Computer methods in applied mechanics and engineering
Comput. Methods Appl. Mech. Engrg. 190 (2001) 6359–6372
www.elsevier.com/locate/cma

Solution of stochastic partial differential equations using Galerkin finite element techniques

Manas K. Deb [a,*], Ivo M. Babuška [b], J. Tinsley Oden [b]

[a] *TIBCO Software, 6400 Harrogate Drive, Austin, TX 78759, USA*
[b] *Texas Institute for Computational and Applied Mathematics, The University of Texas, Austin TX, USA*

Received 12 September 2000

Abstract

This paper presents a framework for the construction of Galerkin approximations of elliptic boundary-value problems with stochastic input data. A variational formulation is developed which allows, among others, numerical treatment by the finite element method; a theory of a posteriori error estimation and corresponding adaptive approaches based on practical experience can be utilized. The paper develops a foundation for treating stochastic partial differential equations (PDEs) which can be further developed in many directions. © 2001 Published by Elsevier Science B.V.

1. Introduction

To date, most computer simulations are based on *deterministic* mathematical models, where all input data are assumed to be *perfectly known*. In fact, this is never the case. All data contain a certain level of uncertainty: material properties, loading scenarios, boundary conditions, domain geometry, etc., have smaller or larger uncertainties which influence the quantities of interest. This resulting uncertainty in the solution, can, of course, be larger or smaller than that in the input data.

The uncertainties of the input data can have different characteristics which must be taken into account, and must be related to the aims of the analysis. They can, for example, show stochastic probabilistic character when the material properties are obtained experimentally. Unfortunately, probabilistic data are often difficult to obtain and are generated by methods which themselves have additional uncertainties. The input data can be based on a "worst scenario" approach when bounds for the quantity of interest or in a probabilistic framework are desired. If the uncertainties are small, perturbation theory is a very valuable tool for analyzing their effects. If they are larger, then perturbation theory is not applicable.

In this paper we address problems characterized by linear partial differential equations for which the input data are stochastic; for example, the coefficients or the right-hand side (RHS) of the partial differential equation (PDE) are the stochastic functions. The aim of the paper is to transform the stochastic PDE problem into a deterministic problem where finite element methods can be used for obtaining useful numerical approximations. It is possible to use the theory of finite elements to obtain several useful results, including a posteriori error estimates, adaptive approaches, superconvergence computations of functionals, etc. (see [1,4]). The formulation of the stochastic boundary-value problem developed here also provides a basis for interpreting and analyzing numerous approaches suggested in the FEM literature from a unified point of view.

* Corresponding author.

0045-7825/01/$ - see front matter © 2001 Published by Elsevier Science B.V.
PII: S 0 0 4 5 - 7 8 2 5 (0 1) 0 0 2 3 7 - 7

This paper will concentrate on the stochastic input data in problems which are not white noise, i.e., there are significant correlations. This is typical in many engineering applications. Differential equations with white noise are broadly studied in various contexts in physics, financial models, etc. For more on this subject see, e.g. [10]. One obvious way to treat stochastic PDEs is the Monte Carlo Method. This method is expensive especially when some higher-order accuracy for mean values, variation, etc. are sought. In addition, errors in the numerical approximation of the exact solution must be characterized in a probabilistic way. It is worthwhile to mention our approach when adaptive finite element methods are used, has a clear relation to an adapative Monte Carlo Method.

Various numerical methods to solve stochastic partial differential equations (SPDEs) have been proposed in the literature. The work of Ghanem [9] and Ghanem and Spanos [8] advocate a hybrid finite element-spectral approach, while the monograph of Kleiber and Hien [11] utilizes a perturbation approach. Elishakoff and Ren [7] examine engineering finite element methods for structures with large stochastic variations and point to limitations of some approaches. A fuller description of work on computational methods for SPDEs used to model stochastic behavior in problems of mechanics can be found in the survey of Schueller and Pradwarter [14], in the books of Kleiber and Hien [11], Ghanem and Spanos [8], and in Deb [5,6] and the references therein. This paper is closely related to the Ph.D. thesis [6].

2. Model problems

We begin by considering two model problems of SPDEs of elliptic type.

Problem 1.

$$\begin{aligned} &\nabla \cdot \widetilde{a}(x)\nabla\widetilde{u}(x) = f(x) \quad && \text{on } D, \\ &\widetilde{u}(x) = 0 && \text{on } \partial D. \end{aligned} \tag{2.1}$$

Here $\widetilde{a}$ and $\widetilde{u}$ are stochastic functions while f is deterministic.

Problem 2.

$$\begin{aligned} &\nabla \cdot a(x)\nabla\widetilde{u}(x) = \widetilde{f}(x) \quad && \text{on } D, \\ &\widetilde{u}(x) = 0 && \text{on } \partial D. \end{aligned} \tag{2.2}$$

Here $\widetilde{f}$ and $\widetilde{u}$ are stochastic functions while a is deterministic. We assume that $D \in \mathbb{R}^d, d = 1,2,3$, is a bounded Lipschitz domain. We formulated these two problems separately because they have different structures. Of course, it is possible to analyze the problem when both $\widetilde{a}$ and $\widetilde{f}$ are stochastic, but considering each case separately simplifies slightly the notation etc., and, in addition, allows us to exploit special properties of Problem 2. We also can treat analogously, the stochastic problem for other differential equations; for example, those of linear elasticity, nonhomogeneous boundary conditions, stochastic boundaries ∂D etc. We refer to Problem 1 as a problem of the left-hand side (LHS) type of SPDE and analogously Problem 2 will be referred to as the RHS type of SPDE.

3. Mathematical formulation

Let (Ω, F, P) be a probability space, where Ω, F, P are the set of random events, the σ-algebra of subsets of Ω and P the probability measure, respectively. If $\widetilde{X}$ is a real random variable in (Ω, F, P) with $\widetilde{X} \in L^1(\Omega)$, we denote its expected value by

$$E[\widetilde{X}] = \int_\Omega X(\omega)\,\mathrm{d}P(\omega) = \int_{\mathbb{R}} x\,\mathrm{d}\mu(x).$$

Here μ is the distribution probability measure for $\widetilde{X}$, defined on the Borel set B and $\mathbb{R}$ given by

$$\mu(B) = P(X^{-1}(B)).$$

We will assume that $\mu(B)$ is absolutely continuous with respect to Lebesque measure; then there exists a density function for $\widetilde{X}, \rho : \mathbb{R} \to \mathbb{R}^+$ such that

$$E[\widetilde{X}] = \int_{\mathbb{R}} \rho(x)\,\mathrm{d}x.$$

Let us now define a random function. A function $\widetilde{\psi}(x) = \psi(x,\omega) : D \times \Omega \to \mathbb{R}$ will be called a random function when it is jointly measurable (on the Borel sets $(B(D) \oplus (\Omega))$) and

$$E\left[\int_D \psi^2(x,\omega)\,\mathrm{d}x\right] < \infty.$$

Remark 3.1. The last condition is not essential, but in this paper we will consider only such functions.

