Set of trajectories, conjugate priors and metrics

Lennart Svensson

Signal Processing group
Chalmers University of Technology, Sweden

Joint work with Ángel García-Fernández, Mark Morelande, Karl Granström, Jason Williams, Abu Sajana Rahmathullah, Maryam Fatemi and Yuxuan Xia.

MTT and MTF

- We consider two closely related problems[1]:

Multi-target tracking (MTT)

In MTT, the objective is to estimate **target trajectories**, including when targets appear and disappear.

Multi-target filtering (MTF)

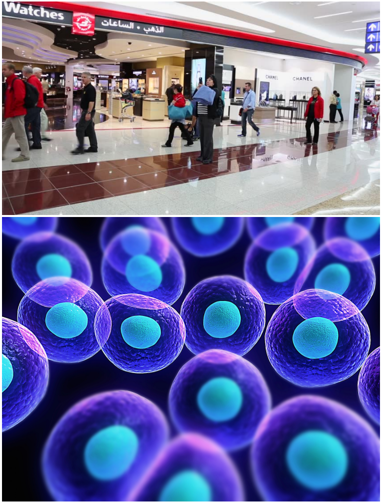
In MTF, the objective is to estimate states of **targets** that are **currently present**.

[1] Ristic, B. et al, "A metric for performance evaluation of multi-target tracking algorithms", *IEEE Trans. of Sign. Proc.*, 59(7), 2011.

Why multi-target tracking (MTT)?

MTT is important in many contexts:

- Airport/domestic surveillance:
  - e.g., to analyze who dropped the bag by the entrance.
- Trajectory information can provide information about:
  - properties of cells,
  - object types (birds/UAVs/...)
  - how much a football player runs.

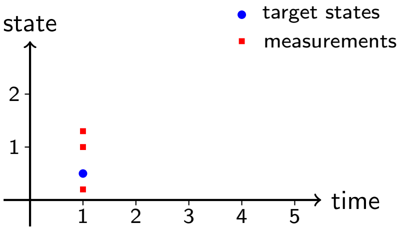


From goo.gl/sckyBQ.

An MTT illustration

- **Standard MTT models** for point objects:

  - Targets move and may appear/disappear with time.
  - Measurements:
    1. 0/1 detections per target,
    2. possibly also false detections,
    3. unknown associations targets–detections.



Objective

- Estimate trajectories from sequence of detections.

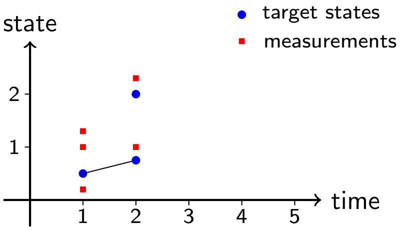
An MTT illustration

- **Standard MTT models** for point objects:

– Targets move and may appear/disappear with time.

– Measurements:

1. 0/1 detections per target,
2. possibly also false detections,
3. unknown associations targets–detections.



Objective

- Estimate trajectories from sequence of detections.

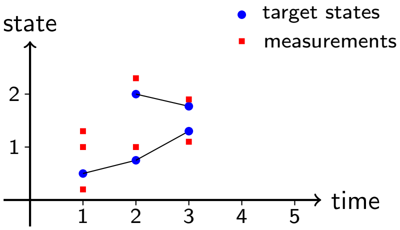
An MTT illustration

- **Standard MTT models** for point objects:

  – Targets move and may appear/disappear with time.

  – Measurements:

    1. 0/1 detections per target,
    2. possibly also false detections,
    3. unknown associations targets–detections.



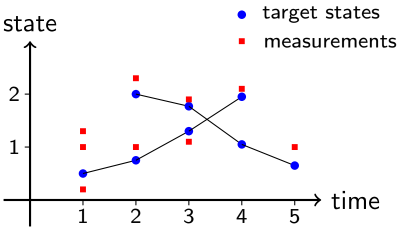
Objective

- Estimate trajectories from sequence of detections.

An MTT illustration

- **Standard MTT models** for point objects:

  - Targets move and may appear/disappear with time.
  - Measurements:
    1. 0/1 detections per target,
    2. possibly also false detections,
    3. unknown associations targets–detections.

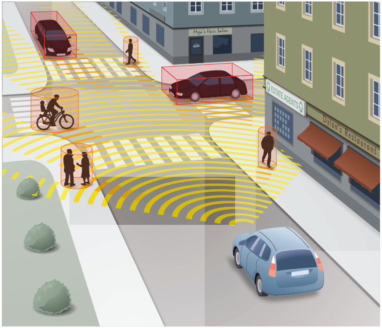


Objective

- Estimate trajectories from sequence of detections.

Why multi-target tracking (MTF)?

- We work with **self-driving vehicles:**
  - estimate states of nearby road users,
  - enables us to avoid collisions,
  - target trajectories often not important.



- Other applications include positioning of
  1. airplanes,
  2. human cells,
  3. space debris.

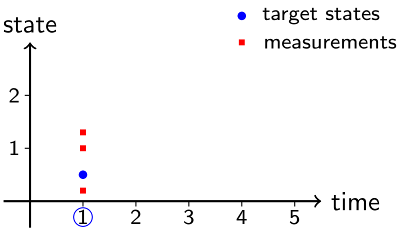
An MTF illustration

- **Identical models** for MTT and MTF!
- Standard models for point objects:

  – Targets move and may appear/disappear with time.

  – Measurements:

    1. 0/1 detections per target,
    2. possibly also false detections,
    3. unknown associations targets–detections.



Objective

- Estimate current target states from sequence of detections.

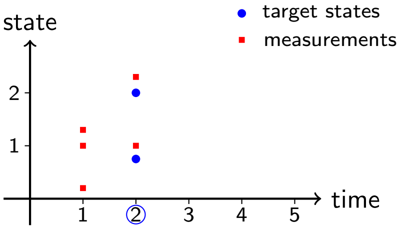
An MTF illustration

- **Identical models** for MTT and MTF!
- Standard models for point objects:

– Targets move and may appear/disappear with time.

– Measurements:

1. 0/1 detections per target,
2. possibly also false detections,
3. unknown associations targets–detections.



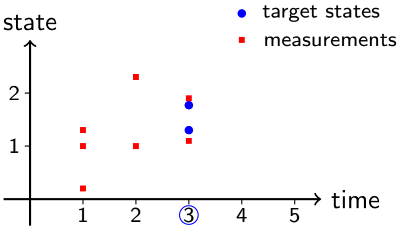
Objective

- Estimate current target states from sequence of detections.

An MTF illustration

- **Identical models** for MTT and MTF!
- Standard models for point objects:

  - Targets move and may appear/disappear with time.
  - Measurements:
    1. 0/1 detections per target,
    2. possibly also false detections,
    3. unknown associations targets–detections.



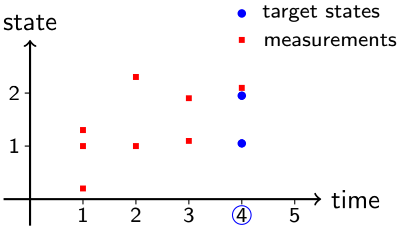
Objective

- Estimate current target states from sequence of detections.

An MTF illustration

- **Identical models** for MTT and MTF!
- Standard models for point objects:

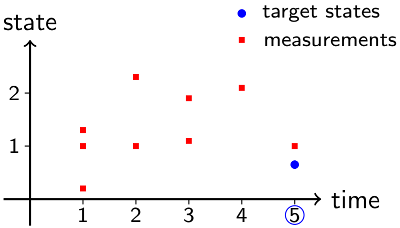
  - Targets move and may appear/disappear with time.
  - Measurements:
    1. 0/1 detections per target,
    2. possibly also false detections,
    3. unknown associations targets–detections.



Objective

- Estimate current target states from sequence of detections.

- **Identical models** for MTT and MTF!
- Standard models for point objects:
  - Targets move and may appear/disappear with time.
  - Measurements:
    1. 0/1 detections per target,
    2. possibly also false detections,
    3. unknown associations targets–detections.



Objective

- Estimate current target states from sequence of detections.

MTT and MTF solutions?

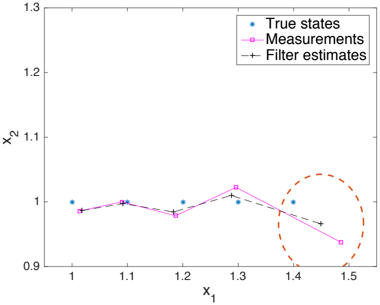
- **How to approach MTT and MTF**?
- Most attempts are **Bayesian**:

– optimal trade-off between information

1. about realistic trajectories,
2. from measurements.

