Polynomial chaos expansions part I: Method Introduction

Jonathan Feinberg and Simen Tennøe

Kalkulo AS

January 14, 2015

Lecture will include many examples using the Chaospy software



Installation instructions:

http://github.com/hplgit/chaospy/

Interactive sessions:

path/to/ipython/notebook/sessions

Practicle application involving bloodflow simulations





In colaboration with V. Eck and L. Hellevik

Modelling require uncertainty quantification







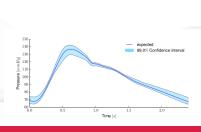
Modelling require uncertainty quantification











ntnu/results/sensitivity-poin

Introducing a naive testcase as a working example

Introducing a naive testcase as a working example

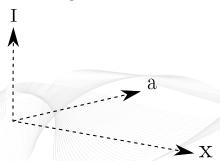
$$\frac{du(x)}{dx} = -au(x) \qquad \qquad u(0) = I$$

- u The quantity of interest
- x Spatio-temperal locations
- a, I Parameters containting uncertainties

Introducing a naive testcase as a working example

$$\frac{du(x)}{dx} = -au(x) \qquad \qquad u(0) = I$$

- u The quantity of interest
- x Spatio-temperal locations
- a, I Parameters containting uncertainties



$$u(x; a, I) = Ie^{-a(x)}$$

$$u(x; a, I) = Ie^{-a(x)}$$

Initially assume model parameters:

$$a \sim \text{Uniform}(0, 0.1)$$

$$I = 1$$

$$u(x; a, I) = Ie^{-a(x)}$$

Initially assume model parameters:

$$a \sim \text{Uniform}(0, 0.1)$$
 $I = 1$

$$E(u) = \int_0^{0.1} e^{-ax} \frac{1}{10} da = -\frac{1}{10x} \left(e^{-0.1x} - 1 \right) = \frac{1 - e^{-0.1x}}{10x}$$

$$u(x; a, I) = Ie^{-a(x)}$$

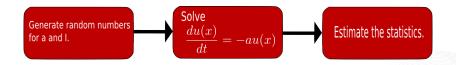
Initially assume model parameters:

$$a \sim \text{Uniform}(0, 0.1)$$
 $I = 1$

$$E(u) = \int_0^{0.1} e^{-ax} \frac{1}{10} da = -\frac{1}{10x} \left(e^{-0.1x} - 1 \right) = \frac{1 - e^{-0.1x}}{10x}$$

$$Var(u) = \int_0^{0.1} \left(e^{-ax} \right)^2 \frac{1}{10} da - E(u)^2 = \frac{1 - e^{-0.2ax}}{20x} - \left(\frac{1 - e^{-0.1x}}{10x} \right)^2$$

Non-trivial models can be analysed using Monte Carlo integration



```
import chaospy as cp
import numpy as np
def u(x, a):
  return np.exp(-a*x)
```

```
import chaospy as cp
import numpy as np
def u(x, a):
  return np.exp(-a*x)
a = cp. Uniform (0.0.1)
samples_a = a.sample(1000)
```

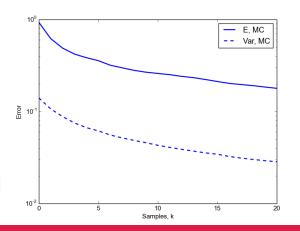
```
import chaospy as cp
import numpy as np
def u(x, a):
  return np.exp(-a*x)
a = cp.Uniform(0.0.1)
samples_a = a.sample(1000)
x = np.linspace(0, 10, 100)
```

```
import chaospy as cp
import numpy as np
def u(x, a):
  return np.exp(-a*x)
a = cp.Uniform(0.0.1)
samples_a = a.sample(1000)
x = np.linspace(0, 10, 100)
U = [u(x,q) \text{ for } q \text{ in } samples_a]
```

