Uncertainty quantification in computer experiments with polynomial chaos

J. KO[1] with J. GARNIER[2], D. LUCOR[3] & A. DIPANKAR[4]

1. j o r d a n . k o @ m a c . c o m
2. Laboratoire de Probabilités et Modèles Aléatoires, Université de Paris VII, France
3. L'Institut Jean Le Rond d'Alembert, Université de Paris VI, France
4. Max Planck Institute for Meteorology, Hamburg, Germany

Workshop on uncertainty quantification, risk and decision-making
Centre for the analysis of time series, LSE
May 23, 2012

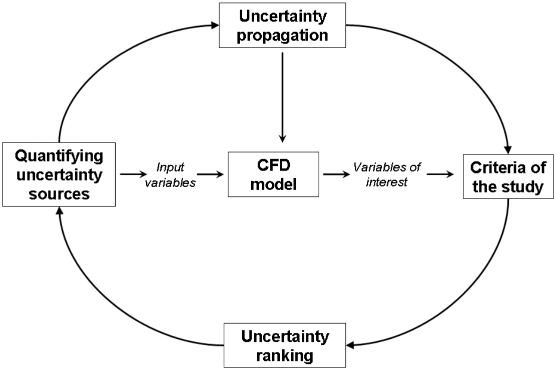
Uncertainty quantification (UQ) in computer experiments

- Context: Deterministic and complex numerical simulator are used to model real dynamic systems and they can be computationally expensive to run
- We are interested to study the effect of epistemic (lack of knowledge) and aleatoric (inherent to system) uncertainties on the model outputs
- Sources include initial condition, boundary condition & model parameters
- Example: drug clearance in circulation as an exponential decay response $\frac{d\theta}{dt} = -C\theta$ with C as a r.v. that represents the population response
- Conventional approaches such as MC are not practical in studying these expensive simulators
- Goal: PC construct a metamodel that mimics the complex model's behaviour and conduct UQ, SA, quantile estimation, optimization, calibration, *etc*.

Probabilistic framework

The UQ of a computer experiment follows the following iterative steps:

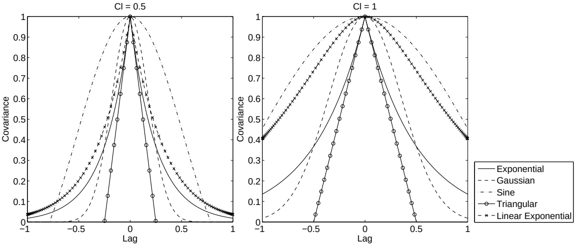
1. representation of input uncertainties - random variable or process
2. uncertainty propagation - MC, GP or gPC
3. quantification of solution uncertainty - mean, variance, pdf or sensitivity



De Rocquigny (2006)

Stochastic input representation: stochastic process

Any second order random process $\kappa(x, \omega)$, with continuous and bounded covariance kernel $C(x_1, x_2) = \mathbb{E}(\kappa(x_1, \omega) \otimes \kappa(x_2, \omega))$, can be represented as an infinite sum of random variables. It is real, symmetric and positive–definite.



- Karhunen-Loève (KL) expansion represents the random process with an orthogonal set of deterministic functions with random coefficients as

$$\kappa(x, \omega) = \mu_\kappa(x) + \sum_{n=1}^{N} \sqrt{\lambda_n} \psi_n(x) \xi_n(\omega).$$

- For a continuous kernel, the convergence of the KL expansion is uniform as $N \to \infty$. Karhunen (1948) & Loève (1977)
- $\psi_n(x)$ and λ_n solved from Fredholm integral equation of 2nd kind with $C(x_1, x_2)$.

Stochastic input representation: random variables

- Represent the random variable, $\kappa(\omega)$, with orthogonal functions of the stochastic variable with deterministic coefficients

$$\kappa(\omega) = \sum_{m=0}^{\infty} \kappa_m \phi_m(\boldsymbol{\xi}(\omega)).$$

- Wiener-Chaos: representation of a Gaussian random variable using Hermite polynomials with L^2 convergence as $M \to \infty$. Wiener (1938), Ghanem & Spanos (1991) and Cameron & Martin (1947)
- generalized Polynomial Chaos: generalized representation to non-Gaussian random variables with polynomials from the Wiener–Askey scheme. Xiu & Karniadakis (2002)
- if $\kappa(\omega)$ follows a normal distribution, it can be represented exactly as $\kappa(\omega) = \mu_\kappa + \sigma_\kappa \boldsymbol{\xi}$ where $\boldsymbol{\xi}$ is the linear term in Hermite

Selection of orthogonal basis

- In the propagation step, we need to evaluate the inner product w.r.t. the probability space measure, $\rho(\boldsymbol{\xi})d\boldsymbol{\xi}$ as

$$\langle \phi_i(\boldsymbol{\xi}), \phi_j(\boldsymbol{\xi}) \rangle = \int_\Gamma \phi_i(\boldsymbol{\xi})\phi_j(\boldsymbol{\xi})\rho(\boldsymbol{\xi})d\boldsymbol{\xi}.$$

- Correspondence between the *pdf* of ξ, $\rho(\xi)$, and the weighting function of classical orthogonal polynomials, $w(\xi)$, determines the polynomial basis

Distribution	Random variable, ξ	Wiener-Askey PC, $\phi(\xi)$	Support, Γ
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	$\{0, 1, \dots, N\}$
Periodic	uniform	Fourier-chaos*	$[-\pi, \pi)$

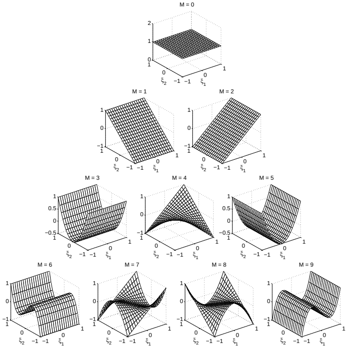
Multivariate basis

Multivariate basis is the tensor products of 1D polynomials

$$
\begin{aligned}
\phi_m(\boldsymbol{\xi}) &= \phi^{\alpha_{m,n=1}}(\xi_1) \otimes \phi^{\alpha_{m,n=2}}(\xi_2) \otimes \cdots \otimes \phi^{\alpha_{m,n=N}}(\xi_N), \quad \text{for } m = 0, \cdots, M, \\
&= \phi^{\boldsymbol{\alpha}_m}(\boldsymbol{\xi}), \quad \text{for } m = 0, \cdots, M.
\end{aligned}
$$

Truncation depends on input dimension, N, and output nonlinearity, P

m	P	$\prod$ Notation	Legendre Polynomials
0	0	$P^0(\xi_1)P^0(\xi_2)$	1
1	1	$P^1(\xi_1)P^0(\xi_2)$	ξ_1
2		$P^0(\xi_1)P^1(\xi_2)$	ξ_2
3	2	$P^2(\xi_1)P^0(\xi_2)$	$3/2\xi_1^2 - 1/2$
4		$P^1(\xi_1)P^1(\xi_2)$	$\xi_1\xi_2$
5		$P^0(\xi_1)P^2(\xi_2)$	$3/2\xi_2^2 - 1/2$
6	3	$P^3(\xi_1)P^0(\xi_2)$	$5/2\xi_1^3 - 3/2\xi_1$
7		$P^2(\xi_1)P^1(\xi_2)$	$3/2\xi_2\xi_1^2 - 1/2\xi_2$
8		$P^1(\xi_1)P^2(\xi_2)$	$3/2\xi_1\xi_2^2 - 1/2\xi_1$
9		$P^0(\xi_1)P^3(\xi_2)$	$5/2\xi_2^3 - 3/2\xi_2$



