

An Optimization-Based Approach for Analog Circuit Technology Migration

1st Kostas Touloupas

National Technical University of Athens

Zografou, Greece

ktouloupas@mail.ntua.gr

2nd Paul Peter Sotiriadis

National Technical University of Athens

Zografou, Greece

pps@ieee.org

Abstract—In this paper a new approach for automatic technology migration of analog Integrated Circuits (ICs) is described. The proposed approach uses a software interface between commercial simulators and a Bayesian Optimization (BO) algorithm to automatically size a fixed topology in the target technology node, given a sized schematic in a source technology as golden standard. By using a new acquisition function for BO, we are able to simulate and examine multiple candidate solutions in parallel, thereby reducing the total runtime of the procedure. The proposed approach is demonstrated on two real world circuits, by migrating their schematics from a TSMC 90nm to a TSMC 40nm technology.

Index Terms—migration, optimization, analog circuits

I. INTRODUCTION

MODERN era systems require high levels of integration and low power utilization, rendering circuit design a tedious procedure. In the case of analog and Radio-Frequency (RF) circuits, new applications such as Internet-of-Things impose strict specifications and require extensive verification prior to tape-out. Automated tools to assist manual labour for analog circuit design could provide a remedy to this situation. However, they have not yet reached maturity, unlike their digital counterparts. Therefore, the development of tools to assist analog designers is a research direction that can have a crucial impact on the semiconductor industry.

A particular task that designers frequently come across is the fabrication technology migration one; A preselected circuit topology that has already been sized in an initial (source) technology must be re-designed in another (target) one. The performance of the initial design must be preserved during this procedure. Although sizing the initial circuit in the source technology provides a set of empirical rules to guide the migration, this procedure is often time consuming and requires many verification trials to complete successfully. A tool that automatically performs circuit migration can reduce the design-to-market time and thereby the design costs.

Automated technology migration of analog circuits is addressed in the literature by two main approaches, simulation-based ones and model based ones. The latter includes [1], where the devices are scaled to preserve the transconductance g_m in the target technology. This procedure is followed by a tuning step that requires a user-provided qualitative dependency matrix to complete. In [2], device scaling factors are derived using a MOSFET compact model and the

transconductance-to-current ratio g_m/I_d methodology. Scaling factors constitute the core of the approaches in [3], [4]. Though simple, the aforementioned approaches have the following disadvantage; they try to preserve circuit performances using simplified equations and small signal parameters and do not yield accurate results for the combined small and large signal behavior of the target circuit.

In the case of simulation-based approaches, a computer software automates the procedure of simulating parametrized testbenches and feeding the simulation outputs in an optimization algorithm. This approach yields accurate results since the candidate designs are evaluated with commercial simulators and they require no further inputs from the designer. In this category fall the population-based Anaconda [5] and the multi-start local search method in [6]. The main drawback of these approaches is the computational cost. They typically require many evaluations to reach acceptable solutions, which in the case of analog circuit simulations are time consuming.

Motivated by the above, in this paper we propose a new method for analog circuit technology migration using simulation-based evaluation. By using a Bayesian Optimization (BO) [7] algorithm as the core of our approach, we are able to balance the exploration-exploitation tradeoff of the loss landscape and reduce the number of simulations required. Also, a Thompson sampling acquisition function is used, that enables the selection of multiple points for evaluation, therefore allowing for parallelized simulations to take place. The proposed approach is applied on two real world circuits and manages to re-size them from a TSMC 90nm to a TSMC 40nm fabrication technology within minutes.

The paper is structured as follows. Section II formulates the schematic migration procedure as an optimization problem and provides with information about its automation. Section III presents BO and its mathematic foundations, along with the proposed acquisition function. Section IV demonstrates the application of the overall approach on a Two-Stage and a Four-Stage CMOS amplifier. Finally, Section V concludes the paper.

