

# Reduction of Epistemic Model Uncertainty in Simulation-Based Multidisciplinary Design

**Zhen Jiang<sup>1</sup>**

Department of Mechanical Engineering,  
Northwestern University,  
Evanston, IL 60208  
e-mail: zhenjiang2015@u.northwestern.edu

**Shishi Chen<sup>1</sup>**

School of Aerospace Engineering;  
Key Laboratory of Dynamics  
and Control of Flight Vehicle,  
Ministry of Education of P.R. China,  
Beijing Institute of Technology,  
Beijing 100081, China  
e-mail: shishi.chen@northwestern.edu

**Daniel W. Apley**

Department of Industrial Engineering  
and Management Sciences,  
Northwestern University,  
Evanston, IL 60208  
e-mail: dapley@northwestern.edu

**Wei Chen<sup>2</sup>**

Department of Mechanical Engineering,  
Northwestern University,  
Evanston, IL 60208  
e-mail: weichen@northwestern.edu

*Model uncertainty is a significant source of epistemic uncertainty that affects the prediction of a multidisciplinary system. In order to achieve a reliable design, it is critical to ensure that the disciplinary/subsystem simulation models are trustworthy, so that the aggregated uncertainty of system quantities of interest (QOIs) is acceptable. Reduction of model uncertainty can be achieved by gathering additional experiments and simulations data; however, resource allocation for multidisciplinary design optimization (MDO) and analysis remains a challenging task due to the complex structure of the system, which involves decision makings about where (sampling locations), what (disciplinary responses), and which type (simulations versus experiments) for allocating more resources. Instead of trying to concurrently make the above decisions, which would be generally intractable, we develop a novel approach in this paper to break the decision making into a sequential procedure. First, a multidisciplinary uncertainty analysis (MUA) is developed to identify the input settings with unacceptable amounts of uncertainty with respect to the system QOIs. Next, a multidisciplinary statistical sensitivity analysis (MSSA) is developed to investigate the relative contributions of (functional) disciplinary responses to the uncertainty of system QOIs. The input settings and critical responses to allocate resources are selected based on the results from MUA and MSSA, with the aid of a new correlation analysis derived from spatial-random-process (SRP) modeling concepts, ensuring the sparsity of the selected inputs. Finally, an enhanced pre-posterior analysis predicts the effectiveness of allocating experimental and/or computational resource to answer the question about which type of resource to allocate. The proposed method is applied to a benchmark electronic packaging problem to demonstrate how epistemic model uncertainty is gradually reduced via resource allocation for data gathering. [DOI: 10.1115/1.4033918]*

**Keywords:** multidisciplinary design optimization, resource allocation, epistemic uncertainty, multidisciplinary uncertainty analysis, multidisciplinary statistical sensitivity analysis

## 1 Introduction

With the increasing accuracy and complexity of simulation models, the cost of running high-fidelity computer simulations becomes more and more prohibitive in simulation-based design. Accordingly, emulators (also known as metamodels or surrogate models) have been widely used as a replacement of original simulation models and are especially useful in searching for the optimal design. However, due to the lack of simulation/experimental data, the accuracy of emulators can be very poor, which introduces *epistemic* model uncertainty that, in contrast to the *aleatory* uncertainty from natural/physical randomness, can be reduced by collecting additional data. We refer to the process of planning additional experiments/simulations for reducing epistemic model uncertainty as a *resource allocation* problem. Resource allocation is even more important in MDO, which, unlike traditional design optimization problems, is notoriously complicated due to the fact that it requires analyses in multiple disciplines and/or involves a number of subsystems and components. In particular, resource allocation for MDO requires an optimal scheme to distribute resources to different disciplines and to enhance the prediction of system QOIs. However, this is not an easy problem, because the

coupling relationship between different disciplines leads to a nested structure and substantial complexity.

Model uncertainty, as a major source of epistemic uncertainty in multidisciplinary systems, stems from either having a limited knowledge of the true reality and/or using a simplified model that does not satisfactorily represents the reality (represented as “model discrepancy” [1]), or from a lack of computer simulation data (termed “code uncertainty” [1] or “interpolation uncertainty” [2]) when replacing the expensive computer model using a surrogate model (metamodel). Even though the sources of these two types of uncertainty are different, they all belong to epistemic model uncertainty due to the lack of data—model discrepancy uncertainty due to the lack of experimental data and interpolation uncertainty due to the lack of computer simulation data. Both sources of epistemic uncertainty can be addressed simultaneously using the Bayesian Gaussian process approach presented by Kennedy and O’Hagan [1]. The epistemic model uncertainty of system QOIs is an aggregation of the uncertainties in the disciplinary subsystems, and the purpose of resource allocation is to reduce the aggregated prediction uncertainty of system QOIs. Nevertheless, it is not straightforward to provide a good resource allocation scheme, due to the amount of decisions involved and their dynamic nature. The decisions to make include:

- (1) *Where* in the input space of a multidisciplinary system should we allocate additional resources? (That is, what are the input locations where system QOIs are most influenced by epistemic model uncertainty?)

<sup>1</sup>Z. Jiang and S. Chen contributed equally to this work.

<sup>2</sup>Corresponding author.

Contributed by the Design Automation Committee of ASME for publication in the JOURNAL OF MECHANICAL DESIGN. Manuscript received October 8, 2015; final manuscript received June 5, 2016; published online June 30, 2016. Assoc. Editor: Nam H. Kim.

- (2) To *what* disciplinary response(s) shall we allocate more resources?
- (3) *Which* type of resources shall we allocate? Is adding more simulations sufficient, or should we conduct more physical experiments?

The first question requires uncertainty quantification of disciplinary subsystems and an efficient way to propagate their impact to the system QOIs. Although there has been extensive research on uncertainty propagation of a single discipline [1,3–8], or a multidisciplinary system but with only aleatory uncertainty (e.g., using Monte Carlo simulation [9–11], Taylor series expansion [12–15], decomposition strategies [16–22], etc.), existing work for multidisciplinary systems with epistemic model uncertainty is scarce. The problem is often oversimplified by either predefining model bias [23–27] or assuming a simple form of interdisciplinary couplings (restricted to feed-forward couplings [28], for example). Furthermore, even after developing a convenient uncertainty propagation scheme, one must still decide where to choose the input sample locations so that they sufficiently cover all regions with large epistemic model uncertainty. Consequently, the second question requires an approach to distinguish the relative contributions (which can vary over the input space) of the different disciplinary simulation models on the uncertainty of system QOIs. Statistical sensitivity analysis (SSA) is commonly used in engineering design to gain knowledge of complex model behavior and to facilitate designers' decision making on where to spend engineering effort [29]. However, SSA of model uncertainty, which is different from SSA of random variables in the more traditional literature, has rarely been studied [30–32]. The difficulty lies in the fact that uncertain models are essentially functional inputs rather than scalars, and hence, their SSA would be fundamentally different than traditional routines. Allaire et al. [30] developed a pure sampling-based approach, whose applicability is practically limited to single-disciplinary systems and hierarchical systems without strong couplings due to its computational cost. SSA for a fully coupled multidisciplinary system was never studied by other researchers. The third question is perhaps the least studied of the three. Previous works focused on allocating a single type of resource, typically experimental (e.g., Ref. [28]). Indeed, relative to conducting simulations, conducting physical experiments may be a more direct and effective way to learn physical reality, but it is also typically more expensive. For that reason, there exists a tradeoff between most reducing uncertainty (by conducting experiments) versus least increasing computational budget (by conducting simulations).

On top of these three individually challenging questions, the complexity of concurrently answering them adds another layer of difficulty. To simplify the problem, the pioneering work by Sankararaman et al. [28] had to assume no coupling relationship between disciplines, and that the model uncertainty is purely due to some unknown model parameters, which makes SSA much easier.

To manage the complexity in the decision-making process, this paper presents a new strategy and mathematical framework for resource allocation in a multidisciplinary system, one that breaks the decision making into a sequential procedure. An MUA method [33,34] we developed earlier is applied to make the first decision (i.e., *where*). A multidisciplinary sensitivity analysis (MSSA) [31,32], which we developed to handle uncertain *functional responses* as the stochastic inputs in the sensitivity analysis (the stochastic inputs are scalar variables in the traditional sensitivity analysis literature), is applied to investigate the relative contributions of disciplinary responses to the aggregated uncertainty of system QOIs for the second decision (i.e., *what*). A type of correlation analysis derived from SRP modeling concepts is also developed as an aid to ensure the sparsity of the selected input settings. Finally, an enhanced preposterior analysis predicts the effectiveness of allocating experimental or computational resource, or a combination of both, to make the last decision (i.e.,

*which*). The remainder of the paper is organized as follows. Section 2 reviews the SRP-based model bias correction approach to quantify epistemic model uncertainty and its extension in multidisciplinary systems. Section 3 details the proposed framework for resource allocation, which is applied in Sec. 4 to a benchmark electronic packaging problem to illustrate how the uncertainty of system QOIs is reduced over the input space via a few iterations of resource allocation. Concluding remarks are provided in Sec. 5.

For clarity of the discussions in this paper, we make the following remarks:

- (1) The objective of this paper is to improve the *global* modeling capability of a multidisciplinary system, such that the epistemic model uncertainty of system QOIs is acceptable over the input space. A globally accurate model can be useful in many design activities other than direct optimization; for example, the simulation/emulation results of a model can be passed into a classification analysis for designers to understand the potentially feasible/infeasible input settings and to be more focused on the “regions of interest.” This line of research [35,36] is different from the objective-oriented sampling [37–39] which specifically targets a *local* region where potential optimal design solution lies.
- (2) When referring to *resources*, we mean *experiments* and *simulations*. We use “experiments” in a broad context: it can be real physical experiments, or extremely high-fidelity simulations that (almost) can represent the reality. “Simulations” in this paper generally refer to results from physics-based models that are less accurate than experiments.
- (3) *Emulators* are nonphysics-based models that are built upon data from experiments and simulations. They are much less costly than direct simulations. In this work, an SRP modeling technique is used to create disciplinary emulators.
- (4) Experiments and simulations are conducted for individual disciplinary responses, but not for the system QOIs (which we consider practically unaffordable). Any full systemwise analysis is conducted using emulators.
- (5) The framework proposed in this paper is targeted at a multidisciplinary problem, although it can be simplified to work for a single discipline as well.

## 2 Review of Model Bias Correction in Multidisciplinary Systems

Model uncertainty is a significant type of epistemic uncertainty in a multidisciplinary system. To quantify model uncertainty, data from the computer simulations are first collected and then compared with the experimental data. Based on the differences in the response values over the two datasets, the original simulation model is “updated” to a more accurate emulator; meanwhile, the model uncertainty is quantified. We refer to this process as *model bias correction*. Subsequently, an extra experimental dataset is provided to validate the emulator. Model bias correction is conducted iteratively by refining the model and/or collecting additional data, until the emulator is trustworthy. In this section, we briefly review the SRP-based model bias correction technique and its application in multidisciplinary systems.

For a functional response  $y(\mathbf{x})$  that depends on a  $p$ -dimensional set of input variables  $\mathbf{x} = (x_1, \dots, x_p)$ , the following formulation [1,3–8] presents the relationship between the simulations and the experiments:

$$y^e(\mathbf{x}) = y^m(\mathbf{x}) + \delta(\mathbf{x}) + \varepsilon \quad (1)$$

where the superscript “*e*” denotes the experimental response of  $y(\mathbf{x})$ , and the superscript “*m*” denotes the simulation model of the same response.  $\delta(\mathbf{x})$  is the discrepancy function that represents the

model bias, and  $\varepsilon \sim \mathcal{N}(0, \lambda)$  is a random error accounting for the experimental variability, assumed to be normally distributed with unknown variance  $\lambda$ . The situation where  $\varepsilon$  does not have zero mean (i.e., a so-called experimental bias or “systematic error”) would be more complicated and is beyond the scope of this paper.