We assume that in (2.1) $\widetilde{a}(x) = a(x,\omega)$ is such that

$$0 < \alpha_1 \leqslant a(x,\omega) < \alpha_2 < \infty \quad \text{a.e. on } D \times \Omega. \tag{3.1}$$

Let $v(x,\omega)$ be defined on $D \times \Omega$. Then we define

$$V = \left\{ |v(x,\omega)| \|v\|_V < \infty, \quad v(x,\omega) = 0 \quad \text{on } \partial D \right\}, \tag{3.2}$$

where $\|\cdot\|_V$ is the energy norm

$$\|v\|_V^2 = E\left[\int_D a(x,\omega)\,|\nabla_x v|^2\,\mathrm{d}x\right] = \int_D (E[a(x,\omega)\,|\nabla_x v|^2])\,\mathrm{d}x. \tag{3.3}$$

The natural inner product in V is the bilinear form $\mathscr{B} : V \times V \to R$

$$\mathscr{B}(u,v) = E\left[\int_D a\nabla_x u \cdot \nabla_x v\,\mathrm{d}x\right] = \int_D (E[a\nabla_x u \cdot \nabla_x v])\,\mathrm{d}x. \tag{3.4}$$

Obviously, V is a Hilbert space of random functions.

Theorem 3.2. *Let $f(x) \in L^2(D)$. There exists unique weak solution $u_0(x,\omega) \in V$ (i.e., a random function) of the problem* (2.1) *which satisfies*

$$\mathscr{B}(u_0,v) = \int_D (E[fv])\,\mathrm{d}x = \mathscr{L}(v) \quad \forall v \in V. \tag{3.5}$$

Proof. The theorem follows immediately from the Lax–Milgram lemma. □

Remark 3.3. We address here the LHS problem (2.1) when f is deterministic. The RHS problem is addressed in Section 7.

Let us assume now that

$$a(x,\omega) = (E(a))(x) + \sum_{i=1}^{M} \sqrt{\lambda_i} a_i(x) A_i(\omega), \tag{3.6}$$

where $A_i(\omega)$, $i = 1,2,\ldots,M$ are real mutually independent random variables with mean value zero $(E(\widetilde{A}_i) = 0)$, variance one $(E(\widetilde{A}_i^2) = 1)$, and bounded images Γ_i of Ω, $\Gamma_i = A_i(\Omega) \in \mathbb{R}^1$, $i = 1,\ldots,M$.

Further, we assume that each $\widetilde{A}_i$ has a probability density function $\rho_i : \Gamma_i \to R^+$, $i = 1,\ldots,M$, and $0 < \beta_1 \leqslant \rho_i \leqslant \beta_2 < \infty$ and $a_i(x) \in L^\infty(D)$, $\lambda_i > 0$, $i = 1,\ldots,M$, and (3.1) holds.

Remark 3.4. In (3.6), we can have $M = \infty$ provided that the series converges in $L^\infty(D)$. Then, using (3.6) as the truncated series, we can estimate the error caused by the truncation, but we will not address this estimation here.

The expansion (3.6) is known as the Karhunen–Loeve (K–L) expansion [12]. Then λ_i and $a_i(x)$ are the eigenvalues and eigenfunctions associated with the given covariance function $C(x,\bar{x})$ of $\widetilde{a}(x) = a(x,\omega)$ and $a_i(x)$ have orthonormalities properties which can be utilized in the implementation. The probability densities ρ_i are arbitrary. For any of them, we get $E[a(x,\omega)a(\bar{x},\omega)] = C(x,\bar{x})$.

Using (3.6), we can write

$$a(x,\omega) = a(x, A_1(\omega), \ldots, A_M(\omega)). \tag{3.7}$$

Let $\Gamma = \Gamma_1 \times \Gamma_2 \times \cdots \times \Gamma_M \subset \mathbb{R}^M$,

$$\rho(y) = \rho_1(y_1)\rho_2(y_2)\cdots\rho_M(y_M), \quad y_i \in \Gamma_i, \ y = (y_1,\ldots,y_M) \in \Gamma, \ 0 < \beta_1 \leqslant \rho(y_i) \leqslant \beta_2 < \infty.$$

Then $a(x,\omega) = a(x,y)$ with $y_i = A_i(\omega)$.

Using the Doob–Dynkn lemma (see [13], p. 9), the solution $u_0(x,\omega)$ of (3.6) has the form

$$u_0(x,\omega) = u_0(x, A_1(\omega), \ldots, A_M(\omega)) = u_0(x, y_1, \ldots, y_M). \tag{3.8}$$

Let $v(x,y), x \in D, y \in \Gamma$ and

$$W(D,\Gamma) = \left\{ v(x,y) \,\middle|\, \int_\Gamma \rho(y) \int_D a(x,y)\,|\nabla_x v(x,y)|^2 \,\mathrm{d}x\,\mathrm{d}y = \|v\|_W < \infty, v(x,y) = 0 \text{ on } \partial D \text{ for all } y \right\}. \tag{3.9}$$

Then the space, W is the same as (equivalent to) V when $a(x,\omega)$ is given in (3.6). Because of our assumptions on a and ρ, W is Hilbert space. Further, (3.4) becomes

$$\mathscr{B}(u,v) = \int_\Gamma \rho(y) \int_D a(x,y) \nabla_x u(x,y) \cdot \nabla_x v(x,y)\,\mathrm{d}x\,\mathrm{d}y \tag{3.10}$$

which will be used in the sequel. Further, we have

$$\mathscr{L}(v) = \int_\Gamma \rho(y) \int_D v(x,y) f(x)\,\mathrm{d}x\,\mathrm{d}y, \quad v \in W. \tag{3.11}$$

Then $u_0(x,y) \in W$ defined in (3.5) can be written in the form (3.8) and satisfies

$$\mathscr{B}(u_0, v) = \mathscr{L}(v) \quad \forall v \in W. \tag{3.12}$$

Because of our assumptions, $u_0(x,y) \in W$ is uniquely determined from (3.12); we transformed the stochastic problem (2.1) (resp. (3.5)) into the deterministic problem (3.12).

