AIAA SciTech Forum
23-27 January 2023, National Harbor, MD & Online
AIAA SCITECH 2023 Forum

10.2514/6.2023-2203

Multi-Objective Bayesian Optimization using an Active Subspace-based Approach

Danial Khatamsaz*, Douglas Allaire†
J. Mike Walker '66 Department of Mechanical Engineering, Texas A&M University, College Station, TX, USA

Multi-objective optimization is often a difficult task owing to the need to balance competing objectives. Solution to a multi-objective optimization problem is represented by a set of non-dominated designs that form Pareto frontier in objective space. Quality approximation of Pareto frontier requires performing thousands of experiments that makes the process computationally demanding. In this work, we propose two active subspace-based Bayesian optimization methods to handle expensive multi-objective design tasks. The Active Subspace Method (ASM) enables recognition of highly informative lower dimensional spaces in design space in terms of objective function variability. Defining a problem on lower dimensional spaces ease learning process and enables better approximations of Pareto frontier at lower computational expenses. The approaches are demonstrated on different real-world engineering problems.

I. Introduction

Many design problems in engineering involve optimization of expensive black-box objective functions with unknown characteristics. Often, limited computational resources necessitate the use of efficient algorithms that can both explore and exploit the unknown design space. In such circumstances, Bayesian optimization (BO) algorithms [1, 2] have been effectively used and offer significant improvements in efficiency in comparison to other design algorithms [3, 4]. A Bayesian optimization algorithm works with minimal data, incorporates prior knowledge about an objective function, calculates the posterior distribution and looks for the potentially best next experiment to iteratively improve the current best design and increase the potential to improve it further on the next iteration. A Bayesian optimization framework uses surrogate models, usually Gaussian processes (GPs) [5] to model underlying objective functions. An acquisition, or utility function is also used to search for the best decision to make at every iteration to maximize the improvement of system's estimation of the optimal design.

A concern with any optimization algorithms, including Bayesian optimization frameworks is loss of efficiency as the dimensionality of the design space increases. As the number of design variables goes up, searching the space gets more challenging and numerical aspects of many algorithms can deteriorate, adversely affecting the learning process. Although Bayesian optimization frameworks are known to work with minimal data, their performance diminishes in high-dimensional design spaces. To address the issue, some dimensionality reduction techniques have been proposed, for example, subspace approximation techniques [6] are introduced to project high-dimensional design problems onto a lower dimensional space. This in fact can ease the learning process of an objective function and increase the efficiency of a design framework [7–9]. The active subspace method (ASM) is one of subspace approximation techniques introduced in Refs. [10–12]. The idea behind this technique is to define an objective function on a lower dimensional space such that the highest amount of information about the objective function is preserved. In other words, the active subspace is a subspace of the high-dimensional design space which an objective function varies the most. The active subspace is formed using directions that an objective function has the largest variability and those directions that the objective function is less sensitive to are discarded. Once the active subspace is formed, optimization can be performed over this lower dimensional space to avoid wasting resource to search less important regions of the design space.

The key difference between the active subspace method and some other dimensionality reduction techniques, like, principal component analysis (PCA) [13], is that the active subspace method aims to preserve the most information about the objective function in a lower dimensional space to assist the optimization process. In contrast, techniques such as PCA seek to keep as much information as possible about the samples only in the input space, which is not necessarily helping with finding the optimal design and actually, it may result in loss of information about the objective function. The active subspace method has been used in works such as design optimization [14–16], shape optimization [17–19], and uncertainty quantification [20, 21]. Recently, a multifidelity Bayesian optimization framework employing this

*Ph.D. Candidate, khatamsaz@tamu.edu, AIAA Student Member.
†Associate Professor, dallaire@tamu.edu, AIAA Senior Member.

technique has been proposed in Ref. [22] where they have shown how to apply the active subspace method in presence of different models with varying active subspaces representing the same quantity of interest. However, in all of these works, single objective functions are being optimized. To the best of our knowledge, there is no work done toward employing a subspace approximation technique to optimize multi-objective functions.

Some studies have proposed other dimensionality reduction techniques to solve multi-objective optimization problems. In Refs. [23, 24], they introduce algorithms to divide design variable into diversity related clusters and deal with sub-groups of variables one by one, considering the linkage between design variables. In Ref [25], a strategy is proposed that take into account the mapping relationships between design variables and objective functions. They divide the decision space into several subspaces based on the obtained relationships and take local search strategies. Another work presented in Ref. [26] clusters the design variables into convergence related and diversity related groups. Then, principal component analysis is used to reduce the dimensionality of the those design variables affecting the convergence of the evolution. Our contribution in this study is to propose approaches to allow use of active subspace method as a subspace approximation technique in Bayesian optimization frameworks to handle multi-objective optimization problems. The challenge here is that every objective function has a different active subspace and varies differently with respect to the change in design variables. Basically, we aim to explore different active subspaces and count for their impact on improving the solution estimation of multi-objective design problems. Two different approaches are introduced in this work and we investigate how using the active subspace method can improve the performance of multi-objective Bayesian optimization frameworks.

The rest of the paper proceeds as follows. First, background is provided to discuss each ingredient of the proposed framework. Gaussian process regression is introduced as surrogates to model objective functions, followed by presenting a discussion on multi-objective optimization and available techniques to solve such design problems. The active subspace method is introduced as the last part of the background section. We then describe proposed approaches to implement the active subspace method within a multi-objective Bayesian optimization framework. Finally, we apply our proposed approaches on some real-world design problems to examine the performance of each approach. We then provide concluding remarks and discuss avenues of future work.

