Sam Foreman's Résumé

Sam Foreman

2025-04-26

Table of contents

About

Computational Scientist at Argonne National Laboratory.
Scaling AI for science on supercomputers.
samforeman.me GitHub • Google Scholar • ORCID • Twitter

Education

- **Ph.D., Physics**
  University of Iowa | 2015–2019
  - *Learning Better Physics: A Machine Learning Approach to Lattice Gauge Theory*
- **B.S. in Engineering Physics**
  University of Illinois at Urbana-Champaign | 2010–2015
  - Energy Storage in Quantum Resonators (US Patent #US9741492B2)
- **B.S. in Applied Mathematics**
  University of Illinois at Urbana-Champaign | 2010–2015

Professional Experience

- **Assistant Computational Scientist**
  - *Argonne National Laboratory*, Leadership Computing Facility (ALCF) Lemont, IL | 2022–Present
    * Research lead on scaling large language models (LLMs) and generative AI for science on supercomputers (Aurora, Frontier, LUMI, Leonardo, …).
      · Co-lead the Models and Pretraining team of the AuroraGPT project
    * Optimize large-scale training of foundation models and language models for scientific applications.
    * Collaborate with interdisciplinary teams to enhance simulation efficiency and scalability
    * Focus on AI and HPC for scientific applications, including:
      · Training large language models on supercomputers
      · Genome scale language models (GenSLMs) for studying SARS-CoV-2 evolutionary dynamics
      · Direct Preference Optimization (DPO) for multimodal protein design workflows
      · Climate modeling and weather forecasting using foundation models
      · Developing improved sampling algorithms for lattice quantum chromodynamics (QCD)
    * https://www.alcf.anl.gov/about/people/sam-foreman
- **Postdoctoral Researcher**
  - *Argonne National Laboratory*, Leadership Computing Facility (ALCF) Lemont, IL | 2019 – 2022
    * Applied deep learning to lattice gauge theory and quantum field simulations.
    * Developed ML-enhanced Monte Carlo methods for QCD (l2hmc-qcd).
    * Engaged in AI-for-Science collaborations with national labs and university partners.
- **Graduate Researcher (DOE SCGSR Fellowship)**
  - *Argonne National Laboratory*, Mathematics and Computer Sciences Division (MCS) Lemont, IL | 2018 – 2019
    * Development of l2hmc-qcd in collaboration with ALCF for my PhD Thesis research

Publications

> Note
>
> You can find a full list of my publications on my Google Scholar

1. **AERIS: Argonne Earth Systems Model for Reliable and Skillful Predictions** (Hatanpää et al. (2025))
    - *2025 ACM Gordon Bell Prize for Climate Modeling Finalist*
2. Aurora: Architecting Argonne's First Exascale Supercomputer for Accelerated Scientific Discovery (Allen et al. (2025))
3. HiPerRAG: High-Performance Retrieval Augmented Generation for Scientific Insights (Gokdemir et al. (2025))
4. Automated Tuning for HMC Mass Ratios (Torsiello et al. (2025))
5. MOFA: Discovering Materials for Carbon Capture with a GenAI and Simulation-Based Workflow (Yan et al. (2025))
6. **MProt-DPO: Breaking the ExaFLOPS Barrier for Multimodal Protein Design with DPO** (Dharuman et al. (2024))
    - *2024 ACM Gordon Bell Finalist*
7. Intro to HPC Bootcamp: Engaging New Communities Through Energy Justice Projects (Leung et al. (2024))
8. Thorough Characterization and Analysis of Large Transformer Model Training At-Scale (Cheng et al. (2024))
9. MLMC: Machine Learning Monte Carlo for Lattice Gauge Theory (Foreman et al. (2023))
10. Protein Generation via Genome-scale Language Models with Bio-physical Scoring (Dharuman et al. (2023))
11. DeepSpeed4Science Initiative: Enabling Large-Scale Scientific Discovery (Song et al. (2023))
    - DeepSpeed4Science.ai Blog Post
    - Looooooooong Sequence Lengths
12. Comprehensive Performance Study of LLMs on Novel AI Accelerators (Emani et al. (2023))
13. Exploratory Analysis of Climate Data with `ClimRR`, Intro to HPC Bootcamp @ NERSC (Foreman (2023))
14. **GenSLMs: Genome-scale language models reveal SARS-Cov-2 evolutionary dynamics** (Zvyagin et al. (2023))
    - Winner of the *ACM Gordon Bell Special Prize for High Performance Computing-Based COVID-19 Research*
15. Lattice QCD and Particle Physics (Kronfeld et al. (2022))