Stochastic Galerkin method: intrusive approach

PC represent the stochastic solution $u(\mathbf{x}, \boldsymbol{\xi})$ with the same orthogonal basis as the input, *i.e.* $u(\mathbf{x}, \boldsymbol{\xi}) = \sum u_m(\mathbf{x})\phi_m(\boldsymbol{\xi})$
Substitute the expansions into the system of equations, $\mathcal{L}(\mathbf{x}, \boldsymbol{\xi}; u) = f(\mathbf{x}, \boldsymbol{\xi})$.
Take the Galerkin projection, *i.e.*

$$\langle \mathcal{L}\left(\mathbf{x}, \boldsymbol{\xi}; \sum u_m(\mathbf{x})\phi_m(\boldsymbol{\xi})\right), \phi_m(\boldsymbol{\xi})\rangle \quad = \quad \langle f\left(\mathbf{x}, \boldsymbol{\xi}\right), \phi_m(\boldsymbol{\xi})\rangle, \quad \text{for } m = 0, \ldots, M.$$

- $u_m(\mathbf{x})$ are solved from the system of $(M+1)$ coupled equations.
- The system is deterministic and can be solved using a standard discretization technique.
- Extensive modification on the simulator is needed.

Stochastic Galerkin method: intrusive approach

Example

First-order linear ODE: $\dot{\Theta}(t, \boldsymbol{\xi}) = -C(\boldsymbol{\xi})\Theta(t, \boldsymbol{\xi})$ with rate of decay as a normal r.v., *i.e.* $C(\boldsymbol{\xi}) = \sum_{i=0}^{M} C_i \phi_i(\boldsymbol{\xi})$. The gPC expansions of $C(\boldsymbol{\xi})$ and $\Theta(t, \boldsymbol{\xi})$ are substituted into the ODE to give

$$\sum_{k=0}^{M_\theta} \dot{\Theta}_k(t)\phi_k(\boldsymbol{\xi}) \quad = \quad -\sum_{i=0}^{M_C} \sum_{j=0}^{M_\theta} C_i \Theta_j(t)\phi_i(\boldsymbol{\xi})\phi_j(\boldsymbol{\xi}).$$

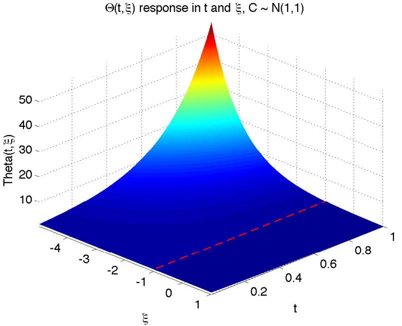
The Galerkin projection of the expanded ODE with orthogonal polynomial:

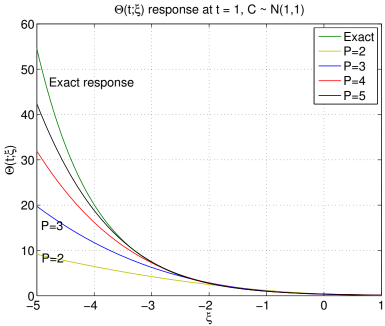
$$\dot{\Theta}_k(t) = -\sum_{i=0}^{M_C} \sum_{j=0}^{M_\theta} \frac{\langle \phi_i \phi_j \phi_k \rangle}{\langle \phi_k^2 \rangle} C_i \Theta_j(t), \quad \text{for k} = 0, \ldots, M_\theta.$$

This coupled deterministic system of equations is solved with an initial condition $\Theta(t=0) = \sum \Theta_m(t=0)\phi_m(\boldsymbol{\xi})$. With increasing t, the modal coefficients are propagated from the lower Θ_m to higher Θ_m, *i.e.* propagation of uncertainty as increasing non–linear response in the random space.

Surface response of the linear ODE

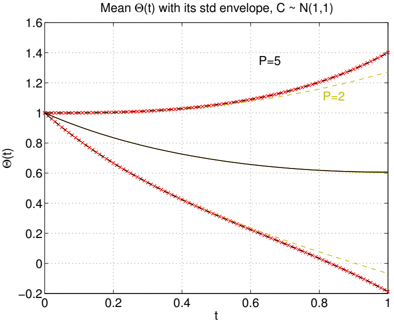
- $\dot{\Theta}(t, \boldsymbol{\xi}) = -C(\boldsymbol{\xi})\Theta(t, \boldsymbol{\xi})$
- $\Theta(t, \xi)$ response is exponential in t with $\Theta(t = 0) = 1$.
- Treating the coefficient of decay as a random variable, $C(\xi) \sim \mathcal{N}(1, 1)$
- We represent the univariate stochastic output $\Theta(t; \xi)$ as a linear combination of Hermite polynomials $\Theta(t; \boldsymbol{\xi}) = \sum \Theta_m(t)\phi_m(\boldsymbol{\xi})$.
- Uncertainty propagation visualized as solution response surface evolution in random space, ξ

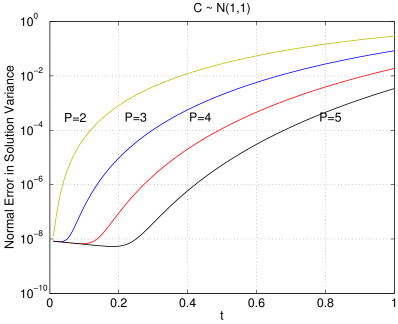




The choice of polynomial chaos truncation

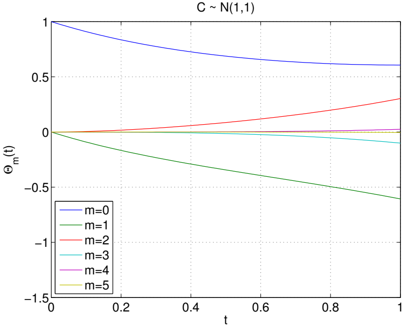
- As response in ξ becomes more non–linear with t, the higher order P in $\phi_m(\xi)$ are needed in gPC expansion
- Estimation of higher order statistics also require higher P
- Premature truncation leads to large error in the response surface and the solution statistics

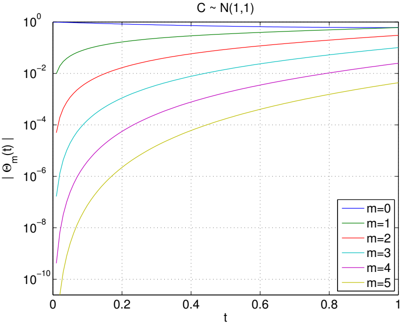




Evolution of the PC coefficients

- Increasing t propagates the initial uncertainty from lower order coefficients to higher order coefficients





- The task now is to determine the coefficients of expansion, $\Theta_m(t)$ in the representation.
- This simple system of equation easily solved with the intrusive approach
- Complex numerical solvers can benefit from a non–intrusive approach

Probabilistic collocation method (PCM)

Projecting directly the stochastic solution, $u(\mathbf{x}, \boldsymbol{\xi}) = \sum u_m(\mathbf{x})\phi_m(\boldsymbol{\xi})$, onto the orthogonal basis, $\phi_m(\boldsymbol{\xi})$, we obtain the following $(M+1)$ decoupled equations:

$$u_m(\mathbf{x}) = \frac{\langle u(\mathbf{x}, \boldsymbol{\xi}), \phi_m(\boldsymbol{\xi})\rangle}{\langle \phi_m^2(\boldsymbol{\xi})\rangle}, \quad \text{for } m = 0, ..., M.$$

The inner–product can be evaluated using Monte Carlo or related methods. We investigate a numerical quadrature approach to approximate the inner product where the numerical solver is treated as a black box from which samples are repeated taken.

