Tau decay identification in ILD

Trong Hieu TRAN
Laboratoire Leprince-Ringuet,
Ecole polytechnique, CNRS/IN2P3

Outline:

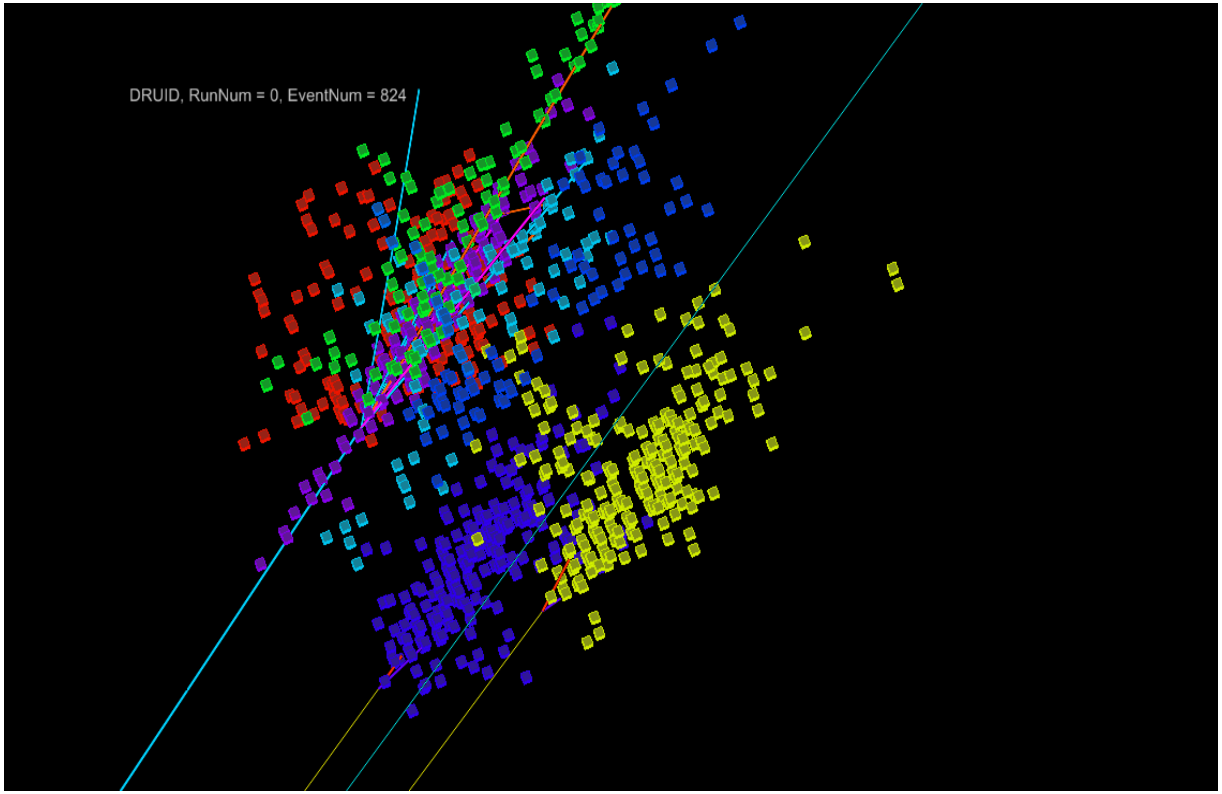
- Tau decay modes
- Samples & analysis procedures
- Application of multivariate analysis technique
  - photon identification
  - tau decay mode identification
- Comparison between ILD models
  - baseline, reduced radius, change of B-field

Introduction

Tau decay reconstruction: a key to probe detector performance

- High mass compared to other leptons
  - → *tau can decay hadronically (65% of cases)*
- Light enough compared to hadrons → decay final state is clean
  → *can provide precise test of the SM (via both leptonic & hadronic sectors)*
- Tau jet is highly collimated → tau identification is a crucial test for ECAL
  - *based essentially on the separation of the photons from π^- interaction or between photons themselves*

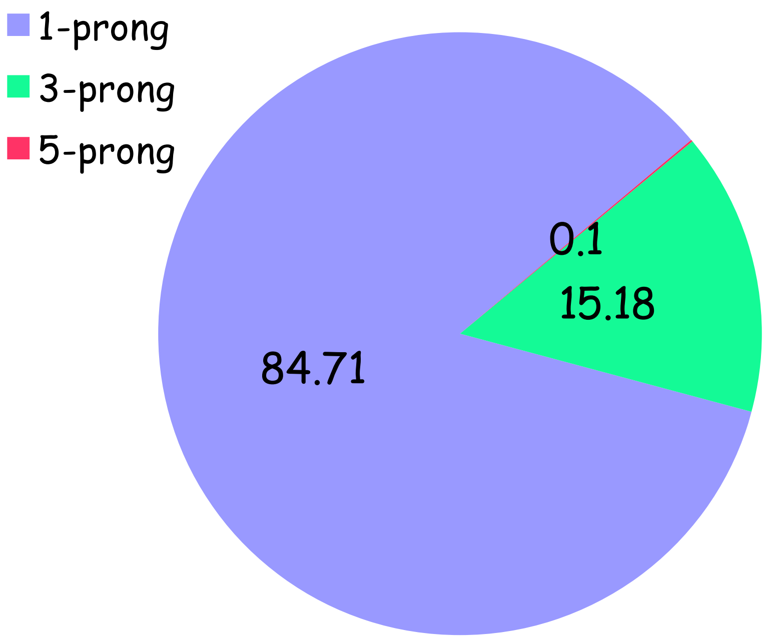
Tau decay studied in this analysis to investigate the possibility of reducing the detector size.



Tau decay modes

Tau jet is compact. Separation of photons in final state is challenging, especially for a small detector.

5, 7 tracks: very infrequent



Topologically: 3 decay modes (1,3,5-prong)
hadronic 1-prong: single charged pion (kaon) and any number of π^0

Final state	Branching fraction
$e^-\bar{\nu}_e\nu_\tau$	$17.85 \pm 0.05\%$
$\mu^-\bar{\nu}_\mu\nu_\tau$	$17.36 \pm 0.05\%$
$\pi^-\nu_\tau$	$10.91 \pm 0.07\%$
$\rho^-\nu_\tau$ ($\rho^- \rightarrow \pi^-\pi^0$)	$25.52 \pm 0.10\%$
$a_1^-\nu_\tau$ ($a_1^- \rightarrow \pi^-\pi^0\pi^0$)	$9.27 \pm 0.12\%$
$a_1^-\nu_\tau$ ($a_1^- \rightarrow \pi^-\pi^+\pi^-$)	$8.99 \pm 0.06\%$
24 other modes	10.10%

Branching fraction of main decays

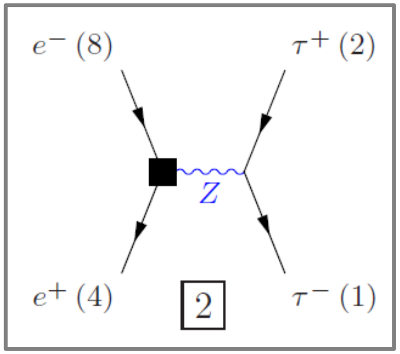
This analysis:
consider only one charged pion decay

$\pi^-\nu_\tau$	0 photon
$\rho^-\nu_\tau$ ($\rho^- \rightarrow \pi^-\pi^0$)	2 photons
$a_1^-\nu_\tau$ ($a_1^- \rightarrow \pi^-\pi^0\pi^0$)	4 photons

Samples

DBD generators $e^+e^- \rightarrow Z \rightarrow \tau^-\tau^+$
at 250 GeV C.M. energy
(mixed with $e^+e^- \rightarrow Z \rightarrow \mu^-\mu^+$
→ preselection of τ events using generator information)

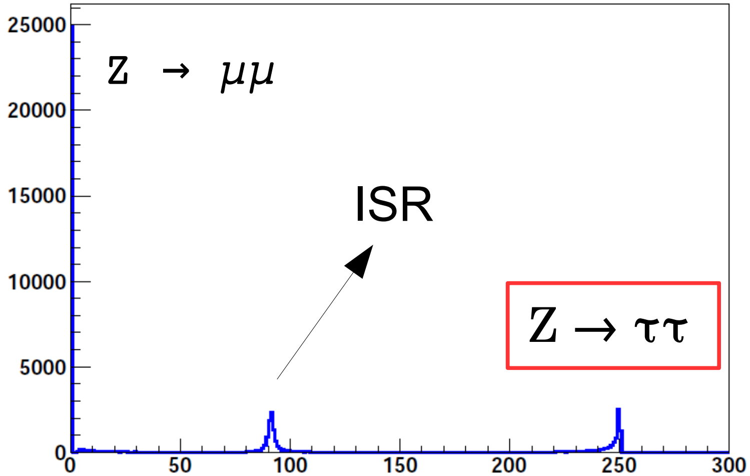
τ energy ~ 125 GeV



This analysis:

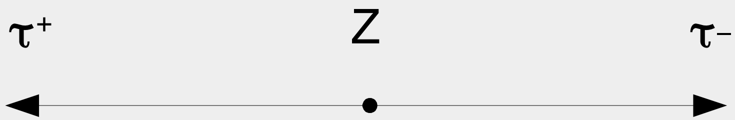
consider only one charged pion decay

$\pi^-\nu_\tau$	0 photon
$\rho^-\nu_\tau$ $(\rho^- \rightarrow \pi^-\pi^0)$	2 photons
$a_1^-\nu_\tau$ $(a_1^- \rightarrow \pi^-\pi^0\pi^0)$	4 photons



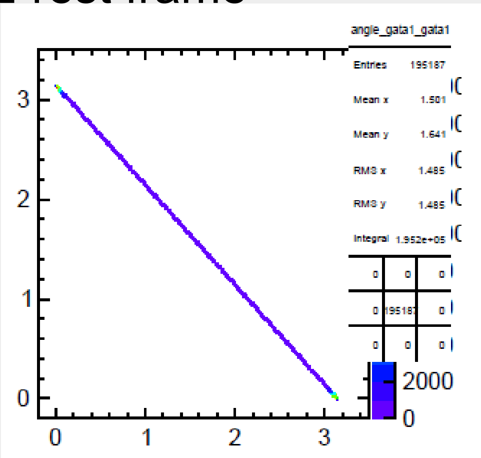
Tau-tau invariant mass

- Two independent decays are used (double statistics)



The two tau's are back-to-back in the Z-rest frame

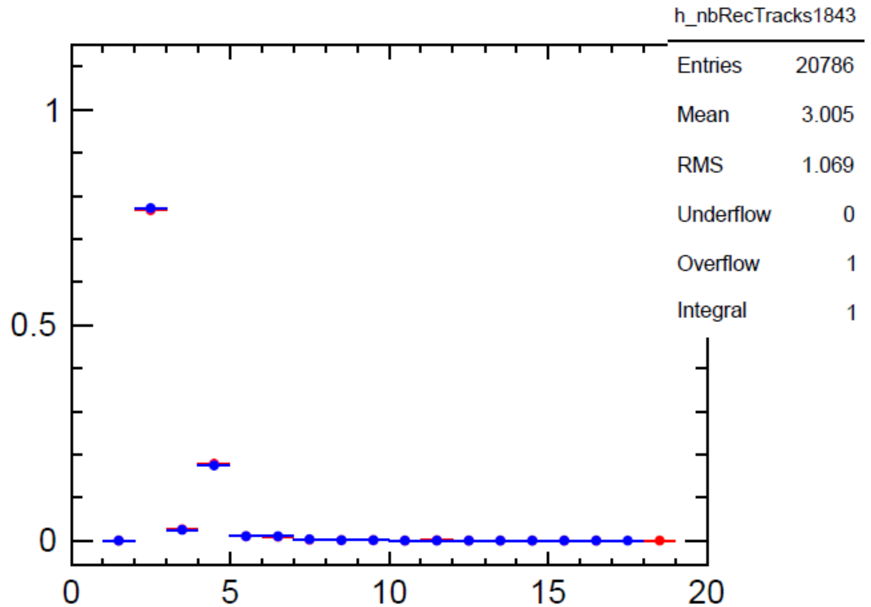
Angle cluster-tau1
vs angle cluster-tau2



Simulation & reconstruction

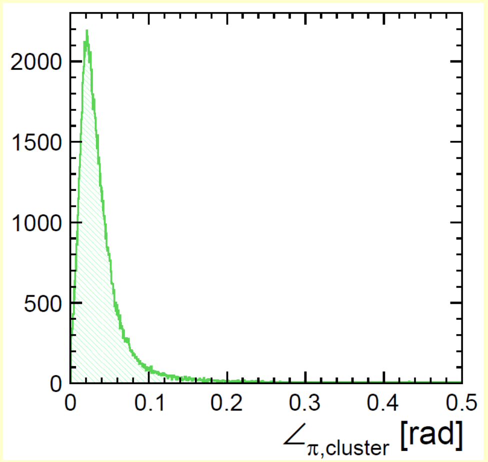
Softwares
Ilcsoft v01-17-06, Mokka-08-04
Garlic v3.0.3

- Baseline ILD design (DBD): **SiW ECAL**, $\mathbf{R_{ECAL}^{inner} = 1843\ mm}$ (ILD_o2_v05)
- Alternative setups: $\mathbf{R_{ECAL}^{inner} = 1615, 1450\ mm}$
- Reduced TPC radius
  → ECAL, HCAL, Yoke, … radii are reduced
- Keep same aspect ratio: Radius/Length
  (for a reduced radius, the length is reduced as well)
- Other configurations unchanged
  (cell size **5x5 mm²**, thicknesses)



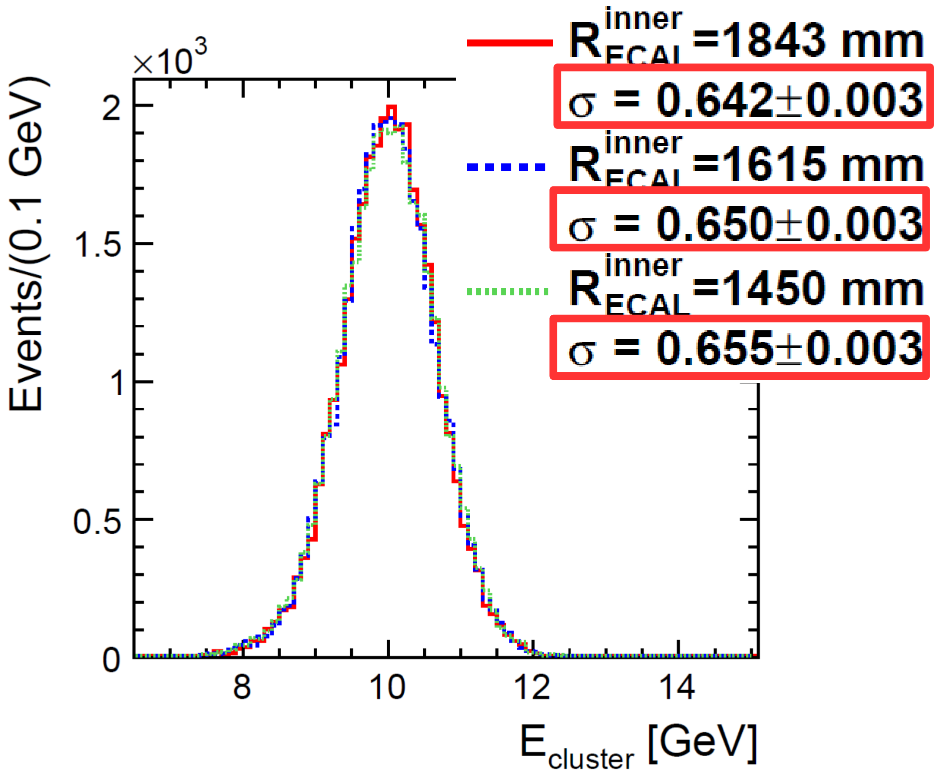
Nb of reconstructed tracks with baseline ILD and with reduced TPC radius
→ **no difference**

- Garlic (v3.0.3) is used for photon reconstruction
- Preselection:
  - preselection based on generator info: choose decays with only one charged pion
  - $|\cos(\theta)|^{tau} < 0.7$
  - sample with **one and only one reconstructed track** in tau direction



Calibration

- Single photon samples are used for determination of calibration constant
- Calibration done for each of ECAL radii
- Very small difference observed for different sizes



