

Physics-Constrained Bayesian Optimization for Optimal Actuators Placement in Composite Structures Assembly

Areej AlBahar¹, Inyoung Kim², Xing Wang, and Xiaowei Yue³, *Senior Member, IEEE*

Abstract—Complex constrained global optimization problems such as optimal actuators placement are extremely challenging. Such challenges, including nonlinearity and nonstationarity of engineering response surfaces, hinder the use of ordinary constrained Bayesian optimization (CBO) techniques with standard Gaussian processes as surrogate models. To overcome those challenges, we propose a physics-constrained Bayesian optimization with multi-layer deep structured Gaussian processes, *MGP-CBO*. Specifically, we introduce a surrogate model with a multi-layer deep Gaussian process (MGP) mean function. The hierarchical structure of our model enables the applicability of constrained Bayesian optimization to complex nonlinear and nonstationary processes. The deep Gaussian process regression model, *MGP*, can efficiently and effectively represent the response surface function between actuators and dimensional deformations, thus yielding a better estimated global optimum in a shorter computational time. The proposed *MGP-CBO* model can realize faster convergence to the global optimum with lower constraint violations. Through extensive evaluations carried out on synthetic problems and a real-world engineering design problem, we show that *MGP-CBO* outperforms existing benchmarks. Although we use the optimal actuators placement as a demonstration example, the proposed *MGP-CBO* model can be applied to other complex nonstationary engineering optimization problems.

Note to Practitioners—Bayesian optimization is a widely used sequential design strategy for engineering optimization because it does not rely on functional forms of response surfaces. This paper helps address two questions in practice: (i) how to incorporate physics constraints into Bayesian optimization. (ii) How to do Bayesian optimization when the systems have hierarchical structures. In practice, the hierarchical system structure is

ubiquitous, and the engineering optimization is constrained by physical laws or special requirements. Therefore, the proposed physics-constrained Bayesian optimization with a multi-layer Gaussian process could provide a new tool for engineering design optimization problems. The computational convergence and complexity have been investigated. The proposed method is applicable to broad complex and nonstationary engineering optimization problems.

Index Terms—Constrained Bayesian optimization, nonstationary processes, physical constraints, hierarchical systems, composite structures assembly.

I. INTRODUCTION

COMPLEX engineering design optimization problems are often nonlinear, nonstationary, expensive-to-evaluate, and constrained. Such challenges inhibit the use of statistical nonlinear global optimization algorithms and existing constrained Bayesian optimization techniques. To tackle those challenges, a physics-constrained Bayesian optimization model with a deep Gaussian process surrogate model is proposed. The proposed model is used to infer the global optimum under a set of physical and expensive-to-evaluate constraint functions.

Constrained Bayesian optimization is a class of Bayesian optimization where the black-box objective function to be optimized is subjected to a set of constraints. In ordinary constrained Bayesian optimization, Gaussian processes, which are probabilistic surrogate models parameterized by mean and kernel functions, are used to model and represent the objective and constraints functions. In this setting, both the objective function and set of constraints are assumed to be independent and follow Gaussian processes with constant mean functions. In practice, it is usually assumed that the kernel function is the most critical component of a Gaussian process model. However, in this paper, we prove that the mean function of a GP is at least as important and critical as its kernel function.

Gaussian processes are nonparametric probabilistic models used to represent complex and expensive-to-evaluate functions. Despite their great advantages in uncertainty quantification, interpretability, and extensive uses in inferential modeling, they cannot efficiently represent the nonstationarity and anisotropy of engineering processes. These drawbacks limit their capabilities in modeling complex spatial nonstationary response functions and in estimating their global optimum. A Gaussian process mean function has a significant impact on the performance of the surrogate model. Besides the constant mean function, linear and nonlinear mean functions

Manuscript received 7 May 2022; accepted 1 August 2022. This article was recommended for publication by Associate Editor K. Paynabar and Editor J. Li upon evaluation of the reviewers' comments. The work of Areej AlBahar was supported by Kuwait University. The work of Xing Wang was supported by NSF under Grant 2153329. The work of Xiaowei Yue was supported in part by the USA National Science Foundation (NSF) under Grant CMMI-2035038, in part by the Department of Defense (DoD) MEEP Program under Award N00014-19-1-2728, and in part by the Grainger Frontiers of Engineering Grant Award from the National Academy of Engineering. (Corresponding author: Xiaowei Yue.)

Areej AlBahar is with the Department of Industrial and Systems Engineering, Virginia Tech, Blacksburg, VA 24061 USA, and also with the Department of Industrial and Management Systems Engineering, Kuwait University, Kuwait City 12037, Kuwait (e-mail: areejaa3@vt.edu).

Inyoung Kim is with the Department of Statistics, Virginia Tech, Blacksburg, VA 24061 USA (e-mail: inyoungk@vt.edu).

Xing Wang is with the Department of Mathematics, Illinois State University, Normal, IL 61790 USA (e-mail: xwang70@ilstu.edu).

Xiaowei Yue is with the Department of Industrial and Systems Engineering, Virginia Tech, Blacksburg, VA 24061 USA (e-mail: xwy@vt.edu).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TASE.2022.3200376>.

Digital Object Identifier 10.1109/TASE.2022.3200376

1545-5955 © 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.

See <https://www.ieee.org/publications/rights/index.html> for more information.

have been proposed. Although these mean functions may improve Gaussian process model representations and inference, they still have a challenge in modeling complex nonstationary functions. In this regard, we enhance the efficacy and efficiency of Gaussian processes by modeling the mean function as a multi-layer deep Gaussian process. Although deep and nonstationary Gaussian processes have been previously presented and applied in Bayesian optimization [1], [2], [3], they are still inefficient in modeling spatial nonstationary functions. Detailed literature on deep and nonstationary Gaussian processes is provided in Section II. In this paper, to clearly show the importance of the mean function of a GP model, we limit our analysis and comparisons to Gaussian processes with nonstationarities in either the mean or the kernel functions.

Constrained Bayesian optimization has been widely used in optimizing black-box constrained objective functions [4], robotics [5], automotive [6], and chemical design [7]. However, little research has been done on the application of constrained Bayesian optimization in aerospace engineering design. Most engineering design optimization problems incur physical constraints that limit design data acquisition and response function optimization. We take aerospace assembly as an example. Aerospace engineering systems use advanced materials such as composite structures that impose critical design optimization challenges. One significant aerospace engineering design problem is the optimal shape control of composite fuselages in which actuators are installed to control the dimensions of the composite structure and improve the quality of the assembly process.

The placement of actuators in composite structures assembly is a critical aerospace process design problem. Composite materials such as carbon fiber, carbon sandwich and fiber-reinforced matrix systems, are nowadays the building blocks of many aircrafts. For instance, composite structures account for more than 50% of weight in the Boeing 787 airplane and 35% in the Lockheed Martin F-35 fighter. Although they have advantages such as light weight and higher strength and stiffness compared to traditional aluminum alloys, composite materials are more expensive and brittle than their counterparts. Composite materials impose challenges that hinder the use of existing Bayesian optimization techniques, such as having complex nonlinear, nonstationary, and anisotropic response surface functions [8], [9]. More precisely, the process of efficiently assembling two composite structures is critical and often needs a quality control preprocessing that uses actuators to reduce the supplier-inherited dimensional gap between the two composite structures. Therefore, precise inferential and predictive modeling of the dimensional deviations is indispensable to the ultra-high precision quality control of the assembled composite structures.

Incorporating domain knowledge is significant to the success and accuracy of predictive models. Therefore, to overcome the aforementioned limitations, a different approach is proposed to enhance the modeling capabilities of Gaussian processes. More specifically, we propose a physics-constrained Bayesian optimization technique with a deep multi-layer Gaussian process surrogate model. The integration of the

deep embedded hierarchical structure in the mean function of the proposed Gaussian process is sought to aid in superior modeling of nonstationary processes, thus improving process representations and function approximations.

Our contributions are summarized as follows.

1. We propose a novel Multi-layer deep Gaussian Process based Constrained Bayesian Optimization algorithm, called *MGP-CBO*, that efficiently optimizes physics-constrained complex nonstationary functions for optimal actuators placement in composite structures assembly.
2. We develop a Gaussian process surrogate model with a multi-layer deep Gaussian process mean function (*MGP*) and provide its approximate posterior predictive distribution.
3. We provide computational investigations that analyze the convergence rate to the global optimum, constraints violation rates, and time complexity properties of our proposed *MGP-CBO* model.
4. The proposed *MGP-CBO* model is evaluated on two 2-D synthetic functions optimization and a case study of an aerospace engineering design optimization problem. We show that *MGP-CBO* outperforms stationary and nonstationary CBO models in the two experiments and case study.