II. Background

A multi-objective optimization problem is defined as

$$\text{minimize} \ \ \{f_1(\mathbf{x}), ..., f_n(\mathbf{x})\}, \mathbf{x} \in \mathcal{X} \tag{1}$$

where $f_1(\mathbf{x}), \ldots, f_n(\mathbf{x})$ are the objectives, usually in form of black-box functions, and $\mathcal{X}$ is the feasible design space. Although we develop an unconstrained approach, but in can be extended to a constrained approach by considering penalty terms for constraint handling. For instance, in some of the applications presented in this work, a new objective is introduced to encapsulates all constraint's violation penalty terms.

A. Gaussian Process Regression

A Gaussian process (GP) conditions a probabilistic model to training data to represent an objective function as a stochastic process. GPs allow for estimating an underlying function using previously evaluated samples from a model at negligible computational cost. The popularity of GPs for modeling purposes in engineering applications stems from their flexibility and easy manipulation characteristics. A GP maps every point in the input space to a normal distribution in the objective space defined by mean and covariance functions [5]. Assume there are N previously evaluated samples from an objective function that forms our prior knowledge about that objective function denoted by $\{\mathbf{X}_N, \mathbf{y}_N\}$, where $\mathbf{X}_N = (\mathbf{x}_1, \ldots, \mathbf{x}_N)$ are N input designs and $\mathbf{y}_N = (f(\mathbf{x}_1), \ldots, f(\mathbf{x}_N))$ are the corresponding objective values. Then, the posterior distribution of the objective function at any input location $\mathbf{x}$ is given as

$$f_{\text{GP}}(\mathbf{x}) \mid \mathbf{X}_N, \mathbf{y}_N \sim \mathcal{N}\left(\mu(\mathbf{x}), \sigma^2_{\text{GP}}(\mathbf{x})\right) \tag{2}$$

where

$$\begin{aligned} \mu(\mathbf{x}) &= K(\mathbf{X}_N, \mathbf{x})^T [K(\mathbf{X}_N, \mathbf{X}_N) + \sigma_n^2 I]^{-1} \mathbf{y}_N \\ \sigma^2_{\text{GP}}(\mathbf{x}) &= k(\mathbf{x}, \mathbf{x}) - K(\mathbf{X}_N, \mathbf{x})^T \\ &\qquad [K(\mathbf{X}_N, \mathbf{X}_N) + \sigma_n^2 I]^{-1} K(\mathbf{X}_N, \mathbf{x}) \end{aligned} \tag{3}$$

with k as a real-valued kernel function over the input space. Then, $K(\mathbf{X}_N, \mathbf{X}_N)$ is a $N \times N$ matrix with m, n entry as $k(\mathbf{x}_m, \mathbf{x}_n)$, and $K(\mathbf{X}_N, \mathbf{x})$ is a $N \times 1$ vector with m^{th} entry as $k(\mathbf{x}_m, \mathbf{x})$. The term σ_n^2, incorporates model observation error based on experiments or expert's opinion. Among different choices for the kernel function, we use the squared exponential function:

$$k(\mathbf{x}, \mathbf{x}') = \sigma_s^2 \exp\left(-\sum_{h=1}^{d} \frac{(x_h - x'_h)^2}{2l_h^2}\right) \tag{4}$$

where d is the dimensionality of the input space, σ_s^2 is the signal variance, and l_h, where $h = 1, 2, \ldots, d$, is the characteristic length-scale that determine the degree of correlation between samples within the dimension h of the input space. A maximum likelihood method [5] may be employed to obtain the optimum hyperparameters for a GP, but as the dimensionality of the space increases, there are chances that these methods fail since there are not sufficient data available. Therefore, expert's opinion can be a better choice in such cases.

B. Multi-Objective Optimization

In multi-objective design problems, it is usually the case that some objectives are in contradiction with each other and there is no single solution optimizing all objectives simultaneously. To manage such design problems, one approach is to create a single scalar objective, for example, weighted sum of all objective values, to transform the problem in to a single objective optimization problem. Another approach is to discover non-dominated solutions converging to the Pareto frontier. In this work, we take the latter technique and try to find the best approximation of the Pareto frontier. Here, non-dominated designs, $\mathbf{y}$, to a n-objective design problem are given as

$$\{\mathbf{y} : \mathbf{y} = (y_1, y_2, \ldots, y_n),\ y_i \le y'_i\ \ \forall\, i \in \{1, 2, \ldots, n\},\quad \exists\, j \in \{1, 2, \ldots, n\} : y_j < y'_j\} \tag{5}$$

where $\mathbf{y'} = (y'_1, y'_2, \ldots, y'_n)$ denotes possible objective outputs. The set of $\mathbf{y}$ is the Pareto front of the problem. Figure. 1 illustrates the Pareto frontier of a 2-objective design problem if the goal is to minimization of both objectives.

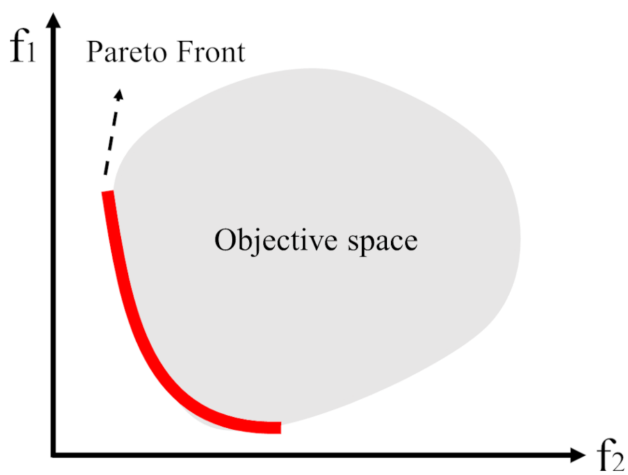

Fig 1. All points on the red line are non-dominated and constitute the solution set.