One–dimensional quadrature rules

Integrals are approximated as the weighted sum of function evaluations on deterministic quadrature points, *i.e.*

$$\begin{aligned}\langle u(\mathbf{x},\boldsymbol{\xi}),\phi_m(\boldsymbol{\xi})\rangle &= \int_{\boldsymbol{\Gamma}} u(\mathbf{x},\boldsymbol{\xi})\phi_m(\boldsymbol{\xi})\rho(\boldsymbol{\xi})d\boldsymbol{\xi},\\ &\approx \sum_{j=0}^{N_q} w_j u(\mathbf{x},\mathbf{z}_j)\phi_m(\mathbf{z}_j).\end{aligned}$$

The accuracy of the method depends on the selection of the quadrature approach, *i.e.* constructions of w_j and $\mathbf{z}_j$.

	Γ	P	N_q	Nestedness
Gauss-Legendre	(-1,1)	$2L-1$	L	No
Clenshaw-Curtis	[-1,1]	$L-1$	$2^{L-1}+1$	Yes
Gauss-Laguerre	$[0,\infty)$	$2L-1$	L	No
Gauss-Hermite	$(-\infty,\infty)$	$2L-1$	L	No
Hermite Kronrod-Patterson	$(-\infty,\infty)$	$2m+n-1^{*}$	1-3-9-19-35 or 1-4-18-30	Yes

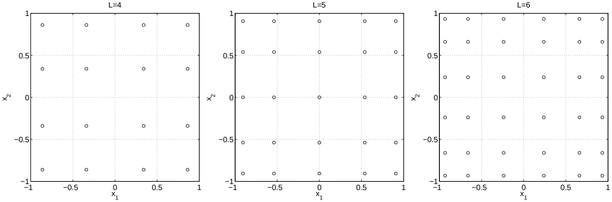
Multi–dimensional quadrature rules are constructed from 1D quadrature rules.

Full–tensor quadrature

Multi–dimensional full–tensor quadrature relies on tensor product of 1D quadrature rules, *e.g.* N–dimensional quadrature points are

$$\mathcal{Q}_L^N(f) = (\mathcal{U}^{i_1} \otimes \cdots \otimes \mathcal{U}^{i_N})(f).$$

Example: Two–dimensional Gauss–Legendre quadrature:



Accuracy: Theoretical polynomial exactness $P = 2L - 1$ in each dimension where L is the number of quadrature points in each dimension
Cost: Number of quadrature points grows as $\mathcal{O}(L^N)$ and error converges as $\epsilon(Z) = \mathcal{O}(Z^{-r/N})$. – *"curse of dimensionality"*

Sparse quadrature: the Smolyak approach

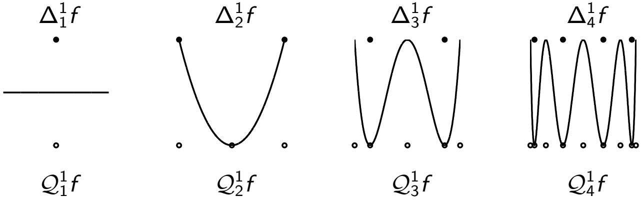
"Curse of dimensionality" could be 'broken' with the sparse grid. Its construction is based on the following three steps: Gerstner & Griebel (1998)

1. Constructed from 1D difference grid
2. Tensor product of 1D difference grids: cost reduction
3. Linear combination of the tensor products: embeddedness → refinement cost reduction

Accuracy: Theoretical polynomial exactness at least $P \leq 2L - 1$ where L is the quadrature level. Smolyak (1963), Novak & Ritter (1996)
Cost: Error converges as $\epsilon(Z) = \mathcal{O}(Z^{-r}(\log(Z)^{(N-1)(r+1)}))$. Novak & Ritter (1996)

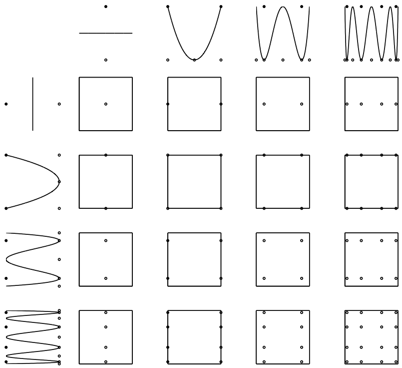
Sparse quadrature: with nested Clenshaw-Curtis quadrature rule

1D difference grid: $\triangle_k^1 f := \left(\mathcal{Q}_k^1 - \mathcal{Q}_{k-1}^1\right) f$



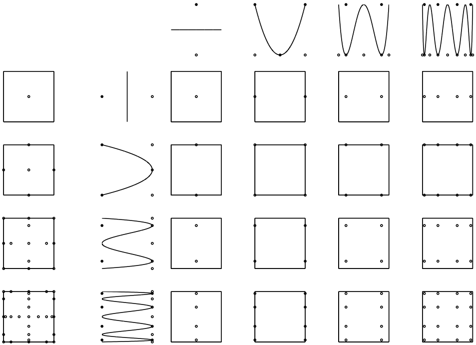
Sparse quadrature: with nested Clenshaw-Curtis quadrature rule

1D difference grid: $\triangle_k^1 f := \left(\mathcal{Q}_k^1 - \mathcal{Q}_{k-1}^1\right) f$
Tensor product: $\left(\triangle_{k_1}^1 \otimes \cdots \otimes \triangle_{k_N}^1\right) f$



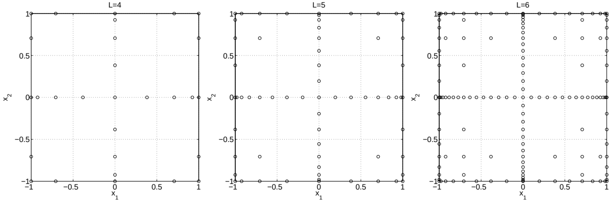
Sparse quadrature: with nested Clenshaw-Curtis quadrature rule

1D difference grid: $\triangle_k^1 f := \left(\mathcal{Q}_k^1 - \mathcal{Q}_{k-1}^1\right) f$
Tensor product: $\left(\triangle_{k_1}^1 \otimes \cdots \otimes \triangle_{k_N}^1\right) f$
Linear combination: $\mathcal{Q}_L^N[f] := \sum \left(\triangle_{k_1}^1 \otimes \cdots \otimes \triangle_{k_N}^1\right) f$

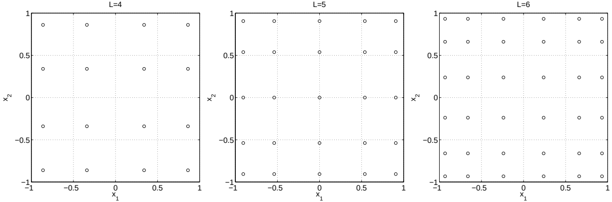


Sparse quadrature: comparison with full–tensor quadratures

Sparse Clenshaw-Curtis Chebyshev: P=7, P=9 & P=11



Full Gauss–Legendre Quadrature: P=7, P=9 & P=11



Canonical, maximum and anisotropic expansions

M is determined by the accuracy of the quadrature approach. If the quadrature has a polynomial accuracy of P or $\mathbf{P}$, there are the following expansions for

$$f_r(\mathbf{x}) = \sum_{\boldsymbol{\alpha}\in\mathbb{N}^N} f_{\boldsymbol{\alpha}} \phi_{\boldsymbol{\alpha}}(\mathbf{x})$$

- Canonical: *total* degrees not greater than P, *i.e.* $\{\phi_{\boldsymbol{\alpha}} \ / \ |\boldsymbol{\alpha}| \leq P\}$
- Maximum: degree in *each* n not greater than P, *i.e.* $\{\phi_{\boldsymbol{\alpha}} \ / \ \boldsymbol{\alpha} \leq P\}$.
- Anisotropic: degree in *each* n not greater than P_n, *i.e.* $\{\phi_{\boldsymbol{\alpha}} \ / \ \boldsymbol{\alpha} \leq \mathbf{P}\}$.

