

# Multifidelity Uncertainty Propagation in Coupled Multidisciplinary Systems

Anirban Chaudhuri\* and Karen Willcox†

*Massachusetts Institute of Technology, Cambridge, MA, 02155, USA*

Fixed point iteration is a common strategy to handle interdisciplinary coupling within a coupled multidisciplinary analysis. For each coupled analysis, this requires a large number of disciplinary high-fidelity simulations to resolve the interactions between different disciplines. When embedded within an uncertainty analysis loop (e.g., with Monte Carlo sampling over uncertain parameters) the number of high-fidelity disciplinary simulations quickly becomes prohibitive, since each sample requires a fixed point iteration and the uncertainty analysis typically involves thousands or even millions of samples. This paper develops a method for uncertainty analysis in feedback-coupled black-box systems that leverages adaptive surrogates to reduce the number of cases for which fixed point iteration is needed. The multifidelity coupled uncertainty propagation method is an iterative process that uses surrogates for approximating the coupling variables and adaptive sampling strategies to refine the surrogates. The adaptive sampling strategies explored in this work are residual error, information gain, and weighted information gain. The surrogate models are adapted in a way that does not compromise accuracy of the uncertainty analysis relative to the original coupled high-fidelity problem.

## I. Introduction

MULTIDISCIPLINARY analysis and optimization is an extensive area of research, intended to take into account the interactions between multiple disciplines working towards building an efficient engineering system.<sup>1–10</sup> Optimization of such complex systems often pushes the designs to the brink of failure, which makes accounting for inherent system uncertainties paramount. Historically, uncertainties are accounted for by using safety factors; however, the drive to enhance efficiency and robustness of a system requires more rigorous uncertainty characterization methods. To be practically applicable, these methods must also be computationally efficient. This has led to significant research in uncertainty analysis for multidisciplinary optimization; a review of such methods can be found in Yao et al.<sup>10</sup> The task of uncertainty analysis—forward propagation of uncertainty through a system—is particularly challenging for a multidisciplinary system due to coupling between different disciplines. This coupling could be feed-forward (one-directional) or feedback (bidirectional). Feed-forward coupling is usually easier to deal with; it has been tackled using approximations such as surrogates<sup>11</sup> and decomposition combined with recomposition through importance sampling.<sup>12</sup> Here, we focus on the more complicated case of a feedback-coupled system. We treat all disciplinary analyses as black-boxes (i.e., they can be viewed in terms of their inputs and outputs, and knowledge of their internal mechanisms is not needed) and we develop a non-intrusive approach that does not modify the disciplinary analysis.

A generic feedback-coupled multidisciplinary system is shown in Figure 1. The reader is referred to Cramer et al.<sup>13</sup> for a more detailed discussion on problem formulation for multidisciplinary systems. For the example depicted, there are three disciplines, with discipline 1 and discipline 2 being feedback-coupled through coupling variables  $C_1$  and  $C_2$  (here, the number of coupling variables is  $N_C = 2$ ). The coupling variable  $C_1$  is an input for discipline 2 and an output for discipline 1. The coupling variable  $C_2$  is an input for discipline 1 and an output for discipline 2. Let  $C_1^*$  and  $C_2^*$  be the multidisciplinary feasible solution of the coupling variables for the feedback-coupled system, which is achieved when the interdisciplinary

---

\*Postdoctoral Associate, Department of Aeronautics and Astronautics, anirbanc@mit.edu, AIAA member.

†Professor, Department of Aeronautics and Astronautics, kwillcox@mit.edu, AIAA Associate Fellow.

coupling constraints are satisfied.<sup>13</sup> The multidisciplinary system is described by the set of high-fidelity models  $\mathbf{H} = \{H_{C_1}(C_2, \mathbf{x}), H_{C_2}(C_1, \mathbf{x}), H_{y^{(1)}}(C_1, C_2, \mathbf{x}), H_{y^{(2)}}(C_1, C_2, \mathbf{x}), H_S(\mathbf{y}, C_1, C_2, \mathbf{x})\}$  with input vector  $\mathbf{x} \in \mathbb{R}^{dim}$ , coupling variables  $C_1$  and  $C_2$ , disciplinary output vector  $\mathbf{y} = [y^{(1)}, y^{(2)}]^T$ , and system output  $S$ . The inputs could be independent or shared inputs for different disciplines. Note that in this work we are treating the coupling variables and outputs of each discipline as scalars but in general they could be multidimensional vectors. All theories developed in this work can be extended to such cases. Also note that while Figure 1 shows only two disciplines, our approach extends to the case of multiple disciplines.

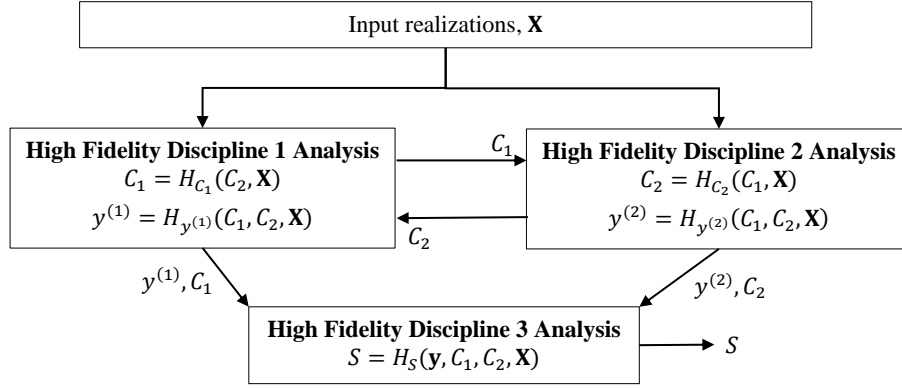


Figure 1. Feedback coupled multidisciplinary system

The aim of this work is to quantify uncertainty in disciplinary outputs  $\mathbf{y}$  and system output  $S$  due to uncertainty in inputs  $\mathbf{x}$ . We consider uncertainty in the inputs defined by some probability distribution function,  $\pi_{\mathbf{x}}$ . To estimate the corresponding uncertainty in  $\mathbf{y}$ , we use Monte Carlo simulation. Consider  $N_{total}$  input realizations  $\mathbf{x}_1, \dots, \mathbf{x}_{N_{total}}$  drawn randomly from  $\pi_{\mathbf{x}}$ . The Monte Carlo simulation propagates each of these samples through the system analysis and generates the corresponding output sample  $\mathbf{y}_1, \dots, \mathbf{y}_{N_{total}}$ . Let  $\mathbf{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_{N_{total}}\}$  be the set of  $N_{total}$  input realizations. For each input realization, we must solve the coupled multidisciplinary system. A common method to do this is through fixed point iteration (FPI).<sup>13–15</sup> In FPI, the outputs from one discipline are fed as inputs to the coupled discipline and this process iterates until a multidisciplinary feasible solution is reached. For the depicted example, at a given realization  $\mathbf{x}_i$ , the FPI is initialized with an arbitrary value of one of the coupling variables. For example, an initial guess for  $C_1$  is used for discipline 2 to calculate the value of  $C_2$  and feed that value of  $C_2$  to discipline 1 to calculate  $C_1$ . The output  $C_1$  from discipline 1 is again fed into discipline 2 and the process continues until the multidisciplinary feasible solution  $[C_1^*, C_2^*]^T$  is reached. Note that here we assume  $\exists [C_1^*, C_2^*]^T \forall \mathbf{x}_i$ . Each iteration within the FPI requires disciplinary high-fidelity solves, thus the method becomes computationally intensive as the number of iterations for FPI convergence increases. When FPI is embedded within Monte Carlo simulation for uncertainty propagation, the number of high-fidelity simulations can quickly become computationally prohibitive.

Past work has dealt with uncertainty analysis of feedback-coupled systems by using decoupling approaches such as collaborative reliability analysis,<sup>16</sup> the first order reliability method (FORM)<sup>17</sup> and a likelihood-based approach,<sup>18</sup> and polynomial chaos expansion based method.<sup>19</sup> FORM and the likelihood-based decoupling approach cast the feedback-coupled system as a feed-forward system, thus avoiding coupled system analyses and their associated FPI, but at the cost of neglecting the dependence between the inputs and the coupling variables. This can be an effective approach when the sensitivity of the system output to coupling variables is low, but can lead to poor results when sensitivity to coupling variables is high.

In this paper, we introduce an approach that uses surrogate models to reduce the number of cases for which FPI is executed in the Monte Carlo simulation of a feedback-coupled black-box system, while not compromising the accuracy of the uncertainty analysis relative to the original coupled high-fidelity problem. In particular, we formulate an iterative multifidelity coupled uncertainty propagation method that leverages surrogate models for approximating the coupled variables and adaptive sampling for refining the surrogates. The adaptive sampling strategies explored in this work are maximizing residual error, information gain, and weighted information gain. Entropy-based or information-gain-based approaches have been used previously for global optimization<sup>20–22</sup> and optimal experimental design.<sup>23, 24</sup> Here, we develop an information-gain-

based metric in the context of uncertainty propagation in multidisciplinary systems.

