Universidade Federal do Rio de Janeiro

17th International workshop on
Advanced Computing and Analysis Techniques in Physics research

Ring-shaped Calorimetry Information for a Neural eGamma Identification with ATLAS Detector

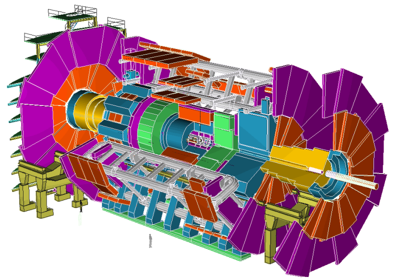
João Victor da Fonseca Pinto on behalf of ATLAS Collaboration

Universidade Federal do Rio de Janeiro
UFRJ

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

Outline

- ATLAS Trigger system;
- Selection algorithms;
- Ringer approach:
  - Ringer algorithm;
  - Neural Network;
  - Ringer hypothesis;
- Implementation in details;
- Ringer operation for the trigger;
- Summary and ongoing for 2016.

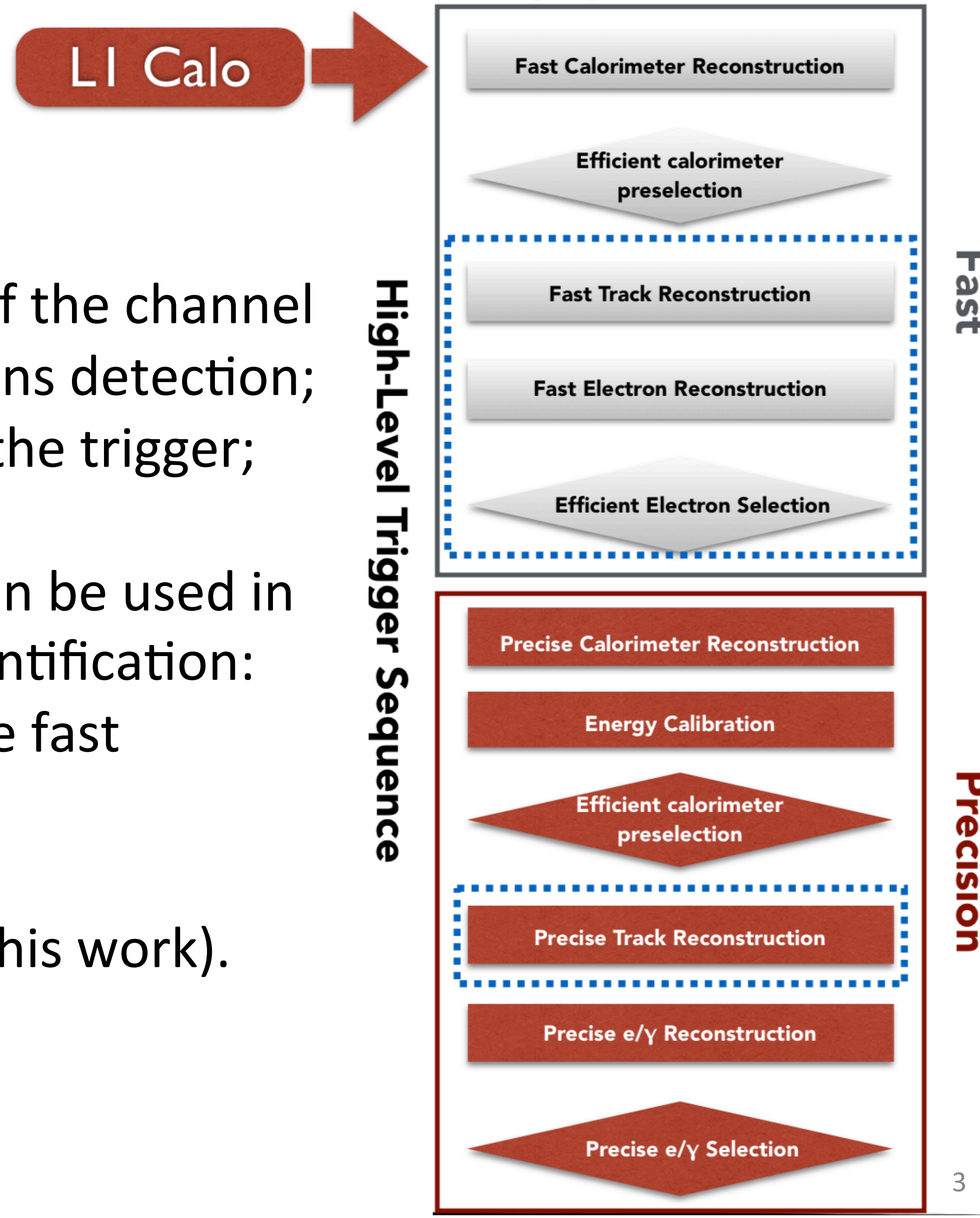


The e/γ trigger goal:

- Maximize the efficiency of the channel e/γ : Electrons and Photons detection;
- Reduce the fake rates of the trigger;

Three types of algorithms can be used in the electron and photon identification:

- Cut-based (used in the fast preselection);
- Likelihood;
- Ringer (proposed by this work).

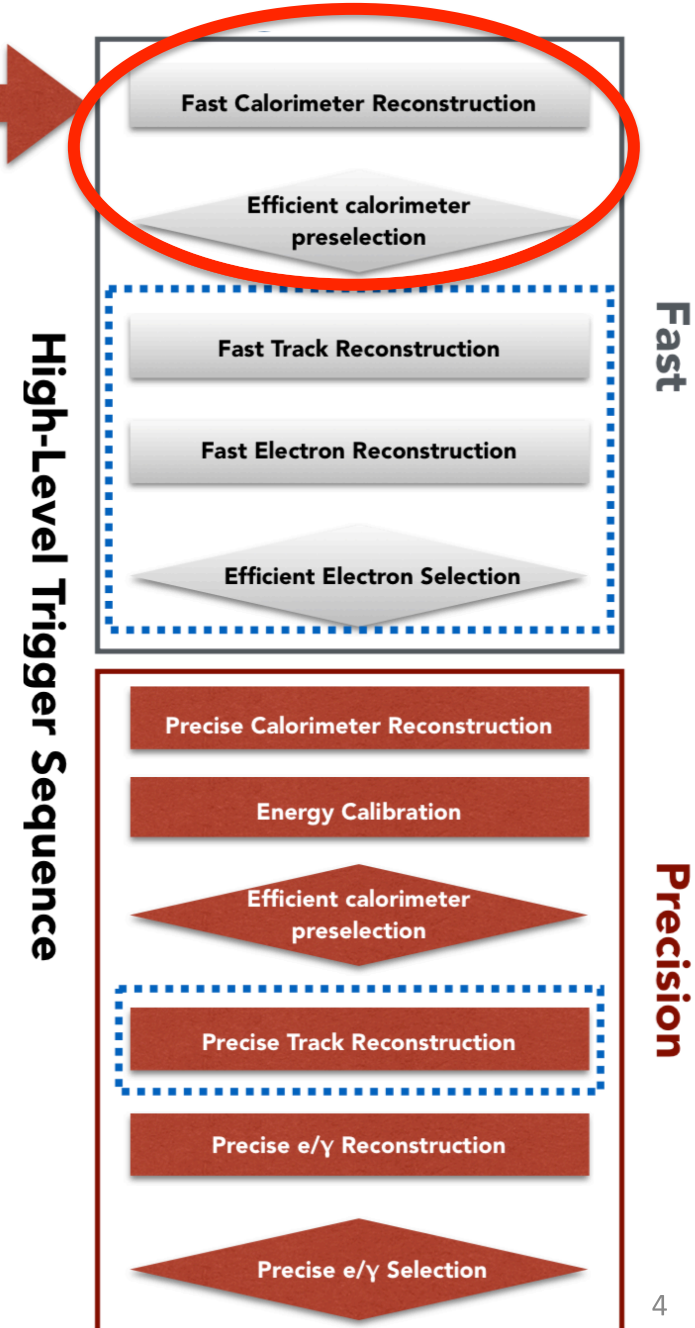


The e/γ trigger goal:

- Maximize the efficiency of the channel e/γ : Electrons and Photons detection;
- Reduce the fake rates of the trigger;

Three types of algorithms can be used in the electron and photon identification:

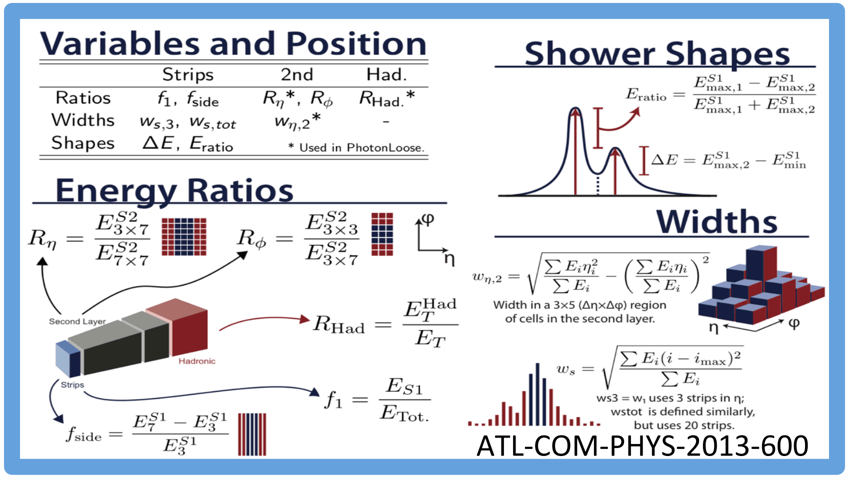
- Cut-based (used in the fast preselection);
- Likelihood;
- Ringer (proposed by this work).