M	P	Legendre polynomial	Canonical, $P = 2$	Maximum, $P = 2$	Anisotropic $\mathbf{P} = [3, 1]$
0	0	1	$P^0(x_1)P^0(x_2)$	$P^0(x_1)P^0(x_2)$	$P^0(x_1)P^0(x_2)$
1	1	x_1	$P^1(x_1)P^0(x_2)$	$P^1(x_1)P^0(x_2)$	$P^1(x_1)P^0(x_2)$
2		x_2	$P^0(x_1)P^1(x_2)$	$P^0(x_1)P^1(x_2)$	$P^0(x_1)P^1(x_2)$
3	2	$3/2x_1^2 - 1/2$	$P^2(x_1)P^0(x_2)$	$P^2(x_1)P^0(x_2)$	$P^2(x_1)P^0(x_2)$
4		x_1x_2	$P^1(x_1)P^1(x_2)$	$P^1(x_1)P^1(x_2)$	$P^1(x_1)P^1(x_2)$
5		$3/2x_2^2 - 1/2$	$P^0(x_1)P^2(x_2)$	$P^0(x_1)P^2(x_2)$	$P^0(x_1)P^2(x_2)$
6	3	$5/2x_1^3 - 3/2x_1$	$P^3(x_1)P^0(x_2)$	$P^3(x_1)P^0(x_2)$	$P^3(x_1)P^0(x_2)$
7		$(3/2x_1^2 - 1/2)x_2$	$P^2(x_1)P^1(x_2)$	$P^2(x_1)P^1(x_2)$	$P^2(x_1)P^1(x_2)$
8		$x_1(3/2x_2^2 - 1/2)$	$P^1(x_1)P^2(x_2)$	$P^1(x_1)P^2(x_2)$	$P^1(x_1)P^2(x_2)$
9		$5/2x_2^3 - 3/2x_2$	$P^0(x_1)P^3(x_2)$	$P^0(x_1)P^3(x_2)$	$P^0(x_1)P^3(x_2)$
10	4	$35/8x_1^4 - 15/4x_1^2 + 3/8$	$P^4(x_1)P^0(x_2)$	$P^4(x_1)P^0(x_2)$	$P^4(x_1)P^0(x_2)$
11		$(5/2x_1^3 - 3/2x_1)x_2$	$P^3(x_1)P^1(x_2)$	$P^3(x_1)P^1(x_2)$	$P^3(x_1)P^1(x_2)$
12		$(3/2x_1^2 - 1/2)(3/2x_2^2 - 1/2)$	$P^2(x_1)P^2(x_2)$	$P^2(x_1)P^2(x_2)$	$P^2(x_1)P^2(x_2)$
13		$x_1(5/2x_2^3 - 3/2x_2)$	$P^1(x_1)P^3(x_2)$	$P^1(x_1)P^3(x_2)$	$P^1(x_1)P^3(x_2)$
14		$35/8x_2^4 - 15/4x_2^2 + 3/8$	$P^0(x_1)P^4(x_2)$	$P^0(x_1)P^4(x_2)$	$P^0(x_1)P^4(x_2)$

gPC as a Uncertainty Quantification (UQ) & Sensitivity Analysis (SA) tool

Statistical moments:

$$\begin{aligned}\mu_u(\mathbf{x}) &= \int_\Gamma u_r(\mathbf{x};\omega)\phi_0(\boldsymbol{\xi})\rho(\boldsymbol{\xi})d\boldsymbol{\xi} = u_0(\mathbf{x}),\\ \sigma^2_{u,gPC}(\mathbf{x}) &= \int_\Gamma \left[\sum_{m=0}^{M} u_m(\mathbf{x})\phi_m(\boldsymbol{\xi}) - u_0(\mathbf{x})\right]^2 \rho(\boldsymbol{\xi})d\boldsymbol{\xi} = \sum_{m=1}^{M} u_m^2(\mathbf{x})\langle\phi_m^2(\boldsymbol{\xi})\rangle.\end{aligned}$$

Solution sensitivity: Partial differentiation wrt ξ_n Agarwal (2008)

$$\mathbf{S}_{\xi_n}(\mathbf{x}) = \frac{\partial u_r(\mathbf{x};\boldsymbol{\xi})}{\partial \xi_n}.$$

Sensitivity analysis: partial variances Sobol' (1993)

$$\sigma_u^2(\mathbf{x}) = \sum_{i_1=1}^{N} D_{i_1}(\mathbf{x}) + \sum_{i_1=1}^{N}\sum_{i_2=1}^{i_1} D_{i_1 i_2}(\mathbf{x}) + \sum_{i_1=1}^{N}\sum_{i_2=1}^{i_1}\sum_{i_3=1}^{i_2} D_{i_1 i_2 i_3}(\mathbf{x}) \cdots + D_{i_1 i_2 \ldots i_N}(\mathbf{x}).$$

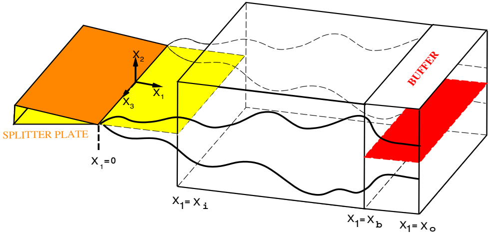
Probability density function (PDF): numerical computation from the histogram of a large MC sample of $u_r(\mathbf{x},\boldsymbol{\xi})$ based on the distribution of $\boldsymbol{\xi}$

Application of gPC to some examples

Examples	Tasks	N	R.V / Representations
Mixing layer magnitude	UQ & SA	2 & 3	Uniform/Legendre
Mixing layer phase	UQ & SA	1 & 2	Periodic/Fourier
Toy models	QE	1 to 10	Gauss.&Uni./Herm.&Leg.
Global circulation model	SA & CAL.	5	Log-uni.&Uni./Leg.

