LithoSim: A Large, Holistic Lithography Simulation Benchmark for AI-Driven Semiconductor Manufacturing

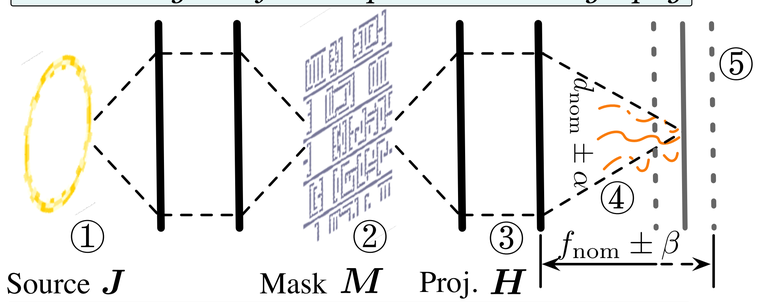
Hongquan He[1], *Zhen Wang*[1], *Jingya Wang*[1], *Tao Wu*[1], *Xuming He*[1], *Bei Yu*[2], *Jingyi Yu*[1], ***Hao Geng***[1]

[1]*ShanghaiTech University,* [2]*The Chinese University of Hong Kong*

NEURAL INFORMATION PROCESSING SYSTEMS

Introduction & Problem

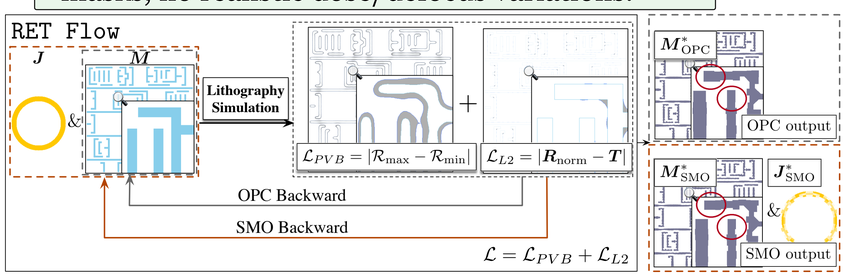
The Challenge: AI for Computational Lithography



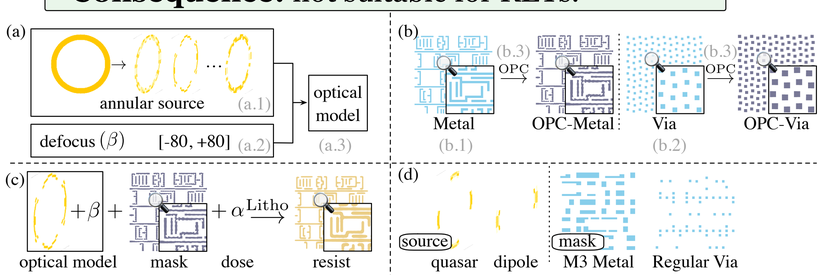
- **Key Problem:** Traditional physical simulators are too slow (>10^3 CPU-hours/mm^2), making design optimization impractical.
- **ML Promise:** Fast, data-driven surrogate models can learn the mapping **M** → **R**.

The Data Gap: Why Current Benchmarks Fail

- ❌ **Outdated:** 32-45nm nodes, not sub-28nm.
- ❌ **Small Scale:** missing proximity effects.
- ❌ **Lack of Variations:** fixed sources, no OPC masks, no realistic dose/defocus variations.

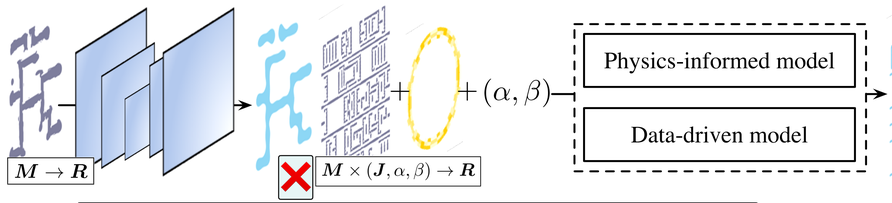


- **Consequence:** not suitable for RETs.



