Searching For Long-Lived Particles with HCAL Segmentation in CMS at the Large Hadron Collider

Katherine Avanesov[1,3], Harvey Newman[1,3], Kiley Kennedy[2,3], Gillian Kopp[2,3]

[1]The Division of Physics, Mathematics, and Astronomy, California Institute of Technology, [2]Department of Physics, Princeton University, [3]CMS

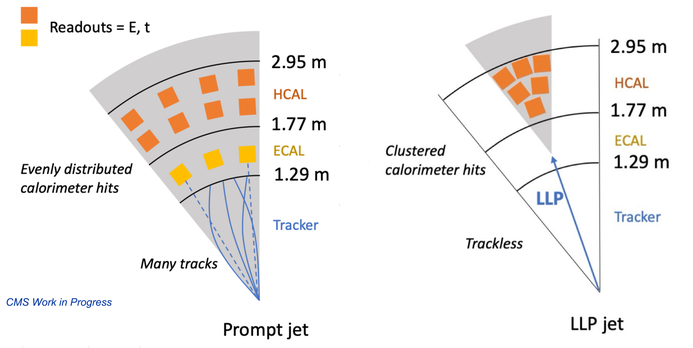
Introduction

Searches for Long-Lived Particles (LLPs) with lifetimes > 0.1 ns, motivated by both Standard Model and BSM theories (dark matter, matter–antimatter asymmetry, supersymmetry).

Focus on Higgs decays H→XX→~~b~~~~b~~~~b~~~~b~~, leveraging the new depth segmentation of the CMS Hadronic Calorimeter (HCAL) for identifying LLP decays within HCAL.

LLP Signatures

- Trackless due to decays in the HCAL
- Displaced vertex
- Narrower energy spread
- Longer lived particles will deposit less energy in the earlier layers of HCAL
- Delayed due to longer lifetime and path length

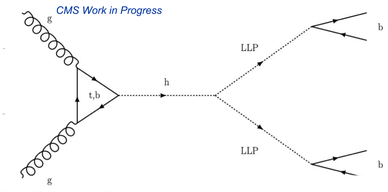


LLP Search Strategy

- Searching for H→ss→bbbb that decays within the HCAL (≈2-3m depth)

Online Selection

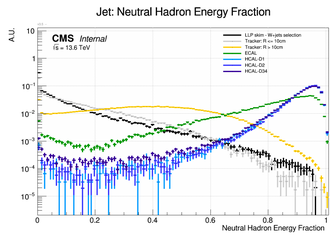
- Already developed L1 Trigger to select LLP-matching events (Higher energy fraction in HCAL segmentation, Time delay, Jet Energy threshold 40GeV)

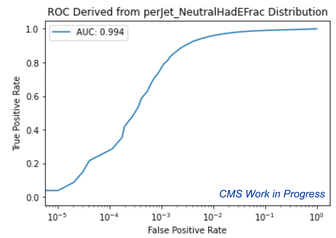


Offline Selection

- Developing a powerful LLP event tagger using jet metrics and 3D energy deposition patterns within the HCAL

Motivation and Data generation





LLP signatures differ significantly from prompt jets in several raw variables (e.g., jet energy depth fraction, neutral hadron fraction). Even a single variable can yield a high ROC AUC. However, with ~100 million events and <1% signal, we need high sensitivity at very low background efficiency, motivating the use of powerful classification techniques.

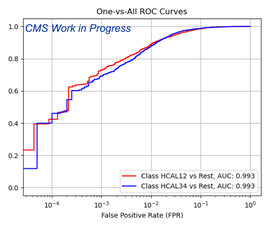
Signal

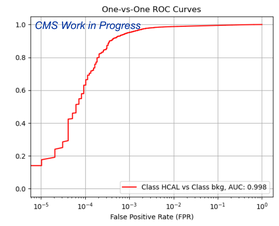
- MC-generated LLP events, based on different combinations of Higgs Mass and LLP Mass; Each file contains roughly 100k entries (reduces to 30-70k after pre-selections)

Background

- Run3 Skim for W+ Jets ≈ 80M events
- Training performed against W+ jets (other sources of background include QCD and Z+ jets)

