FIRE: THE FIRST-YEAR INNOVATION & RESEARCH EXPERIENCE

SIMULATING PARTICLE DETECTION

Effects of Pile-Up in Photon Reconstruction

Matt Bilger, Peter Kozlov, Will Scott, Arya Teymourlouei

Abstract

When CMS experiments first began, the idea of pile-up was not largely important. Now, due to higher luminosity of the collider, it has become more of an issue. Due to the high volume of the collisions, noise is created, and incorrect collisions are detected (pile-up). Pile-up effects the reconstruction and identification of particles, such as photons. Figure 1 displays the data events of interest in correlation to the beam pipe. All other surrounding hits are pile-up. It is clear that pile-up has a significant impact on the quality of the data. Our goal is to analyze the effects of pile-up by using the latest CMSSW data and using Python analysis code.

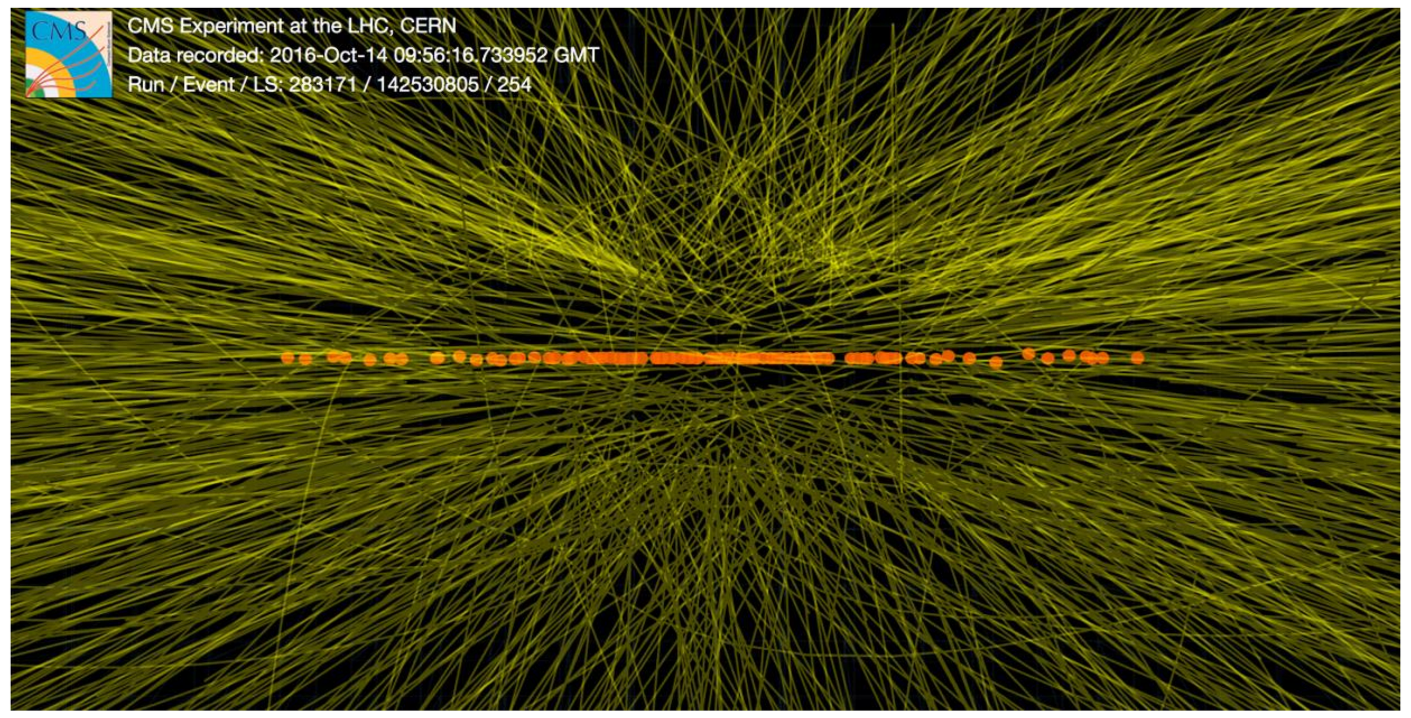

Fig. 1. Pile-up Visualization (Kahle, 2019)

