

Identification Performance of Leptons in Jets



ILD Analysis/Software Meeting
Dan YU, Manqi RUAN



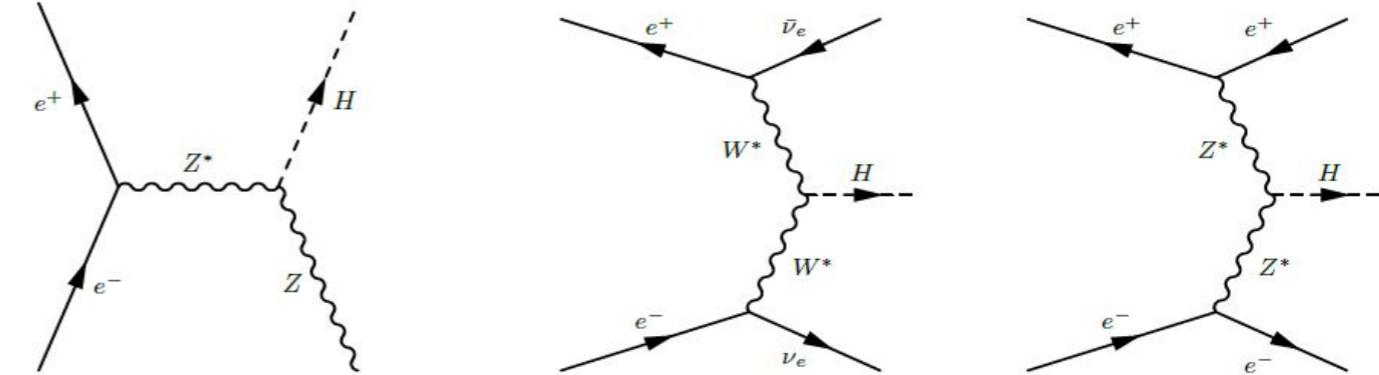
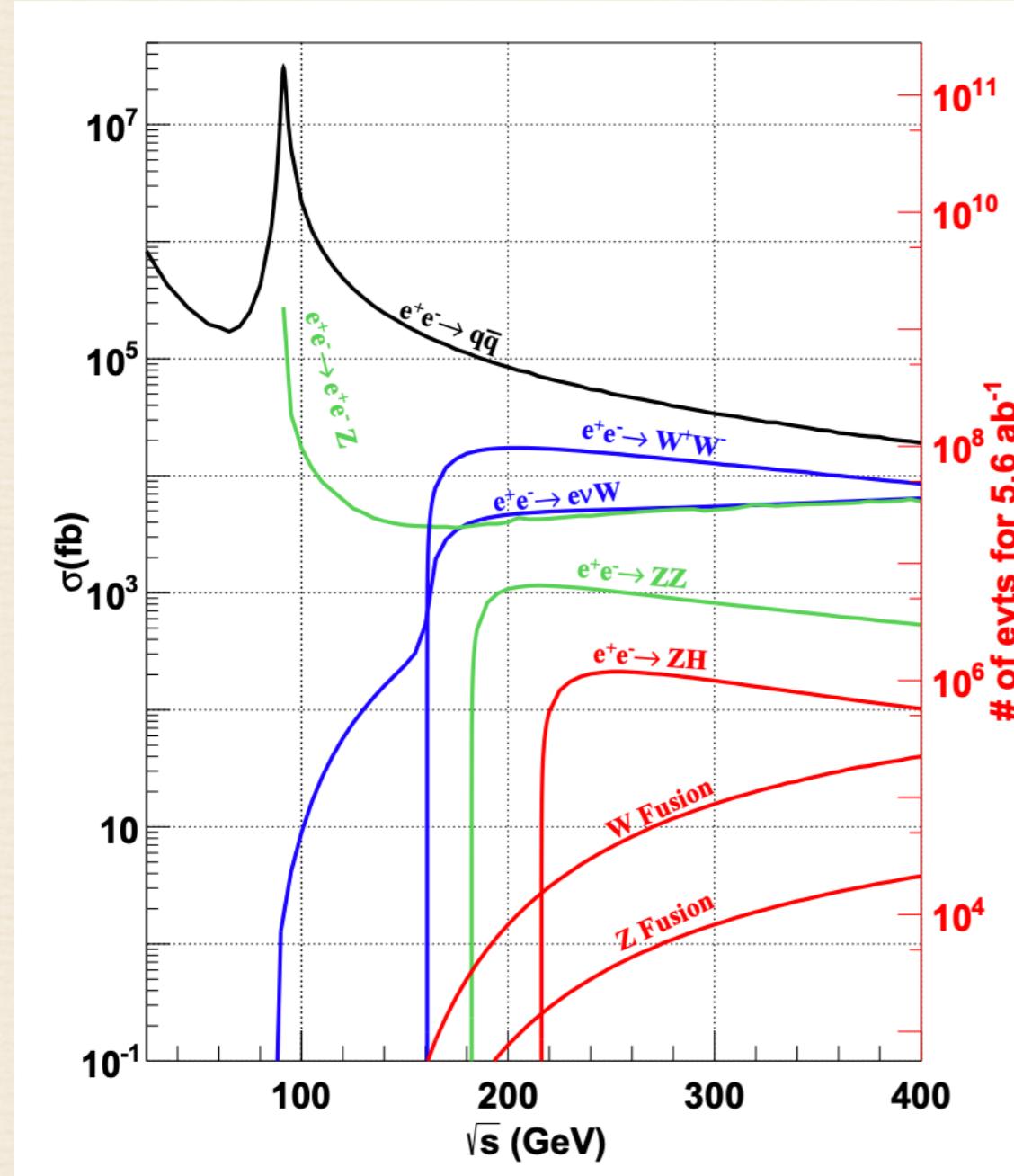
中国科学院高能物理研究所
Institute of High Energy Physics Chinese Academy of Sciences



Plan

- ❖ Introduction
- ❖ Lepton Identification
 - ❖ Single lepton
 - ❖ Lepton in jets
- ❖ τ Identification
- ❖ Summary

