Status of the muon reconstruction

Grigory Safronov
(JINR)

Baikal-GVD collaboration meeting, JINR, 28 - 30 November 2017

Outline

- Available track reconstruction methods
- Results on data-MC comparison for runs 110-150, 2016
- Track quality cuts and background rejection
- New, "loose", method of the hit selection
- New reconstruction performance
- Muon reconstruction status and plans

Muon track reconstruction

Track parametrisation: $\vec{R(t)} = \vec{R_0} + c(t - t_0)\vec{V}$

$\vec{V}$ is unit vector in polar coordinates, $\vec{R_0}$ and t_0 are taken in plane Z=0

Initial approximation: $\vec{V_{init}} = (1/|\sum_{t_i<t_j} \vec{R_{ij}}|) \sum_{t_i<t_j} \vec{R_{ij}}$ where *i* and *j* are ordered in time and belong to different strings

Initial values for $\vec{R_0}$ and t_0 are obtained using hit with largest amplitude

Minimisation function:

$$Q = \sum_{i=1}^{N_{hit}} \left[\frac{(t_i^{est} - t_i)^2}{\sigma_t^2} + \frac{A(a_i)D(d_i)}{d_0} \right]$$

inspired by ANTARES paper: https://arxiv.org/abs/1105.4116

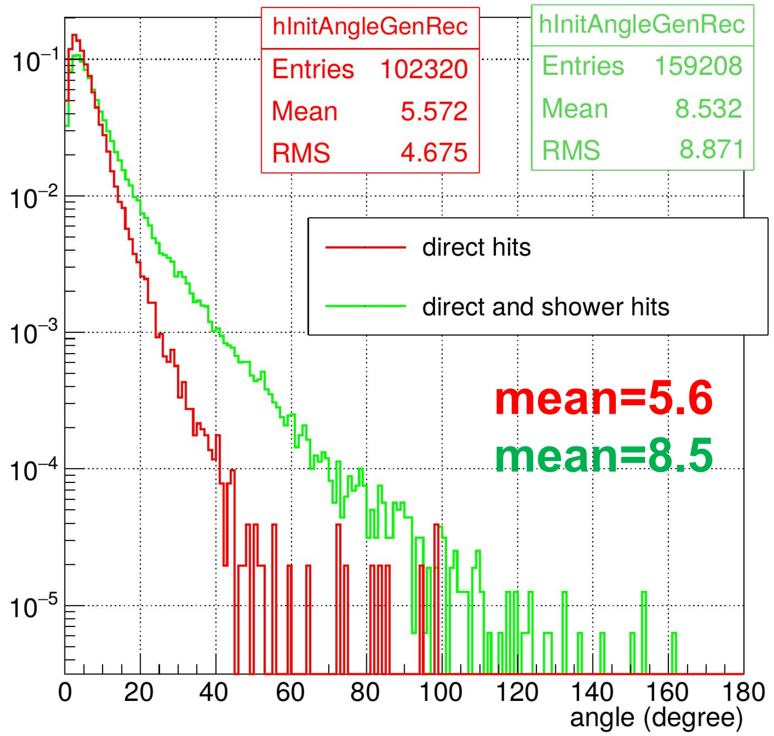
Direct cerenkov approximation for the time estimation
A and *D* are amplitude and distance functions

Performance with MC truth

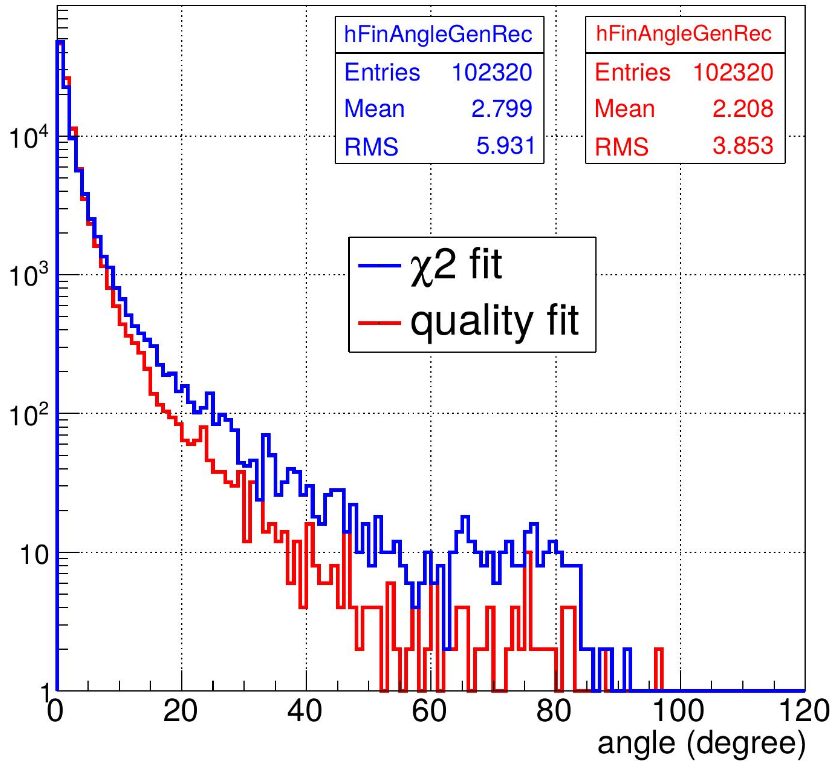
Events with at least 6 hits on 3 strings are selected (here and further in this talk)

MC truth -based selection of direct cerenkov hits and shower hits

mismatch angle for initial approximation



fit results for direct hits



median angle for quality fit: 1.3

Hotspot hit selection I

Pairs of adjacent hits (hotspots) are selected if:

- Two hits in adjacent modules fall into causality window: $|\Delta T|<(n\Delta R/c)+t_S$, ΔR=15 m, t_s - "safety time window" =30ns

- Hits 1 and 2 should have total change deposit $>A_1$ and $>A_2$

It is important to adjust parameters t_s , A_1 and A_2 to have low noise contribution

- Main parameters for noise suppression are A_1 and A_2
- $A_{1,2}$ > 1.5 p.e. provide purity better than 98%

Hotspot hit selection II

Hotspots are merged as follows:

1. All possible pairs are constructed at each string (same impulse may belong to 2 hotspots)
2. Combinations of hotspots (groups) of not less than 6 OM's in total are considered
3. Initial trajectory approximation is reconstructed for each group
4. Quality function is calculated for initial trajectory approximation
5. Group which gives the smallest [quality function/nHits] ratio is selected for further analysis

Hotspot hit selection III

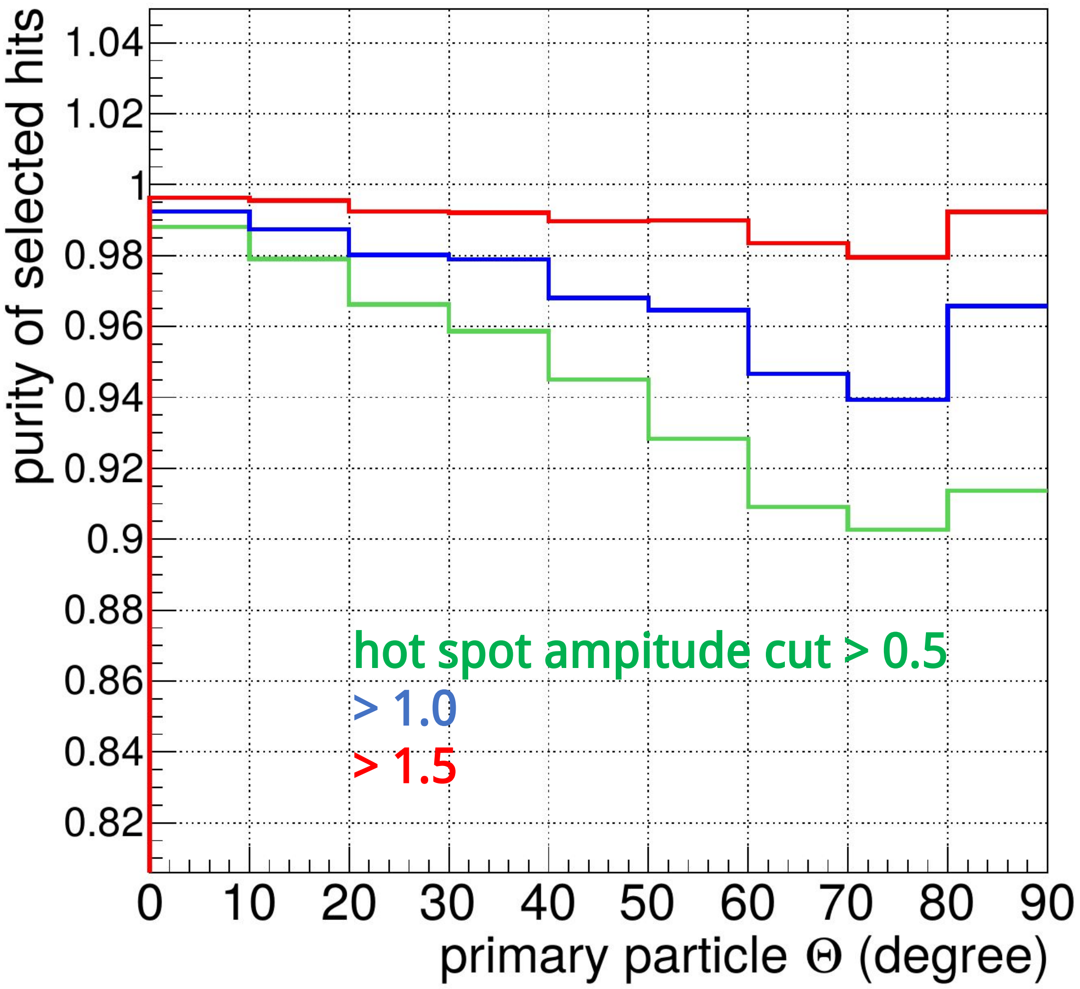
Purity of merged groups:

- Fraction of signal hits (direct or shower) in obtained global groups

Few symmetric amplitude cuts:

$A_{1,2} > 0.5, 1.0, 1.5$ (p.e.)

With $A_{1,2} > 1.5$ purity of > 98% is achieved

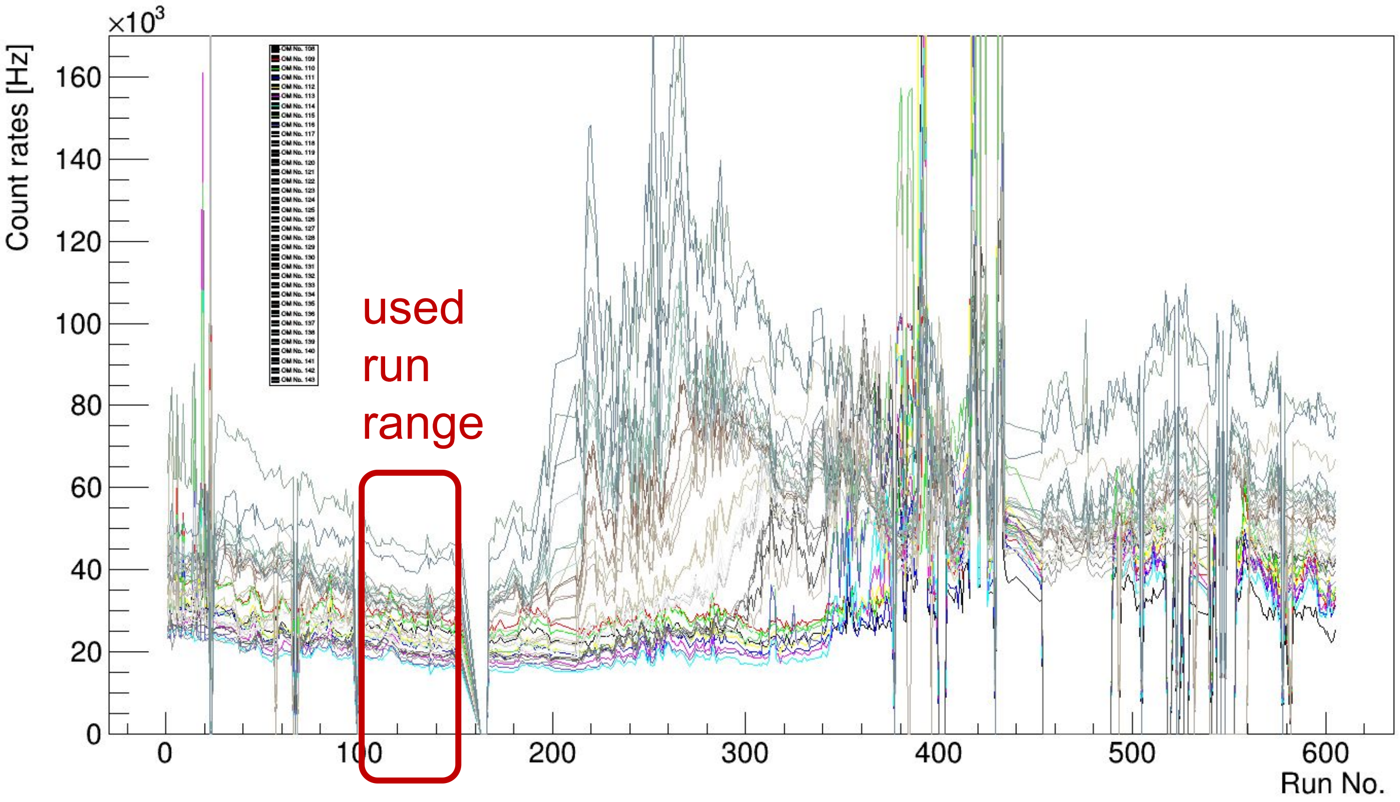


Atmospheric muon reconstruction in data and MC

Analysed datasets

DATA: year 2016, run period 110-150

- 27 runs from this period were processed
- ~272h (11 days) exposition time
- relevant time, amplitude and OM position calibrations are applied



Monte Carlo:

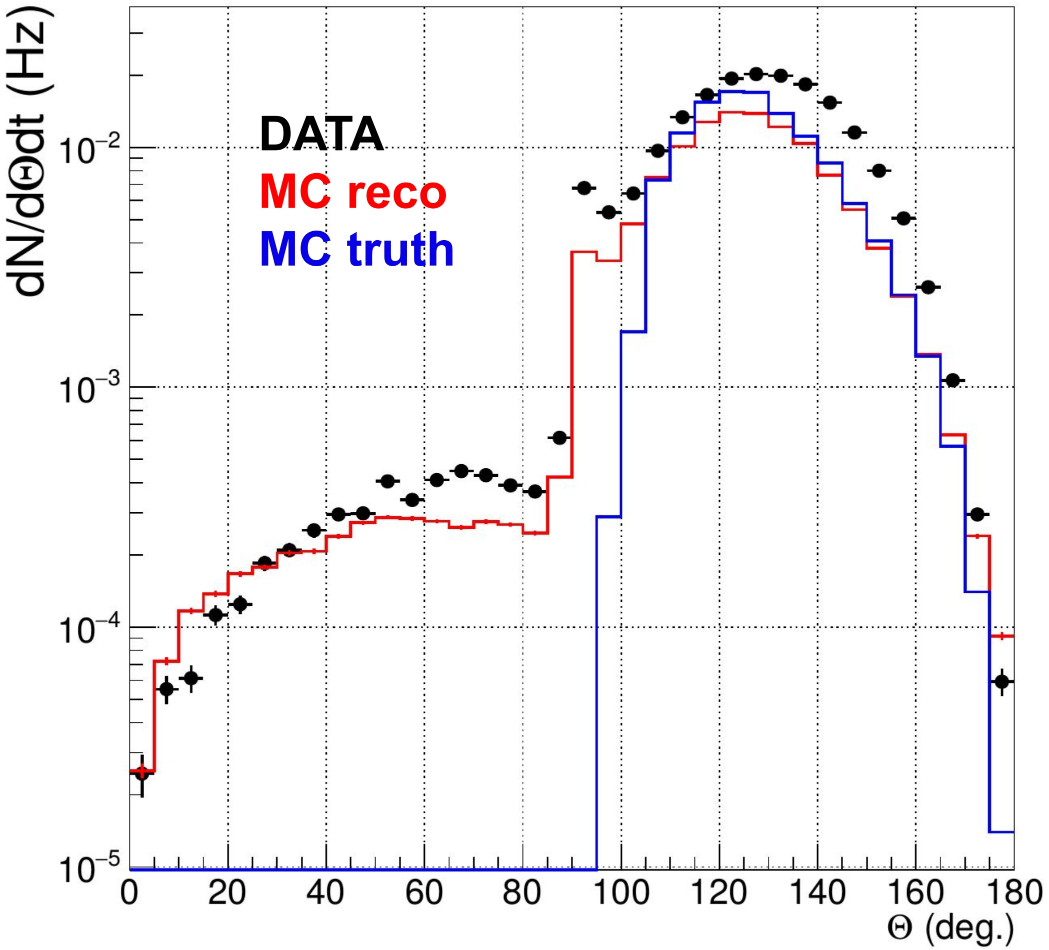
upgoing atmospheric neutrino MC

- neutrino energy > 500 GeV
- exposition time 1.26e10 sec
- ideal geometry and calibration,
- realistic OM noise

downgoing atmospheric muon MC

- Muon flux from CORSIKA
- 8.3e6 sec (~90 days) exposition time
- ideal geometry and calibration,
- realistic OM noise

Rate of reconstructed tracks



Absolute rates for data and MC:

~ 11 days of exposition in data

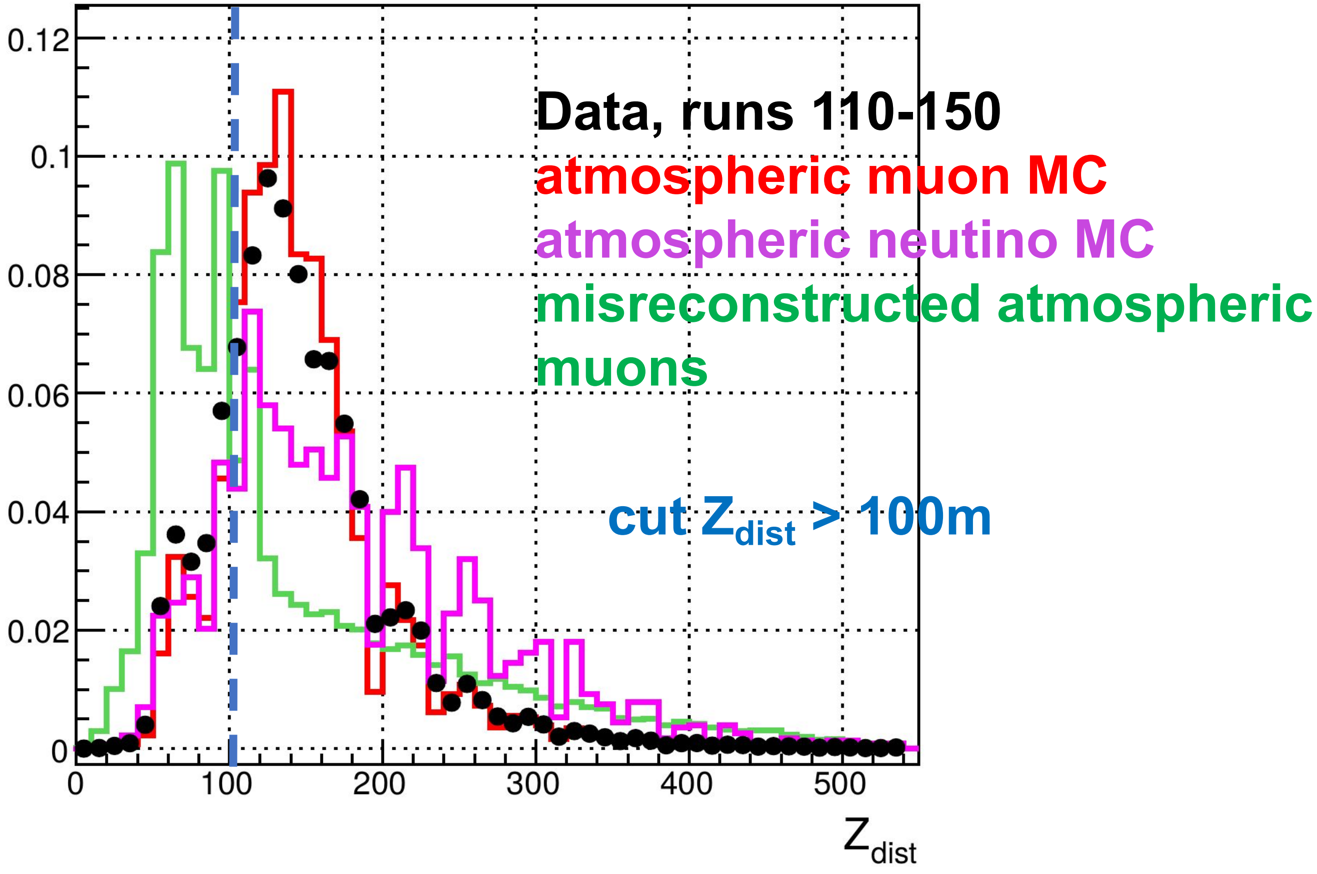
~ 90 days of exposition in MC

Total data rate is 1.5 times larger than MC (to be understood)

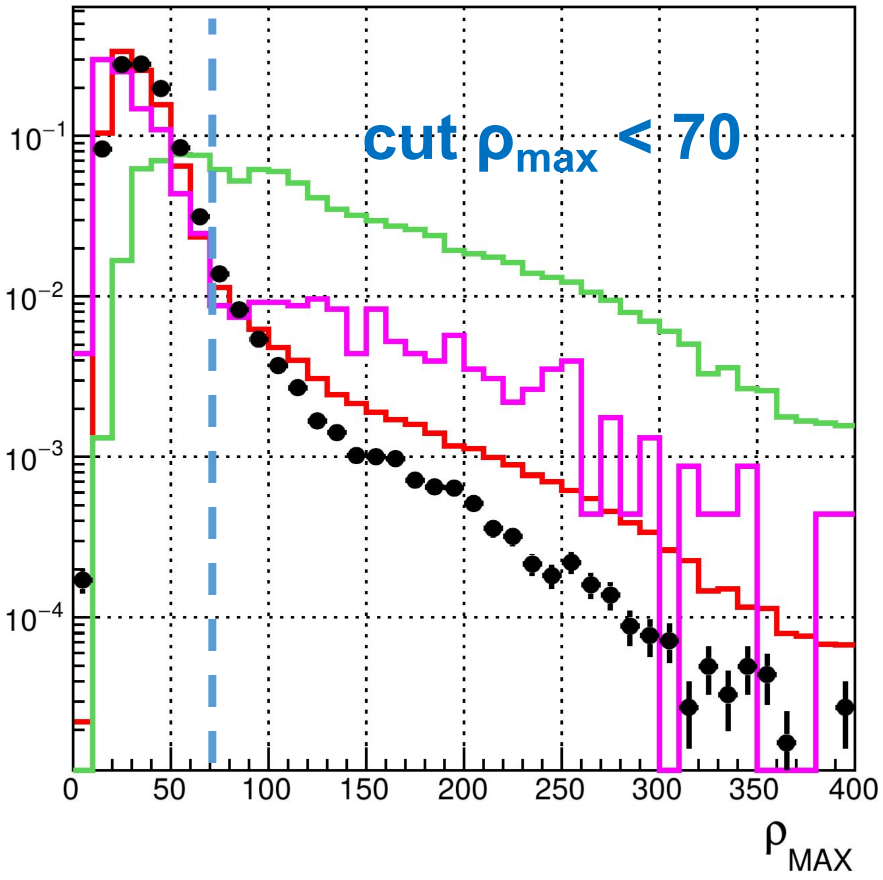
Fraction of misreconstructed as upgoing tracks is ~ 3.3% both for data and MC

Track quality variables I

Z_{dist} - largest distance between projections of used OM's on the track



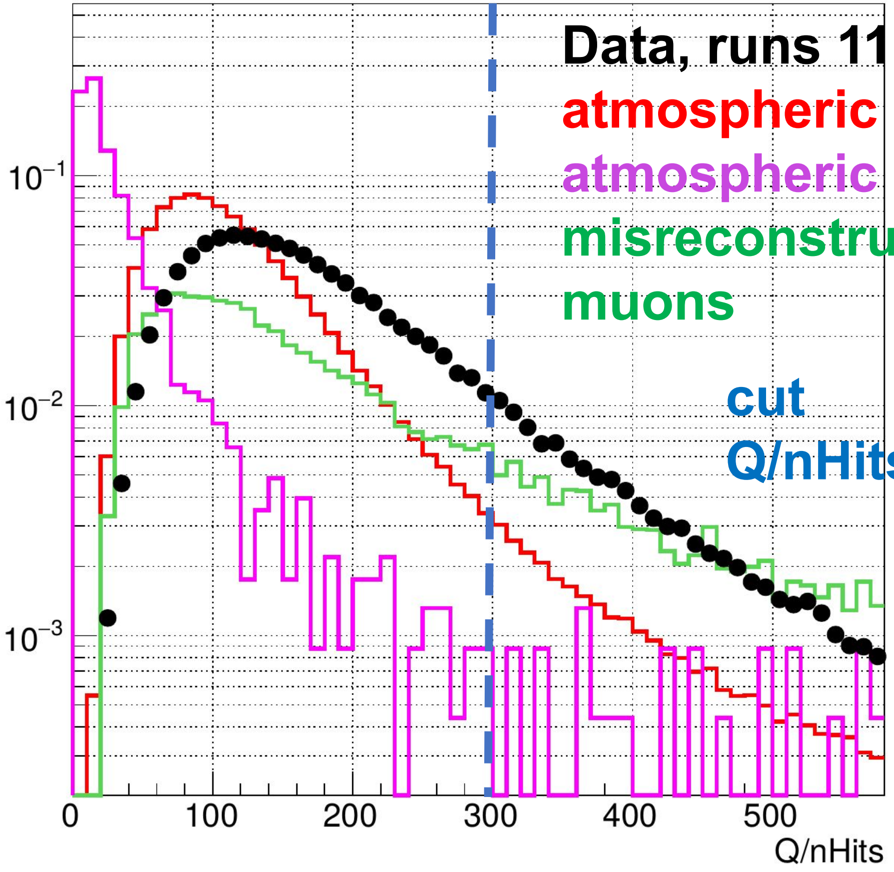
ρ_{max} - maximum distance from the track to the used OM



