Simplified Monte Carlo based HL-LHC Calorimeter with Multi-Event Validation

Project Report

January 2026

https://github.com/chanduyasodar/monte_carlo_calorimeter_V1

Abstract

This report details the Monte Carlo based simulation of a simplified calorimeter modeled as 100x100 grid. Gaussian, Poisson and Exponential distributions were incorporated into modeling the noise, pileup and signal shower components of energy deposition. Stochastic equations were used to describe the time deviation and energy correlations. Energy–time correlations and timing performance under High-Luminosity LHC (HL-LHC) pile-up conditions has been established. The study focuses on per-cell energy deposition, stochastic timing resolution, and the impact of pile-up on timing observables. A multi-event statistical analysis is performed to validate the mathematical consistency of the model prior to introducing trigger logic. We report the current results,critically assess their limitations, and outline corrections planned for the next version.

Contents

List of Figures

1 Introduction

A calorimeter is a destructive detector used to measure the kinetic energy of the particles from the collision. As the particle enters the calorimeter, it interacts with the dense matter inside and initiates a shower. This shower is converted into recordable signal.

In this project, the simulation is performed by modeling the calorimeter into a simple grid. The calorimeter detects two things, energy deposition and timing measurement. Each of these has 3 components noise, pileup and the actual signal which we are interested to record and study.

In the hadron collider, protons collide in bunch, as a result, not all of them undergo a perfect head on, high energy collision. Most of the collisions observed are not very interesting, and these do release energy which is referred to as "Pileup". These disturb the data obtained. Noise is just electronic jitter. So, in order to not to let both of these interfere with the actual signal, time measurement is brought in, and we shall see how they are used in filtering the data effectively.

2 Theoretical Framework and Statistical Modeling

The Calorimeter is structured as a two dimensional grid of cells, each of them recording energy deposition and timing measurement for each event. Each event represents one proton-proton collision.

2.1 Energy Deposition and Measurement.

For each event, the energy of each cell(i, j) of calorimeter is calculated by:

$$E_{ij}^{total} = E_{ij}^{signal} + E_{ij}^{PU} + E_{ij}^{noise} \tag{1}$$

2.1.1 Noise

Modeled via a **Gaussian Distribution**

$$E^{\text{noise}} \sim \mathcal{N}(\mu, \sigma) \qquad \mu = 0, \sigma \approx 0.5$$

Noise arises from many, small independent fluctuations in voltage and current. According to the *Central Limit Theorem*, the aggregate of these fluctuations follows a Gaussian distribution.

2.1.2 Pileup

Pileup is the unwanted energy deposition due to low energy collisions

- **Energy distribution**:. Modeled via **Uniform Distribution**, and is added to random cells in a loop repeating N^{PU} times.

  $$E^{PU} \sim Uniform[0, 5] \quad looping\ N^{PU}\ times$$

  .

- **Number of Interactions**: n^{PU} is defined as the no.of interactions per bunch crossing, formulated by

  $$n^{PU} = \frac{L_{inst} \cdot \sigma_{inel}}{f_{rev} \cdot n_b} \tag{2}$$

  Where:

  $L_{inst} = 5.0 \times 10^{34}$ cm^{-2}s$^{-1}\sim$ Instantaneous Luminosity (the "brightness" of the beam).

$\sigma_{inel} \approx 80$ mb $= 80 \times 10^{-27}$ cm^2 $\sim$ Inelastic proton-proton cross-section (the "target size" of a proton).

$f_{rev} \sim$ Revolution frequency of the LHC

$n_b \sim$ Number of bunches in the ring.

$f_{rev}.n_b = 140$ MHz

Substituting all the values, we get $n^{PU} \approx 100$. Due to the bunches not being perfectly uniform, a fill factor is brought in the formula, increasing $n^{PU} \approx 140, 200$. Any of these values can be used.

Every bunch crossing doesn't guarantee $n^{PU} \approx 140, 200$ collisions. So, the actual number of collisions,N^{PU},is modeled via **Poisson Distribution** for each event.

$$N^{PU} \sim Poisson(n^{PU})$$

2.1.3 Signal

Single signal system is simulated in this project in order to avoid the complexity.

