Report from Group5 : part2

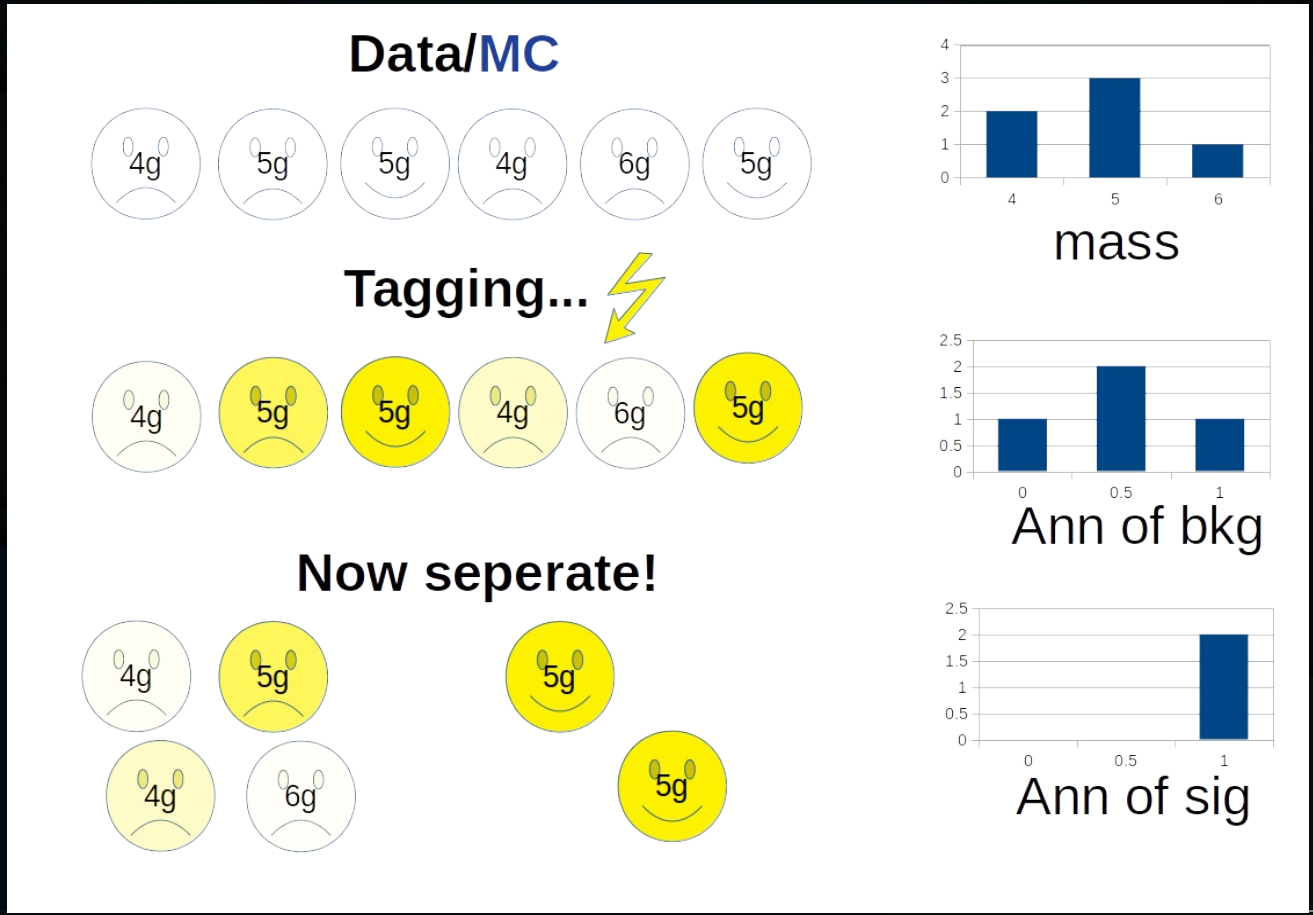
Problem

- Given **MC(sig)**, **MC(bkg)** and **data** 3 datasets
- Each dataset has *mass* and *ann* 2 variables
- Use **MC** to get the **Cut** so that Eff(sig)=90%
- Apply **Cut** on **data**
- Find Signal

Idea

- Use bisection method to find the Cut
- Apply cut on data
- Fit to find signal
- Validation
- Optimization

For short: what is cut?



