Updates on ν_τCC search based on kinematics for the τ—>ρ channel
—
Improving the ρ identification

ν_τ meeting group

-

3rd December 2020

Thomas Kosc — kosc@ipnl.in2p3.fr

3rd year PhD student at Institut de Physique des Deux Infinis (Lyon, France)

Supervisor : Dario AUTIERO

DEEP UNDERGROUND NEUTRINO EXPERIMENT

Abstract

- One month ago I presented a ν_τ identification analysis based on kinematic criteria at DUNE FD, using the simulation files (CAFANA NTuples) of the DUNE TDR. I focused on the $\tau^- \to \rho^-$ decay mode (large BR, resonance).

The ρ decays into a pair of pions $\rho^- \to \pi^-\pi0$ and the neutral pion itself decays into two photons (assuming isotropic decay in the $\pi0$ rest frame), thus $\rho^- \to \pi^-\pi0 \to \pi^-\gamma_1\gamma_2$.

Obviously, within a single ν_τCC events where the $\tau^- \to \rho^-$, the daughter particles of the ρ^- can be blurred by by the combinatorics due to the pions coming from the hadronic system of the event. The first step of the analysis is then to develop a method to tag correctly the ρ and assess its performance. Later, the tagging method will be done blindly.

- First, I'll recall the key points of the $\tau^- \to \rho^-$ analysis and then will I develop possible improvements.

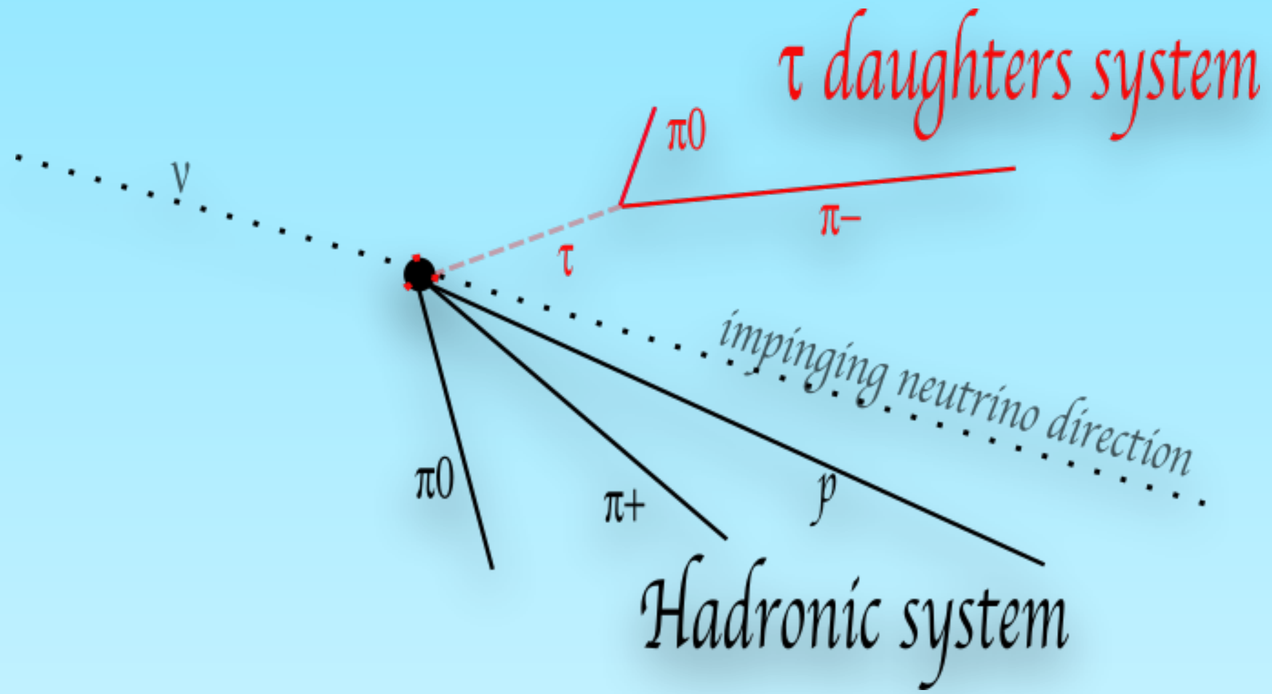
- Framework : neutrino mode, individual particle energy smearing, 100% correct particle idendification (but no charge identification, for instance to us $\pi^- = \pi\pm$).

Semantica - « true ρ » and « fake ρ »

In this slide I don't represent the photons from the π0 decay, for simplicity.

Defining a ρ candidate is equivalent to defining a pair (π0;π±) (assuming no charge identification for the charged pions). Doing so, I could:

—> Pick the pions system that is the true ρ daughter decay system. This choice corresponds to the **true** ρ.
—> Pick one pion from the τ system, and one pion from the hadronic system. In this case I define a **fake** ρ.
—> Pick a pions system 100% hadronic, thus having also a **fake** ρ candidate.