16. Applications of ML to Lattice QFT (Boyda et al. (2022))
17. LeapFrogLayers: Trainable Framework for Effective Sampling (Foreman, Izubuchi, et al. (2021))
18. HMC with Normalizing Flows [slides] (Foreman, Izubuchi, et al. (2021))
19. Deep Learning Hamiltonian Monte Carlo [+ poster] (Foreman, Jin, et al. (2021))
20. Machine Learning and Neural Networks for Field Theory (Foreman et al. (2020))
21. Examples of renormalization group transformations for image sets (Samuel Foreman et al. (2018))
22. RG inspired Machine Learning for lattice field theory (Sam Foreman et al. (2018))
23. Large Energy Density in Three-Plate Nanocapacitors due to Coulomb Blockade (Hubler et al. (2018))
24. Superconductivity of In and Sn Samples (Deamont and Foreman (2014))

References

Allen, Benjamin S., James Anchell, Victor Anisimov, et al. 2025. *Aurora: Architecting Argonne's First Exascale Supercomputer for Accelerated Scientific Discovery*. https://arxiv.org/abs/2509.08207.

Boyda, Denis, Salvatore Calı̀, Sam Foreman, et al. 2022. "Applications of Machine Learning to Lattice Quantum Field Theory." *arXiv Preprint arXiv:2202.05838*. https://arxiv.org/abs/2202.05838.

Cheng, Scott, Jun-Liang Lin, Murali Emani, et al. 2024. "Thorough Characterization and Analysis of Large Transformer Model Training at-Scale." *Proc. ACM Meas. Anal. Comput. Syst.* (New York, NY, USA) 8 (1). https://doi.org/10.1145/3639034.

Deamont, George, and Sam Foreman. 2014. *Superconductivity of in and Sn Samples*.

Dharuman, Gautham, Kyle Hippe, Alexander Brace, et al. 2024. "MProt-DPO: Breaking the ExaFLOPS Barrier for Multimodal Protein Design Workflows with Direct Preference Optimization." *Proceedings of the International Conference for High Performance Computing, Networking, Storage, and Analysis* (Atlanta, GA, USA), SC '24. https://doi.org/10.1109/SC41406.2024.00013.

Dharuman, Gautham, Logan Ward, Heng Ma, et al. 2023. "Protein Generation via Genome-Scale Language Models with Bio-Physical Scoring." *Proceedings of the SC'23 Workshops of the International Conference on High Performance Computing, Network, Storage, and Analysis*, 95–101.

Emani, Murali, Sam Foreman, Varuni Sastry, et al. 2023. "A Comprehensive Performance Study of Large Language Models on Novel AI Accelerators." *arXiv Preprint*

arXiv:2310.04607. https://arxiv.org/abs/2310.04607.

Foreman, Sam. 2023. "Energy Justice Analysis of Climate Data with ClimRR." August 7. https://saforem2.github.io/climate-analysis.

Foreman, Sam, Joel Giedt, Yannick Meurice, and Judah Unmuth-Yockey. 2018. "RG-inspired machine learning for lattice field theory." *European Physical Journal Web of Conferences*, European physical journal web of conferences, vol. 175 (March): 11025. https://doi.org/10.1051/epjconf/201817511025.

Foreman, Sam, Taku Izubuchi, Luchang Jin, Xiao-Yong Jin, James C Osborn, and Akio Tomiya. 2021. "HMC with Normalizing Flows." *arXiv Preprint arXiv:2112.01586*. https://arxiv.org/abs/2112.01586.

Foreman, Sam, Xiao-Yong Jin, and Osborn James C. 2021. *Deep Learning Hamiltonian Monte Carlo*. https://arxiv.org/abs/2105.03418.

Foreman, Sam, Xiao-Yong Jin, and James C Osborn. 2020. *Machine Learning and Neural Networks for Field Theory*.

Foreman, Sam, Xiao-Yong Jin, and James C. Osborn. 2023. *MLMC: Machine Learning Monte Carlo for Lattice Gauge Theory*. https://arxiv.org/abs/2312.08936.

Foreman, Samuel, Joel Giedt, Yannick Meurice, and Judah Unmuth-Yockey. 2018. "Examples of Renormalization Group Transformations for Image Sets." *Physical Review E* 98 (5): 052129.

Gokdemir, Ozan, Carlo Siebenschuh, Alexander Brace, et al. 2025. *HiPerRAG: High-Performance Retrieval Augmented Generation for Scientific Insights*. https://arxiv.org/abs/2505.04846.

