AIAA JOURNAL
Vol. 57, No. 3, March 2019

Adaptive Sampling with Varying Sampling Cost for Design Space Exploration

Yiming Zhang,* Aravind Neelakantan,† Chanyoung Park,‡ Nam H. Kim,§ Herman Lam,¶ and Raphael T. Haftka**
University of Florida, Gainesville, Florida 32611

DOI: 10.2514/1.J057470

Surrogate models have been developed to infer the response of engineering systems based on scattered tests/simulations. An effective sampling scheme enables surrogates to have a desirable accuracy while balancing the sampling budget. Most sampling methods implicitly assume that all samples have the same cost to produce. In some applications, however, the cost to obtain samples may substantially vary in the input variable space because some configurations are more expensive to test or simulate than others. As an initial effort to incorporate with varying sampling cost, this paper explores an adaptive sampling strategy in which the sampling cost varies (AS-C). The proposed scheme adopts the Gaussian process for design space exploration, which is based on space filling. Two surrogates are constructed: one for the target function (quantity of interest) and the other for the sampling cost. Then a value metric is defined to estimate the uncertainty reduction per cost. A new sample is added per iteration at the point with the maximum value metric. The proposed AS-C is evaluated using 1D and 2D analytical functions. Four different cost functions and 100 sets of initial samples are produced for evaluation. For a fixed sampling budget, the AS-C adds more samples in an inexpensive region and thus provides a better accuracy than the standard adaptive sampling strategy (AS). As a case study, the AS-C is applied to the design space exploration of behavioral emulation (BE). BE is a coarse-grained simulation method, which predicts the runtime of a given simulation using high-performance computing. Because the cost/runtime of BE varies by the orders of magnitude, the AS-C adds many more samples in the inexpensive region and greatly outperforms the AS for a given sampling budget.

Nomenclature

$A_{\max}$	=	area metric for maximum error with increasing sampling budget
A_{R^2}	=	area metric for R-squared with increasing sampling budget
$C(\boldsymbol{x})$	=	cost function for sampling at $\boldsymbol{x}$
$\hat{C}(\boldsymbol{x})$	=	estimated cost function
$\mathrm{Cov}(\cdot)$	=	covariance function
$e_{\max}$	=	maximum error
$\hat{f}(\boldsymbol{x})$	=	surrogate model at the input $\boldsymbol{x}$
$\bar{f}(\boldsymbol{x}_j)$	=	mean of the function values at the grids
$H[n]$	=	Heaviside step function
p	=	dimensionality of the input variable space
$\hat{V}(\boldsymbol{x})$	=	estimated value function
$\mathrm{Var}(\boldsymbol{x})$	=	prediction variance of $\hat{f}(\boldsymbol{x})$
$\boldsymbol{x}$	=	a point in the multidimensional space
$\boldsymbol{x}_i$	=	the ith point in the input variable space
$Z(\boldsymbol{x})$	=	Gaussian process
μ	=	mean value of samples
θ_m	=	kriging hyperparameter vector with $m = 1, 2, \ldots, p$
σ^2	=	variance of the Gaussian process
$\hat{\sigma}(\boldsymbol{x})$	=	standard deviation of the prediction

Presented as Paper 2018-1163 at the 2018 AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, Kissimmee, FL, 8–12 January 2018; received 3 May 2018; revision received 13 September 2018; accepted for publication 21 October 2018; published online 24 January 2019.

*Graduate Student, Department of Mechanical and Aerospace Engineering. Student Member AIAA.
†Graduate Student, Department of Electrical and Computer Engineering.
‡Research Scientist, Department of Mechanical and Aerospace Engineering. Member AIAA.
§Professor, Department of Mechanical and Aerospace Engineering. Associate Fellow AIAA.
¶Associate Professor, Department of Electrical and Computer Engineering.
**Distinguished Professor, Department of Mechanical and Aerospace Engineering. Fellow AIAA.

I. Introduction

DESIGN optimization of engineering systems usually requires extensive simulations and tests to achieve desirable performances. Surrogate models are often introduced as an efficient tool to approximate the response of engineering systems from scattered simulations/tests in the input variable space. An effective surrogate model enables an inexpensive prediction of system responses at a given input configuration/design. Design optimization based on the surrogate model has been adopted for numerous engineering systems with significantly reduced simulation costs or test period [1–3]. Besides design optimization, surrogate models have been applied to a various engineering analysis. For example, advanced materials or innovative designs are emerging, which might lack effective theoretical models such as lattice structure and composite material [4]. Surrogate models may serve as an empirical model for the fundamental mechanics to enable multiscale analysis. Surrogates model is also a key technology for verification, validation, and uncertainty quantification such as the multifidelity surrogates, which is developed to compensate for the discrepancy between experiments and simulations [5,6]. For example, Alexandrov et al. [7] showed that different fidelity models can be used to save computational cost for optimization. Chaudhuri et al. [8] presented adaptive sampling (AS) with multifidelity surrogates for the multidisciplinary application.

The accuracy of surrogate models strongly depends on the location and number of samples (experiments and simulations). Methods to select samples are termed *design of experiments* (DOE). Desirable DOE enables accurate surrogate models while balancing the sampling budget. AS is a popular strategy of DOE, which adds samples iteratively with updated surrogates. Various AS strategies [9] were developed and proved effective for different applications, such as reduced order modeling [10,11], structural optimization [12], and reliability-based design optimization [13]. Efforts have been made to understand the major components of AS, such as surrogate types, prediction uncertainty metrics, updating schemes, and stopping criteria [9,14]. The AS could be more effective than all-at-once sampling based on the reported studies [14,15]. Toolboxes have been developed for the general usage of AS schemes [16,17].

One major component of AS is uncertainty metrics. Two types of uncertainty metrics are mainly used for AS: model-based and data-driven prediction variances. The prediction error at a point could be estimated by prediction variance. The prediction variance quantifies the variation of the potential system response, which is assumed a statistical distribution. The points with a large prediction variance imply a risky approximation and could be improved by adding more samples. The commonly used prediction variance is usually available for a few types of surrogates by inherently assuming a distribution for the samples. For example, the Gaussian process [18] assumes a multivariate normal distribution for the samples, while polynomial response surface assumes a normal distribution for the residual errors [19]. The Gaussian process-based AS has proved to be effective for various engineering applications [14,20]. Recent efforts have been made to obtain the uncertainty metric for an arbitrary surrogate based on data-driven approach such as cross-validation (CV). Jin et al. [21] estimated the prediction uncertainty using the difference between the surrogates from leaving-one-out CV and the surrogate using full samples. Ben Salem et al. [17] proposed a series of surrogates from leaving-one-out CV. Then a weighted scheme was used to fit the empirical cumulative distribution function (ECDF) from the predictions of the surrogates. The prediction variance is then generated from the ECDF to indicate the estimated uncertainty [22]. Xu et al. [23] associated surrogate prediction with the Voronoi diagram to determine the uncertainty of surrogate predictions.

The current practice of AS for single-fidelity surrogates implicitly assumes that all samples have the same cost, which is applicable for many applications. The number of samples is therefore used to indicate the cost/budget of sampling. However, in some applications, the cost to obtain samples may substantially vary in the input variable space. For example, the computational resource of a computational fluid dynamics (CFD) analysis relies heavily on the mesh density, which changes with Reynolds number and Mach number. The runtime of the CFD could also vary while using the warm-start strategy for accelerated computation [24]. The effect of sampling cost might be significant in the high-performance computing (HPC) environment. The HPC architecture and algorithm might change in the input variable space for optimum performance such as power consumption, an important part of sampling cost, for a given CFD simulation task [25,26].

This paper explores a DOE strategy with varying sampling costs. The cost/budget of sampling is reflected by the runtime of simulations. We examine the AS for adding samples iteratively. We take a cue from multifidelity and multisource optimization approaches, where estimated gain is maximized per unit cost [24,25]. A value metric is proposed to maximize the expected gain from an added sample per unit cost. In addition, unlike papers on multisource optimization, we assume that the actual cost of a sample is not known in advance. Therefore, two sets of surrogates are developed: one for the target function and the other for the sampling cost. The value metric is defined as the ratio between the standard deviation of prediction and the sampling cost estimated from the two surrogates. A new sample is then added at the point where the value metric is maximized (adaptive sampling with varying sampling cost [AS-C]). As analytical examples, algebraic test functions with algebraic cost functions are examined. We evaluated the proposed approach based on 1D and 2D algebraic functions, which are convenient for visualization and discussion.

As a case study, the proposed AS-C has been applied to the design space exploration of behavioral emulation (BE). BE predicts the run time of a given CFD simulation using HPC. As we move toward exascale computing, it is important for application developers and system architects to perform co-design to develop an optimized, energy-efficient application code and machine [27]. To speed up this co-design process and to enable architectural design space exploration, system architects build simulator models to study the performance of the application on various underlying conditions. BE [28] is one such coarse-grained approach for simulation of extreme-scale systems and application. Because the cost/run time of BE for different input configurations varies by orders of magnitude, it serves as an excellent case study for AS-C.

In the remainder of the paper, details on the standard AS using kriging are introduced in Sec. II. The proposed AS strategy with varying cost is presented in Sec. III. Section IV introduces the multivariate algebraic test function and four algebraic cost functions. Section V investigates the numerical performance of the proposed approach. Effect of different cost functions with increasing complexity and different initial samples are discussed. Section VI applies the AS-C to the approximation of BE with a comparison to AS, followed by conclusions in Sec. VII.