The remainder of this paper is organized as follows: Section II illustrates the background of constrained Bayesian optimization and its variants, and some related work on composite structures assembly. Section III describes the proposed deep Gaussian process and presents illustrative experiments. Section IV describes the problem setting and computational algorithm of our proposed *MGP-CBO* model. Section V discusses the properties of the *MGP-CBO* model, such as convergence, constraints violations, and computational complexity. Section VI conducts the numerical study and compares the proposed *MGP-CBO* model against the benchmarks. Section VII presents the case study of actuator placements in composite structures assembly. Finally, a brief summary is provided in Section VIII.

II. LITERATURE REVIEW

In this section, we provide literature on constrained Bayesian optimization, deep and nonstationary Gaussian processes, and composite structures assembly in aerospace industry.

A. Constrained Bayesian Optimization

Constrained Bayesian optimization is a type of global optimization where the optimization process is restricted by a set of constraints. In constrained Bayesian optimization, surrogate models (e.g., Gaussian processes) can still be used to represent the objective and constraint functions. However, standard acquisition functions such as the expected improvement, lower confidence bound, and entropy search, to name a few, have to be constrained by some form of a feasibility measure. Thus, different constrained acquisition functions such as constrained expected improvement [10], constrained

lower confidence bound [6], and constrained predictive entropy search [11] have been proposed to accommodate the feasibility of the constraints functions and limit the search space design when acquiring data.

B. Deep and Nonstationary Gaussian Processes

Deep Gaussian processes were first introduced in [12]. Their deep structure is built by composing multiple GPs; where each GP's output is modeled as the input to the next. Nested computer codes [13] were the first to apply this deep structure in Bayesian optimization. In their work, they further extended the use of Bayesian optimization to 2-layer composite functions, where the inner and outer functions are assumed to be black-box and are modeled as Gaussian processes. Modeling the GP mean function as a deep GP was introduced in [14]. Although their proposed approach has similar characteristics to ours, two main significant differences, in terms of the research problem and the algorithm, are present. First, our multi-layer deep Gaussian process mean function is developed for Bayesian optimization and surrogate modeling, while their approach was developed for integrating multi-fidelity data. Second, our algorithm is based on directly optimizing the posterior distribution, while theirs was based on variational Bayes sampling which was obtained from the distribution of the approximated posterior distribution. Since we need to accurately estimate the inputs that maximize the posterior distribution, we need to directly optimize the correct posterior distribution. Hence, variational Bayes is not efficient for Bayesian optimization. Another direction is the use of nonstationary and composite kernel functions in Gaussian processes to represent complex nonstationary functions [15], [16], [17]. Although this direction may improve objective function approximations, the resulting model is complex, statistically intractable, and still lacks the flexibility of accurately representing non-stationary processes. Input warping has also been presented to accommodate nonstationary functions in Bayesian optimization [18]. Partitioned GPs [19], composite GPs [20], [21], nonstationary GPs [22], convolutional GPs [23], and Neural network GPs [24] have also been proposed. However, modeling complex nonstationary functions for composite structures assembly is still a major challenge.

C. Actuator Placement for Composite Structures Assembly

The problem of assembling two composite structures such as assembling two fuselages in an aircraft has gained a lot of interest in the past few years. One important step to control the assembly process is the placement of actuators. Actuators are placed on composite fuselages to bring the composite structure closer to its design dimensions prior to the assembly process. A finite element simulation model has been developed for evaluating the shape control of composite fuselages [8]. Surrogate model based control has been developed to mitigate the naturally occurring gap between two pre-assembled composite structures [25], [26]. Du *et al.* proposed a constrained sparse learning model to optimize actuators placement [27], where Alternating Direction Method of Multipliers (ADMM) was used for parameter estimation. Jiao and Djurdjanovic applied

the generic algorithm to deal with the combinatorial optimization in multi-station manufacturing quality control [28].

III. MULTI-LAYER DEEP GAUSSIAN PROCESS (MGP)

In this section, we introduce the proposed multi-layer deep Gaussian process, *MGP*, we theoretically show its posterior predictive mean and variance, and perform multiple regression experiments and analysis. The proposed MGP model is built to efficiently learn the underlying nonstationarity of the process and improve global optimum function approximations. The deepness of the proposed Gaussian process is illustrated in its mean function. The proposed MGP deep Gaussian process is structured as follows:

Suppose we have n samples, then, each response variable y_i is assumed to:

- 1) follow a normal distribution,
- 2) be nonparametrically dependent on x_i ,

where x_i^T is a $1 \times p$ vector of the i th observation. More specifically, we model the response variable y_i as $y_i = f(x_i) + \epsilon_i$, where $f(x)$ is modelled as a GP with the proposed multi-layer deep Gaussian process (i.e., MGP) as its mean function, and $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2 I_n)$.

In general, a Gaussian process model with an L -layer MGP model as its mean function can be written, under the Bayesian hierarchical model framework, as follows:

$$\begin{aligned} [y|f(x)] &\sim \mathcal{N}(f(x), \sigma_\epsilon^2 I_n), \\ [f(x)|\mu_L(x)] &\sim \mathcal{GP}(\mu_L(x), k(x, x')), \\ [\mu_l(x)|\mu_{l-1}(x)] &\sim \mathcal{GP}(\mu_{l-1}(x), k_l(x, x')), \quad l = 1, \dots, L, \end{aligned}$$

where L represents the number of layers of the MGP model, $k(x, x')$ is the kernel (i.e., covariance) function, and $\mu_0(x)$ is a constant mean function. From the Bayesian point of view, the Gaussian process distributions of the L layers can be considered as the distributions of the hyperparameters. These hyperparameter distributions can be automatically solved for identifiability issues in the model.

Tuning of Model Parameters: L and $\mu_0(x)$. In this paper, we consider, for demonstration purposes, two values of L (i.e., two and three layers), however, for any given scenario or application, the hyperparameter L may be optimized using different statistical model selection criteria such as Akaike's Information Criteria (AIC), Bayesian Information Criteria (BIC), or even modeled as a hyperparameter in the Bayesian optimization model (e.g., L is modeled as a Poisson distribution). For the mean function of the zeroth layer, $\mu_0(x)$, we use a zero mean function.

The multi-layer deep structure can provide Gaussian processes more flexibility in representing nonstationary and complex response surface functions. In subsection III-A, we will explore the marginal likelihood and hyperparameter estimation of the MGP model.

A. Marginal Likelihood and Hyperparameter Estimation

The marginal likelihood denoted as $p(y|X)$ is the integral of the likelihood multiplied by the prior of the proposed L -layer

MGP model as shown in Equation 1.

$$p(y|X) = \int \cdots \int p(y|f)p(f|\mu_L) \times \prod_{l=1}^L p(\mu_l|\mu_{l-1})d\mu_L \cdots d\mu_1. \quad (1)$$

Although scalable inference in this construction is a challenge, at each layer l , the output μ_l only depends on the corresponding input μ_{l-1} . This property is a direct consequence of setting every layer upon, in the proposed MGP model, exactly a standard GP.

The mean and variance of the marginal distribution of y under the L -layer MPG model can be expressed as follows:

$$\begin{aligned} E[y] &= E[E[y|f(x)]] = E[f(x)] = E[E[f(x)|\mu_L(x)]] \\ &= E[\mu_L(x)] \\ V[y] &= V[E[y|f(x)]] + E[V[y|f(x)]] \\ &= V[f(x)] + \sigma_\epsilon^2 \mathbf{I}_n \\ &= V[E[f(x)|\mu_L(x)]] + E[V[f(x)|\mu_L(x)]] + \sigma_\epsilon^2 \mathbf{I}_n \\ &= V[\mu_L(x)] + K + \sigma_\epsilon^2 \mathbf{I}_n, \end{aligned}$$

where K is the *Gram matrix* of the kernel function $k(x, x')$.

Proposition 1: If a Gaussian process model has an L -layer MGP as its mean function, then, the output of the l^{th} layer mean function denoted as $\mu_l(x)$ depends only on the corresponding input $\mu_{l-1}(x)$, $l = 2, \dots, L$.

Proposition 2: Assume we build a model via a Gaussian process with an L -layer MGP as the mean function, then, the predictive function value f_* and L^{th} layer mean value μ_{L*} corresponding to the test input x_* , can be sampled from the normal probability distributions $\mathcal{N}(\bar{f}_*, \text{cov}(f_*))$ and $\mathcal{N}(\bar{\mu}_{L*}, \text{cov}(\mu_{L*}))$, respectively, where

$$\begin{aligned} \bar{f}_* &= k(x_*, x)[K + \sigma_\epsilon^2 \mathbf{I}_n]^{-1}(y - \mu_L(\mu_{L-1}(x))), \\ \text{cov}(f_*) &= k(x_*, x_*) - k(x_*, x)[K + \sigma_\epsilon^2 \mathbf{I}_n]^{-1}k(x, x_*), \end{aligned}$$

and

$$\begin{aligned} \bar{\mu}_{L*} &= k_L(x_*, x)[K_L + K]^{-1}[f(x) - \mu_{L-1}(\mu_{L-2}(x))], \\ \text{cov}(\mu_{L*}) &= k_L(x_*, x_*) - k_L(x_*, x)[K_L + K]^{-1}k_L(x, x_*). \end{aligned}$$

These two properties make the approach scalable by enabling efficient probabilistic back-propagation.