- **Event Occurrence**: A **Binomial Distribution** determines if a signal exists in the current frame
$$\mathcal{P}(Event) \sim Binomial(n = 1, p = 0.01)$$
- **Signal Type**: A second **Binomial Trial** determines whether the signal is of the Electromagnetic or Hadronic shower
$$\mathcal{P}(Event_{EM}) \sim Binomial(n = 1, p = 0.3) \quad else, Signal = Event_{Hd}$$
- **Energy Spectrum**:
  - **Electromagnetic Jet**: Modeled via **Exponential Distribution**, ensuring low frequency of high energy signal
$$A_{EM}^{signal} \sim Expoential(Amplitude = 30GeV)$$
  - **Hadron Shower**: Modeled via **Gaussian Distribution** as a hadron jet consists of many hadrons and their combination of their energies results in energy measurement fluctuating around a central mean symmetrically.
$$A_{Hd}^{signal} \sim \mathcal{N}(\mu = 50, \sigma = 5) \quad \text{where, } \mu\text{=Amplitude=50 GeV}$$
- **Spatial Profile**: Random coordinate is chosen on the grid and a **2D-Gaussian packet** is applied around that coordinate.
$$Shape \sim e^{-d^2/2\sigma_{signal}}$$
, where $\sigma_{EM} = 3$ (narrow and compact)and $\sigma_{Hd} = 9$(broad).
- **Signal Profile**: Signals observed in colliders are not perfect Gaussian shower. A realistic signal peak is modeled via **Poisson Distribution** of product of the Energy spectrum and shower shape
$$E^{signal} \sim Poisson(A^{signal} * Shape)$$

2.1.4 Threshold

$$E_{min} = n * \sigma_{noise} \tag{3}$$

where n is chosen such that the detector ignores noise energy and notices signal energy.

$$E_{ij}^{total} = E_{ij}^{signal} + E_{ij}^{PU} + E_{ij}^{noise}$$

2.2 Timing Measurement

The final cellular time is formulated as a sum of t_0 and **Gaussian Distribution** of mean time and time resolution. Here,we assign $t_0 \approx 0$ implying that t_{ij} of signal cells is 0.

$$t_{ij} = t_0 + \mathcal{N}(\mu_{ij}^t, \sigma^t(E_{ij})) \tag{4}$$

2.2.1 Time Resolution

Timing uncertainty is correlated to Energy by relation

$$\sigma_{ij}^t = \sqrt{\left(\frac{a}{\sqrt{E_{ij} + E_{min}}}\right)^2 + b^2} \tag{5}$$

where a=75 ps (Stochastic Term) accounting for sampling fluctuation and b=25 ps (Constant term) accounting for electronic jitter. These values are chosen to reflect the target performance of the CMS High Granularity Calorimeter under HL-LHC conditions.E_{min} was added to the stochastic term for computation purposes

2.2.2 Pileup Time

The window in which the N^{PU} collisions occur is calculated by

$$\Delta t = \frac{\text{Lenght of Proton bunch} \approx 7.5 \; to \; 10 \; cm}{\text{Speed of light}} \approx 180 \; to \; 200 \; ps$$

Pileup time is modeled via **Gaussian Distribution** as per *Central Limit Theorem* and the bunch shape, which is also Gaussian.

$$t^{PU} \sim \mathcal{N}(\mu = 0, \sigma = \Delta t)$$

2.2.3 Mean Time

We assign $t^{signal} \approx 0$ for high energy collisions.

$$\mu_{ij}^t = \frac{(E_{ij}^{signal} * t^{signal}) + (E_{ij}^{PU} * t_{ij}^{PU})}{E_{ij}^{total} + E_{min}} \tag{6}$$

