Lecture 3: Multi-Fidelity Uncertainty Quantification

Markus Peer Rumpfkeil

University of Dayton

3rd AIAA Workshop on MF Methods for Design and UQ

June 7, 2026

Outline

Uncertainty Quantification I

- Uncertainty quantification (UQ) is important since computations typically assume perfect knowledge of all parameters. In reality there is much uncertainty due to
  - manufacturing tolerances
  - in-service wear-and-tear
  - approximate modeling parameters
- Two types of uncertainty:
  1. Aleatory uncertainty (or type A, or irreducible uncertainty) is characterized by inherent randomness
  2. Epistemic uncertainty (or type B, or reducible uncertainty) represents a lack of knowledge about the appropriate value
- UQ consists of three phases:
  1. characterization of input parameter variability from observations and physical evidence
  2. propagation of input variabilities through the model
  3. calculation of statistical properties of output

Uncertainty Quantification II

- Easiest and most accurate method for propagating aleatory uncertainties through the model is a full non-linear Monte Carlo (MC) simulation which is prohibitively expensive for high-fidelity computations
- For a MC simulation N_t random samples x are drawn from the underlying distribution of the input parameters with mean, μ_D, and standard deviations, σ_{D_j}. Then the mean and variance of the non-linear objective function, f can be computed from

$$\mu_f = \frac{1}{N_t}\sum_{i=1}^{N_t} f(x^{(i)}) \qquad \mathrm{Var}_f = \left(\frac{1}{N_t}\sum_{i=1}^{N_t} f(x^{(i)})^2\right) - \mu_f^2$$

Uncertainty Quantification III

- For epistemic uncertainties MC methods can be employed as well, but the results can only be interpreted with regards to the interval produced on the output functional
- Other approaches, such as Dempster-Shafer evidence theory also require large numbers of function evaluations

Claim

The construction of an accurate surrogate model is one of the most effective options for propagating both aleatory and epistemic uncertainties through a high-fidelity simulation

Unknown Characterizations of Uncertainty I

- Romero from SANDIA discusses the "Engineer's versus Mathematician's UQ problem"

  V. Romero "A Systems Approach to Effective Treatment of Aleatory and Epistemic Uncertainties Involving Typical Information Limitations in Engineering Project", AIAA SciTech 2024-0791

- Mathematician's problem is to solve the posed UQ problem as discussed, however, the engineer knows that the posed UQ problem is not as well-defined as maybe required
- The model being used to propagate the uncertainty may not be validated to any rigorous degree thus having an unknown bias error
- Complex engineering physics models often have quantities of interest (QOIs) that are non-monotonic over input variable uncertainty ranges

Unknown Characterizations of Uncertainty II

- Accuracy beyond a certain amount does not add real value and can even be counter-productive if it diverts resources away from activities which could add more overall value: **Does one spend limited UQ resources on elaborate UQ methodologies or trying to reduce input uncertainties in the UQ problem?**
- Romero: "Much more research is needed by the UQ community to more fully investigate and characterize the performance of the various approaches and methods in the sparse-data probabilistic UQ realm. Nevertheless, mounting empirical evidence exists to corroborate analytical reasoning and arguments that when very sparse sample data are involved, more control and predictability of conservative but not overly conservative tail percentile and probability

Unknown Characterizations of Uncertainty III

estimates may generally exist with Discrete Direct (DD) and Simultaneous DD (SDD) approaches than with other established calibration-propagation-UQ approaches."

- In the real world most input uncertainties are usually interval ranges such as "we believe the value of this input quantity or parameter lies between these upper and lower bounds" or "our best estimate is x and our uncertainty is $\pm 5\%$" and are somewhat speculative in most cases
- Therefore, although the physical nature may be aleatory variability, it is effectively epistemic uncertainty

Unknown Characterizations of Uncertainty IV

- Romero suggests that a reasonable truncation of the strict interval QOI results can be pursued by temporarily representing the model parameter uncertainties with uniform probability distribution enabling the application of a probabilistic mathematical/computational propagation approach
- A reasonable interpretation of posed uncertainty ranges is that they are "conservative" or "safe" ranges outside of which the expert believes little chance exists that the true or most effective value of the parameters lie

Unknown Characterizations of Uncertainty V

- If the assumed uniform distributions had legitimate probabilistic significance, then one might seek, say, the 5th and 95th percentiles of the QOI response. Highly reliably conservative estimates of the central 90% range can be obtained with relatively few MC samples (model runs), for example five or so, through the use of 90/95 tolerance intervals (TIs)
- If significantly more (but still moderate) numbers of UQ samples can be afforded, then the sometimes egregious over-conservatism of small-sample TI methods can be avoided by using other approaches such as
  - Agresti-Couli confidence intervals (AGCIs)
  - Wilks one-tail confidence method
  - Pearson and Johnson four-parameter distribution families

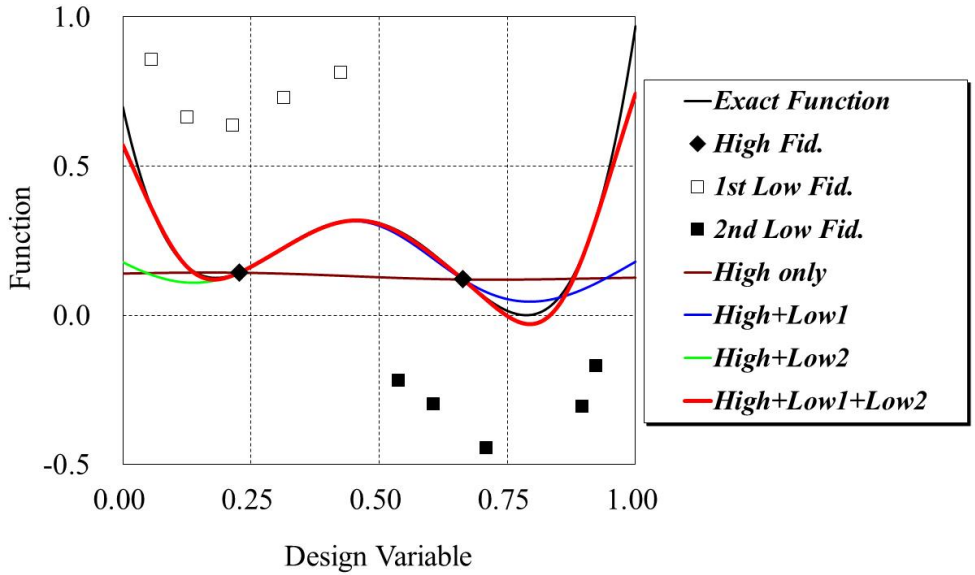
Unknown Characterizations of Uncertainty VI

- Non-parametric kernel-density methods specifically designed for sparse data
- Bayesian model averaging and hypothesis testing
- Multi-distribution Bayesian methods
- Compressive sensing approaches

Multi-Fidelity Surrogate Modeling

Construct more accurate surrogate model by using the trends of lower fidelity data:

High-fidelity Model	Low-fidelity Model
Experimental data	CFD results
RANS	Euler
Finer mesh CFD results	Coarser mesh CFD results
Fully converged solutions	Partly converged solutions



Potential Impact

- Multifidelity methods have potential to
  - Accelerate and reduce cost of design
  - Increase confidence in predictions before downselection
- Two biggest multifidelity issues:
  - Achieving the right fidelity level for the right application
  - Combining models with different levels of fidelity
- Pre-Milestone A Systems Engineering Study (2008): 70–75% of life-cycle cost locked in at analysis of alternatives (downselection) where design freedom is greatest, but design confidence is lowest
- Identification of Critical Flight Loads program survey (1999): Two-thirds of unanticipated structural response/damage incidents due to inadequate loads analysis because of lack of accurate and fully representative aeroelastic data
- Impacts: costly redesign, flight limitations, reduced service life

Construction of Multi-Fidelity Surrogates (MFS)

Premise

Build highly accurate non-intrusive surrogate models at small overall computational cost assuming that obtaining high-fidelity training point information is the dominant cost factor

Strategy: Enhance surrogate modeling techniques by

- Utilizing lower fidelity information
- Adaptively selecting training points
- Utilizing sparsity information and combining surrogate strategies

Benchmark Problems

Scalable and interesting but not too complex
(in terms of implementation and computational cost)

1. Non-polynomial test function (1d)
2. A polynomial test function (any dimension)
3. Coupled Spring-Mass-System leading to a linear system of ODEs (any dimension)

			Cost per Fidelity Level			
Function	**Dimension, D**	**Comp. Budget**	1	2	3	4
Forrester	1	100	1	0.5	0.1	0.05
Rosenbrock	2	200	1	0.5	0.1	-
	5	500	1	0.5	0.1	-
	10	1000	1	0.5	0.1	-
Spring-Mass	2 (springs)	200	1	1/60	-	-
	4 (springs + masses)	400	1	1/60	-	-

Analytic Test Functions I

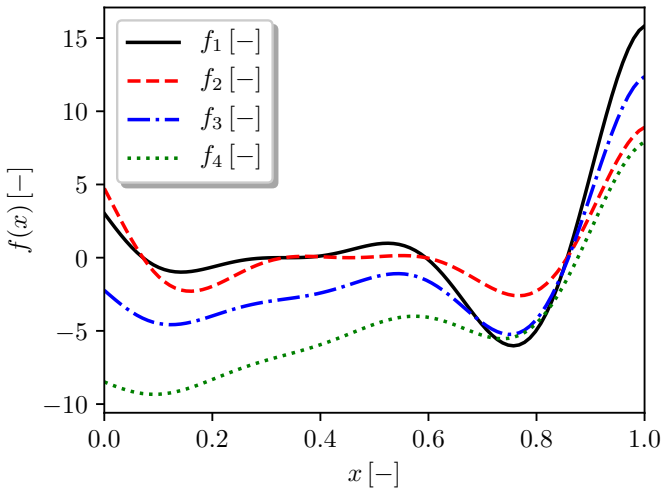
1. Forrester function on domain $[0, 1]$:

$$f_1(x) = (6x - 2)^2 \sin(12x - 4)$$

$$f_2(x) = (5.5x - 2.5)^2 \sin(12x - 4)$$

$$f_3(x) = 0.75 f_1(x) + 5(x - 0.5) - 2$$

$$f_4(x) = 0.5 f_1(x) + 10(x - 0.5) - 5$$



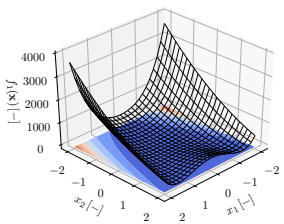
Analytic Test Functions II

2. Rosenbrock function on domain $[-2, 2]^D$:

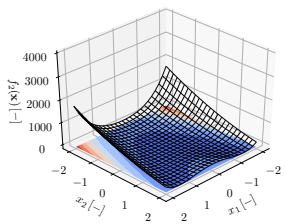
$$f_1(\mathbf{x}) = \sum_{i=1}^{D-1} 100\left(x_{i+1} - x_i^2\right)^2 + (1 - x_i)^2$$

$$f_2(\mathbf{x}) = \sum_{i=1}^{D-1} 50\left(x_{i+1} - x_i^2\right)^2 + (-2 - x_i)^2 - \sum_{i=1}^{D} 0.5 x_i$$