Corresponding calibration constants

46.58

46.84

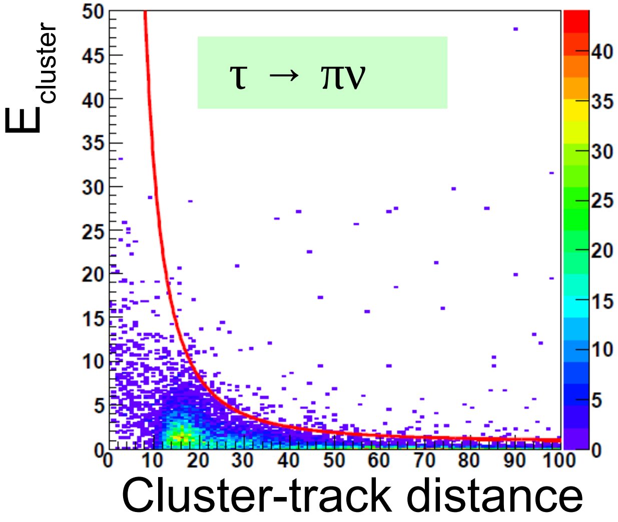
46.85

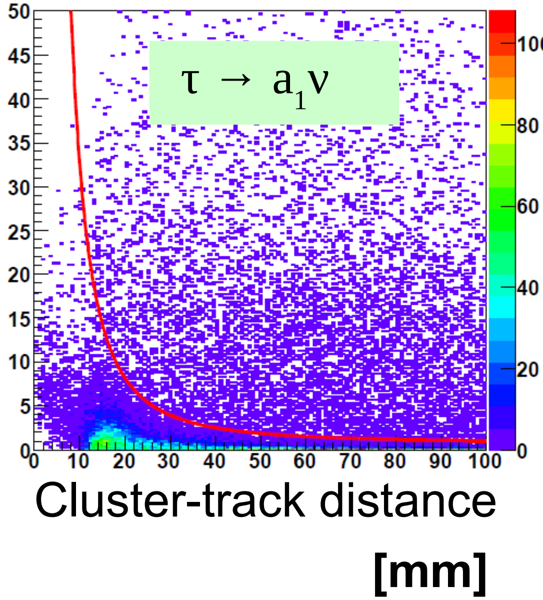
Performance for single particle depends only on intrinsic calorimeter resolution + dead zones

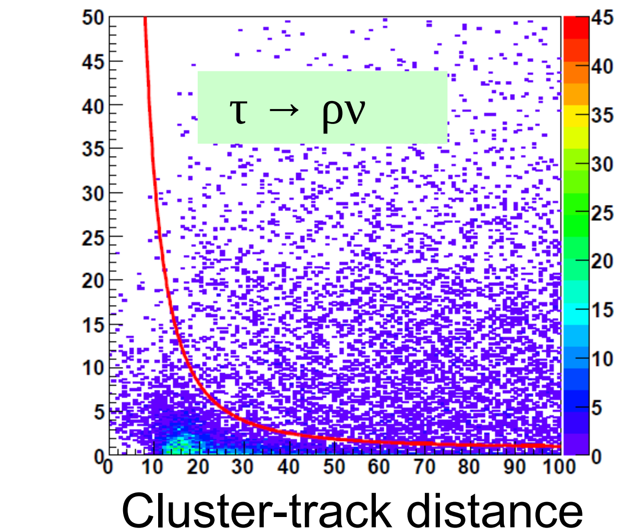
Reconstructed photon energy using Garlic (simulated at 10 GeV)

Photon selection:

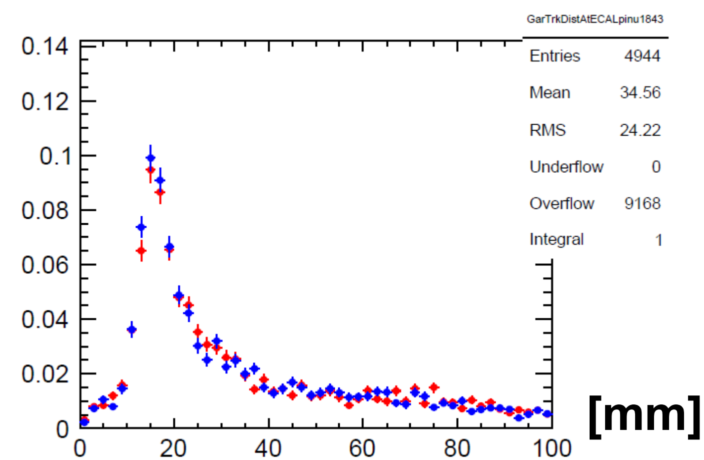
photonE vs distance to track







Aim to remove fake clusters at low energy or from pion interaction



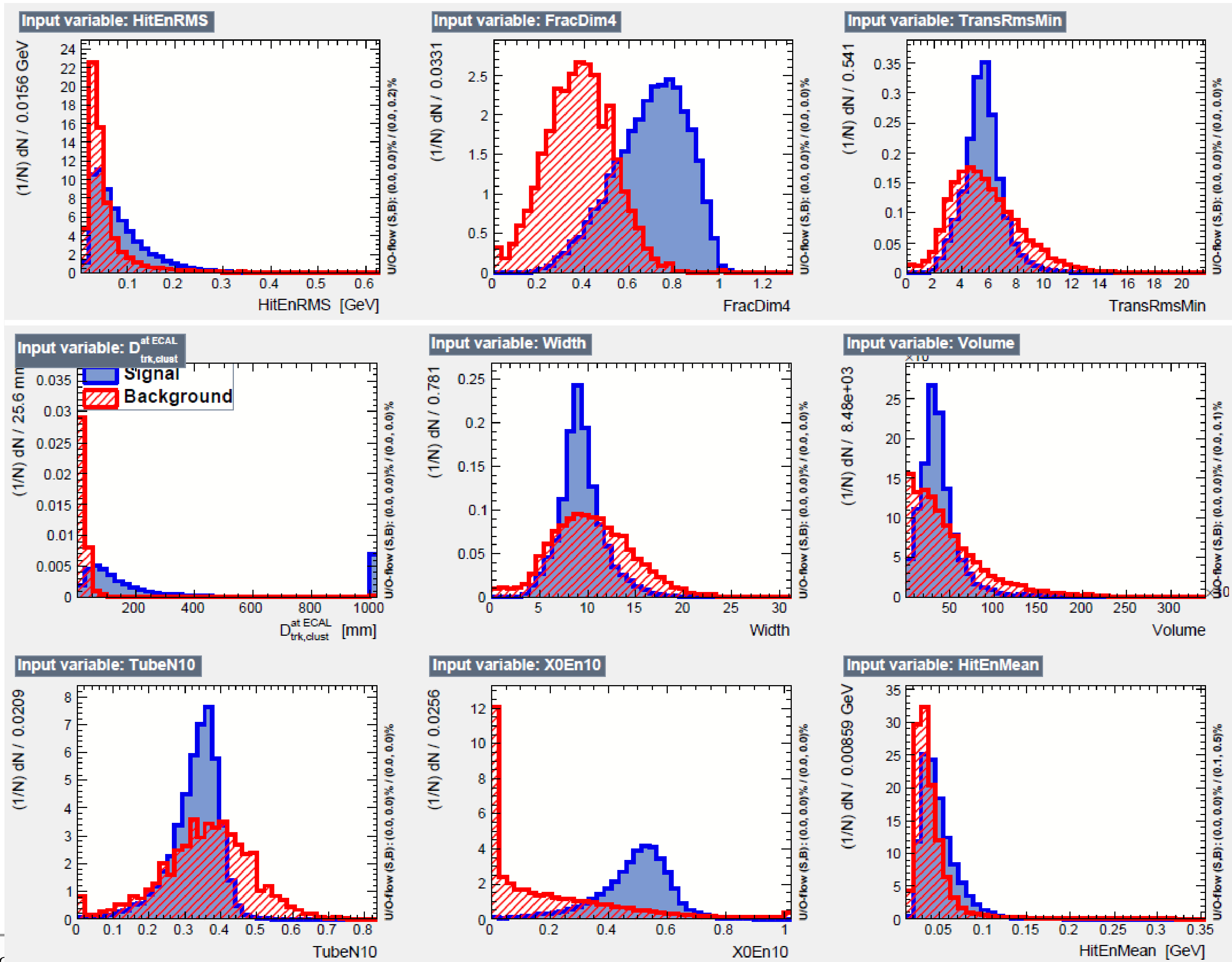
Comparison: **track-cluster distance** at ECAL surface
1843 vs 1450

Clusters in track direction are already removed by Garlic

Photon identification

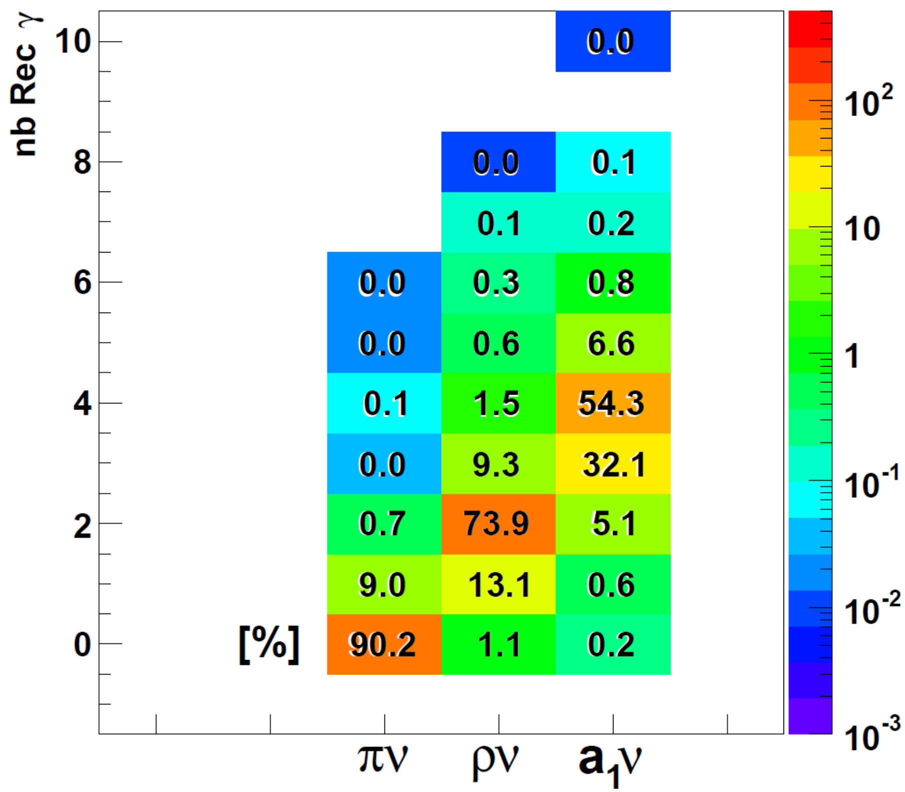
MVA Level 1

- **Boosted Decision Tree** with gradient boost (BDTG) as implemented in TMVA is used.

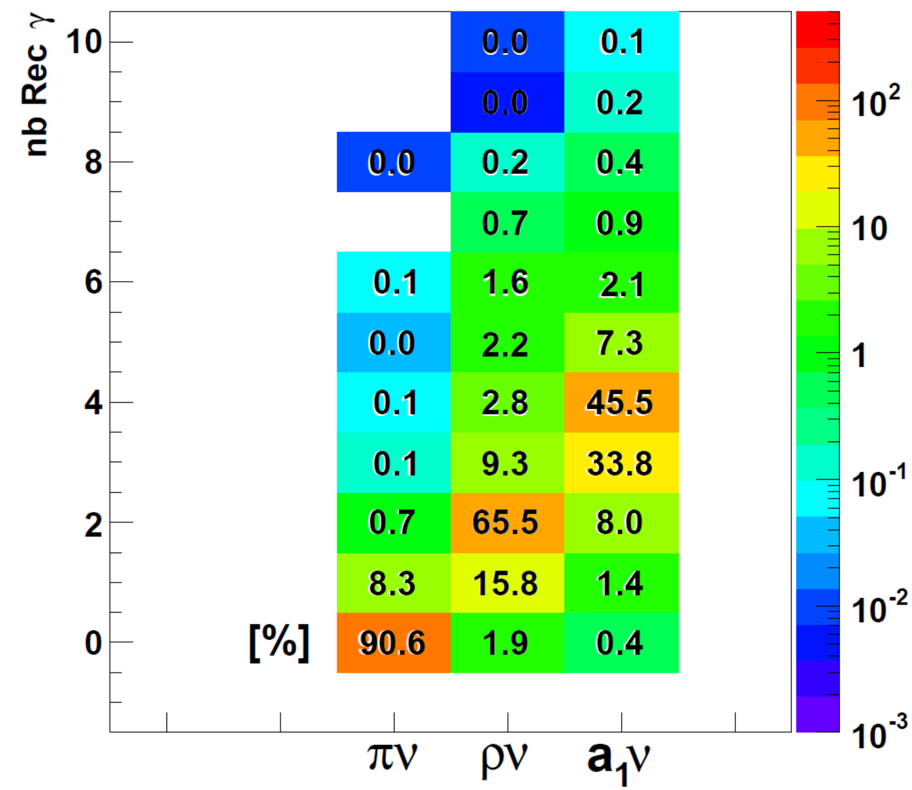


Number of rec photons

Reduction of radius degrades significantly the photon identification



$R_{ECAL}^{(inner)}$ = 1843 mm

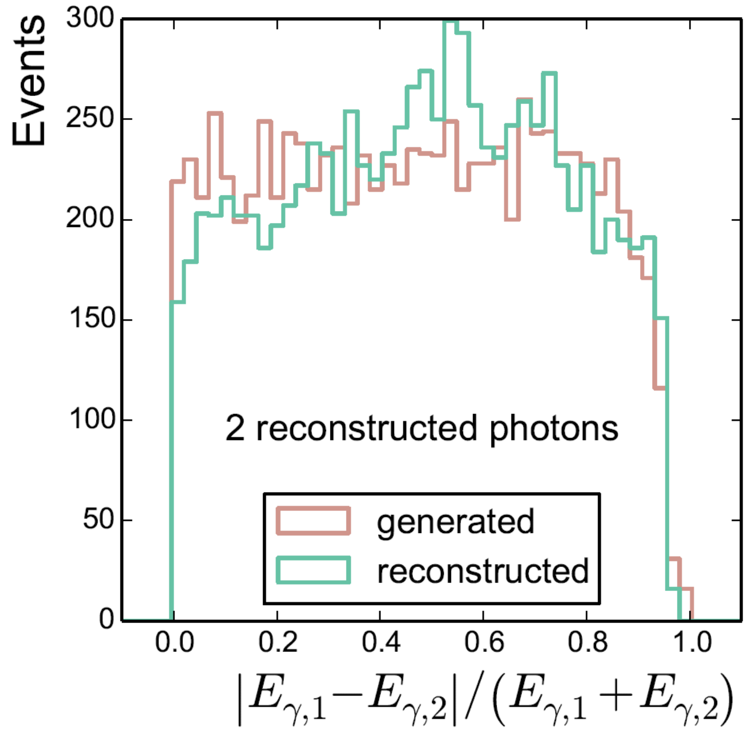


$R_{ECAL}^{(inner)}$ = 1450 mm

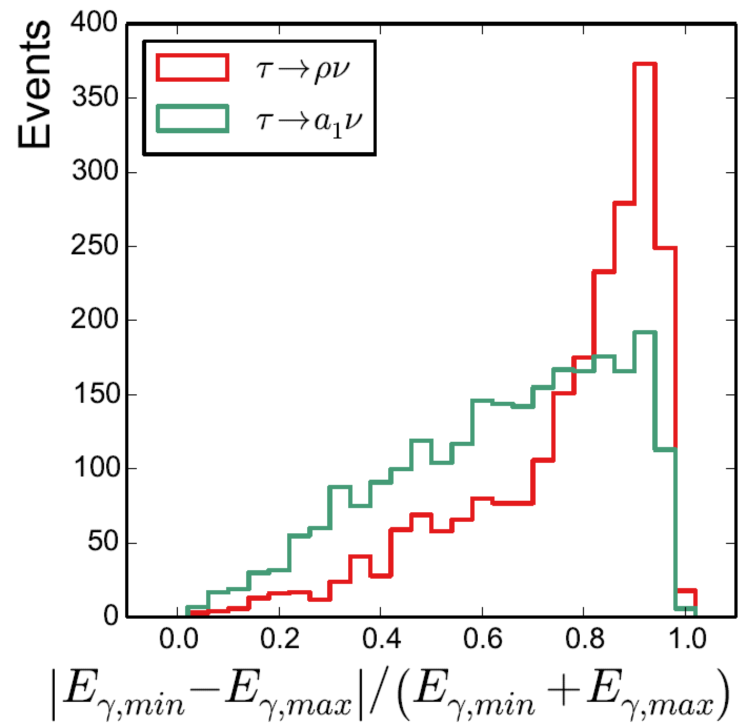
Special case: 3 reconstructed photons

$$A = \frac{\left| E_{\gamma_{\min}} - E_{\gamma_{\max}} \right|}{E_{\gamma_{\min}} + E_{\gamma_{\max}}}$$

