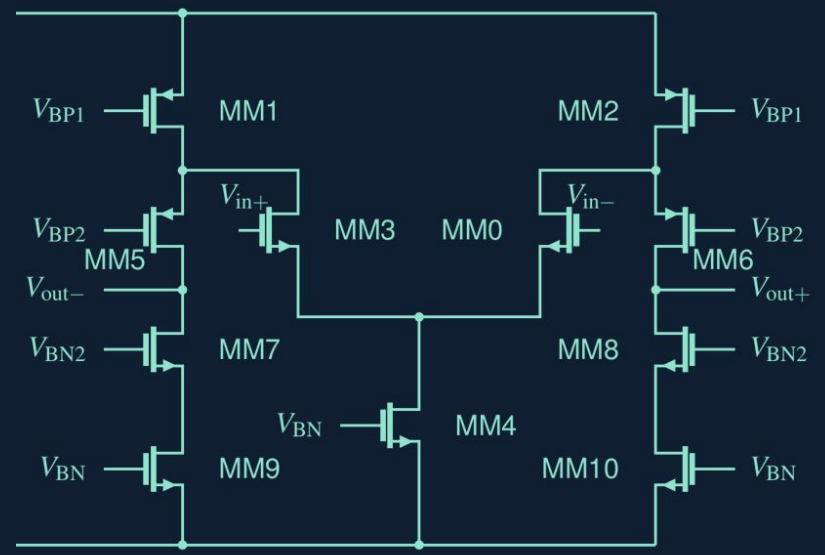
NeurIPS 2026

SLIC

Small LM Iterating on Circuits

Reinforcement Fine-Tuning Small LMs for Multi-Turn Analog Circuit Optimization

Local 4B language model for simulator-guided optimization



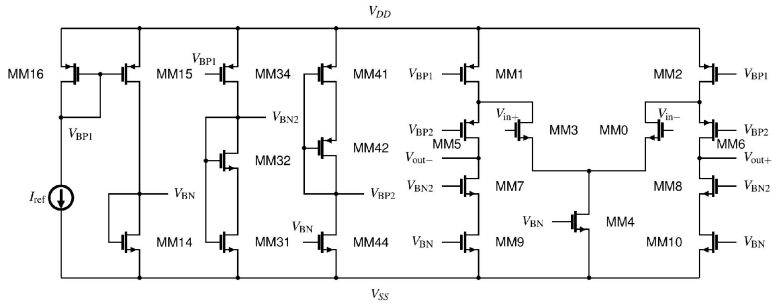
FOLDED-CASCODE AMPLIFIER CORE

Yan Xu, Zhengqi Gao, Ching-Yun Ko, Zhang-Wei Hong, Wenjie Lu, Tao Yu, Xin Zhang, Duane S. Boning and Ruonan Han

MIT / **Analog Devices, Inc.** / **IBM Research**

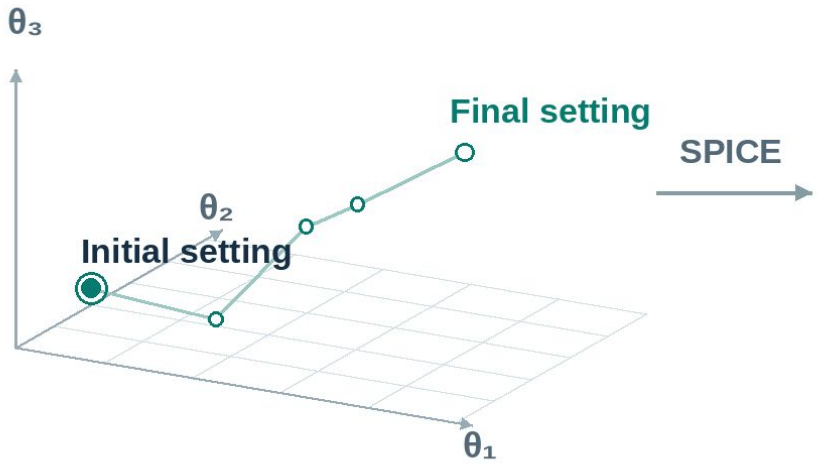
Tuning an amplifier to meet four coupled targets

