A Likelihood Fitting Technique for Nuclear Physics Data Analysis

Priyashree Roy

Nuclear Physics Seminar

10/24/2014

Outline

Motivation

Good fitting techniques are essential to extract maximum and most accurate information from experiments.

Motivation

Good fitting techniques are essential to extract maximum and most accurate information from experiments.

"I can prove it or disprove it! What do you want me to do?"

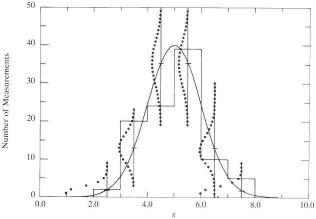
- Method of least squares (χ^2 fits)- Most common fitting technique for data with huge statistics.
- Event-based maximum likelihood fitting (ML fits)- very useful for extracting maximum information from low statistics datasets.

The Method of Least Squares

Bin the data and plot histograms. Based on the assumption that each datapoint is Gaussian distributed, minimize

$$\chi^2 = \sum\left(\frac{y_i^{experiment} - y_i^{theory}(\alpha_j)}{\sigma_i}\right)^2$$

and get the best estimate of the fit parameters α_j.



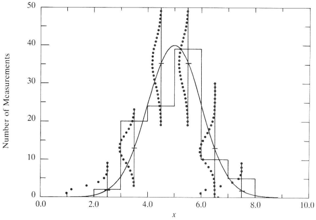
P.R. Bevington, 'Data Reduction and Analysis for the Physical Sciences' (McGraw-Hill, 1992)

The Method of Least Squares

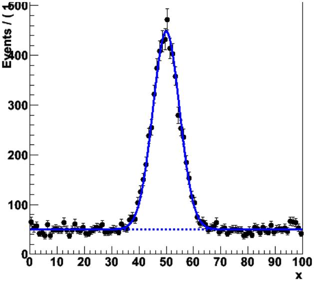
Bin the data and plot histograms. Based on the assumption that each datapoint is Gaussian distributed, minimize

$$\chi^2 = \sum\left(\frac{y_i^{experiment} - y_i^{theory}(\alpha_j)}{\sigma_i}\right)^2$$

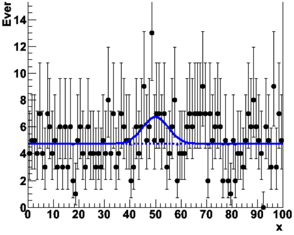
and get the best estimate of the fit parameters α_j.



High statistics - easy to fit



Bad fit - datapoints not Gaussian distributed.

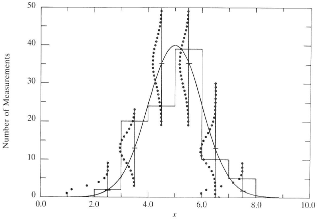


The Method of Least Squares

Bin the data and plot histograms. Based on the assumption that each datapoint is Gaussian distributed, minimize

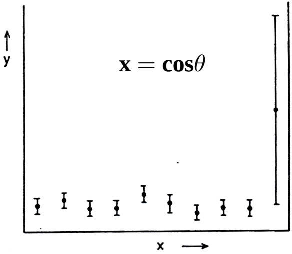
$$\chi^2 = \sum\left(\frac{y_i^{experiment} - y_i^{theory}(\alpha_j)}{\sigma_i}\right)^2$$

and get the best estimate of the fit parameters α_j.



For example: fitting an angular distribution-

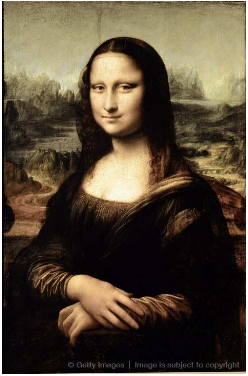
$y^{theory}(a, b) = a + bcos^2(\theta)$
Calculate χ^2 and minimize it to get the best values of a and b.



L. Lyons, 'Statistics for Nuclear and Particle Physicists' (CUP, 1986)

The Method of Least Squares

Disadvantage - Each bin should have enough statistics so that it is Gaussian distributed. Need coarse binning for data with low statistics. This can lead to loss of information.