Electron Identification

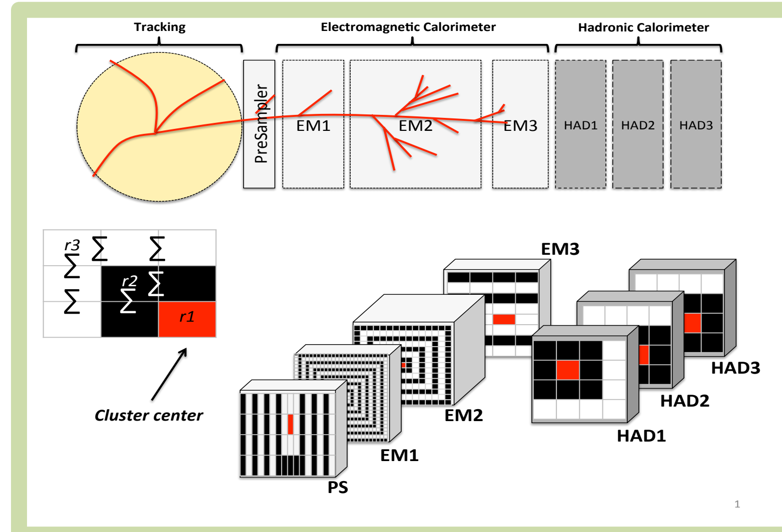
Calorimeter Features

CutID
Likelihood



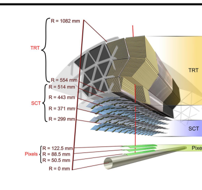
OR

Ringer



Other Features

If Electron
Pure tracking features
+
Combined Tracking & Calorimeter features



Classifiers

Cut-based
- Applies linear cuts over the over shower shape

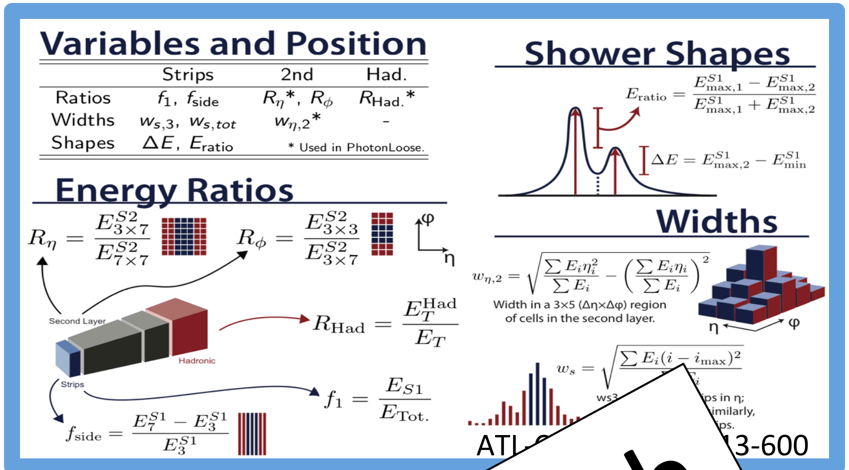
Likelihood
- Applies Naïve Bayes like approach over shower shape

Ringer
- Applies Neural Networks fed with rings.

Electron Identification

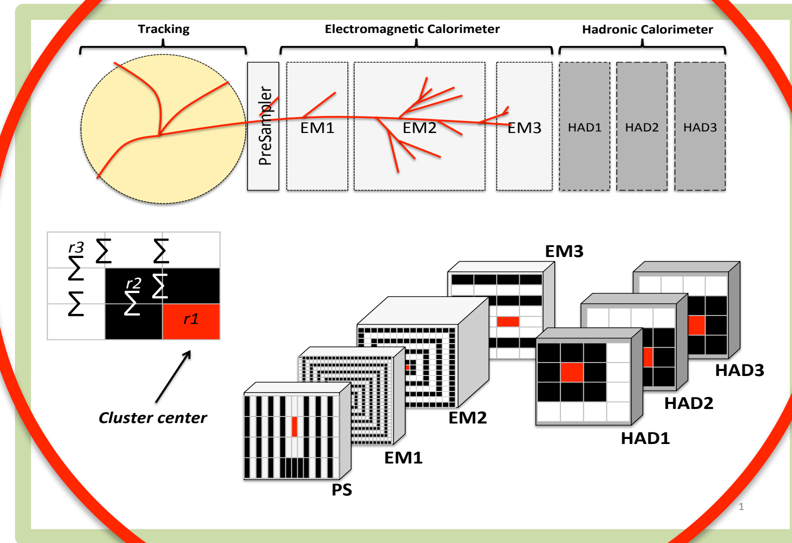
Calorimeter Features

CutID
Likelihood



OR

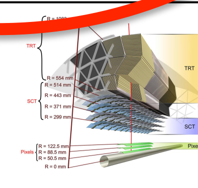
Ringer



New approach

Other Features

If Electron
Pure tracking features
+
Combined Tracking & Calorimeter features



Classifiers

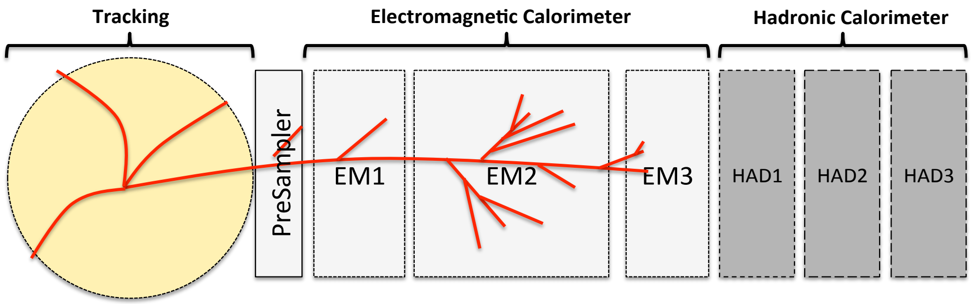
Cut-based
- Applies linear cuts over the over shower shape

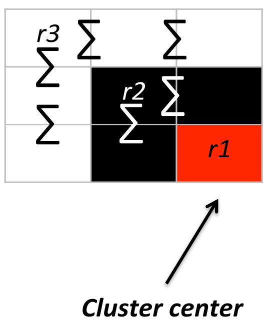
Likelihood
- Applies Naïve Bayes like approach over shower shape

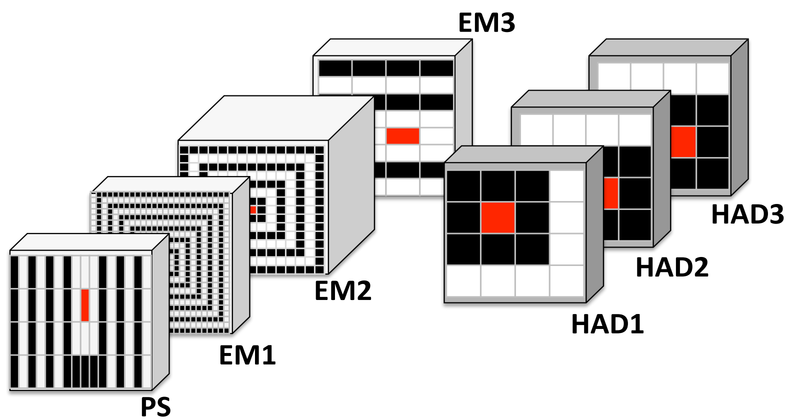
Ringer
- Applies Neural Networks fed with rings.

Calorimeter Ring concept

Algorithm skept:



