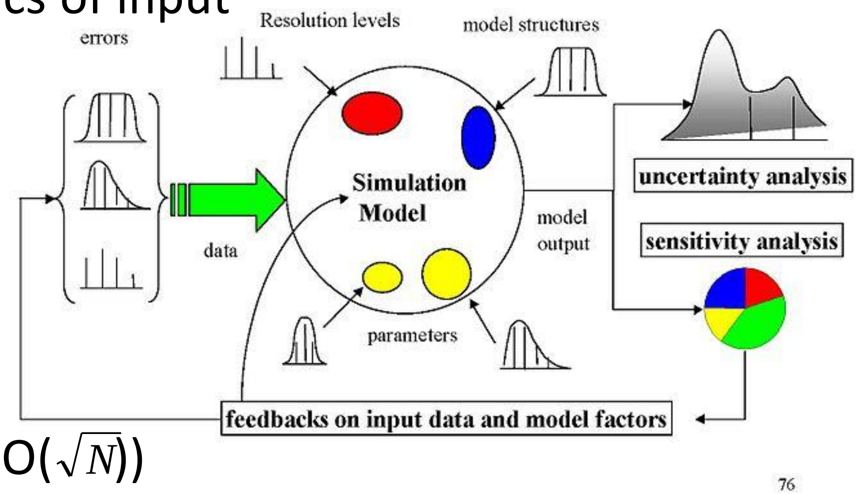
UNIVERSITY OF WYOMING

Adjoint Methods for Uncertainty Quantification

Dimitri Mavriplis
Department of Mechanical Engineering
University of Wyoming
Laramie WY USA

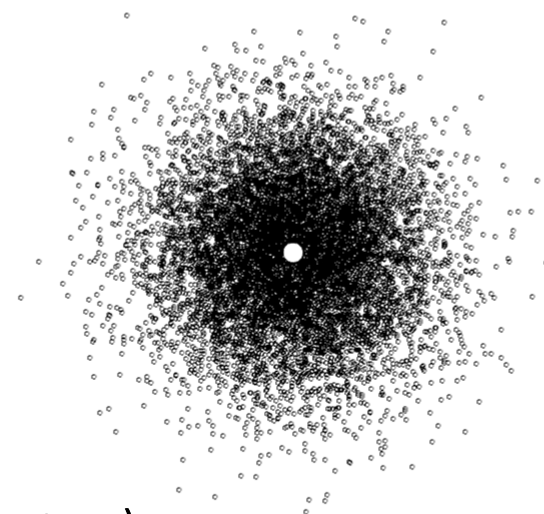
Motivation

- Motivation
  - Deterministic prediction inadequate for many engineering applications
  - Statistics of output given statistics of input
- Types of Uncertainty
  - Aleatory
  - Epistemic
  - Combined Aleatory/Epistemic
- Monte Carlo methods
  - Sampling
  - Expensive to build up statistics (O($\sqrt{N}$))
- Surrogate Models
  - Curse of dimensionality (number of parameters)
- Explore ways in which adjoint methods can reduce cost of UQ
  - Principal advantage: N derivatives for 1 adjoint solve (per output)

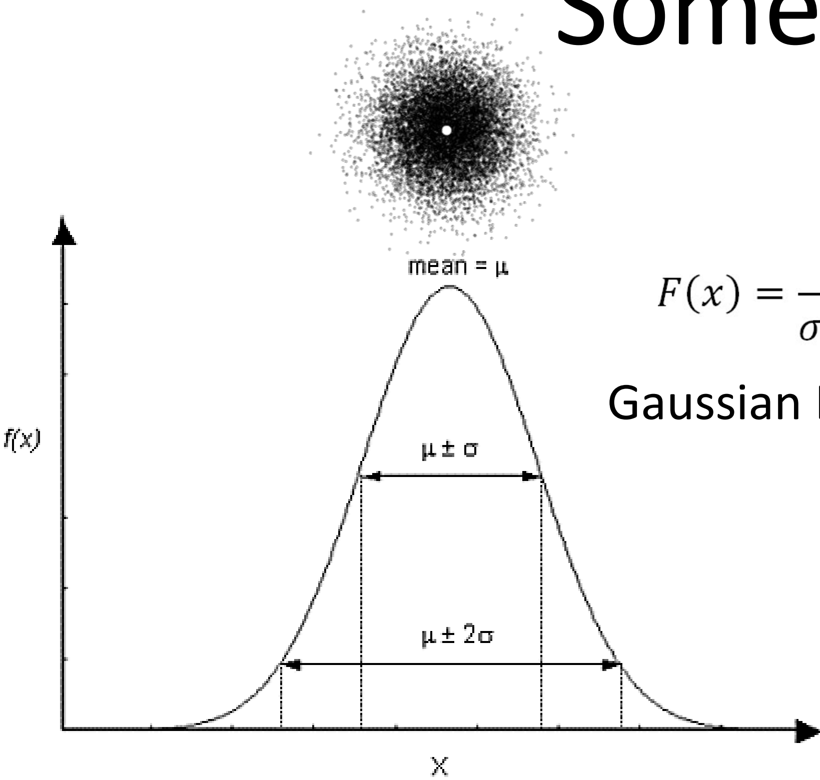


Motivation

- Motivation
  - Deterministic prediction inadequate for many engineering applications
  - Statistics of output given statistics of input
- Types of Uncertainty
  - Aleatory
  - Epistemic
  - Combined Aleatory/Epistemic
- Monte Carlo methods
  - Sampling
  - Expensive to build up statistics (O($\sqrt{N}$))
- Surrogate Models
  - Curse of dimensionality (number of parameters)
- Explore ways in which adjoint methods can reduce cost of UQ
  - Principal advantage: N derivatives for 1 adjoint solve (per output)



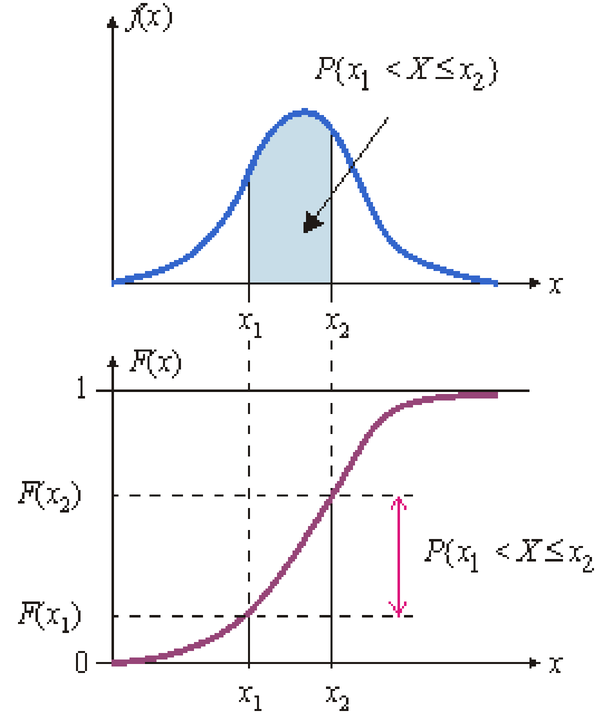
Some Definitions



Probability Distribution Function (PDF)

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2 / 2\sigma^2}$$

Gaussian Distribution



Cumulative Distribution Function (CDF)

Some Definitions

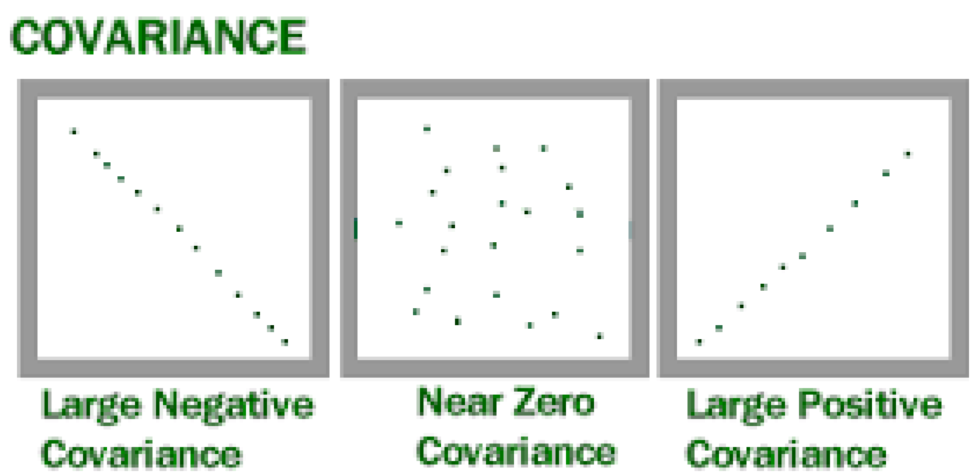
<u>Sample Variance</u>

$$s^2 = \frac{\sum(x - \bar{x})^2}{n - 1}$$

<u>Sample Standard Deviation</u>

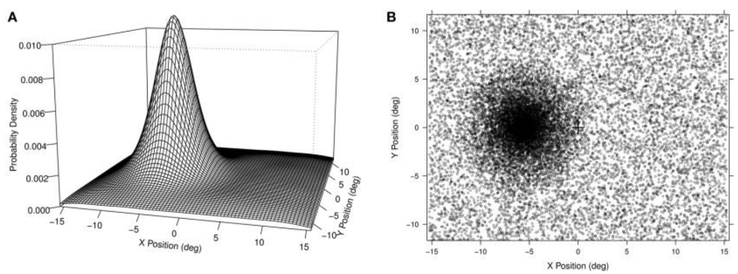
$$s = \sqrt{\frac{\sum(x - \bar{x})^2}{n - 1}}$$

$$Cov_{xy} = \frac{\Sigma(x - \bar{x})(y - \bar{y})}{(n-1)} = \frac{\Sigma xy - n\overline{xy}}{(n-1)}$$



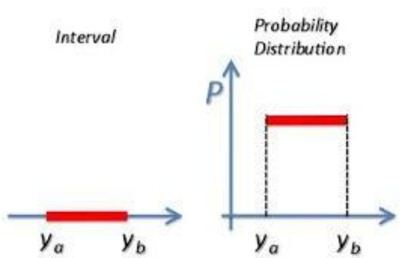
Some Definitions

- Aleatory uncertainty
  - Inherently variable, defined by PDF



- Monte Carlo sampling should follow prescribed input distribution
  - Markov Chain Monte Carlo

- Epistemic uncertainty
  - Lack of knowledge, defined by interval



- Monte Carlo sampling should follow prescribed input distribution
  - Latin Hypercube sampling (uniform)

Different UQ Requirements

- Different Forms of Uncertainty:
  1. **Aleatory**:
     - Due to inherent randomness
     - Specified with probability distribution
     - Quantified using Monte Carlo Sampling ($\sim 10^3 - 10^4$)
  2. **Epistemic**:
     - Due to lack of knowledge about exact value
     - Specified by interval
     - Quantified using Latin Hypercube sampling ($\sim 3^d$)
  3. **Mixed**:
     - Inputs have different forms
     - Quantified using Mixed Sampling ($\sim 3^{d+8}$)
     - Output distribution has interval
- Each form extremely expensive to quantify for complex simulations (Aleatory $\lll$ Epistemic $\lll$ Mixed)
- Different Gradient-based strategies used for each

Outline

- Tangent and Adjoint for First-order sensitivities
- Tangent and Adjoint for Second-order sensitivities (Hessian)
- Preliminary examples of adjoint/Hessian in UQ
  - Extrapolation about Mean
  - Method of Moments
  - Inexpensive Monte Carlo
- Surrogate Model Construction
  - Gradient/Hessian Enhanced Polynomial Regression
  - Gradient/Hessian Enhanced Kriging Model
- Epistemic Uncertainty Quantification
  - Intervals
  - Gradient-based bound optimization
- Example: Hypersonic Flow UQ
  - Combined aleatory/epistemic uncertainties

Adjoint Formulation for Parameter Sensitivity

$$L = L(D_j, U(D_j))$$

$$\frac{dL}{dD_j} = \frac{\partial L}{\partial D_j} + \frac{\partial L}{\partial U}\frac{\partial U}{\partial D_j}$$

$$R(U(D_j), D_j) = 0$$

$$\left[\frac{\partial R}{\partial U}\right]\frac{\partial U}{\partial D_j} = -\frac{\partial R}{\partial D_j}$$

$$\frac{dL}{dD_j} = \frac{\partial L}{\partial D_j} + \Lambda^T \frac{\partial R}{\partial D_j}$$

$$\left[\frac{\partial R}{\partial U}\right]^T \Lambda = -\left[\frac{\partial L}{\partial U}\right]^T$$

- Single adjoint solution gives sensitivity of L with respect to all parameters D_j

Second Order Sensitivities: Hessian

$$\frac{d^2 L}{dD_j dD_k} = \mathfrak{D}_{ik} L + \frac{\partial L}{\partial \mathbf{U}} \frac{\partial^2 \mathbf{U}}{\partial D_j \partial D_k}$$

$$\mathfrak{D}_{ik} L = \frac{\partial^2 L}{\partial D_j \partial D_k} + \frac{\partial^2 L}{\partial \mathbf{U} \partial D_k} \frac{\partial \mathbf{U}}{\partial D_j} + \frac{\partial^2 L}{\partial \mathbf{U} \partial D_j} \frac{\partial \mathbf{U}}{\partial D_k} + \frac{\partial^2 L}{\partial \mathbf{U}^2} \frac{\partial \mathbf{U}}{\partial D_j} \frac{\partial \mathbf{U}}{\partial D_j}$$

$$\left[\frac{\partial \mathbf{R}}{\partial \mathbf{U}}\right] \frac{\partial^2 \mathbf{U}}{\partial D_j \partial D_k} = -\mathfrak{D}_{ik} \mathbf{R}$$

