SAND2011-4104C

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

Tim Wildey | Sandia National Labs

Troy Butler
Clint Dawson | The University of Texas at Austin

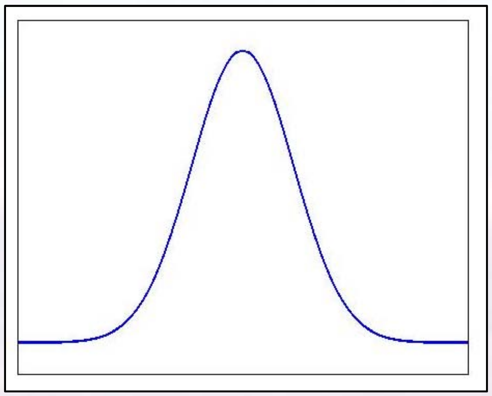
Uncertainty Quantification and Multiscale Materials Modeling Workshop

Santa Fe, NM June 15, 2011

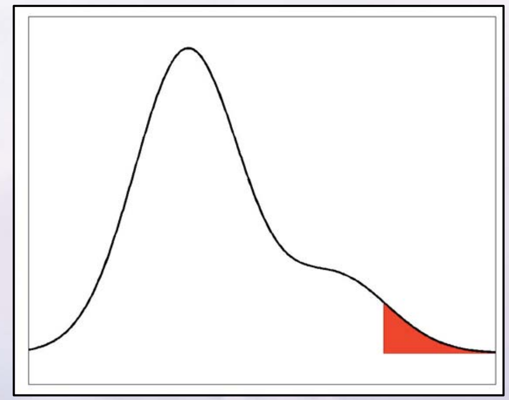
Sandia National Laboratories is a multi program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Motivation

Uncertainty Quantification



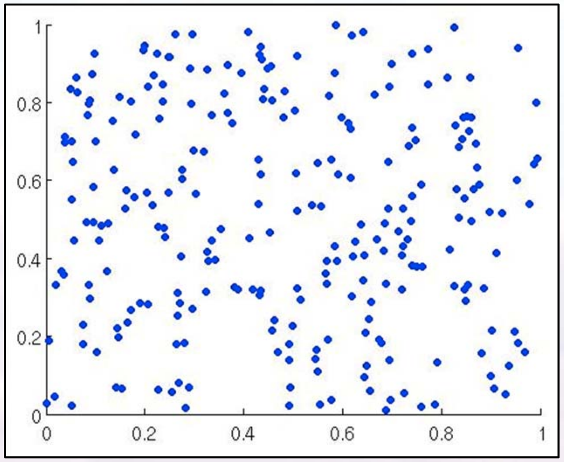
Model:
$u = M(\lambda)$



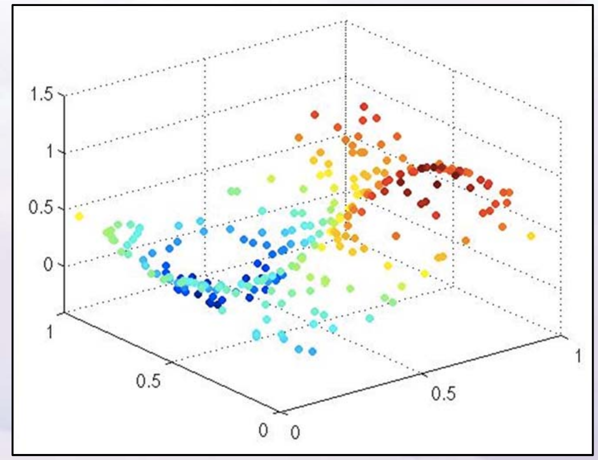
Given parameter(s) λ with distribution(s).

Calculate probabilities on quantities of interest $q(u(\lambda))$.

Monte Carlo Sampling (LHS, IS)



Model:
$u = M(\lambda)$



$\mathrm{P}[q(u(\lambda)) > 1] = 0$

Do we trust Monte Carlo (LHS, IS) to compute probabilities of rare events given a small number of samples?[1]

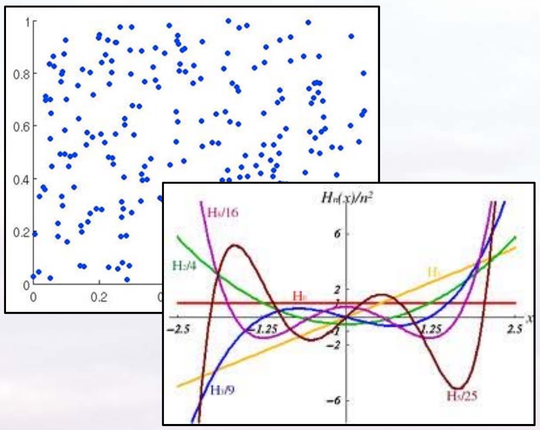
[1] "Importance Sampling: Promises and Limitations". L. P. Swiler and N. J. West. 2010.

Monte Carlo

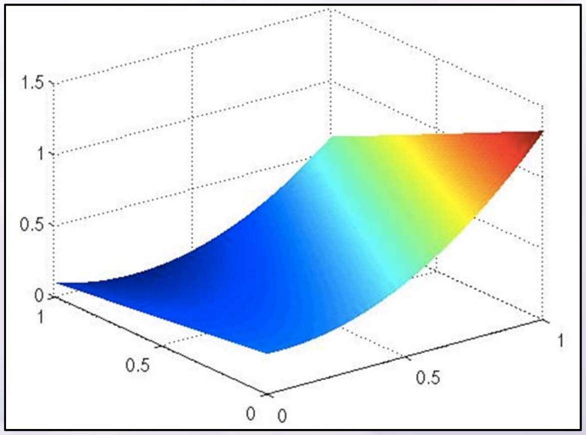
Monte Carlo

Sandia National Laboratories

Surrogate Models and Sampling



Model:
$u = M(\lambda)$



$$\mathrm{P}[q(u(\lambda)) > 1] = 0.035$$

Do we trust surrogate methods to compute probabilities of rare events given (virtually) unlimited samples?

Surrogate Models and Sampling

Spectral Methods for Uncertainty Quantification. O. P. Le Maitre and O. M. Knio, pgs. 39-40:

> ... the statistics of the random variable can be estimated by means of sampling strategies ... evaluation of the PC series at the sample points. We shall rely heavily on such sampling procedure to estimate densities, cumulative density functions, probabilities, etc.

"Stochastic spectral methods for efficient Bayesian solution of inverse problems". Y. Marzouk, H. Najm, and L. Rahn. Journal of Comp. Phys. 224 (2007) 560-586:

> Indeed, the per-sample cost is three orders of magnitude smaller for PC evaluations than for direct evaluations...

"Evaluation of failure probability via surrogate models". J. Li and D. Xiu. Journal of Comp. Phys., 229 (2010) 8966-8980:

> ... the straightforward sampling of a surrogate model can lead to erroneous results, no matter how accurate the surrogate model is.

Model Problem and Approximations

Model Problem

Model for nonlinear stochastic diffusive transport:

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

where S is a convex polygonal domain, $Q = S \times [0,T]$ and $(\cdot,\cdot)_S$ is the L^2 inner product.

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

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

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

Polynomial Chaos Expansions

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,

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

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

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

Sandia National Laboratories

Variational Formulation

Find $u_k \in L^2([0,T];H^1(S))$ such that for $k = 0, 1, \ldots, P$,

$$\int_0^T (\partial u_k/\partial t, v)_S \; dt$$

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

$$+ \frac{1}{\|\Phi_k\|^2} \int_0^T \left(\left\langle g\left(x, t; \sum_{j=0}^{P} u_j \Phi_j \right), \Phi_k \right\rangle, v \right)_S \; dt$$

$$= \int_0^T (f_k(x,t), v)_S \; dt$$

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

Discretization

Let

- $\mathcal{T}_h$ be a quasiuniform triangulation of S,
- $0 = t_0 < t_i < \cdots < t_N = T$ discretize $[0, T]$ with intervals $I_n = (t_{n-1}, t_n)$.
- V_h denote the space of continuous piecewise linear polynomials on $\mathcal{T}_h$.
- $W_n^{(q)} = V_h \times \mathbb{P}^{(q)}(I_n)$ where $\mathbb{P}^{(q)}(I_n)$ is the space of polynomials of degree q on I_n.

We compute $U_k \in W_n^{(q)}$ for $n = 1, 2, \ldots$ such that the variational formulation holds for all $v \in W_n^{(q)}$.

Adjoint-Based A Posteriori Error Analysis

Adjoint-Based Error Estimation

Consider the linear algebraic equation: $Ax = b$.

Let $X \approx x$ and define $e = x - X$ and $R = b - AX$.

Let ϕ solve the *adjoint problem*: $A^T\phi = \psi$.

Error representation:

$$(e, \psi) = (e, A^T\phi) = (Ae, \phi) = (R, \phi).$$

Not computable

Computable

Differential Operators

Linear differential equation:

$$\mathcal{L}u = -\nabla \cdot K\nabla u + \boldsymbol{b} \cdot \nabla u + cu = f,$$

with appropriate boundary conditions.