The hyperparameters of the multi-layer deep Gaussian process, MGP model, are optimized, at each layer, by maximizing the log marginal likelihood under a standard GP. The posterior predictive mean and variance are also obtained using this efficient probabilistic back-propagation approach. The optimization process is carried out as follows:

Firstly, we start with $\mu_0(x)$ which is set as a constant zero mean function.

Secondly, we sequentially estimate $\mu_{l-1}(x)|\mu_{l-2}(x)$ for $l = 2, \dots, L$ via back-propagation [29].

Thirdly, we estimate $\mu_L(x)$ and obtain the GP model of $f(x)$.

Finally, we obtain the predictive distribution of the response variable y .

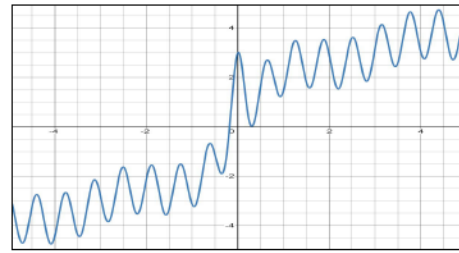


Fig. 1. Visualization of the 1-D regression function.

In this paper, for demonstration purposes, we use only two- and three-layer deep Gaussian processes and update the sequence efficiently. The advantage of sequential updating is that the updated GP is fed as an initial/prior model for the next GP.

B. Multi-Layer Deep Gaussian Process Regression

We perform a one-dimensional regression task to evaluate the effectiveness and efficiency of our proposed Gaussian process with a multi-layer deep Gaussian process mean function, MGP. The regression model is as follows:

$$y_i = f(x_i) + \epsilon_i, \quad (2)$$

where ϵ is the noise parameter and is assumed to follow a Gaussian distribution (i.e., $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2 \mathbf{I}_n)$). We use a one-dimensional function, shown in Equation 3, for the regression task.

$$\begin{aligned} f(x) &= x + \sin(x) + 2 * e^{-30*x^2} \\ &\quad + \cos(10x) + \sin(x) * \cos(x), \end{aligned} \quad (3)$$

where $x \in [-5, 5]$ and the true global minimum is $y_{\min} = -4.751$. The one-dimensional regression function can be visualized in Fig. 1.

1) Experimental Setup: In this experiment, we compare between six regression models, namely: stationary GP (SGP), Polynomial nonstationary GP (Polynomial NSGP), ArcCosine nonstationary GP (ArcCosine NSGP), Neural Network GP (NN GP), two-layer deep Gaussian process (2-layer MGP), and three-layer deep Gaussian process (3-layer MGP). The SGP regression model has a Matern5/2 kernel function, whereas the two nonstationary GP models, namely Polynomial NSGP and ArcCosine NSGP, have a Polynomial kernel function and an ArcCosine (i.e., a neural network) kernel function, respectively. The NN GP regression model has a Matern5/2 kernel function and a neural network mean function. The two variants of our proposed multi-layer deep Gaussian process model (i.e., 2-layer and 3-layer MGP models) have a Matern5/2 kernel function and two-layer and three-layer Gaussian process mean functions, respectively.

2) Metrics: For comparison purposes, we use two metrics, namely: the mean squared error (MSE) 4 and the average optimization time in seconds.

$$MSE = \sum_{v=1}^V (y_v^* - y_{\min})^2, \quad (4)$$

where y_v^* is the estimated global minimum of the v^{th} run and $y_{\min}$ is the true global minimum.

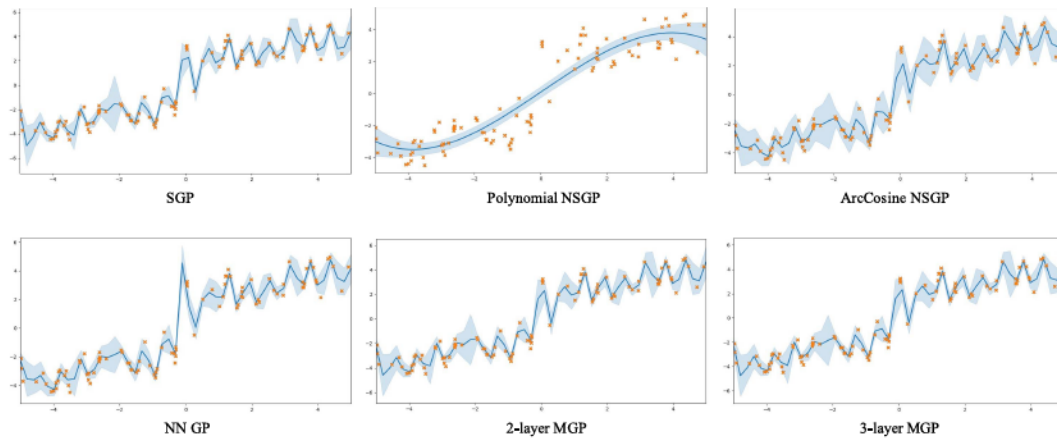


Fig. 2. Posterior predictive distributions of SGP, Polynomial NSGP, ArcCosine NSGP, NN GP, 2-layer MGP, and 3-layer MGP regression models.

TABLE I

MEAN SQUARED ERROR (MSE) AND OPTIMIZATION TIME IN SECONDS OF SIX BENCHMARK GP REGRESSION MODELS WHEN THE NOISE PARAMETER FOLLOWS A GAUSSIAN AND STUDENT-T DISTRIBUTIONS. THE BOLD UNDERLINED RESULTS REPRESENT THE BEST OUTCOME IN EACH METRIC, WHILE THE UNDERLINED RESULTS REPRESENT THE SECOND BEST OUTCOME IN EACH METRIC

Model	Gaussian Noise		Student-t Noise	
	MSE	Optimization Time	MSE	Optimization Time
SGP	0.1185	0.2893	0.1608	0.2820
Polynomial NSGP	0.9196	<u>0.2728</u>	0.8558	<u>0.2744</u>
ArcCosine NSGP	0.1633	0.5131	0.4810	0.4713
NN GP	0.1547	3.4859	0.1794	3.4575
2-layer MGP	<u>0.1015</u>	0.3121	<u>0.1155</u>	0.3056
3-layer MGP	<u>0.0983</u>	<u>0.2750</u>	<u>0.1090</u>	<u>0.2662</u>

3) *Results*: We run the six aforementioned GP regression models for 1000 optimization iterations. We show the results of the regression task in Table I. From Table I, in terms of the MSE metric, the two variants of our proposed model (i.e., 2-layer and 3-layer MGP models) outperform the other four benchmarks. In terms of the optimization time, the 3-layer MGP and Polynomial NSGP are more efficient than the other four benchmark models.

To visualize the performance of the six Gaussian process regression benchmark models, we show the predictive posterior distributions of each benchmark model in Fig. 2. The 95% confidence intervals of the predictive posterior distributions of the 2-layer MGP and 3-layer MGP models are narrower than the predictive posterior distribution of the SGP, ArcCosine NSGP, and NN GP models. Meanwhile, the predictive posterior distribution of the Polynomial NSGP model does not represent the true regression function very well.

a) *Sensitivity analysis*: We test the effectiveness and robustness of our proposed multi-layer deep Gaussian process, MGP, when the noise parameter ϵ is assumed to follow a Student-t distribution (i.e., $\epsilon \sim ST(df, 0, \sigma_\epsilon^2 I_n)$, where df is the degree of freedom) in the regression model shown previously in Equation 2. For comparison purposes, we evaluate the six Gaussian process regression benchmarks used in the previous regression task. We use the regression function in Equation 3 for this regression task. We use the mean squared error (MSE) and average optimization time in seconds as two comparison metrics. Table I shows the results of

the regression task when the noise term follows a Student-t distribution. From Table I, in terms of the MSE metric, the two variants of our proposed model (i.e., 2-layer and 3-layer MGP models) outperform the other four benchmarks. In terms of the optimization time, the 3-layer MGP and Polynomial NSGP are more efficient than the other four benchmark models. Attaining equivalent conclusions, in the two regression tasks, indicates our proposed MGP model is flexible, robust, and effective.

b) *Computational convergence and training time*: Two of the most important properties of an effective and efficient regression model are the computational convergence and optimization time. Our proposed multi-layer deep Gaussian process model, MGP, has shown remarkable results in the regression task with the Gaussian noise as well as the regression task with the Student-t noise. The two variants of our proposed model (i.e., 2-layer MGP and 3-layer MGP) outperform the SGP, Polynomial and ArcCosine NSGP, and the NN GP benchmark regression models in terms of the mean squared error metric. Obtaining a lower MSE corresponds to a better convergence rate to the true global minimum. Moreover, besides obtaining a lower MSE, the 3-layer MGP regression model is more efficient at attaining the estimated minimum than the other benchmarks. Although the Polynomial NSGP has a comparable optimization time to the 3-layer MGP, its MSE value is larger which corresponds to inaccurate function estimations. The proposed MGP model has verified its ability in efficiently optimizing the response surface function through requiring a significantly shorter total optimization time and in effectively obtaining reliable predictive models with a better estimated global minimum.