There are several methods to approximate Pareto frontiers of multi-objective optimization problems such as the weighted sum approach [27], the adaptive weighted sum approach [28], normal boundary intersection methods [29], hypervolume indicator methods [30–36], and others. In this work, we employ a hypervolume indicator approach since it is a better fit for expected improvement-based algorithms. In hypervolume indicator approaches, the concept of hypervolume is introduced as the quality measure of Pareto front approximation. By defining a fixed reference point in the objective space, the hypervolume is computed as the volume between this reference point and approximated non-dominated designs. A larger hypervolume indicates a better approximation of Pareto front, thus, a hypervolume indicator algorithm seeks to maximize the hypervolume. Mathematically, the hypervolume and its uncertainty can be calculated using, for example, prior predictive distributions obtained from GPs for each objective value. In this work, we follow Ref. [37] for fast and exact computation of expected hypervolume improvement (EHVI) as the utility function in a Bayesian optimization framework to balance the exploration of the space versus exploitation of the current system's knowledge of optimum designs. Mathematical expressions and details about the computation of EHVI and the implementation in Bayesian optimization frameworks are discussed in Refs. [37–42].

C. The Active Subspace Method

Dimensionality reduction techniques are used to represent data in lower dimensional spaces, mainly to reduce the associated computational costs of working with high-dimensional data, for example, in machine learning objectives. The active subspace method (ASM) is one of such techniques that is used to approximate an objective function in lower dimensional spaces. In contrast to other dimensionality reduction techniques that seek to project data to lower dimensions such that they preserve the most information about the data, for instance, principal component analysis (PCA), the active subspace method uses mapping information from input to output space and considers information form output space as well. In other words, the dimensionality of the input space is reduced such that the highest amount of information on the output side is preserved. The contribution of objective space's information makes active subspace method well-suited for learning purposes in machine learning objectives [10, 22, 43]. Although using approaches such as PCA can ease the learning process by reducing the dimensionality of the input space to address the curse of dimensionality in some degree, it is usually the case that in machine learning design tasks, there are not enough observations initially available, plus, preserving the information in the input space is not necessary linked to the information needed for optimizing an objective function.

The active subspace is considered a highly informative lower dimensional space representing the objective function. Figure 2 illustrates an example of a 2-dimensional space and a possible active subspace which is essentially 1-dimensional in this case. By defining an objective function in a lower dimensional space, we speed up the design process significantly and reduce associated computational costs required for probing the space looking for the best next experiment in Bayesian optimization frameworks.

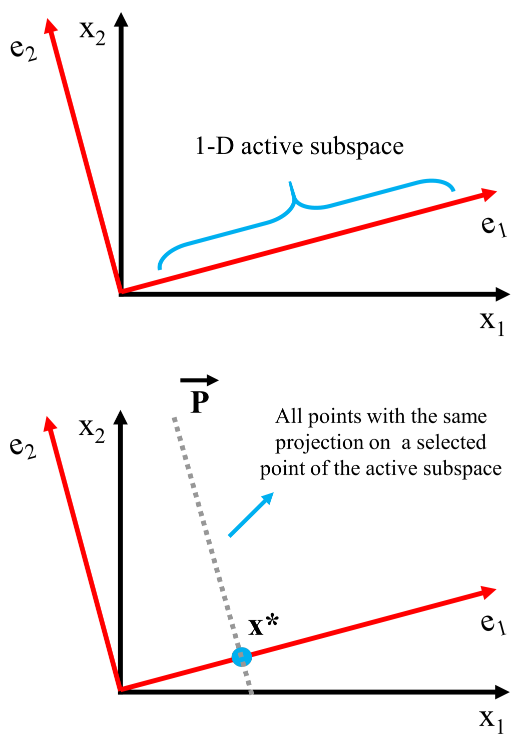

Fig 2. An example of determining the active subspace and inverse mapping problem solutions of a 2-dimensional design space.

Following the Refs. [11, 43, 44], assuming a scalar function f that takes m dimensional input $\mathbf{x}$ from the design space $\mathcal{X}$, $\nabla_x f$ represents the gradient of the objective function at location $\mathbf{x}$ in $\mathcal{X}$. First, we compute the covariance of the gradient, $\mathbf{C}$:

$$\mathbf{C} = \mathbb{E}[\nabla_{\mathbf{x}} f(\mathbf{x}) \nabla_{\mathbf{x}} f(\mathbf{x})^{\mathrm{T}}] \tag{6}$$

If the objective function has a black-box nature, as it does in many engineering design problems, the gradient can be approximated by Monte Carlo approaches. Assuming there are M samples available from previous function evaluations, the covariance matrix is formed as

$$\mathbf{C} \approx \frac{1}{M}\sum_{i=1}^{M} \nabla_{\mathbf{x}} f(\mathbf{x}_i) \nabla_{\mathbf{x}} f(\mathbf{x}_i)^{\mathrm{T}} \tag{7}$$

By computing the eigenvalues and eigenvectors of the covariance matrix, the effectiveness of directions defined by eigenvectors can be compared by the respective eigenvalues. Based on the eigenvalue decomposition, the covariance matrix can be written as

$$\mathbf{C} = \mathbf{W}\lambda\mathbf{W}^{\mathrm{T}} \tag{8}$$

where $\mathbf{W}$ is the matrix formed by eigenvectors and λ is a diagonal matrix of eigenvalues. To build an n-dimensional active subspace, the eigenvectors corresponding to the first n largest eigenvalues are selected.

$$\mathbf{W} = [\mathbf{U} \;\; \mathbf{V}], \quad \lambda = \begin{bmatrix} \lambda_1 & \\ & \lambda_2 \end{bmatrix} \tag{9}$$

