Zichun Hao

zichunhao.com • LinkedIn GitHub Google Scholar

EDUCATION

California Institute of Technology *Sep 2023 – Present*
- Ph.D. candidate in Physics, Division of Physics, Mathematics and Astronomy (PMA)
- Master of Science in Physics, received June 2026
- Thesis advisor: Prof. Maria Spiropulu; research in machine learning and experimental high-energy physics

University of California, San Diego *Sep 2020 – June 2023*
- Bachelor of Science, double major in Physics and Computer Science; Phi Beta Kappa

University of California, Santa Barbara *Sep 2018 – Aug 2020*
- Undergraduate study in Physics

RESEARCH EXPERIENCE

Caltech High Energy Physics Group | Graduate Research Assistant *Sep 2023 – Present*
- Advisor: Prof. Maria Spiropulu
- Research at the Compact Muon Solenoid (CMS) Experiment at CERN and on deep learning methods for physics: self-supervised foundation models, neural-network tagger calibration, trigger development, and detector R&D

Fermi National Accelerator Laboratory (Fermilab) | Visitor *July – Nov 2024; July – Sep 2025*
- Supervisor: Prof. Si Xie
- Self-supervised learning methods for high-energy physics; search for double Higgs (HH) production with CMS data at the Large Hadron Collider

Duarte Lab | Undergraduate Researcher *Jan 2021 – June 2023*
- Mentors: Prof. Javier Duarte and Dr. Raghav Kansal
- Geometric deep learning for experimental high-energy physics: Lorentz group equivariant jet autoencoders (LGAE, first author) and a ParticleNet-based tagger for boosted Higgs boson decays to W boson pairs for the CMS experiment

RESEARCH PROJECTS

RINO: Renormalization Group Invariance with No Labels *Nov 2024 – Present*
First author
- Improved out-of-distribution accuracy and generalization to unseen Higgs decay channels over the strongest fully-supervised baseline, without retraining
- Pretrained transformer encoders with DINO/DINOv2-style self-distillation (iBOT masked-token prediction, KoLeo regularization), using multi-scale k_{T}-clustering views as a physically grounded, renormalization-group-inspired source of self-supervised augmentations, which ablations confirm drive the gains
- Selected as a **spotlight** paper at the ML4PS workshop @ NeurIPS 2025 (offered to ∼1% of accepted papers), arXiv:2509.07486

Search for Non-Resonant HH Production in the $\mathrm{b\bar{b}b\bar{b}}$ Final State *Apr 2024 – Present*
Core contributor, CMS Collaboration
- One of the main contributors to the boosted HH→ $\mathrm{b\bar{b}b\bar{b}}$ search, responsible for BDT trainings, datacard creation, and the final statistical fits in the continuing analysis of the 2024 and 2025 CMS data
- Developing sensitivity improvements for this iteration, including a transformer-based approach and data-driven QCD background estimation
- Performed calibration studies of transformer- and graph-based b-taggers (GloParT, ParticleNet), including data/simulation scale factors via the sfBDT method and $\mathrm{Z} \to \mathrm{b\bar{b}}$ control-region cross-checks
- Measured trigger efficiencies and corrections for τ_3/τ_2 and tagger score distributions on $\mathrm{t\bar{t}}$ MC samples; processed and managed data and simulation events for the calibration studies
- Presented calibration results at the CMS B-Tagging and Vertexing (BTV) Working Group [link] and CMS Higgs Physics Group [link] meetings
- Contributed the trigger and neural-network jet tagger calibrations to the public analysis CMS-HIG-24-010 (*Phys. Rev. D*, 2026), which sets the most stringent constraints on HH production in the 4b final state to date

Development of the Barrel Timing Layer (BTL) at the CMS Experiment *Oct 2023 – Present*

- Contributed to the quality assurance and quality control (QA/QC) procedure
- Assembled detector modules and trays; performed thermal cooling studies of the module
- Test beam results published in *Nucl. Instrum. Meth. A* **1081**, 170823 (2026)

Search for Highly Energetic HH Production in the $b\bar{b}VV$ Final State *Mar – Apr 2024*

- Optimized event selection based on signal and matching efficiencies; designed a new physics-inspired strategy for choosing signal events
- Analysis summary CMS-PAS-HIG-23-012 made public

Development of Muon Detector Shower (MDS) Cross Triggers *Nov 2023 – Apr 2024*

- Developed the electron+MDS and muon+MDS cross triggers, accepted into the CMS trigger menu in 2023; processed large datasets with parallel workflows
- Trigger system published in *JINST* **21**, P07040 (2026), arXiv:2606.27165

Lorentz Group Equivariant Jet Autoencoders (LGAE) *Jan 2021 – June 2023*

First author; with Prof. Javier Duarte (UC San Diego), Dr. Raghav Kansal (UC San Diego, Fermilab), and Dr. Nadezda Chernyavskaya (CERN)

- Built an autoencoder fully equivariant to the Lorentz group, with a latent space in Lorentz-group irreducible representations, using equivariant message passing with Clebsch–Gordan tensor products
- Outperformed graph- and CNN-based autoencoder baselines on data compression, reconstruction, and unsupervised anomaly detection, with robust performance when trained on only 1% of data
- Published in *Eur. Phys. J. C* **83**, 485 (2023); ∼35 citations

