Journal Article

An Efficient Method for Antenna Design Based on a Self-Adaptive Bayesian Neural Network-Assisted Global Optimization Technique

Liu, Y., Liu, B., Ur-Rehman, M., Imran, M. A., Akinsolu, M.O., et al.

An Efficient Method for Antenna Design Based on a Self-Adaptive Bayesian Neural Network-Assisted Global Optimization Technique

Yushi Liu, *Student Member, IEEE*, Bo Liu, *Senior Member, IEEE*, Masood Ur-Rehman, *Senior Member, IEEE*, Muhammad Ali Imran, *Senior Member, IEEE*, Mobayode O. Akinsolu, *Senior Member, IEEE*, Peter Excell, *Life Senior Member, IEEE*, and Qiang Hua

***Abstract*— Gaussian process (GP) is a very popular machine learning method for online surrogate-model-assisted antenna design optimization. Despite many successes, two improvements are important for the GP-based antenna global optimization methods, including: 1) the convergence speed (i.e., the number of necessary electromagnetic (EM) simulations to obtain a high-performance design) and 2) the GP model training cost when there are several tens of design variables and/or specifications. In both aspects, the state-of-the-art GP-based methods show practical but not desirable performance. Therefore, a new method, called the self-adaptive Bayesian neural network surrogate-model-assisted differential evolution (DE) for antenna optimization (SB-SADEA), is presented in this article. The key innovations include: 1) the introduction of the Bayesian neural network (BNN)-based antenna surrogate modeling method into this research area, replacing GP modeling, and 2) a bespoke self-adaptive lower confidence bound (LCB) method for antenna design landscape making use of the BNN-based antenna surrogate model. The performance of SB-SADEA is demonstrated by two challenging design cases, showing considerable improvement in terms of both convergence speed and machine learning cost compared with the state-of-the-art GP-based antenna global optimization methods.**

***Index Terms*— Antenna design, antenna optimization, Bayesian neural network (BNN), computationally expensive optimization, differential evolution (DE), lower confidence bound (LCB), surrogate modeling.**

I. Introduction

EVOLUTIONARY algorithms (EAs) are widely used in antenna design [1], [2]. Due to their ability to jump out of local optima, without the need of an initial design and generality, they are showing advantages for many design cases. The differential evolution (DE) [3] and particle swarm optimization (PSO) [4] algorithms are arguably playing the leading role in EA-driven antenna design [2], [5], [6]. However, considering full-wave electromagnetic (EM) simulations are often needed to obtain accurate performance of a candidate design, and EAs often need tens of thousands of such EM simulations to obtain the optimal design, the optimization time can be prohibitive.

Therefore, surrogate models, constructed by the machine learning techniques, are trained to approximate the antenna performance obtained by EM simulations [7]. Hence, in the optimization process, many computationally expensive EM simulations can be replaced by computationally cheap surrogate model predictions. The optimization time can, therefore, be largely reduced. Note that this research focuses on online surrogate-model-assisted antenna global optimization, where the surrogate model keeps being updated in each iteration.

Three critical factors for online surrogate-model-assisted antenna global optimization are the surrogate modeling method, the search operators, and the model management method. The surrogate modeling method refers to the machine learning core and its associated bespoke operators for antennas. Search operators refer to the optimization engine. The model management method refers to the framework making surrogate modeling and optimization work harmoniously. Since prediction uncertainty is unavoidable, which may lead to wrong convergence, identifying high potential candidate designs under uncertainty to maintain correct convergence and optimal update of the surrogate model are the goals of model management. The three factors are strongly interconnected.

Among online antenna global optimization methods, the surrogate-model-assisted DE for antenna optimization (SADEA) series [8], [9], [10], [11] is one of the state-of-the-art approaches. For design cases with fewer than 20 design variables and a few specifications, the latest P-SADEA method [10], [12] integrates three DE mutation operators with different characteristics, which are organized by a new model management method. It shows success in challenging antenna cases where DE and PSO fail to obtain feasible designs, e.g., [13] while improving the speed by up to 30 times without parallel computing. For antenna cases with several tens of design variables and/or specifications, machine learning cost becomes a new challenge, which may be even higher than the EM simulation cost [11]. To the best of our knowledge,

Yushi Liu, Bo Liu, Masood Ur-Rehman, and Muhammad Ali Imran are with the James Watt School of Engineering, University of Glasgow, G12 8LU Glasgow, U.K. (e-mail: y.liu.8@research.gla.ac.uk; bo.liu@glasgow.ac.uk; masood.urrehman@glasgow.ac.uk; muhammad.imran@glasgow.ac.uk).

Mobayode O. Akinsolu and Peter Excell are with the Faculty of Computing and Engineering, Wrexham Glyndwr University, LL11 2AW Wrexham, U.K. (e-mail: m.o.akinsolu@glyndwr.ac.uk; p.excell@glyndwr.ac.uk).

Qiang Hua is with the Faculty of Engineering and Technology, Liverpool John Moores University, L69 3GJ Liverpool, U.K. (e-mail: q.hua@ljmu.ac.uk).

the TR-SADEA method first addresses this challenge by a Gaussian process (GP) model sharing method to reduce the necessary number of GP models and a self-adaptive surrogate-model-assisted local optimization scheme to improve the convergence speed. It shows success for complex base station antennas [11].

In the evolution of the SADEA series, the main innovations locate in the model management method and the search operators. Considering the surrogate modeling method, to the best of our knowledge, not only the SADEA series but also GP machine learning is used by most online state-of-the-art surrogate-model-assisted antenna optimization methods [14], [15], [16], [17]. For the often highly multimodal antenna design landscape [10], a strong machine learning method is essential. GP stands out due to its strong learning ability, with very few empirical parameters, and can provide a rigorous prediction uncertainty for each candidate design.

However, GP has its drawbacks, i.e., its training cost [18]. GP's training time grows cubically with the number of training data points, which is highly related to the number of design variables and specifications. Considering normal desktop computers (e.g., Intel i7 3.0 GHz CPU), for antennas with a few design variables, GP training time is short. However, for antennas with more than 20 and even around 50 design variables with more than 10 specifications, TR-SADEA uses one to two days for GP model training [11] (note that 90% of the expected training time compared with standard SADEA is reduced). This training cost is practical but not desirable.

For design cases with very challenging specifications, e.g., [13], most surrogate-model-assisted antenna global optimization methods still need a few thousand EM simulations, sometimes costing more than a week. Hence, much space is left for improving the convergence speed (i.e., the necessary number of EM simulations). The long GP training time restricts the number of surrogate models to be built, which affects the convergence speed. For example, an idea like deep supervision [19], [20] in image recognition is suitable for many antenna cases to improve the convergence speed at the cost of training more surrogate models. However, considering GP modeling time, it is difficult to be used for many cases whose number of design variables/specifications is not small.

Owing to this, this research aims to seek a new machine learning core to replace the GP-based antenna modeling and propose a method using the new surrogate model. The goal is to considerably improve both the convergence speed and training cost and provide a universal method for antennas with various numbers of design variables and specifications. Hence, a new method, called the self-adaptive Bayesian neural network surrogate-model-assisted DE for antenna optimization (SB-SADEA), is proposed. The key innovations include: 1) the introduction of the Bayesian neural network (BNN)-based antenna surrogate modeling method into this area to replace GP and 2) a bespoke self-adaptive lower confidence bound (LCB) method for antenna design landscape making use of the BNN results as a part of model management.

The remainder of this article is organized as follows. Section II presents the background knowledge. Section III elaborates on the SB-SADEA method. Section IV presents the advantages of SB-SADEA using a compact ultrawideband (UWB) monopole antenna and a compact multiband 5G mm-wave antenna, both with challenging specifications. The concluding remarks are provided in Section V.

II. Background Knowledge

A. Online Surrogate-Model-Assisted Antenna Design Optimization

Machine learning is attracting much attention in EM design recently [21], and surrogate-model-assisted antenna optimization is an essential part, which can be classified into offline and online. In the offline methods, a high-quality surrogate model is first built and there are no or few updates of the surrogate model in the optimization process [22]. The advantage for the offline methods is that the resulting effective antenna surrogate model is useful on many occasions, e.g., antenna circuit co-design and multiobjective Pareto optimization, showing excellent results [14], [23]. Particularly, some state-of-the-art machine learning techniques are proposed for EM device modeling and inverse design (e.g., Fourier subspace-based deep learning [23]) in recent years, addressing important bottlenecks.

The limitation for the offline methods is the "curse of dimensionality" [24]. When there are more than a few design variables and/or the modeling range is not narrow, the necessary number of EM simulations needed for preparing sufficient training data points to build a high-quality surrogate model could be many, which grows exponentially with the number of design variables. If an ad hoc training data preparation is needed, carrying out those simulations could cost tremendous time, canceling out the time saved using the surrogate models. Hence, design optimization is often a by-product of state-of-the-art antenna surrogate modeling methods, or the antenna/EM device model is largely reusable.

Online methods, in contrast, keep improving the surrogate model quality in the optimization process. In each iteration, (a) surrogate model(s) is/are built using the available simulated candidate designs. New candidate designs are generated by search operators, which are then predicted by the surrogate model. Using the prediction result, candidate designs with high potential are simulated and used to update the surrogate model for the next iteration. Hence, the surrogate model quality is not always high but is gradually improved. Particularly, in the beginning stage, the surrogate model quality may be poor due to the lack of training data points.

Unlike assuming an accurate surrogate model like offline methods, prediction quality and uncertainty largely affect the optimization. When the prediction quality is insufficient and the prediction uncertainty is not appropriately considered, the optimization is highly likely to converge to a local optimum far away from the design specifications [24], [25]. Therefore, appropriate collaboration of the surrogate modeling method, the search operators, and the model management method (i.e., the three key factors in Section I) is essential.

An important task in model management is to co-use the prediction uncertainty together with the predicted value to judge the potential of a candidate antenna design. This is also

called the prescreening or acquisition function [26], which will be discussed in Section II-B. Explicitly providing the prediction uncertainty for each candidate design (instead of the overall uncertainty of the surrogate model) is necessary for most prescreening methods.