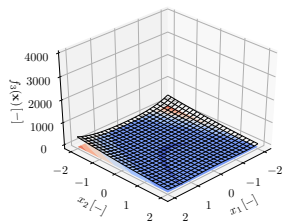
$$f_3(\mathbf{x}) = \frac{f_1(\mathbf{x}) - 4 - \sum_{i=1}^{D} 0.5 x_i}{10 + \sum_{i=1}^{D} 0.25 x_i}$$



f_1



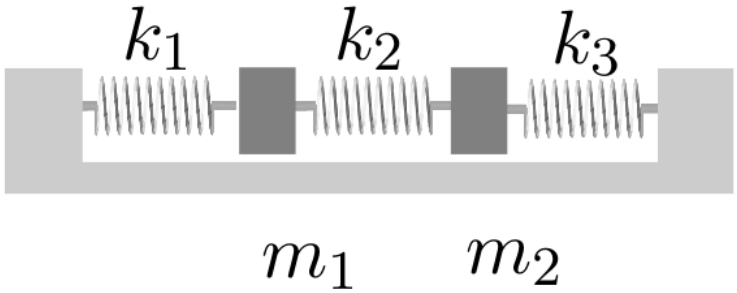
f_2



f_3

Coupled Spring-Mass-System I

- Two masses are attached to each other by three springs



Mass Constants	The masses m_1 and m_2 are assumed to be point masses concentrated at their center of gravity.
Spring Constants	The mass of each spring is negligible. The springs operate according to Hooke's law and the constants k_1, k_2, and k_3 denote the Hooke's constants. The springs restore after compression and extension.
Position Variables	$x_1(t)$ and $x_2(t)$ denote the mass positions along the horizontal surface, measured from their equilibrium positions, positive right and negative left.
Fixed Ends	The first and last spring are attached to fixed walls and have the same spring constant.

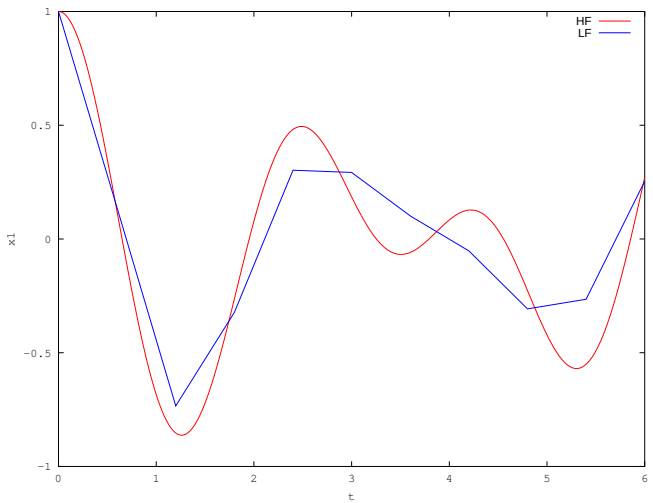
Coupled Spring-Mass-System II

- The equations of motion are given by

$$M\ddot{\mathbf{x}}(t) = K\mathbf{x}(t)$$

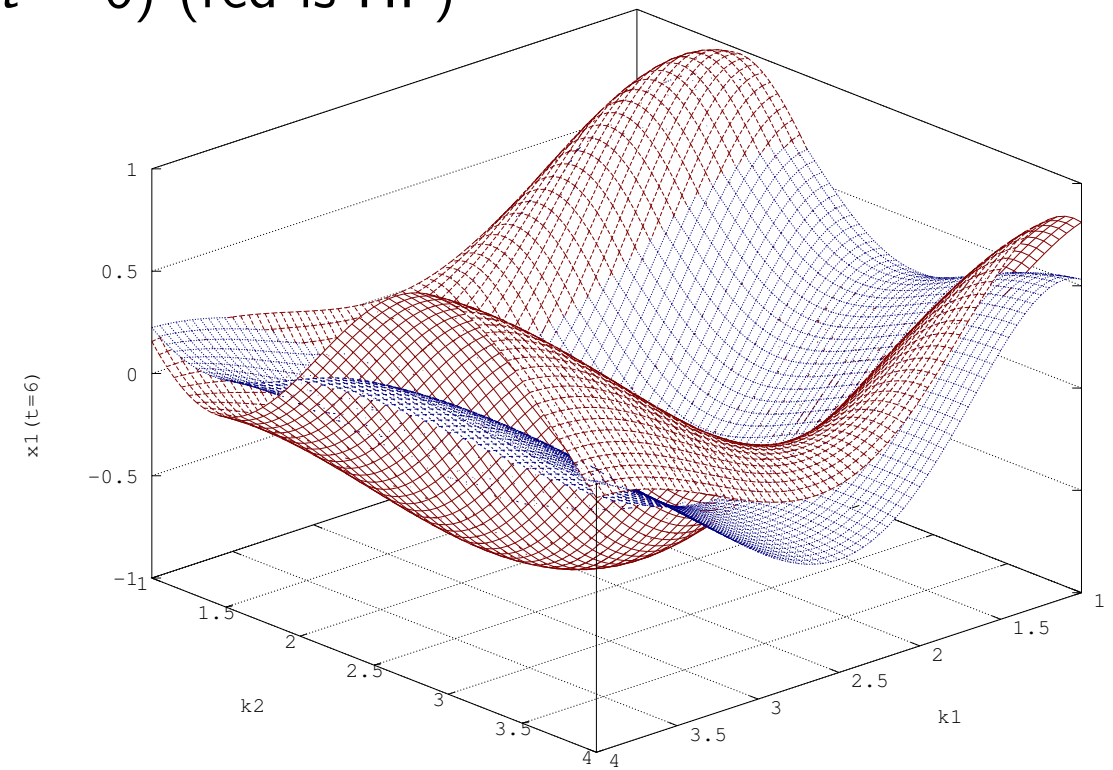
$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \mathcal{M} = \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix} \quad K = \begin{pmatrix} -k_1 - k_2 & k_2 \\ k_2 & -k_2 - k_3 \end{pmatrix}$$

- Let $m_1 = m_2 = 1$, $k_1 = k_2 = k_3 = 1$ as well as $\mathbf{x}_0 = (1 \;\; 0)^T$ and $\dot{\mathbf{x}}_0 = (0 \;\; 0)^T$ and using RK4 time-marching with $\Delta t_1 = 0.01$ and $\Delta t_2 = 0.6$ yields



Coupled Spring-Mass-System III

- Treating the two spring constants as independent input variables where $1 \leq k_1 = k_3, k_2 \leq 4$ while $m_1 = m_2 = 1$ yields for $x_1(t = 6)$ (red is HF)



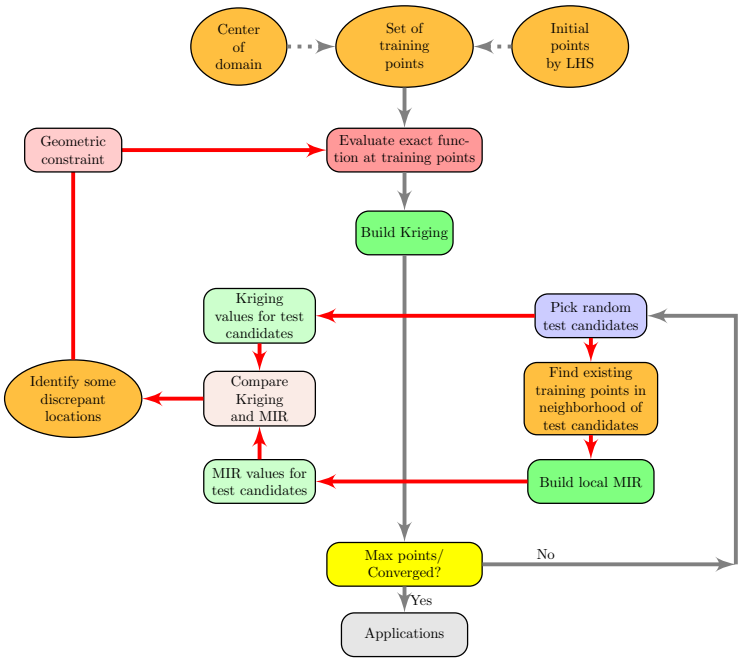
Multi-Fidelity Kriging

- Kriging model approach originated in geological statistics and predicts the function value by using stochastic processes
- Ability to use a gradient and Hessian enhanced cokriging (correlation between func-grad, func-Hess, grad-grad, etc.)
- Try to counteract the "curse of dimensionality":
  - As the number of inputs, D, increases, using the output function and its derivative provides $D + 1$ pieces of information for roughly the cost of two function evaluations
  - Similarly, the Hessian provides $D \cdot (D + 1)/2$ pieces of information for roughly the cost of D function evaluations

$\rightarrow$ One can reasonably expect to have to compute the output function overall far fewer times to obtain a good surrogate

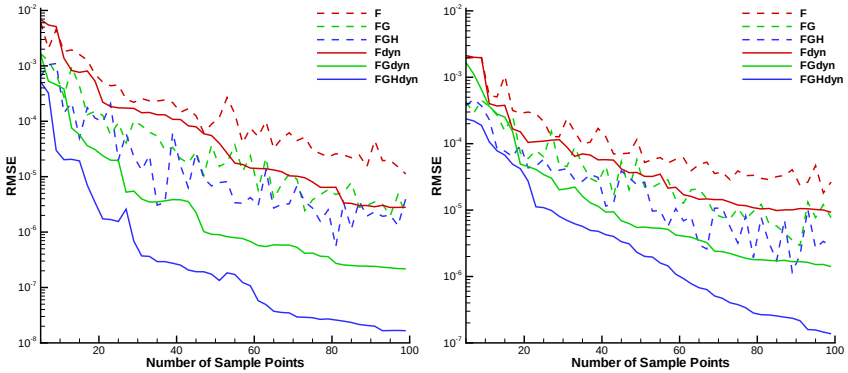
Adaptive Training Point Selection

Idea: Select training points adaptively rather than randomly (typically space filling) at the beginning. That way learned information can be used to hone in on interesting areas of the design space by placing additional training points there.



