Matthew Buchholz

Website: buchholzmd.github.io
Email: buchholz.md@gmail.com
LinkedIn: matt-buchholz
GitHub: github.com/buchholzmd

Personal Statement

My research focuses on the mathematics of deep learning, with the goal of building a clear theoretical understanding of how neural networks approximate functions, learn features and representations, and generalize to unseen data and tasks. I use tools from applied harmonic and functional analysis, high-dimensional probability, and stochastic analysis to study approximation, optimization dynamics, and generalization in neural networks. My current research focuses on developing a mathematical understanding of modern deep networks along four interconnected directions. First, I study the function spaces naturally associated with wide and deep architectures and how these spaces lead to natural notions of regularization and inductive bias. Second, I aim to understand approximate convexity in neural networks as width and depth grow, and how such structure contributes to efficient optimization and stable convergence, particularly in large language and vision models (LLMs/LVMs). Third, I investigate the interplay of feature learning with the dynamics of gradient based stochastic optimization like SGD and ADAM with momentum, learning-rate schedules, and iterate-averaging. Finally, I am interested in quantifying the limitations of uniform convergence in capturing generalization and identifying statistical quantities that explain generalization in large-scale, self-supervised, and multi-modal settings. These directions share a common goal: to develop rigorous mathematical characterizations of training dynamics and feature learning within overparameterized models.

Education

University of British Columbia Sep 2025 – Present
Ph.D. Candidate in Computer Science
Advised by Dr. Mark Schmidt, Dr. Danica J. Sutherland, and Dr. Geoff Pleiss

New York University: Courant Institute of Mathematical Sciences Sep 2022 – May 2025
M.Sc. in Mathematics
Thesis: *On Approximation by Shallow Neural Networks*
Advised by Dr. Sinan Güntürk and Dr. Joan Bruna

- GPA: 3.4
- PhD Coursework: Foundations of Machine Learning, Bayesian Machine Learning, Mathematical Statistics, Probability Theory, Wavelets and Approximation Theory, Real Variables, Partial Differential Equations
- Master's Coursework: Linear Algebra, Multivariable Analysis, Complex Variables

California State University, Long Beach Jan 2018 – May 2021
B.Sc. Computer Science; B.A. Physics; Minor in Mathematics

- GPA: 3.9
- Graduate Coursework: Theory of Computation, Statistical Physics, Probability Theory and Random Processes for Electrical Engineering
- Undergraduate Coursework: Machine Learning, Quantum Physics, Electricity and Magnetism, Analytical Mechanics, Dynamical Systems, Abstract Algebra, Mathematical Optimization, Real Analysis, Programming Languages, Networks and Distributed Computing, Database Fundamentals, Operating Systems, Computer Security, Software Engineering

Publications

[1] Y.-M. Pun, M. Buchholz, and R. M. Gower, *Schedulers for schedule-free: Theoretically inspired hyperparameters*, 2025. arXiv: 2511.07767 [cs.LG].

[2] M. Buchholz, W. Zhang, E. N. Meese, Y. Yang, C. G. Lowe, and H.-G. Yeh, "Recurrent attentive kernel learning for shark activity recognition", *7th SIGKDD Workshop on Mining and Learning from Time Series (MiLeTS)*, 2021.

[3] S. Ho, W. Zhang, W. Young, M. Buchholz, S. Al Jufout, K. Dajani, L. Bian, and M. Mozumdar, "Dlam: Deep learning based real-time porosity prediction for additive manufacturing using thermal images of the melt pool", *IEEE Access*, vol. 9, pp. 115 100–115 114, 2021.

Honors and Awards

- **President's Academic Excellence Initiative PhD Award**: Winter 2025
- **President's list**: Spring 2018, Fall 2018, Spring 2019, Spring 2020
- **Dean's list**: Fall 2019
- **Tau Beta Pi Engineering Honor Society** (Top $1/8^{th}$ of Junior Class - CSULB Chapter)

Research Experience

Convergence Analysis of Schedule-Free SGD May 2025 – Present

Collaborators: Robert Gower and Mandy Pun *Flatiron Institute*

- Derived last-iterate convergence rate for schedule-free free SGD with a convex objective and demonstrated empirical performance of the method with optimal choice of hyperparameter with respect to the derived upper bound for deep learning. Demonstrated the predictive power of this bound in the convex case on (non-convex) deep learning problems.

Approximation Theory of Deep Neural Networks Sep 2023 – May 2025

Advisors: Dr. Sinan Güntürk and Dr. Joan Bruna *New York University*

- Literature review of neural network approximation: proof techniques for approximation upper and lower bounds, depth separations, rank separations in transformers, infinite-width networks, Barron spaces, and connections to convex geometry.

