The Initialization Effect: Refining Bayesian Optimization for Circuit Design

Anusha James*, Zhi Hui Kong†, Sezin Kircali Ata*, Khin Mi Mi Aung*,
Kiat Seng Yeo†, Chuan Sheng Foo*§ and Ashish James*
*Institute for Infocomm Research (I^2R), Agency for Science, Technology and Research (A*STAR), Singapore
§Centre for Frontier AI Research (CFAR), Agency for Science, Technology and Research (A*STAR), Singapore
†Singapore University of Technology and Design (SUTD)

***Abstract*—Bayesian optimization (BO) has emerged as a powerful sample efficient technique for optimizing expensive and time-consuming design of analog circuits. The BO framework leverages probabilistic models to efficiently explore design spaces, allowing for a streamlined approach to identify optimal designs that meet specific performance targets. Traditionally, BO framework is initialized with predefined or randomly generated samples, thereby failing to fully exploit the latent insights concealed within the data leading to suboptimal results and lengthy convergence time. This paper delves into the significance of initialization in BO for analog circuit design, investigating its impact on the efficiency and effectiveness of the optimization process. The study demonstrates that the model can achieve optimal design outcomes using as few as 16 data points, highlighting the potential for data-efficient learning in complex design tasks. Through empirical analysis and experimentation, we explore various initialization strategies and their influence on optimization performance, considering factors such as convergence speed, solution quality etc. The results highlight the influence of initialization strategies, with approaches such as Latin hypercube sampling (LHS) and K-center greedy achieving faster convergence compared to other methods like K-means and random initialization. This study demonstrates how effective initialization strategies in BO can reduce the need for extensive data collection, by optimizing design outcomes with fewer iterations and samples. Furthermore, we elucidate the underlying mechanisms through which different initialization approaches affect the optimization process, providing insights into the interplay between initialization, exploration, and exploitation in analog circuit design optimization. Our findings shed light on the importance of thoughtful initialization strategies in harnessing the full potential of BO for analog circuit design, offering valuable guidelines for circuit designers.**

***Index Terms*—Regression, Analog Circuit Design, Operational Amplifier, Folded Cascode Operational Amplifier, Bayesian Optimization**

I. Introduction

In the realm of engineering design, the pursuit of optimal solutions often entails navigating complex and high-dimensional parameter spaces to achieve desired performance objectives. In this context, inverse design methodologies and Bayesian optimization (BO) [1] have emerged as powerful tools for efficiently exploring and optimizing design spaces, particularly in the domain of analog circuit design.

Analog circuit design poses unique challenges in comparison to digital design due to the inherent intricacies of dealing with continuous signals and multitude of other factors influencing their behavior. These challenges are further exacerbated with the advent of emerging technology nodes, the incessant demand for low-power and high-performance design, and the escalating complexities of electronic devices. Variables in analog circuits are highly interdependent, and the overall design space is considerably larger, necessitating a deep understanding and a careful balancing act to achieve optimal performance. Consequently, long design cycles are the main bottleneck in analog circuit design due to its heavy reliance on the experience and skill of the circuit designer, compounded by a notable absence of advanced design automation. Furthermore, the complex and inherent uncertainty in analog design can lead to multiple solutions that fulfill a specified predetermined set of requirements and initial conditions. Hence, the circuit designers are clearly tasked with navigating an iterative process, continuously making critical decisions to refine the design towards optimal performance.

To aid designers in this pursuit, a diverse array of methods and approaches have been proposed, spanning from traditional employment of matrix theory to more contemporary utilization of inverse design methodologies [2]. Several studies demonstrate the practical viability of evolutionary algorithms in analog circuit design, showcasing their potential in aspects such as time control, achieving efficient and adaptable design solutions, and optimization [3]. The domain of machine learning driven design for analog circuits is currently a subject of extensive exploration which could lead to reductions in design time, enhancements in design quality, and alleviation of the workload for both novice and experienced engineers.

Inverse design methodologies aim to reverse engineer solutions by specifying desired performance objectives and iteratively refining designs to meet these objectives. Within this framework, BO stands out as a data-driven approach that requires an initial set of samples to train the surrogate model, which serves as the foundation for the entire data-driven optimization process. This starting set of samples (traditionally from historical runs or randomly generated), serves as the initialization for the optimization process and its quality significantly impacts the trajectory and ultimate outcome [4]. In particular, to achieve strong performance on a specific task, the data should be accurate, representative of the problem domain, diverse and informative. Nonetheless, the impact of initialization has gained prominence as a relatively unexplored area in recent times.

Recognizing the critical role of initialization in BO for

circuit design [2], this paper investigates how different initialization strategies influence optimization performance. Through empirical analysis and experimentation, we explore the mechanisms by which initialization affects BO efficiency and effectiveness in analog circuit design. Our findings aim to offer valuable insights and practical guidelines for researchers and practitioners seeking to harness the full potential of BO in circuit design.

II. Literature Review

Over the past decade, a substantial body of research has been dedicated to advancing the automation level of analog integrated circuit (IC) design using machine learning (ML). Numerous methodologies involving neural networks (NNs) [5] [6] [7] [8], multi-objective optimization techniques [1] [9], and hybrid technologies combining global optimization with NNs [10] [11] [3] [12] [13] [14] have been developed. NNs utilizing supervised and reinforcement learning [15] [16], have gained attention for their ability to model complex problems with the advancements in high performance computing. Multi-objective optimization techniques, employing heuristic and stochastic strategies, efficiently explore the design space, while hybrid methods offer enhanced performance with lower computation cost [2].

Initial validation of NN's application dates back to 2003, with notable success in operational amplifiers (OpAmp) design [17]. Within ML approaches, supervised learning stands out, involving the compilation of input-output datasets to approximate solutions. However, the effectiveness of supervised learning hinges on the availability of extensive labeled training datasets, posing a challenge in scenarios with limited data. Recent studies have demonstrated the utility of supervised learning in designing basic circuits and modeling power consumption dynamics in complex circuits like relaxation oscillators. Prior works have addressed dataset generation techniques [5], feature selection [6], NN complexity, IC fabrication processes, circuit types [2], and validation methods [11].

For automation of IC design, a dataset of circuit simulations are required, encompassing various design parameters and performance specifications [18]. Analog designers define the performance specifications and carry out circuit simulations across a significant range of design parameters using CAD tools or SPICE circuit simulators [5], which is essentially a data acquisition process requiring significant computational cost. While more data generally enhance model accuracy, designers generate a reasonable amount to maintain practicality and manage time effectively.

