Deep Neural Network Algorithms in the CMS Level-1 Trigger

Anthony Vizcaíno Aportela on behalf of the CMS Level-1 trigger group

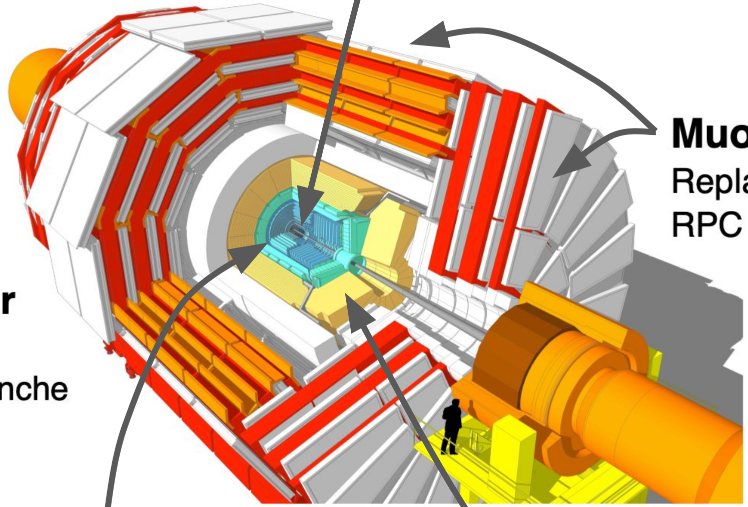
Basic Idea

- The high-luminosity upgrade of the CMS Level-1 trigger has enabled new capabilities,
  - use of inputs from the outer tracker
  - high-granularity calorimeter
- With richer inputs means more ways of applying machine learning to better discriminate, estimate, and learn the physics of each event.

High Luminosity Upgrade (a Problem)

- Each event produces ~1MB of data
- At 40 MHz, data rate becomes astronomical
- We need data reduction at a microsecond latency
- These are what the triggers are for
- We need to know what signatures correspond to particles, and reduce
- FPGAs provide the latency and flexibility
- ML algos for solving this and have potential improve on algos we have