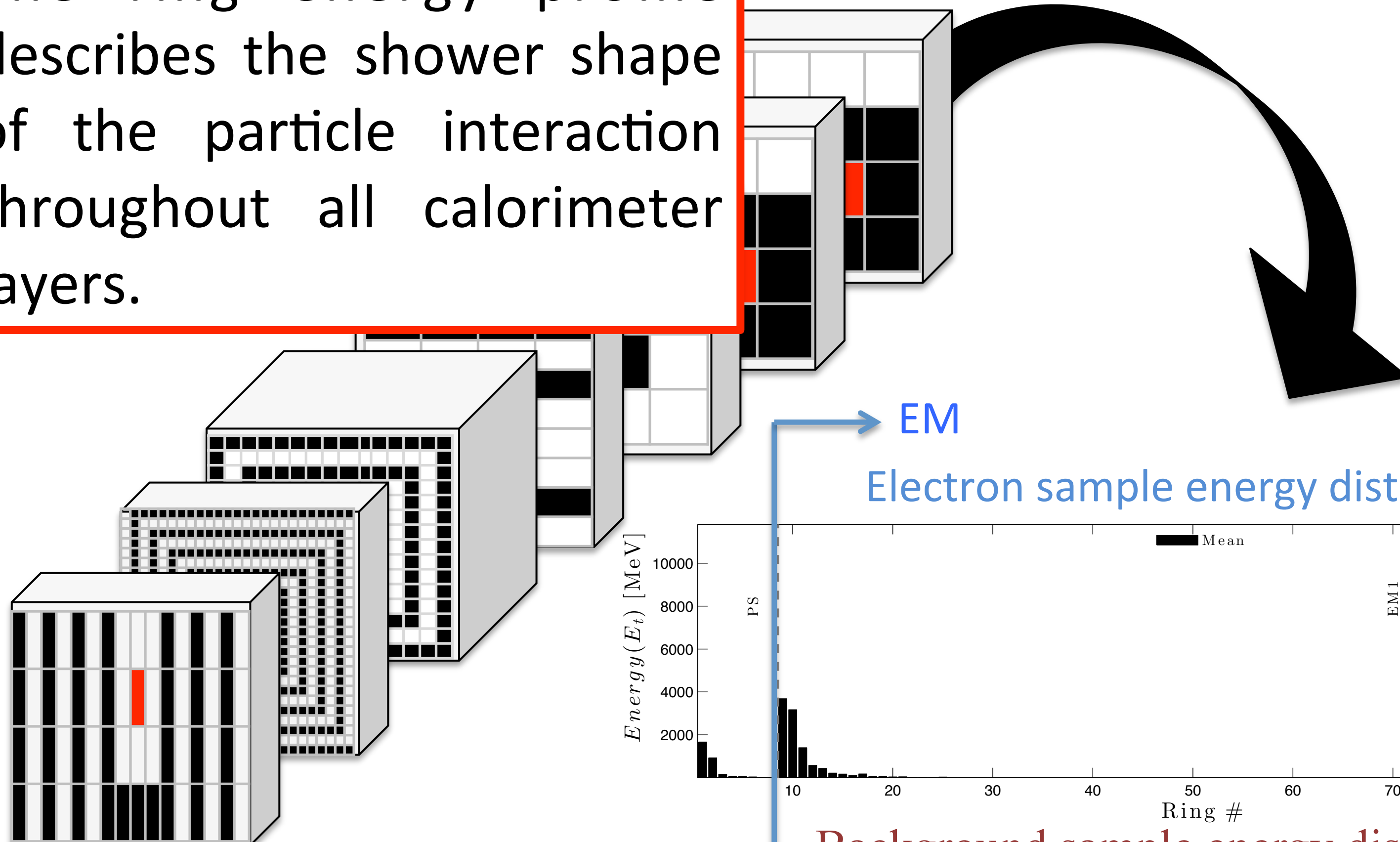


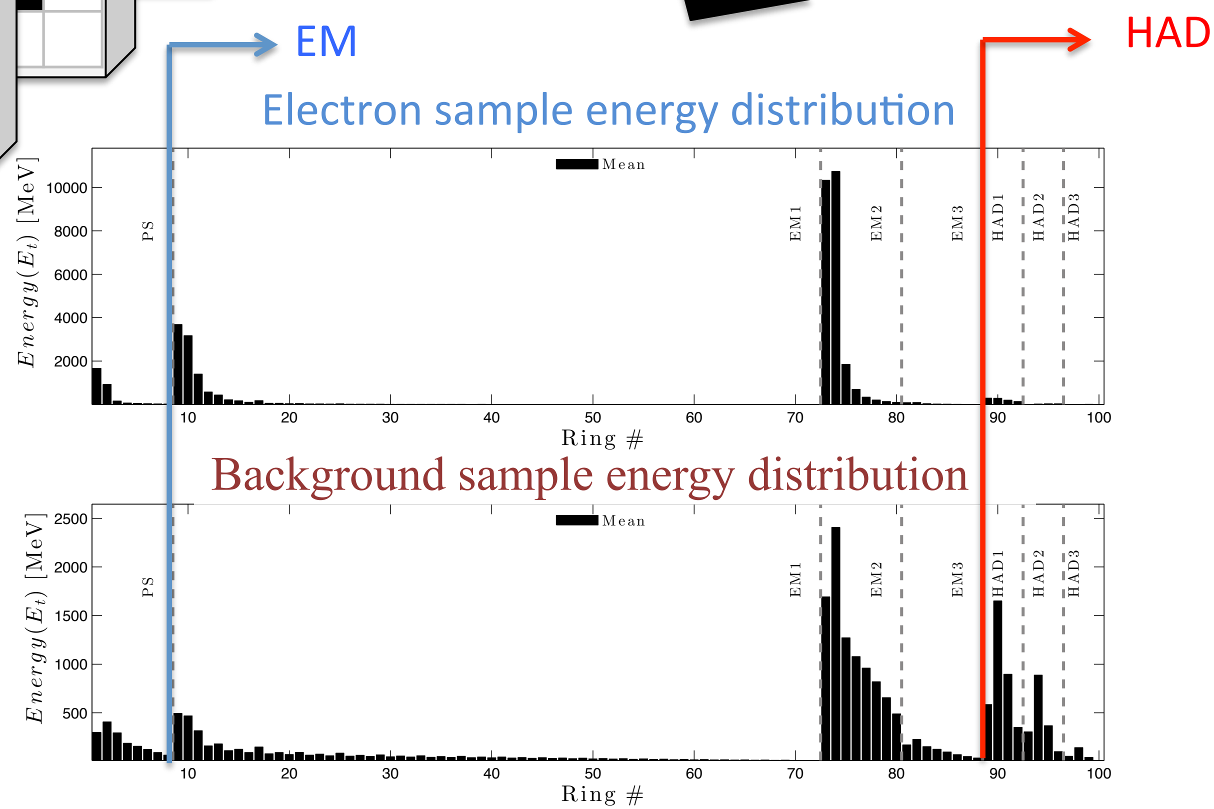


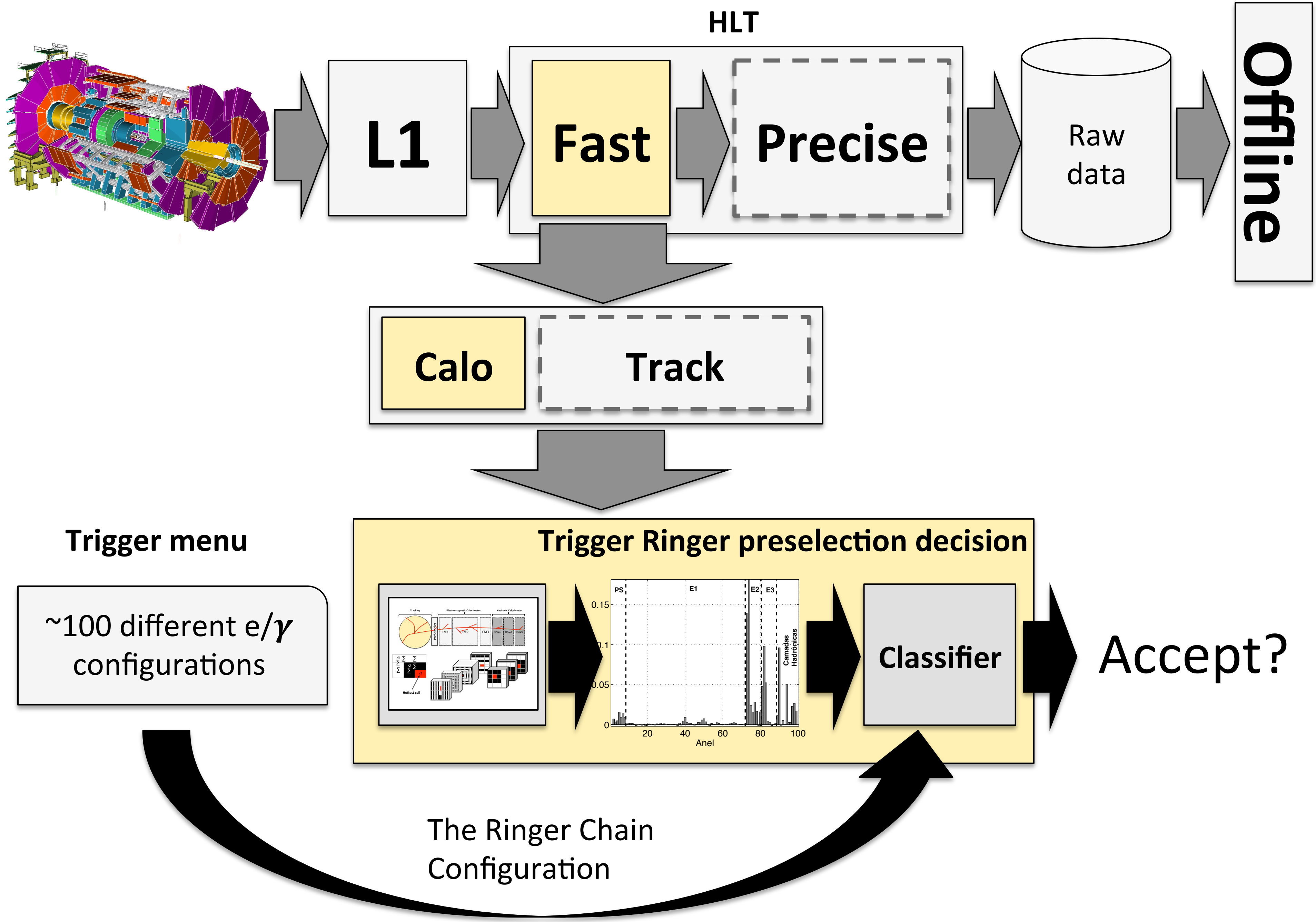
Total number of Rings per layer (covering 0.4 x 0.4 region in η x φ)						
PS	EM1	EM2	EM3	HAD1	HAD2	HAD3
8	64	8	8	4	4	4

