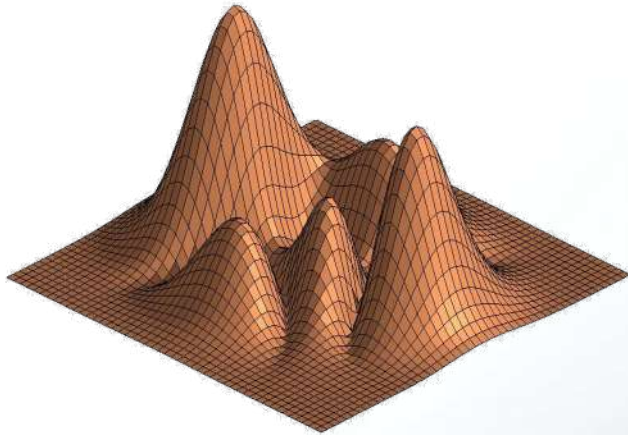


Inference with Multi-Object Hidden Markov Models



Ba-Ngu Vo <http://ba-ngu.vo-au.com>

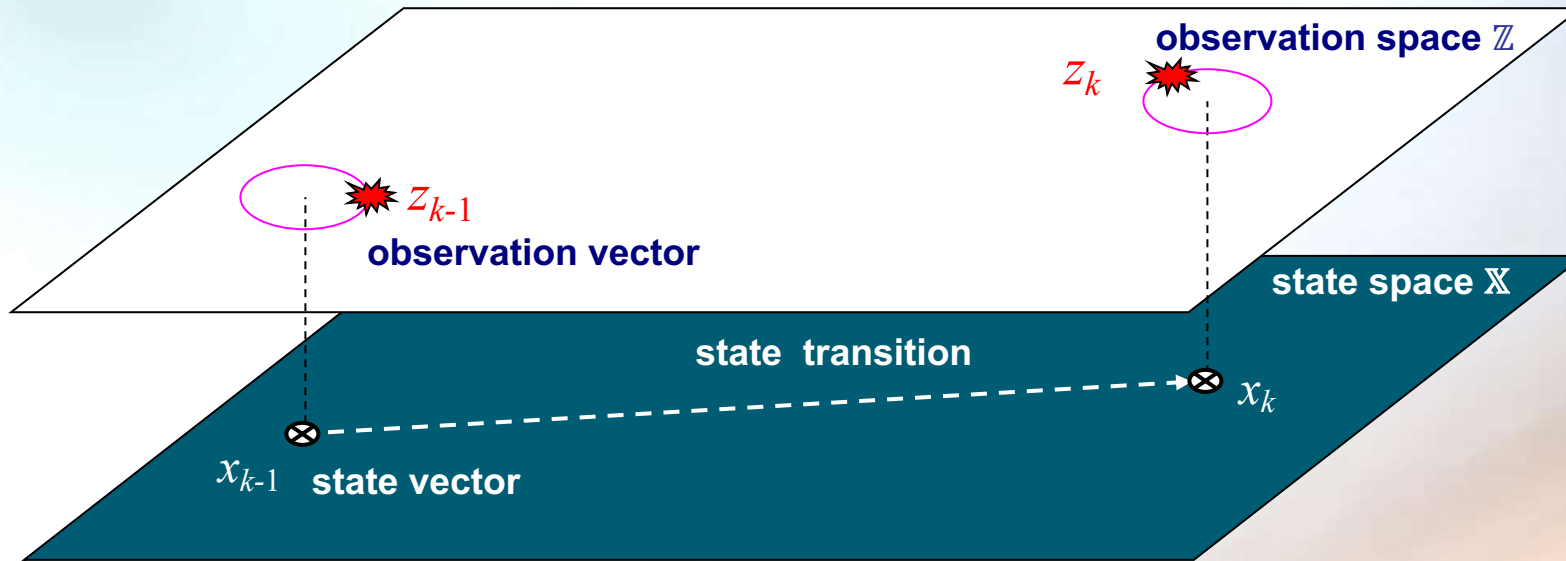
**Dept. Electrical & Computer Engineering
Curtin University
Perth, Australia**

Outline

- ▶ **Multi-Object Hidden Markov Model (HMM)**
- ▶ **Stochastic Geometry**
- ▶ **Inference for Multi-Object HMM**
- ▶ **Generalized Labeled Multi-Bernoulli**
- ▶ **Applications**
- ▶ **Conclusions**

Multi-Object Hidden Markov Model (HMM)

Hidden Markov Model (HMM) aka State Space Model of a Dynamical System



- ▶ State Estimation
- ▶ System ID
- ▶ Control

State equation

$$x_k = \Phi(x_{k-1}, u_{k-1}, n_k)$$



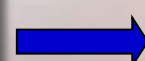
$$f_{k|k-1}(x_k | x_{k-1}, u_{k-1})$$

Markov transition density

control noise

Observation equation

$$z_k = \Psi(x_k, u_k, v_k)$$

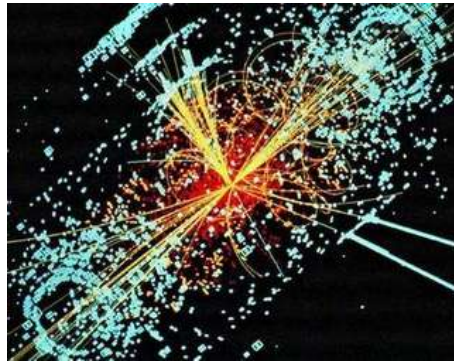


$$g_k(z_k | x_k, u_{k-1})$$

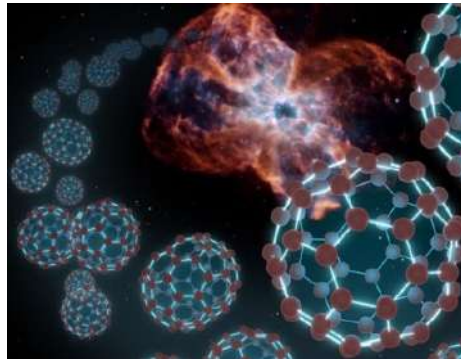
Observation likelihood

Multi-Object Hidden Markov Model (HMM)

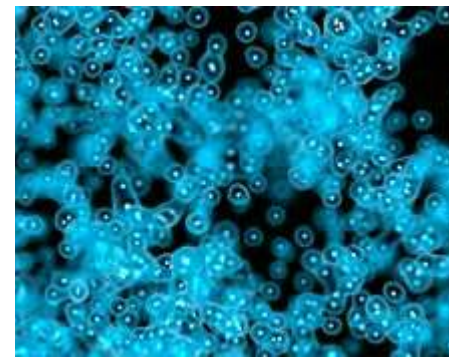
- What if the state & observation are not vectors but point patterns, i.e., finite sets?



Particles



Molecules



Cells



People



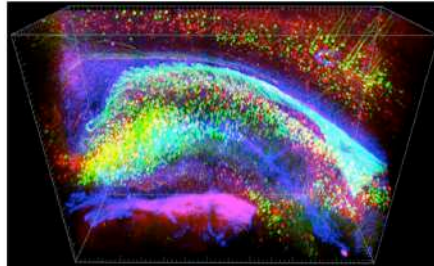
Orbital Debris



Astronomy

Multi-Object Hidden Markov Model (HMM)

► Point patterns in data science



Neuron point cloud



Tree locations

5	0	0	0	0	0	0	0	→	{{1,5}}
0	0	0	0	0	0	0	37	→	{{8,37}}
6	5	0	0	0	0	0	6	→	{{1,6}, {2,5}, {8,6}}
80	0	0	0	0	0	0	0	→	{{1,80}}
0	0	0	6	0	0	0	0	→	{{4,6}}
48	0	0	0	0	0	0	26	→	{{1,48}, {8,26}}

Sparse data



SIFT key points



Bag of words



Transaction record

- Multiple Instance (MI) Learning: Machine Learning for point patterns [Amores 13]

Stochastic Geometry

Bayesian: $p(X | \text{Data})$ Posterior PDF

including finite sets

Non-Bayesian: $p(\text{Data} | X)$ Data likelihood function

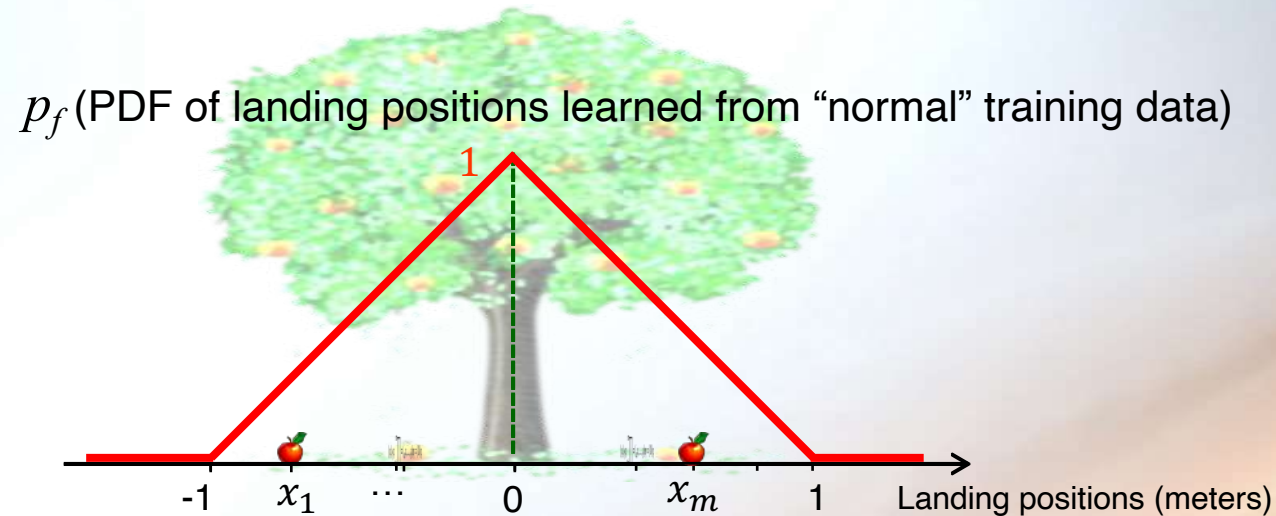
► What's the big deal about working with sets?

Most widely adopted practice (in engineering/computer science):

$X = \{x_1, \dots, x_m\}$ $\Rightarrow$ $p(X) = p(x_1, \dots, x_m)$
Finite Set Probability Density Function (PDF)

Stochastic Geometry

► What's the big deal about working with sets?



