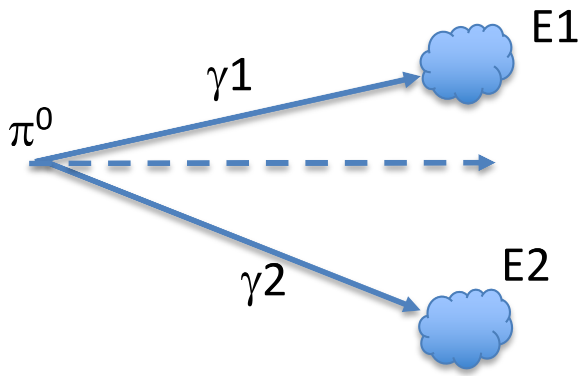
Event selection - Kinematic fit

1) Determine or improve knowledge of kinematic quantities
2) Define a test statistics to select the event

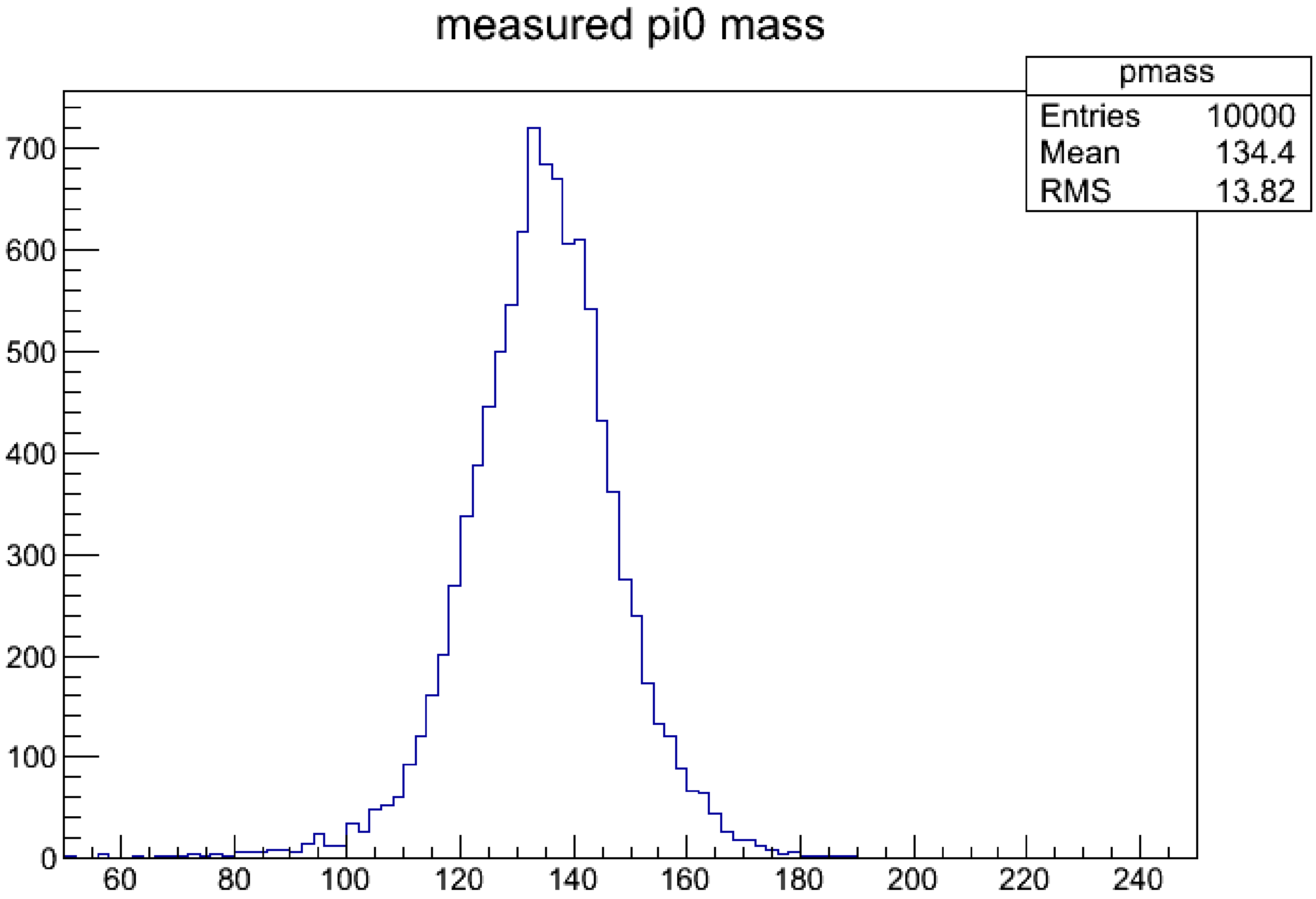
Impose kinematical constraints on measured variables

Example: Decay of Neutral Pion into Two Photons

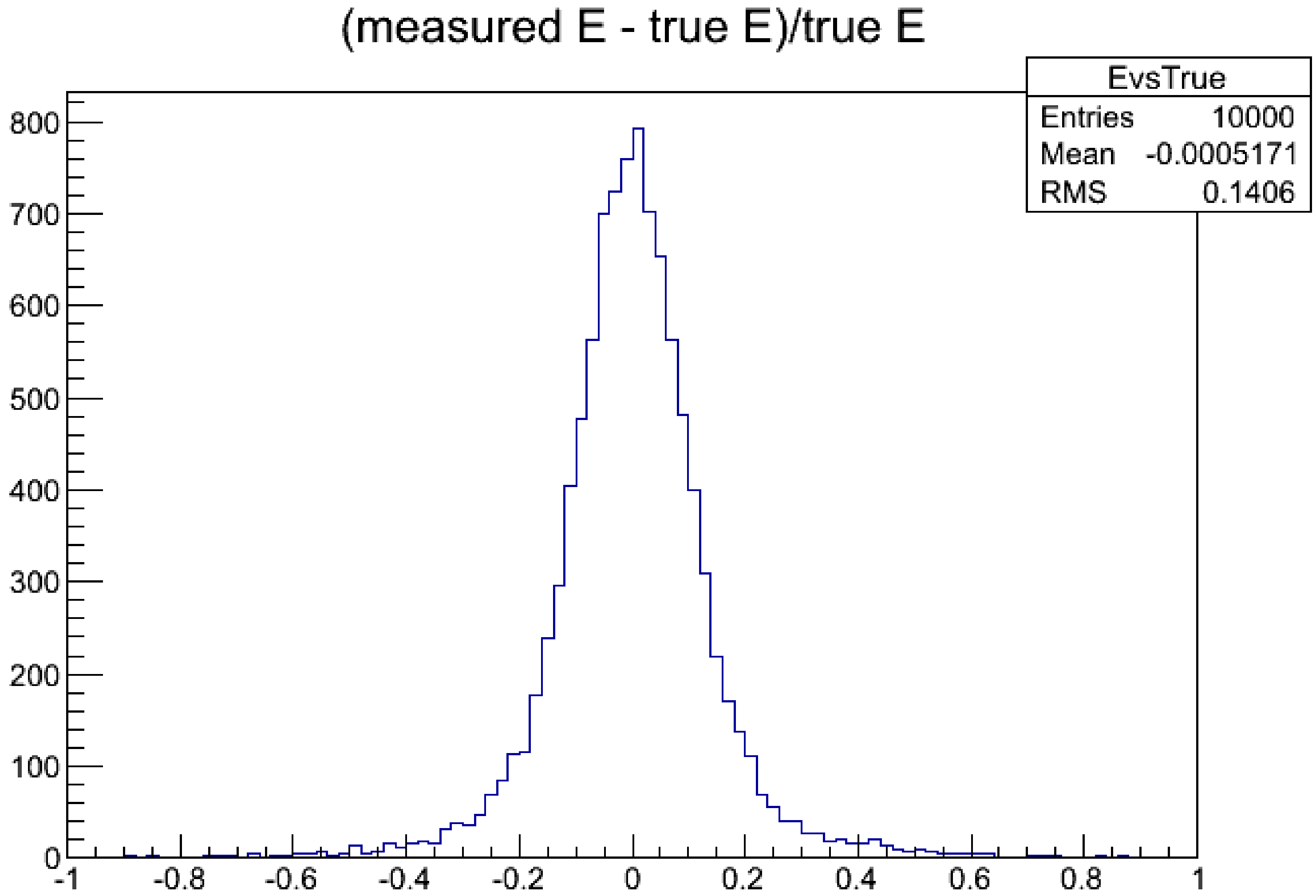
- $\pi^0 \rightarrow \gamma\gamma$
- both photons detected
- assume photon directions are known precisely
- energies have relative uncertainty $\sigma_E/E = 5\%/\sqrt{E}$
- for simplicity: look at 500 MeV/c π^0 moving in z-direction
- but, will not assume 500 MeV/c nor z-direction for pion



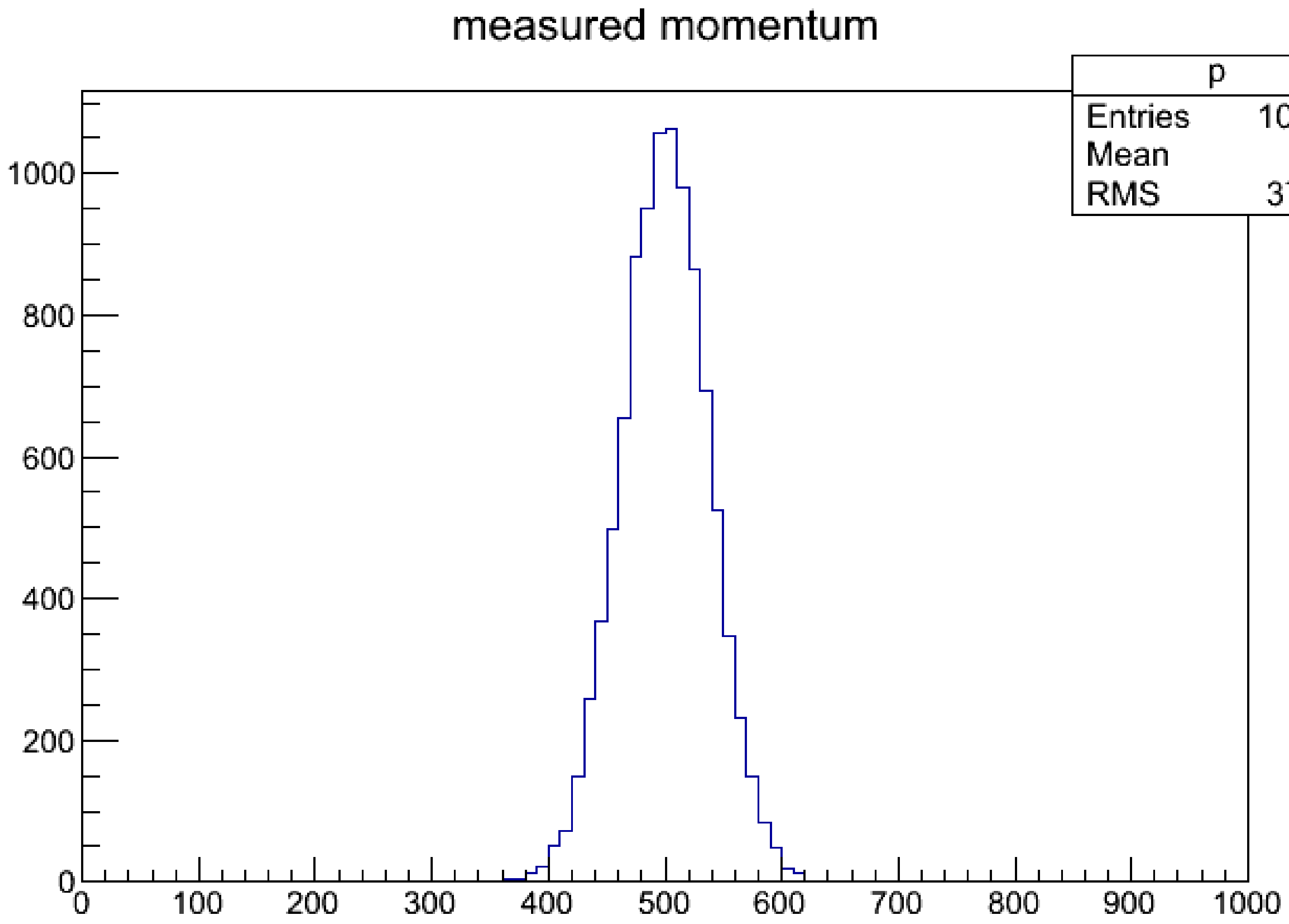
two-photon invariant mass



relative error on single photon energy



two-photon measured momentum



What is the problem?

- want to improve resolution
- assume that photons came from π^0, so mass is known
- can we use this information?
  - could adjust one photon (why just one?)
  - could scale them both (high energy γ: better E measurement)
- could minimize

$$\chi^2 = \left(\frac{E_{1,\mathrm{fit}} - E_{1,\mathrm{meas}}}{\sigma_1}\right)^2 + \left(\frac{E_{2,\mathrm{fit}} - E_{2,\mathrm{meas}}}{\sigma_2}\right)^2$$

- but minimum is clear: $E_{\mathrm{fit}} = E_{\mathrm{meas}}$ (something is missing!)
- must introduce constraint: $(k_1 + k_2)^2 = m_\pi^2$ gives

$$2E_1 E_2(1 - \cos\theta) = m_\pi^2$$

- Problem: minimize χ^2 while simultaneously satisfying constraint

Minimization Strategy

- multi-variable minimization with constraints: Lagrange multipliers
- instead of minimizing over two variables, minimize over three $E_{1,\mathrm{fit}}$, $E_{2,\mathrm{fit}}$, and λ

$$\chi^2 = \left(\frac{E_{1,\mathrm{fit}} - E_{1,\mathrm{meas}}}{\sigma_1}\right)^2 + \left(\frac{E_{2,\mathrm{fit}} - E_{2,\mathrm{meas}}}{\sigma_2}\right)^2$$

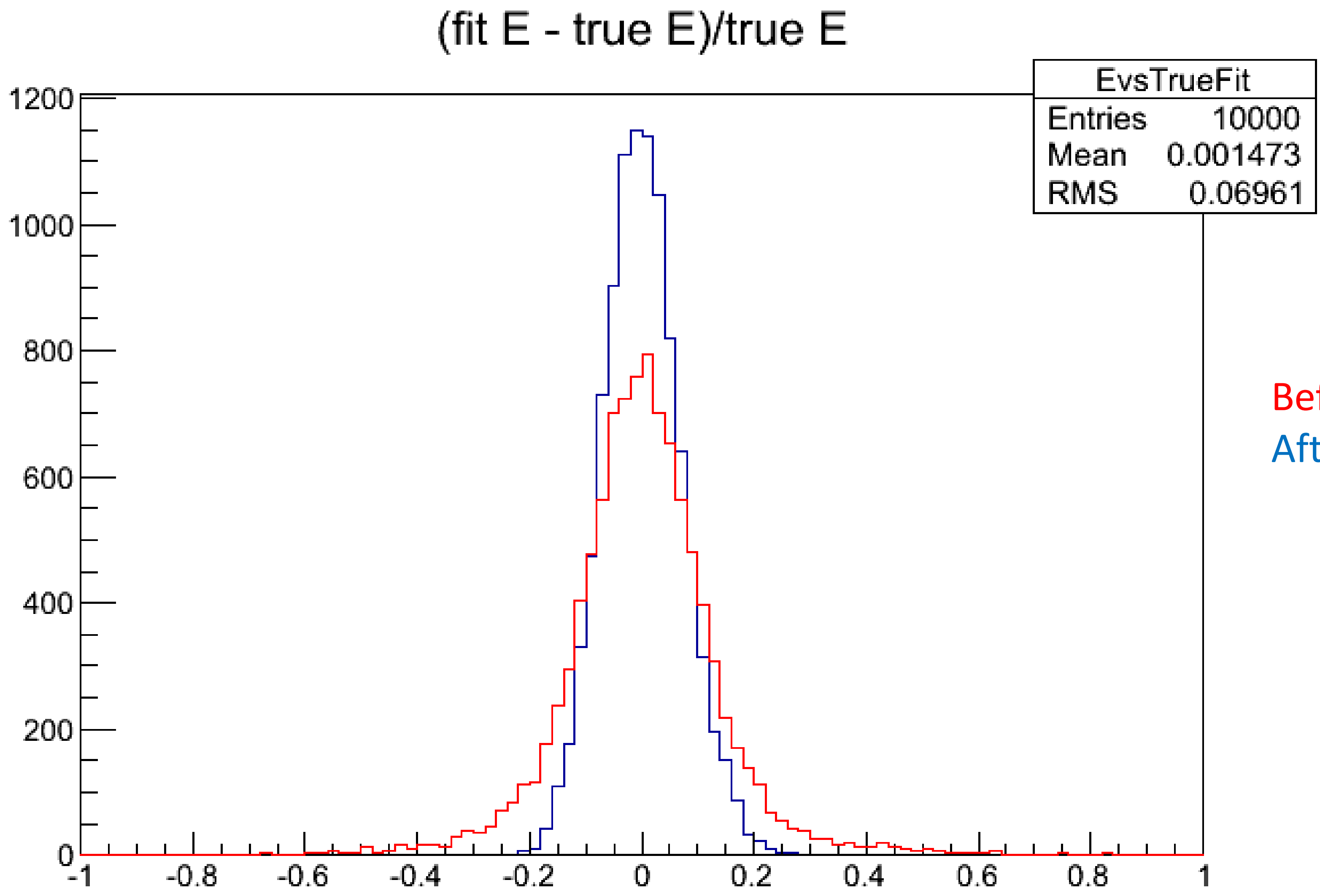
$$+2\lambda\left[2E_{1,\mathrm{fit}}E_{2,\mathrm{fit}}(1-\cos\theta) - m_\pi^2\right]$$

Comments:

1 d.o.f.

X^2 evaluated for each single event

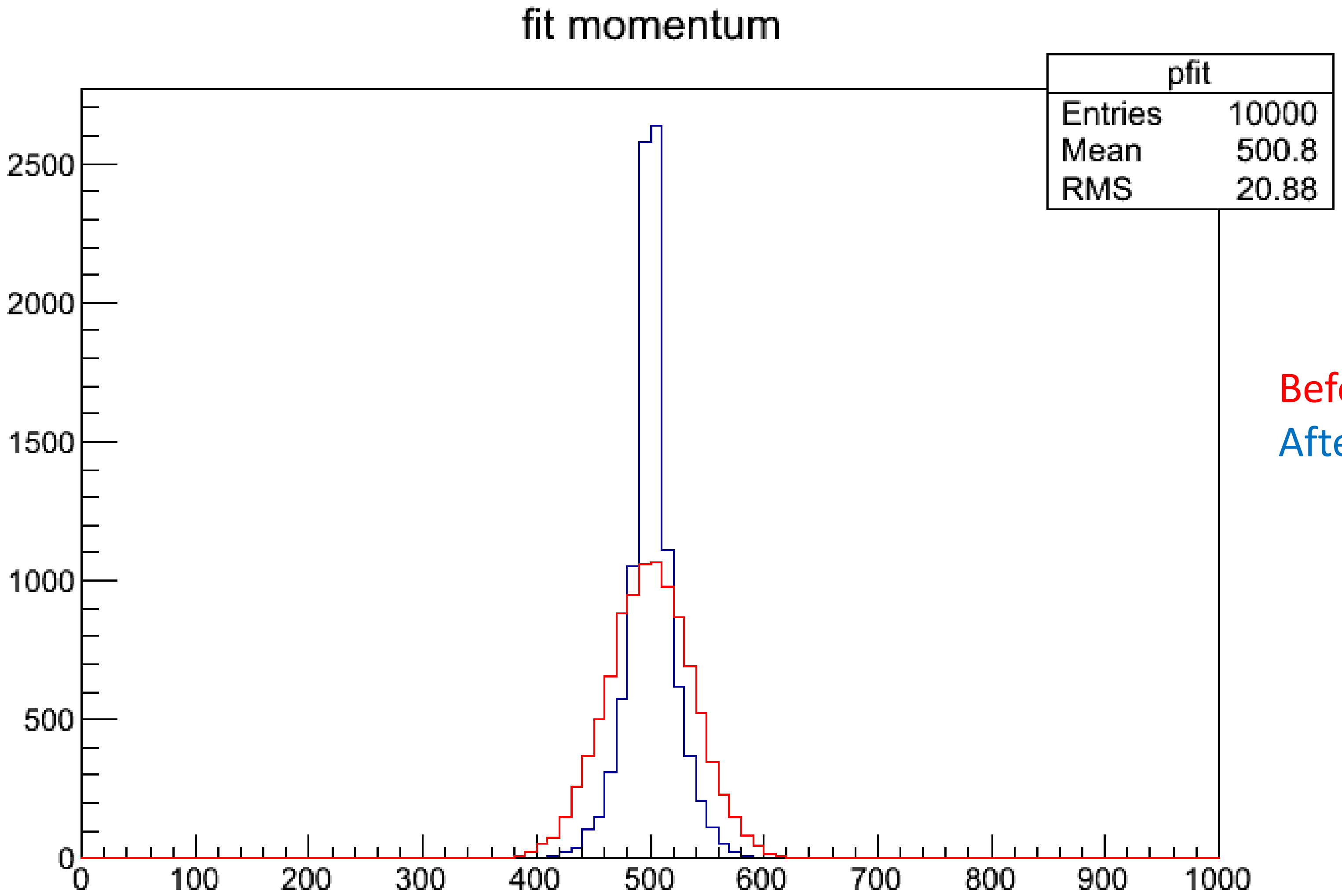
fit relative error on single photon energy



Before fit

After fit

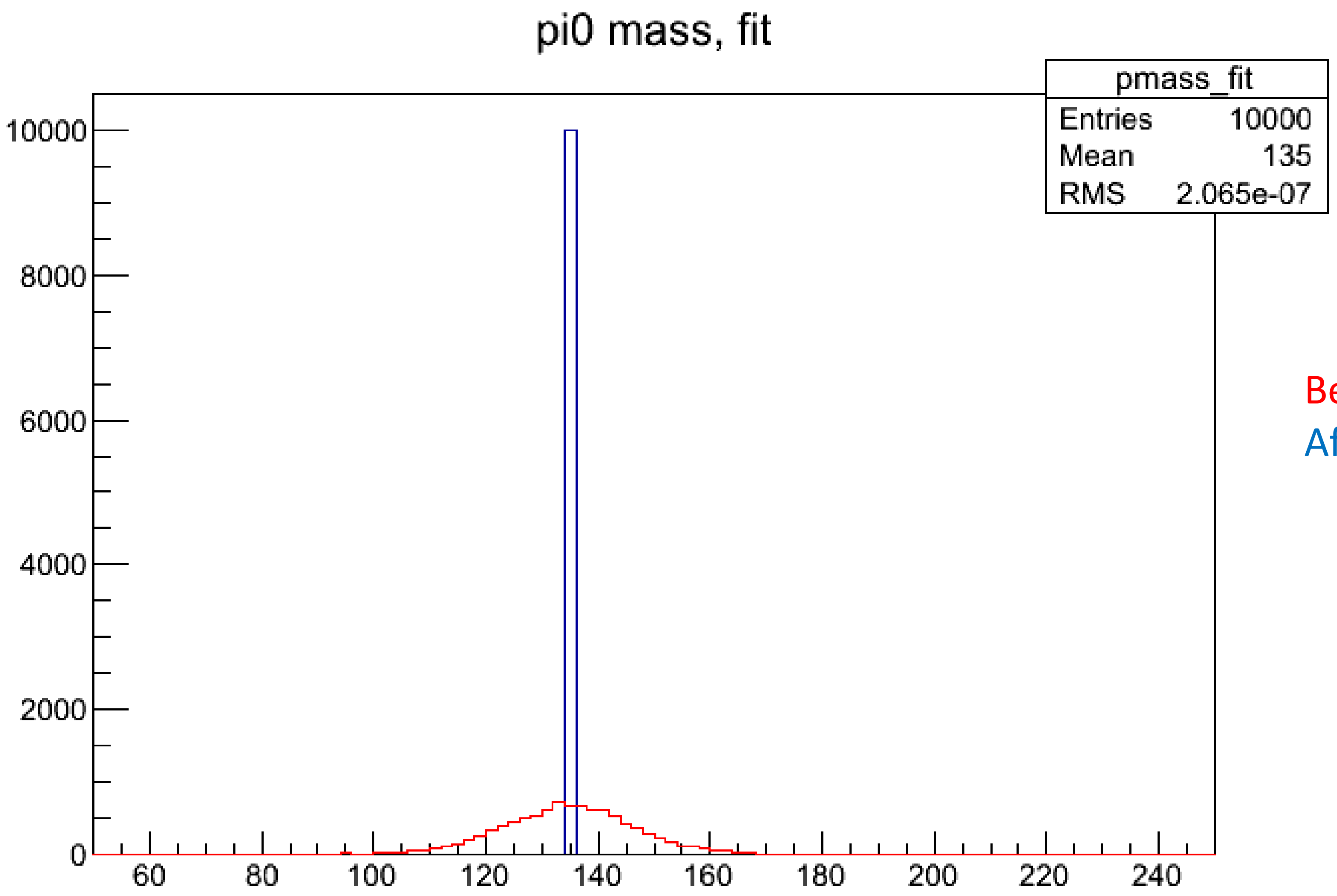
fit two-photon measured momentum



Before fit

After fit

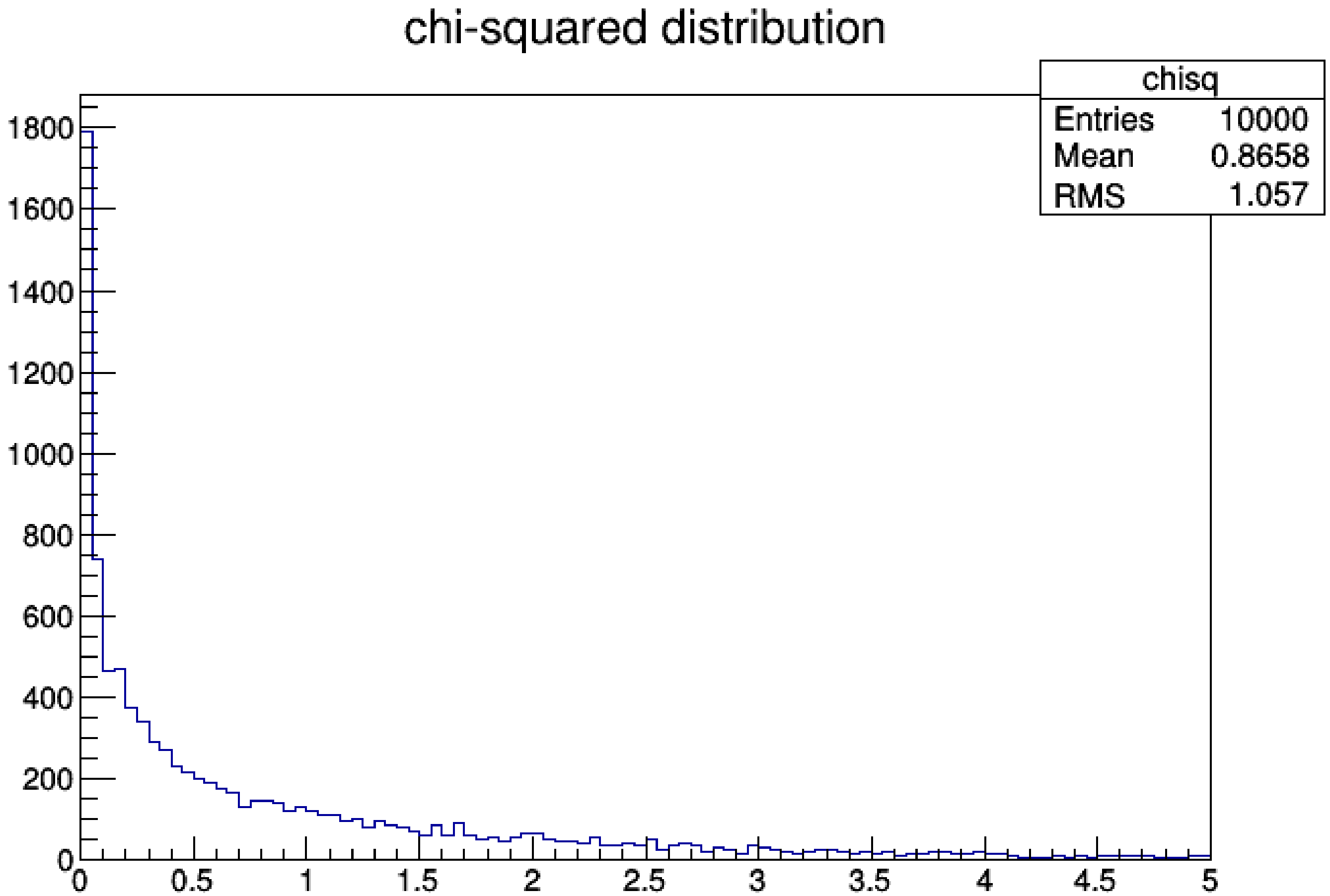
fit two-photon invariant mass



Before fit

After fit

χ^2 Distribution



Variable Definitions I

- measured variables

  N number of measured variables (input)
  y vector of measurements, N-dimensional (input)
  V covariance matrix, $N \times N$ (input)
  η vector of fit values of measured variables, N-dimensional (output)

$$\chi^2 = (y - \eta)^T V^{-1} (y - \eta)$$

- unmeasured variables

  J number of unmeasured variables (input)
  ξ vector of unmeasured variable values, J-dimensional (input)

Variable Definitions II

- constraints

  K number of constraints (input)

  f vector of constraint functions, K dimensional (input)

Each constraint a function of measured and unmeasured variables. When constraint is satisfied

$$f_k(\eta_1, \ldots, \eta_N, \xi_1, \ldots, \xi_J) = 0 \quad \text{for} \quad k = 1, \ldots, K$$

- Lagrange multipliers

  λ vector of mulipliers, K-dimensional

Extended χ^2 to be minimized:

$$\chi^2(\eta, \xi, \lambda) = (y - \eta)^T V^{-1}(y - \eta) + 2\lambda^T f(\eta, \xi)$$

Minimization Condition

Set all partial derivatives to zero:

$$\frac{\partial \chi^2}{\partial \eta_n} = \left[-2V^{-1}(y-\eta) + 2F_\eta^T \lambda\right]_n = 0, \quad n = 1, \ldots, N$$

$$\frac{\partial \chi^2}{\partial \xi_j} = \left[2F_\xi^T \lambda\right]_j = 0, \quad j = 1, \ldots, J$$

$$\frac{\partial \chi^2}{\partial \lambda_k} = [2f]_k = 0, \quad k = 1, \ldots, K$$

where

$$(F_\eta)_{kn} = \frac{\partial f_k}{\eta_n} \quad \text{and} \quad (F_\xi)_{kj} = \frac{\partial f_k}{\xi_j}$$

In general, a system of non-linear equations, $N + J + K$ equations with $N + J + K$ unknowns.

Stretch Functions or "Pulls"

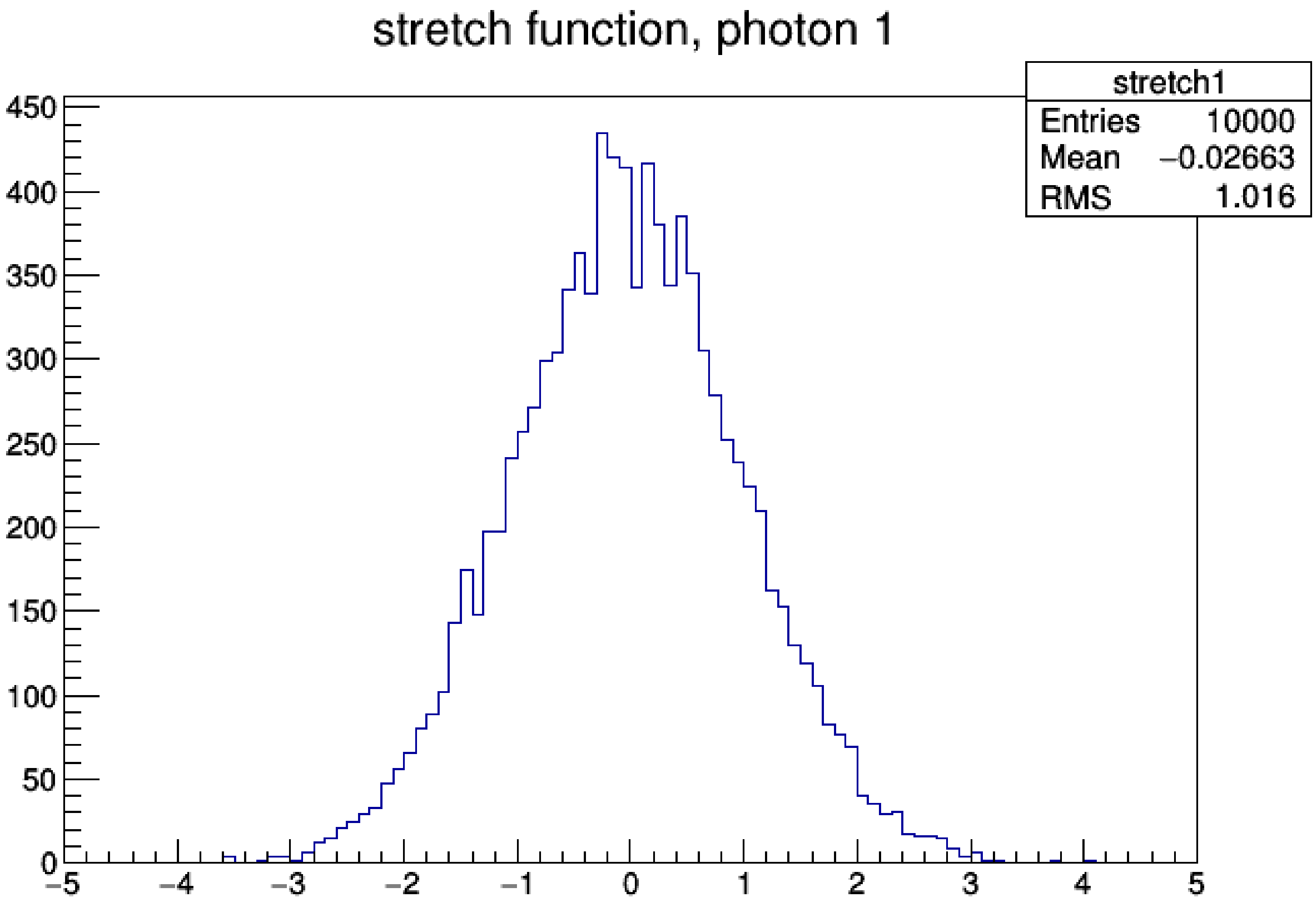
How to tell if the thing is working?

- look at use these N quantities:

$$z_n = \frac{y_n - \eta_n}{\sqrt{\sigma^2(y_n) - \sigma^2(\eta_n)}} \quad n = 1, \ldots, N$$

- Gaussian with mean at 0, σ of 1
- If not there are problems:
  - offset mean: measurements biased
  - wrong width: errors not correct
  - tails: non-Gaussian tails in measurements, background in sample

stretch function



Unmeasured Variables, Number of Constraints

- go back to π^0 decay
- could have introduced unmeasured variables: $p_{\pi,x}$, $p_{\pi,y}$, $p_{\pi,z}$
- but then would have to apply 3-momentum conservation
- now have 4 constraints with 3 unmeasured variables
- used to have 1 constraint with 0 unmeasured variables
- same problem recast: 1-C fit
- $C = K - J$, the number of degrees of freedom

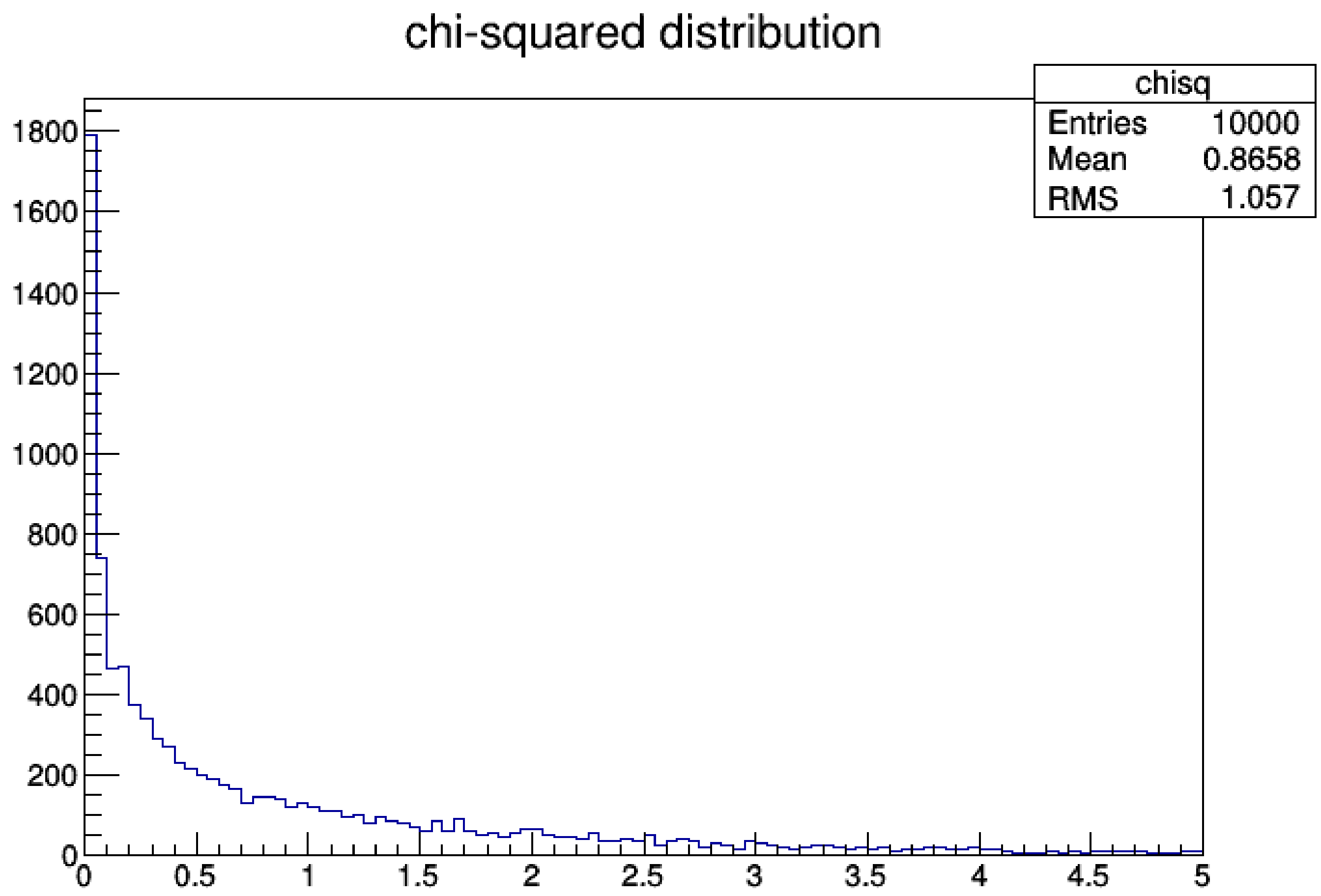
Check χ^2

- χ^2 should have a standard probablility density distribution: $f(\chi^2)$
- Convenient to check χ^2 probability:

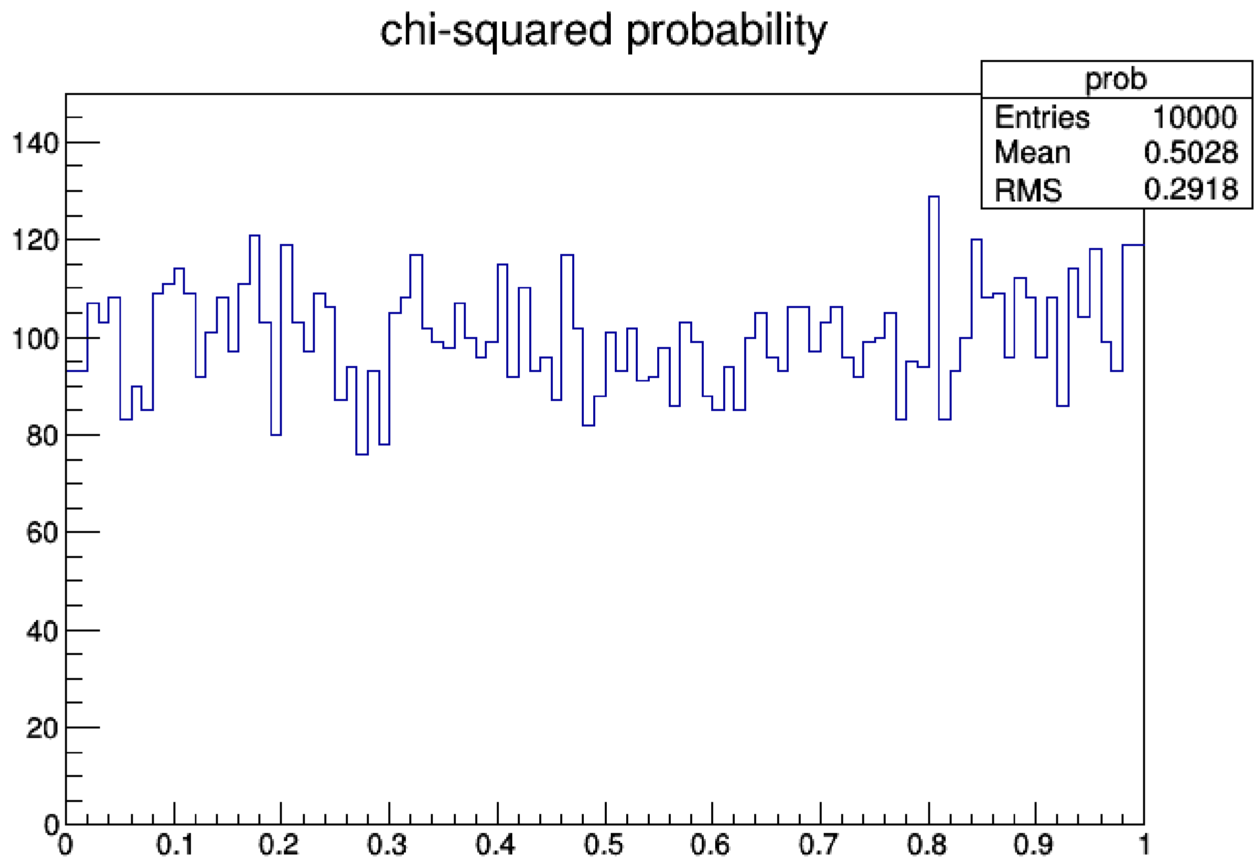
$$P(\chi_0^2) = \int_{\chi_0^2}^{\infty} f(\chi^2) d\chi^2$$

- P runs from 0 to 1
- for nominal χ^2 distribution P will be uniform
- non-uniformity: problem with errors, check the pulls
- often see peaks near 0: bad χ^2, background in sample

χ^2 Distribution



χ^2 Probability Distribution



Summary

- measured variables, with or without statistical correlation, may have physical relationships
- kinematic fit varies values of measured quantities to satisfy relationships
- minimize χ^2 with constraints
- improved measurements
- diagnostics about bias and errors in measurements are generated
- goodness of fit a handle on correctness of physical relationships assumed

$\phi\to\eta\gamma$ with $\eta\to\gamma\gamma$

Event selection of 3 photons final state - Kinematic fit

Impose kinematical constraints on measured variables

$$
\begin{aligned}
E_{\gamma_1} + E_{\gamma_2} + E_{\gamma_3} &= \sqrt{s} \\
\vec{p}_{\gamma_1} + \vec{p}_{\gamma_2} + \vec{p}_{\gamma_3} &= 0
\end{aligned}
$$

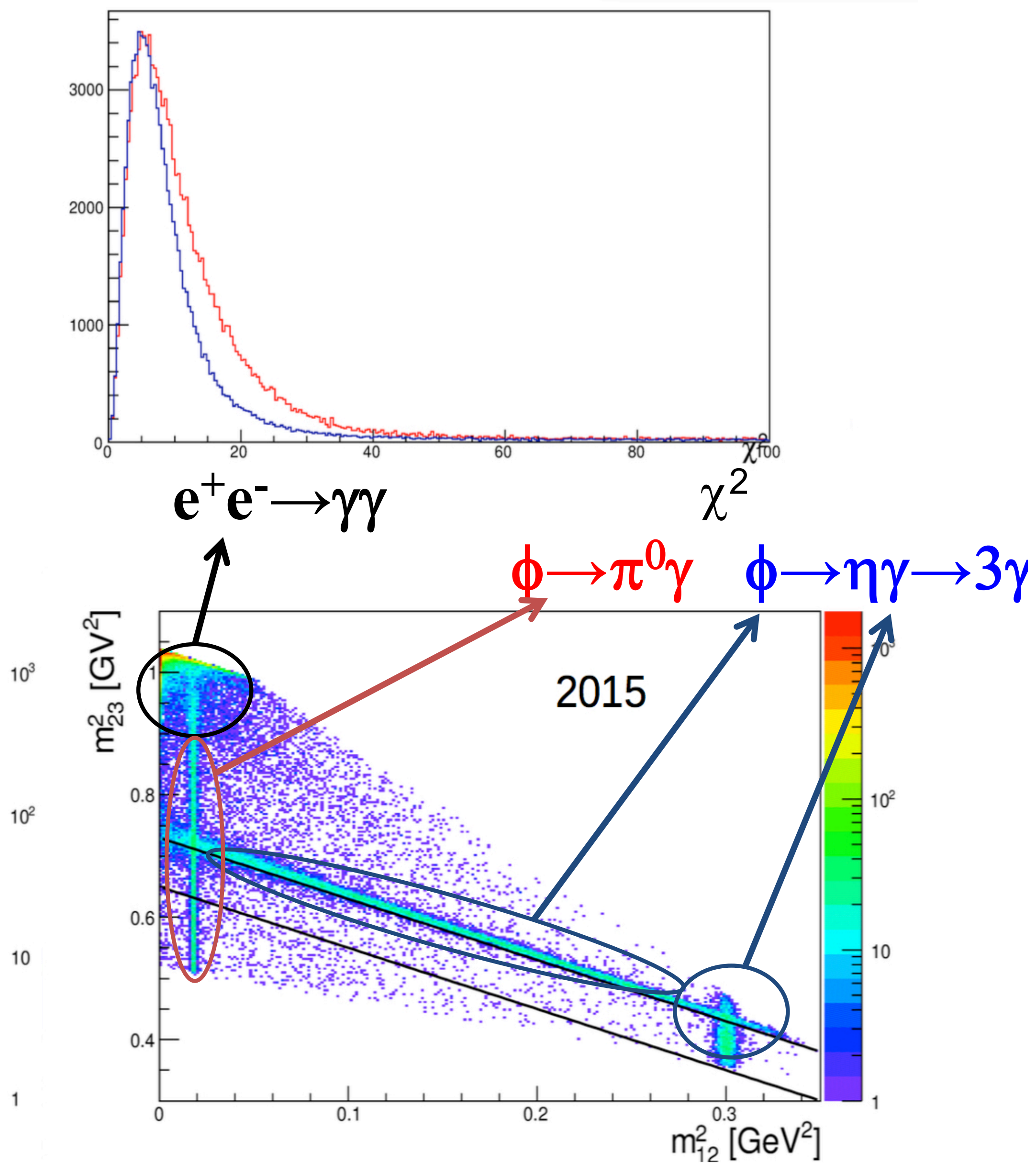
3 photons
in the final state:
7 constraints

Ti-Ri/c=0

$\phi \to \eta\gamma$ with $\eta \to \gamma\gamma$

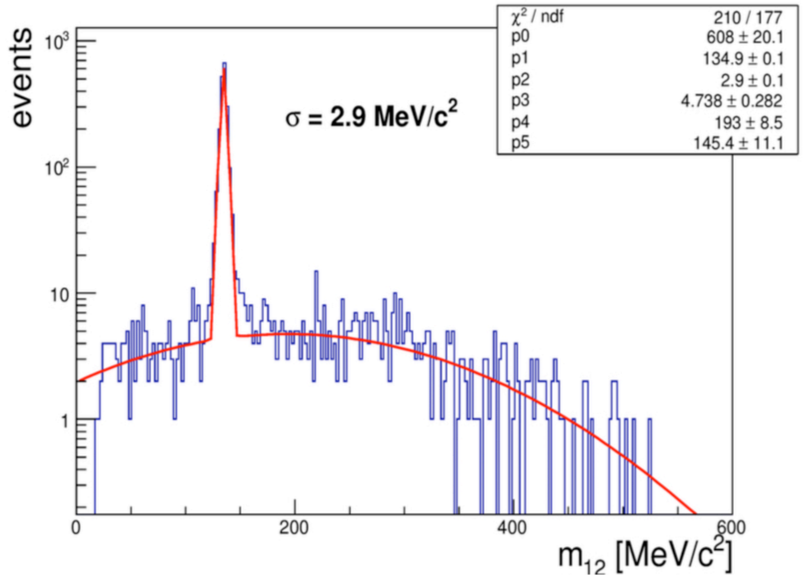
- E1 < E2 < E3
- Kinfit $\chi^2 < 35$

- Cuts to select eta and pion peaks



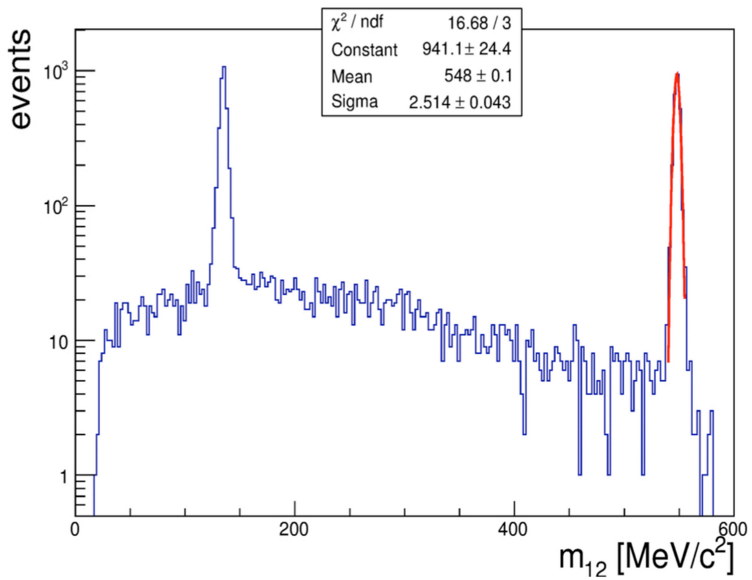
$\phi \to \eta\gamma$ with $\eta \to \gamma\gamma$

π^0 peak



$m_{23}^2 < 0.65\ GeV^2 - m_{12}^2$

η peak



$m_{23}^2 < 0.73\ GeV^2 - m_{12}^2$

Event selection - Kinematic fit

1) Determine or improve knowledge of kinematic quantities
2) Define a test statistics to select the event

Impose kinematical constraints on variables

$$\begin{aligned} E_{\gamma_1} + E_{\gamma_2} + E_{\gamma_3} &= \sqrt{s} \\ \vec{p}_{\gamma_1} + \vec{p}_{\gamma_2} + \vec{p}_{\gamma_3} &= 0 \end{aligned}$$

3 photons
in the final state:

1+3+3=7 constraints

Ti-Ri/c=0

Imposing the constraint:

$$E_{\gamma_i} E_{\gamma_j} \left(1 - \cos \Delta\alpha_{ij}\right) = M_\eta^2$$

would limit the study of the background or other similar contributions (e.g. $\pi^0 \to \gamma\gamma$)

8.2. **Typical constraints.** A list of the most common constraints used in kinematic fits is given here. In the following with N_c we indicate the number of constraints.

- *Quadri-momentum conservation* ($N_c = 4$). To apply this constraint the initial state has to be known. In e^+e^- collisions the initial state is known (apart from initial state radiation effects) while in pp collisions the initial state can be to a good approximation known only in the transverse plane. In fact the interaction takes place between 2 partons, so that the longitudinal momentum of the initial state is not defined at all.
- *Mass constraint* ($N_c = 1$). When several combinations are possible, the constraint allows to determine the "good" combination.
- *Vertex constraint* ($N_c = 2N_p - 3$ N_p is the number of particles). Two or more particles are constrained to converge in the same point, the vertex. Several methods have been developed to apply the vertex constraint.
- *Velocity constraint* ($N_c = N_p$). If the particle time of flight is measured, and the β of the particle is independently measured (or the particle is a photon so that $\beta = 1$), the constraint $T - L/(c\beta)$ can be applied to each particle.

Method of the Lagrange multipliers

8.3. **The method of the Lagrange Multipliers: an example.** The most widely used implementation of the kinematic fit is based on the *Lagrange Multipliers*.

We consider here a purely "mathematical" example to illustrate the main features of the method. Suppose that two variables, a and b are measured, the values $a_0 \pm \sigma_a$ and $b_0 \pm \sigma_b$ are obtained. We assume for simplicity that the a and b are not correlated and that the two uncertainties are equal, $\sigma_a = \sigma_b = \sigma$.

Method of the Lagrange multipliers

On the other hand we know that the sum of the two variable should satisfy the relation:

$$a + b = s \tag{233}$$

with s a known fixed number. We apply the Lagrange Multiplier method to this very elementary example.

The following χ^2 variable is introduced:

$$\chi^2 = \frac{(a - a_0)^2}{\sigma^2} + \frac{(b - b_0)^2}{\sigma^2} + 2\lambda\,(a + b - s) \tag{234}$$

where to the usual χ^2 an additional term has been added multiplied by a new parameter λ. The meaning of such an additional term is clear: it imposes directly the constraint 233. The χ^2 variable is now minimized with respect to the three parameters: a, b and λ. From the system we get:

$$\hat{a} = \frac{s}{2} + \frac{a_0 - b_0}{2} \tag{235}$$

$$\hat{b} = \frac{s}{2} - \frac{a_0 - b_0}{2} \tag{236}$$

$$\hat{\lambda} = -\frac{1}{2\sigma^2}\,(s - a_0 - b_0) \tag{237}$$

$\hat{a}$ and $\hat{b}$ are the best estimates of a and b taking into account the constraint. It is useful to rewrite the solutions for $\hat{a}$ and $\hat{b}$ in the following form:

$$\hat{a} = a_0 + \frac{s - a_0 - b_0}{2} \tag{238}$$

$$\hat{b} = b_0 + \frac{s - a_0 - b_0}{2} \tag{239}$$

as the sum of the measured quantities a_0 and b_0 and a term that vanishes if the constraint is satisfied by the measurements. In other words we see that the kinematic fit pulls a and b away from the measured values by a quantity depending on the constraint.

Since the two estimates $\hat{a}$ and $\hat{b}$ are functions of the measured a_0 and b_0, in order to evaluate the covariance matrix of $\hat{a}$ and $\hat{b}$, the formula for the uncertainty propagation is used[45]. We get:

$$\sigma_{\hat{a}} = \frac{\sigma}{\sqrt{2}} \tag{240}$$

$$\sigma_{\hat{b}} = \frac{\sigma}{\sqrt{2}} \tag{241}$$

$$cov\left[\hat{a}, \hat{b}\right] = -\frac{\sigma^2}{2} \tag{242}$$

---

[45]In case of M functions y_i depending on N variables x_k we have

$$cov\left[y_i, y_j\right] = \sum_{k,h} \frac{\partial y_i}{\partial x_k} \frac{\partial y_j}{\partial x_h} cov\left[x_k, x_h\right]$$

or, expressing it as a covariance matrix:

$$\begin{pmatrix} \frac{\sigma^2}{2} & -\frac{\sigma^2}{2} \\ -\frac{\sigma^2}{2} & \frac{\sigma^2}{2} \end{pmatrix}$$

The results are very interesting and illustrate the main features of the kinematic fit.

As already said, the constraint pulls the estimates of a and b from the measured values a_0 and b_0 to other values depending on the constraint. The uncertainties on the parameters decrease with respect to the measurement uncertainties and the estimates have a correlation even if the original measurements are not correlated.

By substituting the values of a and b in eq.234 with $\hat{a}$ and $\hat{b}$ given in eqs.241 and 242, the following χ^2 is obtained:

$$\chi^2 = \frac{2}{\sigma^2}\left(\frac{s}{2} - \frac{a_0 + b_0}{2}\right) \tag{243}$$

Since the uncertainty on $(a_0 + b_0)/2$ is $\sigma/\sqrt{2}$, it is a χ^2 with one degree of freedom, as expected since we have posed a single constraint.

If an additional variable c not measured (a sort of "neutrino") is introduced, it can be verified that with a single constraint only a trivial solution is obtained :

$$\hat{a} = a_0 \tag{244}$$

$$\hat{b} = b_0 \tag{245}$$

$$\hat{c} = s - a_0 - b_0 \tag{246}$$

with χ^2 identically equal to 0. No fit is obtained clearly, the number of unknowns being equal to the number of constraints. Additional constraints are needed in this case.

8.4. **The method of the Lagrange Multipliers: general formulation.** Let's assume that the final state we are analyzing depends on N variables α_i[46]. All these variables have been measured and the values α_{i0} have been obtained, with V_{ij} being the experimental covariance matrix of the measurements. Then we suppose to have R constraints, each of the form $H_k(\vec{\alpha}) = 0$[47], with the Hs being general functions. The χ^2 function including the Lagrange multipliers is:

$$\chi^2 = \sum_{ij} (\alpha_i - \alpha_{i0}) V_{ij}^{-1} (\alpha_j - \alpha_{j0}) + 2 \sum_k \lambda_k H_k(\vec{\alpha}) \tag{247}$$

The constraints can be expanded around a certain N-dimensional point $\vec{\alpha}_A$

$$H_k(\vec{\alpha}) = H_k(\vec{\alpha}_A) + \sum_j \frac{\partial H_k}{\partial \alpha_j} (\alpha_j - \alpha_{jA}) \tag{248}$$

in such a way that eq.247 becomes:

$$\chi^2 = \sum_{ij} (\alpha_i - \alpha_{i0}) V_{ij}^{-1} (\alpha_j - \alpha_{j0}) + 2 \sum_k \lambda_k \left(H_k(\vec{\alpha}_A) + \sum_j \frac{\partial H_k}{\partial \alpha_j} (\alpha_j - \alpha_{jA}) \right) \tag{249}$$

---

[46] If the final state consists of K particles, in the most general case $N = 7K$ since each particle have to be described in the most complete form by 7 variables: 3 coordinates of a point, three components of a vector and a mass.

[47] In this section we use the vecto symbol $\vec{\alpha}$ to identify vectors and the notation $\underline{V}$ to identify matrices.

Method of the Lagrange multipliers

The linearization of the constraints allows to have an analytically solvable system. The details of the derivation of the solution are not given here, the final results are shown.

Using a matrix formalism the following vectors and matrices are defined:

$$\Delta\vec{\alpha} = \vec{\alpha} - \vec{\alpha_A} \tag{250}$$

$$\vec{d} = \vec{H}(\vec{\alpha_A}) \tag{251}$$

$$D_{ki} = \left.\frac{\partial H_k}{\partial \alpha_j}\right|_{\alpha_j=\alpha_{jA}} \tag{252}$$

where the first is a vector of dimension N, the second of dimension R, the third is a $R \times N$ matrix. The χ^2 can be written as

$$\chi^2 = (\vec{\alpha} - \vec{\alpha}_0)^T \underline{V}^{-1}(\vec{\alpha} - \vec{\alpha}_0) + 2\vec{\lambda}^T(\underline{D}\Delta\vec{\alpha} + \vec{d}) \tag{253}$$

The minimization gives the following solution for the variables $\vec{\alpha}$:

$$\hat{\vec{\alpha}} = \vec{\alpha}_0 - \underline{V}\,\underline{D}^T(\underline{D}\underline{V}\,\underline{D}^T)^{-1}(\underline{D}\Delta\vec{\alpha}_0 + \vec{d}) \tag{254}$$

and the covariance matrix of the estimates is

$$\underline{V}' = \underline{V} - \underline{V}\,\underline{D}^T(\underline{D}\underline{V}\,\underline{D}^T)^{-1}\underline{D}\underline{V} \tag{255}$$

Finally the χ^2 can be expressed as the sum of R terms:

$$\chi^2 = \vec{\lambda}^T(\underline{D}\Delta\vec{\alpha}_0 + \vec{d}) \tag{256}$$

one per constraint.

Eq.254 shows that the best estimate of the kinematic variables of the event are equal to the measured values minus terms that depend on the constraints. The variables are "pulled" from the measured values. The covariance matrix of the estimated variables is also pulled (see eq.255) from the measurement covariance matrix. It can be demonstrated that the diagonal terms of V' are always smaller than the corresponding diagonal terms of V, so that the outcome of the kinematic fit is an improved kinematic reconstruction of the event.

Finally the so called **pulls** are defined as measures of how each single variable is pulled away from the measured values:

$$pull_i = \frac{\hat{\alpha}_i - \alpha_{i0}}{\sqrt{\sigma^2_{\alpha_{i0}} - \sigma^2_{\alpha_i}}} \tag{257}$$

the denominator is the uncertainty on the difference between the two variables. If the kinematic fit is working correctly, the distribution of the pulls should have a standardized gaussian shape.