Energy asymmetry between the most and the least energetic clusters.
One EM cluster is fake → A is close to 1.
The two clusters are then merged together if A>0.8



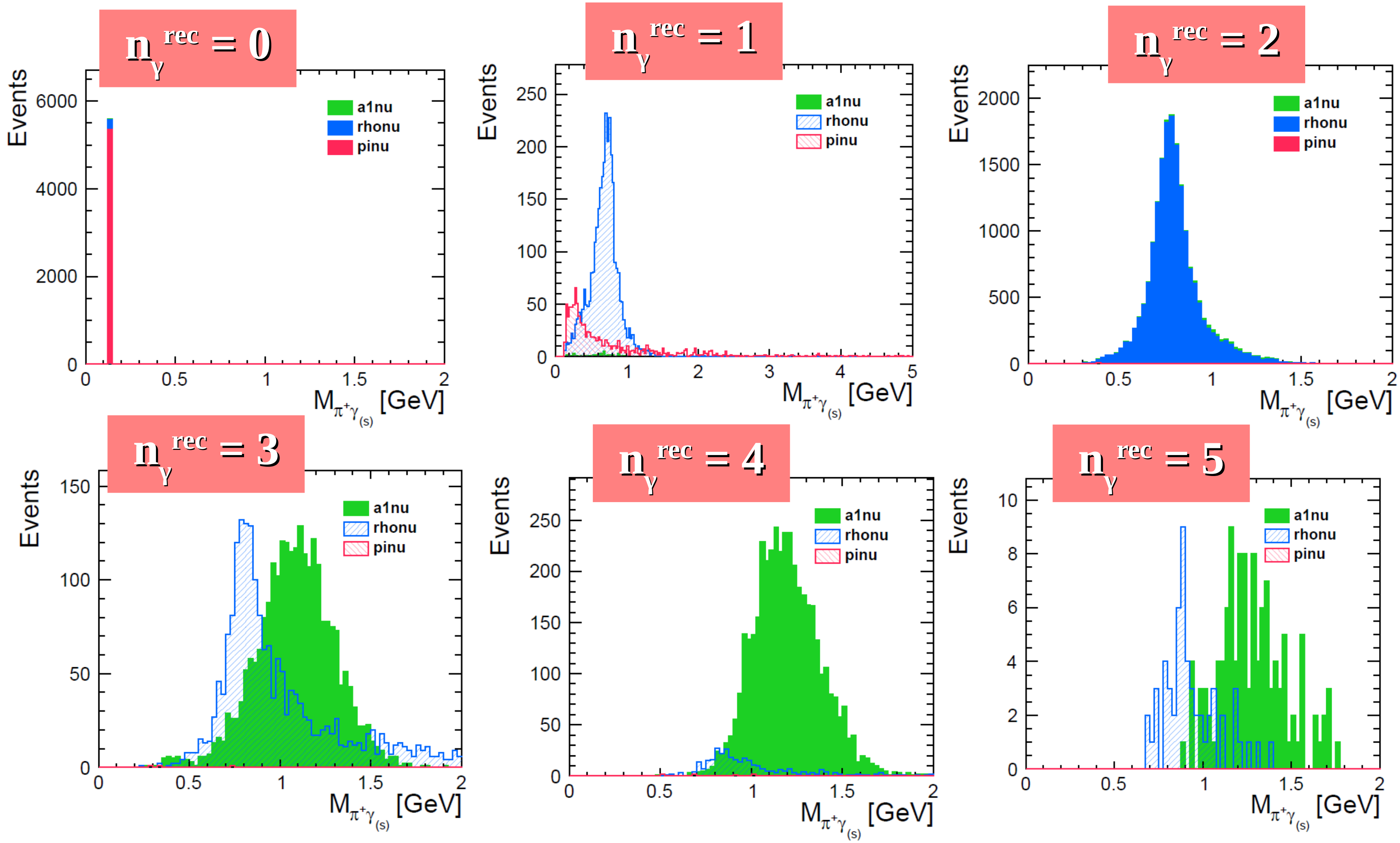
(a) The photon-photon energy asymmetry for the case of two reconstructed photons: comparison between generated and reconstructed values, the distributions are rather flat.



(b) The photon-photon energy asymmetry in the case of three reconstructed photons for the two decay modes $\rho\nu$ and $a_1\nu$.

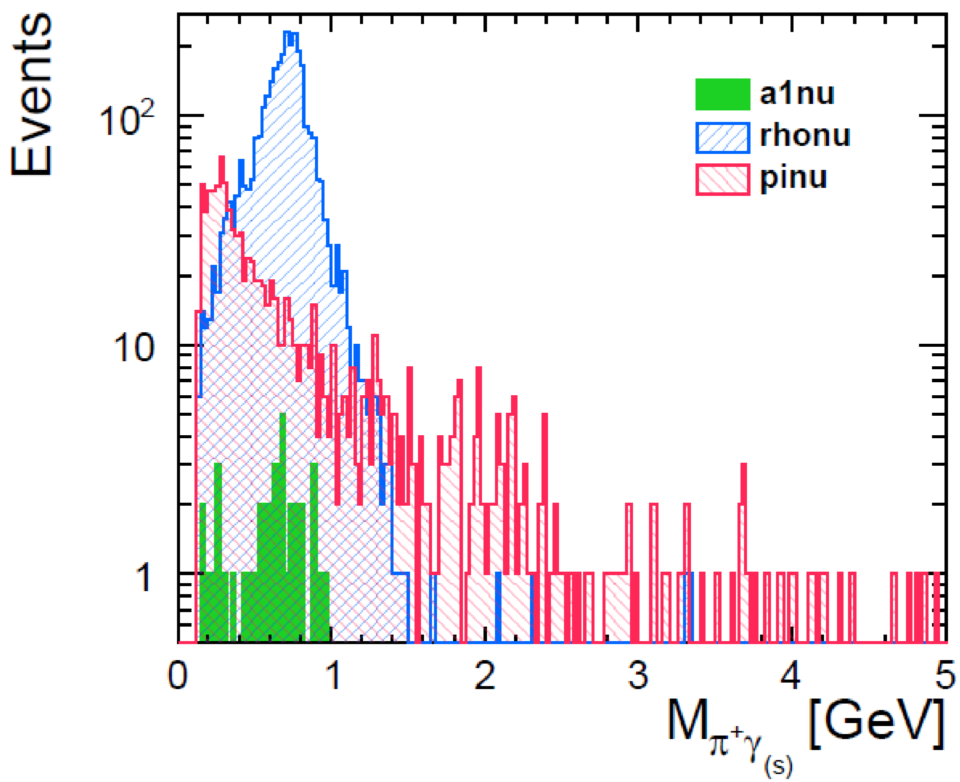
Mass peaks

- Mass of the system $\pi^{\pm}\gamma$(s) can be used to distinguish different decay mode

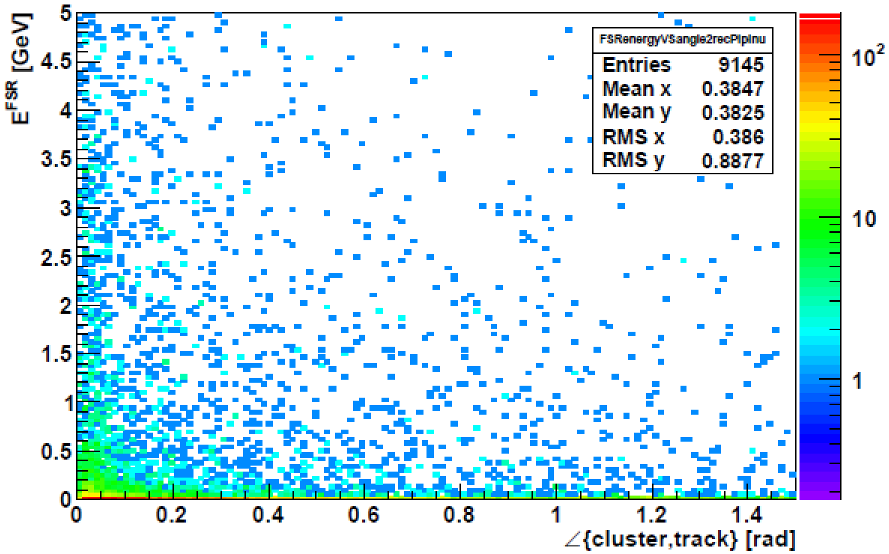


Solid color histograms are additive,
transparent ones are not.

Special case: $\tau \rightarrow \pi\nu$ decay with one rec. photon

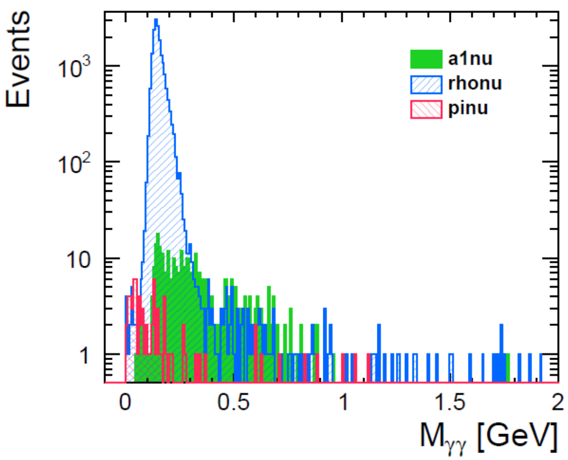


Photon can come from **final state radiation** (**FSR**, tau) with any energy or from interaction between pi and detector
$\rightarrow$ for $\pi\nu$ mode, invariant mass can be > m_τ

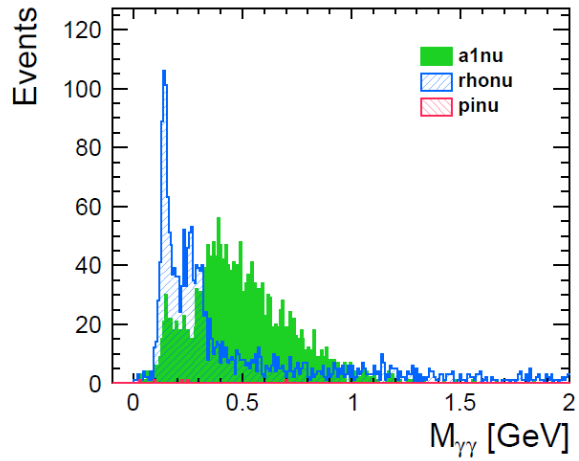


Photons: invariant mass

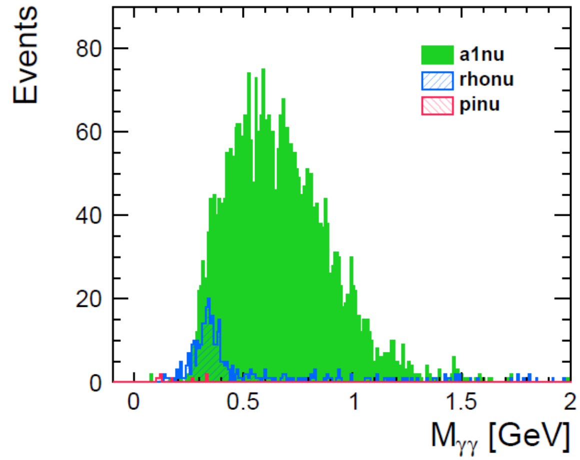
2 rec photons



3 rec photons

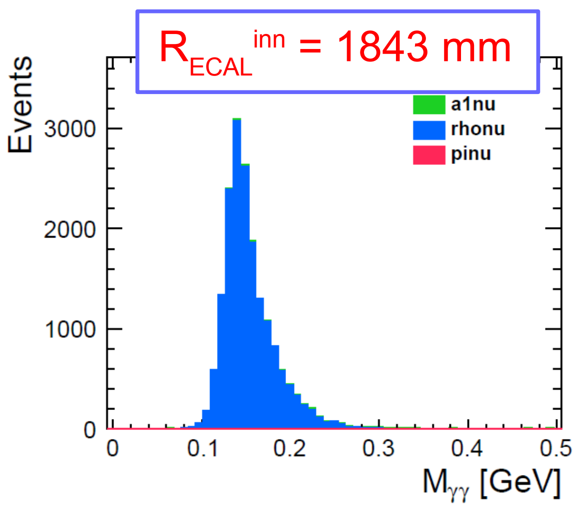


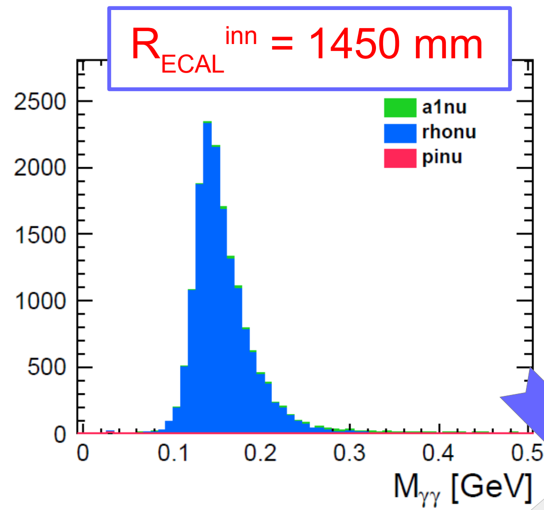
4 rec photons



- **2 reconstructed photons:**
  - $a_1\nu$: 2 photons have been lost.
- **3 reconstructed photons:**
  - $\rho\nu$: at least one photon is fake
  - $a_1\nu$: 1 photon is lost, however, photon invariant mass is mostly far from π^0 mass
- **4 reconstructed photons:**
  - small contribution from $\rho\nu$

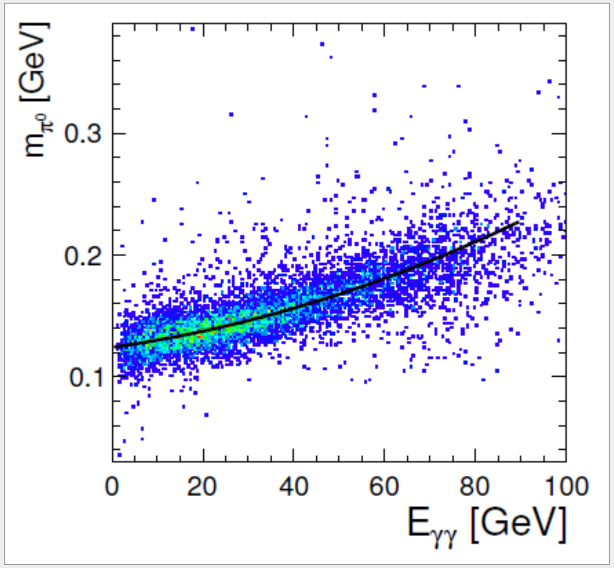
π^0 mass reconstruction

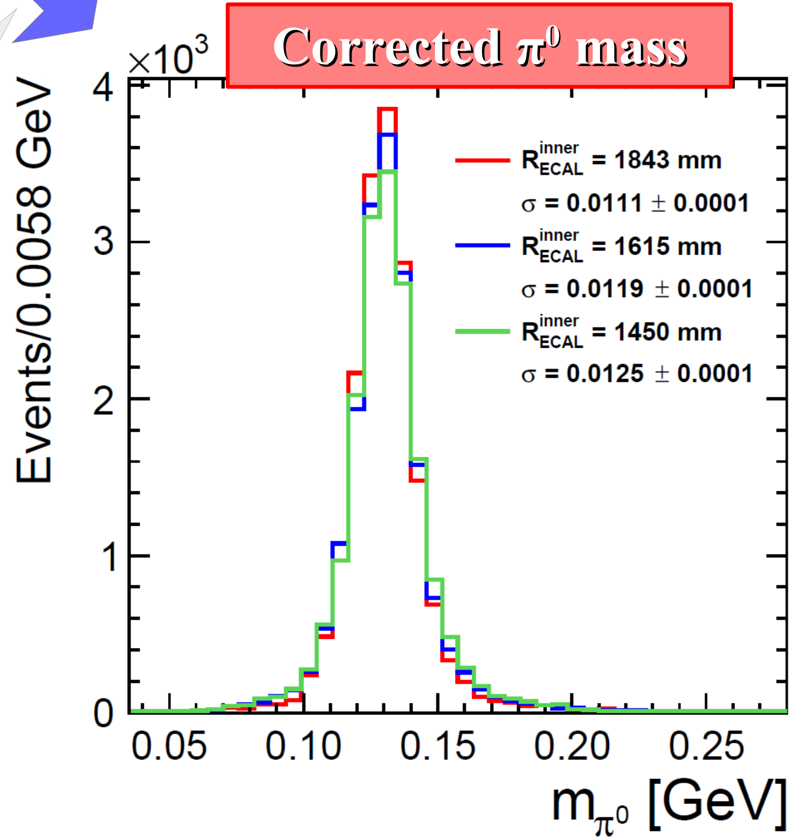




- Nice peak photon-photon (π^0) invariant mass
- tail due to high π^0 energies
- Good ratio signal/background