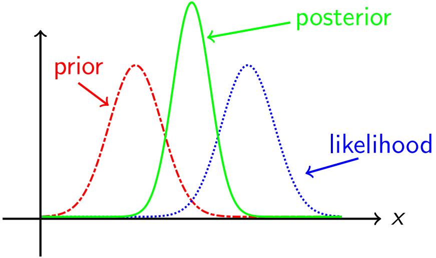
– gives a posterior density
$\Rightarrow$ we can compute

1. probabilities of different events,
2. optimal decisions/estimates (e.g., MMSE).



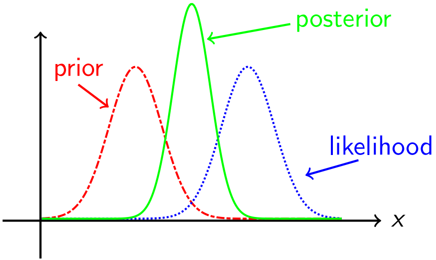
In Bayesian statistics:

- we compute **posterior densities** of, x,
- posterior density summarizes **what we know** about x,
- Very useful! E.g., can compute **optimal estimates**.



In Bayesian statistics:

- we compute **posterior densities** of, x,
- posterior density summarizes **what we know** about x,
- Very useful! E.g., can compute **optimal estimates**.



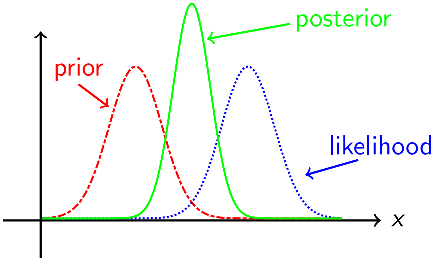
Outline:

1) **Sets of trajectories**: suitable x in MTT and MTF? Which are our quantities of interest?

Bayesian statistics and outline

In Bayesian statistics:

- we compute **posterior densities** of, x,
- posterior density summarizes **what we know** about x,
- Very useful! E.g., can compute **optimal estimates**.



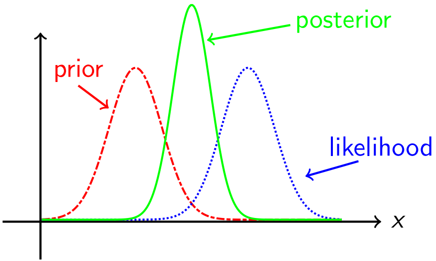
Outline:

1) **Sets of trajectories**: suitable x in MTT and MTF? Which are our quantities of interest?

2) **Conjugate prior densities**: reasonable priors and likelihoods to obtain tractable posteriors?

In Bayesian statistics:

- we compute **posterior densities** of, x,
- posterior density summarizes **what we know** about x,
- Very useful! E.g., can compute **optimal estimates**.



Outline:

1) **Sets of trajectories**: suitable x in MTT and MTF? Which are our quantities of interest?
2) **Conjugate prior densities**: reasonable priors and likelihoods to obtain tractable posteriors?
3) **Metrics**: how can we measure performance in MTT and MTF?

Some notation

- **Target states:** (for a single target)
  - are denoted x_k where k is our time index.
  - often contain position, velocity, etc.
  - may also contain other properties: color, age, size, etc.

- **Measurements:**
  - $\mathbf{z}^k = \{z_1^k, z_2^k, \dots,\}$ is the set of measurements at time k.
  - $\mathbf{Z}^k = (\mathbf{z}^1, \dots, \mathbf{z}^k)$ denotes the sequence of measurements up to and including time k.

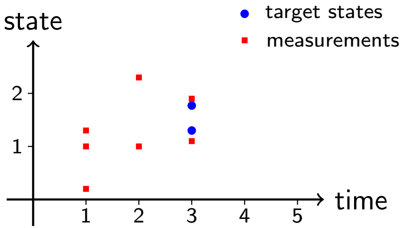
- **MTF:** the **set of target states** is a suitable state

$$\mathbf{x}^k = \left\{ x_1^k, x_2^k, \ldots, x_{n_k}^k \right\}$$

where n_k is ♯targets present at time k.

- **Example:**

$$\mathbf{x}^3 = \{1.3,\ 1.8\}$$
$$n_3 = 2.$$



Objective in MTF

Recursively compute $p(\mathbf{x}^k | \mathbf{Z}^k)$.

- Why is

$$\mathbf{x}^k = \{x_1^k, x_2^k, \ldots, x_{n_k}^k\}$$

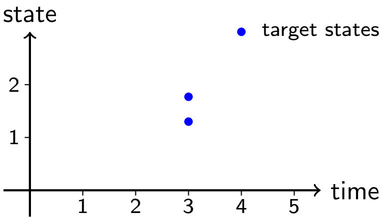
a suitable state?

1. $\mathbf{x}^k$ captures/is our quantity of interest,
2. $\mathbf{x}^k$ is a minimal representation:

$$\mathbf{x}^k \overset{\text{1-1}}{\longleftrightarrow} \text{physical quantities of interest.}$$

- **Example:**

$$\mathbf{x}^3 = \{1.3,\ 1.8\}$$
$$n_3 = 2.$$



MTF algorithms

- We can use conjugate priors (see 2nd part) to recursively approximate
$$p(\mathbf{x}^k \middle| \mathbf{Z}^k)$$
using PMB and PMBM filters.

- Many other filters have also been developed, including
  1. Probability Hypothesis Density (PHD)
  2. Cardinalized PHD (CPHD)

  filters.

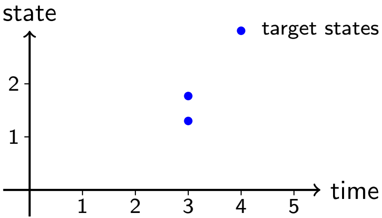
- This state representation can also be used to motivate Multiple Hypothesis Tracking (MHT) algorithms from a Bayesian perspective.

State in MTF

- Why not use, e.g., an ordered vector
$$\tilde{\mathbf{x}}^k = \left[x_1^k, x_2^k, \ldots, x_{n_k}^k\right]?$$
The ordering
  1. does not convey relevant information,
  2. cannot (generally) be inferred from physical reality.

- **Example:**

$$\tilde{\mathbf{x}}^3 = [1.3,\ 1.8] \quad \text{or}$$
$$\tilde{\mathbf{x}}^3 = [1.8,\ 1.3]?$$



- Both the transition model and posterior would need to handle uncertainties in the ordering
$\rightsquigarrow$ arbitrary choices and irrelevant uncertainties!

- **MTT:** we argue that the **set of trajectories** is a suitable state

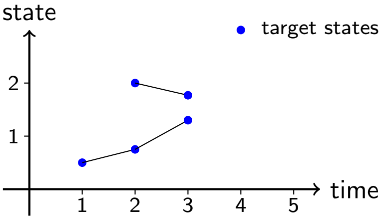
$$\mathbf{X}^k = \{X_1^k, X_2^k, \ldots, X_{N_k}^k\},$$

where X_i^k is a trajectory and N_k is ♯targets present until time k.

- We denote trajectories as $X = (t, x^{1:i})$, where
  - t: start time,
  - i duration,
  - $x^{1:i} = (x^1, x^2, \ldots, x^i)$ sequence of target states.

- **Example:**
  $\mathbf{X}^3 = \{X_1^3, X_2^3\}$ where

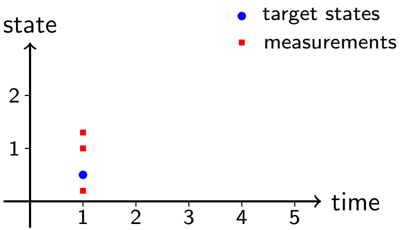
$$X_1^3 = (1, (0.5, 0.8, 1.3))$$
$$X_2^3 = (2, (2, 1.8)).$$



- **Note:** three possible trajectory types during prediction:
  1. persist/survive: extended by including the new state at the end of the trajectory.
  2. ended/dead: remain the same.
  3. newly born: obtain a trajectory $X = (k, x^k)$.
- Trajectories are never removed from the set.

- **Example:**
  $\mathbf{X}^1 = \{X_1^1\}$ where

  $$X_1^5 = (1, (0.5)).$$



Objective in MTT

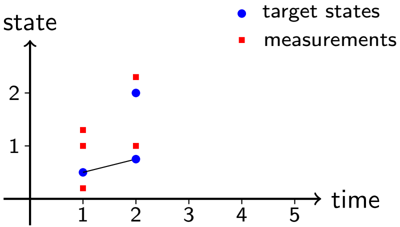
Recursively compute $p(\mathbf{X}^k | \mathbf{Z}^k)$.

State in MTT

- **Note:** three possible trajectory types during prediction:
  1. persist/survive: extended by including the new state at the end of the trajectory.
  2. ended/dead: remain the same.
  3. newly born: obtain a trajectory $X = (k, x^k)$.
- Trajectories are never removed from the set.

