

FERNANDO TORALES ACOSTA

✉ fernandotta@gmail.com
🌐 fernando-torales-acosta

🌐 ftoralesacosta

EXPERIENCE (YRS 5+)

Postdoctoral Research Fellow

Lawrence Berkeley National Lab

📅 Feb 2022 – Present 📍 Berkeley, CA

- Spearheaded a diverse team of researchers to leverage DeepSets and Graph Neural Networks for detector design
- Emphasized MLOps and AutoML practices to streamline design process
- Achieved a 50% improvement compared to traditional methods
- Established a comprehensive data pipeline from the ground up, facilitating the generation of synthetic data for training state-of-the-art machine learning models. Implemented efficient job distribution on HPC systems, rigorous data quality assurance procedures, independent model validation, and standalone inference using final models.
- Investigated optimal use of state-of-the-art generative diffusion models. Compared models made up of CNNs trained on images to models consisting of DeepSets trained on Point Clouds. Resulted in 150x reduction in disk space, and a 300% reduction in sampling time, all while maintaining cutting-edge performance.
- Generalized bayesian data correction to use DNNs for multidimensional, scalable correction. Trained thousands of classifiers on HPC systems.

PyTorch TensorFlow HDF5 Python Git C++

Nuclear and Particle Physics Graduate Researcher

University of California, Berkeley & CERN, LHC

📅 Aug 2016 – May 2021 📍 Berkeley, CA

- Presented methodology to international collaborators.
- Refactored codebase from procedural to functional paradigm. Parallelized refactored code using OpenMP. Implemented novel data correction algorithm, decreasing processing time from 2 weeks to just 1 hour.
- Trained classifier for background rejection, obtaining 30% improvement over previous algorithm shown to be robust for 15 years.
- Modernized data pipeline converting petabytes of data to HDF5 format. Sped up analysis team's workflows by 4-10x.

C C++ Python Jupyter TensorFlow Git HDF5

SKILLS

Programming & Other Languages

Python C C++ \LaTeX Lua YAML JSON Fortran

Frameworks, Libs, & Databases

Tensorflow PyTorch NumPy HDF5 Jupyter Keras
Onnx Pandas OpenMP MPI

Development Tools & Environment

Git UNIX/Linux BASH Vim GNU

EDUCATION

University of California, Berkeley

PhD Nuclear and Particle Physics

📅 2016 – 2021 📍 Berkeley, CA

Stony Brook University

B.S. Physics

📅 2012 – 2016 📍 Stony Brook, NY

PROJECTS

GNNs_4_Pions

Particle Feature Regression with Graph Neural Networks

📅 Feb 2023 🌐 ftoralesacosta/GNNs_4Pions_EIC

- Leverage DeepSets and Graph Neural Networks in the development of physical detectors
- Systematically varies input features, from floats to bidirectional graphs
- Established full data pipeline, data QA, model validation, and inference scripts for teammates

TensorFlow Python HDF5 C++

Diffusion: Images vs Point Clouds

Benchmarking Stable Diffusion Models

📅 Jul 2023 🌐 ftoralesacosta/GSGM_for_EIC_Calo

- Compared diffusion models based Images/CNNs to models based on Point Clouds/GNNs
- Created point cloud to image converter to compare generated samples
- Trained classifier on generated and original images to benchmark performance

Tensorflow Python PyTorch HDF5

Deconvolution for Collider Data

Training Classifiers for Data Correction

📅 Jul 2023 🌐 ftoralesacosta/h1_asymmetry

- Using ML-based deconvolution to measure moments defining properties of physics collisions
- Use classifiers as likelihood-ratio estimators, reweighting synthetic data to match real data
- Trained thousands of classifiers on hundreds of NVIDIA A100 GPUs, visualizing weights and final outputs of models

Tensorflow Python HOROVOD

RELEVANT PAPERS

- Point Cloud vs. Image based generative diffusion: **NeurIPS 2023 #192**
- Optimal design of detectors using DeepSets and GNNs: **arxiv: 2310.04442**
- Isolated Photon-Hadron Correlation in ALICE: **Phys. Rev. C 102, 044908**