Uncertainty Quantification and Optimization Under Uncertainty Using Surrogate Models

Master's Thesis Defense

Komahan Boopathy

University of Dayton
Department of Mechanical and Aerospace Engineering

March 25, 2014

Outline

Background I

Analysis & Optimization:

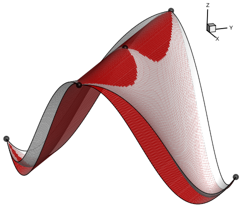
- Many design iterations – can be very expensive
- Highly coupled with several disciplines
- Time consuming to do physical testing and infeasibility

Advances in Computation:

- Hardware (processor speed, multi-core systems)
- Software (parallel programming)
- Algorithms and other tools (sophisticated methods)

Surrogate/ Meta models/ Response surfaces

- Approximation of the exact function using interpolation and/or extrapolation
- Some Applications:
  - Optimization
  - Database creation
  - Uncertainty quantification



Background II

Choice of Training Points:

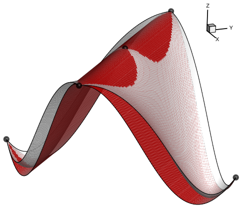
- Accuracy depends on choice of training points
- Optimal training is difficult (no defined criteria)
- Spacing and other heuristics

Surrogate Approximation Error:

- Need to know the model's accuracy
- Warrants exact function evaluations

Curse of Dimensionality

- Dramatic rise in number of training points with the number of input variables
- Good Tendencies:
  - Higher-order derivative information (Gradients, Hessian)
  - Variable-fidelity modeling
  - Piecewise approximation (polynomial surrogates)



Research Objectives I

1. to develop a training point selection framework for surrogate models
   1. absence of derivative information (function values only)
   2. presence of derivative information (function, gradient and Hessian values)
2. to propose a surrogate model error estimate
3. to show the framework's applicability on different surrogate models (kriging and polynomial chaos),
4. to advance gradient-enhanced polynomial chaos to Hessian-enhanced polynomial chaos methods,
5. to compare kriging and polynomial chaos surrogate models
6. apply to uncertainty quantification and optimization under uncertainty (mixed epistemic/aleatory)

Kriging

- The basic formulation of Kriging is given as,

$$\tilde{f} = f(x)^T \beta + Z(x)$$

  - $f(x)^T \rightarrow$ models the mean behavior
  - $Z(x) \rightarrow$ models the local variation from the mean behavior using a Gaussian process
- Predicts the function by stochastic processes
- Uses spatial correlation between data

Polynomial Chaos I

- Spectral expansion of orthogonal polynomials
- Intrusive/Non-intrusive forms
- Response surface:

$$\widehat{f}(\boldsymbol{x}) = \sum_{k=0}^{P} u_k \psi_k(\boldsymbol{x}), \tag{1}$$

- $\widehat{f}(\boldsymbol{\xi}) \rightarrow$ approximated function value
- $\boldsymbol{u} \rightarrow$ Expansion coefficients
- $\psi(\boldsymbol{\xi}) \rightarrow$ Orthogonal basis function

Polynomial Chaos II

Linear system:

$$\begin{bmatrix} \psi_0(\boldsymbol{x}^{(1)}) & \psi_1(\boldsymbol{x}^{(1)}) & \cdots & \psi_P(\boldsymbol{x}^{(1)}) \\ \psi_0(\boldsymbol{x}^{(2)}) & \psi_1(\boldsymbol{x}^{(2)}) & \cdots & \psi_P(\boldsymbol{x}^{(2)}) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_0(\boldsymbol{x}^{(N)}) & \psi_1(\boldsymbol{x}^{(N)}) & \cdots & \psi_P(\boldsymbol{x}^{(N)}) \end{bmatrix} \begin{Bmatrix} u_0 \\ u_1 \\ \vdots \\ u_P \end{Bmatrix} = \begin{Bmatrix} f(\boldsymbol{x}^{(1)}) \\ f(\boldsymbol{x}^{(2)}) \\ \vdots \\ f(\boldsymbol{x}^{(N)}) \end{Bmatrix}$$

- Data fitting at N points to find T coefficients
- Size: $N \times T$, where $T = P + 1$
- N=T $\rightarrow$ Interpolation, $N > T \rightarrow$ Regression
- Oversampling factor of 2

Polynomial Chaos III

With Gradients:

$$\left[\begin{array}{c}\begin{pmatrix} \psi_0(\mathbf{x}^{(1)}) & \psi_1(\mathbf{x}^{(1)}) & \cdots & \psi_P(\mathbf{x}^{(1)}) \\ \frac{\partial \psi_0(\mathbf{x}^{(1)})}{\partial x_1} & \frac{\partial \psi_1(\mathbf{x}^{(1)})}{\partial x_1} & \cdots & \frac{\partial \psi_P(\mathbf{x}^{(1)})}{\partial x_1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \psi_0(\mathbf{x}^{(1)})}{\partial x_M} & \frac{\partial \psi_1(\mathbf{x}^{(1)})}{\partial x_M} & \cdots & \frac{\partial \psi_P(\mathbf{x}^{(1)})}{\partial x_M} \end{pmatrix} \\ \vdots \end{array}\right] \left\{\begin{array}{c} u_0 \\ u_1 \\ \vdots \\ u_P \end{array}\right\} = \left\{\begin{array}{c}\begin{pmatrix} f(\mathbf{x}^{(1)}) \\ \frac{\partial f(\mathbf{x}^{(1)})}{\partial x_1} \\ \vdots \\ \frac{\partial f(\mathbf{x}^{(1)})}{\partial x_M} \end{pmatrix} \\ \vdots \end{array}\right\}$$

- Size $N' \times T$, where $N' = N \cdot (1 + M)$.

  $M \rightarrow$ Number of dimensions/variables
- Generally over-determined (least-squares)

Polynomial Chaos IV

With Hessian:

$$\left[\begin{pmatrix} \psi_0(\mathbf{x}^{(1)}) & \psi_1(\mathbf{x}^{(1)}) & \cdots & \psi_P(\mathbf{x}^{(1)}) \\ \frac{\partial\psi_0(\mathbf{x}^{(1)})}{\partial x_1} & \frac{\partial\psi_1(\mathbf{x}^{(1)})}{\partial x_1} & \cdots & \frac{\partial\psi_P(\mathbf{x}^{(1)})}{\partial x_1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial\psi_0(\mathbf{x}^{(1)})}{\partial x_M} & \frac{\partial\psi_1(\mathbf{x}^{(1)})}{\partial x_M} & \cdots & \frac{\partial\psi_P(\mathbf{x}^{(1)})}{\partial x_M} \\ \frac{\partial^2\psi_0(\mathbf{x}^{(1)})}{\partial^2 x_1} & \frac{\partial^2\psi_1(\mathbf{x}^{(1)})}{\partial^2 x_1} & \cdots & \frac{\partial^2\psi_P(\mathbf{x}^{(1)})}{\partial^2 x_1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2\psi_0(\mathbf{x}^{(1)})}{\partial x_1\partial x_M} & \frac{\partial^2\psi_1(\mathbf{x}^{(1)})}{\partial x_1\partial x_M} & \cdots & \frac{\partial^2\psi_P(\mathbf{x}^{(1)})}{\partial x_1\partial x_M} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2\psi_0(\mathbf{x}^{(1)})}{\partial^2 x_M} & \frac{\partial^2\psi_1(\mathbf{x}^{(1)})}{\partial^2 x_M} & \cdots & \frac{\partial^2\psi_P(\mathbf{x}^{(1)})}{\partial^2 x_M} \end{pmatrix} \\ \vdots \right] \begin{Bmatrix} u_0 \\ u_1 \\ \vdots \\ u_P \end{Bmatrix} = \begin{Bmatrix} \begin{pmatrix} f(\mathbf{x}^{(1)}) \\ \frac{\partial f(\mathbf{x}^{(1)})}{\partial x_1} \\ \vdots \\ \frac{\partial f(\mathbf{x}^{(1)})}{\partial x_M} \\ \frac{\partial^2 f(\mathbf{x}^{(1)})}{\partial^2 x_1} \\ \vdots \\ \frac{\partial^2 f(\mathbf{x}^{(1)})}{\partial x_1\partial x_M} \\ \vdots \\ \frac{\partial^2 f(\mathbf{x}^{(1)})}{\partial^2 x_M} \end{pmatrix} \\ \vdots \end{Bmatrix}$$

Size: $N' \times T$, where $N' = N \cdot (1 + M + \frac{M(M+1)}{2})$

Multivariate Interpolation and Regression

- Based on Taylor series expansion
- Mathematically,

$$\tilde{f} = \sum_{i=1}^{N_v} a_{vi}(x) f(x_{vi}) + \sum_{i=1}^{N_g} a_{gi}(x) \nabla f(x_{gi})$$

  - N_v, N_g is the number of function and func-grad data points
  - a_{vi} and a_{gi} are the basis functions
  - f and ∇f are the function f and gradient values

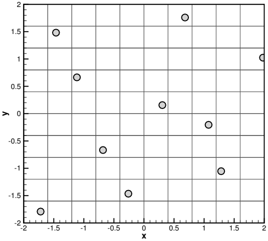
Training Point Selection

Domain based training

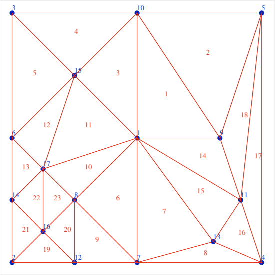
- Monte-Carlo
- Latin Hypercube
- Delaunay Triangulation
- Quadrature nodes
- Quasi-random sequences (Sobol, Halton)

Response based training

- Function values
- Kriging MSE and Expected Improvement
- Trust region



LHS



Delaunay Triangulation

Surrogate Validation

Needs Additional Exact Function Evaluations

- Root Mean Square Error (RMSE)
- Maximum Absolute Error (MAE)
- Relative Maximum Absolute Error (RMAE)
- Relative Average Absolute Error (RAAE)
- Split Sampling
- ...