Folded-cascode amplifier



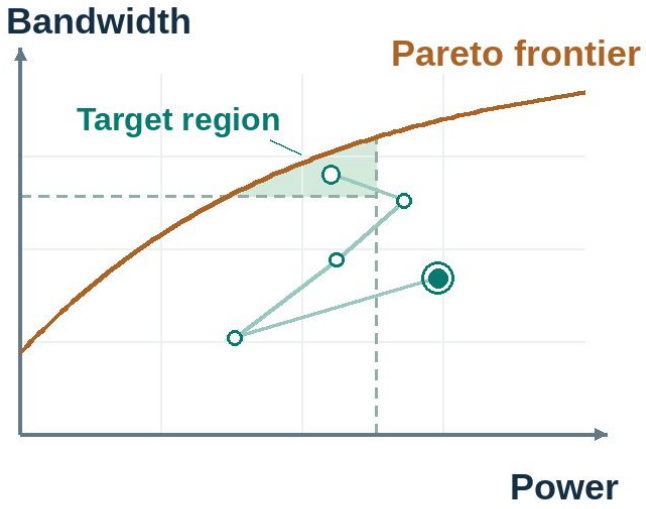
20 transistors, fixed topology

31 tunable parameters



Schematic projection

Performance trade-offs



Illustrative trade-off

Design targets

Gain ↑	Bandwidth ↑	Phase margin ↑	Power ↓

Engineer's intuition and experience → Choose parameters → Simulate → Review and refine

Why train a small language model with RL?

What makes expert tuning hard to automate	SLIC design choices
Scarce expert design trajectories	**Teacher demonstrations → SFT → RL** DeepSeek + SPICE / Qwen3.5-4B student
Limited topology and fabrication process coverage	**Reasoning through iterative refinement** Measurements → diagnosis → next adjustment
Proprietary PDKs and circuit data	**Local deployment** 4B policy + in-house simulator

> **Our hypothesis: tuning strategies learned on some circuit topologies and open sourced PDKs can transfer to new topologies and fabrication processes.**

Training data and tests of transfer

Hindsight goal relabeling

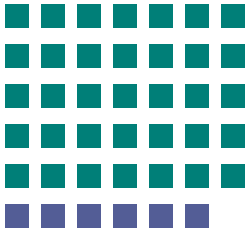
DeepSeek-V4.1-Flash explores in simulator

Each rollout's best observed performance becomes its goal

6,233 valid, improving examples for SFT

Qwen3.5-4B then trains with RL

41 amplifier topologies



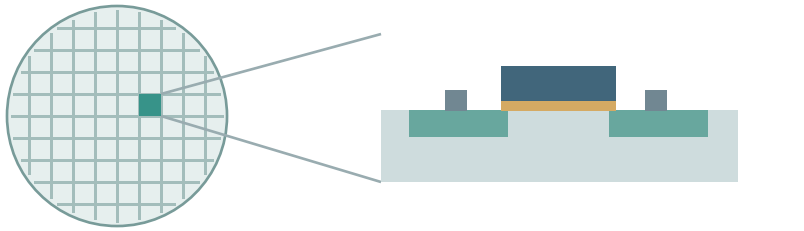
35 training topologies

840 seen-topology test tasks

6 held-out topologies

240 held-out test tasks

Fabrication processes



Silicon wafer

Transistor cross-section

Training on open source predictive technology model (PTM) 65 nm

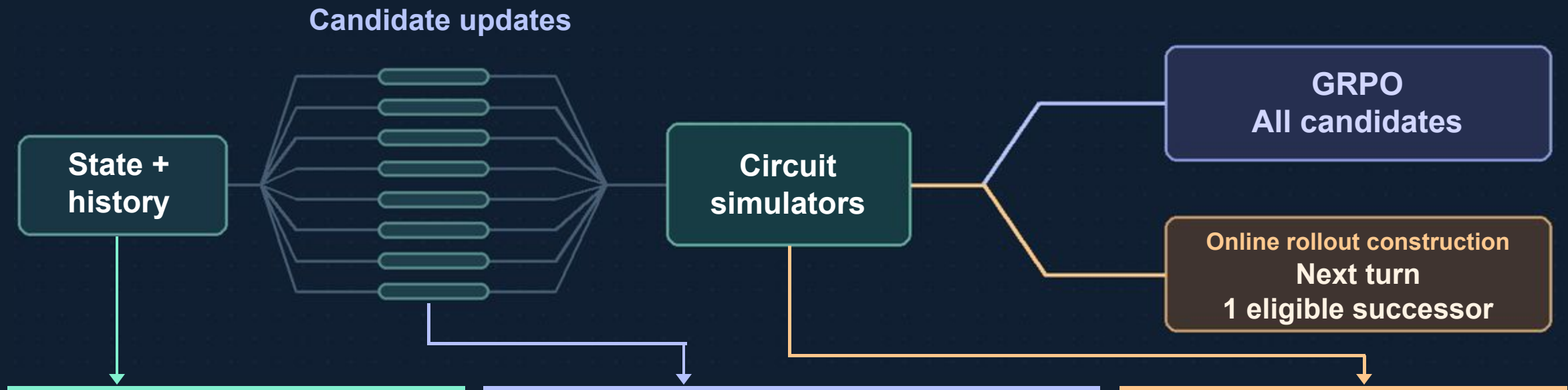
Five unseen PTM nodes: 22, 32, 45, 90, 130 nm

