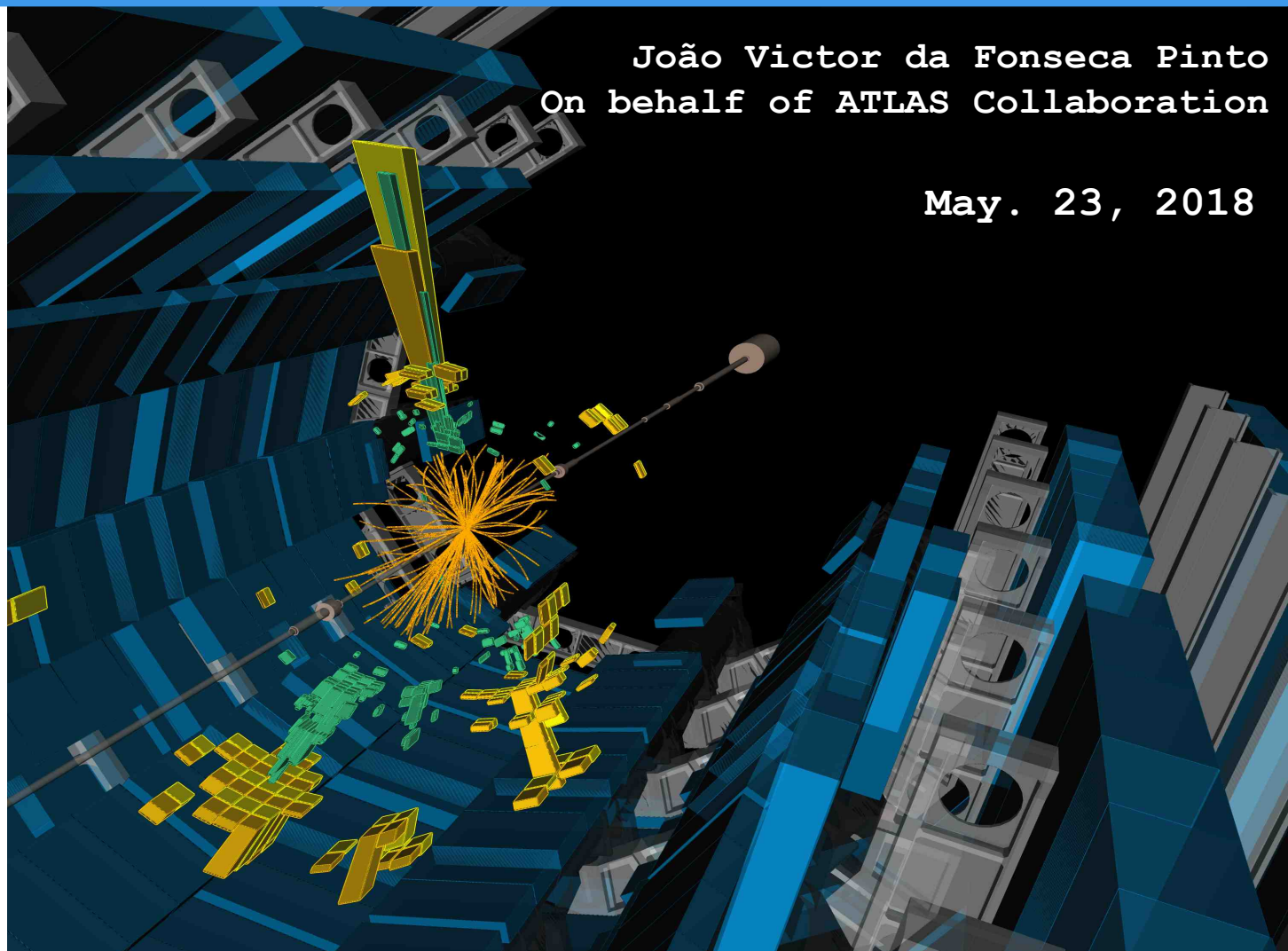
# An Ensemble of Neural Networks for Online Electron Filtering at the ATLAS Experiment

## CALOR 2018, Eugene OR

João Victor da Fonseca Pinto
On behalf of ATLAS Collaboration

May. 23, 2018



COPPE/Poli/UFRJ
Laboratório de Processamento de Sinais
Inteligência Computacional. Inovação

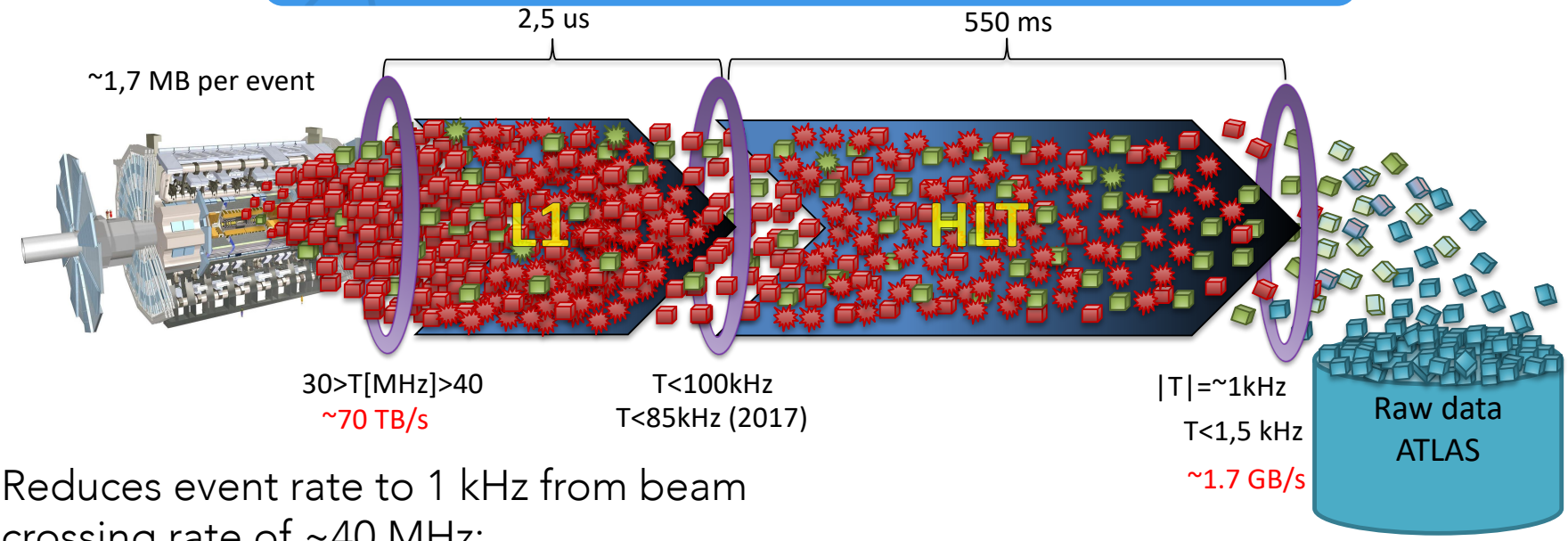
ATLAS EXPERIMENT

# Outline

- Basics:
  - ATLAS Trigger System;
    - HLT Trigger Optimization;
  - Ring-shaped Calorimetry Extraction;

