

RASNIL: PVT-Robust Many-Objective Analog Sizing via Nested Hybrid Fidelity Framework with Incremental Learning

Xingyu Tang, Sen Yin, Zhujun Yao, Bingzhang Huang, Xiaosen Liu, and Yan Wang
School of Integrated Circuits, Tsinghua University, Beijing, China
wangy46@tsinghua.edu.cn

Abstract—Yield-driven analog circuit design under process, supply voltage, and operating temperature (PVT) variations remains a major challenge, particularly as technology advances and design goals diversify. Traditional yield analysis relies on time-consuming Monte Carlo simulations, while PVT-aware sizing often depends on hybrid fidelity model-based methods that suffer from slow training and limited efficiency. We propose an efficient algorithm to overcome these limitations. First, Monte Carlo simulations are replaced by a sensitivity-based fast yield estimation technique. Second, a hybrid fidelity Kriging model based on incremental learning with a self-adaptive training strategy greatly reduces training costs. Third, an efficient nested optimization framework incorporates prescreening to lower prediction time and a nested-selection mechanism based on Nondomination Rank and dynamic weighted Local Outlier Factor to enhance convergence and diversity. Finally, a general many-objective optimization strategy enables effective trade-offs among four or more design goals. Experiments on two real-world analog circuits show that our algorithm reduces model training time by up to 99%, achieves 4.68× acceleration in total runtime, improves Hypervolume by up to 430%, and consistently produces high-yield designs (>97%), outperforming state-of-the-art approaches in yield-driven, PVT-aware, many-objective sizing problems.

Index Terms—Yield-driven, many-objective optimization, hybrid fidelity model, incremental learning, analog circuit.

I. INTRODUCTION

For the past few decades, the increasing demands of high-performance, low-power, and low-area circuits have made the manual analog Integrated Circuit (IC) design much more difficult. Automated analog circuit sizing has attracted more research interest [1]. With the scaling down of IC technology, process variation combined with the fluctuations in supply voltage and operating temperature (PVT) deviates circuit performance from expected, increasing the threat of functional failures [2]. Therefore, PVT-aware analog sizing has become an interesting and urgent field of research.

Analog circuit sizing methodologies can be broadly divided into two categories, depending on the optimization objective. Single-metric optimization focuses on a single figure of merit, yielding the best design solution [3], [4]. In contrast, multiobjective optimization simultaneously considers multiple performance metrics and generates multiple Pareto fronts and Pareto solutions [5], [6]. This empowers designers to select the design that most closely aligns with their specific application requirements. In this work, we focus on the latter.

In real-world analog circuit design, there exist many objectives (≥ 4), rather than merely multiobjective (≤ 3) [7]. Most current sizing strategies revolve around the nondominated sorting genetic algorithm (NSGA-II) [8]. However, as the number of objectives grows, NSGA-II identifies an increasing number of Pareto solutions, making it difficult to distinguish truly promising solutions and thus degrading solution quality, i.e., the Dominance Resistance (DR) phenomenon [7]. These shortcomings highlight the urgent need for sizing techniques that remain effective in genuinely many-objective scenarios.

Existing PVT-aware analog sizing methods trade fidelity for efficiency. Ref. [6] conducts High-Fidelity (HF) evaluation for each candidate design, which conducts simulation on all PVT corners and uses the worst-case performance as the design metric, but this exhaustive evaluation is prohibitively time-consuming. To alleviate this, [9] and [10] employ Bayesian optimization with hybrid fidelity Kriging surrogate models: candidates satisfying a variance-based criterion trigger expensive HF evaluations, while others are assessed using nominal Low-Fidelity (LF) runs, greatly reducing runtime. However, two major issues persist. First, model training can dominate total optimization time [6], and hybrid frameworks involve training different fidelity models, multiplying the time costs. Second, the optimizer in [9] is highly sensitive to the hyperparameter γ , requiring problem-specific tuning [11], and although [10] is less sensitive to hyperparameters, its Monte Carlo (MC)-based HF candidate decision rule incurs extra Kriging model prediction time, resulting in higher computational overhead.

Most optimization algorithms still use discrete process corners (e.g., TT, FF) to characterize process variation and design for worst-case performance across all corners. This approach is often overly conservative [12]. Conversely, designing for optimal performance often fails to guarantee optimal yield [13]. To effectively balance performance and yield, yield-driven analog circuit sizing becomes essential. However, traditional yield analysis techniques such as MC simulations and others remain computationally expensive and are thus impractical for fast evaluation [14]–[16]. Consequently, developing a fast and efficient yield-driven, PVT-aware analog sizing method remains a critical challenge.

In this work, we propose a PVT-Robust many-objective Analog Sizing algorithm via Nested hybrid fidelity framework

with Incremental Learning (RASNIL). The primary contributions are listed as follows.

1) We introduce a sensitivity-based fast evaluation method that rapidly extracts analog circuit yield along with various performance metrics without requiring MC simulations, replacing traditional worst-case process corner approaches and steering optimization toward high-yield designs, accelerating yield-driven, PVT-aware analog sizing.

2) We develop a hybrid fidelity Kriging surrogate model based on incremental learning with a self-adaptive training strategy, dramatically reducing time costs of model training.

3) We propose an efficient nested hybrid fidelity Bayesian optimization framework that reduces model prediction time via a prescreening technique and promotes both convergence and diversity by employing a nested-selection strategy based on Nondomination Rank (NR) and dynamic weighted Local Outlier Factor (LOF).

4) We employ a general many-objective optimization strategy that consistently identifies a small set of promising candidate designs across varying numbers of objectives, enabling effective trade-offs among various design goals (≥ 4) while avoiding the DR phenomenon.

