H-OSIRIS: HIERARCHICAL LAYOUT-AWARE ANALOG CIRCUIT SIZING AND FOOTPRINT OPTIMIZATION USING DEEP REINFORCEMENT LEARNING

Anonymous authors
Paper under double-blind review

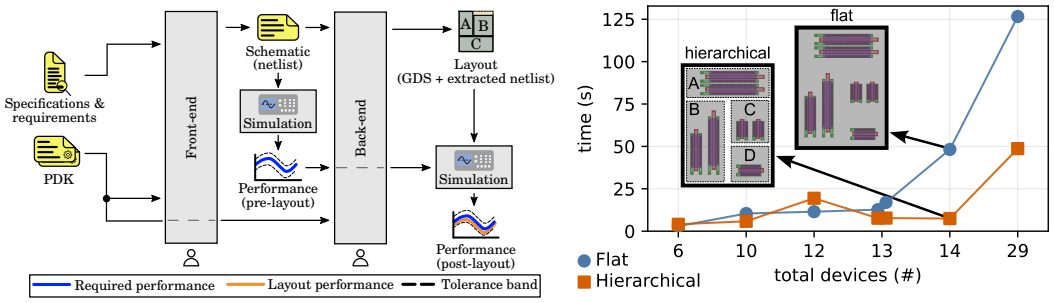
ABSTRACT

The automation of analog integrated-circuit layout remains challenging because circuit performance depends on coupled interaction among device sizing, parasitics, layout-dependent effects, and physical area. These effects are difficult to model accurately with rule-based or template-driven methods, particularly as circuits grow from individual functional blocks into multi-block systems. While machine learning has shown promising results for isolated stages of the analog design flow, an automated back-end optimization methodology that exploits the multi-block hierarchy of a layout while leveraging parasitic-aware feedback remains largely unexplored. This paper presents H-OSIRIS, a deep reinforcement learning methodology for post-layout optimization of multi-block analog circuits. H-OSIRIS operates at two optimization levels *(i)* it allocates area across circuit blocks globally and *(ii)* performs local ϵ-sizing of block devices. Guided by post-layout feedback, these mechanisms align post-layout functional metrics with their pre-layout references while reducing the overall layout footprint. Across four analog topologies and six metric sets, H-OSIRIS achieves 87% joint feasibility on held-out test samples while reducing the median layout area by 5.5%.

1 INTRODUCTION

Analog integrated circuits (ICs) are critical in a wide range of modern electronics applications, including telecommunications, power electronics, audio engineering, biomedical engineering, and instrumentation. They enable precise sensing, amplification, and filtering, serving as a critical bridge between digital systems and the physical environment. This capability ensures robust data acquisition and signal processing across various applications. However, unlike digital ICs, which benefit from highly automated and increasingly ML-driven design flows Razavi (2005); Carusone et al. (2011); Cui et al. (2024); Fang et al. (2025); Liu et al. (2025a); Zhao et al. (2025); Liu et al. (2025b), analog circuit design remains largely manual due to *(i)* complexity, unlike digital design based on simple logic gates, analog circuits use diverse components (transistors, resistors, capacitors) with intricate interconnections, *(ii)* sensitivity, small changes can drastically affect functionality, leading to a vast search space. In particular, the analog design flow is divided into front- and back-end phases. Figure 1a reports an overview of the analog design flow. The front-end phase targets schematic-level optimization, including component parameter tuning (e.g., transistor sizing) and topology selection. The back-end phase tackles the complex task of translating the optimized schematic into a manufacturable physical layout, encompassing layout generation and feasibility verification. In the front-end phase, designers determine the circuit's functional structure by selecting a topology and sizing devices to meet electrical specifications under idealized or schematic-level models. These tasks operate on abstract representations. The back-end phase then converts the validated schematic into a manufacturable physical layout. This requires solving placement and routing under strict geometric constraints, where small spatial modifications can alter matching, introduce parasitics, or break critical symmetries. Designers must manage layout-dependent effects, process variations, and wiring-induced parasitics, ensuring that the final layout passes DRC, matches the schematic through LVS, and undergoes parasitic extraction followed by post-layout validation. Because analog performance is extremely sensitive to layout geometry, this stage is iterative and expert-driven,

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.



(a) Overview of the analog design flow. Design specifications and target PDK are first translated into a Schematic that captures the required functionality. The Schematic is then realized as a Layout through the back-end stage. The layout must meet the same specifications as the schematic.

(b) Time (seconds) to generate a layout from a flat or hierarchical netlist as the number of devices increases. Flat netlists are unpartitioned, whereas the hierarchical representation preserves its decomposition into blocks.

Figure 1: (a) overview of the analog design flow and (b) netlist structure comparison.

constituting a major bottleneck. Recent research has predominantly focused on front-end analog design automation, including topology generation and sizing with GNNs Dong et al. (2023), LLMs Lai et al. (2025a;b); Chang et al. (2024); Wei et al. (2025), RL Vijayaraghavan et al., and decoder-only transformers Gao et al. (2025b). For example, AnalogCoder Lai et al. (2025a;b) uses prompt engineering and feedback optimization to generate PySpice code, whereas LaMAGIC Chang et al. (2024) fine-tunes Flan-T5 to map textual specifications directly to optimized designs. More recent GPT-based approaches, including AnalogGenie Gao et al. (2025a) and AnalogFed Li et al. (2025), model circuit connectivity at the device-pin level and support topology discovery from distributed private datasets. In contrast, back-end automation remains comparatively underexplored. Parasitics, technology-specific constraints, and manufacturability requirements strongly affect post-layout performance Noori Zadeh & Elamien (2025). Existing learning-based approaches address individual back-end tasks, including placement Li et al. (2020); Ahmadi & Zhang (2021), routing Zhu et al. (2019); Chen et al. (2023), and pre-layout performance prediction Ren et al. (2020). However, analog layout optimization remains largely human-centric, and the tight coupling between front-end design choices and post-layout behavior is often overlooked. Addressing this coupling requires an automated optimization process that uses physical-layout feedback throughout design refinement, making the cost of automatic layout generation an important consideration. Many analog circuits are naturally organized as interconnected blocks, each serving a distinct role in the overall design. A hierarchical circuit representation exploits this organization, reducing the computational burden of layout generation (see Figure 1b) while providing an optimization structure aligned with the circuit's composition. A flat representation treats the complete circuit as a single, unpartitioned layout and optimization problem. In contrast, a hierarchical representation preserves the decomposition into subcircuits, each optimized within its local context. This decomposition enables local decisions to refine individual subcircuits, while global decisions coordinate the circuit's overall requirements. This work introduces H-OSIRIS, a novel methodology that bridges the gap between phases of the analog design flow by enabling automated, iterative, and layout-aware hierarchical exploration driven by deep reinforcement learning (DRL), ultimately producing layouts whose post-layout performance meets pre-layout specifications while compacting layout footprint. H-OSIRIS provides three main contributions to the state of the art:

- **H-OSIRIS**, a hierarchical methodology that iteratively refines analog layouts using post-layout feedback. A circuit-level Manager handles area allocation across functional blocks, while a reconfigurable Worker performs local ϵ-sizing on functional blocks. Repeated layout generation and post-layout evaluation guide these coordinated global and local decisions toward quality metrics satisfaction and reduced total layout area.
- **Extensive evaluation** across four analog circuit topologies and six metric sets, each combining global area with two to six electrical metric assignments. Across 600 held-out episodes, H-OSIRIS achieves terminal joint feasibility in 88% of the cases, while reducing median layout area by 3-8% across the evaluated topology-metric sets. Held-out trajec-

Table 1: Comparison of analog design automation approaches with respect to their target tasks, support for block-level device sizing, global area control, and use of post-layout feedback.

Work	Target task	Block sizing	Global area control	Post-layout feedback
Budak et al. (2021)	Sizing	○	○	○
Gao et al. (2024b)	Sizing	○	○	●
Li & Carusone (2024)	Sizing	○	○	●
Gao et al. (2024a)	Placement	○	●	○
Basso et al. (2025)	Placement	○	●	○
Xu et al. (2026)	Placement	○	⊙	⊙
Mehradfar et al. (2026)	Sizing	○	⊙	⊙
H-OSIRIS	Post-layout sizing and area optimization	●	●	●

● Support, ⊙ Partial support, ○ No support.

tories and objective-conditioned action distributions further show adaptation to both the selected performance targets and different initial layouts.

- **A benchmark collection** (available in the Supplementary Material) for the held-out evaluation episodes, including initial and terminal schematic netlists, extracted netlists, GDS layouts, and calibrated baseline characterizations, enabling future methods to be evaluated from the same initial circuit states.

2 Related Work

The optimization of analog IC physical design is particularly challenging due to tight performance constraints, sensitivity to parasitics, and complex design rules that traditional techniques often fail to handle effectively, requiring time-consuming intervention of expert engineers.

Back-end. Recent research has begun to address the complexities of physical layout and verification. ACDC Esfandiari et al. (2025) frames transistor placement as a sequence-prediction problem, fine-tuning transformers to predict device coordinates from a netlist. SOLOMON Wen & Zhang (2025) adapts general-purpose LLMs to generate Python scripts for complex semiconductor-layout structures. Regarding verification, LLM-HD Chen et al. (2024) processes GDSII binaries to detect lithography hotspots, whereas DRC-Coder Chang et al. (2025) generates standard verification rule format (SVRF) code from textual design rules. These approaches target placement prediction or verification rather than iterative analog-layout optimization. Other work focuses specifically on analog layout automation. BAG Crossley et al. (2013); Chang et al. (2018) introduced parameterized generators for analog and mixed-signal circuits, whereas ALIGN Kunal et al. (2019) and MAGICAL Xu et al. (2019) provide algorithmic layout-synthesis flows. Recent learning-based methods target distinct optimization tasks. DNN-Opt Budak et al. (2021) performs schematic sizing, while Gao et al. (2024b); Li & Carusone (2024) extend sizing optimization to post-layout performance. At the placement level, Basso et al. (2025); Gao et al. (2024a); Xu et al. (2026) optimize layout geometry, area, routability, or predicted post-layout performance. FALCON Mehradfar et al. (2026) further combines topology selection and parameter inference with a differentiable layout cost. These methods address post-layout sizing, placement, or layout-constrained synthesis, but do not jointly combine block-local device sizing, explicit global footprint control, and extracted post-layout feedback. H-OSIRIS formulates this combination as a hierarchical decision process: a Manager controls the footprint boundaries of individual circuit blocks, while one reconfigurable Worker environment loads the selected block-specific policy bundle and performs local ϵ-sizing of its matched device groups. Each resulting design is evaluated through full-circuit parasitic extraction and post-layout simulation, enabling the learned policies to align circuit quality metrics with pre-layout references while reducing total layout area. Table 1 positions H-OSIRIS with respect to representative prior approaches.

3 The H-OSIRIS Methodology

H-OSIRIS adopts a hierarchical DRL framework with a circuit-level Manager and one reconfigurable Worker environment that loads the configuration bundle of the functional block selected by the Manager. Let $\mathcal{B}$ denote the set of functional blocks, with $b \in \mathcal{B}$, and let $t \in \{0, 1, \ldots\}$ index

Manager steps. For a given block, the bundle comprises its policy and target-network state, replay buffer, sizing design space, device matching groups, metric targets and done bounds, and reward parameters. The loaded state combines the current sizing and post-layout metrics of the selected block, its allocated area, and Manager-provided global context. The Worker then performs discrete ϵ-sizing exclusively within that functional block. The Manager represents the complete circuit as a graph of interconnected functional blocks, with node features encoding block constraints, post-layout performance, and area information. At each Manager step, it selects the next functional block to optimize and modifies that block's allocated geometric layout constraints. A two-block example of the Manager graph state and the state of the selected Worker is provided in Appendix A.1.2. For block b, $A_b^{(t)}$ is its realised layout area and $(W_{\max,b}^{(t)}, H_{\max,b}^{(t)})$ are its allowed footprint bounds. Metric set $\mathcal{Q}_b$ contains the metrics assigned to block b, whereas $\mathcal{Q}_{\text{glob}}$ contains circuit-level metrics.

3.1 Optimization Mechanisms

H-OSIRIS combines two complementary optimization mechanisms that operate at different levels of circuit hierarchy. A global mechanism adjusts the footprint boundaries of functional blocks, whereas a local mechanism performs fine-grained device ϵ-sizing within each block.