IV. PHYSICS-CONSTRAINED BAYESIAN OPTIMIZATION WITH DEEP GAUSSIAN PROCESSES

In this section, we start with an illustration of the generalized constrained Bayesian optimization framework, subsection IV-A, and then we propose our multi-layer Gaussian process based physics-constrained Bayesian optimization model in subsection IV-B.

A. Constrained Bayesian Optimization

In the constrained Bayesian optimization framework, we seek to minimize the response surface objective function

subject to a set of inequality constraints. In this setting, the constrained Bayesian optimization problem is formulated as:

$$\min_{x \in D_x} f(x) \text{ s.t. } c_r(x) \leq 0 \quad \forall r \in [1, \dots, R],$$

where x denotes the input design vector, r is the constraint index, and the domain of all design variables is referred to as D_x . The objective function $f(x)$ and constraint functions $c_r(x)$ are assumed to follow standard Gaussian processes as shown below.

$$f(x) \sim \mathcal{GP}(\mu, \Sigma) \text{ and } c_r(x) \sim \mathcal{GP}(\mu_r, \Sigma_r). \quad (5)$$

A standard Bayesian optimization model consists of a surrogate model, usually a Gaussian process, and an acquisition function, usually the expected improvement. The optimization process starts with the acquisition function, a simple and cheap-to-evaluate function used to acquire the next set of points to be evaluated based on specific exploitation/exploration trade-offs. In constrained Bayesian optimization, standard acquisition functions have been manipulated to include some feasibility measures of the set of constraint functions. One of the most used constrained acquisition functions is the constrained expected improvement (cEI) [10] which integrates constraints feasibility through the cumulative distribution function, CDF, of each individual constraint's GP. The constrained expected improvement function is shown in Equation 6.

$$cEI(x) = EI(x) \times \prod_{r=1}^R \Phi\left(\frac{\mu_r}{\sigma_r}\right) \quad (6)$$

where Φ is the constraint's cumulative distribution function (i.e., CDF) of the Gaussian process model and $EI(x)$ is the expected improvement acquisition function [30]. The Gaussian process models of the constraint functions are assumed to be independent and therefore their CDF's are multiplied together.

B. Physics-Constrained Bayesian Optimization With Multi-Layer Deep Gaussian Processes (MGP-CBO)

The physical constraints in engineering design optimization represented by materials properties and natural laws impose limitations on the physical design of a specific system. Physical constraints defined by materials properties include but are not limited to strength, compression/tension, density, and fatigue. Physical constraints defined by natural laws include but are not limited to laws of forces, motions, gravity, and stress. Engineering design response surface functions are usually highly complex and nonstationary. In this paper, we present a physics-constrained Bayesian optimization model with embedded deep Gaussian process structure. Current state-of-the-art Bayesian optimization methods have strong assumptions about the mean function of the Gaussian process (i.e., zero/constant mean functions). This strong assumption affects the inferential model through misspecifications of the mean of the GP as well as disregarding any dynamical nonstationarity of the true process. Thus, modeling the objective function as a deep Gaussian process better represents the response function and enables superior function estimations and lower constraints violations.

Algorithm 1 Physics-Constrained Bayesian Optimization With Multi-Layer Deep Gaussian Processes

Input: A vector of z initial points $X_z = [x_1, \dots, x_z]$, the initial constrained Bayesian optimization model and the maximum number of iterations (T).

- 1: Find the initial objective function values $f(x_1), \dots, f(x_z)$ and constraint functions values $c_r(x_1), \dots, c_r(x_z)$ for every $r \in [1, \dots, R]$.
- 2: Find $\mu_0(x)$.
- 3: **for** $l = 2, \dots, L$ **do**
- 4: Estimate the hyperparameters set θ_{l-1} of the $\mu_{l-1}(x)$ GP by maximizing the log marginal likelihood.
- 5: Update the mean and variance of $\mu_{l-1}(x)$.
- 6: **end for**
- 7: Estimate the hyperparameters set θ of the $f(x)$ GP model.
- 8: Update the mean and variance of $f(x)$ GP model.
- 9: **for** $t = z + 1, \dots, T$ **do**
- 10: Select the next point to be evaluated, x_t , such that $x_t = \argmax cEI(x)$ for $x \in \mathcal{D}_X$.
- 11: Evaluate $f(x_t)$ and $c_r(x_t)$ at x_t .
- 12: Redo steps 2 to 8.
- 13: Find $y_t = f(x_t) + \epsilon_t$
- 14: Update the mean and variance of the GP model of y .
- 15: Update the so far optimal input point x^* and the corresponding y^*
- 16: **end for**
- 17: **Return** The response variable's Gaussian process posterior predictive mean and variance, and the global optimum y^* at x^* .

1) Problem Setting: A constrained global optimization problem often consists of an objective function (i.e., $f(x)$) and a set of constraint functions (i.e., $c_r(x)$'s). The constraint functions restrict the global optimization process by applying some limits on the search space available for data acquisition. In our problem setting, we assume that our objective function is complex, nonlinear, nonstationary, and is subjected to a set of constraints. The constraint functions are also assumed to be expensive-to-evaluate. We use the proposed physics-constrained Bayesian optimization algorithm, *MGP-CBO*, to estimate the global optimum. Both the objective and constraint functions are modeled as Gaussian processes as shown in Equation 5. More precisely, the Gaussian process of the objective function is embedded with the proposed multi-layer deep structured Gaussian process as its mean function. We include theoretical investigations and computational properties of the proposed MGP-CBO model in the next section.

2) Bayesian Optimization Algorithm: Our proposed physics-constrained Bayesian optimization model is shown in detail in Algorithm 1. To add more flexibility, each layer in the proposed MGP multi-layer deep Gaussian process model can have a different kernel function. Algorithm 1 represents the step-by-step framework of the proposed MGP-CBO model.

An illustration of the pseudo code in Algorithm 1 is as follows. The constrained Bayesian optimization process starts

with a set of initial points that are randomly sampled from the input search space. The initial points are used to initialize the GP models of the objective, $f(x)$, and constraint functions, $c_r(x)$'s, respectively. Then, the process of updating the proposed multi-layer deep Gaussian process (MGP) is carried out as follows: (1) the zeroth layer of the mean function, $\mu_0(x)$, is found (i.e., a zero mean function), (2) for each layer of the MGP model, ($l = 2, \dots, L$), the mean function is estimated given the mean function of the previous layer (i.e., $\mu_{l-1}(x)|\mu_{l-2}(x)$), and the prior distribution of each mean function is updated into the posterior. Then, the Bayesian optimization iterations start by maximizing the constrained expected improvement acquisition function and selecting the next point to be evaluated. The objective function and constraint functions are then evaluated at the selected point, and their GP models are updated into the posterior distribution. Using the previously selected point, the GP models of each layer of the MGP model and the response surface variable, y , are updated. After each optimization iteration, the so far optimal y^* at the input point x^* is updated. The previous steps are repeated until no further optimization iterations are available. After exceeding all the available iterations, the estimated global optimum y^* is returned along with the posterior predictive mean and variance of the response variable y .

V. PROPERTIES OF THE MGP-CBO MODEL

In this section, we investigate the computational properties of our proposed physics-constrained Bayesian optimization model. More specifically, we analyze the computational convergence to the global optimum, the computational violations of the constraint functions, and computational complexity properties of MGP-CBO.

A. Analysis of the Computational Convergence

We computationally evaluate the convergence of our proposed MGP-CBO model to the global optimum. To show the convergence rate of our proposed physics-constrained Bayesian optimization model with the multi-layer deep Gaussian process surrogate model (i.e., MGP-CBO) in comparison with the baseline stationary constrained Bayesian optimization (i.e., SCBO) model, we optimize a 2-D synthetic function, *Branin-Hoo*, shown in Equation 9. We use the 2-layer MGP-CBO variant of our proposed model and denote it as MGP-CBO. We compare the proposed MGP-CBO and SCBO models in optimizing Equation 9 for 500 iterations (i.e., objective function evaluations). After each iteration, the best objective function value so far is recorded and plotted against the number of iterations. We show the convergence rates of MGP-CBO and the baseline SCBO in Fig. 3. From Fig. 3, we conclude that the MGP-CBO model has a faster convergence rate to the global optimum than its counterpart the SCBO model.