II. Adaptive Sampling Using Surrogate

A. Basic Steps in Adaptive Sampling

Adaptive/sequential sampling refers to the systematic procedure to add samples iteratively in order to improve the accuracy of surrogate prediction. Among different ways of AS, we consider here AS that relies on an uncertainty model of the surrogate and that seeks to improve the surrogate everywhere rather than toward a target. The basic steps for such AS are provided in Fig. 1. The initial samples are first generated with a portion of the total sampling budget. Then an initial surrogate model is built with prediction variance in the input variable space. One additional sample $\boldsymbol{x}_{\text{new}}$ is added at the point with maximum prediction variance as shown in Eq. (1), where $\text{Var}(\boldsymbol{x})$ is the prediction variance at the input $\boldsymbol{x}$.

$$\boldsymbol{x}_{\text{new}} = \arg\max_{\boldsymbol{x}} \text{Var}(\boldsymbol{x}) \tag{1}$$

Instead of maximum prediction variance, other criteria are also available to determine the new sampling point, such as expected improvement (EI), which is used to search for the possible minimum function value during optimization [29]. Multiple sampling points can be added per iteration to take advantage of parallel computation and to reduce the number of iterations. The stopping criterion plays an important role in the performance of AS [30,31]. The AS and modeling iterate until a prescribed stopping criterion based on total cost or estimated accuracy of the surrogate are met [31]. In this paper, kriging is used for surrogate modeling, where the prediction variance is based on its inherent assumption for multivariate normal distribution. The stopping criterion adopted here is the total sampling cost. Namely, the AS procedure stops just before the total sampling cost exceeds a prescribed sampling budget.

B. Kriging Surrogate

Kriging [19] with a constant trend is adopted in the paper for surrogate modeling as given in Eq. (2):

$$\hat{f}(\boldsymbol{x}) = \mu + Z(\boldsymbol{x}) \tag{2}$$

where μ is the mean value of the samples and $Z(\boldsymbol{x})$ is assumed to be a Gaussian process. The covariance function of the Gaussian process is

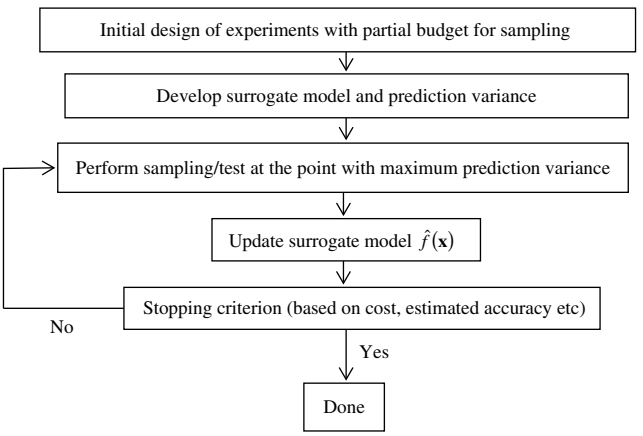

Fig. 1 Basic steps of classical adaptive sampling for global approximation. One additional sample is added per iteration.

set to be the squared anisotropic exponential function and given in Eq. (3):

$$\mathrm{Cov}(Z(\boldsymbol{x}_i), Z(\boldsymbol{x}_j)) = \sigma^2 \prod_{m=1}^{p} \exp(-\theta_m (x_{i,m} - x_{j,m})^2) \tag{3}$$

where $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ denote two points in the p-dimensional space, σ^2 is the process variance, and θ_m is the hyperparameter with $m = 1, 2, \ldots, p$. The parameters of kriging are obtained from maximum likelihood estimation. The kriging predictions interpolate samples, and therefore the prediction variance is zero at sample points. The kriging is based on the assumption of multivariate normal distribution and naturally provides the prediction variance at an untested point. Kriging has been proved effective to approximate the response of various systems and has been used as a major surrogate for AS. The implementation of kriging is based on the surrogate toolbox from Viana [32].

III. Adaptive Sampling Strategy with Varying Sampling Cost

A. Proposed Methodology

The classical AS strategy (shown in Fig. 1) is revised to incorporate the effect of varying sampling cost. For this purpose, a second surrogate $\hat{C}(x)$ is constructed to approximate the sampling cost. Then a value metric $\hat{V}(x)$ is defined as the ratio between the standard deviation of the prediction $\hat{\sigma}(\boldsymbol{x})$ and the cost prediction $\hat{C}(\boldsymbol{x})$ as shown in Eq. (4). The $\hat{V}(\boldsymbol{x})$ is evaluated at a grid of 100^p test points in the input variable space. One additional sample, $\boldsymbol{x}_{\text{new}}$, is determined at the point with the maximum value $\hat{V}(\boldsymbol{x})$ as in Eq. (5). The $\hat{V}(\boldsymbol{x})$ indicates the uncertainty reduction per unit cost. $\hat{\sigma}(\boldsymbol{x})$ and $\hat{V}(\boldsymbol{x})$ decrease to 0 after adding the sample at $\boldsymbol{x}$ because prediction uncertainty is zero at a sample point. The $\hat{\sigma}(\boldsymbol{x})$ is used to define $\hat{V}(\boldsymbol{x})$, considering that $\hat{\sigma}(\boldsymbol{x})$ is proportional to the confidence interval, which is the key measure of estimated prediction error. Note that AS-C becomes classical AS when the cost function is a constant. The flowchart of AS-C is summarized in Fig. 2.

$$\hat{V}(\boldsymbol{x}) = \frac{\hat{\sigma}(\boldsymbol{x})}{\hat{C}(\boldsymbol{x})} \tag{4}$$

$$\boldsymbol{x}_{\text{new}} = \arg\max_{\boldsymbol{x}} \hat{V}(\boldsymbol{x}) \tag{5}$$

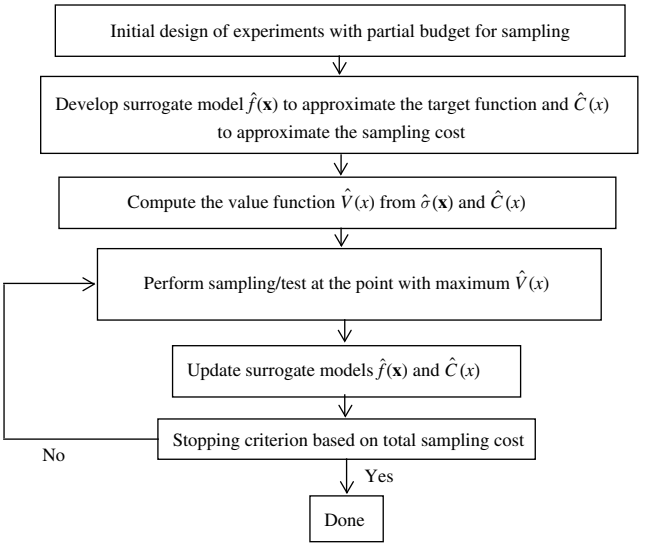

Fig. 2 Flowchart of the adaptive sampling with varying sampling cost. One additional sample is added per iteration at the point with maximum value metric to improve the global approximation.

B. Illustration of AS-C Using 1D Algebraic Function

The Forrester function [33] is selected for illustration and preliminary investigation of the proposed AS-C. The Forrester function is a 1D algebraic function on $x \in [0, 1]$ as given in Eq. (6):

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

A linear cost function is introduced in Eq. (7). The cost function needs to be positive in the input variable space. The response of Forrester function and associated cost function are visualized in Fig. 3.

$$C(x) = x + 0.1 \tag{7}$$

To approximate the Forrester function, the initial DOE is 3 samples at $x = 0$, 0.5, and 1. The initial samples and surrogates are shown in Fig. 4a for the target function and Fig. 4b for the cost function. The green-colored area represents 95% confidence intervals due to prediction uncertainty. It is clear that the accuracy of $\hat{f}(x)$ is poor due to the sparse initial samples. The cost function has a simple linear trend, and $\hat{C}(x)$ matches the true cost function even with just three samples. For a complicated cost function, $\hat{C}(x)$ is expected to be less accurate. Compared with AS, the AS-C is based on two surrogates and expected to suffer from larger uncertainty when the cost function has a complicated response.

After constructing the initial surrogate, one sample is added per iteration and a new surrogate is built using kriging. The sampling budget is set to be up to 4.5 in this example. Final predictions using AS resulted in 7 samples as shown in Fig. 5a, while AS-C in 8 samples as shown in Fig. 5b because AS-C chose samples in the low-cost region. Details of surrogate results are summarized in Table 1. Compared with AS, the AS-C enabled more samples in the input variable space for a fixed budget. Based on the maximum error, AS-C was more accurate than AS. In the following section, the AS-C is investigated further using a 2D test function with different cost functions and different initial samples.

C. Application of AS-C with Other Value Metrics

The prediction variance of GP is based on only the hyperparameters and spatial location, not a direct reflection on the goodness of the fit. For isotropic GP, AS with maximum variance is essentially a uniform spatial sampling scheme like Latin Hypercube Sampling (LHS). There have been proposals for improving on the prediction variance for error estimates and/or AS. Examples are the universal prediction distribution (UPD) [17] and the sequential sampling for global metamodeling [21]. The proposed AS-C could be used with all these uncertainty estimations. The effect of different uncertainty measures could be explored further.