II. PROBLEM FORMULATION

Consider an already sized circuit in the source technology which is parametrized by d variables arranged in a vector \mathbf{x} . These variables may include transistor widths, transistor lengths, capacitances, etc. This circuit provides a set of scalar

performance metrics $\{P_i^s\}_{i=1}^m$ that describe its behavior. Migrating the circuit to a target fabrication technology necessitates the target performance metrics, $\{P_i^t\}_{i=1}^m$, to be equal or better than the initial ones. For instance, in the case of an operational amplifier, DC Gain in the target technology should be equal or greater than in the source technology (case 1), whereas power dissipation in the target technology must equal or less than in the source technology (case 2). To account for both cases provided above, we introduce a weighting function

$$W(x) := \begin{cases} -1 & \text{for case 1} \\ 1 & \text{for case 2} \end{cases}. \quad (1)$$

For the following we assume that the target fabrication technology is smaller than the source one. Besides the aforementioned constraints for performance metrics, in our formulation we explicitly optimize for target circuit power dissipation. We consider that power consumption, P_{dc} , is the most important factor when re-sizing a schematic and opt to exploit any benefits in terms of low power utilization the target technology has to offer. The migration procedure can now be formulated as

$$\begin{aligned} \min \quad & P_{dc}(\mathbf{x}), \quad \mathbf{x} = [x_1, x_2, \dots, x_d] \\ \text{s.t.} \quad & g_j(\mathbf{x}) \leq 0, \quad j = 1, \dots, m \\ & L_i \leq x_i \leq U_i, \quad i = 1, \dots, d \end{aligned} \quad (2)$$

where L_i and U_i are the lower and upper bounds of the i -th variable and the j -th constraint is defined as follows

$$g_j(\mathbf{x}) = W(P_j^t(\mathbf{x}) - P_j^s). \quad (3)$$

The variable space is $\mathbb{S} = \prod_{i=1}^d [L_i, U_i]$ and the degree of constraint violation for a given \mathbf{x} is defined as

$$\text{CV}(\mathbf{x}) = \sum_j \max[0, g_j(\mathbf{x})]. \quad (4)$$

The above minimization problem requires circuit performances to be computed for different parameter vectors \mathbf{x} . The evaluation is done using the commercial simulator Cadence Spectre. After the simulation, the performances of each candidate parametrization of the circuit in the target technology are parsed and fed to an algorithm that determines a new set of candidate vectors, until a certain termination criterion is met. This simulation-in-the-loop procedure is conceptually illustrated in Fig. 1.

In our case, the automation procedure is operated by a software tool written in Python. It implements the optimization algorithm (see Section III) and provides an interface to Cadence Spectre for simulation automation and result parsing. Furthermore, it utilizes multiple threads and takes advantage of the batched mode of the simulator, to speed up the evaluation procedure.

III. BAYESIAN OPTIMIZATION

In this section the BO algorithm for automated circuit technology migration is presented. Prior to discussing the functionality of BO, we provide with details about the models that constitute its core, Gaussian Processes (GPs).

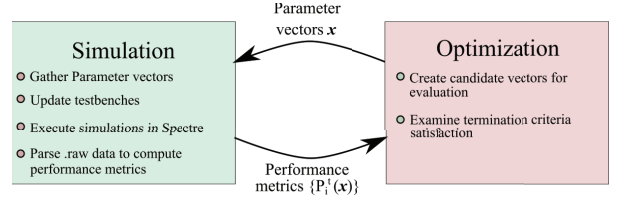


Fig. 1. Simulation-in-the-loop optimization.

A. Gaussian Processes

Consider a variable space \mathbb{S} and let $\mathbf{X} = \{\mathbf{x}_i\}_{i=1}^N$ be the inputs in \mathbb{S} of an unknown function $f: \mathbb{S} \rightarrow \mathbb{R}$. A GP is a non-parametric regression model that approximates f . In contrast to deterministic models, GPs are probabilistic in nature and provide with uncertainty estimates about their predictions, by defining a probability distribution over functions in \mathbb{S} [9].

