LPC CMS DAS:
Tau Leptons

Abdollah Mohammadi, Isobel Ojalvo, Andrew Melo

January 9th 2018

Tau Detection at CMS

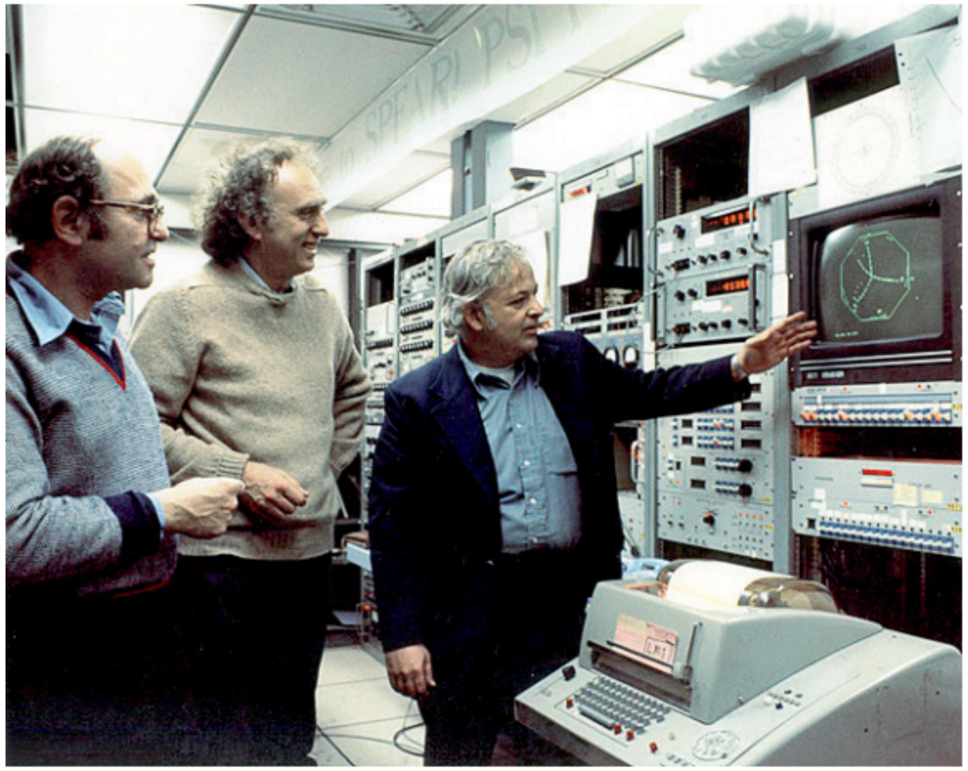
1.) Overview of Tau Identification and Reconstruction Techniques at CMS
2.) Fake Probability and Tau ID Efficiency Definitions
3.) Hands on Example

Goals:
1.) Understand what are the primary challenges one needs to consider when Identifying Taus at CMS
2.) Write a module which will measure various Tau ID Efficiency and other Object Fake Probability
3.) Create a ROC curve!

A Brief History…

Discovered by Martin L. Perl et al. (1974/1975)

- Stanford Positron Electron Accelerating Ring (SPEAR) — at SLAC
    - e+e- collider (CoM 7.4 GeV)
- Discovered via $e^+e^- \rightarrow \tau\,\tau \rightarrow \mu^\pm e^\mp \nu\nu$
- Search for a heavier lepton already underway at ADONE (Italy, 1.5 GeV CoM)
- M.L. Perl recognized the potential for discovery of a heavier lepton using SPEAR



Martin in the SPEAR Control Room in November 1974, following discovery of the J/Psi. Left to right, Gerson Goldhaber (LBL), Martin Perl, and Burton Richter.)

Tau Lepton was first of the **third generation** fermions discovered

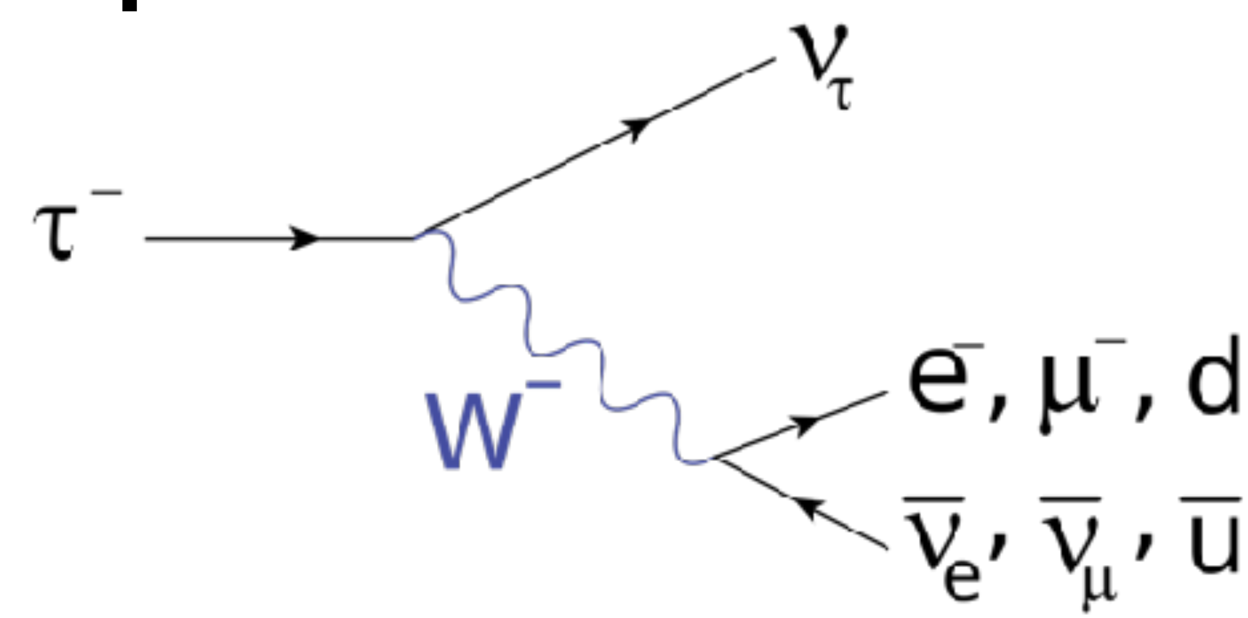
- Followed by the discovery of the **bottom quark** (1977), the **top quark** (1995) and the **tau neutrino** (2000)

http://www.nasonline.org/publications/biographical-memoirs/memoir-pdfs/perl-martin.pdf

Tau Leptons

Third generation lepton

* **charge**: +/-1
* **mass**: 1776.86 ± 0.12 MeV
* **spin**: 1/2 (fermion)
* **mean lifetime**: 2.9×10^{-13} s