- **Example:**
  $\mathbf{X}^2 = \{X_1^2, X_2^2\}$ where

$$X_1^5 = (1, (0.5, 0.8))$$
$$X_2^5 = (2, (2)).$$



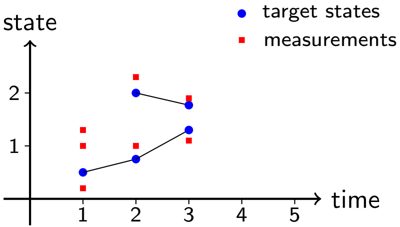
Objective in MTT

Recursively compute $p(\mathbf{X}^k | \mathbf{Z}^k)$.

- **Note:** three possible trajectory types during prediction:
  1. persist/survive: extended by including the new state at the end of the trajectory.
  2. ended/dead: remain the same.
  3. newly born: obtain a trajectory $X = (k, x^k)$.
- Trajectories are never removed from the set.

- **Example:**
  $\mathbf{X}^3 = \{X_1^3, X_2^3\}$ where

$$X_1^3 = (1, (0.5, 0.8, 1.3))$$
$$X_2^3 = (2, (2, 1.8)).$$



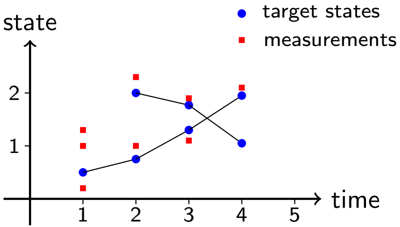
Objective in MTT

Recursively compute $p(\mathbf{X}^k | \mathbf{Z}^k)$.

- **Note:** three possible trajectory types during prediction:
  1. persist/survive: extended by including the new state at the end of the trajectory.
  2. ended/dead: remain the same.
  3. newly born: obtain a trajectory $X = (k, x^k)$.
- Trajectories are never removed from the set.

- **Example:**
  $\mathbf{X}^4 = \{X_1^4, X_2^4\}$ where

$$X_1^4 = (1, (0.5, 0.8, 1.3, 1.9))$$
$$X_2^4 = (2, (2, 1.8, 1.0)).$$



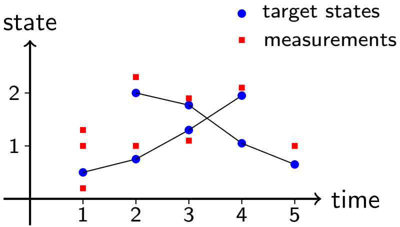
Objective in MTT

Recursively compute $p(\mathbf{X}^k | \mathbf{Z}^k)$.

- **Note:** three possible trajectory types during prediction:
  1. persist/survive: extended by including the new state at the end of the trajectory.
  2. ended/dead: remain the same.
  3. newly born: obtain a trajectory $X = (k, x^k)$.
- Trajectories are never removed from the set.

- **Example:**
  $\mathbf{X}^5 = \{X_1^5, X_2^5\}$ where

$$X_1^5 = (1, (0.5, 0.8, 1.3, 1.9))$$
$$X_2^5 = (2, (2, 1.8, 1.0, 0.6)).$$



Objective in MTT

Recursively compute $p(\mathbf{X}^k | \mathbf{Z}^k)$.

- Why is

$$\mathbf{X}^k = \{X_1^k, X_2^k, \ldots, X_{n_k}^k\}$$

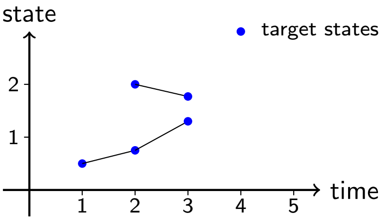
a suitable state?

1. $\mathbf{X}^k$ captures/is our quantity of interest,
2. $\mathbf{X}^k$ is a minimal representation:

$$\mathbf{X}^k \overset{\text{1-1}}{\longleftrightarrow} \text{ physical quantities of interest.}$$

- **Example:**
$\mathbf{X}^3 = \{X_1^3, X_2^3\}$ where

$$X_1^3 = (1, (0.5, 0.8, 1.3))$$
$$X_2^3 = (2, (2, 1.8)).$$

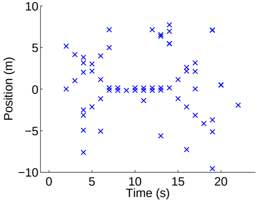


MTT using sets of trajectories

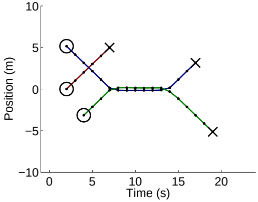
- One can use **conjugate priors for sets of trajectories** to develop algorithms.
- For standard models, the result resembles MHT:
  1. hypothesis trees that grow rapidly,
  2. merge/prune branches to reduce complexity.
- **Complexity**?
  - Suppose we are given a data association hypothesis, θ.
  - To compute $p(\mathbf{X}^k | \mathbf{Z}^k, \theta)$ we should smooth our estimates at all times, $1, 2, \ldots, k$. **Unfeasible!**
  - In practice, we often only update estimates for last L steps, $k - L + 1, \ldots, k$, where L is a design variable.

Example: standard measurement model

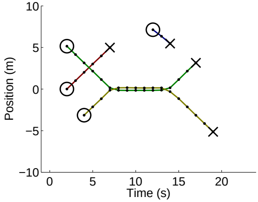
- Each target produces a noisy measurement with a probability of detection. There is additional clutter.



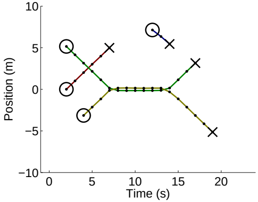
(a) Measurements



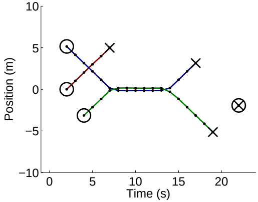
(b) 0.42



(c) 0.12



(d) 0.12



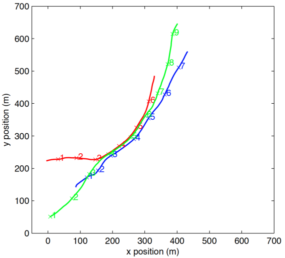
(e) 0.04

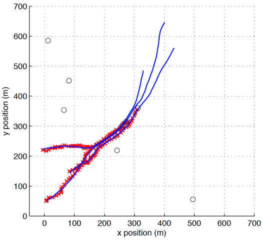
TPHD: a PHD filter for trajectories

- One can also use sets of trajectories to **extend the PHD filter to trajectories**.
- **Idea:** recursively approximate $p(\mathbf{X}^k | \mathbf{Z}^k)$ as a Poisson multitrajectory PDF,

$$\nu(\{X_1, \ldots, X_N\}) = e^{-\lambda_\nu} \lambda_\nu^N \prod_{i=1}^{N} \breve{\nu}(X_i).$$

- Scenario and results assuming standard model for point targets:





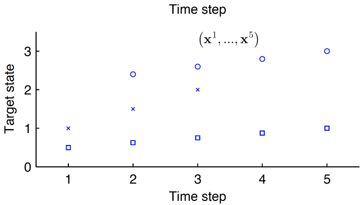
(a) $k = 50$

- The algorithm efficiently estimates trajectories in a principled manner.

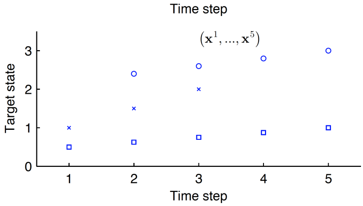
Labeled sets as state

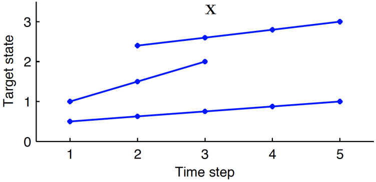
- Another approach to MTT is to use **sets of labelled target states**
  - append target states with a static label, ℓ,
$$\tilde{x} = (x, \ell).$$
  - two targets cannot share the same label,
  - normally, labels **lack physical interpretation** and are unobservable,
  - $\mathbf{x}^k = \{\tilde{x}_1^k, \ldots, \tilde{x}_{n_k}^k\}$ is used as state variable.

- Labels enable us to **connect state estimates across time**.
- **Labels are arbitrary**; we could replace $\circ$ with $\Diamond$ and $\square$ with $\circ$ without changing trajectories.



- A sequence of sets of labeled states (left figure) fully describe a set of trajectories (right figure).





- However, labels are arbitrary

$$(\mathbf{x}^1, \ldots, \mathbf{x}^k) \overset{\text{1-1}}{\not\leftrightarrow} \text{ physical quantities of interest.}$$

- In theory, one could compute $p(\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^k | \mathbf{Z}^k)$, but this is essentially a more involved version of $p(\mathbf{X}^k | \mathbf{Z}^k)$.

