation function. This study can be extended to predict the expected behavior of the general case when there is a system of a reservoir with injection/production wells with uncertain parameters and/or boundary conditions.

**Author Contributions:** H.A. was responsible for the computations, documentation and literature review. A.A.-J. was responsible for administration, supervising and reviewing the manuscript. M.E.-B. was responsible for the scientific content, novelty and analysis of the results. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Shenoy, A.V. Non-Newtonian Fluid Heat Transfer in Porous Media. In *Advances in Heat Transfer*; Hartnett, J.P., Irvine, T.F., Cho, Y.I., Eds.; Elsevier: Amsterdam, The Netherlands, 1994; Volume 24, pp. 101–190.
- El-Beltagy, M.A. A practical comparison between the spectral techniques in solving the SDEs. *Eng. Comput.* **2019**, *36*, 2369–2402.
- Le Maître, O.; Knio, O. *Spectral Methods for Uncertainty Quantification: With Applications to Computational Fluid Dynamics*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2010.
- Giles, M.B. Multilevel Monte Carlo Path Simulation. *Op. Res.* **2008**, *56*, 607–617. [[CrossRef](#)]
- Doucet, A.; de Freitas, N.; Gordon, N. An Introduction to Sequential Monte Carlo Methods. In *Sequential Monte Carlo Methods in Practice*; Doucet, A., de Freitas, N., Gordon, N., Eds.; Springer: New York, NY, USA, 2001; pp. 3–14.
- Caflich, R.E. Monte Carlo and quasi-Monte Carlo methods. *Acta Numer.* **1998**, *7*, 1–49. [[CrossRef](#)]
- Gelhar, L.W. Stochastic subsurface hydrology from theory to applications. *Water Resour. Res.* **1986**, *22*, 135S–145S. [[CrossRef](#)]
- Ketabchi, H.; Ataie-Ashtiani, B. Review: Coastal groundwater optimization—advances, challenges, and practical solutions. *Hydrogeol. J.* **2015**, *23*, 1129–1154. [[CrossRef](#)]
- Razavi, S.; Tolson, B.A.; Burn, D.H. Review of surrogate modeling in water resources. *Water Resour. Res.* **2012**, *48*, W07401. [[CrossRef](#)]

10. Frangos, M.; Marzouk, Y.; Willcox, K.; van Bloemen Waanders, B. Surrogate and Reduced-Order Modeling: A Comparison of Approaches for Large-Scale Statistical Inverse Problems. In *Computational Methods for Large-Scale Inverse Problems and Quantification of Uncertainty*; John Wiley & Sons: Hoboken, NJ, USA, 2010; pp. 123–149.
11. Ghanem, R. Ingredients for a general purpose stochastic finite elements implementation. *Comput. Methods Appl. Mech. Eng.* **1999**, *168*, 19–34. [[CrossRef](#)]
12. Xiu, D.; Em Karniadakis, G. Modeling uncertainty in steady state diffusion problems via generalized polynomial chaos. *Comput. Methods Appl. Mech. Eng.* **2002**, *191*, 4927–4948. [[CrossRef](#)]
13. Cameron, R.H.; Martin, W.T. The Orthogonal Development of Non-Linear Functionals in Series of Fourier-Hermite Functionals. *Ann. Math.* **1947**, *48*, 385–392. [[CrossRef](#)]
14. Xiu, D.; Karniadakis, G.E. The Wiener-Askey Polynomial Chaos for Stochastic Differential Equations. *SIAM J. Sci. Comput.* **2002**, *24*, 619–644. [[CrossRef](#)]
15. Askey, R.; Wilson, J. Some basic hypergeometric orthogonal polynomials that generalize Jacobi polynomials. *Mem. Am. Math. Soc.* **1985**, *54*, 1–63. [[CrossRef](#)]
16. Paulson, J.; Buehler, E.; Mesbah, A. Arbitrary Polynomial Chaos for Uncertainty Propagation of Correlated Random Variables in Dynamic Systems. *IFAC PapersOnLine* **2017**, *50*, 3548–3553. [[CrossRef](#)]
17. Oladyshkin, S.; Nowak, W. Data-driven uncertainty quantification using the arbitrary polynomial chaos expansion. *Reliab. Eng. Syst. Saf.* **2012**, *106*, 179–190. [[CrossRef](#)]
18. Loève, M. *Probability Theory I*; Springer: Berlin/Heidelberg, Germany, 1987.
19. Sapsis, T.P.; Lermusiaux, P.F.J. Dynamically orthogonal field equations for continuous stochastic dynamical systems. *Phys. D Nonlinear Phenom.* **2009**, *238*, 2347–2360. [[CrossRef](#)]
20. Cheng, M.; Hou, T.Y.; Zhang, Z. A dynamically bi-orthogonal method for time-dependent stochastic partial differential equations I: Derivation and algorithms. *J. Comput. Phys.* **2013**, *242*, 843–868. [[CrossRef](#)]
21. Chavent, G.; Jaffre, J. *Mathematical Models and Finite Elements in Reservoir Simulation*; Elsevier: Amsterdam, The Netherlands, 1986.
22. Hasle, G.; Lie, K.; Quak, E. *Geometric Modelling, Numerical Simulation, and Optimization: Applied Mathematics at SINTEF*; Springer: Heidelberg, Germany, 2007.
23. Zhang, D. *Stochastic Methods for Flow in Porous Media: Coping with Uncertainties*; Academic Press: San Diego, CA, USA, 2001.
24. El-Beltagy, M.A.; Wafa, M.I. Stochastic 2D incompressible Navier-Stokes solver using the vorticity-stream function formulation. *J. Appl. Math.* **2013**, *2013*, 14. [[CrossRef](#)]
25. Augustin, F.; Gilg, A.; Paffrath, M.; Rentrop, P.; Wever, U. Polynomial chaos for the approximation of uncertainties: Chances and limits. *Eur. J. Appl. Math.* **2008**, *19*, 149–190. [[CrossRef](#)]
26. Gentle, J.E. *Numerical Linear Algebra for Applications in Statistics*; Springer: Berlin/Heidelberg, Germany, 1998.
27. Lie, K.-A. *An Introduction to Reservoir Simulation Using MATLAB*; Cambridge University Press: Cambridge, UK, 2015.
28. Pollock, D.W. Semianalytical Computation of Path Lines for Finite-Difference Models. *Groundwater* **1988**, *26*, 743–750. [[CrossRef](#)]
29. Noor, A.; Barnawi, A.; Nour, R.; Assiri, A.; El-Beltagy, M. Analysis of the Stochastic Population Model with Random Parameters. *Entropy* **2020**, *22*, 562. [[CrossRef](#)]
30. Paleologos, E.K.; Avaniidou, T.; Mylopoulos, N. Stochastic analysis and prioritization of the influence of parameter uncertainty on the predicted pressure profile in heterogeneous, unsaturated soils. *J. Hazard. Mater.* **2006**, *136*, 137–143. [[CrossRef](#)]
31. Kourakos, G.; Harter, T. Parallel simulation of groundwater non-point source pollution using algebraic multigrid preconditioners. *Comput. Geosci.* **2014**, *18*, 851–867. [[CrossRef](#)]
32. Crevillén-García, D.; Leung, P.K.; Rodchanarowan, A.; Shah, A.A. Uncertainty Quantification for Flow and Transport in Highly Heterogeneous Porous Media Based on Simultaneous Stochastic Model Dimensionality Reduction. *Transp. Porous Media* **2019**, *126*, 79–95. [[CrossRef](#)] [[PubMed](#)]