The remainder of the paper is organized as follows. Section II provides the details of the multifidelity coupled uncertainty propagation method. Section III presents numerical experiments that compare the efficiency of the proposed method to the standard FPI approach. Section IV provides the conclusions.

## II. Multifidelity Coupled Uncertainty Propagation

We propose a multifidelity coupled uncertainty propagation method that reduces the number of cases for which we need to perform FPI and thus reduces the number of required high-fidelity disciplinary simulations. This section gives an overview of the approach and describes the adaptive sampling strategies.

### A. Approach Overview

Figure 2 shows the flowchart detailing the multifidelity coupled uncertainty propagation method. This flowchart uses the same definitions from the example described in Figure 1, although again note that the approach could be extended to multiple disciplines and multi-dimensional coupling variables. The method is an iterative process that uses surrogate models for approximating the coupling variables, denoted by  $\tilde{C}$ . Here we use an interpolating Kriging surrogate,<sup>25,26</sup> but one could use any surrogate that is equipped with an uncertainty estimate.

Given the set  $\mathbf{X}$  of  $N_{total}$  input realizations, we first select a subset of  $N_{sur}$  initial samples,  $\mathbf{X}_{sur}^0 \subset \mathbf{X}$ . The superscript 0 represents the 0<sup>th</sup> cycle of the method to denote initial samples for building the surrogates. The subset selection can be conducted in any way, but here we maximize the minimum distance between the samples. Second, FPI is conducted for the coupled system at each of the  $N_{sur}$  samples to obtain the multidisciplinary feasible solutions of the coupling variables at those samples. We denote the resulting vector of multidisciplinary feasible solutions of the  $i^{th}$  coupling variable as  $\mathbf{C}_i^*(\mathbf{X}_{sur}^0)$ . Third, the dataset  $\{\mathbf{X}_{sur}^0; \mathbf{C}_i^*(\mathbf{X}_{sur}^0)\}$  is used to build the surrogates for the  $i^{th}$  coupling variable. Then the set of remaining input realizations,  $\mathbf{X} \setminus \mathbf{X}_{sur}^0$ , and the surrogate predictions for the coupling variables are propagated through the respective high-fidelity disciplinary analysis to estimate the outputs. Note that FPI is not performed at this point and the high-fidelity simulations of the coupled disciplines can be decoupled as seen in Figure 2.

A normalized residual error metric  $e^t$  at the current cycle  $t$  for the approximation of the coupling variables is defined as

$$e^t(\mathbf{x}) = \sum_{i=1}^{N_C} \frac{|\tilde{C}_i^t(\mathbf{x}) - C_i^{out}(\mathbf{x})|}{\kappa_i}, \quad (1)$$

where  $\tilde{C}_i^t$  is the surrogate model prediction at the current cycle  $t$ ,  $C_i^{out}$  is the output from the high-fidelity analysis of the  $i^{th}$  coupling variable,  $\kappa_i$  is the normalization constant for the residual errors of the  $i^{th}$  coupling variable, and  $N_C$  is the number of coupling variables. The normalization constants are problem specific and will be specified with the respective test problems. The outputs for the  $\mathbf{x}$  realizations that satisfy the residual error tolerance criterion  $e^t(\mathbf{x}) \leq \varepsilon$ , where  $\varepsilon$  is a user-defined residual error tolerance, are accepted.

Let  $\mathbf{X}_{rem}^t = \{\mathbf{x} \in \mathbf{X} \setminus \mathbf{X}_{sur}^t : e^t(\mathbf{x}) > \varepsilon\}$  be the set of  $N_{rem}^t$  realizations that did not satisfy the error tolerance. An adaptive sampling strategy (described in detail in the next sub-section) is used to select one of these realizations,  $\mathbf{x}_t^* \in \mathbf{X}_{rem}^t$ . For the selected realization, FPI is used to solve for the multidisciplinary feasible solution,  $\mathbf{C}_i^*(\mathbf{x}_t^*)$ , for  $i = 1, \dots, N_C$ . The surrogate model for the  $i^{th}$  coupling variable is then refined using  $\mathbf{X}_{sur}^{t+1} = \mathbf{X}_{sur}^t \cup \mathbf{x}_t^*$  and the updated dataset  $\{\mathbf{X}_{sur}^{t+1}; \mathbf{C}_i^*(\mathbf{X}_{sur}^{t+1})\}$ . No additional high-fidelity simulations are required during the selection of  $\mathbf{x}_t^*$  with the adaptive sampling strategy. In the next cycle,  $\mathbf{X}_{rem}^t$  realizations are propagated through the layers of surrogates and high-fidelity models as described above to find the realizations for which the error in Equation 1 is not acceptable,  $\mathbf{X}_{rem}^{t+1}$ . This proceeds iteratively till  $\mathbf{X}_{rem}^{t+1} = \emptyset$ .

The major advantage of this method is that the accepted solutions correspond to those that would be obtained from a high-fidelity simulation (to within the tolerance) without having performed FPI for all  $\mathbf{X}$  realizations. FPI is performed only for the initial  $\mathbf{X}_{sur}^0$  realizations and the selected  $\mathbf{x}_t^*$  realizations in each cycle through the adaptive sampling process. As the results will show, this can lead to substantial computational savings without compromising accuracy. Another advantage of the method is that, although the high-fidelity coupled disciplinary analyses for FPI cannot be decoupled, the remainder of the high-fidelity disciplinary analyses can be decoupled as shown in the flowchart (Figure 2). This would be favorable when

communication between the disciplines is difficult due to their development by different working groups or the analyses running on separate platforms.

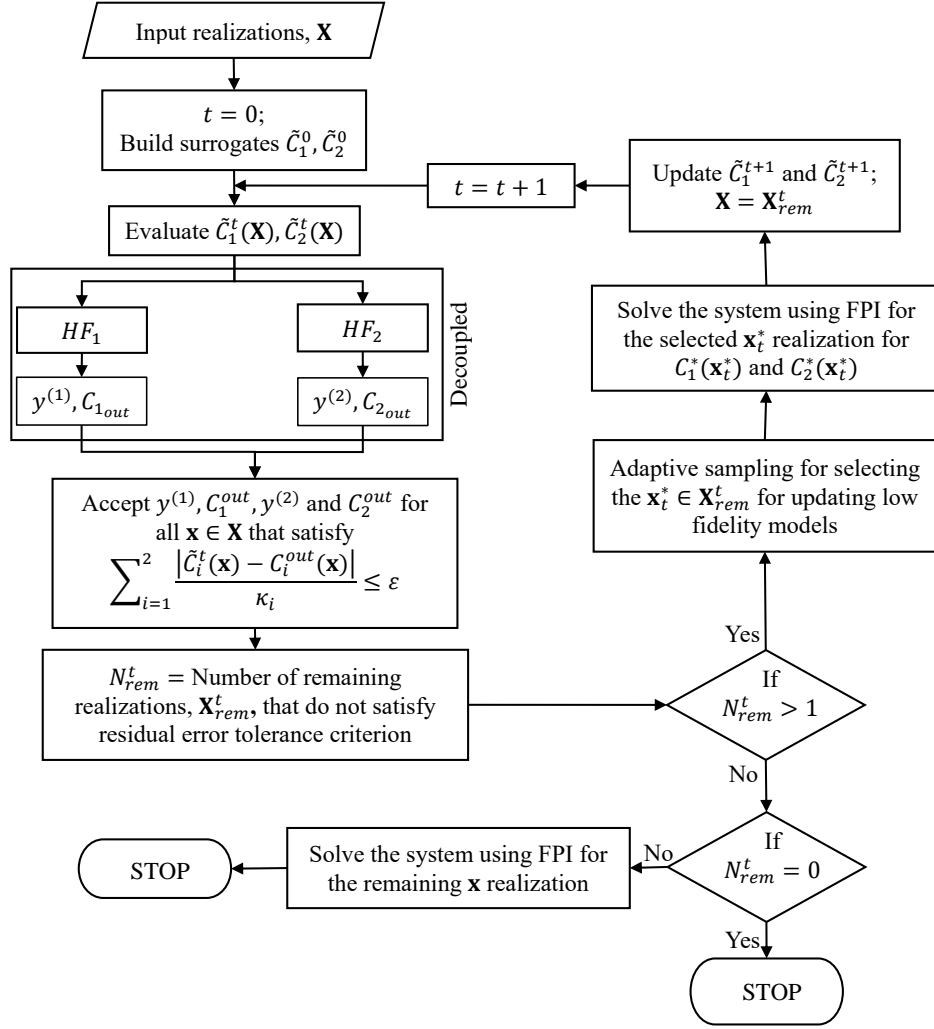


Figure 2. Multifidelity coupled uncertainty propagation method ( $HF_1$  and  $HF_2$  denotes high-fidelity analysis of disciplines 1 and 2, respectively, and  $t$  denotes the current cycle).

