ALICE
Statistical Wish-List

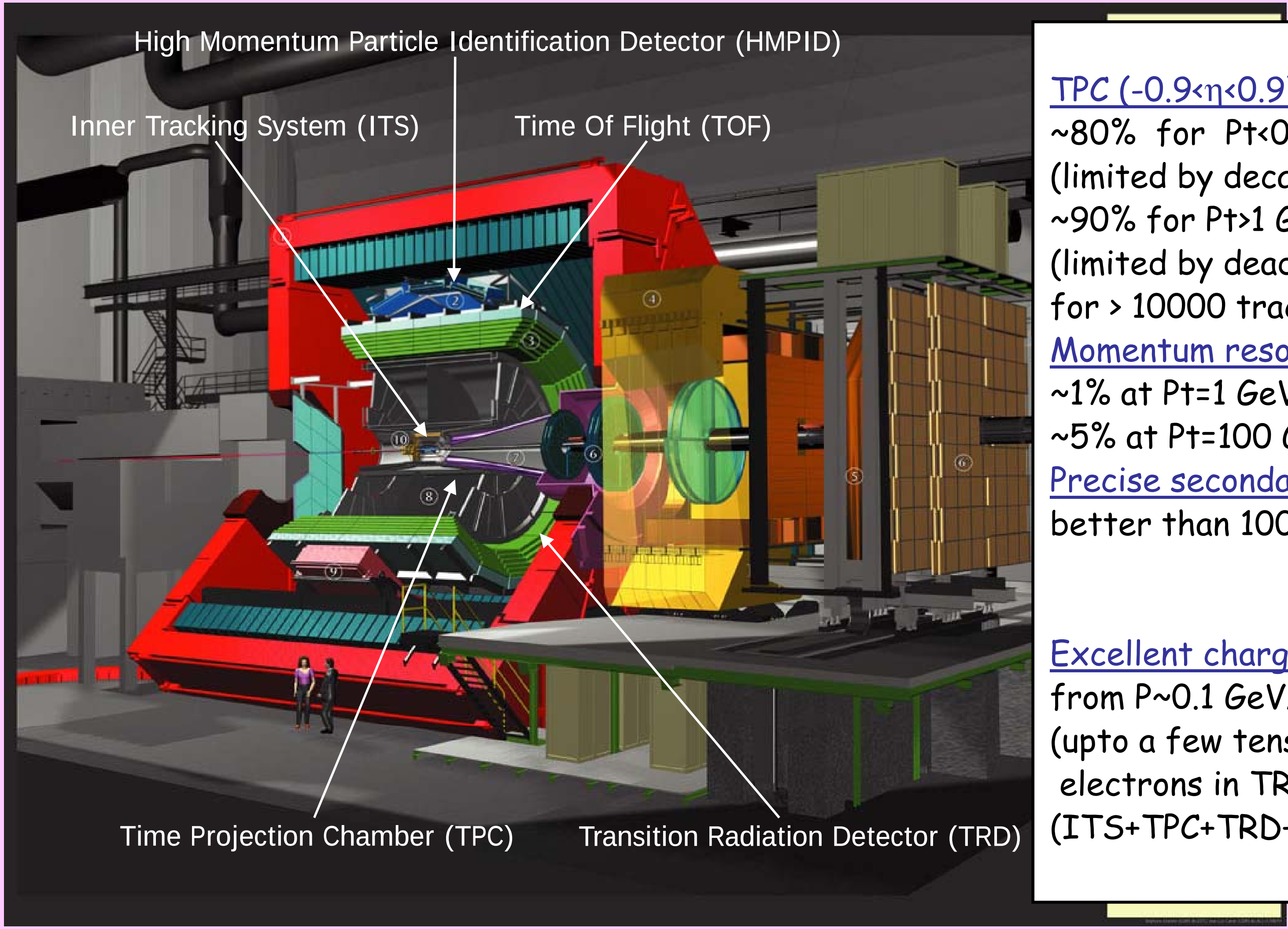
I. Belikov, on behalf of the ALICE collaboration

PHYSTAT2007, CERN Geneva, 27-29 June 2007

Outline

- ALICE experiment at LHC.
- Statistical problems with track finding in ITS.
  - Ad-hoc extensions of the Kalman filter.
- Statistical problems with particle identification.
  - Contribution of the prior probabilities to the Bayesian decision.
  - Ad-hoc treatment of mismatching (PID mismeasurements).
- The wish-list.