Because $u_0(x,\omega) \in V$ (resp. $u_0(x,y) \in W$), we have

$$\begin{aligned}
E[u_0(x,\omega)] &= \int_\Gamma \rho(y) u_0(x,y)\,\mathrm{d}y \in H^1(D),\\
E[u_0^2(x,\omega)] &= \int_\Gamma \rho(y) u_0^2(x,y)\,\mathrm{d}y \in L^1(D),\\
E\left[\frac{\partial u_0}{\partial x_i}(x,\omega)\right] &= \int_\Gamma \rho(y) \frac{\partial u_0}{\partial x_i}(x,y)\,\mathrm{d}y \in L^2(D),\\
E\left[\left(\frac{\partial u_0}{\partial x_i}(x,\omega)\right)^2\right] &= \int_\Gamma \rho(y) \left(\frac{\partial u_0}{\partial x_i}(x,y)\right)^2 \mathrm{d}y \in L^1(D),\\
C(x,\bar{x}) &= E[u_0(x,\omega)u_0(\bar{x},\omega)] \in L^2(D \times D).
\end{aligned} \tag{3.13}$$

Equations (3.13) shows that the mean value, variance, standard deviation and covariance are well-defined functions. Their smoothness depends on $a_i(x)$ in (3.6), $f(x)$, and smoothness of ∂D.

4. The finite element solution of the LHS type of SPDE

We have seen in Section 3 that the LHS type of SPDE can be cast in the form (3.10). This form is very suitable for finite element approximation, and allows us to use the theories of FEM, the a posteriori error estimation, adaptivity, etc.

Let $\Delta(D)$ be a mesh on D which satisfies all of the usual assumptions, common in the finite element theory and $\tau(\Delta(D))$ be the elements with $h(D) = \max(\operatorname{diam} \tau(\Delta(\Gamma))$. Then we denote by $\Delta(\Gamma) = \Delta(\Gamma_1) \times \cdots \times \Delta(\Gamma_M)$ the rectangular mesh on Γ and let $\tau(\Delta(\Gamma)) = \tau(\Delta(\Gamma_1)) \times \cdots \times \tau(\Delta(\Gamma_M))$ the associated elements. We denote by $\Delta(D, \Gamma)$ the mesh on $D \times \Gamma$, and $\tau(\Delta(D, \Gamma)) = \tau(\Delta(D)) \times \tau(\Delta(\Gamma))$ the elements on $D \times \Gamma$. Let

$$S^{p,q}(D,\Gamma) = \Big\{ v(x,y) \in W | v(x,y)|_{\tau(D,\Gamma)} \text{ is polynomial of degree } p \text{ in } x, \ \forall y \in \Gamma \text{ and of degree } q \text{ in } y_1, y_2, \ldots y_m, \forall x \in D \Big\}. \tag{4.1}$$

If $\tau(D)$ are quadrilaterals, then $v(x,y)$ are polynomials of degree p in every variable x_i, $i = 1, \ldots, d$. If the mesh on D is curvilinear, then we use the standard pull-back polynomials of degree p.

Remark 4.1. The finite element approximation then reads: Find $u_{S^{p,q}(D,\Gamma)} \in S^{p,q}(D,\Gamma)$ such that

$$\mathscr{B}\big(u_{S^{p,q}(D,\Gamma)}, v\big) = \mathscr{L}(v) \quad \forall v \in S^{p,q}(D,\Gamma) \tag{4.2}$$

where $\mathscr{B}$ and $\mathscr{L}$ are defined in (3.10) and (3.11). From standard finite element theory, we get the basic result:

Theorem 4.2. *The finite element solution $u_{S^{p,q}(D,\Gamma)} \in S^{p,q}(D,\Gamma)$ exists, is unique, and*

$$\|u_0 - u_{S^{p,q}(D,\Gamma)}\|_{W(D,\Gamma)} \leqslant \inf_{\chi \in W} \|u_0 - \chi\|_{W(D,\Gamma)}. \tag{4.3}$$

Hence

$$\|u_0 - u_{S^{p,q}(D,\Gamma)}\|_{W(D,\Gamma)} \to 0 \quad \text{as } h(\Delta(D)) \text{ and } h(\Delta(\Gamma)) \to 0 \tag{4.4}$$

and

$$\|u_0 - u_{S^{p,q}(D,\Gamma)}\|_{W(D,\Gamma)} \to 0 \quad \text{as } p, q \to \infty. \tag{4.5}$$

Remark 4.3. Condition (4.4) expresses convergence of the h version while (4.5) that of the p version of the finite element method.

Utilizing the smoothness of the data and of the solution, we can, in standard ways, prove a priori estimates for the error in the energy norm, L^2, norm and the error in the data at interest for the mean value, standard derivation, covariance etc. When data are smooth, the following result follows from standard FEM procedures.

Theorem 4.4. *Let the input data $a_i(x), f(x)$ and $\partial\Omega$ be sufficiently smooth. Then we have*

$$\big\|u_0 - u_{S^{p,q}(D,\Gamma)}\big\|_W \leqslant C_1(p,q)\Big((h(D))^p + (h(\Gamma))^{q+1}\Big), \tag{4.6}$$

$$\big\|u_0 - u_{S^{p,q}(D,\Gamma)}\big\|_{L^2(D,\Gamma)} \leqslant C_2(p,q)(h(D)^{p+1}) + (h(\Gamma)^{q+1}), \tag{4.7}$$

$$\big\|E[u_0] - E\big[u_{S^{p,q}(D,\Gamma)}\big]\big\|_{L^2(D)} \leqslant C_3(p,q)\Big(h(D)^{p+1} + (h(\Gamma))^{q+1}\Big). \tag{4.8}$$

5. A simple model problem: algebraic stochastic equations

We next consider a very simple problem which has many of the same essential properties as the LHS type of SPDE. To show the major ideas, we consider the stochastic problem

$$\widetilde{a}\widetilde{u} = 1, \tag{5.1}$$

where $\widetilde{a} = a(\omega)$ is a stochastic function (in fact, a random variable) with

$$0 < a_{\min} \leqslant a(\omega) \leqslant a_{\max} < \infty \tag{5.2}$$

and $\widetilde{u} = u(\omega)$ is the stochastic solution of (5.1).