Adjoint differential equation:

$$\mathcal{L}^*\phi = -\nabla \cdot K^T\nabla\phi - \nabla \cdot (\boldsymbol{b}\phi) + c\phi = \psi,$$

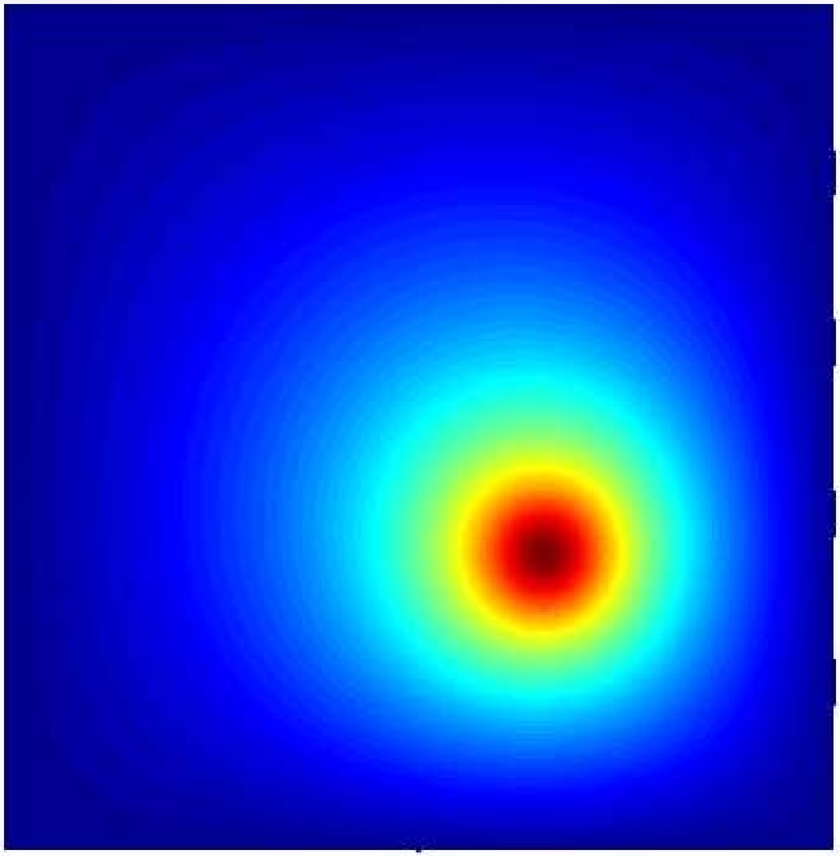
with adjoint boundary conditions.

Let u_h be a finite element approximation of u.

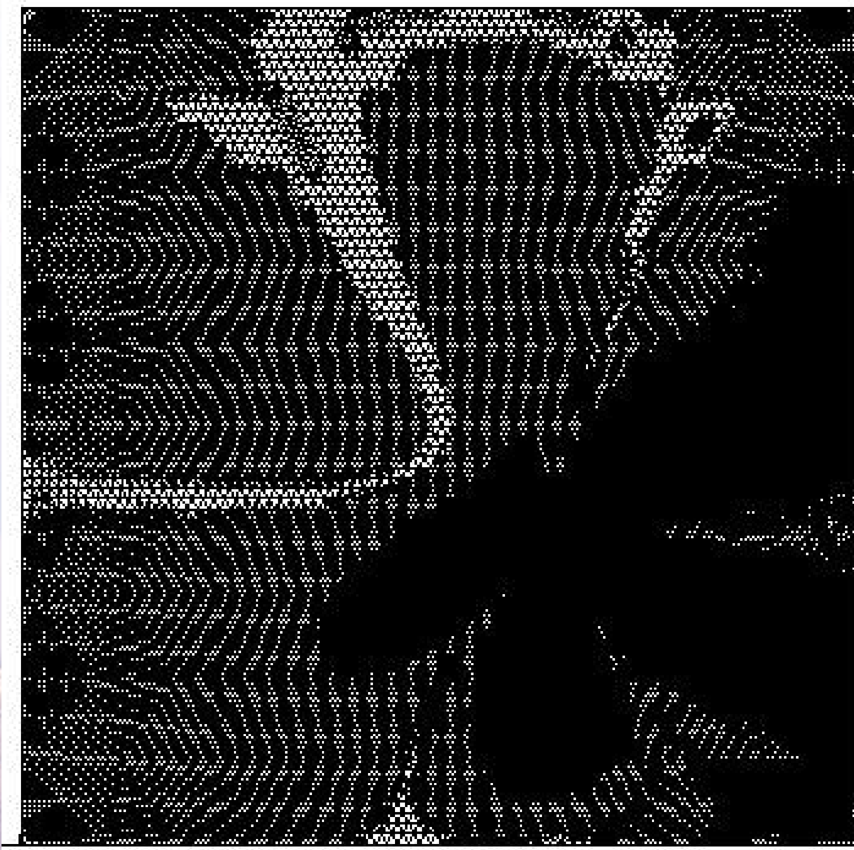
Error representation:

$$\begin{aligned}(e,\psi) = (f, \phi - \pi_h\phi) &- (K\nabla u_h, \nabla\phi - \nabla\pi_h\phi) \\ &- (\boldsymbol{b} \cdot \nabla u_h, \phi - \pi_h\phi) - (cu_h, \phi - \pi_h\phi)\end{aligned}$$

Steady State: No Convection

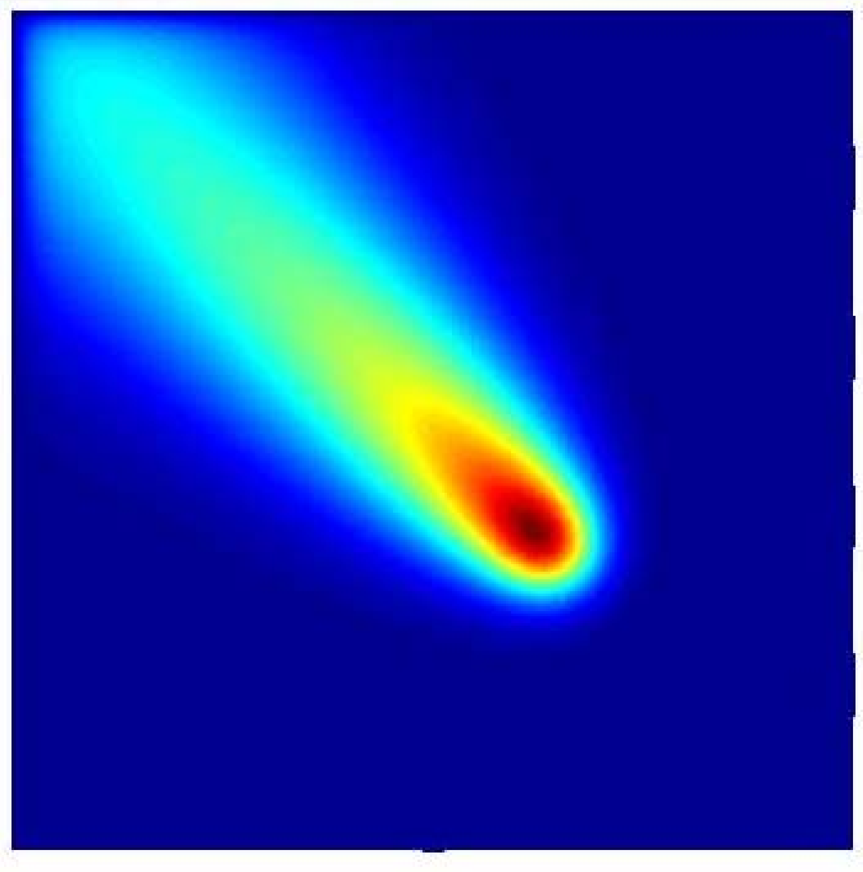


Adjoint solution

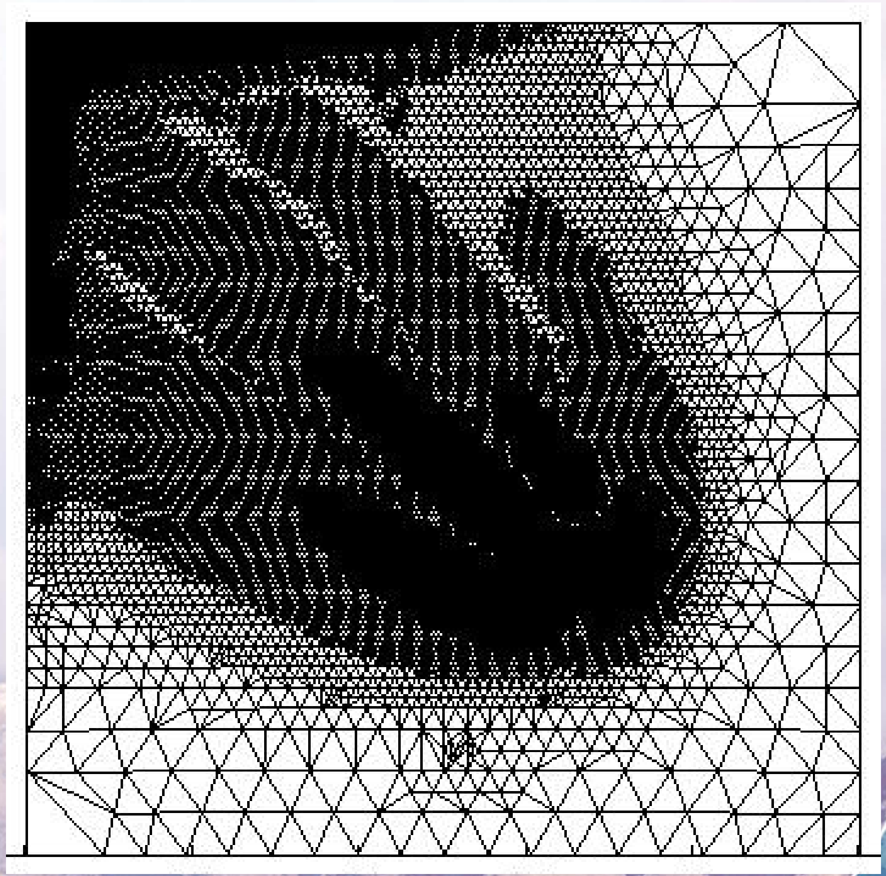


Adaptive mesh

Steady State: Strong Convection



Adjoint solution



Adaptive mesh

Time Dependent Adjoints

Time dependent linear differential equation:

$$\frac{\partial u}{\partial t} + \mathcal{L}u = f(x,t),$$

subject to appropriate initial and boundary conditions.

Adjoint differential equation:

$$-\frac{\partial \phi}{\partial t} + \mathcal{L}^* \phi = 0,$$

with adjoint boundary conditions and "initial condition": $\phi(x,T) = \psi$.

Error representation has the form:

$$\begin{aligned}(e(T),\psi) = &\int_0^T (\text{space residual}, \text{space adjoint weight})\, dt \\ &+ \int_0^T (\text{time residual}, \text{time adjoint weight})\, dt\end{aligned}$$

Adjoints for Nonlinear Operators

Nonlinear differential equation: $\mathcal{D}(u) = f(x)$.

