Generalized Polynomial Chaos Solution for Differential Equations with Random Inputs [1]

G.E. Karniadakis[2], C.-H. Su[3], D. Xiu[4], D. Lucor[5],
C. Schwab and R.A. Todor

Research Report No. 2005-01
January 2005
Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

[1] This manuscript is under review in SIAM Review. The work of Karniadakis, Su, Xiu and Lucor was supported by DOE, AFOSR, ONR and NSF. The work of Schwab and Todor was supported by IHP network "Breaking Complexity" of the EC under contract number HPRN-CT-2002-00286, and by the Swiss Federal Office for Science and Education under grant BBW02.0418

[2] Corresponding author, Division of Applied Mathematics, Brown University, Providence, RI 02912, USA, (email:gk@dam.brown.edu)

[3] Division of Applied Mathematics, Brown University, Providence, RI 02912, USA

[4] Department of Chemical Engineering, Princeton University, Princeton, NJ 08544, USA

[5] Department of Ocean Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Generalized Polynomial Chaos Solution for Differential Equations with Random Inputs [1]

G.E. Karniadakis[2], C.-H. Su[3], D. Xiu[4], D. Lucor[5],
C. Schwab and R.A. Todor
Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

Research Report No. 2005-01 January 2005

Abstract

Stochastic modeling by differential equations with random inputs is reviewed. Recent developments in deterministic solution algorithms based on generalizations of the homogeneous Hermite Polynomial Chaos (PC) expansions of N. Wiener and on sparse tensor product approximations of spatial correlation functions in perturbation expansions are presented. Mathematical issues in the formulation, PC discretization, and numerical analysis of PC solution algorithms are discussed.

Keywords: Polynomial chaos, orthogonal polynomials, random inputs, differential equations, noisy systems, uncertainty quantification

Subject Classification: 65C20, 65C30, 65N30

[1] This manuscript is under review in SIAM Review. The work of Karniadakis, Su, Xiu and Lucor was supported by DOE, AFOSR, ONR and NSF. The work of Schwab and Todor was supported by IHP network "Breaking Complexity" of the EC under contract number HPRN-CT-2002-00286, and by the Swiss Federal Office for Science and Education under grant BBW02.0418

[2] Corresponding author, Division of Applied Mathematics, Brown University, Providence, RI 02912, USA, (email:gk@dam.brown.edu)

[3] Division of Applied Mathematics, Brown University, Providence, RI 02912, USA

[4] Department of Chemical Engineering, Princeton University, Princeton, NJ 08544, USA

[5] Department of Ocean Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

1. Introduction.

1.1. The need for stochastic simulations. This paper addresses numerical solutions of differential equations with random inputs. The source of random inputs includes uncertainty in system parameters, boundary and initial conditions, material properties, source and interaction terms, geometry, etc. Such types of uncertainty are ubiquitous in engineering applications, and are often modeled as random fields. The classical approach is to model the random inputs as functionals of idealized processes such as Wiener processes, Poisson processes, etc. This approach has led to elegant mathematical analysis and corresponding numerical methods by using stochastic calculi, as e.g., Ito or Stratonovich calculus for ordinary differential equations and the stochastic calculus of variations of Malliavin for partial differential equations.

The resulting differential equations are broadly termed "stochastic ordinary/partial differential equations"; see, for example, [15, 21, 24, 41]. Recently, more research efforts have been devoted to studying problems with more realistic random inputs, e.g., correlated random processes ("colored noise" in the language of physics) and, in case of random parameters, fully correlated random processes, i.e., *random variables*. In this context, the classical stochastic calculi do not readily apply and hence other approaches are required. This class of problems is the focus of this paper.

Efficient numerical solution of differential equations with stochastic inputs has direct impact on simulations of physical and biological processes in at least the following three areas:

1. Uncertainty quantification,
2. Stability of noisy systems,
3. Coarse-grained and multiscale representation.

In the following, we will elaborate on each of these areas.

There has been an increasing interest recently in studying *uncertainty quantification* in large-scale numerical simulations to assess modeling uncertainty, see [39, 1, 17]. In deterministic numerical simulations we can now routinely monitor accuracy of the computed results by *a posteriori* error bounds and thus assess discretization errors. In contrast, *a posteriori* estimation of *modeling errors*, due to adoption of simplified working models as basis for numerical simulation, is still at an early stage of development. Often there are discrepancies between computational results and experimental observations, and errors from any stage of the numerical simulations can contribute to such discrepancies. However, if we assume that the discretization and modeling errors are appropriately controlled (which can be achieved for many problems of engineering interest), then the validity of the adopted mathematical model, e.g., a constitutive law, becomes the leading cause of the discrepancies. The validity of the underlying mathematical model can be established *only if* uncertainty in numerical predictions due to uncertain input data (e.g., transport and material coefficients, source and interaction terms, geometric irregularities as surface roughness) can be quantified.

Uncertainty quantification requires the propagation of uncertainty through a given mathematical model and affects all stages of numerical simulation. Specifically, partial differential equations widely used to formulate mathematical models of physical systems must be reformulated as stochastic partial differential equations. This poses new challenges for mathematics (pure and applied alike): parsimonious parametric description of stochastic input data and model calibration requires methods from statistics, while the numerical solution of stochastic partial differential equations will equally impact numerical analysis and scientific computing. Due to limitations of space, in the present paper only some of these issues will be addressed.

Noisy nonlinear systems are encountered in many applications, from the nanoscale, e.g. in self-assembly processes [63], to the macroscale, e.g. large sudden disturbances in flow past an aircraft. Bifurcations and chaotic transitions in stochastic dynamical systems can be very different from instabilities in deterministic dynamical systems [53]. The counter-intuitive phenomenon of stochastic resonance for noisy signals, where the signal-to-noise ratio can be improved by increasing the noise level, is unique to stochastic systems. (An extensive review on stochastic resonance can be found in [14].) How exactly *extrinsic* stochasticity interacts with *intrinsic* stochasticity caused by the systems' nonlinear interactions is intriguing and generally not well understood at present.

In fluid flows, for example, at high Reynolds number where a wide range of small scales exists the *mean* flow seems to be totally unaffected even in the presence of substantial background turbulence [36]. In contrast, for other flow systems at low Reynolds number even small amounts of noise can have a dramatic effect on the structure of the *mean* flow. In Figure 1.1 we present an example of such response. A uniform flow past a circular cylinder oscillating in the crossflow direction is considered. A small amount of noise is superimposed on the uniform inflow velocity. In the absence of noise, a vortex street is formed in the wake characterized by the shedding of three vortices per shedding cycle – the so-called (P+S) pattern [66]. However, when noise is introduced the vortex street *switches* to another state above a certain noise threshold, characterized now by only two vortices per shedding cycle, as in the standard von Karman street for stationary cylinders (2S pattern). This phenomenon is discussed in detail in [33] and independent experimental validation can be found in [44].

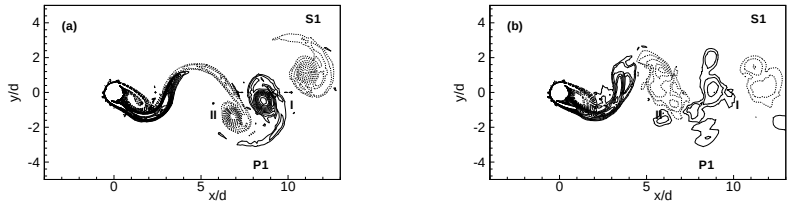

FIG. 1.1. *Comparison between deterministic (a) and stochastic* mean *(b) instantaneous vorticity field in a 2D cylinder wake at identical time. The "clean" wake shows a* $(P+S)$ *shedding-mode while the "noisy" wake shows a* $2S$ *shedding-mode similar to the standard von Karman vortex street. A* 10% *noise level is superimposed to the uniform inflow at Reynolds number 400.*

The third area in the aforementioned list contains complex systems with an extremely large number of degrees of freedom, e.g., models of turbulent flows at very high Reynolds number or atomistic simulations of mesoscopic size systems. Often, a coarse-grained procedure is employed to remove degrees of freedom carrying relatively small contributions to the overall system's energy. The removed degrees of freedom are usually accounted for by stochastic terms in the dimensionally reduced 'effective' or 'upscaled' equations. For example, in the dissipative particle dynamics method [10], coarse-graining of the molecular dynamics method leads to a system of stochastic ODEs that need to be solved efficiently for a large number of particles [51].

1.2. Statistical versus non-statistical approaches. As mentioned, from the mathematical point of view we need to reformulate mathematical models as stochastic ordinary or partial differential equations, while from the computational point of view new efficient numerical solution methods must be developed. From the perspective

of the computational scientist, familiarity with existing concepts as well as physical interpretation and implementation friendliness of the stochastic models obtained by such reformulation are also important factors. For the numerical solution of stochastic differential equations, we broadly distinguish *statistical* and *deterministic methods*.

The statistical approach includes among others Monte Carlo simulation, stratified sampling, Latin hypercube sampling (cf. [11]). Such statistical methods involve sampling and estimation and in most cases are straightforward to apply to partial differential equations if efficient deterministic PDE solvers are available. However, since the accuracy of the sampling techniques depends on the sample size, in accordance with the 'weak law of large numbers', simulations can become prohibitively expensive, especially for systems that are already computationally complex even in their deterministic version. In addition, attempting to resolve high-order moments associated with the low probability domain of the stochastic response may require a very large number of realizations, beyond the limits of even massively parallel computers. To accelerate convergence of statistical simulation methods, several modifications have been introduced. We mention here only Latin hypercube sampling [31], the Quasi-Monte Carlo (QMC) method [12], the Markov Chain Monte Carlo method (MCMC) [37], and the Response Surface method (RSM) [45]. These methods improve the efficiency of statistical, 'brute-force' sampling methods. Additional restrictions on the statistics of the random input are imposed, thus limiting the generality and applicability of these 'accelerated' sampling methods.

Alternatively, one can develop deterministic methods for stochastic PDEs. The most widely used method is the perturbation method, where input random fields are represented as infinite perturbation expansions of fluctuations around their mean fields. These perturbation expansions are truncated at a certain order in order to obtain coupled finite systems of differential equations for corresponding expansions of the random solution. We emphasize that *a priori* truncation of moment expansions of random data and solutions amounts *de facto* to the imposition of *moment closure hypotheses* on perturbation expansions.

In practical applications, at most second-order expansions are employed – the so-called '*first-order - second moment analysis*' [29] that has been used extensively in engineering applications [23, 28, 58, 74]. This is due to the following reasons: (a) first and second moments of random solutions are of main interest in applications; (b) the system of equations resulting from truncations at higher order becomes quite complicated beyond second-order; and (c) the dimensionality of the (deterministic) differential equations for the higher order moments of the random solution increases with the order of the moment to be computed: if the physical problem is posed in a (space or time) domain $D \subset \mathbb{R}^d$ of dimension d, the kth moment of the random solution involves the solution of a deterministic problem on the domain $D^k \subset \mathbb{R}^{kd}$. An inherent limitation of such first-order perturbation methods is that the uncertainties must not be too large, i.e., the fluctuations of the random fields should be small compared to their mean values (typically less then 10%). This requirement needs to be satisfied not only by the stochastic inputs but also by the stochastic outputs. This is especially difficult to verify *a priori* for nonlinear problems, as small fluctuations in random inputs may result in large fluctuations in the systems' responses (see an example in [72]). Also, higher order statistics of input data are not readily available.

A related approach is based on manipulation of the stochastic operators. Such methods include the Neumann expansion, which is based on expanding the inverse of the stochastic operator in a Neumann series [52, 73], and the weighted integral method

[8, 9, 55]. These methods have limitations on the type of model equations they can address, and, if truncated after second moment terms, they are also restricted to small uncertainties. In the present paper, we sketch, following [48, 49], an efficient approach for the deterministic approximation of second and higher order spatial correlations of random solutions, based on sparse tensor products of hierarchic finite element spaces. It allows, in principle, for the fast evaluation of perturbation expansions to any fixed order in *log-linear* complexity with respect to N, the number of degrees of freedom used to approximate the mean field problem in the physical domain D.

Another approach to deterministic numerical solution of stochastic PDEs is based on introducing *geometry* and *coordinates* in the probability space on which input and solution uncertainty are modeled. It is based on two observations:

(i) The space of random fields with finite second moments can be equipped with norm and innerproduct, thereby turning it into a Hilbert space and allowing for 'Galerkin projection' of a random field onto a (computationally convenient) family of parametric approximations of it – very much akin to what is done in the variational formulation of finite element methods for deterministic problems.

(ii) The efficiency of this approach depends crucially on the judicious choice of "coordinates" in probability space. Ghanem and Spanos in [54] pioneered the computational use of the *polynomial chaos* (PC) expansion method, and have successfully applied it to various problems in solid mechanics [16]. PC expansions are based on the homogeneous chaos theory of N. Wiener [65] and are essentially spectral expansions of Gaussian random fields into Hermite polynomials. PC expansions allow high-order deterministic approximation of random fields and appear to exhibit spectral convergence in many cases as we will show.

