Uncertainty Quantification for Complex Systems

Paul W. Talbot[1], Anil K. Prinja[1], Cristian Rabiti[2]

[1]University of New Mexico
[2]Idaho National Laboratory

BYU-Idaho Physics Colloquium, February 26th 2015

Outline

Discussion Points

Introduction

Who is this guy?

Current:

- Ph.D Nuclear Engineering student, UNM
- Idaho National Laboratory (RAVEN, MOOSE)

Past:

- M.S. Nuclear Engineering, Oregon State University
- B.S. Physics, BYU-Idaho (2010)

Uncertainty

Two Types

- Aleatory - True Randomness
  - Quantum effects
  - Particle-Material interactions (gold foil)
  - Brownian Motion

- Epistemic - Unmeasured Uncertainty
  - Tool Accuracy
  - Complicated Dependencies (arrow, double pendulum)
  - Documentation

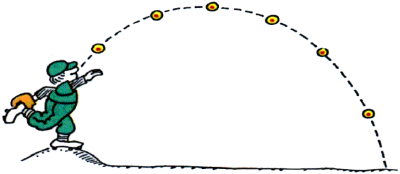
Idaho National Laboratory

Example Stochastic Problem

Projectile Motion

$$y(t) = y_i + v\sin(\theta)t - \frac{1}{2}gt^2,$$

$$x(t) = v\cos(\theta)t.$$



Solution: $x_f = \frac{v\cos\theta}{g}\left(v\sin\theta + \sqrt{v^2\sin^2\theta + 2gy_i}\right)$

Example Stochastic Problem

Solved!

$$x_f = \frac{v\cos\theta}{g}\left(v\sin\theta + \sqrt{v^2\sin^2\theta + 2gy_i}\right)$$

- initial height $y_i = 2$ m
- initial velocity $v = 35$ m/s
- initial trajectory $\theta = 35^o$
- accel. gravity $g = -9.81$ m/s/s

Solution: $x_f \approx 120$ m

Example Stochastic Problem

Uncertainty

$$x_f = \frac{v\cos\theta}{g}\left(v\sin\theta + \sqrt{v^2\sin^2\theta + 2gy_i}\right)$$

- initial height $y_i = 1 \pm 1$ m
- initial velocity $v = 35.5 \pm 2.5$ m/s
- initial trajectory $\theta = 45 \pm 10^o$
- accel. gravity $g = 9.7988 \pm 0.0349$ m/s/s

Solution: x_f=?

Uncertainty Quantification

Methods

Some way to quantify uncertainty:

Uncertainty Quantification
Methods

Some way to quantify uncertainty:

- Min-Max
  - Good for monotonic problems

Uncertainty Quantification
Methods

Some way to quantify uncertainty:

- Min-Max
    - Good for monotonic problems
- Sandwich Formula
    - Good for analytic solutions

Uncertainty Quantification

Methods

Some way to quantify uncertainty:

- Min-Max
  - Good for monotonic problems
- Sandwich Formula
  - Good for analytic solutions
- Perturbation
  - Valid for small uncertainty

Uncertainty Quantification

$$x_f = \frac{v\cos\theta}{g}\left(v\sin\theta + \sqrt{v^2\sin^2\theta + 2gy_i}\right)$$

Min-Max

$$x_{f,\text{min}} = \frac{(33)(0.5736)}{9.8337}\left((33)(0.8192) + \sqrt{(33)^2(0.8192)^2 + 2(9.8337)(0)}\right) = 105.46 \text{ m}$$

$$x_{f,\text{max}} = \frac{(55)(0.8192)}{9.7369}\left((55)(0.5736) + \sqrt{(55)^2(0.5736)^2 + 2(9.8337)(2)}\right) = 142.17 \text{ m}$$

Result: $x_f \approx 124 \pm 18.3$ m

Idaho National Laboratory

UNM

Uncertainty Quantification

$$x_f = \frac{v\cos\theta}{g}\left(v\sin\theta + \sqrt{v^2\sin^2\theta + 2gy_i}\right)$$

Min-Max

$$x_{f,\text{min}} = \frac{(33)(0.5736)}{9.8337}\left((33)(0.8192) + \sqrt{(33)^2(0.8192)^2 + 2(9.8337)(0)}\right) = 105.46 \text{ m}$$

$$x_{f,\text{max}} = \frac{(55)(0.8192)}{9.7369}\left((55)(0.5736) + \sqrt{(55)^2(0.5736)^2 + 2(9.8337)(2)}\right) = 142.17 \text{ m}$$

Result: $x_f \approx 124 \pm 18.3$ m

Flawed Reasoning

- θ not monotonic!
- Does increasing θ make a longer or shorter range?