The matrix $\mathbf{U}$ has n eigenvectors corresponding to the first n largest eigenvectors forming λ_1 and is called the transformation matrix. Any design point in $\mathcal{X}$ can be transformed to the active subspace using the transformation matrix:

$$\mathbf{z} = \mathbf{U}^{\mathrm{T}}\mathbf{x} \tag{10}$$

and the function g represents the original function f in the active subspace:

$$g(\mathbf{z}) = g(\mathbf{U}^{\mathrm{T}}\mathbf{x}) \approx f(\mathbf{x}) \tag{11}$$

Once the transformation matrix U is found, all evaluated design points from the objective function f are projected to the active subspace

$$\mathrm{Z}_N = \mathrm{U}^{\mathrm{T}}\mathrm{X}_N \tag{12}$$

and a new Gaussian process is built using the projected design points in the active subspace

$$g(\mathbf{z}) \mid \mathbf{Z}_N, \mathrm{y}_N \approx \mathcal{N}\left(\mu(\mathbf{z}), \sigma^2(\mathbf{z})\right) \tag{13}$$

where $g(\mathbf{z})$ is the posterior distribution of the objective function in the active subspace. Now, we seek to learn the objective function g in the active subspace instead of the original objective function f.

When using the active subspace method, Bayesian optimization should be completed twice at every iteration. The first step is applied by probing the active subspace to suggest the best next experiment, while the second step must be completed over the original input space. The trick here is, for the second step, we only search the subspace of the full high-dimensional space that has the same projection on the active subspace as the selected design in the first step. As shown in Fig. 2, the point $\mathbf{x}^*$ is selected in the active space, then the second step is to search the subspace defined by vector $\mathbf{P}$. Detailed discussion on how to find the complementary subspace for completing the second step and solve the mapping back problem is provided in Refs. [16, 22].

III. Approach

In this section, we introduce a framework to employ the active subspace method in optimization of multi-objective functions.

A challenge when employing the active subspace method in multi-objective design problems is that each objective function varies differently when moving toward different directions in the design space. Therefore, there are multiple active subspaces available depending on which objective value is used for calculations. Here, we aim to use all possible active subspaces and seek to maximize the acquisition function (EHVI in this work). Here, we propose 2 different approaches:

- **Approach 1**: suggest one best next experiment per objective function using its active subspace
- **Approach 2**: suggest one best next experiment chosen among the best next experiments selected in the first step of Bayesian optimization using the active subspace associated to each of the objective functions

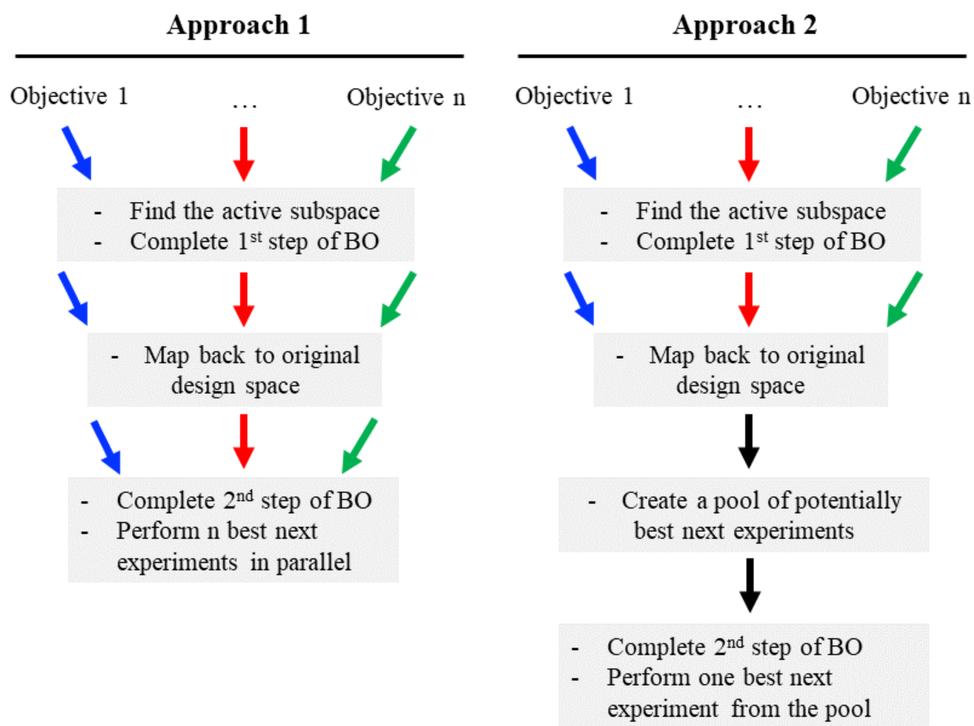

Fig 3. Illustration of both approaches to employ the active subspace method in multi-objective optimization framework.

Here, we explain each approach in more details.

The first approach accounts for improving the approximation of Pareto front with respect to all objective functions in turn. Basically, the optimization is done multiple times at every iteration using different active subspaces. Note that, we still use EHVI as the utility function to value each potential next experiment in both steps of Bayesian optimization, but we define the problem on different active subspaces to explore the impact of all objective functions on the hypervolume improvement. Taking this approach, it is possible to employ parallel computations, exploring all active subspaces at no additional simulation wall-time.

In other approach, the second step of Bayesian optimization is completed over a pool of best experiments created by storing all suggested best experiments in active subspace of each objective function. In other words, once the first step of Bayesian optimization is completed over an active subspace and the selected design is mapped back to the original design space, the resulted solution set is added to a pool. Then, the second step of Bayesian optimization searches this pool to suggest the best next experiment (best of best next experiments). This approach is still exploring all active subspaces but only one experiment is selected to be completed, regardless of the active subspace searched in the first steps of Bayesian optimization. Figure 3 is summarizing the steps in both approaches.

