The Lepton&Tau Identification at CEPC

LCWS2019 @ Sendai

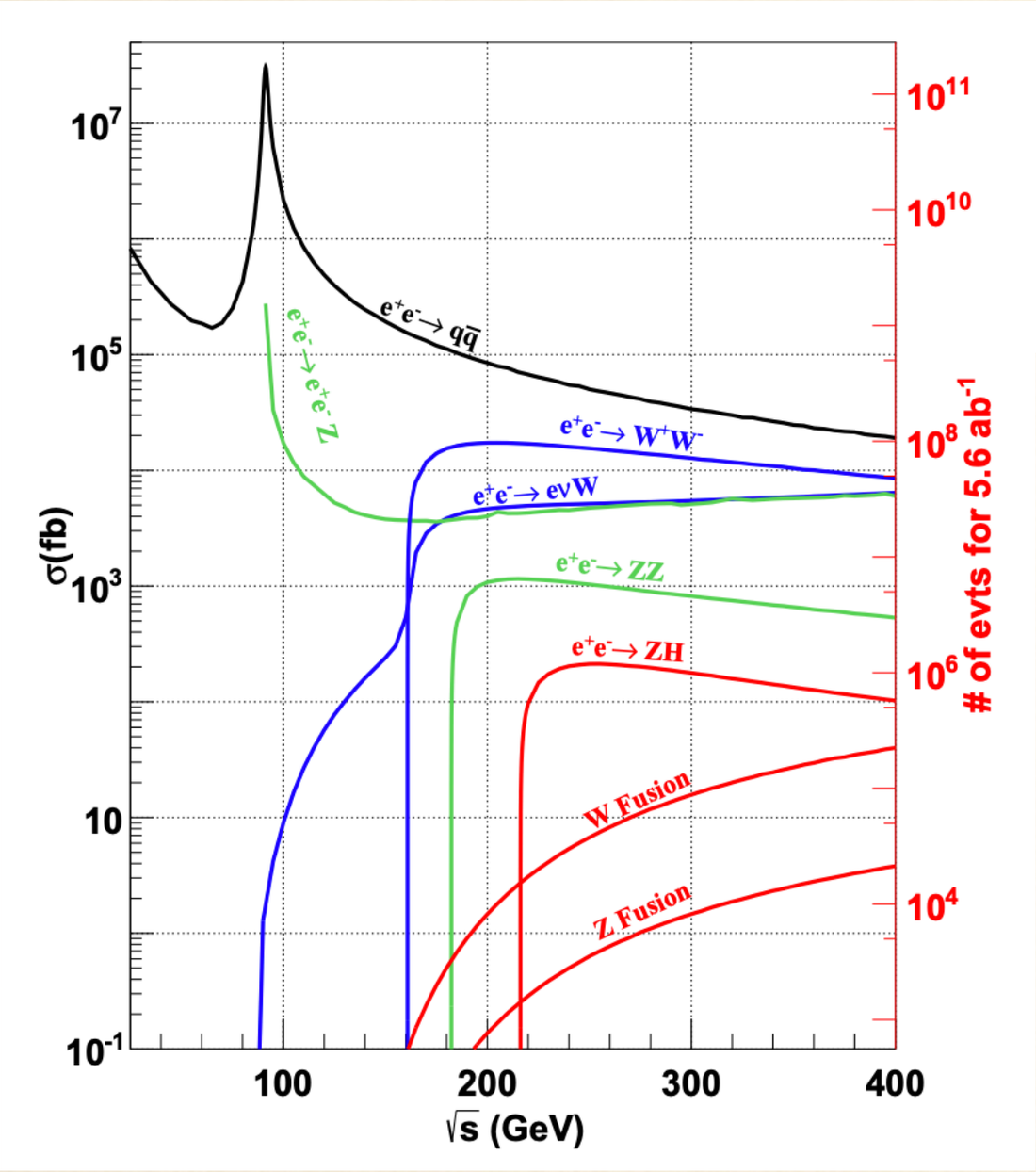
Dan YU

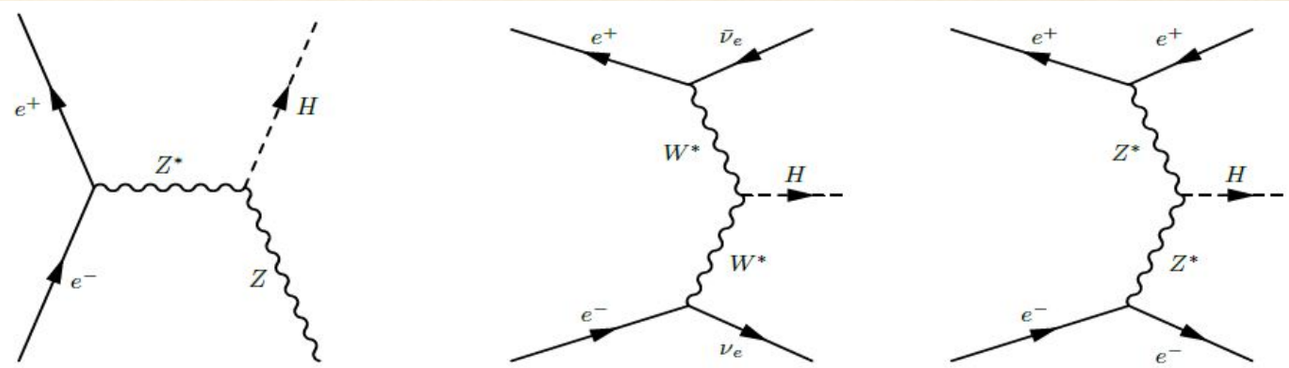
中国科学院高能物理研究所
Institute of High Energy Physics Chinese Academy of Sciences

Plan

- Introduction
- Lepton Identification
  - Single lepton
  - Lepton in jets
- Tau Identification
- Summary

CEPC





- Higgs factory: 240 GeV, 10^6 Higgs,
  - Advantage: Clean, Known initial states
  - Measurements: Higgs boson mass, cross section, decay modes, branching ratio
- Z factory: 91 GeV, 6×10^{11}
  - EW precision physics
- WW threshold runs, ~160GeV, 10^8
  - W mass/width measurement
- PFA Oriented detector

Light Lepton (Isolated)