Adjoint operator is defined such that:

$$\left(\overline{\mathcal{D}(u, u_h)}e, \phi\right) = \left(e, \overline{\mathcal{D}(u, u_h)}^* \phi\right)$$

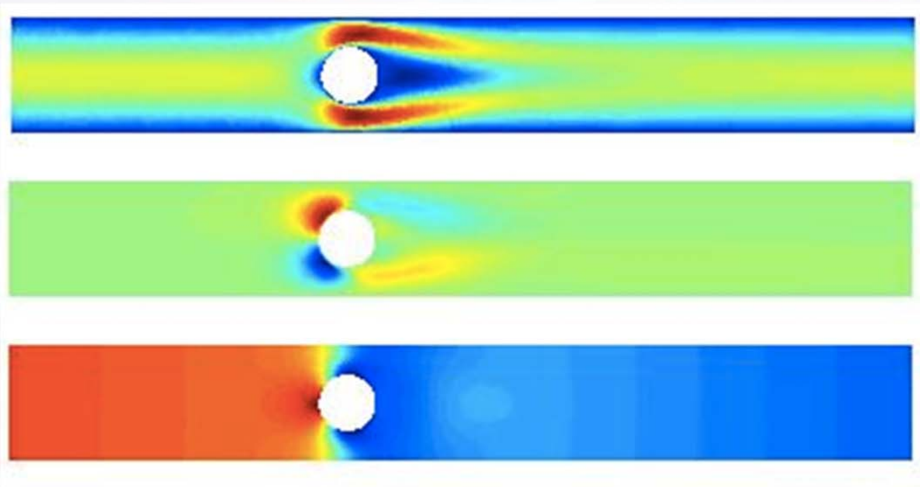
where the linearized operator, $\overline{\mathcal{D}(u, u_h)}$, satisfies

$$\overline{\mathcal{D}(u, u_h)} = \int_0^1 \partial_u \mathcal{D}(su + (1 - s)u_h)\, ds.$$

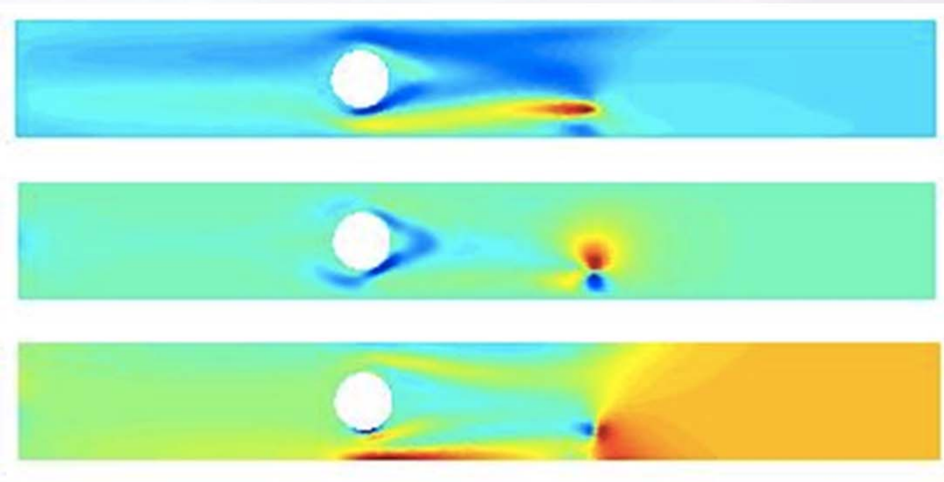
In practice, u is unavailable so the operator is linearized around u_h.

The effect on the error representation is usually higher order.

Navier-Stokes: Re = 100



Forward solution (u_x, u_y, p)



Adjoint solution (ϕ_x, ϕ_y, w)

Adaptive mesh

A Posteriori Error Analysis for Polynomial Chaos Approximations

The Adjoint Operator

The strong from of the adjoint to the nonlinear stochastic diffusion transport problem is,

$$\begin{cases} \frac{\partial \phi}{\partial t} - \nabla \cdot \left(A^T(x,t,\lambda) \nabla \phi \right) + \overline{g(u,U;\lambda)}^T \phi = \psi_1, & x \in S, T > t \geq 0, \\ A^T \nabla \phi \cdot \boldsymbol{n} = 0, & x \in \partial S, T > t \geq 0, \\ \phi(x,T) = \psi_2, & x \in S, \end{cases}$$

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

Either ψ_1 or ψ_2 is usually zero depending on the quantity of interest.

We approximate ϕ using a PC expansion:

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

The Error Representation

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

$$
\begin{aligned}
&\int_0^1 (e(x,t;\lambda),\psi_1)_S \; dt + (e(x,T;\lambda),\psi_2)_S = \\
&(e(x,0;\lambda),\phi(x,0;\lambda))_S - \sum_{n=1}^{N} \int_{I_n} (\partial U(x,t;\lambda)/\partial t, \phi(x,t;\lambda))_S \; dt \\
&+ \sum_{n=2}^{N} ([U(x,t;\lambda)], \phi(x,t;\lambda))_S + \sum_{n=1}^{N} \int_{I_n} (f, \phi(x,t;\lambda))_S \; dt \\
&- \sum_{n=1}^{N} \int_{I_n} (A(x,t;\lambda)\nabla U(x,t;\lambda), \nabla\phi(x,t;\lambda))_S \; dt \\
&- \sum_{n=1}^{N} \int_{I_n} (g(x,t;U), \phi(x,t;\lambda))_S \; dt
\end{aligned}
$$

Numerical Results

Problem Description

Consider the contaminant source problem[2]:

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

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

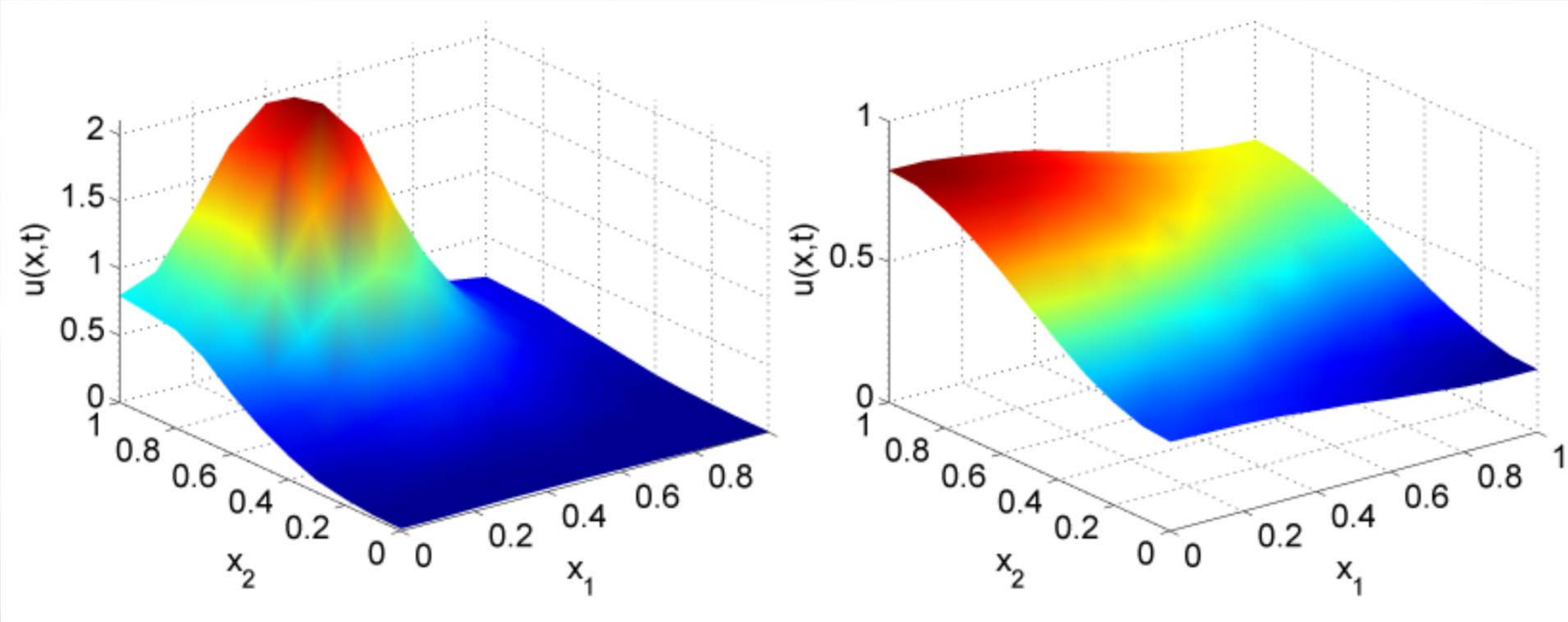
Random variable λ uniformly distributed on $[0,1]^2$.

Quantities of interest: Concentration at $t = 0.05$ and $t = 0.15$ at 9 measurement locations.

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

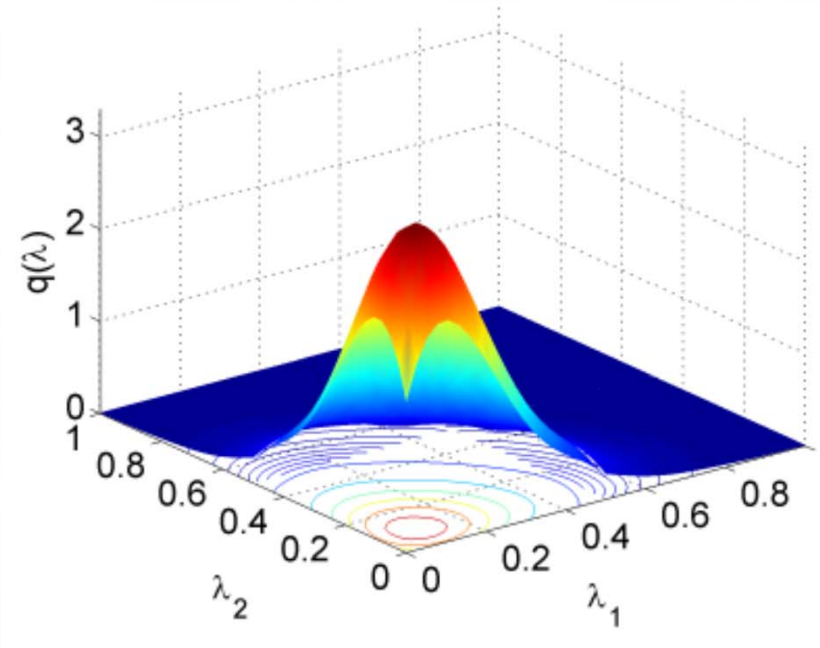
[2]See "Stochastic spectral methods for efficient Bayesian solution of inverse problems". Y. Marzouk, H. Najm and L. Rahn. 2007.