- Fit mass dependence on γγ energy with a parabola
- correct to mean value



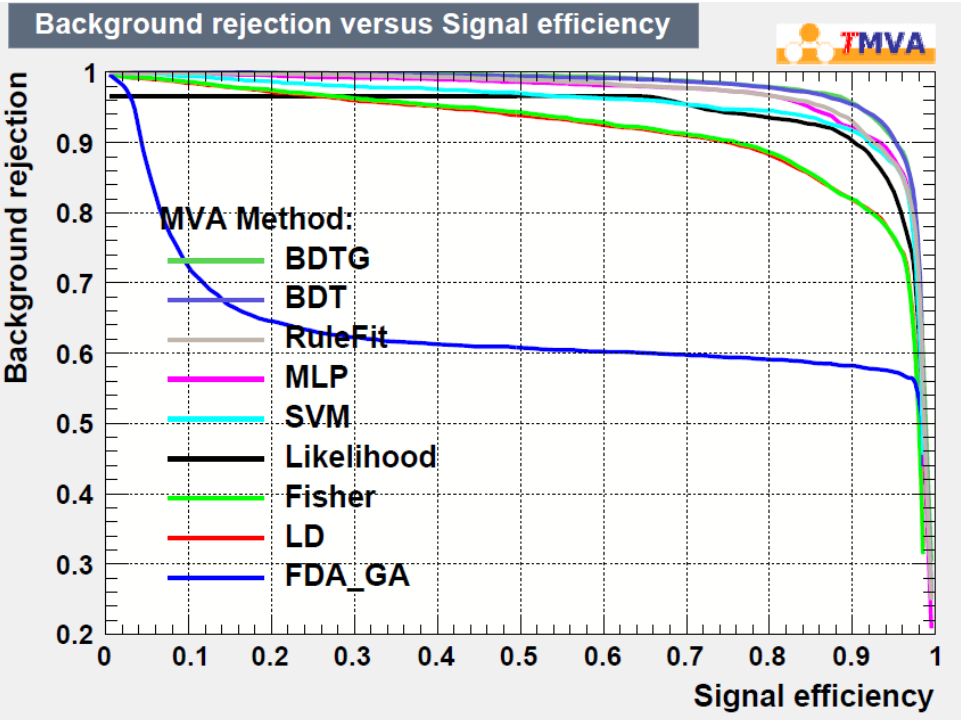


Decay mode identification

using Boosted decision tree

MVA Level 2

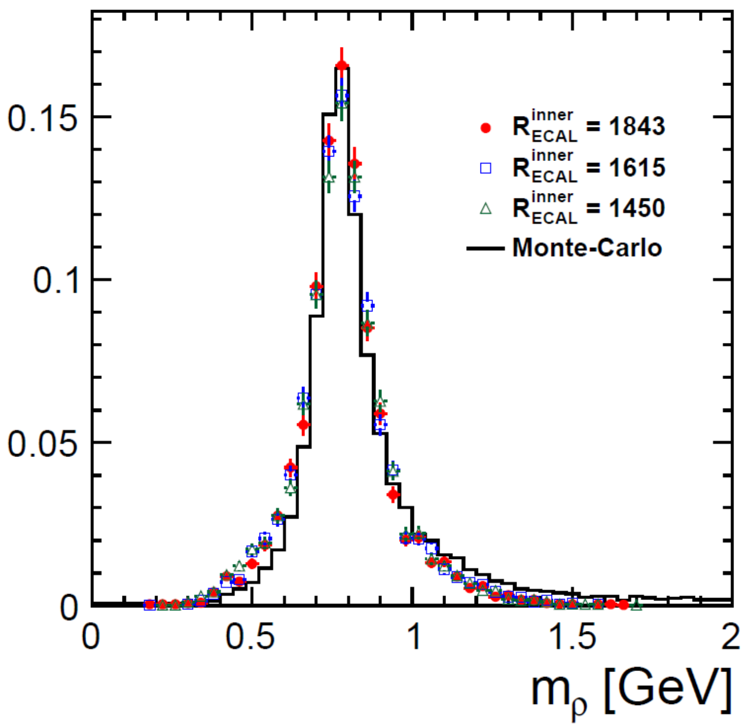
- Each decay mode ($\tau \to \pi\nu$, $\tau \to \rho\nu$, $\tau \to a_1\nu$) is used one by one as signal, the other two are used as background
- Input variables:
  - Reconstructed jet invariant mass
  - Photon invariant mass
  - Number of photons, photons' energy
- Boosted Decision Tree seems to give best performance
- Additional cuts based on jet mass ($m_{jet} < m_{tau}$, etc.)



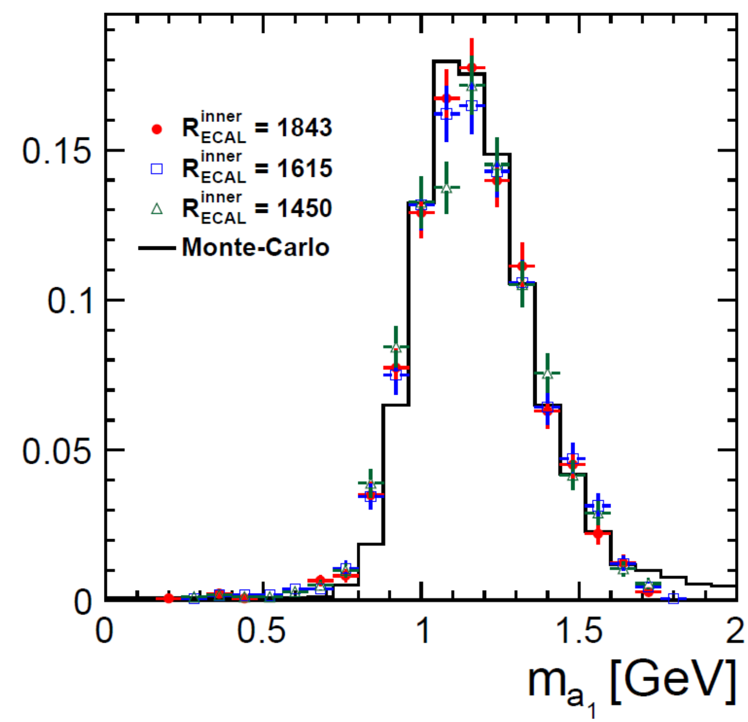
Example of using $\tau \to \rho\nu$ as signal

Reconstructed mass

- Reconstructed mass for events which are identified as **rho** or **a_1** decays compared to MC



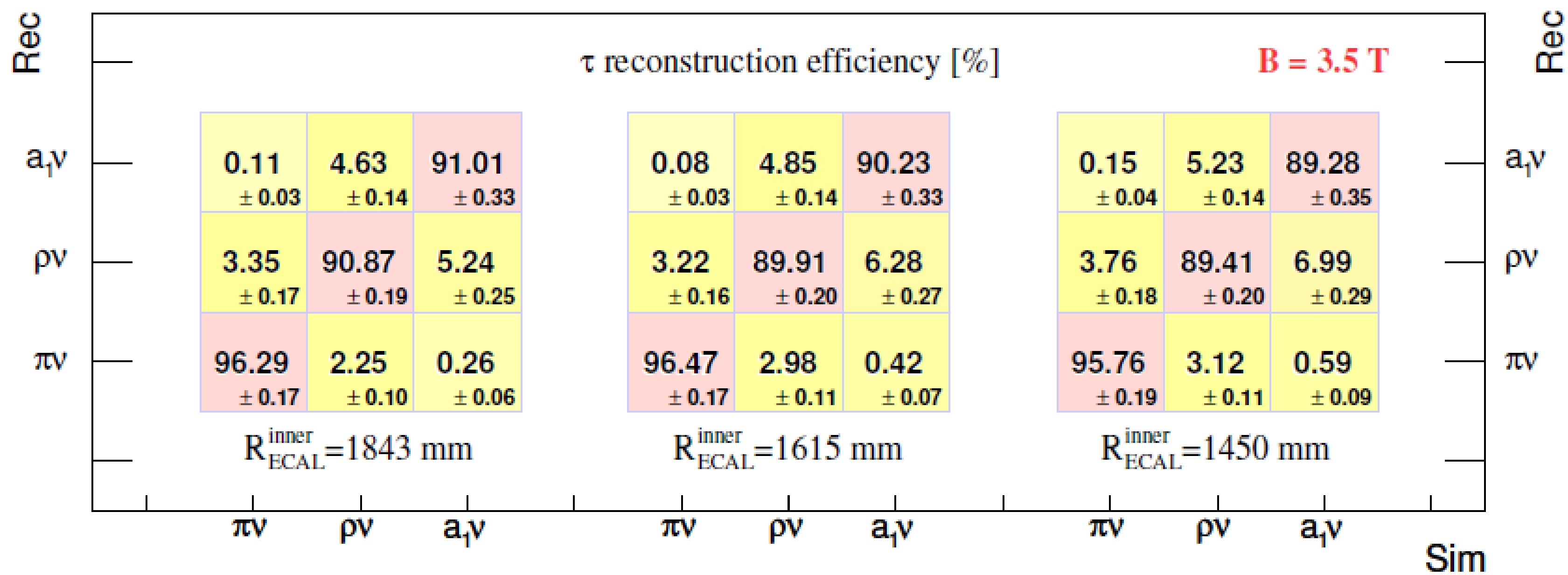
Histograms are normalised to 1.



- Very good separation of ρ and a_1 masses
- Slight difference observed for different ECAL sizes, especially for a_1 mass

Reconstruction efficiency & purity

Tau decay mode reconstruction efficiency
for R_{ECAL} = 1843, 1615 and 1450 mm



Slight difference in terms of efficiency for three ECAL models.

This is mostly due to smaller distance between photons and between photon&track for reduced radii.

Effect of magnetic field

- Baseline magnetic field B = 3.5 T
- Increase B to 4.0 Tesla
- Performance improves slightly
- MVA done for each detector setup

τ reconstruction efficiency [%] — B = 4.0 T

R_{ECAL}^{inner}=1615 mm

REC \ SIM	πν	ρν	a_1ν
a_1ν	0.08 ± 0.03	4.56 ± 0.14	89.97 ± 0.34
ρν	3.61 ± 0.17	90.57 ± 0.19	6.19 ± 0.27
πν	96.04 ± 0.18	2.48 ± 0.10	0.43 ± 0.08

R_{ECAL}^{inner}=1450 mm

REC \ SIM	πν	ρν	a_1ν
a_1ν	0.18 ± 0.04	4.92 ± 0.14	89.97 ± 0.34
ρν	3.79 ± 0.18	90.05 ± 0.19	6.30 ± 0.28
πν	95.68 ± 0.19	2.67 ± 0.10	0.46 ± 0.08

τ reconstruction efficiency [%] — B = 3.5 T

R_{ECAL}^{inner}=1843 mm

Rec \ Sim	πν	ρν	a_1ν
a_1ν	0.11 ± 0.03	4.63 ± 0.14	91.01 ± 0.33
ρν	3.35 ± 0.17	90.87 ± 0.19	5.24 ± 0.25
πν	96.29 ± 0.17	2.25 ± 0.10	0.26 ± 0.06

R_{ECAL}^{inner}=1615 mm

Rec \ Sim	πν	ρν	a_1ν
a_1ν	0.08 ± 0.03	4.85 ± 0.14	90.23 ± 0.33
ρν	3.22 ± 0.16	89.91 ± 0.20	6.28 ± 0.27
πν	96.47 ± 0.17	2.98 ± 0.11	0.42 ± 0.07

R_{ECAL}^{inner}=1450 mm

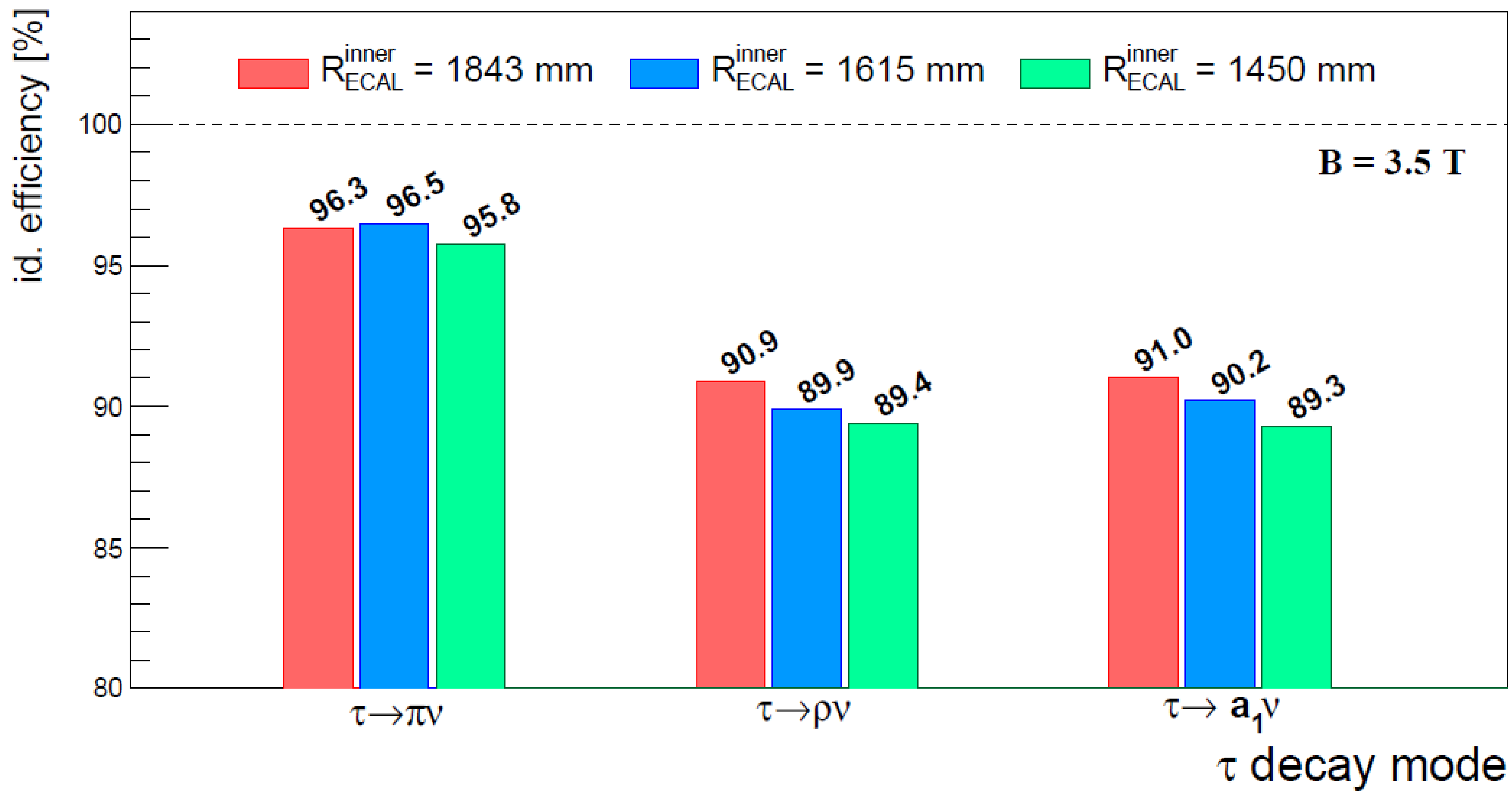
Rec \ Sim	πν	ρν	a_1ν
a_1ν	0.15 ± 0.04	5.23 ± 0.14	89.28 ± 0.35
ρν	3.76 ± 0.18	89.41 ± 0.20	6.99 ± 0.29
πν	95.76 ± 0.19	3.12 ± 0.11	0.59 ± 0.09

Summary

- **Tau decay mode reconstruction** (**E_{tau} ~ 125 GeV** *which is equivalent to τ's energy in ZH, H → τ⁺τ at 500 GeV cms*) is investigated using Garlic v3.0.3 (ilcsoft v01-17-06).
- Multivariate analysis technique is used for both photon and tau decay mode identification → **significant improvement compared to cuts based analysis**
- **Nice mass peaks observed (π^0, ρ, a_1) for different radii**
- High reconstruction efficiency even with a reduced detector size
- Comparison between ILD with ECAL of radii 1843, 1615 and 1450 mm **shows slight difference (overall less than 2% in terms of reconstruction efficiency and contamination)**
- No significant improvement for reduced ECAL size if choose to increase magnetic field from 3.5 to 4 Tesla
- Result for small ILD sizes is quite comparable to M. Reinhard's analysis → with a reduced ILD size we would still be able to measure the Higgs CP state via decay H → τ τ.

