MATHUSLA Fixed Beta Kalman Filter Calculations

August 2022

1 Coordinate Definitions - Orthogonal Angular

State velocity variables will be changed to θ_x the polar angle from the vertical of the track velocity along x, and θ_z the polar angle from the vertical of the track velocity along z.

$$x = \begin{pmatrix} x \\ t \\ z \\ \tan\theta_x \\ \tan\theta_z \end{pmatrix} \tag{1}$$

The measurement variables (information about each Digi hit) will be unchanged

$$m = \begin{pmatrix} x \\ t \\ z \end{pmatrix} \tag{2}$$

In CMS coordinates, y is along the vertical. Hence our polar coordinates are non-canonical. Consider the displacement between hits in the detector on separate layers, with coordinate differences given by $\Delta\vec{r} \doteq [\Delta x, \Delta y, \Delta z]^T$, in the standard basis. We define our angular coordinates by

$$\tan\theta_x \equiv \frac{\Delta x}{\Delta y} = \frac{v_x}{v_y} \tag{3}$$

$$\tan\theta_z \equiv \frac{\Delta z}{\Delta y} = \frac{v_z}{v_y} \tag{4}$$

1.1 Parametrization

As y is the most precisely measured parameter in the experiment, we parametrize our tracks by it. Using $v = \beta c$ we find

$$\beta^2 c^2 = v_x^2 + v_y^2 + v_z^2 = v_y^2(1 + \tan^2\theta_x + \tan^2\theta_z) \equiv v_y^2 N^2 \tag{5}$$

Where N is a normalization constant we define for convenience.

$$v_y = \frac{\beta c}{N} \qquad \Longrightarrow \qquad \Delta t = \frac{\Delta y}{v_y} = \frac{\Delta y N}{\beta c} \tag{6}$$

$$v_x = \frac{\beta c}{N}\tan\theta_x \qquad \Longrightarrow \qquad \Delta x = v_x \Delta t = \Delta y \tan\theta_x \tag{7}$$

$$v_z = \frac{\beta c}{N}\tan\theta_z \qquad \Longrightarrow \qquad \Delta z = v_z \Delta t = \Delta y \tan\theta_z \tag{8}$$

Note that there is no need to consider relativistic kinematics since all of this is measured in the lab frame, and hence no Lorentz boosts come into play even in the highly relativistic case. Ie. there should be no factors of γ hiding anywhere.

1.2 Propagation and Measurement

Definitions of Measurement and Process Matrices (F_{k-1} and H_k in Fruhwirth paper). These determine how the propagation of the state between layers is done, and how the comparison between the propagated state and the hits to be added to the track is done.

Measurement matrix H_k (C in the code) will be

$$H_k = C = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix} \tag{9}$$

This projects out the angular components of the state such that the following is satisfied

$$m = C\ x \tag{10}$$

The propagation matrix F_{k-1} (A in the code) will be

$$F_{k-1} = A = \begin{pmatrix} 1 & 0 & 0 & \Delta y & 0 \\ 0 & 1 & \Delta t/z & 0 & 0 \\ 0 & 0 & 1 & 0 & \Delta y \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{11}$$

which propagates each of the coordinates in x by an amount determined by the seperation between the current and next layer in the detector, as well as the direction of propogation. Note that t has been propagated in the z column to avoid singularities in the division by the coordinate.

$$x_k^{k-1} = F_{k-1} x_{k-1} = \begin{pmatrix} x + \Delta x \\ t + \Delta t \\ z + \Delta z \\ \tan\theta_x \\ \tan\theta_z \end{pmatrix} = x_{k-1} + \Delta y \begin{pmatrix} \tan\theta_x \\ \frac{1}{\beta c}\sqrt{1 + \tan^2\theta_x + \tan^2\theta_z} \\ \tan\theta_z \\ 0 \\ 0 \end{pmatrix} \tag{12}$$

Where the indexing follows the conventions of [1], and we have used (6-8) in the last equality.

2 Process Covariance Matrix

We now derive the Multiple Scattering Process Covariance Matrix for use with a Kalman Filter tracking algorithm designed for use on MATHUSLA. The prescription is described in detail in [2]. To summarize, we calculate the partial derivatives of scattered state vector elements w.r.t. uncorrelated scattering angles in a track coordinate system, rotate to a reference coordinate system, and use the following formula for the co-variance matrix for multiple scattering.

$$Q_{ij}^{\text{scatt}} = \langle P_i, P_j \rangle = \sigma^2(\theta_{\text{proj}}) \left(\frac{\partial P_i}{\partial \theta_1}\frac{\partial P_j}{\partial \theta_1} + \frac{\partial P_i}{\partial \theta_2}\frac{\partial P_j}{\partial \theta_2} \right) \tag{13}$$

Where P_i and P_j are components of the state vector describing the system in a chosen parametrisation.