Background

Beginning in 2027, the Large Hadron Collider will increase the luminosity of the collisions by a factor of 10, meaning there will be more proton-proton collisions being analyzed. When analyzing a collision event in the CMS, detectors often pickup information from other collisions, which is known as pileup. There are two types of pileup: in-time and out-of-time. The first (intime) comes from the same proton bunch crossing. You can determine the interaction points by identifying vertices, and comparing their distances to the Primary Vertex. The second occurs when other bunches cross and your detector has not yet recorded the signal completely (Mitrevski 2016). This causes algorithms to incorrectly identify proton-proton collisions which are important.

We collectively chose this topic due to our group interest in generated pile-up and noise within the detector environment. We also wanted a research topic that was more data analytical rather than computational, with more plotting and active analysis. We simulated the photon energies of a p-p collision and analyzed how pileup affects the reconstructed energies using a simple cone algorithm.

Results

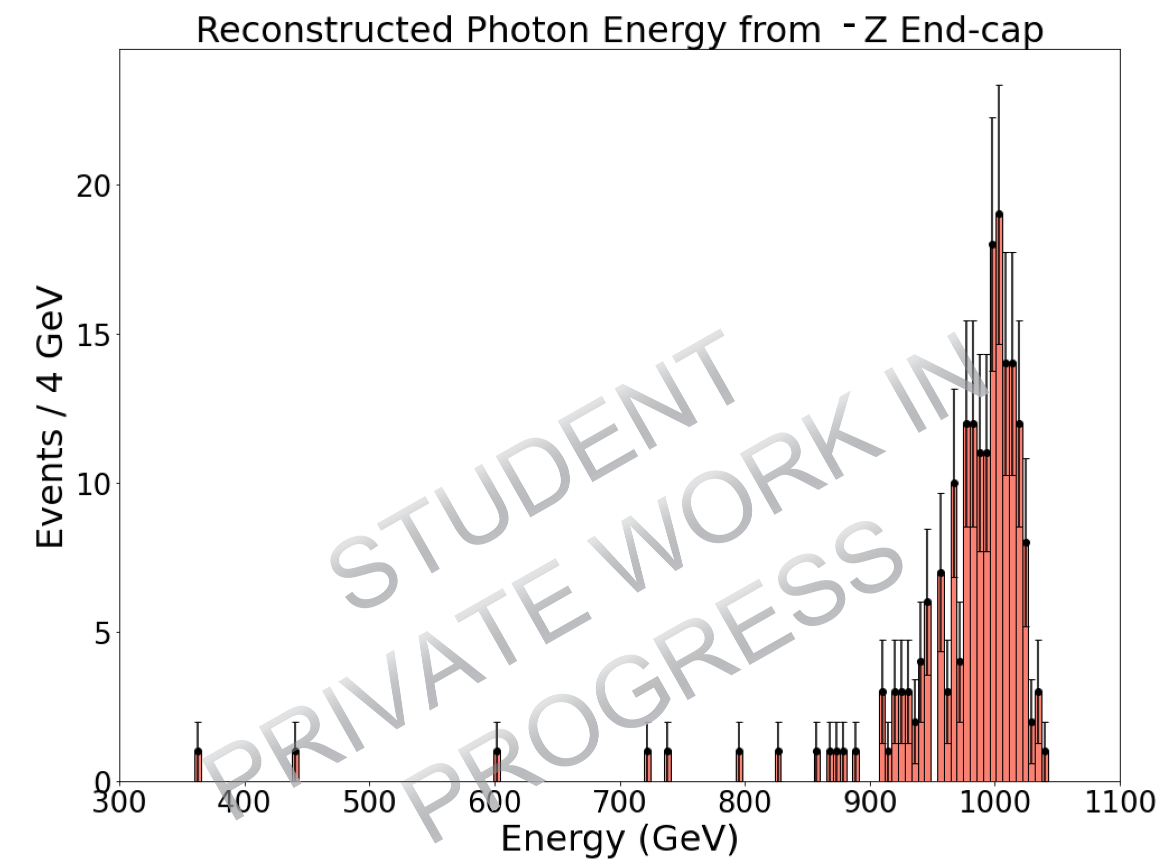

Fig. 2. This plot shows the reconstructed energy of photons using the -Z end-cap. The data for this plot was **400 events** of back-to-back photon collisions at 1 TeV (1000 GeV). This dataset was generated with the older framework (CMSSW 11_3_0). The graph was generated using an algorithm that looped through each event. The reconstruction of photon energies took approximately 2 hours, so each event took around 30 seconds to process. The data is analyzed using a Gaussian distribution. Note there is no pile-up in this plot, or in figure 3.