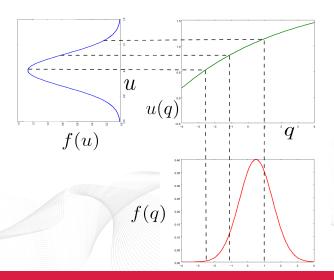
```
import chaospy as cp
import numpy as np
def u(x, a):
  return np.exp(-a*x)
a = cp.Uniform(0.0.1)
samples_a = a.sample(1000)
x = np.linspace(0, 10, 100)
U = [u(x,q) \text{ for } q \text{ in } samples_a]
E = np.mean(U)
Var = np.var(U)
```

Convergence of Monte Carlo is slow

$$\varepsilon_E = \int_0^{10} |E(u) - E(\hat{u})| dx$$
 $\varepsilon_{Var} = \int_0^{10} |Var(u) - Var(\hat{u})| dx$



Monte Carlo is based on the idea of indirect sampling



Using Lagrange polynomials to approximate the model

$$u(x; a) \approx \hat{u}_M(x; a) = \sum_{n=0}^{N} c_n(x) P_n(a)$$
 $N = M + 1,$

where

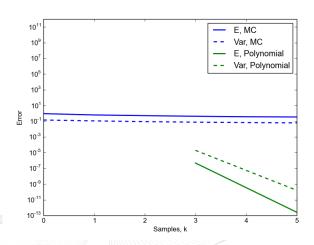
 c_n are model evaluations $u(x, a_n)$

 P_n are Lagrange polynomials:

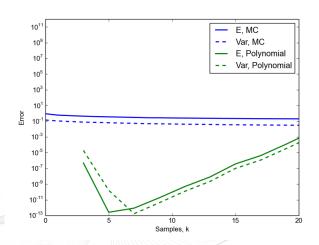
$$P_n(a) = \prod_{\substack{m=0\\m\neq n}}^N \frac{a - a_n}{a_m - a_n}$$

an are collocation nodes

Much better convergence properties than Monte Carlo



Stochastic analysis of model approximations can be dangerous



Combining polynomial approximation with uncertainty quantification well requires a better underlying theory

$$\langle u, v \rangle_Q = E(u \cdot v)$$
 $||u||_Q = \sqrt{\langle u, u \rangle_Q}$

where Q is a random vector, i.e. (a, I).

Combining polynomial approximation with uncertainty quantification well requires a better underlying theory

$$\langle u, v \rangle_Q = E(u \cdot v) \qquad ||u||_Q = \sqrt{\langle u, u \rangle_Q}$$

= $\int f_Q(q) u(x, q) v(x, q) dq$

where Q is a random vector, i.e. (a, I).

Combining polynomial approximation with uncertainty quantification well requires a better underlying theory

$$\langle u, v \rangle_Q = E(u \cdot v) \qquad ||u||_Q = \sqrt{\langle u, u \rangle_Q}$$

= $\int f_Q(q)u(x, q)v(x, q)dq$

where Q is a random vector, i.e. (a, I).

Orthogonality:

$$\langle P_n, P_m \rangle = \begin{cases} \|P_n\|_Q^2 & n = m \\ 0 & n \neq m \end{cases}$$

$$u = \sum_{n=0}^{N} c_n P_n$$

$$u = \sum_{n=0}^{N} c_n P_n$$

$$\langle u, P_k \rangle_Q = \left\langle \sum_{n=0}^{N} c_n P_n, P_k \right\rangle_Q \qquad k = 0, \dots, N$$

$$u = \sum_{n=0}^{N} c_n P_n$$

$$\langle u, P_k \rangle_Q = \left\langle \sum_{n=0}^{N} c_n P_n, P_k \right\rangle_Q$$

$$= \sum_{n=0}^{N} c_n \langle P_n, P_k \rangle_Q \qquad k = 0, \dots, N$$

$$u = \sum_{n=0}^{N} c_n P_n$$

$$\langle u, P_k \rangle_Q = \left\langle \sum_{n=0}^{N} c_n P_n, P_k \right\rangle_Q$$

$$= \sum_{n=0}^{N} c_n \langle P_n, P_k \rangle_Q$$

$$= c_k \langle P_k, P_k \rangle_Q \qquad k = 0, \dots, N$$

$$u = \sum_{n=0}^{N} c_n P_n$$

$$\langle u, P_k \rangle_Q = \left\langle \sum_{n=0}^{N} c_n P_n, P_k \right\rangle_Q$$