The rest of the paper is organized as follows. We provide the background information in Section II. We detail the proposed RASNIL method in Section III. The experimental results are illustrated in Section IV. Finally, we conclude in Section V.

II. BACKGROUND

In this section, we introduce the definition of multiobjective optimization, yield analysis, and Kriging model.

A. PVT-Aware Multiobjective Optimization

Without loss of generality, we assume all objectives are cast as minimization problems—an assumption that likewise underpins the ensuing theoretical developments. Consequently, the analog sizing under PVT variations can be posed as a robust constrained multiobjective optimization problem:

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^d} & \left(\max_{s \in \mathcal{S}_{\text{PVT}}} f_1(\mathbf{x}, s), \dots, \max_{s \in \mathcal{S}_{\text{PVT}}} f_m(\mathbf{x}, s) \right) \\ \text{s.t.} & \max_{s \in \mathcal{S}_{\text{PVT}}} g_i(\mathbf{x}, s) \leq 0, \quad i = 1, 2, \dots, k. \end{aligned} \quad (1)$$

Here, $\mathbf{x} \in \mathbb{R}^d$ is the vector of d design variables, $f_i(\mathbf{x}, s)$ is the i -th objective evaluated under PVT corner s , $g_i(\mathbf{x}, s)$ is the i -th constraint evaluated under PVT corner s , \mathcal{S}_{PVT} denotes the set of all considered PVT corners, m and k are the numbers of objectives and constraints, respectively.

Consider two candidate solutions $\mathbf{x}_1, \mathbf{x}_2$. We say \mathbf{x}_1 dominates \mathbf{x}_2 (denoted $\mathbf{x}_1 \prec \mathbf{x}_2$) if $\forall i \in \{1, \dots, m\}, f_i(\mathbf{x}_1) \leq f_i(\mathbf{x}_2)$ and $\exists i \in \{1, \dots, m\}, f_i(\mathbf{x}_1) < f_i(\mathbf{x}_2)$. Multiobjective optimization aims to find a uniform set of non-dominated solutions, i.e., Pareto solutions, instead of a single best design. The constrained-domination [8] is employed to handle constraints.

B. Yield Analysis

In the statistical device model provided by the foundry, process variation are usually expressed as some process parameters \mathbf{p} that satisfy particular distributions, such as the

threshold voltage V_{th} , the saturation velocity v_{sat} , the gate oxide thickness T_{ox} , etc. Yield Y is the probability that the objective functions corresponding to circuit performances $\phi(\mathbf{p})$ belongs to the set of successful events \mathcal{Z} .

In the brute-force MC method, the yield is approximated as the proportion of samples that meet the conditions to the total samples N_{MC} , because all samples $\mathbf{p}_i, i = 1, 2, \dots, N_{\text{MC}}$ are generated according to the same distribution.

$$\hat{Y} = \frac{1}{N_{\text{MC}}} \sum_{i=1}^{N_{\text{MC}}} \mathbf{I}(\phi(\mathbf{p}_i) \in \mathcal{Z}), \quad (2)$$

where \mathbf{I} is the indicator function.

$$\mathbf{I}(\phi(\mathbf{p}) \in \mathcal{Z}) = \begin{cases} 1 & \phi(\mathbf{p}) \in \mathcal{Z} \\ 0 & \phi(\mathbf{p}) \notin \mathcal{Z}. \end{cases} \quad (3)$$

C. Kriging Model

The Kriging model treats the response at any design point as a Gaussian random variable, which is fully characterized by the mean function and correlation function. A widely adopted choice for the latter is the Gaussian correlation kernel:

$$R(\mathbf{x}, \mathbf{x}') = \exp\left(-\sum_{i=1}^d \theta_i (x_i - x'_i)^2\right). \quad (4)$$

Here, $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_d)$ is the vector of hyperparameters and \mathbf{x}, \mathbf{x}' are two distinct points.

The goal of training the Kriging model is to find the best hyperparameters $\hat{\boldsymbol{\theta}}$ for the observed data, and consequently building \mathbf{R} , which is the $n \times n$ correlation matrix whose (i, j) -th element is $R(\mathbf{x}^i, \mathbf{x}^j; \boldsymbol{\theta})$. Given n sample points $\{\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^n\}$ and their corresponding objective values $\mathbf{y} = \{y^1, y^2, \dots, y^n\}$, the hyperparameters are estimated by maximizing the likelihood of the n observations. A commonly used equivalent formulation is as follows [17]:

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta} \in \mathbb{R}^d}{\text{argmax}} \left(-\frac{n}{2} \ln \hat{\sigma}^2 - \ln |\mathbf{R}| \right), \quad (5)$$

$$\hat{\boldsymbol{\mu}} = \frac{\mathbf{1}^\top \mathbf{R}^{-1} \mathbf{y}}{\mathbf{1}^\top \mathbf{R}^{-1} \mathbf{1}}, \quad (6)$$

$$\hat{\sigma}^2 = \frac{(\mathbf{y} - \mathbf{1}\hat{\boldsymbol{\mu}})^\top \mathbf{R}^{-1} (\mathbf{y} - \mathbf{1}\hat{\boldsymbol{\mu}})}{n}. \quad (7)$$

Here, $\mathbf{1}$ is an n -dimensional column vector of ones. By leveraging a trained Kriging model, we maximize an acquisition function at each iteration to select promising designs for simulation, augment the dataset with the new observations, and repeat the cycle until convergence on the optimal solution—an approach known as Bayesian optimization [18].