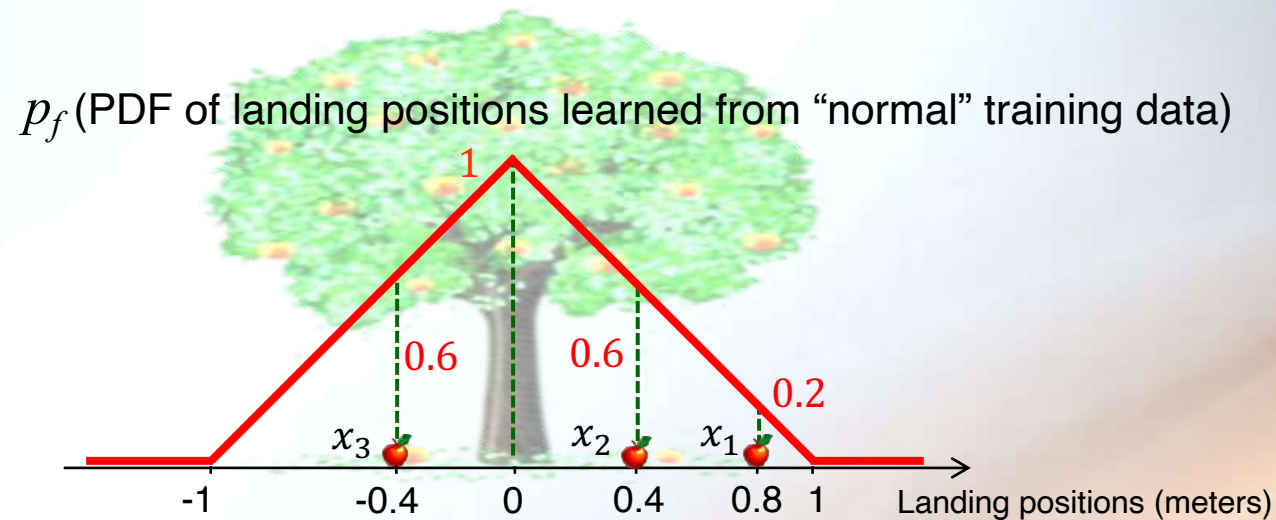
- Apples land on the ground independently from each other

$$p(X) = p(x_1, \dots, x_m) = \prod_{i=1}^m p_f(x_i)$$

- Daily landing patterns are independent from each other
- **Novelty Detection:** find unlikely daily landing patterns

Stochastic Geometry

► What's the big deal about working with sets?



Day 1: x_1

$$p(x_1) = p_f(x_1) = 0.2,$$

Day 2: x_2, x_3

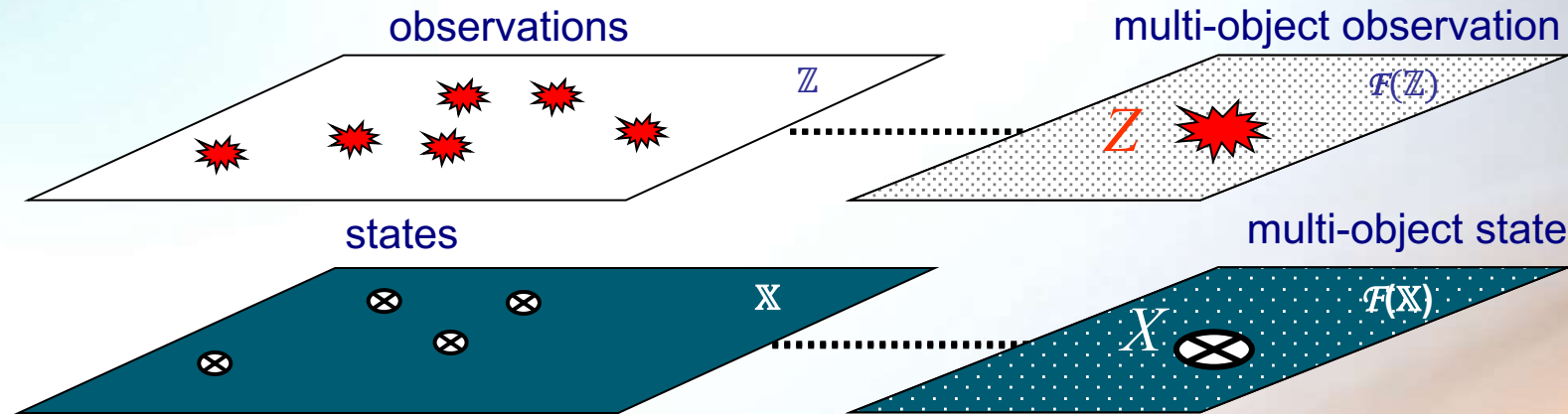
$$p(x_2, x_3) = p_f(x_2) p_f(x_3) = 0.36$$

- Q: Which pattern is less likely? Ans: day 1 less likely than day 2.
- Change unit of measurement from m to cm and ...

$$p(x_1) = 0.002 > p(x_2, x_3) = 0.000036$$

Stochastic Geometry

- ▶ **Multi-Object HMM:** HMM where the state is a finite set - multi-object state



- ▶ **Model multi-object state as a random finite set (RFS) ...**

- Needs: Markov transition density & observation likelihood for finite-set-valued **state**
- Can't treat a (random) finite set as if it were a (random) vector
- PDF of random finite set not the same as PDF of random vector
- **Need PDFs (+ suitable notion of density & integration) for Random Finite Sets**

Stochastic Geometry

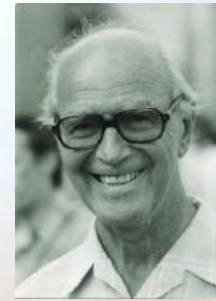
- ▶ Mathematical tools for dealing with Random Sets
- ▶ Foundation (1960s-1970s): mostly due to independent work by **Matheron** and **Kendall**, both of whom gave credits to earlier work by **Choquet**



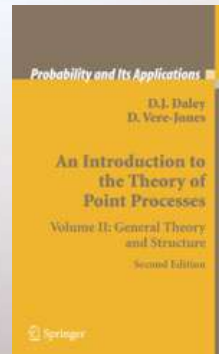
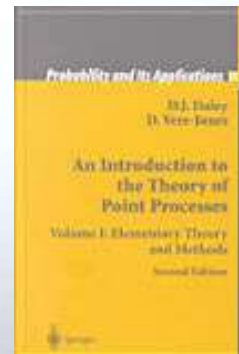
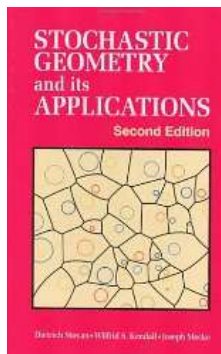
G. Matheron (1930-2000)



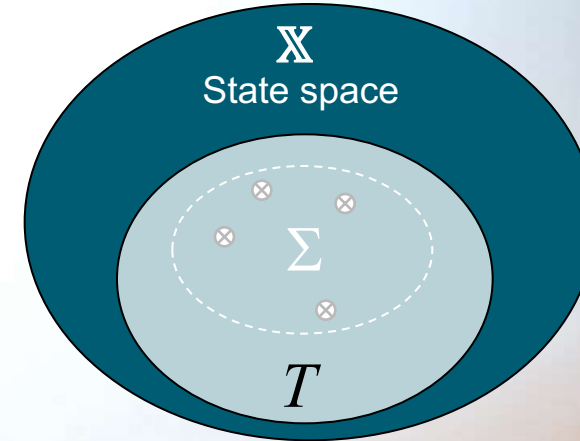
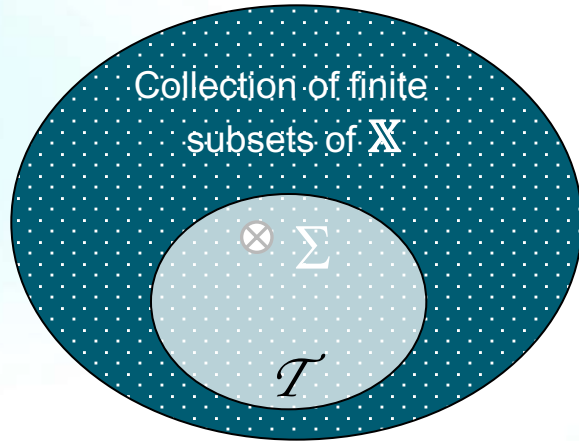
D. Kendall (1918-2007)



G. Choquet (1915-2006)



Stochastic Geometry



Probability distribution of Σ
 $P_{\Sigma}(\mathcal{T}) = P(\Sigma \in \mathcal{T}), \mathcal{T} \subseteq \mathcal{F}(\mathbb{X})$

Choquet
(1968)

Belief "distribution" of Σ
 $\beta_{\Sigma}(T) = P(\Sigma \subseteq T), T \subseteq \mathbb{X}$

Point Process Theory (1950-1960's)

Mahler's Finite Set Statistics (1994)

Probability density of Σ
 $p_{\Sigma} : \mathcal{F}(\mathbb{X}) \rightarrow [0, \infty)$
 $P_{\Sigma}(\mathcal{T}) = \int_{\mathcal{T}} p_{\Sigma}(X) \mu(dX)$

VSD
(2005)

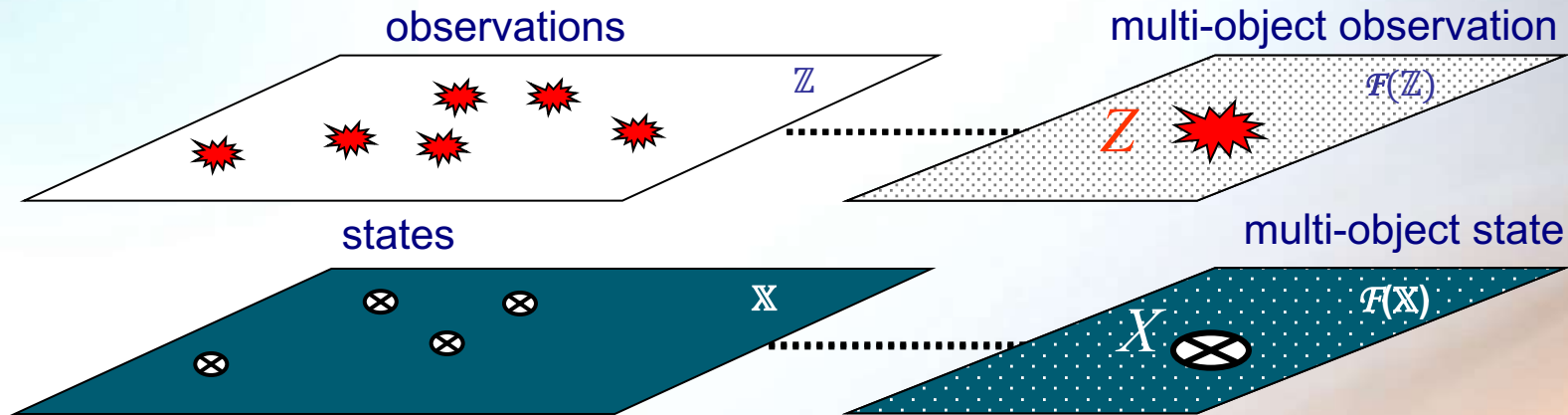
Belief "density" of Σ
 $f_{\Sigma} : \mathcal{F}(\mathbb{X}) \rightarrow [0, \infty)$
 $\beta_{\Sigma}(T) = \int_T f_{\Sigma}(X) \delta X$

