An Efficient Bayesian Optimization Approach for Automated Optimization of Analog Circuits

Wenlong Lyu, Pan Xue, Fan Yang, *Member, IEEE*, Changhao Yan, *Member, IEEE*, Zhiliang Hong, Xuan Zeng, *Member, IEEE*, and Dian Zhou, *Senior Member, IEEE*

Abstract—The computation-intensive circuit simulation makes the analog circuit sizing challenging for large-scale/complicated analog/RF circuits. A Bayesian optimization approach has been proposed recently for the optimization problems involving the evaluations of black-box functions with high computational cost in either objective functions or constraints. In this paper, we propose a weighted expected improvement-based Bayesian optimization approach for automated analog circuit sizing. Gaussian processes (GP) are used as the online surrogate models for circuit performances. Expected improvement is selected as the acquisition function to balance the exploration and exploitation during the optimization procedure. The expected improvement is weighted by the probability of satisfying the constraints. In this paper, we propose a complete Bayesian optimization framework for the optimization of analog circuits with constraints for the first time. The existing GP model-based optimization methods for analog circuits take the GP models as either offline models or as assistance for the evolutionary algorithms. We also extend the Bayesian optimization algorithm to handle multi-objective optimization problems. Compared with the state-of-the-art approaches listed in this paper, the proposed Bayesian optimization method achieves better optimization results with significantly less number of simulations.

Index Terms—Analog circuit sizing, Bayesian optimization, Gaussian process, weighted expected improvement, multi-objective optimization.

I. INTRODUCTION

MANUAL analog circuit design becomes increasingly difficult due to the continuous scaling of the IC technology and the ever-increasing demands for high-performance, lower-power circuits [1]. Automated analog circuit sizing attracted many research interests. It is mainly because the analog circuit sizing can be well formulated as constrained nonlinear optimization problems, which could be solved by the well-developed optimization algorithms.

The classical analog circuit sizing approaches can be classified into two categories, i.e., the model-based and simulation-based approaches. For the model based methods, the analytical expressions are used to model the circuit performances with respect to the design parameters. Designers encode their knowledge into mathematical expressions. If the manually derived expressions are not available, regression models could be used instead. One of the best-known model based approaches is the geometric programming based method [2]. In this approach, the circuit performances are modeled as posynomial and monomial functions. The analog circuit sizing problems can then be formulated as a convex optimization problem. Once the posynomial models are available, the global optimum could be found very efficiently, even for large-scale problems such as the Phase Locked Loops (PLL) [3] and the analog-digital converters (ADC) [4]. In [5], the posynomial models can be obtained automatically from simulation data.

In [6], general polynomials are used to model the circuit performances, and the general polynomial optimization problem is transformed to convex programming by Semidefinite-Programming (SDP) relaxations. Other modeling techniques including support vector machine (SVM), artificial neural network (ANN) or Gaussian process (GP, also known as kriging in geostatistics) [7], [8] are also used. A comprehensive review of the modeling techniques can be found in [1], and [9] gives an empirical comparison of different models.

One of the advantages of model-based approaches is that the optimization can be efficiently solved with the performance models. The computational cost of evaluating the models is much lower than that of circuit simulations. The global optimum can be guaranteed with posynomial models or even general polynomial models. Another advantage is that the models can be reused. Once a model is built, it can be saved and reused for further optimization. The disadvantage of model based method is the deviation between the predicted performances of the circuit models and the real circuit performances. The deviation would make the optimal points obtained from the models diverge from the real optima.

For simulation-based methods, the optimization is driven directly by the circuit simulations. The objective functions and constraints are viewed as black-box functions. They are obtained directly from circuit simulations. Since the circuit simulations are invoked on-the-fly, the number of required simulations would even be much lower than those of model-based approaches. A variety of global optimization algorithms,

including simulated annealing (SA) [10], particle swarm intelligence (PSO) algorithm [11] and evolutionary algorithm [12], have been proposed to better explore the solution space. Stochastic characteristics of these algorithms help them to avoid getting stuck in the local optimum. However, they also suffer from relatively low convergence rate. Recently, gradient based local search with multiple starting points (MSP) [13]–[15] have been reported to be very efficient for analog circuit sizing.

However, the computation-intensive circuit simulation makes the analog circuit sizing challenging for large-scale/complicated analog/RF circuits. For example, one transient simulation could take several hours for complicated circuits like a PLL. The passive components such as inductors and transformers are widely used in RF and microwave circuits. However, these components are usually characterized by Electro-Magnetic (EM) simulations, which could take even several hours for one configuration. Furthermore, if the process variations are taken into account, PVT corner enumerations or Monte-Carlo yield estimations are required. The computational cost would thus be greatly increased. To address this problem, a hybrid methodology which combines both the model-based approach and simulation-based approach has been proposed recently [16]–[20]. Instead of using pre-built offline models, online evolving models are built. The circuit simulations are invoked during the optimization procedure to update the online models incrementally. The initial model is constructed firstly using the data gathered from random sampling or Design of Experiments (DoE). The model is used to guide the selection of the next point towards higher performance. Once a new data point is obtained through simulation, the model is updated with the new data point. For example, GASPAD was proposed in [20] for microwave circuit synthesis.

Multi-objective optimization has also been introduced for automated analog circuit design [21], as analog circuits usually have several conflicting performances. It would be helpful to provide the designers a set of designs with the best trade-off between performances. Unlike the aforementioned single-objective optimization algorithms, a multi-objective optimization algorithm tries to find a set of *non-dominated* designs that approximate the Pareto front of design objectives. For multi-objective analog circuits optimization, meta-heuristic based algorithms are commonly used. The Non-dominated Sorting Genetic Algorithm (NSGA-II) [22], the multi-objective evolutionary algorithm based on decomposition (MOEA/D) [23], and the particle swarm intelligence (PSO) algorithm [11] with weighted sum aggregation serve as the optimization kernel of [24]–[28].

In this paper, we propose a Weighted Expected Improvement based Bayesian Optimization (WEIBO) approach for the automated analog circuit sizing. Based on the prior belief of the objective functions, the probabilistic surrogate models of the objective functions are refined incrementally via Bayesian posterior updating when new data are observed. The probabilistic models of the objective functions are used to guide the optimization procedure. Gaussian processes (GP) are used as the online surrogate models for circuit performances in our

proposed approach. It is important to balance the *exploration* and *exploitation* during the optimization. The *exploration* means that the next point tends to explore the unknown regions where the uncertainty in the surrogate model is large. The *exploitation* means that the next point tends to be the optimum which is predicted by the surrogate model with high probability. Acquisition functions are introduced in Bayesian optimization to balance the exploration and exploitation. Expected Improvement (EI) is selected as the acquisition function in our proposed approach. The expected improvement is weighted by the probability of satisfying the constraints. The weighted EI (wEI) can well balance the exploration and exploitation, and thus the wEI based Bayesian optimization framework can significantly reduce the number of simulations while achieving better optimization results.

The Bayesian optimization algorithm is further extended to handle multi-objective optimization problems in this paper. We propose to transform the multi-objective problems into single-objective problems that can be solved by our proposed WEIBO algorithm by Tchebysheff scalarization [21] with smooth relaxation.

The main contributions of the paper are summarized as follows. Firstly, we propose a complete Bayesian optimization framework for the optimization of analog circuits with constraints for the first time. The existing GP model based optimization methods for analog circuits take GP models as either offline models or as assistance for evolutionary algorithms. Secondly, we extend the Bayesian optimization algorithm to deal with the multi-objective problems through Tchebysheff scalarization and smooth relaxation.

The proposed WEIBO method was tested with several test cases including two passive components working at 60GHz, a low power operational amplifier, a power amplifier, a voltage controlled oscillator (VCO), and a charge pump. Compared with the state-of-the-art approaches listed in the paper, the proposed WEIBO method can achieve better optimization results with significantly less number of simulations.

The rest of the paper is organized as follows. In Section II, we will present the problem formulations of single- and multi-objective analog circuit optimizations. In Section III, we will present the proposed Bayesian optimization approach. In Section IV, the multi-objective Bayesian optimization algorithm is proposed. In Section V, we review the GP-related methods to clarify our contributions. In Section VI, the efficiency of the proposed approach will be demonstrated with several test cases. We conclude the paper in Section VII.

II. PROBLEM FORMULATION

A. Single-Objective Optimization

If we consider optimizing a single objective of the analog circuits and impose constraints to the other performances, the analog circuit optimization problem can be formulated as a constrained optimization problem as follows.

$$\text{minimize } f(\pmb{x})$$
$$\text{s.t. } c_i(\pmb{x}) < 0$$
$$\forall i \in 1 \ldots N_c, \tag{1}$$

where $x \in R^d$ represents d design variables, $f(x)$ denotes the Figure of Merit (FOM) of the analog circuit. Generally, for each design variable x_i, it is limited in a range $[p_i^-, p_i^+]$ due to the fabrication process or design requirements. We denote the design space as X. $f(x)$ could be a combination of circuit performances which should be optimized. For simplification, we only consider the minimization problem in the paper. A maximization problem can be easily converted to a minimization problem by taking the negative of the objective function. The optimization problem (1) is constrained by the specifications of other performances. Without loss of generality, we formulate these constraints as $c_i(x) < 0, \forall i \in 1 \dots N_c$, where a $c_i(x) < 0$ corresponds to the i-th constraints.

B. Multi-Objective Optimization

If we aim to find the best trade-off between multiple performances, the formulation in (1) should be extended to multi-objective optimization as follows.

$$\text{minimize } f_1(x), \quad \dots \quad , f_m(x)$$
$$\text{s.t. } c_i(x) < 0$$
$$\forall i \in 1 \dots N_c, \tag{2}$$

Unlike the single-objective problem, there is usually no *best* design for a multi-objective problem, as objectives can be conflicting, and it is unlikely to optimize all of them simultaneously. The goal of a multi-objective optimization algorithm is to find a set of solutions to approximate the *Pareto front*. A solution x_1 is said to *dominate* a solution x_2 if $\forall i \in \{1 \dots m\}$, $f_i(x_1) \leq f_i(x_2)$ and $\exists j \in \{1 \dots m\}$, $f_j(x_1) < f_j(x_2)$. A point x_* is *weakly Pareto-optimal* if $\forall x \in X$, x_* is not dominated by x. A point x_* is *Pareto-optimal* if $\forall x \in X$, x_* is not dominated by x and $\exists i \in \{1 \dots m\}$, $f_i(x_*) < f_i(x)$. The set of Pareto optimal outcomes is called the *Pareto front*.