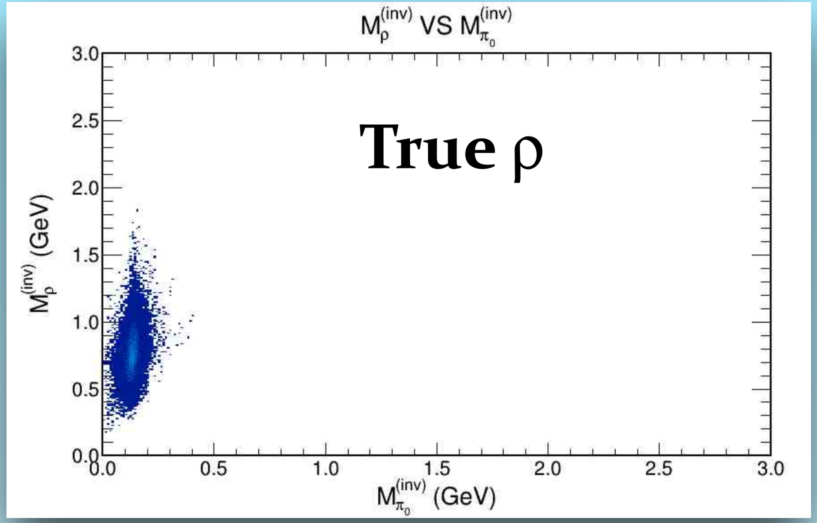
ν_τCC illustration

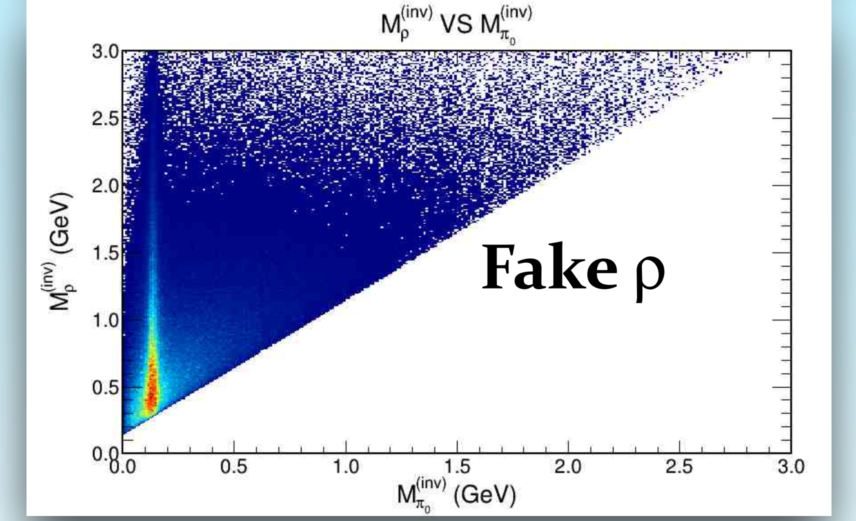
True and **fake** ρ are tagged **using the MCTruth**. The purpose of this denomination is for the ν_τCC(τ—>ρ) only, to assess our ability to recover the correct ρ when there is one.
In a NC event, there are only fake ρ candidates !

Developing a ρ tagging method with the invariant masses

• Let's exploit the scatter plot of the **invariant masses** ($\pi0$;ρ) (the $\pi0$ momentum is thus present in both invariant masses).

• We observe that the true ρ decay particles (top right) are distributed around (0.135 ; 0.776) GeV, the $\pi0$ and ρ true masses. All other ρ candidates are gathered in a single plot (bottom right), which we observe to be way more spread in the plane. Motivation for a **ranking method** based on the 2D cartesian distance:





$$d = \sqrt{\left(M_{\pi_0}^{(inv)} - m_{\pi_0}\right)^2 + \left(M_{\rho}^{(inv)} - m_{\rho}\right)^2}$$

Within an event, I refer as the **best** ρ, the ρ candidate which scored the smallest **d** . The best ρ can be either the true ρ, either a fake ρ.

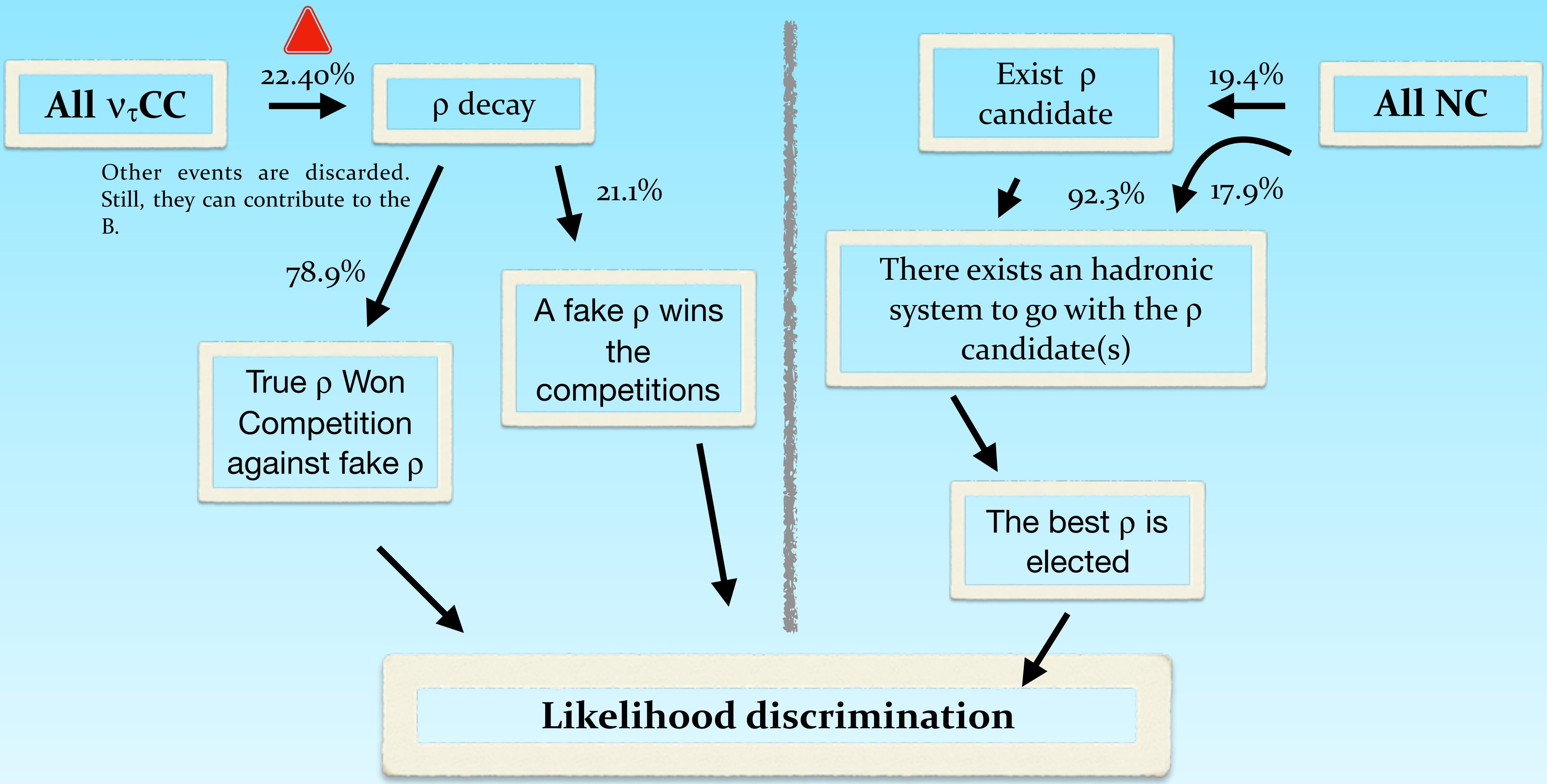
The ρ tagging method - performance

Among a sample of ν_τCC(τ—>ρ), based on the invariant mass ranking method, I found that :