$$= \sum_{n=0}^{N} c_n \langle P_n, P_k \rangle_Q$$

$$= c_k \langle P_k, P_k \rangle_Q$$

$$c_k = \frac{\langle u, P_k \rangle_Q}{\|P_k\|_Q^2}$$

$$k = 0, \ldots, N$$

$$(c_0,\ldots,c_N) = \underset{c_0,\ldots,c_N}{\operatorname{argmin}} \|u - \hat{u}_M\|_Q$$

$$(c_0, \dots, c_N) = \underset{c_0, \dots, c_N}{\operatorname{argmin}} \|u - \hat{u}_M\|_Q$$
$$= \underset{c_0, \dots, c_N}{\operatorname{argmin}} \|u - \hat{u}_M\|_Q^2$$

$$(c_0, \dots, c_N) = \underset{c_0, \dots, c_N}{\operatorname{argmin}} \|u - \hat{u}_M\|_Q$$
$$= \underset{c_0, \dots, c_N}{\operatorname{argmin}} \|u - \hat{u}_M\|_Q^2$$
$$= \underset{c_0, \dots, c_N}{\operatorname{argmin}} E((u - \hat{u}_M)^2)$$

$$(c_0, \dots, c_N) = \underset{c_0, \dots, c_N}{\operatorname{argmin}} \| u - \hat{u}_M \|_Q$$

$$= \underset{c_0, \dots, c_N}{\operatorname{argmin}} \| u - \hat{u}_M \|_Q^2$$

$$= \underset{c_0, \dots, c_N}{\operatorname{argmin}} \mathsf{E}((u - \hat{u}_M)^2)$$

$$= \underset{c_0, \dots, c_N}{\operatorname{argmin}} \mathsf{Var}(u - \hat{u}_M)$$

The mean and variance has a simpler form

Assumption:
$$P_0 = 1$$

$$\mathsf{E}(\hat{u}_M) = \mathsf{E}\left(\sum_{n=0}^N c_n P_n\right)$$

$$E(\hat{u}_M) = E\left(\sum_{n=0}^{N} c_n P_n\right)$$
$$= \sum_{n=0}^{N} c_n E(P_n)$$

$$E(\hat{u}_M) = E\left(\sum_{n=0}^{N} c_n P_n\right)$$
$$= \sum_{n=0}^{N} c_n E(P_n)$$
$$= \sum_{n=0}^{N} c_n \langle P_n, P_0 \rangle_Q$$

$$E(\hat{u}_M) = E\left(\sum_{n=0}^N c_n P_n\right)$$
$$= \sum_{n=0}^N c_n E(P_n)$$
$$= \sum_{n=0}^N c_n \langle P_n, P_0 \rangle_Q$$

$$\mathsf{E}(\hat{u}_{M}) = c_{0}$$

$$\begin{aligned} \mathsf{E}(\hat{u}_M) &= \mathsf{E}\bigg(\sum_{n=0}^N c_n P_n\bigg) & \mathsf{Var}(\hat{u}_M) &= \mathsf{Var}\bigg(\sum_{n=0}^N c_n P_n\bigg) \\ &= \sum_{n=0}^N c_n \mathsf{E}(P_n) \\ &= \sum_{n=0}^N c_n \langle P_n, P_0 \rangle_Q \end{aligned}$$

$$\mathsf{E}(\hat{u}_M) = c_0$$

$$\begin{split} \mathsf{E}(\hat{u}_M) &= \mathsf{E}\bigg(\sum_{n=0}^N c_n P_n\bigg) \qquad \mathsf{Var}(\hat{u}_M) = \mathsf{Var}\bigg(\sum_{n=0}^N c_n P_n\bigg) \\ &= \sum_{n=0}^N c_n \mathsf{E}(P_n) \qquad \qquad = \sum_{n=0}^N c_n c_m \left(\mathsf{E}(P_n P_m) - \mathsf{E}(P_n)\mathsf{E}(P_m)\right) \\ &= \sum_{n=0}^N c_n \left\langle P_n, P_0 \right\rangle_Q \end{split}$$