GPs are uniquely defined by two components; a mean function $m: \mathbb{S} \rightarrow \mathbb{R}$ and a kernel function $k: \mathbb{S} \times \mathbb{S} \rightarrow \mathbb{R}$. Consider N noise corrupted observations from f , $\mathbf{y} = \{y_i\}_{i=1}^N$, with $y_i = f(\mathbf{x}_i) + \epsilon_i$, where $\epsilon_i \sim \mathcal{N}(0, \sigma_n^2)$. In GP regression we say that f follows a GP

$$f(\mathbf{x}) \sim \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')), \quad (5)$$

to denote the probabilistic nature of the model. A fundamental property of GP models is that vector $\mathbf{f} = [f(\mathbf{x}_i)]_{i=1}^N$ follows a Multivariate Gaussian distribution for any positive integer N , such that

$$\mathbf{f} | \mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, K), \quad (6)$$

Here, the $(N \times 1)$ mean vector is $\boldsymbol{\mu} = [m(\mathbf{x}_i)]_{i=1}^N$ and the $(N \times N)$ covariance matrix K is defined such that $K_{ij} = k(\mathbf{x}_i, \mathbf{x}_j) + \sigma_n^2 \delta_{ij}$, where δ_{ij} is the Kronecker delta.

The mean function $m(\mathbf{x})$ describes the shape of the unknown function f . In cases when this information is not available, we set $m(\mathbf{x})$ to zero. The kernel function describes how similar the outputs of any two input points are. In our work, we use the Matérn 5/2 kernel to account for non-smooth functions,

$$k(\mathbf{x}_i, \mathbf{x}_j) = \sigma^2 \left(1 + \sqrt{5}r + \frac{5}{3}r^2 \right) e^{-\sqrt{5}r} \quad (7)$$

where

$$r = \left(\sum_{k=1}^d \frac{(x_{i,k} - x_{j,k})^2}{\lambda_k^2} \right)^{1/2}. \quad (8)$$

The hyperparameters λ_k , σ_n and σ are arranged in a vector $\boldsymbol{\theta}$. To adjust a GP model to the observations \mathbf{y} , one must discover the values of $\boldsymbol{\theta}$ by minimizing the negative log marginal likelihood

$$L(\boldsymbol{\theta}) = \frac{1}{2} \mathbf{y}^T K^{-1} \mathbf{y} + \frac{1}{2} \log(|K|) + \frac{N}{2} \log(2\pi), \quad (9)$$

where the size of \mathbf{y}^T is $(1 \times N)$.

Consider a point $\mathbf{x}^* \notin \mathbf{X}$. To predict $f(\mathbf{x})$ at \mathbf{x}^* , we resort to the predictive distribution $p(f(\mathbf{x}^*)|\mathbf{y})$. This is a Gaussian distribution with mean and variance

$$\begin{aligned}\mu_{f|\mathbf{y}}(\mathbf{x}^*) &= \mathbf{k}^T K^{-1} \mathbf{y} \\ \sigma_{f|\mathbf{y}}^2(\mathbf{x}^*) &= c - \mathbf{k}^T K^{-1} \mathbf{k}\end{aligned}\quad (10)$$

Here, \mathbf{k}^T is a $(1 \times N)$ vector with values $k(\mathbf{x}_i, \mathbf{x}^*)$ for $i = 1, \dots, N$ and $c = k(\mathbf{x}^*, \mathbf{x}^*)$.

To compute the joint predictive distribution $p([f(\mathbf{x}_1^*), \dots, f(\mathbf{x}_k^*)]|\mathbf{y})$ over k unseen points in \mathbb{S} , one uses a Multivariate Gaussian distribution with mean such as in Eq. 10. The covariance matrix entry between any two points is given by

$$\text{Cov}(\mathbf{x}^*, \mathbf{x}^{\prime}) = k(\mathbf{x}^*, \mathbf{x}^{\prime}) - \mathbf{k}_{\mathbf{X}, \mathbf{x}^*}^T K^{-1} \mathbf{k}_{\mathbf{X}, \mathbf{x}^{\prime}}, \quad (11)$$

where the $(1 \times N)$ vector $\mathbf{k}_{\mathbf{X}, \mathbf{x}^*}^T = [k(\mathbf{x}_i, \mathbf{x}^*)]_{i=1}^N$.

