Minimising of functions with constraints

It is often necessary to minimise a function $f(x_1,x_2,...x_n)$ which is subject to constraints:

$$
\begin{aligned}
&g_1(x_1,x_2,...x_n) = 0 \\
&g_2(x_1,x_2,...x_n) = 0 \\
&. \\
&. \\
&g_m(x_1,x_2,...x_n) = 0 \quad \text{where } m < n.
\end{aligned}
\tag{1}
$$

A straightforward method to solve this problem is to incorporate the constraints into the fitting function by elimination of variables. This method works, however, only very simple problems.

Ex. Take a measurement of two angles that we know should add up to 90^0. The function to minimise is then

$$\chi^2(\eta_1,\eta_2)=\sum_{i=1}^{2}\left(\frac{y_i-\eta_i}{\sigma_i}\right)^2 = \text{mimimum}$$

with

$$g=\sum_{i=1}^{2}\eta_i-90^0=0^0$$

as constraint.

Substitution of η_2 using the constraint then gives

$$\chi^2=\left(\frac{y_1-\eta_1}{\sigma_1}\right)^2+\left(\frac{y_2-(90^0-\eta_1)}{\sigma_2}\right)^2 = \text{minimum}$$

This has the solution (assuming $\sigma_1 = \sigma_2$)

$$\hat{\eta}_1=\frac{1}{2}\left(90^0+y_1-y_2\right)=y_1+\frac{1}{2}\left(90^0-y_1-y_2\right)$$
$$\hat{\eta}_2=\frac{1}{2}\left(90^0-y_1+y_2\right)=y_2+\frac{1}{2}\left(90^0-y_1-y_2\right)$$

The uncertainties in these angles is found by applying the error propagation:

$$V(\hat{\bar{\eta}}) = SV(\bar{y})S^T$$

This gives

$$V(\hat{\bar{\eta}}) = \begin{pmatrix} \frac{d\eta_1}{dy_1} & \frac{d\eta_1}{dy_2} \\ \frac{d\eta_2}{dy_1} & \frac{d\eta_2}{dy_2} \end{pmatrix} \begin{pmatrix} \sigma_y^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix} \begin{pmatrix} \frac{d\eta_1}{dy_1} & \frac{d\eta_2}{dy_1} \\ \frac{d\eta_1}{dy_2} & \frac{d\eta_2}{dy_2} \end{pmatrix} =$$

$$\begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \sigma_y^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} = \sigma_y^2 \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

=> The accuracy of the fitted values for the angles are a factor of $\sqrt{2}$ smaller than the measured ones.

=> The constraints introduce an (anti)correlation between the fitted values.

Elimination of variables becomes very quickly (too) complicated. A widely used method for minimising functions with constraints is **the Method of Lagrangian Multipliers**.

Lagrangian Multipliers: Method

Take a function $f(x,y,z) = w$ which is to be minimised subject to the constraints:

$$g_1(x,y,z) = 0$$
$$g_2(x,y,z) = 0 \qquad (2)$$

Here g_1 and g_2 describe a surface in the 3-d space.

=> their intersection is a curve in space and we should minimise the function, $f(x,y,z)$, along this intersection curve.

This means that

$$\nabla f = \frac{\partial f}{\partial x}\bar{i} + \frac{\partial f}{\partial y}\bar{j} + \frac{\partial f}{\partial z}\bar{k} \qquad (3)$$

must lie in a plane normal to this curve at a minimum. ∇g_1 and ∇g_2 must also lie in the same plane.

Generally, if three vectors are coplanar there exist two scalars λ_1 and λ_2 such that

$$\nabla f + \lambda_1 \nabla g_1 + \lambda_2 \nabla g_2 = 0 \tag{4}$$

In 3-d space this represents three scalar equations:

$$\begin{cases} \dfrac{\partial f}{\partial x} + \lambda_1 \dfrac{\partial g_1}{\partial x} + \lambda_2 \dfrac{\partial g_2}{\partial x} = 0 \\ \dfrac{\partial f}{\partial y} + \lambda_1 \dfrac{\partial g_1}{\partial y} + \lambda_2 \dfrac{\partial g_2}{\partial y} = 0 \\ \dfrac{\partial f}{\partial z} + \lambda_1 \dfrac{\partial g_1}{\partial z} + \lambda_2 \dfrac{\partial g_2}{\partial z} = 0 \end{cases} \tag{5}$$

These three equations, together with the constraints (2) represent five equations that can be solved with respect to the five unknowns: x^0, y^0, z^0, λ_1 and λ_2. These λ's are called **Lagrangian multipliers**.