Suppose that a set of  $M$  simulation response observations  $\{\mathbf{x}_i^m, y^m(\mathbf{x}_i^m) : i = 1, \dots, M\}$  and a set of  $N$  experimental response observations  $\{\mathbf{x}_i^e, y^e(\mathbf{x}_i^e) : i = 1, \dots, N\}$  have been collected. Since  $\mathbf{x}_i^m$  ( $i = 1, \dots, M$ ) and  $\mathbf{x}_i^e$  ( $i = 1, \dots, N$ ) are not necessarily the same input settings, and likewise,  $y^m(\mathbf{x}_i^m)$  and  $y^e(\mathbf{x}_i^e)$  may not be directly compared to obtain the discrepancy function, an SRP modeling technique is adopted to bridge the gap and enable the comparison. An SRP can be viewed as a collection of random variables distributed over the input space that collectively have certain multivariate spatial correlation structure. Specifically, Gaussian process (GP) modeling, which assumes the random variables to follow a multivariate normal distribution, has emerged as one of the most popular metamodeling approaches due to its ability to capture functional nonlinearity, perfectly interpolate deterministic surfaces, etc. [40,41]. In this study, we adopt GP modeling and model the simulation and the discrepancy function as GPs such that

$$\begin{aligned} y^m(\mathbf{x}) &\sim \mathcal{GP}(\mathbf{h}^m(\mathbf{x})^T \boldsymbol{\beta}^m, V^m(\mathbf{x}, \mathbf{x}')) \\ V^m(\mathbf{x}, \mathbf{x}') &= \sigma_m^2 \exp \left\{ -\sum_{k=1}^p \omega_k^m (x_k - x'_k)^2 \right\} \\ \delta(\mathbf{x}) &\sim \mathcal{GP}(\mathbf{h}^\delta(\mathbf{x})^T \boldsymbol{\beta}^\delta, V^\delta(\mathbf{x}, \mathbf{x}')) \\ V^\delta(\mathbf{x}, \mathbf{x}') &= \sigma_\delta^2 \exp \left\{ -\sum_{k=1}^p \omega_k^\delta (x_k - x'_k)^2 \right\} \end{aligned} \quad (2)$$

In the preceding,  $\mathbf{h}^m(\mathbf{x})$  and  $\mathbf{h}^\delta(\mathbf{x})$  denote vector-valued functions whose elements are some user-specified regression basis functions (e.g., constant, linear, quadratic, etc.), and  $\boldsymbol{\beta}^m$  and  $\boldsymbol{\beta}^\delta$  are two vectors of regression coefficients associated with  $\mathbf{h}^m(\mathbf{x})$  and  $\mathbf{h}^\delta(\mathbf{x})$ , respectively. Their products,  $\mathbf{h}^m(\mathbf{x})^T \boldsymbol{\beta}^m$  and  $\mathbf{h}^\delta(\mathbf{x})^T \boldsymbol{\beta}^\delta$ , construct the prior mean functions of  $y^m(\mathbf{x})$  and  $\delta(\mathbf{x})$ .  $\sigma_m$  and  $\sigma_\delta$  are the prior standard deviations, and  $\boldsymbol{\omega}^m = [\omega_1^m, \omega_2^m, \dots, \omega_p^m]^T$  and  $\boldsymbol{\omega}^\delta = [\omega_1^\delta, \omega_2^\delta, \dots, \omega_p^\delta]^T$  are the spatial correlation parameters used to capture the nonlinearity of the functions.  $V^m(\mathbf{x}, \mathbf{x}')$  and  $V^\delta(\mathbf{x}, \mathbf{x}')$ , which depend on  $\sigma_m$ ,  $\sigma_\delta$ ,  $\boldsymbol{\omega}^m$ , and  $\boldsymbol{\omega}^\delta$ , are the covariance functions of  $y^m(\mathbf{x})$  and  $\delta(\mathbf{x})$ , respectively.

Combining Eqs. (1) and (2) yields that the collected data follow a multivariate normal distribution

$$\mathbf{d} \sim \mathcal{N}(\mathbf{H}\boldsymbol{\beta}, \mathbf{V}_d) \quad (3)$$

where  $\mathbf{d} = \{y^m(\mathbf{x}_1^m), \dots, y^m(\mathbf{x}_M^m), y^e(\mathbf{x}_1^e), \dots, y^e(\mathbf{x}_N^e)\}^T$  is the collected response data from simulations and experiments, and  $\boldsymbol{\beta} = [(\boldsymbol{\beta}^m)^T, (\boldsymbol{\beta}^\delta)^T]^T$  is the collection of regression coefficients.  $\mathbf{H}$  is a matrix generated by  $\mathbf{h}^m(\cdot)$  and  $\mathbf{h}^\delta(\cdot)$ , and  $\mathbf{V}_d$  is generated by  $V^m(\cdot, \cdot)$  and  $V^\delta(\cdot, \cdot)$  (see Refs. [8] and [41] for full derivation and discussion).

Let  $\boldsymbol{\phi} = \{\boldsymbol{\beta}^m, \sigma_m, \boldsymbol{\omega}^m, \boldsymbol{\beta}^\delta, \sigma_\delta, \boldsymbol{\omega}^\delta, \lambda\}$  denote the collection of unknown parameters of the GP models, referred to as *hyperparameters*. The key to estimating the discrepancy function and subsequently enabling a better model prediction is to estimate these hyperparameters. They are usually estimated via *maximum likelihood estimation*, i.e., maximizing the likelihood function  $p(\boldsymbol{\phi}|\mathbf{d})$ . Subsequently, we can obtain a prediction of  $y^e(\mathbf{x})$  at any untested design of interest  $\mathbf{x}$ . The relationship between  $y^e(\mathbf{x})$  and its prediction can be written as

$$y^e(\mathbf{x}) = \hat{y}^e(\mathbf{x}) + Z(\mathbf{x}) \quad (4)$$

where  $\hat{y}^e(\mathbf{x})$  denotes the updated mean prediction that takes into account the discrepancy function  $\delta(\mathbf{x})$ .  $Z(\mathbf{x})$  is a zero-mean GP,

i.e.,  $Z(\mathbf{x})|\mathbf{x} \sim \mathcal{N}(0, \sigma_Z^2(\mathbf{x}))$ , which accounts for the uncertainty in the prediction.  $\sigma_Z^2(\mathbf{x})$  is also known as the mean square error of the mean prediction  $\hat{y}^e(\mathbf{x})$ . The merit of using GP models is that we are able to obtain analytical equations for calculating  $\hat{y}^e(\mathbf{x})$  and  $\sigma_Z^2(\mathbf{x})$ . For example,

$$\hat{y}^e(\mathbf{x}) = \mathbf{h}(\mathbf{x})\boldsymbol{\beta} + \mathbf{t}(\mathbf{x})\mathbf{V}_d^{-1}(\mathbf{d} - \mathbf{H}\boldsymbol{\beta}) \quad (5)$$

where

$$\begin{aligned} \mathbf{h}(\mathbf{x}) &= [\mathbf{h}^m(\mathbf{x}) \quad \mathbf{h}^\delta(\mathbf{x})] \\ \mathbf{t}(\mathbf{x}) &= [V^m(\mathbf{x}, \mathbf{x}^m) \quad V^m(\mathbf{x}, \mathbf{x}^e) + V^\delta(\mathbf{x}, \mathbf{x}^e)] \end{aligned} \quad (6)$$

Again, Refs. [8] and [41] provide full details of calculating  $\hat{y}^e(\mathbf{x})$  and  $\sigma_Z^2(\mathbf{x})$ .

Figure 1 illustrates a typical result from GP-based model bias correction, in which there is originally a significant model bias when comparing the simulation data (the triangles) and the experimental data (the circles), especially on the left and right boundaries of the  $x$  domain. Using GP modeling, we are able to obtain an updated model (the solid purple curve) that matches the experimental data, while maintaining the general functional trend learned from the simulation data. Besides, GP modeling enables the quantification of interpolation uncertainty at the locations that have not yet been simulated, represented by a prediction interval (PI) (the shaded region) calculated from  $\sigma_Z^2(\mathbf{x})$ .

GP-based model bias correction can be applied to multidisciplinary systems. A notional multidisciplinary system is depicted in Fig. 2. Suppose it involves a total of ND disciplines associated with a collection of individual disciplinary input variables  $\mathbf{x}_{\text{ind}} = \{\mathbf{x}_i : i = 1, \dots, \text{ND}\}$  and some input variables  $\mathbf{x}_s$  that are shared across at least two disciplines. The disciplines are coupled via linking variables  $\mathbf{u}_{ij}$  ( $i, j = 1, \dots, \text{ND}$ ), which serve as output from the  $i$ th discipline and input to the  $j$ th discipline. If both  $\mathbf{u}_{ij}$  and  $\mathbf{u}_{ji}$  are nonempty, the relation between the  $i$ th and  $j$ th disciplines is referred to as *feedback coupling*, and otherwise, *feed-forward coupling*. We denote the collection of all linking variables that are output from and input to the  $i$ th discipline as  $\mathbf{u}_i$  and  $\mathbf{u}_{-i}$ , respectively, i.e.,  $\mathbf{u}_i = \{\mathbf{u}_{ij} : j = 1, \dots, \text{ND}, j \neq i\}$  and  $\mathbf{u}_{-i} = \{\mathbf{u}_{ji} : j = 1, \dots, \text{ND}, j \neq i\}$ . The system QOIs  $\mathbf{y}_{\text{sys}}$  depend on individual disciplinary responses  $\mathbf{y}_{\text{ind}} = \{\mathbf{y}_i : i = 1, \dots, \text{ND}\}$  collected from the responses of ND disciplines, through the system analysis model.

Epistemic model uncertainty widely exists in multidisciplinary systems. For the  $i$ th discipline, the simulated response  $\mathbf{u}_i$  may have uncertainty, which will introduce uncertainty to  $\mathbf{u}_{-i}$  and further influence  $\mathbf{u}_i$  itself. Such “convolution” accumulates uncertainty in the disciplinary level and passes it on to the disciplinary

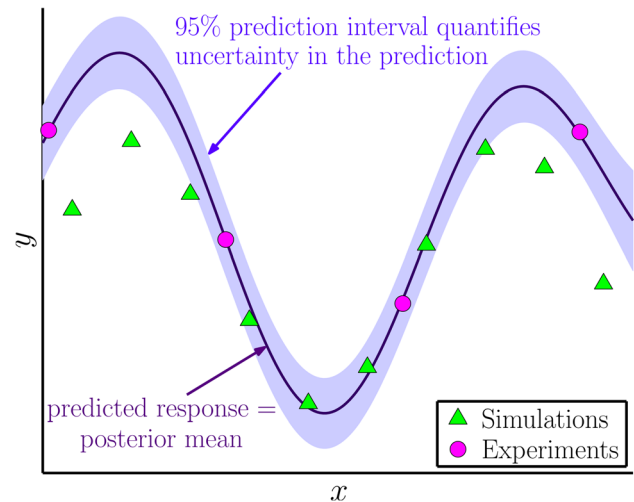
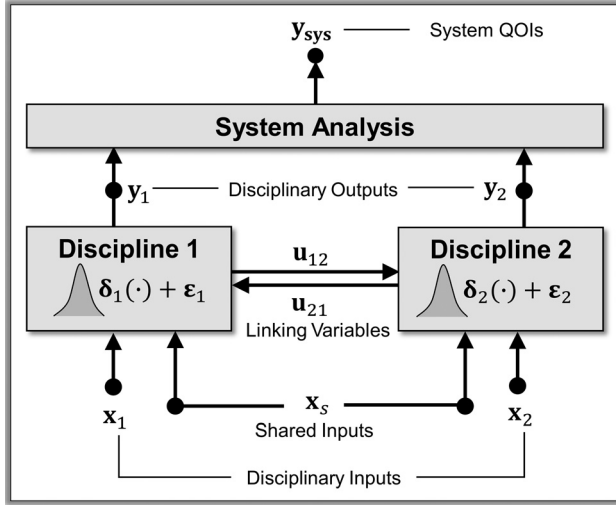


Fig. 1 Illustration of GP-based model bias correction





**Fig. 2 A notional multidisciplinary system with epistemic model uncertainty**

outputs  $y_i$ . Another complication is that the simulation model for  $y_i$  may also involve epistemic model uncertainty. That is, even if  $u_{.i}$  has no uncertainty, simulating the response  $y_i$  still could involve uncertainty. Based on Eq. (4), we present the predictions of all the linking variables and the disciplinary output variables, respectively, as (for  $i = 1, 2, \dots, ND$ )

$$u_i^e(\mathbf{x}_i, \mathbf{x}_s, \mathbf{u}_{.i}^e) = \hat{u}_i^e(\mathbf{x}_i, \mathbf{x}_s, \mathbf{u}_{.i}^e) + \mathbf{Z}_{ui}(\mathbf{x}_i, \mathbf{x}_s, \mathbf{u}_{.i}^e) \quad (7)$$

and

$$y_i^e(\mathbf{x}_i, \mathbf{x}_s, \mathbf{u}_{.i}^e) = \hat{y}_i^e(\mathbf{x}_i, \mathbf{x}_s, \mathbf{u}_{.i}^e) + \mathbf{Z}_{yi}(\mathbf{x}_i, \mathbf{x}_s, \mathbf{u}_{.i}^e) \quad (8)$$

Assessing the cumulative uncertainty of system QOIs calls for an uncertainty propagation approach. Our recent development of a GP-based MUA method [33,34] addresses the issues of efficient uncertainty propagation in multidisciplinary systems, which is generically suitable for situations with both aleatory uncertainty (from input variables  $\mathbf{x}_{ind}$  and  $\mathbf{x}_s$ ) and epistemic uncertainty (from models). A version of the method that only considers epistemic uncertainty (which is the main focus of this paper) will be provided in Sec. 3.1.

