THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Poisson Multi-Bernoulli Mixtures for Multiple Object Tracking

From point objects to extended objects and from sets of objects to sets of trajectories

Yuxuan Xia

CHALMERS
UNIVERSITY OF TECHNOLOGY

Signal Processing Group
Department of Electrical Engineering
Chalmers University of Technology
Gothenburg, Sweden, 2022

Poisson Multi-Bernoulli Mixtures for Multiple Object Tracking

From point objects to extended objects and from sets of objects to sets of trajectories

Yuxuan Xia
ISBN 978-91-7905-690-2

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr 5156
ISSN 0346-718X

Department of Electrical Engineering
Chalmers University of Technology
SE-412 96 Gothenburg, Sweden
Phone: +46 (0)31 772 1000

Cover:
An example of multiple extended object tracking in an urban area.

Printed by Chalmers Reproservice
Gothenburg, Sweden, August 2022

To my family.

Every cloud has a silver lining.

Poisson Multi-Bernoulli Mixtures for Multiple Object Tracking

From point objects to extended objects and from sets of objects to sets of trajectories

Yuxuan Xia
Department of Electrical Engineering
Chalmers University of Technology

Abstract

Multi-object tracking (MOT) refers to the process of estimating object trajectories of interest based on sequences of noisy sensor measurements obtained from multiple sources. Nowadays, MOT has found applications in numerous areas, including, e.g., air traffic control, maritime navigation, remote sensing, intelligent video surveillance, and more recently environmental perception, which is a key enabling technology in automated vehicles. This thesis studies Poisson multi-Bernoulli mixture (PMBM) conjugate priors for MOT.

Finite Set Statistics provides an elegant Bayesian formulation of MOT based on random finite sets (RFSs), and a significant trend in RFSs-based MOT is the development of conjugate distributions in Bayesian probability theory, such as the PMBM distributions. Multi-object conjugate priors are of great interest as they provide families of distributions that are suitable to work with when seeking accurate approximations to the true posterior distributions.

Many RFS-based MOT approaches are only concerned with multi-object filtering without attempting to estimate object trajectories. An appealing approach to building trajectories is by computing the multi-object densities on sets of trajectories. This leads to the development of many multi-object filters based on sets of trajectories, e.g., the trajectory PMBM filters.

In this thesis, [Paper A] and [Paper B] consider the problem of point object tracking where an object generates at most one measurement per time scan. In [Paper A], a multi-scan implementation of trajectory PMBM filters via dual decomposition is presented. In [Paper B], a multi-trajectory particle smoother using backward simulation is presented for computing the multi-object posterior for sets of trajectories using a sequence of multi-object filtering densities and a multi-object dynamic model. [Paper C] and [Paper D] consider the problem of extended object tracking where an object may generate multiple measurements per time scan. In [Paper C], an extended object Poisson multi-Bernoulli (PMB) filter is presented, where the PMBM posterior density after the update step is approximated as a PMB. In [Paper D], a trajectory PMB filter for extended object tracking using belief propagation is presented, where the efficient PMB approximation is enabled by leveraging the PMBM conjugacy and the factor graph formulation.

Keywords: Bayesian filtering, random finite sets, conjugate prior, multi-object tracking, extended object, trajectory estimation.

List of Publications

This thesis is based on the following publications:

[A] **Yuxuan Xia**, Karl Granström, Lennart Svensson,
Ángel F. García-Fernández, Jason L. Williams, "Multi-scan implementation of the trajectory Poisson multi-Bernoulli mixture filter". Published in Journal of Advances in Information Fusion, Dec. 2019.

[B] **Yuxuan Xia**, Lennart Svensson, Ángel F. García-Fernández,
Jason L. Williams, Daniel Svensson, and Karl Granström, "Multiple object trajectory estimation using backward simulation". Published in IEEE Transactions on Signal Processing, Jun. 2022.

[C] **Yuxuan Xia**, Karl Granström, Lennart Svensson,
Maryam Fatemi, Ángel F. García-Fernández, Jason L. Williams, "Poisson multi-Bernoulli approximations for multiple extended object filtering". Published in IEEE Transactions on Aerospace and Electronic Systems, Apr. 2022.

[D] **Yuxuan Xia**, Ángel F. García-Fernández, Florian Meyer,
Jason L. Williams, Karl Granström, and Lennart Svensson, "Trajectory PMB filters for extended object tracking using belief propagation". To be submitted.

Other publications by the author, not included in this thesis, are:

[E] **Y. Xia**, P. Wang, K. Berntorp, L. Svensson, K. Granström, H. Mansour, P. Boufounos, P. Orlik, "Learning-based extended object tracking using hierarchical truncation measurement model with automotive radar". *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, n. 4, pp. 1013–1029, Feb. 2021.

[F] **Y. Xia**, L. Svensson, Á. F. García-Fernández, J. L. Williams, K. Granström, "Backward simulation for sets of trajectories". *23nd International Conference on Information Fusion (FUSION)*, Rustenburg, South Africa, Jul. 2020.

[G] **Y. Xia**, P. Wang, K. Berntorp, P. Boufounos, P. Orlik, L. Svensson, K. Granström, "Extended object tracking with automotive radar using learned

structural measurement model". *IEEE Radar Conference (RadarConf)*, Florence, Italy, Sep. 2020.

[H] **Y. Xia**, P. Wang, K. Berntorp, H. Mansour, P. Boufounos, P. Orlik, "Extended object tracking using hierarchical truncation model with partial-view measurements". *IEEE 11th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, Hangzhou, China, Jun. 2020.

[I] **Y. Xia**, P. Wang, K. Berntorp, T. Koike-Akino, H. Mansour, M. Pajovic, P. Boufounos, P. Orlik, "Extended object tracking using hierarchical truncation measurement model with automotive radar". *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Barcelona, Spain, May. 2020.

[J] **Y. Xia**, K. Granström, L. Svensson, Á. F. García-Fernández, J. L. Williams, "Extended target Poisson multi-Bernoulli mixture trackers based on sets of trajectories". *22rd International Conference on Information Fusion (FUSION)*, Ottawa, Canada, Jul. 2019.

[K] **Y. Xia**, K. Granström, L. Svensson, Á. F. García-Fernández, "An implementation of the Poisson multi-Bernoulli mixture trajectory filter via dual decomposition". *21st International Conference on Information Fusion (FUSION)*, Cambridge, United Kingdom, Jul. 2018.

[L] **Y. Xia**, K. Granström, L. Svensson, Á. F. García-Fernández, "Performance evaluation of multi-Bernoulli conjugate priors for multi-target filtering". *20th International Conference on Information Fusion (FUSION)*, Xi'an, China, Jul. 2017.

[M] W. Xiong, J. Liu, **Y. Xia**, T. Huang, B. Zhu, W. Xiang, "Contrastive learning for automotive mmWave radar detection points based instance segmentation". *IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*, Macau, China, Oct. 2022.

[N] Á. F. García-Fernández, **Y. Xia**, L. Svensson, "A comparison between PMBM Bayesian track initiation and labelled RFS adaptive birth". *25th International Conference on Information Fusion (FUSION)*, Linköping, Sweden, Jul. 2022.

