A Posteriori Error Analysis of Stochastic Differential Equations Using Polynomial Chaos Approximations

Mathematical and Computational Tools for Predictive Simulation of Complex Coupled Systems Under Uncertainty

Paul Constantine, Eric Phipps and Tim Wildey: Sandia National Laboratories, Albuquerque, NM 87123
Troy Butler and Clint Dawson: The University of Texas at Austin, Austin, TX 78723

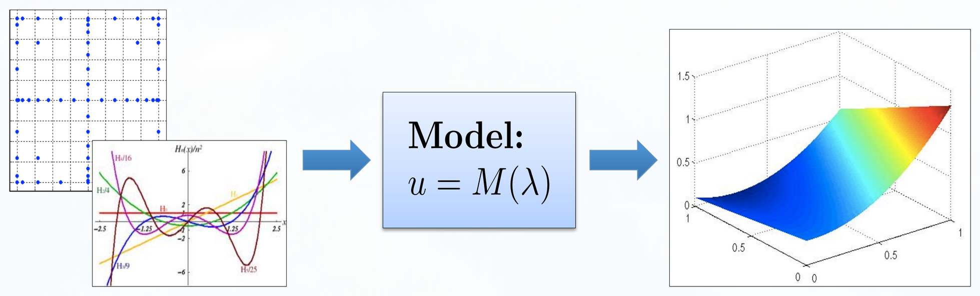
Motivation

We are oftentimes interested in using computationally intense simulations to compute statistical properties (moments, probabilities, etc.) for a quantity of interest (QoI).

Error in $P[q(\lambda) > T]$ = discretization error + sampling error

Do we trust Monte Carlo to compute statistical properties given a small number of samples?

One alternative is to use a surrogate model (polynomial chaos, stochastic collocation, gaussian process, etc.).



Trade off: Smaller sampling error for a larger discretization error.

Do we trust surrogate models to compute statistical properties given (virtually) unlimited samples?

The error in the statistical property can be quantified if we can estimate the error in each sample of the surrogate model.

The goal is to estimate the error in samples of a quantity of interest computed from a polynomial chaos (PC) approximation.

We take the following approach:

- We compute a PC approximation of the forward problem.
- We compute a PC approximation of a properly defined adjoint problem.
- For each sample of the QoI, we sample the adjoint approximation and produce an estimate of the error.

A Posteriori Error Analysis

Let X and Y be Banach spaces and consider $L : X \to Y$.

The adjoint operator $L^* : Y^* \to X^*$ is defined such that $\langle Lx, y^* \rangle = \langle x, L^* y^* \rangle$

Let x solve $Lx = f$, let $\tilde{x} \approx x$ and define $e = x - \tilde{x}$ and $R = f - L\tilde{x}$.

Let ϕ solve the adjoint problem, $L^*\phi = \psi$.

We derive the error representation, $\langle \psi, e \rangle = \langle L^*\phi, e \rangle = \langle \phi, Le \rangle = \langle \phi, R \rangle$.

Not Computable   Computable

The error analysis for nonlinear operators is handled by an appropriate linearization and the extension to systems of equations is straightforward [3].

Example 1: Standard Analysis for a Nonlinear Coupled System

Consider the coupled model for neutron diffusion, u_1 and temperature, u_2,

$$\begin{cases} -\nabla \cdot (D(u_2)\nabla u_1) + (\Sigma_a(u_2) - \nu\Sigma_f(u_2))u_1 = s, & x \in \Omega, \\ u_1 = 0, & x \in \partial\Omega, \\ -\nabla \cdot (K\nabla u_2) + Hu_2 - E_f \nu \Sigma_f(u_2) u_1 = Hu_{2,\infty}, & x \in \Omega, \\ K\nabla u_2 \cdot \mathbf{n} = \mathbf{0}, & x \in \partial\Omega. \end{cases}$$

where

$$\nu\Sigma_f(u_2) = 0.0162\sqrt{\frac{u_{2,\infty}}{u_2}}, \quad \Sigma_a(u_2) = 0.02\sqrt{\frac{u_{2,\infty}}{u_2}}, \quad D(u_2) = 2.2\sqrt{\frac{u_2}{u_{2,\infty}}},$$

Let $u_{h,1}$ and $u_{h,2}$ be finite element approximations to u_1 and u_2 respectively.

