LI WEI (NEWAY) 李巍

weili3@andrew.cmu.com
Ph.D. Candidate ⋄ Department of Electrical and Computer Engineering ⋄ CMU

RESEARCH INTERESTS

- Multi-modal ML and LLMs. [DAC'20, ASPDAC'21, TCAD'22, MLCAD'23, LAD'25]
- Scalable methods for physical design and testing. [ITC'22, DAC'24]
- Optimization methods for VLSI design and testing. [TCAD'21, DAC'23]

EDUCATION

Carnegie Mellon University, PA, USA Aug. 2021 – Present
Doctor of Philosophy, Department of Electrical and Computer Engineering
Supervisor: Professor Shawn Blanton & Professor Jose Moura

The Chinese University of Hong Kong, Hong Kong
Master of Philosophy, Department of Computer Science and Engineering Aug. 2019 – July. 2021
Bachelor of Science, ELITE Stream, Department of Computer Science and Engineering Aug. 2014 – Aug. 2018
Supervisor: Professor Bei Yu

RESEARCH EXPERIENCE - IN UNIVERSITY

PhD candidate, Carnegie Mellon University, United States Sep. 2021 – Now

Multi-modal agentic LLMs for EDA

- Introduces a new paradigm where EDA agents move beyond rigid, predefined APIs and interact with a unified, multi-modal representation of the design environment.
- Outlines two challenging case studies—automated RTL debugging and logic diagnosis—to demonstrate the framework's effectiveness.
- Investigate the new problem-solving strategies that emerge from this increased agent autonomy.
- Study how this more complex interaction model impacts inference-time scaling laws, particularly the efficiency of repeated sampling.

DEFT: Differentiable Automatic Test Pattern Generation

- Reformulated ATPG as a differentiable optimization with a new reparameterization
- Custom CUDA kernel achieves 4×–26× speedup on industrial circuits
- Improved HTD detection by 21%–49% under the same pattern budget
- Support practical ATPG features such as partial assignment, generating patterns with 19.3% fewer 0/1 bits while detecting 35% more faults

Graph modality in LLMs [ICLAD'25, Best Paper Honorable Mention Award]

- Explore the graph modality integration into LLMs for VLSI
- An fully automated data collection pipeline
- Collect more than 10 billion training tokens

Global Floorplanning using Semidefinite Programming [DAC'23]

- A SDP-based method for finding the best locations of modules in a chip
- The average wirelength is reduced by at least from 3.02% to 20.01%
- The industrial case study shows **500%** quality improvement compared to the industrial tool.

Pseudo-Exhaustive Physically-Aware Region Testing [ITC'22]

- Comprehensively analyze both the physical layout and the logic netlist to identify single- or multi-output sub-circuits.
- Implemented a novel tensor-based representation of layout polygon coordinates that enables a neighborhood search strategy that reduces computational complexity from $O(n^2)$ to $O(dn)$.
- Implemented a GPU-based algorithm the physical sub-circuit extraction containing billions of sub-circuits.

GNN study in logic locking [MLCAD'23]

- Modeled their ability to identify circuit changes that stem from a logic lock as the ability to decide the isomorphism between logic netlists.
- Showed that GNNs are always upper bounded by heterogeneous Weisfeiler Lehman test in deciding the netlist isomorphism, and gave the conditions when GNNs reach the bound.

MPhil Student, The Chinese University of Hong Kong, Hong Kong Aug. 2019 – June. 2019

Routing Tree Construction [ASP-DAC'21, Best Paper Award]

- Formalized special properties of the point cloud for the routing tree construction with theoretical proof.
- Proposed an adaptive flow, which used the cloud embedding obtained by a specifically-designed model based on special properties, to select the best approach and predict the best parameter.
- Outperformed previous methods by a large margin yet being extensible and flexible.

Adaptive Layout Decomposition [DAC'20, TCAD'21]

- Proposed an adaptive workflow for efficient decomposer selection and graph matching using graph embeddings.
- Designed a graph library construction algorithm based on graph embeddings for small graphs excluding isomorphic ones.
- Reduced the runtime by 87.7% while still preserving the optimality compared with the ILP-based decomposer.

Research Assistant, The Chinese University of Hong Kong, Hong Kong Feb. 2019 – July. 2019

Open-source Layout Decomposition Framework [TCAD'21]

- Presented an open-source layout decomposition framework, with efficient implementations of various state-of-the-art simplification and decomposition algorithms.
- Discovered a set of issues of previous algorithms and proposed corresponding solutions.

Acceleration and Compression of DNNs [ICTAI'19, Best Student Paper Award]

- Proposed a unified framework to compress CNNs by combining both lowrankness and sparsity.
- Compressed the model with up to 4.9× reduction of parameters at a cost of little loss.

