Off-line combined PID

(I.Belikov on behalf of the detector groups)

Outline

- PID in the framework of the ALICE off-line reconstruction.
- Obtaining PID signals at the level of detectors.
- Bayesian PID with a single detector.
  - Obtaining the conditional probability density functions.
  - Obtaining the *a priori* probabilities.
- PID combined over several detectors.
- Conclusions.

PID in the framework the off-line reconstruction

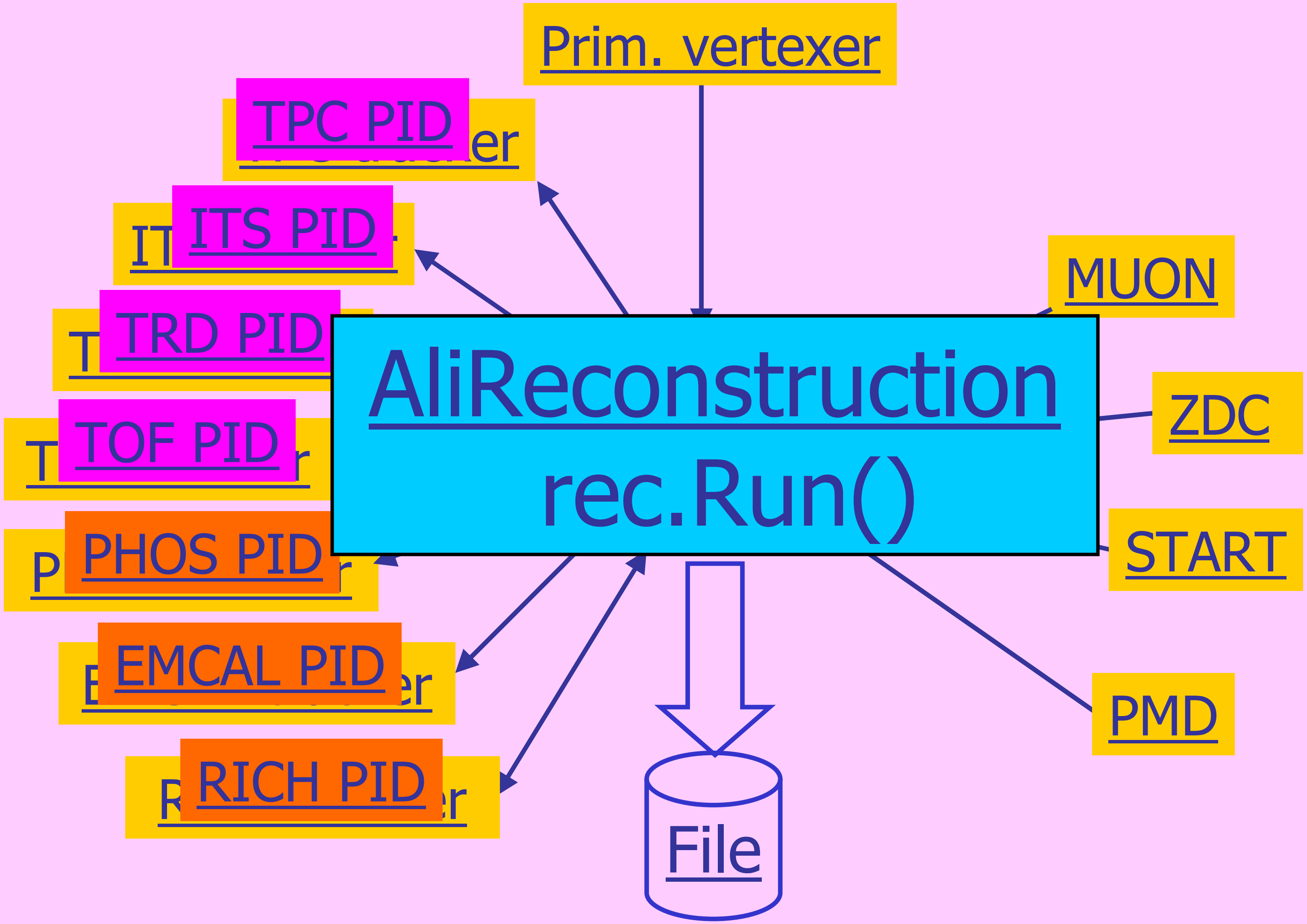
- The off-line reconstruction performs the "transformation":

digits (RAW/simulated) -> ESD

(ESD: tracks+PID+V0/Cascades+"general event information")

