AIAA-2001-1677

On the Theoretical Foundations of Stochastic Reduced Basis Methods

Prasanth B. Nair*
University of Southampton, Highfield, Southampton SO17 1BJ, U.K

Abstract

Stochastic reduced basis methods (SRBMs) are a class of numerical techniques for approximately computing the response of stochastic systems. The basic idea is to approximate the response using a linear combination of stochastic basis vectors with undetermined coefficients. In this paper, we examine the theoretical foundations of SRBMs by exploring their relationship with Krylov subspace methods for deterministic systems. The mathematical justification for employing the terms of the stochastic Krylov subspace as basis vectors is presented. It is shown that SRBMs are a stochastic generalization of preconditioned Krylov subspace methods. Subsequently, some approaches for stochastic generalization of the Bubnov-Galerkin scheme are analyzed. We also address the issue of computing *a posteriori* error estimates of SRBMs. Some preliminary numerical studies are presented for examining the accuracy of the error estimates. The paper concludes with a discussion of ongoing work on algebraic random eigenvalue problems.

Introduction

The field of computational stochastic mechanics in general is concerned with the development of efficient numerical schemes for analysis of systems governed by stochastic partial differential equations (SPDEs). A fundamental topic in this area involves addressing issues related to realistic modeling of parametric uncertainty. Given the representation of the uncertain system parameters and the environment either in terms of random fields or random variables, it becomes possible to integrate discretization methods for the response and the random fields to arrive at a system of coupled ordinary differential equations (ODEs) with random coefficients; see, for example, Ghanem and Spanos[1]. In such approaches, randomness is embodied as an additional dimension of the problem. For this line of approach to be successful in practice, it is crucial to have general-purpose numerical schemes for the solution of random algebraic equations.

Over the last few decades, researchers in stochastic mechanics have mostly used perturbation or series expansion methods for solving a finite-dimensional approximation of the governing SPDE; see, for example, Kleiber and Hien[2]. These methods although general-purpose in scope, have limited accuracy for large stochasticity. Few papers in the literature have focused on alternative approaches which may potentially lead to *physics-based*, general purpose methods in the long term future. Some notable exceptions include the so-called spectral stochastic finite element method (SSFEM)[1,3] and variational principles for stochastic systems.[4] An excellent discussion on the motivation for additional work in the area of computational stochastic mechanics has been presented by Elishakoff and Ren[5].

More recently, Nair and Keane[6,7] introduced stochastic reduced basis methods (SRBMs) for numerical solution of systems of linear algebraic equations with random coefficients. SRBMs are essentially a family of numerical schemes for computing an approximate solution to a system of linear algebraic equations with a random coefficient matrix and right hand side. It was proposed that the solution process (say $\mathbf{x}(\boldsymbol{\theta}) \in \mathbb{R}^n$) can be approximated in a subspace spanned by a set of stochastic basis vectors as

$$\mathbf{x}(\boldsymbol{\theta}) = \xi_0(\boldsymbol{\theta})\boldsymbol{\psi}_0(\boldsymbol{\theta}) + \ldots \xi_{m-1}(\boldsymbol{\theta})\boldsymbol{\psi}_{m-1}(\boldsymbol{\theta}), \quad (1)$$

where $\boldsymbol{\psi}_0(\boldsymbol{\theta})$, $\boldsymbol{\psi}_1(\boldsymbol{\theta})$, $\ldots$, $\boldsymbol{\psi}_{m-1}(\boldsymbol{\theta}) \in \mathbb{R}^n$ denotes a set of m stochastic basis vectors, and $\xi_0(\boldsymbol{\theta})$, $\xi_1(\boldsymbol{\theta})$, $\ldots$, $\xi_{m-1}(\boldsymbol{\theta}) \in \mathbb{R}^1$ denotes the m random undetermined coefficients in the reduced basis. Note that for the sake of computational efficiency we will choose $m <<< n$. Equation (1) will be referred to as the stochastic reduced basis representation (SRBR)

*Research Fellow, Computational Engineering and Design Center, School of Engineering Sciences, Member AIAA. P.B.Nair@soton.ac.uk.

throughout this paper.

In reference[6], it was proposed that the terms of the Neumann expansion series can be deployed as stochastic basis vectors. An alternative choice involving the use of the first-order Taylor series was also suggested. The undetermined coefficients in the SRBR were then computed using stochastic variants of the Bubnov-Galerkin (BG) scheme. Some ideas on the extension of SRBMs to nonlinear stochastic systems were also presented without numerical validation. In summary, the key ingredients of SRBMs are: (a) the choice of stochastic basis vectors, and (b) the BG scheme employed to compute the undetermined coefficients in the SRBR. Note that there also exists an alternative approach for computing the coefficients of the SRBR; see Appendix A for details.

This paper is concerned with the theoretical foundations of SRBMs. Much of the theoretical analysis presented in this paper stems from the observation that the terms of the Neumann series form a preconditioned stochastic Krylov subspace (see the next section for the definition of a stochastic Krylov subspace). In numerical linear algebra, methods based on the Krylov subspace have a history of nearly 50 years of existence, and they continue to be an area of extensive research; see Saad and Van der Vorst[8] for a historical overview. The literature on this topic dedicated to the solution of deterministic linear systems is vast; the interested reader is referred to the books by Saad[9], Axelsson[10], and Golub and Van Loan[11]. Not surprisingly, there are intimate connections between SRBMs and preconditioned solvers for deterministic systems. This observation provides us access to a solid theoretical foundation using which the theoretical frontiers of *numerical stochastic linear algebra* can be extended.

Another motivation for the present research comes from a recent theoretical study presented by Fokkema et al.[12] It was shown that a wide class of Krylov subspace methods can be viewed as instances of an accelerated inexact Newton scheme. This is a very important observation which may potentially aid us in the task of design and analysis of mathematically rigorous extensions of SRBMs for analysis of nonlinear stochastic systems.

In this paper, we present the mathematical justification for employing the terms of the stochastic Krylov subspace as basis vectors to compute the solution of a system of linear random algebraic equations. A stochastic version of Arnoldi's method is then presented for computing a set of orthonormal basis vectors spanning the stochastic Krylov subspace. This version of Arnoldi's method also helps to clarify the links between SRBMs and Krylov methods for deterministic linear systems. Subsequently, the idea behind using the preconditioned stochastic Krylov subspace is discussed.

Various approaches for stochastic generalization of the BG scheme are then analyzed. This analysis shows that the BG scheme can only be implemented in an approximate sense for stochastic problems. As discussed in reference[6], the BG scheme can be implemented exactly (in a computationally efficient fashion) only for special cases involving 2 to 3 stochastic basis vectors. We show that the BG scheme commonly used in the literature (see, for example, Ghanem and Spanos[1]) is a zero-order approximation. Further, the present analysis suggests that, in practice, the order of the approximation can be potentially improved by including higher-order terms. Conditions under which this improvement is possible are postulated.

Next, we consider the problem of efficiently computing *a posteriori* error estimates for SRBMs. In contrast to deterministic problems, the error estimates are themselves random functions. Formulations are hence developed for approximately computing error estimates for the first and second order statistics of the computed solution. Some preliminary studies on the accuracy of the error estimates are presented for static response analysis of a stochastic structural system. Motivated by the hybrid stochastic finite element formulation recently proposed by Ghanem[17], a simulation scheme is then presented for error estimation as well improvement of the accuracy of SRBM formulation.

We conclude this paper with a discussion of ongoing work on solving algebraic random eigenvalue problems using a stochastic reduced basis representation. The idea of using a preconditioned stochastic subspace to achieve this goal is outlined.

Preliminaries

We use the following notation throughout this paper. For the sake of generality, all the vector spaces considered are complex, unless otherwise stated. Random quantities are indicated explicitly as a function of $\boldsymbol{\theta}$ or $\boldsymbol{\eta}$, and the ensemble average is denoted as $\langle . \rangle$. We denote vectors and matrices by lower case and upper case bold letters, respectively. We use the notation $\mathbf{x}^*$ to denote the complex conjugate transpose of a vector/matrix (if it is complex), or the transpose (if it is real). We now introduce the following definitions which are extensively used throughout this paper.

Definition 1: Two random vectors $\mathbf{x}_1(\boldsymbol{\theta})$ and $\mathbf{x}_2(\boldsymbol{\theta}) \in \mathbb{C}^n$ are said to be orthogonal in the Hilbert space of random variables if $\langle \mathbf{x}_1^*(\boldsymbol{\theta})\mathbf{x}_2(\boldsymbol{\theta})\rangle = 0$.