The Event-based Maximum Likelihood Method

We calculate the probability density for observing **each event** as a function of the fit parameter, $P(y_i(\alpha))$, and construct a likelihood L -
$L = \prod_i P(y_i(\alpha))$ over all events.
Maximize L as a function of the fit parameter α to find the best value for α.

The Event-based Maximum Likelihood Method

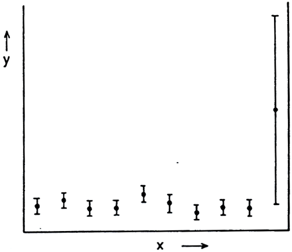
We calculate the probability density for observing **each event** as a function of the fit parameter, $P(y_i(\alpha))$, and construct a likelihood L - $L = \prod_i P(y_i(\alpha))$ over all events.

Maximize L as a function of the fit parameter α to find the best value for α.

Example : Fitting the angular distribution (slide 2 data) -

Probability density, $y_i(\frac{b}{a}) = \frac{1}{2[1+(\frac{b}{3a})]}[1 + (\frac{b}{a})cos^2\theta_i]$ for the i^{th} event.

Likelihood, $L(\frac{b}{a}) = \prod_i \frac{1}{2[1+(\frac{b}{3a})]}[1 + (\frac{b}{a})cos^2\theta_i]$



Event-based - no loss of information due to binning !

L. Lyons, 'Statistics for Nuclear and Particle Physicists' (CUP, 1986)

ML or χ^2 - which one to use ?

χ^2 fitting

Pros-

- fastest and easiest.
- Gives goodness-of-fit indication.

Cons -

- Makes (incorrect) Gaussian error assumption on low statistics bins.
- Binning problem - misses information for feature size $<$ bin size.

ML fitting

Pros-

- Most robust.
- No Gaussian error assumption.
- No loss of information.

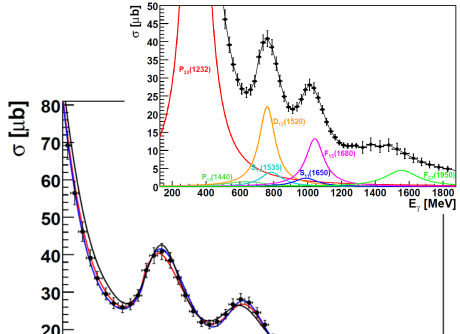
Cons -

- No goodness-of-fit indication. Needs Monte-Carlo studies for verification.
- Computationally expensive for large N.

Outline

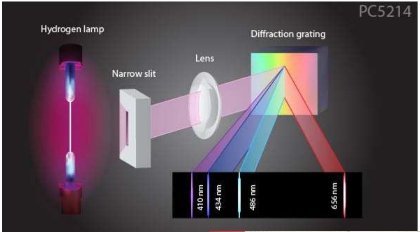
Why are Polarization Observables Important?

Volker Crede, Nuclear Physics Seminar Talk, Sept 26, 2014



$\gamma p \to p\pi^0$

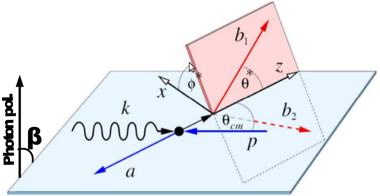
Atomic Spectrum of Hydrogen



CLAS@JLab

ELSA
MAMI
GRAAL
SPring-8
...

