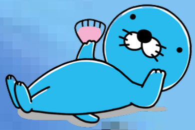


Notes on Kinematic Fit

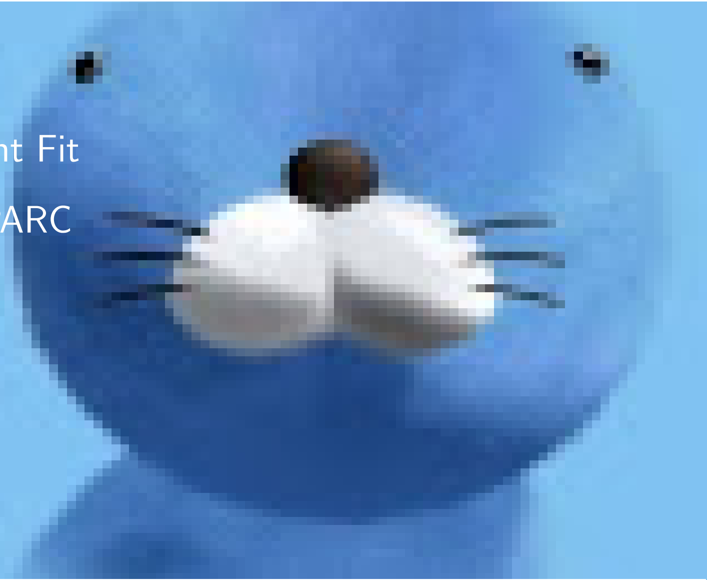
Bono, Bono[1]

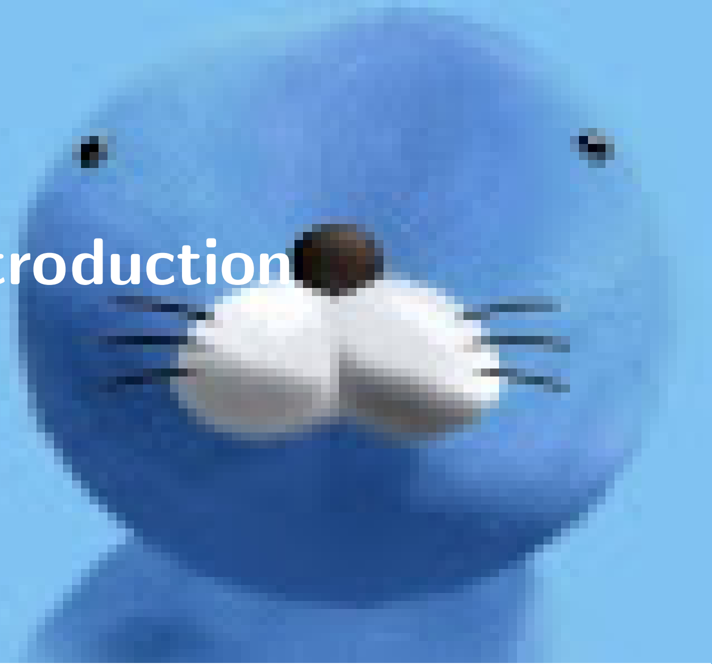
October 3, 2025

[1]Bono University

Outline

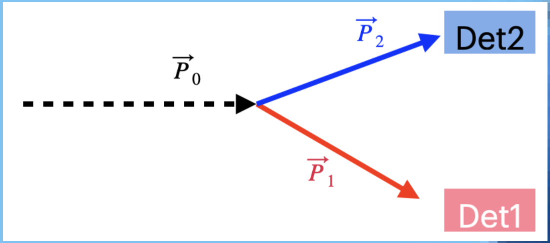
1. Introduction
2. Example: Mass-Constraint Fit
3. Example: Ξ decay at J-PARC
4. Tricks
5. Calculations





Introduction

Measurement Error



Assume a beam with momentum $\vec{P}^{-}$ decays into $\vec{P}_1$ and $\vec{P}_2$. Measured momentum are smeared due to detector resolution, leading to unbalance in the momentum conservation.

$$\vec{P}_0 = \vec{P}_1 + \vec{P}_2; \quad \vec{P}_{1,meas} + \vec{P}_{2,meas} \neq \vec{P}_0 \tag{1}$$