Definition 2: Two random vectors $\mathbf{x}_1(\boldsymbol{\theta})$ and $\mathbf{x}_2(\boldsymbol{\theta}) \in \mathbb{C}^n$ are said to be orthogonal in an exact sense if $P[\mathbf{x}_1^*(\boldsymbol{\theta})\mathbf{x}_2(\boldsymbol{\theta}) = 0] = 1$.

The reader is referred to reference[1] for a more formal review of the origins of Definition 1. Definition 2 (which is rather restrictive) is a trivial probabilistic interpretation of the definition of orthogonality between two deterministic vectors.

Definition 3: The stochastic Krylov subspace of order m is defined as

$$\mathcal{K}_m(\mathbf{B}(\boldsymbol{\theta}), \mathbf{y}(\boldsymbol{\theta})) = span\{\mathbf{y}(\boldsymbol{\theta}), \mathbf{B}(\boldsymbol{\theta})\mathbf{y}(\boldsymbol{\theta}), \mathbf{B}^2(\boldsymbol{\theta})\mathbf{y}(\boldsymbol{\theta}), \ldots, \mathbf{B}(\boldsymbol{\theta})^{m-1}\mathbf{y}(\boldsymbol{\theta})\}$$

where $\mathbf{B}(\boldsymbol{\theta}) \in \mathbb{C}^{n\times n}$ is a random matrix, and $\mathbf{y}(\boldsymbol{\theta}) \in \mathbb{C}^n$ is either a random or deterministic vector.

Definition 4: The stochastic Hessenberg reduction of a random matrix $\mathbf{A}(\boldsymbol{\theta}) \in \mathbb{C}^{n\times n}$ is defined as $\mathbf{V}^*(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})\mathbf{V}(\boldsymbol{\theta})$, such that $\mathbf{H} = \langle \mathbf{V}^*(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})\mathbf{V}(\boldsymbol{\theta})\rangle$ is a Hessenberg matrix*, where $\mathbf{V}(\boldsymbol{\theta}) \in \mathbb{C}^{n\times m}$ is a matrix of orthonormal random vectors in the sense of Definition 1, i.e., $\langle \mathbf{V}^*(\boldsymbol{\theta})\mathbf{V}(\boldsymbol{\theta})\rangle = \mathbf{I}_m$; where $\mathbf{I}_m \in \mathbb{R}^{m\times m}$ is the identity matrix.

Why Employ a Stochastic Krylov Subspace ?

In this section, we briefly address the question why a stochastic Krylov subspace is appropriate for computing the solution of a linear random algebraic system of equations. For a detailed overview of this argument for a deterministic matrix, the reader is referred to Ipsen and Meyer.[13]

Consider the problem, where given a random matrix $\mathbf{A}(\boldsymbol{\theta})$, it is required to compute the vector $\mathbf{A}(\boldsymbol{\theta})^{-1}\mathbf{f}(\boldsymbol{\theta})$. Let us first consider the notion of the minimal polynomial q of a random matrix. For a random matrix, $q(t, \boldsymbol{\theta})$ can be defined as the unique monic random polynomial of minimal degree such that $q(\mathbf{A}(\boldsymbol{\theta}), \boldsymbol{\theta}) = 0$.

Let the d distinct eigenvalues of $\mathbf{A}(\boldsymbol{\theta})$ be denoted by $\lambda_1(\boldsymbol{\theta}), \lambda_2(\boldsymbol{\theta}), \ldots \lambda_d(\boldsymbol{\theta})$, and let m_j denote the index of the jth eigenvalue (i.e., the size of a largest Jordan block associated with it). Then, it follows that

$$m = \sum_{j=1}^{d} m_j, \quad \text{and} \quad q(t, \boldsymbol{\theta}) = \prod_{j=1}^{d} (t - \lambda_j(\boldsymbol{\theta}))^{m_j} \tag{2}$$

Using the preceding equation, we can write

$$q(t, \boldsymbol{\theta}) = \sum_{j=0}^{m} \alpha_j(\boldsymbol{\theta}) t^j \tag{3}$$

where the term $\alpha_0(\boldsymbol{\theta}) = \prod_{j=1}^{d}(-\lambda_j(\boldsymbol{\theta}))^{m_j}$ is nonzero if the random matrix $\mathbf{A}(\boldsymbol{\theta})$ is nonsingular.

The minimal polynomial can hence be written as

$$q(\mathbf{A}(\boldsymbol{\theta}), \boldsymbol{\theta}) = \sum_{j=0}^{m} \alpha_j(\boldsymbol{\theta})\mathbf{A}^j(\boldsymbol{\theta}) = 0 \tag{4}$$

Since $\alpha_0(\boldsymbol{\theta}) \neq 0$, the inverse of $\mathbf{A}(\boldsymbol{\theta})$ can be written as

$$\mathbf{A}(\boldsymbol{\theta})^{-1} = -\frac{1}{\alpha_0(\boldsymbol{\theta})} \sum_{j=0}^{m-1} \alpha_{j+1}(\boldsymbol{\theta})\mathbf{A}^j(\boldsymbol{\theta}) \tag{5}$$

It can be clearly seen from (5) that the inverse of a nonsingular random matrix lies in the space spanned by the terms of the minimal random polynomial. This implies that desired solution vector $\mathbf{A}(\boldsymbol{\theta})^{-1}\mathbf{f}(\boldsymbol{\theta})$ belongs to the stochastic Krylov subspace defined earlier (see Definition 3). The dimension of the subspace depends on the degree of the minimal polynomial of the random matrix. This observation can be formally stated as follows.

Theorem 1 *If the minimal random polynomial of a nonsingular random matrix $\mathbf{A}(\boldsymbol{\theta})$ has degree m, then the solution to $\mathbf{A}(\boldsymbol{\theta})\mathbf{u}(\boldsymbol{\theta}) = \mathbf{f}(\boldsymbol{\theta})$ lies in the stochastic Krylov space $\mathcal{K}_m(\mathbf{A}(\boldsymbol{\theta}), \mathbf{f}(\boldsymbol{\theta}))$.*

Theorem 1 clearly suggests that $\mathcal{K}_m(\mathbf{A}(\boldsymbol{\theta}), \mathbf{f}(\boldsymbol{\theta}))$ provides a stochastic subspace using which a good approximation for the random vector $\mathbf{A}(\boldsymbol{\theta})^{-1}\mathbf{f}(\boldsymbol{\theta})$ can be computed. The next section presents a stochastic version of Arnoldi's method for computing an orthogonal basis of the stochastic Krylov subspace.

Stochastic Variant of Arnoldi's Method

Arnoldi's method[14] is a version of the Gram-Schmidt orthogonalization procedure tailored to the Krylov subspace. This algorithm was introduced in 1951 for reducing a general dense deterministic matrix into Hessenberg form. Note that this algorithm reduces to the well-known Lanczos algorithm for the case of symmetric/Hermitian matrices, i.e., the Hessenberg form becomes tridiagonal. In this section, we

*A Hessenberg matrix is a upper triangular matrix with an additional off-diagonal below the diagonal; i.e., $H_{ij} = 0, i > j + 1$.

develop a stochastic version of Arnoldi's method for reducing a random matrix into the so called *stochastic Hessenberg* form defined earlier (see Definition 4). The objective is to illustrate connections with Krylov methods for deterministic systems and to improve the numerical properties of SRBMs.