- PID is an integral part of the off-line reconstruction.
  - Importance for the ALICE physics program.
- Any detector PID procedure has (minimum) two parts:
  - matching/tracking;
  - calculating some "PID weights".
- Feedback PID -> tracking:
  - corrections for the energy losses and multiple scattering;
  - secondary vertex reconstruction "on flight".

PID in the framework the off-line reconstruction



Obtaining PID signals at the level of detectors

- PID signal from the ITS/TPC/TRD: truncated mean dE/dx
  - There are some detector specific “tricks”.
- PID signal from the TOF: difference between measured and “predicted” time-of-flight.
  - uses time “predictions” estimated by the Kalman tracking.
- PID signal from the RICH: Θ *vs* momentum.
  - momentum is estimated by the Kalman tracking.
- PID signal from the PHOS/EMCAL : shower, shape of the shower, signals from the CPV …
  - rather independent from the rest of the reconstruction.
- General requirements:
  - Interface with the ESD;
  - PID in the form of the “detector response probabilities”.

Bayesian PID with a single detector

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

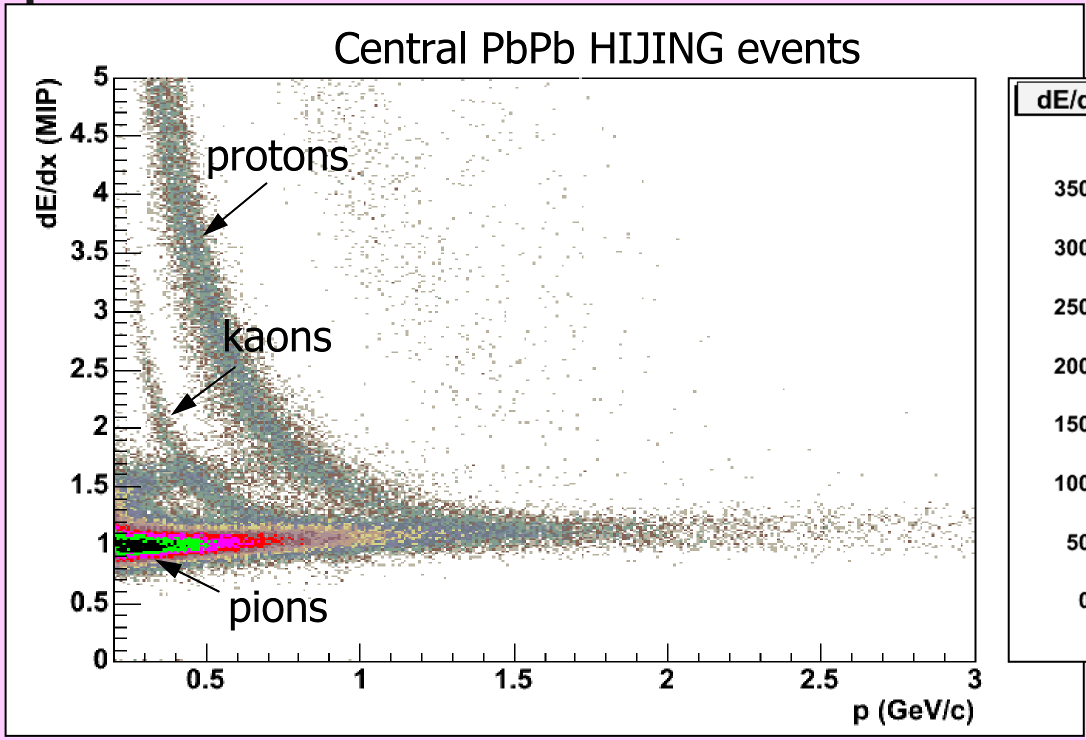
$$w(i\,|\,s) = \frac{C_i r(s|i)}{\sum\limits_{k=e,\mu,\pi,\ldots} C_k r(s\,|\,k)}$$