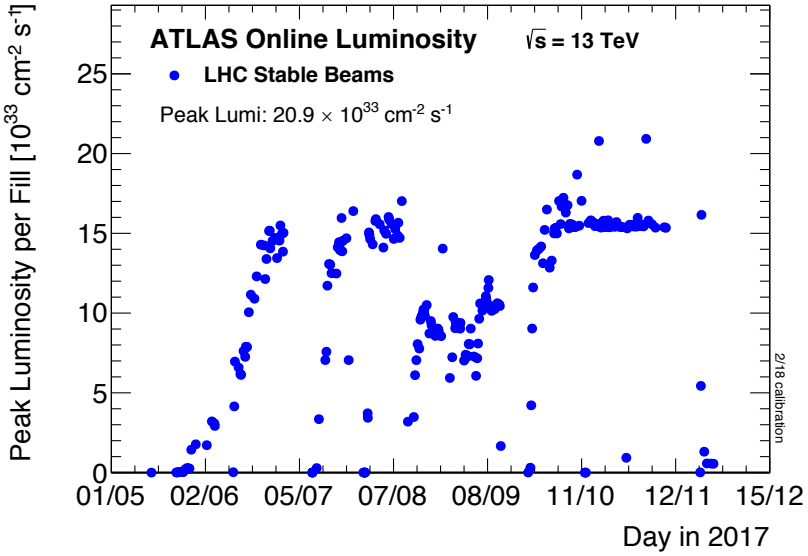
- Neural Ringer Operation in 2017;
  - Trigger Efficiency after switching to ringer;
- Impact studies;
- Conclusion.

# ATLAS Trigger System



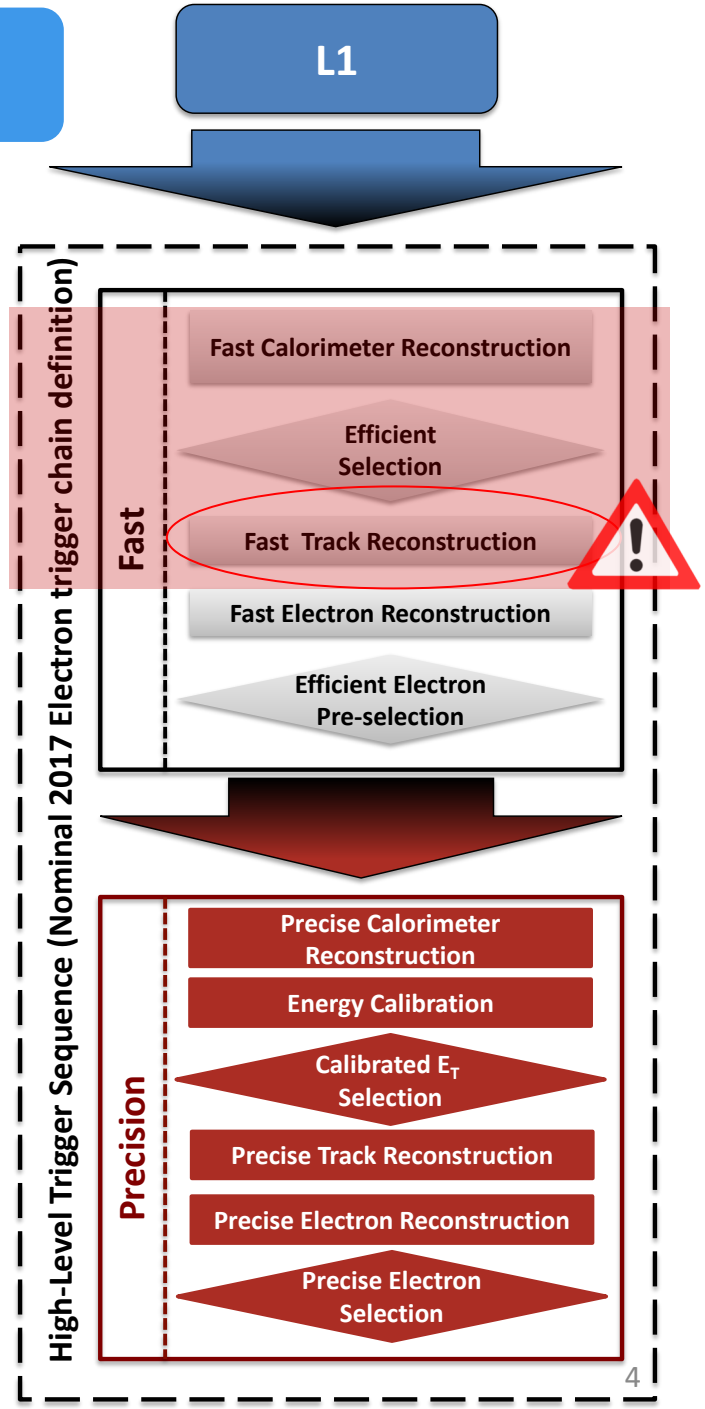

- Reduces event rate to 1 kHz from beam crossing rate of ~40 MHz;
- Around 20% allocated to e/**γ** ;
- In Run 2 the peak lumi is 2X larger than Run 1;

- The trigger system was designed to record only ~1KHz;
  - Need to keep the rates under control in high luminosity scenario;
  - Upgrades were implemented during Run 2



# HLT Trigger Optimization

- Fast Calo Intervention:
  - Use a new event **calorimetry** description (concentric rings);
  - This information will be used to fed a multivariate discriminator;
    - An ensemble of neural networks;
  - High rejection power when compared to the old paradigm (cut-based selection);
    - Fake rate reduction **before the track reconstruction**;
  - Pileup correction to keep up the efficiency;
  - Only when triggering electrons **above 15 GeV**.



