Decay Chain Fitting with a Kalman Filter

Walter Ikegami Andersson

Uppsala University

PANDA Collaboration Meeting
November 5-9, 2018
GSI

Outline

- Introduction to traditional kinematic fitting formalism
- Motivation for alternative tool
- Decay Chain Fitting
  - Measurement constraints
  - Exact constraint
  - Examples
- Summary

Standard Kinematic Fit Formalism

Standard kinematic fit procedure

$$\chi^2 = (x - x_{\mathrm{meas}})^T C_{\mathrm{meas}}^{-1} (x - x_{\mathrm{meas}}) + 2\lambda^T g(x)$$

- x - vector of $(p_x, p_y, p_z, E, x, y, z)$ of all particles being fitted
- C_{meas} - covariance matrix of measured x_{meas}
- $g(x)$ - vector of exact constraints, should satisfy $g_i(x) = 0$
  e.g. mass constraint, vertex constraint, four-momentum constraint...
- λ - Lagrange multipliers, one introduced for each constraint g_i

Standard Least Squares Fit

A successful fit requires $\partial\chi^2/\partial x = 0$ (here, λ added to x list). Using *e.g.* Newton-Raphsons method, improved parameters are obtained from

$$x^{(1)} = x^{(0)} - \left(\frac{\partial^2\chi^2}{\partial x^2}\right)^{-1} \frac{\partial\chi^2}{\partial x}$$

The final covariance of the fitted x is given by

$$C(x) = 2\left(\frac{\partial^2\chi^2}{\partial x^2}\right)^{-1}$$

Standard Least Squares Fit

A successful fit requires $\partial\chi^2/\partial x = 0$ (here, λ added to x list). Using *e.g.* Newton-Raphsons method, improved parameters are obtained from

$$x^{(1)} = x^{(0)} - \left(\frac{\partial^2\chi^2}{\partial x^2}\right)^{-1}\frac{\partial\chi^2}{\partial x}$$

The final covariance of the fitted x is given by

$$C(x) = 2\left(\frac{\partial^2\chi^2}{\partial x^2}\right)^{-1}$$

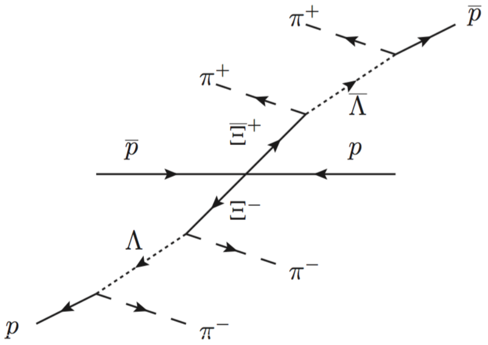
Computationally demanding to invert square matrices of dimension x

Standard Least Squares Fit

Example of complicated decay tree
$\overline{p}p \to \overline{\Xi}^+\Xi^-$,
$\Xi^- \to \Lambda\pi^-$, $\Lambda \to p\pi^-$ + c.c.

Traditional leaf-by-leaf fitting:

- Vertex fit $p\pi^-$
- Mass constraint Λ
- Vertex fit $\Lambda\pi^+$
- Repeat for Ξ^+
- 4C fit with constraint to initial system

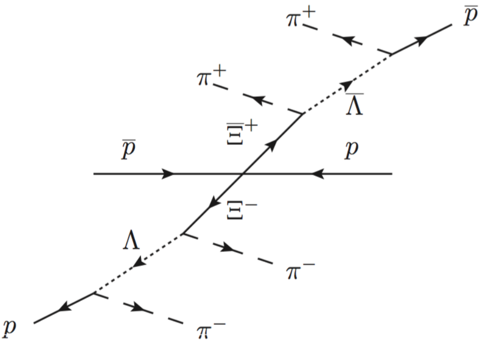


Standard Least Squares Fit

Example of complicated decay tree
$\overline{p}p \to \overline{\Xi}^{+}\Xi^{-}$,
$\Xi^{-} \to \Lambda\pi^{-}$, $\Lambda \to p\pi^{-}$ + c.c.

Traditional leaf-by-leaf fitting:

- Vertex fit $p\pi^{-}$
- Mass constraint Λ
- Vertex fit $\Lambda\pi^{+}$
- Repeat for Ξ^{+}
- 4C fit with constraint to initial system



Introduce alternative tool to apply all constraints simultaneously!

Progressive Fit with Kalman Filter

Motivation

- Fit whole event topology in one step, all constraints applied simultaneously *e.g.* mass constraint, vertex, constraint to p_{ini}
- Computationally expensive due to large number of fitting parameters

$\rightarrow$ Kalman filter for faster performance

Decay Chain Fit

Tool developed by Wouter D. Hulsbergen for BaBar analysis.

