Data association techniques in multiple object tracking

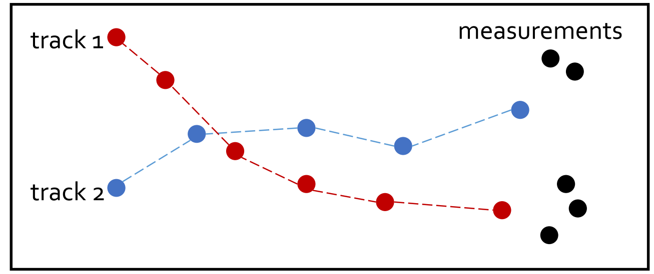
April 27th 2018

Hyojoon Park

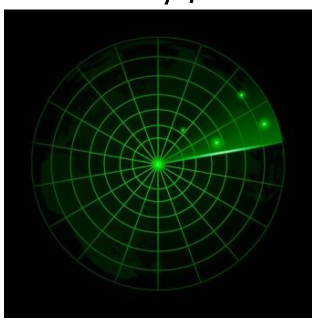
Contents

Introduction

- **What is data association?**
  - Process of associating uncertain measurements to known tracks.
  - **Tracking = Data association + Filtering**

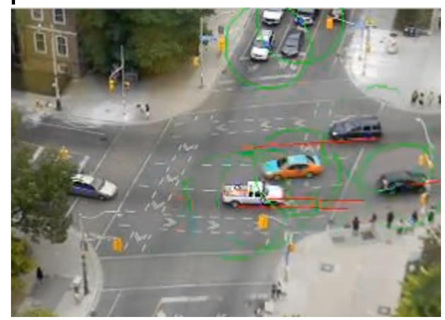


- **When is it used?**
  - Initially ,radars → now, all kinds of tracking problems.

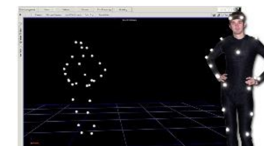


radar

laboratory



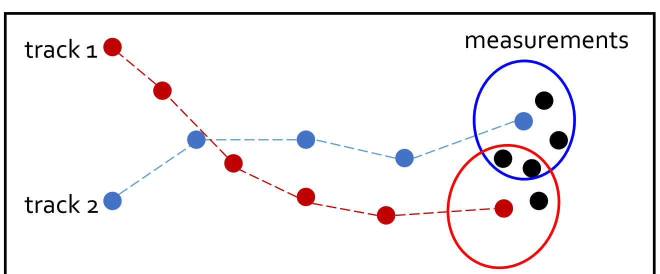
surveillance

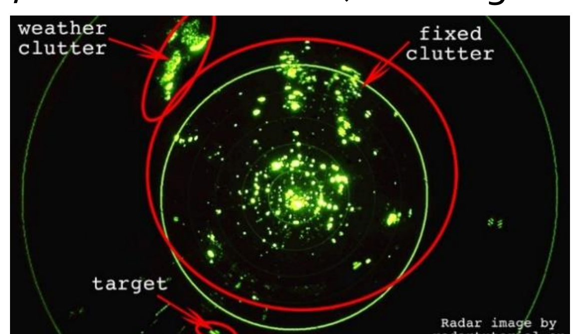


motion capture

fleet management

- **Problem?**
  - multiple tracks → association ambiguity
  - false alarm/clutter → shouldn't detect, but detected (false positive)
  - detection uncertainty → should detect, but not detected (false negative)





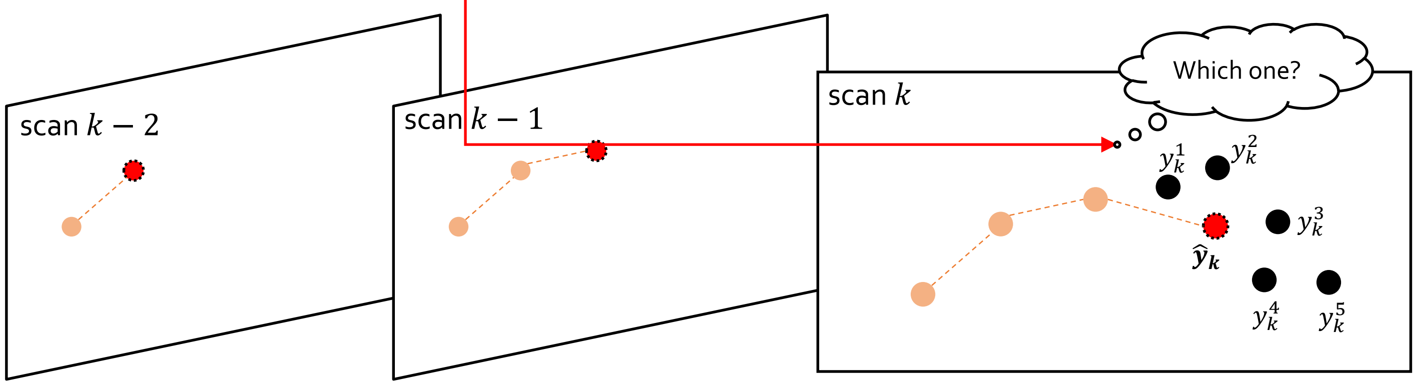
Introduction

- System is assumed to behave as:
  - state vector: $x(k) = F(k-1)x(k-1) + v(k-1) \in \mathbb{R}^{n_1}$
  - measurement vector: $y(k) = H(k)x(k) + w(k) \in \mathbb{R}^{n_2}$
    - $v(k-1) \sim \mathcal{N}[0, Q(k-1)], w(k) \sim \mathcal{N}[0, R(k)]$: Gaussian noise

1. Predict state, measurement, and state covariance
   - a *prior* state estimate: $\hat{x}_{k|k-1} = F_{k-1}\hat{x}_{k-1|k-1}$
   - predicted measurement: $\hat{y}_k = H_k \hat{x}_{k|k-1}$
   - a *priori* state covariance: $P_{k|k-1} = F_{k-1}P_{k-1|k-1}{F_{k-1}}^T + Q_{k-1}$

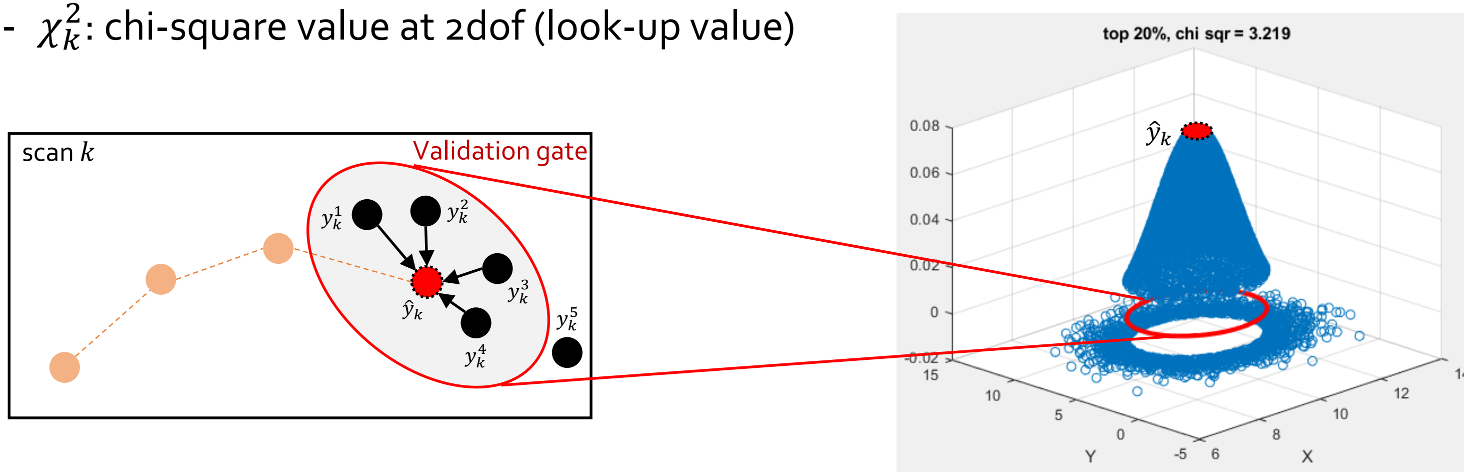
Validation gating and **data association** process

2. Update state using measurement
   - innovation vector: $z_k = \boxed{y_k} - \hat{y}_k$
   - a *posterior* state estimate: $\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k z_k$



Introduction

- **Validation Gating**
  → Rule out measurements that are geometrically unlikely from the start.
  - Assuming Gaussian probability density for the residual z_k follow:
    $$p(z_k) = p(y_k - \hat{y}_k) = \mathcal{N}\left(z_k; 0, \hat{S}_k\right) = \frac{1}{2\pi\sqrt{|\hat{S}_k|}} e^{-\frac{1}{2}{z_k}^T {\hat{S}_k}^{-1} z_k}$$
  - Validation gate using **Mahalanobis distance**:
    $$\left(y_k^i - \hat{y}_k\right)^T {\hat{S}_k}^{-1}\left(y_k^i - \hat{y}_k\right) < \chi_k^2$$
    where,
    - $\hat{S}_k = H_k P_{k|k-1} {H_k}^T$ = residual covariance (uncertainty in $\hat{y}_k$)
    - χ_k^2: chi-square value at 2dof (look-up value)



  - In implementation perspective, use **Euclidean distance** to avoid computing ${\hat{S}_k}^{-1}$.
    - Gate calculation costs a great part of the total MHT calculation time. (discussed later)

Data association techniques

I. Non-Bayesian
   1. Nearest Neighbor Standard Filter (NNSF)
   2. Global Nearest Neighbor (GNN)

II. Bayesian
   1. Probabilistic Data Association Filter (PDAF)
   2. Joint Probabilistic Data Association Filter (JPDAF)
   3. Multiple Hypothesis Tracking (MHT)

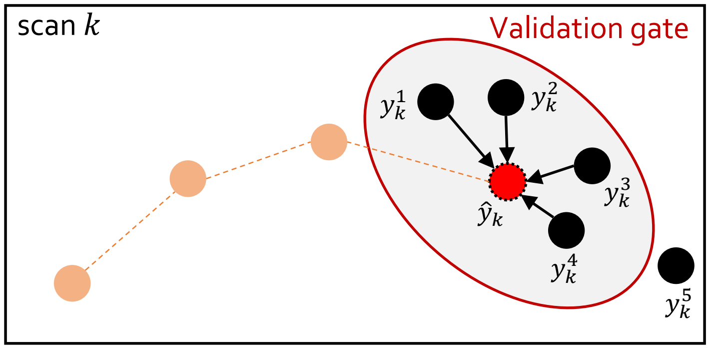
Data association techniques

I. Non-Bayesian
   1. **Nearest Neighbor Standard Filter (NNSF)**
   2. Global Nearest Neighbor (GNN)

II. Bayesian
   1. Probabilistic Data Association Filter (PDAF)
   2. Joint Probabilistic Data Association Filter (JPDAF)
   3. Multiple Hypothesis Tracking (MHT)

Nearest Neighbor Standard Filter (NNSF)

- Key idea
  - choose one **closest** measurement.



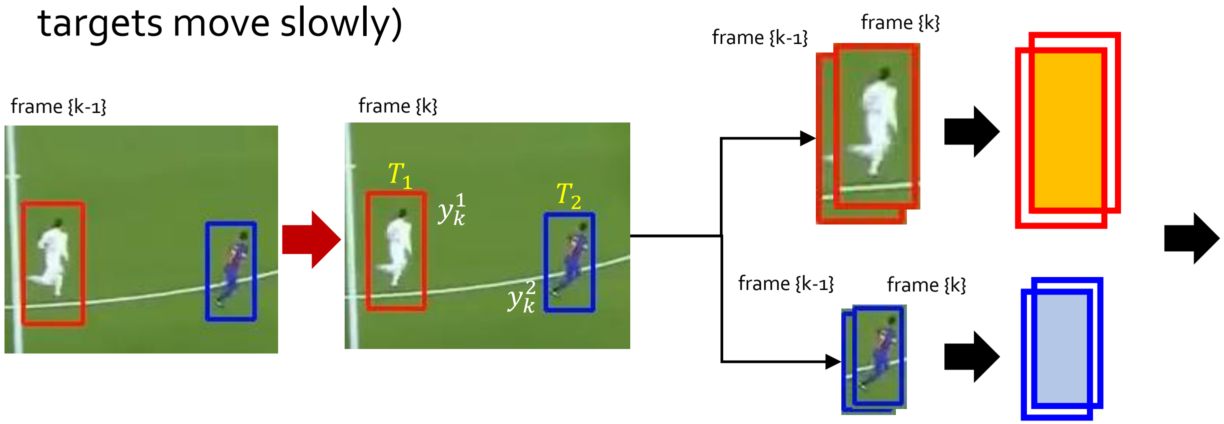
Assignment matrix

$$\begin{array}{cc} & T_{i1} \\ \begin{array}{c} y_k^1 \\ y_k^2 \\ y_k^3 \\ y_k^4 \end{array} & \begin{bmatrix} 8.0 \\ 5.0 \\ 4.0 \\ \boxed{2.0} \end{bmatrix} \end{array} \leftarrow \left(y_k^i - \hat{y}_k\right)^T \hat{S}_k^{\,-1} \left(y_k^i - \hat{y}_k\right)$$

  - update state estimate using y_k^4

    → a *posterior* state estimate: $\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k\left(y_k^4 - \hat{y}_k\right)$

- Can use other scoring criteria other than distance.
  - Affinity scores such as size, shape, appearance, etc.
  - e.g. location, size and shape similarity can be scored based on bounding box **overlap** (if targets move slowly)

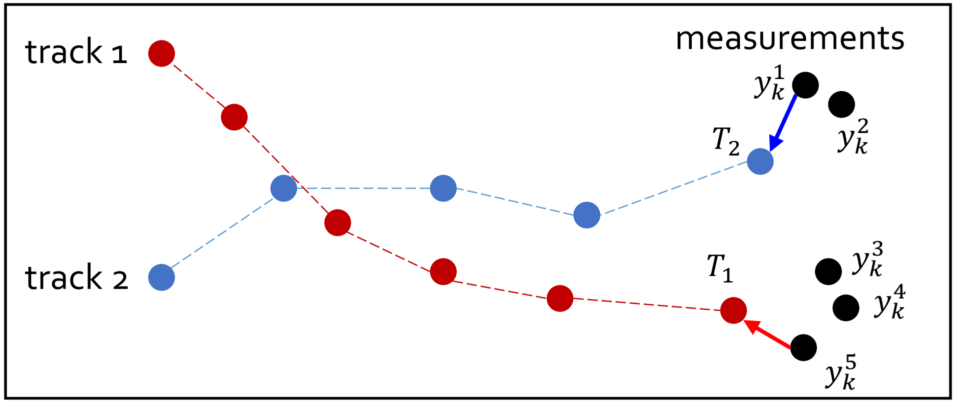


$$\begin{array}{ccc} & T_{i1} & Ti2 \\ \begin{array}{c} y_k^1 \\ y_k^2 \end{array} & \multicolumn{2}{c}{\begin{bmatrix} \boxed{0.8} & 0 \\ 0 & \boxed{0.9} \end{bmatrix}} \end{array}$$

Nearest Neighbor Standard Filter (NNSF)

- Implementation:
  1. Build assignment matrix, each element:

$$d_{ij}^{\,2} = \left(y_k^i - \hat{y}_k\right)^T \hat{S}_k^{\,-1}\left(y_k^i - \hat{y}_k\right)$$



