

PRACHURJYA PRAN HAZARIKA

Research Scholar in Experimental High Energy Physics
Indian Institute of Science Education and Research, Pune



CONTACT

Main Building, Room #A367
Department of Physics, IISER-Pune
Dr. Homi Bhabha Road, Pune, India, 411008

(+91) 8136028266
prachurjyahazarika@gmail.com
prachurjya.hazarika@cern.ch
GitHub: github.com/phazarik

RESEARCH INTERESTS

- **Physics analysis:** Broad interest in experimental particle physics, including searches for physics beyond the standard model, precision data analyses and projection studies for future high-energy experiments, spanning collider physics, neutrino experiments, and astroparticle physics.
- **Machine learning and computing:** Application of machine learning and statistical methods to high-energy physics problems, including event classification, object identification, and workflow automation for large-scale distributed analyses. Interests also include the use of agentic AI and generative models in HEP.
- **Detector R&D and software development:** Development of reconstruction systems and analysis tools for large-scale particle physics experiments, including algorithm design, data processing workflows, and performance optimization. This includes applications of machine learning methods and GPU/Alpaka-based computing techniques for fast and efficient data processing.

EDUCATION

2020-present	PhD in Physics , IISER-Pune, Maharashtra, India, 411008 Supervisor: Prof. Sourabh Dube Member of the CMS Collaboration since February, 2021.
2017-2019	MSc in Physics , Tezpur University, Assam, India, 784028 First class, first rank holder (gold medalist) with CGPA 8.92
2014-2017	BSc in Physics , Jorhat Institute of Science and Technology, Jorhat Affiliated to Dibrugarh University, Assam, India, 786004 First class, second rank with percentage 90.64

RESEARCH EXPERIENCE

ANALYSIS

2020-Present (ongoing)	Search for vector-like Leptons in 2L final state using Run 2 and Run 3 data in CMS Led and developed a search for vector-like leptons in proton-proton collisions using a dataset corresponding to 200 fb^{-1} collected between 2016 and 2023 at collision energies of 13 and 13.6 TeV, as part of a multilepton analysis in the CMS Collaboration. The analysis focuses on the scenario where these VLLs couple to the standard model light leptons (electrons and muons) with a mass hypothesis ranging from 100 GeV to 1.6 TeV. The analysis focuses on final states containing exactly two light leptons, where a dedicated machine learning strategy based on deep neural networks (DNNs) is implemented to define signal-enriched regions. In combination with the complementary multilepton final states, the search is sensitive to vector-like lepton doublet models with masses up to 1.1 TeV and singlet models beyond 300 GeV. The analysis is currently undergoing internal review within the CMS Collaboration.
---------------------------	--

SERVICE WORK FOR CMS

November,
2025

Detector-on-call for ECAL during Pb-Pb collisions

Served as the on-call person for the CMS Electromagnetic Calorimeter (ECAL) and the preshower during the heavy-ion collision period for two weeks.

- **Monitored** real-time ECAL and preshower activities to ensure data integrity, focusing on data quality, trigger rate stability, and identifying detector anomalies in real-time in the high-multiplicity heavy-ion environment.
- **Acted** as the primary interface between the central CMS data-taking team and the ECAL subsystem experts.
- **Coordinated** prompt interventions for hardware or software issues by contacting the right experts, minimizing detector downtime during critical physics runs.

2023-Present
(ongoing)

MC contact for the EGamma group

Serving as the liaison between the EGamma group, which handles electron and photon reconstruction in CMS, and the Physics Data and Monte-Carlo Validation (PdmV) group, which oversees large-scale Monte Carlo sample production and validation for CMS. Responsibilities include:

- **Reported** updates on the progress, availability, and timelines of central MC samples to the EGamma group.
- **Managed** specialized MC requests for performance studies and detector calibration. This includes creating new gridpacks or cloning existing requests and performing technical validation within the CMSSW framework prior to submission.
- **Monitored** the end-to-end life-cycle of official MC requests to ensure they satisfy the technical requirements and goals of the EGamma group.

2021-2022

Development of a DNN-based ID for photons in the Run 3 dataset

Developed a deep neural network (DNN)-based identification (ID) for photons within the EGamma POG as a high-performance alternative to the traditional cut-based Particle Flow ID.

- **Designed** and trained a multivariate DNN model utilizing ECAL shower shapes and cluster variables to effectively discriminate prompt photons from hadronic backgrounds and other fakes.
- **Optimized** and validated a specific working point that provides simultaneous improvements in both signal efficiency and background rejection compared to the cut-based ID.

This DNN-based ID is integrated into the central CMSSW_12+ framework, ready to be used in all photon-based physics analyses throughout Run 3.

MASTER'S THESIS

2018-2019

A Study on the Potential of Detection of NSI in Neutrino-Oscillation Experiments

Conducted a phenomenological investigation into the Non-Standard Interactions (NSI) in the neutrino sector and their impact on oscillation probabilities. The research utilized the Super-Kamiokande detector framework as a benchmark to quantify sensitivity toward exotic couplings beyond the standard model. By performing a χ^2 analysis across the energy-azimuthal angle phase space, the study identified specific regions of maximum sensitivity to NSI deviations, separately accounting for both normal and inverted mass hierarchy scenarios. This work mapped the potential discovery regions for NSI signatures in the context of the Super-K detector. This work was part of the thesis submitted in partial fulfillment of the requirements for the Master's degree.

COLLABORATIONS

2020-Present

CMS Collaboration, at the Large Hadron Collider, CERN, Geneva

SKILLS

Technical Skills

Programming languages

Proficient in C++, Python, and scientific scripting for large-scale data analysis and software development in high energy physics. Comfortable working with APIs, automation pipelines, HTML/CSS, PHP, scientific visualization and workflow optimization tools. Basic knowledge of JavaScript, GPU programming using CUDA and Alpaka.

