Kalman Filtering

Martin Völkl

2021-02-11

UNIVERSITÄT
HEIDELBERG
ZUKUNFT
SEIT 1386

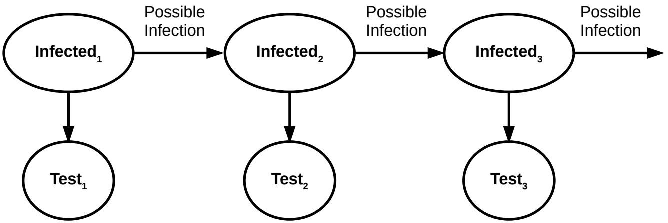
Reminder: Updating knowledge with data

- Example: Daily (independent) tests for disease (e.g. for medical professionals)
- Want to know sick (s) or healthy (h) from positive ($+$) or negative ($-$) test result
- Likelihood e.g. $p(+|s) = p(-|h) = 0.9$
- We know how to update: Bayes rule (e.g. negative result, 0.5 prior)

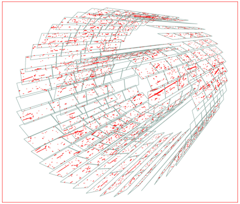
$$p(s|-) = \frac{p(-|s)p(s)}{p(-|s)p(s) + p(-|h)p(h)} = 0.1$$

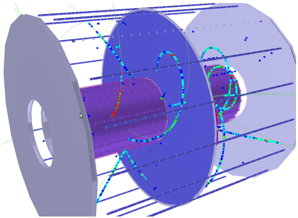
- Prior for the next (independent) measurement and so on
- But it is also possible to become sick from one day to the next

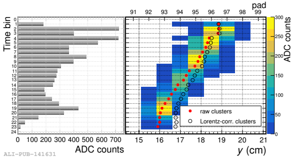
Measuring a changing state



- What if there is a 5% chance to become sick every day?
- This means the true state can change – aiming for a moving target
- Still straightforward calculation
- If $p_{yesterday}(s) = 0.1$ and there is a 5% chance to become sick, then $p_{today}(s) = 0.1 + (1 - 0.1) \cdot 0.05 = 0.145$
- This is now the prior for todays test
- General idea: Update state (possibility of infection), apply knowledge (test) using Bayes, repeat







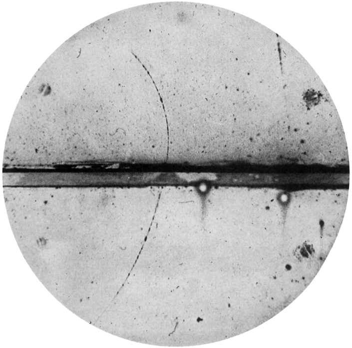
- Silicon Pixel Detector
- Rectangular active semiconductor detector regions
- Signal essentially true/false

- Time Projection Chamber
- Ionization drifts towards readout pads at the end
- Reconstruct 3D clusters of signal using drift time since event

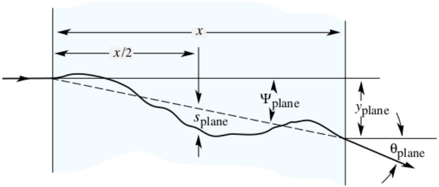
- Transition Radiation Detector
- Signal from ionization + potentially Transition radiation
- Drift in radial direction, reconstruct from arrival time

- Different types of measurements and measurements of different quantities need to be combined for tracking
- Different dimension of the information and different coordinate systems