Commercial TSMC 65 nm CMOS

Transfer evaluation without additional fine-tuning

Policy learning from simulator feedback

One turn of amplifier tuning



MODEL INPUT

Folded-cascode topology

Devices and connections
Previous action and measurements
Bandwidth 15 kHz, target ≥ 43.1 kHz

MODEL OUTPUT

Proposed Actions (excerpt)

```
<action>decrease @10 l_MM30 70%</action>
<action>increase @10 w_MM30 10%</action>
```

ENVIRONMENT FEEDBACK

Bandwidth: 15 → 37 kHz

Input to the next turn

Continuation sampling SLIC - Random: uniform sampling / Softmax: reward-weighted sampling

Reward design and rollout expansion

Reward for extra headroom

All four targets met, power limit 100 µW

Design A — **99 µW**

Design B — **80 µW**

Binary reward ties A and B. More headroom gives B a higher slack bonus.

Eight-turn Pass@1

Random continuation with dynamic filtering [1]

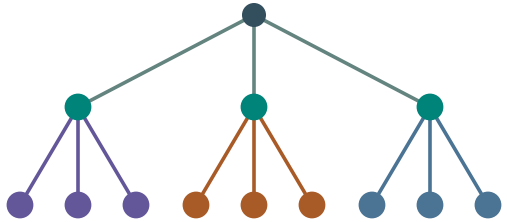
Evaluation	Without	With bonus
Seen	76.4%	84.9%
Held-out	**56.7%**	**68.8%**

Candidate expansion

Same color = candidates sampled from the same state

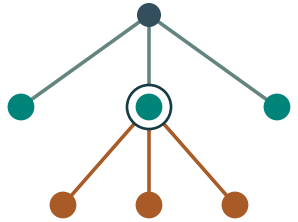
Tree MT-GRPO [2]

Expand all: exponential growth



SLIC

Advance one: linear growth



Candidate evaluations per task

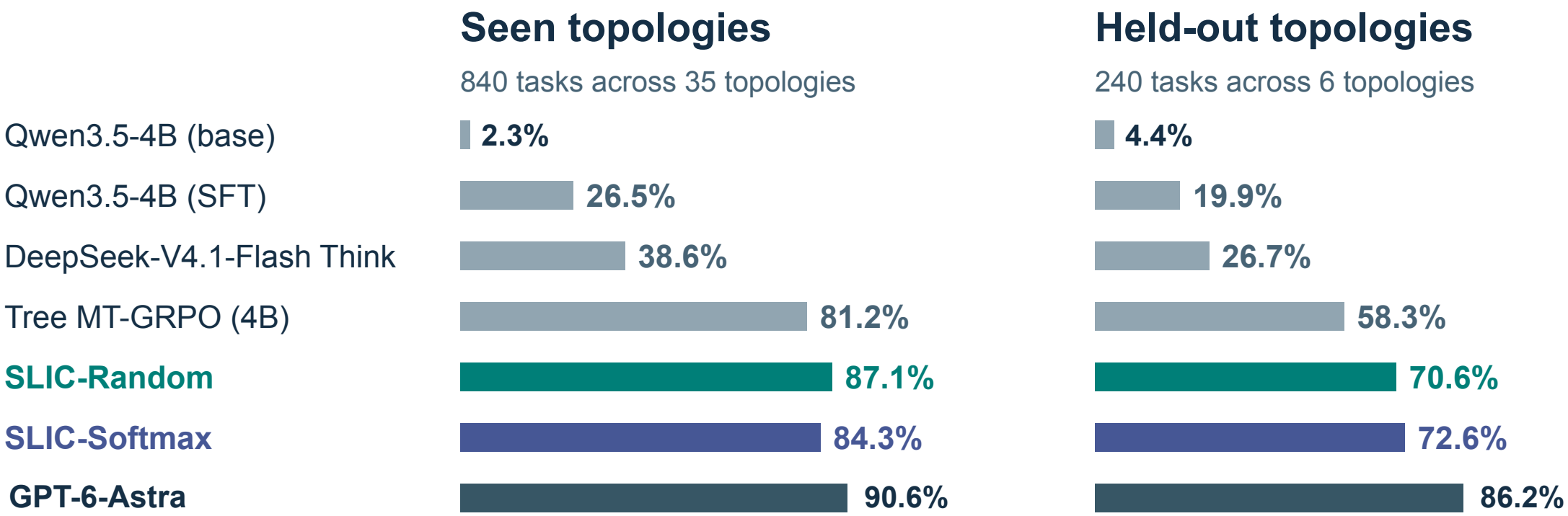
G = 3 for both methods, assuming no early stopping

