Flavor identification in ORCA

Thomas, Heid

MANTS Garching, 14th-15th October 2013

ERLANGEN CENTRE
FOR ASTROPARTICLE
PHYSICS

FRIEDRICH-ALEXANDER
UNIVERSITÄT
ERLANGEN-NÜRNBERG

ERLANGEN CENTRE FOR ASTROPARTICLE PHYSICS

Outline

- Motivation
- Methods
  - Random decision forests
  - Used features (Examples)
- Results
- Conclusion

Motivation

- Best signal for ORCA in muon neutrinos

Track-like events ⟷ Shower-like events

ν_μ (CC)

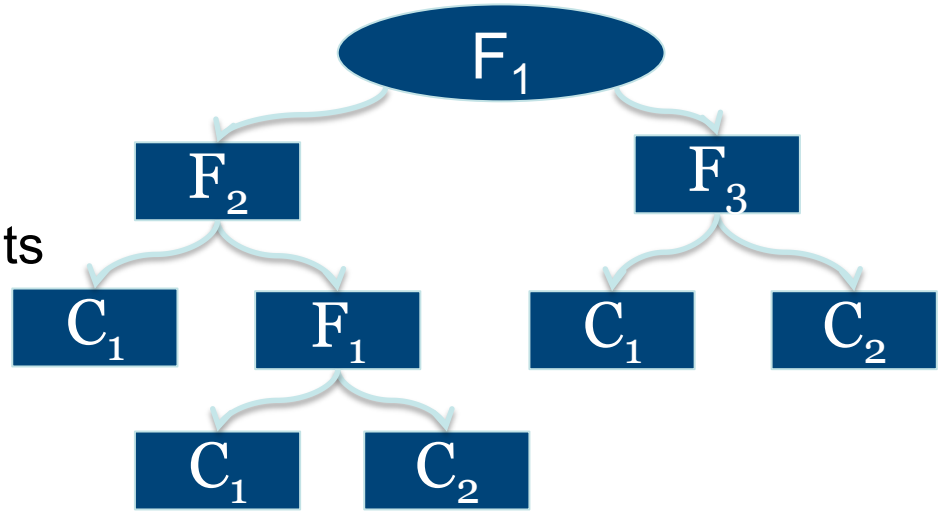
ν_μ (NC)

ν_e (CC)

ν_e (NC)

Random decision forests

- Random Decision Tree
  - Cut on feature F_x
  - Classify as class C_x
- Trained with Monte Carlo Events