Phase-II Upgrades of the CMS Detector



Trigger/HLT/DAQ

Level-1 trigger: 12.5 µs latency, 750kHz out

HLT output: 7.5kHz

New tracker

Increased granularity, coverage to |η|<4

40 MHz readout (p_T>2 GeV) outer tracker

Muon systems

Replace DT&CSC FE/BE

RPC coverage from 1.5<|η|<2.4

MIP Timing detector

Barrel: Crystal+SiPM

Endcap: Low-gain avalanche diode

Barrel ECal/HCal

Replace FE/BE electronics

ECal single-crystal granularity @40MHz

New endcap calorimeters

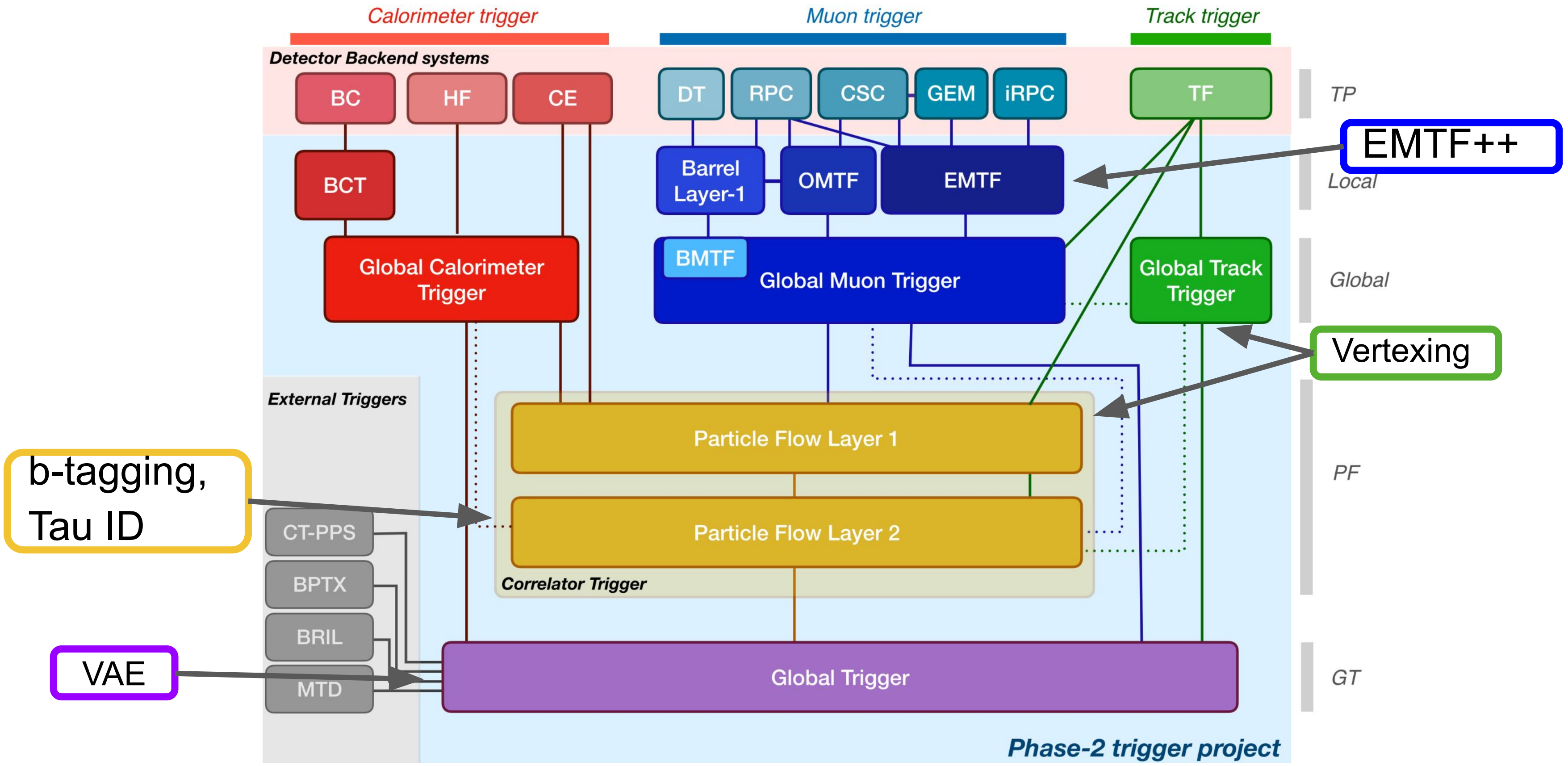
3d showers, precise timing

Si/Scint+SIPM in Pb/W

Recent ML developments CMS

1. Muon Transverse Momentum Estimation
2. b-tagging
3. End-to-End Vertexing NN
4. Tau Identification

Architecture of the Phase-II L1 Trigger



1. Muon Transverse Momentum Estimation

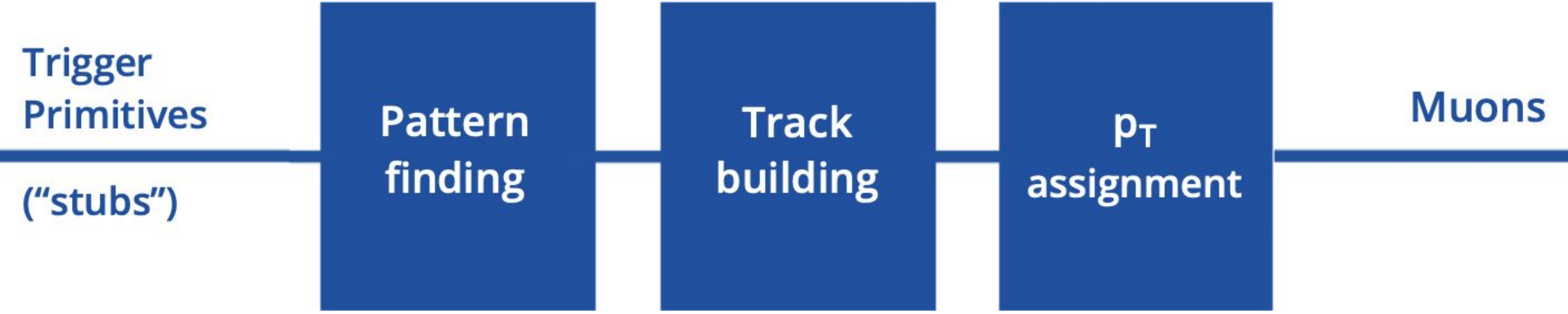
Endcap Muon Track Finder (EMTF)