# Electron Identification (Fast step)

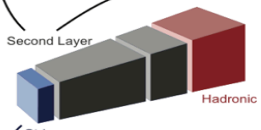
Old

## Variables and Position

| | Strips | 2nd | Had. |
|---|---|---|---|
| Ratios | $f_1$, $f_{side}$ | $R_\eta$*, $R_\phi$ | $R_{Had.}$* |
| Widths | $w_{s,3}$, $w_{s,tot}$ | $w_{\eta,2}$* | - |
| Shapes | $\Delta E$, $E_{ratio}$ | * Used in PhotonLoose. | |

## Energy Ratios

$$R_\eta = \frac{E^{S2}_{3\times7}}{E^{S2}_{7\times7}} \qquad R_\phi = \frac{E^{S2}_{3\times3}}{E^{S2}_{3\times7}}$$




$$R_{Had} = \frac{E^{Had}_T}{E_T}$$

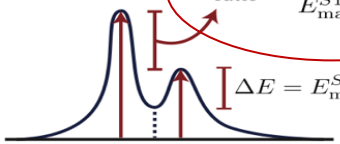
$$f_1 = \frac{E_{S1}}{E_{Tot.}}$$

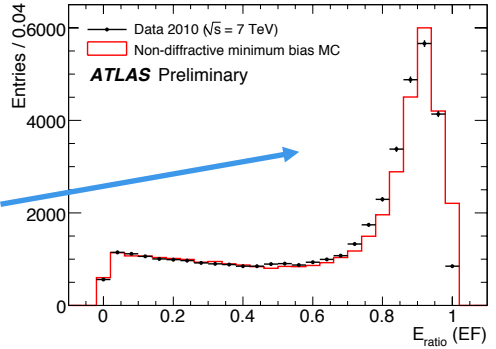
$$f_{side} = \frac{E^{S1}_7 - E^{S1}_3}{E^{S1}_3}$$

## Shower Shapes

$$E_{ratio} = \frac{E^{S1}_{max,1} - E^{S1}_{max,2}}{E^{S1}_{max,1} + E^{S1}_{max,2}}$$

$$\Delta E = E^{S1}_{max,2} - E^{S1}_{min}$$

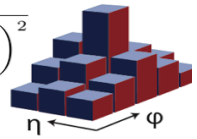





## Widths

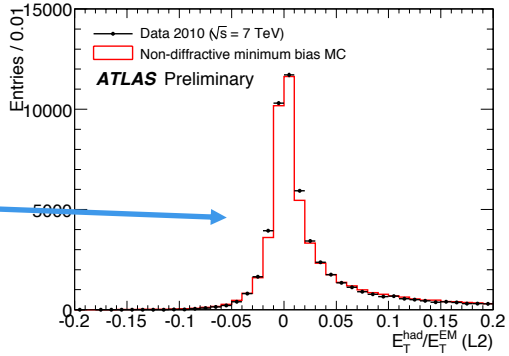
$$w_{\eta,2} = \sqrt{\frac{\sum E_i \eta_i^2}{\sum E_i} - \left(\frac{\sum E_i \eta_i}{\sum E_i}\right)^2}$$

Width in a 3×5 (Δη×Δφ) region of cells in the second layer.

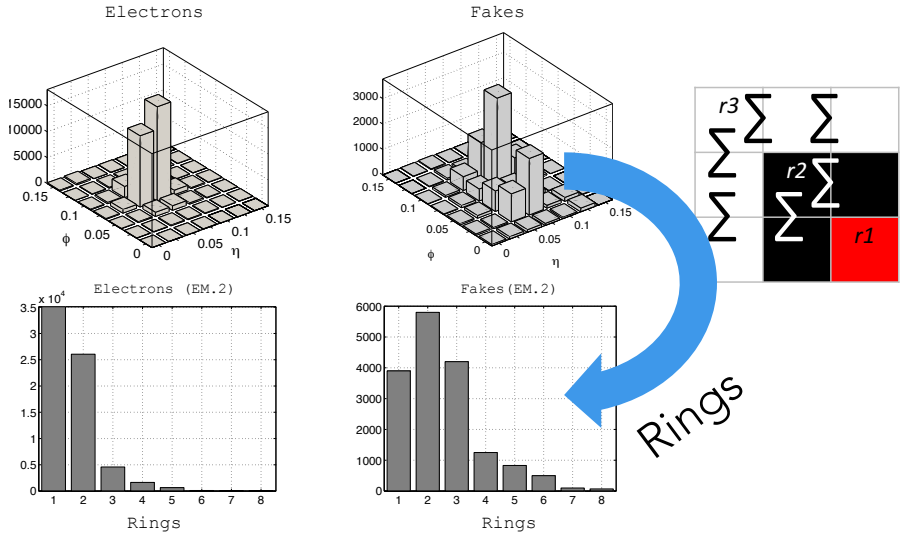



$$w_s = \sqrt{\frac{\sum E_i (i - i_{max})^2}{\sum E_i}}$$

ws3 = $w_1$ uses 3 strips in η; wstot is defined similarly, but uses 20 strips.



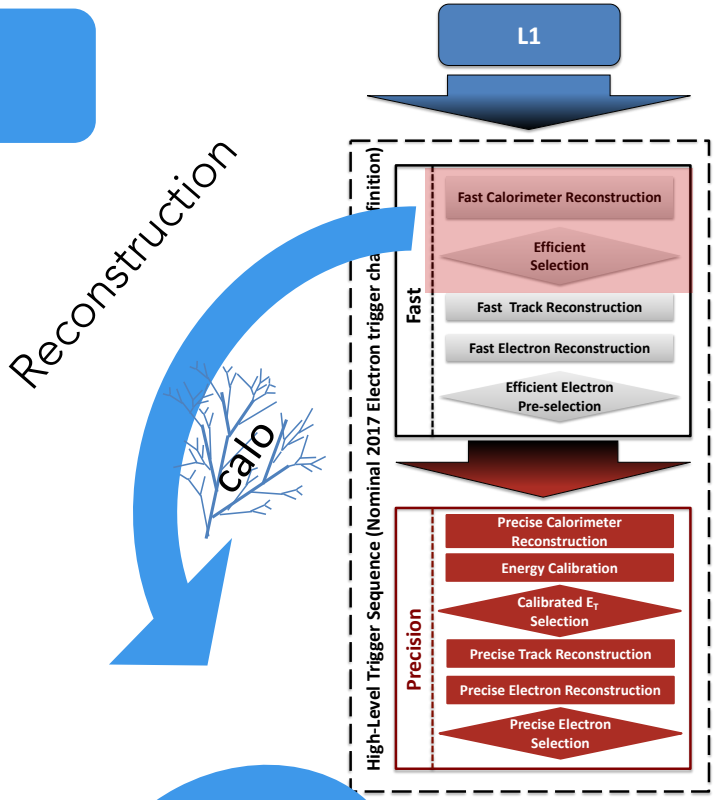

New

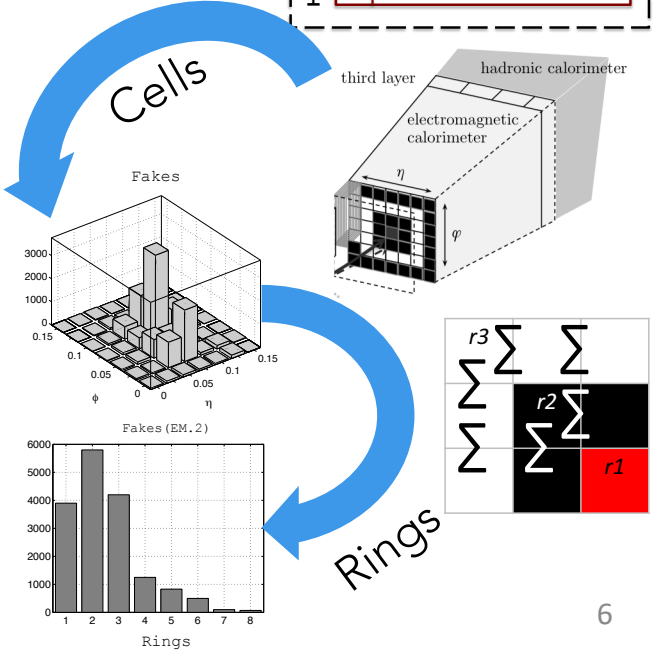



- Ringer Shape:
  - Concentric rings are built for all layers;
  - Compact cell information used to describe the event throughout of the calorimeter

