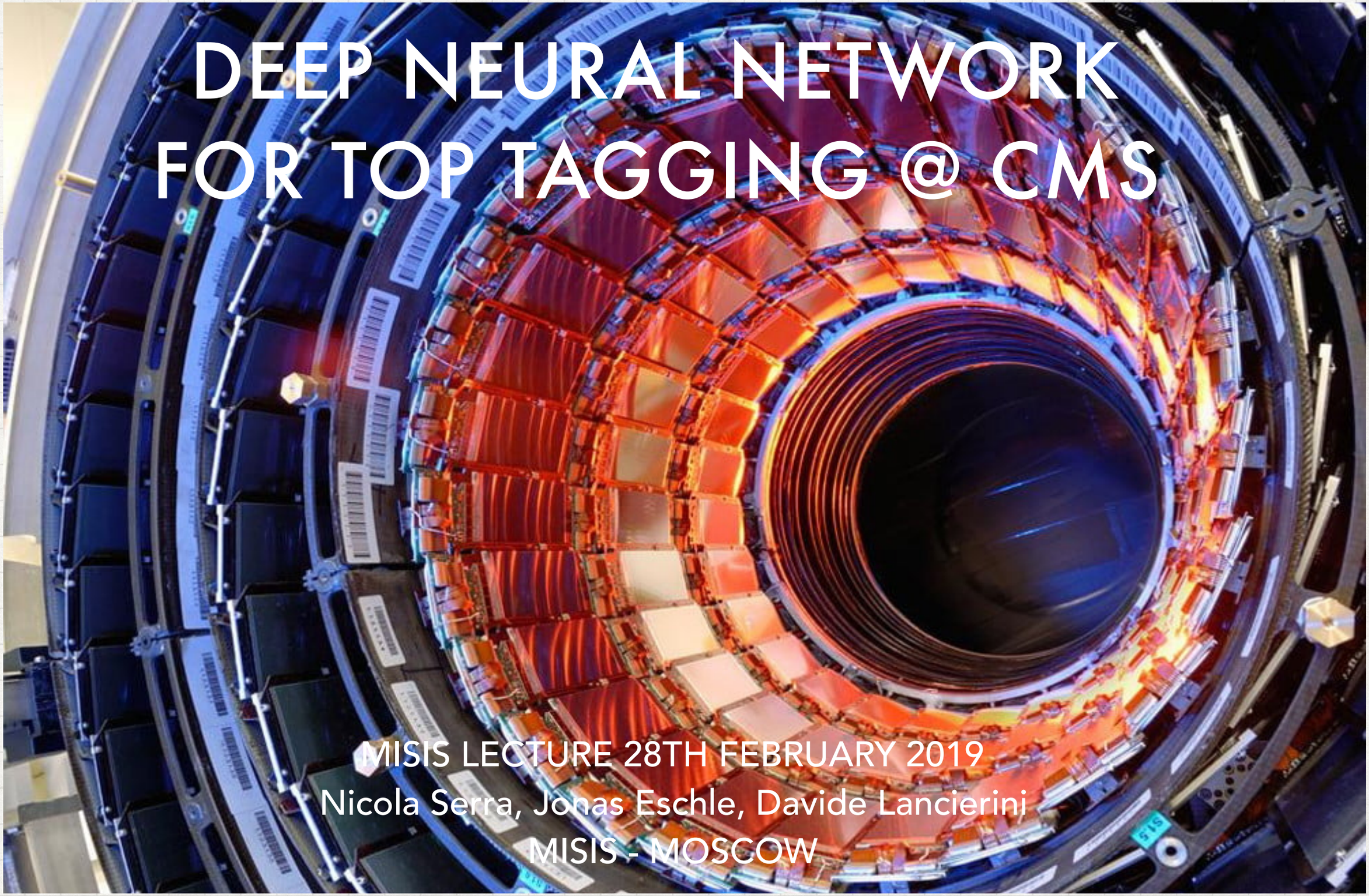
DEEP NEURAL NETWORK FOR TOP TAGGING @ CMS