2.1 Calculation of Multiple Scattering Variance

In order to calculate the Multiple Scattering variance, $\sigma^2(\theta_{\text{proj}})$, as tracked particles pass through detector layers, we use the following approximate formula found in [2–4].

$$\sigma(\theta_{\text{proj}}) = \frac{13.6z}{p\beta}\sqrt{\frac{X}{X_0}}\left[1 + 0.038\log_{10}\left(\frac{Xz^2}{X_0\beta^2}\right)\right] \tag{14}$$

$$= \frac{13.6}{p\beta}\sqrt{\frac{X}{X_0}}\left[1 + 0.038\log_{10}\left(\frac{X}{X_0\beta^2}\right)\right] \tag{15}$$

Where X is the width of the layer, X_0 is it's radiation length, p is the magnitude of the particle's 3-momentum, and the second equality follows since we are only concerned with muons (singly charged $\implies z = 1$). X and X_0 are determined from the geometry and material of the detector layer. When more than one material is propagated through, the prescription suggested by [3] is the following.

$$\sigma(\theta_{\text{proj}}) = 13.6z\sqrt{\sum_i \frac{X_i}{p_i^2\beta_i^2 X_{0,i}}}\left[1 + 0.038\log_{10}\left(\sum_i \frac{X_i z^2}{\beta_i^2 X_{0,i}}\right)\right] \tag{16}$$

$$= \frac{13.6}{p\beta}\sqrt{\sum_i \frac{X_i}{X_{0,i}}}\left[1 + 0.038\log_{10}\left(\frac{1}{\beta^2}\sum_i \frac{X_i}{X_{0,i}}\right)\right] \tag{17}$$

Where i indexes the material, X_i is the width of material i, and $X_{0,i}$ is the radiation length (in units of distance) for material i. In the last line we have assumed that the momentum (and hence β) is unaffected by passing through the material. We simply choose a representative value for p as no momentum information is available prior to the track fit in the absence of a strong magnetic field. The radiation lengths for scintillator and aluminum are found in [5] and [6] respectively.

$$\begin{aligned} X_{\text{sc}} &= 2 \text{ cm} & X_{\text{al}} &= 1 \text{ cm} \\ X_{0,\text{sc}} &= 43 \text{ cm} & X_{0,\text{al}} &= \rho_{\text{al}}^{-1}(24.0111 \text{ g cm}^{-2}) = 8.9 \text{ cm} \\ & & \rho_{\text{al}} &= 2.7 \text{ g cm}^{-3} \end{aligned}$$

This gives

$$\omega = \sum_i \frac{X_i}{X_{0,i}} = 0.16 \tag{18}$$

Where ω is width in radiation lengths of the detector layer. We notice that this corresponds to the number of radiation lengths of the path of the particle through the layer if it is moving orthogonal to the surface of the layer. Hence we divide $\cos\theta$ (ie. β_y).

$$\sigma(\theta_{\text{proj}}) = \frac{13.6}{p\beta}\sqrt{\frac{\omega}{\cos\theta}}\left[1 + 0.038\log_{10}\left(\frac{\omega}{\beta^2\cos\theta}\right)\right] \tag{19}$$

It is pointless to estimate uncertainty in the calculated values as they will invariably be negligible compared to the uncertainties p and θ.

2.2 Setting up the calculation

The propagated state vector to the next layer is given in terms of θ_x and θ_z below.

$$P = \left\{x_0 + \Delta y\tan(\theta_x) \quad t_0 + \frac{\Delta y\sqrt{1+\tan(\theta_x)^2+\tan(\theta_z)^2}}{c\beta} \quad z_0 + \Delta y\tan(\theta_z) \quad \tan(\theta_x) \quad \tan(\theta_z)\right\} \tag{20}$$

