

# Soo Min Kwon

---

## CONTACT INFORMATION

Email: [soominkwon0402@gmail.com](mailto:soominkwon0402@gmail.com)  
Website: [soominkwon.github.io](https://soominkwon.github.io)

Github: [github.com/soominkwon](https://github.com/soominkwon)  
LinkedIn: [linkedin.com/in/soominkwon](https://linkedin.com/in/soominkwon)

## EDUCATION


**University of Michigan** Ann Arbor, MI  
Ph.D., Electrical Engineering and Computer Science Sept. 2022 – April 2026 (Expected)

- Thesis: “Deep Learning through Low-Dimensional Representations: Theory and Algorithms”
- Advisors: Prof. Laura Balzano and Prof. Qing Qu


**Rutgers University** New Brunswick, NJ  
M.S., B.S., Electrical and Computer Engineering Sept. 2020 – May 2022

- Advisor: Prof. Anand D. Sarwate
- Minor: Mathematics


## WORK EXPERIENCE

**Student Researcher** Aug. 2025 – Nov. 2025  
 Google New York, NY

- Developed CoDistill-GRPO, a novel variant of GRPO aimed at improving small language model performance via co-distillation with larger models
- Validated the effectiveness of CoDistill-GRPO on the Qwen and Llama model families for models up to 8B parameters

**Applied Scientist Intern** Aug. 2024 – Nov. 2024  
 Amazon Seattle, WA

- Developed a causal inference framework using deep learning methods for the SCOT team that reduced variance estimates by over 10%

**Applied Research Data Science Intern** May 2022 – Aug. 2022  
 LinkedIn Corporation Sunnyvale, CA

- Productionized a machine learning pipeline for the infrastructure team, reducing MAPE by over 15% in forecasting hardware needs for the next calendar year

## PUBLICATIONS († EQUAL CONTRIBUTION)

- [1] **S. M. Kwon**<sup>†</sup>, A. S. Xu<sup>†</sup>, C. Yaras, L. Balzano, Q. Qu. “Out-of-Distribution Generalization of In-Context Learning: A Low-Dimensional Subspace Perspective”. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2026. [[Online](#)]
- [2] L. Balzano, T. Ding, B. D. Haeffele, **S. M. Kwon**, Q. Qu, P. Wang, Z. Wang, C. Yaras. “An Overview of Low-Rank Structures in the Training and Adaptation of Large Models”. In *IEEE Signal Processing Magazine*, 2025 ( $\alpha$ - $\beta$  Order). [[Online](#)]
- [3] A. Ghosh<sup>†</sup>, **S. M. Kwon**<sup>†</sup>, R. Wang, S. Ravishankar, Q. Qu. “Learning Dynamics of Deep Matrix Factorization Beyond the Edge of Stability”. In *International Conference on Learning Representations (ICLR)*, 2025. [[Online](#)]
- [4] C. Lee, **S. M. Kwon**, Q. Qu, H. Lee. “BLAST: Block-Level Adaptive Structured Matrices for Efficient Deep Neural Network Inference.” In *Neural Information Processing Systems (NeurIPS)*, 2024. [[Online](#)]
- [5] **S. M. Kwon**, L. Ding, L. Balzano, Q. Qu. “On the Relationship Between Small Initialization and Flatness in Deep Networks.” In *International Conference on Learning Representations (ICLR) Workshop on Bridging the Gap Between Practice and Theory in Deep Learning*, 2024.

- [6] **S. M. Kwon**, Z. Zhang, D. Song, L. Balzano, Q. Qu. “Efficient Compression of Overparameterized Deep Models.” In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2024. [[Online](#)]
- [7] B. Song<sup>†</sup>, **S. M. Kwon**<sup>†</sup>, Z. Zhang, X. Hu, Q. Qu, L. Shen. “Solving Inverse Problems with Latent Diffusion Models via Hard Data Consistency.” In *International Conference on Learning Representations (ICLR)*, 2024 (**Spotlight, Top 5%**). [[Online](#)]
- [8] D. K. Saha, V. Calhoun, **S. M. Kwon**, A. D. Sarwate, R. Saha, S. Plis. “Federated, Fast, and Private Visualization of Decentralized Data”. In *International Conference on Machine Learning (ICML) Workshop on Federated Learning*, 2023. [[Online](#)]
- [9] **S. M. Kwon**, X. Li, A. D. Sarwate. “Low-Rank Phase Retrieval with Structured Tensor Models.” In *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2022. [[Online](#)]
- [10] D. K. Saha, V. D. Calhoun, Y. Du, Z. Fu, R. Panta, **S. M. Kwon**, A. D. Sarwate, S. M. Plis. “Privacy-Preserving Quality Control of Neuroimaging Datasets in Federated Environments”. In *Organization for Human Brain Mapping (OHBM)*, 2021. [[Online](#)]
- [11] **S. M. Kwon**, A. D. Sarwate. “Learning Predictors from Multidimensional Data with Tensor Factorizations”. In *Rutgers University Aresty Undergraduate Research Journal*, 2021. [[Online](#)]
- [12] **S. M. Kwon**, S. Yang, J. Liu, X. Yang, W. Saleh, S. Patel, C. Mathews, Y. Chen. “Hands-Free Human Activity Recognition Using Millimeter-Wave Sensors”. In *IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*, 2019. [[Online](#)]

PREPRINTS

- [1] **S. M. Kwon**, Himanshu Jain, Ziteng Sun, Ananda Theertha Suresh, Sanjiv Kumar. “CoDistill-GRPO: A Co-Distillation Recipe for Efficient Group Relative Policy Optimization”. Submitted to *Conference on Language Modeling (COLM)*, 2026. [[Online](#)]
- [2] **S. M. Kwon**, Alec S. Xu, Can Yaras, Dogyoon Song, Laura Balzano, Qing Qu. “The Effect of Training Task Diversity on In-Context Learning through the Lens of Low-Dimensional Subspaces”. Submitted to *Journal of Machine Learning Research (JMLR)*, 2026.
- [3] **S. M. Kwon**<sup>†</sup>, C. Blocker<sup>†</sup>, H. Raja, J. Fessler, L. Balzano. “Dynamic Subspace Estimation from Undersampled Data using Grassmannian Geodesics”. Submitted to *Transactions on Machine Learning Research (TMLR)*, 2026.
- [4] X. Li, **S. M. Kwon**, I. Alkhouri, S. Ravishankar, Q. Qu. “Decoupled Data Consistency for Solving General Inverse Problems with Diffusion Models.” Submitted to the *IEEE Journal of Selected Topics in Signal Processing (JSTSP)*, 2025. [[Online](#)]

AWARDS & HONORS

Harvey G. and Joyce H. Behner Graduate Fellowship	2024
University of Michigan PhD Rackham Merit Fellowship	2023
Rutgers ECE Outstanding Master’s Student Award	2022
Rutgers ECE Outstanding Teaching Assistant Award	2021
Rutgers ECE Departmental Leadership & Service Award	2020
Rutgers WINLAB GA/TA Grant	2020 – 2020
Rutgers University Dean’s List	2018 – 2020

INVITED TALKS

2025 INFORMS Annual Meeting	2025
Math Machine Learning Seminar @ MPI MIS + UCLA	2025