III. PROPOSED ALGORITHM

In this section, we elaborate RASNIL algorithm in detail.

A. Yield-Driven PVT-Aware Analog Circuit Sizing

In PVT-aware analog sizing, designing for optimal performance does not guarantee optimal yield. To ensure circuit robustness, yield-driven analog circuit sizing is critical. However, classical yield analysis methods require thousands of transistor-level simulations [14]–[16]. To enable a rapid optimization flow, we perform a rapid yield evaluation using a

sensitivity-based technique [19], [20]. From a single sensitivity analysis of design \mathbf{x} , we can extract all key performance metrics and their gradients with respect to parameters \mathbf{p} , then build linear surrogate model for each performance metric:

$$\phi(\mathbf{p}) = \phi_0 + \mathbf{J}_0^T \mathbf{p}. \quad (8)$$

Here, ϕ_0 and \mathbf{J}_0 are the nominal performance and its gradient at $\mathbf{p}_0 = [0, 0, \dots, 0]$. Performances at \mathbf{p} are obtained directly from model predictions instead of costly MC simulations, and then the yield are estimated based on (2) with a relatively high accuracy [20]. Thus, a single sensitivity analysis provides all performance metrics and yield without extra simulations. As shown in the upper-left region of Fig. 1, the corner set-based strategy is applied to voltage and temperature variations, while yield under process variation is estimated at each corner using surrogate models.

Another critical issue is the strong trade-off between an analog circuit's various performance metrics and its yield. Optimizing all of them simultaneously as objectives leads to suboptimal overall performances [21]. Therefore, we treat yield as a constraint in analog sizing.

B. Incremental Hybrid Fidelity Kriging Model

In general, HF evaluations are prohibitively expensive, while LF evaluations are cheaper but less accurate. Hybrid fidelity model leverages multi-fidelity datasets to exploit the correlations between these two levels to improve overall efficiency. Assume model ϕ_h and ϕ_l represent the approximation of HF and LF functions, respectively. For simplicity, based on the trained model ϕ_l , ϕ_h can be expressed as follows [9]:

$$\phi_h(\mathbf{x}) = \rho \phi_l(\mathbf{x}) + \delta(\mathbf{x}). \quad (9)$$

Here, ρ is a scale factor and $\delta(\mathbf{x})$ is a residual Kriging model whose hyperparameters must be learned from HF dataset.

However, the hybrid fidelity model introduces multiple training costs. Assume that a Kriging model has been trained with n samples and that one additional sample becomes available. Because the goal of updating model is obtaining the new \mathbf{R}^{-1} , we employ incremental learning technique to update the inverse directly [22], re-using \mathbf{R}^{-1} from the previous model and thus avoiding a full retraining.

Let the enlarged correlation matrix after inserting the $(n+1)^{\text{st}}$ sample be $\tilde{\mathbf{R}}$. The original matrix \mathbf{R} forms the upper-left block of $\tilde{\mathbf{R}}$, which can be partitioned as

$$\tilde{\mathbf{R}} = \begin{bmatrix} R_{11} & \dots & R_{1n} & R_{1(n+1)} \\ R_{21} & \dots & R_{2n} & R_{2(n+1)} \\ \vdots & & \vdots & \vdots \\ R_{n1} & \dots & R_{nn} & R_{n(n+1)} \\ R_{(n+1)1} & \dots & R_{(n+1)n} & R_{(n+1)(n+1)} \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{A} \\ \mathbf{A}^\top & \mathbf{B} \end{bmatrix}, \quad (10)$$

where \mathbf{A} and \mathbf{B} are the block matrices of $\tilde{\mathbf{R}}$. Applying the block matrix inversion formula yields

$$\tilde{\mathbf{R}}^{-1} = \begin{bmatrix} \mathbf{R}^{-1} + \mathbf{R}^{-1} \mathbf{A} \mathbf{C}^{-1} \mathbf{A}^\top \mathbf{R}^{-1} & -\mathbf{R}^{-1} \mathbf{A} \mathbf{C}^{-1} \\ -\mathbf{C}^{-1} \mathbf{A}^\top \mathbf{R}^{-1} & \mathbf{C}^{-1} \end{bmatrix}. \quad (11)$$

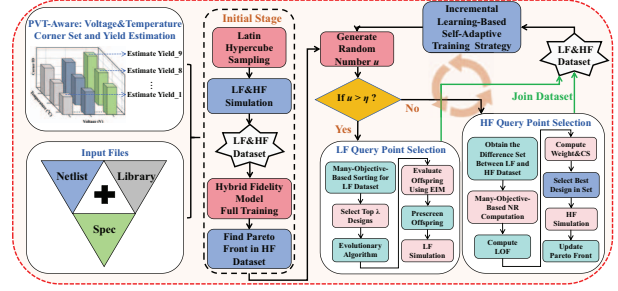


Fig. 1. Framework of RASNIL.

Here, $\mathbf{C} = \mathbf{B} - \mathbf{A}^\top \mathbf{R}^{-1} \mathbf{A}$. Consequently, $\tilde{\mathbf{R}}^{-1}$ can be derived directly via matrix operations instead of retraining, markedly reducing computational costs.

With the incremental learning strategy in place, the hybrid fidelity model no longer has to be retrained from scratch whenever new data arrives. When a HF sample is added, we incrementally update the residual term $\delta(\mathbf{x})$, while holding both ρ and ϕ_l constant; if the new sample is LF-type, we instead update the LF component ϕ_l , while fixing both ρ and $\delta(\mathbf{x})$ constant. In addition, we adopt a self-adaptive strategy to reduce computational costs: the hybrid fidelity model is fully retrained every H simulations; otherwise, it is updated incrementally. The value of H is adaptively adjusted according to H_{min} and H_{max} , following the method in [6].