Uncertainty Quantification

Sandwich Formula (simplified):

$$\sigma_{x_f} = \sqrt{\left(\frac{\partial x_f}{\partial y_i}\right)^2 \sigma_{y_i}^2 + \left(\frac{\partial x_f}{\partial v}\right)^2 \sigma_v^2 + \left(\frac{\partial x_f}{\partial g}\right)^2 \sigma_g^2 + \left(\frac{\partial x_f}{\partial \theta}\right)^2 \sigma_\theta^2}$$

Works well for simple functions

- Simple derivatives
- Analytic solution
- Assumes mean is reference value

Uncertainty Quantification

$$x_f = \frac{v\cos\theta}{g}\left(v\sin\theta + \sqrt{v^2\sin^2\theta + 2gy_i}\right)$$

Sandwich Formula:

$$\sigma_{x_f} = \sqrt{\left(\frac{\partial x_f}{\partial y_i}\right)^2 \sigma_{y_i}^2 + \left(\frac{\partial x_f}{\partial v}\right)^2 \sigma_v^2 + \left(\frac{\partial x_f}{\partial g}\right)^2 \sigma_g^2 + \left(\frac{\partial x_f}{\partial \theta}\right)^2 \sigma_\theta^2}$$

Result: $x_f = 120 \pm 62.44$ m

A More Difficult Problem

Air Resistance

With Air Resistance:

$$y(t) = \frac{v_T}{g}(v\sin\theta + v_T)\left(1 - e^{-gt/v_T}\right) - v_T t,$$

$$x(t) = \frac{v v_T \cos\theta}{g}\left(1 - e^{-gt/v_T}\right).$$

$$v_T = \frac{mg}{D}, \qquad D = \frac{\rho C A}{2}, \qquad A = \pi r^2$$

Solve numerically to get x_f (Forward Euler).

A More Difficult Problem

Aside: Forward Euler

Take small Δ_t time steps `while` $y^t > 0$: $t = t + \Delta_t$,

$$a_x^{(t+\Delta_t)} = \frac{-D}{m} v^{(t)}\ v_x^{(t)}, \qquad a_y^{(t+\Delta_t)} = -g - \frac{D}{m} v^{(t)}\ v_x^{(t)},$$

$$v_x^{(t+\Delta_t)} = v_x^{(t)} + a_x^{(t+\Delta_t)} \Delta_t, \qquad v_y^{(t+\Delta_t)} = v_y^{(t)} + a_y^{(t+\Delta_t)} \Delta_t,$$

$$x^{(t+\Delta_t)} = x^{(t)} + v_x^{(t+\Delta_t)} \Delta_t + \frac{1}{2} a_x^{(t+\Delta_t)} \Delta_t^2,$$

$$y^{(t+\Delta_t)} = y^{(t)} + v_y^{(t+\Delta_t)} \Delta_t + \frac{1}{2} a_y^{(t+\Delta_t)} \Delta_t^2.$$

(video)

A More Difficult Problem

Uncertainty Summary

$$
\begin{aligned}
y_i &= 1 \pm 1 \text{ m}, \\
v &= 35.5 \pm 2.5 \text{ m/s}, \\
\theta &= 45 \pm 10^o, \\
g &= 9.7988 \pm 0.0349 \text{ m/s/s}, \\
m &= 0.145 \pm 0.0725 \text{ kg}, \\
r &= 0.0336 \pm 0.00336 \text{ m}, \\
C &= 0.5 \pm 0.5, \\
\rho_{\text{air}} &= 1.2 \pm 0.1 \text{ kg/m}^3.
\end{aligned}
$$

Sensitivity Study

A More Difficult Problem

Equation Summary

$$y(t) = \frac{v_T}{g}(v \sin\theta + v_T)\left(1 - e^{-gt/v_T}\right) - v_T t,$$

$$x(t) = \frac{v v_T \cos\theta}{g}\left(1 - e^{-gt/v_T}\right).$$

$$v_T = \frac{mg}{D}, \qquad D = \frac{\rho C A}{2}, \qquad A = \pi r^2$$

Uncertainty Quantification

Complicated Problems

How do we quantify uncertainty for problems without simple analytic solutions?