B. Bayesian Optimization

The BO optimization framework is a model-based approach to global optimization of black-box functions. It is particularly useful in cases when the evaluation of the unknown function to be minimized is time consuming, since it manages to approximate the global optimum by using relatively few evaluations. This is achieved by using the uncertainty information provided by the employed GP models to guide the search towards promising parts of the variable space, by balancing exploration and exploitation. BO therefore is the most promising approach to low budget optimization problems and has found numerous applications in various fields of studies, such as machine learning and robotics [7].

BO is comprised of two main components; the surrogate models, which are GPs that approximate the unknown function f and the constraints functions, and an acquisition function $\alpha(\cdot)$. Acquisition functions use GP model predictions and uncertainty estimates to provide with a score of goodness, or utility score, about any point in the search domain \mathbb{S} . In particular, acquisition functions such as expected improvement (EI), probability of improvement (PI) and lower confidence bound (LCB) [7] make use of pointwise GP distributions to assign the utility for expensive evaluation to each point in the search space. Selecting a query point \mathbf{x}^* for evaluation therefore reduces to the maximization of the employed acquisition function.

The complete functionality of the BO algorithm is shown in Algorithm 1. Starting from a random sampling of the variable space and after evaluating the initial samples, an archive of past evaluations is created and stored in sets \mathbf{X} , \mathbf{y} . The GP models that approximate the objective and constraint functions are trained and then the iterative procedure of optimizing $\alpha(\cdot)$ and evaluating the query points begins. The procedure terminates when a certain number of iterations has been reached.

C. Employed Acquisition Function

The functionality of the basic BO described in Algorithm 1 assumes that the maximization of the acquisition function

Algorithm 1: BO Algorithm

Input : Initial samples N_{init} , number of iterations T_{max} , variable space \mathbb{S}

Output: Global minimum \mathbf{x}_{best}

- 1 Create randomly a set \mathbf{X} of N_{init} initial samples from \mathbb{S}
 - 2 Evaluate \mathbf{X} to acquire observations \mathbf{y}
 - 3 **for** $i = 1, \dots, T_{max}$ **do**
 - 4 Adjust GP models using Eq. 9
 - 5 $\mathbf{x}^* \leftarrow \text{argmax}_{\mathbf{x} \in \mathbb{S}} \alpha(\mathbf{x}, \mathbf{y})$
 - 6 Evaluate \mathbf{x}^* to acquire y^*
 - 7 Update archive $\mathbf{X} \leftarrow \mathbf{X} \cup \{\mathbf{x}^*\}$, $\mathbf{y} \leftarrow \mathbf{y} \cup \{y^*\}$
 - 8 **end**
 - 9 Find \mathbf{x}_{best} from \mathbf{X} , \mathbf{y}
-

results in a single query point \mathbf{x}^* . In fact, this is the case for the most popular acquisitions functions such EI and LCB and constitutes a disadvantage of BO. By relying on a single query point at each iteration, modern hardware systems that enable parallelized evaluations of multiple candidate points are not exploited. In addition, GP training, which is a relatively time consuming procedure, takes place after each evaluation rendering the overall procedure time consuming.

Taking the above into consideration, in this work we employ a parallelizable acquisition function that provides with multiple query points. By examining many points in parallel, we are able to gather more information in the same time-frame. The proposed acquisition function is based on Thompson sampling (TS) [11], which is a randomized selection strategy. Instead of relying on GP pointwise predictive distributions, TS uses the joint predictive distributions of the models over a quantization of the search domain \mathbb{S} . To generate the quantization of \mathbb{S} , at each iteration, we use a quasi-random number generator [12] to produce k candidate points $\{\mathbf{x}\}_{i=1}^k$. Then, we sample from the joint predictive Multivariate Gaussian distribution to produce a sample of the unknown objective or constraint function modeled by each GP.