In [5], a method where initial values for design parameters were determined by a circuit designer and simulated using SPICE circuit simulator is introduced. These values were then varied randomly within a range of $\pm 30\%$ to produce 100 different circuit patterns. Each circuit was simulated, and a score was calculated based on performance metrics, selecting the best circuit for the dataset. Although this approach enhanced exploration, the selection process relied on a single score, potentially overlooking valuable information. In contrast, [6] proposed two innovative approaches: first, utilizing an existing dataset from previous optimizations to ensure high-quality training data; and second, augmenting the dataset by including circuits that had lower or higher performance, thus artificially increasing dataset size.

In [8], a special software framework (MaxFit GA-SPICE) based on a genetic multi-objective optimization algorithm which relied on a set of predefined circuit equations developed by circuit designers is employed to generate the dataset. This ensured high-confidence and accurate circuit performances but relied heavily on the availability and accuracy of the software. Additionally, outlier values were filtered out from the dataset. A different approach is taken in [7], by generating a fully random dataset from a normal distribution on all design parameters, with added restrictions to ensure quality examples and enhance model accuracy, thus widening the prediction range for the target performance.

Following dataset generation, a pivotal aspect of NN training for analog IC design automation revolves around feature selection, crucial for predicting design parameters from performance specifications. Techniques such as nth-order polynomial combinations have been employed to enrich the feature space, as demonstrated in [6]. Experienced analog circuit designers play a pivotal role in this process, leveraging their deep understanding of the intricate relationships between performance metrics and design parameters to ensure accurate model convergence and mitigate operational issues.

Within the domain of hybrid IC design automation methods, novel approaches have emerged to streamline the design process. A two-step method for redesigning analog circuits in new contexts, defining a context as a set of high-level conditions influencing circuit performance is introduced in [13]. First, a multivariate polynomial regressor estimated the performance trade-offs and predicted circuit performance for new contexts, then an artificial NN predicted device sizes corresponding to the new performance. This approach leveraged existing designs efficiently, providing a swift method for sizing solutions in fresh circuit contexts.

In [14], separate NNs were used for each performance specification to reduce convergence time of an evolutionary algorithm predicting circuit design parameters, yielding rapid circuit performance estimation, replacing slow SPICE circuit simulators. Additionally, [12] adopted a time-efficient approach combining GA global optimization with NNs for local minimum searches, yielding significant speed improvements in circuits like rail-to-rail Op-amps. Lastly, [11] utilized a Bayesian neural network (BNN) within a BO framework to model multiple circuit specifications efficiently, capturing performance trade-offs and reducing computation cost across various CMOS circuits.

To build a model capable of predicting the optimal design, the initial step in supervised learning is to generate a dataset that covers the design space sufficiently. Examining the methods of starting sample generation are thus insightful. The notion of choosing initial samples for annotation, a practice referred to as initialization, has attracted growing attention in the field of active learning in recent years [4]. The choice of the initialization technique in a data centric framework, can greatly impact the overall performance and efficiency of the learning

process. However, a critical gap lies in understanding the impact of initialization techniques on optimization outcomes, particularly in achieving faster convergence with less data. Despite the advancements in automating IC design, no prior studies have explicitly addressed this aspect. We recognize the significance of initialization techniques in influencing the overall performance and efficiency of the learning process since inadequate initialization practice can lead to being stuck into sub-optimal space, resulting in slower convergence, higher annotation costs and poor model performance [2].

III. Bayesian Optimization (BO) for Analog Circuit Design

Figure 1 illustrates a BO framework for finding the optimal circuit parameters to achieve a desired performance where the impact of initial dataset generation techniques are also highlighted. Such a framework maximizes the information gained from experiments, reduces resource wastage, and accelerates the optimization process by making informed decisions based on historical data and predictive modeling.

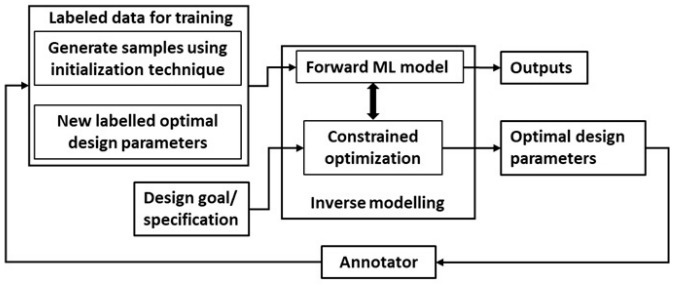


Fig. 1: Block diagram of Bayesian optimization framework with initial data generation techniques

In BO as illustrated in Fig. 1, forward machine learning (ML) models known as surrogate models are constructed to approximate the relationship between circuit parameters (inputs) and the performance parameters (outputs). In our study, the input parameters comprise primarily the gate widths of the transistors, which are varied across all circuit topologies as these are the primary tuning knob in CMOS analog circuit design for adjusting key performance metrics. In certain cases, additional input parameters such as compensation capacitance and biasing currents are also included when they have a significant impact on circuit behavior and offer meaningful tuning flexibility. This approach is grounded in both practical circuit design principles and is aligned with prior literature on analog IC design automation [19] [20]. Other parameters, such as transistor gate lengths and supply voltage were held constant to reduce simulation costs, manage design space complexity, and enhance interpretability of results. Also, parameters that are typically fixed by technology constraints (e.g. threshold voltages, oxide thickness) or not directly involved in optimization at the schematic level were excluded from the input space. Specific guidelines for analog circuit designers who wish to achieve a broader design space are also provided towards the end of Section IV-A3.

The output performance parameters considered in this study include phase margin (PM $\geq 45^o$), DC Gain $\geq$ 45 dB, gain-bandwidth product (GBW $\geq$ 20 MHz), and power consumption $\leq$ 200 µW. These are widely recognized as critical metrics in analog circuit design, as they encapsulate the fundamental trade-offs between stability, accuracy, speed, and energy efficiency that analog designers must routinely balance. These target values reflect realistic constraints commonly encountered in modern low-power, moderate-performance applications, such as wearable biomedical devices, energy-efficient front-ends, and portable sensing systems. Phase margin thresholds of 45^o and above are commonly adopted to ensure robust loop stability and prevent oscillations [21] [22]. Similarly, DC gain levels of 45 dB or more are typical to preserve signal integrity in low-voltage environments, whereas GBW, which combines gain and speed, is critical for circuits operating in the MHz range, and the target of $\geq$ 20 MHz is typical for mid-speed analog processing [23] [24]. Power consumption remains a dominant constraint in today's battery-operated and implantable systems. The chosen upper bound of 200 µW aligns with recent ultra-low-power analog front-end designs [25]. By modeling these outputs, the proposed BO framework can effectively navigate the complex design space and converge toward optimal solutions that satisfy stringent analog design constraints.

