

Causal-Guided Dimension Reduction for Efficient Pareto Optimization

Dinithi Jayasuriya, Divake Kumar, Sureshkumar Senthilkumar, Devashri Naik, Nastaran Darabi and Amit Ranjan Trivedi
University of Illinois at Chicago, IL, USA, email: {dkasth2, dkumar33, ssenth7, dnaik6, ndarab2, amitrtr}@uic.edu

Abstract—Multi-objective optimization of analog circuits is hindered by high-dimensional parameter spaces, strong feedback couplings, and expensive transistor-level simulations. Evolutionary algorithms such as Non-dominated Sorting Genetic Algorithm I (NSGA-II) are widely used but treat all parameters equally, hereby wasting effort on variables with little impact on performance, which limits their scalability. We introduce CaDRO, causal-guided dimensionality reduction framework that embeds causal discovery into the optimization pipeline. CaDRO builds a quantitative causal map through a hybrid observational-interventional process, ranking parameters by their causal effect on the objectives. Low-impact parameters are fixed to values from high-quality solutions, while critical drivers remain active in the search. The reduced design space enables focused evolutionary optimization without modifying the underlying algorithm. Across amplifiers, regulators, and RF circuits, CaDRO converges up to $0\times$ faster than NSGA-II while preserving or improving Pareto quality. For instance, on the Folded-Cascode Amplifier, hypervolume improves from 0.56 to 0.94, and on the LDO regulator from .65 to 0.81, with large gains in non-dominated solutions.

Index Terms—Analog Circuit Design, Multi-Objective Optimization, Causal Inference, Design Automation

I. INTRODUCTION

Analog circuit design remains one of the most intricate problems in EDA: designers must simultaneously satisfy conflicting objectives such as gain, bandwidth, power, noise, and area across dozens of interdependent device and bias parameters. These parameters interact nonlinearly through feedback loops, parasitics, and biasing networks, making the search space difficult to navigate. Brute-force or black-box optimization is prohibitively expensive for complex circuits and does not reveal which parameters truly drive performance, limiting critical design insights and their reuse across topologies, domains, and application constraints [1], [2].

To reduce simulation cost, prior work can be grouped into four main directions: *First*, surrogate and Bayesian optimization methods for analog circuit sizing, with recent efforts tackling high-dimensional BO via subspace or truncated sampling [3]. *Second*, ML-based mapping and heuristic methods predict parameters directly from specifications or guide search via learned domain knowledge (e.g., LEDRO’s use of LLMs for design space reduction) [4]. *Third*, RL and variation-aware optimization address real-world PVT variation and multi-task scenarios; RobustAnalog employs multi-task RL with pruning [5], while ROSE-Opt combines BO+RL with domain knowledge for robust optimization [6]. *Fourth*, dimensionality or search-space reduction techniques shrink parameter count or

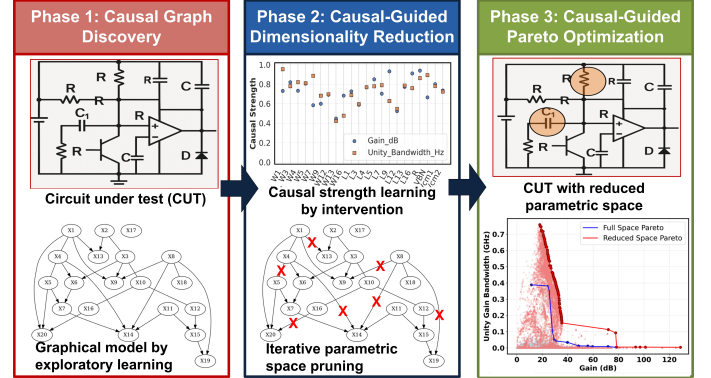


Fig. 1. **CaDRO framework**: Phase 1 learns the causal graph of the circuit under test (CUT). Phase 2 prunes weak parameters via causal strength estimation. Phase 3 performs Pareto optimization on the reduced space, achieving near full-space performance with lower complexity.

ranges via statistical methods, subspace selection, or sparse regression, often guided by feature importance or gradient approximations (e.g., Sparse Regression & Error Margining [7], LinEasyBO [8], LEDRO [4]).

While prior works have improved efficiency, they share a key limitation: parameter importance is typically inferred from correlation or sensitivity. In analog circuits, where dense interconnections and feedback loops are pervasive, such measures are often misleading. Variables may appear predictive through indirect coupling while having little direct effect on performance. As a result, optimizers waste simulations on parameters that do not meaningfully affect the Pareto front.

We address this gap by introducing *causal reasoning* as a systematic basis for dimensionality reduction (Fig. 1). Causal analysis provides a principled way to separate true design drivers from spurious associations. Unlike deep learning models that capture correlations without interpretability, causal reasoning makes dependencies explicit and actionable for design.

With this rationale, this paper introduces CaDRO: a causal-guided dimensionality reduction framework that embeds causal discovery into multi-objective evolutionary search for analog circuit design. We integrate CaDRO into the Non-dominated Sorting Genetic Algorithm II (NSGA-II) to assess its impact on multi-objective analog circuit optimization. In CaDRO, an initial exploratory run generates a diverse dataset, from which we construct a causal map using a hybrid observational-interventional approach. Each parameter is assigned a causal strength, allowing us to separate critical design drivers from low-impact variables. The latter are fixed to values from

high-performing solutions. The evolutionary algorithm then operates only on the reduced set, yielding order-of-magnitude simulation savings while preserving Pareto quality across analog and RF benchmarks. Across diverse analog circuits including amplifiers, regulators, and RF circuits, CaDRO converges up to $10\times$ faster than baseline NSGA-II while preserving or improving Pareto front quality. Beyond efficiency, causal maps reveal which parameters truly govern circuit behavior, making the approach scalable and interpretable.

