A new projection scheme for linear stochastic problems

Sachin K. Sachdeva*, Prasanth B. Nair, Andy J. Keane

Computational Engineering and Design Group, School of Engineering Sciences, University of Southampton, Highfield, Southampton SO17 1BJ, UK

Abstract

In this paper we present a new projection scheme for solving linear stochastic partial differential equations. The solution process is approximated using a set of basis vectors spanning a preconditioned stochastic Krylov subspace. We propose a strong Galerkin condition which ensures that the stochastic residual error is orthogonal to the approximating subspace with probability one. We present numerical studies for a model problem in stochastic structural mechanics to demonstrate that the proposed strong Galerkin projection scheme gives better results than the weak Galerkin scheme.

Keywords: Stochastic projection schemes; Polynomial chaos; Krylov subspace

1. Introduction

Linear stochastic partial differential equations (PDEs) arise in a number of areas, including structural mechanics, heat transfer and flow through porous media. If the governing stochastic operator is elliptic and self-adjoint, then discretization in space and the random dimension leads to a linear random algebraic system of equations of the form

$$\mathbf{A}(\boldsymbol{\theta})\mathbf{x}(\boldsymbol{\theta}) = \mathbf{b} \qquad (1)$$

where $\mathbf{A} \in \mathbb{R}^{n\times n}$ is a symmetric positive definite matrix and $\boldsymbol{\theta} \in \mathbb{R}^{p}$ is a vector of random variables.

It is well known that a stochastic projection scheme can be employed to solve Eq. (1). An orthogonal Galerkin projection scheme extracts the approximate solution $\mathbf{x}(\boldsymbol{\theta}) \in \mathbb{R}^{n}$ from an arbitrary subspace $\mathcal{K} \in \mathbb{R}^{n\times m}$ by imposing m orthogonality constraints, where $\dim(\mathcal{K}) = m$. Clearly, the choice of the search subspace $\mathcal{K}$ is critical to the accuracy and efficiency of the numerical scheme employed to solve Eq. (1).

A variety of subspaces have been proposed in the literature for approximating the solution to Eq. (1) as well as a wider class of stochastic operator problems. Ghanem and Spanos [1] proposed an approach referred to as the spectral stochastic finite element method (SSFEM), which makes use of the subspace spanned by multi-dimensional Hermite polynomials. More recently, Xiu and Karniadakis [2] proposed a generalized polynomial chaos (PC) approach where basis functions from the Askey family of hypergeometric polynomials are used. More recently, stochastic reduced basis methods (SRBMs) were introduced that make use of basis vectors spanning the stochastic Krylov subspace defined below [3,4]:

$$\mathcal{K}_m(\mathbf{A}(\boldsymbol{\theta}), \mathbf{b}) = \text{span}\{\mathbf{b}, \mathbf{A}(\boldsymbol{\theta})\mathbf{b}, \mathbf{A}(\boldsymbol{\theta})^2\mathbf{b}, \cdots, \mathbf{A}(\boldsymbol{\theta})^{m-1}\mathbf{b}\} \qquad (2)$$

Recent numerical studies suggest that SRBMs give accuracy levels comparable or better than the SSFEM approach at a significantly lower computational cost for a class of linear stochastic PDEs [5]. This is primarily due to the fact that when a suitably preconditioned stochastic Krylov subspace is used to approximate the solution process, highly accurate results can be obtained using three–four basis vectors (i.e. by solving a small reduced-order deterministic system of equations) even for large coefficients of variation. In contrast, the SSFEM approach involves the solution of a large-scale deterministic system of equations with dimensionality nP, where P is the number of basis functions used in the PC expansion of the solution process. For a detailed exposition of the theoretical foundations of SRBMs, the reader is referred to [6].

The present paper focuses on the application of Galerkin projection schemes in conjunction with the stochastic Krylov subspace to solve linear stochastic

* Corresponding author. Tel.: +44 (7821) 420500; Fax: +44 (2380) 593230; E-mail: sachin@soton.ac.uk

Computational Fluid and Solid Mechanics 2005
K.J. Bathe (Editor)

PDEs. We introduce a strong form of the Galerkin projection scheme in which the stochastic residual is enforced to be orthogonal with respect to the approximating subspace $\mathcal{K}$ with probability one. We present numerical studies to demonstrate that the proposed scheme gives better results than techniques based on the standard weak Galerkin condition.