Taus Decay Weakly

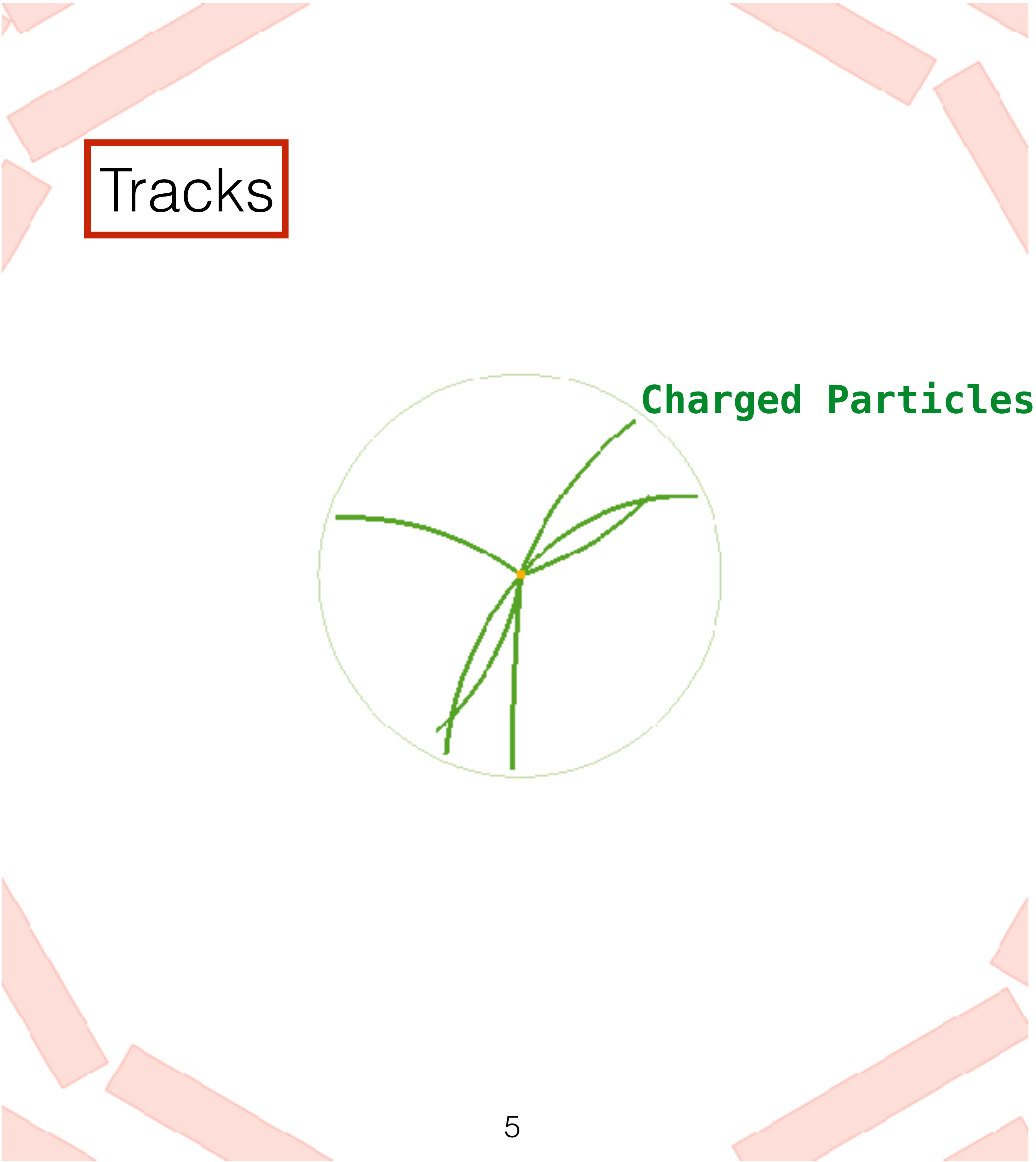
- Leptonic e/muon + 2ν
- Hadronic + ν

Decay Mode	Resonance	$\mathcal{B}$ [%]
$\tau^- \to e^- \bar{\nu}_e \nu_\tau$		17.8
$\tau^- \to \mu^- \bar{\nu}_\mu \nu_\tau$		17.4
$\tau^- \to h^- \nu_\tau$		11.5
$\tau^- \to h^- \pi^0 \nu_\tau$	$\rho(770)$	26.0
$\tau^- \to h^- \pi^0 \pi^0 \nu_\tau$	$a_1(1260)$	10.8
$\tau^- \to h^- h^+ h^- \nu_\tau$	$a_1(1260)$	9.8
$\tau^- \to h^- h^+ h^- \pi^0 \nu_\tau$		4.8
Other hadronic modes		1.8
All hadronic modes		64.8

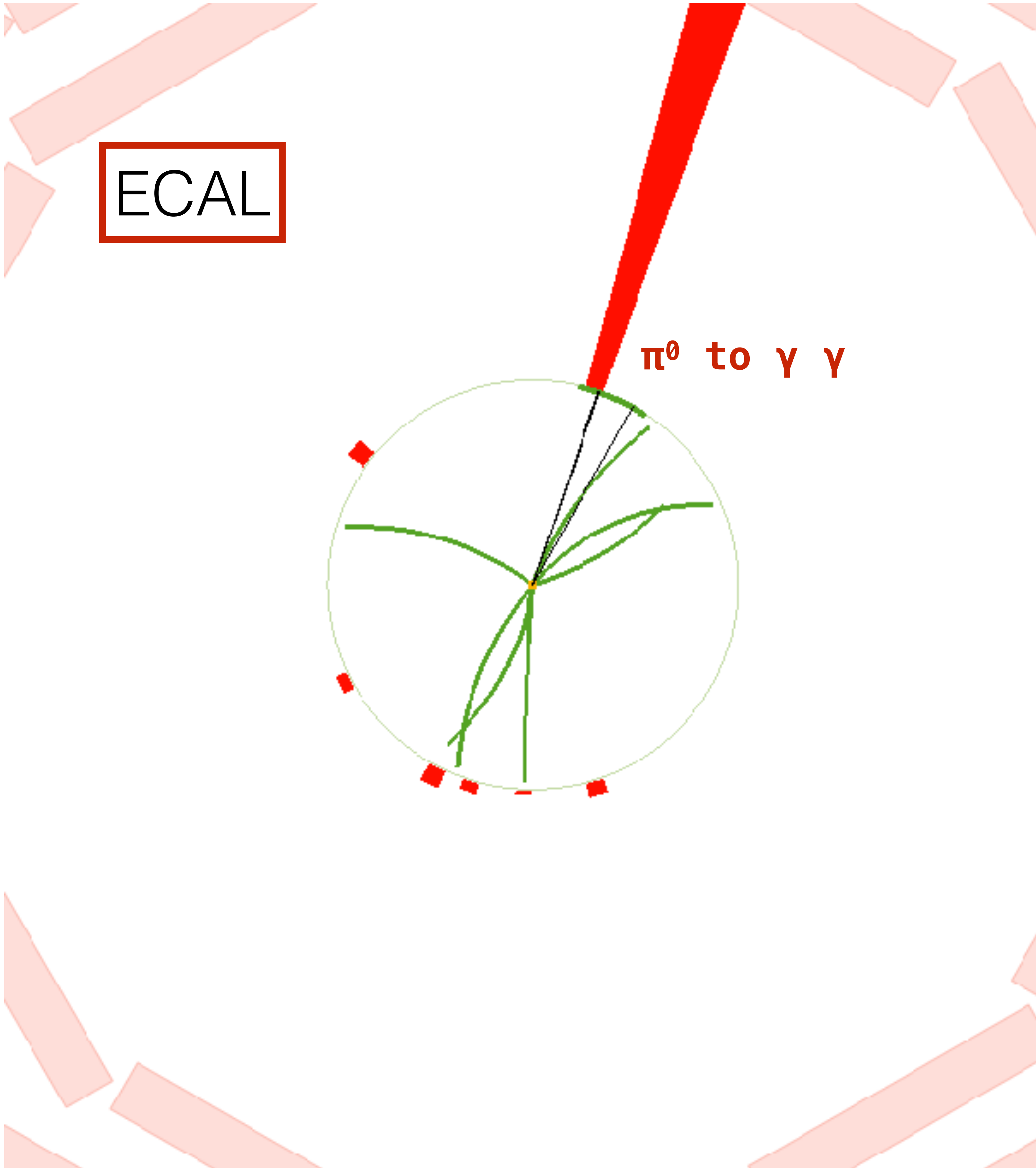
Remember: At CMS Taus are never Fully Reconstructed due to the presence of Neutrinos, instead many studies of Tau Leptons only make use of the **'visible decay products'**