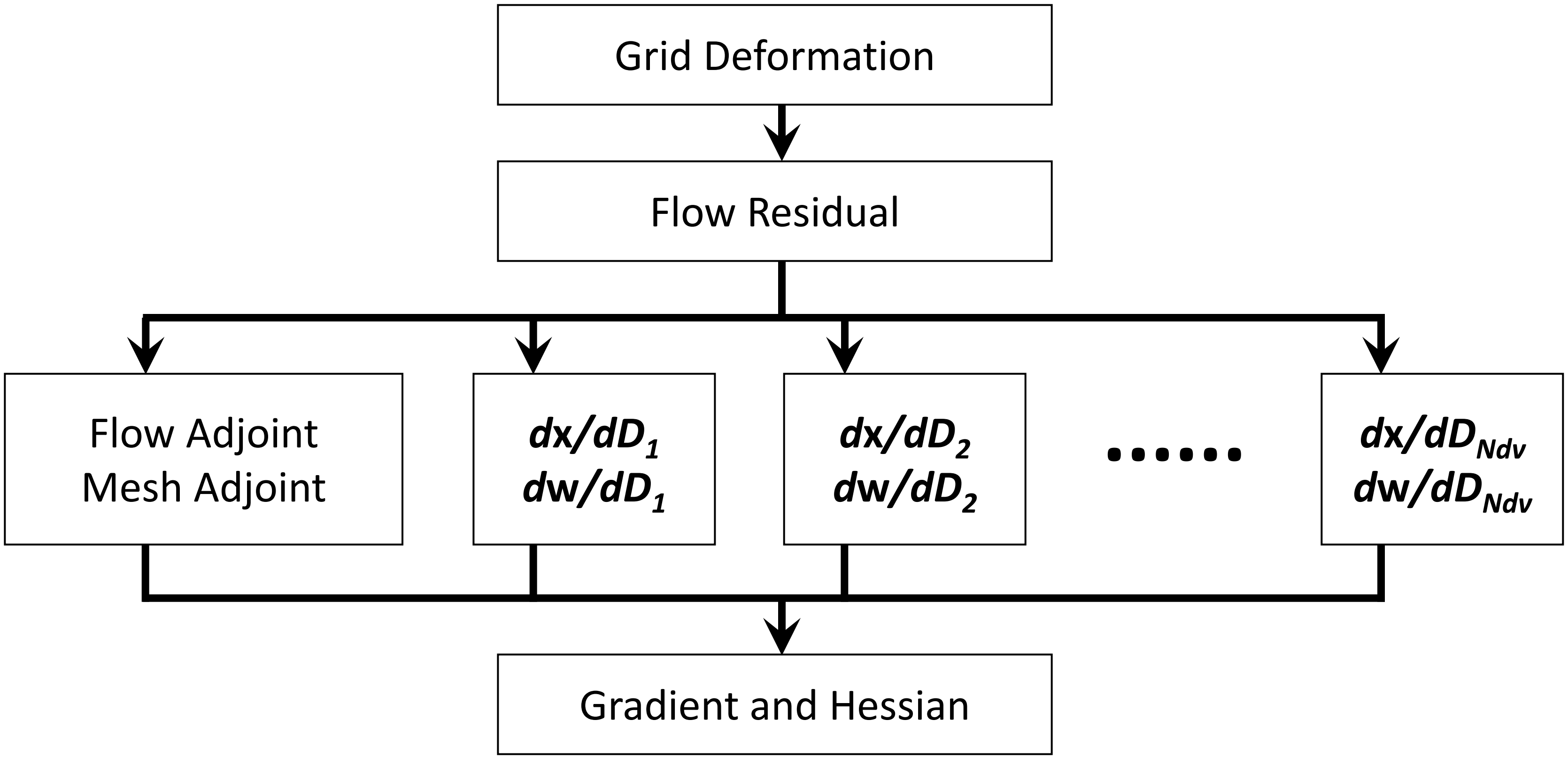
$$\mathfrak{D}_{ik} \mathbf{R} = \frac{\partial^2 \mathbf{R}}{\partial D_j \partial D_k} + \frac{\partial^2 \mathbf{R}}{\partial \mathbf{U} \partial D_k} \frac{\partial \mathbf{U}}{\partial D_j} + \frac{\partial^2 \mathbf{R}}{\partial \mathbf{U} \partial D_j} \frac{\partial \mathbf{U}}{\partial D_k} + \frac{\partial^2 \mathbf{R}}{\partial \mathbf{U}^2} \frac{\partial \mathbf{U}}{\partial D_j} \frac{\partial \mathbf{U}}{\partial D_j}$$

$$\frac{d^2 L}{dD_j dD_k} = \mathfrak{D}_{ik} L + \mathbf{\Lambda}^T \mathfrak{D}_{ik} \mathbf{R}$$

- Hessian can be computed with one adjoint solution and N forward sensitivity solutions $\frac{\partial U}{\partial D_j}$ j=1,…,N

Efficient CFD Hessian Calculation

An efficient CFD Hessian calculation method
by Adjoint method and Automatic Differentiation (AD)



Extrapolation Model using First and Second-Order Sensitivities

$$L_{\text{Lin}}\Big(D, \mathbf{x}(D), \mathbf{U}(D)\Big) = L\Big(D_0, \mathbf{x}(D_0), \mathbf{U}(D_0)\Big) + \sum_{j=1}^{N} \frac{dL}{dD_j}\bigg|_{D_0} \cdot \Delta D_j$$

$$L_{\text{Quad}}\Big(D, \mathbf{x}(D), \mathbf{U}(D)\Big) = L_{\text{Lin}}\Big(D, \mathbf{x}(D), \mathbf{U}(D)\Big) + \sum_{j=1}^{N} \sum_{k=1}^{N} \frac{1}{2} \frac{\partial^2 L}{\partial D_j \partial D_k}\bigg|_{D_0} \cdot \Delta D_j \Delta D_k$$

- Calculate functional value near nominal location D_o using nominal function value and its derivatives
  - Low cost one derivatives obtained
  - Use Monte Carlo sampling of input parameters D_j

Method of Moments

- Linear

$$\mu_L^{(1)} = L(D_0)$$

Mean output = output of mean

$$\sigma_L^{(1)} = \sqrt{\sum_{j=1}^{M} \left(\left.\frac{dL}{dD_j}\right|_{D_0} \sigma_{D_j} \right)^2}$$

- Quadratic

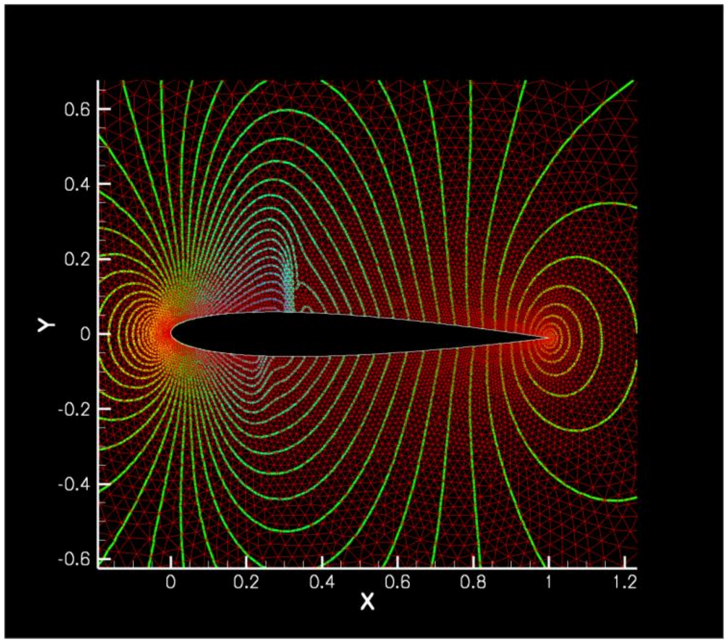
$$\mu_L^{(2)} = \mu_L^{(1)} + \frac{1}{2} \sum_{j=1}^{M} \left(\left.\frac{d^2L}{dD_j^2}\right|_{D_0} \sigma_{D_j}^2 \right)$$

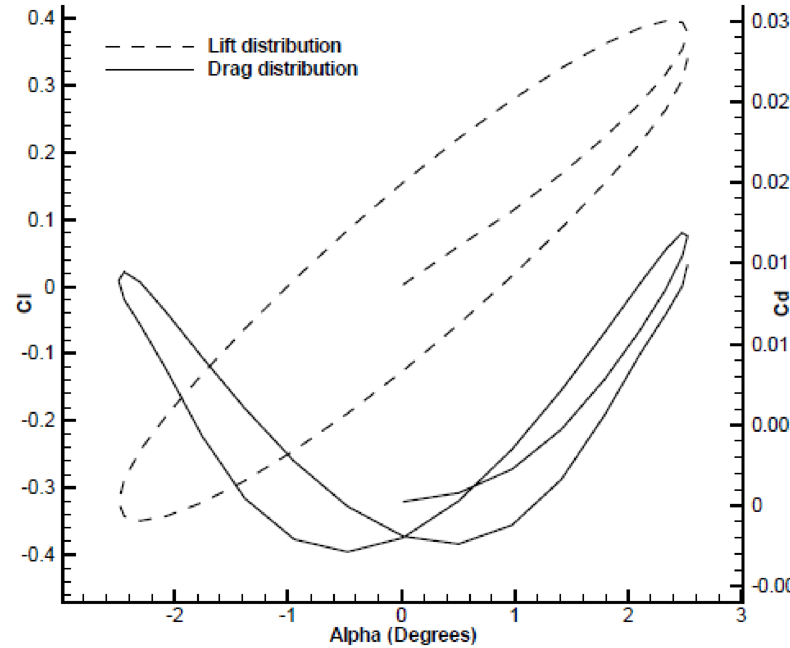
Non-linear shift in mean

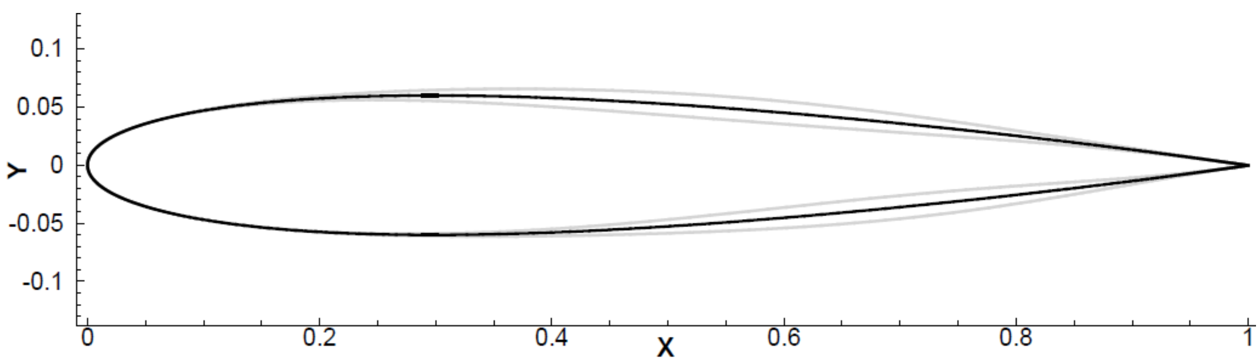
$$\sigma_L^{(2)} = \sqrt{\sum_{j=1}^{M} \left(\left.\frac{dL}{dD_j}\right|_{D_0} \sigma_{D_j} \right)^2 + \frac{1}{2} \sum_{j=1}^{M} \sum_{k=1}^{M} \left(\left.\frac{d^2L}{dD_j dD_k}\right|_{D_0} \sigma_{D_j} \sigma_{D_k} \right)^2}$$

- Propagate mean and variance using first and second-order derivatives
- Only provides output mean and variance (no probability distributions)

Uncertainty Quantification using Gradients and Hessian Information







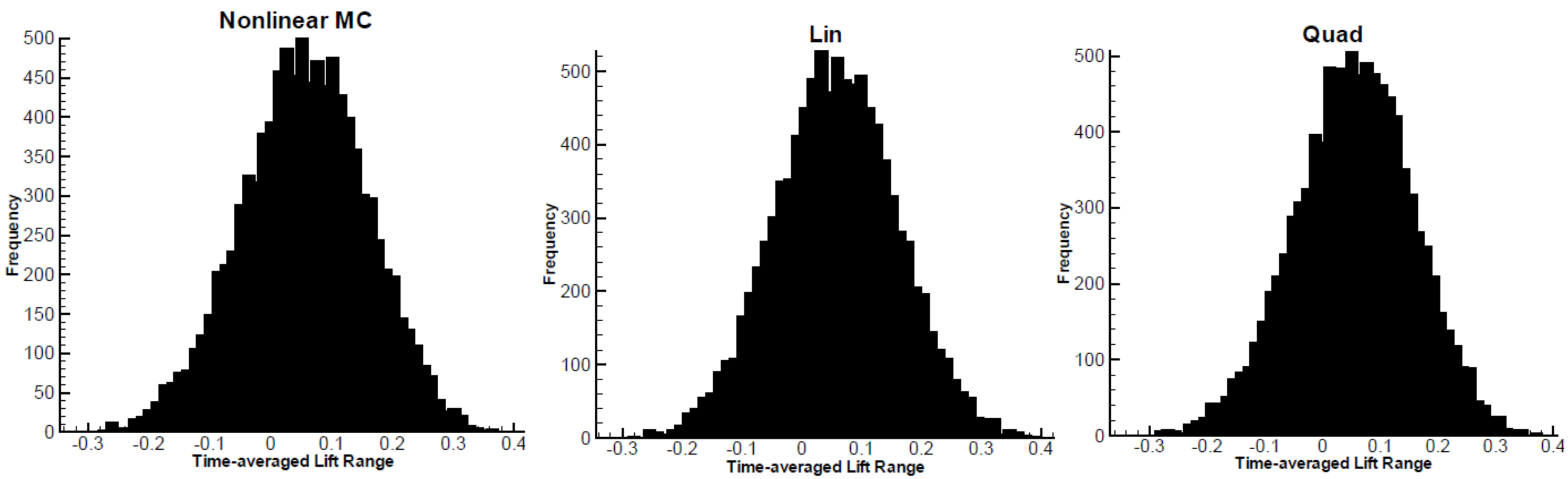
2 uncertain shape parameters

Uncertainty Quantification using Gradients and Hessian Information

- Prescribed input (D) statistical distributions
  - Mean values (nominal NACA0012)
  - Standard deviation = 0.01
  - Normal (Gaussian) distribution
- Full Monte Carlo: Compute Lift mean and Standard Deviation by sampling input distribution and running CFD for each sample of inputs
  - Compare with Method of Moments
  - Compare with Monte Carlo using extrapolation instead of CFD runs