II. BACKGROUND AND RELATED WORK

A. Causal Discovery

Causal discovery uncovers directed cause–effect relationships among variables, moving beyond correlations. A causal graph $G = (V, E)$ represents parameters and performance objectives as nodes V , with directed edges E denoting direct influences. Unlike correlation graphs, which capture statistical associations, causal graphs encode asymmetric, interventionally testable dependencies [9]–[11]. Common methods include:

- *Constraint-based algorithms* such as PC [12] and FCI [12], [13] infer causal directions from conditional independencies.
- *Score-based searches* (e.g., GES [14]) optimize a scoring criterion such as Bayesian information criterion (BIC).
- *Functional causal models* exploit asymmetries in data-generating processes: LiNGAM [15] assumes linear non-Gaussian models, while NOTEARS [16] casts discovery as continuous optimization.

Observational data alone are often insufficient due to confounding: two parameters may appear correlated not because one causes the other, but because both are driven by hidden factors. To resolve this, *interventions*, i.e., actively perturbing one parameter while holding others fixed, are used to confirm edges and quantify causal effect sizes [17]. Distinguishing true design drivers from spurious correlations is critical in analog circuits, where dense feedback loops and shared current paths often create misleading associations. Causal maps that quantify each parameter’s effect and confidence provide a principled basis for dimensionality reduction before costly multi-objective optimization. *Yet*, prior works have not closed this loop. E.g., Jiao et al. [18] extracted parameter–performance dependencies from simulations but did not integrate causal knowledge into optimization. Other approaches, such as surrogate-assisted Bayesian optimization (MACE [19]), reinforcement learning under PVT variation (RobustAnalog [20]), deterministic group-based search [21], and hybrid evolutionary–surrogate schemes [22] improve efficiency and Pareto quality, but still treat all variables as equally important; thereby facing scalability constraints as parametric space expands.

B. Non-Dominated Sorting Genetic Algorithm II (NSGA-II)

The Non-Dominated Sorting Genetic Algorithm II (NSGA-II) [23] is a widely used evolutionary algorithms for multi-objective optimization. Like other genetic algorithms, it maintains a population of candidate solutions and evolves them through selection, crossover, and mutation, with three features:

- *Fast non-dominated sorting*: Ranks solutions by Pareto dominance for scalable multi-objective selection.
- *Crowding distance*: Preserves diversity by favouring solutions in less crowded regions of the Pareto front.
- *Elitism*: Ensures the best non-dominated solutions are retained across generations.

NSGA-II has been extended to other variants, including NSGA-III [24] for many-objective problems and MOEA/D [25] for decomposition-based optimization. In analog circuit design, NSGA-II has been widely applied to trade off gain, bandwidth, noise, and power [2], [26], but its efficiency deteriorates as the number of variables grows, since all parameters are perturbed and evolved regardless of their true influence. This motivates dimensionality reduction techniques of CaDRO that shrink the search space while preserving solution quality.

III. CADRO: CAUSAL-GUIDED DIMENSIONALITY REDUCTION FOR MULTI-OBJECTIVE OPTIMIZATION

CaDRO is a three-phase pipeline that accelerates multi-objective optimization by reducing the effective dimensionality of the design space before expensive evolutionary search. Its core idea is to discover and exploit the causal structure linking design parameters to performance metrics by analyzing true cause–effect relationships. By quantifying and ranking these relationships, CaDRO directs computation to a small subset of influential parameters while safely fixing those with negligible impact. This selective focus yields substantial simulation savings without compromising Pareto quality. The pipeline consists of three phases: (i) Causal Discovery and Strength Analysis, (ii) Causal-Based Dimensionality Reduction, and (iii) Focused Multi-Objective Optimization. We discuss each below:

A. Phase 1: Causal Discovery and Strength Analysis

CaDRO begins by constructing a high-confidence map of cause–effect relationships within the circuit. To this end, we generate a comprehensive dataset through an exploratory run of NSGA-II that evolves a population of candidate solutions via selection, crossover, and mutation, while employing non-dominated sorting and crowding distance to balance convergence and diversity [23]. *Notably*, the goal here is not to find optimal solutions but to leverage NSGA-II’s exploration capability. By sampling a broad range of parameter combinations, the algorithm produces a dataset containing good, bad, and mediocre designs with solution diversity that is essential for capturing statistical relationships across the design space. In our framework, causal discovery is performed using 8k simulations, which also serve as the initial population for subsequent optimization.

With this dataset, we begin the observational analysis. In complex analog circuits, parameter–performance relationships are often highly non-linear, so a single analytical method risks missing key dependencies. To address this, we employ an ensemble approach combining Pearson correlation for linear trends, Random Forest models for non-linear effects, and mutual information for general dependencies. An initial confidence score is synthesized from this evidence: it starts with a base score from the statistical significance (p -value) of the Pearson

correlation, then is incrementally boosted if the link is also supported by other methods, such as high feature importance from tree-based models. This yields a robust preliminary score reflecting consensus across the ensemble.