- Monte Carlo sampling
- Stochastic Collocation

Uncertainty Quantification

Monte Carlo

- Let $u(Y)$ be any system, like $x_f(y_i, v, \theta, g, m, r, C, \rho)$
- Randomly sample input parameters, record outputs
- Repeat M times
- Calculate moments (mean, variance, skew, kurtosis)

$$\text{Mean: } \bar{u} \approx \frac{1}{M}\sum u\left(Y^{(m)}\right)$$

(video)

Idaho National Laboratory

UNM

Uncertainty Quantification

Stochastic Collocation

- Let $u(Y)$ be any system, like $x_f(y_i, v, \theta, g, m, r, C, \rho)$
- Represent original model with polynomials
- Calculate moments (mean, variance, skew, kurtosis)

$$u(Y) \approx \sum_{k\in\Lambda} c_k \Phi_k(Y),$$

$$\Phi_k(Y) = \phi_{k_1}(Y_1)\cdot\phi_{k_2}(Y_2)\cdot\ldots\cdot\phi_{k_N}(Y_N)$$

Uncertainty Quantification

Stochastic Collocation

Our case:

$$x_f(y_i, v, \theta, g, m, r, C, \rho) \approx \sum_{k \in \Lambda} c_k \Phi_k(y_i, v, \theta, g, m, r, C, \rho),$$

$$\Phi_k(y_i, v, \theta, g, m, r, C, \rho) = \phi_{y_i}(y_i) \cdot \phi_v(v) \cdot \ldots \cdot \phi_\rho(\rho).$$

$$c_k = \frac{\int\int\int\int\int\int\int\int x_f \Phi \, d(y_i, v, \theta, g, m, r, C, \rho)}{\int\int\int\int\int\int\int\int \Phi^2 \, d(y_i, v, \theta, g, m, r, C, \rho)}.$$

Uncertainty Quantification

Combining polynomials?

For example, let ϕ be monomials $(1, x, x^2, x^3, x^4, ...)$.

$$\Lambda = \left\{ \begin{array}{c} (0,0,0,0,0,0,0,0), \\ (1,0,0,0,0,0,0,0), \\ (0,1,0,0,0,0,0,0), \\ \dots \\ (1,2,3,4,5,6,7,8), \\ \dots \end{array} \right\}$$

k	polynomial Φ
(0,0,0,0,0,0,0,0)	$y_i^0 \cdot v^0 \cdot \theta^0 \cdot g^0 \cdot m^0 \cdot r^0 \cdot C^0 \cdot \rho^0 = 1$
(1,2,3,4,5,6,7,8)	$y_i^1 \cdot v^2 \cdot \theta^3 \cdot g^4 \cdot m^5 \cdot r^6 \cdot C^7 \cdot \rho^8$
(1,1,1,1,1,1,1,1)	$y_i \cdot v \cdot \theta \cdot g \cdot m \cdot r \cdot C \cdot \rho$

Idaho National Laboratory

UNM

Uncertainty Quantification

Stochastic Collocation

Comparison

Monte Carlo	Stochastic Collocation
Dimension-independent	Calculations grow with dimension*
Slow converging	Very fast convergence*
	Can replace original model