Contaminant Approximation: λ = (0.4,0.8)

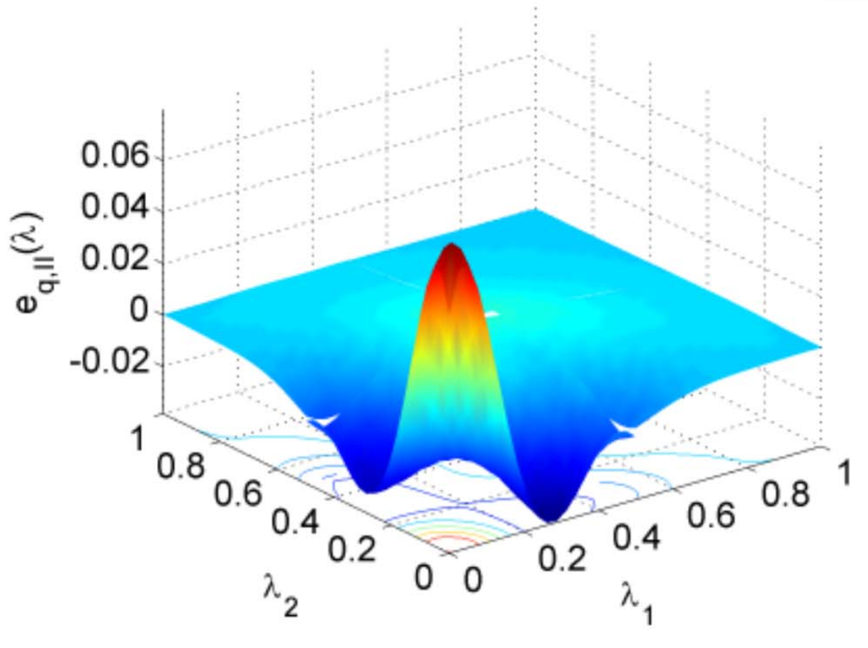


$t = 0.05$ $t = 0.15$

First Quantity of Interest at t = 0.05

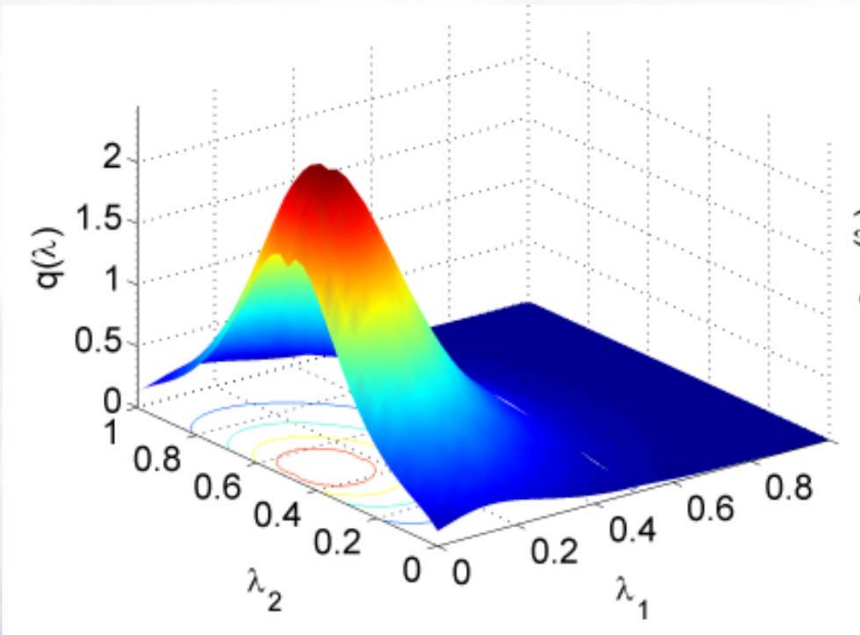


PC approximation

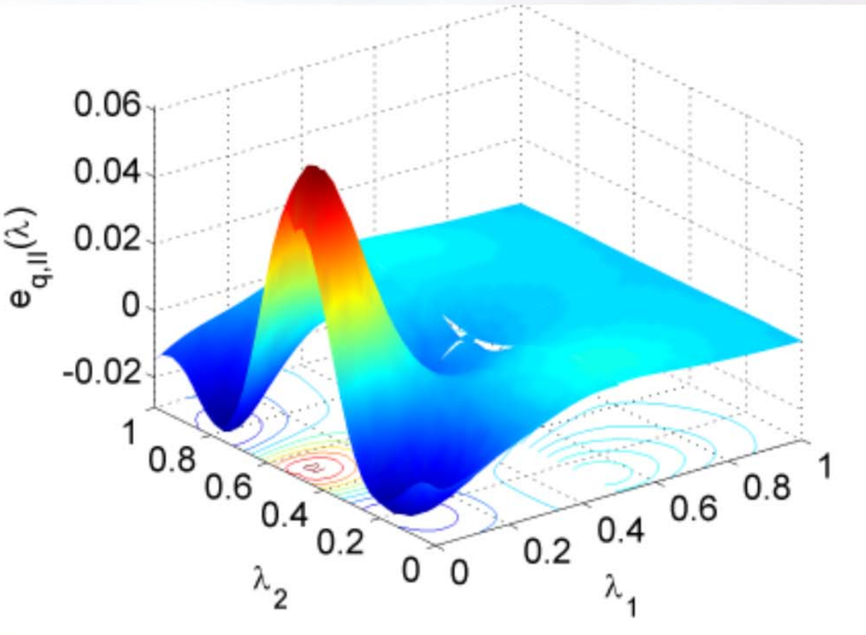


A posteriori error estimate

Fourth Quantity of Interest at t = 0.05

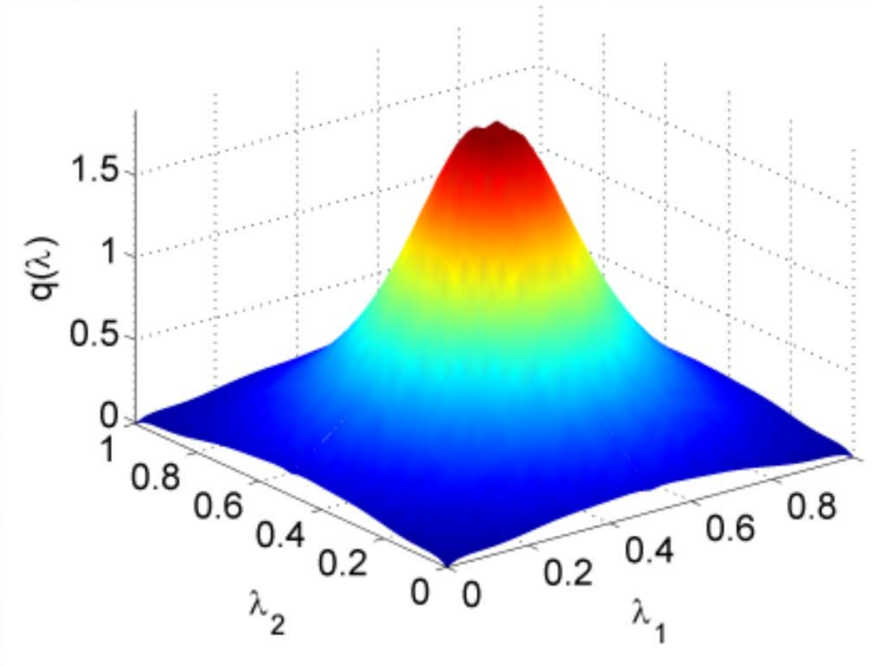


PC approximation

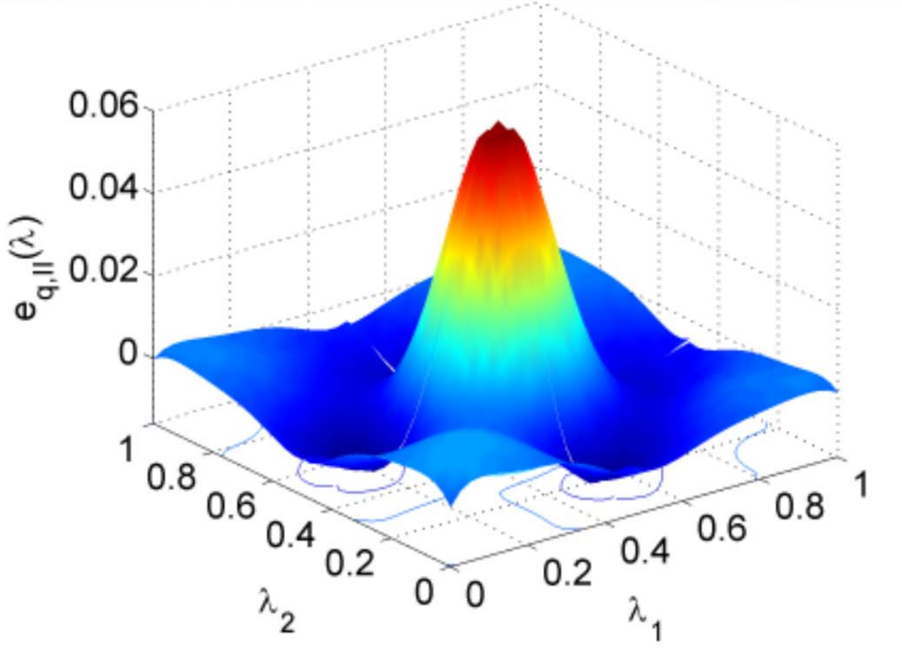


A posteriori error estimate

Fifth Quantity of Interest at t = 0.05

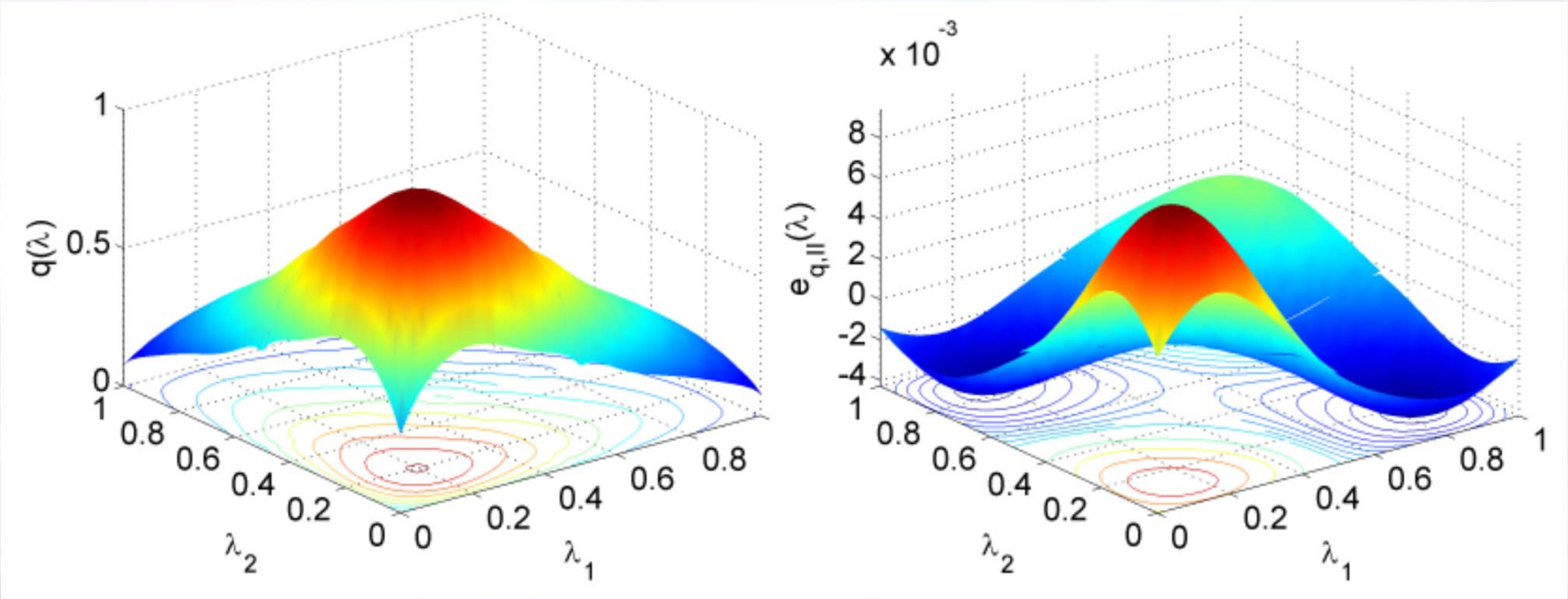


PC approximation



A posteriori error estimate

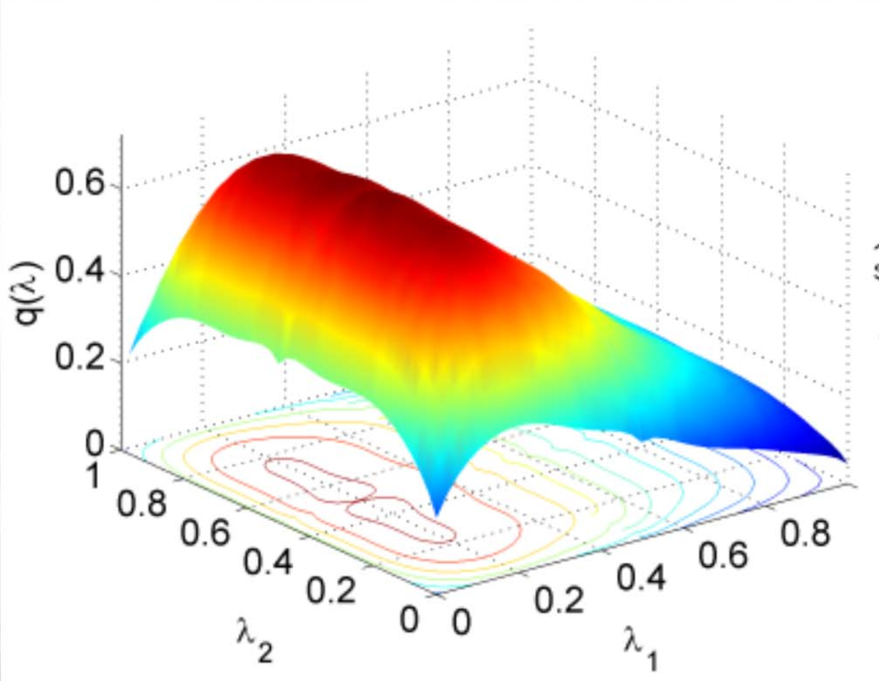
First Quantity of Interest at t = 0.15



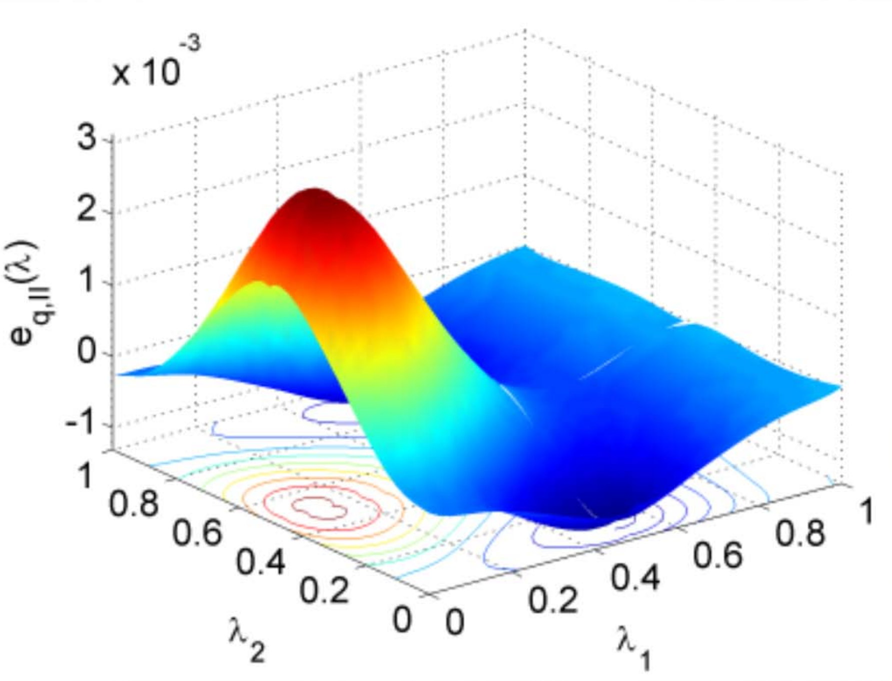
PC approximation

A posteriori error estimate

Fourth Quantity of Interest at t = 0.15

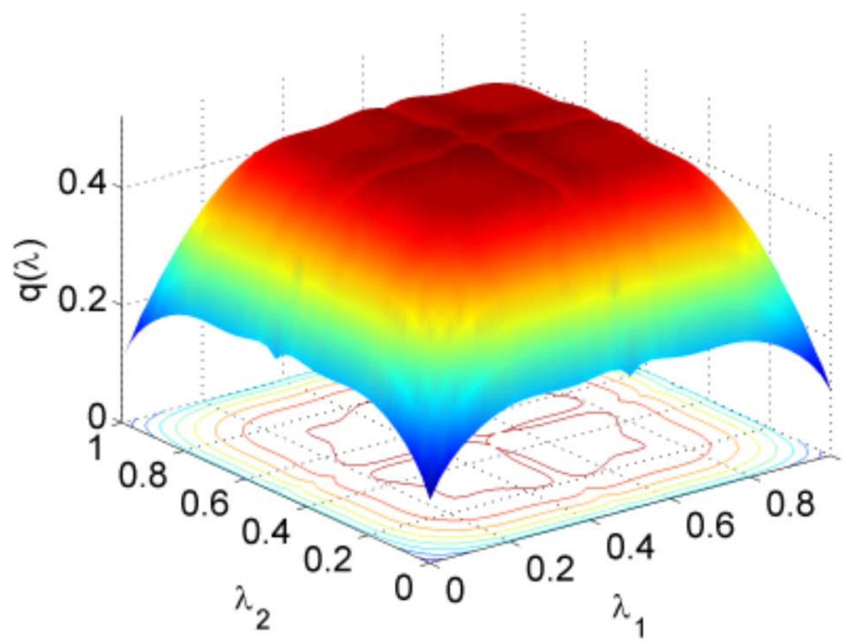


PC approximation

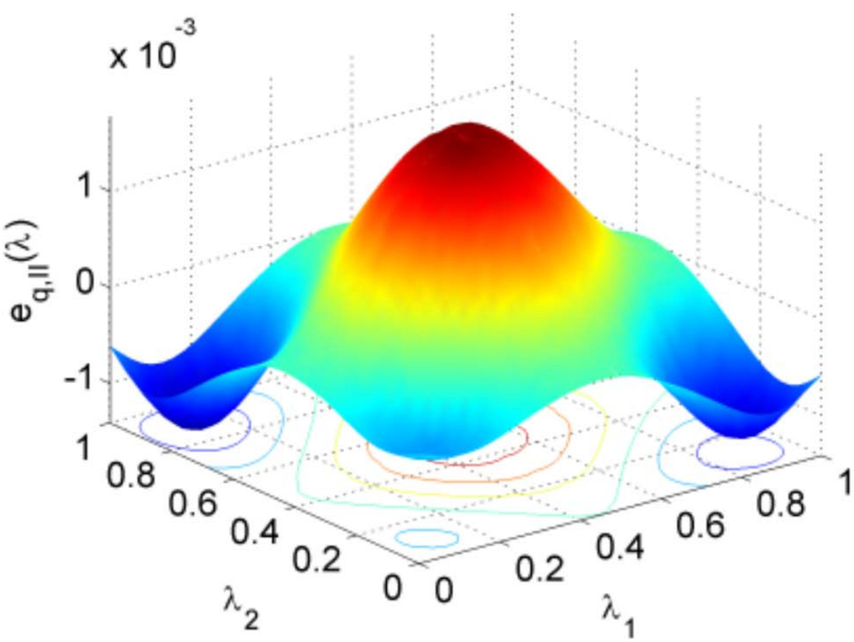


A posteriori error estimate

Fifth Quantity of Interest at t = 0.15



PC approximation



A posteriori error estimate

Effectivity of Error Estimate

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

Effectivity of Error Estimate

Time	λ	Std Err Est $u(x^{\{4\}}, t)$	PC Err Est $u(x^{\{4\}}, t)$	Ratio
0.05	$(0.25, 0.25)$	$-5.477E-03$	$-5.936E-03$	1.084