Spin Observables for $\vec{\gamma}\vec{p} \rightarrow p\pi^+\pi^-$ at CLAS, JLAB



5 independent kinematic variables needed -
E_γ, $\phi^*_{\pi^+}$, $\cos(\theta^*_{\pi^+})$, $\cos(\theta^{c.m.}_p)$, $m_{\pi^+\pi^-}$

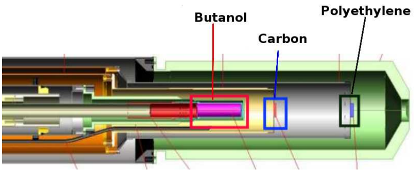
For linearly polarized photon beam and transversely polarized protons, reaction rate-

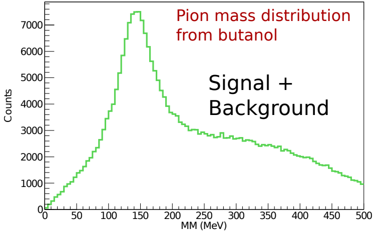
W. Roberts *et al.*, Phys. Rev. C **71**, 055201 (2005)

$$\sigma = \sigma_0\{(1 + \Lambda_x \mathbf{P_x} + \Lambda_y \mathbf{P_y}) \\ + \delta_l[sin2\beta(\mathbf{I^s} + \Lambda_x \mathbf{P^s_x} + \Lambda_y \mathbf{P^s_y}) \\ + cos2\beta(\mathbf{I^c} + \Lambda_x \mathbf{P^c_x} + \Lambda_y \mathbf{P^c_y})]\}$$

δ_l : deg. of beam polarization, Λ : deg. of target polarization

Event-Based Qfactor Method for Signal-Bkg Separation

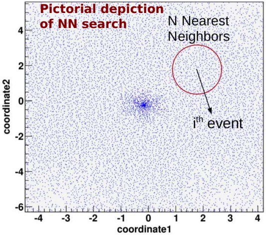


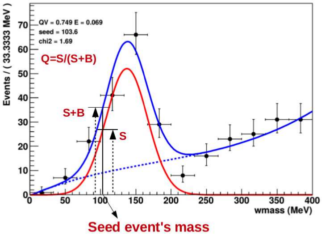


- **Polarized Target: Butanol**
  has free polarized p + unpolarized bound p and n.
  Reactions from bound nucleons contribute to the background.
- C and CH_2 to study background.

Signal-background separation -
Assign an event-based dilution factor or "Qfactor" to each event which shows the chance that it came from the signal distribution.

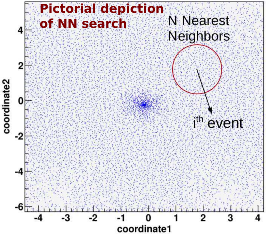
Event-Based Qfactor Method with Likelihood Fits

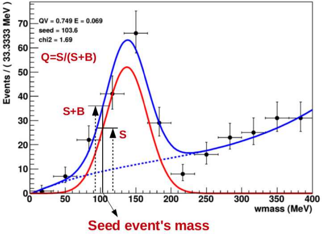




- **A multivariate analysis -** For each event ("seed event"), find N nearest neighbors in 4-D kinematic phase space ($E_\gamma, \theta^*, \phi^*, cos(\theta_p)^{c.m.}$). Plot mass distribution of the $N+1$ events and fit.

Event-Based Qfactor Method with Likelihood Fits



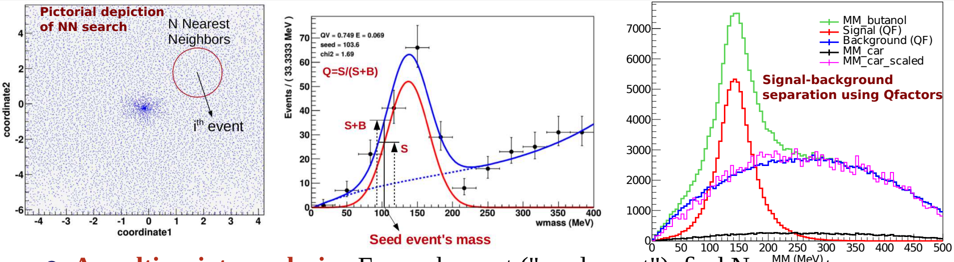


- **A multivariate analysis -** For each event ("seed event"), find N nearest neighbors in 4-D kinematic phase space $(E_\gamma, \theta^*, \phi^*, cos(\theta_p)^{c.m.})$. Plot mass distribution of the $N+1$ events and fit.
- Since N is small (300), use ML method to fit the mass distribution.

$$L = \prod_i [f^{Signal}(m_i, \alpha) + f^{Bkg}(m_i, \beta)]$$

$$\mathbf{Q}_{\mathbf{seed-event}} = \frac{f^{Signal}(m_0, \alpha^{best})}{[f^{Signal}(m_0, \alpha^{best}) + f^{Bkg}(m_0, \beta^{best})]},$$ m_0- seed event's mass.