The relation between the track system and the reference system is given by the matrix

$$R = \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{pmatrix} \tag{21}$$

The scattered track vector is given in the track system in terms of $\begin{pmatrix} x & y & z \end{pmatrix}$ forcing $y = 1$, by

$$\begin{pmatrix} \tan(\theta_1) \\ 1 \\ \tan(\theta_2) \end{pmatrix} \tag{22}$$

Accounting for the constant value of y, the scattered vector in the reference system is then given by

$$\begin{pmatrix} \frac{\alpha_2+\alpha_1\tan(\theta_1)+\alpha_3\tan(\theta_2)}{\beta_2+\beta_1\tan(\theta_1)+\beta_3\tan(\theta_2)} \\ 1 \\ \frac{\gamma_2+\gamma_1\tan(\theta_1)+\gamma_3\tan(\theta_2)}{\beta 2+\beta_1\tan(\theta_1)+\beta_3\tan(\theta_2)} \end{pmatrix} = \begin{pmatrix} \tan(\theta_x) \\ 1 \\ \tan(\theta_z) \end{pmatrix} \tag{23}$$

From equations (10)-(13) of [2], we can use equation (6) of Wolin and Ho to note that the direction cosines are 3 orthogonal unit vectors, and hence satisfy the algebra for SO(3). Ie. they are cross products of each other.

$$\frac{\partial \tan\theta_x}{\partial\theta_1}\bigg|_{\theta_1,\theta_2\to 0} = \frac{\alpha_1\beta_2-\alpha_2\beta_1}{(\beta_2)^2} = \frac{(\vec{\beta}\times\vec{\alpha})_3}{(\beta_2)^2} = \frac{-\gamma_3}{(\beta_2)^2} \tag{24}$$

$$\frac{\partial \tan\theta_x}{\partial\theta_2}\bigg|_{\theta_1,\theta_2\to 0} = \frac{\alpha_3\beta_2-\alpha_2\beta_3}{(\beta_2)^2} = \frac{(\vec{\alpha}\times\vec{\beta})_1}{(\beta_2)^2} = \frac{\gamma_1}{(\beta_3)^2} \tag{25}$$

$$\frac{\partial \tan\theta_z}{\partial\theta_1}\bigg|_{\theta_1,\theta_2\to 0} = \frac{\beta_2\gamma_1-\beta_1\gamma_2}{(\beta_2)^2} = \frac{(\vec{\beta}\times\vec{\gamma})_3}{(\beta_2)^2} = \frac{\alpha_3}{(\beta_2)^2} \tag{26}$$

$$\frac{\partial \tan\theta_z}{\partial\theta_2}\bigg|_{\theta_1,\theta_2\to 0} = \frac{\beta_2\gamma_3-\beta_3\gamma_2}{(\beta_2)^2} = \frac{(\vec{\gamma}\times\vec{\beta})_1}{(\beta_2)^2} = \frac{-\alpha_1}{(\beta_2)^2} \tag{27}$$

We then substitute equations (5)-(8) into $\frac{\partial P}{\partial\theta_1}$ and $\frac{\partial P}{\partial\theta_2}$ to find their simplified forms.

$$\frac{\partial P}{\partial\theta_1} = \left\{ \frac{-\gamma_3\Delta y}{\beta_2^2} \qquad \frac{\Delta y(\tan(\theta_z)\alpha_3-\tan(\theta_x)\gamma_3)}{\beta_2 c\beta\sqrt{1+\tan(\theta_x)^2+\tan(\theta_z)^2}} \qquad \frac{\alpha_3\Delta y}{\beta_2^2} \qquad \frac{-\gamma_3}{\beta_2^2} \qquad \frac{\alpha_3}{\beta_2^2} \right\} \tag{28}$$

$$\frac{\partial P}{\partial\theta_2} = \left\{ \frac{\gamma_1\Delta y}{\beta_2^2} \qquad \frac{\Delta y(\tan(\theta_x)\gamma_1-\tan(\theta_z)\alpha_1)}{\beta_2 c\beta\sqrt{1+\tan(\theta_x)^2+\tan(\theta_z)^2}} \qquad \frac{-\alpha_1\Delta y}{\beta_2^2} \qquad \frac{\gamma_1}{\beta_2^2} \qquad \frac{-\alpha_1}{\beta_2^2} \right\} \tag{29}$$