Conventional integral

Set integral

Inference for Multi-Object HMM

► Multi-Object HMM: Finite-set-valued HMM



► Inference for Multi-Object HMM: State Estimation on the space $\mathcal{F}(\mathbb{X})$ of finite sets of $\mathbb{X}$

► Fundamental difference from classical dynamical system:

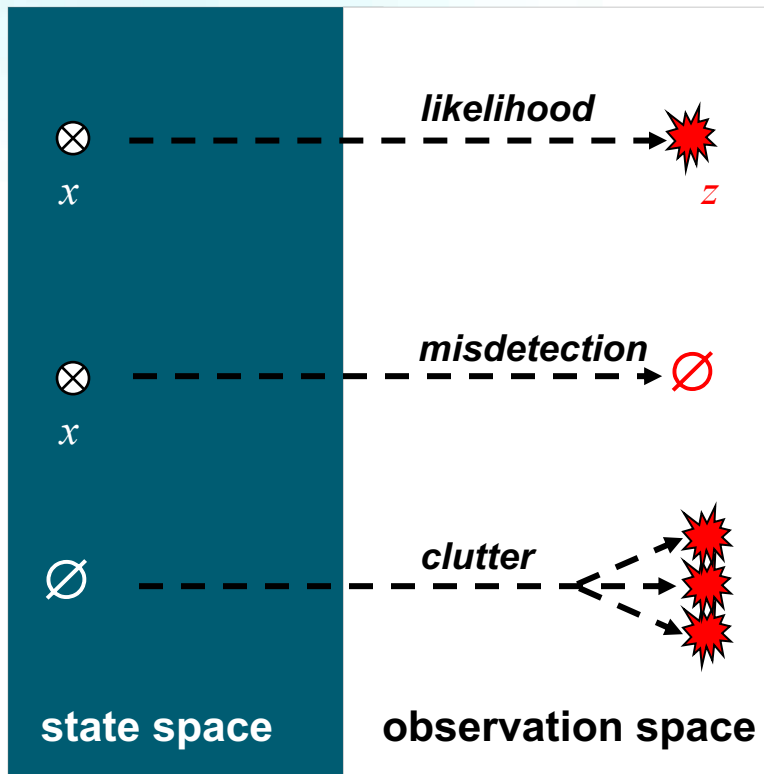
- Random time-varying number of states and measurements
- False negatives, false positives, association uncertainty
- **Much more challenging computationally!**

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

R. Mahler, *Advances in Statistical Multisource-Multitarget Information Fusion*, Artech House, 2014.

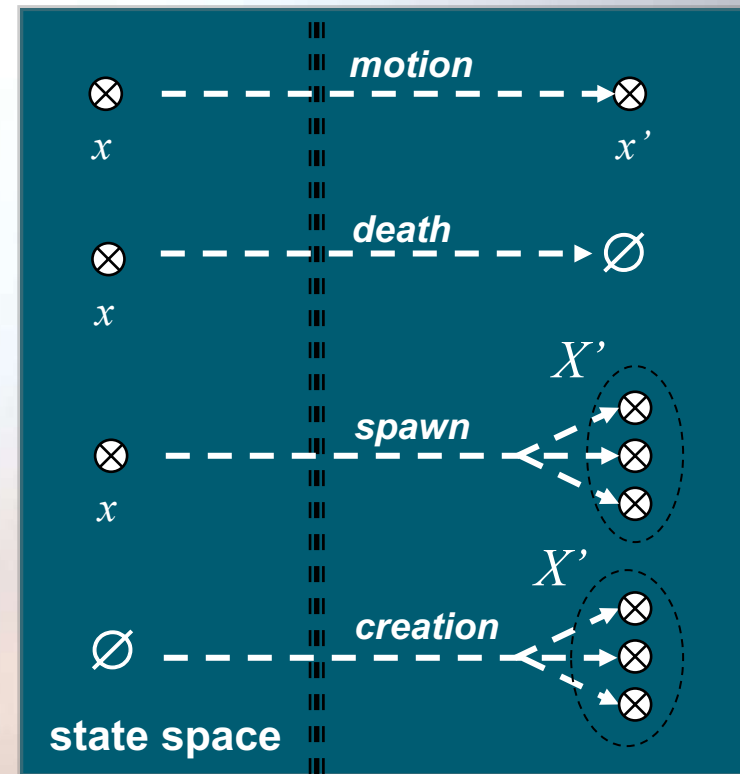
Inference for Multi-Object HMM

A Standard Multi-Object HMM



Observation likelihood

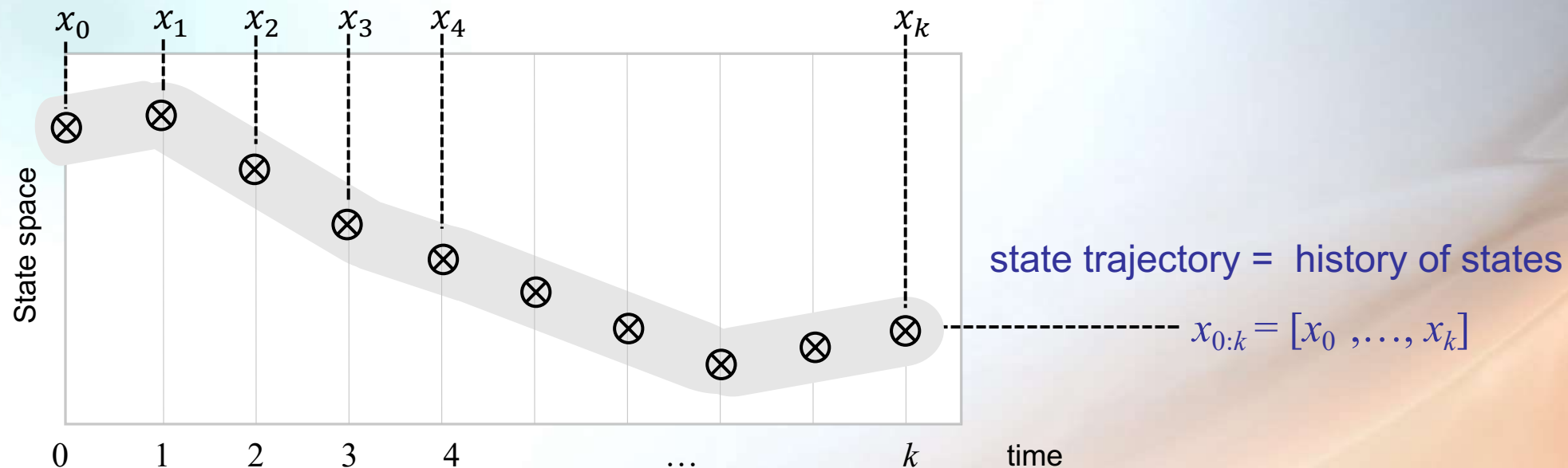
$$g_k(Z_k | X_k)$$



State transition kernel

$$f_{k|k-1}(X_k | X_{k-1})$$

Inference for Multi-Object HMM



- ▶ **State Estimation:** estimate state trajectory
- ▶ **Online Operation:** need **fixed complexity per time step** to be useful in practice
- ▶ **Filtering:** $\hat{x}_0, \dots, \hat{x}_k$, **suitable** for **online** – Kalman & particle filters
- ▶ **Smoothing:** $\hat{x}_{0:k}$, **not suitable** for **online** – Kalman & particle smoothers

Inference for Multi-Object HMM

smoothing-while-filtering

Bayes smoother

$$\frac{g_k(z_k | x_k) f_{k|k-1}(x_k | x_{k-1}) p_{0:k-1}(x_{0:k-1} | z_{1:k-1})}{\int g_k(z_k | x_k) f_{k|k-1}(x_k | x_{k-1}) p_{0:k-1}(x_{0:k-1} | z_{1:k-1}) dx_{0:k}}$$

$$\dots \longrightarrow p_{0:k-1}(x_{0:k-1} | z_{1:k-1}) \longrightarrow p_{0:k}(x_{0:k} | z_{1:k}) \longrightarrow \dots$$

- Complexity per time step increases with time - **not suitable** for **online**
- Smooth over a fix-length moving window - fixed complexity per time step

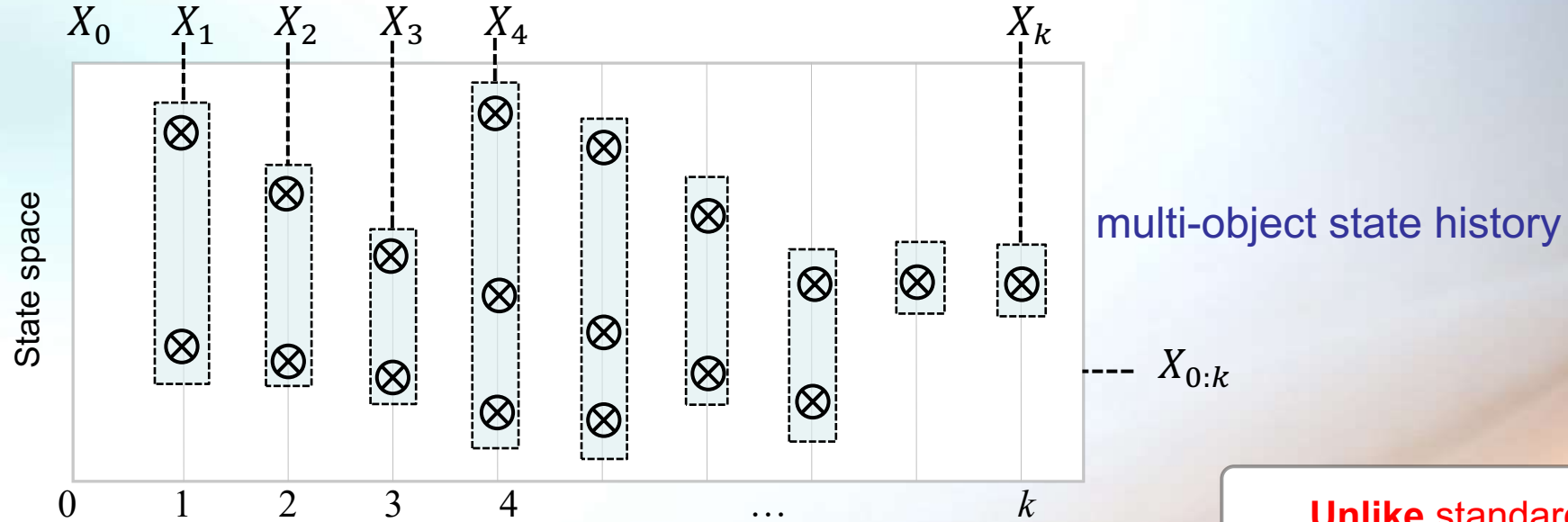
Bayes filter

$$\begin{array}{c} \int p_{k-1}(x_{k-1} | z_{1:k-1}) f_{k|k-1}(x_k | x_{k-1}) dx_{k-1} \\ \hline \dots \longrightarrow p_{k-1}(x_{k-1} | z_{1:k-1}) \xrightarrow{\text{prediction}} p_{k|k-1}(x_k | z_{1:k-1}) \xrightarrow{\text{data-update}} p_k(x_k | z_{1:k}) \longrightarrow \dots \end{array}$$

