Kinematic Fitting and Vertex Fitting

Kanglin He, Xueyao Zhang, Beijiang Liu

email: hekl@mail.ihep.ac.cn

November 4, 2006

Contents

1 Kinematic fitting **1**
1.1 General algorithm for kinematic fitting 1
1.2 Track representation 3
1.3 Kinematic Constraints 3

2 Vertex Fitting **4**
2.1 Formulas for Vertex Fitting 5
2.2 Vertex constraint to a fixed position 6
2.3 Vertex constraints to an unknown position 6
2.4 Swimming the Track Parameters to Vertex Position 6

3 The Decay Vertex and Lifetime Determination **7**

Kinematic/Vertex fitting is a mathematical procedure in which one uses the physical law governing a particle interaction or decay to improve the measurements describing the process. For example, considering the decay chains, $\psi(3770) \to D^+D^-$, then $D^+ \to K^-\pi^+\pi^+$. The fact that the three particles coming from a D^+ decay must come from a common space point can be used to improve the momentum vector of daughter particles, the total energy of D^+ must equal to the beam energy, thus improving the mass and momentum resolution of D^+. These resolution improvements translate to larger signal to background ratio. Kinematic fitting is used in all high energy physics experiments today.

1 Kinematic fitting

1.1 General algorithm for kinematic fitting

The fitting technique is straightforward and is based on the well-known Lagrange multiplier method[1]. It is assumed that the constraint equations can be linearized and summarized in two matrices, **D** and **d**. Let α represent the parameters for a set of n tracks. It has the form of a column vector

$$\alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \tag{1.1}$$

Initially the track parameters have the unconstrained values α_0, obtained from the reconstruction. The **r** functions describing the constraints can be written generally as

$$\mathbf{H}(\alpha) \equiv 0, \quad \text{where } \mathbf{H} = \begin{pmatrix} H_1 & H_2 & \cdots & H_r \end{pmatrix} \tag{1.2}$$

Expanding (1.2) around a convenient point α_A yields the linearized equations

$$0 = \frac{\partial H(\alpha_A)}{\partial \alpha}(\alpha - \alpha_A) + H(\alpha_A) = \mathbf{D}\delta\alpha + \mathbf{d} \tag{1.3}$$

where $\delta\alpha = \alpha - \alpha_A$. Thus we see that

$$\mathbf{D} = \begin{pmatrix} \frac{\partial H_1}{\partial \alpha_1} & \frac{\partial H_1}{\partial \alpha_2} & \cdots & \frac{\partial H_1}{\partial \alpha_n} \\ \frac{\partial H_2}{\partial \alpha_1} & \frac{\partial H_2}{\partial \alpha_2} & \cdots & \frac{\partial H_2}{\partial \alpha_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial H_r}{\partial \alpha_1} & \frac{\partial H_r}{\partial \alpha_2} & \cdots & \frac{\partial H_r}{\partial \alpha_n} \end{pmatrix} \quad \mathbf{d} = \begin{pmatrix} H_1(\alpha_A) \\ H_2(\alpha_A) \\ \vdots \\ H_r(\alpha_A) \end{pmatrix} \tag{1.4}$$

or $D_{ij} = \dfrac{\partial H_i}{\partial \alpha_j}$ and $d_i = H_i(\alpha_A)$. The constraints are incorporated using the method of Lagrange multipliers in which the χ^2 is written as a sum of two term

$$\chi^2 = (\alpha - \alpha_0)^T V_{\alpha_0}^{-1}(\alpha - \alpha_0) + 2\lambda^T(\mathbf{D}\delta\alpha + \mathbf{d}) \tag{1.5}$$

where λ is a vector of **r** unknowns, the Lagrange multipliers. Minimizing the χ^2 with respect to α and λ yields two vector equations which can be solved for parameters α and their covariance matrix:

$$\begin{gathered} V_{\alpha_0}^{-1}(\alpha - \alpha_0) + \mathbf{D}^T\lambda = 0 \\ \mathbf{D}\delta\alpha + \mathbf{d} = 0 \end{gathered} \tag{1.6}$$