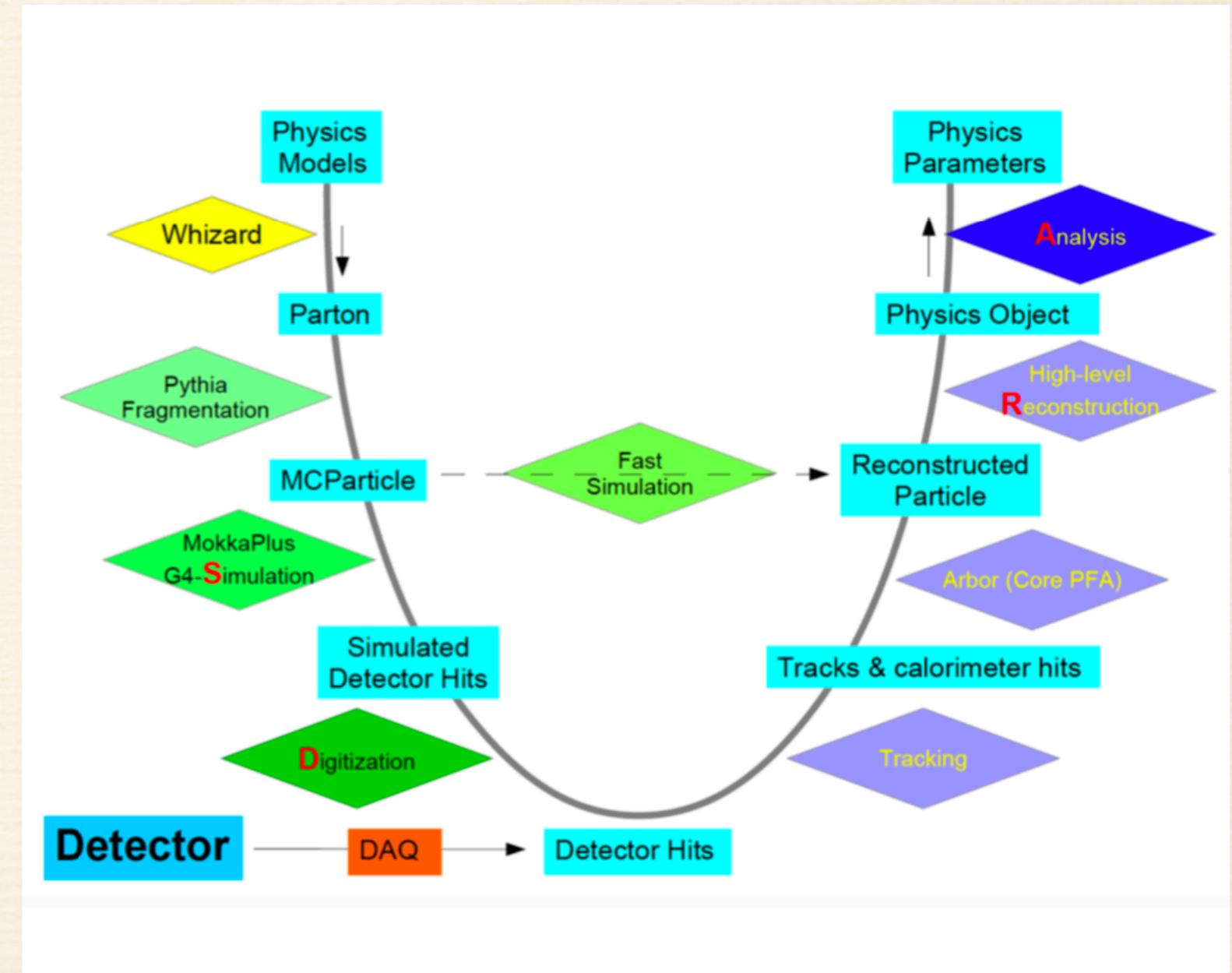
CEPC



- ❖ 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, $\sim 160 \text{ GeV}$, 10^8
 - ❖ W mass/width measurement
 - ❖ PFA Oriented detector
 - ❖ Lepton identification is essential to the precise Higgs measurements/jet flavor tagging and the jet charge measurement
 - ❖ Flavor physics in Z factory

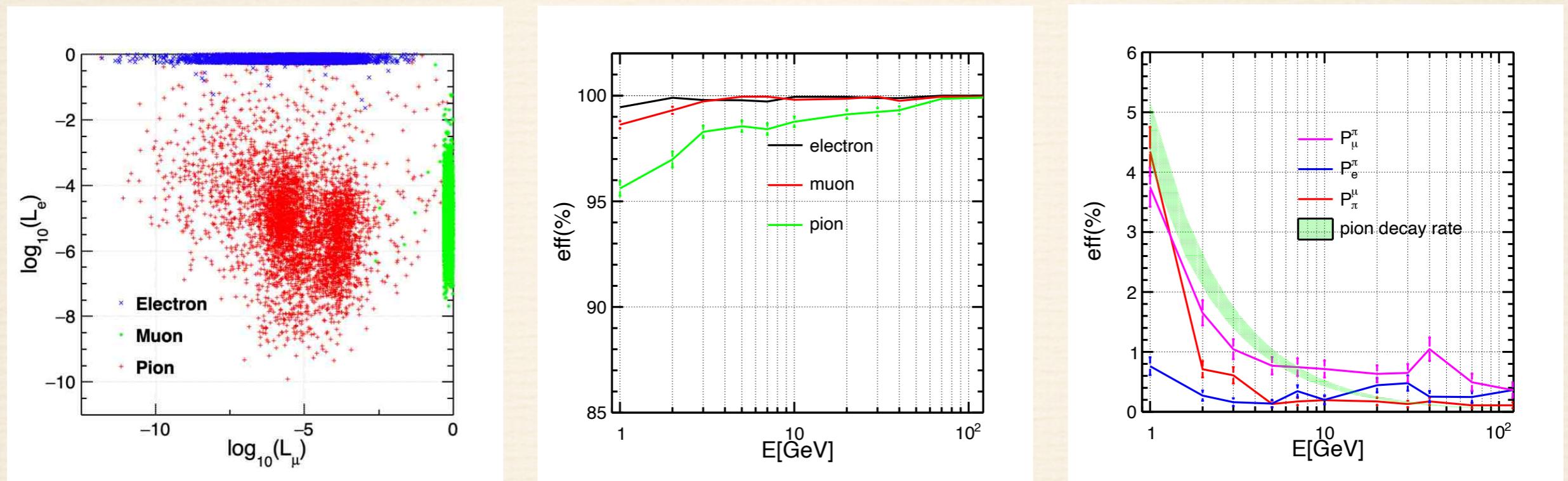
CEPC Full Simulation

- ❖ Software chain developed from ilcsoft
- ❖ CDR Samples:
 - Full simulated Higgs signal
 - small cross-section(<20 fb): simulated to a minimal statistic of 100k
 - 4 fermion background Full simulated
 - 2 fermion background: 20% simulated



Isolate Leptons

- 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 $\sim 98\%$.



Migration Matrix at 40GeV (LICH)