# Ringer Reconstruction

- Ringer reconstruction setup in the Fast Calorimeter Reconstruction:
  - Built from all calorimeter layers, centered in a window from the cluster barycenter;
  - First ring in each layer is the cell closest to cluster barycenter;
  - The next ring is the collection of cells around the previous one; ring value is the sum $E_T$ of all cells composing the ring;
  - This process reduces the amount of information (w.r.t. using all cells), but keeps the physics interpretation (typical EM object shower shape);



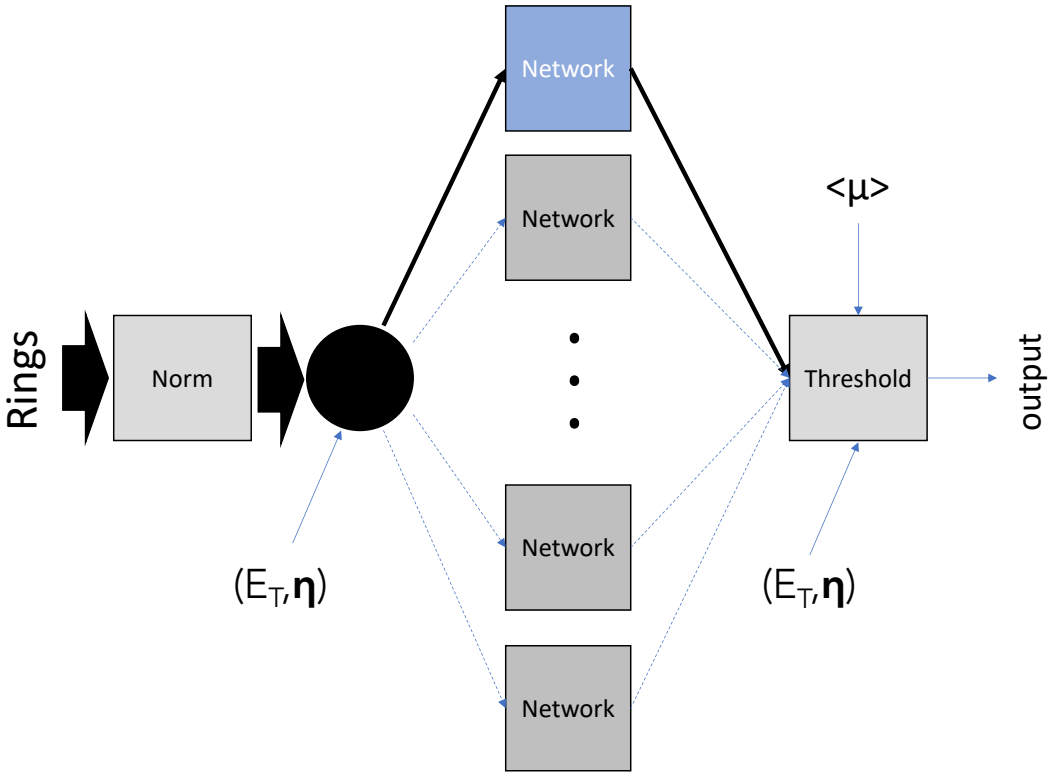




| Total number of Rings per layer (covering 0.4 x 0.4 region in η x φ) | | | | | | |
|---|---|---|---|---|---|---|
| PS | EM1 | EM2 | EM3 | HAD1 | HAD2 | HAD3 |
| 8 | 64 | 8 | 8 | 4 | 4 | 4 |

# Ringer Ensemble

- As the same way that the standard shower shapes quantities are subject to distortions according to the particle interaction position and energy in ATLAS, as well are the rings;

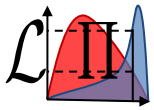
- To deal with these distortions, as chosen from the offline analysis, the ringer process online data through an ensemble that is **defined in bins of eta and energy**;
  - Specific models for defined regions to minimize distortions;

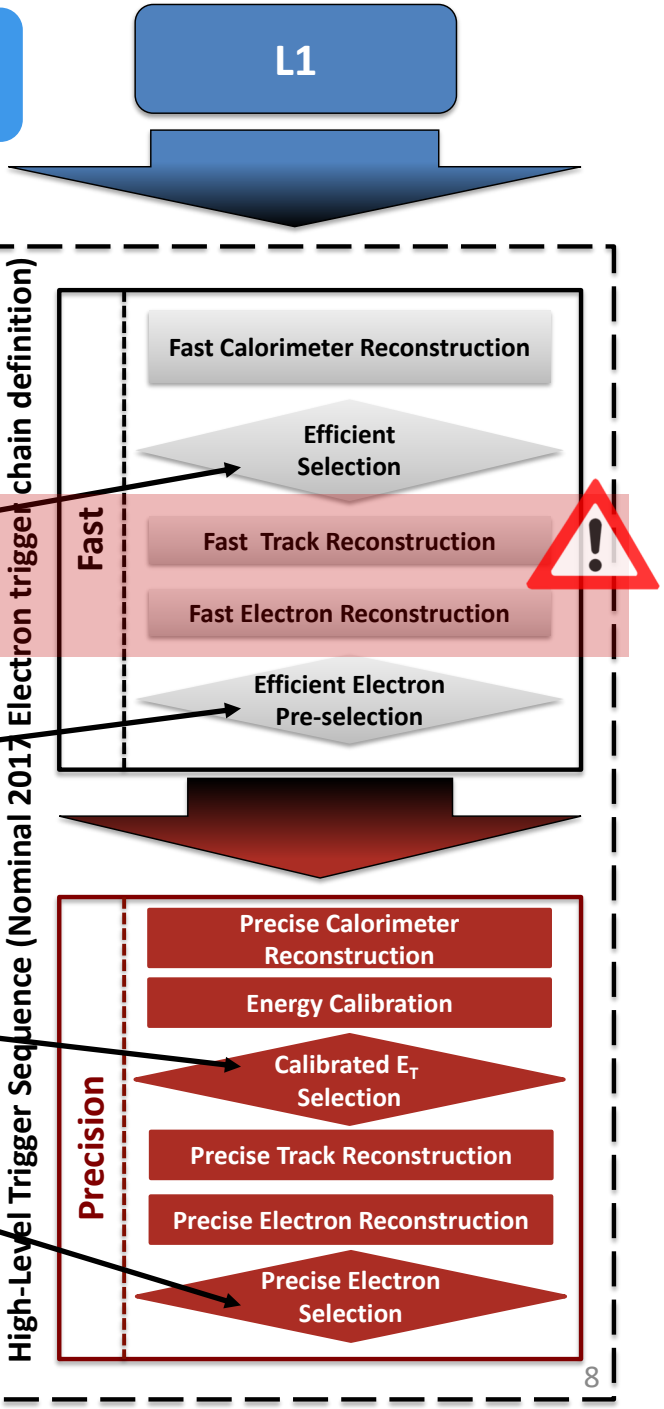



- The ensemble is build from single-layer **MLP neural networks**;
- Ensemble operation is set to reach high detection efficiency as defined by the HLT Precision step (Likelihood);
  - Best trained models are those that optimize the fake rate reduction.