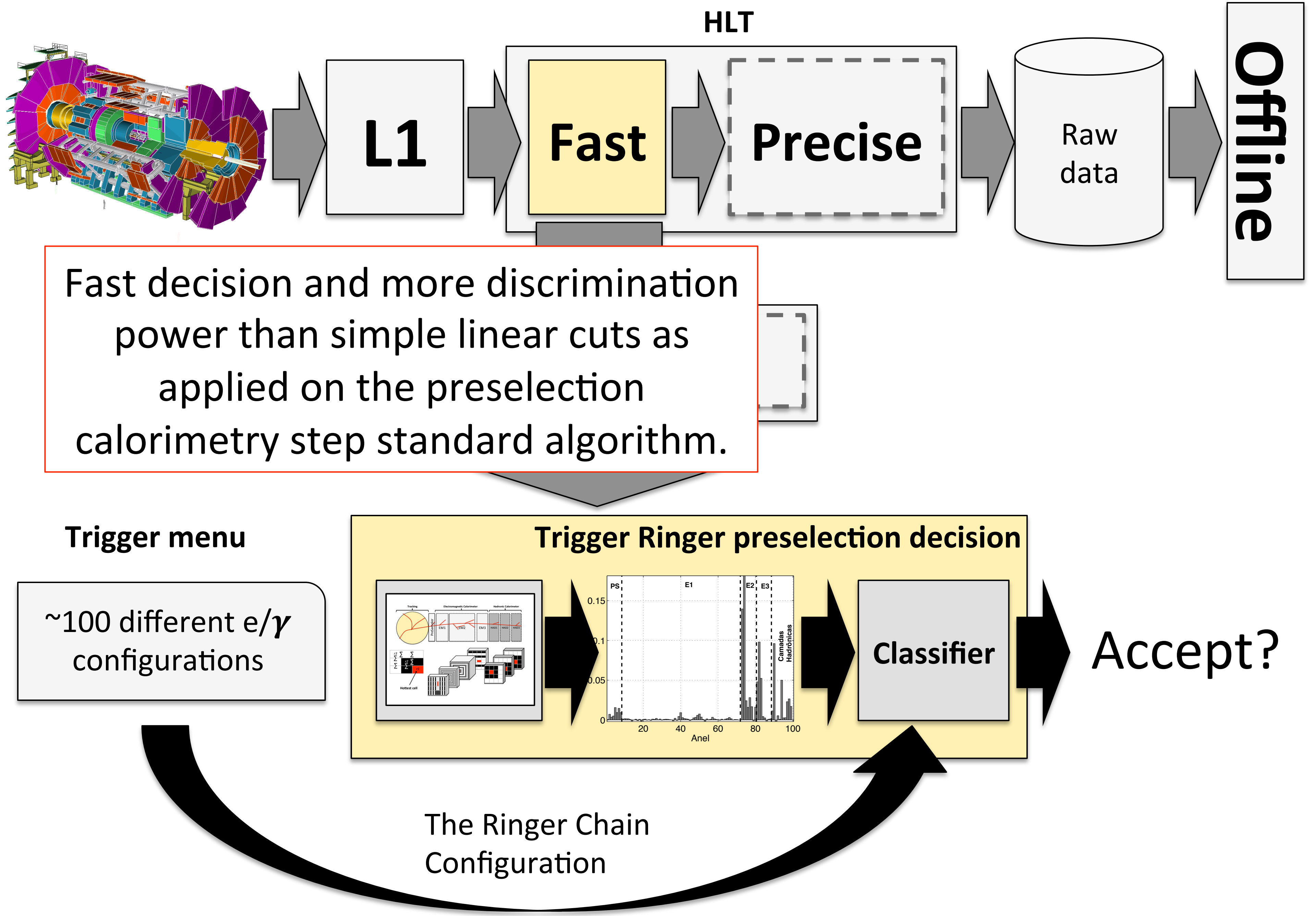
- Build using all calorimeter layers, centered in a window (0.4 x 0.4 in η x ϕ) at the hottest cell of each layer;
- Hottest cell: first "ring" (on each layer);
- Next ring: collection of cells around the previous one;
- The ring "value" is the sum of the E_T of all cells composing the ring;
- Provides input data reduction for the neural processing (w.r.t using all cells);
- Keeps the physics interpretation (typical EM object shower shape).

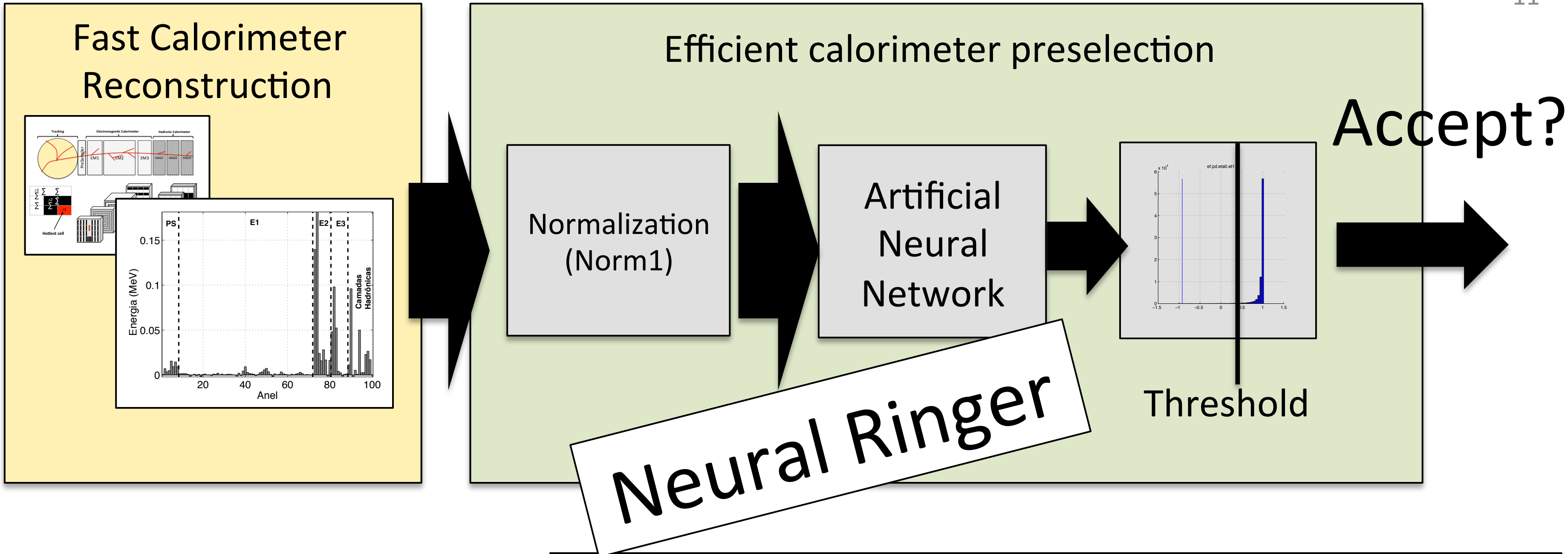
The ring energy profile describes the shower shape of the particle interaction throughout all calorimeter layers.

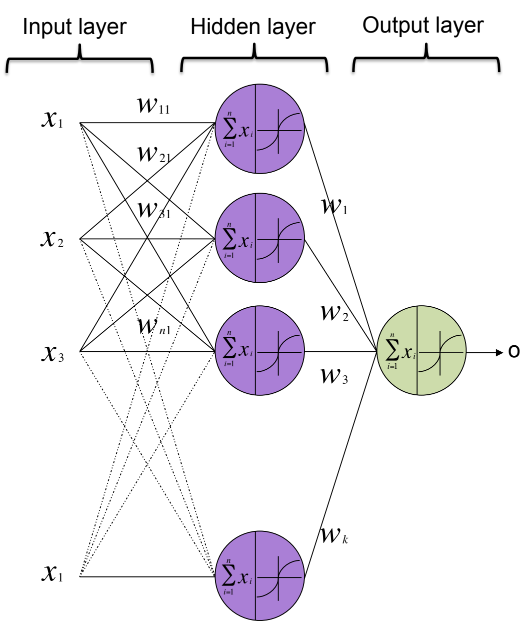





HLT
L1
Fast
Precise
Raw data
Offline
Calo
Track
Trigger menu
~100 different e/γ configurations
Trigger Ringer preselection decision
PS
E1
E2
E3
0.15
0.05
0
20
40
60
80
100
Anel
Camadas Hadrônicas
Classifier
Accept?
The Ringer Chain Configuration


HLT
L1
Fast
Precise
Raw data
Offline
Fast decision and more discrimination power than simple linear cuts as applied on the preselection calorimetry step standard algorithm.
Trigger menu
~100 different e/γ configurations
Trigger Ringer preselection decision
PS
E1
E2
E3
Camadas Hadrônicas
0.15
0.1
0.05
0
20
40
60
80
100
Anel
Classifier
Accept?
The Ringer Chain Configuration





- A standard feed forward neural network is used for the particle identification task;
- Normalized ring information (norm-1 of the rings energy in all calorimeter layers) is fed into the input nodes of the neural network;
- The network output is compared to a threshold for final decision;
- Cross-Validation: used to evaluate the statistical fluctuations from the dataset.
- Network topology determination (number of neurons in the hidden layer): see backup slides.

Performance Indexes

Signal efficiency or Detection Probability:

$$PD(\%) = \frac{N_{e|e}}{N_e} \times 100\%$$

Background Efficiency or Fake Rate Probability:

$$FR(\%) = \frac{N_{e|b}}{N_b} \times 100\%$$

- N_e = Number of candidates on Zee, which **passed offline Likelihood tight** (aims at assessing sample with high purity).
- N_b = Number of candidates on background, **not labeled as isolated electron by truth (Z mother, W mother…)**.
- $N_{e|e}$ = Number of trigger events classified as electrons, given that they are counted as N_e.
- $N_{e|b}$ = Number of trigger events classified as electrons, given that they are counted N_b.

Tag and Probe (T&P): is a method to measure signal efficiency. Z→ee events with one electron passing default trigger and second (probe) reconstructed electron by offline are selected and used to check if probe electron has also passed trigger.