Labeled sets and trajectory estimates

- **Standard approach:** recursively compute

$$p(\mathbf{x}^k | \mathbf{Z}^k),$$

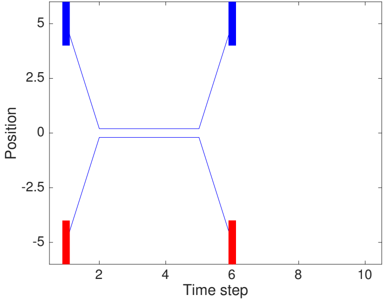
  and extract estimates $\hat{\mathbf{x}}^k$.

- In many cases, this yields reasonable estimates.
- In simple scenarios, it resembles approximating $p(\mathbf{X}^k | \mathbf{Z}^k)$ and use $L = 1$ in that there is no smoothing.
- However, let us look at these marginal densities

$$p(\mathbf{x}^1 | \mathbf{Z}^1),\ p(\mathbf{x}^2 | \mathbf{Z}^2), \ldots, p(\mathbf{x}^k | \mathbf{Z}^k),$$

  in a more challenging example.

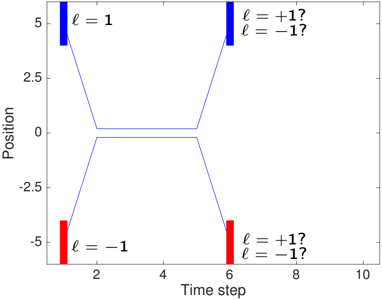
- **Toy example:** Suppose we are tracking two people, who occasionally meet to chat together. Labels are ± 1 at $k = 1$.
  - **k=1:** $\Pr[\mathbf{x}^1 = \{(+5,+1),(-5,-1)\}] = 1$
  - **k=6:** we are now confused about labels
    $\Pr[\mathbf{x}^6 = \{(+5,+1),(-5,-1)\}] =$
    $\Pr[\mathbf{x}^6 = \{(+5,-1),(-5,+1)\}] = 0.5.$



- Are label uncertainties/mixed labeling a problem?

Label uncertainties: example

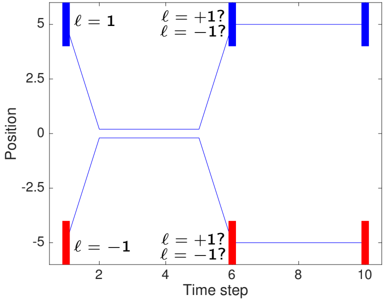
- **Toy example:** Suppose we are tracking two people, who occasionally meet to chat together. Labels are ± 1 at $k = 1$.
  - **k=1:** $\Pr[\mathbf{x}^1 = \{(+5, +1), (-5, -1)\}] = 1$
  - **k=6:** we are now confused about labels $\Pr[\mathbf{x}^6 = \{(+5, +1), (-5, -1)\}] = \Pr[\mathbf{x}^6 = \{(+5, -1), (-5, +1)\}] = 0.5.$



- Are label uncertainties/mixed labeling a problem?

Label uncertainties: example

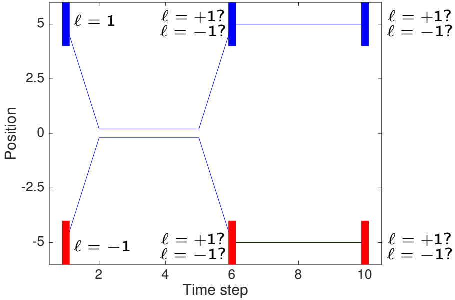
- **Toy example:** Suppose we can tell that the persons do not move from time 6 to 10.



- What do we know about the labels at time 10?

Label uncertainties: example

- **Toy example:** Suppose we can tell that the persons do not move from time 6 to 10.

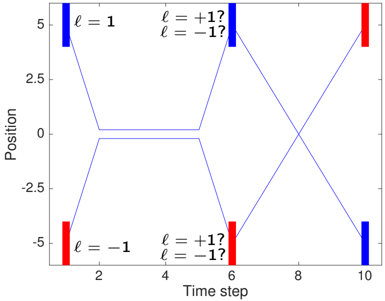


- What do we know about the labels at time 10?
- Still equally confused about labels:

  $\Pr[\mathbf{x}^{10} = \{(+5, +1), (-5, -1)\}] = \Pr[\mathbf{x}^{10} = \{(+5, -1), (-5, +1)\}] = 0.5.$

Label uncertainties: example

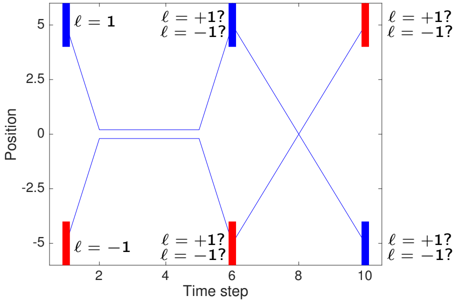
- **Toy example:** Suppose we can instead tell that the persons change place from time 6 to 10.



- What do we know about the labels at time 10?

Label uncertainties: example

- **Toy example:** Suppose we can instead tell that the persons change place from time 6 to 10.



- What do we know about the labels at time 10?
- Still equally uncertain about labels:

$$\Pr[\mathbf{x}^{10} = \{(+5,+1),(-5,-1)\}] = \Pr[\mathbf{x}^{10} = \{(+5,-1),(-5,+1)\}] = 0.5.$$

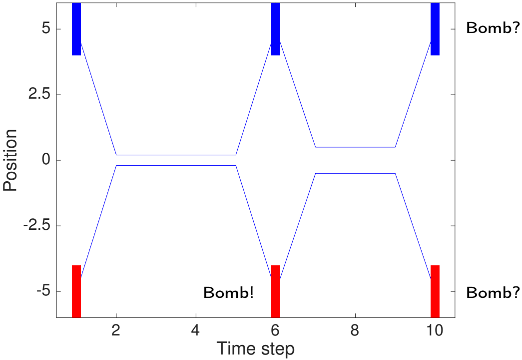
Conclusions?

- **Trajectories are assumed known** from time 6 to 10.
- Still, $p(\mathbf{x}^{10}|\mathbf{Z}^{10})$ is identical for static and crossing trajectories.

  $\Rightarrow$ $p(\mathbf{x}^{6}|\mathbf{Z}^{6})$ and $p(\mathbf{x}^{10}|\mathbf{Z}^{10})$ **provide no information** about how to "connect the dots".
- In fact, once we have "total mixed labeling", a labeled set contains as much trajectory information as an unlabeled set.
- Using **sets of trajectories**, we are able to tell if they stayed in the same place or not.

Label uncertainties: remark

- Note that trajectory information is often important and non-trivial to extract.

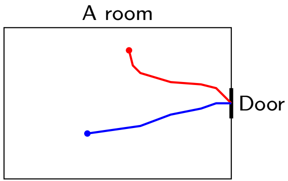


Label uncertainties: birth process

- Appearing targets are modeled using a birth process, and label **uncertainties** may arise already **at target birth**.

Example:

- Imagine tracking two people who enter a room together.
- In many cases, this leads to mixed labeling, even when trajectories can be estimated accurately.



- **Poisson processes are commonly used birth processes**, but the density of a labeled Poisson RFS satisfies, e.g.,

$$\pi(\{(x_1, \ell_1)(x_2, \ell_2)\}) = \pi(\{(x_1, \ell_2)(x_2, \ell_1)\}),$$

which **leads to mixed labeling** when two targets appear simultaneously, even if they are far apart.

Sets of trajectories: concluding remarks

- **MTT:** we argue that the **set of trajectories** is a suitable state

$$\mathbf{X}^k = \left\{X_1^k, X_2^k, \ldots, X_{N_k}^k\right\},$$

where X_i^k is a trajectory.

1. $\mathbf{X}^k$ is our quantity of interest,
2. $\mathbf{X}^k$ is a minimal representation:

$$\mathbf{X}^k \overset{\text{1-1}}{\longleftrightarrow} \text{physical quantities of interest.}$$

Objective

Compute

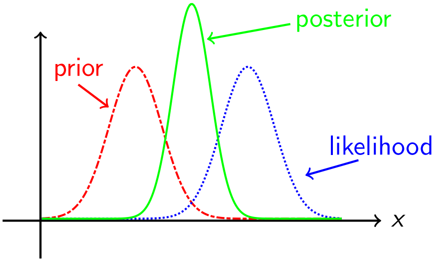
$$p(\mathbf{X}^k | \mathbf{Z}^k).$$

That is, compute **posterior density of our quantity of interest**.

- Sets of trajectories, $\mathbf{X}^k$, enable
  1. Development of novel algorithms, e.g., TPHD.
  2. Straightforward extraction of trajectory information in a fully Bayesian manner.

In Bayesian statistics:

- we compute **posterior densities** of, x,
- posterior density summarizes **what we know** about x,
- Very useful! E.g., can compute **optimal estimates**.