Initialization

$$A = \begin{array}{c} \text{measurements} \\ \begin{bmatrix} d_{11}^{\,2} & d_{12}^{\,2} & d_{13}^{\,2} & d_{14}^{\,2} & d_{15}^{\,2} \\ d_{21}^{\,2} & d_{22}^{\,2} & d_{23}^{\,2} & d_{24}^{\,2} & d_{25}^{\,2} \end{bmatrix} \end{array} \text{tracks}$$

$$= \begin{bmatrix} 200 & 170 & 36.25 & 52 & 28.5 \\ 25 & 25 & 56.25 & 89 & 127.25 \end{bmatrix}$$

Loop

2. Find the closest pairing in A.

$$= \begin{bmatrix} 200 & 170 & 36.25 & 52 & 28.5 \\ \boxed{25} & 26 & 56.25 & 89 & 127.25 \end{bmatrix}$$

3. Remove row & column of that pairing.

$$= \begin{bmatrix} \cancel{200} & 170 & 36.25 & 52 & 28.5 \\ \boxed{25} & \cancel{26} & \cancel{56.25} & \cancel{89} & \cancel{127.25} \end{bmatrix}$$

4. Find next closest pairing in A.

$$= \begin{bmatrix} \cancel{200} & 170 & 36.25 & 52 & \boxed{28.5} \\ \boxed{25} & \cancel{26} & \cancel{56.25} & \cancel{89} & \cancel{127.25} \end{bmatrix}$$

5. Remove row & column of that pairing.

$$= \begin{array}{c} \begin{bmatrix} \cancel{200} & \cancel{170} & \cancel{36.25} & \cancel{52} & \boxed{28.5} \\ \boxed{25} & \cancel{26} & \cancel{56.25} & \cancel{89} & \cancel{127.25} \end{bmatrix} \begin{array}{c} \boldsymbol{T_1} \\ \boldsymbol{T_2} \end{array} \\ \begin{array}{ccccc} \boldsymbol{y_k^1} & \boldsymbol{y_k^2} & \boldsymbol{y_k^3} & \boldsymbol{y_k^4} & \boldsymbol{y_k^5} \end{array} \end{array}$$

Update

6. No elements left → done.
7. Update all tracks with associated measurements.

- Generally does **not** find the global minimum.
- Cannot handle association ambiguities.

→ Instead of NNSF, for multiple tracks use **Global Nearest Neighbor(GNN)**

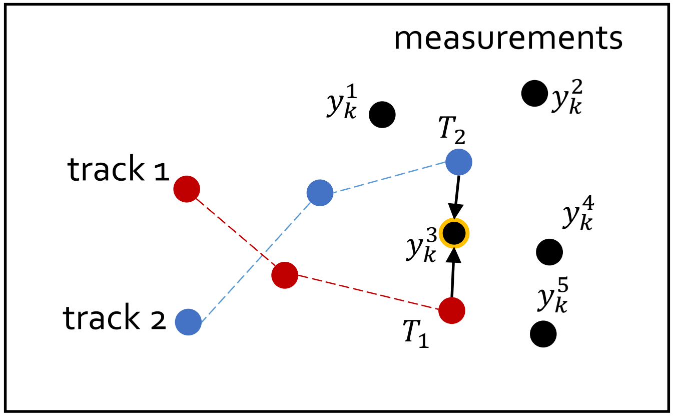
Data association techniques

I. Non-Bayesian
   1. Nearest Neighbor Standard Filter (NNSF)
   2. **Global Nearest Neighbor (GNN)**

II. Bayesian
   1. Probabilistic Data Association Filter (PDAF)
   2. Joint Probabilistic Data Association Filter (JPDAF)
   3. Multiple Hypothesis Tracking (MHT)

Global Nearest Neighbor (GNN)

- Problem: both tracks want y_k^3
- Linear Assignment Problem
  → joint data association, finds global minimum.
- "Best assignment of n different workers to n jobs"



Given a cost matrix $C = \{c_{ij}\}$,

$$min \sum c_{ij} x_{ij}\,, \quad x_{ij} \in \{1, 0\}$$

$$s.t. \boxed{\textstyle\sum_i x_{ij} = 1}, \boxed{\textstyle\sum_j x_{ij} = 1}$$

sum of each row = 1

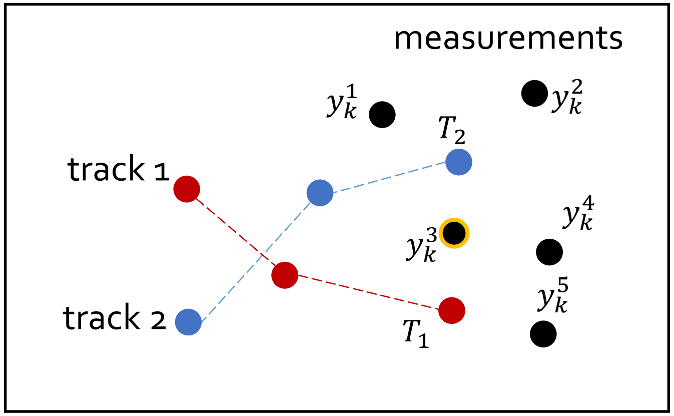
sum of each column = 1

e.g. $\begin{array}{cc} & \begin{array}{ccc} T_1 & T_2 & T_3 \end{array} & \\ & \begin{bmatrix} 108 & 125 & \mathbf{150} \\ 150 & \mathbf{135} & 175 \\ \mathbf{122} & 148 & 250 \end{bmatrix} & \begin{array}{c} y_k^1 \\ y_k^2 \\ y_k^3 \end{array} \end{array} \rightarrow \begin{bmatrix} 0 & 0 & \mathbf{1} \\ 0 & \mathbf{1} & 0 \\ \mathbf{1} & 0 & 0 \end{bmatrix}$, with minimum cost: $\sum c_{ij} x_{ij} = 122 + 135 + 150 = \mathbf{407}$

- **Hungarian Algorithm** to solve.
  For non-square cost matrix, augment it with zeros to make it square.
- Developed by American mathematician Harold Kuhn in 1955.
- Time complexity $O(n^4)$

Global Nearest Neighbor (GNN)

- Implementation:



1. Build cost(assignment) matrix, each element:

$$c_{ij} = \left(y_k^i - \hat{y}_k\right)^T \hat{S}_k^{-1} \left(y_k^i - \hat{y}_k\right)$$

2. $$C = \begin{array}{c} \text{measurements} \\ \begin{bmatrix} 100 & 105 & \boxed{20} & 32 & 28 \\ 30 & 39 & \boxed{20} & 47 & 93 \end{bmatrix} \end{array} \text{tracks}$$

wanted by both tracks.

3. Subtract row minima → reduce the rows

$$\begin{bmatrix} 100 & 105 & 20 & 32 & 28 \\ 30 & 39 & 20 & 47 & 93 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 100-20 & 105-20 & \mathbf{20}-20 & 32-20 & 28-20 \\ 30-20 & 39-20 & \mathbf{20}-20 & 47-20 & 93-20 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 80 & 85 & 0 & 12 & 8 \\ 10 & 19 & 0 & 27 & 73 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

4. Subtract column minima → reduce the columns

$$\begin{bmatrix} 80 & 85 & \mathbf{0} & 12 & 8 \\ 10 & 19 & \mathbf{0} & 27 & 73 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \rightarrow \begin{bmatrix} 80 & 85 & 0 & 12 & 8 \\ 10 & 19 & 0 & 27 & 73 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Global Nearest Neighbor (GNN)

5. Cover zeros using minimum N number of lines.

$$\begin{bmatrix} 80 & 85 & 0 & 12 & 8 \\ 10 & 19 & 0 & 27 & 73 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 80 & 85 & 0 & 12 & 8 \\ 10 & 19 & 0 & 27 & 73 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow N = 4 \text{ (i.e. 4 lines used)}$$

6. $N = 4 == 5$? No → Create more zeros by:

1) Subtracting minima of uncovered elements from uncovered elements.

$$\begin{bmatrix} 80 & 85 & 0 & 12 & \mathbf{8} \\ 10 & 19 & 0 & 27 & 73 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 80-\mathbf{8} & 85-\mathbf{8} & 0 & 12-\mathbf{8} & \mathbf{8}-\mathbf{8} \\ 10-\mathbf{8} & 19-\mathbf{8} & 0 & 27-\mathbf{8} & 73-\mathbf{8} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 72 & 77 & 0 & 4 & \mathbf{0} \\ 2 & 11 & 0 & 19 & 65 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

2) Leave once-covered elements as they are.

$$\begin{bmatrix} 72 & 77 & 0 & 4 & 0 \\ 2 & 11 & 0 & 19 & 65 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 72 & 77 & {\color{red}\mathbf{0}} & 4 & 0 \\ 2 & 11 & {\color{red}\mathbf{0}} & 19 & 65 \\ {\color{red}\mathbf{0}} & {\color{red}\mathbf{0}} & {\color{blue}\mathbf{0}} & {\color{red}\mathbf{0}} & {\color{red}\mathbf{0}} \\ {\color{red}\mathbf{0}} & {\color{red}\mathbf{0}} & {\color{blue}\mathbf{0}} & {\color{red}\mathbf{0}} & {\color{red}\mathbf{0}} \\ {\color{red}\mathbf{0}} & {\color{red}\mathbf{0}} & {\color{blue}\mathbf{0}} & {\color{red}\mathbf{0}} & {\color{red}\mathbf{0}} \end{bmatrix}$$

3) Add minima to all twice covered elements.

$$\begin{bmatrix} 72 & 77 & {\color{red}\mathbf{0}} & 4 & 0 \\ 2 & 11 & {\color{red}\mathbf{0}} & 19 & 65 \\ {\color{red}\mathbf{0}} & {\color{red}\mathbf{0}} & {\color{blue}\mathbf{0}} & {\color{red}\mathbf{0}} & {\color{red}\mathbf{0}} \\ {\color{red}\mathbf{0}} & {\color{red}\mathbf{0}} & {\color{blue}\mathbf{0}} & {\color{red}\mathbf{0}} & {\color{red}\mathbf{0}} \\ {\color{red}\mathbf{0}} & {\color{red}\mathbf{0}} & {\color{blue}\mathbf{0}} & {\color{red}\mathbf{0}} & {\color{red}\mathbf{0}} \end{bmatrix} \rightarrow \begin{bmatrix} 72 & 77 & 0 & 4 & 0 \\ 2 & 11 & 0 & 19 & 65 \\ 0 & 0 & {\color{blue}\mathbf{8}} & 0 & 0 \\ 0 & 0 & {\color{blue}\mathbf{8}} & 0 & 0 \\ 0 & 0 & {\color{blue}\mathbf{8}} & 0 & 0 \end{bmatrix}$$

Global Nearest Neighbor (GNN)

7. Cover zeros using minimum N number of lines

$$\begin{bmatrix} 72 & 77 & 0 & 4 & 0 \\ 2 & 11 & 0 & 19 & 65 \\ 0 & 0 & 8 & 0 & 0 \\ 0 & 0 & 8 & 0 & 0 \\ 0 & 0 & 8 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 72 & 77 & 0 & 4 & 0 \\ 2 & 11 & 0 & 19 & 65 \\ 0 & 0 & 8 & 0 & 0 \\ 0 & 0 & 8 & 0 & 0 \\ 0 & 0 & 8 & 0 & 0 \end{bmatrix}$$ → $N = 5$ (i.e. 5 lines used)

8. $N = 5 == 5$? Yes → Do association.

9. Assign tastk → Start with lines with only one zero.

$$\begin{bmatrix} 72 & 77 & 0 & 4 & \boxed{0} \\ 2 & 11 & 0 & 19 & 65 \\ 0 & 0 & 8 & 0 & 0 \\ 0 & 0 & 8 & 0 & 0 \\ 0 & 0 & 8 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 72 & 77 & 0 & 4 & \boxed{0} \\ 2 & 11 & \boxed{0} & 19 & 65 \\ 0 & 0 & 8 & 0 & 0 \\ 0 & 0 & 8 & 0 & 0 \\ 0 & 0 & 8 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 72 & 77 & 0 & 4 & \boxed{0} \\ 2 & 11 & \boxed{0} & 19 & 65 \\ \boxed{0} & 0 & 8 & 0 & 0 \\ 0 & \boxed{0} & 8 & 0 & 0 \\ 0 & 0 & 8 & \boxed{0} & 0 \end{bmatrix}$$

10. Calculate the corresponding cost

$$C = \begin{bmatrix} 100 & 105 & 20 & 32 & 28 \\ 30 & 39 & 20 & 47 & 93 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 100 & 105 & 20 & 32 & \boxed{28} \\ 30 & 39 & \boxed{20} & 47 & 93 \\ \boxed{0} & 0 & 0 & 0 & 0 \\ 0 & \boxed{0} & 0 & 0 & 0 \\ 0 & 0 & 0 & \boxed{0} & 0 \end{bmatrix}$$ → $20 + 28 + 0 + 0 + 0 =$ **48 (minimum)**

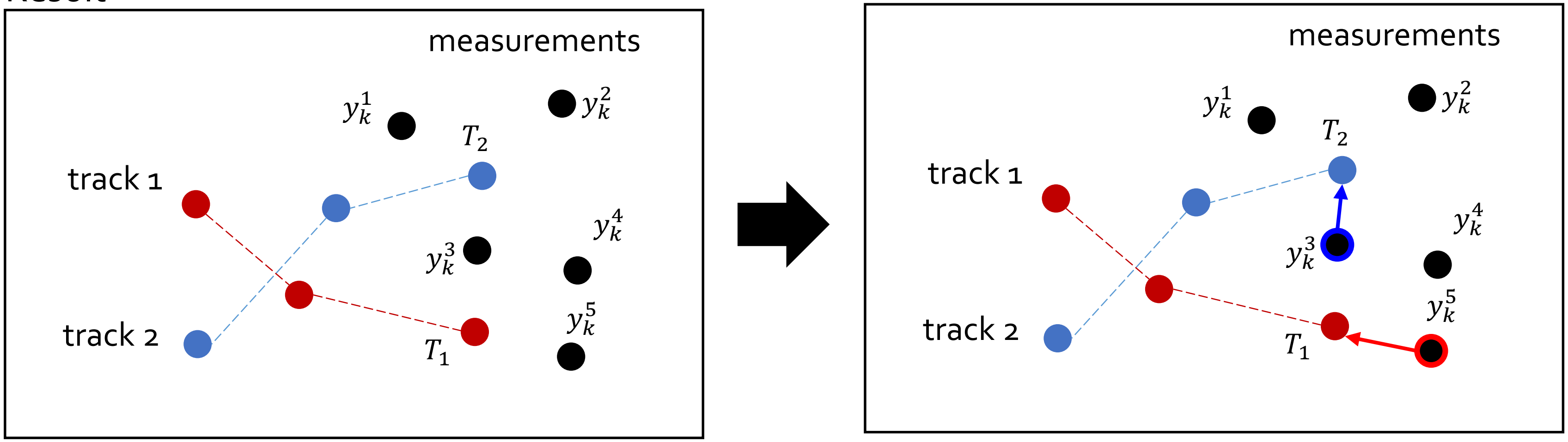
Global Nearest Neighbor (GNN)

$$C = \begin{bmatrix} 100 & 105 & 20 & 32 & 28 \\ 30 & 39 & 20 & 47 & 93 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{array}{c} \begin{array}{ccccc} \boldsymbol{y_k^1} & \boldsymbol{y_k^2} & \boldsymbol{y_k^3} & \boldsymbol{y_k^4} & \boldsymbol{y_k^5} \end{array} \\ \begin{bmatrix} 100 & 105 & 20 & 32 & \boxed{28} \\ 30 & 39 & \boxed{20} & 47 & 93 \\ \boxed{0} & 0 & 0 & 0 & 0 \\ 0 & \boxed{0} & 0 & 0 & 0 \\ 0 & 0 & 0 & \boxed{0} & 0 \end{bmatrix} \begin{array}{c} \boldsymbol{T_1} \\ \boldsymbol{T_2} \\ \\ \\ \\ \end{array} \end{array} \rightarrow \text{output: } \begin{bmatrix} 0 & 0 & 0 & 0 & \mathbf{1} \\ 0 & 0 & \mathbf{1} & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} = \text{permutation matrix}$$

→ cost : $20 + 28 + 0 + 0 + 0 =$ **48 (minimum)**

Given a cost matrix $C = \{c_{ij}\}$,
$min \sum c_{ij}x_{ij}, \quad x_{ij} \in \{1, 0\}$
$s.t. \boldsymbol{\sum_i x_{ij} = 1}, \sum_j x_{ij} = 1$

Result



Solution may not be unique → same cost, but different matching pairs.