We can define the χ^2 to quantitatively represent our measurement error. However, we can't derive meaningful expressions from this χ^2.

$$\chi^2 = \frac{(P_1 - P_{1,meas})^2}{\sigma_1^2} + \frac{(P_2 - P_{2,meas})^2}{\sigma_2^2} \tag{2}$$

Constrained Optimization with The Lagrange Multiplier

By incorporating the *Kinematic Constraints*, specifically *momentum conservation*, we involve additional knowledge to (2). This is known as the *Lagrange Multiplier*

$$\chi^2 = \frac{(P_{1,KF} - P_{1,meas})^2}{\sigma_1^2} + \frac{(P_{2,KF} - P_{2,meas})^2}{\sigma_2^2} + 2\lambda(\mathbf{P_{1,KF}} + \mathbf{P_{2,KF}} - \mathbf{P_0}) \tag{3}$$

Now we have meaningful expressions to minimize χ^2, hence get better estimations for the measurement.

$$\frac{1}{2}\frac{\partial \chi 2}{\partial P_{1,KF}} = \frac{(P_{1,KF} - P_{1,meas})}{\sigma_1^2} + \lambda = 0 \tag{4}$$

$$\frac{1}{2}\frac{\partial \chi 2}{\partial P_{2,KF}} = \frac{(P_{2,KF} - P_{2,meas})}{\sigma_2^2} + \lambda = 0 \tag{5}$$

$$\frac{1}{2}\frac{\partial \chi 2}{\partial \lambda} = (P_{1,KF} + P_{2,KF} - P_0) = 0 \tag{6}$$

Why Better Resolution?

By solving the equations 4,5,6 and defining $\delta_i = P_{i,meas} - P_i$, we obtain the following expressions:

$$\lambda = \frac{P_{1,meas} + P_{2,meas} - P_0}{\sigma_1^2 + \sigma_2^2} = \frac{\delta_1 + \delta_2}{\sigma_1^2 + \sigma_2^2} \tag{7}$$

$$P_{1,KF} = P_{1,meas} - \sigma_1^2 \lambda \tag{8}$$

$$P_{2,KF} = P_{2,meas} - \sigma_2^2 \lambda \tag{9}$$

$$\begin{aligned} < P_{1,KF} - P_1 > &=< P_{1,KF} - P_{1,meas} + \delta_1 >=< -\sigma_1^2 \lambda + \delta_1 > \\ &=< \frac{-\sigma_1^2}{\sigma_1^2 + \sigma_2^2}(\delta_1 + \delta_2) + \delta_1 >=< \frac{\sigma_2^2 \delta_1 - \sigma_1^2 \delta_2}{\sigma_1^2 + \sigma_2^2} > \end{aligned} \tag{10}$$

$$\sigma_{1,KF}^2 =< (P_{1,KF} - P_1)^2 >= \frac{\sigma_2^4 \cancelto{\sigma_1^2}{< \delta_1^2 >} + \sigma_1^4 \cancelto{\sigma_2^2}{< \delta_2^2 >}}{(\sigma_1^2 + \sigma_2^2)^2} = \frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2} < \sigma_1^2 \tag{11}$$

The Covariance After Kinematic Fit

$$cov(P_1, P_2)_{KF} =< \delta_{1,KF}\delta_{2,KF} >=< (\delta_1 - \sigma_1^2\lambda)(\delta_2 - \sigma_2^2\lambda) >$$

$$= \sigma_1^2\sigma_2^2 \cancelto{\frac{1}{\sigma_1^2+\sigma_2^2}}{<\lambda^2>} - \frac{\sigma_1^2 < \delta_2^2 > + \sigma_2^2 < \delta_1^2 >}{\sigma_1^2 + \sigma_2^2} = -\frac{\sigma_1^2\sigma_2^2}{\sigma_1^2 + \sigma_2^2} \tag{12}$$

$$V = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix} \rightarrow V_{KF} = \begin{pmatrix} \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2} & -\frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2} \\ -\frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2} & \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2} \end{pmatrix} \tag{13}$$

- Improved momentum resolution
- Negative correlation between P_1 and P_2

Generalization to Multi-Variables

Assume that we have a set of measured data $\mathbf{m^0}$, unknown parameters $\mathbf{u^0}$ and constraints $\mathbf{f}^0$.

