Exercise/Lesson #4

Scientific Data Analysis Lab course

Alexis Pompili - UniBA

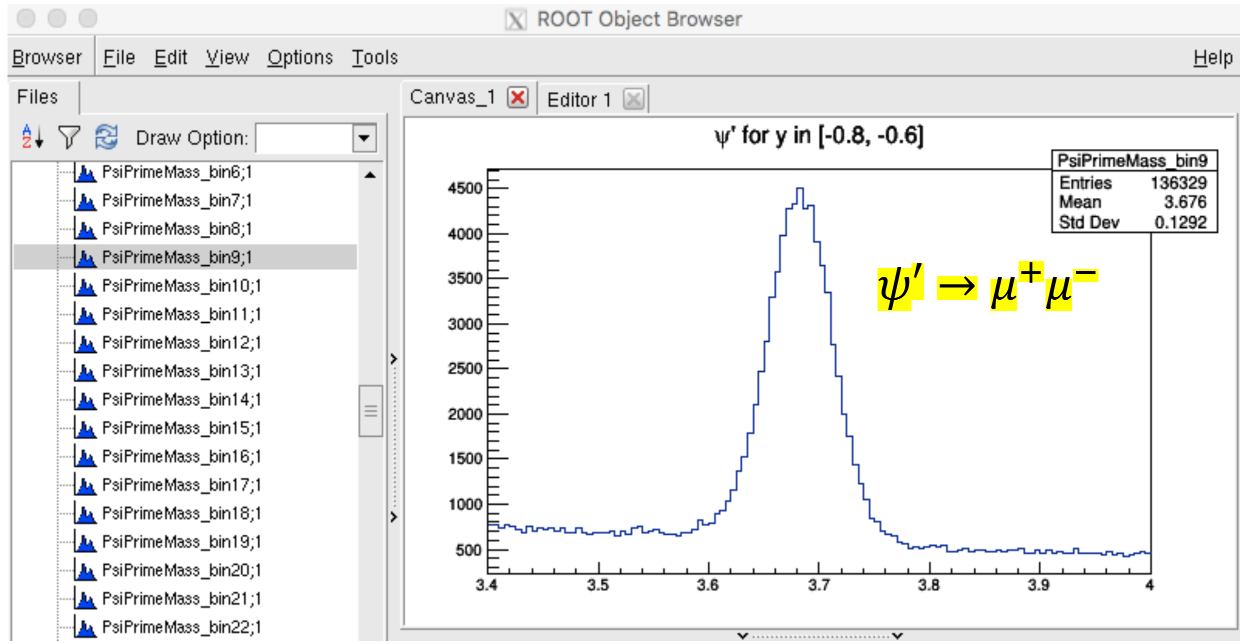
Preliminarily, let us visualize the distribution that we are going to interpolate:

```
[pompili@pompilic7 Esercitazione-4]$ root -l hlt_5_newSoftMuon_alsoInPsiPrimeWind.root
root [0]
Attaching file hlt_5_newSoftMuon_alsoInPsiPrimeWind.root as _file0...
(TFile *) 0x2c6dd10
root [1] TBrowser a;
```

ROOT file di input (data are from CMS)

This is the dimuon invariant mass associated to the $\psi' \rightarrow \mu^+\mu^-$ decay
The signal peak lies upon the combinatorial background.
The candidates entering the plot are characterized by a rapidity

$$y_{\psi'} \in [-0.8, -0.6]$$



$m(\mu^+\mu^-)$

To execute the fit you just need to do:

root [0] *.x psiprime_fit.C*

macro file (in C++)

Let us analize the macro :