$$t_{ij} = t_0 + \mathcal{N}(\mu_{ij}^t, \sigma^t(E_{ij}))$$

3 Signal Event Analysis

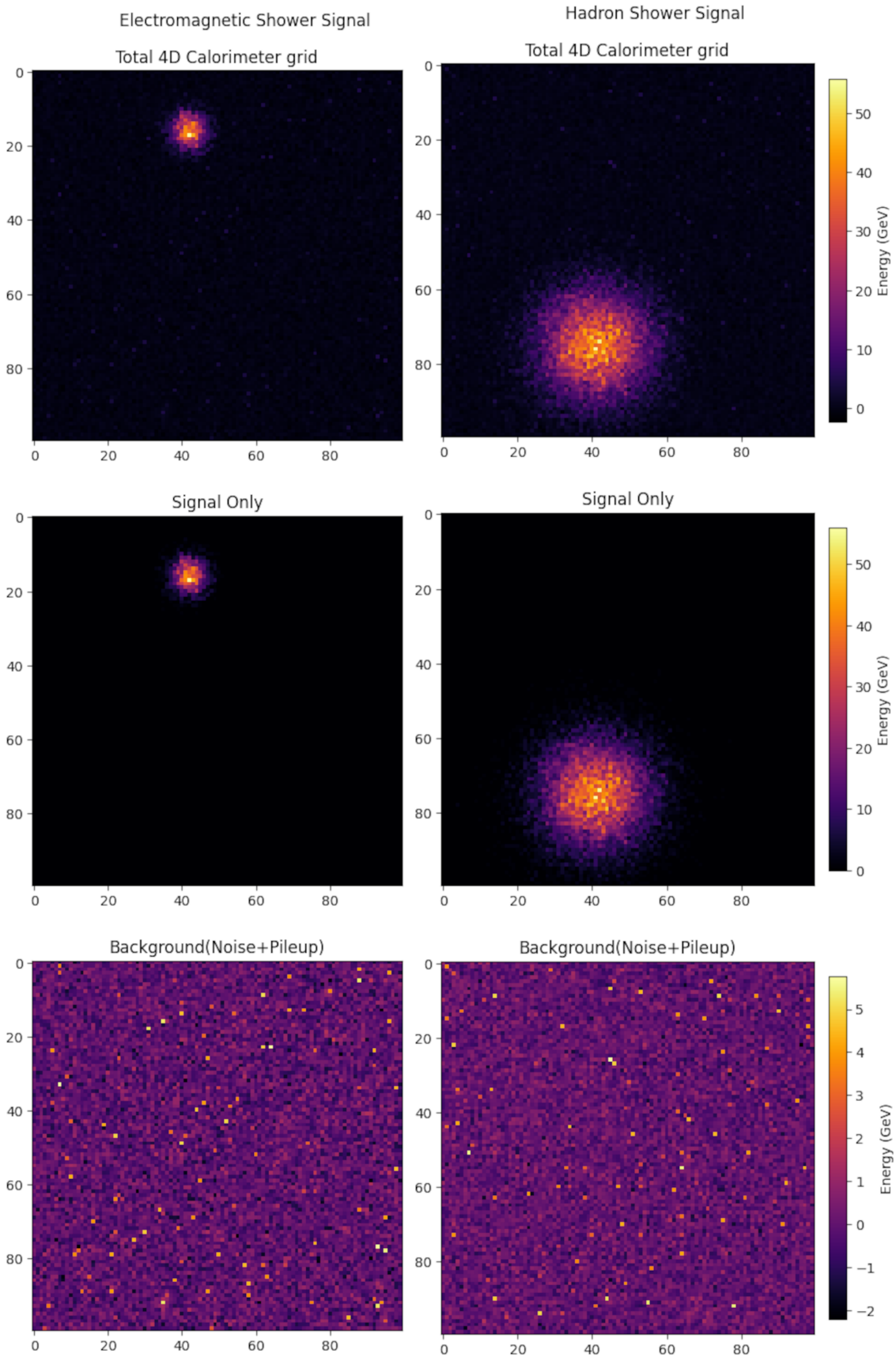

Figure 1: Heat Map

3.1 Spatial Observation: Heat Maps

3.1.1 Observation

- **Electron Shower**: Heatmap displays a highly concentrated longitudinal energy profile. The lateral spread (Molière radius) is narrow and well-defined.

  This behavior closely matches experimental EM showers observed in ATLAS or CMS ECAL calorimeters, where electrons and photons produce narrow showers governed by the Molière radius. The compact spatial extent and sharp core are realistic.

- **Hadron Shower**: Compared to the electromagnetic signal, this shower is significantly deeper and more diffuse. Energy deposition is spread across more layers and shows several "sub-clusters," representing the stochastic nature of hadronic interactions and nuclear fragmentation.

  This agrees well with experimental hadronic showers, which are wider due to nuclear interactions, invisible energy losses, and large event-by-event fluctuations.

3.1.2 Limitations and Planned Model Upgrades

- **Over-symmetric shower shapes $\longrightarrow$ Incident-angle and geometry modeling**: The near-radial symmetry of simulated showers will be addressed by introducing non-normal incidence angles and realistic detector granularity to reproduce shower elongation observed in real calorimeters.

- **Lack of longitudinal structure $\longrightarrow$ Multi-layer calorimeter simulation**: The absence of depth-dependent shower development will be resolved by extending the model to multiple longitudinal layers, allowing electromagnetic and hadronic showers to peak at different depths as in real detectors.

- **Simplified hadronic fluctuations $\longrightarrow$ Stochastic sub-shower modeling**: The overly smooth hadronic shower structure will be improved by incorporating stochastic secondary interactions and non-Gaussian energy fluctuations to generate realistic cell-to-cell variations and hot spots.