No Additional Exact Function Evaluations

- Boot Strapping
- Cross Validation
- Inbuilt estimates (Kriging/GPR $\rightarrow$ MSE, MIR$\rightarrow$ sigma)

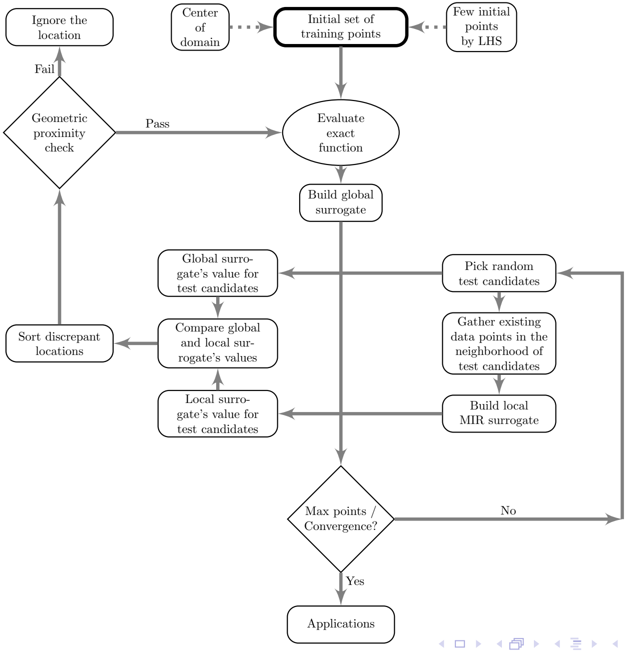
RMSE or L_2-norm

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(f^{(i)} - \widehat{f}^{(i)})^2},$$

MAE or L_∞-norm

$$\mathrm{MAE} = \mathbf{max}\{|f^{(i)} - \widehat{f}^{(i)}|\} \qquad i = 1, \ldots, n$$

Proposed Framework for Training and Validation



Proposed Error Estimates

Root Mean Square Discrepancy

$$\mathrm{RMSD} = \sqrt{\frac{1}{N_{test}} \sum_{j=1}^{N_{test}} (\widehat{f}^{(j)}_{global} - \widehat{f}^{(j)}_{local})^2} = \sqrt{\frac{1}{N_{test}} \sum_{j=1}^{N_{test}} (\delta^{(j)})^2},$$

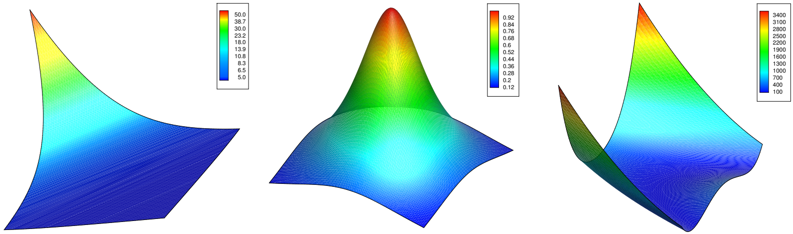
Approximate the actual root mean square error (RMSE or L_2-norm)

Maximum Absolute Discrepancy

$$\mathrm{MAD} = \mathbf{max}\{|\widehat{f}^{(j)}_{global} - \widehat{f}^{(j)}_{local}|\} \qquad j = 1, \ldots, N_{test}$$

Emulate the actual maximum absolute error (MAE or L_∞-norm)

Analytical Test Functions

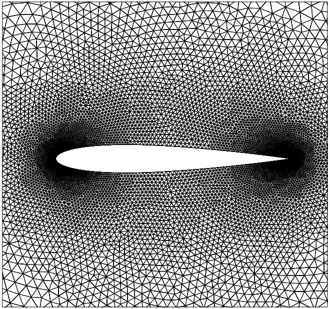


(a) Exponential (b) Runge (c) Rosenbrock

Contour plots of analytical test functions in two dimensions where the contours are colored by function values.

1. $f_1(x_1, \ldots, x_M) = \mathrm{e}^{(x_1 + \ldots + x_M)}$
2. $f_2(x_1, \ldots, x_M) = \frac{1}{1 + x_1^2 + \ldots + x_M^2}$
3. $f_3(x_1, \ldots, x_M) = \sum_{i=1}^{M-1} \left[(1 - x_i)^2 + 100(x_{i+1} - x_i^2)^2\right]$

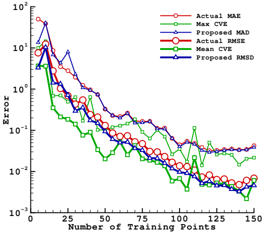
Aerodynamic Test Case



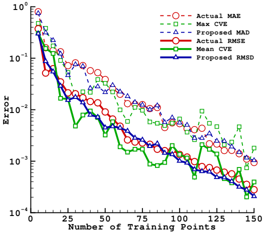
Problem Setup

- NACA0012 airfoil
- Eulerian flow solver
- Cell-centered second-order accurate finite-volume approach
- $0.5 < M < 1.5$ and $0^\circ < \alpha < 5^\circ$
- Mesh $19,548$ elements

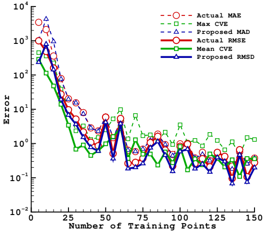
Error Estimate I



Exponential



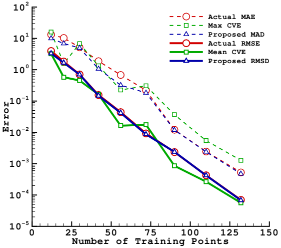
Runge



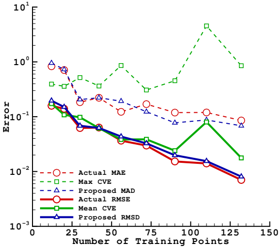
Rosenbrock

Figure : Kriging

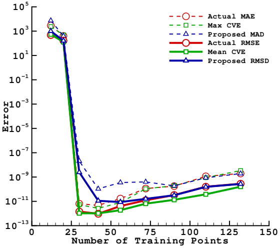
Error Estimate II



Exponential



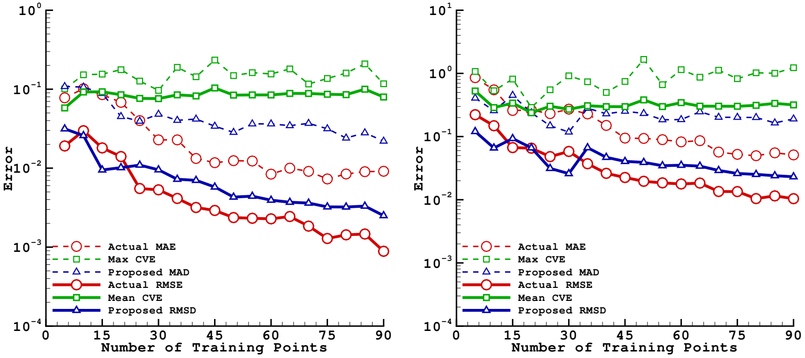
Runge



Rosenbrock

Figure : Polynomial Chaos

Error Estimate III



Drag Coefficient

Lift Coefficient

Figure : Kriging

Error Estimate IV

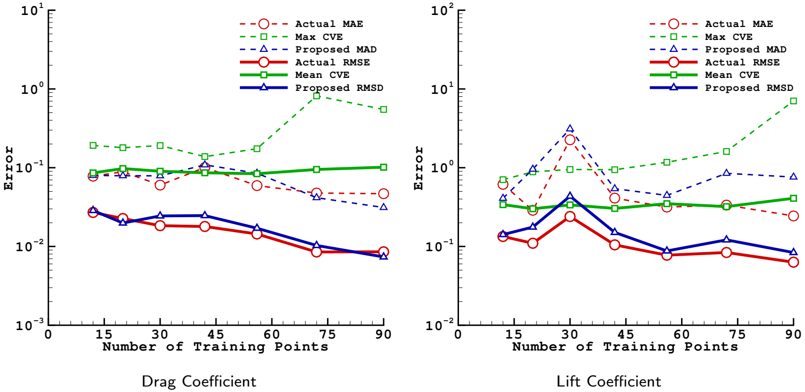
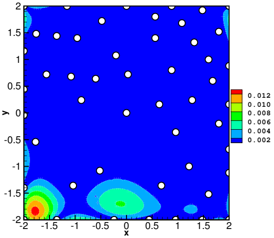
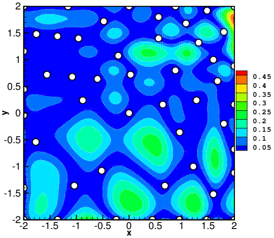


Figure : Polynomial Chaos

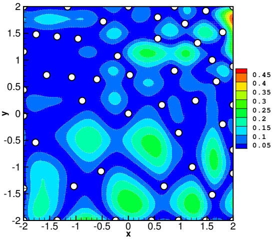
Error Estimate V



Actual error distribution (local)



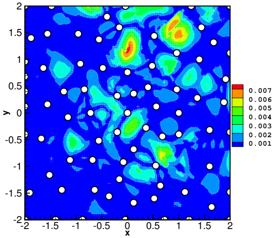
Actual error distribution (global)



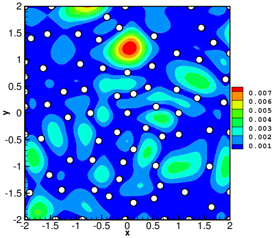
Proposed error distribution

Figure : Exponential test function.

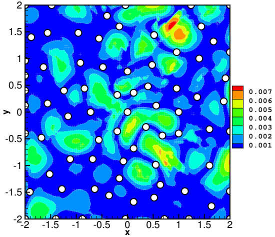
Error Estimate VI



Actual error distribution (local)



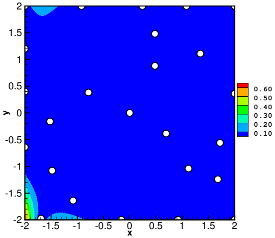
Actual error distribution (global)



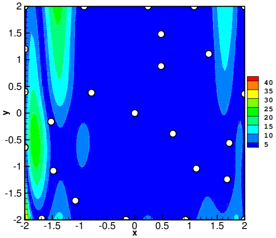
Proposed error distribution

Figure : Runge test function.

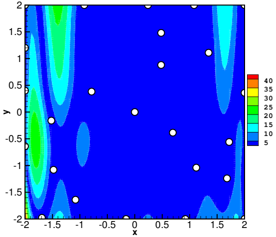
Error Estimate VII



Actual error distribution (local)



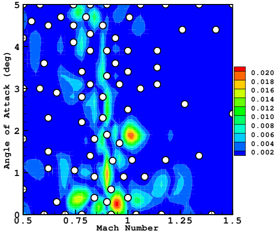
Actual error distribution (global)



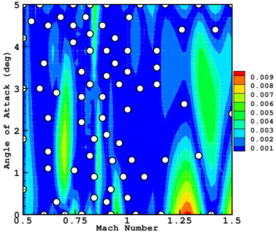
Proposed error distribution

Figure : Rosenbrock test function.