0.05	$(0.75, 0.25)$	$2.352E-03$	$2.352E-03$	1.000
0.05	$(0.25, 0.75)$	$-1.211E-03$	$-1.833E-03$	1.513
0.05	$(0.75, 0.75)$	$1.953E-03$	$1.943E-03$	0.995
0.05	$(0.4, 0.375)$	$-3.628E-03$	$-3.883E-03$	1.070
0.15	$(0.25, 0.25)$	$4.951E-04$	$5.018E-04$	1.013
0.15	$(0.75, 0.25)$	$-5.848E-04$	$-5.959E-04$	1.019
0.15	$(0.25, 0.75)$	$6.266E-04$	$6.301E-04$	1.006
0.15	$(0.75, 0.75)$	$-5.766E-04$	$-5.778E-04$	1.002
0.15	$(0.4, 0.375)$	$4.516E-04$	$4.447E-04$	0.985

Sandia National Laboratories

Effectivity of Error Estimate

Time	λ	Std Err Est $u(x^{\{5\}}, t)$	PC Err Est $u(x^{\{5\}}, t)$	Ratio
0.05	$(0.25, 0.25)$	$2.387E-03$	$2.238E-03$	0.938
0.05	$(0.75, 0.25)$	$-8.675E-03$	$-8.803E-03$	1.015
0.05	$(0.25, 0.75)$	$-8.377E-03$	$-8.383E-03$	1.051
0.05	$(0.75, 0.75)$	$-1.499E-03$	$-1.707E-03$	1.139
0.05	$(0.4, 0.375)$	$1.918E-02$	$1.860E-02$	0.970
0.15	$(0.25, 0.25)$	$2.782E-04$	$2.815E-04$	1.012
0.15	$(0.75, 0.25)$	$-2.299E-04$	$-2.286E-04$	0.994
0.15	$(0.25, 0.75)$	$-3.882E-04$	$-3.893E-04$	1.003
0.15	$(0.75, 0.75)$	$3.887E-04$	$3.935E-04$	1.012
0.15	$(0.4, 0.375)$	$1.439E-03$	$1.439E-03$	1.000

Problem Description

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 $\sigma = 0.1$.

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