Classical Wiener-Hermice PC expansions are based on the *Hermite* polynomial functionals in terms of *Gaussian* random variables. In theory, they converge to any L_2 functional on the random space [6]. However, in practice they converge slowly for non-Gaussian random fields and do not apply to random fields with *discrete* distributions. Accordingly, for fast convergence in PC expansions and, hence, for computational efficiency, the "coordinates in probability space" in which PC expansions of the random solution are sought should be adapted to the statistics of the input data and of the random solution. This can be done in at least two ways:

(a) By employing *generalized PC expansions* (gPC expansions) that are orthogonal with respect to non-Gaussian probability measures. Such expansions, referred to as 'Wiener-Askey' chaos expansions were first employed in computational algorithms in [70], following developments in probability in [40, 46], and on orthogonal polynomials in [2, 25].

(b) By optimally separating deterministic and stochastic components of random input data with prescribed spatial correlation through Karhunen-Loève (KL) decomposition [30]. Apart from being a theoretical tool, we show how KL decompositions can be efficiently computed in general domains D for a wide class of spatial correlation functions by generalized Fast Multipole Methods for efficient computational spectral approximation of the covariance operator.

In gPC, the polynomials are chosen from the hypergeometric polynomials of the *Askey family* where the underlying random variables are not restricted to Gaussian random variables. In fact, there exists a unique correspondence between the probability distribution function (PDF) of the stochastic input and the weighting function of the orthogonal polynomials. The convergence properties of different trial bases were studied in [70] and exponential convergence rate was demonstrated computationally

for model problems. The aforementioned correspondence can be extended to arbitrary PDFs with the orthogonal polynomials constructed on-the-fly; this extension was presented in [60]. In essence, gPC approximations of random fields correspond to the spectral/hp element method, see [22, 47]. Depending on the stochastic regularity of the random field, it may by advantageous to combine mesh refinement with increase of the polynomial degree, leading to an *hp*-generalization of gPC approximations, see [59, 27, 4].

The paper is organized as follows: We first review basic concepts and formulation of gPC approximations of stochastic differential equations in Section 2. We discuss in some detail the representation of stochastic inputs in general domains and with prescribed spatial correlation using multipole-based Karhunen-Loève expansions in Section 3. In Sections 4 and 5 we present gPC solutions to prototype ordinary and partial differential equations. The exposition of the gPC-based methods in Sections 2 to 5 is formal. In Section 6 we present the main ideas on combining perturbation expansions and sparse grids. Finally, in Section 7 we address some outstanding mathematical and computational issues associated with stochastic modeling in general, and with gPC based methods in particular.

2. Generalized Polynomial Chaos. Stochastic mathematical models are based on a *probability space* $(\Omega, \mathcal{A}, \mathcal{P})$ where Ω is the event space, $\mathcal{A} \subset 2^\Omega$ its σ-algebra, and $\mathcal{P}$ its probability measure.

Data and solutions of stochastic differential equations are *random fields* $X(\omega)$, i.e. mappings $X : \Omega \to V$ from the probability space into a function space V. If $V = \mathbb{R}$, we speak of random variables, and if V is a function space over a time and/or space interval, of random fields or stochastic processes. For stochastic partial differential equations, V is often a space of generalized functions in a physical domain $D \subset \mathbb{R}^d$, $d = 2, 3$.

In all examples considered here, V is a Hilbert space with dual V', norm $\|\circ\|$ and inner product $(\cdot,\cdot) : V \times V \to \mathbb{R}$. As V is densely embedded in V', we abuse notation and denote by $(\cdot,\cdot)$ also the $V \times V'$ duality pairing.

A random field $X : \Omega \to V$ is a *second-order random field over a Hilbert space* V, if

$$\mathbb{E}\|X\|^2 = \mathbb{E}(X, X) < \infty,$$

where $\mathbb{E}$ denotes the expectation of a random variable $Y \in L^1(\Omega, \mathbb{R})$ defined by

$$\mathbb{E}Y = \int_{\omega\in\Omega} Y(\omega) dP(\omega).$$

Generalized polynomial chaos (gPC) is a means of representing second-order stochastic processes $X(\omega)$ parametrically through a set of random variables $\{\xi_j(\omega)\}_{j=1}^N, N \in \mathbb{N}$, through the events $\omega \in \Omega$:

$$X(\omega) = \sum_{k=0}^{\infty} a_k \Phi_k(\boldsymbol{\xi}(\omega)). \tag{2.1}$$

Here $\{\Phi_j(\boldsymbol{\xi}(\omega))\}$ are orthogonal polynomials in terms of a zero-mean random vector $\boldsymbol{\xi} := \{\xi_j(\omega)\}_{j=1}^N$, satisfying the orthogonality relation

$$\langle \Phi_i \Phi_j \rangle = \langle \Phi_i^2 \rangle \delta_{ij}, \tag{2.2}$$

where δ_{ij} is the Kronecker delta and $\langle \cdot, \cdot \rangle$ denotes the ensemble average. The number of random variables $N \in \mathbb{N}$ is in general infinite, so is the index in (2.1). In practice, however, we need to retain a finite set of random variables, i.e., to $\{\xi_j\}_{j=1}^N$ with $N < \infty$, and a finite-term truncation of (2.1).

The inner product in (2.2) is in the Hilbert space determined by the measure of the random variables

$$\langle f(\boldsymbol{\xi}) g(\boldsymbol{\xi}) \rangle = \int_{\omega \in \Omega} f(\boldsymbol{\xi}) g(\boldsymbol{\xi}) dP(\omega) = \int f(\boldsymbol{\xi}) g(\boldsymbol{\xi}) w(\boldsymbol{\xi}) d\boldsymbol{\xi} \tag{2.3}$$

with $w(\boldsymbol{\xi})$ denoting the density of the law $dP(\omega)$ with respect to the Lebesgue measure $d\boldsymbol{\xi}$ and with integration taken over a suitable domain, determined by the range of the random vector $\boldsymbol{\xi}$.

In the discrete case, the above orthogonal relation takes the form

$$\langle f(\boldsymbol{\xi}) g(\boldsymbol{\xi}) \rangle = \sum_{\boldsymbol{\xi}} f(\boldsymbol{\xi}) g(\boldsymbol{\xi}) w(\boldsymbol{\xi}). \tag{2.4}$$

In equation (2.1), there is a one-to-one correspondence between the type of the orthogonal polynomials $\{\Phi\}$ and the law of the random variables $\boldsymbol{\xi}$. This is determined by choosing the type of orthogonal polynomials $\{\Phi\}$ in such a way that their weighting function $w(\boldsymbol{\xi})$ in the orthogonality relation (2.3) has the same form as the probability distribution function of the underlying random variables $\boldsymbol{\xi}$. For example, the weighting function of Hermite orthogonal polynomials is $\frac{1}{\sqrt{(2\pi)^n}} \exp(-\frac{1}{2}\boldsymbol{\xi}^T \boldsymbol{\xi})$, and is the same as the probability density function of the $N-$dimensional Gaussian random variables $\boldsymbol{\xi}$. Hence, the classical Wiener polynomial chaos is an expansion of Hermite polynomials in terms of Gaussian random variables. A correspondence between orthogonal polynomials and random variables was first established in [40, 46]. Here, we list some types of generalized polynomial chaos corresponding to the commonly known distributions in Table 2.1. The flexibility of generalized polynomial

TABLE 2.1
Correspondence of the type of Wiener-Askey polynomial chaos and their underlying random variables ($N \geq 0$ is a finite integer).

	Random variables $\boldsymbol{\xi}$	Wiener-Askey chaos $\{\Phi(\boldsymbol{\xi})\}$	Support
Continuous	Gaussian	Hermite-chaos	$(-\infty, \infty)$
	gamma	Laguerre-chaos	$[0, \infty)$
	beta	Jacobi-chaos	$[a, b]$
	uniform	Legendre-chaos	$[a, b]$
Discrete	Poisson	Charlier-chaos	$\{0, 1, 2, \dots\}$
	binomial	Krawtchouk-chaos	$\{0, 1, \dots, N\}$
	negative binomial	Meixner-chaos	$\{0, 1, 2, \dots\}$
	hypergeometric	Hahn-chaos	$\{0, 1, \dots, N\}$

chaos can be seen in Figure 2.1, where we use Hermite-chaos to approximate a uniform random variable. We observe that although the Hermite-chaos converges to the target distribution as the expansion order increases, the approximation is not accurate and suffers from Gibbs-like oscillations. On the other hand, the generalized polynomial chaos corresponding to a uniform distribution, the Legendre-chaos (see Table 2.1), can represent uniform random variables exactly by just first-order expansions. For more details on approximation of an arbitrary random variable via generalized polynomial chaos, see [70, 60].

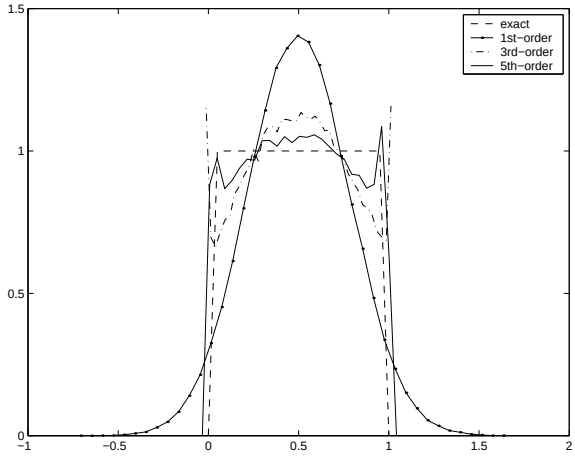

FIG. 2.1. *Approximations of a uniform distribution by Hermite-Chaos*

The expansion in equation (2.1) can be truncated by reducing the expansion to the finite-dimensional space, i.e., expansion of finite-dimensional random variables $\boldsymbol{\xi}$, according to the nature of random inputs; we also set the highest order of the polynomials $\{\Phi\}$ according to accuracy requirements. The finite-term expansion takes the form

$$X(\omega) = \sum_{j=0}^{M} a_j \Phi_j(\boldsymbol{\xi}(\omega)), \tag{2.5}$$

where $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_N)^T$ is an $N-$dimensional random vector with ξ_i independent of ξ_j for all $1 \leq i \neq j \leq N$. If we denote the highest order of polynomial $\{\Phi\}$ as P, then the total number of expansion terms $(M+1)$ is,

$$(M+1) = (N+P)!/(N!P!). \tag{2.6}$$

The multi-dimensional generalized polynomial chaos expansion is constructed as the tensor product of the corresponding one-dimensional expansion. Note in one-dimensional expansions $(N=1)$, $M=P$.

Let us now consider a general setting for a differential equation with random inputs

$$\mathcal{L}(\mathbf{x}, t, \omega; u) = f(\mathbf{x}, t; \omega), \qquad x \in D(\Lambda), t \in (0, T), \omega \in \Omega, \tag{2.7}$$

where $\mathcal{L}$ is a differential operator, $D(\Lambda) \in \mathbb{R}^d (d = 1, 2, 3)$ a bounded domain with diameter $\Lambda > 0$, and $T > 0$. $(\Omega, \mathcal{A}, \mathcal{P})$ is an appropriately defined complete probability space, where $\mathcal{A} \subseteq 2^{\Omega}$ is the σ-algebra and $\mathcal{P}$ the probability measure; $u := u(\mathbf{x}, t; \omega)$ is the solution and $f(\mathbf{x}, t; \omega)$ is the source term. The general procedure of applying the generalized polynomial chaos consists of the following steps:

1. Express the random inputs by a finite number of random variables $\boldsymbol{\xi}(\omega) = \{\xi_1(\omega), \cdot, \xi_N(\omega)\}$, and rewrite the stochastic problem parametrically as

$$\mathcal{L}(\mathbf{x}, t, \boldsymbol{\xi}; u(x, t; \boldsymbol{\xi})) = f(\mathbf{x}, t; \boldsymbol{\xi}). \tag{2.8}$$

This step is trivial when the random inputs already take the form of random variables. When the random inputs are random fields, a decomposition technique is needed. One popular choice of such decomposition is the Karhunen-Loève expansion, which will be discussed in the following section.