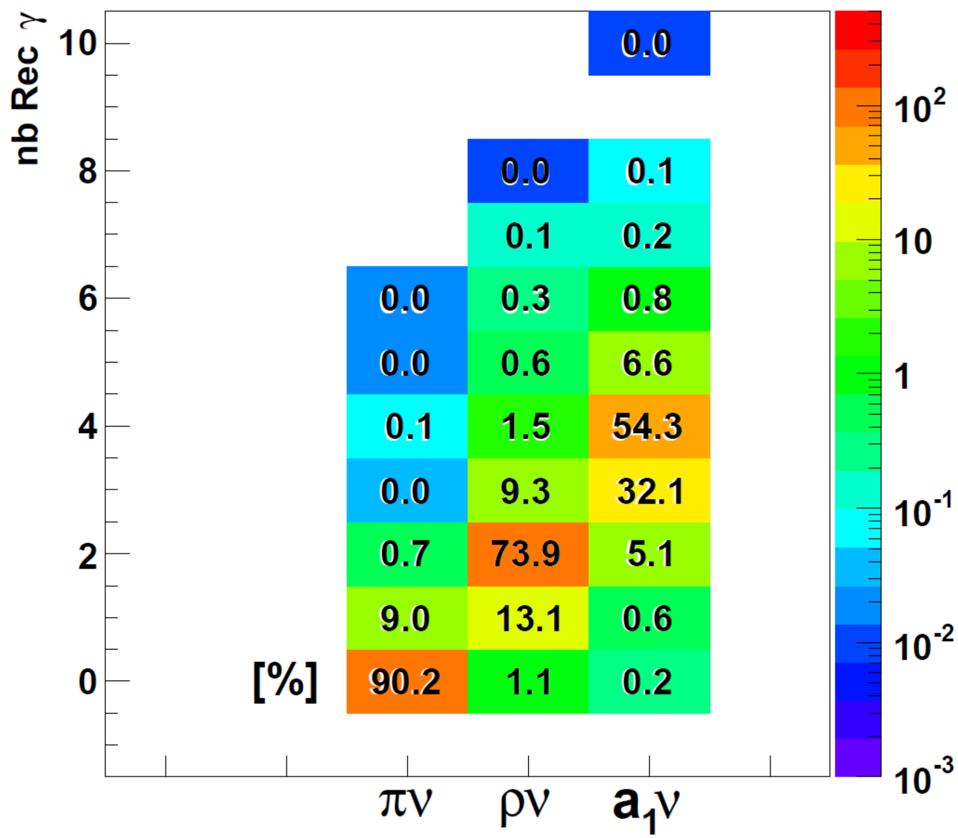
Extra slides

Reconstruction efficiency (cont'd)

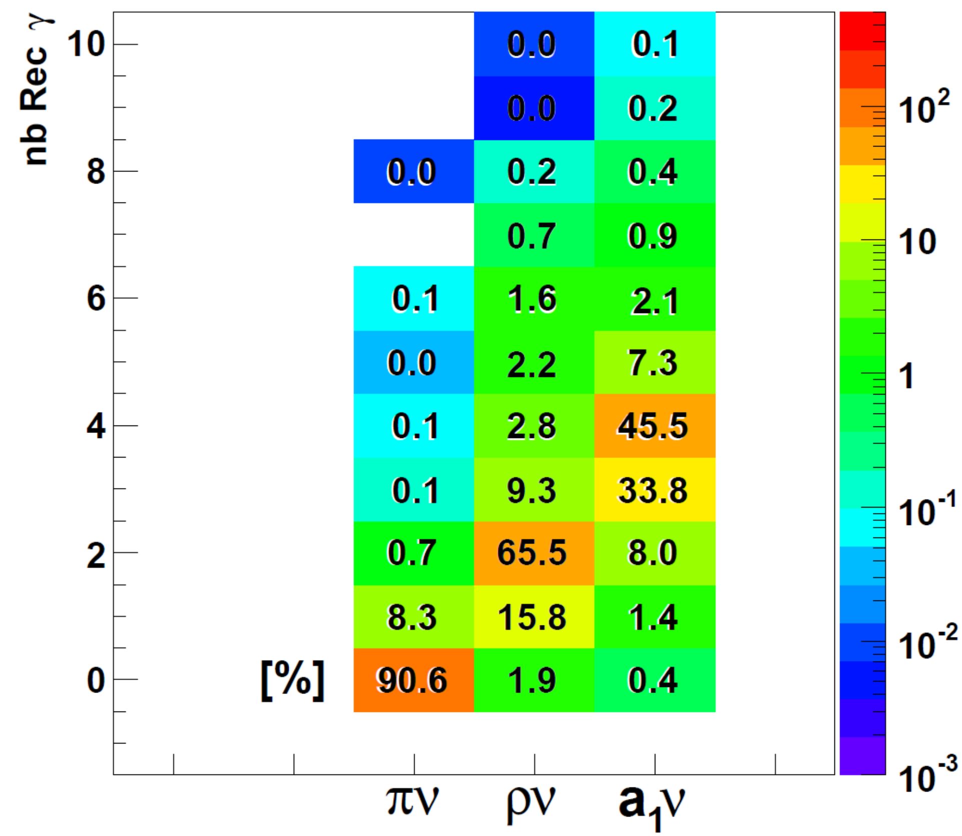


Comparison of efficiency for different ECAL sizes: 1843, 1615 and 1450 mm.
Slight difference observed.

Number of rec photons [%]



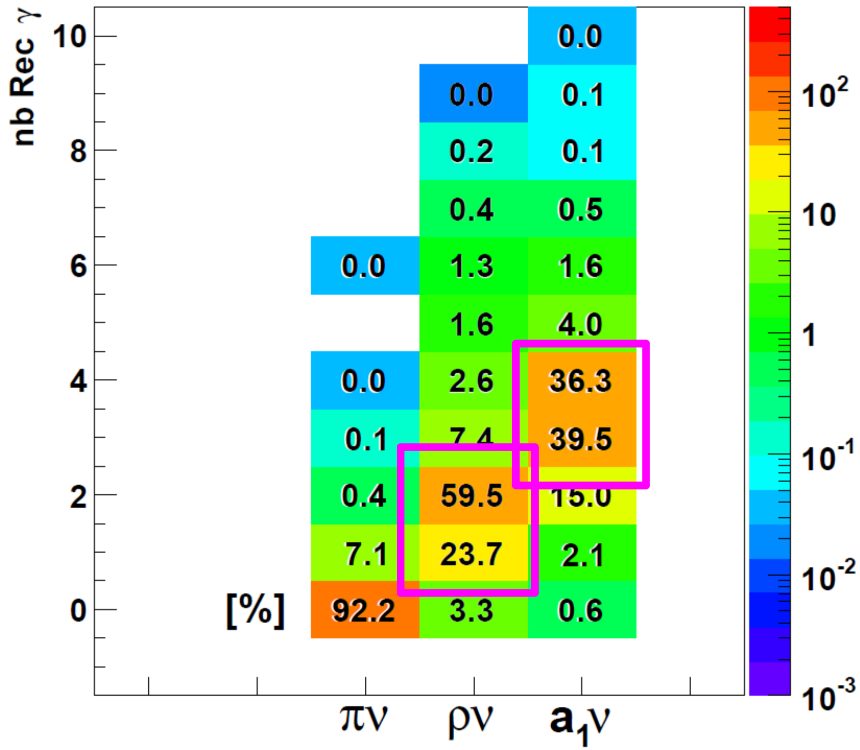
$R_{ECAL}^{(inner)}$ = 1843 mm



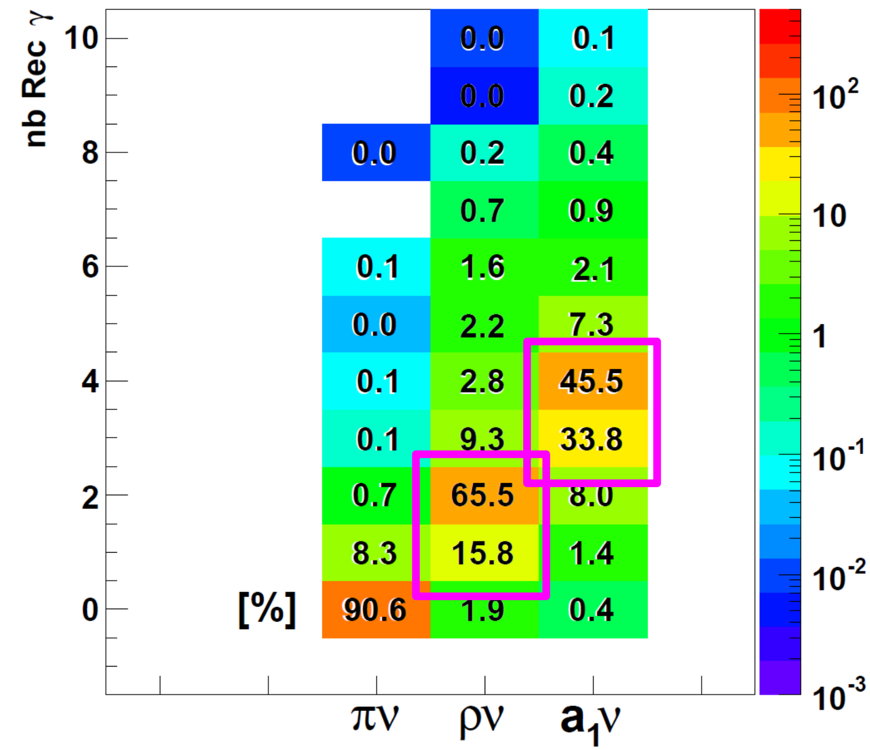
$R_{ECAL}^{(inner)}$ = 1450 mm

Number of reconstructed photons

Decay mode known from MC info.
Look at samples with different number of reconstructed photons.
If everything is fine: **πν**: 0 photon, **ρν**: 2 photons, **a_1ν**: 4 photons.



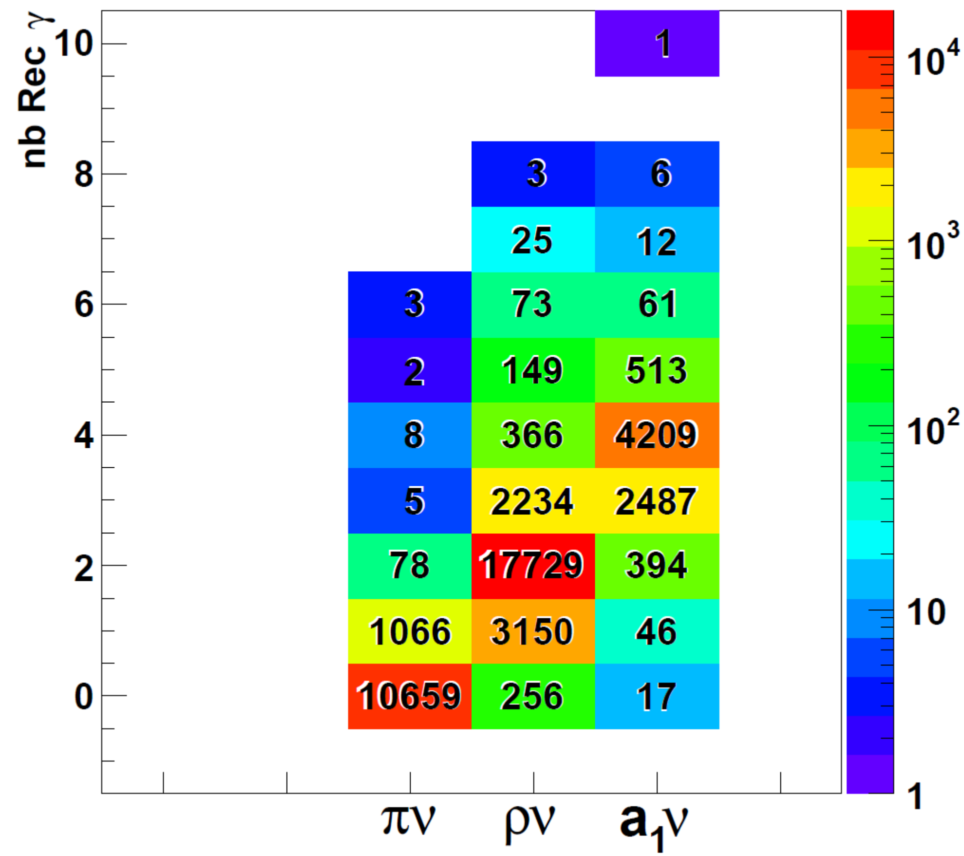
$R_{ECAL}^{(inner)}$ = 1450 mm
Cut-based photon selection



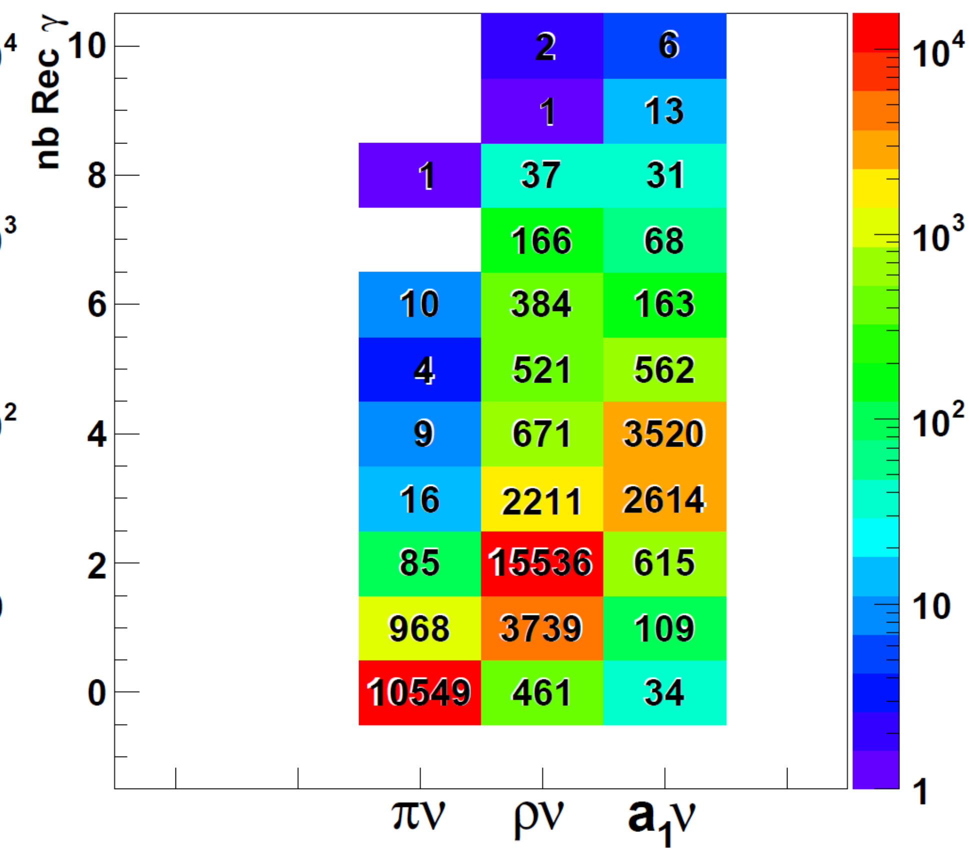
$R_{ECAL}^{(inner)}$ = 1450 mm
MVA photon selection

MVA does significantly help for photon selection against fake EM clusters compared to cut-based analysis and at the same time keep photon clusters which are not far from track.

