2020

# Adaptive Multi-Fidelity Modeling for Efficient Design Exploration Under Uncertainty

Atticus J. Beachy
*Wright State University*

# ADAPTIVE MULTI-FIDELITY MODELING FOR EFFICIENT DESIGN EXPLORATION UNDER UNCERTAINTY

A Thesis submitted in partial fulfillment of the
requirements for the degree of
Master of Science in Mechanical Engineering

by

ATTICUS J. BEACHY

B.S.M.E., Cedarville University, 2018

2020

Wright State University

WRIGHT STATE UNIVERSITY

GRADUATE SCHOOL

July 30, 2020

I HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER MY SUPERVISION BY Atticus J. Beachy ENTITLED Adaptive Multi-Fidelity Modeling for Efficient Design Exploration Under Uncertainty BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF Master of Science in Mechanical Engineering.

_____

Harok Bae, PhD
Thesis Director

_____

Raghavan Srinivasan, PhD
Chair, Mechanical and Materials Engineering Department

Committee on Final Examination:

_____

Edwin Forster, PhD

_____

Joy Gockel, PhD

_____

Barry Milligan, PhD, PE
Interim Dean of the Graduate School

# ABSTRACT

Beachy, Atticus J., M.S.M.E., Mechanical and Materials Engineering Department, Wright State University, 2020. Adaptive Multi-Fidelity Modeling for Efficient Design Exploration Under Uncertainty.

This thesis work introduces a novel multi-fidelity modeling framework, which is designed to address the practical challenges encountered in Aerospace vehicle design when 1) multiple low-fidelity models exist, 2) each low-fidelity model may only be correlated with the high-fidelity model in part of the design domain, and 3) models may contain noise or uncertainty. The proposed approach approximates a high-fidelity model by consolidating multiple low-fidelity models using the localized Galerkin formulation. Also, two adaptive sampling methods are developed to efficiently construct an accurate model. The first acquisition formulation, expected effectiveness, searches for the global optimum and is useful for modeling engineering objectives. The second acquisition formulation, expected usefulness, identifies feasible design domains and is useful for constrained design exploration. The proposed methods can be applied to any engineering systems with complex and demanding simulation models.

TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

Table Page

ACKNOWLEDGEMENTS

I would like to thank my thesis advisor, Dr. Harok Bae, for the opportunity to conduct research under his guidance. His mentoring and encouragement have enabled me to become a better researcher. Without his support and knowledge this work would not have been possible.

I also want to thank the members of the thesis committee, Dr. Edwin Forster and Dr. Joy Gockel. Dr. Forster provided extensive feedback on this thesis document.

This research project would not have been possible without the funding provided by DAGSI/SOCHE, AFOSR-SFFP, and an AFOSR-RQVC lab task monitored by Dr. Fariba Fahroo. I would like to thank Dr. Forster and Dr. Ramana Grandhi for serving as Air Force Lab advisers for AFOSR-SFFP.

I want to thank Dr. Daniel Clark, who provided the FEA model for the thermoelastic panel example contained herein as a fellow member of the Wright State Research Group.

Finally, I would like to thank those who provided comments and feedback on the research, including Dr. Edwin Forster and Dr. Philip Beran from AFRL-RQVC, as well as Dr. Marcus Rumpfkeil from the University of Dayton.

Dedicated to my parents

# I. RESEARCH BACKGROUND AND TECHNICAL NEEDS

This thesis lays out novel methods to reduce the time and cost of engineering design exploration when using computer simulations. The main approach is to build and use surrogate models, which inexpensively approximate computer model responses using data from a limited number of simulation runs. Multi-Fidelity (MF) surrogate modeling allows for multiple data sources of various accuracies and cost to be leveraged, allowing for increased modeling flexibility and decreased overall cost without sacrificing prediction accuracy. Adaptive sampling methods sequentially select new data samples in regions of the design space where increased accuracy is important. A novel MF surrogate modeling method is introduced, as well as two adaptive sampling methods, one for global optimization and the other for determining contours and boundaries of design feasibility. The former introduces Expected Effectiveness (EE) and is useful for capturing engineering design objectives by exploiting MF data sources, while the latter defines Expected Usefulness (EU) for modeling engineering feasibility in the design domain of interest.

## 1.1 Surrogate Modeling in Engineering Design Exploration

Computational simulations and analysis have been widely used to reduce the cost and time of engineering design exploration. To streamline the design process, the design optimization and uncertainty quantification approaches can be used to examine and mature multiple design concepts in the early stage of design development. These design exploration studies require many iterations of model evaluations, which may incur

intractable computational costs. To alleviate the computational costs, many surrogate-based design exploration methods [1-3] have been proposed. A surrogate model is a mathematical model that is constructed using data sampled from the original model. The surrogate model is then used as an inexpensive replacement of the original model for accelerated analysis. While the computational costs incurred after the surrogate model is built are typically manageable, collecting the samples to construct an accurate surrogate is often computationally prohibitive. Data-fit surrogate models include the response surface method [4], Taylor series-based approximation [5], neural network [6], reduced order modeling [7] and kriging [8-10]. However, these data-fit methods typically require many simulation samples to achieve the desired level of accuracy. The computational demands of generating many simulation samples may be challenging. The high computational costs associated with sampling complex, non-linear responses of high dimension models has motivated the development of multi-fidelity modeling.

### 1.2 Multi-Fidelity Modeling Approaches

MF modeling methods [11-18] leverage mixed data from multiple sources of different cost and accuracy to build a reliable surrogate model with reduced computational cost. The basic strategy is to use many samples from the Low Fidelity (LF) data to find the general trend of the model, while correcting the trend using a small number of High Fidelity (HF) data points. It is assumed that the HF model predicts the true system response of interest with the desired level of accuracy for the current modeling and simulation purpose. HF data can come from expensive physical tests or fully-integrated multi-physics simulations, while LF data, which are typically much cheaper than HF data, can be generated from simplified or decoupled physics-based simulations, empirical regression models, or

reduced sub-system tests. There are many different MF approaches that can be classified based on the types of sources of LF data, strategies for combining data, and applications of MF models. Peherstorfer et al. [17] divided the MF approaches of combining fidelity data into three categories: adaptation, fusing, and filtering.

The proposed MF modeling method in this thesis is based on the adaptation approach that uses surrogate models to correct the LF models using a small number of HF samples. The model corrections can be defined as multiplicative, additive, hybrid/comprehensive or space mapping. In the context of design optimization, the multiplicative corrections are often given by either constant factors or low-order regression functions to capture the global trend of the HF model [19, 20]. As for additive corrections, surrogate models such as kriging are constructed and used to compensate for the local discrepancies from the HF model. As a general approach, hybrid or comprehensive MF methods [11-16] have been developed that use both multiplicative and additive corrections. Adaptive hybrid methods [11, 12, 21] in which the additive and multiplicative corrections are combined by using a constant weight factor were developed for the applications of design optimization. The weighting factors are determined by using the previously evaluated data point within a local trust region. Han et al. [15] proposed the Generalized Hybrid Bridge Function (GHBF) to build an MF kriging model that can cover the global domain. In GHBF, the regression term formulated as a multiplicative correction is coupled with the stochastic process for the additive correction, which is determined via the usual Maximum Likelihood Estimation (MLE) method. Adopting GHBF, Rumpfkeil and Beran [22] developed the dynamic MF modeling approach that can address non-stationary HF model behaviors with an adaptive sampling scheme.

Most existing MF modeling methods assume three things [11-16]: globally correlated LF models, known hierarchical rankings for the LF models, and deterministic HF and LF data. First, it is assumed that the trend of the LF models is well correlated with the HF model over the entire design domain of interest. However, there are often more than two LF models that may provide valid correlations within different local ranges of the design domain. For example, different buckling models can be used based on different ranges of the slenderness ratio, and different flutter equations are used for subsonic, supersonic, and transonic speed ranges. The localized valid domains of LF models can be disjointed or partially overlapped. In many situations, before performing any model evaluations, it is hard to decide which LF models should be used in which local domains. Second, there are several methods of combining more than two LF models by using either sequential adaptation [22] or co-kriging regression [23]. However, it is often required that either the stationary hierarchical rank of model accuracy among the LF models be user-defined or enough samples of both HF and LF models are available to construct a valid correlation structure. The rank of accuracy is simply regarded as the same as the rank of the model fidelity, which is not always true depending on the application of the models. Lastly, the data from HF and LF models or sources are assumed to be deterministic. However, in practice almost all measurements and estimations carry some degree of uncertainty sourced from measurement randomness, modeling error, or noise in the operational conditions.

### 1.3 Adaptive Sampling of Models

To minimize the required number of samples for surrogate modeling, many studies have been performed to develop adaptive sampling and variance reduction techniques [9, 24, 25]. These methods maximize a metric called the acquisition function to determine the next

sample location for sequential surrogate modeling. The acquisition function used depends on the goal of the model, whether to develop a model accurate everywhere in the design space, to find the global optimum, or to find a contour boundary. For instance, the problem of finding a function's global optimum is addressed by the Expected Improvement (EI) concept [9, 25] which updates a kriging surrogate model by adding adaptively selected samples within the Efficient Global Optimization (EGO) framework. EI is defined as the expected value by which a stochastic kriging prediction surpasses the current best sample. This approach provides a balance between improving the kriging model's prediction while exploiting its approximation and has been successful in many applications of adaptive kriging refinement and global optimization. However, the EGO method needs user specified stopping criteria to avoid numerical overfitting. The performance and quality of EGO can vary significantly with the stopping criteria. As a variation of EI, Clark et al. [26] proposed an adaptive infill criteria that considers both aleatory and modeling epistemic uncertainties within the framework of Non-Deterministic Kriging (NDK) [27] to successfully perform EGO on uncertain data and to achieve stable convergence.

Recently, efforts have been made to develop methods that enable adaptive sampling of MF models. For example, multi-fidelity expected improvement based multi-criteria adaptive sampling has been proposed and applied to the shape optimization of a NACA hydrofoil [28, 29]. Chandhuri et al. [30] proposed an adaptive sampling strategy considering residual error, information gain, and weighted information gain. In these methods, however, the adaptive sample selection only focused on improvement of prediction model accuracy. While optimization methods for adaptively sampling MF models exist [31-34], the MF surrogates perform poorly when LF models do not have

stationary ranks of accuracy and individual LF models only capture the true trend in local regions of the design space.

The overall goal of adaptive MF sampling is allocation of limited computational resources among variable fidelity models with different computational costs in the way that best improves prediction accuracy. When performing optimization, lower fidelity models can be effective in domains with less expectation of an optimum solution and at the beginning of the sequential sampling, whereas higher fidelity models should be selected at locations of higher expectation or towards the final stages of sequential sampling. The Value-based Global Optimization (VGO) method [35] has been proposed to address this problem by using utility metrics of variable fidelity models with different costs. The VGO method uses the expected value of information in the adaptive sampling selection instead of EI. However, based on the kriging formulation, VGO needs to fit many hyperparameters to combine samples from multiple fidelity models. The fitting process requires the solution of a multi-dimensional optimization problem to find unknown hyperparameters simultaneously unlike conventional kriging in which only one-dimensional problems are needed for individual hyperparameters. This numerical fitting of multiple VGO hyperparameters can pose numerical challenges of overfitting and non-uniqueness of the fitting solution. In contrast, the multi-fidelity modeling method proposed in this thesis work has only one kernel length hyperparameter that needs to be set.

For a related problem, Multi-Information Source Optimization (MISO), a surrogate modeling method that flexibly varies LF model bias across the domain while remaining robust to noise was introduced [36]. The misoKG algorithm adaptively samples the fidelity

and location that maximizes the knowledge gradient. This method avoids the assumptions of global LF accuracy, known rank-ordered accuracies of LF models, and noiseless data.

An adaptive sampling method for contour estimation [37] provides an efficient way of predicting failure boundaries and determining feasible and infeasible regions of the design space under uncertainty. The Efficient Global Reliability Analysis (EGRA) [24] method includes an adaptive sampling scheme called the Expected Feasibility Function (EFF) with the ability to sample multiple constraints simultaneously to determine the composite feasible region. This allows for computational savings when a design can fail in multiple ways. However, the above methods only work for a single fidelity of data. A multi-fidelity contour estimation method, Contour Location via Entropy Reduction (CLoVER) [38], uses the same surrogate model as the misoKG [36] method. This method therefore avoids the assumptions of global LF accuracy, known rank-ordered accuracies of LF models, and noiseless data. While the method performs well, it is designed to handle only a single constraint at a time.

# II. RESEARCH GOALS

Based on the technical difficulties and limitations of the existing MF modeling methods for engineering design exploration, the research goals of the thesis are identified as follows:

1. Develop a MF surrogate modeling approach that performs well when individual LF models only capture the true trend in local regions of the design space, LF models do not have stationary ranks of accuracy, and noise may exist in the HF and LF data.
2. Enable adaptive MF modeling for global design optimization considering the balance between information gained and data cost.
3. Enable adaptive MF modeling for capturing the composite feasible region when multiple constraints exist.

To achieve these goals, novel modeling approaches are developed and proposed in this work including the Localized Galerkin Multi-Fidelity (LGMF) method, the Expected Effectiveness (EE) adaptive sampling method for global design optimization, and the Expected Usefulness (EU) adaptive sampling method for determining constraint failure boundaries. The LGMF method enables exploitation of an arbitrary number of non-hierarchical LF information sources and can handle noise in both LF and HF data. Additionally, the method returns uncertainty bounds and dominance information which can be used by adaptive sampling methods. The EE adaptive sampling method determines where to generate data by selecting the LF model that improves LGMF optimally in an iterative process for efficient global optimization. EE is essentially a composite metric of EI, modeling dominance, modeling uncertainty, and the cost of generating data from a LF model. The EU adaptive sampling method enables efficient updating of a composite

feasibility boundary model within the design domain of interest. EU represents data usefulness as measured by EFF, modeling dominance, and modeling uncertainty balanced against the costs of generating data from a LF model. The method can ignore inactive constraint boundaries while simultaneously considering multiple active constraints, further reducing the required number of HF samples and increasing efficiency.

Within this thesis, existing surrogate modeling methods are discussed in Chapter III. The novel methods mentioned previously are built on these exiting methods, which include Kriging, EGO and EI, EGRA and EFF, and correction-based adaptation methods. In Chapter IV, the proposed LGMF, EE, and EU methodologies are introduced in detail and demonstrated with multiple numerical examples. Finally, the summary and discussion of promising directions for future work are presented in Chapter V.

# III. EXISTING SURROGATE MODELING METHODS

This section lays out the Kriging formulation and existing adaptive sampling approaches, which are later extended to a multi-fidelity context. The samples are used to construct a Kriging surrogate, which is then used to determine the next location to sample. The EGO method is reviewed for optimization, while the EGRA method is discussed for the problem of feasibility contour estimation.

## 3.1 Kriging Formulation

Kriging was originally developed for use in geostatistics as a means of estimating the distribution of ore using samples taken from a limited number of bore holes [39]. When a Gaussian kernel is used for the kriging model, as in this thesis, it is also known as Gaussian Process Regression (GPR).

When a function is estimated from $m$ data samples, the sample locations are given by $\boldsymbol{S} = [s_1, s_2, \dots, s_m]^T$ and the sample responses are given by $\boldsymbol{Y} = [y_1, y_2, \dots, y_m]^T$. The true function $y(x)$ is treated as a realization of a stochastic process $\hat{y}(x)$, which includes a regression term $\boldsymbol{f}(x)^{\boldsymbol{T}}\boldsymbol{b}$ and a stochastic process $z(x)$,

$$\hat{y}(x) = \boldsymbol{f}(x)^{\boldsymbol{T}}\boldsymbol{b} + \mathrm{z(x)} \tag{1}$$

where $\boldsymbol{f}(x) = \left[f_1(x), f_2(x), \dots, f_p(x)\right]^T$ is the basis vector of $p$ regression functions and $\boldsymbol{b}$ is the coefficient vector of the basis functions. The stochastic process $z(x)$ is used to fit the residuals of the regression term and is assumed to have a mean of 0. The reason a stochastic process is used to model the deterministic deviations of the regression model from the true responses is that those deviations are assumed to resemble white noise for a well-chosen

regression model. The random process $z(x)$ describes epistemic uncertainty about the true deviation value and is modeled with covariance

$$COV\left[z(s_i), z\left(s_j\right)\right] = \sigma^2 R(\boldsymbol{\theta}, s_i, s_j) \tag{2}$$

where $\sigma^2$ is the mean squared error of the regression term, $\boldsymbol{\theta}$ is the model hyperparameter vector and $R$ is the correlation among sample points. This work uses a Gaussian correlation function,

$$R\left(\boldsymbol{\theta}, s_i, s_j\right) = \prod_{k=1}^{N_d} \exp\left(-\theta_k d_k^2\right) \tag{3}$$

where $d_k$ is the distance between the sample points along the $k^{th}$ dimensional direction and $N_d$ is the number of dimensions of the problem. The regression coefficients $\boldsymbol{b}$ are calculated using the least-squares method, i.e., by minimizing the mean squared error of the regression term, defined as

$$\sigma^2 = E[\left(\widehat{\boldsymbol{Y}}_{\boldsymbol{reg}} - \boldsymbol{Y}\right)^2] \tag{4}$$

where $\widehat{\boldsymbol{Y}}_{reg} = \boldsymbol{f}(\boldsymbol{S})^{\boldsymbol{T}}\boldsymbol{b}$ is the vector of predicted regression responses at the sample locations. The regression coefficients can then be derived as

$$\widehat{\boldsymbol{b}} = (\boldsymbol{F^T}\, \boldsymbol{R}^{-1}\boldsymbol{F})^{-1}\boldsymbol{F^T}\boldsymbol{R}^{-1}\boldsymbol{Y} \tag{5}$$

where $\boldsymbol{R}$ is the $m \times m$ matrix of stochastic-process correlations between z responses at the sample locations, given as

$$\boldsymbol{R}_{ij} = R\left(\boldsymbol{\theta}, s_i, s_j\right), \quad i, j = 1, \ldots, m \tag{6}$$

and where $\boldsymbol{F}$ is the $m \times p$ regression design matrix at the sample locations, given as

$$\boldsymbol{F} = [\boldsymbol{f}(\boldsymbol{s_1}),\ \boldsymbol{f}(\boldsymbol{s_2}), \ldots \boldsymbol{f}(\boldsymbol{s_m})]^{\boldsymbol{T}} \tag{7}$$

The prediction response at any point $x$ is then given by

$$\hat{y}(x) = \boldsymbol{f}(\boldsymbol{x})^{\boldsymbol{T}}\, \widehat{\boldsymbol{b}} + \boldsymbol{r}(\boldsymbol{x})^{\boldsymbol{T}}\boldsymbol{R}^{-1}\left(\boldsymbol{Y} - \boldsymbol{F}\widehat{\boldsymbol{b}}\right) \tag{8}$$

where $\boldsymbol{r}(\boldsymbol{x})$ is the correlation between the prediction location $x$ and the sample points $\boldsymbol{S}$. Under this formulation the kriging response of Eq. 8 is dependent on the regression coefficients of Eq. 5, which are dependent on the correlations $\boldsymbol{R}$ among samples of Eq. 6, which is dependent on the model hyperparameters $\boldsymbol{\theta}$. These hyperparameters must be selected to fully determine the model. Using the Maximum Likelihood Estimation (MLE) approach, the optimal correlation model parameter $\boldsymbol{\theta}^*$ for the Gaussian process is computed by solving the optimization problem

$$\boldsymbol{\theta}^* = \underset{\boldsymbol{\theta}}{\operatorname{argmax}}\, L \tag{9}$$

Where the likelihood $L$ is given by

$$L = -\frac{1}{2}\left(N_s \ln(2\pi) + \ln(|\boldsymbol{R}|) + \left(\boldsymbol{Y} - \boldsymbol{F}\hat{\boldsymbol{b}}\right)^T \boldsymbol{R}^{-1}\left(\boldsymbol{Y} - \boldsymbol{F}\hat{\boldsymbol{b}}\right)\right) \tag{10}$$

Once $\boldsymbol{\theta}^*$ is found, the mean squared error can be predicted as

$$\hat{\sigma}^2(x) = \sigma^2\left(1 + \boldsymbol{u}^T(\boldsymbol{F}^T\boldsymbol{R}^{-1}\boldsymbol{F})\boldsymbol{u} - \boldsymbol{r}(\boldsymbol{x})^T\boldsymbol{R}^{-1}\boldsymbol{r}(\boldsymbol{x})\right) \tag{11}$$

where

$$\boldsymbol{u} = \boldsymbol{F}^T\boldsymbol{R}^{-1}\boldsymbol{r}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}) \tag{12}$$

Model predictions are represented as Gaussian distributions at each point $x$, with both a prediction mean $\hat{y}(x)$ (Eq. 8) and a standard deviation $\hat{\sigma}(x)$ (Eq. 11) representing the epistemic uncertainty about the true response. This uncertainty information is invaluable when performing adaptive sampling. More detailed descriptions of the theory behind kriging, as well as the optimization process for model fitting, can be found in [8, 10, 40]. The Kriging models in this theses were built using the DACE Toolbox [40] with some modifications to the source code.

### 3.2 EGO and EI

EGO [9] was introduced to use the prediction and uncertainty information from a kriging fit to balance exploration and exploitation efficiently for global optimization. It works by sampling at the point with maximum EI, where EI is the value by which a point taken at a given sampling location can be expected to improve over the current best sample, where a worse or equal value yields an improvement of 0. This can be calculated by integrating over the portion of the prediction probability density function that extends below the current optimum, as illustrated in Fig. 1.

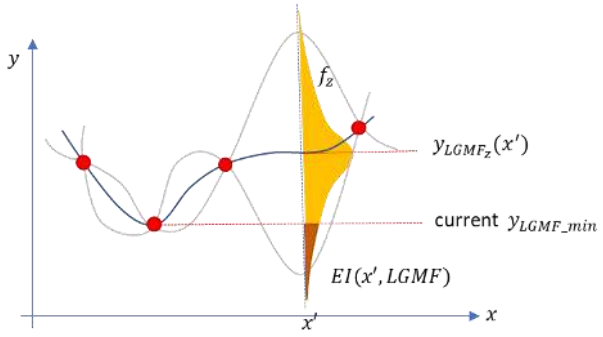



**Figure 1.** Illustration of Expected Improvement metric for adaptive sampling.

For a Gaussian distribution, the integral can be solved and the expected improvement given as a closed-form expression,

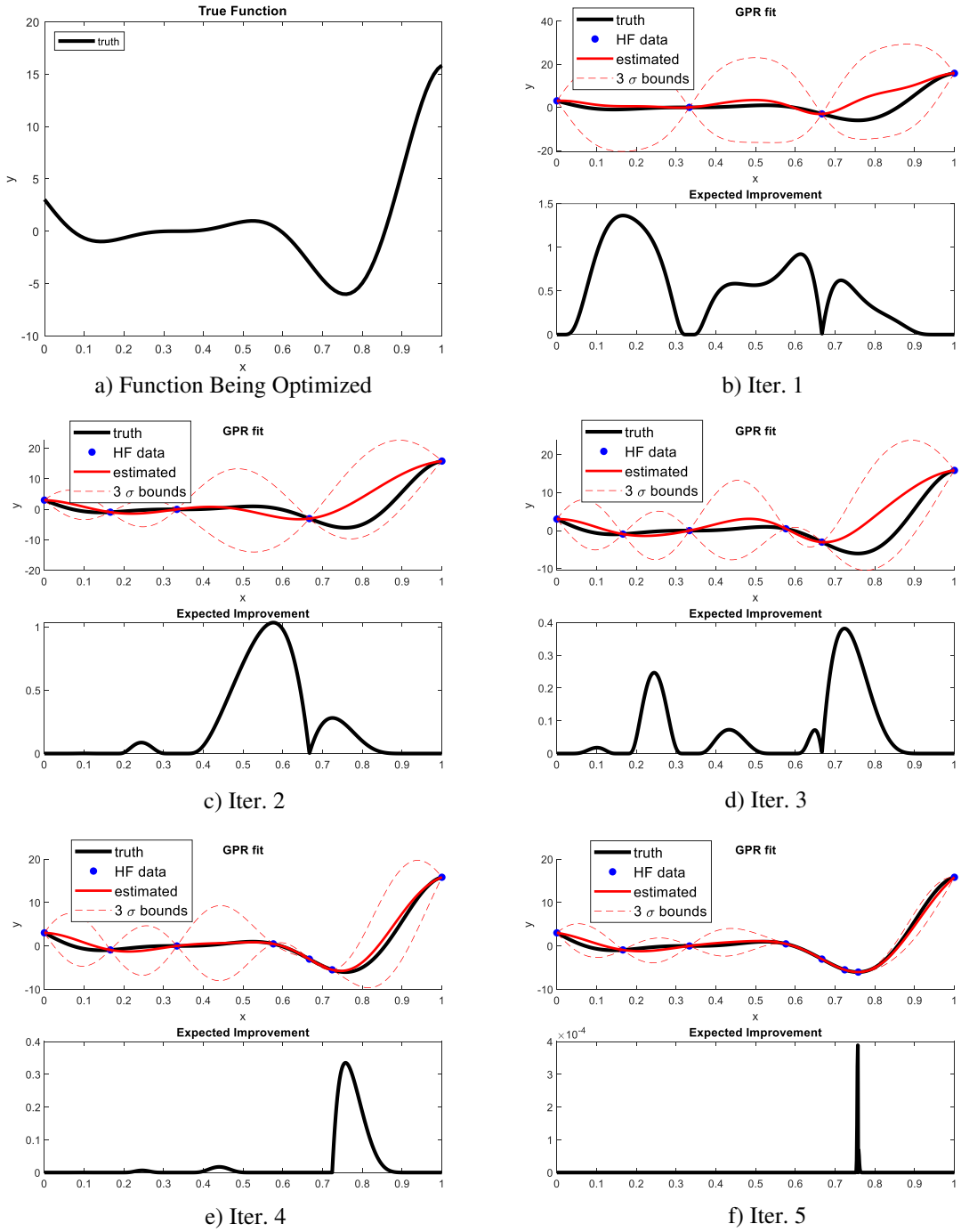
$$EI(x) = (f_{min} - \hat{y}(\mathrm{x})) * \Phi\left(\frac{f_{min}-\hat{y}(x)}{\sigma(x)}\right) + \sigma(x) * \phi\left(\frac{f_{min}-\hat{y}(x)}{\sigma(x)}\right) \quad (13)$$

where $\sigma$ is the standard deviation of the kriging estimation, $f_{min}$ is the minimum sample point found so far, and $\hat{y}$ is the kriging estimate. Also, $\phi(\cdot)$ and $\Phi(\cdot)$ are the standard normal density and cumulative distribution functions, respectively.