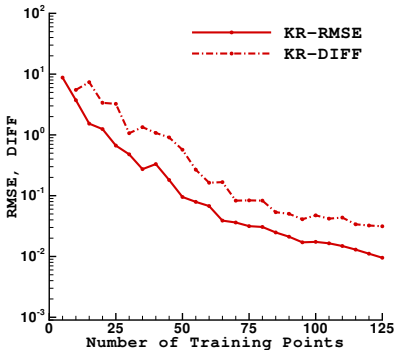
Root-mean square error (RMSE) in 2d

Training points are either selected through latin hypercube (dashed) or dynamic sampling (solid)



Cosine function Runge function

Accuracy Estimate

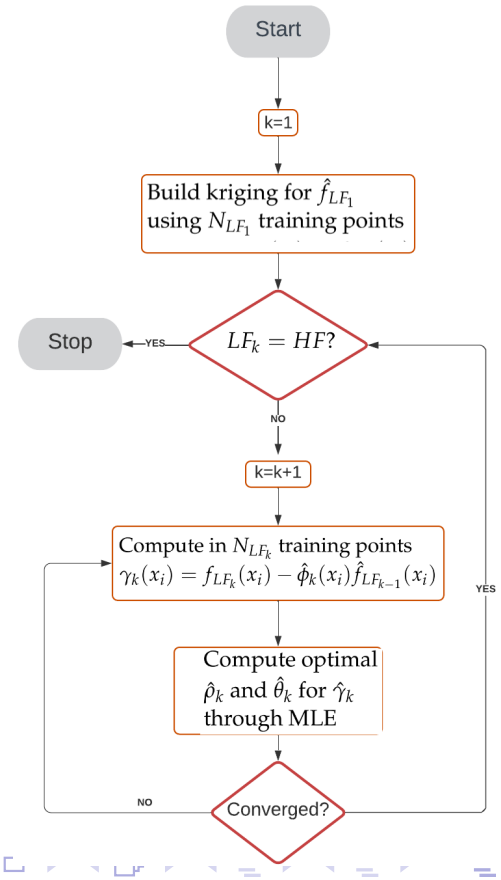


K. Boopathy and M.P. Rumpfkeil "A Unified Framework for Training Point Selection and Error Estimation for Surrogate Models", AIAA Journal, 2015.

Hybrid Bridge Function

$$\hat{f}_{HF}(x) = \hat{\phi}(x)\hat{f}_{LF}(x) + \hat{\gamma}(x)$$
$$\hat{\phi}(x) = g^T(x)\hat{\rho} \quad \text{low-order polynomial}$$

1. Build kriging model for lowest fidelity data, $\hat{f}_{LF_1}$, using N_{LF_1} lowest fidelity training points.
2. Build another kriging model for the additive bridge function, $\hat{\gamma}_2$, where $\gamma_2(x_i) = f_{LF_2}(x_i) - \hat{\phi}_2(x_i)\hat{f}_{LF_1}(x_i)$ in N_{LF_2} next fidelity level training points. Compute an optimal $\hat{\rho}_2$ (and $\hat{\theta}_2$) during the maximum likelihood estimation updates for $\hat{\gamma}_2$.
3. If $LF_2 = HF$ stop, otherwise repeat.



Rumpfkeil, M. P. and Beran, P., "Construction of Multi-Fidelity Surrogate Models for Aerodynamic Databases," Proceedings of ICCFD9, 2016.

Multi-fidelity Sparse Polynomial Chaos Expansions

- Usually, multi-fidelity PCE implementations take a multi-step approach, separating the estimation of corrective functions from the fitting of the model of interest
- Typically, corrections are additive and/or multiplicative, and one must choose a priori what form the correction should take or estimate a blend of both forms
- All-at-once approach developed by Bryson and Rumpfkeil simultaneously determines additive and multiplicative corrections in a least-squares sense
- Approach can also be augmented with the addition of gradient and Hessian information

Bryson, D. E. and Rumpfkeil, M. P., "All-at-Once Approach to Multifidelity Polynomial Chaos Expansion Surrogate Modeling," Aerospace Science and Technology, Vol. 70C, 2017, pp. 121–136.

Training Point Selection

Number of training points for several grid levels and dimensions of **slow-growth, sparse, Gauss-Patterson** (optimal for Legendre basis polynomials) grids

Grid Level	Dimension, D							
	2	3	4	5	6	10	20	40
0	1	1	1	1	1	1	1	1
1	5	7	9	11	13	21	41	81
2	9	19	33	51	73	201	801	3201
3	17	39	81	151	257	1201	10001	82401
4	33	87	193	391	737	5281	90561	1557121
5	33	135	385	903	1889	19105	641409	huge
6	65	207	641	1743	4161	60225	huge	huge
7	97	399	1217	3343	8481	169185	huge	huge

Compressed Sensing (CS) I

- CS is a growing subject in the fields of image or signal processing, computer science and statistics
- CS offers a framework to accurately recover a sparse signal, $\hat{\beta}$, from a set of incomplete observations, $\mathbf{b}$, that is, CS aims at selecting a few basis polynomials with great impact on the model response
- One popular method is to solve this optimization problem

$$\min_{\hat{\beta}} 0.5||\mathbf{A}\hat{\beta} - \mathbf{b}||_2^2 + \lambda||\hat{\beta}||_1$$

- A greedy solution can be found by a least angle regression (LARS) or orthogonal matching pursuit (OMP) algorithm
- LARS and OMP do not allow a column of $\mathbf{A}$ (PCE basis) to leave the active set

Compressed Sensing (CS) II

- Removing this restriction yields LASSO (least absolute shrinkage and selection operator) algorithm which can provably solve optimization problem (e.g., library mlpack)
- Selecting proper value for λ is very important for an accurate computation of the sparse PCE coefficients
  - If λ is too large reconstructed signal may not be accurate enough, while very small values may result in over-fitting
  - Following Salehi *et al.* error or the leave one-out (LOO) cross-validation procedure is used to determine an optimal $\hat{\lambda}$: LASSO is solved for multiple λ and error is calculated for each case; $\hat{\lambda}$ is set to the value of λ for which error is smallest

Rumpfkeil, M. P., Bryson, D., and Beran, P., "Multi-Fidelity Sparse Polynomial Chaos Surrogate Models for Flutter Database Generation," AIAA Paper, 2019-1998.

Multi-Fidelity SPCE-Kriging I

- Kriging is interpolatory in nature and PCE is a regression $\rightarrow$ kriging captures the local variations well, whereas PCE captures the global behavior better
- Combine the advantages of both into a SPCE-Kriging model
- The construction of a single-fidelity SPCE-Kriging consists of the following three main steps:
  1. Construct PCE orthogonal bases corresponding with the distributions of random inputs
  2. Solve for the PCE coefficients using the LASSO algorithm to determine which bases are most important
  3. Use only those important bases as trend functions in the universal kriging mean value
- For the construction of a multi-fidelity SPCE-Kriging the hybrid bridge function is employed

Multi-Fidelity SPCE-Kriging II

- SPCE-Kriging is only used in the training of the lowest fidelity surrogate model
- For the additive bridge function an ordinary kriging model is constructed
- The reason for that is that in training the additive bridge function the genetic algorithm requires a lot of log-likelihood evaluations to compute the optimal $\hat{\rho}_2$ (and $\hat{\theta}_2$)
- Hence, using SPCE-Kriging would require a lot of SPCE evaluations which are much more expensive than simple ordinary kriging constant trend computations

M.P. Rumpfkeil, Dean Bryson and P. Beran "Multi-Fidelity Sparse Polynomial Chaos and Kriging Surrogate Models Applied to Analytical Benchmark Problems", Algorithms, 15(3), 2022.

MF Surrogates Applied to UQ

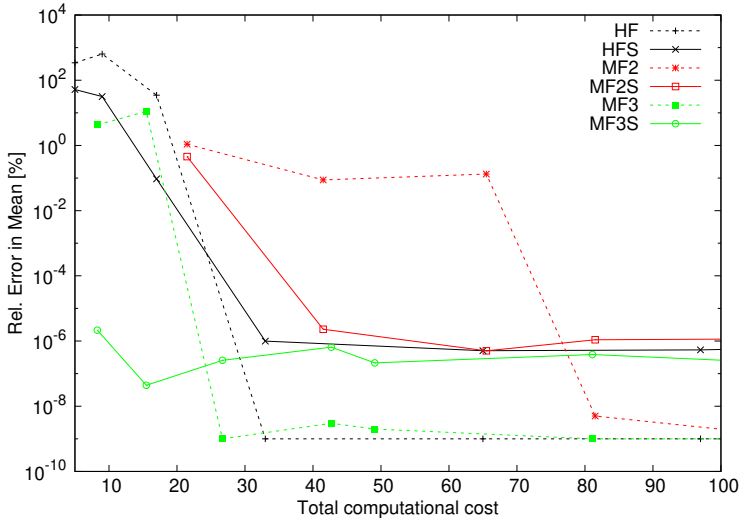
Methodology

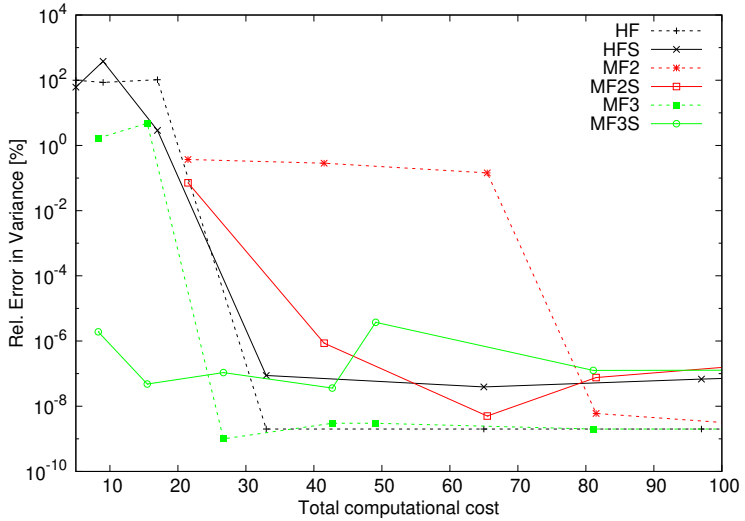
For UQ $50,000$ normally distributed MC sample points created via LHS with $\mu_{D_i} = 0.5$ and $\sigma_{D_i} = 0.15 \rightarrow 99\%$ of all samples fall within computational domain.

The relative errors (in percent) for the mean and variance predictions between full and inexpensive MC simulations are shown.