The creation of a predictive model from inputs x_C to output y_C commences with the initial dataset generation, which necessitates the comprehensive coverage of the entire design space. Traditionally, this involved manual and heuristic approaches such as historical data, expert intuition, trial and error, random sampling etc. However, such approaches may lead to suboptimal sample selection, insufficient diversity, and slow convergence, motivating the need for more systematic and data-efficient initialization strategies.

A. Initial Sample Generation

In this study, we investigate three different data-driven initialization techniques for BO: K-means, K-center greedy, Latin hypercube sampling (LHS), alongside traditional random sampling. These methods are specifically chosen for their ability to promote uniform coverage of the search space, ensuring diverse and well-distributed initial samples that can enhance the efficiency and effectiveness of the optimization process. This section provides comprehensive descriptions of each method, highlighting their advantages and limitations.

1) Random Sampling: This is utilized as a conventional baseline approach, characterized by its method of selecting samples at random to form the initial dataset. This technique is advantageous with sufficiently large datasets, as it is broadly effective across diverse scenarios, efficiently spanning the problem space without undue concentration on localized regions or excessive inclusion of potential outliers, such as points near boundaries. However, its lack of precision in navigating complex parameter spaces can lead to the selection of outlier points or redundant points clustered in local regions. This could lead to suboptimal solutions and slow down the convergence of the optimization algorithm.

2) K-Means Sampling: A data sampling method employing the K-means clustering algorithm exploits the underlying structure of all possible circuit configurations and group them with similar characteristics into K clusters. The objective of K-Means clustering is to minimize the sum of squared distances represented by:

$$J = \sum_{i=1}^{|U|} \sum_{j=1}^{K} ||x^i - c^j||^2 \tag{1}$$

where U represents all possible circuit configurations, K denotes the number of clusters, x^i signifies the ith data point in the multi-dimensional circuit configuration space, and c^j the jth cluster centroid. Equation 1 minimizes the sum of squared distances between each data point and the centroid of its assigned cluster, effectively organizing the circuit configuration space into regions of similarity. This allows K-means sampling to systematically adhere to the principle of diversity, offering a more representative selection of the circuit design space compared to random sampling. However, its effectiveness may be contingent on the chosen value of K and the data's suitability for meaningful clustering. Under circumstances where these conditions are not met, alternative sampling techniques might be more appropriate.

3) K-Center Greedy Sampling: The goal is to select K sets of points in a way that maximizes their ability to represent various regions within the data space [26]. Initially, this method incorporates the mean point of the data space as a selected point and subsequently adds points to the set that are the furthest from all currently selected points. Although this algorithm promotes diversity within the data space, its effectiveness is significantly affected by the data space's characteristics and their relation to performance. Typically, this approach tends to select samples from the boundaries of the feature space, where these boundary points are often outliers that adversely affect performance and convergence speed. Furthermore, it shares a common limitation with the K-Means initialization regarding sensitivity to the choice of K and its deterministic nature, which can make it difficult to uniformly explore different regions of the data space. Formally, we begin by including the mean point of x_C into an initially empty set of selected points S, while U corresponds to all other possible circuit configurations. This method iteratively selects $K(= N)$ points to add to S using the following criterion:

$$x^* = \arg\max_{x \in U} \min_{z \in S} |x - z| \tag{2}$$

Here, x^* represents the chosen data point. The objective of this selection process is to maximize the minimum distance between the points in U and those already in S thereby ensuring that the initial samples encompass a broad range of the data space.

4) Latin Hypercube Sampling (LHS): The prior initialization methods discussed necessitate generating all possible circuit configurations, denoted as U, which becomes computationally intensive and often impractical as the number of circuit parameters increases. In contrast, LHS segments the circuit configuration space into stratified and evenly spaced intervals across each of the D dimensions [27]. This method ensures that no more than one sample occupies each interval, achieving a uniform distribution across the entire circuit design space. LHS facilitates a structured approach to selecting initial samples, effectively addressing the need for diversity and comprehensive coverage of the design space. It achieves this by uniformly generating samples within the predefined range of $[x_C^{min}, x_C^{max})$ for the design parameters x_C.

B. Optimization

After generating the initial samples through the above discussed techniques, forward machine learning/surrogate models are trained that takes the input as the circuit parameters x_C and predicts the performance parameters y_C. Fundamentally such models should learn the complex mapping between input and output space from the data provided; which is ideal for the purpose of modeling nonlinear experimental design and decision support. We choose feed-forward NNs as our function approximator due to their capacity to capture intricate, nonlinear, and high-dimensional relationships. Additionally, NNs are differentiable, enabling us to employ efficient gradient-based optimization methods to identify an optimal design. Also, the NNs can effectively balance the exploration of the input space and the exploitation of promising regions, which contributes to efficient convergence toward the optimal solution.

In this context, we employ an ensemble of NNs as the surrogate model. To ensure consistency across the ensemble, each NN is constructed using the best configuration identified through a BO based hyperparameter search. The hyperparameters explored include the number of layers, the number of neurons per layer, the activation function (e.g., rectified linear unit (ReLU), sigmoid, or tanh), the number of training epochs, learning rate, batch size, and the optimizer (e.g., Adam, SGD, or RMSprop). This approach identifies the best configuration that enhances predictive accuracy while maintaining computational efficiency. The ensemble of NNs is then utilized to optimize the acquisition function, which guides the selection of new experimental points expected to provide the most valuable information for model refinement and optimization. In this paper, we have chosen the expected improvement (EI) as the acquisition function, as it is widely used in optimization problems. The EI acquisition function assists us in determining the next optimal point for evaluating the true objective function. EI is usually preferred over RMSE as an acquisition function since it considers both the mean and uncertainty of the surrogate model. Also, it balances the exploration-exploitation trade-off by selecting points that are both close to the current best point (exploitation) and in regions where the objective function is uncertain (exploration). It is defined as below:

$$\mathbb{E}[Improvement(x)] = \mathbb{E}\left[\max(y_t - f(x), 0)\right] \tag{3}$$

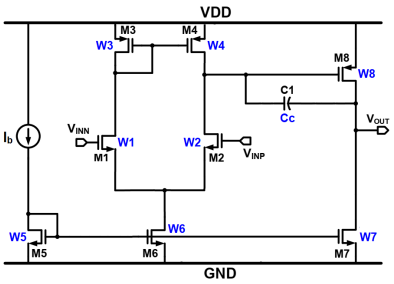
where y_t denotes the specified target, and $f(x)$ is the predicted mean of the objective function at x (starts from init guess). Equation 3 basically provides the improvement over the best observed value that would be obtained by evaluating the objective function at x. Efficient optimization of the acquisition

function selects the point that maximizes its expected improvement over the current best observation, indicating the most promising location for evaluation. A new experiment is then conducted based on this optimal point and the observed data is incorporated into the surrogate model, updating its predictions and uncertainty estimates. The acquisition function is then re-optimized to select the next best point for experimentation. This iterative process of BO is continued until all the design goals are satisfied.