2. Approximate the solution and inputs by finite-term polynomial chaos expansions (2.5), and substitute the expanded variables into the variational equation

$$\mathcal{L}\left(\mathbf{x}, t, \boldsymbol{\xi}(\omega); \sum_{i=0}^{M} u_i \Phi_i(\boldsymbol{\xi}(\omega)\right) = f(\mathbf{x}, t; \boldsymbol{\xi}(\omega)).$$

3. Perform a Galerkin projection onto each of the polynomial basis

$$\left\langle \mathcal{L}\left(\mathbf{x}, t, \boldsymbol{\xi}; \sum_{i=0}^{M} u_i \Phi_i(\boldsymbol{\xi})\right), \Phi_k(\boldsymbol{\xi}) \right\rangle = \langle f, \Phi_k(\boldsymbol{\xi}) \rangle, \qquad k = 0, 1, \cdots, M.$$

This procedure results in a set of $(M+1)$ *deterministic*, in general coupled, differential equations which can be solved by conventional discretization methods. Importantly, this successive Galerkin discretization "in probability" (by gPC approximation) and "in physical space" (by finite differences or finite elements) yields generally large systems of algebraic equations which carry a tensor product block structure. For linear problems, the structure of the probabilistic part of this system is determined by the orthogonal polynomials used in the gPC discretization and is, in particular, independent of the differential operator under consideration. This observation underlies the application of gPC type methods described in Sections 4 and 5 ahead. Let us next turn to the problem of parametric representation of random field input data.

3. Representation of Stochastic Inputs: Karhunen-Loève Expansion. Karhunen-Loève (KL) expansion is a way of representing a random process [30]. It is based on the spectral expansion of the covariance function of the process. Let us denote the process as $h(x, \omega)$ and its covariance function as $R_h(x_1, x_2)$, where x_1 and x_2 are spatial or temporal coordinates. By definition, the covariance function is real, symmetric and positive-definite. It has an orthogonal set of eigenfunctions which forms a complete basis; $\phi_i(x)$ and λ_i are the eigenfunctions and eigenvalues of the covariance function, respectively, i.e.,

$$\int R_h(x_1, x_2)\phi_m(x_2)dx_2 = \lambda_i \phi_m(x_1), \qquad m = 1, 2, \cdots. \tag{3.1}$$

The finite-term Karhunen-Loève expansion then takes the following form:

$$h_N(x, \omega) = \bar{h}(x) + \sigma_h \sum_{m=1}^{N} \sqrt{\lambda_m}\phi_m(t)\xi_m(\omega), \tag{3.2}$$

where $\bar{h}(x)$ is the mean of the random process, σ_h the standard deviation, and $\xi_m(\omega)$ a set of uncorrelated random variables with zero mean and unit variance. These random variables are determined by

$$\xi_m(\omega) = \int h(t, \omega)\phi_m(x)dx, \quad m = 1, 2, \cdots. \tag{3.3}$$

3.1. Properties of Karhunen-Loève expansion. Among other possible decompositions of a random process, the Karhunen-Loève expansion is optimal in the sense that the mean-square error of the finite-term representation $h_N(x,\omega)$ of the process $h(x,\omega)$ is minimized [30]. The covariance function $R_h(x_1,x_2)$ plays a central role in Karhunen-Loève expansion, as it is the kernel function in the eigenvalue problem defined in (3.1). The decay of the eigenvalues and the regularity of the eigenfunctions depend strongly on the regularity of the covariance kernel function. It can be shown that (see [50] for the proofs)

- if R_h is piecewise analytic on $D \times D$, we have exponential KL eigenvalue decay and convergence of the KL expansion. More precisely,

$$\lambda_m \leq c_1 \exp(-c_2 m^{1/d}), \qquad \forall m \geq 1, \tag{3.4}$$

  where $c_1, c_2 > 0$ are constants independent of m;
- if R_h is piecewise $H^{k,0} \equiv H^k \otimes L^2$ with $k > 0$, i.e., R_h has finite differentiability, then

$$\lambda_m \leq c_3 m^{-k/d}, \qquad \forall m \geq 1, \tag{3.5}$$

  where $c_3 > 0$ is a constant.

For example, two of the most commonly used covariance kernels are the *Gaussian* covariance function

$$R_h(x_1,x_2) = e^{-(x_1-x_2)^2/\gamma^2}, \qquad (x_1,x_2) \in D(\Lambda) \times D(\Lambda), \tag{3.6}$$

and the *exponential* covariance function

$$R_h(x_1,x_2) = \exp(-|x_1-x_2|/\gamma), \qquad (x_1,x_2) \in D(\Lambda) \times D(\Lambda), \tag{3.7}$$

where $\gamma > 0$ is the correlation length. The Gaussian kernal (3.6) is analytic, and a sharper estimate of eigenvalue decay than (3.4) can be obtained, i.e., $0 < \lambda_m \lesssim \sigma_h^2 \frac{(1/\gamma)^{m^{1/d}+2}}{\Gamma(0.5m^{1/d})}, \forall m \geq 1$, where Γ is the gamma function. In Fig. 3.1, we show the decay of eigenvalues for the Gaussian kernel function. On the left of Fig. 3.1, the decay in a three-dimensional ($d = 3$) domain is shown, along with the theoretical estimate, for up to $2,000$ eigenvalues. We observe good agreement between the estimate and the actual eigenvalue decay. On the right of Fig. 3.1, the first 21 eigenvalues are plotted for the Gaussian kernel in 1D, 2D and 3D ($d = 1,2,3$) spatial domains D. We observe exponential decay of eigenvalues in all three cases, with smaller decay rate at higher dimensions. This is consistent with the theoretical estimate. In Fig. 3.2, we show the decay of eigenvalues for the covariance kernel function $R_h = \exp(-|x_1-x_2|^{1+\delta})$, where $0 \leq \delta < 1$ is a parameter to control the differentiability of the function. (Note when $\delta = 1$ we obtain the Gaussian kernel.) The first 14 eigenvalues are shown, along with the theoretical estimate, and good agreement is obtained. Also, for smaller values of δ, the eigenvalues decay slower, as predicted by the estimate.

In the above results, we have used the notions of "piecewise analytic" and "piecewise $H^{k,0}$". For precise definitions of such terms and mathematical analysis of KL expansions, see [50]. Numerical studies of the convergence of truncated KL expansions can be found in [19]. In another study ([34]), the convergence rate of the N-term truncated Karhunen-Loève expansion for the exponential covariance function (3.7) in one-dimensional ($d = 1$) space $D(\Lambda)$ was estimated as

$$\varepsilon \equiv \frac{\langle h^2(x;\omega)\rangle - \langle h_N^2(x;\omega)\rangle}{\langle h^2(x;\omega)\rangle} \sim 0.4053 \frac{\Lambda}{\gamma}\frac{1}{N}. \tag{3.8}$$

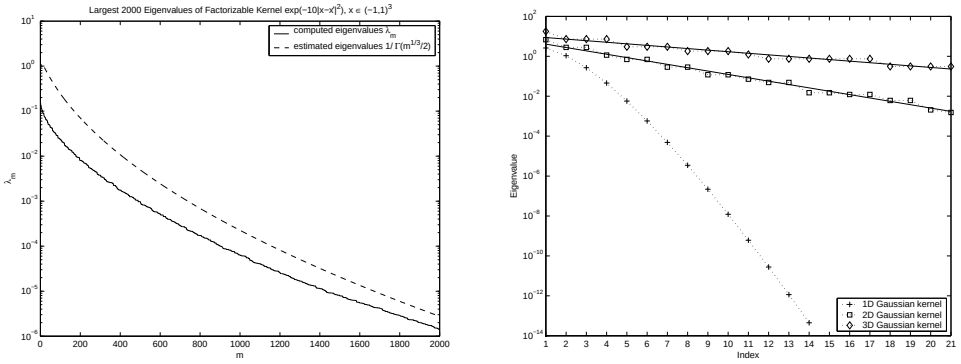

FIG. 3.1. *Eigenvalue decay for analytic covariance kernel. Left: Eigenvalues in 3D domain and theoretical estimate; Right: Eigenvalue decays in 1D, 2D, and 3D domains.*

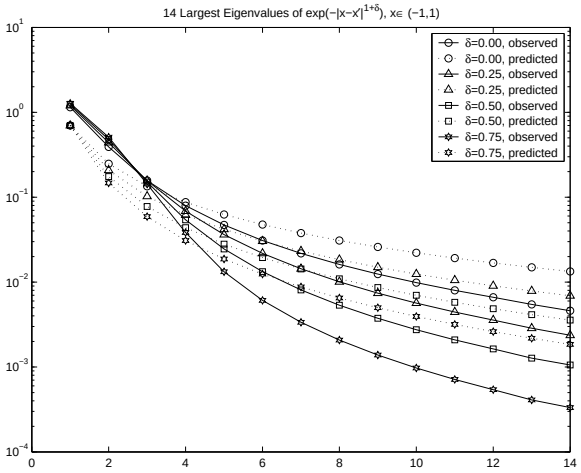

FIG. 3.2. *Decay of eigenvalues of covariance functions with varying differentiability.*

Numerical simulations for random oscillators established that this indeed is a very sharp error estimate [34].

3.2. Choices of covariance kernel functions. The kernel function R_h in (3.1) is the covariance function of the random process $h(x;\omega)$, and it can be constructed from repetitive observations of the process. Such measurement procedure can be costly, and in computations we often assume the form of the covariance function. The most commonly used covariance function is the "exponential kernel" defined in (3.7). Such a covariance function can be generated through a first-order Markov process

$$h_i = \alpha h_{i-1} + \beta \xi_i, \quad \alpha = e^{-1/\gamma}, \quad \alpha^2 + \beta^2 = 1, \tag{3.9}$$

where $\{h_i\}$ are discrete points on a uniform stencil indexed by integers i, $\{\xi_i\}$ are i.i.d. random variables with zero mean value and unit variance. Note that such a process has a preference in one direction and is more suitable to model random time series. For *spatial* random processes, it is more appropriate to take into account the dependence on all spatial directions. Several choices are presented in Table 3.1, along with the corresponding covariance functions generated by them.

TABLE 3.1

Correspondence of random processes and their covariance functions. Here i, j, k are integers, $r = |x_1 - x_2|$ is the distance between two points in $D \subset \mathbb{R}^d, d = 1, 2, 3$, and $\{\xi\}$ are i.i.d. random variables with zero mean and unit variance. $K_1(x)$ is the first-order modified Bessel function of the second kind. For the $d = 3$ case both the plus and minus terms (e.g. $h_{i\pm1,j,k}$) should be included.

Dimension d	Random process $h(x;\omega)$	Covariance function R_h
$d = 1$	$h_i = \alpha h_{i-1} + \beta\xi_i$	$\exp\left(-\frac{r}{\gamma}\right)$
$d = 1$	$h_i = \frac{b}{2}(h_{i-1} + h_{i+1}) + \xi_i$	$(1 + r/\gamma)\exp(-r/\gamma)$
$d = 2$	$h_{i,j} = \frac{b}{4}(h_{i\pm1,j} + h_{i,j\pm1}) + \xi_{i,j}$	$\frac{r}{\gamma}K_1\left(\frac{r}{\gamma}\right)$
$d = 3$	$h_{i,j,k} = \frac{b}{6}(h_{i\pm1,j,k} + h_{i,j\pm1,k} + h_{i,j,k\pm1}) + \xi_{i,j,k}$	$\exp\left(-\frac{r}{\gamma}\right)$
$d = 1, 2, 3$	N/A	$\exp(-r^2)$

It is worth mentioning that the spatial processes with dependence in *all* directions in Table 3.1 are obtained as solutions to Helmholtz equation ([64, 35]). For example, in two dimensions ($d = 2$), the process with covariance kernel $\left(\frac{r}{\gamma}\right)K_1\left(\frac{r}{\gamma}\right)$ corresponds to the equation

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} - \frac{1}{\gamma^2}\right) h(x, y) = \xi(x, y).$$

On the other hand, the exponential kernel $\exp(-r/\gamma)$ in two dimensions ($d = 2$) corresponds to

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} - \frac{1}{\gamma^2}\right)^{3/4} h(x, y) = \xi(x, y).$$

It is difficult to find a physical analog which would lead to such a relation although the exponential kernel is used in the literature indiscriminately in all dimensions.

The appropriate relationship between such kernels and discrete dynamical systems for a variety of conditions was studied in [35]. Here we give a specific example for a spatially periodic process on one-dimensional equidistant grid with spacing Δx. The discrete dynamical system is given by

$$h_i = \frac{\alpha}{2}(h_{i-1} + h_{i+1}) + \beta\xi_i, \quad i = 1, 2, 3, \ldots, (n-1)$$

where $h_1 = h_n = 0$ and n is an even integer. The corresponding continuous process satisfies (within $\mathcal{O}(\Delta x^2)$ approximation) the following Helmholtz equation

$$\frac{\partial^2 h}{\partial x^2} - \frac{h}{\gamma^2} = 2\frac{\beta}{\Delta x^2}\xi.$$

Solving analytically the above equation with periodic boundary conditions produces the solution $h(x, \xi)$ from which the covariance can then be computed, see [35].