Number of rec photons

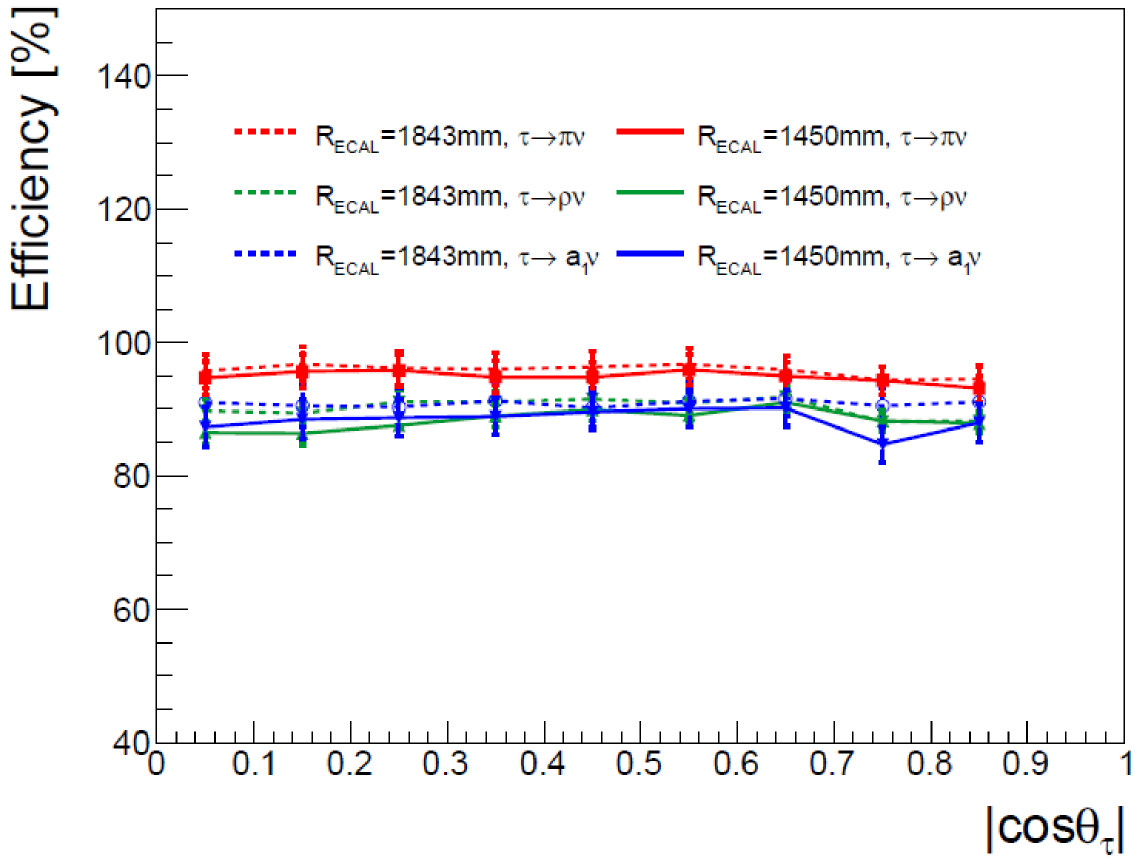


$R_{ECAL}^{(inner)}$ = 1843 mm



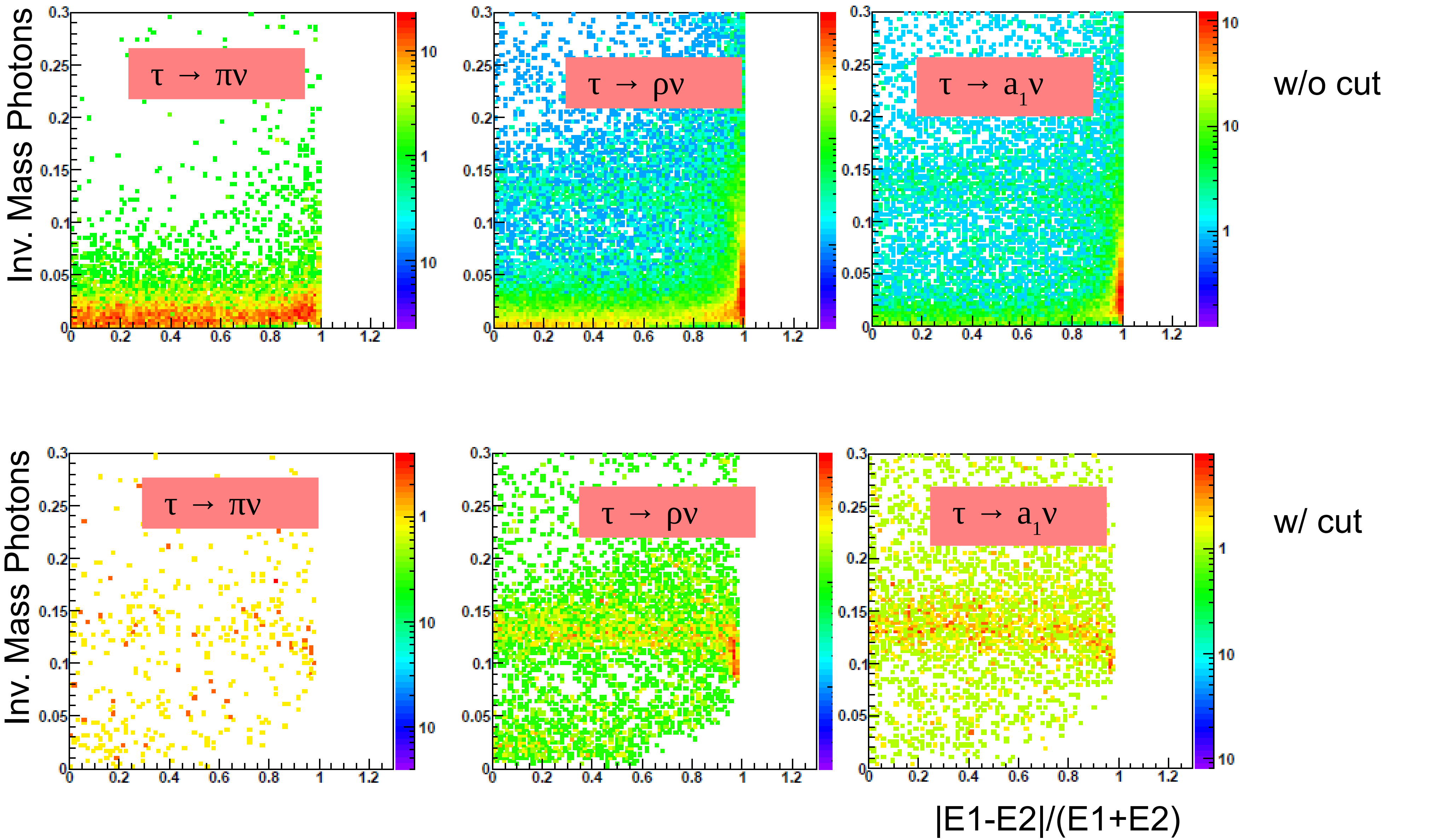
$R_{ECAL}^{(inner)}$ = 1450 mm

Rec efficiency vs cosTheta



Slight dependence on $|\cos\theta|$

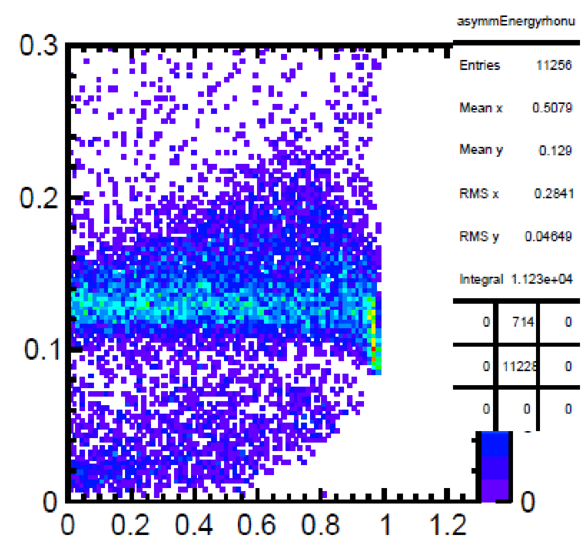
Effect of distance-energy cut



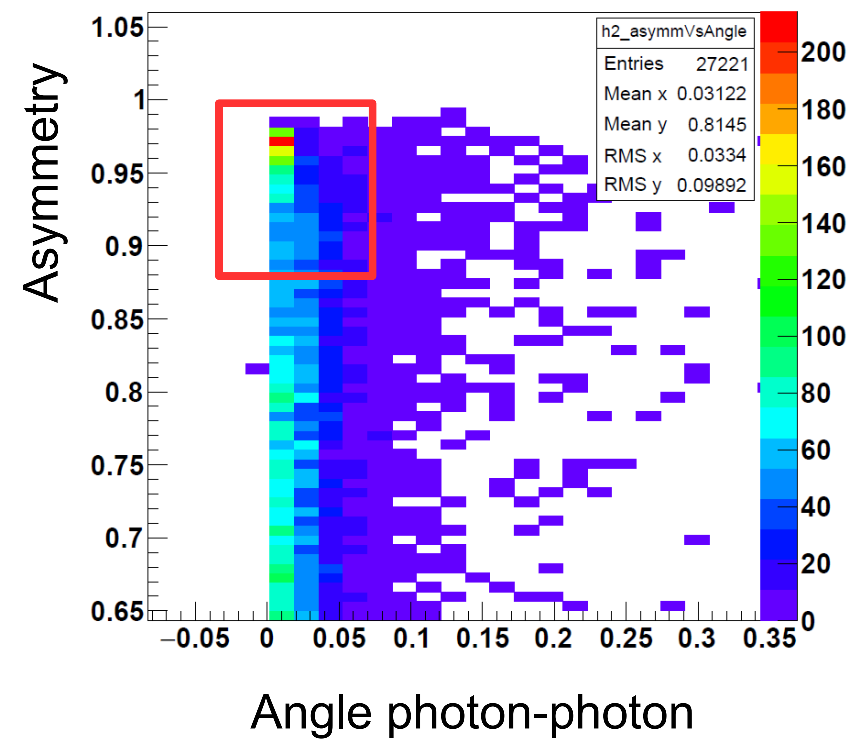
Photon id: 3 rec photons case

- Fake clusters created from interaction with detector
- "Asymmetry" of energy very close to 1

$$\text{Asymmetry} = |E_{\gamma 1} - E_{\gamma 2}| / (E_{\gamma 1} + E_{\gamma 2})$$

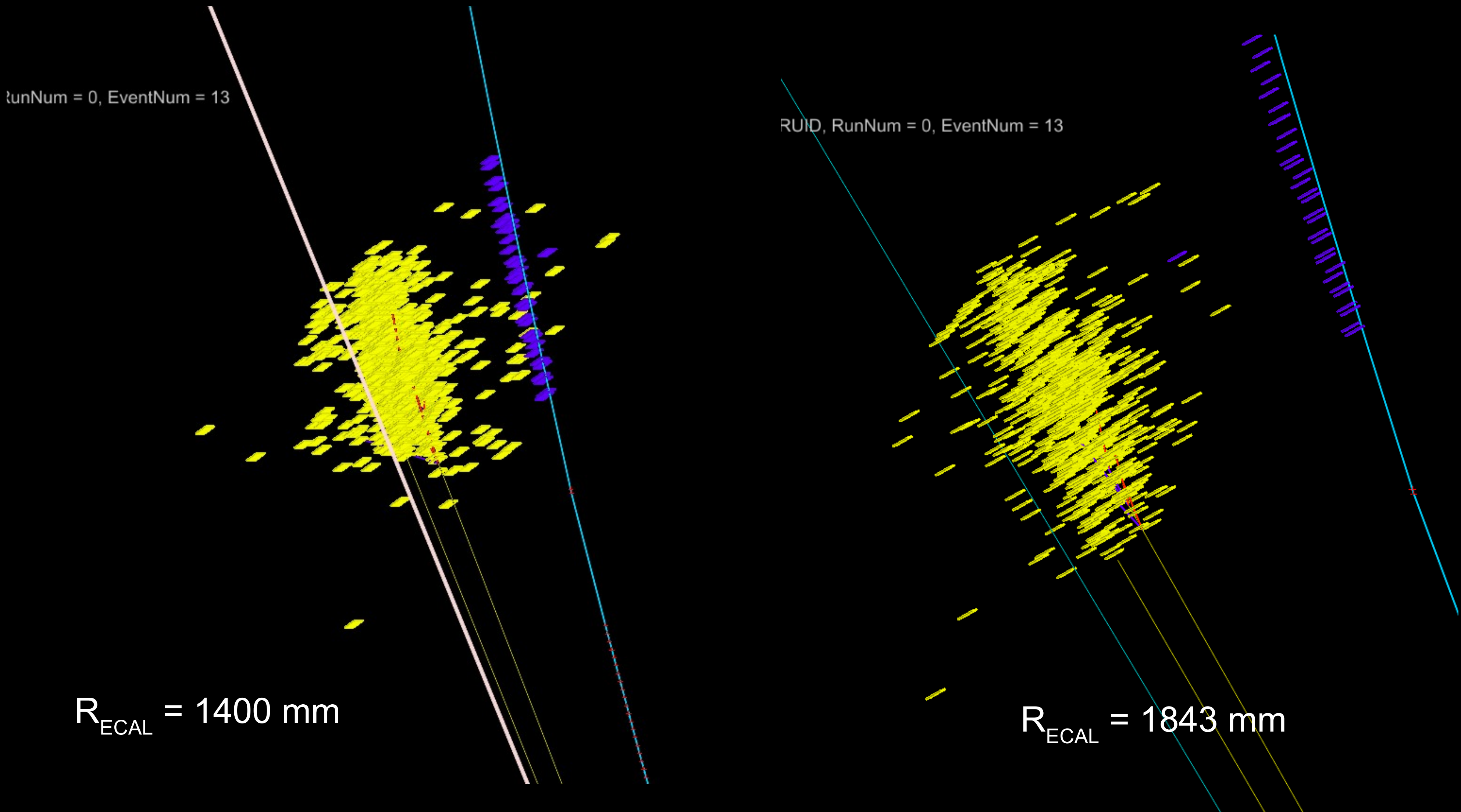


Example of photon invariant mass vs asymmetry



Choose to merge closest clusters with asymmetry close to 1.

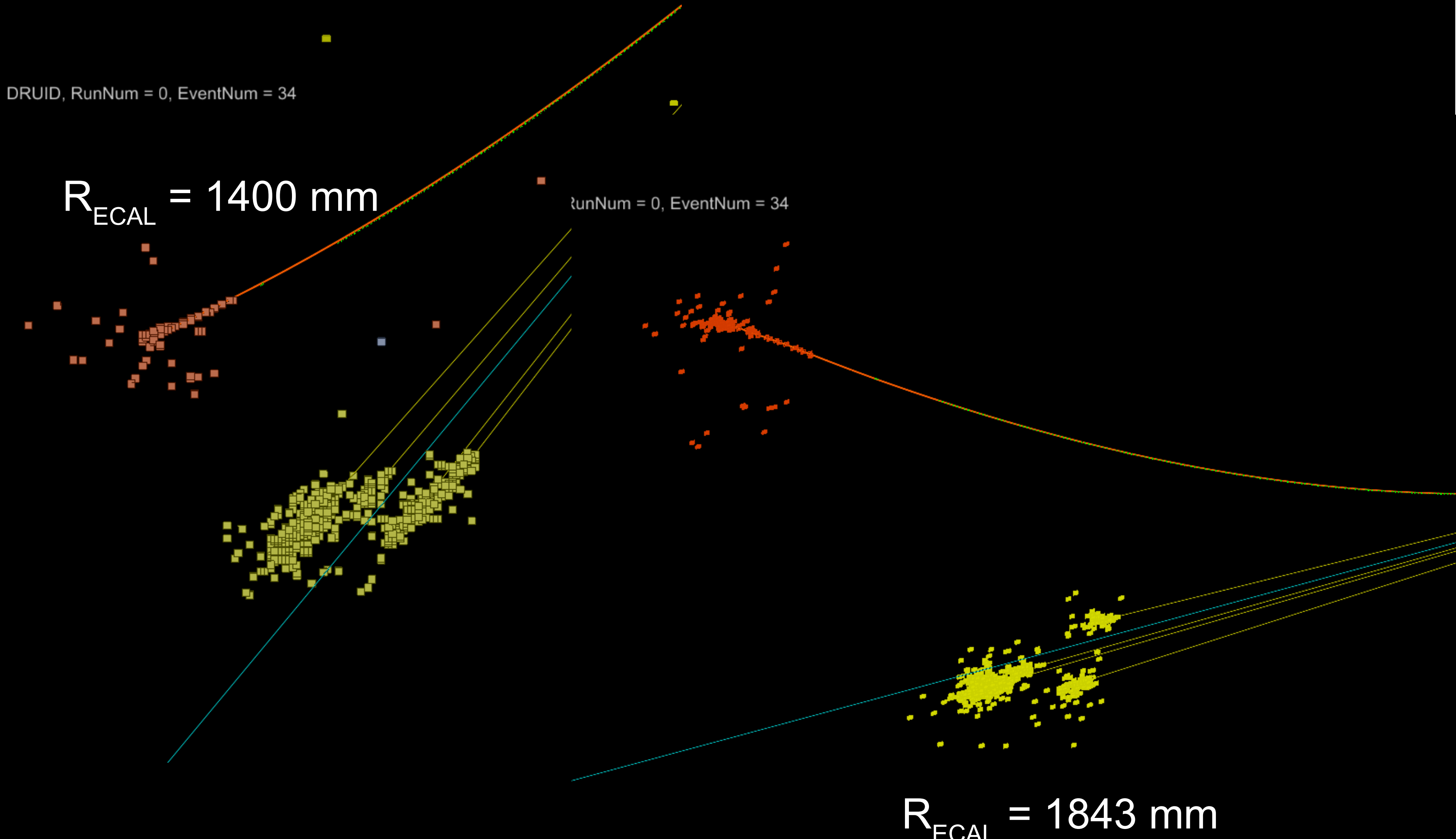
Example (1)