IV. Applications

To assess the performance of our framework, we have implemented it to solve a few real-world engineering design problems. Although there are available some synthetic test functions for testing multi-objective optimization approaches, the problem is, most often, these synthetic functions, such as DTLZ function family [45] or ZDT function family [46], have unusual properties and unrealistic characteristics that do not usually appear in engineering applications. Therefore, we believe real-world applications and available physical models are a better choice to examine and compare the performance of evolutionary multi-objective optimization algorithms. In the following, we present the results of implementing both suggested approaches and compare it to the case without employing the active subspace method to solve multiple real-world engineering design problems from Ref. [47], where they introduce several easy to implement design problems suitable for testing multi-objective optimization frameworks. Via representing each design problem by simple mathematical expressions, for example, defining parameters using response surface methods, it is possible to model different engineering applications without running specific simulation software or experiment. Readers are referred to [47] for details of the each design problem and formulation of objective functions and constraints.

All the results are averaged over 50 replications of a simulation.

A. Four bar truss design problem

This is a 2-objective design problem defined on a 4-dimensional design space. Design variables represent the length of four bars and the goal is to minimize the structural volume and the joint displacement of the four bar truss.

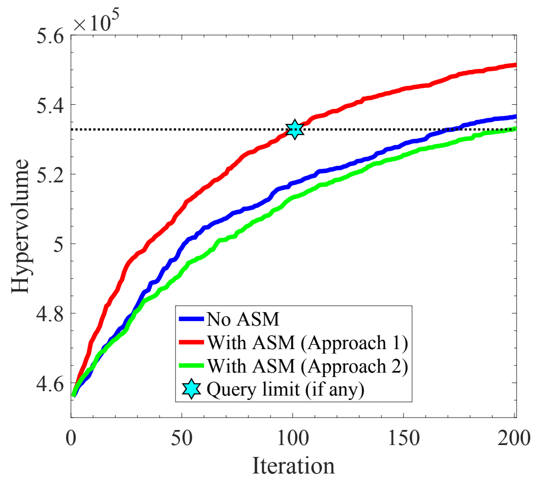

Fig 4. Four bar truss design problem. Hypervolume values averaged over 50 replications of a simulation for all approaches.

Figure 4 shows the hypervolume at different iterations. Note that the value of the hypervolume itself does not provide any specific information as it depends on the reference point selected for computing the hypervolume, however, the improvements are indications of finding better approximations of the Pareto front. Employing the first approach introduced in Sec. III, at any iterations, a better Pareto front is found in comparison to the case where no active subspace method is used. On the other hand, the second approach is not suggesting any improvements to the results. One argument here is, the first approach is making 1 evaluation per objective, so the total number of experiments is twice the other approaches. Although the evaluations can be made in parallel, without increasing the simulation time, the budget may have been defined over total number of allowed experiments and not the number of iterations. In this condition, by employing the first approach, all budget is exhausted at iteration 100, marked in Fig. 4. As evident, in this scenario, the final hypervolume is almost the same as the approach with no ASM, but the optimization time is halved. This points to the fact that taking advantage of the active subspace method is beneficial in the sense that it makes it possible to discover the optimal design region faster.

B. Reinforced concrete beam design problem

This is a constrained 3-dimensional design problem and the first objective is to minimize the total cost of concrete and reinforcing steel of the beam. To formulate the second objective, we use sum of all constraint violations and seek to minimize it.

In Fig. 5, it has been shown that taking the first approach is successfully discovering a significantly better approximation of the Pareto front as the hypervolume values are considerably larger. Regardless of the scenario defined to solve this design problem, either limiting the number of iterations or number of allowed evaluations, employing the active subspace method is suggesting better solutions in both proposed approaches in comparison to the case where the active subspace method tool is not utilized.

C. Hatch cover design problem

In this constrained 2-dimensional design problem, the first objective is to minimize the weight of the hatch cover and the second objective is to minimize constraint violations. Flange thickness and beam height of the hatch cover are the

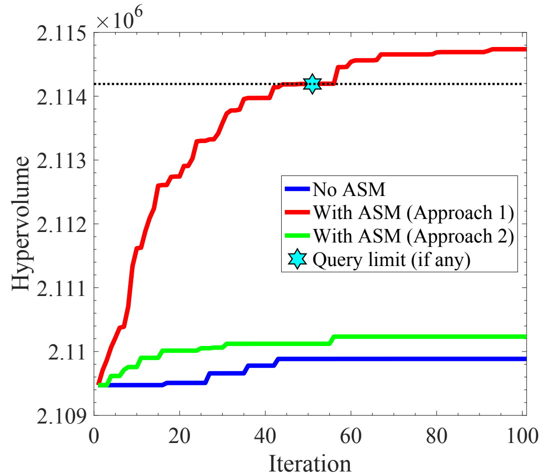

Fig 5. Reinforced concrete beam design problem. Hypervolume values averaged over 50 replications of a simulation for all approaches.

design variables.