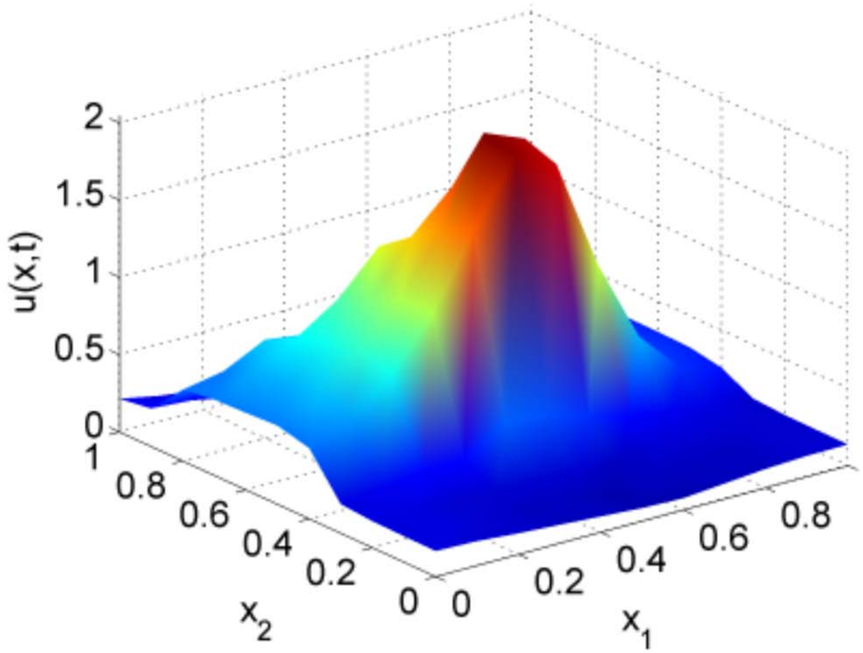
Random variable λ uniformly distributed on $[0.5, 1.5]$.

Quantities of interest: Concentration at $t = 0.05$ and $t = 0.15$ at 9 measurement locations.

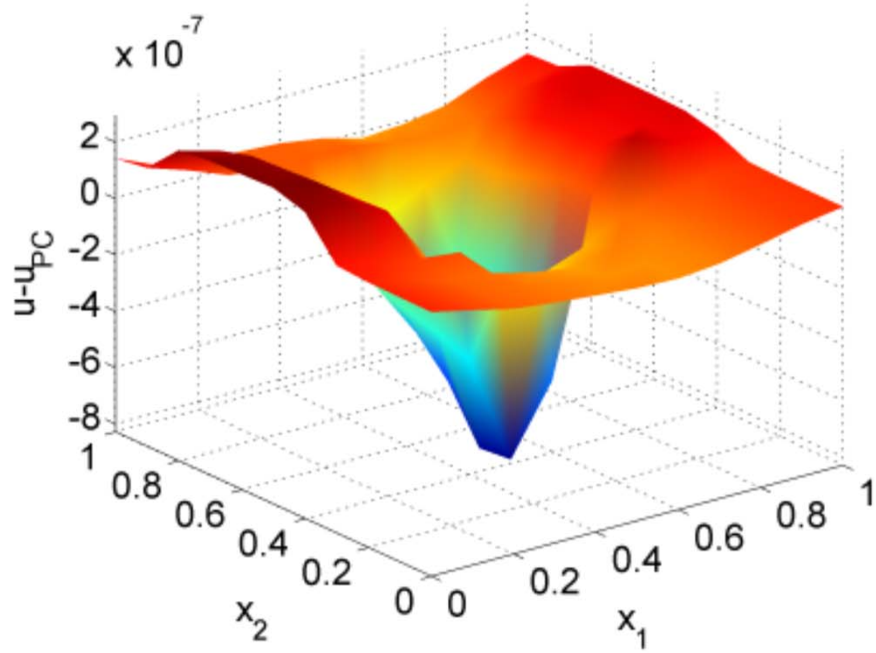
Discretization: $h = 0.1$, $\Delta t = 0.005$ and $6^{\text{th}}-$order PC expansion.

Sandia National Laboratories

Approximation and PC Error at λ = 1

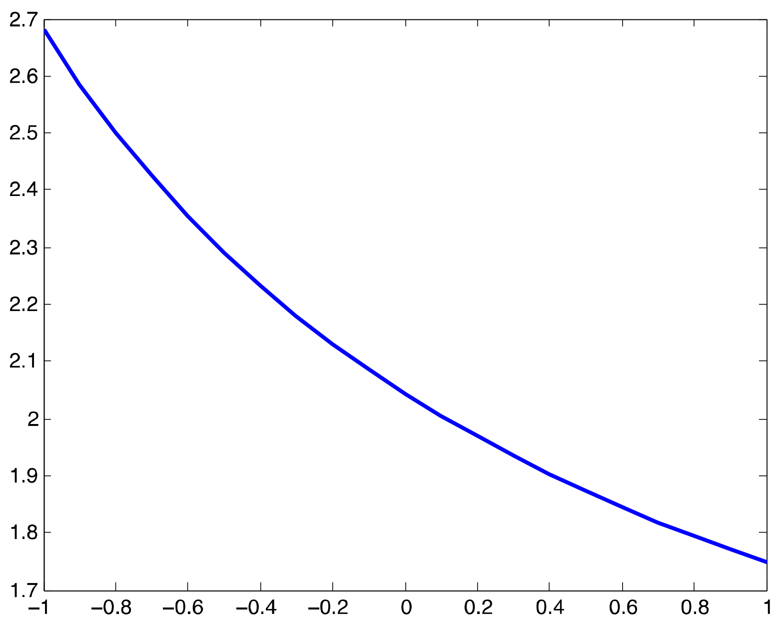


PC approximation

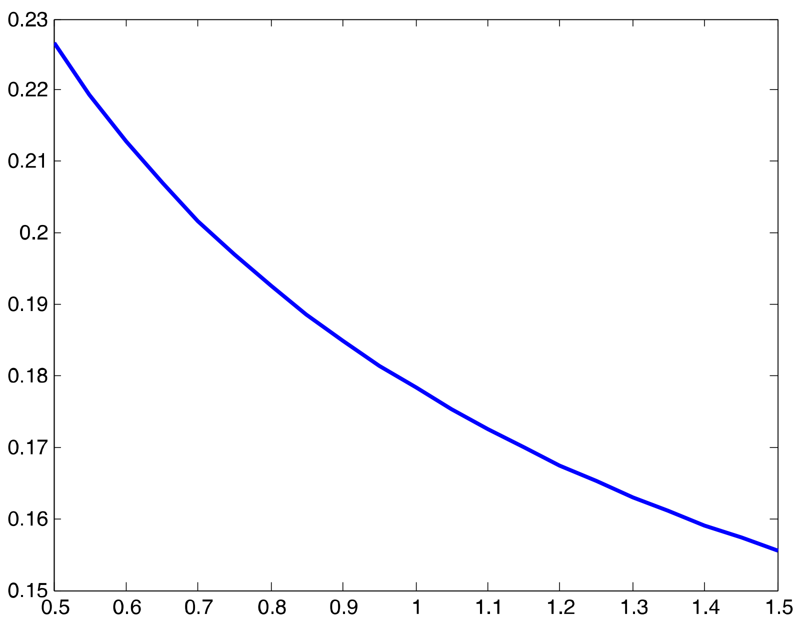


PC truncation error

Fifth Quantity of Interest at t = 0.05



PC approximation



A posteriori error estimate

Effectivity of Error Estimate

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

Sandia National Laboratories

Improved Linear Functionals for Parameterized Linear Systems

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,

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

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

Error Analysis

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

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

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

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

where $\phi_M(s)$ is some approximation of $\phi(s)$.