An example of adaptive sampling using EI is included for Eq. 14:

$$\mathrm{y}(x) = (6x-2)^2 sin(12x-4) \quad (14)$$

EGO takes five iterations to converge a kriging model initialized with four samples. The true function, along with the iterative history of the sampling and convergence, is shown in Fig. 2.




a) Function Being Optimized

b) Iter. 1

c) Iter. 2

d) Iter. 3

e) Iter. 4

f) Iter. 5

**Figure 2.** Iteration History of EGO, steps 1 to 5. (Top) Kriging estimation and confidence bounds (Bottom) Expected Improvement values across design domain

Convergence occurs when the maximum EI value drops below tolerance, in this case set to 0.001. For the first iteration (Fig. 2b) data is sparse so the uncertainty bounds are quite wide. The maximum EI value occurs between $x$ values of 0.1 and 0.2, so that is where the next sample is added. In the second iteration (Fig. 2c) the added sample has significantly reduced uncertainty in that region, and the maximum EI value occurs just below the $x$ value of 0.6. Therefore, that is where the next sample is added. This process continues until the final iteration (Fig. 2f), where the maximum EI value has dropped below $10^{-3}$. This indicates the optimum has been located and further samples are unlikely to significantly improve it.

### 3.3 EGRA and EFF

The Efficient Global Reliability Analysis (EGRA) methodology [24] was developed to evaluate the reliability of systems for engineering design. The method uses the Expected Feasibility Function (EFF) metric as an acquisition function for adaptive sampling. The samples are used to construct a kriging surrogate model, which is then used to accurately evaluate the contour boundary and therefore the feasible region. The metric balances sampling locations that are predicted to be near the failure boundary with sampling locations that have high uncertainty. The next location to be sampled is the location that maximizes the EFF. For a single constraint, the EFF is given by

$$\begin{aligned} EFF(x) = \left(\mu_g - \bar{z}\right) * \left[2\Phi\left(\frac{\bar{z}-\mu_g}{\sigma_g}\right) - \Phi\left(\frac{z^- - \mu_g}{\sigma_g}\right) - \Phi\left(\frac{z^+ - \mu_g}{\sigma_g}\right)\right] \\ - \sigma_g\left[2\phi\left(\frac{\bar{z}-\mu_g}{\sigma_g}\right) - \phi\left(\frac{z^- - \mu_g}{\sigma_g}\right) - \phi\left(\frac{z^+ - \mu_g}{\sigma_g}\right)\right] \\ + \varepsilon\left[\Phi\left(\frac{z^+ - \mu_g}{\sigma_g}\right) - \Phi\left(\frac{z^- - \mu_g}{\sigma_g}\right)\right] \end{aligned} \quad (15)$$

where $\bar{z}$ is the contour level (in this case 0), $\mu_g$ is the mean kriging estimate, $\sigma_g$ is the kriging standard deviation, $\varepsilon \propto \sigma$ (in this case set to $\varepsilon = 2\sigma$) and $z^+$ and $z^-$ are equal to $\bar{z} \pm \varepsilon$, respectively.

When multiple constraints exist, it may not be necessary to find the contours of each constraint function at all locations. Parts of the contours which appear in the infeasible regions of other constraints do not need to be accurately found. The constraints $g$ only need to be sampled until their composite failure contour is known, at which point the feasible region is fully understood. This leads to the concept of a composite expected feasibility function (CEFF), given by

$$CEFF(x) = \left(\mu_g^* - \bar{z}\right) * \left[2\Phi\left(\frac{\bar{z}-\mu_g^*}{\sigma_g^*}\right) - \Phi\left(\frac{z^- - \mu_g^*}{\sigma_g^*}\right) - \Phi\left(\frac{z^+ - \mu_g^*}{\sigma_g^*}\right)\right]$$
$$- \sigma_g^* \left[2\phi\left(\frac{\bar{z}-\mu_g^*}{\sigma_g^*}\right) - \phi\left(\frac{z^- - \mu_g^*}{\sigma_g^*}\right) - \phi\left(\frac{z^+ - \mu_g^*}{\sigma_g^*}\right)\right]$$
$$+ \varepsilon\left[\Phi\left(\frac{z^+ - \mu_g^*}{\sigma_g^*}\right) - \Phi\left(\frac{z^- - \mu_g^*}{\sigma_g^*}\right)\right] \quad (16)$$

where $\mu_g^*$ is the mean and $\sigma_g^*$ is the standard deviation of the kriging prediction that is closest to failing at $x$, that is

$$\mu_g^* = \max\left[\mu_g^i\right], \quad i = 1, \ldots, I \quad (17)$$

and $\sigma_g^*$ is the corresponding uncertainty.

### 3.4 Correction-Based Adaptation Methods for Multi-Fidelity Modeling

In many cases, there are multiple choices of simulation models to predict the response of interest with different levels of model fidelity and computational cost. It is assumed that the computational cost of an HF model evaluation is significantly higher than that of the LF models. In the adaptation approach, the correction functions, also called bridge

functions or scaling functions, can be divided into three categories: additive, multiplicative and hybrid/comprehensive corrections. The additive correction δ can be expressed as

$$\delta(x) = y_H(x) - y_L(x) \tag{18}$$

After the surrogate models of the correction function and LF model are constructed, the HF response can be approximated as a MF model by

$$y_{MF\_add}(x) = \hat{y}_L(x) + \hat{\delta}(x) \tag{19}$$

where the diacritic hat ( ˆ ) indicates a surrogate model of the function.

Similarly, the multiplicative correction $\rho$ is obtained as

$$\rho(x) = \frac{y_H(x)}{y_L(x)} \tag{20}$$

and the HF response can be approximated as a MF model by

$$y_{MF_{mult}}(x) = \hat{y}_L(x) * \hat{\rho}(x) \tag{21}$$

where the diacritic " ˆ " again indicates a surrogate of the associated function.

Popular choices for surrogate models of additive and multiplicative corrections are typically low-order response surface models and kriging under the assumption that the LF model is correlated to the HF model well enough to capture its global trend. An additive correction is effective when the majority of a LF model's prediction error is described as a translational deviation from an HF model. On the other hand, a multiplicative correction is capable of correcting incorrect trends of a LF model by scaling its response negatively. However, Gano et al. [11] found that the qualities of model adaptation via either additive or multiplicative corrections can vary depending on the problem, which motivated the development of hybrid methods [12-15]. Notionally, the two corrections are combined by using a weight factor w in the hybrid methods as,

$$y_{MF}(x) = (1 - w)\, y_{MF\_mult}(x) + w y_{MF\_add}(x) \tag{22}$$

Eldred et al. [12] proposed to determine the weight factor $w$ by matching the Multi-Fidelity model to the HF data at a nearby point $x_{old}$, such as a previous design point explored during a design optimization iteration:

$$w = \frac{y_H(x_{old}) - y_{MF\_mult}(x_{old})}{y_{MF\_add}(x_{old}) - y_{MF\_mult}(x_{old})} \tag{23}$$

When $w$ is close to 1, it means the additive correction is more accurate based on the previous design iteration history. To improve the convergence rate of an optimized design, Fischer et al. [21] proposed the Bayesian posterior updating approach to determine the weight factors of the additive and multiplicative corrections individually,

$$w_{add}^{k} = \frac{w_{add}^{k-1}\psi_{add}(x)}{w_{add}^{k-1}\psi_{add}(x) - w_{mult}^{k-1}\psi_{mult}(x)} \tag{24}$$

$$w_{mult}^{k} = \frac{w_{mult}^{k-1}\psi_{mult}(x)}{w_{add}^{k-1}\psi_{add}(x) - w_{mult}^{k-1}\psi_{mult}(x)} \tag{25}$$

where $k$ is the design iteration number and $\psi_{add}(x)$ and $\psi_{mult}(x)$ denote the model likelihood of the respective correction functions. The update of the weight factor starts with $w_{add}^{0} = w_{mult}^{0} = 1/2$. The model likelihood is defined using Eqs. 26 and 27,

$$\psi_i(x) = \left(\frac{1}{2\pi\hat{\sigma}_{i,mle}^2}\right)^{n/2} e^{-n/2} \tag{26}$$

$$\hat{\sigma}_{i,mle}^2 = \frac{\sum_{j=1}^{n}\left(y_H(x_j) - y_{MF_i}(x_j)\right)^2}{n} \tag{27}$$

where $i$ stands for either the additive or multiplicative case and $n$ denotes the number of data points available within the current trust region of design exploration. This approach was applied to fundamental mathematical and aerodynamic airfoil shape optimization problems and showed promising computational advantages over conventional optimization in terms of the required number of high-fidelity evaluations. Also, the approach

demonstrated its ability to capture the descent behavior of a HF model even when the LF model exhibited weak similarity.

Another form of hybrid method, called comprehensive correction [15], is expressed as

$$y_{MF}(x) = \alpha(x)\hat{y}_L(x) + \hat{\gamma}(x) \tag{28}$$

where $\alpha(x)$ is the generalized multiplication correction and $\hat{\gamma}(x)$ is the generalized additive correction. The additive correction $\hat{\gamma}(x)$ is constructed as a kriging model using the discrepancy samples defined by

$$\gamma_k = y_H(x_k) - \alpha(x_k)y_L(x_k), \quad \mathrm{k} = 1, \dots, N_h \tag{29}$$

In many approaches, the multiplicative correction is either a simple regression coefficient [12, 13] or kriging function [15, 21, 22]. In the comprehensive Bayesian MF method [13, 15], the multiplicative correction term also includes calibration parameters. Using the Generalized Hybrid Bridge Function (GHBF), Han et al. [15] coupled the two correction terms and determined the multiplicative low-order regression coefficients and additive hyper parameters of kriging simultaneously via the Maximum Likelihood Estimation (MLE) method. Essentially, GHBF can be viewed as universal kriging with a trend function for the multiplicative correction in the form of a low-order polynomial regression model and a stochastic process for the additive correction. The same information or data is used for GHBF as for the additive, multiplicative, and hybrid corrections. No additional information is needed, and only the formulation of the comprehensive correction is different than that of other corrections. GHBF also demonstrated its promising performance in some analytical and airfoil aerodynamic design problems. Rumpfkeil and Beran [22] developed a dynamic MF modeling approach in which both GHBF and an

adaptive sampling method are integrated to address non-stationary system responses with variable fidelity models.

# IV. PROPOSED METHODS

To address general and practical situations, the non-deterministic Localized-Galerkin Multi-Fidelity (LGMF) modeling methodology is proposed [41-42]. The method is based on two main technical processes: the consolidation of multiple LF models and the refined adaptation of the consolidated model. Non-Deterministic Kriging (NDK) [10] is also employed for the variable fidelity modeling under uncertainty. The proposed non-deterministic LGMF method is demonstrated in multiple analytical examples and a thermally coupled structural design problem.

As an extension of EI, the Expected Effectiveness (EE) [43] adaptive sampling approach is proposed for accelerated global design optimization using multi-fidelity information sources. While adaptive sampling of the HF model will be done using EI, adaptive sampling of LF models will be done using EE. EE performs sequential LGMF modeling, selecting which fidelity model to evaluate every iteration and where to achieve computational cost savings and alleviate computational challenges. This is achieved by basing EE on EI, while also accounting for the Modeling Uncertainty (MU), Dominance under Uncertainty (DU) and cost of each LF model.

For a design problem with multiple failure modes and constraints, the existing Expected Feasibility Function (EFF) [24] performs well for adaptive sampling of HF data only. To exploit MF information sources using adaptive sampling, the Expected Usefulness (EU) method is proposed. During contour estimation, the EFF will be used to sample the HF model while EU will be used to sample the LF models. EU is based on the Expected Feasibility Function (EFF), but like EE accounts for MU, DU, and cost. In this chapter, the

three proposed methodologies are introduced and demonstrated with fundamental and application examples to show how the aforementioned technical gaps can be addressed.

## 4.1 Localized Galerkin Multi-Fidelity (LGMF) Modeling

The proposed Localized-Galerkin Multi-Fidelity (LGMF) modeling methodology is based on two main technical processes: the consolidation of multiple LF models and the refined adaptation of the consolidated model. Non-Deterministic Kriging (NDK) [10] is also employed for the variable fidelity modeling under uncertainty. The proposed non-deterministic LGMF method is demonstrated in multiple analytical examples and a thermally coupled structural design problem. The following sections review the existing correction-based adaptation methods, describe the proposed framework of non-deterministic LGMF, and present numerical examples to demonstrate the characteristics and prediction performance of the proposed method.

### *4.1.1 Proposed Localized Galerkin Multi-Fidelity (LGMF) Modeling*

#### A. Framework of LGMF

The proposed LGMF prediction model is expressed as a weighted sum of basis functions,

$$y_{LGMF}(x) = \sum_{i=1}^{M} c_i(x)\,\eta_i(x) \tag{30}$$

where $M$ is the total number of basis functions of consideration, $\eta_i(x)$ is the $i^{\text{th}}$ basis function and $c_i(x)$ is the participation function of the $i^{\text{th}}$ basis function. Generally, the basis functions can be derived from the LF models using additive, multiplicative, hybrid, or any other correction. For example, by using the multiplicative correction, a basis function can be derived from a LF model as

$$\eta(x) = \hat{\rho}(x)\hat{y}_L(x) \tag{31}$$

The expansion form allows as many models as are available to be considered in the MF model adaptation. To determine the model participation function at a prediction location $x_p$, the localized Galerkin equations are formulated as

$$\int \phi_j(x, x_p)\left(y_{LGMF}(x) - y_H(x)\right)dD = 0, \quad j = 1, \dots, M \tag{32}$$

where D $\in \mathbb{R}^n$ is the design domain of interest and $\phi_j(x, x_p)$ is the $j^{th}$ locally weighted test function at $x_p$ defined as

$$\phi_j(x, x_p) = \omega(x, x_p, h)\eta_j(x) \tag{33}$$

Here, $\omega(x, x_p, h) = e^{-\frac{1}{2}\left(\frac{x-x_p}{h}\right)^2}$ is the Gaussian kernel in which the shape parameter $h$ of the kernel function is determined by the density of HF samples and expected HF nonlinearity within the design domain. By replacing $y_{LGMF}$ and $\phi_j$ with Eqs. 30 and 33, the Galerkin equations become

$$\int \omega(x, x_p, h)\eta_j(x)\left(\sum_{i=1}^{M} c_i(x)\eta_i(x) - y_H(x)\right)dD = 0, \quad j = 1, \dots, M \tag{34}$$

Since $y_H$ is known only at $N_h$ HF sample locations, the integral of Eq. 34 can only be evaluated and aggregated at those locations as

$$\sum_{k=1}^{N_h} \omega(x_k, x_p, h)\eta_j(x_k) \sum_{i=1}^{M} c_i(x_k)\eta_i(x_k) \; - \;$$
$$\sum_{k=1}^{N_h} \omega(x_k, x_p, h)\eta_j(x_k)\mathrm{y}_H(x_k) = 0, \;\; j = 1, \dots, M \tag{35}$$

The equation above can be expressed in matrix form as

$$\boldsymbol{Hc} = \boldsymbol{y} \tag{36}$$

where

$$\boldsymbol{H} \in \mathbb{R}^{M\times M} \text{ with } H_{ji} = \sum_{k=1}^{N_h}\sum_{i=1}^{M} \omega(x_k, x_p, h)\eta_i(x)\eta_j(x_k), \;\; i, j = 1, \dots, M \tag{37}$$

$$\boldsymbol{y} \in \mathbb{R}^{M\times 1} \text{ with } y_j = \sum_{k=1}^{N_h} \omega(x_k, x_p, h)\eta_j(x)y_H(x_k), \;\; j = 1, \dots, M \tag{38}$$

and $c \in \mathbb{R}^{M\times 1}$ is the participation vector at $x_p$. By solving Eq. 36, the participation factors are determined at the prediction location and they are plugged back into Eq. 30 to complete the LGMF prediction. Without requiring user-defined ranks of fidelity or accuracy, the degrees of local dominance of LF models are estimated based on the participation factors that are obtained mathematically by solving the locally weighted Galerkin equations. In this study, to address the possible practical situations aforementioned, the proposed framework is applied to build the MF model by the following two main stages:

1. ***Consolidation of multiple LF models***: Consider multiple LF models that are valid in different local ranges of the design domain of interest. Each of the LF models captures the HF trend within a local range better than the other LF models. In this stage, the goal is to consolidate the multiple LF models into a single representative model that can capture the global trend of the HF model, while identifying individual correlations of the LF models to the HF model. To achieve this goal, the LF functions are corrected to obtain basis functions, which are then consolidated into a single function. The basis functions are defined with a single type of correction function, either additive or multiplicative. In this study, additive corrections of the LF models are selected, and the basis functions are defined as

$$\eta_{add,i}(x) = \hat{y}_{Li}(x) + \hat{\delta}_i(x) \qquad i = 1\sim M \tag{39}$$

where $M$ becomes the total number of available LF models. By setting up and solving the localized Galerkin equations (Eqs. 36-38) for $c_i(x)$, a single Consolidated LF (CLF) model is obtained:

$$\hat{y}_{CLF}(x) = \sum_{i=1}^{M} c_i(x)\, \eta_{add,i}(x) \tag{40}$$

Since only the additive corrections are used as the basis functions, the CLF model can be viewed as a combination of linearly translated LF models based on their local correlations to the HF samples. The differences of the participation functions $c_i(x)$ of the multiple LF models can be directly interpreted as a map of LF model dominance within the design domain. It is possible to use the other form of corrections such as multiplicative or even combined, but it will need an additional conversion function to extract the LF model dominance information. The quality of $y_{CLF}(x)$ depends on how many HF samples are available and how well the LF models capture the global trend of HF in a combined way.

2. ***Refined Adaptation of the consolidated model as the resulting LGMF model***: As pointed out by many previous researchers [11, 15], additive corrections are not always good enough without multiplicative ones. Unlike other hybrid or comprehensive MF models, the CLF model from the previous stage does not interpolate the HF samples exactly because the participation factors are obtained by minimizing the residual between HF and LGMF with the locally weighted test function $\phi$ in Eq. 33. Therefore, in this second stage, the CLF model is used as a new single LF model and further refined. The new basis functions are derived from the CLF model, i.e., multiplicative and additive or hybrid corrections. In this study, the basis functions are the multiplicative and additive corrections derived from the CLF model.

$$\eta_{add}(x) = \hat{y}_{CLF}(x) + \hat{\delta}_{CLF}(x) \tag{41}$$

$$\eta_{mult}(x) = \hat{\rho}_{CLF}(x)\hat{y}_{CLF}(x) \tag{42}$$

Here $\hat{\delta}_{CLF}(x)$ and $\hat{\rho}_{CLF}(x)$ are surrogates for the correction functions defined in Eqs. 18 and 20, and $\hat{y}_{CLF}(x)$ is defined in Eq. 40. The resulting LGMF model is obtained as,

$$y_{LGMF}(x) = c_{add}(x)\eta_{add}(x) + c_{mult}(x)\eta_{mult}(x) \tag{43}$$

Since the major adaptations were already performed during the previous stage, only small refinements are needed to finalize the resulting LGMF model.

A critical factor for the creation of the LGMF model is selection of a suitable shape parameter $h$ for the Gaussian kernel function, which may vary based on the number and layout of the HF and LF samples. Like kriging, where the stochastic process is defined by hyper-parameters that must be optimized, LGMF seeks to determine the optimal parameter $h$ to build the non-deterministic prediction model that will best capture the HF behavior. The Maximum Likelihood Estimation (MLE) method is a popular approach to fit a process model parameter to non-deterministic data. However, unlike kriging, the LGMF function does not have an explicit function form with the kernel function. Also, since multiple basis functions derived from individual LF models are involved, the underlying true kernel process can be regarded as non-ergodic, which makes the MLE approach inappropriate. Instead of MLE, the Cross Validation (CV) approach was used to optimize the shape parameter with the Leave-One-Out (LOO) criteria by formulating an optimization problem in which the sum of the squared errors at each HF point during the LOO process is minimized

$$h_{cv} \in \underset{h \in H}{\operatorname{argmin}} \sum_{k=1}^{Nhi} \left(y_k - y_{LGMF,k,-k}\right)^2 \tag{44}$$

subject to the constraint that no more than one outlier per point left out, or 1% outliers, whichever is greater, are allowed from the LOO process as

$$\frac{\sum_{k}^{Nhi}\sum_{i}^{Nh} I(i,k)}{Nh*Nhi} \leq \max\left(\frac{1}{N_h}, 0.01\right) \tag{45}$$

where $y_{LGMF,k,-k}$ is the conditional estimation at $x_k$ with LGMF built by leaving the $k^{\text{th}}$ HF sample out, $N_{hi}$ is the number of interior high fidelity samples (as boundary samples are not left out), and $I(i,k)$ is an indicator function that determines whether a point is an outlier (falls outside of the ±3 standard deviation uncertainty bounds) and is defined by

$$I(i,k) = \begin{cases} 1 & if\ \left|y_k - y_{LGMF,k,-i}\right|^2 - 9\sigma^2_{LGMF,k,-i} > 0 \\ 0 & otherwise \end{cases} \tag{46}$$

The optimization via CV LOO could be nontrivial because the objective function is often multimodal. In this study, a general multiple starting point search strategy was used by calling the sequential quadratic programming solver in Matlab at each starting point. The optimized shape parameter enables the LGMF model to capture the HF samples within the prediction uncertainty bounds while making an accurate mean prediction. It was found [44] that CV achieves better and more robust fitting than MLE especially when the underlying process is not well matched with the Gaussian-family covariance structure.

**B. Non-Deterministic Kriging for LGMF**

When the evaluation costs are trivial, the LF models can be used directly. However, in most cases it is more computationally efficient to build surrogate models such as kriging from a finite number of LF training samples. In this study, it is assumed that the samples from both HF and LF models can carry some degree of uncertainty sourced from either modeling uncertainty or natural randomness in the environmental and operational conditions. When the samples are under non-stationary uncertainty, deterministic kriging is prone to fail to model physically meaningful behaviors due to the interpolation requirement. Counterintuitively, the modeling failure gets worse as more samples are

added. Non-Deterministic Kriging (NDK) [10] provides a flexible framework that can properly capture both the means and non-stationary variances of prediction from data samples under uncertainty. To accommodate non-deterministic samples, NDK is formulated as in Eq. 47 by combining the global trend function with the realizations of two stochastic processes, $z_E(x)$ and $z_A(x)$ that represent epistemic and aleatory uncertainties respectively.

$$\hat{y}_{nd}(x) = \boldsymbol{f}(\boldsymbol{x})^{\boldsymbol{T}}\boldsymbol{\beta} + z_E(x) + z_A(x) \tag{47}$$

Here, $\boldsymbol{f}(\boldsymbol{x}) = [f_1(x), f_2(x), \dots, f_p(x)]$ are the vector of known $p$ basis trend functions of $\boldsymbol{x}$, and $\boldsymbol{\beta}$ is the regression coefficient vector. Epistemic uncertainty ($z_E$) comes from lack of confidence in interpolation modeling due to limited or missing data, which can be reduced by adding more data and information. On the other hand, natural and irreducible randomness, such as a measurement error or statistical distribution of material property, is modeled as aleatory uncertainty ($z_A$). It is noted that when random samples are too small to obtain accurate statistical inference, one can say that $z_A$ has both epistemic and aleatory uncertainties. In this case, adding more samples will reduce the epistemic uncertainty in $z_A$ and make the statistical distribution more accurate. When statistical information is available along with the training samples, the statistical information can be directly used as aleatory uncertainty in NDK. In the NDK framework, the first step is to estimate the aleatory variances $z_A$ at each data point using local polynomial regression. Then, the epistemic modeling uncertainty is determined by fitting hyper-parameters in $z_E$ through the MLE approach. Within the LGMF framework, both the LF and CLF models are modeled with NDK by generating and using a finite number of training samples under

uncertainty. Assuming that the aleatory uncertainties are independent and identically distributed among the multiple LF models, the aleatory uncertainty of the resulting LGMF model is estimated as

$$Z_{A\_LGMF}(\mathrm{x}) = \left(\sum_{i=1}^{M}\left(nc_i(x)\, Z_{A\_\eta i}(x)\right)^2\right)^{1/2} \tag{48}$$

where $Z_{A\_\eta i}$ is the aleatory uncertainty of the $i^{th}$ basis function and $nc_i$ is the normalized participation ratio. Therefore, the resulting LGMF model can provide not only the expected prediction mean but also the uncertainty bounds of the prediction. According to the application or goal of MF modeling, the prediction model can be designed to provide either optimistic risk-taking or conservative risk-averse predictions. Also, the quantified uncertainty bounds can be useful in an adaptive sampling strategy for efficient model updating.