- **State update**
  - update track 1 state estimate using y_k^5
    →a *posterior* state estimate: $\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k(y_k^5 - \hat{y}_k)$
  - update track 2 state estimate using y_k^3
    → a *posterior* state estimate: $\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k(y_k^3 - \hat{y}_k)$

Data association techniques

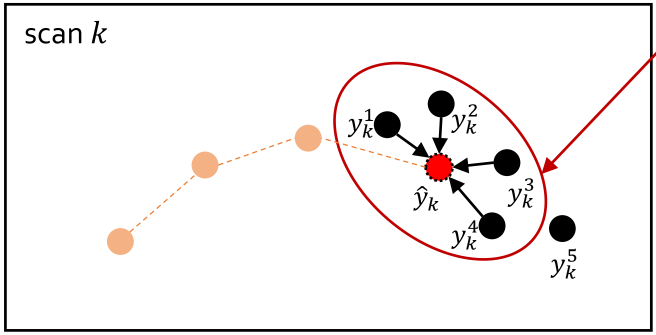
I. Non-Bayesian
   1. Nearest Neighbor Standard Filter (NNSF)
   2. Global Nearest Neighbor (GNN)

II. Bayesian
   1. **Probabilistic Data Association Filter (PDAF)**
   2. Joint Probabilistic Data Association Filter (JPDAF)
   3. Multiple Hypothesis Tracking (MHT)

Probabilistic Data Association Filter (PDAF)

- Key idea
  - Update state using ***all*** validated measurements weighted by their likelihoods.
    ① ②
  - $\hat{x}_{k|k} = \int x_k p(x_k|Y^k) dx_k$, where $Y_k = \{y_k^i\}_{i=1}^{m_k}$: a set of measurements at k^{th} scan:
    The best estimate, in MMSE sense, considers all the measurements and their probability that they are originated from the same track.
  - $\hat{x}_{k|k} = \sum_{i=0}^{m_k} E\{x_k|\chi_{k,i}, Y^k\} \boxed{P(\theta_{k,i}|Y^k)}$

$P(\theta_{k,i}|Y^k) \triangleq \beta_k^i$
= Association probability
= Probability that y_k^i is associated with track.



① Validated measurements:
→ Validation gate using Mahalanobis distance:

$$(y_k^i - \hat{y}_k)^T \hat{S}_k^{-1} (y_k^i - \hat{y}_k) < \chi_k^2$$

where,

$\hat{S}_k = H_k P_{k|k-1} H_k^T$

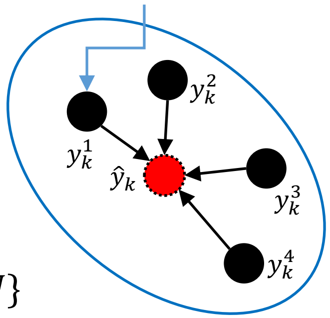
= Residual variance (uncertainty in $\hat{y}_k$)

② weighted by their likelihoods:
→ condition the update on the association events.

1) False alarm model
2) Association events

Probabilistic Data Association Filter (PDAF)

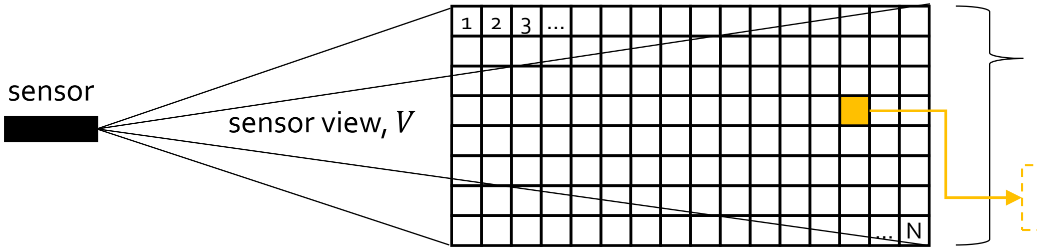
actual measurement or false alarm?



1) False alarm (false positive) due to sensor imperfections, backgrounds, etc.

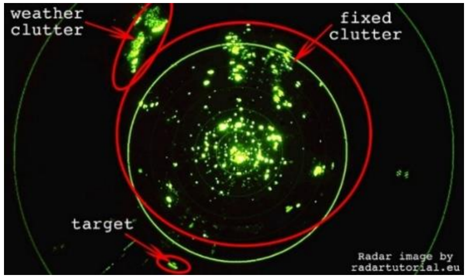
- Assume we know spatial density λ of false alarm.

  Uniformly distributed across V: spatial density $\lambda = \frac{N \cdot P_F}{V} = \frac{N_p}{V}$



N discrete cells c_i, $i \in \{1, \dots, N\}$

In each cell, false alarms occur with probability P_F

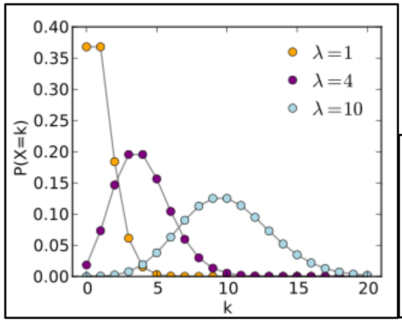


- Follows **Bernoulli process** (appear / not appear) with probability: P_F.

  → **Binomial distribution**: $P(m_F) = \binom{N}{m_F} {P_F}^{m_F} (1 - P_F)^{N - m_F}$

- Average number of false alarms in the sensor view V each scan: N_p.
- Let $N \to \infty$, then **Binomial distribution** with $P_F \ll 1$ approximates to: **Poisson distribution**(mean = variance =λV_k) → easier to compute for large N.

$$\lim_{N \to \infty} \binom{N}{m_F} {P_F}^{m_F} (1 - P_F)^{N - m_F} = \boxed{e^{-\lambda V_k} \frac{(\lambda V_k)^{m_F}}{m_F!} =: \mu_F(m_F)}$$



Poisson distribution is a model for k number of events:

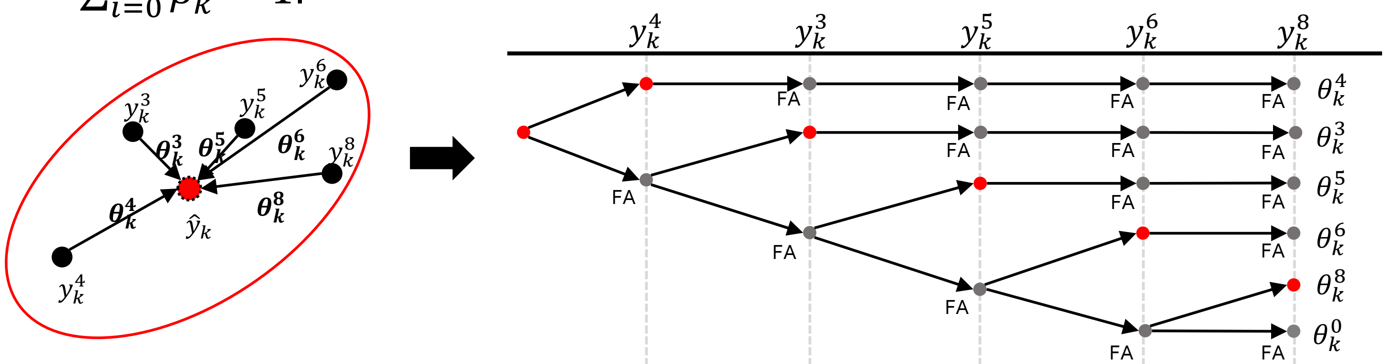
$P(k) = e^{-\lambda} \frac{\lambda^k}{k!}$

→ mean = variance = λ.

Probabilistic Data Association Filter (PDAF)

2) Association event θ_k^i

- $\theta_k^i = \begin{cases} \text{measurement } y_k^i \text{ belongs to the track.} & i = 1, \dots, m_k \\ \text{none of the measurement belongs to the track.} & i = 0 \end{cases}$
- β_k^i is the **association probability,** i.e. a posteriori probability $\mathrm{P}\left(\theta_k^i | Y^k, m_k\right) = \beta_k^i$, such that $\sum_{i=0}^{m_k} \beta_k^i = 1$.



- P_D = detection probability = the target from the track is detected.
  - e.g. "the target we are tracking has emitted a signal, and we detected it and is among the measurements inside the gate."
  - assumed to be known.
    - dependent on sensors and scenarios
    - can be tested via offline experiments.
- P_G = validation probability = probability corresponding to χ_k^2.
- **Assumptions:**
  - at most one of the validated observations is target-originated.
  - remaining measurements are false alarms.

Probabilistic Data Association Filter (PDAF)

β_k^i = Association probability given measurements up to k, Y^{k}:

$$\beta_k^i = P(\theta_k^i|Y^k) = P(\theta_k^i|Y_k, m_k, Y^{k-1}) = \frac{1}{c}\underbrace{f(Y_k|\theta_k^i, m_k, Y^{k-1})}_{1.}\ \underbrace{P(\theta_k^i|m_k, Y^{k-1})}_{2.}$$

where, $\mathrm{i} = 0, \ldots, \mathrm{m(k)}$

Baye's formula:

$$P(\theta|Y) = \frac{f(Y|\theta)P(\theta)}{\sum f(Y|\theta_i)P(\theta_i)}$$

1. $f(Y_k|\theta_k^i, m_k Y^{k-1})$= pdf of the measurements.

$$\begin{cases} \frac{1}{V_k}^{m_{k-1}} \frac{\mathcal{N}\left(z_k^i; 0, \hat{S}_k\right)}{\sum_{j=1}^{m_k} \mathcal{N}\left(z_k^j; 0, \hat{S}_k\right)}, i = 1, \ldots, m_k \\ \frac{1}{V_k}^{m_k} \qquad , i = 0 \end{cases}$$

pdf of incorrect measurements assumed *uniform* in th e validation region V_k. i.e. $\frac{1}{V_k}$

2. $P(\theta_k^i|m_k, Y^{k-1})$= prior probability of the association events, conditioned on the number of measurements.

$$= \begin{cases} P(\theta_k^i|when\ 1\ measurement\ is\ to\ be\ associated, m_k - 1\ are\ FA) \times \mathrm{P}(m_k - 1\ are\ \mathrm{FA}), i = 1, \ldots, m_k \\ P\left(\theta_k^i \middle| when\ 0\ measurement\ is\ to\ be\ associated, m_k\ are\ FA\right) \times \mathrm{P}(m_k\ are\ \mathrm{FA}) \qquad , i = 0 \end{cases}$$

$$= \begin{cases} P\left(\theta_k^i \middle| FA = m_k - 1\right)\boxed{P(FA = m_k - 1|Y_k)}, i = 1, \ldots, m_k \\ P\left(\theta_k^i \middle| FA = m_k\right)\boxed{P(FA = m_k|Y_k)} \qquad , i = 0 \end{cases}$$

$$where \begin{cases} \boxed{\mathrm{P}(FA = m_k - 1|Y_k)} = \frac{P(Y_k|FA = m_k - 1)P(FA = m_k - 1)}{P(Y_k)} = \frac{P_D P_G \times \mu_F(m_k - 1)}{P(Y_k)}, i = 1, \ldots, m_k \\ \boxed{\mathrm{P}(FA = m_k|Y_k)} = \frac{P(Y_k|FA = m_k)(FA = m_k)}{P(Y_k)} = \frac{(1 - P_D P_G) \times \mu_F(m_k)}{P(Y_k)} \qquad , i = 0 \end{cases}$$

$$where, P(Y_k) = P_D P_G \times \mu_F(m_k - 1) + (1 - P_D P_G) \times \mu_F(m_k)$$

$$= \begin{cases} \frac{1}{m_k} P_D P_G \times \mu_F(m_k - 1) P(Y_k)^{-1}, i = 1, \ldots, m_k \\ (1 - P_D P_G) \times \mu_F(m_k) P(Y_k)^{-1} \quad , i = 0 \end{cases}$$

Probabilistic Data Association Filter (PDAF)

$$\beta_k^i = \begin{cases} \frac{1}{c_k}\left[\boxed{\frac{1}{V_k}^{m_{k-1}} \frac{\mathcal{N}\left(z_k^i;0,\hat{S}_k\right)}{\sum_{j=1}^{m_k}\mathcal{N}\left(z_k^j;0,\hat{S}_k\right)}}\boxed{\frac{P_D P_G}{m_k}\mu_F(m_k-1)}\right] & ,i=1,\dots,m_k \\ \frac{1}{c_k}\left[\boxed{\frac{1}{V_k}^{m_k}}\boxed{(1-P_D P_G)\mu_F(m_k)}\right] & ,i=0 \end{cases}$$

where, $c_k = \frac{1}{V_k}^{m_k}(1-P_D P_G)\mu_F(m_k) + \sum_{i=1}^{m_k}\frac{1}{V_k}^{m_{k-1}}\frac{\mathcal{N}\left(z_k^i;0,\hat{S}_k\right)}{\sum_{j=1}^{m_k}\mathcal{N}\left(z_k^j;0,\hat{S}_k\right)}\frac{P_D P_G}{m_k}\mu_F(m_k-1)$

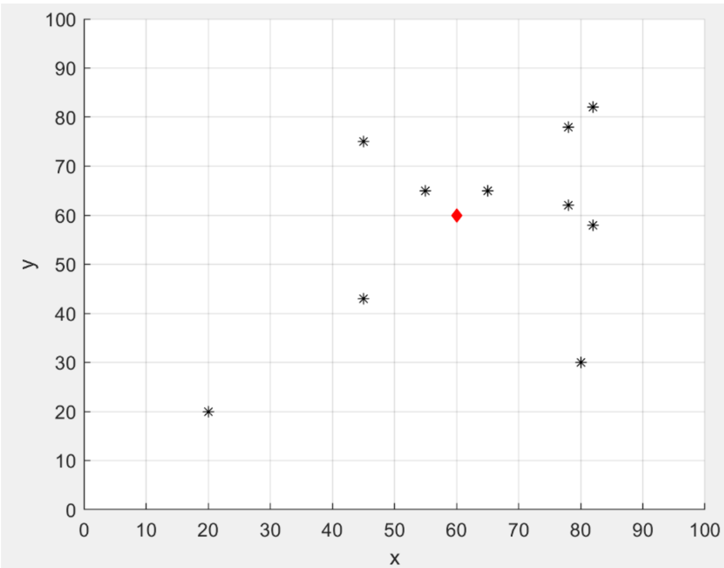
z_k^i is from the track:

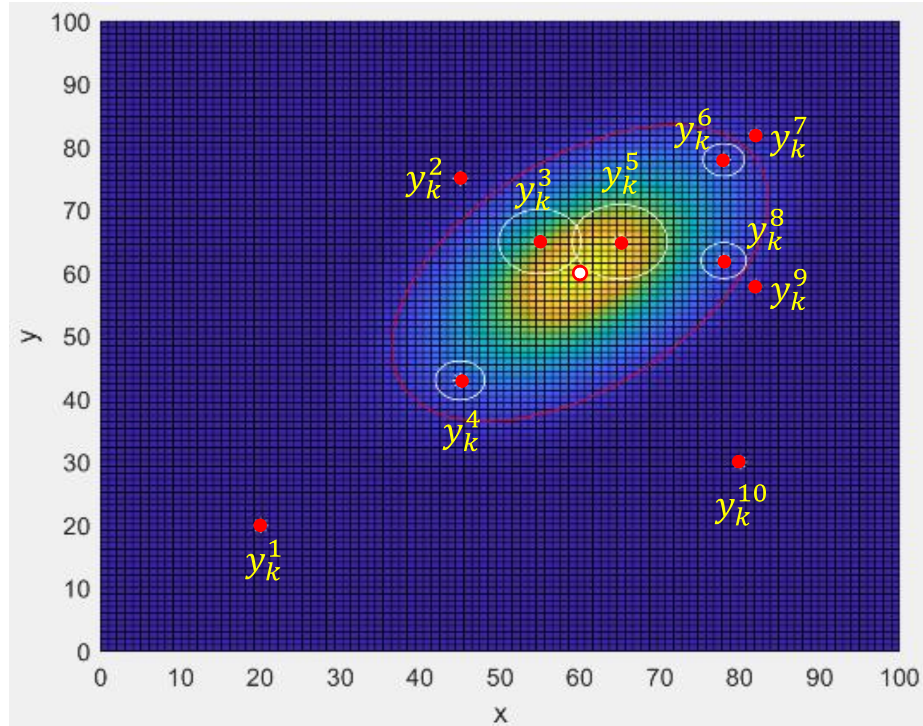
- pdf of 1 measurement and $m_k - 1$ false alarms.
- Probability of 1 measurement being associated to the track x probability of $m_k - 1$ false alarms.

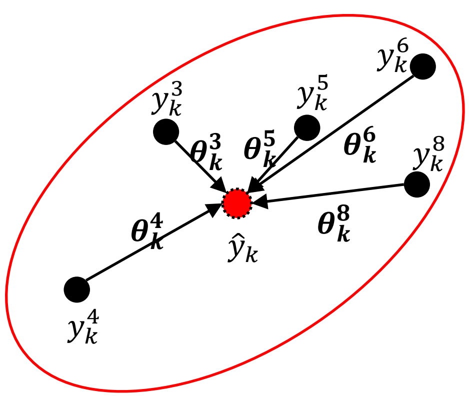
None of the measurements are from the track:

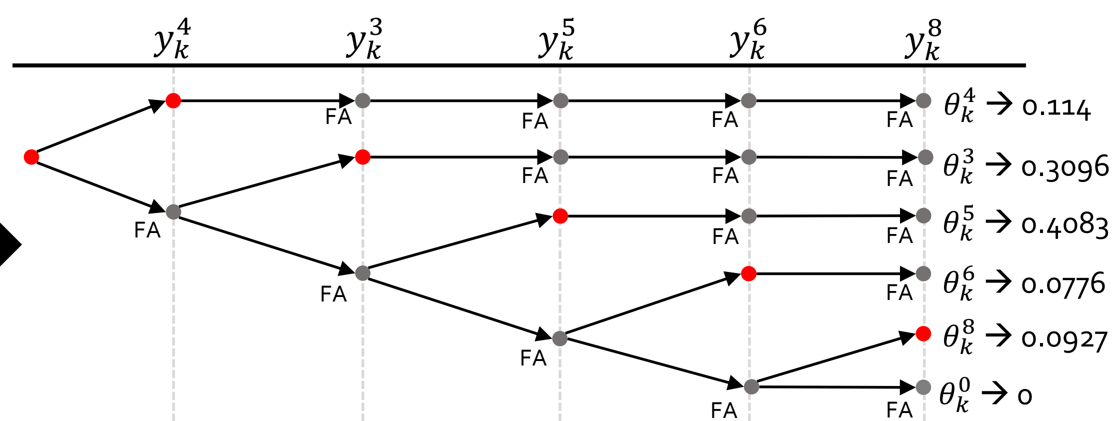
- pdf of m_k false alarms.
- Probability that no measurements being associated to the track x probability of m_k false alarms.

	y_k^0	y_k^1	y_k^2	y_k^3	y_k^4	y_k^5	y_k^6	y_k^7	y_k^8	y_k^9	y_k^{10}
$\boldsymbol{\beta_k^i}$	0	0	0	0.3096	0.1114	0.4083	0.0776	0	0.0927	0	0









Probabilistic Data Association Filter (PDAF)

- Key idea
  - Update state using all validated measurements weighted by their likelihoods.
    ① ②

- Use weighted combination of the measurements:
  - posterior state estimate:
    $\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k \boxed{\sum_{i=1}^{m_k} \beta_k^i}(y_k - \hat{y}_k)$
  - posterior state uncertainty covariance:

- K_k = Kalman gain
- H_k = state-to-measurement mapping matrix
- $\hat{S}_k = H_k P_{k|k-1} {H_k}^T$ = Residual variance
- $\boldsymbol{Z}_k = \sum_{i=1}^{m_k} \beta_k^i z_k = \sum_{i=1}^{m_k} \beta_k^i (y_k - \hat{y}_k)$ = innovation mtx

$$P_{k|k} = \boxed{\beta_k^0}\boxed{P_{k|k-1}} + \boxed{(1-\beta_k^0)}\boxed{[(I - K_k H_k) P_{k|k-1} (I - K_k H_k)^T + K_k \hat{S}_k {K_k}^T]} + \boxed{K_k \left[\sum_{i=1}^{m_k} (\beta_k^i z_k^i {z_k^i}^T) - \boldsymbol{Z}_k {\boldsymbol{Z}_k}^T\right] {K_k}^T}$$

if $\beta_k^0 = 1$, i.e. 100% sure that **none** of the measurements belong to the track
→ covariance of state does not get updated.

$= P_{k|k}$: cov. of the update if only 1 measurement is used.

Effects of incorrect measurement assignments
→ increases covariance of the update. ($\succcurlyeq 0$, PSD)

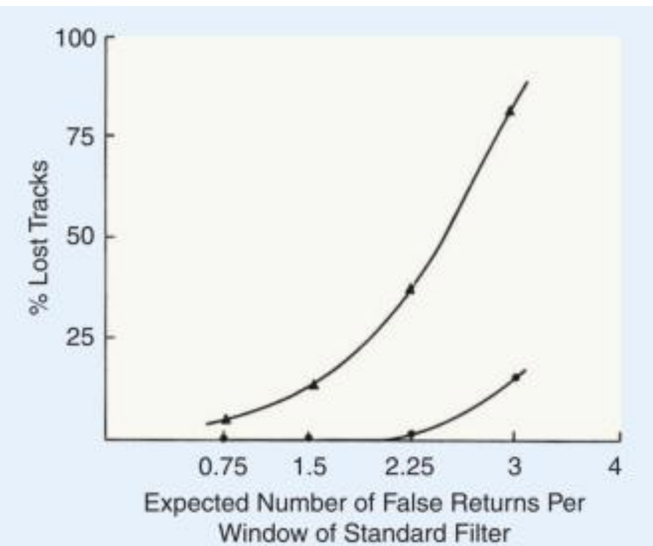
c.f.) When using a single best measurement (e.g. GNN):

- $\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k (y_k - \hat{y}_k)$
- $\boxed{P_{k|k} = (I - K_k H_k) P_{k|k-1} (I - K_k H_k)^T + K_k \hat{S}_k {K_k}^T}$ ← compare
  or
  $P_{k|k} = (I - K_k H_k) P_{k|k-1}$ with optimal Kalman gain for MMSE estimates, $K_k = P_{k|k-1} {H_k}^T {S_k}^{-1}$

- Although state estimation is linear, the filter is nonlinear because in state estimation update the association probability, β_k^i, is nonlinearly defined.

Probabilistic Data Association Filter (PDAF)

<table>
<tr><th></th><th>Nearest Neighbor Standard Filter (NNSF)</th><th>Probabilistic Data Association Filter (PDAF)</th></tr>
<tr><td>state prediction $\boldsymbol{\hat{x}_{k|k-1}}$</td><td colspan="2">$\hat{x}_{k|k-1} = F_{k-1}\hat{x}_{k-1|k-1}$</td></tr>
<tr><td>state cov. prediction $\boldsymbol{P_{k|k-1}}$</td><td colspan="2">$P_{k|k-1} = F_{k-1}P_{k-1|k-1}{F_{k-1}}^T + Q_{k-1}$</td></tr>
<tr><td>innovation covariance $\boldsymbol{\hat{S}_k}$</td><td colspan="2">$\hat{S}_k = H_k P_{k|k-1}{H_k}^T$</td></tr>
<tr><td>Kalman gain $\boldsymbol{K_k}$ (optimal)</td><td colspan="2">$K_k = P_{k|k-1}{H_k}^T{\hat{S}_k}^{-1}$</td></tr>
<tr><td>innovation $\boldsymbol{z_k}$</td><td>$z_k = y_k - H_k\hat{x}_{k|k-1}$</td><td>$z_k = \sum_{i=1}^{m_k}\beta_k^i\,(y_k^i - H_k\hat{x}_{k|k-1})$</td></tr>
<tr><td>state cov. update $\boldsymbol{P_{k|k}}$</td><td>$P_{k|k} = (I - K_k H_k)P_{k|k-1}$</td><td>$P_{k|k} = \beta_k^0 P_{k|k-1} + \left(1-\beta_k^0\right)P_{k|k} + K_k\left[\sum_{i=1}^{m_k}(\beta_k^i\, z_k^i {z_k^i}^T) - \boldsymbol{Z}_k{\boldsymbol{Z}_k}^T\right]{K_k}^T$</td></tr>
</table>



- **Computational requirement:**
  PDAF $\approx$ 2 $\times$ Standard KF
- **Robustness against false alarm in validation gate:**
  PDAF $\gg$ NNSF
- **Suboptimal strategy**
  never totally wrong, but never totally right.
- **Multiple tracks?** Joint PDAF (JPDAF)

Data association techniques

I. Non-Bayesian

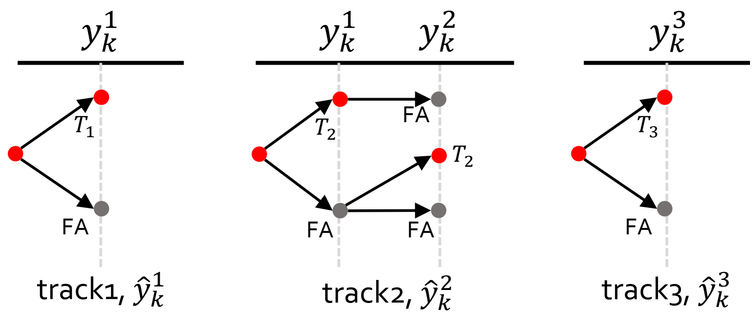
1. Nearest Neighbor Standard Filter (NNSF)
2. Global Nearest Neighbor (GNN)

II. Bayesian

1. Probabilistic Data Association Filter (PDAF)
2. **Joint Probabilistic Data Association Filter (JPDAF)**
3. Multiple Hypothesis Tracking (MHT)

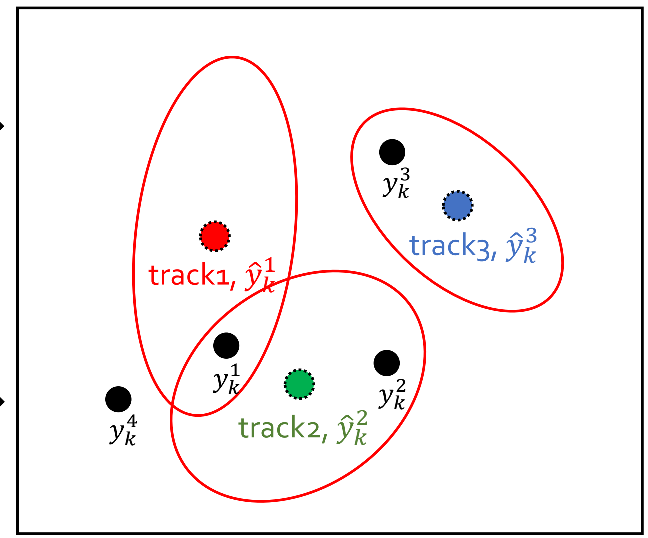
Joint Probabilistic Data Association Filter (JPDAF)

- Differs in the definition of association events and association probability.
- **PDAF**: disjoint trees for each track

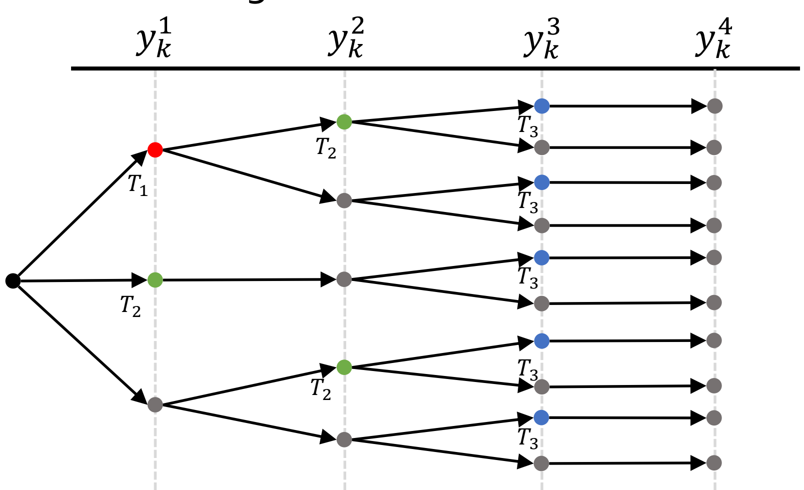


Association events θ_k^i decoupled.

$$\theta_k^i = \begin{cases} y_k^i \text{ belongs to the track.} & i = \{1, \dots, m_k\} \\ \text{none belongs to the track.} & i = 0 \end{cases}$$



- **JPDAF**: single tree for all measurements



Association events $\theta_k^{T_j^i}$ coupled.

$$\theta_k = \cap_{i=1}^{m_k} \theta_k^{T_j^i}, j = 1, \dots, N_T$$

Joint Probabilistic Data Association Filter (JPDAF)

- Probability of joint association event θ_k
- $P(\theta_k|Y^k) = P(\theta_k|Y_k, m_k, Y^{k-1}) = \frac{1}{c}\underbrace{f(Y_k|\theta_k, m_k, Y^{k-1})}_{1.}\underbrace{P(\theta_k|m_k)}_{2.}$