Ringer studies for the Trigger L2 Calorimetry optimization

ATLAS

CERN

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO

UFRJ

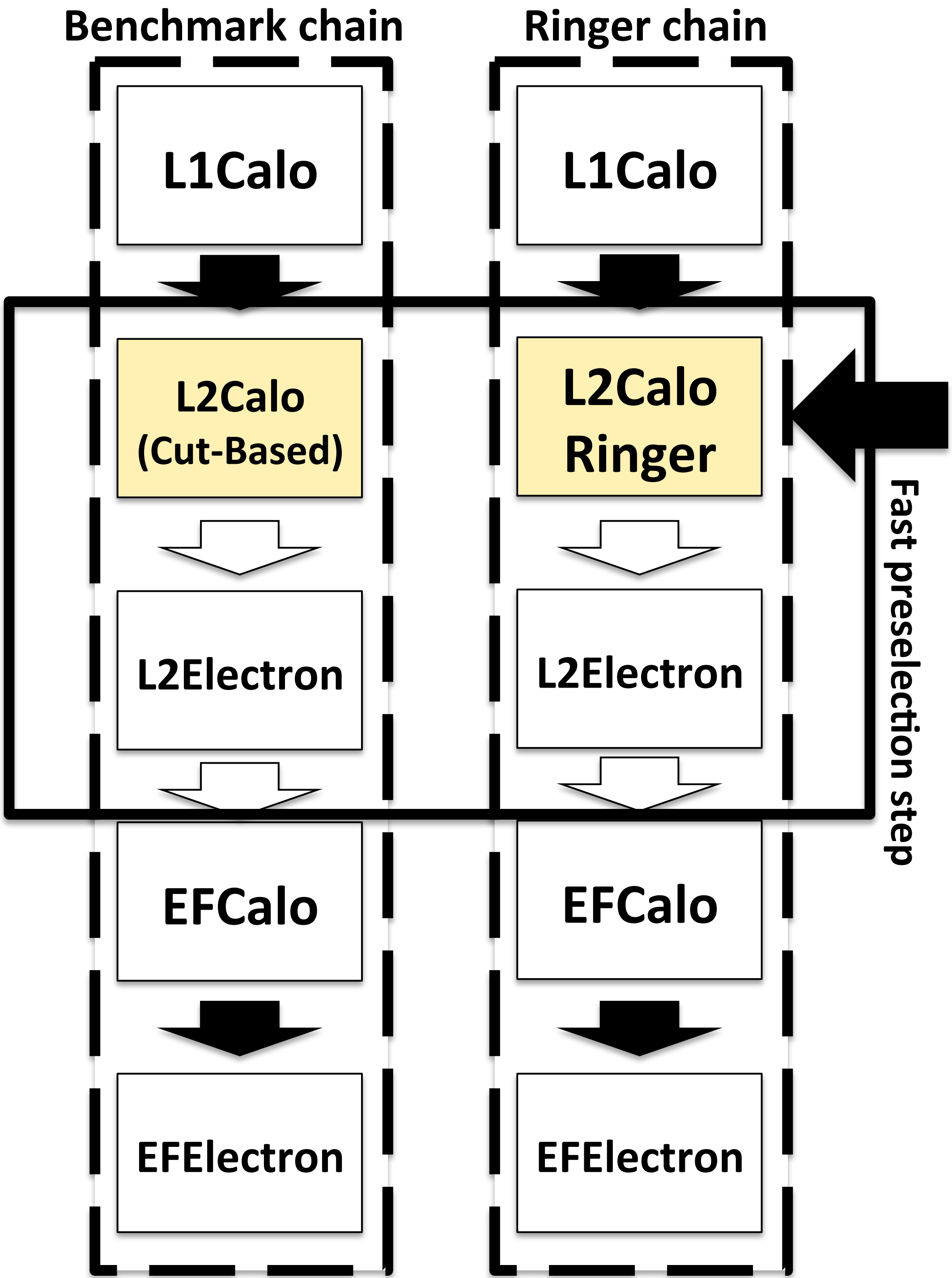
lps

COPPE/Poli/UFRJ

Laboratório de Processamento de Sinais

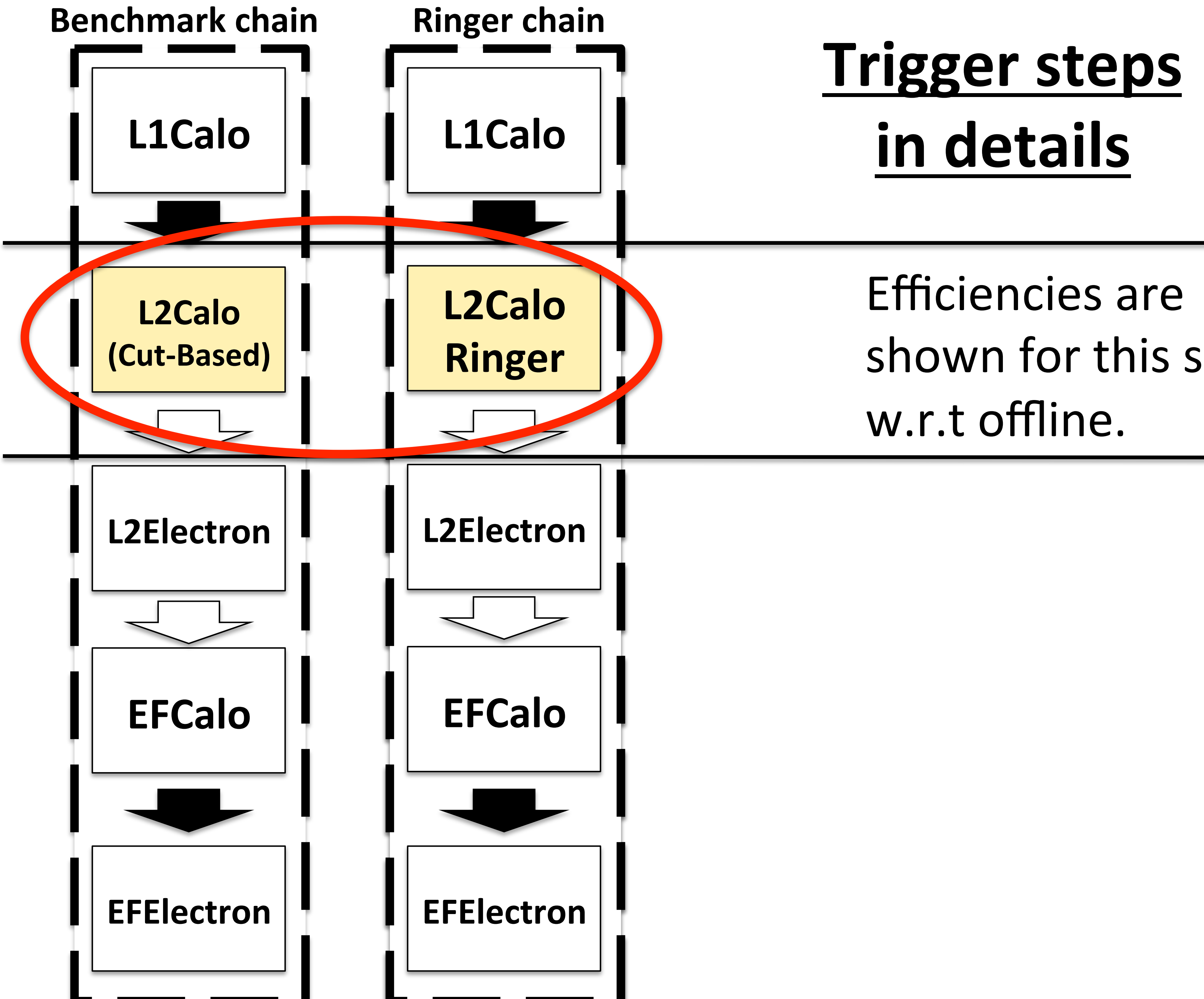
Inteligência Computacional. Inovação

Trigger steps in details

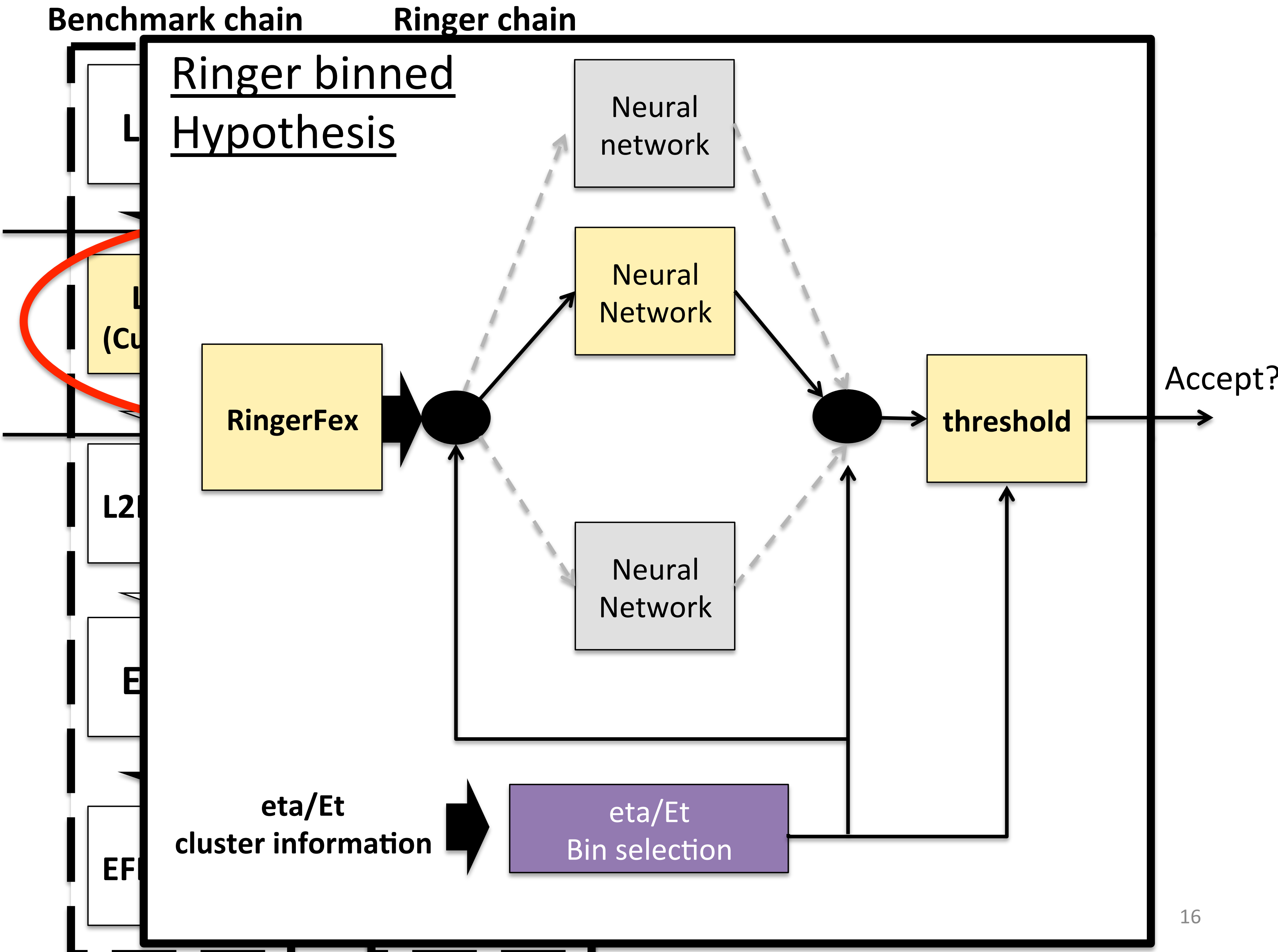


Calorimetry pre-selection step.