Interaction of tracks and the detector

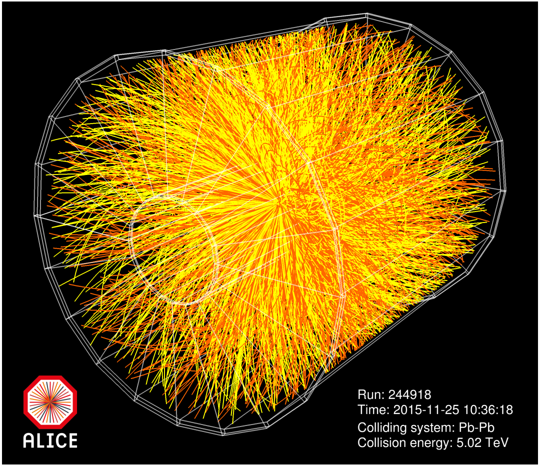


- Tracks interact with detector material
- "State" changes



- At high momenta: Changes in direction through Coulomb scattering $\sim \sqrt{x}$
- $\theta_{rms} = \frac{13.6\ \mathrm{MeV}}{\beta cp} z \sqrt{\frac{x}{X_0}} \left(1 + 0.088 \log_{10} \frac{xz^2}{X_0 \beta^2}\right)$
- Track parameters are a moving target

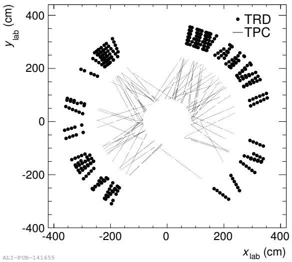
Another Complication



- Often more than one track creating signals – need way to separate contributions

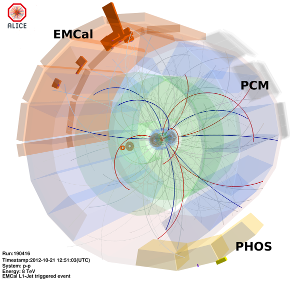
Ideal Tracks

- Useful to parametrize tracks with some time variable



Cosmic muons without magnetic field (ALICE)

- Straight line
- Parametrization $\vec{x}(t) = \vec{x}_0 + \vec{v}_0 t$

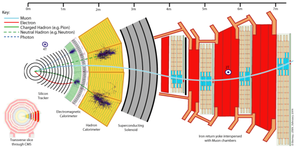


- Homogeneous magnetic field – helix shape

$$x = x_0 + r\cos(\omega t + \phi_0)$$

$$y = y_0 + r\sin(\omega t + \phi_0)$$

$$z = v_z t$$



CMS with muon track

- More complex fields give more complex track shapes

The state vector

- Locally a 3D track can usually be described by 5 parameters
- e.g. A track at Radius r has two free position parameters and three for the momentum vector
- For ALICE two positions, track curvature, helix center position, dip angle
- In two dimensions described by 2 (line) or 3 (circle) parameters
- Uncertainty expressed by covariance matrix of parameters
- This is the state vector of the particle for a given time variable

Defining the problem

- Similar as in introduction problem: Track state changes between measurements
- First step is to change problem to something easier:
  - Assume linear state change (by matrix multiplication)
  - Assume that all state information can be characterized completely by their state vector and its covariance matrix
  - Assume measurements are unbiased and can be completely described by the measurement covariance matrix (in an appropriate coordinate system)
  - Assume processes affecting the track state are unbiased and completely described by the process covariance matrix
  - Different measurements and process noises are all independent
- If all distributions can be considered Gaussian and they only propagate via linear transformations, all results must also be Gaussian
- This suggests that all Bayesian inference should be expressable in terms of linear algebra
- The iterative algorithm for this is called the **Kalman filter**
- Which input information is needed?

Ingredients

- The state vector $\vec{s}$ and its covariance matrix $\mathbf{S}$
- A covariance matrix for the process noise $\mathbf{W}$ in the coordinates of the state vector
- The measurement vector $\vec{m}$ in the coordinate system of the measurement (can be different dimensionality)
- A covariance matrix for the measurement uncertainty $\mathbf{V}$
- A Matrix $\mathbf{H}$ transforming from the state vector into the measurement space
  - If $\vec{s}$ is the current state, then $\mathbf{H}\vec{s}$ is the expectation value for the signal
- A Matrix $\mathbf{F}$ transforming the state vector at the position of the last measurement to the one at the next measurement

The Kalman Filter

- Assuming we know the state $\vec{s}_{k-1|k-1}$ at the position of the previous measurement with all information of the measurements until then

First: Update state and covariance to current position including process noise

$$\vec{s}_{k|k-1} = \mathbf{F}_k \vec{s}_{k-1|k-1}$$

$$\mathbf{S}_{k|k-1} = \mathbf{F}_k \mathbf{S}_{k|k-1} \mathbf{F}_k^T + \mathbf{W}_k$$