Training horizon	Tree	SLIC
4 turns	**66**	**12**
8 turns	**5,466**	**24**

[1] Q. Yu et al. DAPO: An Open-Source LLM Reinforcement Learning System at Scale. NeurIPS 2025.
[2] Q. Wei et al. Reinforcing Multi-Turn Reasoning in LLM Agents via Fine-Grained Reward Structure and Credit Assignment. arXiv:2505.11821v3, 2026.

Eight-turn success on seen and held-out topologies

Four-specification Pass@1, one rollout, up to 8 turns, 5% tolerance



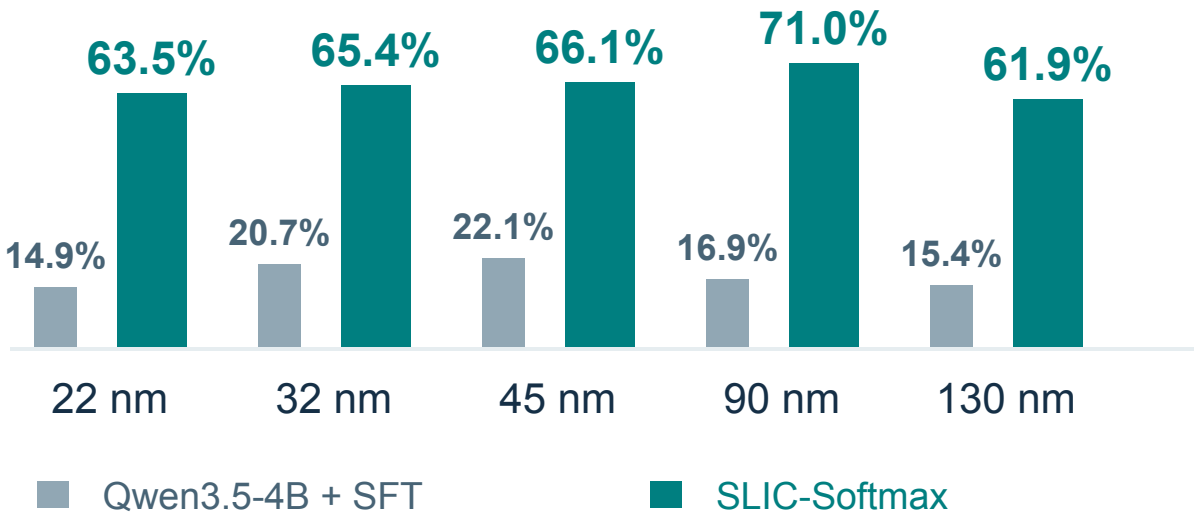
On held-out tasks, 4B SLIC-Random reaches 70.6% versus 26.7% for 763B DeepSeek Think.