As an initial effort to incorporate varying cost with sampling, this paper is limited to design space exploration that focuses on the global accuracy everywhere. We have performed a preliminary study on EI-based optimization with varying cost. For optimization, the quantity of interest is on the optima rather the whole design space. Adding more cheap runs far from optima might be not as valuable as the case for design space exploration.

The AS-C might fail (i.e., be worse than a regular AS) for highly nonlinear cost functions and sparse samples. As discussed later in Fig. 6, the AS-C deteriorated for the highly nonlinear cost functions due to the large error from the surrogate on cost estimation. But with increasing number of samples, the AS-C would be more accurate than AS. The essence of AS-C is adding more cheap runs instead of a few expensive ones to benefit the global accuracy everywhere.

IV. Multivariate Test Function and Algebraic Cost Functions

The complexity of cost function has a significant effect on the AS-C as discussed in Sec. III. In this section, a 2D test function is selected for further evaluation of AS-C. Four different cost functions are introduced to imitate the effect of varying sampling cost.

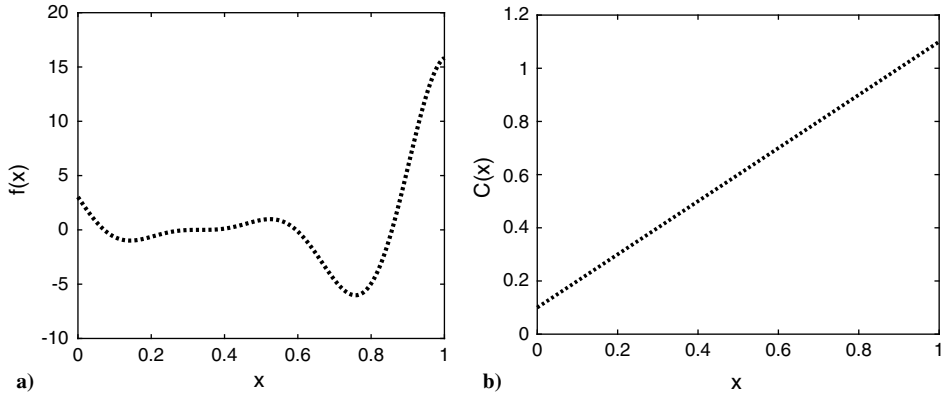

Fig. 3 One-dimensional algebraic example: a) Forrester function; b) cost function for sampling.

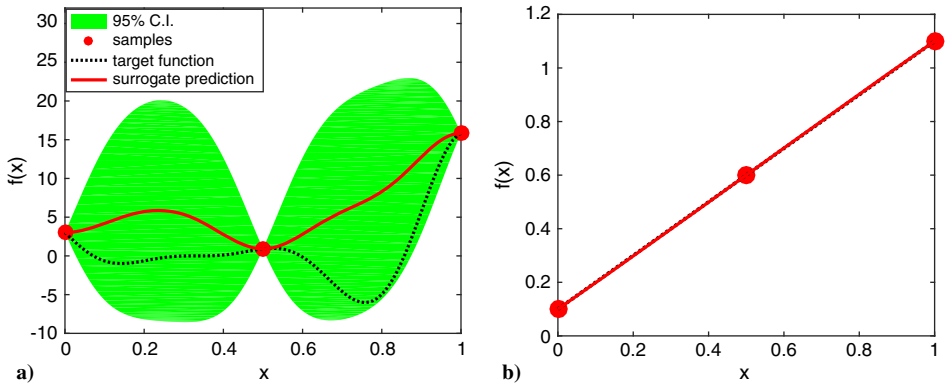

Fig. 4 Initial samples and surrogate prediction for a) the Forrester function and b) linear cost function.

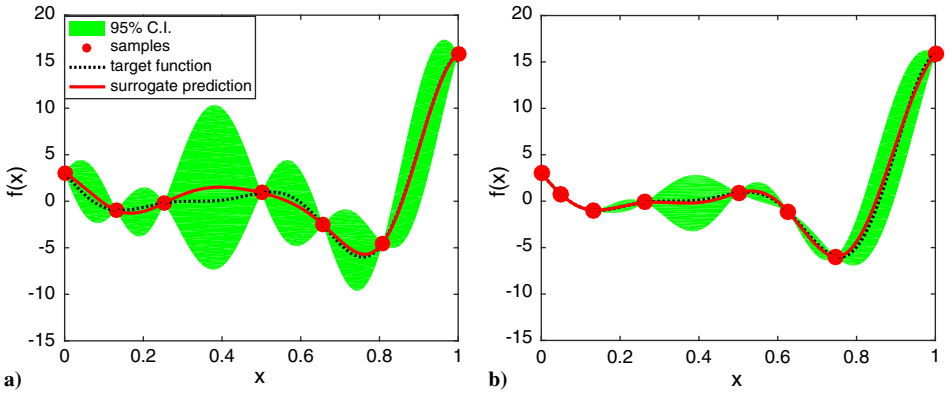

Fig. 5 Surrogate predictions for adaptive sampling plans for Forrester function example using a) AS and b) AS-C.

A. Normalized Branin Function

The Branin function is selected as the target function to evaluate AS-C. The function is in the 2D input variable space, which is convenient for visualization and inspection. The original Branin function is given in Eq. (8) and defined in $x_1 \in [-5, 10]$, $x_2 \in [0, 15]$. For the study of varying sampling cost, the input variable space of Branin function and cost function is normalized within $x_1, x_2 \in [0, 1]$ for consistency. The response of Branin function is normalized within $f(x) \in [0.1, 1.1]$ for a convenient quantitative study. The mapping/scaling of input variable space and function value is performed through a linear transformation. The response of the normalized Branin function is shown in Fig. 7.

Table 1 Details of AS and AS-C for the approximation of Forrester function with a budget of up to 4.5 total cost

	Number of samples	Total cost	Max error
AS	7	4.0	1.46
AS-C	8	4.1	0.87

$$f_{\text{Branin}}(\boldsymbol{x}) = \left(x_2 - \frac{5.1}{4\pi^2}x_1^2 + \frac{5}{\pi}x_1 - 6\right)^2 + 10\left(1 - \frac{1}{8\pi}\right)\cos(x_1) + 10 \tag{8}$$

B. Four Algebraic Cost Functions

Four different cost functions have been adopted to illustrate different cases of varying cost in the input variable space. They are a linear function, an exponential function, the Rosenbrock function, and the Damper function. The linear function is shown in Eq. (9). The exponential function shown in Eq. (10) increases more drastically in the input variable space. The Rosenbrock function is given in Eq. (11), which has a response with moderate nonlinearity. The Damper function [34] is given in Eq. (12) to imitate the resonance effect. The original input variable spaces of the four test functions are provided in Table 2. For the study of varying sampling cost, the input variable spaces of cost functions are normalized within $x_1, x_2 \in [0, 1]$ for consistency with the target function. The value of cost function is normalized within $f(x) \in [0.1, 1.1]$ for

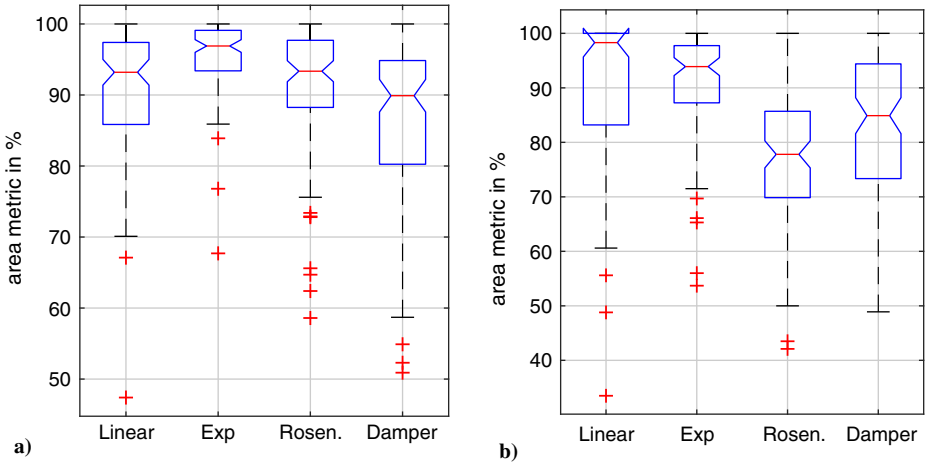

Fig. 6 The population of area metrics for the approximation of the Branin function with four cost functions: a) area metric for *R*-squared; b) area metric for max error.