[O] J. Liu, W. Xiong, L. Bai, **Y. Xia**, T. Huang, W. Ouyang, "Deep instance segmentation with automotive radar detection points". *IEEE Transactions on Intelligent Vehicles*, Apr. 2022.

[P] J. Pinto, **Y. Xia**, L. Svensson, H. Wymeersch, "An uncertainty-aware performance measure for multi-object tracking". *IEEE Signal Processing Letters*, vol. 28, pp. 1689–1693, Aug. 2021.

[Q] Á. F. García-Fernández, J. L. Williams, L. Svensson, **Y. Xia**, "A Poisson multi-Bernoulli mixture filter for coexisting point and extended targets". *IEEE Transactions on Signal Processing*, vol. 69, pp. 2600–2610, Apr. 2021.

[R] J. Pinto, G. Hess, W. Ljungbergh, **Y. Xia**, L. Svensson, H. Wymeersch, "Next generation multitarget trackers: random finite set methods vs transformer-based deep learning". *24th International Conference on Information Fusion (FUSION)*, Sun City, South Africa, Nov. 2021.

[S] Á. F. García-Fernández, L. Svensson, J. L. Williams, **Y. Xia**, K. Granström, "Trajectory Poisson multi-Bernoulli filters". *IEEE Transactions on Signal Processing*, vol. 68, pp. 4933–4945, Aug. 2020.

[T] Á. F. García-Fernández, L. Svensson, J. L. Williams, **Y. Xia**, K. Granström, "Trajectory multi-Bernoulli filters for multi-target tracking based on sets of trajectories". *23nd International Conference on Information Fusion (FUSION)*, Rustenburg, South Africa, Jul. 2020.

[U] K. Granström, L. Svensson, **Y. Xia**, Á. F. García-Fernández, J. L. Williams, "Spatiotemporal constraints for sets of trajectories with applications to PMBM densities". *23nd International Conference on Information Fusion (FUSION)*, Rustenburg, South Africa, Jul. 2020.

[V] Á. F. García-Fernández, **Y. Xia**, K. Granström, L. Svensson, J. L. Williams, "Gaussian implementation of the multi-Bernoulli mixture filter". *22rd International Conference on Information Fusion (FUSION)*, Ottawa, Canada, Jul. 2019.

[W] K. Granström, L. Svensson, **Y. Xia**, Á. F. García-Fernández, J. L. Williams, "Poisson multi-Bernoulli mixture trackers: continuity through random finite

sets of trajectories". *21st International Conference on Information Fusion (FUSION)*, Cambridge, United Kingdom, Jul. 2018.

[X] K. Granström, L. Svensson, S. Reuter, **Y. Xia**, M. Fatemi, "Likelihood-based data association for extended object tracking using sampling methods". *IEEE Transactions on Intelligent Vehicles*, vol. 3, no. 1, pp. 30–45, Dec. 2017.

Acknowledgments

Finally, my long Ph.D. journey has come to its end! There have been many ups and downs. It took about three years to have my first journal paper published, but in the end I am happy with the research outcomes that I have obtained so far. I would like to take this opportunity to express my appreciation to the people, without whom this thesis would not have been possible.

First and foremost I am extremely grateful to my examiner and supervisor Prof. Lennart Svensson and my previous supervisor Dr. Karl Granström now at Embark Trucks for drawing up the research project and their immense help and support on my research. Their enthusiasm, encouragement and patience stimulate me to be a better researcher.

I would like to express my sincere gratitude to Dr. Ángel F. García-Fernández at Liverpool University for all the stimulating discussions as well as the timely and detailed feedback on the many papers we have collaborated on. I would like to extend my sincere thanks to Dr. Jason L. Williams at CSIRO Data61 for his research ideas and insightful comments on the manuscripts. I am also deeply grateful to Dr. Daniel Svensson now at Nvidia and Dr. Maryam Fatemi at Zenseact for being the coordinators of my Ph.D. project and for proofreading of my paper and thesis.

I would also like to thank Dr. Pu (Perry) Wang for being my host and for all the support I received from you during my internship at MERL. Additionally, I would like to thank Dr. Karl Berntorp, Dr. Petros Boufounos, Dr. Philip Orlik, and many other people I met at MERL for all the interesting discussions. I also appreciate Prof. Florian Meyer at University of California San Diego for your expertise on graph models for multi-object tracking.

My appreciation also goes to Jan Krejči for proofreading my thesis and the interesting discussions we had during your internship in the signal processing group at Chalmers. I would like to offer my special thanks to Prof. Lars Hammarstrand. It was a nice experience being a teaching assistant of the Sensor Fusion and Nonlinear Filtering course that you teach. I am also very grateful to all the current and former colleagues in the signal processing group for jointly maintaining a pleasant research atmosphere. Many thanks also to my Chinese friends in the department for all the joyful time we have had together.

Last but not the least, my deepest thanks go to my family for their love, understanding and support all the way in my journey towards a Ph.D. degree.

Acronyms

BP:	Belief Propagation
CPHD:	Cardinalized probability hypothesis density
EKF:	Extended Kalman filter
EOT:	Extended object tracking
GGIW:	Gamma Gaussian inverse-Wishart
GLMB:	Generalized labelled multi-Bernoulli
GNN:	Global nearest neighbour
GOSPA:	Generalized optimal sub-pattern assignment
GWD:	Gaussian Wasserstein distance
JPDA:	Joint probabilistic data association
KLD:	Kullback-Leibler divergence
MB:	Multi-Bernoulli
MBM:	Multi-Bernoulli mixture
MHT:	Multiple hypothesis tracker
MOT:	Multi-object tracking
OSPA:	Optimal sub-pattern assignment
PGFL:	Probability generating functionals
PHD:	Probability hypothesis density
PMB:	Poisson multi-Bernoulli
PMBM:	Poisson multi-Bernoulli mixture
PLF:	Posterior linearized filter

PPP:	Poisson point process
RFS:	Random finite set
RTSS:	Rauch-Tung-Striebel smoother
SPA:	Sum Product Algorithm

Contents

Part I

Overview

Part I Overview is a modified version of Part I Overview in Yuxuan Xia's Licentiate thesis "Conjugate priors for Bayesian object tracking", Chalmers University of Technology, 2020.

CHAPTER 1

Introduction

1.1 Background

Object tracking refers to the problem of using sensor measurements to determine the location, trajectory, and characteristics of objects of interest [1], [2]. Initially driven by aerospace and defence applications, object tracking has a long history spanning over decades. In recent times, with the advances in object tracking techniques as well as sensing and computing technologies, there has been an explosion in the use of object tracking technology in numerous research venues as well as many application areas. Typical examples include air traffic control, maritime navigation, remote sensing, biomedical research, intelligent video surveillance, and more recently environmental perception, which is a key enabling technology in automated vehicles. Specifically, environmental perception in automated driving involves processing data collected from vehicle sensors into an understanding of the world around the vehicle.

This thesis studies Bayesian object tracking algorithms. Bayes's theorem provides an elegant and powerful probabilistic framework to solve the object tracking problem. The goal of Bayesian estimation in object tracking is to compute the posterior density of the random variables of interest, which en-

capsulates all the information contained in the measurements [3]. In object tracking, the objects may not always be detected, and the sensor measurements are noisy and contain clutter. Conventional tracking algorithms consider the problem of point object tracking with the assumption that an object gives rise to at most one measurement per scan. However, objects may occupy multiple sensor resolution cells, depending on their distance to the sensor and the sensor resolution. This gives rise to multiple measurements per scan. The tracking of such an object is called extended object tracking (EOT) [4], and it has found more applications with the development of high-resolution sensors, such as automotive radar and lidar [5]–[9].