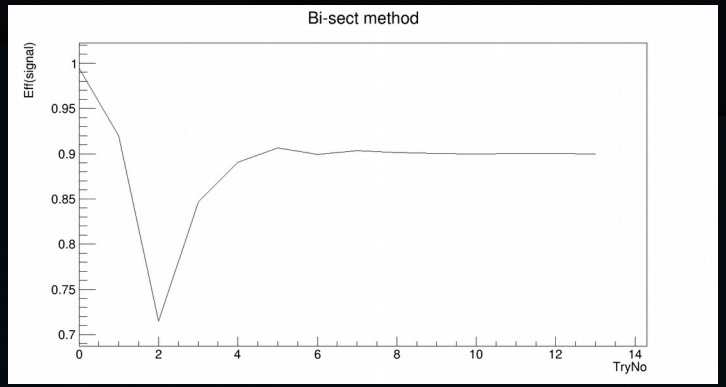
Definition: Metric

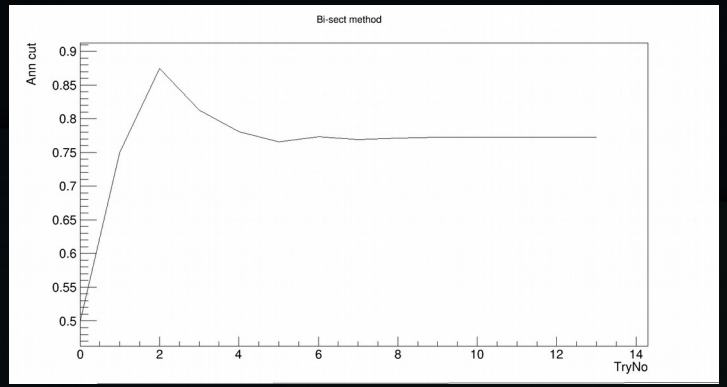
- Eff(sig) = Sig_after_cut / Sig_all = Sig alive
- Rej(bkg)=1 - Bkg_after_cut / Bkg_all = Bkg drop
- “Power” = Eff*Rej

Find the cut

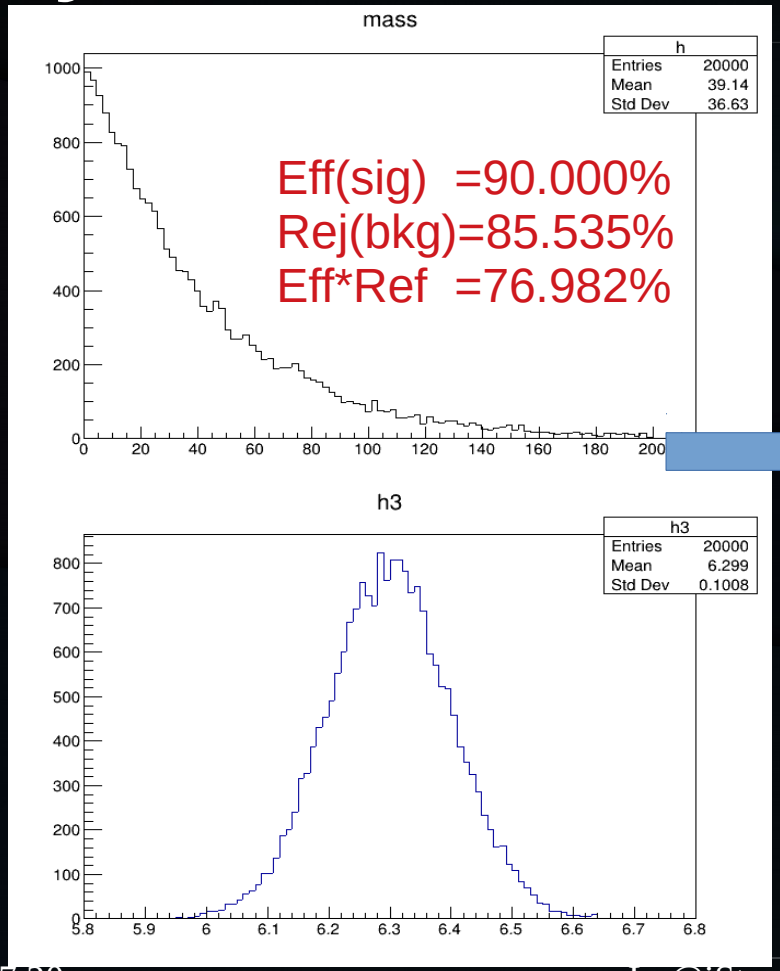
```
for(;1;){
tryNo++;
cond=Form("ann>%f",thre);
Sig=mc_sig->GetEntries(cond);
Ratio=1.0*Sig/All;
if(tryNo>MAXTRY) break;
if(Ratio-require>eps) {minThre=thre;thre=(maxThre+minThre)/2;continue;}
else if(Ratio-require<-eps) {maxThre=thre;thre=(maxThre+minThre)/2;continue;}
else break;
}
```