$$\mathbf{m^0} = \{m_1^0, m_2^0 \ldots m_N^0\}; \quad \mathbf{u^0} = \{u_1^0, u_2^0 \ldots u_J^0\}$$
$$\mathbf{f} = \{f_1(m_1^0, m_2^0, \ldots m_N^0, u_1^0, u_2^0, \ldots u_N^0), f_2^0, \ldots f_K^0\} \tag{14}$$

Let $\mathbf{m}^0$ denote our initial measured data, and $\mathbf{m}$ represent the 'guess' of the data in each iterative step, just alike P_{KF}s in the previous example. Equation (3) is generalized to:

$$\chi^2(\mathbf{m}) = (\mathbf{m}^0 - \mathbf{m})^\dagger V^{-1}(\mathbf{m}^0 - \mathbf{m}) + 2\lambda^\dagger \mathbf{f}(\mathbf{m}, \mathbf{u}). \tag{15}$$

Here, the Lagrange multiplier $\lambda = \{\lambda_1, \lambda_2, \ldots \lambda_K\}$ is not just a number but a column vector with k elements, corresponding to each kinematic constraint in $\mathbf{f}$.

χ^2 Minimization

We want to solve the equation

$$\vec{\nabla}\chi^2 = 0 \tag{16}$$

to obtain the minimized state. The differential term are listed within three groups.

$$\nabla_{\mathbf{m}} = -2V^{-1}(\mathbf{m}^0)(\mathbf{m}^0 - \mathbf{m}) + 2\mathbf{F}^{\dagger}_{\mathbf{m}}(\mathbf{m},\mathbf{u})\lambda = 0 \tag{17}$$

$$\nabla_{\mathbf{u}} = 2\mathbf{F}^{\dagger}_{\mathbf{u}}(\mathbf{m},\mathbf{u})\lambda = 0 \tag{18}$$

$$\nabla_{\lambda} = \mathbf{f}(\mathbf{m},\mathbf{u}). \tag{19}$$

Here, the subscripts denote partial derivatives. i.e. $((\mathbf{F}_m)_{ki} \equiv \frac{\partial f_k}{\partial m_i})$.

User Should Define...

$\mathbf{m}$	$\mathbf{u}$	$\mathbf{f}$	$\mathbf{V}$
Measured Data	Unknown parameters	Constraints	Covariance Matrix

Example: Mass-Constraint Fit

Example: $\Lambda \to p\pi$, Defining Variables and Constraints

Assume a decay of $\Lambda \to p\pi^-$. We define the measurements and unknowns as:

$$\mathbf{m} = \{P_p, \theta_p, \phi_p, P_\pi, \theta_\pi, \phi_\pi\}; \quad \mathbf{u} = \{P_\Lambda, \theta_\Lambda, \phi_\Lambda\} \tag{20}$$

Then we define the energy-momentum constraint equation as:

$$\begin{pmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{pmatrix} = \begin{pmatrix} -P_\Lambda \sin\theta_\Lambda \cos\phi_\Lambda + P_p \sin\theta_p \cos\phi_p + P_\pi \sin\theta_\pi \cos\phi_\pi \\ -P_\Lambda \sin\theta_\Lambda \sin\phi_\Lambda + P_p \sin\theta_p \sin\phi_p + P_\pi \sin\theta_\pi \sin\phi_\pi \\ -P_\Lambda \cos\theta_\Lambda + P_p \cos\theta_p + p_\pi \cos\theta_\pi \\ -\sqrt{P_\Lambda^2 + m_\Lambda^2} + \sqrt{P_p^2 + m_p^2} + \sqrt{P_\pi^2 + m_\pi^2} \end{pmatrix}. \tag{21}$$

where the mass constraint is naturally implemented in energy term.
Since we have 3 unmeasured variable with 4 kinematical constraints, this is a 4-3 = 1-Constrained fit.