good data-MC agreement

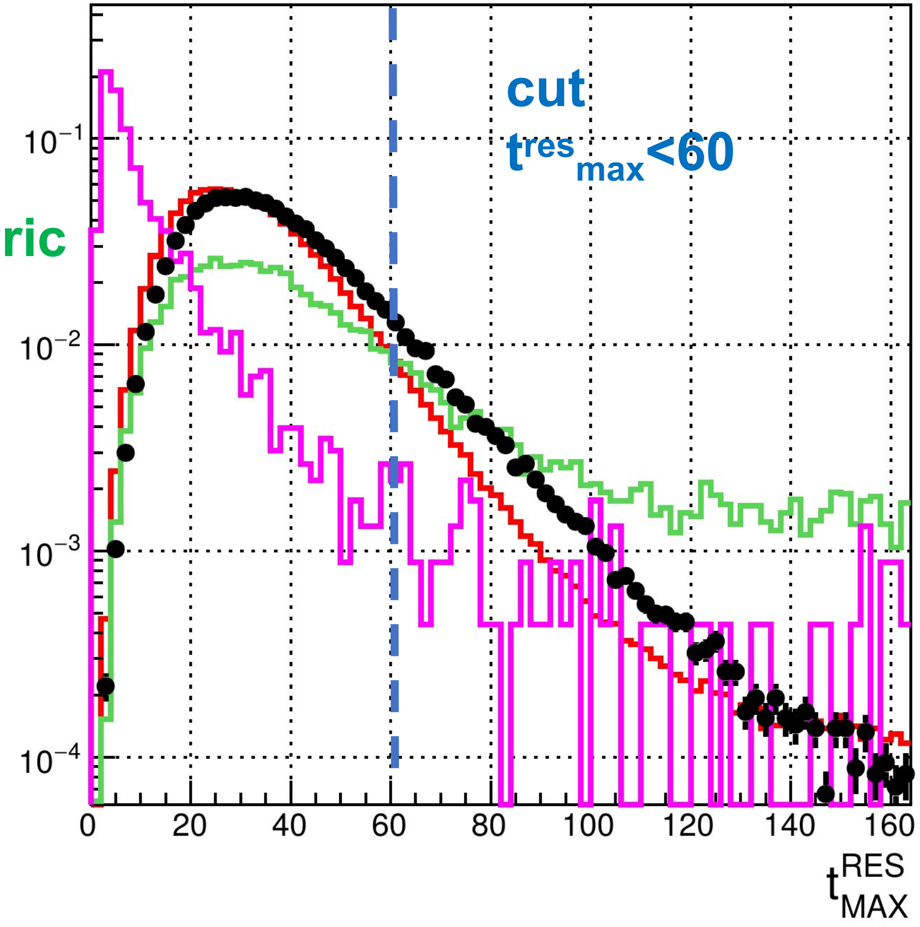
Track quality variables II

Q/nHits - quality function divided by the number of used hits

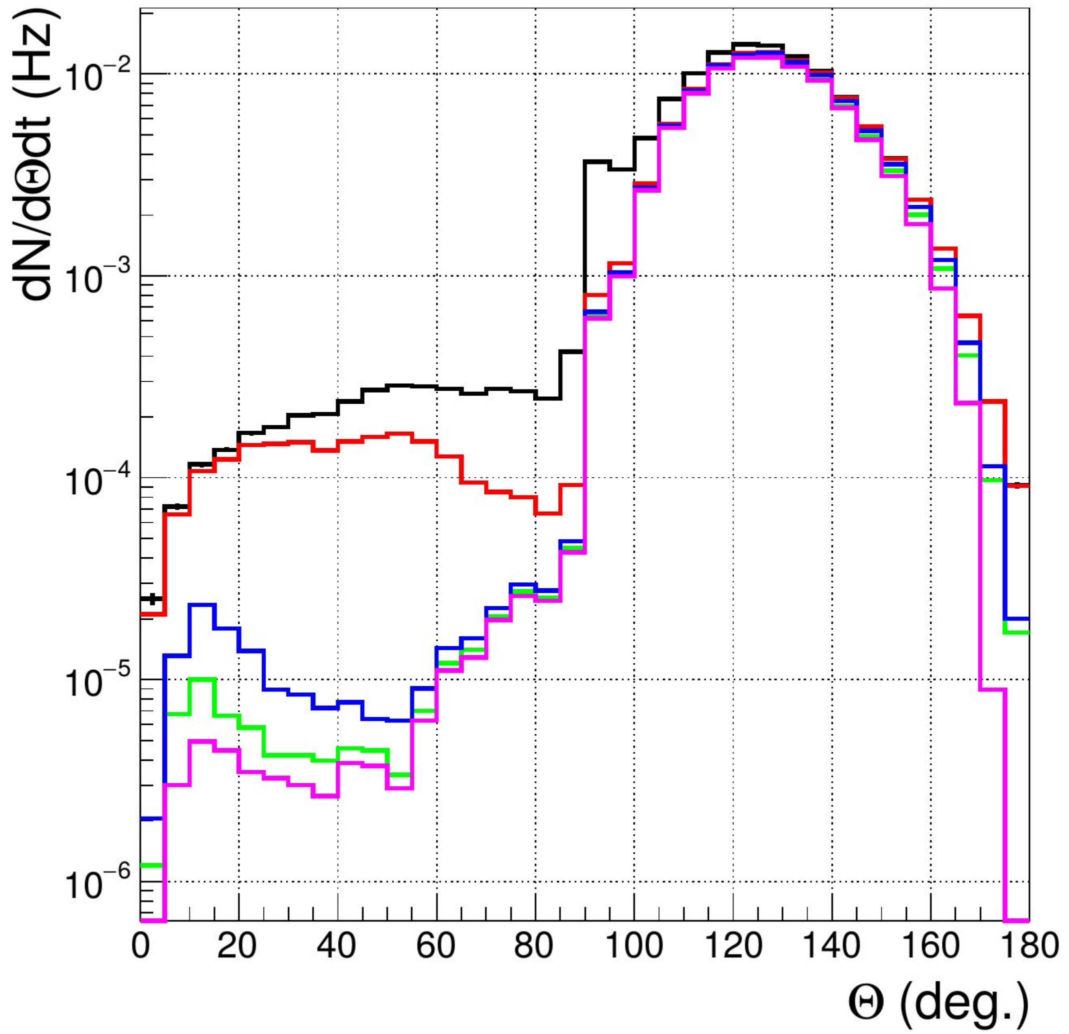


poor data-MC agreement for quality

t^{res}_{max} - maximum residual time among used hits



Track quality variables III



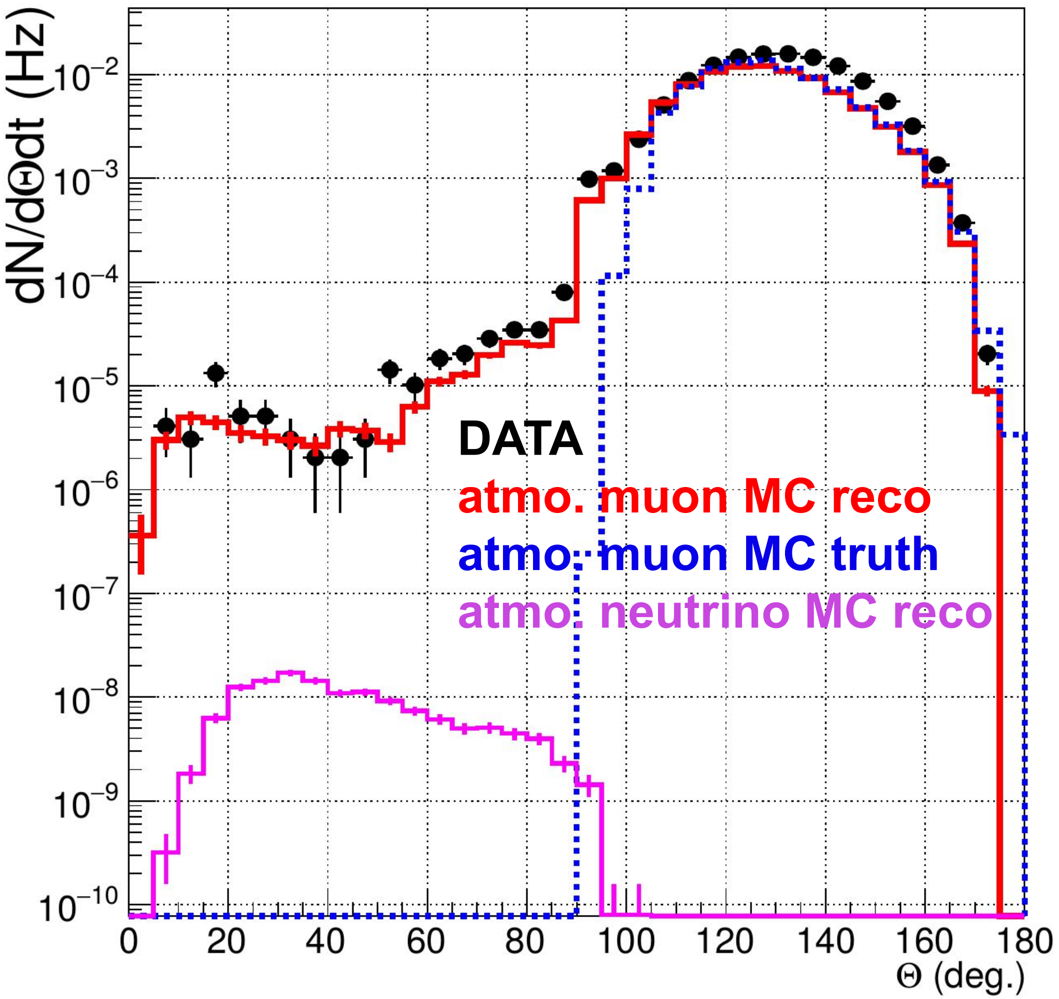
Effect of cuts on polar angle distribution

all tracks
Z_{dist} > 100m
ρ_{max} < 70 && Z_{dist}>100m
t^{res}_{max}<60 && ρ_{max} < 70 && Z_{dist} > 100m
Q/nHits<300 && t^{res}_{max}<60 && ρ_{max} < 70 && Z_{dist} > 100m

Fraction of misreconstructed as upgoing tracks is reduced to ~ 0.2% (MC)

76% of downgoing tracks have survived

Rate after quality cuts



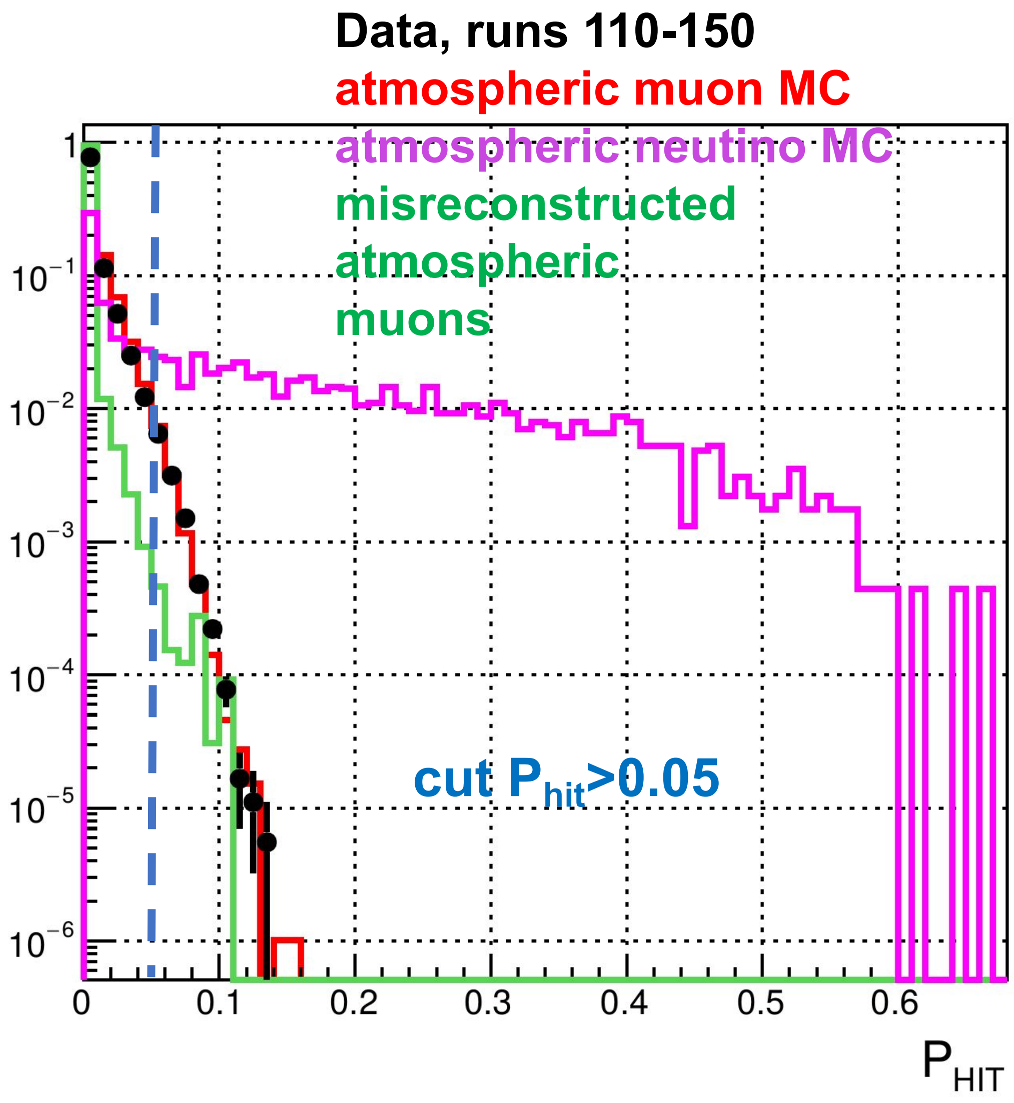
Absolute muon rates in comparison to atmospheric neutrino rate

Total rate of muon tracks ~ 0.15 Hz

Background from misreconstructed groups is still >100 times larger than neutrino signal

Neutrino signal efficiency ~85%

P_{hit} variable



P_{hit} - probability that muon with current trajectory has fired the collection of OMs used in the fit

In the assumption of direct cerenkov signal

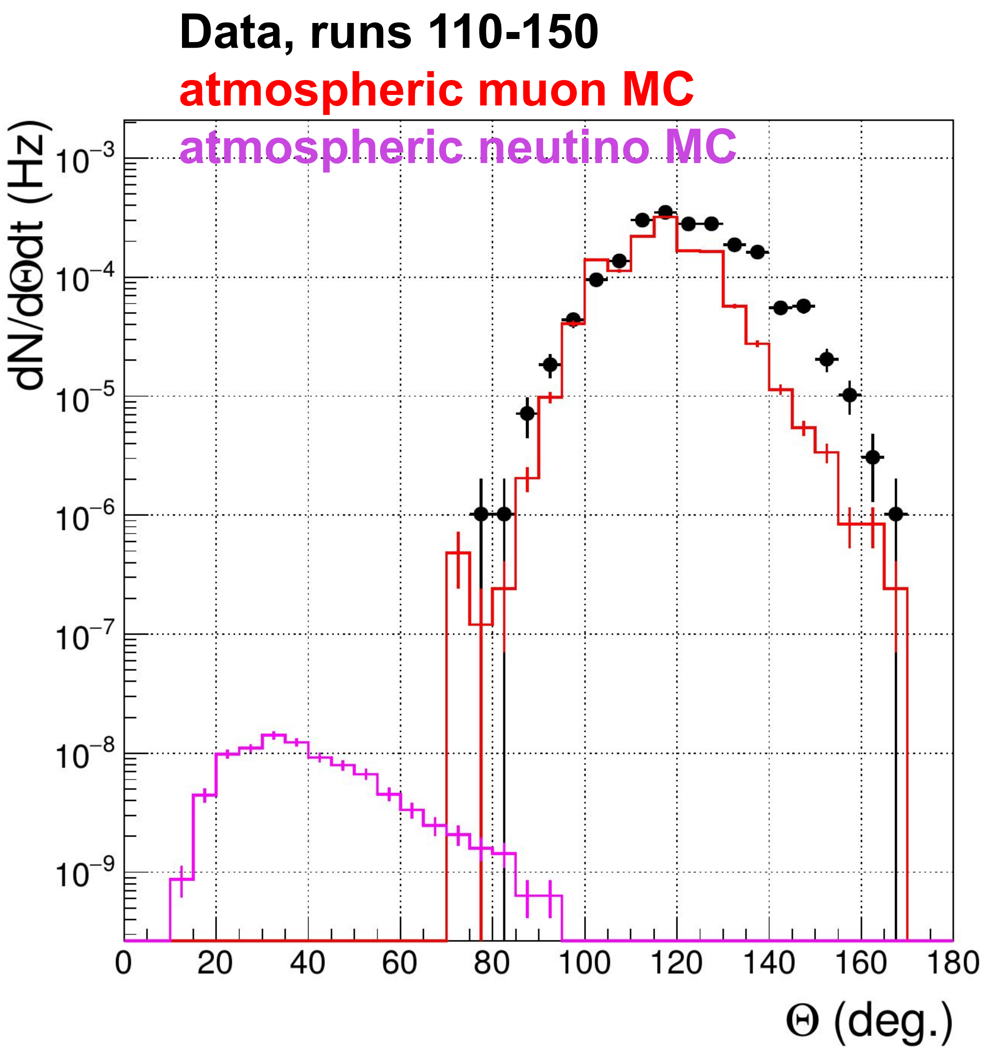
$$P_{hit} = \left(\prod_{i=1}^{N_{hit}} p_i(r)\right)^{1/N_{hit}}$$

Good agreement between data and MC

Strong discriminating power between downgoing groups and upgoing tracks

Cut of **P_{hit}** > 0.05 was applied to data and MC sets

Results



Misreconstructed tracks are removed

99% of muon groups is removed

Neutrino signal efficiency: 55%

No upgoing events in data

Need to increase the data sample and loosen the hit selection

Further development of the reconstruction

New method for hit selection

Available hit selection procedure is too tight
Sensitivity is small - candidate neutrino events were rejected

Need to modify hit selection to obtain good purity of hits used for the fit with as high as possible efficiency of the selection