- 53% had no ρ confusing candidate (no degeneracy), the true ρ is the only ρ candidate/combination.
- 26% had the true ρ scoring the best score (I don't specify the number of fake ρ candidates defeated in the process).
- 9% had the ρ scoring the 2nd score.
- 3% had the ρ scoring the 3rd score.
- 6% had the ρ scoring 4th score and more.
- 2% were discarded (it happens, for instance, when one of the ρ daughters undergoes smearing and gets a null energy, thus considered unreconstructed).

The correct ρ combination is selected 79% of the time, 21% contamination a this stage.

An overview of the S/B analysis



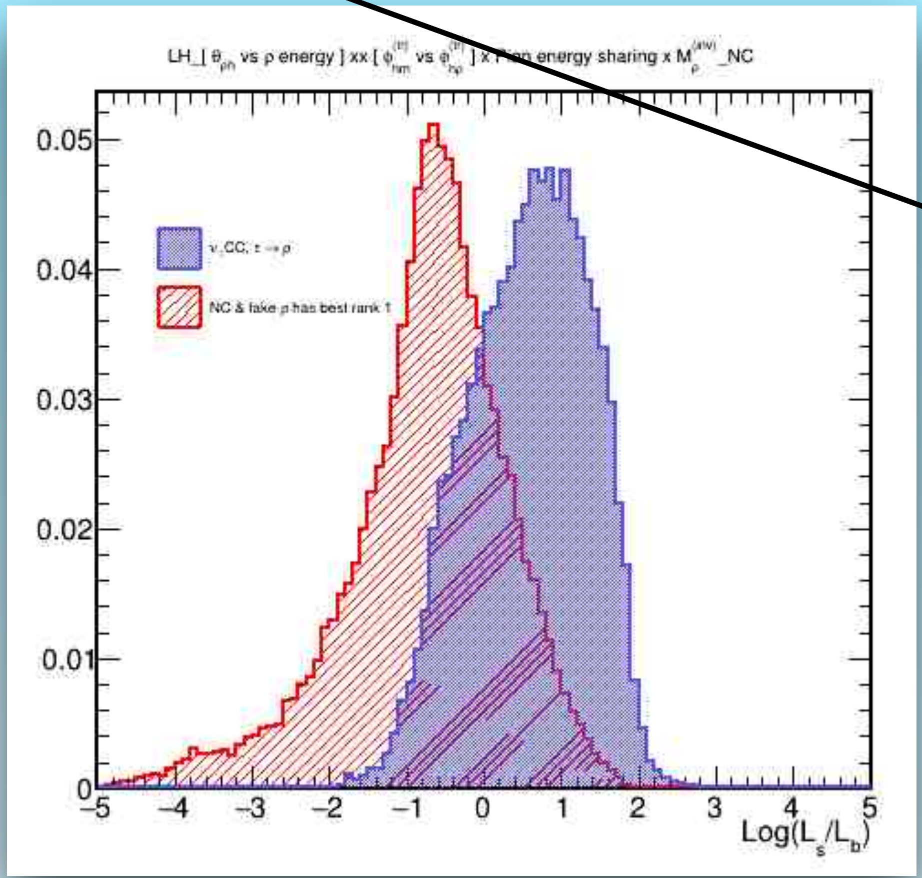
The final likelihood helps us decide whether the neutrino event, with its ρ candidate, is tagged as a signal (ν_τCC) or as a background.

This BR decay isn't in agreement with the PDG, expected 25.49%, stat. fluctuation of my number is of the order 0.1%. I will reweight it, but a specific investigation is ongoing.

Optimised likelihood search - no individual paticle energy smearing

$$\left[\theta_{\rho h};\rho_K\right]\times\left[\phi_{hm}^{(tr)};\phi_{h\rho}^{(tr)}\right]\times M_\rho^{(inv)}\times r_\pi$$

$$\left[\theta_{\rho h};\rho_K\right]\times\left[p_{miss}^{(tr)};p_{\rho}^{(tr)}\right]\times M_\rho^{(inv)}\times r_\pi$$