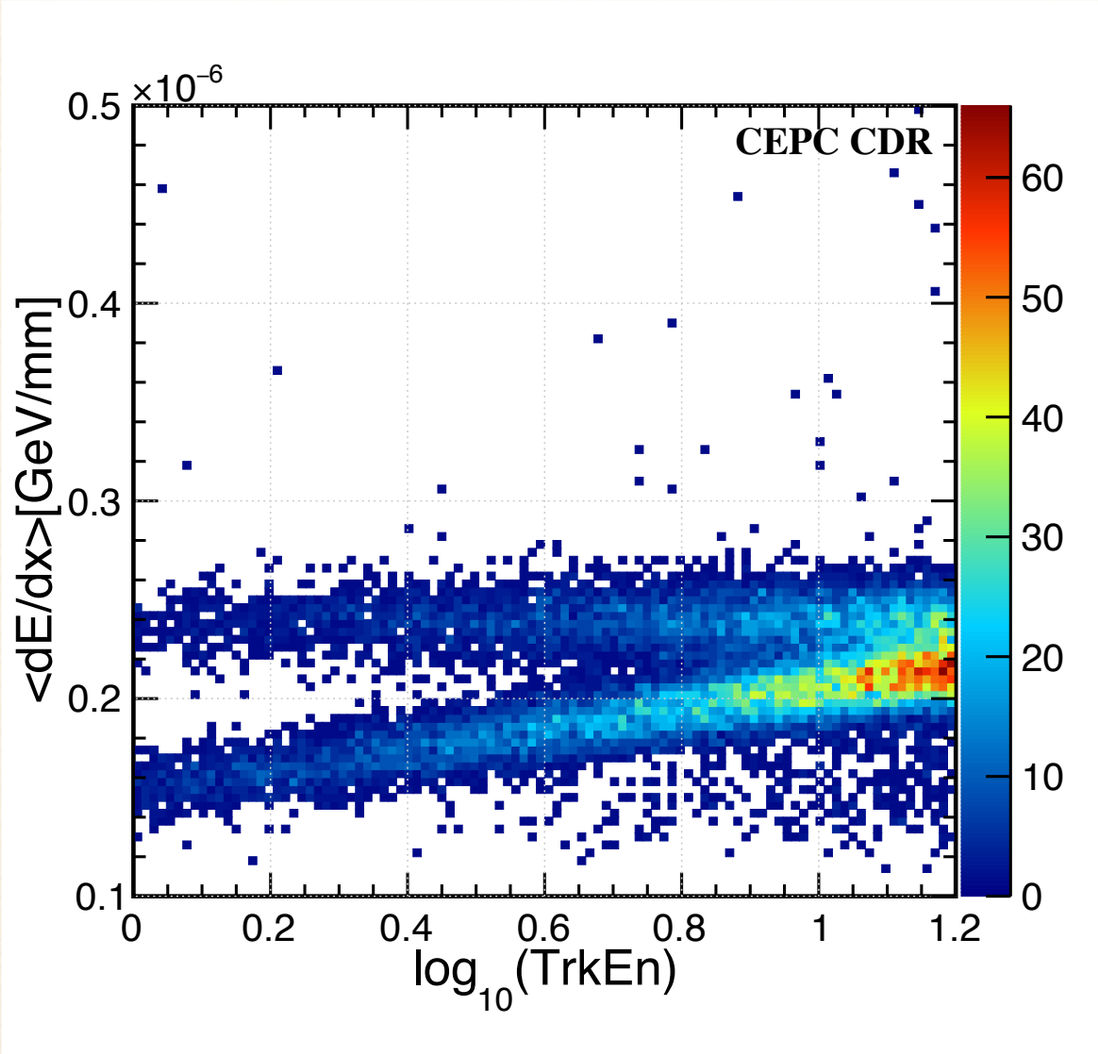
Essential to the precise Higgs measurements
jet flavor tagging and the jet charge measurement

Sample

- LICH (Lepton Identification for Calorimeter with High granularity)
  - Input: 24 variables from reconstructed charged particle
  - Tool: TMVA
  - Training samples: Single particle: e, μ, π (1 GeV ~ 120 GeV) at different regions (endcap, barrel, overlap)
  - Output: likelihood

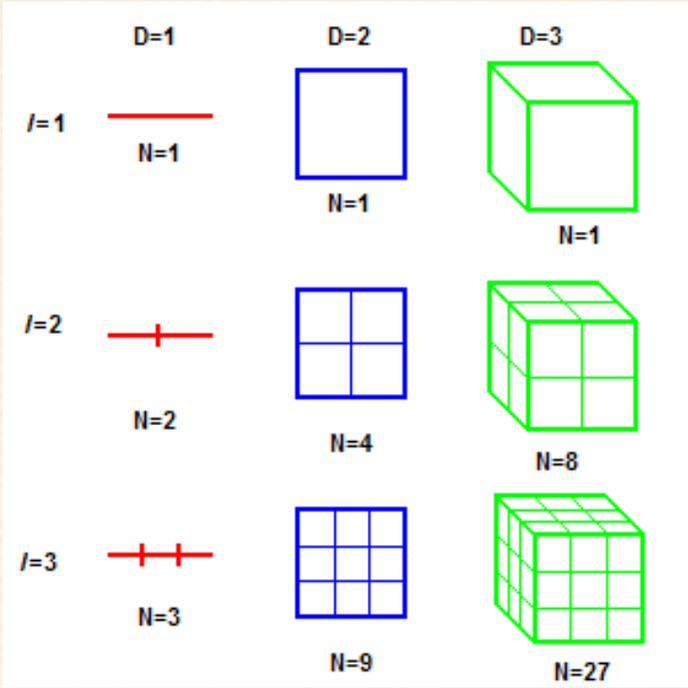
dE/dx

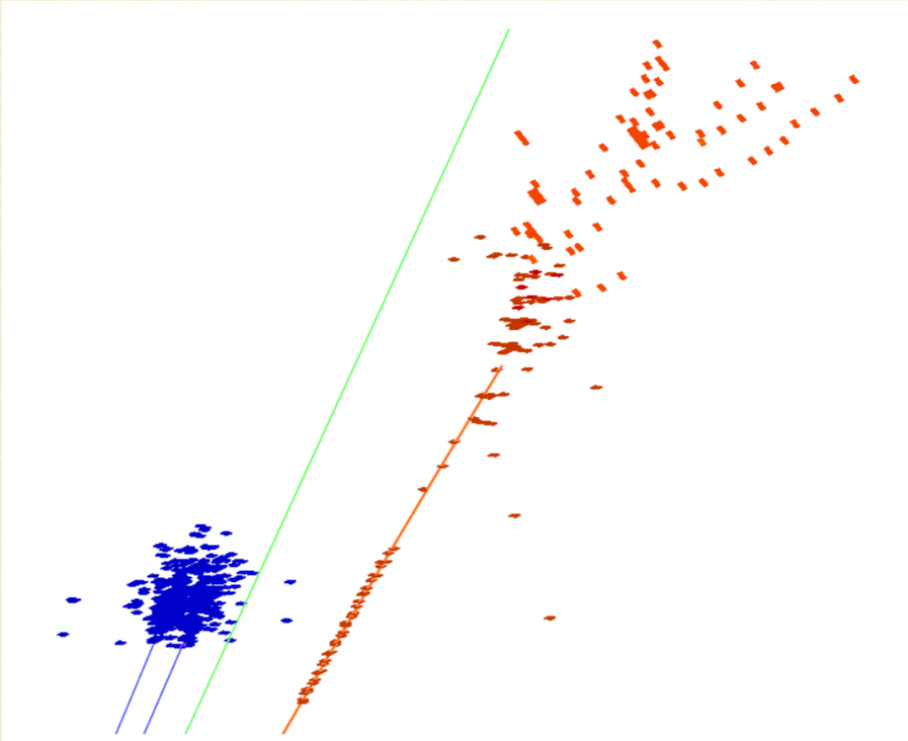
- For a track in TPC, the distribution of energy loss per unit of depth follows an approximately Landau distribution.