Uncertainty Quantification

Results: pdf

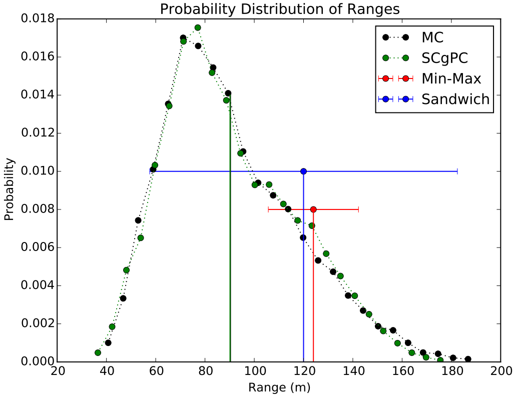


Figure: Probability Distributions

Uncertainty Quantification

Results: Expected Value, Values

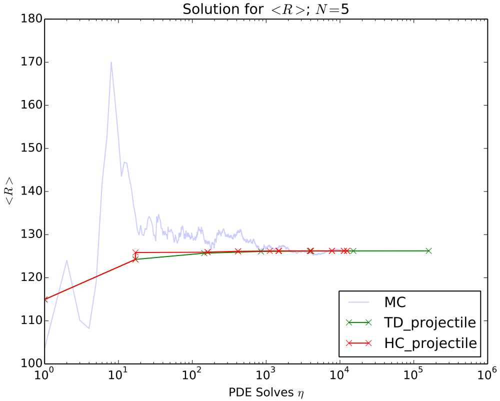


Figure: $\mathbb{E}[x_f]$

Uncertainty Quantification

Results: Expected Value, Errors

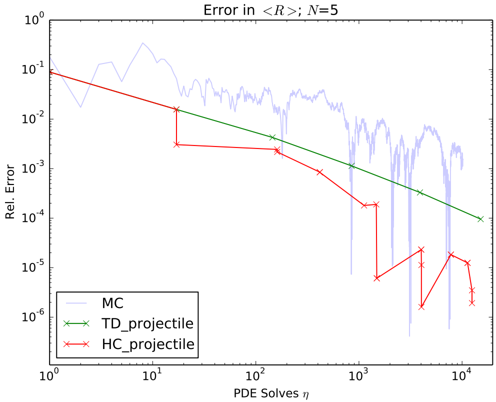


Figure: Error in $\mathbb{E}[x_f]$

Sensitivity Analysis

Polynomial Expansion Revisited

Recall: $u(Y) \approx \sum_{k\in\Lambda} c_k \Phi_k(Y)$, so for $x_f(y_i, v, \theta, g, m, r, C, \rho)$:

$$
\begin{aligned}
x_f &\approx c_{(0,0,0,0,0,0,0,0)} \\
&+ c_{(1,0,0,0,0,0,0,0)} y_i + c_{(0,1,0,0,0,0,0,0)} v + c_{(0,0,1,0,0,0,0,0)} \theta + \dots \\
&+ c_{(2,0,0,0,0,0,0,0)} y_i^2 + c_{(1,1,0,0,0,0,0,0)} y_i \cdot v + c_{(1,0,1,0,0,0,0,0)} y_i \cdot \theta + \dots \\
&+ c_{(3,0,0,0,0,0,0,0)} y_i^3 + c_{(1,1,0,0,0,0,0,0)} y_i \cdot v \cdot \theta + \dots \\
&\dots
\end{aligned}
$$

Sensitivity Analysis

Polynomial Expansion Revisited

Rearrange:

$$
\begin{aligned}
x_f \approx\ & c_{(0,0,0,0,0,0,0,0)} \\
& + c_{(1,0,0,0,0,0,0,0,0)} y_i + c_{(2,0,0,0,0,0,0,0,0)} y_i^2 + c_{(3,0,0,0,0,0,0,0,0)} y_i^3 + \dots \\
& + c_{(0,1,0,0,0,0,0,0,0)} v + c_{(0,2,0,0,0,0,0,0,0)} v^2 + c_{(0,3,0,0,0,0,0,0,0)} v^3 + \dots \\
& \dots \\
& + c_{(1,1,0,0,0,0,0,0,0)} y_i \cdot v + c_{(1,2,0,0,0,0,0,0,0)} y_i \cdot v^2 + \dots \\
& \dots
\end{aligned}
$$

Sensitivity Analysis

ANOVA

ANOVA: [AN]alysis [O]f [VA]riance

- How much does each input contribute to the variance?