- Was hard coded into ~1GB LUT (Bad)

Upgrade (EMTF++)

- Utilizing more advanced ML

1. EMTF++



Trigger Primitives

(“stubs”)

Pattern finding

Track building

p_T assignment

Muons

from:
- CSC
- RPC
- GEM
- iRPC
- ME0

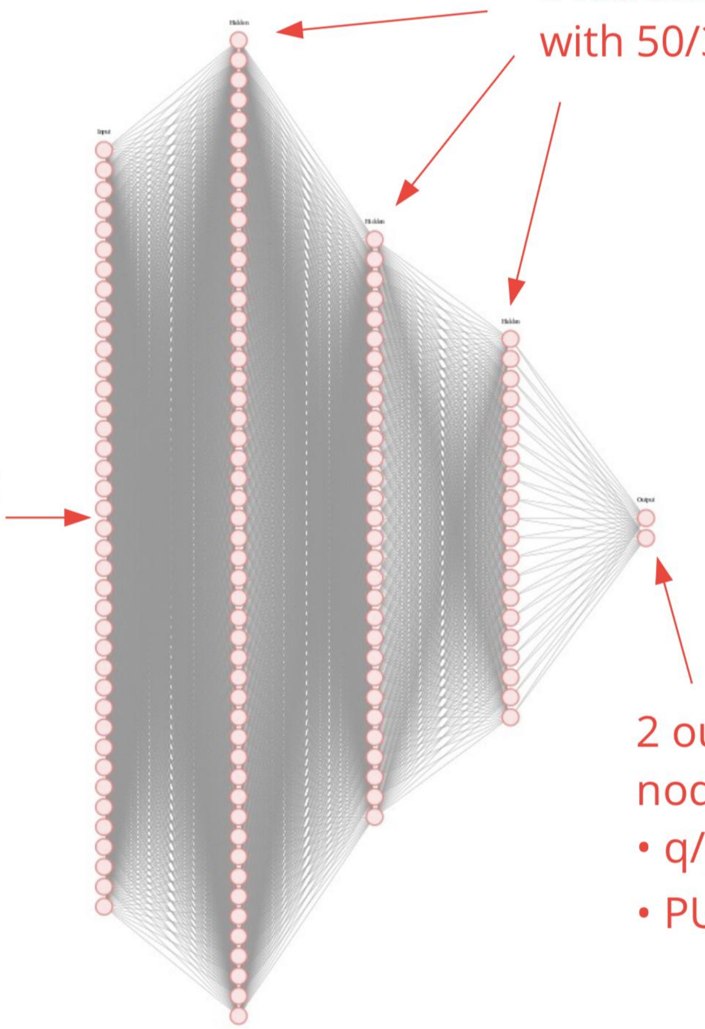
Use patterns to find stubs in stations

Build track candidates by selecting a unique stub from each muon station.

Use ML to determine the pT from the track candidate

1. EMTF++

- Dense network
- 39 inputs muon system params – phi, bend, etc
- 2 outputs, q/pT and pileup discriminator



3 hidden layers with 50/30/20 nodes

39 input nodes

2 output nodes:

- q/p_T
- PU discr

2. b-tagging

- **L1** will include **tracking** and a higher granularity calorimeter (**HGcal**)
- Including **tracking makes it possible** to develop algorithms to identify jets from bottom quarks (**b-tagging**) **in the L1** stage
- Taking advantage of new systems, CMS will implement special versions of PF and PUPPI algorithms in L1.
- The particles reconstructed using the PUPPI algorithm are used to construct higher level objects, such as jets.
- **Neural network** is implementable on **current trigger hardware**
- Runs on PUPPI particles and **discriminates** between **b-quark jets** and light quark / gluon jets