### 3 Resource Allocation for Uncertainty Reduction in Multidisciplinary Systems

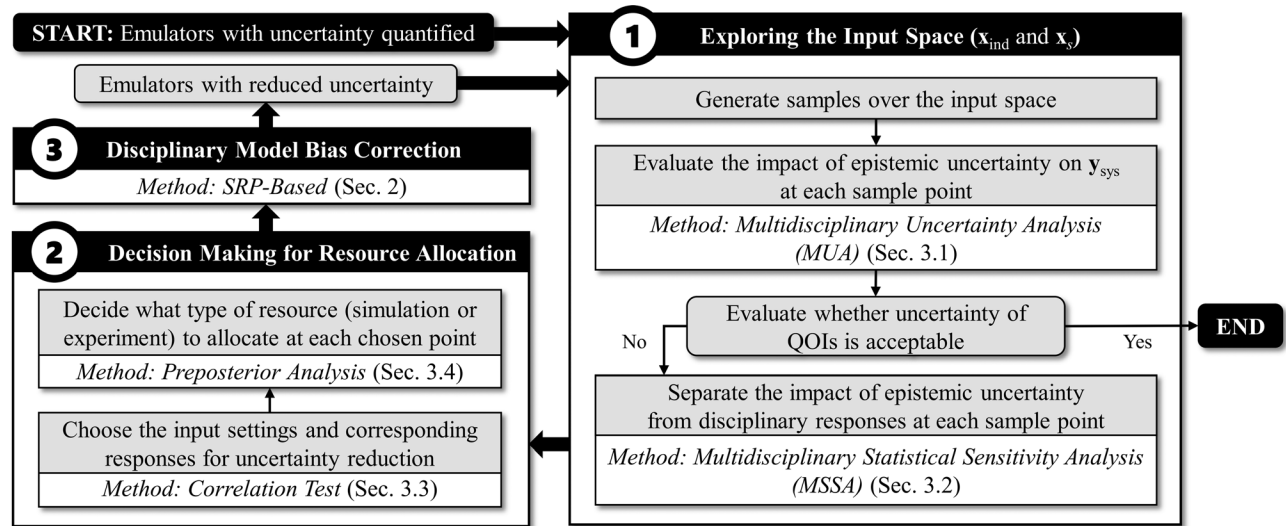
In this section, we present in detail the proposed resource allocation approach for reducing epistemic model uncertainty in a multidisciplinary system. An overview is provided in Fig. 3. In the preliminary step, model bias correction is conducted based on some existing data to construct emulators for disciplinary responses (linking the input variables  $u_{ij}$  and  $x_j$  to the disciplinary outputs  $y_i$ , for each  $i$ ). In step 1, a space-filling strategy is used to explore the input space  $\{\mathbf{x}_{ind}, \mathbf{x}_s\}$  by generating sufficient sample points and identifying the locations with unacceptable amounts of uncertainty with respect to the system QOIs  $y_{sys}$ . The MUA method is applied to assess the aggregated epistemic model uncertainty of  $y_{sys}$  at each sample point (Sec. 3.1). An MSSA then separates the impact of epistemic model uncertainty from different responses at each sample point (Sec. 3.2). Decisions about resource allocation (*where*, *what*, and *which*) are made in sequence in step 2, where we integrate the information from MUA, MSSA, and a preposterior analysis. Finally in step 3 as resources are allocated, another round of model bias correction is performed.

Note that it may require a few iterations of the procedure in Fig. 3 before the epistemic model uncertainty is reduced to a satisfactory level. Since the variance of system QOIs  $y_{sys}$  is an indicator of the amount of uncertainty, we specify the following stopping criterion in this paper:

$$\gamma(\mathbf{x}_{ind}, \mathbf{x}_s) \triangleq \frac{\sqrt{\text{Var}[y_{sys}(\mathbf{x}_{ind}, \mathbf{x}_s)]}}{\iint \|y_{sys}(\mathbf{x}_{ind}, \mathbf{x}_s)\| d\mathbf{x}_{ind} d\mathbf{x}_s / \iint d\mathbf{x}_{ind} d\mathbf{x}_s} \leq \alpha\%, \quad \text{for } \forall \mathbf{x}_{ind}, \mathbf{x}_s \quad (9)$$

The numerator is the standard deviation of  $y_{sys}$  (at any given input setting  $\{\mathbf{x}_{ind}, \mathbf{x}_s\}$ ), and the denominator is the mean of the absolute value of  $y_{sys}$  over the input space. Thus,  $\gamma$  is an intuitive measure of the prediction uncertainty that resembles the “coefficient of variation” in traditional statistics. The user can choose other variance-based stopping criteria, independent from the rest of procedure presented.

**3.1 Exploring the Input Space and Evaluating the Impact of Epistemic Model Uncertainty.** In step 1, to find where in the input space is most influenced by epistemic model uncertainty, a straightforward treatment is to apply a space-filling strategy (for



**Fig. 3 Flowchart of the proposed resource allocation approach**

example, Monte Carlo uniform sampling, Latin hypercube sampling (LHS), Halton sequence, etc.) and generate  $N_s$  samples over the input space:  $(\mathbf{x}_{\text{ind}}^{(k)}, \mathbf{x}_s^{(k)}) = (\mathbf{x}_1^{(k)}, \dots, \mathbf{x}_{ND}^{(k)}, \mathbf{x}_s^{(k)})$ ,  $k = 1, \dots, N_s$ . Next, for each sample point, the GP-based MUA method [33,34] is applied to evaluate the mean and variance of system QOIs  $\mathbf{y}_{\text{sys}}$ . The merit of the MUA method is that it utilizes the structure of GP emulators and enables an efficient analytical uncertainty analysis. The following discussion in this subsection is for every sample input that is generated, and hence for notational simplicity, we omit the superscript “(k).”

Given a specific input setting  $(\mathbf{x}_{\text{ind}}, \mathbf{x}_s)$ , a first-order Taylor expansion of  $\hat{\mathbf{u}}_i^e(\cdot, \cdot, \cdot)$  and  $\hat{\mathbf{y}}_i^e(\cdot, \cdot, \cdot)$  in Eqs. (7) and (8) is applied, which leads to (for  $i = 1, \dots, ND$ )

$$\mathbf{u}_i^e(\mathbf{x}_i, \mathbf{x}_s, \mathbf{u}_i^e) \approx \hat{\mathbf{u}}_i^e(\mathbf{x}_i, \mathbf{x}_s, \boldsymbol{\mu}_{u-i}) + \sum_{j=1, j \neq i}^{ND} \frac{\partial \hat{\mathbf{u}}_i^e}{\partial \mathbf{u}_j^e} (\mathbf{u}_j^e - \boldsymbol{\mu}_{uj}) + \mathbf{Z}_{ui}(\mathbf{x}_i, \mathbf{x}_s, \mathbf{u}_i^e) \quad (10)$$

$$\mathbf{y}_i^e(\mathbf{x}_i, \mathbf{x}_s, \mathbf{u}_i^e) \approx \hat{\mathbf{y}}_i^e(\mathbf{x}_i, \mathbf{x}_s, \boldsymbol{\mu}_{u-i}) + \sum_{j=1, j \neq i}^{ND} \frac{\partial \hat{\mathbf{y}}_i^e}{\partial \mathbf{u}_j^e} (\mathbf{u}_j^e - \boldsymbol{\mu}_{uj}) + \mathbf{Z}_{yi}(\mathbf{x}_i, \mathbf{x}_s, \mathbf{u}_i^e) \quad (11)$$

where  $\boldsymbol{\mu}_{ui}$ ,  $\boldsymbol{\mu}_{u-i}$ , and  $\boldsymbol{\mu}_{yi}$  denote the means of  $\mathbf{u}_i$ ,  $\mathbf{u}_{-i}$ , and  $\mathbf{y}_i$ , respectively. The derivatives on the right-hand side, while not explicitly written out, are taken at the mean of the variables, i.e.,  $(\mathbf{x}_{\text{ind}}, \mathbf{x}_s, \boldsymbol{\mu}_u)$ . Therefore, the means and covariance matrices of the linking variables (i.e., all  $\mathbf{u}_{ij}$ 's) and the disciplinary outputs (i.e., all  $\mathbf{y}_i$ 's) can be estimated by (for  $i = 1, \dots, ND$ )

$$\boldsymbol{\mu}_{ui} \approx \hat{\mathbf{u}}_i^e(\mathbf{x}_i, \mathbf{x}_s, \boldsymbol{\mu}_{u-i}), \quad \boldsymbol{\Sigma}_u \approx (\mathbf{A}^{-1}) \boldsymbol{\Sigma}_{Zu} (\mathbf{A}^{-1})^T \quad (12)$$

$$\boldsymbol{\mu}_{yi} \approx \hat{\mathbf{y}}_i^e(\mathbf{x}_i, \mathbf{x}_s, \boldsymbol{\mu}_{u-i}), \quad \boldsymbol{\Sigma}_y \approx (\mathbf{B} \mathbf{A}^{-1}) \boldsymbol{\Sigma}_{Zu} (\mathbf{B} \mathbf{A}^{-1})^T + \boldsymbol{\Sigma}_{Zy} \quad (13)$$

where  $\boldsymbol{\Sigma}_u$ ,  $\boldsymbol{\Sigma}_{Zu}$ ,  $\boldsymbol{\Sigma}_y$ , and  $\boldsymbol{\Sigma}_{Zy}$  denote the covariance matrices of  $\mathbf{u}^e$ ,  $\mathbf{Z}_u$ ,  $\mathbf{y}^e$ , and  $\mathbf{Z}_y$ , respectively. The matrices  $\mathbf{A}$  and  $\mathbf{B}$  are

$$\mathbf{A} = \begin{bmatrix} \mathbf{I} & -\frac{\partial \hat{\mathbf{u}}_1^e}{\partial \mathbf{u}_2^e} & \dots & -\frac{\partial \hat{\mathbf{u}}_1^e}{\partial \mathbf{u}_{ND}^e} \\ -\frac{\partial \hat{\mathbf{u}}_2^e}{\partial \mathbf{u}_1^e} & \mathbf{I} & \dots & -\frac{\partial \hat{\mathbf{u}}_2^e}{\partial \mathbf{u}_{ND}^e} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{\partial \hat{\mathbf{u}}_{ND}^e}{\partial \mathbf{u}_1^e} & -\frac{\partial \hat{\mathbf{u}}_{ND}^e}{\partial \mathbf{u}_2^e} & \dots & \mathbf{I} \end{bmatrix} \quad (14)$$

$$\mathbf{B} = \begin{bmatrix} 0 & \frac{\partial \hat{\mathbf{y}}_1^e}{\partial \mathbf{u}_2^e} & \dots & \frac{\partial \hat{\mathbf{y}}_1^e}{\partial \mathbf{u}_{ND}^e} \\ \frac{\partial \hat{\mathbf{y}}_2^e}{\partial \mathbf{u}_1^e} & 0 & \dots & \frac{\partial \hat{\mathbf{y}}_2^e}{\partial \mathbf{u}_{ND}^e} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \hat{\mathbf{y}}_{ND}^e}{\partial \mathbf{u}_1^e} & \frac{\partial \hat{\mathbf{y}}_{ND}^e}{\partial \mathbf{u}_2^e} & \dots & 0 \end{bmatrix} \quad (15)$$

Since GP-based model bias correction provides analytical formulas to calculate  $\hat{\mathbf{u}}_i^e(\cdot, \cdot, \cdot)$  and  $\hat{\mathbf{y}}_i^e(\cdot, \cdot, \cdot)$  (similar to Eq. (5)), the entries of matrices  $\mathbf{A}$  and  $\mathbf{B}$  also have analytical forms, which enables an efficient calculation.