2. Stochastic projection schemes

Prior to the application of stochastic projection schemes, the governing stochastic PDEs need to be discretized in space and the random dimension. For example, discretization along the random dimension can be carried out using a random field discretization technique such as the Karhunen–Loéve expansion technique or optimal linear estimation; see, for example, [1,7,8,9]. When a lognormal random field is used for modeling uncertainty, it is convenient to use the Karhunen-Loéve expansion in conjunction with the PC decomposition technique [10]. Subsequently, spatial discretization of the governing stochastic PDEs can be carried out using standard techniques such as the finite element method. Irrespective of the particular choice of the spatial and random field discretization technique employed, the coefficient matrix in equation (1) can be written in the general form:

$$\mathbf{A}(\boldsymbol{\theta}) = \sum_{i=0}^{P_1} \mathbf{A}_i \Gamma_i(\boldsymbol{\theta}) \tag{3}$$

where $\mathbf{A}_i \in \mathbb{R}^{n\times n}$ are deterministic matrices, $\Gamma_i(\boldsymbol{\theta})$ are multi-dimensional Hermite polynomials in $\boldsymbol{\theta}$ and P_1 depends on the order of the PC decomposition scheme used to represent the input uncertainty.

A stochastic reduced basis approximation for the solution of Eq. (1) can be written as:

$$\hat{\mathbf{x}}(\theta) = \xi_0\psi_0(\boldsymbol{\theta}) + \xi_1\psi_1(\boldsymbol{\theta}) + \cdots + \xi_m\psi_m(\boldsymbol{\theta}) = \Psi + \boldsymbol{\xi} \tag{4}$$

where $\Psi(\boldsymbol{\theta}) = \{\psi_0(\boldsymbol{\theta}), \psi_1(\boldsymbol{\theta}), \ldots, \psi_m(\boldsymbol{\theta})\} \in \mathbb{R}^{n\times(m+1)}$ is a set of basis vectors spanning the stochastic Krylov subspace $\mathcal{K}_{m+1}$ defined in Eq. (2) and $\boldsymbol{\xi} = \{\xi_0, \xi_1, \ldots, \xi_m\}^T \in \mathbb{R}^{m+1}$ is a vector of undetermined coefficients. In practice, a preconditioned stochastic Krylov subspace is used to ensure accurate approximations using a few basis vectors [4,5,6].

2.1. Weak Galerkin scheme

Substituting Eqs (3) and (4) into Eq. (1), we arrive at the following stochastic residual error vector:

$$\boldsymbol{\varepsilon}(\boldsymbol{\theta}) = \left(\sum_{i=0}^{P_1} \mathbf{A}_i \Gamma_i\right) \Psi(\boldsymbol{\theta})\boldsymbol{\xi} - \mathbf{b} \tag{5}$$

In the Galerkin scheme, $\boldsymbol{\xi}$ is computed by enforcing the condition $\boldsymbol{\varepsilon}(\boldsymbol{\theta}) \perp \psi_i(\boldsymbol{\theta})$, $i = 0, 1, 2, \ldots, m$. In the weak Galerkin scheme we use the standard definition that two random vectors $\mathbf{x}_1(\boldsymbol{\theta})$ and $\mathbf{x}_2(\boldsymbol{\theta})$ are orthogonal to each other if $\langle \mathbf{x}_1^{\mathrm{T}}(\boldsymbol{\theta})\mathbf{x}_2(\boldsymbol{\theta}) \rangle = 0$, where $<\cdot>$ denotes the expectation operator. We refer to the formulation based on this definition of orthogonality as the weak Galerkin scheme.