Many different Trees ⟶ Random decision forest

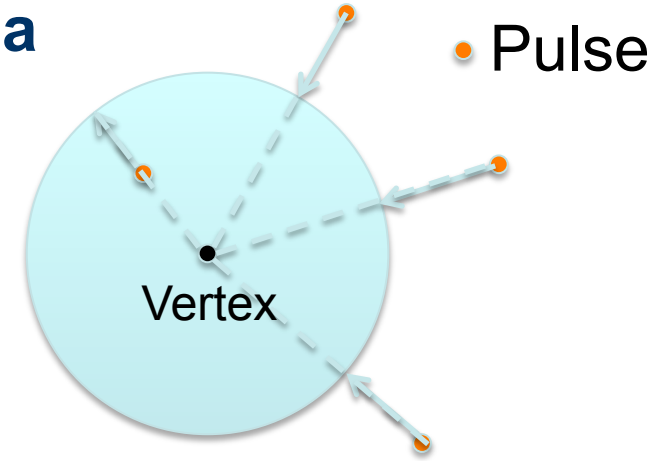
- Define a majority for the decision

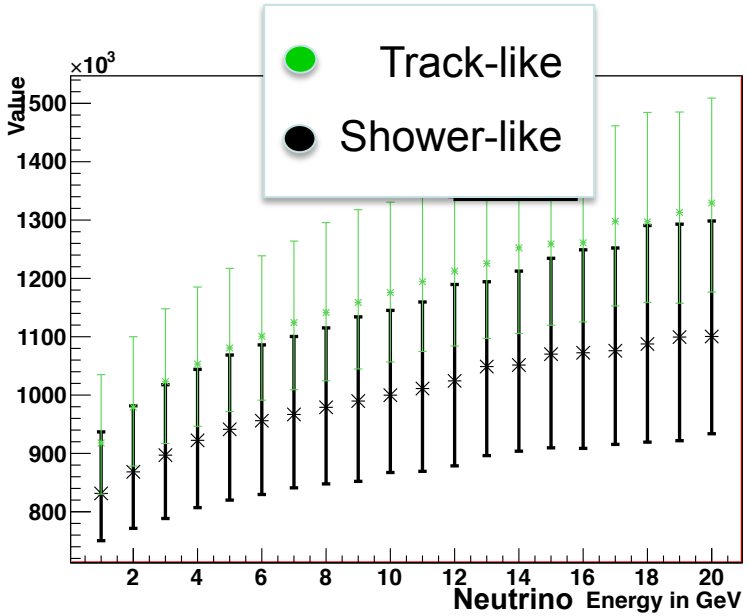
Used Monte Carlo Events

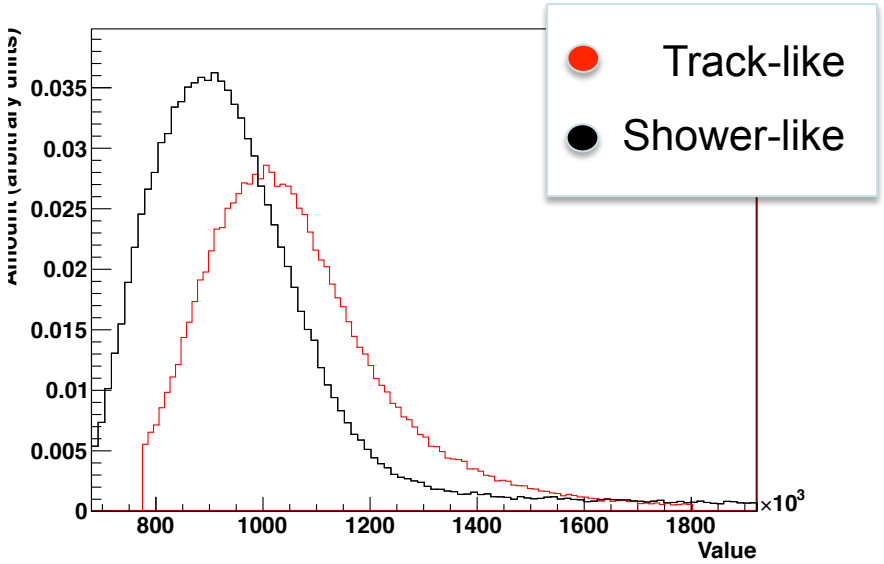
- Physical Events
  - Catania MC Production
  - Genie
  - Event selection:
    - Minimum 5 MC Hits
    - Vertex in detector
      $|z| < 50$; $r<50$
  - No noise
  - Around 10000 Events
- Premium Events
  - Catania MC Production
  - Genie
  - Nearly all light is contained
  - km3v4r5
  - Geasimv4r12
  - Modk40seawiet
  - 300000 Events

Features (Example): Tensor Of Inertia

- Eigenvalues of tensor of inertia
- Projection of all pulses to sphere around reconstructed vertex



