Patrick T. Komiske III

pkomiske@mit.edu energyflow.network | pkomiske.com | github.com/pkomiske +1 (443) 690-3299

Summary

Theoretical physicist seeking research positions involving machine learning, data science, and quantitative problem solving. Significant research experience at the interface of machine learning and particle physics. Highly skilled at solving complex quantitative problems, developing custom ML-based solutions, as well as designing and deploying effective software.

Education

Ph.D. Candidate in Physics – Massachusetts Institute of Technology Cambridge, MA
Graduate Researcher, Center for Theoretical Physics, advised by Prof. Jesse Thaler – GPA: 5.0/5.0 *May 2021 (expected)*

A.M., A.B. (*summa cum laude*) – Harvard University Cambridge, MA
Master of Arts in Physics – GPA: 3.90/4.00 *May 2016*
Bachelor of Arts in Physics (highest honors) and Mathematics, secondary field in Computer Science – GPA: 3.93/4.00 *May 2016*

Research/Work Experience

Massachusetts Institute of Technology Cambridge, MA
Graduate Researcher in the Center for Theoretical Physics *September 2016 - Present*

- Conducted particle physics research with focus on Large Hadron Collider phenomenology and data analysis strategies. Developed and applied novel machine learning strategies for high-energy physics including custom neural network architectures [1,2,6], weakly-supervised training paradigms [4, 5], non-parametric methods [7, 11], and improved high-dimensional unfolding [10]. Calculated distributions of jets physics observables in QCD and Soft-Collinear Effective Theory.
  - Synthesized insights from quantum field theory and machine learning to develop Energy/Particle Flow Networks to process jets as (weighted) point clouds in a highly-efficient, performant, and simple architecture [6]. EFNs are particularly robust and interpretable due to respecting "infrared safety," an important constraint imposed by quantum field theory.
  - Adapted optimal transport techniques for use in high-energy physics as the "Energy Mover's Distance," defining geometric spaces of "events" and "theories" [7, 11]. This work was featured in Physical Review Letters – a top journal in physics – Physics Magazine, MIT News, and other news outlets, and was presented in multiple invited talks and seminars.
- Designed software libraries to provide high-performance and easy-to-use implementations of novel quantitative strategies
  - Lead developer of the EnergyFlow Python package, a suite of particle physics tools including dense, convolutional, and point cloud (Energy/Particle Flow) neural network architectures in Keras/Tensorflow; vectorized implementations of Energy Flow Polynomials/Moments, collider event utilities, and other observables; and simple interfaces to automatically load particle physics datasets from Zenodo. Fully documented with mkdocs/readthedocs, tested via pytest and Travis CI.
  - Lead developer of the Wasserstein package for computing Wasserstein distances (a.k.a. the Earth/Energy Mover's Distance), via the network simplex algorithm. Wasserstein is 50-100% faster than the Python Optimal Transport library on small computations such as those in jet physics and allows for greater control over numerical accuracy. Implemented in C++ with a Python wrapper generated by SWIG, it is designed to seamlessly interface with the FastJet library.
  - Lead developer of the EnergyEnergyCorrelators Package for computing the eponymous particle physics quantities. Written in C++, Cython, and Python, utilizes the BOOST Histogram library, designed to interface easily with FastJet.
- Processed and released complex datasets previously usable only by physicists into formats easily-utilized by non-domain experts
  - Reprocessed multi-TB datasets of events from the CMS Open Data 2011 Jet Primary Dataset and associated simulated datasets into a custom HDF5-based format easily usable via the EnergyFlow package, thereby enabling analysis of real jets recorded by the CMS detector in just a few lines of Python code [8]. Primary dataset hosted as Zenodo record 3340205.
  - Generated high-quality datasets of simulated jets using the Pythia, Herwig, FastJet, and Delphes physics libraries, enabling comparison of different machine learning models and other techniques on standardized datasets [2, 6, 10]. Easily accessible via EnergyFlow Python package; hosted as Zenodo records 3164691, 3066475, 3548091, 2652034.

Jane Street Capital New York, NY
Winter Intern – Trading *January 2015, 2016*

- Studied financial markets, wrote bash program to analyze novel type of options trade, participated in mock trading

Northrop Grumman Electronic Systems Baltimore, MD
Summer Intern – Superconducting Electronics Group, Quantum Computing Collaboration *Summer 2014*

- Wrote MATLAB program to improve fidelity of high-speed, precision microwave pulses for qubit control via deconvolution

Johns Hopkins University Applied Physics Laboratory Laurel, MD
Summer Intern – Asymmetric Operations and Research and Exploratory Development Departments *Summer 2012, 2013*

- Investigated electromagnetic properties of high-impedance Sievenpiper metamaterial structures for low-profile RF antennas

PUBLICATIONS

All authors alphabetical; 702 total citations on Google Scholar as of January 1, 2021.
Referee for PRL, PRD, Machine Learning for the Physical Sciences Workshop at NeurIPS 2020

[11] P. Komiske, E. Metodiev, J. Thaler. *The Hidden Geometry of Particle Collisions*. JHEP 07 (2020) 006 [arXiv:2004.04159]

[10] A. Andreassen, P. Komiske, E. Metodiev, B. Nachman, J. Thaler. *OmniFold: A Method to Simultaneously Unfold All Observables*. PRL 124 (2020) 182001 [arXiv:1911.09107]