In this section the compact notation $\mathcal{K}_m(\boldsymbol{\theta})$ is used to denote the mth order stochastic Krylov subspace $\mathcal{K}_m(\mathbf{A}(\boldsymbol{\theta}), \mathbf{v}(\boldsymbol{\theta}))$. Further, the $\mathcal{L}_2$ norm of a random vector $\mathbf{x}(\boldsymbol{\theta})$ is defined as

$$||\mathbf{x}(\boldsymbol{\theta})||_2 = (\langle \mathbf{x}^*(\boldsymbol{\theta})\mathbf{x}(\boldsymbol{\theta})\rangle)^{\frac{1}{2}} \tag{6}$$

Let the starting normalized random vector be denoted by $\mathbf{v}_1(\boldsymbol{\theta}) = \tilde{\mathbf{v}}_1(\boldsymbol{\theta})/||\tilde{\mathbf{v}}_1(\boldsymbol{\theta})||$, which is a basis for $\mathcal{K}_1(\boldsymbol{\theta})$. Arnoldi's method proceeds by recursively orthogonalizing the vector $\tilde{\mathbf{v}}_{j+1}(\boldsymbol{\theta}) = \mathbf{A}(\boldsymbol{\theta})\mathbf{v}_j(\boldsymbol{\theta})$ with respect to $\mathbf{v}_j(\boldsymbol{\theta})$. This can be written as

$$\tilde{\mathbf{v}}_{j+1}(\boldsymbol{\theta}) = \mathbf{A}(\boldsymbol{\theta})\mathbf{v}_j(\boldsymbol{\theta}) - (H_{1j}\mathbf{v}_1(\boldsymbol{\theta}) + \ldots + H_{jj}\mathbf{v}_j(\boldsymbol{\theta})) \tag{7}$$

where $H_{ij} = \langle \mathbf{v}_i^*(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})\mathbf{v}_j(\boldsymbol{\theta})\rangle$. The new basis vector is then normalized as

$$\mathbf{v}_{j+1}(\boldsymbol{\theta}) = \tilde{\mathbf{v}}_{j+1}(\boldsymbol{\theta})/||\tilde{\mathbf{v}}_{j+1}(\boldsymbol{\theta})||_2 \tag{8}$$

Let $\mathbf{V}_j(\boldsymbol{\theta})$ denote the matrix formed from the j orthonormal stochastic basis vectors. Then the decomposition associated with the stochastic version of Arnoldi's method can be written as

$$\mathbf{A}(\boldsymbol{\theta})\mathbf{V}_j(\boldsymbol{\theta}) = \mathbf{V}_{j+1}(\boldsymbol{\theta})\mathbf{H}_j \tag{9}$$

where $\mathbf{H}_j \in \mathbb{C}^{(j+1)\times j}$ is an upper Hessenberg matrix.

The expectation operations for computing the $\mathcal{L}_2$ norm in the preceding equations can be efficiently conducted when the terms of the stochastic Krylov subspace are written explicit functions of the random parameter vector $\boldsymbol{\theta}$ (see next section).

Stochastic Basis Vectors

In section, we discuss the idea behind preconditioning, and examine issues involved in efficiently computing the terms of the preconditioned stochastic Krylov subspace. Consider a system of linear random algebraic equations of the form

$$\mathbf{A}(\boldsymbol{\theta})\mathbf{x}(\boldsymbol{\theta}, \boldsymbol{\eta}) = \mathbf{f}(\boldsymbol{\eta}), \tag{10}$$

where $\mathbf{A}(\boldsymbol{\theta}) \in \mathbb{C}^{n\times n}$ is a random matrix, $\mathbf{x}(\boldsymbol{\theta}, \boldsymbol{\eta})$, and $\mathbf{f}(\boldsymbol{\eta}) \in \mathbb{C}^n$ are random vectors. $\boldsymbol{\theta} \in \mathbb{C}^p$ and $\boldsymbol{\eta} \in \mathbb{C}^q$ denotes the vector of random variables in the coefficient matrix and the right hand side, respectively.

A SRBR of the response process $\mathbf{x}(\boldsymbol{\theta}, \boldsymbol{\eta})$ stated earlier in equation (1) can be written in matrix form as

$$\hat{\mathbf{x}}(\boldsymbol{\theta}, \boldsymbol{\eta}) = \Psi(\boldsymbol{\theta}, \boldsymbol{\eta})\Xi(\boldsymbol{\theta}, \boldsymbol{\eta}), \tag{11}$$

where $\Psi(\boldsymbol{\theta}, \boldsymbol{\eta}) = [\psi_0(\boldsymbol{\theta}, \boldsymbol{\eta}), \psi_1(\boldsymbol{\theta}, \boldsymbol{\eta}), \ldots \psi_{m-1}(\boldsymbol{\theta}, \boldsymbol{\eta})] \in \mathbb{C}^{n\times m}$, and $\Xi(\boldsymbol{\theta}, \boldsymbol{\eta}) = \{\xi_0, \xi_1, \ldots, \xi_{m-1}\} \in \mathbb{C}^m$ denotes the matrix of stochastic basis vectors and the vector of undetermined coefficients, respectively.

A straight forward choice of basis vectors would be the mth order stochastic Krylov subspace $\mathcal{K}_m(\mathbf{A}(\boldsymbol{\theta}), \mathbf{f}(\boldsymbol{\eta}))$, i.e.,

$$\psi_0 = \mathbf{f}(\boldsymbol{\eta}), \psi_1 = \mathbf{A}(\boldsymbol{\theta})\mathbf{f}(\boldsymbol{\eta}), \ldots, \psi_{m-1} = \mathbf{A}(\boldsymbol{\theta})^{m-1}\mathbf{f}(\boldsymbol{\eta}). \tag{12}$$

However, according to Theorem 1, the number of stochastic basis vectors required to approximate the solution vector could be as high as n. For the sake of computational efficiency, it is desirable to choose a small number of basis vectors. In order to arrive at a richer stochastic subspace, the concept of preconditioning approach becomes useful. A good choice[†] of preconditioner would be the deterministic matrix $\langle \mathbf{A}(\boldsymbol{\theta})\rangle^{-1}$. The left preconditioned version of (3) can hence be written as $\langle \mathbf{A}(\boldsymbol{\theta})\rangle^{-1}\mathbf{A}(\boldsymbol{\theta})\mathbf{x}(\boldsymbol{\theta}, \boldsymbol{\eta}) = \langle \mathbf{A}(\boldsymbol{\theta})\rangle^{-1}\mathbf{f}(\boldsymbol{\eta})$.

The stochastic Krylov subspace associated with the left preconditioned version of (10) coincides with the terms of the Neumann series

$$\tilde{\mathbf{x}}(\boldsymbol{\theta}, \boldsymbol{\eta}) = \sum_{i=0}^{\infty}(-1)^i \left(\langle \mathbf{A}(\boldsymbol{\theta})\rangle^{-1}\mathbf{A}(\boldsymbol{\theta}) - \mathbf{I}_m\right)^i \times \langle \mathbf{A}(\boldsymbol{\theta})\rangle^{-1}\mathbf{f}(\boldsymbol{\theta}, \boldsymbol{\eta}). \tag{13}$$

This clearly shows that the method presented in reference[6] is a preconditioned stochastic Krylov method.

Let us now consider the case when the random coefficient matrix can be written in the form $\mathbf{A}(\boldsymbol{\theta}) = \langle \mathbf{A}(\boldsymbol{\theta})\rangle + \sum_{i=1}^{p}\theta_i\mathbf{A}^i$, and the right hand side (rhs) vector can be written as $\mathbf{f}(\boldsymbol{\eta}) = \langle \mathbf{f}(\boldsymbol{\theta})\rangle + \sum_{i=1}^{q}\eta_i\mathbf{f}^i$; where $\mathbf{A}^i \in \mathbb{C}^{n\times n}$ and $\mathbf{f}^i \in \mathbb{C}^n$ are deterministic. Since, we are primarily interested in physical systems governed by stochastic PDEs, such a representation of the random coefficient matrix and the rhs is readily achievable in terms of the random physical parameters of the system. Note that this representation is exact when the coefficient matrix and

[†]The matrix $\mathbf{M}$ is a good preconditioner if $\mathbf{M}^{-1}\mathbf{A}$ is close to an identity matrix or a matrix with highly clustered eigenvalues.

the rhs are linear functions of the random variables; e.g., this representation is exact when the Young's modulus of a beam member is modeled as a random variable. Note that the formulation presented here can also tackle the case when the chaos decomposition scheme[1] is employed to expand the coefficient matrix and the rhs in terms of orthogonal random polynomials.