Detailed derivations can be found in Refs. [33] and [34]. After deriving the means and covariance matrices of the disciplinary responses  $\mathbf{y}_{\text{ind}}$ , the derivation of the mean and variance of the system QOIs  $\mathbf{y}_{\text{sys}}$  can be treated as a single-disciplinary uncertainty

propagation problem and solved using any conventional method. From the variance of  $\mathbf{y}_{\text{sys}}$ , it can be easily discovered which regions have large uncertainty.

**3.2 Separating the Impact of Epistemic Model Uncertainty From Disciplinary Responses.** Apart from assessing the aggregated impact of epistemic model uncertainty on  $\mathbf{y}_{\text{sys}}$  over the  $N_s$  sample points, one can actually separate the total impact into the contributions from different responses ( $\mathbf{u}_{ij}$ 's and  $\mathbf{y}_i$ 's) using SSA. This information is particularly important for resource allocation in a multidisciplinary system, as it helps decision makers understand what the most important factors are that influence the system performance. Moreover, it will be shown in Sec. 3.3 that it also helps in choosing the most appropriate input settings that are most in need of additional resources.

The most widely used measure of variability in SSA is variance, because it is usually interpretable for practitioners. Building on the Sobol' method [42,43], which is a variance-based Monte Carlo method for aleatory input variability, we developed a more comprehensive MSSA in Refs. [31] and [32] that will assess the impact of aleatory and epistemic uncertainties. Different from the Sobol' method, MSSA has the capability of measuring the impact of *functional responses* that serve as stochastic inputs in the sensitivity indices (the stochastic inputs are scalar variables in the traditional Sobol' method, although the output is a functional response). Particularly, we use *local* sensitivity analysis in this subsection, which focuses on examining the impact of uncertainties from different sources at a single input setting.

In the Sobol' method, the variance of a random output variable  $Y$  can be decomposed into the contributions of a set of random input variables  $X_1, \dots, X_p$ . The *main sensitivity index* (MSI) and the *total sensitivity index* (TSI) of a particular input  $X_i$ , which measure the main effect and the overall effect of  $X_i$  on the output  $Y$ , respectively, are given by

$$\text{MSI}(X_i) = \frac{\text{Var}_{X_i}(\mathbb{E}_{\mathbf{X}_{\sim i}}(Y|X_i))}{\text{Var}(Y)} \quad (16)$$

$$\text{TSI}(X_i) = 1 - \frac{\text{Var}_{\mathbf{X}_{\sim i}}(\mathbb{E}_{X_i}(Y|\mathbf{X}_{\sim i}))}{\text{Var}(Y)} \quad (17)$$

where  $\mathbf{X}_{\sim i}$  denotes all the input variables except  $X_i$ . The subscript in “Var” and “ $\mathbb{E}$ ” denotes the source of uncertainty; for example,  $\text{Var}_{X_i}(\cdot)$  is the variance of the variable in the brackets due to the variation of  $X_i$ . Both MSI and TSI are between 0 and 1, with 0 indicating no effect and 1 the most dominant effect. The same idea can be extended to examining the relative contributions of epistemic model uncertainty. If we view any  $Z_i \in \{\mathbf{Z}_{ui}, \mathbf{Z}_{yi}; i = 1, \dots, ND\}$  (the random term that quantifies epistemic model uncertainty of a particular response in Eqs. (7) and (8)) in a similar way as  $X_i$  is treated in Eqs. (16) and (17), then the MSI and TSI of  $Z_i$  on the system QOIs  $\mathbf{y}_{\text{sys}}$  can be defined as

$$\text{MSI}(Z_i) = \frac{\text{Var}_{Z_i}(\mathbb{E}_{\mathbf{Z}_{\sim i}}(\mathbf{y}_{\text{sys}}|Z_i))}{\text{Var}(\mathbf{y}_{\text{sys}})} \quad (18)$$

$$\text{TSI}(Z_i) = 1 - \frac{\text{Var}_{\mathbf{Z}_{\sim i}}(\mathbb{E}_{Z_i}(\mathbf{y}_{\text{sys}}|\mathbf{Z}_{\sim i}))}{\text{Var}(\mathbf{y}_{\text{sys}})} \quad (19)$$

where  $\mathbf{Z}_{\sim i} = \{\mathbf{Z}_{ui}, \mathbf{Z}_{yi}; i = 1, \dots, ND\} \setminus Z_i$ . The Sobol' variance decomposition assumes that the uncertainty sources are uncorrelated, which is reasonable in our case because the epistemic uncertainty of the individual disciplinary models (i.e.,  $Z_i$ 's) is quantified separately and not correlated with each other.

It is worth noting that although Eqs. (18) and (19) resemble the standard Sobol' indices (Eqs. (16) and (17)), the calculation is much more complex due to the fact that the  $\mathbf{Z}$ 's are the stochastic “input variables” in our indices, but they are really stochastic

functional responses over the input variables (i.e., random processes rather than random variables). In standard Sobol' indices, the output is a functional response over the stochastic inputs, but the stochastic inputs are scalar variables and not functional responses themselves. For example, a sampling-based treatment to calculate the MSI of  $Z_i$  as defined in Eq. (18) is a trilevel procedure: in the outer level, one needs to generate multiple Monte Carlo realizations of  $Z_i$  (which are, by the nature, realizations of a function); in the intermediate level, for each realization of  $Z_i$ , one needs to generate realizations of  $\mathbf{Z}_{\sim i}$  (realizations of a set of functions); and finally in the inner level, given the realizations of all  $\mathbf{Z}$ 's, computations are needed to evaluate  $\mathbb{E}_{\mathbf{Z}_{\sim i}}(\mathbf{y}_{\text{sys}}|Z_i)$ . Therefore, the overall computation could be extremely expensive. Likewise for the process of calculating the TSI of  $Z_i$ .

The MUA method developed in Sec. 3.1 significantly speeds up the analysis. By applying Eqs. (12) and (13),  $\mathbb{E}_{\mathbf{Z}_{\sim i}}(\mathbf{y}_{\text{sys}}|Z_i)$  is directly obtained without having to sample  $\mathbf{Z}_{\sim i}$ , which eliminates the need of the outer level aforementioned. If necessary, one can further simplify the analysis and obtain rough approximations of MSI and TSI by linearizing  $\mathbb{E}_{\mathbf{Z}_{\sim i}}(\mathbf{y}_{\text{sys}}|Z_i)$  and  $\mathbb{E}_{Z_i}(\mathbf{y}_{\text{sys}}|\mathbf{Z}_{\sim i})$  as  $[\mathbf{y}_{\text{sys}}|\mathbf{Z}_{\sim i} \equiv 0, Z_i]$  and  $[\mathbf{y}_{\text{sys}}|Z_i \equiv 0, \mathbf{Z}_{\sim i}]$ , respectively, i.e., fixing the randomly varying  $\mathbf{Z}$  terms to their zero means, which then eliminates the need of the intermediate level as well. Such linearization would certainly lead to a less accurate sensitivity assessment, but it is very useful for a fast analysis in a large complex system.

**3.3 Choosing Input Settings and Corresponding Responses for Resource Allocation.** To decide where to allocate more resources, a simple approach is to choose a few samples (up to the

budget) with the largest variance of  $\mathbf{y}_{\text{sys}}$  among the  $N_s$  samples generated in Sec. 3.1. However, this is not the best approach, since multiple sample points may happen to fall inside the same region that has a large epistemic model uncertainty, especially when  $N_s$  is large. In such case, allocating resource to a single point inside this region should be sufficient for reducing the uncertainty. A good decision-making criterion is needed to detect whether two points with large uncertainty are "sufficiently far away" from each other, so that our final selection of samples would be widely spread over the input space without dense clustering. In this subsection, we propose a form of correlation analysis that utilizes information from the GP-based correlation, which measures the spatial proximity between two input settings, to select the best subset of input settings to allocate more resources.

Consider two sample points, denoted by  $(\mathbf{x}_{\text{ind}}^{(k)}, \mathbf{x}_s^{(k)})$  ( $k = 1, 2$ ) without loss of generality, from the  $N_s$  samples we generated in Sec. 3.1. We start by calculating their sensitivity indices and detecting the main contributing sources of uncertainty, respectively, at these two input locations. If their major sources of uncertainty are different, i.e., if  $\text{Var}[\mathbf{y}_{\text{sys}}(\mathbf{x}_{\text{ind}}^{(k)}, \mathbf{x}_s^{(k)})]$  ( $k = 1, 2$ ) are mostly induced by model uncertainty from two different responses, then it is concluded that  $(\mathbf{x}_{\text{ind}}^{(1)}, \mathbf{x}_s^{(1)})$  and  $(\mathbf{x}_{\text{ind}}^{(2)}, \mathbf{x}_s^{(2)})$  are "far away" from each other and reside in different regions. On the other hand, if  $\text{Var}[\mathbf{y}_{\text{sys}}(\mathbf{x}_{\text{ind}}^{(k)}, \mathbf{x}_s^{(k)})]$  ( $k = 1, 2$ ) turn out to be induced by uncertainty of the same response (for example, some response  $L(\mathbf{x}_L, \mathbf{x}_s, \mathbf{u}_L) \in \{\mathbf{u}_i, \mathbf{y}_i: i = 1, \dots, \text{ND}\}$  associated with model uncertainty term  $Z_i$ ), then the two points may or may not be far away. We need to proceed to evaluate their correlation coefficient in the GP emulator of  $L$ , defined as

$$\rho_{1,2} = \frac{V^{mL}(\left(\mathbf{x}_{\text{ind}}^{(1)}, \mathbf{x}_s^{(1)}, \boldsymbol{\mu}_{u-L}^{(1)}\right), \left(\mathbf{x}_{\text{ind}}^{(2)}, \mathbf{x}_s^{(2)}, \boldsymbol{\mu}_{u-L}^{(2)}\right)) + V^{\delta L}(\left(\mathbf{x}_{\text{ind}}^{(1)}, \mathbf{x}_s^{(1)}, \boldsymbol{\mu}_{u-L}^{(1)}\right), \left(\mathbf{x}_{\text{ind}}^{(2)}, \mathbf{x}_s^{(2)}, \boldsymbol{\mu}_{u-L}^{(2)}\right))}{\sigma_{mL}^2 + \sigma_{\delta L}^2 + \lambda_L} \quad (20)$$

In the denominator,  $\sigma_{mL}^2$ ,  $\sigma_{\delta L}^2$ , and  $\lambda_L$  are the prior variances of the simulation model, the discrepancy function, and the experimental variance of response  $L$ , respectively; their sum, according to Eq. (2), is the prior variance of experimental response  $L^e$  at  $(\mathbf{x}_{\text{ind}}^{(k)}, \mathbf{x}_s^{(k)})$  ( $k = 1, 2$ ). In the numerator,  $V^{mL}(\cdot, \cdot)$  and  $V^{\delta L}(\cdot, \cdot)$  are the covariance functions of the simulation model and the discrepancy function of  $L$ , respectively, which can be directly calculated from Eq. (2); their sum is the prior covariance of the two sample points.  $\boldsymbol{\mu}_{u-L}^{(k)}$  ( $k = 1, 2$ ) are the estimated values of linking variables (as inputs to the response  $L$ ). Therefore,  $\rho_{1,2}$  is the GP-based correlation coefficient, which is an indicator of how much impact the two input settings have over each other; in other words, how likely it is that they would yield a similar value of response and/or share a similar amount of uncertainty. It can be seen that this type of correlation coefficient is easy to compute and yet comprehensively combines the information of the absolute spatial distance between two sample points, the roughness of the response (quantified by  $\omega^m$  and  $\omega^\delta$  in Eq. (2)), and the general trend of the response (quantified by  $\beta^m$ ,  $\beta^\delta$ ,  $\sigma_m$ , and  $\sigma_\delta$  in Eq. (2)). If  $\rho_{1,2}$  is close to 1 (or larger than a user-defined limit  $\beta$ ), the two sample points are considered to be in the same local region, and we conclude that allocating resource to one of them should be sufficient.