For more information about Taus
http://www.springer.com/us/book/9783662156537

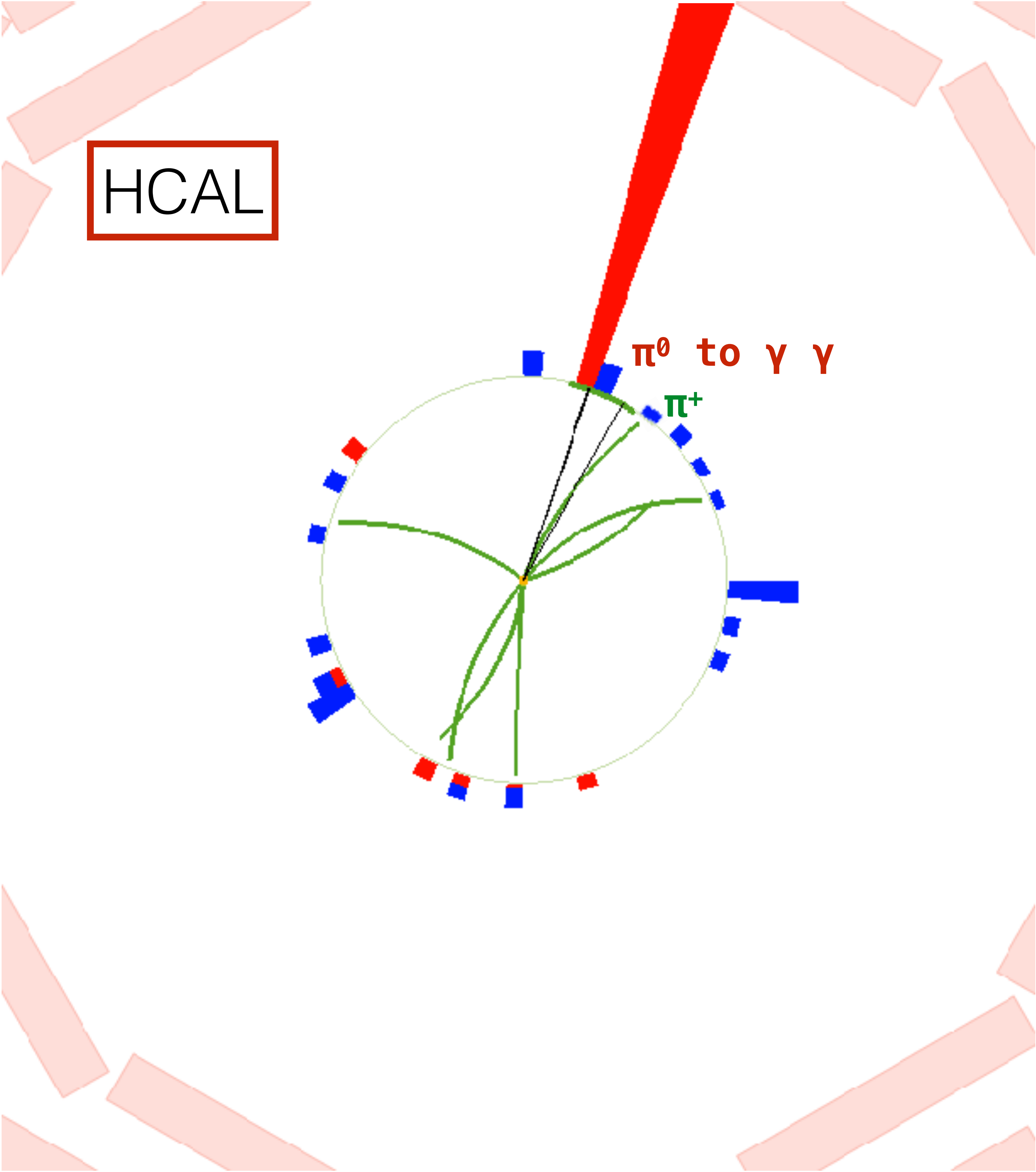
Tau Detection at CMS



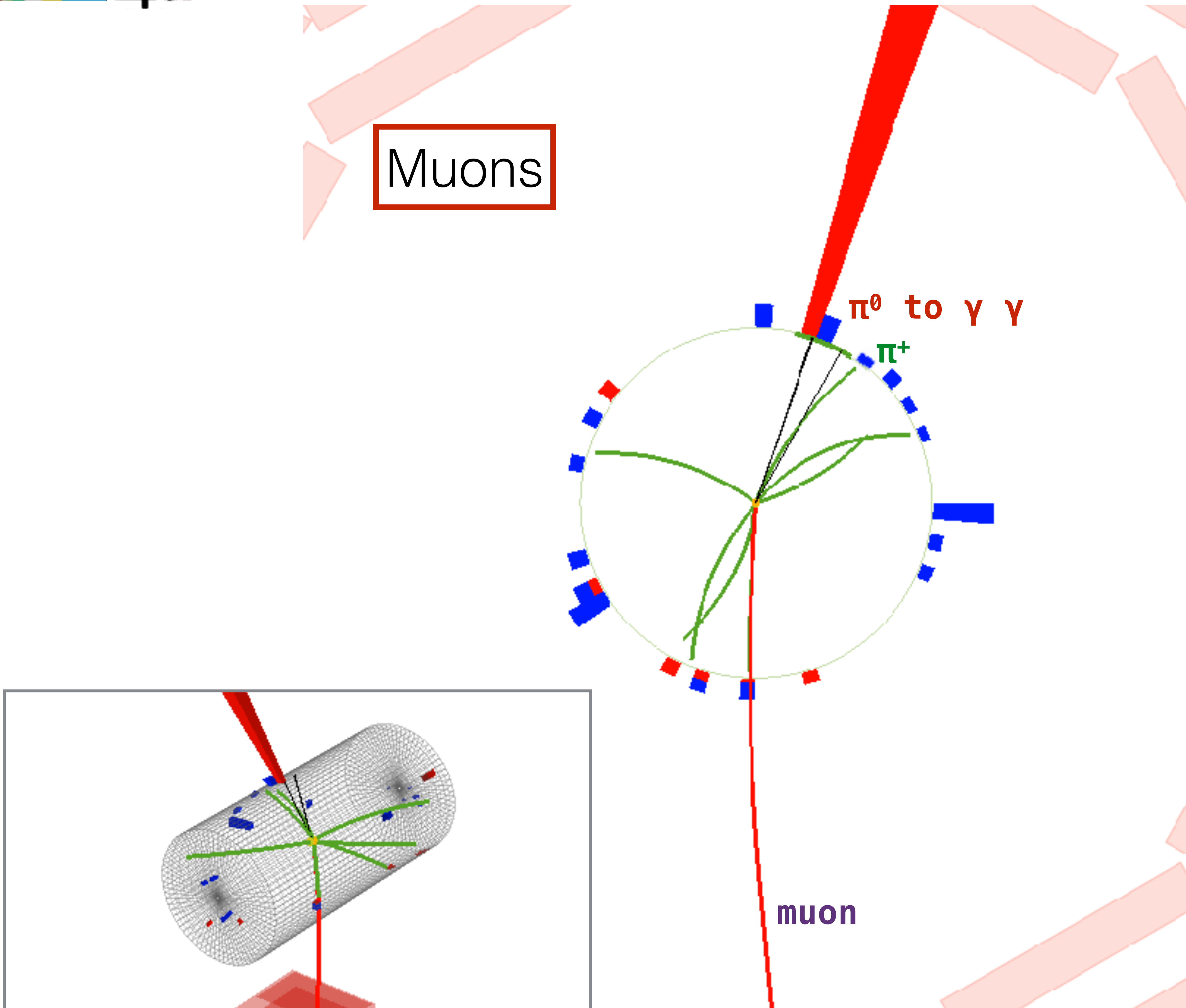
Particle Detection at CMS



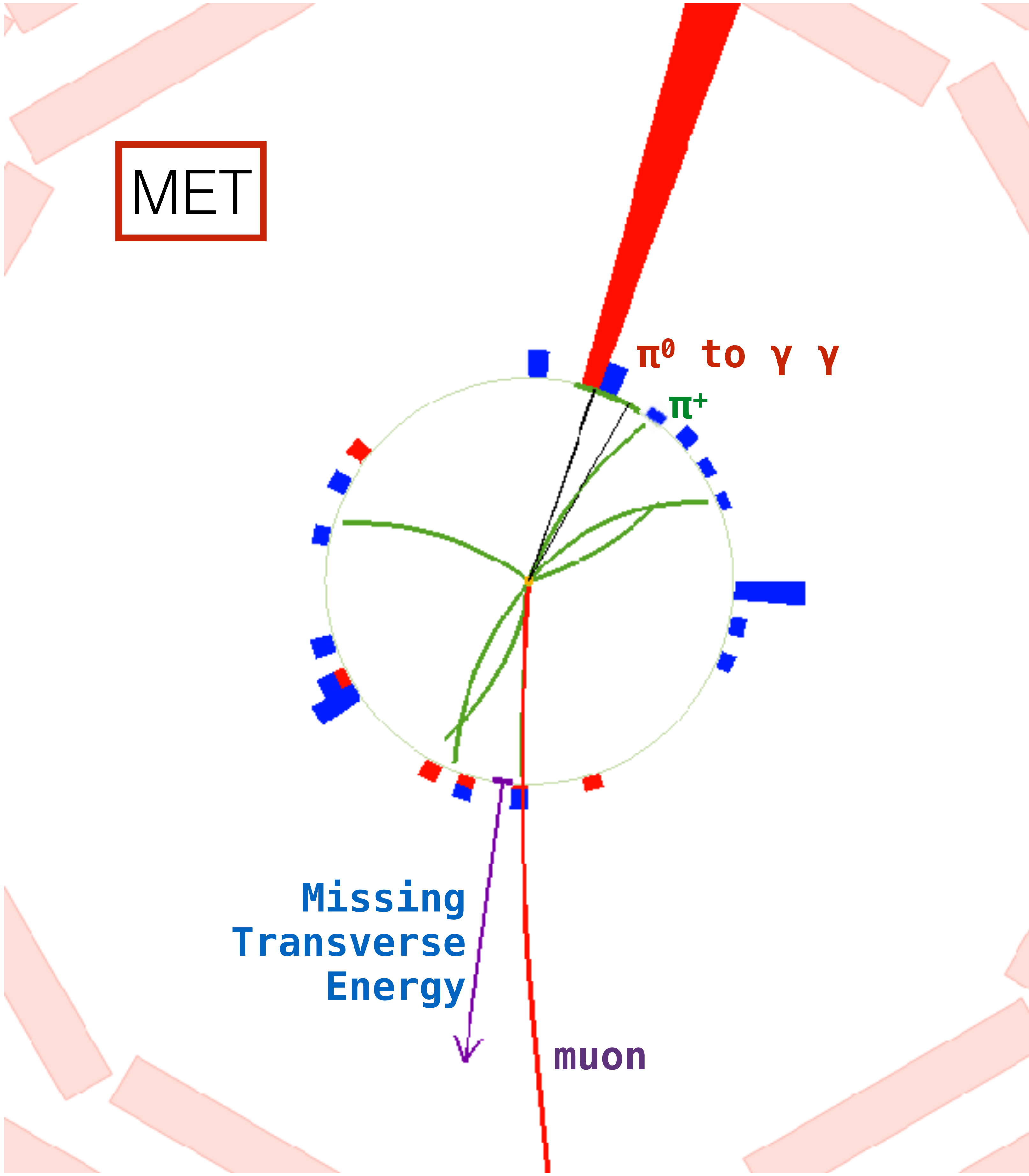
Particle Detection at CMS



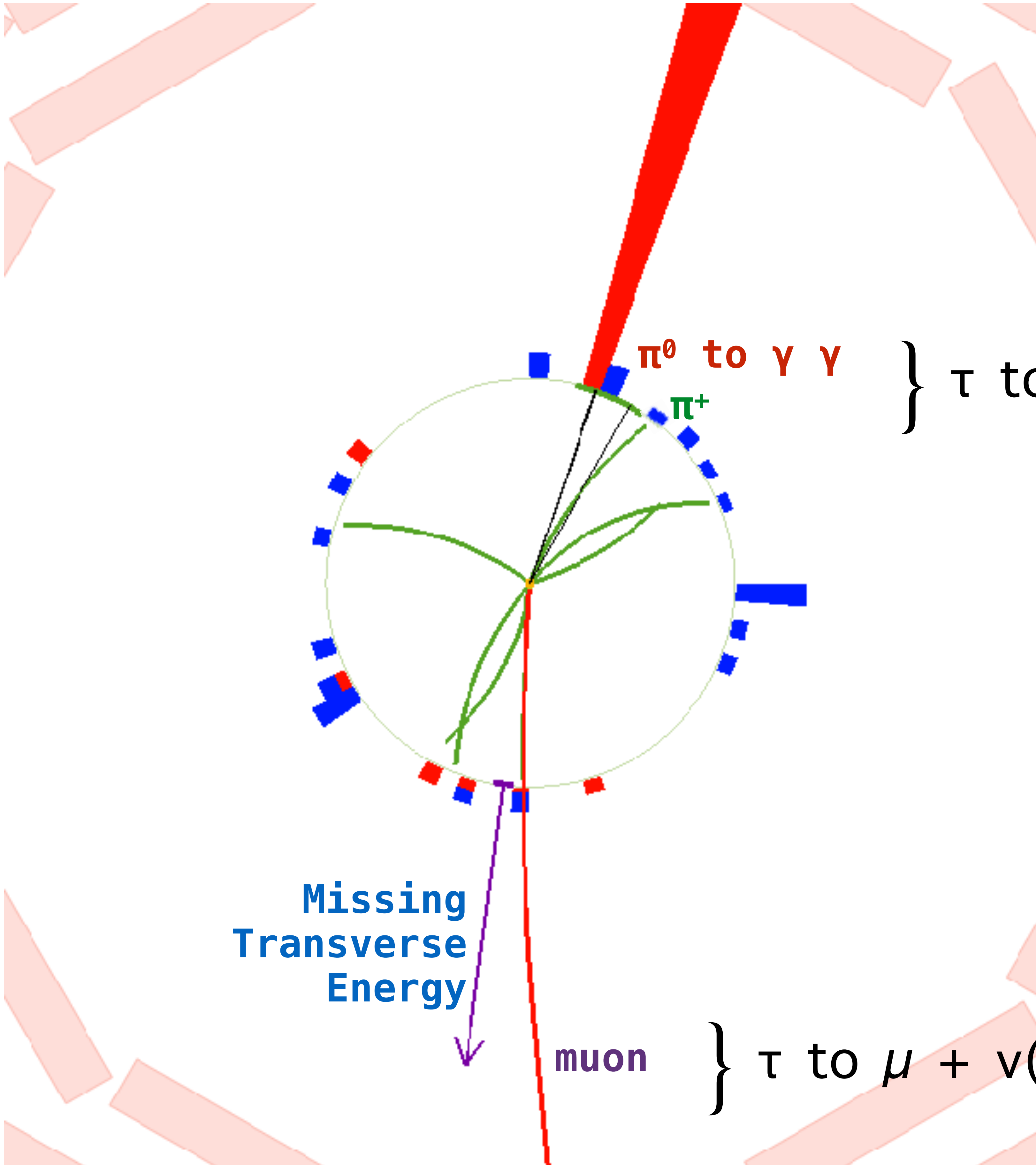
Particle Detection at CMS



Particle Detection at CMS



Particle Detection at CMS



} τ to $\pi^0\pi^+$

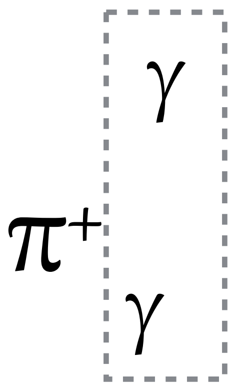
} τ to μ + ν('s)

Tau ID: Hadron+Strips

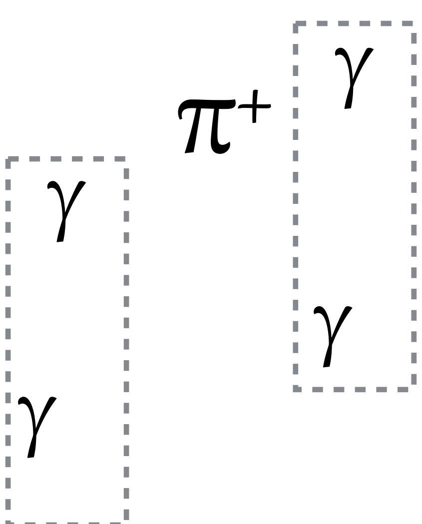
Tau Algorithm Seeded by a Particle Flow Jet

$\pi^o \rightarrow \gamma\gamma$

1 prong + 1pi0



1 prong + 2pi0



Special attention paid to photon conversions in the tracker material which decay into "Strips"

- $\gamma \rightarrow e^+e^-$ undergoes bending of e^+,e^- tracks in B-field of CMS, broadens the signature of neutral pions
- Strips are then built out of EM Particles (PF photons and electrons)
- Strip reconstruction starts by centering on the most energetic EM particles within the PF Jet
- The algorithm searches for other particles within a window of eta 0.05 phi 0.20
- If other EM particles are found within that window, the most energetic one is associated with the strip and the strip four-momentum is recalculated
- pTStrip >1 GeV