Global: area boundaries. The physical footprint available to each functional block is defined by maximum width and height constraints, $(W_{\max,b}^{(t)}, H_{\max,b}^{(t)})$. These constraints do not directly modify device dimensions. They specify the geometric region available to the place-and-route engine. Increasing or decreasing ($W_{\max,b}^{(t)}$ and $H_{\max,b}^{(t)}$) allows the methodology to explore geometric trade-offs among layout area and post-layout functional metrics without modifying device dimensions. Appendix A.1.1 reports an example of $W_{\max}$ variation.

Local: ϵ-sizing. Each device (i.e., resistor, capacitor, transistor) can undergo fine-grained adjustments to its dimensions within user-defined bounds. ϵ-sizing allows for alterations in width and length, as well as to the number of fingers. These controlled perturbations enable the exploration of local improvements to ensure that the layout matches original circuit specifications. Moreover, within circuit topologies, certain transistors must share matched sizing (e.g., current mirrors and differential pairs). These transistors are grouped and adjusted as a single unit.

3.2 Evaluation of Analog Layouts

H-OSIRIS assesses a layout on post-layout quality metrics and layout area.

Quality metrics. Analog circuits are characterized by a set of performance metrics that depend on the circuit class, including, for example, phase margin, common-mode rejection ratio, and open-loop gain for amplifiers, or passband gain and stopband attenuation for filters. These metrics collectively define the design specifications and serve as optimization objectives. In H-OSIRIS, their evaluation relies on post-layout, parasitic-aware simulations to accurately capture layout-induced effects.

Layout area. The area of a generated layout is computed as the smallest rectangle that contains all the placed primitives of the circuit, as defined in Equation 1:

$$A_{\text{layout}} = \big(\max_{p\in\{1,\dots,P\}} x_p^{ur} - \min_{p\in\{1,\dots,P\}} x_p^{ll}\big) \cdot \big(\max_{p\in\{1,\dots,P\}} y_p^{ur} - \min_{p\in\{1,\dots,P\}} y_p^{ll}\big) \tag{1}$$

where x^{ur} and x^{ll} are the x coordinates of the upper right and lower left corners, and y^{ur} and y^{ll} are the y coordinates of the upper right and lower left corners of each primitive, respectively. They are expressed in μm. P is the total number of devices. Layout area is computed from the GDSII file.

3.3 Reward Function

H-OSIRIS assigns rewards across two hierarchy levels: *(i)* a local reward to the functional block's network loaded in the Worker, and *(ii)* a global reward to the Manager, which alters block area boundaries. Although actions are local, every reward is evaluated after assembling, placing, extracting, and simulating the complete circuit. Local rewards therefore account for inter-block electrical and parasitic coupling. For a quality metric q_i with measured value v_i, target interval $[\ell_i^{\text{tar}}, u_i^{\text{tar}}]$, and

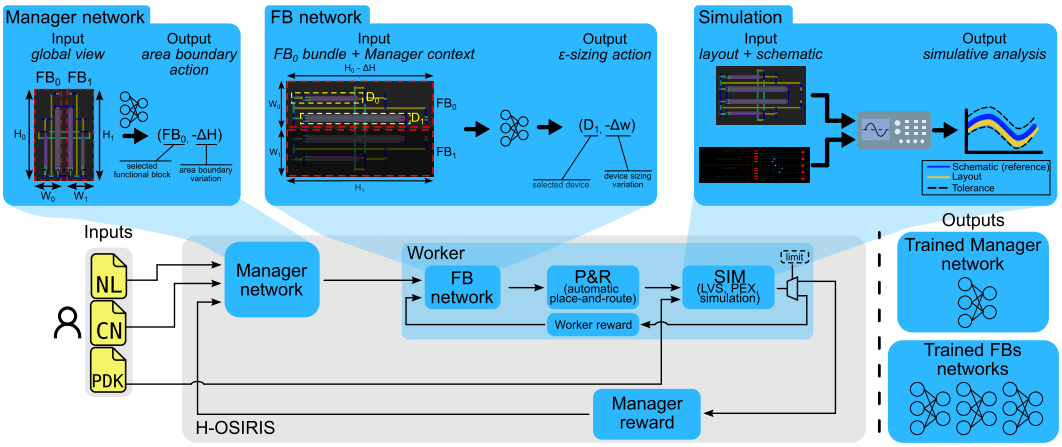


Figure 2: Overview of the H-OSIRIS training pipeline. At each Manager step, the Manager selects a functional block (FB) and updates its area boundary. The reconfigurable Worker environment then loads that block's policy, replay buffer, design constraints, metric configuration, and Manager context, and executes a local ϵ-sizing episode. Each Worker action is evaluated through place-and-route and post-layout simulation. Local transitions update the loaded block-specific bundle, whereas the terminal full-circuit state provides the Manager reward and transition.

scale τ_i, we define the target violation (e_i), normalized error ($\hat{e}_i$), and shaped metric reward (R_i) as:

$$e_i = \max(0, v_i - u_i^{\text{tar}}) + \max(0, \ell_i^{\text{tar}} - v_i), \qquad \hat{e}_i = \frac{e_i}{\tau_i}, \qquad R_i = \exp(-\hat{e}_i). \tag{2}$$

Bandwidth and cutoff-frequency violations use the same linear-domain definition and metric-specific scale in Equation 2. Let b denote the functional block currently loaded in the Worker. The Worker optimizes its assigned metric set $\mathcal{Q}_b$ and local layout area A_b. Its dense reward combines shaped metric rewards and a piecewise-linear area reward $R_{\text{area}}^{(b)}$:

$$r_{\text{dense},t}^{(b)} = \frac{\sum_{i\in\mathcal{Q}_b} \tilde{w}_i R_i + \tilde{w}_{\text{area}} R_{\text{area}}^{(b)}}{\sum_{i\in\mathcal{Q}_b} \tilde{w}_i + \tilde{w}_{\text{area}}}, \qquad r_t^{(b)} = r_{\text{dense},t}^{(b)} + B_b d_t^{(b)}. \tag{3}$$

Here, $\tilde{w}_i = w_i(1+\alpha\hat{e}_i)$ emphasizes the most violated objectives. Local area remains a soft objective, whereas global area is part of Manager terminal feasibility. The Manager jointly evaluates $\mathcal{Q}_{\text{M}} = \left(\bigcup_{b\in\mathcal{B}} \mathcal{Q}_b\right) \cup \mathcal{Q}_{\text{glob}}$ and receives:

$$r_t^{\text{M}} = \frac{\sum_{i\in\mathcal{Q}_{\text{M}}} R_i + R_{\text{area}}^{\text{glob}}}{|\mathcal{Q}_{\text{M}}| + 1} + B_M d_t^{\text{M}} + r_{\text{progress},t}^{\text{M}}. \tag{4}$$

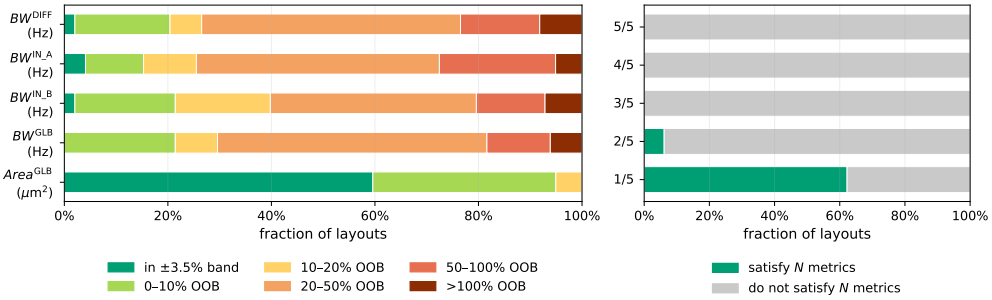
The first term rewards electrical and area feasibility, while $r_{\text{progress},t}^{\text{M}}$ rewards improvements in normalized electrical error and total area after each Manager action. Further details on area rewards, terminal indicators, and progress-reward definition are provided in Appendix A.1.3.

3.4 Training Pipeline

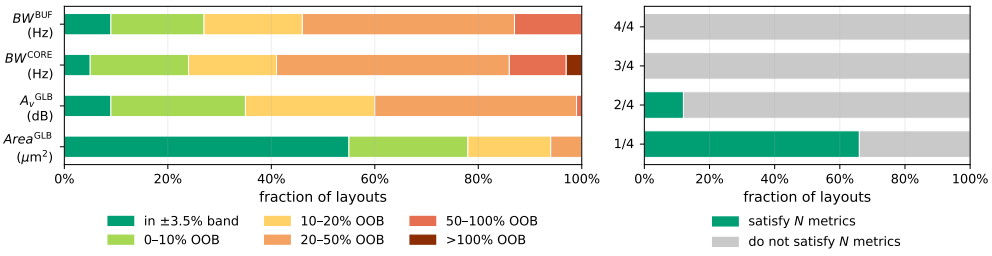
H-OSIRIS is a hierarchical DRL framework that closes the gap between schematic-level reference and post-layout circuit quality. As illustrated in Figure 2, the training pipeline comprises a circuit-level Manager and one reconfigurable Worker environment. The framework receives the schematic netlist (NL), design constraints (CN), and process-design-kit (PDK) information. CN specifies the admissible area-boundary variations, ϵ-sizing intervals, device matching groups, and metric target and done values. At each Manager step, the Manager observes the complete circuit layout, selects a functional block (FB_i), and applies an area-boundary action, such as $\Delta H_{i,t} = -\delta_H$ to decrement the height assigned to FB_i's area. The Worker then loads the persistent FB_i-specific bundle: its policy and target-network state, replay buffer, netlist, design-space constraints, metric configuration, and the current Manager context. The loaded Worker performs an ϵ-sizing episode on FB_i

Table 2: Functional-block partitioning and target quality metrics. Each row is one metric set. Each column identifies the metrics assigned to a functional block. GLB denotes global-circuit metrics.

Metric set	2-Stage Amp			3-OpAmp InAmp				2-OTA Comp.			Amp-Filter Comp.			
	BUF	CORE	GLB	DIFF	IN_A	IN_B	GLB	FT_0	FT_1	GLB	FT	HPF	MIL	GLB
$\mathcal{S}_\text{A}$	BW	BW	A_v, Area	–	–	–	–	–	–	–	–	–	–	–
$\mathcal{S}_\text{B}$	–	–	–	BW	BW	BW	BW, Area	–	–	–	–	–	–	–
$\mathcal{S}_\text{C}$	–	–	–	–	–	–	–	A_v	A_v	Area	–	–	–	–
$\mathcal{S}_\text{D}$	–	–	–	–	–	–	–	CMRR	CMRR, PSRR	Area	–	–	–	–
$\mathcal{S}_\text{E}$	–	–	–	–	–	–	–	CMRR, PSRR, A_v	CMRR, PSRR, A_v	Area	–	–	–	–
$\mathcal{S}_\text{F}$	–	–	–	–	–	–	–	–	–	–	A_v	f_c	A_v	Area



(a) 3-OpAmp InAmp, $\mathcal{S}_\text{B}$.



(b) 2-Stage Amp, $\mathcal{S}_\text{A}$.

Figure 3: Characterization of post-layout evaluations over 100 randomly sampled configurations within the design-space intervals. Left panels report, for each quality metric, the fraction of layouts within its $\pm 3.5\%$ done bound. Colors partition out-of-bound (OOB) layouts by normalized violation magnitude. Right panels report cumulative joint feasibility: each N/K row gives the fraction of layouts satisfying at least N of the K configured metrics, with K/K corresponding to joint feasibility.

only. All other functional blocks remain fixed. Each Worker action is evaluated through automatic place-and-route (P&R), LVS, parasitic extraction, and post-layout simulation (SIM). The resulting simulation analysis produces a local transition that is stored in the replay buffer of the currently loaded FB_i bundle. When the block episode terminates, the resulting full-circuit state produces the Manager transition and reward. The Worker then releases the runtime configuration. When the Manager selects FB_k, it loads the persistent FB_k bundle instead. Thus, one Worker environment executes all local episodes, while each functional block retains independent policy and replay state. At every Worker and Manager decision, action masks exclude updates that would violate the corresponding sizing or layout-allocation bounds. This ensures that exploratory actions remain within the configured discrete design space. The Manager and block-specific Worker policies are trained with independent Double-DQN updates. Worker policies use fully connected Q-networks, whereas the Manager uses a relational GNN. Further DDQN details are provided in Appendix A.1.4.