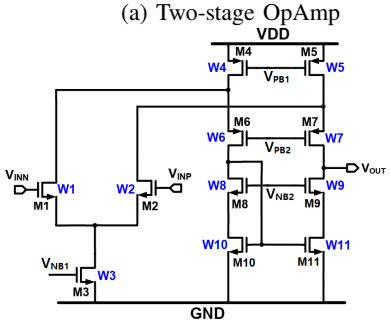
IV. Results & Discussion

In this section, we analyze the impact of the aforementioned initialization strategies on BO based on three conventional operational amplifier (OpAmp) circuit topologies of increasing complexities, as depicted in Figure 2. The selected OpAmp circuits comprise 9, 11, and 15 input design parameters, respectively, providing a progressively scalable framework for analyzing BO performance in higher-dimensional design spaces. These OpAmp topologies are widely used and well-understood in analog circuit design, making them a practical and relevant basis for our analysis. More importantly, the key reason for focusing on these circuits is to address the recognized challenge of scaling BO to high-dimensional analog design spaces [28]. Including more analog circuit types, such as RF circuits, would introduce entirely new performance metrics, design constraints, and simulation setups, which would shift the emphasis away from our focus of demonstrating that BO, when coupled with principled initialization, can be effectively scaled to more complex analog circuit optimization tasks.

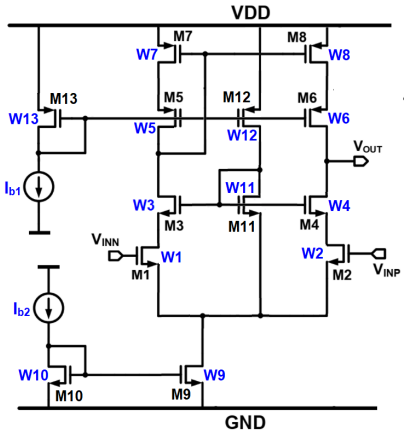
The test datasets for training the model for the BO framework are generated using Cadence Virtuoso analog design environment (ADE) explorer and assembler with Maestro simulation setup. They are obtained based on the implementation of the respective circuits in 45 nm TSMC CMOS technology.



(a) Two-stage OpAmp



(b) Folded Cascode OpAmp



(c) Telescopic Cascode

Fig. 2: Schematics of analog circuits

A. Analog Circuits

1) Two-Stage Miller-Compensated Operational Amplifier (OpAmp): Figure 2(a) depicts the schematic of a two-stage Miller-compensated OpAmp design [29] with differential inputs and single-ended output. In this example, there are 9 design parameters ($x_C \in \mathbb{R}^9$), of which 8 are dedicated to the sizing of the gate widths of individual transistors and 1 corresponds to the compensation capacitance. These design parameters, indicated in blue in Figure 2(a), can be varied within predefined ranges while adhering to predetermined step sizes as given in Table I. The step sizes are chosen to keep the design space computationally feasible while still allowing for meaningful exploration. The permutation of these design parameters results in around 27 million possible circuit configurations. The biasing conditions remain constant throughout the experiment to ensure reasonable predictability in circuit behaviour and consistency in performance analysis. As mentioned in Section III, the performance of the circuit is measured by design metrics such as phase margin (PM), DC gain, gain bandwidth product (GBW) and power, which are the targeted design specifications of optimization, i.e. $y_C \in \mathbb{R}^4$. The design objectives are respectively set as: PM $\geq 45^o$, DC Gain $\geq$ 45 dB, GBW $\geq$ 20 MHz, and power $\leq$ 200 µW, and all need to be met concurrently.

TABLE I: Range and step sizes for Two-Stage OpAmp

	W1(m)	W2(m)	W3(m)	W4(m)	W5(m)	W6(m)	W7(m)	W8(m)	Cc(F)
Min.	6μ	6μ	9μ	9μ	6μ	6μ	15μ	9μ	3p
Max.	21μ	21μ	42μ	42μ	21μ	21μ	21μ	42μ	6p
Step	3μ	3μ	3μ	3μ	3μ	3μ	3μ	3μ	1p

2) Folded Cascode Operational Amplifier (OpAmp): The schematic diagram of the second candidate utilized in our analysis, namely the folded cascode OpAmp [30] [31] is portrayed in Figure 2(b). In this circuit, we employ 11 design parameters arising from a transistor that acts as a current source with five transistor pairs with matched aspect ratios (W1=W2, W4=W5, W6=W7, W8=W9, W10=W11). Each parameter varies across a larger range of 1μm to 50μm in

steps of 0.5μm ($x_C \in \mathbb{R}^{11}$). This leads to around 900 billion possible circuit configurations.

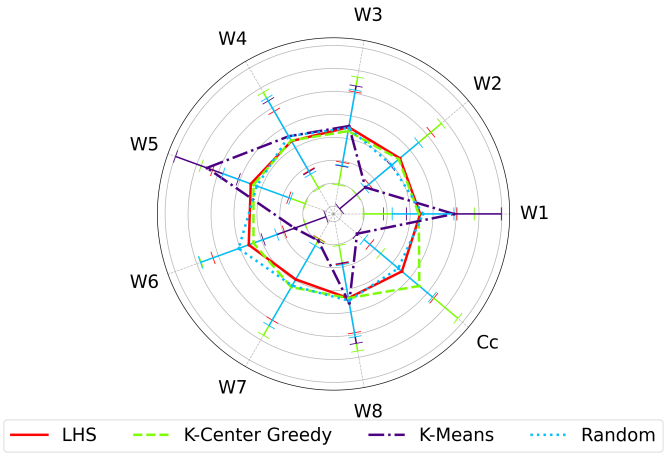
Similar to its Miller-compensated counterpart, biasing conditions are also kept unchanged over the course of the experiment. The performance metrics of PM, DC Gain, GBW, and power, are retained along with their associated target values as those of the previous circuit ($y_C \in \mathbb{R}^4$). It is evident that experiments concerning this circuit topology are significantly more complex compared to the previous circuit, primarily due to the increased number of design parameters involved.