In a multi-object scenario, the measurements may originate from one of the various objects, and the number of objects is generally time-varying due to objects appearing in and disappearing from the surveillance area [10]. Fundamental to this problem is the estimation of both the number of objects and their trajectories by partitioning measurements into the sets of measurements originating from different objects and clutter. The major approaches to multi-object tracking (MOT) include the global nearest neighbour (GNN) filter, the joint probabilistic data association (JPDA) filter [11], the multiple hypothesis tracker (MHT) [12], and random finite sets (RFSs) based multi-object filters.

Finite Set Statistics [13] provides a theoretically elegant Bayesian formulation of MOT based on RFSs where the multi-object state is represented as a finite set of single-object states [14], [15]. Exact closed-form solutions of RFS-based MOT Bayes filters are captured using multi-object conjugate priors [16], which means that if we start with the proposed conjugate initial prior, then all subsequent predicted and posterior distributions have the same form as the initial prior. MOT conjugate priors are of great interest as they provide families of distributions that are suitable to work with when seeking accurate approximations to the true posterior distributions. Two well-established MOT conjugate priors are the Poisson multi-Bernoulli mixture (PMBM) [17], based on unlabelled RFSs, and the generalized labelled multi-Bernoulli (GLMB) [16], based on labelled RFSs. The difference between these two MOT conjugate priors mainly lies in 1) whether the elements in RFSs are (uniquely) labelled, and 2) the modelling of newborn objects, i.e. objects appearing in the surveillance area.

For the standard multi-object models with Poisson point process (PPP) birth, the posterior density is a PMBM [17]. The PMBM density has a com-

pact representation of global hypotheses with probabilistic object existence in each Bernoulli component and undetected objects represented by a PPP. The PMBM filtering recursions have been established for point objects [18], for extended objects [19], [20], as well as for coexisting point and extended objects [21]. The PMBM filter and its approximations have been successfully applied not only to tracking of moving objects, but also to the mapping of stationary objects [22], joint tracking and sensor localization [23], decentralized sensor fusion [24], sensor management [25], as well as simultaneous localization and mapping for 5G applications [26].

If the birth model is multi-Bernoulli (MB) instead of Poisson, the filtering density is a multi-Bernoulli mixture (MBM). The MBM filtering recursion corresponds to the PMBM filtering recursion by setting the intensity of the Poisson process to zero and adding Bernoulli components for newborn objects in the prediction step [27]. The MBM filter can be further extended to consider MBs with deterministic object existence, which we refer to as the MBM_{01} filter, at the expense of increasing the number of global hypotheses [18]. Both MBM and MBM_{01} filters can consider object states with labels, and the (labelled) MBM_{01} filtering recursion is analogous to the delta GLMB filtering recursion [28].

Vector-type MOT methods, e.g. the JPDA filter and the MHT, describe the multi-object states and measurements by random vectors. They explicitly estimate trajectories; i.e. they associate a state estimate with a previous state estimate or declare the appearance of a new object. For MOT methods based on set representations, time sequences of tracks cannot be constructed easily as the multi-object states are order independent. For this reason, many RFS-based MOT approaches, e.g. the probability hypothesis density (PHD) filter [29] and the cardinalized PHD (CPHD) filter [30], are only concerned with multi-object filtering, in which one aims to estimate the current set of objects, without attempting to estimate object trajectories. The PMBM filter has a hypothesis structure similar to MHT [31], but track continuity in the form of trajectories is not explicitly established as the posterior itself only provides information about the current set of objects. A categorization of the multi-object filters mentioned so far is illustrated in Fig. 1.1.

One approach to building trajectories from posterior densities is to add unique labels to the object states and form trajectories by linking object state estimates with the same label [16], [32], [33]. Sequential track building ap-

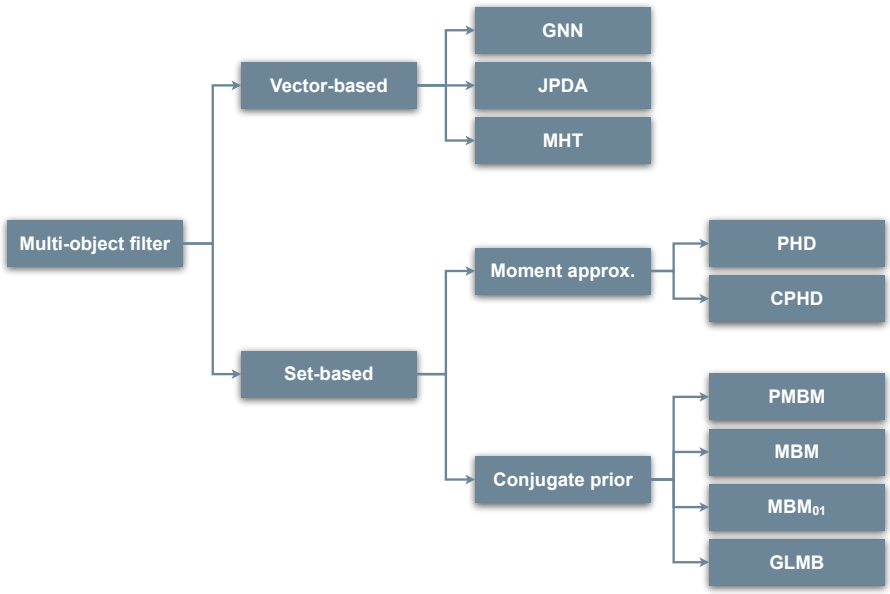

Figure 1.1: A chart categorizing different multi-object filters.

proaches based on labelling work well in many cases, but they are not always adequate due to ambiguity in object-to-label associations, e.g. when object birth is independent and identically distributed, and when objects get in close proximity and then separate. The above track building problems can be solved by computing multi-object densities on sets of trajectories [34]. This leads to the development of trajectory filters including, e.g. the trajectory PHD filter [35], the trajectory CPHD filter [35], the trajectory PMBM filter [36]–[39], its approximation to the trajectory Poisson multi-Bernoulli (PMB) filter [40], the trajectory MBM filter [41], and its approximation to the trajectory MB filter [42].

Bayes filters use the measurements obtained up until and including the current time step for computing the estimate of the current object state. However, sometimes it is also interesting to exploit the entire measurement history to arrive at more accurate object state estimates at all the preceding time steps. This problem can be solved with Bayesian smoothing. The multi-object generalization of a Bayesian smoother algorithm computes the multi-object den-

sities at all the preceding time steps given the entire batch of measurements. Existing literature on multi-object smoothing [15, Chapter 14] only focuses on computing the multi-object smoothing densities at each time step, which, even if labelled, may not be enough to provide trajectory information. Multi-object trajectory filters compute the filtering densities of sets of trajectories and can therefore directly produce smoothed trajectory estimates using, e.g. the accumulated state densities [43]. Nevertheless, many MOT methods efficiently estimate the object states, but cannot easily produce trajectory estimation in a principled manner. Therefore, how to leverage multi-object filters that do not keep trajectory information to compute the posterior density of sets of trajectories needs further investigation.