III. WEIGHTED EXPECTED IMPROVEMENT BASED BAYESIAN OPTIMIZATION APPROACH FOR ANALOG CIRCUIT SIZING

In this section, we present a Weighted Expected Improvement based Bayesian Optimization (WEIBO) approach for analog circuit sizing.

A. Overview of Bayesian Optimization

The Bayesian optimization approach consists of two essential ingredients. The first ingredient is the probabilistic surrogate model of the objective function, which is refined incrementally when new data is observed. With the probabilistic surrogate model, the model uncertainty can also be evaluated. The second ingredient is an acquisition function, which is used to explore the state space based on the surrogate model optimally. The acquisition function aims to balance the exploration and exploitation during the optimization. The model uncertainty plays a key role here to balance the exploration and exploitation. A typical acquisition function is the Expected Improvement (EI) based on the current probabilistic surrogate model. A comprehensive overview of Bayesian optimization can be found in [29].

We summarize the overall Bayesian optimization approach in Algorithm 1. The initial models are built firstly using the data gathered from random sampling or design of experiment (DoE). In each iteration, the probabilistic surrogate models are updated using the new data point. The next data point is then generated by maximizing the acquisition function.

Algorithm 1 Bayesian Optimization

1: Initial Sampling
2: Construct probabilistic surrogate model
3: **for** t = 1, 2, … **do**
4:     Find x_t that optimizes the acquisition function
5:     Sample $y_t = f(x_t)$
6:     Update probabilistic surrogate model
7: **end for**
8: **return** best $f(x)$ recorded during iterations

B. Gaussian Process Model

The *de facto* surrogate model used in Bayesian optimization is the Gaussian process (GP) model. We refer readers to [30] for more details of the GP model. The motivation of using GP model is that its model prediction and model uncertainty have closed forms. Let $f : \chi \to R$ be a black-box function (e.g., circuit performances in circuit sizing problem) to be inferred. For a Gaussian process, it assumes that for any finite samples $\{x_1, \dots, x_n\} \in \chi^n$, the vector $(f(x_1), \dots, f(x_n))^T$ follows joint multivariate Gaussian distribution

$$\begin{bmatrix} f(x_1) \\ \vdots \\ f(x_n) \end{bmatrix} \sim N(\mu, K), \tag{3}$$

where μ is an $n \times 1$ mean vector and K is an $n \times n$ covariance matrix.

The Gaussian process can be fully characterized by its mean function $m(x)$ and its covariance function $k(x_1, x_2)$. In (3), the mean vector μ is determined by $m(x)$ such that $\mu_i = m(x_i)$, $\forall i \in 1 \dots n$, and the covariance matrix is determined by the covariance function such that $K_{ij} = k(x_i, x_j)$.

The mean and covariance functions can be viewed as our prior belief in the unknown scalar-valued function $y = f(x)$. The mean function can be any function, e.g., zero, constant, linear, or functions obtained by other regression techniques. For simplicity, we use the constant mean function $m(x) = \mu$ in our proposed approach, where μ is a hyper-parameter to be determined. Any function that generates a symmetric and semi-definite covariance matrix can be a valid covariance function. The squared exponential covariance function k_{SE} is used in our approach.

$$k_{SE}(x_i, x_j) = \sigma_f^2 \exp\left(-\frac{1}{2}(x_i - x_j)^T \Lambda^{-1}(x_i - x_j)\right), \tag{4}$$

where σ_f is the variance, $\Lambda = \text{diag}(l_1^2, l_2^2, \dots, l_d^2)$ is a $d \times d$ diagonal matrix, l_i is the length scale in the i-th dimension. Λ and σ_f are hyper-parameters to be determined.

Let training set $D = \{X, y\}$, where $X = \{x_1, \dots, x_N\}$ and $y = \{f(x_1), \dots, f(x_N)\}$. Given a new point x_*, the function

value $y_* = f(\mathbf{x}_*)$ and $\mathbf{y}$ follow the joint Gaussian distribution

$$\begin{bmatrix} y_* \\ \mathbf{y} \end{bmatrix} \sim N\left(\begin{bmatrix} m(\mathbf{x}_*) \\ \mathbf{m} \end{bmatrix}, \begin{bmatrix} [l]k(\mathbf{x}_*, \mathbf{x}_*) & k(\mathbf{x}_*, X) \\ k(X, \mathbf{x}_*) & K_N \end{bmatrix} \right), \quad (5)$$

where $\mathbf{m} = (m(\mathbf{x}_1), \ldots, m(\mathbf{x}_N))^T$ is the $N \times 1$ mean vector, and

$$k(\mathbf{x}_*, X) = (k(\mathbf{x}_*, \mathbf{x}_1), \ldots, k(\mathbf{x}_*, \mathbf{x}_N))$$
$$k(X, \mathbf{x}_*) = k(\mathbf{x}_*, X)^T$$
$$K_N = \begin{bmatrix} k(\mathbf{x}_1, \mathbf{x}_1) & \cdots & k(\mathbf{x}_1, \mathbf{x}_N) \\ k(\mathbf{x}_2, \mathbf{x}_1) & \cdots & k(\mathbf{x}_2, \mathbf{x}_N) \\ \vdots & \ddots & \vdots \\ k(\mathbf{x}_N, \mathbf{x}_1) & \cdots & k(\mathbf{x}_N, \mathbf{x}_N) \end{bmatrix}.$$

Note that y_* conditioned by $\mathbf{y}$ also follows the normal distribution [31]. The distribution can be expressed as

$$y_* \sim N(\mu_{y_*|\mathbf{y}}, \sigma^2_{y_*|\mathbf{y}})$$

where

$$\begin{cases} \mu_{y_*|\mathbf{y}} = m(\mathbf{x}_*) + k(\mathbf{x}_*, X)K_N^{-1}(\mathbf{y} - \mathbf{m}) \\ \sigma^2_{y_*|\mathbf{y}} = k(\mathbf{x}_*, \mathbf{x}_*) - k(\mathbf{x}_*, X)K_N^{-1}k(X, \mathbf{x}_*). \end{cases} \quad (6)$$

Equation (6) gives the probabilistic prediction of y_*. The mean $\mu_{y_*|\mathbf{y}}$ can be viewed as the prediction, while the variance $\sigma^2_{y_*|\mathbf{y}}$ can be viewed as the confidence of the prediction.

A Gaussian noise $\epsilon \sim N(0, \sigma_n^2)$ can be added to $f(\mathbf{x})$ to model the observation noise, where σ_n^2 is the variance of the Gaussian noise. For the analog circuit sizing problem, $f(\mathbf{x})$ is obtained from simulation, and it can be viewed as noise-free. Nevertheless, the noise term should still be added to avoid overfitting and improve the numerical stability. If we take noise into consideration. The original covariance function becomes

$$k_n(\mathbf{x}_i, \mathbf{x}_j) = k(\mathbf{x}_i, \mathbf{x}_j) + \sigma_n^2 \delta_{ij}, \quad (7)$$

where $\delta_{ij} = 0$ if $i \neq j$, $\delta_{ij} = 1$ if $i = j$.

With the noise, the prediction of GP model becomes

$$\begin{cases} \mu_{y_*|\mathbf{y}} = m(\mathbf{x}_*) + k(\mathbf{x}_*, X)(K_N + \sigma_n^2 I)^{-1}(\mathbf{y} - \mathbf{m}) \\ \sigma^2_{y_*|\mathbf{y}} = k(\mathbf{x}_*, \mathbf{x}_*) - k(\mathbf{x}_*, X)(K_N + \sigma_n^2 I)^{-1}k(X, \mathbf{x}_*). \end{cases} \quad (8)$$

In our algorithm, we use the constant mean function, and the covariance function (4) with noise term. For a d-dimensional black-box function, there are $d+3$ hyper-parameters in the GP model, including d length scales, μ for the mean function, variance σ_f for the covariance function, and σ_n for the white noise. One commonly used method to determine the hyper-parameters is the maximum likelihood estimation (MLE). For simplicity, we introduce a new vector $\boldsymbol{\theta}$ to denote the vector of hyper-parameters, i.e., $[\mu \ \sigma_f \ \sigma_n \ l_1 \ \cdots \ l_d]^T$. Given training set $D = \{X, \mathbf{y}\}$ with N data points, the likelihood of the training set can be expressed as

$$p(\mathbf{y}|X, \boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi^N |K_\theta|}}$$
$$\times \exp\left(-\frac{1}{2}(\mathbf{y} - \boldsymbol{\mu}_\theta)^T K_\theta^{-1}(\mathbf{y} - \boldsymbol{\mu}_\theta) \right), \quad (9)$$

where $\boldsymbol{\mu}_\theta$ is the N-dimensional mean vector, and $\boldsymbol{\mu}_\theta = [\mu, \cdots, \mu]^T$. K_θ is $N \times N$ covariance matrix, and $K_\theta = K_N + \sigma_n^2 I$. We can derive the hyper-parameters by maximizing the likelihood in (9).

C. Acquisition Function

An acquisition is used to balance the exploration and exploitation. We omit the constraints of (1), and consider the unconstrained optimization problem as follows firstly.

$$\text{minimize} f(\mathbf{x}). \quad (10)$$

The objective function $y = f(\mathbf{x})$ is approximated by a GP model. Equation (6) gives the probabilistic prediction of y_* based on the training data. The mean $\mu_{y_*|\mathbf{y}}$ can be viewed as the prediction, while the variance $\sigma^2_{y_*|\mathbf{y}}$ can be viewed as the confidence of the prediction. Based on this model, it is possible to optimally balance the exploration and exploitation by introducing the acquisition function.

