

EDUCATION

- **University of Michigan, Ann Arbor** Ann Arbor, MI
Doctor of Philosophy in Statistics; *Sep. 2016 – Aug. 2022*
 - **Thesis:** On Some Approximate Inference Approaches in Population Genetics
- **Stanford University** Stanford, CA
Master of Science in Computational and Mathematical Engineering; *Sep. 2013 – July. 2015*
- **Harbin Institute of Technology** Harbin, China
Bachelor of Science in Mathematics; *Sep. 2009 – July. 2013*

RESEARCH EXPERIENCE

- **University of Michigan-Ann Arbor** Ann Arbor, MI
Graduate Researcher *Sep. 2016 - Aug. 2022*
 - **Bayesian Asymptotics of the Sequentially Markov Coalescent:** Examined the well-known pairwise sequentially Markov coalescent (PSMC) algorithm for inferring local ancestry using genetic variation data. Proved that the posterior distribution of (a slight variant of) PSMC contracts, up to log factors, at a nearly parametric rate to the underlying true parameter. Studied the ergodicity properties of the underlying SMC process itself. Proved via a coupling argument that the underlying discrete jump chain of the process is geometrically ergodic.
 - **The Solution Surface of the Li-Stephens Haplotype Copying Model:** Gave a new interpretation of the recombination rate and the mutation rate in the Li-Stephens (LS) model by appealing to the concept of machine learning. Treated the recombination rate and the mutation rate as two tuning parameters. Designed a new, efficient algorithm for finding the complete solution surface of both haploid and diploid variants of the LS algorithm by appealing to a result in the changepoint detection literature and several results from convex analysis. Demonstrated our algorithm can be viewed as characterizing the trade-off between recombination and mutation through simulation. Manuscript is available at: <https://www.biorxiv.org/content/10.1101/2022.08.03.502674v1>, Python code is available at: <https://github.com/jthlab/lss>.
 - **A Shrinkage Estimator of Population Size History:** Provided a shrinkage estimator for estimating population size history. The estimator based on sample frequency spectrum and PSMC are two popular estimators for estimating demography, however, these two estimators both have their own deficiencies. Derived a novel shrinkage estimator with variable component-wise shrinkage coefficients for estimating population size history. Demonstrated this estimator dominates two popular population size history estimators under various simulation schemes. Added a smoothness penalty to prevent overflow and make optimization easier. Using the automatic differentiation of Jax to get the gradient and the Hessian.
- **Stanford University** Stanford, CA
Research Assistant *Sep. 2014 - July. 2015*
 - **Convergence of Contrastive Divergence Algorithm:** Investigate the well-known contrastive divergence algorithm developed by Geoffrey Hinton. Proved the convergence property of the contrastive divergence algorithm which remained unknown for 20 years. Paper is published in *Annals of Statistics*: <https://projecteuclid.org/euclid.aos/1536307243>

WORK EXPERIENCE

- **Bytedance** Mountain View, CA
Research Intern *May 2020 - Aug 2020*
 - **Distributed Bayes Procedure:** Developed a novel distributed Bayes procedure that combines information from each local machine by training a deep discriminative neural network that can be treated as a black box. Compared with existing methods, it does not require distributional assumption, and data on each local machine do not need to be identically distributed. Implemented the algorithm using PyTorch.

TECHNICAL SKILLS

- **Scripting Languages and Analytics:** Python, R, Matlab, SQL