Contributions

This thesis investigates Bayesian object tracking methods for both point and extended objects, with a particular focus on Bayesian MOT methods based on sets of trajectories. [Paper A] and [Paper B] consider the problem of point object tracking. In [Paper A], the filtering recursions for the trajectory MBM filter using an MB birth model are presented. In addition, the multi-scan implementations of trajectory PMBM, MBM, and MBM_{01} filters using dual decomposition and N-scan pruning are proposed. In [Paper B], a general solution for multi-object smoothing and trajectory estimation is presented, along with its particle implementation using backward simulation. The proposed multi-object smoother computes the posterior of the set of trajectories from a sequence of multi-object filtering densities and the multi-object dynamic model.

[Paper C] and [Paper D] consider the problem of extended object tracking. In [Paper C], an extended object PMB filter is presented, where the PMBM posterior density after the update step is approximated as a PMB. Two approximation methods are presented: one is based on the track-oriented multi-Bernoulli (MB) approximation, and the other is based on the variational MB approximation via Kullback-Leibler divergence (KLD) minimization. In [Paper D], the PMBM conjugacy for standard multi-object models with a generalized measurement, in which each object generates an independent set of measurements, is generalized to sets of trajectories. In addition, a trajectory PMB filter for extended object tracking using belief propagation is presented, where the efficient PMB approximation is enabled by leveraging the PMBM

conjugacy and the factor graph formulation.

1.2 Thesis outline

The remainder of Part I of the thesis is organized as follows. Chapter 2 reviews Bayesian filtering and smoothing in dynamical systems, including a brief introduction to factor graphs and belief propagation. Chapter 3 covers the basic concepts and properties of random finite sets as well as metrics for tracking performance evaluation. Chapter 4 introduces the multi-object dynamic and measurement models used in this thesis. Chapter 5 presents the single-object and multi-object conjugacy for object tracking. Chapter 6 is about multi-object tracking based on sets of trajectories. Chapter 7 provides a summary of the included papers in this thesis. Finally, Chapter 8 summarizes the conclusions and possible directions for future work. Part II of the thesis includes the appended papers.

1.3 Notation

This section introduces the notations used in Part I of the thesis. Vectors are generally represented by lower-case letters (e.g. x). Matrices are generally represented by upper-case letters (e.g. X). Sets of vectors are represented by bold lower-case letters (e.g. $\mathbf{x}$). Sets of trajectories are represented by bold upper-case letters (e.g. $\mathbf{X}$). Classes of distributions are represented by calligraphy letters (e.g, $\mathcal{X}$). Spaces are generally represented by blackboard bold letters. For example, the n-dimensional Euclidean space is denoted by $\mathbb{R}^n$ and the space of positive integers is denoted by $\mathbb{N}$. The more general spaces are represented by Fraktur letters (e.g. $\mathfrak{X}$). The set of finite subsets of a space $\mathfrak{X}$ is represented by $\mathcal{F}(\mathfrak{X})$.

CHAPTER 2

Bayesian inference

Bayesian inference in dynamical systems refers to a class of methods that can be used for estimating the state of a dynamical system which is indirectly observed through noisy measurements. This chapter covers the basic aspects of Bayesian inference in dynamical systems. The reader is referred to [3] and [44] for further readings on these topics.

2.1 Bayesian inference in dynamical systems

The state of the dynamical system at time step k is denoted as $x_k \in \mathbb{R}^{n_x}$ where n_x is the dimension of the state. In the context of object tracking, x_k may represent the object's position, velocity, and any other motion or extent parameters of interest of the object at time step k. The state has an initial prior density $p(x_0)$ at time 0, and it evolves in time according to a Markov process with transition density $p(x_k|x_{k-1})$, where a Markov process is a random process in which the future is independent of the past, given the present. At each time step k, the state is observed through a noisy measurement $z_k \in \mathbb{R}^{n_z}$ whose likelihood is $p(z_k|x_k)$. Let $x_{0:k} = (x_0, x_1, \ldots, x_k)$ denote the sequence of states from time step 0 to k. Note that in object tracking the sequence

$x_{0:k}$ is used to represent the object trajectory up to time step k. Also, let $z_{1:k} = (z_1, \ldots, z_k)$ denote the sequence of measurements from time step 1 to k. The joint density of all the states and measurements up to time step k is given by

$$p(x_{0:k}, z_{1:k}) = p(x_0) \prod_{j=1}^{k} p(x_j|x_{j-1}) p(z_j|x_j). \tag{2.1}$$

In the Bayesian framework, all the information of interest in the state sequence $x_{0:k}$ is given by the posterior density $p(x_{0:k}|z_{1:k})$, which denotes the density of $x_{0:k}$ given the measurement sequence $z_{1:k}$. This density can be computed by applying Bayes' rule on (2.1)

$$\begin{aligned} p(x_{0:k}|z_{1:k}) &= \frac{p(x_{0:k}, z_{1:k})}{p(z_{1:k})} \\ &= \frac{p(x_0) \prod_{j=1}^{k} p(x_j|x_{j-1}) p(z_j|x_j)}{p(z_{1:k})}, \end{aligned} \tag{2.2}$$

where $p(z_{1:k})$ is the normalization constant defined as

$$p(z_{1:k}) = \int p(x_{0:k}, z_{1:k}) dx_{0:k}. \tag{2.3}$$

For long time sequences, it is usually intractable to compute the posterior density for state sequences, described in (2.2), without approximations. Bayesian inference in dynamical systems therefore often focuses on the following simpler problems:

- Filtering: the objective is to compute the density $p(x_k|z_{1:k})$ of the current state x_k given the measurements up to and including the current time step.
- Smoothing: the objective is to compute the density $p(x_j|z_{1:k})$ of a past state x_j, where $0 \leq j < k$, given the measurements up to and including the current time step k.
- Prediction: the objective is to compute the density $p(x_j|z_{1:k})$ of a future state x_j, where $j > k$, given the measurements up to and including the current time step k.

However, sometimes it is necessary to approximate the posterior density on

the entire state sequence rather than focusing on simpler problems, such as filtering or smoothing. Typical examples include backward simulation particle smoother [45], view-based simultaneous localization and mapping using delayed-state filters [46] and MOT where the objective is to infer object trajectories [34], [47], [48].