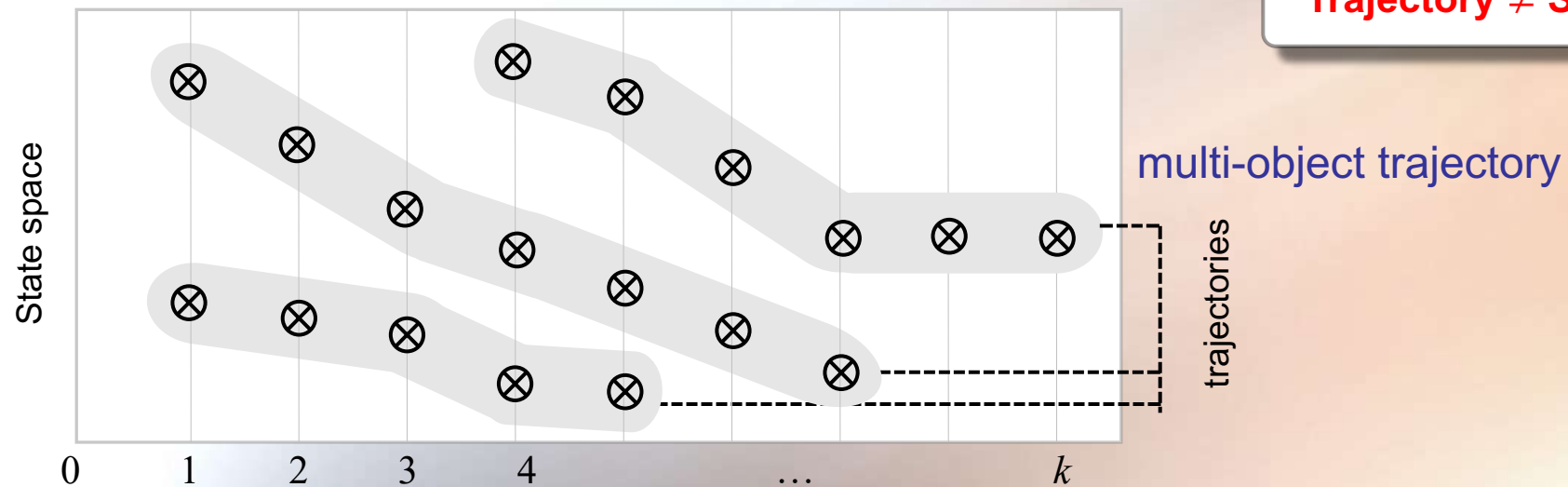
$$\frac{g_k(z_k | x_k) p_{k|k-1}(x_k | z_{1:k-1})}{\int g_k(z_k | x_k) p_{k|k-1}(x_k | z_{1:k-1}) dx_k}$$

- Fixed complexity per time step - **suitable** for **online**, widely used

Inference for Multi-Object HMM

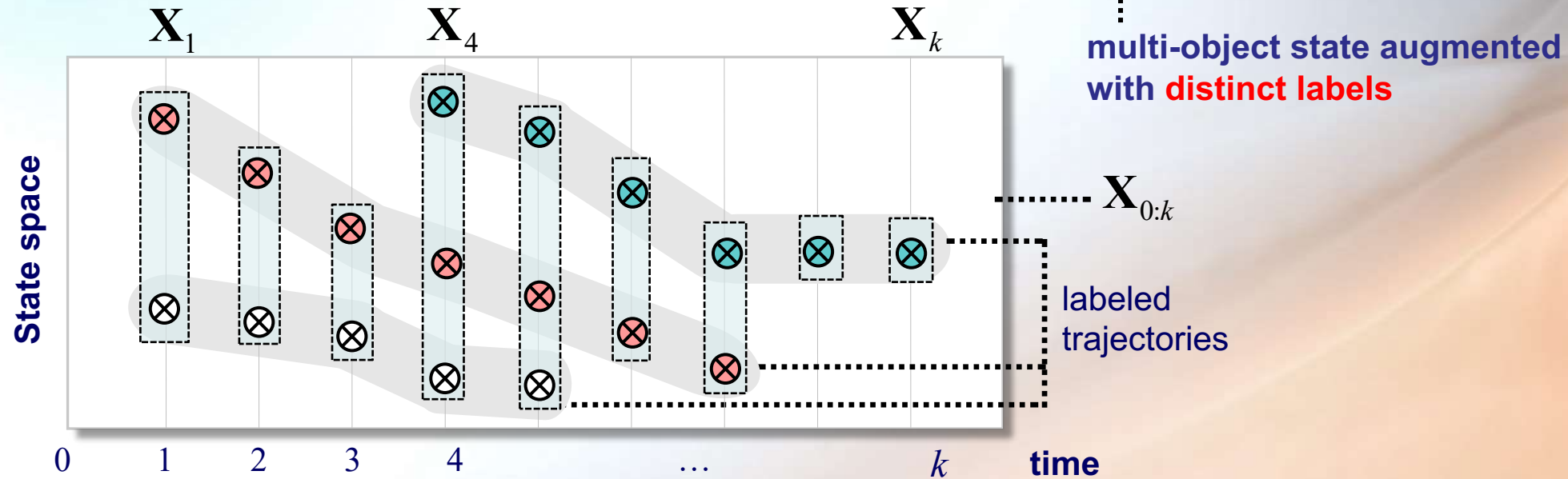


Unlike standard HMM
Trajectory $\neq$ State History



Inference for Multi-Object HMM

- ▶ multi-object trajectory = history of labeled multi-object states



- ▶ Labels provides multi-object trajectory estimate (even from a single scan)

- filtering with labeled multi-object states: $\hat{X}_{0,\dots,t}$
- smoothing with labeled multi-object states: $\hat{X}_{0:k}$

Inference for Multi-Object HMM

- ▶ Labels provides multi-object trajectory estimate (even from a single scan)

- filtering with labeled multi-object states: $\hat{\mathbf{X}}_0, \dots, \hat{\mathbf{X}}_k$
- smoothing with labeled multi-object states: $\hat{\mathbf{X}}_{0:k}$

- ▶ Labels admits closure under posterior truncation

multi-object likelihood

$$g_k(Z|\mathbf{X}) \propto \sum_{\theta} \left[\Psi_{Z,k}^{(\theta(\mathcal{L}(\cdot)))}(\cdot) \right]^{\mathbf{X}}$$

Vo & Vo "Labeled RFSs and Multi-Object Conjugate Priors," IEEE Trans. SP, 61(13): 3460-3475, 2013

- **Each term** of the multi-object likelihood is **symmetric**
- **Truncated** labeled posterior/filtering density is a **function of sets**
- ▶ **2 birds with one stone:** provides **trajectories & closure under truncation**

Generalized Labeled Multi-Bernoulli (GLMB)

► **GLMB filter**: Multi-object Analogue of Kalman Smoother/Filter

► **Closure under Bayes recursion** - Analytic solutions to:

- Multi-object Bayes Posterior Recursion (Smoothing-while-filtering)

$$\pi_{0:k}(\mathbf{X}_{0:k} | Z_{1:k}) \propto g_k(Z_k | \mathbf{X}_k) \mathbf{f}_{k|k-1}(\mathbf{X}_k | \mathbf{X}_{k-1}) \pi_{0:k-1}(\mathbf{X}_{0:k-1} | Z_{1:k-1})$$

estimated **labeled** multi-object state history $\widehat{\mathbf{X}}_{0:k}$

Vo & Vo "A Multi-Scan Labeled RFS Model for Multi-object State Estimation," IEEE Trans. SP, 67(19):4948-4963, 2019

- Multi-object Bayes Filtering Recursion

$$\begin{aligned} \pi_{k|k-1}(\mathbf{X}_k) &= \int \mathbf{f}_{k|k-1}(\mathbf{X}_k | \mathbf{X}) \pi_{k-1}(\mathbf{X}) \delta \mathbf{X}, \\ \pi_k(\mathbf{X}_k) &= \frac{g_k(Z_k | \mathbf{X}_k) \pi_{k|k-1}(\mathbf{X}_k)}{\int g_k(Z_k | \mathbf{X}) \pi_{k|k-1}(\mathbf{X}) \delta \mathbf{X}}, \end{aligned}$$

estimated **labeled** multi-object states $\widehat{\mathbf{X}}_0, \dots, \widehat{\mathbf{X}}_k$ forms a set of tracks

Vo & Vo "Labeled RFSs and Multi-Object Conjugate Priors," IEEE Trans. SP, 61(13): 3460-3475, 2013

Generalized Labeled Multi-Bernoulli (GLMB)

- ▶ GLMB density - multi-object analogue of exponential mixture:

$$\pi(\mathbf{X}) = \underbrace{\Delta(\mathbf{X})}_{\delta_{|\mathbf{X}|}(|\mathcal{L}(\mathbf{X})|)} \sum_{\xi \in \Xi} \underbrace{w^{(\xi)}(\mathcal{L}(\mathbf{X}))}_{\text{labels of } \mathbf{X}} \underbrace{\left[p^{(\xi)} \right]^{\mathbf{X}}}_{\prod_{\mathbf{x} \in \mathbf{X}} p^{(\xi)}(\mathbf{x})}$$

distinct label indicator associations history multi-object exponential

- **Weight Normalization**

$$\sum_{L \subseteq \mathbb{L}} \sum_{\xi \in \Xi} w^{(\xi)}(L) = 1 \quad \int p^{(\xi)}(x, \ell) dx = 1$$

- **Cardinality Distribution & 1st moment**

$$\rho(n) = \sum_{\xi \in \Xi} \sum_{L \subseteq \mathbb{L}} \delta_n(|L|) w^{(\xi)}(L)$$

$$v(x, \ell) = \sum_{\xi \in \Xi} p^{(\xi)}(x, \ell) \sum_{L \subseteq \mathbb{L}} 1_L(\ell) w^{(\xi)}(L)$$

Generalized Labeled Multi-Bernoulli (GLMB)

Labeled Multi-Object Density Approximations

► Truncation of GLMB

- Any truncation of a GLMB is a GLMB – closure under truncation
- L1-norm of truncation error can be computed analytically
- Minimum L1-norm truncation error: truncate smallest weights

Vo et. al. "Labeled RFS and the Bayes Multi-Target Tracking Filter," IEEE Trans. SP, 62(24):6554-6567, 2014

► Approximation of GLMB by a 1-term GLMB (or LMB) with

- Same PHD & same Cardinality distribution

Reuter et. al. "The labelled multi-Bernoulli filter," IEEE Trans. SP, 62(12):3246-3260, 2014

► Approximation of labeled multi-object density by a GLMB with:

- Same PHD & same Cardinality distribution
- Minimal Kullback-Leibler divergence

Papi et. al., "GLMB approximation of Multi-object densities," IEEE Trans. SP, 63(20):5487-5497, 2015

Generalized Labeled Multi-Bernoulli (GLMB)

- ▶ GLMB Filtering/Posterior sum grows in no. terms, requires reduction of terms
- ▶ Truncate terms with smallest weights $\Rightarrow$ minimum L1-norm truncation error
- ▶ **Implementation:** How to truncate without exhaustive enumeration of the terms?