ALICE experiment at LHC



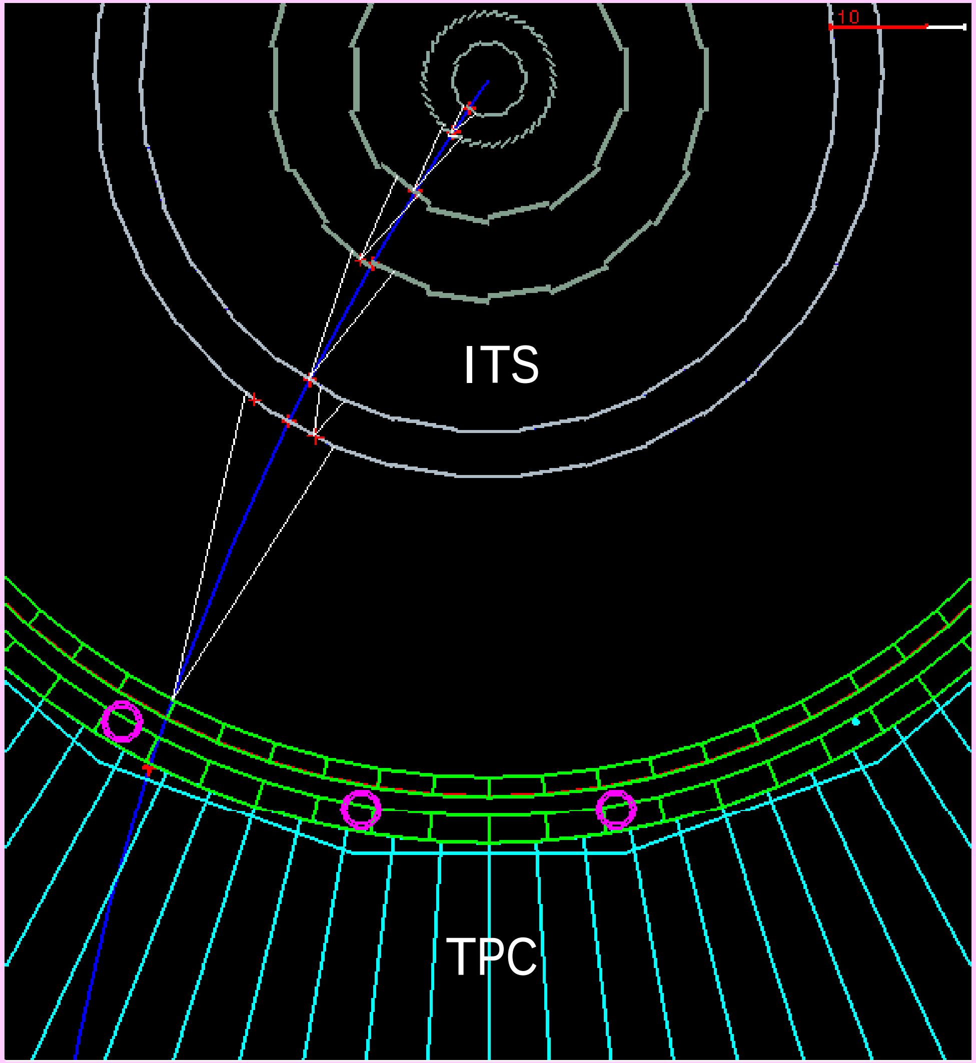
TPC ($-0.9<\eta<0.9$) tracking efficiency:
~80% for Pt<0.2 GeV/c
(limited by decays),
~90% for Pt>1 GeV/c
(limited by dead zones),
for > 10000 tracks in the TPC.