To distinguish true causality from spurious correlations, we perform *interventional refinement*, an active learning process with targeted experiments on the most uncertain links [27], [28]. Each intervention runs two small sets of simulations: in the first, a single input parameter is fixed to a “low” value (e.g., its 25th percentile), and in the second to a “high” value (e.g., its 75th percentile), while other parameters vary. A statistical test compares the output distributions of the two groups: if a significant difference is observed, the causal link is confirmed and its confidence score upgraded; if not, the link is classified as spurious with confidence set to zero. The causal strength (effect size) is then quantified as the normalized difference in the output mean between the “high” and “low” groups, providing a direct measure of the parameter’s influence. In this way, we experimentally confirm or reject each causal hypothesis. After several interventions, the final output is a weighted directed graph, with each edge annotated by causal strength and a refined confidence score based on combined observational and interventional evidence [29].

B. Phase 2: Causal-Based Dimensionality Reduction

With the validated causal map from Phase 1, we obtain a reliable guide to which parameters truly drive circuit performance. Phase 2 uses this knowledge to simplify the optimization problem by focusing computational effort on influential parameters while filtering those with negligible impact. *First*, we compute an overall *importance score* for each parameter. For a given input P_i , the score $S(P_i)$ is defined as

$$S(P_i) = \sum_j |E(P_i, O_j) \times C(P_i, O_j)|,$$

where $E(P_i, O_j)$ is the causal effect size of P_i on objective O_j , and $C(P_i, O_j)$ is the corresponding confidence score. This formulation holistically captures total influence: the product weights effects by their reliability, the absolute value treats negative and positive effects equally, and the summation across objectives identifies parameters with broad influence as more critical than those with narrow but strong effects. Based on these scores, design parameters are partitioned into an *Active set*, containing the most influential parameters to be optimized, and a *Pruned set*, containing low-importance parameters removed from active search. Our framework supports both fixed (*top-k*) and adaptive pruning strategies, with the cutoff determined by the distribution of importance scores. Crucially, pruned parameters are not discarded or set arbitrarily, which could shift the design into suboptimal regions. Instead, they are fixed to values drawn from high-performing, non-dominated solutions obtained in Phase 1, anchoring the search in a proven region and providing a strong foundation for final optimization.

C. Phase 3: Focused Multi-Objective Optimization

The final phase performs multi-objective optimization on the reduced problem. NSGA-II is initialized to operate ex-

clusively in the low-dimensional search space defined by the active parameters from Phase 2. All evolutionary mechanisms, population management, crossover and mutation, and selection, are thus focused solely on this critical subset.

For each candidate solution, only the active parameters are generated. To enable simulation, the framework reconstructs a full parameter vector by combining these active values with the fixed values of pruned parameters. This ensures evaluation in a valid, high-performance region while avoiding exploration of inconsequential dimensions. By shrinking the search space, the optimizer sidesteps flat, low-impact regions, leading to an exponential reduction in volume and much faster convergence. This focused approach reduces both the number of simulations and the total time needed to obtain high-quality Pareto solutions.

IV. RESULTS AND EVALUATION

A. Benchmark Circuits and Design Objectives

We evaluate CaDRO on six benchmark circuits spanning amplifiers, regulators, and RF blocks. The Active-Load Differential Amplifier (ALDA) has seven parameters and aims to maximize low-frequency gain and UGBW while minimizing power (Fig. 2a). The Two-Stage OTA (TSOTA) includes 13 parameters, targeting high DC gain and UGBW with low power (Fig. 2d). The Folded-Cascode Amplifier (FCA) has 22 parameters and seeks high gain and bandwidth at low power (Fig. 2b). The Low-Dropout (LDO) Regulator, with nine parameters, minimizes load regulation and quiescent current while maximizing PSRR at 10 kHz (Fig. 2f). The Voltage-Controlled Oscillator (VCO) optimizes seven parameters of a cross-coupled LC tank to maximize oscillation frequency and minimize power (Fig. 2e). Finally, the Two-Stage Voltage Amplifier (TSVA) has six parameters and targets high gain and UGBW with low power (Fig. 2c). Mainly ablating the utility of causal reasoning for automated analog EDA, our experiments address two crucial questions: (i) how does the degree of causal-guided pruning trade off computational efficiency and solution quality, and (ii) how do the final optimization results from the pruned search space compare to traditional optimization?

B. Discovered Causal Strengths

In the causal discovery phase, we quantify the influence of each design parameter on performance objectives. We visualize the causal discovery results as per-parameter scatter panels as shown in Fig. 3. In each panel, the x-axis lists the input parameters, the y-axis reports the normalized causal strength in [0, 1], and different objectives are distinguished by legend-coded markers. A higher point on the y-axis means that parameter has a stronger direct influence on that objective; points near 0 indicate little to no influence. In FCA and TSOTA, most parameters show strong causal links, so aggressive pruning degrades performance; in FCA nearly all transistor dimensions affect both Gain and UGBW. In contrast, ALDA exhibits more selective dependencies: L1 and L5 (differential pair and load) dominate Gain, while the tail current source (L3) has weak influence and is prunable. The other benchmarks also match design intuition. In the LDO, the pass device width (W_{p2})