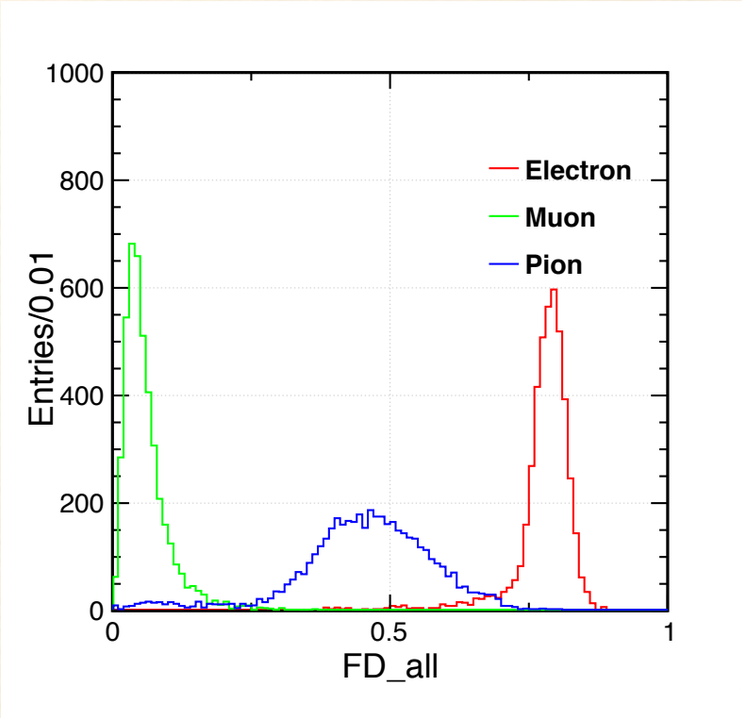


Fractal Dimension

- Describe the self-similar behavior of shower spatial configurations (compactness of the particle shower)
  - $FD_{\beta} = \langle \log(R_{\alpha,\beta})/\log\alpha \rangle + 1$
    where $R_{\alpha,\beta} = N_{\beta}/N_{\alpha}$, α and ß are scales at which the shower is analyzed.
  - Average over range: 1cm - 120cm

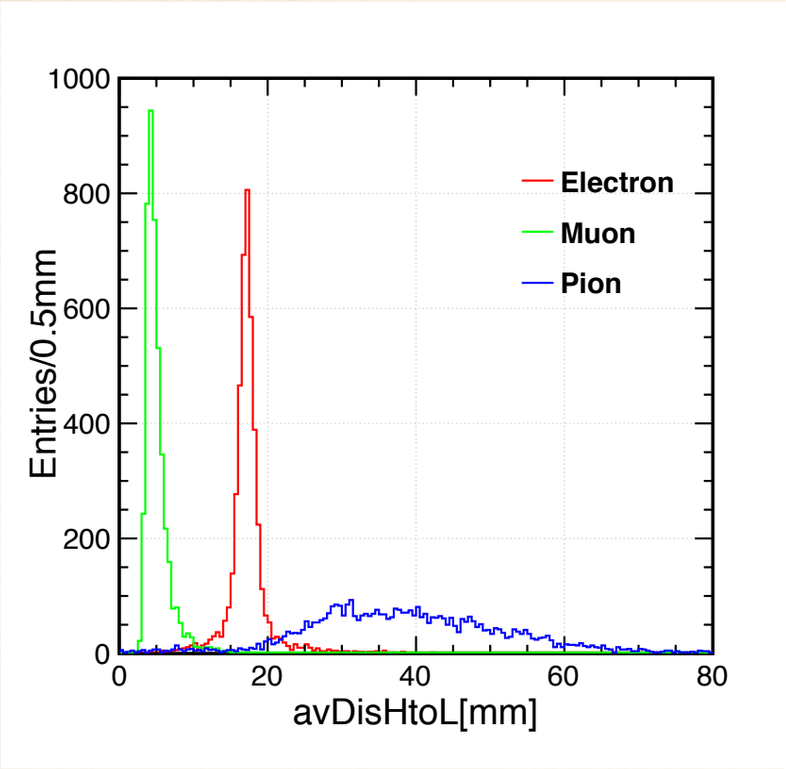


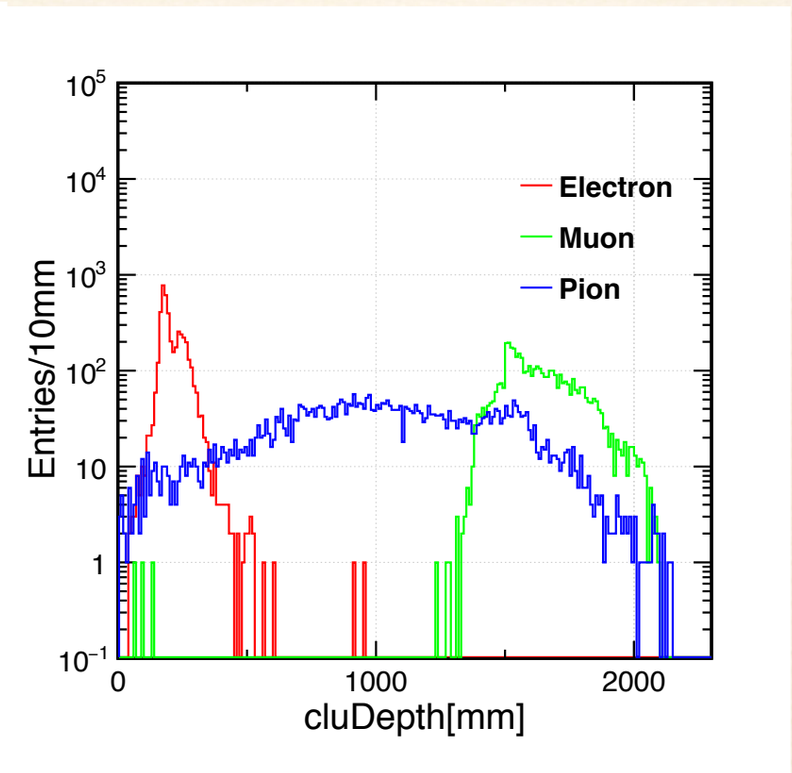


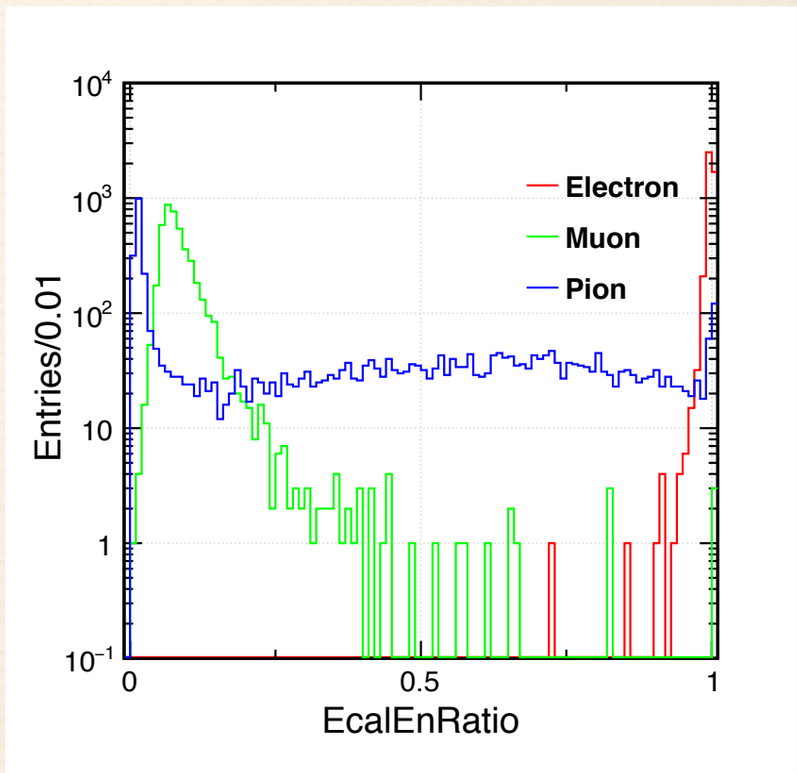


Other Parameters

- Proportion of energy: Energy deposit in the first 10 layers in ECAL to the entire ECAL, or the energy deposit in a cylinder around the incident direction with a radius of 1 and 1.5 Moliere radius.
- Distance(max, min, avr) between hit and track / axis
- Number of hits / number of layers hit by the shower
- Depth
- …