Hatanpää, Väinö, Eugene Ku, Jason Stock, et al. 2025. *AERIS: Argonne Earth Systems Model for Reliable and Skillful Predictions*. https://arxiv.org/abs/2509.13523.

Hubler, A, S Foreman, J Liu, and L Wortsmann. 2018. "Large Energy Density in Three-Plate Nanocapacitors Due to Coulomb Blockade." *Journal of Applied Physics* 123 (10).

Kronfeld, Andreas S, Tanmoy Bhattacharya, Thomas Blum, et al. 2022. "Lattice QCD and Particle Physics." *arXiv Preprint arXiv:2207.07641*. https://arxiv.org/abs/2207.07641.

Leung, Mary Ann, Katharine Cahill, Rebecca Hartman-Baker, et al. 2024. "Intro to HPC Bootcamp: Engaging New Communities Through Energy Justice Projects." *Journal of Computational Science Education* 15 (1). https://doi.org/10.22369/issn.2153-4136/15/1/10.

Song, Shuaiwen Leon, Bonnie Kruft, Minjia Zhang, et al. 2023. "DeepSpeed4Science Initiative: Enabling Large-Scale Scientific Discovery Through Sophisticated AI System Technologies." *arXiv Preprint arXiv:2310.04610*. https://arxiv.org/abs/2310.04610.

Torsiello, J., G. T. Fleming, S. Foreman, X.-Y. Jin, and J. C. Osborn. 2025. "Automated Tuning for HMC Mass Ratios." In *PoS*. Argonne, ALCF; Argonne National Laboratory (ANL), Argonne, IL (United States); Temple U.; Fermi National Accelerator Laboratory (FNAL), Batavia, IL (United States). https://doi.org/10.22323/1.466.0052.

Yan, Xiaoli, Nathaniel Hudson, Hyun Park, et al. 2025. *MOFA: Discovering Materials for Carbon Capture with a GenAI- and Simulation-Based Workflow*. https://arxiv.org/abs/2501.10651.

Zvyagin, Maxim, Alexander Brace, Kyle Hippe, et al. 2023. "GenSLMs: Genome-Scale Language Models Reveal SARS-CoV-2 Evolutionary Dynamics." *The International Journal of High Performance Computing Applications* 37 (6): 683–705.

Awards and Honors

- Member of the DeepSpeed Technical Steering Commiittee, 2025 – Present
  - Contributing to the development and direction of the DeepSpeed library for large-scale model training.
- Nominated to serve on the US **Coordinating Panel for Software and Computing** by the Division of Particles and Fields of the American Physical Society (APS).
- **Finalist, ACM Gordon Bell Prize in Climate Modeling**, 2025
  - Recognized for our work on
    AERIS (Hatanpää et al. (2025)): The first billion-parameter pixel-level diffusion model for global weather and subseasonal-to-seasonal forecasting. Trained efficiently at scales from 1.3–80B parameters with our sequence-window parallelism (SWiPe) strategy, we achieve a sustained mixed-precision performance of 10.21 ExaFLOPS and peak performance of 11.21 ExaFLOPS, scaling to 10,080 nodes (120,960 GPUs) on the Aurora supercomputer.
- **Finalist, ACM Gordon Bell Prize**, 2024

  - Acknowledged for the MProt-DPO (Dharuman et al. (2024)) project, which achieved over 4 ExaFLOP sustained performance in multimodal protein design workflows using Direct Preference Optimization.
    * Argonne team breaks new ground in AI-driven protein design – Argonne @ SC

- **ACM Gordon Bell Special Prize for High Performance Computing-Based COVID-19 Research**, 2022
  - Recognized for contributions to the GenSLMs (Zvyagin et al. (2023)) project, which developed genome-scale language models to study SARS-CoV-2 evolutionary dynamics.
    * ACM Gordon Bell Special Prize for HPC-Based COVID-19 Research Awarded to Team for Modelling How Pandemic-Causing Viruses, Especially SARS-CoV-2, are Identified and Classified

- **DOE Office of Science Graduate Student Research Fellow**, 2018
  - Awarded by the Department of Energy for outstanding research contributions during graduate studies.

