

EDUCATION

- University of Colorado Boulder** Boulder, CO
M.S. in Applied Mathematics 2019–2022
Ph.D. in Machine Learning and Scientific Computing, Advisor: Jed Brown 2019–2024
– Dissertation: “Integrating Numerical and Physical Insights for Scientific Deep Learning”
- Northwestern University** Evanston, IL
B.A. in Economics and International Studies 2009–2013
– Capstone: “Analyzing the Impact of *Growth in a Time of Debt* on the Public Discourse on Austerity with Document Similarity Methods”

WORK EXPERIENCE

- Flatiron Institute**, New York, NY, *Research Analyst* 6/2023–Current
– Leads efforts on multiple physics pretraining as part of Polymathic AI collaboration. These efforts are geared towards developing foundation models for physical dynamics by developing architectures and training approaches for learning to forecast in multiple spatiotemporal physical systems simultaneously.
- University of Colorado Boulder**, Boulder, CO, *Research Assistant* 8/2019–6/2024
– Performs research at the intersection of Machine Learning and Scientific Computing. Current projects focus on integrating prior knowledge in the form of differential equations, conservation laws, and invariances into scientific deep learning methods to improve data efficiency and extrapolation capabilities.
- Lawrence Berkeley National Lab**, Berkeley, CA, *Graduate Researcher* 5/2022–1/2023
– Examined the failure scenarios in deep learning based global weather models from the perspective of numerical methods and spectral analysis. Identified several key modes of failure and developed architectural improvements to address them resulting in significantly more stable forecasts enabling architectures originally designed for medium-range forecasts (10 days) to be used coherently for seasonal forecasts (3 months) or longer without significant computational cost increases.
- Argonne National Lab**, Lemont, IL, *Givens Associate* 5/2020–7/2020
– Developed multi-device RL-based framework for learning to optimize in mixed-precision in JAX.
- Ayasdi**, Menlo Park, CA, *Data Scientist* 11/2016–7/2019
– Designed and implemented Machine Learning and Topological Data Analysis pipelines for healthcare, government, and financial services clients using both Ayasdi software and the scientific Python ecosystem.
– Collaborated with product and engineering teams in research and development capacity to develop new platform capabilities and prioritize candidate features for development.
– Developed an improved optimization framework for a non-convex unsupervised learning problem that exploited the structure of the problem to reduce run-time of a widely used software component to a quarter of the original execution time.
- Eccella**, New York, NY, *Senior Consultant* 2/2015–11/2016
– Architected large scale data integration projects involving 30+ heterogeneous systems and multiple business organizations from scoping through design, development, testing, and production support.
– Developed automation tools in Python and Java to integrate with enterprise ETL platforms in order to reduce development time on certain data flow mappings by approximately 80%.

- Implemented custom C# solutions to tailor software to specific customer needs, particularly with regard to Health Information Exchange integration and cross-enterprise secure messaging.
- Completed more than twenty projects as a lead consultant within first year including managing product betas at multiple hospitals working closely with engineering team to identify potential defects and enable clients to hit Meaningful Use metrics with minimal disruption of existing workflows.

PUBLICATIONS

- [1] S. Golkar, M. Pettee, M. Eickenberg, A. Bietti, M. Cranmer, G. Krawezik, F. Lanusse, **M. McCabe**, R. Ohana, L. Parker, B. R.-S. Blancard, T. Tesileanu, K. Cho, and S. Ho, “Xval: A continuous number encoding for large language models”, in *NeurIPS 2023 AI for Science Workshop*, 2023.
- [2] F. Lanusse, L. Parker, S. Golkar, A. Bietti, M. Cranmer, M. Eickenberg, G. Krawezik, **M. McCabe**, R. Ohana, M. Pettee, B. R.-S. Blancard, T. Tesileanu, K. Cho, and S. Ho, “AstroCLIP: Cross-modal pre-training for astronomical foundation models”, in *NeurIPS 2023 AI for Science Workshop*, 2023.
- [3] **M. McCabe**, B. R.-S. Blancard, L. Parker, R. Ohana, M. Cranmer, A. Bietti, M. Eickenberg, S. Golkar, G. Krawezik, F. Lanusse, M. Pettee, T. Tesileanu, K. Cho, and S. Ho, “Multiple physics pretraining for physical surrogate models”, in *NeurIPS 2023 AI for Science Workshop*, Accepted Oral, 2023, https://github.com/PolymathicAI/multiple_physics_pretraining.
- [4] **M. McCabe**, P. Harrington, S. Subramanian, and J. Brown, “Towards stability of autoregressive neural operators”, *Transactions on Machine Learning Research*, 2023, ISSN: 2835-8856, https://github.com/mikemccabe210/stabilizing_neural_operators.
- [5] R. Koki, A. Sheehan, E. Bradley, **M. McCabe**, and J. Brown, “Improving deep learning seismic arrival pickers using wavelet transforms”, in *Fall Meeting 2022*, AGU, 2022.
- [6] R. Koki, J. Stewart, C. Kumler, and **M. McCabe**, “A deep learning approach for detection, semantic segmentation and density classification of smoke in satellite imagery”, in *Fall Meeting 2022*, AGU, 2022.
- [7] **M. McCabe** and J. Brown, “Learning to assimilate in chaotic dynamical systems”, in *Thirty-Fifth Conference on Neural Information Processing Systems*, 2021, <https://github.com/mikemccabe210/amortizedassimilation>.
- [8] **M. McCabe** and J. Brown, “Using Differentiable Physics for Self-Supervised Assimilation of Chaotic Dynamical Systems”, presented at the 2020 Workshop on Differentiable Vision, Graphics, and Physics in Machine Learning at NeurIPS, 2020.
- [9] **M. McCabe**, *Mapper Comparison with Wasserstein Metrics*, 2018. arXiv: 1812.06232 [cs.CG], https://github.com/mikemccabe210/mapper_comparison.

PRESENTATIONS

- **Challenges and Opportunities in Learning Physical Dynamics** at Georgia Tech University October 2023
AI/ML in Physics
- **Data Science in the Healthcare Industry** at Colorado State University February 2020
CSU Data Science Seminar
- **Graph Neural Networks in the Physical Sciences** at University of Colorado Boulder November 2019
CU Optimization and ML Seminar
- **Fine-grained Segmentation for Population Risk** at HIMSS 2018 March 2018
HIMSS Solutions Lab

TECHNICAL SKILLS

Languages: Python, SQL, C++, Java, Scala, Matlab
Frameworks: PyTorch, JAX, TensorFlow, SciPy Ecosystem, Spark
Databases: Oracle, MySQL, MS SQL Server, SQLite, Postgres, HDFS

SCHOLARSHIPS AND AWARDS

- CU Boulder CS Early Career Professional Development Fellowship 2019
- Ayasdi “Flare” Performance Award 2018
- Eccella “Eccelarator” Performance Award 2015
- Siemens “You Answered” Recognition 2014
- Nominated for Guetzkow Prize for Outstanding Senior Research Paper in International Studies 2013
- Design Award for Engineering Design and Communication 2009
- Siemens Foundation Merit Scholarship 2009–2013

SERVICE

- Co-organizer “AI-driven discovery in physics and astrophysics” Workshop 2023 2024
- Reviewer at NeurIPS 2021, 2023
- Reviewer at ICML 2021, 2022
- CU Boulder Numerical Computing Reading Group Coordinator 2020–2022
- CU Boulder CS Department Peer Mentor 2020–2021
- CU Boulder CS Department Application Feedback Mentor 2020, 2021
- Volunteer at NeurIPS 2020
- Volunteer at ICML 2020
- Volunteer at ICLR 2020