▶ Filter implementation:

- K-shortest path prediction & ranked assignment update (cubic in no. detections)
- Gibbs sampling + Joint Prediction & Update (linear in no. detections)

Vo et. al. "Labeled RFS and the Bayes Multi-Target Tracking Filter," IEEE Trans. SP, 62(24):6554-6567, 2014.

Vo et. al. "An Efficient Implementation of the GLMB Filter," IEEE Trans. SP, 65(8):1975-1987, 2017.

▶ Multi-sensor filter implementation:

- Multi-dimensional assignment (NP-Hard): Gibbs sampling (linear in total no. of detections)

Vo et. al. "Multi-sensor multi-object tracking with the GLMB filter," IEEE Trans. SP, 67 (23):5952-5967, 2019.

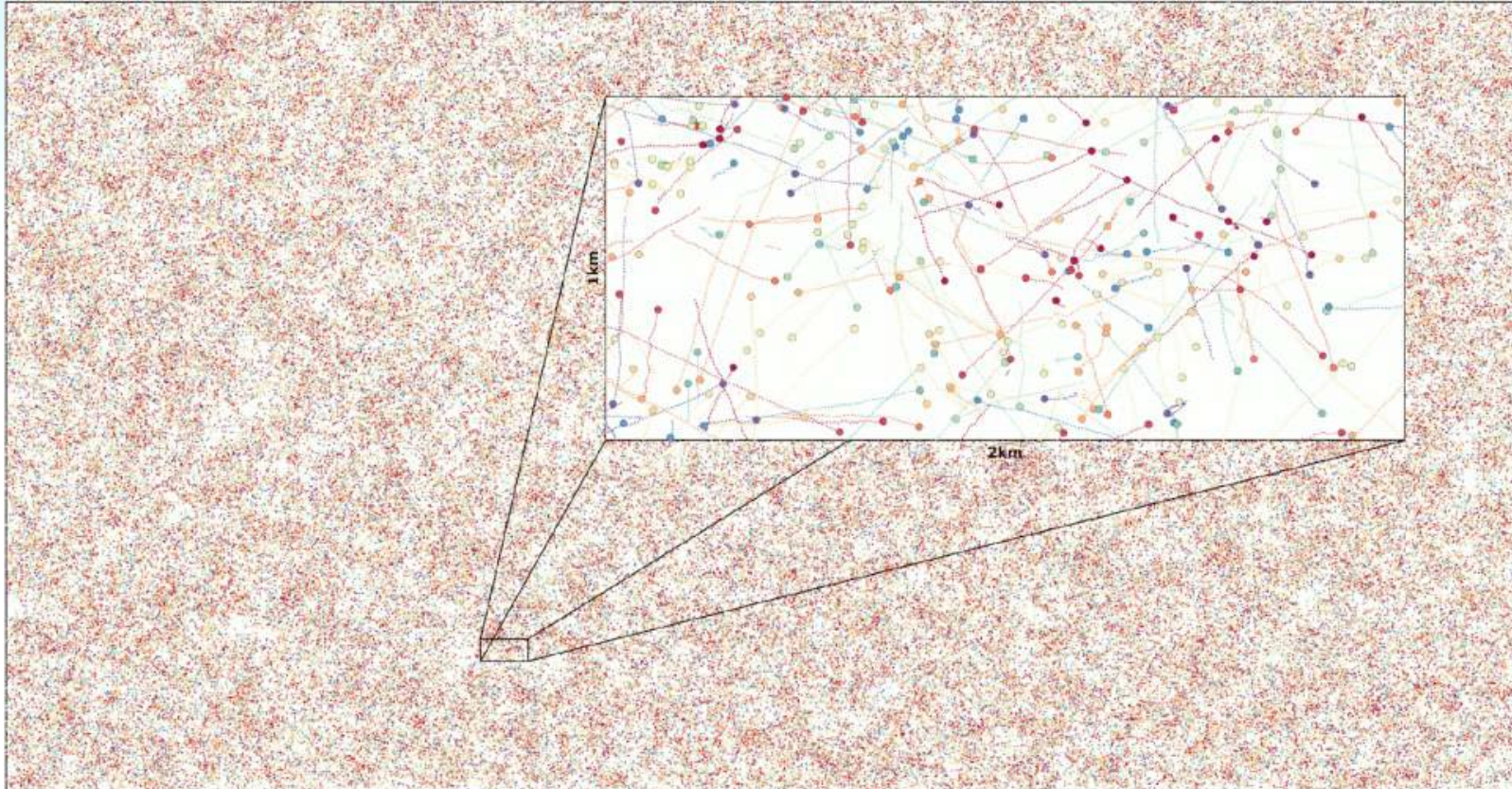
▶ Smoother implementation:

- Multi-dimensional assignment (NP-Hard): Gibbs sampling (linear in total no. of time steps)
- On-line: smooth over fixed-length window and link trajectories with the same labels

Vo & Vo "A Multi-Scan Labeled RFS Model for Multi-object State Estimation," IEEE Trans. SP, 67(19):4948-4963, 2019

Generalized Labeled Multi-Bernoulli (GLMB)

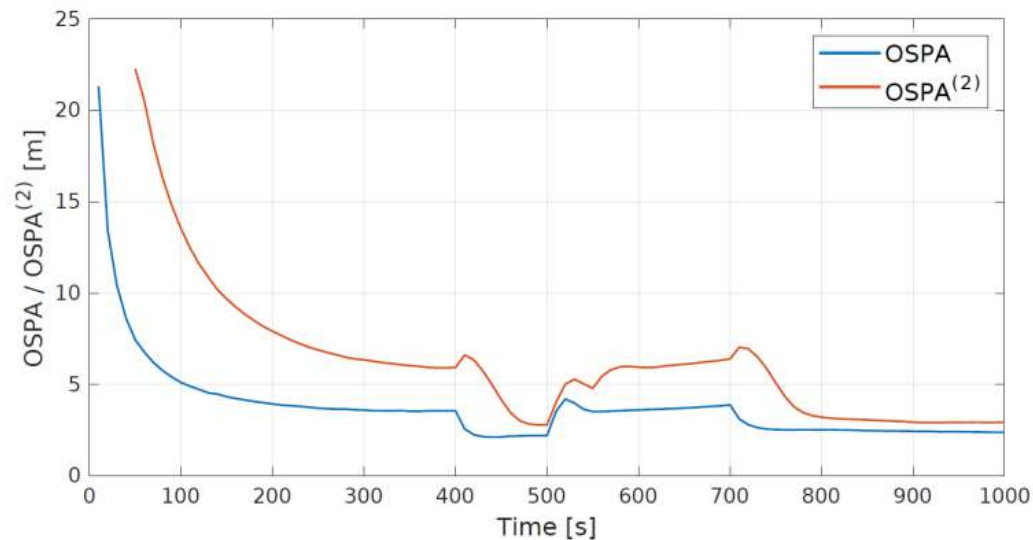
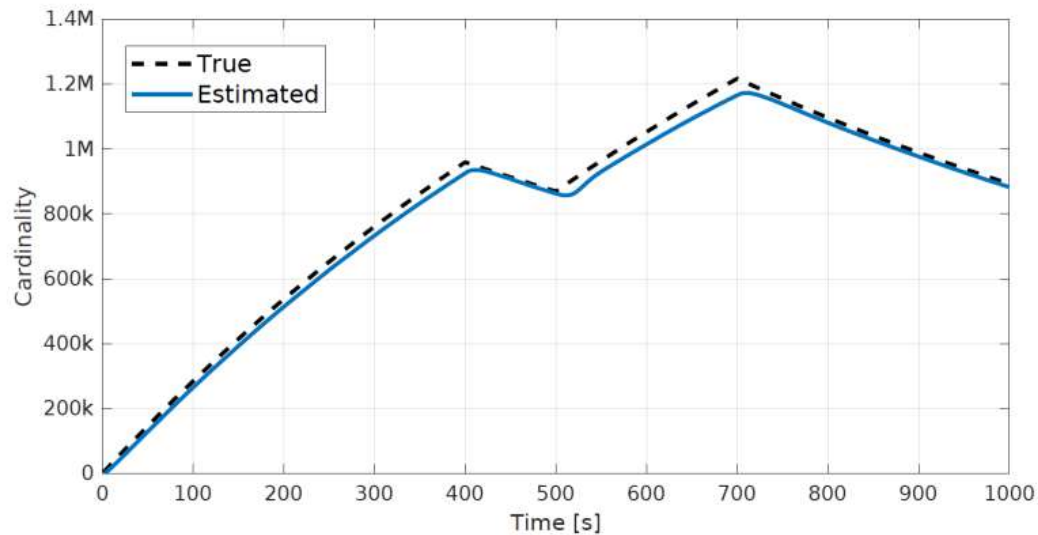
► Computational Tractability?



estimated trajectories on a 64km by 32km area

M. Beard et. al. "A Solution for Large-scale Multi-object Tracking," IEEE Trans. SP, 68:2754–2769, 2020.

Generalized Labeled Multi-Bernoulli (GLMB)



Computational Tractability?

Advanced algorithms: at best hundreds of objects/frame

Problem size = no. objects or observations/frame

GLMB

- Over **1 million** objects per frame
- Peak cardinality: 1,217,531 objects/frame (at time 700)
- Peak object density: 520 km⁻²
- Duration 1000 instances ... ~ 1 billion data points
- OSPA⁽²⁾: metric for sets of tracks,
- Generalization of OSPA that accounts for:
 - Localization & Cardinality error;
 - Track fragmentation;
 - ID switches

Beard et. al. "A Solution for Large-scale MOT," IEEE Trans. SP, 68:2754–2769, 2020.

