Asynchronous Batch Constrained Multi-Objective Bayesian Optimization for Analog Circuit Sizing

Xuyang Zhao[12], Zhaori Bi[12]*, Changhao Yan[12], Fan Yang[12], Ye Lu[13], Dian Zhou[4], Xuan Zeng[12]*

[1]State Key Laboratory of Integrated Chips & System, Fudan University, China
[2]School of Microelectronics, Fudan University, China
[3]School of Information Science and Technology, Fudan University, China
[4]Department of Electrical Engineering, University of Texas at Dallas, USA

Introduction

- The schematic design of analog IC includes two main steps:
  - Topology selection
  - Design parameter optimization (Sizing)
- For analog sizing automation, the problems are commonly formulated as single-objective optimization tasks by the Figure of Merit (FOM).

$$\mathrm{FOM} = \sum_{i=1}^{N} w_i f^{(i)}(\boldsymbol{x})$$

- When specifications change, the single-objective optimization has to be re-run.
- Multi-objective optimization provides a set of optimal solutions.

Multi-Objective Optimization

- Methods:
- Evolutionary Algorithm (EA)-based
  - NSGA-II (TEVC 2002), MOEA/D (TEVC 2007)
  - Generate candidate population and refine the outcomes through crossover and mutation operations
  - Need a substantial number of simulations
- Bayesian Optimization (BO)-based
  - Utilize surrogate models to approximate circuit performances
  - Sequential: MOBO (DAC 2018)
  - Synchronous batch: LoCoMOBO (TCAD 2022)
  - Asynchronous batch: AEIM (TCAS-II 2022)

ABCMOBO

- This paper proposes a novel algorithm called Asynchronous Batch Constrained Multi-Objective Bayesian Optimization (ABCMOBO) for analog circuit sizing.
  - We introduce the Expected HyperVolume Improvement (EHVI) acquisition function into multi-objective analog circuit sizing problems and propose a novel constraint-handling strategy for MOBO.
  - We propose an asynchronous batch MOBO framework based on constrained EHVI and Cached Box Decomposition (CBD).
- The experimental results show that ABCMOBO outperforms the state-of-the-art methods, achieving a speed-up of 3.49x~8.18x with comparable results in two real-world analog ICs.

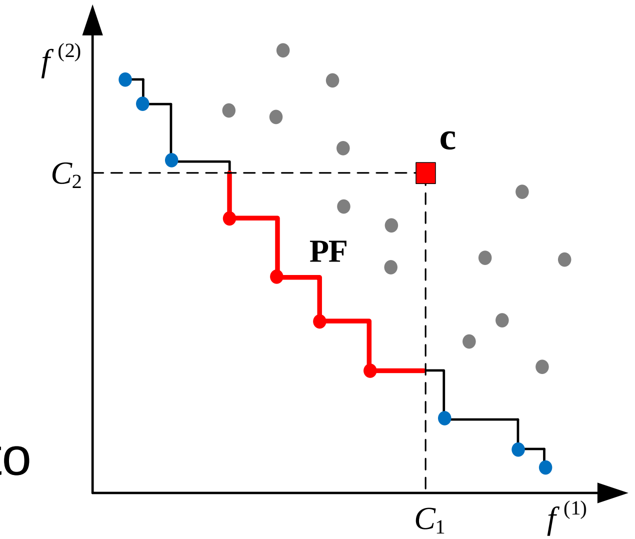
Problem Formulation

- A constrained multi-objective analog circuit sizing problem can be defined as equation (1).

$$\begin{aligned} \text{minimize} \quad & f^{(1)}(\boldsymbol{x}), \cdots, f^{(M)}(\boldsymbol{x}). \\ \text{s.t.} \quad & f^{(i)}(\boldsymbol{x}) \le C_i, \qquad i = 1, \cdots, M. \end{aligned} \tag{1}$$

- For arbitrary data points $\boldsymbol{a}$ and $\boldsymbol{b}$, $\boldsymbol{a} \prec \boldsymbol{b}$ ($\boldsymbol{a}$ *dominates* $\boldsymbol{b}$) if

$$\begin{aligned} & \forall i \in \{1, \cdots, M\} \quad f^{(i)}(\boldsymbol{a}) \le f^{(i)}(\boldsymbol{b}), \\ \text{and} \quad & \exists i \in \{1, \cdots, M\} \quad f^{(i)}(\boldsymbol{a}) < f^{(i)}(\boldsymbol{b}). \end{aligned} \tag{2}$$