C. Efficient Hybrid Fidelity Optimization Framework

In hybrid fidelity optimization, both the fidelity level and the candidate solution should be determined in each iteration. Traditional methods determine fidelity level via a variance-based criterion [9], [10], which suffers from several drawbacks as demonstrated in Section I. Moreover, they rely on surrogate models to estimate variance—forcing these models to be highly accurate (and thus costly to train)—and risk misclassification whenever their predictions fall short of precision. To reduce this dependence, Ref. [11] introduces a randomized scheme: draw $u \sim \mathcal{U}[0, 1]$; if $u \leq \eta$, evaluate at HF-level, otherwise at LF-level. Although η is a hyperparameter, it is problem-agnostic and requires no per-problem tuning. However, the candidate determination strategy in [11] still relies on numerous model predictions and computes Hypervolume (HV) at each iteration, incurring high time costs, especially in many-objective scenarios [6], [23]. Therefore, we adopt the randomized scheme but replace the candidate determination procedure with a more efficient alternative.

For LF candidate solution determination, assuming λ is the population size, LF dataset is first ranked by the many-objective-based strategy that will be elaborated in Section III-D; the top λ designs become parents and produce λ offspring by employing differential evolution and polynomial mutation algorithms [24]. Offspring are then prescreened with the Expected Improvement Matrix (EIM) acquisition function [25] to select the best candidates. In contrast to traditional internal optimization for maximizing the acquisition function, which solves a sub-optimization problem and therefore requires numerous model predictions, and to the strategy in [11],

Algorithm 1 RASNIL Algorithm

Require: Max number of iterations N_{max} , threshold parameter η , population size λ ;

- 1: Initialize datasets D_h, D_l via Latin Hypercube Sampling;
- 2: Build hybrid fidelity models for objectives and constraints;
- 3: Compute initial Pareto Front from D_h ;
- 4: **for** $i = 1$ to N_{max} **do**
- 5: Draw a random number $u \sim \mathcal{U}[0, 1]$;
- 6: **if** $u > \eta$ **then**
- 7: Rank D_l by many-objective-based strategy; select top λ designs;
- 8: Generate λ offspring via Differential Evolution and Polynomial Mutation algorithm; prescreen best via EIM $\rightarrow \mathbf{x}_l^{new}$;
- 9: Conduct LF evaluation on \mathbf{x}_l^{new} ; update D_l ;
- 10: **else**
- 11: Let $DS = D_l \setminus D_h$; compute NR, LOF on DS ;
- 12: Compute weight; compute CS on DS ;
- 13: Select $\mathbf{x}_h^{new} = \arg \max_{\mathbf{x} \in DS} CS(\mathbf{x})$;
- 14: Conduct HF evaluation on \mathbf{x}_h^{new} ; update D_h and Pareto Front;
- 15: **end if**
- 16: **if** self-adaptive criterion met **then**
- 17: Fully retrain hybrid fidelity models;
- 18: **else**
- 19: Incrementally update hybrid fidelity models;
- 20: **end if**
- 21: **end for**
- 22: **return** Final Pareto Front;

the prescreening technique requires only λ model predictions in each iteration, reducing time costs.

For HF candidate solution determination, we employ the nested-selection strategy [11]: each new candidate is chosen from the set difference between the LF and HF datasets, leveraging existing data. To rank the designs in this difference set DS , we compute a Composite Score (CS) that merges NR [8] with a dynamic weighted LOF [26]. Assume \mathbf{x} is a design in difference set, the t -objects in HF dataset closest to \mathbf{x} are called t -nearest neighbors of \mathbf{x} and are denoted as $N_t(\mathbf{x})$. LOF is an outlier detection method based on local density:

$$\text{LOF}(\mathbf{x}) = \frac{1}{t} \sum_{\mathbf{x}' \in N_t(\mathbf{x})} \frac{\rho_t(\mathbf{x}')}{\rho_t(\mathbf{x})}. \quad (12)$$

Here, $\rho_t(\mathbf{x})$ is the local reachability density of \mathbf{x} [26]. From (12), $\text{LOF}(\mathbf{x})$ quantifies the outlierness of \mathbf{x} by comparing its local density $\rho_t(\mathbf{x})$ with the average density of its t -nearest neighbors $N_t(\mathbf{x})$. If \mathbf{x} is isolated from $N_t(\mathbf{x})$, then $\text{LOF}(\mathbf{x}) \gg 1$; if it lies within the neighbor cluster, then $\text{LOF}(\mathbf{x}) \approx 1$. With the help of LOF, we can find outliers and enhance the diversity of HF dataset. Based on LOF, the CS of candidate design \mathbf{x} can be expressed as follows:

$$\text{CS}(\mathbf{x}) = \left(\frac{1}{\text{NR}(\mathbf{x})}\right)' + w(\text{LOF}(\mathbf{x}))'. \quad (13)$$

This type of score simultaneously captures convergence (via NR) and diversity (via LOF) without requiring model prediction or HV evaluation. We apply min–max normalization [27] to both NR and LOF to remove any order-of-magnitude effects, resulting in $(1/\text{NR})'$ and LOF' . Moreover, to prevent a large LOF weight from hindering convergence in the late optimization phase, we employ a dynamic weighted strategy:

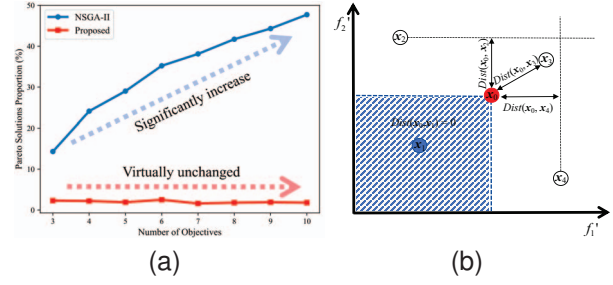


Fig. 2. General many-objective optimization strategy. (a) The proportion of Pareto solutions selected from 100 candidate solutions under different numbers of objectives. (b) Illustration of $Dist$. The blue rectangle indicates that all solutions in this region dominate \mathbf{x}_0 .

the weight w is gradually reduced every several simulations. For simplicity, we align this update schedule with that in Section III-B, i.e., we decrease w after every H simulations.