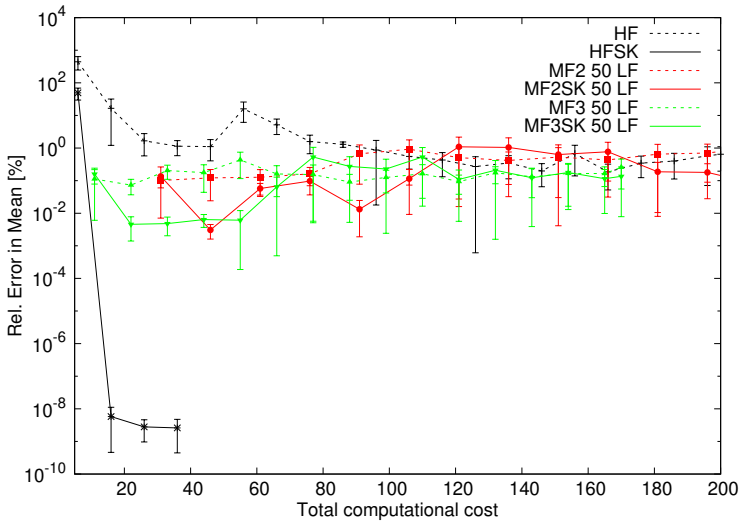
Rosenbrock 2d Results

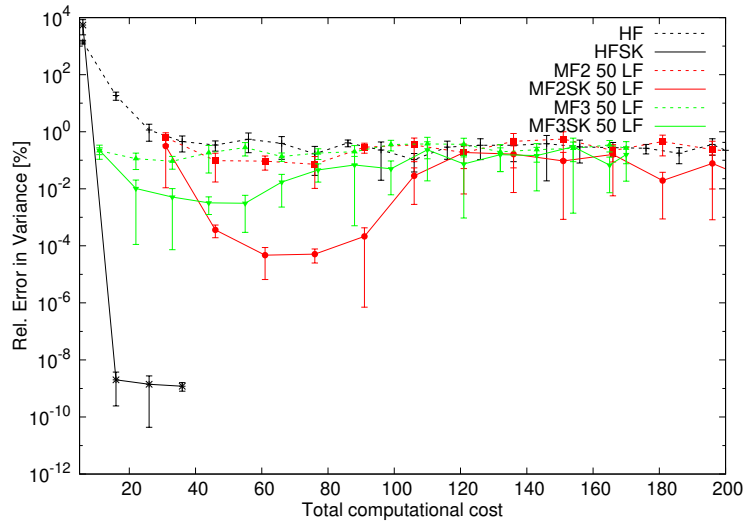
MFSPCE





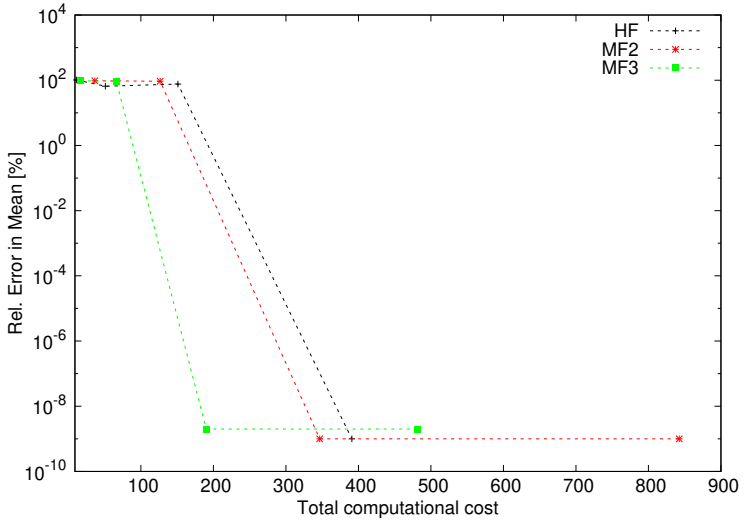
MF Kriging

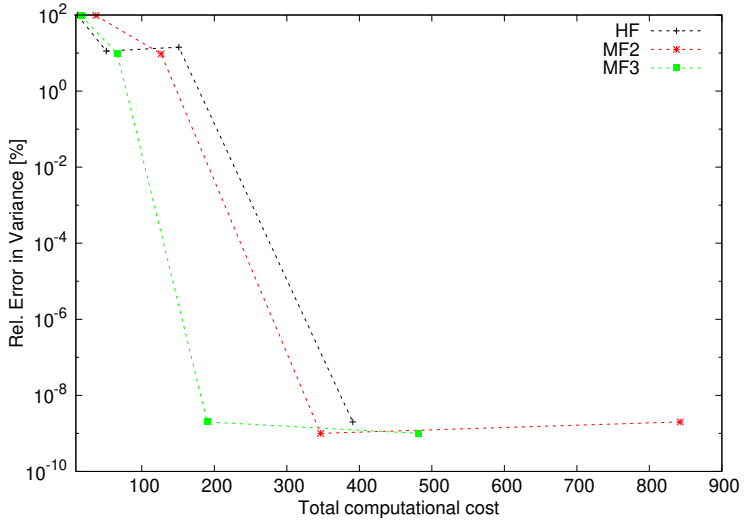




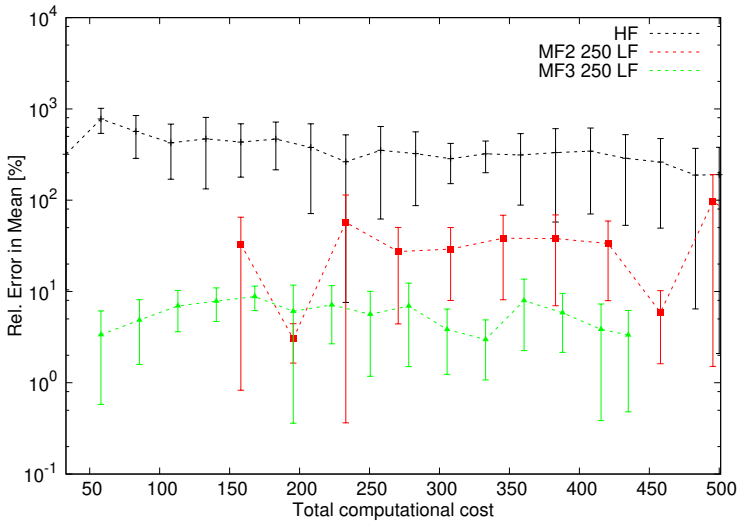
Rosenbrock 5d Results

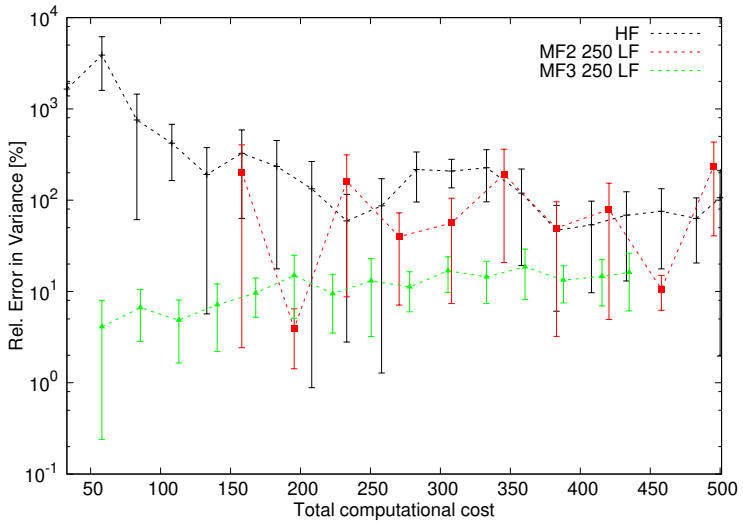
MFSPCE





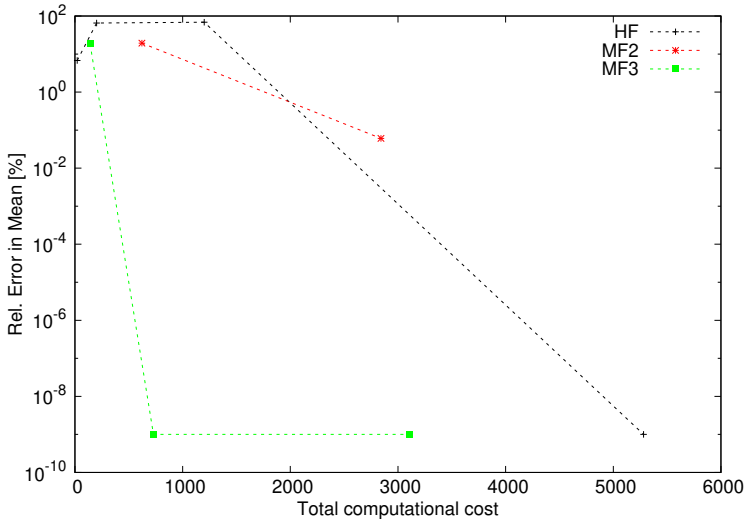
MF Kriging

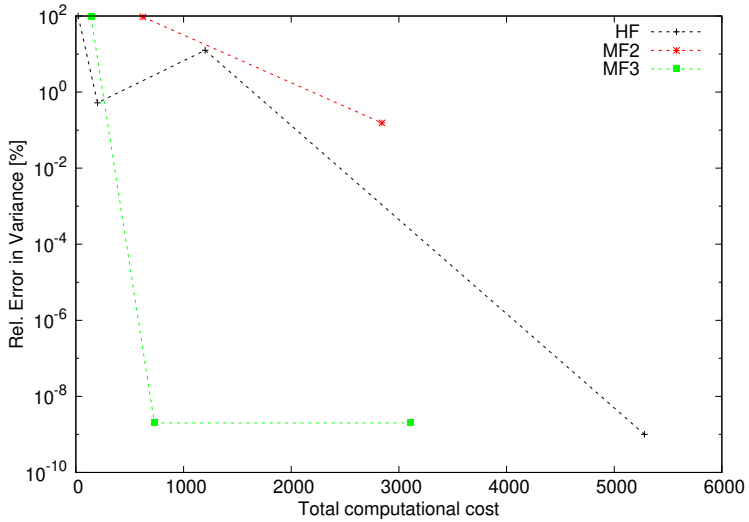




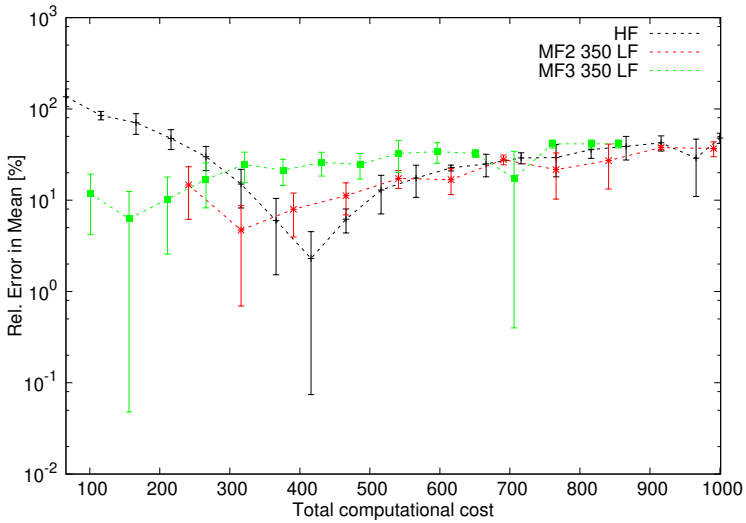
Rosenbrock 10d Results

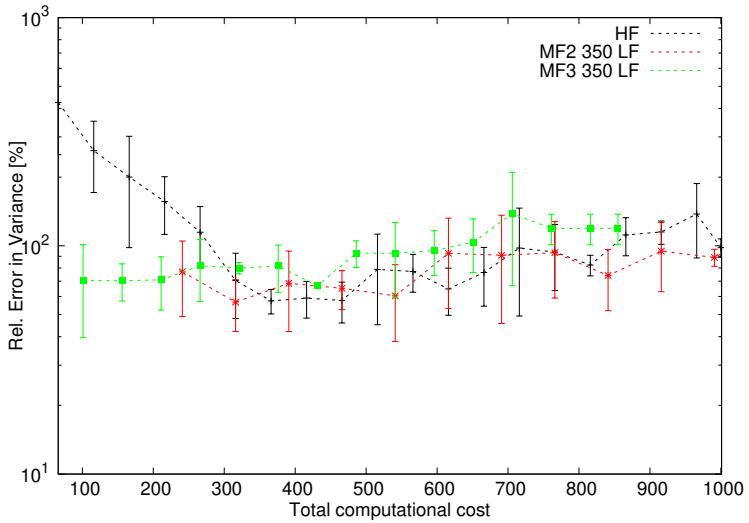
MFSPCE





MF Kriging





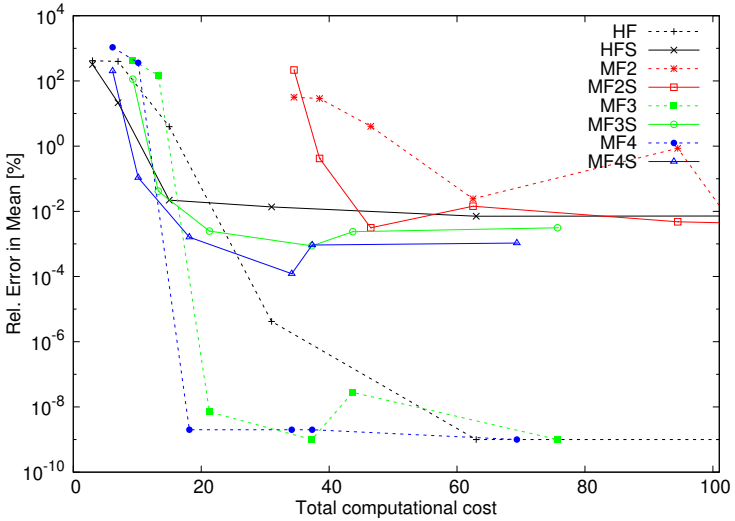
Forrester Results I

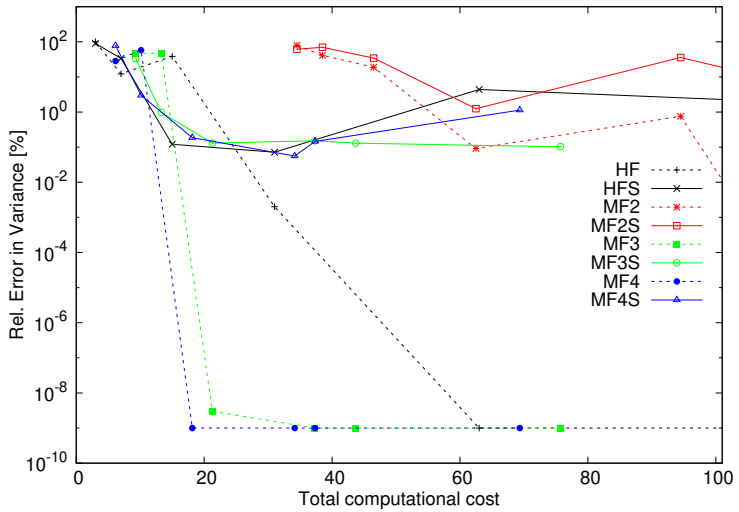
Best order and λ values

# of HF points	# of LF3 points	λ	Order, P	Additive, R	Multiplicative, Q
3	0	10^{0}	3	–	–
7	0	10^{-2}	9	–	–
15	0	10^{-8}	12	–	–