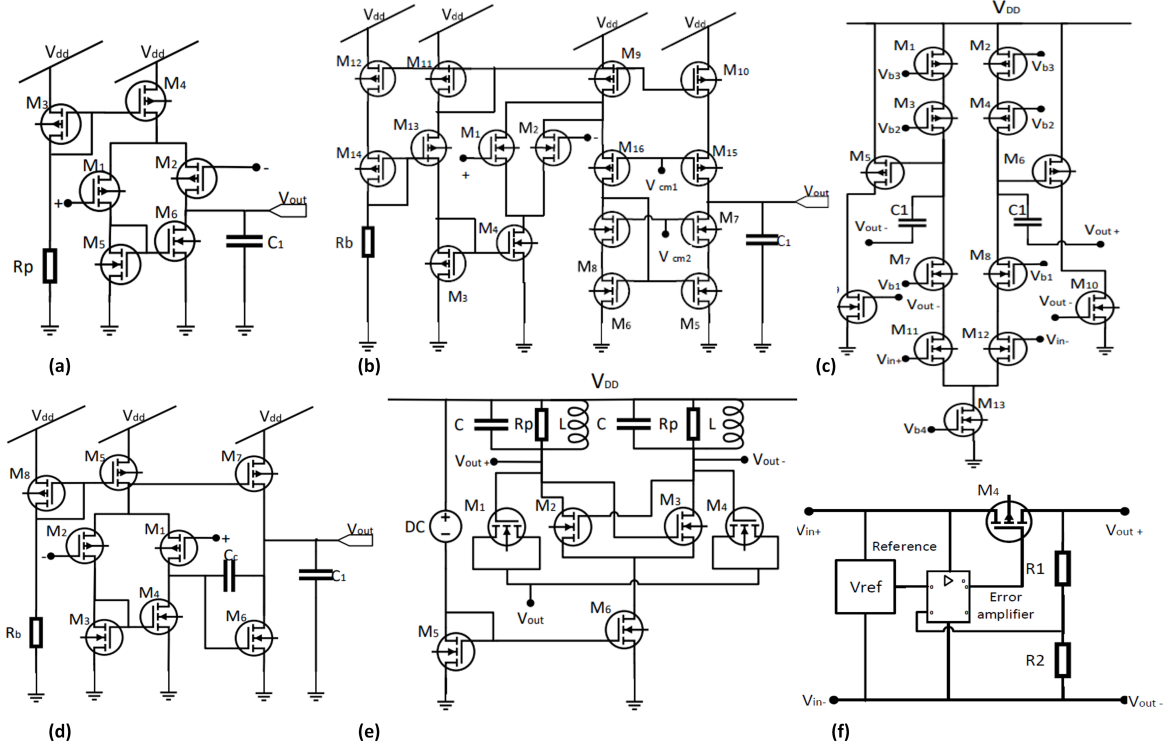


Fig. 2. **Schematics of benchmark circuits used to evaluate CaDRO:** (a) Active-Load Differential Amplifier, (b) Folded-Cascode Amplifier, (c) Two-Stage Voltage Amplifier, (d) Two-Stage Operational Transconductance Amplifier, (e) Voltage-Controlled Oscillator, and (f) Low-Dropout Regulator.

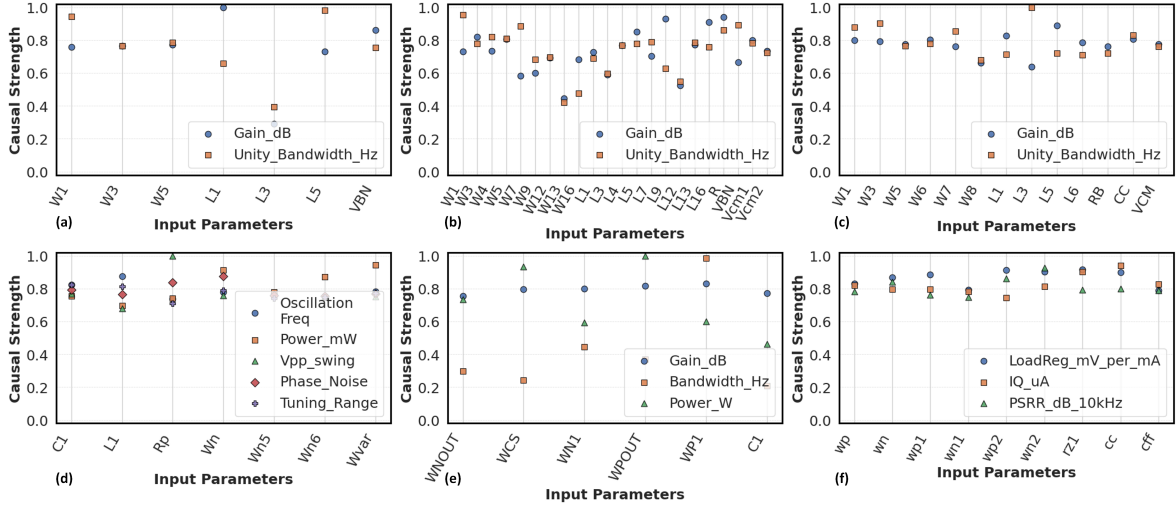


Fig. 3. **Causal strength of input parameters across benchmark circuits:** Each subplot shows normalized causal strengths from CaDRO’s observational–interventional analysis. (a) ALDA, (b) FCA, (c) TSOTA, (d) VCO, (e) TSVA, and (f) LDO. Markers denote circuit objectives (e.g., Gain, UGBW, Power, PSRR). The profiles highlight dominant drivers and low-impact parameters that can be pruned safely.

strongly impacts both LoadReg and PSRR, consistent with its role in output regulation. In the VCO, the tank inductor (L_1) directly determines oscillation frequency. In TSVA, the input pair width (W_{P1}) drives first-stage transconductance and bandwidth, while the output device (W_{POUT}) primarily sets power and moderately increases bandwidth. These causal maps confirm known dependencies while identifying low-impact parameters.