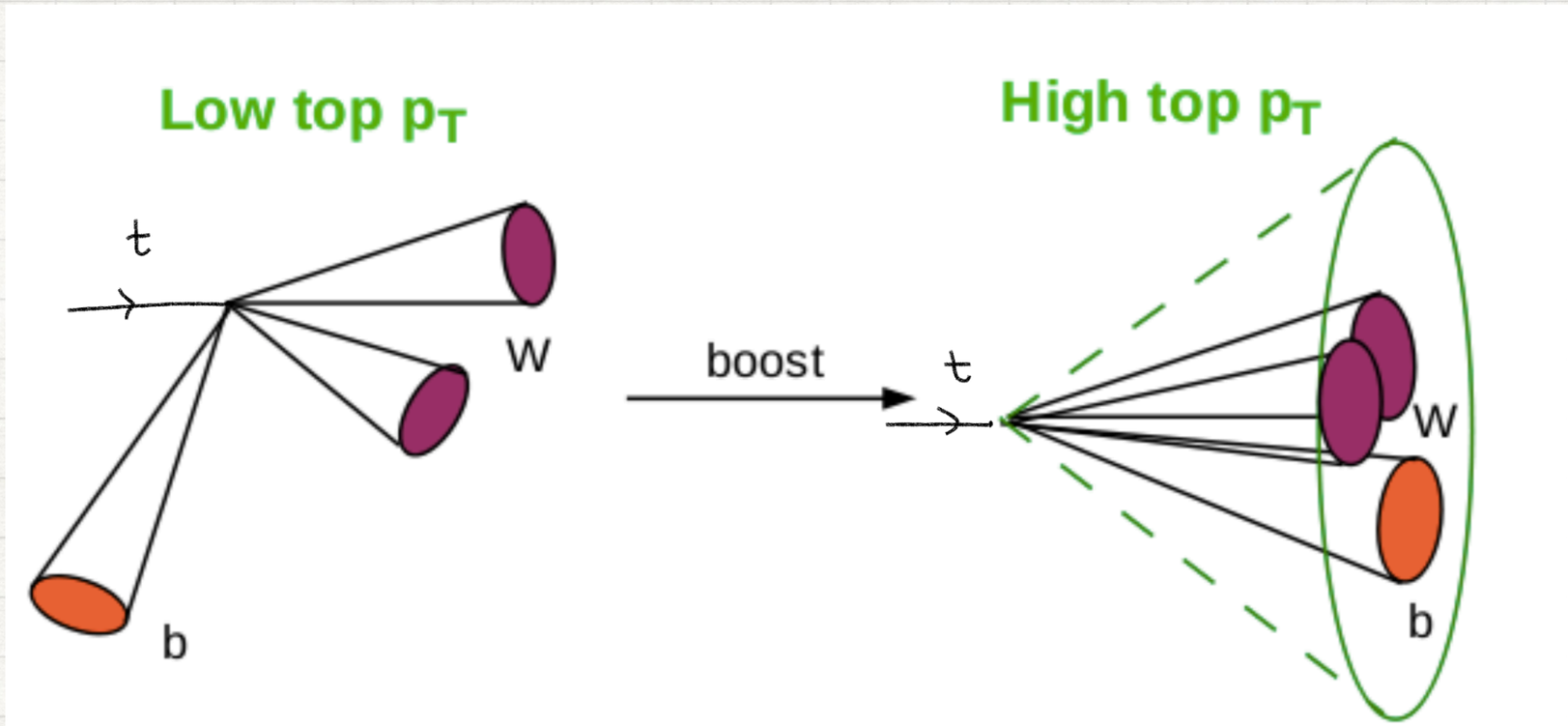


MISIS LECTURE 28TH FEBRUARY 2019

Nicola Serra, Jonas Eschle, Davide Lancierini
MISIS - MOSCOW

Based on the dataset provided by Gregor Kasieczka (gregor.kasieczka@uni-Hamburg.de)