We assume that

$$a(\omega) = a_0 + \sum_{i=1}^{M} a_i A_i(\omega), \tag{5.3}$$

where $A_i(\omega),\ i = 1, \ldots, M$ are independent random variables with the same assumptions as in Section 2, $a_i \in \mathbb{R}$. Conditions (5.2) are the expressions (3.6) when $a(x, \omega)$ is x independent. Let $\Gamma_i = A_i(\Omega) = (-\alpha_i, \alpha_i) = I_i, \alpha_i > 0,\ A_i(\omega) = y_i \in I_i,\ i = 1, \ldots, M,\ \Gamma = \Gamma_i \times \cdots \times \Gamma_M,\ \rho_i(y_i)$ be the probability density of A_i and $\rho(y) = \rho_1(y_1)\rho_2(y_2)\cdots\rho(y_M),\ y = (y_1, \ldots, y_M)$.

We introduce the space W. Let $v(y), y \in \Gamma = I_i \times I_2 \times \cdots \times I_M$. Then

$$W = \left\{ v(y) \middle| \int_\Gamma \rho(y)a(y)v^2(y)\,\mathrm{d}y = \|v\|_W^2 < \infty \right\}, \tag{5.4}$$

where $a(y) = a_0 + \sum_{i=1}^{M} a_i y_i, a_j \in \mathbb{R}, j = 0, \ldots, M$. Further, for $u \in W,\ v \in W$ we define the bilinear form

$$\mathscr{B}(u, v) = \int_\Gamma \rho(y)a(y)u(y)v(y)\,\mathrm{d}y \tag{5.5a}$$

and the functional

$$\mathscr{L}(v) = \int_\Gamma \rho(y)v(y)\,\mathrm{d}y. \tag{5.5b}$$

The problem (5.1) can be formulated as follows: Find $u_0(y) \in W$ such that

$$\mathscr{B}(u_0, v) = \mathscr{L}(v), \quad \forall v \in W. \tag{5.6}$$

Let on I_i, the mesh $\Delta(I_i) : -\alpha_i = y_i^0 < y_i^1 < \cdots < y_i^{n(i)} = \alpha_i,\ h_i^j = y_i^j - y_i^{j-1},\ h_i = \max_j\ h_i^j = 1, \ldots, n(i)$. Then by $\Delta(\Gamma)$ we denote the rectangular mesh $\Delta(\Gamma) = \Delta(\Gamma_1) \times \Delta(\Gamma_2) \times \Delta(\Gamma_M)$ and $\tau(\Delta(\Gamma)) = \tau(\Delta(\Gamma_i)) \times \tau(\Delta(\Gamma_2)) \times \cdots \times \tau(\Delta(\Gamma_M))$ be the elements on Γ. Then we let

$$S^q = \left\{ v(y) \in W | v|_{\tau(\Delta(\Gamma))} \text{ is polynomial of degree } q \text{ separately in } y_1, \ldots, y_M \right\}, \tag{5.7}$$

and the finite element $S^q(\Gamma) \in S^q$ satisfies

$$\mathscr{B}(u_{S^q}, v) = \mathscr{L}(v) \quad \forall v \in S^q. \tag{5.8}$$

Obviously, the exact solution of (5.1) is

$$u_0(\omega) = u_0(y) = \frac{1}{a(y)}. \tag{5.9}$$

Hence, because of (5.2) and (5.3), it is easy to see that $u(y)$ is smooth on Γ and hence we have

$$\|u_0 - u_{S^q}\|_W \leqslant C \left(\sum_{i=1}^{M} h_i^{2(q+1)} \right)^{1/2}. \tag{5.10}$$

Further, as an example, consider

$$E[u_0] - E[u_{S^q}] = E[u_0 - u_{S^q}] \tag{5.11}$$

which is the error in the data of interest, namely $E[u_0]$. Proceeding analogously as in the classical finite element method, we define the influence function $G \in W$ by

$$\mathscr{B}(u, G) = \int_\Gamma u\rho \, \mathrm{d}y \quad \forall u \in W. \tag{5.12}$$

Because of the symmetry of $\mathscr{B}(u, v)$ and $G(y) = 1/a(y), G(y)$ is smooth.

$$E[u_0] = \mathscr{B}(u_0, G), \tag{5.13a}$$

$$E[u_{S^q}] = \mathscr{B}(u_{S^q}, G), \tag{5.13b}$$

and

$$|E[u_0] - E[u_{S^q}]| = |\mathscr{B}(u_0 - u_{S^q}, G)| = |\mathscr{B}(u_0 - u_{S^q}, G - G_{S^q})|$$
$$\leqslant \|u_0 - u_{S^q}\|_W \|G - G_{S^q}\|_W \leqslant C\left(\sum_{i=1}^{M} h_i^{2q+1}\right)^{2(q+1)}. \tag{5.14}$$

Next consider the error of the variance,

$$E[u_0^2] - E[u_{S^q}^2] = E[u_0^2 - u_{S^q}^2]. \tag{5.15}$$

Then we have

$$\begin{aligned}
\left|E\left[u_0^2 - u_{S^q}^2\right]\right| &= |E[(u_0 - u_{S^q})(u_0 + u_{S^q})]| \\
&= |E[(u_0 - u_{S^q})(2u_0 - (u_0 - u_{S^q}))]| \\
&= |2E[(u_0 - u_{S^q})u_0] - E[(u_0 - u_{S^q})(u_0 - u_{S^q})]| \\
&\leqslant |2\mathscr{B}(u_0 - u_{S^q}, G_0) + E[(u_0 - u_{S^q})(u_0 - u_{S^q})]| \\
&\leqslant |2\mathscr{B}(u_0 - u_{S^q}, G_0 - G_{0,S^q})| + E[(u - u_{S^q}), (u_0 - u_{S^q})] \\
&\leqslant C\left[\sum_{i=1}^{M} h_i^{2(q+1)}\right].
\end{aligned} \tag{5.16}$$