In the general case, one forms the N equations for the function

$$\frac{\partial f}{\partial x_i}\left\{f(\overline{x})+\lambda_1 g_1(\overline{x})+\lambda_2 g_2(\overline{x})+\ldots+\lambda_M g_M(\overline{x})\right\}=0 \quad i=1,2,\ldots,N \tag{6}$$

and the M equations for the constraints

$$g_j(\overline{x})=0 \qquad j=1,2,\ldots\ldots,M \tag{7}$$

to solve the N + M unknowns: , λ_j (j = 1,2,...,M).

Matrix notation:

$$\begin{cases} \chi^2\left(\overline{\theta}\right)=\left(\overline{y}-A\overline{\theta}\right)^{\mathrm{T}} V^{-1}\left(\overline{y}-A\overline{\theta}\right)= \text{ minimum} \\ G\overline{\theta}-\overline{g}=0 \end{cases} \tag{8}$$

The introduction of the Lagrangian multipliers, $\overline{\lambda}$, gives

$$\chi^2(\overline{\theta},\overline{\lambda})=\left(\overline{y}-A\overline{\theta}\right)^T V^{-1}\left(\overline{y}-A\overline{\theta}\right)+2\overline{\lambda}^T\left(G\overline{\theta}-\overline{g}\right)= \text{ minimum} \tag{9}$$

N. B. The term with the Lagrangian multipliers adds zero to the χ^2 if the constraint equations are fulfilled.

The minimum is found by setting the derivatives of $\chi^2(\overline{\theta},\overline{\lambda})$ to zero:

$$\begin{cases} \nabla_\theta \chi^2 = -2\left(A^T V^{-1}\overline{y} - A^T V^{-1} A\overline{\theta}\right) + 2G^T\overline{\lambda} = 0 \\ \nabla_\lambda \chi^2 = 2\left(G\overline{\theta}-\overline{g}\right) = 0 \end{cases} \tag{10}$$

Introduce $C \equiv A^T V^{-1} A$ and $\overline{c} \equiv A^T V^{-1}\overline{y}$ and rewrite (10) as

$$\begin{cases} C\overline{\theta} + G^T\overline{\lambda} = \overline{c} \\ G\overline{\theta} = \overline{g} \end{cases} \tag{11}$$

$$\begin{cases} C\bar{\theta} + G^T\bar{\lambda} = \bar{c} \\ G\bar{\theta} = \bar{g} \end{cases} \tag{11}$$

Multiply the upper Eq. (11) by GC^{-1} from the left and use the lower one:

$$\bar{g} + GC^{-1}G^T\bar{\lambda} = GC^{-1}\bar{c} \tag{12}$$

Introducing $V_G \equiv GC^{-1}G^T$ gives the solution, $\hat{\lambda}$, for the Lagrangian multipliers:

$$\hat{\lambda} = V_G^{-1}\left(GC^{-1}\bar{c} - \bar{g}\right) \tag{13}$$

(13) substituted into into the upper Eq. (11) gives the solution for $\hat{\theta}$ as

$$\hat{\theta} = C^{-1}\bar{c} - C^{-1}G^T V_G^{-1}\left(GC^{-1}\bar{c} - \bar{g}\right) \tag{14}$$

measures how much the constraint Eqs. are violated by the measurement

An expression for $V(\hat{\theta})$ is obtained from the error propagation formula:

$$V(\hat{\theta}) = C^{-1} - \left(GC^{-1}\right)^T V_B^{-1} \left(GC^{-1}\right) = C^{-1}\left(I_N - G^T V_B^{-1} G C^{-1}\right) \qquad (15)$$

where I_N is the unitary matrix. Equations (14) and (15) provide an exact solution since all matrices and vectors are known.

N.B. For the unconstrained problem is $C^{-1}\overline{c}$ the solution and C^{-1} the covariance matrix. The uncertainties in the estimated parameters are usually smaller in the constrained case.

Lagrangian multipliers: Iterative procedure

$\bar{x}$ = a vector of N observables for which we have the first approximation
i. e. a measurement $\bar{y}$ with its uncertainties contained in the covariance matrix $V(\bar{y})$.

$\bar{\xi} = \{\xi_1, \xi_2, ..., \xi_J\}$ = a set of unmeasured variables which are deduced from the measurement via the constraints. The N measured and J unmeasured variables are related and have to satisfy a set of K constraint equations:

$$g_k(x_1, x_2, ..., x_N, \xi_1, \xi_2, ... \xi_J) = 0 \quad k = 1,2,...,K \qquad (16)$$

According to the principle of Least Squares we should minimise (c. f. Eq. (8))

$$\begin{cases} \chi^2(\bar{x}) = (\bar{y} - \bar{x})^T V^{-1}(\bar{y})(\bar{y} - \bar{x}) = \text{minimum} \\ \bar{g}(\bar{x}, \bar{\xi}) = 0 \end{cases} \qquad (17)$$

Eq. (17) is solved by the method of Lagrangian multipliers
=> Introduce K additional unknowns $\bar{\lambda} = (\lambda_1, \lambda_2, ..., \lambda_K)$ and rewrite (17) requiring (c.f. (9))

$$\chi^2(\bar{x}, \bar{\xi}, \bar{\lambda}) = (\bar{y} - \bar{x})^T V^{-1}(\bar{y})(\bar{y} - \bar{x}) + 2\bar{\lambda}^T \bar{g}(\bar{x}, \bar{\xi}) = \text{minimum} \qquad (18)$$

We have a total of $N+J+K$ unknowns. To find the minimum we take the derivative of χ^2 with respect to these unknowns and take the result to be equal to zero:

$$\begin{cases} \nabla_x \chi^2 = -2V^{-1}\left(\bar{y} - \bar{x}\right) + 2G_x^T \bar{\lambda} = 0 \\ \nabla_\xi \chi^2 = 2G_\xi^T \bar{\lambda} = 0 \\ \nabla_\lambda \chi^2 = 2\bar{g}_\xi^T = 0 \end{cases} \qquad (19)$$

The elements of the matrices G_x $(K \times N)$ and G_ξ $(K \times J)$ are defined as

$$\left(G_x\right)_{ki} \equiv \frac{\partial g_k}{\partial x_i} \quad ; \quad \left(G_\xi\right)_{kj} \equiv \frac{\partial g_k}{\partial \xi_j} \qquad (20)$$

The solution of (19) for the $N+J+K$ equations must generally be found by iteration.

Assume that iteration number ν has been made and that want to make one more iteration. We then make a Taylor expansion of the constraint equation around the point $(\overline{x}^{\nu},\overline{\xi}^{\nu})$:

$$g_k(\overline{x},\overline{\xi})=g_k^{\nu}(\overline{x}^{\nu},\overline{\xi}^{\nu})+\sum_{i=1}^{N}\left(\frac{\partial g_k}{\partial x_i}\right)^{\nu}\left(x_i^{\nu+1}-x_i^{\nu}\right)+\sum_{j=1}^{J}\left(\frac{\partial g_k}{\partial \xi_j}\right)^{\nu}\left(\xi_j^{\nu+1}-\xi_j^{\nu}\right)+H.O.=0 \quad (21)$$

Eq. (21) can be rewritten using Eq. (20) (neglecting the higher order terms):

$$\overline{g}^{\nu}+G_x^{\nu}\left(\overline{x}^{\nu+1}-\overline{x}^{\nu}\right)+G_{\xi}^{\nu}\left(\overline{\xi}^{\nu+1}-\overline{\xi}^{\nu}\right)=0 \quad (22)$$

The first two equations of (19) now read for the (ν+1):th iteration

$$\begin{cases} V^{-1}\left(\overline{x}^{\nu+1}-\overline{y}\right)+\left(G_x^T\right)^{\nu}\overline{\lambda}^{\nu+1}=0 \\ \left(G_{\xi}^T\right)^{\nu}\overline{\lambda}^{\nu+1}=0 \end{cases} \quad (23)$$

These equations make it possible to express all the unknowns of the (ν+1):th iteration.

- Multiply the first Eq. of (23) by V from the left to get an expression for $\bar{x}^{\nu+1}$.
- Substitute the result into Eq. (22):

$$\bar{g}^{\nu} + G_x^{\nu}\left(\left(\bar{y} - V\left(G_x^T\right)^{\nu} \bar{\lambda}^{\nu+1}\right) - \bar{x}^{\nu}\right) + G_{\xi}^{\nu}\left(\bar{\xi}^{\nu+1} - \bar{\xi}^{\nu}\right) = 0 \tag{24}$$

This may be rewritten in short form as

$$\bar{r} + G_{\xi}^{\nu}\left(\bar{\xi}^{\nu+1} - \bar{\xi}^{\nu}\right) = S\bar{\lambda}^{\nu+1} \tag{25}$$

using

$$\bar{r} \equiv \bar{g}^{\nu} + G_x^{\nu}\left(\bar{y} - \bar{x}^{\nu}\right) \text{ and } S \equiv G_x^{\nu} V\left(G_x^T\right)^{\nu}$$

Multiplying (25) with S^{-1} from the left gives an expression for $\bar{\lambda}^{\nu+1}$.

Substituting this expression into the lower Eq. of (23) gives an equation where only $\bar{\xi}^{\nu+1}$ is unknown:

$$\left(G_{\xi}^{T}\right)^{\nu} S^{-1}\left(\bar{r}+G_{\xi}^{\nu}\left(\bar{\xi}^{\nu+1}-\bar{\xi}^{\nu}\right)\right)=0 \tag{26}$$

This equation can be solved for $\bar{\xi}^{\nu+1}$ and the result can be substituted back into (25) to obtain $\bar{\lambda}^{\nu+1}$ and finally get $\bar{x}^{\nu+1}$ from (23):

$$\bar{\xi}^{\nu+1}=\bar{\xi}^{\nu}-\left(G_{\xi}^{T} S^{-1} G_{\xi}\right)^{-1} G_{\xi}^{T} S^{-1} \bar{r} \tag{27}$$

$$\bar{\lambda}^{\nu+1}=S^{-1}\left(\bar{r}+G_{\xi}\left(\bar{\xi}^{\nu+1}-\bar{\xi}^{\nu}\right)\right) \tag{28}$$

$$\bar{x}^{\nu+1}=\bar{y}-V G_{x}^{T} \bar{\lambda}^{\nu+1} \tag{29}$$

Good starting values $\bar{x}^{0}$, $\bar{\xi}^{0}$ are important for the convergence of the iterations:

$\bar{x}^{0}$: start with $\bar{x}^{0}$ = $\bar{y}$, i. e. the measured values

$\bar{\xi}^{0}$: to be evaluated from the most convenient constraint equation inserting the measured $\bar{x}^{0}$ for $\bar{x}$.

Summary: The minimisation procedure with constraints using Lagrangian multipliers should proceed in the following steps:

1) Evaluate the vector $\bar{r}$ and the matrix S from the definitions in Eq. (25) using the starting values for $\bar{x}^0$ and $\bar{\xi}^0$.

2) Find the new vector for the unknowns $\bar{\xi}^{\nu+1}$ from Eq. (27).

3) Find the new vector $\bar{\lambda}^{\nu+1}$ of the Lagrangian multipliers from Eq. (28).

4) Find the new vector $\bar{x}^{\nu+1}$ of measured quantities from Eq. (29).

5) Calculate the new chi-square value $\left(\chi^2\right)^{\nu+1}$ from Eq. (18).

6) Compare the result from 5) with the previous iteration. Stop if satisfactory results are obtained otherwise proceeds to 1).

Lagrangian Multipliers: Error calculation

The uncertainties in the final estimates of the measured and unmeasured variables are found by applying the law of error propagation.

Write the estimators $\hat{\overline{x}}$ (= $\overline{x}^{\nu+1}$) and $\hat{\overline{\xi}}$ (= $\overline{\xi}^{\nu+1}$) as functions of the measurements:

$$\begin{cases} \hat{\overline{x}} = \overline{a}(\overline{y}) \\ \hat{\overline{\xi}} = \overline{b}(\overline{y}) \end{cases} \qquad (30)$$

The forms of $\overline{a}$ and $\overline{b}$ are given by the Equations (27 - 29). and by using the definition of $\overline{r}$ in Eq. (25).

$$\begin{cases} \overline{a} = \overline{y} - VG_x^T \underbrace{S^{-1}\left[I_K - G_\xi\left(G_\xi^T S^{-1} G_\xi\right)^{-1} G_\xi^T S^{-1}\right]\left[\overline{g} + G_x\left(\overline{y} - \overline{x}\right)\right]}_{\lambda^{\nu+1}} \\ \overline{b} = \overline{\xi} - \left(G_\xi^T S^{-1} G_\xi\right)^{-1} G_\xi^T S^{-1} \underbrace{\left[\overline{g} + G_x\left(\overline{y} - \overline{x}\right)\right]}_{\overline{r}} \end{cases} \qquad (31)$$

where I_k is the unity matrix.

In these formulae (Eq. 31) are $\overline{x}$ and $\overline{\xi}$, as well as $\overline{g}$ and the matrices G and S evaluated in the last iteration.

The covariance matrices for $\hat{\overline{x}}$ and $\hat{\overline{\xi}}$ are given by the error propagation law as (assuming a linear dependence on $\overline{x}$)

$$\begin{cases} V(\hat{\overline{x}}) = \left(\frac{d\overline{a}}{d\overline{y}} \right) V(\overline{y}) \left(\frac{d\overline{a}}{d\overline{y}} \right)^T \\ V(\hat{\overline{\xi}}) = \left(\frac{d\overline{b}}{d\overline{y}} \right) V(\overline{y}) \left(\frac{d\overline{b}}{d\overline{y}} \right)^T \\ \operatorname{cov}(\hat{\overline{x}}, \hat{\overline{\xi}}) = \left(\frac{d\overline{a}}{d\overline{y}} \right) V(\overline{y}) \left(\frac{d\overline{b}}{d\overline{y}} \right)^T \end{cases} \tag{32}$$

The derivatives can be calculated using Eq. 31:

$$\begin{cases} \frac{d\overline{a}}{d\overline{y}} = I_N - V(\overline{y}) \left[G_x^T S^{-1} G_x - G_x^T S^{-1} G_\xi \left(G_\xi^T S^{-1} G_\xi \right)^{-1} G_\xi^T S^{-1} G_x \right] \\ \frac{d\overline{b}}{d\overline{y}} = -\left(G_\xi^T S^{-1} G_\xi \right)^{-1} G_\xi^T S^{-1} G_x \end{cases} \tag{33}$$

By defining

$$A \equiv G_x^T S^{-1} G_x, \; B \equiv G_x^T S^{-1} G_\xi \text{ and } U^{-1} \equiv G_\xi^T S^{-1} G_\xi \tag{34}$$

one obtain after "some" algebra

$$\begin{cases} V(\hat{\bar{x}}) = V(\bar{y})\left[I_N - \left(A - BUB^T \right) V(\bar{y}) \right] \\ V(\hat{\bar{\xi}}) = U \\ \operatorname{cov}(\hat{\bar{x}}, \hat{\bar{\xi}}) = -V(\bar{y}) BU \end{cases} \tag{35}$$

These formulae imply

i) The uncertainties of the constrained fitted quantities are generally smaller than the uncertainties in the observations of $\bar{y}$.

ii) The fitted quantities will be correlated even if the measurements are uncorrelated.

A useful test of the quality of a constrained fit is to use the "pull" or "stretch" variable:

$$\frac{\left(x_i - \hat{x}_i\right)}{\sqrt{\sigma_{x_i}^2 - \sigma_{\hat{x}_i}^2}} \qquad (36)$$

It can be shown that this quantity is normally distributed with a mean of 0 and a variance of 1. If the pulls are not normally distributed it is an indication of that either the uncertainties are badly estimated or that they are inherently non-normal. In the latter case is the χ^2 not a reliable measure of the goodness of the fit.

Lagrangian Multipliers: Example of kinematical fitting

A frequently used application of minimisation with constraints is kinematical fitting, i. e. a fit of the kinematical variables of a measured reaction/decay using energy and momentum conservation to improve the result.

As an example, take the two-body decay of a particle:

$$\Lambda \rightarrow p + \pi^-$$

This type of decay is often called V^0-decay because it involves the decay of a neutral particle (unobserved) into two charged ones. The topology of such decay looks as a Vee in the detector.

Assume that we have measured the momenta and directions of the decay particles but the origin and momentum of the Λ-particle is unmeasured. We then have six measured quantities

$$\overline{x} = \left\{ p_p, \theta_p, \phi_p, p_\pi, \theta_\pi, \phi_\pi \right\}$$

and three unknown variables

$$\overline{\xi} = \left\{ p_\Lambda, \theta_\Lambda, \phi_\Lambda \right\}$$

Spherical coordinates: p = momentum
θ = polar angle
Φ = azimuthal angle

There are four constraint equations of momentum and energy conservation:

$$g_1 = -p_\Lambda \sin\theta_\Lambda \cos\phi_\Lambda + p_p \sin\theta_p \cos\phi_p + p_\pi \sin\theta_\pi \cos\phi_\pi = 0 \quad (p_x)$$
$$g_2 = -p_\Lambda \sin\theta_\Lambda \sin\phi_\Lambda + p_p \sin\theta_p \sin\phi_p + p_\pi \sin\theta_\pi \sin\phi_\pi = 0 \quad (p_y)$$
$$g_3 = -p_\Lambda \cos\theta_\Lambda + p_p \cos\theta_p + p_\pi \cos\theta_\pi = 0 \quad (p_z)$$
$$g_4 = -\sqrt{p_\Lambda^2 + m_\Lambda^2} + \sqrt{p_p^2 + m_p^2} + \sqrt{p_\pi^2 + m_\pi^2} = 0 \quad (E)$$

=> There are four constraints and three unmeasured quantities.

# of constraints - # of unknowns = 1, i. e. one overconstraint.
Such a fit is called a 1C fit (C = # of overconstraints).

N.B. $\chi^2_{\#\text{ of overconstraints}} = \chi^2_\nu$

From Eq. (20) we have that the matrices $G_x\left(4\times 6\right)$ and $G_\xi\left(4\times 3\right)$ are obtained from taking the derivatives of the constraint equations as

$$\left(G_x\right)_{ki} \equiv \frac{\partial g_k}{\partial x_i} \quad ; \quad \left(G_\xi\right)_{kj} \equiv \frac{\partial g_k}{\partial \xi_j}$$

giving

$$G_x = \begin{pmatrix} \sin\theta_p \cos\phi_p & p_p \cos\theta_p \cos\phi_p & -p_p \sin\theta_p \sin\phi_p & \sin\theta_\pi \cos\phi_\pi & p_\pi \cos\theta_\pi \cos\phi_\pi & -p_\pi \sin\theta_\pi \sin\phi_\pi \\ \sin\theta_p \sin\phi_p & p_p \cos\theta_p \sin\phi_p & p_p \sin\theta_p \cos\phi_p & \sin\theta_\pi \sin\phi_\pi & p_\pi \cos\theta_\pi \sin\phi_\pi & p_\pi \sin\theta_\pi \cos\phi_\pi \\ \cos\theta_p & -p_p \sin\theta_p & 0 & \cos\theta_\pi & -p_\pi \sin\theta_\pi & 0 \\ \frac{p_p}{\sqrt{p_p^2 + m_p^2}} & 0 & 0 & \frac{p_\pi}{\sqrt{p_\pi^2 + m_\pi^2}} & 0 & 0 \end{pmatrix}$$

and

$$G_{\xi} = \begin{pmatrix} -\sin\theta_{\Lambda}\cos\phi_{\Lambda} & -p_{\Lambda}\cos\theta_{\Lambda}\cos\phi_{\Lambda} & p_{\Lambda}\sin\theta_{\Lambda}\sin\phi_{\Lambda} \\ -\sin\theta_{\Lambda}\sin\phi_{\Lambda} & -p_{\Lambda}\cos\theta_{\Lambda}\sin\phi_{\Lambda} & -p_{\Lambda}\sin\theta_{\Lambda}\cos\phi_{\Lambda} \\ -\cos\theta_{\Lambda} & p_{\Lambda}\sin\theta_{\Lambda} & 0 \\ \frac{-p_{\Lambda}}{\sqrt{p_{\Lambda}^2 + m_{\Lambda}^2}} & 0 & 0 \end{pmatrix}$$

To start the iterations we take the measurement as the initial $\bar{x}^0$:

$$\bar{y} = \left\{ p_p^0, \theta_p^0, \phi_p^0, p_\pi^0, \theta_\pi^0, \phi_\pi^0 \right\}$$

For $\bar{\xi}^0$ we can take, for example, the value

$$\bar{\xi}^0 = \left\{ p_\Lambda^0, \theta_\Lambda^0, \phi_\Lambda^0 \right\}$$

by demanding that the first three constraint equations are fulfilled, i. e. momentum conservation satisfied. The fourth constraint equation will then, in general, not be satisfied.

From this latter consideration we then have an initial value of the vector $\overline{r}$ from Eq. (25) as

$$\overline{r} = \{0,0,0,g_4\}$$

The matrices G_x and G_ξ can now be calculated using Eq. (20) and the approximations $(\overline{x}^0, \overline{\xi}^0)$ and to obtain the (4×4)matrix S from Eq. (25)

$$S = G_x^0 V \left(G_x^0\right)^T$$

By inverting this matrix we can find the next iterative values, $\overline{\xi}^1, \overline{\lambda}^1, \overline{x}^1$ in succession from Equations (27 - 29).

$$p+p \rightarrow p+p+\eta$$
$$\hookrightarrow 2\gamma$$

Here 5 constraints: 4 from energy - momentum conservation
1 from invariant γγ mass = η mass

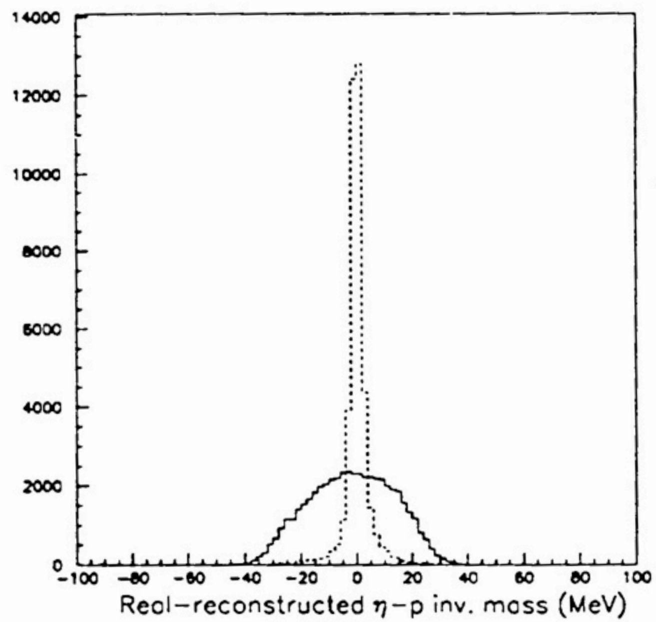

Figure 59. Improvement of the η-p invariant mass resolution with the kinematic fit routine. The η-p invariant mass is reconstructed from 1352 MeV Monte Carlo data, and the difference between the real and reconstructed invariant mass is plotted for each event. The solid line corresponds to reconstruction before the kinematic fit, and the dashed line how the mass is reconstructed after the kinematic fit has been applied. The FWHM resolution is improved from 40 to 4 MeV.

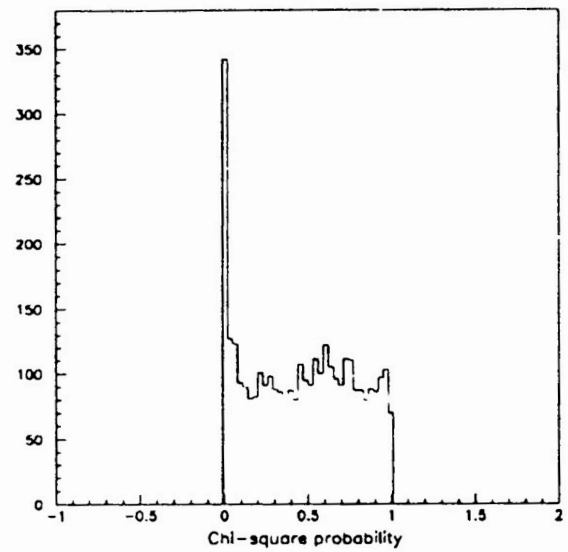

Figure 66. Chi-square probability distribution of the 3-C fit applied to 1352 MeV data.

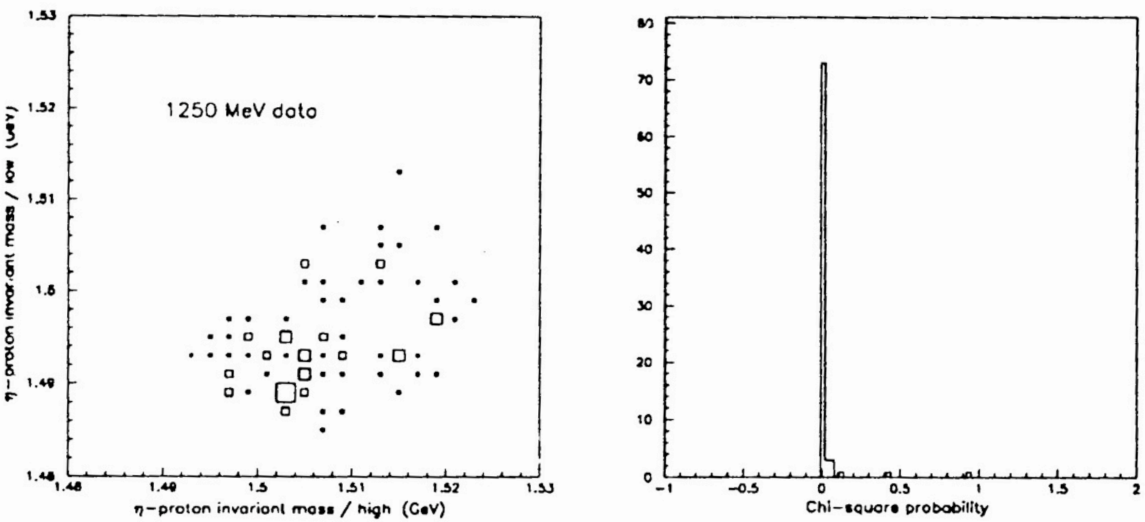

Figure 67. (a) Dalitz plot obtained with 1250 MeV data (below threshold) treated as 1352 MeV η production data. In (b), the corresponding chi-square probability distribution is shown.

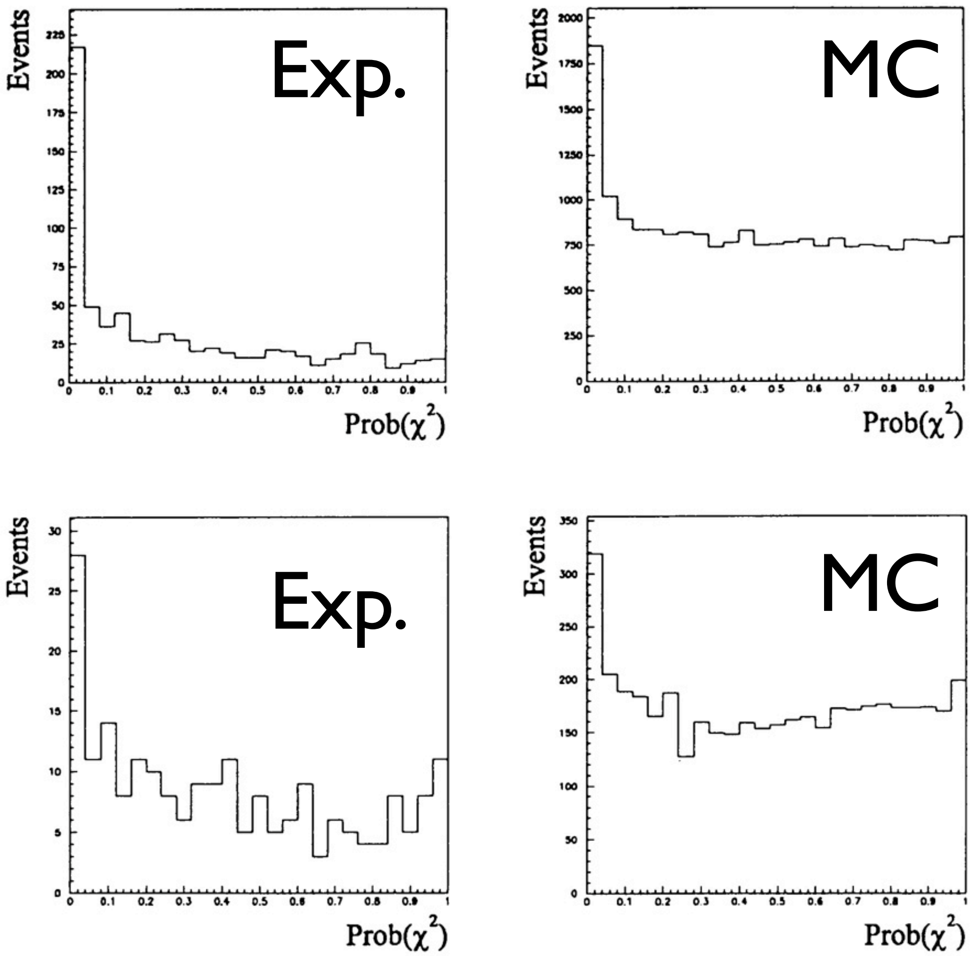

Figure 5.4: The probability distributions for the kinematical fit for η-data. Upper left: data 1350 MeV, upper right: MC 1350 MeV; lower left: data 1296 MeV and lower right: Monte Carlo 1296 MeV.

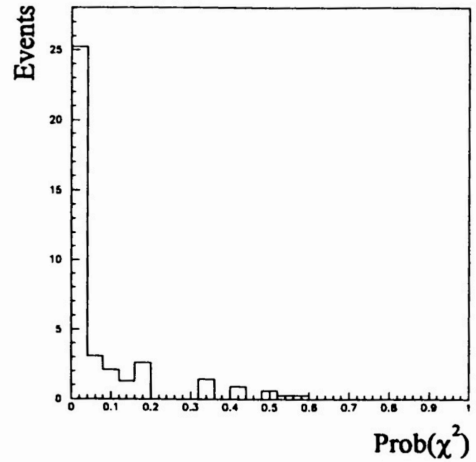

Figure 5.3: The probability distribution for $2\pi^0$-data (generated from Monte Carlo) for the kinematical fit with constraints matching η-data.