In the new version we consider the root file `hlt_5_newSoftMuon_alsoInPsiPrimeWind.root`

```
//////////////////////////////////////
// run with root: .x psiPrime_fit.C
//////////////////////////////////////
#include <vector>

gROOT->Reset();
gROOT->Clear();

using namespace RooFit;

void psiPrime_fit() {
  gROOT->ForceStyle();
  gStyle->SetTitleOffset(1.4, "Y");
  gStyle->SetOptFit(1);

  TFile* f1 = TFile::Open("./Select/selected_histo.root","read");

  TH1F* hPsiPrime;
  hPsiPrime = (TH1F*) f1->Get("PsiPrimeMass_bin8");

  TCanvas *myC = new TCanvas("myC","PsiPrimeMassPlot", 700, 700);

  Double_t xMin = hPsiPrime->GetXaxis()->GetXmin();
  Double_t xMax = hPsiPrime->GetXaxis()->GetXmax();
  Int_t nBins = hPsiPrime->GetNbinsX();

  RooRealVar xVar("xVar", "m(#mu^{+}#mu^{-}) [GeV/c^{2}]", xMin, xMax);
  xVar.setBins(nBins);

  RooDataHist* MuMuHist = new RooDataHist("#mu#mu_hist", hPsiPrime->GetTitle(), RooArgSet(xVar), Import(*hPsiPrime,kFALSE));
```

Needed to use the RooFit workspace

In the new version we consider bin9 !

Apre rootupla esterna e ne prende l'istogramma d'interesse

Definisce variabile reale (massa invariante μμ) di RooFit: $m_{\mu\mu}$

Definisce istogramma di RooFit associato alla variabile reale precedentemente introdotta

```
RooRealVar mG("mean", "mean", 3.7, 3.67, 3.73);
RooRealVar sigma1("#sigma_{1}", "sigma1", 0.02, 0.001, 0.1);

RooGaussian sigPDF("sigPDF", "Signal", xVar, mG, sigma1);
```

modello per il segnale : PDF gaussiana

$$G_{SIG}(m_{\mu\mu})$$

```
RooRealVar c1("c_{1}", "c1", -0.1 ,-10, 10);
RooRealVar c2("c_{2}", "c2", -0.1 ,-10, 10);
RooChebychev bkgPDF("bkgPDF", "bkgPDF", xVar, RooArgSet(c1,c2));
```

un modello per il fondo (assunto lineare): polinomiale di ord.1 (con polinomi di Chebyshev) :

$$C_{BKG}(m_{\mu\mu})$$

```
RooRealVar nSig("nSig", "Number of signal candidates ", 2e+5, 1., 1e+6);
RooRealVar nBkg("nBkg", "Bkg component", 120e+3, 1., 1e+6);

RooAddPdf* totalPDF = new RooAddPdf("totalPDF", "totalPDF", RooArgList(sigPDF, bkgPDF), RooArgList(nSig, nBkg));
```

modello complessivo per segnale + fondo : combinazione lineare di segnale e fondo

$$n_{SIG} \cdot G_{SIG}(m_{\mu\mu}) + n_{BKG} \cdot C_{BKG}(m_{\mu\mu})$$

```
totalPDF->fitTo(*MuMuHist, Extended(kTRUE));
```

Qui viene eseguito il fit della distribuzione *binnata* della variabile

Per capire esattamente cosa significa *extended likelihood function in the case of binned data*, vedere G.Cowan 6.10 (e 6.9) !

A schermo si ottengono informazioni sul fit:

```
[pompili@cmssusy esercitazione-5]$ root -l
root [0] .x psiPrime_fit.C
```