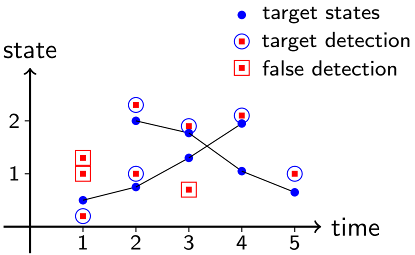
Outline:

1) Sets of trajectories: suitable x in MTT and MTF? Which are our quantities of interest?
2) **Conjugate prior densities**: reasonable priors and likelihoods to obtain tractable posteriors?
3) Metrics: how can we measure performance in MTT and MTF?

Models in MTF

Things to model:

- **Prior:**
  1. target birth
  2. target motion and death
- **Likelihood:**
  1. target detections
  2. false detections.
- **Posterior:** $p(\mathbf{x}^k | \mathbf{Z}^k)$



- **Key components:** multi-Bernoulli (MB) and Poisson processes.

Definition: Bernoulli

- $\mathbf{x}$ is Bernoulli random finite set (RFS) if

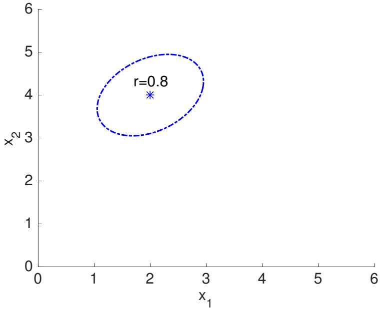
$$p(\mathbf{x}) = \begin{cases} 1-r & \text{if } \mathbf{x} = \emptyset \\ r\, p_1(x) & \text{if } \mathbf{x} = \{x\}, \end{cases}$$

that is, it can only contain zero or one object states.

- Here, r is an existence probability and $p_1(x)$ is a density.

Things it can model:

- Set of detections from a target
  $r =$prob. of detection, P_d.
- Distribution of a single target
  $r =$probability of existence



multi-Bernoulli RFS

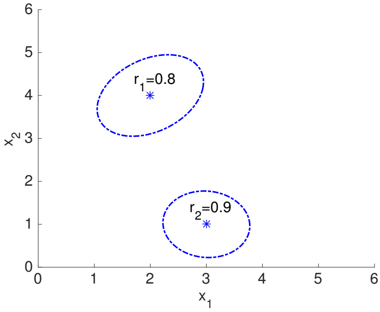
- $\mathbf{x}$ is **multi-Bernoulli** (MB) RFS if

$$\mathbf{x} = \bigcup_{i=1}^{n} \mathbf{x}_i$$

where $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are independent Bernoulli RFS.

Things MBs can model:

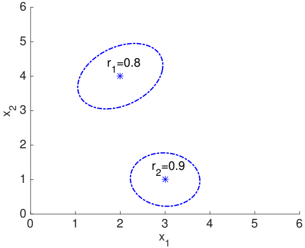
- Set of target detections, conditioned on the set of targets.
- Posterior distribution of targets.
- Distribution of appearing targets; the birth process.



- If $f_{ij}(\mathbf{x}_j)$ are Bernoulli densities, then

$$f_i^{mb}(\mathbf{x}) = \sum_{\mathbf{x}_1 \uplus \cdots \uplus \mathbf{x}_n = \mathbf{x}} \prod_{j=1}^{n} f_{ij}(\mathbf{x}_j)$$

is a MB density.
Models: n potential targets (tracks).

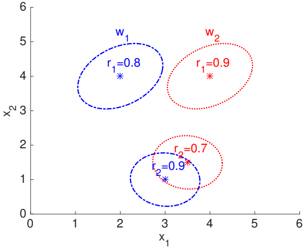


- **x is a MBM** RFS if its density is

$$f^{\mathrm{mbm}}(\mathbf{x}) = \sum_i w_i f_i^{mb}(\mathbf{x}),$$

where $\sum_i w_i = 1$.
Models:

1. w_i probability of data association hypothesis i
2. $f_i^{mb}(\mathbf{x})$: distribution of potential targets given ith hypothesis.



Definition: Poisson processes

- $\mathbf{x}$ is Poisson RFS with intensity $\lambda(x)$ if

$$p(\mathbf{x}) = exp\left[-\int \lambda(x)\,\mathrm{d}x\right] \prod_{x \in \mathbf{x}} \lambda(x)$$

- **Interesting properties:**

1. We can generate $\mathbf{x} = \{x_1, \ldots, x_n\} \sim p(\mathbf{x})$ by
   i) generating $n \sim \text{Poisson}(\bar{\lambda})$
   ii) generating $x_1, \ldots, x_n \sim \frac{\lambda(x)}{\bar{\lambda}}$
   where $\bar{\lambda} = \int \lambda(x)\,\mathrm{d}x$.
2. If $\mathbf{A}$ and $\mathbf{B}$ are two disjoint regions, $\mathbf{x} \bigcap \mathbf{A}$ and $\mathbf{x} \bigcap \mathbf{B}$ are independent Poisson processes.

Definition: Poisson processes

- $\mathbf{x}$ is Poisson RFS with intensity $\lambda(x)$ if

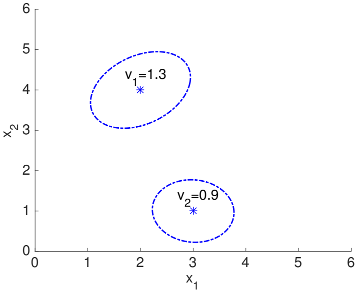
$$p(\mathbf{x}) = exp\left[-\int \lambda(x)\,\mathrm{d}x\right] \prod_{x \in \mathbf{x}} \lambda(x)$$

Things a Poisson RFS can model:

- Set of false detections (clutter measurements).
- Distribution of appearing targets; the birth process.
- Set of target detections from a single extended target.

Example:

$$\lambda(x) = 1.3\mathcal{N}(x; \mu_1, P_1) + 0.9\mathcal{N}(x; \mu_2, P_2)$$

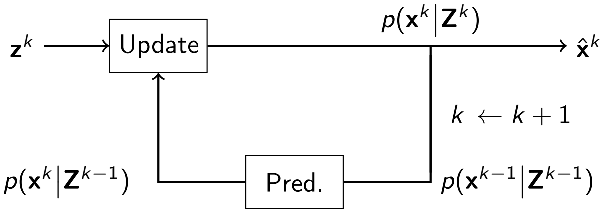


Filtering recursions

Objective in MTF

Recursively compute $p(\mathbf{x}^k | \mathbf{Z}^k)$.

- Most Bayesian filters rely on prediction and update steps:



Another key property

- $p(\mathbf{x}_{k-1} | \mathbf{Z}^{k-1})$ and $p(\mathbf{x}^k | \mathbf{Z}^k)$ are the same type of density
  $\Rightarrow$ we (may) have a recursive algorithm!

- **Example:** in a Kalman filter, both $p(\mathbf{x}^{k-1} | \mathbf{Z}^{k-1})$ and $p(\mathbf{x}^k | \mathbf{Z}^k)$ are Gaussian.

Filtering recursions

Standard approach to filtering

1. Select a density parameterization $p(\mathbf{x}; \boldsymbol{\theta})$.
2. Start from
$$p(\mathbf{x}^{k-1} \big| \mathbf{Z}^{k-1}) \approx p(\mathbf{x}^{k-1}; \boldsymbol{\theta}_{k-1|k-1})$$
and find $\boldsymbol{\theta}_{k|k}$ such that
$$p(\mathbf{x}^{k} \big| \mathbf{Z}^{k}) \approx p(\mathbf{x}^{k}; \boldsymbol{\theta}_{k|k}).$$

- $p(\mathbf{x}; \boldsymbol{\theta}) = \mathcal{N}(\mathbf{x}; \boldsymbol{\mu}, \mathbf{P})$ — KF, EKF, UKF, CKF, PDA, NN, GNN, PDA, JPDA.
- $p(\mathbf{x}; \boldsymbol{\theta}) = \sum_{i=1}^{N} w^{(i)} \delta(\mathbf{x} - \mathbf{x}^{(i)})$ — Particle filters.

- What is a **suitable parameterization $p(\mathbf{x}; \theta)$ for MTF**?

Conjugate priors for filtering

- In MTF and MTT we can use conjugate priors.

Conjugate priors in multi-object filtering

A family of distributions, $\{p(\mathsf{x};\theta)\}_\theta$ is conjugate (to certain motion and measurement models) if

$$p(\mathsf{x}^{k-1}|\mathsf{Z}^{k-1}) = p(\mathsf{x}^{k-1};\theta_{k-1|k-1})$$
$$\Rightarrow \exists\theta_{k|k-1},\theta_{k|k} : \begin{cases} p(\mathsf{x}^{k}|\mathsf{Z}^{k-1}) = p(\mathsf{x}^{k};\theta_{k|k-1}) \\ p(\mathsf{x}^{k}|\mathsf{Z}^{k}) = p(\mathsf{x}^{k};\theta_{k|k}) \end{cases}$$

- **Example:** the family of Gaussian densities is conjugate to linear and Gaussian state space models.
- **Note 1:** conjugate families seem to enable exact filtering.
- **Note 2:** computing $p(\mathsf{x}^{k}|\mathsf{Z}^{k})$ may still be intractable.