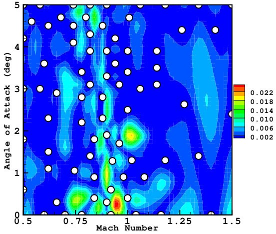
Error Estimate VIII



Actual error distribution (local)



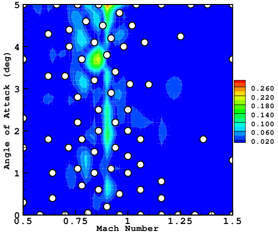
Actual error distribution (global)



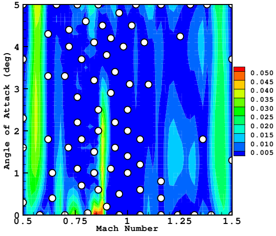
Proposed error distribution

Figure : Drag Coefficient

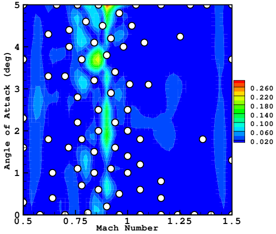
Error Estimate IX



Actual error distribution (local)



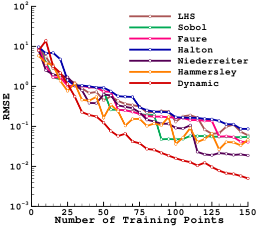
Actual error distribution (global)



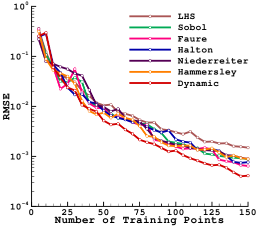
Proposed error distribution

Figure : Lift Coefficient

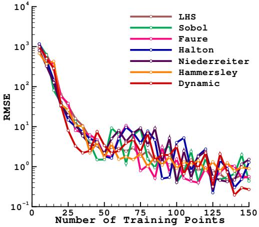
Quasi-random Sequences I



Exponential



Runge



Rosenbrock

Figure : Dynamic method versus quasi-random sequences using kriging

Comparing with LHS using PCE I

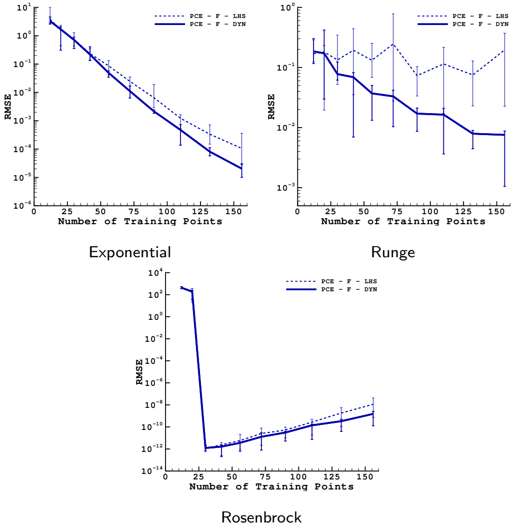
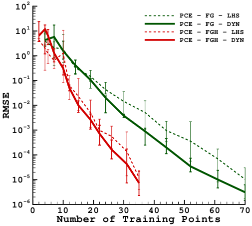
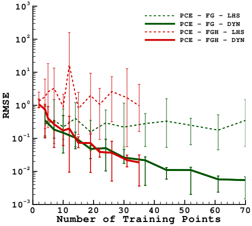


Figure : Dynamic method versus LHS using PCE in 2D (F only).

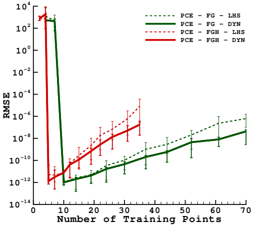
Comparing with LHS using PCE II



Exponential



Runge



Rosenbrock

Figure : Dynamic method versus LHS using PCE in 2D (FG and FGH).

Comparing with LHS using PCE III

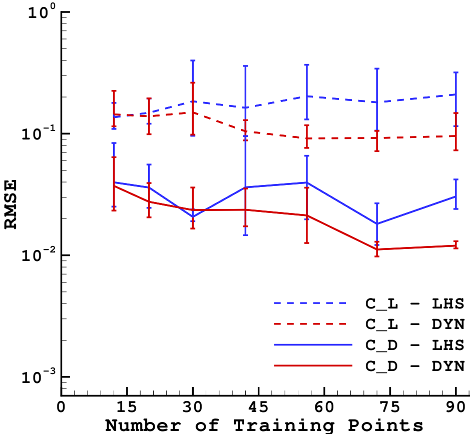
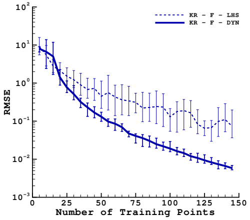
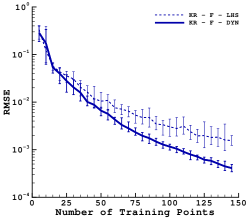


Figure : Drag and lift coefficients using kriging.

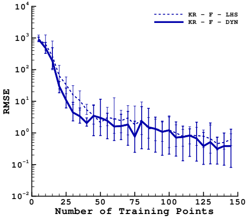
Comparing with LHS using Kriging I



(a) Exponential



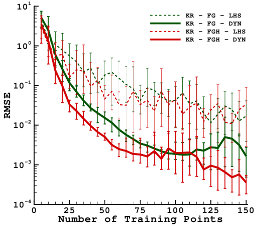
(b) Runge



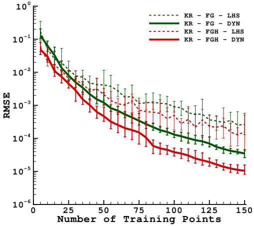
(c) Rosenbrock

Figure : Dynamic method versus LHS using kriging in 2D (F only).

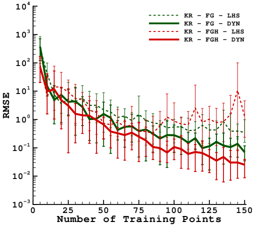
Comparing with LHS using Kriging II



Exponential



Runge



Rosenbrock

Figure : Dynamic method versus LHS using kriging in 2D (FG and FGH).

Comparing with LHS using Kriging III

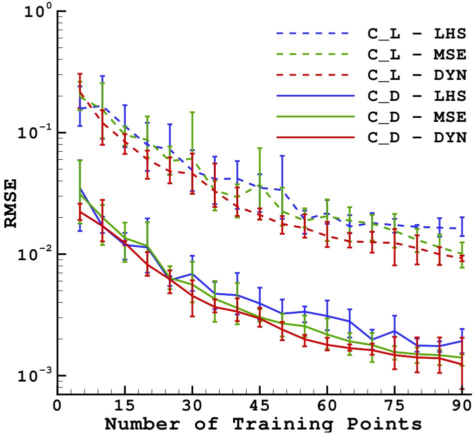


Figure : Drag and lift coefficients using kriging.

Comparing with LHS using Kriging IV

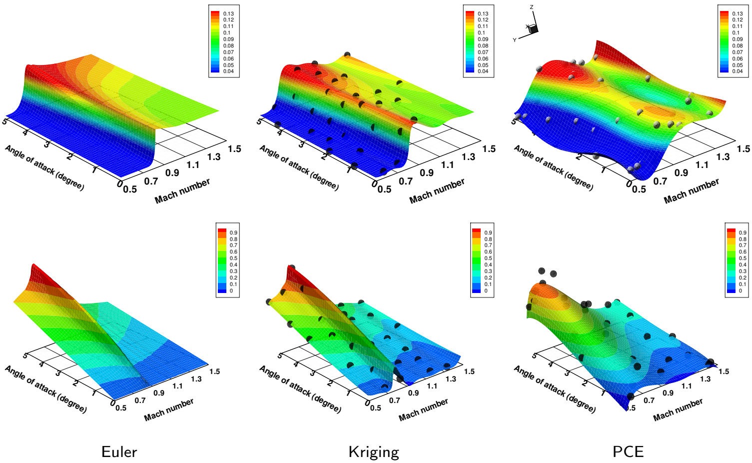


Figure : Contours of exact database (left), kriging (middle) and PCE (right) for drag (top) and lift coefficients (bottom) with 30 training points chosen with dynamic training point selection.

Variable Fidelity Kriging I

Variable Fidelity Kriging

- Even reduced simulation requirements by surrogate models can be expensive
- Idea is to combine trends from low-fidelity data (e.g., coarser meshes, less sophisticated models) with interpolations of high-fidelity data (e.g., finer meshes, better models, experimental data)
- Low-fidelity data from Euler evaluations with high-fidelity data from Navier-Stokes evaluations.
- Fine mesh $19,548$ elements Coarse mesh $4,433$ elements

1. Han, Z. H., Goertz, S., and Zimmermann, R., "Improving variable-fidelity surrogate modeling via gradient-enhanced kriging and a generalized hybrid bridge function," Aerospace Science and Technology, 2012.
2. Yamazaki, W., "Uncertainty Quantification via Variable Fidelity Kriging Model," Japan Society of Aeronautical Space Sciences, Vol. 60, 2012, pp. 80–88.

Variable Fidelity Kriging II

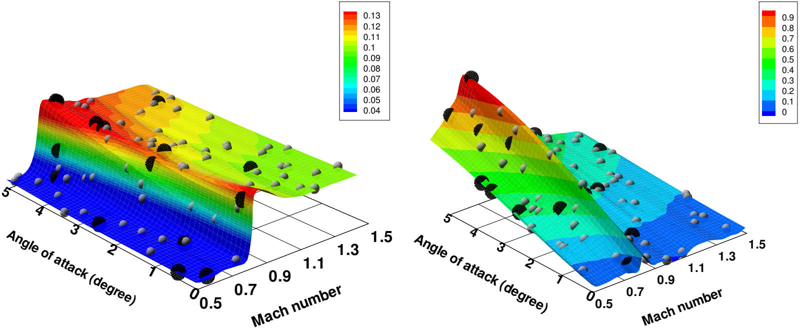


Figure : Kriging contour plots demonstrating the use of variable-fidelity data for drag (left) and lift (right) coefficients.