2.2 Bayesian filtering

In Bayesian filtering, the objective is to compute the filtering density $p(x_k|z_{1:k})$, which can be done using the Bayesian filtering recursion steps: prediction and update. Given the filtering density $p(x_{k-1}|z_{1:k-1})$ and the transition density $p(x_k|x_{k-1})$ at time step $k-1$, the predicted density $p(x_k|z_{1:k-1})$ is given by the Chapman-Kolmogorov equation

$$p(x_k|z_{1:k-1}) = \int p(x_k|x_{k-1})p(x_{k-1}|z_{1:k-1})dx_{k-1}. \tag{2.4}$$

In the update step, given the predicted density $p(x_k|z_{1:k-1})$ and the measurement distribution $p(z_k|x_k)$ at time step k, the filtering density at time k is given by Bayes' rule

$$p(x_k|z_{1:k}) = \frac{p(z_k|x_k)p(x_k|z_{1:k-1})}{p(z_k|z_{1:k-1})}, \tag{2.5}$$

where the normalization constant is

$$p(z_k|z_{1:k-1}) = \int p(z_k|x_k)p(x_k|z_{1:k-1})dx_k. \tag{2.6}$$

Kalman filtering

The Kalman filter is a closed-form solution to the Bayesian filtering equations for linear Gaussian dynamic and measurement models:

$$p(x_k|x_{k-1}) = \mathcal{N}(x_k; Fx_{k-1}, Q), \tag{2.7a}$$
$$p(z_k|x_k) = \mathcal{N}(z_k; Hx_k, R), \tag{2.7b}$$

where $F \in \mathbb{R}^{n_x,n_x}$ is the transition matrix, $Q \in \mathbb{R}^{n_x,n_x}$ is the covariance matrix of the process noise, $H \in \mathbb{R}^{n_z,n_x}$ is the observation matrix, and $R \in \mathbb{R}^{n_z,n_z}$

is the covariance matrix of the measurement noise. The prior density of the state at time 0 is $p(x_0) = \mathcal{N}(x_0; \bar{x}_{0|0}, P_{0|0})$, where $\bar{x}_{0|0}$ and $P_{0|0}$ are the a priori mean and covariance matrix of the state at time step 0.

The prediction and filtering densities at time step k are Gaussian and denoted as

$$p(x_k|z_{1:k-1}) = \mathcal{N}(x_k; \bar{x}_{k|k-1}, P_{k|k-1}), \tag{2.8a}$$
$$p(x_k|z_{1:k}) = \mathcal{N}(x_k; \bar{x}_{k|k}, P_{k|k}), \tag{2.8b}$$

where $\bar{x}_{k|k-1}$ and $P_{k|k-1}$ are the mean and the covariance of the predicted density and $\bar{x}_{k|k}$ and $P_{k|k}$ are the mean and the covariance of the filtering density. The parameters of the distributions in (2.8) can be computed with the following Kalman filter prediction and update steps.

- The prediction step is

$$\bar{x}_{k|k-1} = F\bar{x}_{k-1|k-1}, \tag{2.9a}$$
$$P_{k|k-1} = FP_{k-1|k-1}F^{\mathrm{T}} + Q. \tag{2.9b}$$

- The update step is

$$\bar{x}_{k|k} = \bar{x}_{k|k-1} + K_k(z_k - \bar{z}_k), \tag{2.10a}$$
$$P_{k|k} = P_{k|k-1} - K_k\Psi_k^{\mathrm{T}}, \tag{2.10b}$$
$$\bar{z}_k = H\bar{x}_{k|k-1}, \tag{2.10c}$$
$$\Psi_k = P_{k|k-1}H^{\mathrm{T}}, \tag{2.10d}$$
$$S_k = HP_{k|k-1}H^{\mathrm{T}} + R, \tag{2.10e}$$
$$K_k = \Psi_k S_k^{-1}, \tag{2.10f}$$

where K_k and S_k are usually referred to as the Kalman gain and the innovation covariance, respectively. The recursion is started from the prior mean $\bar{x}_{0|0}$ and covariance $P_{0|0}$.

Extended Kalman filter

The Kalman filter is not appropriate when the dynamic and/or measurement models are non-linear. However, the filtering distributions of non-linear

models can often be approximated by Gaussian distributions. The extended Kalman filter (EKF) is a "non-linear Kalman filter" based on Taylor series expansions. For dynamical systems with additive Gaussian noise, the transition and measurement densities have the form [3, Chapter 5]

$$p(x_k|x_{k-1}) = \mathcal{N}(x_k; f(x_{k-1}), Q), \tag{2.11a}$$
$$p(z_k|x_k) = \mathcal{N}(z_k; h(x_k), R), \tag{2.11b}$$

where $f(\cdot)$ and $h(\cdot)$ are possibly non-linear dynamic and measurement model functions, respectively.

The idea of the EKF is to assume Gaussian approximations

$$p(x_k|z_{1:k-1}) \approx \mathcal{N}(x_k; \bar{x}_{k|k-1}, P_{k|k-1}), \tag{2.12a}$$
$$p(x_k|z_{1:k}) \approx \mathcal{N}(x_k; \bar{x}_{k|k}, P_{k|k}), \tag{2.12b}$$

to the prediction and filtering densities, and to use first-order Taylor series approximations to the non-linear functions $f(\cdot)$ and $h(\cdot)$ around $\bar{x}_{k-1|k-1}$ and $\bar{x}_{k|k-1}$, respectively. Let $F(\bar{x}_{k-1|k-1})$ be the Jacobian matrix of $f(\cdot)$ evaluated at $\bar{x}_{k-1|k-1}$ and let $H(\bar{x}_{k|k-1})$ be the Jacobian matrix of $h(\cdot)$ evaluated at $\bar{x}_{k|k-1}$. The parameters of the distributions in (2.12) can then be computed with the following EKF prediction and update steps.

- The prediction step is

$$\bar{x}_{k|k-1} = f(\bar{x}_{k-1|k-1}), \tag{2.13a}$$
$$P_{k|k-1} = F(\bar{x}_{k-1|k-1})P_{k-1|k-1}F(\bar{x}_{k-1|k-1})^{\mathrm{T}} + Q. \tag{2.13b}$$

- The update step is

$$\bar{x}_{k|k} = \bar{x}_{k|k-1} + K_k(z_k - \bar{z}_k), \tag{2.14a}$$
$$P_{k|k} = P_{k|k-1} - K_k\Psi_k^{\mathrm{T}}, \tag{2.14b}$$
$$\bar{z}_k = h(\bar{x}_{k|k-1}), \tag{2.14c}$$
$$\Psi_k = P_{k|k-1}H(\bar{x}_{k|k-1})^{\mathrm{T}}, \tag{2.14d}$$
$$S_k = H(\bar{x}_{k|k-1})P_{k|k-1}H(\bar{x}_{k|k-1})^{\mathrm{T}} + R, \tag{2.14e}$$
$$K_k = \Psi_k S_k^{-1}. \tag{2.14f}$$

There are other non-linear Kalman filters based on Gaussian approximations in the literature, such as the posterior linearized filter (PLF) [49], unscented Kalman filter [50], cubature Kalman filter [51], and iterated approaches, such as iterative EKF and iterative PLF [49], which may be better alternatives to the EKF in some situations.

2.3 Bayesian smoothing

In Bayesian smoothing, the objective is to compute the smoothing density $p(x_k|z_{1:K})$, which is the distribution of the state x_k at time step k after receiving the measurements up to and including a time step K where $K > k$. The backward recursive equation for computing the smoothed densities $p(x_k|z_{1:K})$ for any $k < K$ is given by

$$p(x_k|z_{1:K}) = p(x_k|z_{1:k}) \int \frac{p(x_{k+1}|x_k)p(x_{k+1}|z_{1:K})}{p(x_{k+1}|z_{1:k})} dx_{k+1}, \tag{2.15}$$

where $p(x_k|z_{1:k})$ is the filtering density at time step k and $p(x_{k+1}|z_{1:k})$ is the predicted density at time step $k+1$.

