STATISTICAL ANALYSIS IN EXPERIMENTAL PARTICLE PHYSICS

Kai-Feng Chen
National Taiwan University

EXERCISE 3

EXAMPLE DATA SAMPLES

- Today we are going to have a "chained" exercises and the example data samples are given below:
  - **"Experimental data":**
    http://hep1.phys.ntu.edu.tw/~kfjack/lecture/hepstat/03/example_exdata.root
  - **"J/ψK$^+$ Monte Carlo":**
    http://hep1.phys.ntu.edu.tw/~kfjack/lecture/hepstat/03/example_signal.root
  - **"J/ψπ$^+$ Monte Carlo":**
    http://hep1.phys.ntu.edu.tw/~kfjack/lecture/hepstat/03/example_psipi.root
  - **"J/ψ+hadrons Monte Carlo":**
    http://hep1.phys.ntu.edu.tw/~kfjack/lecture/hepstat/03/example_psihadrons.root
- Let's practice model building and structure your own maximum likelihood fitter with these samples!

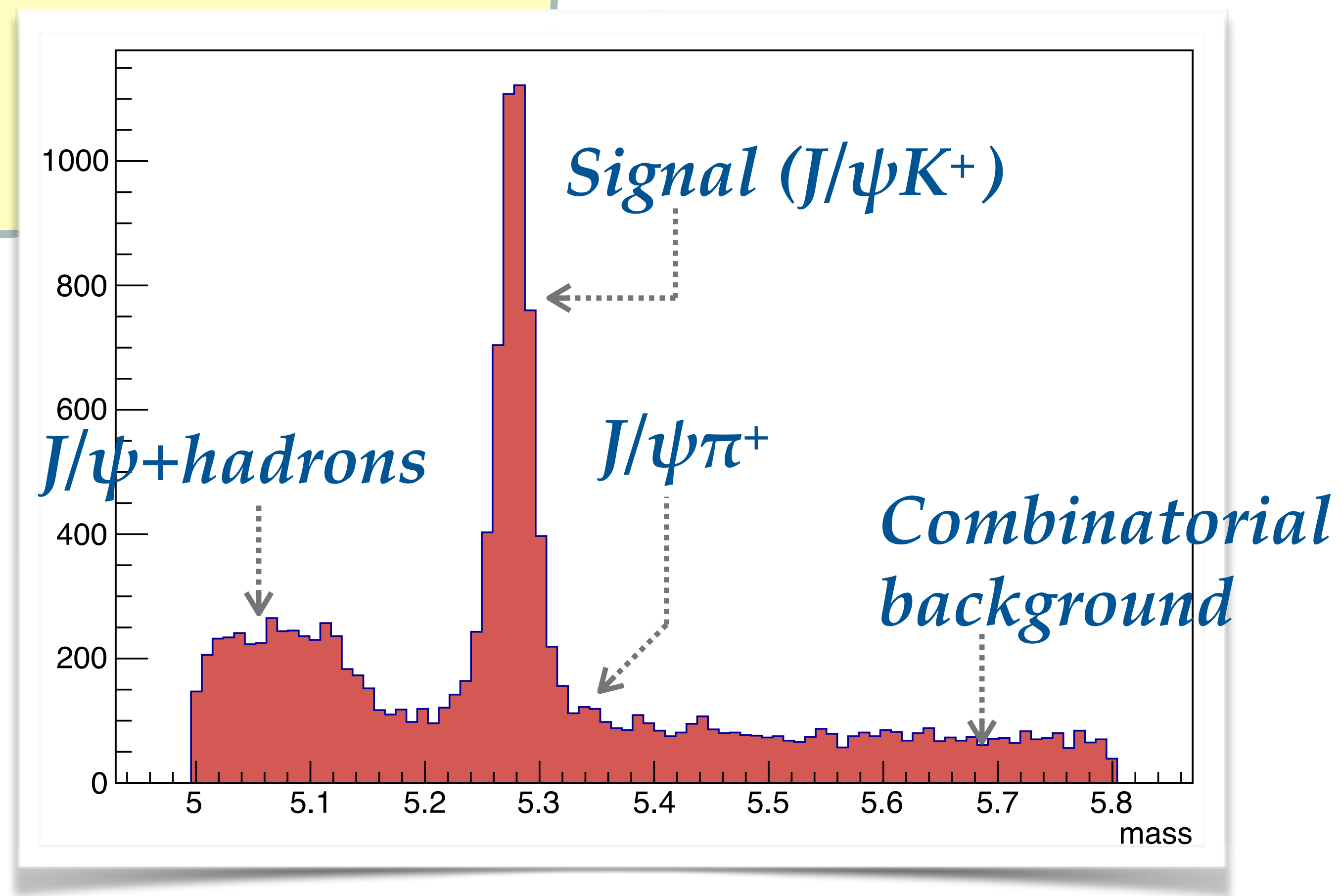
THE DATA FILES CONTAIN…

- The data files contain some "$B^+ \to J/\psi K^+$" reconstructed events, and in this exercise you are requested to perform a proper maximum likelihood fit to the events, and extract the parameter(s) of interests.

```
{
    TFile *fin = new TFile("example_exdata.root");
    TNtupleD *nt = (TNtupleD*)fin->Get("nt");

    nt->SetFillColor(50);
    nt->Draw("mass");
}
```