To account for the constrained optimization formulation defined in Eq. 2, the query point selection must take into account the constraint and objective function values simultaneously. This is done by using the feasibility rule [8], which compares the k candidate points in pairs and selects a single one based on the following

- Feasible candidate solutions are preferred than infeasible ones
- amongst feasible solutions, the ones with better fitness function are preferred and,
- amongst infeasible solutions, the ones with the least constraint violation are preferred.

This procedure extends to multiple query points by simply drawing $N_s > 1$ samples from the GP joint predictive distributions and selecting a single query point from each one of them. Fig. 2 shows a case where $N_s = 30$ samples are drawn from a quantization (40 samples) of the variable space [3, 7]. The locations of the query points associated with each sample are also shown.

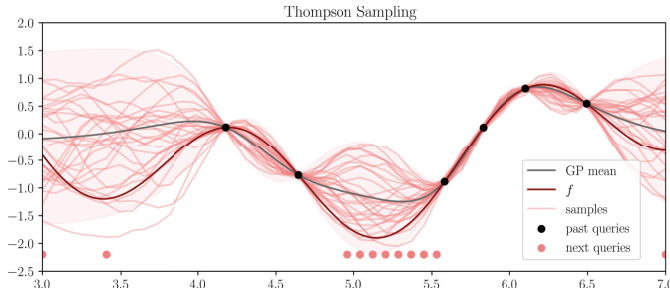


Fig. 2. A GP model's 95% confidence bounds, predictive mean and past query points are shown here. The true function f along with 30 samples drawn from the posterior distribution are also shown, with the next query point locations marked as pink dots.

IV. APPLICATIONS

In this section the BO algorithm described previously is used to migrate two CMOS amplifiers based on the formulation of Eq. 2. Both circuits are designed in Cadence Virtuoso in a TSMC 90nm technology, and Cadence Spectre is used for simulation in the target TSMC 40nm technology. The experiments were executed on a 8 core linux workstation.

A. Two Stage Amplifier

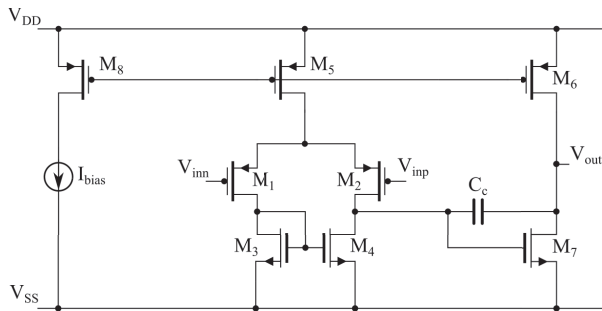


Fig. 3. Two stage amplifier.

Fig. 3 depicts the topology of the already sized circuit in the source technology. It consists of a differential input stage with active load and a common source output stage, with a capacitor for frequency compensation. For this example, the variables that parametrize the topology include three transistor lengths for transistor sets $\{M_8, M_5, M_6\}$, $\{M_1, M_2\}$, $\{M_3, M_4, M_7\}$, six transistor widths, the capacitance of C_c and the biasing current. In both technologies it holds $V_{DD} = 1.2V$.

For this experiment, we examine the following the specifications: DC Gain, Phase Margin (PM), Unity Gain Frequency (UGF), average Slew Rate (SR_{avg}) and power consumption P_{dc} . The variable space \mathcal{S} is as follows: the variable range for transistor lengths is $[0.1, 2]\mu m$, for transistor widths $[1, 100]\mu m$, for C_c $[0.5, 5]pF$ and for I_{bias} $[1, 40]\mu A$. The golden standard specifications from the source technology are given in detail in Table I. To acquire these metrics for the target circuit, two separate testbenches are used, one for AC and DC analysis and one for transient analysis.