The solution cab be written

$$\begin{gathered} \alpha = \alpha_0 - V_{\alpha_0}\mathbf{D}^T\lambda \\ \lambda = V_D(\mathbf{D}\delta\alpha_0 + \mathbf{d}) \\ V_D = \left(\mathbf{D}V_{\alpha_0}\mathbf{D}^T\right)^{-1} \\ V_\alpha = V_{\alpha_0} - V_{\alpha_0}\mathbf{D}^T V_D \mathbf{D} V_{\alpha_0} \\ \chi^2 = \lambda^T V_D^{-1}\lambda = \lambda^T(\mathbf{D}\delta\alpha_0 + \mathbf{d}) \end{gathered} \tag{1.7}$$

where $\delta\alpha_0 = \alpha_0 - \alpha_A$. In the above solution, only a single matrix must be inverted the $\mathbf{r} \times \mathbf{r}$ matrix V_D. The changes to α caused by the constraints are propagated by matrix multiplication. The χ^2 is a sum of $\mathbf{r}$ distinct terms, one per constraint. The contribution of each constraints is correlated with all others through V_D.

It is useful to compute how far the parameters have to move to satisfy a particular constraint j. The initial "distance from satisfaction" can be characterized by the quantity $(\mathbf{D}\delta\alpha_0 + \mathbf{d})_j$ and the number of standard deviations away from the satisfying the constraint is easily calculated to be

$$\sigma_j = \frac{D_{ji}\delta\alpha_{0i} + d_j}{\sqrt{(V_D^-)_{jj}}} \tag{1.8}$$

This information can be used to provide criteria for rejecting background in addition to the overall χ^2.

1.2 Track representation

For kinematic fitting it is important to choose a track representation which uses physically meaningful quantities and is complete. Here, 7-parameter W format, defined as $\alpha_W = (p_x, p_y, p_z, E, x, y, z)$ a 4-momentum and a point where the 4-momentum is evaluated, is adopted in kinematic fitting software package. It's easy to transfer the parameters and their covariance to W representation for neutral tracks. The W format is much simpler to transport particles in a magnetic field, and will be helpful to vertex fitting. It's noted that the W formats also have enough information to represent the general decays of particles.

1.3 Kinematic Constraints

In this section, we'll compute the explicit form of the $\mathbf{D}$ and $\mathbf{d}$ matrices for constraints commonly encountered in high energy physics. If multiple constraints are desired then one just extends the matrices by adding rows to them, one row per constraint. This allow many constraints to be used simultaneously in the fit.

1) Invariant mass constraint

$$\begin{gathered} H = E^2 - p_x^2 - p_y^2 - p_z^2 - m_c^2 = 0 \\ \mathbf{D} = \begin{pmatrix} -2p_x & -2p_y & -2p_z & 2E & 0 & 0 & 0 \end{pmatrix} \\ \mathbf{d} = E^2 - p_x^2 - p_y^2 - p_z^2 - m_c^2 \end{gathered} \tag{1.9}$$

2) Total energy constraint

$$\begin{gathered} H = E - E_c = 0 \\ \mathbf{D} = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \\ \mathbf{d} = E - E_c \end{gathered} \tag{1.10}$$

3) Total momentum constraint

$$\begin{gathered} H = \sqrt{p_x^2 + p_y^2 + p_z^2} - p_c = 0 \\ \mathbf{D} = \begin{pmatrix} p_x/p & p_y/p & p_z/p & 0 & 0 & 0 & 0 \end{pmatrix} \\ \mathbf{d} = \sqrt{p_x^2 + p_y^2 + p_z^2} - p_c \end{gathered} \tag{1.11}$$

4) Total 3-vector constraint

$$
\begin{aligned}
H &= \begin{pmatrix} p_x - p_{cx} \\ p_y - p_{cy} \\ p_z - p_{cz} \end{pmatrix} = 0 \\
\mathbf{D} &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \\
\mathbf{d} &= \begin{pmatrix} p_x - p_{cx} \\ p_y - p_{cy} \\ p_z - p_{cz} \end{pmatrix}
\end{aligned}
\tag{1.12}
$$

5) Total 4-vector constraint

$$
\begin{aligned}
H &= \begin{pmatrix} p_x - p_{cx} \\ p_y - p_{cy} \\ p_z - p_{cz} \\ E - E_c \end{pmatrix} = 0 \\
\mathbf{D} &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \\
\mathbf{d} &= \begin{pmatrix} p_x - p_{cx} \\ p_y - p_{cy} \\ p_z - p_{cz} \\ E - E_c \end{pmatrix}
\end{aligned}
\tag{1.13}
$$