C. Impact of Causal Pruning on Pareto Trade-Offs

Fig. 4 compares reduced-space CaDRO (red) against full-space NSGA-II (blue) for three amplifier circuits: TSOTA,

FCA, and ALDA. The top row shows Pareto front overlays in Gain vs. Unity-Gain Bandwidth (UGBW), while the bottom row reports best and mean values with percentage changes. These views show not only where CaDRO improves or shifts the fronts, but also *how causal pruning redistributes search effort across competing objectives*. In the Two-Stage OTA (TSOTA, Fig. 4a,d), the reduced-space front covers the trade-off curve more densely, reflecting more efficient exploration. The best-case gain drops (−24%), while mean and best UGBW values improve slightly (+2%). This indicates that CaDRO favours balanced designs with more reliable bandwidth at the

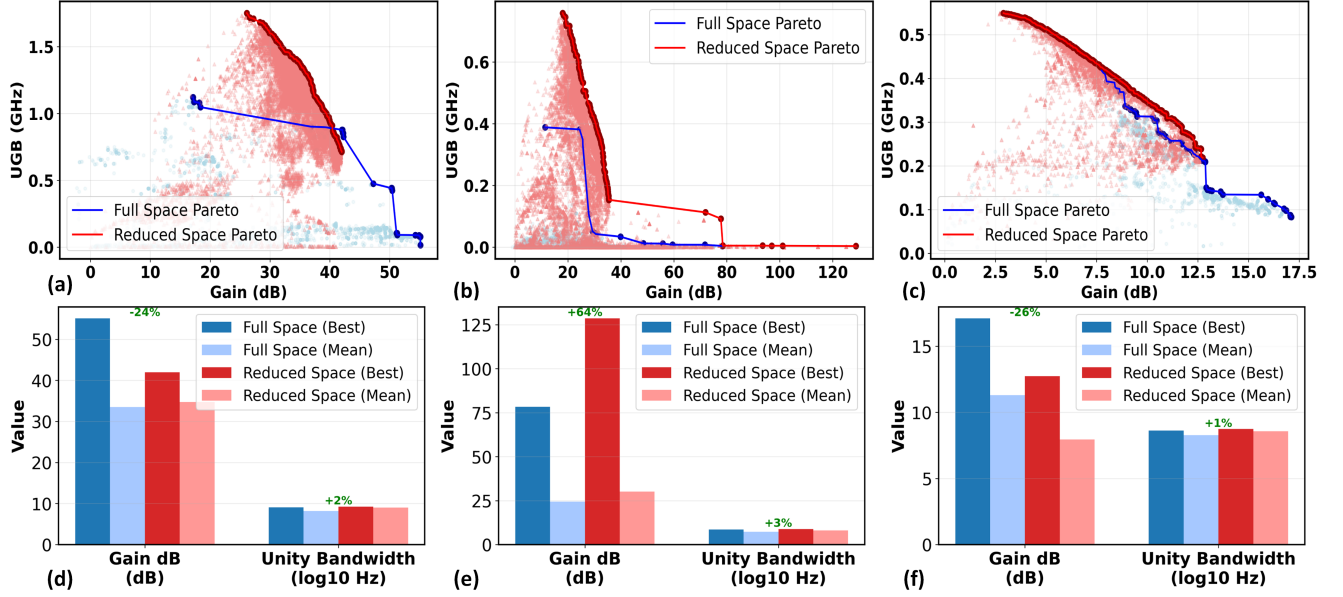


Fig. 4. **Amplifier benchmarks under CaDRO vs. full-space NSGA-II:** (a–c) Pareto fronts for the Two-Stage OTA (TSOTA), Folded-Cascode Amplifier (FCA), and Active-Load Differential Amplifier (ALDA), comparing reduced-space CaDRO (red) against full-space NSGA-II (blue). CaDRO consistently produces outward-shifted and denser Pareto fronts. (d–f) Best and mean values of Gain and Unity-Gain Bandwidth from the corresponding fronts. In TSOTA (d), CaDRO sacrifices some gain (–24%) but slightly improves UGBW (+2%). In FCA (e), CaDRO achieves a large gain improvement (+64%) together with higher UGBW (+3%). In ALDA (f), gain decreases (–26%) while UGBW is preserved (+1%). These plots show that causal pruning redistributes search effort: in circuits where causal drivers strongly govern both objectives (FCA), fronts expand dramatically, while in others (TSOTA, ALDA), fronts densify and trade-offs shift.

cost of extreme gain, often preferable in practical analog design where stability margins matter.

For the Folded-Cascode Amplifier (FCA, Fig. 4b,e), CaDRO achieves the most dramatic improvement: the Pareto front shifts outward along both axes, hypervolume increases from 0.56 to 0.94, and best-case gain improves substantially (+64%) while UGBW also rises (+3%). The denser frontier shows that pruning low-impact parameters prevents wasted evaluations and concentrates sampling where true trade-offs exist. In the Active-Load Differential Amplifier (ALDA, Fig. 4c,f), CaDRO maintains full frontier coverage but with a noticeable trade-off shift: gain decreases (–26%) while UGBW is essentially preserved (+1%). The result is a denser but lower-gain frontier, showing that in smaller circuits with fewer strong causal drivers, pruning can emphasize bandwidth consistency at the expense of peak gain. Importantly, mean performance remains competitive, indicating that CaDRO avoids collapsing the search space.

Fig. 5 extends this comparison to the Voltage-Controlled Oscillator (VCO) and Two-Stage Voltage Amplifier (TSVA). In the VCO (Fig. 5a–c), CaDRO preserves the full set of trade-offs in oscillation frequency, power, phase noise, and tuning range. The close overlap of the red and blue fronts shows that pruning weakly causal parameters does not distort achievable performance, while efficiency improves substantially; fronts are reached with less than half the evaluations (1160 → 478). For the TSVA (Fig. 5d–f), the impact is even clearer. Causal analysis revealed that only the input pair and compensation capacitor drive behavior. Once the remaining parameters are fixed, CaDRO reproduces the same gain, bandwidth, and power trade-offs as the baseline, but with an order-of-magnitude fewer evaluations (49 → 4).