Poisson multi-Bernoulli mixture (PMBM)

A PMBM is a conjugate prior to standard models for MTF.

That is, if $p(\mathbf{x}^1)$ is a PMBM, so are all future densities $p(\mathbf{x}^k|\mathbf{Z}^{k-1})$ and $p(\mathbf{x}^k|\mathbf{Z}^k)$.

This holds for both point and extended targets.

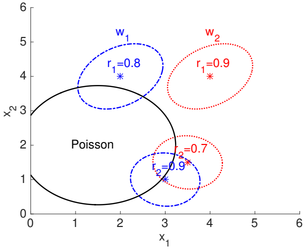
Definition: PMBM

$\mathbf{x}$ is a PMBM RFS if

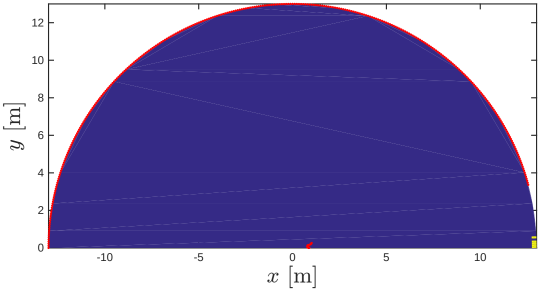
$$\mathbf{x} = \mathbf{x}_{\mathrm{p}} \uplus \mathbf{x}_{\mathrm{mbm}},$$

where $\mathbf{x}_{\mathrm{p}}$ a Poisson RFS and $\mathbf{x}_{\mathrm{mbm}}$ is an MBM RFS.

- **Poisson**: set of undetected targets.
- **MBM:** set of detected targets.



A Poisson process to model undetected targets

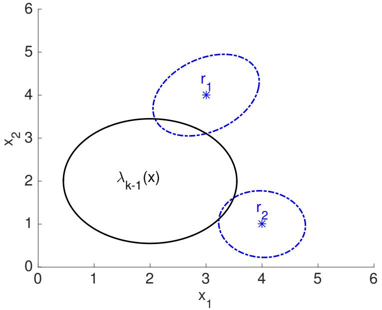


- Poisson intensity increases in occluded areas where we may have undetected objects.

PMB prediction

- Let us illustrate the prediction and update for a PMB.
- **Prediction events:**

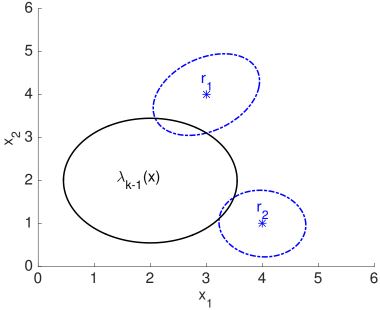
Previous posterior, $p(\mathbf{x}^{k-1} | \mathbf{Z}^{k-1})$:



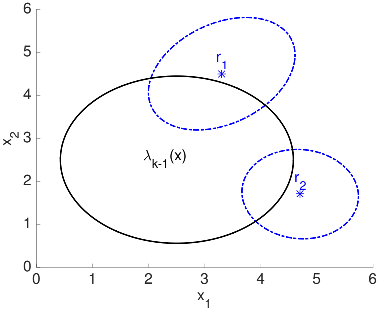
PMB prediction

- Let us illustrate the prediction and update for a PMB.
- **Prediction events:**
  1. existing targets may move,

Previous posterior, $p(\mathbf{x}^{k-1}|\mathbf{Z}^{k-1})$:



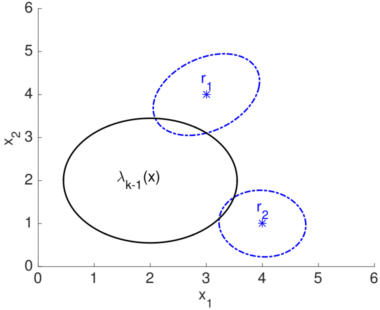
Predicted density, $p(\mathbf{x}^{k}|\mathbf{Z}^{k-1})$:



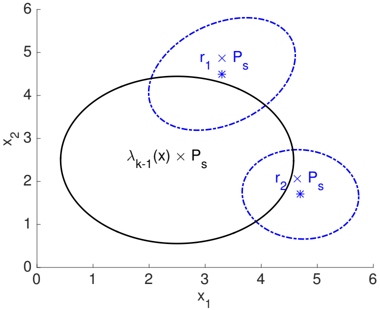
PMB prediction

- Let us illustrate the prediction and update for a PMB.
- **Prediction events:**
  1. existing targets may move,
  2. or die (disappear),

Previous posterior, $p(\mathbf{x}^{k-1}|\mathbf{Z}^{k-1})$:



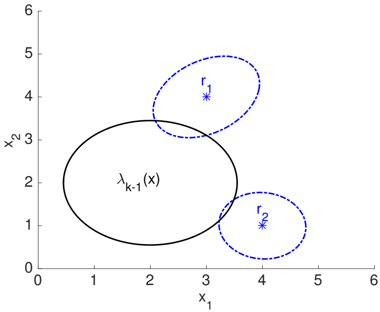
Predicted density, $p(\mathbf{x}^{k}|\mathbf{Z}^{k-1})$:



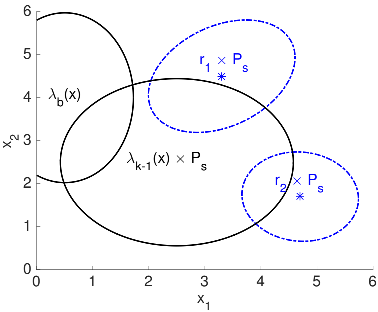
PMB prediction

- Let us illustrate the prediction and update for a PMB.
- **Prediction events:**
  1. existing targets may move,
  2. or die (disappear),
  3. new targets may arrive: Poisson birth process, $\lambda_b(x)$.

Previous posterior, $p(\mathbf{x}^{k-1} | \mathbf{Z}^{k-1})$:



Predicted density, $p(\mathbf{x}^{k} | \mathbf{Z}^{k-1})$:



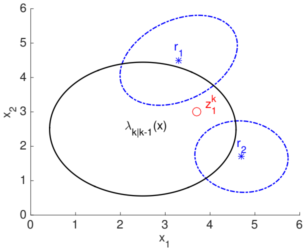
PMB update

- The MB contains Bernoulli components that we often call **tracks**. Suppose we have n tracks ($n = 2$ in illustrations).

- **Update step:**
  For each measurements we have $n + 1$ hypotheses:

  1. For $i = 1, 2, \ldots, n$: measurement was **generated by track** i. $\Rightarrow r_i = 1$.
  2. Measurement was generated by **previously undetected target or clutter**.

     $\Rightarrow$ create a new track!



- **Note 1:** tracks are initiated based on measurements.

- **Note 2:** posterior is a mixture, due to the many different hypotheses.

PMBM recursions

- Number of hypotheses grows quickly with each update.

- **Basic idea:** use pruning and merging to reduce ♯hypotheses. Two versions:

  1. PMBM filters: reduce ♯hypotheses to a manageable number.
  2. PMB filters: reduce ♯hypotheses to one.

- **Novel techniques:**

  1. **variational merging**: enables improved merging across different tracks (similar to SJPDA).
  2. **recycling**: low-probability Bernoullis approximated as Poisson.

- Why use a **conjugate prior**? ($\not\Rightarrow$ tractability)

  1. PMBMs can approximate true posterior arbitrarily well by maintaining many hypotheses.
  2. The best PMB is better than the best cluster processes in Kullback-Leibler sense.
     (Current proof only valid for point targets).

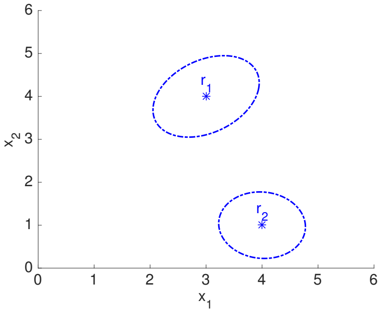
Include labels?

- Let us try to understand the relation to a labeled approach.
- Can we **augment our state with an implicit (unobservable) label**? Yes, this does not change the above results.
- However, to obtain reasonable labels we should make sure that:
  1. birth process generates unique labels,
  2. labels do not change with time.
- Birth process?
  1. labeled Poisson $\Rightarrow$ may yield total mixed labeling already at birth,
  2. **labeled multi-Bernoulli** can avoid this problem.

- Both (labeled) MBM and (labeled) PMBM are **conjugate priors** for this birth process.
  Why is **MBM a conjugate prior?**

- Let us illustrate the prediction and update for a MB.
- **Prediction events:**

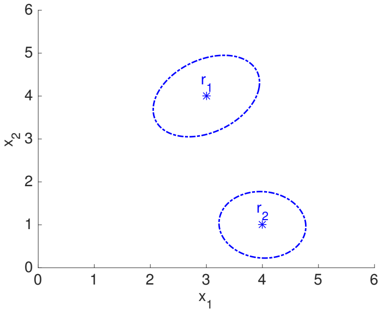
Previous posterior, $p(\mathbf{x}^{k-1}|\mathbf{Z}^{k-1})$:



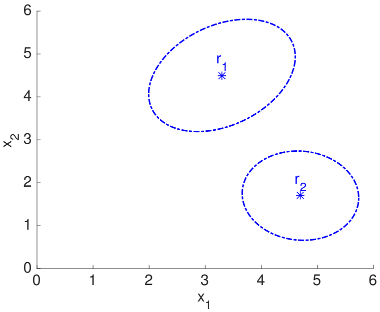
MB prediction

- Let us illustrate the prediction and update for a MB.
- **Prediction events:**
  1. existing targets may move,

Previous posterior, $p(\mathbf{x}^{k-1} | \mathbf{Z}^{k-1})$:



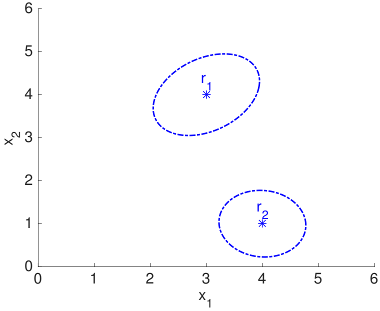
Predicted density, $p(\mathbf{x}^{k} | \mathbf{Z}^{k-1})$:



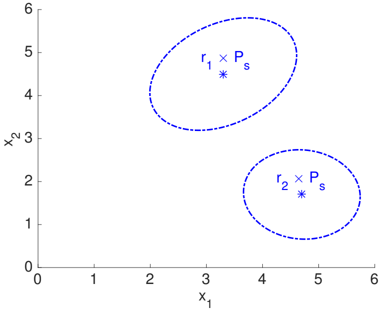
MB prediction

- Let us illustrate the prediction and update for a MB.
- **Prediction events:**
  1. existing targets may move,
  2. or die (disappear),

Previous posterior, $p(\mathbf{x}^{k-1}|\mathbf{Z}^{k-1})$:



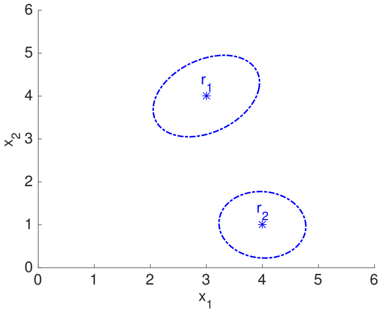
Predicted density, $p(\mathbf{x}^{k}|\mathbf{Z}^{k-1})$:



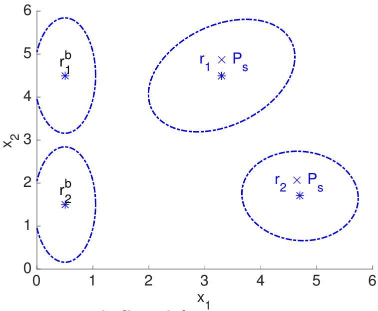
MB prediction

- Let us illustrate the prediction and update for a MB.
- **Prediction events:**
  1. existing targets may move,
  2. or die (disappear),
  3. new targets may arrive: MB birth process.

Previous posterior, $p(\mathbf{x}^{k-1}|\mathbf{Z}^{k-1})$:



Predicted density, $p(\mathbf{x}^{k}|\mathbf{Z}^{k-1})$:



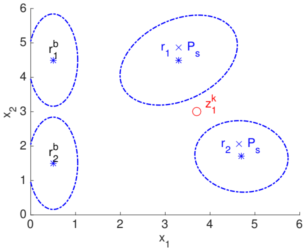
- The **MB birth creates new tracks** at pre-defined locations.

LMBM update

- The MB contains Bernoulli components that we often call **tracks**. Suppose we have n tracks ($n = 2$ in illustrations).

- **Update step:**
  For each measurements we have $n + 1$ hypotheses:
  1. For $i = 1, 2, \dots, n$: measurement was **generated by track** i. $\Rightarrow r_i = 1$.
  2. Measurement is **clutter**.

  $\Rightarrow$ no new tracks!



- **Note 1:** update is identical to PMBM with $\lambda(x) = 0$.
- **Note 2:** no new tracks during update.

δ-GLMB

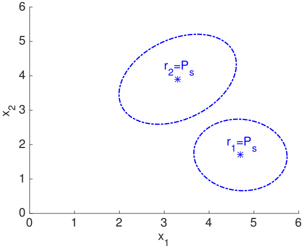
- As you have seen, labels can be handled using LMBM, which is essentially a special case of a PMBM.
- However, the standard **conjugate prior for labelled RFS** is the **δ-GLMB distribution**.
- Yet another conjugate prior? Not really.
- The δ-GLMB is a special type of LMBM where all existence probabilities are 0 or 1.

δ-GLMB

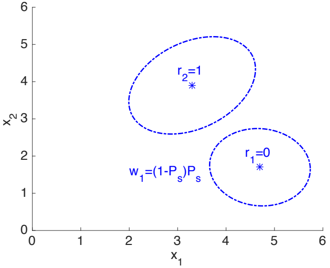
- How can we restrict the existence probabilities to $r \in \{0, 1\}$? By creating **more hypotheses!**

- Suppose posterior at time $k-1$ and is an LMB with $r = 1$ for all Bernoulli components.
- After prediction, their existence probabilities are P_s, but we can also express this using 2^n hypotheses with $r_{ij} \in \{0, 1\}$:

- An LMBM representation:



- A δ-GLMB representation:



δ-GLMB and LMB algorithms

- Two popular algorithms for labeled MTF are:

  1. **the δ-GLMB filter**: maintains several/many hypotheses; all correspond to LMBs with $r \in \{0, 1\}$.
  2. **the LMB filter**: reduces the δ-GLMB posterior to a single LMB with general existence probabilities.

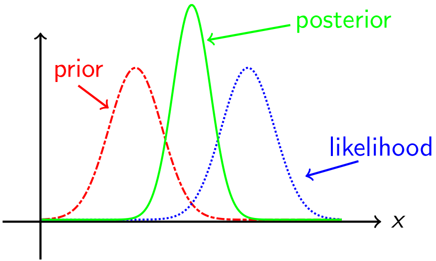
Conjugate priors: summary

- PMBM, LMBM and δ-GLMB are all conjugate priors for MTF.
- Conjugate priors are useful to develop powerful algorithms.
- Using a Poisson birth process and a PMBM posterior has several advantages:
  1. tracks are initiated by measurements,
  2. fewer hypotheses,
  3. enables recycling (approximating low-probability tracks as Poisson).

Bayesian statistics and outline

In Bayesian statistics:

- we compute **posterior densities** of, x,
- posterior density summarizes **what we know** about x,
- Very useful! E.g., can compute **optimal estimates**.



Outline:

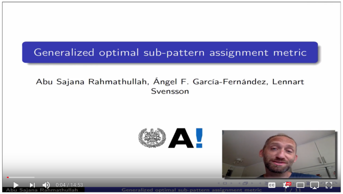
1) Sets of trajectories: suitable x in MTT and MTF?
   Which are our quantities of interest?
2) Conjugate prior densities: reasonable priors and likelihoods to obtain tractable posteriors?
3) **Metrics**: how can we measure performance in MTT and MTF?

- **Metrics** are useful to
  1. evaluate performance of algorithms,
  2. derive optimal estimators.

  We have developed metrics for MTF and MTT.

1) **Generalized OSPA**: a metric for MTF, i.e., a metric between sets of targets.

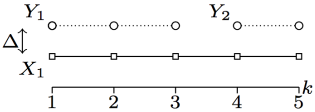
- GOSPA paper received best paper award at Fusion, 2017,
- YouTube video where the paper is carefully explained.



Generalized optimal sub-pattern assignment metric (GOSPA)

2) **A metric for MTT**, i.e., a metric between sets of trajectories.

- Trajectory version of GOSPA that also penalizes "track switches".



- What is GOSPA?
  - A metric on sets of targets, useful to evaluate performance and design estimators.
  - An alternative to OSPA!