Table : RMSE comparisons for different kriging models.

RMSE	High-fidelity (30 high-fidelity points)	Variable-fidelity (15 high-fidelity and 60 low-fidelity points)
Drag Coefficient C_D	0.39×10^{-2}	0.31×10^{-2}
Lift Coefficient C_L	0.35×10^{-1}	0.18×10^{-1}

Why Uncertainty Quantification? I

- Design variables and input parameters are always subject to variations
  - Uncertain operating conditions (weather, ice accumulation on wing)
  - Uncertainties in boundary conditions/problem parameters
  - Uncertainties from lack of knowledge about a quantity (manufacturing tolerances)
  - Modeling inaccuracies (Navier-Stokes/Euler)
  - Random elements in a simulation
- Allowances must be made to accommodate likely variations/uncertainties

Why Uncertainty Quantification? II

- Traditionally we use **factor of safety** based on heuristics/expert opinion

A Typical Stress Constraint

$$g(\boldsymbol{d}) = \frac{\sigma}{\sigma_{max}} - 1 \leq 0 \Longrightarrow g(\boldsymbol{d}) = F_s \cdot \frac{\sigma}{\sigma_{max}} - 1 \leq 0$$

- What is an adequate or good factor of safety?
- Assumed **Factor of Safety** can be:
  - Adequate as well as over-conservative
  - Inadequate and prone to failure
- Increasingly difficult to come up with a factor of safety for radically new designs

Why Uncertainty Quantification? III

Why Quantify Uncertainties?

- Determine the real effects of uncertainties on the design (robust or vulnerable)
- Obtain confidence intervals for results (range of possible outcomes)
  - 95% probability (confidence) that the target C_L is achieved
  - 1% probability of violation of constraint #10
- Identify the limitations of the design (and improve)
- Reliability analysis for certification and quality assurance purposes

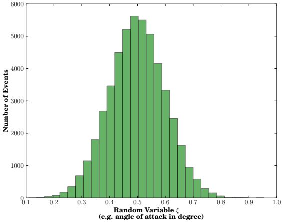
Uncertainty Types

- Aleatory / Irreducible / Type A
- Epistemic / Reducible / Type B
- Mixed

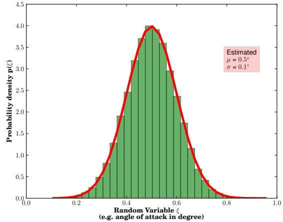
Aleatory Uncertainties I

Characteristics

- Inherent randomness or variations:
  - input parameters (Youngs modulus, shear force)
  - design variables
  - operating environment (cruise settings, temperature)
- Input probability distributions are known (sometimes assumed)
- Goal is to determine the output distribution



Available data

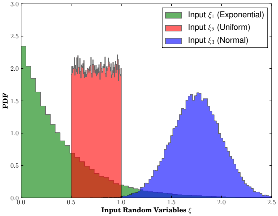


Fitted/Assumed distribution

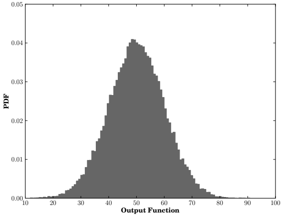
Aleatory Uncertainties II

Quantifying Aleatory Uncertainties

- Input data is available (mean, standard dev., distribution type)
- Need to know the input–output relationship of uncertainties
- Use Monte Carlo Sampling (MCS)
- Need thousands of simulations
- Use **surrogate models** to approximate the simulation output (kriging, polynomial chaos)



Input distributions



Output distribution

Aleatory Uncertainties III

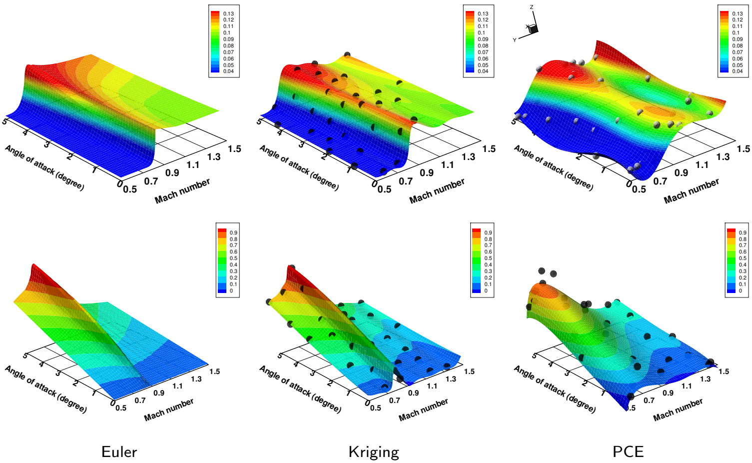
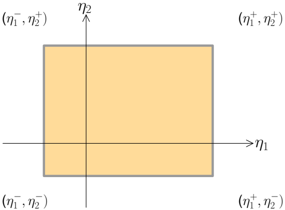


Figure : Contours of exact database (left), kriging (middle) and PCE (right) for drag (top) and lift coefficients (bottom) with 30 training points chosen with dynamic training point selection.

Epistemic Uncertainties I

Characteristics

- Lack of knowledge about the appropriate value
- Only bounds can be specified $I(\boldsymbol{\eta}) = [\boldsymbol{\eta}^-, \boldsymbol{\eta}^+] = [\bar{\boldsymbol{\eta}} - \boldsymbol{\tau}, \bar{\boldsymbol{\eta}} + \boldsymbol{\tau}]$
- Goal: determine the worst and best scenarios within the interval $I(\boldsymbol{\eta})$



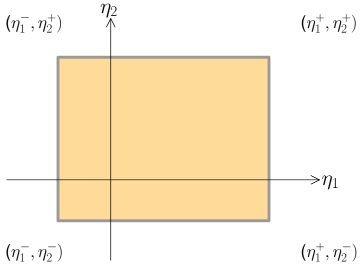
Bounds on epistemic variables

Epistemic Uncertainties II

Goal: determine the worst and best scenarios within the bounds

1. Extensive Sampling

- Need $10^3 - 10^6$ simulations
- Prohibitively expensive for bigger problems



Bounds on epistemic variables

Epistemic Uncertainties III

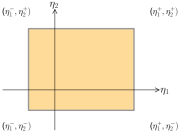
Goal: determine the worst and best scenarios within the bounds

2. Bound Constrained Optimization

- Optimization problem:

$$\begin{aligned} \underset{\boldsymbol{\beta}}{\text{minimize/maximize}} \quad & f = f(\boldsymbol{\eta}), \\ \text{subject to} \quad & \boldsymbol{\beta} \in I(\boldsymbol{\eta}) = [\bar{\boldsymbol{\eta}} - \boldsymbol{\tau}, \bar{\boldsymbol{\eta}} + \boldsymbol{\tau}]. \end{aligned}$$

- L-BFGS optimizer (needs gradients)
- Attractive even for bigger problems (scales linearly)



Bounds on epistemic variables

Mixed Aleatory & Epistemic Uncertainties

Quantifying Mixed Uncertainties

- Comprise of both aleatory ξ and epistemic uncertainties η
  - Naive approach: Nested Sampling
    - Very expensive (millions of function evaluations)
    - Not computationally affordable
  - Our approach: IMCS+BCO
    - Surrogate models for aleatory uncertainties
    - Bound constrained optimization for epistemic uncertainties
    - Few hundred (or thousand) function evaluations (manageable)

Optimization Problem Formulation I

Deterministic Optimization

$$
\begin{aligned}
\underset{\boldsymbol{d}}{\text{minimize}} \quad & J = J(f, \boldsymbol{q}, \boldsymbol{d}), \\
\text{subject to} \quad & R(\boldsymbol{q}, \boldsymbol{d}) = 0, \\
& g(f, \boldsymbol{q}, \boldsymbol{d}) \leq 0.
\end{aligned}
$$

Optimization Under Uncertainty

$$
\begin{aligned}
\underset{\boldsymbol{\xi}, \boldsymbol{\eta}}{\text{minimize}} \quad & \mathcal{J} = \mathcal{J}(\mu_{f*}, \sigma^2_{f*}, \boldsymbol{q}, \boldsymbol{\xi}, \boldsymbol{\eta}), \\
\text{subject to} \quad & R(\boldsymbol{q}, \boldsymbol{\xi}, \boldsymbol{\eta}) = 0, \\
& g^r = g(\mu_{f*}, \boldsymbol{q}, \boldsymbol{\xi}, \boldsymbol{\eta}) + k\sigma_{f*} \leq 0.
\end{aligned}
$$

Optimization Problem Formulation II

Lift constrained drag minimization

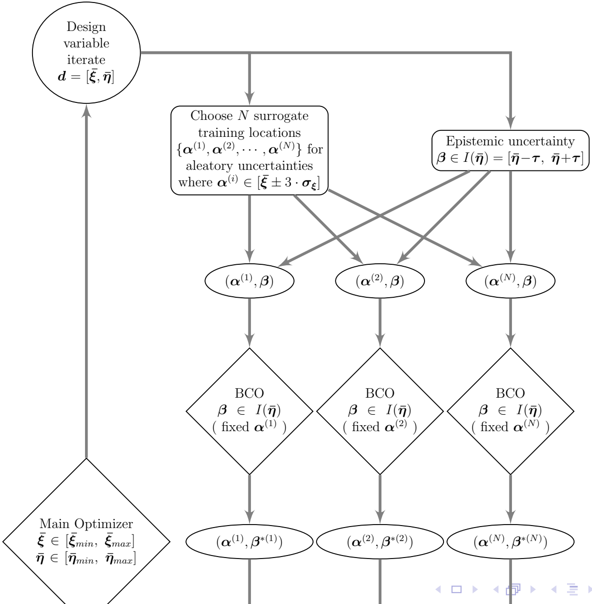
Deterministic Problem

$$\begin{aligned} \underset{\boldsymbol{d}}{\text{minimize}} \quad & \mathcal{J} = C_d, \\ \text{subject to} \quad & g = C_l - {C_l}^{+} \geq 0, \end{aligned}$$