Input	Variance	% Variance	Weight
C	0.523	0.6657	1
θ	0.236	0.3006	1/2
r	0.00868	0.0111	1/5
m	0.00862	0.0110	1/5
y_i	0.00671	0.0085	1/5
ρ	0.00209	0.0027	1/6
v	0.000348	0.0004	1/7
g	2.83×10^{-6}	3.601×10^{-6}	1/12

Idaho National Laboratory

Uncertainty Ranges

UNM

Sensitivity Analysis

Results: Expected Value, Values

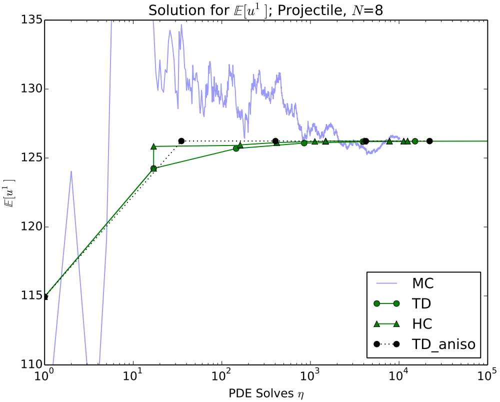


Figure: $\mathbb{E}[x_f]$

Uncertainty Quantification

Results: Expected Value, Errors

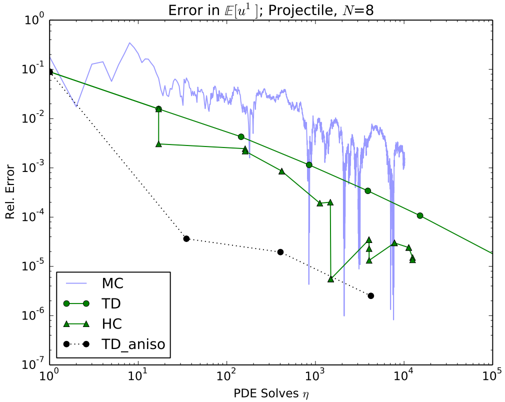


Figure: Error in $\mathbb{E}[x_f]$

Uncertainty Quantification

Results: Second Moment, Values

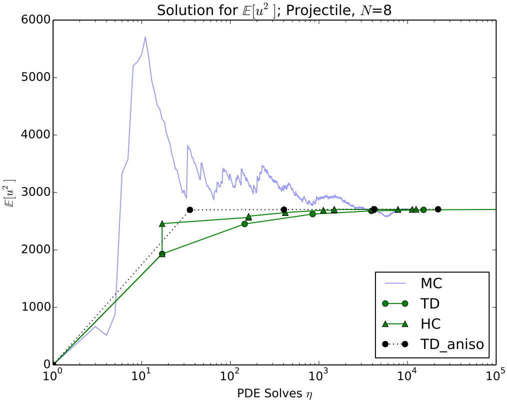


Figure: $\mathbb{E}[x_f^2]$

Uncertainty Quantification

Results: Second Moment, Errors

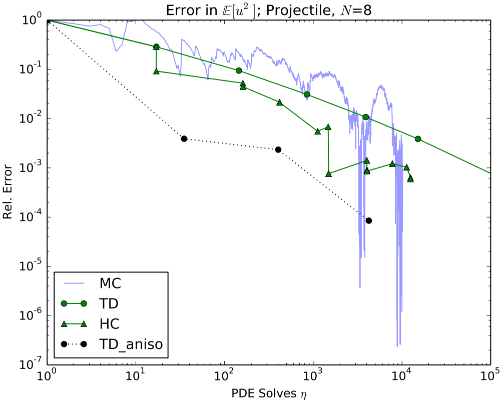


Figure: Error in $\mathbb{E}[x_f^2]$

Conclusions

- Uncertainty Quantification methods
  - Analytic methods: good, but possible deceiving
  - Numerical methods: expensive, but robust
- Sensitivity Analysis
  - Reveals importance of parameters
  - Tighten uncertainty in experiment/model
- Areas of study
  - Adaptive sampling
  - Sparse quadrature integration
  - Improved Monte Carlo methods
  - Efficient statistics algorithms