Generalized Labeled Multi-Bernoulli (GLMB)

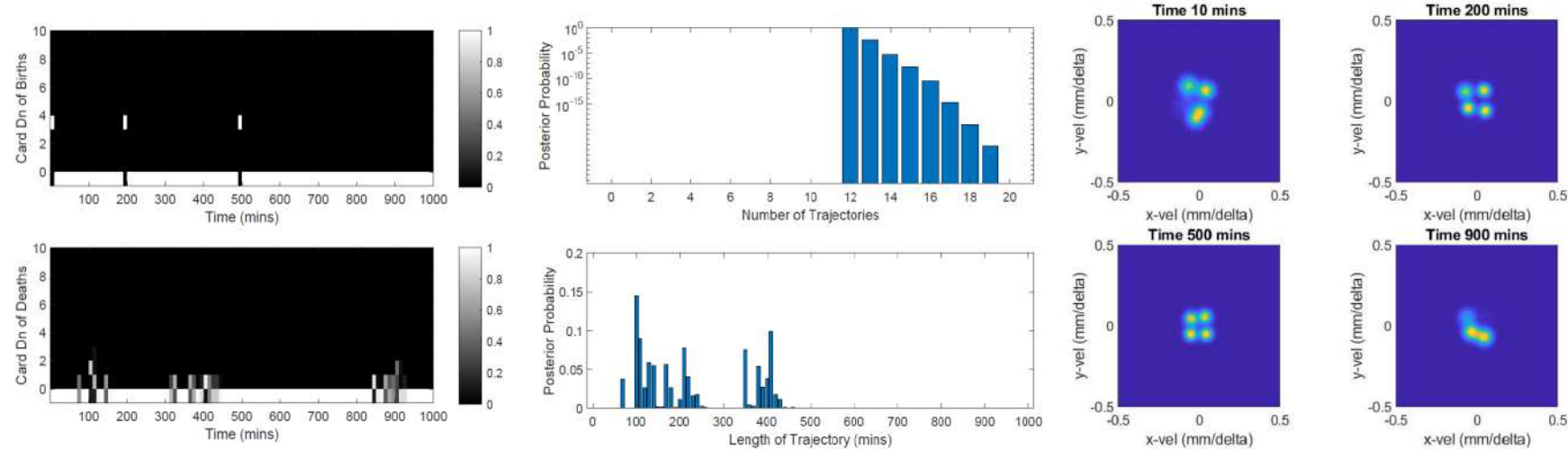
► Smoother: posterior statistics example in cell-microscopy

- Duration 1000mins, sampling period $\Delta=10$ mins
- Clutter 0.3/scan,
- Detection probability 0.33 (very low)
- Observation noise sigma 0.3mm
- Dynamic noise sigma $0.01\text{mm}/\Delta^2$

Scenario:

- min 001: 4 cells appear, live for 100mins
- min 200: 4 cells appear, live for 200mins
- min 500: 4 cells appear, live for 400mins

Statistics on births/deaths, cell-life, migration pattern

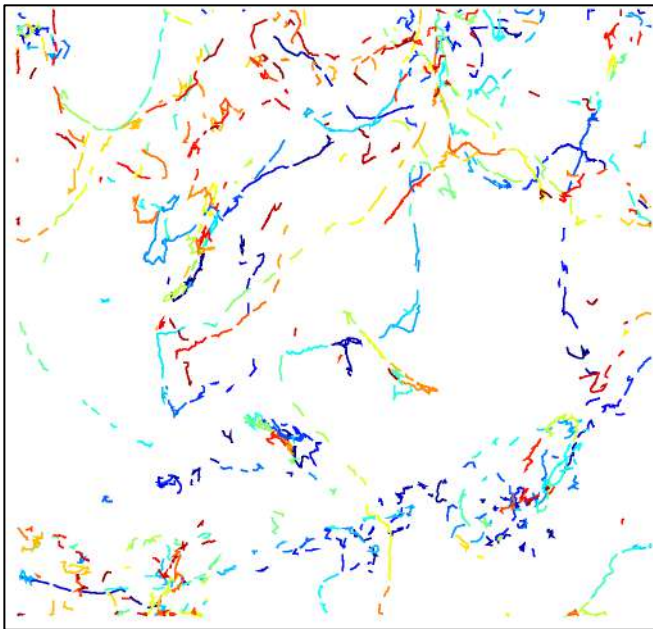


Application: Cell-Microscopy

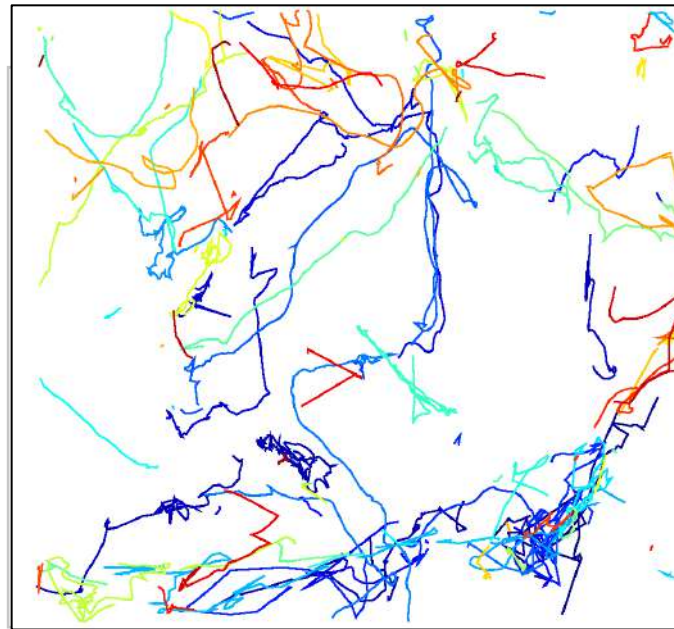
Cell Migration Analysis

- 4320 mins stem cell migration sequence, 1 image every 16 min

MHT (tree depth=5) [Chenouard et. al.13]



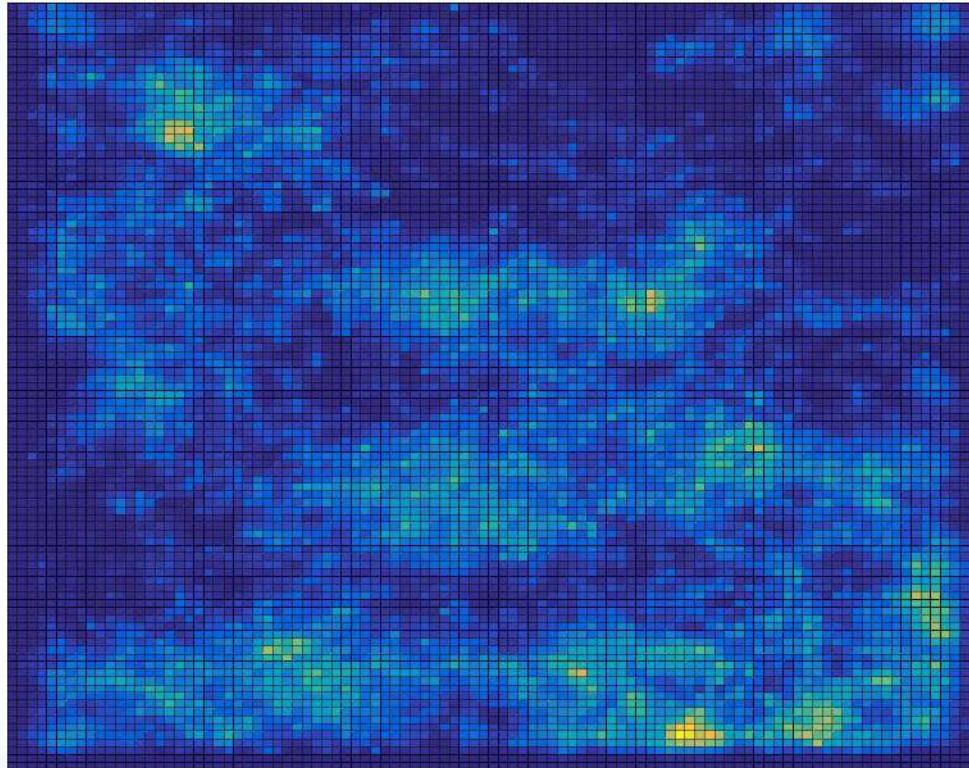
GLMB filter: online & less parameters



Application: Cell-Microscopy

Cell Migration Analysis

- Provides more than just tracking results: Statistics from the posterior distribution of the cells:



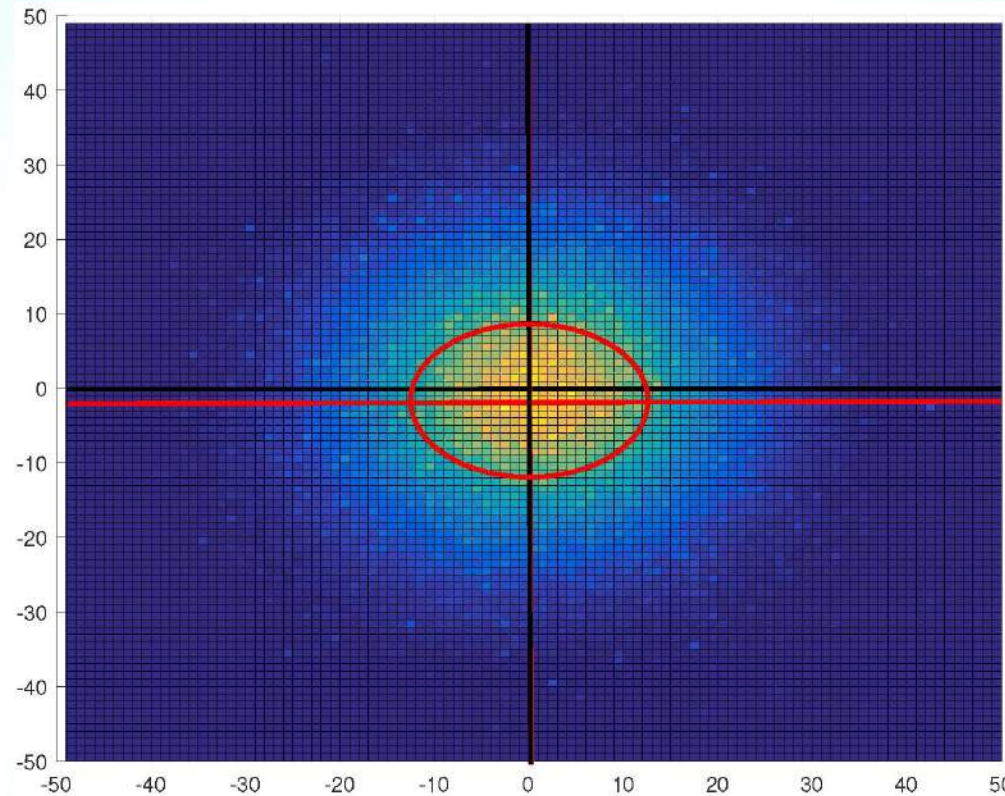
Net migration Top –Bottom!