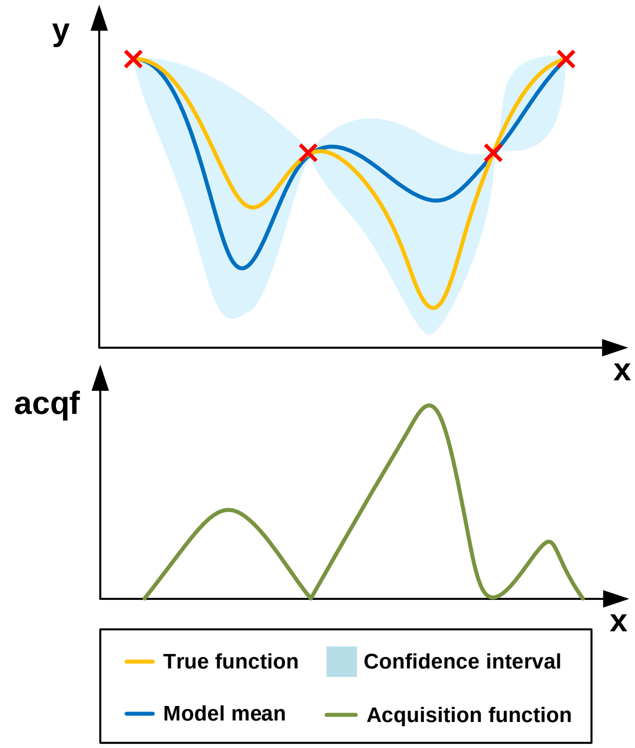


- The multi-objective optimization returns a Pareto Set PS and Pareto Front PF.

$$\begin{aligned} PS &= \{\, \boldsymbol{x} \in X_C \mid \nexists \boldsymbol{x}' \in X_C : \boldsymbol{x}' \prec \boldsymbol{x} \}, \\ X_C &= \{\, \boldsymbol{x} \in X \mid f^{(i)}(\boldsymbol{x}) \le C_i, i = 1, \cdots, M \}, \\ PF &= \{\, \boldsymbol{f}(\boldsymbol{x}) \mid \boldsymbol{x} \in PS \}. \end{aligned} \tag{3}$$

Bayesian Optimization

- Bayesian Optimization (BO) comprises two fundamental components:
  - Surrogate model
  - Acquisition function
- Gaussian Process Regression (GPR) model is a commonly used surrogate model.
- It can provide posterior distribution for an arbitrary $\boldsymbol{x}^*$ as follow.



$$\begin{cases} \mu(\boldsymbol{x}^*) = m(\boldsymbol{x}) + k(\boldsymbol{x}^*, X)[K + \sigma_n^2 I]^{-1}(\boldsymbol{y} - m(\boldsymbol{x})) \\ \sigma^2(\boldsymbol{x}^*) = k(\boldsymbol{x}^*, \boldsymbol{x}^*) - k(\boldsymbol{x}^*, X)[K + \sigma_n^2 I]^{-1} k(X, \boldsymbol{x}^*) \end{cases} \tag{4}$$

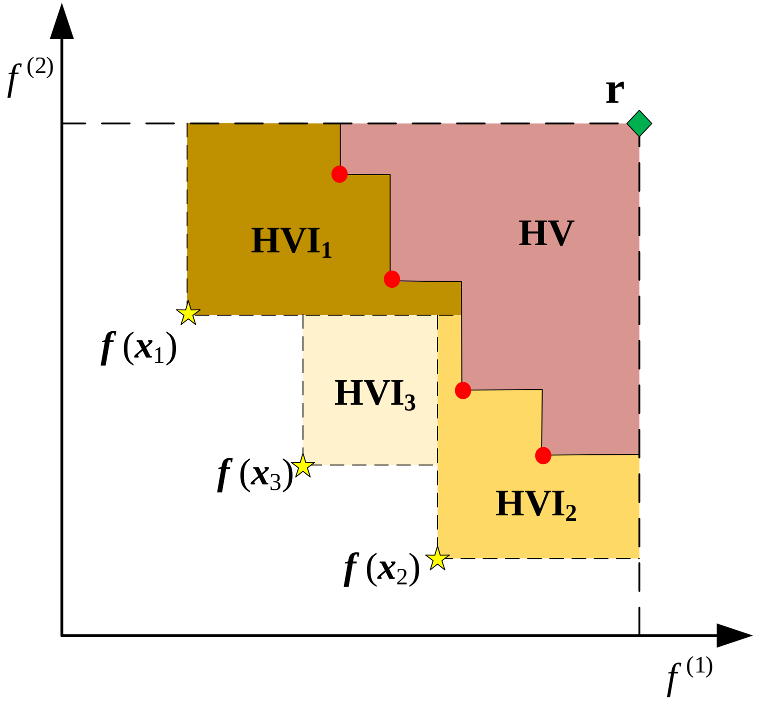
Hypervolume

- Hypervolume (HV) is a criterion for assessing the quality of Pareto set PS in multi-objective optimization.

$$\mathrm{HV}(PS, \boldsymbol{r}) = \lambda_M\left(\bigcup_{i=1}^{|PS|} [\boldsymbol{r}, \boldsymbol{y}_i]\right) \quad (5)$$

- The hypervolume improvement (HVI) is defined by (6).

$$\mathrm{HVI}(\boldsymbol{f}(\boldsymbol{x}_*)|PS, \boldsymbol{r}) = \mathrm{HV}(PS \cup \boldsymbol{x}_*, \boldsymbol{r}) - \mathrm{HV}(PS, \boldsymbol{r}) \quad (6)$$



EHVI Acquisition Function

- EHVI is the expectation of HVI over the posterior of the GPR models.