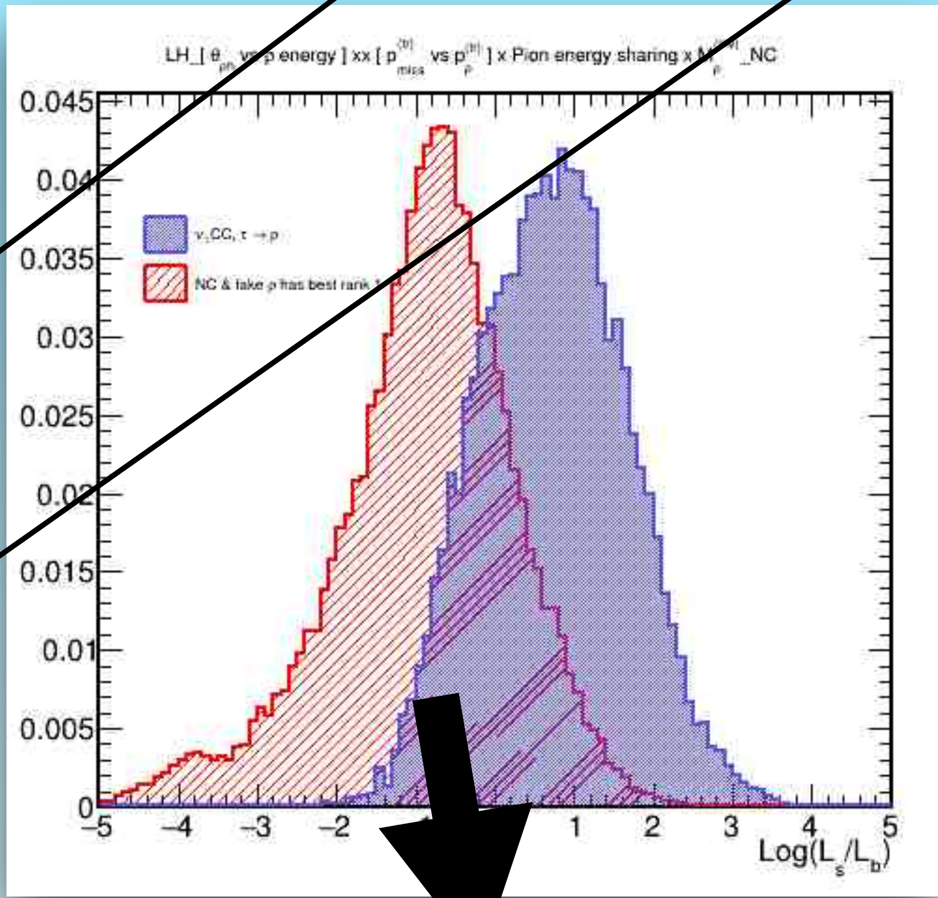


space angle between ρ daughters system and hadronic system

ρ energy

Transverse momenta

ratio of the pions energy sharing



Likelihood plots are normalised to 1. They inform on the discrimating power of the variables, but don't provide normalised S/B estimations.

LH Cut	0.2	0.4	0.6	0.8	1.0	1.2
S (%) fake ρ (%)	70.1 (6.3)	63.1 (5.1)	55.3 (4.0)	47.4 (3.0)	39.2 (2.3)	31.2 (1.6)
B (%)	15.1	11.0	7.8	5.3	3.2	2.1
S/sqrt(B)	1.80	1.90	1.98	2.06	2.16	2.15

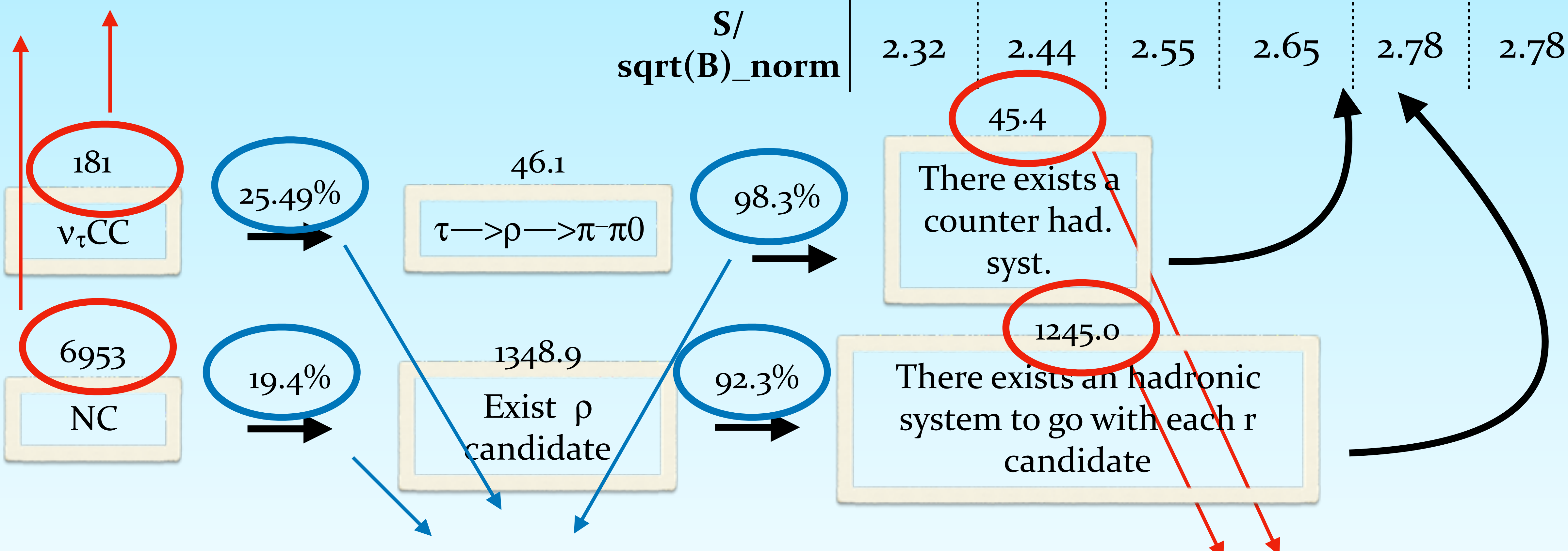
Optimised likelihood results - no individual particle energy smearing

$$\left[\theta_{\rho h};\rho_{energy}\right]\times\left[p_{miss}^{(tr)};p_{\rho}^{(tr)}\right]\times M_{\rho}^{inv}\times ratio$$

- 3.5 years staged
- 1.1e21 POT
- 1.2 MW beam
- DUNE CP violation optimised flux

LH Cut	0.2	0.4	0.6	0.8	1.0	1.2
S_norm (fake r)	31.8 (2.9)	28.6 (2.3)	25.1 (1.8)	21.5 (1.4)	17.8 (1.0)	14.2 (0.7)
B_norm	187.9	136.9	97.1	66.0	41.1	26.1
S/sqrt(B)_norm	2.32	2.44	2.55	2.65	2.78	2.78

expected DUNE FD number of events



These numbers don't correspond to kinematical/likelihood cuts. They merely reflect the neutrino events behaviors For instance, only 19.4% of the NC do contribute to the background, others don't provide ρ candidates.

