DAVID H. BROOKES

(240) 499-4647 ⋄ david.brookes@berkeley.edu ⋄ Berkeley, CA

`dhbrookes.github.io`

EDUCATION

Ph.D., Biophysics Graduate Group   August 2016 - May 2021
Berkeley AI Research Lab
University of California, Berkeley

B.S., Bioengineering   August 2011 - May 2015
Concentration in Computational Biology
University of California, Berkeley

Relevant coursework: Graduate Statistical Mechanics I & II, Graduate Quantum Mechanics I (audit), Introduction to Machine Learning, Efficient Algorithms and Intractable Problems, Numerical Analysis, Introduction to Artificial Intelligence.

PROFESSIONAL EXPERIENCE

AI Resident, *Google X*   May 2019 - August 2019
· Combined machine learning expertise with software engineering skills to bolster an early-stage confidential project, as well as shape the project's future direction.
· Effectively communicated advanced modeling ideas to a team of people with vastly different expertise.

RESEARCH EXPERIENCE

Graduate Student Researcher, *Berkeley AI Research Lab*   February 2018 - Present
Supervisor: Prof. Jennifer Listgarten
· Developed theory to determine number of experimental measurements required to accurately estimate biological fitness functions with sparse recovery algorithms ([1]).
· Developed probabilistic modeling approaches for biological sequence design. Specifically, a rigorous probabilistic framework for sampling sequences that satisfy desired properties using joint supervised and generative modeling approaches ([4] and [5], below).
· In ongoing work, we are applying our design methods in conjunction with novel modeling strategies to design a library of viral capsid sequences to be used in a downstream Directed Evolution experiment run by collaborators.

Student Researcher, *UC Berkeley College of Chemistry*   October 2014 - February 2018
Supervisor: Prof. Teresa Head-Gordon
· Developed a Bayesian inference procedure for determining likely ensembles of intrinsically disordered protein structures given a set of experimental data ([8]).
· Lead developer of software suite to solve for electrostatic interactions on the surface of biological macromolecules. ([6] and [7], below, with code available at `github.com/dhbrookes/pbsolvers`)
· Solved a linear model of physically relevant basis functions to find the distribution of possible structures of liquid water that satisfy a set of experimental data ([9]).

RESEARCH

Papers

[1] **D. H. Brookes**, A. Aghazadeh, and J. Listgarten. On the sparsity of fitness functions and implications for learning. *bioRxiv*, 2021.

[2] A. Aghazadeh, H. Nisonoff, O. Ocal, **D. H. Brookes**, Y. Huang, O. O. Koyluoglu, J. Listgarten, and K. Ramchandran. Sparse Epistatic Regularization of Deep Neural Networks for Inferring Fitness Functions. *bioRxiv*, 2021.

[3] **D. H. Brookes**, A. Busia, C. Fannjiang, K. Murphy, and J. Listgarten. A view of Estimation of Distribution Algorithms through the lens of Expectation Maximization. *Proceedings of The Genetic and Evolutionary Computation Conference*, 2020.

[4] **D. H. Brookes**, H. Park, and J. Listgarten. Conditioning by adaptive sampling for robust design. *Proceedings of ICML*, 2019. **Selected for a 20 minute oral presentation ($< 5\%$ of submissions)**.

[5] **D. H. Brookes** and J. Listgarten. Design by adaptive sampling. *NeurIPS Workshop on Machine Learning for Molecules and Materials*, 2018.

[6] E. Jurrus, D. Engel, K. Star, K. Monson, J. Brandi, L. E. Felberg, **D. H. Brookes**, L. Wilson, J. Chen, K. Liles, M. Chun, P. Li, T. Dolinsky, R. Konecny, D. Koes, J. E. Nielsen, T. Head-Gordon, W. Geng, R. Krasny, M. Gunner, G.-W. Wei, M. J. Holst, J. A. McCammon, N. A. Baker. Improvements to the APBS biomolecular solvation software suite. *Protein Sci.* 27 (1), pp. 112-128, 2018.

[7] L. E. Felberg, **D. H. Brookes**, E. Jurrus, N. Baker, and T. Head-Gordon. PB-AM: An open-source, fully analytical linear Poisson-Boltzmann solver. *J. Comp. Chem.* 38 (15), pp. 1275-1282, 2017.

[8] **D. H. Brookes** and T. Head-Gordon. Experimental inferential structure determination of ensembles for intrinsically disordered proteins *J. Am. Chem. Soc.* 138 (13), pp. 4530-4538, 2016.

[9] **D. H. Brookes** and T. Head-Gordon. The family of oxygen-oxygen radial distribution functions for water. *J. Phys. Chem. Lett.* 6 (15), pp. 2938-2943, 2015

[10] L. E. Felberg, **D. H. Brookes**, J. E. Rice, T. Head-Gordon, and W. Swope. Role of hydrophilicity and length of diblock arms for determining star polymer physical properties. *J. Phys. Chem. B.* 119 (3), pp. 944-957, 2015.

Invited talks and posters

1. A view of Estimation of Distribution Algorithms through the lens of Expectation Maximization *The Genetic and Evolutionary Computation Conference*, Jul. 2020.

2. Conditioning by adaptive sampling for robust design. *ICML*, Jun. 2019.

3. Design by Adaptive Sampling. *Google AI, Mountain View*, Jan. 2019.

4. Design by Adaptive Sampling. *NeurIPS Workshop on Machine Learning for Molecules and Materials*, Dec. 2018.

5. Sequence Design by Adaptive Sampling. *Berkeley Center for Computational Biology Retreat*, Nov. 2018.