Application of the weak Galerkin scheme results in the following reduced-order system of deterministic linear algebraic equations:

$$\left\langle \sum_{i=0}^{P_1} \Gamma_i \Psi(\boldsymbol{\theta})^T \mathbf{A}_i \Psi(\boldsymbol{\theta}) \right\rangle \boldsymbol{\xi} = \left\langle \Psi(\boldsymbol{\theta})^T \mathbf{b} \right\rangle \tag{6}$$

Equation (6) can be solved for $\boldsymbol{\xi}$, which can be subsequently substituted into Eq. (4) to obtain the final expression for the solution process.

2.2. Strong Galerkin scheme

We now derive a strong Galerkin scheme in which orthogonality is imposed in a stricter sense. In other words, we seek to compute the vector of undetermined coefficients $\boldsymbol{\xi}$ such that the random functions $\boldsymbol{\varepsilon}^T(\boldsymbol{\theta})\boldsymbol{\psi}_i(\boldsymbol{\theta}) = 0, \forall_i = 0, 1, 2, \ldots, m$ for any realization of the random vector $\boldsymbol{\theta}$. This condition will ensure that the stochastic residual error vector is orthogonal to the m basis vectors with probability one. As shown earlier by Nair and Keane [4], the strong orthogonality condition will be satisfied only when $\boldsymbol{\xi}$ is computed by solving the following reduced-order *random* algebraic system of equations:

$$\left[\sum_{i=0}^{P_1} \Gamma_i \Psi^T(\boldsymbol{\theta}) \mathbf{A}_i \Psi(\boldsymbol{\theta})\right] \boldsymbol{\xi} = \Psi^T(\boldsymbol{\theta})\mathbf{b} \tag{7}$$

As can be seen from the preceding equation, in order to satisfy the strong Galerkin condition, we need to model the undetermined coefficients $\xi_1, \xi_2, \ldots, \xi_m$ as functions of $\boldsymbol{\theta}$, i.e. the stochastic reduced basis approximation in Eq. (4) has to be rewritten as:

$$\hat{\mathbf{x}}(\boldsymbol{\theta}) = \xi_1(\theta)\psi_1(\boldsymbol{\theta}) + \xi_2(\boldsymbol{\theta})\psi_2(\boldsymbol{\theta}) + \cdots + \xi_m(\boldsymbol{\theta})\psi_m(\boldsymbol{\theta}) = \Psi(\boldsymbol{\theta})\boldsymbol{\xi}(\boldsymbol{\theta}) \tag{8}$$

In general, it is not possible to solve (7) explicitly unless further assumptions are made. Our idea is to relax the strong Galerkin condition by employing a PC decomposition of $\boldsymbol{\xi}$ as follows:

$$\xi(\theta) = \sum_{i=0}^{P_2} \xi_i \Gamma_i(\theta) \tag{9}$$

where $\xi_i \in \mathbb{R}^m$, $i = 0, 1, 2, \ldots, P_2$ are undetermined vectors.

Further, since the matrix of stochastic basis vectors $\Psi(\theta)$ are functions of random variables, they can be represented using a PC expansion as:

$$\hat{\Psi}(\theta) = \left[\sum_{i=0}^{P_3} \Pi_i \Gamma_i\right] \tag{10}$$

where $\Pi_i \in \mathbb{R}^{n\times(m+1)}$, $i = 1, 2, \ldots, P_3$ are known deterministic matrices.

Substituting Eqs (9) and (10) into (7) and rearranging terms, we have:

$$\sum_{i=0}^{P_3}\sum_{j=0}^{P_1}\sum_{k=0}^{P_3}\sum_{l=0}^{P_2} \Pi_i^T \mathbf{A}_j \Pi_k \Gamma_i \Gamma_j \Gamma_k \Gamma_l \xi_l = \sum_{i=0}^{P_3} \Pi_i^T \mathbf{b} \Gamma_i \tag{11}$$

