

# Ho Fung Tsoi

Postdoctoral Researcher

Department of Physics and Astronomy, University of Pennsylvania

[hftsoi@sas.upenn.edu](mailto:hftsoi@sas.upenn.edu)

Profile: [hftsoi.github.io](https://hftsoi.github.io)

Last updated: Dec 2025

## Research Interests

---

Experimental particle physics at the CERN [Large Hadron Collider](#) (ATLAS experiment) for new physics searches beyond the Standard Model; development of novel machine learning algorithms: self-supervised pretraining, anomaly detection, symbolic regression; low-latency (sub-microsecond) machine learning algorithms on FPGAs.

## Education

---

**University of Wisconsin–Madison**, Ph.D. Physics 2024  
Thesis: [Search for exotic Higgs boson decays with CMS and fast machine learning solutions for the LHC](#)  
Advisor: Sridhara Dasu

**The Chinese University of Hong Kong**, B.Sc. Physics 2018  
Thesis: First passage time problem of the time-dependent Ornstein-Uhlenbeck process  
Advisor: Chi-Fai Lo  
Visiting student at University of California Berkeley & Berkeley Lab (Jan – Aug 2017)

## Professional Experience

---

**Postdoctoral Researcher**, University of Pennsylvania (Supervisor: Dylan Rankin) 2024 – Present

## Awards and Honors

---

CERN, [CMS PhD Thesis Award](#) 2024  
UWisconsin, Cornelius P. and Cynthia C. Browne Fellowship 2024  
CUHK, CN Yang Scholarship 2017  
CUHK, Wei Lun Foundation Exchange Scholarship 2017  
CUHK, Professor Dennis Yam Kuen Lo Physics Award 2016

## Selected Research Experience

---

### New Physics Search at the LHC

- Lead analyst of the CMS Run 2 search for exotic Higgs boson decays into pseudoscalars in the final state of two b quarks and two  $\tau$  leptons, and of the combination with the  $\mu\mu b\bar{b}$  final state [8] (2020–2024).

### Novel ML Algorithms for Particle Physics

- Self-supervised pretraining methods for jet tagging [1] and mass regression (2025–Present).
- Symbolic regression for automating parametric modeling of binned distributions in LHC physics analyses [6] (2024–2025).

### Low-Latency (Sub-Microsecond) ML Algorithms on FPGAs

- Neuromorphic spiking neural networks (SNNs) for  $dN/dx$  clustering counting in drift chambers (2025–Present).
- Sparse CNNs for spatially sparse image data on FPGAs [4] (2024–Present).
- Compression of neural networks using symbolic regression for fast inference [7][9] (2022–2024).

### Trigger System at the LHC

- Commissioning of the GNN-based tau trigger (GNTau) at the ATLAS High-Level Trigger for Run 3 [3] (2024–Present).
- Development of the CICADA anomaly detection trigger algorithm at the CMS Level-1 Trigger [2][10] (2022–2024).
- Development of the Data Quality Monitoring (DQM) software for the Calorimeter Layer-1 subsystem at the CMS Level-1 Trigger (2020–2024).
- HL-LHC trigger sensitivity projection for the CMS searches for exotic Higgs boson decays [12] (2020–2021).

## Selected Publications

---

Listed below are selected publications, conference proceedings, and technical reports to which I was a primary author or made significant contributions (\*: refereed). I also co-authored papers by [CMS](#) and [ATLAS](#) as a collaboration member. For the complete list see [INSPIRE](#). Preprints at [arXiv](#).

### 2025

- [1] **“jBOT: Semantic Jet Representation Clustering Emerges from Self-Supervised Pre-Training”**  
[H.F. Tsoi](#), D. Rankin, *manuscript in preparation*.
- [2] **“Anomaly Detection in the CMS Level-1 Trigger in Run 3”**  
CMS Collaboration, *manuscript in preparation*.
- [3] **“Performance of the ATLAS Tau Trigger in Run 3”**  
ATLAS Collaboration, *manuscript in preparation*.
- [4] **“SparsePixels: Efficient Convolution for Sparse Data on FPGAs”**  
[H.F. Tsoi](#), D. Rankin, V. Loncar, P. Harris, submitted to *ACM Trans. Reconf. Tech. Syst.* [[arXiv:2512.06208](#)].
- [5] **“hls4ml: A Flexible, Open-Source Platform for Deep Learning Acceleration on Reconfigurable Hardware”**  
J.F. Schulte et al., submitted to *ACM Trans. Reconf. Tech. Syst.* [[arXiv:2512.01463](#)].
- [6\*] **“SymbolFit: Automatic Parametric Modeling with Symbolic Regression”**  
[H.F. Tsoi](#), D. Rankin, C. Caillol, M. Cranmer, S. Dasu, J. Duarte, P. Harris, E. Lipeles, V. Loncar, *Comput Softw Big Sci* **9**, 12 (2025) [[arXiv:2411.09851](#)].
- [7\*] **“SymbolNet: Neural Symbolic Regression with Adaptive Dynamic Pruning for Compression”**  
[H.F. Tsoi](#), V. Loncar, S. Dasu, P. Harris, *Mach. Learn.: Sci. Technol.* **6**, 015021 (2025) [[arXiv:2401.09949](#)].