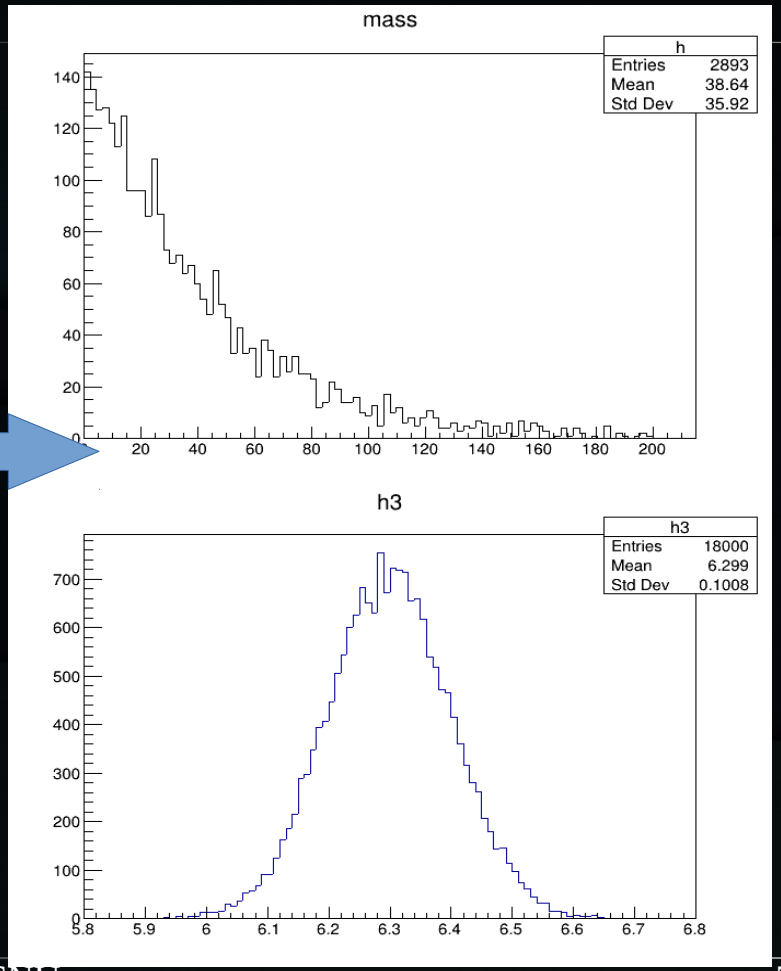
Ann>0.772888 -> Eff=90%





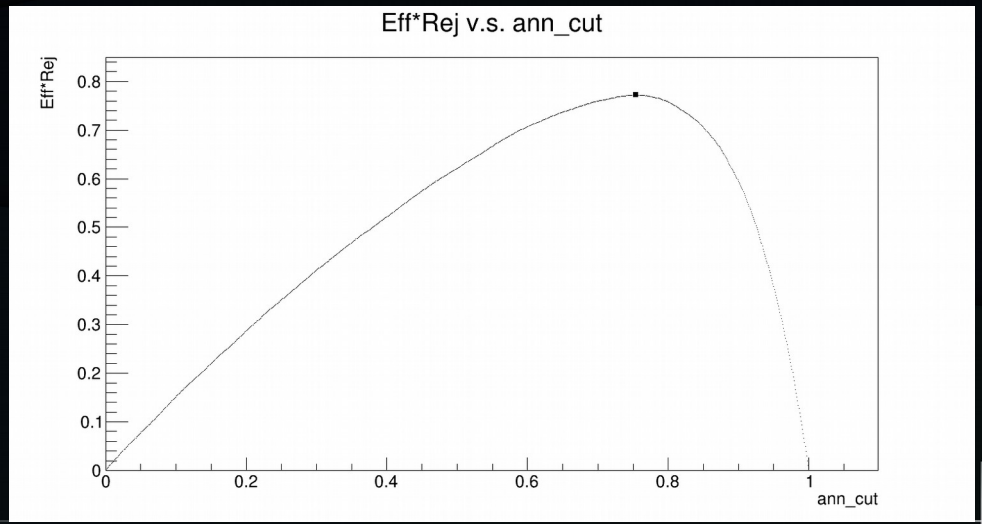
Apply Cut: MC


mass
h
Entries 20000
Mean 39.14
Std Dev 36.63
Eff(sig) =90.000%
Rej(bkg)=85.535%
Eff*Ref =76.982%
h3
h3
Entries 20000
Mean 6.299
Std Dev 0.1008