D. Quantitative Benchmarking

Table I reports a detailed comparison of CaDRO and full-space NSGA-II across amplifiers and the LDO regulator. Each row lists the circuit, method, and simulation budget, followed by metrics: hypervolume (overall volume of the dominated region, ↑), Generational Distance (GD, average distance from obtained solutions to the true front, ↓), Inverted Generational Distance (IGD, average distance from reference Pareto points to the obtained set, ↓), additive ε indicator (worst-case dominance gap, ↓), spacing S (distribution uniformity, ↓), coverage (fraction of baseline solutions dominated, ↑), and cardinality (size of the non-dominated set) [30]–[32].

In the amplifier benchmarks, CaDRO consistently improves hypervolume while lowering GD and IGD. For the Folded-Cascode Amplifier (FCA), hypervolume rises from 0.56 to 0.94, and the non-dominated set expands from 5 to 48, reflecting both better coverage and richer sampling. In the Two-Stage OTA (TSOTA), hypervolume improves from 0.48 to 0.61, while spacing shrinks (1.25 → 0.77), yielding a more uniform distribution of solutions. The Active-Load Differential Amplifier (ALDA) highlights another strength: although hypervolume increases only modestly (0.29 → 0.31), the number of non-dominated solutions grows nearly fivefold (85 → 410).

The Low-Dropout Regulator (LDO) provides perhaps the clearest evidence of causal pruning’s effect. Hypervolume increases from 0.65 to 0.81, GD is halved (0.125 → 0.061), and coverage rises from 0.13 to 0.61, while spacing decreases relative to baseline. These quantitative gains mirror the visual expansions seen earlier: CaDRO consistently reallocates simulation effort away from low-impact parameters and into

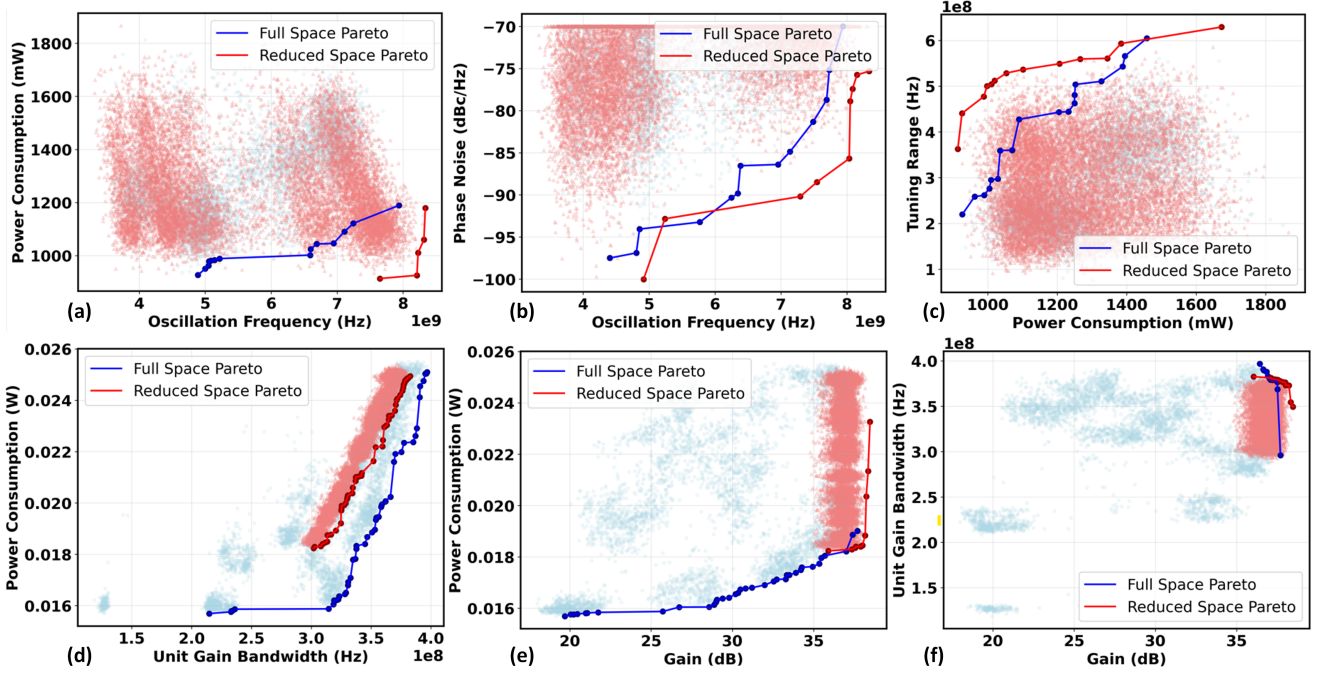


Fig. 5. **Oscillator and Regulator benchmarks under CaDRO vs. full-space NSGA-II:** (a–c) Voltage-Controlled Oscillator (VCO) and (d–f) Two-Stage Voltage Amplifier (TSVA). Reduced-space CaDRO (red) achieves fronts nearly identical to full-space NSGA-II (blue) but with far fewer evaluations. In the VCO, CaDRO maintains trade-offs across oscillation frequency, power, phase noise, and tuning range, while avoiding wasted exploration of weakly causal parameters. In the TSVA, fronts in gain, bandwidth, and power consumption closely track the baseline yet are reached with an order-of-magnitude fewer simulations.