## B. Adaptive Sampling Strategies

Three different adaptive sampling strategies for selecting the  $\mathbf{x}_t^*$  realization, which is used to update the surrogate models, are explored for the multifidelity coupled uncertainty propagation method.

### 1. Maximum Residual Error

In this sampling strategy, the  $\mathbf{x}_t^* \in \mathbf{X}_{rem}^t$  with maximum normalized residual relative error is selected as the next sample. The optimization problem for the maximum residual error adaptive sampling strategy for cycle  $t$  is given by

$$\mathbf{x}_t^* = \underset{\mathbf{x} \in \mathbf{X}_{rem}^t}{\operatorname{argmax}} \sum_{i=1}^{N_C} \frac{|\tilde{C}_i^t(\mathbf{x}) - C_i^{out}(\mathbf{x})|}{\kappa_i}. \quad (2)$$

## 2. Maximum Information Gain

Using information gain as the sampling strategy enables us to use the surrogate prediction as well as the surrogate prediction standard deviation to make a decision on where to sample next. Note that both of these quantities are available from the Kriging (or a Gaussian process) fit that is used in this work. The Kriging surrogate prediction at any location is defined by a Gaussian distribution with mean prediction,  $\mu$ , and prediction standard deviation,  $\sigma$ . This adaptive sampling strategy chooses the next sampling location such that there will be maximum information gain in the surrogate prediction at the locations corresponding to the realizations in  $\mathbf{X}_{rem}^t$ .

We denote the present surrogate predicted Gaussian distribution as  $G_P \sim \mathcal{N}(\mu_P, \sigma_P^2)$  and a possible future surrogate predicted Gaussian distribution as  $G_F \sim \mathcal{N}(\mu_F, \sigma_F^2)$ . The Kullback-Liebler divergence between  $G_P$  and  $G_F$  at any realization  $\mathbf{x}$ , given the data  $d = \{\mathbf{X}_{sur}^t; \mathbf{C}_i^*(\mathbf{X}_{sur}^t)\}$  and possible future data  $\{\mathbf{x}_F; \tilde{\mathbf{C}}_i^F\}$  for the  $i^{th}$  coupling variable, is defined by

$$D_{KL}(G_P(\mathbf{x}|d) \parallel G_F(\mathbf{x}|d, \mathbf{x}_F, \tilde{\mathbf{C}}_i^F)) = \log \left( \frac{\sigma_F(\mathbf{x}|d, \mathbf{x}_F, \tilde{\mathbf{C}}_i^F)}{\sigma_P(\mathbf{x}|d)} \right) + \frac{\sigma_P^2(\mathbf{x}|d) + (\mu_P(\mathbf{x}|d) - \mu_F(\mathbf{x}|d, \mathbf{x}_F, \tilde{\mathbf{C}}_i^F))^2}{2\sigma_F^2(\mathbf{x}|d, \mathbf{x}_F, \tilde{\mathbf{C}}_i^F)} - \frac{1}{2}. \quad (3)$$

$D_{KL}$  is used as the information gain metric.  $\mu_P$  and  $\sigma_P$  are predictions from the present surrogate fit, and  $\mu_F$  and  $\sigma_F$  are predictions from a possible future surrogate fit. The total information gain at any  $\mathbf{x}$  realization can be calculated by integrating over all possible values of  $\tilde{\mathbf{C}}_i^F$  (defined by  $G_P(\mathbf{x}_F|d)$ ) for  $\mathbf{x}_F \in \mathbf{X}_{rem}^t$ . Gauss-Hermite quadrature is used for the integration and  $\tilde{\mathbf{C}}_{i,k}^F$  is the  $k^{th}$  Gauss-Hermite quadrature points (in this work, 5 Gauss-Hermite quadrature points are used) for the  $i^{th}$  coupling variable. The dataset  $d$  combined with the possible future data  $\{\mathbf{x}_F; \tilde{\mathbf{C}}_{i,k}^F\}$ ,  $d \cup \{\mathbf{x}_F; \tilde{\mathbf{C}}_{i,k}^F\}$ , is used to obtain the future surrogate fit  $G_F$  for the  $i^{th}$  coupling variable. This leads to  $k$  possible future surrogate fits for any  $\mathbf{x}_F$ . Then the total information gain at any  $\mathbf{x}$  realization for  $i^{th}$  coupling variable,  $D_i(\mathbf{x}|\mathbf{x}_F)$ , is calculated by integrating through Gauss-Hermite quadrature using the weights,  $w_k$ , for the  $k^{th}$  Gauss-Hermite quadrature point, as defined by

$$D_i(\mathbf{x}|\mathbf{x}_F) = \sum_{k=1}^5 w_k D_{KL}(G_P(\mathbf{x}|d) \parallel G_F(\mathbf{x}|d, \mathbf{x}_F, \tilde{\mathbf{C}}_{i,k}^F)). \quad (4)$$

The optimization problem for finding  $\mathbf{x}_t^*$  through the information gain based adaptive sampling criterion combines the information gain at all  $\mathbf{x}_{rem,j} \in \mathbf{X}_{rem}^t$  for all  $N_C$  coupling variables, as defined by

$$\mathbf{x}_t^* = \underset{\mathbf{x}_F \in \mathbf{X}_{rem}^t}{\operatorname{argmax}} \sum_{i=1}^{N_C} \sum_{j=1}^{N_{rem}^t} D_i(\mathbf{x}_{rem,j}|\mathbf{x}_F). \quad (5)$$

The information gain based strategy does not scale well if we increase the number of samples. As seen in Equation 5, a combinatorial problem needs to be solved to select the  $\mathbf{x}_t^*$  that maximizes the information gain. In order to tackle this scaling issue, we used a continuous optimization algorithm (Differential Evolution<sup>27</sup>). In order to ensure  $\mathbf{x}_F \in \mathbf{X}_{rem}^t$ , the  $\mathbf{x}_F$  input from the optimizer was rounded-off to the nearest neighbor in  $\mathbf{X}_{rem}^t$  while evaluating the objective function.

## 3. Maximum Weighted Information Gain

The weighted information gain criterion uses the present normalized residual error at  $\mathbf{x}_F \in \mathbf{X}_{rem}^t$  as the weight for the information gain. The weight at  $\mathbf{x}_F$  for the current cycle  $t$  is defined by

$$W_F^t(\mathbf{x}_F) = \frac{e^t(\mathbf{x}_F)}{\sum_{m=1}^{N_{rem}^t} e^t(\mathbf{x}_{rem,m})}, \quad (6)$$

where  $\mathbf{x}_{rem,m} \in \mathbf{X}_{rem}^t$  and  $e^t(\mathbf{x})$  is given by Equation 1. This attaches more importance to the information gain from sampling the input realizations that have higher associated residual errors. The optimization problem for the weighted information gain based adaptive sampling strategy is given by

$$\mathbf{x}_t^* = \underset{\mathbf{x}_F \in \mathbf{X}_{rem}^t}{\operatorname{argmax}} W_F^t(\mathbf{x}_F) \sum_{i=1}^{N_C} \sum_{j=1}^{N_{rem}^t} D_i(\mathbf{x}_{rem,j}|\mathbf{x}_F). \quad (7)$$

The maximization problem in Equation 7 is also solved using a continuous optimization problem definition with nearest neighbor round-off in the objective function as discussed for the maximum information gain criterion.

### III. Numerical Experiments

In this section, we run numerical experiments on two test problems for the three adaptive sampling strategies for the multifidelity coupled uncertainty propagation method. We compare the developed method to Monte Carlo simulation with FPI for every realization.

#### A. Test Problems

Test problem 1 is the Sellar's problem with three input random variables and two coupling variables as seen in Figure 3. The input random variables,  $x_1, x_2$  and  $x_3$ , are assumed to be normally distributed with a mean of 0, 5 and 5, and standard deviation of 3, 1.5 and 1.5, respectively. The normalization constants used for this problem are  $\kappa_1 = 8.4215$  and  $\kappa_2 = 7.8927$ , where these values correspond to the multidisciplinary feasible solution for the coupling variables at the mean of the input random variables. The residual error tolerance,  $\varepsilon$ , is set as 0.01. The number of iterations required for FPI convergence for test problem 1 is  $\sim 3$  on average. Thus, FPI convergence rate is fairly quick for test problem 1.