# HLT e/g Workflow

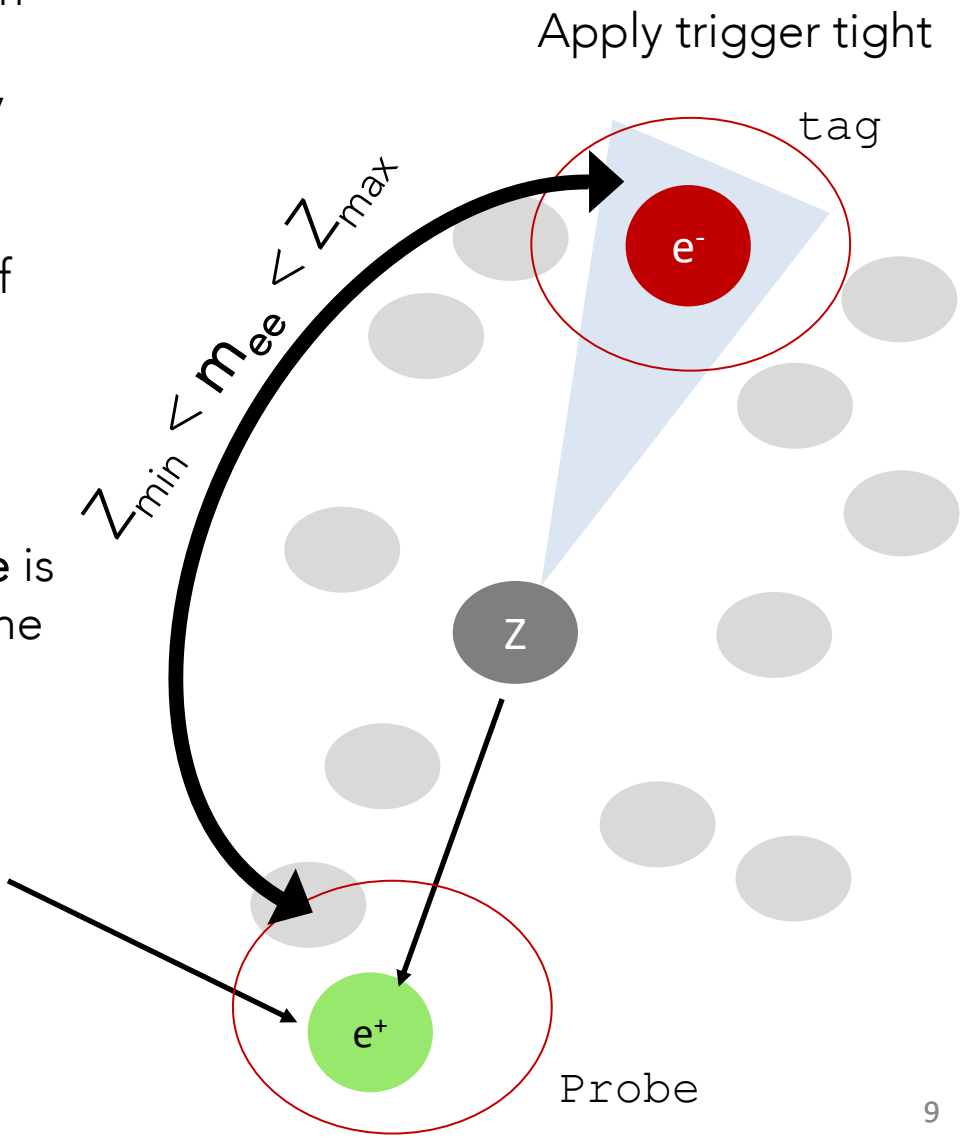
| Old | New |
|---|---|
| Fast Shower Shapes reconstruction | Ringer Reconstruction |
| Cut-Based Selection | Ensemble of neural networks + Pileup correction |
| Track reconstruction | |
| Track cuts | |
| Precise Shower Shapes Reconstruction | |
| Energy Calibration | |
| Precise $E_T$ cut | |
| Precise track reconstruction | |
| Electron Identification based on the Likelihood at relevant quantities (calo+track) + Pileup correction | |



# The Tag And Probe Method

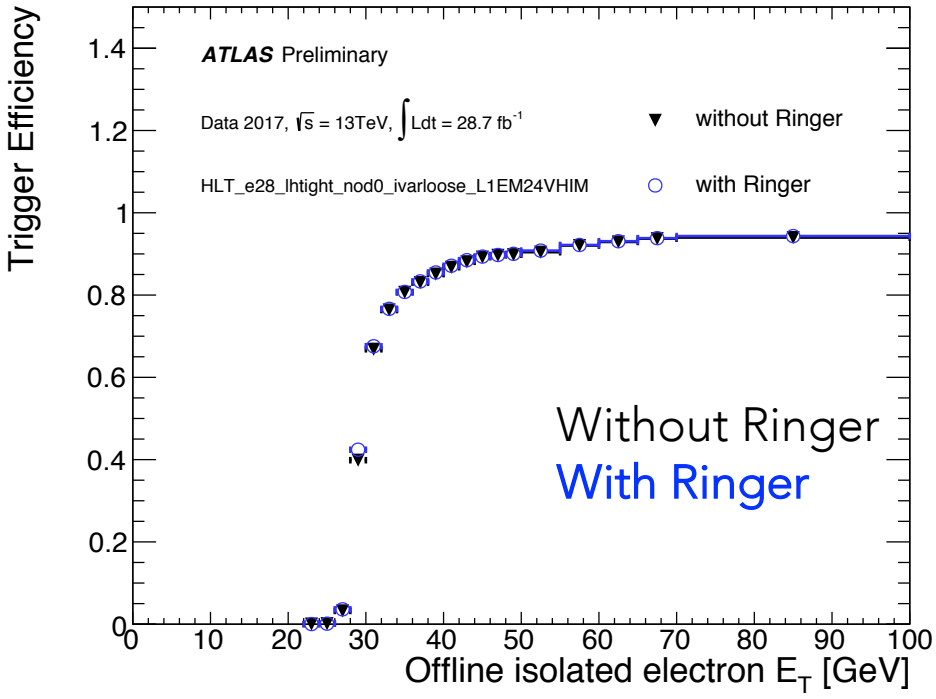
- Need a clean unbiased offline electron sample for efficiency measurement;
  - Use Z → ee / J/**ψ** → ee / W → e**ν** characteristic decays;
- Apply strict selection criteria to one of the decay electrons (**tag**);
  - Usually apply a tight trigger;
  - Isolation;
- The second decay electron, the **probe** is identified with the tag by $m_{ee}$ within the mass window;
  - **Probe electrons** are used for the efficiency measurement;
  - The probe electron is a **trigger unbiased.**