Second: Taking this as the prior, update knowledge with information from measurement

Define:

$$\mathbf{K}_k = \mathbf{S}_{k|k-1} \mathbf{H}_k^T \left(\mathbf{V}_k + \mathbf{H}_k \mathbf{S}_{k|k-1} \mathbf{H}_k^T \right)^{-1}$$

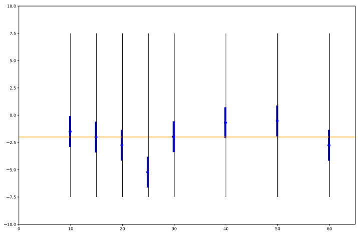
$$\vec{s}_{k|k} = \vec{s}_{k|k-1} + \mathbf{K}_k \left(\vec{m} - \mathbf{H}_k \vec{s}_{k|k-1} \right)$$

$$\mathbf{S}_{k|k} = \left(\mathbf{I} - \mathbf{K}_k \mathbf{H}_k \right) \mathbf{S}_{k|k-1}$$

- $\mathbf{K}$ is called the **Kalman gain matrix**
- Change to state depends on difference of measurement to expectation

Simple example: Track y position

- Known track source, detector far away – horizontal line with unknown $s = y$, time variable is x position
- Measure position y_k, with measurement uncertainty $\sigma^2_{meas,k}$ assume no process noise, $F = 1$



State update:

$$s_{k|k-1} = s_{k-1|k-1}, \sigma^2_{k|k-1} = \sigma^2_{k-1|k-1}$$

$$K = \frac{\sigma^2_{k-1}}{\sigma^2_{k-1} + \sigma^2_{meas,k}}$$

Measurement update:

$$s_{k|k} = s_{k|k-1} + \frac{(y_k - s_{k|k-1})}{\sigma^2_{k|k-1} + \sigma^2_{meas,k}}$$

$$\sigma^2_{k|k} = \left(1 - \frac{\sigma^2_{k-1}}{\sigma^2_{k-1} + \sigma^2_{meas,k}}\right) \sigma^2_{k|k-1}$$

Simple example: Track y position (2)

$$s_{k|k} = s_{k|k-1} + \frac{(y_k - s_{k|k-1})}{\sigma^2_{k|k-1} + \sigma^2_{meas,k}}$$

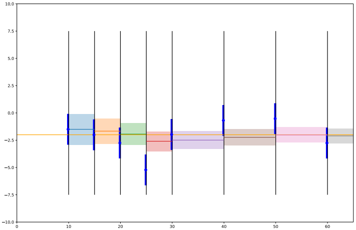
$$\sigma^2_{k|k} = \left(1 - \frac{\sigma^2_{k-1}}{\sigma^2_{k-1} + \sigma^2_{meas,k}}\right)\sigma^2_{k|k-1}$$

- Iteratively updates position knowledge
- Update can be rewritten as:

$$s_{k|k} = \frac{y_k/\sigma^2_{meas,k} + s_{k|k-1}/\sigma^2_{k|k-1}}{1/\sigma^2_{k|k-1} + 1/\sigma^2_{meas,k}}$$

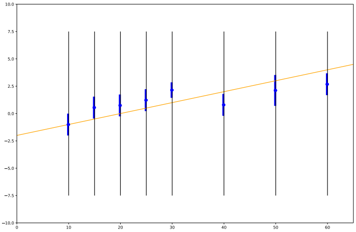
$$\sigma^2_{k|k} = \frac{1}{1/\sigma^2_{k|k-1} + 1/\sigma^2_{meas,k}}$$

- This is exactly the result from the weighted mean!