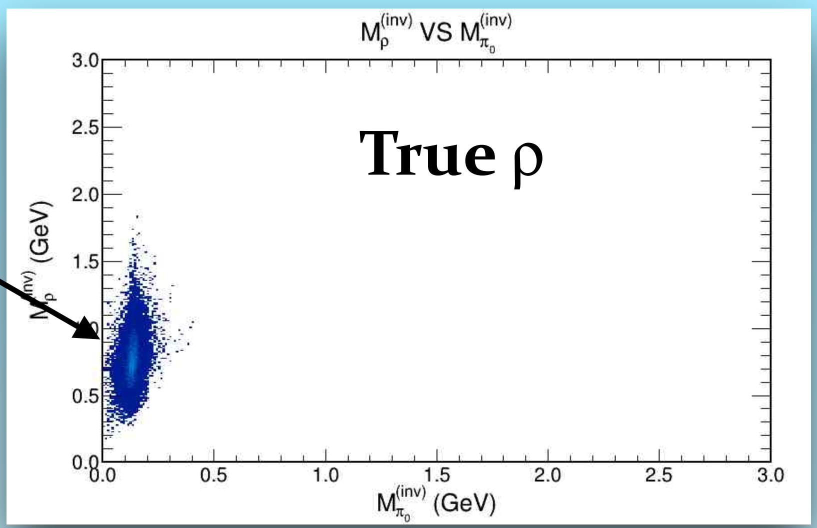
Actual S/B competition in my analysis.

Improvement possible ?

The likelihood does good work, however, about 20% of the ρ are confused before the likelihood discrimination, thus propagating a fake ρ in the analysis, that in the end represents ~[5 ; 10]% of the selected signal after likelihood cut.

Our ρ ranking method is based on the invariant masses of the pions system (π0;ρ), distributed around (0.135 ; 0.776) GeV for the signal.

There is more information available about the ($\pi^-\gamma_1\gamma_2$) that could help improve the ranking efficiency.



Invariant mass ranking method performance

Rank	-1	0	1	2	3	>3
%	2.3	52.9	26.0	9.4	2.9	6.3

52.9%+26.0% = 78.9% of the true ρ are correctly selected.

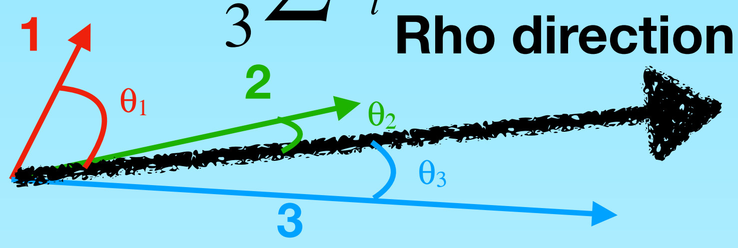
9.4%+2.9% = 12.3% of the true ρ didn't get that far from winning.

We can work on this fraction to improve.

Ranking Improvement : dispersion around ρ direction

• I observed that in most of the cases where the true ρ got defeated in the invariant mass ranking, the fake ρ candidate winning the competition had two out of three « true » particles.
For instance, the two photons were the actual daughter photons of the ρ, but the charged pion came from the hadronic system.

$$\bar{\theta} = \frac{1}{3}\sum \theta_i$$



- Thus, I compute mean(θi) to account for the fact that hybrid fake ρ candidates should be more scattered around the ρ direction.

• For each ντCC event, when the true ρ got ranked 1, 2 or 3, compare the 3 first ρ candidates with mean theta.

True ρ invariant mass ranking performance

Rank	-1	0	1	2	3	>3
%	2.3	52.9	26.0	9.4	2.9	6.3

Sample of true ρ which got 1st with the invariant mass ranking

Theta mean rank	1	2	3
%	69.8	24.9	5.3

Sample of true ρ which got 2nd with the invariant mass ranking

Theta mean rank	1	2	3
%	69.6	24.6	5.8

Sample of true ρ which got 3rd with the invariant mass ranking

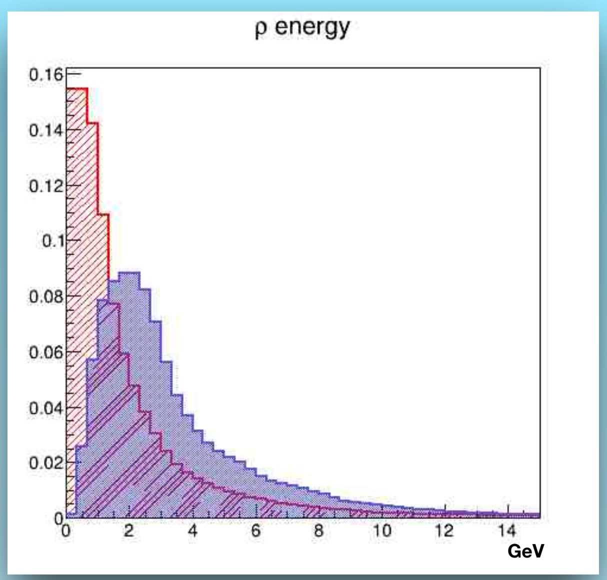
Theta mean rank	1	2	3
%	64.3	20.6	15.1

In about 2/3 of the cases, the true ρ defeats the two other fake ρ candidates by using the theta mean criteria.

Ranking Improvement : ρ energy

Previous observations showed that the true pair of pions had higher energy than the fake ones (blue vs red).

$$\rho_K = E_{K_{\pi 0}} + E_{K_{\pi \pm}}$$



Again, I look at the ρ energy competition between the three best ρ candidates in the cases the true ρ got ranked 1, 2 or 3 by the invariant mass ranking.