*4.1.2 Numerical Examples*

In this section, numerical examples are presented and the performance of the proposed LGMF method is discussed. Under the assumption that the evaluation cost of an HF model is significantly higher than that of a LF model, the computational costs of building an MF model are determined mainly by the costs of generating the HF evaluation samples. Throughout the examples presented in this section, kriging models utilize the stationary Gaussian kernel function, and the hyperparameters are determined based on MLE. When samples are under uncertainty, NDK models are constructed using a finite number of samples for the corrected basis and CLF functions in LGMF modeling. As in practical situations, multiple LF models are considered in demonstrations which 1) may be only locally correlated to the HF model, 2) may provide estimations with different levels of

uncertainty, and 3) may show inconsistent or indefinite ranks of fidelity and accuracy within the design domain.

Example 1: Deterministic One-Dimensional Mathematical Problem with one LF model

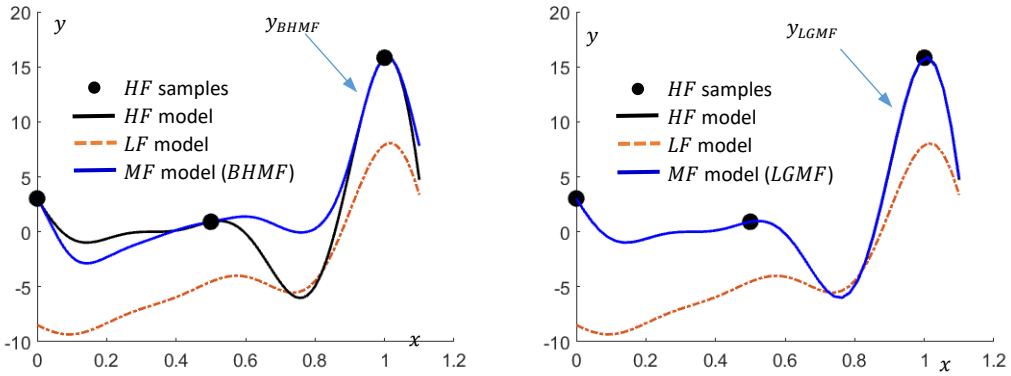
The proposed LGMF method is demonstrated with the deterministic mathematical problem shown in Fig. 1. This problem, which has been often discussed in the literature [1, 15, 18, 22], is presented to compare the proposed LGMF and existing Bayesian Hybrid Multi-Fidelity (BHMF) [21] methods. The LF model is created by scaling the HF value by a constant term and adding a linear deviation as

$$y_H(x) = (6x - 2)^2 sin(12x - 4) \tag{49}$$

$$y_L(x) = 0.5y_H(x) + 10(x - 0.5) - 5 \tag{50}$$

where $x$ is a design variable with the range [0~1.1]. As an extreme case, only three samples of the HF model at $x = \{0, 0.5, 1.0\}$ are considered. In this example, the BHMF method combines the additive and multiplicative corrections by obtaining the constant weighting factors at the fixed location $x$=0.5 as a previously evaluated data point. In BHMF, Gradient Enhanced Kriging (GEK) is also used for both adaptation functions to improve the model accuracy, assuming that the gradient information is available along with the function evaluation value with small additional cost. As for LGMF, the conventional kriging models without using the gradient information are built for the basis functions and CLF. Since there is only one LF model, the first stage of LGMF can be skipped and the second stage begins directly. The results are shown in Fig. 1. The BHMF model (Fig. 3a) shows large discrepancies even through the gradient information was included while the proposed LGMF model (Fig. 3b) is almost overlapped with the HF model. It was expected BHMF would have larger discrepancies than LGMF because BHMF uses constant weighting

factors. On the other hand, the proposed LGMF method updates the participation factors of the basis functions at each prediction location, allowing the linear deviation term of the LF function to be corrected exactly. Similar performance to LGMF was observed when the Bayesian Compressive Model (BCM) or Generalized Hybrid Bridge Model (GHBM) were applied [15, 18]. Note that simple additive and multiplicative corrections were used in this example, but any other adaptation models, such as the BCM or GHBM, can be incorporated as additional basis functions in the proposed LGMF framework.




a) BHMF with GEK b) Proposed LGMF with kriging

**Figure 3.** One dimensional example comparing the BHMF and LGMF methods

Example 2: One-Dimensional Mathematical Problem with Two Locally Correlated LF Models

This example introduces the practically possible situation of two locally-correlated LF models for the same HF model as the previous example. The two LF models are given as $y_{LF1}(x) = 1.5\sin(8x-4) + 5(x-0.5) - 5$ and $y_{LF2}(x) = -6\sin(8x-4) - 7$, as shown in Fig. 4.

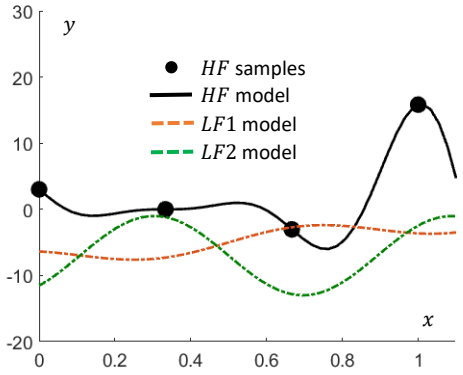
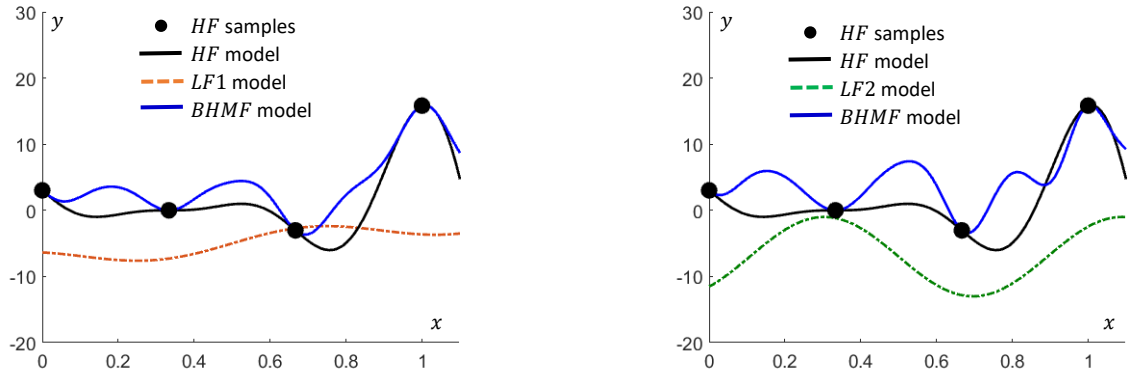



**Figure 4.** One dimensional HF model with two locally correlated LF models

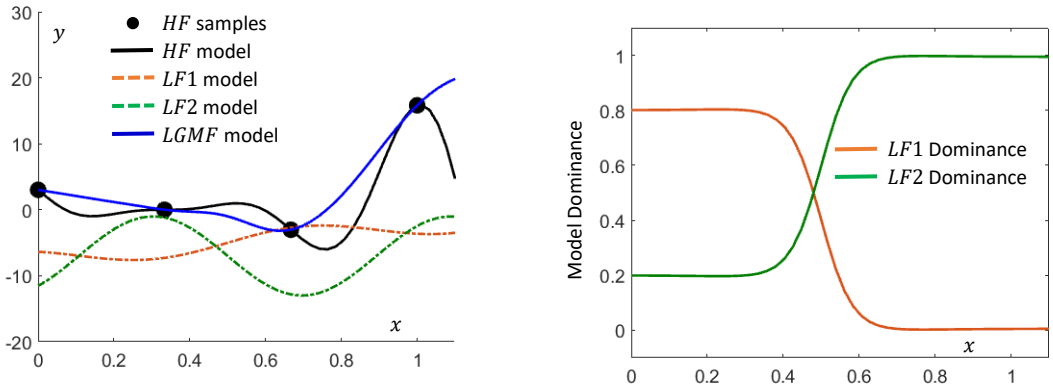
Neither the LF1 nor LF2 model is globally correlated to the HF model. However, based on the similarity of the global trends, LF1 is observed to have better correlation to HF than LF2 in the first half of the domain, $x \in [0.0{\sim}0.6]$ while LF2 correlates better than LF1 in the second half of the domain, $x \in [0.6{\sim}1.0]$. The LF models have their own valid application ranges, which are partially overlapped. Here, four samples of the HF model, collected at $x = \{0, 0.333, 0.667, 1.0\}$, are considered. This example illustrates a practical situation where 1) the boundaries of the valid application ranges of the multiple LF models are unknown and 2) the ranks of accuracy among the LF models are not given. For comparison, the BHMF models are built by choosing only one of the LF models as shown in Fig. 5. It is obvious that none of the BHMF models are accurate because the LF models are only partially correlated, not globally. On the other hand, the proposed LGMF model shown in Fig. 6a is constructed by combining both LF models and provides a more meaningful prediction model than BHMF. The shape parameter in the first stage of LGMF was optimized to $h$=1.1. Along with the prediction model, the LF model dominance information is obtained from LGMF as shown in Fig. 6b.




a) BHMF with LF1 b) BHMF with LF2

**Figure 5.** BHMF models with two individual LF models with four HF samples

The model dominance, which is essentially derived from the modal participation functions, can be interpreted as describing how the correlations of the LF models change over the design domain in a quantitative way. The model dominance is assessed between zero-dominance (0) and full-dominance (1). The model dominance plot in Fig. 6b shows the behavior of local correlations between the two LF models, which is consistent with the earlier observations about the LF and HF models in Fig. 4. Starting from the position $x$=0, LF1 is better correlated to HF and more heavily weighted in the LGMF model than LF2, with partial dominance ratios of 0.6 to 0.4, respectively. After the tipping point (x=0.55), LF2 becomes more dominant than LF1, becoming almost totally dominant after around x=0.7.

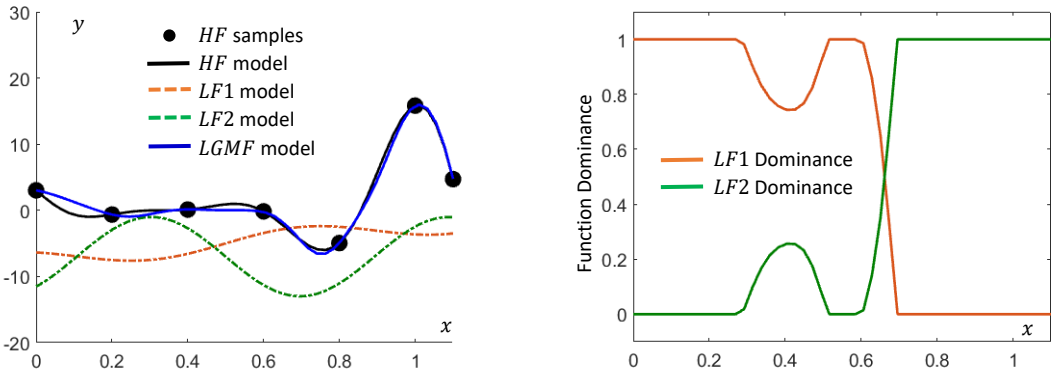



a) LGMF with four HF samples

b) LF model dominance from LGMF

**Figure 6.** LGMF modeling with two LF models that are locally correlated to HF

With the number of HF samples increased to seven and using the optimum shape parameter of $h$=0.8, the LGMF model and LF dominance become accurate and clear as shown in Fig. 7. The dominance information can be useful in understanding the characteristics of the HF model behavior in terms of the fundamental local behaviors described by the LF models. Along with other data classification or clustering algorithms which are purely based on data samples, the model dominance information can be used to enable a physics-based data clustering. In another imminent application, the dominance information can guide us in identifying which local models or transition ranges need to be interrogated with more samplings in an adaptive sampling and model updating scheme.

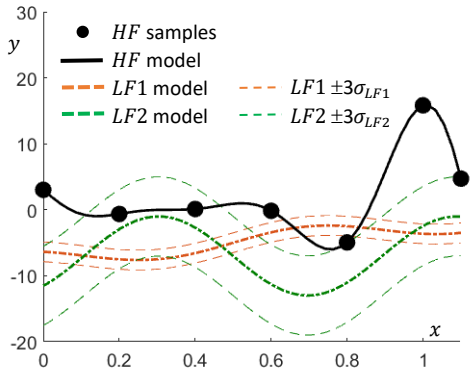



a) LGMF with seven HF samples

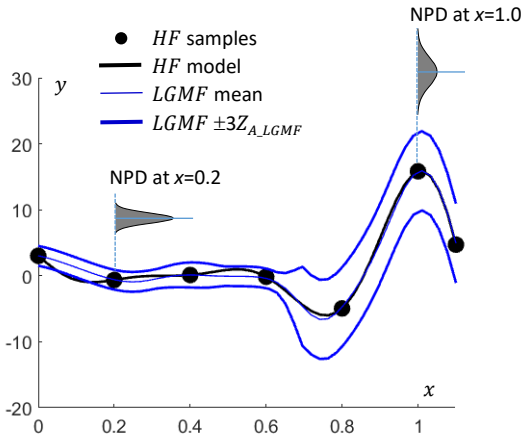
b) LF model dominance from LGMF

**Figure 7.** LGMF modeling with seven HF samples

As a case with modeling under uncertainty, consider non-deterministic LF models with provided prediction means and constant standard deviations ($\sigma_{LF1} = 0.5$ and $\sigma_{LF2} = 2.0$ ) as shown in Fig. 8a. Taking $\sigma_{LF1}$ and $\sigma_{LF2}$ as the aleatory uncertainties of the LF models, the aleatory uncertainty of LGMF, $Z_{A\_LGMF}$ is estimated based on Eq. 48, which is used to suggest the prediction uncertainty bounds as $\pm 3Z_{A\_LGMF}$ as shown in Fig. 8b. As expected, the combined uncertainty bounds are narrow in the range of LF1 dominance and become larger in the LF2 dominance range, which presents the possible prediction errors of LGMF with a non-stationary normal distribution.




a) Non-deterministic LF models




b) Prediction uncertainty bounds and Non-stationary Prediction Distribution (NPD) of LGMF

**Figure 8.** Non-deterministic LGMF model with prediction uncertainty bounds

Example 3: Non-Stationary Mathematical Problem with Non-Deterministic LF Models

This mathematical example addresses an HF model with a non-stationary response, i.e., the response behavior changes abruptly at the $C^0$ continuity. The HF model shown in Fig. 9 is defined as

$$y_H(x_1, x_2) = \max\left[f_1,\ f_2\right] \tag{51}$$

where $f_1$ and $f_2$ are given by

$$f_1(x_1, x_2) = \exp({x_1}^2 + \ x_2) - 7 \tag{52}$$

$$f_2(x_1, x_2) = -(1.7(x_1 + 0.3) - 2)^2 \sin(x_1 x_2 \pi) * (x_2 + 1) \tag{53}$$

As shown in Fig. 9, the HF behavior changes between $f_1$ and $f_2$ along a nonlinear boundary, resulting in discontinuous derivatives.

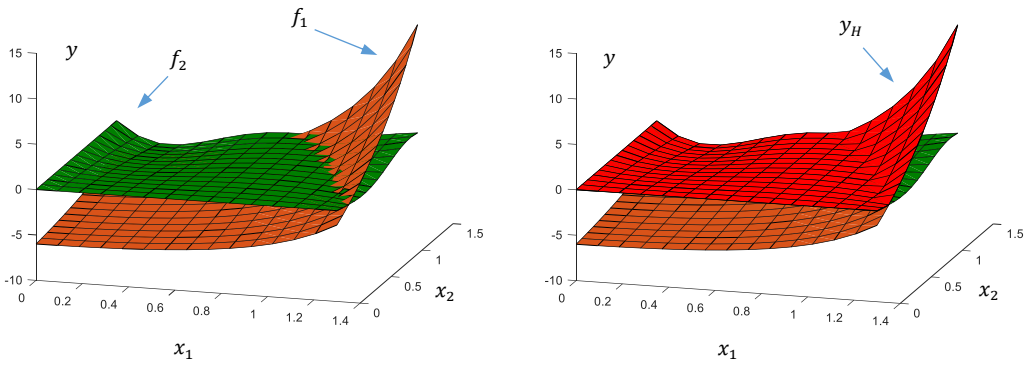



**Figure 9.** Non-stationary HF model

Low-fidelity models are available, which are non-deterministic to account for the effects of natural randomness and incomplete modeling. The two non-deterministic LF models are defined with both linear and nonlinear deviation terms, along with random function terms, as

$$y_{LF1}(x_1, x_2) = 3 + f_1(x_1, x_2) - x_1 - x_2 + x_1 x_2 + 0.2\xi(x_1 + x_2) \quad (54)$$

$$y_{LF2}(x_1, x_2) = 1 + f_2(x_1, x_2) + x_1 x_2 - (x_2 - 0.3)^2 + 0.25\xi \quad (55)$$

where $\xi$ is the standard normal random variable. LF1 is designed to have increasing uncertainty bounds with increasing $x_1$ and $x_2$, while LF2 has constant randomness within the domain. In this example, to simulate a realistic situation, only limited numbers of random samples are taken from the LF models. Using noisy LF samples, NDK models for LF1 and LF2 were constructed for use in the LGMF process.

NDK provides a flexible framework for handling both randomly collected samples directly and statistical samples that come with the estimated mean and standard deviation at every data location. Here, the NDK of LF1 is built with 26 statistical samples with prior uncertainty information, while the NDK of LF2 is modeled with 377 randomly collected

samples over the entire space. Since the model evaluation cost is significantly cheaper for LF models than for HF, it is assumed that the large number of LF2 samples is manageable.

In Figs. 10 and 11, the distributed samples are shown along with the LF mean surfaces. Using the non-deterministic samples, NDK models are created and compared against the mean and uncertainty bounds ($\pm 3\sigma$) of the true LF models. NDK of LF2 seems more accurate in predicting both mean and uncertainty bounds than LF1 because of the well-populated random samples. As for LF1, there are some areas with a lower density of samples, which causes the inaccuracy in the NDK predictions. Adding more samples of the LF model will improve the prediction accuracy.

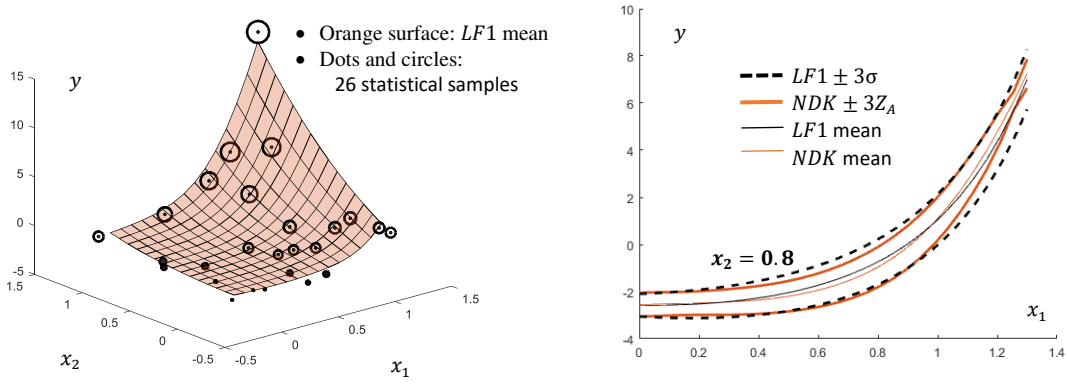



**Figure 10.** NDK with twenty-six statistical samples from LF1

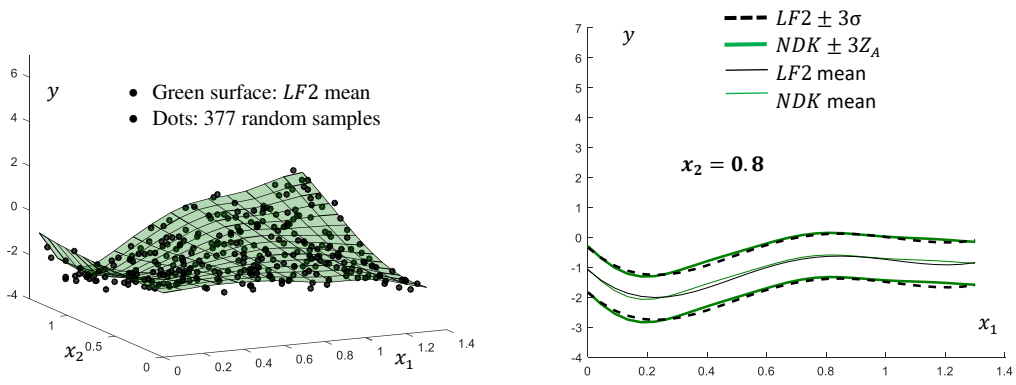



**Figure 11.** NDK with 377 random samples from LF2

The first trial of building the LGMF model is performed with seven Latin Hypercube Sampling (LHS) HF samples and the LF NDK models as shown in Fig. 12. The LGMF model obtained and shown in Fig. 13 is clearly more accurate than the kriging model built using only the HF samples.

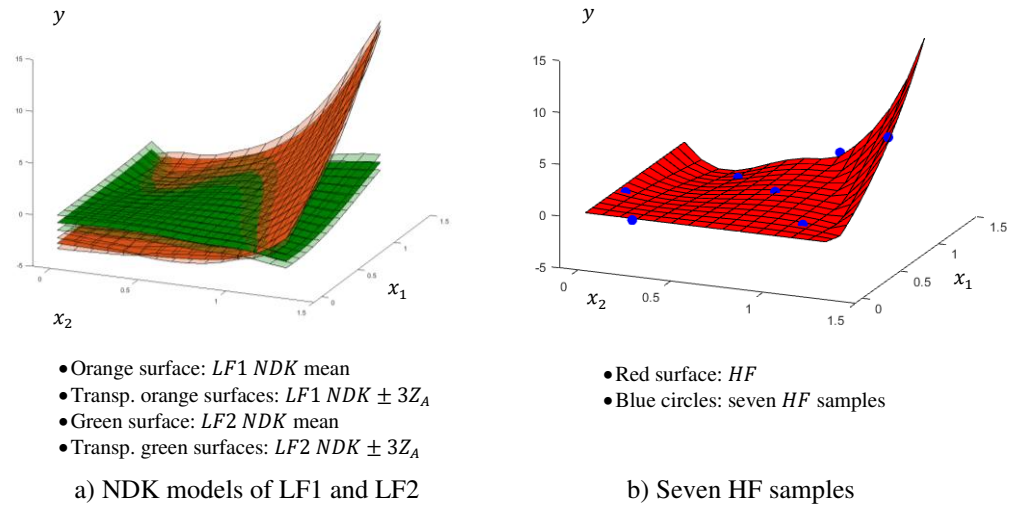



**Figure 12.** LF NDK models and HF function with seven samples

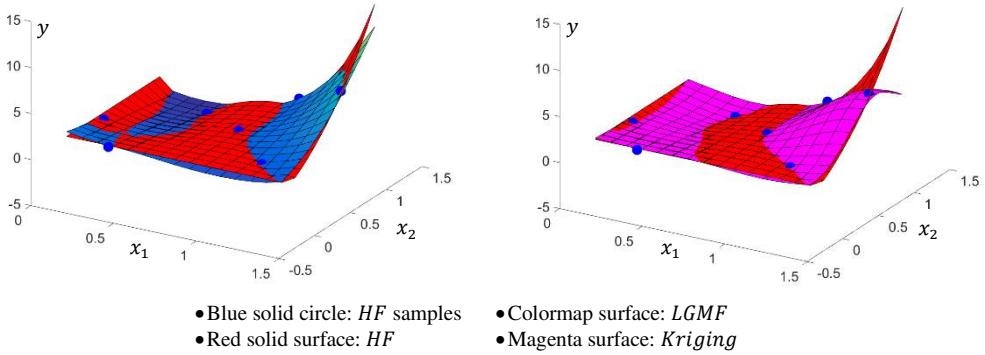
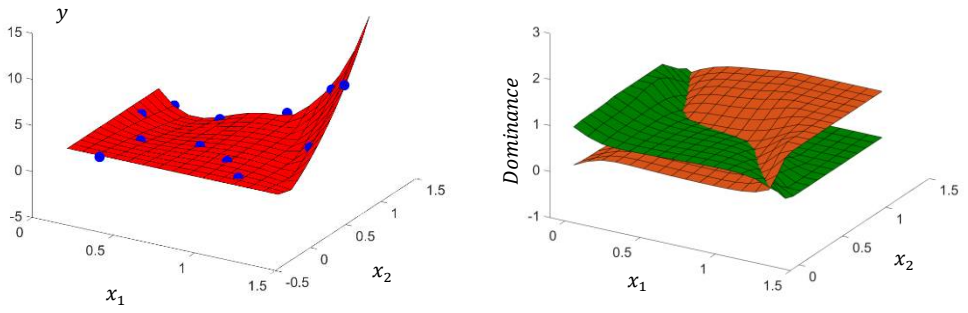



**Figure 13.** LGMF and kriging models built using seven HF samples

As another case, the LGMF model is created with twelve HF samples shown in Fig. 14a. It is found that the LF1-NDK model is overall more dominant than the LF2-NDK model. The dominance of LF1-NDK was expected because the global trend of LF1-NDK is better correlated to HF than LF2-NDK. Increasing the number of samples enables us to

capture the dominance boundary more accurately as shown in Fig. 14b. Comparing the resulting LGMF and kriging models shown in Fig. 15, it is apparent that the LGMF model is more accurate than the kriging model. Due to the noise in the data, the accuracy of the kriging model is degraded when increasing the number of points.




a) 12 HF samples b) Model dominance obtained from LGMF

**Figure 14.** HF model with 12 samples and LF dominance

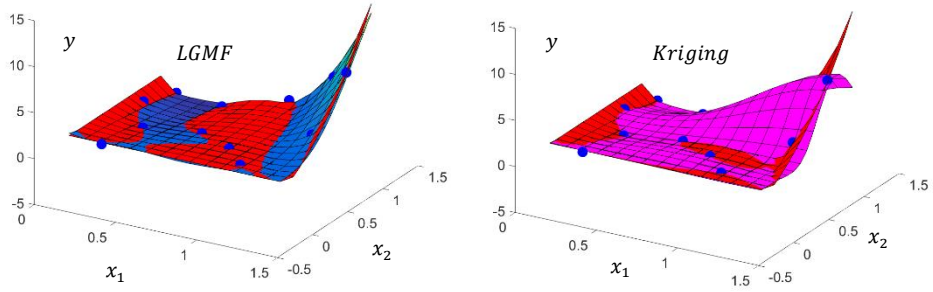



**Figure 15.** LGMF and kriging models with 12 HF samples

The model dominance information shown in Fig. 14b captures the general shape of the non-stationarity boundary shown in Fig. 12. When the LF dominance boundary is known a priori, LGMF can be built accurately with a small number of HF samples by skipping the LF consolidation stage. This is essentially the same as having a single LF model that is

globally well correlated to the HF model. To capture nonlinear deviation, it will still be necessary to determine the optimum locations of samples from the HF model to maximize the LGMF performance. There are many adaptive sampling schemes [45] in which Expected Improvement (EI), mean squared errors, or confidence bounds are measured to determine the sequential sample locations in an iterative process.