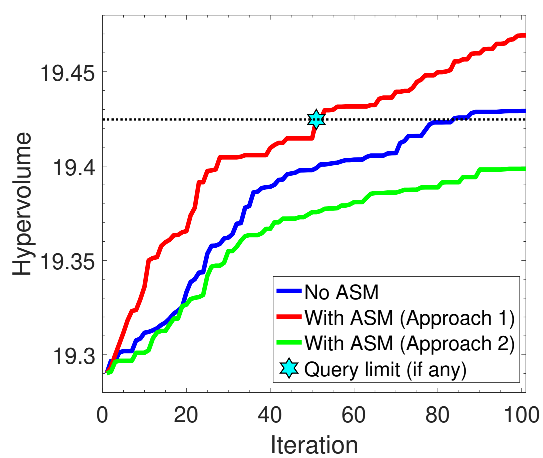

Fig 6. Hatch cover design problem. Hypervolume values averaged over 50 replications of a simulation for all approaches.

similarly, the first proposed approach is successfully discovering a better Pareto front in comparison to other approaches. In case of limitation over allowable number of experiment, it suggests a faster discovery of the optimal region. The second proposed approach however, is not suggesting any improvements when compared to no ASM approach. A possible issue with this approach can be if it is not exploring all active subspaces associated to objective functions. In other words, it is possible that in every iteration, the selected experiment is coming from exploring only one of the active subspaces, so that the potential improvements regarding the other active subspaces are missed.

D. Disc brake design problem

This is a constrained design problem defined over a 4-dimensional design space with 3 objectives to be minimized: mass of the brake and minimum stopping time. The third objective is formulated as sum of constraint violations. Design variables in this problem introduce inner and outer radii of the discs, engaging force and number of friction surfaces.

Similar to reinforced concrete beam design problem, both approaches taking advantage of the active subspace method are outperforming the Bayesian optimization approach with no ASM. In Fig. 7, it is shown that how quickly the

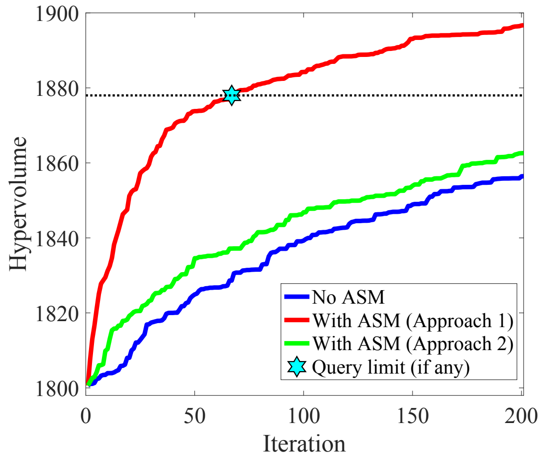

Fig 7. Disc brake design problem. Hypervolume values averaged over 50 replications of a simulation for all approaches.

first approach is discovering optimal non-dominated design regions. Significantly larger hypervolumes are achieved in only a few iterations.

E. The vehicle crashworthiness design problem

The objective in this 5-dimensional design problem is to minimize the weight, acceleration characteristics and toe-board instruction of the vehicle design. Design variables are the thickness of reinforced members of the frontal structure.

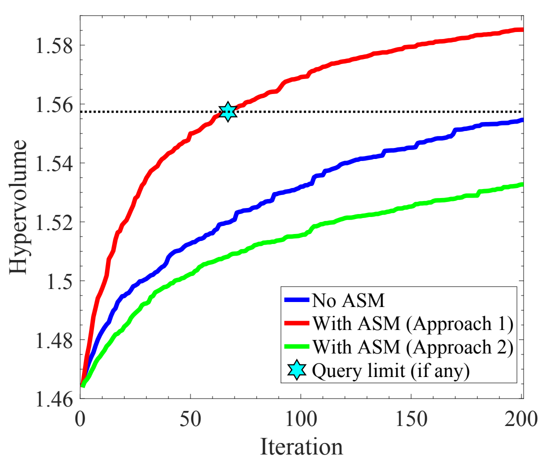

Fig 8. The vehicle Crashworthiness design problem. Hypervolume values averaged over 50 replications of a simulation for all approaches.

Finally, the results of applying both approaches to solve the vehicle crashworthiness design problem are presented in Fig. 8. Again, taking the first approach, larger hypervolumes are quickly obtained in comparison to other cases. As mentioned earlier, the second approach where perform the second step of Bayesian optimization over a pool of best possible experiments selected from the first step of Bayesian optimization step is not able to suggest any improvements which can be due to the fact that it misses exploring all active subspaces.

V. Conclusions and Future Work

In this study, we proposed two approaches to employ the active subspace method in optimization of multi-objective functions in Bayesian optimization framework. By using the active subspace method , it is possible to define the design problem in a lower dimensional space to ease the learning process. Since in multi-objective design problems, each objective function is associated to a different active subspace, we seek to explore all active subspaces to make improvements to the hypervolume of the approximated Pareto front. The results suggest that, by exploring all active subspaces and make evaluations accordingly (that can be done in parallel), better approximations of Pareto frontier is obtained in less amount of computational time, indicating efficiency gains of this approach in comparison to the case which the active subspace tool is not used. The results for the other proposed approach are mixed and is not clear how it interacts with different active subspaces. There is a possibility that it gets stuck in exploring the same objective function's active subspace, so that it is not able to provide significant improvements to the results after some point. Overall, the first proposed approach that explores all objective function's active subspaces is dominating in terms of providing better approximation of a Pareto front. Although the preliminary results are promising and suggesting possible efficiency gains via employing the active subspace method in Bayesian optimization of multi-objective functions, we continue to investigate how different active subspaces in a design problem interact with each other and what are other potential solutions to further increase the efficiency of our proposed approaches. A subject of future work is to implement the active subspace method in multifidelity settings. In such cases, there exist multiple sources (models or experiments) estimating the same set of objective functions varying in fidelity level and computational cost. Via employing an information fusion technique such as reification or cokriging [48, 49], pieces of information from different sources are fused to mitigate the need for querying more expensive models. Therefore, there will be the additional interaction of different models active subspaces as well as different objective functions.