3.3. Fast computation of Karhunen-Loève expansion. In one spatial dimension, the eigenvalue problem (3.1) with the exponential covariance kernel function (3.7) can be solved analytically ([57, 16]). Such analytical solutions, however, are not available for most other cases, and numerical procedures are required to solve equation (3.1). Numerical methods are typically based on projection methods, such as Galerkin methods and collocation methods (cf. [5]). In projection methods, we

choose a sequence of finite-dimensional approximating subspaces $V_n \subseteq L^2(D), n \geq 1$, with V_n having dimension $\kappa_n = \dim(V_n)$ and a set of basis $\{v_k\}_{k=1}^{\kappa_n}$. We then seek $\phi_m^n \in V_n, m = 1, 2, \cdots$ to approximate the eigenfunctions in (3.1)

$$\phi_m^n(x) = \sum_{k=1}^{\kappa_n} a_{m,k} v_k(x), \qquad x \in D, m = 1, 2, \cdots .$$

The Galerkin approximation of (3.1) reads: find $\lambda_m \neq 0$ and $\phi_m^n \in V_n$ such that

$$\int_{D\times D} R_h(x_1, x_2)\phi_m^n(x_2)v(x)dxdx_2 = \lambda_m \int_D \phi_m^n(x)v(x)dx, \quad \forall v \in V_n. \tag{3.10}$$

This problem is a matrix eigenvalue problem

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{M}\mathbf{v}, \tag{3.11}$$

where matrices $\mathbf{A}$ and $\mathbf{M}$ are symmetric and positive definite, with $\mathbf{M}$ being diagonal if we choose an $L^2(D)$-orthogonal basis. Usually, V_n are finite element spaces of piecewise polynomial functions. No continuity of V_n is required since the approximation (3.11) is in $L^2(D)$. In [43], a wavelet-Galerkin method with better resolution is presented.

The size of the matrices is $\kappa_n \times \kappa_n$, and it can be large for practical problems, especially in three-dimensional space ($d = 3$). For example even a modest size problem with about 200,000 degrees of freedom required solution on 256 processors in [61]. To avoid the $O(\kappa_n^2)$ computational and memory cost for the Galerkin solution of the KL-eigenvalue problem (3.11), several methods have been proposed. A kernel-independent Fast Multipole Method (FMM) with $O(\kappa_n \log(\kappa_n))$ operations and memory is discussed in [50]. For Gaussian covariance kernels R_h as in (3.6), in [61] a Fast Gauss Transform (FGT) with $O(\kappa_n)$ operations and memory complexity is applied and sharp estimates on the error induced by the Gauss Transform is presented.

4. Ordinary Differential Equations. In this section, we illustrate the solution procedure of gPC for a simple ordinary differential equation, and present error convergence both through numerical examples and theoretical estimates. The model ODE we consider takes the following form

$$\frac{dy}{dt}(t, \omega) = -k(\omega)y, \qquad y(0) = \hat{y}, t \in (0, T), \tag{4.1}$$

where the decay rate coefficient $k(\omega)$ is a random variable with certain continuous distribution function $f(k)$ and zero mean value. The solution takes a simple form of

$$y(t, \omega) = \hat{y}e^{-k(\omega)t}. \tag{4.2}$$

We apply the gPC expansion (2.5) to the solution y and random input k

$$y(t, \omega) = \sum_{i=0}^{M} y_i(t)\Phi_i(\xi(\omega)), \qquad k(\omega) = \sum_{i=0}^{M} k_i\Phi_i(\xi(\omega)). \tag{4.3}$$

Note here the only random input is $k(\omega)$ and a one-dimensional gPC is needed, i.e., $N = 1$ in (2.6) and $M = P$, where P is the highest order of expansion. By substituting the expansions into the governing equation, we obtain

$$\sum_{i=0}^{M} \frac{dy_i(t)}{dt}\Phi_i = -\sum_{i=0}^{M}\sum_{j=0}^{M} \Phi_i\Phi_j k_i y_j(t). \tag{4.4}$$

We then project the above equation onto the random space spanned by the orthogonal polynomial basis $\{\Phi_i\}$ by taking the inner product of the equation with each basis. By utilizing the orthogonality condition (2.2), we obtain:

$$\frac{dy_l(t)}{dt} = -\frac{1}{\langle \Phi_l^2 \rangle} \sum_{i=0}^{M} \sum_{j=0}^{M} e_{ijl} k_i y_j(t), \qquad l = 0, 1, \ldots, M, \tag{4.5}$$

where $e_{ijl} = \langle \Phi_i \Phi_j \Phi_l \rangle$. Note that the coefficients are smooth and thus any standard ODE solver can be employed here.

In Figure 4.1, the computational results of Jacobi-chaos expansion is shown, where the random input $k(\omega)$ is assumed to have a *beta distribution* with PDF of the form

$$f(k; \alpha, \beta) = \frac{(1-k)^\alpha (1+k)^\beta}{2^{\alpha+\beta+1} B(\alpha+1, \beta+1)}, \qquad -1 < k < 1, \ \alpha, \beta > -1, \tag{4.6}$$

where $B(\alpha, \beta)$ is the Beta function defined as $B(p, q) = \Gamma(p)\Gamma(q)/\Gamma(p+q)$. An important special case is $\alpha = \beta = 0$ when the distribution becomes the uniform distribution and the corresponding Jacobi-chaos becomes the Legendre-chaos. We observe exponential convergence rate of the errors in *mean* and *variance* on the right of Figure 4.1, with different sets of parameter values α and β. In particular, we mote that the asymptotic convergence rate seems to be the same for the variance and the mean unlike in Monte Carlo methods.

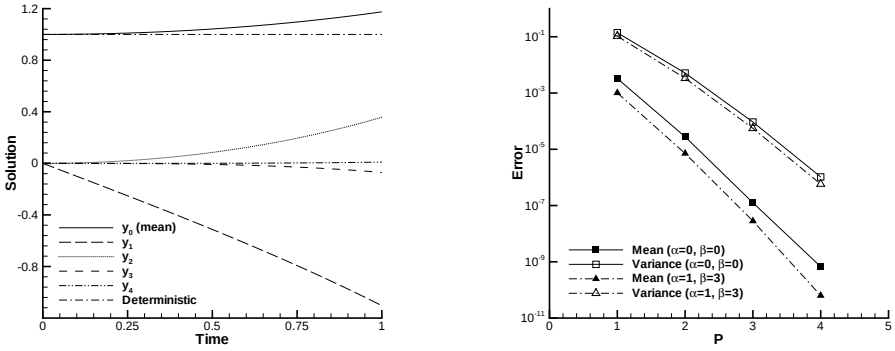

FIG. 4.1. *Solution with beta random input by 4th-order Jacobi-Chaos; Left: Solution of each mode ($\alpha = \beta = 0$: Legendre-Chaos), Right: Error convergence of the mean and the variance with different α and β.*

For this simple ODE we can also perform error analysis for different types of distributions. To this end, we define the relative *mean-square error* as $\epsilon_P = \langle (y(t) - y_P(t))^2 \rangle / \langle y^2(t) \rangle$, where $y_P(t)$ is the finite-term expansion (4.3). The following results have been obtained in [34]:

- If $k(\omega)$ is a *Gaussian* random variable with zero mean and standard deviation $\sigma > 0$ and *Hermite-chaos* is used, then

$$\epsilon_P \le \frac{(\sigma t)^{2(P+1)}}{e^{(\sigma t)^2} - 1} \left[(P+1)! \left(1 - \frac{(\sigma t)^2}{P+1} \right) \right]^{-1}. \tag{4.7}$$

- If $k(\omega)$ is an *exponential* random variable with zero mean and standard deviation $\sigma > 0$ and *Laguerre-chaos* is used, then

$$\epsilon_P = \left(\frac{\sigma t}{1+\sigma t}\right)^{2P}. \tag{4.8}$$

- If $k(\omega)$ is a *uniform* by distributed random variable and *Legendre-chaos* is used, no explicit formula for the error is available. However, the error can be readily evaluated via the three-term recurrence formula of the Legendre polynomials.

In Figure 4.2 we plot the number $(P+1)$ of Legendre expansion terms that is needed to ensure that the error reaches a prescribed value. In particular, we fix the value at $\epsilon = 10^{-7}$. It can be seen that as time increases, the number of terms required grows, and the rate of such growth is different for the three cases; the Legendre-chaos has the slowest growth rate and the Hermite-chaos the fastest. Note that the time axis is scaled with the variance σ for each process. More numerical examples for the first-order ODEs can be found in [70] and the detailed error analysis in [34].

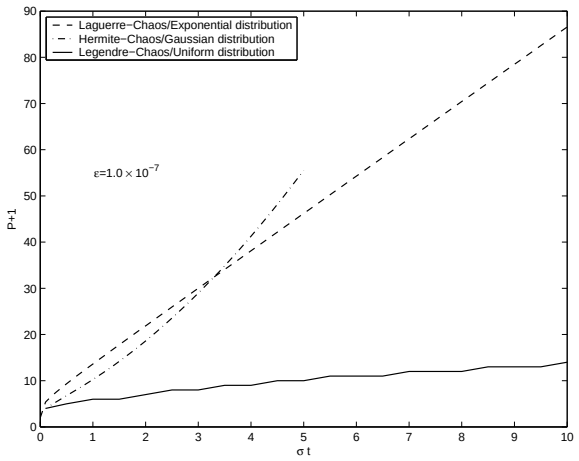

FIG. 4.2. *Number of expansion terms needed to reach a prescribed error level of* $\epsilon_P = 10^{-7}$. *The three curves are Hermite-chaos, Laguerre-chaos, and Legendre-chaos, corresponding to* $k(\omega)$ *being Gaussian, exponential, and uniform random variables, respectively.*

5. Partial Differential Equations. In this section, we present applications of generalized polynomial chaos to some prototype partial differential equations. In particular, we consider some cases for which analytical results can be derived so that such cases can be used for verification studies.

5.1. Linear advection equation. First, we consider the linear advection equation with uncertain transport velocity field, i.e.,

$$\frac{\partial u}{\partial t}(t,x,\omega) + V(t,x,\omega)\frac{\partial u}{\partial x} = 0 \qquad (t,x,\omega) \in [0,T]\times[-1,1]\times\Omega, \tag{5.1}$$

with initial condition $u(0,x) = \sin \pi x$, and periodic boundary conditions. We assume that the transport velocity $V(t,x,\omega)$ is a given random process with mean value $\bar{V}(x) = 1$ and finite variance $\sigma^2 > 0$. If the transport velocity $V \equiv V(t,\omega)$ in (5.1) is

not a function of space, the initial value problem can be solved exactly by the method of characteristics (see [20]). The mean solution is given by

$$\bar{u}(x,t) = \sin \pi(x + 1 - \bar{V}t)e^{-(\pi\tau\sigma)^2/2}, \tag{5.2}$$

where τ characterizes the correlation structure of $V(t,\omega)$ in time

$$\tau^2 = \begin{cases} t^2, & \text{fully correlated,} \\ (\Delta t)t, & \text{uncorrelated,} \\ 2\gamma[t - \gamma(1 - e^{-t/\gamma})], & \text{partially correlated,} \end{cases} \tag{5.3}$$

where γ is the correlation length and Δt denotes the sampling interval.

The variance of the solution $u(x,t)$, when $V(t,\omega)$ is a Gaussian random field, is:

$$\mathrm{Var}(u(x,t)) = \frac{1}{2}(1 - e^{-(\pi\sigma\tau)^2})[1 + \cos 2\pi(x + 1 - \bar{V}t)e^{-(\pi\sigma\tau)^2}]. \tag{5.4}$$

Detailed computational results for linear advection equations can be found in [20]; for advection-diffusion equations, see [62]. We note here that the numerical solution of the gPC equations is rather trivial using standard discretization methods. However, following a Monte Carlo approach special care is required in handling the spatial discretization in the case that $V(x,t,\omega)$ is a stochastic process depending on the spatial or temporal coordinate.

5.2. Elliptic equation. Next, we consider diffusion problems with random diffusivity $\kappa(x,\omega)$ which are prototype equations for subsurface flow problems and heat conduction:

$$\begin{cases} -\nabla \cdot [\kappa(x,\omega)\nabla u(x,\omega)] = f(x,\omega), & (x,\omega) \in D \times \Omega \\ u(x,\omega) = g(x,\omega), & (x,\omega) \in \partial D \times \Omega \end{cases} \tag{5.5}$$

Here D is a bounded domain in $\mathbb{R}^d$ $(d = 1,2,3)$ and Ω is a probability space, and f, g and κ are $\mathbb{R}$-values functions on $D \times \Omega$. This also can be considered as a model of steady state diffusion problems subject to internal (diffusivity κ) and/or external (source term f and/or Dirichlet boundary condition g) uncertainties. For theoretical analysis on the convergence of polynomial chaos expansions for elliptic problems, see [4, 49].