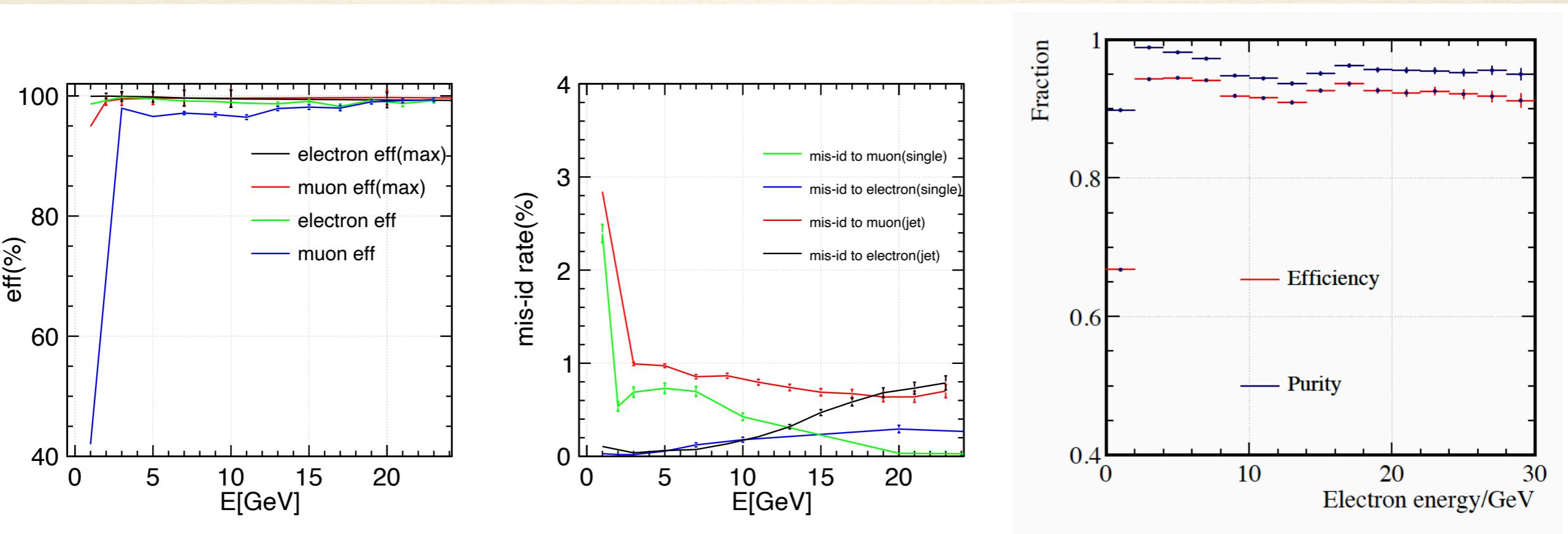
Type	e^- -like	μ^- -like	π^+ -like
e^-	99.71 ± 0.08	< 0.07	0.21 ± 0.07
μ^-	< 0.07	99.87 ± 0.08	0.05 ± 0.05
π^+	0.14 ± 0.05	0.35 ± 0.08	99.26 ± 0.12

Migration Matrix for ALEPH PID (> 2 GeV) (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

Lepton in jets

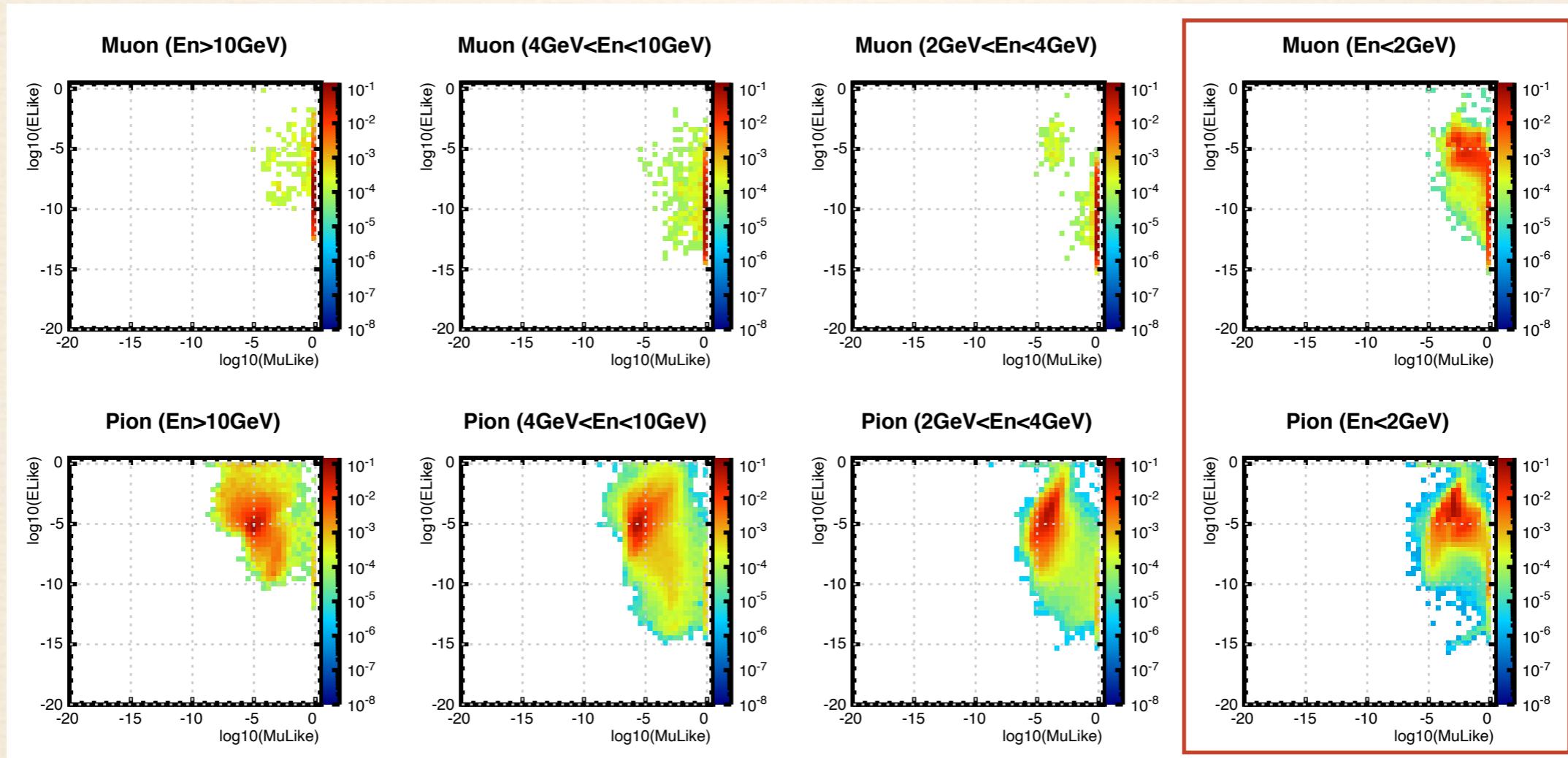
- ❖ The performance for lepton in jets degrades comparing to the single particle results
- ❖ Application: $Bc \rightarrow \tau\nu$