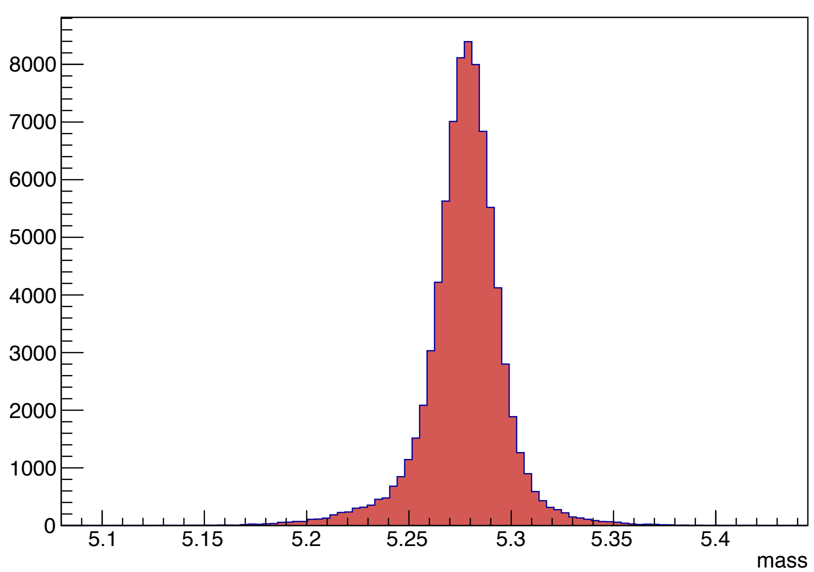
Note there are also other variables in the ntuple:
charge = candidate charge
pt = transverse momentum (in GeV)
eta = pseudorapidity
tau = decay time (in ps)



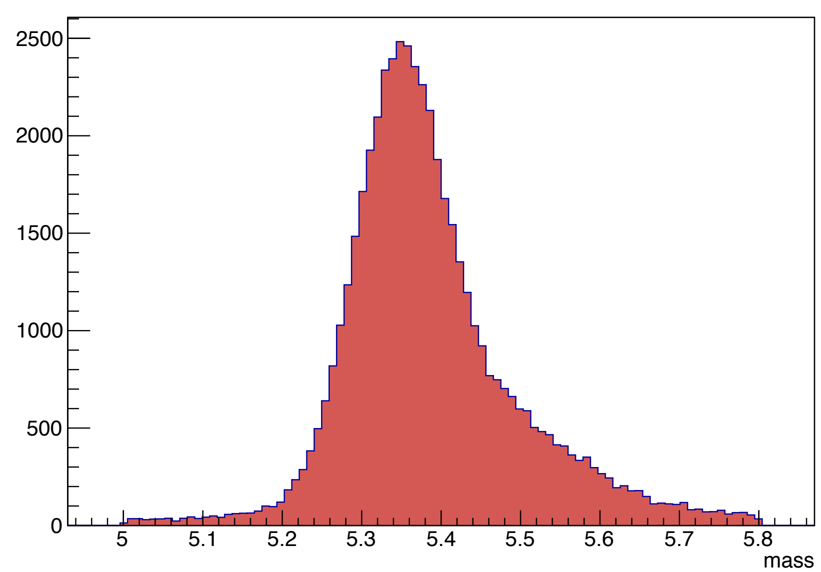
#1: PREPARE MASS MODELS FOR EACH COMPONENT

- Please model the mass distribution of the three components which have "Monte Carlo" samples: **signal**, **J/ψπ$^+$**, **J/ψ+hadrons**.
- You choose (*guess*) what is the best function to model them.
- Present your fit results and fit projection plots!