In order to demonstrate numerically the convergence of generalized polynomial chaos, we consider the following simple benchmark problem

$$\frac{d}{dx}\left[\kappa(x,\omega)\frac{du}{dx}(x,\omega)\right] = 0, \qquad u(0,\omega) = 0, \qquad u(1,\omega) = 1. \tag{5.6}$$

The random diffusivity has the form

$$\kappa(x,\omega) = 1 + \epsilon(\omega)x, \tag{5.7}$$

where $\epsilon(\omega)$ is a random variable, and $\kappa(x,\omega) > 0$. The exact solution to this problem is

$$u_e(x,\omega) = \begin{cases} \ln[1 + \epsilon(\omega)x] \,/\, \ln[1 + \epsilon(\omega)], & \text{for } \epsilon(\omega) \neq 0, \\ x, & \text{for } \epsilon(\omega) = 0, \end{cases} \tag{5.8}$$

and we compute the 'mean-square' error defined as $e_2(x) = \left(\mathbb{E}\left[u_P(x,\omega) - u_e(x,\omega)\right]^2\right)^{1/2}$, where $u_P(x,\omega)$ is the $P-$th order gPC solution. The L_∞ norm is employed in capturing errors due to spatial discretization, which is done using the spectral/hp element method [22].

In Figure 5.1, the convergence of the mean-square error of Jacobi-chaos is shown, where we assume $\epsilon(\omega) = \sigma\xi(\omega)$ in equation (5.7) is a *beta* random variable. Here $\sigma > 0$ measures the magnitude of input uncertainty and $\xi(\omega)$ is a standard beta random variable in $(-1,1)$ with PDF (4.6). It can be seen on the *semi-log* scale that the Jacobi-chaos solution, including the Legendre-chaos for uniform random variables ($\alpha = \beta = 0$), converges exponentially fast as the expansion order P increases. The exponential convergence rate is retained for large input uncertainty such as $\sigma = 0.9$, which is close to the limit of the existence of the solution ($\sigma < 1$). This is in direct contrast to the perturbation-based method which typically works for $\sigma < 0.1$.

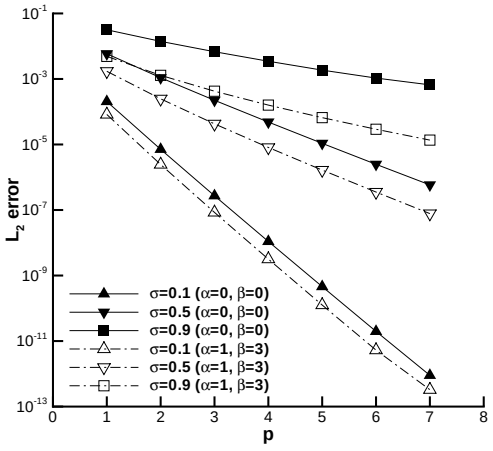

FIG. 5.1. *Convergence of Jacobi-chaos for the one-dimensional elliptic model problem.*

For numerical examples with other types of random inputs, multi-dimensional applications with random fields, and an efficient block-Jacobi iteration solver for the resulting equations, see [69].

5.2.1. Heat equation. We now consider the unsteady stochastic heat equation in a spatially varying medium, i.e.,

$$c(x,\omega)\frac{\partial T}{\partial t}(t,x,\omega) = \nabla\cdot[k(x,\omega)\nabla T] + f(t,x,\omega) \qquad (t,x,\omega)\in(0,\infty)\times D\times\Omega \quad (5.9)$$

subjected to the following initial and boundary conditions

$$T(0,x,\omega) = T_0(x,\omega), \tag{5.10}$$

$$T(t,x,\omega) = T_b, \quad x\in\partial D_1; \qquad -k\frac{\partial T}{\partial n}(t,x,\omega) = q_b, \quad x\in\partial D_2. \tag{5.11}$$

The temperature field $T(t,x,\omega)$ and heat source $f(t,x,\omega)$ are $\mathbb{R}$-valued functions on $[0,\infty)\times D\times\Omega$. The initial condition T_0, the volumetric heat capacity of the medium

c, and the conductivity k are $\mathbb{R}$-valued functions on $D \times \Omega$. ∂D_1 and ∂D_2 denote the subsets of the boundary with fixed temperature and heat flux, respectively.

Here we consider the heat conduction in an electronic chip subject to uncertainties in heat conductivity and capacity (see equation (5.9)). The computational domain D is shown in Figure 5.2 along with the spatial discretization. The boundary of the domain consists of four segments: the top Γ_T, the bottom Γ_B, the two sides Γ_S and the boundaries of the cavity Γ_C, which has a depth of 0.6. Adiabatic boundary conditions are prescribed on Γ_B and Γ_S. The cavity boundary Γ_C is exposed to heat flux $q_b|_{\Gamma_C} = 1$. Two types of conditions on the top Γ_T are considered: one is maintained at constant temperature $T = 0$ (case 1) and the other is adiabatic (case 2). Due to non-zero net heat flux into the domain, there is no steady state in case 2. The initial condition is zero everywhere. Six reference points are placed at the vertices of some chosen elements in the domain, as shown in Figure 5.2. We are interested in the stochastic solution at these points and their cross-correlation coefficients.

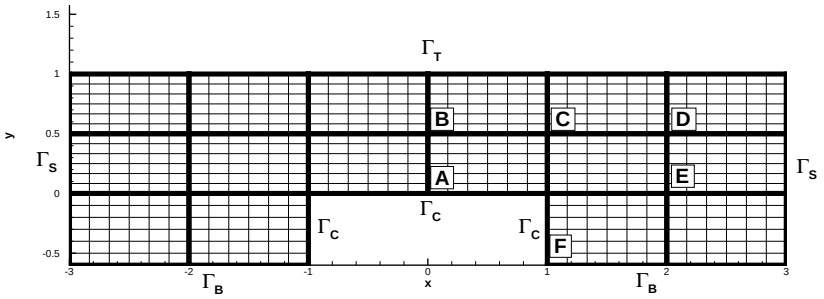

FIG. 5.2. *Schematic of the geometry of the electronic chip. The mesh consists of 16 spectral elements of order 6^{th} (7 points in each direction).*

The uncertain heat conductivity and capacity of the medium are random fields, with mean fields $\bar{k}(x, y; \omega) = 1, \bar{c}(x, y; \omega) = 1$, and with correlation functions of the form $C(r) = \frac{r}{\gamma} K_1\left(\frac{r}{\gamma}\right)$, as defined in Section 3.2. The Karhunen-Loève (KL) decomposition (3.2) is employed to decompose the input random field. In Figure 5.3, the cross-correlation coefficient between point A and B, and A and C, are plotted, along with the results from Monte Carlo simulation. Good agreement is seen between the two approaches. The Legendre-chaos is 4-dimensional ($N = 4$) and third-order ($P = 3$), which results in $(M + 1) = 35$ terms expansion (see equation (2.6)). On the other hand, we utilize $20,000$ and $150,000$ realizations for the steady problem (case 1) and the unsteady problem (case 2), respectively. More applications of stochastic heat conduction can be found in [71].

5.3. Burgers' equation. Here we consider the viscous Burgers' equation,

$$\begin{cases} u_t + uu_x = \nu u_{xx}, & x \in [-1, 1], \\ u(-1) = 1 + \delta, & u(1) = -1, \end{cases} \tag{5.12}$$

where $\delta > 0$ is a small perturbation to the left boundary condition ($x = -1$) and $\nu > 0$ is the viscosity. The presence of viscosity smoothes out the shock discontinuity which will develop otherwise. Thus, the solution has a transition layer, which is a region of rapid variation and extends over a distance $O(\nu)$ as $\nu \downarrow 0$. The location of

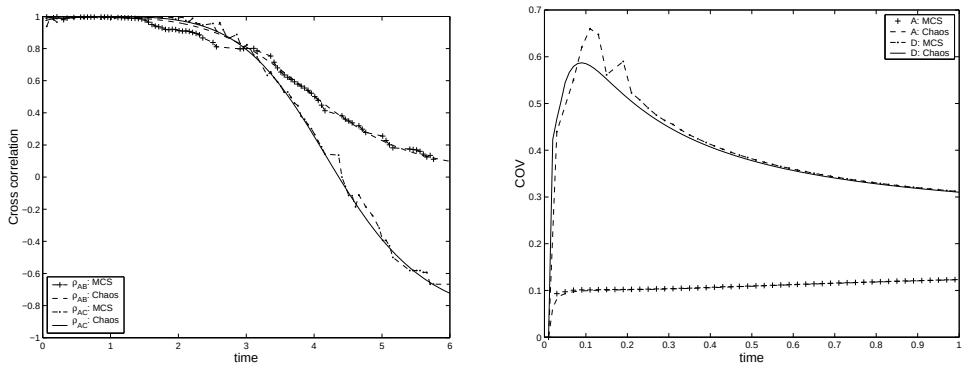

FIG. 5.3. *Comparison of results obtained by Monte Carlo simulation (MCS) and gPC expansion. Left - case 1: evolution of cross-correlation coefficients at reference points for case 1 (20,00 realizations for MCS); Right -case 2: evolution of COVs at reference points (150,000 realizations for MCS).*

the transition layer z, defined as the zero of the solution profile $u(z) = 0$, varies with time, and its eventual location at steady state is extremely sensitive to the boundary data. This phenomenon, termed *supersensitivity* in *deterministic* asymptotic analysis, was first observed by Lorentz [32]. In this section, we will present numerical solutions that exhibit supersensitivity under random perturbations on the boundary condition. In particular, we consider $\delta \sim U(0, \epsilon)$ is a uniform random variable in $(0, \epsilon)$ with $\epsilon \ll O(1)$.

In Table 5.1, the mean location of the transition layer ($\bar{z}$) and its standard deviation (σ_z) are shown. They are obtained by assuming $\delta \sim U(0, 0.1)$ and $\nu = 0.05$. The results of several methods are shown, along with their computational cost normalized by the cost of one deterministic simulation. Specifically, interval analysis deals with the maximal output bounds, and is straightforward to apply to this problem [1]. However, it does not provide any statistical information of the solution. The results from perturbation methods are noticeably different from the accurate solution obtained by Monte Carlo simulation with $10,000$ realizations. In addition, the fourth-order perturbation method does not yield any improvement over the first-order method. This suggests that the perturbation method converges slowly, if at all. The Legendre-chaos method accurately resolves the solution statistics. At fourth-order, its cost is about the same as the fourth-order perturbation method and is *much less* than the Monte Carlo method. On the left of Figure 5.4, the solution profiles of mean and standard deviation are plotted; on the right are the solution PDFs at $x = 0.7$ and $x = 0.8$, obtained by the Legendre-chaos and Monte Carlo simulation with 10,000 realizations. We observe good agreement between the Legendre-chaos solution and Monte Carlo solution. More details of computations of stochastic Burgers' equation can be found in [72], where high resolution numerical solutions convergent to up to seven significant digits are presented.

6. Perturbation Expansion and Sparse Grids. It has been mentioned that in subsurface flow models perturbation expansions are often used to obtain statistics

[1] For this problem, we only need to conduct one simulation corresponding to the maximum input of $\delta = 0.1$ to determine the maximum output. In general, however, such monotonic dependence between input and output does not exist, and a systematic search in the input range is needed to locate the maximum response.

Method	$(\bar{z},\ \sigma_z)$	Cost (unit)
Interval analysis	N/A	≤ 2
First-order perturbation method	$(0.823, 0.349)$	~ 2
Fourth-order perturbation method	$(0.824, 0.328)$	~ 5
Fourth-order Legendre-chaos	$(0.814, 0.414)$	~ 5
Monte Carlo simulation	$(0.814, 0.414)$	$\sim 10,000$

TABLE 5.1

Stochastic solutions and computational cost of different methods for Burgers' equation with uncertain boundary condition. (One unit of cost corresponds to the cost of one deterministic simulation.)