Example: $\Lambda \to p\pi$, The Derivatives

We get $\mathbf{F_u}$ and $\mathbf{F_m}$ as

$$\mathbf{F_u} = \begin{pmatrix} \frac{\partial f_1}{\partial P_\Lambda} & \frac{\partial f_1}{\partial \theta_\Lambda} & \frac{\partial f_1}{\partial \phi_\Lambda} \\ \frac{\partial f_2}{\partial P_\Lambda} & \frac{\partial f_2}{\partial \theta_\Lambda} & \frac{\partial f_2}{\partial \phi_\Lambda} \\ \frac{\partial f_3}{\partial P_\Lambda} & \frac{\partial f_3}{\partial \theta_\Lambda} & \frac{\partial f_3}{\partial \phi_\Lambda} \\ \frac{\partial f_4}{\partial P_\Lambda} & \frac{\partial f_4}{\partial \theta_\Lambda} & \frac{\partial f_4}{\partial \phi_\Lambda} \end{pmatrix}; \quad \mathbf{F_m} = \begin{pmatrix} \frac{\partial f_1}{\partial P_p} & \cdots & \frac{\partial f_1}{\partial \phi_\pi} \\ \frac{\partial f_2}{\partial P_p} & \cdots & \frac{\partial f_2}{\partial \phi_\pi} \\ \frac{\partial f_3}{\partial P_p} & \cdots & \frac{\partial f_3}{\partial \phi_\pi} \\ \frac{\partial f_4}{\partial P_p} & \cdots & \frac{\partial f_4}{\partial \phi_\pi} \end{pmatrix} \tag{22}$$

We have all the matrices to calculate in each step. By applying an appropriate variance matrix and employing χ^2 selection criteria, we can do kinematic fit for the particles.

Example: $\Xi \to \Lambda\pi$, $\Lambda \to p\pi$

We require two mass constraints for $\Xi \to \Lambda\pi$; $\quad \Lambda \to p\pi$. In this case, careful considerations on the selection of variables. We will select

$$\mathbf{u} = \{P_\Xi, \theta_\Xi, \phi_\Xi\}; \quad \mathbf{m} = \{P_p, \theta_p, \phi_p, P_\pi, \theta_\pi, \phi_\pi\} \tag{23}$$

and define five constraints as:

$$\begin{pmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \end{pmatrix} = \begin{pmatrix} -P_{\Xi,x} + P_{p,x} + P_{\pi_\Lambda,x} + P_{\pi_\Xi,x} \\ -P_{\Xi,y} + P_{p,y} + P_{\pi_\Lambda,y} + P_{\pi_\Xi,y} \\ -P_{\Xi,z} + P_{p,z} + P_{\pi_\Lambda,z} + P_{\pi_\Xi,z} \\ -E_\Lambda + E_p + E_{\pi_\Lambda} \\ -E_\Lambda + E_p + E_{\pi_\Lambda} + E_{\pi_\Xi} \end{pmatrix}. \tag{24}$$

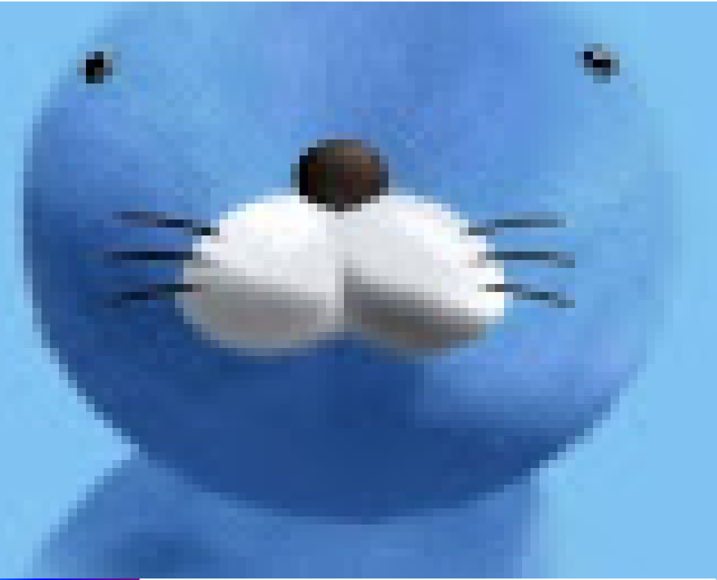
Λ variables are not selected in $\mathbf{u}$ to avoid negative DoF.($N_f - N_u = 5 - 6 = -1$)
Note that **we don't have explicit terms** related to $\vec{P}_\Lambda$, i.e. $-P_{\Lambda,x} + P_{p,x} + P_{\pi_\Lambda,x}$ etc., because $\vec{P}_\Lambda$ are neither unmeasured nor measured variables in our choice of parameters.