B. Computational Analysis of Constraints Violations

We computationally investigate the constraints violations of our proposed MGP-CBO model. To demonstrate the constraint

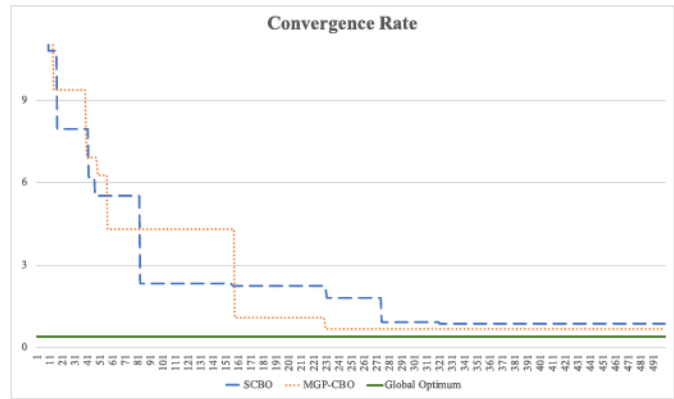


Fig. 3. Convergence rates of MGP-CBO and SCBO models when optimizing the 2-D Branin-Hoo function. The orange dotted line represents the convergence rate of the proposed MGP-CBO model, whereas the dashed blue line represents the convergence rate of the SCBO model. The global optimum is at 0.3979 and is represented by the solid green line.

functions violations of our proposed physics-constrained Bayesian optimization model with a multi-layer deep Gaussian process (i.e., MGP-CBO) in comparison with the baseline stationary constrained Bayesian optimization (i.e., SCBO) model, we optimize a 2-D synthetic function, *Branin-Hoo*, shown in Equation 9. We calculate the constraint violations rate, denoted as CV , for every constraint $r \in [1, \dots, R]$ using the following equation:

$$CV_r = \frac{\sum_{j=1}^T \mathbf{1}[c_r(x_j) > 0]}{T}, \quad (7)$$

where $\mathbf{1}[c_r(x) > 0]$ represents the indicator function and T is the maximum number of optimization iterations. The indicator function is equal to 1 if the statement $c_r(x) > 0$ is true for a given design point x , and zero otherwise.

We use the 2-layer MGP-CBO variant of our proposed model and denote it as MGP-CBO. We compare the proposed MGP-CBO and SCBO models in optimizing Equation 9 for 500 iterations (i.e., objective function evaluations). We visually show the constraint violations of our proposed model, MGP-CBO, in comparison to the SCBO model in Fig. 4. From Fig. 4, we verify that MGP-CBO yields significantly lower constraint function violations in comparison to the SCBO baseline model.

C. Computational Time Complexity

The computational time complexity of a Gaussian process model depends on the number of arithmetic operations (e.g., matrix inversion) applied to its kernel matrix $K + \sigma^2 I_n$. More specifically, a standard 1-layer Gaussian process model needs n^3 operations for its kernel matrix inversion, which corresponds to time complexity of $O(n^3)$. Our proposed MGP-CBO model which uses an L -layer deep Gaussian process mean function, increases the number of Gaussian processes needed to represent the response variable y to $L + 1$. Then, the time complexity of the MGP-CBO model becomes $(L + 1) * O(n^3)$ which is equivalent to $O(n^3)$. This theoretical investigation is equivalent to the results shown in the two regression tasks of section III and will be further verified in the simulation study in section VI.

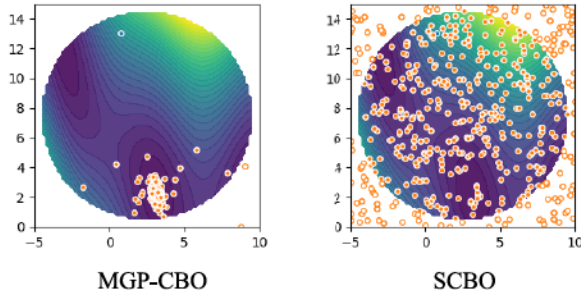


Fig. 4. Demonstration of the constraints violations when MGP-CBO and SCBO models are used to optimize the 2-D Branin-Hoo function. White areas represent infeasible regions. Blue circles are initial points and orange circles are data points acquired through constrained Bayesian optimization.

VI. SIMULATION STUDY

Our proposed physics-constrained Bayesian optimization algorithm, MGP-CBO, is evaluated against state-of-the-art constrained Bayesian optimization models. We use six constrained Bayesian optimization models as benchmarks for evaluation and comparison purposes. The six benchmark models are stationary constrained Bayesian optimization (SCBO), Polynomial nonstationary constrained Bayesian optimization (Polynomial NSCBO), ArcCosine nonstationary constrained Bayesian optimization (ArcCosine NSCBO), stationary constrained Bayesian optimization with a neural network mean function (NN SCBO), and two variants of our proposed model, namely *2-layer MGP-CBO* and *3-layer MGP-CBO*, respectively. We test the six aforementioned models on two 2-D synthetic optimization problems. We include details of the two simulations in the following two subsections.

A. 2-D Synthetic Optimization Problem

A 2-D constrained synthetic optimization problem [10] is used to evaluate SCBO, Polynomial NSCBO, ArcCosine NSCBO, NN SCBO, 2-layer MGP-CBO, and 3-layer MGP-CBO constrained Bayesian optimization benchmark models. The synthetic objective and constraint functions are shown in Equation 8.

$$\begin{aligned} f(x) &= \cos(2x_1)\cos(x_2) + \sin(x_1), \\ c(x) &= \cos(x_1)\cos(x_2) - \sin(x_1)\sin(x_2) - 0.5 \leq 0. \end{aligned} \quad (8)$$

where $f(x)$ represents the objective function and $c(x)$ represents the constraint function. We refer to x as the vector containing the two design variables x_1 and x_2 . The design variables x_1 and $x_2 \in [0, 6]$. The true global minimum is $y_{min} = -2$.

1) *Experimental Setup*: To run the experiments, we make the following design choices: (a) we run the six optimization benchmark models for 100 iterations (i.e., objective function evaluations), (b) we use the kernel functions Matern5/2 for the SCBO model, Polynomial for Polynomial NSCBO, and ArcCosine for ArcCosine NSCBO, (c) we use a neural network mean function and a Matern5/2 kernel function for NN-CBO, (d) we use the Matern5/2 as the kernel function for the 2-layer MGP-CBO and 3-layer MGP-CBO models, (e) we fix the initial points for all benchmark models for fair comparison, (f) we run the experiments ten times.

TABLE II

RESULTS OF THE 2-D SYNTHETIC OPTIMIZATION PROBLEM. THE BOLD UNDERLINED RESULTS REPRESENT THE BEST OUTCOME, WHILE THE UNDERLINED RESULTS REPRESENT THE SECOND BEST OUTCOME IN EACH COMPARISON METRIC

Model	MSE	Constraint Violations	Optimization Time
SCBO	0.2430	29%	69.7702
Polynomial NSCBO	0.6874	40%	61.5974
ArcCosine NSCBO	0.5135	26%	83.4517
NN-CBO	0.0556	33%	187.0564
2-layer MGP-CBO	<u>0.0191</u>	18%	72.3802
3-layer MGP-CBO	0.0138	<u>19%</u>	73.0339

2) *Metrics*: For comparison purposes, we use three metrics, namely the mean squared error shown in Equation 4, the constraint violation rate shown in Equation 7, and the optimization time in seconds.

3) *Results*: We assess the performance of our proposed MGP-CBO physics-constrained Bayesian optimization model against the other four benchmark models in optimizing the 2-D constrained synthetic optimization problem shown in Equation 8. We compare between the six benchmark models using the three previously introduced metrics. We show the experimental results of the six benchmark models in Table II. From Table II, the 3-layer MGP-CBO achieves the lowest mean squared error followed by the 2-layer MGP-CBO. In terms of the constraints violations rate, the 2-layer and 3-layer MGP-CBO models yield the least violations rate followed by the ArcCosine NSCBO model. Although the Polynomial NSCBO benchmark model achieves the lowest total optimization time in seconds, the two variants of our proposed MGP-CBO model achieve comparable results. Fig. 5 shows the results of optimizing the 2-D synthetic constrained objective function when the six competing benchmarks are used. The proposed MGP-CBO significantly improves the performance of constrained Bayesian optimization with better estimated global minimum and lower constraint violations rate.

B. 2-D Constrained Branin-Hoo Function

The 2-D constrained Branin-Hoo optimization problem [31] is used to evaluate six constrained Bayesian optimization benchmark models, namely SCBO, Polynomial NSCBO, ArcCosine NSCBO, NN SCBO, 2-layer MGP-CBO, and 3-layer MGP-CBO. The constrained Branin-Hoo optimization problem has two design variables, x_1 and x_2 , respectively. The objective and constraint functions are shown in Equation 9.

$$\begin{aligned} f(x) &= a(x_2 - bx_1^2 + cx_1 - r)^2 + s(1 - t)\cos(x_1) + s, \\ c(x) &= (x_1 - 2.5)^2 + (x_2 - 7.5)^2 - 50 \leq 0, \end{aligned} \quad (9)$$

where $f(x)$ and $c(x)$ are the objective and constraint functions, respectively. The objective function constants are: $a = 1$, $b = \frac{5.1}{4\pi^2}$, $c = \frac{5}{\pi}$, $r = 6$, $s = 10$, and $t = \frac{1}{8\pi}$. We refer to x as the vector containing the two design variables x_1 and x_2 . The search spaces of the two design variables are $x_1 \in [-5, 10]$ and $x_2 \in [0, 15]$, respectively. The true global minimum is $y_{min} = 0.3979$.