Development of the Tagger for $H \to WW^*$ *June 2021 – Jan 2022*

Supported by the UC San Diego Undergraduate Summer Research Award; with Prof. Javier Duarte (UC San Diego), Dr. Raghav Kansal (UC San Diego, Fermilab), Dr. Cristina Mantilla Suarez (Fermilab), and Prof. Si Xie (Caltech)

- Developed a neural-network tagger for $H \to WW^*$ decays against QCD backgrounds on the ParticleNet architecture (arXiv:1902.08570) for the CMS experiment; responsible for model training on Kubernetes clusters
- Presented at the APS April Meeting 2022

TEACHING

Caltech Ph 22 (Computational Physics Laboratory III) | Teaching Assistant *Mar – Jun 2026*

- Led weekly lab sessions and taught core AI concepts, from neural-network training to policy-gradient methods; graded coursework and mentored students through their final projects
- One of two TAs who designed the course's final project on reinforcement learning for quantum state preparation: built the Gymnasium environment over a Tavis–Cummings cavity-QED simulator, the REINFORCE reference implementation, and PPO/SAC/TQC baselines (Stable-Baselines3) with MLP/LSTM/Transformer policies
- Designed the project questions, the graded ladder of target quantum states, reward shaping, and multi-seed evaluation requirements

PUBLICATIONS

- Zichun Hao, Raghav Kansal, Abhijith Gandrakota, Chang Sun, Jennifer Ngadiuba, Javier Duarte, Maria Spiropulu, "RINO: Renormalization Group Invariance with No Labels," Nov. 2025, arXiv:2509.07486 (ML4PS workshop **spotlight** paper, offered to ∼1% of accepted papers; poster)
- Michael Chen, Raghav Kansal, Abhijith Gandrakota, Zichun Hao, Jennifer Ngadiuba, Maria Spiropulu, "An Evaluation of Representation Learning Methods in Particle Physics Foundation Models," Nov. 2025, arXiv:2511.12829 (ML4PS paper; poster) — served as a mentor on this work
- Raghav Kansal, Carlos Pareja, Zichun Hao, Javier Duarte, "JetNet: A Python Package for Accessing Open Datasets and Benchmarking Machine Learning Methods in High Energy Physics," *J. Open Source Softw.* **8**(90), 5789 (2023), doi:10.21105/joss.05789
- Zichun Hao, Raghav Kansal, Javier Duarte, Nadezda Chernyavskaya, "Lorentz Group Equivariant Autoencoders," *Eur. Phys. J. C* **83**, 485 (2023), doi:10.1140/epjc/s10052-023-11633-5
- F. Addesa et al. (incl. Zichun Hao), "The CMS Barrel Timing Layer: Test Beam Confirmation of Module Timing Performance," *Nucl. Instrum. Meth. A* **1081**, 170823 (2026), arXiv:2504.11209 — contributor
- M. Abbott et al. (incl. Zichun Hao), "High Multiplicity Trigger for Long-Lived Particles in CMS Detector," *JINST* **21**, P07040 (2026), doi:10.1088/1748-0221/21/07/P07040 — contributor

- CMS Collaboration, "Improved results on Higgs boson pair production in the 4b final state," *Phys. Rev. D* (2026), doi:10.1103/w2zp-fckl, arXiv:2604.27044 — core contributor
- CMS Collaboration, "Search for highly energetic double Higgs boson production in the two bottom quark and two vector boson all-hadronic final state," CMS-PAS-HIG-23-012, CERN (2024) — contributor
- CMS Collaboration, "Particle transformers for identifying Lorentz-boosted Higgs bosons decaying to a pair of W bosons," CMS-PAS-JME-25-001, CERN (2025), submitted to *JHEP* — trained the previous ParticleNet-based iteration during undergraduate years

CONFERENCES & WORKSHOPS

- Spotlight talk at the ML4PS workshop, 39th Conference on Neural Information Processing Systems (NeurIPS 2025), on RINO (Dec. 2025, San Diego)
- ML4Jets 2025 on representation learning of jets with physics-informed self-distillations (Aug. 18, 2025, Caltech)
- APS Global Physics Summit 2025 on neural-network tagger calibrations in the search for nonresonant Higgs boson pair production in the four bottom quark decay channel with Run-3 CMS data (Mar. 20, 2025)
- ML4Jets 2024 on Lorentz group equivariant jet autoencoders (Nov. 7, 2024, Paris)
- APS Far West Section Annual Meeting 2023 on Lorentz group equivariant jet autoencoders (Oct. 7, 2023)
- APS April Meeting 2022 on the $H \to WW^*$ tagger (Apr. 9–12, 2022)
- ML4Jets 2021 on Lorentz group equivariant jet autoencoders (July 6, 2021)
- UC San Diego Online Undergraduate Research Symposium 2021 on Lorentz group equivariant jet autoencoders (May 26, 2021)

HONORS & AWARDS

- **Spotlight** paper (first author), ML4PS workshop @ NeurIPS 2025 — offered to ∼1% of the workshop's accepted papers
- Fermilab LPC Guest and Visitor Award (2024 & 2025)
- Phi Beta Kappa Honor Society (Apr. 2022)
- UC San Diego Physical Sciences 2022 Dean's Undergraduate Award for Excellence
- UC San Diego 2021 Undergraduate Summer Research Award
- UC San Diego Provost Honors (Fall 2020, Winter 2021, Spring 2021, Fall 2021, Winter 2022, Spring 2022)
- UC Santa Barbara Dean's Honors (Winter 2019, Spring 2019, Fall 2019, Winter 2020, Spring 2020)