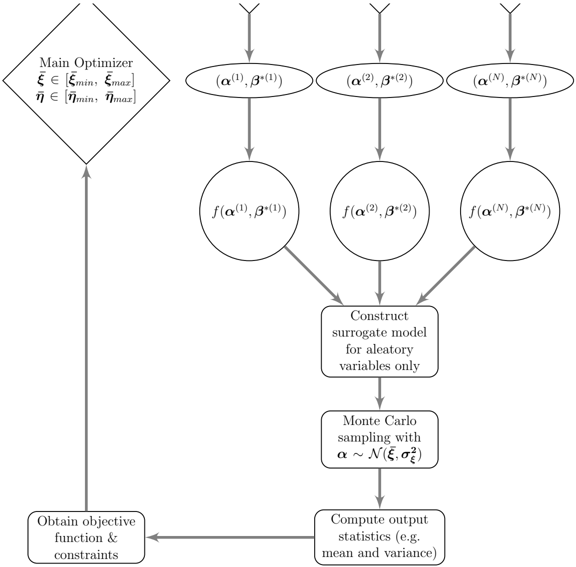
Robust Optimization Problem

$$\begin{aligned} \underset{\boldsymbol{\xi},\boldsymbol{\eta}}{\text{minimize}} \quad & \mathcal{J} = \mu_{C_{d_{max}}} + \sigma^2_{C_{d_{max}}}, \\ \text{subject to} \quad & g = (\mu_{C_{l_{min}}} + k\sigma_{C_{l_{min}}}) - {C_l}^{+} \geq 0, \end{aligned}$$

Mixed OUU Framework: IMCS+BCO I

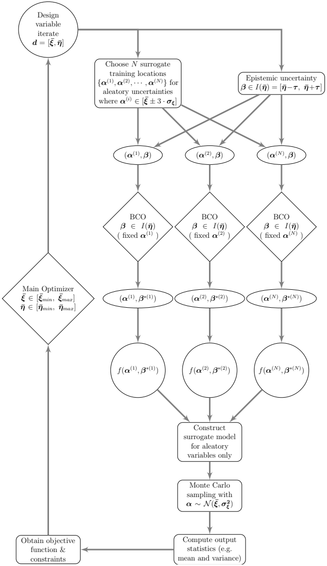


Mixed OUU Framework: IMCS+BCO II



Framework for optimization under mixed aleatory and epistemic uncertainties.

Mixed OUU Framework: IMCS+BCO III



Framework for optimization under mixed aleatory and epistemic uncertainties.

Airfoil Optimization I

Lift constrained drag minimization

Deterministic Problem

$$\begin{aligned} \underset{\boldsymbol{d}}{\text{minimize}} \quad & \mathcal{J} = C_d, \\ \text{subject to} \quad & g = C_l - {C_l}^+ \geq 0, \end{aligned}$$

Robust Optimization Problem

$$\begin{aligned} \underset{\boldsymbol{\xi},\boldsymbol{\eta}}{\text{minimize}} \quad & \mathcal{J} = \mu_{C_{d_{max}}} + \sigma^2_{C_{d_{max}}}, \\ \text{subject to} \quad & g = (\mu_{C_{l_{min}}} + k\sigma_{C_{l_{min}}}) - {C_l}^+ \geq 0, \end{aligned}$$

Airfoil Optimization II

Mean and variance from surrogate

$$\mathcal{J} = w_1 \, \mu_{f*} + w_2 \, \vartheta_{f*} \tag{2}$$

$$\mu_{f*} \approx \frac{1}{\widetilde{N}} \sum_{k=1}^{\widetilde{N}} \widehat{f^*}(\boldsymbol{\alpha}^k) \tag{3}$$

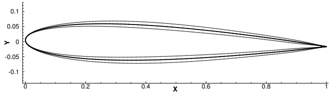
$$\vartheta_{f*} \approx \left(\frac{1}{\widetilde{N}} \sum_{k=1}^{\widetilde{N}} \widehat{f^*}^2(\boldsymbol{\alpha}^k) \right) - \mu_{f*}^2 \tag{4}$$

- w_1 and w_2 are user specified weights
- The Monte Carlo samples $\boldsymbol{\alpha}^{(k)}$, $k = 1, \ldots, \widetilde{N}$ are chosen based on their underlying probability distribution
- $\widehat{f^*}$ represents the surrogate approximated value of exact function f^*

Airfoil Optimization III

Data for robust optimization of airfoil

Random Variable	Description	Uncertainty Type	τ_{min}	τ_{max}	Standard Deviation
$\eta_{1,2,13,14}$	Shape design variables	Epistemic	-0.00125	0.00125	-
η_{3-12}	Shape design variables	Epistemic	-0.01	0.01	-
ξ_α	Angle of attack	Aleatory	-	-	0.1°
ξ_M	Mach number	Aleatory	-	-	0.01



The NACA 0012 airfoil (in black) and airfoils resulting from perturbations of ± 0.0025 (in gray).

- Seven shape design variables at 20%, 30%, 40%, 50%, 60%, 80%, and 90% chord
- Flow variable bounds: $0^\circ \leq \alpha \leq 4^\circ$ and $0.6 \leq M \leq 0.78$

Optimization Results

Optimization results for airfoil

Type	k	P_k	$\mu_{c_{d_{max}}}$	$\sigma^2_{c_{d_{max}}}$	$\mu_{c_{l_{min}}}$	$\sigma_{c_{l_{min}}}$	α	M	No. of F/FG Evals. & Iterations
Initial	-	-	$4.72 \cdot 10^{-4}$	-	0.335	-	2.000°	0.650	
Deterministic	-	-	$1.17 \cdot 10^{-3}$	-	0.600	-	2.510°	0.600	49/49 – 24
Robust-KR	0	0.5000	$2.72 \cdot 10^{-3}$	$2.03 \cdot 10^{-7}$	0.600	$1.84 \cdot 10^{-2}$	2.013°	0.600	844/844-23
Robust-PC	0	0.5000	$2.62 \cdot 10^{-3}$	$5.80 \cdot 10^{-8}$	0.600	$1.82 \cdot 10^{-2}$	2.389°	0.600	675/6751-16
Robust-KR	1	0.8413	$2.93 \cdot 10^{-3}$	$3.07 \cdot 10^{-7}$	0.619	$1.86 \cdot 10^{-2}$	2.065°	0.600	434/434-13
Robust-PC	1	0.8413	$2.73 \cdot 10^{-3}$	$2.50 \cdot 10^{-7}$	0.618	$1.84 \cdot 10^{-2}$	3.058°	0.600	434/434-15
Robust-KR	2	0.9772	$3.10 \cdot 10^{-3}$	$4.46 \cdot 10^{-7}$	0.637	$1.88 \cdot 10^{-2}$	2.179°	0.600	831/831-19
Robust-PC	2	0.9772	$3.20 \cdot 10^{-3}$	$8.58 \cdot 10^{-7}$	0.637	$1.89 \cdot 10^{-2}$	2.193°	0.600	710/710-22
Robust-KR	3	0.9986	$3.28 \cdot 10^{-3}$	$6.23 \cdot 10^{-7}$	0.657	$1.90 \cdot 10^{-2}$	2.301°	0.600	650/650-21
Robust-PC	3	0.9986	$3.25 \cdot 10^{-3}$	$9.83 \cdot 10^{-7}$	0.658	$1.92 \cdot 10^{-2}$	2.352°	0.600	1145/1145-21
Robust-KR	4	0.9999	$3.56 \cdot 10^{-3}$	$9.50 \cdot 10^{-7}$	0.677	$1.93 \cdot 10^{-2}$	2.421°	0.600	620/620-15
Robust-PC	4	0.9999	$3.65 \cdot 10^{-3}$	$1.25 \cdot 10^{-6}$	0.677	$1.93 \cdot 10^{-2}$	2.427°	0.600	2104/2104-36

Iteration History

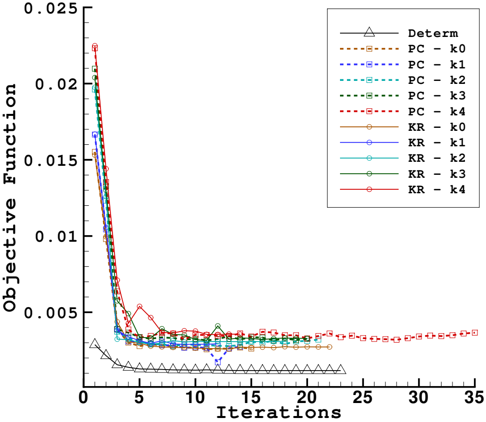
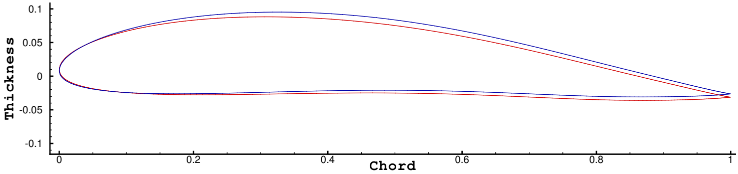
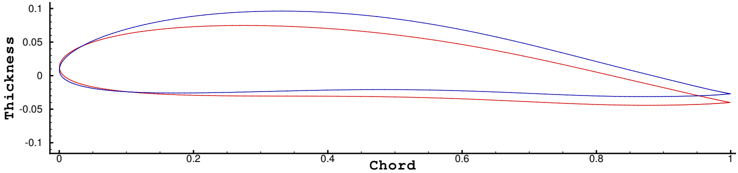


Figure : Optimizer iteration history for airfoil design problem.