3) Telescopic Cascode Operational Amplifier (OpAmp): The third analog circuit, namely the telescopic cascode OpAmp [32], [33] adopted as use case in our analytical study is illustrated in Fig. 2(c). As compared to the Miller-compensated and folded-cascode designs, each with 9 and 11 design parameters respectively, this circuit topology incorporates 15 design parameters, which include symmetrical widths for transistor pairs (W1=W2, W3=W4, W5=W6, W7=W8, W9=W10). Additionally, the widths of the biasing devices formed by diode-connected transistors W11 and W13, and current source device W12 and the biasing currents I_{b1} and I_{b2} are also included into the analysis to attain a wider range of configurations ($\approx 2.5 \times 10^{21}$), thereby allowing for a more comprehensive evaluation and demonstration of the proof of concept. Each width parameter varies from 1μm to 50μm in steps of 0.5μm and each bias current ranges from 1μA to 50μA in steps of 1μA, collectively representing a parameter space $x_C \in \mathbb{R}^{15}$. The performance metrics of the telescopic cascode are kept the same as its folded cascode counterpart.

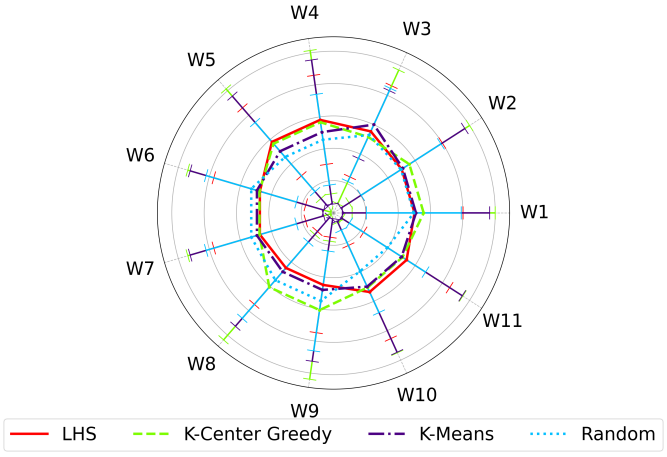
To further train the BO models in terms of its scalability and efficiency, designers can expand the design space by increasing the number of input parameters considered during optimization. For instance, in addition to varying the widths of individual transistors, they can also adjust other parameters such as biasing conditions (as in the case of the telescopic cascode OpAmp) and gate lengths. This approach broadens the design search space, enabling exploration of a wider range of configurations and potentially uncovering superior solutions that might not be evident in a more constrained design scope. Designers may also refine the optimization process by increasing the upper and lower boundaries of the variables to further enlarge the search space. Additionally, adjusting the step sizes or resolution of the variables allows for more granular exploration of the design space, facilitating finer control over trade-offs between performance metrics such as power, gain, and bandwidth. Incorporating these strategies enables a more comprehensive evaluation of the proof of concept and enhances the robustness of the final design.

B. Analysis of Initial Samples on BO

In this paper, our emphasis is on designing data-efficient circuits and thereby the initialization phase in BO framework commences with a relatively small sample set size of 16 for all the circuits. Figure 3 depicts the distribution of mean values of the initial samples for two-stage and folded cascode OpAmp circuits using the initialization techniques discussed in Section III. The vertical bars indicate the standard deviation among the initial 16 samples. It is evident that clustering approaches like K-means do not evenly explore the design space, unlike LHS, which inherently aim for more representations and uniformity among samples. The higher standard deviation in Figure 3 for K-center greedy signifies its inclination for sampling from the design space boundaries. Further, random sampling leads to non-uniform coverage of the circuit design space.



(a) Two-stage OpAmp



(b) Folded Cascode OpAmp

Fig. 3: Initial sample distribution for active learning

In order to evaluate the diversity among different initialization techniques, the average inter-sample Euclidean distance is calculated for two-Stage and folded cascode OpAmp, as shown in Table II. Among the different initialization techniques, K-center greedy sampling, which is specifically designed to maximize the distance between samples, demonstrated the highest diversity by achieving the largest inter-sample distances. LHS and random sampling followed, exhibiting relatively uniform distributions. In contrast, K-means sampling produced localized clusters, resulting in the lowest diversity characterized by shorter inter-sample distances. These findings highlight the trade-offs between diversity and clustering tendencies inherent in different sampling strategies, emphasizing the need to select an appropriate method based on the specific requirements of the design process.

The next objective is to investigate the impact of different initialization techniques on the speed of BO convergence

TABLE II: Average Euclidean distance across initializations

$d = \|x_i - x_j\|$	Random	K-Means	K-Center Greedy	LHS
Two-Stage OpAmp	3.00E-05	2.30E-05	4.40E-05	2.90E-05
Folded Cascode OpAmp	6.30E-05	5.30E-05	10.8E-05	6.10E-05

towards predefined output targets specified by the circuit designers. The BO platform utilized in this paper is based on the BoTorch framework [34] along with the commercial EDA tool Cadence in a closed-loop configuration, enabling a seamless iterative BO process. BoTorch offers a modular and flexible framework for implementing BO algorithms, leveraging ensemble of NNs (as discussed in Section III) as surrogate models to efficiently navigate design spaces. When coupled with Cadence, this closed-loop setup facilitates automated exploration and optimization of circuit parameters. Specifically, BO platform suggests candidate design points based on prior evaluations utilizing the acquisition function defined in equation 3. These points are then simulated or synthesized in Cadence to assess their performance metrics, which are subsequently fed back into the BO loop to refine the surrogate model, enabling convergence toward optimal solutions with minimal simulations.

To study the impact of initialization strategies on the optimization process, only the initial samples are varied, while the BO framework remains unchanged. Random initialization, widely regarded as the state-of-the-art baseline [35], is employed as a benchmark to assess the performance of alternative initialization methods. By varying the initial sample distribution, the closed-loop system assesses how different starting points influence convergence rates and final design quality. To quantify the influence of these initialization techniques, a metric termed - design goals achieved, defined as the average percentage of each design objective that has been achieved, is used for comparison. As an example, if the achieved design goals are: $\hat{Y}_C \in \{\text{PM} = 46^o; \text{DC Gain} = 47dB; \text{GBW} = 10MHz; \text{Power} = 185\mu W\}$, the percentage of design goal achieved is $((46 > 45) + (47 > 45) + (10e^6 > 20e^6) + (185e^{-6} < 200e^{-6}))/4 \times 100\% = 75\%$. Only when all design objectives are satisfied will the design goal achieved be 100%.