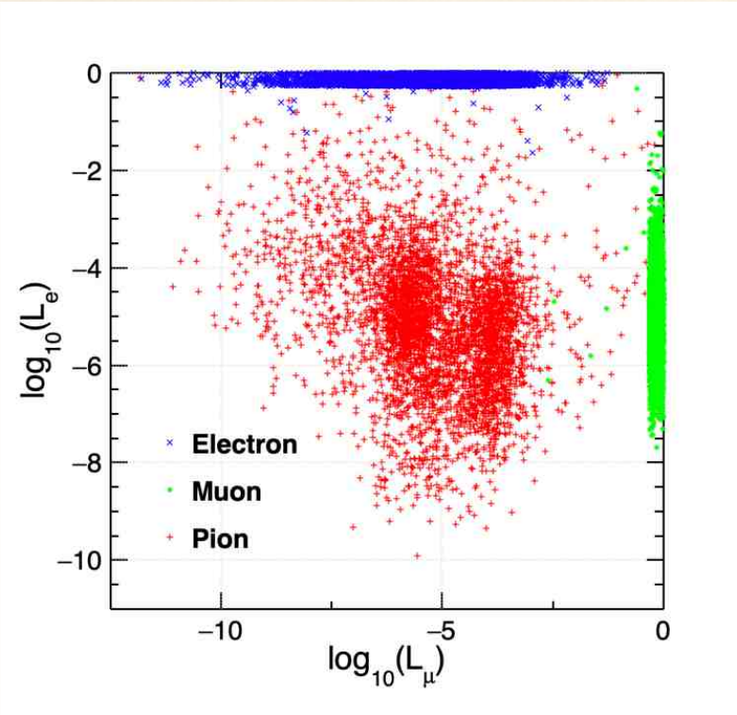


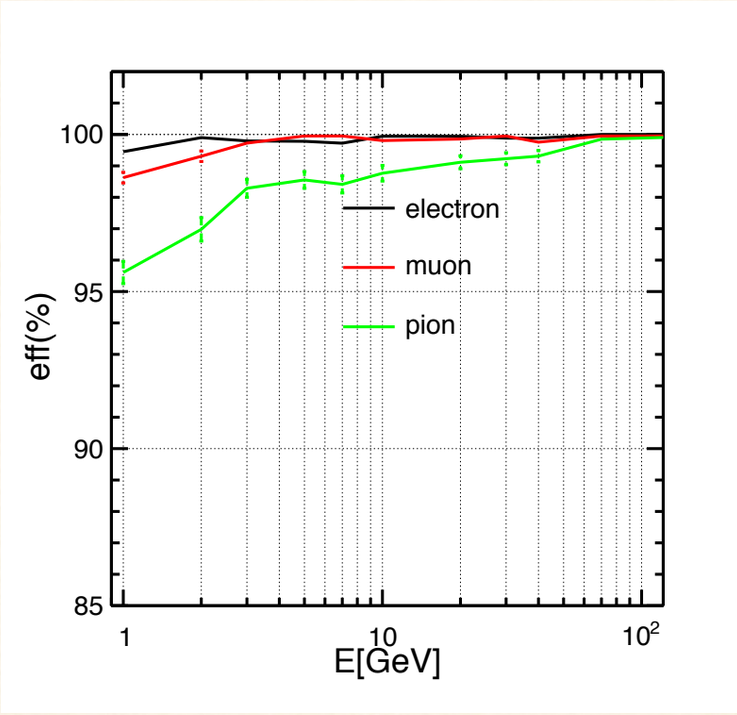


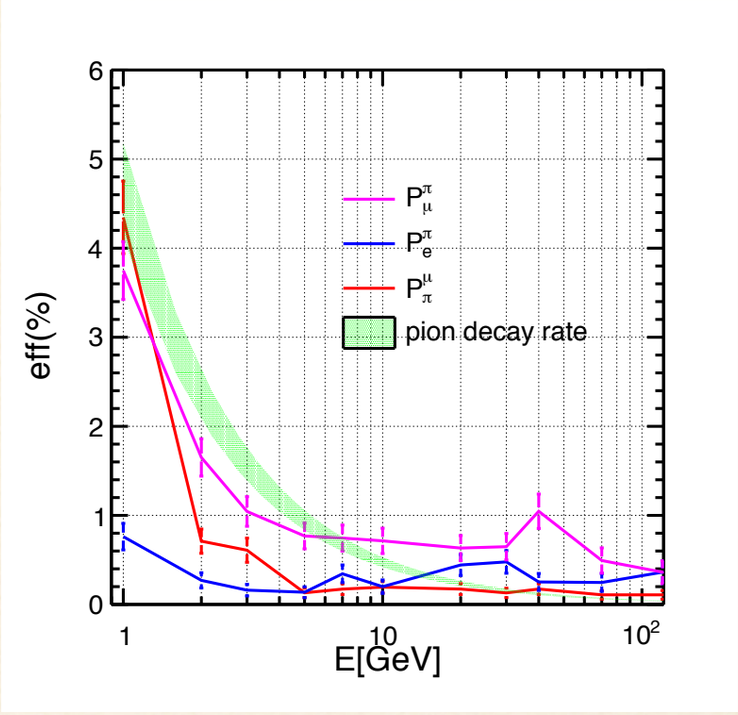


Result

- LICH uses TMVA methods to summarize 24 input variables into two likelihoods, corresponding to electrons and muons.
- The efficiency for electron and muon is higher than 99.5% (E>2 GeV). Pion efficiency ~ 98%.







Migration Matrix at 40GeV (LICH)

Type	e^- *like*	μ^- *like*	π^+ *like*
e^-	99.71 ± 0.08	< 0.07	0.21 ± 0.07
μ^-	< 0.07	99.87 ± 0.08	0.05 ± 0.05
π^+	0.14 ± 0.05	0.35 ± 0.08	99.26 ± 0.12

Migration Matrix for ALEPH PID ($> 2GeV$)(*Eur.Phys.J.C20:401-430,2001*)