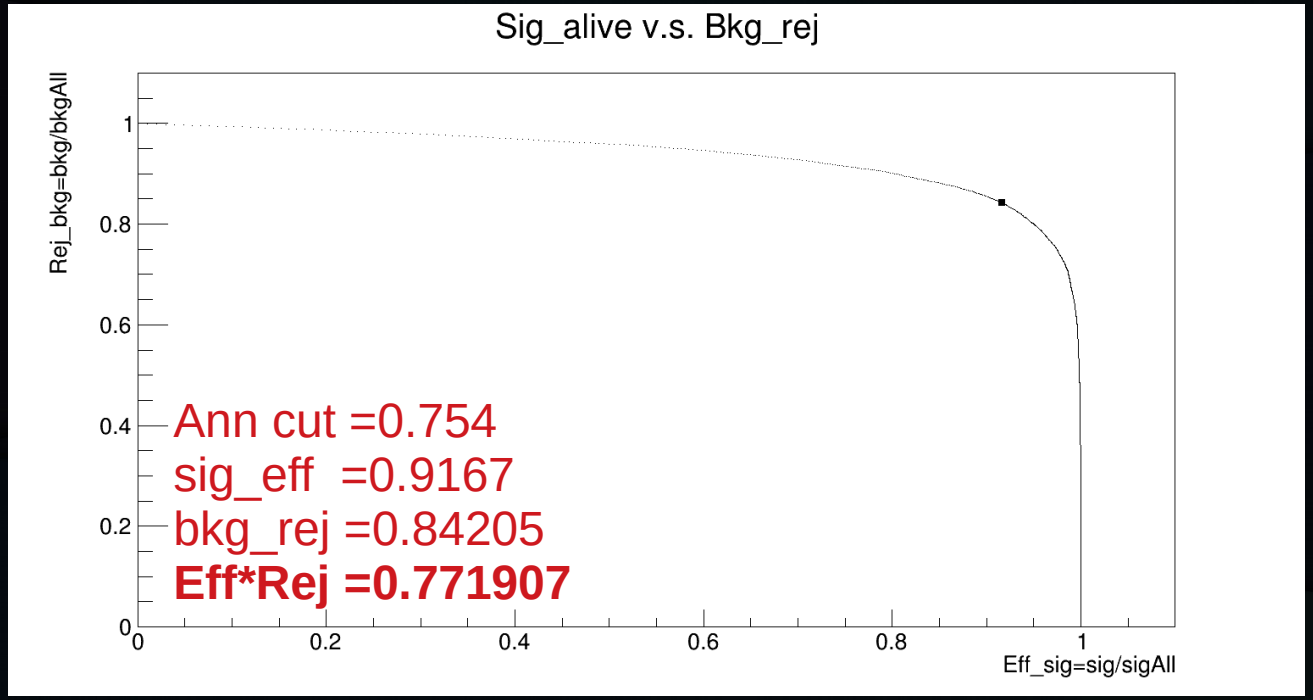

mass
h
Entries 2893
Mean 38.64
Std Dev 35.92
h3
h3
Entries 18000
Mean 6.299
Std Dev 0.1008

Cut optiz

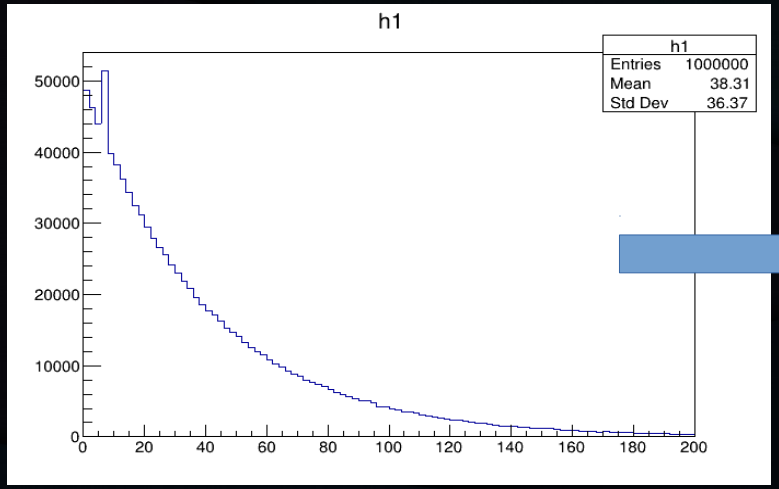
- Use “Power”=Eff*Rej => MAX
  - Ann>0.754

