ADA SHAW

629 Armenise
210 Longwood Ave,
Boston, MA 02115

ashaw3895.github.io
ayshaw@g.harvard.edu
(916)-412-5899

EDUCATION

Ph.D., Applied Math, Harvard University **Expected Graduation: 9/2023**
Thesis: Optimizing Protein Fitness and Function with Sparse Experimental Data
Thesis Advisor: Debora Marks

M.S., Engineering Sciences, Harvard University **May 2019**

B.S., Chemical Engineering, **May 2017**
University of California, Berkeley

SKILLS

Programming Languages: Python, C, C++, MATLAB, R, Bash, HTML, CSS, FORTRAN, Javascript, Javascript React
Proficient knowledge of pattern recognition algorithms, Neural Networks, Machine Learning, Unix/Linux based systems, Handling terabyte size data, Parallel Computing, AWS Cloud based computing, Web Design, COMSOL Multiphysics

RESEARCH EXPERIENCE

Graduate Research in Computational Biology **Advisor: Debora Marks**
Harvard Medical School, Boston, MA **April 2019 - Present**

- ***Removing Bias in Sequence Models of Protein Fitness***
  - Developed statistical methods to improve unsupervised sequence model predictions and incorporate labels with evolutionary models. Designed methods and conducted all computational tests and experiments. Method significantly improves unsupervised model predictions and introduces methods to design multi-functional proteins.
  - **First author paper will be submitted in July 2023.**
- ***Protein Design using Structure-based Residue Preferences***
  - Developed an unsupervised design approach that learns residue mutation preferences from local structural dependencies. Major contributor.
  - **Co-author paper submitted to Nature Structural & Molecular Biology June 2023.** https://doi.org/10.1101/2022.10.31.514613
- ***Deep Learning Prediction of Enzyme Optimum pH***
  - Developed large language modeling methods to infer biological drivers of pH tolerance in enzymes. Major contributor.
  - **Co-author paper in review at Nature Machine Intelligence.**
  - https://doi.org/10.1101/2023.06.22.544776
- Developed machine learning methods to learn and predict from sparse, disparate enzyme activity. Work in collaboration with National Research Energy Lab (NREL) to develop plastic-eating enzymes. **First co-author paper will be submitted in July 2023.**
- ***An in silico method to assess antibody fragment polyreactivity***
  - Used AWS servers to host an online machine learning model to predict antigen poly-specificity and visualize sequence biometrics. Users can visit http://18.224.60.30:3000/ to input antigen sequences and get predictions. **Work published in Nature Communications:**
  - Harvey, E.P., Shin, JE., Skiba, M.A. *et al.* An in silico method to assess antibody fragment polyreactivity. *Nat Commun* **13**, 7554 (2022). https://doi.org/10.1038/s41467-022-35276-4

Graduate internship: *Global modeling and Assimilation Office* **August 2017 - November 2018**
NASA, Goddard, MD

- Collaborated between Jacob lab and NASA GEOS-5 model development to evaluate chemistry and dynamics of new 132 vertical level atmospheric general circulation model
- Used python to analyze terabytes of NetCDF output from GEOS-5 simulations
- Identified GEOS-5 model bugs, physical inconsistencies and collaborated with NASA scientists to fix and ensure valid model performance and output

Graduate Research **Advisor: Daniel Jacob**

Harvard University, Cambridge, MA **August 2017 - November 2018**

- Worked on improving the vertical resolution of GEOS-Chem model to predict transpacific pollution influenced high ozone pollution days in western US
- Used Harvard Supercomputing cluster to run GEOS-Chem Chemical Transport Model
- Used R to create an algorithim to detect ozone laminae in ozone profiles
- Analyzed time series outputs to find statistically significant trends in Harvard Forest CO levels
- *Detecting Ozone Layers from Ozonesondes, Harvard University, Fall 2017*
  Developed an algorithm to detect ozone laminae off the coast of Northern California, using data from Trinidad Head, CA ozonesondes. The algorithm was able to filter out high frequency noise, define the free troposphere, recognize high ozone peaks that fit the criteria of free tropospheric ozone laminae.
- *Analysis of Advection Schemes for Application in a Turbulent Propeller Wake, Fall 2018*
  Coded and tested three advection schemes: Essentially Non-Oscillating (ENO), Superbee, and Monotonic upwind Scheme for Conservation Laws (MUSCL) using 1-D and 2-D standard testing methods. We applied the lowest error schemes to a steady state velocity field produced by a weather balloon propeller in the stratosphere.

Data Science and Parallel Computing Class Projects

- *Dog Recognition Neural Network, Data Science Final Project, Harvard University, Fall 2018*
  Optimized and compared ResNet's Neural Networks, Convolutional Neural Networks and Artificial Neural Networks capabilities to predict the breeds of 20,000+ purebred dogs. We used Keras machine learning to implement and test our networks.
- *Verification of Goldbach's Conjecture, Harvard University, Spring 2018*
  Designed an algorithm in C for verifying Goldbach's conjecture and developed several parallel implementations of the code to identify the best strategies for tackling the problem as integer size increases. We tested the following forms of parallelism: OpenMP shared memory parallelism, MPI distributed memory parallelism, Hybrid MPI-OpenMP parallelism, OpenACC GPU accelerated computing.
- *Keystone Pipeline Mapping Project, UC Berkeley, Fall 2016*
  Worked with a team of students to data mine and develop a web interface that visually relates Keystone pipeline leakages, national watersheds, First Nation Reservation territories. Presented and submitted during Berc CleanWeb Hackathon

Undergraduate research **Advisor: Lisa Alvarez-Cohen**

UC Berkeley Department of Civil and Environmental Engineering **May 2016- May 2017**

- Worked on Anammox carbon-fixation pathway project and used ion-exchange chromatography and high pressure liquid chromatography to monitor levels of nitrite, nitrate, ammonia and carbon sources
- Cultivated anaerobic Anammox bacteria on anaerobic media
- Worked with qPCR, 16s ribosomal RNA, and electrolysis to quantify cell culture species in Anammox consortium
- *Using a Graphite Cathode as an Electron Donor in Anammox Electrolysis Cell to Investigate Extracellular Electron Transfer, UC Berkeley Honors Research in Alvarez-Cohen Lab*
  Developed a research project designing and building a bioelectrical stimulation reactor to test for extracellular electron transport in Anammox bacteria. I measured nitrate, nitrite and ammonia levels to track Anammox activity at various currents.

FOREIGN LANGUAGES:

- Mandarin Chinese: fluent speaking, reading and writing
- Spanish: limited speaking, reading and writing
- French: limited speaking, reading and writing

HONORS AND ACTIVITIES

UC Davis College of Biological Sciences Dean's List	Fall 2013, Fall 2014
UC Davis College of Engineering, Dean's List	Spring 2015
The Aggie, UC Davis newspaper, features writer	Spring 2015
Honors Research at UC Berkeley	Fall 2016
Biofuels Technology Club	Spring 2017
Telegraph Green Initiative Fund, Outreach Intern	June 2015 - June 2016
ESE 6 Graduate Teaching Fellow	Spring 2019

ESSP90s Graduate Teaching Fellow	Spring 2019
MIT EarthHacks 3rd Place Team	Spring 2019
D.E. Shaw Graduate and Postdoctoral Women's Fellowship	Spring 2023