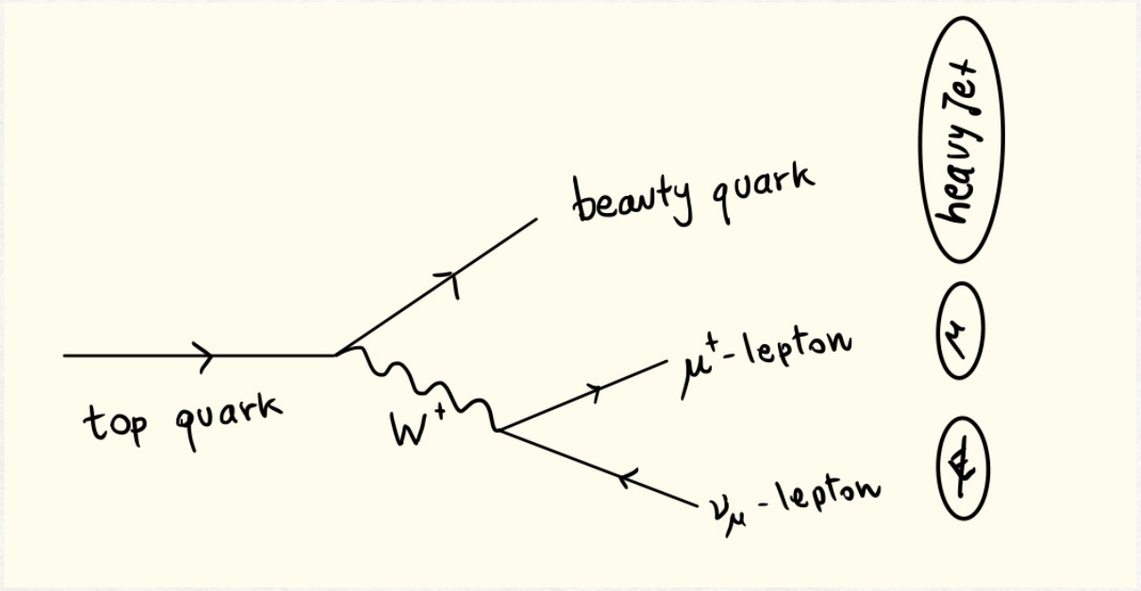
HEAVY RESONANCE TAGGING



- Top/Higgs/W/Z decay hadronically
- When boosted, the decay products are contained in one Jet
- Problem: distinguish between top/H/W/Z (signal) and light quark/ gluon jets (background)

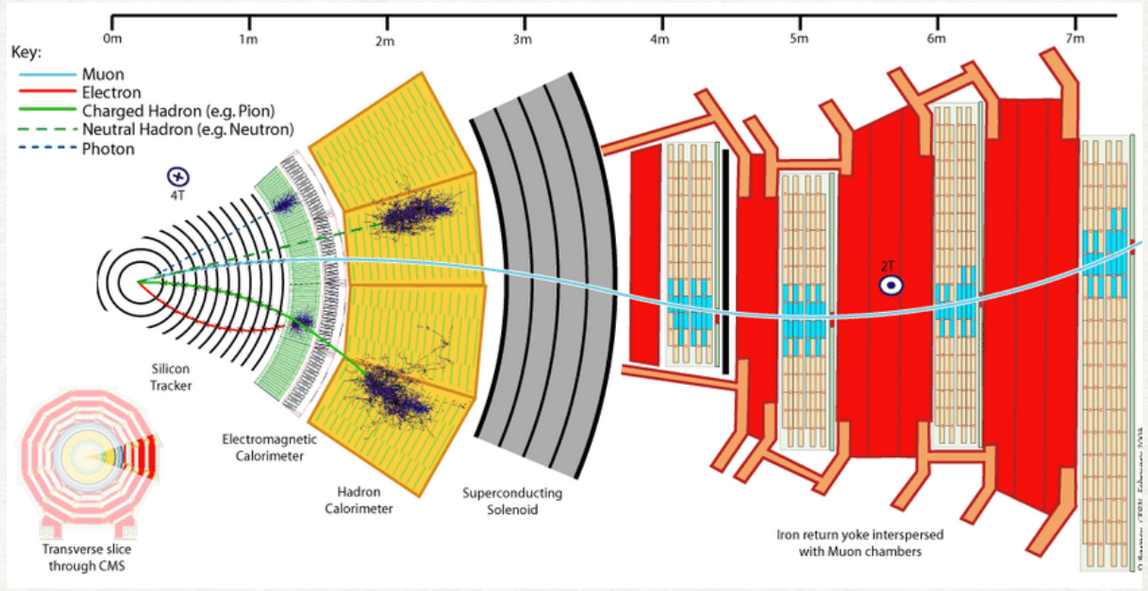
SIGNAL EVENTS

AT THE PARTONIC LEVEL



A signal event at the partonic level

IN THE CMS DETECTOR



https://cds.cern.ch/record/2120661

- Tracking systems use alrgorithms to identify decay
- Energy deposit is measured in the calorimeter
- These energy deposits are clustered into jets

VARIABLES AT HADRON COLLIDERS

- Constituents variables can be cast to new variables (E, η, ϕ)