Example: Ξ decay at J-PARC

Tricks

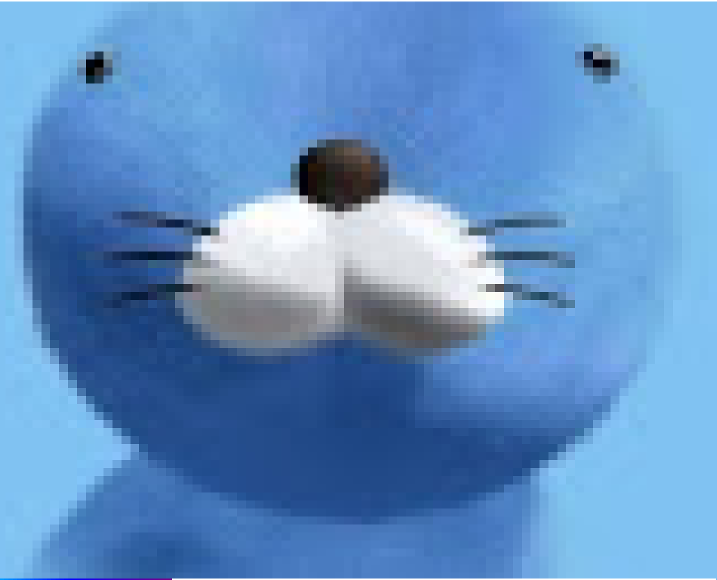
Off-diagonal Reduction

content...



Parameter Normalization

content...



Calculations

Processing Iterative Steps.

We can express the following equations based on the ones provided above:

$$V^{-1}(\mathbf{m}^0)(\mathbf{m}^{\nu+1} - \mathbf{m}^0) + (\mathbf{F}_m^{\dagger})^{\nu}\lambda^{\nu+1} = 0 \tag{25}$$

$$(\mathbf{F}_u^{\dagger})^{\nu}\lambda^{\nu+1} = 0 \tag{26}$$

$$\mathbf{f}^{\nu} + \mathbf{F}_m^{\nu}(\mathbf{m}^{\nu+1} - \mathbf{m}^{\nu}) + \mathbf{F}_u^{\nu}(\mathbf{u}^{\nu+1} - \mathbf{u}^{\nu}) = 0. \tag{27}$$

Equation (27) is not a direct consequence of Equation (18) but rather a *linear approximation* to proceed with our iteration step. Expanding the ∇_λ term with a Taylor series leads to this equation. Note that, as our parameters $\mathbf{m}$ and $\mathbf{u}$ are updated during the step, our constraint matrix $\mathbf{f}$ should also be updated during the iteration. Here, λ should be indexed as $\nu + 1$ since it is a parameter to be guessed in the next step.

Solving the Equation(1)

Multiplying **V** to Equation (25) leads to:

$$\mathbf{m}^{\nu+1} - \mathbf{m}^0 = -V(\mathbf{m}^0)(\mathbf{F}_m^\dagger)^\nu \lambda^{\nu+1}. \tag{28}$$

Substituting Equation (28) into Equation (27), we get:

$$\begin{aligned} \mathbf{F}_u^\nu(\mathbf{u}^{\nu+1} - \mathbf{u}^\nu) &= -\mathbf{f}^\nu - \mathbf{F}_m^\nu(-V(\mathbf{m}^0)(\mathbf{F}_m^\dagger)^\nu \lambda^{\nu+1} + \mathbf{m}^0 - \mathbf{m}^\nu) \\ &= S\lambda^{\nu+1} - R \end{aligned} \tag{29}$$

where$S \equiv \mathbf{F}_\mathbf{m}^\nu V(\mathbf{m}^0)(\mathbf{F}_\mathbf{m}^\dagger)^\nu$ and $R \equiv \mathbf{f}^\nu + \mathbf{F}_\mathbf{m}^\nu(\mathbf{m}^0 - \mathbf{m}^\nu)$. Multiplying $(\mathbf{F}_u^\dagger)^\nu S^{-1}$ and substituting Equation (26), we get:

$$(\mathbf{F}_u^\dagger)^\nu S^{-1} \mathbf{F}_u^\nu(\mathbf{u}^{\nu+1} - \mathbf{u}^\nu) = \cancelto{0}{(\mathbf{F}_u^\dagger)^\nu \lambda^{\nu+1}} - (\mathbf{F}_u^\dagger)^\nu S^{-1} R. \tag{30}$$