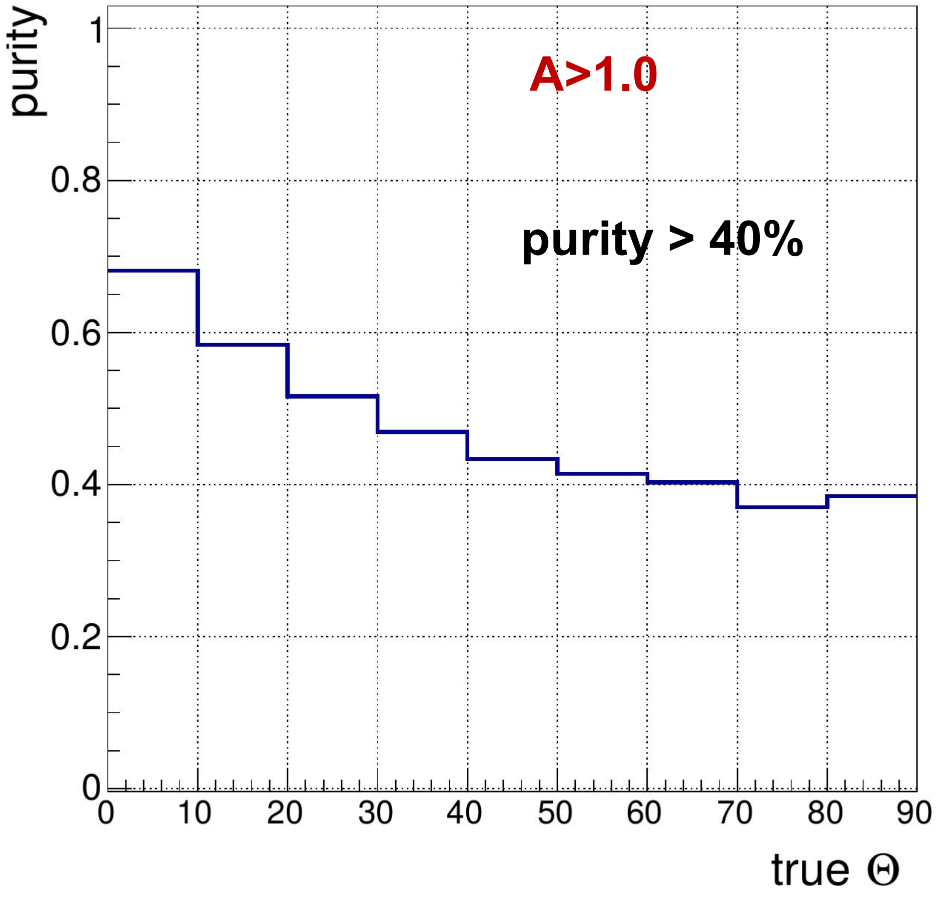
purity - fraction of signal hits in selected group for events with at least 6 selected hits

efficiency - fraction of hits with A>1.6 pe which were picked by the algorithm for events with at least 6 signal hits

The study was done on upgoing atmospheric neutrino MC sample
6 hits on at least 3 strings selection

Causality criterium

First step - apply causality criterium



$|\Delta T| < (n\Delta R/c) + t_s$; $t_s = 30$ ns

Loop over all impulses in the event

- Construct causally connected group starting from each impulse
- All pairs inside the group must be causally connected
- Each group is passed to further processing
- After further processing the group which has the best quality/nHits ratio is selected for trajectory fit

At least 6 hits at 3 strings:
54% of events pass this selection

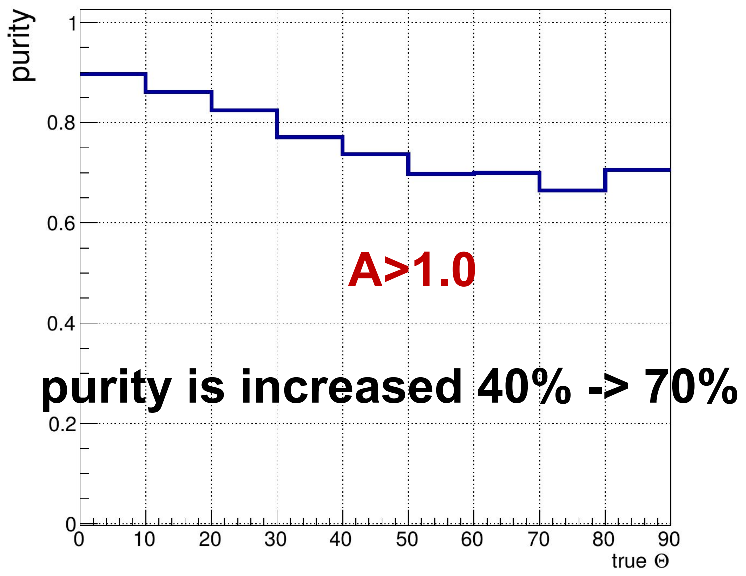
Removal of residuals I

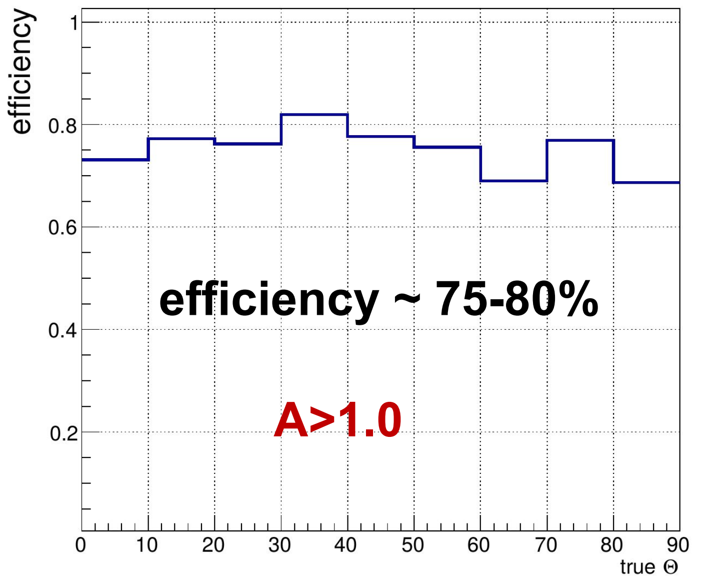
Do the following procedure 7 times with gradual tightening of cuts on $\mathbf{t_{res}}$ and $\boldsymbol{\rho}$

- Loop over the collection of hits inside the group
- Exclude hit and estimate the trajectory direction without it
- Calculate residual time ($\mathbf{t_{res}}$) and distance ($\boldsymbol{\rho}$) for the hit
- Exclude hit from the group if $\mathbf{t_{res}}$ < cut or $\boldsymbol{\rho}$ < cut
- Most tight cuts: **tres**<300 s, $\boldsymbol{\rho}$<60 m

If there are still > 6 hits in the group go to the next stage

With this procedure purity is increased to 70%





Removal of residuals II

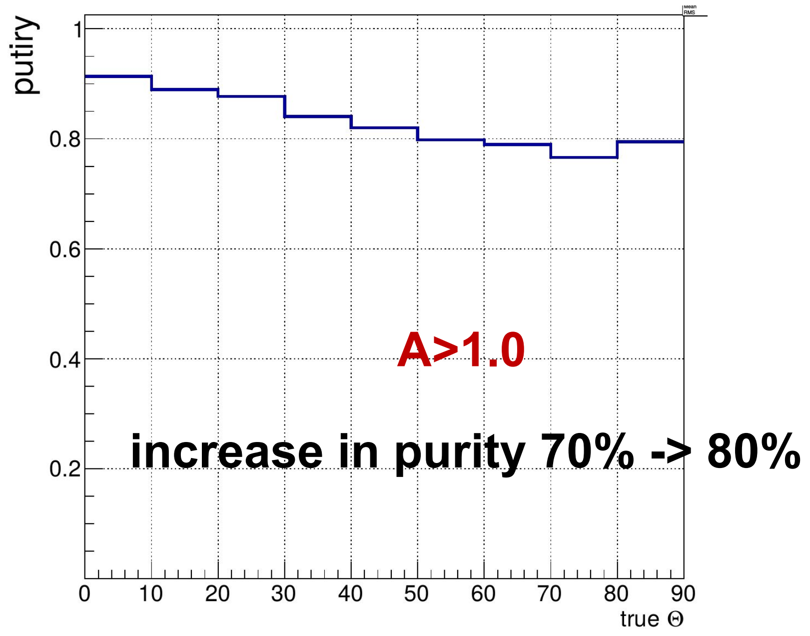
Almost similar procedure using phit variable

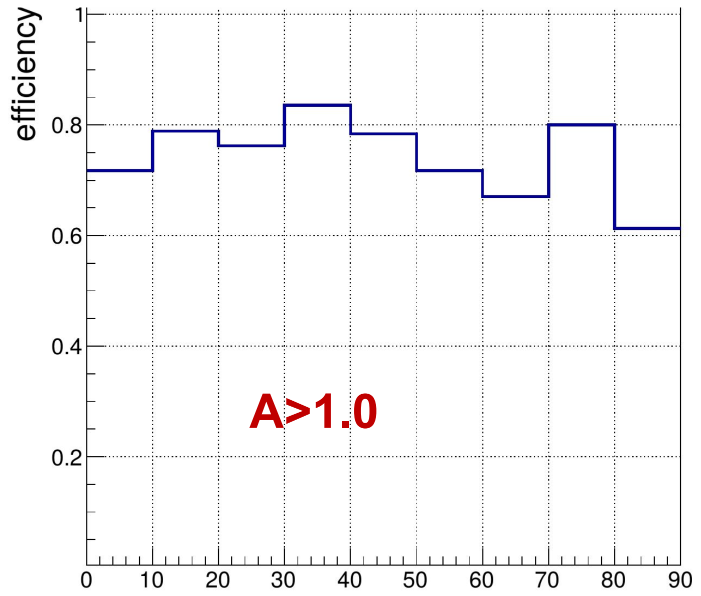
Do the following procedure 5 times
with gradual tightening of cuts on $\mathbf{p_{hit}}$

- Estimate the trajectory direction
- Loop over hits
- Exclude ones with very low probability for the hit
- In 5 steps, tightest cut phit > 1e-5

That's the final step. After that group with the best quality/nHits ratio is selected

10% of events pass 6 hits at 3 strings selection

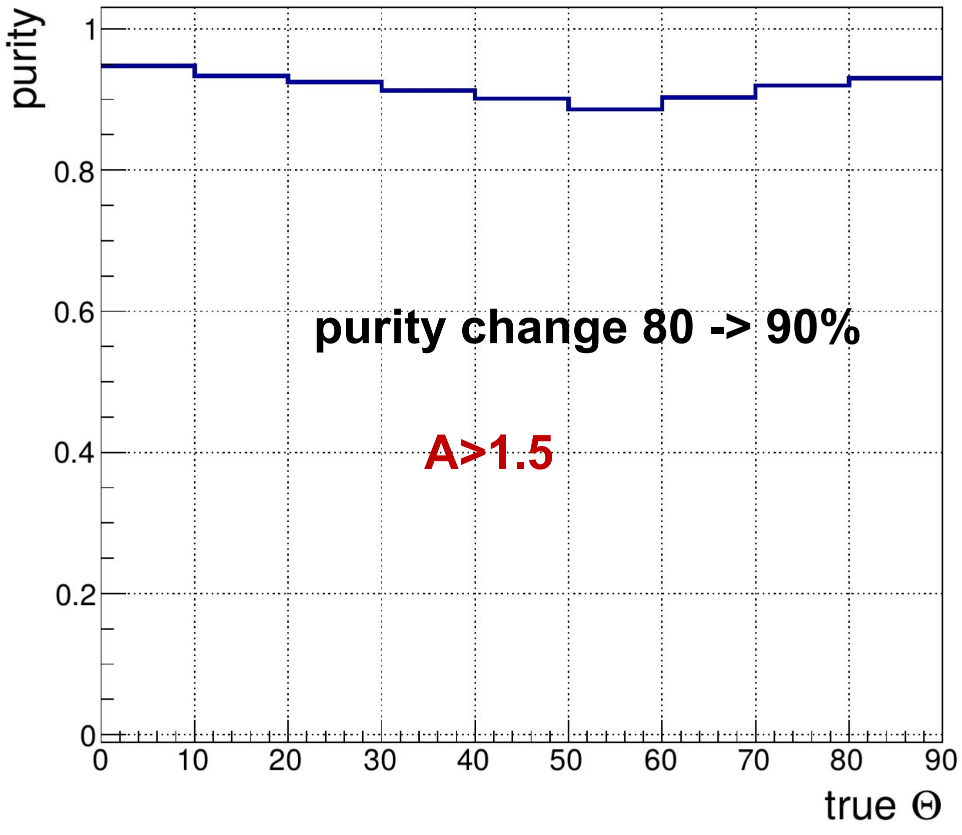


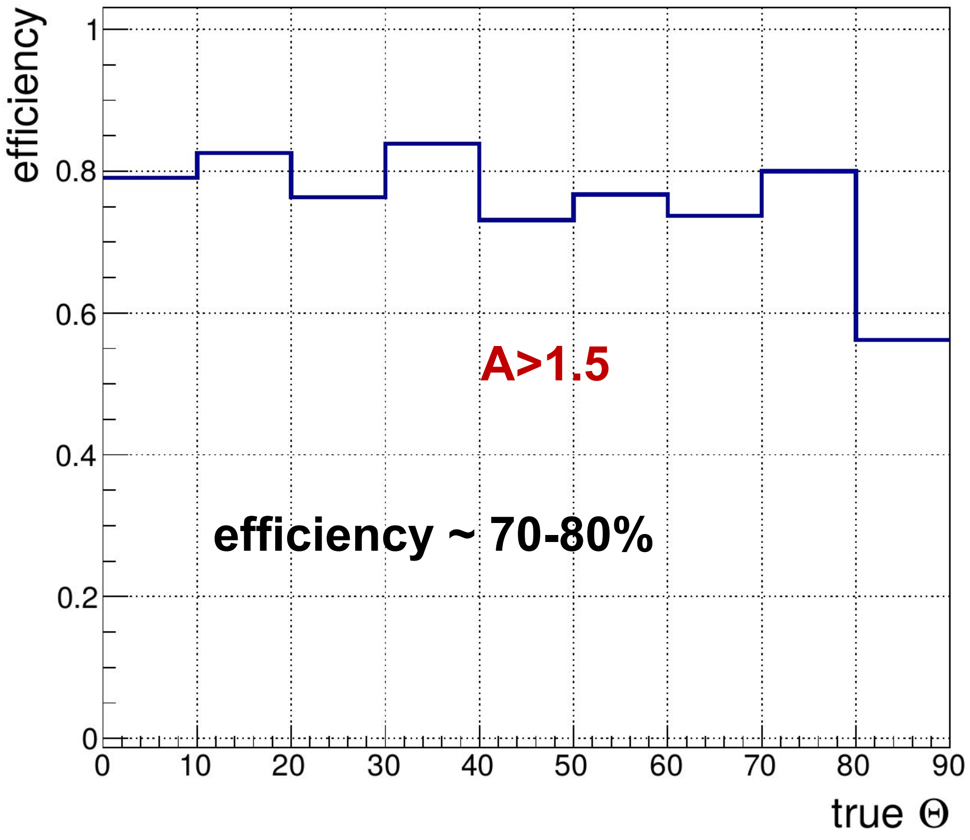


Amplitude cut dependence

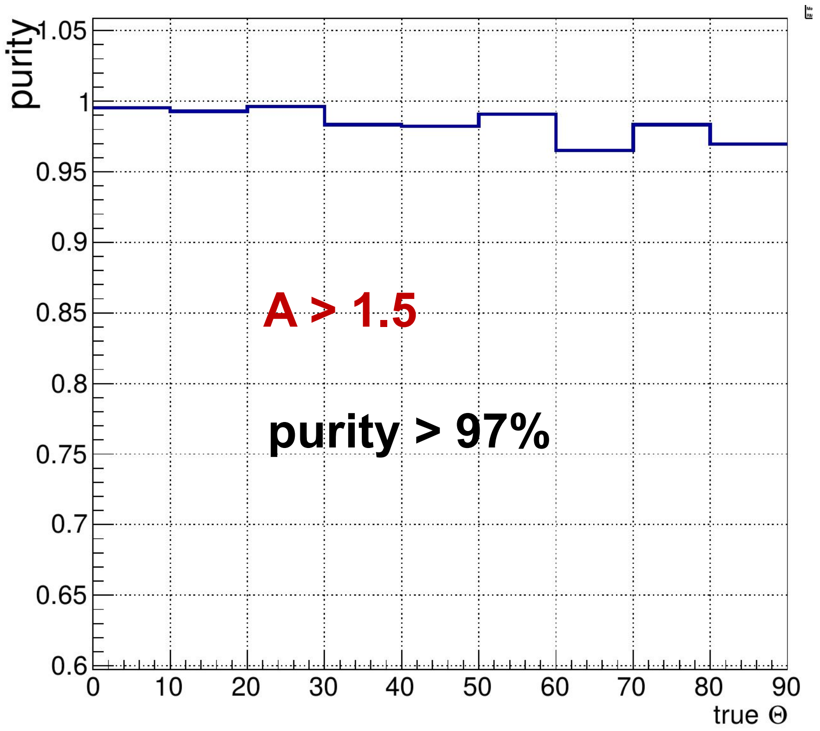
To compare with the old method we increase the threshold from **1.0** to **1.5**