The Appendix gives a pseudocode for choosing the proper sample points for resource allocation. Furthermore, the developed method allows us to decide at each chosen sample point what response to allocate more resource, by picking the response  $L$  that mostly influences  $\mathbf{y}_{\text{sys}}$  at this point.

**3.4 Deciding the Type of Resource to Allocate for Each Chosen Sample Point.** Finally, the one remaining task is to decide which type of resource, simulation or experiment, to allocate for each chosen input setting and the corresponding chosen response at this input setting. To address this, in this subsection, we propose an enhanced *preposterior analysis* approach (see Fig. 4, based upon a similar idea in the earlier work [44,45]), which, prior to allocating the resources, can *predict the posterior variance* of  $\mathbf{y}_{\text{sys}}$  if the resources are actually gathered. Different from the existing work in the literature that focuses only on a single type of resource (usually experimental), the preposterior analysis proposed in this section provides insight for any combination of simulations and experiments.

Following the analyses in Sec. 3.3, suppose we have already made the decision to allocate more resources to response  $L$  at input locations  $(\mathbf{x}_{\text{ind}}^{(k)}, \mathbf{x}_s^{(k)}, \boldsymbol{\mu}_{u-L}^{(k)})$  ( $k = 1, \dots, N_L$ ), and to response  $L'$  at input locations  $(\mathbf{x}_{\text{ind}}^{(k)}, \mathbf{x}_s^{(k)}, \boldsymbol{\mu}_{u-L'}^{(k)})$  ( $k = N_L + 1, \dots, N_L + N_{L'}$ ), etc. The preposterior analysis first generates hypothetical simulation and/or experimental data and then evaluates the influence on the posterior variance of the system QOIs. The detailed procedure is as follows:

- (1) Suggest a resource allocation plan and generate hypothetical data. To assess the impact of conducting experiments at  $(\mathbf{x}_{\text{ind}}^{(k)}, \mathbf{x}_s^{(k)}, \boldsymbol{\mu}_{u-L}^{(k)})$  ( $k = 1, \dots, N_{Le}$ ) and conducting simulations at  $(\mathbf{x}_{\text{ind}}^{(k)}, \mathbf{x}_s^{(k)}, \boldsymbol{\mu}_{u-L}^{(k)})$  ( $k = N_{Le} + 1, \dots, N_L$ ) for response

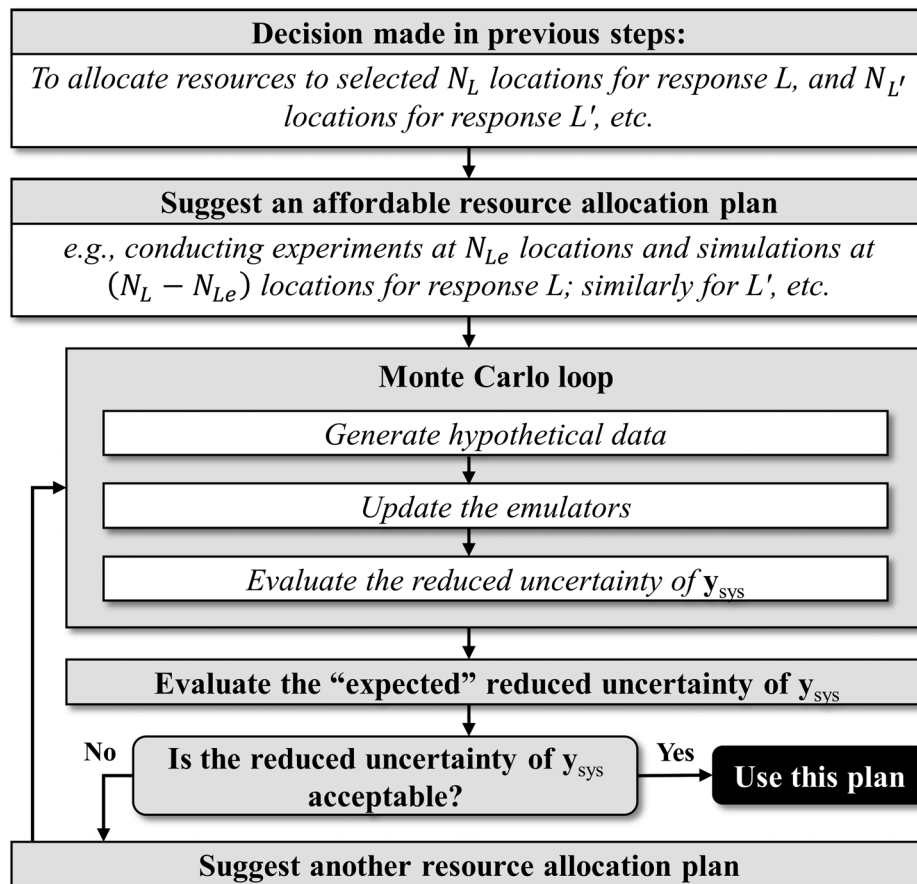


Fig. 4 The preposterior analysis

$L$ , hypothetical experimental and simulation data are generated accordingly at these locations. A preferred way is to generate them simultaneously considering their correlation. It should be noted that  $L^e$  (i.e., the experimental response of  $L$ ) at  $(\mathbf{x}_{ind}^{(k)}, \mathbf{x}_s^{(k)}, \boldsymbol{\mu}_{u-L}^{(k)})$  ( $k = 1, \dots, N_{Le}$ ) and  $L^m$  (i.e., the simulation response of  $L$ ) at  $(\mathbf{x}_{ind}^{(k)}, \mathbf{x}_s^{(k)}, \boldsymbol{\mu}_{u-L}^{(k)})$  ( $k = N_{Le} + 1, \dots, N_L$ ) follow a jointly normal distribution that can be determined from the already estimated hyperparameters of the GP models. The equation for the jointly normal distribution is similar to Eq. (3), i.e.,  $\mathbf{d} \sim \mathcal{N}(\mathbf{H}\boldsymbol{\beta}, \mathbf{V}_d)$ , where  $\mathbf{d}$  here should be interpreted as the hypothetical data to be generated, and  $\mathbf{V}_d$  is a covariance matrix whose entries are the covariances between selected input settings and/or between  $L^e$  and  $L^m$ . Therefore, a set of hypothetical experimental and simulation data can be generated by drawing a sample from this jointly normal distribution.

- (2) Update the emulators after obtaining the hypothetical data. To calculate how the emulators will change after adding the hypothetical data (which we treat as real data in the preposterior analysis), one could do another round of model bias correction using the collection of real data and hypothetical data, although this could be time consuming. Alternatively, by assuming that the estimation of the hyperparameters of the GP models will not significantly change with a few additional data, one can simply obtain a new model prediction by reusing the already estimated values of the hyperparameters in Eq. (5). The symbols now have new meanings; for example,  $\mathbf{d}$  and  $\mathbf{V}_d$  in Eq. (5) are now referred to as the collection of real and hypothetical data and their covariance matrix, respectively.
- (3) Conduct uncertainty propagation using the new model prediction.

After all the emulators for disciplinary responses are updated using the hypothetical data, uncertainty propagation as presented in Sec. 3.1 can provide information about how much the uncertainty of system QOIs  $y_{sys}$  is changed.

- (4) Calculate the "expected" reduced uncertainty of system QOIs  $y_{sys}$ . Note that steps 1–3 in this subsection are done under a specific set of hypothetical data. In order to achieve statistically meaningful results, steps 1–3 should be repeated by generating multiple sets of hypothetical data, which leads to an expected (averaged) reduced uncertainty of  $y_{sys}$ .

The preposterior analysis can be conducted for any suggested resource allocation plan, and users can select a plan that balances effectiveness and cost. Practically, one can first analyze the expected uncertainty reduction under the plan that all additional resources are from simulations, and then check whether the criterion (9) will be satisfied. Those locations where this criterion is not likely to be satisfied after gathering new simulations should be considered for experimental measurements.

#### 4 Case Study

To demonstrate the proposed resource allocation approach, we consider an electronic packaging design problem [46], a benchmark multidisciplinary problem that has been frequently studied in the literature [23,25,47–49]. A circuit consisting of two resistors is mounted on a heat sink, and there are two coupled disciplines, i.e., the electrical and thermal subsystems, as demonstrated in Fig. 5. Eight input variables  $x_1$ – $x_8$ , five linking variables  $y_6$ ,  $y_7$ , and  $y_{11}$ – $y_{13}$ , and a system QOI  $y_1$  are involved in this test problem. The numbering of the variables follows the same as used in the aforementioned works. Table 1 provides the physical



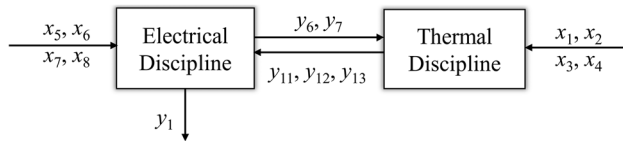


Fig. 5 Electronic packaging system

Table 1 Physical meanings of the variables in the electronic packaging system

$x_1$	Heat sink width (m)
$x_2$	Heat sink length (m)
$x_3$	Fin length (m)
$x_4$	Fin width (m)
$x_5$	Nominal resistance #1 at temperature 20 °C ( $\Omega$ )
$x_6$	Temperature coefficient of electrical resistance #1 ( $K^{-1}$ )
$x_7$	Nominal resistance #2 at temperature 20 °C ( $\Omega$ )
$x_8$	Temperature coefficient of electrical resistance #2 ( $K^{-1}$ )
$y_1$	Negative of watt density ( $W/m^3$ )
$y_6$	Power dissipation in resistor #1 (W)
$y_7$	Power dissipation in resistor #2 (W)
$y_{11}$	Component temperature of resistor #1 (°C)
$y_{12}$	Component temperature of resistor #2 (°C)
$y_{13}$	Heat sink volume ( $m^3$ )

meanings of the variables; a detailed problem statement that includes descriptions of all variables and the bounds of the design variables can be found in the referenced literature.

The original problem is a design optimization formulation to maximize the watt density (i.e., to minimize  $y_1$ ). To test our proposed approach, we modify the original problem as follows. Instead of finding an optimal solution, we now aim at achieving a good predictive model of  $y_1$  over the whole design space of  $x_1$ – $x_8$ . Since the thermal discipline is much more complex than the electrical discipline in that it requires a finite difference strategy to calculate the component temperatures, we assume that the models to evaluate  $y_{11}$ ,  $y_{12}$ , given  $\{x_1, x_2, x_3, x_4, y_6, y_7\}$ , are inaccurate and require model bias correction. In the first iteration of resource allocation, we collect simulation and experimental data and conduct model bias correction for  $y_{11}$  and  $y_{12}$  (Sec. 4.1), and then evaluate the impact of epistemic model uncertainty on the system QOI  $y_1$  (Sec. 4.2), and then apply the proposed resource allocation approach to reduce the aggregated uncertainty on  $y_1$  (Sec. 4.3). The results from subsequent iterations are provided in Sec. 4.4.

**4.1 Model Bias Correction for the Thermal Discipline (First Iteration).** We first collect data comprised of 40 “experimental” observations. Since the original benchmark problem does not provide any experimental data, we generate a set of data for  $y_{11}$  and  $y_{12}$  by adding random noise to the exact model provided in Ref. [46], which we treat as the experimental data. The 40 experimental runs are determined via an LHS over the six-dimensional input space  $\{x_1, x_2, x_3, x_4, y_6, y_7\}$ . We then collect data comprised of 40 simulation observations of  $y_{11}$  and  $y_{12}$  from low-fidelity simulation models that we build, which intentionally have a significant discrepancy  $\delta(\mathbf{x})$  compared with the exact model in Ref. [46]. The 40 simulation runs are determined via a different LHS, so that their locations are completely different than those for the experimental runs. Note that  $\{y_6, y_7\}$  are input variables to the thermal discipline and yet essentially linking variables in the whole system; hence, we do not know beforehand their ranges. We assume in this stage that they are both within  $[0, 8]$  based on some preliminary system-level simulations (using the current “wrong” simulation models). We will show later that their actual bounds are slightly beyond that.