thus obtaining …

```
RooFit v3.56 -- Developed by Wouter Verkerke and David Kirkby
                Copyright (C) 2000-2013 NIKHEF, University of California & Stanford University
                All rights reserved, please read http://roofit.sourceforge.net/license.txt
 ... ... ...
 **********
 **   13 **MIGRAD       3000          1
 **********
 FIRST CALL TO USER FUNCTION AT NEW START POINT, WITH IFLAG=4.
 START MIGRAD MINIMIZATION.  STRATEGY  1.  CONVERGENCE WHEN EDM .LT. 1.00e-03
 ... ... ...
 MIGRAD MINIMIZATION HAS CONVERGED.
 MIGRAD WILL VERIFY CONVERGENCE AND ERROR MATRIX.
 COVARIANCE MATRIX CALCULATED SUCCESSFULLY
 FCN=-2.0678e+06 FROM MIGRAD    STATUS=CONVERGED     198 CALLS         199 TOTAL
                     EDM=3.88053e-05    STRATEGY= 1      ERROR MATRIX ACCURATE
  EXT PARAMETER                                   STEP         FIRST
  NO.   NAME      VALUE            ERROR          SIZE      DERIVATIVE
   1  #sigma_{1}   3.56225e-02   2.17316e-04   3.67024e-03   6.06691e-01
   2  c_{1}       -1.88551e-01   5.22002e-03   5.14039e-04  -2.37797e+00
   3  c_{2}       -1.92706e-02   6.94904e-03   5.52338e-04   2.92534e+00
   4  mean         3.68091e+00   2.03170e-04   8.64724e-03  -5.38998e-01
   5  nBkg         1.11119e+05   4.70821e+02   1.16253e-03   3.01109e+00
   6  nSig         6.43078e+04   4.18322e+02   1.23552e-03  -9.41342e-02
                               ERR DEF= 0.5
 EXTERNAL ERROR MATRIX.    NDIM=  25    NPAR=  6    ERR DEF=0.5
  4.723e-08  4.252e-08  6.647e-07 -1.472e-09 -4.920e-02  4.927e-02
  4.252e-08  2.725e-05 -3.769e-07 -1.218e-07 -8.827e-02  8.832e-02
  6.647e-07 -3.769e-07  4.829e-05  3.331e-08 -1.580e+00  1.580e+00
 -1.472e-09 -1.218e-07  3.331e-08  4.128e-08  1.535e-03 -1.540e-03
 -4.920e-02 -8.827e-02 -1.580e+00  1.535e-03  2.217e+05 -1.106e+05
  4.927e-02  8.832e-02  1.580e+00 -1.540e-03 -1.106e+05  1.750e+05
 PARAMETER  CORRELATION COEFFICIENTS
       NO.  GLOBAL      1      2      3      4      5      6
        1  0.59608   1.000  0.037  0.440 -0.033 -0.481  0.542
        2  0.12945   0.037  1.000 -0.010 -0.115 -0.036  0.040
        3  0.59940   0.440 -0.010  1.000  0.024 -0.483  0.544
        4  0.12587  -0.033 -0.115  0.024  1.000  0.016 -0.018
        5  0.62392  -0.481 -0.036 -0.483  0.016  1.000 -0.561
        6  0.68387   0.542  0.040  0.544 -0.018 -0.561  1.000
```