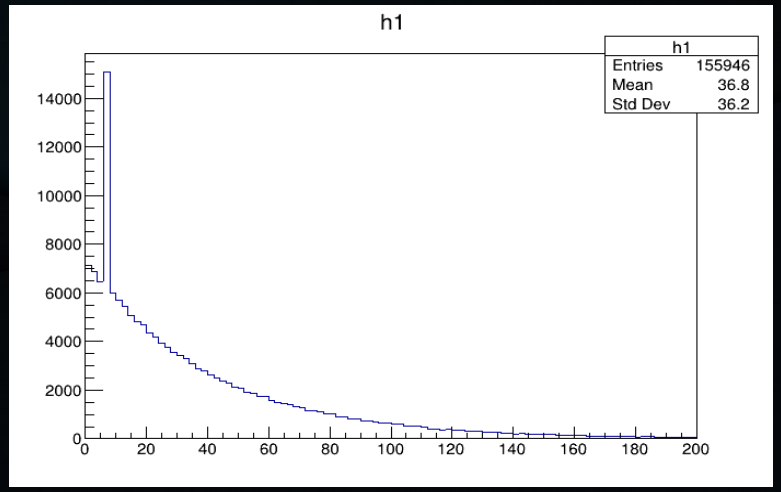


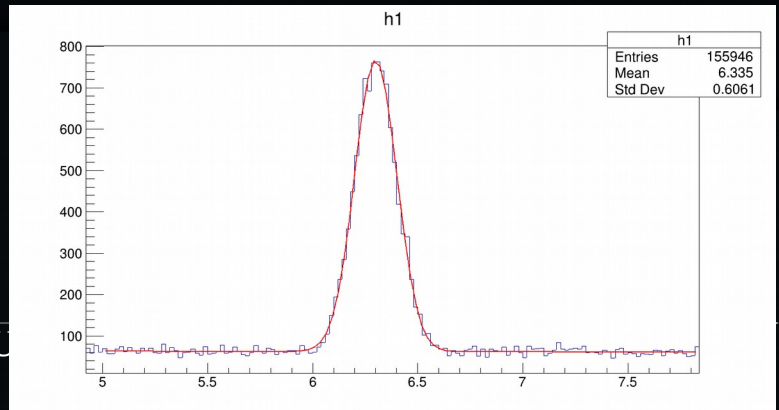
ROC curve: Eff v.s. Rej



Apply Cut: data







Fit

- Method1: use max as mass_suspect, and limit width
- Method2: fit narrow range(5-8)
  - [p0]*exp(-0.5*((x-[p1])/[p2])*((x-[p1])/[p2]))+exp([p3]+[p4]*x)
- Method3: multi-fit
  - Fit bkg region(8-200) with f_bkg (Exp)
  - Fit sig region(5-8) with f_fig (Gaus)
  - Fit All with f_bkg+f_fig, and use below par as initial value.

Concern about Formula (Thanks to Chen Zhizhong)

- gaus(0)+expo(3) => freedom=3+2=5!
- [p0]*exp(-0.5*((x-[p1])/[p2])*((x-[p1])/[p2])) +exp([p3]+[p4]*x)
- k*Gaus(x,u,s)+~~(??)*~~Exp(a+b*x)
- Don’t let computer confusing…
- Look out! STATUS=...bounds outside…

Warning!

```
158071
Info in <ROOT::Math::ParameterSettings>: lower/upper bounds outside current parameter value. The value will be set t
o (low+up)/2
Info in <ROOT::Math::ParameterSettings>: lower/upper bounds outside current parameter value. The value will be set t
o (low+up)/2
 FCN=112.14 FROM MIGRAD    STATUS=CONVERGED     703 CALLS         704 TOTAL
                     EDM=1.04513e-08    STRATEGY= 1      ERROR MATRIX ACCURATE
  EXT PARAMETER                                   STEP         FIRST
  NO.   NAME      VALUE            ERROR          SIZE      DERIVATIVE
   1  p0           1.74965e+03   2.51193e+01   1.04397e-01  -4.63950e-06
   2  p1           6.29959e+00   1.29793e-03   6.73716e-05  -6.89300e-04
   3  p2           1.01286e-01   1.10706e-03   1.32218e-04  -4.27511e-04
   4  p3           5.15939e+00   4.58923e-02   4.35633e-05  -2.30927e-03
   5  p4          -1.53094e-02   5.90890e-03   5.64793e-06  -1.37632e-03
root [1] Info in <TCanvas::Print>: file /home/binus/work/iStep/dayLast/x.png has been created
```