Rank	-1	0	1	2	3	>3
%	2.3	52.9	26.0	9.4	2.9	6.3

Sample of true ρ which got 1st with the invariant mass ranking

Theta mean	1	2	3
%	76.3	18.7	4.9

Sample of true ρ which got 2nd with the invariant mass ranking

Theta mean	1	2	3
%	73.8	20.2	6.0

Sample of true ρ which got 3rd with the invariant mass ranking

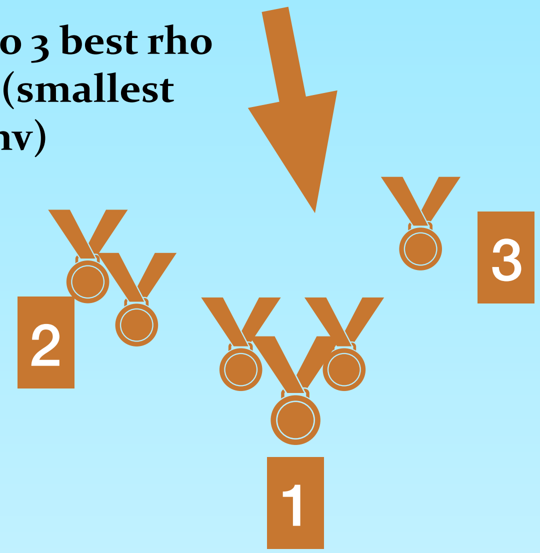
Theta mean	1	2	3
%	65.2	18.7	15.6

Medal Game motivation !

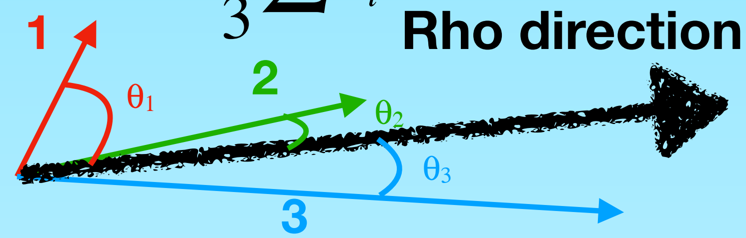
The Medal Game

$$d_{Minv} = \sqrt{(M_{\pi_0}^{(inv)} - m_{\pi_0})^2 + (M_{\rho}^{(inv)} - m_{\rho})^2}$$

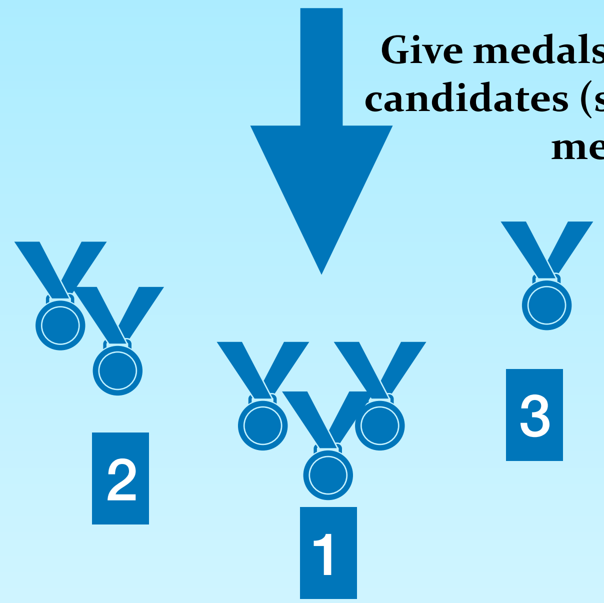
Give medals to 3 best rho candidates (smallest dminv)



$$\bar{\theta} = \frac{1}{3}\sum \theta_i$$

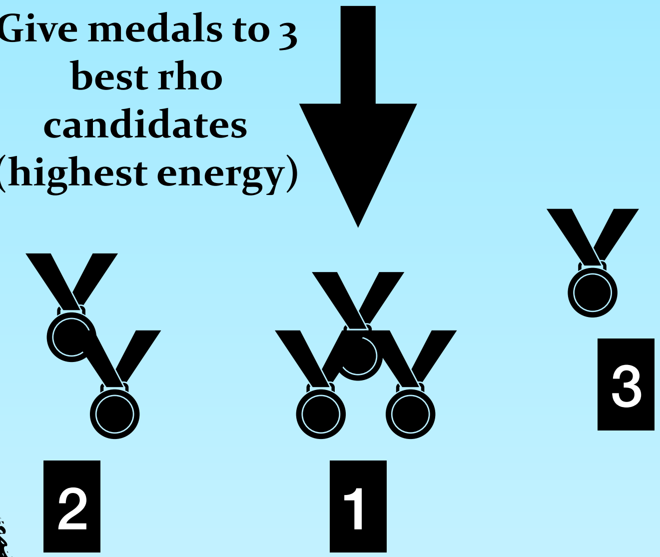


Give medals to 3 best rho candidates (smallest theta mean)



$$\rho_E = E_{K_{\pi 0}} + E_{K_{\pi\pm}}$$

Give medals to 3 best rho candidates (highest energy)



Invariant mass ranking method performance

Rank	-1	0	1	2	3	>3
%	2.3	52.9	26.0	9.4	2.9	6.3

→

Medal Game ranking method performance

Rank	-1	0	1	2	3	>3
%	2.3	52.9	30.7	6.3	1.9	5.9

Rank 1 improvement of 4.7% out of the objective of 12.3%. Slight improvement, correct selection from 78.9% to 83.6%.

Conclusion

• I presented an analysis for the ν_τ search based on kinematics for the ρ resonant decay mode, exploiting its large branching ratio (25%). It follows and extend an analysis I had made for the leptonic decay modes.

• The main background components of this process are the NC with final state pions that can mimic the true ρ signature. We exploited the rich kinematic information to discriminate between S and B. It appears there is the possibility to select a sample of ~18 ντCC events in a 3.5 years staged run, while having a NC contamination of ~41 (S/sqrt(B)=2.78).

• Discussion around the ρ selection performance with respect to combinatorics with the hadronic systems. Various attempts and efforts and finally a slight improvement using the **invariant masses**, the ρ **energy** and some **angle dispersion** (transverse angles didn't prove effective).