- Decay Chain Fit calculates χ_k^2 constributions from constraints k one-by-one
- Once constraint contribution minimized, add to total $\chi^2 = \sum \chi_k^2$

Measurement constraints

Progressive fit utilizes constraints from measurements. Measured parameters m are helix parameters from a track fit or energy deposits in calorimeters.

Define residual

$$r(x) = m - h(x)$$

where $h(x)$ is measurement model of m

Meaurement constraints introduce residual terms in χ^2

$$\chi^2 = r^T(x) V_{\mathrm{meas}}^{-1} r(x)$$

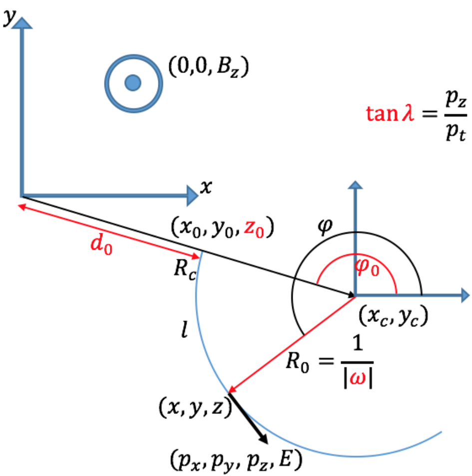
where V_{meas}^{-1} is the covariance of the measurement m

Measurement constraints

Charged track represented by 5-parameter helix

$$m_{\text{charged}} = \begin{pmatrix} d_0 \\ \phi_0 \\ \omega \\ z_0 \\ \tan\lambda \end{pmatrix}$$

$$h_{\text{charged}} = \begin{pmatrix} (R_c - R_0) \\ \operatorname{atan2}(x_c, y_c) \\ qB_z c/p_t \\ z - lp_z/p_t \\ p_z/p_t \end{pmatrix}$$



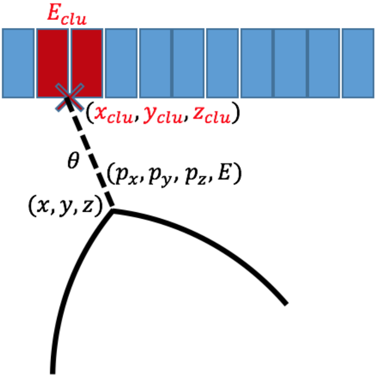
Measurement constraints

Photons represented by cluster position and energy deposit

$$m_{\text{photon}} = \begin{pmatrix} x_{\text{clu}} \\ y_{\text{clu}} \\ z_{\text{clu}} \\ E_{\text{clu}} \end{pmatrix},$$

$$h_{\text{photon}} = \begin{pmatrix} x + \theta p_x \\ y + \theta p_y \\ z + \theta p_z \\ \sqrt{p_x^2 + p_y^2 + p_z^2} \end{pmatrix}$$

where θ is the photon "decay length"



Progressive Least Squares Fit: Measurement Constraint Contribution

Consider a measurement constraint k. A χ^2 containing contributions from $\{0, ..., k-1\}$ constraints has been minimized.

Define χ^2 constribution

$$\chi_k^2 = (x - x_{k-1})^T C_{k-1}^{-1} (x - x_{k-1}) + (h_k(x) - m_k)^T V_k^{-1} (h_k(x) - m_k)$$

Solution for x obtained from $\partial \chi^2 / \partial x = 0$ *i.e.*

$$C_{k-1}^{-1} (x - x_{k-1}) + H_k^T V_k^{-1} (h_x(x) - m_k) = 0$$

where $H_k = \partial h / \partial x |_{x_{k-1}}$

Progressive Least Squares Fit: Measurement Constraint Contribution

Defining the residual of the prediction $r_k^{k-1} = m_k - h_k(x_{k-1})$ and assuming h is linear

$$x_k = x_{k-1} + K_k r_k^{k-1}$$

where

$$K_k = C_{k-1} H_k^T (R_k^{k-1})^{-1}$$

is the gain matrix and

$$R_k^{k-1} = V_k + H_k C_{k-1} H_k^T$$

is the uncertainty in the predicted residual

Progressive Least Squares Fit: Measurement Constraint Contribution

The new covariance matrix of the parameters x is given by

$$C_k = C_{k-1} - K_k(2H_k C_{k-1} - R_k^{k-1} K_k^T)$$

Finally, the χ_k^2 contribution of constraint m_k is given by

$$\chi_k^2 = (r_k^{k-1})^T (R_k^{k-1})^{-1} r_k^{k-1}$$

Progressive Least Squares Fit: Measurement Constraint Contribution

The new covariance matrix of the parameters x is given by

$$C_k = C_{k-1} - K_k(2H_k C_{k-1} - R_k^{k-1} K_k^T)$$

Finally, the χ^2_k contribution of constraint m_k is given by

$$\chi^2_k = (r_k^{k-1})^T (R_k^{k-1})^{-1} r_k^{k-1}$$

The Kalman filter does not require inversion of matrices of dimension x, but rather m_k