Table 3: Hierarchical training and held-out results. N_{train} is the number of training Manager episodes and N_{test} the number of held-out starts. Joint feasibility requires satisfying every configured electrical metric and the global area constraint. $\mathrm{SR}^{(0)}_{\mathrm{joint}}$, $\mathrm{SR}^{(T)}_{\mathrm{joint}}$, and $\mathrm{SR}_{\mathrm{rnd}}$ are initial, terminal, and random-policy joint success rates, respectively. $\tilde{N}_{\mathrm{M}}$ is the median number of Manager actions until feasibility and $\tilde{N}_{\mathrm{W}}$ the median executed Worker actions per Manager decision. $\widetilde{\rho}$ is the median terminal-to-initial area ratio. $\widetilde{d}_{\mathrm{M}}$ and $\widetilde{d}_{\mathrm{W}}$ are median normalized action-sequence edit distances for Manager and Worker actions.

Topology	Metric set	$\mathrm{SR}_{\mathrm{rnd}}$ [%]	N_{train}	N_{test}	$\mathrm{SR}^{(0)}_{\mathrm{joint}}$ [%]	$\mathrm{SR}^{(T)}_{\mathrm{joint}}$ [%]	$\tilde{N}_{\mathrm{M}}$	$\tilde{N}_{\mathrm{W}}$	$\widetilde{\rho}_{\mathrm{area}}$	$\widetilde{d}_{\mathrm{M}}$	$\widetilde{d}_{\mathrm{W}}$
2-Stage Amp	$\mathcal{S}_{\mathrm{A}}$	0	3 000	100	0	74	3.0	10.0	0.94	0.83	0.86
3-OpAmp InAmp	$\mathcal{S}_{\mathrm{B}}$	9	2 800	100	9	99	2.0	5.0	0.96	0.74	0.80
Dual-OTA Comp.	$\mathcal{S}_{\mathrm{C}}$	5	1 000	100	5	93	2.0	3.0	0.97	0.50	0.96
	$\mathcal{S}_{\mathrm{D}}$	2	1 000	100	2	92	2.0	2.0	0.94	0.50	0.90
	$\mathcal{S}_{\mathrm{E}}$	1	1 500	100	1	92	2.0	2.0	0.94	0.50	0.89
Amp-Filter Comp.	$\mathcal{S}_{\mathrm{F}}$	0	1 500	100	0	72	3.0	10.0	0.92	0.83	0.92

4 Experimental Evaluation

4.1 Experimental Setup

Hardware and software configuration. Experiments run on a Rocky Linux 9 machine with 256 GB RAM. H-OSIRIS combines the open-source layout generator blind (blind), Ngspice simulation, and Magic parasitic extraction to generate DRC-clean, LVS-compliant layouts. All experiments use the open-source SkyWater 130 nm PDK.

Hierarchical design space definition. When loaded with a selected functional-block configuration, the Worker applies ϵ-sizing to that block, locally adjusting matched transistor-group length, width, and finger count, as well as available resistor and capacitor parameters. Table 8 (Appendix A.3.2) reports the resulting discrete ranges, increments, action cardinalities, and block-level sizing spaces. The Manager operates on a separate discrete design space: it selects a functional block and changes its allocated width or height area boundary. Table 9 (Appendix A.3.2) reports these boundary ranges, increments, and joint allocation spaces.

Layout evaluation. Post-layout quality metrics and total area are evaluated as described in Section 3.2. The evaluated circuits comprise a two-stage Miller amplifier with source follower (2-Stage Amp), a three-op-amp instrumentation amplifier (3-OpAmp InAmp), a dual-OTA composite (Dual-OTA Comp.), and an amplifier-filter composite (Amp-Filter Comp.). Their functional blocks expose local metrics, while the 2-Stage Amp and 3-OpAmp InAmp additionally include global electrical metrics. Table 2 assigns each metric to its functional block or to global level (GLB). Depending on the experiment, the objectives include gain (A_v), bandwidth BW, CMRR, PSRR, and cutoff frequency (f_c), together with global layout area. Success requires satisfying jointly every configured metric done bound. Each policy is evaluated on 100 sizings held-out from training.

Target and done bands. For each electrical metric, the target interval used by reward shaping is a $\pm 1\%$ band around its pre-layout reference value, whereas the done interval used for terminal feasibility is a $\pm 3.5\%$ band. Thus, for a reference value v_i^{ref}, a two-sided metric is within the target or done band when $|v_i - v_i^{\mathrm{ref}}| \leq \eta |v_i^{\mathrm{ref}}|$, with $\eta = 0.01$ or $\eta = 0.035$, respectively.

4.2 Experimental Results

Initial post-layout gap. A layout does not generally preserve the pre-layout reference behavior of its schematic sizing because of layout-dependent parasitic effects. Figure 3 reports the evaluation of extracted layouts obtained from 100 randomly sampled sizing configurations for the 2-Stage Amp and 3-OpAmp InAmp. These random configurations are used only to characterize the initial post-layout gap; they are distinct from the held-out initial states used to compute $\mathrm{SR}^{(0)}_{\mathrm{joint}}$ in Table 3. In each left panel, a bar reports the fraction of layouts within the $\pm 3.5\%$ done band or at increasing out-of-band (OOB) distances from its reference. For the 2-Stage Amp, only 5 and 9 initial layouts satisfy the CORE BW and GLB A_v bands, respectively, while the BUF BW band is satisfied by 9

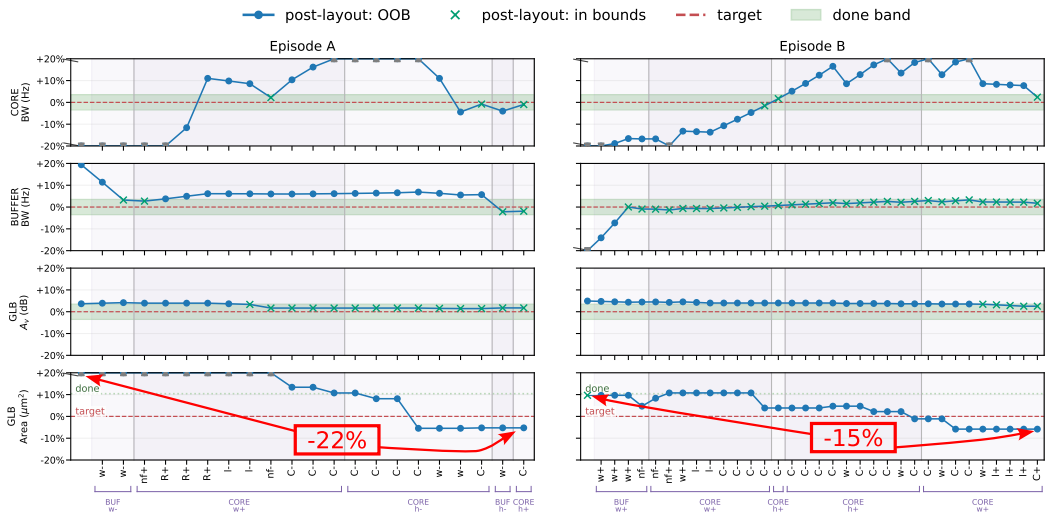


Figure 4: Held-out post-layout trajectories from two distinct initial states of 2-Stage Amp, S_A. Worker actions are grouped according to Manager decisions. Despite different initial states, the learned policies adapt the action choices to drive all quality metrics within the $\pm 3.5\%$ band while reducing layout area by 22% and 15%, respectively.

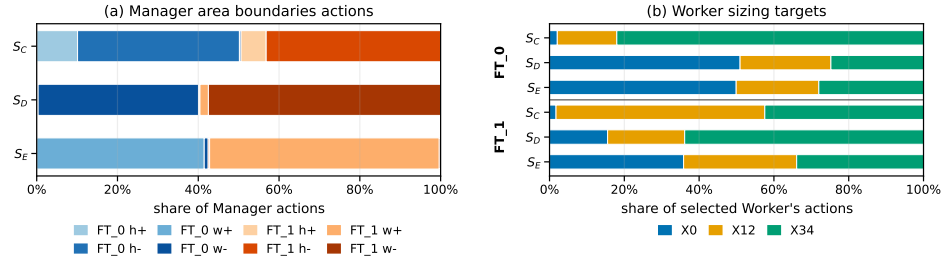


Figure 5: Objective-conditioned action distributions for the Dual-OTA Comp. With the topology and action space fixed, (a) the Manager and (b) dynamically loaded Worker configurations produce different action distributions for the target metric sets ($\mathcal{S}_{\mathrm{C}}$, $\mathcal{S}_{\mathrm{D}}$, and $\mathcal{S}_{\mathrm{E}}$). Legends report area-boundary perturbations and selected device groups in the Manager and Worker panels, respectively.

samples. Similarly, no InAmp initial layout satisfies the global BW band, and each local BW band is satisfied by only 2-4 samples. The right panels report cumulative joint satisfaction. For each N/K row, the green portion gives the fraction of layouts satisfying at least N of the K configured metrics, whereas the grey portion gives the complement. Thus, the K/K row directly reports joint feasibility. Although 55% of the 2-Stage and 59% of the InAmp configurations in this characterization sample already meet the area bound, none satisfies all configured metrics jointly. Optimization is therefore needed to correct coupled electrical deviations while retaining area feasibility.

Training and held-out testing. Table 3 reports the success rate of a random hierarchical policy ($\mathrm{SR}_{\mathrm{rnd}}$) evaluated from the same held-out starts and under the same maximum decision horizon, sizing intervals, and metric constraints, showing that unguided action selection does not yield terminal joint feasibility. Table 3 also summarizes the hierarchical training runs and held-out outcomes across the 2-Stage Amp, 3-OpAmp InAmp, Dual-OTA Comp., and Amp-Filter Comp. Besides held-out initial ($\mathrm{SR}^{(0)}_{\mathrm{joint}}$) and terminal ($\mathrm{SR}^{(T)}_{\mathrm{joint}}$) joint success, the table reports the median number of Manager decisions to feasibility ($\tilde{N}_{\mathrm{M}}$), the median number of Worker actions per Manager decision ($\tilde{N}_{\mathrm{W}}$), the median initial-to-terminal area ratio ($\tilde{\rho}_{\mathrm{area}}$), and the Manager and Worker sequence-diversity measures ($\tilde{d}_{\mathrm{M}}$ and $\tilde{d}_{\mathrm{W}}$). The latter are computed with median pairwise normalized Levenshtein distances. For two sequences $\mathbf{a}$ and $\mathbf{a}'$: $d(\mathbf{a}, \mathbf{a}') = \frac{\mathrm{Lev}(\mathbf{a},\mathbf{a}')}{\max\{|\mathbf{a}|,|\mathbf{a}'|\}}$. Thus, d=0 denotes

Table 4: Hierarchical and flat optimization comparison. N_{train} is the number of completed Manager training episodes. $T_{\text{train}}^{(\text{hier})}$, $T_{\text{train}}^{(\text{flat})}$ and $S_{\text{joint}}^{(T,H)}$, $S_{\text{joint}}^{(T,F)}$ are the training times and success rates for hierarchical and flat optimization, respectively.

Topology	Metrics set	N_{train} (#)	$T_{\text{train}}^{(\text{hier})}$ (h)	$T_{\text{train}}^{(\text{flat})}$ (h)	$S_{\text{joint}}^{(T,H)}$ (%)	$S_{\text{joint}}^{(T,F)}$ (%)
2-Stage Amp	$\mathcal{S}_{\text{A}}$	3 000	61	102	74	10
3-OpAmp InAmp	$\mathcal{S}_{\text{B}}$	2 800	54	465	99	40
Dual-OTA Comp.	$\mathcal{S}_{\text{C}}$	1 000	10	51	93	34
	$\mathcal{S}_{\text{D}}$	1 000	10	56	92	6
	$\mathcal{S}_{\text{E}}$	1 500	12	102	92	0
Amp-Filter Comp.	$\mathcal{S}_{\text{F}}$	1 500	29	279	72	0

identical sequences, whereas values closer to 1 denote more dissimilar action sequences. Notably, $\tilde{N}_{\text{M}}$, $\tilde{N}_{\text{W}}$, $\widetilde{\rho}_{\text{area}}$, $\widetilde{d}_{\text{M}}$, and $\widetilde{d}_{\text{W}}$ refer to held-out test episodes. Together, these quantities highlight the fact that these learned policies reach feasibility by adapting their hierarchical decisions to different initial layouts and objective sets without collapsing to the same repeated sequences of actions. The learned policies raise joint feasibility from 0-9% at initialization to 72-99% at termination. They reach feasibility after a median of 2-3 Manager decisions, with 2-10 Worker actions per Manager decision, while reducing median layout area by 3-8%. The Manager and Worker diversity values, $\widetilde{d}_{\text{M}} = 0.50$-$0.83$ and $\widetilde{d}_{\text{W}} = 0.80$-$0.96$, further indicate that the policies do not collapse to a single repeated action sequence. Appendix A.4.2 reports initial-to-terminal measurements.