risoluzione → $\hat{\sigma} \cong (35.62 \pm 0.22) MeV$

coefficienti di Cebyshev

massa → $\hat{m} \cong (3680.91 \pm 0.20) MeV$

 # candidati di fondo

 # candidati di segnale → $\hat{N}_{sig} \cong 64308 \pm 418$

```
**********
**   18 **HESSE        3000
**********
COVARIANCE MATRIX CALCULATED SUCCESSFULLY
FCN=-2.0678e+06 FROM HESSE     STATUS=OK           40 CALLS          239 TOTAL
                    EDM=3.89377e-05    STRATEGY= 1      ERROR MATRIX ACCURATE
 EXT PARAMETER                                   INTERNAL      INTERNAL
 NO.   NAME      VALUE            ERROR          STEP SIZE       VALUE
  1  #sigma_{1}   3.56225e-02   2.17932e-04   1.46809e-04  -3.05276e-01
  2  c_{1}       -1.88551e-01   5.22000e-03   2.05616e-05  -1.88562e-02
  3  c_{2}       -1.92706e-02   6.95669e-03   2.20935e-05  -1.92706e-03
  4  mean         3.68091e+00   2.03183e-04   3.45890e-04  -6.89538e-01
  5  nBkg         1.11119e+05   4.71806e+02   4.65010e-05  -8.91100e-01
  6  nSig         6.43078e+04   4.19277e+02   4.94208e-05  -1.05802e+00
                               ERR DEF= 0.5
EXTERNAL ERROR MATRIX.    NDIM=  25    NPAR=  6    ERR DEF=0.5
 4.749e-08  4.178e-08  6.707e-07 -1.597e-09 -4.978e-02  4.978e-02
 4.178e-08  2.725e-05 -4.689e-07 -1.215e-07 -8.568e-02  8.568e-02
 6.707e-07 -4.689e-07  4.840e-05  3.327e-08 -1.589e+00  1.589e+00
-1.597e-09 -1.215e-07  3.327e-08  4.128e-08  1.556e-03 -1.556e-03
-4.978e-02 -8.568e-02 -1.589e+00  1.556e-03  2.226e+05 -1.115e+05
 4.978e-02  8.568e-02  1.589e+00 -1.556e-03 -1.115e+05  1.758e+05
PARAMETER  CORRELATION COEFFICIENTS
      NO.  GLOBAL      1      2      3      4      5      6
       1  0.59912   1.000  0.037  0.442 -0.036 -0.484  0.545
       2  0.12939   0.037  1.000 -0.013 -0.115 -0.035  0.039
       3  0.60057   0.442 -0.013  1.000  0.024 -0.484  0.545
       4  0.12644  -0.036 -0.115  0.024  1.000  0.016 -0.018
       5  0.62596  -0.484 -0.035 -0.484  0.016  1.000 -0.564
       6  0.68564   0.545  0.039  0.545 -0.018 -0.564  1.000
[#1] INFO:Minization -- RooMinuit::optimizeConst: deactivating const optimization
[#1] INFO:Plotting -- RooAbsPdf::plotOn(totalPDF) directly selected PDF components: (sigPDF)
[#1] INFO:Plotting -- RooAbsPdf::plotOn(totalPDF) indirectly selected PDF components: ()
[#1] INFO:Plotting -- RooAbsPdf::plotOn(totalPDF) directly selected PDF components: (bkgPDF)
[#1] INFO:Plotting -- RooAbsPdf::plotOn(totalPDF) indirectly selected PDF components: ()
Info in <TCanvas::Print>: png file ./Plots/PsiPrimeMassFit_alt.png has been created
```

Viene ricalcolata la matrice di covarianza.

I valori centrali delle stime dei parametri sono gli stessi ma viene raffinata la stima delle incertezze!

Il resto del codice serve per rappresentare l'istogramma e il risultato dell'interpolazione! (vedi slide seguente per il risultato)

```
RooPlot* xframe = xVar.frame();
xframe->SetTitle( hPsiPrime->GetTitle() );
xframe->SetYTitle("Candidates / 10 MeV/c^{2}");

MuMuHist->plotOn(xframe);
totalPDF->plotOn(xframe);

totalPDF->plotOn(xframe, Components(RooArgSet(sigPDF)), LineColor(kRed));

totalPDF->plotOn(xframe, Components(RooArgSet(bkgPDF)), LineColor(kGreen), LineStyle(kDashed) );

totalPDF->paramOn(xframe, Parameters(RooArgSet(mG,sigma1,nSig)), Layout(0.52,0.99,0.9)); //box con stime parametri

myC->cd();
xframe->Draw();

myC->SaveAs("./Plots/PsiPrimeMassFit_alt.png");

}
```

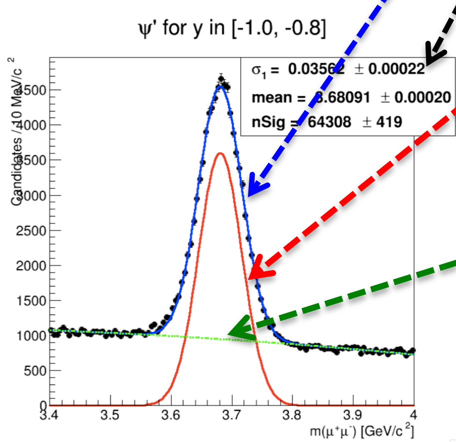
Definisce un *frame* a partire dalla variabile d'interesse

