MATHUSLA Kalman Filter - Cited Calculations

Keegan Humphrey

June 2021

Abstract

In this paper we give the details of the calculations needed for the Kalman Filter Tracking and Vertexing Algorithm used on the MATHUSLA displaced vertex detector. We show that the equality of filtered and predicted chi squared increments are equal at all steps of the filter. We estimate the variance in the scattering angle due to Multiple Coulomb Scatters as charged particles pass through the detector layers. We then use this to calculate the Process Covariance Matrix of the Kalman Filter Tracking Algorithm. We then find approximations to the co-variance matrices of the states used to initialise the Tracking and Vertexing.

1 Proving the Equality of χ^2_P and χ^2_F

We want to show that

$$\chi^2_P = \chi^2_F \tag{1}$$

This result is mentioned in [1], where they also define

$$\chi^2_P \equiv (r_k^{k-1})^T (R_k^{k-1})^{-1} r_k^{k-1} \tag{2}$$
$$\chi^2_F \equiv r_k^T (R_k)^{-1} r_k \tag{3}$$

with

$$R_k^{k-1} \equiv V_k + H_k C_k^{k-1} H_k^T \tag{4}$$
$$R_k \equiv V_k - H_k C_k H_k^T \tag{5}$$
$$K_k \equiv C_k H_k^T V_k^{-1} \tag{6}$$

and

$$r_k = (I - H_k K_k) r_k^{k-1} \tag{7}$$
$$C_k = (I - K_k H_k) C_k^{k-1} \tag{8}$$

Notice first that

$$\begin{aligned} I - H_k K_k &= I - H_k C_k H_k^T V_k^{-1} \\ &= R_k V_k^{-1} \end{aligned}$$

hence

$$\begin{aligned} (I - K_k^T H_k^T)(R_k)^{-1}(I - H_k K_k) &= (V_k^{-1})^T R_k^T V_k^{-1} \\ &= (V_k^T)^{-1} R_k^T V_k^{-1} \\ &= V_k^{-1} R_k^T V_k^{-1} \end{aligned}$$

Since all co-variance matrices are symmetric, and the inverse of the transpose of a matrix is equal to the transpose of the inverse. So (1) is true exactly if

$$(R_k^{k-1})^{-1} = V_k^{-1} R_k^T V_k^{-1}$$

using (2) and (3). In other words, only if

$$\begin{aligned}
I &= V_k^{-1} R_k^T V_k^{-1} R_k^{k-1} \\
&= (V_k^{-1} - V_k^{-1} H_k C_k H_k^T V_k^{-1})(V_k + H_k C_k^{k-1} H_k^T) \\
&= I - V_k^{-1} H_k C_k H_k^T + V_k^{-1} H_k C_k^{k-1} H_k^T - V_k^{-1} H_k C_k H_k^T V_k^{-1} H_k C_k^{k-1} H_k^T
\end{aligned}$$

since C_k is also a covariance matrix, $C_k = C_k^T$. So (1) is true only if

$$\begin{aligned}
V_k^{-1} H_k C_k^{k-1} H_k^T &= V_k^{-1} H_k C_k H_k^T (I + V_k^{-1} H_k C_k^{k-1} H_k^T) \\
H_k C_k^{k-1} H_k^T &= H_k C_k^T H_k^T + H_k K_k H_k C_k^{k-1} H_k^T \\
H_k (I - K_k H_k) C_k^{k-1} H_k^T &= H_k C_k H_k^T
\end{aligned}$$

Which always holds because of (8). Hence (1) is true.

2 Process Covariance Matrix (Q) for Kalman Filter

We first 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.

$$Q_{ij} = \langle P_i, P_j \rangle = \sigma^2(\theta_{\text{proj}}) \left(\frac{\partial P_i}{\partial \theta_1} \frac{\partial P_j}{\partial \theta_1} + \frac{\partial P_i}{\partial \theta_2} \frac{\partial P_j}{\partial \theta_2} \right) \tag{9}$$

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].

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

$$= \frac{13.6}{p} \sqrt{\frac{X}{X_0}} \left[1 + 0.038 \ln \left(\frac{X}{X_0} \right) \right] \tag{11}$$

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 $\implies z = 1$) in the limit of $\beta \to 1$. 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.

$$\sigma(\theta_{\text{proj}}) = 13.6z\sqrt{\sum_i \frac{X_i}{p_i^2\beta_i^2 X_{0,i}}}\left[1+0.038\ln\left(\sum_i \frac{X_i z^2}{\beta_i^2 X_{0,i}}\right)\right] \tag{12}$$

$$= \frac{13.6}{p}\sqrt{\sum_i \frac{X_i}{X_{0,i}}}\left[1+0.038\ln\left(\sum_i \frac{X_i}{X_{0,i}}\right)\right] \tag{13}$$