2. B-tagging algo

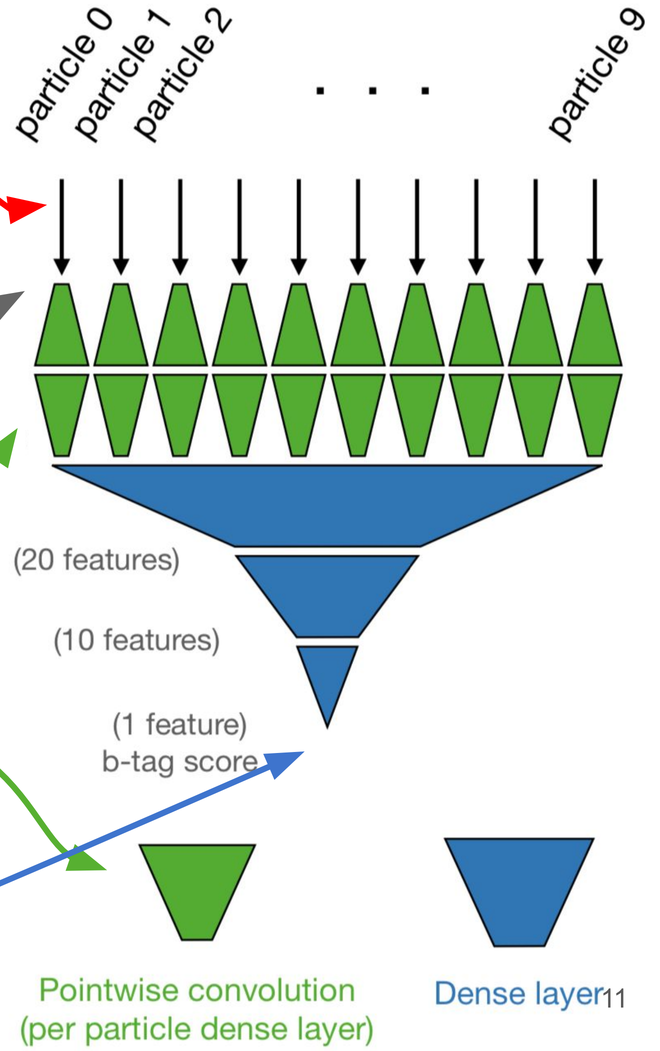
Inputs are top ten PUPPI candidates ranked by jet p_T

Input features

$[b_0,b_1,b_2,b_3,b_4,b_5,b_6,b_7]$	Particle type, one-hot encoded
$[p_T, \eta, \varphi]$	kinematic information scaled relative to jet
$[z_0, d_0]$	z-position and transverse impact parameter with respect to the primary vertex