Next, considering implementational aspects, we assume that $q = 0$, i.e., we use constant shape functions on the mesh $\Delta(\Gamma)$, and we have

$$u_{S^0} = \sum_{\substack{i_j=1 \\ j=1,\ldots,M}}^{n(i)} C_{i_1,i_2,\ldots,i_M} \psi_{i_1}(y_1)\psi_{i_2}(y_2)\cdots\psi_{i_M}(y_M) = \sum_{\substack{i_j=1 \\ j=1,\ldots,M}}^{n(i)} C_{i_1,\ldots,i_M} \psi_{i_1,\ldots,i_M}(y), \tag{5.17}$$

where

$$\begin{cases} \psi_{i_j}(y_j) = \frac{1}{h^{j_i}} & \text{for } y^{i_j - 1} < y_j < y_j^{i_j}, \\ 0 & \text{elsewhere}, \end{cases} \tag{5.18}$$

and

$$\psi_{i_i \ldots, i_M}(y) = \psi_{i_1}(y_1) \cdots \psi_{i_M}(y_M).$$

Obviously,

$$\mathscr{B}(\psi_{k_1,\ldots,k_M}, \psi_{\ell_1,\ldots,\ell_M}) = 0$$

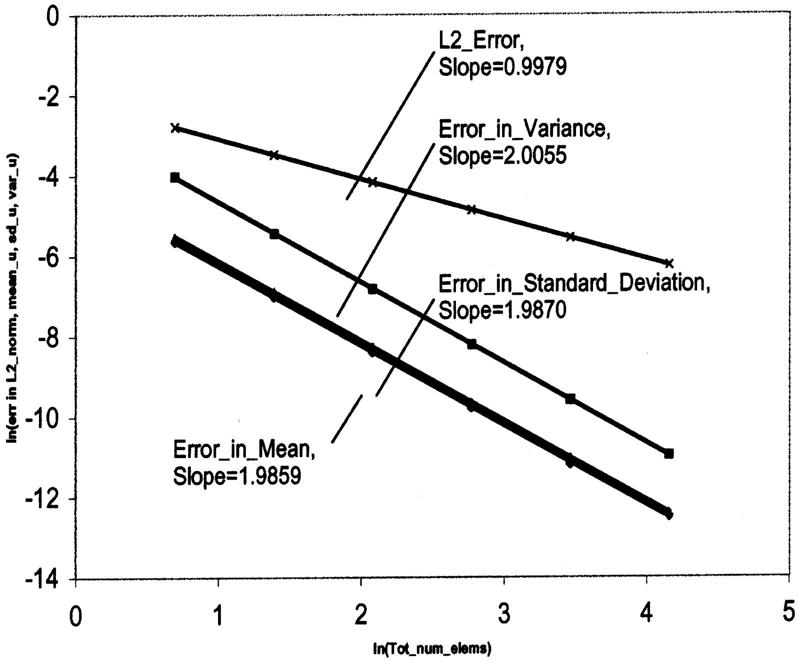

Fig. 1. Convergence rate for $M = 1, a_0 = a_1 = 1, \alpha_{\max}/\alpha_{\min} = 1.5$. The observed rate is very close to the theoretical one also for crude mesh.

for $(k, \ldots, k_M) \neq (\ell, \ldots \ell_M)$ and the system to solve is diagonal. As usual, we call the number of basis functions in (5.15) the number of degrees of freedom and denote it by N. We can of course use $q > 0$ and orthogonalize the shape functions on $\mathscr{T}(\Delta(\Gamma))$; then we solve systems of equations with a diagonal matrix.

Let us consider a numerical example. Assume $M = 1, a_0 = a_1 = 1$ and that α_1 is selected so that $a_{\max}/a_{\min} = 1.5$ and the density ρ_1 is uniform. Further, consider a uniform mesh $\Delta(\Gamma_1)$. Because the exact solution is known, i.e., its probability distribution; we now can compute the error

$$\|u_0 - u_{S^q}\|_W = \|u_0 - u_{S^q}\|_{L^2(I)}$$

as well the errors in the mean standard derivation and variation.

In Fig. 1, we show in the $\ell n - \ell n$ scale the errors and their observed rate obtained by a least-square fit of the data. The theoretical rate for the L_2 norm is 1 and the other data is 2. We see that high accuracy was achieved with small numbers of degrees of freedom M and the rates match perfectly the theoretical estimates.

In Fig. 2, we show analogous results for $M = 2, a_0 = a_1 = a_2,\ \rho_i$ and $\alpha_1 = \alpha_2$ uniform and $a_{\max}/a_{\min} = 199$. We observe that also for crude mesh, convergence is in the asymptotic range.

6. The general case of random LHS of SPD, an example

Let us consider the two-dimensional problem (2.1), respectively (3.10) and (3.11), and its approximate solution based on (4.1).

Let $D = (-0.5, 0.5) \times (-0.5, 0.5), \sigma_a = 0.1$

$$a(x_1, x_2, \omega) = a_0(x_1, x_2) + \sigma_a \sum_{i=1}^{3} \sqrt{\lambda_i} a_i(x_1, x_2) A_i(\omega)$$

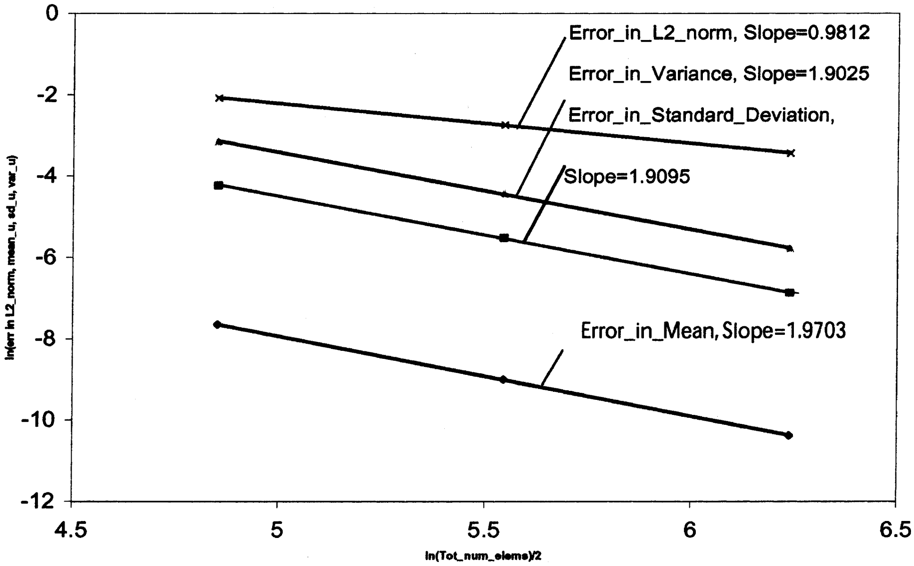

Fig. 2. Convergence for $M = 2, a_0 = a_1 = a_2 = 1, \alpha_{\max}/\alpha_{\min} = 199$. The observed rate is very close to the theoretical one.