Equation (13) can hence be written as

$$\tilde{\mathbf{x}}(\boldsymbol{\theta},\boldsymbol{\eta}) = \sum_{i=0}^{\infty}(-1)^i \left(\langle \mathbf{A}(\boldsymbol{\theta}) \rangle^{-1} \sum_{i=1}^{p} \theta_i \mathbf{A}^i \right)^i \times \langle \mathbf{A}(\boldsymbol{\theta}) \rangle^{-1} \{ \langle \mathbf{f}(\boldsymbol{\theta}) \rangle + \sum_{i=1}^{p} \eta_i \mathbf{f}^i \} \qquad (14)$$

It can be readily seen that each term of the Neumann series (except for the first term) is a vector of homogeneous random polynomials. Each term (say $\tilde{\psi}_i$) can be explicitly written as a function of $\boldsymbol{\theta}$ and $\boldsymbol{\eta}$ as shown below.

$$\tilde{\psi}_0(\boldsymbol{\theta},\boldsymbol{\eta}) = \mathbf{u}_o + \sum_{i=1}^{q} \eta_i \mathbf{u}_i \qquad (15)$$

where $\mathbf{u}_o = \langle \mathbf{A}(\boldsymbol{\theta}) \rangle^{-1} \langle \mathbf{f}(\boldsymbol{\eta}) \rangle \in \mathbb{C}^n$ and $\mathbf{u}_i = \langle \mathbf{A}(\boldsymbol{\theta}) \rangle^{-1} \mathbf{f}^i \in \mathbb{C}^n$ are deterministic vectors.

$$\tilde{\psi}_1(\boldsymbol{\theta},\boldsymbol{\eta}) = \sum_{i=1}^{p} \theta_i \left(\mathbf{b}_i^1 + \sum_{j=1}^{q} \eta_j \mathbf{c}_{ij}^1 \right), \qquad (16)$$

where $\mathbf{b}_i^1 = \langle \mathbf{A}(\boldsymbol{\theta}) \rangle^{-1} \mathbf{A}^i \mathbf{u}_o \in \mathbb{C}^n$ and $\mathbf{c}_{ij}^1 = \langle \mathbf{A}(\boldsymbol{\theta}) \rangle^{-1} \mathbf{A}^i \mathbf{u}_j \in \mathbb{C}^n$ are deterministic vectors.

A general expression for the kth basis vector may be written as

$$\tilde{\psi}_k(\boldsymbol{\theta},\boldsymbol{\eta}) = \sum_{i_1,i_2,\ldots,i_k=1}^{p} \theta_{i_1}\theta_{i_2}\ldots\theta_{i_k} \times \left(\mathbf{b}_{i_1 i_2 \ldots i_k}^k + \sum_{i_{m+1}=1}^{q} \eta_{i_{m+1}} \mathbf{c}_{i_1 i_2 \ldots i_{m+1}}^k \right) \qquad (17)$$

where

$$\mathbf{b}_{i_1 i_2 \ldots i_k}^k = \langle \mathbf{A}(\boldsymbol{\theta}) \rangle^{-1} \mathbf{A}^{i_1} \langle \mathbf{A}(\boldsymbol{\theta}) \rangle^{-1} \mathbf{A}^{i_2} \ldots \langle \mathbf{A}(\boldsymbol{\theta}) \rangle^{-1} \mathbf{A}^{i_k} \mathbf{u}_o \in \mathbb{C}^n$$

and

$$\mathbf{c}_{i_1 i_2 \ldots i_{k+1}}^k = \langle \mathbf{A}(\boldsymbol{\theta}) \rangle^{-1} \mathbf{A}^{i_1} \langle \mathbf{A}(\boldsymbol{\theta}) \rangle^{-1} \mathbf{A}^{i_2} \ldots \langle \mathbf{A}(\boldsymbol{\theta}) \rangle^{-1} \mathbf{A}^{i_k} \mathbf{u}_{i_{m+1}} \in \mathbb{C}^n.$$

Note that the tensors $\mathbf{b}^k$ and $\mathbf{c}^k$ can be recursively computed using the equations

$$\mathbf{b}_{i_1 i_2 \ldots i_{k+1}}^{k+1} = \left(\sum_{k+1=1}^{p} \langle \mathbf{A}(\boldsymbol{\theta}) \rangle^{-1} \mathbf{A}^{i_{k+1}} \right) \mathbf{b}_{i_1 i_2 \ldots i_k}^k$$

$$\mathbf{c}_{i_1 i_2 \ldots i_{k+2}}^{k+1} = \left(\sum_{k+2=1}^{p} \langle \mathbf{A}(\boldsymbol{\theta}) \rangle^{-1} \mathbf{A}^{i_{k+2}} \right) \mathbf{c}_{i_1 i_2 \ldots i_{k+1}}^k$$

The terms of the Neumann series (i.e., our basis vectors) are in general not orthogonal. However, the stochastic version of Arnoldi's method can be readily applied to orthogonalize the basis vectors. Note that the above explicit representation of the preconditioned stochastic Krylov subspace is crucial for efficiently computing the $\mathcal{L}_2$ norms in the stochastic version of Arnoldi's method presented earlier.

A Note on Computational Aspects of Basis Vector Computation

If the decomposed form of the matrix $\langle \mathbf{A}(\boldsymbol{\theta}) \rangle$ is available, then the basis vectors can be readily computed using forward and backward substitutions. However, note that, increasing the order of the preconditioned stochastic Krylov subspace inevitably leads to an exponential increase ($\mathcal{O}(n^k)$) in storage complexity. This is true when the higher-order tensors are computed to arrive at an explicit expression for the response process. Storing the higher-order tensors in memory may be impossible for systems with large number of dof and random variables. Fortunately, if only the first and second-order statistical moments of the solution are of interest, then the storage complexity issue is of no concern in software implementation of SRBMs. This is because only the statistics of the basis vectors are required to be stored.

One way to reduce the memory requirements and computational cost would be to neglect the interaction terms in the basis vectors of order greater than 2. For example, the fourth basis vector may be rewritten as

$$\psi_3(\boldsymbol{\theta}) = \sum_{i=1}^{p} \theta_i^3 \left(\mathbf{b}_i^3 + \sum_{j=1}^{q} \eta_j \mathbf{c}_{ij}^3 \right),$$

where

$$\mathbf{b}_i^3 = \langle \mathbf{A}(\boldsymbol{\theta}) \rangle^{-1} \mathbf{A}^i \langle \mathbf{A}(\boldsymbol{\theta}) \rangle^{-1} \mathbf{A}^i \langle \mathbf{A}(\boldsymbol{\theta}) \rangle^{-1} \mathbf{A}^i \langle \mathbf{A}(\boldsymbol{\theta}) \rangle^{-1} \mathbf{f}^0,$$

$$\mathbf{c}_{ij}^3 = \langle \mathbf{A}(\boldsymbol{\theta}) \rangle^{-1} \mathbf{A}^i \langle \mathbf{A}(\boldsymbol{\theta}) \rangle^{-1} \mathbf{A}^i \langle \mathbf{A}(\boldsymbol{\theta}) \rangle^{-1} \mathbf{A}^i \langle \mathbf{A}(\boldsymbol{\theta}) \rangle^{-1} \mathbf{f}^j$$

Using the simplified basis vectors of the form shown above, the memory requirements as well the computational complexity of the stochastic subspace

projection schemes to be presented later can be significantly reduced. An alternative approach would be to use a block version of SRBMs which utilizes a set of the basis vectors $\psi_0(\boldsymbol{\theta})$ and $\psi_1(\boldsymbol{\theta})$ evaluated at a number of points in the parameter space. These ideas are clearly worth pursuing and the accuracy of the resulting approximations remains to be tested.

Bubnov-Galerkin (BG) Scheme for Stochastic Problems

In previous sections, we examined the theoretical and computational aspects of the preconditioned stochastic Krylov subspace. This section is concerned with stochastic subspace projection schemes for computing the undetermined coefficients in the reduced basis representation. In particular, our focus is on the Bubnov-Galerkin scheme. This projection scheme involves formulating a stochastic residual error of the form

$$\mathbf{r}(\boldsymbol{\theta},\boldsymbol{\eta}) = \mathbf{A}(\boldsymbol{\theta})\Psi(\boldsymbol{\theta},\boldsymbol{\eta})\Xi(\boldsymbol{\theta},\boldsymbol{\eta}) - \mathbf{f}(\boldsymbol{\eta}) \tag{18}$$

When the Bubnov-Galerkin scheme is applied to deterministic problems, the undetermined coefficients are determined by enforcing the condition $\mathbf{r}(\boldsymbol{\theta},\boldsymbol{\eta}) \perp \Psi(\boldsymbol{\theta},\boldsymbol{\eta})$. Hence, this scheme is also referred to as an orthogonal projection scheme. An alternative scheme involving the use of oblique projection is outlined in Appendix A. Using the orthogonal projection scheme, the undetermined coefficient vector $\Xi(\boldsymbol{\theta},\boldsymbol{\eta})$ should be computed by solving the reduced-order system of equations

$$\tilde{\mathbf{A}}(\boldsymbol{\theta},\boldsymbol{\eta})\Xi(\boldsymbol{\theta},\boldsymbol{\eta}) = \tilde{\mathbf{f}}(\boldsymbol{\theta},\boldsymbol{\eta}) \tag{19}$$

where $\tilde{\mathbf{A}}(\boldsymbol{\theta},\boldsymbol{\eta}) = \Psi^*(\boldsymbol{\theta},\boldsymbol{\eta})\mathbf{A}(\boldsymbol{\theta})\Psi(\boldsymbol{\theta},\boldsymbol{\eta}) \in \mathbb{C}^{m\times m}$ and $\tilde{\mathbf{f}}(\boldsymbol{\theta},\boldsymbol{\eta}) = \Psi^*(\boldsymbol{\theta},\boldsymbol{\eta})\mathbf{f}(\boldsymbol{\eta}) \in \mathbb{C}^m$ are the reduced random coefficient matrix and rhs, respectively.