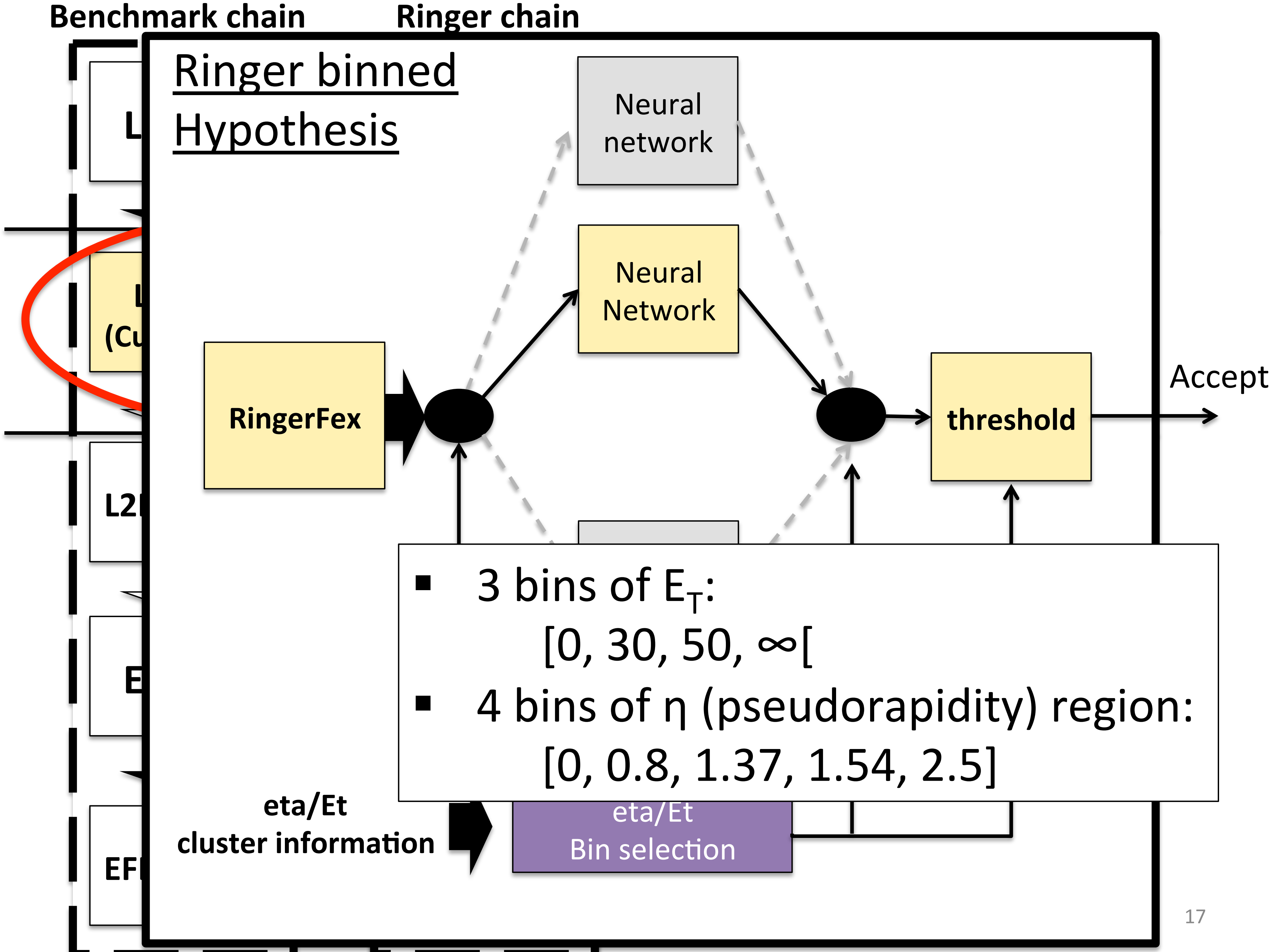
Trigger steps in details



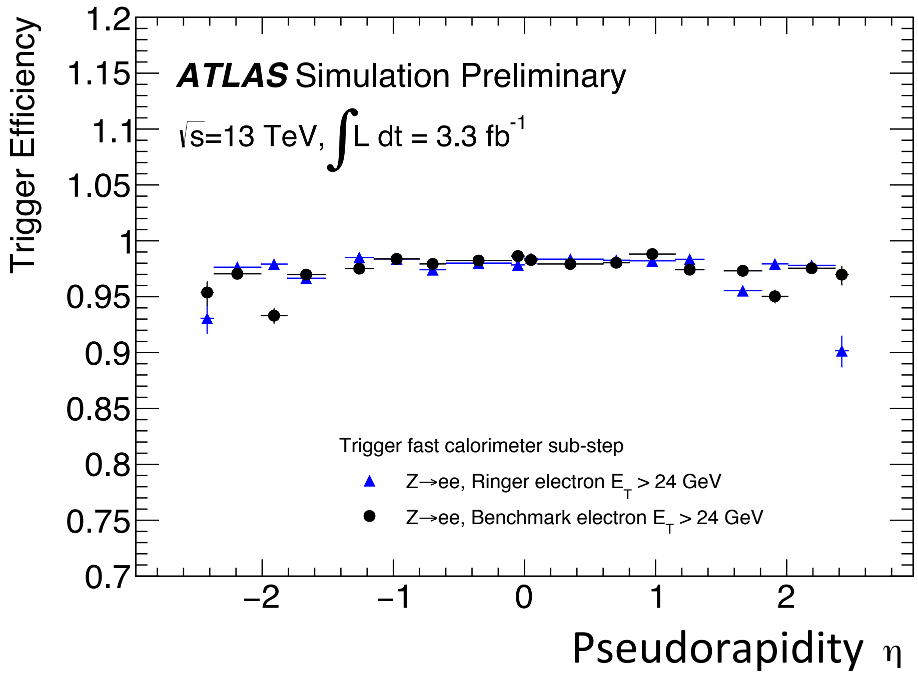
Efficiencies are shown for this step w.r.t offline.

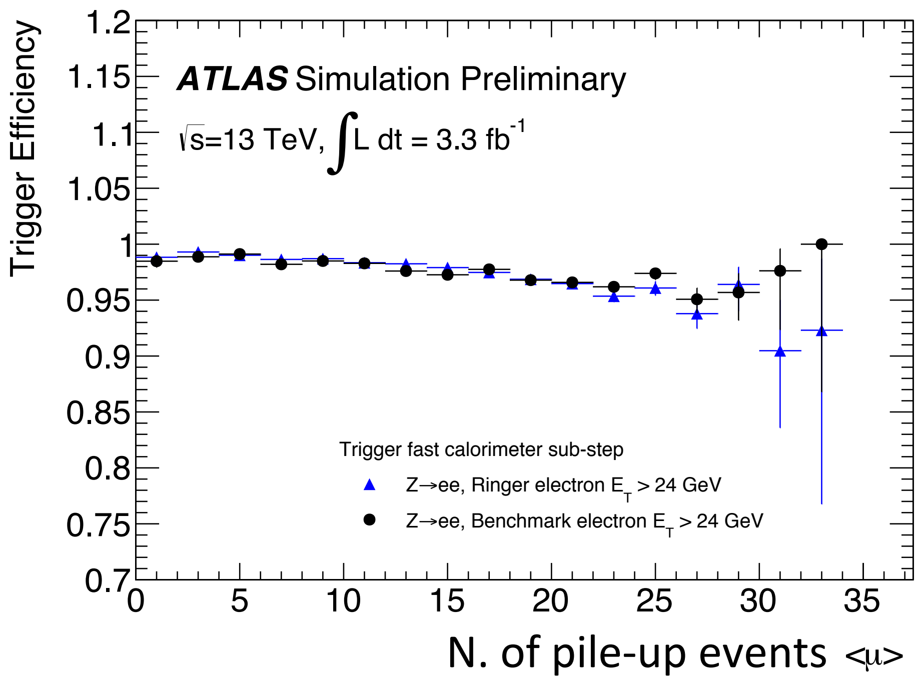

Benchmark chain
Ringer chain
Ringer binned
Hypothesis
Neural
network
Neural
Network
RingerFex
threshold
Accept?
Neural
Network
eta/Et
cluster information
eta/Et
Bin selection

Ringer binned Hypothesis



- 3 bins of E_T:
  [0, 30, 50, ∞[
- 4 bins of η (pseudorapidity) region:
  [0, 0.8, 1.37, 1.54, 2.5]