Sensitivity of spatially developing mixing layer

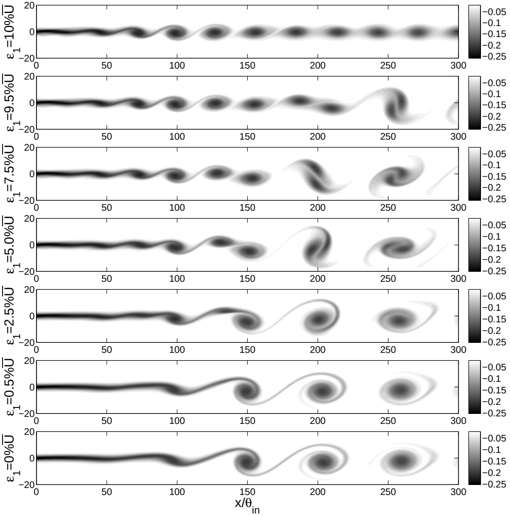
- Coherent vortical structures triggered by inflow forcing Brown & Roshko (1974)
- Shear layer at the inflow approximated as $\overline{U}_{in}(y) = 1 + \lambda \tanh(y/2)$
- Downstream shear layer growth is very sensitive to forcing definition
- Forcing with LST fundamental mode, *i.e.* most unstable, and its subharmonic modes: $u_p(y,t) = \sum \epsilon_n f_n(y) \exp(i(\omega_n t + \gamma_n))$
- 3D flow structure is largely 2D $\rightarrow$ 2D DNS Delville et al. (1999)
- Goal: To generalize the approach to design discrete forcing with random magnitude or phasing



De Brun (2001)

Sensitivity to forcing: magnitude ϵ_n

- Instantaneous vorticity contours with bimodal perturbation
- Vortical structure variation as the relative frequency content in inflow forcing changes



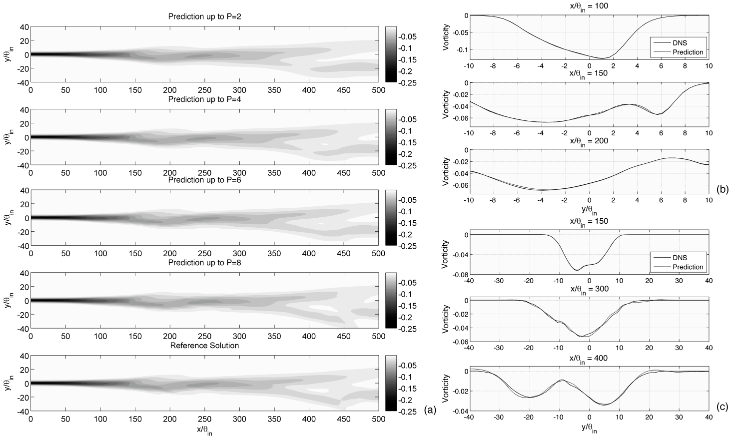
Stochastic mixing layer with random magnitudes, ϵ_n

Treat ϵ_n and γ_n as random variables to determine the most general way to control mixing layer growth with inflow forcing.

- Bimodal forcing and trimodal forcing examined
- Stochastic forcing magnitudes ϵ_n as uniform variables in $[0, 10\%\overline{U}]$
- Legendre-Chaos expansion of stochastic fields
- Mixing layer solutions with 2D spectral/hp DNS solver
- $Re = 100$, $\lambda = 0.5$
- Non–intrusive Probabilistic Collocation Method with full–tensor Gauss–quadrature
- 81 full–tensor quadrature points for bimodal forcing (N=2, L=9, P=8) & 1000 for trimodal (N=3, L=10, P=9)
- Examine time-averaged mixing layer thickness, *e.g.* momentum thickness θ

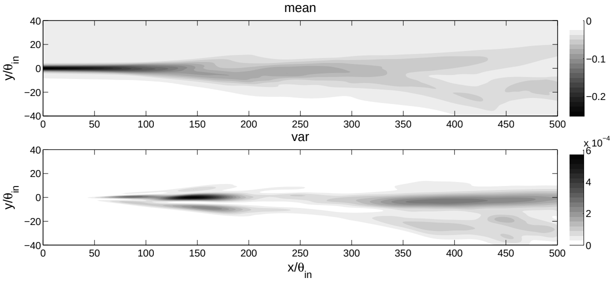
Accuracy of the gPC expansion: solution prediction

With $u(\mathbf{x}, \boldsymbol{\xi}) = u_m(\mathbf{x})\phi_m(\boldsymbol{\xi})$, we can predict the solution at an arbitrary point within $\boldsymbol{\Gamma}$. Accuracy of the prediction increases with increasing M or P.

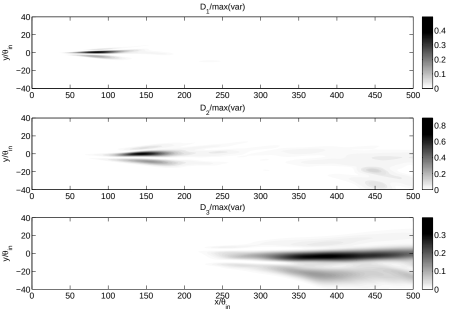


Response variability in trimodal perturbation case

- Initial response up to $x/\theta_{in} = 250$ similar to the bimodal case
- Large local variance at the location associated with the onset of deterministic subharmonic vortex merging

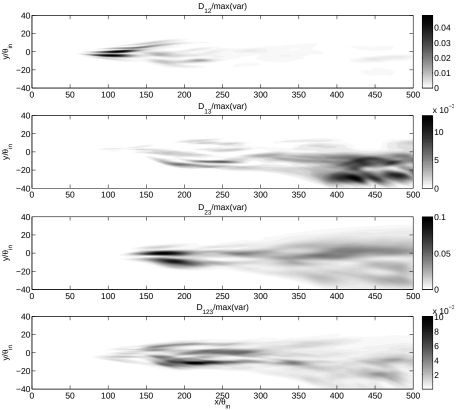


Partial variance contour in trimodal vorticity



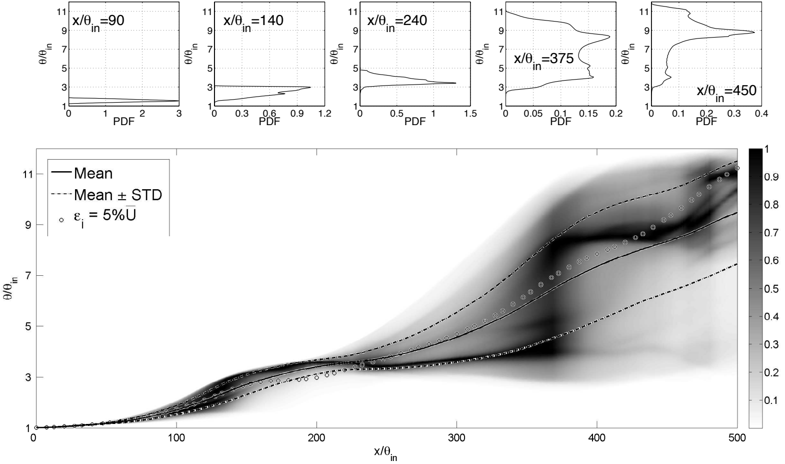
- D_n: sensitivities of the solution to ϵ_n
- Contours of each sensitivity index correspond closely to the deterministic vortex-roll up of each mode

Partial variance in trimodal vorticity contour



- D_{ij}: sensitivities of the solution to interaction between ϵ_i and ϵ_j
- Large D_{12} and D_{23} $\longrightarrow$ interactions between successive modes are dominant Kelly (1967)
- D_{123}: sensitivities of the solution to the mutual interaction amongst all modes