Momentum resolution (B=0.5 T):
~1% at Pt=1 GeV/c,
~5% at Pt=100 GeV/c (ITS+TPC+…).

Precise secondary vertexing
better than 100 μm (ITS).

Excellent charged PID capability:
from P~0.1 GeV/c upto a few GeV/c,
(upto a few tens GeV/c, TPC rel. rise),
electrons in TRD, P>1GeV/c
(ITS+TPC+TRD+TOF+HMPID+…).

Statistical problems with track finding in ITS

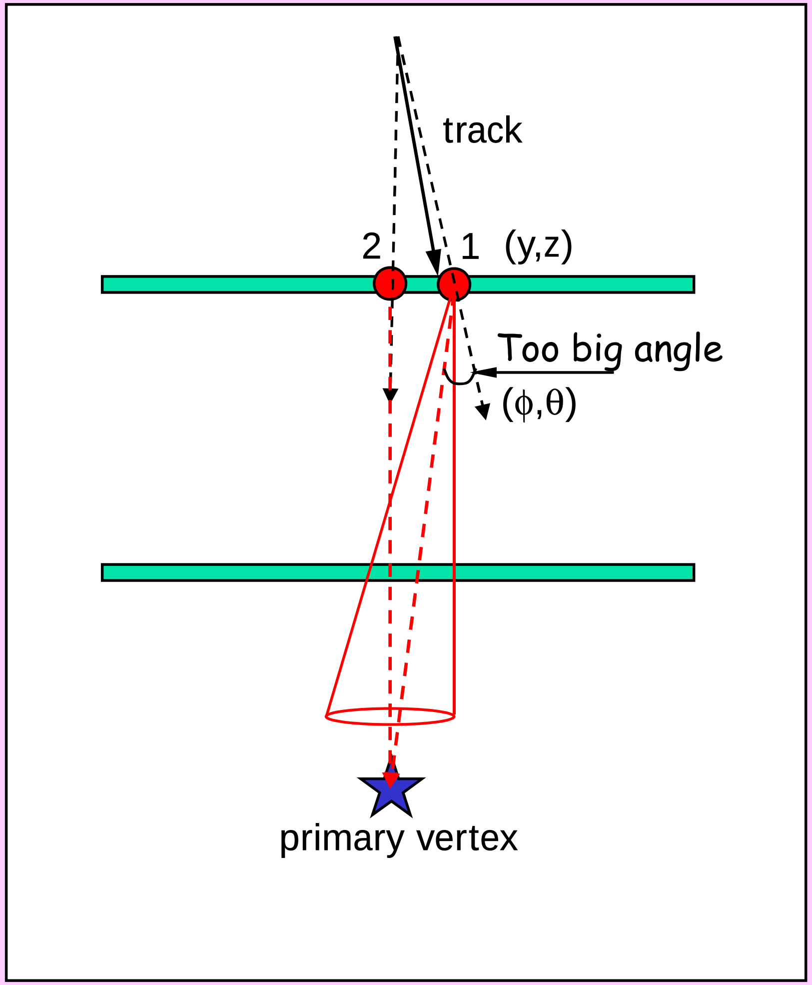


Several clusters within the "road" defined by multiple scattering…

Ad-hoc solutions:

- Investigation of the whole tree of possible prolongations.
- Applying a "vertex constraint" (1st pass).