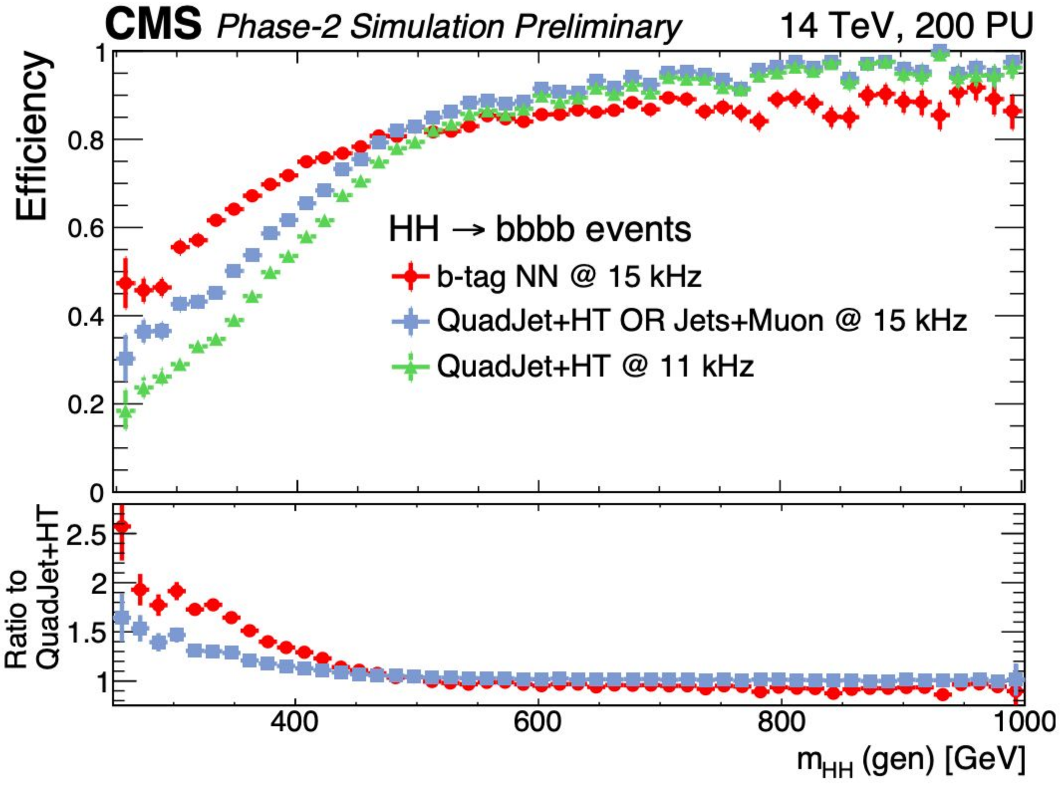
Featurizers

Scores close to 1 indicate jets that are likely to have originated from bottom quarks, while scores close to 0 indicate jets that are likely to have originated from light quarks or gluons.



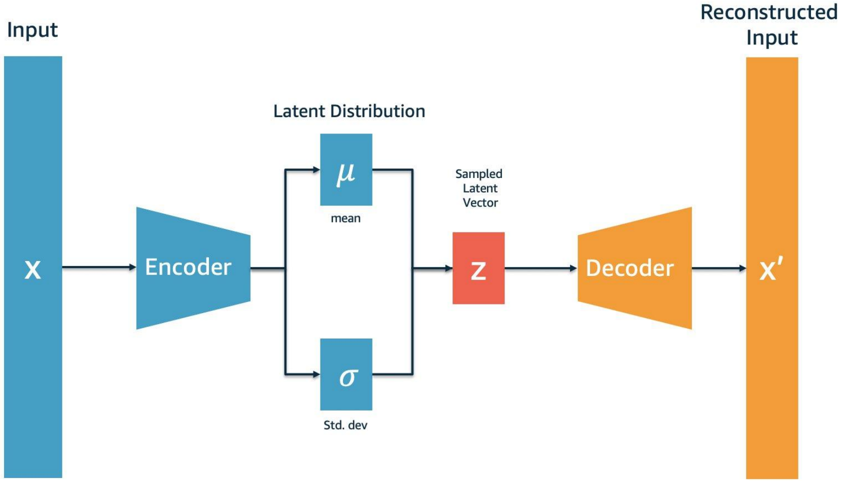
2. Prelim b-tagging results

A b-tag CNN can recover inefficiencies at low m_{HH} seen by conventional strategies