The data points are shown in Fig. 6, with only the  $\{y_6, y_7\}$  plotted for illustration purpose. The GP-based model bias correction

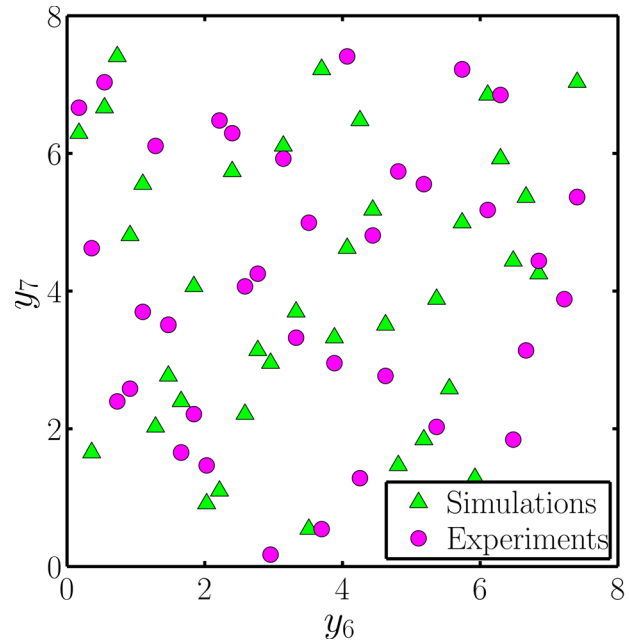


Fig. 6 Design of experiments for model bias correction in the thermal discipline: 40 experiments + 40 simulations sparsely over the six-dimensional space  $\{x_1, x_2, x_3, x_4, y_6, y_7\}$ . The data points are plotted by projecting them onto  $\{y_6, y_7\}$  for illustration.

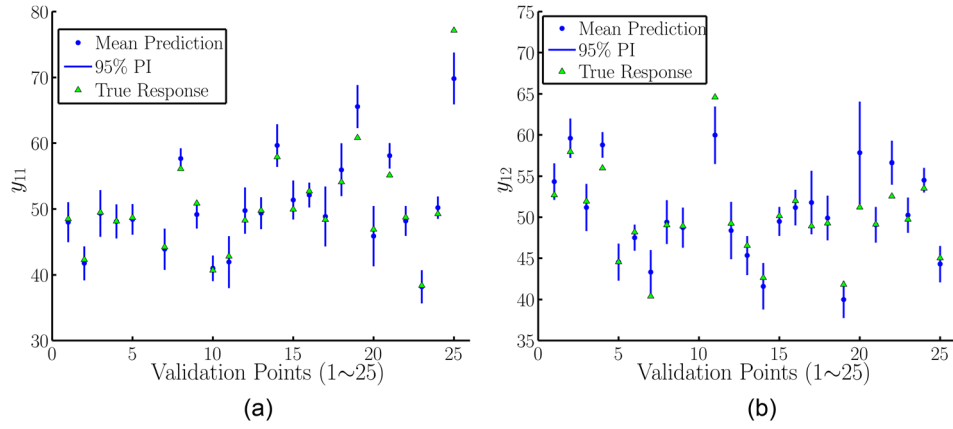
in Sec. 2.1 is applied to integrate the two sets of data and to subsequently obtain the updated emulators of  $y_{11}$  and  $y_{12}$ . An extra validation step is performed with 25 reserved experimental data; a comparison between the updated emulators and the validation data, together with the 95% PIs, is provided in Fig. 7. It can be seen that for both responses, 40 experimental data + 40 simulations are barely sufficient for bias correction: at some validation points, the PIs are large, and a few validation data are even outside the 95% PI. It can be anticipated that such uncertainty will propagate and influence the system QOI  $y_1$ .

**4.2 Selection of Input Settings and Responses for Resources Allocation (First Iteration).** In this step, 2000 LHS samples are generated to explore the input space of  $x_1$ – $x_8$  (see Fig. 8). The MUA method described in Sec. 3.1 evaluates the aggregated uncertainty of  $y_1$  at each sample point. As a reminder, the analysis in this step only relies upon the emulators that we just built in Sec. 4.1, and thus, the computations for 2000 samples are not at all prohibitive.

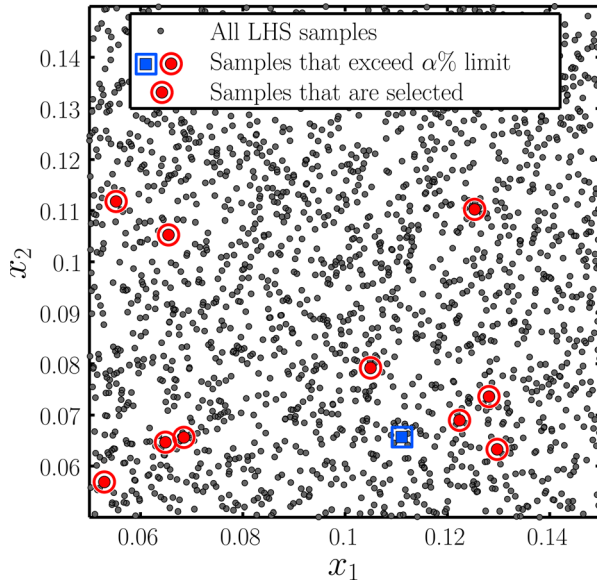
Uncertainty acceptance criterion  $\alpha\%$  in Eq. (9) is set to be 10%. It turns out that there are only 11 samples (the highlighted blue rectangle and red circles in Fig. 8) out of 2000 that violate the acceptance criterion, which indicates an overall acceptable amount of uncertainty in the design space. However, the largest  $\gamma$  in Eq. (9) is 47.31%, which shows that in some local areas, the uncertainty is huge. Using the correlation check presented in Sec. 3.3 (with the correlation limit  $\beta$  set to be 0.95), 10 samples (the red circles) out of 11 are selected for allocating more resources.

The result of MSSA at the ten selected input settings is provided in Fig. 9, based on which we decide to allocate more resources to  $y_{11}$  at input settings #1–3, 5, and 10 (since at these points the MSI of  $y_{11}$  is larger than that of  $y_{12}$ ), and to  $y_{12}$  at the rest of input settings. Note that here we only plot the MSI of  $y_{11}$  and  $y_{12}$ . It is because they are only two uncertainty sources in this problem, and based on the definitions of MSI and TSI in Eqs. (18) and (19), we have

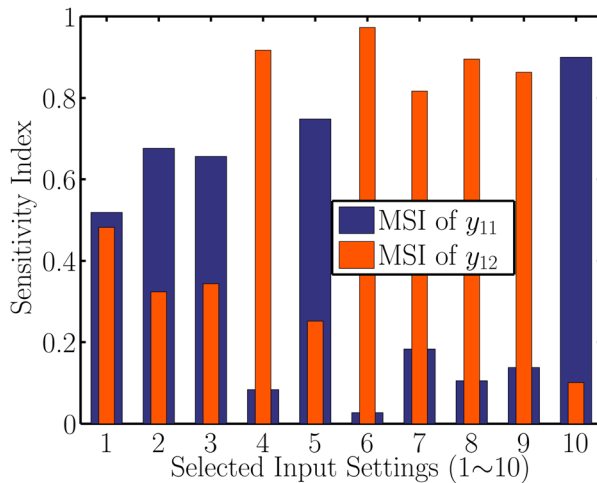




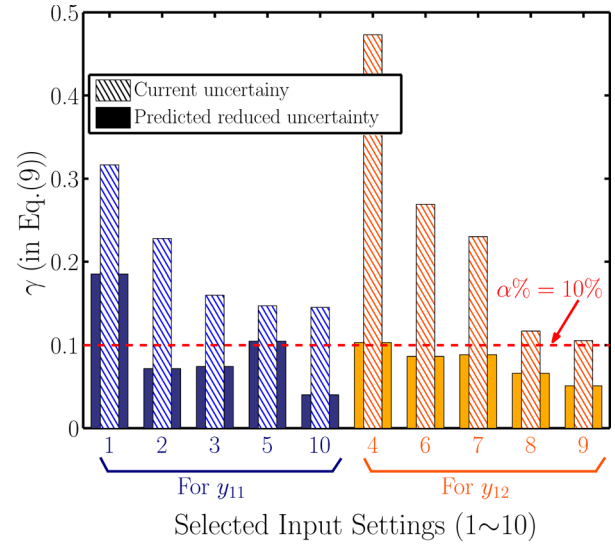
**Fig. 7 Comparison between the validation data and the updated emulators of (a)  $y_{11}$  and (b)  $y_{12}$  after model bias correction**



**Fig. 8 Two thousand LHS samples to explore the eight-dimensional design space  $x_1$ – $x_8$  (projected onto  $\{x_1, x_2\}$  for illustration), and the final selected samples for resource allocation**



**Fig. 9 MSSA of the selected ten input settings**

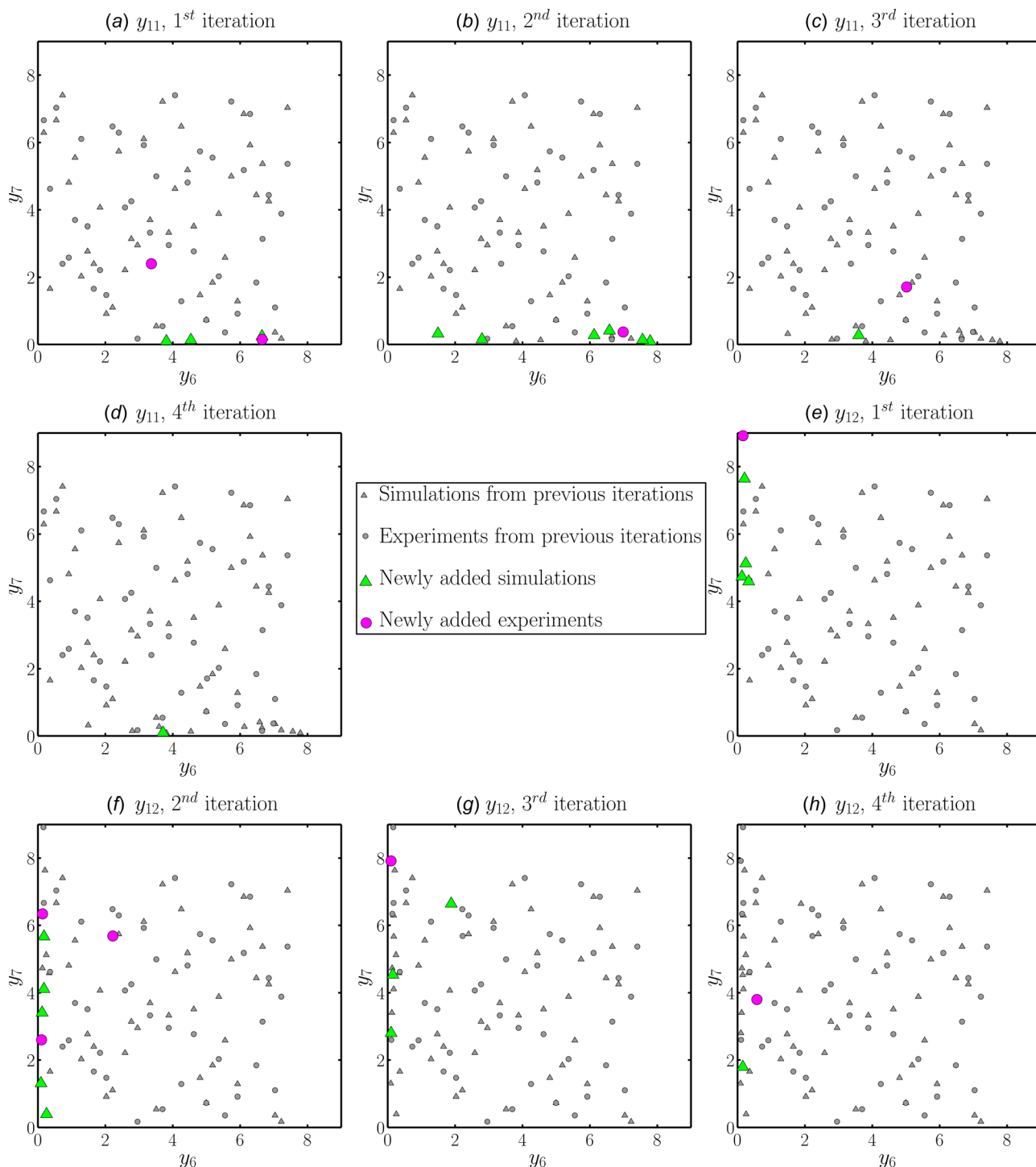


**Fig. 10 Preposterior analysis: comparison between the current uncertainty and the expected reduced uncertainty (calculated by taking the average of the reduced uncertainty under 500 sets of hypothetical simulation data)**

$$\begin{aligned} \text{TSI}(Z_{y_{11}}) &= 1 - \text{MSI}(Z_{y_{12}}) \\ \text{TSI}(Z_{y_{12}}) &= 1 - \text{MSI}(Z_{y_{11}}) \end{aligned} \quad (21)$$

We omit the plots of TSI as MSI offers sufficient information from sensitivity analysis.