D. General Many-Objective Optimization Strategy

As shown in [7], most candidate solutions become mutually nondominated in many-objective scenarios. Using the DTLZ7 [28] benchmark as the objectives, the blue solid line in Fig. 2(a) illustrates the proportion of Pareto solutions selected from 100 candidates by the NSGA-II-based approach under varying numbers of objectives. As the objective count increases, the classical method identifies an increasing number of Pareto solutions, making it difficult to distinguish truly promising designs and thus degrading solution quality, i.e., the DR phenomenon. To address this, we transform the many-objective into multiobjective by employing two indicators [29].

1) *Convergence Indicator*: This indicator is used to estimate the convergence of the solution on all objectives. We first apply min–max normalization to each objective, then sum the resulting normalized values to quantify convergence.

$$I_c(\mathbf{x}) = \sum_{i=1}^m f'_i(\mathbf{x}). \quad (14)$$

Here, f'_i represents the i -th objective after normalization. A smaller $I_c(\mathbf{x})$ corresponds to better overall performances.

2) *Diversity Indicator*: The diversity indicator is used to assess the degree of difference between different solutions:

$$I_d(\mathbf{x}) = \frac{1}{\min_{\mathbf{x}' \in U, \mathbf{x}' \neq \mathbf{x}} \text{Dist}(\mathbf{x}, \mathbf{x}') + 1.0}, \quad (15)$$

$$\text{Dist}(\mathbf{x}, \mathbf{x}') = \sqrt{\sum_{i=1}^m (\min(0, f'_i(\mathbf{x}) - f'_i(\mathbf{x}')))^2}. \quad (16)$$

Here, U is the set consisting of all candidate solutions. Assume solution \mathbf{x}_0 (red circle in Fig. 2(b)) represents \mathbf{x} . If any \mathbf{x}' such as \mathbf{x}_1 falls in the dominating (blue) region of the solution \mathbf{x} , then the $\text{Dist}(\mathbf{x}, \mathbf{x}')$ equals to zero and $I_d(\mathbf{x})$ reaches the maximum value, i.e., 1, indicating that \mathbf{x} is non-preferable. Otherwise, $\text{Dist}(\mathbf{x}, \mathbf{x}')$ is computed from the differences between the normalized objectives. Thus, a smaller $I_d(\mathbf{x})$ both preserves solution diversity and steers the population toward promising regions [29].

Fig. 2(a) further shows the comparison between NSGA-II and the proposed general approach. The results indicate

that the former struggles to discriminate among candidates in high-dimensional objective spaces, whereas the general approach consistently identifies a small set of truly promising solutions regardless of the number of objectives, demonstrating its effectiveness for many-objective scenarios.

We summarize the RASNIL in Fig. 1 and Algorithm 1.

IV. EXPERIMENTAL RESULTS

In this section, we compare RASNIL with several algorithms: (1) NSGA-II [8], (2) MFBO [9], (3) FUZYE [16], and (4) SILE [6]. The results are summarized in Table I. NSGA-II, MFBO, and SILE omit yield/MC in their original forms (marked N/A). FUZYE is a yield-driven and PVT-aware analog sizing method, but each iteration runs multiple full MC simulations, making it slow. In contrast, RASNIL avoids MC simulations via surrogate-based yield estimation for fast sizing.

To further evaluate solution quality and optimization efficiency, we test RASNIL on two real-world analog circuits against three adapted baselines:

- 1) NSGA-II* [8]: a classical simulation-based multiobjective optimizer enhanced with RASNIL's rapid yield analysis.
- 2) MFBO* [9]: a hybrid fidelity Bayesian optimization framework, modified with EIM-based acquisition function for multiobjective tasks and integrated with rapid yield analysis.
- 3) SILE* [6]: a single-fidelity Bayesian optimization framework with acquisition and self-adaptive strategies similar to RASNIL, also integrated with rapid yield analysis.

With this integration, all baseline methods become yield-driven. For HF evaluation, performance and yield are assessed under nine corners (supply voltages and temperatures), requiring nine simulations per HF evaluation, i.e., nine sensitivity analyses, while LF evaluation requires only one simulation. Sensitivity analysis is conducted using the Xyce simulator [30]. All experiments run on a Linux system with an Intel Core i7 CPU and 8 GB RAM.

For RASNIL, we begin with 20 HF and 80 LF evaluations (260 total simulations), set $\lambda = 20$ and $t = 10$, and use a self-adaptive strategy with $H_{min} = 20$, $H_{max} = 100$. The weight of LOF starts at 1 and decays by 10% each time the self-adaptive criterion is met. To conserve HF evaluations before feasibility is achieved, we use $\eta = 0.2$; once a feasible solution is found, η increases to 1/3 to more thoroughly explore diverse HF solutions. RASNIL and MFBO* are limited to 1000 simulations, NSGA-II* to 5000, while SILE*-1k and SILE*-3k denote limits of 1000 and 3000, respectively. All Kriging-based models in RASNIL, SILE*-1k, SILE*-3k, and MFBO* are trained for 50 iterations. In this work, we demonstrate HV and success rate to quantify solution quality. For cases where no feasible solution is found, $HV = 0$.