B. GP Modeling and Prescreening, Advantages and Drawbacks

The basic principle of GP is as follows [18]. In the following, superscript refers to the index of a sample in the dataset and subscript refers to the index of a variable in a sample. Giving n observations [$\boldsymbol{x} = (\boldsymbol{x^1}, \ldots, \boldsymbol{x^n})$ and $\boldsymbol{y} = (y^1, \ldots, y^n)$], GP assumes that $y(x)$ is a sample of a Gaussian distributed stochastic process with mean μ and variance σ. GP then predicts the value of $y(\boldsymbol{x})$ for a new $\boldsymbol{x}$ using its relation with the n observations. For example, a correlation function can be described as follows:

$$\begin{aligned} \mathrm{Corr}\left(\boldsymbol{x^i}, \boldsymbol{x^j}\right) &= \exp\left(-\sum_{l=1}^{d} \theta_l |x_l^i - x_l^j|^{p_l}\right) \\ \theta_l &> 0, \quad 1 \le p_l \le 2 \end{aligned} \tag{1}$$

where d is the dimension of $\boldsymbol{x}$. θ_l and p_l are the hyperparameters, which are determined by maximizing the likelihood function in the following equation:

$$\frac{1}{\left(2\pi\sigma^2\right)^{n/2}\sqrt{\det(\boldsymbol{R})}} \exp\left[-\frac{(\boldsymbol{y}-\mu\boldsymbol{I})^T\boldsymbol{R}^{-1}(\boldsymbol{y}-\mu\boldsymbol{I})}{2\sigma^2}\right] \tag{2}$$

where $\boldsymbol{R}$ is a $n \times n$ covariance matrix, and $\boldsymbol{I}$ is a $n \times 1$ vector having all its elements as unity. By maximizing the likelihood function that $y = y^i$ at $\boldsymbol{x} = \boldsymbol{x^i}(i = 1, \ldots, n)$ and handling the prediction uncertainty based on the best linear unbiased prediction, for a new point $\boldsymbol{x^*}$, the predicted value and prediction uncertainty are $\hat{y}(\boldsymbol{x^*})$ and $\hat{s}^2(\boldsymbol{x^*})$, respectively, which are as follows:

$$\hat{y}\left(\boldsymbol{x^*}\right) = \mu + \boldsymbol{r}^T\boldsymbol{R}^{-1}(\boldsymbol{y} - \boldsymbol{I}\mu) \tag{3}$$

where

$$\begin{aligned} R_{i,j} &= \mathrm{Corr}\left(\boldsymbol{x^i}, \boldsymbol{x^j}\right), \quad i, j = 1, 2, \ldots, n \\ \boldsymbol{r} &= \left[\mathrm{Corr}\left(\boldsymbol{x^*}, \boldsymbol{x^1}\right), \mathrm{Corr}\left(\boldsymbol{x^*}, \boldsymbol{x^2}\right), \ldots, \mathrm{Corr}\left(\boldsymbol{x^*}, \boldsymbol{x^n}\right)\right] \\ \hat{\mu} &= \left(\boldsymbol{I}^T\boldsymbol{R}^{-1}\boldsymbol{y}\right)\left(\boldsymbol{I}^T\boldsymbol{R}^{-1}\boldsymbol{I}\right)^{-1} \end{aligned} \tag{4}$$

$$\hat{s}^2\left(\boldsymbol{x^*}\right) = \hat{\sigma}^2\left[1 - \boldsymbol{r}^T\boldsymbol{R}^{-1}\boldsymbol{r} + \left(\boldsymbol{I} - \boldsymbol{r}^T\boldsymbol{R}^{-1}\boldsymbol{r}\right)^2\left(\boldsymbol{I}^T\boldsymbol{R}^{-1}\boldsymbol{I}\right)^{-1}\right] \tag{5}$$

where

$$\hat{\sigma}^2 = (\boldsymbol{y} - \boldsymbol{I}\hat{\mu})^T\boldsymbol{R}^{-1}(\boldsymbol{y} - \boldsymbol{I}\hat{\mu})n^{-1}. \tag{6}$$

Based on the above, two advantages of GP include: 1) there are almost no empirical parameters in GP modeling except deciding the type of correlation function; A few appropriate correlation functions are already found by antenna surrogate modeling researchers [8], [14], [15], [16]. Hence, overfitting or underfitting like artificial neural networks (ANNs) is less likely to happen, which improves the prediction quality and 2) the prediction uncertainty (5) is statistically grounded, which can play an important role when judging the full potential of a candidate antenna design in prescreening or acquisition function.

With prediction uncertainty, the widely used prescreening methods include expected improvement [26], probability of improvement [27], and LCB [28]. LCB is the fundamental of the new prescreening method in this article and is introduced as follows. Given the objective function $y(\boldsymbol{x})$ has a predictive distribution of $N(\hat{y}(\boldsymbol{x}), \hat{s}^2(\boldsymbol{x}))$, an LCB prescreening of $y(\boldsymbol{x})$ is

$$\begin{aligned} y_{\mathrm{lcb}}(\boldsymbol{x}) &= \hat{y}(\boldsymbol{x}) - \omega\hat{s}(\boldsymbol{x}) \\ \omega &\in [0, 3] \end{aligned} \tag{7}$$

where ω is a constant and is often set to 2 in many algorithms in the AI domain [24], and it is applicable to antenna problems [8].

However, the main drawback of GP is its training cost. In online surrogate-model-assisted antenna global optimization, the total training time of GP models in the optimization process can be estimated as $T_{GP} \times N_{specs} \times N_{pop} \times N_{it}$ [11], where T_{GP} is the training time of each GP model, N_{specs} is the number of specifications, N_{pop} is the number of candidate designs in a population, and N_{it} is the number of iterations in antenna design optimization.

For a GP model, the computational complexity is $O(N_o n^3 d)$ [24], where N_o is the number of iterations spent in hyperparameter optimization (i.e., (2)) and n is the number of training data points. n is highly affected by d to construct a reliable surrogate model. Liu et al. [8] and [11] show that at least $4 \times d$ training data points are needed for antenna problems. Often, when d reaches 20, T_{GP} could be in minutes for a normal computer and then grows cubically [11]. Also, to maintain the exploration ability, N_{pop} is also highly affected by d (e.g., often at least $4 \times d$ when using DE operators). This makes the GP modeling time in antenna optimization become long when d is large and could be even longer than the EM simulation time.

To the best of our knowledge, TR-SADEA [11] is the first method addressing this challenge. A self-adaptive GP model sharing method is proposed aiming to highly decrease the number of GP modeling while maintaining correct convergence. Although this method decreases the GP model training time to be practical, the time consumption (e.g., one to two days for complex base station antennas with several tens of design variables and specifications) is not desirable.

C. Antenna Design Optimization and the DE Algorithm

Antenna design optimization can be classified into local optimization [29] and global optimization [1], [6]. The antenna local optimization methods require an initial design, aiming to find an optimal design around it, while for the antenna global optimization methods, only a reasonable search range is needed. Antenna global optimization does not guarantee to obtain the global optimal design. It refers to not searching around the initial design and having the search mechanism to jump out of local optima in the design landscape.

In antenna global optimization, EAs were introduced into this area two decades ago and are widely used. Multistart local search and surrogate-model-assisted multistart local search for antenna global optimization are introduced recently [17], [30], showing successful results. This article focuses on using EAs. As said in Section I, DE and PSO are arguably the most widely used EAs in antenna global optimization [2], [5], [31], [32], and DE operators are used in this work. A brief introduction to DE is as follows [3].

$\boldsymbol{P}$ is a population composed of N_{pop} individual solutions, and each solution is denoted by $\boldsymbol{x} = (x_1, \ldots, x_d) \in R^d$. To create a child solution $\boldsymbol{u} = (u_1, \ldots, u_d)$, first, mutation happens to generate a donor vector

$$\boldsymbol{v}^i = \boldsymbol{x}^i + F \cdot \left(\boldsymbol{x}^{best} - \boldsymbol{x}^i\right) + F \cdot \left(\boldsymbol{x}^{r_1} - \boldsymbol{x}^{r_2}\right) \tag{8}$$

where $\boldsymbol{x}^i$ is the ith vector in the current population; $\boldsymbol{x}^{best}$ is the best candidate solution in the current population $\boldsymbol{P}$; $\boldsymbol{x}^{r_1}$ and $\boldsymbol{x}^{r_2}$ are two mutually exclusive solutions randomly selected from $\boldsymbol{P}$; $\boldsymbol{v}^i$ is the ith mutant vector; $F \in (0, 2]$ is the scaling factor. The mutation strategy in (8) is called DE/current-to-best/1, which is used in SB-SADEA.

The crossover operator then happens to produce the child solution $\boldsymbol{u}$.

1 Randomly select a variable index $j_{rand} \in \{1, \ldots, d\}$,

2 For each $j = 1$ to d, generate a uniformly distributed random number $rand$ from (0, 1) and set

$$u_j = \begin{cases} v_j, & \text{if } (rand \le CR) | j = j_{rand} \\ x_j, & \text{otherwise} \end{cases} \tag{9}$$

where $CR \in [0, 1]$ is the crossover rate.

III. SB-SADEA Algorithm

A. Algorithm Framework

As discussed in Section I, this research aims to seek a new machine learning core to replace GP and propose a new prescreening method for it. The resulting SB-SADEA algorithm is expected to significantly improve both the convergence speed (i.e., the number of EM simulations used to obtain the optimal design) and the surrogate model training cost compared with the GP-based methods. Also, it should be universal for antenna cases with various numbers of design variables and specifications. In the following, the general framework of SB-SADEA is first provided, and the details of the two innovations are then described in Sections III-B and III-C.

The SB-SADEA framework is shown in Fig. 1, and the algorithm works as follows.

1) *Step 1:* Sample α (often a small number of) candidate designs from the design space $[\boldsymbol{LB}, \boldsymbol{UB}]^d$ ($\boldsymbol{LB}$ and $\boldsymbol{UB}$ are the lower and upper bounds of design variables, respectively) using Latin hypercube sampling (LHS) [33]. Carry out EM simulations to obtain their performance values and form the initial database.
2) *Step 2:* If a preset stopping criterion is met (e.g., the computing budget is exhausted, satisfies the specifications), output the best candidate design from the database; otherwise, go to Step 3.