Hierarchical versus flat optimization. The flat baseline matches H-OSIRIS in circuit topology, constraints, number of training episodes, held-out starts, and evaluation budget. The only difference is the representation: the flat formulation treats the complete circuit as one optimization unit, with a single-node Manager graph and one circuit-wide Worker configuration. Table 4 confirms this advantage across all six evaluated topology-metric set pairs. Hierarchical training requires 10-61 h, compared with 51-465 h for flat optimization, corresponding to reductions of $1.7\times$ to $9.6\times$. Moreover, H-OSIRIS achieves 72-99% terminal joint feasibility, whereas the flat baseline reaches 0-40%. These results indicate that factorizing the search into block-level sizing and circuit-level area allocation improves both optimization efficiency and terminal post-layout feasibility.

Policy behavior. Figure 4 reports two held-out 2-Stage Amp $\mathcal{S}_{\text{A}}$ trajectories, denoted Episode A and B. The starts lie in different infeasible regions and consequently induce different Manager block selections and Worker sizing actions. Episode A applies 21 Worker actions and reduces total area by 22%, whereas Episode B applies 33 Worker actions and reduces area by 15%. Both terminate successfully. The Manager reallocates area between the CORE and BUF blocks, while the selected Worker performs local ϵ-sizing, displaying the intended division of responsibility. Figure 5 further isolates objective adaptation on the Dual-OTA Comp. The topology, hierarchy, and available actions are fixed, whereas the three target metric sets differ. Both Worker and Manager action distributions change across these requests, consistently with the high Worker sequence diversity (0.89-0.96) reported in Table 3. Thus, policies associate distinct performance requirements with different device groups and functional blocks rather than collapsing to a generic sizing or area-allocation sequence.

5 Conclusions

This work presents H-OSIRIS, a DRL methodology for hierarchical post-layout optimization of multi-block analog circuits. H-OSIRIS combines Manager-driven allocation of functional-block area boundaries with local ϵ-sizing performed by a reconfigurable Worker environment. This decomposition lets the methodology jointly address layout footprint and post-layout functional objectives while accounting for interactions among interconnected blocks. Across four circuits and six metric sets, the evaluation shows that H-OSIRIS improves joint post-layout feasibility from diverse infeasible initial layouts while reducing terminal layout area. These results support DRL optimization as a practical direction for scalable, layout-aware analog design automation. Future work will extend H-OSIRIS to broader circuit classes, richer layout constraints, and larger hierarchical systems.

6 AI USE STATEMENT

In this work, we used generative AI tools to polish the syntax, grammar, and readability of manuscript text. We have not used generative AI tools to gather, process, or analyze data, to develop the methodology, or to conduct experiments. All AI-assisted text and suggestions were reviewed, verified, and revised by the authors for technical accuracy, fidelity to the implemented methodology, and consistency with the experimental results. We take responsibility for the content of this work.

7 ETHICS STATEMENT

This work does not involve human subjects or personal data. H-OSIRIS targets analog integrated-circuit design automation and is evaluated using open-source tools and the SkyWater 130nm process-design kit. The methodology is intended to reduce manual design effort and does not make decisions affecting individuals or protected groups. We report the experimental setup and evaluation criteria to support reproducibility and responsible interpretation of the results.

8 REPRODUCIBILITY STATEMENT

We support reproducibility through the methodological description in Section 3, which specifies the hierarchical formulation, state and action representations, reward construction, and training pipeline. Section 4 describes the evaluated circuits, software and hardware environment, held-out evaluation protocol, and flat and random baselines. Appendices A.1- A.3 report the reward coefficients, metric-specific targets, bounds, and scales, functional-block assignments, sizing and area-allocation design spaces, and RL architecture and training hyperparameters. Finally, Appendix A.5 describes the supplementary artifact package, which provides the initial and terminal schematic netlists, extracted netlists, GDS layouts, and baseline characterizations for the held-out episodes.

REFERENCES

Mehrnaz Ahmadi and Lihong Zhang. Analog layout placement for finfet technology using reinforcement learning. In *2021 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2021.

Davide Basso, Luca Bortolussi, Mirjana Videnovic-Misic, and Husni Habal. Effective analog ics floorplanning with relational graph neural networks and reinforcement learning. In *2025 Design, Automation & Test in Europe Conference (DATE)*. IEEE, 2025.

blind. blind. *blind*, blind.

Ahmet F Budak, Prateek Bhansali, Bo Liu, Nan Sun, David Z Pan, and Chandramouli V Kashyap. Dnn-opt: An rl inspired optimization for analog circuit sizing using deep neural networks. In *2021 58th ACM/IEEE Design Automation Conference (DAC)*. IEEE, 2021.

Tony Chan Carusone, David A Johns, and Kenneth W Martin. *Analog integrated circuit design*. John Wiley & Sons, 2011.

Chen-Chia Chang, Yikang Shen, Shaoze Fan, Jing Li, Shun Zhang, Ningyuan Cao, Yiran Chen, and Xin Zhang. Lamagic: Language-model-based topology generation for analog integrated circuits. In *International Conference on Machine Learning*. PMLR, 2024.

Chen-Chia Chang, Chia-Tung Ho, Yaguang Li, Yiran Chen, and Haoxing Ren. Drc-coder: Automated drc checker code generation using llm autonomous agent. In *Proceedings of the 2025 International Symposium on Physical Design*. ACM, 2025. doi: 10.1145/3698364.3705347.

Eric Chang, Jaeduk Han, Woorham Bae, Zhongkai Wang, Nathan Narevsky, Borivoje Nikolic, and Elad Alon. Bag2: A process-portable framework for generator-based ams circuit design. In *2018 IEEE Custom Integrated Circuits Conference (CICC)*. IEEE, 2018.

Hao Chen, Kai-Chieh Hsu, Walker J Turner, Po-Hsuan Wei, Keren Zhu, David Z Pan, and Haoxing Ren. Reinforcement learning guided detailed routing for custom circuits. In *Proceedings of the 2023 International Symposium on Physical Design*, 2023.

Yuyang Chen, Yiwen Wu, Jingya Wang, Tao Wu, Xuming He, Jingyi Yu, and Hao Geng. Llm-hd: Layout language model for hotspot detection with gds semantic encoding. In *Proceedings of the 61st ACM/IEEE Design Automation Conference*, DAC '24. ACM, 2024. doi: 10.1145/3649329.3658479.

John Crossley, Alberto Puggelli, H-P Le, B Yang, R Nancollas, Kwangmo Jung, Lingkai Kong, Nathan Narevsky, Yue Lu, Nicholas Sutardja, et al. Bag: A designer-oriented integrated framework for the development of ams circuit generators. In *2013 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*. IEEE, 2013.

Fan Cui, Chenyang Yin, Kexing Zhou, Youwei Xiao, Guangyu Sun, Qiang Xu, Qipeng Guo, Yun Liang, Xingcheng Zhang, Demin Song, et al. Origen: Enhancing rtl code generation with code-to-code augmentation and self-reflection. In *Proceedings of the 43rd IEEE/ACM International Conference on Computer-Aided Design*, 2024.

Zehao Dong, Weidong Cao, Muhan Zhang, Dacheng Tao, Yixin Chen, and Xuan Zhang. Cktgnn: Circuit graph neural network for electronic design automation. In *11th International Conference on Learning Representations, ICLR 2023*, 2023.

Yasaman Esfandiari, Jocelyn Rego, Austin Meyer, Jonathan Gallagher, and Mia Levy. Llms for analog circuit design continuum (acdc), 2025. URL https://arxiv.org/abs/2512.09199.

Wenji Fang, Jing Wang, Yao Lu, Shang Liu, Yuchao Wu, Yuzhe Ma, and Zhiyao Xie. A survey of circuit foundation model: Foundation ai models for vlsi circuit design and eda, 2025. URL https://arxiv.org/abs/2504.03711.

Jian Gao, Weidong Cao, Junyi Yang, and Xuan Zhang. Analoggenie: A generative engine for automatic discovery of analog circuit topologies. In *The Thirteenth International Conference on Learning Representations*, 2025a.

Jian Gao, Weimin Fu, Xiaolong Guo, Weidong Cao, and Xuan Zhang. Eva: An efficient and versatile generative engine for targeted discovery of novel analog circuits. In *2025 62nd ACM/IEEE Design Automation Conference (DAC)*. IEEE, 2025b.

Xiaohan Gao, Haoyi Zhang, Bingyang Liu, Yibo Lin, Runsheng Wang, and Ru Huang. Joint placement optimization for hierarchical analog/mixed-signal circuits. In *Proceedings of the 43rd IEEE/ACM International Conference on Computer-Aided Design*, pp. 1–9, 2024a.

Xiaohan Gao, Haoyi Zhang, Siyuan Ye, Mingjie Liu, David Z Pan, Linxiao Shen, Runsheng Wang, Yibo Lin, and Ru Huang. Post-layout simulation driven analog circuit sizing. *Science China Information Sciences*, 2024b.

Kishor Kunal, Meghna Madhusudan, Arvind K Sharma, Wenbin Xu, Steven M Burns, Ramesh Harjani, Jiang Hu, Desmond A Kirkpatrick, and Sachin S Sapatnekar. Align: Open-source analog layout automation from the ground up. In *Proceedings of the 56th Annual Design Automation Conference 2019*, 2019.

Yao Lai, Sungyoung Lee, Guojin Chen, Souradip Poddar, Mengkang Hu, David Z Pan, and Ping Luo. Analogcoder: Analog circuit design via training-free code generation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2025a.

Yao Lai, Souradip Poddar, Sungyoung Lee, Guojin Chen, Mengkang Hu, Bei Yu, Ping Luo, and David Z Pan. Analogcoder-pro: Unifying analog circuit generation and optimization via multi-modal llms. *arXiv preprint arXiv:2508.02518*, 2025b.

Qiufeng Li, Shu Hong, Jian Gao, Xuan Zhang, Tian Lan, and Weidong Cao. Analogfed: Federated discovery of analog circuit topologies with generative ai. *arXiv preprint arXiv:2507.15104*, 2025.

Yaguang Li, Yishuang Lin, Meghna Madhusudan, Arvind Sharma, Wenbin Xu, Sachin S Sapatnekar, Ramesh Harjani, and Jiang Hu. A customized graph neural network model for guiding analog ic placement. In *Proceedings of the 39th International Conference on Computer-Aided Design*, 2020.

Zonghao Li and Anthony Chan Carusone. An open-source ams circuit optimization framework based on reinforcement learning—from specifications to layouts. *IEEE Access*, 12:150032–150045, 2024.

Mingjie Liu, Yun-Da Tsai, Wenfei Zhou, and Haoxing Ren. Craftrtl: High-quality synthetic data generation for verilog code models with correct-by-construction non-textual representations and targeted code repair. In *The Thirteenth International Conference on Learning Representations*, 2025a.

Shang Liu, Wenji Fang, Yao Lu, Qijun Zhang, and Zhiyao Xie. Towards big data in ai for eda research: Generation of new pseudo circuits at rtl stage. In *Proceedings of the 30th Asia and South Pacific Design Automation Conference*, 2025b.

Asal Mehradfar, Xuzhe Zhao, Yilun Huang, Emir Ceyani, Yankai Yang, Hamidreza Aghasi, Salman Avestimehr, et al. Falcon: An ml framework for fully automated layout-constrained analog circuit design. *Advances in Neural Information Processing Systems*, 38:54450–54483, 2026.

Danial Noori Zadeh and Mohamed B. Elamien. Generative ai for analog integrated circuit design: Methodologies and applications. *IEEE Access*, 2025. doi: 10.1109/ACCESS.2025.3553743.