3. Autoencoder for Anomaly Detection

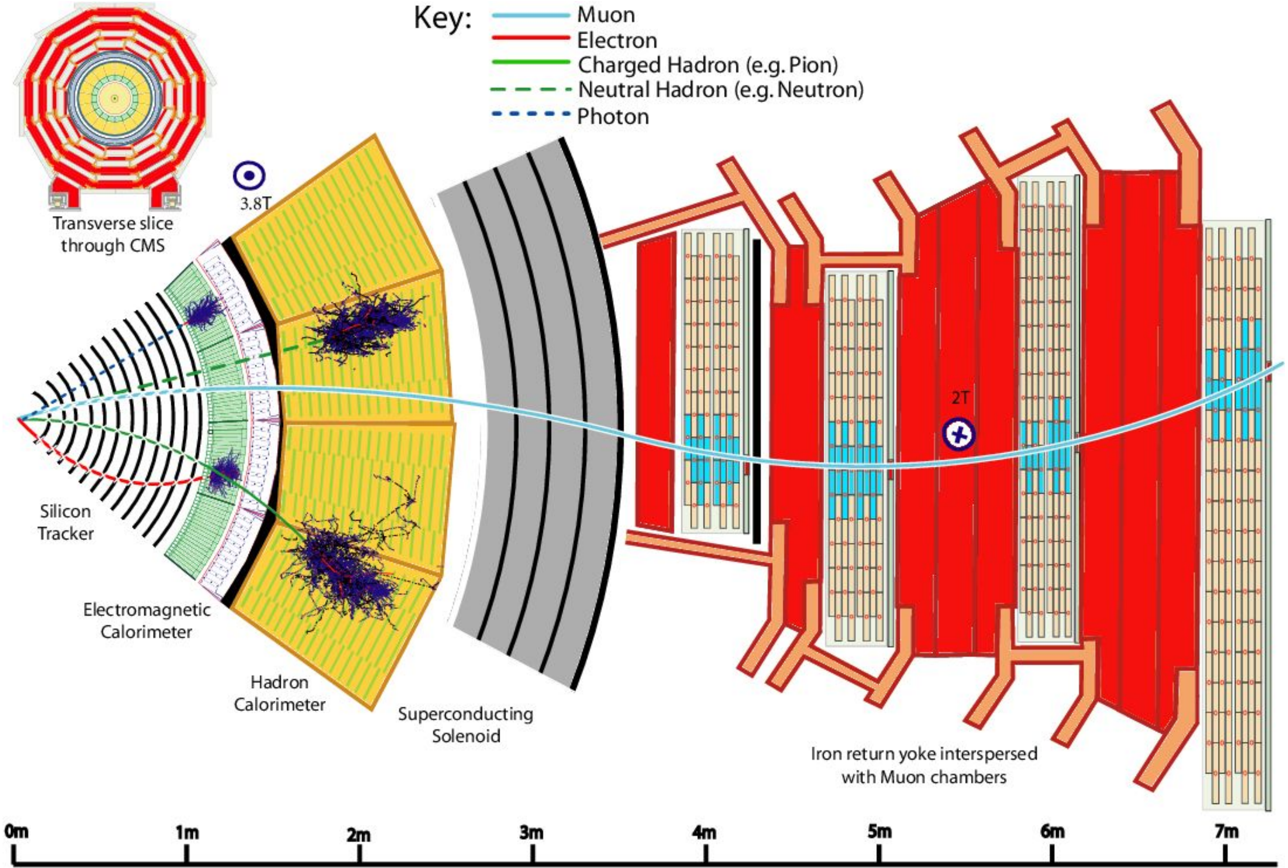
- Possible that new physics events, if any, are systematically rejected by the CMS trigger system
- Requires departure from the usual fully supervised search scheme
- A strategy would be to pose as an anomaly detection task in real time



3. Autoencoder for Anomaly Detection

- The input features consist of (p_T, f) of PUPPI E_{miss}, PUPPI H_T, and (p_T, h, f) of 5 PUPPI jets, for a total of 18 input variables
- The AE is constructed as a multilayer perceptron (MLP) consisting of 8 layers
- Each layer uses ReLU activation function
- objective function is norm$_1$ distance
- Simple functions for use on FPGAs

3. End-to-End Vertexing NN


Key:
Muon
Electron
Charged Hadron (e.g. Pion)
Neutral Hadron (e.g. Neutron)
Photon
Transverse slice through CMS
3.8T
2T
Silicon Tracker
Electromagnetic Calorimeter
Hadron Calorimeter
Superconducting Solenoid
Iron return yoke interspersed with Muon chambers
0m
1m
2m
3m
4m
5m
6m
7m