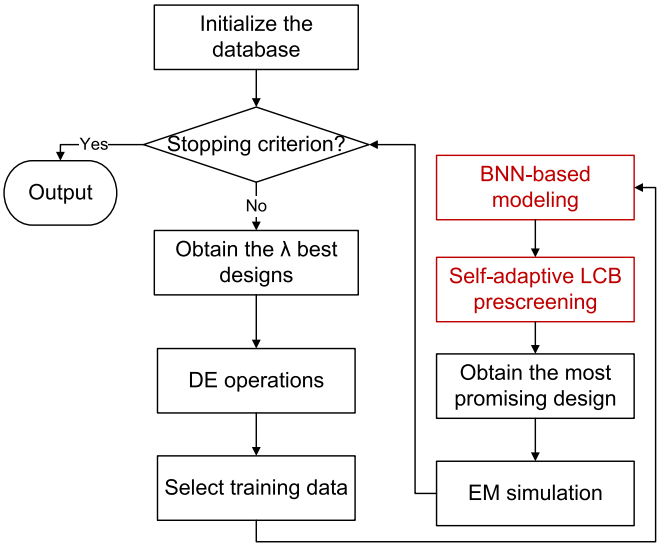

Fig. 1. Flow diagram of SB-SADEA.

3) *Step 3:* Obtain the λ best candidate designs from the database to form a population $\boldsymbol{P}$.
4) *Step 4:* Apply the DE/current-to-best/1 operator (8) to P to create λ new child solutions.
5) *Step 5:* For each child solution, obtain τ nearest samples (based on the Euclidean distance) as the training data points and construct a BNN-based surrogate model (see Section III-B).
6) *Step 6:* Prescreen the child solutions generated in Step 4 using the BNN model and the self-adaptive LCB method (see SectIII-C).
7) *Step 7:* Carry out EM simulation to the estimated best child solution from Step 6. Add this evaluated candidate design and its performance values to the database. Go back to Step 2.

It can be seen that some model management operators are borrowed from standard SADEA [8]. This model management method is attracting much attention in the AI domain [22] and its advantages are detailed in [8], [10], and [22]. The two novel methods, including the BNN-based antenna surrogate modeling (Step 5) and the self-adaptive LCB method (Step 6), which are red blocks in Fig. 1 are introduced in the following subsections. Note that they are compatible with model management frameworks in other SADEA versions and some other online antenna global optimization methods.

B. BNN-Based Antenna Surrogate Modeling Method

To replace GP modeling, the machine learning core must satisfy the following requirements: 1) can provide a high-quality predicted value comparable to GP and has less risk to be overfitted; 2) the prediction uncertainty of each candidate design is statistically grounded; and 3) the training cost is much lower than GP when the targeted antenna has many design variables/specifications.

To meet the above requirements, an alternative is a stochastic neural network, in particular, a BNN. Many fitting-based

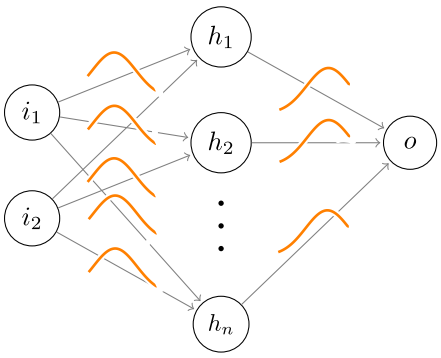

Fig. 2. Illustrative figure of a basic BNN.

machine learning methods, such as most kinds of ANN, radial basis function [34], and response surface models, either cannot provide a prediction uncertainty for each candidate design (not the uncertainty of the whole surrogate model) or the prediction uncertainty is not statistically grounded. However, a statistically grounded prediction uncertainty for each candidate design is essential for prescreening or acquisition function (e.g., (7)). They are, therefore, less suitable for online surrogate-model-assisted antenna global optimization because the acquisition function largely affects the convergence speed and the ability to jump out of local optima. Although researchers suggest using the dropout method in ANN training to provide a prediction uncertainty for each candidate solution [35], [36], [37], our experiments using antenna problems show much worse results than GP. BNN, however, can address this challenge. To the best of our knowledge, BNN has not been used in antenna design optimization and is seldom considered for surrogate-model-assisted optimization even in the AI domain. In the following, BNN concepts are introduced.

Considering the antenna design variables as $\boldsymbol{x}$, and the performance as $\boldsymbol{y}$, for an ANN, the model parameters are $\boldsymbol{\theta} = [w_1, \ldots, w_j, b_1, \ldots, b_k]$, where $\boldsymbol{w}$ are the weights and $\boldsymbol{b}$ are the biases. In a multilayer ANN, each layer is a linear transformation, followed by a nonlinear activation function. The training optimizes the cost function, which is often the log-likelihood of the training data points, i.e., maximize $\sum_{i=1}^{n} \log(p(\boldsymbol{x}^i; \boldsymbol{\theta}))$ with a regularization term. The optimized $\boldsymbol{\theta}$, which are fixed values, is then used for prediction.

For BNN, the network structure does not change compared with a standard ANN, but $\boldsymbol{\theta}$ become stochastic variables with their probability distribution $p(\boldsymbol{\theta})$. Fig. 2 shows an illustrative figure of BNN. i, h, and o represent the neurons on the input, hidden, and output layers of the BNN, respectively. In the following, the training set is denoted by $\boldsymbol{D}$, and the training inputs and training outputs are denoted by $\boldsymbol{D_x}$ and $\boldsymbol{D_y}$, respectively. By applying Bayes' theorem, the Bayesian posterior can be expressed as

$$p(\boldsymbol{\theta}|\boldsymbol{D}) = \frac{p(\boldsymbol{D_y}|\boldsymbol{D_x},\boldsymbol{\theta})p(\boldsymbol{\theta})}{\int_{\boldsymbol{\theta}} p(\boldsymbol{D_y}|\boldsymbol{D_x},\boldsymbol{\theta}')p(\boldsymbol{\theta}')\mathrm{d}\boldsymbol{\theta}'} \propto p(\boldsymbol{D_y}|\boldsymbol{D_x},\boldsymbol{\theta})p(\boldsymbol{\theta}) \tag{10}$$

where $p(\boldsymbol{D_y}|\boldsymbol{D_x},\boldsymbol{\theta})$ is the likelihood, $p(\boldsymbol{\theta})$ is the prior, the denominator is the evidence, and $p(\boldsymbol{\theta}|\boldsymbol{D})$ is the posterior. The posterior is what we acquire, which is used in obtaining the predicted value and prediction uncertainty. Obtaining $p(\boldsymbol{\theta}|\boldsymbol{D})$ by standard sampling method is intractable. Hence, the variational inference method [38] is proposed.

In variational inference, a new distribution $q(\phi)$ (ϕ are the model parameters), called a variational distribution, is proposed to approximate $p(\boldsymbol{\theta}|\boldsymbol{D})$. By minimizing the Kullback–Leibler (KL) divergence D_{KL} between $q(\phi)$ and $p(\boldsymbol{\theta}|\boldsymbol{D})$, the closest distribution can be found to replace the posterior. Compared with the posterior, $q(\phi)$ has a smaller set of parameters, which are often considered as means and variances of a multivariate Gaussian distribution, and are optimized in training.

The cost function to be maximized is

$$E_{\phi\sim q}\big(\log p\big(\boldsymbol{D_y}|\boldsymbol{D_x};\phi\big)\big) - D_{\mathrm{KL}}(q(\phi)||p(\boldsymbol{\theta})). \tag{11}$$

Equation (11) is called the evidence lower bound. The first term E represents the Shannon entropy, which means the sum of the expected log-likelihood of the data. The second term is the regularization loss represented by the KL divergence, which is a closed form for the Gaussian distribution. The first item can be obtained by sampling. After optimization or training, the posterior, $p(\theta|\boldsymbol{D})$, is approximated, and the BNN is ready to be used.

When performing prediction by BNN, given the posterior, $p(\boldsymbol{\theta}|\boldsymbol{D})$, the model's prediction uncertainty can be derived from $p(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{D})$. Mathematically, it can be written as

$$p(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{D}) = \sum_{\boldsymbol{\theta}} p\big(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{\theta}'\big)\, p\big(\boldsymbol{\theta}'|\boldsymbol{D}\big)\mathrm{d}\boldsymbol{\theta}'. \tag{12}$$

In practice, this is done by sampling [39]

$$\boldsymbol{\theta} \sim p(\boldsymbol{\theta}|\boldsymbol{D}). \tag{13}$$

The predicted value is the average of BNN model output samples

$$\hat{\boldsymbol{y}} = \frac{1}{|\boldsymbol{\Theta}|}\sum_{\theta_i\in\boldsymbol{\Theta}} \Phi_{\theta_i}(\boldsymbol{x}) \tag{14}$$

where $\boldsymbol{\Theta}$ is a set containing all $\boldsymbol{\theta}$, $|\boldsymbol{\Theta}|$ is the size of the set, $\Phi_{\boldsymbol{\theta}}(\boldsymbol{x})$ is the BNN model, and $\hat{\boldsymbol{y}}$ is the estimated output. The uncertainty quantification is given by the covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{y}|\boldsymbol{x},\boldsymbol{D}}$, which is

$$\boldsymbol{\Sigma}_{\boldsymbol{y}|\boldsymbol{x},\boldsymbol{D}} = \frac{1}{|\boldsymbol{\Theta}|-1}\sum_{\theta_i\in\boldsymbol{\Theta}} \big(\Phi_{\theta_i}(\boldsymbol{x}) - \hat{\boldsymbol{y}}\big)\big(\Phi_{\theta_i}(\boldsymbol{x}) - \hat{\boldsymbol{y}}\big)^T. \tag{15}$$

Some clarifications in terms of the requirements at the beginning of this section are as follows.

1) BNN has a good potential to provide high-quality prediction results. BNN can be interpreted as a special case of ensemble methods [40]. The ensemble methods are well-known for taking advantage of the fact that the aggregation of multiple averaged and independent predictors may outperform a single expert predictor, given the same training information [41]. BNN's stochastic components similarly improve a normal ANN. Also, BNN can avoid overfitting when learning from a small dataset (i.e., available training data points via EM

simulations) by considering both aleatoric uncertainty and epistemic uncertainty, as evidenced in [42].

2) The prediction uncertainty of BNN is statistically grounded [43], [44]. Intuitively, as in (15), for any input $\boldsymbol{x}$, low uncertainty is given when multiple sample models yield close estimated outputs $\hat{\boldsymbol{y}}$; high, otherwise.
3) BNN has much less training complexity compared with GP. As discussed earlier, the computational complexity of a single GP model is $O(N_o n^3 d)$. In the SADEA series, n is linearly increased with d and there are m specifications. Hence, the complexity is $O(N_o d^4 m)$. For the BNN used in SB-SADEA, which uses two hidden layers, the computational complexity is $O(N_b d(d+m)^2 s)$, where s is the sampling cost, and N_b is the number of iterations in training. More verifications are shown in Section IV.