Tau ID: Hadron+Strips

1 prong

π^+

Single Hadron is assumed to have the mass of the pion

3 prong

π^+

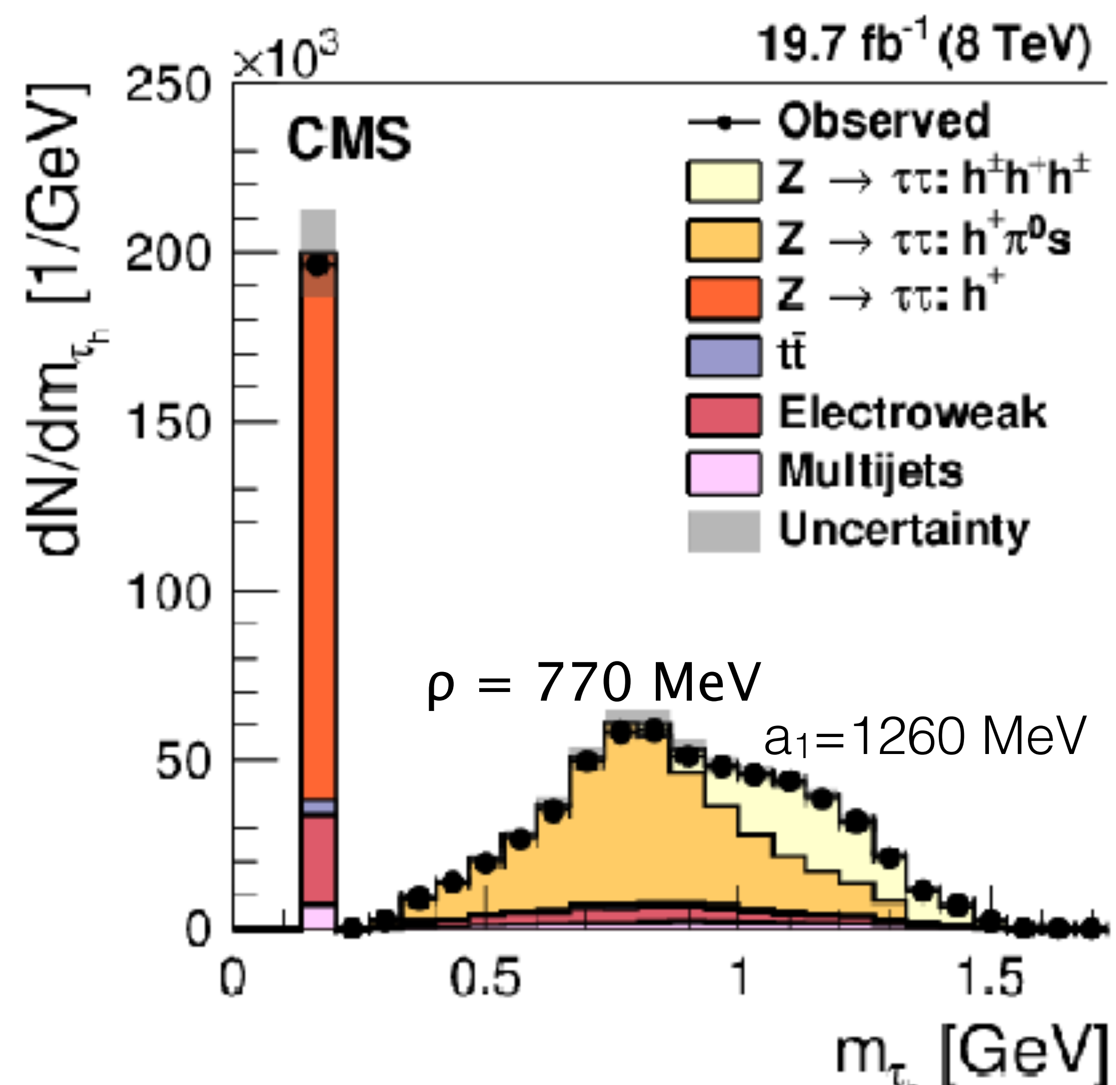
π^-

π^+

Three charged hadrons are required to come from the same secondary vertex created using a kalman vertex fitter

The τ_h mass distribution is used to control the tau energy-scale within 3%

Decay Mode	Resonance	$\mathcal{B}$ [%]
$\tau^- \to e^- \overline{\nu}_e \nu_\tau$		17.8
$\tau^- \to \mu^- \overline{\nu}_\mu \nu_\tau$		17.4
$\tau^- \to h^- \nu_\tau$		11.5
$\tau^- \to h^- \pi^0 \nu_\tau$	$\rho(770)$	26.0
$\tau^- \to h^- \pi^0 \pi^0 \nu_\tau$	$a_1(1260)$	10.8
$\tau^- \to h^- h^+ h^- \nu_\tau$	$a_1(1260)$	9.8
$\tau^- \to h^- h^+ h^- \pi^0 \nu_\tau$		4.8
Other hadronic modes		1.8
All hadronic modes		64.8

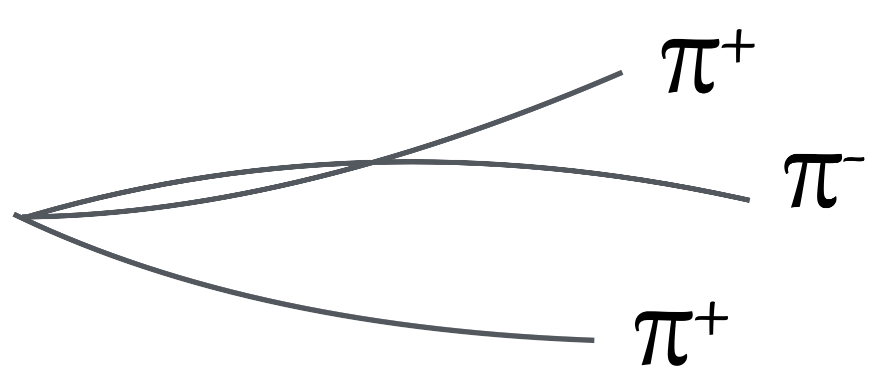


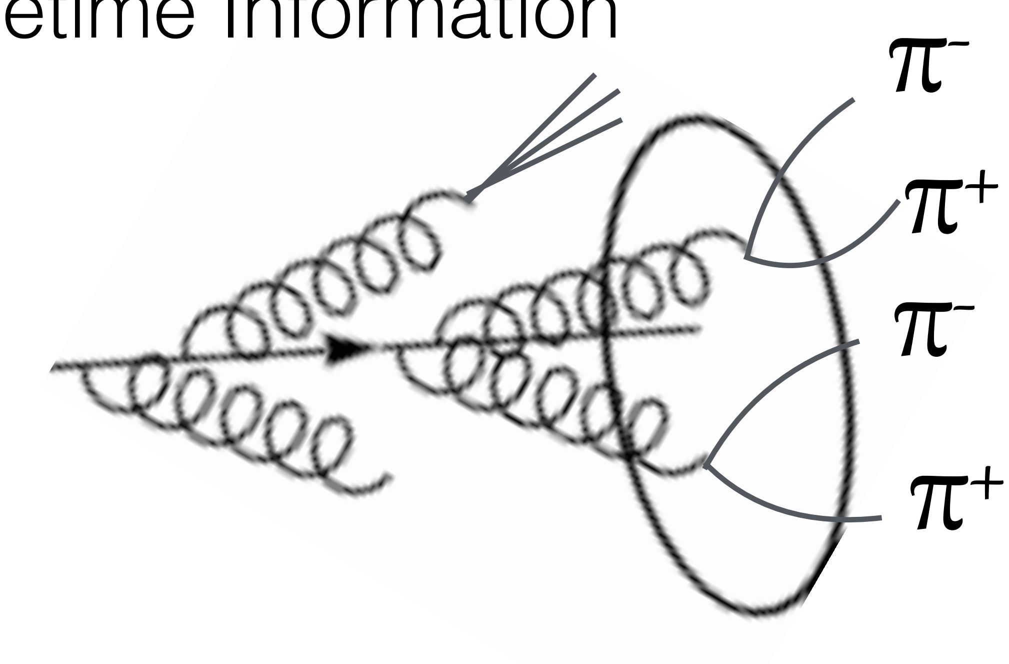
What are Fake Taus? (1)

Jets are produced in large amounts at Hadron Colliders

- QCD, Pileup, other EWK interactions, ect.
- Jets consist of a wide range of charged and neutral particles with various multiplicities
- Able to produce exactly the same decay products as can be found in Hadronically decaying Taus
- **HOWEVER** almost always produce more constituents than are produce in Hadronic Taus

➡ **Key to reduce Jets faking Taus? Isolation**
New for Run II: Also using Tau Lifetime Information