Time averaged intensity function in position space Mean number of cells per unit area

Application: Cell-Microscopy

Cell Migration Analysis

- Provides more than just tracking results: Statistics from the posterior distribution of the cells:



Net migration Top –Bottom!

Time averaged intensity function in velocity space

Interpretation: Velocity heatmap or concentration of cells at different velocities

Application: Multi-Sensor Fusion

- ▶ 3D (state + extent) tracking from multiple camera views
 - 4 Kinect sensors placed high up and facing inwards in each corner
 - YOLO detector produces 2D detections of bounding boxes in pixel space
 - Detections are noisy and subject to false positives/negatives



Ong et. al. "A Bayesian Filter for Multi-view 3D Multi-object Tracking with Occlusion Handling," IEEE Trans. PAMI, 2020.

Application: Multi-Sensor Fusion

- Provides tracks in 3D instead of ground-plane tracks as in existing methods
- No retraining when reconfiguring cameras

CMC3 (Maximum/Average 15 people)													
Detector and Tracker	IDF1 ↑	IDP ↑	IDR ↑	MT ↑	PT ↓	ML ↓	FP ↓	FN ↓	IDs ↓	FM ↓	MOTA ↑	MOIP ↑	OSPA ⁽²⁾ ↓
YOLOv3+MV-GLMB-OC	70.7%	72.3%	69.1%	14	1	0	94	222	45	37	87.2%	52.8%	0.53
YOLOv3+MV-GLMB-OC*	60.8%	65.7%	56.6%	9	6	0	91	481	66	56	77.4%	46.4%	0.60
YOLOv3+MS-GLMB	41.4%	57.3%	32.4%	0	15	0	10	1239	64	60	53.5%	46.7%	0.76
Faster-RCNN(VGG16)+MV-GLMB-OC	63.7%	66.6%	61.1%	12	3	0	97	329	63	41	82.7%	52.8%	0.58
Faster-RCNN(VGG16)+MV-GLMB-OC*	57.3%	61.0%	54.0%	10	5	0	133	460	78	60	76.3%	47.9%	0.66
Faster-RCNN(VGG16)+MS-GLMB	45.7%	61.7%	36.3%	0	15	0	13	1175	61	67	55.8%	46.6%	0.75

CMC5 (Jumping and Falling, Maximum/Average 7 people)													
Detector and Tracker	IDF1 ↑	IDP ↑	IDR ↑	MT ↑	PT ↓	ML ↓	FP ↓	FN ↓	IDs ↓	FM ↓	MOTA ↑	MOTP ↑	OSPA ⁽²⁾ ↓
YOLOv3+MV-GLMB-OC	59.8%	61.0%	60.8%	3	4	0	404	951	67	54	60.6%	45.0%	0.65
YOLOv3+MV-GLMB-OC*	55.9%	54.9%	57.1%	3	3	1	689	1125	80	85	55.3%	43.4%	0.71
YOLOv3+MV-GLMB	49.5%	50.1%	45.0%	3	2	2	715	1750	94	91	49.3%	42.6%	0.78
Faster-RCNN(VGG16)+MV-GLMB-OC	58.1%	60.8%	59.4%	3	4	0	451	1008	72	57	59.9%	43.1%	0.66
Faster-RCNN(VGG16)+MV-GLMB-OC*	55.9%	53.6%	51.6%	3	3	1	569	1519	81	88	51.4%	42.7%	0.75
Faster-RCNN(VGG16)+MS-GLMB	48.8%	45.3%	41.7%	3	3	1	734	1493	96	98	43.3%	43.9%	0.81

CLEAR MOT scores and OSPA⁽²⁾ 3D errors with 3D GloU base-distance

Detectors (monocular): Faster-RCNN(VGG16) and YOLOv3

GLMB Filters: Standard Multi-Sensor (MS), Multi-View with Occlusion Model (MV)

Asterisk e.g. MV-GLMB-OC* indicates tracking while reconfiguring cameras

Application: Autonomous Driving

Autonomous Driving: SLAM + multi-object filtering



Prototype system with E-Class Mercedes-Benz

H. Deusch et. al. "The Labeled Multi-Bernoulli SLAM filter," IEEE Signal Proc. Letters, 22(10), 2015.

Application: Reinforcement Learning

Use drone to autonomously track and localize animals

How biologists track VHF-collared wildlife:

- Trek long distances with VHF radio receivers, directional antennas and battery packs
- Manually home in on radio collar signals from collared wildlife to find and locate them

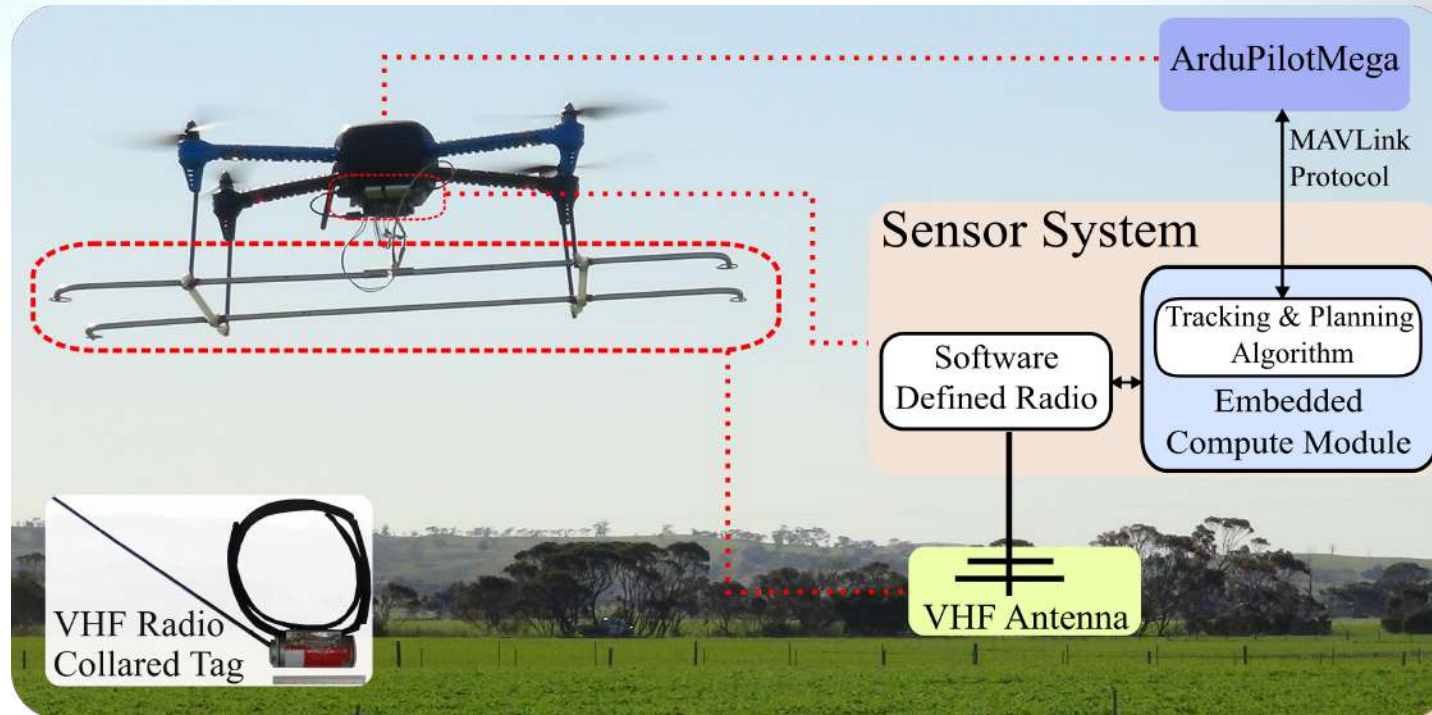


Handheld device

Application: Reinforcement Learning

Autonomous Drone:

- Pros: Low cost, reduces time and operating costs
- Cons: Introduce disturbances to wildlife



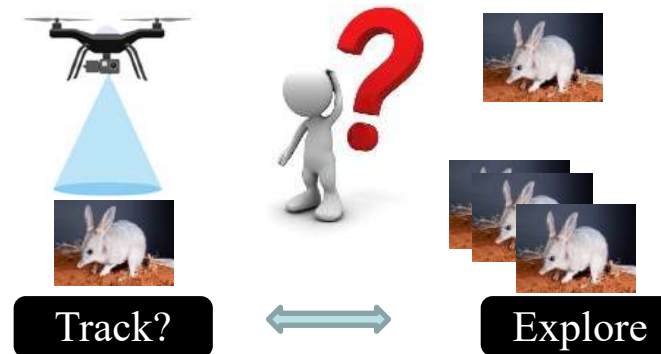
- Need intelligent drones that can perform prescribed mission on their own

Application: Reinforcement Learning

Team of agents to track and search for a time-varying number of mobile objects



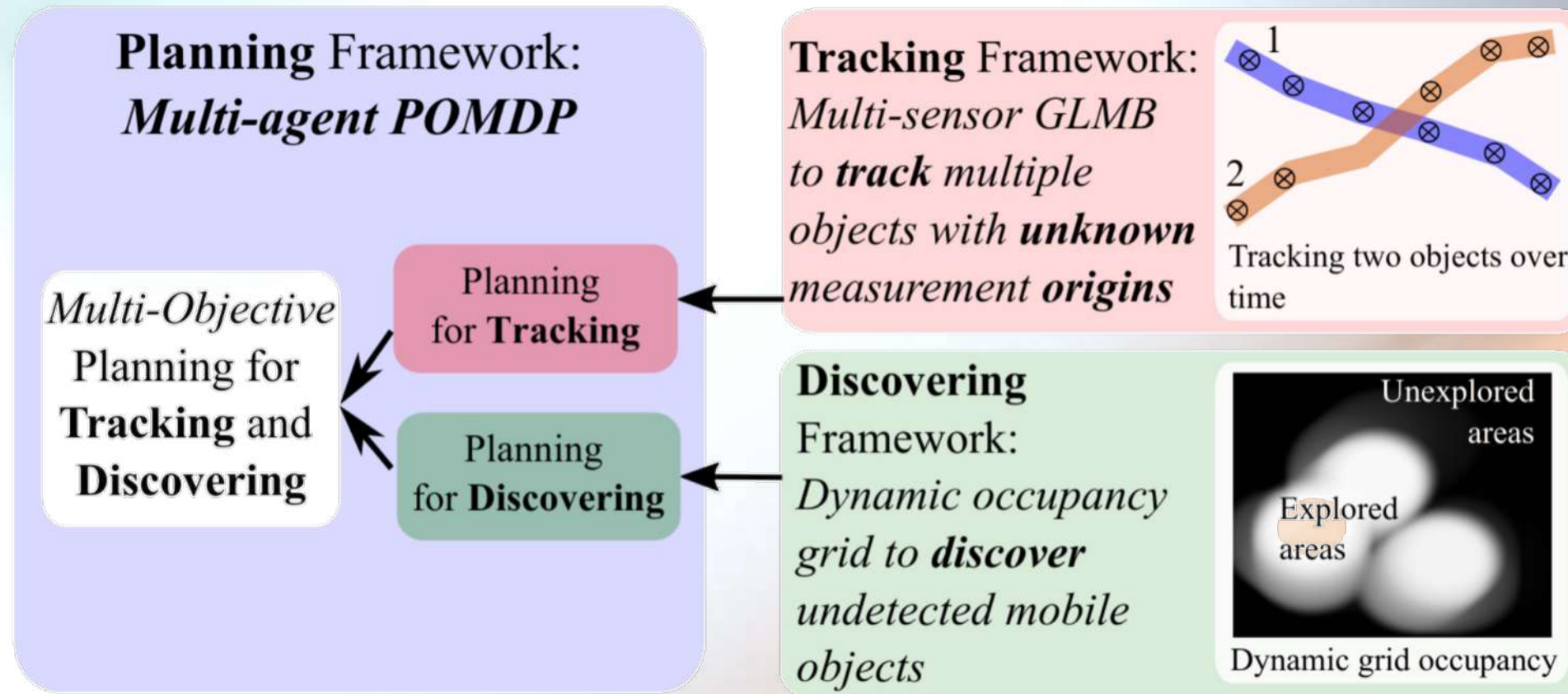
- ▶ Real-time operation
- ▶ Limited FoVs
- ▶ Competing Objectives