Due to the considerably reduced training cost of BNN, an idea inspired by deeply supervised nets [19] for image recognition is proposed, which we call "fine supervision." Often, the antenna response over the operating band is considered as a whole and a maximum or minimum is obtained as the performance (e.g., $\max(|S_{11}|)$). Using a point to represent a whole curve, much information is lost.

In the proposed fine supervision, the response curve is divided by resonances, and for each part of the response curve, its maximum or minimum value is used. In this way, much more information participates in the learning with the cost of increasing the number of specifications (i.e., the number of surrogate models). This is a significant burden to GP modeling when the number of design variables is large because several times more GP models need to be trained [11]. However, for BNN, this is affordable because only the number of neurons in the output layer increases. This conclusion is verified by case study 2 (four-band 5G mm-wave antenna) in Section IV.

The parameter setting of BNN is as follows. The BNN structure has two hidden layers, and the number of neurons is d (input layer), $2d$ (the first hidden layer), $\max([d, 2m])$ (the second hidden layer), and m (the output layer), respectively. The prior standard deviation is defaulted to be 0.1. The Adam optimizer is used for training with an initial learning rate of 0.05 and a decay rate of 0.999 in every step of the model parameter optimization. An early stopping criterion is set within training to stop any training proceeding with an insignificant loss decrease. All the above are based on rules of thumbs or empirical settings and are verified by antennas with various characteristics.

C. Self-Adaptive LCB Method

For a machine learning core (i.e., BNN-based model in this case), a corresponding prescreening method (Step 6 in SB-SADEA) is often needed considering its own data characteristics. Most existing prescreening methods consider the data characteristics of the GP model. In our pilot experiments, the widely used expected improvement [26] and probability of improvement [27] prescreening are used together with the BNN-based model for multiple antenna cases, and more than 50% of the runs are stuck in local optima. This invites us to

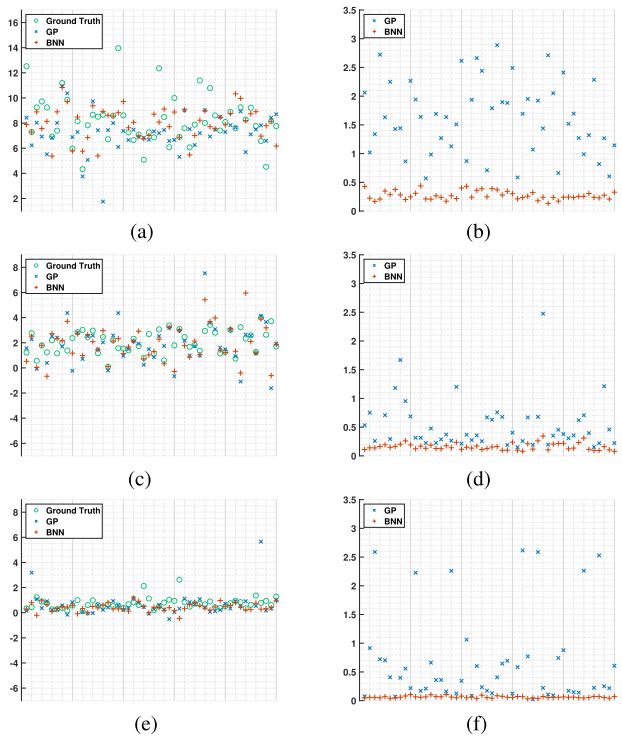

Fig. 3. GP and BNN predicted values and prediction uncertainty during early, middle, and late stages of optimization (ground truth is from EM simulations). (a) Early stage S_{11} predicted values. (b) Early stage S_{11} prediction uncertainties (standard deviation). (c) Middle stage S_{11} predicted values. (d) Middle stage S_{11} prediction uncertainties (standard deviation). (e) Late stage S_{11} predicted values. (f) Late stage S_{11} prediction uncertainties (standard deviation).

study the difference in data characteristics between BNN and GP in terms of predicted values and prediction uncertainty, so as to propose a prescreening method that can jump out of local optima and also improve the convergence speed for the BNN-based model.

Using case study 1 in Section IV (i.e., UWB monopole antenna), Fig. 3 shows the GP's and BNN's predicted values and prediction uncertainty for three sample populations of candidate designs during the early, middle, and late stages in the optimization process. $\max(|S_{11}|)$ is used. It can be observed that: 1) in terms of the predicted values, the BNN-based model and GP-based model are comparable, and both show reasonably low prediction error compared with the simulated values (i.e., ground truth) considering all three sample populations and 2) in terms of the prediction uncertainty, the BNN-based model shows much smaller values than that of GP, and the gap between them is much clearer in the later iterations. For example, when optimization is at its late stage and nearly converges, the BNN prediction uncertainty is at the level of 0.05, while GP prediction uncertainty is at the level of 0.5.

The reason why a surrogate-model-assisted antenna global optimization method falls into local optima is the lack of exploration ability. In the optimization theory, exploration refers to exploring the search region that currently lacks knowledge, while exploitation refers to finding the optimum in the search region with sufficient knowledge. Antenna design landscapes are often highly multimodal, and strong exploration

ability is required [10]. Fully considering prediction uncertainty is important for exploration, which is the reason for prescreening methods. For the popular expected improvement and potential of improvement prescreening methods, there are no hyperparameters controlling the extent of exploration, and the prediction uncertainty obtained by the BNN-based model is small. Hence, it is not a surprise that using a BNN-based model often leads to falling into a local optimum for antenna cases compared with using the GP model.

A solution is using the LCB prescreening method (7) [28], which has a hyperparameter ω to control the extent of exploration. The value of ω can be set empirically using experiments with various antenna design cases, and the recommended value is 14. Using a large value for ω can promote the exploration ability, but high exploration ability inevitably slows down the convergence (i.e., more EM simulations) due to the no-free-lunch theorem. Hence, a novel method to obtain the appropriate trade-off, called the self-adaptive LCB, is proposed. Liu et al. [45] have proposed a self-adaptive LCB method by self-adaptively learning the appropriate ω value from the optimization history. However, the method proposed here targets the prediction data characteristics of BNN and is completely different from the previous method.

Given the λ current best candidate designs, called $\boldsymbol{P_b}$, and a vector called $\boldsymbol{S}$, where $S_i (i = 1, 2, \ldots, k)$ saves the smallest distance between the current predicted best candidate design to all the candidate designs in $\boldsymbol{P_b}$ in each iteration, self-adaptive LCB (Step 6 in SB-SADEA) works as follows.

1) *Step 1:* Obtain the best candidate design $\boldsymbol{x_b}$ in the child population in Step 4 of SB-SADEA using the BNN-based model predicted values in Step 5 of SB-SADEA.
2) *Step 2:* Calculate the distance between $\boldsymbol{x_b}$ and each individual in $\boldsymbol{P_b}$ and obtain the smallest distance, S_{k+1}.
3) *Step 3:* Taking the last ten elements of $\boldsymbol{S}$ to form $\boldsymbol{S}_z$, $z = k-9, \ldots, k$, check whether S_{k+1} from Step 2 is smaller than $\bar{\boldsymbol{S}}_z - 0.5 \times \sigma_{\boldsymbol{S}_z}$, where $\sigma_{\boldsymbol{S}_z}$ is the standard deviation of $\boldsymbol{S}_z$. If yes, go to Step 4; Otherwise, output $\boldsymbol{x_b}$.
4) *Step 4:* Prescreen the child population using the LCB method (7) with the recommended ω value.
5) *Step 5:* Output the best candidate design according to LCB values.

Some clarifications are as follows.

1) The self-adaptive LCB method alternatively uses the BNN model predicted value and the LCB prescreened value for selecting the estimated best candidate design from the child population. The former is for promoting exploitation so as to improve the convergence speed, while the latter is for promoting exploration for jumping out of local optima.
2) Whether the algorithm has sufficient exploration ability or not highly depends on the diversity of $\boldsymbol{P_b}$ (Step 3 of SB-SADEA). Hence, the predicted values are used when the diversity is reasonable, while LCB values are imposed when the diversity is small.
3) The method to judge the extent of introduced diversity is to compare with the smallest distance to any individual in $\boldsymbol{P_b}$ with those in the last ten iterations. Assuming $S_i (i = 1, 2, \ldots, k)$ is Gaussian-distributed, the $0.5 \times \sigma$ value is used as the threshold to find those introducing low diversity to $\boldsymbol{P_b}$ when using them.

D. Parameter Settings

Compared with standard SADEA [8], SB-SADEA only introduces one new parameter, ω, in the self-adaptive LCB method. Using various challenging antennas from fewer than ten design variables to 45 design variables, from a few specifications to more than 20 specifications, ω is suggested to be set to 14 for successfully jumping out of local optima. Although when using a smaller value of ω the two case studies in Section IV can obtain the optimal design much faster and also with a 100% success rate, the success rate for some other cases does not reach 100%. Our experimental results show that $\omega = 14$ is safe to use, while for not so challenging antennas, a smaller value can be used to improve the speed. In BNN modeling, the network parameters are predecided by the rules of thumb and do not need the users to alter. For all the other parameters, the setting rule in other SADEA versions is still applicable to SB-SADEA, which are: $\alpha = 4 \times d$, $\lambda = 4 \times d$, $\tau = 4$, $F = 0.8$, and $CR = 0.8$. They are used in all the test cases in Section IV.

IV. Experimental Results and Verifications

SB-SADEA is tested by seven challenging antennas with various characteristics and the comparisons show the same conclusion. In this section, two typical cases from them are used to demonstrate SB-SADEA's performance in different aspects.

The first case study is a slotted monopole antenna for UWB microwave imaging applications [46]. The antenna has ten design variables and three specifications. The design optimization of this antenna is challenging due to its compact size to ensure proper physical placement and integration of its antenna structure with compact components on the same printed circuit board. For antennas with ten design variables and three specifications, the machine learning cost using most methods is often small. Hence, the purpose of this case study is to test SB-SADEA's convergence speed (i.e., the number of EM simulations needed to obtain the optimal design) when facing stringent design specifications.

