Berlin Chen

bc2188@princeton.edu | GitHub | Google Scholar

EDUCATION

Princeton University | Candidate for Ph.D. in Computer Science   Aug 2023 - Present

University of Cambridge | M.A.St. in Mathematics (Part III)   Oct 2020 - June 2021
- **Part III paper:** Langevin-Type MCMC Algorithms. Supervisors: Richard Nickl and Randolf Altmeyer.

Swarthmore College | B.A. in Honors Mathematics and Computer Science, GPA 3.8/4.0   Aug 2015 - May 2019

University of Oxford | Visiting Student in Computer Science and Mathematics   Jan-June 2018

RESEACH EXPERIENCE

Princeton Neuroscience Institute, Princeton University, Research Assistant   Jan 2022 - July 2023
- Developed and maintained a large (~3TB/month) data ingestion pipeline that performed automatic labeling using trained ML models such as Detectron2 and Google Automatic Speech Recognition.
- Fine-tuned object detection DNNs on new data and achieved increased accuracy upon deployment (~50% to ~80%).

Dept. of Engineering, University of Cambridge, Research Assistant   July-Sep 2021
- Designed a novel algorithm that learns the parameters of manifold-valued hidden Markov models (HMMs) in a fast and consistent manner; implemented the algorithm (code).
- Testing the algorithm on synthetic examples (e.g., Poincaré disk) and real data for Brain-Computer Interface (BCI).

NASA JPL, Caltech, Technologist I   July 2019 - Sep 2020
- Built and maintained simulators for the data acquisition/conversion process of the Atmospheric Infrared Sounder.
- Used machine learning (self-organizing maps) to identify distinct weather regimes in the Tropical Eastern Pacific.

Johns Hopkins Whiting School of Engineering, Undergraduate Researcher   June-Aug 2018
- Contributed to the toolkit (jiant) for transfer/multitask learning on sentence understanding tasks; explored the application of attention mechanism to pre-train tasks such as natural language inference.
- Implemented a novel methodology to probe various contextualized word embeddings (e.g., ELMo).

Max Planck Institute for Mathematics in the Sciences, Visiting Researcher   Mar-April 2018
- Implemented an information-theoretic framework analyzing the sensorimotor loop in the domain of morphological computation.

NASA JPL, Caltech, Caltech Summer Undergraduate Research Fellow   June-Aug 2017
- Implemented neural networks that emulate short-term atmospheric processes to efficiently explore a climate model's parameter space.

PEER-REVIEWED PUBLICATIONS

1. **Chen, B.**; Mostajeran, C.; Said, S. Geometric Learning of Hidden Markov Models via a Method of Moments Algorithm. Phys. Sci. Forum 2022, 5, 10. https://doi.org/10.3390/psf2022005010.

2. Tenney, I.; Xia, P.; **Chen, B.**; Wang, A.; Poliak, A.; McCoy, R.T.; Kim, N.; Durme, B.V.; Bowman, S.; Das, D.; et al. What do you learn from context? Probing for sentence structure in contextualized word representations. In Proceedings of the International Conference on Learning Representations, 2019.

3. Wang, A.; Hula, J.; Xia, P.; Pappagari, R.; McCoy, R.T.; Patel, R.; Kim, N.; Tenney, I.; Huang, Y.; Yu, K.; et al. Can You Tell Me How to Get Past Sesame Street? Sentence-Level Pretraining Beyond Language Modeling. In Proceedings of the Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics; Association for Computational Linguistics: Florence, Italy, 2019; pp. 4465–4476. https://doi.org/10.18653/v1/P19-1439.

OTHER PUBLICATIONS

4. Wang, A.; Tenney, I.F.; Pruksachatkun, Y.; Yeres, P.; Phang, J.; Liu, H.; Htut, P.M.; Yu, K.; Hula, J.; Xia, P.; et al. jiant 1.3: A software toolkit for research on general-purpose text understanding models. http://jiant.info/, 2019.

5. **Chen, B.** Convolutional Neural Networks as Efficient Emulators for Atmospheric Models. In Proceedings of the Caltech Undergraduate Research Journal, 2018. https://curj.caltech.edu/2018/07/11/convolutional-neural-networks-as-efficient-emulators-for-atmospheric-models. Editorially reviewed.

ACADEMIC AWARDS

- Gordon Wu Fellowship, Princeton University, awarded "to the most outstanding incoming doctoral students in engineering," 2023.
- Lockwood Fellowship, Swarthmore College, for graduate studies at Cambridge, 2020.
- Honorable Mention, NSF Graduate Research Fellowship Program, 2019.
- Research Fellowship, the Center for Language and Speech Processing (CLSP) Summer School in Human Language Technology at The Johns Hopkins University, 2018.
- First place, Mid-Atlantic USA Regional at Wilkes. International Collegiate Programming Contest, 2017.

RESEARCH TALKS

1. **Geometry, Statistics, and Hidden Markov Models: An Overview**. Uncertainty Quantification for Remote Sensing Inverse Problems. Virtual Meeting at JPL. Oct 2022.

2. **Geometric Learning of Hidden Markov Models via a Method of Moments Algorithm**. MaxEnt2022. Institut Henri Poincaré. Paris, France. July 2022.

3. **A Convex Optimization Approach to Learning Geometric HMMs**. Cambridge Math Placement Program Symposium. University of Cambridge. Cambridge, UK. Aug 2021.

4. **Transfer Learning for Image Recognition**. Computer Science Senior Poster Session. Swarthmore College. Swarthmore, PA. Nov 2018.

5. **Creating Weather System Ensembles Through Synergistic Process Modeling and Machine Learning**. 2017 AGU Fall Meeting. New Orlean, LA. Dec 2017.

SKILLS

Languages (roughly ordered by decreasing familiarity): Python, LaTeX, MATLAB, C++, C, Bash, OCaml, SQL, Go, R.

Python Packages: PyTorch, Pandas, Scikit-learn, Keras, Ax-platform, Hugging Face, Detectron2, RetinaFace.

Tools: Git, Docker, AWS (EC2, Lambda, etc.), ELK stack, Vim, Jupyter Notebook, Colab, Tensorboard, Label Studio.

Natural Languages: English and Mandarin fluent. Basic French.