arxiv:2007.08234 by Taifan ZHENG

Likelihood vs Energy

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



Clustering Performance

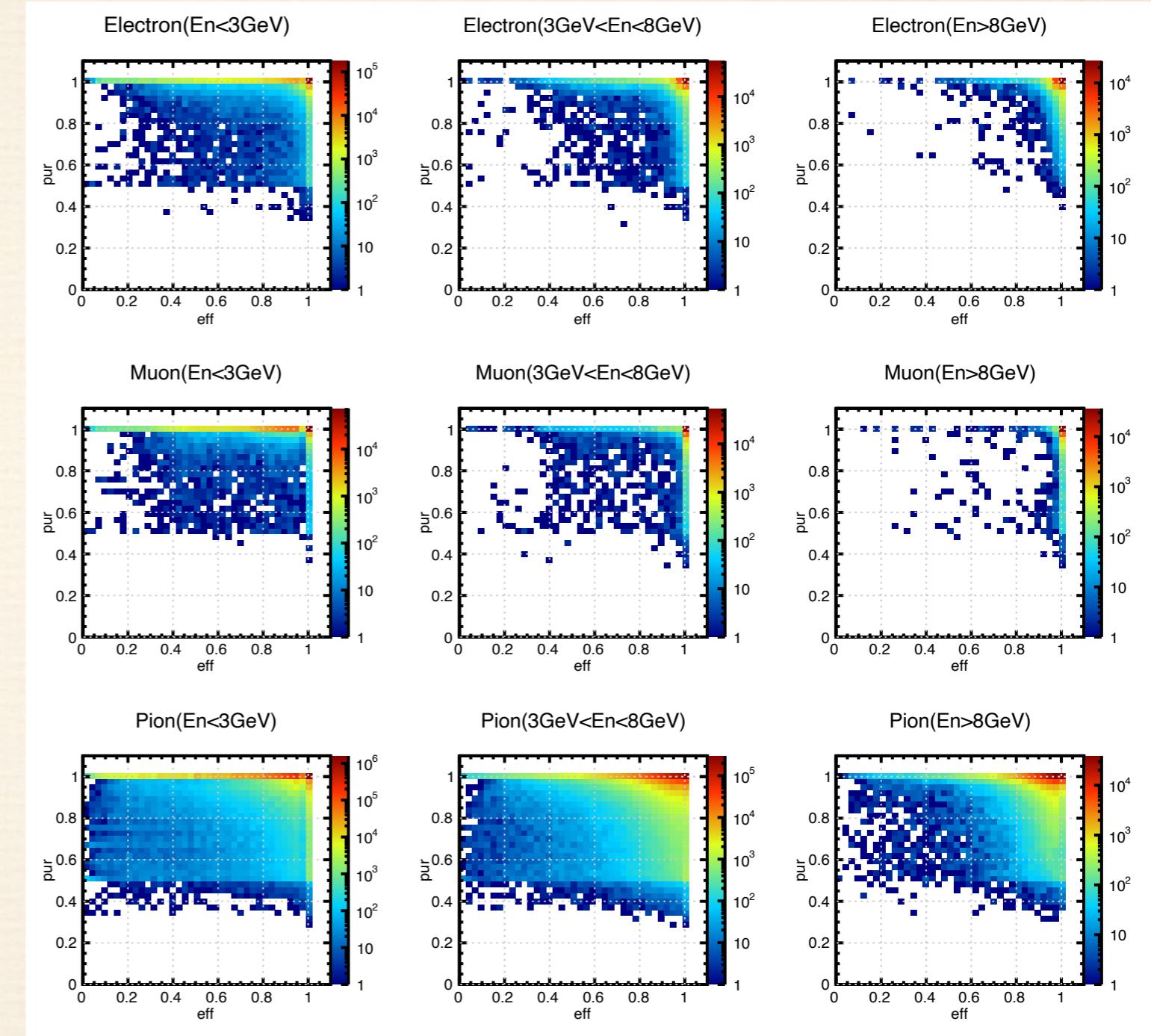
- ❖ Use clustering
 - ❖ efficiency (correct collected hits/particle hits)
 - ❖ purity (correct collected hits/cluster hits)

to characterize clustering performance



Clustering Performance

- ❖ Higher energy, better clustering performance
- ❖ Muon:
 - ❖ 85% perfect
 - ❖ 5% $\text{eff} * \text{pur} < 0.9$
- ❖ Electron:
 - ❖ 64% perfect
 - ❖ 22% $\text{eff} * \text{pur} < 0.9$
- ❖ Pion:
 - ❖ 52% perfect
 - ❖ 24% $\text{eff} * \text{pur} < 0.9$

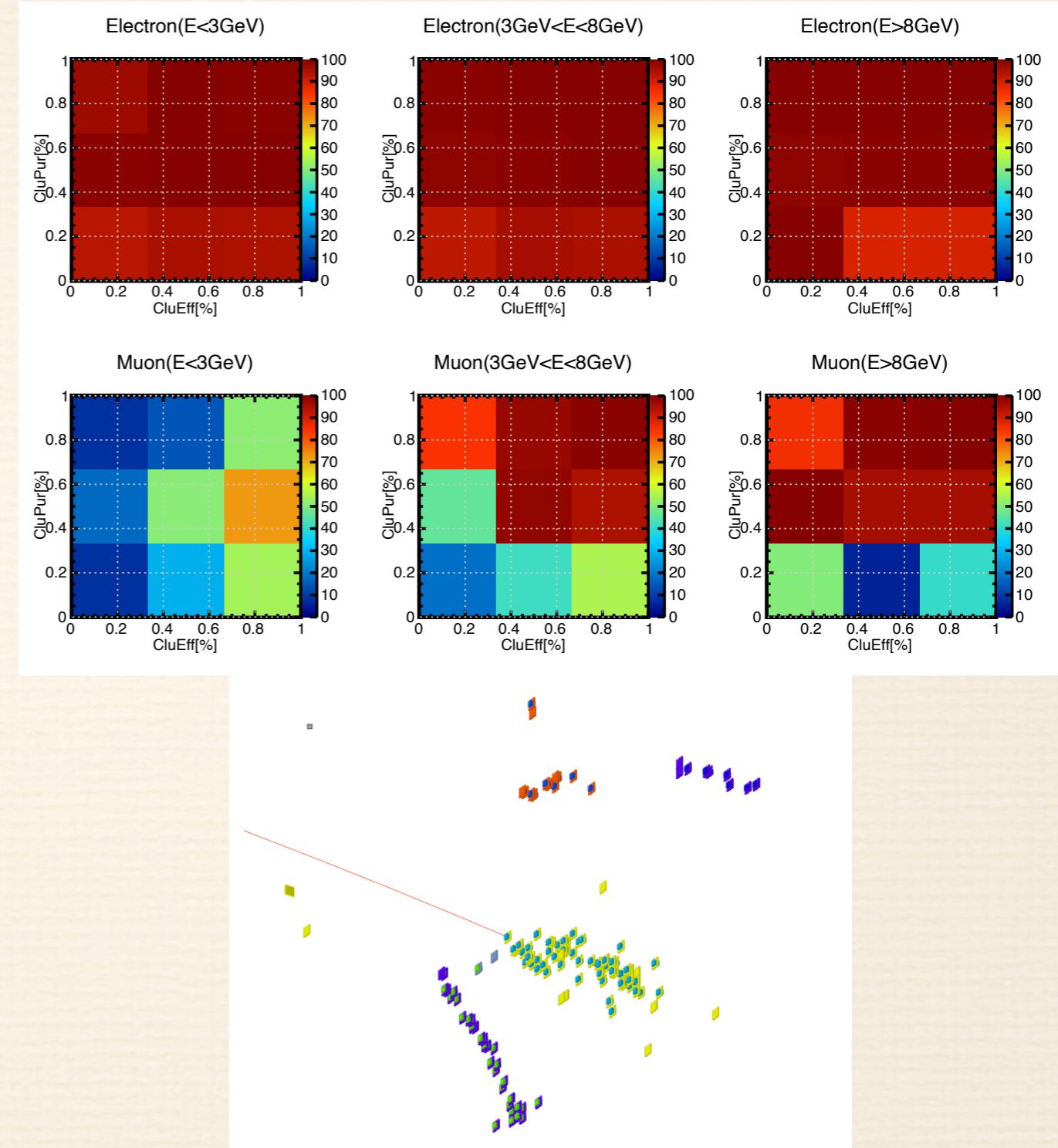


Clustering vs PID

- ❖ Electrons:
 - ❖ low energy: dE/dx dominate
 - ❖ clusters are compact, the splitting clusters still electron-like