The Expected Improvement (EI) [32] is used as the acquisition function in our proposed approach. Denote τ as the minimal value of $f(\mathbf{x})$ we found currently. The improvement of y can be expressed as

$$I(y, \tau) = \begin{cases} \tau - y & y < \tau \\ 0 & \text{otherwise}, \end{cases} \quad (11)$$

which means that the improvement exists only when y is less than τ. From the GP model, we know that for a given training set D, the prediction of y at point $\mathbf{x}$ follows the distribution

$$p(y|D, \boldsymbol{\theta}) = N(\mu(\mathbf{x}), \sigma^2(\mathbf{x})),$$

where $\mu(\mathbf{x})$ and $\sigma(\mathbf{x})$ are the mean and variance of the distribution, $\boldsymbol{\theta}$ is the vector of hyper-parameters. With this predictive distribution, we can derive the expectation of improvement (EI)

$$\begin{aligned} \text{EI}(\mathbf{x}) &= \mathbb{E}[I(y, \tau)] \\ &= \int_{-\infty}^{\infty} I(y, \tau)p(y|D, \boldsymbol{\theta})dy \\ &= \left(\tau - \mu(\mathbf{x}) \right)\Phi\left(\frac{\tau - \mu(\mathbf{x})}{\sigma(\mathbf{x})} \right) + \sigma(\mathbf{x})\phi\left(\frac{\tau - \mu(\mathbf{x})}{\sigma(\mathbf{x})} \right), \end{aligned} \quad (12)$$

where $\Phi(.)$ is the CDF of standard normal distribution, and $\phi(.)$ is the PDF function of standard normal distribution.

Besides the improvement-based acquisition functions, there are some other types of acquisition functions. For example, the methods in [20], [33], and [34] use lower confidence bound (LCB) as acquisition function. For a predictive distribution $y \sim N(\mu(\mathbf{x}), \sigma(\mathbf{x})^2)$, LCB is defined as

$$\text{LCB}(\mathbf{x}) = \mu(\mathbf{x}) + \kappa\sigma(\mathbf{x}), \quad (13)$$

where κ is a value used to explicitly control the trade-off between exploration and exploitation.

In Figure 1a and Figure 1b, we show an illustrative example of the GP model and the EI acquisition function using the real data gathered from an operational amplifier. We swept the bias current I_b and recorded the corresponding Phase Margins (PM). The 180° - PM curve with respect to the bias current was thus plotted. Several points were randomly sampled on the curve to train a GP model. The real curve, the training points, the mean and 3σ confidence bound given by the GP model are plotted in Figure 1a, and the curve of EI values are plotted in Figure 1b. Also, the lower 3σ confidence bound can be viewed as the LCB function with $\kappa = 3$.
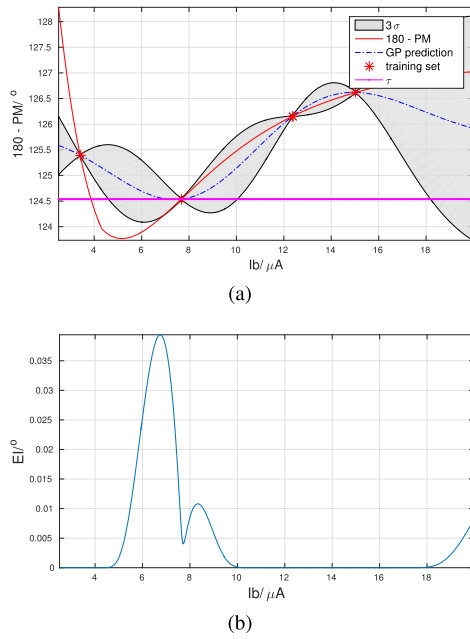
Fig. 1. Illustration of GP model and EI acquisition function for an amplifier's phase margin. (a) Illustration of the GP model. (b) Illustration of the EI acquisition function.

D. Weighted Expected Improvement (wEI)

Most of the existing acquisition functions do not handle nonlinear constraints. As far as we know, Expected Improvement (EI) and predictive entropy search [35], [36] are the acquisition functions that can deal with constraints. According to [36], the advantage of an entropy-based acquisition function is that it can handle the decoupled constraints, i.e., the constraints and the objective functions are evaluated separately, so this approach might be useful if objectives and constraints are obtained from different measures from different testbenches. However, the entropy-based acquisition function does not have a closed form, we thus use EI based constrained acquisition function in our proposed approach.

We consider the optimization problem (1) with constraints. The weighted Expected Improvement (wEI) [37], [38] is proposed to deal with the constraints.

For each constraint $u_i = c_i(x)$, a single GP model is also constructed similarly to (6) to approximate the constraint $c_i(x)$. The GP model gives a distribution of u_i as prediction

$$u_i \sim N(\mu_i, \sigma_i^2),$$

where μ_i and σ_i are the corresponding mean and variance, respectively. For the constraint $c_i(x) < 0$, the probability of the constraint being satisfied can be expressed as

$$PF_i(x) = \Pr(c_i(x) < 0) = \Phi(\frac{-\mu_i}{\sigma_i}), \qquad (14)$$

where $\Phi(.)$ is the CDF of standard normal distribution.

To deal with the constraints, the improvement function is changed. The improvement is non-zero if and only if the

objective function $f(x) < \tau$ and the point x is feasible

$$I_c(y, c, \tau) = \begin{cases} \tau - y & y < \tau \text{ and } c < 0 \\ 0 & \text{otherwise.} \end{cases} \qquad (15)$$

Assume that the objective function $f(x)$ and all the constraints $c_i(x)$ are mutually independent. The new expected improvement can thus be expressed as

$$wEI(x) = \mathbb{E}[I_c(x)] = EI(x) \prod_{i=1}^{N_c} \Pr(c_i(x) < 0), \qquad (16)$$

where N_c is the number of constraints. The EI in (12) is now weighted by the probability of feasibility of all the constraints. The weighted Expected Improvement (wEI) in (16) is a fully probabilistic acquisition function as it considers the uncertainties of both objective functions and constraints. Note that τ is the best result we found. However, it is possible that there are no feasible points found at the beginning state of optimization, so that the wEI in (16) is undefined. In such cases, we try to find the feasible points firstly. The probability of feasibility (PF) is thus used as acquisition function

$$PF(x) = \prod_{i=1}^{N_c} PF_i(x) = \prod_{i=1}^{N_c} \Pr(c_i(x) < 0). \qquad (17)$$

The expression of wEI is based on the assumption that the constraints are independent so that they can be independently modeled and the probabilities could be multiplied. Although the objective functions and constraints are mapped from different specifications of circuits and they may be correlated, we found the wEI function works well in practice.

E. Multiple Starting Point Strategy for wEI Maximization

As the GP model and the wEI acquisition function have closed-forms, its analytical gradient can be derived directly. Once the analytical gradient of acquisition function is obtained, the cost of evaluating the gradient is negligible.

A Multiple Starting Point (MSP) strategy is employed to maximize the wEI function. It has been reported in [13]–[15] that MSP strategy is a very effective strategy for global optimization. The starting point selection strategy is borrowed from [39]. Firstly, 50,000 points are randomly sampled. The best 200 points are selected as the starting points. Multiple gradient-based local searches are then performed starting from these points. The optimal solution is finally selected from the optimization results of the multiple local searches. We use the BFGS method [40] for local search. BFGS is an iterative quasi-newton method for solving the unconstrained nonlinear optimization problems.

F. Summary of the Proposed Approach

We summarize the proposed approach in Figure 2. As shown in the figure, starting from the initial training set, we built the GP models for objective and constraint functions firstly. The hyper-parameters of the GP models are determined by maximum likelihood estimation. The acquisition function (EI for unconstrained problem, and wEI for constrained problem) is then constructed based on the GP models.
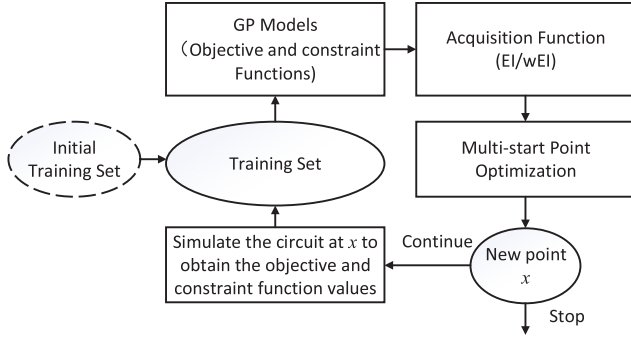
Fig. 2.   Illustration of the proposed approach.

The acquisition function is then maximized to find the next data point x. By simulating the circuits, we get the objective and constraint function values at x. The new data point is added to the training set. The GP models are updated and the iterations continue until the maximum number of simulations is reached or other stop criteria are met.

IV. MULTI-OBJECTIVE EXTENSION OF BAYESIAN OPTIMIZATION

In this section, we will present our proposed multi-objective Bayesian optimization algorithm for analog circuits. We transform the multi-objective optimization problem (2) to a weighted min-max formulation (or weighted Tchebysheff formulation) firstly [21].

$$\text{minimize} \quad \max_{i \in \{1...m\}} \omega_i (f_i(x) - f_i^*)$$
$$\text{s.t. } c_j(x) < 0,$$
$$\forall j \in \{1 \ldots N_c\}, \quad (18)$$

where the $(f_1^*, \ldots, f_i^*, \cdots f_m^*)$ in (18) is called the *utopia point*. f_i^* denotes the minimum of the i-th objective. The utopia point is generally unattainable. It could be set to certain unreachable values according to design experiences. For example, for the utopia point of a power amplifier, we can set the output efficiency to 100%. The vector $\boldsymbol{\omega} = (\omega_1, \cdots, \omega_i, \cdots, \omega_m)$ denote the weights, and we typically set $\sum_{i=1}^{m} \omega_i = 1$ and $\boldsymbol{\omega} > 0$. It can be proven that the optimal solution of (18) corresponds to a Pareto-optimal point of the original multi-objective optimization problem (2). More importantly, the complete Pareto optimal set can be obtained by varying the weights. By changing the weights, we can get a series of Pareto-optimal points and thus the approximation of the Pareto front of (2).

After the transformation, the multi-objective optimization problem is transformed to several single-objective optimization problems, and the standard Bayesian optimization can be applied. However, in the GP model, the objective functions are supposed to be smooth functions since we use the squared-exponential covariance function k_{SE} in (4) to model the correlations of objective/constrained functions. However, the objective function in (18) has a min-max form, which is no longer smooth. Although GP can model non-smooth functions, it requires more training points to guarantee the accuracy.