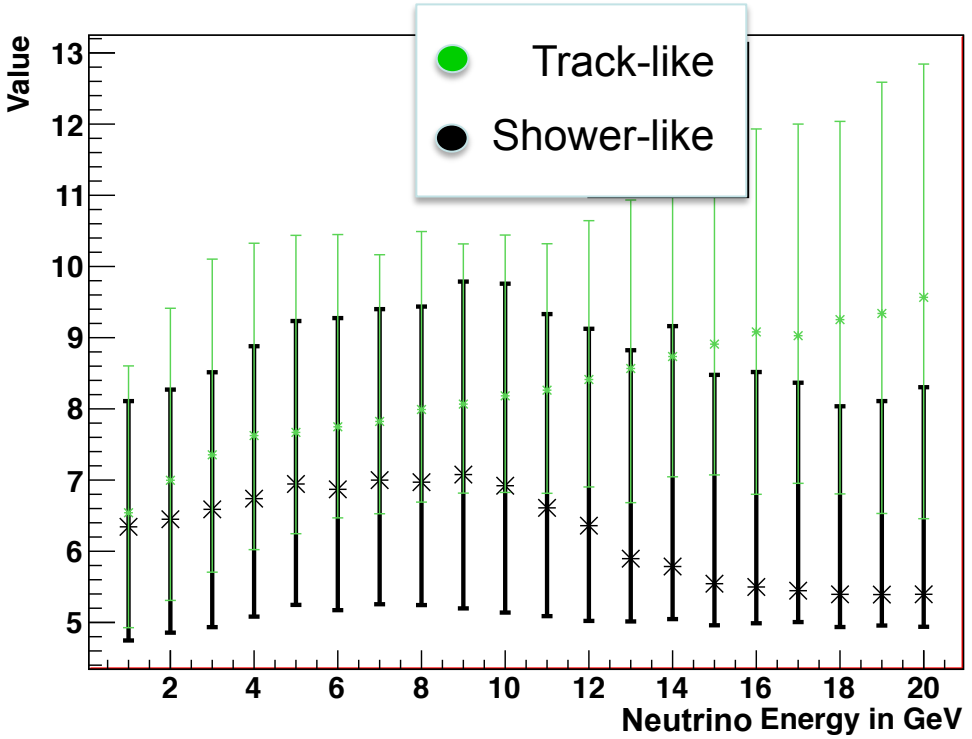




Features (Selection): QStrategy Quality Parameter

- Qstrategy shower reconstruction algorithm (ANTARES-PHYS-2012-11)
  - Based on time residuals under shower hypothesis

$$M = \frac{1}{N}\sum_{i}^{N} 2 * \sqrt{1 + \frac{(\Delta T)^2}{2 * \sigma^2}} + 2$$



ERLANGEN CENTRE FOR ASTROPARTICLE PHYSICS

Best of features

Fraction of number of hits in shower selection and track selection
QStrategy quality parameter
RMS of time residuals
FWHM of time differences
Number of Hits
Mean distance of hit from reconstructed vertex
Smallest eigenvalue of tensor of inertia
FilteringFit rlogl
…