- ❖ Muon:
 - ❖ cluster is not MIP-like if mixed with other hits
 - ❖ muon likeliness is lost when the muon cluster splits into small pieces

- ❖ Pion:
 - ❖ likely to be a EM cluster with some branches
 - ❖ more likely to be mis-identified as an electron for lower clustering efficiency

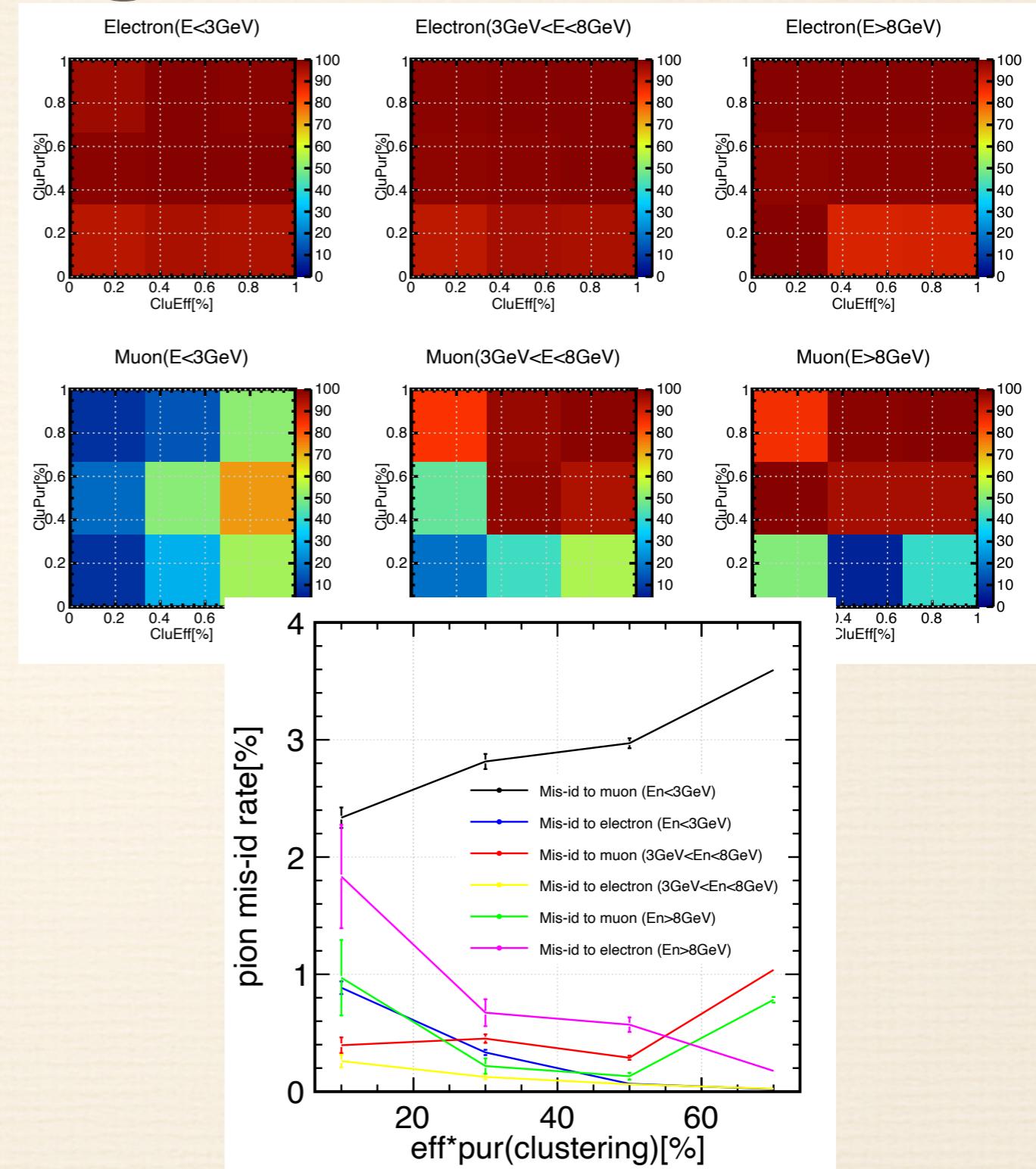


Clustering vs PID

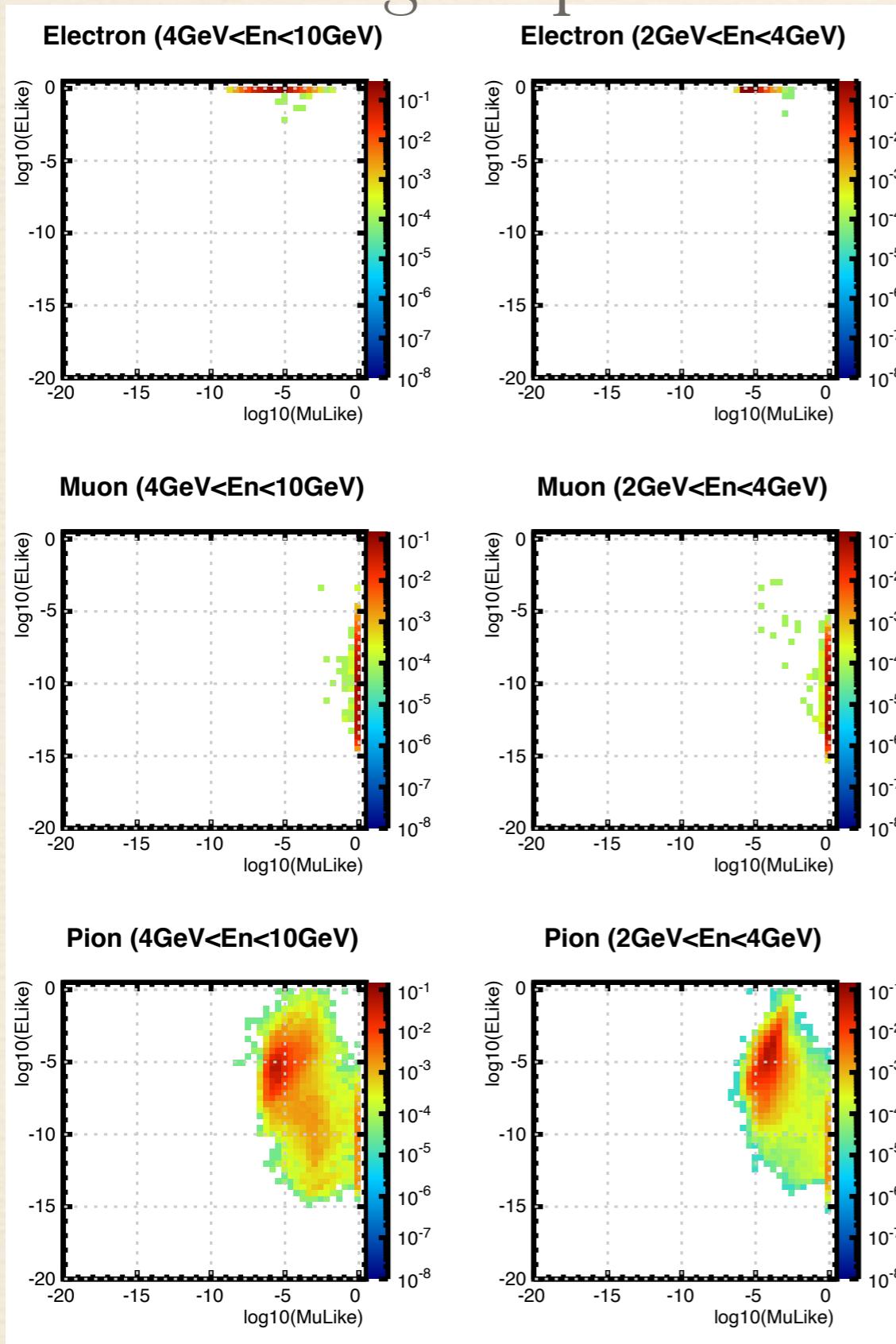
- ❖ Electrons:
 - ❖ low energy: dE/dx dominate
 - ❖ clusters are compact, the splitting clusters still electron-like

