Alex Infanger

email: alexdinfanger[at]gmail[dot]com, website: https://alexinfanger.github.io

Education

Stanford University. (09/2016-06/2022)
PhD in Computational and Mathematical Engineering.
Dissertation, *Truncation Algorithms for Markov Chains and Processes.*
Honorable Mention for Gene Golub Dissertation Award.
ICME Teaching Assistant Award (2020-2021).

Research Topics: My thesis focuses on the problem of approximating an infinite or very large state space Markov chain $X = (X_n : n \geq 0)$ on a smaller subset of the state space A. A well-known approach to this problem is to re-route transitions of the original chain that attempt to leave A into A^c back into A. We give new conditions under which such an approximation is good for estimating the stationary distribution π of X. We also provide a new approximation for estimating π on A that comes with error bounds. More generally, I've worked a lot on the structured numerical linear algebra and optimization problems that arise when computing/bounding expected values in the context of Markov chain modeling.

University of California, Santa Cruz. (08/2012-09/2016)
BS in Physics, **highest honors**.
Minor in Mathematics.
Summa cum laude, **Phi Beta Kappa**.
Senior Thesis, *The Existence of Terrestrial Gamma-Ray Flashes that Paralyze RHESSI*, awarded the Dean's and Chancellor's Awards.

Selected Research & Work

Long-Term Future Fund (05/2023-present)
Grant Recipient

- Initiated and co-led a project on sampling the reverse dynamics of large language models for AI safety (automated redteaming) applications. Led to a spotlight paper at the Neurips 2023 SoLaR workshop.

AGI Safety Fundamentals Reading Group Facilitator (10/2022-12/2022)

- Facilitated multiple reading groups that went through a variation of the AGI Safety Curriculum (with modifications by Sam Marks) for the MIT AI Alignment Club.

Adobe Systems Incorporated (07/2018-09/2018)
Data Science Intern

- Estimated Markov model for Creative Cloud customers in PySpark and Pandas.

Selected Publications and Preprints

"Eliciting language model behaviors using reverse language models." (2023). J. Pfau, A. Infanger, A. Sheshadri, A. Panda, C. Huebner, J. Michael. Neurips 2023 SoLaR Workshop **(Spotlight)**.

"A new truncation algorithm for Markov chain equilibrium distributions with computable error bounds." (2022). A. Infanger, P. Glynn. arXiv:2208.14406

"On convergence of general truncation-augmentation schemes for approximating stationary distributions of Markov chains." (2022). A. Infanger, P. Glynn, Y. Liu. arXiv:2203.15167.

"Solutions of Poisson's equation for stochastically monotone Markov chains". P. Glynn, A. Infanger. *Queueing Models and Service Management.* (2022)

"On the spectral radius and stiffness of Markov jump process rate matrices" P. Glynn, A. Infanger. *Stochastic Models.* (2020)

Programming

Python, PyTorch, Tensorflow, Keras, Julia, PySpark, Pandas, SQL, Matlab.