	Mean	Standard deviation	Run time (minutes)
Nonlinear	5.55×10^{-2}	1.03×10^{-1}	$150,000$
MM1	5.81×10^{-2}	1.02×10^{-1}	30
MM2	5.39×10^{-2}	1.02×10^{-1}	60
Lin	5.82×10^{-2}	1.02×10^{-1}	30
Quad	5.39×10^{-2}	1.03×10^{-1}	60

Probability Distribution of Lift

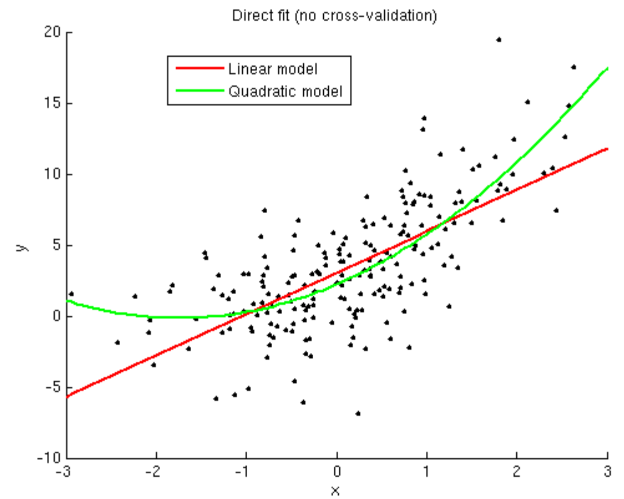


- Monte Carlo gives full statistics of output
  - Inexpensive Monte Carlo using extrapolation gives similar PDF
  - Only valid for small regions near mean
  - Method of Moments gives no PDF information

Surrogate Models

- Extrapolation is a simple example of a surrogate model
  - Requires 1 functional and adjoint evaluation
    - Additional forward linear problems at same location for Hessian
  - Only valid near evaluated location
- More sophisticated surrogate models make use of information at many locations in parameter space
  - Polynomial Regression
  - Kriging Models
  - Number of samples grows exponentially with dimension of parameter space (number of individual parameter)
- These can also be enhanced with
  - Gradient information
  - Hessian information
- Pros and Cons of Adjoint enhancement
  - ++ More information for low cost (N values for 1 adjoint)
  - -- All information is local

Polynomial Regression



- Fitting of statistical data with polynomial model $$y(D) = \sum_s \beta_s \Psi_s(D)$$
  - Best fit in least-square sense
  - Multi-dimensional
  - High dimensional (each parameter/input corresponds to a dimension)

$$\begin{bmatrix} \Psi_1(D_1) & \Psi_2(D_1) & \cdots & \Psi_s(D_1) \\ \Psi_1(D_2) & \Psi_2(D_2) & \cdots & \Psi_s(D_2) \\ \vdots & \ddots & \ddots & \vdots \\ \Psi_1(D_{N-1}) & \Psi_2(D_{N-1}) & \cdots & \Psi_s(D_{N-1}) \\ \Psi_1(D_N) & \Psi_2(D_N) & \cdots & \Psi_s(D_N) \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_s \end{bmatrix} = \begin{bmatrix} y(D_1) \\ y(D_2) \\ \vdots \\ y(D_{N-1}) \\ y(D_N) \end{bmatrix}$$

$$H\beta = Y \qquad H^T H\beta = A\beta = H^T Y \qquad \beta = A^{-1} H^T Y$$

Gradient Enhanced Polynomial Regression

- Incorporate gradient information by differentiating polynomial bases
- Require least squares best match of function values and gradients

$$y(D) = \sum_s \beta_s \Psi_s(D)$$

$$\frac{\partial y(D)}{\partial D_k} = \sum_s \beta_s \frac{\partial \Psi_s(D)}{\partial D_k}$$

$$\begin{bmatrix} \Psi_1(D_1) & \Psi_2(D_1) & \cdots & \Psi_s(D_1) \\ \frac{\partial \Psi_1(D_1)}{\partial D_1} & \frac{\partial \Psi_2(D_1)}{\partial D_1} & \cdots & \frac{\partial \Psi_s(D_1)}{\partial D_1} \\ \vdots & \ddots & \ddots & \vdots \\ \frac{\partial \Psi_1(D_1)}{\partial D_d} & \frac{\partial \Psi_2(D_1)}{\partial D_d} & \cdots & \frac{\partial \Psi_s(D_1)}{\partial D_d} \\ \vdots & \ddots & \ddots & \vdots \\ \Psi_1(D_N) & \Psi_2(D_N) & \cdots & \Psi_s(D_N) \\ \frac{\partial \Psi_1(D_N)}{\partial D_1} & \frac{\partial \Psi_2(D_N)}{\partial D_1} & \cdots & \frac{\partial \Psi_s(D_N)}{\partial D_1} \\ \vdots & \ddots & \ddots & \vdots \\ \frac{\partial \Psi_1(D_N)}{\partial D_d} & \frac{\partial \Psi_2(D_N)}{\partial D_d} & \cdots & \frac{\partial \Psi_s(D_N)}{\partial D_d} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_s \end{bmatrix} = \begin{bmatrix} y(D_1) \\ \frac{\partial y(D_1)}{\partial D_1} \\ \vdots \\ \frac{\partial y(D_1)}{\partial D_d} \\ \vdots \\ y(D_N) \\ \frac{\partial y(D_N)}{\partial D_1} \\ \vdots \\ \frac{\partial y(D_N)}{\partial D_d} \end{bmatrix}$$

- Match d derivatives at each sample point for cost of 1 adjoint

Gradient Enhanced Polynomial Regression

- Minimum number of sample points for function only regression

$$S = \frac{(d+p)!}{d!p!}$$

- d = number of parameters (dimension)
- p = polynomial order

- Minimum number of sample points for function/gradient regression

$$N \geq \left\lceil \frac{(d+p)!}{d!p!(d+1)} \right\rceil$$

- Number of sample points for linear/quadratic gradient enhanced regression

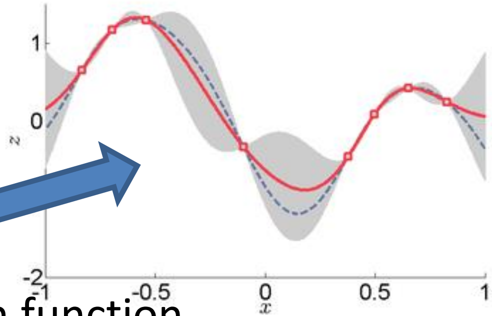
$$N \geq \begin{cases} 1 & \text{for } p = 1, \\ \left\lceil \frac{(d+2)}{2} \right\rceil & \text{for } p = 2. \end{cases}$$

Typically use X2 more points (or more)

Kriging Models

Consider a random process model estimating a function value by a linear combination of function values (training points)

$$\hat{y}_{(\mathbf{x})} = \sum_{i=1}^{n} w_i(x) y_i$$



Modeled as stochastic process with Gaussian correlation function dependent on distance between points

Where y_i represents the training point functional values and $w_i(x)$ are the coefficients determined by solving the problem that minimizes the Variance of the error in the approximation

Kriging, Gradient-enhanced Kriging

Kriging model approach - originally in geological statistics

Two gradient-enhanced Kriging (cokriging or GEK)

- ✓ **Direct Cokriging**

  Gradient information is included in the formulation

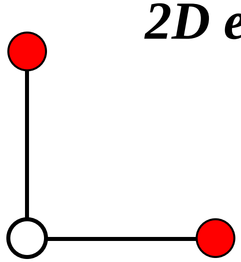
  (correlation between func-grad and grad-grad)

- ✓ **Indirect Cokriging**

  Same formulation as original Kriging

  Additional samples are created by using the gradient info

  Kriging model by both real and additional pts



2D example

○ ***: Real Sample Point***

● ***: Additional Sample Point***

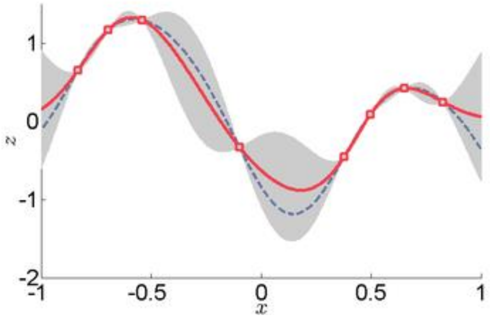
$$\mathbf{x}_{add} = \mathbf{x}_i + \Delta\mathbf{x}$$

$$y_{add} = y_{(\mathbf{x}_i)} + \Delta\mathbf{x}^T \frac{\partial y_{(\mathbf{x}_i)}}{\partial \mathbf{x}}$$

Gradient Enhanced Kriging Models

Consider a random process model estimating a function value
by a linear combination of function and gradient values (training points)

$$\hat{y}_{(\mathbf{x})} = \sum_{i=1}^{n} w_i y_i + \sum_{i=1}^{n'} \lambda_i y_i'$$



Where y'$_i$ represents the training point gradient values and λ_i(x) are the coefficients determined by solving the problem that minimizes the Variance of the error in the approximation

Gradient/Hessian-enhanced Kriging

Indirect Approach

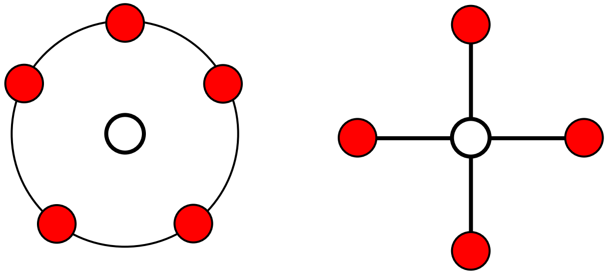
$$\mathbf{x}_{add} = \mathbf{x}_i + \Delta\mathbf{x}$$

$$y_{add} = y_{(\mathbf{x}_i)} + \Delta\mathbf{x}^T\mathbf{G} + \frac{1}{2}\Delta\mathbf{x}^T\mathbf{H}\Delta\mathbf{x}$$

2D example

○ ***: Real Sample Point***

● ***: Additional Sample Point***



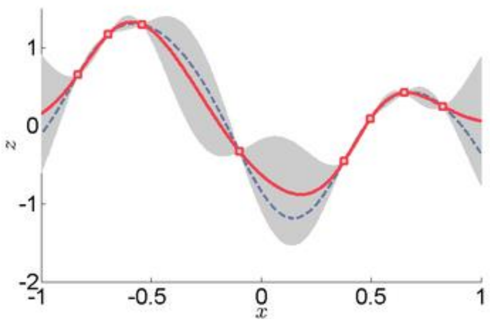
Arrangements to Use Full Hessian / Diagonal Terms

- **Major parameters :**
  - **distance between real / additional pts**
  - **number of additional pts per real pt**
- **Worse matrix conditioning with smaller distance, larger number of additional pts**
- **Severe tradeoffs for these parameters**

Gradient/Hessian Enhanced Kriging Models

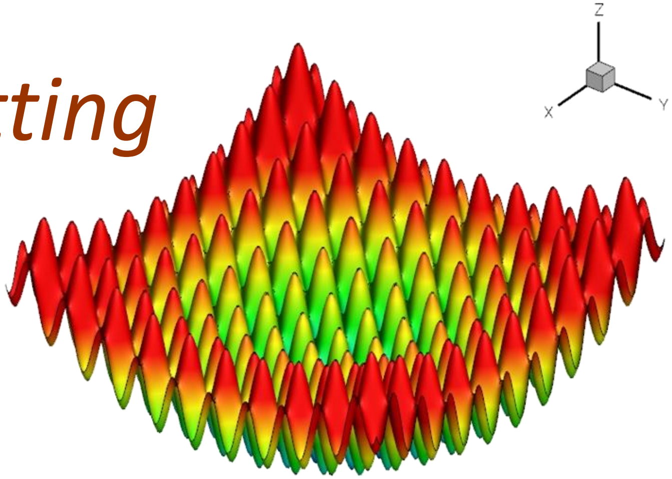
Consider a random process model estimating a function value by a linear combination of function values (training points)

$$\hat{y}_{(\mathbf{x})} = \sum_{i=1}^{n} w_i y_i + \sum_{i=1}^{n'} \lambda_i y_i^{'} + \sum_{i=1}^{n''} \phi_i y_i^{''}$$



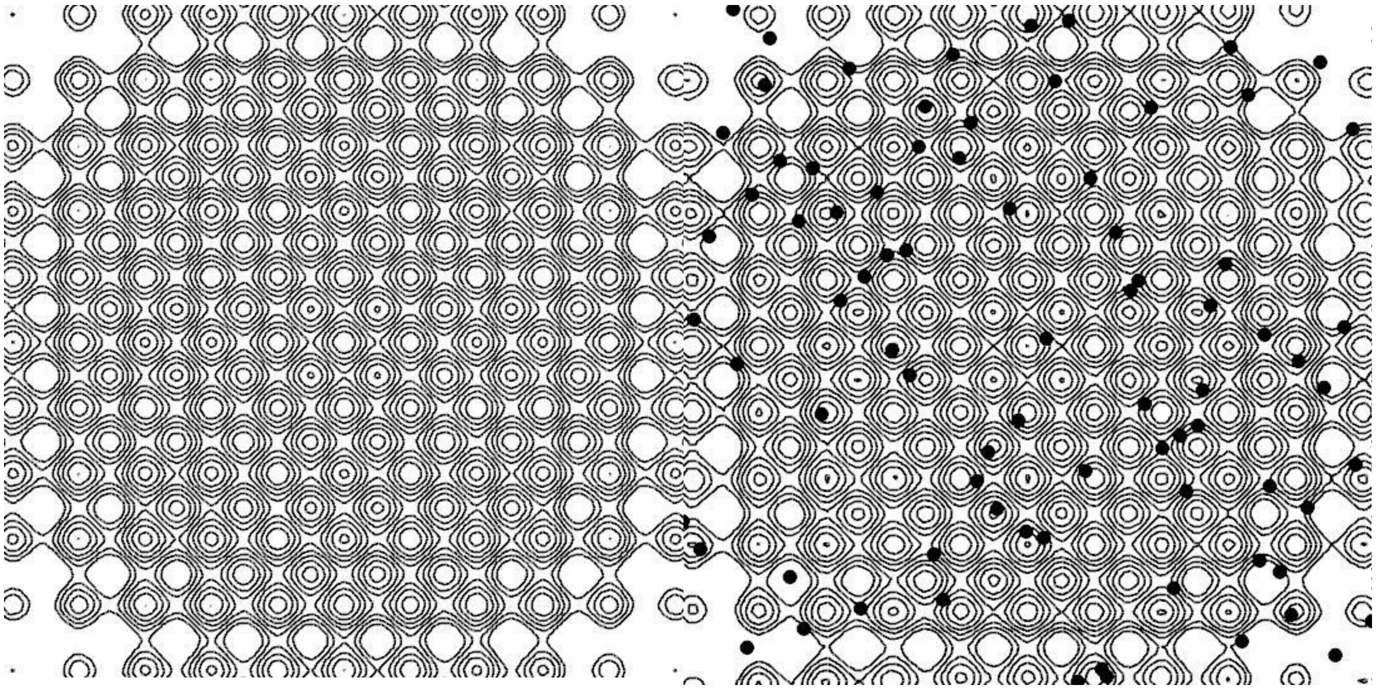
Where y''_i represents the training point Hessian values and $\phi_i(x)$ are the coefficients determined by solving the problem that minimizes the Variance of the error in the approximation