Tracking in two dimensions – State update

- Now slope and position are unknown, track model $y = \alpha x + \beta$; x is the time variable
- State vector defined as (β, α)
- β_k is the y-Position of the track at detector position x_k
- No process noise $\mathbf{W} = 0$
- State update $\alpha \to \alpha$, $\beta \to \beta + \alpha \cdot \Delta x$ with $\Delta x = x_k - x_{k-1}$



State update:

$$\mathbf{F}_k = \begin{pmatrix} 1 & \Delta x_k \\ 0 & 1 \end{pmatrix}$$

$$\vec{s}_{k|k-1} = \mathbf{F}_k \vec{s}_{k-1|k-1}$$

$$\mathbf{S}_{k|k-1} = \mathbf{F}_k \mathbf{S}_{k|k-1} \mathbf{F}_k^T$$

Tracking in two dimensions – Initial guess

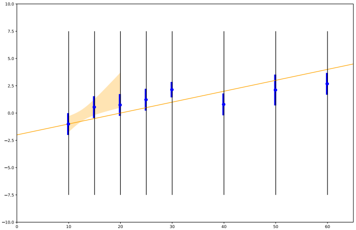
State update:

$$\mathbf{F}_k = \begin{pmatrix} 1 & \Delta x_k \\ 0 & 1 \end{pmatrix}$$

$$\vec{s}_{k|k-1} = \mathbf{F}_k \vec{s}_{k-1|k-1}$$

$$\mathbf{S}_{k|k-1} = \mathbf{F}_k \mathbf{S}_{k|k-1} \mathbf{F}_k^T$$

- For the initial guess, make a minimum χ^2 fit through the first two points (the *track seed*)
- Result is parameters for $x = 0$, can then propagate to second point
- First step of Kalman filter is then to propagate to third point – the first new measurement



Tracking in two dimensions – Initial guess

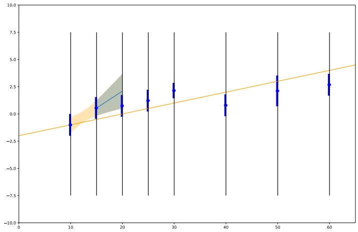
State update:

$$\mathbf{F}_k = \begin{pmatrix} 1 & \Delta x_k \\ 0 & 1 \end{pmatrix}$$

$$\vec{s}_{k|k-1} = \mathbf{F}_k \vec{s}_{k-1|k-1}$$

$$\mathbf{S}_{k|k-1} = \mathbf{F}_k \mathbf{S}_{k|k-1} \mathbf{F}_k^T$$

- For the initial guess, make a minimum χ^2 fit through the first two points (the *track seed*)
- Result is parameters for $x = 0$, can then propagate to second point
- First step of Kalman filter is then to propagate to third point – the first new measurement



Tracking in two dimensions – Measurement Update

- Measurement is the y position
- $\mathbf{H}_k \vec{s}_{k|k-1}$ gives the expected measurement
- $\vec{s} = (\beta, \alpha)$, thus $\mathbf{H}_k = (1\ 0)$ projects out the position

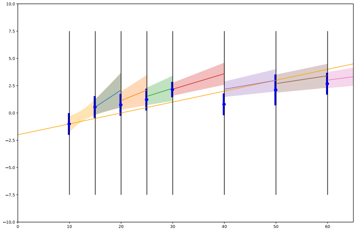
Define:

$$\mathbf{K}_k = \mathbf{S}_{k|k-1} \mathbf{H}_k^T \left(\mathbf{V}_k + \mathbf{H}_k \mathbf{S}_{k|k-1} \mathbf{H}_k^T \right)^{-1}$$