6 hits at 3 strings selection: 3.5% of the sample is selected





Comparison with the old method



6 hits on 3 strings: 0.02% have survived

change in performance:

~15 times increase in event yield

worse purity: 97% -> 90%

Less stable fit -> less tracks will be passing the quality criteria

Robust fit method should be developed

Track reconstruction results

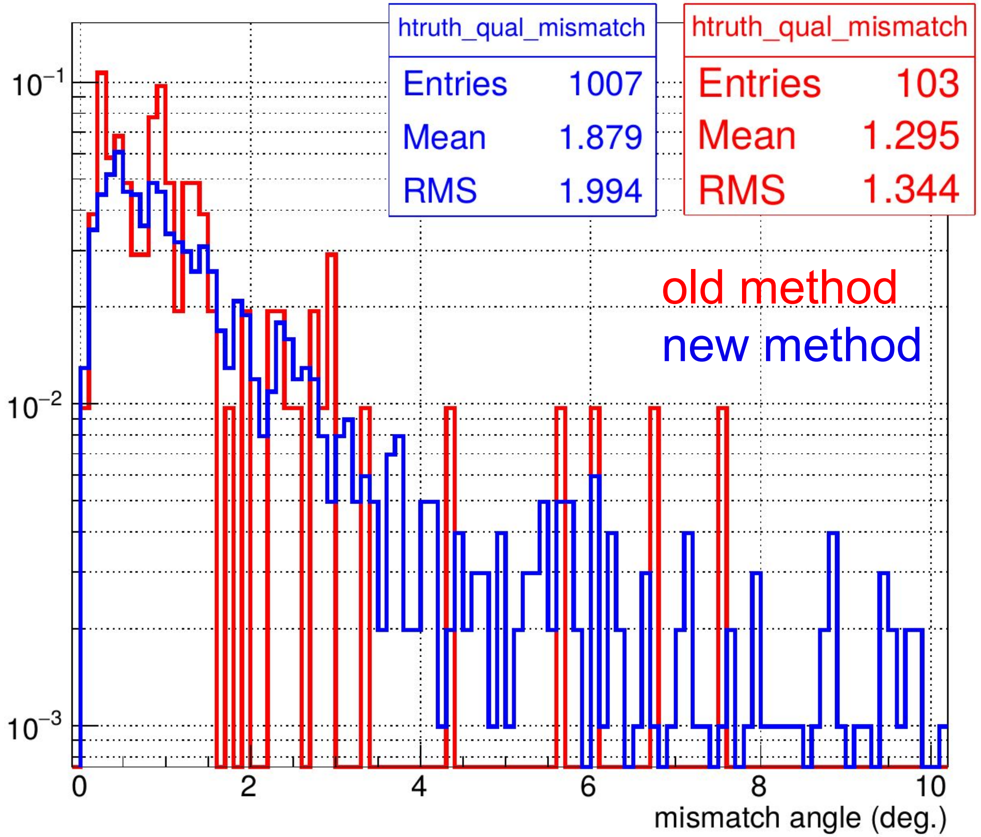
Mismatch angle after the quality cuts

New method, median: 1.3 degrees

Old method, median: 1.0 degree

10 times increase in sensitivity

Tail:
new method - 4.5% of tracks have mismatch angle > 10 degrees



These methods should be complimentary and run in parralel with different thresholds

Fit method shoud be improved - use robust fitting (not sensitive to outliers)

Another way - exclude outliers and refit (slower)

Code status

First version of the reconstruction is implemented in BARS

- Hotspot method for hit selection
- Track reconstruction
- Almost no quality parameters

Current priority: commit recent developments to BARS

- Hit selection modules:
  - New method for hit selection
- Reconstruction modules:
  - More minimisation functions
  - Functions to exclude outliers
  - Fill various quality data members in BRecParameters
  - Add references to used impulses to BRecParameters
- Monitoring package:
  - Hit selection monitor
  - Reconstruction monitor (theta, phi, quality, etc...)
- Automation of processing

Muon reconstruction action items I

- Automatic processing
  - New programs/reconstruct-muon program from Alexander
  - Should also include monitoring
  - Two settings - MC and data
- Neutrino events bank
  - Process as much data as possible
  - Select good candidates to neutrino events using cut-based analysis
- Precise vertical track reconstruction
- Understand the data-MC discrepancy
  - Polar angle distributions
  - Quality distribution, etc..
- Processing time optimisation

Muon reconstruction action items II

- Background rejection
  - An MVA-based (NN, BDT) rejection of background would enhance the sensitivity
  - Several variables with background rejection potential have been identified
- Optimisation
  - There are many parameters in hit selection and minimisation (many are hardcoded)
  - Current parameters are chosen "by eye"
  - An optimisation for range of energies and polar angles should be done to maximise the sensitivity
- Study the sensitivity to miscalibrations and misalignment
- Likelyhood fit
  - Second stage of the fit should be added - likelyhood maximisation

BACKUP

Fit with quality function

Quality function: $Q = \sum_{i=1}^{N_{hit}} \left[\frac{(t_i^{est} - t_i)^2}{\sigma_t^2} + \frac{A(a_i)D(d_i)}{d_0} \right]$ [inspired by https://arxiv.org/abs/1105.4116]

- **σ_t=3 ns** - jitter of PMT signal development
- **t_i^{est}** - estimation of arrival time in direct cerenkov light approximation
- **t_i** and **d_i** - actual time of hit and distance of OM from the track
- Amplitude function, $A(a_i) = a_0 a_i / \sqrt{a_0^2 + a_i^2}$, uses amplitude **a_i corrected for the angular sensitivity of the OM**
- Distance function, $D(d_i) = \sqrt{d_1^2 + d_i^2}$
- Chosen parameter values: **σ_t=3 ns, a_0=500 p.e., d_0 = 5 m , d_1 = 10 cm**

Minimisation is performed with MIGRAD package from ROOT

Merging of groups

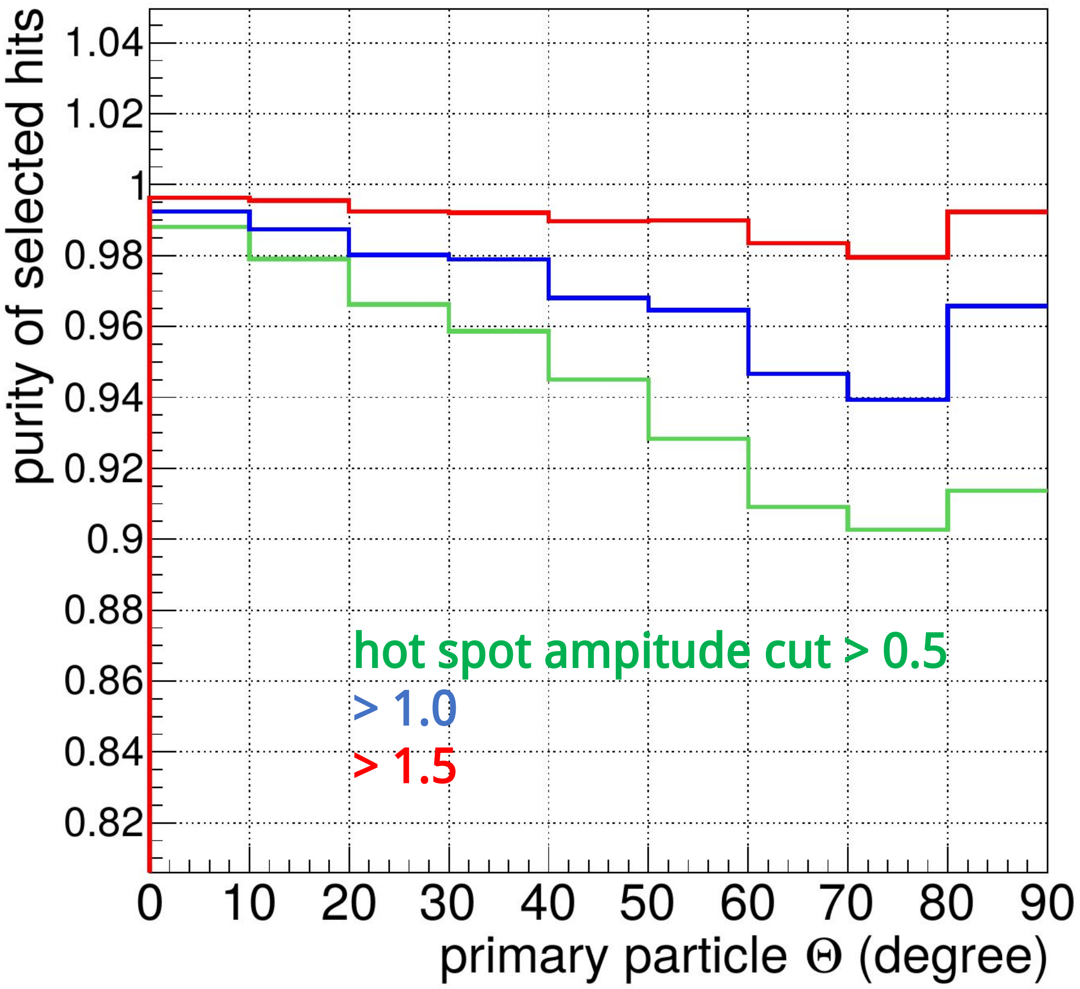
Purity of merged groups:

- Fraction of signal hits (direct or shower) in obtained global groups

Few symmetric amplitude cuts:

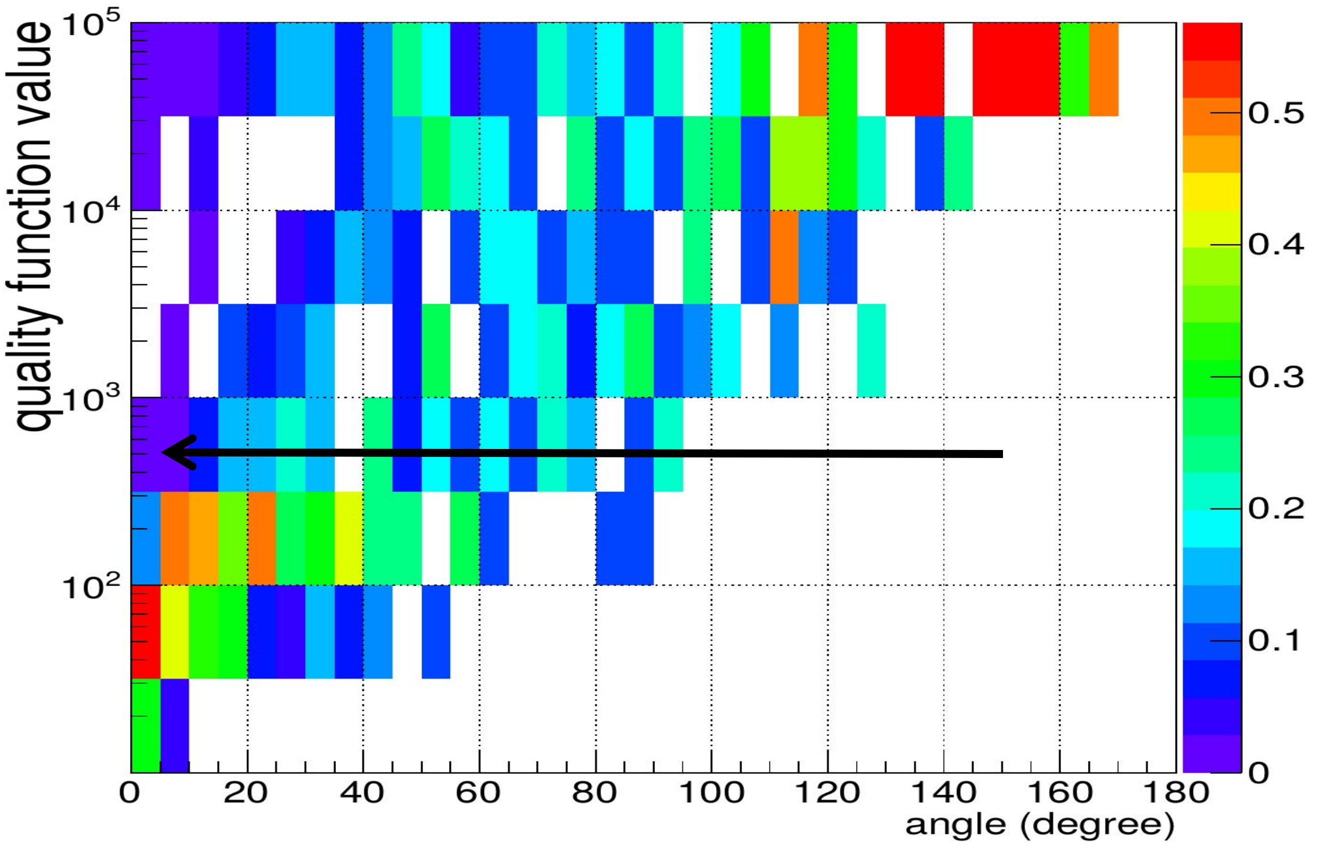
$A_{1,2} > 0.5, 1.0, 1.5$ (p.e.)