2D Rastrigin Function Fitting



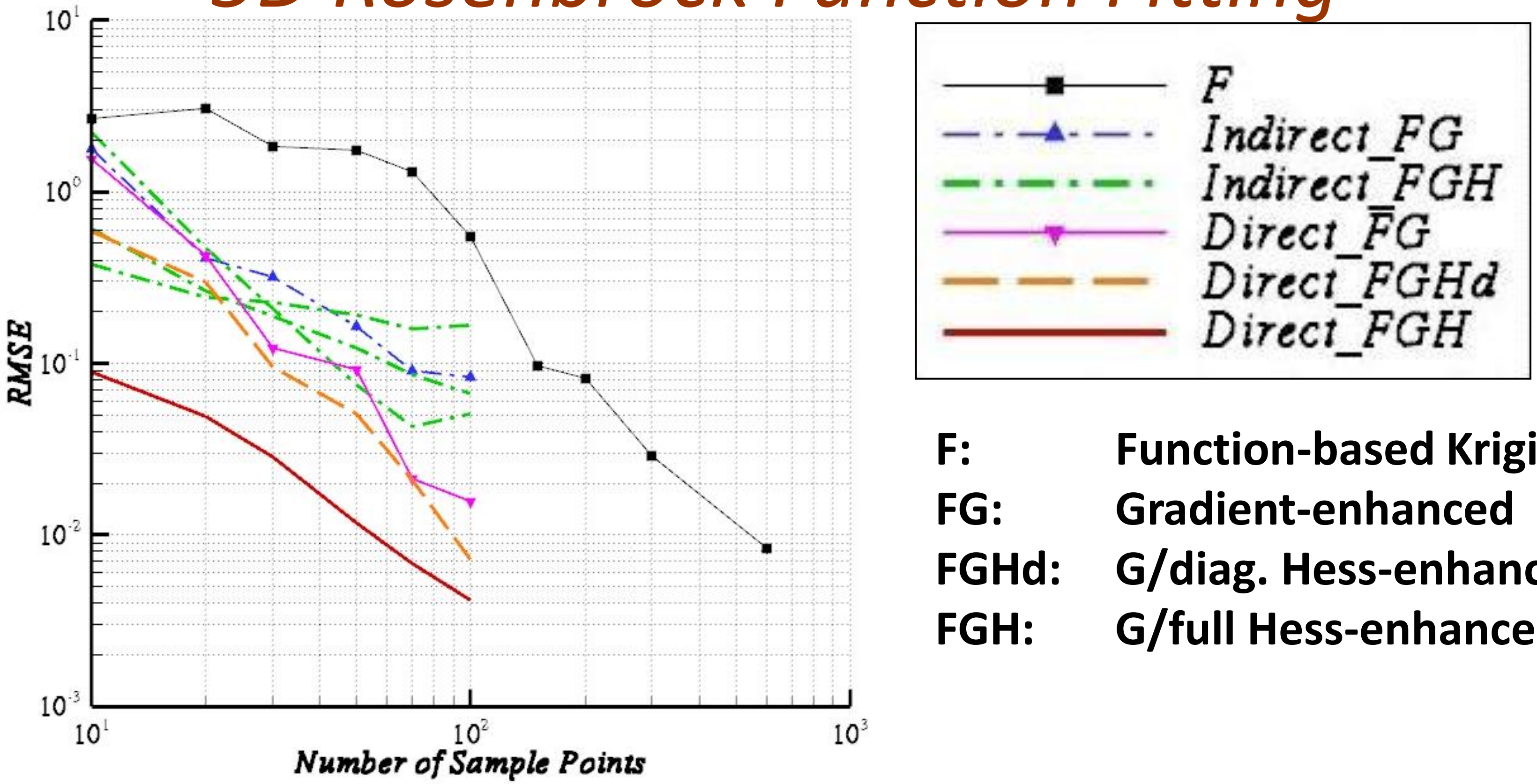
$$y_{(\mathbf{x})} = 20 + x_1^2 + x_2^2 - 10\left(\cos\left(2\pi x_1\right) + \cos\left(2\pi x_2\right)\right)$$

80 samples by Latin Hypercube Sampling
Direct Kriging approach



Exact Rastrigin Function ***Gradient/Hessian-enhanced***

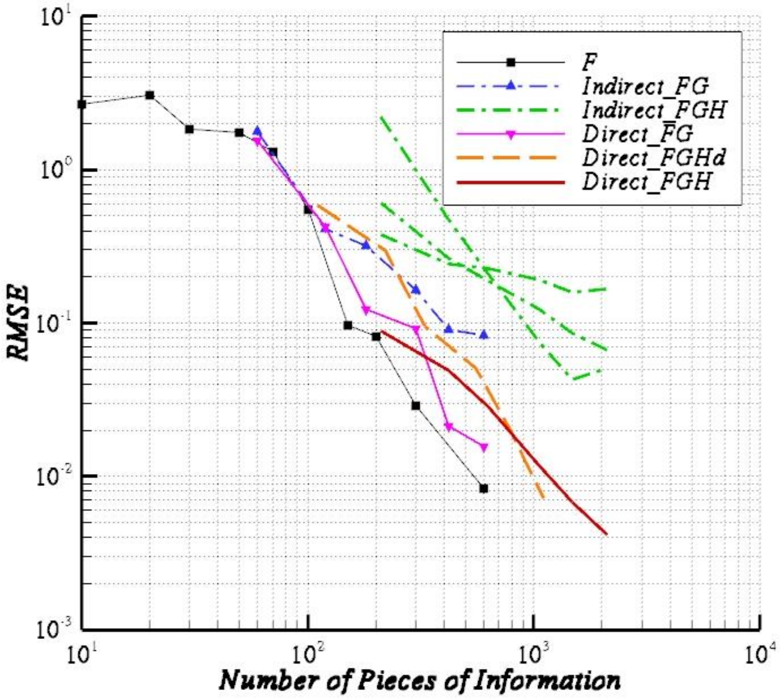
5D Rosenbrock Function Fitting

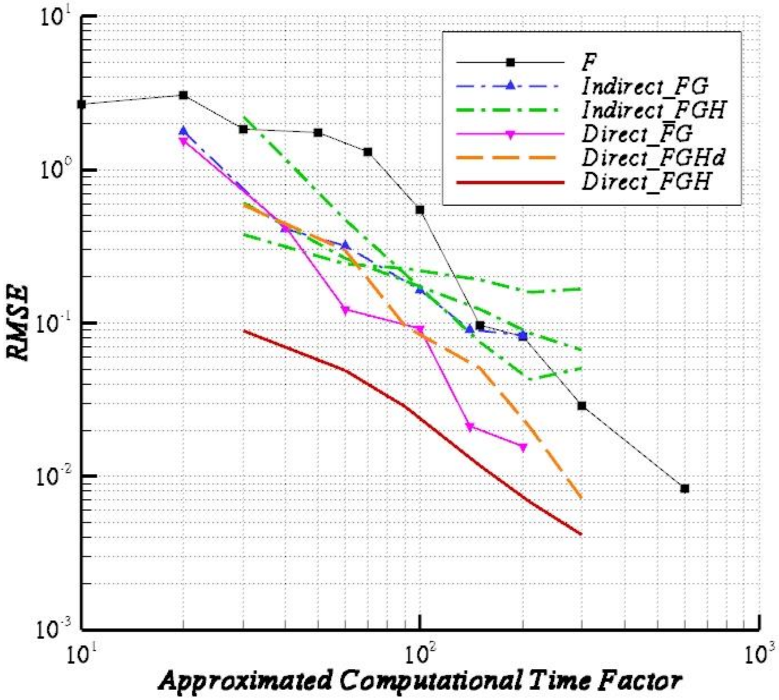


F: **Function-based Kriging**
FG: **Gradient-enhanced**
FGHd: **G/diag. Hess-enhanced**
FGH: **G/full Hess-enhanced**

- **RMSE .vs. Number of sample points**
- **Superiority in direct Kriging approaches due to exact enforcement of derivative information and better conditioning of correlation matrix**

5D Rosenbrock Function Fitting



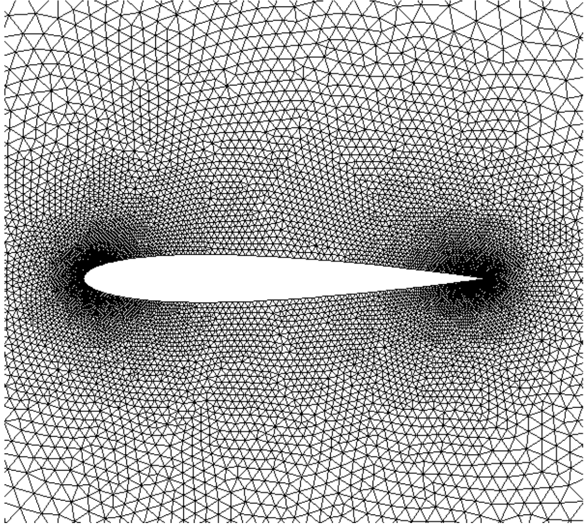


- **# of pieces of information = sum of # of F/G/H net components**
- **To scatter samples is better than concentration at limited samples**
- **Approximated computational time factor**

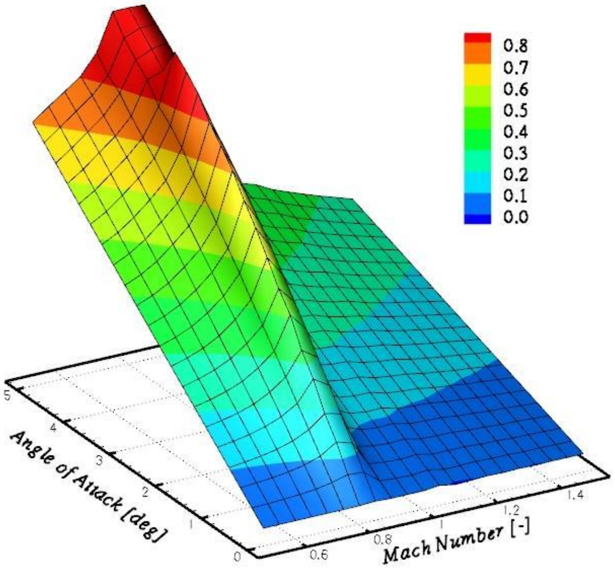
$$TF = \sum_{i=1}^{N_{sample}} T_i, \quad T_i = 1/2/3, \quad if \quad i \quad has \quad F/FG/FGH$$

- **G/H-enhanced surrogate model provides better performance with efficient Gradient/Hessian calculation methods**

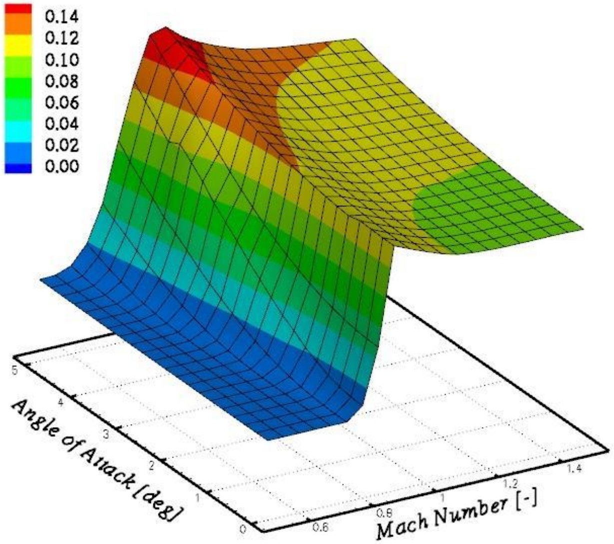
Aerodynamic Data Modeling



- Unstructured mesh CFD
- Steady inviscid flow, NACA0012
- 20,000 triangle elements
- Mach Number [0.5, 1.5]
- Angle of Attack$_{[deg]}$ [0.0, 5.0]
- 21x21=441 validation data