We linearize the problem around $\boldsymbol{u}_h = (u_{h,1}, u_{h,2})^T$ to obtain

$$J(\mathbf{u_h})\boldsymbol{\delta} := \begin{cases} \mathcal{L}_{11}(\boldsymbol{u}_h)\delta_1 + \mathcal{L}_{12}(\boldsymbol{u}_h)\delta_2 \\ \mathcal{L}_{21}(\boldsymbol{u}_h)\delta_1 + \mathcal{L}_{22}(\boldsymbol{u}_h)\delta_2, \end{cases}$$

The adjoint operator is given by,

$$J(\boldsymbol{u}_h)^*\boldsymbol{\phi} := \begin{cases} \mathcal{L}_{11}^*(\boldsymbol{u}_h)\phi_1 + \mathcal{L}_{21}^*(\boldsymbol{u}_h)\phi_2 \\ \mathcal{L}_{12}^*(\boldsymbol{u}_h)\phi_1 + \mathcal{L}_{22}^*(\boldsymbol{u}_h)\phi_2, \end{cases}$$

ψ_1	ψ_2	Value	Error	η_1	η_2	Effect.
1/100	0	7.3312E13	5.6029E9	5.3300E9	2.7284E8	0.9999
0	1/100	5.9482E2	1.3470E-2	1.3112E-2	3.5832E-4	1.0000
$p(50;10)$	0	1.0526E14	2.6428E8	−1.6916E6	2.6608E8	1.0004
0	$p(25,10)$	5.9763E2	1.4962E-2	1.2693E-2	2.2690E-3	0.9999
0	$p(65,10)$	6.1128E2	1.3760E-2	1.0739E-2	3.0209E-3	0.9999

Table: The approximate value, error, and contributions to the error from the neutron diffusion residual, η_1, and from the temperature residual, η_2, for a variety of quantities of interest.

Polynomial Chaos Approximations

Let $\{\Omega, \mathcal{F}, P\}$ be a probability space.

Let $Z(\omega)$ be a random variable and let $\{\Phi_i(Z)\}_{i=1}^{\infty}$ be a set of polynomials orthogonal w.r.t. density of Z.

Model parameter as a random variable $\lambda = \Lambda(\omega)$ with finite variance,

$$\Lambda(\omega) = \sum_{i=0}^{\infty} \lambda_i \Phi_i(Z(w)), \quad \text{where } \lambda_i = \frac{\langle \Lambda, \Phi_i \rangle}{\langle \Phi_i, \Phi_i \rangle}.$$

Truncate expansion at order p, giving the total number of terms,

$$P + 1 = \frac{(d+p)!}{d!p!}.$$

Parameterized Linear Systems

Let $x(s) \in \mathbb{R}^n$ solve the parameterized linear system,

$$A(s)x(s) = b(s), \quad s \in \Omega,$$

for a given $A(s) \in \mathbb{R}^n \times \mathbb{R}^n$ and $b(s) \in \mathbb{R}^n$.

Let x_N be a surrogate approximation and define, $e(s) = x(s) - x_N(s)$.

We assume the following point-wise error estimate holds,

$$\|e(s)\|_{L^\infty(\Omega; l^2(\mathbb{R}^n))} \le C\epsilon_1(N)$$

for some $\epsilon_1(N) \ge 0$.

Let $\phi(s)$ solve the adjoint problem,

$$A^T(s)\phi(s) = \psi, \quad \forall s \in \Omega.$$

Lemma [2]

Let ϕ_M be a PC approximation of ϕ.

At each $\hat{s} \in \Omega$ we derive the error representation:

$$\begin{aligned} \langle \psi, e(\hat{s}) \rangle &= \langle R(\hat{s}), \phi(\hat{s}) \rangle \\ &= \langle R(\hat{s}), \phi_M(\hat{s}) \rangle + \langle R(\hat{s}), \phi(\hat{s}) - \phi_M(\hat{s}) \rangle \end{aligned}$$

Computable   Higher Order

Allows us to define an *improved linear functional*,

$$g(x_N(s), \phi_M(s)) = \langle \psi, x_N(s) \rangle + \langle R(s), \phi_M(s) \rangle.$$

Theorem [2]

If the pointwise error in the adjoint solution satisfies,

$$\|\phi(s) - \phi_M(s)\|_{L^\infty(\Omega; l^2(\mathbb{R}^n))} \le \epsilon_2(M),$$

then the pointwise error in the improved linear functional is bounded by,

$$\| \langle \psi, x(s) \rangle - g(x_N(s), \phi_M(s)) \|_{L^\infty(\Omega)} \le C\epsilon_1(N)\epsilon_2(M),$$

where $C > 0$ depends only on $A(s)$.