The simulation iterations required to achieve the design goals for different initializations for two-stage OpAmp is shown in Fig. 4. Notably, LHS and K-center greedy methods achieved convergence at the 12^{th} iteration, while K-means reached convergence at the 15^{th} iteration for Two-stage OpAmp circuit. However, K-center greedy approach tend to sample extreme or outlier data points while minimizing the maximum distance and might not be suitable across a range of problems when compared to LHS that ensures even coverage across the complex parameter space. On the contrary, when employing the random initialization method, not all of the predefined design objectives were achieved successfully. The input parameters corresponding to various initialization strategies that successfully achieved all the design goals are presented in Table III, while the resulting circuit performance metrics are summarized in Table IV. This further illustrates that different initializations will result in varying convergence patterns and careful consideration of initialization strategies is crucial to achieve the desired optimization results.

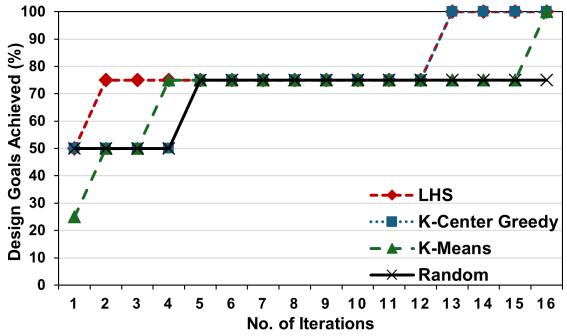


Fig. 4: Comparison of iterations for design closure of Two-stage OpAmp

The same trend is observed for folded cascode OpAmp as depicted in Fig. 5, where K-center greedy and LHS methods achieved convergence at the 5^{th} and 6^{th} iteration, respectively. While K-means and random initialization could not achieve all the predefined design objectives even after 4 additional iterations. It is also interesting to note that, although random initialization achieved 3 design objectives the fastest, it failed to converge for 4th design objective.

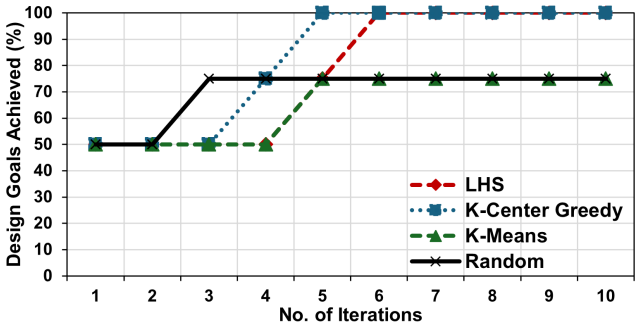


Fig. 5: Comparison of iterations for design closure of Folded Cascode OpAmp

To further demonstrate the impact of the initial training data on BO convergence, LHS identified as the best initialization method is applied to the large design space of Telescopic Cascode OpAmp. The number of iterations required to achieve design closure is compared against baseline approach, such as random initialization, as shown in Fig. 6. The figure highlights that with LHS initialization, design closure is reached in just six iterations, whereas random initialization fails to converge even after more than five times the number of iterations. This stark difference underscores the importance of selecting an effective initialization strategy, as it directly influences convergence efficiency and computational cost. This significant difference highlights the critical role of selecting an effective initialization strategy, as it directly influences both convergence efficiency and computational cost. Notably, the proposed approach requires only 21 samples in total (16 for initialization and 5 for convergence) demonstrating superior data efficiency resulting in minimizing design time and computational resources.

TABLE III: Optimal points obtained for different initialization techniques

Initialization	W1	W2	W3	W4	W5	W6	W7	W8	Cc
LHS	1.92E-05	1.58E-05	2.10E-05	1.06E-05	9.74E-06	7.26E-06	3.63E-05	9.22E-05	8.92E-13
K-Center Greedy	1.63E-05	1.13E-05	2.26E-05	1.23E-05	1.04E-05	6.37E-06	4.82E-05	8.82E-05	6.92E-13
K-Means	1.18E-05	1.59E-05	3.56E-05	1.40E-05	1.21E-05	1.25E-05	4.69E-05	9.89E-05	8.40E-13

TABLE IV: Circuit performance of optimal points obtained for different initialization techniques

Initialization	PM	DC Gain	GBW	Power
LHS	46.06	48.29	2.10E+07	1.50E-04
K-Center Greedy	46.77	47.36	2.29E+07	1.75E-04
K-Means	45.24	46.96	2.38E+07	1.66E-04

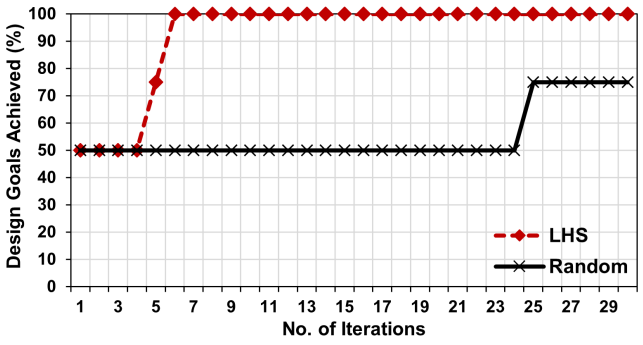


Fig. 6: Comparison of iterations for design closure of Telescopic Cascode OpAmp

In Table V, a comparison of the proposed approach with prior studies [18], [28], [36] on analog circuit design optimization using BO are performed. While all approaches target similar circuit types, such as OpAmps, they differ in terms of circuit complexity. To emphasize the impact of initialization in low to high-dimensional design spaces, this work focuses on circuits with increasing number of design parameters. Using only 16 initial samples, the proposed approach achieves convergence within 5–12 iterations, demonstrating significantly higher sample efficiency compared to existing methods. Overall, the proposed approach achieves comparable optimization performance with substantially fewer simulations, highlighting its effectiveness and practical applicability in analog design automation. Such advancements are particularly valuable in complex analog circuit designs, where accelerated convergence can substantially impact development timelines and performance verification. These findings underscore the benefits of adopting structured initialization methods like LHS to not only enhance convergence but also foster a more robust and efficient optimization framework.

V. Future Directions

The circuits under test in this research primarily focus on a specific class of circuit design, i.e. OpAmps, to investigate the impact of initialization strategies in BO-based design workflows. The selected topologies span a range of increasing design complexity, with up to 15 design parameters. This allows for an in-depth yet realistic evaluation of BO scalability in high-dimensional settings.