To address this problem, we propose to exploit a smooth relaxation to transform (18) to the following relaxed smooth

---

Algorithm 2 Constrained Multi-Objective Bayesian Optimization Algorithm

1: Initial Sampling based on design of experiment (DoE)
2: **for** t = 1, 2, ..., N_p **do**
3:     Randomly generate a weight vector $\boldsymbol{\omega}$
4:     **for** every (x, y) in data set **do**
5:         Randomly generate λ
6:         Transform (x, y) to (x_t, y_t)
7:     **end for**
8:     Train the GP models using the newly transformed data set, get the GP hyper-parameters
9:     Find x_t that maximizes the wEI using the GP models of objectives and constraints
10:     Simulate the circuits to obtain the values of the objectives and constraints at x
11:     Update the data set using the newly obtained data
12: **end for**
13: **return** the Pareto front and the Pareto set extracted from the data set (x, y)

---

formulation by introducing a new variable λ.

$$\text{minimize } \lambda$$
$$\text{s.t. } \omega_i (f_i(x) - f_i^*) < \lambda, \quad \forall i \in 1 \ldots m,$$
$$c_j(x) < 0, \quad \forall j \in 1 \ldots N_c. \quad (19)$$

For the problem in (19), besides the original design variables x, we introduce a new "pseudo" design variable λ. The design vector can thus be transformed to $x_t = (x, \lambda)$. The GP models will be built for the objective function λ and the constraints of (19), i.e., for each element of the following vector y_t

$$y_t = \begin{bmatrix} \lambda \\ \omega_1(f_1(x) - f_1^*) - \lambda \\ \vdots \\ \omega_m (f_m(x) - f_m^*) - \lambda \\ c_1(x) \\ \vdots \\ c_{N_c}(x) \end{bmatrix}. \quad (20)$$

y_t is a transformation of the vector y composed of the original objective and constraints

$$y = [f_1(x), \cdots, f_m(x), c_1(x), \cdots, c_{N_c}(x)]^T. \quad (21)$$

We propose an efficient heuristic approach for the multi-objective optimization as shown in Algorithm 2. Starting with the initial samples, at each iteration, a random weight vector $\boldsymbol{\omega}$ is generated. For this weight vector $\boldsymbol{\omega}$, we can construct a wEI acquisition function for the optimization problem (19) based on the GP models of the vector y_t. By maximizing the wEI, we can find a new data point, and this point is then added to the training set. Note that we would not find a new Pareto-optimal point with just one wEI maximization for the randomly generated weight vector $\boldsymbol{\omega}$. This new data point is possibly located in the vicinity of the true Pareto front. With the new data points, the accuracy of the GP models of y_t near

TABLE I
SUMMARY OF TEST CASES

Circuit	Process	# Transistors	# Design Variables	# PVT corners	Simulator	# Objective
60GHz Inductor	TSMC 65nm	/	3	1	ADS momentum	1
60GHz Transformer	TSMC 65nm	/	4	1	ADS momentum	2
Power Amplifier	TSMC 65nm	8192	5	1	Spectre	1, 2, 3
Operational Amplifier	TSMC 0.35μm	40	24	1	HSPICE	1
VCO	SMIC 40nm	131	20	10	HSPICE	1
Charge Pump	SMIC 40nm	41	36	18	HSPICE	1

the Pareto front would be improved as the iteration continues. The new data points found by maximizing the wEIs would be closer to the true Pareto-optimal front gradually. As a result, we can extract an approximation of the Pareto front from the data set generated by our algorithm.

V. RELATED WORKS

GP model has been introduced for analog circuit optimization in previous works. In [7] and [8], GP is combined with the Meta-heuristic optimization algorithms for variation-aware circuit optimizations. However, GP is only used as an offline model in those works. Also, only the predicted values of GP models are utilized in these works while the uncertainty information provided by the GP models are not well exploited.

In [16]–[20] and [33], the GP models are built online. Especially, the GASPAD method is proposed in [20]. Our proposed approach differs with these approaches in terms of the employment of the models and the dealing of the constraints.

The acquisition functions are introduced in GASPAD method [20] and those methods proposed in [16]–[19] and [33]. In the Bayesian optimization framework, we try to find the maximum of acquisition functions to optimally explore the solution space. However, in GASPAD [20] as well as [16]–[18], [33], they follow the framework of evolutionary algorithms. They just perform one mutation operation and one crossover operation on the population to generate the children, and the child with maximal acquisition value is then selected as the next data point to explore. In [19], they use DoE sampling to generate the new data points, and the data point with maximal acquisition value is selected as the next data point. Obviously, it is impossible to maximize the acquisition functions with just one iteration of these operations. One should note once the GP models are built, the acquisition functions are cheap to evaluate. Thus, we can employ the well-developed optimization techniques to maximize the acquisition functions. The optimal solutions that optimally balance the exploitation and exploration could thus be found. This is just what we proposed in our approach. We follow the framework of Bayesian optimization. Once the GP models are built, the acquisition function is constructed. We employ the MSP algorithm to maximize the acquisition function. The data point that maximizes the acquisition function is selected as the next data point to explore.

The constraints are handled in different ways in our proposed approach and the existing GP-based methods. In GASPAD [20] and [18], [19], [33], the tournament-based selection rule [41] is used to deal with the constraints, which is commonly used in evolutionary algorithms. Although GASPAD used the LCB acquisition function, the κ in (13) is set to zero for models of constraints. It thus cannot make use of the uncertainty information of constraints. In [16] and [17], simple static penalty function with manually specified weights is used to handle constraints. In our proposed approach, we use the weighted EI to handle the constraints. wEI takes into account the uncertainties of both the objective and the constraints. Furthermore, the existence of closed-forms of the wEI enables fast gradient-based optimization.

Also, the methods in [16]–[20] and [33] can only deal with single-objective optimization problems. We extend our method to handle the multi-objective optimization problems.

VI. EXPERIMENTAL RESULTS

We implemented the proposed Weighted Expected Improvement based Bayesian Optimization (WEIBO) algorithm in C++ language. The BFGS algorithm from the nlopt library [42] is used for local search. All the experiments were conducted on a Linux workstation with two Intel Xeon CPUs and 128G memory.

Several circuits are used to demonstrate the efficiency of the proposed approach. We list the information of these test circuits in Table I, including the technology nodes, the numbers of transistors, the numbers of the design variables, the PVT corners to be considered, the corresponding simulators and the numbers of objectives. In the circuits listed in Table I, the 60GHz transformer is used to test the multi-objective optimization algorithms, and the power amplifier is used to test both single- and multi-objective algorithms.

For single-objective circuit optimization, we compared our proposed WEIBO algorithm with the following algorithms.

1) The GASPAD method proposed in [20];
2) Sequential quadratic programming(SQP) local search with multiple starting points [13]–[15], the gradient is approximated via finite difference (MSP for short);
3) The differential evolution algorithm (DE for short) [12];
4) The particle swarm intelligence algorithm (PSO for short) [11];
5) The simulated annealing algorithm (SA for short) [10].

The following algorithms were compared with our multi-objective Bayesian optimization algorithm (MO-WEIBO) as they are representative multi-objective algorithms and have been widely used for multi-objective analog circuit synthesis [24], [25], [27], [43]–[45].

1) The non-dominated sorting genetic algorithm II (NSGA-II) [22];
2) The multi-objective evolutionary algorithm based on decomposition (MOEA/D) [23].

For constrained optimization problems, we compared our proposed MO-WEIBO algorithm only with NSGA-II, because the current code provided by the authors of MOEA/D cannot deal with the constraints. For unconstrained problems, we compared our proposed MO-WEIBO algorithm with both NSGA-II and MOEA/D algorithms.

Unlike the single-objective optimization, it is hard to directly compare the two approximations of the Pareto front. A variety of performance metrics have been proposed. In this paper, the inverted generational distance (IGD) metric and the C-metric [46] are used to compare different algorithms. To achieve a low IGD value, the approximation of Pareto front must be close to the reference Pareto front and not miss any details of the reference Pareto front. The C-metric gives an intuitive comparison between a pair of Pareto fronts.

The IGD metric is defined as

$$\text{IGD}(\mathbf{P}^*, \mathbf{P}) = \frac{\sum_{v \in \mathbf{P}^*} d(v, \mathbf{P})}{|\mathbf{P}^*|}, \tag{22}$$

where $\mathbf{P}^*$ is a set of uniformly distributed Pareto-optimal points in the true PF, and $\mathbf{P}$ is an approximation of the Pareto front (i.e., the output of an algorithm). $d(v, \mathbf{P})$ is the minimum Euclidean distance between v and all the points in $\mathbf{P}$. $|\mathbf{P}^*|$ is the cardinality of $\mathbf{P}^*$.

In (22), a true Pareto front $\mathbf{P}^*$ is needed to calculate the IGD value, For a real-world problem, however, the true Pareto front is unknown. In our experiments, we combine all the sampled points to extract a reference Pareto front.

The C-metric compares the coverage between two approximated Pareto fronts. Given two approximated Pareto fronts $\mathbf{P_A}$ and $\mathbf{P_B}$, the C-metric $C(\mathbf{P_A}, \mathbf{P_B})$ is defined as the proportion of the points in $\mathbf{P_B}$ that is dominated by at least one point in $\mathbf{P_A}$.

$$C(\mathbf{P_A}, \mathbf{P_B}) = \frac{|\{\mathbf{x} \in \mathbf{P_B} | \exists \mathbf{y} \in \mathbf{P_A} : \ \mathbf{y} \text{ dominates } \mathbf{x}\}|}{|\mathbf{P_B}|}. \tag{23}$$

If $C(\mathbf{P_A}, \mathbf{P_B}) = 1$, it means that all solutions in $\mathbf{P_B}$ are dominated by at least one solution in $\mathbf{P_A}$. Also, it should be noted that $C(\mathbf{P_A}, \mathbf{P_B})$ does not necessarily equal to $1 - C(\mathbf{P_B}, \mathbf{P_A})$. As a result, both $C(\mathbf{P_A}, \mathbf{P_B})$ and $C(\mathbf{P_B}, \mathbf{P_A})$ should be calculated.

A. 60GHz Inductor

The first example is an inductor working at 60GHz shown in Figure 3. As the working frequency is very high, the Electro-Magnetic (EM) simulator ADS momentum is used for simulation.