Rauch-Tung-Striebel smoother

The Rauch-Tung-Striebel smoother (RTSS) [52] gives the closed-form smoothing solution to linear Gaussian models. The smoothing density at time step k is Gaussian and denoted as

$$p(x_k|z_{1:K}) = \mathcal{N}(x_k; \bar{x}_{k|K}, P_{k|K}), \tag{2.16}$$

where $\bar{x}_{k|K}$ and $P_{k|K}$ are the mean and the covariance of the smoothed density, respectively. The parameters of the distribution in (2.16) can be computed with the following backward recursion equations [52]:

$$G_k = P_{k|k} F^{\mathrm{T}} P_{k+1|k}^{-1}, \tag{2.17a}$$

$$\bar{x}_{k|K} = \bar{x}_{k|k} + G_k(\bar{x}_{k+1|K} - \bar{x}_{k+1|k}), \tag{2.17b}$$

$$P_{k|K} = P_{k|k} + G_k(P_{k+1|K} - P_{k+1|k})G_k^{\mathrm{T}}, \tag{2.17c}$$

where G_k is the smoothing gain, $\bar{x}_{k|k}$ and $P_{k|k}$ are the mean and covariance computed by the Kalman update, and $\bar{x}_{k+1|k}$ and $P_{k+1|k}$ are the mean and covariance computed by the Kalman prediction. The recursion is started from the last time step K with $\bar{x}_{K|K}$ and $P_{K|K}$, computed using Kalman filtering.

Backward simulation

Bayesian inference in dynamical systems often requires generating samples from the posterior density $p(x_{0:K}|z_{1:K})$. This problem can be addressed using backward simulation [45]. An alternative recursion for the posterior density $p(x_{k:K}|z_{1:K})$ without marginalizing out the states after time step k is

$$\begin{aligned} p(x_{k:K}|z_{1:K}) &= p(x_k|x_{k+1}, z_{1:K})p(x_{k+1:K}|z_{1:K}), \\ &= p(x_k|x_{k+1}, z_{1:k})p(x_{k+1:K}|z_{1:K}), \end{aligned} \tag{2.18}$$

where $p(x_k|x_{k+1}, z_{1:k})$ is usually referred to as the backward kernel. The recursion (2.18) evolves backward in time and starts with the filtering density $p(x_K|z_{1:K})$ at time step K.

Using (2.18), the posterior density $p(x_{0:K}|z_{1:K})$ can be factorized as

$$p(x_{0:K}|z_{1:K}) = p(x_0) \left(\prod_{j=1}^{K-1} p(x_j|x_{j+1}, z_{1:j}) \right) p(x_K|z_{1:K}). \tag{2.19}$$

Initially, a sample is generated from the filtering density $p(x_K|z_{1:K})$ at time step K,

$$x_K \sim p(x_K|z_{1:K}). \tag{2.20}$$

The so-called backward trajectory is then constructed by successively augmenting x_K with samples generated from $p(x_k|x_{k+1}, z_{1:k})$,

$$x_k \sim p(x_k|x_{k+1}, z_{1:k}), \tag{2.21}$$

for $k = K-1, \ldots, 1$. At time step 0, the backward trajectory is augmented with a sample generated from the initial prior density $p(x_0)$. After a complete backward sweep, the backward trajectory $x_{0:k}$ can be regarded as a realization from the posterior density $p(x_{0:K}|z_{1:K})$.

For linear Gaussian models, the backward kernel density $p(x_k|x_{k+1}, z_{1:k})$ is

a Gaussian

$$p(x_k|x_{k+1}, z_{1:k}) = \mathcal{N}(x_k; \mu_k, M_k), \tag{2.22}$$

with

$$\mu_k = \bar{x}_{k|k} + G_k(x_{k+1} - F_k\bar{x}_{k|k}), \tag{2.23a}$$
$$M_k = P_{k|k} - G_kFP_{k|k}, \tag{2.23b}$$

where G_k is the smoothing gain (2.17a) in RTSS, and $\bar{x}_{k|k}$ and $P_{k|k}$ are the mean and covariance of the filtering density $p(x_k|z_{1:k})$ computed by a Kalman filter.

2.4 Bayesian inference in factor graphs

In Bayesian statistics, factor graphs are used to represent factorizations of a probability distribution function, enabling efficient computations [44, Chapter 8]. For inference in factor graphs, the objective is to compute the posterior distributions of some variables of interest.

Consider a joint distribution over variables $\mathbf{x} = (x_1, \ldots, x_n)$ that can be factorized into a product of factors

$$p(\mathbf{x}) = \prod_s f_s(\mathbf{x}_s), \tag{2.24}$$

where $\mathbf{x}_s$ represents variables involved in factor $f_s(\cdot)$. The factor graph representation of the joint distribution (2.24) has one variable node for each variable x_i with $i \in \{1, \ldots, n\}$ and one factor node for each factor $f_s(\cdot)$. There are undirected links connecting each factor node to all the variable nodes on which that factor depends. Every factor graph is bipartite since all links go between nodes of opposite types. For example, consider a distribution that is expressed in terms of the factorization

$$p(x_1, x_2, x_3) = f_a(x_1, x_2)f_b(x_1, x_2)f_c(x_2, x_3). \tag{2.25}$$

Its factor graph representation is shown in Fig. 2.1.

State space models can be represented using factor graphs. For example, for the dynamical system described in Section 2.1, the joint distribution of

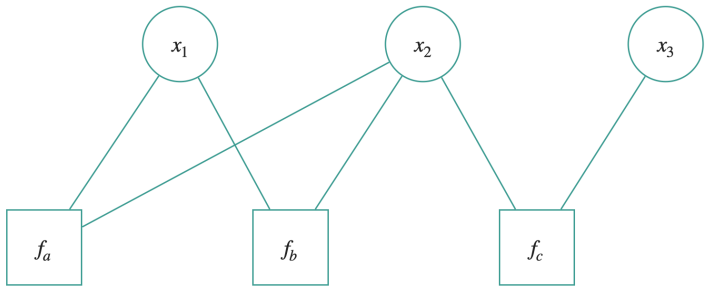

Figure 2.1: Factor graph representation of (2.25).

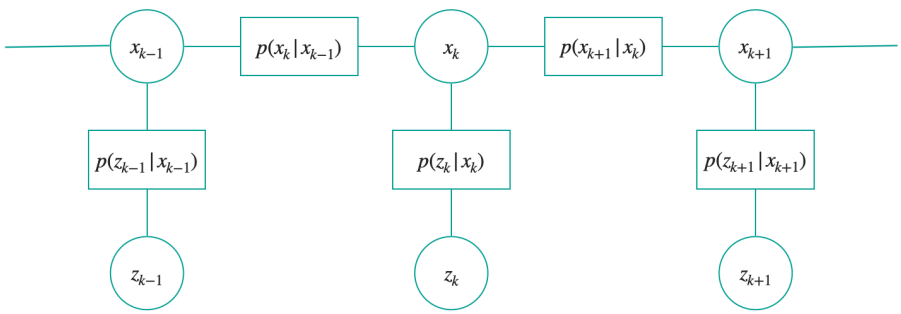

Figure 2.2: Factor graph representation of (2.26) for time step $k-1$, k and $k+1$.

states and measurements is

$$p(x_0, x_1, \ldots, x_K, z_1, \ldots, z_K) = p(x_0) \prod_{k=1}^{K} p(x_k|x_{k-1}) \prod_{k=1}^{K} p(z_k|x_k), \qquad (2.26)$$

and its corresponding factor graph representation is illustrated in Fig. 2.2.