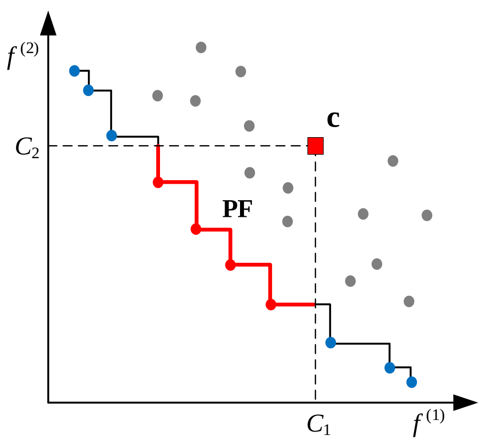
$$\alpha_{EHVI}(X_{cand}|PS, \boldsymbol{r}) = \mathbb{E}[\mathrm{HVI}(\boldsymbol{f}(X_{cand})|PS, \boldsymbol{r})] \qquad (7)$$

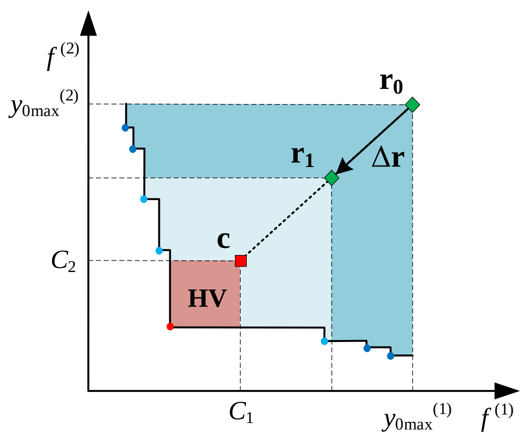
- Usually, EHVI can be approximated with MC integration.

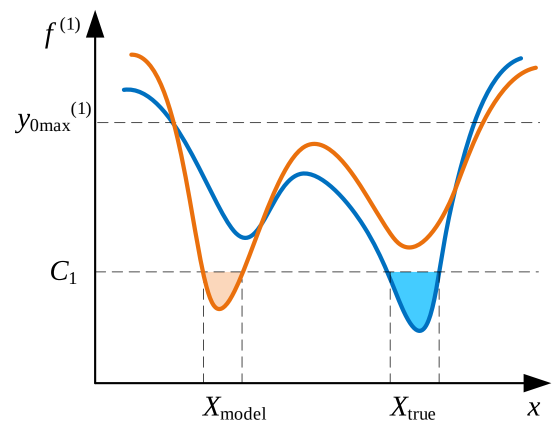
$$\hat{\alpha}_{EHVI}^{N}(X_{cand}|PS, \boldsymbol{r}) = \frac{1}{\mathrm{N}}\sum_{t=1}^{N} \mathrm{HVI}(\boldsymbol{f}_t(X_{cand})|PS, \boldsymbol{r}) \qquad (8)$$

Constraint Handling

- IC designs that fail to meet specifications are worthless.
- Solutions beyond the reference point don't contribute to the hypervolume.
- Specifications can be set as the reference point in ABCMOBO.
- Selecting strict constraints as the reference point in the early stage may lead to the omission of potential optimal regions.