Results

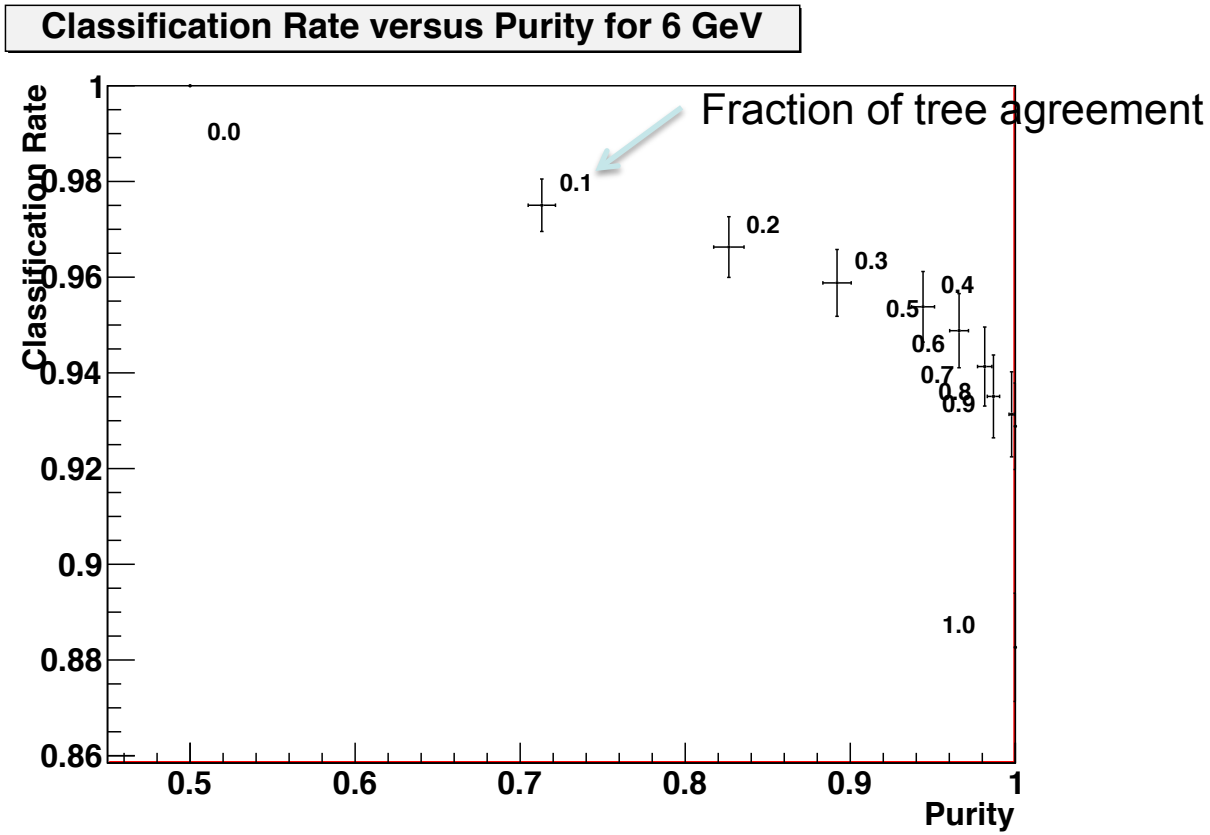
Definitions for Evaluation

$$Purity = \frac{\text{N(correctly classified track-like events)}}{\text{N(events classified as track-like)}}$$

$$Classification\ rate = \frac{\text{N(correctly classified track-like events)}}{\text{N(track-like events)}}$$

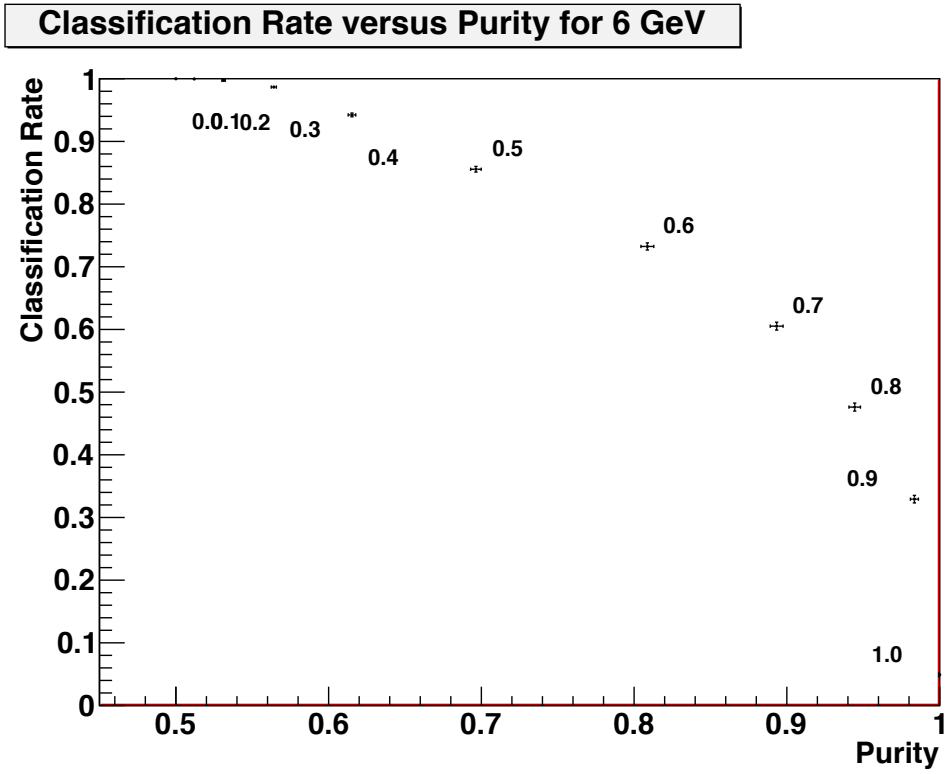
Physical Events without noise

- Event selection:
  - Minimum 5 MC Hits
  - Vertex in detector
    |z| < 50; r<50
- No noise