Behzad Razavi. *Design of analog CMOS integrated circuits*. 2005.

Haoxing Ren, George F. Kokai, Walker J. Turner, and Ting-Sheng Ku. Paragraph: Layout parasitics and device parameter prediction using graph neural networks. In *2020 57th ACM/IEEE Design Automation Conference (DAC)*, 2020. doi: 10.1109/DAC18072.2020.9218515.

Prashanth Vijayaraghavan, Luyao Shi, Ehsan Degan, Vandana Mukherjee, and Xin Zhang. Autocircuit-rl: Reinforcement learning-driven llm for automated circuit topology generation. In *Forty-second International Conference on Machine Learning*.

Ziming Wei, Zichen Kong, Yuan Wang, David Z Pan, and Xiyuan Tang. Toposizing: An llm-aided framework of topology-based understanding and sizing for ams circuits. *arXiv preprint arXiv:2509.14169*, 2025.

Bo Wen and Xin Zhang. Enhancing reasoning to adapt large language models for domain-specific applications, 2025. URL https://arxiv.org/abs/2502.04384.

Biying Xu, Keren Zhu, Mingjie Liu, Yibo Lin, Shaolan Li, Xiyuan Tang, Nan Sun, and David Z Pan. Magical: Toward fully automated analog ic layout leveraging human and machine intelligence. In *2019 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*. IEEE, 2019.

Peng Xu, Yuan Pu, Keren Zhu, Tinghuan Chen, Tsungyi Ho, and Bei Yu. Paplace: Performance-driven differentiable analog placement. *ACM Transactions on Design Automation of Electronic Systems*, 31(5):1–23, 2026.

Zhuorui Zhao, Ruidi Qiu, Chao Lin, Grace Li Zhang, Bing Li, and Ulf Schlichtmann. Vrank: Enhancing verilog code generation from large language models via self-consistency. In *2025 26th International Symposium on Quality Electronic Design (ISQED)*. IEEE, 2025.

Keren Zhu, Mingjie Liu, Yibo Lin, Biying Xu, Shaolan Li, Xiyuan Tang, Nan Sun, and David Z Pan. Geniusroute: A new analog routing paradigm using generative neural network guidance. In *2019 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*. IEEE, 2019.

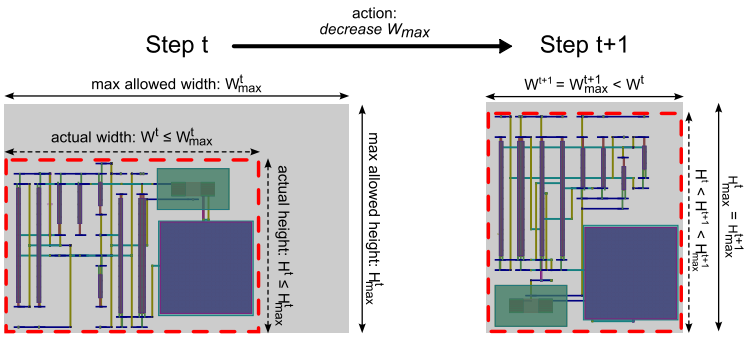


Figure 6: Example of a Manager area-boundary action. At step t, the selected functional block is placed within maximum dimensions $(W^{(t)}_{\max,b}, H^{(t)}_{\max,b})$. The Manager decreases only the maximum allowed width, producing $(W^{(t+1)}_{\max,b}, H^{(t+1)}_{\max,b})$ with $W^{(t+1)}_{\max,b} < W^{(t)}_{\max,b}$ and $H^{(t+1)}_{\max,b} = H^{(t)}_{\max,b}$. The subsequent $t+1$ place-and-route pass reallocates the layout within the updated boundary. Grey rectangles denote the allowed boundary, whereas red dashed rectangles denote the generated layout footprint. Actual dimensions need to satisfy the corresponding maximum bounds. In this example, the $t+1$ layout saturates the width bound while remaining just below the height bound.

A Appendix

A.1 Additional Methodology Details

A.1.1 Area-Boundary Action Example

Figure 6 illustrates a width-reduction action. At step t, the block is generated within the boundary $(W^{(t)}_{\max,b}, H^{(t)}_{\max,b})$. The Manager then reduces the maximum allowed width while keeping the maximum allowed height fixed, yielding $W^{(t+1)}_{\max,b} < W^{(t)}_{\max,b}$ and $H^{(t+1)}_{\max,b} = H^{(t)}_{\max,b}$. The layout regenerated at step $t+1$ consequently adopts a different spatial arrangement within the narrower boundary. Its realized dimensions need not equal the allowed values, but must satisfy $W^{(t+1)}_b \le W^{(t+1)}_{\max,b}$ and $H^{(t+1)}_b \le H^{(t+1)}_{\max,b}$. In this example, because $W^{(t+1)}_{\max,b} < W^{(t)}_b$, the layout uses the available excess height to accommodate the narrower width constraint.

A.1.2 Manager and Worker States

We illustrate the hierarchical state representation using the Dual-OTA composite, which contains two connected functional blocks, FT_0 and FT_1. Each block is a five-transistor OTA comprising a tail device (X_0), a matched differential pair (X_{12}), a matched current mirror (X_{34}), and a compensation capacitor (X_{C5}). This example clarifies the distinct information used by the circuit-level Manager and by the Worker loaded for one selected block. The Manager represents the circuit as a graph with one node per functional block. Hence, the example graph has nodes FT_0 and FT_1, with edges representing their electrical connectivity through shared nets. For a block b, its node contains its current allowed boundary, realized area, and assigned post-layout metrics:

$$\mathbf{x}^{\mathrm{M}}_b = \left[W_{\max,b},\ H_{\max,b},\ A_b,\ \{v_i, \hat{e}_i\}_{i \in \mathcal{Q}_b},\ d^{(b)}_t \right]. \tag{5}$$

Here, v_i is the measured post-layout value of metric q_i and $\hat{e}_i$ is its target violation, as defined in Equation 2. The Manager graph also contains circuit-level information, including the total realized area A_t and the values of globally assigned metrics. Thus, the Manager can jointly identify whether a block is electrically limiting, geometrically constrained, or contributing disproportionately to total area. For example, if FT_0 has an unsatisfied gain target while its allowed boundary is narrow, the Manager can select FT_0 and apply a width or height action. This action changes $(W_{\max,\mathrm{FT}_0}, H_{\max,\mathrm{FT}_0})$, but does not directly modify transistor parameters. The resulting boundary is passed to the Worker loaded for FT_0. The corresponding Worker state contains the current sizing

of the selected block, its local post-layout results, and the relevant Manager context:

$$\mathbf{s}_t^{(\mathrm{FT}_0)} = \left[\mathbf{z}_t^{(\mathrm{FT}_0)},\ \{v_i, \hat{e}_i\}_{i\in\mathcal{Q}_{\mathrm{FT}_0}},\ A_{\mathrm{FT}_0},\ W_{\max,\mathrm{FT}_0},\ H_{\max,\mathrm{FT}_0},\ A_t\right]. \tag{6}$$

In this example,

$$\mathbf{z}_t^{(\mathrm{FT}_0)} = \left[\mathbf{z}_{X_0},\ \mathbf{z}_{X_{12}},\ \mathbf{z}_{X_{34}},\ s_{C5}\right], \tag{7}$$

where $\mathbf{z}_{X_0}$, $\mathbf{z}_{X_{12}}$, and $\mathbf{z}_{X_{34}}$ contain the controllable length, width, and finger count of the tail-device, differential-pair, and current-mirror matching groups, respectively, while s_{C5} is the capacitor side length. Consequently, the Worker can select a local sizing action using both its present electrical state and the area boundary imposed by the Manager. For instance, after the Manager reduces $W_{\max,\mathrm{FT}_0}$, the Worker may resize the differential pair or current mirror so that the regenerated layout remains feasible within the new boundary while improving the metrics in $\mathcal{Q}_{\mathrm{FT}_0}$. All resulting electrical quantities are still evaluated after assembling and simulating the complete extracted circuit. Therefore, the local Worker decision remains informed by inter-block coupling.

A.1.3 Reward Definition

This section specifies the area rewards, terminal indicators, and Manager progress reward used in Equations 3 and 4. For the active functional block b, the local area reward is:

$$R_{\mathrm{area}}^{(b)} = \begin{cases} 1, & 0 < A_b \le A_{\mathrm{tar}}^{(b)}, \\ 0, & A_b \ge A_{\max}^{(b)}, \\ 1 - \dfrac{A_b - A_{\mathrm{tar}}^{(b)}}{A_{\max}^{(b)} - A_{\mathrm{tar}}^{(b)}}, & \text{otherwise.} \end{cases} \tag{8}$$

Thus, layouts at or below the target area receive unit reward, whereas layouts at or above the maximum allowed area receive zero reward. We define $\hat{e}_{\mathrm{area}}^{(b)} = 1 - R_{\mathrm{area}}^{(b)}$ and use $\tilde{w}_{\mathrm{area}} = w_{\mathrm{area}}(1+\alpha\hat{e}_{\mathrm{area}}^{(b)})$ in Equation 3. The active functional-block episode succeeds when every assigned metric to satisfy its done bound:

$$d_t^{(b)} = \prod_{i\in\mathcal{Q}_b} \mathbf{1}\left\{\ell_i^{\mathrm{done}} \le v_i \le u_i^{\mathrm{done}}\right\}. \tag{9}$$

The Manager uses the same area-reward function with global area references, yielding $R_{\mathrm{area}}^{\mathrm{glob}}$. Its terminal indicator additionally requires total area to satisfy the global done bound:

$$d_t^{\mathrm{M}} = \left(\prod_{i\in\mathcal{Q}_{\mathrm{M}}} \mathbf{1}\left\{\ell_i^{\mathrm{done}} \le v_i \le u_i^{\mathrm{done}}\right\}\right) \mathbf{1}\left\{A_t \le A_{\mathrm{done}}^{\mathrm{glob}}\right\}. \tag{10}$$

To provide immediate feedback after each Manager action, we compute the mean normalized electrical violation:

$$\bar{e}_t = \frac{1}{|\mathcal{Q}_{\mathrm{M}}|} \sum_{i\in\mathcal{Q}_{\mathrm{M}}} \hat{e}_i \tag{11}$$

and the normalized area change:

$$\delta_{A,t} = \frac{A_{t-1} - A_t}{\max(A_{t-1}, A_t)}. \tag{12}$$

The Manager progress reward is:

$$r_{\mathrm{progress},t}^{\mathrm{M}} = \mathrm{clip}_{[-1,1]}\left(\beta_q(\bar{e}_{t-1} - \bar{e}_t) + \begin{cases} \beta_A\delta_{A,t}, & \delta_{A,t} \ge 0, \\ \beta_G\delta_{A,t}, & \delta_{A,t} < 0. \end{cases}\right). \tag{13}$$

Therefore, electrical improvement and area reduction receive positive feedback, whereas area growth is penalized. The fixed reward parameters and topology-specific normalization scales are reported in Appendix A.2.1.