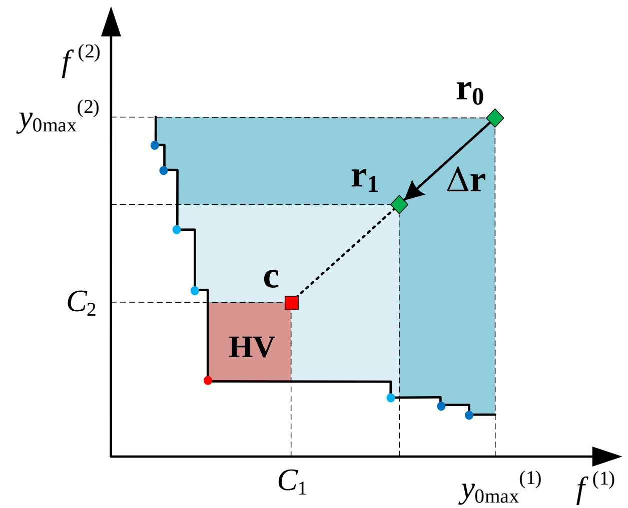


Dynamic Reference Point Selection

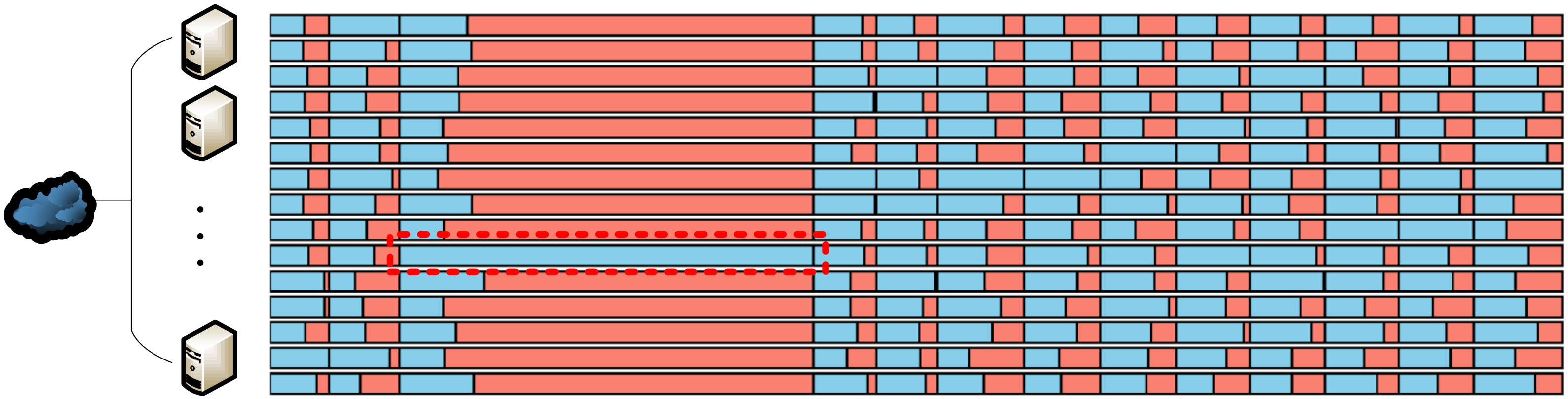
Algorithm 1 Dynamic Reference Point Selection

Input: The dataset D_t, the constraints $\boldsymbol{c}$, the hyperparameter k_s and the maximum number of iterations N_{iter}

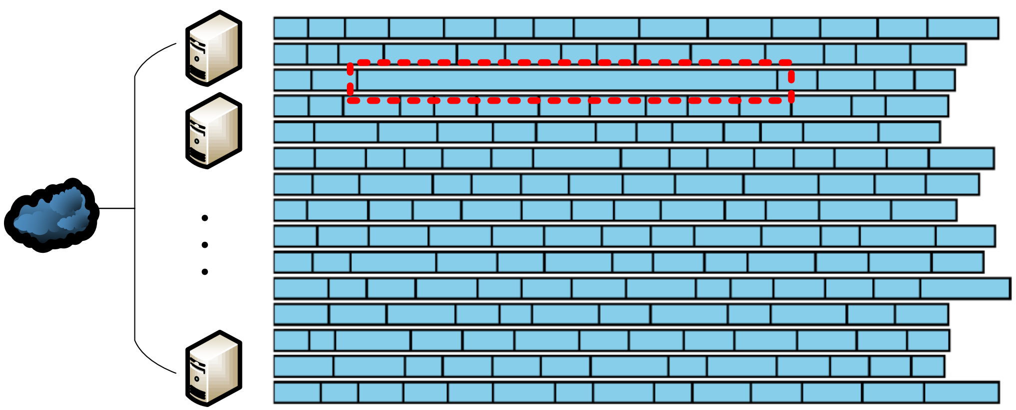
$r_0^{(i)} = y_{0max}^{(i)} = \max(y^{(i)} | \boldsymbol{y} \in D_0),\ i = 1, \cdots, M$
$\Delta r^{(i)} = k_s(C_i - y_{0max}^{(i)}),\ i = 1, \cdots, M$
$t = 1$
while $t < N_{iter}$ **and** $\boldsymbol{r}_t \neq \boldsymbol{c}$ **do**
    if $\exists \boldsymbol{y}^* \in D_t, \boldsymbol{y}^* \prec \boldsymbol{r}_{t-1} + \Delta \boldsymbol{r}$ **then**
        $\boldsymbol{r}_t = \boldsymbol{r}_{t-1} + \Delta \boldsymbol{r}$
        if $\boldsymbol{r}_t \prec \boldsymbol{c}$ **then**
            $\boldsymbol{r}_t = \boldsymbol{c}$
    $t = t + 1$