If (19) is solved exactly for any realization of $\boldsymbol{\theta}$ and $\boldsymbol{\eta}$, then $P[\Psi^*(\boldsymbol{\theta},\boldsymbol{\eta})\mathbf{r}(\boldsymbol{\theta},\boldsymbol{\eta}) = 0] = 1$ (see Definition 2). However, as shown in reference[6], this involves the symbolic inversion of $\tilde{\mathbf{A}}(\boldsymbol{\theta},\boldsymbol{\eta})$ to explicitly represent the coefficients of the reduced basis as random polynomials. This is readily possible only when 2 or 3 basis vectors are used. A more practical fix to this problem is to employ simulation schemes for efficiently computing the statistics of the coefficients of the SRBR.

In the computational stochastic mechanics literature (see, for example, Ghanem and Spanos[1]), the undetermined coefficients in (19) are considered as deterministic scalars which are computed by solving the deterministic system of equations

$$\left\langle \tilde{\mathbf{A}}(\boldsymbol{\theta},\boldsymbol{\eta}) \right\rangle \Xi = \left\langle \tilde{\mathbf{f}}(\boldsymbol{\theta},\boldsymbol{\eta}) \right\rangle \tag{20}$$

The preceding equation is derived by using the measure of orthogonality defined in the Hilbert space of random variables (see Definition 2). Note that the deterministic matrix $\left\langle \tilde{\mathbf{A}}(\boldsymbol{\theta},\boldsymbol{\eta}) \right\rangle$ is a square block partition of the Hessenberg matrix which arises as a by-product of the stochastic version of Arnoldi's method.[‡]

As shown in reference[7], equation (19) gives better approximations and converges more rapidly (as the number of basis vectors are increased) as compared to (20). Our objective in this section is to study the relationship between equation (20) and the exact stochastic Bubnov-Galerkin scheme given in (19). We show that equation (20) is a zero-order approximation for $\Xi(\boldsymbol{\theta},\boldsymbol{\eta})$. An approximate stochastic Bubnov-Galerkin scheme is then presented to improve the accuracy of (20). Conditions under which the proposed scheme converges to the solution of equation (19) are postulated. Henceforth, we do not show the dependence of the random quantities on $\boldsymbol{\eta}$ explicitly unless it becomes important for the sake of clarity.

Convergence Analysis

Equation (19) can be rewritten as

$$\Xi(\boldsymbol{\theta}) = \left(\left\langle \tilde{\mathbf{A}}(\boldsymbol{\theta}) \right\rangle + \tilde{\mathbf{A}}(\boldsymbol{\theta}) - \left\langle \tilde{\mathbf{A}}(\boldsymbol{\theta}) \right\rangle \right)^{-1} \tilde{\mathbf{f}}(\boldsymbol{\theta}) \tag{21}$$

The Neumann expansion series for $\Xi(\boldsymbol{\theta})$ can be written as

$$\Xi(\boldsymbol{\theta}) = \sum_{i=0}^{\infty} (-1)^i \left(\left\langle \tilde{\mathbf{A}}(\boldsymbol{\theta}) \right\rangle^{-1} \tilde{\mathbf{A}}(\boldsymbol{\theta}) - \mathbf{I}_m \right)^i \left\langle \tilde{\mathbf{A}}(\boldsymbol{\theta}) \right\rangle^{-1} \tilde{\mathbf{f}}(\boldsymbol{\theta}) \tag{22}$$

where $\mathbf{I}_m \in \mathbb{R}^{m\times m}$ is the identity matrix.

Setting $i = 0$ and assuming that $\tilde{\mathbf{f}}(\boldsymbol{\theta})$ can be replaced by its ensemble average, it can be clearly seen that equation (22) becomes identical to (20). This suggests that the stochastic Bubnov-Galerkin scheme used in the literature is a zero-order approximation. Henceforth, we will use the compact notation BG_j to denote BG approximation of order j.

Interestingly, equation (22) suggests how the zero-order approximation available in the literature can

[‡]The reader is referred to the text of Saad[9] for a detailed exposition of this point

be improved. In particular, we can improve the zero-order approximation by rewriting (20) as

$$\Xi(\boldsymbol{\theta}) = \left\langle \tilde{\mathbf{A}}(\boldsymbol{\theta}) \right\rangle^{-1} \tilde{\mathbf{f}}(\boldsymbol{\theta}) \tag{23}$$

Equation (23) may provide a better approximation as compared to (20) when the coefficient of variation of $\tilde{\mathbf{f}}(\boldsymbol{\theta})$ is large. Further, the undetermined coefficients in the reduced basis will be random scalars when (23) is used instead of (20).

The conditions under which equation (22) converges for a given realization of $\boldsymbol{\theta}$ are stated below.

Lemma 1: Equation (26) will converge if $\rho(\langle\tilde{\mathbf{A}}(\boldsymbol{\theta})\rangle^{-1}\tilde{\mathbf{A}}(\boldsymbol{\theta}) - \mathbf{I}_m) < 1$ or $||\langle\tilde{\mathbf{A}}(\boldsymbol{\theta})\rangle^{-1}\tilde{\mathbf{A}}(\boldsymbol{\theta}) - \mathbf{I}_m|| < 1$; where $\rho(\mathbf{B})$ denotes the spectral radius of the matrix $\mathbf{B}$.

Remark: In practice, it may be difficult to confirm when the condition imposed by Lemma 1 is satisfied.

A Posteriori Error Estimation

Using the procedures outlined in the previous section, the basis vectors and the undetermined coefficients in the reduced basis can be efficiently computed. This ultimately leads to an approximation for the solution vector of the form $\hat{\mathbf{x}}(\boldsymbol{\theta}) = \Psi(\boldsymbol{\theta})\Xi$. This section examines the issue of computing *a posteriori* error estimates for the reduced basis approximation $\hat{\mathbf{x}}(\boldsymbol{\theta})$. In contrast to stochastic systems, error estimates for deterministic systems are readily available in the literature; see, for example, Golub and Van Loan[11].

Two commonly used error norms will be considered in this paper - (1) the residual vector $\mathbf{r}(\boldsymbol{\theta}) = \mathbf{A}(\boldsymbol{\theta})\hat{\mathbf{x}}(\boldsymbol{\theta}) - \mathbf{f}(\boldsymbol{\theta})$ and the error vector $\Delta\mathbf{x}(\boldsymbol{\theta}) = \mathbf{x}(\boldsymbol{\theta}) - \hat{\mathbf{x}}(\boldsymbol{\theta})$.