Solving the Equation(2)

Then we naturally obtain:

$$\mathbf{u}^{\nu+1} = \mathbf{u}^{\nu} - ((\mathbf{F}_u^{\dagger})^{\nu} S^{-1} \mathbf{F}_u^{\nu})^{-1} (\mathbf{F}_u^{\dagger})^{\nu} S^{-1} R. \tag{31}$$

and from Equation (29)

$$\lambda^{\nu+1} = S^{-1}(\mathbf{F}_u^{\nu}(\mathbf{u}^{\nu+1} - \mathbf{u}^{\nu}) + R) \tag{32}$$

. For a summary, we have obtained all equations to proceed to the next step. All other matrices in the equation can be calculated from parameters of the current step, and χ^2 can be evaluated from (15) .

$$\begin{cases} \mathbf{u}^{\nu+1} = \mathbf{u}^{\nu} - ((\mathbf{F}_u^{\dagger})^{\nu} S^{-1} \mathbf{F}_u^{\nu})^{-1} (\mathbf{F}_u^{\dagger})^{\nu} S^{-1} R & (31) \\ \lambda^{\nu+1} = S^{-1}(\mathbf{F}_u^{\nu}(\mathbf{u}^{\nu+1} - \mathbf{u}^{\nu}) + R) & (32) \\ \mathbf{m}^{\nu+1} = \mathbf{m}^0 - V(\mathbf{m}^0)(\mathbf{F}_m^{\dagger})^{\nu} \lambda^{\nu+1} & (28) \end{cases}$$

Evolution of the Variance Matrix

Take a look at Eq.(28). We see that $m^{\nu+1}$ is an addition(subtraction) of some parameters to the initially measured data. As we already know the error, i.e. Variance matrix, of initial data, we can estimate how error propagates through the fitting process.

$$V(m) = J_{m,m^0} V(m^0) J^{\dagger}_{m,m^0} \tag{33}$$

We need to calculate the Jacobian,

$$J_{m,m^0(i,j)} = \frac{\partial m_i}{\partial m^0_j} \tag{34}$$

Evolution of the Variance Matrix.

To begin with, let us express Eq (28) in terms of m^0. At the moment we will drop the superscript ν. As $\mathbf{f}(\mathbf{m},\mathbf{u})$is a constant on m^0, $\mathbf{F_m}$ also will be a constant to m^0. Then we only need to consider the derivatives of λ. By substituting (31) ,

$$\lambda = S^{-1}(-\mathbf{F}_u(((\mathbf{F}_u^\dagger)S^{-1}\mathbf{F}_u)^{-1}(\mathbf{F}_u^\dagger)S^{-1}R) + R) \tag{35}$$

and we have

$$R \equiv \mathbf{f} + \mathbf{F_m}(\mathbf{m}^0 - \mathbf{m}) \rightarrow \frac{\partial R}{\partial m^0} = \mathbf{F}_m \tag{36}$$

so that

$$\frac{\partial \lambda}{\partial m^0} = S^{-1}(-\mathbf{F}_u((\mathbf{F}_u^\dagger S^{-1}\mathbf{F}_u)^{-1}\mathbf{F}_u^\dagger S^{-1}\mathbf{F}_m) + \mathbf{F}_m). \tag{37}$$

