SBumps

Analysis framework for automatic search and identification of peaks

S.V.Chekanov (ANL)

HERA-LHC workshop
(DESY June 2, 2004)

Test version:

http://www.desy.de/~chekanov/sbumps/

Introduction

- To search peaks in invariant masses is a tedious task (especially if you do not know that your are looking for)
  - need to check many mass assumptions
  - 2,3,4 etc. body decays should be looked at
  - reflections from known states should be removed

Example: Look at 2 and 3 and 4 body decays using
pion, **proton**, **kaon** mass assumption –
6 two-particle mass non-identical combinations
10 three-particle non-identical combinations
20 four-particle non-identical combinations
+ various many charge combinations !

Sbumps:
automatic search and identification of peaks

- written in C++ using ROOT libraries
- Takes any input (3-momenta + probabilities for each particle)
- For given mass mass assumptions, automatically creates necessary histograms
- Fills histograms
- Automatically searches for peaks
- Identifies known PDG states and reflections
- Makes reports on unknown states

Inputs

- Root ntuple with Px,Py,Pz
- (optional) vector with probabilities that particle is pion, photon, etc (up to 98 states)
- Define which mass assumptions to use to calculate invariant mass
- Run over 2-particle, 3-particle decays (4-particle decays – in future)
- Define (naïve) significance level for final peaks
- Define bin width, expected 2-particle resolution and maximum value for invariant mass (min- done automatically)