Motivates defining an *improved linear functional*:

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

Sandia National Laboratories

An Improved Linear Functional

If the pointwise error in the adjoint solution satisfies,

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

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

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

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

Example: Smooth Solution

Parameterized linear system:

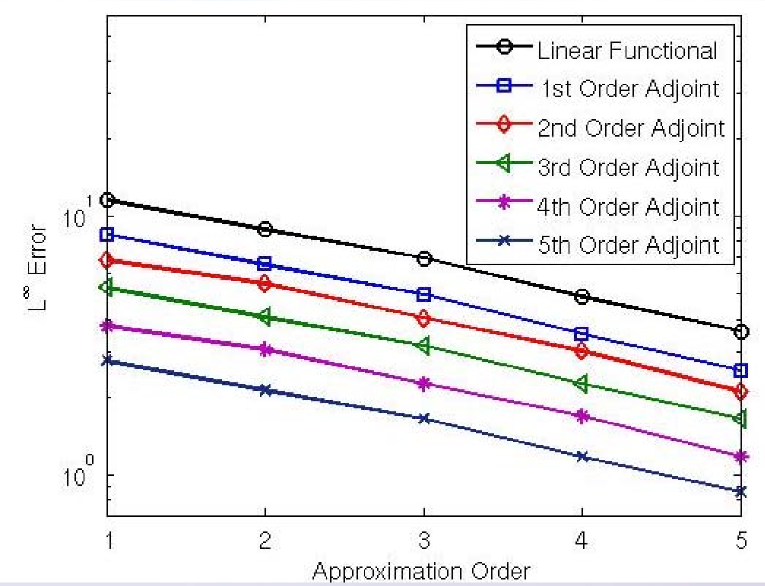
$$\begin{bmatrix} 1+\epsilon & s \\ s & 1 \end{bmatrix} \begin{bmatrix} x_1(s) \\ x_2(s) \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

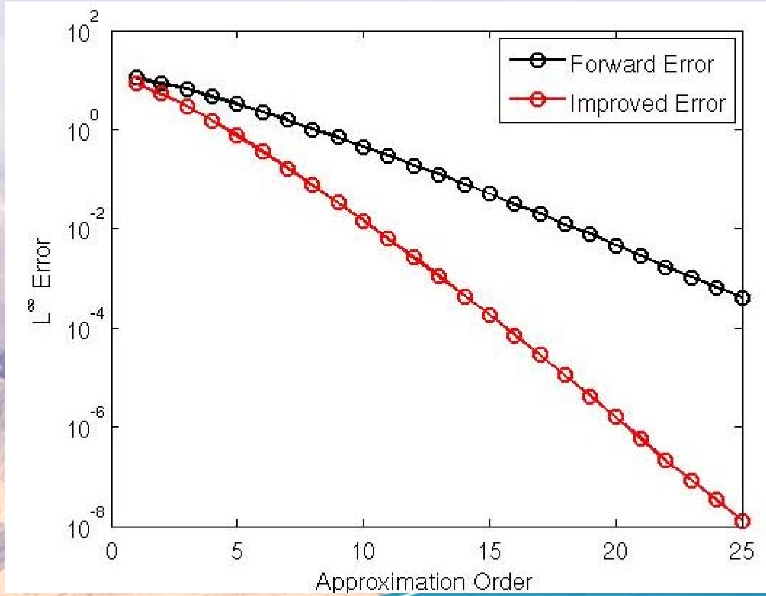
with $s \in [-1, 1]$ and $\epsilon = 0.8$.

Quantity of interest is $x_1(s)$.

Parameterized adjoint system:

$$\begin{bmatrix} 1+\epsilon & s \\ s & 1 \end{bmatrix} \begin{bmatrix} \phi_1(s) \\ \phi_2(s) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$





Example: Discontinuous Solution

Parameterized linear system:

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

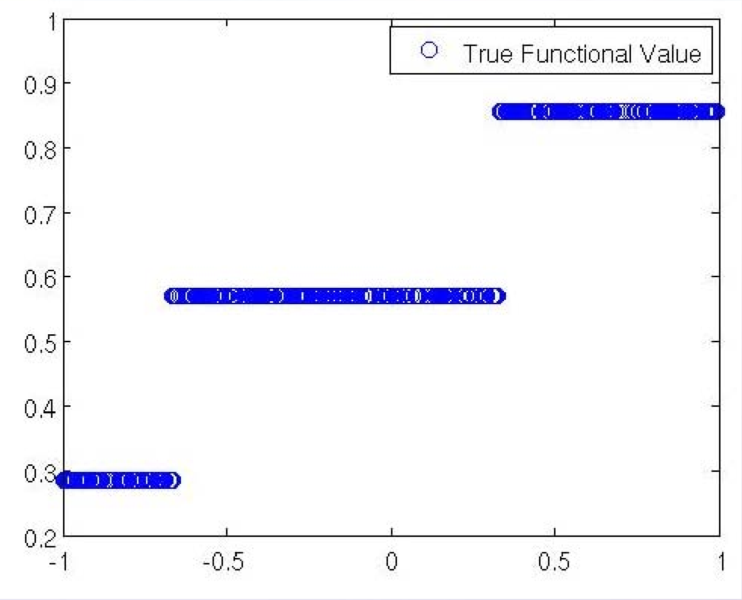
where $\lceil \cdot \rceil$ is the ceiling operator and $s_i \in [-1, 1]$.

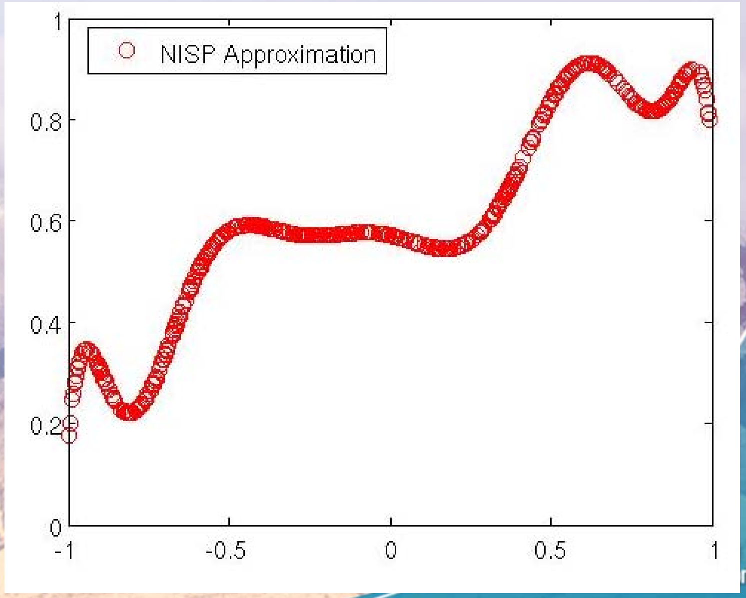
Discontinuity at $s_3 = -2/3, 1/3$

Quantity of interest is $x_1(s)$.

Parameterized adjoint system:

$$\begin{bmatrix} 2 & -s_2 \\ -s_1 & 1 \end{bmatrix} \begin{bmatrix} \phi_1(s) \\ \phi_2(s) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$





Example: Discontinuous Solution

Parameterized linear system:

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

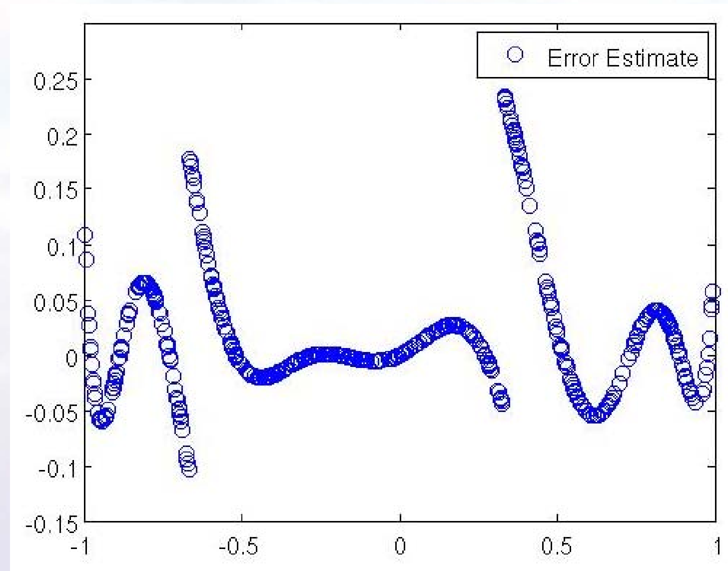
where $\lceil \cdot \rceil$ is the ceiling operator and $s_i \in [-1, 1]$.