In contrast to deterministic problems, both error estimates will be random quantities. Hence, error estimation entails computing the statistics of $\mathbf{r}(\boldsymbol{\theta})$ and $\Delta\mathbf{x}(\boldsymbol{\theta})$. Error estimates are derived here for the case when BG_0 (see (20)) is used to compute the coefficients in the SRBR. The approximate solution is hence available in the form of a vector of random polynomials. The mean of $\mathbf{r}(\boldsymbol{\theta})$ can then be readily computed as

$$\langle\mathbf{r}(\boldsymbol{\theta})\rangle = \langle\mathbf{A}(\boldsymbol{\theta})\Psi(\boldsymbol{\theta})\rangle\Xi - \langle\mathbf{f}(\boldsymbol{\theta})\rangle \tag{24}$$

$||\langle\mathbf{r}(\boldsymbol{\theta})\rangle||_2$ can be used to determine whether additional basis vectors must be employed to improve the accuracy of the solution. A note of caution is advisable here. The residual error only indicates the extent to which the governing equation is satisfied. This may not necessarily be the same as finding an approximate solution which is close to the exact one. Small residual need not indicate small value of $\Delta\mathbf{x}(\boldsymbol{\theta})$ or *vice-versa.*

A more useful error estimate would be an estimate of $\Delta\mathbf{x}(\boldsymbol{\theta})$, which can be computed by solving the stochastic equations

$$\mathbf{A}(\boldsymbol{\theta})\Delta\mathbf{x}(\boldsymbol{\theta}) = \mathbf{r}(\boldsymbol{\theta}) \tag{25}$$

It can be clearly seen that equation (25) cannot be solved exactly in a computationally efficient fashion. This not surprising - if we can can compute the error exactly, we can also solve the governing equations exactly ! Hence, we have to settle for an approximation to the error estimate. This practice is quite common in numerical analysis wherein first-order estimates are often used. Clearly, the best way out would be to compute an upper bound on the error estimate. On the lines of the standard error analysis procedures available in the literature[11], it can be readily shown that the following stochastic inequality is valid

$$||\Delta\mathbf{x}(\boldsymbol{\theta})|| \leq ||\mathbf{A}(\boldsymbol{\theta})^{-1}||||\mathbf{r}(\boldsymbol{\theta})|| \tag{26}$$

This stochastic inequality is intractable, and would require an approximation. For example, an estimate of the upper bound of the tail of the random right hand side could be used. It is felt that the theory of *stochastic majorization* [15] could be useful in this context. An alternative approach we are currently pursuing is to extend some fundamental results on the pdf of the $\mathcal{L}_2$ norm of the inverse of random matrices; see, for example, Edelman[16], where it is assumed that the elements of the random matrix are uncorrelated and Gaussian. It would be useful if this result can be extended to establish rigorous upper bounds on $\Delta\mathbf{x}(\boldsymbol{\theta})$, which hold with a pre-specified probability.

From a practical viewpoint, it is important to first examine the context in which an error estimate might be used. In SRBM, error estimates may be used to decide whether more basis vectors should be used to improve the accuracy. Furthermore, the error estimates may be applied to the task of computing error bounds on the statistical moments of the response process. Let us examine the case when it is aimed to approximate the stochastic equality in equation (25) rather than an complicated inequality.

A zero-order estimate of $\Delta\mathbf{x}(\boldsymbol{\theta})$ can be computed as

$$\Delta\mathbf{x}(\boldsymbol{\theta}) \approx \langle\mathbf{A}(\boldsymbol{\theta})\rangle^{-1}\mathbf{r}(\boldsymbol{\theta}) \tag{27}$$

From this approximate equality, we can directly derive error estimates for the mean and covariance of

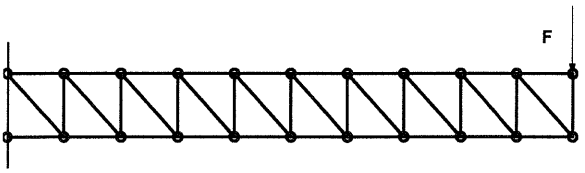

Figure 1: 40 Member Frame Structure with Random Youngs Modulus

the solution computed by SRBM. For example, the error in the mean is approximately $\langle \Delta \mathbf{x} \rangle$. The error in the covariance matrix Δ_{cov} can be written as

$$\Delta_{cov} \approx \langle \Delta \mathbf{x}(\boldsymbol{\theta}) \Delta \mathbf{x}^*(\boldsymbol{\theta}) \rangle + 2 \langle \Delta \mathbf{x}(\boldsymbol{\theta}) \hat{\mathbf{x}}^*(\boldsymbol{\theta}) \rangle \quad (28)$$

Note that it is not a trivial exercise to establish the probability that the estimates in (27) and (28) are correct. However, numerical experience suggests that they are correct with high (albeit unquantifiable) probability. Note that the order of the estimate can be readily increased. From the computational viewpoint, the expectation operations can be easily carried out since $\hat{\mathbf{x}}(\boldsymbol{\theta})$ and $\mathbf{r}(\boldsymbol{\theta})$ are available as a vector of random polynomials.

Some preliminary studies are presented to examine the convergence of the SRBM when the preconditioned stochastic Krylov subspace of increasing order is employed in conjunction with BG_0. The example problem considered (see Figure 1) involves a network of Euler-Bernoulli beams with random Youngs modulus subjected to a deterministic force. This problem has a total of 180 degrees of freedom and 40 random system parameters. The nominal material properties for this system have been taken from reference.[7]

Two cases are taken up to study the convergence rate of SRBM. In case 1 and 2, the standard deviation of the Youngs modulus of each element are taken as 0.05 and 0.10, respectively. Figure 2 shows the convergence of the $\mathcal{L}_2$ norm of the mean residual vector as the order of the Krylov subspace is increased from 2 to 5. It can be clearly seen that the mean residual error decreases rapidly as the number of basis vectors are increased. This trend also clearly suggests that it may be rather difficult to judge the accuracy of the approximated solution by examining only the residual error statistics. For example, in Case 1, when 2 basis vectors are used, the maximum error in the mean and standard deviation is

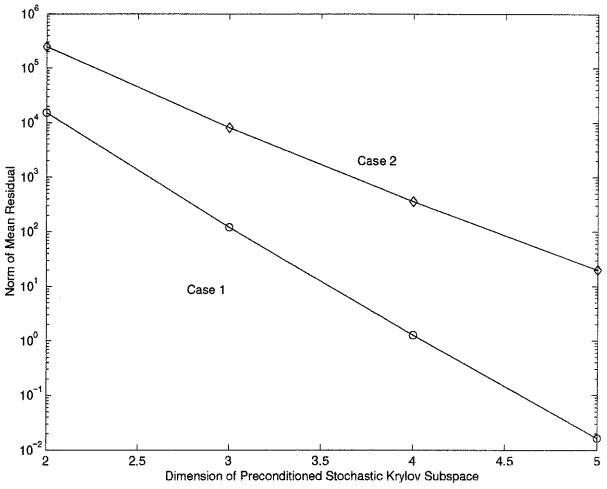

Figure 2: Convergence of Norm of Mean Residual Error as a Function of Dimension of Preconditioned Stochastic Krylov Subspace

Table 1: Comparison of Accuracy of the Zero-order Error Estimate

Order	*Case 1*		*Case 2*	
	Appr.	**Exact**	**Appr.**	**Exact**
1	8.0e-05	1.3e-05	1.0e-04	2.2e-04
2	9.8e-07	5.6e-07	1.9e-05	1.8e-05

just of the order of 1%. However, the corresponding mean residual error norm is of the order of 10^4.

Next, some studies were conducted to study the accuracy of the zero-order error estimate for $\Delta \mathbf{x}$. The following measure was used as an indicator of the accuracy of the approximate solution computed using SRBM

$$\delta = ||\langle \Delta \mathbf{x}(\boldsymbol{\theta}) \rangle||_\infty / ||\langle \hat{x}(\boldsymbol{\theta}) + \Delta \mathbf{x}(\boldsymbol{\theta}) \rangle||_\infty$$

The accuracy of the zero-order error estimate is compared with a Monte Carlo simulation scheme (sample size 20,000) which is employed to compute the preceding error indicator exactly. The results for case 1 and 2 are summarized in Table 1. The results indicate that the zero-order estimate is an 'order of magnitude correct', and hence it can be applied to decide whether more basis vectors should be included to improve the accuracy of the approximation.

Coupling SRBMs with Monte Carlo Simulation

In the earlier section, we examined issues involved in *a posteriori* error estimation. In this section, we

outline a simple procedure for coupling SRBMs with the Monte Carlo simulation scheme. This procedure is motived by the hybrid stochastic finite element approach presented earlier by Ghanem.[17] To motivate the hybrid procedure, consider the expression for the covariance matrix of the "true" response in terms of the approximate covariance matrix computed using SRBMs and the error vector $\Delta\mathbf{x}(\boldsymbol{\theta})$

$$Cov^{true} = Cov^{SRBM} + \langle \Delta\mathbf{x}(\boldsymbol{\theta})\Delta\mathbf{x}^*(\boldsymbol{\theta})\rangle + 2\langle \Delta\mathbf{x}(\boldsymbol{\theta})\hat{\mathbf{x}}^*(\boldsymbol{\theta})\rangle \tag{29}$$

From previous experience, it is known that SRBMs give reasonably accurate approximations for the response covariance matrix. Hence, the contribution of the terms involving $\Delta\mathbf{x}(\boldsymbol{\theta})$ is expected to be small. This naturally motivates the use of a simulation scheme to estimate this term, since a few number of samples will typically be required to compute the statistics of a random variable with small variance. In summary, on the lines of Ghanem[17], a hybrid procedure coupling SRBMs with the Monte Carlo simulation scheme can be developed.