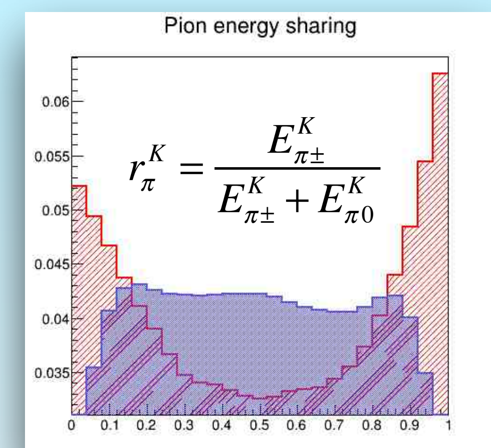
• Likelihood analysis must be run again using the improved ρ ranking method, since for the background (NC), the selection might bias differently the distributions of the variables used, thus affecting likelihood cuts.

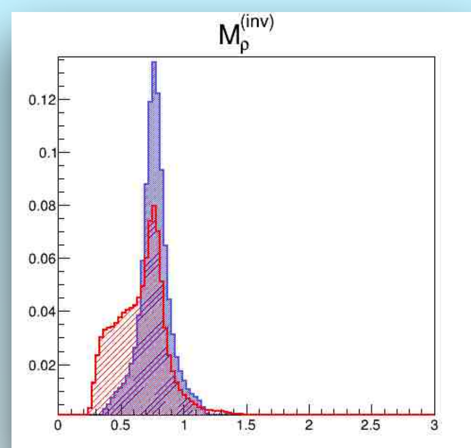
BACK UP

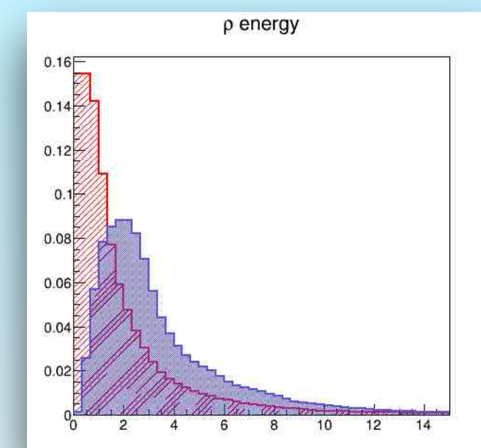
The set of kin. variables used ντCC true ρ VS NC best ρ

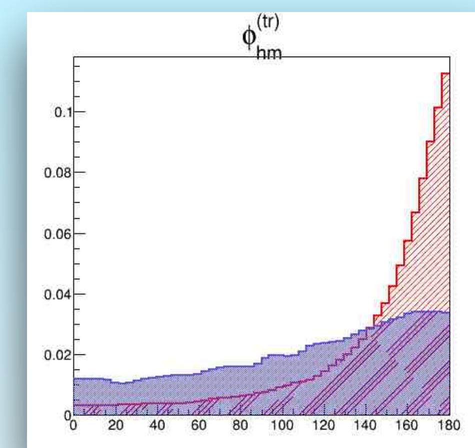
I use **16 kinematical variables**, including:

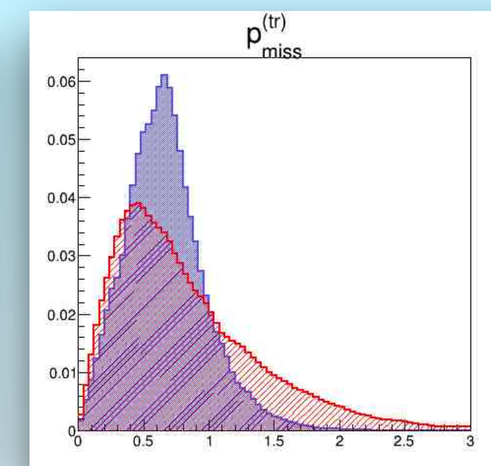
- Pions energie the sum of both (that I call the ρ energy, a terrible name), + pion energy sharing. $r_\pi^K = \frac{E_{\pi\pm}^K}{E_{\pi\pm}^K + E_{\pi 0}^K}$
- Invariant masses for π0 and (π0π±) systems. $E_{\pi_0}^K; E_{\pi^\pm}^K; \rho_K; r_\pi^K$ $M_{\pi_0}^{(inv)}; M_\rho^{(inv)}$
- Various space angles (θ) between system momenta : ρ, h(hadronic), total, ν (beam direction). Some of these angles are representative of the isolation of the rho candidate with respect to the hadronic system. $\theta_{\rho h}; \theta_{\rho tot}; \theta_{h\nu}; \theta_{\rho\nu}$
- Transverse plane information of had. syst., ρ syst. and missing component (modulus of the momentum, plus relative direction with angle ϕ). $p_\rho^{(tr)}; p_{had}^{(tr)}; p_{miss}^{(tr)}; \phi_{h\rho}^{(tr)}; \phi_{hm}^{(tr)}; \phi_{m\rho}^{(tr)}$