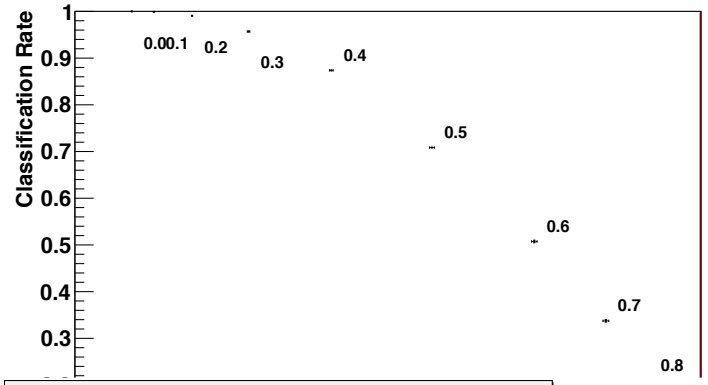
Premium Events with noise

- Nearly all light is contained

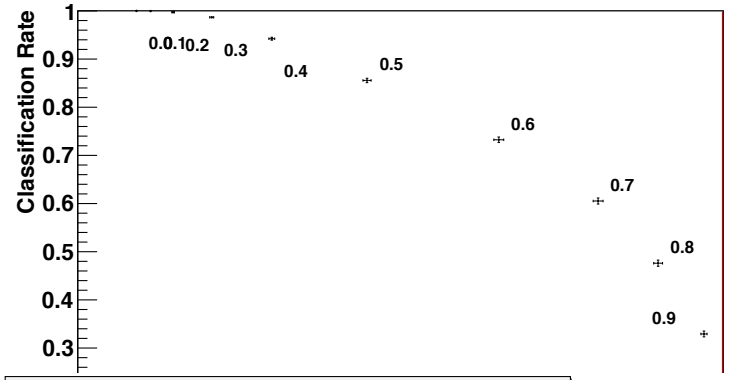


Premium Events with noise

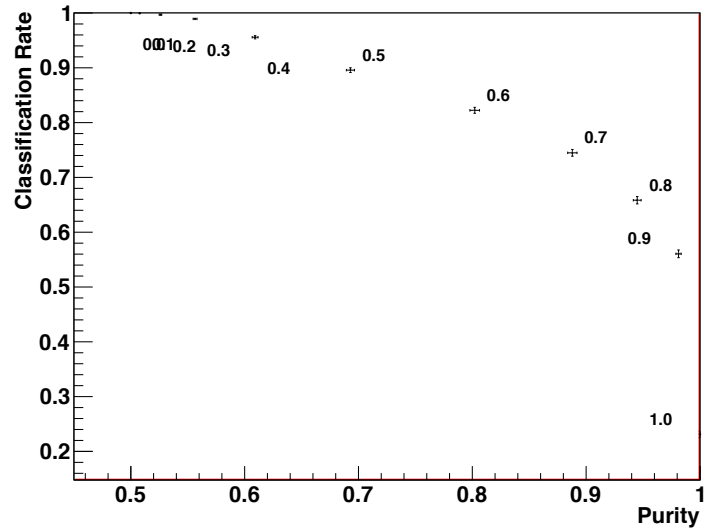
Classification Rate versus Purity for 3 GeV



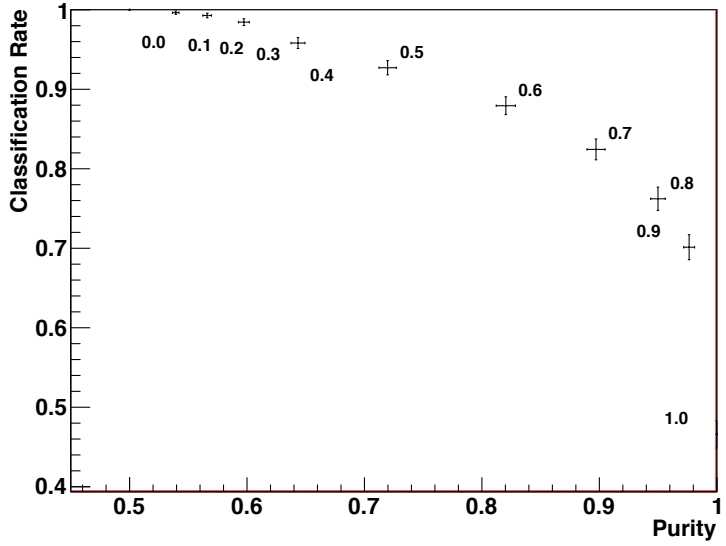
Classification Rate versus Purity for 6 GeV