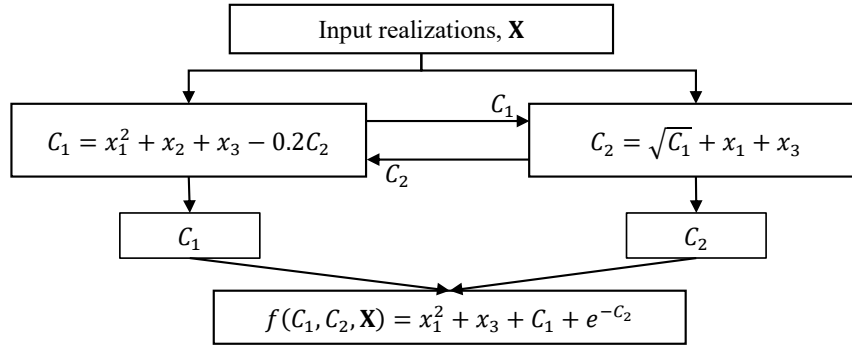


Figure 3. Test problem 1 (Sellar's problem)

Test problem 2 has five input random variables and two coupling variables as seen in Figure 4. All the input random variables are assumed to be normally distributed with a mean of 1 and standard deviation of 0.1. This problem uses  $\kappa_1 = 6.6259$  and  $\kappa_2 = 7.5370$ , where these values correspond to the multidisciplinary feasible solution for the coupling variables at the mean of the input random variables.  $\varepsilon$  is set as  $1.5\text{e-}4$ . The number of iterations required for FPI convergence for test problem 2 is  $\sim 18$  on average. Thus, it has slower FPI convergence rate than test problem 1 and conducting Monte Carlo simulation with the embedded FPI is expensive.

#### B. Results

We assess the efficiency of the multifidelity coupled uncertainty propagation method for  $N_{total}$  values of  $10^3$  and  $10^4$ , and an  $N_{sur}$  value of 20, for both problems. Tables 1 and 2 present the results for the comparison of different adaptive sampling strategies and Monte Carlo simulation with FPI for test problems 1 and 2, respectively.

The results for test problem 1 (Table 1) indicate that for  $N_{total} = 10^3$  and  $10^4$ , the total number of high-fidelity simulations required (considering both the disciplines) for each of the adaptive sampling strategies is worse than just using Monte Carlo simulation with FPI. This is because the rapid convergence of FPI for this problem ( $\sim 3$  iterations) makes it difficult to beat Monte Carlo simulation with FPI in terms of computational efficiency. The weighted information gain adaptive sampling strategy performs the best out of the three adaptive sampling strategies. For a problem such as this, where the FPI convergence is rapid, the standard Monte Carlo approach is most likely the best option.

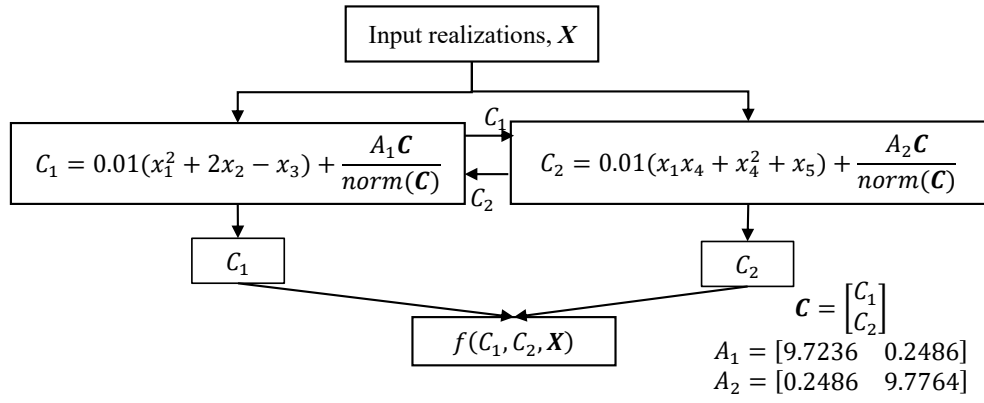


Figure 4. Test problem 2

The results for test problem 2 (Table 2) indicate that for  $N_{total} = 10^3$ , the total number of high-fidelity simulations required (considering both the disciplines) for all the adaptive sampling strategies outperforms using standard Monte Carlo simulation with FPI by a factor of  $\sim 1.8 - 2.1$ . The weighted information gain strategy performs the best, followed by the residual error and the information gain based adaptive sampling strategy. The gains in computational efficiency for test problem 2 are larger than for test problem 1 because in this case FPI convergence takes longer ( $\sim 18$  iterations for a FPI in test problem 2 as compared to  $\sim 3$  for test problem 1). For  $N_{total} = 10^4$ , all adaptive sampling strategies show a reduction in the number of high-fidelity simulations as compared to Monte Carlo simulation with FPI by a factor of  $\sim 2.2 - 3.4$ . The residual error sampling strategy performs the best, followed by the weighted information gain and the information gain based sampling strategies.

Table 1. Comparison of efficiency of different methods for test problem 1

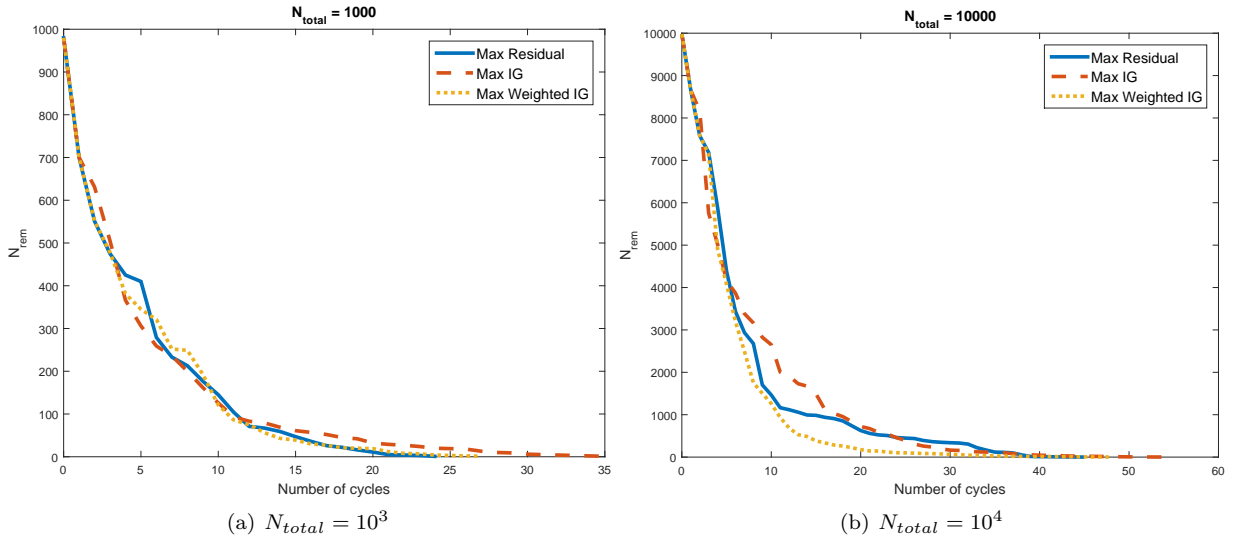
$N_{total}$	Sampling strategy	Initial $N_{rem}^0$	Total number of cases for FPI	Number of high-fidelity simulations
$10^3$	Monte Carlo Simulation with FPI	–	$10^3$	5,934
	Max Residual Error	700	44	10,346
	Max Information gain	700	55	10,818
	Max Weighted Information gain	700	47	10,286
$10^4$	Monte Carlo Simulation with FPI	–	$10^4$	58,938
	Max Residual Error	8,654	64	142,550
	Max Information gain	8,654	74	152,632
	Max Weighted Information gain	8,654	68	116,736

The convergence of the algorithm in terms of the number of remaining realizations,  $N_{rem}^t$ , that did not satisfy the residual error tolerance after each cycle is shown in Figures 5 and 6 for test problems 1 and 2, respectively (where “IG” stands for information gain). The residual error adaptive sampling strategy is generally the fastest converging, i.e., it requires the least number of FPI executions. Only for test problem 2 with  $N_{total} = 10^3$ , does the weighted information gain strategy perform better. We successfully decreased the number of realizations where FPI was employed in the multifidelity coupled uncertainty propagation method ( $\ll N_{total}$ ) as seen in Tables 1 and 2 for all the cases.