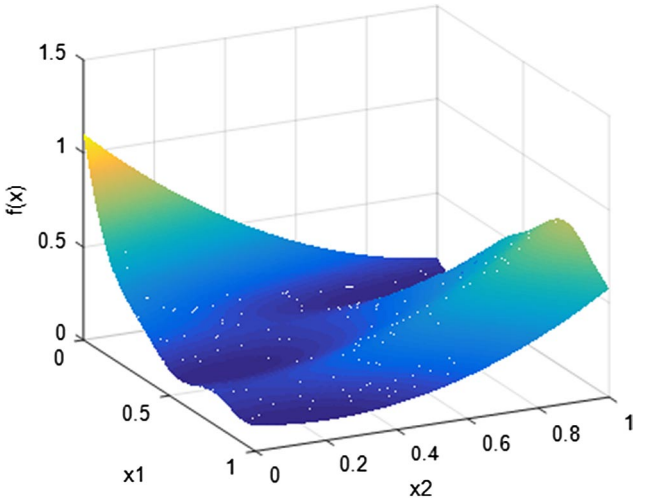

Fig. 7 Normalized Branin function in the input variable space.

convenient quantitative comparison. The lower bound 0.1 indicates the basic sampling cost. This has the effect of creating a cost ratio of 1/11 for all the cost functions between the cheapest and most expensive points. The response of normalized cost functions is visualized in Fig. 8.

$$C_{\text{Linear}}(\boldsymbol{x}) = 100x_1 + 100x_2 + 50 \tag{9}$$

$$C_{\text{Exp}}(\boldsymbol{x}) = \exp(5x_1) + \exp(5x_2) + 50 \tag{10}$$

$$C_{\text{Rosen.}}(\boldsymbol{x}) = 100(x_2 - x_1^2)^2 + (x_1 - 1)^2 \tag{11}$$

$$C_{\text{Damper}}(\boldsymbol{x}) = \frac{|1 - (1/x_2)^2|}{\sqrt{[1 - R(1/x_1)^2 - (1/x_1)^2 - (1/\beta_2)^2 + (1/x_1^2x_2^2)]^2 + 4\zeta^2[(1/x_1) - (1/x_1x_2^2)]^2}} \tag{12}$$

Table 2 Original input variable space for the four cost functions

Variable	$C_{\text{Linear}}(\boldsymbol{x})$	$C_{\text{Exp}}(\boldsymbol{x})$	$C_{\text{Rosen.}}(\boldsymbol{x})$	$C_{\text{Damper}}(\boldsymbol{x})$
x_1	[0, 1]	[0, 1]	[−1.5, −0.5]	[1, 1.1]
x_2	[0, 1]	[0, 1]	[2, 3]	[1, 1.1]

V. Numerical Performance of AS-C

A. Evaluation Plan

The proposed AS-C was compared with AS using the Branin function. For AS, only one surrogate was developed to approximate the target function, and the true cost functions in Eqs. (9–12) were used to evaluate the total cost of sampling. The stopping criterion of AS-C and AS was based on the total sampling cost. The sampling procedure stopped just before the total sampling cost exceeds 10. For a given set of samples, kriging surrogate with the constant trend is built as shown in Sec. II for both the target function and the cost function. The six initial samples were generated by the LHS with 5000 iterations. One hundred sets of initial samples were generated to account for the effect of sampling uncertainty. The key factors of the evaluation plan are summarized in Table 3.

The prediction accuracy of the surrogate was examined at a grid of 100×100 test points, $\boldsymbol{x}_j$, $j = 1, 2, \ldots, 10{,}000$. For a given set of initial samples, two prediction metrics were adopted to estimate the global and local performance of the surrogates. The global accuracy was quantified by the coefficient of determination R^2 as given in Eq. (13), where $f(\boldsymbol{x}_j)$ denotes the function true value at the jth grid point, $\hat{f}(\boldsymbol{x}_j)$ denotes the surrogate prediction, and $\bar{\boldsymbol{f}}$ is the mean of $f(\boldsymbol{x}_j)$. An R^2 of 1 indicates that the surrogate perfectly fits all the test points. The local performance of surrogates was quantified by the maximum error $e_{\max}$ in Eq. (14). The worst prediction/maximum error is usually of critical interest in addition to the overall accuracy.

$$R^2 = 1 - \frac{\sum_{j=1}^{100\times100}[f(\boldsymbol{x}_j) - \hat{f}(\boldsymbol{x}_j)]^2}{\sum_{j=1}^{100\times100}[f(\boldsymbol{x}_j) - \bar{\boldsymbol{f}}]^2} \tag{13}$$

$$e_{\max} = \max_{j\in[1,10000]}\Big[|f(\boldsymbol{x}_j) - \hat{f}(\boldsymbol{x}_j)|\Big] \tag{14}$$

R^2 and $e_{\max}$ measured the prediction accuracy at a given cost. The relative performance (higher accuracy) of AS and AS-C might change with increasing samples as illustrated in Fig. 9, where c_{initial} is the cost of initial samples. The performance of an AS scheme should be interpreted over the entire range of sampling cost. The area metrics A_{R^2} and $A_{\max}$ were proposed to measure the relative performance of

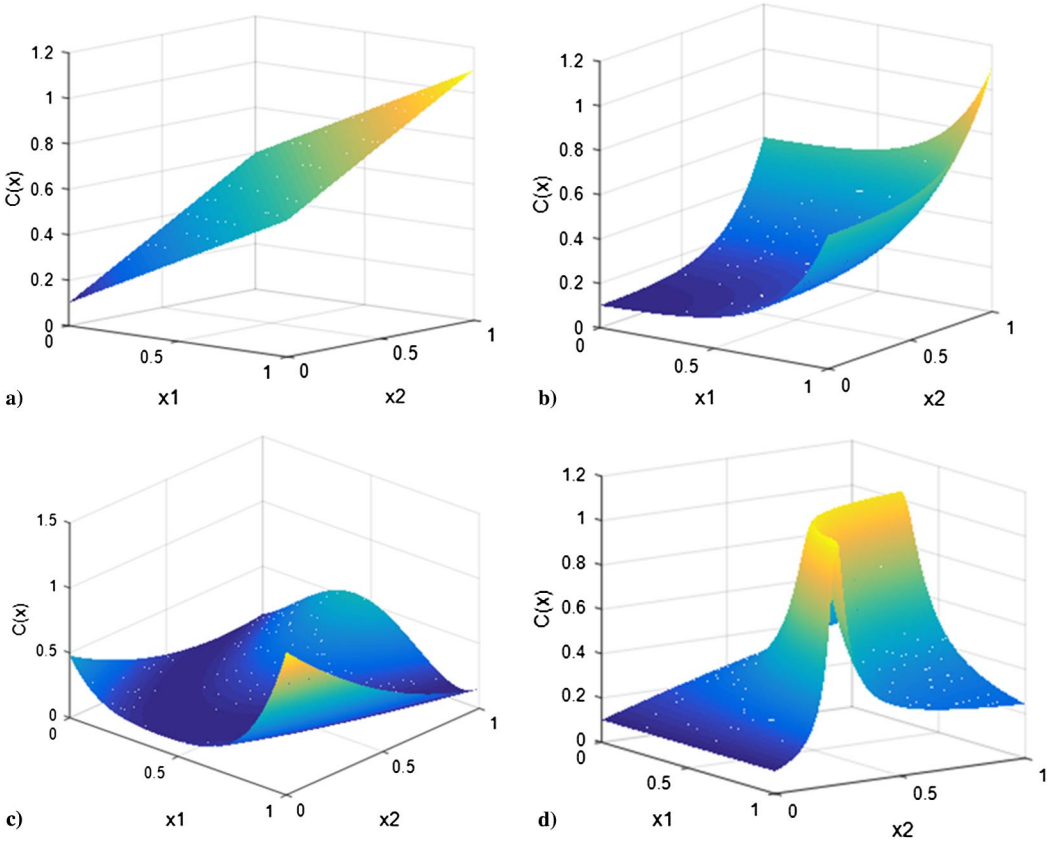

Fig. 8 Response of the four cost functions in normalized input variable space: a) linear cost function; b) exponential cost function; c) Rosenbrock cost function; d) Damper cost function.

AS and AS-C as given in Eqs. (15) and (16), where $H[\cdot]$ is the Heaviside step function defined in Eq. (17). The evolution of AS with increasing total sampling cost is interpolated between samples. A_{R^2} measures the percentage of times when AS-C is more accurate than AS over increasing samples with accuracy measured by R^2. $A_{\max}$ measured the percentage of times when AS-C dominates AS over increasing samples with accuracy measured $e_{\max}$. An area metric equal to 50% indicates a similar performance of AS and AS-C. The median and variation of A_{R^2} and $A_{\max}$ with different initial samples are also presented in the following sections.

$$A_{R^2} = \frac{\int_{c_{\text{initial}}}^{c_{\text{final}}} H[R^2_{\text{AS-}C}(c) - R^2_{\text{AS}}(c)]\,\mathrm{d}c}{c_{\text{final}} - c_{\text{initial}}} \times 100\% \tag{15}$$

$$A_{\max} = \frac{\int_{c_{\text{initial}}}^{c_{\text{final}}} H[e_{\max,\text{AS}}(c) - e_{\max,\text{AS-}C}(c)]\,\mathrm{d}c}{c_{\text{final}} - c_{\text{initial}}} \times 100\% \tag{16}$$

$$H[n] = \begin{cases} 0, & n < 0 \\ 1, & n \geq 0 \end{cases} \tag{17}$$

Table 3 Key factors of the evaluation plan

Key factor	Plan
Initial sample	100 sets of 6 samples generated from LHS
Cost function	Four cost functions with varying complexity
Evaluation plan	Evolution for R-squared and a maximum error of AS and AS-C at different total sampling cost
Prediction metric	R-squared to measure overall accuracy; max error to measure local accuracy; area metric to measure the effect of total sampling cost
Comparisons between AS and AS-C	1) Details of a typical case, 2) median performance with different initial samples, and 3) variation with different initial samples

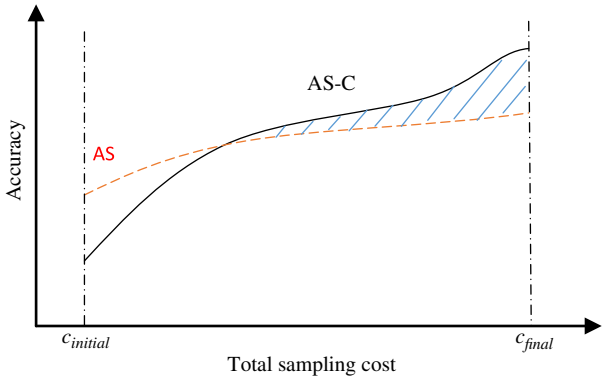

Fig. 9 An illustration of the evolution of accuracy for adaptive sampling with increasing total sampling cost.

