

Education

Harvard University

PhD Computer Science (Optimization & Machine Learning)

Supervised by Melanie Weber

Expected 2027

- NSERC CGS-D PhD Fellowship: \$120k
- Teaching Fellow (Graduate Course in Geometric Machine Learning)

McGill University & MILA

Thesis-Based Master's in Mathematics (Optimization & Machine Learning)

Supervised by Courtney and Elliot Paquette

2023

- **cGPA: 4.00/4.00** | **Coursework:** Advanced Probability 1 & 2, Random Matrix Theory, Convex Optimization, High-Dimensional Probability
- Graduate Excellence Scholarship (MILA): \$30,000 | FQRNT Scholarship: \$17,500 | MITACS Scholarship: \$30,000

McGill University

BSc. Joint Honours Statistics and Computer Science & Minor in Physics

2021

- **cGPA: 3.80** | **Major GPA: 3.90**
- **Master's coursework:** Time Series Analysis; Sampling Theory; Matrix Computations; Honours Linear Regression; Generalized Linear Models; Applied Machine Learning
- **PhD-level coursework:** Mathematical Techniques in Machine Learning; Computation Intensive Statistics; Probabilistic Analysis of Algorithms (achieved highest grade in all 9 graduate courses during undergrad)

Publications & Preprints

1. K. Lee, **A. Cheng**, E. Paquette, and C. Paquette. Trajectory of mini-batch momentum: Batch size saturation and convergence in high dimensions. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, editors, *Advances in Neural Information Processing Systems*, volume 35, pages 36944-36957. Curran Associates, Inc., 2022. URL
2. **A. Cheng** and M. Weber. Structured regularization for spd optimization with side information. In *2024 60-th Annual Allerton Conference on Communication, Control, and Computing*, pages 1 – 8, 2024. IEEE: <https://ieeexplore.ieee.org/abstract/document/10735290>
3. **A. Cheng**, K. Lee, C. Paquette. Exact Dynamics of First Order Stochastic Algorithms in High-Dimensions (2024) (accepted to NeurIPS 2024 MLOPT workshop and submitted to SIAM Optimization).
4. **A. Cheng**, M. Weber. *Structured Optimization on the PSD Riemannian Manifold* (2024) (accepted to NeurIPS 2024 MLOPT workshop and submitted to Mathematical Programming). <https://arxiv.org/abs/2410.09660>
5. **A. Cheng**, V. Dixit, and M. Weber. Disciplined geodesically convex programming. Manuscript under review, 2024. <https://arxiv.org/abs/2407.05261>
6. **A. Cheng**, B. Kiani, A. Radhakrishnan, and M. Weber. Emergence of low rank structure in group convolutional neural networks. Manuscript in preparation, 2025.

Talks

1. Presented two workshop papers: *High Dimensional First Order Mini-Batch Algorithms* and *Structured Regularization on the SPD Manifold* at the NeurIPS 2024 MLOPT workshop
2. Presented the link between stochastic optimization of quadratic functions and polynomials to the Montreal Optimization Seminar (2021)
3. Presented at 2022 INFORMS Annual Meeting, NeurIPS 2022, and the Montreal Optimization Seminar

Work Experience

Optimization for ML Graduate Intern

Lawrence Livermore National Lab

May - August 2025

- Large-scale optimization and theory for machine learning

Deep Learning Research Intern

Valence Labs (Start-Up Acquired in June 2023)

May 2022 - July 2022 & June 2023 - August 2023

- Designed unsupervised learning methods inspired by the notion of disentanglement to navigate the molecular space to generate novel molecular compounds
- Applied transformers and variational autoencoders in PyTorch to molecular (NLP) data sets