A.1.4 DDQN Learning Target

H-OSIRIS uses one reconfigurable Worker environment and one circuit-level Manager. Each functional block b owns a persistent Worker bundle containing its online and target Q-networks, replay buffer, design-space configuration, and metric configuration. When the Manager selects b, the Worker loads this bundle and executes the corresponding local episode. Its state $\mathbf{s}_t^{(b)}$ contains the current block sizing vector, the local post-layout metrics $\{v_i\}_{i\in\mathcal{Q}_b}$, the block area A_b, and the context received from the Manager. A Worker action $\mathbf{a}_t^{(b)} = (g_t, p_t, \delta_t)$ selects a device group g_t, a sizing parameter $p_t \in \{w, \ell, n_f\}$, and an increment $\delta_t \in \{-\delta_{g_t,p_t}, +\delta_{g_t,p_t}\}$. The resulting sizing update is evaluated in the assembled full circuit, yielding the Worker reward $r_t^{(b)}$ defined in Section 3.3. Each local transition is stored in the replay buffer of the currently loaded functional-block bundle:

$$(\mathbf{s}_t^{(b)}, \mathbf{a}_t^{(b)}, r_t^{(b)}, \mathbf{s}_{t+1}^{(b)}, d_t^{(b)}), \tag{14}$$

where $d_t^{(b)}$ is the Worker episode-termination indicator; in this work it equals the local done indicator in Equation 9. The Manager acts at the physical-constraint level. Its state $\mathcal{G}_t$ is a graph with one node per functional block. Node and edge features encode the current block constraints, local performance and area information, and circuit connectivity. Its action a_t^{M} selects a block and a valid geometry operation, such as $\Delta W_{b,t} \in \{-\delta_W, +\delta_W\}$ or $\Delta H_{b,t} \in \{-\delta_H, +\delta_H\}$. The subsequent full-circuit layout and simulation provide the Manager reward r_t^{M}, which combines absolute circuit quality with the improvement relative to the preceding Manager step. The Manager stores transitions:

$$(\mathcal{G}_t, a_t^{\mathrm{M}}, r_t^{\mathrm{M}}, \mathcal{G}_{t+1}, d_t^{\mathrm{M}}), \tag{15}$$

where d_t^{M} is the Manager episode-termination indicator; in this work it equals the global terminal-feasibility indicator in Equation 10. The Manager and each persistent block-specific Worker bundle maintain separate Double-DQN online and target networks and an independent replay buffer. The Worker environment loads only the bundle associated with the Manager-selected block. For an agent $j \in \{\mathrm{M}\} \cup \mathcal{B}$, let $x_t^{(j)}$ and $a_t^{(j)}$ denote its state and action. The DDQN target is:

$$y_t^{(j)} = r_t^{(j)} + \gamma(1 - d_t^{(j)}) Q_{\bar{\boldsymbol{\theta}}_j}\left(x_{t+1}^{(j)}, \arg\max_{a'} Q_{\boldsymbol{\theta}_j}(x_{t+1}^{(j)}, a')\right), \tag{16}$$

where $Q_{\boldsymbol{\theta}_j}$ selects the next action and $Q_{\bar{\boldsymbol{\theta}}_j}$ evaluates it. The discount factor is $\gamma = 0.99$. The online network minimizes the Huber loss between $Q_{\boldsymbol{\theta}_j}(x_t^{(j)}, a_t^{(j)})$ and $y_t^{(j)}$. Target-network parameters are updated using Polyak averaging:

$$\bar{\boldsymbol{\theta}}_j \leftarrow \tau_p \boldsymbol{\theta}_j + (1 - \tau_p)\bar{\boldsymbol{\theta}}_j, \tag{17}$$

with $\tau_p = 0.005$. Each block-specific Worker policy uses a fully connected ReLU DQN. The Manager, instead, uses a relational graph neural network each functional block represented as a node and their connections as edges. This architecture is needed because the Manager must reason jointly about a variable set of functional blocks, their electrical connectivity, and the coupling between per-block geometry constraints and full-circuit performance.

A.2 Network Architectures and Hyperparameters

Each functional block owns a persistent Worker bundle containing its block-specific Double-DQN online and target networks and replay buffer. The single reconfigurable Worker environment loads the bundle selected by the Manager. Its Q-network is a fully connected multilayer perceptron (MLP): the input dimension equals the loaded block-specific state dimension, the output dimension equals its discrete action-space cardinality, and the network contains three hidden layers of 256 units with ReLU activations. Thus, input and output dimensions vary with the loaded functional block, whereas the hidden-layer architecture is shared across block-specific Worker bundles. The Manager is also trained with Double-DQN, but represents the circuit as a relational graph whose nodes correspond to functional blocks and whose edges represent circuit connectivity. Its online and target networks contain three relational graph message-passing layers with hidden dimension 128, ReLU activations, and dropout probability 0.1. The Manager Q-head produces values for the valid block-geometry operation pairs. At both hierarchy levels, invalid actions are masked before action selection and target evaluation. Table 5 reports the network architectures and optimization hyperparameters shared across the reported experiments.

Table 5: Network architectures and hyperparameters. Worker input and output dimensions depend on the selected functional block. Manager graph dimensions depend on the circuit hierarchy.

Category	Setting	Value
Worker network	Learning algorithm	Double-DQN
	Network architecture	Fully connected MLP with ReLU activations
	Hidden-layer widths	$[256, 256, 256]$
	Input dimension	Block-specific Worker-state dimension
	Output dimension	Block-specific discrete action-space cardinality
	Online/target networks	Separate networks, updated by Polyak averaging
Manager network	Learning algorithm	Double-DQN
	Network architecture	Relational graph neural network with ReLU activations
	Message-passing layers	3
	Hidden dimension	128
	Dropout probability	0.1
	Input/output dimensions	Hierarchy-dependent graph and action dimensions
	Online/target networks	Separate networks, updated by Polyak averaging
Shared optimization	Optimizer	Adam
	Initial learning rate	10^{-4}
	Final learning rate	10^{-5}
	Learning-rate schedule	Linear decay over 2 500 episodes
	Discount factor γ	0.99
	Target-network Polyak rate τ_p	0.005
	Loss function	Huber loss
	Replay minibatch size	64
	Gradient-norm clipping	0.5
Exploration and updates	Worker ϵ schedule	Linear decay from 1.0 to 0.05 over 50% of expected Worker decisions
	Manager ϵ schedule	Linear decay from 1.0 to 0.05 over 75% of Manager episodes
	Worker update frequency	Every 150 completed Worker transitions
	Worker updates per barrier	8
	Manager update frequency	Every 100 completed Manager transitions
	Manager updates per barrier	4
	Manager pretraining	20 epochs on Phase-0 logs

Table 6: Reward-shaping parameters shared across the evaluated hierarchies.

Parameter	Value	Role
Worker metric weight w_i	2	Base weight of each optimized electrical metric
Worker area weight w_{area}	1	Base weight of the local area-reward branch
Adaptive-weight coefficient α	1	Emphasizes larger normalized target violations
Manager electrical/area branch weight	1	Uniform average over declared Manager objectives
Manager quality-progress coefficient β_q	0.5	Weights reduction in average normalized error
Manager area-reduction coefficient β_A	0.2	Rewards a reduction in total layout area
Manager area-growth coefficient β_G	0.1	Penalizes an increase in total layout area
Manager progress range	$[-1, 1]$	Clipping interval of the progress term

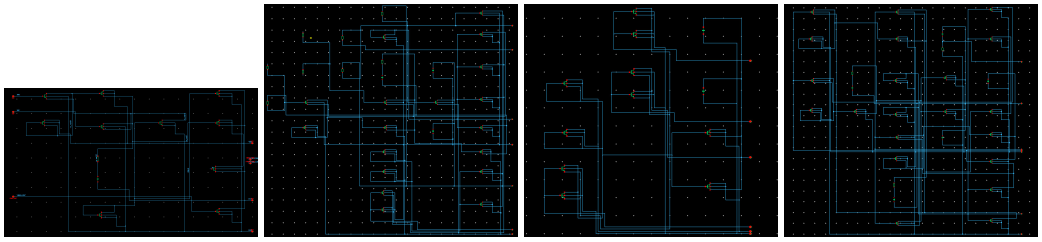
A.2.1 Reward Parameters

This section summarizes the fixed parameters used by the hierarchical reward defined in Section 3.3. The functional-block scopes and active metric sets are listed in Table 2.

Electrical target and done bounds are specified for each metric set before training. They are evaluated from post-layout simulation of the complete extracted circuit, including when a metric is assigned to an individual functional block. For an electrical metric q_i, the reward uses the normalized target violation $\hat{e}_i = e_i/\tau_i$ and the shaped term $R_i = \exp(-\hat{e}_i)$. The scale τ_i controls how rapidly the reward decreases outside the target interval and is expressed in the physical unit of the corresponding metric. Table 6 summarizes the reward-shaping parameters shared across the evaluated hierarchies. Worker metric weights are $w_i = 2$, whereas the area branch has unit weight. The adaptive Worker weighting coefficient is $\alpha = 1$. The Manager instead averages the shaped rewards of all declared local and global metrics together with the global area reward. Table 7 reports the electrical normalization scales used by each active metric of the evaluated topologies. Each τ_i is expressed in the unit of its associated electrical metric and is used in Equation 2. Area references are calibrated offline and separately for each hierarchy before RL training. The calibration uses up to 60 legal layouts drawn from the corresponding design space and produces three fixed, topology-specific values:

Table 7: Metric-specific reward scales. τ_i is expressed in the unit of the associated electrical metric.

Topology	Metric set	Functional block	Electrical metric	τ_i
2-Stage Amp	$\mathcal{S}_{\mathrm{A}}$	CORE	BW	5.0×10^6 Hz
		BUF	BW	2.0×10^5 Hz
		GLB	A_v	0.5 dB
3-OpAmp InAmp	$\mathcal{S}_{\mathrm{B}}$	IN_A	BW	5.0×10^6 Hz
		IN_B	BW	5.0×10^6 Hz
		DIFF	BW	5.0×10^6 Hz
		GLB	BW	5.0×10^6 Hz
Dual-OTA Comp.	$\mathcal{S}_{\mathrm{C}}$	FT_0	A_v	5 dB
		FT_1	A_v	5 dB
	$\mathcal{S}_{\mathrm{D}}$	FT_0	CMRR	6 dB
		FT_1	CMRR	6 dB
		FT_1	PSRR	5 dB
	$\mathcal{S}_{\mathrm{E}}$	FT_0	CMRR	6 dB
		FT_0	PSRR	5 dB
		FT_0	A_v	5 dB
		FT_1	CMRR	6 dB
		FT_1	PSRR	5 dB
		FT_1	A_v	5 dB
Amp-Filter Comp.	$\mathcal{S}_{\mathrm{F}}$	MIL	A_v	3 dB
		FT	A_v	2 dB
		HPF	f_c	2.0×10^6 Hz



(a) 2-Stage Amp. (b) 3-OpAmp InAmp. (c) 2-OTA Comp. (d) Amp-Filter Comp.

Figure 7: Schematics of the four evaluated circuit topologies.

$A^{(b)}_{\mathrm{tar}}$, $A^{(b)}_{\mathrm{done}}$, and $A^{(b)}_{\mathrm{max}}$ for each block, with the corresponding global references for the Manager. They respectively represent a compact-layout reference, the maximum area accepted for terminal feasibility, and the upper limit beyond which the dense area reward vanishes. The references remain fixed throughout training and held-out testing. For a layout with area A, the dense area term rewards layouts at or below $A^{(b)}_{\mathrm{tar}}$, decreases linearly between $A^{(b)}_{\mathrm{tar}}$ and $A^{(b)}_{\mathrm{max}}$, and is zero at or above $A^{(b)}_{\mathrm{max}}$. The global done-area reference $A^{\mathrm{glob}}_{\mathrm{done}}$ is additionally included in the Manager terminal condition. Therefore, Manager success requires all configured electrical metrics and the global area constraint to be satisfied jointly. Area remains a dense objective even after it is feasible, encouraging compact layouts rather than merely layouts below the terminal limit.

A.3 Additional Design-Space Details and Held-Out Results

A.3.1 Evaluated Circuit Topologies

Figure 7 shows the circuit topologies evaluated in Section 4. Table 10 reports the device composition of each topology organized by functional block.

A.3.2 Hierarchical Design Spaces

Table 8 details the discrete local design spaces loaded by the Worker for the selected block. Rows are grouped first by topology and functional block (FB), and then by matched device group. The parameter columns separately report the allowed intervals and increments for transistor length (ℓ_g), width (w_g), and finger count ($n_{f,g}$), as well as resistor (R) and capacitor side length ($s_{C,g}$) where applicable. The final two columns give the number of valid local actions, $|\mathcal{A}_b|$, and the number of discrete sizing configurations, $|\mathcal{X}_b|$, for each block. Matched groups ensure that devices requiring

Table 8: Block-specific Worker configuration sizing spaces. Each row is one matched device group. Ranges and increments are expressed in μm except R (Ω) and $n_{f,g}$ (units). $|\mathcal{A}_b|$ and $|\mathcal{X}_b|$ are the action and discrete sizing-state cardinalities of the Worker when loaded with block b.