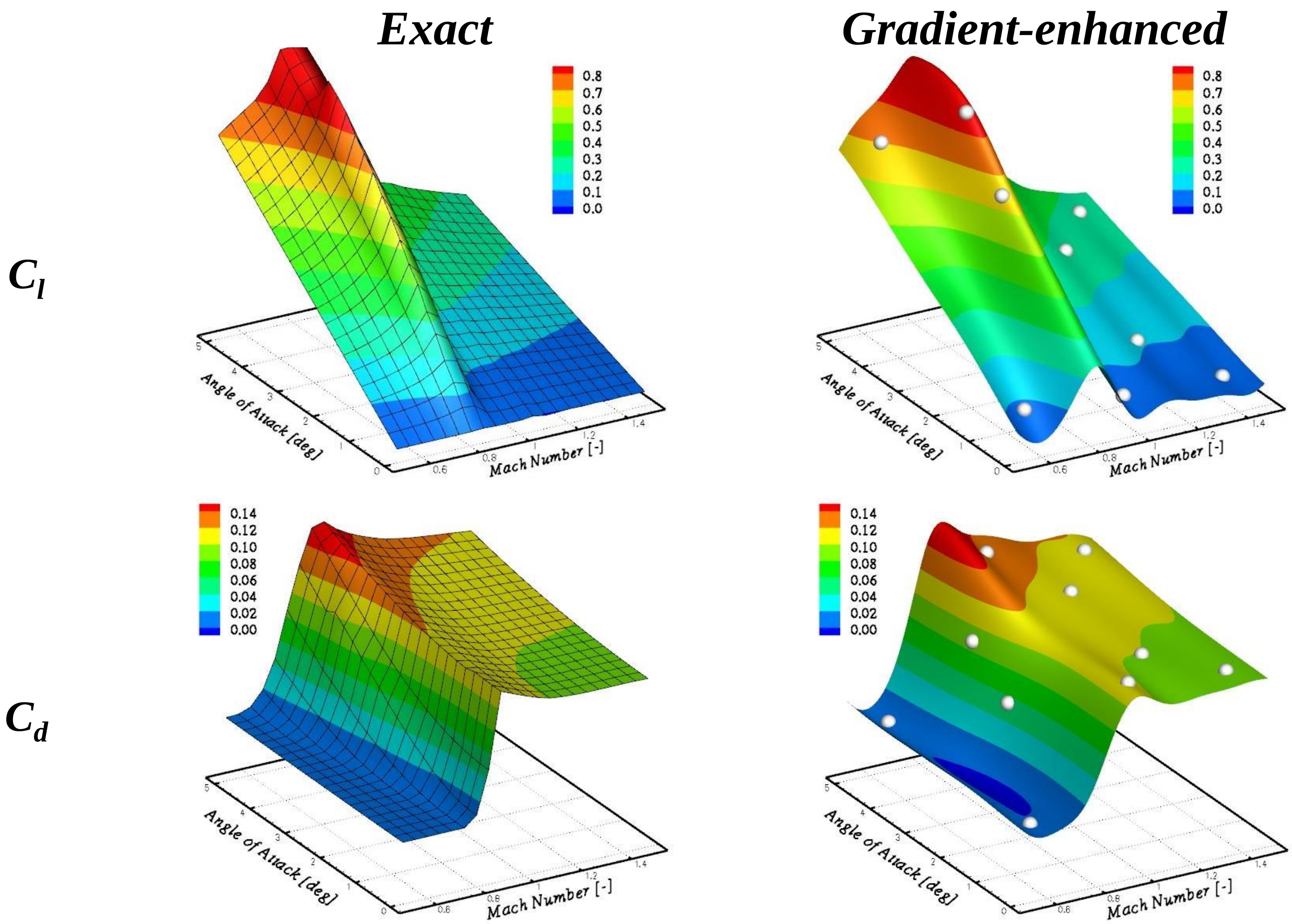


Exact Hypersurface of Lift Coefficient



Exact Hypersurface of Drag Coefficient

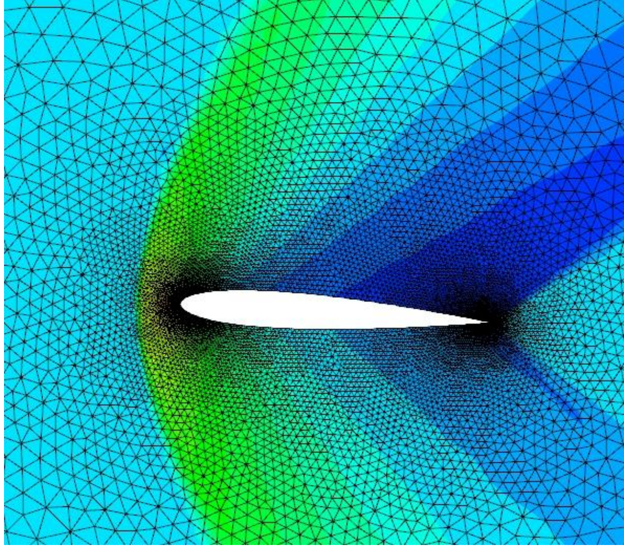
Aerodynamic Data Modeling

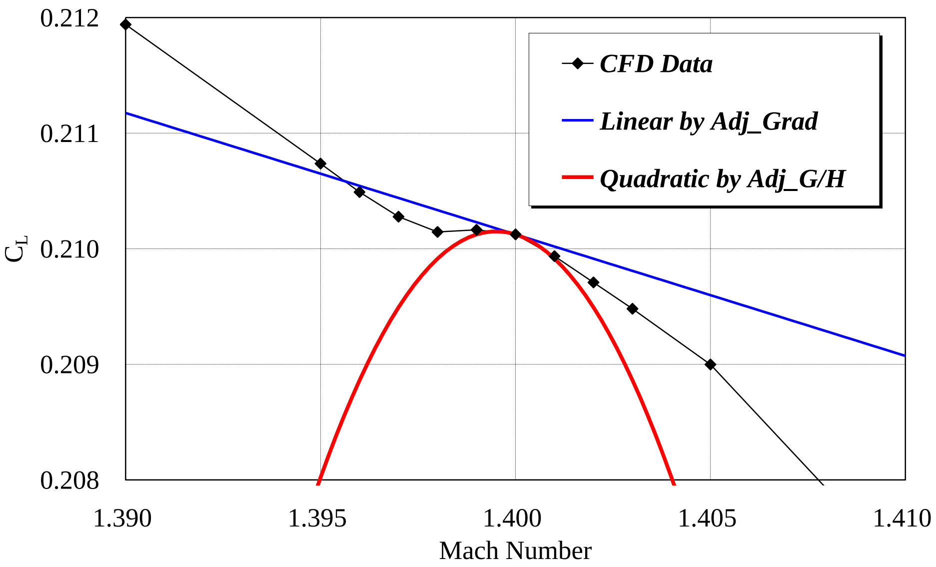


- Adjoint gradient is helpful to construct accurate surrogate model
- CFD Hessian is not helpful due to noisy design space

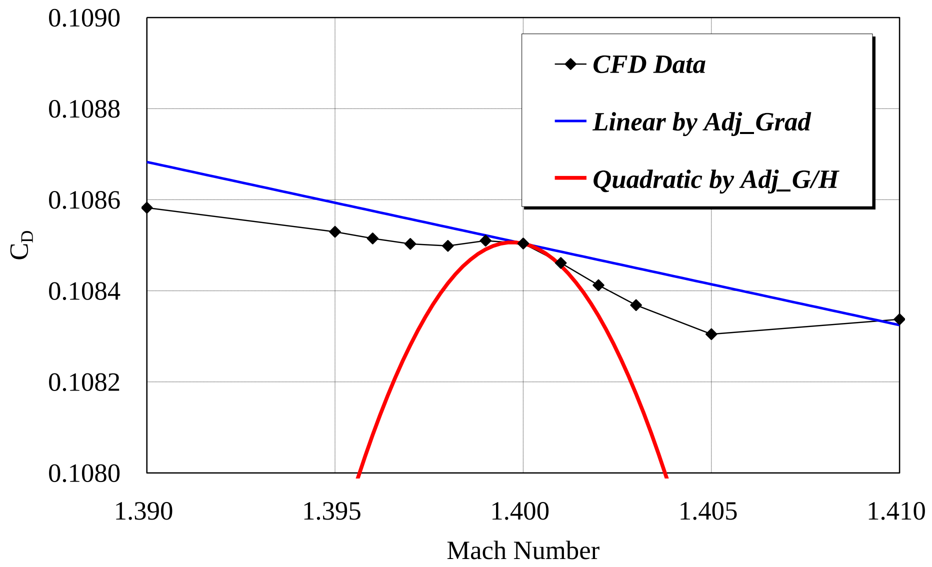
Aerodynamic Data Modeling

- ✓ NACA0012
- ✓ M=1.4
- ✓ AoA=3.5[deg]
- ✓ Noisy in Mach number direction



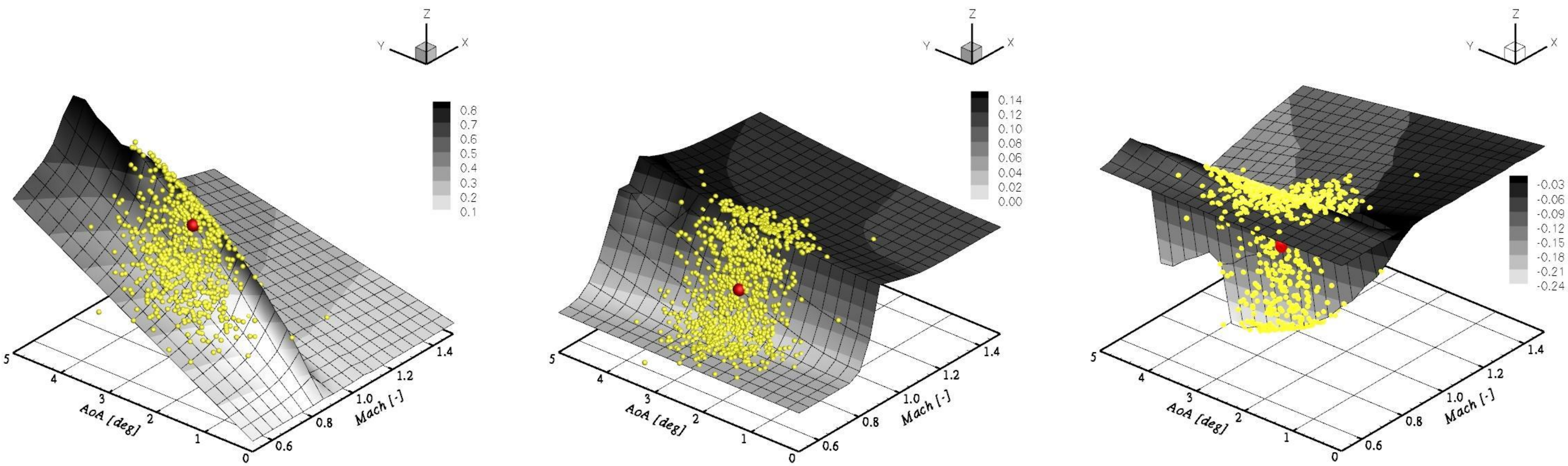


C_l



C_d

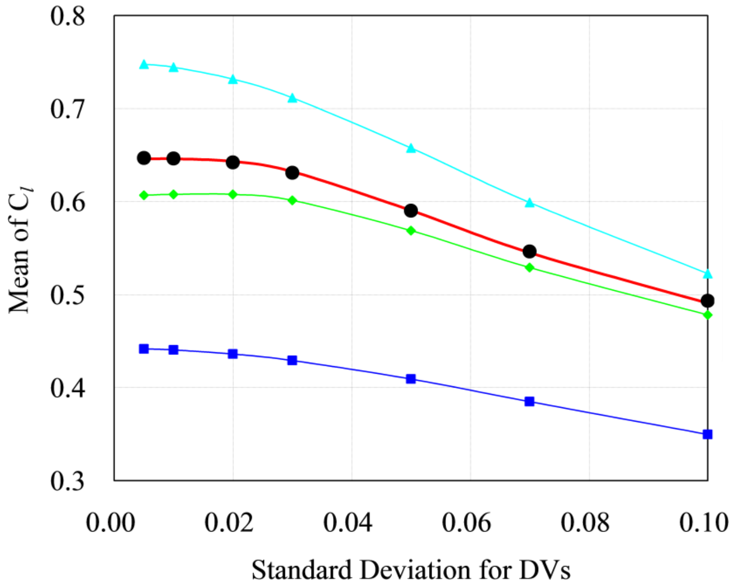
UQ using Kriging Surrogate Model

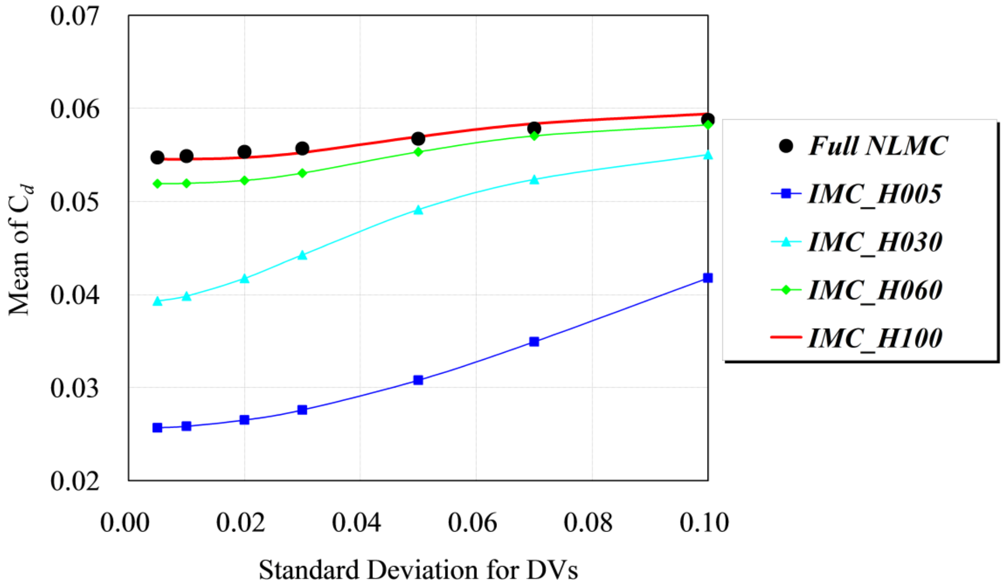


<u>Full-MC results for σ=0.1</u>

- ✓ Uncertainty analysis at M=0.8, AoA=2.5 for both Mach/AoA
- ✓ 1000 CFD evaluations for a specified σ value
- ✓ In total 7000 CFD evaluations (= 1000 x 7) for full-MC

Mach-AoA Hypersurfaces





Mean of C_m

Variance of C_m

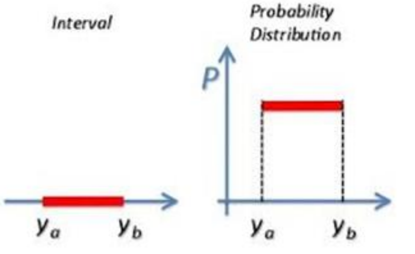
✓ More accurate uncertainty analysis by Inexpensive MC using Kriging model