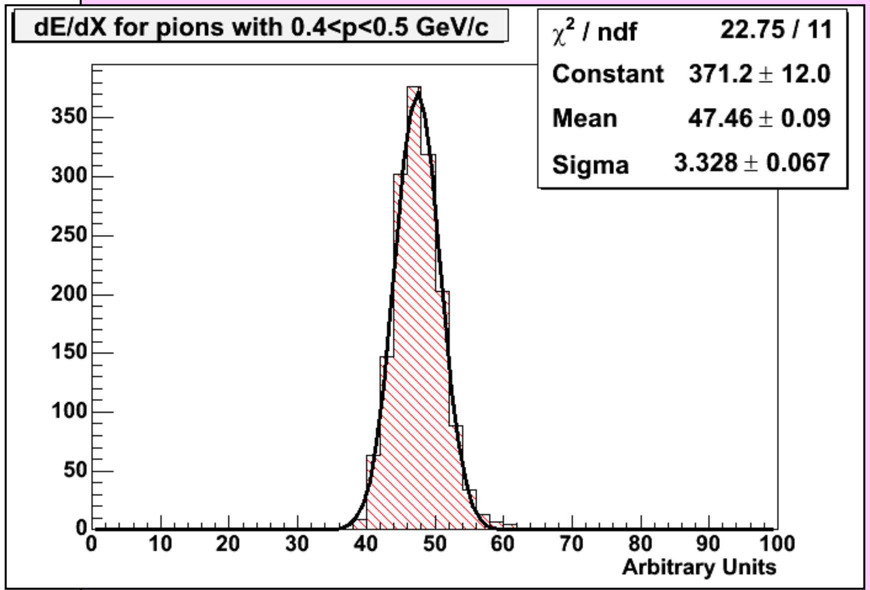
- C_i - *a priori* probabilities to be a particle of the i-type. "Particle concentrations", that depend on the track selection.
- $r(s|i)$ – conditional probability density functions to get the signal s, if a particle of i-type hits the detector.
  "Detector response functions", that depend on properties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

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

Bayesian PID with a single detector

Example of obtaining the conditional PDFs: "TPC response function"