Progressive Least Squares Fit: Exact Constraint Contribution

For an exact constraint $g_k(x)$, the χ^2 contribution is given by

$$\chi_k^2 = (x - x_{k-1})^T C_{k-1}^{-1} (x - x_{k-1}) + 2\lambda_k^T g_k(x)$$

Solution given by solving $\partial\chi^2/\partial x = 0$ $\partial\chi^2/\partial\lambda = 0$. Linearizing constraint around x_{k-1}

$$C_{k-1}^{-1}(x - x_{k-1}) + G_k^T \lambda_k = 0$$

$$g_k(x_{k-1}) + G_k(x - x_{k-1}) = 0$$

where

$$G_k = \partial g/\partial x$$

Progressive Least Squares Fit: Exact Constraint Contribution

The solution is given by, after some rewriting

$$x_k = x_{k-1} - K_k g_k(x_{k-1})$$

where the gain matrix is

$$K_k = C_{k-1} G_k^T (G_k C_{k-1} G_k^T)^{-1}$$

Updated covariance is given by

$$C_k = (1 - K_k G_k) C_{k-1} (1 - K_k G_k)^T$$

And the χ^2 contribution for the exact constraint is

$$\chi_k^2 = g_k(x_{k-1})^T (G_k C_{k-1} G_k^T)^{-1} g_k(x_{k-1})$$

Progressive Least Squares Fit: Exact Constraint Contribution

The solution is given by, after some rewriting

$$x_k = x_{k-1} - K_k g_k(x_{k-1})$$

where the gain matrix is

$$K_k = C_{k-1} G_k^T (G_k C_{k-1} G_k^T)^{-1}$$

Updated covariance is given by

$$C_k = (1 - K_k G_k) C_{k-1} (1 - K_k G_k)^T$$

And the χ^2 contribution for the exact constraint is

$$\chi_k^2 = g_k(x_{k-1})^T (G_k C_{k-1} G_k^T)^{-1} g_k(x_{k-1})$$

The Kalman filter eliminates λ. Additional fit parameters in standard case

Ordering Constraints

Progressive fit calculates χ^2 constributions from one constraint at the time.

Order of the constraints plays a role if they are non-linear

- Reconstructed tracks and clusters
- Constraint to interaction point
- Vertex constraints
- Geometric constraints
- Mass constraints

Ordering Constraints

Progressive fit calculates χ^2 constributions from one constraint at the time.

Order of the constraints plays a role if they are non-linear

- Reconstructed tracks and clusters
- Constraint to interaction point
- Vertex constraints
- Geometric constraints
- Mass constraints

General rule-of-thumb: Add linear constraints first, add non-linear ones last!

Fit Initialization

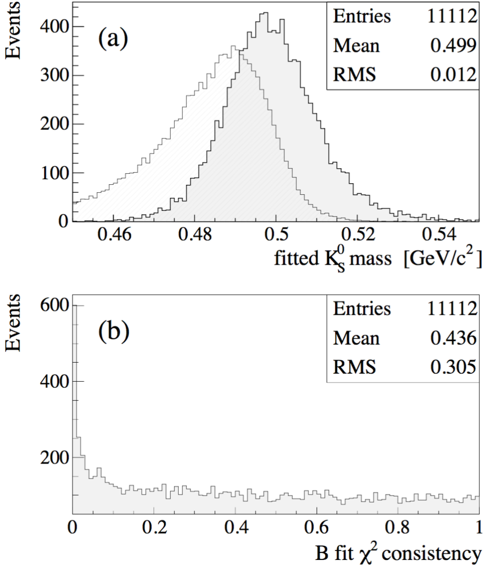
The parameters x_0 and corresponding covariance matrix C_0 must be initialized in the fit. Initialized in the following order:

- Vertex intitalized with POCA of daughter tracks
- Momenta of daughter tracks evaluated at vertex
- Photon momenta calculated using initial vertex as origin
- Composite particle momenta initialized by adding momenta of daughters
- Decay time parameters initialized using initial vertex and momentum

The covariance matrix C_0 is initialized as a diagonal matrix with elements 1000 times larger than typical resolutions of corresponding parameters.