Algebraic Random Eigenvalue Problems

In this section, we briefly examine the application of SRBMs to symmetric algebraic random eigenvalue problems of the form

$$\mathbf{K}(\boldsymbol{\theta})\mathbf{x}_i(\boldsymbol{\theta}) = \lambda_i(\boldsymbol{\theta})\mathbf{M}(\boldsymbol{\theta})\mathbf{x}_i(\boldsymbol{\theta}), \tag{30}$$

where $\mathbf{x}_i(\boldsymbol{\theta}) \in \mathbb{C}^n$ and $\lambda_i(\boldsymbol{\theta}) \in \mathbb{C}$ denotes the eigenvector and eigenvalue and mode i, respectively. Further, the random parameterization of the coefficient matrices are assumed to be of the form

$$\mathbf{K}(\boldsymbol{\theta}) = \langle \mathbf{K}(\boldsymbol{\theta})\rangle + \theta_1\mathbf{K}_1 + \theta_2\mathbf{K}_2 + \ldots + \theta_p\mathbf{K}_p,$$

$$\mathbf{M}(\boldsymbol{\theta}) = \langle \mathbf{M}(\boldsymbol{\theta})\rangle + \theta_1\mathbf{M}_1 + \theta_2\mathbf{M}_2 + \ldots + \theta_p\mathbf{M}_p,$$

where $\mathbf{K}_i, \mathbf{M}_i \in \mathbb{R}^{n\times n}$ are deterministic symmetric matrices, θ_i are assumed to be random variables or orthogonal random functions.

Recently, Nair and Keane[19] proposed a stochastic reduced basis method for approximating the random eigenvalues and eigenvectors of such randomly parameterized eigenvalue problems. This method involves the use of two independent basis vectors to approximate each random eigenvalue and eigenvector of interest. The eigenparameter statistics were then computed by solving a sequence of reduced-order 2×2 random eigenproblem for each mode of interest. Encouraging results were presented for free-vibration and frequency response analysis of a network of stochastic Euler-Bernoulli beams.

More recent studies by the author indicate that the accuracy of this approach may deteriorate when the structural system has high statistical overlap factor.[§] Interestingly, the eigenvalue statistics turn out to be more difficult to approximate as compared to that of the eigenvectors. This difficulty primarily arises due to mode switching phenomena typical of structures with high modal density.

These observations suggest that that formulations which use a global set of stochastic basis vectors may lead to better approximations for the eigenvalue statistics. The fundamental idea here is to approximate the first m eigenvectors of Eqn. (30) in a subspace spanned by a set of stochastic basis vectors, i.e.,

$$\hat{\mathbf{X}}(\boldsymbol{\theta}) = \boldsymbol{\Psi}(\boldsymbol{\theta})\mathbf{C},$$

where $\boldsymbol{\Psi}(\boldsymbol{\theta}) = [\psi_1(\boldsymbol{\theta}), \psi_2(\boldsymbol{\theta}), \ldots \psi_m(\boldsymbol{\theta})] \in \mathbb{C}^{n\times m}$ denotes the matrix of m stochastic basis vectors, and $\mathbf{C} = [\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_m] \in \mathbb{C}^{m\times m}$ is a matrix of undetermined coefficients, where $\mathbf{c}_i \in \mathbb{C}^m$.

The matrix of undetermined coefficients can be computed by imposing the Bubnov-Galerkin condition

$$\left(\mathbf{K}(\boldsymbol{\theta}) - \lambda(\boldsymbol{\theta})\mathbf{M}(\boldsymbol{\theta})\right)\boldsymbol{\Psi}(\boldsymbol{\theta})\mathbf{C} \perp \boldsymbol{\Psi}(\boldsymbol{\theta}).$$

Using the definition of inner product in the Hilbert space of second-order random variables, i.e., $(\mathbf{x}_1, \mathbf{x}_2) = \langle \mathbf{x}_1^*\mathbf{x}_2\rangle$, we arrive at a reduced-order deterministic generalized eigenvalue problem

$$\mathbf{K}_R\mathbf{C} = \mathbf{M}_R\mathbf{C}\boldsymbol{\Omega}, \tag{31}$$

where $\mathbf{K}_R = \langle \boldsymbol{\Psi}^*(\boldsymbol{\theta})\mathbf{K}(\boldsymbol{\theta})\boldsymbol{\Psi}(\boldsymbol{\theta})\rangle$, $\mathbf{M}_R = \langle \boldsymbol{\Psi}^*(\boldsymbol{\theta})\mathbf{M}(\boldsymbol{\theta})\boldsymbol{\Psi}(\boldsymbol{\theta})\rangle \in \mathbb{C}^{m\times m}$ are reduced-order matrices, and $\boldsymbol{\Omega} \in \mathbb{C}^{m\times m}$ denotes the diagonal matrix of eigenvalues of Eqn. (31).

An important problem which arises here is how to choose a set of stochastic basis vectors. One approach which we are currently pursuing uses the preconditioned stochastic subspace

$$\Psi(\boldsymbol{\theta}) = \langle \mathbf{K}(\boldsymbol{\theta})\rangle^{-1}\left(\mathbf{M}(\boldsymbol{\theta})\mathbf{X}_o\Lambda^o - \sum_{j=1}^{p}\theta_j\mathbf{K}_j\mathbf{X}_o\right).$$

where $\mathbf{X}_o \in \mathbb{R}^{n\times m}$ and $\Lambda^o \in \mathbb{R}^{n\times m}$ are the matrices of eigenvectors and eigenvalues of the deterministic eigenvalue problem $\langle \mathbf{K}(\boldsymbol{\theta})\rangle\mathbf{x}_i^o = \lambda_i^o\langle \mathbf{M}(\boldsymbol{\theta})\rangle\mathbf{x}_i^o$.

An attractive feature of this subspace is that the basis vectors turn out to be linear functions of the random variables appearing in the parameterized coefficient matrices, i.e., the matrix of basis vectors can

[§]The statistical modal overlap factor can be defined as $\sigma_i/\langle\lambda_{i+1} - \lambda_i\rangle$, where σ_i denotes the standard deviation of the ith eigenvalue.

be written in block form as

$$\Psi(\boldsymbol{\theta}) = \mathbf{X}_o + \sum_{j=1}^{p} \theta_j \boldsymbol{\Delta}_j,$$

where $\boldsymbol{\Delta}_j \in \mathbb{C}^{n\times n}$ is a deterministic matrix which can be written as

$$\boldsymbol{\Delta}_j = \langle \mathbf{K}(\boldsymbol{\theta}) \rangle^{-1} \left(\mathbf{M}_j \mathbf{X}_o \Lambda^o - \mathbf{K}_j \mathbf{X}_o \right).$$

It can be seen that it will require modest memory $\mathcal{O}(mnp)$ to store the basis vectors in explicit form. Further, the elements of the reduced-order matrices in Eqn. (31) can be efficiently computed for this choice of basis vectors.

If the basis vectors are rich (i.e., the first m eigenvectors of Eqn. (30) has significant components along it), then the solution of Eqn. (31) leads to a good approximation for the stochastic invariant subspace $\hat{\mathbf{X}}(\boldsymbol{\theta}) = \boldsymbol{\Psi}(\boldsymbol{\theta})\mathbf{C}$. We can hence write approximations for the first m random eigenvalues of Eqn. (30) as

$$\hat{\lambda}_i(\boldsymbol{\theta}) = \frac{\mathbf{c}_i^* \boldsymbol{\Psi}^*(\boldsymbol{\theta}) \mathbf{K}(\boldsymbol{\theta}) \boldsymbol{\Psi}(\boldsymbol{\theta}) \mathbf{c}_i}{\mathbf{c}_i^* \boldsymbol{\Psi}^*(\boldsymbol{\theta}) \mathbf{M}(\boldsymbol{\theta}) \boldsymbol{\Psi}(\boldsymbol{\theta}) \mathbf{c}_i}, \quad i = 1, 2, \ldots m,$$

where $\mathbf{c}_i \in \mathbb{C}^m$ denotes the eigenvector of Eqn. (31) corresponding to its ith eigenvalue.