B. Typical Case Using the AS and AS-C

A typical case of the AS and AS-C to approximate the Branin function and linear cost function is presented in this section. Six initial samples were generated from LHS. One additional sample was added per iteration until the total cost of samples was about to exceed 10. The initial samples and final samples are shown in Fig. 10 using AS and AS-C. The AS-C allocated more samples around the origin due to the cheap sampling cost. Details of AS and AS-C results are summarized in Table 4. AS ended with 16 samples and the total cost was 9.5, whereas AS-C ended with 18 samples and the total cost was 9.3. Adding one more sample at the location requested by the sampling algorithm would exceed the total sampling budget of 10. However, it is clear that cheaper samples can still be added.

Both AS and AS-C approximated the global response well (R^2 were 0.96 and 0.99, respectively), with AS-C being more accurate. The difference between $e_{\max}$ was significant (0.1 vs 0.05) considering the function value varies between [0.1, 1]. Evolution of prediction accuracy with respect to the total sampling cost is shown in Fig. 11. A_{R^2}

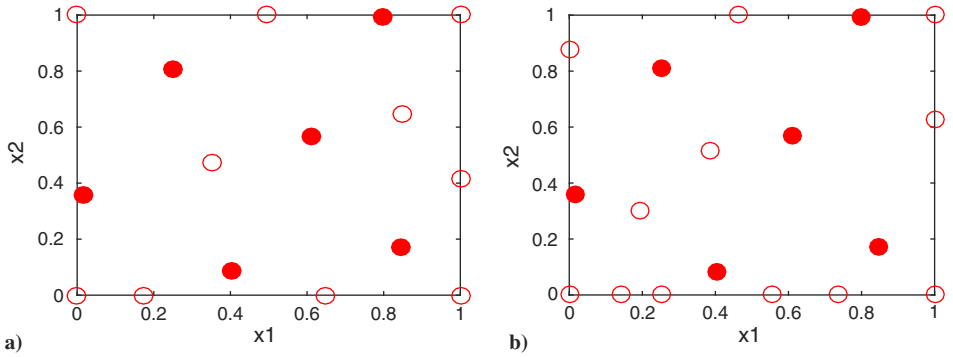

Fig. 10 Final samples for a typical DOE the approximation of Branin function with linear cost function using a) AS and b) AS-C. The solid circles denote the initial samples.

Table 4 Details for a typical DOE of AS and ASC for the approximation of Branin function and a linear cost function with up to 10 total cost

	R^2	$e_{\max}$	A_{R^2}	$A_{\max}$	Total sampling cost	Number of samples
AS	0.96	0.1	99.7%	100%	9.5	16
AS-C	0.99	0.05			9.3	18

was 99.7% and $A_{\max}$ was 100%; that is, AS-C was almost always better than AS. The performance of AS-C oscillated at the beginning in Fig. 11. This was mainly due to the large uncertainty of the Branin surrogate from scarce samples. The cost function was estimated accurately even with the initial samples.

C. Different Cost Functions and Different Initial Samples

The performance of AS heavily depends on the quality of surrogate and initial samples. As mentioned before, when the cost function is complicated, AS-C might suffer from uncertainty in the cost model too. AS and AS-C were evaluated using the four cost functions, as shown in Fig. 8. To consider the initial sampling uncertainty, AS and AS-C were repeated 100 times with different initial samples. Therefore, 100 sets of A_{R^2} and $A_{\max}$ were generated. The median value of them is used to indicate the average performance of AS and AS-C.

When A_{R^2} or $A_{\max}$ is equal to 50%, this indicates similar performance between AS and AS-C with different total sampling cost. A_{R^2} and $A_{\max}$ larger than 50% indicate the higher accuracy of AS-C over AS-C. As seen in Table 5, A_{R^2} was close to 100 % for all four cost functions, and therefore AS-C proved to be accurate overwhelmingly. The lowest A_{R^2} came from the Damper cost function that has a complicated response and leads to a large uncertainty of estimation. The values of $A_{\max}$ were also high and indicated preference of AS-C over AS. Similarly, the $A_{\max}$ was lower for Rosenbrock cost function and Damper cost function as they have complicated responses.

Table 5 Median value of area metrics A_{R^2} and $A_{\max}$ (in %) using 100 sets of initial samples

Area metric	$C_{\text{Linear}}(\boldsymbol{x})$	$C_{\text{Exp}}(\boldsymbol{x})$	$C_{\text{Rosen.}}(\boldsymbol{x})$	$C_{\text{Damper}}(\boldsymbol{x})$
A_{R^2}	93.2	96.9	93.4	89.9
$A_{\max}$	98.3	93.9	77.8	84.9

Table 6 Median values of total samples for AS and AS-C using 100 sets of initial samples

Sampling scheme	$C_{\text{Linear}}(\boldsymbol{x})$	$C_{\text{Exp}}(\boldsymbol{x})$	$C_{\text{Rosen.}}(\boldsymbol{x})$	$C_{\text{Damper}}(\boldsymbol{x})$
AS	16.0	25.0	36.0	27.0
AS-C	19.0	31.0	44.0	33.5

The collection of A_{R^2} and $A_{\max}$ was visualized using boxplot in Fig. 6. Almost all the values of A_{R^2} and $A_{\max}$ were larger than 50%; that is, AS-C was preferred for the approximation of the Branin function. The boxplot also reveals significant variability of A_{R^2} and $A_{\max}$. The performance of AS might be more robust by increasing the number of initial samples or introducing other advanced schemes for the initial DOE.

AS-C might lead to more total samples than the AS, as discussed in Sec. III. The total number of samples was collected for the 100 sets of evaluations to validate this observation. Table 6 summarizes the median value of total samples for AS and AS-C from 100 repetitions. The AS-C resulted in more samples in general. The difference between the numbers of samples increased when more samples were used. Compared with AS, AS-C adopted three more samples using the linear cost function and eight more samples using the Damper cost function. The collection of the total number of samples was visualized using boxplot in Fig. 12. The AS-C lead to more samples than AS almost all the time.

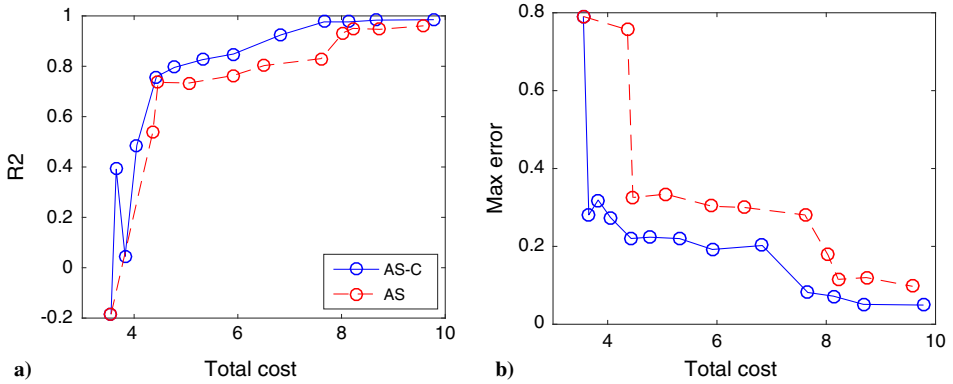

Fig. 11 Evolution of adaptive sampling with increasing samples for a typical DOE for the approximation of Branin function and linear cost function: a) *R*-squared; b) max error.

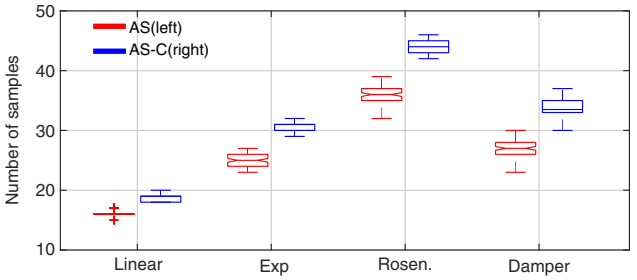

Fig. 12 The number of samples used for the approximation of the Branin function with four cost functions using AS and AS-C.

VI. Adaptive Sampling for Behavioral Emulation of Exascale Computing

At the PSAAP-II Center for Compressible Multiphase Turbulence (CCMT) at the University of Florida, CMT-nek [35], a large-scale parallel application to be run on future exascale systems, is being developed to perform a simulation with instabilities, turbulence, and mixing in particulate-laden flows under conditions of extreme pressure and temperature. CMT-nek is being developed from a production release of petascale code Nek5000 [36], a Gordon Bell prize–winning open-source software for simulating unsteady incompressible flow with thermal and passive scalar transport.