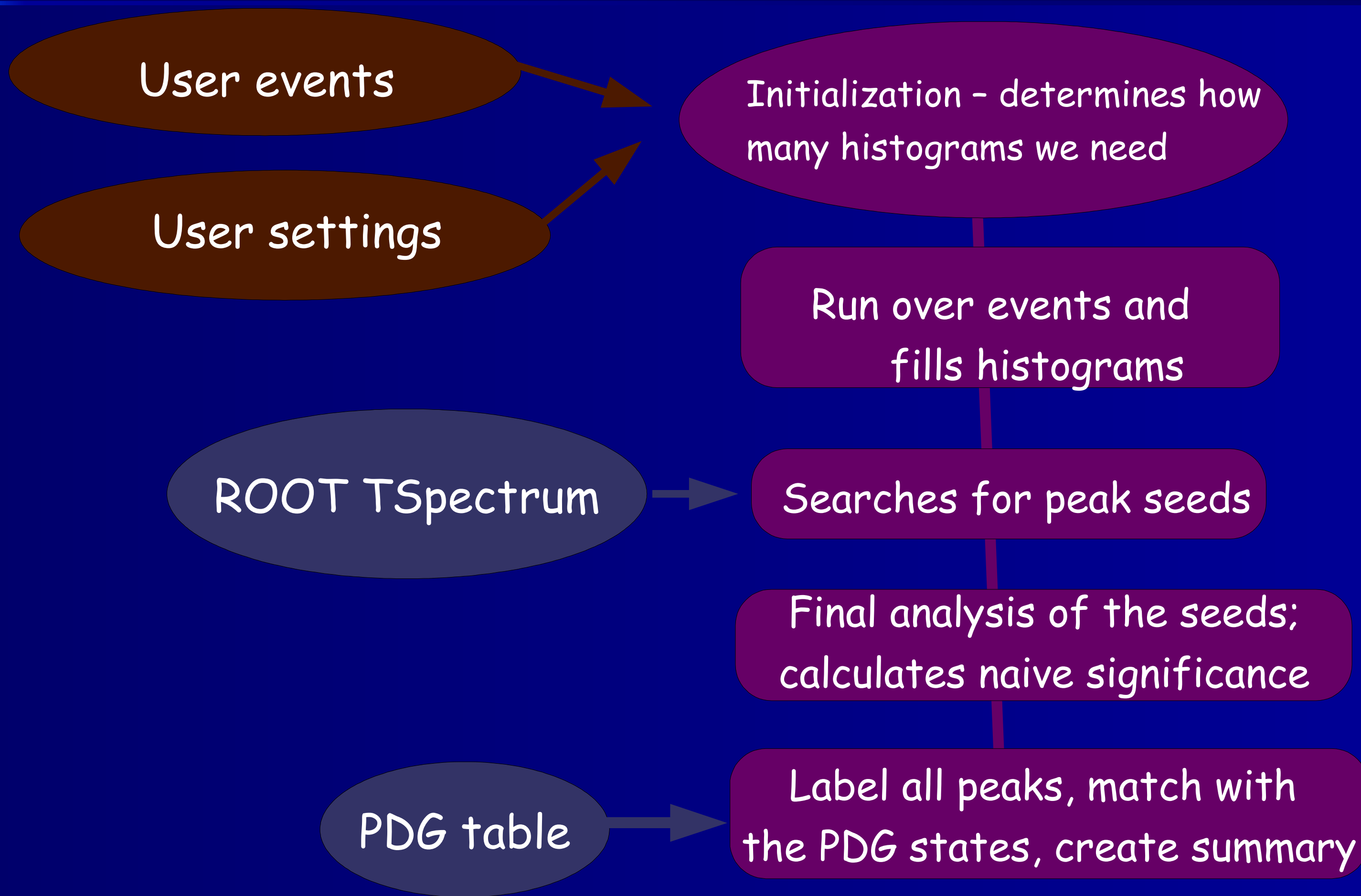
Outputs

- Summary of observed peaks
- Root histograms with invariant masses
- PDG states are identified and labeled automatically (taking into account PDG and expected experimental mass uncertainties)

Peak searching

- Based on ROOT TSpectrum class
- Fast algorithm using Markov approach for peak searching in presence of background and statistical noise
- This was developed for gamma-ray physics and usually does not work correctly for searches in invariant masses
- Therefore, this algorithm was used only to create "seeds" with suspected peaks
- Final peaks were identified after analysis of the seed peaks