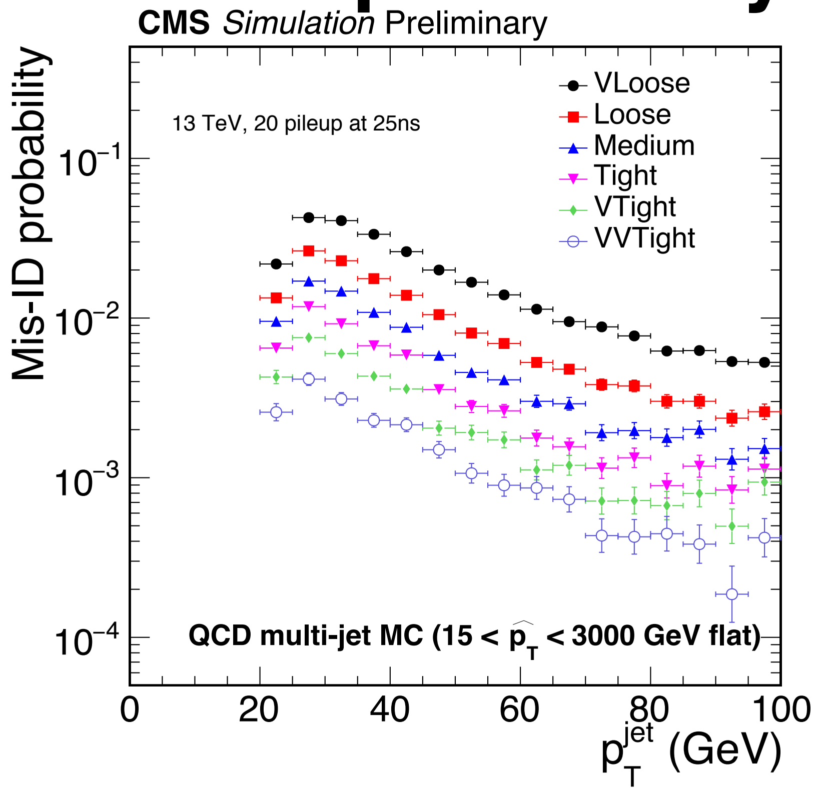




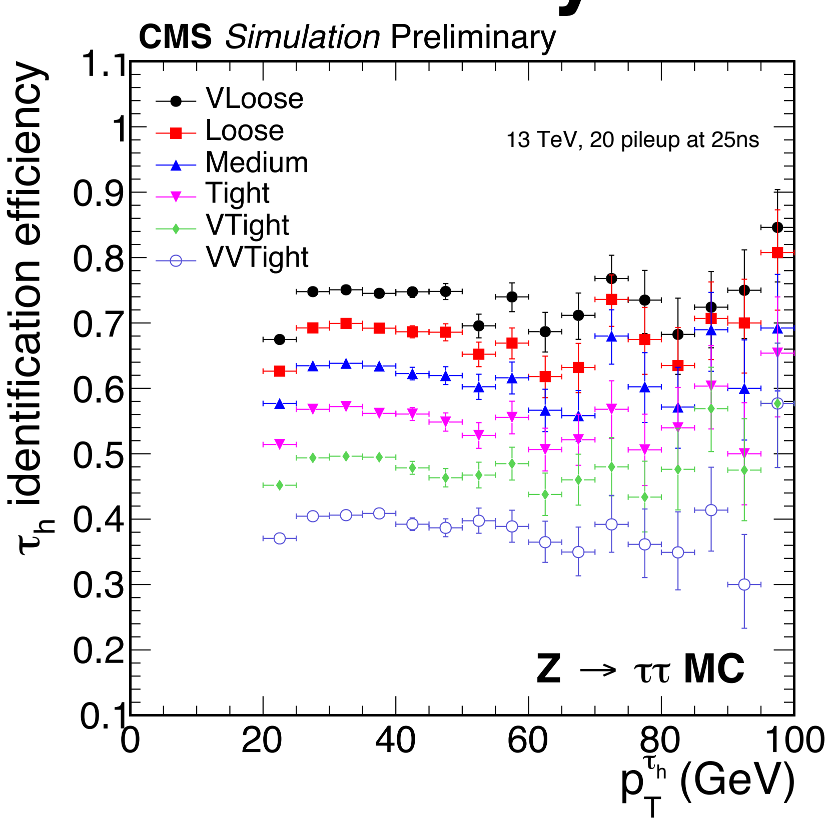
What are Fake Taus? (1)

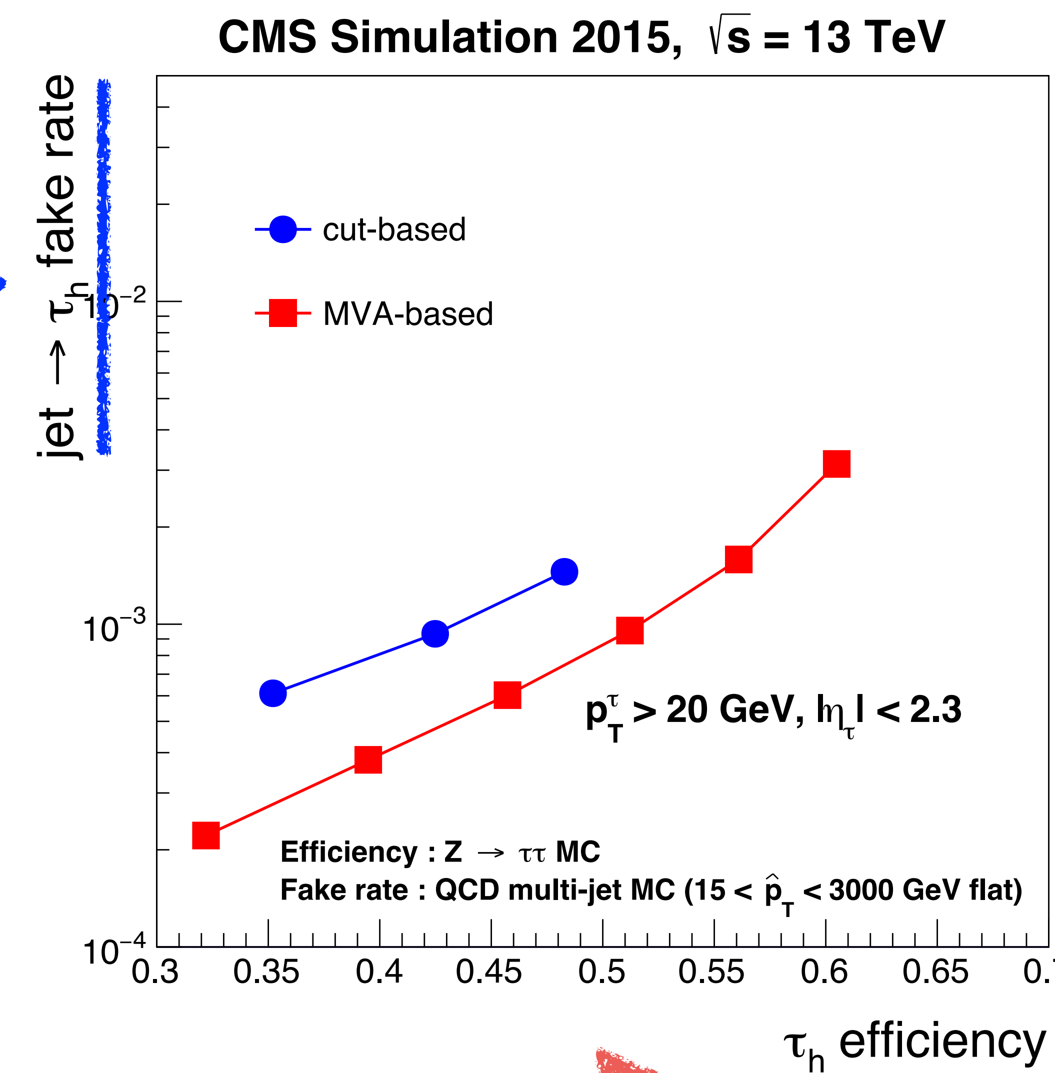
Isolation Discriminator Performance

Tau Mis-ID probability



Tau ID Efficiency





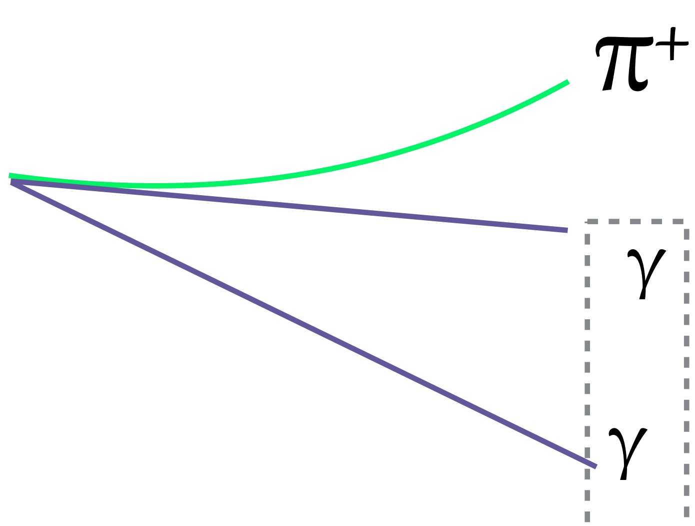
What are Fake Taus? (2)