- **Uniform background distribution $\longrightarrow$ Correlated pile-up and detector non-uniformities**: The spatial uniformity of the background will be refined by modeling time- and space-correlated pile-up vertices and detector-specific noise variations.

- **Overly clean dynamic range separation $\longrightarrow$ Realistic thresholds and zero-suppression**: The clear separation between signal and background will be made more realistic by introducing readout thresholds, zero-suppression, and digitization effects that increase overlap between low-energy signal and background.

3.2 Signal vs Background Energy Distribution

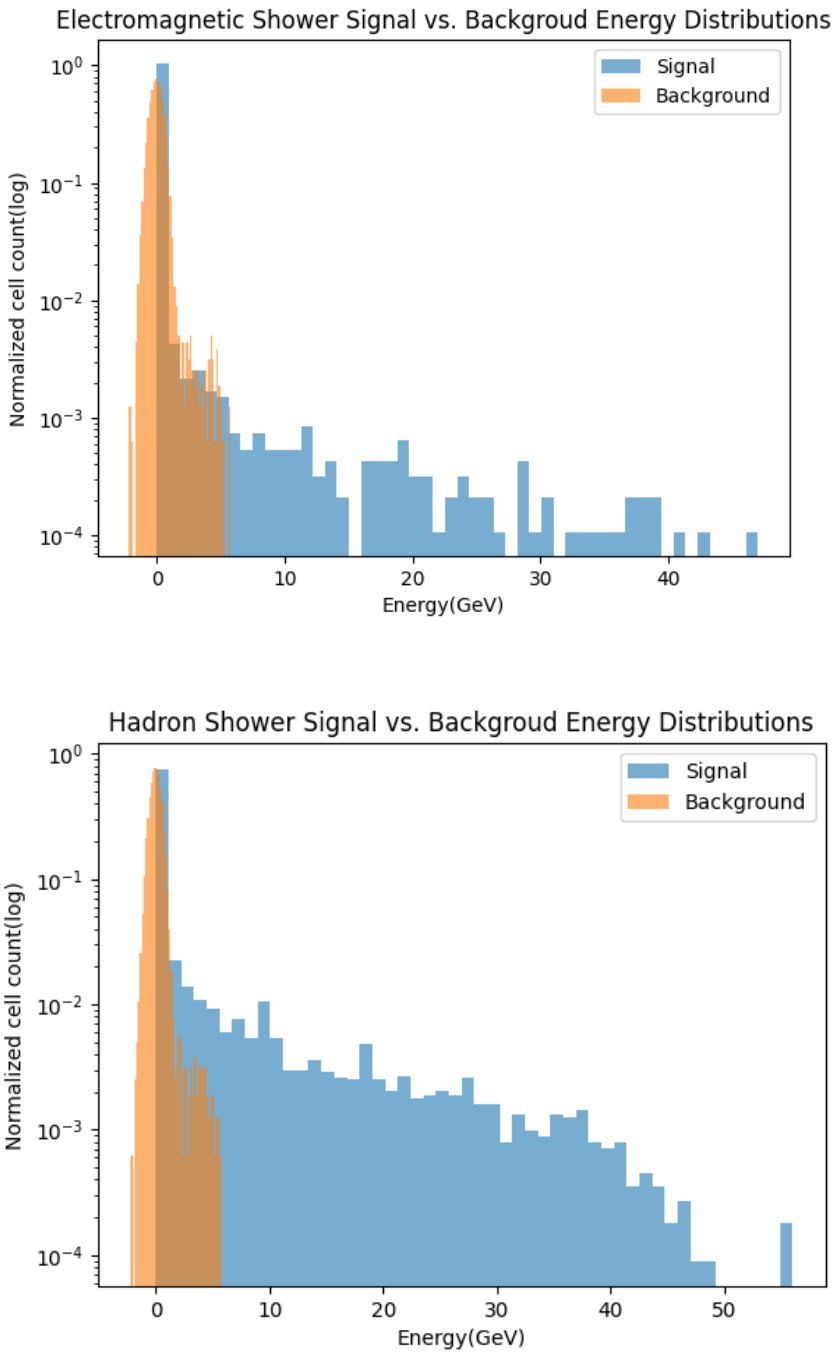

Figure 2: Signal vs Background Histogram

3.2.1 Oberservation

- **Electromagnetic Showers**:

  For electromagnetic showers, the background energy distribution is sharply peaked at low energies and decreases rapidly, consistent with expectations from electronic noise and low-energy pile-up. The signal distribution exhibits a steeply falling spectrum with a pronounced high-energy tail extending to approximately 40–45 GeV. At moderate and high energies, the signal dominates the background by several orders of magnitude.