Research Assistant, Southern University of Science and Technology, China June. 2018 – Jan. 2019

Testing of Auto-driving Systems [ICSE'20]

- Introduced a joint optimization method to systematically generate adversarial perturbations to mislead steering of an autonomous driving system physically.
- Demonstrated the possibility of continuous physical-world tests for auto-driving scenarios as the first study.

Fault Localization [ISSTA'19, Distinguished Paper Award]

- Proposed a hierarchical DL approach to automatically learn the most effective features for precise fault localization.
- Significantly outperformed state-of-the-art with over 20% improvement.

RESEARCH EXPERIENCE - IN INDUSTRY

Intern, SoC Physical Design Group, Apple, United States May. 2024 – Aug. 2024

Floorplan Encoder

- Propose a novel floorplan encoder for the floorplanning task, the encoder is capable of encode the floorplan state, which is multi-modal, and includes multi-objects.
- Achieves 95% accuracy and shows 3X speedup to achieve the same quality compared to the industrial tool.

Research Intern, Nvidia, United States May. 2023 – Aug. 2023

Differentiable Global Routing [DAC'24]

- A differentiabl eglobal router capable of concurrent optimization for millions of nets
- Reduced nets with overflow by more than 80%

Intern, SoC Physical Design Group, Apple, United States June. 2022 – Sep. 2022

Exploration of GNNs for Physical Design

- Implemented a basic GNN model for predicting holder buffer before routing.
- Tried different methods: path-based, sub-circuit based, sub-graph based.

Perfect Rectilinear Floorplanning

- A Simulated Annealing based algorithm for perfect rectilinear floorplanning.
- Reinforcement learning, and supervised-learning that guides SA are also explored.
- **Employed in the industrial tool for all Apple chips starting from 2023.**

TEACHING ASSISTANT

Spring 2023	CMU	18202 Mathematical Foundations of Electrical Engineering
Fall 2022,2023,2024,2025	CMU	18765 Digital System Testing and Testable Design
Spring 2020	CUHK	CENG3420 Computer Organization and Design
Spring 2021	CUHK	CENG2030 Fundamentals of Embedded Systems

SELECTED AWARDS AND HONORS

Best Paper Honorable Mention Award	ICLAD	2025
Jack and Mildred Bowers Scholarship	CMU	2025
Qualcomm Innovation Fellowship	Qualcomm Inc.	2024
Apple PhD fellowship in Integrated Systems	Apple Inc.	2024
Apple PhD fellowship in Integrated Systems	Apple Inc.	2022
Faculty Outstanding Thesis Award	Engineering Faculty, CUHK	2021
Dean's Fellowship	CMU	2021
Talent Development Scholarship	HKSAR Goverment	2021
Best Paper Award	ASP-DAC	2021
1st Place Award in EDA elite challenge	Chinese Institute of Electronics	2020
Richard Newton Young Student Fellow	DAC	2020
Best Student Paper Award	ICTAI	2019
Distinguished Paper Award	ISSTA	2019
Full Postgraduate Studentship	CUHK	2019-2021
2nd Place Award in CAD Contest	ICCAD	2018
ELITE Stream Student Scholarship	Faculty of Engineering, CUHK	2018
Undergraduate Admission Scholarship	Soong Ching Ling Foundation	2015-2018

PATENTS

[p2] **Wei Li**, Rongjian Liang, Haoxing Ren. "Differentiable Global Router". U.S. Patent Application 18/965,415, Filed Dec 2, 2024. Assigned to NVIDIA Corp. - Patent Pending.

[P1] **Wei Li**, Ronald Blanton, Jose M. F. Moura. "System and Method for Global Floorplanning via Semidefinite Programming". U.S. Patent Application 18/910,938 (Filed: Oct 9, 2024) —Patent Pending.

PUBLICATIONS

Journal Papers

[J2] **Wei Li**, Yuzhe Ma, Yibo Lin, Bei Yu, "Adaptive Layout Decomposition with Graph Embedding Neural Networks", IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (**TCAD**).

[J1] **Wei Li**, Yuzhe Ma, Qi Sun, Zhang Lu, Yibo Lin, Iris Hui-Ru Jiang, Bei Yu, David Z. Pan, "OpenMPL: An Open Source Layout Decomposer", IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (**TCAD**).

Conference Papers

[18] **Wei Li**, Yang Zou, Yixin Liang, Shawn Blanton and José Moura, "DEFT: Differentiable Framework for Circuit Testing", submitted to ACM/IEEE Design Automation Conference (DAC), 2026.

[C17] **Wei Li**, Yang Zou, Christopher Ellis, Ruben Purdy, Shawn Blanton and José Moura, "BRIDGE: Bridging Graph and Large Language Models in EDA", IEEE International Conference on LLM-Aided Design (**ICLAD**), 2025. **(Best Paper Honorable Mention Award, 1/94)**

[C16] Chris Nigh, Ruben Purdy, **Wei Li**, Subhasish Mitra, R.D. Blanton, " IC-PEPR: PEPR Testing Goes Intra-Cell", IEEE International Test Conference (**ITC**), 2025.