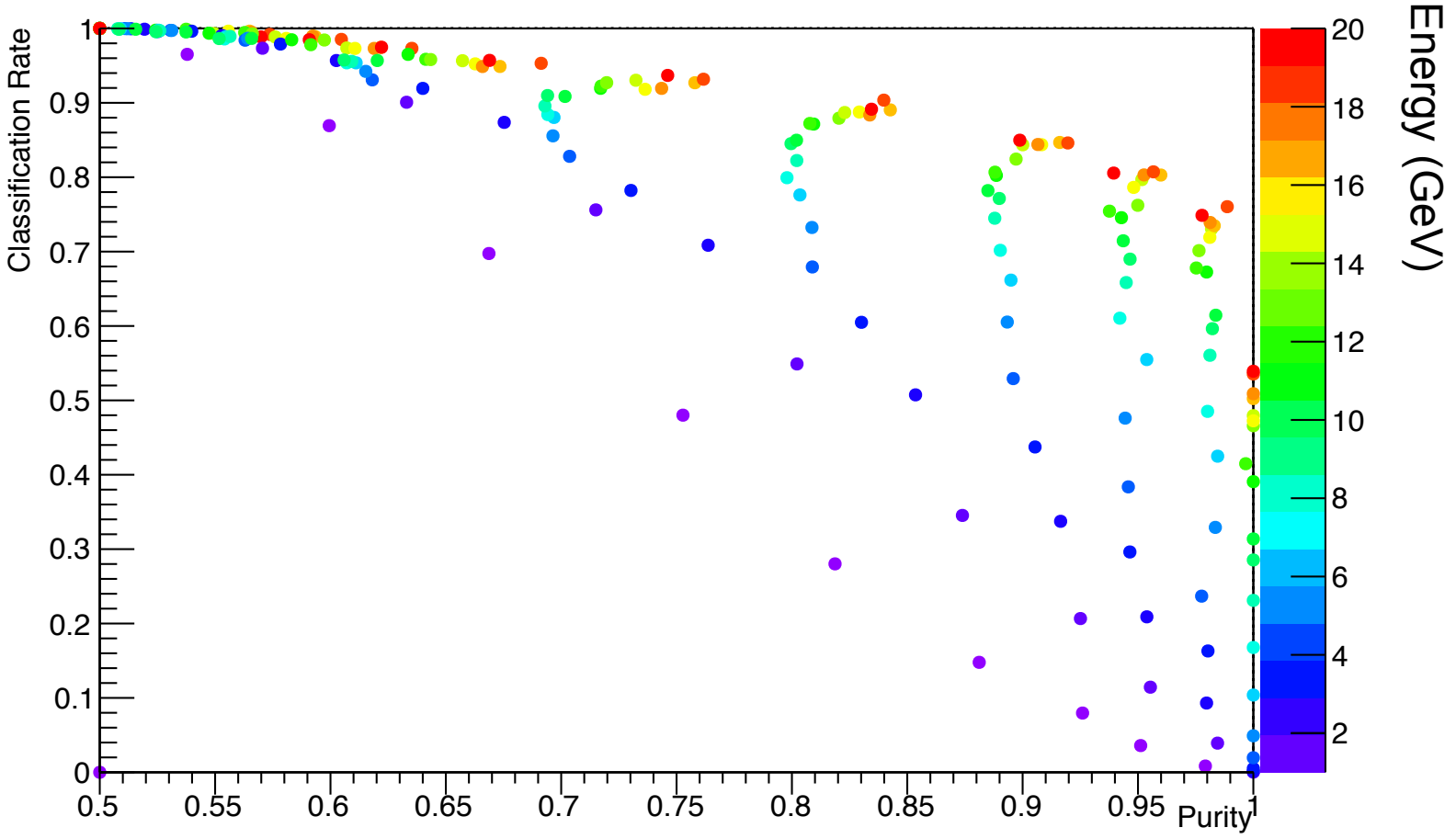
Classification Rate versus Purity for 9 GeV



Classification Rate versus Purity for 14 GeV



Premium Events with noise



Summary and Conclusion

- Random Decision Forest with about 30 Features
  - Properties of distribution of time residual
  - Tensor of inertia
  - Quality parameter of reconstruction
  - …
- It is possible to identify track-like events
  - 80% Classification rate with over 80% purity rate
  - Without noise 95% classification rate with over 90% purity

Thanks for your attention