$$\mathsf{E}(\hat{u}_{M}) = c_{0}$$

$$\begin{split} \mathsf{E}(\hat{u}_M) &= \mathsf{E}\left(\sum_{n=0}^N c_n P_n\right) \qquad \mathsf{Var}(\hat{u}_M) = \mathsf{Var}\left(\sum_{n=0}^N c_n P_n\right) \\ &= \sum_{n=0}^N c_n \mathsf{E}(P_n) \qquad \qquad = \sum_{\substack{n=0\\m=0}}^N c_n c_m \left(\mathsf{E}(P_n P_m) - \mathsf{E}(P_n)\mathsf{E}(P_m)\right) \\ &= \sum_{n=0}^N c_n \langle P_n, P_0 \rangle_Q \qquad \qquad = \sum_{\substack{n=0\\m=0}}^N c_n c_m \langle P_n, P_m \rangle_Q - c_0^2 \end{split}$$

$$\mathsf{E}(\hat{u}_M) = c_0$$

$$\begin{split} \mathsf{E}(\hat{u}_M) &= \mathsf{E}\left(\sum_{n=0}^N c_n P_n\right) \qquad \mathsf{Var}(\hat{u}_M) = \mathsf{Var}\left(\sum_{n=0}^N c_n P_n\right) \\ &= \sum_{n=0}^N c_n \mathsf{E}(P_n) \qquad \qquad = \sum_{n=0}^N c_n c_m \left(\mathsf{E}(P_n P_m) - \mathsf{E}(P_n) \mathsf{E}(P_m)\right) \\ &= \sum_{n=0}^N c_n \left\langle P_n, P_0 \right\rangle_Q \qquad \qquad = \sum_{n=0}^N c_n c_m \left\langle P_n, P_m \right\rangle_Q - c_0^2 \\ &= \mathsf{E}(\hat{u}_M) = c_0 \qquad \qquad \mathsf{Var}(\hat{u}_M) = \sum_{n=0}^N c_n^2 \left\| P_n \right\|_Q \end{split}$$

$$v_0, v_1, ..., v_N = 1, q, ..., q^N$$

$$v_0, v_1, ..., v_N = 1, q, ..., q^N$$

The Gram Schmidt method is

$$P_0 = v_0$$

$$v_0, v_1, ..., v_N = 1, q, ..., q^N$$

The Gram Schmidt method is

$$P_0 = v_0$$

$$P_n = v_n - \sum_{m=0}^{n-1} \frac{\langle v_n, P_m \rangle_Q}{\|P_m\|_Q^2}$$

$$v_0, v_1, ..., v_N = 1, q, ..., q^N$$

The Gram Schmidt method is

$$P_{0} = v_{0}$$

$$P_{n} = v_{n} - \sum_{m=0}^{n-1} \frac{\langle v_{n}, P_{m} \rangle_{Q}}{\|P_{m}\|_{Q}^{2}}$$

$$= v_{n} - \sum_{m=0}^{n-1} \frac{E(v_{n}P_{m})}{E(P_{m}^{2})}$$

Gram-Schmidt with chaospy

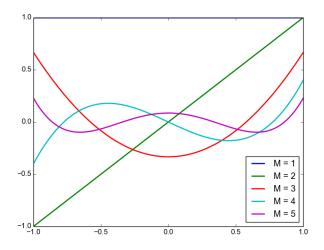
Gram-Schmidt with chaospy

```
M = 5; D = 1; N = M - 1
a = cp.Uniform(0,0.1)
v = cp.basis(0,M,1)
P = [v[0]]
```

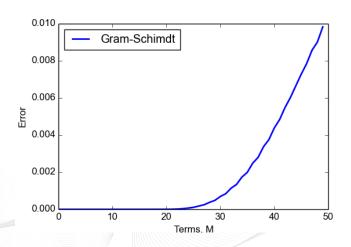
Gram-Schmidt with chaospy

```
M = 5; D = 1; N = M - 1
a = cp. Uniform(0,0.1)
v = cp. basis (0, M, 1)
P = [v[0]]
for n in xrange(1,N):
    p = v[n]
    for m in xrange(0,n):
         p = P[m] * cp. E(v[n] * P[m], a)
                               /cp.E(P[m]**2,a)
    P.append(p)
P = cp.Polv(P)
```