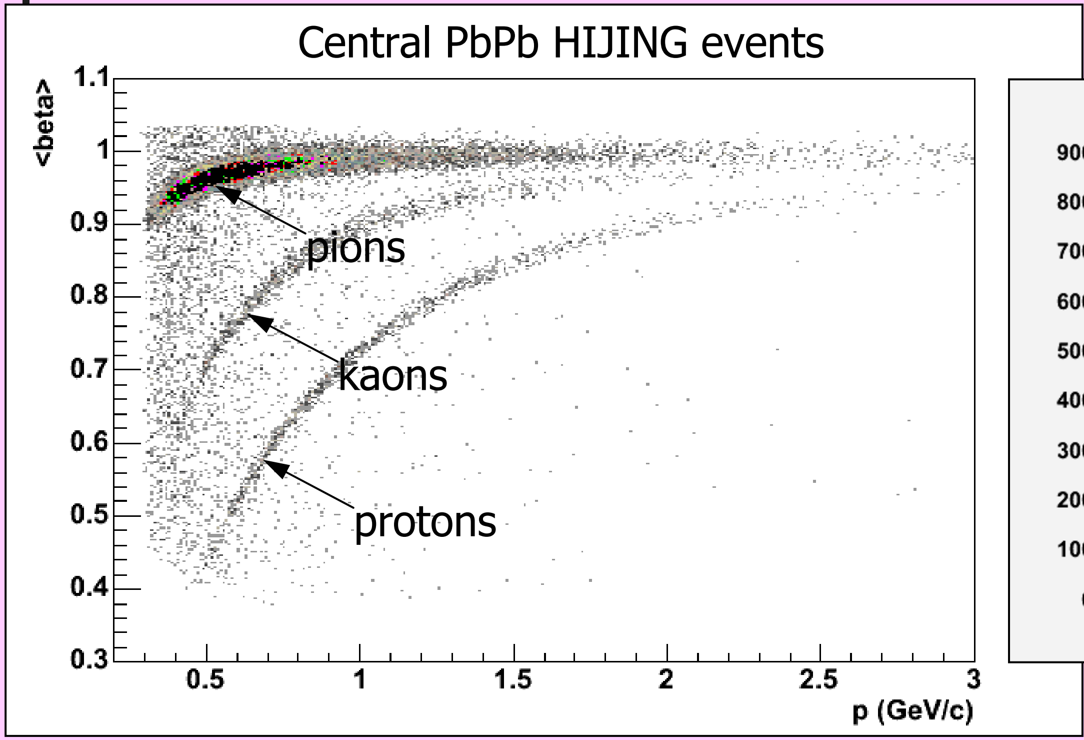


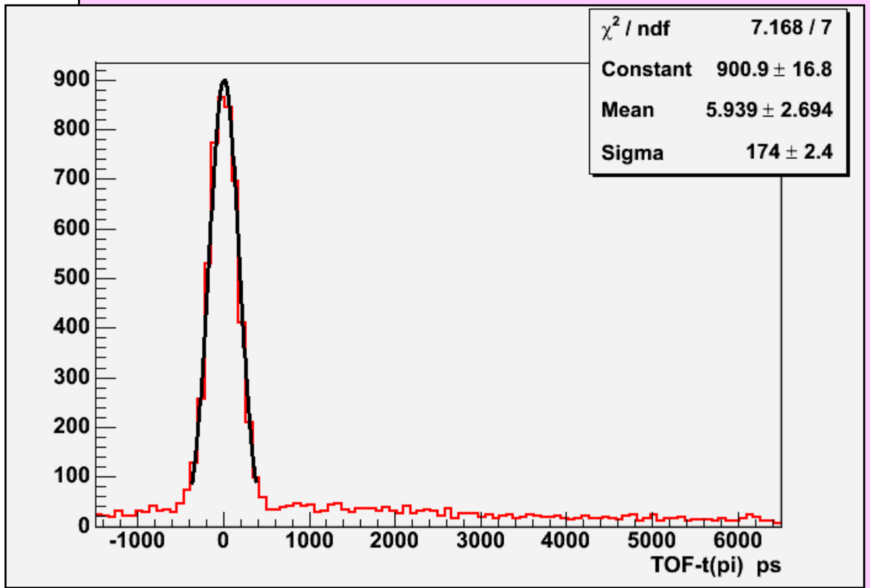


For each momentum p the function $r(s|i)$ is a Gaussian with centroid $<dE/dx>$ given by the Bethe-Bloch formula and sigma $\sigma = 0.08<dE/dx>$

Bayesian PID with a single detector

Example of obtaining the conditional PDFs: "TOF response function"