With $A_{1,2} > 1.5$ purity of > 98% is achieved



Muon reconstruction for upgoing neutrino II

Quality function demonstrates correlation with mismatch angle

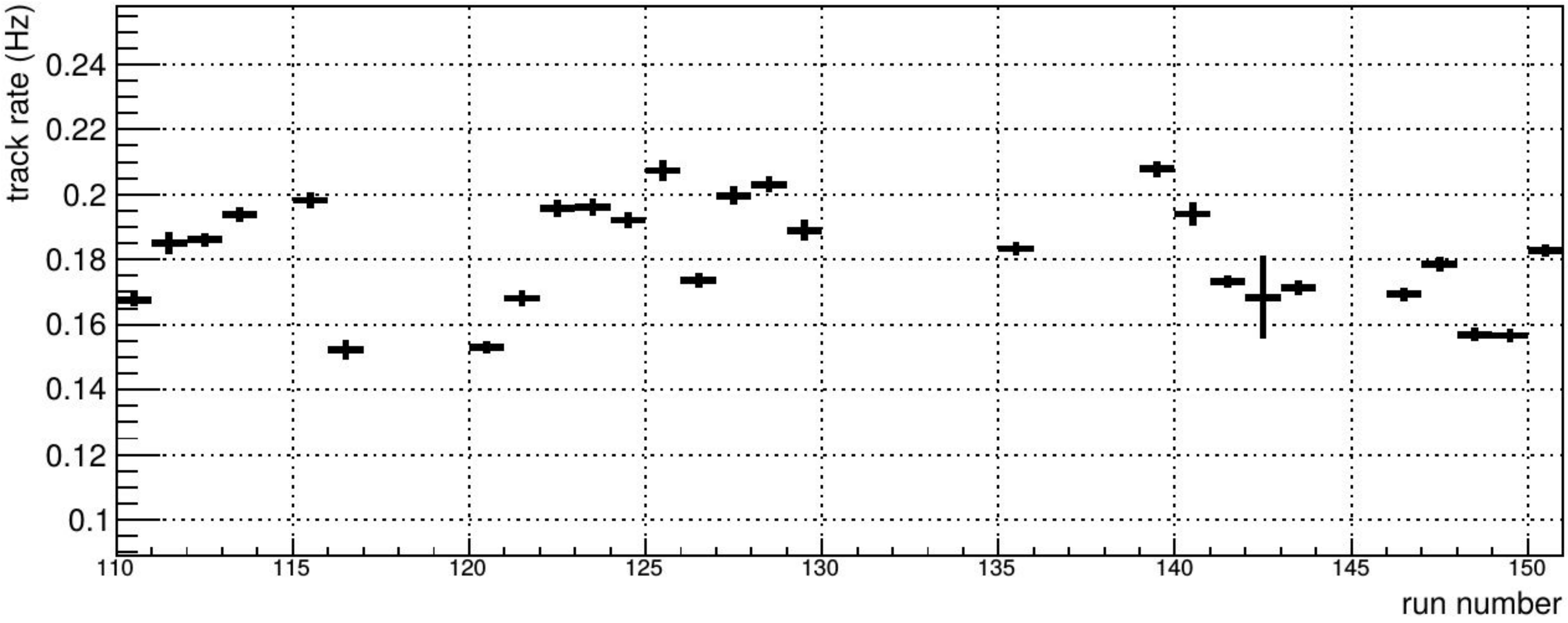


Plots shows distribution of quality function for series of mismatch angle values

Tracks of good quality are defined as: QF < 500

Results III

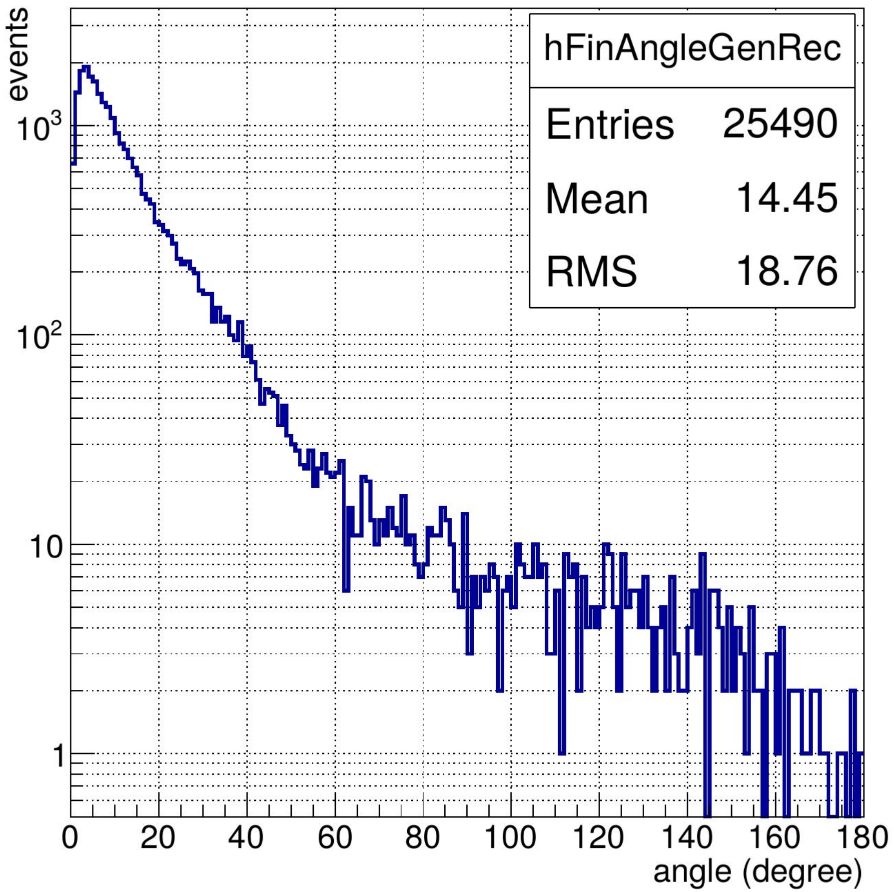
resonstructed track rate vs. run number



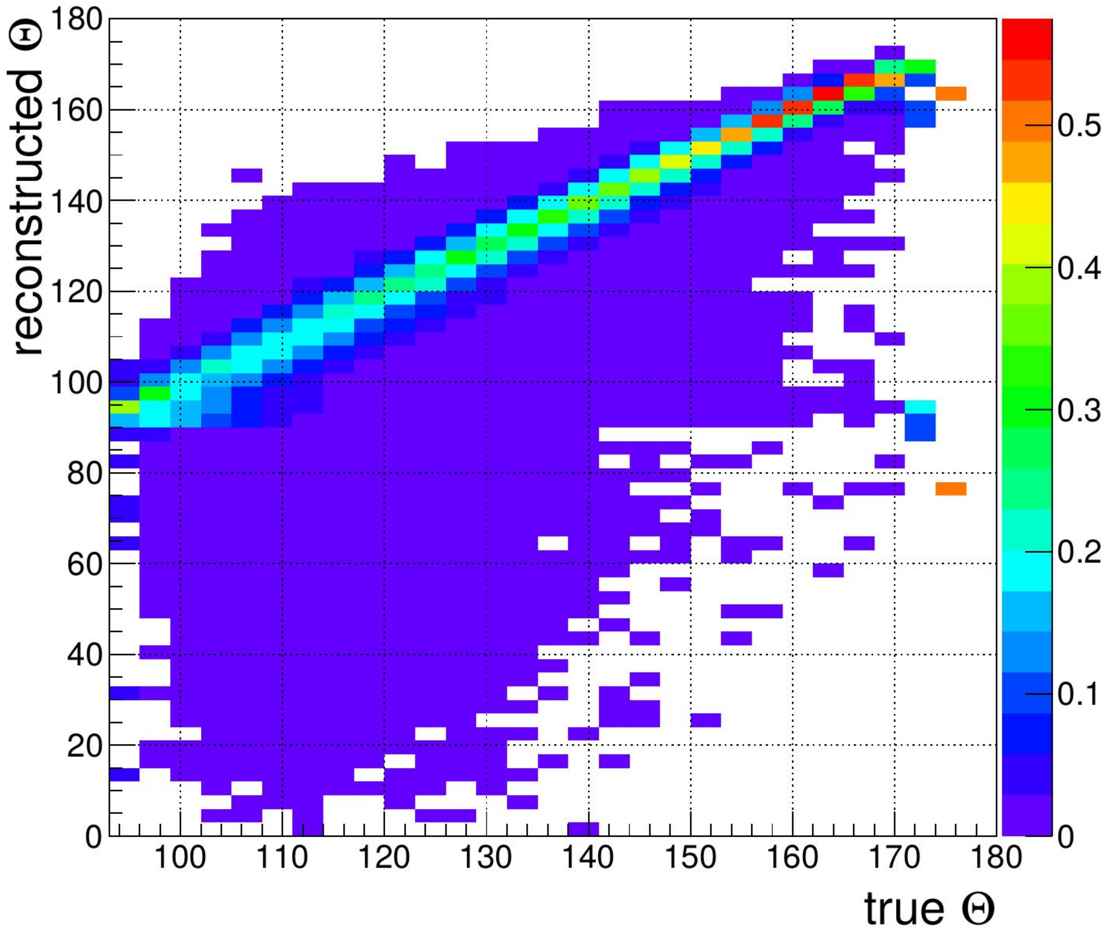
Performance of reconstruction in MC

Selection on further slides: $A_{1,2}$ > 1.5/1.5; **at least 6 hits on 3 strings**

mismatch angle



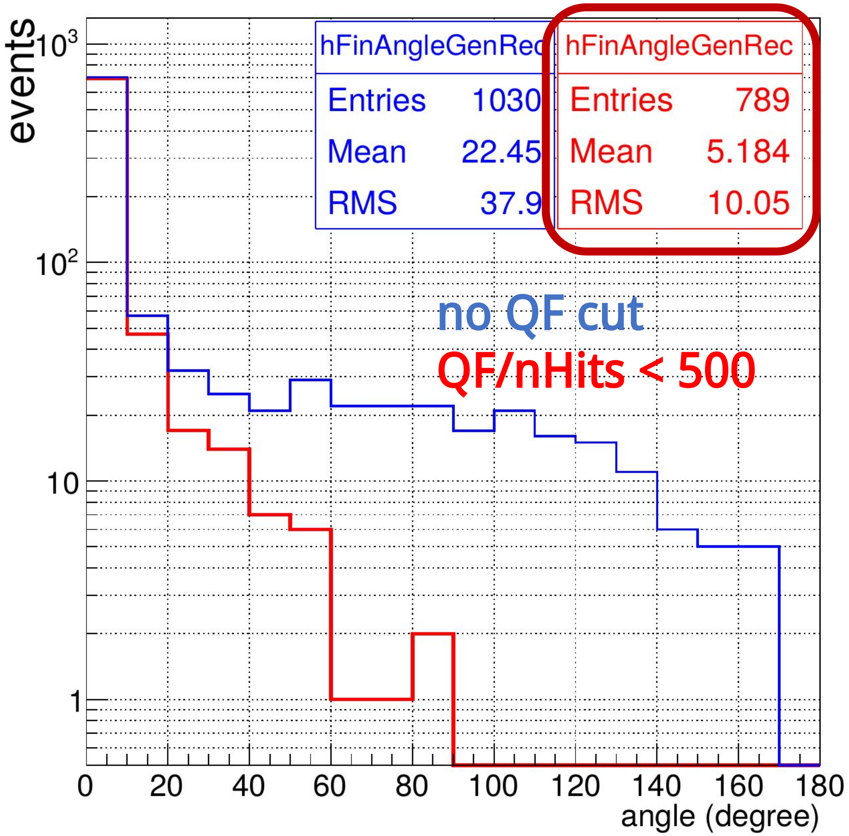
response matrix



Upgoing muons : applying quality function cut

Event selection: $A_{1,2}$ > 1.5 p.e, t_s = 30ns, **at least 6 hits on 3 strings**

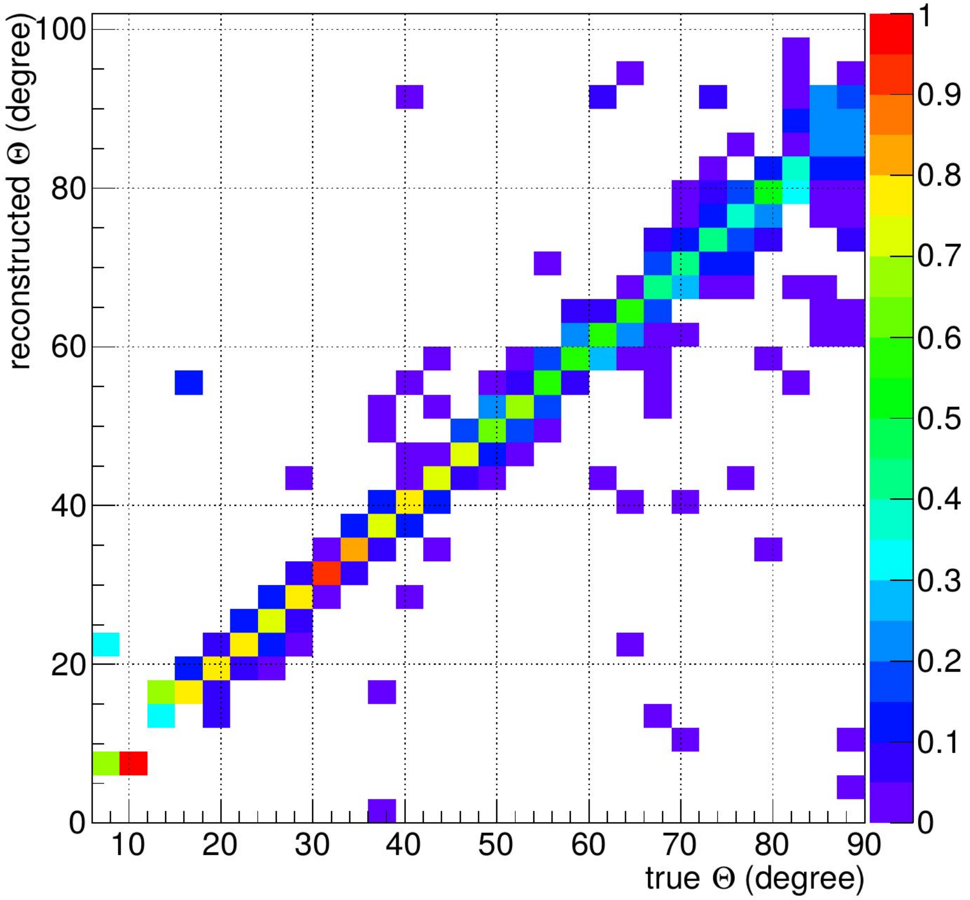
mismatch angle



median without quality cut: **2.9**

median with quality cut: **1.8**

response matrix



Muon track reconstruction

Track parametrisation: $\vec{R(t)} = \vec{R_0} + c(t - t_0)\vec{V}$

$\vec{V}$ is unit vector in polar coordinates

Initial approximation: $\vec{V_{init}} = (1/|\sum_{t_i<t_j} \vec{R_{ij}}|) \sum_{t_i<t_j} \vec{R_{ij}}$ where *i* and *j* are ordered in time and belong to different strings