# Trigger Efficiency

Full 2017 period

- Efficiency measurements in Run2;
- Clean unbiased samples give by the tag & probe method;
- Same signal efficiency w.r.t the old paradigm (cut-based at the fast calo step);
  - Primary chain latency reduction: 200 ms to ~100 ms;
  - High rejection power (~2-3X);
- Electron + photon slice: ~1/4 latency reduction;




- Kept operating backup trigger sequence with the previously cut-based selection to assess;
  - Efficiencies changes;
  - Offline impact.

https://twiki.cern.ch/twiki/bin/view/AtlasPublic/EgammaTriggerPublicResults#Performance_of_Ringer_in_Trigger

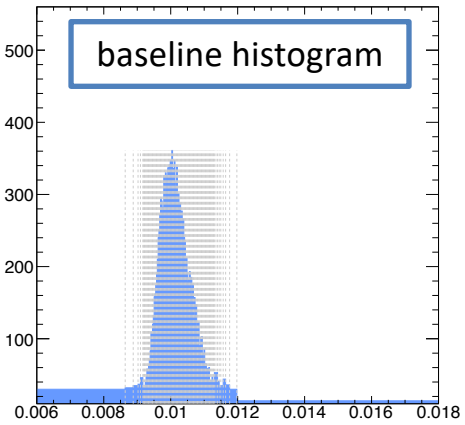
# Offline Impact

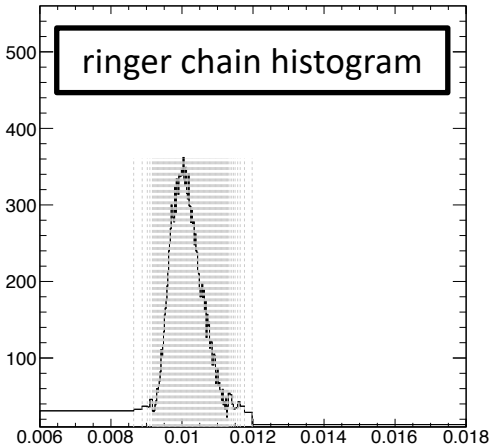
- We are also interested in assessing whether:
  - Is there a bias in the collected probe samples when we change single-lepton triggers (**the tag trigger**)?
  - e.g. would the offline standard quantities (**shower shapes**) be biased by the ringer chains?

- To evaluate this, we apply statistical tests comparing histograms built with the quantity profiles of the probes distribution:
  - **Comparing the shape of the histograms**;
  - One histogram is built with the monitoring chain (previous paradigm) applied to the tag;
  - Other histogram has tag passing equivalent ringer chain;

| noringer | ringer |
| --- | --- |
| Standard T&P procedure with tag passing single_lepton trigger list e28_lhtight_nod0(_noringer)_ivarloose | Standard T&P procedure with tag passing single_lepton trigger list e28_lhtight_nod0(_ringer)_ivarloose |

**Check for distortions (currently using histograms)**

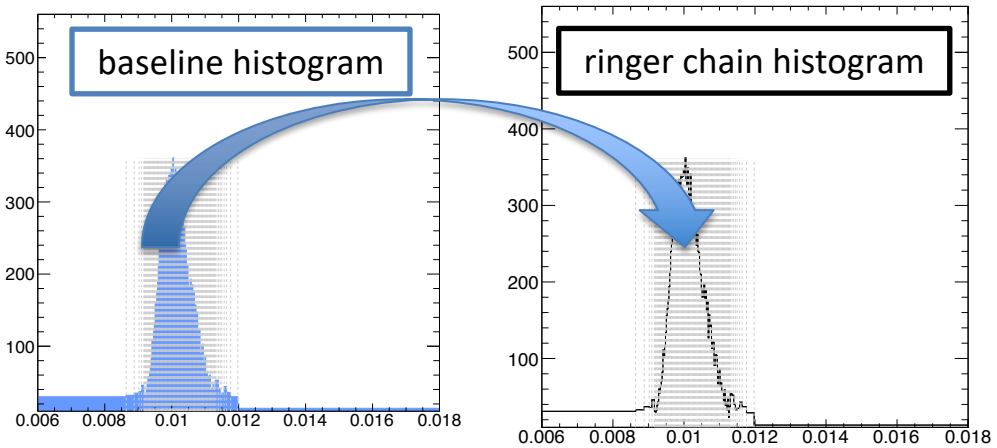



This process is applied for all standard quantities and phase space regions.

# Offline Impact

- As the number of T&P pairs in each phase space bin are not the same to small differences in the chains operations points:
  - The total histogram entries are not the same
  - we remove samples at random in the histogram with higher counts;
  - Force both histograms to have the same number of counts;