**4.3 Preposterior Analysis to Decide the Type of Resources to Allocate (First Iteration).** We now apply the preposterior analysis to decide which type of resources to allocate, for each of the ten selected input settings. A total of 500 random sets of data are drawn from the joint normal distribution of  $y_{11}^m$  at input settings #1–3, 5, and 10, and  $y_{12}^m$  at input settings #4 and 6–9, which compose the hypothetical simulation data (a  $500 \times 10$  matrix). This joint distribution comes from the GP models of  $y_{11}^m$  and  $y_{12}^m$  built in Sec. 4.1. By adding one set of hypothetical data (one row of the matrix), the emulators for  $y_{11}$  and  $y_{12}$  are both updated (hypothetically), based on which the aggregated uncertainty of system QOI  $y_1$  (presented by  $\gamma$  in Eq. (9)) is reevaluated (again, for each of the ten selected input settings). The predicted reduced uncertainty of  $y_1$  after gathering more simulation data is taken to be the average of the reevaluated  $\gamma$  over the 500 hypothetical sets of simulation data.



**Fig. 11** Locations and types of the allocated resources for  $y_{11}$  and  $y_{12}$ : (a)  $y_{11}$ , first iteration, (b)  $y_{11}$ , second iteration, (c)  $y_{11}$ , third iteration, (d)  $y_{11}$ , fourth iteration, (e)  $y_{12}$ , first iteration, (f)  $y_{12}$ , second iteration, (g)  $y_{12}$ , third iteration, and (h)  $y_{12}$ , fourth iteration. Data are over the six-dimensional space  $\{x_1, x_2, x_3, x_4, y_6, y_7\}$  but projected onto  $\{y_6, y_7\}$  for illustration.

The result of the preposterior analysis is shown in Fig. 10. The uncertainty of  $y_1$  at most of the input settings (#2, 3, and 6–10) is expected to be reduced to the acceptable level ( $\alpha\% = 10\%$  in Eq. (9)); however, points #1, 4, and 5 still do not satisfy the uncertainty criterion. Therefore, instead of allocating simulation data to all the ten points, we should rather allocate experimental data to #1, 4, and 5.

To sum up, our decision for resource allocation after the first iteration of analysis is:

- (1) conduct simulations of  $y_{11}$  at points #2, 3, and 10 and  $y_{12}$  at points #6–9
- (2) conduct physical experiments of  $y_{11}$  at points #1 and 5 and  $y_{12}$  at point #4

**4.4 Decisions for Resource Allocation (Second–Fourth Iterations).** The process of resource allocation is iterative. The 2000 LHS samples in each iteration are different, and hence help

**Table 2 Number of total samples for model bias correction after each iteration**

Iteration #	$y_{11}^e$	$y_{11}^m$	$y_{12}^e$	$y_{12}^m$
0	40	40	40	40
1	42	43	41	44
2	43	49	44	49
3	44	50	45	52
4	44	51	46	53

to thoroughly explore the whole design space. It turns out that after four iterations, all of the LHS samples generated satisfy the uncertainty limit specified in Eq. (9). The locations and types of the added resources and the corresponding responses after each iteration are provided in Fig. 11. The number of total samples (for model bias correction) after each iteration is provided in Table 2.

In Fig. 11, it is seen that although  $\{y_6, y_7\}$  are initially assumed to be within the range of  $[0, 8]^2$  (discussed in Sec. 4.1), the actual range for  $\{y_6, y_7\}$  is slightly beyond that and there are some new samples selected in the later iterations that reach the region where  $y_6 > 8$  or  $y_7 > 8$ . We consider it a merit of the proposed method that it can gradually detect the region where we might have overlooked as the models get improved.

Beyond considering only physical experiments in existing literature, we treat computer simulations as yet another useful tool to reduce the epistemic model uncertainty, and this example showcases the merit of allocating computational resources. As shown in Table 2, more than 70% of the allocated resources (24 out of 34) are from computer simulations. We believe that it is because computer simulations help reduce the uncertainty due to the lack of data, and potentially also help discover the fundamental features of disciplinary responses. One can imagine that in an ideal situation when simulation models are perfect, uncertainty reduction can be achieved by conducting only simulations without help from physical experiments.

Another observation from this case study is that, to achieve a globally satisfactory model for QOIs in a multidisciplinary system does not require the disciplinary subsystem emulators to be globally satisfactory in their own input domains. In other words, the allocated resources to disciplinary subsystems do not necessarily cover the input space of disciplinary models. For example, it can be seen from Fig. 11 that all the resources allocated to  $y_{11}$  are clustered in a narrow band where  $y_7$  is small, and that the resources allocated to  $y_{12}$  are clustered in a narrow band where  $y_6$  is small. It is due to the complex coupling relationship between different disciplines in the system; many combinations of  $\{y_6, y_7\}$  are not achievable when the two disciplines couple together, and hence, there is no need to improve the modeling capability over those unachievable regions.

## 5 Conclusions

In this paper, we develop a new approach for resource allocation in multidisciplinary systems to reduce the aggregated epistemic model uncertainty of system QOIs. The novelty of the approach lies in breaking down the complex resource allocation problem that involves multiple decisions into a sequential set of decisions to answer questions of *where* (sampling locations), *what* (disciplinary responses), and *which* (simulations versus experiments). The proposed sequential decision-making strategy manages the complexity in creating globally accurate surrogate models for each subsystem in multidisciplinary analysis and design optimization by allocating resources (samples) that have the largest impact on reducing the uncertainty in predicting system QOI. After quantifying disciplinary epistemic model uncertainty using GP-based model bias correction, we first explore the input space of the whole system to identify the locations (*where*) with unacceptable amounts of uncertainty with respect to

the system QOIs. In this step, the MUA method provides a fast and analytical uncertainty analysis. Next, the identification of critical responses (*what*) at the selected inputs is done using an MSSA we developed earlier to evaluate the relative contribution of each functional response. Meanwhile, the proposed correlation analysis uses the information from GP-based correlation to determine the proximity of input settings and to ensure that the selected input settings are sparsely located. Finally, decisions are made about *which* type of resources to allocate to the critical responses at the chosen input locations, via an enhanced preposterior analysis that predicts the effect of gathering more resources prior to the actual resource allocation. The preposterior analysis, unlike any preceding work in the literature, is applicable to any combination of experiments and simulations. The proposed method is applied to a benchmark electronic packaging problem to reduce the aggregated epistemic model uncertainty of the system QOI.

The proposed approach has the following advantages. First, it considers not only the physical experiments but also the computer simulations. We demonstrate in the case study that in many cases, the physical experiments are not really necessary, if the epistemic model uncertainty mainly comes from lack of data. Second, because the proposed approach strategically breaks resource allocation into a sequential process, the decision making is much more tractable. Third, the method is efficient for complex multidisciplinary analysis. Except for the stage when the simulation or physical experiments are collected, all the analyses are conducted based on inexpensive GP emulators. The MUA and MSSA methods further accelerate the procedure by providing analytical formulas. Last but not least, while we do not elaborate it, the proposed framework can be simplified to work for a single discipline as well. In that case, MUA and MSSA may not be useful as a single discipline typically involves a single response; however, the GP-based correlation analysis and the preposterior analysis are still applicable.

The limitation of the work is that it does not explicitly consider the costs of simulations and experiments, although users can adjust the tunable preference parameters of the uncertainty acceptance level and correlation coefficient of samples based on the total budget. Our future research will include the budget as a constraint in the decision-making process of resource allocation.

## Acknowledgment

The grant support from the National Science Foundation (CMMI-1233403 and CMMI-1537641) is greatly acknowledged. Shishi Chen's predoctoral visit at Northwestern University is sponsored by the China Scholarship Council.

## Nomenclature

- $\mathbf{d}$  = collected data
- $\mathbf{h}^m(\cdot), \mathbf{h}^\delta(\cdot)$  = predefined basis function for polynomial regression of mean function of simulation model GP/discrepancy function GP
- MSI = main sensitivity index
- ND = number of disciplines
- TSI = total sensitivity index
- $\mathbf{u}_{ij}$  = linking variables, output from the  $i$ th discipline and input to the  $j$ th discipline
- $V^m(\cdot, \cdot), V^\delta(\cdot, \cdot)$  = covariance function of simulation model GP/discrepancy function GP
- $\mathbf{x} = (x_1, \dots, x_p) = p$ -dimensional input variable
- $\mathbf{x}_i, \mathbf{x}_{\text{ind}}$  = disciplinary input variables of the  $i$ th discipline/all ND disciplines
- $\mathbf{x}_s$  = shared input variables across two or more disciplines
- $\mathbf{y}_i, \mathbf{y}_{\text{ind}}$  = disciplinary outputs of the  $i$ th discipline/all ND disciplines
- $\mathbf{y}_{\text{sys}}$  = system QOIs
- $y^m(\cdot), y^e(\cdot)$  = simulation/experimental response



$\hat{y}^e(\cdot)$  = mean prediction of the experimental response  
 $Z(\cdot)$  = a random quantity representing model uncertainty  
 $Z_l$  = model uncertainty term for model  $L$   
 $\mathbf{Z}_{\sim l}$  = model uncertainty terms for models except  $L$   
 $\alpha\%$  = predefined uncertainty limit for  $\gamma(\cdot, \cdot)$   
 $\beta$  = predefined limit for  $\rho$   
 $\beta^m, \beta^\delta$  = coefficients for polynomial regression of mean function of simulation model GP/discrepancy function GP  
 $\gamma(\cdot, \cdot)$  = measure of uncertainty at a particular input setting  
 $\delta(\cdot)$  = model discrepancy function  
 $\varepsilon$  = experimental error  
 $\mu$  = mean vector; subscript, if available, denotes a specific random quantify, e.g.,  $\mu_{ui}$  refers to the mean of  $\mathbf{u}_i$ .  
 $\rho$  = correlation coefficient  
 $\sigma_Z^2(\cdot)$  = variance of  $Z(\cdot)$   
 $\Sigma$  = covariance matrix; subscript, if available, denotes a specific random quantify, e.g.,  $\Sigma_{uu}$  refers to the covariance  $\mathbf{u}^e$   
 $\phi$  = hyperparameters of GP

## Appendix

### Pseudocode for Choosing the Sample Points in Sec. 3.3.

---

Given:  $(\mathbf{x}_{\text{ind}}^{(k)}, \mathbf{x}_s^{(k)})$ ,  $k = 1, \dots, N_s$  from Section 3.1

1. Sort  $k$  by  $\text{Var}[\mathbf{y}_{\text{sys}}(\mathbf{x}_{\text{ind}}^{(k)}, \mathbf{x}_s^{(k)})]$  in descending order
2. Initialize  $\alpha, \beta$ ,  $\text{list1} = \{\xi | (\mathbf{x}_{\text{ind}}^{(\xi)}, \mathbf{x}_s^{(\xi)}) \text{ violates Eq. (9)}\}$ ,  $\text{list2} = \{1\}$
3. for all  $\xi \in \text{list1}$
4.   for all  $\eta \in \text{list2}$
5.     find response  $L$  that contributes most to  $\text{Var}[\mathbf{y}_{\text{sys}}(\mathbf{x}_{\text{ind}}^{(\xi)}, \mathbf{x}_s^{(\xi)})]$
6.     find response  $L'$  that contributes most to  $\text{Var}[\mathbf{y}_{\text{sys}}(\mathbf{x}_{\text{ind}}^{(\eta)}, \mathbf{x}_s^{(\eta)})]$
7.     if  $L \neq L'$ , then
8.       set  $\text{flag}(\xi, \eta) = 1$
9.     elseif  $\rho_{\xi, \eta} < \beta$  [Eq. (20)], then
10.       set  $\text{flag}(\xi, \eta) = 1$
11.     else
12.       set  $\text{flag}(\xi, \eta) = 0$
13.     endif
14.   endfor
15.   if  $\text{flag}(\xi, \eta) = 1, \forall \eta \in \text{list2}$ , then
16.     add  $\xi$  to  $\text{list2}$
17.   endif
18.   endfor

---

The chosen sample points are stored in “list2.”