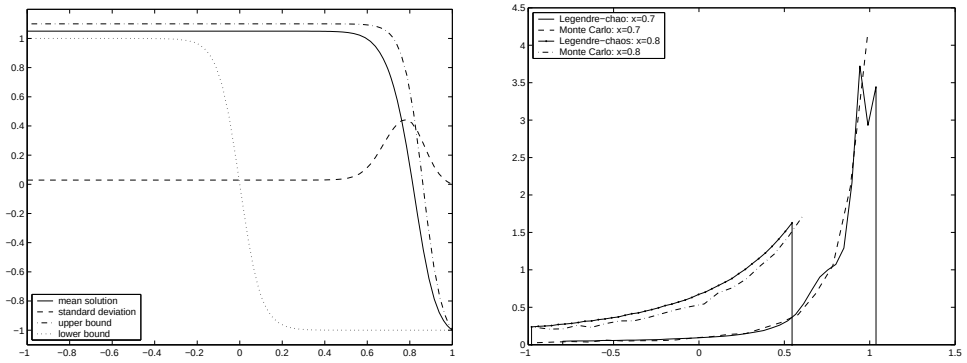

FIG. 5.4. *Stochastic solutions by Legendre-chaos with* $\delta \sim U(0, 0.1)$ *and* $\nu = 0.05$. *Left: Solution profiles of mean and variance. The upper and lower bounds are the deterministic solutions corresponding to the bounds of the random inputs,* $\delta = 0.1$ *and* $\delta = 0$, *respectively. Right:Probability density functions at* $x = 0.7$ *and* $x = 0.8$.

of the random solution. These techniques are different from gPC approaches since they work via *moments* of the random solution: for

$$u(x, \omega) : D \times \Omega \to \mathbb{R}$$

we define the *moment of order* k by

$$\mathcal{M}_u^k : D^k = D \times D \times \cdots \times D \to \mathbb{R}$$
$$\mathcal{M}_u^k(x_1, x_2, \ldots, x_k) := \int_\Omega u(x_1, \omega) u(x_2, \omega) \cdots u(x_k, \omega)\, dP(\omega)$$

Assuming in (5.5) the coefficient κ to be deterministic and the source term $f(x, \omega)$ to be stochastic, we have a *deterministic* problem of order $2k$ for $\mathcal{M}_u^k$:

To define it, we require *anisotropic Sobolev spaces* of mixed highest derivatives:

$$H^{\mathbf{s}}(D^k) \quad := \quad H^s(D) \otimes H^s(D) \otimes \cdots \otimes H^s(D) \quad \forall s \in [-1, \infty[$$

$$H_0^{\mathbf{1}}(D^k) \quad := \quad H_0^1(D) \otimes H_0^1(D) \otimes \cdots \otimes H_0^1(D)$$

e.g. $H^{\mathbf{1}}(D^2) \simeq \{v(x, y) \in L^2(D \times D) \mid \nabla_x v, \nabla_y v, \nabla_x \nabla_y v \in L^2\}$ and define on these spaces the following product linear operators

$$\Delta_\kappa := \operatorname{div}(\kappa \nabla) : H_0^1(D) \to H^{-1}(D),$$

$$\Delta_\kappa \otimes \Delta_\kappa \otimes \cdots \otimes \Delta_\kappa : H_0^{\mathbf{1}}(D^k) \to H^{-\mathbf{1}}(D^k).$$

Then, a deterministic equation for $\mathcal{M}_u^k$ in terms of $\mathcal{M}_f^k$ is

$$(-1)^k(\Delta_\kappa \otimes \Delta_\kappa \otimes \cdots \otimes \Delta_\kappa)\mathcal{M}_u^k = \mathcal{M}_f^k \quad \text{in } H^{-\mathbf{1}}(D^k)$$

This deterministic problem for the kth moment of u in terms of the kth moment of f shows that we can trade, as in gPC, randomness in the boundary value problem at the expense of high dimension in the problem domain. In [48, 49], it was shown that *sparse grids* can be used to solve the high dimensional kth moment problems in work essentially proportional to that of the mean field equation, i.e. for $k = 1$.

The fast numerical solution of the kth moment equations with any $k > 1$ by sparse grid techniques is also significant for problems with stochastic coefficient $\kappa(x, \omega)$ in the context of a perturbation approach. The computation proceeds in terms of the moments of the data (κ, f) and gives directly the moments of the solution u. We present here the mean field computation algorithm (see e.g. [56] for more details).

The idea is now, as in the widely used first order - second moment approach, to use a decomposition

$$\kappa(x, \omega) = e(x) + r(x, \omega) = \text{'deterministic expectation'} + \text{'random fluctuation'},$$

where the fluctuation is assumed to be smaller than the expectation,

$$0 < \frac{\|r\|_{L^\infty(D\times\Omega)}}{\inf_{x\in D} e(x)} < 1,$$

and analytic in the physical domain $D \subset [-1, 1]^d$ (this condition can be relaxed to only finite Sobolev regularity),

$$r \in \mathcal{A}([-1, 1]^d, L^\infty(\Omega)).$$

The analyticity assumption (which is satisfied e.g. if the covariance of κ is Gaussian) ensures the existence of a fluctuation representation as a fast convergent series separating the deterministic and stochastic variables,

$$r(x, \omega) = \sum_{m\in\mathbb{N}_+} \phi_m(x)X_m(\omega), \quad \|\phi_m \otimes X_m\|_{L^\infty(D\times\Omega)} \lesssim e^{-O(m^{1/d})}.$$

Note that such a representation is of course not unique; for example, one can choose e.g. the Karhunen-Loève expansion of r, or $(\phi_m)_{m\in\mathbb{N}_+}$ to be the Legendre polynomials in $[-1, 1]^d$. In both cases the deterministic part $(\phi_m)_{m\in\mathbb{N}_+}$ can be computed explicitly and statistical information on the stochastic part $(X_m)_{m\in\mathbb{N}_+}$ follows by testing r against $(\phi_m)_{m\in\mathbb{N}_+}$,

$$X_m(\omega) = \int_\Omega r(x, \omega)\phi_m(x)\, dP(\omega).$$

Choosing a truncation order M for the fluctuation series (r_M denotes the truncated fluctuation) and a finite element (FE) discretization level L (corresponding to the FE space V_L of dimension N_L), we define recursively, with $\Delta_e = \operatorname{div}(e\nabla)$, the sequence $(u_{m,L})_{m\in\mathbb{N}}$ by

$$\begin{aligned} -\Delta_e u_{0,L}(\cdot, \omega) &= f(\cdot, \omega) \\ -\Delta_e u_{m,L}(\cdot, \omega) &= \operatorname{div}(r_M(\cdot, \omega)u_{m-1,L}(\cdot, \omega)) \quad \forall m \geq 1 \end{aligned} \qquad \text{in } V_L^*, \quad P\text{-a.e. } \omega \in \Omega.$$

The Neumann series $\sum_{m\in\mathbb{N}} u_{m,L}$ is then easily shown to converge uniformly in ω like a geometric series. Moreover, assuming the data f to be sufficiently regular in D as to ensure FE convergence rate $\Phi(N_L)$ uniformly in $\omega\in\Omega$, we have

$$\|\mathcal{M}_u^1 - \sum_{j=0}^{n} \mathcal{M}^1_{u_{j,L}}\|_{H_0^1(D)} \lesssim e^{-O(M^{1/d})} + \Phi(N_L) + e^{-O(n)}.$$

In order to balance the contributions of the three discretization steps (truncation of fluctuation series, FE in D, truncation of Neumann series) and achieve an accuracy of order $\varepsilon > 0$ we choose

$$M \sim |\log\varepsilon|^d, \quad L \text{ s.t. } \Phi(N_L) \le \varepsilon, \quad n \sim |\log\varepsilon|.$$

However, the computation of $\mathcal{M}^1_{u_{j,L}}$ for $0 \le j \le n$ is not an obvious task. It can be shown that $\mathcal{M}^1(u_{j,L})$ can be computed *exactly* using any deterministic solver in D at FE discretization level L, starting from the *mixed moment*

$$\mathcal{M}^{(j,1)}_{r_M,f}(\mathbf{x},x) := \int_\Omega r_M(x_1,\omega)r_M(x_2,\omega)\cdots r_M(x_j,\omega)f(x,\omega)\,dP(\omega) \quad \mathbf{x}\in D^j, x\in D$$

via a composition of $(j+1)$ bounded operators defined in terms of div, ∇, the trace operator on the diagonal set in $D\times D$ and the FE solution operator $\Delta_{e,L}^{-1}$ of the diffusion problem with coefficient e at level L,

$$\begin{array}{ccccccccc}
\mathcal{A}(D^j, H^{-1}(D)) & \xrightarrow{\mathrm{Id}_{L^\infty(D^j)}\otimes\Delta_{e,L}^{-1}} & \mathcal{A}(D^j,V_L) & \to & \mathcal{A}(D^{j-1},V_L) & \to & \cdots & \to & V_L \\
\mathcal{M}^{(j,1)}_{r_M,f} & \xrightarrow{\mathrm{Id}_{L^\infty(D^j)}\otimes\Delta_{e,L}^{-1}} & \mathcal{M}^{(j,1)}_{r_M,u_{0,L}} & \to & \mathcal{M}^{(j-1,1)}_{r_M,u_{1,L}} & \to & \cdots & \to & \mathcal{M}^1_{u_{j,L}}
\end{array}$$

The alternative representation of the $(j,1)$ mixed moment of (r_M,f),

$$\mathcal{M}^{(j,1)}_{r_M,f} = \sum_{\substack{1\le m_k\le M\\ 1\le k\le j}} \underbrace{\left(\int_\Omega X_{m_1}(\omega)\cdots X_{m_j}(\omega)f(x,\omega)\,dP(\omega)\right)}_{\in H^{-1}(D)} \underbrace{\phi_{m_1}\otimes\cdots\otimes\phi_{m_j}}_{\in L^\infty(D^j)}$$

is here processed and transformed into $\mathcal{M}^1_{u_{j,L}}$ at a cost which is essentially equal to the computational effort needed to run the first step for $j=n$ (definition of the the following j operators in the diagram above thus omitted). Taking into account the moment symmetry, the resulting algorithm consists therefore in solving as many diffusion problems at FE discretization level L as terms in the alternative representation of the mixed moment of (r_M,f), that is,

$$M^n/n! \sim \varepsilon^{-O(1)}|\log\varepsilon|^{O((d-1)|\log\varepsilon|)} \text{ deterministic problems in } D \text{ with accuracy } \varepsilon.$$

This super-algebraic complexity (which renders the algorithm inefficient, compared to a standard Monte Carlo simulation, see discussion below) can be substantially reduced by noting that not *exact* but only *approximate*, $\varepsilon-$accurate knowledge of $\mathcal{M}^1_{u_{j,L}}$ for $0\le j\le n$ is required. Based on the fast decay of the fluctuation series, many terms in the alternative representation of the mixed moment of (r_M,f) can be discarded without affecting the computation accuracy. By doing so the complexity is reduced to a *nearly optimal* one, that is to solving

$$\sim |\log\varepsilon|^{O(|\log\varepsilon|^{d/(d+1)})} \lesssim \varepsilon^{-o(1)} \text{ deterministic problems in } D \text{ with accuracy } \varepsilon.$$

Note the higher efficiency of the perturbation algorithm as contrasted with the *Monte Carlo* method. Sampling κ, f requires knowledge of the data distribution *everywhere in* D and solving for each sample the diffusion problem at discretization level L produces an

$$\text{accuracy } \sim \Phi(N_L) + \frac{1}{\sqrt{|\Omega_0|}},$$

where Ω_0 denotes the (finite) sample set. The complexity of the Monte Carlo method is therefore as high as that of solving

$$\sim \varepsilon^{-2} \text{ deterministic problems in } D \text{ with accuracy } \varepsilon.$$

The *spatial regularity* of the random fluctuation can be therefore used to construct alternative algorithms which perform better than the *Monte Carlo* simulation.

We conclude by noting that the perturbation algorithm presented above combined with the sparse grid solution algorithm for the higher order moment problem make possible also the computation of $\mathcal{M}_u^k$ for $k \geq 2$ in *nearly optimal* complexity. For details, we refer to [56].

7. Open Issues.

7.1. Mathematical framework. So far, we presented the variational formulation underlying stochastic Galerkin projections *formally*, and in particular *assuming* the existence of a probability space $(\Omega, \mathcal{A}, \mathcal{P})$ on the function spaces V in which data and/or solutions are sought. As is well known since the pioneering work of N. Wiener, however, the construction of a probability measure $\mathcal{P}$ on a function space is nontrivial. One way to build such measures is via Kolmogoroff's construction whereby $\mathcal{P}$ is obtained as suitable limit of product measures on cylinders of increasing dimension. Kolmogoroff's construction is important in the context of gPC discretizations of stochastic PDE, since a gPC discretization is, in a sense, the converse to the limiting process in Kolmogoroff's construction.

In a series of papers, Babuška and his co-workers [4] gave a mathematical framework for gPC approximation of the elliptic boundary value problem with stochastic coefficients, see Section 5.2, for homogeneous boundary conditions. The diffusion coefficients are represented by a truncated KL expansion at sufficiently high-order N. The error incurred in truncating KL expansions of input data on the accuracy of the solutions of stochastic elliptic equations can be estimated using standard numerical analysis of elliptic boundary value problems, in particular, the Strang Lemma. In accordance with the remarks made at the end of Section 2, and as illustrated in the examples in Sections 4 and 5, gPC discretization is fairly independent of the type of differential equation under consideration. Combining continuous dependence results of solutions on the coefficients of parabolic and hyperbolic PDEs with decay estimates on KL expansions should therefore allow justification of the gPC approximations of *linear* differential equations presented in Sections 4 and 5 above.

7.2. Stochastic regularity and short correlation length. A key step in the error analysis and the understanding of the performance of gPC approximations of stochastic differential equations is the *stochastic regularity* of the solution. The principle underlying efficiency of gPC approximations for *linear* stochastic differential equations is that this dependence on ξ is *analytic*; this, in turn, implies exponential convergence rates $\exp(-bP)$ where $b > 0$ and P denotes the degree of the gPC approximation.