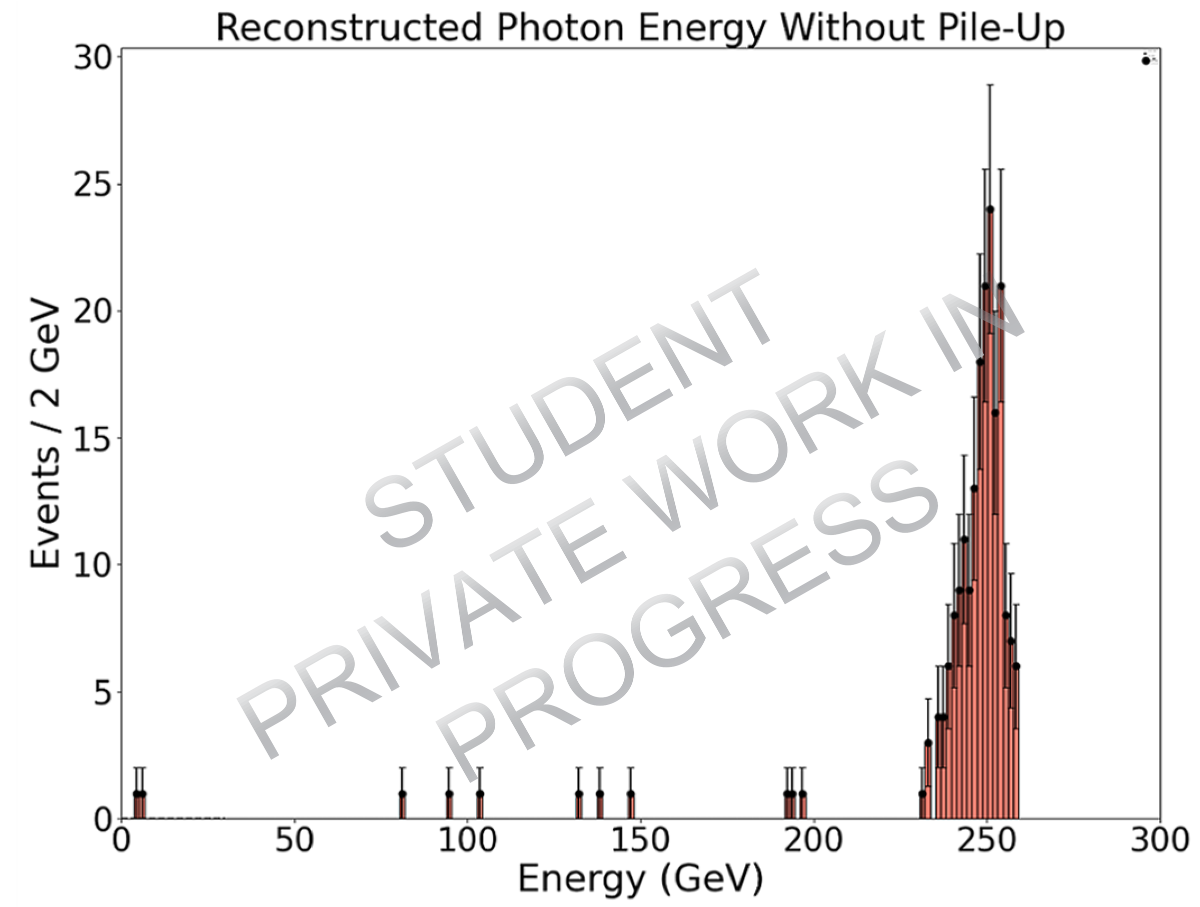

Fig. 3. This graph plots the reconstructed energy of photons using the new CMSSW framework (12_3_0). In this plot, both end-caps have 200 events each, totaling 400 events. The energy level is 250 GeV. The data file was generated by combining the events from the negative and positive end-caps. Ideally, this plot and the one above should be fairly similar. Both distributions look in agreement with the generated input levels, with a Gaussian smearing around the peak. Note the time for 400 events took around 40 minutes to analyze.

Acknowledgements

We are very grateful and thank Dr. Muge Karagoz, our faculty leader, for continuous support and guidance. We also want to thank Fred Garcia for the Python analysis. We thank CMS for providing the dataset framework (CMSSW). We also want to thank our PRM, Jacques Deroin.

References

- Brondolin, E. (2020). Silicon sensors for the CMS HGCAL upgrade: challenges, sensor design & electrical characterization. *Journal of Instrumentation*, 15(05), C05068.
- Kahle, Kate. (2019). "From capturing collisions to avoiding them". *CERN*, https://home.cern/news/news/knowledge-sharing/capturing-collisions-avoiding-them. Accessed 17 April 2022.
- Maier, B., Narayanan, S. M., de Castro, G., Goncharov, M., Paus, C., & Schott, M. (2021). Pile-Up Mitigation using Attention. *arXiv preprint arXiv:2107.02779*.
- Mitrevski, J., & Atlas Collaboration. (2016). Electron and photon reconstruction with the atlas detector. *Nuclear and Particle Physics Proceedings*, 273, 2539-2541.

Methods

The Python analysis code is a particle gun simulation (MC). We use the newest CMS Software (CMSSW) to generate datasets for the analysis. We use a dataset of photons. The main Python script was previously in a Jupyter notebook (written by Fred Garcia), but we re-configured it to two separate scripts. The first script generates an output file of the events. The second script uses the generated output file to plot the events of the reconstructed photon energies. The main Python function that generates the reconstructed hit energies is a for loop that traverses each event.