Example 4: Non-Deterministic Prediction Model of Thermally Coupled Aircraft Structure Response

In this example, an idealized model of a thin exhaust-washed structure is considered. The structure is located aft of an embedded aircraft engine and can exhibit geometric nonlinear responses due to extreme thermal loads. It was found by Deaton and Grandhi [46] that the geometric nonlinearity resulting from an elevated thermal load makes a significant contribution to the overall structural response through stress stiffening behavior and deformation-dependent load redistributions. To investigate the characteristic behaviors, the fundamental curved strip model was developed as an idealization of a thin-shell thermal structure as shown in Fig. 16.

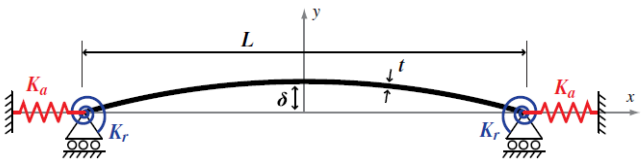
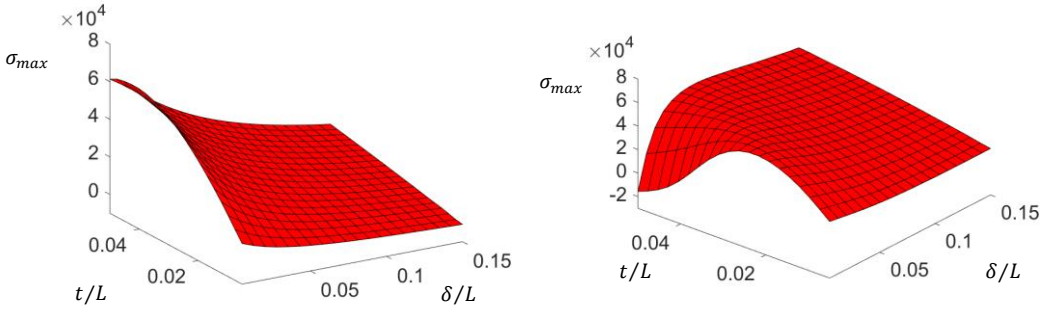



**Figure 16.** Fundamental curved strip model under extreme thermal load

The material properties are $E = 12.5 \times 10^6 psi$ and $\alpha = 5.5 \times 10^6/°F$ for Young's modulus and the coefficient of thermal expansion, respectively. The model is parameterized using the thickness-to-span-length ratio $t/L$ and the curvature-to-span-length ratio $\delta/L$. The finite element model is created with 250 two-node beam elements

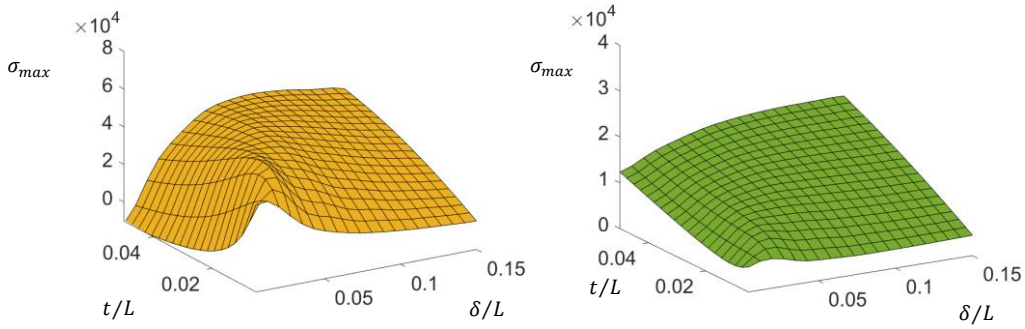
and analyzed using MSC Nastran under a thermal load of 900°F applied uniformly over the structure. To capture the effects of thermal expansion on the boundary conditions, linear elastic boundaries consisting of axial ($K_a$) and rotational ($K_r$) elastic members are used. By changing the stiffness values of the elastic members, the boundary conditions can be adjusted to any case between fully clamped and weakly constrained in a continuous manner. The maximum stress responses for two example boundary conditions are shown in Fig. 17. The following characteristic behaviors were discussed in [46]: 1) the effect of varying thickness on geometric nonlinearity appears to be small compared to the other factors, 2) increasing the curvature reduces the effect of geometric nonlinearity, and 3) the effect of geometric nonlinearity becomes sensitive when both $K_a$ and $K_r$ increase. Most importantly, it was observed that the trend predicted by a linear analysis is the opposite of the trend of the nonlinear structural responses, especially for a small curvature structural configuration. For more details of the curved strip model and discussions of its nonlinear responses, the reader is referred to [46].




a) Fixed BC ($K_a = \infty$, $K_r = \infty$)  b) Rotation-free BC ($K_a = \infty$, $K_r = 0$)

**Figure 17.** Mean surfaces of maximum stress of the HF nonlinear model with fixed and rotation-free BCs

As an example of the proposed LGMF method, a prediction model of the maximum stress of the curved strip model under extreme thermal load is built within the design domain of interest. The design parameters are $t/L$ and $\delta/L$ with ranges of $[0.005\sim0.050]$ and $[0.005\sim0.150]$ respectively. Assuming that geometric nonlinear analysis is capable of capturing the actual physics and nonlinear responses, HF samples are obtained by performing a nonlinear analysis of the curved strip model. Considering that in practical large-scale problems linear analysis is generally much cheaper than nonlinear analysis, a linear model is used as the LF model in this demonstration problem. It was observed [46] that the linear model makes well-correlated predictions to the nonlinear one within some ranges of the design domain, but becomes inaccurate and fails to capture even the trend of the nonlinear behavior in other ranges. Therefore, two different LF linear models are selected with non-zero finite stiffness ratios $k_a$ and $k_r$ (defined in [46]), as shown in Fig. 18. Unlike the HF models, the LF models are assumed to be deterministic.




a) LF01 ($k_a = 1.0, k_r = 500$) b) LF02 ($k_a = 0.5, k_r = 400$)

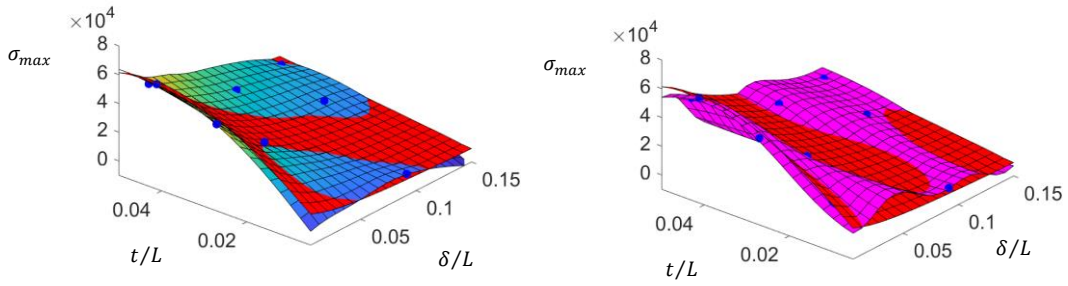
**Figure 18.** Maximum stress of the two selected linear LF models with finite stiffness ratios

In the following demonstrations of the proposed LGMF method, two different Boundary Conditions (BCs) are considered for the HF model: Fixed BCs with $K_a = \infty$ and $K_r = \infty$

and Rotation-free BCs with $K_a = \infty$ and $K_r = 0$. The maximum stress surface of the nonlinear model for each boundary condition is shown in Fig. 17. In either case, to simulate a practical situation it is assumed that HF samples are non-deterministic and randomly perturbed due to aleatory uncertainty in the BCs and modeling uncertainty of the geometric nonlinear model. The HF samples could also be obtained from actual physical tests which are subject to random variations in test conditions and operations. In this example, random perturbations to the HF stress response are accounted for by adding normal distributions to the rotational stiffness ratios $k_r$ and maximum stress $\sigma_{max}$ with standard deviations of 5 and 3000 psi respectively. The random variation of the HF stress will increase as $\delta/L$ reduces and $t/L$ increases due to the randomness in the rotational stiffness ratio. The goal is to use a small number of non-deterministic HF samples and the deterministic LF models to build a useful and physically meaningful prediction model that can be used for a design exploration study.

*Case 1: Fixed BCs with 12 HF non-deterministic samples*

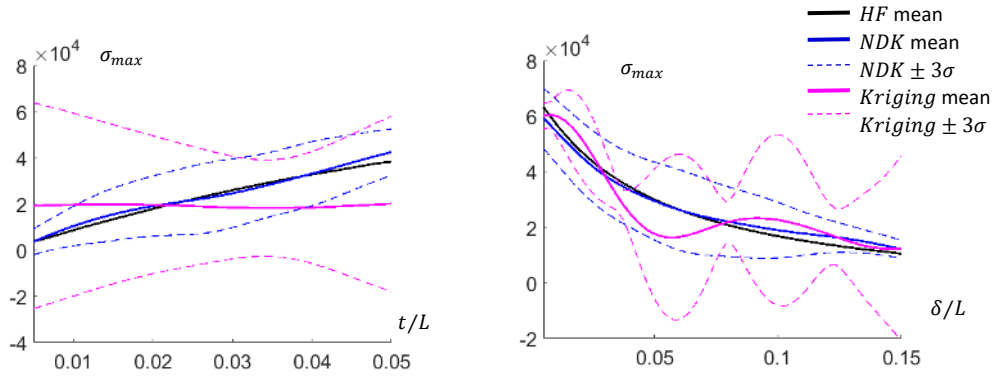
As a first case, the fixed BCs are set up using elastic elements with $K_a = \infty$, $K_r = \infty$. The geometric nonlinear response is plotted with the 18 HF samples in the design domain of $t/L \in [0.005{\sim}0.050]$ and $\delta/L \in [0.005{\sim}0.150]$ as shown in Fig. 19. Since they are non-deterministic, the HF samples are not exactly on the HF mean surface. The resulting LGMF and conventional kriging models are shown in Figs 19a and 19b, respectively. Here, kriging is built using the HF samples only while LGMF leverages the LF models along with the HF samples to build a prediction model that is insensitive to the randomness in HF samples.




a) LGMF (colormap surface) b) Kriging (magenta surface)

**Figure 19.** Maximum stress from HF (red surface), LGMF and kriging with 12 HF samples (blue circles)

Fig. 20 compares the maximum stress predictions along two cross-sections of the design domain, where $\delta/L = 0.048$ is the 30% point within its range, and $t/L = 0.036$ is the 80% point in its range. It is obvious based on Figs. 19 and 20 that LGMF is more accurate (closer to HF) than kriging. It is often more important in design exploration to make physically meaningful estimations of the response gradients than to precisely estimate the response itself. As shown in Fig. 20, kriging can mislead the design exploration by giving an incorrect gradient at a design point, while LGMF can more reliably estimate physically meaningful design gradients as well as the prediction uncertainty bounds (i.e., $\pm 3\sigma$).

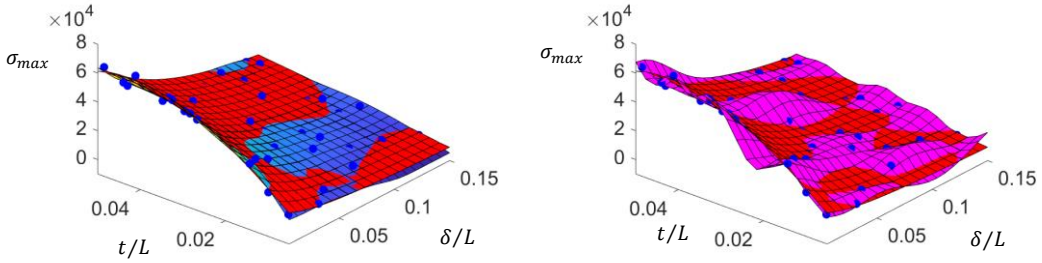



a) Stress along $t/L$ at $\delta/L = 0.048$ b) Stress along $\delta/L$ at $t/L = 0.036$

**Figure 20.** Case 1: Comparisons of the maximum stress responses from LGMF and kriging against HF

*Case 2: Fixed BCs: $K_a = \infty$, $K_r = \infty$ with 46 HF samples*

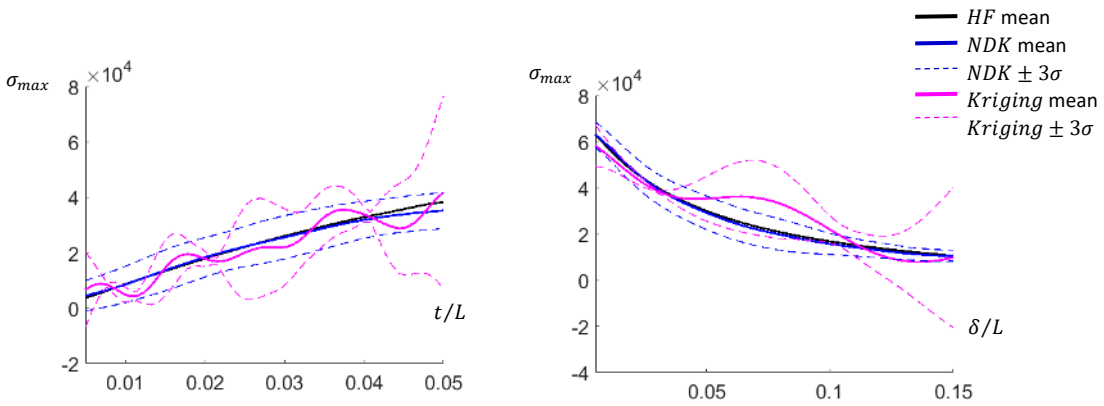
By increasing the number of samples from 12 to 46, the prediction model of LGMF is improved as shown in Fig. 21a. compared to Fig. 19a. However, it is observed that kriging shows more local bumps despite the increased number of samples. The local bumps of kriging depend on the random fluctuations of the samples and the layout of the data points. Unlike kriging, LGMF is immune to the local random fluctuations, and as more samples are added, the prediction model of LGMF becomes more accurate as expected.




a) LGMF (colormap surface) b) Kriging (magenta surface)

**Figure 21.** Maximum stress from HF (red surface), LGMF and kriging with 46 HF samples (blue circles)

It can be seen clearly in the comparison of the 1D predictions in Fig. 22 that the kriging model is more accurate than the previous case, although the frequency of local bumps is more severe. In this case, one should be cautious when interpreting the kriging prediction since the local bumps are due to numerical instability with no physical meaning. On the other hand, the mean and uncertainty bounds of the LGMF predictions shown in Fig. 22 accurately capture the true global trend.

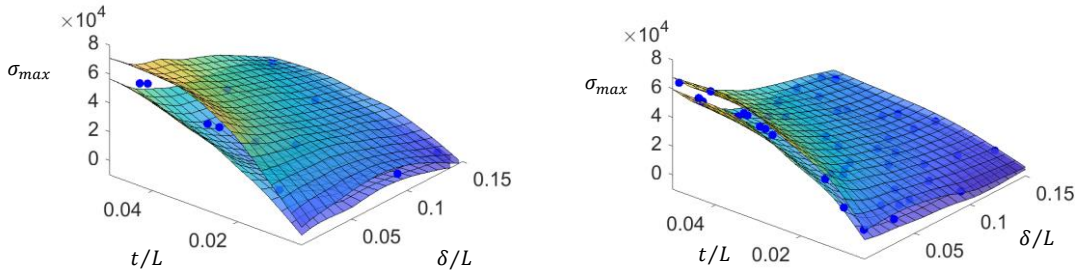



a) Stress along $t/L$ at $\delta/L = 0.048$

b) Stress along $\delta/L$ at $t/L = 0.036$

**Figure 22.** Case 2: Comparisons of the maximum stress responses from LGMF and kriging against HF

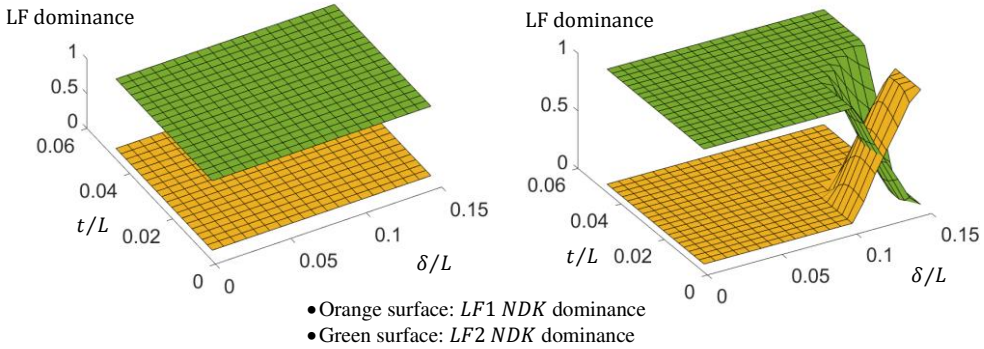
With non-deterministic HF samples, LGMF can provide the ±3σ uncertainty bounds of the prediction as shown in Fig. 23 along with the mean predictions shown in Figs. 19a and 21a. The assessed uncertainty bounds include both aleatory uncertainty due to the random perturbations of HF samples and modeling uncertainty caused by insufficient samples. The LGMF prediction uncertainty bounds shown in Fig. 23a are conservatively wide enough to include all the given HF random samples. Fig. 23 shows how the conservative bounds of Case 1 improve to the ones of Case 2 as the number of samples increases. Since the modeling uncertainty reduces when more samples are included, the prediction uncertainty bounds get closer to the true uncertainty bounds of the random perturbations of the HF samples.




a) Case 1 with 12 samples b) Case 2 with 46 samples

**Figure 23.** Uncertainty bounds (±3σ) of LGMF predictions for Case 1 and Case 2

As a byproduct of LGMF, the model dominance information is obtained in a quantitative manner as shown in Fig. 24. With a small number of samples as in Case 1, the model dominance information is only valid in an approximated global sense. However, with a large number of samples as in Case 2, the model dominance information becomes higher-resolution and more accurate. According to the dominance information, LF2 is better correlated to HF than LF1 overall, but for a high $\delta/L$ value and a low $t/L$ value LF1 better captures HF than LF2.

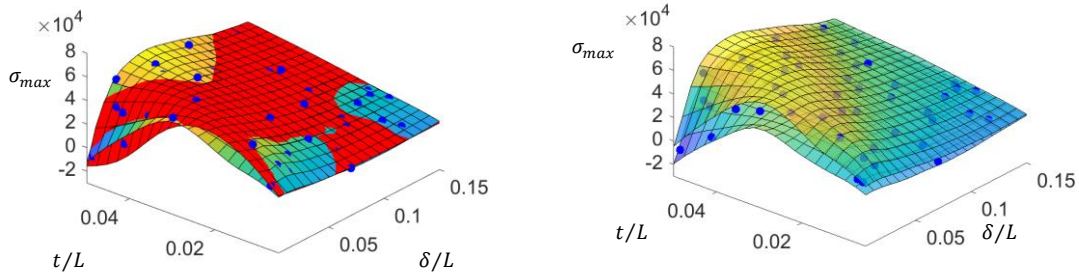



a) Case 1 with 12 samples b) Case 2 with 46 samples

**Figure 24.** Model dominance information from LGMF for Case 1 and Case 2

*Case 3: Rotation-free BCs: $K_a = \infty$, $K_r = 0$ with 46 HF samples*

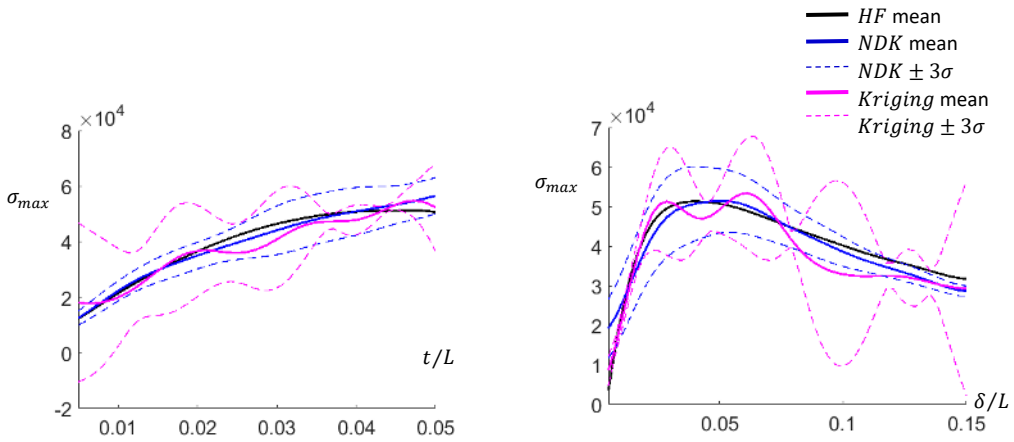
It is shown in Fig. 17 that geometric nonlinearity with rotation-free BCs is far different from the other cases with the fixed BCs, especially for a strip with low curvature and high thickness. Like the previous cases, LGMF provides a reliable prediction model with uncertainty bounds of prediction with 46 random HF samples as shown in Figs. 25 and 26.




a) LGMF (colormap surface)

b) Uncertainty bounds (±3σ)

**Figure 25.** Case 3: LGMF prediction mean and uncertainty bounds for the rotation-free BCs case with 46 HF samples (blue circles)




a) Stress along $t/L$ at $\delta/L = 0.048$

b) Stress along $\delta/L$ at $t/L = 0.036$

**Figure 26.** Case 3: Comparisons of the maximum stress responses from LGMF and kriging against HF

Comparing the HF responses with the LF1 and LF2 models, it is observed that in the design domain LF1 is better correlated to HF within the ranges of small curvature and high thickness, but in ranges of small curvature and small thickness, LF2 captures the HF behavior more accurately. This observation is clearly reflected in the model dominance information shown in Fig. 27. One can see how the HF model behavior makes transitions from one fundamental behavior explained by a LF model to the other behavior captured better by another LF model within the design domain in a quantitative way.