There are three design variables in this design, i.e.,

$$\begin{cases} R \in [15, 50]\,\mu\text{m}, \\ W \in [3, 10]\,\mu\text{m}, \\ S \in [3, 8]\,\mu\text{m}, \end{cases} \tag{24}$$

where R is the inner radius, W is the metal width, and S is the metal spacing. The number of turns is fixed to 1.5.
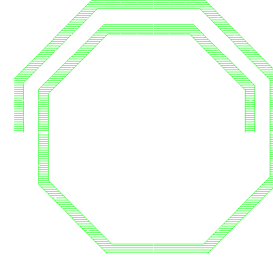

Fig. 3. Layout of 60GHz inductor.

TABLE II
THE SINGLE-OBJECTIVE OPTIMIZATION RESULTS
OF THE 60 GHz INDUCTOR

Algo	WEIBO	GASPAD	MSP	DE	PSO	SA
L@60GHz/nH	0.450	0.451	0.453	0.451	0.454	0.453
SRF > 100GHz	Yes	Yes	Yes	Yes	Yes	Yes
Q@60GHz(mean)	**15.17**	15.15	14.75	15.04	14.73	14.53
Q@60GHz(median)	**15.16**	15.12	14.81	15.01	14.77	14.53
Q@60GHz(best)	**15.31**	15.39	15.19	15.14	14.95	14.66
Q@60GHz(worst)	**15.09**	15.01	14.35	14.98	14.41	14.41
Avg. # Sim	**98**	321	437	611	384	156
# Success	**12/12**	12/12	8/12	4/12	3/12	2/12

The objective is to maximize quality factor at 60GHz, with the constraints of inductance L and self-resonance frequency (SRF) as listed below.

$$\text{maximize } Q$$
$$\text{s.t. } L \in [0.45, 0.5]\,\text{nH},$$
$$SRF > 100\,\text{GHz}. \tag{25}$$

We ran all the optimization algorithms for 12 times to average out the random fluctuations. In our experiments, we first ran our proposed WEIBO algorithm with a budget of 150 simulations, of which 24 points are the randomly sampled initial points. For the other algorithms, the simulation budget is set to 750. In Table II, we show the average values of the constraints and statistics of the FOM.

In Table II, '# Success' means the number of runs where feasible designs were found, the 'Avg. # Sim' records the average iterations used by one algorithm to converge to its final solution. It can be seen that our proposed WEIBO algorithm and GASPAD gave feasible solutions in all their 12 runs, while the other algorithms cannot guarantee to find feasible solutions within 750 simulations. Even they did find feasible solutions in several runs, they used much more number of simulations than our proposed WEIBO algorithm and GASPAD. Furthermore, the quality factor reached by them are also worse than WEIBO and GASPAD.

Although GASPAD also found high-quality designs, it still needed far more simulations than our proposed WEIBO algorithm. The number of simulations consumed by our proposed WEIBO algorithm is about 30% of that consumed by GASPAD.

B. 60GHz Transformer

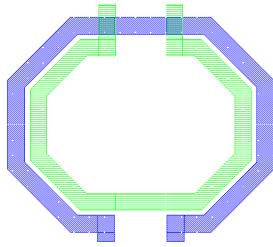A 60GHz transformer shown in Figure 4 was used to test the proposed multi-objective Bayesian optimization algorithm

Fig. 4. 60GHz Transformer.

TABLE III
THE STATISTICS OF IGD VALUES AND THE STATISTICS OF BEST
OBJECTIVES OF THE TRANSFORMER TEST CASE

	MO-WEIBO	NSGA-II
best IGD	**0.54384**	2.2594
worst IGD	**4.0503**	24.544
median IGD	**1.92725**	8.9051
mean IGD	**1.77367**	12.0485
best PTE(%)	**83.85**	83.35
best area(μm^2)	**953.9**	1051.9

MO-WEIBO. This transformer is implemented in TSMC 65nm technology. ADS momentum is used as the EM simulator.

There are four design variables, they are the radii of the metal-8 and metal-9 layers, and the widths of metal-8 and metal-9 layers. The ranges of the design variables are listed in (26).

$$\begin{cases} R_{m8} \in [10, 75]\, \mu\text{m}, \\ R_{m9} \in [10, 75]\, \mu\text{m}, \\ W_{m8} \in [5, 10]\, \mu\text{m}, \\ W_{m9} \in [5, 10]\, \mu\text{m}. \end{cases} \qquad (26)$$

For this transformer, the goal of the optimization is to maximize the power transfer efficiency (PTE) while minimizing the area. Five constraints are imposed on the following performance: the coupling efficiency k, the quality factor of the primary inductor Q_1 and quality factor of the secondary inductor Q_2, the real and imaginary part of the input impedance Z_{in}. The specifications are listed as follows.

$$\begin{aligned} \text{minimize } & (-PTE, \ area) \\ \text{s.t. } & k > 0.8, \\ & Q_1 > 10, \\ & Q_2 > 10, \\ & \text{Re}(Z_{in}) \in [10, 20]\, \Omega, \\ & \text{Im}(Z_{in}) \in [10, 25]\, \Omega. \end{aligned} \qquad (27)$$

The proposed MO-WEIBO algorithm was compared with NSGA-II. The maximum number of simulation for MO-WEIBO is 180, including 40 initial points. The maximum number of simulation for NSGA-II is 300. We run both algorithms for ten times. The statistics of the IGD values are shown in Table III.

From Table III, we can find that our proposed approach can achieve much lower IGDs than those of NSGA-II, which means our proposed approach can approximate the ideal Pareto front better.
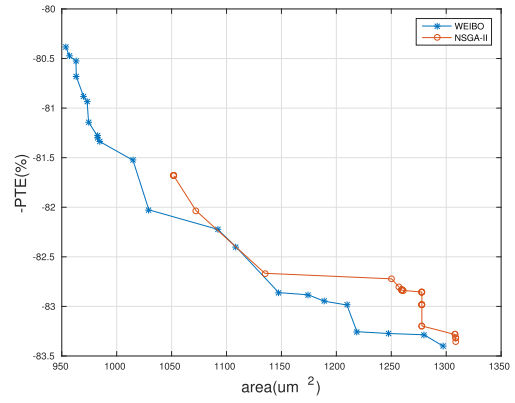

Fig. 5. The best Pareto fronts of the transformer test case from MO-WEIBO and NSGA-II.

We summarize the cover relationship of the two algorithms using a cover matrix as follows.

$$\begin{array}{cc} & \begin{array}{cc} \text{MO-WEIBO} & \text{NSGA-II} \end{array} \\ \begin{array}{c} \text{MO-WEIBO} \\ \text{NSGA-II} \end{array} & \left(\begin{array}{cc} 0 & 0.9 \\ 0.044 & 0 \end{array} \right) \end{array}$$

where the (i, j) element of the matrix denotes the proportion that the points generated by the j-th algorithm dominated by the i-th algorithm. On average, 90% of NSGA-II generated solutions are dominated by MO-WEIBO, while only less than 5% of solutions from MO-WEIBO are dominated by NSGA-II.

To illustrate the superiority of MO-WEIBO, we plot the best Pareto fronts of MO-WEIBO and NSGA-II in Figure 5. It can be seen from the figure that most of the solutions generated by our approach dominate those generated by NSGA-II.

C. Power Amplifier

The third example is a power amplifier. The circuit works at 2.4GHz and is designed using a TSMC 65nm process. It is an array-based design with 2^{11} duplicate cells. Each cell consists of 4 transistors. There are in total 8192 transistors in this design. We constructed three experiments using this example, including one single-objective experiment and two multi-objective experiments.

There are 5 design parameters, as shown in (28).

$$\begin{cases} C_p \in [0.01, 20]\, \text{pF} \\ C_s \in [0.01, 20]\, \text{pF} \\ V_{dd} \in [1, 1.5]\, \text{V} \\ V_b \in [1, 2]\, \text{V} \\ W \in [0.12, 5]\, \mu\text{m} \end{cases} \qquad (28)$$

For the single-objective optimization, we aim to maximize the output efficiency (Eff) while the constraints of the output power (Pout) and the total harmonic distortion (thd) being satisfied. The specifications are shown in (29).

$$\begin{aligned} \text{maximize } & \textit{Eff} \\ \text{s.t. } & Pout > 23\, \text{dBm} \\ & thd < 12.65\, \text{dB}. \end{aligned} \qquad (29)$$

TABLE IV

THE SINGLE-OBJECTIVE OPTIMIZATION RESULTS OF
THE ARRAY-BASED POWER AMPLIFIER

Algo	WEIBO	GASPAD	MSP	DE	PSO	SA
thd/dB	8.87	13.05	11.68	12.39	11.98	13.29
Pout/dBm	23.11	24.23	24.53	24.25	24.48	23.97
Eff(mean)/%	**60.29**	31.63	29.83	31.54	28.28	21.22
Eff(median)/%	**62.23**	22.58	20.78	33.35	20.09	17.32
Eff(best)/%	**64.02**	62.02	60.83	53.69	54.35	45.31
Eff(worst)/%	**48.86**	22.07	11.21	14.07	14.29	13.44
Avg. # Sim	**82**	257	228	234	243	170
# Success	**12/12**	12/12	12/12	12/12	12/12	6/12

TABLE V

THE STATISTICS OF IGD VALUES AND BEST OBJECTIVES OF
THE THREE-OBJECTIVE POWER AMPLIFIER CASE

	MO-WEIBO	NSGA-II	MOEA/D
best IGD	**1.8535**	2.9656	2.6532
worst IGD	**6.8316**	13.139	13.26
median IGD	**2.76505**	7.59895	3.5756
mean IGD	**2.99335**	7.48404	5.92897
best Eff(%)	**65.37**	62.35	63.08
best thd(dB)	**5.602**	6.047	6.045
best Pout(dBm)	**26.89**	26.58	25.71

TABLE VI

THE STATISTICS OF IGD VALUES AND BEST OBJECTIVES OF
THE TWO-OBJECTIVE POWER AMPLIFIER CASE

	MO-WEIBO	NSGA-II
best IGD	**0.031**	5.421
worst IGD	**3.089**	27.970
median IGD	**0.529**	19.076
mean IGD	**1.058**	18.472
best Eff(%)	**64.0**	59.12
best thd(dB)	**5.532**	7.52
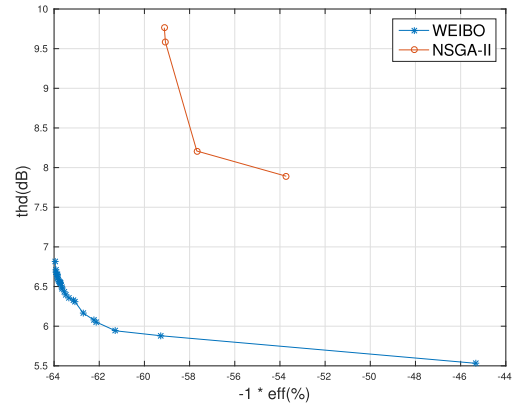
Fig. 6. The best Pareto fronts from MO-WEIBO and NSGA-II

For WEIBO, we limited the maximum allowed number of simulations to 150, and 40 out of the 150 simulations are initial samples. For the other algorithms, the simulation budget was set to 300. We also ran all the optimization algorithms for 12 times to average out the random fluctuations. The optimization results are shown in Table IV. We can see from Table IV that our proposed WEIBO approach can achieve significantly better results than the other algorithms on average. GASPAD did not work well on this circuit, actually, 9 out of the 12 runs fell into the same local optimum with $Eff \simeq 22\%$.