6) Equal mass constraint

$$
\begin{aligned}
H &= (E_1^2 - p_{1x}^2 - p_{1y}^2 - p_{1z}^2) - (E_2^2 - p_{2x}^2 - p_{2y}^2 - p_{2z}^2) \\
\mathbf{D}_1 &= \begin{pmatrix} -2p_{1x} & -2p_{1y} & -2p_{1z} & 2E_1 & 0 & 0 & 0 \end{pmatrix} \\
\mathbf{D}_2 &= \begin{pmatrix} 2p_{2x} & 2p_{2y} & 2p_{2z} & 2E_2 & 0 & 0 & 0 \end{pmatrix} \\
\mathbf{d} &= (E_1^2 - p_{1x}^2 - p_{1y}^2 - p_{1z}^2) - (E_2^2 - p_{2x}^2 - p_{2y}^2 - p_{2z}^2)
\end{aligned}
\tag{1.14}
$$

Once these matrices are known the tracks can be kinematically fit using the procedure described in the previous section.

2 Vertex Fitting

Consider a set of **n** tracks forced to pass through a common vertex $\mathbf{x} = (x, y, z)$. The covariance matrix of the vertex may be known in advance, the overall χ^2 can be written as a general form

$$
\chi^2 = (\alpha - \alpha_0)^T V_{\alpha_0}^{-1} (\alpha - \alpha_0) + (x - x_0)^T V_{x0}^{-1} (x - x_0) + 2\lambda^T (\mathbf{D}\delta\alpha + \mathbf{E}\delta x + \mathbf{d}) \tag{2.1}
$$

where the terms represent, respectively, the contribution from tracks, vertex and the constraints, x_0 and V_{x0} represent the initial vertex position and its covariance matrix, $\delta\alpha = \alpha - \alpha_A$ and $\delta x = x - x_A$, $E_{ij} = \dfrac{\partial H_i}{\partial x_j}$.

2.1 Formulas for Vertex Fitting

For each track i there are two constraint equations, corresponding to the bend and non-bend plane. The momentum and position components of a charged particle in a magnetic field can be written as:

$$\begin{gathered}
\hat{p} = -(\hat{p_0} \times \hat{h}) \times \hat{h} \cos \rho s - \hat{p_0} \times \hat{h} \sin \rho s + (\hat{p_0} \cdot \hat{h})\hat{h} \\
\hat{x} = \hat{x_0} - \frac{(\hat{p_0} \times \hat{h}) \times \hat{h}}{a} \sin \rho s - \frac{\hat{p_0} \times \hat{h}}{a}(1 - \cos \rho s) + \frac{(\hat{p_0} \cdot \hat{h})s}{p}\hat{h} \\
\hat{p} = \hat{p_0} - a(\hat{x} - \hat{x_0}) \times \hat{h}
\end{gathered} \tag{2.2}$$

where $\hat{p} = (p_x, p_y, p_z$, $\hat{x} = (x, y, z)$ and $\hat{h}$ is a unit vector pointing along the direction of the magnetic field, $a = -0.00299792458Bq$, B is the magnetic field strength in Tesla, q is the charge, $\rho = a/p$, s is the arc length measured from point $\hat{x_0}$ to $\hat{x}$.

From (2.2), we can get the vertex constraints

$$\begin{gathered}
0 = (\hat{p_i} \times \Delta\hat{x_i}) \cdot \hat{h} - \frac{a_i}{2}\left(\Delta\hat{x_i}^2 - (\Delta\hat{x_i} \cdot \hat{h})^2\right) \\
0 = \Delta\hat{x_i} \cdot \hat{h} - \frac{\hat{p_i} \cdot \hat{h}}{a_i} \sin^{-1}\left[a_i \frac{\left(\hat{p_i} \cdot \Delta\hat{x_i} - (\hat{p_i} \cdot \hat{h})(\Delta\hat{x_i} \cdot \hat{h})\right)}{\left|\hat{p_i} \times \hat{h}\right|^2}\right]
\end{gathered} \tag{2.3}$$

where $\Delta\hat{x_i} = \hat{x_x} - \hat{x_i}$. The **E** and **D** matrices have the simple form

$$\mathbf{E} = \begin{pmatrix} E_1 \\ E_2 \\ \vdots \\ E_n \end{pmatrix} \quad \mathbf{D} = \begin{pmatrix} D_1 & & & \\ & D_2 & & \\ & & \ddots & \\ & & & D_n \end{pmatrix} \tag{2.4}$$

where E_i is a 2×3 matrix and D_i is a 2×7 matrix. **E** and **D** have this particular block diagonal form because the vertex constraints for each track only involve the parameters for that track. It can greatly simplify the solution of vertex fitting, the inversion of $2n \times 2n$ matrix V_D can be factored into n 2×2 inversion.