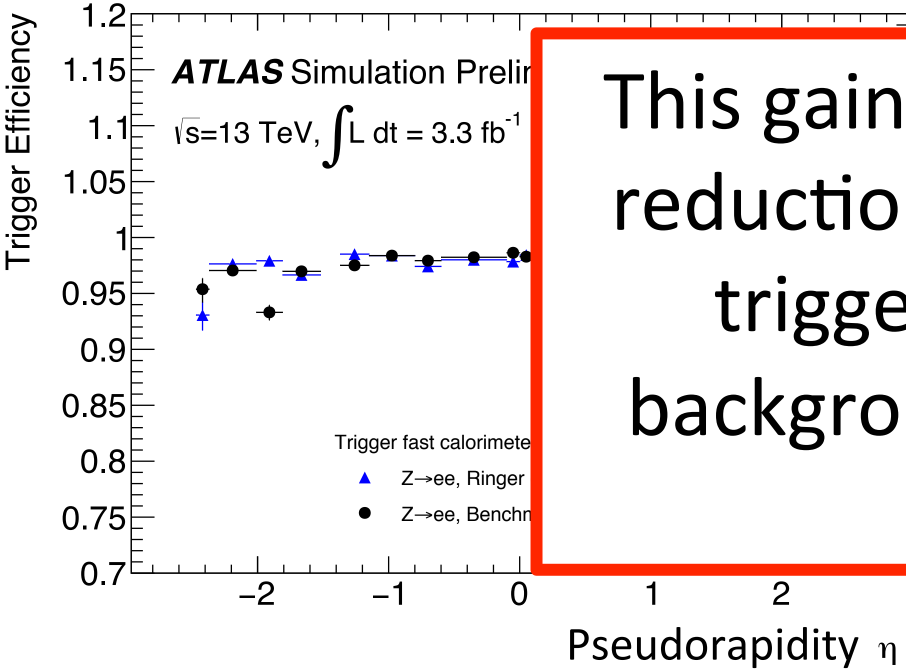




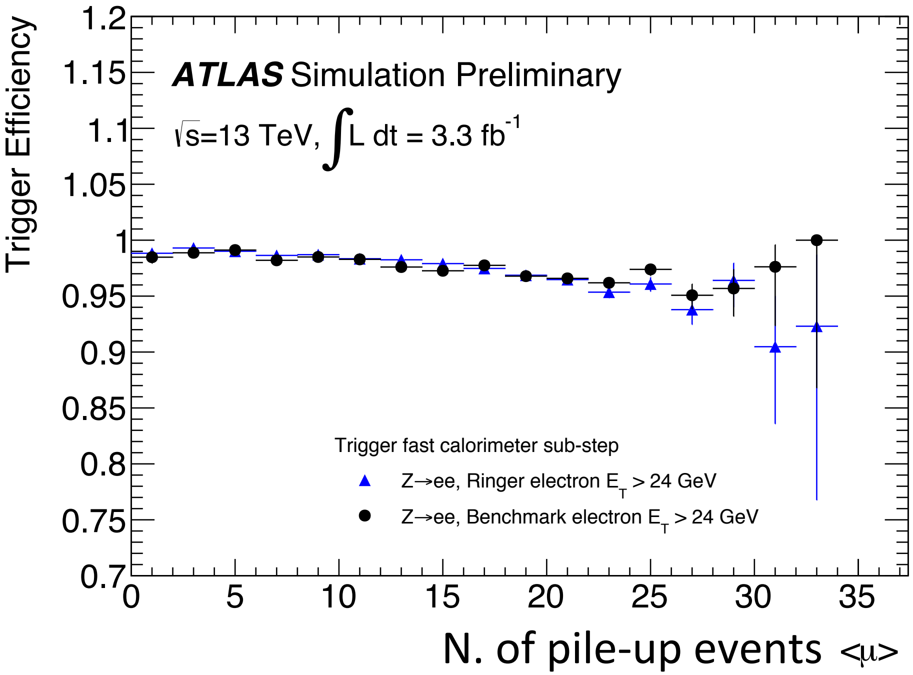
Benchmark: L2Calo@e24_lhmedium_L1EM20VH

Medium (P_D) T&P efficiency

	T&P PD%	FR%
Ringer	97.78	5.37
L2Calo	97.66	12.73



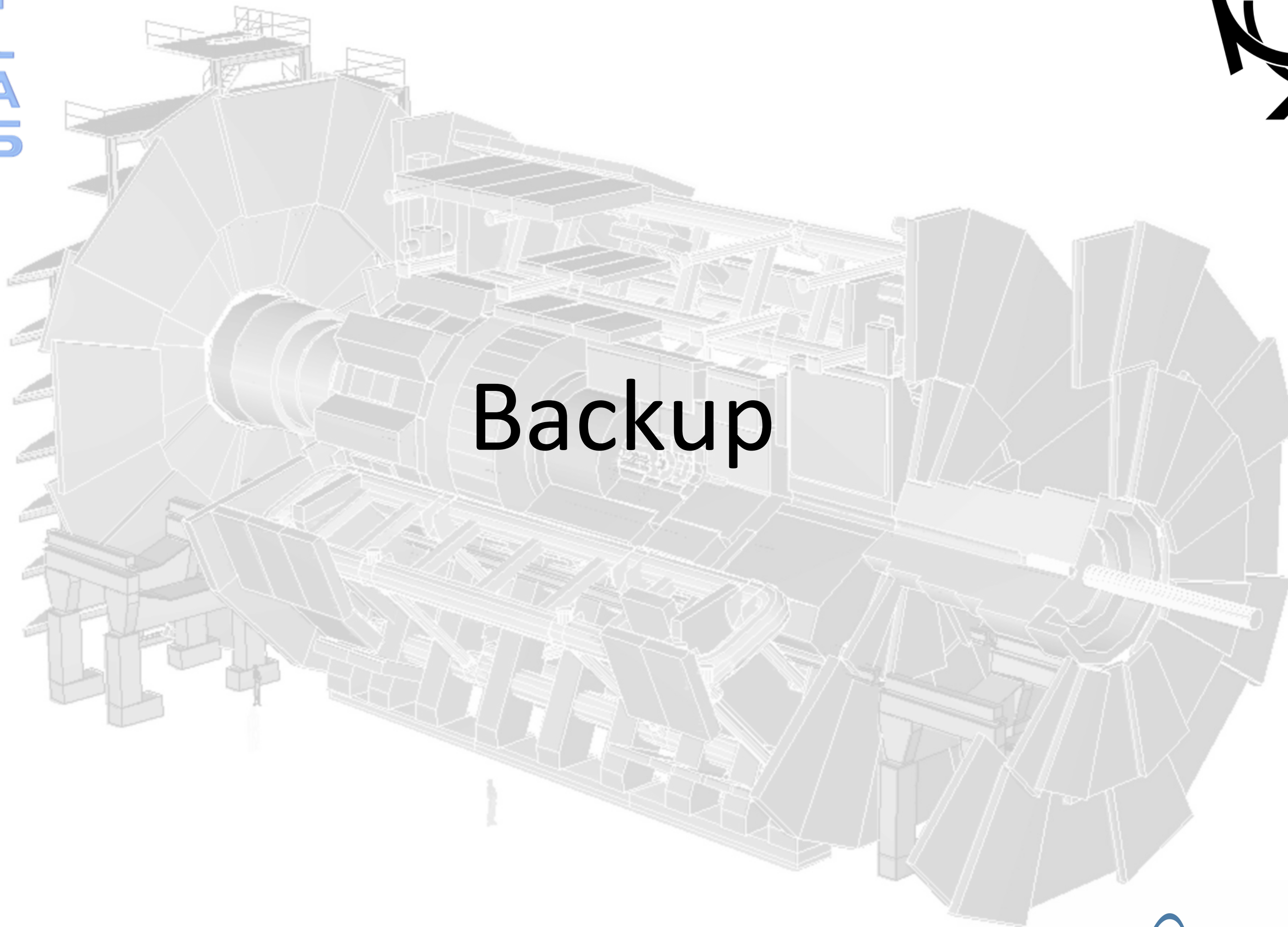
This gain in the fake rate means a reduction of a factor of ~2 in the trigger rate due to the high background/signal cross-section ratio.



	T&P PD%	FR%
Ringer	97.78	5.37
L2Calo	97.66	12.73

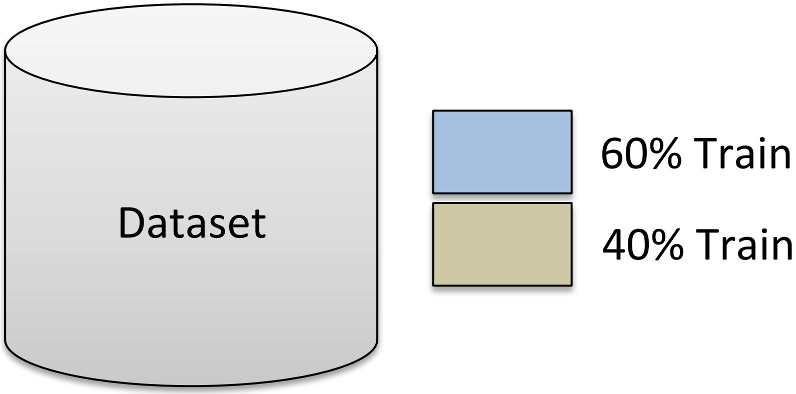
Summary and ongoing work

- Implemented Trigger Hypothesis algorithm:
  - Hypothesis verification in bins and candidate E_T for more stable performance;
- Superior performance as compared to cut-based algorithm:
  - Factor of ~2 reduction of the fake rate.
- Implementation and studies on complete trigger chains in progress;
- A complete Ringer Offline version was implemented and its analysis is on-going.