We also used this power amplifier to test our multi-objective Bayesian optimization algorithm. Two experiments were conducted. The first one is a three-objective optimization problem without constraints, the objectives are expressed as follows.

$$\text{minimize } (-Eff, -Pout, thd) \qquad (30)$$

We limited the number of simulations of NSGA-II and MOEA/D to 300. The number of simulations of our proposed MO-WEIBO was limited to 200. Each algorithm was run ten times, and the statistics of IGD values and the best objectives found are shown in Table V.

From Table V, we can see that our proposed approach can achieve much lower IGDs than those of NSGA-II and MOEA/D.

The cover matrix was also calculated, as shown below.

$$
\begin{array}{c|ccc}
 & \text{MO-WEIBO} & \text{MOEA/D} & \text{NSGA-II} \\
\hline
\text{MO-WEIBO} & 0 & 0.815 & 0.786 \\
\text{MOEA/D} & 0.001 & 0 & 0.383 \\
\text{NSGA-II} & 0.005 & 0.111 & 0
\end{array}
$$

It can be seen that on average, 81.5% of the final solutions of MOEA/D and 78.6% of the final solutions of NSGA-II are dominated by at least one solution generated by MO-WEIBO.

Only less than one percent of solutions of MO-WEIBO are dominated by MOEA/D or NSGA-II. As for the comparison between MOEA/D and NSGA-II, we can see that 38.8% NSGA-II solutions are dominated by MOEA/D while 11.1% MOEA/D solutions are dominated by NSGA-II.

To test the ability to handle constraints of MO-WEIBO, the *Pout* in (30) is used as constraints, while *Eff* and *thd* are taken as the two objectives. The specifications are listed as follows.

$$\text{minimize } (-Eff, \ thd)$$
$$\text{s.t. } Pout > 23 \, \text{dBm}. \qquad (31)$$

MO-WEIBO was compared with NSGA-II. We set the maximum number of simulations to 200 for MO-WEIBO, and 300 for NSGA-II. The statistics of the IGD values and best objectives are summarized in Table VI. From Table VI, we can see that our proposed approach can achieve much lower IGDs than those of NSGA-II.

The calculated cover matrix is listed as below.

$$
\begin{array}{c|cc}
 & \text{MO-WEIBO} & \text{NSGA-II} \\
\hline
\text{MO-WEIBO} & 0 & 1 \\
\text{NSGA-II} & 0 & 0
\end{array}.
$$

As can be seen, all solutions generated by NSGA-II are dominated by the solutions generated by MO-WEIBO.

We plot the Pareto fronts generated by MO-WEIBO and NSGA-II with the lowest IGD values in Figure 6. It can be seen that the Pareto front generated by NSGA-II is far from optimal. They are all dominated by the Pareto front generated by MO-WEIBO.
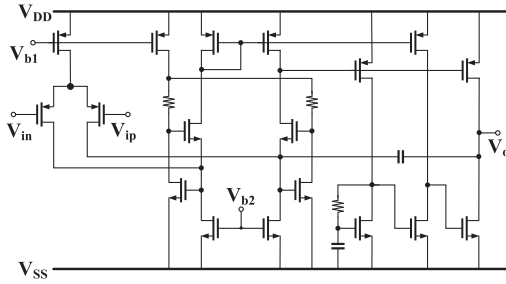
Fig. 7.   Schematic of the low power three-stage amplifier.

TABLE VII

THE SINGLE-OBJECTIVE OPTIMIZATION RESULTS OF THE
LOW POWER THREE-STAGE AMPLIFIER

Algo	WEIBO	GASPAD	MSP	DE	PSO	SA
GAIN/dB	100.57	101.6	100.81	102.73	102.39	102.49
UGF/MHz	0.922	0.924	0.982	0.956	0.963	1.046
PM/$^\circ$	52.52	52.60	53.22	54.62	54.25	56.70
GM/dB	19.76	20.05	22.30	20.62	21.32	20.92
SR+(V/μs)	0.20	0.24	0.23	0.21	0.23	0.25
SR-(V/μs)	0.46	0.55	0.51	0.54	0.51	0.49
Iq(mean)/μA	**27.87**	31.55	49.60	41.26	44.22	59.78
Iq(median)/μA	**27.65**	31.47	48.29	40.70	43.16	53.56
Iq(best)/μA	**26.70**	28.36	32.06	37.32	37.18	46.34
Iq(worst)/μA	**29.30**	34.47	75.32	46.09	59.56	83.48
Avg. # Sim	**798**	2744	2163	2400	2417	620
# Success	**12/12**	12/12	9/12	12/12	12/12	5/12

D. Low-Power Three-Stage Amplifier

The three-stage amplifier shown in Figure 7 was pro-posed in [47]. This amplifier is capable of driving large load (up to 15nF). There are 24 design variables, including the lengths and widths of transistors, the capacitance and resistance and the bias current. We use the simulation results of the manual design in [47] as the specifications, which are listed in (32). In (32), Iq means the static current, GAIN denotes the DC gain, UGF denotes the unit gain frequency, PM denotes the phase margin, GM denotes the gain margin, SRR and SRF means the rising and falling slew rate, respectively.

$$\text{minimize } Iq$$
$$\text{s.t. } GAIN > 100\,\text{dB}$$
$$UGF > 0.92\,\text{MHz}$$
$$PM > 52.5\,^\circ$$
$$GM > 19.5\,\text{dB}$$
$$SRR > 0.18\,\text{V}/\mu\text{s}$$
$$SRF > 0.2\,\text{V}/\mu\text{s}. \qquad (32)$$

In this experiment, we ran all the optimization algorithms for 12 times to average out the random fluctuations. We first ran our proposed WEIBO algorithm and limited the maximum allowed number of simulations to 1000, and the number of the initial samples is 200. For the other algorithms, we limited the maximum number of simulations to 3000.

The optimization results are shown in Table VII. As can be seen from Table VII, all algorithms can find feasible solutions. However, WEIBO is the only algorithm that reached $I_q < 30\mu A$ on average. Although the optimal I_q found by WEIBO and GASPAD are close on average, the number of simulations consumed by WEIBO is only about 29% of that consumed by GASPAD on average.
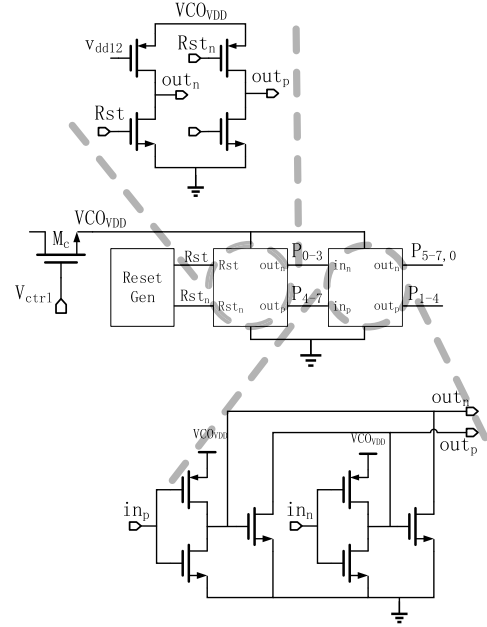

Fig. 8.   Schematic of the voltage controlled oscillator.

E. Voltage Controlled Oscillator

A Voltage Controlled Oscillator (VCO) circuit is shown in Figure 8. It consists of 131 transistors, and it is implemented in a SMIC 40nm process. There are 20 design variables in this circuit. The simulation of VCOs is very time-consuming. Even worse, the process variations and different control voltages are considered in the design specifications, which makes the simulation more time-consuming. For this circuit, a total number of 10 PVT corners were taken into account.

For this circuit, we aim to minimize the variation of the output frequency, with constraints of the linearity and the slope of the frequency-voltage curve. For each corner, we draw the frequency-voltage curve, measure the linear correlation R between the output frequency and the input control voltage. The slope k of frequency-voltage is obtained via least squares linear regression. We also measure the lowest linear frequency f_l and the highest linear frequency f_h. The specifications are listed as (33).

$$\text{minimize } FOM$$
$$\text{s.t. } k_{min} > 300\,\text{MHz/V},$$
$$R_{min} > 0.95. \qquad (33)$$

where

$$\begin{cases} k_{min} = \min_{\forall PVT}(k), \\ R_{min} = \min_{\forall PVT}(R), \\ FOM = 10 \times \left(\dfrac{\sigma(f_l)}{\mu(f_l)} + \dfrac{\sigma(f_h)}{\mu(f_h)} \right). \end{cases} \qquad (34)$$

In (34), k_{min} is the minimum of measured slopes k in all corners; R_{min} is the minimum of the measured correlations. $\mu(\cdot)$ and $\sigma(\cdot)$ in the Figure of Merit (FOM) expression represent the standard deviation and mean function, respectively. The FOM consists of the standard deviation and the mean of the f_l and f_h, which are used to model the variation of the output frequency.
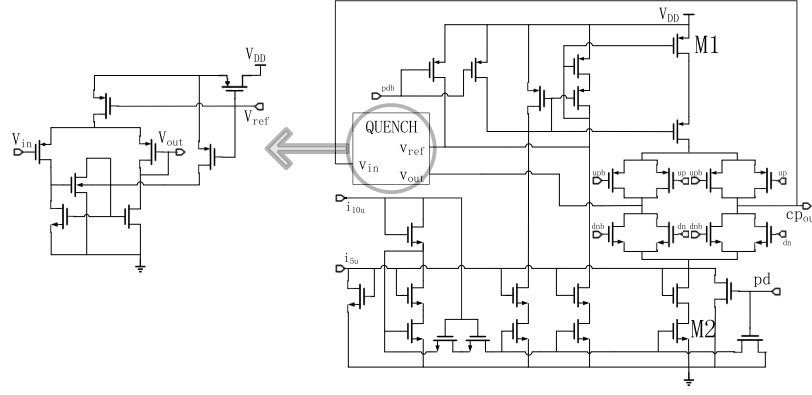
Fig. 9.   Schematic of the charge pump.