Topology	FB	Matched group (devices)	ℓ_g range	$\delta_{g,\ell}$	w_g range	$\delta_{g,w}$	$n_{f,g}$ range	δ_{g,n_f}	R range	$\delta_{g,R}$	$s_{C,g}$ range	δ_{g,s_C}	$\lvert\mathcal{A}_b\rvert$	$\lvert\mathcal{X}_b\rvert$
2-Stage Amp	CORE	M_{02} (M0/M2)	[0.48, 0.52]	0.01	[6.35, 10.35]	1	[1, 2]	1	–	–	–	–	34	6.56×10^{10}
		M_{110} (M1/M10)	[0.48, 0.52]	0.01	[3, 7]	1	[1, 2]	1	–	–	–	–		
		M_{34} (M3/M4)	[0.48, 0.52]	0.01	[23, 27]	1	[1, 2]	1	–	–	–	–		
		M_5 (M5)	[0.48, 0.52]	0.01	[48, 52]	1	[1, 4]	1	–	–	–	–		
		M_6 (M6)	[0.48, 0.52]	0.01	[23, 27]	1	[1, 2]	1	–	–	–	–		
		R_0 (R0)	–	–	–	–	–	–	[1 500, 3 500]	500	–	–		
		C_0 (C0)	–	–	–	–	–	–	–	–	[18, 38]	1		
	BUFFER	M_{BUF} (M0)	[0.46, 0.54]	0.01	[34, 45]	1	[1, 4]	1	–	–	–	–	12	93 312
		M_{SINK} (M1)	[0.46, 0.54]	0.01	[5, 16]	1	[1, 2]	1	–	–	–	–		
3-OpAmp InAmp	IN_A	M_0 (M0)	[0.98, 1.02]	0.01	[3, 7]	1	[1, 2]	1	–	–	–	–	20	1.05×10^{7}
		M_{12} (M1/M2)	[0.98, 1.02]	0.01	[18, 22]	1	[1, 4]	1	–	–	–	–		
		M_{34} (M3/M4)	[0.98, 1.02]	0.01	[38, 42]	1	[1, 4]	1	–	–	–	–		
		C_5 (C5)	–	–	–	–	–	–	–	–	[10, 30]	1		
	IN_B	M_0 (M0)	[0.98, 1.02]	0.01	[3, 7]	1	[1, 2]	1	–	–	–	–	20	1.05×10^{7}
		M_{12} (M1/M2)	[0.98, 1.02]	0.01	[18, 22]	1	[1, 4]	1	–	–	–	–		
		M_{34} (M3/M4)	[0.98, 1.02]	0.01	[38, 42]	1	[1, 4]	1	–	–	–	–		
		C_5 (C5)	–	–	–	–	–	–	–	–	[10, 30]	1		
	DIFF	M_{02} (M0/M2)	[0.48, 0.52]	0.01	[6.4, 10.4]	1	[1, 2]	1	–	–	–	–	34	2.62×10^{11}
		M_{110} (M1/M10)	[0.48, 0.52]	0.01	[3, 7]	1	[1, 2]	1	–	–	–	–		
		M_{34} (M3/M4)	[0.48, 0.52]	0.01	[23, 27]	1	[1, 4]	1	–	–	–	–		
		M_5 (M5)	[0.48, 0.52]	0.01	[48, 52]	1	[1, 4]	1	–	–	–	–		
		M_6 (M6)	[0.48, 0.52]	0.01	[23, 27]	1	[1, 4]	1	–	–	–	–		
		R_0 (R0)	–	–	–	–	–	–	[1 500, 3 500]	500	–	–		
		C_0 (C0)	–	–	–	–	–	–	–	–	[18, 38]	1		
Dual-OTA Comp.	FT_0	M_0 (M0)	[0.8, 1.2]	0.1	[3, 7]	1	[1, 2]	1	–	–	–	–	20	2.62×10^{6}
		M_{12} (M1/M2)	[0.3, 0.7]	0.1	[8, 12]	1	[1, 2]	1	–	–	–	–		
		M_{34} (M3/M4)	[0.3, 0.7]	0.1	[18, 22]	1	[1, 2]	1	–	–	–	–		
		C_5 (C5)	–	–	–	–	–	–	–	–	[1, 21]	1		
	FT_1	M_0 (M0)	[0.8, 1.2]	0.1	[3, 7]	1	[1, 2]	1	–	–	–	–	20	2.62×10^{6}
		M_{12} (M1/M2)	[0.3, 0.7]	0.1	[8, 12]	1	[1, 2]	1	–	–	–	–		
		M_{34} (M3/M4)	[0.3, 0.7]	0.1	[18, 22]	1	[1, 2]	1	–	–	–	–		
		C_5 (C5)	–	–	–	–	–	–	–	–	[1, 21]	1		
Amp-Filter Comp.	MIL	M_{02} (M0/M2)	[0.15, 0.17]	0.01	[3, 7]	1	[1, 2]	1	–	–	–	–	34	2.55×10^{9}
		M_{110} (M1/M10)	[0.15, 0.17]	0.01	[1, 5]	1	[1, 2]	1	–	–	–	–		
		M_{34} (M3/M4)	[0.15, 0.17]	0.01	[13, 17]	1	[1, 2]	1	–	–	–	–		
		M_5 (M5)	[0.15, 0.17]	0.01	[58, 62]	1	[1, 2]	1	–	–	–	–		
		M_6 (M6)	[0.15, 0.17]	0.01	[13, 17]	1	[1, 2]	1	–	–	–	–		
		R_0 (R0)	–	–	–	–	–	–	[500, 2 500]	500	–	–		
		C_0 (C0)	–	–	–	–	–	–	–	–	[1, 21]	1		
	FT	M_0 (M0)	[0.8, 1.2]	0.1	[3, 7]	1	[1, 2]	1	–	–	–	–	20	2.62×10^{6}
		M_{12} (M1/M2)	[0.3, 0.7]	0.1	[8, 12]	1	[1, 2]	1	–	–	–	–		
		M_{34} (M3/M4)	[0.3, 0.7]	0.1	[18, 22]	1	[1, 2]	1	–	–	–	–		
		C_5 (C5)	–	–	–	–	–	–	–	–	[1, 21]	1		
	HPF	M_{112} (M1/M2)	[0.15, 0.2]	0.01	[11, 19]	1	[1, 2]	1	–	–	–	–	40	9.33×10^{14}
		M_{34} (M3/M4)	[0.15, 0.2]	0.01	[8, 16]	1	[1, 2]	1	–	–	–	–		
		M_{07} (M0/M7)	[0.15, 0.2]	0.01	[20, 28]	1	[1, 2]	1	–	–	–	–		
		M_5 (M5)	[0.15, 0.2]	0.01	[26, 34]	1	[1, 2]	1	–	–	–	–		
		M_6 (M6)	[0.15, 0.2]	0.01	[26, 34]	1	[1, 2]	1	–	–	–	–		
		R_0 (R0)	–	–	–	–	–	–	[500, 2 500]	500	–	–		
		C_0 (C0)	–	–	–	–	–	–	–	–	[1, 11]	1		
		C_1 (C1)	–	–	–	–	–	–	–	–	[1, 11]	1		
		C_2 (C2)	–	–	–	–	–	–	–	–	[1, 21]	1		
		R_1 (R1)	–	–	–	–	–	–	[500, 2 500]	500	–	–		

identical sizing, such as differential pairs and current mirrors, are adjusted together. The resulting Worker action spaces remain compact, ranging from 12 actions for BUF to 40 for HPF, despite large discrete sizing spaces: for example, the INAMP DIFF block has 2.62×10^{11} configurations and the HPF block has 9.33×10^{14}. Table 9 reports the complementary Manager area allocation space. Each non-GLB row specifies the legal width (W) and height (H) intervals, their grid step (δ_W, δ_H), the number of block-boundary configurations ($|\mathcal{C}_b|$), and the four local geometry actions (local $|\mathcal{A}_M|$). The GLB row aggregates the Manager action cardinality (GLB $|\mathcal{A}_{\mathrm{M}}|$), the product allocation space ($\prod_{b\in\mathcal{B}} |\mathcal{C}_b|$), and the product block-specific sizing space ($\prod_{b\in\mathcal{B}} |\mathcal{X}_b|$). For example, the 3-OpAmp InAmp combines 2.94×10^8 legal block-boundary allocations with a 2.89×10^{25} discrete Worker sizing space. The Amp-Filter Comp. reaches 6.25×10^{30} joint configurations. These values highlight that the full search problem remains combinatorial, while the proposed hierarchy factorizes its decisions into block sizing policies and block-level area-allocation decisions.

A.4 Hierarchical-Flat Comparison Protocol

Table 11 summarizes the controlled comparison between H-OSIRIS and the flat baseline. The two formulations use the same circuit topology, electrical and global-area constraints, training and held-out sizing distributions, and post-layout evaluation flow. They differ only in the optimization representation: H-OSIRIS factorizes decisions over functional blocks, whereas the flat baseline treats the

Table 9: Manager discrete area-allocation spaces. $|\mathcal{C}_b|$ is the number of legal width-height bounds for block b. The GLB row gives the joint Manager action and allocation cardinalities, together with the product Worker sizing space. Dimensions are in μm.

Topology	Target	W range	δ_W	H range	δ_H	$\lvert\mathcal{C}_b\rvert$	local $\lvert\mathcal{A}_{\mathrm{M}}\rvert$	GLB $\lvert\mathcal{A}_{\mathrm{M}}\rvert$	$\prod_{b\in\mathcal{B}}\lvert\mathcal{C}_b\rvert$	$\prod_{b\in\mathcal{B}}\lvert\mathcal{X}_b\rvert$
2-Stage Amp	CORE	[110, 180]	10	[150, 210]	10	56	4	–	–	–
	BUF	[32, 56]	4	[84, 128]	4	84	4	–	–	–
	GLB	–	–	–	–	–	–	8	4 704	6.12×10^{15}
3-OpAmp InAmp	IN_A	[100, 140]	2	[100, 140]	2	441	4	–	–	–
	IN_B	[90, 120]	2	[90, 120]	2	256	4	–	–	–
	DIFF	[220, 320]	2	[220, 320]	2	2 601	4	–	–	–
	GLB	–	–	–	–	–	–	12	2.94×10^{8}	2.89×10^{25}
Dual-OTA Comp.	FT_0	[50, 100]	2	[76, 100]	2	338	4	–	–	–
	FT_1	[50, 100]	2	[76, 100]	2	338	4	–	–	–
	GLB	–	–	–	–	–	–	8	1.14×10^{5}	6.89×10^{12}
Amp-Filter Comp.	MIL	[150, 170]	2	[150, 170]	2	121	4	–	–	–
	FT	[50, 100]	2	[76, 100]	2	338	4	–	–	–
	HPF	[150, 170]	2	[150, 170]	2	121	4	–	–	–
	GLB	–	–	–	–	–	–	12	4.95×10^{6}	6.25×10^{30}

Table 10: Functional block composition of the evaluated circuit topologies. Counts refer to schematic device instances.

Topology	Functional block	PMOS (#)	NMOS (#)	RES (#)	CAP (#)	Total (#)
2-Stage Amp	CORE	3	5	1	1	10
	BUFFER	0	2	0	0	2
	Circuit total	3	7	1	1	12
3-OpAmp InAmp	IN_A	2	3	2	1	8
	IN_B	2	3	1	1	7
	DIFF	3	5	5	1	14
	Circuit total	7	11	8	3	29
Dual-OTA Comp.	FT_0	2	3	0	1	6
	FT_1	2	3	0	1	6
	Circuit total	4	6	0	2	12
Amp-Filter Comp.	MIL	3	5	1	1	10
	FT	2	3	0	1	6
	HPF	3	5	2	3	13
	Circuit total	8	13	3	5	29

complete circuit as a single optimization unit. Table 4 reports the corresponding training-time and held-out feasibility results.

A.4.1 Initial Post-Layout Gaps

Figure 8 extends the initial-gap analysis to the Amp-Filter Comp. ($\mathcal{S}_{\mathrm{F}}$) and the three Dual-OTA Comp. objective sets ($\mathcal{S}_{\mathrm{C}}$, $\mathcal{S}_{\mathrm{D}}$, and $\mathcal{S}_{\mathrm{E}}$). As in Figure 3, these are independently sampled characterization configurations rather than the held-out starts summarized in Table 3. In each left panel, a horizontal row represents one displayed quality metric. The green segment gives the fraction of sampled layouts within the $\pm3.5\%$ done band, while the remaining colors partition out-of-band (OOB) layouts by increasing normalized violation. The right panel of each subfigure reports cumulative multi-metric satisfaction. For a row labelled N/K, the green portion gives the fraction of layouts satisfying at least N of the K displayed constraints, whereas the grey portion gives the complement. Consequently, the K/K row directly reports joint feasibility. This representation distinguishes layouts that already satisfy several individual requirements from layouts satisfying the complete multi-objective request. The panels confirm that the initial post-layout gap persists across both topology and objective-set changes. Under the uniform displayed done-band criterion, no valid sampled layout is jointly feasible for the Amp-Filter Comp. ($0/100$) or Dual-OTA $\mathcal{S}_{\mathrm{E}}$ ($0/100$). Only $5/100$ layouts are jointly feasible for Dual-OTA $\mathcal{S}_{\mathrm{C}}$, and $2/100$ for $\mathcal{S}_{\mathrm{D}}$. Together with Figure 3, these results show that post-layout infeasibility is not specific to one circuit family or electrical objective set, but is a general obstacle that the optimization procedure must overcome.