Epistemic Uncertainty Quantification

- Method of intervals
  - Simply sum intervals for each parameter using constant gradient

$$y_o = f(x_o)$$

$$\Delta_y = \sum_{i=1}^{d} \left| \frac{\partial f}{\partial x_i} \Delta_{x_j} \right|$$



- Bound constrained optimization problem
  - Determine minimum and maximum output values over all possible input values.

$$y_{min} = \min_{x \in I} f(x)$$

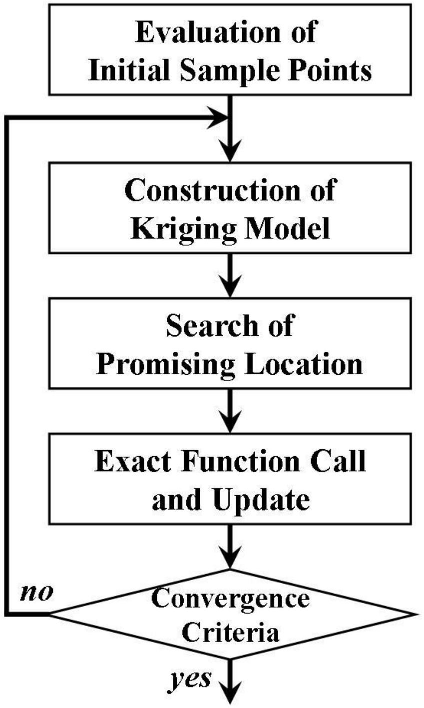
$$y_{max} = \max_{x \in I} f(x)$$

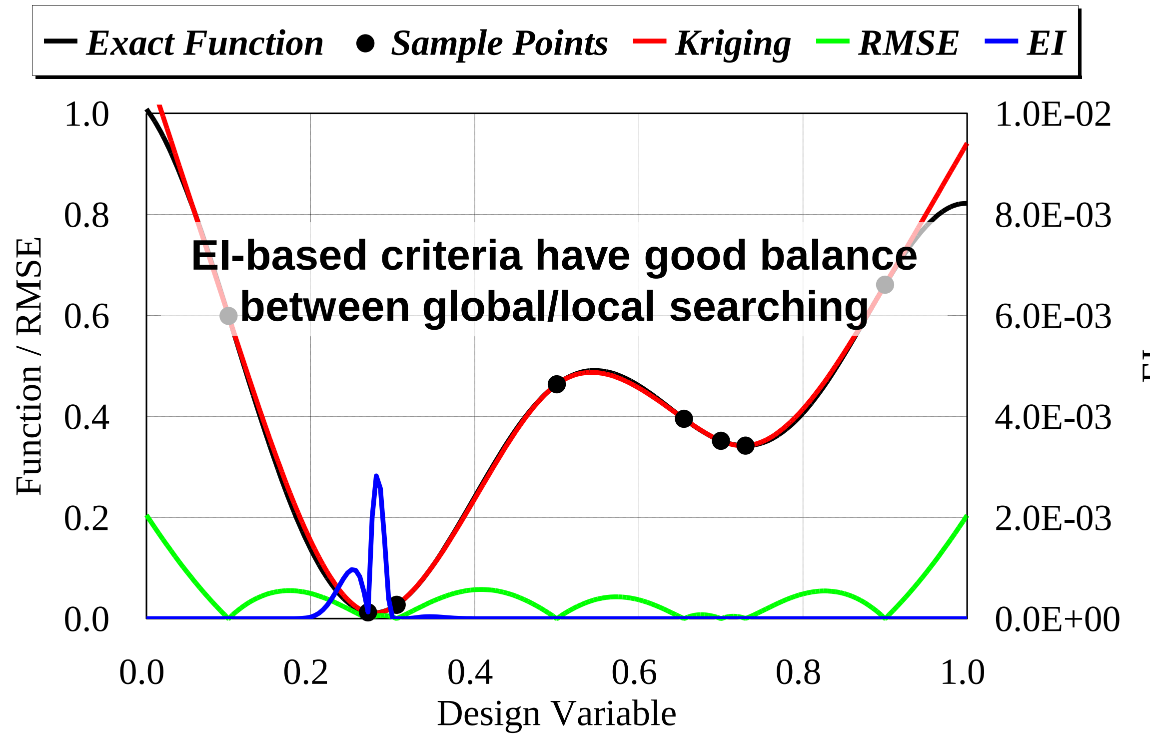
  - Use gradient-based optimization (adjoint enabled)
  - Global optimum required
    - Global optimization/Surrogate models

Kriging Sampling Criteria for Global Optimization

- ✓ How to find promising location on surrogate model ?
- ✓ Expected Improvement (EI) value
- ✓ Potential of being smaller than current minimum (optimal)
- ✓ Consider both estimated function and uncertainty (RMSE)

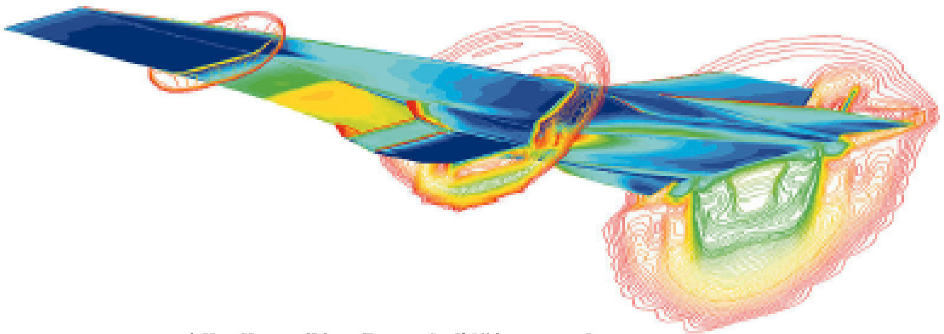
$$EI(\mathbf{x}) = (y_{min} - \hat{y}(\mathbf{x}))\Phi\left(\frac{y_{min} - \hat{y}(\mathbf{x})}{s}\right) + s\phi\left(\frac{y_{min} - \hat{y}(\mathbf{x})}{s}\right) \qquad \left(\frac{\partial EI}{\partial \hat{y}} < 0, \quad \frac{\partial EI}{\partial s} > 0\right)$$





Demonstration: Hypersonic Flow UQ

- Hypersonic Flow roughly defined as $M > 5$
- Characterized by:
  - Strong Shocks
  - Internal Energy Modes (Rotational, Vibrational, Electronic)
  - Chemical Reactions
- Non-equilibrium chemistry requires each species to be modeled
- Thermal non-equilibrium requires individual energy modes to be solved independently
- Models can require hundreds of parameters to define (Arrhenius Reaction Coefficients, Curve fits, etc.)



http://en.wikipedia.org/wiki/Hypersonic

Real Gas Model

- Five Species, Two Temperature Real Gas Model for Air
  - Accounts for Molecular dissociation: N_2, O_2, N, O, NO
  - Energy described by translation-rotational temperature and vibrational-electronic temperature
- Compressible Navier Stokes Equations:

$$\frac{\partial \rho_s}{\partial t} + \nabla \cdot (\rho_s \vec{U}) = -\nabla \cdot (\rho_s \vec{V}_s) + \omega_s$$

$$\frac{\partial \rho U}{\partial t} + \nabla \cdot (\rho \vec{U} \otimes \vec{U}) = -\nabla P + \nabla \cdot \underline{\tau}$$

$$\frac{\partial \rho e_t}{\partial t} + \nabla \cdot (\rho h_t U) = \nabla \cdot (\underline{\tau} \vec{u}) - \nabla \cdot \vec{q} - \nabla \cdot \vec{q}_v - \nabla \cdot \left(\sum_s h_{t,s} \rho_s \tilde{V}_s \right)$$

$$\frac{\partial \rho e_v}{\partial t} + \nabla \cdot (\rho e_v U) = Q_{T-V} + \sum_s e_{v,s} \omega_s - \nabla \cdot \left(\sum_s h_{v,s} \rho_s \tilde{V}_s \right) - \nabla \cdot \vec{q}_v$$

Real Gas Model

- Constitutive Law's:

$$\rho_s \tilde{V}_s = -\rho D_s \nabla c_s \qquad \text{Fick's Law}$$

$$\underline{\tau} = \mu(\nabla \vec{u} + \vec{u}\nabla) - \frac{2}{3}\mu \nabla \cdot \vec{u}\underline{I} \qquad \text{Newtonian Fluid}$$

$$\vec{q} = -k \nabla T \qquad \text{Fourier's Law}$$

$$\vec{q_v} = -k_v \nabla T_v$$

- Equations of State:

$$\frac{C_v^s(T)M_s}{\bar{R}} = A_{o,s}^i + A_{1,s}^i T + A_{2,s}^i T^2 + A_{3,s}^i T^3 + A_{4,s}^i T^4 \qquad \text{(Caloric)}$$

$$P(\rho, T) = \rho \sum_s c_s \frac{\bar{R}}{M_s} T \qquad \text{(Thermal)}$$

Real Gas: Transport Model Parameters

- Defines: $\mu = \mu(T, \rho_s)$, $k = k(T, \rho_s)$, $k_v = k_v(T, \rho_s)$, $D_s = D_s(T, \rho_s)$
- Calculated using Collision integrals (cross-sections) for each interaction $\Omega_{s,r}^{k,k}$
- Specified at 2000 K and 4000 K and interpolated using:

$$log_{10}(\Omega_{s,r}^{k,k}) = log_{10}(\Omega_{s,r}^{k,k})_{2000} + \left[log_{10}(\Omega_{s,r}^{k,k})_{4000} - log_{10}(\Omega_{s,r}^{k,k})_{2000} \right] \frac{ln(T) - ln(2000)}{ln(4000) - ln(2000)}$$

- 15 interactions possible giving 60 total model parameters
- Effect of curve shifts accounted for using parameter $A_{s,r}^{k}$:

30 parameters assuming values at 2000K and 4000K are correlated

$$\Omega_{s,r}^{k,k}(T) = A_{s,r}^{k} \hat{\Omega}_{s,r}^{k,k}(T)$$

Real Gas: Reaction Model Parameters

- Net creation/destruction of each species ω_s:

$$\omega_s = M_s \sum_r (\beta_{s,r} - \alpha_{s,r})(R_{f,r} - R_{b,r})$$

- Reaction Rates specified using Law of Mass Action:

$$R_{f,r} = 1000 \left[k_{f,r} \prod_s (0.001\rho_s / M_s)^{\alpha_{s,r}} \right]$$

- Rate Coefficients $k_{f,r}$ and $k_{b,r}$ given by Arrhenius relation (Dunn-Kang Model)

$$k_{f,r} = C_{f,r} T_a^{\eta_{f,r}} e^{-\frac{E_{f,r}}{k_B T_a}} \qquad k_{b,r} = C_{b,r} T_a^{\eta_{b,r}} e^{-\frac{E_{b,r}}{k_B T_a}}$$

- 17 reactions total, 34 parameters:

Flow Solver

- Equations solved numerically in two dimensions using in-house developed finite-volume solver
- Capable of solving on unstructured triangles/quadrilaterals
- Solution marched to steady state using implicit pseudo-time stepping

$$\mathbf{J}(\mathbf{U}^n, \mathbf{U}^{n-1}) = \frac{\mathbf{U}^n - \mathbf{U}^{n-1}}{\Delta t} + \mathbf{R}(\mathbf{U}^n)$$

- Newton's Method used to solve nonlinear equation at each time-step:

$$\delta \mathbf{U}^k = -\left[P\right]^{-1} \mathbf{J}(\mathbf{U}^k, \mathbf{U}^{n-1})$$
$$\mathbf{U}^{k+1} = \mathbf{U}^k + \lambda \delta \mathbf{U}^k$$

- Jacobi or line-preconditioned GMRES used to invert Jacobian

Flow Solver

- Gradient reconstruction of primitives
- Green-Gauss contour integration used to calculate gradients
- Smooth Van Albada Limiter with Pressure Switch used:

$$\Psi_k = max(0, 1 - K\nu_k)\frac{1}{\Delta^-}\frac{(\Delta^{+^2} + \varepsilon^2)\Delta^- + 2\Delta^{-^2}\Delta^+}{\Delta^{+^2} + 2\Delta^- + \Delta^-\Delta^+ + \varepsilon^2}$$

$$\nu_i = \frac{\sum_k |P_R - P_L|}{\sum_k P_R + P_L}$$

- Face based Gradients calculated using averaging and correction term:

$$\nabla \mathbf{V}_k = \tilde{\nabla} \mathbf{V} + \frac{\mathbf{V}_R - \mathbf{V}_L - \tilde{\nabla}\mathbf{V} \cdot \Delta \vec{T}}{|\Delta \vec{T}|}\frac{\Delta \vec{T}}{|\Delta \vec{T}|}$$

- Inviscid Flux Calculated Using AUSM+UP flux function with Frozen Speed of Sound

Code Validation

- 5 km/s cylinder test case
- Fixed Wall temperature
- Super-catalytic Wall
- Results compared with LAURA (Same Mesh)
- Park Chemical Kinetics Model