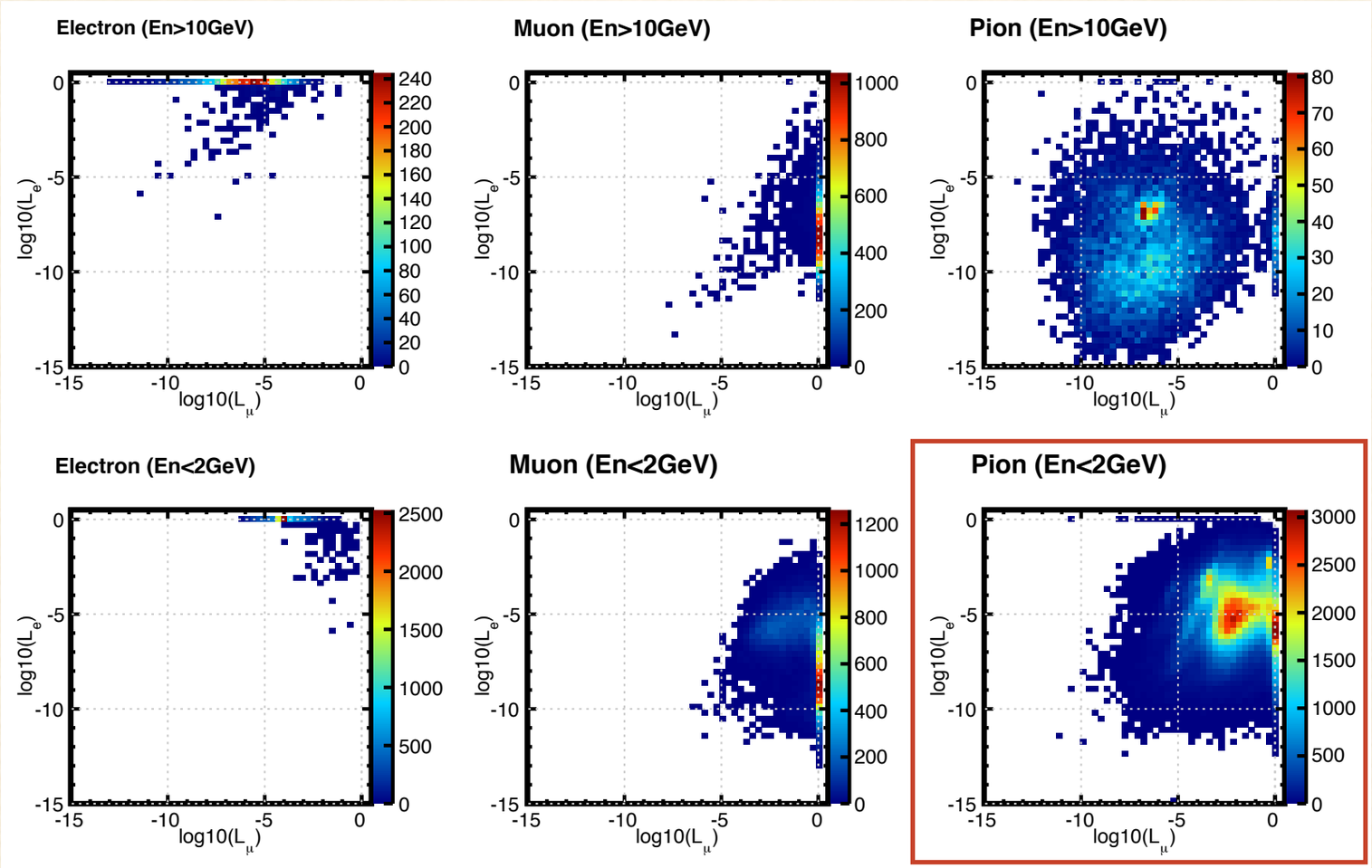
Type	e^- *like*	μ^- *like*	π^+ *like*	undefined
e^-	99.57 ± 0.07	< 0.01	0.32 ± 0.0	0.09 ± 0.04
μ^-	< 0.01	99.11 ± 0.08	0.88 ± 0.08	0.01 ± 0.01
π^+	0.71 ± 0.04	0.72 ± 0.04	98.45 ± 0.06	0.12 ± 0.03

Light Lepton (in Jets)

The performance for lepton in jets degrades comparing to the single particle results because of the high statistics of background and the cluster overlap

Likelihood vs Energy

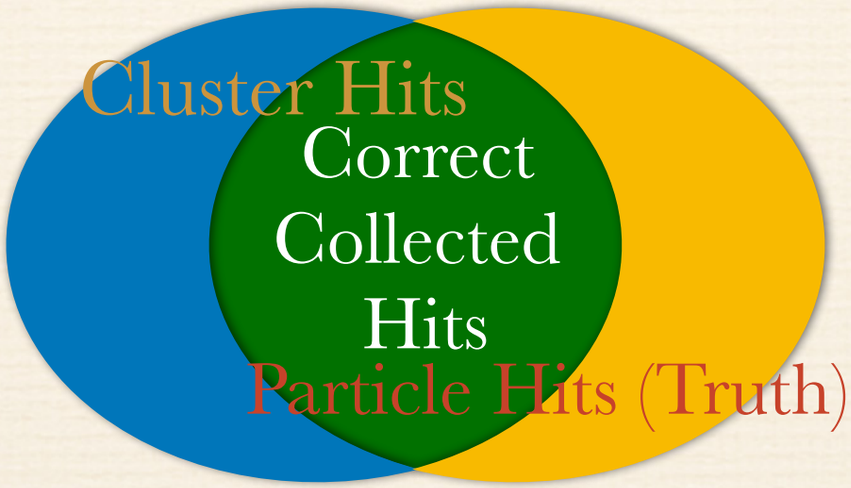
- For higher energy, still nice separation
- For lower energy, pion mixed with muon



Clustering Performance

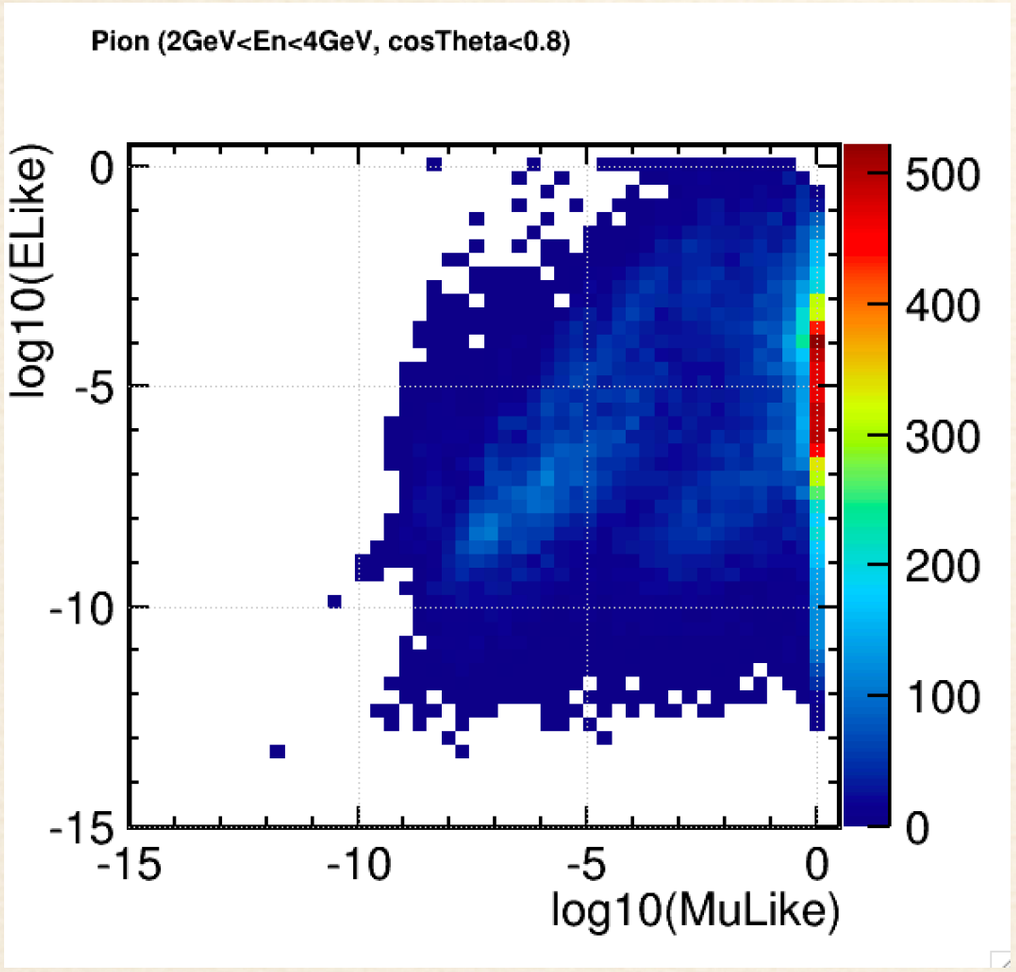
- Use clustering
  - efficiency (correct collected hits/particle hits)
  - purity (correct collected hits/cluster hits)

