Ibrahim Elsharkawy

ML Researcher & Physicist

Toronto, ON | ibra.shark@gmail.com | +1 832-247-6842

ibrahime.org | ibrahimelsharkawy | Google Scholar

Research interests. Interpretability, uncertainty quantification, AI for Science, and physics-inspired ML theory. My work spans two complementary directions, *Physics for AI* (physics theory and data to make headway in AI interpretability and predictability) and *AI for Physics* (generative and foundation models as tools for scientific discovery).

EDUCATION

University of Toronto & Vector Institute — Toronto, ON | 2025 - Present
Ph.D. in Physics — Y. Kahn, C. Madison, D. Curtin

University of Illinois Urbana-Champaign — Urbana, IL | 2023 - 2025
M.Sc. in Physics — Y. Kahn, B. Hooberman

Rice University — Houston, TX | 2019 - 2023
B.S. Computational Physics | B.A. Applied Math & Philosophy — K. Ecklund

RESEARCH

Interpretability of Generative Models for Physical Systems — Toronto, ON | Jan 2026 - Present
University of Toronto & Vector Institute
Doctoral Researcher

- *Physics for AI:* Building generative models that are **interpretable by construction** and constrained to satisfy physical symmetries and conservation laws. We find that **physical data is the unique testbed** by which interpretability methods can be developed and verified **[P2]**.
- *AI for Physics:* Using the same interpretability tools validated on simulated systems **to extract novel physics from generative models** trained on real data.

Foundation Models for Scientific Point Clouds — Berkeley, CA | Mar 2025 - Present
Lawrence Berkeley National Laboratory, NERSC
Doctoral Researcher

- Building **cross-domain** foundation models for scientific point clouds (small irregular graphs) spanning high-energy physics, cosmology, and molecular dynamics **[A2] [A4]**. We find that pre-training in a domain **with cheap simulation** (e.g. High Energy Physics Proton Collisions) **benefits domains with limited or expensive data** (e.g., molecular dynamics, proteins) due potentially to shared physics.
- Researching optimal pre-training objectives for scientific foundation models when accurate simulators can produce **labeled data at scale**. Is self-supervision superior for downstream tasks when not necessary?

Neural Uncertainty Quantification — Urbana, IL | 2023 - 2025
University of Illinois Urbana Champaign
Doctoral Researcher

- Derived and empirically validated an **effective-field-theory description of neural networks** that predicts initialization-driven ensemble variance from first principles **without training an ensemble [P1]**.
- Developed a contrastive-normalizing-flow methodology for **uncertainty-aware parameter estimation** robust to systematic sources of error. **Won 1st place** in the Higgs Uncertainty Challenge at NeurIPS 2024 and CERN IML **[A1] [A3]**.

INDUSTRY

Research Geophysicist Intern — Houston, TX | 2024 - 2025
ExxonMobil
AI for Seismic Inversion and Anomaly Detection

- Researched and developed **ML parameter-estimation methods for 3D seismic inversion**, targeting the inversion of traditionally degenerate earth properties.
- Built novel ML-based **anomaly detection tools for 4D seismic** data.

Research and Software Engineering Intern — Houston, TX | 2018 - 2023
Petroleum GeoServices
Machine Learning and Cloud Infrastructure for Seismic

- Researched ML methods for seismic first-break picking, contributed to cloud migration and microservices.

PUBLICATIONS

Full list of preprints at ibrahime.org.

Physics for AI

P1. **I. Elsharkawy**, et al. *Uncertainty Quantification from Ensemble Variance Scaling Laws in Deep Neural Networks.* **Machine Learning: Science and Technology**, 6(3):035040, 2025.

P2. Bogorad, **I. Elsharkawy**, Kahn, et al. *Generative Models on Phase Space.* arXiv:2604.02415, 2026.

AI for Physics

A1. **I. Elsharkawy** and Y. Kahn. *Contrastive Normalizing Flows for Uncertainty-Aware Parameter Estimation.* arXiv:2505.08709, 2025.

A2. **I. Elsharkawy**, et al. *OmniMol: Transferring Particle Physics Knowledge to Molecular Dynamics with Point-Edge Transformers.* arXiv:2601.10791, 2026.

A3. W. Bhimji, et al. *FAIR Universe HiggsML Uncertainty Dataset and Competition.* **NeurIPS 2025 Datasets & Benchmarks Track**.

A4. V. Mikuni, **I. Elsharkawy**, et al. *OmniCosmos: Transferring Particle Physics Knowledge across the Cosmos.* arXiv:2512.24422, 2025.

TALKS & POSTERS

Invited Talk. *1st-Place Competition Milestone: Ensembles and Uncertainty Quantification.* The Challenge of Handling Uncertainties in Fundamental Science, NeurIPS 2024.

Invited Talk. *Contrastive Normalizing Flows for Uncertainty-Aware Parameter Estimation.* 7th Inter-Experimental LHC Machine Learning Workshop, CERN, 2025.

Poster. *Pre-Training for Science: A Study on Foundation Model Training Objectives.* Stanford Center for Decoding the Universe Forum, 2025.

Poster. *FAIR Universe HiggsML Uncertainty Dataset and Competition.* Competition Track, NeurIPS 2025.

Poster. *Uncertainty Quantification from Scaling Laws in Deep Neural Networks.* Machine Learning and the Physical Sciences Workshop, NeurIPS 2024.

AWARDS

2026	NSERC Doctoral Canadian Graduate Research Scholarship (CGRS D)
2025	Connaught International Fellowship, University of Toronto
2024	1st Place, Higgs Uncertainty Challenge, NeurIPS 2024
2023	UIUC Graduate College 3-Year Fellowship
2021	J. and M. Graham Physics Scholar, Rice University
2016	Eagle Scout

TEACHING

TA and Guest Lecturer PHY 2108/2109, *The Physics of Machine Learning*, University of Toronto, 2026.

TA. PHY 252, *Thermal Physics*, University of Toronto, 2025.

TA. PHYS 413, *Optics Laboratory*, UIUC, 2024.

SKILLS

Languages	Python (proficient), C/C++, MATLAB, Fortran, Bash, HTML/CSS, LaTeX
Preferred ML Libraries	PyTorch, Lightning, Weights & Biases, Hugging Face, TensorFlow
Methods	Generative models (flow matching, normalizing flows, diffusion), transformers (Point-Edge Transformers, ViTs, LLMs), Fourier Neural Operators, foundation model pre-training, uncertainty quantification and Bayesian inference
Compute/Infrastructure	Large Scale Distributed training (running on $\mathcal{O}(500)$ nodes on NERSC Perlmutter), Slurm, CUDA, Docker, Git