Implement

```
TString cond=Form("ann>%f",0.772888);
TH1D* h1=new TH1D("h1","h1",10000,0,200);
data->Draw("mass>>h1",cond);
```

```
Int_t binmax = h1->GetMaximumBin();
Double_t mass_suspect = h1->GetXaxis()->GetBinCenter(binmax);
```

```
TF1* f1=new TF1("f1","gaus(0)+expo(3)",0,200);
const Double_t width_suspect=1;
f1->SetParLimits(1,mass_suspect-
width_suspect,mass_suspect+width_suspect);
f1->SetParLimits(2,eps,width_suspect*2);
h1->Fit("f1","R","",5,8);
```

Implement

```
TF1* f_sig=new TF1("f_sig","gaus(0)",0,200);
TF1* f_bkg=new TF1("f_bkg","expo(0)",0,200);
TF1* f_data=new TF1("f_data","gaus(0)+expo(3)",0,200);
h1->Fit("f_sig","R","",5,8);
h1->Fit("f_bkg","R","",8,200);
Double_t* gaus_par=f_sig->GetParameters();
Double_t* expo_par=f_bkg->GetParameters();
f_data->SetParameters(gaus_par[0],gaus_par[1],gaus_par[2],expo_par[0],expo_par[1]);
h1->Fit("f_data","R");
```

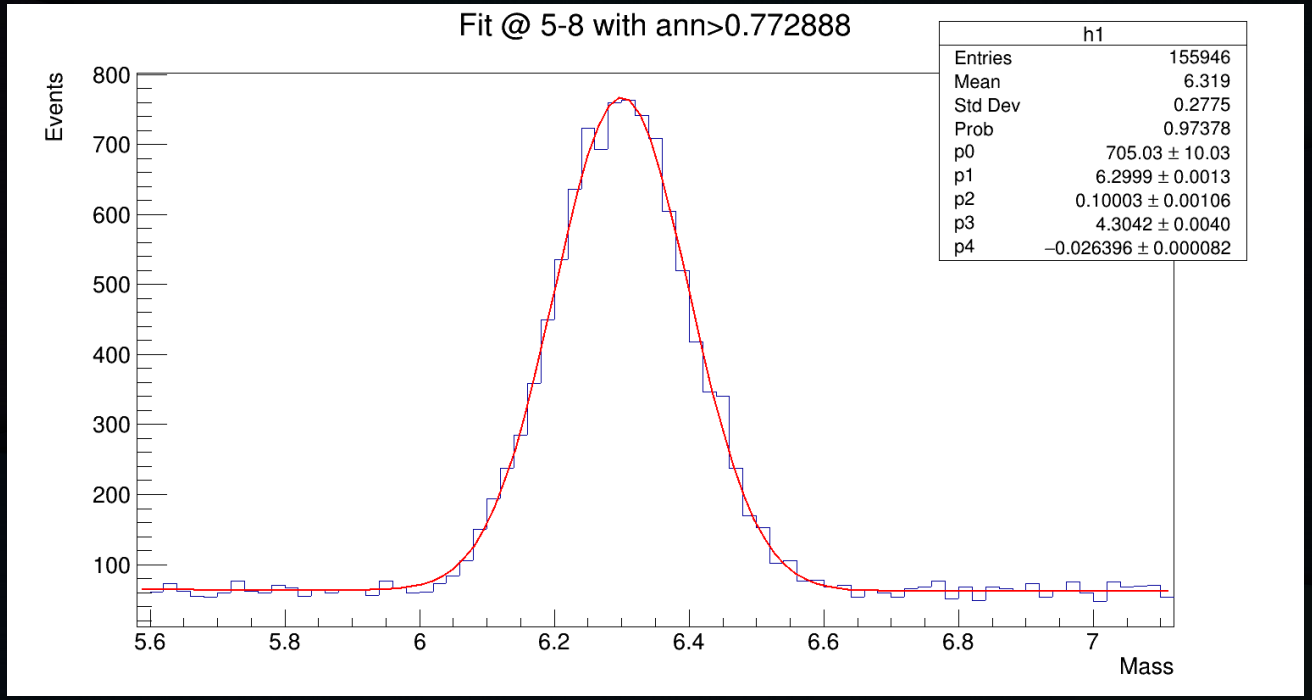
Result

name	mass	ann	Eff	Rej	Eff*Ref
Eff_90_mfit	0=>200	0.773	0.900	0.854	0.770
Eff_max_mfit	0=>200	0.754	0.917	0.842	0.772
Eff_90_5-8_fit	5=>8	0.773	0.900	0.870	0.783