In a solenoidal field, the matrices $\mathbf{D}_i$, $\mathbf{E}_i$ and $\mathbf{d}_i$ for each track are given by

$$\begin{gathered}
\mathbf{D}_i = \begin{pmatrix} \Delta y_i & -\Delta x_i & 0 & 0 & p_{yi} + a_i\Delta x_i & -p_{xi} + a_i\Delta y_i & 0 \\ -p_{zi}S_iR_{xi} & -p_{zi}S_iR_{yi} & -\dfrac{\sin^{-1} J_i}{a_i} & 0 & p_{xi}p_{zi}S_i & p_{yi}p_{zi}S_i & -1 \end{pmatrix} \\
\mathbf{E}_i = \begin{pmatrix} -(p_{yi} + a_i\Delta x_i) & p_{xi} - a_i\Delta y_i & 0 \\ -p_{xi}p_{zi}S_i & -p_{yi}p_{zi}S_i & 1 \end{pmatrix} \\
\mathbf{d}_i = \begin{pmatrix} \Delta y_i p_{xi} - \Delta x_i p_{yi} - \dfrac{a_i}{2}(\Delta x_i^2 + \Delta y_i^2) \\ \Delta z_i - \dfrac{p_{zi}}{a_i} \sin^{-1} J_i \end{pmatrix}
\end{gathered} \tag{2.5}$$

where the auxiliary quantities J, R_x, R_y and S are defined as follows:

$$\begin{aligned} J &= a\frac{\Delta x p_x + \Delta y p_y}{p_T^2} \\ R_{x(y)} &= \Delta x(y) - 2p_{x(y)}\frac{\Delta x p_x + \Delta y p_y}{p_T^2} \\ S &= \frac{1}{p_T^2\sqrt{1-J^2}} \end{aligned} \tag{2.6}$$

The solution of (2.1) is straightforward, but algebraically tedious:

$$\begin{aligned} \delta\alpha &= \delta\alpha_0 - V_{\alpha_0}\mathbf{D}^T\lambda \\ \delta x &= \delta x_0 - V_{x0}\mathbf{E}^T\lambda = V_x V_{x0}^{-1}\delta x_0 - V_x\mathbf{E}^T\lambda_0 \\ \lambda &= V_{\tilde{D}}(\mathbf{D}\delta\alpha_0 + \mathbf{E}\delta x_0 + \mathbf{d}) = \lambda_0 + V_D\mathbf{E}\delta x \\ \lambda_0 &= V_D(\mathbf{D}\delta\alpha_0 + \mathbf{d}) \\ V_{\tilde{D}} &= V_D - V_D\mathbf{E}V_x\mathbf{E}^T V_D \\ V_{Di} &= (\mathbf{D}_i V_{0i}\mathbf{D}_i^T)^{-1} \end{aligned} \tag{2.7}$$

where $\delta\alpha = \alpha - \alpha_A$ and $\delta x = x - x_A$. The covariance matrices are

$$\begin{aligned} V_x &= (V_{x_0}^{-1} + \mathbf{E}^T V_D\mathbf{E})^{-1} \\ V_\alpha &= V_{\alpha_0} - V_{\alpha_0}\mathbf{D}^T V_D\mathbf{D}V_{\alpha_0} + V_{\alpha_0}\mathbf{D}^T V_D\mathbf{E}V_x\mathbf{E}^T V_D\mathbf{D}V_{\alpha_0} \\ \mathrm{cov}(\alpha, x) &= -V_{\alpha_0}\mathbf{D}^T V_D\mathbf{E}V_x \end{aligned} \tag{2.8}$$

The vertex err matrix V_x is the weighted average of its initial covariance matrix and the errors determined from the tracks. The track error matrix V_α has an initial piece that is decreased by the constraints applied per track and is increased by the wiggle of the vertex. $\mathrm{cov}(\alpha, x)$ is the correlation between tracks introduced by the vertex constraint. The χ^2 is given by

$$\chi^2 = \lambda^T(\mathbf{D}\delta\alpha_0 + \mathbf{E}\delta x_0 + \mathbf{d}) \tag{2.9}$$

2.2 Vertex constraint to a fixed position

In this case, the vertex position is fixed and the solution can be obtained by setting $\delta x_0 = 0$, $\mathbf{E}$=0 and $V_{x_0} = 0$ in (2.7),(2.8) and (2.9). The solution factors into n pieces, one per track.