A key motivation for this focused approach is grounded in the fact that BO techniques tend to perform well primarily in low-dimensional design problems (typically fewer than 10 variables). Scaling BO to more complex, high-dimensional analog design spaces remains a recognized challenge in the field. By evaluating our proposed initialization strategies across circuits with 9–15 design variables, this work directly addresses that limitation. The inclusion of more analog circuit types in this study would introduce new metrics, constraints, and simulation setups, thereby shifting focus away from the core objective of rigorously evaluating BO performance under high-dimensional conditions. For these reasons, the current circuit set provides sufficient complexity and practical relevance to establish a strong methodological foundation.

Future studies can extend this work to other circuit classes, including RF and mixed-signal blocks, which may present different optimization behaviors and architectural challenges. This expansion would allow for a comprehensive understanding of how the optimization techniques perform across different architectures and applications.

Future works of this research could also entail studying the issue of potential variability in circuit performance, particularly in relation to process, voltage, and temperature (PVT) variations, by integrating variability modeling into the BO process. This can be done by generating datasets that reflect PVT variations and designing optimization objectives that account for performance metrics under varying conditions, which could involve developing and implementing methods to evaluate performance variability, such as sensitivity analyses, to quantify how fluctuations in these parameters impact circuit behavior. These enhancements would strengthen the framework's practical applicability and resilience.

VI. Conclusion

In summary, this paper has underscored the pivotal role played by initialization strategies in Bayesian optimization (BO) for analog circuit design. We have substantiated that the selection of an appropriate initialization method such as LHS, significantly enhances the efficiency and efficacy of the optimization procedure. Through an examination of diverse initialization strategies and their effects on optimization performance including considerations such as convergence speed and solution quality, we have underscored the necessity of deliberate initialization strategies, such as LHS, for fully unlocking BO's potential in analog circuit design. By elucidating the significance of initialization strategies, our study furnishes valuable guidance for circuit designers endeavoring to employ BO for the proficient and effective design of analog

TABLE V: Benchmarking AI techniques in Analog IC design

	This Work	TCAD 2022 [28]	SOCC 2022 [18]	DAC 2020 [36]
Circuit Types	Two-stage OpAmp Folded cascode OpAmp Telescopic cascode OpAmp	Two-stage OpAmp Voltage Reference (VR) Bandgap VR	Two-stage OpAmp Folded cascode OpAmp Integrated voltage regulator (IVR)	OpAmp Power Amplifier
No. of design variables	9-15	7-10	7-13	10-12
Initial Samples	16	250	65-2000	20
BO Iterations	5-12	800	800	150 (Batch)
Acquisition Function (AF)	EI	EI	RMSE	EasyBO
Simulator	Cadence	HSPICE	Cadence	HSPICE

circuits. Therefore, while this study deliberately limits its scope to a targeted yet sufficiently complex set of OpAmp designs, it addresses a core scalability challenge in BO and lays the groundwork for broader applications in analog circuit design automation. Ultimately, our research contributes to the advancement of analog circuit design by equipping practitioners with actionable insights to refine their optimization methodologies and achieve superior design outcomes.

ACKNOWLEDGEMENT

This project was supported in part by Singapore National Research Foundation under Grant NRF-CRP20-2017-0006 and in part by Agency for Science, Technology and Research (A*STAR) under its AME Programmatic Funds under Grant A19E3b0099 and Grant A20H6b0151.

REFERENCES

[1] W. Lyu, P. Xue, F. Yang, C. Yan, Z. Hong, X. Zeng, and D. Zhou, "An efficient bayesian optimization approach for automated optimization of analog circuits," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 65, no. 6, pp. 1954–1967, 2018.

[2] R. Mina, C. Jabbour, and G. E. Sakr, "A review of machine learning techniques in analog integrated circuit design automation," *Electronics*, vol. 11, no. 3, p. 435, 2022.

[3] K. Hakhamaneshi, N. Werblun, P. Abbeel, and V. Stojanović, "Bagnet: Berkeley analog generator with layout optimizer boosted with deep neural networks," in *IEEE/ACM Intl. Conf. Computer-Aided Design (ICCAD)*, 2019, pp. 1–8.

[4] Y. Yang and M. Loog, "To actively initialize active learning," *Pattern Recognition*, vol. 131, p. 108836, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S003132032200317X

[5] M. Fukuda, T. Ishii, and N. Takai, "Op-amp sizing by inference of element values using machine learning," in *Intl. Symp. Intelligent Signal Processing and Communication Systems (ISPACS)*, 2017, pp. 622–627.

[6] N. Lourenço, J. Rosa, R. Martins, H. Aidos, A. Canelas, R. Póvoa, and N. Horta, "On the exploration of promising analog ic designs via artificial neural networks," in *15th Intl. Conf. Synthesis, Modeling, Analysis and Simulation Methods and Applications to Circuit Design (SMACD)*, 2018, pp. 133–136.

[7] S. D. Murphy and K. G. McCarthy, "Automated design of cmos operational amplifier using a neural network," in *32nd Irish Signals and Systems Conference (ISSC)*, 2021, pp. 1–6.

[8] H. M.V. and B. P. Harish, "Artificial neural network model for design optimization of 2-stage op-amp," in *24th Intl. Symp. VLSI Design and Test (VDAT)*, 2020, pp. 1–5.

[9] W. Lyu, F. Yang, C. Yan, D. Zhou, and X. Zeng, "Batch Bayesian optimization via multi-objective acquisition ensemble for automated analog circuit design," in *Proceedings of the 35th International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, J. Dy and A. Krause, Eds., vol. 80. PMLR, 10–15 Jul 2018, pp. 3306–3314. [Online]. Available: https://proceedings.mlr.press/v80/lyu18a.html

[10] E. Dumesnil, F. Nabki, and M. Boukadoum, "Rf-lna circuit synthesis by genetic algorithm-specified artificial neural network," in *2014 21st IEEE International Conference on Electronics, Circuits and Systems (ICECS)*, 2014, pp. 758–761.

[11] Z. Gao, J. Tao, F. Yang, Y. Su, D. Zhou, and X. Zeng, "Efficient performance trade-off modeling for analog circuit based on bayesian neural network," in *IEEE/ACM Intl. Conf. Computer-Aided Design (ICCAD)*, 2019, pp. 1–8.

[12] Y. Li, Y. Wang, Y. Li, R. Zhou, and Z. Lin, "An artificial neural network assisted optimization system for analog design space exploration," *IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems*, vol. 39, no. 10, pp. 2640–2653, 2020.