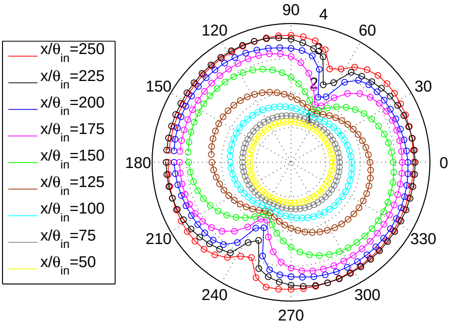
θ PDF in trimodal perturbation case



Stochastic mixing layer with random phase γ_n

- Bimodal forcing and trimodal forcing examined
- Stochastic phase shifts γ_n as uniform random variables in $[0, 2\pi)$
- Forcing magnitudes maintained at $\sum \epsilon_n = 10\%\overline{U}$
- SCM with Newton-Cotes quadrature
- Fourier-Chaos expansion of stochastic fields
- Discrete Fourier transformation (DFT) speeds up coefficient computations
- 72 equidistant quadrature samples are used (nested points)
- Examine time-averaged mixing layer thickness, *e.g.* momentum thickness θ

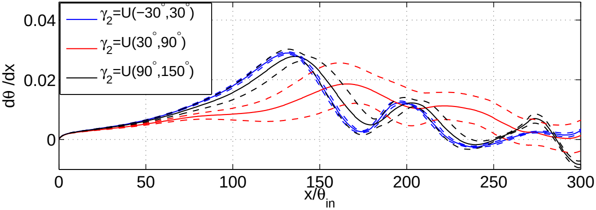
Response of momentum thickness



- Symmetry observed as $\gamma_2 \in [0, 2\pi]$ includes two periods of fund. forcing
- Mixing layer growth strongly delayed over small γ_2 range near 70° Inoue (1995)
- Delayed growth reported for $\gamma_2 = 0$ at merging locations Stanley & Sarkar (1997)
- 45° difference between inflow forcing formulations
- Phase shift at inflow does not correspond to phase shift at merging locations

Mixing layer growth rate statistics

- $\partial\theta/\partial x$ examined for:
  Normal growth: $\gamma_2 = U(-30^\circ, 30^\circ)$ & $\gamma_2 = U(90^\circ, 150^\circ)$
  Delayed growth: $\gamma_2 = U(30^\circ, 90^\circ)$



- 'Normal growth': Fast growth near inflow followed by sharp drop in $\partial\theta/\partial x$. Drop or contraction of the mixing layer Oster & Wygnansk (1982)
- 'Delayed growth': Slower growth with less $\partial\theta/\partial x$ fluctuation. Large variance due to solution sensitivity in $\gamma_2 \in [45^\circ, 80^\circ]$. Range of sensitivity is small Stanley & Sarkar (1997)

PC as a quantile estimation tool

Empirical quantile: estimated from $\hat{Y}_\alpha = \inf\{y; \hat{F}(y) \geq \alpha\}$ which gives

$$\hat{Y}_\alpha = Y_{(\lceil \alpha Z \rceil)}, \tag{1}$$

where $\{Y_{(i)}\}_{i=1}^{Z}$ are the ordered set of the Z MC samples.
The metamodel accurately determines the statistical moments but fails in extreme quantile estimations, *i.e.* α near 0 or 1.
We propose a multi–element refinement approach: global gFC metamodel is complimented by local metamodel constructed around design points $\boldsymbol{\xi}_\alpha$.
Design point: *most likely* random input that corresponds to $u_r(\mathbf{x}, \boldsymbol{\xi}) = u_\alpha(\mathbf{x})$.
This gives a constraint nonlinear minimization problem , *i.e.*

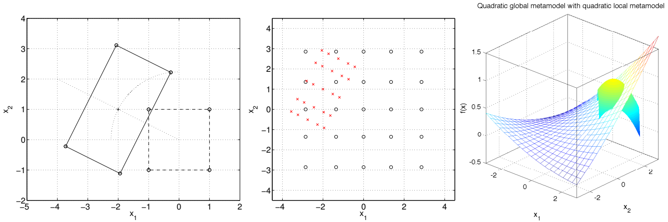
$$\min \|\boldsymbol{\xi}\|, \quad \text{s.t.} \quad \sum_{m=0}^{M} u_m(\mathbf{x})\phi_m(\boldsymbol{\xi}) - \hat{Y}_\alpha = 0.$$

The above problem is solved by the method of **Lagrangian multipliers**.

Multi–Element Monte Carlo simulation

Local gPC metamodels are created around the design points. The multielement solution is used as the metamodel, *i.e.*

$$D_{\mathrm{ME}} = \begin{cases} D_{\mathrm{global}} = D \setminus D_{\mathrm{local}}, & \text{domain of global gPC,} \\ D_{\mathrm{local}} = \cup D_{\beta_i}, & \text{domains of refinement about } \hat{\boldsymbol{\xi}}_{\alpha_i}, \text{ for } i = 1, \ldots, N_\beta. \end{cases}$$



The final multi–element gPC (MEgPC) metamodel is

$$f_{\mathrm{ME}}(\mathbf{x}) = \begin{cases} \sum_{m=0}^{M} f_m \phi_m(\mathbf{x}), & \text{if } \mathbf{x} \in D_{\mathrm{global}}, \\ \sum_{m=0}^{M_i^*} f_{m,i}^* \psi_{m,i}(\mathbf{T}_i^{-1}(\mathbf{x})), & \text{if } \mathbf{x} \in D_{\beta_i}. \end{cases}$$

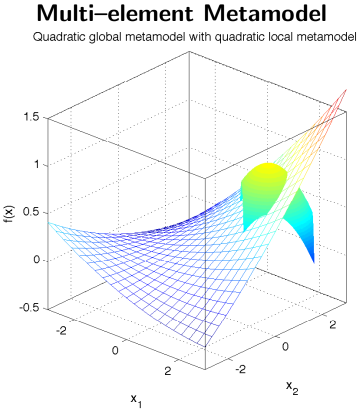
where $\mathbf{T}_i$ a transformation operator that maps a point in the uniform bounded support $\mathbf{x}^* \in [-1, 1]^N$ to the local domain $\mathbf{x} \in D_{\beta_i}$.

Example: Gaussian–like response

We examine the quantile of the output of a Gaussian-like function:

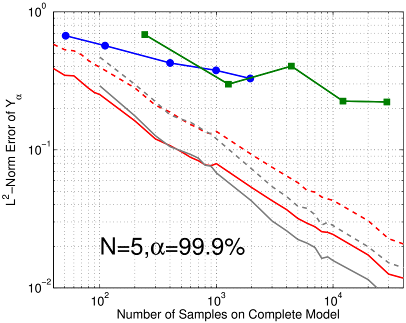
$$f(\mathbf{x}) = \sum_{i=1}^{N_\alpha} \prod_{n=1}^{N} \exp\left(\frac{-(x_n - \mu_{n,i})^2}{2\sigma_{n,i}^2} \right), \qquad (2)$$

where $\|\boldsymbol{\mu}\| = 2$, $\boldsymbol{\sigma} = 1$, $\mathbf{x}$ are i.i.d. random variables and $x_n \in \mathcal{N}(0,1)$.