Airfoil Shapes I



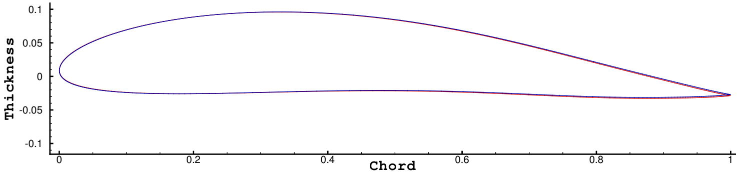
(a) Robust Airfoils $k = 0$



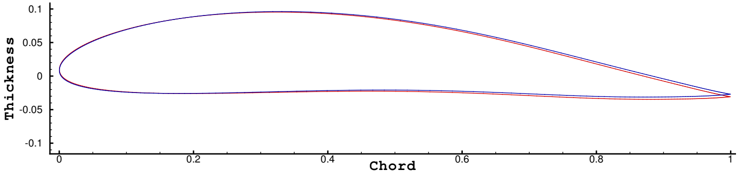
(b) Robust Airfoils $k = 1$

Figure : Red=Polynomial Chaos, Blue=Kriging

Airfoil Shapes II



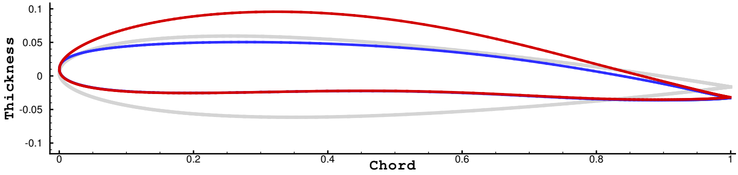
(a) Robust Airfoils $k = 2$



(b) Robust Airfoils $k = 3$

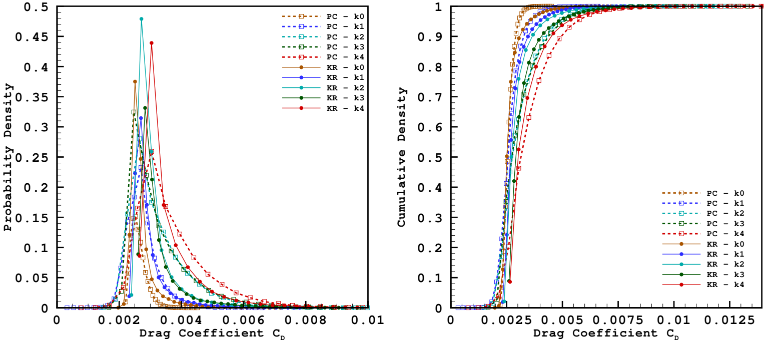
Figure : Red=Polynomial Chaos, Blue=Kriging

Airfoil Shapes III



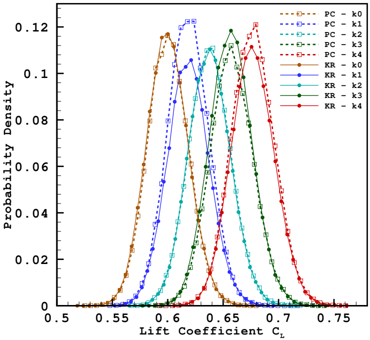
NACA 0012, Deterministic, Robust Airfoils corresponding to $k = 4$.

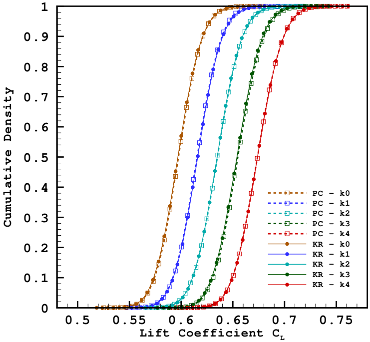
Output Distributions I



PDF and CDF drag coefficient at the optimum design.

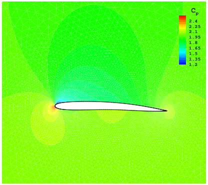
Output Distributions II



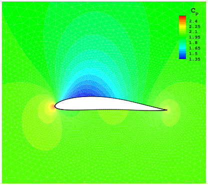


PDF and CDF lift coefficient at the optimum design.

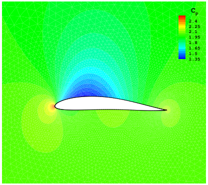
Pressure Distributions I



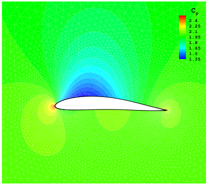
(a) Deterministic



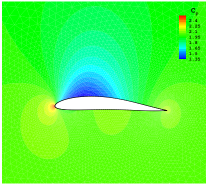
(b) Robust $k = 0$



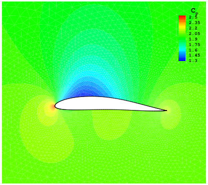
(c) Robust $k = 1$



(d) Robust $k = 2$



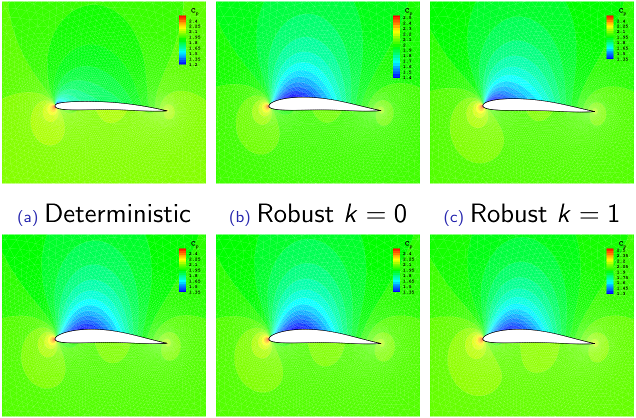
(e) Robust $k = 3$



(f) Robust $k = 4$

Figure : Contour plots of pressure coefficients C_p at different optimum designs using kriging.

Pressure Distributions II



(a) Deterministic (b) Robust $k = 0$ (c) Robust $k = 1$

(d) Robust $k = 2$ (e) Robust $k = 3$ (f) Robust $k = 4$

Figure : Contour plots of pressure coefficients C_p at different optimum designs using polynomial chaos.

Three Bar Truss I

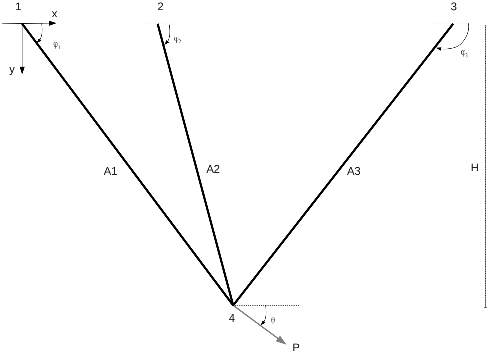


Figure : A schematic of the three-bar truss structure.

- Minimum weight truss design
- 8 constraints (6 stress, 2 displacement)
- Design variables (areas A_i and orientations ϕ_i)

Three Bar Truss II

Mathematical Formulation

$$
\begin{aligned}
\underset{\boldsymbol{d}}{\text{minimize}} \quad & W = \frac{A_1 \gamma H}{\sin(\phi_1)} + \frac{A_2 \gamma H}{\sin(\phi_2)} + \frac{A_3 \gamma H}{\sin(\phi_3)}, \\
\text{subject to} \quad & g_1 = \frac{\sigma_1}{\sigma_{1_{max}}} - 1 \le 0, \\
& g_2 = \frac{\sigma_2}{\sigma_{2_{max}}} - 1 \le 0, \\
& g_3 = \frac{\sigma_3}{\sigma_{3_{max}}} - 1 \le 0, \\
& g_4 = -\frac{\sigma_1}{\sigma_{1_{max}}} - 1 \le 0, \\
& g_5 = -\frac{\sigma_2}{\sigma_{2_{max}}} - 1 \le 0, \\
& g_6 = -\frac{\sigma_3}{\sigma_{3_{max}}} - 1 \le 0, \\
& g_7 = \frac{Q_{4x}}{Q_{4x_{max}}} - 1 \le 0, \\
& g_8 = \frac{Q_{4y}}{Q_{4y_{max}}} - 1 \le 0.
\end{aligned}
$$

Bounds

$$
\begin{aligned}
0.25\ in^2 \le \ & A_1,\ A_2,\ A_3 \ \le 5.0\ in^2, \\
30^\circ \le \ & \phi_1 \ \le 60^\circ, \\
60^\circ \le \ & \phi_2 \ \le 120^\circ, \\
120^\circ \le \ & \phi_3 \ \le 150^\circ.
\end{aligned}
$$

Solver

- Stresses and displacements using hand-coded FEA procedure

Three Bar Truss III

Table : Design data for three-bar truss.

Quantity	Description	Value	Unit
P	Load	30000	*lb*
θ	Loading angle	50	*deg*
E	Young's modulus	10^7	*psi*
γ	Weight density	0.1	lb/in^3
H	Reference length (projection on $y-$axis)	10	*in*
$\sigma_{1_{max}}$	Allowable axial stress on bar 1	5000	*psi*
$\sigma_{2_{max}}$	Allowable axial stress on bar 2	10000	*psi*
$\sigma_{3_{max}}$	Allowable axial stress on bar 3	5000	*psi*
$u_{4x_{max}}$	Allowable x-displacement at 4	0.005	*in*
$u_{4y_{max}}$	Allowable y-displacement at 4	0.005	*in*
ϵ_1	Constraint violation tolerance	10^{-3}	-
ϵ_2	Norm of design change $\|\|\Delta \boldsymbol{d}\|\|$	10^{-3}	-

Three Bar Truss IV

Robust Optimization Problem

$$\begin{aligned} \underset{\boldsymbol{\xi},\boldsymbol{\eta}}{\text{minimize}} \quad & \mathcal{J} = \mu_W + \vartheta_W, \\ \text{subject to} \quad & g_i^r = \mu_{g_i} + k\sigma_{g_i} \leq 0, \quad \text{for } i = 1, \ldots, 8 \end{aligned} \tag{5}$$

- Area design variables A_i (epistemic with $\tau_i = 0.1 \ in^2$)
  - Propagated via BCO
- Orientation design variables ϕ_i (aleatory with $\sigma_i = 0.1°$)
  - Propagation via surrogate sampling
  - Kriging and PCE built with 70 training points

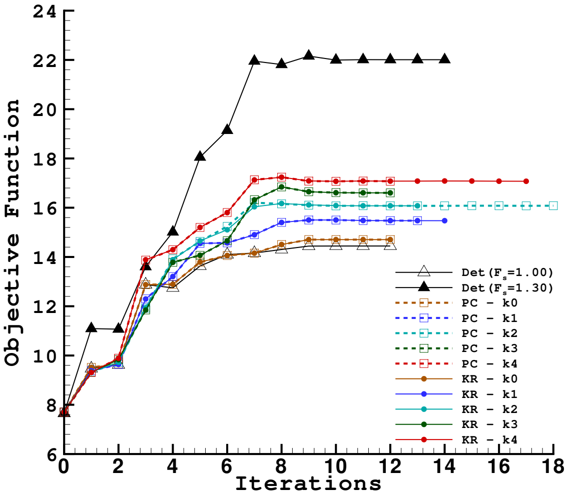
Three Bar Truss V

Table : Optimization results for three-bar truss problem.