The "vertex constraint"



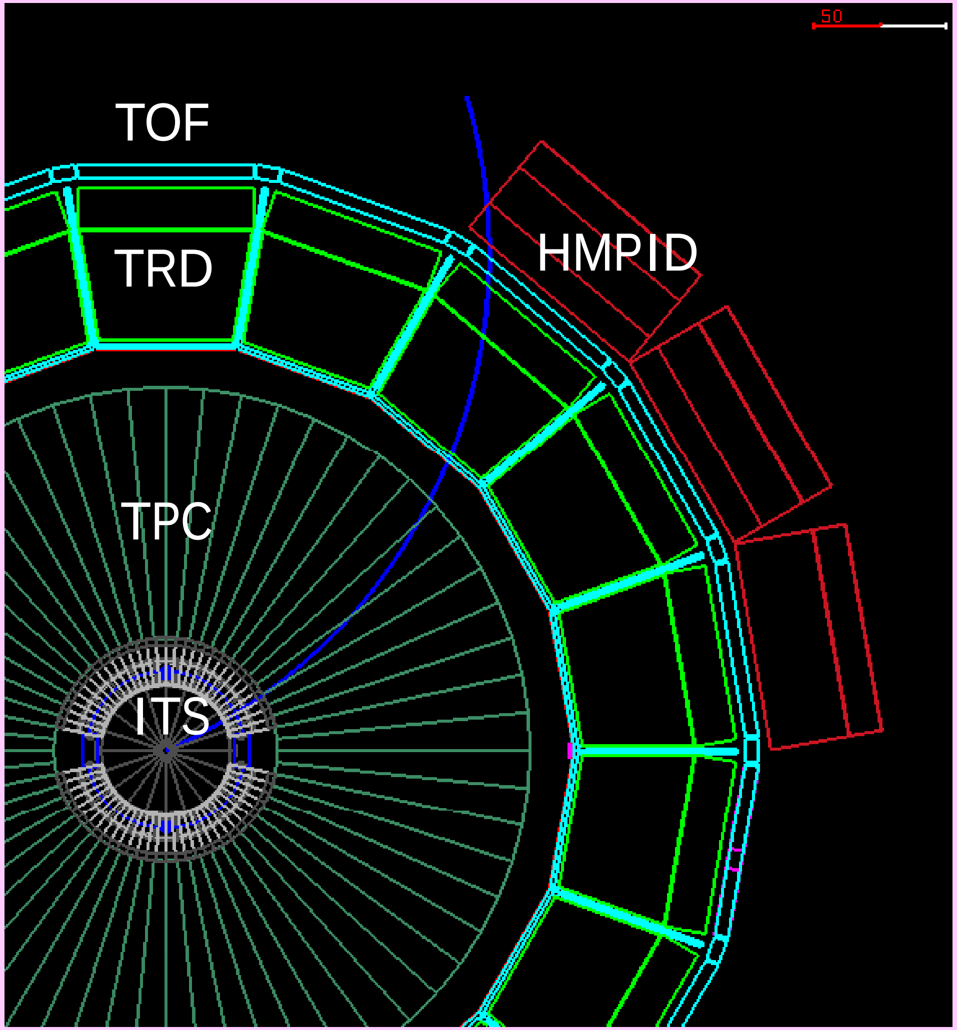
Looking at the cluster position only, the cluster #1 is "better".

But, if also taking into account the direction towards the primary vertex, the cluster #2 becomes more preferable…

Technically, this is done by extending the "measurement" $(y,z) \to (y,z,\phi,\theta)$

<u>A question:</u>
Can all this be justified ? Improved ?
(especially, if applied the same track repeatedly ?)

Particle identification in ALICE



The same track can simultaneously be registered by 5 detectors that

- have quite different PID response,
- are efficient in complementary momentum ranges.

Clearly, the final PID decision depends

- not only on the measurements by the detectors,
- but also on the particle production ratios and/or track selection (ex: particle spectra, Λ reconstruction)