Synchronous VS Asynchronous



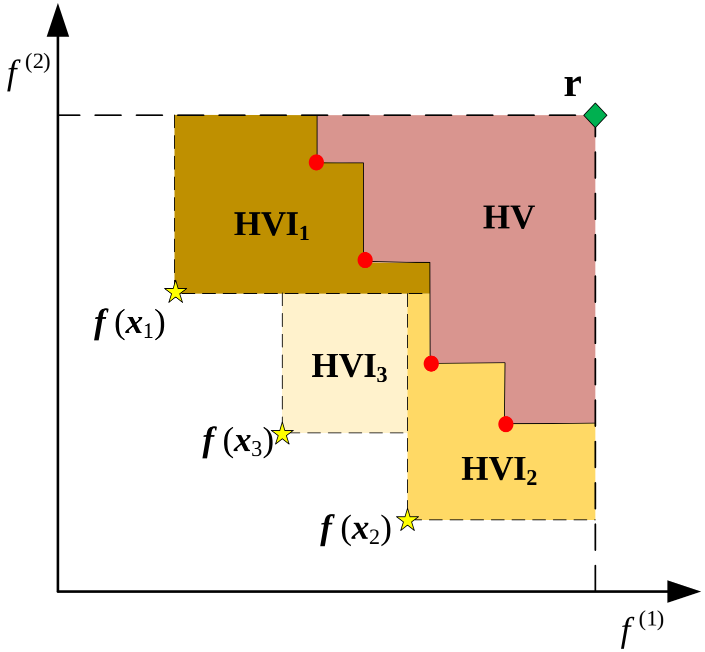
Synchronous



Asynchronous

Asynchronous Batch MOBO

- The current simulation point $\boldsymbol{x}_i$ may not return before the subsequent optimization.
- We put the simulation point $\boldsymbol{x}_i$ into a pending point set and form a pseudo Pareto set.
- Because of the definition of EHVI, the next candidate $\boldsymbol{x}_{i+1}$ must be different from existing points.



ABCMOBO Framework

Algorithm 2 ABCMOBO Algorithm

Input: The size of the initial dataset N_{init}, the maximum number of iterations N_{iter} and the batch size B

1: Generate M initial datasets $D^{(i)} = \{X, \boldsymbol{y}^{(i)}\}$ by randomly sampling, $i = 1, \cdots, M$

2: **for** $t = 0 \rightarrow N_{iter}$ **do**

3: Wait for a worker to be available.

4: Update the training datasets $D_t^{(i)} = D_{t-1}^{(i)} \cup \{\boldsymbol{x}_t, y_t^{(i)}\}$, where $\{\boldsymbol{x}_t, y_t^{(i)}\}$ is the newly observed data.

5: Update the reference point $\boldsymbol{r}_t$ by DRS (Algorithm 1).

6: Construct and train M GPR models with $D_t^{(i)}$.

7: Obtain the predictive means $\hat{\boldsymbol{y}}^{(i)} = \{\hat{y}_j^{(i)}, \cdots, \hat{y}_l^{(i)}\}$ of the already selected query points $\hat{X} = \{\hat{\boldsymbol{x}}_j, \cdots, \hat{\boldsymbol{x}}_l\}$ which are still under simulation.

8: Select the next candidate point $\boldsymbol{x}_{t+1}$ by optimizing the EHVI acquisition function $\alpha_{EHVI}(\boldsymbol{x}|D_t \cup \{\hat{X}, \hat{\boldsymbol{y}}\}, \boldsymbol{r}_t)$.

9: Simulate the candidate point $\boldsymbol{x}_{t+1}$ at the idle worker.

10: **return** Pareto front of $\{\boldsymbol{y}^{(i)}\}_{i=1}^{M}$.

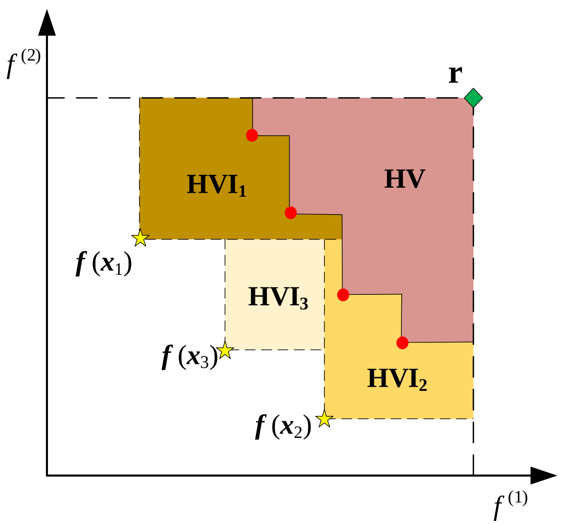
Cached Box Decompositions

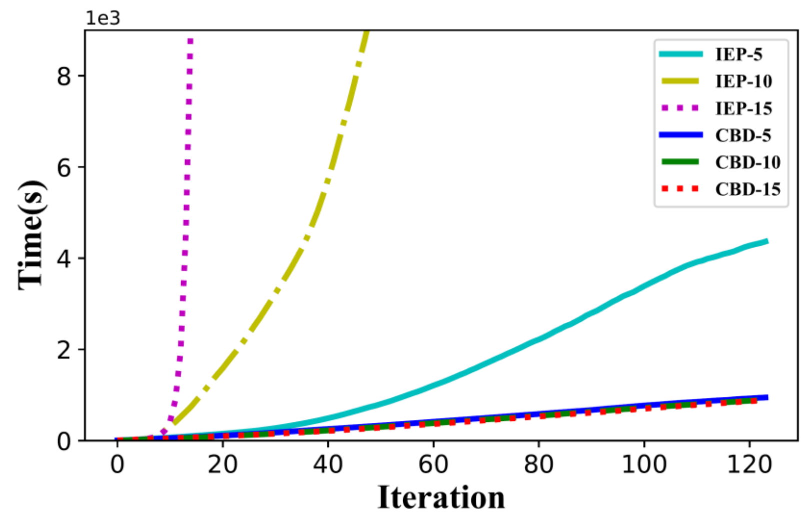
- Inclusion-Exclusion Principle (IEP)-based:

$$\hat{\alpha}^{N}_{EHVI_{IEP}}(X_{cand}) = \frac{1}{N}\sum_{t=1}^{N}\sum_{j=1}^{B}\sum_{X_j \in \mathcal{X}_j}(-1)^{j+1}\mathrm{HVI}(\boldsymbol{f}_t(X_j)) \tag{9}$$

- Cached Box Decompositions (CBD)-based:

$$\hat{\alpha}^{N}_{EHVI_{CBD}}\left(\{\boldsymbol{x}_j\}_{j=1}^{i}\right) = \frac{1}{N}\sum_{t=1}^{N}\mathrm{HVI}\left(\{\tilde{\boldsymbol{f}}_t(\boldsymbol{x}_j)\}_{j=1}^{i-1}\middle|PF\right) + \frac{1}{N}\sum_{t=1}^{N}\mathrm{HVI}(\tilde{\boldsymbol{f}}_t(\boldsymbol{x}_i)|PF \cup \{\tilde{\boldsymbol{f}}_t(\boldsymbol{x}_j)\}_{j=1}^{i-1}) \tag{10}$$



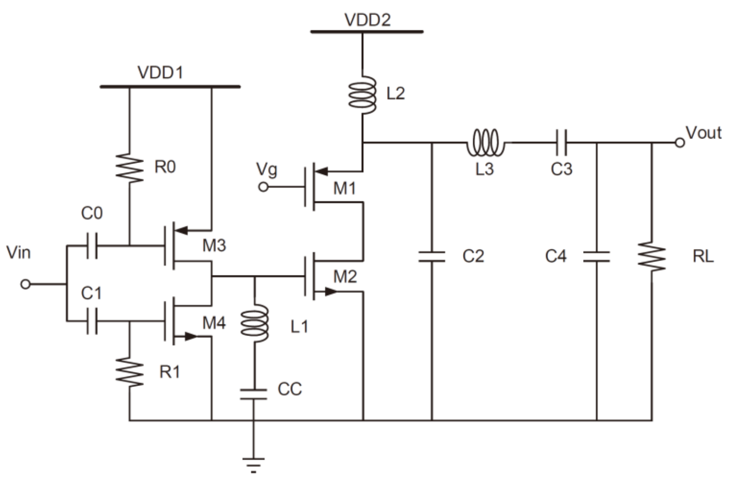


Experimental Setup

- Two real-world analog circuits:
  - Class-E power amplifier
  - Two-stage operational amplifier
- Baselines:
  - Sequential:
    - NSGA-II (TEVC 2002)
    - MOEA/D (TEVC 2007)
    - MOBO (DAC 2018)
  - Batch:
    - LoCoMOBO (TCAD 2022), synchronous
    - AEIM (TCAS-II 2022), asynchronous
  - EHVI-based synchronous batch methods:
    - EHVI_0, which sets the original point $\boldsymbol{o}$ as the reference point
    - EHVI_C, which sets the constraint $\boldsymbol{c}$ as the reference point
    - BCMOBO, which incorporates dynamic reference point selection
- Batch size: 15

Class-E Power Amplifier

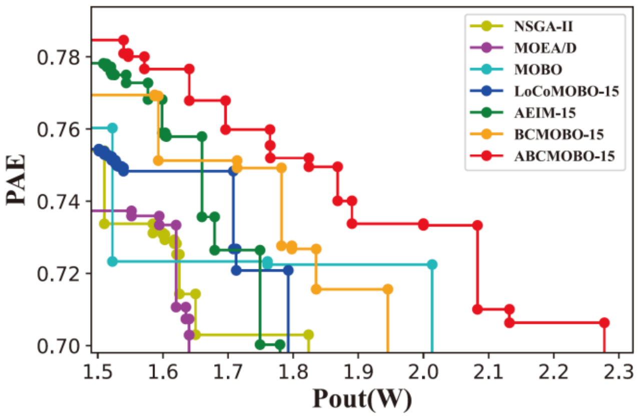
- 12 design variables
- 2 circuit performances:
  - Power-added efficiency (PAE)
  - Output power (P_{out})