Type	k	P_k	A_1 in^2	A_2 in^2	A_3 in^2	ϕ_1 *deg*	ϕ_2 *deg*	ϕ_3 *deg*	μ_W *lb*	σ_W *lb*	C_v -	No. of F/FG Evals. & Iterations
Initial design	-	-	2.0	2.0	2.0	45.0	90.0	135.0	7.66	-	-	-
Det $F_s = 1.0$	-	-	5.00	1.42	2.30	37.6	60.0	150.0	14.45	-	-	108/108-12
Det $F_s = 1.3$	-	-	5.00	4.95	5.00	39.5	60.0	143.6	22.00	-	-	126/126-14
Robust-KR	0	0.5000	5.00	1.45	2.37	37.7	60.0	150.0	14.65	0.24	0.0162	17559/17559-12
Robust-PC	0	0.5000	5.00	1.45	2.37	37.7	60.0	150.0	14.65	0.24	0.0162	17615/17615-12
Robust-KR	1	0.8413	5.00	1.66	2.66	37.5	60.0	149.3	15.41	0.24	0.0159	21963/21963-14
Robust-PC	1	0.8413	5.00	1.66	2.66	37.5	60.0	149.3	15.41	0.24	0.0159	20555/20555-13
Robust-KR	2	0.9772	5.00	1.84	2.92	37.5	60.0	148.6	16.02	0.25	0.0155	23594/23594-13
Robust-PC	2	0.9772	5.00	1.84	2.92	37.5	60.0	148.6	16.02	0.25	0.0155	33555/33555-18
Robust-KR	3	0.9986	5.00	1.99	3.15	37.5	60.0	148.2	16.54	0.25	0.0153	20771/20771-12
Robust-PC	3	0.9986	5.00	1.99	3.15	37.5	60.0	148.2	16.54	0.25	0.0153	17938/17938-12
Robust-KR	4	0.9999	5.00	2.13	3.36	37.6	60.0	147.9	17.00	0.26	0.0151	31178/31178-17
Robust-PC	4	0.9999	5.00	2.13	3.36	37.6	60.0	147.9	17.00	0.26	0.0151	19500/19500-12

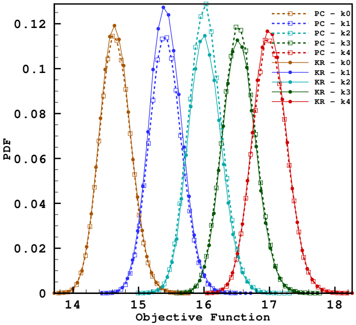
- A deterministic design with no F_s is 15% lighter than a robust design specified by $k = 4$.
- A deterministic design with F_s of 1.3 is 29% heavier than a robust design specified by $k = 4$.

Three Bar Truss VI

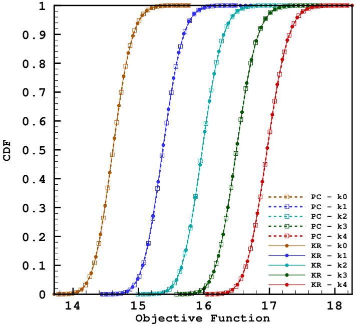


Change in objective function with the number of optimizer iterations.

Objective Function Distribution:

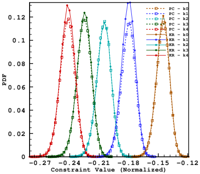


PDF

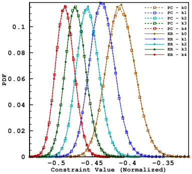


CDF

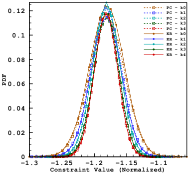
Three Bar Truss VIII



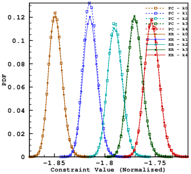
(a) Constraint 1



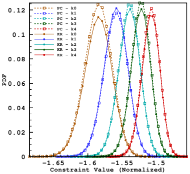
(b) Constraint 2



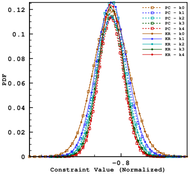
(c) Constraint 3



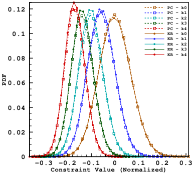
(d) Constraint 4



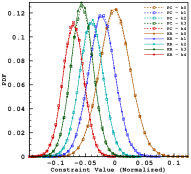
(e) Constraint 5



(f) Constraint 6



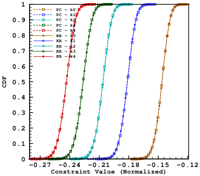
(g) Constraint 7



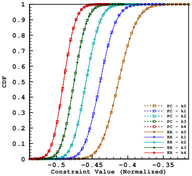
(h) Constraint 8

Probability density function of objective and constraint functions at robust and deterministic optimum designs.

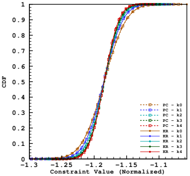
Three Bar Truss IX



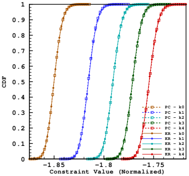
(a) Constraint 1



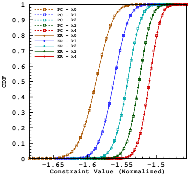
(b) Constraint 2



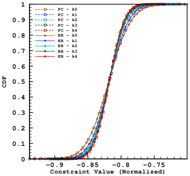
(c) Constraint 3



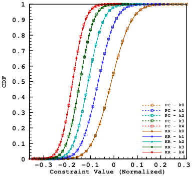
(d) Constraint 4



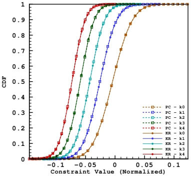
(e) Constraint 5



(f) Constraint 6



(g) Constraint 7



(h) Constraint 8

Cumulative distribution function of objective and constraint functions at robust and deterministic optimum designs.

Conclusion

- Training point selection:
  - Spreads the points and adds data in regions of larger uncertainty (measured by the discrepancy function)
  - More accurate than conventional approaches
  - Monotonicity in convergence
  - Selection in the presence/absence of derivative information
- Error estimate (discrepancy function, RMSD, MAD)
  - Shows promise for effective validation
  - Excellent matching of tendencies
  - No additional evaluations
- Application to Kriging and PCE (any surrogate model)
- Engineering application $\rightarrow$ robust optimization
  - Aleatory uncertainties using surrogate models
  - Epistemic uncertainties using bound constrained optimization
  - Mixed uncertainties using IMCS+BCO

Future Work

- Suitability of training point selection for surrogate-based optimizations
- Study other candidates for local surrogate models
- Apply the framework to other surrogate models
- Apply the OUU framework for engineering problems of practical interest (e.g. wing design)
- Study correlated and non-normally distributed variables

Publications

1. K. Boopathy and M.P. Rumpfkeil, "A Unified Framework for Training Point Selection and Error Estimation for Surrogate Models", AIAA Journal. In Revision.
2. K. Boopathy and M.P. Rumpfkeil, "Robust Optimizations of Structural and Aerodynamic Designs", 15th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, Atlanta, June 2014. Accepted.
3. K. Boopathy and M.P. Rumpfkeil, "A Multivariate Interpolation and Regression Enhanced Kriging Surrogate Model", 21st AIAA Computational Fluid Dynamics Conference, San Diego, June 2013. AIAA Paper 2013-2964.
4. K. Boopathy and M.P. Rumpfkeil, "Building Aerodynamic Databases Using Enhanced Kriging Surrogate Models", AIAA Region III Student Conference, Chicago, April 2013.

Acknowledgments

1. Wataru Yamazaki – Kriging surrogate
2. Karthik Mani – Euler Solver
3. Qiqi Wang – MIR Model

References

1. Arora, J. S., "Optimization of Structural and Mechanical Systems", World Scientific Publishing Co. Pte. Ltd., 2007.
2. Keane, A. and Nair, P., "Computational Approaches for Aerospace Design", John Wiley & Sons, 2005
3. Wang, Q., Moin, P., and Iaccarino, G., "A High-Order Multi-Variate Approximation Scheme for Arbitrary Data Sets," Journal of Computational Physics, Vol. 229, No. 18, 2010, pp. 6343–6361.
4. Sacks, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P., "Design and Analysis of Computer Experiments," Statistical Science, Vol. (4), 1989, pp. 409–423.
5. Yamazaki, W. and Mavripilis, D. J., "Derivative-Enhanced Variable Fidelity Surrogate Modeling for Aerodynamic Functions," AIAA Journal , Vol. 51, No. 1, 2013, pp. 126–137.
6. Helton, J. C., Oberkampf, J. D. J. W. L., and Sallaberry, C. J., "Representation of Analysis Results Involving Aleatory and Epistemic Uncertainty," Tech. Rep.SAND2008-4379, Sandia National Laboratories, 2008.

Any Questions?

Cantilever Beam Design I

Problem Formulation

$$
\begin{aligned}
\underset{b,d}{\text{minimize}} \quad & A(b,d) = bd, \\
\text{subject to} \quad & g_1(b,d,\mathcal{M}) = \frac{6\mathcal{M}}{bd^2\sigma_{allow}} - 1 \le 0, \\
& g_2(b,d,\mathcal{V}) = \frac{3\mathcal{V}}{2bd\tau_{allow}} - 1 \le 0, \\
& g_3(b,d) = \frac{d}{2b} - 1 \le 0, \\
\text{bounds} \quad & 100\ mm \ \le \ b, d \ \le \ 600\ mm,
\end{aligned}
$$

Cantilever Beam Design II

Table : Data and assumed uncertain parameters for cantilever beam design problem.

Random Variable	Description	Uncertainty Type	τ_{min}	τ_{max}	Mean	Standard Deviation	Unit
b	Breadth	Epistemic	-10	10	-	-	mm
d	Width	Epistemic	-10	10	-	-	mm
$\mathcal{M}$	Bending Moment	Aleatory	-	-	$40 \cdot 10^6$	40000	$N \cdot mm$
$\mathcal{V}$	Shear Force	Aleatory	-	-	$150 \cdot 10^3$	1500	N

Robust Optimization Problem

$$\begin{aligned} \underset{b,d}{\text{minimize}} \quad & A(b,d) = \mu_A + \sigma_A^2, \\ \text{subject to} \quad & g_1^r(b,d,\mathcal{M}) = \mu_{g_1} + k\sigma_{g_1} \leq 0, \\ & g_2^r(b,d,\mathcal{V}) = \mu_{g_2} + k\sigma_{g_2} \leq 0, \\ & g_3^r(b,d) = \mu_{g_3} + k\sigma_{g_3} \leq 0. \end{aligned}$$

Cantilever Beam Design III

Table : Optimization results for cantilever beam design problem.