The separation between signal and background improves rapidly with increasing energy, indicating that energy-based selections are effective for electromagnetic objects at the cell level. The residual overlap at low energies is attributed to peripheral shower deposits and noise contributions, as commonly observed in experimental electromagnetic calorimeter data.

- **Hadronic Showers**:

  For hadronic showers, the signal energy distribution is broader and extends to higher energies than in the electromagnetic case, with a tail reaching beyond 50 GeV. The background distribution remains concentrated at low energies and is similar in shape to that observed in the electromagnetic case.

  Compared to electromagnetic showers, the separation between signal and background emerges at higher energies. The increased overlap at low and intermediate energies reflects the wider spatial extent and larger intrinsic fluctuations of hadronic showers, consistent with the stochastic nature of hadronic interactions in calorimeters.

3.2.2 Comparison Between Electromagnetic and Hadronic Signals

The electromagnetic and hadronic signal distributions exhibit qualitatively different behaviors. Electromagnetic showers produce a narrower energy spectrum with a more rapid fall-off, while hadronic showers show a broader distribution and a longer high-energy tail. These differences are consistent with established calorimeter performance and demonstrate that the simulation reproduces the expected contrast between electromagnetic and hadronic energy deposition.

3.2.3 Discussion

The logarithmic representation highlights the rarity of large background fluctuations and the significant dynamic range over which signal dominates background. While energy alone provides strong discrimination for electromagnetic showers, the reduced separation observed for hadronic showers at low energies motivates the inclusion of additional observables, such as timing or shower topology, for improved background rejection.

Overall, the observed energy distributions are in qualitative agreement with experimental calorimeter measurements and provide validation of the signal and background modeling used in this study.

4 Noise Analysis

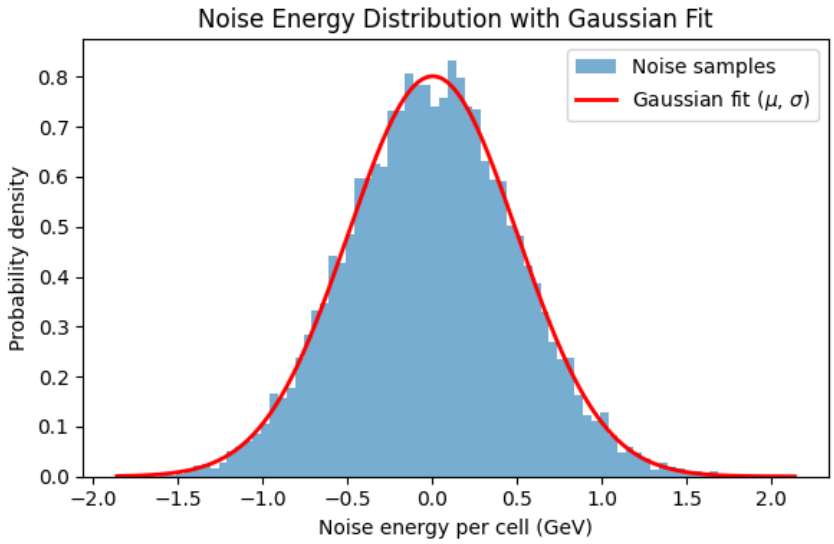

Figure 3: Caption

- **Observation**: The noise energy distribution is symmetric about zero and well described by a single-peaked, bell-shaped curve extending approximately from -2 to $+2$ GeV, with the fitted Gaussian closely following the histogram across the full range.
- **Statistical interpretation**: The agreement between the histogram and the Gaussian fit indicates that the noise fluctuations are consistent with a normal distribution characterized by mean $\mu \approx 0$ and finite standard deviation, as expected from the superposition of many independent electronic noise sources under the *Central Limit Theorem*.
- **Limitation**: The visual agreement alone does not quantify goodness-of-fit, and potential non-Gaussian tails or small biases in the mean cannot be excluded without formal statistical tests or higher-statistics samples.
- **Planned model upgrade**: Perform quantitative goodness-of-fit tests (e.g., χ^2, Kolmogorov–Smirnov) and study noise stability across detector regions and event samples to assess deviations from Gaussian behavior and potential systematic effects.

5 Multi Event Observation and Analysis

The event generation loop was run for 300 times,whose data was recorded and the following graphs were plotted for model validation.