Initial values for $\vec{R_0}$ and t_0 are obtained using first (in time) hit

Minimisation function:

$$Q = \sum_{i=1}^{N_{hit}} \left[\frac{(t_i^{est} - t_i)^2}{\sigma_t^2} + \frac{A(a_i)D(d_i)}{d_0} \right]$$

inspired by ANTARES paper: https://arxiv.org/abs/1105.4116

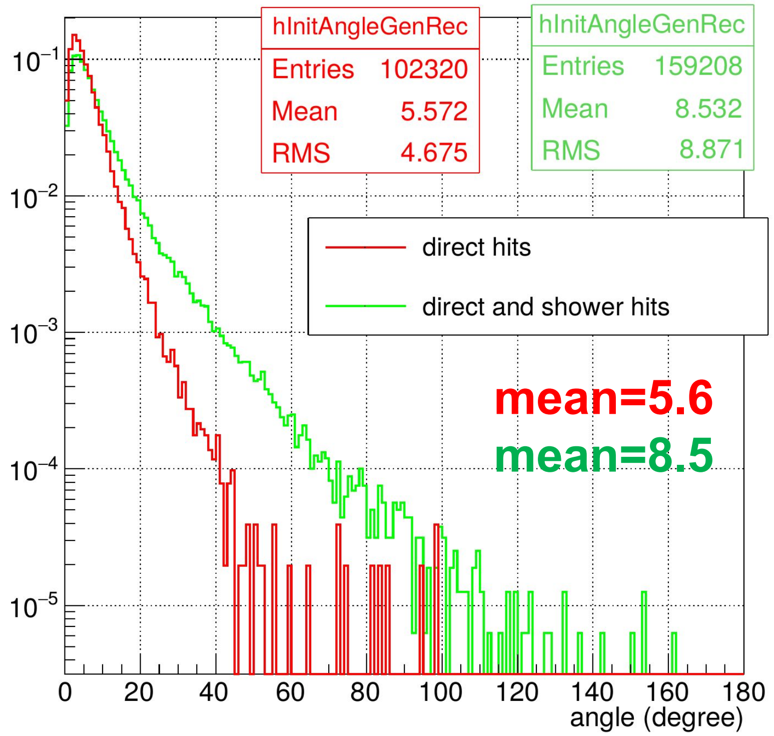
Direct cerenkov approximation for the time estimation
A and *D* are amplitude and distance functions

Performance with MC truth

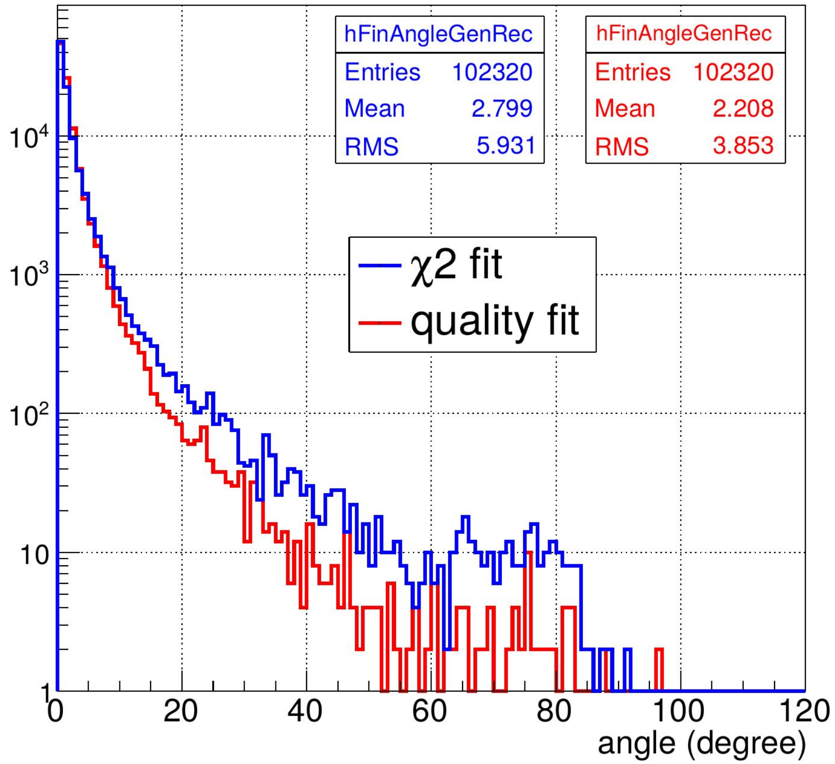
Events with at least 6 hits on 3 strings are selected (here and further in this talk)