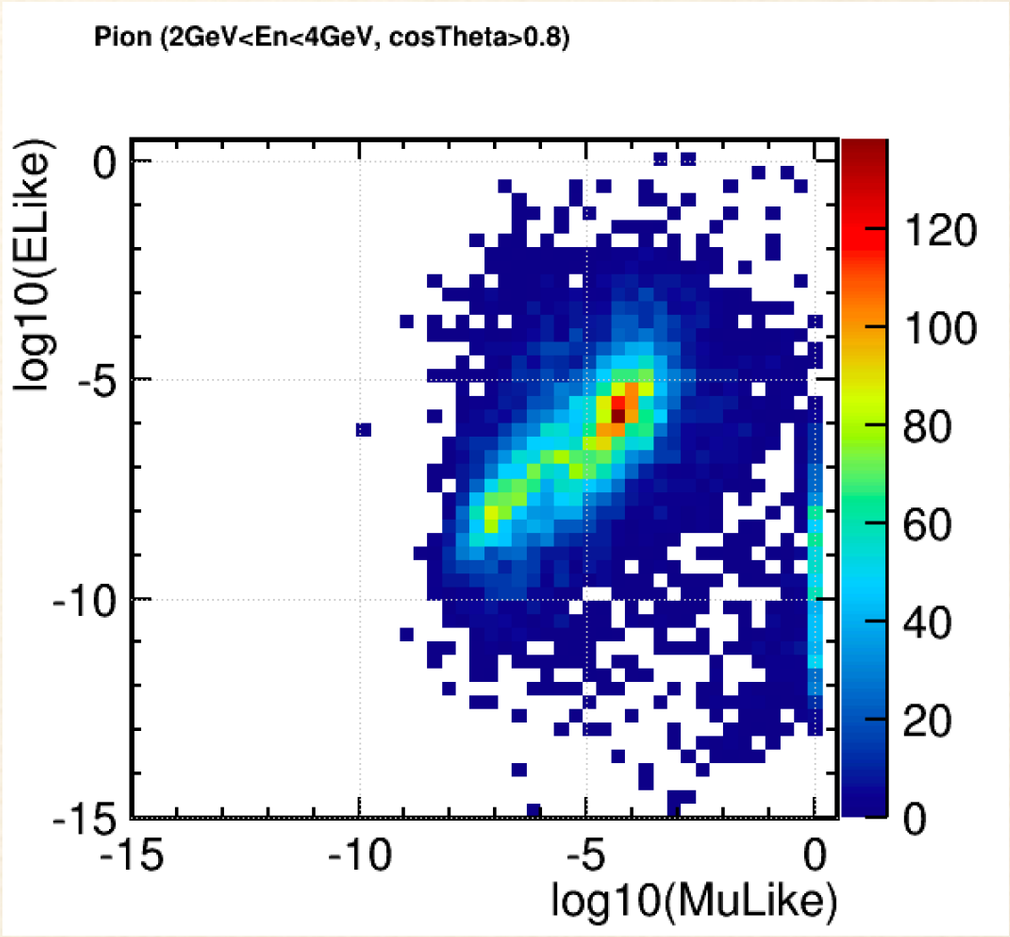
  to characterize clustering performance
- We look into “nice” clusters (efficiency*purity>0.92) and “poor” clusters (efficiency*purity<0.44)



Angular Dependence

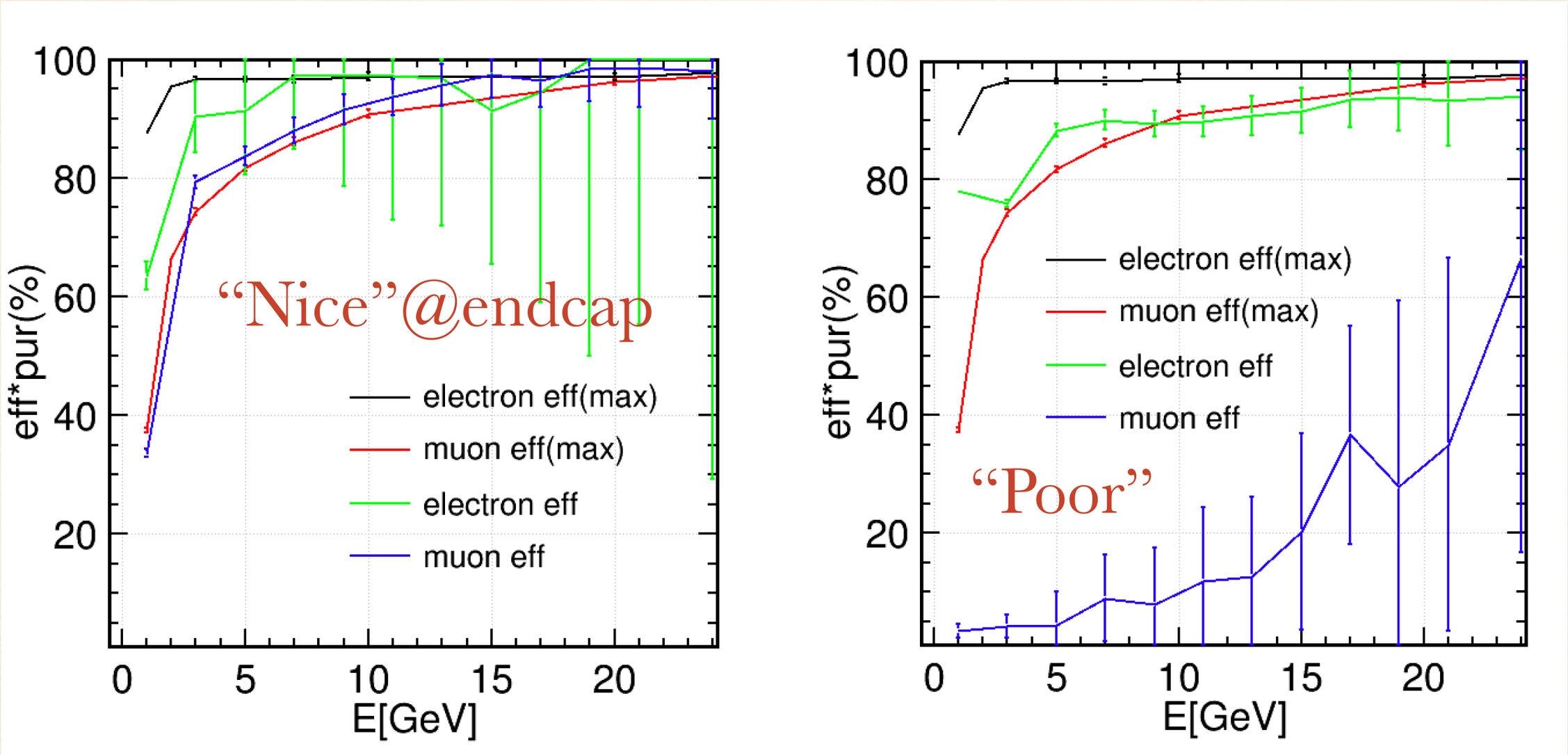
- Low energy pions mixed with muons: better on endcap