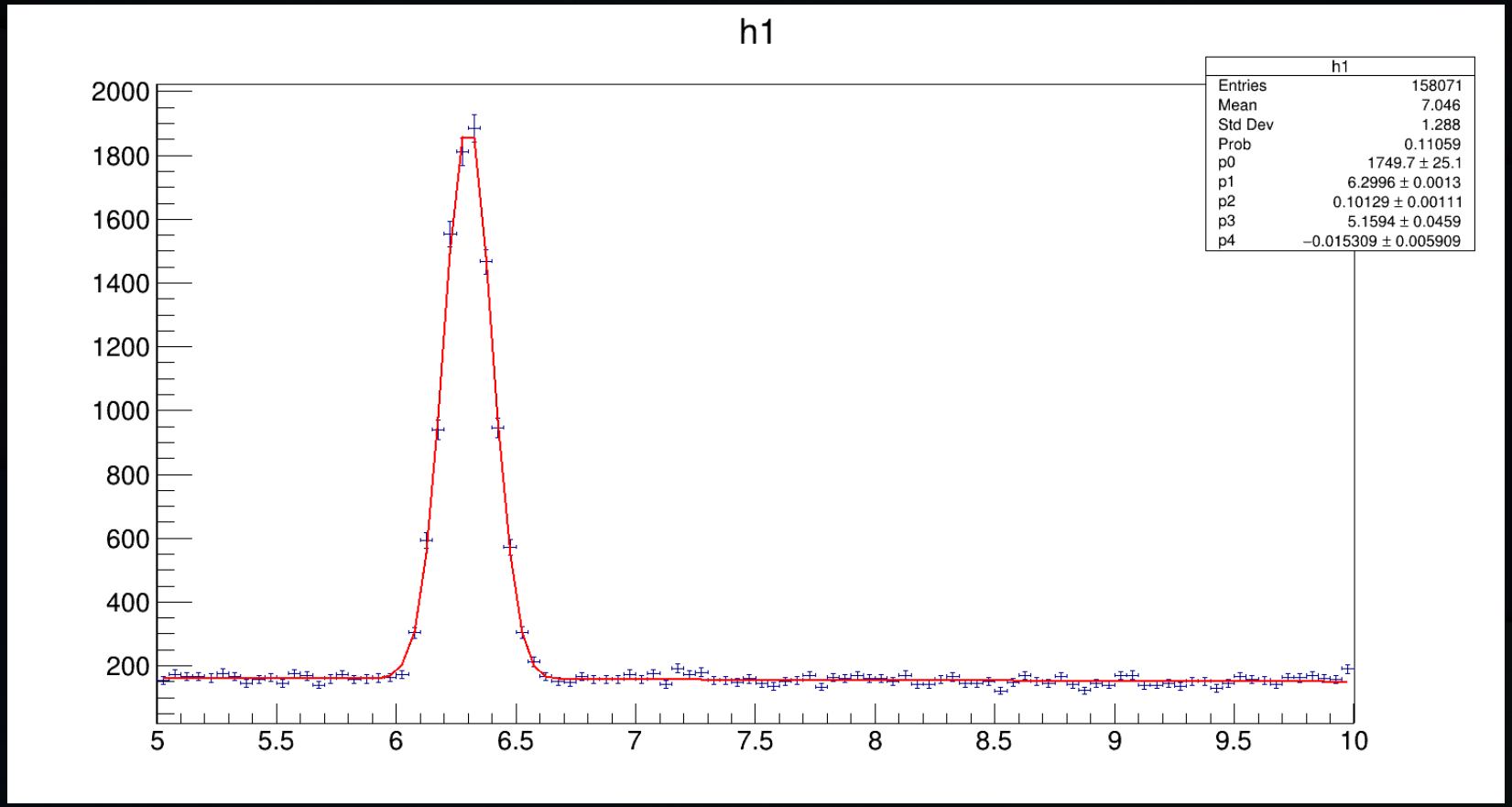
name	mass	miu	miuE	sigma	sigmaE
Eff_90_mfit	0=>200	6.29993	0.00125	0.10003	0.00106
Eff_max_mfit	0=>200	6.29955	0.00129	0.10014	0.00107
Eff_90_5-8_fit	5=>8	6.29986	0.00128	0.10022	0.00109

name	mass	sig	bkg	sig+bkg	data	percent
Eff_90_mfit	0=>200	8838	139473	148311	155946	0.951
Eff_max_mfit	0=>200	8995	153013	162008	169907	0.954
Eff_90_5-8_fit	5=>8	8857	9288	18145	18288	0.992