Plot of all generated polynomials



Most constructors of orthogonal polynomials are illposed



The only numerically stable method for calculating orthogonal polynomials is through the discretized Stiltjes method

The only numerically stable method for calculating orthogonal polynomials is through the discretized Stiltjes method

Three terms recursion relation:

$$P_{n+1} = (x - A_n)P_n - B_nP_{n-1}$$
 $P_{-1} = 0$ $P_0 = 1$,

The only numerically stable method for calculating orthogonal polynomials is through the discretized Stiltjes method

Three terms recursion relation:

$$P_{n+1} = (x - A_n)P_n - B_nP_{n-1}$$
 $P_{-1} = 0$ $P_0 = 1$,

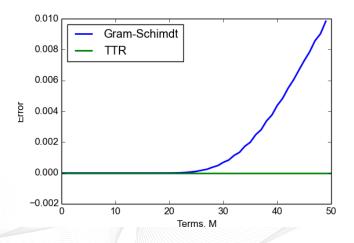
where

$$A_{n} = \frac{\langle qP_{n}, P_{n}\rangle_{Q}}{\|P_{n}\|_{Q}^{2}} \qquad B_{n} = \begin{cases} \frac{\|P_{n}\|_{Q}^{2}}{\|P_{n-1}\|_{Q}^{2}} & n > 0\\ \|P_{n}\|_{Q}^{2} & n = 0 \end{cases}$$

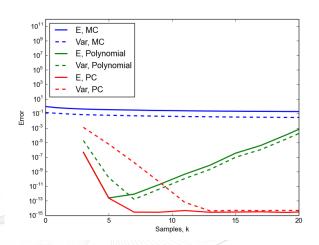
Three terms recursion in Chaospy

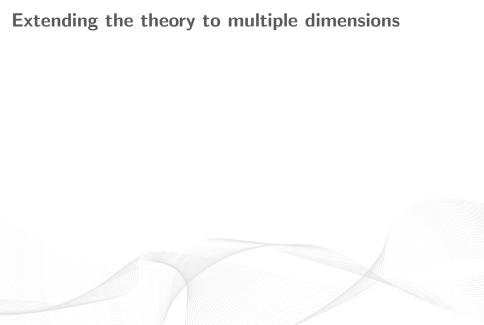
```
import chaospy as cp
M = 3
a = cp.Normal()
P = cp.orth_ttr(M, a)
print P
[1.0, q0, q0^2-1.0, q0^3-3.0q0]
```

Discretized Stiltjes method is numerically stable



Convergence of orthogonal polynomial approximation

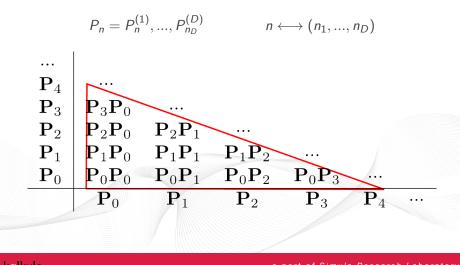




Extending the theory to multiple dimensions

$$P_n = P_n^{(1)}, ..., P_{n_D}^{(D)}$$
 $n \longleftrightarrow (n_1, ..., n_D)$

Extending the theory to multiple dimensions



Construct polynomial approximation

Multi-index

 $\begin{array}{cccc} & P_{00} \\ & P_{10} & P_{01} \\ & P_{20} & P_{11} & P_{02} \\ P_{30} & P_{21} & P_{12} & ... \end{array}$

Single-index

$$\langle \mathbf{P}_n, \mathbf{P}_m \rangle_Q = \mathsf{E} \Big(P_{n_1}^{(1)} \cdots P_{n_D}^{(D)} \cdot P_{m_1}^{(1)} \cdots P_{m_D}^{(D)} \Big)$$

$$\langle \mathbf{P}_{n}, \mathbf{P}_{m} \rangle_{Q} = \mathbb{E} \left(P_{n_{1}}^{(1)} \cdots P_{n_{D}}^{(D)} \cdot P_{m_{1}}^{(1)} \cdots P_{m_{D}}^{(D)} \right)$$