Result

- Comparison of lepton identification performance for "nice"/"poor" clusters and the extrapolated performance using single particle results and the statistics (up limit to be achieved)



Tau Identification

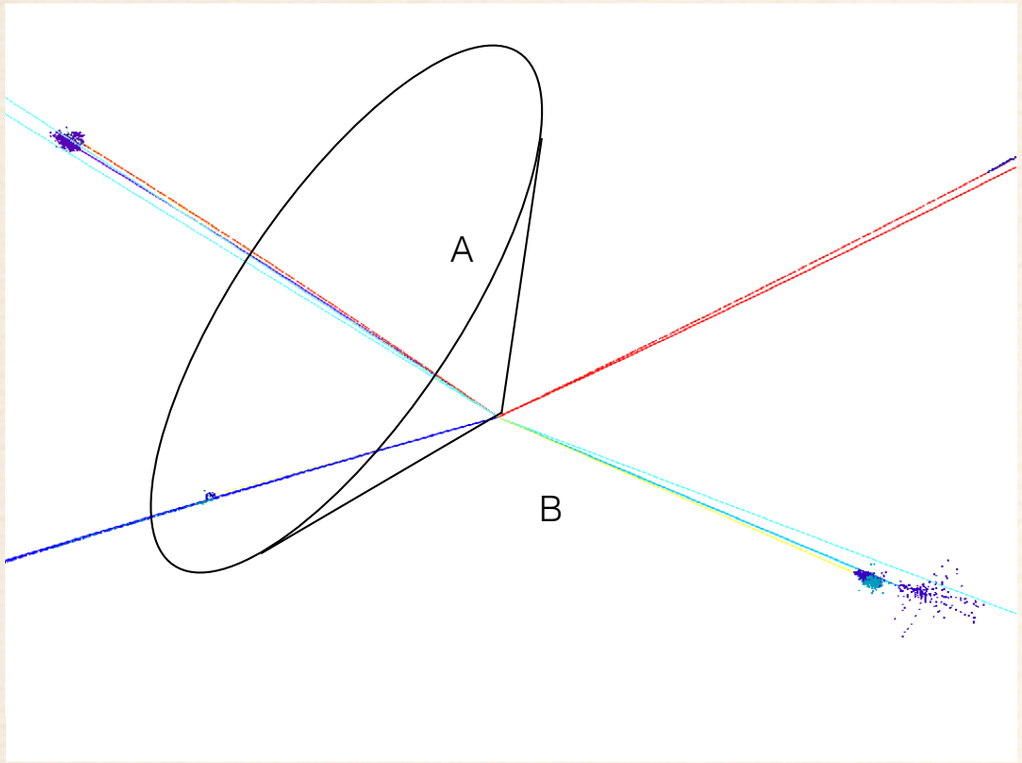
Tau is the heaviest SM lepton - large coupling to Higgs boson Br($H \rightarrow \tau\tau$): 6.27%