- ❖ Muon:
 - ❖ cluster is not MIP-like if mixed with other hits
 - ❖ muon likeliness is lost when the muon cluster splits into small pieces

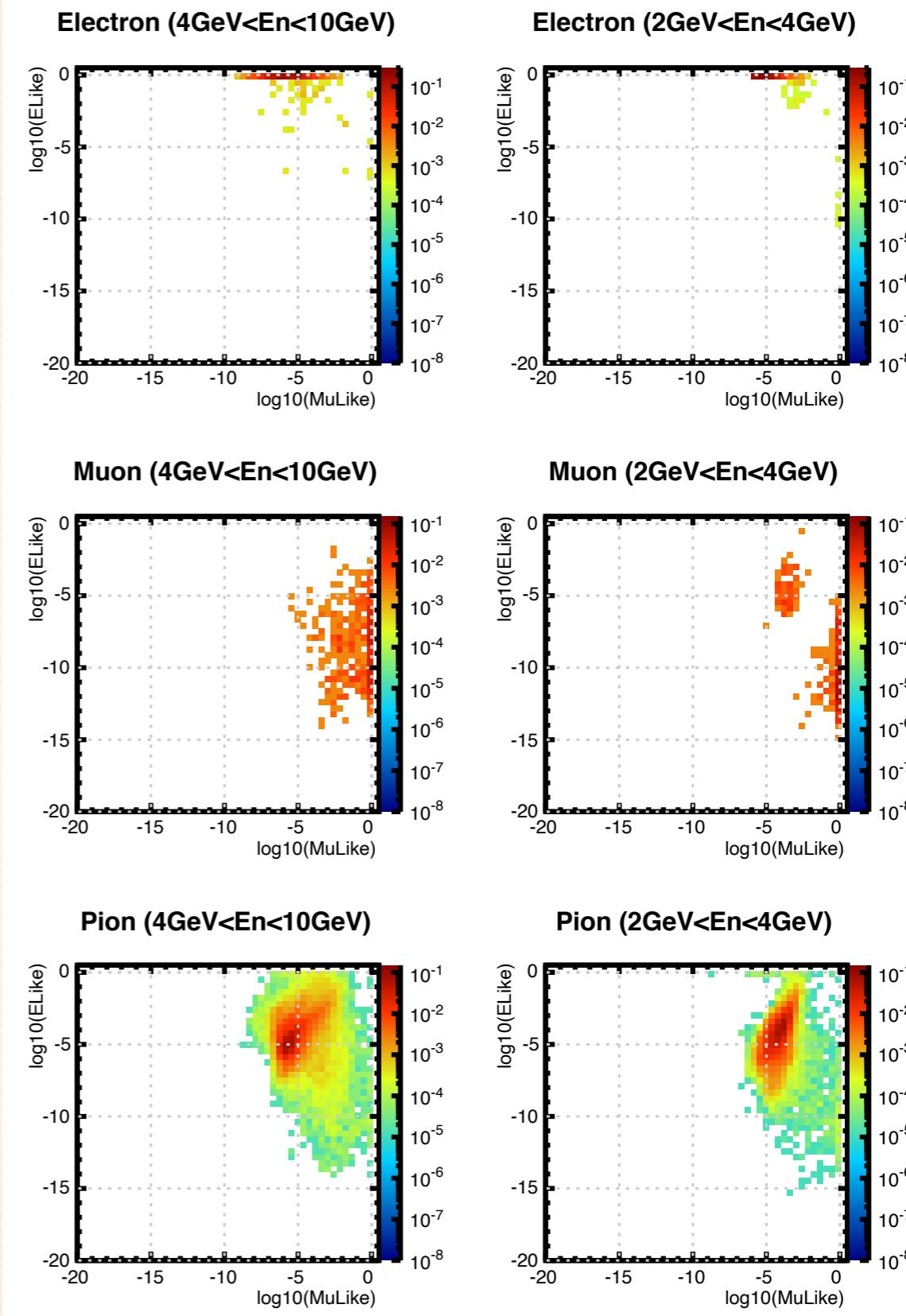
- ❖ Pion:
 - ❖ likely to be a EM cluster with some branches
 - ❖ more likely to be mis-identified as an electron for lower clustering efficiency