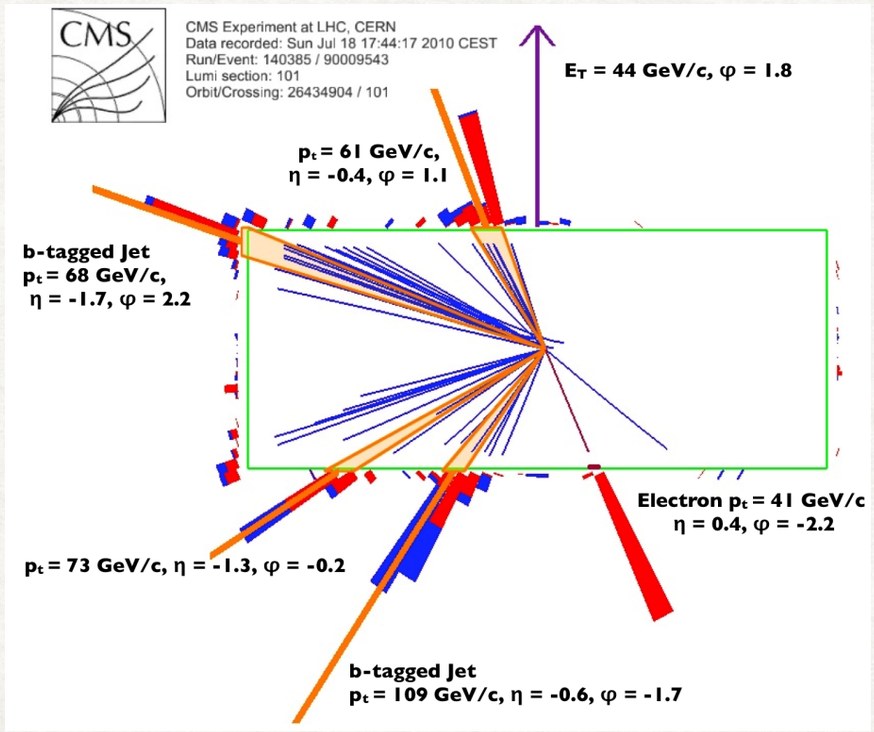
$$(E, PX, PY, PZ)_i \rightarrow (E, \eta, \phi)_i$$

- η is related to the longitudinal momentum PZ

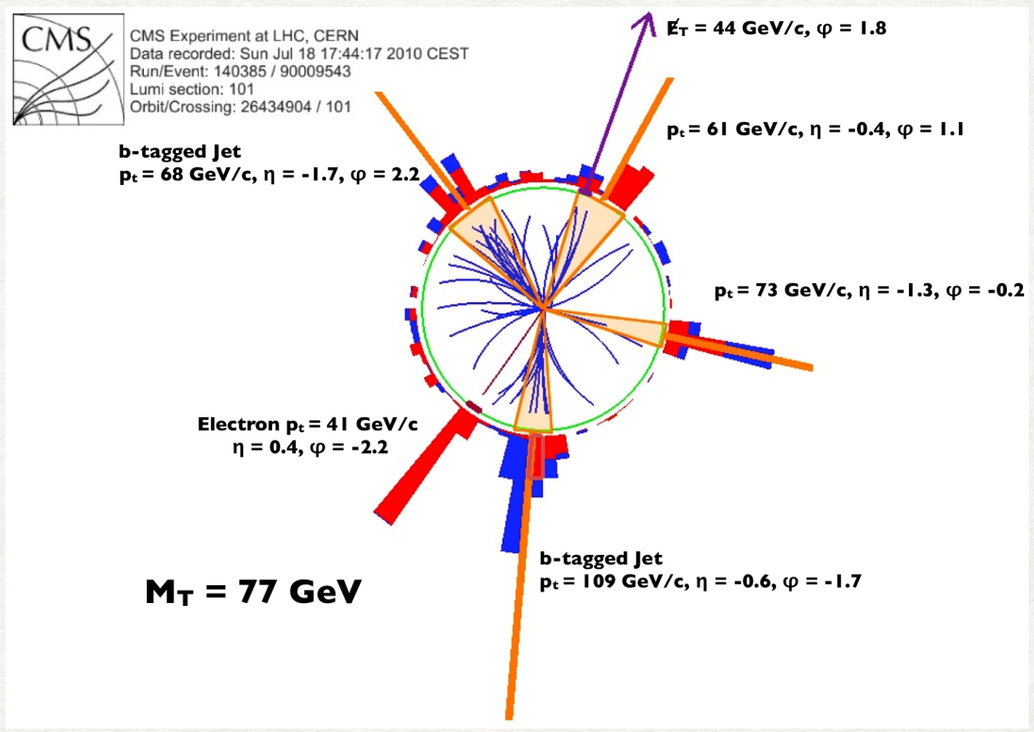
$$\eta = \frac{1}{2}\left(\frac{E - PZ}{E + PZ}\right)$$