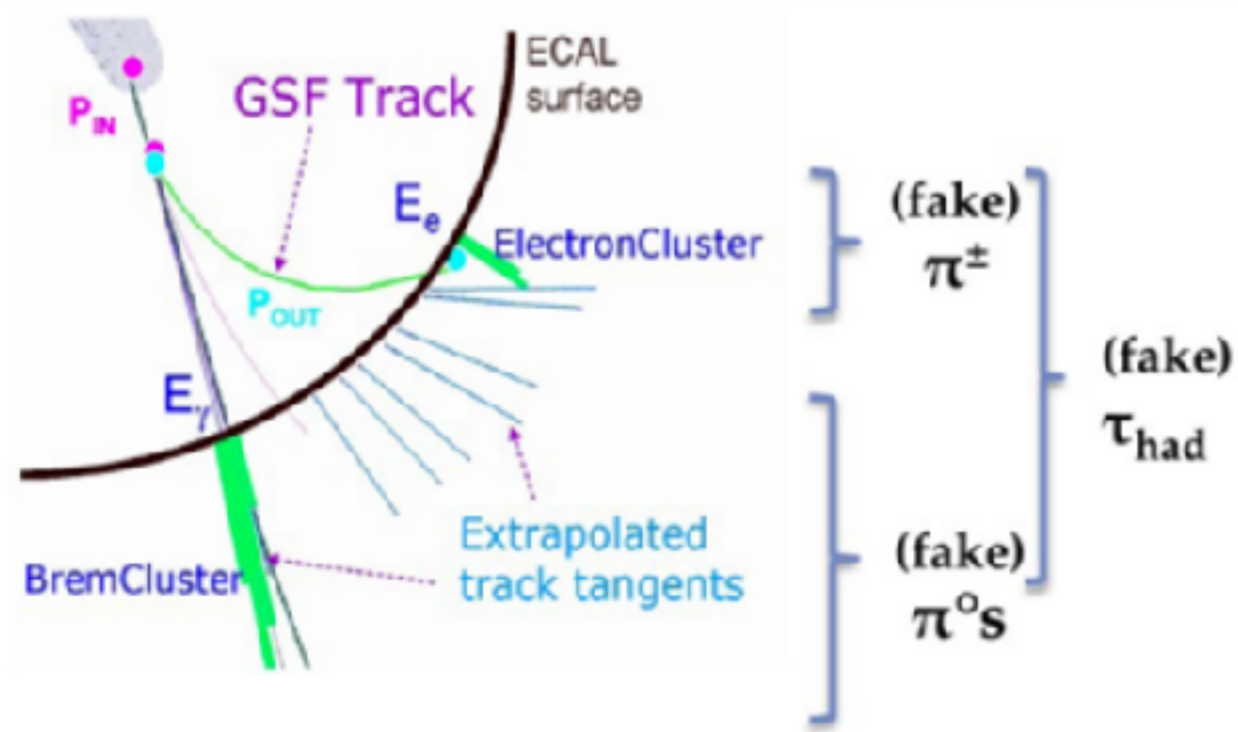
Electrons Faking Taus

- a “perfect” electron (1 GSF track + ECAL deposit) can pass the tau-ID reconstruction algorithm
- Will pass isolation requirements

* However, charged hadrons have a higher HCAL to ECAL ratio
* Electron ID also uses the shape of energy deposits in the ECAL