Informal definition

$$\text{GOSPA} = \text{localisation error} + \frac{c}{2}\left(\sharp\text{missed targets} + \sharp\text{false targets}\right)$$

- Why GOSPA instead of OSPA?
  - We often want few false and missed targets.
    $\rightsquigarrow$ GOSPA measures this, OSPA doesn't

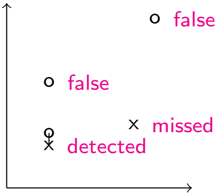


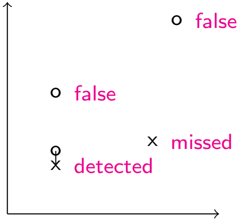
Figure: Detected, missed and false targets

x-truth, o-estimate

How to compute GOSPA?

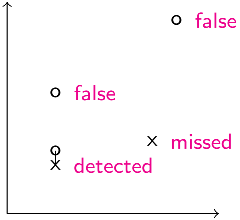
- Computing GOSPA ($\alpha = 2, p = 1$):

  1) Find optimal assignments between sets.
     Remark 1: pairs are left unassigned if $d(x, y) > c$.
     Remark 2: we refer to unassigned elements as false/missed targets.
  2) Assigned pairs cost $d(x, y)$.
  3) Unassigned elements cost $c/2$.



How to compute GOSPA?

- Computing GOSPA ($\alpha = 2, p = 1$):
  1) Find optimal assignments between sets.
     Remark 1: pairs are left unassigned if $d(x, y) > c$.
     Remark 2: we refer to unassigned elements as false/missed targets.
  2) Assigned pairs cost $d(x, y)$.
  3) Unassigned elements cost $c/2$.



Formal definition, GOSPA, $\alpha = 2$

$$\left[\min_{\gamma\in\Gamma}\left(\sum_{(i,j)\in\gamma} d(x_i, y_j)^p + \frac{c^p}{2}\left(\underbrace{|X| - |\gamma|}_{\sharp\text{missed}} + \underbrace{|Y| - |\gamma|}_{\sharp\textit{false}}\right)\right)\right]^{\frac{1}{p}}$$

where X : set of targets, Y : set of estimates and Γ : set of possible assignments.

Application of GOSPA

- The GOSPA metric is a sum of three terms:

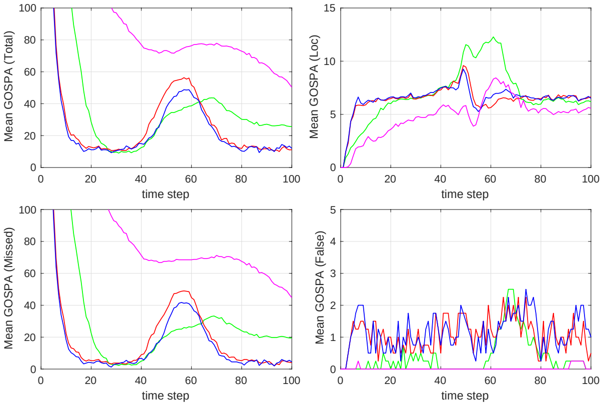
$$\text{GOSPA} = \text{local. error} + \frac{c}{2}\,(\sharp\text{missed targets} + \sharp\text{false targets})$$

- In [Xia2017][2], the performance of different multi-Bernoulli filters evaluated using GOSPA.
  - δ generalised labelled multi-Bernoulli (δ GLMB)
  - Labelled multi-Bernoulli (LMB)
  - Poisson multi-Bernoulli mixture (PMBM)
  - Poisson multi-Bernoulli (PMB)
- Scenario
  - Challenging scenario involving six targets in close proximity at the mid-point of the simulation.

[2] Xia et. al, "Performance Evaluation of Multi-Bernoulli Conjugate Priors for Multi-Target Filtering", 20th Inter. Conf. on Information Fusion, July 2017.

GOSPA results for challenging scenario

- Performance of algorithms compared using GOSPA: localisation error, # missed and # false targets

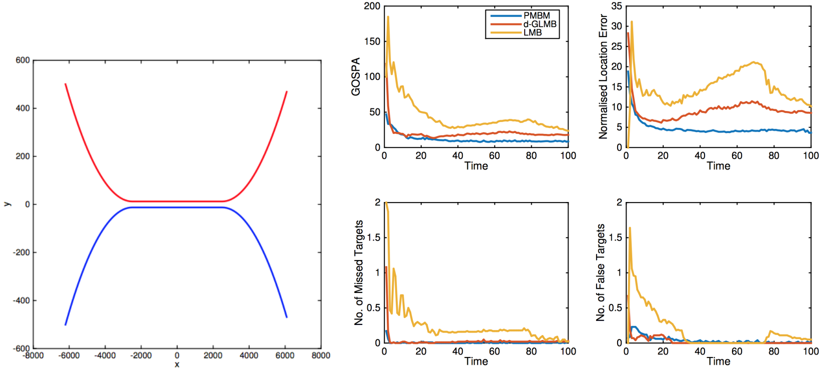


δ-GLMB (green), LMB (magenta) PMBM (red), PMB (Murty) (blue)

- GOSPA clarifies: most errors are due to missed targets!

Extended target tracking

- **Scenario**[3]**:** Two extended targets are well separated, but move closer to each other before they separate again.
- PMBM achieves the lowest GOSPA.
  The PMBM is much faster than δ-GLMB, but slower than LMB.



[3] Granström, K., et. al, 'Poisson multi-Bernoulli conjugate prior for multiple extended object estimation". arxiv.org/abs/1605.06311.

Further Reading I

Part I: sets of trajectories

- L. Svensson and M. Morelande,
  "Target tracking based on estimation of sets of trajectories"
  in *in Proc. 17th International Conference on Information Fusion*, July 2014.

- A. F. García-Fernández, L. Svensson and M. Morelande,
  "Multiple target tracking based on sets of trajectories"
  arXiv pre-print. [Online]. Available: arxiv.org/abs/1605.08163.

- A. F. García-Fernández and L. Svensson,
  "Trajectory probability hypothesis density filter"
  arXiv pre-print. [Online]. Available: arxiv.org/abs/1605.07264.

Further Reading II

Random finite sets (RFSs) and labelled RFSs

R. Mahler.
Statistical Multisource-Multitarget Information Fusion.
Artech House, Inc., 2007.

B. T. Vo and B. N. Vo,
"Labeled random finite sets and multi-object conjugate priors"
IEEE Transactions on Signal Processing, 61(13), 2013.

B. N. Vo, B. T. Vo and D. Phung,
"Labeled random finite sets and the Bayes multi-target tracking filter"
IEEE Transactions on Signal Processing, 62(24), 2014.

S. Reuter, B. T. Vo, B. N. Vo and K. Dietmayer,
"The labeled multi-Bernoulli filter"
IEEE Transactions on Signal Processing, 62(12), 2014.

Further Reading III

Part II: conjugate prior densities

J. L. Williams,

"Marginal multi-bernoulli filters: RFS derivation of MHT, JIPDA, and association-based member"

IEEE Transactions on Aerospace and Electronic Systems, 51(13), 2015.

J. L. Williams,

"An efficient, variational approximation of the best fitting multi-Bernoulli filter"

IEEE Transactions on Signal Processing, 63(1), 2015.

J. L. Williams,

"Hybrid Poisson and multi-Bernoulli filters"

in *in Proc. 15th International Conference on Information Fusion*, July 2012.

Further Reading IV

K. Granström, M. Fatemi and L. Svensson,
"Poisson multi-Bernoulli conjugate prior for multiple extended object estimation"
arXiv pre-print. [Online]. Available: arxiv.org/abs/1703.04264.

K. Granström, M. Fatemi and L. Svensson,
"Gamma Gaussian inverse-Wishart Poisson multi-Bernoulli filter for extended target tracking"
in *in Proc. 19th International Conference on Information Fusion*, 2016.

M. Fatemi, et al.,
"Poisson Multi-Bernoulli Mapping Using Gibbs Sampling"
IEEE Transactions on Signal Processing, 65(11), 2017.

- A. F. García-Fernández, J. Williams, K. Granström and L. Svensson,
  "Poisson multi-Bernoulli mixture filter: direct derivation and implementation"
  arXiv pre-print. [Online]. Available: arxiv.org/abs/1703.04264.

- Y. Xia, K. Granström, L. Svensson and A. F. García-Fernández,
  "Performance Evaluation of Multi-Bernoulli Conjugate Priors for Multi-Target Filtering"
  in *in Proc. 20th International Conference on Information Fusion*, July 2017.

Further Reading VI

Part III: metrics

- A. S. Rahmathullah, A. F. García-Fernández and L. Svensson,
  "Generalized optimal sub-pattern assignment metric"
  in *in Proc. 20th International Conference on Information Fusion*, July 2017.

- A. S. Rahmathullah, A. F. García-Fernández and L. Svensson,
  "A metric on the space of finite sets of trajectories for evaluation of multi-target tracking algorithms"
  arXiv pre-print. [Online]. Available: arxiv.org/abs/1605.01177.