3. End-to-End Vertexing NN

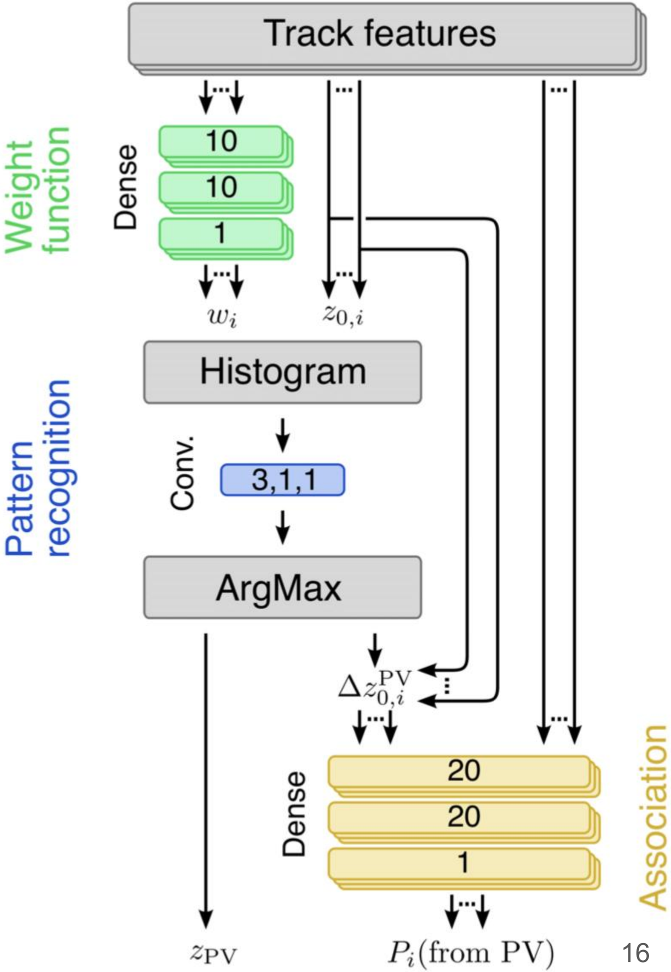
1. Filtering input tracks
2. binning them on z axis
3. performing regression on the histogram to reconstruct vertex z_0 position
4. and finally to associate tracks to the vertex

are combined using “end-to-end ML” as a potential improvement to the vertexing performance.

Uses a 1D CNN to

1. weight input tracks
2. accumulate tracks in histogram

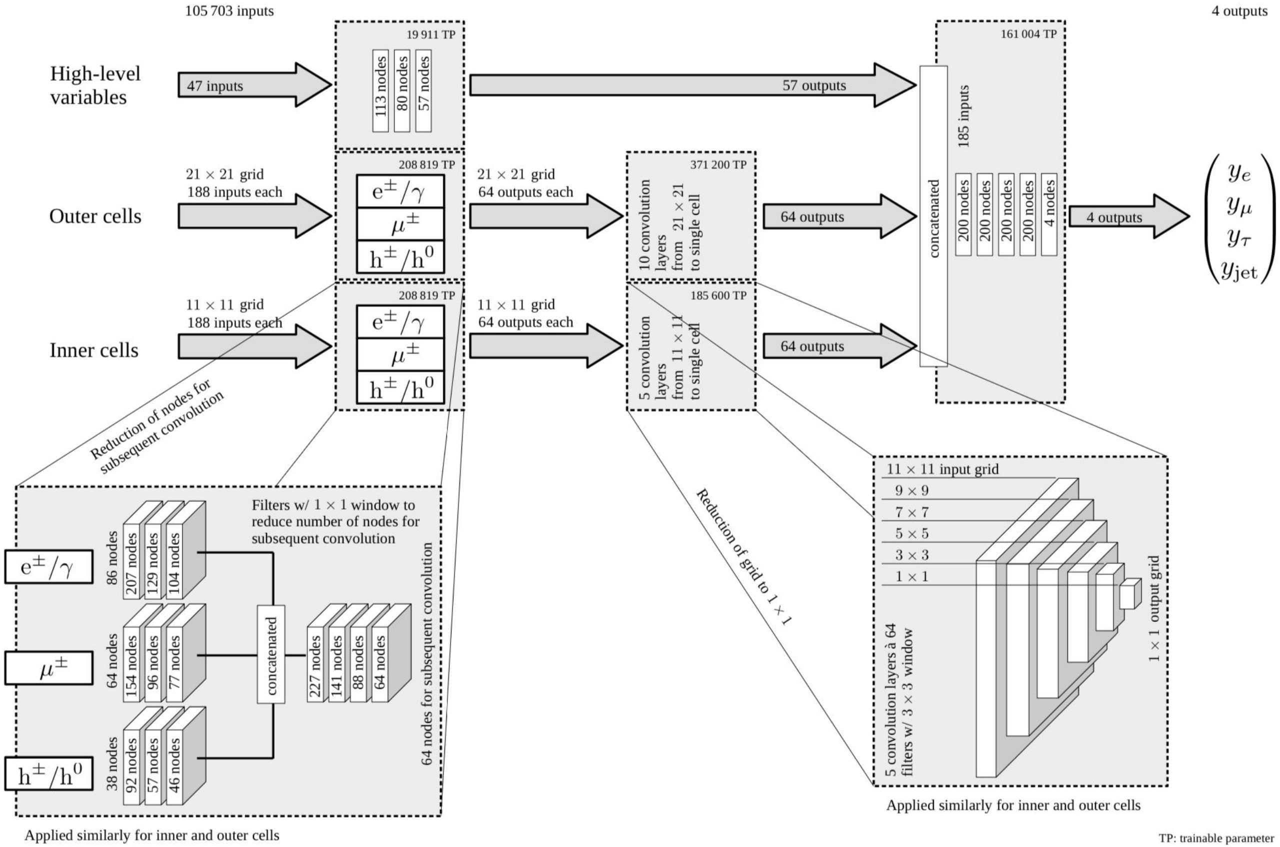
and use another set of convolutional layers to perform peak finding within the histogram