Nguyen et al., "Multi-Objective Multi-Agent Planning for Jointly Discovering and Tracking Mobile Objects," AAAI-2020.

Nguyen et al., "Multi-Objective Multi-Agent Planning for Discovering and Tracking Unknown and Varying Number of Mobile Objects," 2022 (arXiv:2203.04551)

Application: Reinforcement Learning



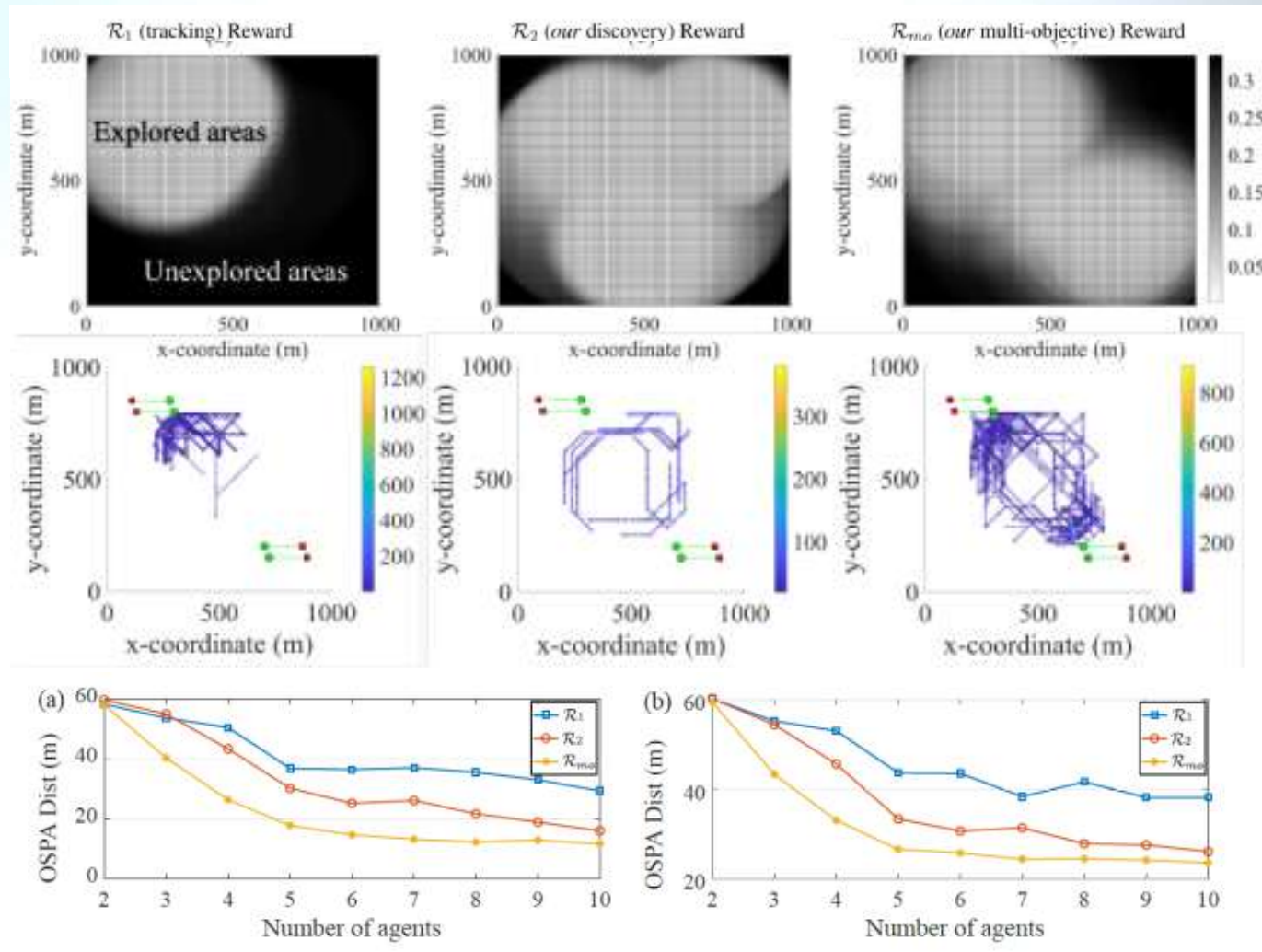
Inspired by the multi-objective RFS control solution in [Zhu et. al. 2019]

Nguyen et al., "Multi-Objective Multi-Agent Planning for Jointly Discovering and Tracking Mobile Objects," AAAI-2020.

Nguyen et al., "Multi-Objective Multi-Agent Planning for Discovering and Tracking Unknown and Varying Number of Mobile Objects," 2022 (arXiv:2203.04551)

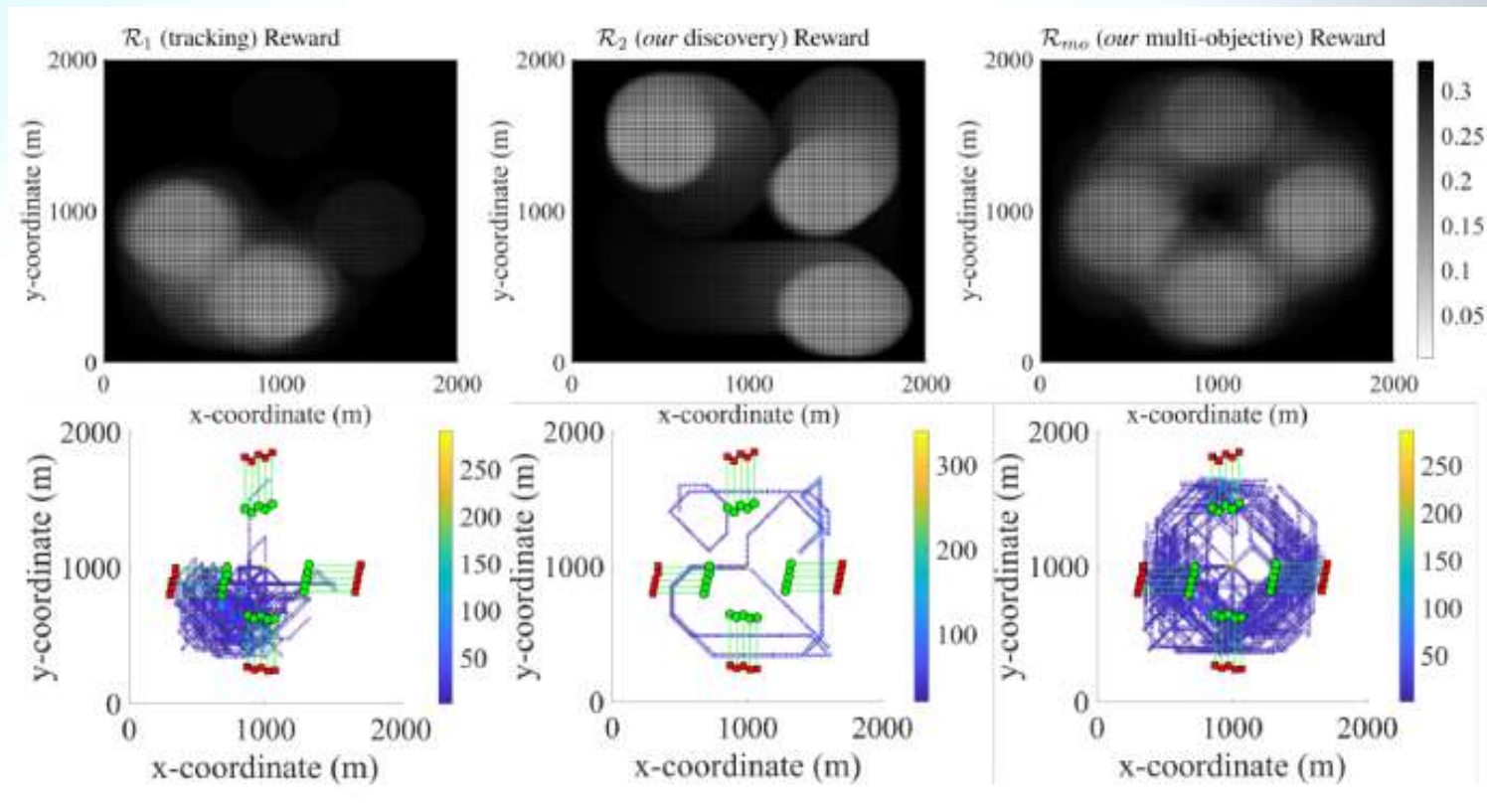
Application: Reinforcement Learning

Multi-objective strategies are more effective ...



Application: Reinforcement Learning

Multi-objective strategies are more effective ...



Conclusions

Multi-Object HMM provides:

- insights into the foundations of applications involving multiple objects &
- efficient and scalable solutions not possible previously

Many interesting problems in

- Artificial intelligence, machine learning, data mining
- Communications, Astro-dynamics (Space Debris)
- Bio-medical research: cell microscopy, brain imaging ...

Preprints/latest works: <http://ba-ngu.vo-au.com/publications.html>

Matlab code: <http://ba-tuong.vo-au.com/codes.html>

Thank You!