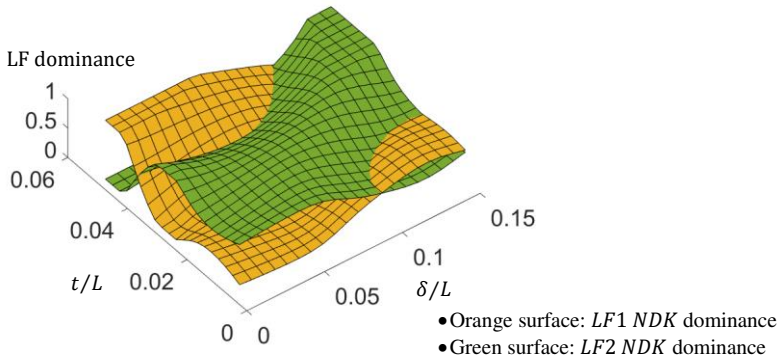



**Figure 27.** Model dominance information from LGMF fit of Case 3, for LF1 (orange surface) and LF2 (green surface)

*4.1.3 Summary of the Proposed LGMF Modeling Method*

This section introduced multi-fidelity modeling using the non-deterministic localized-Galerkin approach to address potential practical challenges such as how to combine models with various fidelities, how to deal with Low Fidelity (LF) models that show localized correlations to the high-fidelity (HF) model, and how to consider samples from simulations and physical data under uncertainty. The Localized-Galerkin Multi-Fidelity (LGMF) model is expressed as an expansion function with an arbitrary number of selected basis functions. The basis functions can be derived as multiplicative, additive, or hybrid/comprehensive corrections, which are consolidated using participation functions in

the proposed LGMF framework. First, the localized-Galerkin approach is applied to determine the participation functions of the multiple LF models. Once the participations are determined, all LF models are consolidated into a single globally correlated model, which will be refined in the second stage of the proposed LGMF framework, resulting in the final MF model. LGMF does not require user-defined ranks of fidelity or accuracy to combine multiple models with differing fidelities. The degree of local correlation or dominance is estimated based on available sample data from the HF model.

The proposed LGMF modeling method has been successfully demonstrated using fundamental mathematical one- and two-dimensional problems with two LF models, which show localized correlations within different local domains. The dominance information can be useful in understanding the characteristics of the HF model behavior in terms of fundamental LF models. If the dominance boundary is available *a priori*, the LF consolidation stage can be skipped since this is essentially the same case as having a globally well-correlated LF model. Potential applications and technical contributions include: 1) along with other existing data classification or clustering algorithms, the model dominance information can enable a physics-based data clustering, 2) the non-deterministic LGMF model can provide aleatory and epistemic uncertainty bounds along with the MF mean prediction, which can be used to enable an adaptive sampling scheme to optimize the LF and HF sample evaluations, and 3) based on the goal of the MF model application, the MF prediction model can be designed to provide either optimistic (risk-taking) or conservative (risk-averse) predictions.

## 4.2 Expected Effectiveness (Adaptive Sampling for Global Optimization)

In the previous section introducing LGMF, HF training samples were added using the LHS method while LF training samples were usually considered computationally trivial and directly sampled at each prediction point. Adaptive sampling yields a more efficient way of collecting HF and LF training samples. Which adaptive sampling method is appropriate depends on the desired information, which in this section is determination of the global optimum. The Expected Improvement (EI) metric works well for adaptively adding HF samples to an LGMF fit. However, adaptively adding LF samples requires a new methodology. In this section the Expected Effectiveness (EE) acquisition function for adaptive sampling of LF models is introduced. Surrogates are created from the limited number of LF samples and used to construct the LGMF fit. To accommodate aleatory uncertainty in the LF samples, the surrogates used are NDK models. If the data are known to be deterministic, NDK will behave like conventional kriging.

### *4.2.1 Changes to LGMF Implementation for adaptive sampling*

The LGMF methodology used here follows the procedures described in the previous section with minor changes. First, lack of data from sparse initialization at the start of adaptive sampling can result in poor selection of the $h$ parameter during LOO optimization. In these examples the $h$ parameter is set to a user-defined constant. Second, a final stage of filtering using NDK helps improve the accuracy of the final LGMF model.

***Filtering of final LGMF model using NDK***: Because of lack of data, the LGMF response in Eq. 43 can lead to some discontinuities even with an optimal $h$ parameter, especially with a small number of data samples in the beginning of the adaptive sampling process. This problem is addressed by using a low-pass filtering process, in which the LGMF Stage

2 responses are resampled and applied to build an NDK model. The filtering parameter for the low-pass frequency can be determined based on the minimum distance of an expected stationary response. This NDK model is used as the final LGMF fit.

*4.2.2 Proposed EE Adaptive Sampling for LGMF*

EGO [9] was introduced to use the estimation and uncertainty bounds of a kriging fit to balance exploration and exploitation efficiently for global optimization. It works by sampling at the point with maximum EI, where EI is the value by which a point taken at a given sampling location can be expected to improve over the current best sample, where a worse or equal value yields an improvement of 0. The formulation of expected improvement was given in Eq. 13 in Chapter III. Based on expected improvement, the new Expected Effectiveness (EE) method for adaptive sampling of multiple fidelities is introduced. The EE method makes use of the EI of the LGMF surrogate, given by

$$EI_{LGMF}(x) = \left(\hat{y}_{min} - \hat{y}_{\mathrm{LGMF}}(\mathrm{x})\right) * \Phi\left(\frac{\hat{y}_{min} - \hat{y}_{LGMF}(x)}{\sigma_{LGMF}(x)}\right) + \sigma_{LGMF}(x) * \phi\left(\frac{\hat{y}_{min} - \hat{y}_{LGMF}(x)}{\sigma_{LGMF}(x)}\right) \quad (56)$$

where $\hat{y}_{min}$ is the current LGMF predicted optimum, $\hat{y}_{LGMF}(x)$ is the LGMF prediction at x, and $\sigma_{LGMF}(x)$ is the standard deviation of the LGMF prediction at $x$. Also, $\phi(\cdot)$ and $\Phi(\cdot)$ are the standard normal density and cumulative distribution functions, respectively. An example using EI is included in the "Existing Surrogate Modeling Methods" section of this thesis.

The EE method combines the EI of the LGMF surrogate with the Modeling Uncertainty (MU), Dominance under Uncertainty (DU), and evaluation cost of the model being evaluated. That is, the EE of the $m^{th}$ LF model is given by

$$EE(x,m) = EI_{LGMF}(x, LGMF) \times DU(x,m) \times MU(x,m)/Cost(m) \quad (57)$$

where MU is the ratio of Epistemic to Aleatory uncertainty in the NDK model of the LF function:

$$MU(x,m) = \sigma_{LFm\ NDK\ Epistemic}(x)/\sigma_{LFm\ NDK\ Aleatory}(x) \quad (58)$$

As more data points are added, the model will become saturated, the epistemic uncertainty will trend toward 0 and sampling of the LF function will cease. DU is defined as the dominance of the LF model plus the change in dominance that resulted from the last adaptive sample, that is

$$DU(x,m) = Dominance\_LGMF\ (x,m) + \Delta Dominance\_LGMF(x,m) \quad (59)$$

where the change in dominance for the $k^{th}$ iteration is calculated as

$$\Delta Dominance\ LGMF(x,m) = Dominance\_LGMF^{k}(x,m) - Dominance\_LGMF^{k-1}(x,m) \quad (60)$$

Each iteration of the adaptive sampling, the LF model $m$ and location $x$ with the highest EE value are sampled, and the LGMF fit is updated. If all LF models are converged below tolerance, the HF model is evaluated at the location of maximum EI. The flowchart for the algorithm's behavior is shown in Fig. 28.

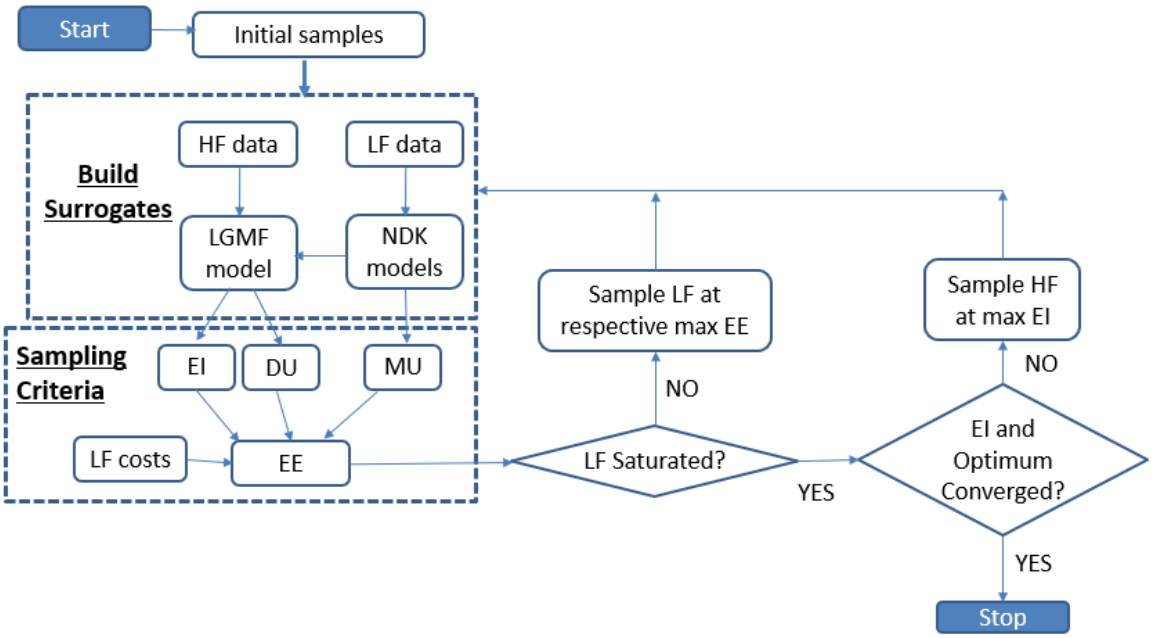



**Figure 28.** Flowchart for behavior of EE adaptive sampling method for LGMF models

*4.2.3 Numerical Examples*

This section presents numerical examples and discusses the behaviors of the proposed EE adaptive sampling method. The EE-based approach is demonstrated with multiple fundamental equations, as well as a fundamental cantilever beam example representing a long-span aircraft wing with store masses under dynamic loads. Since all examples start with small numbers of initial samples, the LGMF kernel shape parameters are fixed to cover sufficiently large distances of 0.8, 1.1, and 1.3 within normalized design spaces for 1D, 2D and 3D examples respectively.

Example 1:

Non-Deterministic One-Dimensional Optimization Problem leveraging two LF models

In this design optimization example, it is desired to find a design $x$ that minimizes the cost function $f(x)$ within a given design domain D=$\{x|\ 0<x<1.1\}$ as in Eq. 61

$$x^* = argmin_{x\in D} f(x) \tag{61}$$

where $f(x) = (6x-2)^2 sin(12x-4)$. Here, the true function $f(x)$ is unknown, and only its measurement can be sampled. The measurement is modeled as the HF model with a stochastic variation of 10% Coefficient of Variance ($CoV$) from the true mean $f(x)$ as

$$f_H(x) \sim Normal\ (f(x), 10\%\ CoV) \tag{62}$$

Instead of using only the HF model, an adaptive MF model is built iteratively within the iterative optimum design search by leveraging two LF models. The LF models are derived as abstracted models of the HF model. It is assumed that they are inexpensive but inaccurate having estimation uncertainties of $\sigma_{LF1} = 0.2$ and $\sigma_{LF2} = 1.0$ as well as nonlinear deviations from the mean of the HF model

$$f_{LF1}(x) = 1.5sin(8x-4) + 5(x-0.5) - 5 \tag{63}$$

$$f_{LF2}(x) = -6\sin(8x-4) - 7 \tag{64}$$

The HF and LF functions and HF uncertainty are shown along with the true optimum solution marked with the red star symbol in Fig. 29.

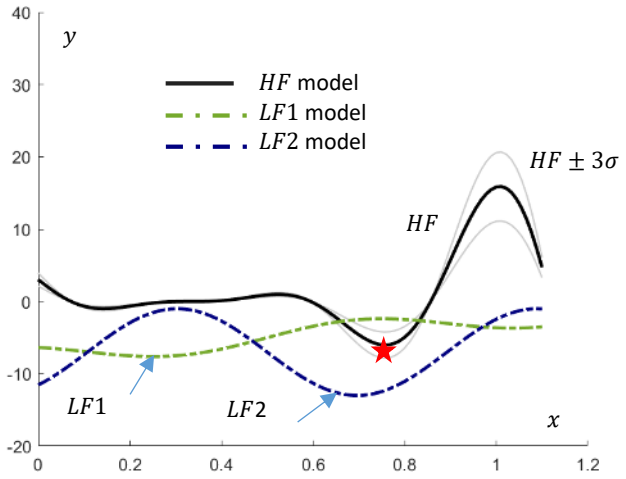
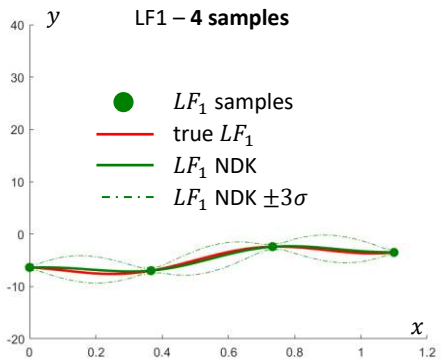


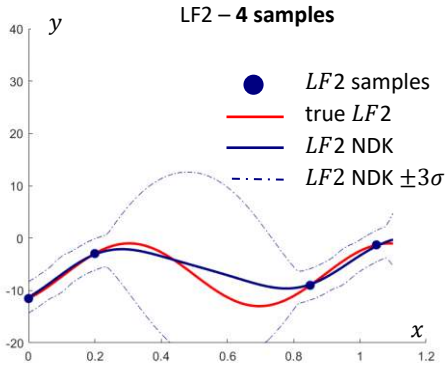

**Figure 29.** Surrogate models used in Example 1. Optimum denoted by star.

As shown in Fig. 30, five HF samples are generated while both LF NDK models are prepared with four samples at the beginning. Here, the ±3σ bounds of NDK models include

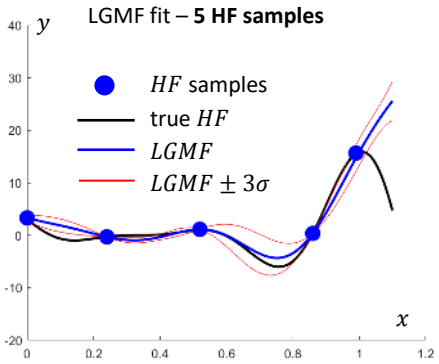
both the aleatoric random variations of $\sigma_{LF1} = 0.2$ and $\sigma_{LF2} = 1.0$ and modeling uncertainty due to lack of samples.




a) LF function 1



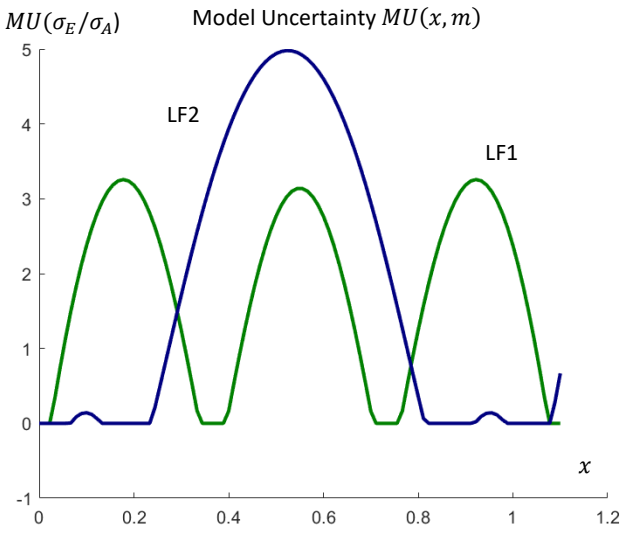

b) LF function 2



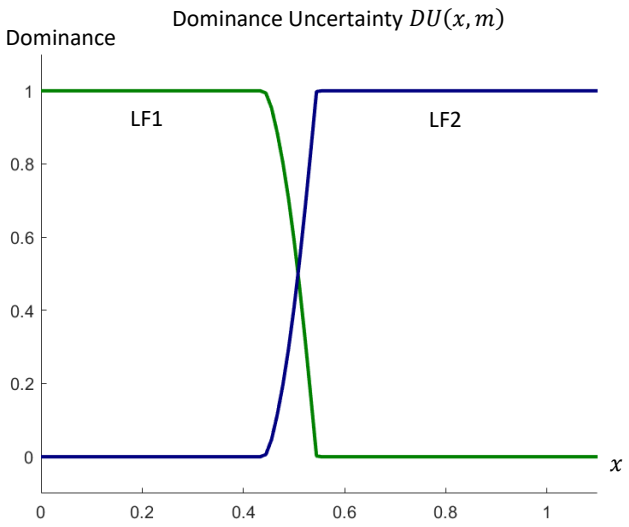

c) HF samples

**Figure 30**. Initial LF samples and surrogates.

The MU and DU terms during this first iteration are shown in Fig. 31. Notice how the lack of LF2 data in the range of between x=2.5 and x=0.8 translates to a high Model Uncertainty shown in Fig. 31a. Note also from Figs. 29 and 30 how NDK $LF_1$ better follows the trend of the HF function in the first half of the design domain, while NDK $LF_2$ better follows the trend in the second half, even though they are still premature. This is reflected in the Dominance under Uncertainty plot (Fig. 31b).



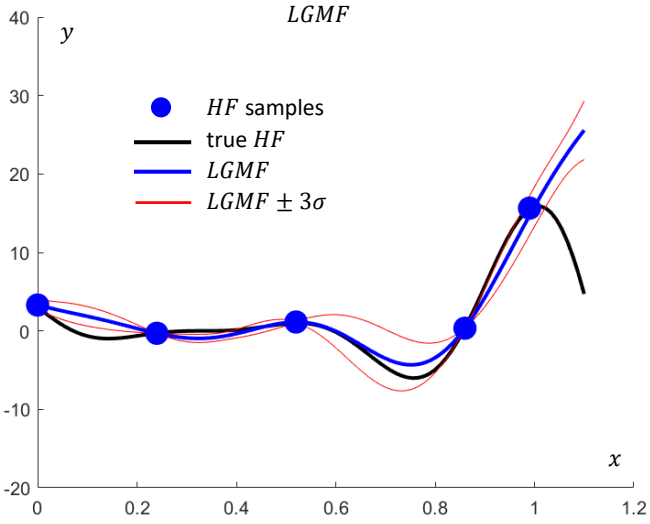

a) Modeling Uncertainty for each LF function.



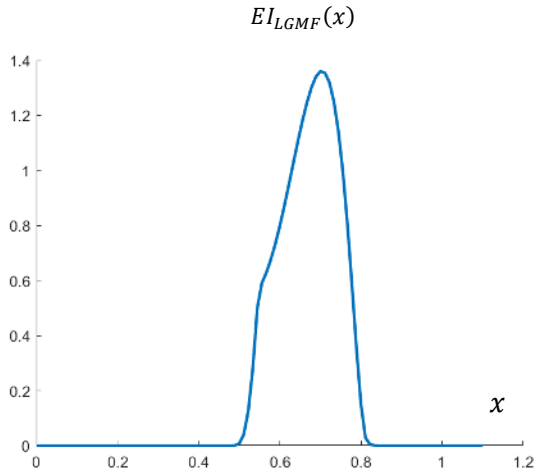

b) Dominance under Uncertainty for each LF function.

**Figure 31.** Information used for adaptive sampling during first iteration.

The resulting LGMF fit, as well as the EI of the LGMF fit, are shown in Fig. 32. Notice how the high uncertainty and low prediction between x=0.6 and x=0.8 results in a high EI value.




a) HF samples and LGMF fit using LF information.




b) Expected Improvement of LGMF fit.

**Figure 32.** LGMF fit and corresponding EI values.

Finally, the EE values for the identical-cost LF functions are shown in Fig. 33. The locations of maximum EE value for each LF function are marked with stars. Unsurprisingly, LF2 has a significantly higher Expected Effectiveness considering both model and dominance uncertainties. In a case where only one adaptive sample is allowed, the next sample will be added for LF2 NDK modeling at the maximum EE location. New samples are added for both LF NDK models every iteration as long as the EE values are above the tolerance value of $10^{-6}$.

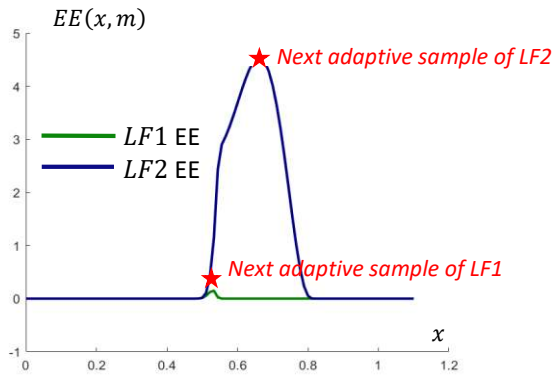



**Figure 33.** EE value for each LF function across the domain.

After sampling both LF functions, the NDK of the LF models and the LGMF surrogate are updated and the new locations with maximum EE are calculated. Both LF functions are sampled each step until their EE values become insignificant or less than the EE tolerance. In this example $LF_1$ converges quickly in one step. Once both LF NDK exhibit insignificant EE, it means that there is no more information to be gained from the LF models with the current set of HF samples. When this is observed, a new non-deterministic HF sample is added at the location of maximum EI. When a new HF sample is added, not only LGMF, but also the dominance uncertainties of the LF models are updated, which can

make the LF EEs significant again. Figs. 34-35 show the updated LF NDKs, MU, DU, EI, EE, and LGMF with the two adaptive samples of both LF models.

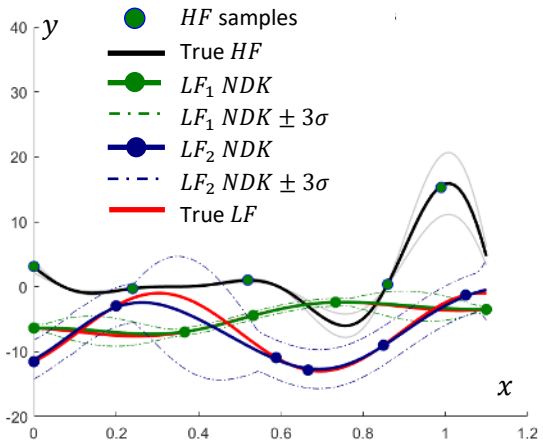
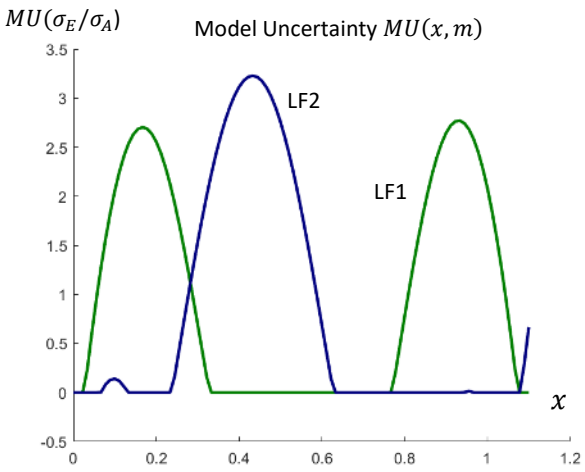
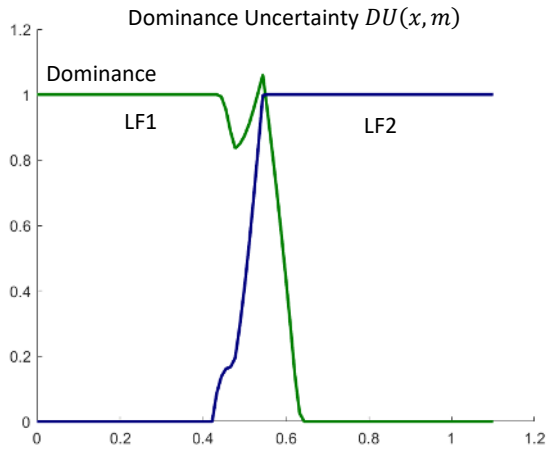


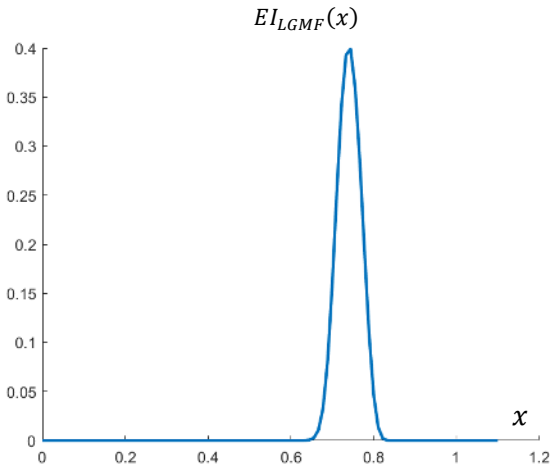

**Figure 34.** Data samples and LF NDK surrogates after models are updated.



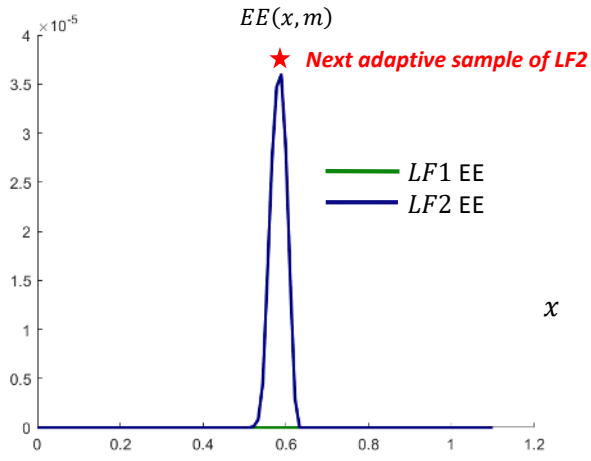

a) Modeling Uncertainty for each LF function.




b) Dominance under Uncertainty for each LF function.




c) Expected Improvement of LGMF fit.