Baye's formula:
$P(\theta|Y) = \frac{f(Y|\theta)P(\theta)}{\sum f(Y|\theta_i)P(\theta_i)}$

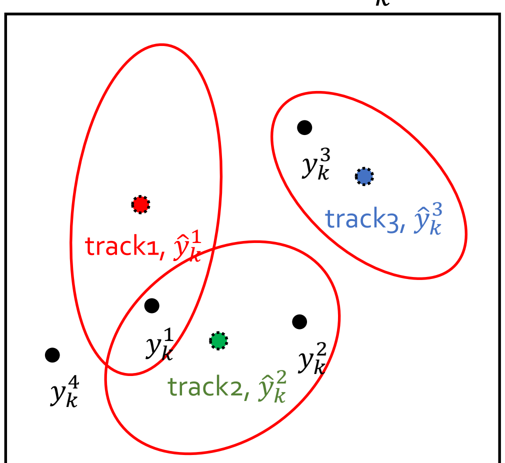
1. $f(Y_k|\theta_k, m_k, Y^{k-1})$= conditional pdf of measurements
   - $V^{-m_k^F}\prod_{T\in\mathcal{T}_D}\mathcal{N}\left(z_k^{j_T}; 0, \hat{S}_k\right)$,
     where m_k^F= number of false alarms, V= scan area
2. $P(\theta_k|m_k)$= prior probability of a joint association event θ_k
   - $\frac{\prod_{T\in\mathcal{T}_D} P_D^T \prod_{T\in\mathcal{T}_{ND}}(1-P_D^T)\mu_F(m_k^F)}{\boxed{{}_{m_k}\mathrm{P}_{m_k-m_k^F}}} = \boxed{\frac{m_k^F!}{m_k!}}\prod_{T\in\mathcal{T}_D} P_D^T \prod_{T\in\mathcal{T}_{ND}}(1-P_D^T)\,\mu_F(m_k^F)$

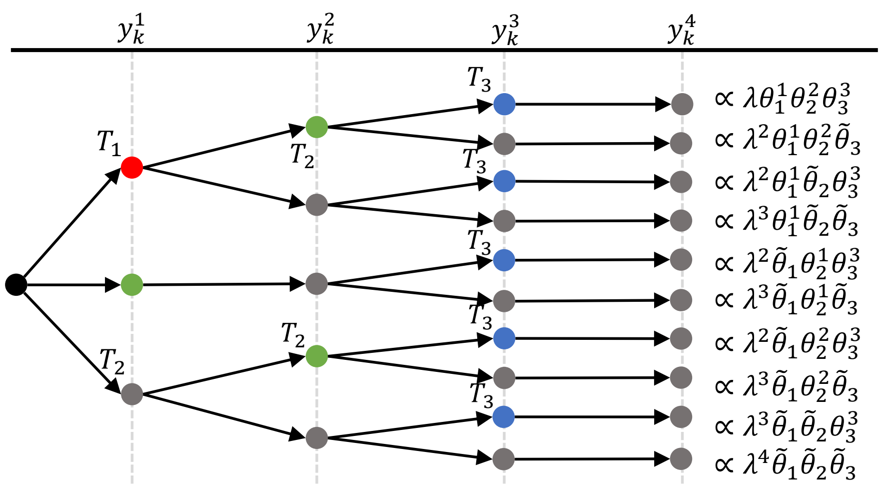
Number of combinations where non-false-alarm $(m_k - m_k^F)$ measurements
out of total m_k measurements assigned to existing tracks, order matters.
$\rightarrow {}_{m_k}\mathrm{P}_{m_k-m_k^F} = \frac{m_k!}{(m_k-(m_k-m_k^F))!} = \frac{m_k!}{m_k^F!}$

- c = normalizing factor
- P_D^T = track-specific detection probability of track T
- P_G^T = known gate probability of track T
- $\mathcal{N}\left(z_k^{j_T}; 0, \hat{S}_k\right)$ = $j_T{}^{th}$ measurement likelihood given track T

- Result:

$$P(\theta_k|Y^k) = \frac{1}{c}\frac{m_k^F!}{m_k!}\mu_F(m_k^F)V^{-m_k^F}\prod_{T\in\mathcal{T}_D} P_D^T\,\mathcal{N}\left(z_k^{j_T}; 0, \hat{S}_k\right)\prod_{T\in\mathcal{T}_{ND}}(1-P_D^T)$$

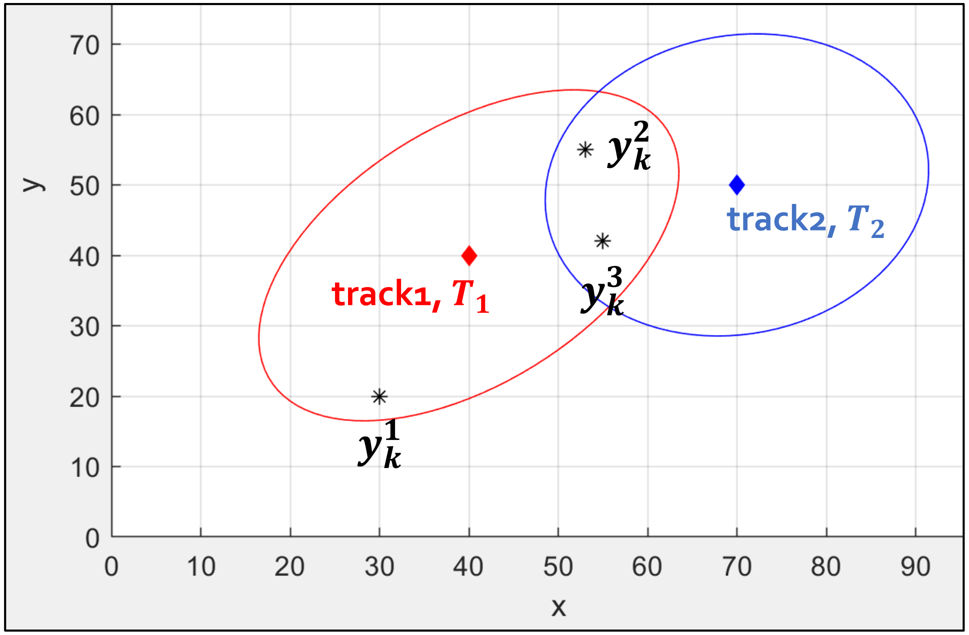


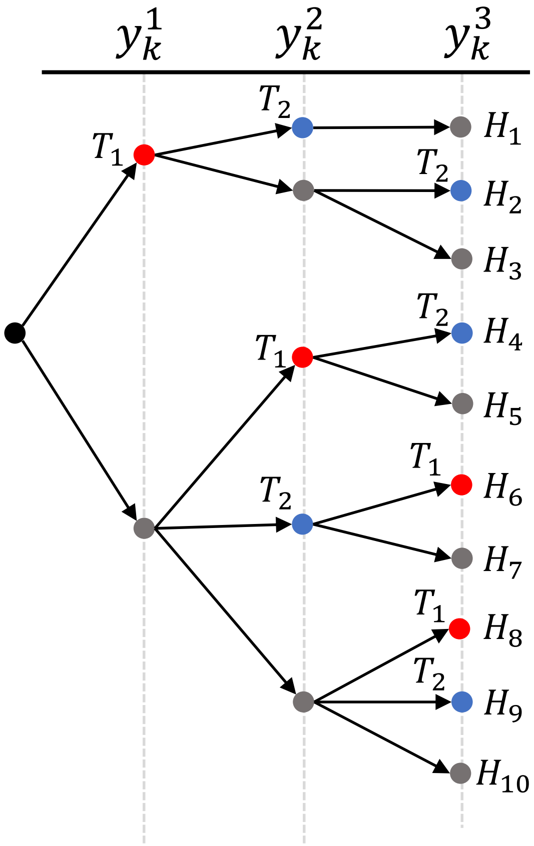


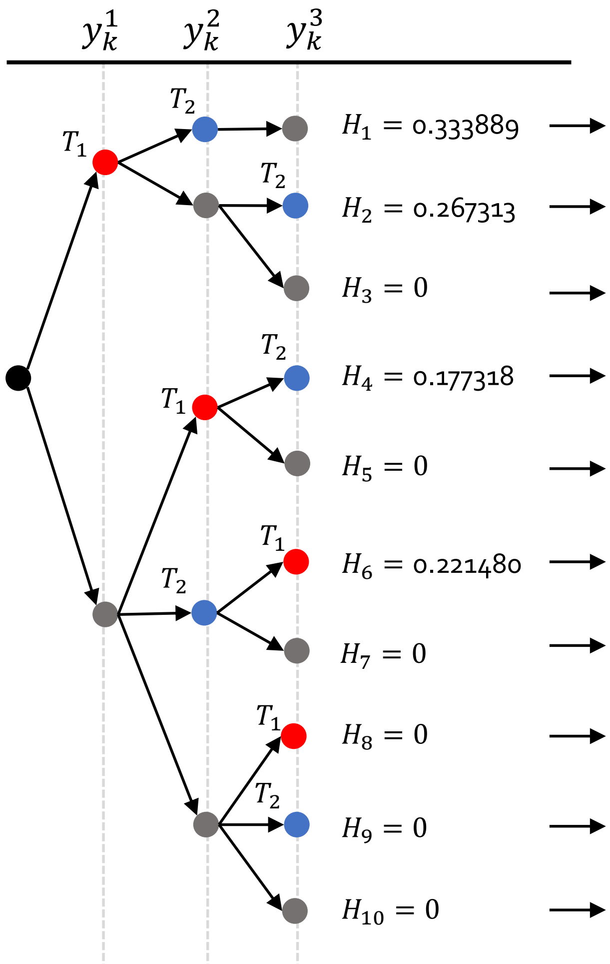
Joint Probabilistic Data Association Filter (JPDAF)

$$P(\theta_k|Y^k) = \frac{1}{c}\frac{m_k^F!}{m_k!}\mu_F(m_k^F)V^{-m_k^F}\prod_{T\in\mathcal{T}_D} P_D^T\,\mathcal{N}\left(z_k^{j_T};0,\hat{S}_k\right)\prod_{T\in\mathcal{T}_{ND}}(1-P_D^T)$$

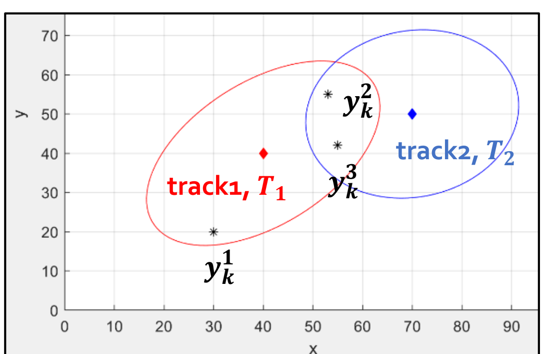
- Example

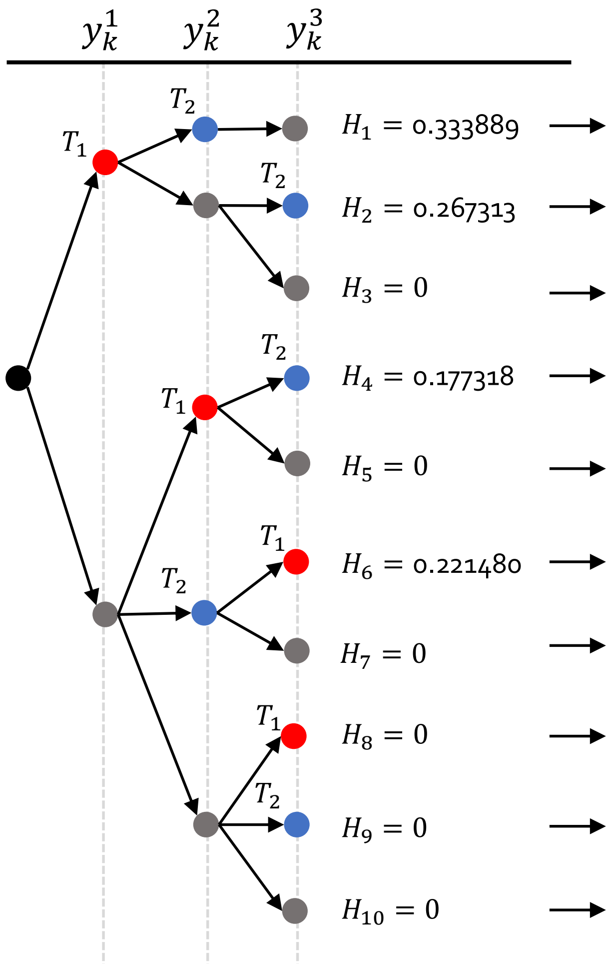




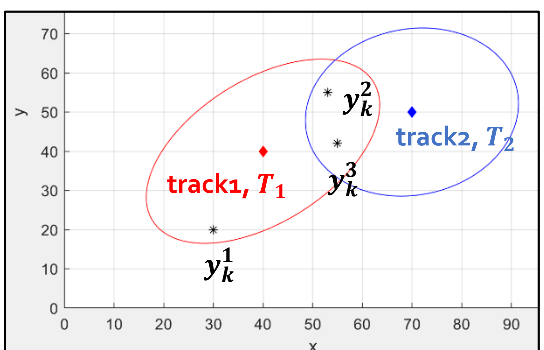


H_i	T_1	T_2	p{H_i}	p{H_i} normalized	
H_1	1	2	$\lambda\theta_1^1\theta_2^2$	$\mu_F(1)V^{-1}\mathcal{N}_1^1\mathcal{N}_2^2P_{D,1}P_{D,2}$	0.333889
H_2	1	3	$\lambda\theta_1^1\theta_2^3$	$\mu_F(1)V^{-1}\mathcal{N}_1^1\mathcal{N}_2^3P_{D,1}P_{D,2}$	0.267313
H_3	1	-	$\lambda^2\theta_1^1\tilde{\theta}_2$	$\mu_F(2)V^{-2}\mathcal{N}_1^1P_{D,1}(1\text{-}P_{D,2})$	0.0
H_4	2	3	$\lambda\theta_1^2\theta_2^3$	$\mu_F(1)V^{-1}\mathcal{N}_1^2\mathcal{N}_2^3P_{D,1}P_{D,2}$	0.177318
H_5	2	-	$\lambda^2\theta_1^2\tilde{\theta}_2$	$\mu_F(2)V^{-2}\mathcal{N}_1^2P_{D,1}(1\text{-}P_{D,2})$	0.0
H_6	3	2	$\lambda\theta_1^3\theta_2^2$	$\mu_F(1)V^{-1}\mathcal{N}_1^3\mathcal{N}_2^2P_{D,1}P_{D,2}$	0.22148
H_7	-	2	$\lambda\tilde{\theta}_1\theta_2^2$	$\mu_F(1)V^{-1}\mathcal{N}_2^2(1-P_{D,1})P_{D,2}$	0.0
H_8	3	-	$\lambda^2\theta_1^3\tilde{\theta}_2$	$\mu_F(2)V^{-2}\mathcal{N}_1^3P_{D,1}(1\text{-}P_{D,2})$	0.0
H_9	-	3	$\lambda^2\tilde{\theta}_1\theta_2^3$	$\mu_F(2)V^{-2}\mathcal{N}_2^3(1-P_{D,1})P_{D,2}$	0.0
H_{10}	-	-	$\lambda^3\tilde{\theta}_1\tilde{\theta}_2$	$\mu_F(3)V^{-3}(1\text{-}P_{D,1})(1\text{-}P_{D,2})$	0.0