In order to compute the vectors of undertermined coefficients ξ_l, $l = 0, 1, 2, \ldots, P_2$, the stochastic residual error vector of the preceding equation is enforced to be orthogonal with respect to the subspace of multi-dimensional Hermite polynomials Γ_e, $e = 1, 2, \ldots, P_2$ in the weak sense, which gives:

$$\left[\sum_{i=0}^{P_3}\sum_{j=0}^{P_1}\sum_{k=0}^{P_3}\sum_{l=0}^{P_2} \Pi_i^T \mathbf{A}_j \Pi_k \langle \Gamma_i \Gamma_j \Gamma_k \Gamma_l \Gamma_e \rangle\right] \xi_l = \sum_{i=0}^{P_3} \Pi_i^T \mathbf{b} \langle \Gamma_i \Gamma_e \rangle, \quad e = 1, 2, \ldots, P_2 \tag{12}$$

Equation (12) is a system of $(m+1)\ P_2 \times (m+1)\ P_2$ deterministic linear algebraic equations which can be solved for ξ_i, $i = 0, 1, \ldots, P_2$. This solution can be subsequently used in conjunction with Eqs.(8), (9) and (10) to arrive at the final expression for the solution process. The resulting expression for $\hat{\mathbf{x}}(\theta)$ can be readily post-processed to compute the complete statistics of the solution process.

It is worth noting here if we set $P_2 = 0$ in Eq. (9) then the strong Galerkin scheme reduces to the weak Galerkin scheme presented earlier. Hence, the strong Galerkin scheme proposed here can be viewed as a generalization of standard SRBMs proposed earlier in the literature [3,4]. Also due to Theorem 2 of Nair and Keane [4], it follows that the strong Galerkin scheme results in a lower value of the **A**-norm of the error compared to the weak scheme. As a consequence, the strong formulation is guaranteed to be more accurate in the sense of this error norm.

3. Numerical studies

We consider a thin square plate of unit length clamped at one edge and subjected to uniform inplane tension at the opposite edge; see [1,7] for details. The Youngs modulus of the plate is modeled using both

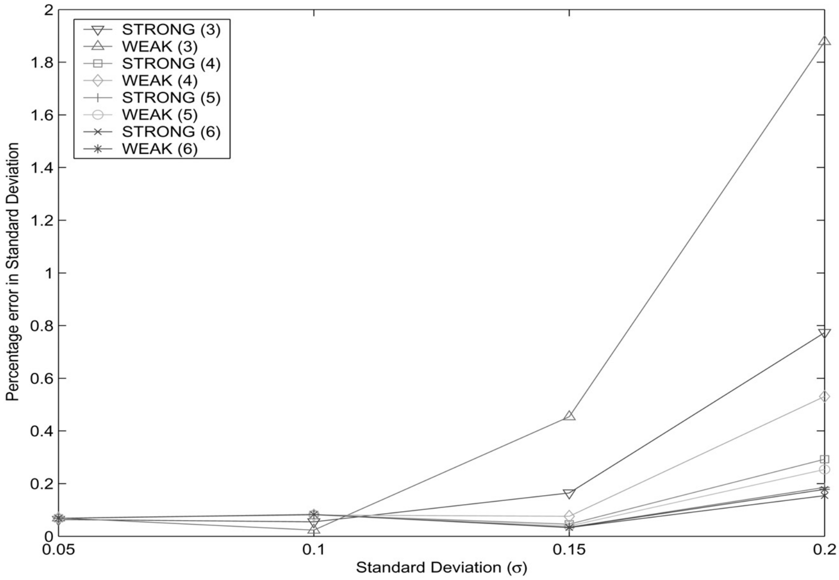

Fig. 1. Percentage error in the standard deviation of the displacement using weak and strong Galerkin schemes (Gaussian model). The numbers within brackets in the legend denote the number of basis vectors.