with $\Gamma = (-\sqrt{3}, \sqrt{3})^3, \rho(y) =$ constant, and λ_i and $a_i(x)$ are obtained by taking products of K–L expansion for single dimension with (one-dimensional) covariance function $C = \sigma^2 e^{-|-x_1 - x_2|}$, $\sigma = 0.1$, and $f(x_1, x_2) = 2(0.5 - x_1^2 - x_2^2)$. Because the problem is symmetric, we need address only a quarter of the domain D. We consider on D uniform mesh of squares; particularly we use an 8×8 mesh on the quarter of the domain D. Also on Γ a uniform square mesh is considered; particularly we use $8 \times 8 \times 8$ mesh on Γ. On D we use bilinear elements and on Γ the constant elements. Denoted by $u_0(x, y)$ and $u(x, y; h(D), h(\Gamma))$ the exact and the finite element solution. We have

$$\left(\int_\Gamma \rho \int_D a(x) |\nabla_x (u_0(x, y) - u(x, y; h(D), h(\Gamma)))|^2 \right)^{1/2} \leqslant C[h(D) + h(\Gamma)]. \tag{6.1a}$$

Further,

$$\begin{aligned} &\|E[u_0] - E[u(h(\Delta), h(\Gamma)]\|_{L^2} \\ &\quad \leqslant C\big(h^2(D) + h(\Gamma)\big) \Big\| \Big(E[u_0^2] - (E[u])^2\Big) - \Big(E(u(h(D), h(\Gamma)) - (E(u(h(\Delta), h(\Gamma)))^2\Big) \Big\|_{L_1(D)} \\ &\quad \leqslant C\big(h^2(D) + h(\Gamma)\big). \end{aligned} \tag{6.1b}$$

In Figs. 3(a) and (b) we show the mean and the variance computed numerically. In Figs. 4(a) and (b) we show the mean and variance using 65,536 realizations using a uniform mesh in D.

7. Solution of random RSH type SPDE

In the previous sections, we addressed the problem of random LHS type in a SPDE. The analysis of RHS type is very similar. Nevertheless, in this case we can, in addition, directly compute the mean value and the covariance. As before we will assume that

$$f(x, \omega) = f_0(x) + \sum_{i=1}^{M} \sqrt{\lambda_i} f_i(x) A_i(\omega) \tag{7.1}$$

and assume about the random variables $A_1(\omega)$, the same as before. Then we can set

$$f(x, \omega) = f(x, y), \quad x \in D \text{ and } y \in \Gamma \tag{7.2}$$

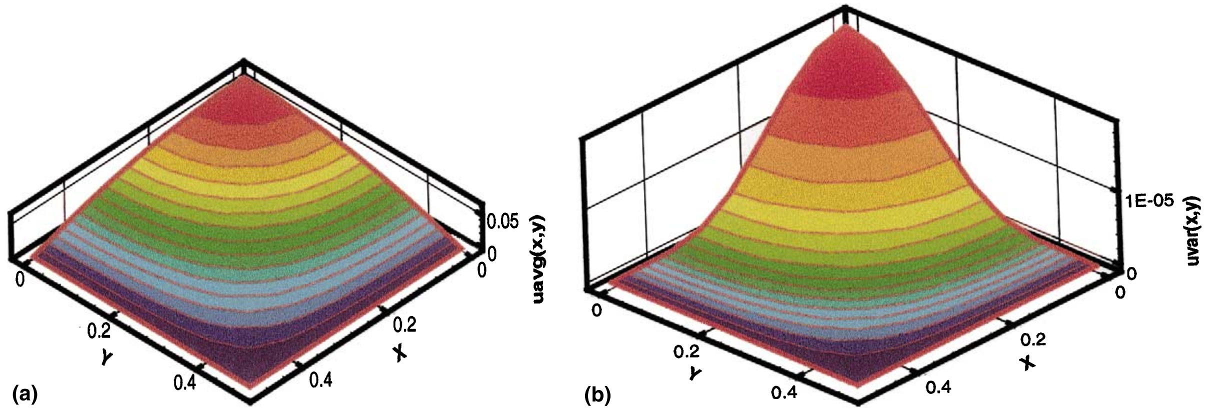

Fig. 3. Computation based on (4.1) and (4.2). (a) The mean value of the solution, max value is 0.06323. (b) the variance of the solution, is max value $0.1876E-04$.

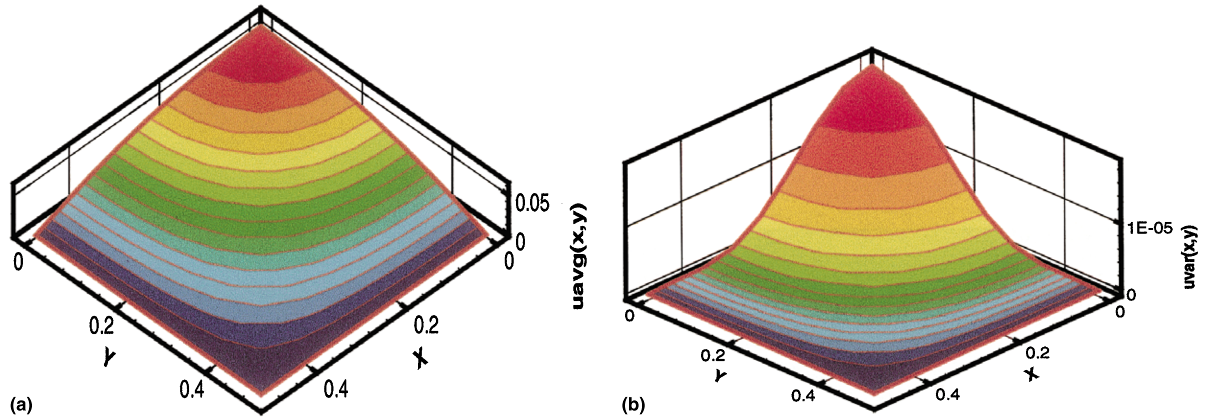

Fig. 4. (a) Computation based on Monte Carlo method: the mean value of the solution, max value is 0.06324. (b) The variance of the solution is max value is $0.1881E-0.4$.

and the solution has the form

$$u_0(x,\omega) = u_0(x,y). \tag{7.3}$$

We define

$$W(D,\Gamma) = \left\{ v(x,y) \middle| \int_\Gamma \rho(y) \int_D a(x) |\nabla_x v(x,y)|^2 \,\mathrm{d}x\,\mathrm{d}y < \infty \right\}. \tag{7.4}$$

The major difference between (3.9) and (7.4) is that $a(x)$ in (7.4) is independent of y. The exact solution $u_0 \in W$ satisfies $\forall v \in W$,

$$\mathscr{B}(u_0, v) = \mathscr{L}(v), \tag{7.5}$$

where

$$\mathscr{B}(u_0, v) = \int_\Gamma \rho(y) \int_D a(x)(\nabla_x u_0 \cdot \nabla_x v) \,\mathrm{d}x\,\mathrm{d}y, \tag{7.6a}$$

$$\mathscr{L}(v) = \int_\Gamma \rho(y) \int_D f(x,y) v(x,y) \,\mathrm{d}x\,\mathrm{d}y. \tag{7.6b}$$

We can use the finite elements as before and taking advantage of the fact that $a(x)$ is now y independent.