Figures 7 to 12 show the progress of the algorithm through a slice of the design space (input random variables 1 and 2 for both the test problems). The figures show the accepted realizations and sampling locations according to the choice of adaptive sampling strategy for selective cycles for both the test problems for  $N_{total} = 10^3$ . “Initial samples” refers to  $\mathbf{X}$ , “Accepted samples” refer to the set of all the realizations for which residual error tolerance criterion was satisfied after that cycle, “Adaptive sampling location” refers to

**Table 2. Comparison of efficiency of different methods for test problem 2**

$N_{total}$	Sampling strategy	Initial $N_{rem}^0$	Total number of cases for FPI	Number of high-fidelity simulations
$10^3$	Monte Carlo Simulation with FPI	—	$10^3$	35,984
	Max Residual Error	952	61	18,648
	Max Information gain	952	73	20,320
	Max Weighted Information gain	952	47	16,814
$10^4$	Monte Carlo Simulation with FPI	—	$10^4$	353,608
	Max Residual Error	7,808	67	103,562
	Max Information gain	7,808	93	158,534
	Max Weighted Information gain	7,808	76	121,098



**Figure 5. Number of realizations of  $X$  that do not satisfy the error tolerance after each cycle for test problem 1.**

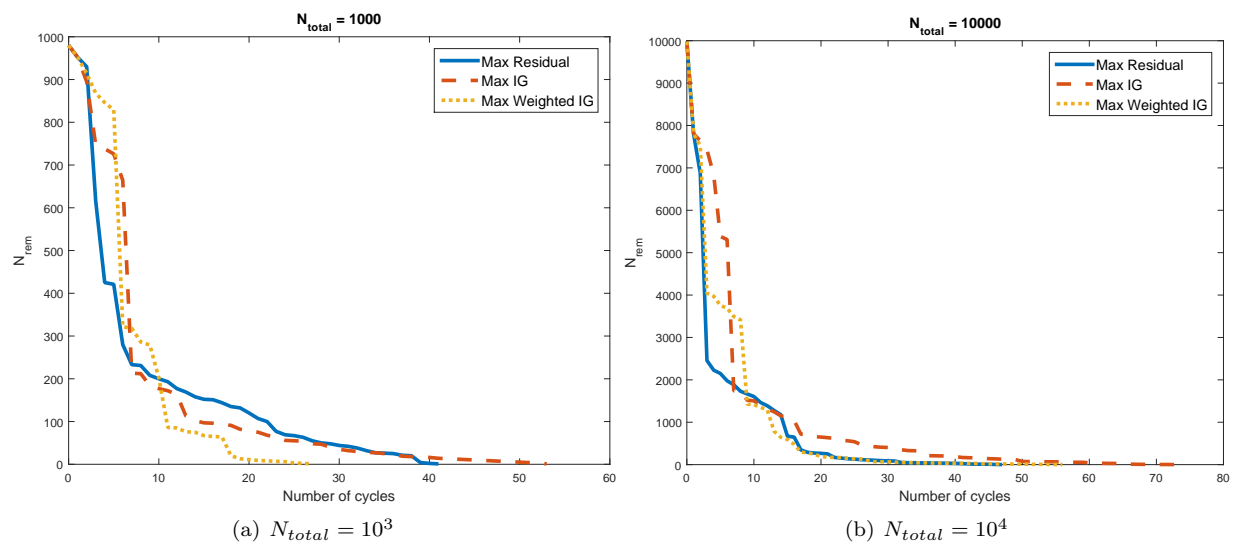
$\mathbf{x}_t^*$ , and “FPI locations” refers to all the realizations where FPI was used in order to get the true values of coupling variables to build or refine the surrogates.

## IV. Concluding Remarks

This paper has developed a new multifidelity coupled uncertainty propagation method for feedback-coupled multidisciplinary black-box systems. Instead of using fixed point iteration with Monte Carlo simulation for uncertainty propagation, the proposed method works to reduce the number of realizations for which FPI is employed. An essential feature of this method is that it maintains the same level of accuracy in the results as the original coupled high-fidelity system. Another advantage of the proposed method is the ability to partially decouple the process, which is helpful when communication between different disciplines is cumbersome.

The multifidelity coupled uncertainty propagation method uses surrogates to approximate the coupling variables and iteratively refines these surrogates using adaptive sampling strategies. We introduced the use of information gain as an adaptive sampling strategy for uncertainty propagation in coupled multidisciplinary systems. A residual error based adaptive sampling strategy is also explored in this work. The information gain weighted with the residual errors exhibits the best computational efficiency in most of the cases tested





**Figure 6.** Number of realizations of  $\mathbf{X}$  that do not satisfy the error tolerance after each cycle for test problem 2.

in this work.

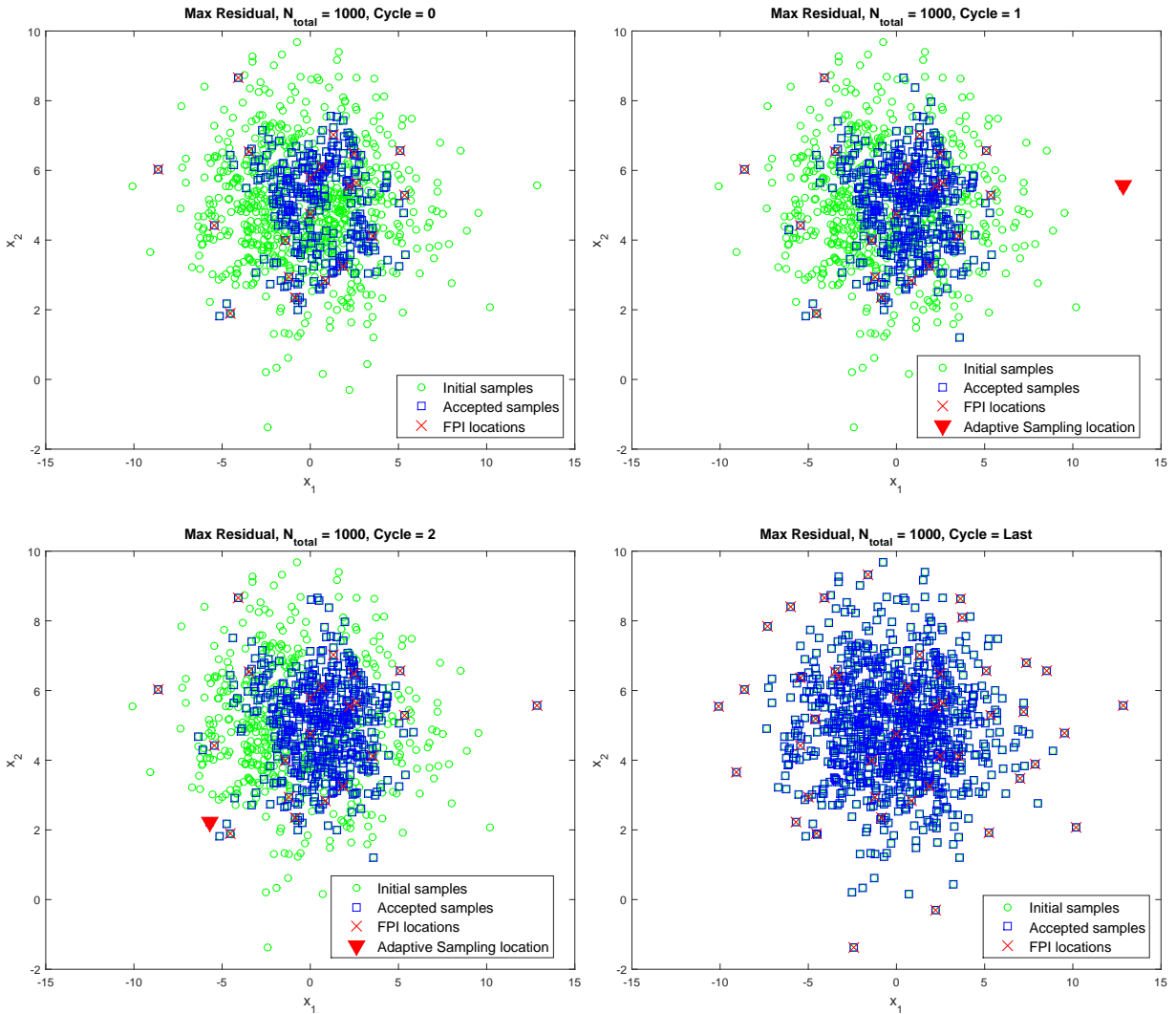
Ongoing work is modifying the approach to screen the realizations based on relative changes in surrogate predictions from one cycle to the next, and only execute a high-fidelity disciplinary analysis for those realizations where the change is beyond a certain threshold. This will potentially help reduce the total number of high-fidelity evaluations for the multifidelity coupled uncertainty propagation method and improve the efficiency gains.

## Acknowledgement

This work has been supported in part by the Air Force Office of Scientific Research (AFOSR) MURI on managing multiple information sources of multi-physics systems, Program Office Jean-Luc Cambier, Award Number FA9550-15-1-0038.