Dense NN (Multi-)Classifier: MH350-MS160

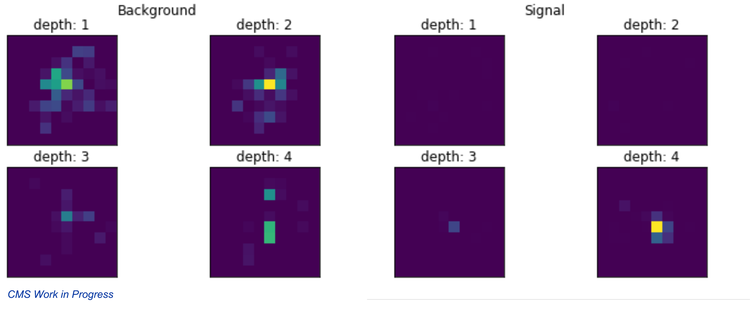


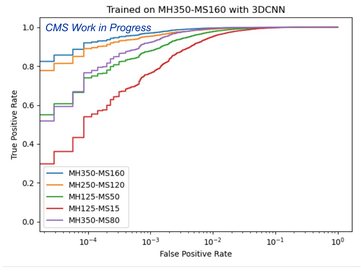


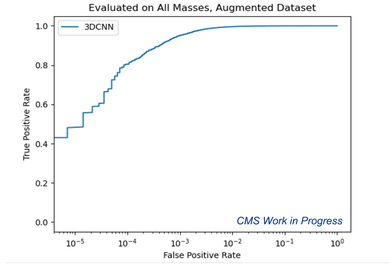
Rank	Variable Importance
1	perJet_NeutralHadEFrac
2	perJet_Track0Pt
3	perJet_EnergyFrac_Depth1
4	perJet_ChargedHadEFrac
5	perJet_Track1Pt
6	perJet_Track2Pt
7	perJet_S_phiphi
8	perJet_Track0dR
9	perJet_EnergyFrac_Depth2
10	perJet_S_etaphi

Training on 3D Jet Images

3D Image Sample from Signal and Background





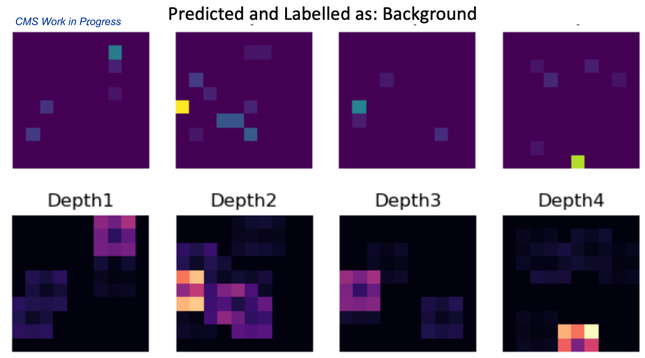


Explainability of 3DCNN: Gradient-Weighted Class Activation Mapping (GradCAM)

GradCAM extracts the learned features of a CNN by mapping the importance of each pixel based on its influence on the final output score.

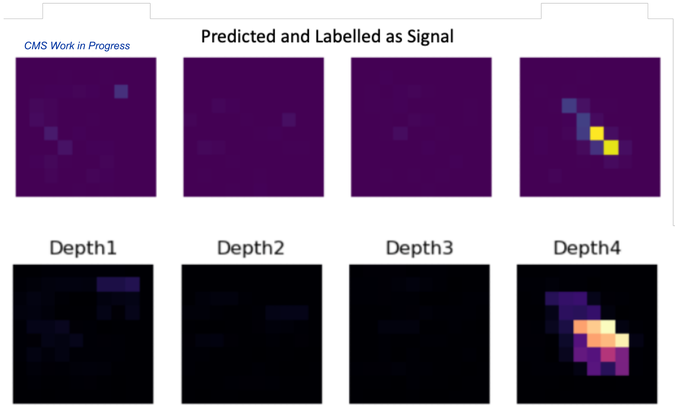
Background GradCAM:

- Model sees only individual, localized hot towers rather than groups of energy clusters
- Even when towers are close-by, an overlap of the 3x3 kernel shape is seen



Signal GradCAM:

- Activation pattern mimics the shape of the jet cluster
- Model perceives nearby hot towers as one energy cluster



Conclusion and Future Steps

- Promising exploration of machine learning-based methods for tagging delayed jets and identifying long-lived particles.
- Each model exhibits some degree of interpretability.
- Most notable achievement so far is a 90% signal efficiency at just 0.01% background efficiency (3D Jet Image tagger).