Type	k	P_k	Width b mm	Depth d mm	Area A $\cdot 10^3$ mm^2	No. of F/FG Evals. & Iterations
Initial Design	-	-	300	300	90.0	-
Det ($F_s = 1.0$)	-	-	335.5	335.4	112.5	33/33-7
Det ($F_s = 1.5$)	-	-	595.5	283.4	168.7	45/45-8
Robust-KR	0	0.5000	347.4	343.4	126.3	7046/3523-7
Robust-PC	0	0.5000	347.4	343.4	126.3	7917/7917-8
Robust-KR	1	0.8413	349.7	344.5	127.5	7146/3573-7
Robust-PC	1	0.8413	349.7	344.5	127.5	8037/8037-8
Robust-KR	2	0.9772	398.5	305.4	128.8	7686/3843-7
Robust-PC	2	0.9772	398.5	305.4	128.8	9661/9661-9
Robust-KR	3	0.9986	386.5	317.8	130.0	8694/4347-8
Robust-PC	3	0.9986	386.5	317.8	130.0	11669/11669-10
Robust-KR	4	0.9999	356.6	347.5	131.1	7286/3643-7
Robust-PC	4	0.9999	356.6	347.5	131.1	8196/8196-8

Cantilever Beam Design IV

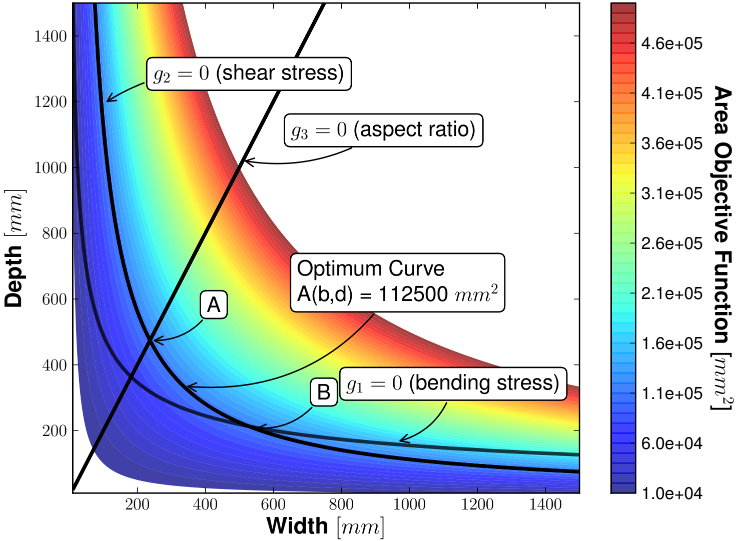


Figure : Graphical solution to the minimum area beam design problem.

Kriging Vs. PCE I

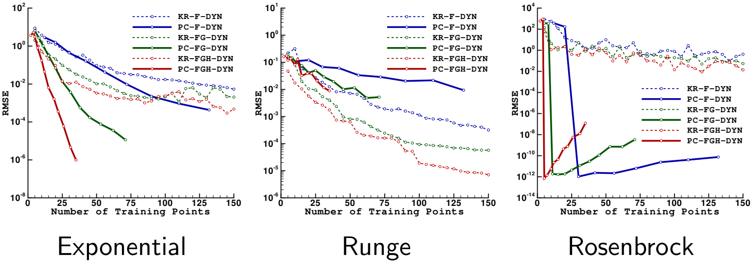
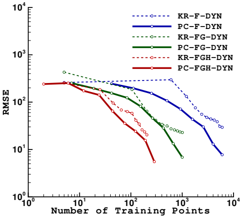
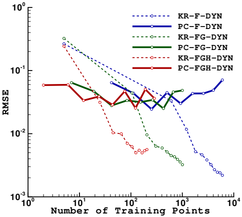


Figure : Kriging versus PCE in 2D.

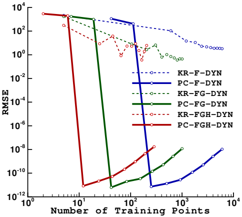
Kriging Vs. PCE II



Exponential



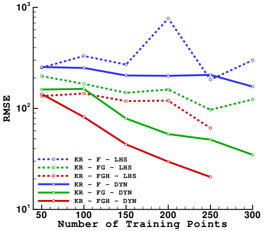
Runge



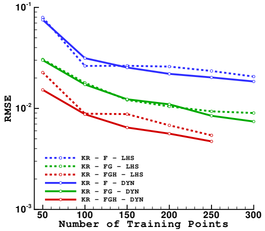
Rosenbrock

Figure : Kriging versus PCE in 5D.

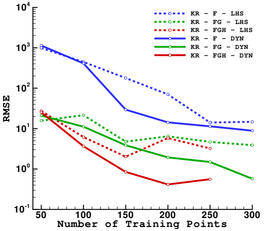
Five Dimensional Results I



Exponential



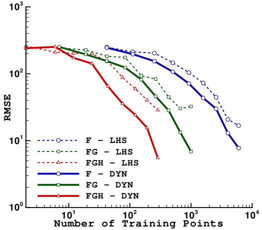
Runge



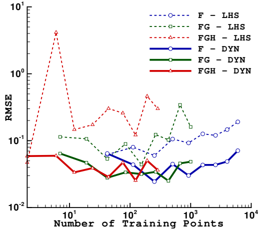
Rosenbrock

Figure : Kriging 5D

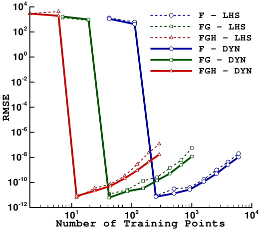
Five Dimensional Results II



Exponential



Runge



Exponential

Figure : PCE in 5D

Aleatory Gradients

Aleatory gradients

$$\frac{d\mathcal{J}}{d\xi} = \frac{\partial \mathcal{J}}{\partial \mu_{f*}} \frac{d\mu_{f*}}{d\xi} + \frac{\partial \mathcal{J}}{\partial \vartheta_{f*}} \frac{d\vartheta_{f*}}{d\xi} = w_1 \boxed{\frac{d\mu_{f*}}{d\xi}} + w_2 \boxed{\frac{d\vartheta_{f*}}{d\xi}} \tag{6}$$

$$\boxed{\frac{d\mu_{f*}}{d\xi}} \approx \frac{1}{\widetilde{N}} \sum_{k=1}^{\widetilde{N}} \frac{d\widehat{f^*}(\alpha^k)}{d\alpha^k} \frac{d\alpha^k}{d\xi} = \frac{1}{\widetilde{N}} \sum_{k=1}^{\widetilde{N}} \frac{d\widehat{f^*}(\alpha^k)}{d\alpha^k} \tag{7}$$

$$\boxed{\frac{d\vartheta_{f*}}{d\xi}} \approx \left(\frac{2}{\widetilde{N}} \sum_{k=1}^{\widetilde{N}} \widehat{f^*}(\alpha^k) \frac{d\widehat{f^*}(\alpha^k)}{d\alpha^k} \right) - 2\mu_{f*} \frac{d\mu_{f*}}{d\xi} \tag{8}$$

Epistemic Gradients I

Epistemic gradients

$$\frac{d\mathcal{J}}{d\boldsymbol{\eta}} = \frac{\partial \mathcal{J}}{\partial \mu_{f*}} \frac{d\mu_{f*}}{d\boldsymbol{\eta}} + \frac{\partial \mathcal{J}}{\partial \vartheta_{f*}} \frac{d\vartheta_{f*}}{d\boldsymbol{\eta}} = w_1 \frac{d\mu_{f*}}{d\boldsymbol{\eta}} + w_2 \frac{d\vartheta_{f*}}{d\boldsymbol{\eta}} \tag{9}$$

Approximations

$$\frac{d\mu_{f*}}{d\boldsymbol{\eta}} \approx \left.\frac{df^*}{d\boldsymbol{\eta}}\right|_{(\boldsymbol{\xi}=\bar{\boldsymbol{\xi}},\boldsymbol{\eta}=\bar{\boldsymbol{\eta}})} \quad \text{and} \quad \frac{d\vartheta_{f*}}{d\boldsymbol{\eta}} \approx 0 \tag{10}$$

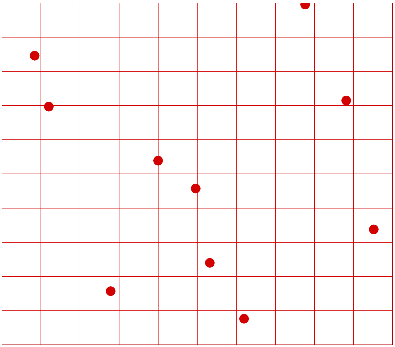
Training Point Selection I

1. **Domain based**
    - Monte-Carlo
    - Latin Hypercube
    - Delaunay Triangulation
2. **Response based (adaptive)**
    - Distance / Function values / Gradients / Physics

Monte-Carlo

- Random number generator
- Very simple to program
- No control over locations

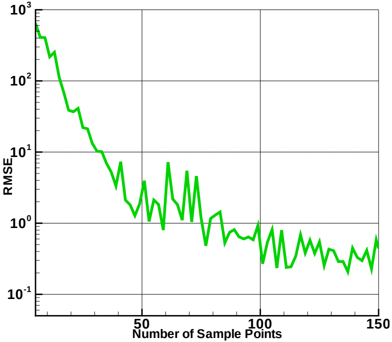
Training Point Selection II



Latin Hypercube

- McKay - while designing computer experiments
- Equal probability
- N^M bins in the design space
- No two points lie in the same bin

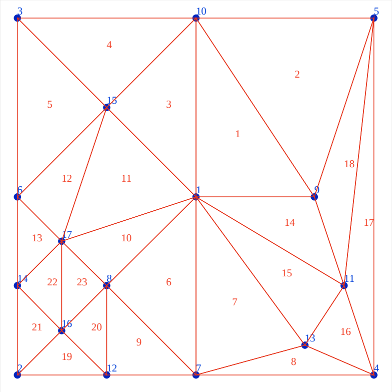
Training Point Selection III



Latin Hypercube

- McKay - while designing computer experiments
- Equal probability
- N^M bins in the design space
- No two points lie in the same bin

Training Point Selection IV



Delaunay Triangulation

- Geometrical method
- Split into hyper triangles
- Poor scaling to higher dimensions