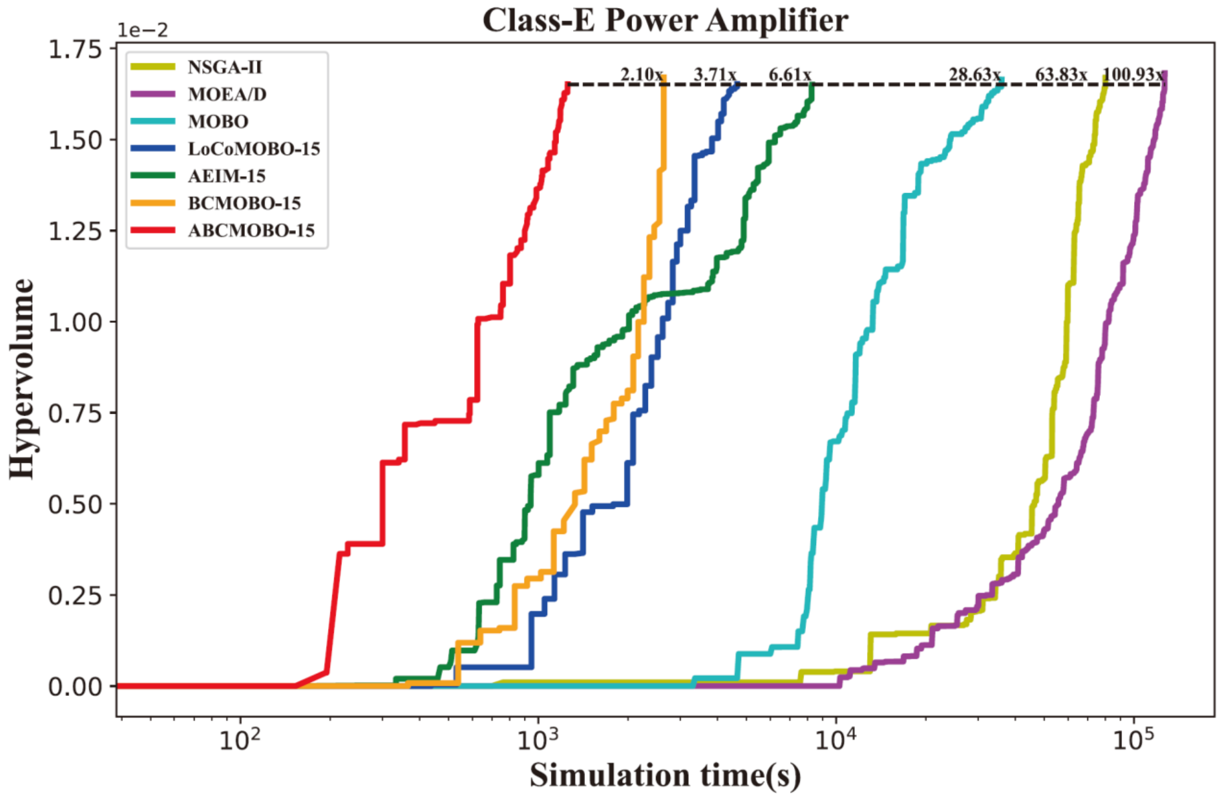


$$
\begin{aligned}
\text{maximize} \quad & PAE, Pout(\mathrm{W}) \\
\text{s.t.} \quad & PAE \geq 0.7, \\
& P_{out} \geq 1.5\ \mathrm{W}.
\end{aligned}
\tag{11}
$$

Class-E Power Amplifier

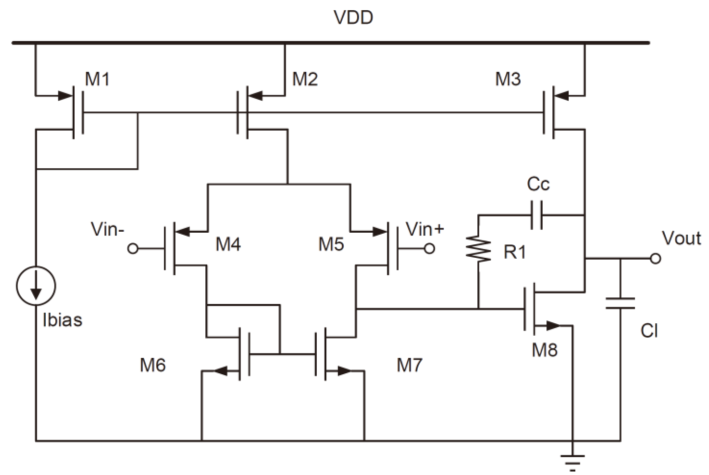
Algorithm	#Succ.	HV($\times 10^{-2}$)	Sim. Time(s)	Speed up
MOEA/D	10/10	1.6385±0.8121	126620	0.07×
NSGA-II	10/10	1.6429±0.3879	80073	0.10×
MOBO	10/10	1.6645±0.3808	35917	0.23×
AEIM-15	5/10	1.6504±0.4141	8295	1.00×
LoCoMOBO-15	7/10	1.6521±0.4702	4650	1.78×
EHVI_0-15	4/10	0.0897±0.0214	3378	-
EHVI_C-15	4/10	0.6196±0.4677	3379	-
BCMOBO-15	10/10	1.6590±0.5682	2637	3.15×
ABCMOBO-15	10/10	**1.6717±0.1659**	**1254**	**6.61×**