Bayesian approach in PID

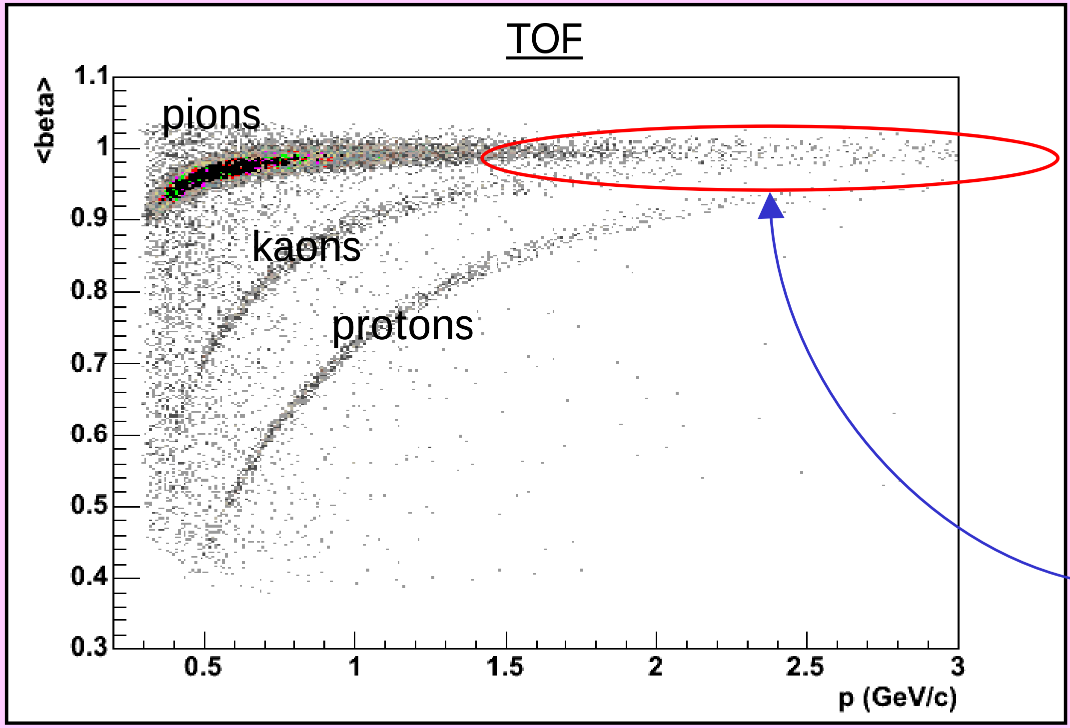
Probability to be a particle of i-type ($i = e, \mu, \pi, K, p, \ldots$), if the PID signal in the detector is s:

$$w(i\,|\,s) = \frac{C_i r(s|i)}{\sum\limits_{k=e,\mu,\pi,\ldots} C_k r(s\,|\,k)}$$

- C_i - *a priori* probabilities to be a particle of the i-type.
  "Particle concentrations", that depend on the track selection.
- $r(s|i)$ – conditional probability density functions to get the signal s, if a particle of i-type hits the detector.
  "Detector response functions", depend on properties of the detector.

Both the "particle concentrations " and the "detector response functions" can in principal be extracted from the data.

Statistical problems with particle identification (1)