The second case study is a four-band mm-wave antenna for wearable 5G and beyond applications [47]. It has 20 design variables and 12 specifications. The design optimization of a high-performance 5G mm-wave antenna is often challenging [48], and this case study has particular challenges due to its compact size, lightweight, low profile, and low maintenance with a simple off-centered microstrip feeding structure. Moreover, maintaining a multiband, high-gain operation in wearable scenarios for body-centric wireless communications at mm-wave frequencies increases the design complexity and sensitivity. Considering the number of design variables

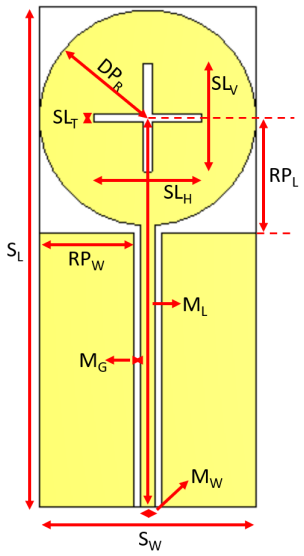

Fig. 4. Layout of the compact UWB slotted monopole antenna.

and specifications, the machine learning cost can be considered computationally expensive for the GP-based methods. Although antennas with more design variables and specifications can make the advantages of SB-SADEA even clearer, considering the time to draw statistical conclusions (i.e., using sufficient runs), this antenna is selected as a representative. Hence, the purpose of this case study is to test SB-SADEA's performance in terms of both convergence speed and machine learning cost.

For both the antennas, because no reasonably good initial designs can be provided, the search ranges provided by the antenna designers are relatively wide, although restricted by the compact size. The antennas are optimized in a workstation with an AMD Ryzen[1] Threadripper[1] PRO (2.7 GHz) and an NVIDIA[2] RTX[1] A4000 GPU. In all, 80 MATLAB parallel workers are activated for GP/BNN-based surrogate model training.

The SADEA series are stochastic algorithms and ten independent runs are carried out for case study 1 to draw statistical conclusions. Ten runs of design optimization are expected to be over a month for case study 2, and five independent runs are carried out. P-SADEA [10], [12] is selected as the reference method for case study 1 since it is one of the state-of-the-art methods for antenna design global optimization with fewer than 20 design variables with a few specifications. PSO, as one of the most popular EAs for antenna global optimization, is also used as a reference.

TR-SADEA [11] is selected as the reference method for case study 2, since to the best of our knowledge, it is the only published method for antenna global optimization with many design variables and specifications, addressing the challenge in machine learning cost. DE, as one of the most popular EAs for antenna global optimization, is also used as a reference. In terms of parameter setting, SB-SADEA follows Section III-D, P-SADEA follows [10], and TR-SADEA follows [11]. The PSO optimizer with default parameters in Computer Simulation Technology—Microwave Studio (CST-MWS) is used. The setting of the DE optimizer follows [3].

[1]Trademarked.
[2]Registered trademark.

TABLE I

SEARCH RANGES OF THE DESIGN VARIABLES AND THE OPTIMAL DESIGN BY SB-SADEA (ALL SIZES IN MM) (CASE STUDY 1)

Parameters	Lower bound	Upper bound	SB-SADEA Optimum
Circular patch radius (DP_R)	2	25	7.14
Substrate width (S_W)	$2\times DP_R$	$3\times DP_R$	14.40
Width of slot throat (SL_T)	0	$2\times DP_R$	5.77
Vertical slots' depth (SL_V)	0	$2\times DP_R$	5.36
Horizontal slots' depth (SL_H)	0	$2\times DP_R$	0.79
Microstrip length (M_L)	RP_L	50	26.21
Partial ground plane length (RP_L)	DP_R	M_L	8.21
Microstrip width (M_W)	0.50	7.50	1.20
Microstrip gap (M_G)	0	DP_R-0.5$\times$ M_W	0.34
Feed guide width (P_W)	$6\times M_W$	$10\times M_W$	10.68
Substrate width (S_L) $= M_L + 2\times DP_R + 0.2$ (mm)			
Partial ground plane width (RP_W) $= (S_W - 2\times M_G - M_W) \div 2$ (mm)			

TABLE II

DESIGN SPECIFICATIONS AND THE PERFORMANCE OF A TYPICAL OPTIMAL DESIGN OBTAINED BY SB-SADEA (CASE STUDY 1)

Item	Specification	SB-SADEA Optimum
Maximum Reflection Coefficient ($\lvert S_{11}\rvert$) (3.1 to 10.6 GHz)	$\leq$ -10 dB	-10.13 dB
Maximum Realized Gain ($G_{\max}$)	$\leq$ 3 dB	2.90 dB
Minimum Realized Gain ($G_{\min}$)	$\geq$ 1 dB	1.19 dB

A. Case Study 1: A Compact UWB Slotted Monopole Antenna

The layout of the slotted monopole antenna is shown in Fig. 4. The antenna is implemented on an FR-4 substrate with a thickness of 0.8 mm, a relative permittivity of 4.4, and a loss tangent of 0.02. It consists of a driven circular patch radiator and two uniform rectangular metal planes separated by the microstrip line. Two slots are fused at the center of the driven circular patch radiator to form a quasi-cross slot, and the geometry of the slot helps control the surface current distribution. Meanwhile, the rectangular planes act as a coplanar partial ground.

The slotted monopole antenna is modeled and discretized in CST-MWS with over 162 000 mesh cells in total. Each EM simulation costs about 45 s on average. For the optimization of the slotted monopole antenna, the design variables shown in Fig. 4 and their search ranges in Table I are considered. The optimization goal is to minimize the fitness function, F_{monopole}, to satisfy the design specifications shown in Table II, mathematically

$$\begin{aligned} F_{\text{monopole}} = {} & w_1 \times \max(|S_{11}| + 10, 0) \\ & + w_2 \times \max(G_{\max} - 3, 0) \\ & + w_3 \times \max(1 - G_{\min}, 0) \end{aligned} \tag{16}$$

where w_1, w_2, and w_3 are the penalty coefficients set to 1, 50, and 50, respectively. When all the design specifications in

TABLE III

NUMBER OF EM SIMULATIONS (AVERAGE NUMBER) USED TO SATISFY THE SPECIFICATIONS FOR DIFFERENT METHODS (CASE STUDY 1)

	SB-SADEA	GP-ALCB	FBN-LCB	BN-ALCB	GP-LCB
ML models	BNN	GP	BNN	BNN	GP
Fine-supervision	Yes	No	Yes	No	No
Prescreening	AdapLCB	AdapLCB	LCB	AdapLCB	LCB
number of EM simulations	**924**	1262	1329	1104	1991

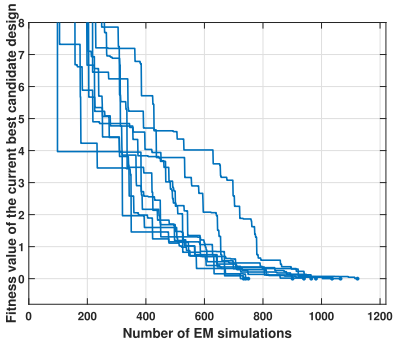

Fig. 5. Convergence trends of SB-SADEA (case study 1, ten runs).

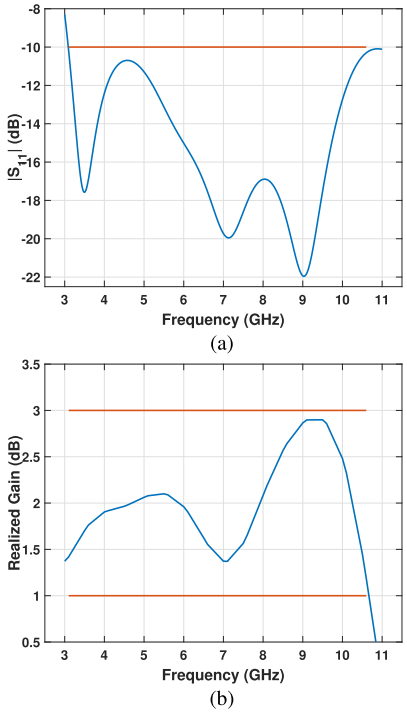

Fig. 6. Response of the optimal design obtained by SB-SADEA (case study 1). (a) Reflection coefficient of the optimal design. (b) Realized gain of the optimal design.

Table II are satisfied, F_{monopole} is equal to 0. Ten independent runs are carried out for SB-SADEA and all the other reference methods except PSO. Three runs are carried out for PSO because more runs are not affordable.

In all the ten runs, SB-SADEA satisfies the design specifications shown in Table II using an average of 924 EM simulations (12 h). Fig. 5 shows the convergence trends. Fig. 6 shows the reflection coefficient and the realized gain of a typical optimal design mentioned in Table II. The size of the antenna shrinks to about 60% compared with a state-of-the-art reference design [49].

As discussed earlier, as one of the state-of-the-art methods for antennas with stringent specifications but without many design variables and specifications, P-SADEA is considered as the reference method. P-SADEA also shows a 100% success rate but uses an average of 1574 EM simulations to satisfy all the specifications. Therefore, SB-SADEA saves 40% of the EM simulations compared with P-SADEA. Note that compared with the standard SADEA [8], P-SADEA improves the convergence speed at the cost of more GP modeling [10], [12] by its new model management framework. SB-SADEA, on the other hand, only uses the model management framework of the standard SADEA [8], and the comparison result shows the effectiveness of the BNN-based modeling and self-adaptive LCB techniques. Moreover, they are compatible with the model management framework of P-SADEA, forming an even faster method.

To verify the effectiveness of the BNN-based antenna modeling, including the fine supervision, and the self-adaptive LCB-based prescreening, more comparisons are shown in Table III. When the GP model is used, the ω value for LCB is set to 2 as other SADEA versions, instead of 14 for the BNN-based model.

The following conclusions can be drawn from Table III.