TABLE I
GOLD STANDARD FOR TWO STAGE AMPLIFIER

Performance Metrics	Specifications
DCGain	67.46dB
P_{dc}	107 μW
SR_{avg}	3.6V/ μs
UGF	15.06MHz
PM	45.9 $^\circ$
V_{DD}	1.2V

For the optimization procedure we use $N_{init} = 100$ initial random samples and the maximum number of iterations is 500. To speedup the optimization procedure, we derive $N_s = 4$ query points from the employed acquisition function. The number of points produced by the quasi-random generator to quantize the variable space is 2000. The optimization took approximately 15 minutes and the performances of the circuit in target technology are given in Table II.

TABLE II
TWO STAGE AMPLIFIER - 40NM PERFORMANCE

Performance Metrics	Specifications
DCGain	67.86dB
P_{dc}	29 μW
SR_{avg}	3.71V/ μs
UGF	16.05MHz
PM	46.6 $^\circ$
V_{DD}	1.2V

It is seen that the optimization algorithm manages to find a solution with equal or better performances than the source circuit specifications. A considerable reduction in power consumption is shown, which is mainly due to the problem formulation.

B. Four Stage Amplifier

Fig. 4 depicts the Four Stage amplifier [10] examined in this subsection. It employs an active zero sub-circuit, a slew-rate enhancer sub-circuit and four gain stages. Similar to the previous case, we wish to re-size the circuit from a gold standard version in TSMC 90nm to a new version in TSMC 40nm. Besides the core amplifier shown in Fig. 4, a biasing sub-circuit which is responsible for $V_{bn1}, V_{bn2}, V_{bp1}, V_{bp2}$ is also included in the testbenches. In total, 35 transistors, 2 capacitors, a single resistor and a current source are employed. We use two testbenches, one for the slew rate and one for AC and DC analysis.

In total, there are 43 parameters, which include 20 transistor widths, 19 transistor lengths, a bias current and C_Z, C_M and R_Z . Transistor length and width ranges are the same as in the previous example, the biasing current range is $[0.5, 10]\mu A$, the range of the resistor is $[0.1, 300]k\Omega$ and the capacitance range is the same as in the previous example. The gold standard performance metrics are shown in Table III.

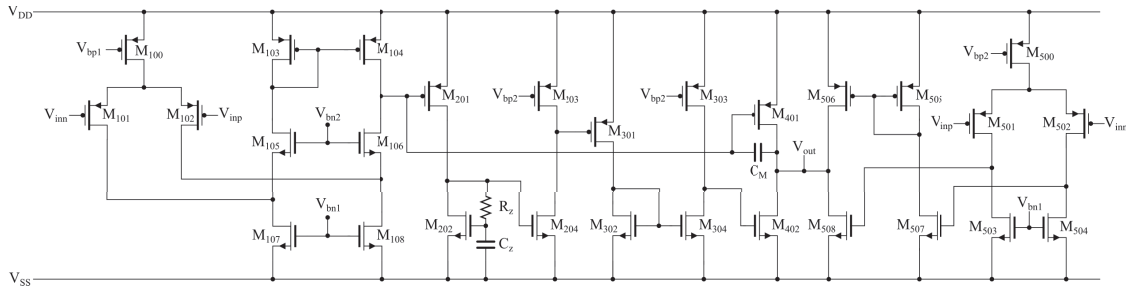


Fig. 4. Four stage amplifier of [10].

TABLE III
GOLD STANDARD FOR FOUR STAGE AMPLIFIER

Performance Metrics	Specifications
DCGain	102.3dB
P_{dc}	$230\mu W$
SR_{avg}	$0.189V/\mu s$
UGF	2.08MHz
PM	49°
GM	16.3dB
V_{DD}	1.2V

For the optimization procedure we use $N_{init} = 300$ initial samples and set the maximum number of iterations to 800. The number of query points per iteration is again $N_s = 4$. In this case we used more evaluations than in the previous example due to the large variable space. The optimization took approximately 42 minutes to complete and the performance metrics of the migrated schematic are given in Table IV.