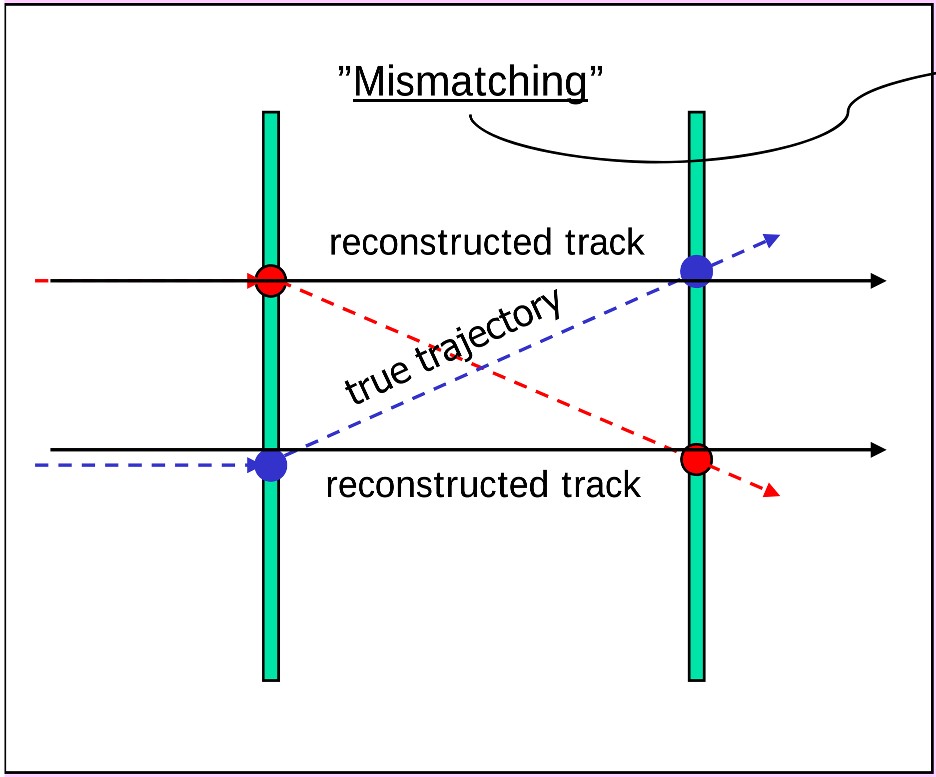


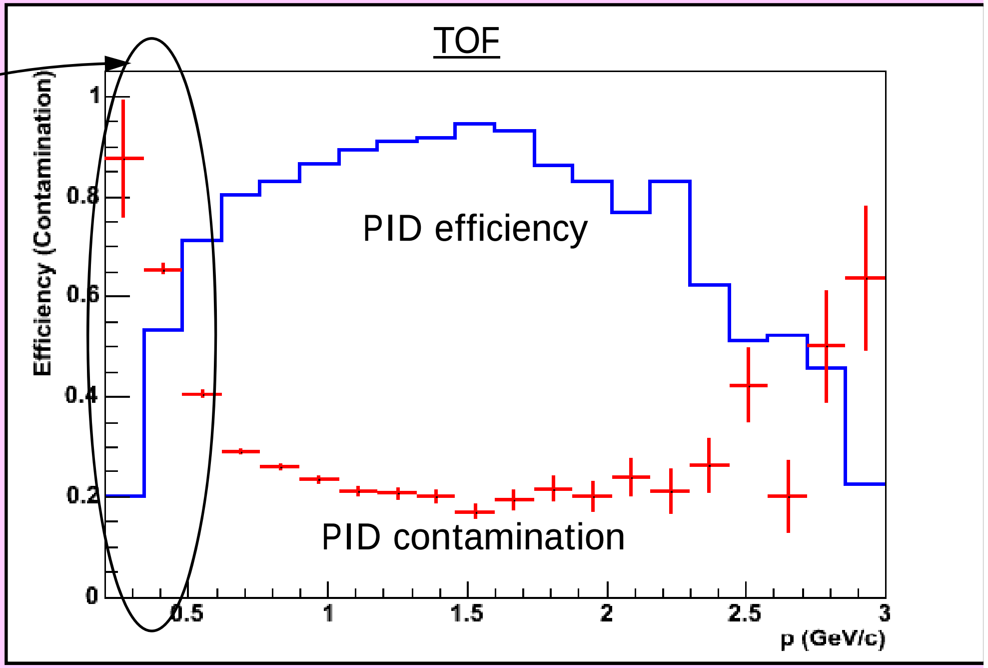
The Bayesian calculations nicely glue together the momentum sub-ranges, but, as the momentum goes up, the "separation power" vanishes, and…

We are left with the bare priors ☹

A question:
The influence of the priors on the final result: Can it be somehow quantified ?