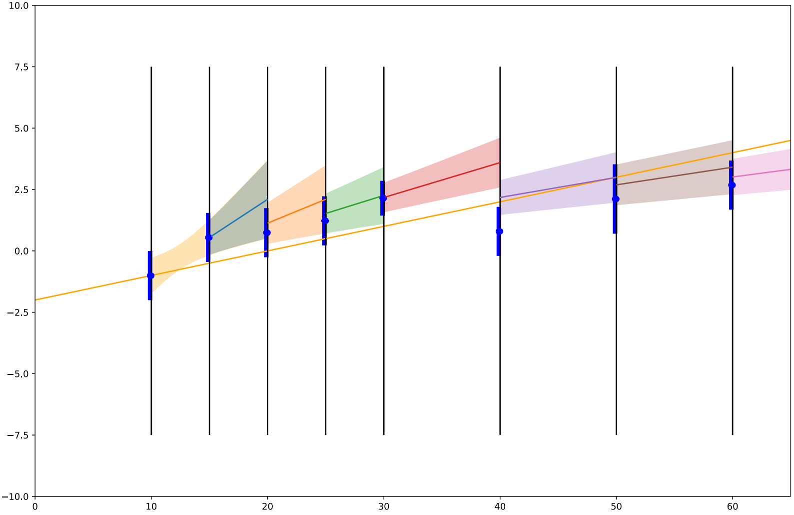
$$\mathbf{K}_k = \left(\mathbf{S}_{\beta,\beta}\ \mathbf{S}_{\beta,\alpha} \right) / \left(\sigma^2_{meas,k} + \mathbf{S}_{\beta,\beta} \right)$$

$$\vec{s}_{k|k} = \vec{s}_{k|k-1} + \mathbf{K}_k \left(y_{meas,k} - \beta_{k|k-1} \right)$$

$$\mathbf{S}_{k|k} = \left(\mathbf{I} - \mathbf{K}_k \mathbf{H}_k \right) \mathbf{S}_{k|k-1}$$



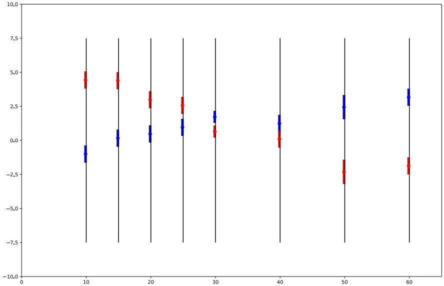
- If measurement uncertainty small, then **K** becomes of order 1, lots of additional information



- Result equivalent to least squares fit

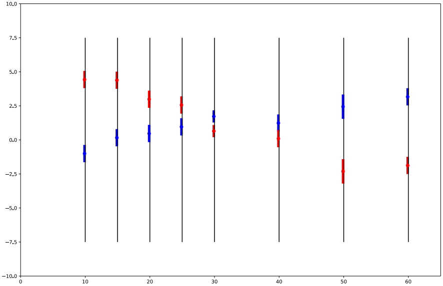
Why not just use least squares?

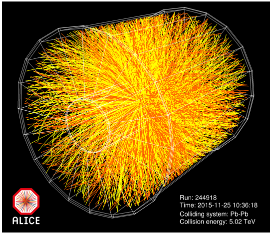
- If it is equivalent, why not just use a least squares fit?
- If there are several tracks, which signal to use for a particular detector layer?
- Possible approach: Find combination with best χ^2 and fit, then remove associated signals; repeat
- For two tracks and 8 detectors: Need to try 2^8 combinations



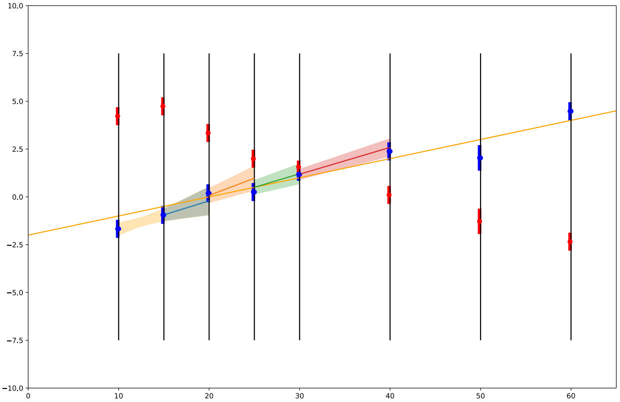
Why not just use least squares?