Transfer across fabrication processes

Trained on PTM 65 nm, evaluated without further fine-tuning

Five unseen predictive process models

Held-out Pass@1 at 8 turns



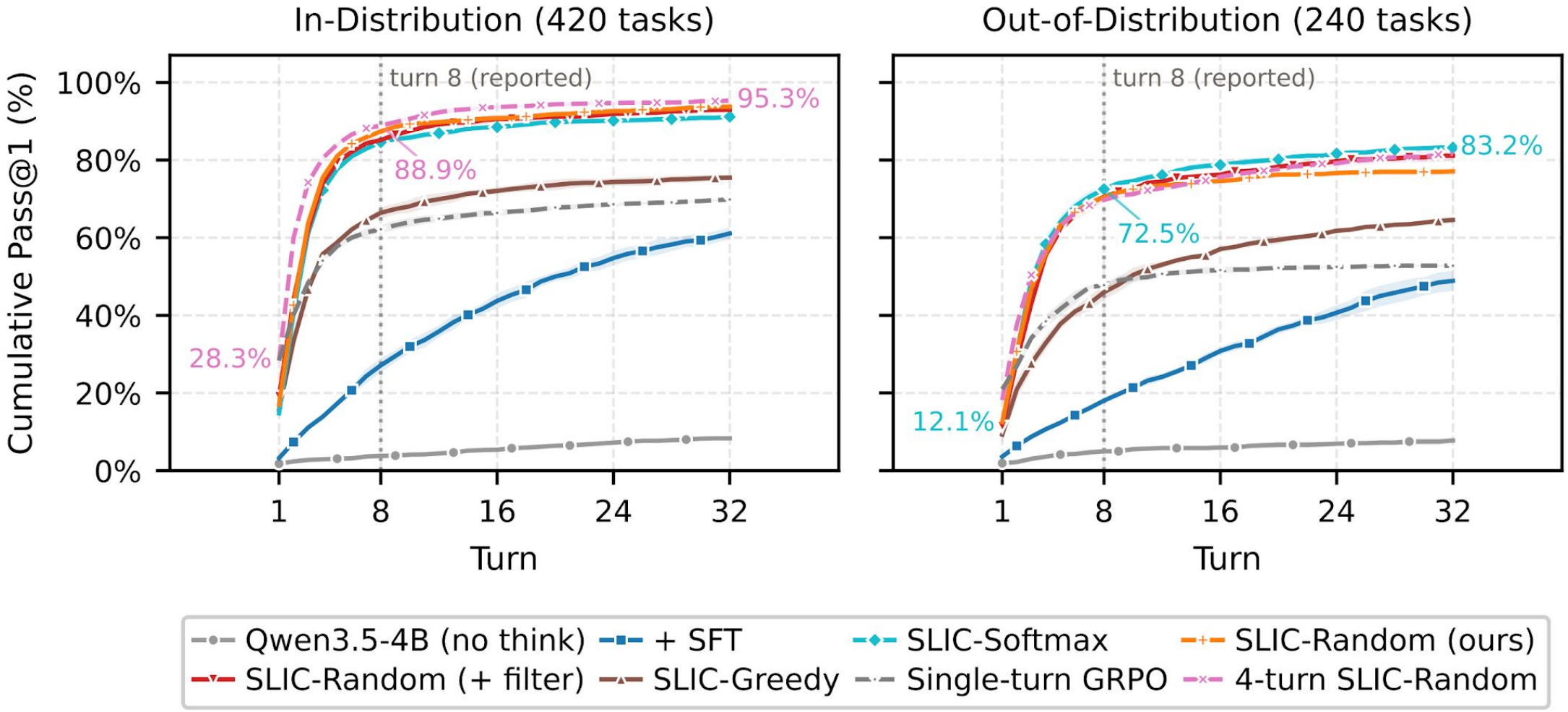
Commercial TSMC 65 nm

49.9%

Held-out Pass@1 for SLIC-Random
Qwen3.5-4B + SFT: **21.0%**
Matched PTM 65 nm baseline: 70.6%

The TSMC PDK captures device effects that PTM simplifies.
These differences may shift trade-offs and make the four PTM targets harder to meet together.

Training horizon and test-time scaling



Single-turn training saturates earlier on held-out topologies.
Four-turn SLIC-Random keeps improving at test time and remains competitive with eight-turn training.

Limitations and future work

Current scope

- Parameter optimization within a fixed amplifier topology
- Four evaluation specifications (gain, bandwidth, power, phase margin)

Future work

- Broader circuits and targets, e.g., regulators, RF amplifiers, noise, offset, and more
- Full design flow, including topology generation, optimization, layout, and verification across process, voltage, and temperature

Acknowledgements

We gratefully acknowledge primary support from **Analog Devices (ADI)** and additional funding from the MIT Cadence Fellowship and the JLO/MACOM Endowment Fund.
We thank IBM, MIT ORCD, and the AICR cluster for computing and storage resources, and MIT CICS for its support.

Paper

arxiv.org/abs/2611.13452

Code

github.com/slic-agent/slic-agent