Talks

> Note
>
> You can see all of my talks online at https://samforeman.me/talks/

- 2025-:
  - 12: Training Foundation Models on Supercomputers @ Argonne National Laboratory
  - 10: Training Foundation Models on Supercomputers @ University of Illinois at Urbana-Champaign
  - 10: Training Foundation Models on Supercomputers @ Georgia Institute of Technology
  - 10: AERIS: Argonne's Earth Systems Model @ 2025 ALCF Hands-On HPC Workshop
  - 09: Scientific AI at Scale: AI for Science @ Open SkAI 2025
  - 09: Scientific AI at Scale: Distributed Training @ Open SkAI 2025
  - 07: Large Scale Training on Diverse Accelerators @ Scalable Deep Learning, SIAM AN2025
  - 05: LLMs on Aurora: AuroraGPT @ 2025 ALCF INCITE GPU Hackathon
  - 05: LLMs on Aurora: ezpz @ 2025 ALCF INCITE GPU Hackathon
  - 02: AuroraGPT: Foundation Models for Science @ Foundation Models for the Electric Grid

- 2024-:
  - 11: Parallel Training Methods @ AI-for-Science on Supercomputers
  - 10: AuroraGPT @ 2024 ALCF Hands-On HPC Workshop
  - 10: Machine Learning and Foundation Models at Scale @ 2024 ALCF Hands-On HPC Workshop
  - 09: AuroraGPT @ HPC User Forum, 2024
  - 08: Training LLMs at Scale @ ATPESC, 2024
  - 07: LLMs on Polaris @ Center for Scientific Foundation Models, Summer School 24'
  - 03: Parallel Training Techniques @ AI-4-Science Training Series
  - 02: LLMs from Scratch @ LLM Tutorial Workshop
- 2023-:
  - 11: Creating Small(-ish) LLMs @ LLM Tutorial Workshop (1)
  - 10: Exascale Science on Aurora @ Intel oneAPI Workshop @ UIC
  - 10: LLM Lunch Talk @ ALCF Hands On HPC Workshop
  - 08: Scaling LLMs for Science @ Data-Intensive Computing + AI/ML at Scale
  - 07: MLMC: Machine Learning Monte Carlo @ Lattice 2023
  - 07: Generative Modeling and Efficient Sampling @ PASC23
  - 04: Efficient Sampling for LGT @ Deep Fridays @ U. Bologna
- 2022-:
  - 11: Large Scale Training @ AI4Science on Supercomputers (ALCF)
  - 10: Hyperparameter Management @ ALCF SDL Workshop
  - 08: Statistical Learning @ ATPESC 2022
  - 05: Scientific Data Science: An Emerging Symbiosis @ ANL (05/2022)
  - 03: Machine Learning in HEP @ UNC Greensboro
- 2021-:
  - 12: Accelerated Sampling Methods for LGT, @ DWQ @ 25 [BNL]
  - 09: Training Topological Samplers for LGT @ ML4HEP, ECT* Trento
  - 05: Deep Learning HMC for Improved Gauge Generation @ ML in LQCD Workshop [2021]
- 2020:
  - 02: Machine Learning for Lattice QCD @ U. Iowa [2020]

Events

- Organizer for:
  - SC25 Workshop: High Performance Python for Science at Scale (HPPSS), November 2025

- SC25 Tutorial: Accelerating and Scaling Python for HPC
- SC24 Workshop: High Performance Python for Science at Scale (HPPSS), November 2024
- SC23 Workshop: High Performance Python for Science at Scale (HPPSS), November 2023
- Machine Learning and Quantum Computing for Earth Sciences at 17th U. S. National Congress on Computational Mechanics, July 2023

References

References

Allen, Benjamin S., James Anchell, Victor Anisimov, et al. 2025. *Aurora: Architecting Argonne's First Exascale Supercomputer for Accelerated Scientific Discovery*. https://arxiv.org/abs/2509.08207.

Boyda, Denis, Salvatore Calì, Sam Foreman, et al. 2022. "Applications of Machine Learning to Lattice Quantum Field Theory." *arXiv Preprint arXiv:2202.05838*. https://arxiv.org/abs/2202.05838.

Cheng, Scott, Jun-Liang Lin, Murali Emani, et al. 2024. "Thorough Characterization and Analysis of Large Transformer Model Training at-Scale." *Proc. ACM Meas. Anal. Comput. Syst.* (New York, NY, USA) 8 (1). https://doi.org/10.1145/3639034.

Deamont, George, and Sam Foreman. 2014. *Superconductivity of in and Sn Samples*.

Dharuman, Gautham, Kyle Hippe, Alexander Brace, et al. 2024. "MProt-DPO: Breaking the ExaFLOPS Barrier for Multimodal Protein Design Workflows with Direct Preference Optimization." *Proceedings of the International Conference for High Performance Computing, Networking, Storage, and Analysis* (Atlanta, GA, USA), SC '24. https://doi.org/10.1109/SC41406.2024.00013.