We further utilize this feature. First, we prove that the mean value $E[u_0]$ satisfies the deterministic equation with $E[f]$ at the RHS of f.

Theorem 7.1. *Let*

$$U = \left\{ v(x) \middle| v(x) = 0 \text{ on } \partial\Omega \text{ and } \|v\|_U = \left(\int_D a(x)|\nabla v|^2 \,\mathrm{d}x \right)^{1/2} < \infty \right\}$$

and

$$\mathscr{B}_0(u, v) = \int_D a(x)\nabla u \cdot \nabla v \,\mathrm{d}x$$

be the bilinear form defined on $U \times U$. Then

$$E[u_0] = \int_\Gamma \rho u_0(x, y)\,\mathrm{d}y \in U, \tag{7.7}$$

$$\mathscr{B}_0(E[u_0], v(x)) = \int_D E[f]v(x)\,\mathrm{d}x = \int_D f_0(x)v(x)\,\mathrm{d}x. \tag{7.8}$$

Proof. Equality (7.7) follows immediately from the fact that $u_0(x, y) \in W$.
Further, $u_0(x, y)$ satisfies

$$\int_\Gamma \rho(y) \int_D a(x)\nabla_x u_0(x, y) \cdot \nabla_x v(x, y)\,\mathrm{d}x\,\mathrm{d}y = \int_\Gamma \rho(y) \int_D f(x, y)v(x, y)\,\mathrm{d}y \quad \forall v \in W.$$

Selecting $v(x, y) = v(x)$ and changing the order of integration, we have

$$\int_D a(x) \int_\Gamma \rho(y)\nabla_x u_0(x, y) \cdot \nabla_x v(x)\,\mathrm{d}x\,\mathrm{d}y = \int_D a(x)\nabla_x(E[u_0]) \cdot \nabla_x v(x)\,\mathrm{d}x = \int_D E[f]v(x)\,\mathrm{d}x = \int_D f_0 v(x)\,\mathrm{d}x$$

and (7.8) is proven. □

Let us now analyze the covariance. Because of Theorem 7.1, we can assume that $E[f] = f_0(x) = 0$.

Theorem 7.2. *Let $v(x, \bar{x})$ be defined on $D \times D$ and*

$$\mathscr{U} = \Big\{ v(x, \bar{x}) | v(x, \bar{x}) = 0 \quad \text{for } x \in \partial D, \bar{x} \in D \quad \text{and for } x \in D, \bar{x} \in \partial D, \nabla_x(\nabla_{\bar{x}} v) \in L_2(D \times D),$$
$$\|v\|^2_{\mathscr{U}} = \int_D \int_D a(x)a(\bar{x})|\nabla_x(\nabla_{\bar{x}})v|^2 < \infty \Big\},$$

$$\mathscr{B}_c(u, v) = \int_D \int_D a(x)a(\bar{x})(\nabla_x u)^{\mathrm{T}}[\nabla_x \nabla_{\bar{x}} v](\nabla_{\bar{x}} u)\,\mathrm{d}x\,\mathrm{d}\bar{x},$$

$$F(x, \bar{x}) = \int_\Gamma \rho(y)f(x, y)f(\bar{x}, y)\,\mathrm{d}y.$$

Then

$$C_0(x, \bar{x}) = \int_\Gamma \rho(y)u(x, y)u(\bar{x}, y) \in \mathscr{U} \tag{7.9}$$

and

$$\mathscr{B}_c(C(x, \bar{x}), v(x, \bar{x})) = \int_D \int_D F(x, \bar{x})v(x, \bar{x})\,\mathrm{d}x\,\mathrm{d}\bar{x}. \tag{7.10}$$

Proof. Relation (7.9) follows from the fact that $u(x,y) \in W$ and the Schwarz inequality.
Let us now prove (7.10). Let $v(x,\bar{x}) \in \mathscr{U}$. Then

$$\int_D \int_D F(x,\bar{x}) v(x,\bar{x})\, \mathrm{d}x\, \mathrm{d}\bar{x} = \int_\Gamma \int_D \int_D \rho(y) f(x,y) f(\bar{x},y) v(x,\bar{x})\, \mathrm{d}x\, \mathrm{d}\bar{x}\, \mathrm{d}y$$

$$\begin{aligned}
&\int_D \int_\Gamma \int_D \rho(y) a(x) \nabla_x u_0(x,y) \cdot \nabla_x v(x,\bar{x}) f(\bar{x},y)\, \mathrm{d}x\, \mathrm{d}\bar{x}\, \mathrm{d}y \\
&\quad = \int_\Gamma \int_D \int_D \rho(y) a(x) a(\bar{x}) (\nabla_x u_0(x,y))^{\mathrm{T}} (\nabla_{\bar{x}} \nabla_x v(x,\bar{x})) \nabla_{\bar{x}} u_0(\bar{x},y)\, \mathrm{d}x\, \mathrm{d}\bar{x}\, \mathrm{d}y \\
&\quad = \mathscr{B}_C(C(x,\bar{x}), v(x,\bar{x}))
\end{aligned}$$

which was to be proven. □

As a a numerical example, consider the one-dimensional problem,

$$\begin{aligned}
&-\frac{\mathrm{d}^2 \widetilde{u}}{\mathrm{d}x^2} = \widetilde{f}(x) \quad \text{on } D = (-0.5, 0.5) \\
&\widetilde{u}(\pm 0.5) = 0.
\end{aligned}$$

Assume that $E[f] = 0$ and the covariance of f to be

$$F(x,\bar{x}) = \sigma^2 \mathrm{e}^{-|x-\bar{x}|}$$

with $\sigma = 0.1$.