1) *Experimental Setup*: To run the experiments, we make the following design choices: (a) we run the six optimization benchmark models for 200 iterations (i.e., objective function

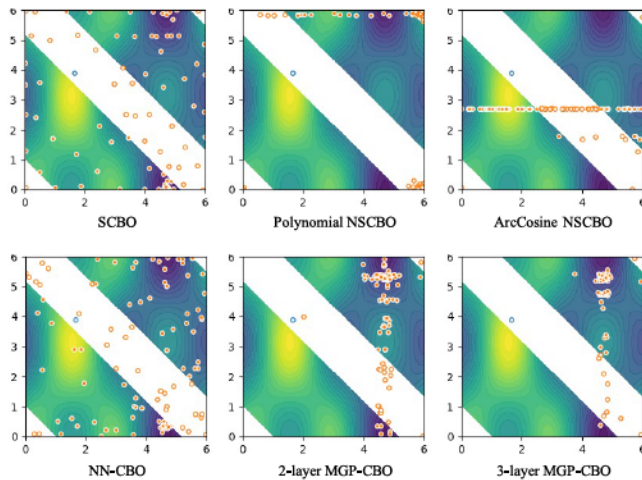


Fig. 5. Optimized constrained 2-D synthetic objective function for six constrained Bayesian optimization benchmark models. White areas represent infeasible regions. Blue circles are initial points and orange circles are data points acquired through constrained Bayesian optimization.

evaluations), (b) we use the kernel functions Exponential for the SCBO model, Polynomial for Polynomial NSCBO, and ArcCosine for ArcCosine NSCBO, (c) we use a neural network mean function and an Exponential kernel function for NN-CBO, (d) we use the Exponential as the kernel function for the 2-layer MGP-CBO and 3-layer MGP-CBO models, (e) we fix the initial points for all benchmark models for fair comparison, (f) we run the experiments five times.

2) *Results*: We assess the performance of our proposed MGP-CBO physics-constrained Bayesian optimization model against the other four benchmark models in optimizing the 2-D constrained Branin-Hoo optimization problem shown in Equation 9. We compare between the six benchmark models using the three previously introduced metrics. We show the experimental results of the six benchmark models in Table III. From Table III, the 3-layer MGP-CBO achieves the lowest MSE value followed by the NN-CBO model. In terms of the constraints violation rate, the 2-layer and 3-layer MGP-CBO models yield the least violation rates. Although the Polynomial NSCBO benchmark model achieves the lowest total optimization time in seconds, the MSE and constraints violation rate are very large. Although, the SCBO model is the second best in terms of the optimization time, the two variants of our proposed MGP-CBO model achieve comparable results. While the NN-CBO model achieves a competing MSE value, its optimization time is almost three times the optimization times of other benchmarks. Fig. 6 shows the optimization results of the six competing benchmarks in optimizing the 2-D Branin-Hoo constrained objective function. The proposed MGP-CBO significantly improves the performance of the constrained Bayesian optimization with better estimated global minimum and lower constraint violations.

VII. CASE STUDY: ACTUATORS PLACEMENT IN COMPOSITE STRUCTURES ASSEMBLY

Composite structures are gaining more and more popularity in advanced manufacturing processes. From aerospace to structural engineering, composite materials establish a strong

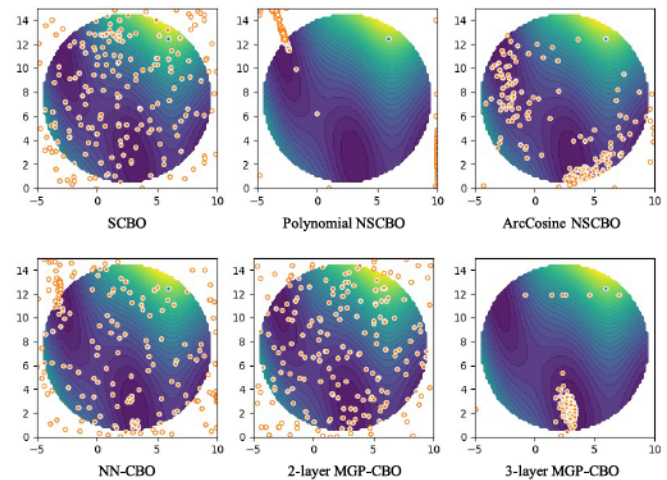


Fig. 6. Optimized constrained 2-D Branin-Hoo objective function for six constrained Bayesian optimization benchmark models. White areas represent infeasible regions. Blue circles are initial points and orange circles are data points acquired through constrained Bayesian optimization.

TABLE III

RESULTS OF THE 2-D BRANIN-HOO OPTIMIZATION PROBLEM. THE BOLD UNDERLINED RESULTS REPRESENT THE BEST OUTCOME, WHILE THE UNDERLINED RESULTS REPRESENT THE SECOND BEST OUTCOME IN EACH COMPARISON METRIC

Model	MSE	Constraint Violations	Optimization Time
SCBO	0.7788	31%	<u>133.5893</u>
Polynomial NSCBO	0.6618	76%	<u>120.0517</u>
ArcCosine NSCBO	0.2447	27%	154.0422
NN-CBO	<u>0.1141</u>	33%	425.5115
2-layer MGP-CBO	0.2479	20%	138.2206
3-layer MGP-CBO	<u>0.0251</u>	20%	137.7406

well-structured base that contributes to their light weight, strength, and stiffness, while excelling at resisting corrosion, fatigue, and wear. Whilst possessing those superior properties, composites have nonlinear and anisotropic characteristics. Composite structures assembly is very critical for the overall quality of final products. A composite fuselage, the main part of the aircraft holding passengers, crew, and cargo, is made mostly of Carbon fiber reinforced composite materials. Hence, careful consideration of the characteristics of composite materials when processing a fuselage part is important and extremely needed to increase productivity and improve the assembly quality of the aircraft.

A. Problem Definition

The shape and quality control of a composite fuselage is crucial for aerospace manufacturing. The quality of composite fuselages is highly dependent on the producer/supplier. A deviation in the design dimensions of a composite fuselage is inevitable, as shown in Fig. 7 (a). In current practices, actuator forces are placed equally-spaced along the lower edge of a fuselage to pull and push the fuselage in order to reduce the deviations to an acceptable value, as shown in Fig. 7 (b-c). However, those practices are far from being optimal and the placed actuators may exert forces more than needed at some locations. Hence, rapid and automatic shape control of composite structures assembly is paramount to ensure the

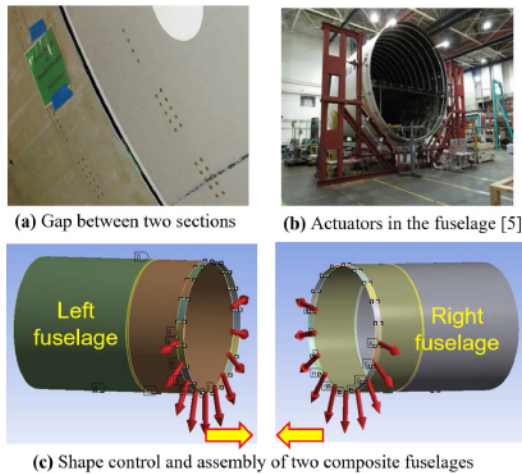


Fig. 7. Dimensional gap and shape control in composite structures assembly.

quality of individual fuselages and the entire assembly process. With the limited number of available actuators and the larger number of possible actuator locations, the optimal placement of actuators becomes a crucial process.

1) *Challenges*: Due to the inherited anisotropic characteristic of composite materials, the process of actuators placement on composite fuselages is very challenging. More particularly, with the different initial distortions of incoming fuselages and constraints on actuators forces and locations, the fuselage shape control process becomes extremely challenging.