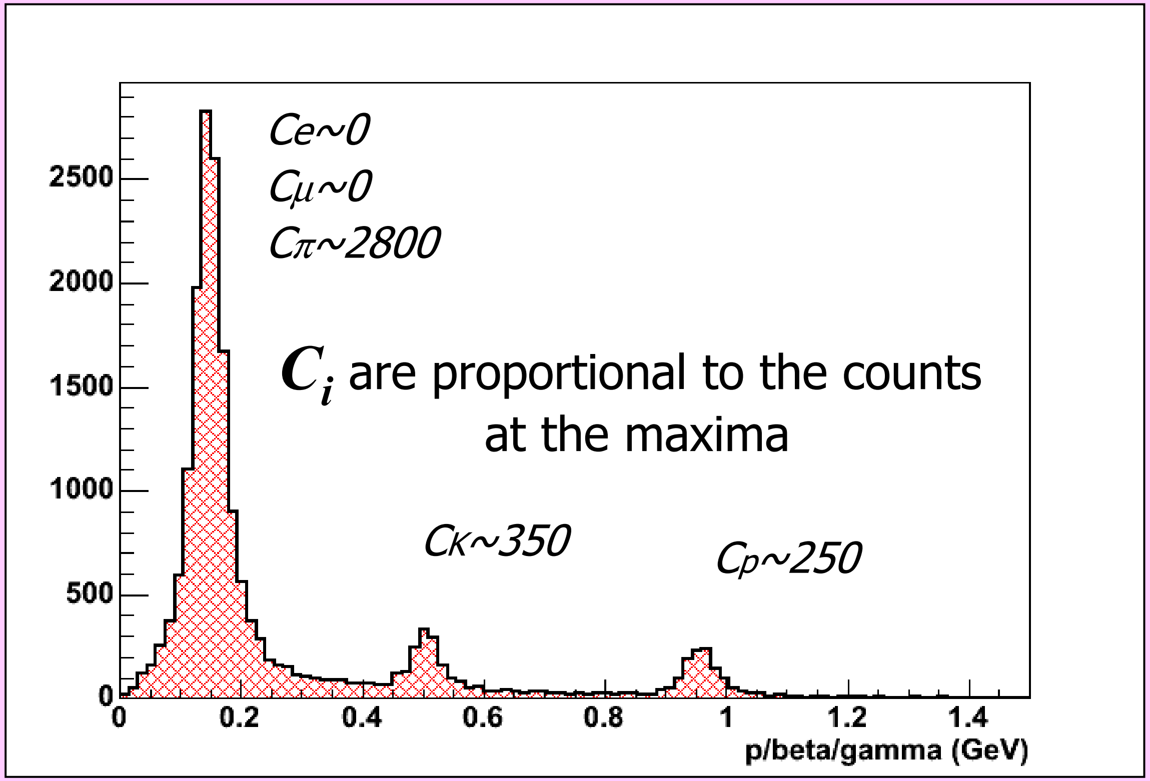




For each momentum p the function $r(s|i)$ is a Gaussian with centroid at 0 and σ given by the distribution of $(TOF\text{-}t_{\pi})$, t_{π}– time calculated by the tracking for the pion mass hypothesis.

Bayesian PID with a single detector

Example of obtaining the *a-priori* probabilities ("particle concentrations")



Selection ITS & TPC & TOF
Central PbPb HIJING events

- *p* – track momentum measured by the tracking
- *beta=L/TOF/c*
  L – track length measured by the tracking

The "particle concentrations" depend on the event and track selection !

PID combined over several detectors

Probability to be a particle of i-type ($i = e, \mu, \pi, K, p, \ldots$),
if we observe a vector $S = \{s_{ITS}, s_{TPC}, s_{TOF}, \ldots\}$ of PID signals:

$$W(i\,|\,S) = \frac{C_i R(S|i)}{\sum\limits_{k=e,\mu,\pi,\ldots} C_k R(S\,|\,i)}$$

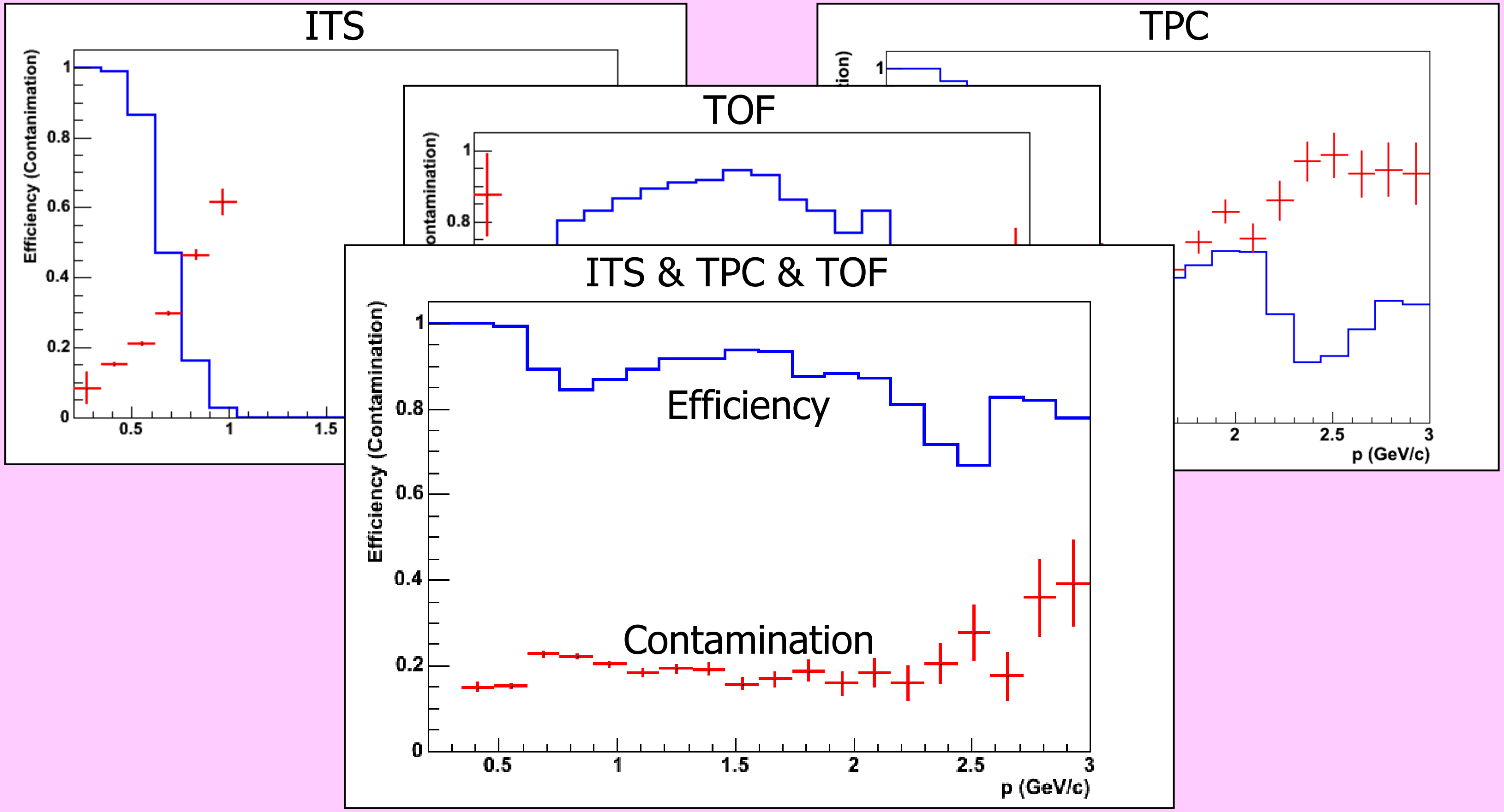
C_i are the same as in the single detector case (or even something reasonably arbitrary like $C_e \sim 0.1$, $C_\mu \sim 0.1$, $C_\pi \sim 7$, $C_K \sim 1$, …)

$$R(S\,|\,i) \approx \prod_{d=ITS,TPC,\ldots} r_d(s_d\,|\,i)$$ are the combined response functions.

The functions $R(S|i)$ are not necessarily "formulas" (can be "procedures").
Some other effects (like mis-measurements) can be accounted for.

PID combined over ITS, TPC and TOF (Kaons)

Selection : ITS & TPC & TOF (central PbPb HIJING events)



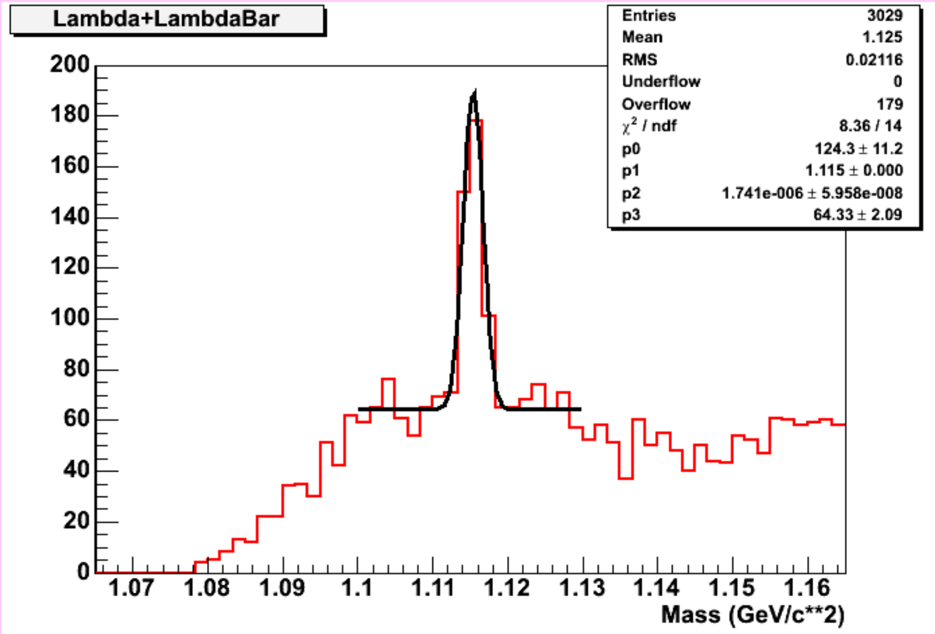
Efficiency of the combined PID is higher (or equal) and the contamination is lower (or equal) than the ones given by any of the detectors stand-alone.