Acknowledgments

The authors acknowledge the support from ARPA-E ULTIMATE Program through Project DE-AR0001427 and from NSF through Grant No. NSF-DEMS-1663130 and NSF-CDSE-2001333. Calculations for the materials science problem were carried out in part at the Texas A&M High-Performance Research Computing (HPRC) Facility.

References

[1] Frazier, P. I., "A tutorial on bayesian optimization," *arXiv preprint arXiv:1807.02811*, 2018.

[2] 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. https://doi.org/10.1023/A:1008306431147.

[3] Močkus, J., "On Bayesian methods for seeking the extremum," *Optimization techniques IFIP technical conference*, Springer, 1975, pp. 400–404. https://doi.org/10.1007/978-3-662-38527-2_55.

[4] Snoek, J., Larochelle, H., and Adams, R. P., "Practical bayesian optimization of machine learning algorithms," *Advances in neural information processing systems*, 2012, pp. 2951–2959.

[5] Rasmussen, C. E., and Williams, C. K. I., *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*, The MIT Press, 2005.

[6] Gill, P. E., Murray, W., and Wright, M. H., *Practical optimization*, SIAM, 2019.

[7] Shan, S., and Wang, G. G., "Survey of modeling and optimization strategies to solve high-dimensional design problems with computationally-expensive black-box functions," *Structural and multidisciplinary optimization*, Vol. 41, No. 2, 2010, pp. 219–241.

[8] Chinesta, F., Huerta, A., Rozza, G., and Willcox, K., "Model order reduction," *Encyclopedia of Computational Mechanics*, 2016.

[9] Benner, P., Gugercin, S., and Willcox, K., "A survey of projection-based model reduction methods for parametric dynamical systems," *SIAM review*, Vol. 57, No. 4, 2015, pp. 483–531.

[10] Russi, T. M., "Uncertainty quantification with experimental data and complex system models," Ph.D. thesis, UC Berkeley, 2010.

[11] Constantine, P. G., Dow, E., and Wang, Q., "Active subspace methods in theory and practice: applications to kriging surfaces," *SIAM Journal on Scientific Computing*, Vol. 36, No. 4, 2014, pp. A1500–A1524.

[12] Constantine, P. G., *Active subspaces: Emerging ideas for dimension reduction in parameter studies*, SIAM, 2015.

[13] Wold, S., Esbensen, K., and Geladi, P., "Principal component analysis," *Chemometrics and intelligent laboratory systems*, Vol. 2, No. 1-3, 1987, pp. 37–52.

[14] Wang, Q., Chen, H., Hu, R., and Constantine, P., "Conditional sampling and experiment design for quantifying manufacturing error of transonic airfoil," *49th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition*, 2011, p. 658.

[15] Dow, E., and Wang, Q., "Output based dimensionality reduction of geometric variability in compressor blades," *51st AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition*, 2013, p. 420.

[16] Khatamsaz, D., and Allaire, D. L., "Materials Design using an Active Subspace Batch Bayesian Optimization Approach," *AIAA SCITECH 2022 Forum*, 2022, p. 0075.

[17] Grey, Z. J., and Constantine, P. G., "Active subspaces of airfoil shape parameterizations," *AIAA Journal*, Vol. 56, No. 5, 2018, pp. 2003–2017.

[18] Lukaczyk, T. W., Constantine, P., Palacios, F., and Alonso, J. J., "Active subspaces for shape optimization," *10th AIAA multidisciplinary design optimization conference*, 2014, p. 1171.

[19] Li, J., Cai, J., and Qu, K., "Surrogate-based aerodynamic shape optimization with the active subspace method," *Structural and Multidisciplinary Optimization*, Vol. 59, No. 2, 2019, pp. 403–419.

[20] Hu, X., Parks, G. T., Chen, X., and Seshadri, P., "Discovering a one-dimensional active subspace to quantify multidisciplinary uncertainty in satellite system design," *Advances in space research*, Vol. 57, No. 5, 2016, pp. 1268–1279.

[21] Constantine, P. G., Emory, M., Larsson, J., and Iaccarino, G., "Exploiting active subspaces to quantify uncertainty in the numerical simulation of the HyShot II scramjet," *Journal of Computational Physics*, Vol. 302, 2015, pp. 1–20.

[22] Khatamsaz, D., Molkeri, A., Couperthwaite, R., James, J., Arróyave, R., Srivastava, A., and Allaire, D., "Adaptive active subspace-based efficient multifidelity materials design," *Materials & Design*, Vol. 209, 2021, p. 110001.

[23] Ma, X., Liu, F., Qi, Y., Wang, X., Li, L., Jiao, L., Yin, M., and Gong, M., "A multiobjective evolutionary algorithm based on decision variable analyses for multiobjective optimization problems with large-scale variables," *IEEE Transactions on Evolutionary Computation*, Vol. 20, No. 2, 2015, pp. 275–298.

[24] Zhang, X., Tian, Y., Cheng, R., and Jin, Y., "A decision variable clustering-based evolutionary algorithm for large-scale many-objective optimization," *IEEE Transactions on Evolutionary Computation*, Vol. 22, No. 1, 2016, pp. 97–112.

[25] Wang, H., Jiao, L., Shang, R., He, S., and Liu, F., "A memetic optimization strategy based on dimension reduction in decision space," *Evolutionary computation*, Vol. 23, No. 1, 2015, pp. 69–100.

[26] Liu, R., Ren, R., Liu, J., and Liu, J., "A clustering and dimensionality reduction based evolutionary algorithm for large-scale multi-objective problems," *Applied Soft Computing*, Vol. 89, 2020, p. 106120.