Example 2: A Discontinuous Quantity of Interest

Consider the parameterized linear system,

$$\begin{bmatrix} 2 & -s_1 \\ -s_2 & 1 \end{bmatrix} \begin{bmatrix} x_1(s) \\ x_2(s) \end{bmatrix} = \begin{bmatrix} 1 \\ \lceil s_3 - 1/3 \rceil \end{bmatrix}$$

where $\lceil \cdot \rceil$ is the ceiling operator and $s_i \in [-1, 1]$.

Quantity of interest is $x_1(s)$.

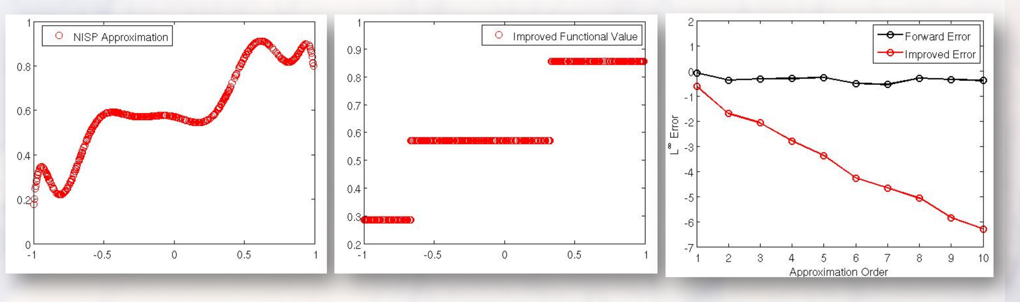


Figure: High-order spectral approximation of the linear functional (left), the improved linear functional (center), and the convergence rates for each (right).

Stochastic Differential Equations

Model for nonlinear stochastic diffusive transport:

$$\begin{cases} \frac{\partial u}{\partial t} - \nabla \cdot (A(x,t,\lambda)\nabla u) + g(x,t;u) = f(x,t,\lambda), & x \in S, 0 < t \le T, \\ A\nabla u \cdot \boldsymbol{n} = 0, & x \in \partial S, 0 < t \le T, \\ u(x,0) = 0, & x \in S, \end{cases}$$

where S is a convex polygonal domain.

Variational formulation for a fixed λ: Find $u \in L^2([0,T]; H^1(S))$ s.t.

$$\int_0^T [(\partial u/\partial t, v)_S + (A(x,t,\lambda)\nabla u, \nabla v)_S + (g(x,t;u), v)_S] \ dt = \int_0^T (f(x,t,\lambda), v)_S \ dt$$

for all $v \in L^2([0,T]; H^1(S))$ with $v(x,0) = 0$.

Seek $u = \sum_{k=0}^P u_k(x,t)\Phi_k(Z)$, such that for $k = 0, 1, \ldots, P$,

$$\begin{aligned} &\int_0^T (\partial u_k/\partial t, v)_S \ dt \\ &+ \frac{1}{\|\Phi_k\|^2} \int_0^T \left(\left\langle A\left(x, t; \sum_{i=0}^P \lambda_i \Phi_i(Z)\right) \sum_{j=0}^P \nabla u_j \Phi_j(Z), \Phi_k \right\rangle, \nabla v \right)_S dt \\ &+ \frac{1}{\|\Phi_k\|^2} \int_0^T \left(\left\langle g\left(x, t; \sum_{j=0}^P u_j \Phi_j\right), \Phi_k \right\rangle, v \right)_S dt \\ &= \int_0^T (f_k(x,t), v)_S \ dt \end{aligned}$$

for all $v \in L^2([0,T]; H^1(S))$.