Similarly, the statistics of the eigenvectors can be computed, which may then be used to compute the forced response in the time or frequency domain. Studies which are currently underway[20] suggest that such a line of approach may hold the key towards developing a general purpose capability for analysis of linear stochastic structural dynamical systems.

Concluding Remarks

In this paper, we have outlined some ongoing research efforts towards developing a theoretical foundation for stochastic reduced basis methods introduced in references.[6,7] The present research is expected to have important ramifications from the theoretical as well as practical point of view. From the theoretical perspective, this paper establishes the connection between Krylov methods for deterministic systems and stochastic reduced basis methods. It was shown that such a line of approach may greatly aid the design and analysis of new algorithms for analysis of stochastic systems.

From the practical perspective, we have shown in this paper that error estimates can be efficiently computed for SRBMs. Further, the stochastic version of Arnoldi's method presented here can be used to improve the numerical stability of the computations. The error estimates developed in this work can be employed to adaptively select the appropriate number of basis vectors required to achieve a desired level of accuracy in the solution. The objective behind this is to provide the analyst with error estimates of the statistical moments, so that the Monte Carlo simulation scheme does not have to be run to judge the accuracy of the results. A simple procedure for coupling simulation schemes with SRBMs has also been outlined.

We have discussed ongoing work focused on extending SRBMs to algebraic random eigenvalue problems. A preconditioned stochastic subspace was presented for approximating the stochastic invariant subspace. The potential advantages offered by this line of approach was briefly outlined.

To summarize, this paper outlines some avenues for further research into physics-based approaches for stochastic system analysis. It appears that much remains to be done before this area reaches the maturity level of its deterministic counterpart. Since the emphasis of the present work is on establishing the theoretical foundations of SRBMs, we have chosen to omit a full-scale numerical validation of the ideas developed here. A future work will present such a systematic computational study to assess the performance of the formulations presented here on realistic test problems. Finally, it is hoped that the understanding gained from this program of research will contribute to a new generation of tools for computational stochastic mechanics.

Acknowledgments

This research was supported by a grant from the Faculty of Engineering and Applied Sciences at the University of Southampton.

References

[1]Ghanem, R., and Spanos, P., *Stochastic Finite Elements : A Spectral Approach*, Springer-Verlag, 1991.

[2]Kleiber, M., and Hien, T. D., *The Stochastic Finite Element Method: Basic Perturbation Technique and Computer Implementation*, Chichester: John Wiley, 1992.

[3]Ghanem, R., "Ingredients for a General Purpose Stochastic Finite Elements Implementation," *Computer Methods in Applied Mechanics and Engineering*, Vol. 168, No.1-4, 1999, pp. 19-34.

[4]Elishakoff, I., Ren, Y. J., Shinozuka, M., " Variational Principles Developed for and Applied to Analysis of Stochastic Beams, " *Journal of Engineering Mechanics*, ASCE, Vol.122, No.6, 1996, pp.559-565.

[5]Elishakoff, I., and Ren, Y. J., "The Bird's Eye View on Finite Element Method for Structures with Large Stochastic Variations," *Computer Methods in Applied Mechanics and Engineering*, Vol. 168, No.1-4, 1999, pp. 51-61.

[6]Nair, P. B., and Keane, A. J., "New Developments in Computational Stochastic Mechanics, Part I: Theory," *Proceedings of the AIAA Non-Deterministic Approaches Forum*, Atlanta, GA, 2000, AIAA-2000-1827.

[7]Nair, P. B., and Keane, A. J., "New Developments in Computational Stochastic Mechanics, Part II: Applications," *Proceedings of the AIAA Non-Deterministic Approaches Forum*, Atlanta, GA, 2000, AIAA-2000-1441.

[8]Saad, Y., and Van der Vorst, H. A., "Iterative Solution of Linear Systems in the 20-th Century, " *Journal of Computational and Applied Mathematics*, 2000, to appear.

[9]Saad, Y., *Iterative Methods for Sparse Linear Systems*, PWS Publishing Company, Boston, 1996 and Second Edition manuscript, 2000.

[10]Axelsson, O., *Iterative Solution Methods*, Cambridge University Press, Cambridge, 1994.

[11]Golub, G., and Van Loan, *Matrix Computations*, John Hopkins Press, Third Edition, 1996.

[12]Fokkema, D. R., Sleijpen, L. G., and Van Der Vorst, H. A., "Accelerated Inexact Newton Schemes for Large Systems of Nonlinear Equations," *SIAM Journal of Scientific Computing*, Vol. 19, No. 2, 1998, pp. 657-674.

[13]Ipsen, I. C. F., and Meyer, C. D., "The Idea Behind Krylov Methods," *American Mathematical Monthly*, Vol. 105, No. 10, 1998, pp. 889-899.

[14]Arnoldi, W. E., "The Principle of Minimized Iterations in the Solution of the Matrix Eigenvalue Problem," *Quarterly Journal of Applied Mathematics*, Vol. 9, 1951, pp. 17-29..

[15]Marshall, A. W., and Olkin, I., *Inequalities: Theory of Majorization and Its Applications*, Academic Press, 1979.

[16]Edelman, A., "Eigenvalues and Condition Numbers of Random Matrices," *SIAM Journal on Matrix Analysis and Applications*, Vol. 9, 1988, pp. 543-560.

[17]Ghanem, R., "Hybrid Stochastic Finite Elements and Generalized Monte Carlo Simulation, *ASME Journal of Applied Mechanics*, Vol. 65, No. 4, 1998, pp. 1004-1009.

[18]Nair, P. B., Choudhury, A., and Keane, A. J., "A Bayesian Framework for Uncertainty Analysis Using Deterministic Black-Box Simulation Codes," *Proceedings of the 42nd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*, Seattle, WA, 2001. AIAA Paper No. 2001-1676.

[19]Nair, P. B., and Keane, A. J., "Reduced Basis Representation of Large-Scale Random Eigenvalue Problems," *Proceedings of the 41st AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Atlanta, GA, 2000, AIAA-2000-1828.

[20]Nair, P. B., "Preconditioned Stochastic Subspace Projection Schemes for Dynamic Analysis of Uncertain Systems," *in preparation.*

Appendix A

An alternative approach for computing the undetermined coefficients in the SRBR is to directly minimize the $\mathcal{L}_2$ norm of the residual error vector $\mathbf{r}(\boldsymbol{\theta})$, which essentially involving enforcing the condition.

$$\mathbf{r}(\boldsymbol{\theta},\boldsymbol{\eta}) \perp \mathbf{A}(\boldsymbol{\theta})\Psi(\boldsymbol{\theta},\boldsymbol{\eta}). \tag{A1}$$

Hence, $\Xi(\boldsymbol{\theta},\boldsymbol{\eta})$ can be computed as

$$\Xi(\boldsymbol{\theta},\boldsymbol{\eta}) = \mathbf{H}(\boldsymbol{\theta},\boldsymbol{\eta})^{+}\tilde{\mathbf{f}}(\boldsymbol{\theta},\boldsymbol{\eta}) \tag{A2}$$

where

$$\mathbf{H}(\boldsymbol{\theta},\boldsymbol{\eta})^{+} = (\Psi^{*}(\boldsymbol{\theta},\boldsymbol{\eta})\mathbf{A}^{*}(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})\Psi(\boldsymbol{\theta},\boldsymbol{\eta}))^{-1} \tag{A3}$$

and

$$\tilde{\mathbf{f}}(\boldsymbol{\theta},\boldsymbol{\eta}) = \Psi^{*}(\boldsymbol{\theta},\boldsymbol{\eta})\mathbf{A}^{*}(\boldsymbol{\theta})\mathbf{f}(\boldsymbol{\eta}) \tag{A4}$$

Using Definition 1, the undetermined coefficients can be computed as

$$\Xi = \langle \mathbf{H}(\boldsymbol{\theta},\boldsymbol{\eta})\rangle^{+}\langle\tilde{\mathbf{f}}(\boldsymbol{\theta},\boldsymbol{\eta})\rangle \tag{A5}$$

It can be readily seen from the above equation that computing $\langle \mathbf{H}(\boldsymbol{\theta},\boldsymbol{\eta})\rangle$ is much more involved as compared to the projected matrix in the Bubnov-Galerkin scheme. However, comparing the accuracy of the residual minimization technique *vis-a-vis* the BG scheme would be an interesting exercise. The interested reader is referred to Saad[9] for a description of the generalized minimal residual method (GMRES) and the quasi-minimal residual method (QMR) for deterministic linear systems.