Recurrent Attentive Kernel Learning Jan 2021 – May 2021

Advisor: Dr. Wenlu Zhang *California State University, Long Beach*

- Developed a novel deep kernel learning method for multivariate time-series data, using attention mechanisms to extract global temporal features from a gated recurrent neural network decoder, where attention weights are used as weights for a mixture of Gaussian process posterior predictions, trained end-to-end via type-II maximum likelihood.
- Implemented the full machine learning pipeline, including data cleaning, feature extraction (DTW, Fourier transform, Wavelet transform, stationarity analysis/decomposition), resampling, augmentation, training, and inference, and was applied to real-world shark data collected by the CSULB Shark Lab for shark behavior classification.
- Paper accepted to KDD 2021 Workshop on Mining and Learning from Time Series.

Deep Learning for Additive Manufacturing Aug 2020 – May 2021

Advisor: Dr. Wenlu Zhang *California State University, Long Beach*

- Designed and implemented machine and deep learning algorithms to predict porosity from infrared thermal images of the melt pool in additive manufacturing, including feature extraction using functional principal component analysis and transfer learning from natural image datasets (CIFAR, ImageNet) with CNNs.

– Paper accepted as journal paper in IEEE Access.

Reviewing

- SIGKDD 2020

Selected Talks

SIGKDD Workshop on Mining and Learning from Time Series (MiLeTS), Virtual — Aug 2021
Recurrent Attentive Kernel Learning for Shark Activity Recognition

Long Beach Deep Learning Lab, Long Beach, CA — Mar 2021
Bayesian Machine Learning

CATIE Research Group, Long Beach, CA — Sep 2020
Functional Principal Component Analysis

The Aerospace Corporation, El Segundo, CA — Aug 2019
Reproducing Kernel Hilbert Space

Long Beach Deep Learning Lab, Long Beach, CA — Jun 2019
Gaussian Visual-Semantic Embedding for Joint Image-Text Representation

Work Experience

Guest Researcher — May 2025 – Present
Flatiron Institute Center for Computational Mathematics — *New York, NY*

– **Empirical performance of schedule-free SGD**: Conducted ablation study comparing momentum SGD, Polyak-Rupert averaging, Adam, and their schedule-free variants with theoretically optimal weights across vision and language modeling tasks.

Machine Learning Engineer — Sep 2021 – Jun 2023
Samsung Research America — *Mountain View, CA*
Worked on the Think Tank Team (TTT), a multi-disciplinary RD team, fast prototyping machine learning solutions, research software and hardware development, design, and product development

– **Representation learning for repetition counting in fitness application**: Used pre-trained Transformer model to learn latent representations of video signals, applying PCA and spectral features with autocorrelation for anomaly detection in latent space

– **Inverse rendering for augmented reality**: Implemented spherical convolutional neural networks with optimal transport loss function for modeling real-world light distributions

– **Point cloud registration in robotics system**: Implemented point cloud registration algorithms, namely iterative closest point (ICP) and large deformation diffeomorphic metric mapping (LDDMM) for robotics system in ROS

Machine Learning Engineer Intern — Jun 2021 – Sep 2021
Samsung Research America — *Mountain View, CA*

– **Generative models for inverse problems**: Designed a generative convolution neural network with spectral regularization for generating inverse image filters to remove hardware specific blur/diffraction. Implemented convex optimization baseline using proximal gradient methods

SDE Intern — Jun 2020 – Aug 2020
Amazon — *Seattle, WA*

– Developed cross-platform (iOS + Android) features for Kindle that improved the client's efficiency within lower-level rendering system (C++)

Software Developer Intern Jun 2019 – Aug 2019
The Aerospace Corporation *El Segundo, CA*

– Improved baseline metric by 24% for resource allocation using Genetic Algorithms

– Increased interpretability of an internal information retrieval system by fine-tuning word2vec model using t-SNE for data visualization (Python)

Application Development Intern May 2018 – Sep 2018
Siemens PLM Software *Cypress, CA*

– Redesigned backend tool that decreased in-flux of support requests by over 70%

Skills

- **Programming Languages:** Python, C++, MATLAB, SQL, JavaScript
- **Packages/Frameworks:** PyTorch, Jax, Pyro, GPyTorch, scikit-learn, NumPy, OpenCV, SciPy, ROS
- **Research Areas:** Deep Learning Theory, Gaussian Processes, Generative Models, Optimization, Inverse Problems, Computational Imaging, Approximation Theory