TABLE VIII
THE OPTIMIZATION RESULTS OF THE VCO

Algo	WEIBO	GASPAD	MSP	DE	PSO	SA
k(MHz/V)	355	368	368	368	362	361
r	0.997	0.998	0.999	0.998	0.999	0.999
FOM(mean)	**3.81**	3.90	4.08	4.08	4.04	4.16
FOM(median)	**3.82**	3.88	4.08	4.08	4.05	4.16
FOM(best)	**3.78**	3.82	3.84	3.96	3.91	3.99
FOM(worst)	**3.82**	4.01	4.19	4.21	4.16	4.30
Avg. # Sim	**181**	917	603	577	813	833
# Success	**12/12**	12/12	12/12	12/12	10/12	9/12

TABLE IX
THE OPTIMIZATION RESULTS OF THE CHARGE PUMP

Algo	WEIBO	GASPAD	MSP	DE	PSO	SA
$diff_1$	6.58	6.83	17.81	17.97	fail	fail
$diff_2$	5.30	5.28	16.82	15.49	fail	fail
$diff_3$	0.24	0.29	1.51	1.84	fail	fail
$diff_4$	0.37	0.40	2.57	3.56	fail	fail
deviation	0.41	0.33	0.38	0.39	fail	fail
FOM	**3.95**	4.00	11.80	11.85	fail	fail
FOM	**3.97**	4.99	11.67	12.31	fail	fail
FOM	**3.48**	3.74	8.26	9.29	fail	fail
FOM	**4.48**	4.43	14.03	13.40	fail	fail
Avg. # Sim	**790**	2328	1599	1538	fail	fail
# Success	**12/12**	12/12	12/12	12/12	0/12	0/12

In this experiment, we also ran all the optimization algorithms for 12 times to average out the random fluctuations. We first ran our proposed WEIBO algorithm to optimize the VCO. We limited the maximum number of simulations to 300 for WEIBO algorithm, and out of the 300 simulations, the number of initial sampled points is 120. For the other algorithms, we limited the maximum number of simulations to 1000.

The optimization results are shown in Table VIII. As can be seen in Table VIII, our proposed WEIBO algorithm and GASPAD can achieve a much better result than the other algorithms. Our proposed WEIBO algorithm can find the best *FOM* with the least number of simulations. Specifically, the number of simulations consumed by our proposed WEIBO algorithm is about 20% of that consumed by GASPAD.

F. Charge Pump

A charge pump circuit is shown in Figure 9, which is implemented in a SMIC 40nm process. There are 36 design variables in this circuit. A total of 18 PVT corners were considered. The design goal is that the current of M1 and M2 should be as close to $40\mu A$ across all PVT corners.

We performed transient simulations of this circuit over all PVT corners. For each corner, we measured the maximum, minimum and average currents of M1 and M2 respectively. The design specifications are given as

$$\text{minimize } FOM$$
$$\text{s.t. } diff_1 < 20\,\mu A,$$
$$diff_2 < 20\,\mu A,$$
$$diff_4 < 5\,\mu A,$$
$$deviation < 5\,\mu A, \tag{35}$$

where

$$\begin{cases} diff_1 = \max_{\forall PVT}(I_{M1,max} - I_{M1,avg}), \\ diff_2 = \max_{\forall PVT}(I_{M1,avg} - I_{M1,min}), \\ diff_3 = \max_{\forall PVT}(I_{M2,max} - I_{M2,avg}), \\ diff_4 = \max_{\forall PVT}(I_{M2,avg} - I_{M2,min}), \\ diff = \Sigma_{i=1}^{4} diff_i, \\ deviation = \max_{\forall PVT}(|I_{M1,avg} - 40\mu A|), \\ + \max_{\forall PVT}(|I_{M2,avg} - 40\mu A|), \\ FOM = 0.3 \times diff + 0.5 \times deviation. \end{cases} \tag{36}$$

In this experiment, we also ran all the optimization algorithms for 12 times to average out the random fluctuations. We first ran our proposed WEIBO algorithm to optimize the charge pump circuit. The maximum number of simulations is set to 1000. The number of initial samples is 120. For the other algorithms, we limited the maximum number of simulations to 2500. We marked the run "fail" if the number of simulations exceeded 2500 while the feasible solution was not found.

The optimization results are shown in Table IX. We can see from Table IX that PSO and SA performed poorly on this circuit, they failed to find feasible solutions in all the 12 runs. Although DE and MSP can find feasible designs, the designs they found in 2500 simulations are much worse than those found by WEIBO and GASPAD. Our proposed WEIBO algorithm can find the best *FOM* with the least number of simulations. The number of simulations consumed by our proposed WEIBO algorithm is about 33.9% of that consumed by GASPAD.
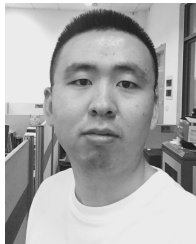
VII. Conclusion

In this paper, we propose a Weighted Expected Improvement based Bayesian Optimization (WEIBO) approach for automated analog circuit sizing. Gaussian processes (GP) are used as the online surrogate models for circuit performances. Weighted Expected improvement (wEI) is selected as the acquisition function to balance the exploration and exploitation during the optimization procedure. The acquisition function is maximized through a multi-start gradient-based approach. We extend the Bayesian optimization algorithm to handle multi-objective optimization problems. The single-objective and multi-objective Bayesian optimization algorithms were tested with several test cases. Compared with the state-of-the-art methods listed in the paper, the experimental results demonstrate that the proposed approach can achieve better optimization results with the significantly fewer number of simulations.

References

[1] R. A. Rutenbar, G. G. E. Gielen, and J. Roychowdhury, "Hierarchical modeling, optimization, and synthesis for system-level analog and RF designs," *Proc. IEEE*, vol. 95, no. 3, pp. 640–669, Mar. 2007.

[2] S. Boyd, S.-J. Kim, L. Vandenberghe, and A. Hassibi, "A tutorial on geometric programming," *Optim. Eng.*, vol. 8, no. 1, p. 67, Mar. 2007.

[3] D. M. Colleran *et al.*, "Optimization of phase-locked loop circuits via geometric programming," in *Proc. IEEE Custom Integr. Circuits Conf.*, Sep. 2003, pp. 377–380.

[4] M. del Mar Hershenson, "Design of pipeline analog-to-digital converters via geometric programming," in *Proc. IEEE/ACM Int. Conf. Comput. Aided Design (ICCAD)*, Nov. 2002, pp. 317–324.

[5] W. Daems, G. Gielen, and W. Sansen, "Simulation-based generation of posynomial performance models for the sizing of analog integrated circuits," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 22, no. 5, pp. 517–534, May 2003.

[6] Y. Wang, M. Orshansky, and C. Caramanis, "Enabling efficient analog synthesis by coupling sparse regression and polynomial optimization," in *Proc. 51st ACM/EDAC/IEEE Design Autom. Conf. (DAC)*, Jun. 2014, pp. 1–6.

[7] O. Okobiah, S. P. Mohanty, and E. Kougianos, "Exploring Kriging for fast and accurate design optimization of nanoscale analog circuits," in *Proc. IEEE Comput. Soc. Annu. Symp. VLSI*, Jul. 2014, pp. 244–247.

[8] O. Okobiah, S. Mohanty, and E. Kougianos, "Fast design optimization through simple Kriging metamodeling: A sense amplifier case study," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 22, no. 4, pp. 932–937, Apr. 2014.

[9] T. McConaghy and G. Gielen, "Analysis of simulation-driven numerical performance modeling techniques for application to analog circuit optimization," in *Proc. IEEE Int. Symp. Circuits Syst.*, vol. 2. May 2005, pp. 1298–1301.

[10] R. Phelps, M. Krasnicki, R. A. Rutenbar, L. R. Carley, and J. R. Hellums, "Anaconda: Simulation-based synthesis of analog circuits via stochastic pattern search," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 19, no. 6, pp. 703–717, Jun. 2000.

[11] R. A. Vural and T. Yildirim, "Analog circuit sizing via swarm intelligence," *AEU-Int. J. Electron. Commun.*, vol. 66, no. 9, pp. 732–740, Sep. 2012.

[12] B. Liu *et al.*, "Analog circuit optimization system based on hybrid evolutionary algorithms," *Integr., VLSI J.*, vol. 42, no. 2, pp. 137–148, Feb. 2009.

[13] G. Huang, L. Qian, S. Saibua, D. Zhou, and X. Zeng, "An efficient optimization based method to evaluate the DRV of SRAM cells," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 60, no. 6, pp. 1511–1520, Jun. 2013.

[14] A. Nieuwoudt and Y. Massoud, "Multi-level approach for integrated spiral inductor optimization," in *Proc. 42nd Design Autom. Conf.*, Jun. 2005, pp. 648–651.

[15] B. Peng, F. Yang, C. Yan, X. Zeng, and D. Zhou, "Efficient multiple starting point optimization for automated analog circuit optimization via recycling simulation data," in *Proc. Design, Autom. Test Eur. Conf. Exhibit. (DATE)*, Mar. 2016, pp. 1417–1422.

[16] B. Liu, D. Zhao, P. Reynaert, and G. G. E. Gielen, "Synthesis of integrated passive components for high-frequency RF ICs based on evolutionary computation and machine learning techniques," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 30, no. 10, pp. 1458–1468, Oct. 2011.