Fit: 5-8


Fit @ 5-8 with ann>0.772888
h1
Entries 155946
Mean 6.319
Std Dev 0.2775
Prob 0.97378
p0 705.03 ± 10.03
p1 6.2999 ± 0.0013
p2 0.10003 ± 0.00106
p3 4.3042 ± 0.0040
p4 −0.026396 ± 0.000082
Events
800
700
600
500
400
300
200
100
5.6
5.8
6
6.2
6.4
6.6
6.8
7
Mass

Fit from Wang Yang



More...

- Scl/Significance?
- Mass cut?
- Mass bin?

Thanks

- 杜玉琳， 高冠峰， 杨云帆， 翟明杰， 王洋， 王亚飞， 冉龙杰， 张明宇， 马震， 陈致中， 关家宝
- Zhang Yulei @ G3
- Wang Zhen @ G1

Backup

Fit: 5-8

[p0]*exp(-0.5*((x-[p1])/[p2])*((x-[p1])/[p2]))+exp([p3]+[p4]*x)

Fit range:5-8

FCN=141.978 FROM MIGRAD STATUS=CONVERGED 126 CALLS 127 TOTAL

EDM=5.18185e-09 STRATEGY= 1 ERROR MATRIX ACCURATE

EXT PARAMETER				STEP	FIRST
NO.	NAME	VALUE	ERROR	SIZE	DERIVATIVE
1	p0	7.05120e+02	1.01083e+01	4.72877e-02	-1.51964e-06
2	p1	6.29986e+00	1.28361e-03	7.48462e-06	3.26167e-02
3	p2	1.00224e-01	1.09015e-03	1.14183e-05	-1.29398e-03
4	p3	4.19331e+00	8.09119e-02	6.52378e-05	-5.50581e-03
5	p4	-1.03715e-02	1.22424e-02	9.90187e-06	-3.04138e-02

Fit: 0-200

Fit in different range

```
 FCN=7408.25 FROM MIGRAD    STATUS=CONVERGED     135 CALLS         136 TOTAL
                EDM=2.54767e-07    STRATEGY= 1      ERROR MATRIX ACCURATE
  EXT PARAMETER                                   STEP         FIRST
  NO.   NAME      VALUE            ERROR          SIZE      DERIVATIVE
   1  Constant     6.74387e+02   9.71343e+00   2.71674e-01   8.82734e-06
   2  Mean         6.30044e+00   1.23533e-03   5.18813e-05   2.05448e-01
   3  Sigma        1.28721e-01   1.38221e-03   4.32478e-05  -3.95988e-01
 FCN=8191.65 FROM MIGRAD    STATUS=CONVERGED      65 CALLS          66 TOTAL
                EDM=1.03548e-10    STRATEGY= 1      ERROR MATRIX ACCURATE
  EXT PARAMETER                                   STEP         FIRST
  NO.   NAME      VALUE            ERROR          SIZE      DERIVATIVE
   1  Constant     4.31254e+00   5.06959e-03   1.31973e-04   2.31124e-04
   2  Slope       -2.65188e-02   9.41179e-05   2.45008e-06  -1.42665e-01
 FCN=8608.59 FROM MIGRAD    STATUS=CONVERGED     112 CALLS         113 TOTAL
                EDM=6.00282e-07    STRATEGY= 1  ERROR MATRIX UNCERTAINTY   2.2 per cent
  EXT PARAMETER                                   STEP         FIRST
  NO.   NAME      VALUE            ERROR          SIZE      DERIVATIVE
   1  p0           7.05031e+02   1.00255e+01   1.24396e-02   3.95064e-05
   2  p1           6.29993e+00   1.25175e-03  -5.51977e-06   7.85272e-01
   3  p2           1.00030e-01   1.06282e-03  -1.62640e-06  -2.87766e-01
   4  p3           4.30419e+00   4.01590e-03  -3.30651e-06  -3.76906e-01
   5  p4          -2.63964e-02   8.21064e-05  -1.16461e-07  -1.73900e+01
```