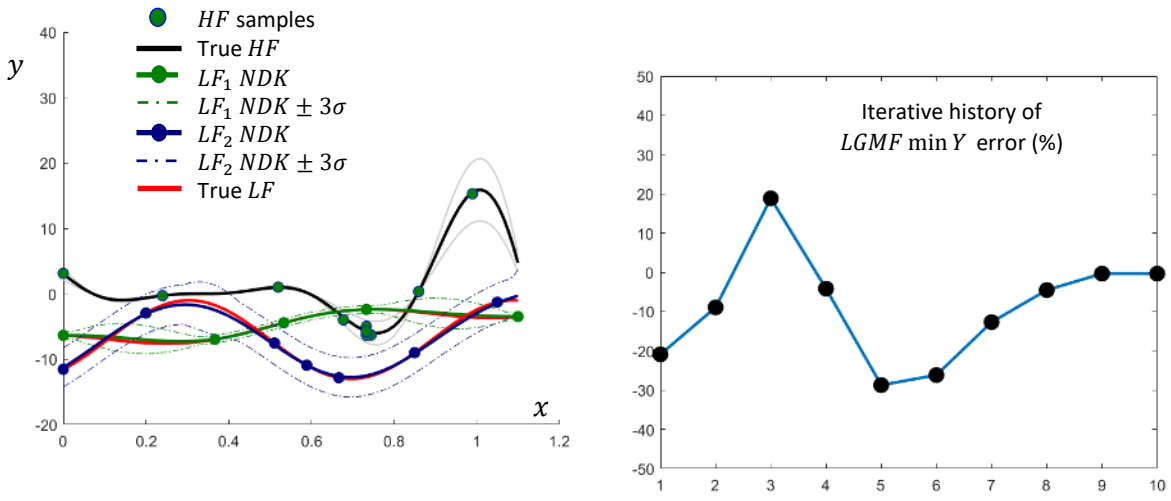



d) EE value for each LF function across the domain.

**Figure 35.** Information used for adaptive sampling during second iteration.

It is observed that DU is changed only slightly while MU, EI, and EE are changed significantly especially in magnitudes and their maximum locations. Over the iterations, it is desirable to see decreasing magnitudes of EE and EI. At the end of adaptive sampling, the maximum EEs of the LF models should be all zero, which means that the LF models are fully exploited around the expected optimum design location. Unlike the conventional EI from kriging, LGMF EI does not become zero, but converges to a finite value at the end

of adaptive sampling due to the non-deterministic nature of the cost function. Therefore, through the iterations, either LF or HF samples are adaptively added until EE becomes insignificant, and both the minimum of HF and LGMF EI are converged within 0.1%. For this example, the samples and NDK surrogates once the adaptive sampling is complete are shown in Fig. 36. An iterative history of the percent error in the optimum response is included. In the end, the algorithm found the converged optimum design $x$=0.75 after adding one $LF_1$ sample, three $LF_2$ samples, and six $HF$ samples. Because it started with four $LF_1$ samples, four $LF_2$ samples, and five $HF$ samples, the final result is obtained with a total of five $LF_1$ samples, seven $LF_2$ samples, and eleven $HF$ samples. Because of the stochastic randomness of HF, it is shown that multiple samples are drawn around the optimum design; five of the HF samples are clustered together.



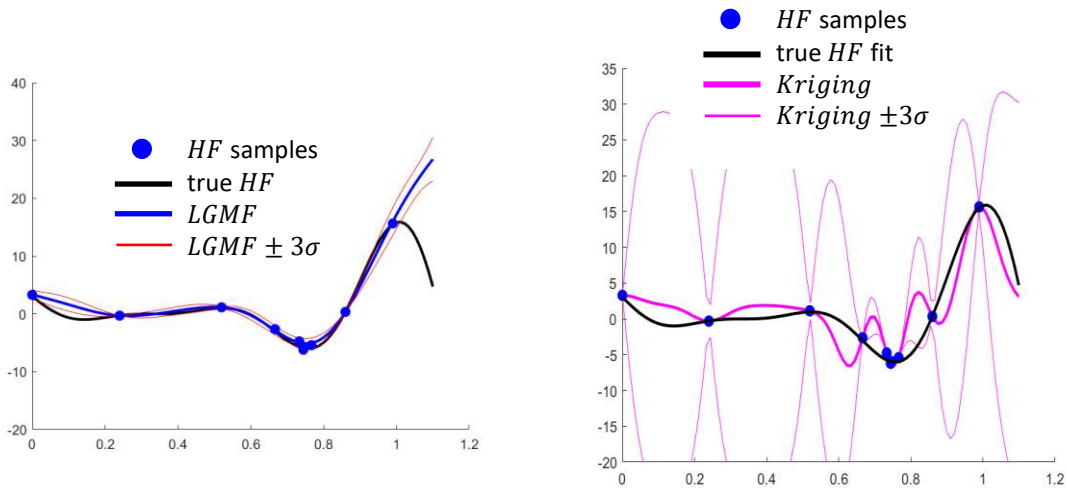

a) Final surrogates and samples.

b) Percent error in the optimum response over the iterations.

**Figure 36.** The completed adaptive sampling process.

A comparison between the final LGMF fit and a kriging fit built by using only the HF samples is shown in Fig. 37. Kriging's interpolation requirement results in significant overfitting and unreasonable kriging uncertainty bounds (±3σ) which will lead to wrong

EI estimations. The overfitting issue can be suppressed by using regression kriging or general Gaussian process with the nugget [10] which basically alleviates the interpolation requirement. However, since those methods capture the stochastic randomness as a random white noise, the measures of EI and EE can mislead adaptive sampling.




a) HF samples and final LGMF surrogate. b) Kriging fit constructed using HF samples.

**Figure 37.** Comparison between LGMF surrogate and Kriging

Example 2: Hartman 3D problem

In this mathematical example, minimization of the 3D Hartman function is considered,

$$f_H(x) = -\sum_{i=1}^{4} c_i \, exp\left[-\sum_{j=1}^{d} \alpha_{ij}\left(x_j - p_{ij}\right)^2\right] \tag{65}$$

where $\alpha$, $c$, and $p$ are matrices defined as

$$\alpha_{ij} = \begin{bmatrix} 3 & 10 & 30 \\ 0.1 & 10 & 35 \\ 3 & 10 & 30 \\ 0.1 & 10 & 35 \end{bmatrix}, c_i = \begin{bmatrix} 1 \\ 1.2 \\ 3 \\ 3.2 \end{bmatrix}, p_{ij} = \begin{bmatrix} 0.37 & 0.11 & 0.27 \\ 0.47 & 0.44 & 0.75 \\ 0.11 & 0.87 & 0.55 \\ 0.04 & 0.57 & 0.88 \end{bmatrix}$$

Having the Hartman equation as the HF model, the LF model is defined in Eq. 66 with a systematic deviation from HF. The deviation function is a second order polynomial function (MA3), introduced in [31] and shown in Eq. 67. The scale factor 7.6 applied to

the MA3 deviation function in Eq. 66 is to make the deviation as large as the full range of the HF response changes.

$$f_{LF}(x) = f_H(x) + 7.6 \times MA3(x) \tag{66}$$

where

$$\begin{aligned} MA3(x) = {} & 0.585 - 0.324x_1 - 0.379x_2 - 0.431x_3 \\ & -0.208x_1x_2 + 0.326x_1x_3 + 0.193x_2x_3 \\ & +0.225x_1^2 + 0.263x_2^2 + 0.274x_3^2 \end{aligned} \tag{67}$$

Additionally, evaluations of both HF and LF models are assumed to have random noise that is normally distributed with standard deviations of 0.01 and 0.02 for HF and LF models, respectively. Contours of the HF model in 3D are shown in Fig. 38. The optimum location is marked by a red star $x_{opt} = (0.1146 \quad 0.5556 \quad 0.8525)$ and the minimum function values $y_{opt}$ is $= -3.8628$.

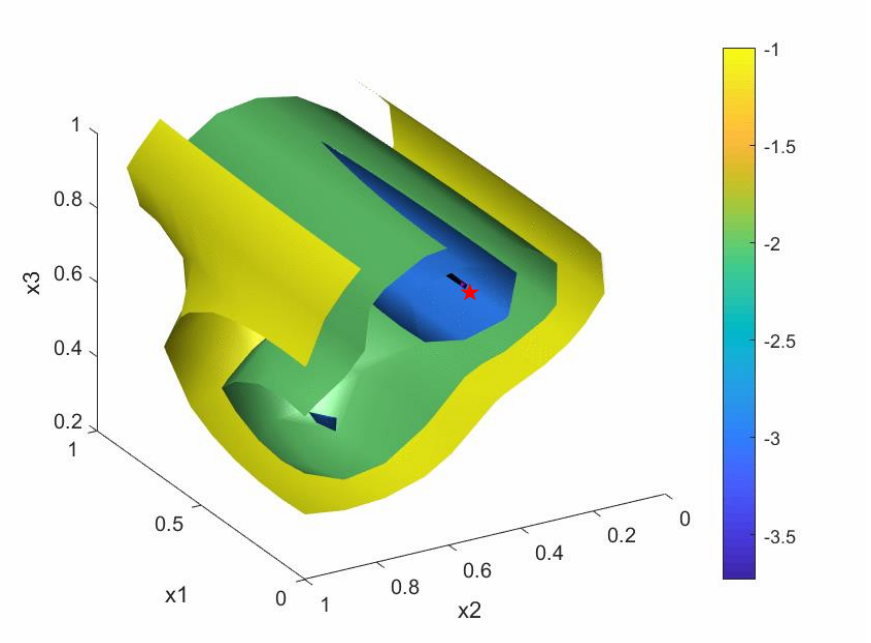



**Figure 38.** Contours of the Hartman 3D function. Optimum denoted by star.

The proposed adaptive sampling iteration starts with the initial 10 HF and 30 LF samples. The initial samples are generated by using the Latin Hypercube Sampling (LHS)

method to minimize the sample clustering. The initial contours are shown in Fig. 39. The true contours are shown as the red surfaces and the initial LGMF contours are shown as the colormap surfaces, which show significant errors.

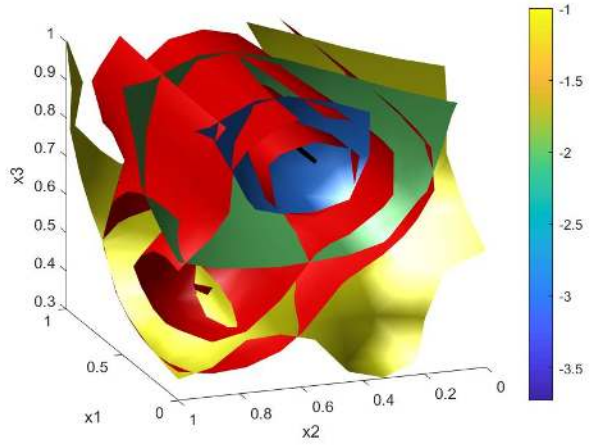



**Figure 39.** Contours of Hartman 3D function (red) and LGMF surrogate (colormap) at the beginning of optimization. (10 total HF evaluations and 30 total LF evaluations)

The adaptive sampling converges after adding an additional 8 HF samples and 23 LF samples. The converged LGMF surrogate (colormap) matches the HF function (red) much more closely around the optimum location as shown in Fig. 40.

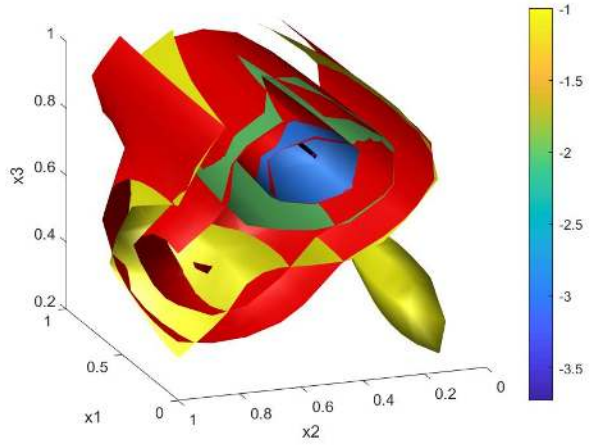



**Figure 40.** Contours of Hartman 3D function (red) and LGMF surrogate (colormap) at the end of optimization. (18 total HF evaluations and 53 total LF evaluations)

The final estimated optimum using the LGMF surrogate is

$$x_{LGMF\ opt} = (0.0000 \quad 0.5609 \quad 0.8511) \tag{68}$$

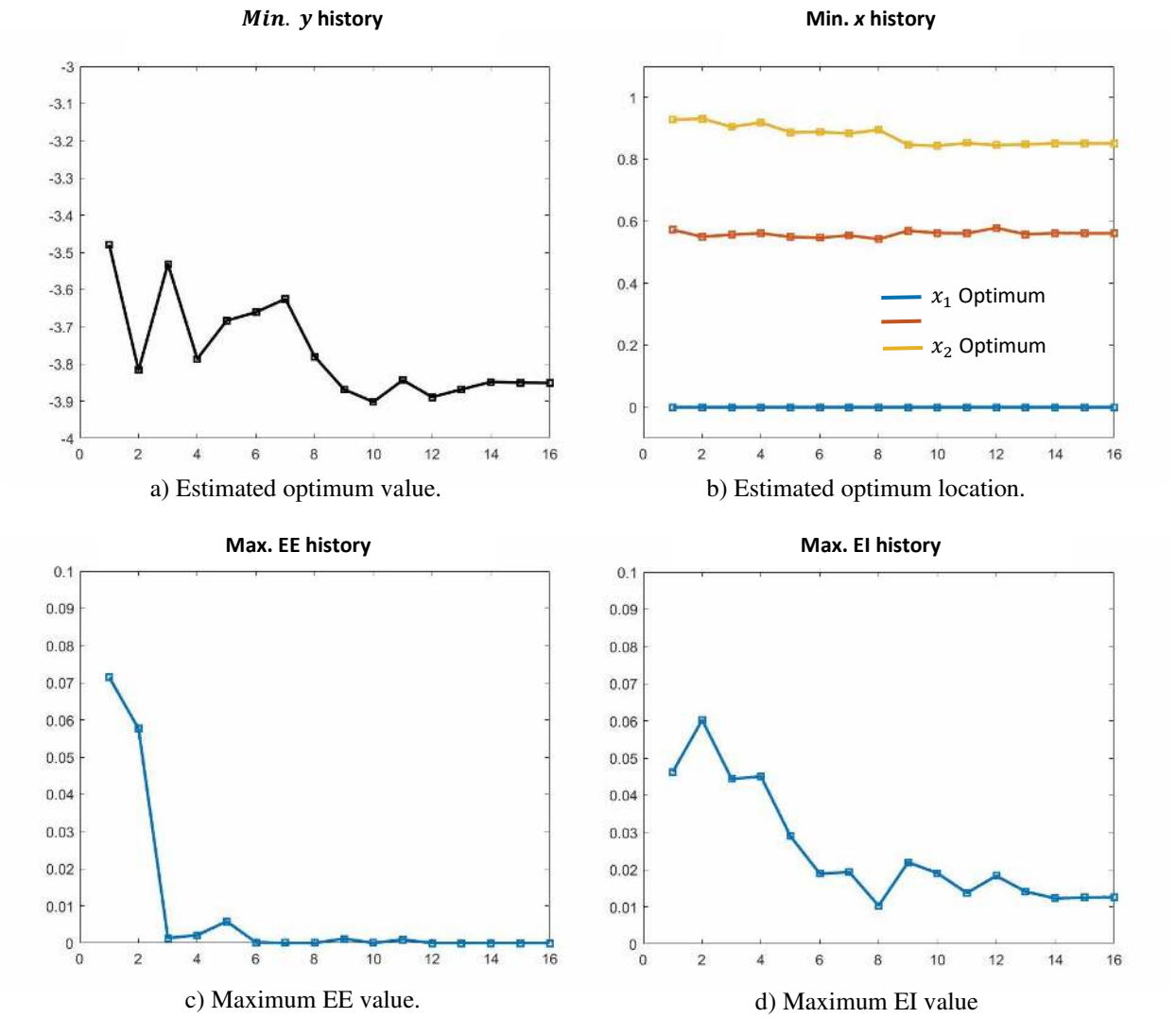
$$y_{LGMF\ opt} = -3.8534 \tag{69}$$

which is close to the true optimum values of

$$x_{opt} = (0.1146 \quad 0.5556 \quad 0.8525) \tag{70}$$

$$y_{opt} = -3.8628 \tag{71}$$

Therefore, in a relatively small number of function evaluations, EE converged to an accurate solution. The iteration history of the estimated optimum value, estimated optimum location, and EE and EI values is shown in Fig. 41.




a) Estimated optimum value.

b) Estimated optimum location.

c) Maximum EE value.

d) Maximum EI value

**Figure 41.** Iteration history of optimization

The iteration history of the expected optimum response shows gradual convergence to the nearly true value. From the iteration history of the estimated optimum location, it is observed that LGMF initially identified the general optimum, and then EE gradually refined its precise location.

As in the previous example, the EE value of the LF function eventually converged towards 0. When it converges, there is no more information to be gained from the LF model with the current set of HF samples. Therefore, a new HF sample is added at the location of

maximum EI. When a new HF sample is added, the LGMF fit and the LF EE values are re-calculated. This can make the LF EEs significant again, as shown as the fluctuations of the Maximum EE history plot in Fig. 41.

Also like the previous example, the EI values converge to a non-zero finite number. This is because of the way LGMF and NDK are designed to avoid interpolating and maintain non-zero uncertainty bounds at data samples in order to handle the white noises in HF and LF models.

Example 3: Fundamental Aircraft Wing Model with a Cantilever Beam under Dynamic Loads

A fundamental aircraft wing structure abstracted as a cantilever beam as shown in Fig. 42 is considered. The cantilever beam is modeled with 12 finite beam elements with two concentrated mass elements (300kg per each) that represent two stores attached to the wing. The length of the beam is 6m and the radius of the circular cross section is 7cm. The Young's modulus and mass density are $20\times10^6$ N/cm$^2$ and 0.02 kg/cm$^3$ respectively. The proportional damping model with the coefficients of $\alpha$=0.01 and $\beta$=0.01 is considered for the dynamic analysis. Sinusoidal excitations with uniformly distributed loads are applied within the frequency range shown in Fig. 43 with the magnitude of 1000kN to simulate aerodynamic loads at a certain flight speed.

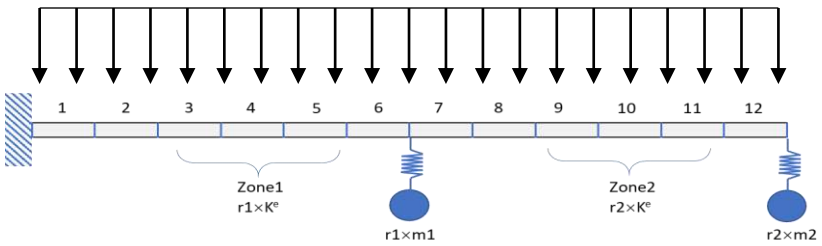



**Figure 42.** Cantilever beam with attached masses and applied forces, used to model an airplane wing with tip stores.

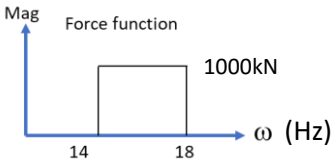



**Figure 43.** Excitation force applied to the cantilever beam.

In this example, it is assumed that the wing structural and store design changes are reflected by varying the $r_1$ and $r_2$ coefficients of the element stiffnesses and concentrated masses, which have the ranges of $r_1 \in [0.1,\ 1.5]$ and $r_2 \in [0.1,\ 1.5]$. The goal is to find the optimum design which will minimize the maximum stress at a predefined critical location i.e., Element 7, as in Eq. 72.

$$x^* = argmin_{x \in D}\ \sigma_{max@\ Element\_7} \tag{72}$$

Here, the design domain $D \in [0,\ 1]^2$ is the normalized space of the design change coefficients with the variable transformation, $r_{1,2} = 1.4x_{1,2} + 0.1$, respectively. It is noted that the dynamic excitation causes non-stationary nonlinear responses (i.e., maximum local stresses) of the cantilever beam due to potential mode switching within the design domain of interest as shown in Fig. 44. Without considering any randomness of the structural design and loading condition, the optimum solution can be located as denoted in the figure.

Instead of design exploration with the single-fidelity model alone, two low-cost LF functions that approximate the maximum stress are included in the proposed adaptive sampling process. These LF models are constructed by using the dynamic sub-structuring method, the Craig-Bampton (CB) Method [47]. This method works by breaking the cantilever beam structure into two components, which are then solved as separate but made to interact each other. The CB method allows us to reduce the computational costs of structural dynamic analysis by ignoring higher-order natural frequencies of components. Fig. 45 illustrates the division of the beam into two components and the interface node.

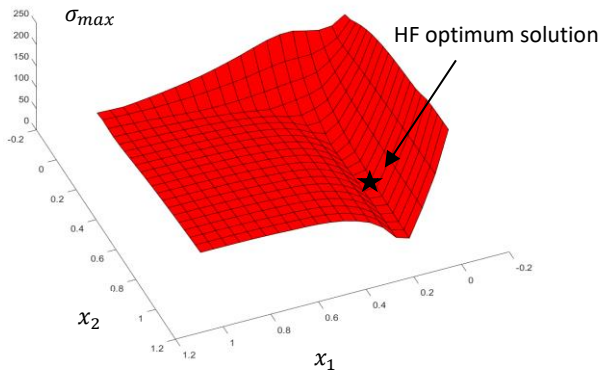



**Figure 44**. Maximum stress responses of Element 7 from the HF model and the optimum solution

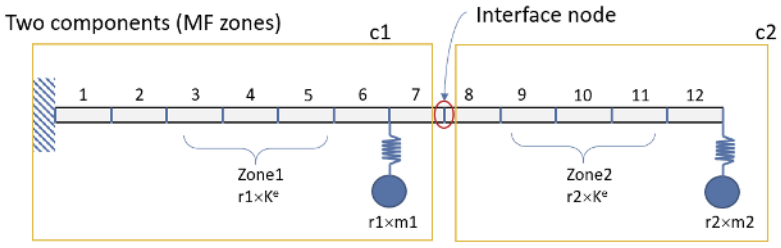



**Figure 45.** Craig-Bampton Method was used to generate LF stress responses at Element 7.

As additional higher order frequencies are dropped from the CB analysis, the cost saving increases with the sacrifice of analysis accuracy. In this problem two different LF models are used that enable different cost savings. For $LF_1$, the computational cost saving is about 25% from the HF model cost by using up to 9$^{th}$ modal information in the CB analysis, while the cost saving of $LF_2$ is about 80% by including only first two modal information. As expected in Fig. 46, the LF1 model is more accurate in predicting the maximum stress with higher costs (more than 3 times) than LF2.

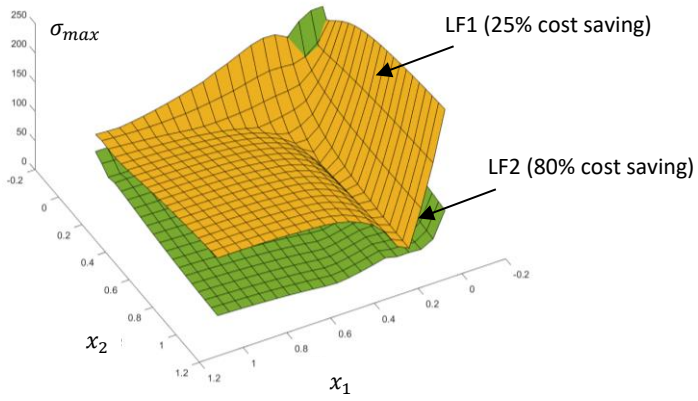



**Figure 46**. Maximum stress responses of Element 7 from the LF models.

The proposed adaptive sampling was initialized with 15 samples from each LF function and 3 samples from the HF function. The initial samples and surrogates are shown in Fig. 47.