H_i	T_1	T_2	$p\{H_i\}$	$p\{H_i\}$ normalized	
H_1	1	2	$\lambda\theta_1^1\theta_2^2$	$\mu_F(1)V^{-1}\mathcal{N}_1^1\mathcal{N}_2^2 P_{D,1}P_{D,2}$	0.333889
H_2	1	3	$\lambda\theta_1^1\theta_2^3$	$\mu_F(1)V^{-1}\mathcal{N}_1^1\mathcal{N}_2^3 P_{D,1}P_{D,2}$	0.267313
H_3	1	-	$\lambda^2\theta_1^1\tilde{\theta}_2$	$\mu_F(2)V^{-2}\mathcal{N}_1^1 P_{D,1}(1\text{-}P_{D,2})$	0.0
H_4	2	3	$\lambda\theta_1^2\theta_2^3$	$\mu_F(1)V^{-1}\mathcal{N}_1^2\mathcal{N}_2^3 P_{D,1}P_{D,2}$	0.177318
H_5	2	-	$\lambda^2\theta_1^2\tilde{\theta}_2$	$\mu_F(2)V^{-2}\mathcal{N}_1^2 P_{D,1}(1\text{-}P_{D,2})$	0.0
H_6	3	2	$\lambda\theta_1^3\theta_2^2$	$\mu_F(1)V^{-1}\mathcal{N}_1^3\mathcal{N}_2^2 P_{D,1}P_{D,2}$	0.22148
H_7	-	2	$\lambda\tilde{\theta}_1\theta_2^2$	$\mu_F(1)V^{-1}\mathcal{N}_2^2(1-P_{D,1})P_{D,2}$	0.0
H_8	3	-	$\lambda^2\theta_1^3\tilde{\theta}_2$	$\mu_F(2)V^{-2}\mathcal{N}_1^3 P_{D,1}(1\text{-}P_{D,2})$	0.0
H_9	-	3	$\lambda^2\tilde{\theta}_1\theta_2^3$	$\mu_F(2)V^{-2}\mathcal{N}_2^3(1-P_{D,1})P_{D,2}$	0.0
H_{10}	-	-	$\lambda^3\tilde{\theta}_1\tilde{\theta}_2$	$\mu_F(3)V^{-3}(1\text{-}P_{D,1})(1\text{-}P_{D,2})$	0.0



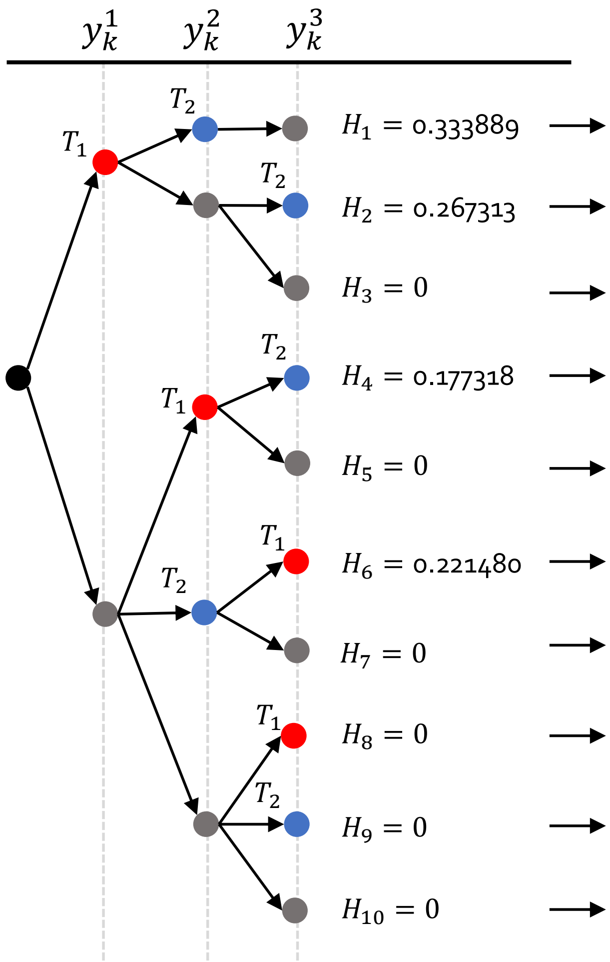
Track 1

P_1^0 = *track 1 not assigned* $= 0 + 0 + 0 = 0$

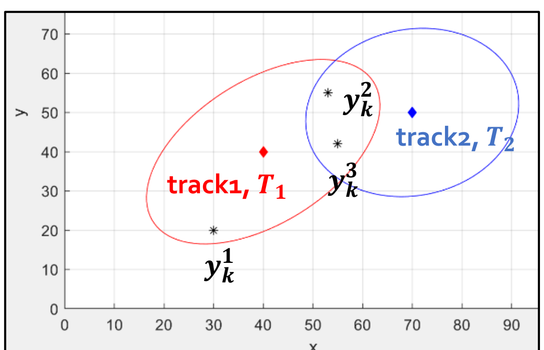
P_1^1 = *track 1 to* $y_k^1 = 0.333889 + 0.267313 + 0 = \mathbf{0.6012}$

P_1^2 = *track 1to* $y_k^2 = 0.177318 + 0 = \mathbf{0.177318}$

P_1^3 = *track 1to* $y_k^3 = 0.22148 + 0 = \mathbf{0.22148}$



H_i	T_1	T_2	p{H_i}	p{H_i} normalized	
H_1	1	2	$\lambda\theta_1^1\theta_2^2$	$\mu_F(1)V^{-1}\mathcal{N}_1^1\mathcal{N}_2^2 P_{D,1}P_{D,2}$	0.333889
H_2	1	3	$\lambda\theta_1^1\theta_2^3$	$\mu_F(1)V^{-1}\mathcal{N}_1^1\mathcal{N}_2^3 P_{D,1}P_{D,2}$	0.267313
H_3	1	-	$\lambda^2\theta_1^1\tilde{\theta}_2$	$\mu_F(2)V^{-2}\mathcal{N}_1^1 P_{D,1}(1\text{-}P_{D,2})$	0.0
H_4	2	3	$\lambda\theta_1^2\theta_2^3$	$\mu_F(1)V^{-1}\mathcal{N}_1^2\mathcal{N}_2^3 P_{D,1}P_{D,2}$	0.177318
H_5	2	-	$\lambda^2\theta_1^2\tilde{\theta}_2$	$\mu_F(2)V^{-2}\mathcal{N}_1^2 P_{D,1}(1\text{-}P_{D,2})$	0.0
H_6	3	2	$\lambda\theta_1^3\theta_2^2$	$\mu_F(1)V^{-1}\mathcal{N}_1^3\mathcal{N}_2^2 P_{D,1}P_{D,2}$	0.22148
H_7	-	2	$\lambda\tilde{\theta}_1\theta_2^2$	$\mu_F(1)V^{-1}\mathcal{N}_2^2(1-P_{D,1})P_{D,2}$	0.0
H_8	3	-	$\lambda^2\theta_1^3\tilde{\theta}_2$	$\mu_F(2)V^{-2}\mathcal{N}_1^3 P_{D,1}(1\text{-}P_{D,2})$	0.0
H_9	-	3	$\lambda^2\tilde{\theta}_1\theta_2^3$	$\mu_F(2)V^{-2}\mathcal{N}_2^3(1-P_{D,1})P_{D,2}$	0.0
H_{10}	-	-	$\lambda^3\tilde{\theta}_1\tilde{\theta}_2$	$\mu_F(3)V^{-3}(1\text{-}P_{D,1})(1\text{-}P_{D,2})$	0.0



Track 2

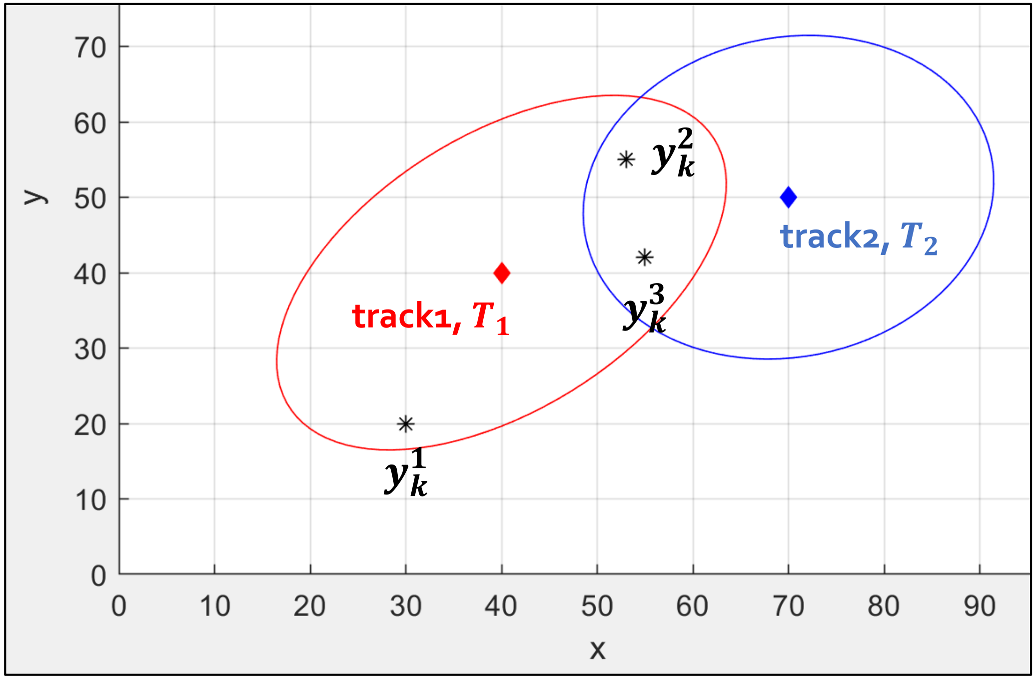
$P_2^0 = track\ 2\ not\ assigned = 0 + 0 + 0 + 0 = \mathbf{0}$

$\cancel{P_2^1 = track\ 2\ to\ y_k^1 = 0}$ ← outside validation region

$P_2^2 = track\ 2 to\ y_k^2 = 0.333889 + 0.22148 + 0 = \mathbf{0.5554}$

$P_2^3 = track\ 2\ to\ y_k^3 = 0.22148 + 0 = \mathbf{0.4446}$

Joint Probabilistic Data Association Filter (JPDAF)



$P_2^0 = 0$
$P_2^1 = 0.6012$
$P_2^2 = 0.177318$
$P_2^3 = 0.22148$

PDAF for track 1

$P_2^0 = 0$
$P_2^1 = 0$
$P_2^2 = 0.5554$
$P_2^3 = 0.4446$

PDAF for track 2

PDAF for each track.

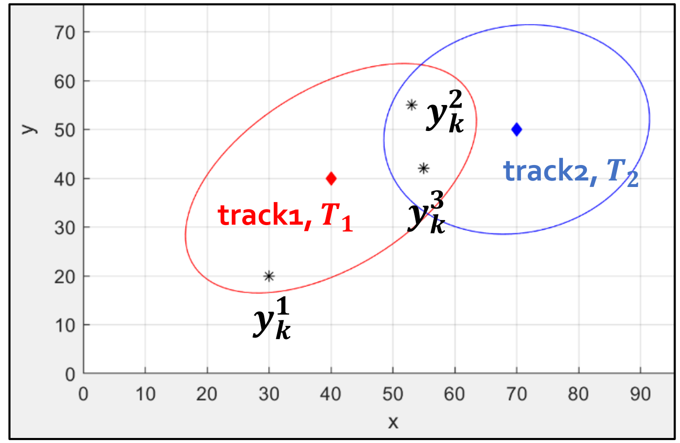
$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k \sum_{i=1}^{m_k} \beta_k^i \left(y_k - \hat{y}_k\right)$$

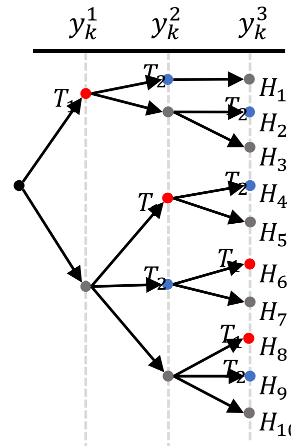
$$P_{k|k} = \beta_k^0 P_{k|k-1} + \left(1-\beta_k^0\right)\left[(I - K_k H_k) P_{k|k-1} (I - K_k H_k)^T + K_k \hat{S}_k {K_k}^T\right] + K_k \left[\sum_{i=1}^{m_k} \left(\beta_k^i\, z_k^i {z_k^i}^T\right) - \boldsymbol{Z}_k {\boldsymbol{Z}_k}^T\right] {K_k}^T$$

- **Implementation perspective**

1. Build Validation Gate matrix $\overline{\Omega}$.
   $(y_k^i - \hat{y}_k)^T \hat{S}_k^{-1} (y_k^i - \hat{y}_k) < \chi_k^2$

$$\rightarrow \overline{\Omega} = \begin{array}{c c} & \begin{array}{ccc} FA & T_1 & T_2 \end{array} \\ & \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{array}{c} y_k^1 \\ y_k^2 \\ y_k^3 \end{array} \end{array}$$





2. Expand to a set of event matrices Ω_X.

$$\begin{array}{cc} & \begin{array}{ccc} FA & T_1 & T_2 \end{array} \\ \Omega_1 = & \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{array}{c} y_k^1 \\ y_k^2 \\ y_k^3 \end{array} \end{array} \quad \begin{array}{cc} & \begin{array}{ccc} FA & T_1 & T_2 \end{array} \\ \Omega_6 = & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{array}{c} y_k^1 \\ y_k^2 \\ y_k^3 \end{array} \end{array}$$

$$\Omega_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{array}{c} y_k^1 \\ y_k^2 \\ y_k^3 \end{array} \quad \Omega_7 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{array}{c} y_k^1 \\ y_k^2 \\ y_k^3 \end{array}$$

$$\Omega_3 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{array}{c} y_k^1 \\ y_k^2 \\ y_k^3 \end{array} \quad \Omega_8 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{array}{c} y_k^1 \\ y_k^2 \\ y_k^3 \end{array}$$

$$\Omega_4 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{array}{c} y_k^1 \\ y_k^2 \\ y_k^3 \end{array} \quad \Omega_9 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{array}{c} y_k^1 \\ y_k^2 \\ y_k^3 \end{array}$$

$$\Omega_5 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{array}{c} y_k^1 \\ y_k^2 \\ y_k^3 \end{array} \quad \Omega_{10} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{array}{c} y_k^1 \\ y_k^2 \\ y_k^3 \end{array}$$

3. For each event matrix, compute joint probability using
   $P(\theta_k | Y^k) = \frac{1}{c} \frac{m_k^F!}{m_k!} \mu_F(m_k^F) V^{-m_k^F} \prod_{T \in \mathcal{T}_D} P_D^T \, \mathcal{N}\left(z_k^{j_T}; 0, \hat{S}_k\right) \prod_{T \in \mathcal{T}_{ND}} (1 - P_D^T)$