31	0	10^{-8}	12	–	–
63	0	10^{-4}	9	–	–
127	0	10^{-6}	11	–	–
3	63	10^{-8}	11	1	4
7	63	10^{-8}	11	3	3
15	63	10^{-8}	11	4	5
31	63	10^{-8}	12	2	4
31	127	10^{-8}	12	3	5
63	127	10^{-8}	12	5	5

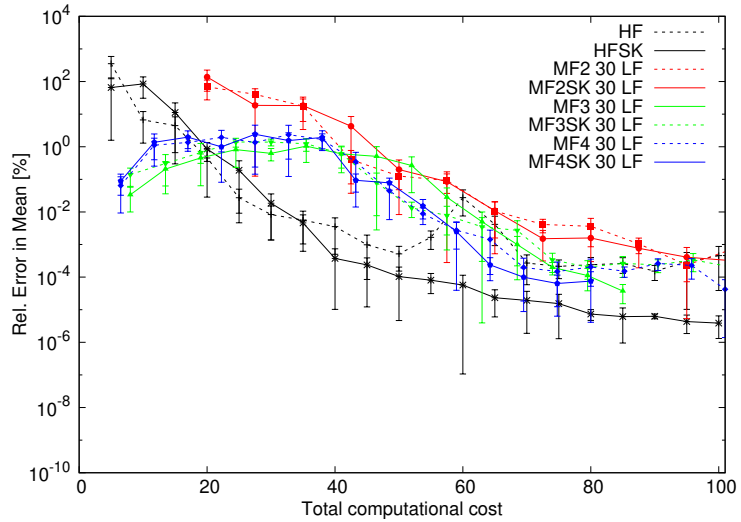
Forrester Results II

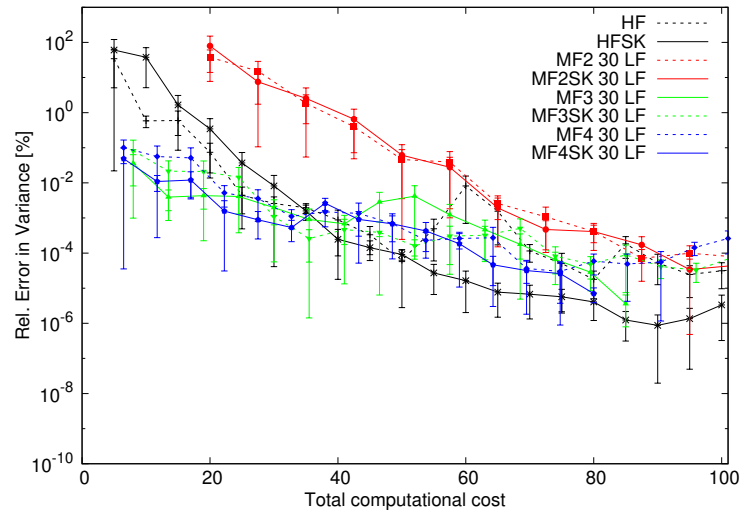
MFSPCE





MF Kriging





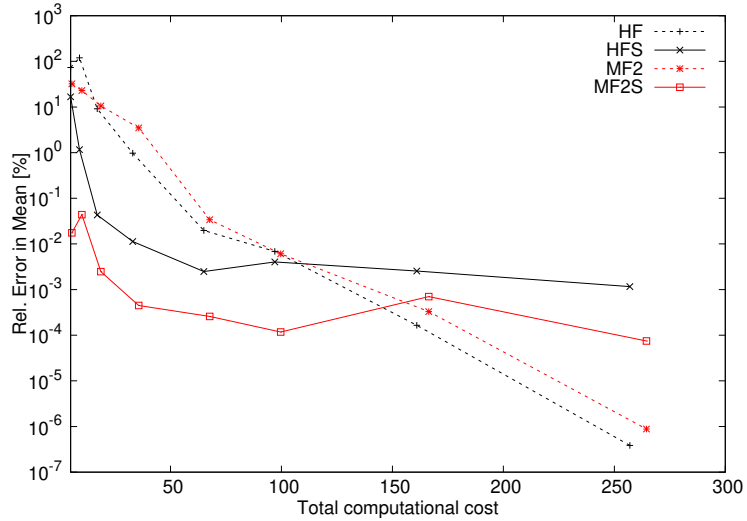
Coupled Spring-Mass-System 2d I

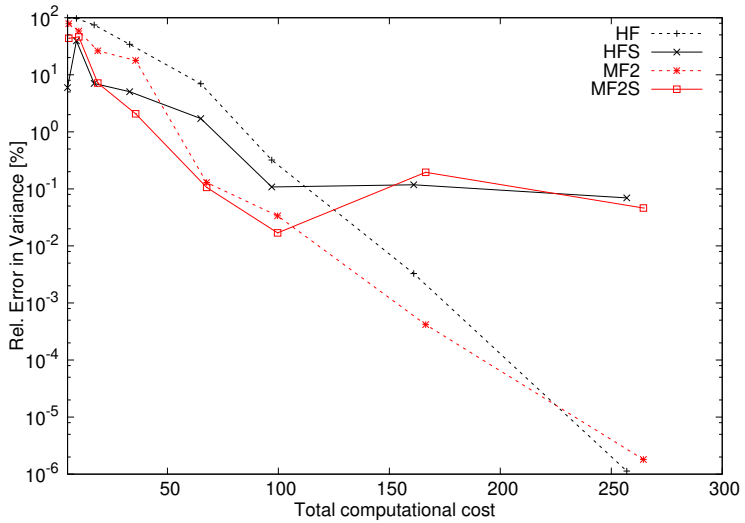
Best order and λ values

# of HF points	# of LF points	λ	Order, P	Additive, R	Multiplicative, Q
5	0	10^{-8}	6	–	–
9	0	10^{-2}	7	–	–
17	0	10^{-4}	9	–	–
33	0	10^{-8}	6	–	–
65	0	10^{-6}	10	–	–
97	0	10^{-8}	12	–	–
161	0	10^{-8}	12	–	–
257	0	10^{-8}	12	–	–
5	33	10^{-8}	10	1	5
9	65	10^{-2}	10	4	5
17	97	10^{-8}	11	4	5
33	161	10^{-6}	11	6	4
65	161	10^{-8}	10	6	4
97	161	10^{-8}	9	6	4
161	321	10^{-8}	11	5	6