The sum-product algorithm

The sum-product algorithm (SPA), also called belief propagation (BP), is a technique to evaluate local marginals over nodes or subsets of nodes in a factor graph [53]. The SPA involves passing messages on the factor graph. If the variables of interest are all discrete, the message from factor node f_s to

variable node x is

$$\mu_{f_s \to x}(x) = \sum_{\mathbf{x}_s \setminus x} f_s(\mathbf{x}_s) \prod_{m \in \mathrm{ne}(f_s) \setminus x} \mu_{x_m \to f_s}(x_m), \tag{2.27}$$

and the message from variable node x to factor node f_s is

$$\mu_{x \to f_s}(x) = \prod_{l \in \mathrm{ne}(x) \setminus f_s} \mu_{f_l \to x}(x), \tag{2.28}$$

where $\mathrm{ne}(f_s)$ denotes the set of variable nodes that are neighbours of the factor node f_s, and $\mathrm{ne}(x)$ denotes the set of factor nodes that are neighbours of the variable node x. Note that if some variables in $\mathbf{x}_s$ are continuous, then the corresponding summations over these variables in (2.27) become integrals.

To initialize the messages in a factor graph without loops (i.e. a tree), we can pick an arbitrary node as a root node and start all the messages at the leaf nodes. In particular, the message from leaf variable node x to factor node f_s is

$$\mu_{x \to f_s}(x) = 1, \tag{2.29}$$

and the message from leaf factor node f_s to variable node x is

$$\mu_{f_s \to x}(x) = f_s(x). \tag{2.30}$$

After sending out all the messages, one can compute the marginal distribution of all variables of interest. The marginal distribution of variables $\mathbf{x}_s$ is proportional to the product of the factor $f_s(\mathbf{x}_s)$ and the messages from the variables $\mathbf{x}_s$

$$p(\mathbf{x}_s) \propto f_s(\mathbf{x}_s) \prod_{m \in \mathrm{ne}(f_s)} \mu_{x_m \to f_s}(x_m), \tag{2.31}$$

where normalization is required to make sure that $p(\mathbf{x}_s)$ is a proper probability mass or density function.

Exact and approximate inference

The SPA provides efficient and exact solution to the marginal distributions in tree-structured factor graphs. For example, for linear Gaussian state space models, the Kalman filter can be derived by applying the SPA to the factor

graph in Fig. 2.2. However, for many practical applications, the factor graph representation of the joint distribution of variables of interest has loops. One simple approach to approximate inference in graphs with loops, known as loopy BP, is to apply the SPA even though there is usually no guarantee of convergence [53]. To initiate the message passing in graphs with loops, one common way is to set the initial messages as unit functions. Loopy BP has proven to be very effective in MOT [54]–[58], cooperative localization [59], [60], simultaneous localization and mapping [61], and many other applications. These methods build on the formulation in [62], for which convergence to a unique solution has been shown to be guaranteed.

CHAPTER 3

Random finite sets and metrics

Random finite sets (RFSs) are set-valued random variables whose elements and cardinality are random. This chapter covers the basic concepts and relevant properties of RFSs that will be used in the rest of the thesis. The reader is referred to [14], [15] for further readings on this topic. Common metrics used for tracking performance evaluation are also introduced.

3.1 Definition

Let $\mathfrak{Y}$ be an underlying space, such as a single-object state space $\mathfrak{X}$. The state of a single-object contains the information of interest in the object, which is usually represented by an n-dimensional vector x in some Euclidean space $\mathbb{R}^n$. In general, $\mathfrak{Y}$ can be any Hausdorff, locally compact, and completely separable topological space [15, Appendix B]. An RFS is a random variable on the set $\mathcal{F}(\mathfrak{Y})$ of all the finite subsets of $\mathfrak{Y}$. In this thesis, the following three types of spaces that meet these properties are considered:

- The single-object state space $\mathfrak{X}$, which usually contains the object's location and any other motion or extent parameters of interest.

- The single-measurement space $\mathbb{R}^{n_z}$ for measurement models based on detections, where n_z is the dimension of the single-measurement vector.
- The single-object trajectory space $\mathfrak{T}$, which contains the information that characterizes the trajectory of an object.

3.2 Multi-object statistics

This section covers basic multi-object concepts and statistics in terms of the theory of RFSs.

Set integral and multi-object densities

Given a real-valued function $f(\cdot)$ on the space $\mathcal{F}(\mathbb{R}^{n_x})$, its set integral is defined as

$$\int f(\mathbf{x})\delta\mathbf{x} = \sum_{i=0}^{\infty} \frac{1}{i!} \int f(\{x_1, \ldots, x_i\}) dx_1 \cdots dx_i. \tag{3.1}$$

The set integral sums over all possible cardinalities and all possible object states for each cardinality. A function $f(\cdot)$ is a multi-object density, if $f(\cdot) \geq 0$ and its set integral is one.

Convolution formula for multi-object densities

Let $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathcal{F}(\mathfrak{X})$ be n statistically independent RFSs with multi-object densities $f_1(\cdot), \ldots, f_n(\cdot)$, respectively. Then the multi-object density $f(\cdot)$ of the union $\mathbf{y} = \mathbf{x}_1 \cup \cdots \cup \mathbf{x}_n$ is given by the convolution formula

$$f(\mathbf{y}) = \sum_{\mathbf{y}=\mathbf{x}_1 \uplus \cdots \uplus \mathbf{x}_n} \prod_{i=1}^{n} f_i(\mathbf{x}_i), \tag{3.2}$$

where $\uplus$ stands for disjoint union and the summation is taken over all mutually disjoint (and possibly empty) subsets $\mathbf{x}_1, \ldots, \mathbf{x}_n$ whose union is $\mathbf{y}$.

Probability generating functionals

Let h be a test function on the single-object space such that $h(x)$ is unitless and $0 \leq h(x) \leq 1$. The probability generating functionals (PGFL) of an RFS

$\mathbf{x}$ is defined as

$$G[h] = \int h^{\mathbf{x}} f(\mathbf{x}) \delta\mathbf{x}, \tag{3.3}$$

where $f(\cdot)$ is the multi-object density of $\mathbf{x}$ and $h^{\mathbf{x}}$ is the power functional defined as

$$h^{\mathbf{x}} = \begin{cases} 1 & \text{if} \quad \mathbf{x} = \varnothing \\ \prod_{x \in \mathbf{x}} h(x) & \text{if} \quad \mathbf{x} \neq \varnothing. \end{cases} \tag{3.4}$$

The PGFL of an RFS completely characterizes its multi-object density, and it is useful for deriving multi-object filters.

Cardinality distribution

The cardinality of a set $\mathbf{x}$, denoted by $|\mathbf{x}|$, is the number of elements in the set. The number of elements of an RFS is a random variable and characterized by a probability mass function, which is referred to as the cardinality distribution. The cardinality distribution of an RFS with multi-object density $f(\cdot)$ is given by

$$\rho(n) = \frac{1}{n!} \int f(\{x_1, \ldots, x_n\}) dx_1 \cdots dx_n. \tag{3.5}$$

The value of $\rho(n)$ is the probability that $\mathbf{x}$ contains n elements.