Table I: Quantitative comparison of CaDRO and full-space NSGA-II across amplifier and regulator benchmarks.

Circuit	Method (Sim. count(%))	Hypervol. ($\uparrow$)	GD ($\downarrow$)	IGD ($\downarrow$)	Additive ε ($\downarrow$)	Spacing S ($\downarrow$)	Coverage	Cardinality
FCA (2D)	Reduced Space (100%)	0.94	0.44	0.41	0.72	1.35	1.00	48.00
	Reduced Space (80%)	0.34	0.03	0.66	0.98	0.87	0.20	8.00
	Full Space (NSGA-II)	0.56	0.14	0.33	0.63	1.32	0.02	5.00
ALDA (2D)	Reduced Space(100%)	0.31	0.02	0.09	0.56	0.71	0.76	410.00
	Reduced Space (80%)	0.14	0.04	0.49	0.75	0.74	0.09	176.00
	Full Space (NSGA-II)	0.29	0.06	0.14	0.36	0.94	0.00	85.00
TSOTA (2D)	Reduced Space (100%)	0.61	0.06	0.13	0.47	0.77	0.26	231.00
	Reduced Space (80%)	0.24	0.28	0.37	0.55	0.97	0.01	192.00
	Full Space (NSGA-II)	0.48	0.08	0.15	0.26	1.25	0.21	31.00
LDO (3D)	Reduced (100%)	0.81	0.06	0.38	0.60	1.45	0.61	602.00
	Reduced (80%)	0.76	0.05	0.49	0.79	1.37	0.45	982.00
	Full Space (NSGA-II)	0.65	0.13	0.57	0.83	1.59	0.13	164.00

Table II: Comparison of CaDRO and full-space NSGA-II for RF circuits (VCO and TSVA) across convergence and diversity metrics.

Circuit	Method	GD	IGD	S	Delta	MS	CP
VCO	Reduced Space	0.11	0.16	0.02	0.39	0.75	478
	Full Space	0.10	0.07	0.01	0.42	0.92	1160
TSVA	Reduced Space	0.03	0.25	0.08	0.88	0.25	4
	Full Space	0.0298	0.11	0.01	0.99	0.88	49

regions of genuine trade-off, producing higher-quality and more comprehensive Pareto sets at lower computational cost.

In Table II, for the Voltage-Controlled Oscillator (VCO), CaDRO achieves nearly the same GD and IGD as the full-space baseline (0.11 vs. 0.10, 0.16 vs. 0.07) but with less than half the evaluations (1160 $\rightarrow$ 478). Spacing (S) and Delta (Δ), which measure uniformity and diversity, remain close to baseline, while Maximum Spread (MS) is slightly reduced, indicating the front is preserved though sampled more

compactly. Computational Cost (CP) highlights the efficiency gain directly. In the Two-Stage Voltage Amplifier (TSVA), the effect is sharper: simulations collapse from 49 to 4, yet GD and IGD remain competitive (0.03 vs. 0.0298, 0.25 vs. 0.11). Here, S and Δ increase modestly, reflecting thinner sampling, but MS and overall coverage remain intact. The near-identical Pareto surface confirms that CaDRO isolates the true performance drivers while eliminating wasted evaluations.

V. CONCLUSION

We presented CaDRO, a causal-guided dimensionality reduction framework for scalable multi-objective optimization of analog and RF circuits. By combining observational-interventional causal discovery with evolutionary search, CaDRO identifies true design drivers, prunes low-impact parameters, and anchors optimization in high-performance re-

gions. Results across amplifiers, regulators, and oscillators show that CaDRO converges up to $10\times$ faster than NSGA-II, consistently improves or preserves Pareto quality, and yields denser, more interpretable fronts.