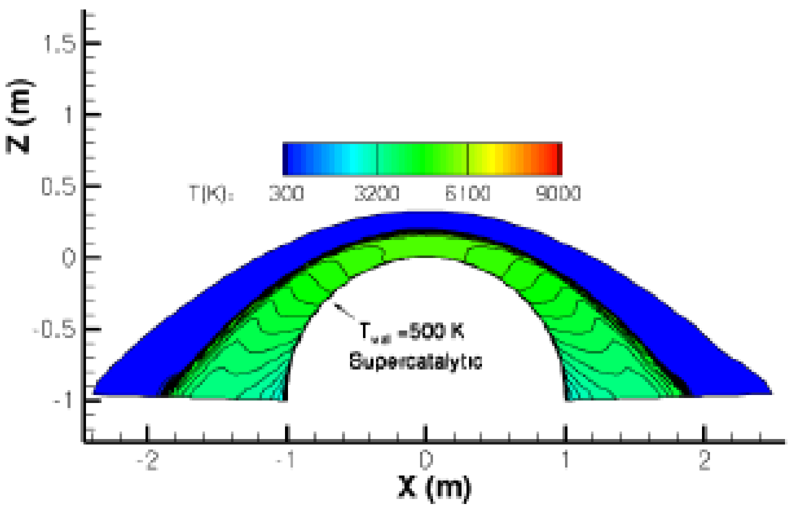


Table: Benchmark Flow Conditions

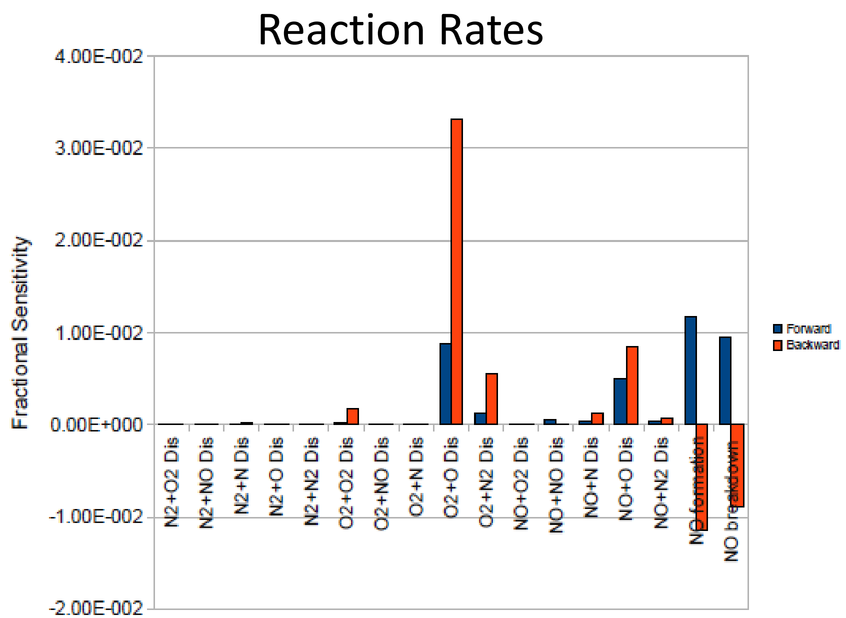
V_∞	=	5 km/s
ρ_∞	=	0.001 kg/m^3
T_∞	=	200 K
T_{wall}	=	500 K
M_∞	=	17.605
Re_∞	=	753,860
Pr_∞	=	0.72

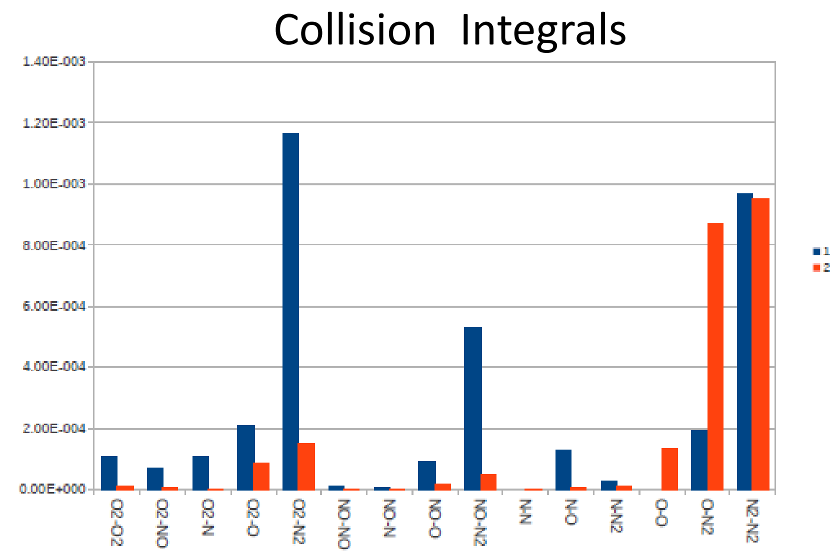
Objective Formulation

- Integrated surface heating objective

$$L = -\frac{\int_{\partial\Omega} k \nabla T \cdot \vec{n} + k_v \nabla T_v \cdot \vec{n} dA}{\frac{1}{2}\rho_\infty V_\infty^3}$$

Local Sensitivity Analysis





- Using adjoint formulation, obtain derivative of surface heating w.r.t. uncertain input parameters
- Classify from most sensitive to least sensitive
- Use to focus on most sensitive parameters

Comparison with Global Sensitivity Analysis

- Local analysis gives effect to infinitesimal change in parameters
- Does not account for interference effects or large perturbations
- Global sensitivity analysis gives average effect over design space
- Calculated via Monte Carlo sampling (6,331 for this case)

$$r_i = \frac{cov(D_i, y)}{\sigma_{D_i}\sigma_y}$$

- Design space given by the uncertainty space of 66 parameters: (Assumed normal distribution)

Number	Variable	Mean	Standard Deviations
1	$\rho_\infty \ (kg/m^3)$	1×10^{-3}	5%
2	$V_\infty(m/s)$	5000	15.42
3-17	A^1_{s-r}	1	5%
18-32	A^2_{s-r}	1	5%
33-49	ξ_f	0	0.25
50-66	ξ_b	0	0.25

Global vs Local Sensitivity Analysis

- Importance ranking and contribution to variance compared
- Variance contribution given by square of correlation coefficient
- Local and Global show significant disagreement

Rank	Variable	Local	Global	Local
1	ρ_∞	1	0.60055	0.43230
2	$O_2 + O \leftrightarrows 2O + O$ (f)	2	1.0610×10^{-1}	1.7490×10^{-1}
3	$NO + O \leftrightarrows N + 2O$ (b)	3	5.1914×10^{-2}	7.7560×10^{-2}
4	O2-N2 (k=1)	7	4.2121×10^{-2}	2.4524×10^{-2}
5	N2-N2 (k=1)	10	3.1617×10^{-2}	1.6956×10^{-2}
6	$O_2 + O_2 \leftrightarrows 2O + O_2$ (b)	13	2.1621×10^{-2}	1.3120×10^{-2}
7	$N_2 + O \leftrightarrows NO + N$ (f)	4	2.0647×10^{-2}	7.2017×10^{-2}
8	N2-N2 (k=2)	11	1.9019×10^{-2}	1.6354×10^{-2}
9	O-N2 (k=2)	12	1.3874×10^{-2}	1.3714×10^{-2}
10	$N_2 + O \leftrightarrows NO + N$ (b)	5	1.2155×10^{-2}	6.8076×10^{-2}

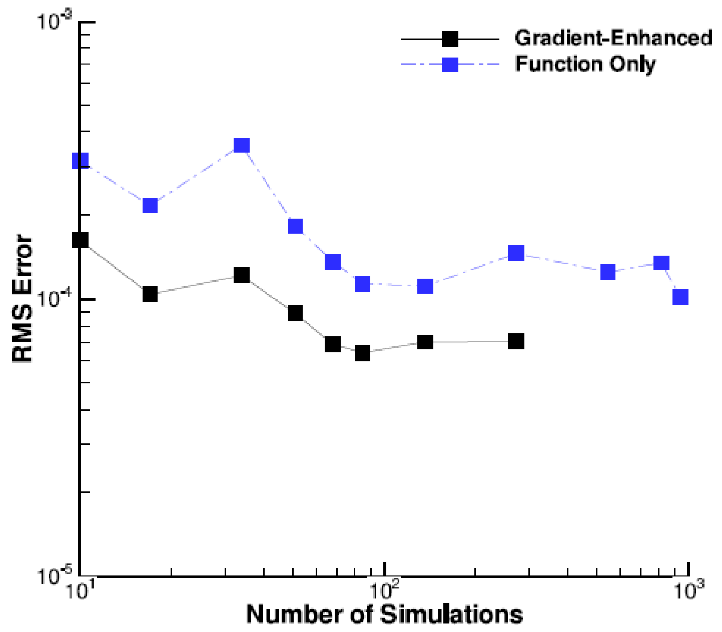
Gradient-Enhanced Regression based Global Sensitivity Analysis

- Global Sensitivity using 68 function/gradients
- Hermite Polynomial basis with maximum order 2
- Correlation calculated by sampling from regression
- Better agreement in terms of ranking and contribution
- Used for dimension reduction for uncertainty quantification

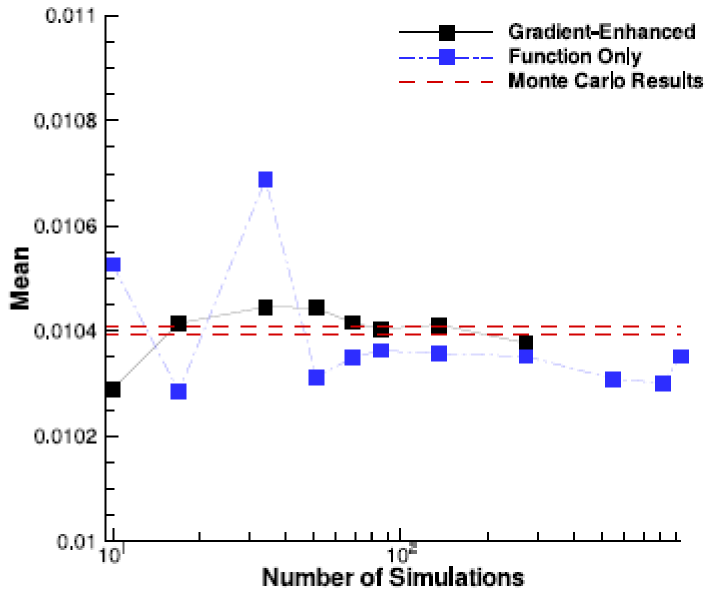
Rank	Variable	Global	Regression	Global
1	ρ_∞	1	0.56879	0.60055
2	$O_2 + O \leftrightarrows 2O + O$ (f)	2	1.0002×10^{-1}	1.0610×10^{-1}
3	$O_2 + O_2 \leftrightarrows 2O + O_2$ (b)	6	5.7669×10^{-2}	2.1621×10^{-2}
4	$NO + O \leftrightarrows N + O + O$ (b)	3	4.0057×10^{-1}	5.1914×10^{-2}
5	N2-N2 (k=1)	5	3.7461×10^{-2}	3.1617×10^{-2}
6	O2-N2 (k=1)	4	3.3299×10^{-2}	4.2121×10^{-2}
7	N2-N2 (k=2)	8	2.1163×10^{-2}	1.9019×10^{-2}
8	O-N2 (k=2)	9	1.7395×10^{-2}	1.3874×10^{-2}
9	V_∞	14	1.3497×10^{-2}	4.8401×10^{-3}
10	$O_2 + O \leftrightarrows 2O + O$ (b)	13	1.1734×10^{-2}	7.4280×10^{-3}

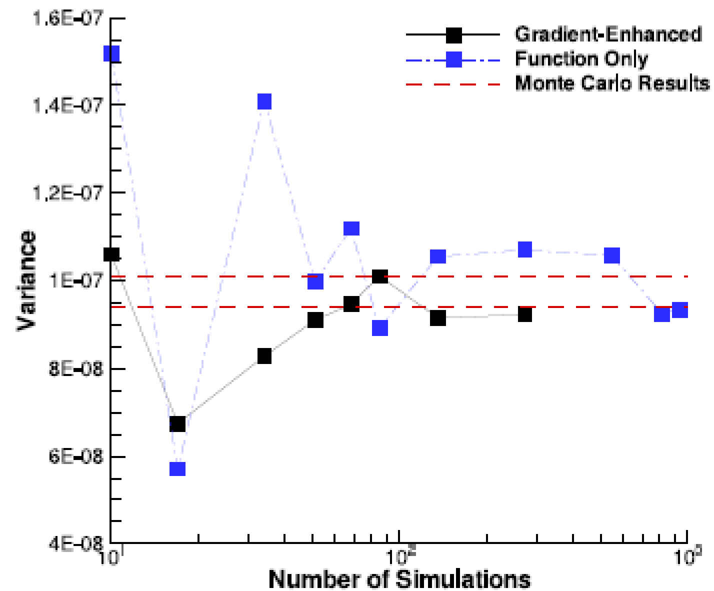
Aleatory Uncertainty using Kriging Model

- Reduce Kriging dimension by using only 15 most "sensitive" parameters as determined from global sensitivity analysis
- Build Function only and Function/Gradient Kriging models
  - Use different number of training points for Kriging model
  - Examine error wrt full Monte Carlo sampling (6331 CFD runs)
- Gradient-enhanced Kriging produces consistently lower error
- Lower limit on error due to omitted parameters
- 68 sample points adequate for Kriging



Aleatory Uncertainty using Kriging Model

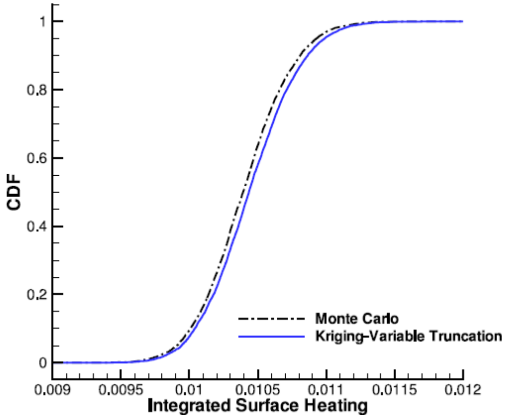




- Gradient-enhanced Kriging superior predictions of mean and variance
- Close to full Monte Carlo results (6331 pts) with only 68 Kriging pts
  - 68 CFD analyses
  - 68 adjoint solutions for gradient-enhanced model

Aleatory Uncertainty Quantification

- Methods compared based on cost and statistic predictior
- Kriging Methods give most accurate results
- Significant Cost reduction possible (6331 f vs. 68 f/g)