5.1 Energy Distribution

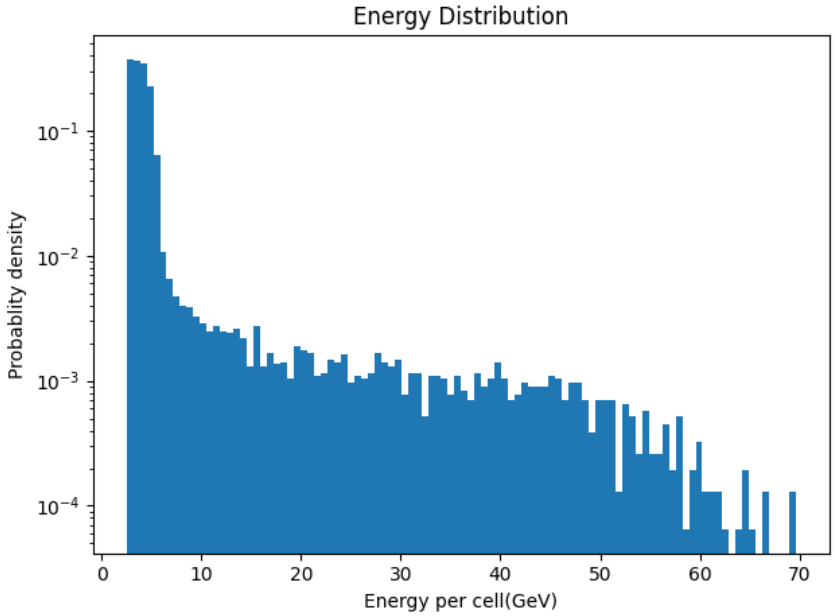

Figure 4: Energy Distribution Histogram

- **Observation**: The per-cell energy distribution is strongly right-skewed, with a pronounced low-energy peak and a long tail extending to high energies, and shows a smooth, monotonic decrease on a logarithmic probability scale.
- **Statistical interpretation**: The low-energy peak is consistent with noise-dominated and peripheral shower cells, while the heavy tail reflects rare, high-energy deposits from shower cores; the absence of secondary modes suggests a single dominant deposition mechanism.
- **Limitation**: The overlap between noise and low-energy signal is not explicitly separated, and no formal goodness-of-fit to an expected functional form (e.g., Gamma/Landau-like) is provided.
- **Planned model upgrade**: Decompose the distribution into noise-only and signal components and perform quantitative fits with statistical tests to assess agreement and overlap.

5.2 Time Distribution

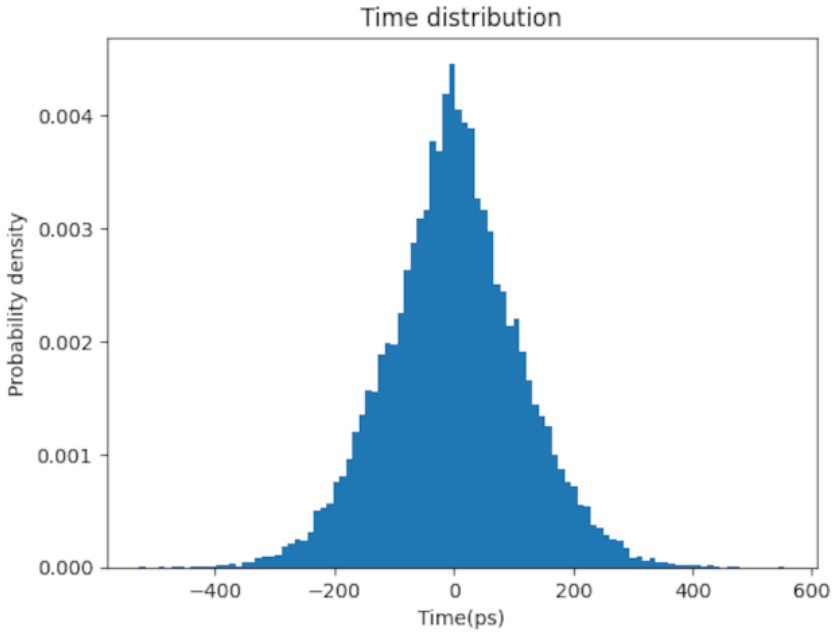

Figure 5: Time Distribution Histogram

- **Observation**: The reconstructed time distribution is approximately Gaussian, centered near zero, with symmetric tails extending to several hundred picoseconds.
- **Statistical interpretation**: The symmetry and unimodality are consistent with Gaussian timing jitter, while the extended tails arise from contributions of low-energy cells with poor timing resolution.
- **Limitation**: Aggregating all energies inflates the overall RMS and obscures the intrinsic timing performance at higher energies.
- **Planned model upgrade**: Produce energy-sliced time distributions and report robust spread measures or truncated RMS values.