- ϕ is related to the transverse momenta (PX, PY)

$$\phi = arctg\left(\frac{PY}{PX}\right)$$



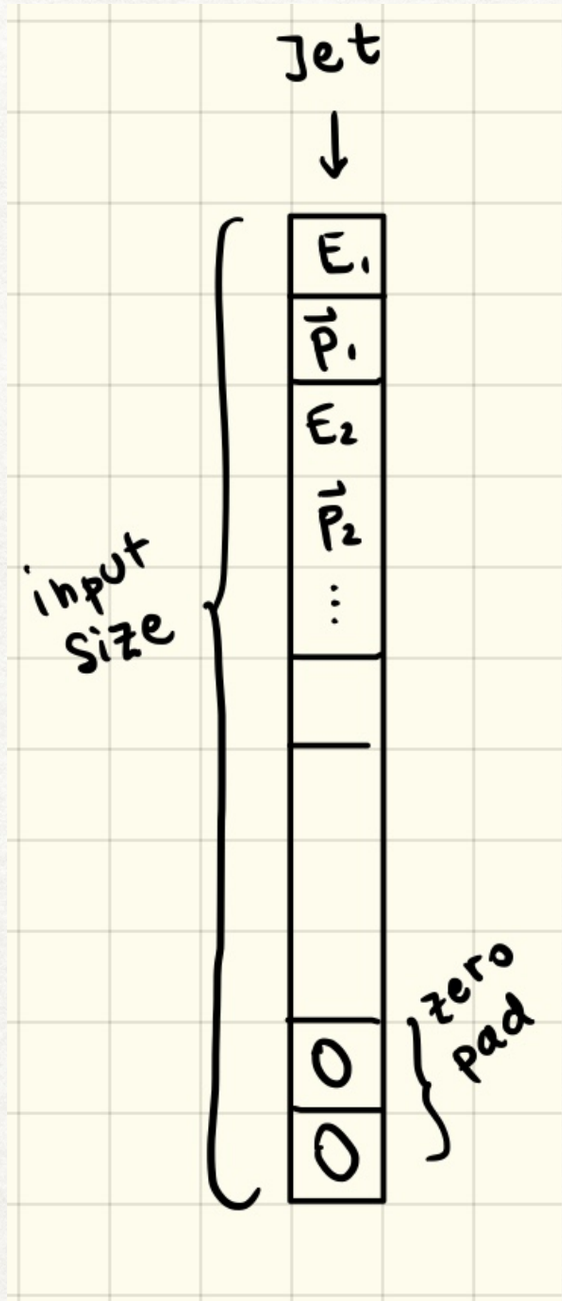
- Side view of the CMS detector



- Front view of the CMS detector

DENSE NETWORK DATASET

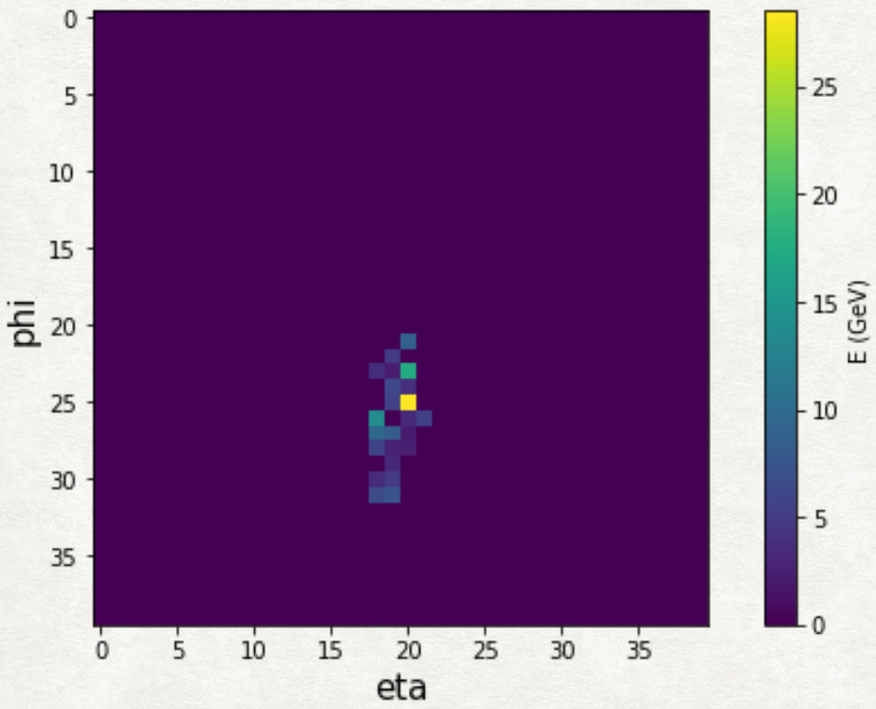
- The input consists of the four-momentum of each jet's constituent (E_i, PX_i, PY_i, PZ_i)
  - The number of constituents per jet is variable $\longrightarrow$ zero-padding for jets with fewer constituents than the input-size