Statistical problems with particle identification (2)

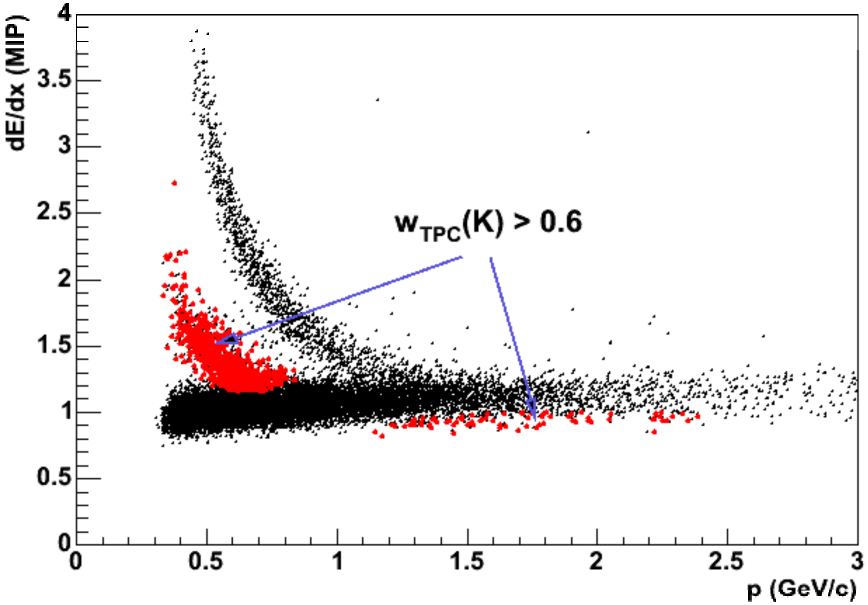




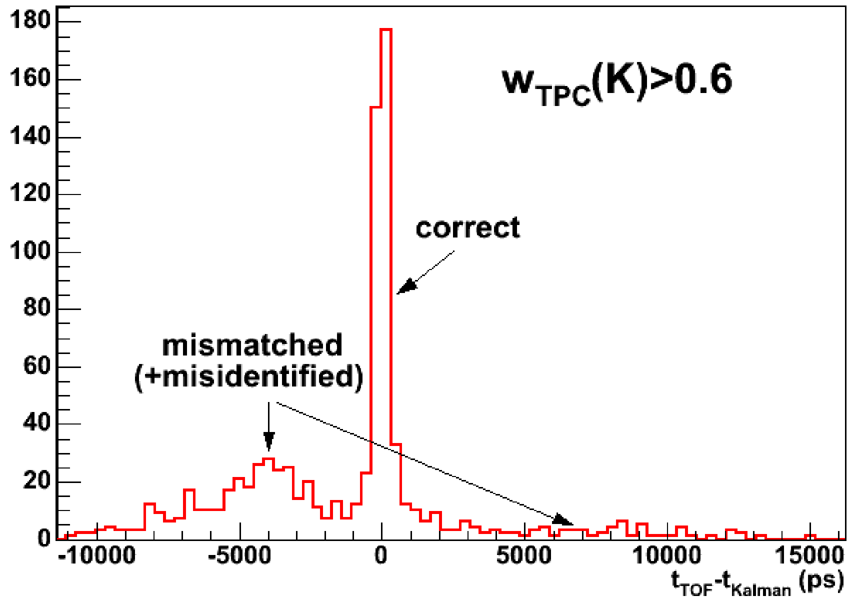
The presence of the mismatching biases the Bayesian combining of the PID information, because one of the main assumptions, that all the detectors register the same particle, is violated.

Ad-hoc correction for the mismatching

TPC



TOF



Observing in one of the detectors the distribution of signals for a clean sample of particles pre-selected in other detectors, we can get the range of signals, where the probability of mismatching is "high" → Veto in the combining…

A question: Can it be somehow generalized ? Made "smooth" ? Optimized ?

Something like $w = (1-p_{12})w_1 + p_{12}w_{12}$ (p_{12} - prob. of a correct matching) ?

ALICE reconstruction: statistical wish-list

- Kalman filter with a constraint.
- Possibility to compare the relative importance of the contributions of the priors and measurements to the Bayesian decision.
- A method to take into account possible mis-measurements in the Bayesian (and likelihood) “combining of information”.