Q is then given by

$$Q = \sigma^2\theta_(proj) \begin{pmatrix} (\frac{\partial P}{\partial\theta_1})_1(\frac{\partial P}{\partial\theta_1})_1 & (\frac{\partial P}{\partial\theta_1})_1(\frac{\partial P}{\partial\theta_1})_2 & (\frac{\partial P}{\partial\theta_1})_1(\frac{\partial P}{\partial\theta_1})_3 & (\frac{\partial P}{\partial\theta_1})_1(\frac{\partial P}{\partial\theta_1})_4 & (\frac{\partial P}{\partial\theta_1})_1(\frac{\partial P}{\partial\theta_1})_5 \\ (\frac{\partial P}{\partial\theta_1})_2(\frac{\partial P}{\partial\theta_1})_1 & (\frac{\partial P}{\partial\theta_1})_2(\frac{\partial P}{\partial\theta_1})_2 & (\frac{\partial P}{\partial\theta_1})_2(\frac{\partial P}{\partial\theta_1})_3 & (\frac{\partial P}{\partial\theta_1})_2(\frac{\partial P}{\partial\theta_1})_4 & (\frac{\partial P}{\partial\theta_1})_2(\frac{\partial P}{\partial\theta_1})_5 \\ (\frac{\partial P}{\partial\theta_1})_3(\frac{\partial P}{\partial\theta_1})_1 & (\frac{\partial P}{\partial\theta_1})_3(\frac{\partial P}{\partial\theta_1})_2 & (\frac{\partial P}{\partial\theta_1})_3(\frac{\partial P}{\partial\theta_1})_3 & (\frac{\partial P}{\partial\theta_1})_3(\frac{\partial P}{\partial\theta_1})_4 & (\frac{\partial P}{\partial\theta_1})_3(\frac{\partial P}{\partial\theta_1})_5 \\ (\frac{\partial P}{\partial\theta_1})_4(\frac{\partial P}{\partial\theta_1})_1 & (\frac{\partial P}{\partial\theta_1})_4(\frac{\partial P}{\partial\theta_1})_2 & (\frac{\partial P}{\partial\theta_1})_4(\frac{\partial P}{\partial\theta_1})_3 & (\frac{\partial P}{\partial\theta_1})_4(\frac{\partial P}{\partial\theta_1})_4 & (\frac{\partial P}{\partial\theta_1})_4(\frac{\partial P}{\partial\theta_1})_5 \\ (\frac{\partial P}{\partial\theta_1})_5(\frac{\partial P}{\partial\theta_1})_1 & (\frac{\partial P}{\partial\theta_1})_5(\frac{\partial P}{\partial\theta_1})_2 & (\frac{\partial P}{\partial\theta_1})_5(\frac{\partial P}{\partial\theta_1})_3 & (\frac{\partial P}{\partial\theta_1})_5(\frac{\partial P}{\partial\theta_1})_4 & (\frac{\partial P}{\partial\theta_1})_5(\frac{\partial P}{\partial\theta_1})_5 \end{pmatrix} \tag{30}$$

3 Seed Covariance Matrix

3.1 Measurement Covariance to State Vector Covariance

Jacobian in equation (27) of the calculation document for the previous algorithm needs to be recalculated and used to initialize the first covariance matrix for the filter with the new coordinates.

We start with a measurement and covariance matrix of the measurement:

$$m_i = [x,t,z,y]_i \qquad V = \mathrm{diag}(\sigma_x^2,\sigma_t^2,\sigma_z^2,\sigma_y^2)$$

A seed consists of a pair of measurements, so our starting coordinates would be:

$$\hat{p} = m_1 \oplus m_2 = [x_1,t_1,z_1,y_1,x_2,t_2,z_2,y_2]$$

$$\hat{V}(\hat{p}) = \mathrm{diag}(\sigma_{1x}^2,\sigma_{1t}^2,\sigma_{1z}^2,\sigma_{1y}^2,\sigma_{2x}^2,\sigma_{2t}^2,\sigma_{2z}^2,\sigma_{2y}^2)$$

Our goal is to calculate the covariance matrix if the measurement is transformed to the following state:

$$\hat{s} = [x_1,t_1,z_1,\tan(\theta_x),\tan(\theta_z)]$$

where

$$\tan\theta_x \equiv \frac{\Delta x}{\Delta y} = \frac{v_x}{v_y} \quad \tan\theta_z \equiv \frac{\Delta z}{\Delta y} = \frac{v_z}{v_y}$$

Under a change of variables, the covariance matrix is of the following form [5]:

$$\hat{V}_{ij}(\hat{s}) = \frac{\partial \hat{s}_i}{\partial \hat{p}_n}\frac{\partial \hat{s}_j}{\partial \hat{p}_m}\hat{V}_{nm}(\hat{p}) = \frac{\partial \hat{s}_i}{\partial \hat{p}_n}\hat{V}_{nm}(\hat{p})\frac{\partial \hat{s}_j}{\partial \hat{p}_m} = J_{in}V_{nm}J^T_{mj} = JVJ^T$$