Evolution of the Variance Matrix

Now define the symmetric matrices $G \equiv \mathbf{F}_{\mathbf{m}}^{\dagger} S^{-1} \mathbf{F}_{\mathbf{m}}$, $U \equiv (\mathbf{F}_{u}^{\dagger} S^{-1} \mathbf{F}_{u})^{-1}$ and $H \equiv \mathbf{F}_{m}^{\dagger} S^{-1} \mathbf{F}_{\mathbf{u}}$. Then we have expressions for $\frac{\partial \lambda}{\partial m^0}$ hence

$$
\begin{aligned}
J_{m,m^0} &= I - V(m^0)\mathbf{F}_{m}^{\dagger} \frac{\partial \lambda}{\partial m^0} = I - V\mathbf{F}_{m}^{\dagger}(-S^{-1}\mathbf{F}_{u}U^{-1}H^{\dagger} + S^{-1}\mathbf{F}_{m}) \\
&= I - V(G - HUH^{\dagger}) \qquad (38)
\end{aligned}
$$

If we let $C = G - HUH^{\dagger}$, we obtain

$$
V(m) = J_{m,m^0} V J_{m,m^0}^{\dagger} = V - 2VCV + VCVCV. \qquad (39)
$$

You might want to neglect higher order term, but please keep 2nd order term at the moment. Some materials like [1] had an error in this part.

Pull distribution

It is better to check the pull distribution to check the quality of a fit. By defining the residual $\epsilon = m - m^0$ and its variance $V(\epsilon)$, pull is defined as:,

$$P(\epsilon) = \epsilon / \sqrt{V(\epsilon)} \tag{40}$$

and

$$V(\epsilon) \equiv V(m) + V(m^0) - 2Cov(m, m^0). \tag{41}$$

We have already calculated J_{m,m^0}. Then we directly get the covariance matrix.

Pull distribution

The covariance can be estimated as:

$$Cov(m, m^0) = J_{m,m^0} V(m) = V - VCV. \tag{42}$$

If we substitute this and Eq.(39) into Eq.(41), we get

$$V(\epsilon) = VCVCV. \tag{43}$$

Note that 2nd order term affects the variance.

Variance of the Unknowns

Just like how we derived Eq.(39) we can estimate the variance matrix of the unknowns.

$$V_U = J_{u,m0} V J_{u,m0}^T \tag{44}$$

$J_{u,m0}$ can be obtained from Eq.(31). Denoting $((\mathbf{F}_u^\dagger)^\nu S^{-1} \mathbf{F}_u^\nu)^{-1} (\mathbf{F}_u^\dagger)^\nu S^{-1}$ as K,

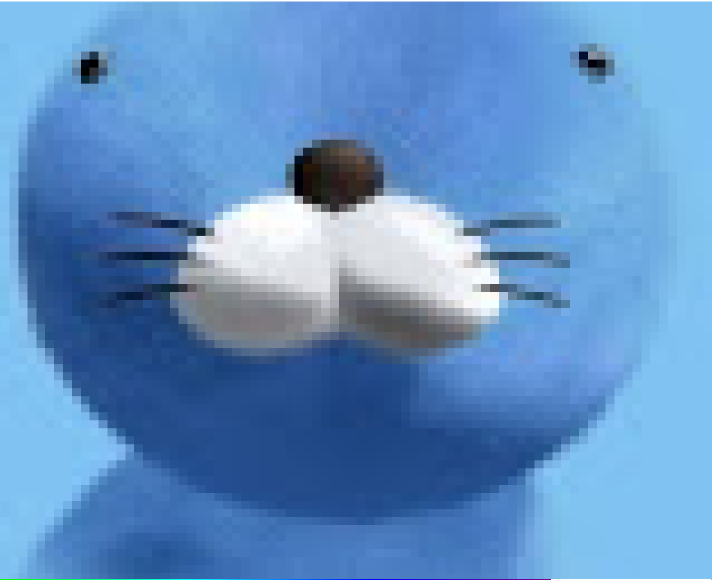
$$J_{u^{\nu+1},m^0} = \frac{\partial u^{\nu+1}}{\partial m^0} = \frac{\partial u^\nu}{\partial m^0} - K \frac{\partial R}{\partial m^0} \simeq -K\mathbf{F}_m. \tag{45}$$

Note that we only have initial "Guess" for the unknowns; In principle, it is not a driven value from measurements. Then, $\frac{\partial u^0}{\partial m_0} = 0$. Also, we approximate that the terms in 2nd or higher iterations are negligible: $J_{u,m^0} \simeq J_{u^0,m^0}$.