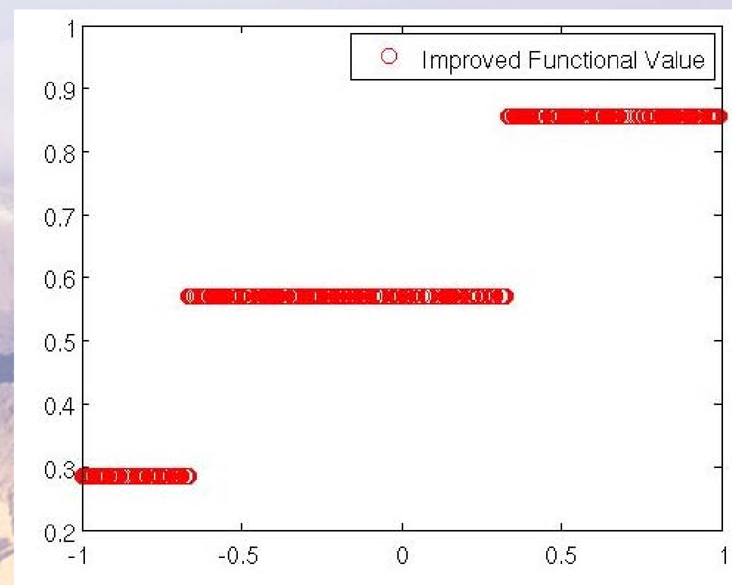
Discontinuity at $s_3 = -2/3, 1/3$

Quantity of interest is $x_1(s)$.

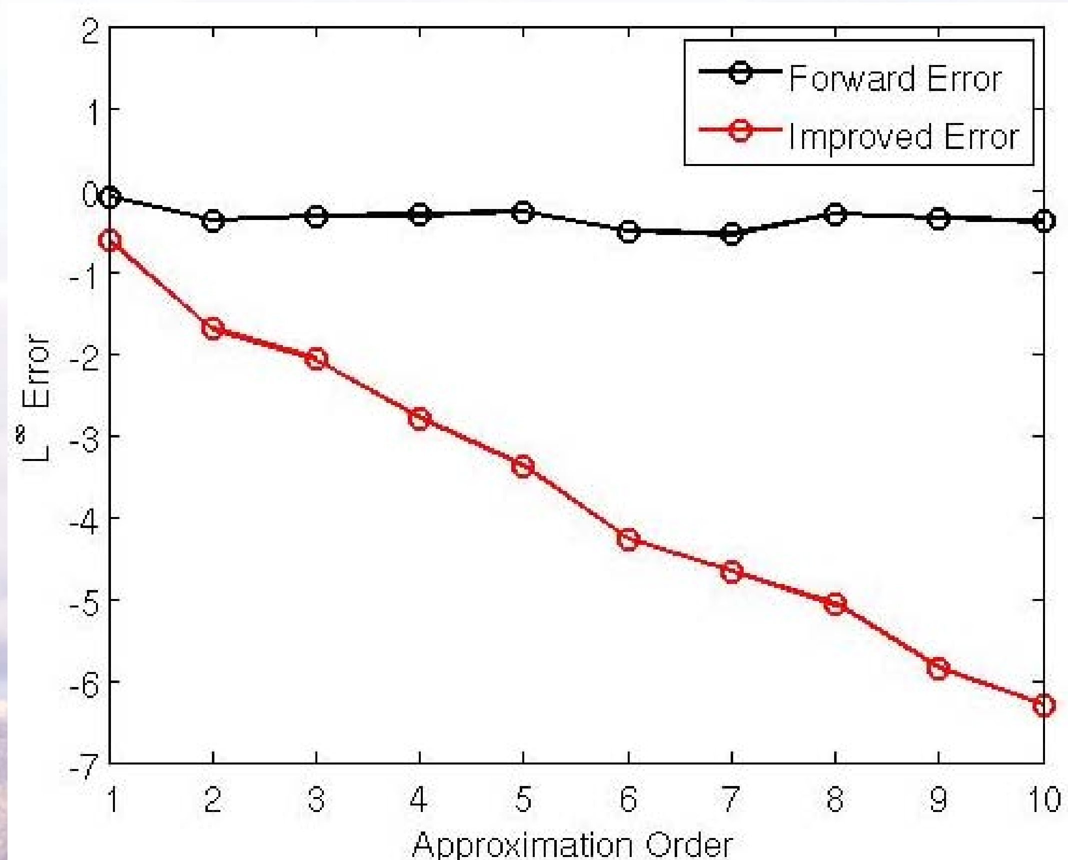
Parameterized adjoint system:

$$\begin{bmatrix} 2 & -s_2 \\ -s_1 & 1 \end{bmatrix} \begin{bmatrix} \phi_1(s) \\ \phi_2(s) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$





Example: Discontinuous Solution



Example: Discretized PDE

Linear partial differential equation,

$$-\nabla \cdot (K(x, y, s)\nabla u) = f(x, y),$$

with $s_i \in [-1, 1], i = 1, 2, \ldots, 8$

Parameterization of $K(x, y, s)$:

$$K(x, y, s) = 15\mathbb{I} + \sum_{k=1}^{6} s_k \left(\sin(k\pi x) + \cos(k\pi y)\right) \mathbb{I},$$

Discretize using Galerkin finite element method.

Parameterized linear system,

$$\left(A_0 + \sum_{k=1}^{6} s_k A_k\right) x(s) = b.$$

Sandia National Laboratories

Example: Discretized PDE

Quantity of interest is solution at $(x, y) = (0.8, 0.6)$

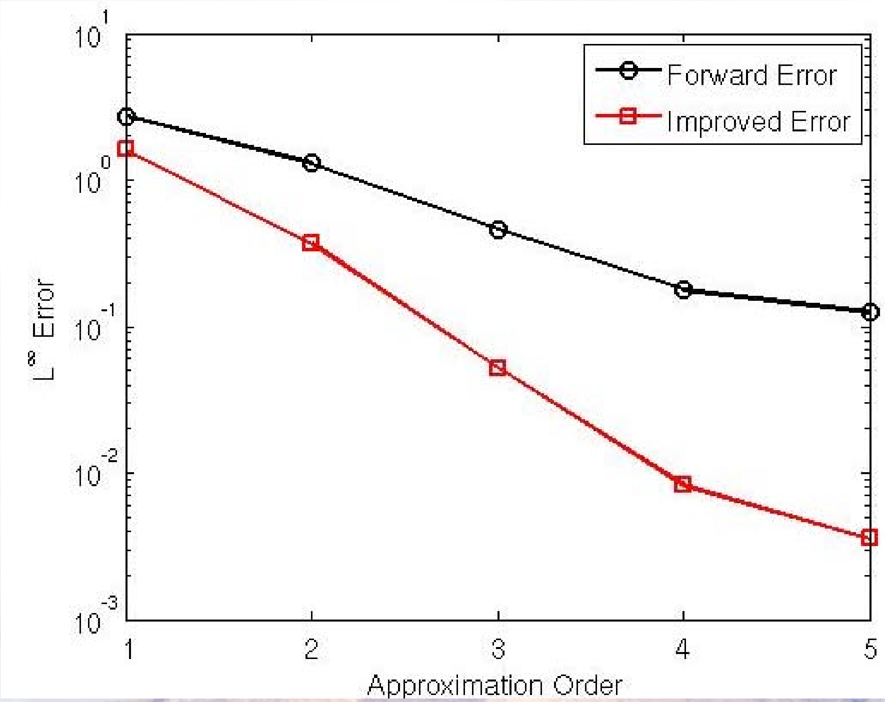
$$\psi(x,y) = \frac{400}{\pi} \exp(-400(x-0.8)^2 - 400(y-0.6)^2),$$

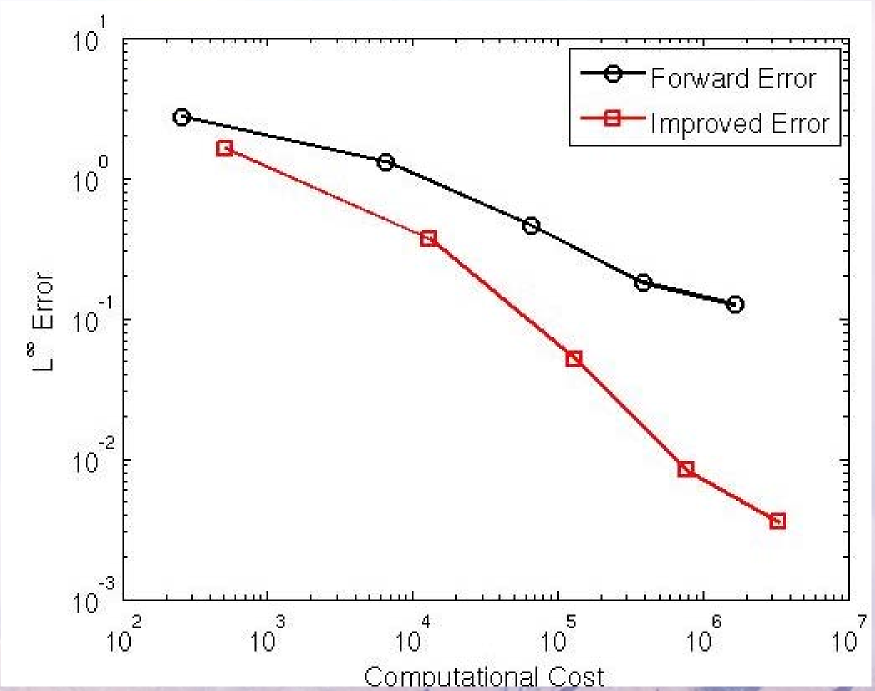
Project onto the finite element space to obtain data for discrete adjoint,

$$\left(A_0 + \sum_{k=1}^{6} s_k A_k \right)^T \phi(s) = \psi.$$

Use the NISP approximation method.

Example: Discretized PDE





Conclusions and Future Work

Conclusions and Future Work

- Surrogate models typically have errors due to:
  - spatial and temporal discretizations
  - truncated stochastic expansions or quadrature
- We can produce a posteriori error estimates for a quantity of interest obtained by sampling a surrogate model.
- Can be used for error estimates, error bounds, defining improved quantities of interest, and adaptivity.
- Future work will include:
  - Error estimates for stochastic PDE's,
  - Stochastic adaptivity,
  - Inverse problems.