- Data is sorted by decreasing PT
- Train_labels = (0 for Background, 1 for signal)



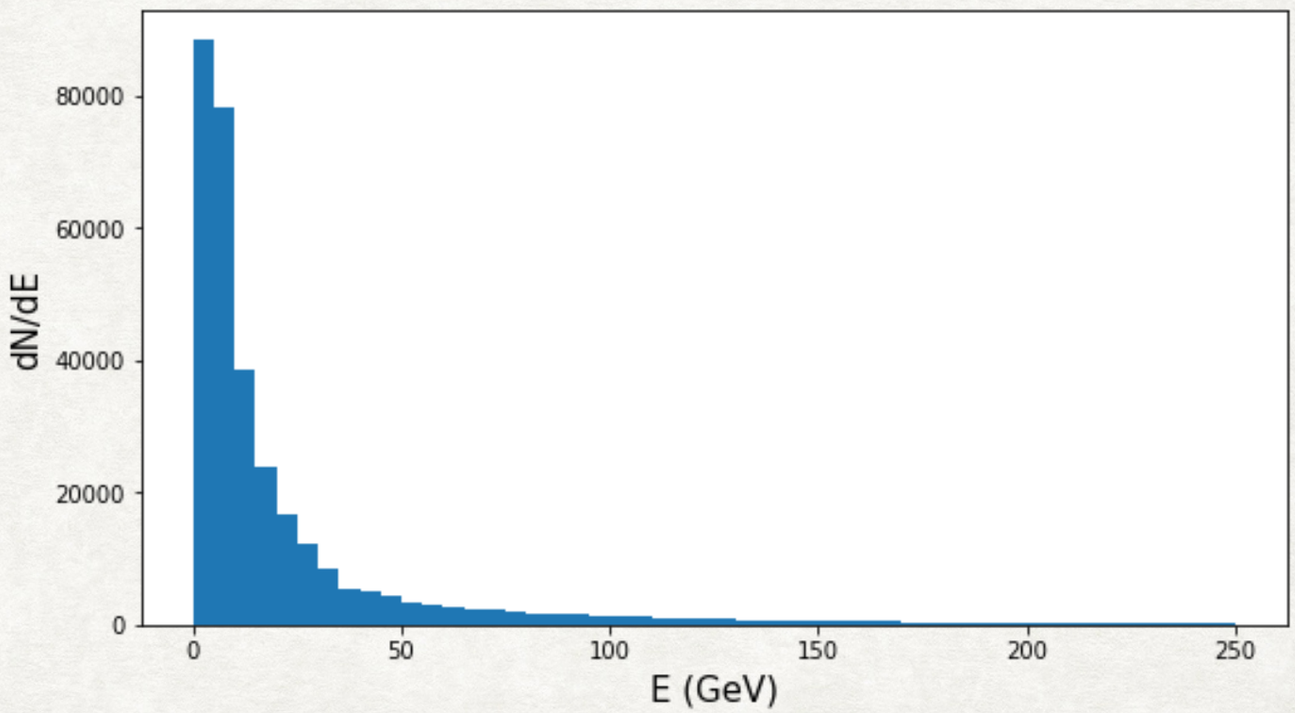
Dense representation of the datset

IMAGES DATASET

- Each event in the dataset is now converted to an image
- Pixels are lit up to the energy recorded by the calorimeter cells
- Use a Convolutional Neural Network approach and compare to the previous dense network performances