   $\rightarrow$ Results: $\mathrm{P}(\Omega_X) =$

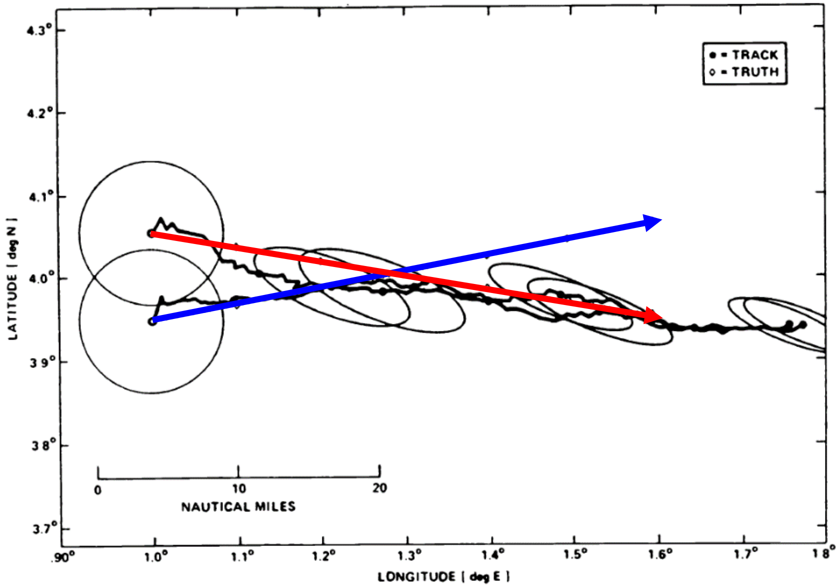
   $[0.333889, 0.267313, 0, 0.177318, 0, 0.22148, 0, 0, 0, 0]$

4. Update state estimation for each track as in PDAF.
   $\beta_{jT}(k) = P\{\theta_{jT}(k) | Y^k\} = \sum_{\theta : \theta_{jT} \in \theta} P\{\theta(k) | Y^k\}$

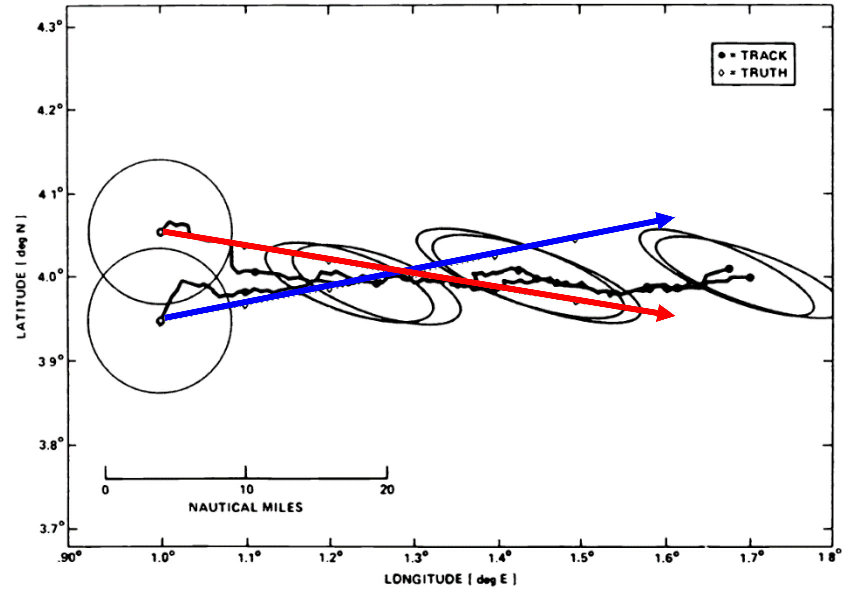
   - $\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k \sum_{i=1}^{m_k} \beta_k^i \, (y_k - \hat{y}_k)$
   - $P_{k|k} = \beta_k^0 P_{k|k-1}$
     $+ (1 - \beta_k^0)\left[(I - K_k H_k) P_{k|k-1} (I - K_k H_k)^T + K_k \hat{S}_k {K_k}^T\right] + K_k \left[\sum_{i=1}^{m_k} (\beta_k^i \, z_k^i {z_k^i}^T) - \boldsymbol{Z}_k {\boldsymbol{Z}_k}^T\right] {K_k}^T$

- **Comparisons**

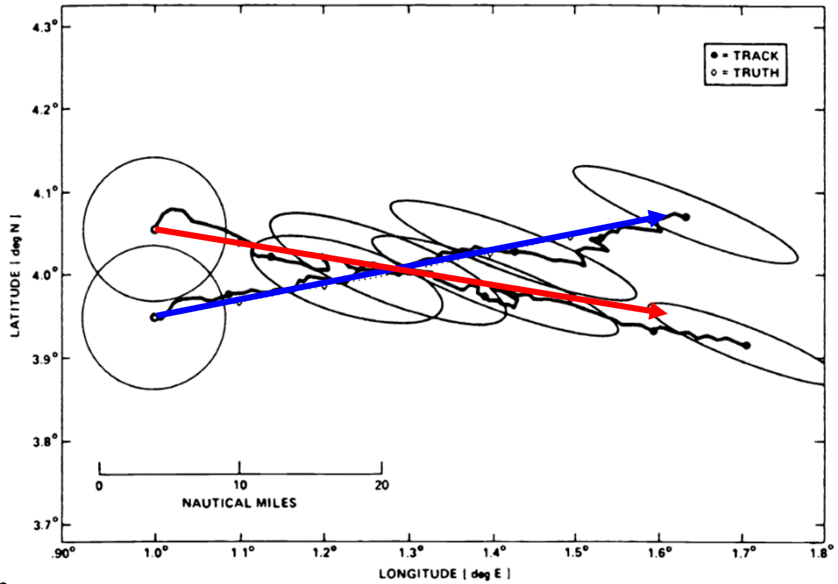
Nearest Neighbor Standard Filter (NNSF)



Probabilistic Data Association Filter (PDAF)



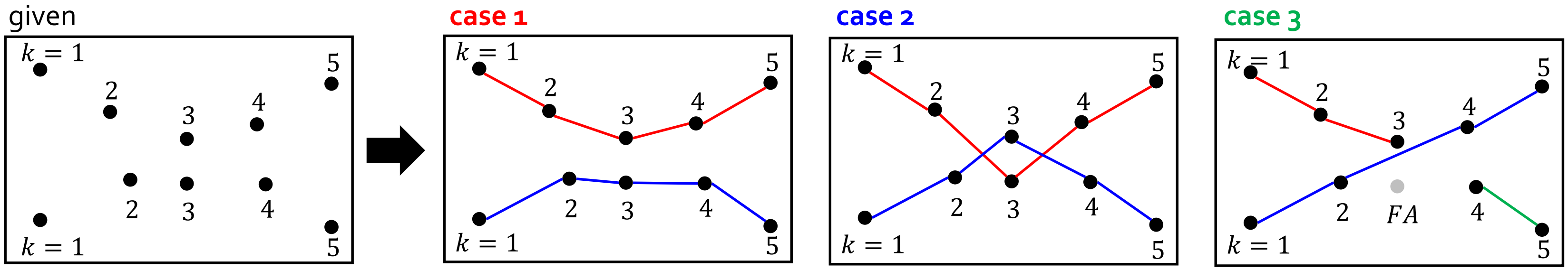
Joint Probabilistic Data Association Filter (JPDAF)

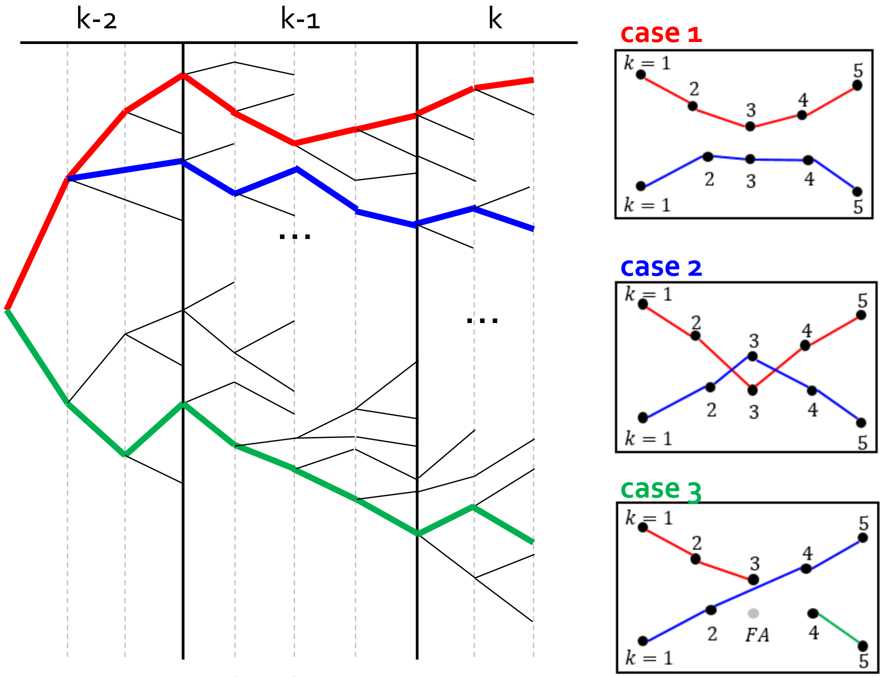


- JPDAF more robust for multiple object tracking.

Multiple Hypothesis Tracking (MHT)

- Resolves assignment ambiguities by delaying measurement-to-track decisions.



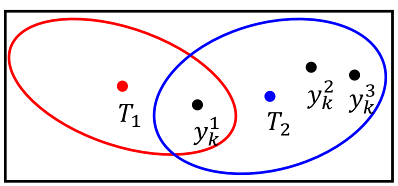


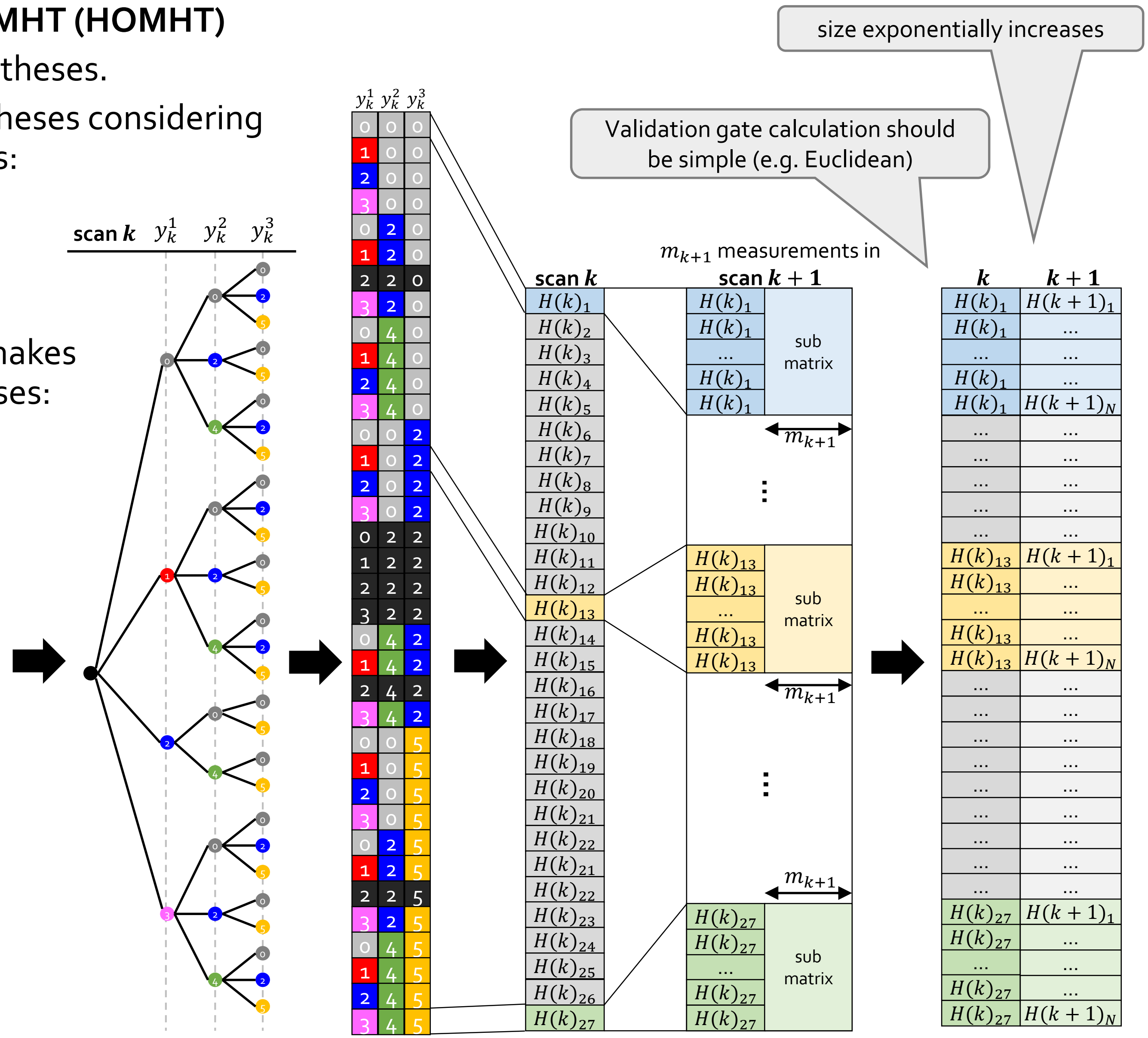
- Two main methods:
  1. Hypotheses Oriented MHT (HOMHT)
  2. Target Oriented MHT (TOMHT) → multi-dimensional assignment problem.

Multiple Hypothesis Tracking (MHT)

1. Hypotheses Oriented MHT (HOMHT)

- Expand existing hypotheses.
- Form different hypotheses considering each measurement as:
  1. new track
  2. false alarm
  3. existing track
- Each measurement makes up at least 2 hypotheses:
  1. new track
  2. false alarm





Multiple Hypothesis Tracking (MHT)

- For each hypothesis,
  1. State and state covariance update is performed in parallel.
     - $\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k(y_k - \hat{y}_k)$
     - $P_{k|k} = (I - K_k H_k) P_{k|k-1} (I - K_k H_k)^T + K_k \hat{S}_k {K_k}^T$
  2. Hypothesis probability is calculated as well.