Idaho National Laboratory

UNM

Polynomial Index Sets

Choosing what polynomial degrees to use

- Tensor Product:

$$\Lambda_{\text{TP}}(L) = \Big\{\bar{p} = [p_1, ..., p_N] : \max_{1 \le n \le N} p_n \le L\Big\}, \eta = (L+1)^N$$

- Total Degree:

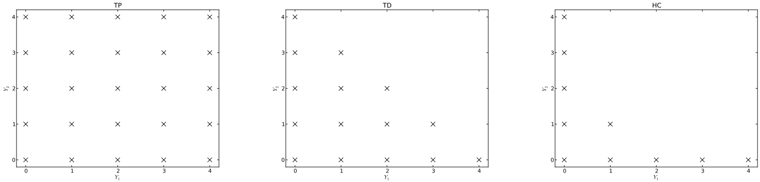
$$\Lambda_{\text{TD}}(L) = \Big\{\bar{p} = [p_1, ..., p_N] : \sum_{n=1}^{N} p_n \le L\Big\}, \eta = \binom{L+N}{N}$$

- Hyperbolic Cross:

$$\Lambda_{\text{HC}}(L) = \Big\{\bar{p} = [p_1, ..., p_N] : \prod_{n=1}^{N} p_n + 1 \le L + 1\Big\}, \eta \le (L+1)(1+\log(L+1))^{N-1}$$

Polynomial Index Sets

2D Example



(a) Tensor Product (b) Total Degree (c) Hyperbolic Cross

Figure: Index Set Examples: $N = 2, L = 4$

Calculating c_k

Where the algorithmic rubber hits the mathematical road.

$$u(Y) \approx \mathcal{S}_{N,\Lambda(L)}[u](Y) = \sum_{\boldsymbol{i}\in\Lambda(L)} c(\boldsymbol{i}) \bigotimes_{n=1}^{N} \mathcal{U}_{n,p(i_n)}[u](Y),$$

$$c(\boldsymbol{i}) = \sum_{\substack{\boldsymbol{j}=\{0,1\}^N,\\ \boldsymbol{i}+\boldsymbol{j}\in\Lambda(L)}} (-1)^{|\boldsymbol{j}|_1},$$

$$\bigotimes_{n=1}^{N} \mathcal{U}_{n,p(i_n)}[u](Y) \equiv \sum_{k_1=0}^{p(i_1)} \cdots \sum_{k_N=0}^{p(i_N)} u_h\left(Y^{(k_1)}, \cdots, Y^{(k_N)}\right) \prod_{n=1}^{N} \mathcal{L}_{k_n}(Y_n),$$

$$= \sum_{k}^{p(\vec{i})} u_h\left(Y^{(k)}\right) \mathcal{L}_k(Y),$$

Calculating c_k

2D Examples

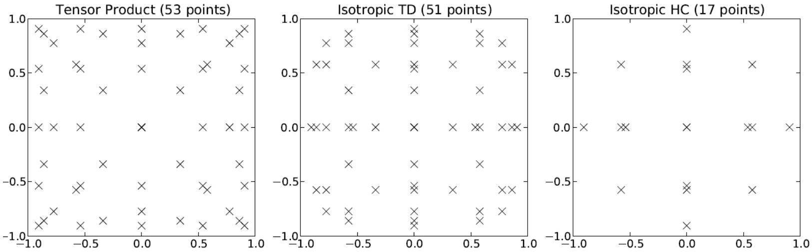


Figure: Sparse Grids, $N = 2, L = 4, p(i) = i$, Legendre points

Calculating c_k

Some Numbers

N	L	TP $\lvert\Lambda(L)\rvert$	TD $\lvert\Lambda(L)\rvert$	TD η	HC $\lvert\Lambda(L)\rvert$	HC η
3	4	125	35	165	16	31
	8	729	165	2,097	44	153
	16	4,913	969	41,857	113	513
	32	35,737	6,545	1,089,713	309	2,181
5	2	293	21	61	11	11
	4	3,125	126	781	31	71
	8	59,049	1,287	28,553	111	481

Table: Index Set and Collocation Size Comparison