Dharuman, Gautham, Logan Ward, Heng Ma, et al. 2023. "Protein Generation via Genome-Scale Language Models with Bio-Physical Scoring." *Proceedings of the SC'23 Workshops of the International Conference on High Performance Computing, Network, Storage, and Analysis*, 95–101.

Emani, Murali, Sam Foreman, Varuni Sastry, et al. 2023. "A Comprehensive Performance Study of Large Language Models on Novel AI Accelerators." *arXiv Preprint arXiv:2310.04607*. https://arxiv.org/abs/2310.04607.

Foreman, Sam. 2023. "Energy Justice Analysis of Climate Data with ClimRR." August 7. https://saforem2.github.io/climate-analysis.

Foreman, Sam, Joel Giedt, Yannick Meurice, and Judah Unmuth-Yockey. 2018. "RG-inspired machine learning for lattice field theory." *European Physical Journal Web of Conferences*, European physical journal web of conferences, vol. 175 (March): 11025. https://doi.org/10.1051/epjconf/201817511025.

Foreman, Sam, Taku Izubuchi, Luchang Jin, Xiao-Yong Jin, James C Osborn, and Akio Tomiya. 2021. "HMC with Normalizing Flows." *arXiv Preprint arXiv:2112.01586*. https://arxiv.org/abs/2112.01586.

Foreman, Sam, Xiao-Yong Jin, and Osborn James C. 2021. *Deep Learning Hamiltonian Monte Carlo*. https://arxiv.org/abs/2105.03418.

Foreman, Sam, Xiao-Yong Jin, and James C Osborn. 2020. *Machine Learning and Neural Networks for Field Theory*.

Foreman, Sam, Xiao-Yong Jin, and James C. Osborn. 2023. *MLMC: Machine Learning Monte Carlo for Lattice Gauge Theory*. https://arxiv.org/abs/2312.08936.

Foreman, Samuel, Joel Giedt, Yannick Meurice, and Judah Unmuth-Yockey. 2018. "Examples of Renormalization Group Transformations for Image Sets." *Physical Review E* 98 (5): 052129.

Gokdemir, Ozan, Carlo Siebenschuh, Alexander Brace, et al. 2025. *HiPerRAG: High-Performance Retrieval Augmented Generation for Scientific Insights*. https://arxiv.org/abs/2505.04846.

Hatanpää, Väinö, Eugene Ku, Jason Stock, et al. 2025. *AERIS: Argonne Earth Systems Model for Reliable and Skillful Predictions*. https://arxiv.org/abs/2509.13523.

Hubler, A, S Foreman, J Liu, and L Wortsmann. 2018. "Large Energy Density in Three-Plate Nanocapacitors Due to Coulomb Blockade." *Journal of Applied Physics* 123 (10).

Kronfeld, Andreas S, Tanmoy Bhattacharya, Thomas Blum, et al. 2022. "Lattice QCD and Particle Physics." *arXiv Preprint arXiv:2207.07641*. https://arxiv.org/abs/2207.07641.

Leung, Mary Ann, Katharine Cahill, Rebecca Hartman-Baker, et al. 2024. "Intro to HPC

Bootcamp: Engaging New Communities Through Energy Justice Projects." *Journal of Computational Science Education* 15 (1). https://doi.org/10.22369/issn.2153-4136/15/1/10.

Song, Shuaiwen Leon, Bonnie Kruft, Minjia Zhang, et al. 2023. "DeepSpeed4Science Initiative: Enabling Large-Scale Scientific Discovery Through Sophisticated AI System Technologies." *arXiv Preprint arXiv:2310.04610*. https://arxiv.org/abs/2310.04610.

Torsiello, J., G. T. Fleming, S. Foreman, X.-Y. Jin, and J. C. Osborn. 2025. "Automated Tuning for HMC Mass Ratios." In *PoS*. Argonne, ALCF; Argonne National Laboratory (ANL), Argonne, IL (United States); Temple U.; Fermi National Accelerator Laboratory (FNAL), Batavia, IL (United States). https://doi.org/10.22323/1.466.0052.

Yan, Xiaoli, Nathaniel Hudson, Hyun Park, et al. 2025. *MOFA: Discovering Materials for Carbon Capture with a GenAI- and Simulation-Based Workflow*. https://arxiv.org/abs/2501.10651.

Zvyagin, Maxim, Alexander Brace, Kyle Hippe, et al. 2023. "GenSLMs: Genome-Scale Language Models Reveal SARS-CoV-2 Evolutionary Dynamics." *The International Journal of High Performance Computing Applications* 37 (6): 683–705.