The strong form of the adjoint for a fixed λ,

$$\begin{cases} -\frac{\partial \phi}{\partial t} - \nabla \cdot \left(A^T(x,t,\lambda)\nabla\phi\right) + \overline{g(u,U;\lambda)}^T \phi = 0, & x \in S, T > t \ge 0, \\ A^T \nabla\phi \cdot \boldsymbol{n} = 0, & x \in \partial S, T > t \ge 0, \\ \phi(x,T) = \psi, & x \in S, \end{cases}$$

where $\overline{g(u,U;\lambda)} = \int_0^1 \partial_u g(x,t; su + (1-s)U) \ ds$.

Lemma [1]

We follow standard steps (substitutions, integration-by-parts, etc.) to derive the error representation:

$$\begin{aligned} (e(T,\lambda), \psi)_S = (e(0;\lambda), \phi(0;\lambda))_S &- \sum_{n=1}^N \int_{I_n} (\partial U(\lambda)/\partial t, \phi(\lambda))_S \ dt \\ &+ \sum_{n=2}^N ([U(\lambda)], \phi(\lambda))_S + \sum_{n=1}^N \int_{I_n} (f - g(U), \phi(\lambda))_S \ dt \\ &- \sum_{n=1}^N \int_{I_n} (A(\lambda)\nabla U(\lambda), \nabla\phi(\lambda))_S \ dt \end{aligned}$$

We approximate ϕ using a PC expansion:

$$\phi(x,t;\lambda) \approx \sum_{i=0}^P \phi_i(x,t)\Phi_i(Z(\omega)).$$

Example 3: Random Source Location

Consider the contaminant source problem:

$$\frac{\partial u}{\partial t} - \nabla \cdot \nabla u = \frac{s}{2\pi\sigma^2} \exp\left(-\frac{|\lambda - x|^2}{2\sigma^2}\right)(1 - H(t - 0.05))$$

with $S = [0,1]^2$, $T = 0.21$, $u(x,0) = 0$, $s = 10$ and $\sigma = 0.1$.

Random variable λ uniformly distributed on $[0,1]^2$.

Discretization: $h = 0.1$, $\Delta t = 0.005$ and 6^{th}−order PC expansion.

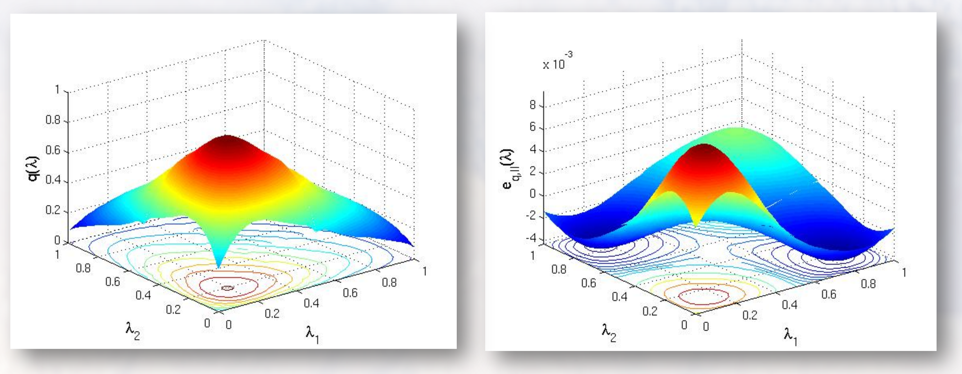


Figure: Polynomial chaos approximation of the quantity of interest (left) and the a posteriori error estimate (right).

Time	λ	Std Err Est $u(x^{\{1\}}, t)$	PC Err Est $u(x^{\{1\}}, t)$	Ratio
0.05	(0.25, 0.25)	$-1.094E-02$	$-1.207E-02$	1.103
0.05	(0.75, 0.25)	$2.142E-03$	$2.144E-03$	1.001
0.05	(0.25, 0.75)	$2.347E-03$	$2.348E-03$	1.001
0.05	(0.75, 0.75)	$1.439E-03$	$1.466E-03$	1.019
0.05	(0.4, 0.375)	$4.273E-03$	$4.508E-03$	1.055
0.15	(0.25, 0.25)	$5.754E-03$	$5.812E-03$	1.010
0.15	(0.75, 0.25)	$-3.637E-03$	$-3.670E-03$	1.009
0.15	(0.25, 0.75)	$-3.511E-03$	$-3.553E-03$	1.012
0.15	(0.75, 0.75)	$1.444E-03$	$1.4376E-03$	0.996
0.15	(0.4, 0.375)	$7.686E-05$	$9.389E-05$	1.222

Table: Comparison of the traditional error estimate with the error estimate using the polynomial chaos approximation of the adjoint.