[13] N. Lourenço, E. Afacan, R. Martins, F. Passos, A. Canelas, R. Póvoa, N. Horta, and G. Dundar, "Using polynomial regression and artificial neural networks for reusable analog ic sizing," in *16th Intl. Conf. Synthesis, Modeling, Analysis and Simulation Methods and Applications to Circuit Design (SMACD)*, 2019, pp. 13–16.

[14] G. İslamoğlu, T. O. Çakici, E. Afacan, and G. Dündar, "Artificial neural network assisted analog ic sizing tool," in *16th Intl. Conf. Synthesis, Modeling, Analysis and Simulation Methods and Applications to Circuit Design (SMACD)*, 2019, pp. 9–12.

[15] K. Settaluri, A. Haj-Ali, Q. Huang, K. Hakhamaneshi, and B. Nikolic, "Autockt: deep reinforcement learning of analog circuit designs," in *Proc. 23rd Conf. Design, Automation and Test in Europe*, 2020, p. 490–495.

[16] H. Wang, J. Yang, H.-S. Lee, and S. Han, "Learning to design circuits," *ArXiv*, vol. abs/1812.02734, 2018.

[17] G. Wolfe and R. Vemuri, "Extraction and use of neural network models in automated synthesis of operational amplifiers," *IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems*, vol. 22, no. 2, pp. 198–212, 2003.

[18] R. Dutta, A. James, S. Raju, Y.-J. Jeon, C. S. Foo, and K. Tshun Chuan Chai, "Automated deep learning platform for accelerated analog circuit design," in *IEEE 35th Intl. System-on-Chip Conf. (SOCC)*, 2022, pp. 1–5.

[19] W. Lyu, P. Xue, F. Yang, C. Yan, Z. Hong, X. Zeng, and D. Zhou, "An efficient bayesian optimization approach for automated optimization of analog circuits," *IEEE Trans. Circuits Syst. I Regular Paper*, vol. 65-I, no. 6, pp. 1954–1967, 2018.

[20] K. Settaluri, A. Haj-Ali, Q. Huang, K. Hakhamaneshi, and B. Nikolic, "AutoCkt: Deep reinforcement learning of analog circuit designs," in *Proceedings of the Design, Automation & Test in Europe Conference & Exhibition (DATE)*. IEEE, 2020, pp. 490–495.

[21] W. Choi, H. Lee, J. Kim, and Y. Chae, "A 70.2-db sndr 20-mhz bandwidth continuous-time $\Delta\Sigma$ modulator using a scalable amplifier topology," *IEEE Journal of Solid-State Circuits*, vol. 55, no. 4, pp. 962–973, 2020.

[22] C.-H. Liao, H.-C. Tsai, C.-H. Yang, and S.-I. Liu, "A 10-ms/s 12-bit energy-efficient sar adc with a dynamic op-amp-based reference buffer," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 69, no. 2, pp. 781–791, 2022.

[23] R. Islam and A. M. A. Ali, "A 16.4-nw 1.2-v 0.16-mm² two-stage class-ab op-amp with adaptive biasing and slew enhancement," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 69, no. 5, pp. 2098–2110, 2022.

[24] K.-M. Choi, H.-J. Yoo, and Y. Chae, "A 1.2-v 12-bit 10-ms/s sar adc using a two-stage dynamic comparator with a gain-boosted opamp," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 66, no. 3, pp. 976–986, 2019.

[25] A. Cevrero, F. Iannuzzo, A. Scuderi, and L. Fanucci, "An ultra-low-power and compact digitally-assisted analog front-end for implantable ecg acquisition," *IEEE Journal of Solid-State Circuits*, vol. 57, no. 10, pp. 3023–3032, 2022.

[26] O. Sener and S. Savarese, "A geometric approach to active learning for convolutional neural networks," *ArXiv*, vol. abs/1708.00489, 2017.

[27] M. Mckay, R. Beckman, and W. Conover, "A comparison of three methods for selecting vales of input variables in the analysis of output from a computer code," *Technometrics*, vol. 21, pp. 239–245, 05 1979.

[28] C. Chen, H. Wang, X. Song, F. Liang, K. Wu, and T. Tao, "High-dimensional bayesian optimization for analog integrated circuit sizing based on dropout and gm/id methodology," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 41, no. 11, pp. 4808–4820, 2022.

[29] G. Palmisano and G. Palumbo, "A compensation strategy for two-stage cmos op-amps based on current buffer," in *IEEE Trans. Circuits and Systems (part I)*, vol. 44, no. 3, 1997, p. 257–262.

[30] S. Mallya and J. Nevin, "Design procedures for a fully differential folded-cascode cmos operational amplifier," *IEEE Journal of Solid-State Circuits*, vol. 24, no. 6, pp. 1737–1740, 1989.

[31] D. B. Ribner, M. A. Copeland, and M. Milkovic, "80MHz low offset fully-differential and single-ended opamps," in *IEEE Custom Integrated Circuits Conf.*, 1983, pp. 74–75.

[32] M. Degrauwe, O. Nys, E. Dijkstra, J. Rijmenants, S. Bitz, B. Goffart, E. Vittoz, S. Cserveny, C. Meixenberger, G. van der Stappen, and H. Oguey, "Idac: An interactive design tool for analog cmos circuits," *IEEE Journal of Solid-State Circuits*, vol. 22, no. 6, pp. 1106–1116, 1987.

[33] K. Gulati and H.-S. Lee, "A high-swing cmos telescopic operational amplifier," *IEEE Journal of Solid-State Circuits*, vol. 33, no. 12, pp. 2010–2019, 1998.

[34] M. Balandat, B. Karrer, D. R. Jiang, S. Daulton, B. Letham, A. G. Wilson, and E. Bakshy, "BoTorch: A Framework for Efficient Monte-Carlo Bayesian Optimization," in *Advances in Neural Information Processing Systems 33*, 2020. [Online]. Available: http://arxiv.org/abs/1910.06403

[35] S. A. YaghoubNejad and M. T. Manzuri, "Exploring efficient and simple initialization strategies for bayesian optimization with SETUP-BO," 2023. [Online]. Available: https://openreview.net/forum?id=TFrzVBZk05g

[36] S. Zhang, F. Yang, D. Zhou, and X. Zeng, "An efficient asynchronous batch bayesian optimization approach for analog circuit synthesis," in *2020 57th ACM/IEEE Design Automation Conference (DAC)*, 2020, pp. 1–6.