A. Two-Stage Operational Amplifier

This example is a Two-Stage Operational Amplifier (OPA) implemented in a 180 nm library at 1.8 V supply voltage as shown in Fig. 3(a). There are 11 design variables, including the lengths and widths of transistors, resistance, and capacitance. We choose the V_{th} of each MOSFET as the process parameter,

TABLE I
COMPARISON OF ALGORITHMS

Methods	NSGA-II	MFBO	FUZYE	SILE	RASNIL
PVT-Aware Optimization	Yes	Yes	Yes	Yes	Yes
Yield-Driven Optimization	No	No	Yes	No	Yes
MC Simulation Requirement	N/A	N/A	Yes	N/A	No

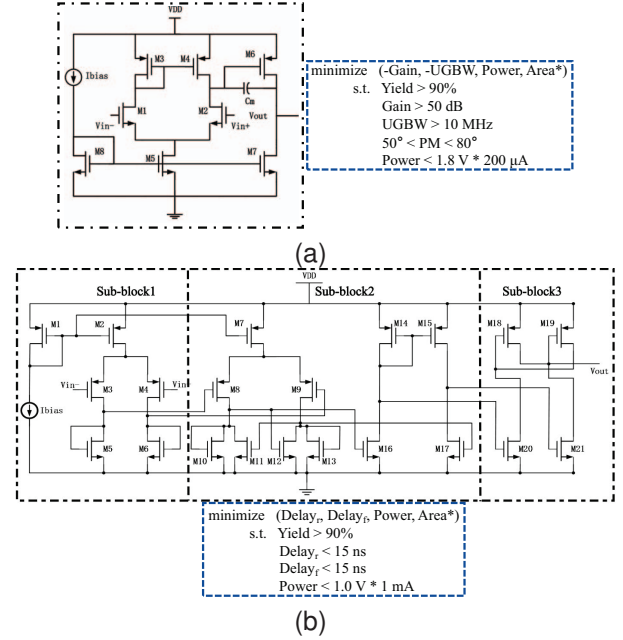


Fig. 3. Schematics and design specifications of two examples. (a) OPA. (b) CMP.

resulting in a total of 16 process parameters in consideration. A total of 9 corners are considered, including three supply voltages (1.62 V, 1.8 V, 1.98 V) and three temperatures (-40 °C, 27 °C, 85 °C). The worst performance metrics and yield under all corners are used for optimization. The optimization problem is also illustrated in Fig. 3(a), where Gain denotes the dc gain, UGBW means the unity-gain bandwidth, and PM is the phase margin. Power is the dc power consumption of the OPA. Area* is used to replace the actual area, which is the product of all design parameters.

We repeat the experiment of each method 10 times to verify the robustness of algorithms. The experimental results are shown in the first part of Table II. With the sensitivity-based fast yield analysis module, all algorithms achieve high-yield final designs ($> 96\%$) without additional MC simulations. Besides, RASNIL surpasses the baselines in both HV and success rate, achieving up to a 430% improvement in HV.

RASNIL achieves a speed up of 4.17 \times compared to NSGA-II* and cuts training time by 99% compared to MFBO*. As shown in Fig. 4(a), hybrid fidelity models retrain fully only when the self-adaptive criterion is met; otherwise, they perform much faster incremental updates. The prescreening technique also reduces model prediction time. SILE*-1k involves only 111 HF evaluations, resulting in a smaller dataset and shorter training time, but also relatively few iterations. In contrast, RASNIL, as a hybrid fidelity optimization algorithm,

TABLE II
STATISTICAL OPTIMIZATION RESULTS OF OPA AND CMP. SILE*-1K
AND SILE*-3K REPRESENT THE MAXIMUM SIMULATION COUNT
SETTINGS OF 1000 AND 3000, RESPECTIVELY

Methods		NSGA-II*	MFBO*	SILE*-1k	SILE*-3k	RASNIL
OPA	Max HV	162.7	113.9	127.7	126.7	200.8
	Min HV	0	0	0	3.7	12.9
	Mean HV	34.0	12.2	39.1	60.3	64.7
	Mean Yield	97.7%	97.6%	96.4%	97.6%	97.8%
	Success Rate	6/10	3/10	7/10	10/10	10/10
	Number of HF	555	82	111	333	79
	Number of LF	0	261	0	0	292
	Training Time/s	N/A	6508	10	152	79
Total Runtime/s	16339	10242	3656	10528	3916	
CMP	Max HV	61.3	61.4	73.8	79.1	80.7
	Min HV	0	0	0	21.9	25.5
	Mean HV	20.9	13.8	14.1	48.4	49.4
	Mean Yield	94.5%	97.0%	96.4%	97.2%	97.3%
	Success Rate	8/10	4/10	3/10	10/10	10/10
	Number of HF	555	83	111	333	81
	Number of LF	0	254	0	0	276
	Training Time/s	N/A	1270	8	129	64
Total Runtime/s	4481	2301	843	2597	958	

employs dozens of HF evaluations and hundreds of LF evaluations. This enables more iterations, which are beneficial for refining the search process, thereby yielding improved solution quality under a comparable total runtime. Furthermore, RASNIL delivers a 2.69 \times acceleration over SILE*-3k without compromising solution quality.