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 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

$$\omega = \sum_i \frac{X_i}{X_{0,i}} = 0.16 \tag{14}$$

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 $\sin\theta$, where θ is the angle between the particle trajectory and the normal to the layer surface.

$$\sigma(\theta_{\text{proj}}) = \frac{13.6}{p}\sqrt{\frac{\omega}{\sin\theta}}\left[1+0.038\ln\left(\frac{\omega}{\sin\theta}\right)\right] \tag{15}$$

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 Calculation of Q

The co-variance matrix is calculated for the following state vector x_{state}, whose components are used in the MATHUSLA Tracking filter.

$$x_{\text{state}} \doteq P_i = \left[\frac{\Delta y\,\alpha_3}{\beta_3} + x_0, \frac{\Delta y}{c\,\beta_3} + t_0, \frac{\Delta y\,\gamma_3}{\beta_3} + z_0, c\,\alpha_3, c\,\beta_3, c\,\gamma_3\right]_i \tag{16}$$

where (x_0, t_0, z_0) is the current filtered estimate for the hit, Δy is the y spacing between the current and next detector layer, c is the speed of light in vacuum, and $\alpha_3, \beta_3, \gamma_3$ are the direction cosines of the un-scattered track. Namely

$$\alpha_3 \equiv \frac{v_{x,0}}{c} \qquad \beta_3 \equiv \frac{v_{y,0}}{c} \qquad \gamma_3 \equiv \frac{v_{z,0}}{c}$$

The current velocity best estimates. This assumes that for all tracks $\beta = 1$. Following [2], we next calculate the derivatives of each element of the state vector with respect to the scattering angles θ_1 and θ_2 which are orthogonal uncorrelated scattering angles for the track that describe the direction of the scattered track in a coordinate system whose z-axis coincides with the track and has it's orgin at (x_0, t_0, z_0). We simply quote the result of [2] which says that for a rotation that is represented by the following matrix R

$$R = \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{pmatrix} \tag{17}$$

that takes the track coordinate system into the reference coordinate system (used by the filter), the derivatives can be evaluated using the following expressions

$$\begin{aligned} \frac{\partial \alpha^*}{\partial \theta_1} &= \alpha_1 & \frac{\partial \beta^*}{\partial \theta_1} &= \beta_1 & \frac{\partial \gamma^*}{\partial \theta_1} &= \gamma_1 \\ \frac{\partial \alpha^*}{\partial \theta_2} &= \alpha_2 & \frac{\partial \beta^*}{\partial \theta_2} &= \beta_2 & \frac{\partial \gamma^*}{\partial \theta_2} &= \gamma_2 \end{aligned}$$

Where * denotes a direction cosine for the *scattered* track in reference (global) coordinates, and all derivatives are taken in the limit $\theta_1,\ \theta_2 \to 0$ (which also implies $(\alpha^*, \beta^*, \gamma^*) \to (\alpha_3, \beta_3, \gamma_3)$). Since R is in the adjoint representation of $SO(3)$, we have the following useful expressions [2].

$$\alpha_i \beta_i = 0 \qquad \alpha_i \gamma_i = 0 \qquad \beta_i \gamma_i = 0 \tag{18}$$

$$\alpha_i \alpha_i = 1 \qquad \beta_i \beta_i = 1 \qquad \gamma_i \gamma_i = 1 \tag{19}$$

$$\alpha_i \alpha_j + \beta_i \beta_j + \gamma_i \gamma_j = \delta_{ij} \tag{20}$$

Where (19) is an immediate consequence of (20) not given by [2]. Since R is arbitrary (20) applies to the transpose of R as well, for any value of it's elements (since it's transpose is also in $SO(3)$ given that $R^{-1} = R^T$ and the inverse of an element of a group is in that group by the closure axiom). Using the above relations and x_{state} we can calculate the partial derivates in (9) by first letting $\alpha_3 \to \alpha^*$, $\beta_3 \to \beta^*$, and $\gamma_3 \to \gamma^*$ in (16).