## References

- <sup>1</sup>Sobieszcanski-Sobieski, J., “Optimization by decomposition: a step from hierarchic to non-hierarchic systems,” *Recent Advances in Multidisciplinary Analysis and Optimization*, 1988, pp. 51.
- <sup>2</sup>Sobieszcanski-Sobieski, J., *Multidisciplinary design optimization: an emerging new engineering discipline*, Springer, 1995.
- <sup>3</sup>Sobieszcanski-Sobieski, J. and Haftka, R. T., “Multidisciplinary aerospace design optimization: survey of recent developments,” *Structural optimization*, Vol. 14, No. 1, 1997, pp. 1–23.
- <sup>4</sup>Alexandrov, N. M. and Hussaini, M. Y., *Multidisciplinary design optimization: State of the art*, Vol. 80, SIAM, 1997.
- <sup>5</sup>Alexandrov, N. M. and Lewis, R. M., “Algorithmic perspectives on problem formulations in MDO,” *Proceedings of the 8th AIAA/USAF/NASA/ISSMO Symposium on MAO, Long Beach, CA*, 2000.
- <sup>6</sup>Barthelemy, J.-F. and Haftka, R. T., “Approximation concepts for optimum structural design review,” *Structural and Multidisciplinary Optimization*, Vol. 5, No. 3, 1993, pp. 129–144.
- <sup>7</sup>Haftka, R., Sobieszcanski-Sobieski, J., and Padula, S., “On options for interdisciplinary analysis and design optimization,” *Structural optimization*, Vol. 4, No. 2, 1992, pp. 65–74.
- <sup>8</sup>Martins, J. R. and Lambe, A. B., “Multidisciplinary design optimization: a survey of architectures,” *AIAA journal*, Vol. 51, No. 9, 2013, pp. 2049–2075.
- <sup>9</sup>Viana, F. A., Simpson, T. W., Balabanov, V., and Toropov, V., “Metamodeling in multidisciplinary design optimization: how far have we really come?” *AIAA Journal*, Vol. 52, No. 4, 2014, pp. 670–690.
- <sup>10</sup>Yao, W., Chen, X., Luo, W., van Tooren, M., and Guo, J., “Review of uncertainty-based multidisciplinary design optimization methods for aerospace vehicles,” *Progress in Aerospace Sciences*, Vol. 47, No. 6, 2011, pp. 450–479.
- <sup>11</sup>Martin, J. D. and Simpson, T. W., “A methodology to manage system-level uncertainty during conceptual design,” *Journal of Mechanical Design*, Vol. 128, No. 4, 2006, pp. 959–968.



**Figure 7. A slice of the design space showing the spread of accepted and remaining  $X$  realizations with the maximum residual error adaptive sampling location for selective cycles for test problem 1**

<sup>12</sup>Amaral, S., Allaire, D., and Willcox, K., “A decomposition-based approach to uncertainty analysis of feed-forward multicomponent systems,” *International Journal for Numerical Methods in Engineering*, Vol. 100, No. 13, 2014, pp. 982–1005.

<sup>13</sup>Cramer, E. J., Dennis, Jr, J., Frank, P. D., Lewis, R. M., and Shubin, G. R., “Problem formulation for multidisciplinary optimization,” *SIAM Journal on Optimization*, Vol. 4, No. 4, 1994, pp. 754–776.

<sup>14</sup>Balling, R. J. and Sobieszcanski-Sobieski, J., “Optimization of coupled systems-A critical overview of approaches,” *AIAA journal*, Vol. 34, No. 1, 1996, pp. 6–17.

<sup>15</sup>Alexandrov, N. M. and Kodiyalam, S., “Initial results of an MDO method evaluation study,” *Proceedings of the 7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, 1998.

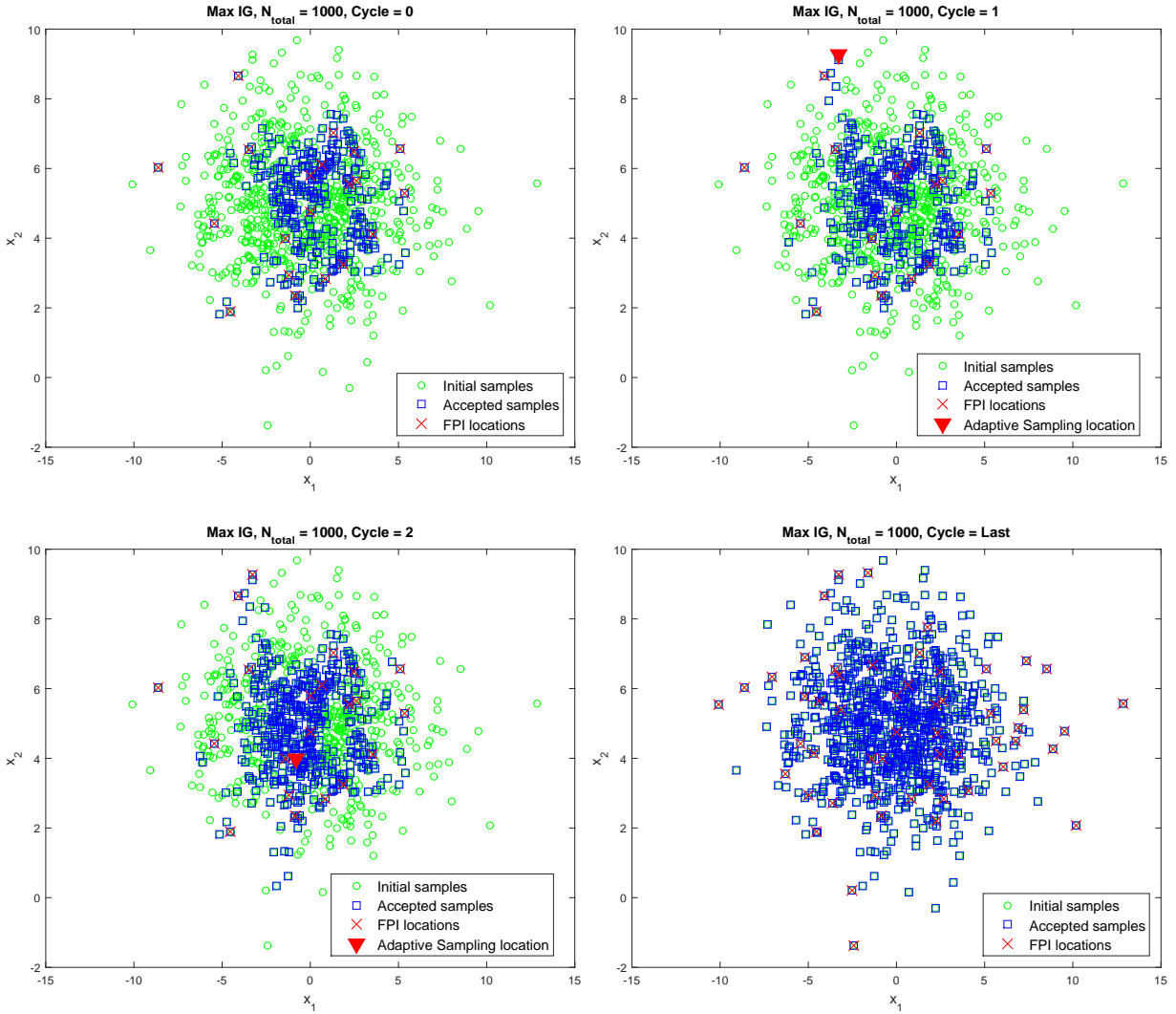
<sup>16</sup>Du, X. and Chen, W., “Collaborative reliability analysis under the framework of multidisciplinary systems design,” *Optimization and Engineering*, Vol. 6, No. 1, 2005, pp. 63–84.

<sup>17</sup>Mahadevan, S. and Smith, N., “Efficient first-order reliability analysis of multidisciplinary systems,” *International Journal of Reliability and Safety*, Vol. 1, No. 1-2, 2006, pp. 137–154.

<sup>18</sup>Sankararaman, S. and Mahadevan, S., “Likelihood-based approach to multidisciplinary analysis under uncertainty,” *Journal of Mechanical Design*, Vol. 134, No. 3, 2012, pp. 031008.

<sup>19</sup>Arnst, M., Ghanem, R., Phipps, E., and Red-Horse, J., “Reduced chaos expansions with random coefficients in reduced-dimensional stochastic modeling of coupled problems,” *International Journal for Numerical Methods in Engineering*, Vol. 97, No. 5, 2014, pp. 352–376.

<sup>20</sup>Villemonteix, J., Vazquez, E., and Walter, E., “An informational approach to the global optimization of expensive-to-evaluate functions,” *Journal of Global Optimization*, Vol. 44, No. 4, 2009, pp. 509–534.