B. Comparator

This example is a Comparator (CMP) implemented in a 65 nm library at 1.0 V supply voltage as shown in Fig. 3(b). There are 24 design variables, including the lengths and widths of transistors. Both V_{th} and v_{sat} are considered, resulting in 42 process parameters. A total of 9 corners are considered including three supply voltages (0.9 V, 1.0 V, 1.1 V) and three temperatures (-40°C , 27°C , 85°C). The worst performance metrics and yield under all corners are used for optimization. The optimization problem is also illustrated in Fig. 3(b), where Delay_r denotes the delay time of the rising edge, Delay_f denotes the delay time of the falling edge, and Power is the dc power consumption of the CMP. Area^* is the product of all design parameters.

We also repeat the experiment of each method 10 times. The experimental results are shown in the second part of Table II. RASNIL algorithm also outperforms the baselines in terms of solution quality, achieving up to a 258% improvement in HV.

RASNIL achieves a speed up of 4.68 \times compared to NSGA-II*, and cuts training time by 95% relative to MFBO*—an improvement made possible by the incremental learning scheme. In addition, compared with SILE*-1k and SILE*-3k, RASNIL achieves significant improvements in solution quality and optimization speed, respectively.

C. Ablation Study

To evaluate the effectiveness of dynamic weighted LOF within the nested hybrid fidelity framework combined with the general many-objective optimization strategy, we conducted ablation studies. RASNIL-Base in Table III denotes

TABLE III
ABLATION STUDY RESULTS

Methods		RASNIL-Base	RASNIL
OPA	Mean HV	36.4	64.7
	Mean Yield	95.6%	97.8%
	Success Rate	8/10	10/10
	Total Runtime/s	3885	3916
CMP	Mean HV	21.1	49.4
	Mean Yield	96.6%	97.3%
	Success Rate	10/10	10/10
	Total Runtime/s	937	958

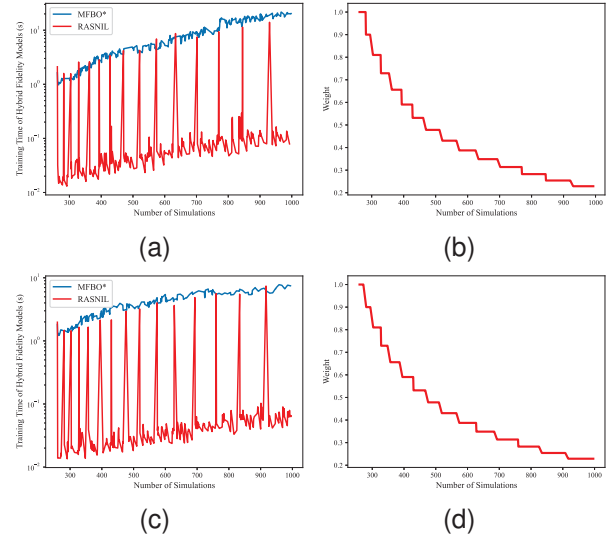


Fig. 4. The comparison of training time and the changes in weight. (a)-(b) OPA. (c)-(d) CMP.

our method without LOF and the many-objective strategy. In the early phase (Fig. 4(b)(d)), the LOF weight remains near 1, guiding the algorithm to balance convergence with diversity, thereby avoiding local optima. As the search progresses, the weight gradually decreases to ensure convergence. Meanwhile, the many-objective strategy efficiently extracts Pareto solutions from numerous candidates while mitigating the DR phenomenon in classical NSGA-II*. Overall, RASNIL achieves higher solution quality without incurring additional time costs, consistent with the experimental results shown in Table III.

V. CONCLUSION

We propose RASNIL that integrates an incremental learning-based training technique, a nested hybrid fidelity framework, and a general many-objective optimization strategy, to deliver fast, efficient PVT-aware analog circuit sizing across numerous design goals—outperforming state-of-the-art methods. Moreover, its sensitivity-based rapid yield analysis makes the optimizer consider the robustness of circuits within the optimization process without requiring MC simulations.

ACKNOWLEDGMENT

This work is supported by the National Key Research and Development Program of China: Design Technology Co-optimization Methodology (granted NO:2023YFB4402701).