 $= \mathbb{E} \left(P_{n_{1}}^{(1)} \cdot P_{m_{1}}^{(1)} \right) \cdots \mathbb{E} \left(P_{n_{D}}^{(D)} \cdot P_{m_{D}}^{(D)} \right)$

$$\langle \mathbf{P}_{n}, \mathbf{P}_{m} \rangle_{Q} = \mathbb{E} \left(P_{n_{1}}^{(1)} \cdots P_{n_{D}}^{(D)} \cdot P_{m_{1}}^{(1)} \cdots P_{m_{D}}^{(D)} \right)$$

$$= \mathbb{E} \left(P_{n_{1}}^{(1)} \cdot P_{m_{1}}^{(1)} \right) \cdots \mathbb{E} \left(P_{n_{D}}^{(D)} \cdot P_{m_{D}}^{(D)} \right)$$

$$= \left\langle P_{n_{1}}^{(1)} \cdot P_{m_{1}}^{(1)} \right\rangle_{Q} \cdots \left\langle P_{n_{D}}^{(D)} \cdot P_{m_{D}}^{(D)} \right\rangle_{Q}$$

$$\langle \mathbf{P}_{n}, \mathbf{P}_{m} \rangle_{Q} = \mathbb{E} \left(P_{n_{1}}^{(1)} \cdots P_{n_{D}}^{(D)} \cdot P_{m_{1}}^{(1)} \cdots P_{m_{D}}^{(D)} \right)$$

$$= \mathbb{E} \left(P_{n_{1}}^{(1)} \cdot P_{m_{1}}^{(1)} \right) \cdots \mathbb{E} \left(P_{n_{D}}^{(D)} \cdot P_{m_{D}}^{(D)} \right)$$

$$= \left\langle P_{n_{1}}^{(1)} \cdot P_{m_{1}}^{(1)} \right\rangle_{Q} \cdots \left\langle P_{n_{D}}^{(D)} \cdot P_{m_{D}}^{(D)} \right\rangle_{Q}$$

$$= \left\| P_{n_{1}}^{(1)} \right\|_{Q} \delta_{n_{1}m_{1}} \cdots \left\| P_{n_{D}}^{(D)} \right\|_{Q} \delta_{n_{D}m_{D}}$$

$$\langle \mathbf{P}_{n}, \mathbf{P}_{m} \rangle_{Q} = \mathbb{E} \left(P_{n_{1}}^{(1)} \cdots P_{n_{D}}^{(D)} \cdot P_{m_{1}}^{(1)} \cdots P_{m_{D}}^{(D)} \right)$$

$$= \mathbb{E} \left(P_{n_{1}}^{(1)} \cdot P_{m_{1}}^{(1)} \right) \cdots \mathbb{E} \left(P_{n_{D}}^{(D)} \cdot P_{m_{D}}^{(D)} \right)$$

$$= \left\langle P_{n_{1}}^{(1)} \cdot P_{m_{1}}^{(1)} \right\rangle_{Q} \cdots \left\langle P_{n_{D}}^{(D)} \cdot P_{m_{D}}^{(D)} \right\rangle_{Q}$$

$$= \left\| P_{n_{1}}^{(1)} \right\|_{Q} \delta_{n_{1}m_{1}} \cdots \left\| P_{n_{D}}^{(D)} \right\|_{Q} \delta_{n_{D}m_{D}}$$

$$= \left\| \mathbf{P}_{n} \right\|_{Q} \delta_{n_{m}}$$

```
a = cp.Uniform(0, 0.1)
I = cp. Uniform(8, 10)
dist = cp.J(a,I)
```

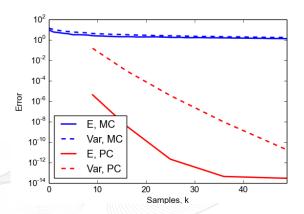
```
a = cp.Uniform(0, 0.1)
I = cp. Uniform(8, 10)
dist = cp.J(a,I)
P = cp.orth_ttr(1, dist)
print P
[1.0, q1-9.0, q0]
```