In parallel, CCMT is performing research on co-design of the CMT algorithms and the computer architecture on which they run. For that purpose, a coarse-grained simulation method called BE is used to predict the execution time of a representative skeleton app of CMT-nek (called CMT-bone-BE). Coarse-grained modeling in BE involves abstraction of the computation and communication operations in the application code. Each of these operations is modeled as an indivisible block with prebuilt performance estimate models or instrumented run data [28]. The BE simulation is performed in a discrete-event fashion, where the predicted execution of the simulated application on simulated hardware is obtained. These simulated results are then validated against results from running the actual application on actual hardware.

Because BE simulations can be computationally demanding, we developed a surrogate based on sample BE simulations. Because the cost/run time of BE for different input configurations varies by orders of magnitude, it serves as an excellent case study for AS-C.

A. Experimental Setup

Again, the application under study is BE simulation of CMT-nek [37]. There are three main application parameters of interest—element size (ES), the number of elements per processor (EPP), and the number of processors (NP). In CMT-nek, fluid flow is broken into small grids called elements. The ES ranges between 5 and 25. NP is the total number of ranks (i.e., number of processors) on which the MPI application code is run. The total number of elements divided by the NP is the number of EPP.

Application performance can be affected by changing any of these parameters. We chose 125 DOE based on a five-level, full-factorial design as shown in Fig. 13a. "Five-level" denotes the 5 points/grids selected along each parameter. The DOE is ES $= [5, 9, 13, 17, 21]$, EPP $= [8, 32, 64, 128, 256]$, and NP $= [16, 256, 2048, 16384, 131072]$. As a result, the experimental runs require up to 131,072 processors, 34 million elements, and 311 billion computational grid points. The collected BE simulations are summarized in Table A1 in the Appendix. The cost of BE simulations varies from 0.01143 to 859.0498 s.

BE simulation run time varies only with the increasing NP. Change in ES and EPP does not affect the simulation time of BE. This is mainly because the problem size of the simulation is the size of the system being simulated. This size is determined through NP. The bigger the simulated system is, the more time it takes to construct the system and then to simulate it on all cores. We accounted for this varying cost through eight runs of varying NP for a fixed number of time steps while keeping ES and EPP fixed. NP was varied from 16 to 131,072 processors as shown in Fig. 13b. The cost (runtime) varies significantly by orders of magnitude. The collected cost of BE simulations is summarized in Table A2 in the Appendix.

The outputs of BE simulations vary from 0.0143 to 110.2280 s. The NP and outputs (objectives) of the BE simulations were analyzed in the logarithmic coordinate due to the large variation.

B. Target Function and Cost Function of BE

Setting up and collecting the results from the BE simulations are time-consuming. For a comprehensive evaluation of AS for BE simulation, the target function and the cost function are represented by the polynomial response surface (PRS) for further tests. Then, the developed PRS is used to produce samples and associated cost for the evaluation of AS. The PRS with different orders were fitted to the 125 BE simulations using leave-one-out CV as shown in Table 7, where the root-mean-square error (RMSE) and maximum error of the CV were calculated. Cubic PRS provided the least error for both RMSE and maximum error. The residual errors fitting to all samples were also shown in Table 7. Considering that the range of BE simulations is 110.2137 s (110.228–0.0143), the cubic PRS was considered to be accurate with small residual errors (RMSE is 0.94; the maximum error is 3.26). Similarly, a cost function was also developed using PRS fitted only with varying NP. The accuracy of the PRS for cost was shown in Table 8. Linear PRS was selected with least CV errors in logarithmic coordinate as shown in Fig. 13b. The RMSE and maximum errors for residual errors decreased with increasing orders in logarithmic coordinate during fitting.

C. Approximation of BE with Adaptive Sampling

AS-C and AS were compared for the approximation of BE simulations with three selected input variables. The samples and associated cost were produced from the fitted PRS. Eight initial samples were generated using LHS. One additional sample was added per iteration. The AS schemes stopped when the budget (given

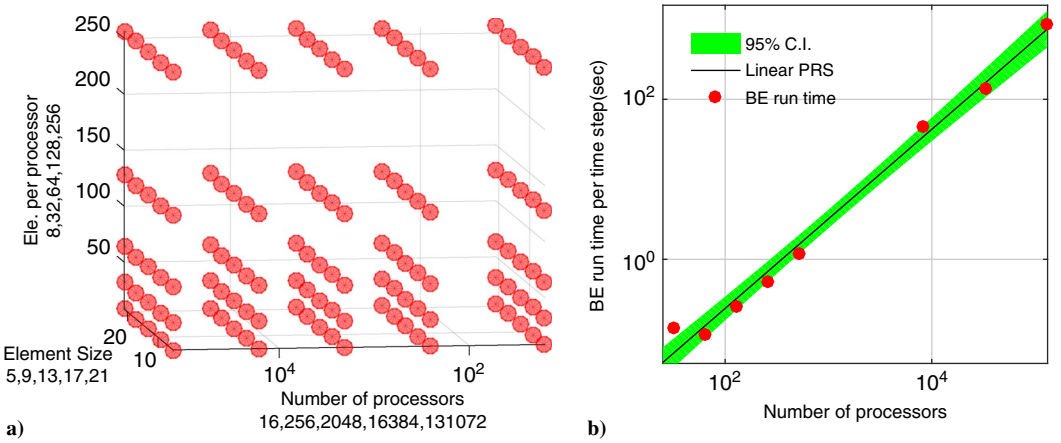

Fig. 13 Design of experiment for BE simulation and runtime of the BE. a) 125 grids for the BE; b) run time of BE with increasing number of processors.

Table 7 Accuracy of the PRS fitting to the BE predictions measured in linear coordinate

Order of PRS fitted in logarithmic coordinate	Leave-one-out cross-validation		Residual errors fitting to all samples	
	RMSE	Max	RMSE	Max
Linear	29.67	148.70	27.23	134.73
Quadratic	4.57	16.80	4.10	14.49
Cubic	1.11	4.15	0.94	3.26

Table 8 Accuracy of the PRS fitting to the runtime of BE measured in linear coordinate

Order of PRS fitted in logarithmic coordinate	Leave-one-out cross-validation		Residual errors fitting to all samples	
	RMSE	Max	RMSE	Max
Linear	70.02	193.89	39.10	106.98
Quadratic	160.84	453.14	34.37	93.16
Cubic	274.54	764.42	43.64	107.58

cost threshold) is met; the budget was set at 600 s for this evaluation. The RMSE and maximum error were adopted as the predictions metrics at the $20 \times 20 \times 20$ grids. Again, for sampling uncertainty, 100 sets of different initial samples were generated for the evaluation.

Typical samples using AS-C were shown in Fig. 14. The total cost for the initial 8 samples was 474 s. It is seen that the initial samples are scattered in the whole design space, whereas the additional samples are mostly allocated at small NP. The total cost for AS-C is 561 s. When AS tries to add one more sample at [ES, NP, EPP] = [21, 131070, 256], the cost becomes 752 s, exceeding the given budget.

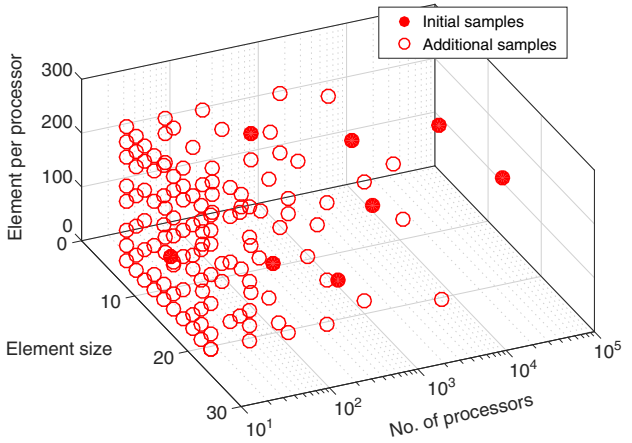

Fig. 14 The design of experiments using AS-C for the approximation of BE with three variables, ended with 130 samples.

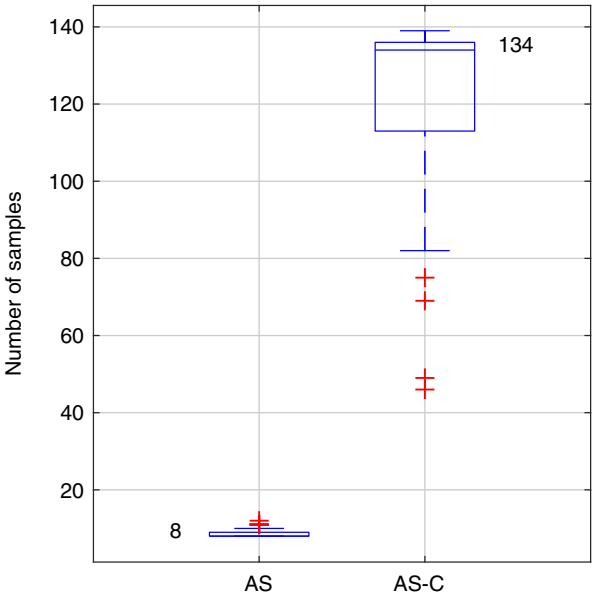

Fig. 16 The total number of samples used for AS and AS-C with 100 sets of initial samples.