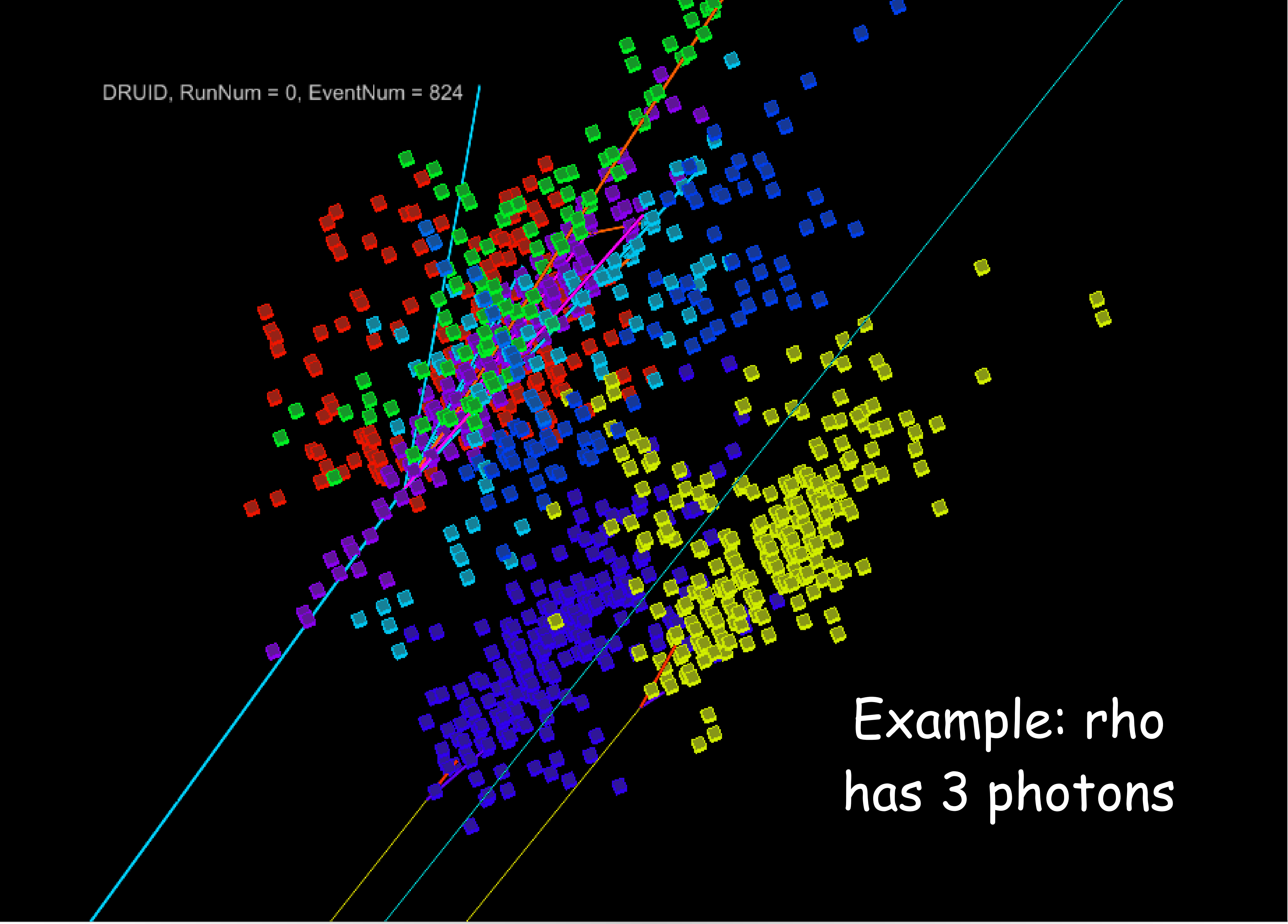


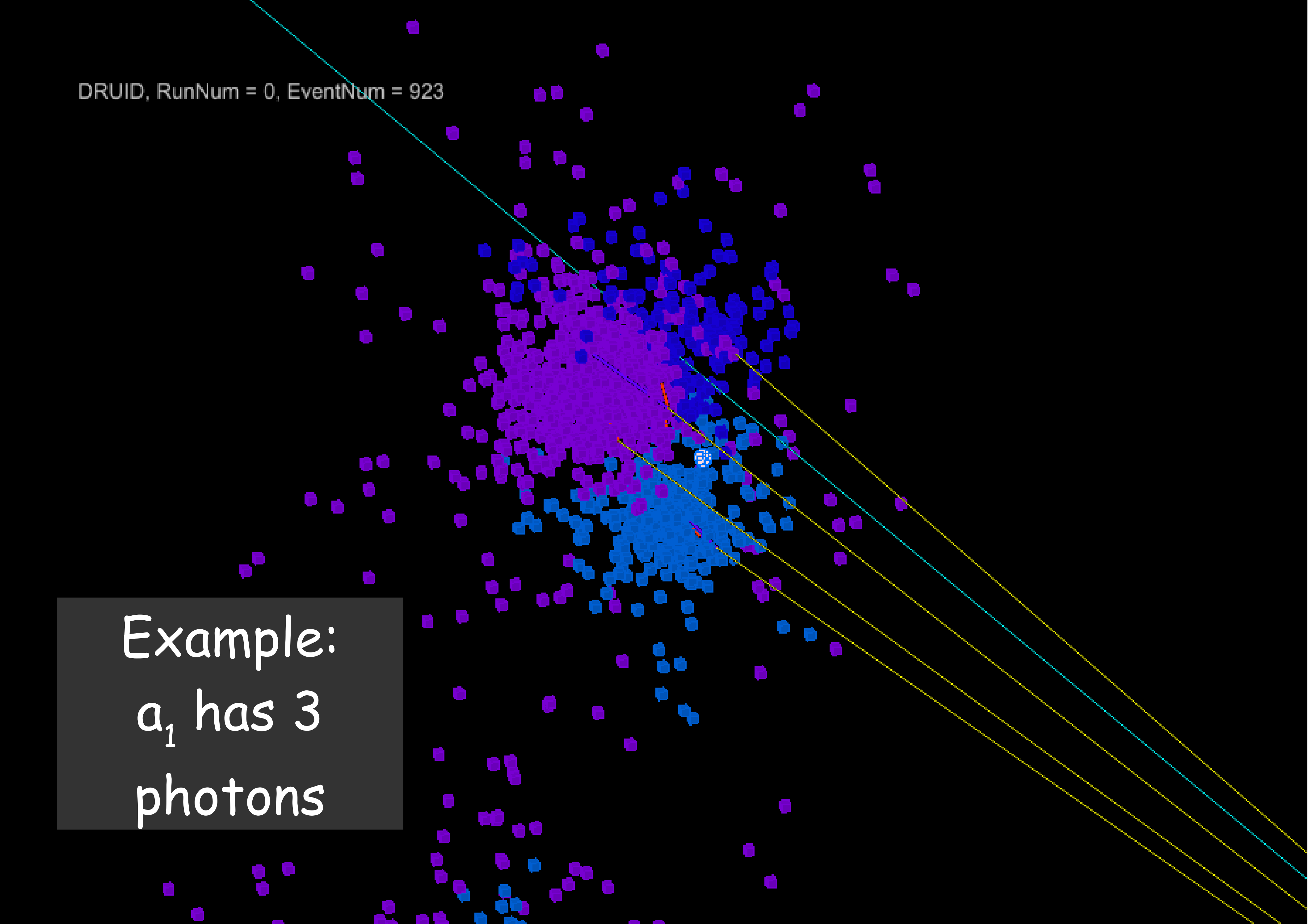
R_{ECAL} = 1400 mm

R_{ECAL} = 1843 mm

Example (2)


DRUID, RunNum = 0, EventNum = 34
R_{ECAL} = 1400 mm
RunNum = 0, EventNum = 34
R_{ECAL} = 1843 mm

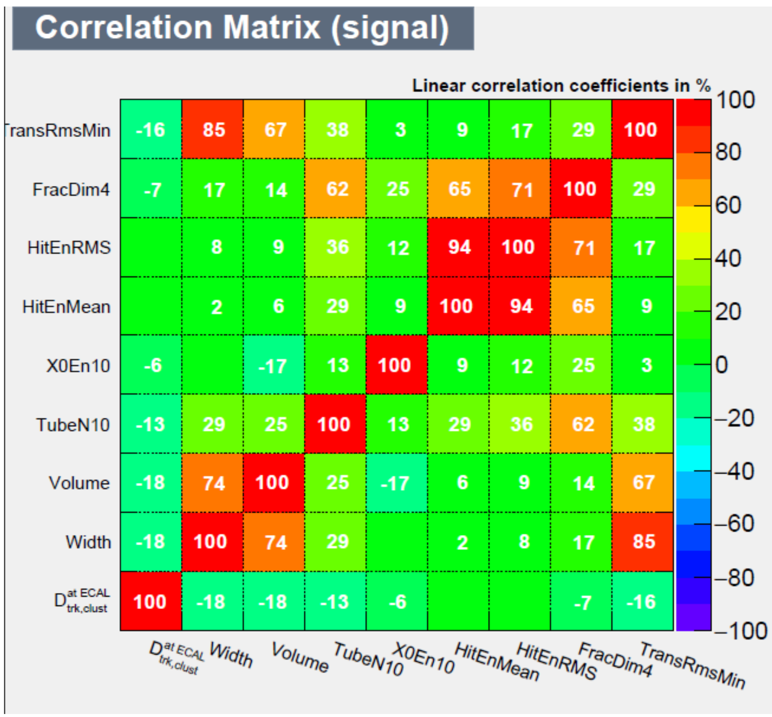

DRUID, RunNum = 0, EventNum = 824
Example: rho
has 3 photons

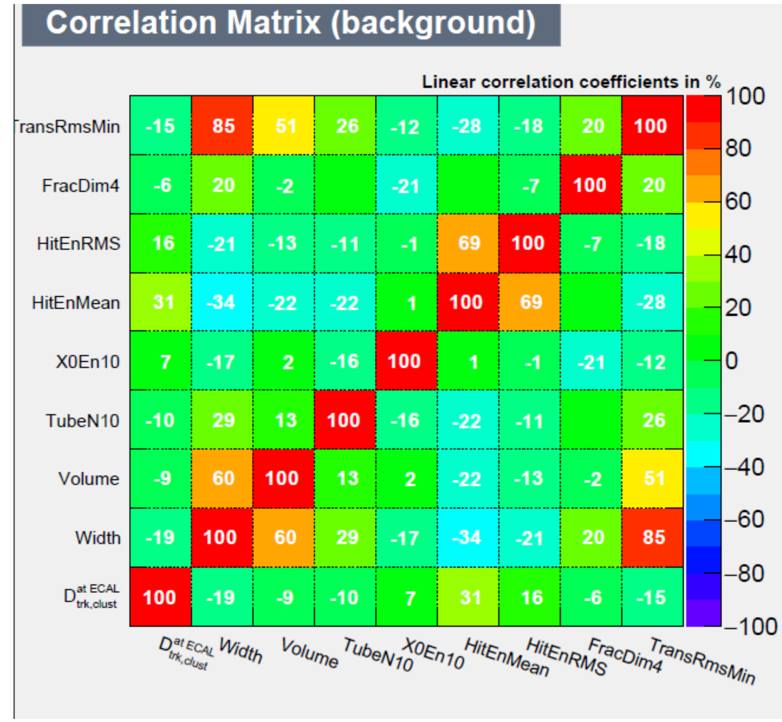


Example: a_1 has 3 photons

Photon identification

MVA Level 1

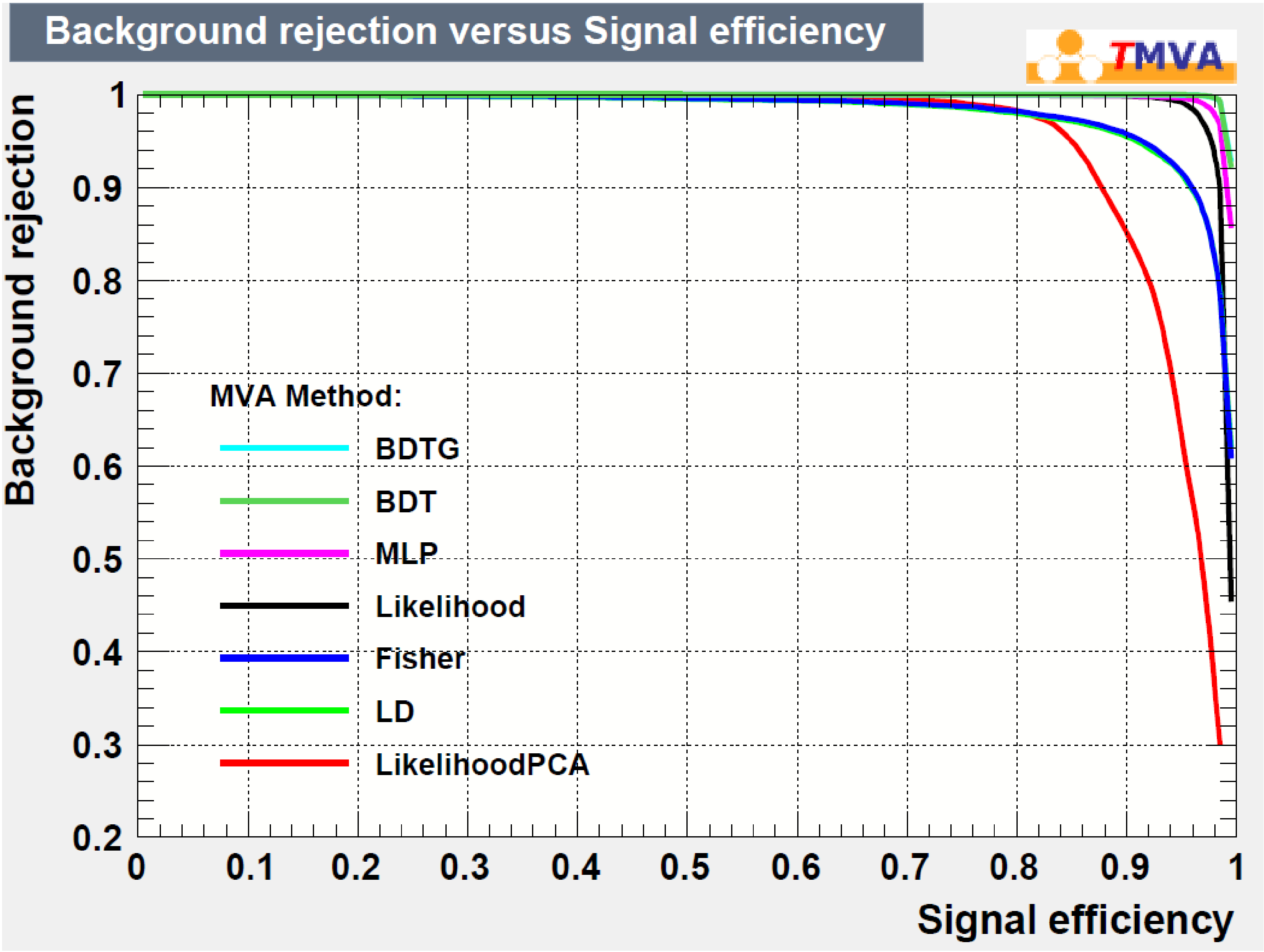




- $D_{trk,clust}$: distance track-cluster at ECAL front face
- Width: cluster width
- X0En10: energy fraction of cluster between 5 and 10 X_0 after the cluster start
- HitEnMean: mean of the cluster's hit energy
- HitEnRMS: RMS of the hit energy distribution
- FracDim4: cluster fractal dimension: log-ratio of the number of hits in the cluster (N1) with the number of hits when cells are grouped into larger cells of 4x4 cells (N_4): FD = $\log_{10}(N_4/N_1)/\log_{10}(4)$.
- TransRmsMin: minimum transverse RMS, cluster hits are projected onto the front face of ECAL along the direction between the CoG and the IP.