Table 11: Implementation comparison between the hierarchical formulation and the flat baseline. Quantities not determined by the representation are held fixed within each topology-metric set pair.

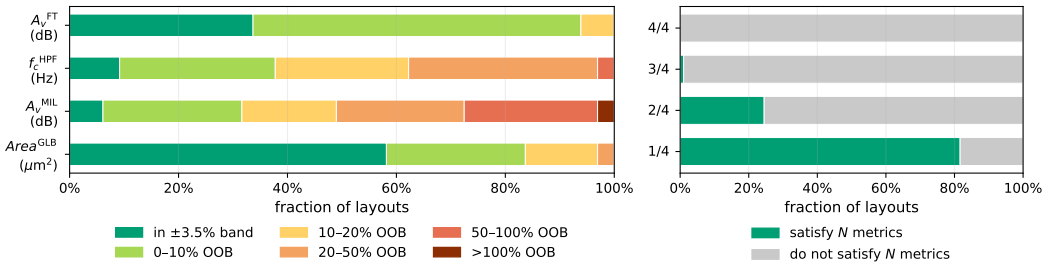
Aspect	H-OSIRIS	Flat baseline
Circuit representation	Circuit decomposed into interconnected functional blocks.	Complete circuit represented as one unpartitioned optimization unit.
Manager representation	Graph with one node per functional block. Manager selects a block and adjusts its allocated width and height bounds.	Single-node Manager graph representing the complete circuit.
Worker configuration	One reconfigurable Worker configuration is loaded for the block selected by the Manager and applies local ϵ-sizing.	One circuit-wide Worker configuration applies ϵ-sizing to the complete circuit.
Electrical and area objectives	Metric assignments and global-area constraint specified by the selected metric set.	Identical metric assignments and global-area constraint.
Training budget	Completed Manager-episode count reported for each metric set in Table 4.	Same completed-episode budget as the corresponding hierarchical run.
Held-out evaluation	100 held-out starting sizings per topology-metric set pair.	The same 100 held-out starting sizings.
Layout evaluation	Complete-circuit place-and-route, LVS, parasitic extraction, and post-layout simulation after each evaluated action.	The same complete-circuit place-and-route, LVS, parasitic extraction, and post-layout simulation flow.
Terminal feasibility	All configured electrical done bounds and the global-area done bound must hold simultaneously.	The identical joint terminal-feasibility criterion.

A.4.2 Initial-to-Terminal Held-Out Results

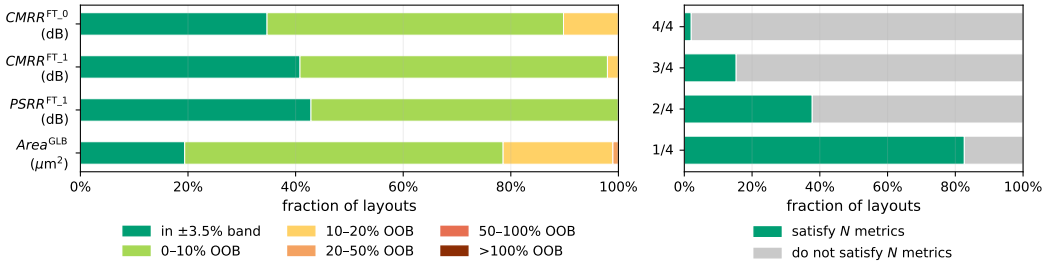
Figure 9 provides the initial-to-terminal post-layout quality metrics. Within each figure, each panel corresponds to one metric or area. Lines pair the initial and terminal extracted states of a held-out episode. Circles denote values outside the done bound and crosses denote values within it. The dashed line marks the pre-layout target. For area, the target and done-area lines are shown without a shaded band. The figures provide evidence for the aggregate outcomes in Table 3. The Amp-Filter Comp. increases joint feasibility from $0/100$ to $72/100$; the Dual-OTA policies increase it from $5/100$, $2/100$, and $1/100$ to $93/100$, $92/100$, and $92/100$ for $\mathcal{S}_{\mathrm{C}}$, $\mathcal{S}_{\mathrm{D}}$, and $\mathcal{S}_{\mathrm{E}}$, respectively. The 3-OpAmp InAmp achieves the largest increase, from $9/100$ initially feasible layouts to $99/100$ terminally feasible layouts. Across the panels, the terminal points concentrate in the done regions while area moves toward or below its target. This behavior shows that the hierarchy improves multiple coupled electrical metrics and the global area constraint jointly, rather than trading one objective for another.

A.5 Supplementary Material

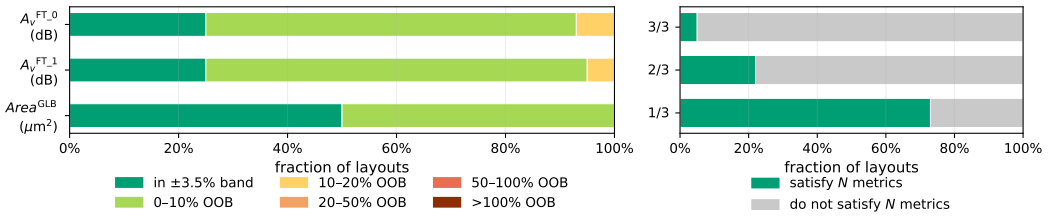
The supplementary material provides artifacts for the held-out test episodes reported in Table 3, covering the six topology-metric set pairs. For every episode, the `episodes/<topology>/<metric_set>/manager_episode_<id>/` directory contains `initial` and `terminal` states. Each state includes the schematic netlist, the layout-extracted netlist, and the generated GDS layout. These artifacts enable inspection of the physical and electrical state transitions underlying the reported held-out outcomes. The `baseline_characterization/` directory additionally provides, for each metric set, the nominal sizing snapshot, calibrated electrical constraints, and block-level and global area references used to define the target and done conditions. Together, these materials provide a common set of initial circuit states, calibrated constraints, and post-layout artifacts that future work can use to benchmark alternative analog-layout optimization methodologies against H-OSIRIS.



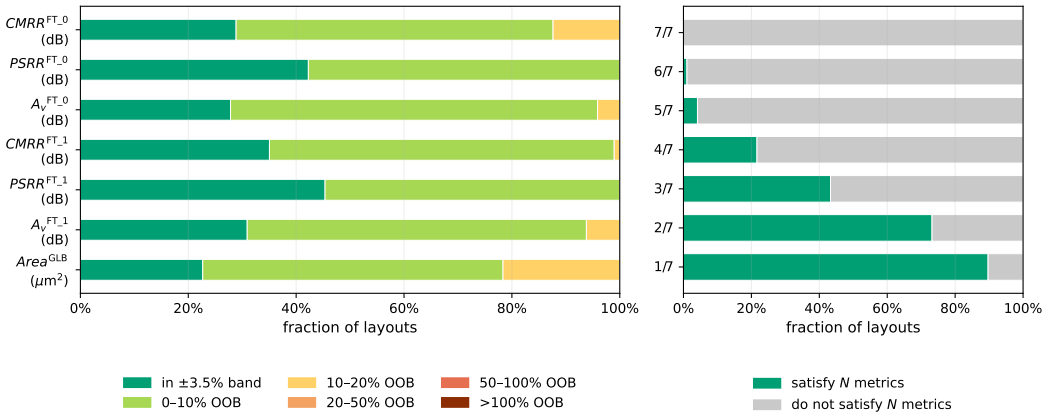
(a) Amp-Filter Comp., $\mathcal{S}_{\mathrm{F}}$.



(b) 2-OTA Comp., $\mathcal{S}_{\mathrm{D}}$.

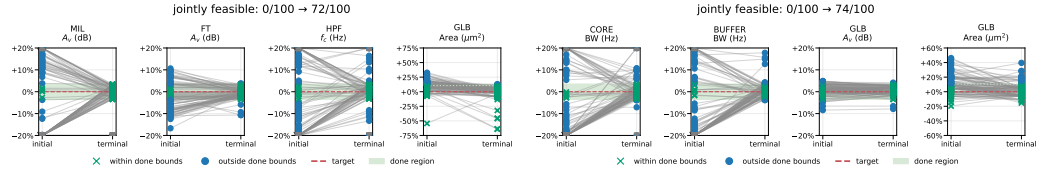


(c) 2-OTA Comp., $\mathcal{S}_{\mathrm{C}}$.



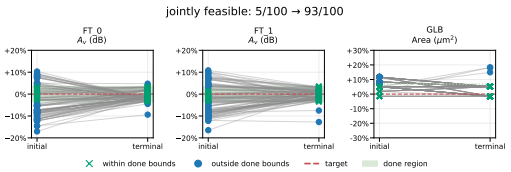
(d) 2-OTA Comp., $\mathcal{S}_{\mathrm{E}}$.

Figure 8: Percentage of post-layout evaluations over 100 randomly sampled sizing configurations within the design-space intervals. Left panels report, for each quality metric, the fraction of layouts within its $\pm 3.5\%$ done bound. Colors partition out-of-bound (OOB) layouts by normalized violation magnitude. Right panels report cumulative joint feasibility: each N/K row gives the fraction of layouts satisfying at least N of the K configured metrics, with K/K corresponding to joint feasibility.

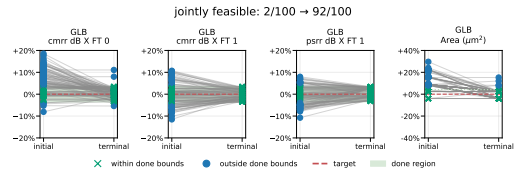


(a) Amp-Filter Comp., $\mathcal{S}_{\mathrm{F}}$

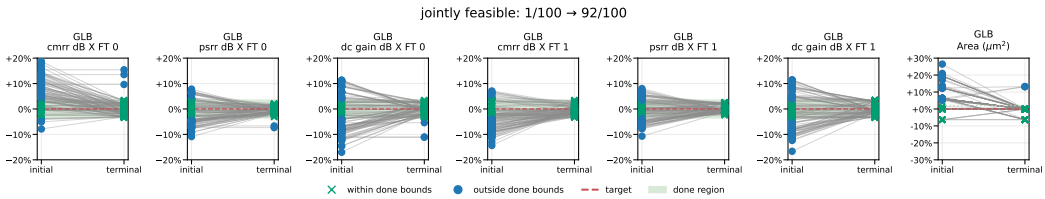
(b) 2-Stage Amp, $\mathcal{S}_{\mathrm{A}}$.



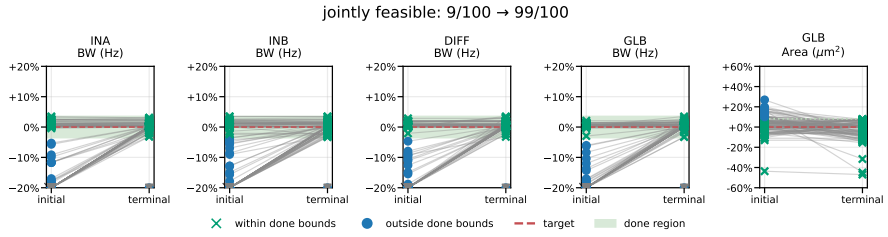
(c) Dual-OTA Comp., $\mathcal{S}_{\mathrm{C}}$



(d) Dual-OTA Comp., $\mathcal{S}_{\mathrm{D}}$



(e) Dual-OTA Comp., $\mathcal{S}_{\mathrm{E}}$



(f) 3-OpAmp InAmp, $\mathcal{S}_{\mathrm{B}}$.

Figure 9: Initial-to-terminal post-layout quality metrics across 100 held-out starts. Crosses denote states within the 3.5% done bounds, circles denote violations. Each line connects one initial to its terminal state.