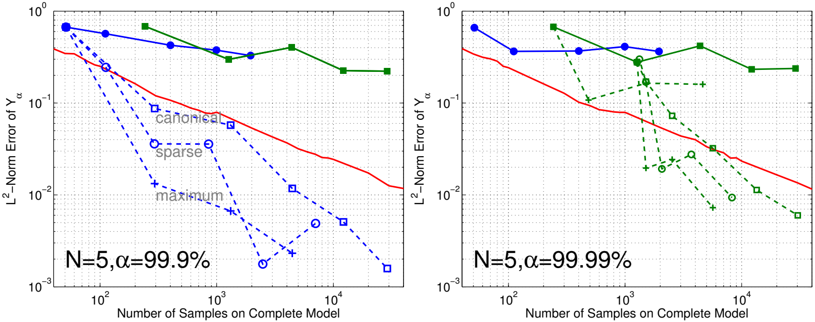
α–quantile estimator convergence for MC, IS and global gPC

- Monte Carlo $\hat{Y}_\alpha$ converges as $1/\sqrt{Z}$
- Importance sampling $\hat{Y}_\alpha$ computed at selected Z: Z/2 MC samples for first estimate of $\hat{Y}_\alpha$, at most $Z/4$ for GPM and the rest for IS
- Global **full** and **sparse** gPC estimations of $\hat{Y}_{\alpha,r}$ (from $L = 3$ to 7) are poor



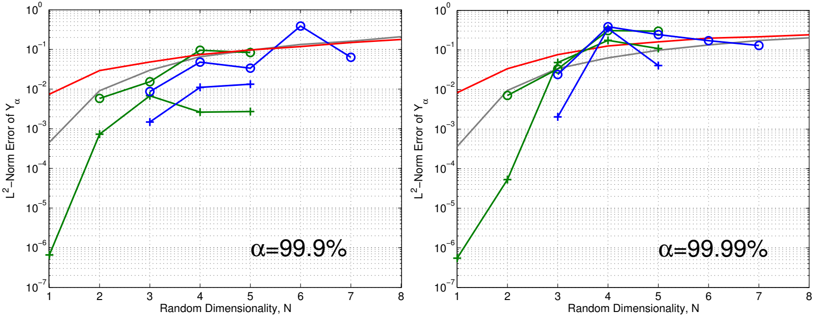
Effects of different local refinements

- Local **full** (canonical & maximum) and **sparse** gPC metamodel refinements
- Maximum expansion improves the accuracy of $\hat{Y}_{\alpha,\mathrm{ME}}$ given the same Z
- Seek best $\hat{\boldsymbol{\xi}}_{\alpha,r}$ estimation by maximizing Z in global gPC metamodel



Target cost study

- An arbitrary target cost that increases linearly with N: $Z_{total} = 100N$
- Monte Carlo and importance sampling $\hat{Y}_\alpha$ with entire sampling budget
- Global **full** and **sparse** + local full maximum (+) and sparse (∘) supplemental metamodels
- Maximize global metamodel cost while not exceeding the entire budget



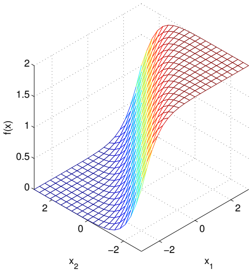
Example: Hypertangent response

We examine the quantile of the output of a hypertangent function:

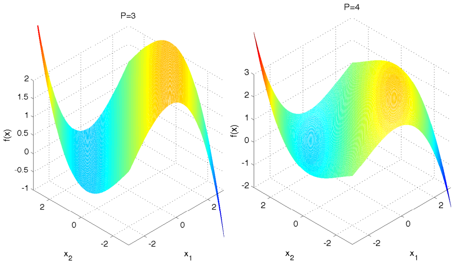
$$Y(\mathbf{x}) = 1 + \tanh\left(\sum_{n=1}^{N} \sigma_n (x_n - \mu_n)\right).$$

where the N–dimensional input are i.i.d. random variables $\mathbf{x} \in \mathcal{N}(0,1)$.

Double Peak Response

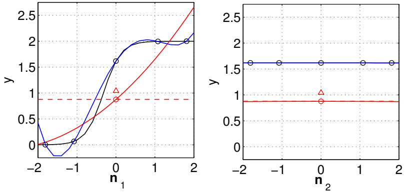


Global gPC Metamodel

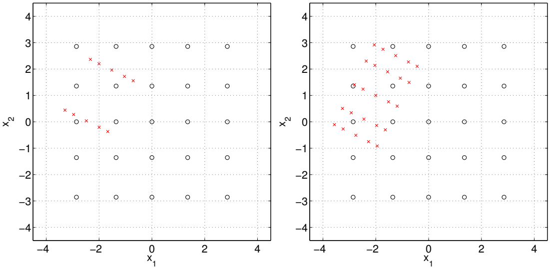


Anisotropic grid

- The dominance of some random variables can be revealed by examining the partial variance of the global gPC metamodel
- One–dimensional metamodels about $\hat{\boldsymbol{\xi}}_{\alpha,r}$ can identify dominant directions

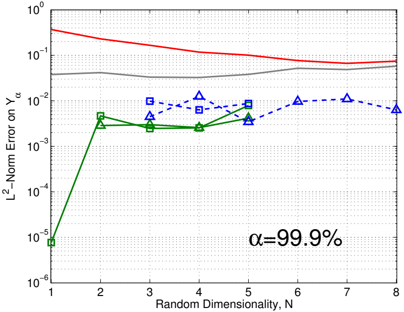


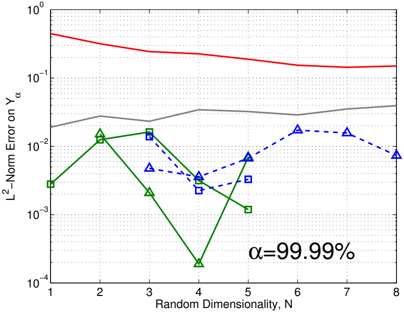
- Anisotropic grids, P in $\hat{\boldsymbol{\xi}}'_{\alpha,r}$ and linear in transverse directions, reduce cost



Target cost study

- An arbitrary target cost that increases linearly with N: $Z_{total} = 100N$
- Monte Carlo and importance sampling $\hat{Y}_\alpha$ with entire sampling budget
- Global **full** and **sparse** + local full canonical (□) and anisotropic (△) supplemental metamodels
- Maximize global metamodel cost while not exceeding the entire budget



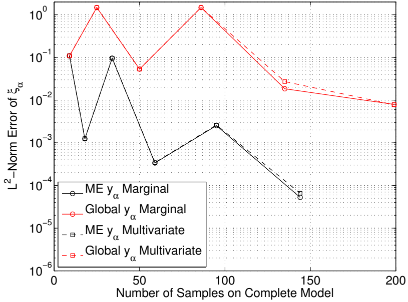


Quantile of multivariate output

We assume that all components of the random output $\mathbf{Y}$ are extreme and define the multivariate α–quantile as the point $\mathbf{y}_\alpha$ where the multivariate and marginal *cdf*'s satisfy the following conditions

$$F(\mathbf{y}_\alpha) = \alpha \quad \text{and} \quad F_1(y_{\alpha,1}) = F_2(y_{\alpha,2}) = \cdots = F_K(y_{\alpha,K}) \tag{3}$$

where K is the number of outputs. Results with N=2 and $\alpha = 99\%$ case for multiple Gaussian peaks:



Calibration and sensitivity analysis GCM

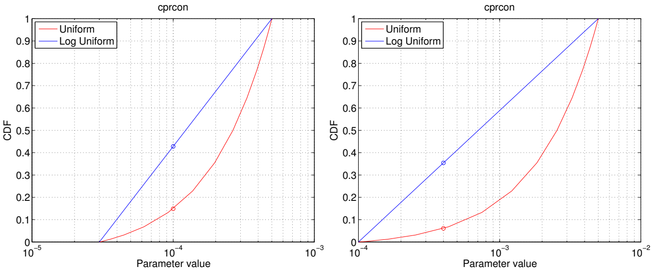
- Examine the AGCM ECHAM6 with uncertain parameters in cloud modeling
- 1977 climatological distributions of sea ice and surface temperature used as initial condition
- Five R.V. in the expert range transformed to the Gaussian space
- Ensemble of model output created for a *single* year run
- Full–tensor quadrature with squadratic accuracy, *i.e.* 243 points