Rappresenta l'istogramma sul *frame*

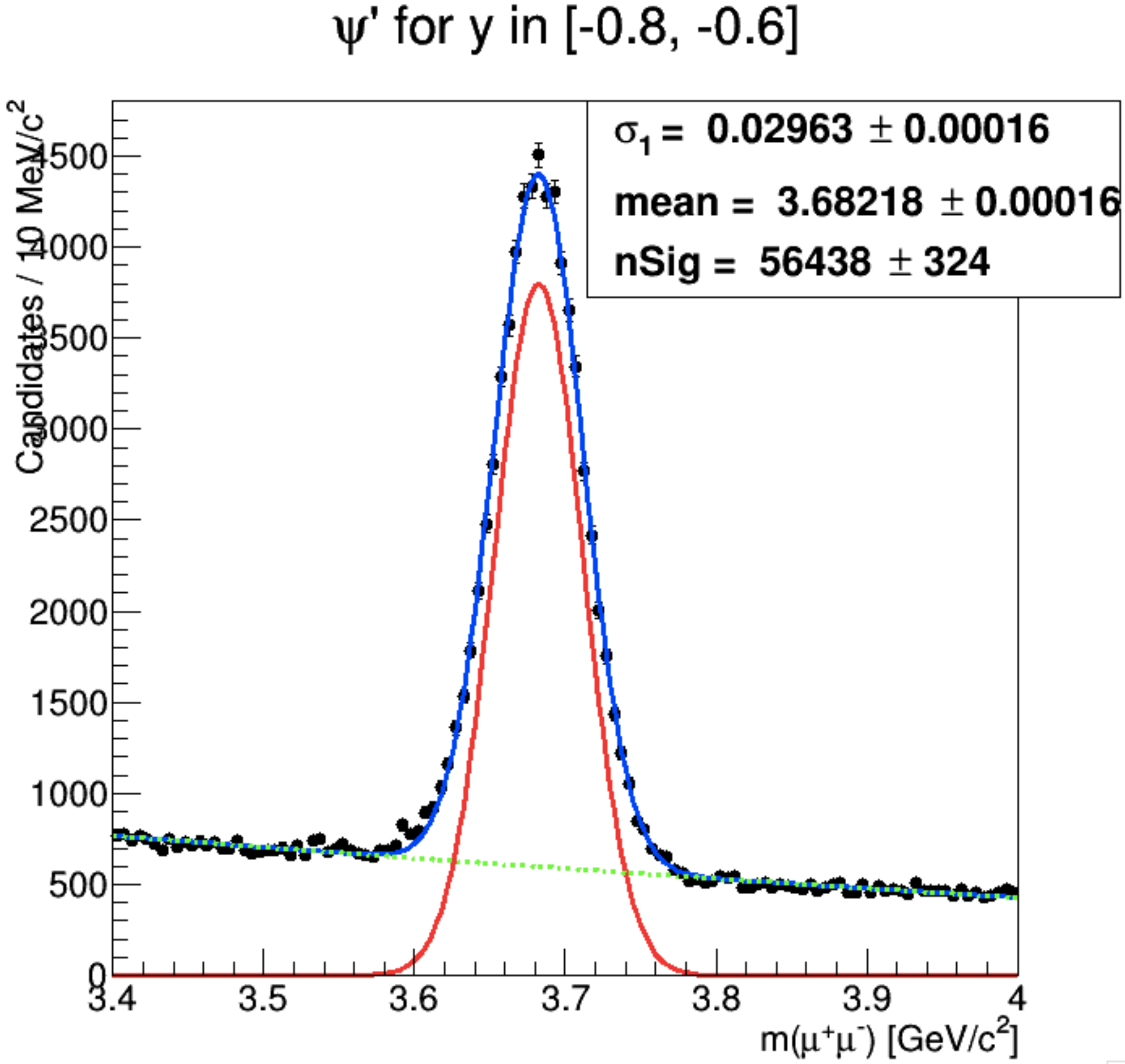
Rappresenta la funzione di fit sul *frame*

Sovrappone sul *frame* la sola componente del segnale

Sovrappone sul *frame* la sola componente del fondo



Finally, we get the plot:



Another plot with a fit in the adjacent rapidity bin: