Kinematic Fitting

Frank Thanh Cao

University of Connecticut

2017

Motivation

We must confront the fact that 4-vectors coming from detectors are not perfect and we may be able to do better.

This presentation will outline kinematic fitting as an answer to this and some results will be presented when applying to EG6 data.

Background

Background

Let $\vec{\boldsymbol{\eta}}$ be a vector of n-measured variables. Then the true vector of the n-variables, $\vec{\boldsymbol{y}}$, will have an associated error vector of n-variables, $\vec{\boldsymbol{\varepsilon}}$. They are related simply by:

$$\vec{\boldsymbol{y}} = \vec{\boldsymbol{\eta}} + \vec{\boldsymbol{\varepsilon}}$$

If there are, say m, unmeasured variables too, then they can be put in a vector, $\vec{\boldsymbol{x}}$. The two vectors, $\vec{\boldsymbol{x}}$ and $\vec{\boldsymbol{y}}$, are then related by r constraint equations, indexed by k:

$$f_k\left(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}}\right) = 0$$

Suppose $\vec{\boldsymbol{x_0}}$ and $\vec{\boldsymbol{y_0}}$ are our best guess (measurements) of the vectors $\vec{\boldsymbol{x}}$ and $\vec{\boldsymbol{y}}$, respectively. Then Taylor expanding to first order each $f_k\left(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}}\right)$ about $\vec{\boldsymbol{x_0}}$ and $\vec{\boldsymbol{y_0}}$ gives:

$$
\begin{aligned}
f_k\left(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}}\right) \approx f_k\left(\vec{\boldsymbol{x_0}}, \vec{\boldsymbol{y_0}}\right) &+ \sum_{i=0}^{m} \left(\frac{\partial f_k}{\partial x_i}\right)\Bigg|_{\left(\vec{\boldsymbol{x_0}}, \vec{\boldsymbol{y_0}}\right)} \left(x_i - x_{0\,i}\right) \\
&+ \sum_{j=0}^{n} \left(\frac{\partial f_k}{\partial y_j}\right)\Bigg|_{\left(\vec{\boldsymbol{x_0}}, \vec{\boldsymbol{y_0}}\right)} \left(y_j - y_{0\,j}\right)
\end{aligned}
\tag{1}
$$

where x_i, y_j are the ith and jth components of $\vec{\boldsymbol{x}}$, $\vec{\boldsymbol{y}}$ and $x_{0\,i}$, $y_{0\,j}$ are the ith and jth components of $\vec{\boldsymbol{x_0}}$, $\vec{\boldsymbol{y_0}}$, respectively.

For convenience, let's introduce

$$\begin{aligned} a_{ij} &:= \left(\frac{\partial f_i}{\partial x_j}\right)\Bigg|_{(\vec{x_0},\vec{y_0})} \quad , \\ b_{ij} &:= \left(\frac{\partial f_i}{\partial y_j}\right)\Bigg|_{(\vec{x_0},\vec{y_0})} \quad , \\ c_i &:= f_i(\vec{x_0},\vec{y_0}) \quad , \end{aligned} \tag{2}$$

and

$$\begin{aligned} \vec{\xi} &:= \vec{x} - \vec{x_0} \quad , \\ \vec{\delta} &:= \vec{y} - \vec{y_0} \quad . \end{aligned}$$

Then, since $f_k\left(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}}\right) \equiv 0 \ \ \forall k$, **Eq. 1** can be written in matrix form as:

$$\vec{\mathbf{0}} \equiv A\vec{\boldsymbol{\xi}} + B\vec{\boldsymbol{\delta}} + \vec{\boldsymbol{c}} \tag{3}$$

where A and B are $(r \times n)$ and $(r \times m)$ matrices with components a_{ij} and b_{ij}, respectively, as defined by **Eqn.'s 2**.

Now, if we have a really good understanding of the correlations between the measured values, then we can construct a covariance matrix, C_η:

$$C_\eta = \vec{\boldsymbol{\sigma}}_\eta^T \rho_\eta \vec{\boldsymbol{\sigma}}_\eta$$

where $\vec{\boldsymbol{\sigma}}_\eta$ is a vector of the resolution errors of η and ρ_η is a symmetric correlation matrix whose components, $\rho_{ij} \in \{-1, 1\}$, house pairwise correlations coefficients, between η_i and η_j.

Note, if there are no correlations, then the ρ is the unit matrix and so the covariance matrix is just a diagonal matrix of the variances of η. In this case, the χ^2 becomes the recognizable:

$$\chi^2 = \sum_{i=0}^{m} \frac{(y_i - y_{0\,i})^2}{\sigma_i^2} = \sum_{i=0}^{m} \frac{\delta_i^2}{\sigma_i^2}$$

This can be generalized, to account for correlations, with:

$$\chi^2 = \vec{\boldsymbol{\delta}}^T C_\eta^{-1} \vec{\boldsymbol{\delta}} \tag{4}$$

Now that we have a χ^2 to work with (to minimize, that is), we can introduce a Lagrangian, $\mathcal{L}$, with Lagrange multipliers $\vec{\boldsymbol{\mu}}$ such that:

$$\mathcal{L} = \vec{\boldsymbol{\delta}}^T C_\eta^{-1} \vec{\boldsymbol{\delta}} + \vec{\boldsymbol{\mu}}^T \left(A \vec{\boldsymbol{\xi}} + B \vec{\boldsymbol{\delta}} + \vec{\boldsymbol{c}} \right) \tag{5}$$

is to be minimized.
Minimization conditions are then:

$$\frac{\partial \mathcal{L}}{\partial \vec{\boldsymbol{\xi}}} = \vec{\boldsymbol{\mu}}^T A \equiv \vec{\mathbf{0}} \tag{6}$$

$$\frac{\partial \mathcal{L}}{\partial \vec{\boldsymbol{\mu}}} = A \vec{\boldsymbol{\xi}} + B \vec{\boldsymbol{\delta}} + \vec{\boldsymbol{c}} \equiv \vec{\mathbf{0}} \tag{7}$$

$$\frac{\partial \mathcal{L}}{\partial \vec{\boldsymbol{\delta}}} = C_\eta^{-1} \vec{\boldsymbol{\delta}} + B^T \vec{\boldsymbol{\mu}} \equiv \vec{\mathbf{0}} \quad . \tag{8}$$

Solving for such $\vec{\boldsymbol{\xi}}, \vec{\boldsymbol{\mu}}, \vec{\boldsymbol{\delta}}$ that satisfy these conditions result in:

$$\begin{aligned}
\vec{\boldsymbol{\xi}} &= -\left(A^T C_B A\right)^{-1} A^T C_B \vec{\boldsymbol{c}} \\
\vec{\boldsymbol{\mu}} &= C_B \left(A\vec{\boldsymbol{\xi}} + \vec{\boldsymbol{c}}\right) \\
\vec{\boldsymbol{\delta}} &= -C_\eta B^T \vec{\boldsymbol{\mu}}
\end{aligned} \quad . \tag{9}$$

where C_B is conveniently defined as

$$C_B := \left(B C_\eta B^T\right)^{-1}$$

With these vectors that satisfies the minimization condition, we can finally form our new fitted vectors $\vec{\boldsymbol{x}}$ and $\vec{\boldsymbol{y}}$:

$$\begin{aligned} \vec{\boldsymbol{x}} &= \vec{\boldsymbol{x_0}} + \vec{\boldsymbol{\xi}} \\ \vec{\boldsymbol{y}} &= \vec{\boldsymbol{y_0}} + \vec{\boldsymbol{\delta}} \end{aligned} \tag{10}$$

with new covariance matrices:

$$\begin{aligned} C_x &= \left(\frac{\partial \vec{\boldsymbol{x}}}{\partial \vec{\boldsymbol{\eta}}}\right) C_\eta \left(\frac{\partial \vec{\boldsymbol{x}}}{\partial \vec{\boldsymbol{\eta}}}\right)^T \\ &= \left(A^T C_B A\right)^{-1} \\ C_y &= \left(\frac{\partial \vec{\boldsymbol{y}}}{\partial \vec{\boldsymbol{\eta}}}\right) C_\eta \left(\frac{\partial \vec{\boldsymbol{y}}}{\partial \vec{\boldsymbol{\eta}}}\right)^T \\ &= C_\eta - C_\eta \left(B^T C_B B\right) C_\eta + C_\eta \left(B^T C_B \left[A C_x A^T\right] C_B B\right) C_\eta \end{aligned} .$$

Quality of Fit

Confidence Levels

To check on the quality of the fit, we look to two sets of distributions: The **Confidence levels** and the **Pull distributions**. Since $\chi^2 := \vec{\delta}^T C_\eta^{-1} \vec{\delta}$ will produce an χ^2 distribution for *ndf* degrees of freedom, let's define the confidence level, CL as:

$$CL := \int_{x=\chi^2}^{\infty} f(x, ndf)\ \mathrm{d}x$$

- **Characteristics**
  - If there is no background in the fit, the distribution is uniform and flat.
  - In the presence of background, there will be a sharp rise as $CL \to 0$.

Cutting out the sharp rise as $CL \to 0$ will cut out the much of the background while keeping much of the signal intact.

Pull Distributions

To see if the covariance matrix is correctly taking into account all pairwise correlations between the variables, we look to the pull distributions. Let's define $\vec{z}$ to house the pulls, z_i, defined as

$$z_i := \frac{y_i - \eta_i}{\sqrt{\sigma_{y_i}^2 - \sigma_{\eta_i}^2}}$$

- <u>**Characteristics**</u>
  Since these are normalized differences, the distributions should be normally distributed with
  - mean 0 and
  - width 1.

Applying to Data

Using Clean Sample

To test if this is a viable option, we apply it to a previously studied dataset that should have mostly coherent π^0 events.

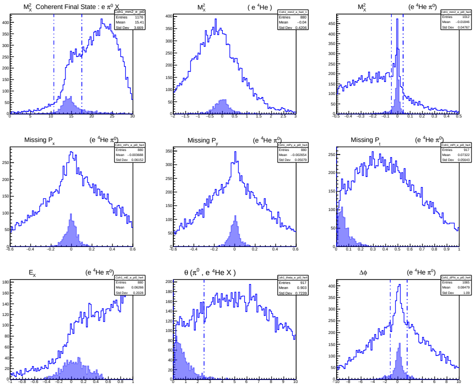


Figure: All events passing the blue line cuts are taken as clean sample (highlighted in blue).

Setting up Inputs

We first want to set up our input vectors and matrices for the fit:

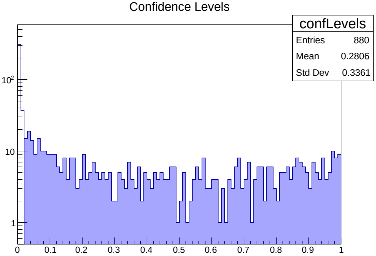
$$\vec{y_0} = \vec{\eta} = \begin{bmatrix} p_e \\ \theta_e \\ \phi_e \\ p_{^4\mathrm{He}} \\ \theta_{^4\mathrm{He}} \\ \phi_{^4\mathrm{He}} \\ p_{\gamma 1} \\ \theta_{\gamma 1} \\ \phi_{\gamma 1} \\ p_{\gamma 2} \\ \theta_{\gamma 2} \\ \phi_{\gamma 2} \end{bmatrix} \quad \vec{c} = \begin{bmatrix} (P_{\mathrm{init}} - P_{\mathrm{fin}})_x \\ (P_{\mathrm{init}} - P_{\mathrm{fin}})_y \\ (P_{\mathrm{init}} - P_{\mathrm{fin}})_z \\ (P_{\mathrm{init}} - P_{\mathrm{fin}})_E \end{bmatrix} \quad B = \begin{bmatrix} \frac{\partial c_1}{\partial \eta_1} & \cdots & \frac{\partial c_1}{\partial \eta_{12}} \\ \vdots & \ddots & \vdots \\ \frac{\partial c_4}{\partial \eta_1} & \cdots & \frac{\partial c_4}{\partial \eta_{12}} \end{bmatrix} \quad . \tag{11}$$

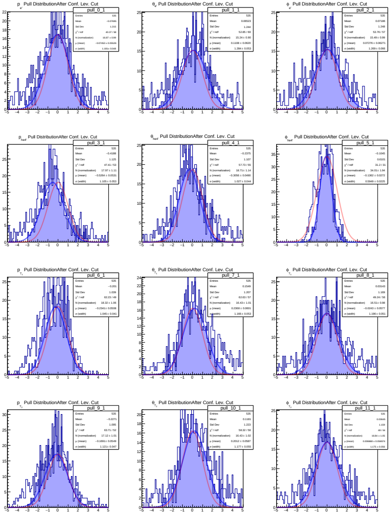
Now, we set up the covariance matrix. Let's start with a simple, uncorrelated matrix:

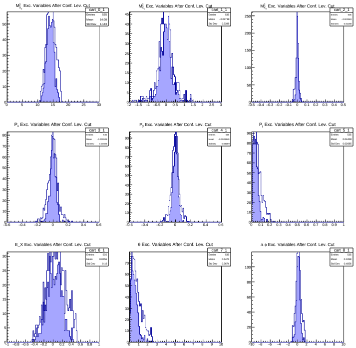
$$C_\eta = \begin{bmatrix} \sigma^2_{\eta_1} & 0 & \dots & \dots & \dots & 0 \\ 0 & \sigma^2_{\eta_2} & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \sigma^2_{\eta_{11}} & 0 \\ 0 & \dots & \dots & \dots & 0 & \sigma^2_{\eta_{12}} \end{bmatrix} \tag{12}$$

where the σ's are the widths extracted from previous Monte-Carlo studies.

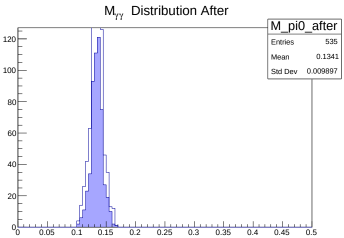
Results for Clean Sample







(a)



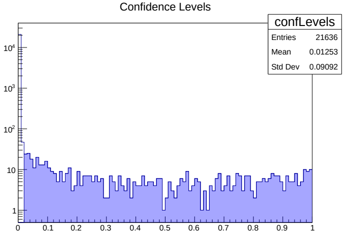
(b)

Figure: Exclusivity Variables (a) and Invariant Mass Distribution (b)

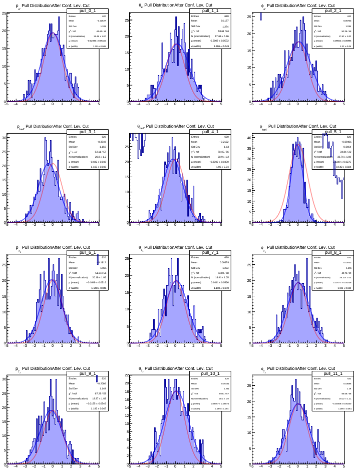
Checking Full Dataset

It looks like the confidence level and pull distributions look okay on the clean sample. Let's see how they look on the full dataset.

Results for Full Dataset

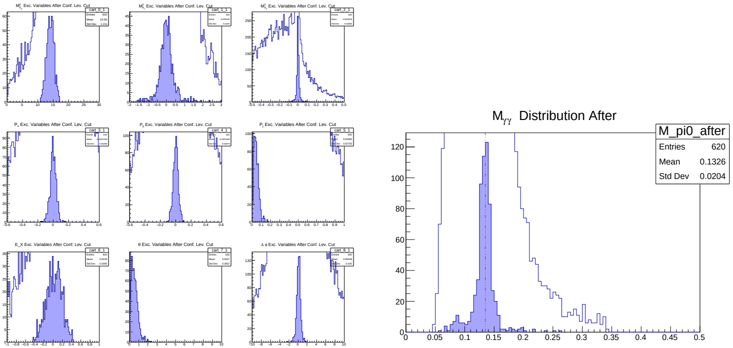


(a)



(b)

Figure: Confidence Levels (a) and Pull Distributions (b)



(a) (b)

Figure: Exclusivity Variables (a) and Invariant Mass Distribution (b)

Conclusion and Outlook

On first order the kinematic fitting looks good.

However, there still needs to be some tweaking done to get as many coherent π^0 events as possible:

- Adding/removing variables
- Adding/removing constraints
- Applying kin. fitting to DVCS events
- Working on a full covariance matrix
- Calibrating each detector individually

Questions?