Two-Stage Operational Amplifier

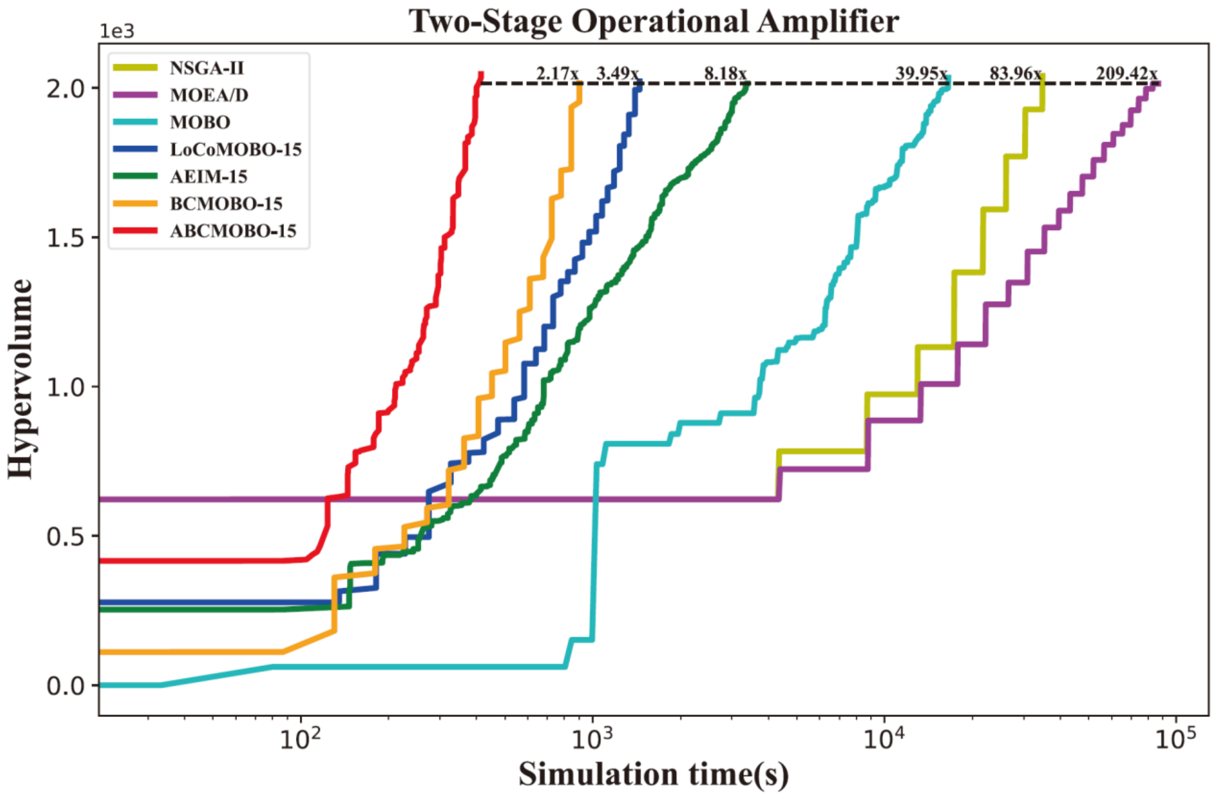
- 10 design variables
- 3 circuit performances:
  - Open-loop gain ($GAIN$)
  - Unity gain frequency (UGF)
  - Phase margin (PM)



$$\begin{aligned} \text{maximize} \quad & GAIN(\text{dB}), UGF(\text{MHz}), PM(°) \\ \text{s.t.} \quad & GAIN \geq 80\ \text{dB}, \\ & UGF \geq 10\ \text{MHz}, \\ & PM \geq 60\ °. \end{aligned} \tag{12}$$

Two-Stage Operational Amplifier

Algorithm	#Succ.	HV($\times 10^3$)	Sim. Time(s)	Speed up
MOEA/D	10/10	2.0146±0.5631	86772	0.04×
NSGA-II	10/10	2.0403±0.3052	34787	0.10×
MOBO	10/10	2.0342±0.2862	16552	0.20×
AEIM-15	10/10	2.0085±0.5653	3391	1.00×
LoCoMOBO-15	10/10	2.0066±0.1674	1448	2.34×
EHVI_0-15	10/10	1.2156±0.1273	1281	-
EHVI_C-15	10/10	0.9271±0.0778	1280	-
BCMOBO-15	10/10	2.0130±0.5980	900	3.76×
ABCMOBO-15	10/10	**2.0465±0.3240**	**414**	**8.18**×



Conclusion

- This paper introduces an innovative approach for multi-objective optimization in analog circuit sizing, specifically targeting constrained optimization problems.
- Our proposed method is based on asynchronous batch optimization and incorporates a DRS strategy within the EHVI acquisition function.
- Furthermore, our acquisition function’s ability to transform constrained multi-objective problems into unconstrained single-objective problems opens up possibilities for exploring various efficient optimization techniques in future research endeavors.

Thanks for listening!

Q&A