2.3 Vertex constraints to an unknown position

In this case the vertex position $\mathbf{x}$ must be determined from the constraints. The simplest approach is to assign large value to V_{x_0}, the initial vertex convariance matrix, and apply the method from (2.7),(2.8) and (2.9).

2.4 Swimming the Track Parameters to Vertex Position

After the vertex fitting, track parameters and their covariance matrix should be updated to the vertex position. The track parameters can be written in matrix form, following (2.2)

$$\alpha_V = \begin{pmatrix} p_V \\ x_V \end{pmatrix} = \begin{pmatrix} \mathbf{A}\alpha + \mathbf{B}x \\ x \end{pmatrix} \tag{2.10}$$

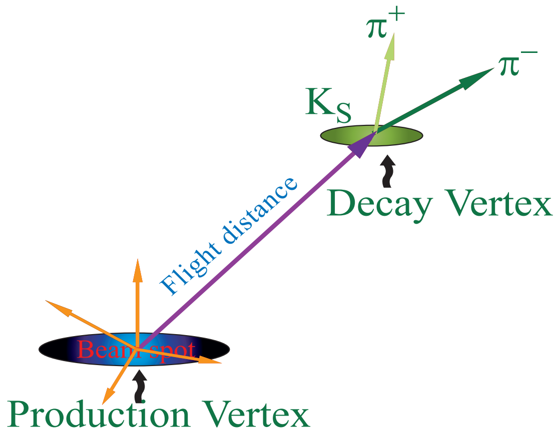

Figure 1: K^0_S travels a certain distance("the flight distance") before decaying into its daughters. These daughters are subsequently measured by the tracking system.

where

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & a_z & -a_y \\ 0 & 1 & 0 & 0 & -a_z & 0 & a_x \\ 0 & 0 & 1 & 0 & a_y & -a_x & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{2.11}$$

The covariance matrix is much complicated, it can be written as

$$\begin{gathered} V_{\alpha_V} = \begin{pmatrix} V_{p_V} & \mathrm{cov}(p_V, x_V) \\ \mathrm{cov}(x_V, p_V) & V_{x_V} \end{pmatrix} \\ V_{p_V} = \mathbf{A} V_\alpha \mathbf{A}^T + \mathbf{A}\mathrm{cov}(\alpha, x)\mathbf{B}^T + \mathbf{B}\mathrm{cov}(x, \alpha)\mathbf{A}^T + \mathbf{B} V_x \mathbf{B}^T \\ \mathrm{cov}(p_V, x_V) = \mathbf{A}\mathrm{cov}(\alpha, x) + \mathbf{B} V_x \\ V_{x_V} = V_x \end{gathered} \tag{2.12}$$

3 The Decay Vertex and Lifetime Determination

To introduce the subject of lifetime determination, consider Figure 1, a K^0_S decays to $\pi^+\pi^-$ at a secondary vertex after being produced in the beam interaction region. An accurate determination of the lifetime requires that both the beginning and endpoint of the K^0_S flight vector be determined accurately. The endpoint is determined by vertex fitting, and its measurement accuracy is controlled purely by the tracking error of the daughter particles. The beginning point is determined by the beam spot size augmented perhaps, as shown in Figure 1, by other tracks produced in IP, or by the average of preliminary vertex for lot of events.

The motion of a neutral track before decay is a simple linear equation. For a charged track, it travels within a helix orbit (2.2) in a solenoidal magnetic field. To eliminate the flight distance (s) measured from the point (x_p, y_p, z_p) to (x_d, y_d, z_d), by the proper lifetime