Selection of the input random variables

Table: Expert parameter range and their default values

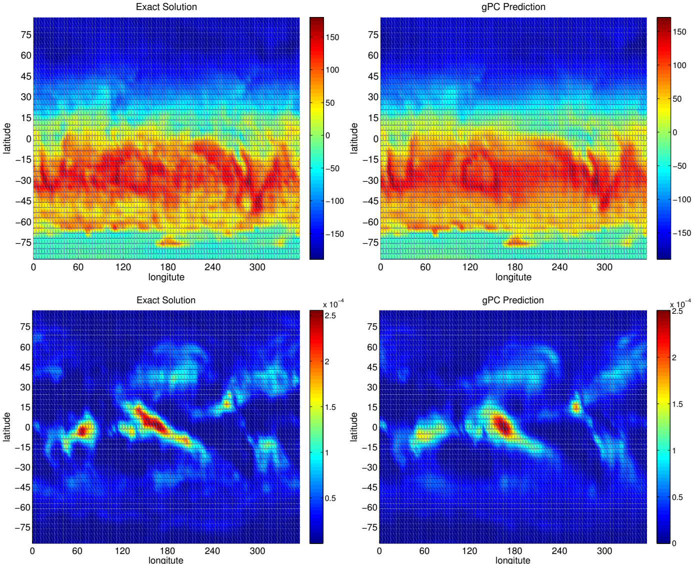
Parameter	Range	Default value
entrainment rate for shallow convection (entrscv)	0.0003-0.001	0.0003
entrainment rate for penetrative convection (entrpen)	0.00003-0.0005	0.0001
inhomogeneities of ice clouds (zinhomi)	0.65-1.0	0.7
inhomogeneities of liquid clouds (zinhoml)	0.65-1.0	0.7
conversion rate of cloud water to rain (cprcon)	0.0001-0.005	0.0004

- zinhomi & zinhoml are treated as uniform r.v.
- entrscv, entrpen & cprcon are treated as uniform r.v or log uniform r.v.
- A dependent parameter, $cmfctop = entrscv \times \frac{1000}{3}$, is included
- A uniform distribution under–weights the entire lower range



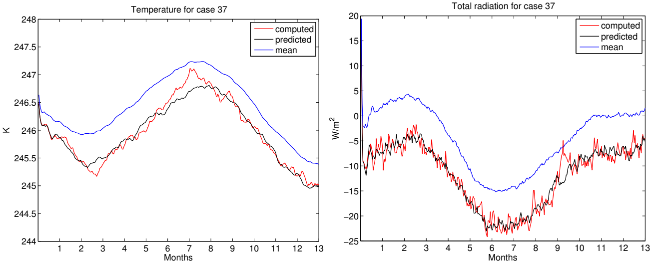
Validation: Comparison of computed global contours and gPC predictions

- Comparison at an arbitrary point within the support
- Exact solution vs gPC prediction for global radiation and precipitation
- For December 1970, large–scale patterns resolved in time-averaged results



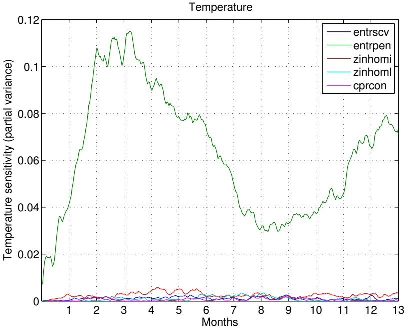
Validation: Comparison of computed global means and gPC predictions

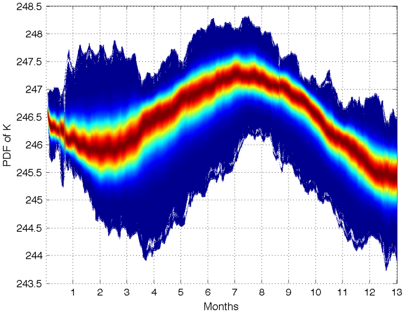
- Global mean should be consider to avoid small eccentric scales



Sensitivity analysis

- Partial variances reveal strong effects from 'entrpen'.
- Couple terms in the partial variance is much smaller
- Temperature PDF generated from the gPC metamodels with 10^5 Monte Carlo samples.





Code calibration

For optimization problem with K objective functions, we seek all the $\boldsymbol{\xi}$ that satisfy the following minimization problem, *e.g.*

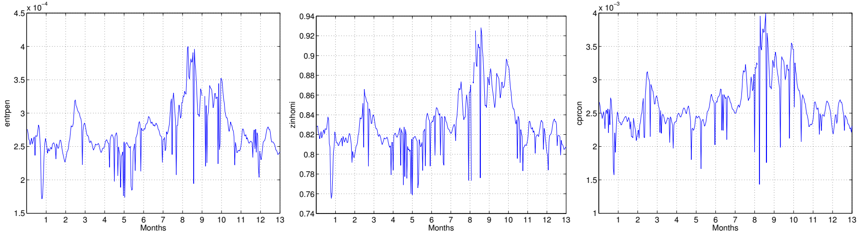
$$\boldsymbol{\xi}^* = \underset{\boldsymbol{\xi}}{\operatorname{argmin}} \sum_{k=1}^{K} \omega_k \left(\sum_{m=0}^{M} u_{m,k}(t)\phi_m(\boldsymbol{\xi}) - u_{\mathrm{obs},k}(t) \right)^2 \quad \text{for t} = 1,\ldots,364$$

The choice of weight vector $\boldsymbol{\omega}$ is arbitrary. Many optimization algorithms exist. So far K=1

- Lagrange multiplier algorithm used to solve the constraint nonlinear minimization problem for global averaged temperature
- u_{obs} are the daily global averaged temperature in 1970 from ECMWF
- the following figures show the daily 'optimal' value for each parameter
- with additional objective functions, there is likely to be non-dominant sets, *i.e.* one cannot make one objective better without worsening the other objectives Neelin (2001)

Calibration results

Parameter	Range	Default value
entrainment rate for shallow convection (entrscv)	0.0003-0.001	0.0003
entrainment rate for penetrative convection (entrpen)	0.00003-0.0005	0.0001
inhomogeneities of ice clouds (zinhomi)	0.65-1.0	0.7
inhomogeneities of liquid clouds (zinhoml)	0.65-1.0	0.7
conversion rate of cloud water to rain (cprcon)	0.0001-0.005	0.0004



Some concluding remarks

- PC and gPC constructs metamodels that accurately mimics the behaviours of complete simulators about the mean of the stochastic inputs
- Initial used as a UQ ans SA tool in engineering problems
- It has potential as a multi–objective optimization tool
- There is no free lunch – it suffers from the "curse of dimensionality"
- Adaptive techniques (multi–element, anisotropic quadrature) can reduce cost
- To investigate anisotropic spare quadrature & sparse gPC representation
- Reduce input dimension via non–dimensional analysis or identification of dominant inputs
- Orphan points (difference between sample budget and quadrature cost) - can we use them in a sequential design – with Hugo?
- Including data assimilation and Bayesian analysis in gPC/PC framework
- Practical issues: need better random input measurement