TABLE IV
FOUR STAGE AMPLIFIER - 40NM PERFORMANCE

Performance Metrics	Specifications
DCGain	102.76dB
P_{dc}	$109\mu W$
SR_{avg}	$0.191V/\mu s$
UGF	2.09MHz
PM	52°
GM	16.4dB
V_{DD}	1.2V

It is seen that the target circuit surpasses the performance of the initial design. In terms of power dissipation, there is a drop in half, which can be attributed to the properties of the target technology, the problem formulation and the sub-optimality of the initial design.

V. CONCLUSION

An optimization-based approach to analog circuit schematic migration was presented in this paper. Using a simulation-in-the-loop approach, we were able to migrate sized circuits from a source to a target fabrication technology using a dedicated software platform. The formulation of the migration procedure

as an optimization problem and the BO algorithm employed in this work were explained. The application of the methodology on two circuits proved the effectiveness of the approach.

ACKNOWLEDGMENT

This research is co-financed by Greece and the European Union (European Social Fund- ESF) through the Operational Programme "Human Resources Development, Education and Lifelong Learning" in the context of the project "Strengthening Human Resources Research Potential via Doctorate Research" (MIS-5000432), implemented by the State Scholarships Foundation (IKY).

REFERENCES

- [1] K. Francken and G. Gielen, "Methodology for analog technology porting including performance tuning," in Proc. IEEE Int. Symp. Circuits Syst., vol. 1, Jul. 1999, pp. 415–418.
- [2] T. Yang, M. Gao, S. Wu, and D. Guo, "A new reuse method of analog circuit design for CMOS technology migration," in Proc. Int. Conf. ASIC, Jul. 2010, pp. 112–115.
- [3] C. Galup-Montoro, M. C. Schneider, and R. M. Coitinho, "Resizing rules for MOS analog-design reuse," IEEE Design Test Comput., vol. 19, no. 2, pp. 50–58, Apr. 2002.
- [4] T. Levi, J. Tomas, N. Lewis, and P. Fouillat, "A CMOS resizing methodology for analog circuits: Linear and non-linear applications," IEEE Design Test Comput., vol. 26, no. 1, pp. 78–87, Jan./Feb. 2009.
- [5] 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 Design Integr. Circuits Syst., vol. 19, no. 6, pp. 703–717, Jun. 2000.
- [6] Qian, Liuxi, et al. "Automated technology migration methodology for mixed-signal circuit based on multistart optimization framework," IEEE Transactions on Very Large Scale Integration (VLSI) Systems 23.11 (2014): 2595-2605.
- [7] 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," Proceedings of the IEEE, vol. 104, no. 1, pp. 148–175, 2015.
- [8] Kalyanmoy Deb, "An efficient constraint handling method for genetic algorithms," Computer Methods in Applied Mechanics and Engineering, Volume 186, Issues 2–4, 2000, 311-338
- [9] C. K. Williams and C. E. Rasmussen, Gaussian processes for machine learning. MIT press Cambridge, MA, 2006, vol. 2, no. 3.
- [10] S. A. Fordjour, J. Riad, and E. S'anchez-Sinencio, "A 175.2-mw 4-stage ota with wide load range (400 pf–12 nf) using active parallel compensation," IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 28, no. 7, pp. 1621–1629, 2020.
- [11] K. Kandasamy, A. Krishnamurthy, J. Schneider, and B. Poczos, "Parallelised bayesian optimisation via thompson sampling," in Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics, ser. Proceedings of Machine Learning Research, A. Storkey and F. Perez-Cruz, Eds., vol. 84. PMLR, 09–11 Apr 2018, pp. 133–142.
- [12] D. Eriksson, M. Pearce, J. Gardner, R. D. Turner, and M. Poloczek, "Scalable global optimization via local bayesian optimization," in Advances in Neural Information Processing Systems, 2019, pp. 5496–5507.