Even with exponential convergence, gPC is computationally effective only if the number N of random variables $\underline{\xi}$ used in the computation is small: for example, if a tensor product gPC space is used, the number M of deterministic problems to be solved after gPC discretization is $M = O(P^N)$ and the exponential gPC convergence expressed in terms of M is $O(\exp(-bM^{1/N}))$. Smallness of N is, therefore, essential; unfortunately, the size of N necessary for good approximation of the stochastic input data is closely related to the decay of the eigenvalues of the covariance operator for the input data; we have given in Section 3 some theoretical estimates for this eigenvalue decay – in particular, for many covariances exponential eigenvalue decay can be achieved.

Computational practice has shown, in accordance with the above heuristic considerations, even exponential gPC convergence rates to be useless if the decay rate in the KL expansion of the input data is too small. Slow KL eigenvalue decay occurs if the spatial correlation length of input data and solution are small compared to the diameter Λ of the physical domain D. This implies an excessively high dimension N of the vector ξ in the gPC discretization and, hence, of the gPC computational domain.

One question of stochastic regularity pertinent to the efficiency of gPC approximation is, therefore, how the correlation length of the input data influences that of the solution; it appears that at least for linear problems, absence of small correlation length in input data ensures this in the random solution. For nonlinear problems the mechanisms which produce singularities (as, e.g., shocks, singularities, etc.) in the deterministic case appear also to generate random solutions with small spatial correlation length, even if the input random fields are not of this type (see [18] for numerical results in this direction). This *problem of small correlation length in input data and solution* is, at present, a major bottleneck in the development of robust and generally applicable gPC-based deterministic solvers for stochastic differential equations.

Based on the above discussion, we see two approaches to tackle the small correlation length problem:

(a) 'coarsening' of the tensor product gPC subspaces; the most direct way to achieve this is to allow for *variable stochastic polynomial degrees* $P_i, i = 1, \ldots, N$. This was used e.g. in [13] and also in [4] and allows to handle (on parallel architectures) expansion orders up to $N = 50$. Numerical experiments appear to indicate, however, that even product spaces with variable polynomial degrees are 'too generous'; if and how further savings can be realized is the topic of ongoing work.

(b) abandoning the KL-type separation of stochastic and deterministic variables – this is rather speculative at present and may be computationally infeasible.

Although analysis of gPC for linear elliptic problems is "relatively" complete, in the sense that exponential convergence under sufficient smoothness assumption has been proved, much remains to be done for nonlinear problems.

7.3. Other expansion basis. The gPC expansion is a p-type expansion that enjoys fast convergence whenever the solution is smooth. However, for some applications, the probability function of the solution may not have sufficient smoothness. Subsequently, the convergence of the p-type expansion deteriorates, as it is a global polynomial approximation in the random space. To circumvent the difficulty, one could use piecewise polynomials. Such approach, termed as k-type expansion, was first proposed in [4]. A similar method, also based on the idea of local approximations, is presented in [27], where a wavelet basis is analyzed. Also, in [59] a multi-element

decomposition of the random space is developed and gPC expansions are employed on each element. An extension of the Wiener-Askey expansions to *arbitrary* pdfs was presented in [60] where orthogonal polynomials were constructed on-the-fly to represent optimally an arbitrary probability distribution. This extension, in fact, is the key to multi-domain random decomposition, otherwise we are only limited to uniform pdfs and correspondingly Legendre polynomials. Another construction of basis worth mentioning is the "double-orthogonal" basis discussed in [4]. When random inputs are represented as KL expansions, the double-orthogonal basis can effectively diagonalize the matrix resulting from the stochastic Galerkin projection. The matrix in doubly orthogonal basis will be block diagonal and can be solved in same complexity as Monte Carlo methods.

7.4. Collocation methods. In addition to the Galerkin methods discussed in this paper, one can adopt a collocation approach. Such an approach typically requires repetitive runs of a deterministic solver on a set of prescribed (deterministic) nodal points and results in a completely uncoupled system of equations. Hence, its implementation is trivial and is parallelization is rather straightforward. The choice of nodal points is critical; although in one random dimension ($N = 1$) there are abundant choices, e.g., quadrature points of orthogonal polynomials ([38, 70]), the cases of multi-dimensional random spaces ($N \gg 1$) are more difficult. In [68], high-order collocation methods are proposed and analyzed. It is shown that the nodal set based on tensor products of one-dimensional nodal sets is inappropriate in high-dimensional random spaces (due to its fast growth of total number of points), and the sets based on Smolyak sparse grid and Stroud's curvature are much more efficient. The collocation method based on the Smolyak sparse grid is able to achieve high-order convergence. Although its total number of nodal points is higher than the number of basis functions of a gPC expansion, the sparse grid collocation method should be investigated as a potential alternative to Galerkin methods.

7.5. The "Tails". The assumption that random inputs follow a Gaussian distribution is employed in many applications, but it is obviously inappropriate for random fields which exhibit significant skewness. For example, the coefficient in the diffusion equation may follow a *lognormal* distribution. Sign properties of input data that are crucial for well-posedness of boundary value problems may be lost in the process of KL truncation, resulting in ill-posedness of the deterministic problems with small probability in gPC approximation. For example, unsolvability of systems resulting from Hermite-chaos expansion for order P of expansions higher than a critical order was observed in numerical experiments [67]. A remedy proposed in [67] was the use of a "truncated" Gaussian distribution. Such a model is based on a truncated Jacobi-chaos expansion. Its distribution approximates a Gaussian closely, as shown in Figure 7.1, and it has bounded support, thereby allowing control of the undershooting which causes ill-posedness.

It has also been shown that the Karhunen-Loève expansion may develop long tails as one increases the number of terms, even though all the random variables in the expansion have bounded support. This will cause the underlying problem, e.g., elliptic equation, to be singularly perturbed, see [3]. It is yet unclear how to resolve this issue from a mathematical point of view, although, in practice, one can control the number of expansion terms to keep the constructed random process well bounded.

7.6. Long-term integration. The results in Section 4 shows that the error of generalized polynomial chaos may grow for long-term integrations. Such observation

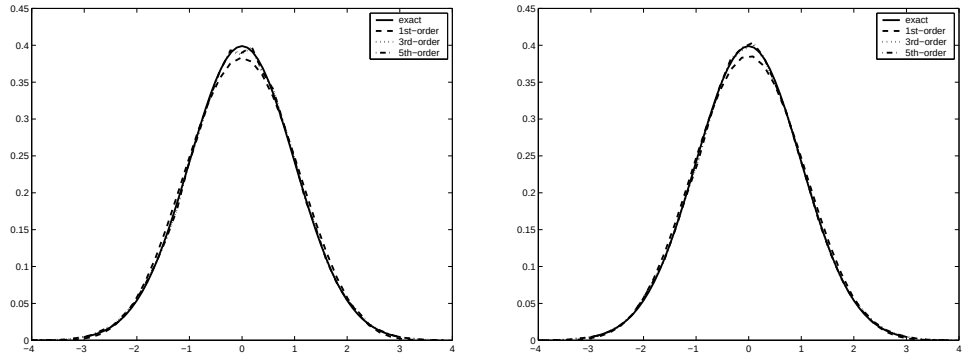

FIG. 7.1. *Approximated Gaussian random variables by Jacobi-chaos; Left:* $\alpha = \beta = 8$, *Right:* $\alpha = \beta = 10$.

presents a challenge on the control of error for time-dependent problems. A possibly remedy may be to regularly change the parameters ξ in terms of which the gPC approximation is sought. This change of parameters can be achieved efficiently from an estimated covariance kernel R_h by the fast approximation of the KL eigenvalues. The long-term integration issue has been addressed in [59], where an adaptive multi-domain gPC approach was developed; specifically, more elements are introduced selectively in order to suppress the fast growth of error as the interval of time integration increases.

7.7. Stochastic discontinuities. It is well known that polynomial chaos fails in a short time for the so-called Kraichnan-Orszag three-mode problem, constructed as a simplified model to represent turbulence interactions [42]. It is a nonlinear three-dimensional ODE system:

$$\frac{dx_1}{dt} = x_2 x_3, \quad \frac{dx_2}{dt} = x_1 x_3, \quad \frac{dx_3}{dt} = -2x_1 x_2$$

subject to stochastic initial conditions. Standard polynomial chaos fails to predict the solution, and in fact truncated expansions may converge to erroneous solutions. The main reason for this is that the frequency of the solution depends on the random variable and also that there is a discontinuity in the initial conditions. Using an adaptive multi-element gPC approach, however, based on the local variance can effectively resolve this problem, see [59]. In figure 7.2(left), we show the evolution of the variance of x_1 using this multi-element approach for different values of an adaptive threshold θ_1. For comparison we include the results given by gPC with polynomial order $p = 30$. It can be seen that comparing to the results given by Monte Carlo with $1,000,000$ realizations, gPC with polynomial order $p = 30$ begins to lose accuracy at $t \approx 8$ and fails beyond this point while the multi-element gPC converges as θ_1 decreases. In figure 7.2(right), we show the corresponding adaptive mesh. We can see that around the point $\xi = 0$ in random space of ξ, where the discontinuity occurs, the random elements are smallest; this means that the discontinuity can be captured effectively by small random elements. These results hold for the Kraichnan-Orszag problem with a discontinuity in only one variable, namely x_1. In three-dimensions, however, the number of elements required to control accuracy is quite high so this multi-element adaptive refinement approach becomes computationally expensive, see [59].

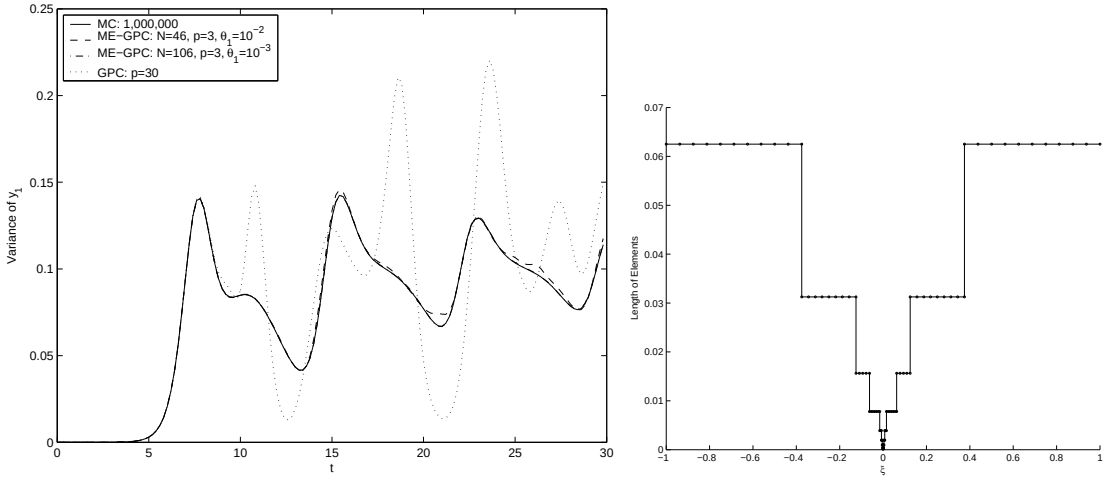

FIG. 7.2. *Evolution of variance (left) and length of elements in the adaptive mesh (right). Refinement occurs in the region around the discontinuity. GPC fails to converge but adaptive ME-GPC converges fast. (Courtesy of Xiaoliang Wan)*

7.8. Computational complexity. Generalized polynomial chaos can be orders of magnitude more efficient than sampling methods such as Monte Carlo method. A comparison of these two approaches for a second-order oscillator is shown in Table 7.1. A similar comparison is also shown in Table 7.2, where we present the computational cost for a two-dimensional diffusion equation with four-dimensional random inputs. The appropriate gPC basis is employed according to the type of input distributions, and all gPC expansions have $M+1 = 35$ terms at third-order $N = 4, P = 3$. Note here this comparison is approximate because the number of realizations in Monte Carlo simulations is determined in such a way that their solutions are reasonably close to the gPC solutions. For details of the computational results, see [69].

Although gPC can be quite efficient, it is still orders of magnitude more expensive than methods for deterministic problems. Massively parallel solution approaches are, therefore, necessary and future focus should be on development of fast and efficient solvers. An important scientific computing issue in this context is thus the clever choice of basis functions in the gPC discretization. The doubly orthogonal polynomials in [3] which perform very well for linear problems, the pseudospectral techniques of [7] to evaluate efficiently nonlinear functionals of gPC expansions, and the adaptive multi-element approaches in [26, 59] are attempts to reduce computational complexity.

REFERENCES

[1] *Decision making under uncertainty*, in Series: The IMA Volumes in Mathematics and its Applications, C. Greengard and A. Ruszczynski, eds., vol. 128, Springer-Verlag, 2002.

[2] R. ASKEY AND J. WILSON, *Some basic hypergeometric polynomials that generalize Jacobi polynomials*, Memoirs Amer. Math. Soc., AMS, Providence RI, 319 (1985).