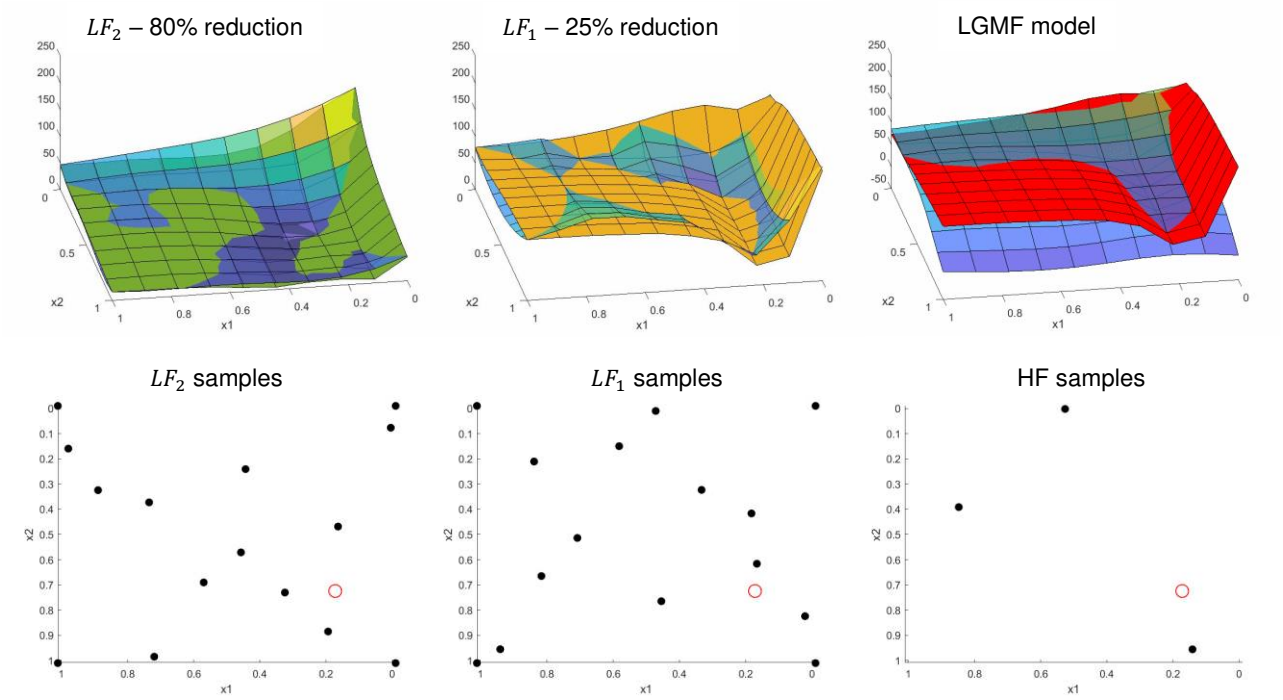



**Figure 47.** Initial stage: LF and HF samples and NDK models. Here, true responses are given as solid color surfaces, while NDK models are transparent color-map surfaces.

After the algorithm sampled an additional 6 HF samples, 11 $LF_1$ samples, and 12 $LF_2$ samples, the algorithm found the solution converged to the optimum design variables as shown in Fig. 48,

$$x_{LGMF\ opt} = (0.1672, \quad 0.7542) \tag{73}$$

$$\sigma_{LGMF\ max\ opt} = 13.002 \tag{74}$$

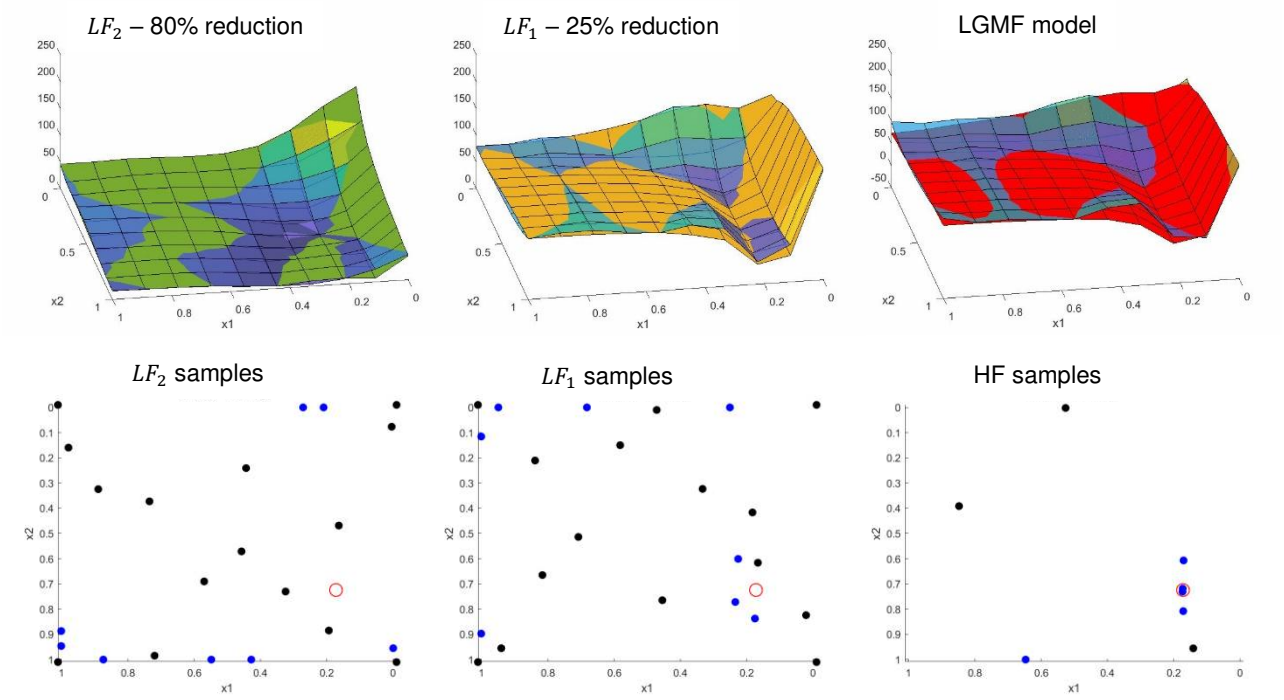



**Figure 48.** Final stage: LF and HF samples and NDK models. Here, true responses are given as solid color surfaces, while NDK models are transparent color-map surfaces. Also, initial and adaptive samples are black and blue respectively while the optimum is marked with red circles.

The discrepancy between the true and predicted response optimum is about 2% compared to the total range of the HF maximum stress within the design domain. As shown in Fig. 48, the EE adaptive sampling algorithm sampled very near the optimum twice. A comparison using EI and kriging is also shown, which took 5 times as many samples to sample the optimum so precisely. The optimum is marked by red circles.

It was observed during the sampling iterations that the proposed method explored and exploited the cheapest model LF2 at the initial stage, and started to have more samples from LF1 and HF gradually to increase the LGMF prediction accuracy even though they are more costly, which is the desirable and expected behavior of the proposed method.

Overall, the final LGMF surrogate had 9 HF samples, 26 $LF_1$ samples, and 27 $LF_2$ samples. Therefore, the optimization had a total cost of $26 * 0.75 + 27 * 0.2 + 9 = 33.9$ HF equivalent samples, as opposed to the 45 HF samples needed by EI based adaptive sampling with the single-fidelity kriging modeling. In general, the cost savings would depend on the selections of LF models and could increase significantly if there are big cost ratios between HF and LF models. The EE iteration history is shown in Fig. 49.

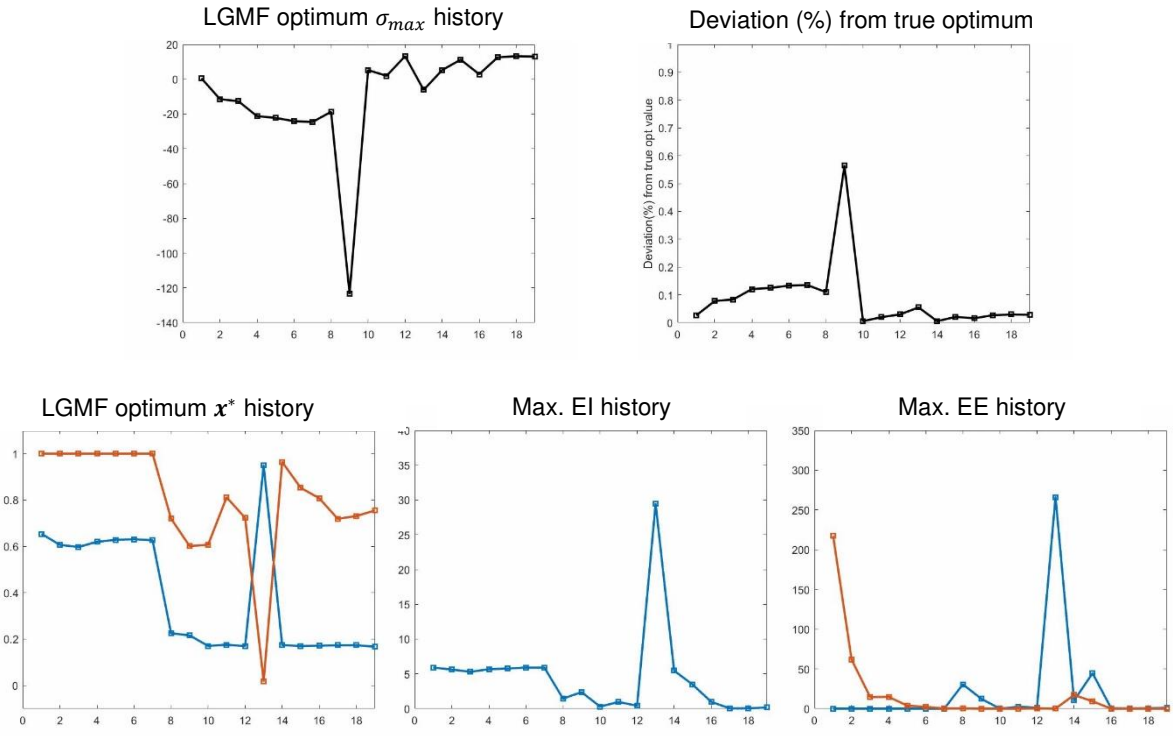



**Figure 49.** Iteration history of EE for LGMF.

The iteration histories for the "LGMF Predicted Optimum" and "Deviation from the Optimum" show eventual convergence to near the true optimum response prediction. The iteration history of the estimated optimum location $\boldsymbol{x}^*$ shows normal convergence towards the optimum, except at iteration 13. At that iteration LGMF produced a poor fit, as shown in Fig. 50. However, additional samples rapidly corrected the fit and the EE values of both

LF functions converged towards 0, while the EI value of the LGMF fit converged to a small constant value, as shown by Fig. 49.

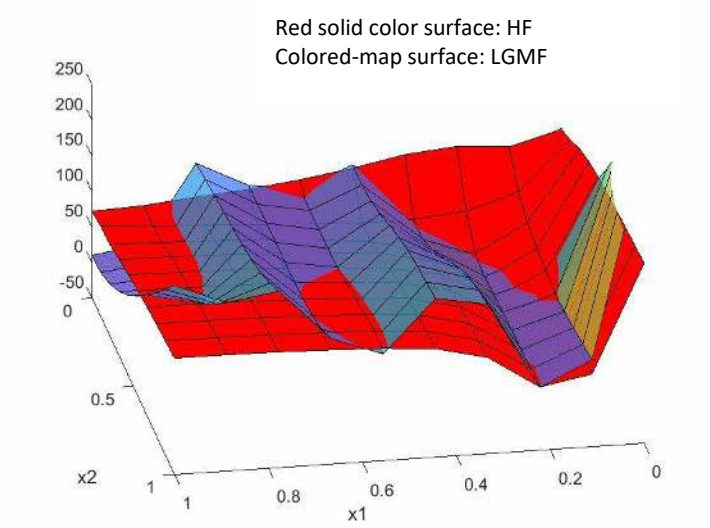
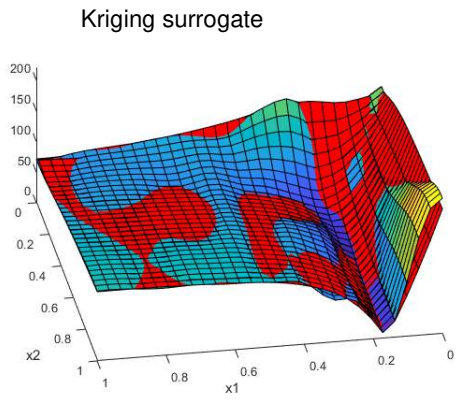
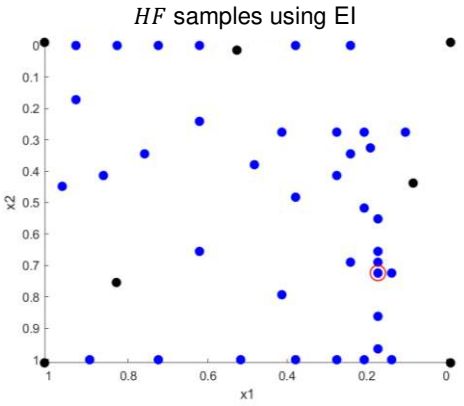



**Figure 50.** Poor LGMF fit at iteration 13.

For comparison, a single-fidelity test is run using EI for kriging, as shown in Fig. 51. EI took 45 HF evaluations instead of the EE method's 9. That is, EE required only 1/5$^{th}$ as many HF samples.




a) Final kriging fit. The true function is the red surface, while the kriging fit is the colormap surface.




b) Final data samples. Initial samples are black, while added samples are blue. The red point is the final sample at the optimum

**Figure 51.** Samples and surrogate for EI adaptive sampling using kriging.

### *4.2.4 Summary of Proposed EE Adaptive Sampling Method*

This section introduced the EE based adaptive sampling method for sequential multi-fidelity modeling approach in the framework of design optimization. This method leverages Multi-Fidelity (MF) models to address the question of how to orchestrate data acquisition from multiple available information sources which provide different approximated predictions with different costs. This adaptive sampling technique built off the Localized Galerkin Multi-Fidelity (LGMF) modeling method, which can provide MF modeling uncertainty and model dominance of the multiple low-fidelity models along with the approximated MF prediction. The Expected Effectiveness is proposed to account for not only the expected improvement, but also modeling uncertainty, dominance uncertainty and costs of all useful LF models by using LGMF and Non-Deterministic Kriging (NDK). The adaptive LGMF-NDK model also does not interpolate data, thereby improving stability and addressing randomness that pose challenges in other existing methods dealing with both physical experimental and high-fidelity computational data exhibiting white-noise and systematic measurement and prediction errors. It is demonstrated successfully that the proposed method enables adaptive MF modeling addressing the practical challenges with data under uncertainty and multiple LF data sources through multiple numerical examples.

## 4.3 Expected Usefulness (Adaptive Sampling of Constraints)

The previous section dealt with an adaptive sampling methodology to determine the global optimum. This section introduces an adaptive sampling method for determining contours, which is useful for finding constraint failure boundaries.

The Expected Feasibility Function (EFF) works well as an acquisition function for adaptively adding HF samples to an LGMF fit. However, adaptively adding LF samples requires a new methodology. In this section the Expected Usefulness (EU) acquisition function for adaptive sampling of LF models is introduced.

As in the section on EE, surrogates are created from a limited number of LF samples and used to construct the LGMF fit.

### *4.3.1 Changes to LGMF Implementation for Adaptive Sampling*

The changes to LGMF in this section are the same as the changes made in the EE section. First, lack of data from sparse initialization at the start of adaptive sampling can result in poor selection of the $h$ parameter during LOO optimization. In these examples the $h$ parameter is set to a user-defined constant. Second, a final stage of filtering using NDK helps improve the accuracy of the final LGMF model.

***Filtering of final LGMF model using NDK***: Because of lack of data, the LGMF response in Eq. 43 can lead to some discontinuities even with an optimal $h$ parameter, especially with a small number of data samples in the beginning of the adaptive sampling process. This problem is addressed by using a low-pass filtering process, in which the LGMF Stage 2 responses are resampled and applied to build an NDK model. The filtering parameter for the low-pass frequency can be determined based on the minimum distance of an expected stationary response. This NDK model is used as the final LGMF fit.

### *4.3.2 Proposed EU Adaptive Sampling for LGMF*

The Efficient Global Reliability Analysis (EGRA) methodology [24] was developed to evaluate the reliability of systems for engineering design. The method uses the Expected

Feasibility Function (EFF) metric as an acquisition function for adaptive sampling of a kriging surrogate model. This is useful for evaluating the failure boundary of a constraint. The metric balances between sampling locations predicted to be near the failure boundary and sampling locations with high uncertainty. The formulation of the EFF was given in Eq. 15.

When multiple constraints exist, it may not be necessary to find the contours of each constraint function everywhere. Contours in the infeasible regions of other constraints are redundant and do not need to be accurately found. The constraints $g$ only need to be sampled until their composite failure contour is known, at which point the feasible region is fully understood. This leads to the concept of a composite expected feasibility function, which was given in Eq. 16.

Based on these concepts, Expected Usefulness (EU) is defined, which combines the Composite EFF of the LGMF models with the individual EFF, Modeling Uncertainty (MU), Dominance under Uncertainty (DU) and evaluation cost of the LF model being evaluated,

$$EU(x,m) = CEFF_{LGMF}(x) \times EFF_{LGMF}(x,c) \times DU(x,m) \times MU(x,m)/Cost(m) \quad (75)$$

where $EFF_{LGMF}(x,m)$ is described by Eq. 15 except the mean and uncertainty of the gaussian process, $\mu_g$ and $\sigma_g$ respectively, are replaced by the mean and standard deviation of an LGMF model, $\mu_{LGMF}$ and $\sigma_{LGMF}$, respectively. Similarly, $CEFF_{LGMF}(x)$ is described by Eq. 16 except that the mean and uncertainty of the gaussian process, $\mu_g^*$ and $\sigma_g^*$ respectively, are replaced by the mean and standard deviation of an LGMF model, $\mu_{LGMF}^*$ and $\sigma_{LGMF}^*$ respectively.

DU is identical to its previous formulation for EE in Eq. 59. Because the given examples used deterministic kriging instead of NDK to model LF data, the MU formulation is changed to be either the Saturation of the LF model or the Scaled Uncertainty of the LF model, whichever is smaller.

$$MU(x,m) = \min\left(Saturation(x,m),\ ScaledUncertainty(x,m)\right) \tag{76}$$

Saturation is the ratio of Epistemic to Aleatory uncertainty in the NDK model of the LF function:

$$Saturation(x,m) = \sigma_{LFm\ NDK\ Epistemic}(x)/\sigma_{LFm\ NDK\ Aleatory}(x) \tag{77}$$

As more data points are added, the epistemic uncertainty will trend toward 0, the model will become saturated, and sampling of the LF function will cease. The Scaled Uncertainty is given by the ratio of the total uncertainty of the LF model to the scaled range of the LF data. The factor of 100 is added so the model does not converge prematurely

$$ScaledUncertainty(x,m) = 100 * \sigma_{LFm\ NDK}\ (x)\ /\ (\max(y_m) - \min(y_m)) \tag{78}$$

DU is defined as the dominance of the LF model plus the change in dominance that resulted from the last adaptive sample, that is

$$DU(x,m) = Dominance_{LGMF}\ (x,m) + \Delta Dominance_{LGMF}(x,m) \tag{79}$$

where the change in dominance for the $k^{th}$ iteration is calculated as

$$\Delta Dominance_{LGMF}(x,m) = Dominance^{k}_{LGMF}\ (x,m) - Dominance^{k-1}_{LGMF}(x,m) \tag{80}$$

Each iteration of the adaptive sampling, the constraint $c$, model $m$ and location $x$ with the highest EU value are sampled, and the corresponding LGMF fit is updated.

*4.3.3 Numerical Examples*

This section presents numerical examples for the proposed EU adaptive sampling method. The EU approach for multiple constraints is demonstrated with fundamental equations, as well as a nonlinear thermoelastic hat-stiffness aircraft panel problem.

Example 1: Estimation of Two 2D Constraints, Each with Two LF Models

This contour estimation example uses two constraints, both of which were used as examples in the EGRA paper [24]. The constraints are given by

$$g_{1_{HF}}(x_1, x_2) = (x_1^2 + 4) * (x_2 - 1)/20 \; - \sin(5/2 * x_1) - 2 \tag{81}$$

$$g_{2_{HF}}(x_1, x_2) = (x_1 + 2)^4 - x_2 + 4 \tag{82}$$

the LF approximations for the first constraint, which include bilinear and nonlinear deviations from the HF model are given by

$$g_{1_{LF1}} = 0.5 * g_{1_{HF}}(x_1, x_2) + x_1 * x_2 \tag{83}$$

$$g_{1_{LF2}}(x_1, x_2) = 2 * g_{1_{HF}}(x_1, x_2) + 0.2\, x_1^2\, x_2 + 0.3\, x_2^2 \tag{84}$$

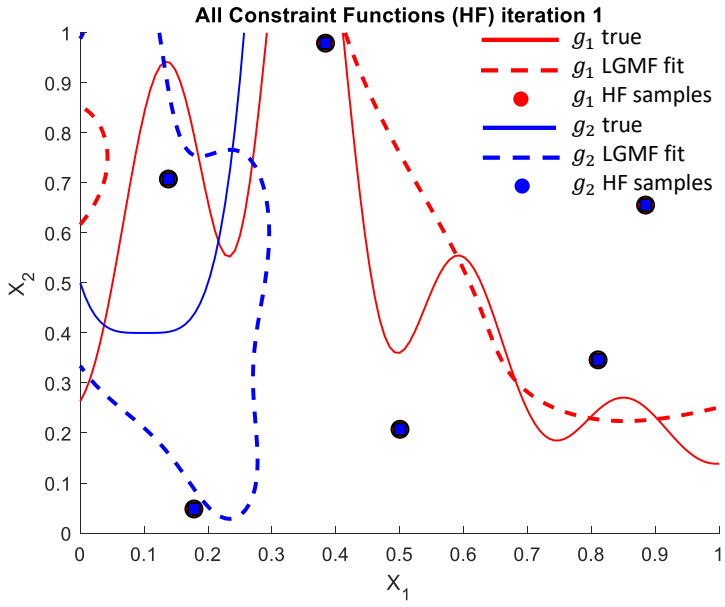
and the LF approximations for the second constraint have the same deviations as the first constraint, given by

$$g_{2_{LF1}} = 0.5 * g_{2_{HF}}(x_1, x_2) + x_1 * x_2 \tag{85}$$

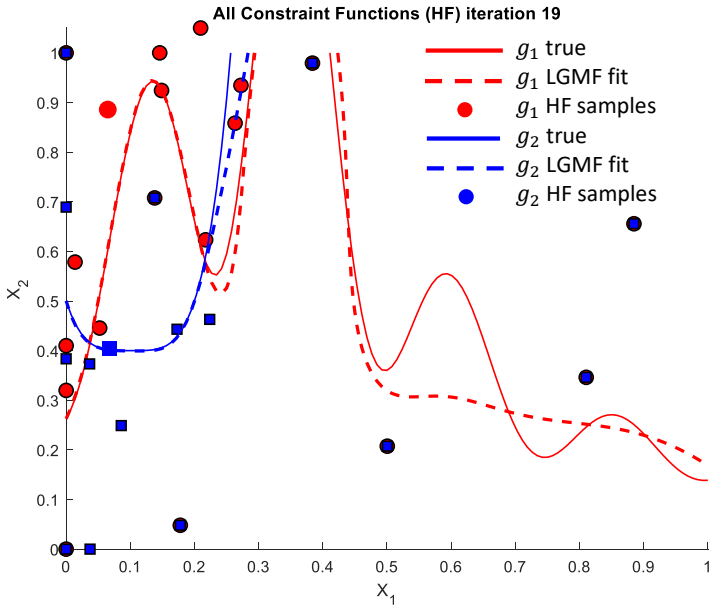
$$g_{2_{LF2}}(x_1, x_2) = 2 * g_{2_{HF}}(x_1, x_2) + 0.2\, x_1^2\, x_2 + 0.3\, x_2^2 \tag{86}$$

Both constraints are initialized with 6 HF samples selected using LHS design, and 12 samples for each LF function (4 at the corners and 8 selected using LHS design), for a total of 12 HF samples and 48 LF samples. The initial and final fits are compared with the true constraints in Fig. 52. The contour is only highly accurate at the boundary of the feasible region, and less accurate further away in the design space. In total 35 HF samples were

required for the accurate feasible region shown. The model also evaluated 66 LF samples total.




a) Initial fit.

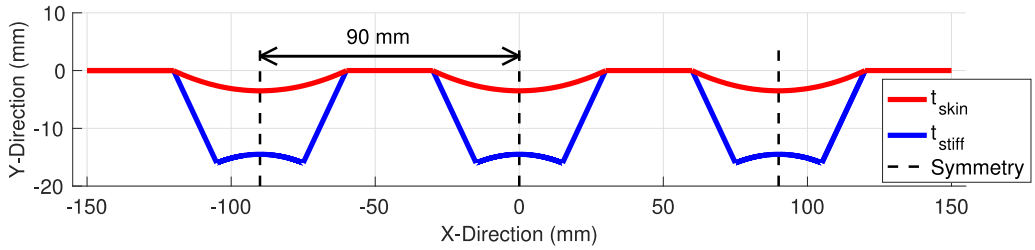



b) Final fit using Expected Usefulness in Example 1.