2) *Problem Formulation*: We formulate our constrained optimization problem as follows: 1) we use the optimal placement of actuators response function [27] as the objective function, 2) we assume the following: we have N available measurement points on the edge of each fuselage and m available locations for the actuators to be placed, 3) we define the search space of each design variable and the set of constraint functions needed to govern the quality of the shape control process of the composite fuselage. The available m feasible actuator locations and N measurement points can be visualized in Fig. 8. The m feasible actuator locations are uniformly distributed from -12° (i.e., location 1) to 192° (i.e., location 18) as shown in Fig. 8. The proposed objective and constraint functions are as follows:

$$\delta = \Psi + UF, \quad (10)$$

$$F_{min} \leq F_j \leq F_{max}, \quad (11)$$

$$\sum_{j=1}^m \omega_j \leq M, \quad (12)$$

$$\omega_j = \begin{cases} 1 & \text{If an actuator is placed at location } j \\ 0 & \text{Otherwise,} \end{cases} \quad (13)$$

where δ represents the $(i \times 1)$ adjusted shape deviations vector, Ψ represents the $(i \times 1)$ initial shape distortions vector, U and $F = [\omega_1 F_1, \omega_2 F_2, \dots, \omega_j F_j]$ represent the $(i \times j)$ displacement matrix, where each element corresponds to the shape correction at measurement point i given a unit actuator force at location j , and $(j \times 1)$ actuator's force vector, respectively. We define i and j as the index of the measurement point and

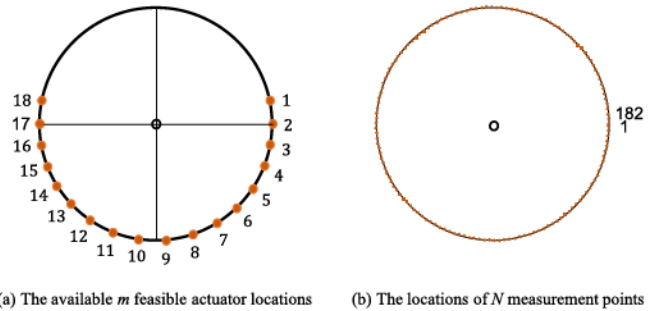


Fig. 8. Demonstration of the optimal shape control of composite fuselages process setup.

the location of the actuator, respectively, and ω as a binary variable that is equal to one if an actuator is placed at location j and zero otherwise. Equation 10 represents the objective function, while Equations 11, 12, and 13 represent the set of constraint functions.

The displacement matrix U is found by solving the least-squares regression model in [27] as follows:

$$Y_i = F_D \beta_i + e_i, \quad (14)$$

where Y_i and F_D represent the design shape distortion vector and actuator force matrix at each measurement point i , respectively. Y_i and F_D are found from the Design of Experiment (DOE) simulations. The vector e_i is the error term of the regression model and is assumed to follow a Gaussian distribution. The β_i vector in Equation 14 is the unknown regressor that represents the i^{th} element of the displacement matrix $U = [\beta_1, \beta_2, \dots, \beta_N]$.

B. Experimental Setup

We use the proposed physics-constrained Bayesian optimization model, MGP-CBO, to optimize the placement of actuators for the optimal shape control of composite fuselages. More specifically, we use the following settings: (a) we use the 2-layer MGP-CBO model to demonstrate the effectiveness of our proposed algorithm, (b) we run the MGP-CBO model for 1000 iterations, (c) we use one initial point prior to starting the optimization process, (d) we use 20 different initial distortion matrices to accommodate the naturally-occurring uncertainty in design deviations of incoming fuselages, (e) we use Matern5/2 as the kernel function of all the GP models. In this section, we denote the 2-layer MGP-CBO as the MGP-CBO model.

1) *Dataset*: In our optimization problem we have known variables that are inputs to our model and unknown decision variables (i.e., design variables) that we seek to find and optimize. The known variables are:

- N : The number of points where distortions are measured.
- m : The feasible locations of actuators.
- M : The maximum number of available actuators.
- Ψ : The vector of initial distortion matrices.
- U : The displacement matrix.
- F_{max} and F_{min} : The maximum and minimum allowed actuator forces. The unknown decision variables are:
- ω : A binary vector representing the location of actuators.
- F : A vector of the optimized actuator forces.

TABLE IV

RESULTS OF THE ACTUATORS PLACEMENT PROCESS FOR THE OPTIMAL SHAPE CONTROL OF COMPOSITE FUSELAGES. THE UNDERLINED RESULTS REPRESENT THE BEST OUTCOME IN EACH COMPARISON METRIC

Benchmark	Mean RMSD	Max RMSD	Mean MF	Max MF	Mean NA	Max NA
Best Fixed Actuator Placement	0.0034	0.0062	291.0714	449.6754	10	10
Sparse ADMM Optimal Placement	0.0028	0.0046	231.2973	397.9796	10	10
MGP-CBO	<u>0.0020</u>	<u>0.0042</u>	<u>178.8052</u>	<u>190.5369</u>	<u>8</u>	<u>9</u>

2) *Benchmarks*: For evaluation purposes, we compare our proposed MGP-CBO constrained Bayesian optimization model against two benchmarks, namely the best fixed actuator placement [25] and the sparse ADMM based optimal placement approach by [27].

3) *Metrics* : For comparison purposes, we use three main metrics, namely the root mean squared deviations denoted as *RMSD*, the maximum absolute forces of applied actuators (i.e., $\max \text{abs}(F_1, F_2, \dots, F_j)$) denoted as *MF*, and the total number of actuators used for the shape control process, denoted as *NA*.

C. Results

We evaluate our MGP-CBO model against the two benchmarks, best fixed actuator placement method and sparse ADMM based optimal placement method. We use the mean and maximum *RMSD*, the mean and maximum *MF*, and the mean and maximum *NA* to compare between those three benchmarks. The goal is to place a maximum of *M* actuators in *m* feasible locations to minimize the total deviations such that the forces of the placed actuators do not exceed the lower and upper limits of actuators forces (i.e., $F_{\min}$ and $F_{\max}$). The results are recorded in Table IV.

From Table IV, we show that our proposed MGP-CBO model outperforms the two benchmark models, best fixed actuator placement and sparse ADMM based optimal placement, respectively. More precisely, MGP-CBO achieves a reduction of 28.6% in mean *RMSD* and 8.7% in maximum *RMSD* in comparison with the sparse ADMM benchmark. Our proposed MGP-CBO model achieves another significant reduction of 22.7% in mean *MF* and 52.1% in maximum *MF* in comparison with the sparse ADMM benchmark. Since residual stresses are highly correlated with maximum forces, the remarkable reduction in maximum *MF*, when the proposed MGP-CBO model is used, is expected to significantly reduce the prospective residual stresses on a fuselage. Interestingly, to achieve these reductions in *RMSD* and *MF*, the MGP-CBO model needs only 8 actuators on average and a maximum of 9 actuators for the optimal shape control of composite fuselages.

VIII. CONCLUSION

Conventional constrained Bayesian optimization methods are inefficient in modeling complex nonstationary functions, especially when modeling physical systems with many constraints and design decision variables. Therefore, a *multi-layer deep Gaussian process based physics-constrained Bayesian optimization* algorithm (MGP-CBO) is proposed. The proposed MGP-CBO model is empowered with a multi-layer deep Gaussian process surrogate model that effectively models complex nonlinear and nonstationary functions. We provided the approximate posterior mean and covariance of the estimated

predictive distribution when using the proposed *multi-layer deep Gaussian process model* (MGP).

We also provided computational investigations that manifest the statistical properties of our proposed MGP-CBO model such as the convergence to the global optimum, constraint violations, and time complexity. We performed extensive evaluations that showcase the effectiveness and efficiency of our proposed MGP-CBO model. More precisely, we evaluated our proposed MGP-CBO model through 2-D simulation experiments and a real-world case study. We compared our proposed model to state-of-the-art stationary and nonstationary constrained Bayesian optimization models and found that the proposed MGP-CBO model significantly outperforms the other benchmarks. We also evaluated the proposed MGP-CBO model on the actuators placement in aerospace manufacturing and showed the outstanding performance of MGP-CBO in comparison to the best fixed actuator placement and sparse ADMM based optimal placement approaches.