5.3 Energy vs Time Scatter Plot

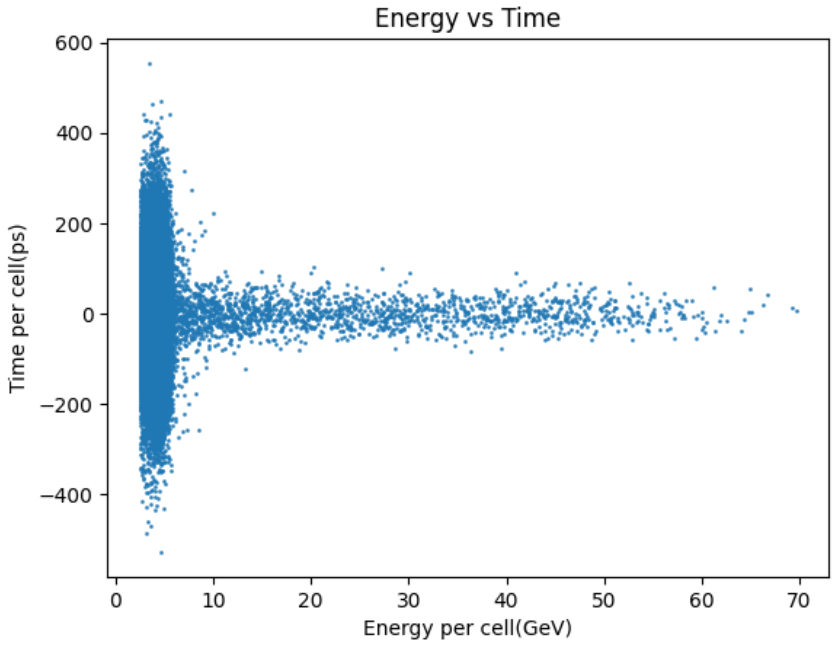

Figure 6: Energy vs Time Scatter Plot

- **Observation**: A funnel-shaped distribution is observed, with large timing spread at low energies and a rapidly narrowing spread as energy increases, while the mean time remains centered near zero.
- **Statistical interpretation**: The decreasing variance with energy indicates heteroscedastic behavior consistent with an energy-dependent timing resolution, and the symmetry around zero suggests no systematic timing bias.
- **Limitation**: Low-energy points dominate the variance and include extreme outliers, indicating that timing information is statistically ill-defined below a certain energy threshold.
- **Planned model upgrade**: Apply minimum energy cuts and quantify timing variance in energy bins to isolate physically meaningful regimes.