## References

- [1] Kennedy, M. C., and O'Hagan, A., 2001, “Bayesian Calibration of Computer Models,” *J. R. Stat. Soc. Ser. B*, **63**(3), pp. 425–450.
- [2] Apley, D. W., Liu, J., and Chen, W., 2006, “Understanding the Effects of Model Uncertainty in Robust Design With Computer Experiments,” *ASME J. Mech. Des.*, **128**(4), pp. 945–958.
- [3] Kennedy, M. C., and O'Hagan, A., 2000, “Predicting the Output From a Complex Computer Code When Fast Approximations Are Available,” *Biometrika*, **87**(1), pp. 1–13.
- [4] Hasselman, T. K., Yap, K., Lin, C.-H., and Cafeo, J. A., 2005, “A Case Study in Model Improvement for Vehicle Crashworthiness Simulation,” 23rd International Modal Analysis Conference, Orlando, FL, pp. 1–12.
- [5] Chen, W., Xiong, Y., Tsui, K.-L., and Wang, S., 2008, “A Design-Driven Validation Approach Using Bayesian Prediction Models,” *ASME J. Mech. Des.*, **130**(2), p. 021101.
- [6] Qian, P. Z. G., and Wu, C. F. J., 2008, “Bayesian Hierarchical Modeling for Integrating Low-Accuracy and High-Accuracy Experiments,” *Technometrics*, **50**(2), pp. 192–204.
- [7] Wang, S. C., Chen, W., and Tsui, K. L., 2009, “Bayesian Validation of Computer Models,” *Technometrics*, **51**(4), pp. 439–451.
- [8] Jiang, Z., Chen, W., Fu, Y., and Yang, R.-J., 2013, “Reliability-Based Design Optimization With Model Bias and Data Uncertainty,” *SAE Int. J. Mater. Manuf.*, **6**(3), pp. 502–516.
- [9] Helton, J. C., Johnson, J. D., Sallaberry, C. J., and Storlie, C. B., 2006, “Survey of Sampling-Based Methods for Uncertainty and Sensitivity Analysis,” *Reliab. Eng. Syst. Saf.*, **91**(10–11), pp. 1175–1209.
- [10] Landau, D. P., and Binder, K., 2005, *A Guide to Monte Carlo Simulations in Statistical Physics*, 2nd ed., Cambridge University Press, New York.
- [11] Robert, C., and Casella, G., 1999, *Monte Carlo Statistical Methods*, Springer, New York.
- [12] Green, L. L., Lin, H.-Z., and Khalessi, M. R., 2002, “Probabilistic Methods for Uncertainty Propagation Applied to Aircraft Design,” *AIAA Paper No. 2002-3140*.
- [13] Cao, H., and Duan, B., 2004, “Uncertainty Analysis for Multidisciplinary Systems Based on Convex Models,” *AIAA Paper No. 2004-4504*.
- [14] Du, X., Guo, J., and Beeram, H., 2008, “Sequential Optimization and Reliability Assessment for Multidisciplinary Systems Design,” *Struct. Multidiscip. Optim.*, **35**(2), pp. 117–130.
- [15] Guo, J., and Du, X., 2010, “Reliability Analysis for Multidisciplinary Systems With Random and Interval Variables,” *AIAA J.*, **48**(1), pp. 82–91.
- [16] Gu, X., and Renaud, J. E., 2002, “Implementation Study of Implicit Uncertainty Propagation (IUP) in Decomposition-Based Optimization,” *AIAA Paper No. 2002-5416*.
- [17] Gu, X. S., Renaud, J. E., and Penninger, C. L., 2006, “Implicit Uncertainty Propagation for Robust Collaborative Optimization,” *ASME J. Mech. Des.*, **128**(4), pp. 1001–1013.
- [18] Du, X., and Chen, W., 2002, “Collaborative Reliability Analysis for Multidisciplinary Systems Design,” *AIAA Paper No. 2002-5474*.
- [19] Du, X., and Chen, W., 2005, “Collaborative Reliability Analysis Under the Framework of Multidisciplinary Systems Design,” *Optim. Eng.*, **6**(1), pp. 63–84.
- [20] Xiong, F., Yin, X., Chen, W., and Yang, S., 2010, “Enhanced Probabilistic Analytical Target Cascading With Application to Multi-Scale Design,” *Eng. Optim.*, **42**(6), pp. 581–592.
- [21] Sankararaman, S., and Mahadevan, S., 2012, “Likelihood-Based Approach to Multidisciplinary Analysis Under Uncertainty,” *ASME J. Mech. Des.*, **134**(3), p. 031008.
- [22] Liang, C., and Mahadevan, S., 2013, “Stochastic Multidisciplinary Analysis With High Dimensional Coupling,” 10th World Congress on Structural and Multidisciplinary Optimization, Orlando, FL, May 19–24, pp. 1–6.
- [23] Du, X., and Chen, W., 2000, “An Efficient Approach to Probabilistic Uncertainty Analysis in Simulation-Based Multidisciplinary Design,” *AIAA Paper No. 2000-0423*.
- [24] Du, X., and Chen, W., 2000, “Methodology for Managing the Effect of Uncertainty in Simulation-Based Design,” *AIAA J.*, **38**(8), pp. 1471–1478.
- [25] Du, X., and Chen, W., 2002, “Efficient Uncertainty Analysis Methods for Multidisciplinary Robust Design,” *AIAA J.*, **40**(3), pp. 545–552.
- [26] Jiang, X. M., and Mahadevan, S., 2009, “Bayesian Hierarchical Uncertainty Quantification by Structural Equation Modeling,” *Int. J. Numer. Methods Eng.*, **80**(6–7), pp. 717–737.
- [27] Jiang, X. M., and Mahadevan, S., 2009, “Bayesian Structural Equation Modeling Method for Hierarchical Model Validation,” *Reliab. Eng. Syst. Saf.*, **94**(4), pp. 796–809.
- [28] Sankararaman, S., McLemore, K., Mahadevan, S., Bradford, S. C., and Peterson, L. D., 2013, “Test Resource Allocation in Hierarchical Systems Using Bayesian Networks,” *AIAA J.*, **51**(3), pp. 537–550.
- [29] Manring, N. D., 2003, “Sensitivity Analysis of the Conical-Shaped Equivalent Model of a Bolted Joint,” *ASME J. Mech. Des.*, **125**(3), pp. 642–646.
- [30] Allaire, D., He, Q., Deyst, J., and Willcox, K., 2012, “An Information-Theoretic Metric of System Complexity With Application to Engineering System Design,” *ASME J. Mech. Des.*, **134**(10), p. 100906.
- [31] Jiang, Z., Chen, W., and German, B. J., 2014, “Statistical Sensitivity Analysis Considering Both Aleatory and Epistemic Uncertainties in Multidisciplinary Design,” *AIAA Paper No. 2014-2870*.
- [32] Jiang, Z., Chen, W., and German, B. J., 2016, “Multidisciplinary Statistical Sensitivity Analysis Considering Both Aleatory and Epistemic Uncertainties,” *AIAA J.*, **54**(4), pp. 1326–1338.
- [33] Jiang, Z., Li, W., Apley, D. W., and Chen, W., 2014, “A System Uncertainty Propagation Approach With Model Uncertainty Quantification in Multidisciplinary Design,” *ASME Paper No. DETC2014-34708*.
- [34] Jiang, Z., Li, W., Apley, D. W., and Chen, W., 2015, “A Spatial-Random-Process Based Multidisciplinary System Uncertainty Propagation Approach With Model Uncertainty,” *ASME J. Mech. Des.*, **137**(10), p. 101402.
- [35] Picheny, V., Ginsbourger, D., Roustant, O., Haftka, R. T., and Kim, N. H., 2010, “Adaptive Designs of Experiments for Accurate Approximation of a Target Region,” *ASME J. Mech. Des.*, **132**(7), p. 071008.
- [36] Liu, Y., Shi, Y., Zhou, Q., and Xiu, R., 2016, “A Sequential Sampling Strategy to Improve the Global Fidelity of Metamodels in Multi-Level System Design,” *Struct. Multidiscip. Optim.*, **53**(6), pp. 1295–1313.
- [37] Jones, D. R., Schonlau, M., and Welch, W. J., 1998, “Efficient Global Optimization of Expensive Black-Box Functions,” *J. Global Optim.*, **13**(4), pp. 455–492.
- [38] Schonlau, M., Welch, W. J., and Jones, D. R., 1998, *Global Versus Local Search in Constrained Optimization of Computer Models* (Lecture Notes—Monograph Series, Vol. 34), Institute of Mathematical Statistics, Bethesda, MD, pp. 11–25.
- [39] Arendt, P. D., Apley, D. W., and Chen, W., 2013, “Objective-Oriented Sequential Sampling for Simulation Based Robust Design Considering Multiple Sour-

- ces of Uncertainty,” *ASME J. Mech. Des.*, **135**(5), p. 051005.
- [40] Sacks, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P., 1989, “Design and Analysis of Computer Experiments,” *Stat. Sci.*, **4**(4), pp. 409–423.
- [41] Rasmussen, C. E., and Williams, C. K. I., 2006, *Gaussian Processes for Machine Learning*, The MIT Press, Cambridge, MA.
- [42] Sobol’, I. M., 1990, “On Sensitivity Estimation for Nonlinear Mathematical Models,” *Mat. Model.*, **2**(1), pp. 112–118 (in Russian).
- [43] Sobol’, I. M., 2001, “Global Sensitivity Indices for Nonlinear Mathematical Models and Their Monte Carlo Estimates,” *Math. Comput. Simul.*, **55**(1–3), pp. 271–280.
- [44] Arendt, P. D., 2012, “Quantification and Mitigation of Multiple Sources of Uncertainty in Simulation Based Design,” Ph.D. dissertation, Mechanical Engineering, Northwestern University, Evanston, IL.
- [45] Arendt, P. D., Apley, D. W., and Chen, W., 2016, “A Preposterior Analysis to Predict Identifiability in Experimental Calibration of Computer Models,” *IIE Trans.*, **48**(1), pp. 75–88.
- [46] Salas, A. O., 1995, “MDO Test Suite,” Last accessed May 1, 2015, <http://www.eng.buffalo.edu/Research/MODEL/mdo.test.orig/class2prob3.html>
- [47] Renaud, J. E., and Gabriele, G. A., 1994, “Approximation in Nonhierarchical System Optimization,” *AIAA J.*, **32**(1), pp. 198–205.
- [48] Padula, S. L., Alexandrov, N., and Green, L. L., 1996, “MDO Test Suite at NASA Langley Research Center,” *AIAA* Paper No. 96-4028.
- [49] Kodiyalam, S., and Sobieszczanski-Sobieski, J., 2001, “Multidisciplinary Design Optimization—Some Formal Methods, Framework Requirements, and Application to Vehicle Design,” *Int. J. Veh. Des.*, **25**(1–2), pp. 3–22.