Event-Based Qfactor Method with Likelihood Fits

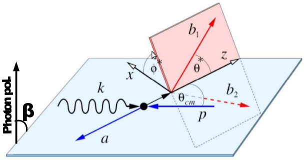


- **A multivariate analysis -** For each event ("seed event"), find N nearest neighbors in 4-D kinematic phase space $(E_\gamma, \theta^*, \phi^*, cos(\theta_p)^{c.m.})$. Plot mass distribution of the $N+1$ events and fit.
- Since N is small (300), use ML method to fit the mass distribution.

  $$L = \prod_i [f^{Signal}(m_i, \alpha) + f^{Bkg}(m_i, \beta)]$$

  $$\mathbf{Q}_{\mathbf{seed-event}} = \frac{f^{Signal}(m_0, \alpha^{best})}{[f^{Signal}(m_0, \alpha^{best}) + f^{Bkg}(m_0, \beta^{best})]}$$, m_0- seed event's mass.
- **Computation time reasonably minimized-** fits 10,000 events in 30 min.

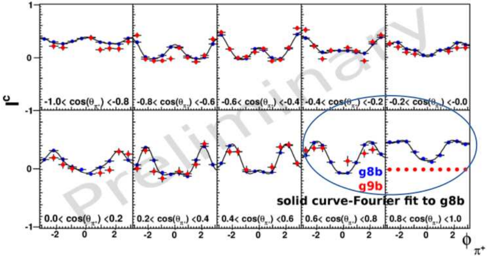
Polarization Observables: χ^2 vs. Unbinned ML Fit

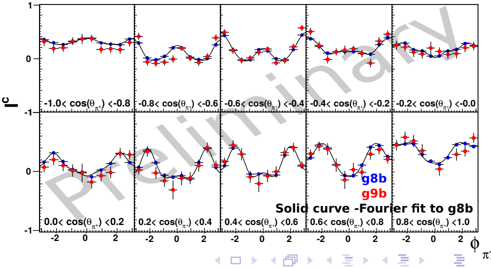


- **For linearly pol. beam and unpolarized target,**

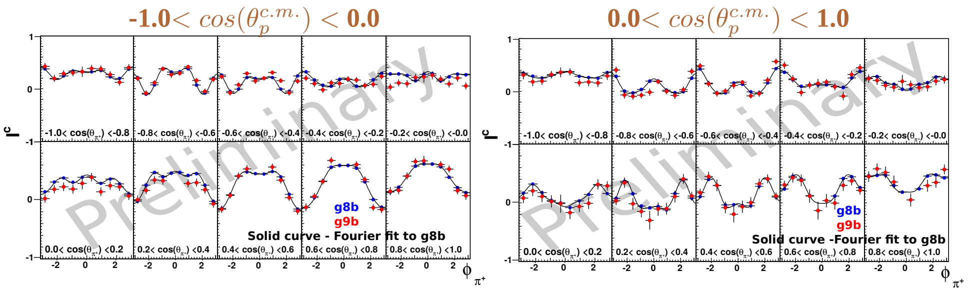
$$\frac{\sigma_\parallel - \sigma_\perp}{\sigma_\parallel + \sigma_\perp} = \frac{2\delta_l^{av}[\mathbf{I^c} cos(2\phi_{lab}) + \mathbf{I^s} sin(2\phi_{lab})]}{2+(\delta_l^\parallel - \delta_l^\perp)[[\mathbf{I^c} cos(2\phi_{lab}) + \mathbf{I^s} sin(2\phi_{lab})]}$$

- Unbinned ML fit to the angular distribution allowed extraction of $I^{s,c}$ **in 4D** $(E_\gamma, \phi^*_{\pi^+}$**,** $\cos(\theta^*_{\pi^+})$**,** $\cos(\theta_p^{c.m.}))$. Not possible with χ^2 fits.