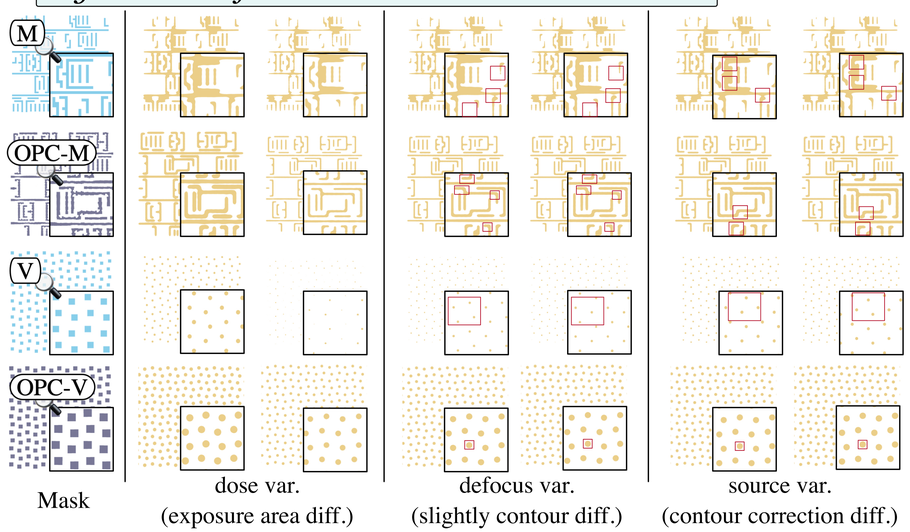
LithoSim Benchmark

Our Contribution: First Holistic Dataset



- >4 Million high-resolution input-output pairs.
- Systematically incorporates real-world variations.

Key Features of LithoSim

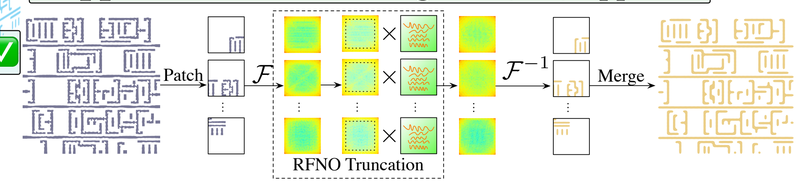


- **Diverse Masks:** Sub-28nm technology nodes.
- **Extensive Source Variations:** Over 600 distinct illumination shapes (Annular, Dipole, Quasar).
- **Realistic Process Variations:** 13 levels dose, 5 offsets defocus.
- **OOD Test Set:** Evaluates generalization to unseen mask layouts and source types.

Experiments & Results

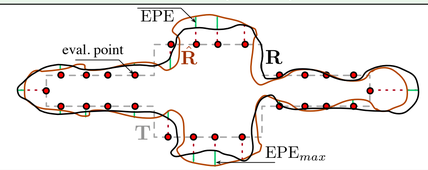
Baseline Models Evaluated

- **Data-Driven:** ED-CNN, ED-Trans
- **Physics-Informed:** FNO-based models, SOCS
- **Upper Bound:** Electromagnetic (EM) Approx.



Baseline Models Evaluated

- **AI Performance:** MSE, PA, IOU.
- **Lithographic Fidelity:** Edge Placement Error.
- **Speed:** Turnaround Time (TAT).



Result: In-Distribution Performance

- Physics-informed models and data-driven models achieve **competitive, high-fidelity simulation**, thanks to LithoSim's scale and diversity.

Result: Superior Generalization (OOD)

- **Big Finding:** SOCS, the most physics-grounded model, demonstrates the best generalization.
- **Takeaway:** physics-informed architectures are crucial for robustness.

Broader Impact

- **Enables** the development of fast, accurate, and robust ML-based lithography simulators.
- **Foundation** for CUDA-based, differentiable RET flows.