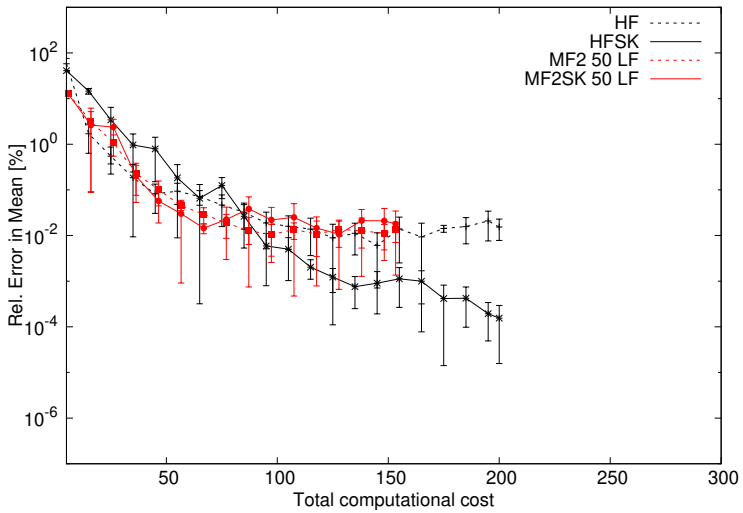
Coupled Spring-Mass-System 2d II

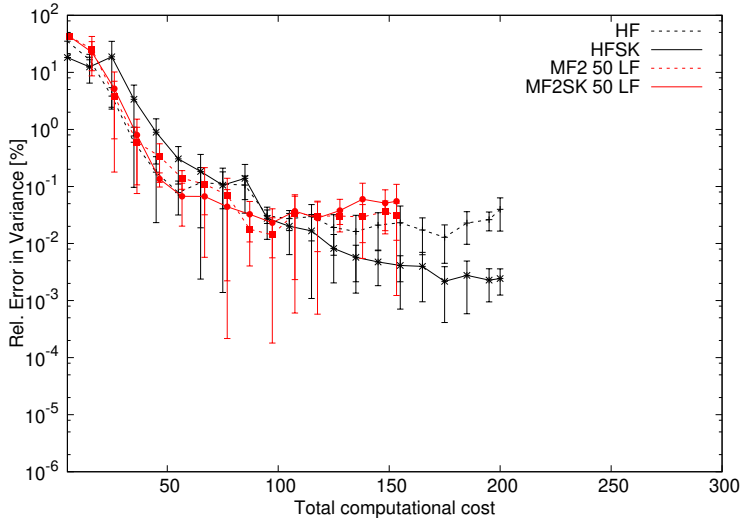
MFSPCE





MF Kriging





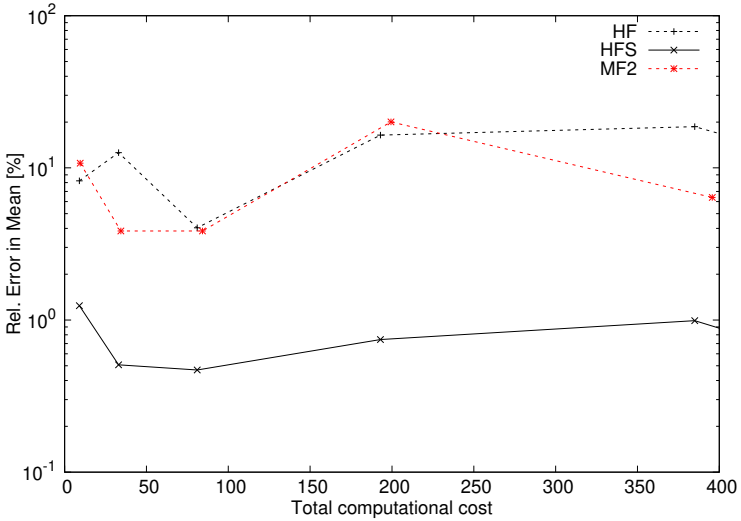
Coupled Spring-Mass-System 4d I

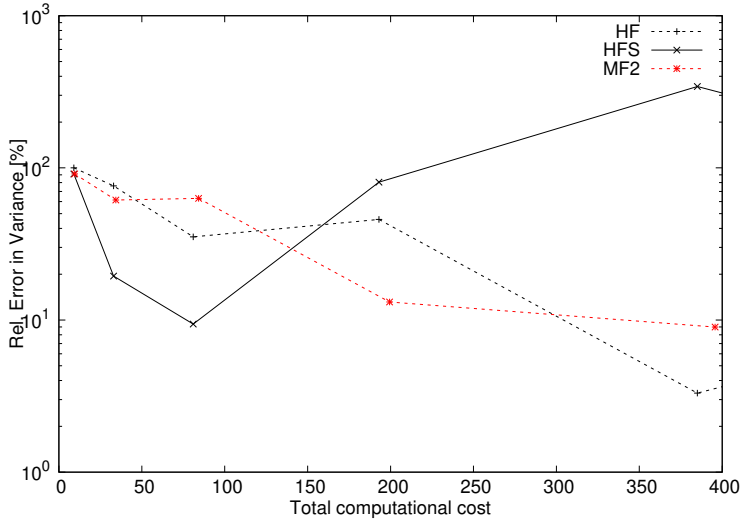
Best order and λ values

# of HF points	# of LF points	λ	Order, P	Additive, R	Multiplicative, Q
9	0	10^{-4}	6	–	–
33	0	10^{-2}	10	–	–
81	0	10^{-2}	6	–	–
193	0	10^{0}	6	–	–
385	0	10^{-8}	9	–	–
641	0	10^{-8}	10	–	–

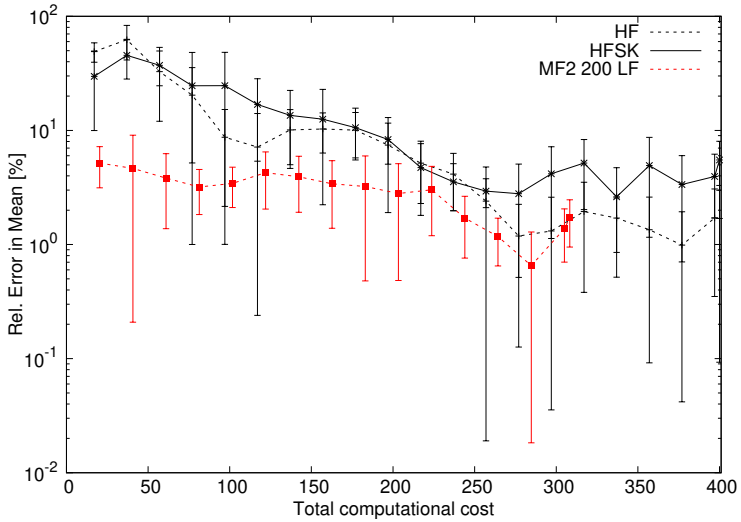
Coupled Spring-Mass-System 4d II

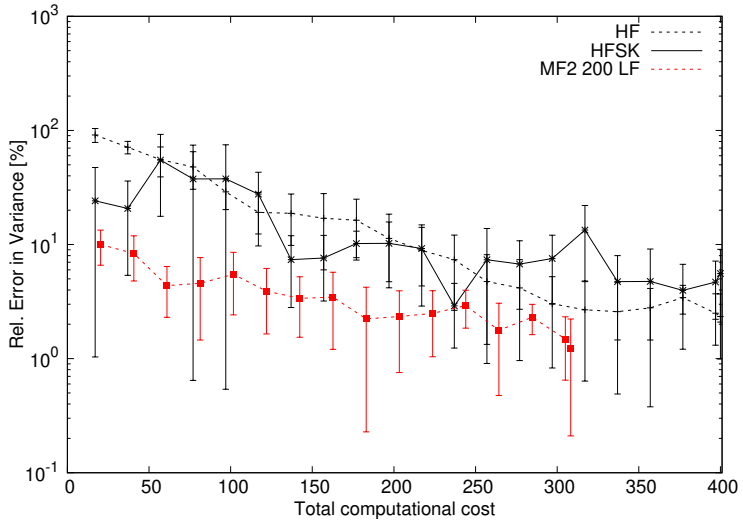
MFSPCE





MF Kriging





Moment Methods I

- If one is only interested in the mean and variance, moment methods can be a good choice
- Moment methods are based on Taylor series expansions of the original non-linear simulation output $f(D)$ about the mean of the input $\bar{D}$ given standard deviations σ_{D_j}
- The resulting mean $\bar{f}$ and variance Var_f of simulation output are given to first order (MM1) by

$$\begin{aligned} \bar{f}^{(1)} &= f(\bar{D}) \\ \mathrm{Var}_f^{(1)} &= \sum_{j=1}^{M} \left(\frac{df}{dD_j}\bigg|_{\bar{D}} \sigma_{D_j} \right)^2 \end{aligned}$$

Moment Methods II

- and to second order (MM2) by

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

$$\mathrm{Var}_f^{(2)} = \mathrm{Var}_f^{(1)} + \frac{1}{2}\sum_{j=1}^{M}\sum_{k=1}^{M}\left(\left.\frac{d^2 f}{dD_j dD_k}\right|_{\bar{D}}\sigma_{D_j}\sigma_{D_k}\right)^2$$

- Note the non-linear shift between the mean of the output and the output of the mean is accounted for by the Hessian diagonal elements
- For the gradient of $\mathrm{Var}_f^{(1)}$ one needs the Hessian as well

Optimization Under Uncertainty I

- A conventional constrained optimization problem:

$$\begin{aligned} \min \quad & J = J(f, q, D) \\ s.t. \quad & R(q, D) = 0 \\ & g(f, q, D) \leq 0. \end{aligned}$$

  with simulation output functionals $f(D) = F(q(D), D)$
- Given uncertainties in D everything is no longer deterministic
- Using a moment matching formulation an optimization under uncertainty (OUU) problem can be expressed as

$$\begin{aligned} \min \quad & \mathcal{J} = \mathcal{J}(\bar{f}, \sigma_f, \bar{q}, \bar{D}) \\ s.t. \quad & R(\bar{q}, \bar{D}) = 0 \\ & g(\bar{f}, \bar{q}, \bar{D}) + k\sigma_g \leq 0 \quad \Longleftrightarrow \quad \mathcal{P}(g \leq 0) \geq P_k \end{aligned}$$

statistically independent
+ normally distributed D

Optimization Under Uncertainty II

- For example, one can define an objective function for robust optimization problems

$$\mathcal{J} = w_1 \bar{f} + w_2 \sigma_f^2 \qquad w_1, w_2 \quad \text{user specified weights}$$

- One could treat this as a multi-objective optimization problem
- For the MM1 approximation the calculation of the standard deviations σ_f and σ_g involves first-order sensitivities
  $\rightarrow$ Gradient-based optimizer requires the Hessian to compute the objective and constraint gradients
- A more accurate estimate of the required means, variances, and standard deviations can be obtained by using a Kriging surrogate, but
  - One has to construct a separate response surface for each simulation output f and system constraint g

Optimization Under Uncertainty III

- Quasi-Newton optimizer still needs a gradient at the mean design variable values in order to calculate a direction of improvement

- One idea might be to just use an approximation of the gradient by using the gradient of the MM1 estimate for the current mean design variable values
- Another, more accurate method, is to use the Kriging model itself to calculate the required gradient
- Kriging predictor $\hat{f}(D)$ for an untried design point D

$$\hat{f}(D) = \underbrace{\hat{\mu}}_{MeanValue} + \underbrace{r^T(D) \overbrace{R^{-1}}^{CorrelationMatrix} (Y - \hat{\mu} I)}_{StationaryRandomProcess} = \hat{\mu} + r^T V$$

Optimization Under Uncertainty IV

- "Gradient predictor" is then given by

$$\frac{d\hat{f}}{dD_j} = \frac{dr^T}{dD_j} R^{-1}(Y - \hat{\mu} I) = \frac{dr^T}{dD_j} V$$

- Sample Kriging $\tilde{N}$ times for inputs $\tilde{D}^k$ chosen based on their underlying probability distribution function [$\tilde{D} = \bar{D} + \sigma_D Z$ with $Z \sim \mathcal{N}(0, 1)$]:

$$\bar{f} \approx \frac{1}{\tilde{N}} \sum_{k=1}^{\tilde{N}} \hat{f}(\tilde{D}^k)$$