$$\begin{aligned}
&\frac{\partial}{\partial \theta_i}\left(\frac{\Delta y\, \alpha^*}{\beta^*} + x_0\right) = \frac{\Delta y\, \alpha_i}{\beta_3} - \frac{\Delta y\, \alpha_3 \beta_i}{{\beta_3}^2} \\
&\frac{\partial}{\partial \theta_i}\left(\frac{\Delta y}{c\, \beta^*} + t_0\right) = -\frac{\Delta y\, \beta_i}{c{\beta_3}^2} \\
&\frac{\partial}{\partial \theta_i}\left(\frac{\Delta y\, \gamma^*}{\beta^*} + z_0\right) = \frac{\Delta y\, \gamma_i}{\beta_3} - \frac{\Delta y\, \gamma_3 \beta_i}{{\beta_3}^2} \\
&\frac{\partial}{\partial \theta_i}\left(c\, \alpha^*\right) = c\, \alpha_i \\
&\frac{\partial}{\partial \theta_i}\left(c\, \beta^*\right) = c\, \beta_i \\
&\frac{\partial}{\partial \theta_i}\left(c\, \gamma^*\right) = c\, \gamma_i
\end{aligned}$$

again in the limit $\theta_1,\ \theta_2 \to 0$. Using (9) we then calculate the Multiple Scattering Process Covariance Matrix, Q.

$$Q = \sigma^2(\theta_{\text{proj}}) \begin{pmatrix}
\frac{\Delta y^2 \left({\beta_3}^2 + {\alpha_3}^2\right)}{{\beta_3}^4} & \frac{\Delta y^2\, \alpha_3}{c\, {\beta_3}^4} & \frac{\Delta y^2\, \alpha_3 \gamma_3}{{\beta_3}^4} & \frac{c\, \Delta y}{\beta_3} & -\frac{c\, \Delta y\, \alpha_3}{{\beta_3}^2} & 0 \\
\frac{\Delta y^2\, \alpha_3}{c\, {\beta_3}^4} & \frac{\Delta y^2 \left(1 - {\beta_3}^2\right)}{c^2\, {\beta_3}^4} & \frac{\Delta y^2\, \gamma_3}{c\, {\beta_3}^4} & \frac{\Delta y\, \alpha_3}{\beta_3} & -\frac{\Delta y \left(1 - {\beta_3}^2\right)}{{\beta_3}^2} & \frac{\Delta y\, \gamma_3}{\beta_3} \\
\frac{\Delta y^2\, \alpha_3 \gamma_3}{{\beta_3}^4} & \frac{\Delta y^2\, \gamma_3}{c\, {\beta_3}^4} & \frac{\Delta y^2 \left({\gamma_3}^2 + {\beta_3}^2\right)}{{\beta_3}^4} & 0 & -\frac{c\, \Delta y\, \gamma_3}{{\beta_3}^2} & \frac{c\, \Delta y}{\beta_3} \\
\frac{c\, \Delta y}{\beta_3} & \frac{\Delta y\, \alpha_3}{\beta_3} & 0 & c^2\, \left(1 - {\alpha_3}^2\right) & -c^2\, \alpha_3 \beta_3 & -c^2\, \alpha_3 \gamma_3 \\
-\frac{c\, \Delta y\, \alpha_3}{{\beta_3}^2} & -\frac{\Delta y \left(1 - {\beta_3}^2\right)}{{\beta_3}^2} & -\frac{c\, \Delta y\, \gamma_3}{{\beta_3}^2} & -c^2\, \alpha_3 \beta_3 & c^2\, \left(1 - {\beta_3}^2\right) & -c^2\, \beta_3 \gamma_3 \\
0 & \frac{\Delta y\, \gamma_3}{\beta_3} & \frac{c\, \Delta y}{\beta_3} & -c^2\, \alpha_3 \gamma_3 & -c^2\, \beta_3 \gamma_3 & c^2\, \left(1 - {\gamma_3}^2\right)
\end{pmatrix}$$

where we have simplified using (18) - (20). It should be noted that $\alpha_k,\ \beta_k,$ and γ_k with $k \in \{1,2\}$ are left arbitrary throughout the calculation.

3 Covariance Matrices Used in the Algorithm

3.1 Covariance Matrices under a Change of Variables

"Suppose we have a set of N random variables y, which may be direct measurements or derived estimators $\hat{\theta}$, and we have a covariance matrix V(y) for these. We can make a transformation to a different set of variables $f_j \equiv f_j(y), j = 1, ..., M (M \leq N)$ and obtain best estimates for the f_j [near $\hat{y}$]" from [5].

$$\hat{f}_j \approx f_j(\hat{y}) + \frac{1}{2}\sum_{n,m}^{N} V_{nm}(\hat{y}) \left[\frac{\partial f_j}{\partial y_n \partial y_m}\right]_{\hat{y}} \tag{21}$$

This has covariance matrix [5]

$$V_{ij}(\hat{f}) \approx \left.\frac{\partial f_i}{\partial y_n}\right|_{\hat{y}} \left.\frac{\partial f_j}{\partial y_m}\right|_{\hat{y}} V_{nm}(\hat{y}) = JV(\hat{y})J^T \tag{22}$$