The detailed prediction metrics for the 100 sets of evaluations are summarized in Fig. 15. The R^2 of AS-C in Fig. 15a concentrates around 1.00, whereas the R^2 of AS is around 0.56 and has much larger variation. AS-C resulted in much more accurate overall. The maximum error of AS-C is around 3.68 as shown in Fig. 15b. Considering that the range of BE simulations is 110.2137 s, AS-C is quite accurate and bested AS regarding the maximum error. The total number of samples used for AS and AS-C are shown in Fig. 16 with 100 sets of initial samples. AS-C resulted in cheaper runs than AS and lead to better accuracy to approximate the input variable space.

VII. Conclusions

This paper studies the effect of varying sampling cost on the DOE. A strategy for AS with varying sampling cost is proposed (AS-C). Two surrogates are developed to approximate the target function and cost function. A value metric is defined as the ratio between the standard deviation of prediction and the estimated cost. The value metric is used to indicate the uncertainty reduction per cost. A new sample is then added at the point with maximum value. The proposed AS-C method was evaluated using 1D and 2D algebraic functions and algebraic cost functions. The AS stops just before the total sampling cost exceeding the prescribed budget. Four different cost functions and 100 sets of different initial samples were produced to evaluate the AS-C. AS-C was compared with standard AS regarding R-squared, maximum error, and evolution with increasing total sampling budget. For the global

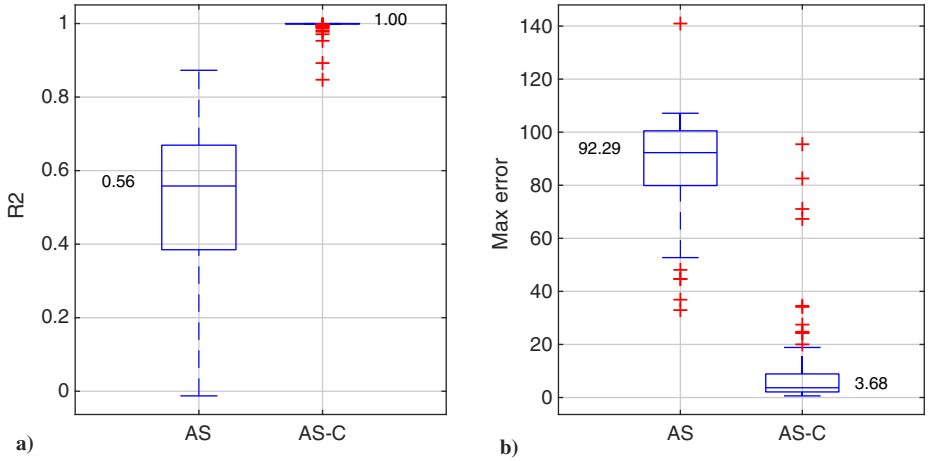

Fig. 15 The 100 sets of predictions developed using AS and AS-C with different initial samples for a) R2 and b) maximum error.

approximation of the test function in the whole input variable space, AS-C led to more samples than AS, and AS-C was more accurate than AS in the large majority of DOEs. As a case study, AS-C was applied for the design space exploration of BE. BE, a coarse-grained simulation method, predicts the run time of a given CFD simulation using a supercomputer. Because the cost/run time of BE varies by the order of magnitude, AS-C resulted in many more samples and greatly outperformed AS for a given sampling budget.

This paper is an initial effort to incorporate varying sampling cost with DOE, the proposed scheme adopted GP for design space exploration that is based on space filling. The proposed AS-C allocated more cheap runs instead of a few expensive runs and improved the global accuracy noticeably for the tested cases. There are alternatives for uncertainty estimations (e.g., universal prediction distribution) and similar schemes could be developed for these alternatives.

Appendix: Behavioral Emulation Predictions and Costs

Table A1 125 runs for the predictions of behavioral emulations (in seconds) for the shock simulation per time step

Index	Element size	Number of processors	Element per processor	BE prediction
1	5	16	8	0.0143
2	5	256	8	0.0143
3	5	2,048	8	0.0143
4	5	16,384	8	0.0144
5	5	131,072	8	0.0144
6	9	16	8	0.1178
7	9	256	8	0.1182
8	9	2,048	8	0.1183
9	9	16,384	8	0.1184
10	9	131,072	8	0.1184
11	13	16	8	0.5417
12	13	256	8	0.5427
13	13	2,048	8	0.5429
14	13	16,384	8	0.5429
15	13	131,072	8	0.5429
16	17	16	8	1.6277
17	17	256	8	1.6304
18	17	2,048	8	1.6307
19	17	16,384	8	1.6309
20	17	131,072	8	1.6308
21	21	16	8	3.4460
22	21	256	8	3.4497
23	21	2,048	8	3.4505
24	21	16,384	8	3.4509
25	21	131,072	8	3.4507
26	5	16	32	0.0559
27	5	256	32	0.0559
28	5	2,048	32	0.0560
29	5	16,384	32	0.0560
30	5	131,072	32	0.0560
31	9	16	32	0.4742
32	9	256	32	0.4778
33	9	2,048	32	0.4780
34	9	16,384	32	0.4781
35	9	131,072	32	0.4781
36	13	16	32	2.1741
37	13	256	32	2.1789
38	13	2,048	32	2.1793
39	13	16,384	32	2.1794
40	13	131,072	32	2.1794
41	17	16	32	6.5119
42	17	256	32	6.4713
43	17	2,048	32	6.5206
44	17	16,384	32	6.5209
45	17	131,072	32	6.5208
46	21	16	32	13.7869
47	21	256	32	13.7997
48	21	2,048	32	13.8010
49	21	16,384	32	13.8012
50	21	131,072	32	13.8011
51	5	16	64	0.1102
52	5	256	64	0.1107
53	5	2,048	64	0.1109
54	5	16,384	64	0.1109
55	5	131,072	64	0.1109
56	9	16	64	0.9498
57	9	256	64	0.9497
58	9	2,048	64	0.9501
59	9	16,384	64	0.9502
60	9	131,072	64	0.9502
61	13	16	64	4.3940
62	13	256	64	4.3991
63	13	2,048	64	4.3999
64	13	16,384	64	4.4000
65	13	131,072	64	4.4000
66	17	16	64	12.5411
67	17	256	64	12.5412
68	17	2,048	64	12.5426
69	17	16,384	64	12.5427
70	17	131,072	64	12.5431
71	21	16	64	27.5157
72	21	256	64	27.5315
73	21	2,048	64	27.5333
74	21	16,384	64	27.5336
75	21	131,072	64	27.5336
76	5	16	128	0.2258
77	5	256	128	0.2271
78	5	2,048	128	0.2272
79	5	16,384	128	0.2273
80	5	131,072	128	0.2273
81	9	16	128	1.8962
82	9	256	128	1.9005
83	9	2,048	128	1.9009
84	9	16,384	128	1.9010
85	9	131,072	128	1.9010
86	13	16	128	8.6678
87	13	256	128	8.6786
88	13	2,048	128	8.6796
89	13	16,384	128	8.6798
90	13	131,072	128	8.6798
91	17	16	128	25.9900
92	17	256	128	26.0086
93	17	2,048	128	26.0101
94	17	16,384	128	26.0104
95	17	131,072	128	26.0103
96	21	16	128	55.0990
97	21	256	128	55.1290
98	21	2,048	128	55.1317
99	21	16,384	128	55.1321
100	21	131,072	128	55.1322
101	5	16	256	0.4500
102	5	256	256	0.4519
103	5	2,048	256	0.4522
104	5	16,384	256	0.4523
105	5	131,072	256	0.4523
106	9	16	256	3.7727
107	9	256	256	3.7727
108	9	2,048	256	3.7802
109	9	16,384	256	3.7803
110	9	131,072	256	3.7803
111	13	16	256	17.2886
112	13	256	256	17.3051
113	13	2,048	256	17.3065
114	13	16,384	256	17.3067
115	13	131,072	256	17.3067
116	17	16	256	51.9621
117	17	256	256	51.9932
118	17	2,048	256	51.9959
119	17	16,384	256	51.9962
120	17	131,072	256	51.9962
121	21	16	256	110.1680
122	21	256	256	110.2240
123	21	2,048	256	110.2260
124	21	16,384	256	110.2280
125	21	131,072	256	110.2270

Table A2 Runtime of behavioral emulations (in seconds) for the shock simulation per time step

No. of Processors	32	64	128	256	512	8,192	32,768	131,072
Runtime per time step	0.1372	0.1143	0.2593	0.5229	1.1964	46.6024	134.4078	859.0498

Acknowledgments

This work was supported in part by the U.S. Department of Energy, National Nuclear Security Administration, Advanced Simulation and Computing Program, as a Cooperative Agreement under the Predictive Science Academic Alliance Program, under Contract No. DE-NA0002378.

References

[1] Queipo, N. V., Haftka, R. T., Shyy, W., Goel, T., Vaidyanathan, R., and Tucker, P. K., "Surrogate-Based Analysis and Optimization," *Progress in Aerospace Sciences*, Vol. 41, No. 1, 2005, pp. 1–28. doi:10.1016/j.paerosci.2005.02.001

[2] Wang, G. G., and Shan, S., "Review of Metamodeling Techniques in Support of Engineering Design Optimization," *Journal of Mechanical Design*, Vol. 129, No. 4, 2007, pp. 370–380. doi:10.1115/1.2429697

[3] Booker, A. J., Dennis, J. E., Frank, P. D., Serafini, D. B., Torczon, V., and Trosset, M. W., "A Rigorous Framework for Optimization of Expensive Functions by Surrogates," *Structural Optimization*, Vol. 17, No. 1, 1999, pp. 1–13. doi:10.1007/BF01197708

[4] Zhang, Y., Meeker, J., Schutte, J., Kim, N., and Haftka, R., "On Approaches to Combine Experimental Strength and Simulation with Application to Open-Hole-Tension Configuration," *Proceedings of the American Society for Composites: Thirty-First Technical Conference*, DEStech Publications, Inc., Lancaster, PA.