HEP software

Extensive experience with ROOT, CMSSW, Monte Carlo production workflows and statistical tools within the CMS Collaboration.

Machine learning

Experience with TensorFlow, Keras, PyTorch, and the scientific Python ecosystem for high-energy physics applications, along with platform-independent model deployment using ONNX.

Languages

English: Fluent.

Assamese: Mother tongue.

Hindi: Fluent.

Bengali: Reading proficiency and strong listening comprehension.

Chinese (Mandarin): Currently learning (early A1 level).

Co-curricular

Science communication

Interested in science outreach and content writing in Assamese and English.

Digital media

Experience with graphic design and video editing using Adobe Photoshop, Premiere Pro, and After Effects.

Creative interests

Painting in oil, acrylic, and digital media, along with acoustic and electric guitar.

SOFTWARE TOOLS & TECHNICAL PROJECTS

CMS Open-Data analysis pipeline

🔗 [GitHub: phazarik/CMS-OpenData-analysis](#)

Pipeline for processing CMS open data, starting with making custom NTuples from CMS data format, to making event-level kinematic distributions.

CMS-Combine tutorial

🔗 [GitHub: phazarik/cms-combine-tutorial](#)

Tutorial repository for offline installation and example usage of the CMS-Combine tool. Utilization of RooStats and RooFit for statistical interpretation in EHEP.

NanoAOD analyzer & CRAB pipeline

🔗 [GitHub: phazarik/nanoAOD_analyzer](#), 🔗 [GitHub: phazarik/MakeSelector-CRAB-setup](#)

Analysis framework for CMS NanoAOD formats utilizing ROOT's MakeSelector() template. Compatible with the CRAB scheduler and capable of per-event, on-the-fly deep neural network evaluation using ONNX Runtime.

Delphes analyzer

🔗 [GitHub: phazarik/delphes_analyzer](#)

Fast-simulation analysis framework for processing Delphes ROOT trees through a customized MakeSelector() template.

AI-powered web application for PostDoc job search

🔗 [GitHub: phazarik/jobhunt-AI](#), [WebApp](#)

Streamlit web application that interfaces with the INSPIRE-HEP API and makes use of GROQ LLMs to intelligently sort, filter, and display postdoctoral opportunities in hep-ex.

CONFERENCES AND WORKSHOPS

- May 4–8 2026 **HSF-India Scientific Computing Workshop**  *IISER Pune, India*
Applications of machine learning for scientific analysis, GPU-based algorithms and Alpaka-based applications for HL-LHC reconstruction in CMS.
- Feb 14–17 2024 **Frontier Topics in Collider Physics**  *IISER Pune, India*
Discussion sessions on EFTs for interpreting collider data, unconventional signatures such as collimated objects, disappearing tracks, inward moving objects, slow moving objects etc. and how to deal with them using ML or specialized detectors.
- Dec 18–20 2023 **Phoenix 2023**  *IIT Hyderabad, India*
Discussion sessions on dark matter models, search proposals for multiple BSM higgs models in HL-LHC and recent anomalies in CMS and ATLAS searches.
- Aug 12–13 2023 **India-CMS collaboration meeting 2023**  *CHEP, IISc Bangalore, India*
Catching up with new detector-level upgrades for Run 3 and status updates on ongoing searches in Run 2 data.
- May 1–5 2023 **HSF-India Software Training Event**  *TIFR Mumbai, India*
Basics of C++, scientific python, GPU programming, machine learning using tensorflow and JAX.
- Jan 6–11 2022 **India-CMS collaboration meeting 2022**  *Online*
Presented my work on the DNN-based photon ID development, which is part of the service work for CMS.

SELECTED PUBLICATIONS

The CMS Collaboration, “*Review of searches for vector-like quarks, vector-like leptons, and heavy neutral leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV at the CMS experiment*,” Phys. Rept. **1115** (2025) 570-677.

The CMS Collaboration, “*Inclusive nonresonant multilepton probes of new phenomena at $\sqrt{s} = 13$ TeV*,” Phys. Rev. D **105**, 112007 (2022).

The CMS Collaboration, “*Electron and photon reconstruction and identification performance at CMS in 2022 and 2023*,” CMS Detector Performance Note CMS-DP-2024-052 (2024).

Full list of publications is available at <https://inspirehep.net/authors/2771643>

OUTREACH

- Jan 27 2025 **Talk: “Machine Learning meets Particle Physics”** *JIST, Assam, India*
Talked on how machine learning and particle physics come together to solve exciting problems, focusing on DNN classifiers and generative models.
- Feb 2024 **YouTube video on Women in Science** 
Edited an interview of some of the women working in the Experimental HEP group in IISER Pune under the guidance of Prof. Sourabh Dube.
- Oct 19 2022 **Introduction to EHEP for high school students**
An Assamese talk on introduction to the standard model and experimental particle physics, intended for 9th and 10th standard students from my high school.
- Dec 18 2021 **Talk: “A brief introduction to Experimental High Energy Physics”** *JIST, Assam, India*
Online talk on Experimental High Energy Physics for 3rd year B.Sc. students in Department of Physics, JIST.

ACHIEVEMENTS AND AWARDS

2025	Breakthrough Prize in Fundamental Physics 2025 (as a member of CMS)
2020	JEST-2020: Qualified with all India rank 631 INSPIRE-DST Fellowship
2019	CSIR-NTA NET, December 2019: Qualified lectureship (India) with rank 046 SLET, 2019 : Qualified lectureship (within N-E India) Gold Medal in M.Sc. Physics from Tezpur University, India

REFERENCES

Prof. Sourabh Dube

Department of Physics

IISER Pune, India

sdube@iiserpune.ac.in