[17] B. Liu, Y. He, P. Reynaert, and G. Gielen, "Global optimization of integrated transformers for high frequency microwave circuits using a Gaussian process based surrogate model," in *Proc. Design, Autom. Test Eur. Conf. Exhibit.*, Mar. 2011, pp. 1–6.

[18] B. Liu, N. Deferm, D. Zhao, P. Reynaert, and G. G. E. Gielen, "An efficient high-frequency linear RF amplifier synthesis method based on evolutionary computation and machine learning techniques," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 31, no. 7, pp. 981–993, Jul. 2012.

[19] C.-A. Tugui, "Design methodology for high-performance circuits based on automatic optimization methods," Ph.D. dissertation, Dept. Signal Process. Electron. Syst., Supélec, Gif-sur-Yvette, France, Jan. 2013.

[20] B. Liu, D. Zhao, P. Reynaert, and G. G. E. Gielen, "GASPAD: A general and efficient mm-wave integrated circuit synthesis method based on surrogate model assisted evolutionary algorithm," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 33, no. 2, pp. 169–182, Feb. 2014.

[21] R. T. Marler and J. S. Arora, "Survey of multi-objective optimization methods for engineering," *Struct. Multidisciplinary Optim.*, vol. 26, no. 6, pp. 369–395, Apr. 2004.

[22] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, Apr. 2002.

[23] Q. Zhang and H. Li, "MOEA/D: A multiobjective evolutionary algorithm based on decomposition," *IEEE Trans. Evol. Comput.*, vol. 11, no. 6, pp. 712–731, Dec. 2007.

[24] N. Lourenço and N. Horta, "GENOM-POF: Multi-objective evolutionary synthesis of analog ICs with corners validation," in *Proc. 14th Annu. Conf. Genet. Evol. Comput.*, 2012, pp. 1119–1126.

[25] R. Martins, N. Lourenço, S. Rodrigues, J. Guilherme, and N. Horta, "AIDA: Automated analog IC design flow from circuit level to layout," in *Proc. Int. Conf. Synth., Modeling, Anal. Simulation Methods Appl. Circuit Design (SMACD)*, Sep. 2012, pp. 29–32.

[26] A. Canelas, R. Martins, R. Póvoa, N. Lourenço, and N. Horta, "Efficient yield optimization method using a variable K-Means algorithm for analog IC sizing," in *Proc. Design, Autom. Test Eur. Conf. Exhibit. (DATE)*, Mar. 2017, pp. 1201–1206.

[27] B. Liu, F. V. Fernández, Q. Zhang, M. Pak, S. Sipahi, and G. Gielen, "An enhanced MOEA/D-DE and its application to multiobjective analog cell sizing," in *Proc. IEEE Congr. Evol. Comput.*, Jul. 2010, pp. 1–7.

[28] M. Fakhfakh, Y. Cooren, A. Sallem, M. Loulou, and P. Siarry, "Analog circuit design optimization through the particle swarm optimization technique," *Analog Integr. Circuits Signal Process.*, vol. 63, no. 1, pp. 71–82, Apr. 2010.

[29] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. de Freitas, "Taking the human out of the loop: A review of Bayesian optimization," *Proc. IEEE*, vol. 104, no. 1, pp. 148–175, Jan. 2016.

[30] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA, USA: MIT Press, 2006.

[31] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2007.

[32] 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.

[33] B. Liu and A. Nikolaeva, "Efficient global optimization of MEMS based on surrogate model assisted evolutionary algorithm," in *Proc. Design, Autom. Test Eur. Conf. Exhibit. (DATE)*, Mar. 2016, pp. 555–558.

[34] S. J. Park, B. Bae, J. Kim, and M. Swaminathan, "Application of machine learning for optimization of 3-D integrated circuits and systems," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 25, no. 6, pp. 1856–1865, Jun. 2017.

[35] J. M. Hernández-Lobato, M. A. Gelbart, R. P. Adams, M. W. Hoffman, and Z. Ghahramani, "A general framework for constrained Bayesian optimization using information-based search," *J. Mach. Learn. Res.*, vol. 17, no. 160, pp. 1–53, 2016.

[36] 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.

[37] M. Schonlau, W. J. Welch, and D. R. Jones, *Global Versus Local Search in Constrained Optimization of Computer Models* (Lecture Notes-Monograph Series), vol. 34. Institute of Mathematical Statistics, pp. 11–25, 1998.

[38] J. Gardner, M. J. Kusner, Z. Xu, K. Q. Weinberger, and J. P. Cunningham, "Bayesian optimization with inequality constraints," in *Proc. The 31st Int. Conf. Mach. Learn.*, 2014, pp. 937–945.

[39] M. A. Gelbart, "Constrained Bayesian optimization and applications," Ph.D. dissertation, Dept. Biophys., Harvard Univ., Cambridge, MA, USA, 2015.

[40] J. Nocedal and S. Wright, *Numerical Optimization*. New York, NY, USA: Springer, 2006.

[41] K. Deb, "An efficient constraint handling method for genetic algorithms," *Comput. Methods Appl. Mech. Eng.*, vol. 186, nos. 2–4, pp. 311–338, Jun. 2000.

[42] S. G. Johnson. (2014). *The NLopt Nonlinear-Optimization Package*. [Online]. Available: http://ab-initio.mit.edu/nlopt

[43] P. Palmers, T. McConnaghy, M. Steyaert, and G. Gielen, "Massively multi-topology sizing of analog integrated circuits," in *Proc. Design, Autom. Test Eur. Conf. Exhibit. (DATE)*, Apr. 2009, pp. 706–711.

[44] T. McConaghy, P. Palmers, G. Gielen, and M. Steyaert, "Simultaneous multi-topology multi-objective sizing across thousands of analog circuit topologies," in *Proc. 44th Annu. Design Autom. Conf.*, Jun. 2007, pp. 944–947.

[45] I. Guerra-Gómez, E. Tlelo-Cuautle, T. McConaghy, and G. Gielen, "Optimizing current conveyors by evolutionary algorithms including differential evolution," in *Proc. 16th IEEE Int. Conf. Electron., Circuits, Syst. (ICECS)*, Dec. 2009, pp. 259–262.

[46] S. Cheng, Y. Shi, and Q. Qin, "On the performance metrics of multiobjective optimization," in *Proc. Int. Conf. Adv. Swarm Intell.*, 2012, pp. 504–512.

[47] Z. Yan, P.-I. Mak, M.-K. Law, and R. P. Martins, "A 0.016-mm^2 144-μW three-stage amplifier capable of driving 1-to-15 nF capacitive load with >0.95-MHz GBW," *IEEE J. Solid-State Circuits*, vol. 48, no. 2, pp. 527–540, Feb. 2013.

Changhao Yan received the B.E. and M.E. degrees from the Huazhong University of Science and Technology, Wuhan, China, in 1996 and 2002, respectively, and the Ph.D. degree in computer science from Tsinghua University, Beijing, China, in 2006.

He is currently an Associate Professor with the Microelectronics College, Fudan University, Shanghai, China. His current research interests include the parasitic parameter extraction of interconnects, parallel algorithms for large scale computation, design for manufacturability, the robust analysis of circuits, and chemical mechanical polishing modeling.

Zhiliang Hong received the B.S. degree from the Chinese University of Science and Technology, Hefei, China, and the Ph.D. degree from the Swiss Federal Institute of Technology (ETHZ), Zurich, Switzerland.

He was the Leader of the Chinese Delegation in co-operation design with Venus System in TU Berlin in 1987, and an Associate Researcher in Berkeley in 1989. He was a Guest Professor with the University of Hannover, Germany, from 1992 to 1994, and a Guest Researcher at ETHZ in 2000. He is currently a Professor with Fudan University, Shanghai, China. He has authored books *Computer Architecture and RISC Design* (Fudan Press, 2005) and *Analog Integrated Circuit Analysis and Design* (Science Press, 2009). He has authored over 300 technical papers in analog, mixed signal, RF integrated circuits, and system-on-chip. He serves as the Editor of the *Journal of Research and Advance of Solid-State Electronics*, and the Program Chair of ASICON in 2005 and 2010.

Wenlong Lyu received the B.S. degree in microelectronics from Fudan University, Shanghai, China, in 2014.

He is currently pursuing the Ph.D. degree with the State Key Laboratory of Application Specific Integrated Circuits and System, Microelectronics Department, Fudan University. His current research interests include analog circuit design automation and optimization.

Pan Xue received the B.S. degree in microelectronics from Xi'an Jiaotong University, Xi'an, China, in 2013, and the M.S. degree in microelectronics from Fudan University, Shanghai, China, in 2016, where he is currently pursuing the Ph.D. degree in microelectronics. His current research interests include analog IC design and RFIC design.

Xuan Zeng (M'97) received the B.S. and Ph.D. degrees in electrical engineering from Fudan University, Shanghai, China, in 1991 and 1997, respectively.

She was a Visiting Professor with the Department of Electrical Engineering, Texas A&M University, and with the Microelectronics Department, Technische Universiteit Delft, Delft, The Netherlands, in 2002 and 2003, respectively. She served as the Director of State Key Laboratory of Application Specific Integrated Circuits and Systems, Fudan University, from 2008 to 2012. She is currently a Full Professor with the Microelectronics Department. Her current research interests include design for manufacturability, high-speed interconnect analysis and optimization, circuit simulation.

Fan Yang (M'08) received the B.S. degree from Xi'an Jiaotong University in 2003, and the Ph.D. degree from Fudan University in 2008.

From 2008 to 2011, he was an Assistant Professor with Fudan University. He is currently an Associate Professor with the Microelectronics Department, Fudan University. His research interests include model order reduction, circuit simulation, high-level synthesis, yield analysis, and design for manufacturability.

Dian Zhou (M'89–SM'07) received the B.S. degree in physics and the M.S. degree in electrical engineering from Fudan University, China, in 1982 and 1985, respectively, and the Ph.D. degree in electrical and computer engineering from the University of Illinois in 1990.

He joined the University of North Carolina at Charlotte as an Assistant Professor in 1990, where he became an Associate Professor in 1995. He joined The University of Texas at Dallas as a Full Professor in 1999. His research interests include high-speed VLSI systems, CAD tools, mixed-signal ICs, and algorithms.