4. Tau ID

- Particle Flow based algorithm
- Iteratively seed from the highest-pT charged particles that have dR > 0.4 from each other.
- For each particle, take all particle-flow or PUPPI candidates within dR < 0.4.
- Of these candidates, take the 10 highest-pT candidates within the cone, compute pT, d_eta_{seed}, d_phi_{seed}, and particle ID (40 inputs total), and input these variables into a dense neural network.
- For all candidates within dR < 0.1 of the seed, assign them to the tau candidate to compute its four-vector

4. Tau ID


105 703 inputs
4 outputs
High-level variables
47 inputs
19 911 TP
113 nodes
80 nodes
57 nodes
57 outputs
161 004 TP
185 inputs
concatenated
200 nodes
200 nodes
200 nodes
200 nodes
4 nodes
4 outputs
y_e
y_μ
y_τ
y_{jet}
Outer cells
21 × 21 grid
188 inputs each
208 819 TP
$e^\pm/\gamma$
$\mu^\pm$
$h^\pm/h^0$
21 × 21 grid
64 outputs each
371 200 TP
10 convolution layers from 21 × 21 to single cell
64 outputs
Inner cells
11 × 11 grid
188 inputs each
208 819 TP
$e^\pm/\gamma$
$\mu^\pm$
$h^\pm/h^0$
11 × 11 grid
64 outputs each
185 600 TP
5 convolution layers from 11 × 11 to single cell
64 outputs
Reduction of nodes for subsequent convolution
$e^\pm/\gamma$
$\mu^\pm$
$h^\pm/h^0$
86 nodes
207 nodes
129 nodes
104 nodes
64 nodes
154 nodes
96 nodes
77 nodes
38 nodes
92 nodes
57 nodes
46 nodes
Filters w/ 1 × 1 window to reduce number of nodes for subsequent convolution
concatenated
227 nodes
141 nodes
88 nodes
64 nodes
64 nodes for subsequent convolution
Applied similarly for inner and outer cells
Reduction of grid to 1 × 1
11 × 11 input grid
9 × 9
7 × 7
5 × 5
3 × 3
1 × 1
5 convolution layers à 64 filters w/ 3 × 3 window
1 × 1 output grid
Applied similarly for inner and outer cells
TP: trainable parameter

At a Glance

Talked about ML techniques in the trigger

1. Muon Transverse Momentum Estimation: In Muon trigger system
2. B-tagging: In Correlator
3. Autoencoder for Anomaly Detection: In global trigger
4. End-to-End Vertexing NN: In correlator and global track trigger
5. Tau Identification: In Correlator