1) By comparing SB-SADEA with GP-ALCB, when both make use of the self-adaptive LCB-based prescreening, nearly 25% fewer EM simulations are saved by the BNN-based surrogate modeling compared with GP.
2) By comparing SB-SADEA with FBN-LCB, when both make use of the BNN-based surrogate modeling, nearly 30% fewer EM simulations are saved. This indicates the effectiveness of the self-adaptive LCB prescreening and its co-working with the BNN-based model. For the BNN-based model, the prediction uncertainty is smaller than that of GP (see Section III), and a larger ω has to be used in LCB prescreening to guarantee the exploration ability, which inevitably slows down the convergence speed. Hence, the self-adaptive LCB technique is essential for the BNN-based model.
3) By comparing SB-SADEA with BN-ALCB, where the only difference is the use of fine supervision, about 15% fewer EM simulations are saved, showing the effect of fine supervision.
4) GP-LCB (i.e., standard SADEA) is the slowest and SB-SADEA decreases 53% of the necessary EM simulations to obtain the optimal design, showing the combined

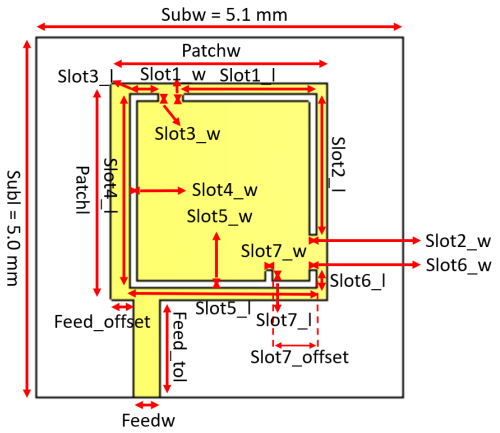

Fig. 7. Layout of the four-band 5G mm-wave antenna.

effect of the BNN-based antenna surrogate model and the self-adaptive LCB method.

In the three PSO runs, the specifications on realized gain are satisfied, but the specification on $\max(|S_{11}|)$ is not, and the average value is -5.2 dB. This can be attributed to the compactness of the structure and the stringency of the design specifications. Considering all these comparisons, this case study verifies the advantages of SB-SADEA in terms of convergence speed.

B. Case Study 2: A Four-Band Mm-Wave Antenna

This case is designed to exhibit a quad-band operation with significant band discrimination and high gain at mm-wave frequencies of 28, 38, 50, and 60 GHz. It aims to achieve a minimum realized gain of 4.5 dB and a total efficiency better than 80% for all the four operating bands. This low-profile antenna uses a patch geometry combining a square patch with an L- and an F-shaped slot on a Rogers RT/Duroid 5880 substrate of 0.254 mm thickness, relative permittivity of 2.2, and loss tangent of 0.0009. This single-layer 5.1 mm $\times$ 5 mm $\times$ 0.254 mm antenna is excited by a 50 Ω off-centered single-feed microstrip line. The slots positioned close to the edges of the patch make the current mostly concentric there and generate inductive and capacitive effects resulting in the multifrequency operation.

The four-band mm-wave antenna is modeled and discretized in CST-MWS with nearly 300 000 mesh cells in total. Each EM simulation costs around 2–2.5 min. For the optimization of the targeted antenna, the design variables shown in Fig. 7 and their search ranges in Table IV are considered. The optimization goal is to minimize the fitness function, F_{mmwave}, to satisfy the design specifications shown in Table V, mathematically

$$\begin{aligned} F_{\text{mmwave}} = &\sum_{i=1}^{4} w_1 \times \max\left(|S_{11}^i| + 10, 0\right) \\ &+ \sum_{i=1}^{4} w_2 \times \max\left(4.5 - G_{\text{min}}^i, 0\right) \\ &+ \sum_{i=1}^{4} w_3 \times \max\left(0.8 - E_{\text{total}}^i, 0\right) \end{aligned} \tag{17}$$

TABLE IV

SEARCH RANGES OF THE DESIGN VARIABLES AND A TYPICAL OPTIMAL DESIGN OBTAINED BY SB-SADEA (ALL SIZES IN MM) (CASE STUDY 2)

Variable	Lower bound	Upper bound	SB-SADEA Optimum
slot1_w	0	3	0.059
slot2_w	0	3	0.72
slot3_w	0	3	0.033
slot4_w	0	3	2.23
slot5_w	-3	0.2	-0.71
slot6_w	-3	0.2	-0.081
slot7_w	-2.2	0.9	-1.34
slot7_offset	0	2.5	2.24
feedw	0.1	0.45	0.18
feed_offset	0	3-feedw	0.017
slot1_l	0	3	0.18
slot2_l	0	3	1.98
slot3_l	0	3	1.95
slot4_l	0	3	0.60
slot5_l	0	3	1.40
slot6_l	0	3	2.15
slot7_l	0	3	0.87
feed_tol	0	3	2.02
patchl	0.5	4.3	4.26
patchw	0.5	5	4.55

TABLE V

DESIGN SPECIFICATIONS AND THE PERFORMANCE OF A TYPICAL OPTIMAL DESIGN OBTAINED BY SB-SADEA (CASE STUDY 2)

Items	Specification	SB-SADEA Optimum
Maximum in-band reflection coefficients ($\|S11\|$) (27.75 to 28.25 GHz)	$\leq$ -10 dB	-12.28 dB
Maximum in-band reflection coefficients ($\|S11\|$) (37.75 to 38.25 GHz)	$\leq$ -10 dB	-13.04 dB
Maximum in-band reflection coefficients ($\|S11\|$) (49.75 to 50.25 GHz)	$\leq$ -10 dB	-10.54 dB
Maximum in-band reflection coefficients ($\|S11\|$) (59.75 to 60.25 GHz)	$\leq$ -10 dB	-16.18 dB
Minimum in-band realized gain (G_{min}) (27.75 to 28.25 GHz)	$\geq$ 4.5 dB	5.67 dB
Minimum in-band realized gain (G_{min}) (37.75 to 38.25 GHz)	$\geq$ 4.5 dB	4.88 dB
Minimum in-band realized gain (G_{min}) (49.75 to 50.25 GHz)	$\geq$ 4.5 dB	6.75 dB
Minimum in-band realized gain (G_{min}) (59.75 to 60.25 GHz)	$\geq$ 4.5 dB	7.01 dB
Minimum in-band total efficiency (E_{tot}) (27.75 to 28.25 GHz)	$\geq$ 80%	82.4%
Minimum in-band total efficiency (E_{tot}) (37.75 to 38.25 GHz)	$\geq$ 80%	86.2%
Minimum in-band total efficiency (E_{tot}) (49.75 to 50.25 GHz)	$\geq$ 80%	84.4%
Minimum in-band total efficiency (E_{tot}) (59.75 to 60.25 GHz)	$\geq$ 80%	89.3%

where i is the index for the current frequency band out of the four frequency bands. w_1, w_2, and w_3 are the penalty coefficients set to 1, 50, and 50, respectively. When all the design specifications in Table V are satisfied, F_{mmwave} is equal to 0.

Five independent runs are carried out to test SB-SADEA. All of them satisfy the design specifications shown in Table V using an average of 1202 EM simulations. Fig. 8 shows the convergence trends. Fig. 9 shows the reflection coefficient,

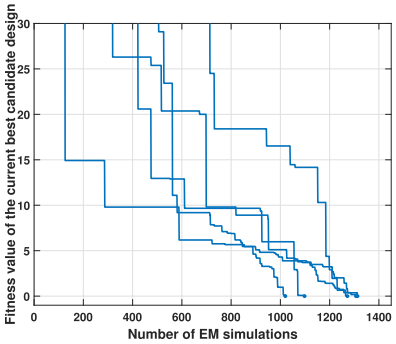

Fig. 8. Convergence trends of SB-SADEA (case study 2, five runs).

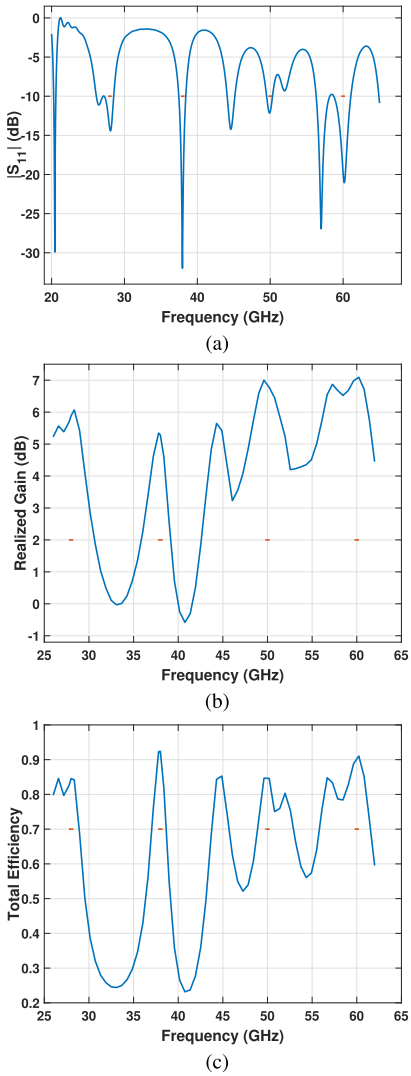

Fig. 9. Response of the optimal design obtained by SB-SADEA (case study 2). (a) Reflection coefficient of the optimal design. (b) Realized gain of the optimal design. (c) Total efficiency of the optimal design.

realized gain, and total efficiency of a typical optimal design in Table IV.

As discussed earlier, TR-SADEA [11] is selected as the reference method. In all the five runs, it also has a 100% success rate but uses an average of 2426 EM simulations. Hence, SB-SADEA decreases the number of EM simulations by more than 50% compared with TR-SADEA in this case study, verifying the advantages in convergence speed again.

TABLE VI
COMPARISON BETWEEN SB-SADEA AND TR-SADEA (CASE STUDY 2, AVERAGE VALUES)

	SB-SADEA	TR-SADEA
ML models	BNN	GP (model sharing)
Fine-supervision	Yes	Yes
Prescreening	AdapLCB	LCB
Number of surrogate models	**102,000**	426,000
Modeling time (hours)	**1.6**	12.0
Number of EM simulations	**1202**	2426
Total optimization time (hours)	**48.5**	104.3

The other aim of this case study is to compare the machine learning cost. TR-SADEA is proposed for antennas with many design variables and specifications, where GP modeling time becomes a challenge. By its GP model sharing method, TR-SADEA often reduces the GP modeling time by 90% [11]. Still, for the targeted antenna, an average of over 426 000 GP surrogate models are built in the optimization using TR-SADEA, taking about 12 h on average. This time consumption is practical but not desirable. With BNN-based surrogate modeling in SB-SADEA, only 1.6 h of surrogate model training time on average is used. Table VI demonstrates the number of EM simulations used, the number of surrogate models trained, and the total time used for the two methods. The average values of the five independent runs are used. The significant improvement in terms of the machine learning cost of SB-SADEA is also shown. The total optimization time decreased by more than a half compared with the reference method, TR-SADEA.