[C15] Ruben Purdy, Chris Nigh, **Wei Li**, R.D. Blanton, "CHEF: CHaracterizing Elusive Logic Circuit Failure" , IEEE VLSI Test Symposium (**VTS**) 2025.

[C14] Chris Nigh, Ruben Purdy,**Wei Li**, Subhasish Mitra, R.D. Blanton, "Faulty Function Extraction for Defective Circuits" , IEEE European Test Symposium (**ETS**) 2024.

[C13] **Wei Li**, Rongjian Liang, Anthony Agnesina, Haoyu Yang, Chia-Tung Ho, Anand Rajaram, Haoxing Ren, "DGR: Differentiable Global Routing", ACM/IEEE Design Automation Conference (**DAC**), San Francisco, 2024.

[C12] **Wei Li**, Ruben Purdy, Jose Moura, Shawn Blanton, "Characterize the ability of GNNs in attacking logic locking", ACM/IEEE Workshop on Machine Learning for CAD (**MLCAD**), Snowbird, Utah, Sep. 11–13, 2023.

[C11] **Wei Li**, Fangzhou Wang, Jose Moura, Shawn Blanton, "Global floorplanning via semidefinite programming", ACM/IEEE Design Automation Conference (**DAC**), San Francisco, July 9-13, 2023.

[C10] **Wei Li**, Chris Nigh, Danielle Duvalsaint, Subhasish Mitra, R.D. Blanton, "PEPR: Pseudo-Exhaustive Physical Region Testing", IEEE International Test Conference (**ITC**), Sep. 25 - Sep. 30, 2022.

[C9] **Wei Li**, Guojin Chen, Haoyu Yang, Ran Chen, Bei Yu, "Learning Point Clouds in EDA", ACM International Symposium on Physical Design (**ISPD**), Mar. 21–Mar. 24, 2021.

[C8] **Wei Li**, Yuxiao Qu, Gengjie Chen, Yuzhe Ma, Bei Yu, "TreeNet: Deep Point Cloud Embedding for Routing Tree Construction", IEEE/ACM Asian and South Pacific Design Automation Conference (**ASP-DAC**), Tokyo, Jan. 18–21, 2021. **(Best Paper Award)**

[C7] **Wei Li**, Jialu Xia, Yuzhe Ma, Jialu Li, Yibo Lin, Bei Yu, "Adaptive Layout Decomposition with Graph Embedding Neural Networks", ACM/IEEE Design Automation Conference (**DAC**), San Francisco, July 19-23, 2020.

[C6] Husheng Zhou, **Wei Li**, Yuankun Zhu, Yuqun Zhang, Bei Yu, Lingming Zhang, Cong Liu, "DeepBillboard: Systematic Physical-World Testing of Autonomous Driving Systems", ACM/IEEE International Conference on Software Engineering (**ICSE**), Seoul, May 23–29, 2020.

[C5] Yuzhe Ma, Zhuolun He, **Wei Li**, Tinghuan Chen, Lu Zhang, Bei Yu, "Understanding Graphs in EDA: From Shallow to Deep Learning", ACM International Symposium on Physical Design (**ISPD**), Taipei, Mar. 25–Apr. 01, 2020.

[C4] Yuzhe Ma, Ran Chen, **Wei Li**, Fanhua Shang, Wenjian Yu, Minsik Cho, Bei Yu, "A Unified Approximation Framework for Deep Neural Networks", The IEEE International Conference on Tools with Artificial Intelligence (**ICTAI**) 2019. **(Best Student Paper Award)**

[C3] **Wei Li**, Yuzhe Ma, Qi Sun, Yibo Lin, Iris Hui-Ru Jiang, Bei Yu, David Z. Pan, "OpenMPL: An Open Source Layout Decomposer", IEEE International Conference on ASIC (**ASICON**), Chongqing, China, Oct. 29–Nov. 1, 2019.

[C2] Xia Li, **Wei Li**, Yuqun Zhang, Yuqun Zhang, Lingming Zhang, "DeepFL: Integrating Multiple Fault Diagnosis Dimensions for Deep Fault Localization",The ACM SIGSOFT International Symposium on Software Testing and Analysis (**ISSTA**), 2019. **(Distinguished Paper Award)**

[C1] Bentian Jiang, Xiaopeng Zhang, Ran Chen, Gengjie Chen, Peishan Tu, **Wei Li**, Evangeline F. Y. Young, Bei Yu, "FIT: Fill Insertion Considering Timing", ACM/IEEE Design Automation Conference (**DAC**), Las Vegas, NV, June 2-6, 2019.

TECHNICAL SKILLS

Languages	Mandarin, Cantonese, English
Programming Languages	C/C++, Python, LaTeX