Example 4: Random Permeability

Consider the contaminant source problem:

$$\frac{\partial u}{\partial t} - \nabla \cdot A(x,t;\lambda)\nabla u = \frac{s}{2\pi\sigma^2} \exp\left(-\frac{|\overline{x} - x|^2}{2\sigma^2}\right)(1 - H(t - 0.05))$$

with $S = [0,1]^2$, $T = 0.21$, $u(x,0) = 0$, $s = 10$ and $\sigma = 0.1$.

$$A(x,t;\lambda) = \begin{pmatrix} \lambda\exp(2\sin(2\pi x)\cos(4\pi y)) & 0 \\ 0 & \exp(2\sin(4\pi y) + 2\cos(2\pi x)) \end{pmatrix}$$

Random variable λ uniformly distributed on $[0.5, 1.5]$.

Discretization: $h = 0.1$, $\Delta t = 0.005$ and 6^{th}−order PC expansion.

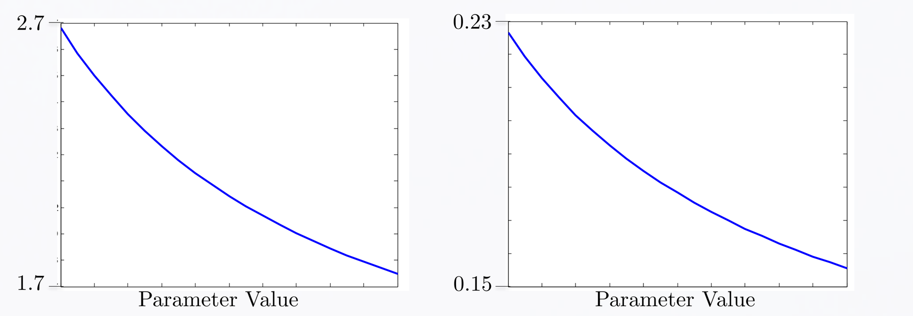


Figure: Polynomial chaos approximation of the quantity of interest (left) and the a posteriori error estimate (right).

λ	Std Err Est $u(x^{\{5\}}, t)$	PC Err Est $u(x^{\{5\}}, t)$	Ratio
0.50	0.22660	0.22667	1.00032
0.75	0.19693	0.19694	1.00006
1.00	0.17823	0.17823	1.00000
1.25	0.16520	0.16519	0.99996
1.50	0.15550	0.15548	0.99983

Table: Comparison of the traditional error estimate with the error estimate using the polynomial chaos approximation of the adjoint.

Conclusions / Future Work

- Statistical properties computed using numerical models have error due to discretizations and sampling.
- High-fidelity models have reduced discretization error, but fewer samples can be taken.
- Surrogate models constructed from high-fidelity simulations can be cheaply sampled, but have larger discretization error.
- A posteriori error analysis can be used to estimate the error in these samples.
- Future works includes an error analysis for coupled systems and an estimation of the effect of measure transformations.

References

[1] T. Butler, C. Dawson, and T. Wildey, *A posteriori error analysis of stochastic differential equations using polynomial chaos expansions.*, SIAM J. Scientific Computing, 33 (2011), pp. 1267-1291.

[2] T. Butler, P. Constantine, and T. Wildey, *A posteriori error analysis of parameterized linear systems using spectral methods.*, Submitted to SIAM Matrix Anal. Appl.

[3] D. Estep, V. Carey, V. Ginting, S. Tavener, and T. Wildey, *A posteriori error analysis of multiscale operator decomposition methods for multiphysics models*, Journal of Physics: Conference Series 125 (2008), pp. 1-16.

[4] L. Mathelin and O. P. Le Maitre, *Dual-based error analysis for uncertainty quantification in a chemical system*, PAMM, 7(1):2010007-2010008, 2007.

Office of Science

Sandia National Laboratories is a multi program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.
.

National Nuclear Security Administration

Sandia National Laboratories