[5] Zhang, Y., Kim, N.-H., Park, C., and Haftka, R. T., "Multi-Fidelity Surrogate Based on Single Linear Regression," *AIAA Journal*, Vol. 56, No. 12, 2018, pp. 4944–4952. doi:10.2514/1.J057299

[6] Fernández-Godino, M. G., Park, C., Kim, N.-H., and Haftka, R. T., "Review of Multi-Fidelity Models," 2016, arXiv preprint arXiv: 1609.07196.

[7] Alexandrov, N. M., Lewis, R. M., Gumbert, C. R., Green, L. L., and Newman, P. A., "Approximation and Model Management in Aerodynamic Optimization with Variable-Fidelity Models," *Journal of Aircraft*, Vol. 38, No. 6, 2001, pp. 1093–1101. doi:10.2514/2.2877

[8] Chaudhuri, A., Lam, R., and Willcox, K., "Multifidelity Uncertainty Propagation via Adaptive Surrogates in Coupled Multidisciplinary Systems," *AIAA Journal*, Vol. 56, No. 1, 2018, pp. 235–249.

[9] Haftka, R. T., Villanueva, D., and Chaudhuri, A., "Parallel Surrogate-Assisted Global Optimization with Expensive Functions—A Survey," *Structural and Multidisciplinary Optimization*, Vol. 54, No. 1, 2016, pp. 3–13. doi:10.1007/s00158-016-1432-3

[10] Bui-Thanh, T., Willcox, K., and Ghattas, O., "Model Reduction for Large-Scale Systems with High-Dimensional Parametric Input Space," *SIAM Journal on Scientific Computing*, Vol. 30, No. 6, 2008, pp. 3270–3288. doi:10.1137/070694855

[11] Guénot, M., Lepot, I., Sainvitu, C., Goblet, J., and Filomeno Coelho, R., "Adaptive Sampling Strategies for Non-Intrusive POD-Based Surrogates," *Engineering Computations*, Vol. 30, No. 4, 2013, pp. 521–547. doi:10.1108/02644401311329352

[12] Hao, P., Wang, B., and Li, G., "Surrogate-Based Optimum Design for Stiffened Shells with Adaptive Sampling," *AIAA Journal*, Vol. 50, No. 11, 2012, pp. 2389–2407. doi:10.2514/1.J051522

[13] Chen, Z., Qiu, H., Gao, L., Li, X., and Li, P., "A Local Adaptive Sampling Method for Reliability-Based Design Optimization Using Kriging Model," *Structural and Multidisciplinary Optimization*, Vol. 49, No. 3, 2014, pp. 401–416. doi:10.1007/s00158-013-0988-4

[14] Mackman, T., Allen, C., Ghoreyshi, M., and Badcock, K., "Comparison of Adaptive Sampling Methods for Generation of Surrogate Aerodynamic Models," *AIAA Journal*, Vol. 51, No. 4, 2013, pp. 797–808. doi:10.2514/1.J051607

[15] Romero, D. A., Amon, C. H., and Finger, S., "On Adaptive Sampling for Single and Multi-Response Bayesian Surrogate Models," *Proceedings of the ASME 2006 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, American Soc. of Mechanical Engineers Paper DETC2006-99210, 2006, pp. 393–404.

[16] Gorissen, D., Couckuyt, I., Demeester, P., Dhaene, T., and Crombecq, K., "A Surrogate Modeling and Adaptive Sampling Toolbox for Computer Based Design," *Journal of Machine Learning Research*, Vol. 11, July 2010, pp. 2051–2055.

[17] Ben Salem, M., Roustant, O., Gamboa, F., and Tomaso, L., "Universal Prediction Distribution for Surrogate Models," *SIAM/ASA Journal on Uncertainty Quantification*, Vol. 5, No. 1, 2017, pp. 1086–1109. doi:10.1137/15M1053529

[18] Rasmussen, C. E., and Williams, C. K. I., *Gaussian Processes for Machine Learning*, The MIT Press, MA, 2006.

[19] Viana, F. A., *Multiple Surrogates for Prediction and Optimization*, Univ. of Florida, Gainesville, FL, 2011.

[20] Viana, F. A., Haftka, R. T., and Watson, L. T., "Efficient Global Optimization Algorithm Assisted by Multiple Surrogate Techniques," *Journal of Global Optimization*, Vol. 56, No. 2, 2013, pp. 669–689. doi:10.1007/s10898-012-9892-5

[21] Jin, R., Chen, W., and Sudjianto, A., "On Sequential Sampling for Global Metamodeling in Engineering Design," *Proceedings of ASME 2002 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, American Soc. of Mechanical Engineers Paper DETC2002/DAC-34092, 2002, pp. 539–548.

[22] Salem, M. B., Roustant, O., Gamboa, F., and Tomaso, L., "Universal Prediction Distribution for Surrogate Models," *SIAM/ASA Journal Uncertainty Quantification*, Vol. 5, No. 1, 2017, pp. 1086–1109.

[23] Xu, S., Liu, H., Wang, X., and Jiang, X., "A Robust Error-Pursuing Sequential Sampling Approach for Global Metamodeling Based on Voronoi Diagram and Cross Validation," *Journal of Mechanical Design*, Vol. 136, No. 7, 2014, Paper 071009. doi:10.1115/1.4027161

[24] Yildirim, E. A., and Wright, S. J., "Warm-Start Strategies in Interior-Point Methods for Linear Programming," *SIAM Journal on Optimization*, Vol. 12, No. 3, 2002, pp. 782–810. doi:10.1137/S1052623400369235

[25] Rudolph, D., and Stitt, G., "An Interpolation-Based Approach to Multi-Parameter Performance Modeling for Heterogeneous Systems," *2015 IEEE 26th International Conference on Proceedings of Application-Specific Systems, Architectures and Processors (ASAP)*, IEEE, Curran Associates, Inc., NY, pp. 174–180.

[26] Banerjee, T., Gadou, M., and Ranka, S., "A Genetic Algorithm Based Approach for Multi-Objective Hardware/Software Co-Optimization," *Sustainable Computing: Informatics and Systems*, Vol. 10, June 2016, pp. 36–47.

[27] Barrett, R. F., Borkar, S., Dosanjh, S. S., Hammond, S. D., Heroux, M. A., Hu, X. S., Luitjens, J., Parker, S. G., Shalf, J., Tang, L., et al., "On the Role of Co-Design in High Performance Computing," *Advances in Parallel Computing*, Vol. 24, Jan. 2013, pp. 141–155.

[28] Kumar, N., Pascoe, C., Hajas, C., Lam, H., Stitt, G., and George, A., "Behavioral Emulation for Scalable Design-Space Exploration of Algorithms and Architectures," *Proceedings of International Conference on High Performance Computing*, Springer, Switzerland, pp. 5–17.

[29] Jones, D. R., Schonlau, M., and Welch, W. J., "Efficient Global Optimization of Expensive Black-Box Functions," *Journal of Global Optimization*, Vol. 13, No. 4, 1998, pp. 455–492. doi:10.1023/A:1008306431147

[30] Chaudhuri, A., and Haftka, R. T., "Efficient Global Optimization with Adaptive Target Setting," *AIAA Journal*, Vol. 52, No. 7, 2014, pp. 1573–1578. doi:10.2514/1.J052930

[31] Chaudhuri, A., "Implications of Optimization Cost for Balancing Exploration and Exploitation in Global Search and for Experimental Optimization," Univ. of Florida, Gainesville, FL, 2014.

[32] Viana, F., "SURROGATES Toolbox User's Guide," Univ. of Florida, Gainesville, FL, 2010.

[33] Forrester, J., Keane, A. J., and Bressloff, N. W., "Design and Analysis of Noisy Computer Experiments," *AIAA Journal*, Vol. 44, No. 10, 2006, pp. 2331–2339. doi:10.2514/1.20068

[34] Ramu, P., Kim, N. H., Haftka, R. T., and Queipo, N. V., "System Reliability Analysis Using Tail Modeling," *Proceedings of Proceedings of the 11th AIAA/ISSMO Multidisciplinary Analysis & Optimization Conference*, Portsmouth, VA, Sept. 2006.

[35] "Center for Compressible Multiphase Turbulence Webpage," 2015, https://www.eng.ufl.edu/ccmt/.

[36] Fischer, P. F., Lottes, J. W., and Kerkemeier, S. G., "nek5000 Web page," 2008, http://nek5000.mcs.anl.gov.

[37] Zhang, Y., Neelakantan, A., Kumar, N., Park, C., Haftka, R. T., Kim, N. H., and Lam, H., "Multi-Fidelity Surrogate Modeling for Application/Architecture Co-Design," *Proceedings of 8th International Workshop on Performance Modeling, Benchmarking and Simulation of High Performance Computer Systems*, Springer, Switzerland, pp. 179–196.

K. E. Willcox
Associate Editor