REFERENCES

- [1] M. M. Dunlop, M. A. Girolami, A. M. Stuart, and A. L. Teckentrup, "How deep are deep Gaussian processes?" *J. Mach. Learn. Res.*, vol. 19, no. 54, pp. 1–46, 2018.
- [2] M. I. Radaideh and T. Kozłowski, "Surrogate modeling of advanced computer simulations using deep Gaussian processes," *Rel. Eng. Syst. Saf.*, vol. 195, Mar. 2020, Art. no. 106731.
- [3] A. Hebbal, L. Brevault, M. Balesdent, E.-G. Talbi, and N. Melab, "Bayesian optimization using deep Gaussian processes with applications to aerospace system design," *Optim. Eng.*, vol. 22, no. 1, pp. 321–361, Mar. 2021.
- [4] D. Eriksson and M. Poloczec, "Scalable constrained Bayesian optimization," in *Proc. Int. Conf. Artif. Intell. Statist.*, 2021, pp. 730–738.
- [5] D. Dries, P. Englert, and M. Toussaint, "Constrained Bayesian optimization of combined interaction force/task space controllers for manipulations," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2017, pp. 902–907.
- [6] A. Pal, L. Zhu, Y. Wang, and G. G. Zhu, "Constrained surrogate-based engine calibration using lower confidence bound," *IEEE/ASME Trans. Mechatronics*, vol. 26, no. 6, pp. 3116–3127, Dec. 2021.
- [7] R.-R. Griffiths and J. M. Hernández-Lobato, *Constrained Bayesian Optimization for Automatic Chemical Design Using Variational Autoencoders*, vol. 11, no. 2. London, U.K.: Royal Society of Chemistry, 2020, pp. 577–586.
- [8] Y. Wen, X. Yue, J. H. Hunt, and J. Shi, "Feasibility analysis of composite fuselage shape control via finite element analysis," *J. Manuf. Syst.*, vol. 46, pp. 272–281, Jan. 2018.
- [9] T. W. Clyne and D. Hull, *An Introduction to Composite Materials*. Cambridge, U.K.: Cambridge Univ. Press, 2019.
- [10] J. R. Gardner, M. J. Kusner, Z. Xu, K. Q. Weinberger, and J. Cunningham, "Bayesian optimization with inequality constraints," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2014, pp. 1–9.
- [11] J. M. Hernández-Lobato, M. Gelbart, M. W. Hoffman, R. P. Adams, and Z. Ghahramani, "Predictive entropy search for Bayesian optimization with unknown constraints," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2015, pp. 1699–1707.
- [12] A. Damianou and N. D. Lawrence, "Deep Gaussian processes," in *Proc. 16th Int. Conf. Artif. Intell. Statist.*, 2013, pp. 207–215.
- [13] Y. Wang, M. Wang, A. AlBahar, and X. Yue, "Nested Bayesian optimization for computer experiments," 2021, *arXiv:2112.09797*.
- [14] J. Ko and H. Kim, "Deep Gaussian process models for integrating multifidelity experiments with nonstationary relationships," *IJSE Trans.*, vol. 54, no. 7, pp. 686–698, 2021.

- [15] S. Remes, M. Heinonen, and S. Kaski, "Non-stationary spectral kernels," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 4645–4654.
- [16] C. Paciorek and M. Schervish, "Nonstationary covariance functions for Gaussian process regression," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 16, Mar. 2004, pp. 1–8.
- [17] H. Xue, Z.-F. Wu, and W.-X. Sun, "Deep spectral kernel learning," in *Proc. 28th Int. Joint Conf. Artif. Intell. (IJCAI)*, Jul. 2019, pp. 4019–4025, doi: [10.24963/ijcai.2019/558](https://doi.org/10.24963/ijcai.2019/558).
- [18] J. Snoek, K. Swersky, R. Zemel, and R. Adams, "Input warping for Bayesian optimization of non-stationary functions," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2014, pp. 1674–1682.
- [19] B.-J. Lee, J. Lee, and K.-E. Kim, "Hierarchically-partitioned Gaussian process approximation," in *Proc. 20th Int. Conf. Artif. Intell. Statist.*, vol. 54, Apr. 2017, pp. 822–831.
- [20] P. Palar, L. Zuhail, and K. Shimoyama, "Gaussian process surrogate model with composite kernel learning for engineering design," *AIAA J.*, vol. 58, pp. 1–17, Feb. 2020.
- [21] S. Ba and V. R. Joseph, "Composite Gaussian process models for emulating expensive functions," *Ann. Appl. Statist.*, vol. 6, no. 4, pp. 1838–1860, Dec. 2012.
- [22] M. Heinonen, H. Mannerström, J. Rousu, S. Kaski, and H. Lähdesmäki, "Non-stationary Gaussian process regression with Hamiltonian Monte Carlo," in *Proc. 19th Int. Conf. Artif. Intell. Statist.*, 2016, pp. 732–740.
- [23] M. van der Wilk, C. Rasmussen, and J. Hensman, "Convolutional Gaussian processes," in *Proc. Adv. Neural Inf. Process. Syst.* Red Hook, NY, USA: Curran Associates, 2017, pp. 2845–2854.
- [24] C. Lee, J. Wu, W. Wang, and X. Yue, "Neural network Gaussian process considering input uncertainty for composite structure assembly," *IEEE/ASME Trans. Mechatronics*, vol. 27, no. 3, pp. 1267–1277, Jun. 2022.
- [25] X. Yue, Y. Wen, J. H. Hunt, and J. Shi, "Surrogate model-based control considering uncertainties for composite fuselage assembly," *J. Manuf. Sci. Eng.*, vol. 140, no. 4, Apr. 2018, Art. no. 041017.
- [26] X. Yue and J. Shi, "Surrogate model-based optimal feed-forward control for dimensional-variation reduction in composite parts' assembly processes," *J. Qual. Technol.*, vol. 50, no. 3, pp. 279–289, 2018.
- [27] J. Du, X. Yue, J. H. Hunt, and J. Shi, "Optimal placement of actuators via sparse learning for composite fuselage shape control," *J. Manuf. Sci. Eng.*, vol. 141, no. 10, pp. 1–12, Oct. 2019.
- [28] Y. Jiao and D. Djurdjanovic, "Allocation of flexible tooling for optimal stochastic multistation manufacturing process quality control," in *Proc. Int. Manuf. Sci. Eng. Conf.*, vol. 48524, 2008, pp. 161–169.
- [29] J. Knoblauch, "Robust deep Gaussian processes," 2019, [arXiv:1904.02303](https://arxiv.org/abs/1904.02303).
- [30] D. R. Jones, M. Schonlau, and W. J. Welch, "Efficient global optimization of expensive black-box functions," *J. Global Optim.*, vol. 13, no. 4, pp. 455–492, 1998.
- [31] M. A. Gelbart, J. Snoek, and R. P. Adams, "Bayesian optimization with unknown constraints," in *Proc. 13th Conf. Uncertainty Artif. Intell.*, 2014, pp. 250–259.



Areej AlBahar received the B.S. degree in industrial and management systems engineering from Kuwait University, Kuwait, in 2012, the M.S. degree in industrial and systems engineering from the University of Florida, Gainesville, USA, in 2016, and the Ph.D. degree in industrial and systems engineering from Virginia Tech, Blacksburg, USA, in 2022.

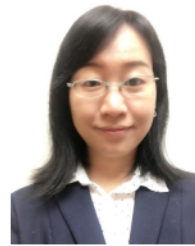
She is currently an Assistant Professor with the Department of Industrial and Management Systems Engineering, Kuwait University. Her research interests include machine learning and advanced statistics

for monitoring, optimization, and quality improvement of manufacturing processes.



Inyoung Kim received the B.S. degree in mathematics from Jeju National University, Jeju, South Korea, in 1994, the M.S. degree in applied statistics from Yonsei University, Seoul, South Korea, in 1996, the Ph.D. degree in statistics from Texas A&M University, College Station, USA, in 2002, and the Ph.D. degree in biostatistics from Yale University, New Haven, USA, in 2015.

She is currently an Associate Professor with the Department of Statistics, Virginia Tech, Blacksburg, USA. Her research interests are focused on developing statistical machine learning and semi/nonparametric statistical methods and theory for high dimensional analysis. She has developed both Frequentist and Bayesian methods. She was a recipient of the Best Paper Award from *Biometrics* in 2016. She is a member of ASA, IBS, and ISBA.



Xing Wang received the B.S. degree in mathematics from Tsinghua University in 2013, the M.S. degree in statistics from the Georgia Institute of Technology in 2014, and the Ph.D. degree in risk management and insurance from the J. Mack Robinson College of Business, Georgia State University in 2018.

She is currently an Assistant Professor with the Department of Mathematics, Illinois State University. Her work focused on statistical inference about risk measures under extreme scenarios. She also served as an instructor for the course risk modeling during her Ph.D. period. Her research interests include risk measure, extreme value theory, and statistical inference. She was the James C. Hickman Scholar. She was a recipient of the FTC Early Career Award and the ASA Early Career Travel Award.



Xiaowei Yue (Senior Member, IEEE) received the B.S. degree in mechanical engineering from the Beijing Institute of Technology, Beijing, China, in 2011, the M.S. degree in power engineering and engineering thermophysics from Tsinghua University, Beijing, in 2013, and the M.S. degree in statistics and the Ph.D. degree in industrial engineering with minor machine learning from the Georgia Institute of Technology, Atlanta, USA, in 2016 and 2018, respectively.

He is currently an Assistant Professor with the Grado Department of Industrial and Systems Engineering, Virginia Tech, Blacksburg, USA. He is also a DoD MEEP Faculty Fellow. His research interests are focused on machine learning for advanced manufacturing.

Dr. Yue is a Senior Member of ASQ and IISE. He is also a member of ASME and SME. He was a recipient of more than ten best paper awards and two best dissertation awards. He was also a recipient of the Sandra Bouckley Outstanding Young Manufacturing Engineer Award from SME and the Grainger Frontiers of Engineering Grant Award from the National Academy of Engineering. He has won several teaching awards (e.g., the Large Class Teaching Award) from the Center for Excellence in Teaching and Learning, Virginia Tech. He serves as an Associate Editor for the *Journal of Intelligent Manufacturing* and the *IISE Transactions*.