**Figure 8. A slice of the design space showing the spread of accepted and remaining X realizations with the maximum information gain adaptive sampling location for selective cycles for test problem 1**

<sup>21</sup>Hennig, P. and Schuler, C. J., “Entropy search for information-efficient global optimization,” *The Journal of Machine Learning Research*, Vol. 13, No. 1, 2012, pp. 1809–1837.

<sup>22</sup>Hernández-Lobato, J. M., Hoffman, M. W., and Ghahramani, Z., “Predictive entropy search for efficient global optimization of black-box functions,” *Advances in Neural Information Processing Systems*, 2014, pp. 918–926.

<sup>23</sup>Huan, X. and Marzouk, Y. M., “Simulation-based optimal Bayesian experimental design for nonlinear systems,” *Journal of Computational Physics*, Vol. 232, No. 1, 2013, pp. 288–317.

<sup>24</sup>Villanueva, D. and Smarslok, B. P., “Using Expected Information Gain to Design Aerothermal Model Calibration Experiments,” *17th AIAA Non-Deterministic Approaches Conference, Kissimmee, FL, USA*, 2015.

<sup>25</sup>Simpson, T. W., Mauery, T. M., Korte, J. J., and Mistree, F., “Kriging models for global approximation in simulation-based multidisciplinary design optimization,” *AIAA journal*, Vol. 39, No. 12, 2001, pp. 2233–2241.

<sup>26</sup>Rasmussen, C. E., “Gaussian processes for machine learning,” MIT Press, 2006.

<sup>27</sup>Price, K., Storn, R. M., and Lampinen, J. A., *Differential evolution: a practical approach to global optimization*, Springer Science & Business Media, 2006.

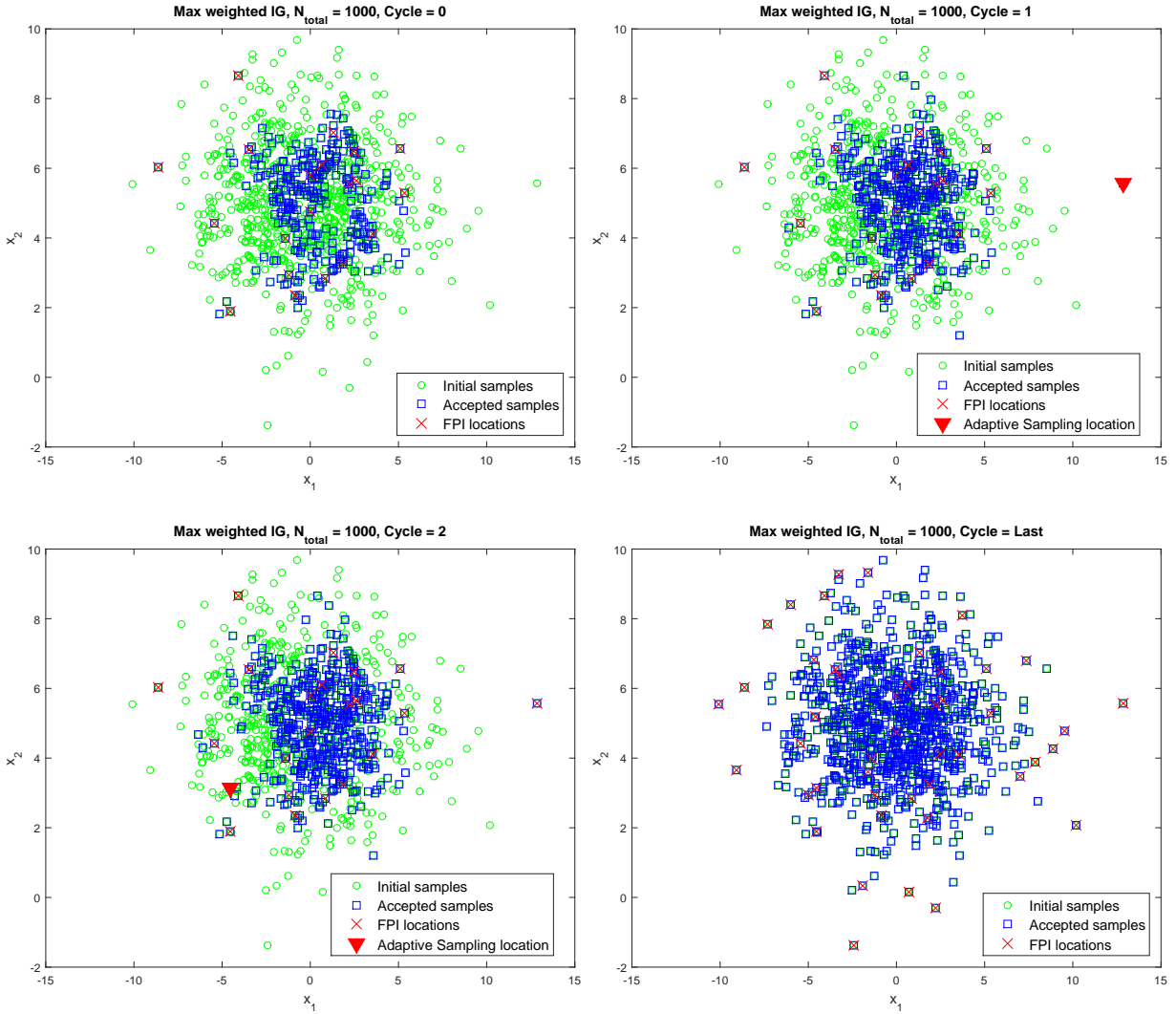


Figure 9. A slice of the design space showing the spread of accepted and remaining X realizations with the maximum weighted information gain adaptive sampling location for selective cycles for test problem 1

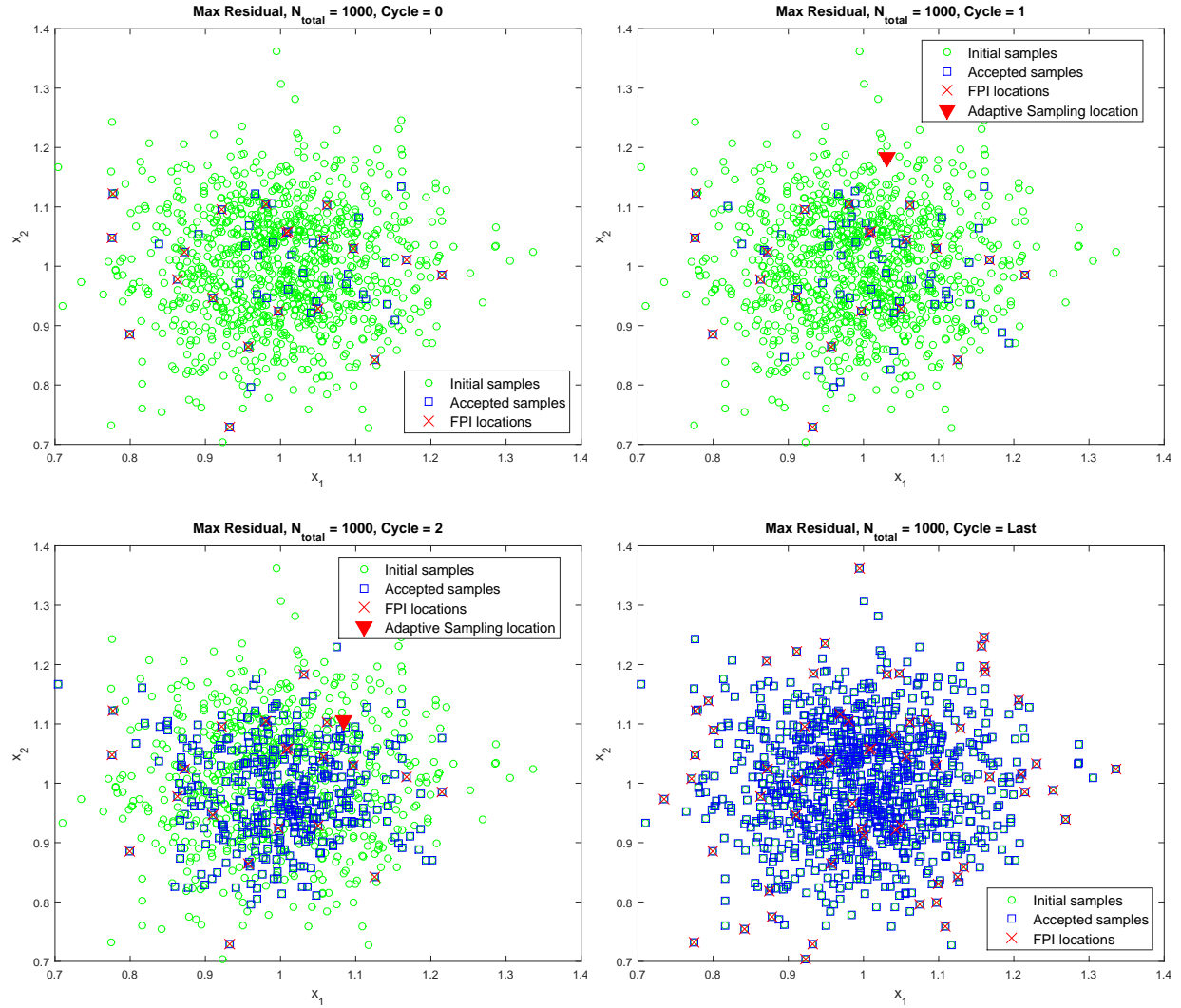


Figure 10. A slice of the design space showing the spread of accepted and remaining X realizations with the maximum residual error adaptive sampling location for selective cycles for test problem 2