- Image representation of an event

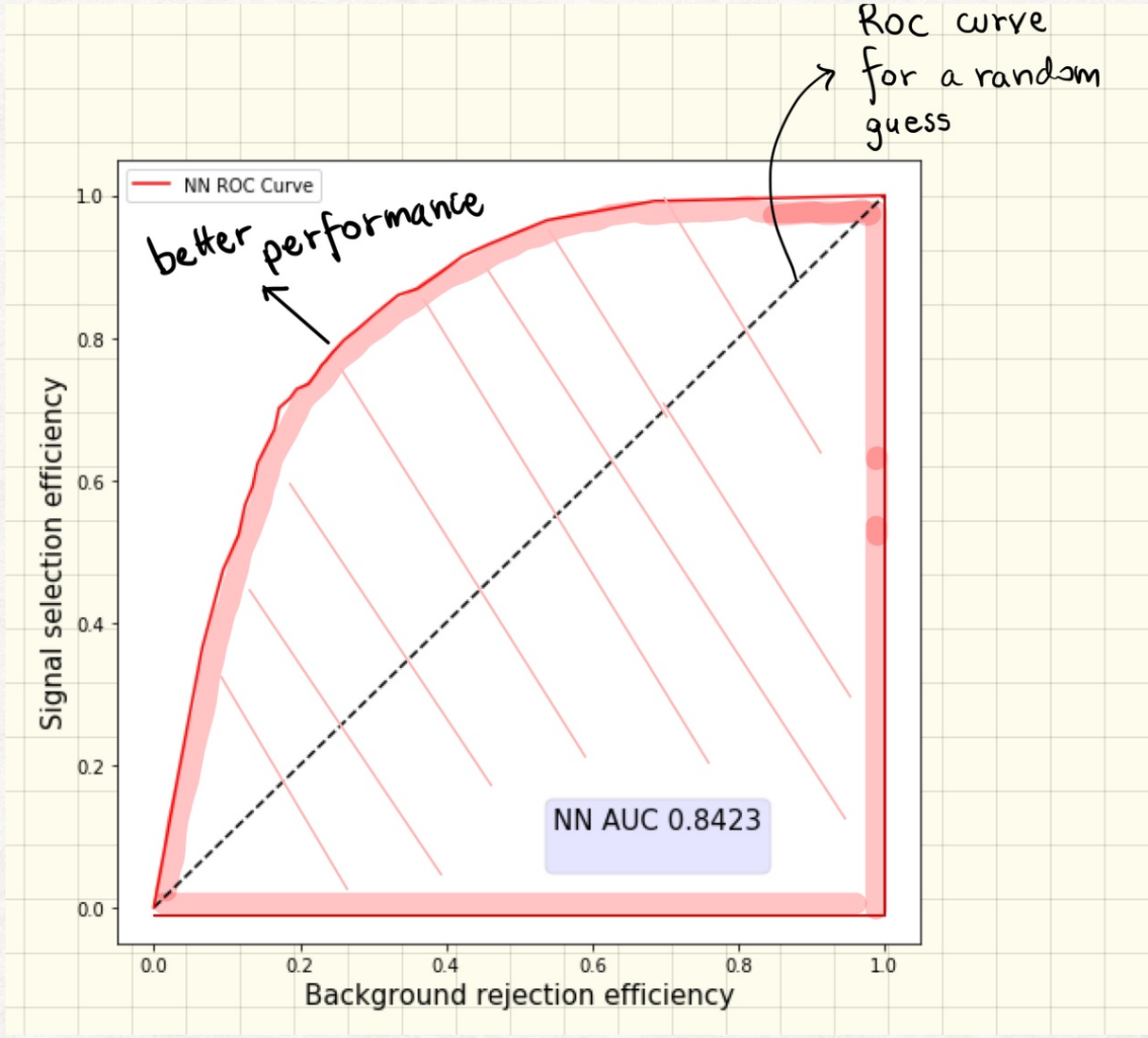


- Energy distribution of the constituents

OBJECTIVE OF THE NOTEBOOKS

- Aim: create a fully connected network to discriminate between jets coming from top decay and QCD jets

- A good network maximises area under ROC curve (AUC)



ROC curve, Area under curve (AUC) plots