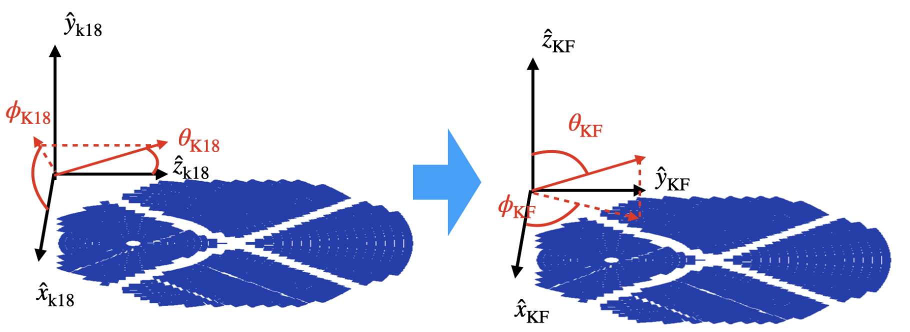
Applications

Example: Chained Mass-Constraint Fit

content...

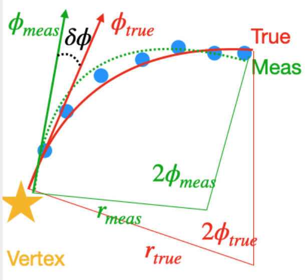


Mass-Constraint Kinematic Fit with HypTPC.



For Kinematic FIt in HypTPC analysis, we want our coordinate system to be aligned with $\vec{B}$, so that our covariance matrix representation fits the representations in KF coordinate. We correlated ϕ angle with p as an feature of helix fit, where ϕ is the angle lying on the circle of the helix.

Covariance Matrix in Helix Track



The cartoon on the left illustrates the correlation between the helical (or circular) track and the azimuth angle at the intersection point (or decay vertex). Since the helical trajectory is defined by measured spatial points, a deviation in radius does not shift the arc itself but moves the center of the circle. It is also worth noting that the arc length, l, remains nearly constant.

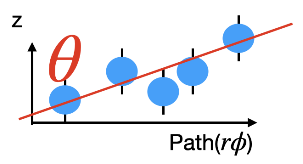
If we let the azimuth angle at the center of the track as ϕ_0, we deduce following relations:

$$\phi = \phi_0 \pm \frac{l}{2r}; \quad \delta\phi = \pm\frac{l}{2r}\frac{\delta r}{r} = \pm\frac{l}{r}\frac{\delta p_T}{p_T} \tag{46}$$

and we naturally obtain:

$$\sigma_\phi^2 = \langle\delta\phi\delta\phi\rangle = \frac{l^2}{r^2 p_T^2}\sigma_{p_T}^2; \quad \mathrm{Cov}(\phi, p_T) = \langle\delta\phi\delta p_T\rangle = \pm\frac{l}{r p_T}\sigma_{p_T}^2 \tag{47}$$

Covariance Matrix in Helix Track



The left figure illustrates the relationship between the pitch(dz) and the helix fit. The pitch is determined by fitting the vertical displacement along transverse path of the helical trajectory. Denoting the polar angle is expressed as $\theta = \frac{\pi}{2} - \arctan(dz)$, we estimate the varaiance of θ based on the fitting error of dz, the slope parameter of linear fit.

$$\sigma_{dz}^2 = \frac{\sum \delta_z^2/(n-2)}{\sum (x-\bar{x})^2} \simeq \frac{n\sigma_z^2/(n-2)}{nL^2/12}; \quad \sigma_\theta = \frac{\partial dz}{\partial \theta}\sigma_{dz} = \frac{1}{1+dz^2}\sigma_{dz}.$$

Note that, the momentum $p_z = p_T dz$ would also have some covariance with θ,

$$\langle \delta p \delta\theta \rangle = dz \cancelto{0}{\langle \delta p_T \delta\theta \rangle} + p_T \langle \delta dz \delta\theta \rangle = \frac{p_T}{1+dz^2}\sigma_{dz}^2.$$

Example: MassVertex-Constraint Fit

To be Updated..

References

O. Skjeggestad Frodesen A.G. **Probability and Statistics in Particle Physics**. Columbia University Press, 1980. ISBN: 8200019063.