BENJAMIN THORNE

me@bthorne.io b-thorne bthorne.io San Francisco, CA

ABOUT

Computational scientist with expertise in scientific machine learning, differentiable programming, probabilistic programming, and high performance computing in Julia and Python. A list of my publications can be found here.

SKILLS

Python: `pytorch`, `tensorflow`, `numpy`, `scipy`, `jax`, `dask`, `mpi4py`, `Cython`, `Numba`
Julia: `CUDA.jl`, `Flux.jl`, `DiffEqFlux.jl`, `ModelingToolkit.jl`, `DifferentialEquations.jl`
Other: Docker, Podman, Kubernetes, Rancher, Quarto, CSS, Javascript, Git, DVC, AWS [S3, EC2]

EXPERIENCE

Senior Machine Learning Engineer — 2023 - present
Atomic Industries — Remote

Physics-based optimization of injection molds

- Developed algorithmic basis for the automated design of cooling systems for injection molding tools. Reduced human expert design time from $\sim$ days to $\sim$ minutes.
- Optimized highly constrained geometric design space, targeting physical properties from Monte Carlo PDE solvers (impelmented in `Julia` and `Jax`).

Machine Learning Engineer — 2023
National Energy Research Scientific Computing Center, Berkeley Lab — Berkeley, CA

Uncertainty-aware machine learning for high energy physics

- Ran project meetings of 10+ scientists and engineers to align goals and deliverables.
- Collaborated on design of machine learning (ML) challenge platform hosted at *NERSC*.
- Designed metrics for scoring uncertainty-aware ML algorithms.
- Designed cosmology-based public ML challenge:
  - Designed dataset derived from 200 TB of simulation data stored on archival tape.
  - Distributed computations across 1000's of processes and ~10's of nodes on *Perlmutter*.
  - Implemented baseline uncertainty-aware convolutional model in PyTorch.

Postdoctoral Researcher — 2019 - 2022
University of California, Davis — Davis, CA

Generative modeling of dust maps using VAEs in Tensorflow

- Created training dataset of 1,000 dust maps from public satellite data.
- Designed and trained convolutional variational autoencoder in Tensorflow.
- Applied trained model to Bayesian inverse problems: data imputation, denoising, inference.

Differentiable likelihoods for cosmology in Julia with CUDA and auto-diff

- Developed differentiable forward model of 10^6-pixel CMB datasets in Julia.
- Used `Zygote.jl` for automatic differentiation and `CUDA.jl` for GPU acceleration.
- Used sparse approximations & preconditioning to improve memory usage and convergence.
- Distributed MCMC analysis of ~100 GB dataset across 10's of A100 GPU nodes on *Perlmutter*.

Python Sky Model Simulation Package

- Developed the *pysm* Python package, widely used by astrophysicists.
- *First author paper* describing the package has ~200 citations.
- Numba and mpi4py used for efficient distributed design, and caching supported for local runs.
- Run on 100's of nodes at NERSC to support *Simons Observatory* and *Stage-4*.

Other duties

- Co-organized the weekly cosmology seminar from January 2020 to December 2022.
- Co-supervised PhD students on various projects in machine learning and physics.
- Regularly delivered seminars and conference talks.

EDUCATION

PhD in Astrophysics 2015 - 2019

University of Oxford, Princeton University & Kavli IPMU Oxford, Princeton, Tokyo

Thesis: "Challenges in probing inflation with primordial gravitational waves" Supervised by Professor Jo Dunkley

Pixel likelihoods for Simons Observatory (SO) forecasting

- Wrote maximum likelihood estimation (MLE) algorithm for pixelized data using Numba.
- Made Monte Carlo (MC) simulations of SO observations (noise, foregrounds, and systematics).
- Applied MLE to MC simulations to forecast constraints on primordial gravitational wave search.

Data analysis for the Atacama Cosmology Telescope (ACT)

- Used public satellite data to compute two-point statistics of polarized dust emission.
- Performed Bayesian fit with MCMC to constrain parametric model of the two-point amplitude.
- Used as priors in ACT's 2018 likelihood analysis, included in two collaboration papers.

Novel observables of SU(2)-Axion inflation

- Analytically derived novel observables in a parity-violating cosmological model.
- Calculated sensitivities of the *LISA* and *LiteBIRD* satellites using Python and Numba.
- *First-author paper* now with 100+ citations.

M.S. & B.S. in Physics 2011 - 2015

University of Oxford, New College Oxford, UK

Photometric decomposition of barred and double-barred galaxies

- Collected dataset of galaxy images from *SDSS*.
- Used χ^2 optimization to determine photometric structure of each galaxy.
- Collaborated with researchers at *St Andrew's University* on two papers using results.