### 2024

- [8\*] **“Search for Exotic Decays of the Higgs Boson to a Pair of Pseudoscalars in the  $\mu\mu b\bar{b}$  and  $\tau\tau b\bar{b}$  Final States”**  
CMS Collaboration, *Eur. Phys. J. C* **84**, 493 (2024) [[arXiv:2402.13358](#)].
- [9\*] **“Symbolic Regression on FPGAs for Fast Machine Learning Inference”**  
[H.F. Tsoi](#), A.A. Pol, V. Loncar, E. Govorkova, M. Cranmer, S. Dasu, P. Elmer, P. Harris, I. Ojalvo, M. Pierini, *EPJ Web of Conferences* **295**, 09036 (2024) [[arXiv:2305.04099](#)].

### 2023

- [10] **“Level-1 Trigger Calorimeter Image Convolutional Anomaly Detection Algorithm”**  
CMS Collaboration, [CMS Detector Performance Summaries \(2023\)](#), CERN-CMS-DP-2023-086.
- [11] **“Searches for exotic Higgs boson decays with the CMS experiment”**  
[H.F. Tsoi](#), [Proceedings of the European Physical Society Conference on High Energy Physics \(PoS EPS-HEP2023 402\)](#).

### 2021

- [12] **“The Phase-2 Upgrade of the CMS Data Acquisition and High Level Trigger”**  
CMS Collaboration, [CMS Technical Design Report \(2021\)](#), CERN-LHCC-2021-007, CMS-TDR-022.

## Conference, Workshop, Seminar Talks

---

Coordinating Panel on Advanced Detectors (CPAD) — UPenn, USA <a href="#">SparsePixels: Efficient Convolution for Sparse Data on FPGAs</a>	Oct 2025
Fast Machine Learning for Science Conference (FastML) — ETH Zurich, Switzerland <a href="#">SparsePixels: Efficient Convolution for Sparse Data on FPGAs</a>	Sep 2025
High-energy physics seminar — UPenn, U.Washington, Fermilab <a href="#">Search for Exotic Higgs Boson Decays with CMS and Fast Machine Learning Solutions for the LHC</a>	Nov 2023, Mar 2024
US LHC Users Association Meeting (US LUA) — Fermilab, USA <a href="#">CICADA: Anomaly Detection for New Physics Searches at the CMS Level-1 Trigger</a>	Dec 2023

Machine Learning at Level-1 Trigger Workshop (ML@L1) — CERN, Switzerland Anomaly Detection – CICADA: Status, Plans, and Prospects for Phase-2	Dec 2023
CMS Machine Learning Town Hall — CERN, Switzerland L1 Anomaly Detection with Calorimeter Inputs: Status and Opportunities	Sep 2023
European Physical Society Conference on High Energy Physics (EPS-HEP) — Hamburg, Germany <a href="#">Searches for exotic Higgs boson decays with the CMS experiment</a>	Aug 2023
International Conference on Computing in High Energy & Nuclear Physics (CHEP) — Norfolk VA, USA <a href="#">Symbolic Regression on FPGAs for Fast Machine Learning Inference</a>	May 2023
International Conference on Applied Mathematics — CityU, Hong Kong First Passage Time Problem of the Time-Dependent Ornstein-Uhlenbeck Process: a Model for Stochastic Decision-Making Process	May 2016

## Professional Service

---

### Membership

• Member, <a href="#">ATLAS Collaboration</a> , CERN	2024 – Present
• Affiliated member, <a href="#">Accelerated AI Algorithms for Data-Driven Discovery (A3D3) Institute</a>	2024 – Present
• Member, <a href="#">CMS Collaboration</a> , CERN	2020 – 2024

### Collaboration Roles

• <b>Co-convener</b> , SUSY Monte Carlo & Interpretation subgroup [CMS]	2023 – 2024
• <b>Contact person</b> , Higgs boson Monte Carlo group [CMS]	2021 – 2023
• <b>Software developer</b> , Data Quality Monitoring system for CaloLayer-1 trigger [CMS]	2020 – 2024

### Other service

- **Journal referee**, *Advanced Theory and Simulations*.

## Teaching and Mentoring

---

### Teaching

- Guest speaker, PHYS 3359 [UPenn]: Data Analysis for the Natural Sciences II: Machine Learning (Spring 2025)
- Teaching assistant, undergraduate physics courses [UW–Madison] (Fall 2018 – Spring 2020)

### Master students

- Kam Wai Lai [UPenn]: Neuromorphic AI with spiking neural networks on FPGAs (2025 – Present)
- Yuyan Wang [UPenn]: Self-supervised learning with JEPA-based pretraining for LHC jet tagging (2025 – Present)

### Undergraduate students

- Alex Yang [UPenn]: Self-supervised learning with SimCLR/VICReg for reconstructing the mass of supersymmetric particles (2025 – Present)
- Ashni Kumar [Drexel]: Acceleration of reinforcement learning agent inference on FPGAs for astronomical observations (2025)