DE is carried out for the 5G mm-wave antenna. After two weeks of optimization, none of the reflection coefficient specifications is satisfied, and only half of the gain and total efficiency specifications are met. Longer run may improve the performance, but the optimization time is too long for practical use. Considering all these comparisons, this case study verifies the advantages of SB-SADEA in terms of both convergence speed and machine learning cost.

V. CONCLUSION

In this article, the SB-SADEA method has been proposed. Its effectiveness and efficiency are demonstrated by two real-world challenging antenna design cases. The main contributions of this article include: 1) introducing the BNN-based surrogate modeling into online antenna global optimization area to replace GP modeling and 2) introducing a new self-adaptive LCB method to co-work with the BNN-based surrogate model, which is essential for it. Hence, significant advantages in terms of both convergence speed (i.e., the number of EM simulations needed to obtain the optimal design) and machine learning cost are obtained. Due to the above innovations, SB-SADEA transforms the SADEA series from the GP-model-based to the BNN-model-based and becomes universal for antennas with various numbers of design variables and specifications, and also with significant performance improvement. Future works will include behavioral analysis of SB-SADEA and its improvement.

REFERENCES

[1] Y. Sato, F. Campelo, and H. Igarashi, "Meander line antenna design using an adaptive genetic algorithm," *IEEE Trans. Magn.*, vol. 49, no. 5, pp. 1889–1892, May 2013.

[2] Z. D. Zaharis and T. V. Yioultsis, "A novel adaptive beamforming technique applied on linear antenna arrays using adaptive mutated Boolean PSO," *Prog. Electromagn. Res.*, vol. 117, pp. 165–179, 2011.

[3] R. Storn and K. Price, "Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces," *J. Global Optim.*, vol. 11, no. 4, pp. 341–359, Dec. 1997.

[4] J. Kennedy, "Particle swarm optimization," in *Encyclopedia of Machine Learning*. Jun. 2010, pp. 760–766. [Online]. Available: https://link.springer.com/referenceworkentry/10.1007/978-0-387-30164-8_630

[5] P. Rocca, G. Oliveri, and A. Massa, "Differential evolution as applied to electromagnetics," *IEEE Antennas Propag. Mag.*, vol. 53, no. 1, pp. 38–49, Feb. 2011.

[6] Z. Zhang, H. C. Chen, and Q. S. Cheng, "Surrogate-assisted quasi-Newton enhanced global optimization of antennas based on a heuristic hypersphere sampling," *IEEE Trans. Antennas Propag.*, vol. 69, no. 5, pp. 2993–2998, May 2020.

[7] S. Koziel and S. Ogurtsov, *Antenna Design by Simulation-Driven Optimization*. Cham, Switzerland: Springer, 2014.

[8] B. Liu, H. Aliakbarian, Z. Ma, G. A. E. Vandenbosch, G. Gielen, and P. Excell, "An efficient method for antenna design optimization based on evolutionary computation and machine learning techniques," *IEEE Trans. Antennas Propag.*, vol. 62, no. 1, pp. 7–18, Jan. 2014.

[9] B. Liu, S. Koziel, and N. Ali, "SADEA-II: A generalized method for efficient global optimization of antenna design," *J. Comput. Des. Eng.*, vol. 4, no. 2, pp. 86–97, 2017.

[10] M. O. Akinsolu, B. Liu, V. Grout, P. I. Lazaridis, M. E. Mognaschi, and P. Di Barba, "A parallel surrogate model assisted evolutionary algorithm for electromagnetic design optimization," *IEEE Trans. Emerg. Topics Comput. Intell.*, vol. 3, no. 2, pp. 93–105, Apr. 2019.

[11] B. Liu et al., "An efficient method for complex antenna design based on a self adaptive surrogate model-assisted optimization technique," *IEEE Trans. Antennas Propag.*, vol. 69, no. 4, pp. 2302–2315, Apr. 2021.

[12] B. Liu, M. O. Akinsolu, N. Ali, and R. Abd-Alhameed, "Efficient global optimisation of microwave antennas based on a parallel surrogate model-assisted evolutionary algorithm," *IET Microw., Antennas Propag.*, vol. 13, no. 2, pp. 149–155, 2018.

[13] M. Alibakhshikenari et al., "Dual-polarized highly folded bowtie antenna with slotted self-grounded structure for sub-6 GHz 5G applications," *IEEE Trans. Antennas Propag.*, vol. 70, no. 4, pp. 3028–3033, Apr. 2022.

[14] Q. Wu, H. Wang, and W. Hong, "Multistage collaborative machine learning and its application to antenna modeling and optimization," *IEEE Trans. Antennas Propag.*, vol. 68, no. 5, pp. 3397–3409, May 2020.

[15] S. Koziel, A. Pietrenko-Dabrowska, and U. Ullah, "Low-cost modeling of microwave components by means of two-stage inverse/forward surrogates and domain confinement," *IEEE Trans. Microw. Theory Techn.*, vol. 69, no. 12, pp. 5189–5202, Dec. 2021.

[16] S. Koziel, A. Bekasiewicz, I. Couckuyt, and T. Dhaene, "Efficient multi-objective simulation-driven antenna design using co-Kriging," *IEEE Trans. Antennas Propag.*, vol. 62, no. 11, pp. 5900–5905, Nov. 2014.

[17] J. Zhou et al., "A trust-region parallel Bayesian optimization method for simulation-driven antenna design," *IEEE Trans. Antennas Propag.*, vol. 69, no. 7, pp. 3966–3981, Jul. 2021.

[18] C. Rasmussen, "Gaussian processes in machine learning," in *Proc. Adv. Lectures Mach. Learn.*, 2004, pp. 63–71.

[19] C.-Y. Lee, S. Xie, P. Gallagher, Z. Zhang, and Z. Tu, "Deeply-supervised nets," in *Artificial Intelligence and Statistics*. San Diego, CA, USA: PMLR, 2015, pp. 562–570.

[20] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: A nested U-Net architecture for medical image segmentation," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Cham, Switzerland: Springer, 2018, pp. 3–11.

[21] D. Erricolo et al., "Machine learning in electromagnetics: A review and some perspectives for future research," in *Proc. Int. Conf. Electromagn. Adv. Appl. (ICEAA)*, Sep. 2019, pp. 1377–1380.

[22] B. Liu, Q. Zhang, and G. G. E. Gielen, "A Gaussian process surrogate model assisted evolutionary algorith for medium scale expensive optimization problems," *IEEE Trans. Evol. Comput.*, vol. 18, no. 2, pp. 180–192, Apr. 2014.

[23] E. Zhu, Z. Wei, X. Xu, and W.-Y. Yin, "Fourier subspace-based deep learning method for inverse design of frequency selective surface," *IEEE Trans. Antennas Propag.*, vol. 70, no. 7, pp. 5130–5143, Jul. 2022.

[24] M. Emmerich, K. Giannakoglou, and B. Naujoks, "Single-and multi-objective evolutionary optimization assisted by Gaussian random field metamodels," *IEEE Trans. Evol. Comput.*, vol. 10, no. 4, pp. 421–439, Aug. 2006.

[25] Y. Jin, "Surrogate-assisted evolutionary computation: Recent advances and future challenges," *Swarm Evol. Comput.*, vol. 1, no. 2, pp. 61–70, Jun. 2011.

[26] 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, Dec. 1998.

[27] H. Ulmer, F. Streichert, and A. Zell, "Evolution strategies assisted by Gaussian processes with improved preselection criterion," in *Proc. Congr. Evol. Comput.*, 2003, pp. 692–699.

[28] J. Dennis and V. Torczon, "Managing approximation models in optimization," in *Multidisciplinary Design Optimization: State-of-the-Art*. Houston, TX, USA: Rice Univ. Press, 1997, pp. 330–347.

[29] A. Bekasiewicz, S. Koziel, and Q. S. Cheng, "Reduced-cost constrained miniaturization of wideband antennas using improved trust-region gradient search with repair step," *IEEE Antennas Wireless Propag. Lett.*, vol. 17, no. 4, pp. 559–562, Apr. 2018.

[30] A. Pietrenko-Dabrowska and S. Koziel, "Expedited gradient-based design closure of antennas using variable-resolution simulations and sparse sensitivity updates," *IEEE Trans. Antennas Propag.*, vol. 70, no. 6, pp. 4925–4930, Jun. 2022.

[31] Y. Youn et al., "PSO-aided ILA methodology for hemispherical beam coverage and scan loss mitigation," in *Proc. Int. Symp. Antennas Propag. (ISAP)*, Jan. 2021, pp. 153–154.

[32] K. Dutta, D. Guha, and C. Kumar, "Theory of controlled aperture field for advanced superstrate design of a resonance cavity antenna with improved radiations properties," *IEEE Trans. Antennas Propag.*, vol. 65, no. 3, pp. 1399–1403, Mar. 2017.

[33] M. Stein, "Large sample properties of simulations using Latin hypercube sampling," *Technometrics*, vol. 29, no. 2, pp. 143–151, May 1987.

[34] G. B. Wright, *Radial Basis Function Interpolation: Numerical and Analytical Developments*. Boulder, CO, USA: Univ. Colorado at Boulder, 2003.

[35] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, Jan. 2014.

[36] Y. Gal and Z. Ghahramani, "Dropout as a Bayesian approximation: Representing model uncertainty in deep learning," in *Proc. Int. Conf. Mach. Learn.*, 2016, pp. 1050–1059.

[37] Y. Gal and Z. Ghahramani, "A theoretically grounded application of dropout in recurrent neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 29, 2016, pp. 1019–1027.

[38] D. M. Blei, A. Kucukelbir, and J. D. McAuliffe, "Variational inference: A review for statisticians," *J. Amer. Stat. Assoc.*, vol. 112, no. 518, pp. 859–877, 2017.

[39] Y. Gal and Z. Ghahramani, "Bayesian convolutional neural networks with Bernoulli approximate variational inference," 2015, *arXiv:1506.02158*.

[40] Z.-H. Zhou, *Ensemble Methods: Foundations and Algorithms*. Boca Raton, FL, USA: CRC Press, 2019.

[41] L. Breiman, "Bagging predictors," *Mach. Learn.*, vol. 24, no. 2, pp. 123–140, 1996.

[42] S. Depeweg, J.-M. Hernandez-Lobato, F. Doshi-Velez, and S. Udluft, "Decomposition of uncertainty in Bayesian deep learning for efficient and risk-sensitive learning," in *Proc. Int. Conf. Mach. Learn.*, 2018, pp. 1184–1193.