Clustering eff*pur=1

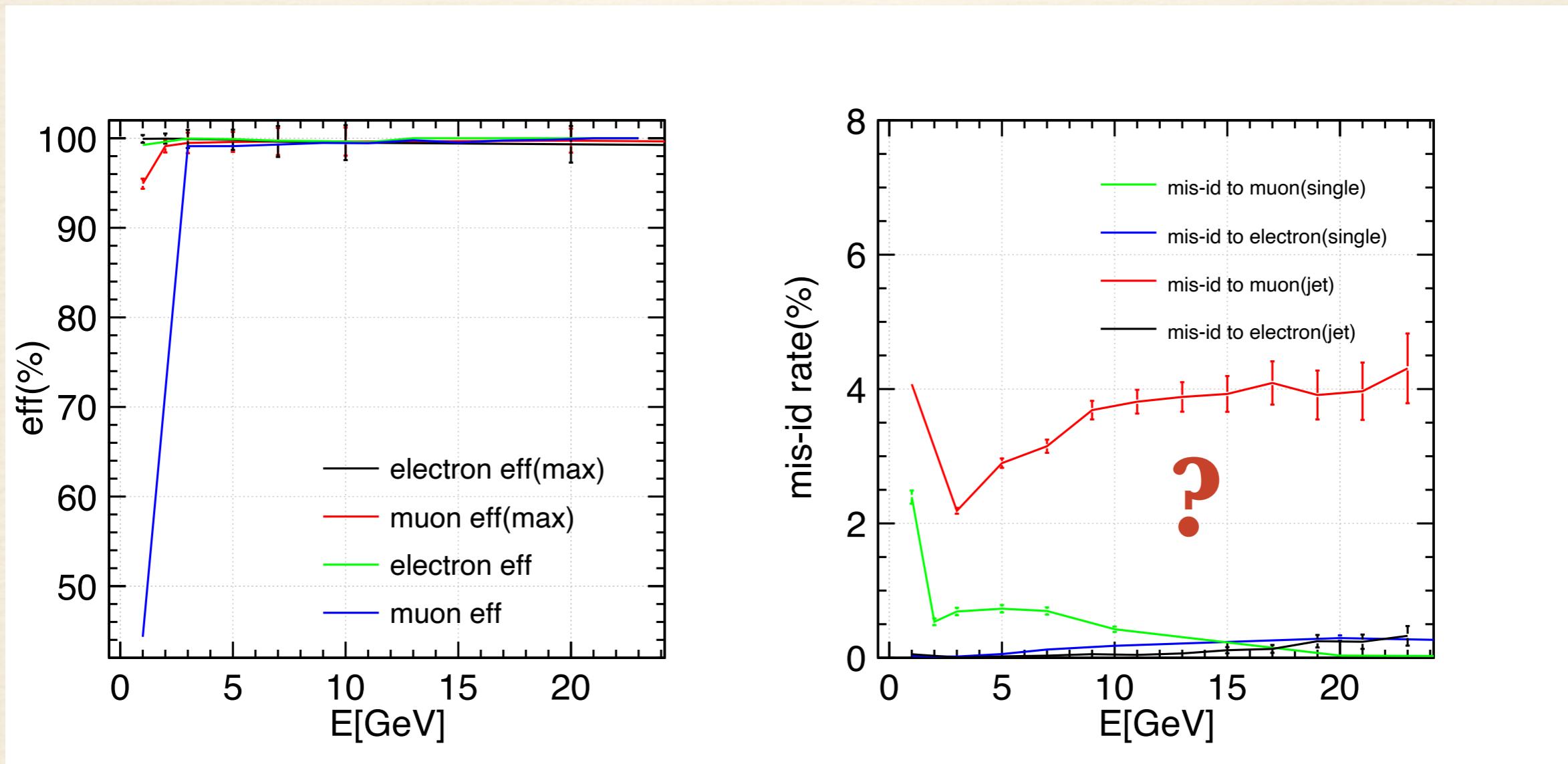


Clustering eff*pur<0.9



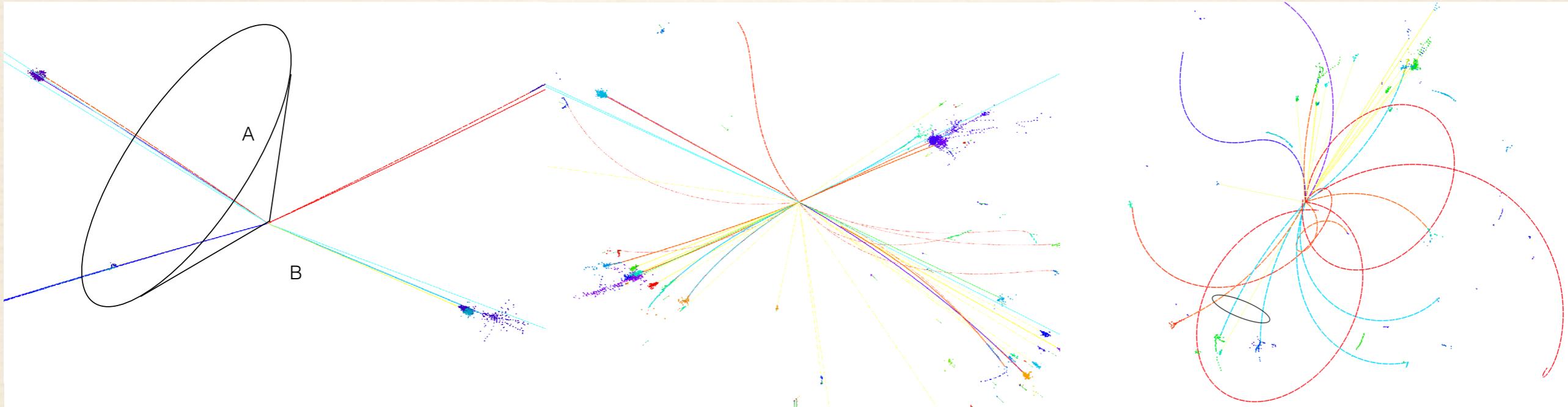
Comparison

- ❖ Comparison of lepton identification performance for perfect clusters and the performance of single particle
- ❖ Pion in jets more likely to be mu-like



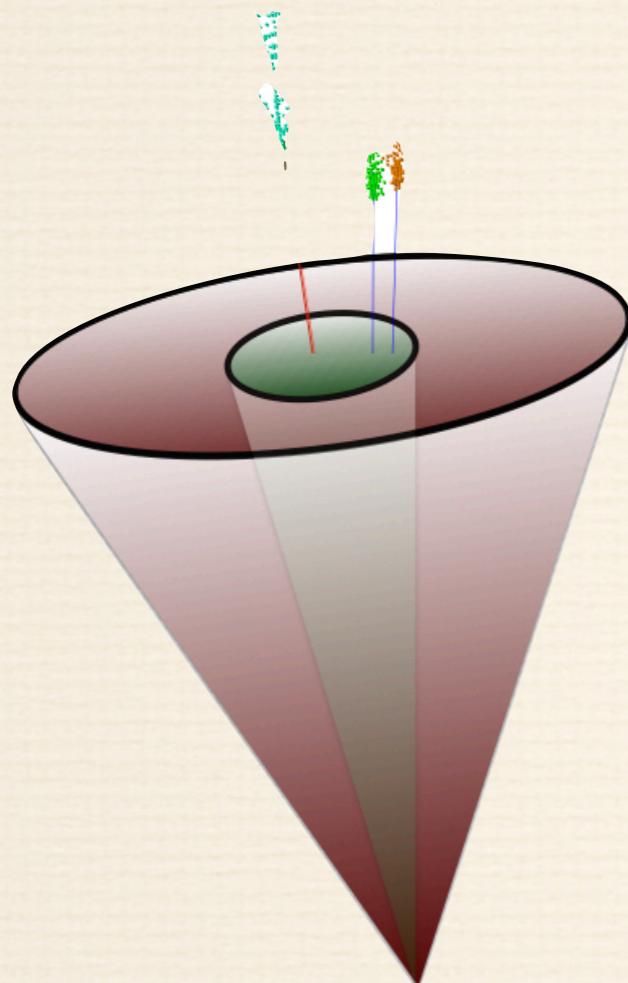
Tau event topology

- ❖ llH channel / $Z \rightarrow \tau\tau$
- ❖ qqH (isolate τ with jets)
- ❖ τ inside jets



- ❖ (Veto the two isolate lepton)
- ❖ Divide the whole space into 2 part
- ❖ Multiplicity & Impact parameter
- ❖ Tau jet reconstruction package: **TAURUS**
- ❖ TAURUS with different parameters