3.2 Calculation Track Seed to Filter Covariance Matrix

A digitized hit measured in a detector layer is described by the vector m with components

$$m_i = [x, t, z, y]_i \tag{23}$$

This has covariance matrix

$$V = \text{diag}(\sigma_x^2,\ \sigma_t^2,\ \sigma_z^2,\ \sigma_y^2) \tag{24}$$

The seed used to initialise the filter is a pair of hits $m^{(1)}$ and $m^{(2)}$. We label the hits such that $m_y^{(1)} < m_y^{(2)}$. Together these are used to construct the initial state vector

$$x_0 \doteq \left[m_x^{(1)}, m_t^{(1)}, m_z^{(1)}, \frac{m_x^{(2)} - m_x^{(1)}}{m_t^{(2)} - m_t^{(1)}}, \frac{m_y^{(2)} - m_y^{(1)}}{m_t^{(2)} - m_t^{(1)}}, \frac{m_z^{(2)} - m_z^{(1)}}{m_t^{(2)} - m_t^{(1)}}\right] \equiv \left[m_x^{(1)}, m_t^{(1)}, m_z^{(1)}, \frac{\Delta x}{\Delta t}, \frac{\Delta y}{\Delta t}, \frac{\Delta z}{\Delta t}\right] \tag{25}$$

So we use (21) and (22) with

$$\hat{y} = m^{(1)} \oplus m^{(2)} \qquad\qquad V(\hat{y}) = V^{(1)} \oplus V^{(2)} \tag{26}$$

Then we have

$$J = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ -\frac{1}{\Delta t} & \frac{\Delta x}{(\Delta t)^2} & 0 & 0 & -\frac{1}{\Delta t} & -\frac{\Delta x}{(\Delta t)^2} & 0 & 0 \\ 0 & \frac{\Delta x}{(\Delta t)^2} & 0 & -\frac{1}{\Delta t} & 0 & -\frac{\Delta x}{(\Delta t)^2} & 0 & \frac{1}{\Delta t} \\ 0 & \frac{\Delta x}{(\Delta t)^2} & -\frac{1}{\Delta t} & 0 & 0 & -\frac{\Delta x}{(\Delta t)^2} & \frac{1}{\Delta t} & 0 \end{pmatrix} \tag{27}$$

3.3 Calculation of Track to Vertex Seed Covariance Matrix

Given the Covariance Matrices of the two state vectors describing two tracks $x_{state,\ i}$, $P_i \equiv Cov\{x_{state,\ i}\}$, we look to calculate the covariance matrix of the vertex seed estimate calculated from two parametric vectors. The vertex seed estimate is taken to be the average predicted position of the vectors (particles) at the time of the closest approach between the two vectors. This is described in detail here [4]. This seed estimate is used to initialise the vertex position vector in the Kalman Filter Vertexing Algorithm.

The state vector for our vertex seed estimate is

$$f_i = [x + v_x \Delta t, y + v_y \Delta t, z + v_z \Delta t, t + \Delta t]_i \tag{28}$$

where

$$x_{state} \doteq [x, y, z, t, v_x, v_y, v_z] \tag{29}$$

is the track parameter state vector at the lowest layer of the first track used in the seed. The vertex position seed estimate is the particle position at the time of closest approach. Where [4]

$$t_{CA} \equiv \frac{(\vec{x}_1 - \vec{x}_0) \cdot (\vec{v}_0 - \vec{v}_1)}{|\vec{v}_0 - \vec{v}_1|^2} \qquad \Delta t = t_{CA} - t \tag{30}$$

is the time of closest approach, and we have labeled the two tracks by 0, and 1.

$$\Delta x_i = (x_1)_i - (x_0)_i \qquad \Delta v_i = (v_0)_i - (v_1)_i \tag{31}$$

in coordinates where the parametric vectors are defined for $t \geq 0$, and the v's and x's are 3-velocities and 3-positions of each track. The Jacobian of the change of variables from

$$y_n = x_{state,n} \tag{32}$$

to f_i is given by

$$\frac{\partial f_i}{\partial y_n} = \frac{1}{v^2} \begin{pmatrix} v^2 + v_x \Delta v_x & -v_x & v_x \Delta v_z & v^2 \Delta t + v_x r_x & v_x r_y & v_x r_z \\ v_y \Delta v_x & -v_y & v_y \Delta v_z & v_y r_x & v^2 \Delta t + v_y r_y & v_y r_z \\ v_z \Delta v_x & -v_z & v^2 + v_z \Delta v_z & v_z r_x & v_z r_y & v^2 \Delta t + v_z r_z \\ \Delta v_x & 0 & \Delta v_z & r_x & r_y & r_z \end{pmatrix} \tag{33}$$

where

$$r_i \equiv (\Delta x_i - 2 \Delta v_i t_{CA}) \tag{34}$$

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