[43] P. Westermann and R. Evins, "Using Bayesian deep learning approaches for uncertainty-aware building energy surrogate models," *Energy AI*, vol. 3, Mar. 2021, Art. no. 100039.

[44] L. V. Jospin, W. Buntine, F. Boussaid, H. Laga, and M. Bennamoun, "Hands-on Bayesian neural networks—A tutorial for deep learning users," 2020, *arXiv:2007.06823*.

[45] B. Liu, Q. Zhang, F. V. Fernández, and G. Gielen, "Self-adaptive lower confidence bound: A new general and effective prescreening method for Gaussian process surrogate model assisted evolutionary algorithms," in *Proc. IEEE Congr. Evol. Comput.*, Jun. 2012, pp. 1–6.

[46] I. M. Danjuma, M. O. Akinsolu, C. H. See, R. A. Abd-Alhameed, and B. Liu, "Design and optimization of a slotted monopole antenna for ultra-wide band body centric imaging applications," *IEEE J. Electromagn., RF Microw. Med. Biol.*, vol. 4, no. 2, pp. 140–147, Jun. 2020.

[47] M. Ur-Rehman, M. Adekanye, and H. T. Chattha, "Tri-band millimetre-wave antenna for body-centric networks," *Nano Commun. Netw.*, vol. 18, pp. 72–81, Dec. 2018.
[48] Z. U. Khan, T. H. Loh, A. Belenguer, and A. Alomainy, "Empty substrate-integrated waveguide-fed patch antenna array for 5G millimeter-wave communication systems," *IEEE Antennas Wireless Propag. Lett.*, vol. 19, no. 5, pp. 776–780, May 2020.
[49] B. Yeboah-Akowuah, P. Kosmas, and Y. Chen, "A Q-Slot monopole for UWB body-centric wireless communications," *IEEE Trans. Antennas Propag.*, vol. 65, no. 10, pp. 5069–5075, Oct. 2017.

Yushi Liu (Student Member, IEEE) received the B.Eng. degree in electronic information engineering from the University of Electronic Science and Technology of China, Chengdu, China, and the B.Eng. degree (Hons.) in electronics and electrical engineering from the University of Glasgow, Glasgow, U.K., in 2017, and the M.Sc. degree in electrical engineering from Tufts University, Medford, MA, USA, in 2019. He is currently pursuing the Ph.D. degree with the AI-Driven Design Laboratory, University of Glasgow, under the supervision of Dr. Bo Liu.

From 2019 to 2020, he was appointed as a Research and Development Engineer at the Institute of Deep-Sea Science and Engineering, Chinese Academy of Sciences, Beijing, China. His research interests include machine learning, engineering optimization, and AI-driven engineering design.

Bo Liu (Senior Member, IEEE) received the B.Eng. degree from Tsinghua University, Beijing, China, in 2008, and the Ph.D. degree from the University of Leuven (Katholieke Universiteit (KU) Leuven), Leuven, Belgium, in 2012.

From 2012 to 2013, he was a Humboldt Research Fellow and was with the Technical University of Dortmund, Dortmund, Germany. In 2013, he was appointed as a Lecturer (Assistant Professor) at Wrexham Glyndwr University, Wrexham, U.K., where he was promoted to a Reader in 2016. He is currently a Reader with the University of Glasgow, Glasgow, U.K. He is also a Senior Honorary Fellow with the University of Birmingham, Birmingham, U.K. He has authored or coauthored one book and more than 80 articles in renowned international journals, edited books, and conference proceedings. His research interests lie in AI-driven design methodologies of analog/RF integrated circuits, microwave devices, MEMS, evolutionary computation, and machine learning. In terms of AI-driven antenna design, he is the main inventor and designer of the SADEA algorithm series.

Masood Ur-Rehman (Senior Member, IEEE) received the B.Sc. degree in electronics and telecommunication engineering from the University of Engineering and Technology at Lahore, Lahore, Pakistan, in 2004, and the M.Sc. and Ph.D. degrees in electronic engineering from the Queen Mary University of London, London, U.K., in 2006 and 2010, respectively.

He has worked as a Post-Doctoral Research Assistant with the Queen Mary University of London, till 2012, before joining the Centre for Wireless Research at the University of Bedfordshire, Luton, U.K., as a Lecturer. He has served briefly at the University of Essex, Colchester, U.K., and then moved to the James Watt School of Engineering, University of Glasgow, Glasgow, U.K., in the capacity of an Assistant Professor, in 2019. He currently works as an Associate Professor with the University of Glasgow. His research interests include compact antenna design, radiowave propagation, satellite navigation system antennas in cluttered environment, electromagnetic wave interaction with human body, wireless sensor networks in healthcare and environmental monitoring, mmWave and nano communications for body-centric networks, and D2D/H2H communications. He has worked on several projects supported by industrial partners and research councils. He has contributed to a patent and authored/coauthored five books, eight book chapters, and more than 145 technical articles (with an H-index of 24) in leading journals and peer reviewed conferences.

Dr. Ur-Rehman is a fellow of the Higher Education Academy (U.K.), a member of the IET and BioEM, and part of the technical program committees and organizing committees of several international conferences, workshops, and special sessions. He is a Committee Member of IEEE APS/SC WG P145, IEEE APS Best Paper Award Committee, and Pearson's focus group on formative assessment. He has been acting as an Editor of *PeerJ Computer Science*; an Associate Editor of IEEE SENSORS JOURNAL, IEEE ACCESS, *IET Electronics Letters*, and *Microwave and Optical Technology Letters*; a Topic Editor of *Sensors* (MDPI); an Editorial Advisor to Cambridge Scholars Publishing; and a Lead Guest Editor of numerous special issues of renowned journals. He is the Vice-Chair of IEEE AP/MTT Scotland Joint Chapter and acted as the Workshop Chair of Workshop on Sustainable and Intelligent Green Internet of Things for 6G and Beyond in IEEE PIMRC2022 and IEEE VTC-S 2022, and the TPC Chair of UCET2020 and BodyNets2021 conferences. He also serves as a Reviewer for EPSRC and international funding bodies, book publishers, IEEE conferences, and leading journals.

Muhammad Ali Imran (Senior Member, IEEE) is currently the Dean of the University of Glasgow, Glasgow, U.K., and the University of Electronic Science and Technology of China (UESTC), Chengdu, China. He is also a Professor of wireless communication systems. He is also the Director of Glasgow–UESTC Centre for Educational Development and Innovation. He is also an Affiliate Professor with The University of Oklahoma, Norman, OK, USA; a Visiting Professor with the 5G Innovation Centre, University of Surrey, Guildford, U.K.; and an Adjunct Research Professor with the Artificial Intelligence Research Center, College of Engineering and Information Technology, Ajman University, Ajman, United Arab Emirates. He is also the Head of the Communications, Sensing, and Imaging Research Group, University of Glasgow, Glasgow, U.K. He has more than 20 years of combined academic and industry experience with several leading roles in multimillion-pound-funded projects. He has filed 15 patents. He has authored/coauthored more than 400 journal articles and conference publications. He has edited seven books. He has authored more than 30 book chapters. He has successfully supervised more than 40 postgraduate students at the doctoral level. His research interests include self-organized networks, wireless networked control systems, the Internet of Things, and wireless sensor systems. He has been a consultant to international projects and local companies in the area of self-organized networks.

Dr. Imran is a fellow of IET and a Senior Fellow of Higher Education Academy (HEA).

Mobayode O. Akinsolu (Senior Member, IEEE) received the M.Sc. degree (Hons.) in electrical and electronic engineering from the University of Bradford, Bradford, U.K., in 2014, after his undergraduate studies and compulsory national service in Nigeria, and the Ph.D. degree from the University of Chester, Chester, U.K., and Wrexham Glyndwr University, Wrexham, U.K., in 2019. His Ph.D. research was focused on electromagnetic design automation using surrogate-model-assisted evolutionary algorithms.

He has worked as a Research Fellow (Industrial Attaché) and a Visiting Researcher with the National Space Research and Development Agency, Nigeria, and the RFID Research Centre, African University of Science and Technology, Nigeria, respectively, until 2016. From 2016 to 2019, he was a Ph.D. Scholarship Awardee in recognition of a joint project between Wrexham Glyndwr University and the University of Birmingham, Birmingham, U.K. He received a commendation from the University of Chester for the publications record (more than 20 publications) relating to his Ph.D. work. His contributions in applying artificial intelligence techniques to automate and expedite engineering design processes have led to the co-authorship of several articles in reputable peer-reviewed journals and presentations at international conferences. He is currently a Senior Lecturer (an Associate Professor) in electronic and communication engineering with Wrexham Glyndwr University, where he leads and teaches modules such as Industry 4.0, machine learning, data analysis and visualization, control system, and others.

Dr. Akinsolu is a Chartered Engineer (CEng) with the Engineering Council, a fellow of the Higher Education Academy (FHEA), a member of the IET (MIET), and a registered Electrical Engineer (R.Engr.) with the COREN.

Peter Excell (Life Senior Member, IEEE) received the B.Sc. degree in engineering science from the University of Reading, Reading, U.K., in 1970, the Ph.D. degree from the University of Bradford, Bradford, U.K., in 1980, and the D.Sc. degree from Wrexham Glyndwr University, Wrexham, U.K., and the University of Chester, Chester, U.K., in 2020.

Until 2007, he was the Associate Dean of research with the School of Informatics, University of Bradford, where he is currently a Visiting Professor. He is also a Professor with the Emeritus of Communications, Wrexham Glyndwr University, where he was the Deputy Vice Chancellor until 2015. He is also a Chartered Engineer and a Chartered IT Professional. He has published more than 500 articles. His academic research interests are concentrated in the areas of wireless technologies, electromagnetics and antennas, plus broader interests in future communications applications, and technological evolution.

Dr. Excell is a fellow of the Institution of Engineering and Technology, the British Computer Society, and the Higher Education Academy; and a member of the Association for Computing Machinery, the Bioelectromagnetics Society, and the Applied Computational Electromagnetics Society.

Qiang Hua received the B.Sc. degree in communication engineering from the University of Liverpool, Liverpool, U.K., in 2016, the M.Sc. degree in digital signal processing from The University of Manchester, Manchester, U.K., in 2017, and the Ph.D. degree from the University of Liverpool in 2022.

He is currently a Post-Doctoral Research Assistant with Liverpool John Moores University, Liverpool. His current research interests include base station antenna design for 5G in wireless communications, new material antennas, smart sensors for health monitoring, wearable devices, and machine learning method for antenna optimization.