Optimization Under Uncertainty V

- Derivative can be calculated at the same time with little computational overhead via

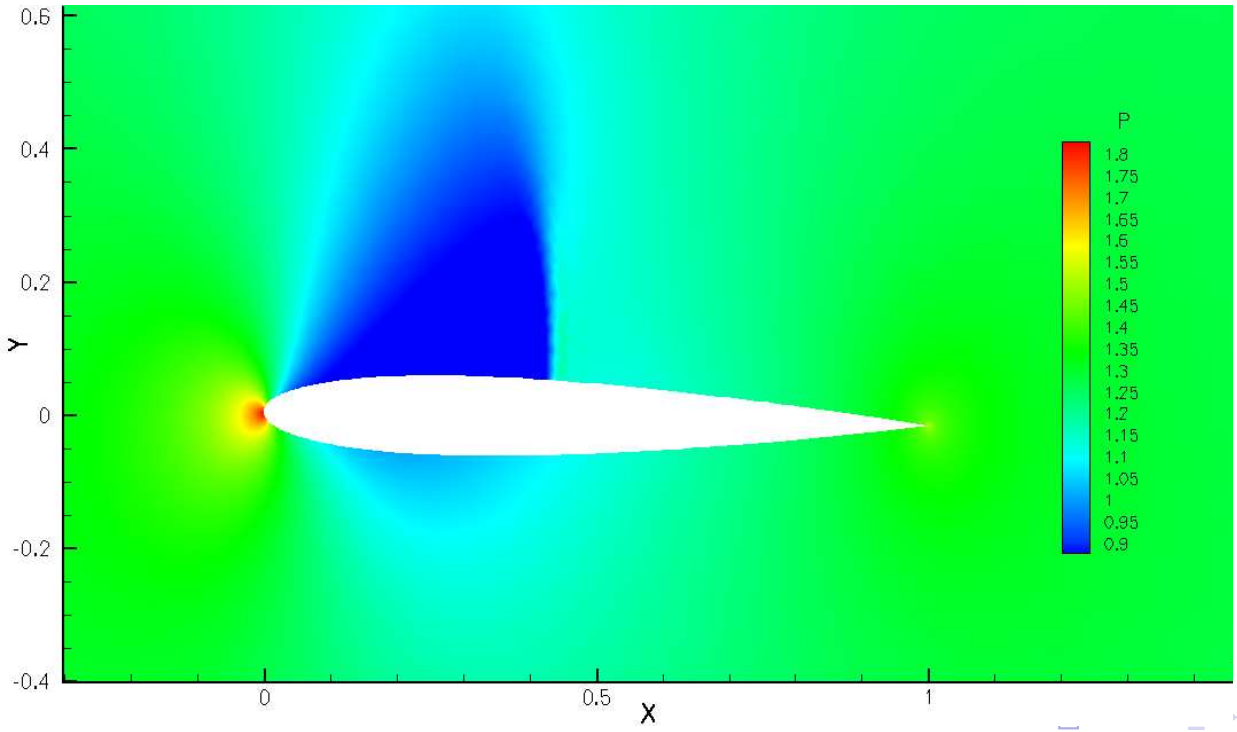
$$\frac{d\bar{f}}{d\bar{D}_j} \approx \frac{1}{\tilde{N}} \sum_{k=1}^{\tilde{N}} \frac{d\hat{f}(\tilde{D}^k)}{d\tilde{D}_i^k} \frac{d\tilde{D}_i^k}{d\bar{D}_j} = \frac{1}{\tilde{N}} \sum_{k=1}^{\tilde{N}} \frac{d\hat{f}(\tilde{D}^k)}{d\tilde{D}_i^k} \delta_{ij} = \frac{1}{\tilde{N}} \sum_{k=1}^{\tilde{N}} \frac{d\hat{f}(\tilde{D}^k)}{d\tilde{D}_j^k}$$

- Similarly, the variance and its derivative can be approximated

$$\begin{aligned} \mathrm{Var}_f &\approx \left(\frac{1}{\tilde{N}} \sum_{k=1}^{\tilde{N}} \hat{f}^2(\tilde{D}^k) \right) - \bar{f}^2 \\ \frac{d\mathrm{Var}_f}{d\bar{D}_j} &\approx \left(\frac{2}{\tilde{N}} \sum_{k=1}^{\tilde{N}} \hat{f}(\tilde{D}^k) \frac{d\hat{f}(\tilde{D}^k)}{d\tilde{D}_j^k} \right) - 2\bar{f} \frac{d\bar{f}}{d\bar{D}_j} \end{aligned}$$

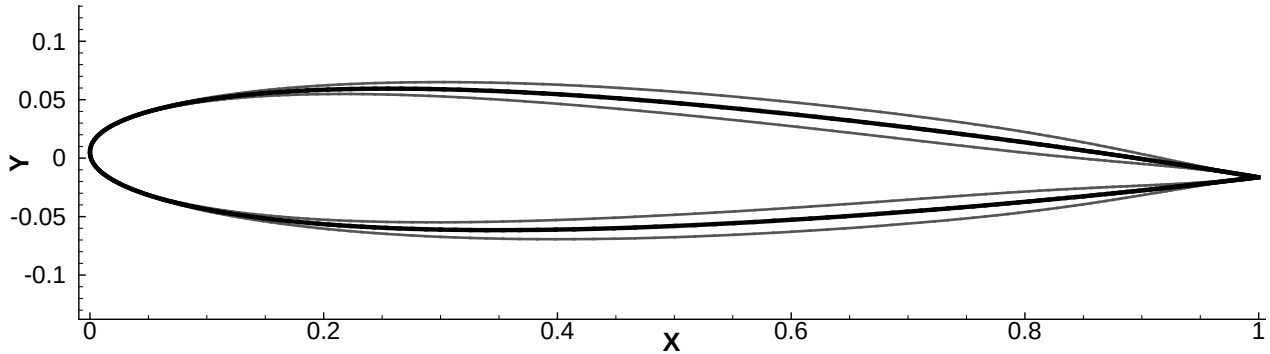
Robustness Optimization of a Transonic Airfoil I

- Consider steady inviscid case of transonic NACA 0012 airfoil
- $M_\infty = 0.755$ with angle of attack of 1.25° leading to $C_l = 0.268$ and $C_d = 0.00521$
- Computational mesh has about $20,000$ triangular elements
- Finite-volume Euler flow solver second-order accurate



Robustness Optimization of a Transonic Airfoil II

- Lift-constrained drag minimization $\mathcal{J} = \bar{C}_d + \sigma^2_{C_d}$
- Vary three shape design variable on the upper and three on the lower surface which control the magnitude of Hicks-Henne sine bump functions
- Deformation of mesh via linear tension spring analogy
- Assume normal distributions for design variables with same standard deviations of ± 0.005



- Deterministic optimization yields a minimal drag of $C_d = 0.00153$ (down from $C_d = 0.00521$)

Comparison of predictions for the original NACA 0012 airfoil at $\alpha = 1.25$

	$\bar{C}_l$	$\sigma^2_{C_l}$	$\bar{C}_d$	$\sigma^2_{C_d}$
NLMC (3000 F)	0.267	$8.7 \cdot 10^{-3}$	$5.95 \cdot 10^{-3}$	$5.8 \cdot 10^{-6}$
MM1	0.268	$6.6 \cdot 10^{-3}$	$5.21 \cdot 10^{-3}$	$3.6 \cdot 10^{-6}$
Kriging (49 F)	0.267	$8.7 \cdot 10^{-3}$	$5.93 \cdot 10^{-3}$	$5.5 \cdot 10^{-6}$
Kriging (97 F)	0.266	$8.6 \cdot 10^{-3}$	$5.95 \cdot 10^{-3}$	$5.7 \cdot 10^{-6}$

Robust optimization results for $\alpha = 1.25$

k	P_k	$\bar{C}_d$	$\sigma^2_{C_d}$	$\bar{C}_l$	σ_{C_l}
0	0.5000	$1.53 \cdot 10^{-3}$	$1.5 \cdot 10^{-6}$	0.268	0.099
1	0.8413	$3.25 \cdot 10^{-3}$	$7.3 \cdot 10^{-6}$	0.383	0.115
2	0.9772	$7.15 \cdot 10^{-3}$	$1.3 \cdot 10^{-5}$	0.605	0.168
3	0.9986	Does not converge			

MM1

k	P_k	$\bar{C}_d$	$\sigma^2_{C_d}$	$\bar{C}_l$	σ_{C_l}
0	0.5000	$2.46 \cdot 10^{-3}$	$2.4 \cdot 10^{-6}$	0.269	0.104
1	0.8413	$3.79 \cdot 10^{-3}$	$3.8 \cdot 10^{-6}$	0.375	0.105
2	0.9772	$5.54 \cdot 10^{-3}$	$6.0 \cdot 10^{-6}$	0.479	0.106
3	0.9986	$8.10 \cdot 10^{-3}$	$8.9 \cdot 10^{-6}$	0.592	0.108

Kriging

NLMC (3000F)	$\bar{C}_d$	$\sigma^2_{C_d}$	$\bar{C}_l$	σ_{C_l}
For MM1 k=2 optimum	$8.69 \cdot 10^{-3}$	$1.2 \cdot 10^{-5}$	0.601	0.108
For Kriging k=2 optimum	$5.54 \cdot 10^{-3}$	$5.9 \cdot 10^{-6}$	0.479	0.105

Comparison of NLMC, Kriging, and MM1 gradients for original NACA 0012 airfoil at $\alpha = 1.25$