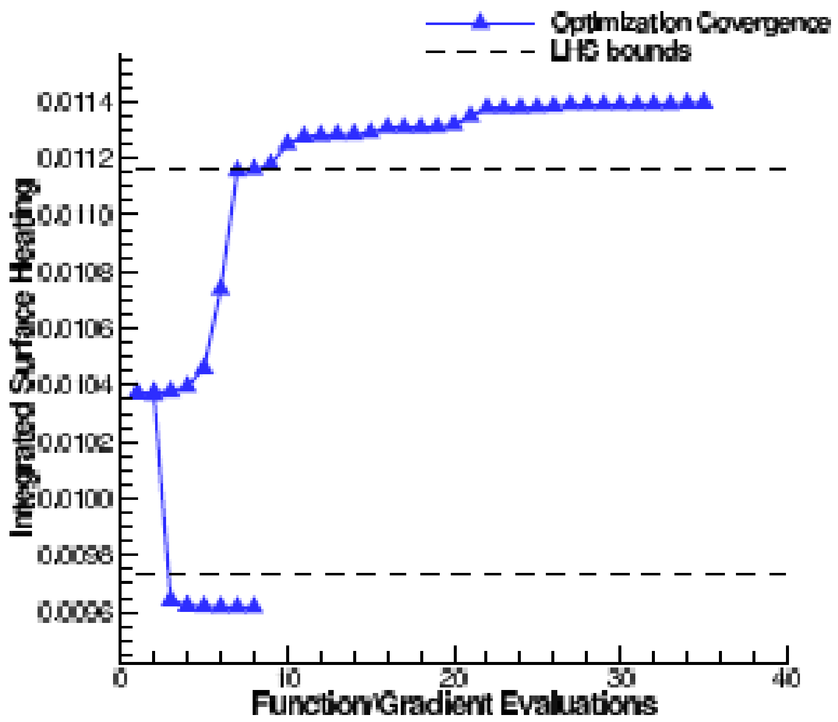
Method	Mean	Variance	95% CI	F/G Cost
Moment Method	1.0370E-002	1.3790E-007	±7.1616%	1
Linear Extrapolation	1.0369E-002	1.3412E-007	±7.0638%	1
P=1 Regression	1.0497E-002	8.8273E-008	±5.6610%	10
P=2 Regression	1.0370E-002	8.6692E-008	±5.6786%	68
Kriging-Trunc.-17D	1.0446E-002	1.0227E-007	±6.1228%	68
Kriging-Reg.-17D	1.0384E-002	9.2394E-008	±5.8543%	68
Monte Carlo-L	**1.0393E-002**	**9.3979E-008**	**±5.8994%**	**6331**

Epistemic Uncertainty Quantification

- Collision integrals treated as epistemic (20% interval width)
- Methods tested using 8 uncertain parameters
- Validated using LHS with 3 points per dimension (6,561 samples)
- Linear (1 f/g) and optimization ($\sim$ 40 f/g) produce more accurate interval

	Linear Method	LHS interval	Optimization
Center	1.0370E-002	1.0449E-002	1.0506E-002
Interval Half Width	8.6634E-004	7.1266E-004	8.8912E-004
Upper	1.1237E-002	1.1161E-002	1.1395E-002
Lower	9.5040E-003	9.7361E-003	9.6168E-003
Percentage	8.35%	6.82%	8.46%

Epistemic Interval using Gradient-Based (Bound) Optimization

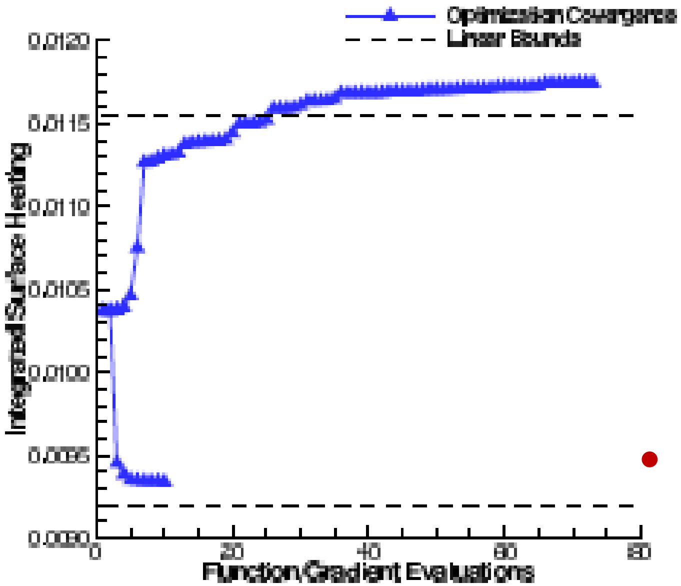


- Optimization more correct result as it satisfies problem statement
- More extensive sampling gives bounds approaching optimization

Epistemic Interval using Gradient-Based (Bound) Optimization

- Optimization/Linear analysis can be applied to large dimension
- Number of parameters expanded to all collision integrals (30 total)
- Methods produce similar interval estimates

LHS sampling unfeasible for 30 parameters >10^{14}



	Linear Method	Optimization
Center	1.0370E-002	1.0543E-002
Half Width	1.1787E-003	1.2031E-003
Upper	1.1549E-002	1.1746E-002
Lower	9.1916E-003	9.3400E-003
Percentage	11.37%	11.41%

- **Linear results give confidence that optimization not stuck in local optimum**
- **May not be true in general case**

Mixed Aleatory-Epistemic

- Variables have either aleatory or epistemic uncertainty
- **Goal**: Determine range containing output with specified probability (P-Box) and separate the contribution from each source
- Typical situation for simulation as complete knowledge rare
- Nested sampling traditionally used; however,
  - For hypersonic flows, number of epistemic variables much greater than number of aleatory variables
  - Expensive of nested sampling increases rapidly with number of epistemic variables
  - Prohibitively expensive for all but explicit functions
- Combine surrogate approaches with gradient-based optimization for rapid mixed UQ

Mixed Aleatory-Epistemic

Define:

- α are aleatory variables
- β are epistemic variables
- $L(\alpha, \beta)$ is simulation output

Nested Sampling:

- Extract β realization for $i = 1, N_r$
    - Sample over α for $j = 1, N_s$
        - Run simulation
        - Compute $L(\alpha, \beta)$
    - Characterize output distribution associated with varying α
- Examine statistics over all realizations (determine worst-case)

Mixed Aleatory-Epistemic

- Nested sampling can be performed inexpensively based on surrogate
- Optimization/Surrogate should scale to higher dimension for large number of epistemic variables
- Two choices for ordering
  - Use optimization to determine min/max of statistic
  - **Use sampling to determine statistic of min/max**
- **Statistics-of-Intervals**
  - Solve multiple optimization problems for different α samples:

$$L_{min}(\alpha) = \min_{\beta} L(\alpha, \beta)$$

$$L_{max}(\alpha) = \max_{\beta} L(\alpha, \beta)$$

  - Construct surrogate (Kriging model) for $L_{min}(\alpha)$ and $L_{max}(\alpha)$
  - Calculate statistics based on sampling over α from surrogate model

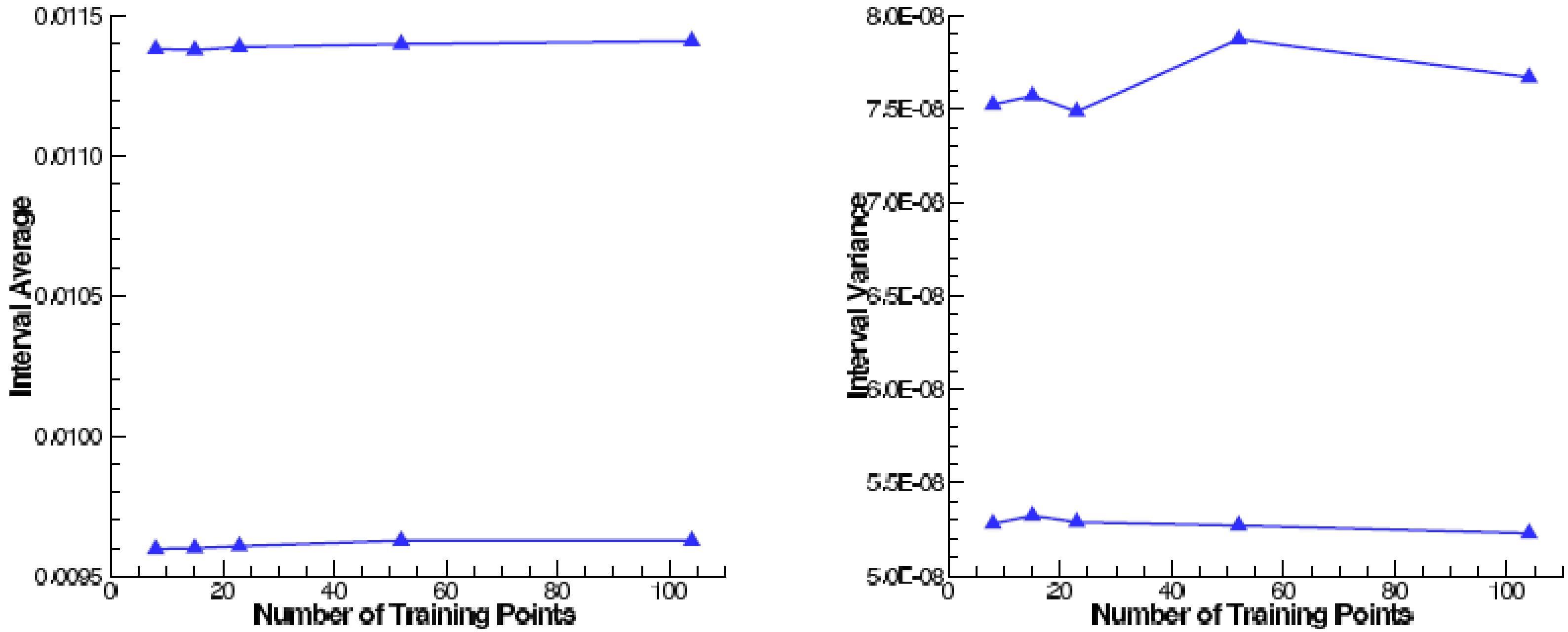
Mixed Aleatory-Epistemic

Uncertain Parameters:

Variable	Type	Uncertainty
ρ_∞ (kg/m^3)	Aleatory	$\pm 10\%$ $(\sigma = 5\%)$
$V_\infty (m/s)$	Aleatory	± 30.84 $(\sigma = 15.42)$
$\Omega^{1,1}_{N2-N2}, \Omega^{2,2}_{N2-N2}$	Epistemic	$\pm 20\%$
$\Omega^{1,1}_{N2-N}, \Omega^{2,2}_{N2-N}$	Epistemic	$\pm 20\%$
$\Omega^{1,1}_{N2-O}, \Omega^{2,2}_{N2-O}$	Epistemic	$\pm 20\%$
$\Omega^{1,1}_{N2-O2}, \Omega^{2,2}_{N2-O2}$	Epistemic	$\pm 20\%$

- 10 total uncertain parameters (2 aleatory, 8 epistemic)
- Nested Sampling used for Validation
- 3 samples per dimension for epistemic variables (6,561 total)
- 5000 samples used for aleatory variables

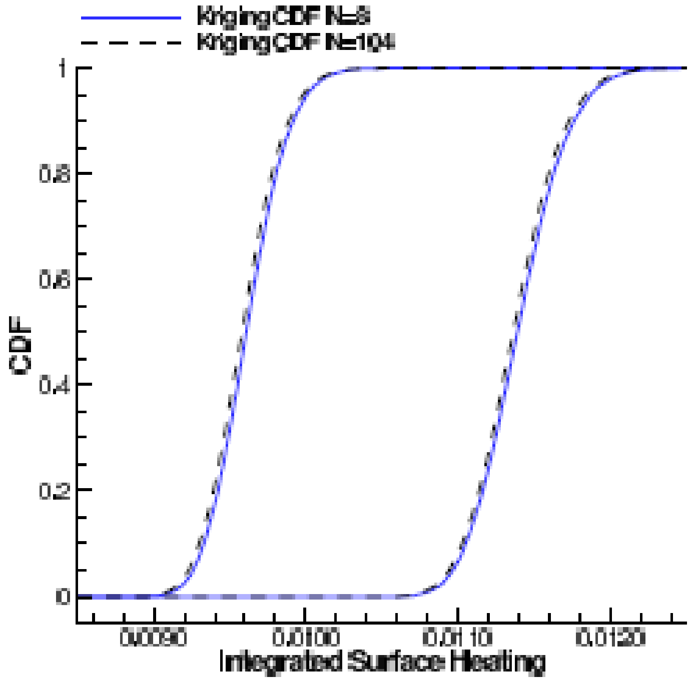
Mixed Aleatory-Epistemic



- Convergence of output interval (min-max) with increasing number of Kriging points

Mixed Aleatory-Epistemic

- CDF for bounds can be created from Kriging Model
- CDF created with Kriging model based on 8 ($\sim$ 500 f/g) and 104 (6176 f/g) pairs of optimizations
- CDF curves virtually identical, implying convergence of Kriging predictions



Mixed Aleatory-Epistemic

- Multiple Optimizations used to approximate combined results
- Kriging model constructed for min and max values
- Monte Carlo performed on Kriging surrogate
- 99*th* percentile of Min/Max predicted

Training Data Size	F/G Evaluations	99^{th} percentile of Min	99^{th} percentile of Max
8	~ 500	1.017556×10^{-2}	1.206949×10^{-2}
15	~ 900	1.016681×10^{-2}	1.207132×10^{-2}
23	~ 1400	1.018928×10^{-2}	1.207939×10^{-2}
52	~ 3000	1.020232×10^{-2}	1.210513×10^{-2}
104	6176	1.020243×10^{-2}	1.210416×10^{-2}

- Statistic converges with handful of optimization results
- SOI method allows mixed UQ when nested strategy prohibitively expensive

Conclusions and Future Work

- Adjoint methods are enabling for Sensitivity Analysis and Uncertainty Quantification
  - Particularly for cases involving one or few objectives
  - Provide entire gradient wrt all uncertain parameters for cost of single adjoint problem
- Demonstrated applications
  - Method of moments
  - Local sensitivity analysis
  - Enhanced surrogate models
    - Polynomial regression
    - Kriging models
  - Gradient-based optimization for epistemic uncertainties
  - Mixed aleatory-epistemic uncertainties

Conclusions and Future Work

- Monte Carlo sampling unfeasible in many cases
  - Mixed aleatory-epistemic uncertainties
- Hessian information can be useful for:
  - Method of moments
  - Further enhanced surrogate models
  - Newton optimization
- Hessian cost must be evaluated vs more global information and effects of non-smooth functionals