5.4 Time Resolution: Simulation vs Design

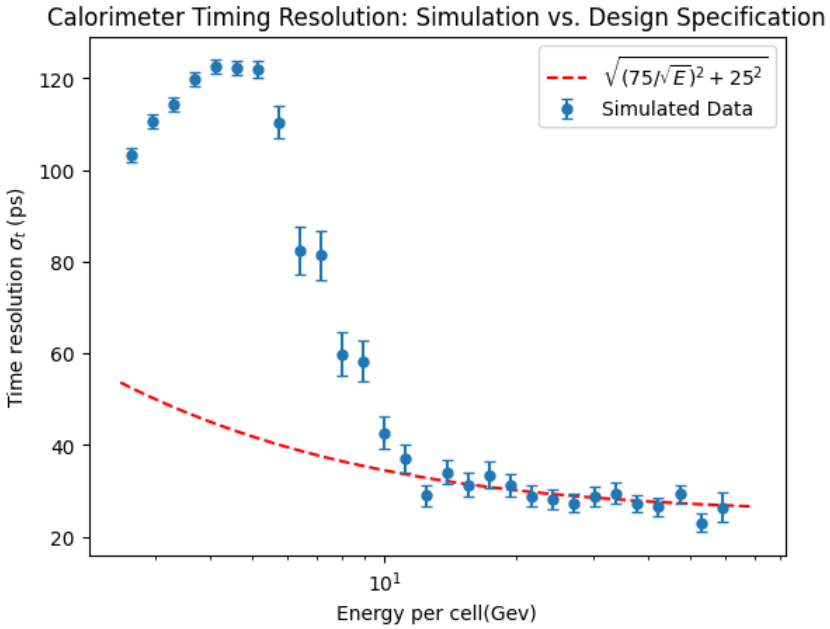

Figure 7: σ^t : Simulation vs Theory

- **Observation**: The simulated timing resolution decreases with increasing energy and approaches the design curve at moderate to high energies, while showing significant degradation at low energies.
- **Statistical interpretation** Agreement at high energies supports the asymptotic behavior of the timing model, whereas low-energy deviations indicate additional variance sources not captured by the nominal parameterization.
- **Limitation**: No quantitative fit quality is reported, and the design curve appears optimistic in the low-energy regime.
- **Planned model upgrade**: Fit the simulated points to extract effective model parameters and include additional low-energy noise terms or thresholds.

5.5 Extracted Time Resolution vs Energy

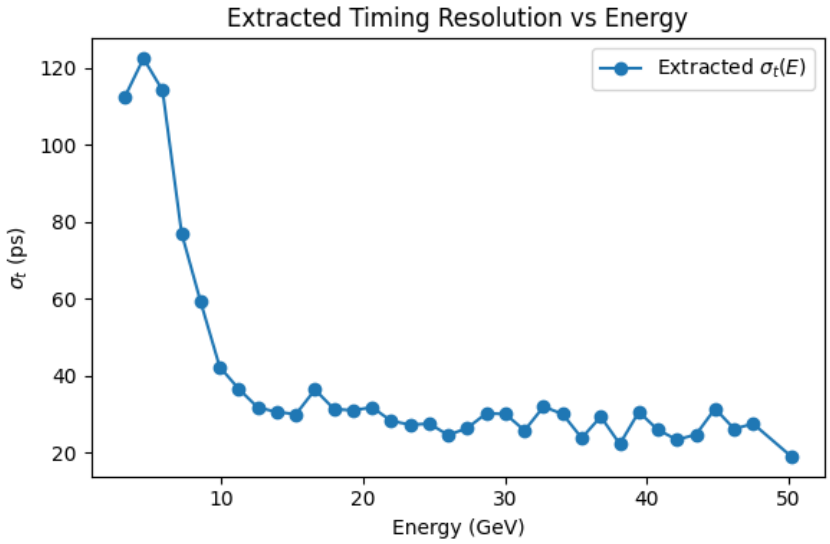

Figure 8: σ^t vs Energy

- **Observation**: The extracted $\sigma^t(E)$ decreases rapidly at low energies and flattens at higher energies, approaching a constant timing floor, with modest bin-to-bin fluctuations.
- **Statistical interpretation**: The trend is consistent with a stochastic term dominating at low energies and a constant term at high energies, while residual scatter suggests limited statistics or residual correlations.
- **Limitation**: Low-energy bins are unstable and dominate the dynamic range, and uncertainties on $\sigma^t(E)$ are not explicitly shown.
- **Planned model upgrade**: Increase statistics per bin, include uncertainty bands, and restrict extraction to energy regions where timing is well defined.

6 Conclusion

This project presents a statistically grounded Monte Carlo framework for studying energy–time correlations in a segmented calorimeter, incorporating electromagnetic and hadronic showers, Gaussian electronic noise, pile-up, and energy-dependent timing resolution. The statistical analyses of energy spectra, timing distributions, and energy–time correlations demonstrate internal consistency and qualitative agreement with established calorimeter behavior, particularly in the trigger-relevant moderate-to-high energy regime. The noise model is validated through Gaussian fits, and the extracted timing resolution follows the expected functional dependence on energy, with deviations at low energies correctly identified as noise-dominated and statistically non-informative.

Overall, the project achieves a decent validation of the core detector physics assumptions while transparently identifying limitations arising from controlled simplifications, establishing a reliable foundation for further development.

7 Future Scope

- Implement FPGA-friendly algorithms using fixed-point arithmetic, pipelined processing, and latency-aware designs suitable for real-time trigger applications.
- Develop local-maxima–based and sliding-window clustering algorithms for robust signal identification in high-pile-up environments.
- Design and integrate a full trigger mechanism incorporating energy- and timing-based selections, pile-up suppression, and real-time decision logic.
- Extend the detector model to include longitudinal segmentation, non-uniform and correlated noise, pile-up vertex structure, calibration uncertainties, and digitization effects.
- Incorporate advanced AI/ML techniques for pattern recognition, anomaly detection, and adaptive thresholding, and benchmark their performance against classical trigger algorithms.
- Evolve the framework into a realistic prototype for next-generation calorimeter trigger studies relevant to the HL-LHC.