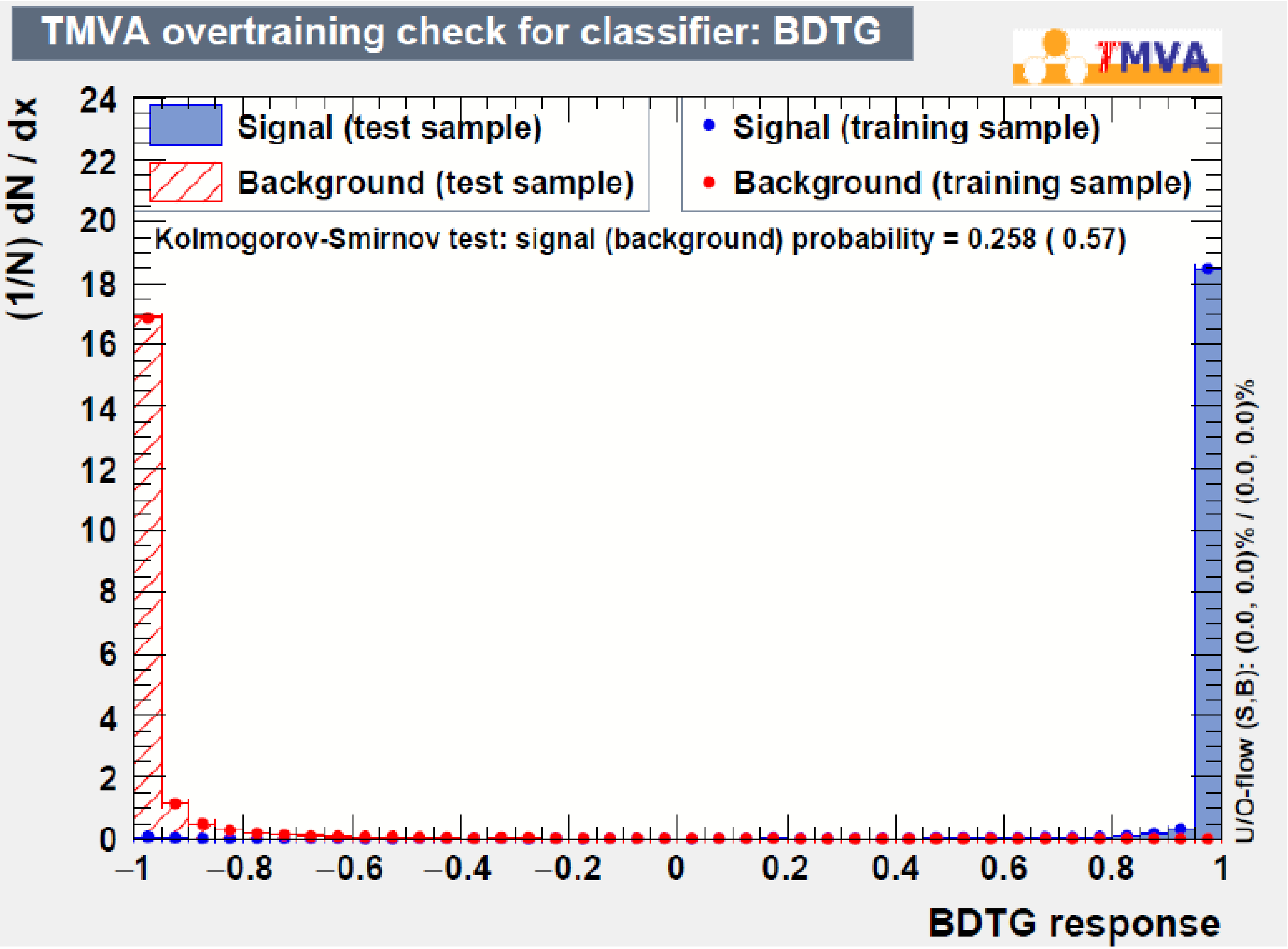
Photon identification

MVA Level 1

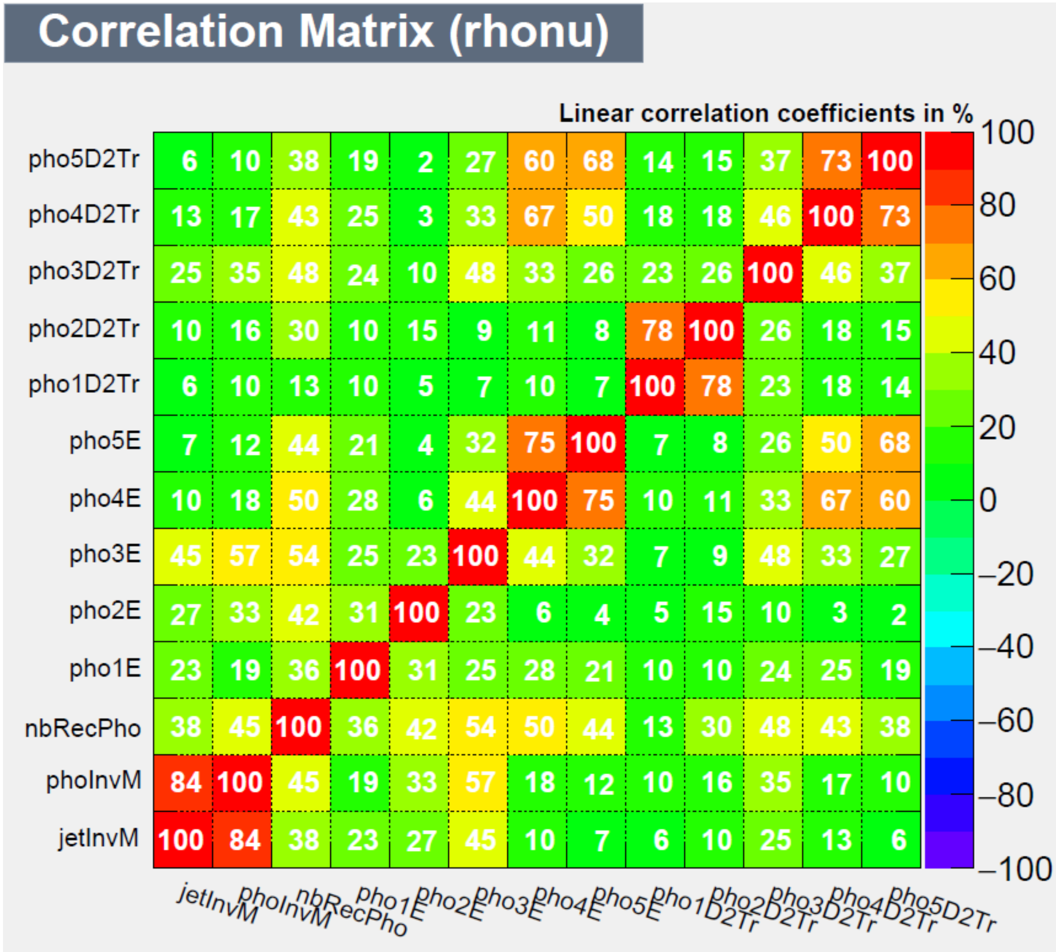


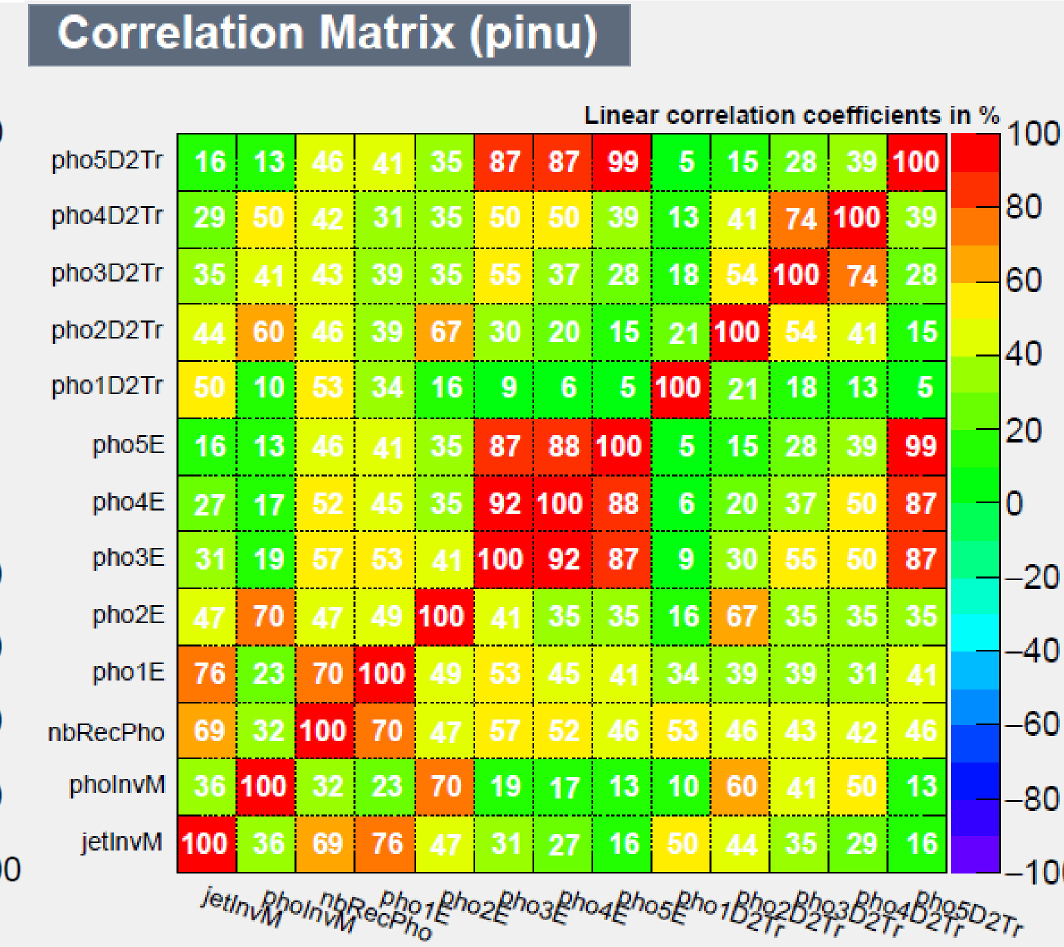
Photon identification

MVA Level 1

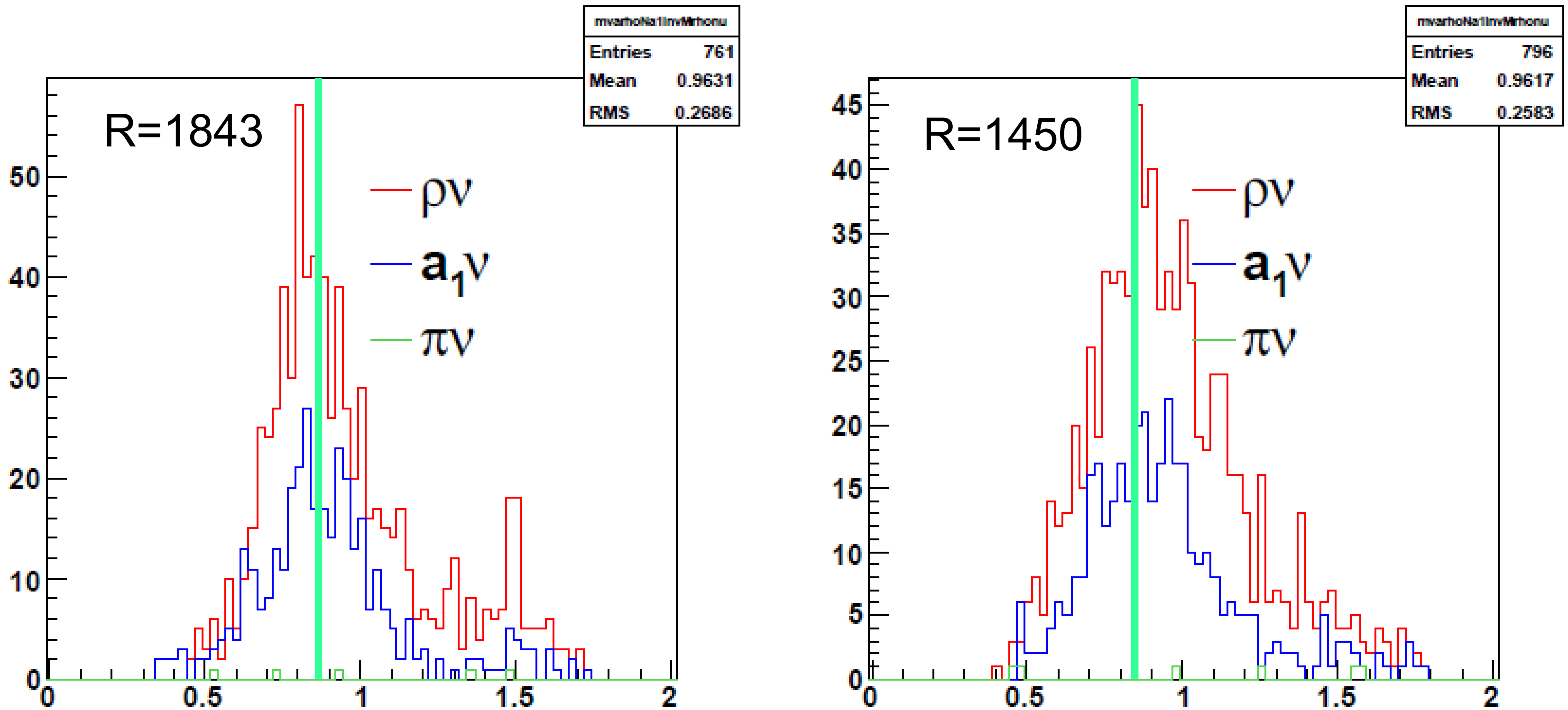


BDT for tau decay mode id.



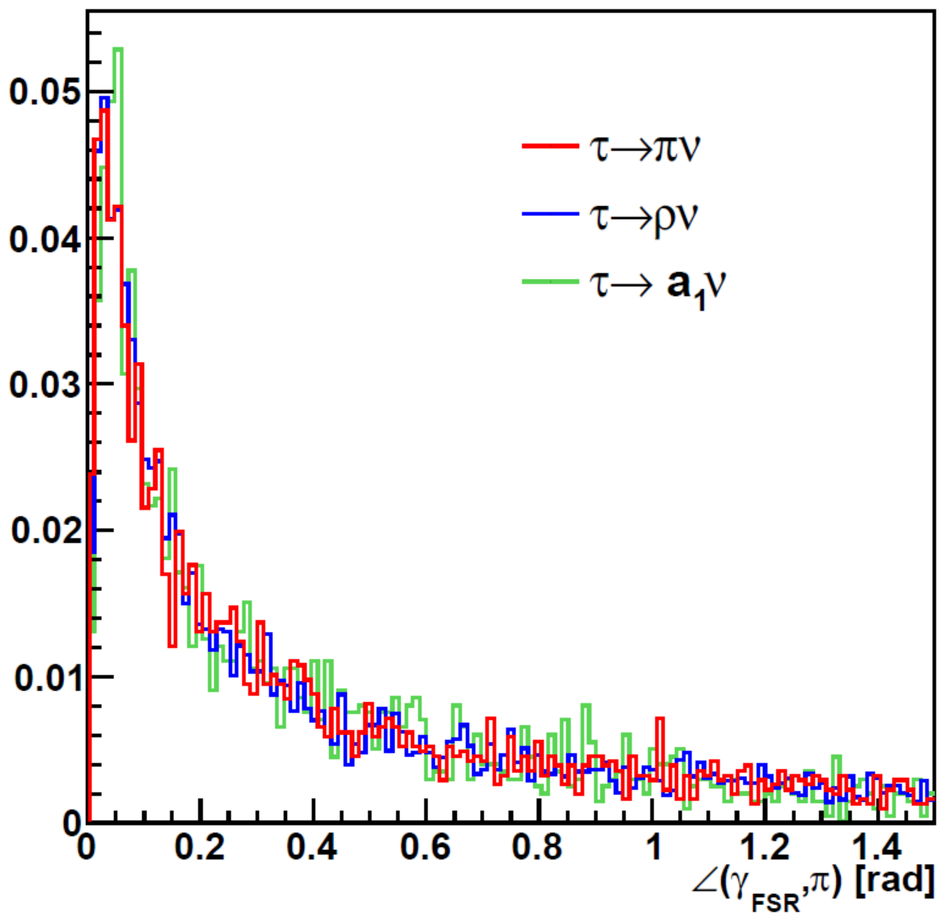


BDT for tau decay mode id.



- Events which pass both 2 identifiers: apply additional criteria based on jet mass

Angle FSR photon to rec. π



Not all of FSR photons can be removed.