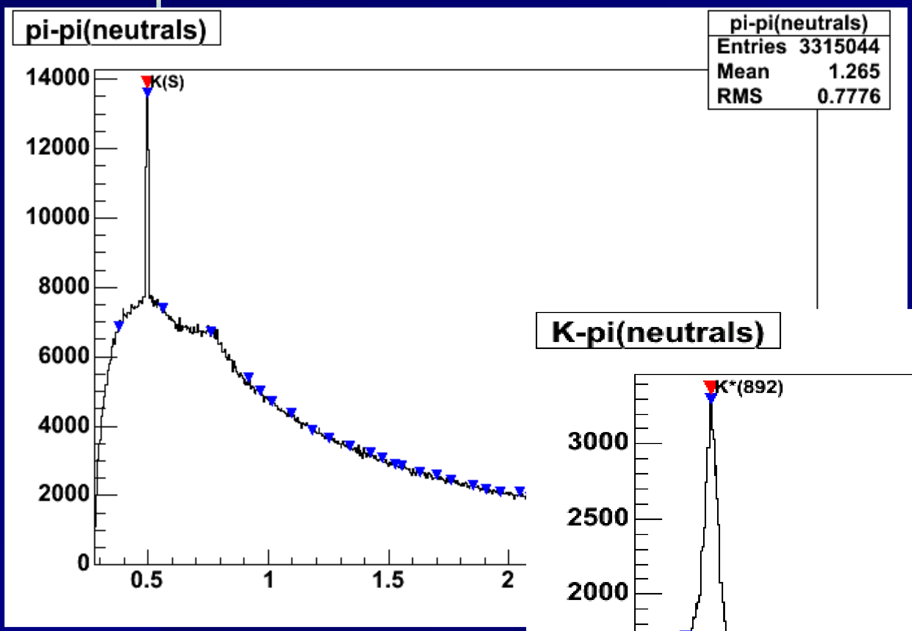
General structure



Simple example

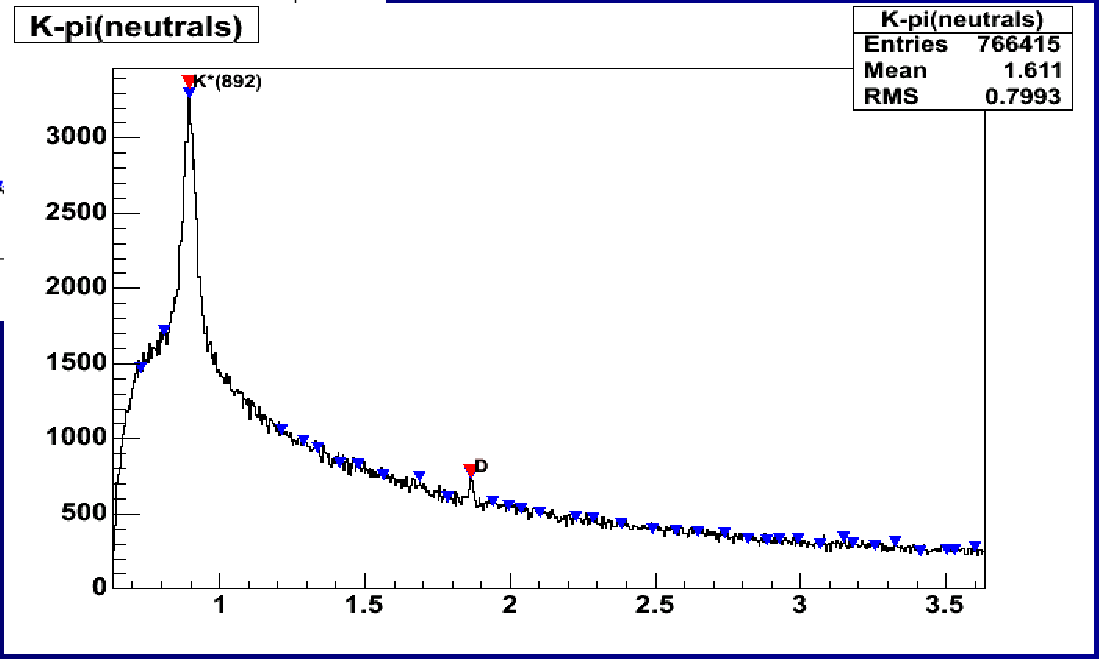
- Read event record from RUNMC (PYTHIA ep)
  http://www.desy.de/~chekanov/runmc
- Smear Px,Py,Pz momenta (i.e. toy tracking simulation)
- For each pion, proton, gamma, K-meson, add vectors containing various probabilities
- Example: For a particle with pion ID:
  - 0.7 probability that this is a pion (using Gaussian smearing with average 0.7 and sigma=0.1)
  - 0.3 probability that this is a proton (using Gaussian)
  - Do similar for protons and gammas
- Ask for peaks with more than 3 sigma and particle probability > 0.7
- Masses of PDG states should not be more than 3 sigma away from the found peak. Information on charge is included

SBumps output: automatically created histograms

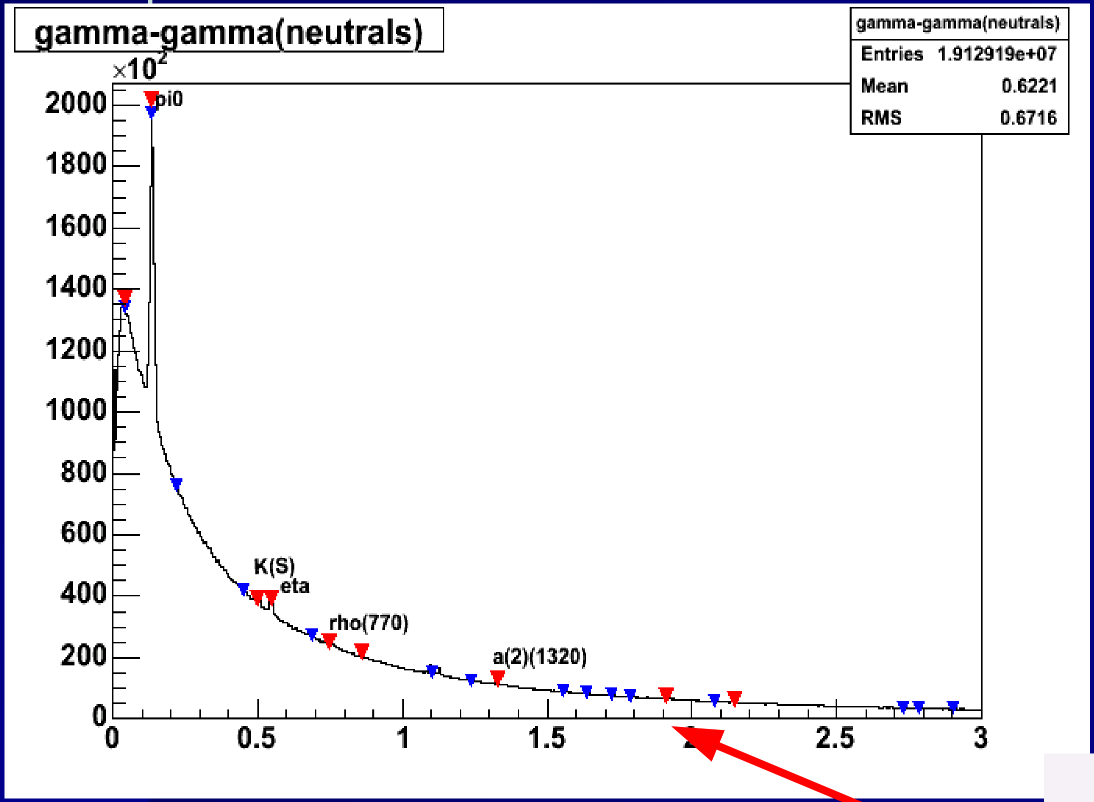


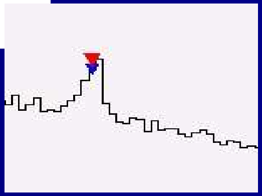
Blue markers – seeds used for future peak searching ("potential peaks")