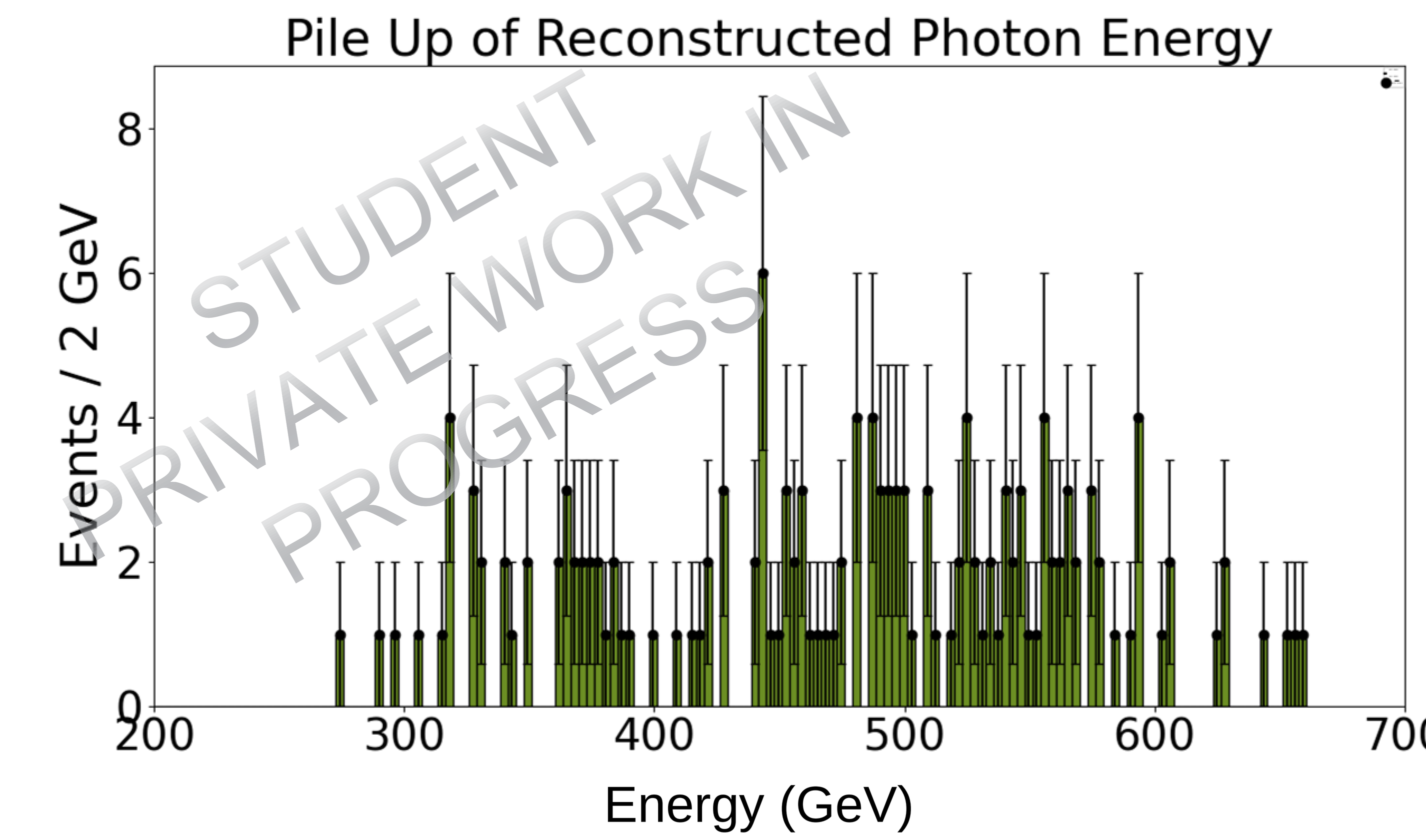

Fig. 4. This plot examines the reconstructed photon energies at 250 GeV with pile-up. There are 160 events (80 eve, and were generated with the new framework (CMSSW 12_3_0). Note the pile-up effect increases the energy scale and range. As a result, the reconstructed hit energies are from the intended 250 GeV. All events took roughly 24 minutes to generate.

Conclusions & Future Work

The generated results show the summed reconstructed hit energy of photons in a simulation. The performance of the analysis of the old data was two hours for 400 events. The new data took around 40 minutes for 400 events. Future work include first further validating the effects of pile-up and then mitigating these effects. Developments in machine learning algorithms are a possible way to mitigate pile-up. For example, one recent algorithm used an attention mechanism, which focuses on the relevant particles in a dataset (Maier et al., 2021). The authors also incorporated transformers and clustering, two other concepts in machine learning. Their algorithm, called PUMA, was successful in mitigating pile-up of jet events. Our group would like to examine the results of the PUMA algorithm on photon reconstruction events. Machine learning can be a useful tool to mitigate the effects of pile-up.