[9] P. Komiske, E. Metodiev, J. Thaler. *Cutting Multiparticle Correlators Down to Size*. PRD 101 (2020) 036019 [arXiv:1911.04491]

[8] P. Komiske, R. Mastandrea, E. Metodiev, P. Naik, J. Thaler. *Exploring the Space of Jets with CMS Open Data*. PRD 101 (2020) 034009 [arXiv:1908.08542] ***Featured* in ML and the Physical Sciences 2019 Workshop at NeurIPS.**

[7] P. Komiske, E. Metodiev, J. Thaler. *The Metric Space of Collider Events*. PRL 123 (2019) 041801 [arXiv:1902.02346] ***PRL Editors' suggestion; Featured in Physics Magazine.***

[6] P. Komiske, E. Metodiev, J. Thaler. *Energy Flow Networks: Deep Sets for Particle Jets*. JHEP 01 (2019) 121 [arXiv:1810.05165]

[5] P. Komiske, E. Metodiev, B. Nachman, M. Schwartz. *An operational definition of quark and gluon jets*. Journal of High Energy Physics JHEP 11 (2018) 059 [arXiv:1809.01140]

[4] P. Komiske, E. Metodiev, B. Nachman, M. Schwartz. *Learning to classify from impure samples with high-dimensional data*. PRD 98 (2018) 011502 [arXiv:1801.10158]

[3] P. Komiske, E. Metodiev, J. Thaler. *Energy flow polynomials: A complete linear basis for jet substructure*. JHEP 04 (2018) 013 [arXiv:1712.07124]

[2] P. Komiske, E. Metodiev, B. Nachman, M. Schwartz. *Pileup Mitigation with Machine Learning (PUMML)*. JHEP 12 (2017) 051 [arXiv:1707.08600]

[1] P. Komiske, E. Metodiev, M. Schwartz. *Deep learning in color: Towards automated quark/gluon jet discrimination*. JHEP 01 (2017) 110 [arXiv:1612.01551]

JHEP = Journal of High Energy Physics
PRD = Physical Review D (Particles and Fields)
PRL = Physical Review Letters

PRESENTATIONS

Invited Talks/Seminars

- *Simultaneously Unfolding All Observables with Deep Learning*. Jefferson Lab Theory Seminar (virtual). *January 2021*
- *Probing QCD with Energy Flow Observables*. CEPC Workshop, Shanghai, China (virtual). *October 2020*
- *Machine Learning – An essential toolkit for particle physics*. Snowmass Computational Frontier Workshop (virtual). *August 2020*
- *The Hidden Geometry of Particle Collisions*. Particle Physics Phenomenology Series, University of Genoa (virtual). *June 2020*
- *The Metric Space of Collider Events*. Particle Physics Seminar, University of Chicago. *May 2019*
- *The (Metric) Space of Collider Events*. Elementary Particle Theory Seminar, University of Maryland. *March 2019*

Plenary Talks

- *OmniFold: Simultaneously Unfolding All Observables*. ML4Jets 2020, New York University. *January 2020*
- *Cutting Multiparticle Correlators Down to Size*. BOOST 2019, MIT. *July 2019*
- *The Metric Space of Collider Events*. Deep Learning in the Natural Sciences, University of Hamburg. *March 2019*
- *Point Cloud Strategies for Boosted Objects*. BSM Forum, CERN. *February 2019*
- *Energy Flow Networks: Deeps Sets for Particle Jets*. Machine Learning for Jet Physics, Fermilab. *November 2018*
- *Energy Flow and Jet Substructure*. BOOST 2018, Paris. *July 2018*
- *Linear Jet Tagging with the Energy Flow Basis*. Machine Learning for Jet Physics, Berkeley National Lab. *December 2017*
- *Quark/Gluon Discrimination with Jet-Images and Deep Learning*. BOOST 2017, University at Buffalo. *July 2017*

HONORS

- *Summa cum laude*, Harvard College *2016*
- Highest Honors, Harvard Physics Department *2016*
- John Harvard Scholarship (top 5% of class) *2014-2015*
- Derek C. Bok Award for Distinction in Teaching *2014*
- University Physics Competition Silver Medal *2014*
- Harvard College Scholarship (top 10% of class) *2013-2014*
- Century High School STEM Academy Award *2012*
- National AP Scholar *2011*

RELATED EXPERIENCE

Cloud and Cluster Computing – Microsoft Azure, Google Cloud Platform, Odyssey Cluster, MIT Supercloud, Cori (at NERSC)
Managed multiple virtual machines in the cloud both with and without GPUs; used SLURM to run high-performance cluster jobs

MIT Physics Teaching Assistant
8.09/8.309 Advanced Classical Mechanics (Prof. Iain Stewart) *Fall 2017, 2018, 2019*

MIT Physics Graduate Student Council – Colloquium Representative and Lunch Organizer *Spring 2017 – Fall 2018*

Harvard Physics Teaching Fellow
Intro. Mechanics and Special Relativity (Prof. Howard Georgi), Quantum Mechanics I (Prof. Matthew Reece) *Fall 2014, 2015*

Harvard-Radcliffe Society of Physics Students *2012 – 2016*
Event Coordinator – co-organizer of first Harvard-MIT SPS Research Conference for undergraduates *2015 – 2016*