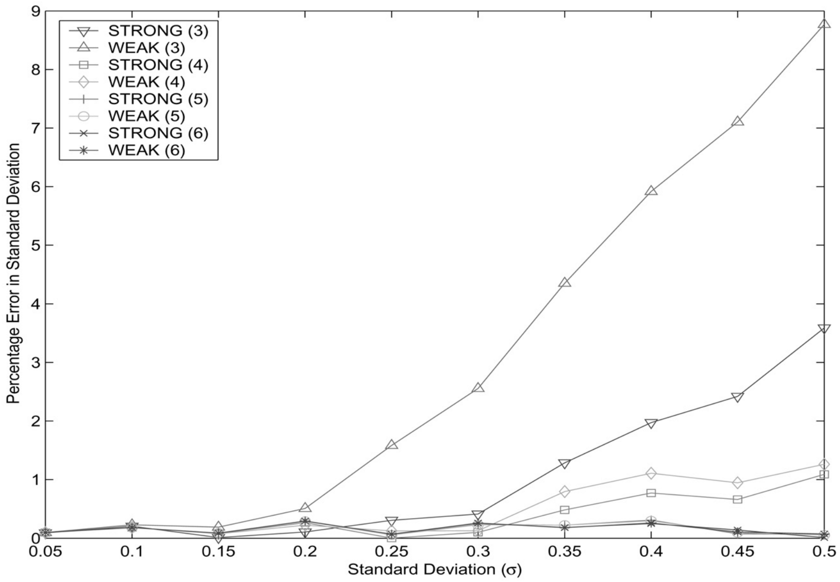

Fig. 2. Percentage error in the standard deviation of the displacement using weak and strong Galerkin schemes (lognormal model). The numbers within brackets in the legend denote the number of basis vectors.

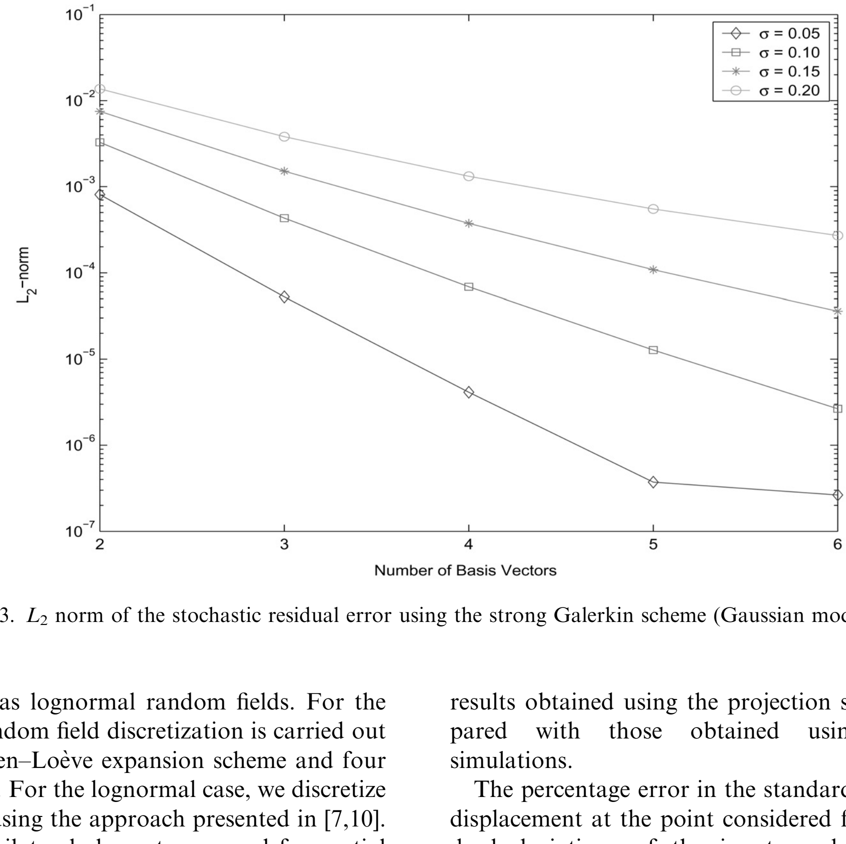

Fig. 3. L_2 norm of the stochastic residual error using the strong Galerkin scheme (Gaussian model).

Gaussian as well as lognormal random fields. For the Gaussian case, random field discretization is carried out using the Karhunen–Loève expansion scheme and four terms are retained. For the lognormal case, we discretize the random field using the approach presented in [7,10]. Four-noded quadrilateral elements are used for spatial discretization. We compute the standard deviation of the displacement at a point with maximum strain. The results obtained using the projection schemes are compared with those obtained using Monte-Carlo simulations.