Where we set

$$J_{in} = \frac{\partial \hat{s}_i}{\partial \hat{p}_n}$$

Using Sympy (and checking some of the values by hand to make sure), we arrive at the folloing form for the Jacobian J:

$$J = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{(y_1-y_2)} & 0 & 0 & \frac{-(x_1-x_2)}{(y_1-y_2)^2} & \frac{-1}{(y_1-y_2)} & 0 & 0 & \frac{(x_1-x_2)}{(y_1-y_2)^2} \\ 0 & 0 & \frac{1}{(y_1-y_2)} & \frac{-(z_1-z_2)}{(y_1-y_2)^2} & 0 & 0 & \frac{-1}{(y_1-y_2)} & \frac{(z_1-z_2)}{(y_1-y_2)^2} \end{bmatrix}$$

or alternatively:

$$J = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \frac{-1}{\Delta y} & 0 & 0 & \frac{\tan(\theta_x)}{\Delta y} & \frac{1}{\Delta y} & 0 & 0 & \frac{-\tan(\theta_x)}{\Delta y} \\ 0 & 0 & \frac{-1}{\Delta y} & \frac{\tan(\theta_z)}{\Delta y} & 0 & 0 & \frac{1}{\Delta y} & \frac{-\tan(\theta_z)}{\Delta y} \end{bmatrix}$$

3.2 State Vector Covariance to Vertex Covariance

Another transformation that must be done is to convert the state vector to the vertex coordinates:

$$\hat{s} = [x, t, z, \tan(\theta_x), \tan(\theta_z)] \rightarrow \hat{v} = [x, y, z, v_x, v_y, v_z, t]$$

where x, t, z is the bottom of the track, and

$$v_y = \frac{\beta c}{N} \qquad v_x = \frac{\beta c}{N}\tan\theta_x \qquad v_z = \frac{\beta c}{N}\tan\theta_z \qquad N = \sqrt{1 + \tan^2(\theta_x) + \tan^2(\theta_z)}$$

Alternatively, we can therefore write $\hat{v}$ as:

$$\hat{v} = [x, y, z, \frac{\beta c}{N}\tan(\theta_x), \frac{\beta c}{N}, \frac{\beta c}{N}\tan(\theta_z), t]$$

This gives the Jacobian

$$J_{ij} = \frac{\partial \hat{v}_i}{\partial \hat{s}_j} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \frac{\beta c(\tan(\theta_z)^2+1)}{N^{3/2}} & \frac{-\beta c\tan(\theta_x)\tan(\theta_z)}{N^{3/2}} \\ 0 & 0 & 0 & \frac{-\beta c\tan(\theta_x)}{N^{3/2}} & \frac{-\beta c\tan(\theta_z)}{N^{3/2}} \\ 0 & 0 & 0 & \frac{-\beta c\tan(\theta_x)\tan(\theta_z)}{N^{3/2}} & \frac{\beta c(\tan(\theta_x)^2+1)}{N^{3/2}} \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

References

[1] Frühwirth, R. (1987). Application of Kalman filtering to track and vertex fitting. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 262(2-3), 444-450. doi:10.1016/0168-9002(87)90887-4

[2] E. J. Wolin and L. L. Ho, Nucl. Instrum. Meth. A **329**, 493-500 (1993) doi:10.1016/0168-9002(93)91285-U

[3] Lynch, G. R., Dahl, O. I. (1991). Approximations to multiple Coulomb scattering. Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms, 58(1), 6-10. doi:10.1016/0168-583x(91)95671-y

[4] Greenberg, S. (2021, January 10). Closest Approach of Two Parametric Vectors [Pdf].

[5] M. Tanabashi et al. (Particle Data Group), Phys. Rev. D 98, 030001 (2018) and 2019 update

[6] PRODUCT DATA SHEET. (n.d.). Retrieved from https://www.crystals.saint-gobain.com/sites/imdf.crystals.com/files/documents/bc440-bc448-series-data-sheet.pdf

[7] Tsai, Y. (1974). Pair production and bremsstrahlung of charged leptons. Reviews of Modern Physics, 46(4), 815-851. doi:10.1103/revmodphys.46.815