MC truth -based selection of direct cerenkov hits and shower hits

mismatch angle for initial approximation



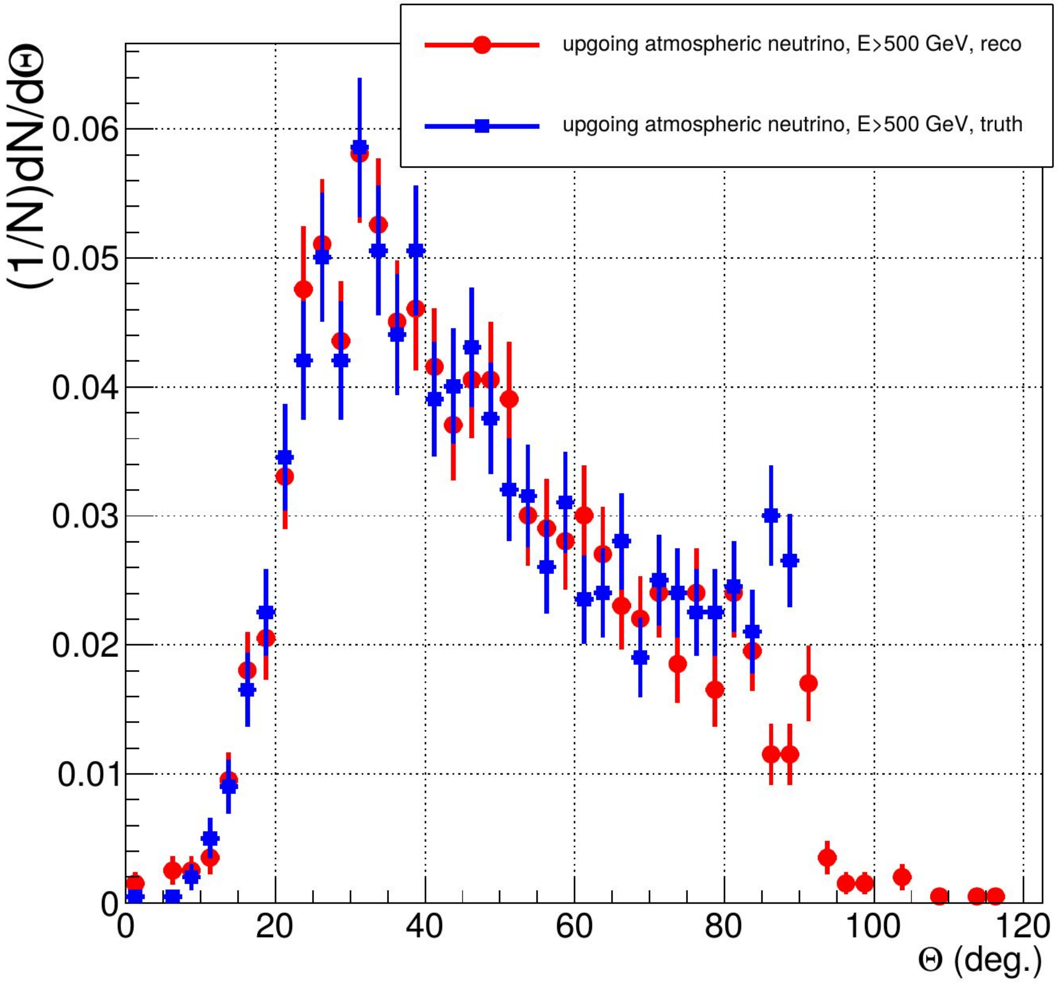
fit results for direct hits



median angle for quality fit: 1.3

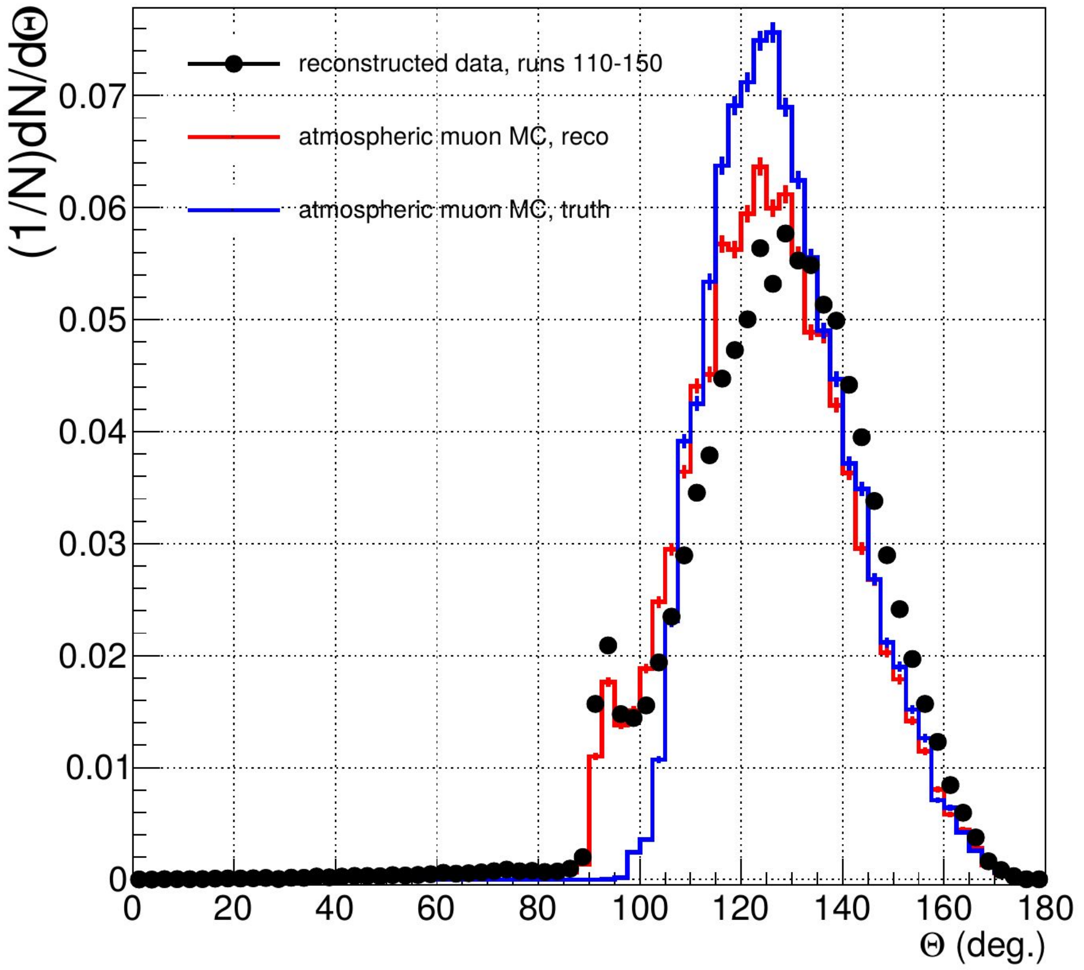
Reconstructed polar angle

MC reconstructed (red)
and
true (blue)
polar angle

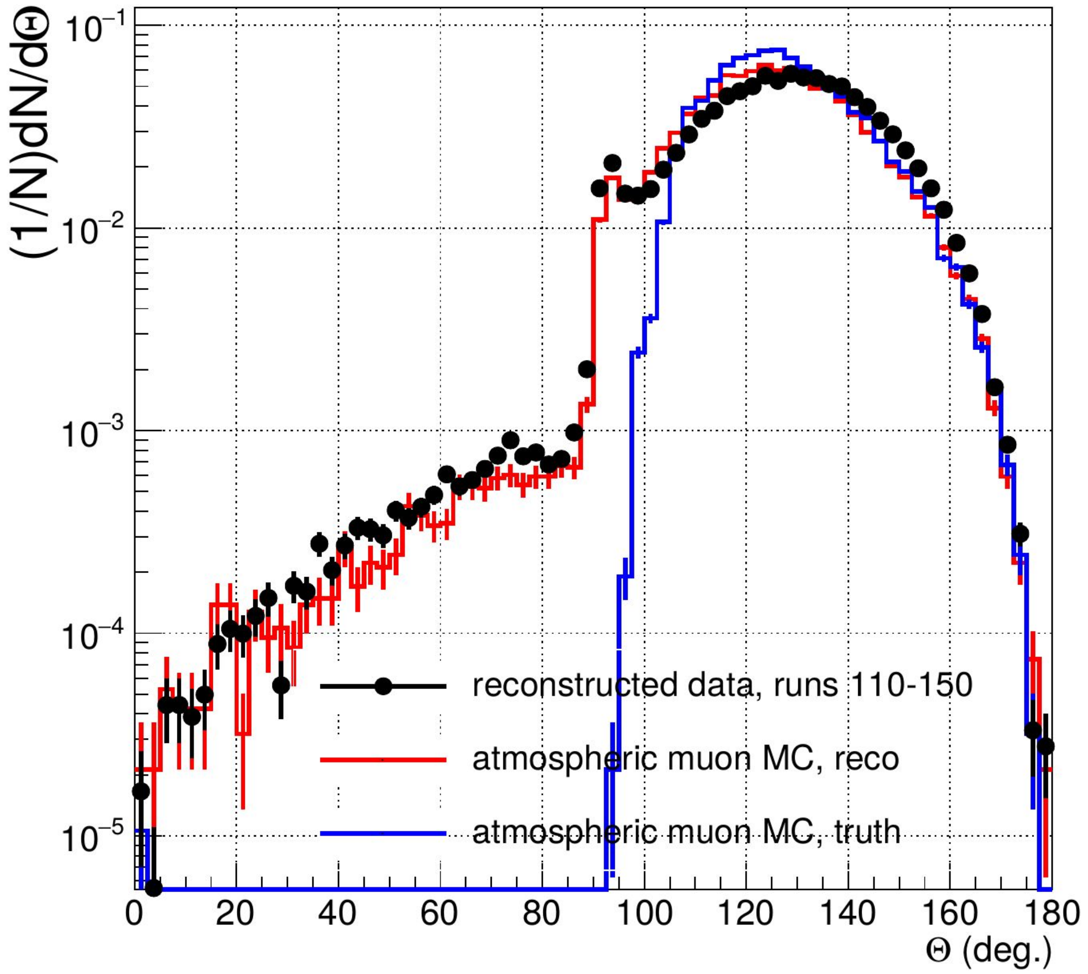


Results I

normal scale

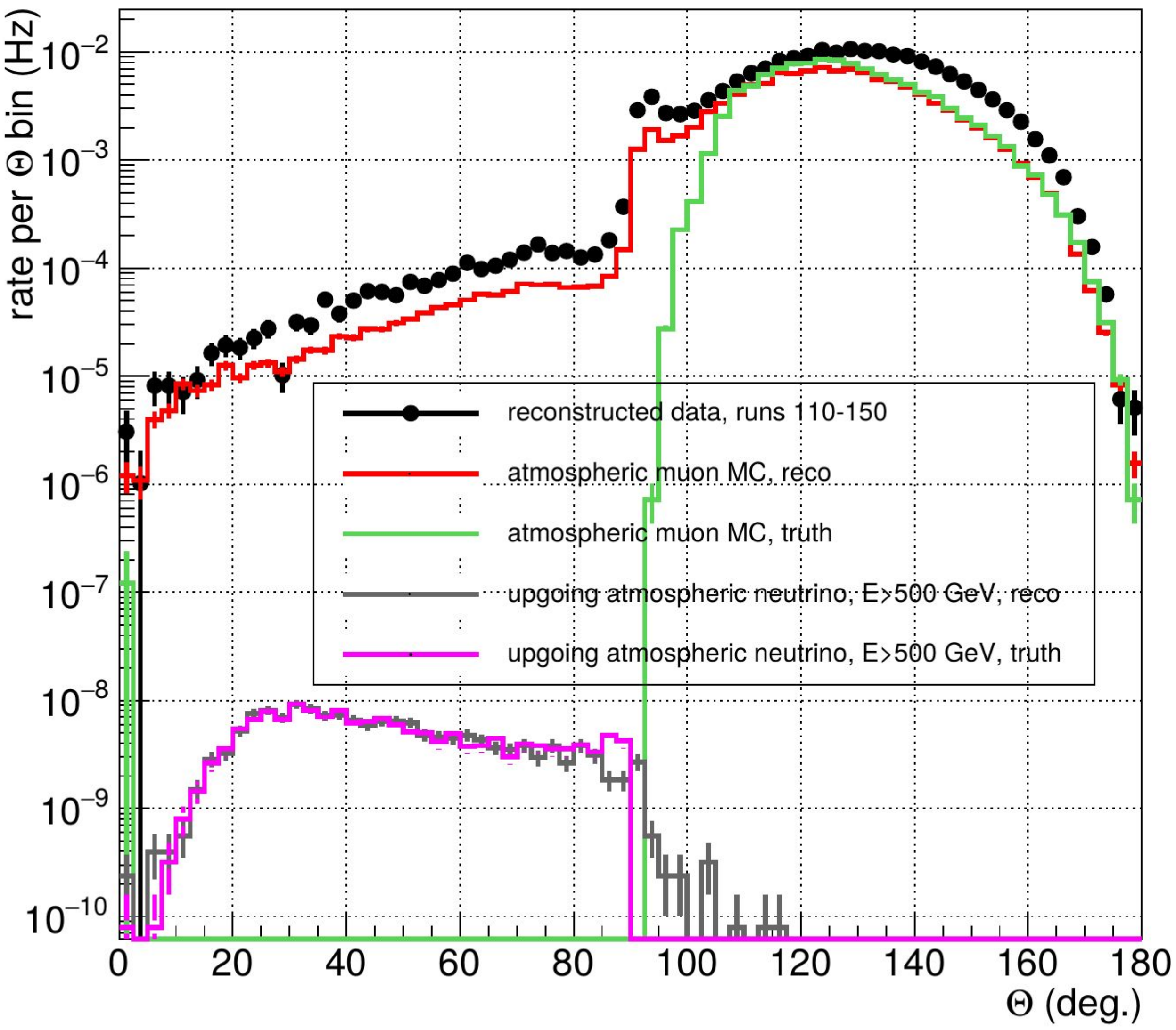


log scale



Fraction of misreconstructed as upgoing tracks is ~ 1.5% both for data and MC

Results IV



Polar angle distributions scaled by the exposition time

Data rate is ~1.5 times larger than MC

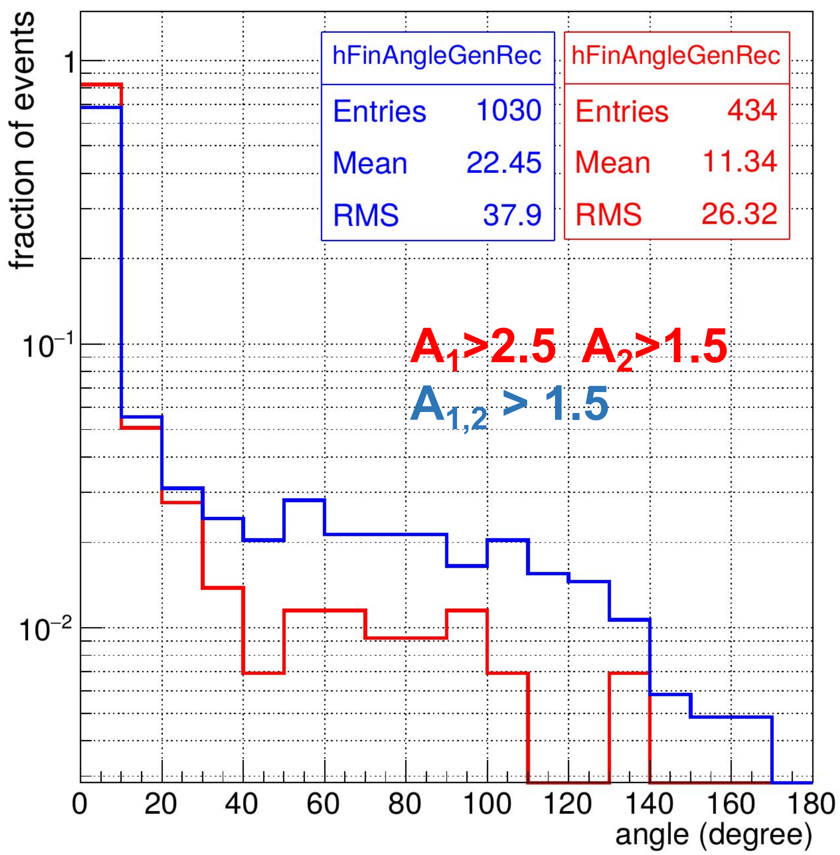
Upgoing atmospheric neutrino rate is 10^3 times smaller than current background from misreconstructed muons

More advanced background supression methods are to be developed

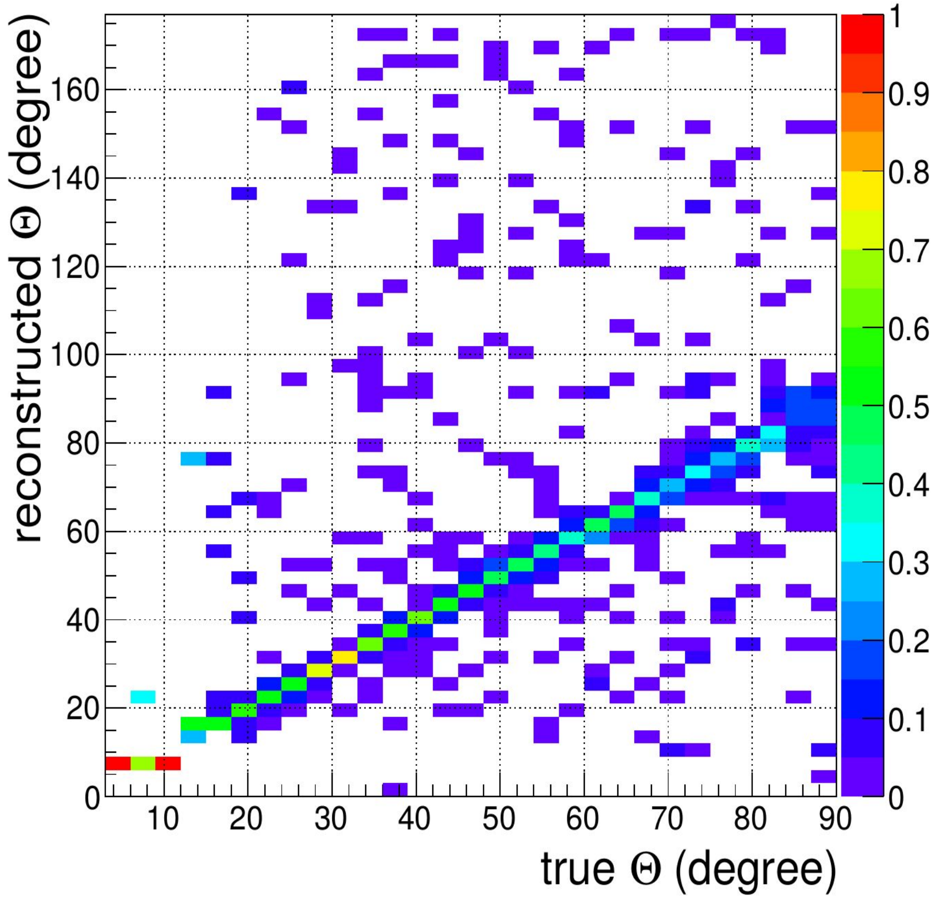
Muon reconstruction for upgoing neutrino

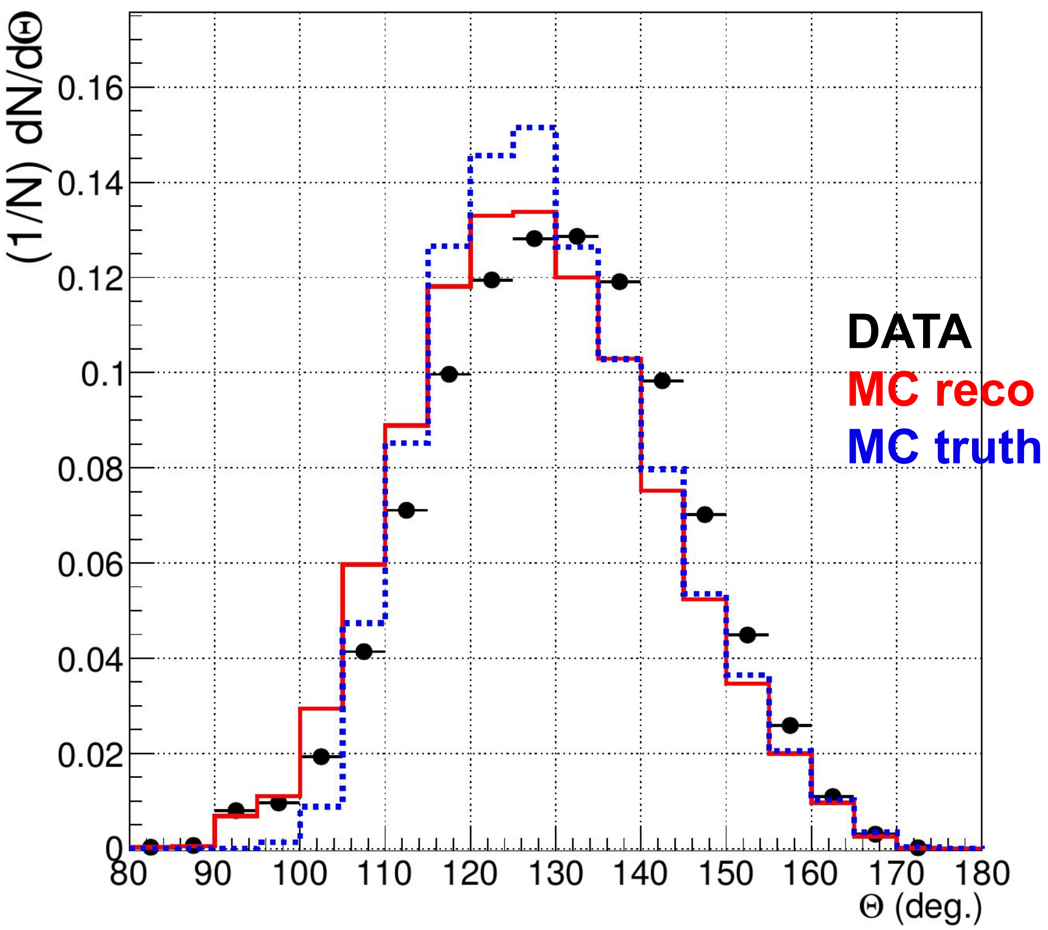
Event selection: $A_{1,2}$ > 1.5 p.e, t_s = 30ns, **at least 6 hits on 3 strings**

mismatch angle

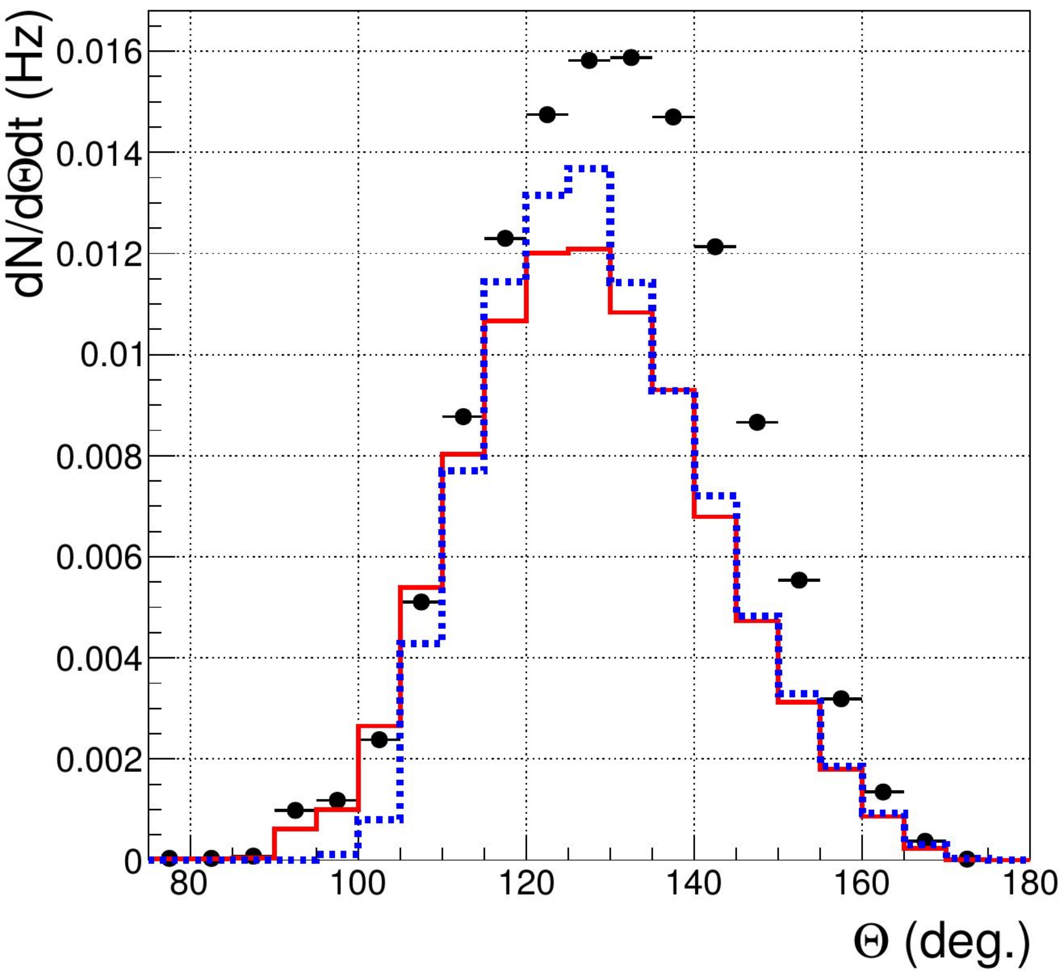


polar angle response matrix





(1/N) dN/dΘ
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
80
90
100
110
120
130
140
150
160
170
180
Θ (deg.)
DATA
MC reco
MC truth



dN/dΘdt (Hz)
0.016
0.014
0.012
0.01
0.008
0.006
0.004
0.002
0
80
100
120
140
160
180
Θ (deg.)