The percentage error in the standard deviation of the displacement at the point considered for different standard deviations of the input random field (σ) are presented in Figs 1 and 2 for Gaussian and lognormal models, respectively. Fig. 3 shows the convergence of the

Table 1
Comparison of the percentage error in standard deviation of displacement computed using various projection schemes; Gaussian uncertainty model with $\sigma = 0.2$

Order	Strong Galerkin scheme	Weak Galerkin scheme	SSFEM PC projection scheme
1	3.459	10.423	10.474
2	0.790	1.877	1.976
3	0.296	0.543	0.543
4	0.197	0.247	0.247
5	0.148	0.197	0.197

L_2 norm of the residual error for the strong Galerkin scheme as the number of basis vectors is increased for different values of σ. In the figures shown, note that the number within brackets in the legends denotes the number of basis vectors used. It can be observed that the error norm converges rapidly when the number of basis vectors is increased. Similar convergence trends are observed when the lognormal model is used to represent uncertainty.

It is evident from the results that the strong Galerkin scheme gives better performance compared to the weak Galerkin scheme for the same number of basis vectors. For the problem considered it was found that the computational cost of the strong formulation is only marginally higher than the weak formulation.

Table 1 compares the results obtained using the weak and strong Galerkin formulations with those computed using the SSFEM approach employing PC expansions [1]. The percentage errors (in standard deviation of displacement) presented in the table were obtained for the case when the Gaussian model is used with $\sigma = 0.2$. It can be seen that the strong Galerkin scheme gives more accurate results than the other projection schemes. Finally, it is also worth noting that the schemes presented here are orders of magnitude more efficient than the SSFEM approach; see [5] for a detailed comparison of the computational efficiency of the weak Galerkin scheme and the SSFEM formulation.

4. Conclusions

We proposed a strong Galerkin projection scheme for approximating the solution of linear random algebraic equations arising from discretization of linear stochastic partial differential equations. It is shown that the proposed formulation is a generalization of stochastic reduced basis methods based on the weak Galerkin condition. Numerical studies are presented to demonstrate that the strong Galerkin projection scheme gives better results compared to the weak scheme for the same number of basis vectors chosen from the stochastic Krylov subspace. The strong condition is shown to work well for both Gaussian as well as lognormal uncertainty models. Finally, a selection of results are presented to illustrate that the proposed formulation gives more accurate results than the SSFEM approach, while incurring significantly lower computational cost.

References

[1] Ghanem R and Spanos P. Stochastic Finite Elements: A Spectral Approach. Berlin: Springer-Verlag, 1991.
[2] Xiu D, Karniadakis GE. The Wiener–Askey polynomial chaos for stochastic differential equations. SIAM Journal of Scientific Computing 2002;24:619–644.
[3] Nair PB. On the theoretical foundations of stochastic reduced basis methods. AIAA Paper 2001–1677, 2001.
[4] Nair PB, Keane AJ. Stochastic reduced basis methods. AIAA J 2002;40:1653–1664.
[5] Sachdeva SK, Nair PB, Keane AJ. Comparative study of projection schemes for stochastic finite element analysis. Computer Methods in Applied Mechanics and Engineering, submitted for review, 2003.
[6] Nair PB. Projection schemes in stochastic finite element analysis. CRC Engineering Design Reliability Handbook, Boca Raton, FL: CRC Press, 2004.
[7] Sudret B, Der Kiureghian A. Stochastic finite elements and reliability: a state-of-the-art report, technical report no. ucb/semm-2000/08, University of California, Berkeley, p. 173, 2000.
[8] Li CC, Der Kiureghian A. Optimal discretization of random fields. Journal of Engineering Mechanics, ASCE 1993;119:1136–1154.
[9] Huang SP, Quek ST, Phoon KK. Convergence study of the truncated Karhunen–Loéve expansion for simulation of stochastic processes. International Journal for Numerical Methods in Engineering 2001;52:1029–1043.
[10] Ghanem R. The nonlinear gaussian spectrum of log-normal stochastic processes and variables. Journal of Applied Mechanics 1999;66:964–973.