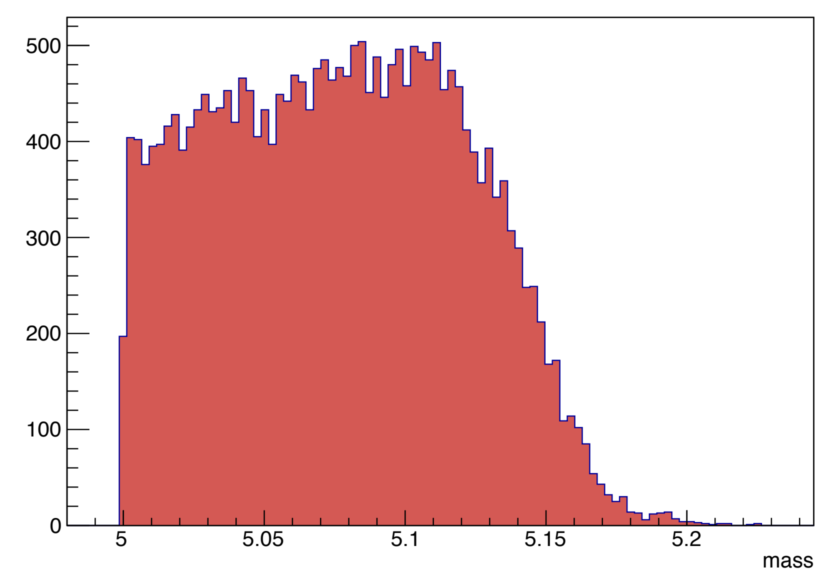
Signal (J/ψK$^+$)