REFERENCES

- [1] R. A. Rutenbar, G. G. E. Gielen, and J. Roychowdhury, "Hierarchical modeling, optimization, and synthesis for system-level analog and rf designs," *Proceedings of the IEEE*, vol. 95, no. 3, pp. 640–669, 2007.
- [2] K. Agarwal, A. Jain, D. Amuru, and Z. Abbas, "Fast and efficient resnn and genetic optimization for pvt aware performance enhancement in digital circuits," in *2022 International Symposium on VLSI Design, Automation and Test (VLSI-DAT)*, 2022, pp. 1–4.
- [3] A. F. Budak, D. Smart, B. Swahn, and D. Z. Pan, "Apostle: Asynchronously parallel optimization for sizing analog transistors using dnn learning," in *2023 28th Asia and South Pacific Design Automation Conference (ASP-DAC)*, 2023, pp. 70–75.
- [4] Y. Choi, M. Choi, K. Lee, and S. Kang, "Ma-opt: Reinforcement learning-based analog circuit optimization using multi-actors," in *2023 Design, Automation Test in Europe Conference Exhibition (DATE)*, 2023, pp. 1–5.
- [5] K. Touloupas and P. P. Sotiriadis, "Locomobo: A local constrained multiobjective bayesian optimization for analog circuit sizing," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 41, no. 9, pp. 2780–2793, 2022.
- [6] S. Yin, R. Wang, J. Zhang, X. Liu, and Y. Wang, "Fast surrogate-assisted constrained multiobjective optimization for analog circuit sizing via self-adaptive incremental learning," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 42, no. 7, pp. 2080–2093, 2023.
- [7] X. Zhang, Y. Tian, R. Cheng, and Y. Jin, "A decision variable clustering-based evolutionary algorithm for large-scale many-objective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 22, no. 1, pp. 97–112, 2018.
- [8] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: Nsga-ii," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.
- [9] S. Zhang, W. Lyu, F. Yang, C. Yan, D. Zhou, X. Zeng, and X. Hu, "An efficient multi-fidelity bayesian optimization approach for analog circuit synthesis," in *2019 56th ACM/IEEE Design Automation Conference (DAC)*, 2019, pp. 1–6.
- [10] B. He, S. Zhang, Y. Wang, T. Gao, F. Yang, C. Yan, D. Zhou, Z. Bi, and X. Zeng, "A batched bayesian optimization approach for analog circuit synthesis via multi-fidelity modeling," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 42, no. 2, pp. 347–359, 2023.
- [11] J. Huang, C. Wang, Y. Yan, C. Tao, F. Yang, C. Yan, W. Hu, D. Zhou, and X. Zeng, "An analog circuit building block generator via nested multi-fidelity modeling," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 70, no. 8, pp. 3280–3293, 2023.
- [12] B. Liu, F. V. Fernandez, and G. G. E. Gielen, "Efficient and accurate statistical analog yield optimization and variation-aware circuit sizing based on computational intelligence techniques," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 30, no. 6, pp. 793–805, 2011.
- [13] C. Jiang, Z. Ye, and Y. Wang, "A systematic design methodology for yield-driven near-threshold sram design," in *2014 12th IEEE International Conference on Solid-State and Integrated Circuit Technology (ICSICT)*, 2014, pp. 1–3.
- [14] E. Afacan, G. Berkol, A. E. Pusane, G. Dündar, and F. Başkaya, "A hybrid quasi monte carlo method for yield aware analog circuit sizing tool," in *2015 Design, Automation Test in Europe Conference Exhibition (DATE)*, 2015, pp. 1225–1228.
- [15] 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 *Design, Automation Test in Europe Conference Exhibition (DATE)*, 2017, pp. 1201–1206.
- [16] A. Canelas, R. Póvoa, R. Martins, N. Lourenço, J. Guilherme, J. P. Carvalho, and N. Horta, "Fuzye: A fuzzy c-means analog ic yield optimization using evolutionary-based algorithms," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 39, no. 1, pp. 1–13, 2020.
- [17] D. R. Jones, M. Schonlau, and W. J. Welch, "Efficient global optimization of expensive black-box functions," *Journal of Global Optimization*, vol. 13, no. 4, pp. 455–492, 1998.
- [18] K. Touloupas, N. Chouridis, and P. P. Sotiriadis, "Local bayesian optimization for analog circuit sizing," in *2021 58th ACM/IEEE Design Automation Conference (DAC)*, 2021, pp. 1237–1242.
- [19] W. Hu, Z. Ye, and Y. Wang, "Adjoint transient sensitivity analysis for objective functions associated to many time points," in *2020 57th ACM/IEEE Design Automation Conference (DAC)*, 2020, pp. 1–6.
- [20] Z. Gao and R. Rohrer, "Efficient non-monte-carlo yield estimation," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 41, no. 5, pp. 1222–1235, May 2022.
- [21] Z. He and Z. Zhang, "Pobo: A polynomial bounding method for chance-constrained yield-aware optimization of photonic ics," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 41, no. 11, pp. 4915–4926, 2022.
- [22] D. Zhan and H. Xing, "A fast kriging-assisted evolutionary algorithm based on incremental learning," *IEEE Transactions on Evolutionary Computation*, vol. 25, no. 5, pp. 941–955, 2021.
- [23] K. Shang, H. Ishibuchi, L. He, and L. M. Pang, "A survey on the hypervolume indicator in evolutionary multiobjective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 25, no. 1, pp. 1–20, 2021.
- [24] X. Yu and M. Gen, *Introduction to Evolutionary Algorithms*. London: Springer London, 2010.
- [25] D. Zhan, Y. Cheng, and J. Liu, "Expected improvement matrix-based infill criteria for expensive multiobjective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 21, no. 6, pp. 956–975, 2017.
- [26] Y. Peng, Y. Yang, Y. Xu, Y. Xue, R. Song, J. Kang, and H. Zhao, "Electricity theft detection in ami based on clustering and local outlier factor," *IEEE Access*, vol. 9, pp. 107 250–107 259, 2021.
- [27] J. Han, M. Kamber, and J. Pei, "3 - data preprocessing," in *Data Mining: Concepts and Techniques (Third Edition)*, third edition ed. Boston: Morgan Kaufmann, 2012, pp. 83–124.
- [28] K. Deb, L. Thiele, M. Laumanns, and E. Zitzler, *Scalable Test Problems for Evolutionary Multiobjective Optimization*. London: Springer London, 2005, pp. 105–145.
- [29] S.-C. Liu, Z.-H. Zhan, K. C. Tan, and J. Zhang, "A multiobjective framework for many-objective optimization," *IEEE Transactions on Cybernetics*, vol. 52, no. 12, pp. 13 654–13 668, 2022.
- [30] E. R. Keiter, K. V. Aadithya, T. Mei, T. V. Russo, R. L. Schiek, P. E. Sholander, H. K. Thornquist, and J. C. Verley, "Xyce parallel electronic simulator users' guide," 2017.