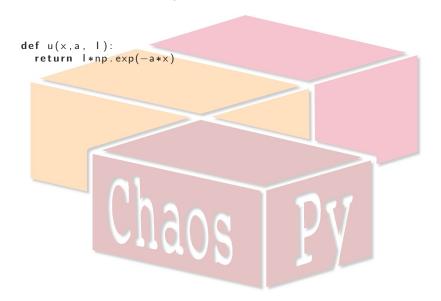
```
a = cp.Uniform(0, 0.1)
I = cp. Uniform(8, 10)
dist = cp.J(a,I)
P = cp. orth_ttr(1, dist)
print P
[1.0, q1-9.0, q0]
P = cp.orth_ttr(3, dist)
print cp.E(P[1]*P[2], dist)
0.0
```

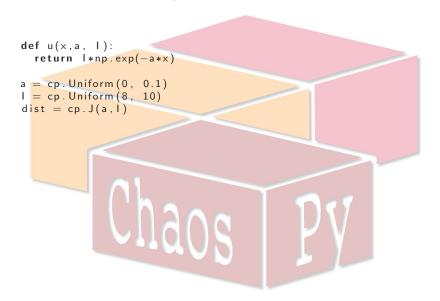
```
a = cp.Uniform(0, 0.1)
I = cp. Uniform(8, 10)
dist = cp.J(a,I)
P = cp. orth_ttr(1, dist)
print P
[1.0, q1-9.0, q0]
P = cp.orth_ttr(3, dist)
print cp.E(P[1]*P[2], dist)
0 0
print cp.E(P[3]*P[3], dist)
0.088888888903
```

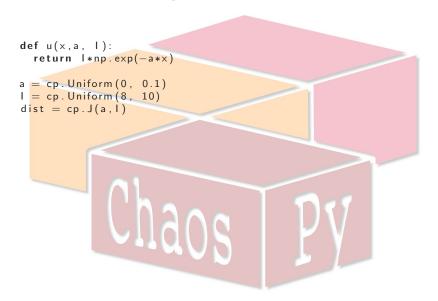
Convergence of a multidimensional problem

$$arepsilon_E = \int_0^{10} |E(u) - E(\hat{u})| \, dx \quad arepsilon_{Var} = \int_0^{10} |Var(u) - Var(\hat{u})| \, dx$$





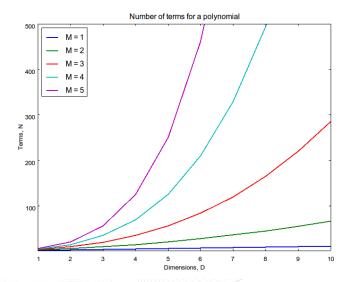




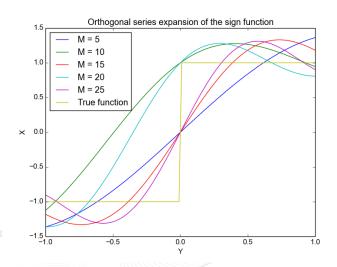
```
def u(x,a, 1):
  return 1*np.exp(-a*x)
a = cp. Uniform(0, 0.1)
I = cp. Uniform(8, 10)
dist = cp.J(a,I)
P = cp.orth_ttr(2, dist)
```

```
def u(x,a,l):
  return 1*np.exp(-a*x)
a = cp. Uniform(0, 0.1)
I = cp. Uniform(8, 10)
dist = cp.J(a,I)
P = cp. orth_ttr(2, dist)
nodes, weights = \
    cp.generate_quadrature(3, dist, rule="G")
x = np.linspace(0, 10, 100)
solves = [u(x, *node) for node in nodes.T]
u_hat = cp.fit_quadrature(P, nodes, weights, solves)
mean, var = cp.E(u_hat, dist), cp.Var(u_hat, dist)
```

The curse of dimensionality



Gibb's Phenomena, discontinues methods are troublesome



Higher number of samples justifies higher number of collocation nodes