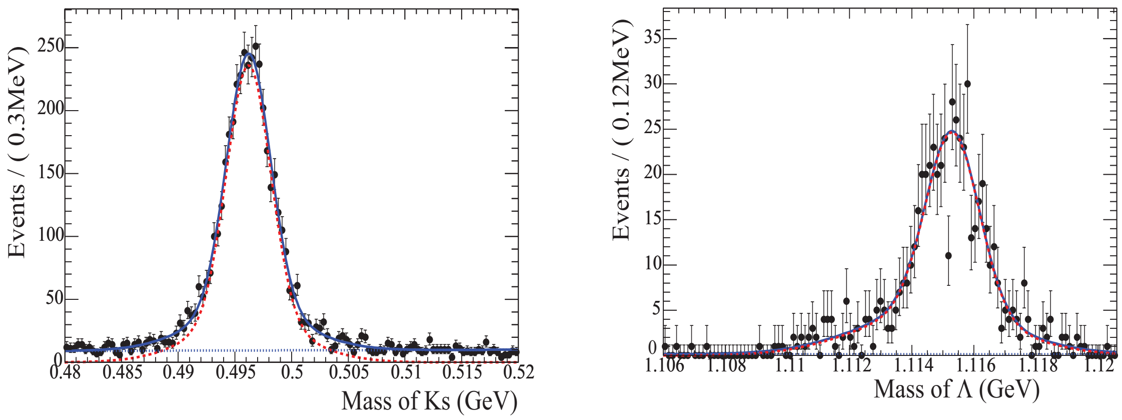

Figure 2: The mass distribution of reconstructed K_S^0 and Λ.

$c\tau$ using $s = \beta ct = \gamma\beta c\tau = (p/m)c\tau$, yielding the new equations

$$\begin{pmatrix} x_p \\ y_p \\ z_p \end{pmatrix} = \begin{pmatrix} x_d - \frac{p_x}{m}\sin\left(\frac{ac\tau}{m}\right) - \frac{p_y}{a}\left(1 - \cos\left(\frac{ac\tau}{m}\right)\right) \\ y_d - \frac{p_y}{m}\sin\left(\frac{ac\tau}{m}\right) + \frac{p_x}{a}\left(1 - \cos\left(\frac{ac\tau}{m}\right)\right) \\ z_d - \frac{p_z}{m}c\tau \end{pmatrix} \quad \begin{pmatrix} x_p \\ y_p \\ z_p \end{pmatrix} = \begin{pmatrix} x_d - \frac{p_x}{m}c\tau \\ y_d - \frac{p_y}{m}c\tau \\ z_d - \frac{p_z}{m}c\tau \end{pmatrix} \tag{3.1}$$

The lifetime $c\tau$ is determined by recognizing (3.1) represent constraint conditions. We can apply

$$\chi^2 = (\alpha - \alpha_0)^T V_{\alpha_0}^{-1}(\alpha - \alpha_0) + 2\lambda^T(\mathbf{D}\delta\alpha + \mathbf{E}\delta c\tau + \mathbf{d}) \tag{3.2}$$

to solve for $c\tau$ and its error, while at the same time improving the track parameters and the start point. The vector $\alpha = (p_x, p_y, p_z, E, x_d, y_d, z_d, x_p, y_p, z_p)^T$ contains 10 variables, the 7 track parameters at the decay point and the 3 for production point.

The solution for $c\tau$ and its covariance matrix is straightforward

$$\begin{gathered} \delta c\tau = -V_{c\tau}\mathbf{E}^T\lambda_0 \\ V_{c\tau} = (\mathbf{E}^T V_D \mathbf{E})^{-1} \\ \lambda_0 = V_D(\mathbf{D}\delta\alpha_0 + \mathbf{d}) \\ V_D = (\mathbf{D}V_{\alpha_0}\mathbf{D}^T)^{-1} \end{gathered} \tag{3.3}$$

The updated α, covariance matrix V_α and the correlation with $c\tau$:

$$\begin{gathered} \alpha = \alpha_0 - V_{\alpha_0}\mathbf{D}^T\lambda \\ \lambda = \lambda_0 + V_D\mathbf{E}\delta c\tau \\ V_\alpha = V_{\alpha_0} - V_{\alpha_0}\mathbf{D}^T V_D \mathbf{D} V_{\alpha_0} + V_{\alpha_0}\mathbf{D}^T V_D \mathbf{E} V_{c\tau}\mathbf{E}^T V_D \mathbf{D} V_{\alpha_0} \\ \mathrm{cov}(\alpha, c\tau) = -V_{\alpha_0}\mathbf{D}^T V_D \mathbf{E} V_{c\tau} \\ \chi^2 = \lambda^T(\mathbf{D}\delta\alpha_0 + \mathbf{d}) \end{gathered} \tag{3.4}$$

Figure 2 shows the mass distribution of K_S^0 and Λ after the secondary vertex reconstruction and selection. The mass resolution is about 3MeV for K_S, and about 1.2MeV for Λ, respectively.

References

[1] Paul Avery, Applied Fitting Theory I: General Least Squares Theory, CBX-91-72.