Fig. 5 shows the covariance of the solution computed from Theorem 7.2.

Theorem 7.2 shows that the covariance of the random RHS type SPDF can be computed as a deterministic problem on $D \times D$. The strong form is

$$L_x L_{\bar{x}} C(x,\bar{x}) = F(x,\bar{x}) \quad \text{on } D \times D, \tag{7.11}$$

where $L_x, L_{\bar{x}}$ are the operators of the LHS of the equation and $C(x,\bar{x}) = 0$ on $\partial D \times D$ and $D \times \partial D$.

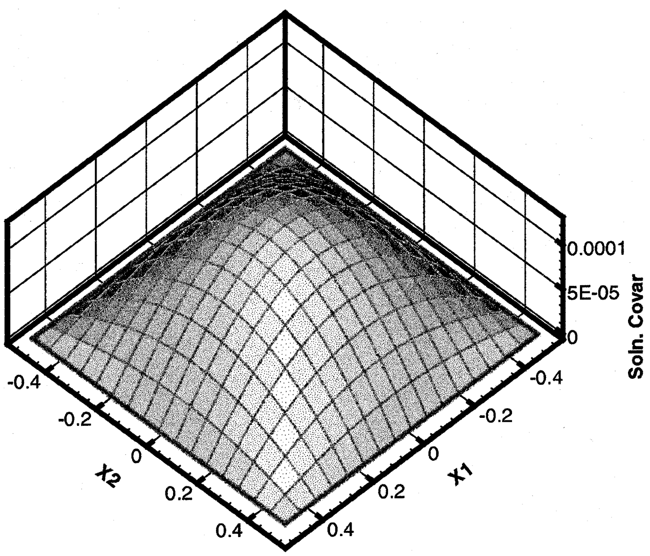

Fig. 5. The covariance of the solution computed by Theorem 7.2.

Remark 7.3. The space $\mathscr{U}$ is the typical space of functions with dominant mixed derivatives. For more, see [2].

Remark 7.4. The proof that the covariance $C(x,\bar{x})$ satisfies (7.11) was proven in [3].

Remark 7.5. Computation of the covariance function directly from (7.6a) and (7.6b) is much more expensive than computation from Theorem 7.2. Using Theorem 7.2 $C(x_1,x_2)$ exists for all $F \in \mathscr{U}'$ and hence the convergence of the series (7.1) can be very weak. Further, in practice, only the covariance function $F(x_1,x_2)$ is available and K–L expansion utilizes it. Hence we should utilize $F(x_1,x_2)$ as much as possible, i.e., to compute directly the covariance by Theorem 7.2.

Remark 7.6. The perturbation approach for dealing with the LHS problem essentially transforms it to the RHS problem.

8. Comments and summary

The aim of the present work was to cast the problem of stochastic PDEs into a framework similar to those familiar in deterministic problems which are suitable for approximation by the finite element method. The FEM is very well-developed, theoretically and practically. A posteriori error estimation is available as well, as are adaptive procedures for h, p and h–p versions of the FEM. The major difficulty with the method described here is the high dimensionality of the problem. This aspect has to be taken into account in the implementation. Obviously, adaptive procedures are necessary for successful solutions. The approach presented in this paper gives a basis for utilizing the large body of known results and methods pertaining to the FEM. Implementational aspects are not addressed here.

It is worthwhile to mention that the use of the Wiener chaos polynomials used e.g., in [8] is essentially the *p*-version in the framework in this paper. We mention only models problems. Analogously, it is possible to analyze stochastic boundary conditions and the problem of stochastic domains. Also the approach presented here can be generalized for other types of differential equations. There are many problems which remain to be addressed. One essential problem is the case in which the K–L series converges very slowly and many terms are needed. It is necessary to weaken the sense of convergence so that the problem with "almost" white noise will also be solvable. We assumed that the data in the stochastic formulation are perfectly known. Of course this is not generally the case and uncertainties are present here too. The influence of these uncertainties has to be analyzed as well. This will be the subject of forthcoming papers.

Acknowledgements

The authors gratefully acknowledge the support of ONR under contract N00014-95-1-0401, NSF Grant DMS 9802367, and support of related work by Sandia National Laboratories under grant BF-2070.

References

[1] M. Ainsworth, J.T. Oden, A Posteriori Error Estimation in Finite Element Analysis, Wiley, New York, 2000.
[2] T.I. Amanov, Spaces of Differentiable Functions with Dominant Mixed Derivatives, Nauka, Moscow, 1976 (in Russian).
[3] I. Babuška, On Randomized Solution of Laplace's Equation, Casopis Pest Mat., 1961.
[4] I. Babuška, T. Strouboulis, The Finite Element Method and its Reliability, Oxford University Press, Oxford, 2001.
[5] M. Deb, I. Babuška, J.T. Oden, Stochastic Finite Element using Galerkin Approximation, Fifth US National Congress on Computational Mechanics, Boulder, 1999.
[6] M. Deb, Solution of stochastic partial differential equations (SPDEs) using Galerkin method: theory and applications, Ph.D. Dissertation, The University of Texas, Austin, 2000.
[7] I. Elishakoff, Y. Ren, The bird's eye view on finite element method for structures with large stochastic variations, Comput. Methods Appl. Mech. Engrg. 168 (1–4) (1999) 51–61.
[8] R. Ghanem, P. Spanos, Stochastic Finite Elements: A Spectral Approach, Springer, Berlin, 1991.

[9] R. Ghanem, Ingredients for a general purpose stochastic finite elements implementation, Comput. Methods Appl. Mech. Engrg. 168 (1–4) (1999) 19–33.
[10] H. Holden, B. Oksendal, J. Uboe, T.S. Zhang, Stochastic Partial Differential Equations – A Modeling White Noise Functional Approach, Birkhauser, Basel, 1966.
[11] M. Kleiber, T.D. Hien, The Stochastic Finite Element Method, Wiley, New York, 1992.
[12] M. Loeve, Probability Theory, fourth ed., Springer, New York, 1977.
[13] B. Oksendal, Stochastic Differential Equations – An Introduction with Application, Springer, Berlin, 1998.
[14] G.I. Schueller and H.J. Pradwarter, Computational stochastic mechanics – current developments and prospects, in: S. Idelsohn, E. Onate, E. Dvorkin (Eds.), Computational Mechanics: New Trends and Applications, CIMNE, Barcelona, 1998.