- If it is equivalent, why not just use a least squares fit?
- If there are several tracks, which signal to use for a particular detector layer?
- Possible approach: Find combination with best χ^2 and fit, then remove associated signals; repeat
- For two tracks and 8 detectors: Need to try 2^8 combinations
- For ALICE $\approx$ 165 signals per track, $\sim$ 2000 tracks in the detectors
- $2000^{165} \sim 10^{544}$ combinations, too much for brute force





Track Finding



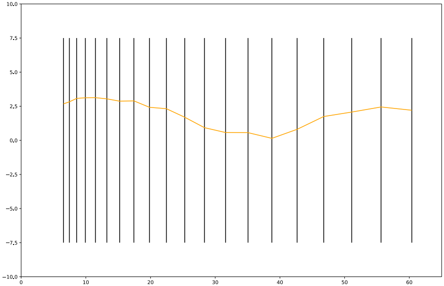
- Search for the next detector signals in the range suggested by the prior after updating the position
- E.g. in the ALICE TPC, clusters are searched for within 4σ
- Use the Kalman filter on several tracks in parallel to resolve ambiguities
- Or explore several possible continutions in parallel and see which one has the best χ^2

Tracks with process noise

- Now include process noise
- Multiple scattering does not change position but direction
- Change $\alpha \to \alpha + n$, where the noise is drawn from a Gaussian with width σ_p
- This also means that the track slowly loses information about its initial state
- Process noise can be quantified by the covariance matrix

$$\mathbf{W} = \begin{pmatrix} 0 & 0 \\ 0 & \sigma_p^2 \end{pmatrix}$$

since it only affects the slope

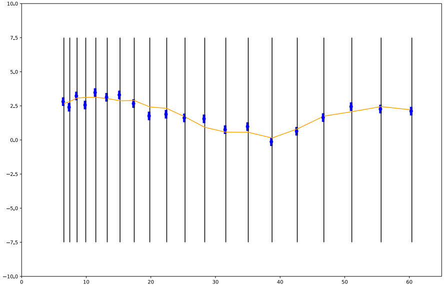


Tracks with process noise

- Now include process noise
- Multiple scattering does not change position but direction
- Change $\alpha \to \alpha + n$, where the noise is drawn from a Gaussian with width σ_p
- This also means that the track slowly loses information about its initial state
- Process noise can be quantified by the covariance matrix

$$\mathbf{W} = \begin{pmatrix} 0 & 0 \\ 0 & \sigma_p^2 \end{pmatrix}$$

since it only affects the slope



Tracking reconstruction with process noise

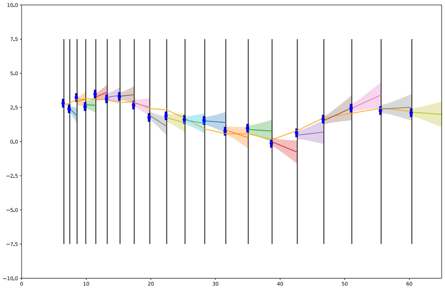
State update:

$$\mathbf{F}_k = \begin{pmatrix} 1 & \Delta x_k \\ 0 & 1 \end{pmatrix}$$

$$\vec{s}_{k|k-1} = \mathbf{F}_k \vec{s}_{k-1|k-1}$$

$$\mathbf{S}_{k|k-1} = \mathbf{F}_k \mathbf{S}_{k|k-1} \mathbf{F}_k^T + \mathbf{W}$$

- The only thing that changes is the addition of the $+\mathbf{W}$
- Measurement update stays the same
- As measurement is in position, the uncertainty in the slope becomes large; visible in propagation

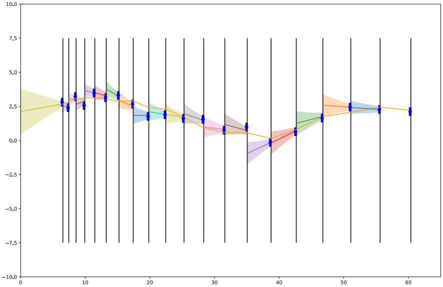


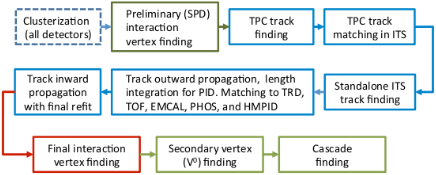
- In the end: Full information about track parameters at the end of the detector
- Of interest: Track parameters at beginning of track (e.g. for invariant mass, decay vertex reconstruction, interaction vertex reconstruction etc.)