Gradient	Method	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$	$j=6$
$\frac{d\bar{C}_d}{d\bar{D}_j}$	NLMC	$2.2 \cdot 10^{-1}$	$1.3 \cdot 10^{-1}$	$7.7 \cdot 10^{-2}$	$2.4 \cdot 10^{-1}$	$-1.8 \cdot 10^{-1}$	$1.5 \cdot 10^{-1}$
	Kriging	$2.2 \cdot 10^{-1}$	$1.3 \cdot 10^{-1}$	$7.5 \cdot 10^{-2}$	$2.3 \cdot 10^{-1}$	$-1.8 \cdot 10^{-1}$	$1.5 \cdot 10^{-1}$
	MM1	$1.9 \cdot 10^{-1}$	$1.1 \cdot 10^{-1}$	$6.5 \cdot 10^{-2}$	$1.0 \cdot 10^{-1}$	$-2.4 \cdot 10^{-1}$	$1.5 \cdot 10^{-1}$
$\frac{d\sigma^2_{C_d}}{d\bar{D}_j}$	NLMC	$2.2 \cdot 10^{-4}$	$1.4 \cdot 10^{-4}$	$9.7 \cdot 10^{-5}$	$8.1 \cdot 10^{-4}$	$2.1 \cdot 10^{-5}$	$8.5 \cdot 10^{-5}$
	Kriging	$2.0 \cdot 10^{-4}$	$1.4 \cdot 10^{-4}$	$9.9 \cdot 10^{-5}$	$7.5 \cdot 10^{-4}$	$5.0 \cdot 10^{-5}$	$8.3 \cdot 10^{-5}$
	MM1	$1.6 \cdot 10^{-4}$	$9.9 \cdot 10^{-5}$	$6.6 \cdot 10^{-5}$	$4.2 \cdot 10^{-4}$	$-1.2 \cdot 10^{-5}$	$9.4 \cdot 10^{-5}$
$\frac{d\bar{C}_l}{d\bar{D}_j}$	NLMC	9.06	6.61	5.43	6.08	8.30	8.57
	Kriging	9.01	6.59	5.42	5.86	8.27	8.52
	MM1	8.25	6.10	5.09	3.44	7.25	8.21
$\frac{d\sigma_{C_l}}{d\bar{D}_j}$	NLMC	$-1.9 \cdot 10^{-1}$	$-1.8 \cdot 10^{-1}$	$-1.3 \cdot 10^{-1}$	$-1.0 \cdot 10^{-3}$	$6.5 \cdot 10^{-1}$	$4.8 \cdot 10^{-1}$
	Kriging	$-1.0 \cdot 10^{-1}$	$-1.9 \cdot 10^{-1}$	$-1.5 \cdot 10^{-1}$	$-1.0 \cdot 10^{-1}$	$5.2 \cdot 10^{-1}$	$4.7 \cdot 10^{-1}$
	MM1	$6.3 \cdot 10^{-2}$	$-3.6 \cdot 10^{-2}$	$-2.7 \cdot 10^{-2}$	1.26	1.56	$4.0 \cdot 10^{-1}$

Robust optimization results for $\alpha = 0.75$

k	P_k	$\bar{C}_d$	$\sigma^2_{C_d}$	$\bar{C}_l$	σ_{C_l}
0	0.5000	$3.44 \cdot 10^{-4}$	$7.5 \cdot 10^{-8}$	0.162	0.095
1	0.8413	$8.65 \cdot 10^{-4}$	$4.3 \cdot 10^{-7}$	0.270	0.108
2	0.9772	$2.20 \cdot 10^{-3}$	$1.2 \cdot 10^{-6}$	0.366	0.102
3	0.9986	$3.68 \cdot 10^{-3}$	$2.5 \cdot 10^{-6}$	0.494	0.111

MM1

k	P_k	$\bar{C}_d$	$\sigma^2_{C_d}$	$\bar{C}_l$	σ_{C_l}
0	0.5000	$9.56 \cdot 10^{-4}$	$5.6 \cdot 10^{-7}$	0.162	0.100
1	0.8413	$1.46 \cdot 10^{-3}$	$1.1 \cdot 10^{-6}$	0.261	0.099
2	0.9772	$2.35 \cdot 10^{-3}$	$2.2 \cdot 10^{-6}$	0.356	0.096
3	0.9986	$3.68 \cdot 10^{-3}$	$3.6 \cdot 10^{-6}$	0.448	0.095

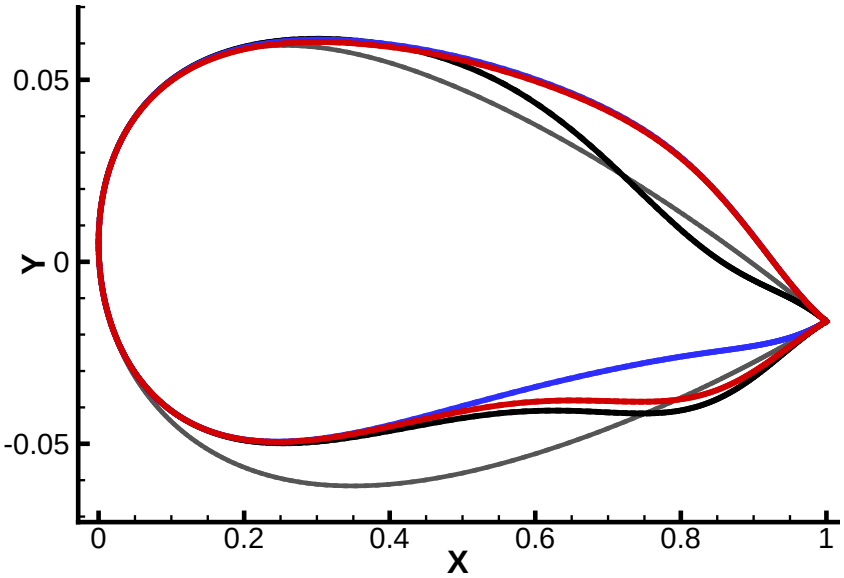
Kriging

NLMC (3000F)	$\bar{C}_d$	$\sigma^2_{C_d}$	$\bar{C}_l$	σ_{C_l}
For MM1 k=2 optimum	$2.81 \cdot 10^{-3}$	$2.2 \cdot 10^{-6}$	0.366	0.096
For Kriging k=2 optimum	$2.34 \cdot 10^{-3}$	$2.1 \cdot 10^{-6}$	0.356	0.096

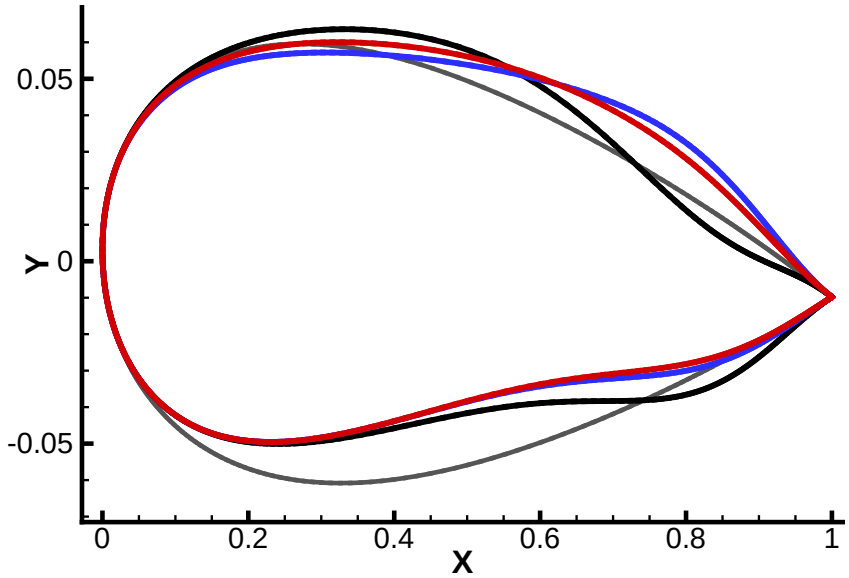
Comparison of NLMC, Kriging, and MM1 gradients for original NACA 0012 airfoil at $\alpha = 0.75$

Gradient	Method	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$	$j=6$
$\frac{d\bar{C}_d}{d\bar{D}_j}$	NLMC	$1.2 \cdot 10^{-1}$	$7.6 \cdot 10^{-2}$	$3.9 \cdot 10^{-2}$	$1.6 \cdot 10^{-1}$	$-1.1 \cdot 10^{-1}$	$8.9 \cdot 10^{-2}$
	Kriging	$1.2 \cdot 10^{-1}$	$7.5 \cdot 10^{-2}$	$4.0 \cdot 10^{-2}$	$1.6 \cdot 10^{-1}$	$-1.1 \cdot 10^{-1}$	$8.8 \cdot 10^{-2}$
	MM1	$1.0 \cdot 10^{-1}$	$6.4 \cdot 10^{-2}$	$3.7 \cdot 10^{-2}$	$8.0 \cdot 10^{-2}$	$-1.3 \cdot 10^{-1}$	$8.7 \cdot 10^{-2}$
$\frac{d\sigma^2_{C_d}}{d\bar{D}_j}$	NLMC	$1.4 \cdot 10^{-4}$	$8.5 \cdot 10^{-5}$	$5.6 \cdot 10^{-5}$	$4.5 \cdot 10^{-4}$	$-4.2 \cdot 10^{-5}$	$6.6 \cdot 10^{-5}$
	Kriging	$1.2 \cdot 10^{-4}$	$9.1 \cdot 10^{-5}$	$5.7 \cdot 10^{-5}$	$4.0 \cdot 10^{-4}$	$-3.3 \cdot 10^{-5}$	$6.4 \cdot 10^{-5}$
	MM1	$1.7 \cdot 10^{-5}$	$1.1 \cdot 10^{-5}$	$7.3 \cdot 10^{-6}$	$-5.0 \cdot 10^{-5}$	$-1.2 \cdot 10^{-4}$	$3.7 \cdot 10^{-5}$
$\frac{d\bar{C}_l}{d\bar{D}_j}$	NLMC	8.96	6.62	5.57	6.00	7.56	8.70
	Kriging	8.97	6.60	5.58	6.11	7.58	8.71
	MM1	8.55	6.35	5.40	4.75	7.27	8.53
$\frac{d\sigma_{C_l}}{d\bar{D}_j}$	NLMC	$-2.7 \cdot 10^{-1}$	$-2.6 \cdot 10^{-1}$	$-1.8 \cdot 10^{-1}$	$4.2 \cdot 10^{-2}$	$4.9 \cdot 10^{-1}$	$3.6 \cdot 10^{-1}$
	Kriging	$-2.2 \cdot 10^{-1}$	$-2.5 \cdot 10^{-1}$	$-2.1 \cdot 10^{-1}$	$1.2 \cdot 10^{-1}$	$4.2 \cdot 10^{-1}$	$3.4 \cdot 10^{-1}$
	MM1	$-2.2 \cdot 10^{0}$	$-1.4 \cdot 10^{0}$	$-9.3 \cdot 10^{-1}$	$-7.1 \cdot 10^{0}$	$-5.1 \cdot 10^{-1}$	$-3.8 \cdot 10^{-1}$

Original NACA 0012 (gray), deterministically (black), MM1 (blue), and Kriging optimized (red) airfoils



$\alpha = 1.25$



$\alpha = 0.75$

Conclusion I

1. Presented a multi-fidelity Kriging surrogate model augmented with dynamic training point selection
2. Presented a multi-fidelity sparse polynomial chaos expansion (MFSPCE) model
3. Combined several surrogate modeling strategies in a somewhat ad hoc manner and used it for inexpensive Monte-Carlo simulations for uncertainty quantification
4. Applied surrogate UQ strategies to analytical test functions and a coupled spring-mass-system
5. Demonstrated benefits from
   - incorporation of lower fidelity data
   - compressed sensing for PCEs
   - combining Kriging and SPCE into a SPCE-kriging model

Conclusion II

6. Described how a first-order moment method and a Kriging surrogate model can be used for OUU problems.
7. Gradient-based optimizations using first-order moment methods require first- and second-order sensitivity derivatives.
8. Kriging surrogate model is much better suited to model non-linearities in the design space, however, the computational cost increases significantly.
9. Performed robust optimizations under uncertainty for a lift-constrained drag minimization of a transonic airfoil
10. Demonstrated that both, a first-order moment method and a Kriging approach, can lead to successful and computationally manageable optimizations under uncertainty.

Any Questions