Example: BaBar

Reconstructing $K_S^0 \to \pi^0\pi^0$ is challenging due to displaced vertex and photons in the final state



BaBar example:
$e^+e^- \to \Upsilon(4S)$
$\Upsilon(4S) \to B^0\overline{B}^0$
Fit hypothesis:
$B^0 \to J/\psi K_S^0$,
$J/\psi \to \mu^+\mu^-$, $K_S^0 \to \pi^0\pi^0$

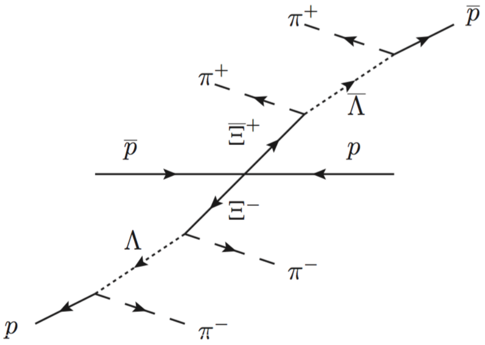
Example: $\overline{\mathrm{P}}$ANDA

$\overline{\mathrm{P}}$ANDA example:
$\bar{p}p \to \overline{\Xi}^{+}\Xi^{-}$, $\Xi^{-} \to \Lambda\pi^{-}$,
$\Lambda \to p\pi^{-}$ + c.c.
Fitting each Ξ separately

- Two displaced vertices
- Mass constraint on Λ
- Three degrees of freedom

Fitting whole decay chain

- Four displaced vertices
- Mass constraints on $\overline{\Lambda}$,Λ
- 4C constraint from $\bar{p}p$
- 11 degrees of freedom

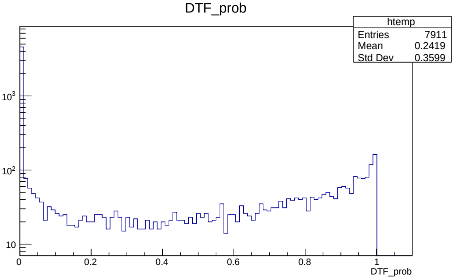


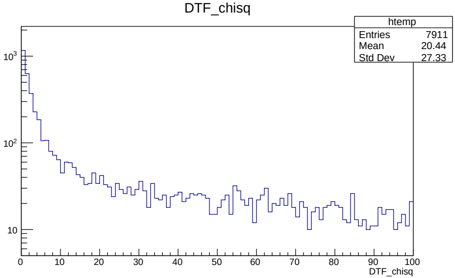
Probability and χ^2 distribution in $\Xi^- \to \Lambda\pi^- \to p\pi^-\pi^-$

- Probability distribution should be flat

There are issues with the fit if:

- Peaks at 0 in probability distribution $\implies$ bad events/poor convergence
- Skews towards higher (lower) values $\implies$ errors over(under)estimated

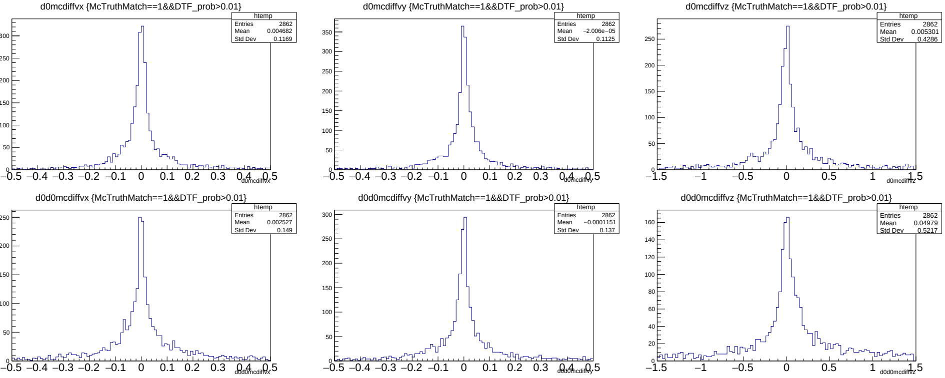




Vertex position resolution in $\Xi^- \to \Lambda\pi^- \to p\pi^-\pi^-$

- Filter out combinatorial background
- Cut on DTF probability

Vertex resolution similar to conventional vertex fits



Summary

- Decay Chain Fitting - tool developed at BaBar for complicated decay chains
  - Can fit any decay tree (as long as d.o.f > 0)
  - Vertex fit, mass constraint, can be done simultaneously
  - Possible to constraint head of tree to *e.g.* initial system
- Due to large number of parameters, Kalman Filter is employed for faster minimization of χ^2

https://arxiv.org/abs/physics/0503191

Summary

- Decay Chain Fitting - tool developed at BaBar for complicated decay chains
  - Can fit any decay tree (as long as d.o.f > 0)
  - Vertex fit, mass constraint, can be done simultaneously
  - Possible to constraint head of tree to *e.g.* initial system
- Due to large number of parameters, Kalman Filter is employed for faster minimization of χ^2

Thank you for your attention!

https://arxiv.org/abs/physics/0503191

Backup

Vertex position resolution in $\Xi^- \to \Lambda\pi^- \to p\pi^-\pi^-$

- Filter out combinatorial background
- Cut on both vertex fit and mass fit probabilities