J/ψπ$^+$

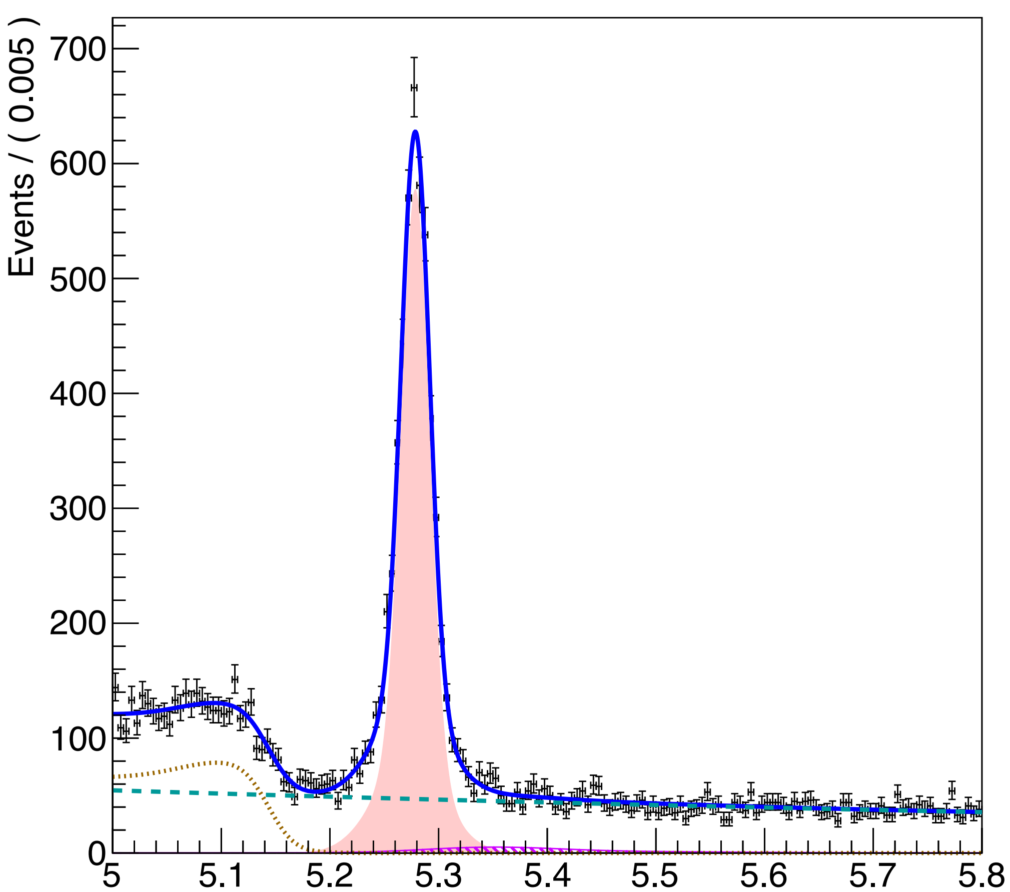


J/ψ+hadrons



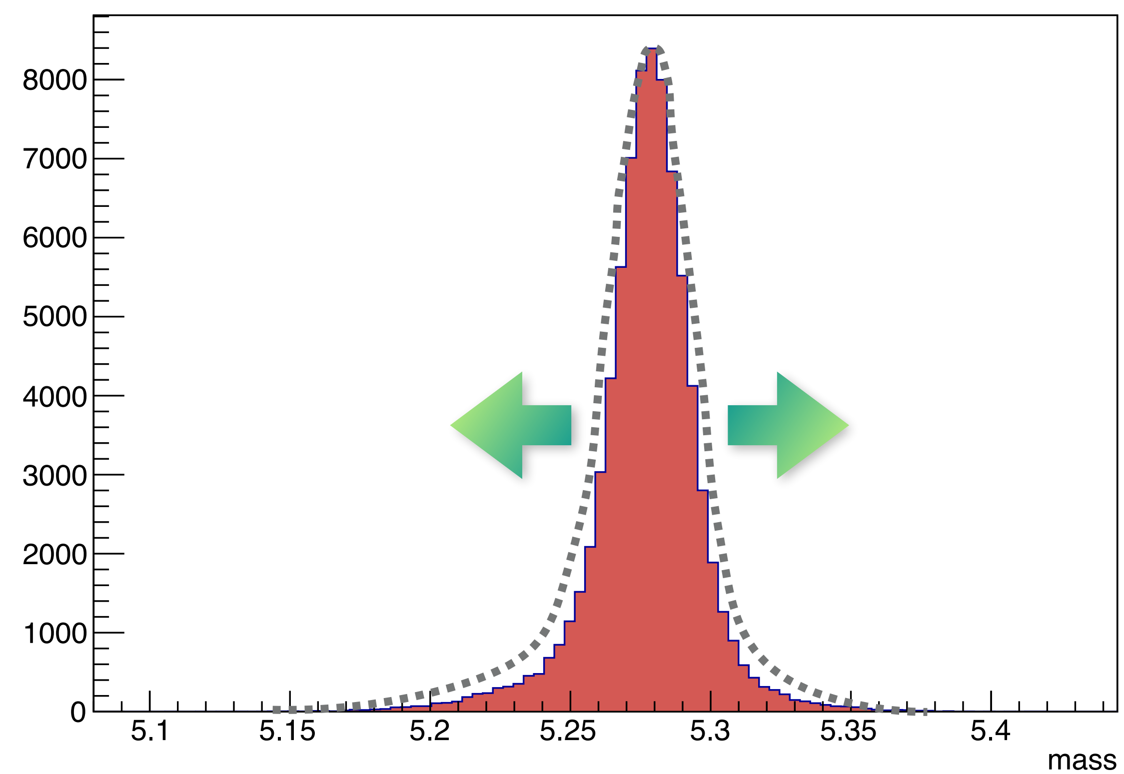
#2: JOINT THEM TOGETHER

- Now you have the PDF for all the components, **signal**, **J/ψπ⁺**, **J/ψ+hadrons**, plus a **combinatorial** background which does not have Monte Carlo samples, into a single **extended likelihood function** and perform a fit to the "experimental data".
- Present your result of the fits (*plot*) and the yields for each component!
- **You can fix most of the PDFs to the Monte Carlo shapes first.**
- This is more-or-less close to you can have:



#3: A CORRECTION TO THE SIGNAL WIDTH

- You may find that the signal width is somewhat not exactly the same between data and MC.
- Please try to apply an **additional correction (*scaling factor*)** to the width of your signal peak model and fit it with data.
- Present your results again.



#4: CONSTRAINING IT

- You may find the $J/\psi\pi^+$ component is hiding under the $J/\psi K^+$ signal peak, and the fit is not so nice.
- But however we can adding more information by constraining it according to the ratio of branching fractions from PDG:

$$\frac{\mathcal{B}(B^+ \to J/\psi K^+)}{\mathcal{B}(B^+ \to J/\psi \pi^+)} = (4.00 \pm 0.39)\%$$

- Hint: you can replace the yield of $J/\psi\pi^+$ by this ratio times the yield of $J/\psi K^+$, and then add a Gaussian constrain term on the given ratio to your fit.

#5: IN BINS OF PT AND ETA

- Now let's study the dependence of **pt** and **eta** (*these variables are also stored in the same ntuple!*)
- Make a plot which shows the **"fitted yields per GeV in pt"** in the following bins of pt:
  ```
  pt =
  [15–16,16–17,17–19,19–21,21–25,25–30,30–40,40–60,60–100]
  ```
- Make another plot which shows the **"fitted yields per 0.2 in eta"**, according to the following binning:
  ```
  eta = [0.0–0.2,0.2–0.4,0.4–0.6,0.6–0.8,0.8–1.0,
  1.0–1.2,1.2–1.4,1.4–1.6,1.6–1.8,1.8–2.0]
  ```
- Note the signal shapes could be different for different bins!

#6: CHARGE ASYMMETRIC?

- Please calculate the following quantity by separating the fits to the events with positive charge and negative charge:

$$A_{CP} = \frac{N_- - N_+}{N_- + N_+}$$

- Try to include this A_{CP} in your fit to the data and extract the value directly within **a single fit**, without separating the samples! (*Hint: this requires a simultaneous fit!*)