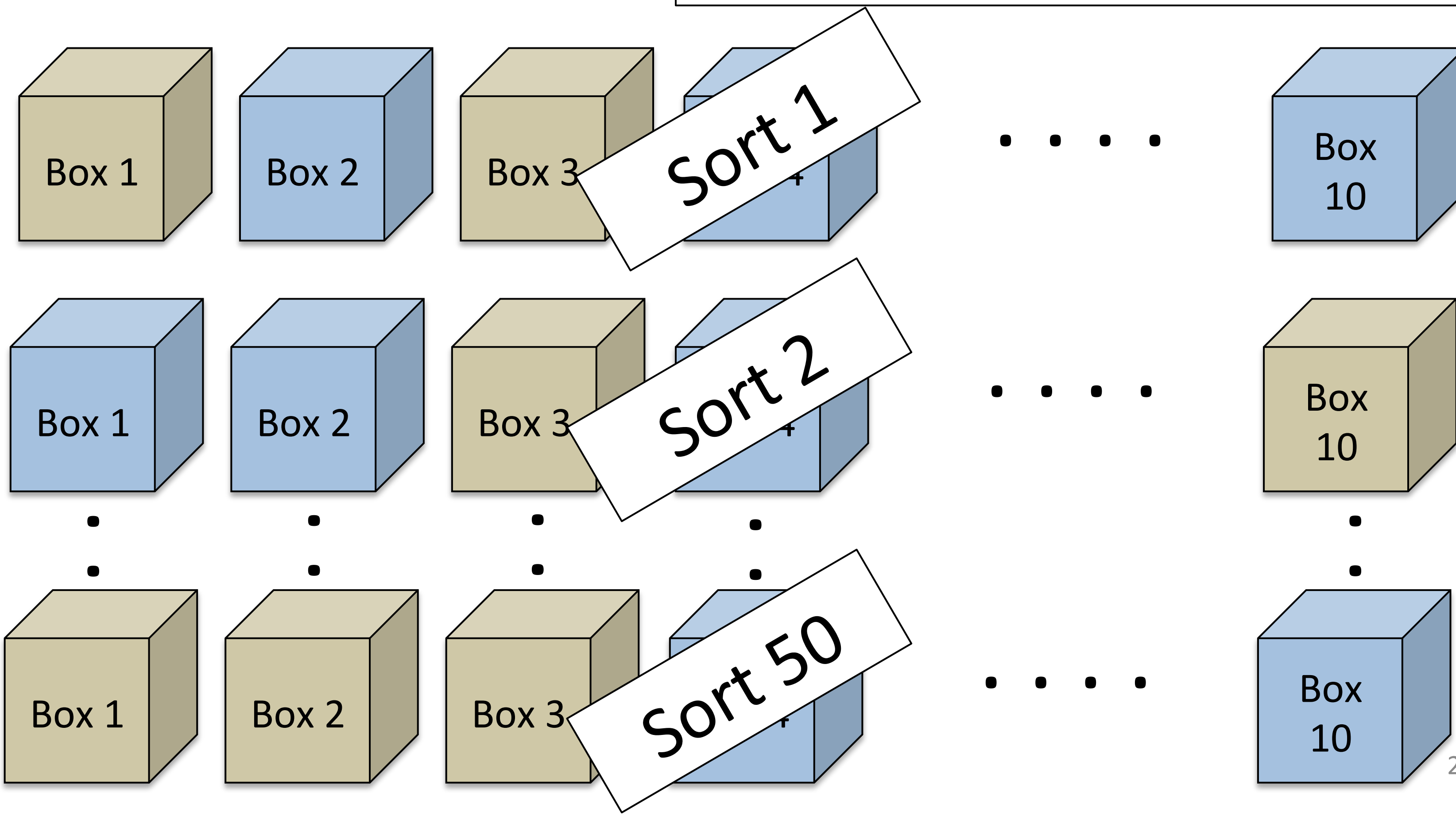


Backup

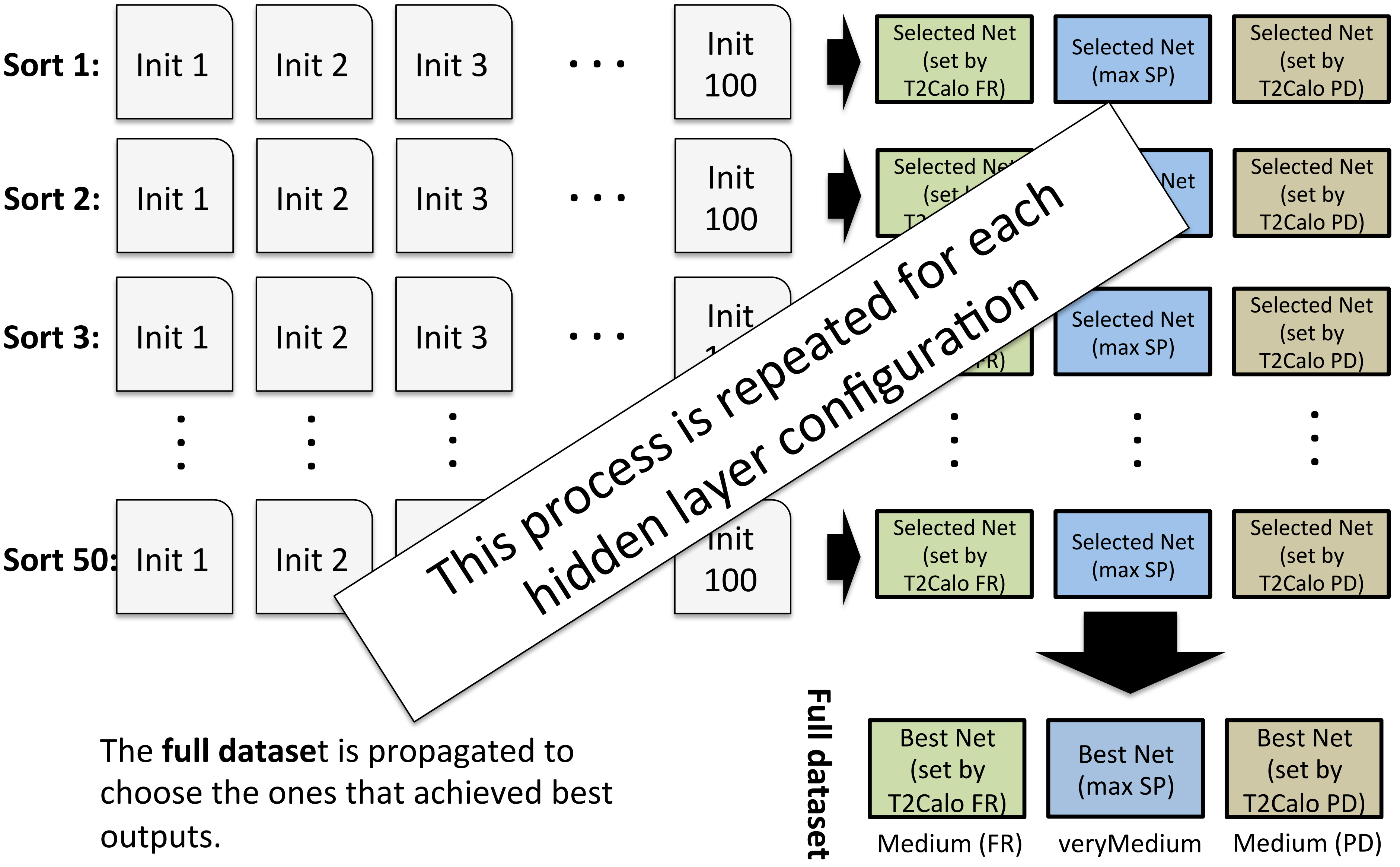
Train Method



Cross Validation (k-fold): 50 sorts using dataset uniformly spread within 10 subsets, usually referred to as boxes. Here, in the particular case: 60% for training phase (model development) and 40% for test. (statistical fluctuations estimation)



Selection Method



The **full dataset** is propagated to choose the ones that achieved best outputs.

Ringer operating points:

"same" Benchmark P_D
"same" Benchmark FR

Et(GeV) ≤ 30

Benchmark: L2Calo@e24_lhmedium_L1EM20VH

Ringer

|eta| < 0.8	
PD (%)	FR (%)
97.0+-0.1	4.4+-0.3 ↓57.9%
99.5-0.1 ↑2.4	8.6+-0.1
97.1	8.6

0.8 < |eta| <= 1.37	
PD (%)	FR (%)
96.9+-0.1	4.6+-0.2 ↓53.2%
99.7+-0.1 ↑2.8	9.8+-0.1
96.9	9.8

1.37 < |eta| <= 1.54	
PD (%)	FR (%)
90.7+-0.0	6.2+-1.2 ↓86.6%
99.7+-0.2 ↑9.1	46.4+-0.2
90.7	46.5

1.54 < |eta| <= 2.5	
PD (%)	FR (%)
95.1+-0.1	10.9+-0.6 ↓38.0%
99.2+-0.1 ↑4.1	17.6+-0.0
95.1	17.6

Benchmark

Ringer operating points:

“same” Benchmark P_D
“same” Benchmark FR

Et(GeV) ≤ 30

Benchmark: L2Calo@e24_lhmedium_L1EM20VH

|eta| < 0.8	
PD (%)	FR (%)
97.0+-0.1	4.4+-0.3 ↓57.9%
99.5-0.1 ↑2.4	8.6+-0.1
97.1	8.6

0.8 < |eta| <= 1.37	
PD (%)	FR (%)
96.9+-0.1	4.6+-0.2 ↓53.2%
99.7+-0.1 ↑2.8	9.8+-0.1
96.9	9.8

1.37 < |eta| <= 1.54	
PD (%)	FR (%)
90.7+-0.0	6.2+-1.2 ↓86.6%
99.7+-0.2 ↑9.1	46.4+-0.2
90.7	46.5

1.54 < |eta| <= 2.5	
PD (%)	FR (%)
95.1+-0.1	10.9+-0.6 ↓38.0%
99.2+-0.1 ↑4.1	17.6+-0.0
95.1	17.6

Ringer operating points:

“same” Benchmark P_D
“same” Benchmark FR

30<Et(GeV)≤50

Benchmark: L2Calo@e24_lhmedium_L1EM20VH

|eta| < 0.8	
PD (%)	FR (%)
99.0+-0.1	5.5+-0.7 ↓54.2%
99.9+-0.0 ↑0.9	11.9+-0.1
99.0	11.9

0.8 < |eta| <= 1.37	
PD (%)	FR (%)
99.1+-0.1	6.3+-0.9 ↓47.5%
99.9+-0.0 ↑0.7	12.0+-0.1
99.2	11.9

1.37 < |eta| <= 1.54	
PD (%)	FR (%)
95.9+-0.1	9.5+-2.1 ↓81.7%
99.9+-0.1 ↑4.0	50.9+-0.2
95.9	50.9

1.54 < |eta| <= 2.5	
PD (%)	FR (%)
97.5+-0.1	11.2+-0.9 ↓51.1%
99.9+-0.1 ↑2.3	22.9+-0.1
97.6	22.8

Ringer operating points:

“same” Benchmark P_D
“same” Benchmark FR

50<Et(GeV)<∞

Benchmark: L2Calo@e24_lhmedium_L1EM20VH

|eta| < 0.8	
PD (%)	FR (%)
99.4+-0.1	2.0+-0.4 ↓72.4%
99.9+-0.1 ↑0.5	7.4+-0.1
99.4	7.4

0.8 < |eta| <= 1.37	
PD (%)	FR (%)
99.7+-0.1	5.2+-1.0 ↓53.8%
99.9+-0.1 ↑0.2	11.3+-0.1
99.7	11.3

1.37 < |eta| <= 1.54	
PD (%)	FR (%)
96.7+-0.1	2.7+-1.9 ↓94.0%
99.9+-0.1 ↑3.2	44.3+-0.6
96.7	44.6

1.54 < |eta| <= 2.5	
PD (%)	FR (%)
98.4+-0.1	4.7+-0.9 ↓67.0%
99.9+-0.1 ↑1.5	14.2+-0.1
98.5	14.1