Observable I^c in $\vec{\gamma}\vec{p} \to p\pi^+\pi^-$, E_γ : 1.1 – 1.2 GeV



FROST g9b run (preliminary) **C. Hanretty *et al.*, CLAS-g8b run, to be published**

Fourier fit to g8b

4-dim. phase space: (E_γ, $\phi^*_{\pi^+}$, $\cos(\theta^*_{\pi^+})$, $\cos(\theta_p^{c.m.})$)

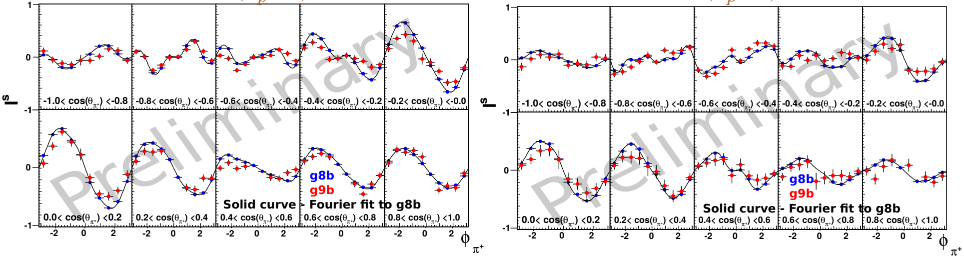
$$I = I_0\{\delta_l[I^s \sin(2\beta) + I^c \cos(2\beta)]\}$$

β : angle between beam pol. and reaction plane.

Good agreement of experimental data.

Observable I^s in $\vec{\gamma}\vec{p} \rightarrow p\pi^+\pi^-$, E_γ : 1.1 – 1.2 GeV

-1.0$< cos(\theta_p^{c.m.}) <$ **0.0** **0.0**$< cos(\theta_p^{c.m.}) <$ **1.0**



FROST g9b run (preliminary) **C. Hanretty *et al.*, CLAS-g8b run, to be published**

—— Fourier fit to g8b

4-dim. phase space: (E_γ, $\phi^*_{\pi^+}$, $\cos(\theta^*_{\pi^+})$, $\cos(\theta_p^{c.m.})$)

$$I = I_0\{\delta_l[I^s \sin(2\beta) + I^c \cos(2\beta)]\}$$

β : angle between beam pol. and reaction plane.

Good agreement of experimental data.

Outline

Outlook

- **Unbinned ML technique -** very useful to optimize fit parameters of any general distribution with low statistics.
  Assumption - the model or fit function has the right form. Works very well for the extraction of polarization observables since the fit function is well-known.
- **Bayesian analysis goes one step further - compare models quantitatively !**
  Example: New particle in the mass spectrum ?
  Model 1 - No new particle, fit function : $f^{Background}$
  Model 2 - A new particle, fit function : $f^{NarrowGaussian} + f^{Background}$
  Construct "evidence ratios" using Bayes' theorem to compare the models.
- **Learn more about Bayesian analysis in the next talk by Raditya !**

Outlook

- **Unbinned ML technique -** very useful to optimize fit parameters of any general distribution with low statistics.

  Assumption - the model or fit function has the right form. Works very well for the extraction of polarization observables since the fit function is well-known.

- **Bayesian analysis goes one step further - compare models quantitatively.**

  Example: New particle in the mass spectrum ?

  Model 1 - No new particle, fit function : $f^{Background}$

  Model 2 - A new particle, fit function : $f^{NarrowGaussian} + f^{Background}$

  Construct "evidence ratios" using Bayes' theorem to compare the models. D.G. Ireland *et al.*, arXiv:0709.3154v2 [hep-ph] 12 Dec 2007

- **Learn more about Bayesian analysis in the next talk by Raditya !**

Thank You