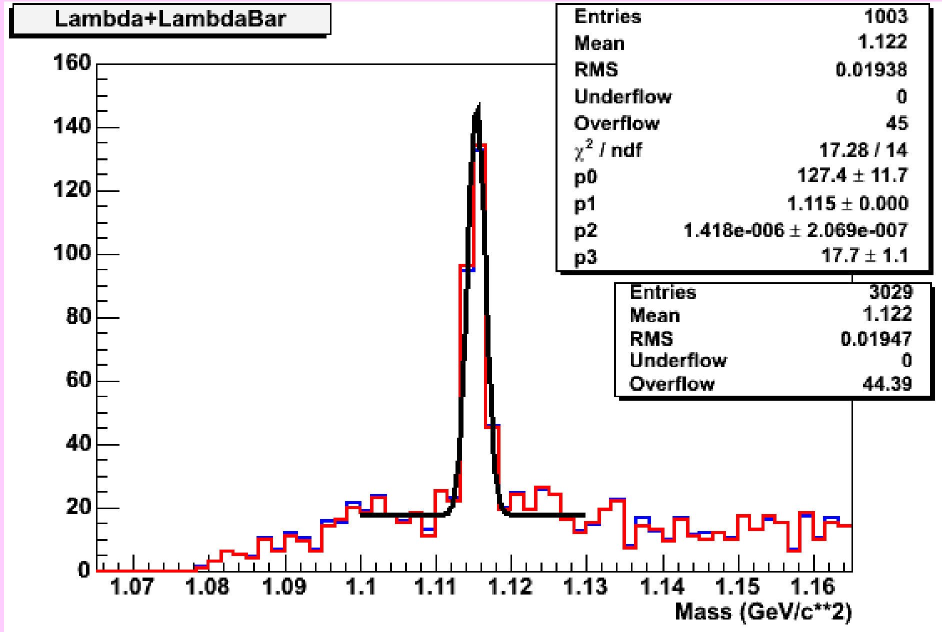
A few results of applying…

(290 events from the jet production)

Fitted with p0*exp(-0.5*(x-p1)*(x-p1)/p2) + p3

Mass=p1=1115 MeV σ=√p1=1.3 MeV





All candidates from the ESD

S ~ p0 = 124

S/B ~ p0/p3 = 1.9

PID with Cπ : CK : Cp = 1 : 0 : 1

S ~ p0=127

S/B ~ p0/p3 = 7.1

Conclusions

- The ALICE off-line PID is done in a Bayesian way. It consists of three parts:
  - Calibration part, performed by the calibration software. Obtaining the single detector response functions.
  - "Constant part", performed by the reconstruction software. Calculating (for each track) the values of detector response functions, combining them and writing the result to the Event Summary Data (ESD).
  - "Variable part", performed by the analysis software. Estimating (for a subset of tracks selected for a particular analysis) the concentrations of particles of each type, calculating the final PID weights by means of Bayes' formula using these particle concentrations and the combined response stored in the ESD.

Conclusions

- The procedure allows to combine PID signals of quite different nature (***dE/dx***,TOF,Cherenkov,…) in a common way.
- It naturally takes into account the fact that the PID depends, due to different event and track selection, on a particular kind of performed physics analysis.
- The procedure is fully automatic. No cuts (graphical/multidimensional) are involved.