- m_k^ν = number of new target (based on hypothesis)
- m_k^ν = number of false alarms

$$\boldsymbol{P}\left(\boldsymbol{\Theta}^{k,H} | \boldsymbol{Y}^{k}\right) = \frac{1}{c} \boxed{\frac{m_k^F!\, m_k^\nu!}{m_k!}} \mu_F\left(m_k^F\right) \boxed{\mu_\nu\left(m_k^\nu\right)} V^{-m_k^F - m_k^\nu} \prod_{i=1}^{m_k} \mathcal{N}\left(z_k^{i_T}; 0, \hat{S}_k\right)^{\{assigned\}} \prod_{T \in \mathcal{T}_D} {P_D^T}^{\{detected\}} \left(1 - P_D^T\right)^{\{\sim detected\}} \boxed{\boldsymbol{P}\left(\boldsymbol{\Theta}^{k-1,H} | \boldsymbol{Y}^{k-1}\right)}$$

Resulting from
${}_{m_k}\mathrm{P}_{(m_k - m_k^F - m_k^\nu)} \times {}_{(m_k^F + m_k^\nu)}\mathrm{C}_{m_k^\nu}$

pdf of new track appearing (assumption)

Hypothesis probability from the last (parent) scan

- Make decision using past N hypotheses ($N = window\ size, e.g. N = 3$)
  1. Reduce hypotheses → if corresponding hypothesis probability is low, discard it.
  2. For each hypotheses from each past scan, loop each column and make decision on track assignment
     - case i)

y_k^1	y_k^2	y_k^3	
1	2	0	H_k^1
1	4	0	H_k^2
1	4	2	H_k^3
1	0	5	H_k^4
1	2	5	H_k^5
1	4	5	H_k^6

→ can confirm y_k^1 belongs to track 1

     - case ii)

y_k^1	y_k^2	y_k^3	
1	4	2	H_k^1
1	2	5	H_k^2
1	0	5	H_k^3
1	0	2	H_k^4
2	0	2	H_k^5
2	0	5	H_k^6

$P_1 = \sum_{i=1}^{4} P(H_k^i)$

$P_2 = \sum_{i=5}^{6} P(H_k^i)$

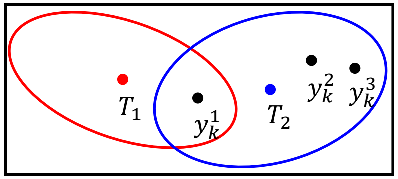
→If $P_1 > P_2$ and $P_1 > thres$
→ can confirm y_k^1 belongs to track 1

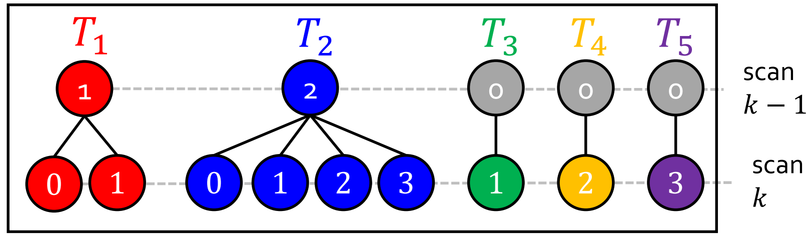
- Update track with:
  1. most probable hypothesis → GNN approach
  2. use weighted sum using track probabilities, in order to combine the estimates from all the hypotheses that contain the track → JPDAF approach

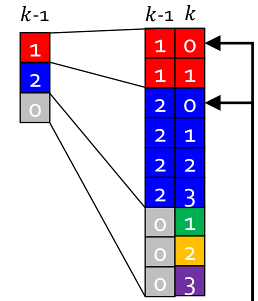
Multiple Hypothesis Tracking (MHT)

2. **Target Oriented MHT (TOMHT)** → door to advanced multiple object tracking techniques

- Instead of expanding on all hypotheses, expand on all tracks.







- For each track
  - begin expansion with false alarm assignment
  - Calculate track probability in parallel

    → similar to HOMHT

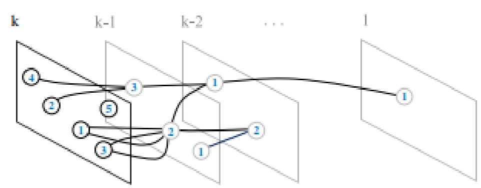
- Find the best measurement-to-track match across a window size:
  → Since 1 measurement cannot be assigned to 2 different tracks at each scan, **multi-dimensional assignment problem** needs to be solved:

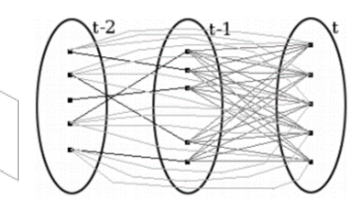
For example,

k-4	k-3	k-2	k-1	k	
1	1	1	0	0	$P(T_1^1)$
1	1	1	0	2	$P(T_1^2)$
1	1	1	1	1	$P(T_1^3)$
1	1	2	1	3	$\boldsymbol{P(T_1^4)}$
1	2	3	3	0	$P(T_1^5)$
1	2	3	3	4	$P(T_1^6)$
2	2	2	4	0	$P(T_2^1)$
2	2	2	4	5	$\boldsymbol{P(T_2^2)}$
2	3	2	4	0	$P(T_2^3)$
2	3	2	4	5	$P(T_2^4)$

Assume, the top two highest track probabilities are:
$\boldsymbol{P(T_1^4)}$ and $\boldsymbol{P(T_2^2)}$
→ conflict occurs in measurement-to-track matching at scan $k-2$.
→ Highest track probability does not mean the best measurement-to-track assignment.

- "**The Lagrangian relaxation method for solving integer programming problems.**", Fisher, 1981
- "**The maximum weight independent set problem for data association in multiple hypothesis tracking.**", Papageorgiou, 2009





Further Researches

- View Multiple Objet Tracking as:
  - **Linear Programming:**
    1. "Application of 0-1 integer programming to multitarget tracking problems.", Morefield, 1977
    2. "An efficient implementation of Reid's multiple hypothesis tracking algorithm and its evaluation for the purpose of visual tracking.", Cox, 1996
    3. "An LP-based algorithm for the data association problem in multitarget tracking.", Storms, , 2003
    4. "A linear programming approach for multiple object tracking.", Jiang, 2007.
  - **Multi-dimensional assignment problem**
    1. "An applications oriented guide to Lagrangian relaxation.", Fisher, 1985
    2. "The maximum weight independent set problem for data association in multiple hypothesis tracking", Dimitri, 2009
  - **Dynamic Programming**
    1. "Finding the best set of K paths through a trellis with application to multitarget tracking.", Wolf, 1989
    2. "Robust people tracking with global trajectory optimization.", Berclaz, 2006.
    3. "Globally-optimal greedy algorithms for tracking a variable number of objects.", Pirsiavash, *2011*
  - **Algorithmic**
    1. "Markov chain Monte Carlo data association for multi-target tracking.", Oh, 2009

MHT Application

- **"Spatially grounded multi-hypothesis tracking of people."** [ICRA2009]
  Luber, Matthias, Gian Diego Tipaldi, and Kai O. Arras. "Spatially grounded multi-hypothesis tracking of people." *Proceedings of the ICRA Workshop on People Detection and Tracking*. Vol. 104. 2009.
- Key idea:
  - MHT assumes that new track and false alarm events are uniformly distributed in the sensor field of view V with fixed Poisson rates, λ.
    → only justifiable for settings which MHT has been originally developed (e.g., radar or underwater sonar). However, in the context of people tracking with vision or laser these models are overly simplified.

$$\boldsymbol{P}\left(\boldsymbol{\Theta}^{\boldsymbol{k},\boldsymbol{H}} \middle| \boldsymbol{Y}^{\boldsymbol{k}}\right) = \frac{1}{c} \frac{m_k^F!\, m_k^\nu!}{m_k!} \boxed{\mu_F\left(m_k^F\right)\mu_\nu\left(m_k^\nu\right)V^{-m_k^F - m_k^\nu}} \prod_{i=1}^{m_k} \mathcal{N}\left(z_k^{i_T}; 0, \hat{S}_k\right)^{\{assigned\}} \prod_{T \in \mathcal{T}_D} P_D^{T\,\{detected\}}\left(1 - P_D^T\right)^{\{\sim detected\}} \boldsymbol{P}\left(\boldsymbol{\Theta}^{\boldsymbol{k-1},\boldsymbol{H}} \middle| \boldsymbol{Y}^{\boldsymbol{k-1}}\right)$$

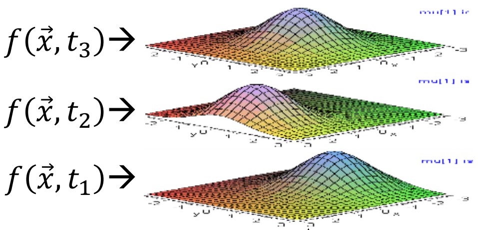
  - Incorporate learned distributions to MHT

- How?
  1. Re-design Poisson distribution from λ to $\lambda(\vec{x}, t)$, where $\vec{x} \in \mathbb{R}^2$.
     → $\lambda(\vec{x}, t)$ = spatio-temporal distribution of events
  2. Learn spatio-temporal distribution, $\lambda(\vec{x}, t)$, of events as tracking continues.

MHT Application

1. Re-design Poisson distribution from λ to $\lambda(\vec{x}, t)$, where $\vec{x} \in X$.
   - Let $N(t)$ be the number of events occurring up to time t with rate λ, then $N(t)$ follows Poisson distribution: $\mu(N(t) = k) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}$
   - Here, let λ be the generalized *rate function* $\lambda(t)$
     - Then, expected number of events between time a and b is:

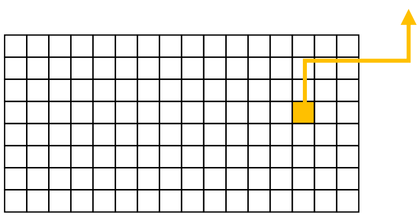
$$\lambda_{a,b} = \int_a^b \lambda(t) dt$$



   - Then, introduce a spatial dependency on the *rate function* $\lambda(\vec{x}, t)$
     - Then, expected number of events at time t in subspace $S \in X$ is:

$$\lambda_S(t) = \int_S \lambda(\vec{x}, t)\, d\vec{x}$$

     - In which case, $\vec{x}$ and t are decoupled in the generalized rate function → $\lambda(\vec{x}, t) = f(\vec{x})\lambda(t)$
       - $\lambda(t)$ = occurrence rate of events
       - $f(\vec{x})$ = probability distribution on where the event occurs in space.
     - Also, give constraint on $f(\vec{x})$ → $\int_X f(\vec{x}) d\vec{x} = 1$
   - Subdivide tracking environment into discrete cells
   - Each cell represents a local homogeneous Poisson process with a fixed-rate over time, $P_{ij}(k) = e^{-\lambda_{ij}} \frac{(\lambda_{ij})^k}{k!}$ where λ_{ij} is constant over time.
   - Finally, the generalized rate function $\lambda(\vec{x}, t)$ using grid approximation: $\lambda(\vec{x}, t) \cong \sum_{(i,j) \in X} \lambda_{ij} \mathbf{1}_{ij}(\vec{x})$, where $\mathbf{1}_{ij}(\vec{x}) = \begin{cases} 1, x \in cell_{ij} \\ 0, x \notin cell_{ij} \end{cases}$



tracking environment

MHT Application

2. Learn Poisson distribution parameters as tracking goes on
   = learn rate function $\lambda(\vec{x}, t)$

   - Model λ using $Gamma$ distribution, as it is a *conjugate prior* of Poisson distribution.
     i.e. λ is distributed according to $Gamma$ density: $Gamma(\lambda;\ \alpha, \beta) = \frac{\beta^{\alpha}}{(\alpha-1)!}\lambda^{\alpha-1}e^{-\beta\lambda}$
   - Learning λ = estimating parameters of Gamma distribution, α and β.
   - Bayes' rule → $\underbrace{P(\lambda_i|K_{1..i})}_{\text{posterior}} \propto \underbrace{P(k_i|\lambda_{i-1})}_{\text{likelihood}}\underbrace{P(\lambda_{i-1})}_{\text{prior}}$, i= time index
     - $P(\lambda_{i-1}) = Gamma(\alpha_{i-1}, \beta_{i-1})$, prior
     - $P(k_i|\lambda_{i-1}) = P(k_i) = e^{-\lambda_i}\frac{(\lambda_i)^{k_i}}{k_i!}$, likelihood

conjugate prior : Poisson distribution and Gamma distribution is in the same family of distributions
e.g. if $\alpha = k + 1$, $\beta = 1$ → Poisson distribution: $e^{-\lambda}\frac{(\lambda)^k}{k!}$

   - By substitution, turns out that
     - $\alpha_i = \alpha_{i-1} + k_i$
     - $\beta_i = \beta_{i-1} + 1$
     - start out with $\alpha_0 = \beta_0 = 1$
   - Expected value (mean) of $Gamma$ distribution, $Gamma(\alpha, \beta)$, is: $\frac{\alpha}{\beta}$
     → i.e. $\mathrm{E}(\lambda) = event\ occuring\ ratio = \frac{number\ of\ events\ occuring}{number\ of\ measurements} = \frac{\alpha}{\beta}$

Thank you

- Amditis, A., et al. (2012). Multiple hypothesis tracking implementation. Laser Scanner Technology, InTech.
- Bar-Shalom, Y. and E. Tse (1975). "Tracking in a cluttered environment with probabilistic data association." Automatica 11(5): 451-460.
- Bar-Shalom, Y., et al. (2007). "Dimensionless score function for multiple hypothesis tracking." IEEE Transactions on Aerospace and Electronic Systems 43(1).
- Bar-Shalom, Y., et al. (2009). "The probabilistic data association filter." IEEE Control Systems 29(6).
- Bar-Shalom, Yaakov, and Xiao-Rong Li. Multitarget-multisensor tracking: principles and techniques. Vol. 19. London, UK:: YBs, 1995.
- Bar-Shalom, Yaakov, X. Rong Li, and Thiagalingam Kirubarajan. Estimation with applications to tracking and navigation: theory algorithms and software. John Wiley & Sons, 2004.
- Bar-Shalom, Yaakov. "Tracking methods in a multitarget environment." IEEE Transactions on automatic control 23.4 (1978): 618-626.
- Beyers, R. J. (1988). Joint Probability Data Association (JPDA) on Tracking Multiple Munitions Fragments, AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH SCHOOL OF ENGINEERING.
- Blackman, Samuel S. "Multiple hypothesis tracking for multiple target tracking." IEEE Aerospace and Electronic Systems Magazine 19.1 (2004): 5-18.
- Cox, Ingemar J., and Sunita L. Hingorani. "An efficient implementation of Reid's multiple hypothesis tracking algorithm and its evaluation for the purpose of visual tracking." IEEE Transactions on pattern analysis and machine intelligence 18.2 (1996): 138-150.
- Fortmann, Thomas, Yaakov Bar-Shalom, and Molly Scheffe. "Sonar tracking of multiple targets using joint probabilistic data association." IEEE journal of Oceanic Engineering 8.3 (1983): 173-184.
- Fortmann, Thomas E., Yaakov Bar-Shalom, and Molly Scheffe. "Multi-target tracking using joint probabilistic data association." Decision and Control including the Symposium on Adaptive Processes, 1980 19th IEEE Conference on. IEEE, 1980.
- Habtemariam, Biruk K., et al. "Multiple detection probabilistic data association filter for multistatic target tracking." Information Fusion (FUSION), 2011 Proceedings of the 14th International Conference on. IEEE, 2011.
- Konstantinova, Pavlina, Alexander Udvarev, and Tzvetan Semerdjiev. "A study of a target tracking algorithm using global nearest neighbor approach." Proceedings of the International Conference on Computer Systems and Technologies (CompSysTech'03). 2003.
- Kuhn, Harold W. "The Hungarian method for the assignment problem." Naval Research Logistics (NRL) 2.1-2 (1955): 83-97.
- Ni, Long-qiang, et al. "Rough Sets Probabilistic Data Association Algorithm and its Application in Multi-target Tracking." Defence Technology 9.4 (2013): 208-216.
- Rasmussen, Christopher, and Gregory D. Hager. "Probabilistic data association methods for tracking complex visual objects." IEEE Transactions on Pattern Analysis and Machine Intelligence 23.6 (2001): 560-576.
- Reid, Donald. "An algorithm for tracking multiple targets." IEEE transactions on Automatic Control 24.6 (1979): 843-854.
- Rothberg, L., et al. "Dynamics of photogenerated solitons in trans-polyacetylene." IEEE journal of quantum electronics24.2 (1988): 311-314.