[27] Marler, R. T., and Arora, J. S., "The weighted sum method for multi-objective optimization: new insights," *Structural and multidisciplinary optimization*, Vol. 41, No. 6, 2010, pp. 853–862. https://doi.org/10.1007/s00158-009-0460-7.

[28] Kim, I. Y., and de Weck, O. L., "Adaptive weighted-sum method for bi-objective optimization: Pareto front generation," *Structural and multidisciplinary optimization*, Vol. 29, No. 2, 2005, pp. 149–158. https://doi.org/10.1007/s00158-004-0465-1.

[29] Das, I., and Dennis, J. E., "Normal-boundary intersection: A new method for generating the Pareto surface in nonlinear multicriteria optimization problems," *SIAM journal on optimization*, Vol. 8, No. 3, 1998, pp. 631–657. https://doi.org/10.1137/S1052623496307510.

[30] Beume, N., "S-metric calculation by considering dominated hypervolume as klee's measure problem," *Evolutionary Computation*, Vol. 17, No. 4, 2009, pp. 477–492. https://doi.org/10.1162/evco.2009.17.4.17402.

[31] Bradstreet, L., While, L., and Barone, L., "A fast many-objective hypervolume algorithm using iterated incremental calculations," *IEEE Congress on Evolutionary Computation*, IEEE, 2010, pp. 1–8. https://doi.org/10.1109/CEC.2010.5586344.

[32] Emmerich, M. T., Deutz, A. H., and Klinkenberg, J. W., "Hypervolume-based expected improvement: Monotonicity properties and exact computation," *2011 IEEE Congress of Evolutionary Computation (CEC)*, IEEE, 2011, pp. 2147–2154. https://doi.org/10.2514/6.2015-0143.

[33] Fonseca, C. M., Paquete, L., and López-Ibánez, M., "An improved dimension-sweep algorithm for the hypervolume indicator," *2006 IEEE international conference on evolutionary computation*, IEEE, 2006, pp. 1157–1163. https://doi.org/10.1109/CEC.2006.1688440.

[34] Russo, L. M., and Francisco, A. P., "Quick hypervolume," *IEEE Transactions on Evolutionary Computation*, Vol. 18, No. 4, 2013, pp. 481–502. https://doi.org/10.1109/TEVC.2013.2281525.

[35] Yang, Q., and Ding, S., "Novel algorithm to calculate hypervolume indicator of Pareto approximation set," *International Conference on Intelligent Computing*, Springer, 2007, pp. 235–244. https://doi.org/10.1007/978-3-540-74282-1_27.

[36] Zitzler, E., and Thiele, L., "Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach," *IEEE transactions on Evolutionary Computation*, Vol. 3, No. 4, 1999, pp. 257–271. https://doi.org/10.1109/4235.797969.

[37] Zhao, G., Arroyave, R., and Qian, X., "Fast Exact Computation of Expected HyperVolume Improvement," 2018.

[38] Feliot, P., Bect, J., and Vazquez, E., "A Bayesian approach to constrained single-and multi-objective optimization," *Journal of Global Optimization*, Vol. 67, No. 1-2, 2017, pp. 97–133. https://doi.org/10.1007/s10898-016-0427-3.

[39] While, L., Bradstreet, L., and Barone, L., "A fast way of calculating exact hypervolumes," *IEEE Transactions on Evolutionary Computation*, Vol. 16, No. 1, 2011, pp. 86–95. https://doi.org/10.1109/TEVC.2010.2077298.

[40] Jaszkiewicz, A., "Improved quick hypervolume algorithm," *Computers & Operations Research*, Vol. 90, 2018, pp. 72–83. https://doi.org/10.1016/j.cor.2017.09.016.

[41] Khatamsaz, D., Peddareddygari, L., Friedman, S., and Allaire, D. L., "Efficient Multi-Information Source Multiobjective Bayesian Optimization," *AIAA Scitech 2020 Forum*, 2020, p. 2127. https://doi.org/10.2514/6.2020-2127.

[42] Khatamsaz, D., Peddareddygari, L., Friedman, S., and Allaire, D., "Bayesian optimization of multiobjective functions using multiple information sources," *AIAA Journal*, Vol. 59, No. 6, 2021, pp. 1964–1974.

[43] Ghoreishi, S. F., Friedman, S., and Allaire, D. L., "Adaptive dimensionality reduction for fast sequential optimization with gaussian processes," *Journal of Mechanical Design*, Vol. 141, No. 7, 2019.

[44] Constantine, P., and Gleich, D., "Computing active subspaces with Monte Carlo," *arXiv preprint arXiv:1408.0545*, 2014.

[45] Deb, K., Thiele, L., Laumanns, M., and Zitzler, E., "Scalable multi-objective optimization test problems," *Proceedings of the 2002 Congress on Evolutionary Computation. CEC'02 (Cat. No. 02TH8600)*, Vol. 1, IEEE, 2002, pp. 825–830. https://doi.org/10.1109/CEC.2002.1007032.

[46] Zitzler, E., Deb, K., and Thiele, L., "Comparison of multiobjective evolutionary algorithms: Empirical results," *Evolutionary computation*, Vol. 8, No. 2, 2000, pp. 173–195.

[47] Tanabe, R., and Ishibuchi, H., "An easy-to-use real-world multi-objective optimization problem suite," *Applied Soft Computing*, Vol. 89, 2020, p. 106078.

[48] Allaire, D., and Willcox, K., "Fusing information from multifidelity computer models of physical systems," *2012 15th International Conference on Information Fusion*, IEEE, 2012, pp. 2458–2465.

[49] Khatamsaz, D., and Allaire, D. L., "A Comparison of Reification and Cokriging for Sequential Multi-Information Source Fusion," *AIAA Scitech 2021 Forum*, 2021, p. 1477.