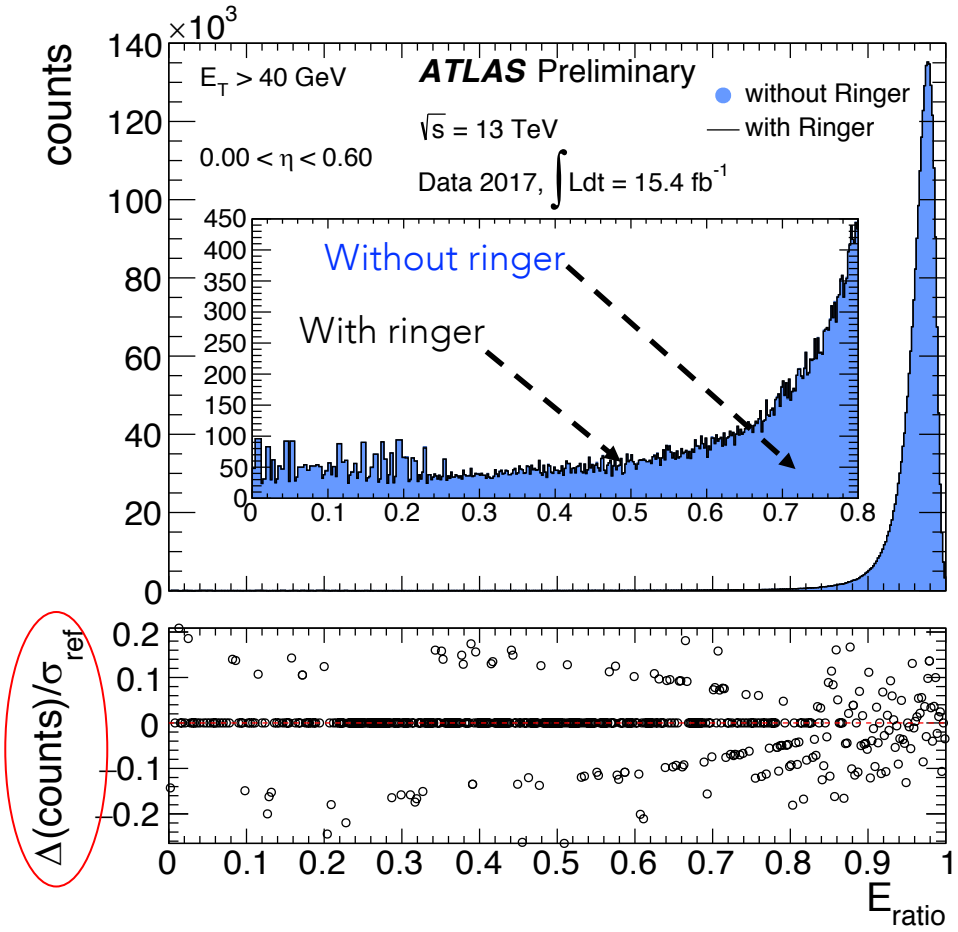



- To reduce the number of bins and profit from Gaussian/Poissonian errors approximation:
  - Adaptive bin grid is calculated in the reference hist;
  - The edges are then propagated to rebin ringer chain hist.

# Offline Impact

$$E_{ratio} = (E^{1}_{max1} - E^{1}_{max2})/(E^{1}_{max1} + E^{1}_{max2})$$

- To verify any **change of shape** after the introduction of the ringer in the trigger sequence:
  - We assess the $\Delta(counts)/\sigma$ (~chi residuals in black markers) where the ringer histogram is used as a model to the baseline histogram (experimental outcome);

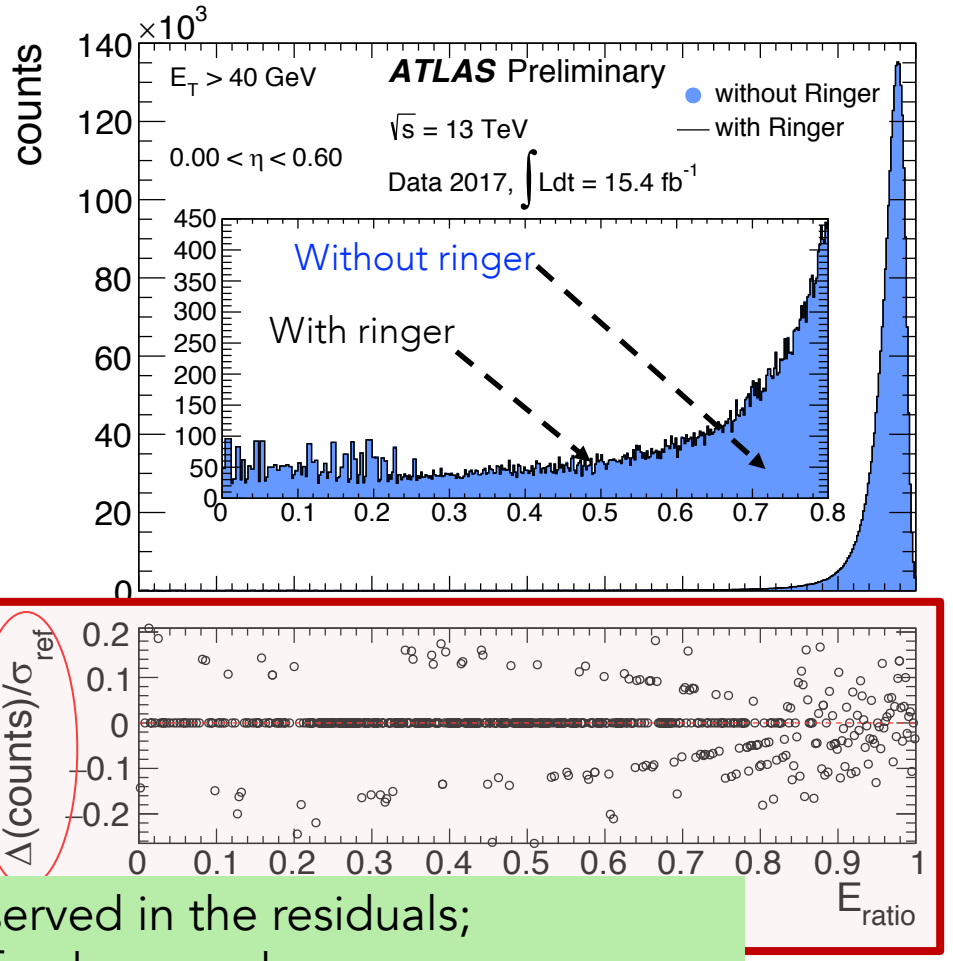



https://twiki.cern.ch/twiki/bin/view/AtlasPublic/EgammaTriggerPublicResults#Performance_of_Ringer_in_Trigger

# Offline Impact

$$E_{ratio} = (E^{1}_{max1} - E^{1}_{max2})/(E^{1}_{max1} + E^{1}_{max2})$$

- To verify any **change of shape** after the introduction of the ringer in the trigger sequence:
  - We assess the $\Delta(counts)/\sigma$ (~chi residuals in black markers) where the ringer histogram is used as a model to the baseline histogram (experimental outcome);




- No clear pattern can be observed in the residuals;
  - They seem to oscillate freely around zero;
- No single residual can be found above 1 sigma deviation for all phase space regions and quantities;