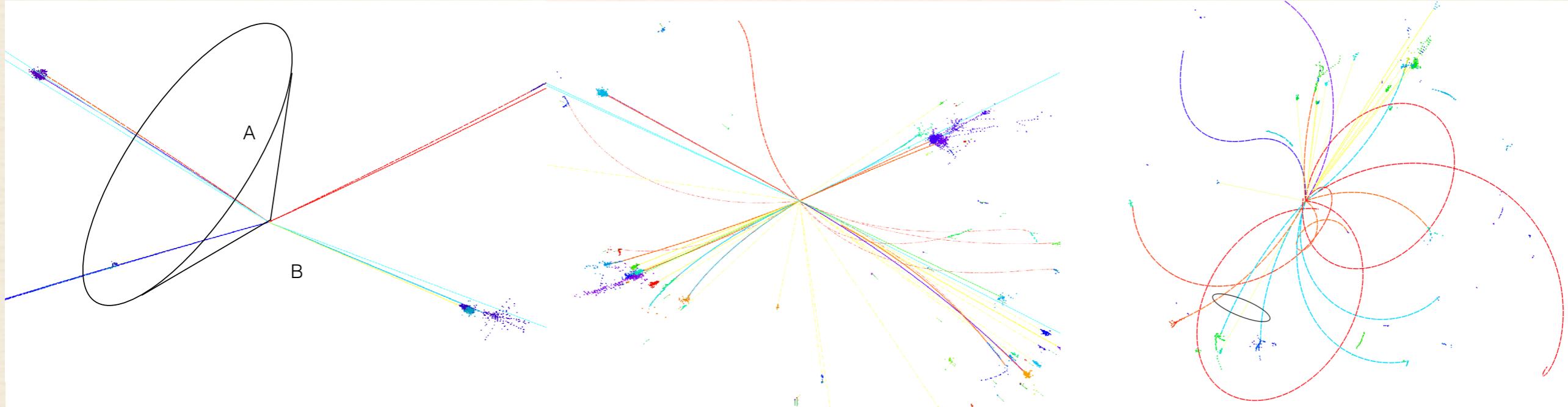
Taurus



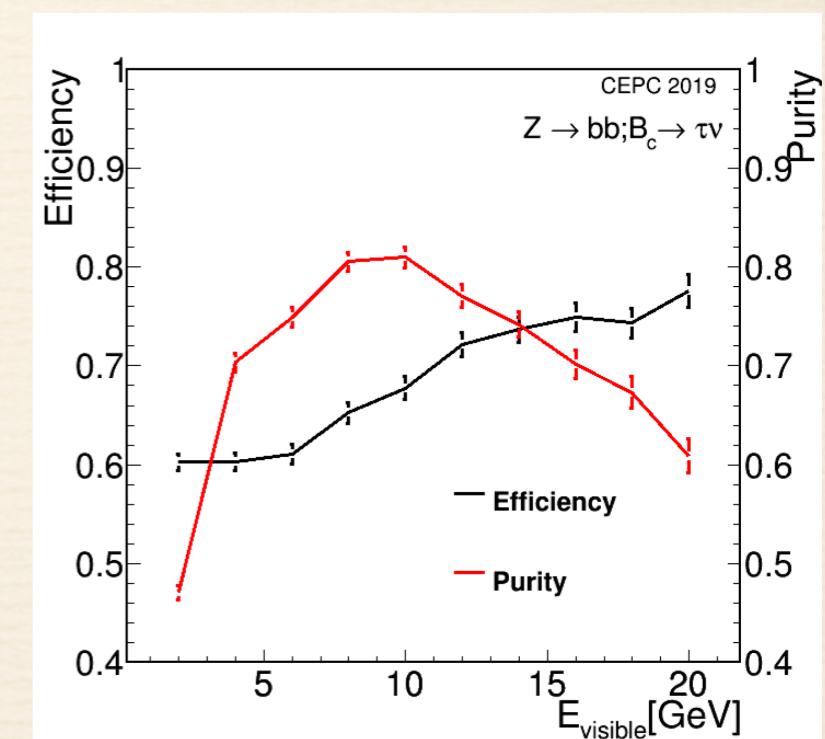
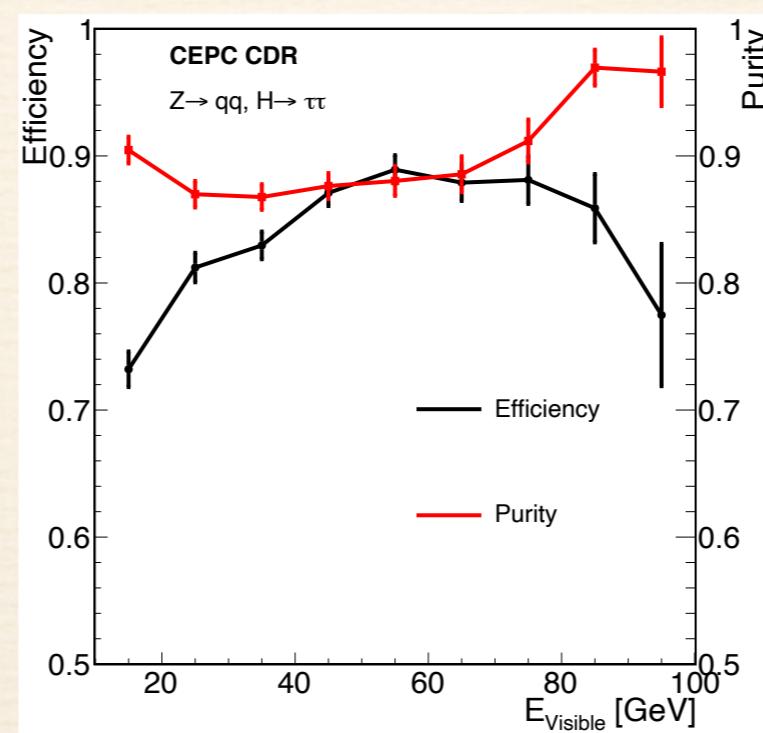
- Double cone based algorithm
 - Find seeds(Tracks with enough energy)
 - Collect particle in two cones
 - Use the multiplicity, energy ratio between two cones, invariant mass for τ tagging

Event topology

- ❖ llH channel / $Z \rightarrow \tau\tau$
- ❖ qqH (isolate τ with jets)
- ❖ τ inside jets



- ❖ (Veto the two isolate lepton)
- ❖ Divide the whole space into 2 part
- ❖ Multiplicity & Impact parameter
- ❖ Efficiency $> 90\%$



Summary

- ❖ TMVA based lepton identification for HGCAL has been developed with high efficiency
 - ❖ For $>2\text{GeV}$ isolate lepton: **99.5%**
 - ❖ For leptons in jets, clustering performance defined (testbed)
 - ❖ At perfect clustering ($\text{eff} \times \text{purity} = 1$), identification performance converge to isolated lepton cases
- ❖ Inclusive τ identification developed
 - ❖ isolate τ efficiency/purity $\sim 80\% / 90\%$
 - ❖ τ in jet efficiency/purity $\sim 70\% / 70\%$
- ❖ Application
 - ❖ Flavor physics: $Bc \rightarrow \tau\nu$
 - ❖ Flavor tag

Thank you for your attention!