[3] I. BABUŠKA AND P. CHATZIPANTELIDIS, *On solving elliptic stochastic partial differential equations*, Comput. Methods Appl. Mech. Engrg., 191 (2002), pp. 4093–4122.

[4] I. BABUŠKA, R. TEMPONE, AND G. ZOURARIS, *Galerkin finite element approximations of stochastic elliptic differential equations*, Tech. Report 02-38, TICAM, 2002.

Speed-up factors				
	$\epsilon_{\mathbf{mean}}$	Monte-Carlo: **N**	Generalized Polynomial Chaos: **P+1**	**S**
Gaussian	2%	350	56	6.25
	0.8%	2,150	120	18
	0.2%	33,200	220	151
Uniform	0.2%	13,000	10	13,000
	0.018%	1,580,000	20	79,000
	0.001%	610,000,000	35	17,430,000

TABLE 7.1

Speed-up factors $\mathcal{S}$ based on relative mean *error ϵ_{mean} of generalized Polynomial Chaos ($(P+1)$ terms) versus Monte-Carlo simulations (N events) for* Gaussian *and* Uniform *distributions.*

input distributions	uniform	Gaussian	Poisson	binomial
σ	0.4	0.2	0.2	0.2
Monte Carlo	50,000	20,000	100,000	50,000

TABLE 7.2

Cost comparison of a two-dimensional ($d = 2$) diffusion equation with four-dimensional ($N = 4$) random inputs with different types of distributions. gPC using 3rd-order expansion ($P = 3$) or equivalently $M + 1 = 35$ expansion terms.

[5] C. BAKER, *The numerical treatment of integral equations*, Oxford University Press, London, 1977.

[6] R. CAMERON AND W. MARTIN, *The orthogonal development of nonlinear functionals in series of Fourier-Hermite functionals*, Ann. Math., 48 (1947), pp. 385–392.

[7] B. DEBUSSCHERE, H. NAJM, P. PEBAY, O. KNIO, R. GHANEM, AND O. LE MAITRE, *Numerical challenges in the use of polynomial chaos representations for stochastic processes*, SIAM Journal of Scientific Computing, to appear (2005).

[8] G. DEODATIS, *Weighted integral method. I: stochastic stiffness matrix*, J. Eng. Mech., 117 (1991), pp. 1851–1864.

[9] G. DEODATIS AND M. SHINOZUKA, *Weighted integral method. II: response variability and reliability*, J. Eng. Mech., 117 (1991), pp. 1865–1877.

[10] P. ESPANOL AND P. WARREN, *Statistical mechanics of dissipative particle dynamics*, Europhys. Lett., 30 (1995), pp. 191–196.

[11] G. FISHMAN, *Monte Carlo: Concepts, Algorithms, and Applications*, Springer-Verlag New York, Inc., 1996.

[12] B. FOX, *Strategies for Quasi-Monte Carlo*, Kluwer Academic Pub., 1999.

[13] P. FRAUENFELDER, C. SCHWAB, AND R. TODOR, *Finite element methods for elliptic problem with stochastic coefficients*, Comput. Meth. Appl. Mech. Engr., in press (2004).

[14] L. GAMMAITONI, P. HANGGI, P. JUNG, AND F. MARCHESONI, *Stochastic resonance*, Rev. Modern Phys., 70 (1998), pp. 223–288.

[15] C. GARDINER, *Handbook of stochastic methods: for physics, chemistry and the natural sciences*, Springer-Verlag, 2nd ed., 1985.

[16] R. GHANEM AND P. SPANOS, *Stochastic Finite Elements: a Spectral Approach*, Springer-Verlag, 1991.

[17] J. GLIMM AND D. SHARP, *Prediction and the quantification of uncertainty*, Physica D, 133 (1999), pp. 152–170.

[18] T. HOU, H. KIM, B. ROZOVSKII, AND H. ZHOU, *Wiener chaos expansions and numerical solutions of randomly forced equations of fluid mechanics*, Preprint, (2004).

[19] S. HUANG, S. QUEK, AND K. PHOON, *Convergence study of the truncated Karhunen-Loeve expansion for simulation of stochastic processes*, Int. J. Numer. Meth. Engng., 52 (2001), pp. 1029–1043.

[20] M. JARDAK, C.-H. SU, AND G. KARNIADAKIS, *Spectral polynomial chaos solutions of the stochastic advection equation*, J. Sci. Comput., 17 (2002), pp. 319–338.

[21] I. KARATZAS AND S. SHREVE, *Brownian motion and stochastic calculus*, Springer-Verlag, 1988.

[22] G. Karniadakis and S. Sherwin, *Spectral/hp Element Methods for CFD*, Oxford University Press, 1999.
[23] M. Kleiber and T. Hien, *The stochastic finite element method*, John Wiley & Sons Ltd, 1992.
[24] P. Kloeden and E. Platen, *Numerical solution of stochastic differential equations*, Springer-Verlag, 1999.
[25] R. Koekoek and R. Swarttouw, *The Askey-scheme of hypergeometric orthogonal polynomials and its q-analogue*, Tech. Report 98-17, Department of Technical Mathematics and Informatics, Delft University of Technology, 1998.
[26] O. Le Maitre, O. Knio, H. Najm, and R. Ghanem, *A stochastic projection method for fluid flow: basic formulation*, J. Comput. Phys., 173 (2001), pp. 481–511.
[27] ———, *Uncertainty propagation using Wiener-Haar expansions*, J. Comput. Phys., 197 (2004), pp. 28–57.
[28] P.-L. Liu and A. Der Kiureghian, *Finite element reliability of geometrically nonlinear uncertain structures*, J. Eng. Mech., 117 (1991), pp. 1806–1825.
[29] W. Liu, T. Belytschko, and A. Mani, *Applications of probabilistic finite element methods in elastic/plastic dynamics*, J. Engrg. Ind., ASME, 109 (1987), pp. 2–8.
[30] M. Loève, *Probability Theory, Fourth edition*, Springer-Verlag, 1977.
[31] W. Loh, *On Latin hypercube sampling*, Ann. Stat., 24 (1996), pp. 2058–2080.
[32] J. Lorentz, *Nonlinear singular perturbation problems and the Engquist-Osher difference scheme*, Tech. Report 8115, University of Nijmegen, Nijmegen, The Netherlands, 1981.
[33] D. Lucor and G. Karniadakis, *Noisy inflows cause a shedding-mode switching in flow past an oscillating cylinder*, Phys. Rev. Lett., 92(15) (2004), p. 154501.
[34] D. Lucor, C.-H. Su, and G. Karniadakis, *Generalized polynomial chaos and random oscillators*, Int. J. Numer. Meth. Engng., 60 (2004), pp. 571–596.
[35] ———, *Karhunen-Loeve representation of periodic second-order auto-regressive processes*, in Proceedings of International Conference on Computational Science, June 6-9, Krakow, Poland, 2004.
[36] P. Maciejewski and R. Moffat, *Heat transfer with very high free-stream turbulence. Part I - Experimental data*, Journal of Heat Transfer, 114 (1992), pp. 827–833.
[37] N. Madras, *Lectures on Monte Carlo methods*, American Mathematical Society, Providence, RI, 2002.
[38] L. Mathelin and M. Hussaini, *A stochastic collocation algorithm for uncertainty analysis*, Tech. Report NASA/CR-2003-212153, NASA Langley Research Center, 2003.
[39] W. Oberkampf, T. Trucano, and C. Hirsch, *Verification, validation, and predictive capability in computational engineering and physics*, Tech. Report SAND2003-3769, Sandia National Laboratories, 2003.
[40] H. Ogura, *Orthogonal functionals of the Poisson process*, IEEE Trans. Info. Theory, IT-18 (1972), pp. 473–481.
[41] B. Oksendal, *Stochastic differential equations. An introduction with applications*, Springer-Verlag, fifth ed., 1998.
[42] S. Orszag and L. Bissonnette, *Dynamical properties of truncated Wiener-Hermite expansions*, Phys. Fluids, 10 (1967), pp. 2603–2613.
[43] K. Phoon, S. Huang, and S. Quek, *Implementation of Karhunen-Loeve expansion for simulation using a wavelet-Galerkin scheme*, Prob. Eng. Mech., 17 (2002), pp. 293–303.
[44] T. Pottebaum, *The relationship between near-wake structure and heat transfer for an oscillating cylinder in cross-flow*, PhD thesis, California Institute of Technology, 2003.
[45] M. Rajaschekhar and B. Ellingwood, *A new look at the response surface approach for reliability analysis*, Struc. Safety, 123 (1993), pp. 205–220.
[46] W. Schoutens, *Stochastic processes in the Askey scheme*, PhD thesis, K.U. Leuven, 1999.
[47] C. Schwab, *p- and hp-Finite Element Methods*, Oxford Science Publications, 1998.
[48] C. Schwab and R. Todor, *Sparse finite elements for elliptic problems with stochastic data*, Numer. Math., 95 (2003), pp. 707–734.
[49] ———, *Sparse finite elements for stochastic elliptic problems-higher order moments*, Computing, 71 (2003), pp. 43–63.
[50] ———, *Karhunen-Loeve approximation of random fields by generalized fast multipole methods*, in preparation, (2004).
[51] T. Shardlow, *Splitting for dissipative particle dynamics*, SIAM J. Sci. Comput., 24 (2003), pp. 1267–1282.
[52] M. Shinozuka and G. Deodatis, *Response variability of stochastic finite element systems*, J. Eng. Mech., 114 (1988), pp. 499–519.
[53] E. Simiu, *Chaotic Transition in Deterministic and Stochastic Systems*, Princeton University Press, 2002.

[54] P. SPANOS AND R. GHANEM, *Stochastic finite element expansion for random media*, J. Eng. Mech., 115 (1989), pp. 1035–1053.
[55] T. TAKADA, *Weighted integral method in stochastic finite element analysis*, Prob. Eng. Mech., 5 (1990), pp. 146–156.
[56] R.A.TODOR, *Numerical analysis of Galerkin FEM for stochastic elliptic PDEs*, PhD thesis, ETHZ, in preparation.
[57] H. VAN TREES, *Detection, estimation and modulation theory, part 1*, Wiley Press, New York, 1968.
[58] E. VANMARCKE AND M. GRIGORIU, *Stochastic finite element analysis of simple beams*, J. Eng. Mech, 109 (1983), pp. 1203–1214.
[59] X. WAN AND G. KARNIADAKIS, *An adaptive multi-element generalized polynomial chaos method for stochastic differential equations*, J. Comput. Phys., submitted (2004).
[60] ———, *Beyond Wiener-Askey expansions: Handling arbitrary PDFs*, J. Sci. Comput., submitted (2004).
[61] ———, *A fast Gauss transform/generalized polynomial chaos solver for 3D random diffusion with high-dimensional stochastic coefficient*, J. Comput. Phys., submitted (2004).
[62] X. WAN, D. XIU, AND G. KARNIADAKIS, *Stochastic solutions for the two-dimensional advection-diffusion equation*, SIAM J. Sci. Comput., in press (2004).
[63] G. WHITESIDES AND B. GRZYBOWSKI, *Self-assembly at all scales*, Science, 295 (2002), pp. 2418–2421.
[64] P. WHITTLE, *On stationary processes in the plane*, Biometrika, 41 (1954), pp. 434–449.
[65] N. WIENER, *The homogeneous chaos*, Amer. J. Math., 60 (1938), pp. 897–936.
[66] C. WILLIAMSON AND A. ROSHKO, *Vortex formation in the wake of an oscillating cylinder*, J. Fluid Mech., 2 (1988), pp. 355–381.
[67] D. XIU, *Generalized (Wiener-Askey) polynomial chaos*, PhD thesis, Brown University, 2004.
[68] D. XIU AND J. HESTHAVEN, *High order collocation methods for differential equations with random inputs*, SIAM J. Sci. Comput., submitted (2004).
[69] D. XIU AND G. KARNIADAKIS, *Modeling uncertainty in steady state diffusion problems via generalized polynomial chaos*, Comput. Methods Appl. Math. Engrg., 191 (2002), pp. 4927–4948.
[70] ———, *The Wiener-Askey polynomial chaos for stochastic differential equations*, SIAM J. Sci. Comput., 24 (2002), pp. 619–644.
[71] ———, *A new stochastic approach to transient heat conduction modeling with uncertainty*, Inter. J. Heat Mass Trans., in press (2003).
[72] ———, *Supersensitivity due to uncertain boundary conditions*, Int. J. Numer. Meth. Engng., in press (2004).
[73] W. ZHU, Y. REN, AND W. WU, *Stochastic FEM based on local average of random vector fields*, J. Eng. Mech., 118 (1992), pp. 496–511.
[74] W. ZHU AND W. WU, *A stochastic finite element method for real eigenvalue problems*, Prob. Eng. Mech., 6 (1991), pp. 228–232.