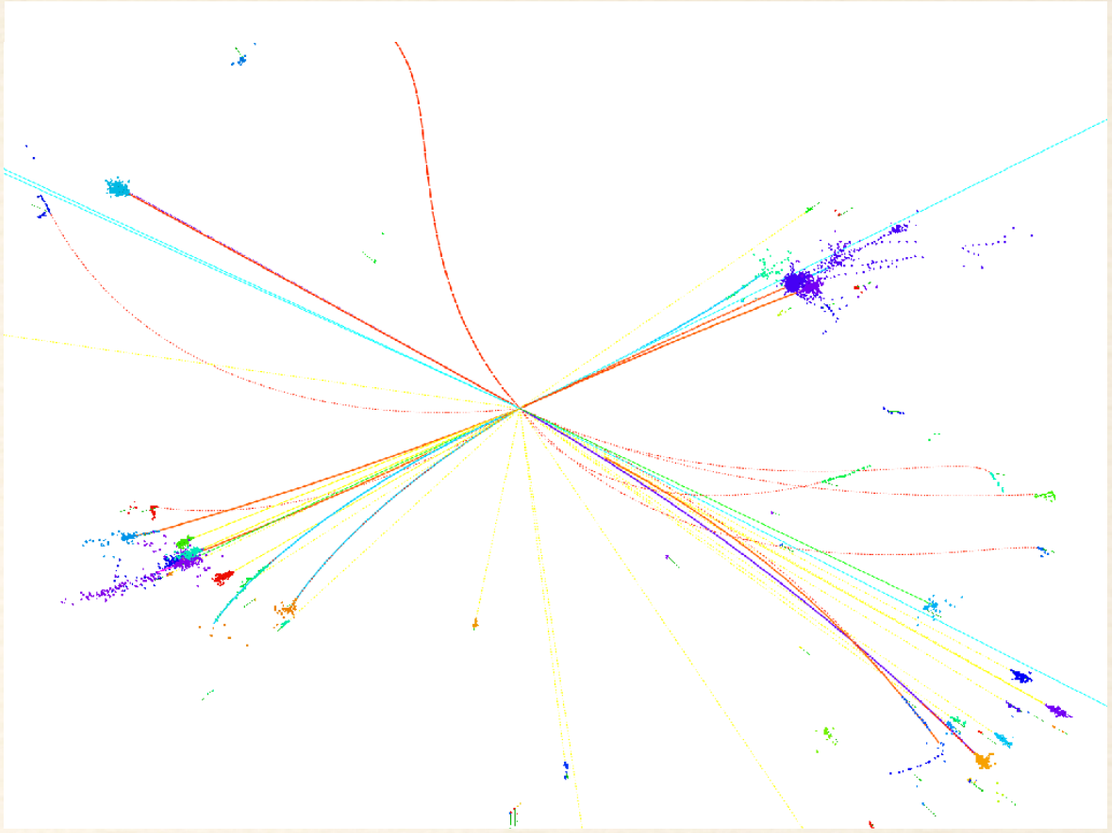
Rich relevant physics

Performance rely on particle separation

Testbed for PFA/Objectives for detector optimization

Event topology

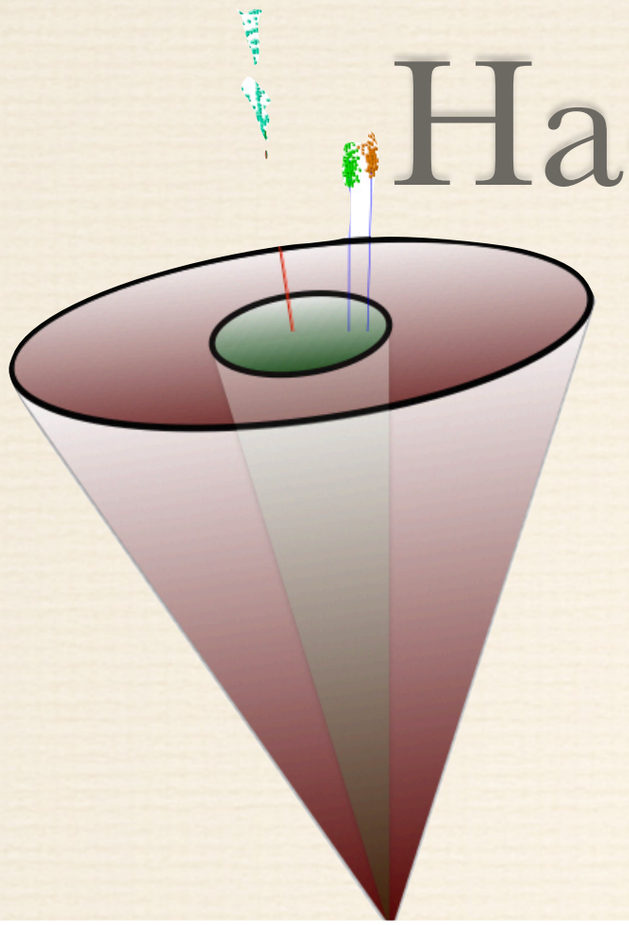


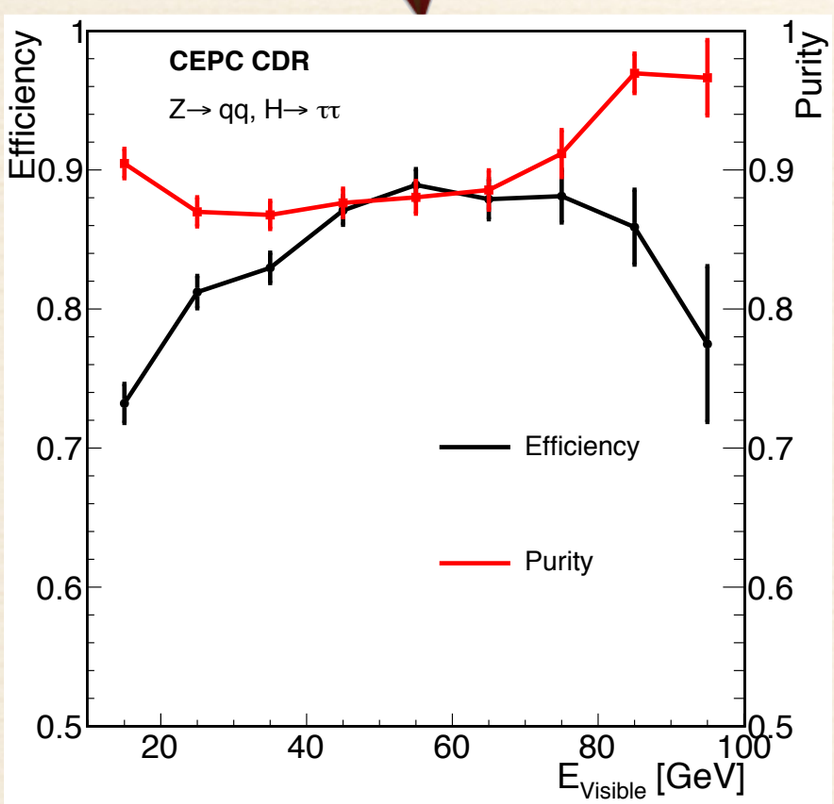


- Veto the two isolate lepton
- Divide the whole space into 2 part
- Use the multiplicity and impact parameter for ττ event selection.
- Fit the ττ mass for signal and background statistics

- qq events selection
- Tau jet reconstruction package: TAURUS
- τ pair selection
- Jet system information
- Fit on impact parameter

Hadronic Channel



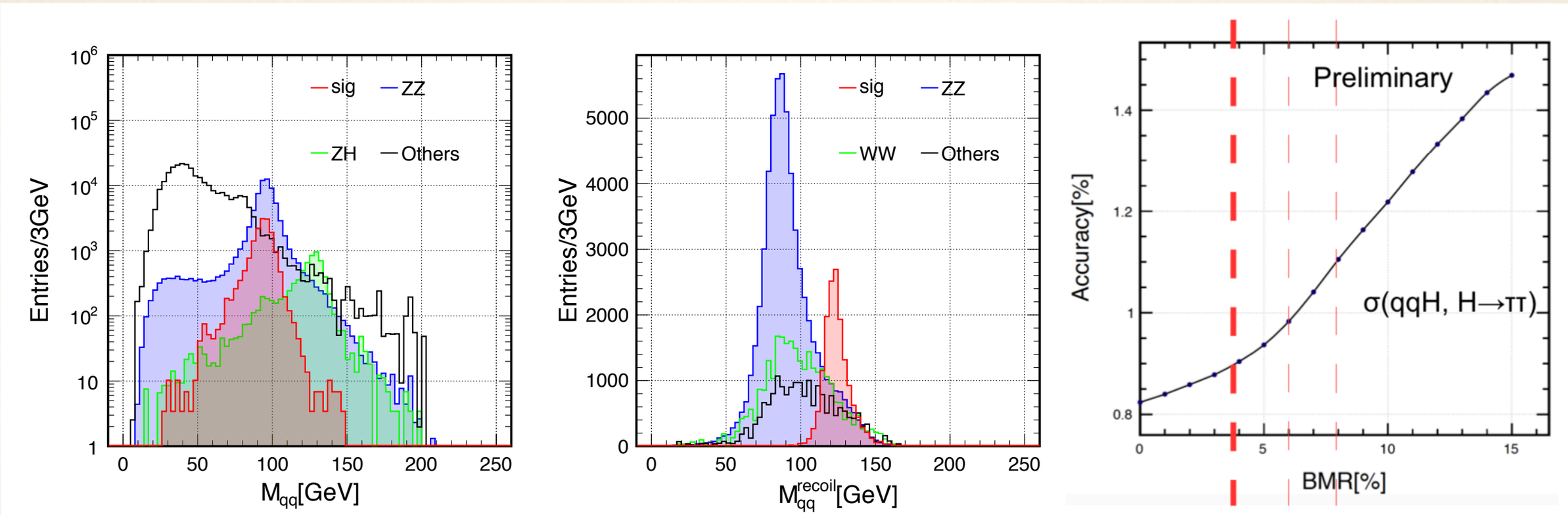


- Double cone based algorithm
  - Find seeds(Tracks with enough energy)
  - Collect particle in two cones
  - Use the multiplicity, energy ratio between two cones, invariant mass for τ tagging
- Event efficiency ~ 60%

Dependence on BMR

	signal	ZZ	ZH conjugation
qq invariant mass [GeV]	91	91	125
qq recoil mass [GeV]	125	91	91

BMR: Boson Mass Resolution



Tau decay mode analysis

	No Trk	1-prong(l)	1-prong(h)	1prong + 1photon	1prong + 2photon	1prong + 3photon	1prong + 4photon	1prong + 5photon	3prong	3prong+ 2photon	other
1-prong(l)	3.58	88.42	3.17	2.58	0.04	0	0	0	0.35	0	Ntrk>1
1-prong(h)	5.90	5.76	78.17	4.49	0.82	0.20	0.06	0	1.16	0	Ntrk>1
1prong + 2photon	2.47	1.31	0.88	29.01	58.34	3.27	0.21	0.01	0.03	1.59	Ntrk>1
1prong + 4photon	1.93	1.23	0.17	1.78	9.75	31.07	45.01	3.24	0	0.19	Ntrk>1
3prong	1.34	1.93	0.34	0.15	0.05	0	0	0	88.44	0.24	Ntrk=2
3prong + 2photon	1.12	1.68	0.14	0.10	0.33	0.10	0.02	0.01	1.08	63.94	Nph=1

Summary

- TMVA based lepton identification has been developed with high efficiency
  - For >2GeV isolate lepton: 99.5%
  - For leptons in jets, degrade due to high statistics, mis-clustering and angular effects
  - "Nice" clusters performance ~ isolate case
- Inclusive τ identification developed with efficiency ~ 80%
  - PFA plays important role in Higgs to ττ analysis (final relative accuracy: 0.8%)
  - Decay modes identification ongoing
  - Better photon/π0 reconstruction needed
- Plan
  - τ in jets
  - CP
  - Exotic decay

Thank you for your attention!