REFERENCES

- [1] E. Tlelo-Cuautle, M. A. Duarte-Villaseñor, and I. Guerra-Gómez, “Evolutionary algorithms in the optimal sizing of analog circuits,” in *Intelligent Computational Optimization in Engineering*, P. Melin, O. Castillo, E. G. Ramírez, and J. Kacprzyk, Eds. Springer, 2010, pp. 205–231.
- [2] R. Rashid, G. Raghunath, V. Badugu, and N. Nambath, “Performance evaluation of evolutionary algorithms for analog integrated circuit design optimisation,” *Microelectronics Journal*, vol. 141, p. 105983, 2023.
- [3] J. Gu, H. Wang *et al.*, “tss-bo: Truncated subspace sampling for high-dimensional bayesian optimization in analog circuit sizing,” in *Design, Automation & Test in Europe (DATE)*, 2024.
- [4] S. Kochar, A. Balaji *et al.*, “Ledro: Llm-enhanced design space reduction and optimization for analog circuits,” *arXiv preprint arXiv:2411.12930*, 2024. [Online]. Available: <https://arxiv.org/abs/2411.12930>
- [5] W. Shi, H. Wang, J. Gu, M. Liu, D. Z. Pan, S. Han, and N. Sun, “Robustanalog: Fast variation-aware analog circuit design via multi-task reinforcement learning,” *arXiv preprint arXiv:2207.06412*, 2022. [Online]. Available: <https://arxiv.org/abs/2207.06412>
- [6] Z. Cao, R. Yan *et al.*, “Rose-opt: Knowledge-infused bayesian optimization and reinforcement learning for analog circuit design,” *arXiv preprint arXiv:2407.19150*, 2024. [Online]. Available: <https://arxiv.org/abs/2407.19150>
- [7] M. B. Alawieh, F. Wang, R. Kanj, X. Li, and R. Joshi, “Efficient analog circuit optimization using sparse regression and error margining,” in *2016 IEEE/ACM International Symposium on Quality Electronic Design (ISQED)*, 2016, pp. 410–417.
- [8] S. Zhang, F. Yang, C. Yan, D. Zhou, and X. Zeng, “Lineasybo: Scalable bayesian optimization approach for analog circuit synthesis via one-dimensional subspaces,” *arXiv preprint arXiv:2109.00617*, 2021. [Online]. Available: <https://arxiv.org/abs/2109.00617>
- [9] B. Schölkopf, “Causality for machine learning,” in *Probabilistic and causal inference: The works of Judea Pearl*, 2022, pp. 765–804.
- [10] L. Klein, M. El-Assady, and P. F. Jäger, “From correlation to causation: Formalizing interpretable machine learning as a statistical process,” *arXiv preprint arXiv:2207.04969*, 2022.
- [11] P. Spirtes and K. Zhang, “Causal discovery and inference: concepts and recent methodological advances,” in *Applied informatics*, vol. 3, no. 1. Springer, 2016, p. 3.
- [12] P. Spirtes, C. Glymour, and R. Scheines, *Causation, Prediction, and Search*, 2nd ed. MIT Press, 2000.
- [13] J. Zhang, “On the completeness of orientation rules for causal discovery in the presence of latent confounders and selection bias,” *Artificial Intelligence*, vol. 172, no. 16–17, pp. 1873–1896, 2008.
- [14] D. M. Chickering, “Optimal structure identification with greedy search,” *Journal of Machine Learning Research*, vol. 3, no. Nov, pp. 507–554, 2002.
- [15] S. Shimizu, P. O. Hoyer, A. Hyvärinen, and A. Kerminen, “A linear non-gaussian acyclic model for causal discovery,” *Journal of Machine Learning Research*, vol. 7, pp. 2003–2030, 2006.
- [16] X. Zheng, B. Aragam, P. Ravikumar, and E. P. Xing, “Dags with no tears: Continuous optimization for structure learning,” in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 31, 2018.
- [17] J. Pearl, *Causality: Models, Reasoning and Inference*, 2nd ed. Cambridge University Press, 2009.
- [18] F. Jiao, “Analog circuit design knowledge mining and circuit causal information modeling,” Ph.D. dissertation, State University of New York at Stony Brook, 2016.
- [19] 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 *International conference on machine learning*. PMLR, 2018, pp. 3306–3314.
- [20] W. Shi, H. Wang, J. Gu, M. Liu, D. Z. Pan, S. Han, and N. Sun, “Robustanalog: Fast variation-aware analog circuit design via multi-task rl,” in *Proceedings of the 2022 ACM/IEEE Workshop on Machine Learning for CAD*, 2022, pp. 35–41.
- [21] Z. Xu, Z. Zhao, and J. Liu, “Deterministic multi-objective optimization of analog circuits,” *Electronics*, vol. 13, no. 13, p. 2510, 2024.
- [22] B. Liu, Y. Wang, Z. Yu, L. Liu, M. Li, Z. Wang, J. Lu, and F. V. Fernández, “Analog circuit optimization system based on hybrid evolutionary algorithms,” *Integration*, vol. 42, no. 2, pp. 137–148, 2009.
- [23] 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.
- [24] K. Deb and H. Jain, “An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part i: solving problems with box constraints,” *IEEE transactions on evolutionary computation*, vol. 18, no. 4, pp. 577–601, 2013.
- [25] Q. Zhang and H. Li, “Moea/d: A multiobjective evolutionary algorithm based on decomposition,” *IEEE Transactions on evolutionary computation*, vol. 11, no. 6, pp. 712–731, 2007.
- [26] G. R. Salgado and C. A. Reyes-Garcia, “Applications of evolutionary algorithms in the design automation of analog integrated circuits,” *J. Appl. Sci.*, vol. 10, no. 17, pp. 1859–1872, 2010.
- [27] B. Varıcı, D. Katz, D. Wei, P. Sattigeri, and A. Tajer, “Interventional causal discovery in a mixture of dags,” *Advances in Neural Information Processing Systems*, vol. 37, pp. 86 574–86 601, 2024.
- [28] A. Hägele, J. Rothfuss, L. Lorch, V. R. Somnath, B. Schölkopf, and A. Krause, “Bacadi: Bayesian causal discovery with unknown interventions,” in *International Conference on Artificial Intelligence and Statistics*. PMLR, 2023, pp. 1411–1436.
- [29] A. Li, A. Jaber, and E. Bareinboim, “Causal discovery from observational and interventional data across multiple environments,” *Advances in Neural Information Processing Systems*, vol. 36, pp. 16 942–16 956, 2023.
- [30] H. Ishibuchi, R. Imada, Y. Setoguchi, and Y. Nojima, “Reference point specification in inverted generational distance for triangular linear pareto front,” *IEEE Transactions on Evolutionary Computation*, vol. 22, no. 6, pp. 961–975, 2018.
- [31] H. Ishibuchi, H. Masuda, Y. Tanigaki, and Y. Nojima, “Modified distance calculation in generational distance and inverted generational distance,” in *International conference on evolutionary multi-criterion optimization*. Springer, 2015, pp. 110–125.
- [32] A. Panichella, “An improved pareto front modeling algorithm for large-scale many-objective optimization,” in *Proceedings of the genetic and evolutionary computation conference*, 2022, pp. 565–573.