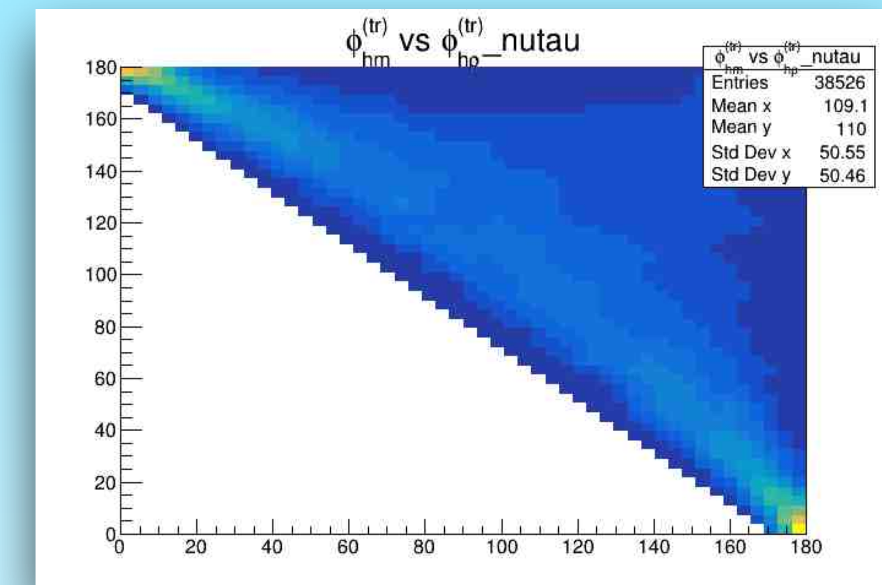


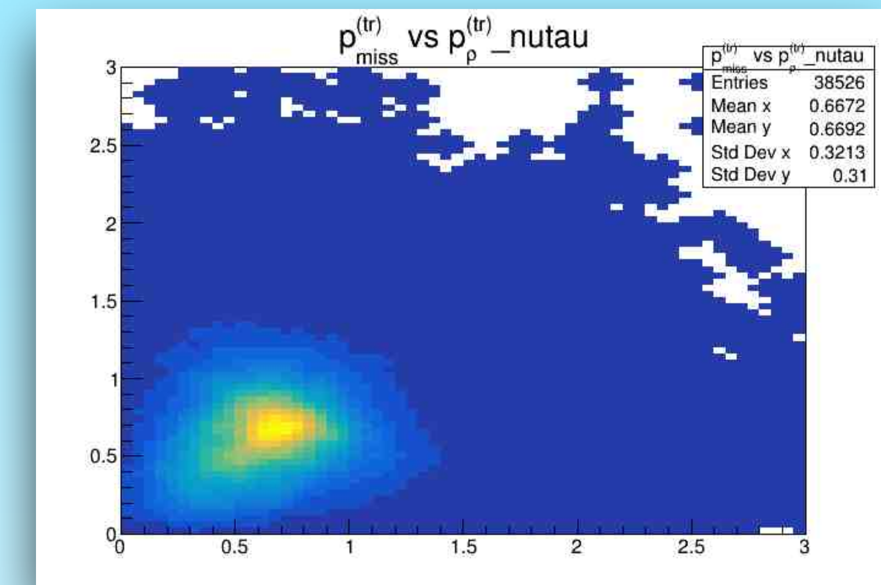
- Pion energy sharing: S has ~flat distribution, while B has asymmetry in the energy sharing.
- Red ρ invariant mass dist. has a strange shape because of the ρ selection criteria (cf. slide 4).
- Pions of S have higher energy than B.
- In the transverse plane, the missing momentum is ~180° with respect to the hadronic momentum for NC (which is what we expect, since their is no true ρ system)

The set of kinematic variable : 2D correlations

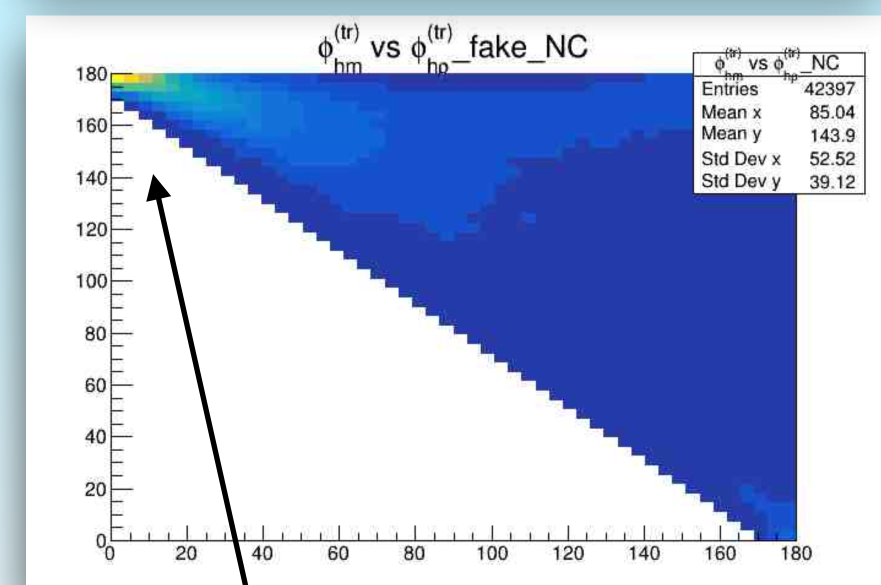
To improve the discrimination power, one can have a look at 2D correlations of variables. Some illustrations

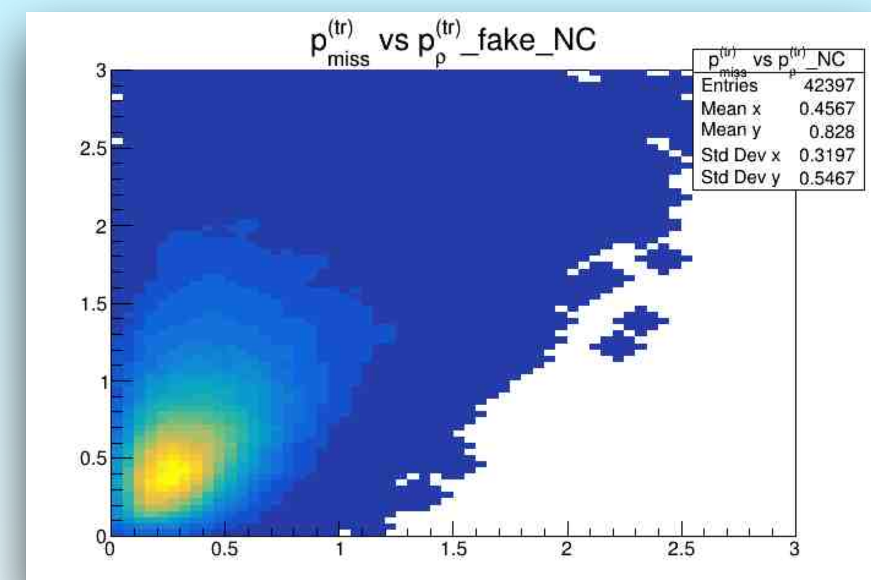
Signal





Background





For NC: ~0° transverse rh angle and ~180°.
Signal is more spread.

Different region of the plane prefered.