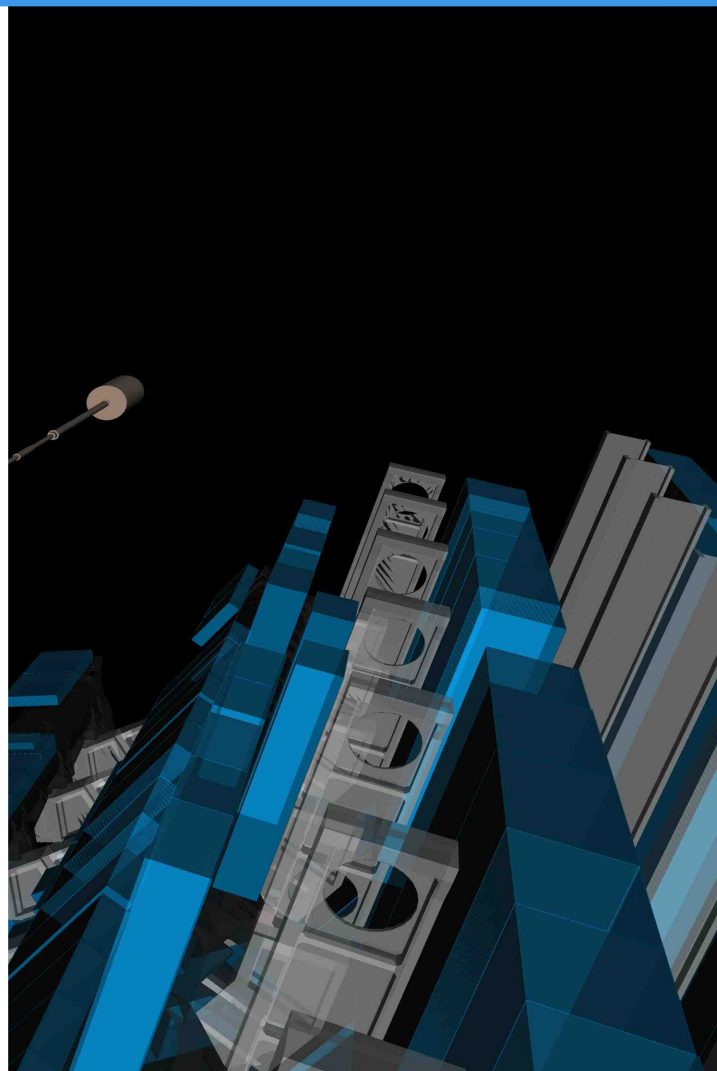
https://twiki.cern.ch/twiki/bin/view/AtlasPublic/EgammaTriggerPublicResults#Performance_of_Ringer_in_Trigger

# Conclusions

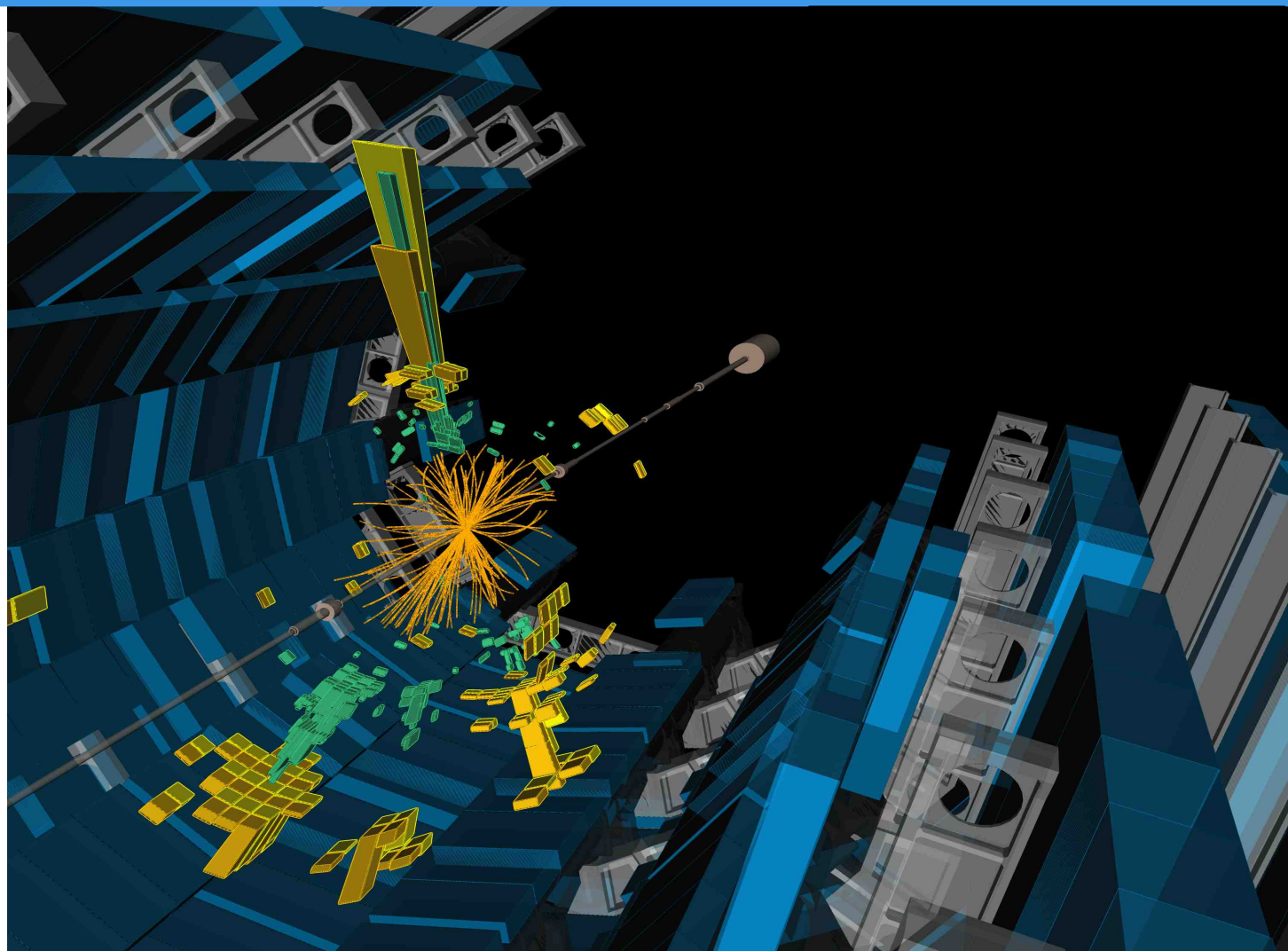
- Ring-shaped Calorimetry: introduced new concept for characterization of the shower development in the ATLAS Trigger System;
  - A complete shower description of the event throughout of the calorimeter;
  - Compact information from the cells.

- Updated at the fast calo step to use an ensemble of neural networks based on calorimetry information;
  - Electron trigger kept operating with similar electron efficiency with large improvement in the processing requirements:
  - 200ms → 100ms, 2-3x reduction in fake rate;
  - Residuals are small and oscillate freely around zero which suggests absence of bias

# Thanks a lot for all support!!!



# Backup



# Ringer Tuning Approach

- Data Extraction:
  - Shapes extracted from the FastCalo rings;
  - Event selection using TrigEgammaAnalysis:
    - Signal: T&P selection + Offline LH Veryloose on probes;
    - Background: Veto Probes;
  - Tuning networks binning configuration:
    - $E_T$ = [15, 20, 30, 40, 50, ∞[ and **η** = [0, 0.8, 1.35, 1.52, 2.37, 2.5], 25 bins;
  - Threshold binning configuration:
    - $E_T$ = [15, 20, 30, 40, 50, ∞[ and **η** = [0, 0.8, 1.35, 1.52, 2,37 2.5], 25 thresholds; (This can be latter adapted);
- Model Extraction:
  - Standard full-connected 1 hidden layer MLP (as usual).