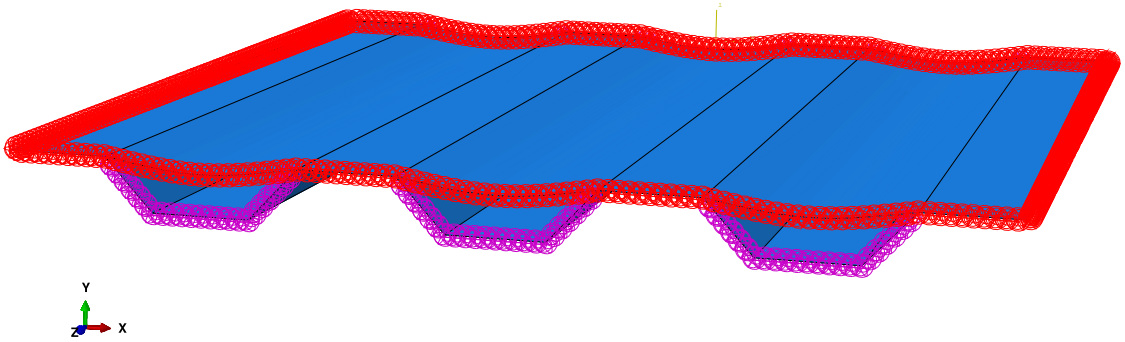
**Figure 52.** Example of adaptive sampling using EE

Example 2:
Feasibility bound study for 3D Nonlinear Thermoelastic Aircraft Panel problem

The nonlinear thermoelastic panel presented within this section is adapted from the hat-stiffened SR-71 like panel by Lee and Bhatia [48], leveraging spring Boundary Conditions (BCs) and TI-6242 following Deaton and Grandhi [46]. The panel is shown in Fig. 53. In this example, the parametric representation of Lee and Bhatia's 300 × 300 mm panel was achieved using five shape parameters: $W_{stiff}$, width of the hat-stiffener, $H_{stiff}$, height of the stiffener, $\eta_{skin}$, curvature of the top skin, $\eta_{stiff}$, curvature of the bottom of the hat, and $r_{ratio}$, the percentage of the bottom-stiffener width that transitions to the top of the panel. There are also two sizing parameters, $t_{skin}$ and $t_{stiff}$, Fig. 53a. This two-dimensional representation is extruded into the z-direction 300 mm to complete the panel, Fig. 53b with the spring BCs indicated by circles. For more details regarding this panel and its validation, see Clark et al. [49].




a) Panel parameterization, Fig. 9a from [46].




b) Panel assembly including spring boundary conditions, Fig. 10b from [46].

**Figure 53.** Parametric thermoelastic aircraft panel representation. a) the red skin region faces the environment and the blue stiffeners are internal.

The upper and lower bounds of each design variable are shown in Table 1.

**Table 1.** Design variables to be included and their descriptions. Distances are in meters.

| Design Variable | $x_1$ Hat Height | $x_2$ Hat Width | $x_3$ Hat Ramp Ratio | $x_4$ Delta Skin (outer skin curvature) | $x_5$ Delta Hat (hat bottom curvature) | $x_6$ Thickness Top | $x_7$ Thickness Bottom |
|---|---|---|---|---|---|---|---|
| Lower bound | 0.012 | 0.04 | 0.05 | -0.0075 | -0.003 | 0.002 | 0.002 |
| Upper bound | 0.02 | 0.08 | 0.45 | -0.0001 | 0.003 | 0.01 | 0.01 |

The panel is subject to two constraints: the stress may not exceed the maximum allowable value of 680.36 MPa, and the lowest natural frequency may not drop below the

minimum allowable value of 706.8 Hz. If the lowest resonant frequency drops below the allowable value, vibrations in flight may cause the panel to flutter and fail. The maximum stress in the panel, and the lowest natural frequency, is calculated using FEA analysis in Abaqus. Two LF models of the hat panel are also considered. For the same FE model, the first LF model uses a linear solver instead of the nonlinear Riks solver. The second LF model also uses the linear solver, while aggressively simplifying the FE model as a thin strip model constrained to the x-y plane, as shown in Fig. 54. To avoid the stress concentration, stress in elements near the ends is not considered.

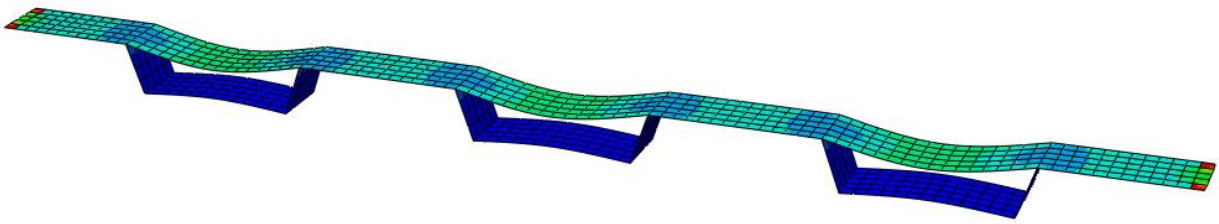


**Figure 54.** Thin strip model used for second LF model. Colors indicate stress values.

The computational cost differences are not significant in this example because the HF model is already defeatured and simplified. In an actual design the FE model may include more details including fasteners, off-set connections, multiple materials, combination of different stiffeners, etc., which would cause a wide range of cost differences for the FE simulations. In this abstract example problem only the first three variables, Hat Height, Hat Width, and Hat Ramp Ratio are considered. The EU-LGMF adaptive model was initialized with 256 samples from each LF model (8 at the corners and 248 LHS samples), for a total of 1024 LF samples. The HF models were initialized with 10 LHS samples each, for a total of 20 HF samples.

For each iteration, the approximate accuracy of the model is estimated using 2000 LHS samples. By comparing the predicted feasibility at these points to the actual feasibility, it is possible to compute the percentage of feasible points that are not predicted as feasible (False Negative), which may result in an overly conservative design. It is also possible to compute the percentage of points which are predicted to be feasible but are not (False Positive), which can result in system failure. These metrics are recorded over the iterations as shown in Fig. 55.

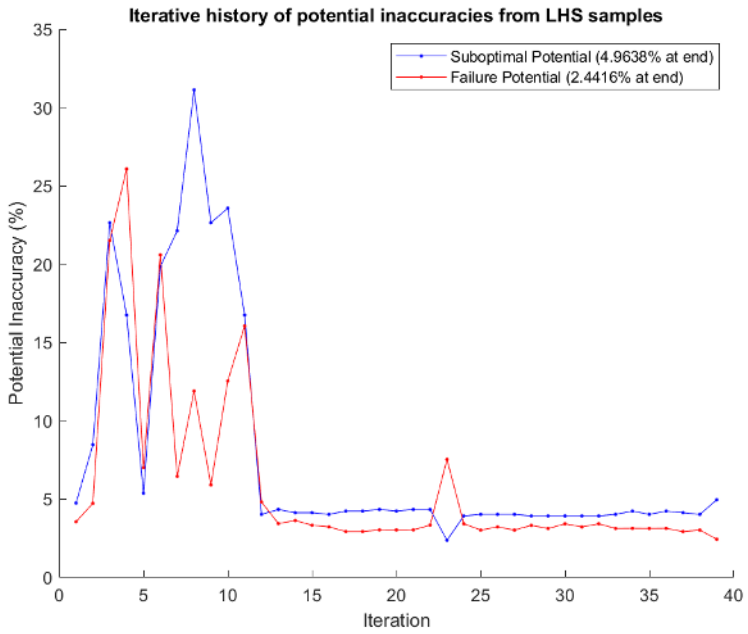



**Figure 55.** Percent of points that were feasible but predicted to be infeasible (blue line) or predicted to be feasible but were not (red line).

The surrogate model is surprisingly accurate from the beginning, with only a 4.8% false negative rate and 3.6% false positive rate. The early accuracy probably occurred by chance, as adding more information causes the model to drop in accuracy before returning to a more accurate solution. The optimization ended with a total of 66 HF samples, 179 LF

stress and frequency evaluations from the linear model, and 177 LF stress and frequency evaluations from the strip model.

### *4.3.4 Summary of Proposed EU Adaptive Sampling Method*

This section introduced the EU acquisition function for sequential multi-fidelity modeling for contour estimation. This method addresses the question of how to orchestrate data acquisition from multiple available information sources which provide different approximated predictions with different costs. The method computes the composite feasible region, i.e. the region that is feasible when all constraints are included. By ignoring redundant constraint boundaries and exploiting low fidelity data sources, the method greatly reduces the required number of high-fidelity samples and by extension the computational cost.

This adaptive sampling technique built off the Localized Galerkin Multi-Fidelity (LGMF) modeling method, which can provide modeling uncertainty and model dominance values of multiple low-fidelity models along with the approximated MF prediction. EU is formulated as a function of composite Expected Feasibility, individual Expected Feasibility, model Dominance under Uncertainty, Modeling Uncertainty, and Cost of evaluation of each LF model. HF models are evaluated using composite Expected Feasibility and individual Expected Feasibility values. The proposed adaptive sampling approach was demonstrated with a numerical example and a three dimensional nonlinear thermoelastic hat-stiffened aircraft panel problem.

# V. CONCLUSIONS

This thesis work introduced a novel multi-fidelity modeling framework designed to reduce the time and cost of engineering design when using computer simulations. The framework leverages surrogate models, which inexpensively approximate computer model outputs using data from a limited number of runs. The proposed localized-Galerkin multi-fidelity surrogate modeling method addresses the practical challenges encountered in Aerospace vehicle design when 1) multiple low-fidelity models exist, 2) each low-fidelity model may only be correlated with the high-fidelity model in part of the design domain, and 3) models may contain noise or uncertainty. The proposed approach consolidates multiple low-fidelity models into a single model by using the localized Galerkin formulation. The method has been successfully demonstrated using fundamental mathematical problems with two LF models, which show localized correlations within different local domains.

Two adaptive sampling methods were also introduced to iteratively select new data samples in regions of the design space where increased accuracy is important. The first acquisition formulation, Expected Effectiveness (EE), searches for the global optimum and is intended to model engineering objectives. EE is used to sample the LF models and accounts not only for Expected Improvement (EI), but also Modeling Uncertainty, Dominance under Uncertainty, and model Cost. EI is used to sample the HF model. In this section the localized-Galerkin multi-fidelity method is combined with Non-Deterministic Kriging (NDK) to form a model that does not interpolate data. This enables the method to handle randomness in both HF and LF data that pose challenges in other existing methods dealing with both physical experimental and computational data exhibiting white-noise

errors. In multiple numerical examples, the method was successfully demonstrated to enable adaptive MF modeling while addressing the practical challenges associated with data under uncertainty and the existence of multiple LF data sources.

The second adaptive sampling formulation, EU, estimates contours to identify feasible design domains and is intended to model engineering constraints. The method computes the composite feasible region, i.e. the region that is feasible when all constraints are included. By ignoring redundant constraint boundaries and exploiting low fidelity data sources, the method greatly reduces the required number of high-fidelity samples and by extension the computational cost. Each constraint is approximated by an LGMF model. EU is formulated as a function of composite Expected Feasibility, individual Expected Feasibility, model Dominance under Uncertainty, Modeling Uncertainty, and Cost of evaluation of each LF model. HF models are evaluated using the composite Expected Feasibility and individual Expected Feasibility values. The proposed adaptive sampling approach was demonstrated with multiple examples.

**Future Work**

A promising area of research for future work is handling multi-fidelity constrained optimization problems where constraints and objectives both require expensive simulations to evaluate. In such situations extraneous regions of the design space may be ignored, i.e., feasibility of a sub-optimal region of the design space is irrelevant, as is optimality of an infeasible region of the design space. Therefore, considering constraints and objectives simultaneously can result in significant cost savings.

Care must be taken to avoid an acquisition function that tends to be overly non-smooth, where most of the design space is unpromising (acquisition value near 0) and sudden

"spikes" exist in a few local regions. Especially in high-dimensional cases, this may make the optimization required to maximize the acquisition function computationally expensive, offsetting the savings gained from fewer simulation runs.

Another area for future study is extension of LGMF to high-dimensional problems. Any Kriging-based modeling including NDK is impractical for modeling non-stationary system responses or problems above 10-20 dimensions, and LGMF itself develops prohibitive memory requirements above around 15-20 dimensions.

Reduced modeling techniques such as dimensionality reduction may be able to cope with this issue. Dimensionality reduction methods are used to reduce a problem from $d$-diminsional space to $n$-dimensional space, where $n < d$, while retaining the major contents of information. Methods include Principal Component Analysis (PCA), Partial Least Squares Regression (PLSR), and autoencoders. Replacing Kriging-based methods with more scalable surrogates such as Deep Gaussian Processes [50-53] or Artificial Neural Networks [6, 54-55] may also alleviate the curse of dimensionality. Finally, data-clustering methods [56] may allow subsets of the design space to be analyzed separately, reducing computational challenges.

# VI. REFERENCES


[1] Forrester, A.I., and Keane, A.J., "Recent advances in surrogate-based optimization," *Progress in Aerospace Sciences*, Vol. 45, No. 1, 2009, pp.50-79.

[2] Queipo, N.V., Haftka, R.T., Shyy, W., Goel, T., Vaidyanathan, R., and Tucker, P.K., "Surrogate-based analysis and optimization," *Progress in aerospace sciences*, Vol. 41, No. 1, 2005, pp.1-28.

[3] Wang, G. G., and Shan, S., "Review of metamodeling techniques in support of engineering design optimization," *Journal of Mechanical Design*, Vol. 129 No. 4, 2007, pp. 370-380.

[4] Wang, G. G., Dong, Z., and Aitchison, P., "Adaptive response surface method—a global optimization scheme for approximation-based design problems," *Engineering Optimization*, Vol. 33, No. 6, 2001, pp. 707-734.

[5] Kim, J. R., and Choi, D. H., "Enhanced two-point diagonal quadratic approximation methods for design optimization," *Computer methods in applied mechanics and engineering*, Vol. 197, No. 6, 2008, pp. 846-856.

[6] Park, J., and Sandberg, I. W., "Universal approximation using radial-basis-function networks," *Neural computation*, Vol. 3, No. 2, 1991, pp. 246-257.

[7] Henry, E. B, Brown, J. M., Beck J. A., and Kaszynski, A. A., "Mistuned Rotor Reduced Order Modeling with Surrogate-Modeled Airfoil Substructures," 58th AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, AIAA SciTech Forum, AIAA 2017-1600, https://doi.org/10.2514/6.2017-1600

[8] Sacks, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P., "Design and analysis of computer experiments," *Statistical Science*, Vol. 4, No 4, 1989, pp. 409–435.

[9] Jones, D. R., Schonlau, M., and Welch, W. J., "Efficient global optimization of expensive black-box functions," *Journal of Global optimization*, Vol. 13, No. 4, 1998, pp.455-492.

[10] Bae, H., Clark, D., and Forster, E., "Non-Deterministic Kriging for Engineering Design Exploration," AIAA Journal, Vol. 57, No. 4, April, 2019.

[11] Gano, S. E., Renaud, J. E., Martin, J. D. , and Simpson, T. W., "Update strategies for kriging models for using in variable fidelity optimization, 46th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference, Structures, Structural Dynamics, and Materials and Co-located Conferences , AIAA Paper 2005–2057, April 2005, https://doi-org.wrs.idm.oclc.org/10.2514/6.2005-2057.

[12] Eldred, M.S., Giunta, A.A., Collis, S.S., Alexandrov, N.A., and Lewis, R.M., "Second-order corrections for surrogate-based optimization with model hierarchies," 10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, Multidisciplinary Analysis Optimization Conferences, AIAA Paper 2004-4457, Aug. 2004. https://doi-org.wrs.idm.oclc.org/10.2514/6.2004-4457.

[13] Kennedy, M. C., and O'Hagan, A., "Bayesian calibration of computer models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Vol. 63, No. 3, 2001 pp. 425-464.

[14] Qian, P. Z., and Wu, C. J., "Bayesian hierarchical modeling for integrating low-accuracy and high-accuracy experiments," *Technometrics*, Vol. 50, No. 2, 2008, pp. 192–204.

[15] 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*, Vol. 25, No. 1, 2013, pp. 177–189. doi:10.1016/j.ast.2012.01.006

[16] Kennedy, M., and O'Hagan, A., "Predicting the output from a complex computer code when fast approximations are available", *Biometrika*, Vol. 87, No. 1, 2000, pp. 1-13, https://doi.org/10.1093/biomet/87.1.1

[17] Peherstorfer, B., Willcox, K., and Gunzburger, M., "Survey of multifidelity methods in uncertainty propagation, inference, and optimization.", ACDL Tech. Report, TR16-1, Preprint, June 2016, pp.1-57.

[18] Park, C., Haftka, R.T., and Kim, N.H., "Remarks on multi-fidelity surrogates," *Structural and Multidisciplinary Optimization*, Vol. 55, No. 3, 2017, pp.1029-1050.

[19] Alexandrov, N. M., Lewis, R. M., Gumbert, C. R., and Green, L. L., and Newman, P. A., "Optimization with variable-fidelity models applied to wing design", 38th Aerospace Sciences Meeting and Exhibit, Aerospace Sciences Meetings, AIAA Paper 2000-0841, https://doi-org.wrs.idm.oclc.org/10.2514/6.2000-841.

[20] Alexandrov, N. M., Dennis, J. E., Lewis, R. M., and Torczon, V., "A trust-region framework for managing the use of approximation models in optimization," *Structural and Multidisciplinary Optimization*, Vol. 15, No. 1, 1998, pp. 16–23.

[21] Fischer, C. C., Grandhi, R. V., and Beran, P. S., "Bayesian Low-Fidelity Correction Approach to Multi-Fidelity Aerospace Design," 58th AIAA/ASCE/AHS/ASC Structural Dynamics, and Materials Conference, AIAA SciTech Forum, AIAA 2017-0133, https://doi.org/10.2514/6.2017-0133.

[22] Rumpfkeil, M. P., and Beran, P., "Construction of Dynamic Multifidelity Locally Optimized Surrogate Models", *AIAA Journal*, Vol. 55, No. 9, September 2017, pp. 3169-3179.

[23] Kuya, Y., Takeda, K., Zhang, X., and Forrester, A. I. J., "Multifidelity Surrogate Modeling of Experimental and Computational Aerodynamic Data Sets", *AIAA Journal*, Vol. 49, No. 2, 2011, pp. 289-298.

[24] Bichon, B. J., Eldred, M. S., Swiler, L. P., Mahadevan, S., and McFarland, J. M., "Efficient global reliability analysis for nonlinear implicit performance functions," *AIAA Journal*, Vol. 46, No. 10, 2008, pp. 2459-2468.

[25] Viana, F. A. C., Haftka, R. T., and Watson, L. T., "Efficient global optimization algorithm assisted by multiple surrogate techniques," *Journal of Global Optimization*, Vol. 56, Iss. 2, June 2013, pp. 669-689.

[26] Clark, D., Bae, H., Deaton, J., and Forster, E., "Adaptive Infill Criteria for Non-Deterministic Kriging Considering Aleatory and Epistemic Uncertainties", *AIAA Scitech 2019 Forum*, 2019, doi: 10.2514/6.2019-1489
[27] Clark, D. L. and Bae, H., "Non-Deterministic Kriging Framework for Responses with Mixed Uncertainty," 19th AIAA Non-Deterministic Approaches Conference, AIAA SciTech Forum, AIAA 2017-0593, https://doi-org.wrs.idm.oclc.org/10.2514/6.2017-0593
[28] Pellegrini, R., Serani, A., Diez, M., Wackers, J., Queutey, P., and Visonneau, M., "Adaptive sampling criteria for multi-fidelity metamodels in CFD-based shape optimization.", *ECCM-ECFD*, Glasgow, Scotland, 2018.
[29] Diez, M., Volpi, S., Serani, A., Stern, F., and Campana, E., "Simulation based design optimization by sequential multi-criterion adaptive sampling and dynamic radial basis functions.", *EUROGEN 2015, International Conference on Evolutionary and Deterministic Methods for Design, Optimization and Control with Applications to Industrial and Societal Problems*, 2015.
[30] Chaudhuri, A., Lam, R., and Willcox, K., "Multifidelity Uncertainty Propagation via Adaptive Surrogates in Coupled Multidisciplinary Systems", *AIAA Journal*, Vol. 56, No. 1, pp. 235-249, 2018, https://dx.doi.org/10.2514/1.J055678
[31] Huang, D., Allen, T., Notz, W., and Miller, R., "Sequential kriging optimization using multiple-fidelity evaluations", *Structural and Multidisciplinary Optimization*, Vol. 32, No. 5, 2006, pp. 369-382, doi: 10.1007/s00158-005-0587-0
[32] Forrester, A., Sóbester, A., and Keane, A., "Multi-fidelity optimization via surrogate modelling", *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, Vol. 463, No. 2088, 2007, pp. 3251-3269, https://doi.org/10.1098/rspa.2007.1900
[33] Le Gratiet, L., and Cannamela, C., "Cokriging-Based Sequential Design Strategies Using Fast Cross-Validation Techniques for Multi-Fidelity Computer Codes", *Technometrics*, Vol. 57, No. 3, 2015, pp. 418-427, doi: 10.1080/00401706.2014.928233
[34] Kandasamy, K., Dasarathy, G., Oliva, J., Schneider, J., and Poczos, B., "Gaussian process bandit optimization with multi-fidelity evaluations.", *Advances in Neural Information Processing Systems*, 2016, https://arxiv.org/pdf/1603.06288v4.pdf
[35] Moore, R., Romero, D., and Paredis, C., 2014. "Value-based global optimization." *Journal of Mechanical Design*, Vol. 136, No. 4, p. 041003.
[36] Poloczek, M., Wang, J., and Frazier, P., "Multi-Information Source Optimization", *Advances in Neural Information Processing Systems*, Vol. 30, Curran Associates, Inc., 2017, pp. 4288-4298, http://papers.nips.cc/paper/7016-multi-information-source-optimization.pdf.
[37] Ranjan, P., Bingham, D., & Michailidis, G., "Sequential experiment design for contour estimation from complex computer codes.", 2008, *Technometrics*, *50*(4), 527-541, doi: 10.1198/004017008000000541
[38] Marques, A. N., Lam, R. R., Willcox, K. E., "Contour location via entropy reduction leveraging multiple information sources", *32nd Conference on Neural Information Processing Systems (NeurIPS 2018),* 2018, URL http://papers.nips.cc/paper/7768-contour-location-via-entropy-reduction-leveraging-multiple-information-sources.pdf
[39] Gómez-Hernández, J. J., "Geostatistics." *Hydrogeophysics*, 2005, pp. 59-83, Springer, Dordrecht, doi: 10.1007/1-4020-3102-5_3
[40] Lophaven, S. N., Nielsen, H. B., Sondergaard, J., "Dace, a MATLAB Kriging Toolbox", *Technical University of Denmark*, 2002, Kongens Lyngby, Technical Report No. IMMTR-2002, 12.
[41] Bae, H., Beachy, A., Clark, D. L., Deaton, J., Forster, E. E. "Multi-Fidelity Modeling using Non-Deterministic Localized-Galerkin Approach", *AIAA Scitech 2019 Forum*, 2019, doi: 10.2514/6.2019-1999.
[42] Bae, H., Beachy, A.J., Clark, D.L., Deaton, J.D., and Forster, E.E., "Multifidelity Modeling Using Nondeterministic Localized Galerkin Approach." *AIAA Journal*, 2020, 58(5), pp.2246-2260, doi: 10.2514/1.J058410
[43] Beachy, A. J., Clark, D. L., Bae, H., and Forster, E. E., "Expected Effectiveness Based Adaptive Multi-Fidelity Modeling for Efficient Design Optimization," *AIAA Scitech 2020 Forum*, 2020, p. 1144. https://doi.org/10.2514/6.2020-1144.
[44] Bachoc, F., "Cross validation and maximum likelihood estimations of hyper-parameters of Gaussian processes with model misspecification," *Computational Statistics & Data Analysis*, Vol. 66, 2013, pp.55-69.
[45] Boopathy, K., and Rumpfkeil,M. P., "A Unified Framework for Training Point Selection and Error Estimation for Surrogate Models," *AIAA Journal*, Vol. 53, No. 1, 2015, pp. 215–234. doi:10.2514/1.J053064.
[46] Deaton, J. D. and Grandhi, R. V., "Significance of Geometric Nonlinearity in the Design of Thermally Loaded Structures," *Journal of Aircraft*, Vol. 52, No. 4, 2014, pp. 1226–1234.
[47] R. Craig, M. Bampton "Coupling of substructures for dynamic analysis" Amer. Inst. Aero. Astro. J., 6 (7) (1968), pp. 1313-1319
[48] Lee, J. and Bhatia, M., "Impact of corrugations on bifurcation and thermoelastic responses of hat-stiffened panels," Thin- Walled Structures, Vol. 140, 2019, pp. 209–221, DOI:10.1016/j.tws.2019.03.027
[49] Clark, D. L., Bae, H., and Forster, E. E., "Gaussian Surrogate Dimension Reduction for Efficient Reliability-Based Design Optimization," *SciTech 2020*, Abstract Accepted.
[50] Damianou, A., and Lawrence, N., "Deep gaussian processes," *Artificial Intelligence and Statistics*, 2013, pp. 207–215.

[51] Hebbal, A., Brevault, L., Balesdent, M., Talbi, E.-G., and Melab, N., "Bayesian optimization using deep Gaussian processes," *arXiv preprint* arXiv:1905.03350, 2019.
[52] Hebbal, A., Brevault, L., Balesdent, M., Talbi, E.-G., and Melab, N., "Multi-objective optimization using Deep Gaussian Processes: Application to Aerospace Vehicle Design," *AIAA Scitech 2019 Forum*, 2019, p. 1973.
[53] Cutajar, K., Pullin, M., Damianou, A., Lawrence, N., and González, J., "Deep gaussian processes for multi-fidelity modeling," *arXiv preprint* arXiv:1903.07320, 2019.
[54] Pearl, J., "Bayesian networks," *UCLA: Department of Statistics*, UCLA, 2011. URL https://escholarship.org/uc/item/53n4f34m.
[55] Mullachery, V., Khera, A., and Husain, A., "Bayesian neural networks," *arXiv preprint* arXiv:1801.07710, 2018.
[56] Scarth, C., Sartor, P. N., Cooper, J. E., Weaver, P. M., and Silva, G. H., "Robust Aeroelastic Design of Composite Plate Wings," *17th AIAA Non-Deterministic Approaches Conference*, 2015. https://doi.org/10.2514/6.2015-0918.