Information about track parameters at other positions

- The initial track parameters can be found by using the Kalman filter backwards
- Result shows, why it is useful to have little material in the inner layers of the detector
- Can be useful to go one way for track finding and then backwards for the parameters
- Information at intermediate stages, by Filtering from both sides and averaging at the intermediate positions, called a **Kalman Smoother**
- Real-life reconstruction can be even more complex depending on e.g. which detector gives the best track seeds, which initial propagation direction works best etc.

Filtering backwards





ALICE event reconstruction scheme

Nonlinear dependencies

- The Kalman filter assumes track propagation and measurement outputs are linear in the track parameters + noise
- Generalization as usual by linear approximation

Track propagation:

$$\vec{s}_{k|k-1} = \mathbf{F}_k \vec{s}_{k-1|k-1}$$

$$\rightarrow \vec{s}_{k|k-1} = \vec{f}_k(\vec{s}_{k-1|k-1}, \vec{w}_{k-1})$$

with process noise $\vec{w}_{k-1}$
Take derivatives wrt. s and w:

$$(\mathbf{A}_k)_{ij} = \frac{\partial (f_k)_i}{\partial (s_k)_j}(\vec{s}_{k-1}, 0)$$

$$(\mathbf{B}_k)_{ij} = \frac{\partial (f_k)_i}{\partial (w_k)_j}(\vec{s}_{k-1}, 0)$$

Modify propagation equations:

$$\vec{s}_{k|k-1} = \vec{f}_k(\vec{s}_{k-1|k-1}, \vec{w}_{k-1})$$

$$\mathbf{S}_{k|k-1} = \mathbf{A}_k \mathbf{S}_{k-1|k-1} \mathbf{A}_k^T + \mathbf{B}_k \mathbf{W}_{k-1} \mathbf{B}_k^T$$

Knowledge updating:

$$\mathbf{H}_k \vec{s}_{k|k-1} \rightarrow \vec{h}_k(\vec{s}_{k|k-1}, \vec{v}_k)$$

with measurement noise $\vec{v}_k$
Take derivatives wrt. s and w

$$(\mathbf{P}_k)_{ij} = \frac{\partial (h_k)_i}{\partial (s_k)_j}(\vec{s}_{k-1}, 0)$$

$$(\mathbf{Q}_k)_{ij} = \frac{\partial (h_k)_i}{\partial (v_k)_j}(\vec{s}_{k-1}, 0)$$

Modify measurement update equations:

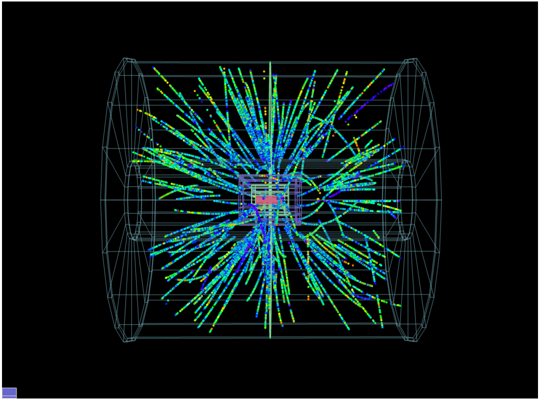
$$\mathbf{K}_k = \mathbf{S}_{k|k-1} \mathbf{P}_k^T \left(\mathbf{Q}_k \mathbf{V}_k \mathbf{Q}_k^T + \mathbf{P}_k \mathbf{S}_{k|k-1} \mathbf{P}_k^T \right)^{-1}$$

$$\vec{s}_{k|k} = \vec{s}_{k|k-1} + \mathbf{K}_k \left(\vec{m} - \vec{h}_k(\vec{s}_{k|k-1}, 0) \right)$$

$$\mathbf{S}_{k|k} = (\mathbf{I} - \mathbf{K}_k \mathbf{P}_k) \mathbf{S}_{k|k-1}$$

Concluding remarks

- Kalman filter computationally efficient way to include tracking information
- Includes state change by interaction with material
- Can include many types of detector signals



- Much complexity also in finding detector signals, separating them from each other, removing bias, and track seed finding, not discussed here