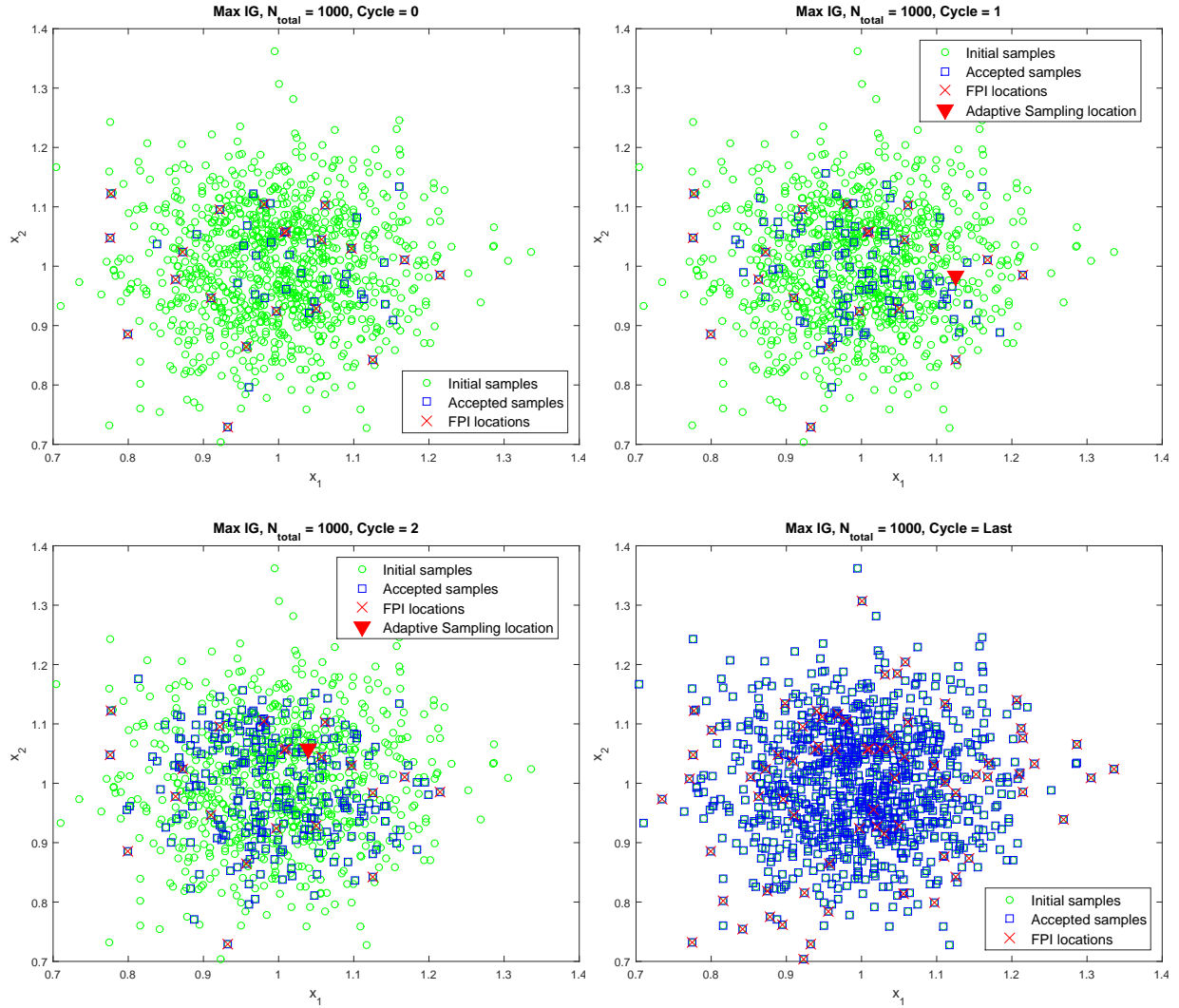


Figure 11. A slice of the design space showing the spread of accepted and remaining X realizations with the maximum information gain adaptive sampling location for selective cycles for test problem 2



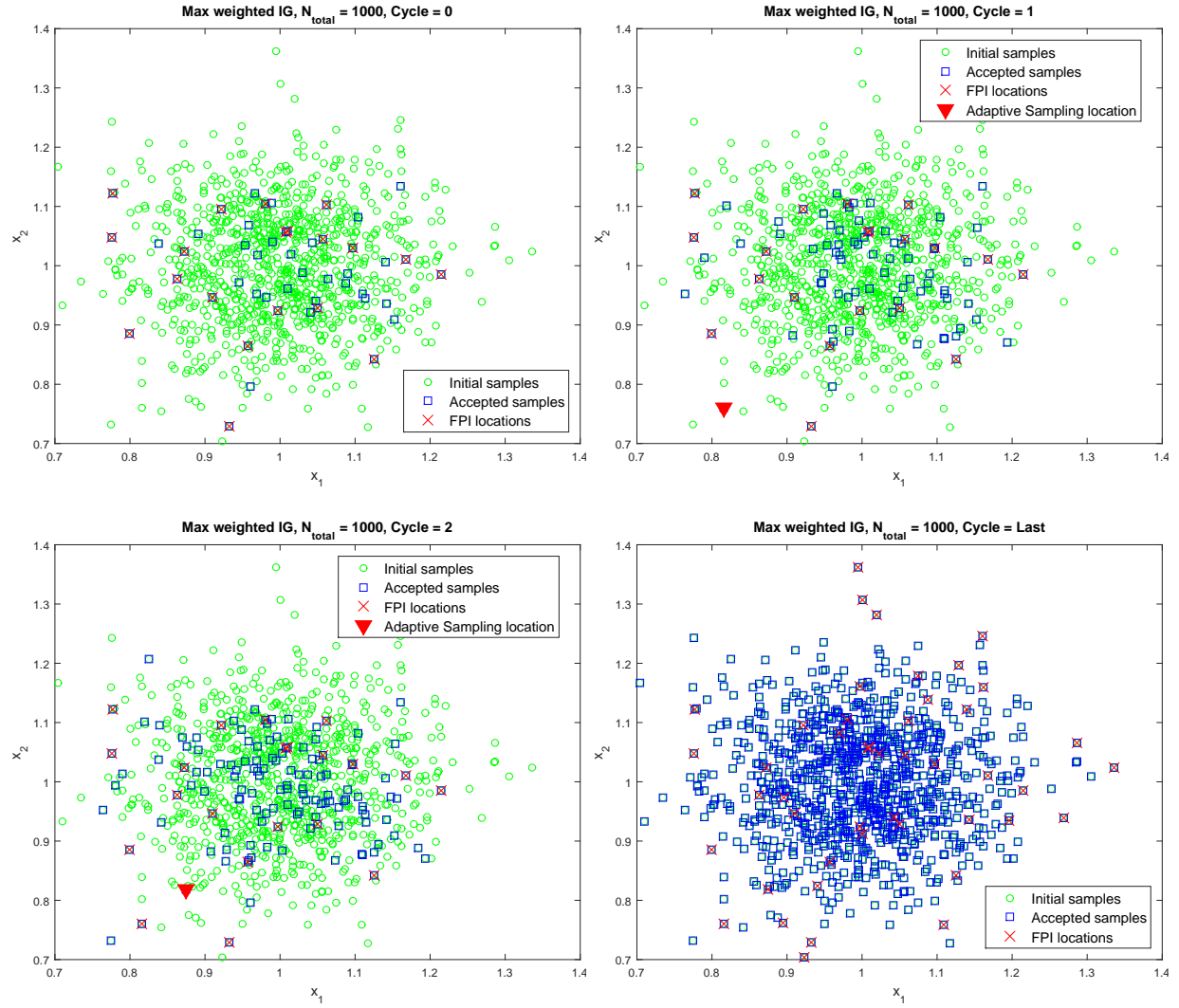


Figure 12. A slice of the design space showing the spread of accepted and remaining X realizations with the maximum weighted information gain adaptive sampling location for selective cycles for test problem 2