Red markers – final peaks (also identified as PDG states)



SBumps output: more complicated example



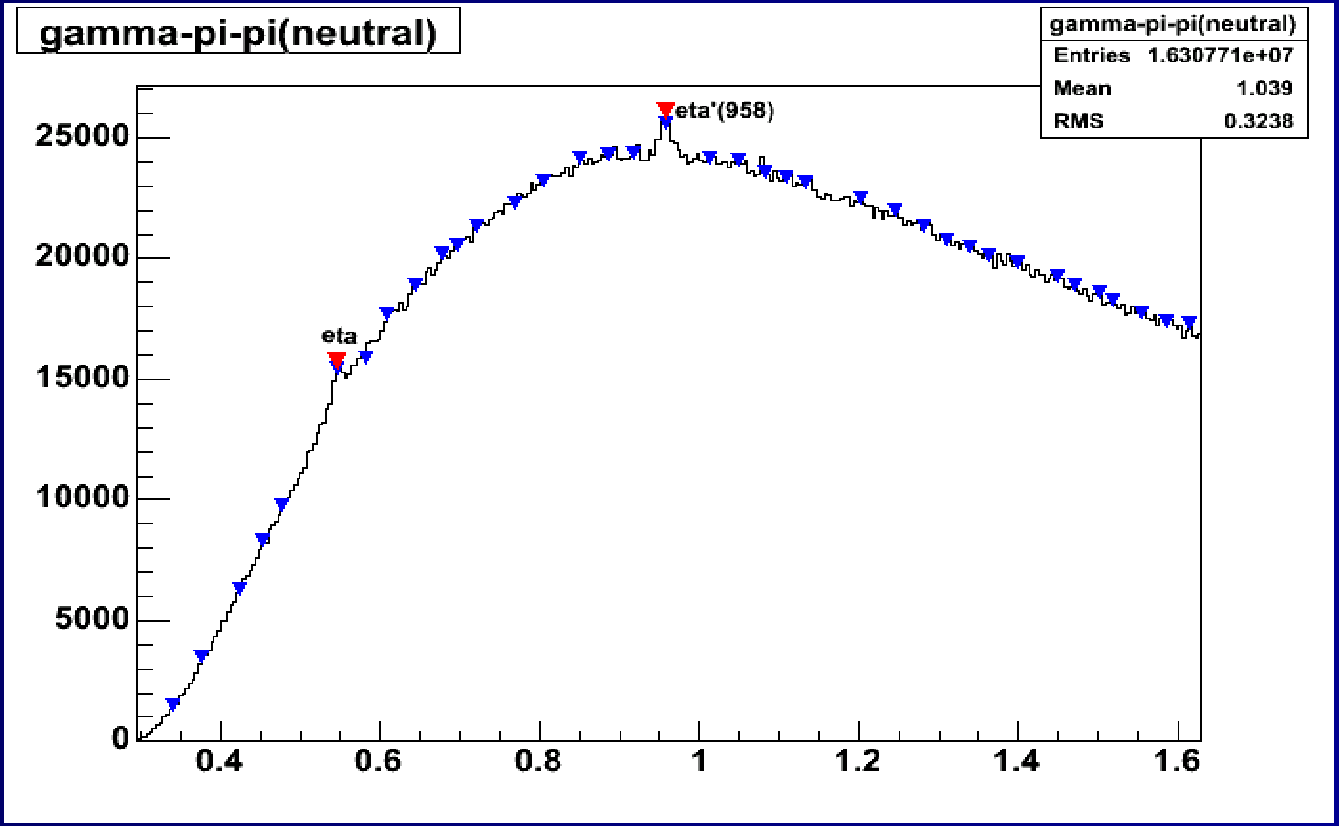


5 peaks are identified!

1 peak – background shape

3 peaks – found, but could not be matched with known PDG states "reflections"?

SBumps output:
3-body decay example



Two peaks are found and identified
but too many seeds used to do this!

Summary

- First step towards automatic peak searching algorithm in presence of continuous background:
  - still need to improve the peak searching algorithms to avoid "fake" peaks
  - PDG information on decays channels will be included to avoid peak misidentification
- It cannot do full physics analysis – it rather helps to identify invariant mass distributions which could be interesting for further studies
  - maybe pentaquarks at LHC?!
- If the input contains previously reconstructed particles – it can go beyond 4-particle decays
- Automatic search for reflections – will be done in future