➡Key to reduce Electrons faking Taus?
Invert Electron ID using an MVA

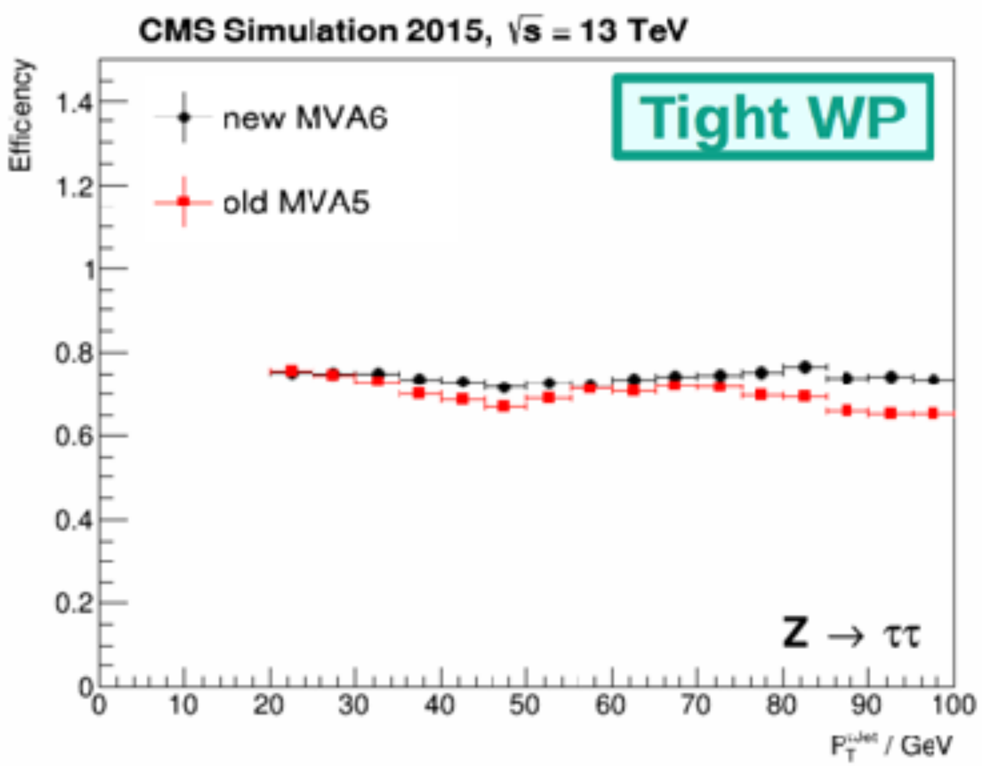


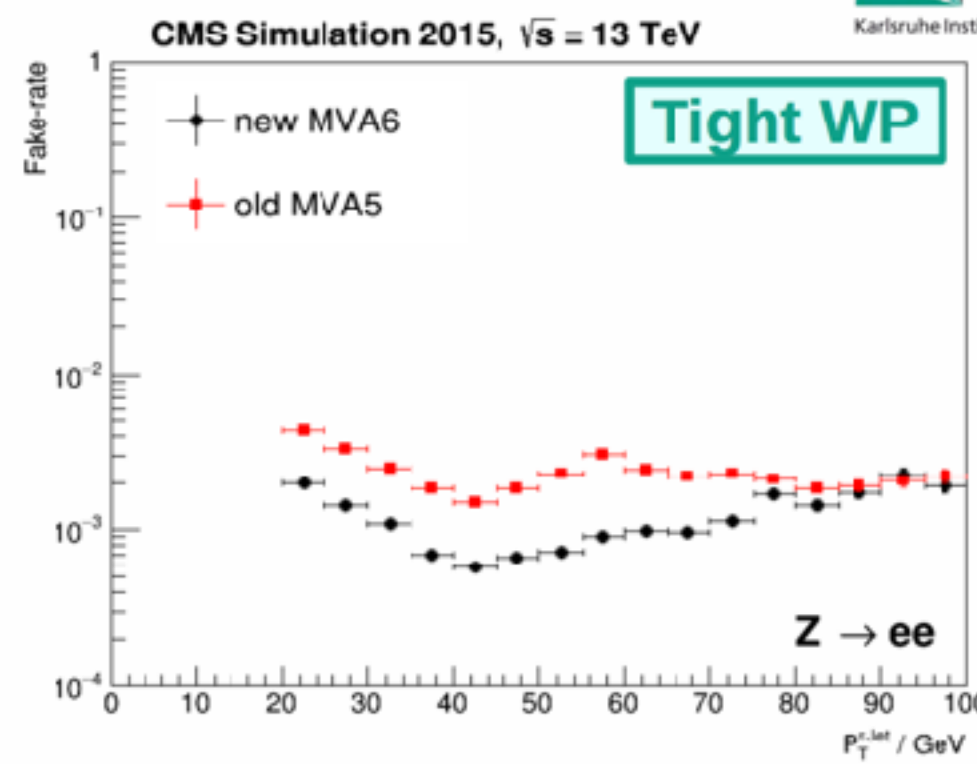


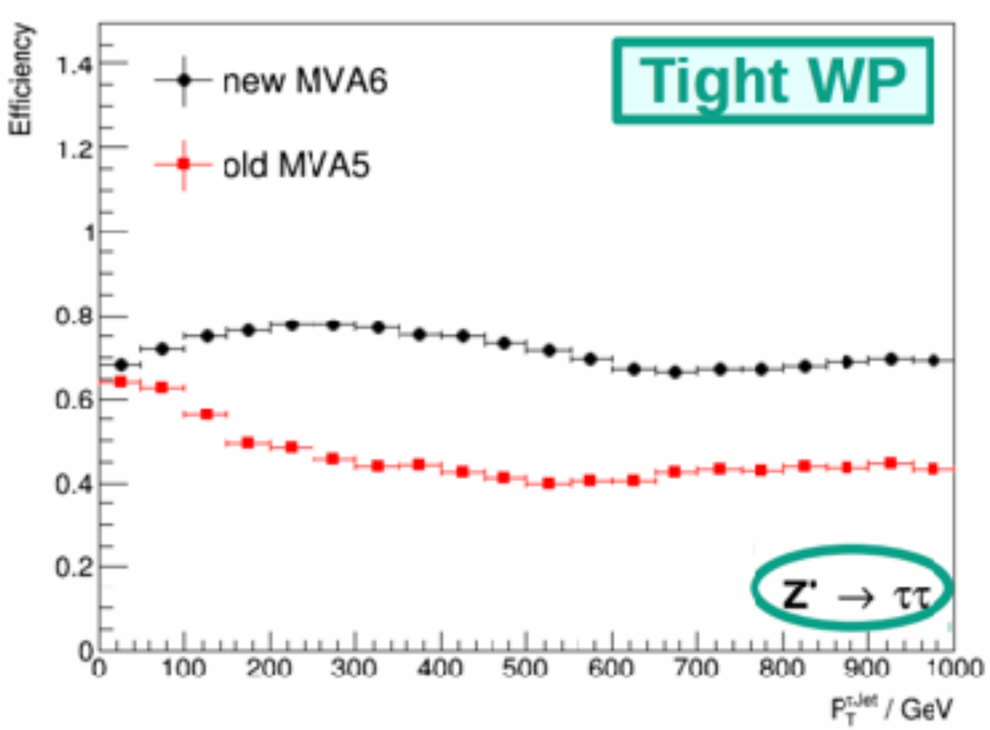
What are Fake Taus? (2)

MVA6 vs. MVA5 comparison (efficiency / fake-rate vs. pT)

Karlsruhe Institute of Technology







- New MVA6 discriminator has the **same efficiency of the old one**
- **Fake-rate is reduced**
- **Drop in efficiency** seen at high-pT with the old MVA5 discriminator **seems to be cured** when using the new one

What are Fake Taus? (3)

Finally, Muons can also Fake Taus

* **Muons can also pass the Tau reconstruction algorithm**

➡ **Key to reduce Electrons faking Taus?**

Require energy not deposited in the Muon System

Conclusion: Electrons, Muons, Jets can all fake a Hadronic Tau! In analysis at CMS we use Variable Cuts and MVA Discriminators to reduce the Fake Probability while maintaining a high Identification Efficiency

Jet to Tau Fake Probability

How often does a Jet fake a Tau?

'Reconstructed' Visible Tau

$$\frac{\tau_h \text{ Reco. + ID.}}{\text{Fake Prob.}} = \frac{\text{denominator} \;\&\; \tau_h\ p_T > 20\text{ GeV} \;\&\; |\,\eta\,| < 2.3 \;\&\; \text{Isolation}}{\text{Jet } p_T > 20\text{ GeV} \;\&\; |\text{ Jet } \eta\,| < 2.3}$$

'Reconstructed' Visible Jet

1.) **Iterate** through all Reconstructed Jets

2.) **Match** Reconstructed Jets to Reconstructed Taus

- Require the distance (ΔR) between the Reconstructed Jet and Reco Level Tau is less than 0.5

3.) Require the Reconstructed Tau passes the **Identification**

Tau Identification Efficiency

How often will a true Tau Lepton be reconstructed at CMS?

'Reconstructed' Visible Tau

$$\tau_h \text{ Reco. + ID. Efficiency} = \frac{\text{Denominator} \; \& \; p_T > 20 \text{ GeV} \; \& \; |\eta| < 2.3 \; \& \; \text{Isolation} \; \& \; \text{Decay Mode Finding}}{\text{Gen. visible } p_T > 20 \text{ GeV} \; \& \; |\text{Gen. visible } \eta| < 2.3}$$

'Generator-Level' Visible Tau

1.) Iterate through all Generator-Level Taus ⇨ **|pdg ID| = 15**

2.) Match Generated Taus to Reconstructed Taus

- Require the distance (ΔR) between the Reconstructed Jet and Reco Level Tau is less than 0.5

3.) Require the Reconstructed Tau passes the **Identification**

Exercise

https://twiki.cern.ch/twiki/bin/viewauth/CMS/SWGuideCMSDataAnalysisSchoolLPC2018TauShortExercise

Introduction

The goal of this exercise is to measure the jet->tau fake rake and the efficiency for several tau identification working points (IDs).

1. Measure jet to tau fake rate from MC or Data.
2. Measure tau ID efficiency from MC.
3. Make a Roc curve with 3 different tau IDs.

In case you have time you can compare the jet to tau fake rate between data and MC. All you need is to repeat Step 1 on Single Muon dataset and compare it with fake rate from W+Jet MC Simulation samples.

Jet to Tau Fake Probability

https://twiki.cern.ch/twiki/bin/viewauth/CMS/SWGuideCMSDataAnalysisSchoolLPC2018TauShortExercise

Exercise 1: Jet to Tau Fake Rate

The jet->tau fake rate can be easily measured both on Data (in this case using a Single Muon dataset) and on Monte Carlo Simulation (a W+Jet Sample)

This exercise aims to answer the question "How likely is a jet to be mis-identified as a tau lepton?" for different tau identification working points (IDs). The general process is as follows:

1. Select W events (W->mu nu events). To do so find a good muon in event and check the Transverse Mass of muon and MET to be compatible with W Transverse Mass (M_T).
2. Find a tau jet which pass decay mode finding and anti muon and anti electron discriminator.
3. Find how often this object can pass different working points of the Tau Isolation.

While doing this exercise, please consider the following:

1. Try all eta ranges on the taus. Remember to always make an eta cut of 2.1 for your muons and 2.3 for your taus! Why is this eta range important? Answer: Show
2. One can parametrize this fake rate v.s. Tau Pt, closest Jet pt to Tau or Tau ete, etc.
3. Compare tau fake rate between data and MC
4. See the fake rate of different working points! Add in other tauIDs from the TreeReader.h
5. After measuring tau ID from ZTT MC sample, make the ROC curve!

Tau Identification Efficiency

https://twiki.cern.ch/twiki/bin/viewauth/CMS/SWGuideCMSDataAnalysisSchoolLPC2018TauShortExercise

Exercise 2: Tau ID Efficiency

This is a MC only exercise because we are accessing the Generator-Level information.

This aims to answer the question "How likely is it that a real tau passes a specific tau ID?".

The general flow of section is as follows:

1. Find real reconstructed taus that pass and do-not pass ID.
2. Save histograms of generator-level taus and of reconstructed taus.
3. The efficiency is the ratio of real taus that pass an ID over all real taus.

Ideas to consider while working through the exercise:

1. Try all eta ranges on the taus. Remember to always make an eta cut of 2.1 for your muons and 2.3 for your taus! Why is this eta range important? Answer: Show
2. Try a Pt cut for the generator level taus, this could be added to the TauPreSelection for example! The gen-Pt of the tau will be higher than the reconstructed pt of the tau, because a hadronic tau-decay will have a neutrino in the decay products. Try plotting the efficiency as a function of the genPt of the tau.
3. See the efficiency of different working points! Add in other tauIDs from the `TreeReader.h`
4. The choice of tau Id should be based on both fake rate and efficiency.

Tau Identification Efficiency

https://twiki.cern.ch/twiki/bin/viewauth/CMS/SWGuideCMSDataAnalysisSchoolLPC2018TauShortExercise

Exercise 3: ROC Curve

With `outputEfficiency.root` and `outputFR.root` in hand with 4 histograms in each (1 denominator and 3 numerator histograms), a ROC curve can be plotted with the 3 IDs.