Probability hypothesis density

The probability hypothesis density (PHD), also referred to as the intensity function, of an RFS with multi-object density $f(\cdot)$ is defined on the single-object space $\mathfrak{X}$ as

$$\begin{aligned} D(x) &= \int f(\{x\} \cup \mathbf{x}) \delta\mathbf{x} \\ &= \sum_{i=0}^{\infty} \frac{1}{i!} \int f(\{x, x_1, \ldots, x_i\}) dx_1 \cdots dx_i. \end{aligned} \tag{3.6}$$

The integral of the PHD in a region $\mathbb{A} \subseteq \mathfrak{X}$ yields the expected number $\hat{N}_{\mathbb{A}}$ of objects in this region

$$\hat{N}_{\mathbb{A}} = \int_{\mathbb{A}} D(x) dx. \tag{3.7}$$

3.3 Important multi-object processes

This section introduces some standard types of multi-object processes that will be used in this thesis. They are the Poisson RFSs, the Bernoulli RFSs, the multi-Bernoulli (MB) RFSs and the MB mixture (MBM) RFSs.

Poisson RFSs

In a Poisson RFS, also referred to as Poisson point process (PPP), the cardinality of the set is Poisson distributed and, for each cardinality, its elements are independent and identically distributed. The multi-object density for a Poisson RFS $\mathbf{x} = \{x_1, \ldots, x_n\}$ is given by

$$f(\mathbf{x}) = e^{-\lambda}\lambda^n \prod_{i=1}^{n} p(x_i), \tag{3.8}$$

where $\lambda \geq 0$ is the parameter of the Poisson cardinality distribution and $p(\cdot)$ denotes a single-object density. A Poisson RFS can be characterized by either its PHD/intensity function $D(x) = \lambda p(x)$ or by λ and $p(\cdot)$. Therefore, the multi-object density of the Poisson RFS $\mathbf{x} = \{x_1, \ldots, x_n\}$ can be alternatively expressed as

$$f(\mathbf{x}) = e^{-\int D(x)dx} \prod_{i=1}^{n} D(x_i). \tag{3.9}$$

Bernoulli RFSs

In a Bernoulli RFS, the cardinality of the set is Bernoulli distributed. The multi-object density for a Bernoulli RFS $\mathbf{x}$ is given by

$$f(\mathbf{x}) = \begin{cases} 1 - r & \text{if} \quad \mathbf{x} = \varnothing \\ rp(x) & \text{if} \quad \mathbf{x} = \{x\}, \end{cases} \tag{3.10}$$

where r is the probability of existence and $p(\cdot)$ is the single-object density of the object if it exists. A Bernoulli RFS is characterized by r and $p(\cdot)$. Every Bernoulli density can be expanded into a Bernoulli mixture density containing two terms with deterministic probability of existence $r = 0$ and $r = 1$, respectively.

Multi-Bernoulli RFSs

An MB RFS corresponds to the union of a finite number n of independent Bernoulli RFSs. The multi-object density of an MB RFS $\mathbf{x}$ is given by

$$f(\mathbf{x}) = \sum_{\mathbf{x}_1 \uplus \cdots \uplus \mathbf{x}_n = \mathbf{x}} \prod_{i=1}^{n} f_i(\mathbf{x}_i), \tag{3.11}$$

where $f_i(\cdot)$ is the ith Bernoulli component characterized by a probability r_i of existence and a single object density $p_i(\cdot)$. Therefore, an MB RFS is characterized by the set of parameters $\{(r_1, p_1(\cdot)), \ldots, (r_n, p_n(\cdot))\}$.

Multi-Bernoulli mixture RFSs

An MB mixture (MBM) RFS is a weighted sum of MB RFSs. The multi-object density for an MBM RFS $\mathbf{x}$ is given by

$$f(\mathbf{x}) = \sum_{h=1}^{\mathcal{H}} w^h \sum_{\mathbf{x}_1 \uplus \cdots \uplus \mathbf{x}_n = \mathbf{x}} \prod_{i=1}^{n} f_i^h(\mathbf{x}_i), \tag{3.12}$$

where $w^1, \ldots, w^{\mathcal{H}}$ are non-negative weights such that $\sum_{i=1}^{\mathcal{H}} w^h = 1$ and $f_i^h(\mathbf{x}_i)$ are Bernoulli RFS densities for $i = 1, \ldots, n$ and $h = 1, \ldots, \mathcal{H}$.

An MBM_{01} is defined as an MBM for which in each mixture component, every Bernoulli component has probability of existence of either 0 or 1. Any MB or MBM distribution can be represented as a MBM_{01} with a greater number of mixture components.

3.4 Metrics on object tracking

This section introduces some standard single object metrics and MOT metrics for sets of objects.

Definition

It is important that the tracking performance is measured consistently. Metrics ensure that the "distance" between the estimate and the ground truth is

mathematically meaningful. A metric d on a space $\mathfrak{Y}$ is a distance function

$$d : \mathfrak{Y} \times \mathfrak{Y} \to [0, \infty), \tag{3.13}$$

where for all $x, y, z \in \mathfrak{Y}$ the following four conditions are satisfied:

$$\text{Non-negativity:} \quad d(x, y) \geq 0. \tag{3.14a}$$

$$\text{Identity of indiscernible:} \quad d(x, y) = 0 \Leftrightarrow x = y. \tag{3.14b}$$

$$\text{Symmetry:} \quad d(x, y) = d(y, x). \tag{3.14c}$$

$$\text{Triangle inequality:} \quad d(x, z) \leq d(x, y) + d(y, z). \tag{3.14d}$$

Single-object metrics

The Euclidean metric of order p between a state x and a corresponding estimate $\hat{x}$ is defined as

$$d_p(x, \hat{x}) = \|x - \hat{x}\|_p, \tag{3.15}$$

where $\|\cdot\|_p$ denotes the L^p-norm.

For performance evaluation of extended object estimates with ellipsoidal extents, a comparison study in [63] has shown that the Gaussian Wasserstein distance (GWD) is a good choice. The GWD between an extended object state $\xi = (x, E)$, where E is positive symmetric matrix used to represent the extent of an elliptic object, and a corresponding estimate $\hat{\xi} = \left(\hat{x}, \hat{E}\right)$ of order 2 is defined as [64]

$$d_{\mathrm{GW}}(\xi, \hat{\xi}) = \left(\|Hx - H\hat{x}\|_2^2 + \operatorname{trace}\left(E + \hat{E} - 2\left(E\hat{E}\right)^{\frac{1}{2}} \right) \right)^{\frac{1}{2}}, \tag{3.16}$$

where the observation matrix H picks out the position from the kinematic state vector and $E^{1/2}$ denotes the matrix square root of E.

Multi-object metrics

In MOT, it is important to measure not only the state estimation error but also the cardinality estimation error due to misdetections and false detections. The different cardinality-related errors include:

- Misdetections: true objects for which there are no corresponding esti-

mates.

- False detections: object estimates for which there are no corresponding true objects.

An MOT metric computes the "distance" between a set of object states $\mathbf{x}$ and a corresponding set estimate $\hat{\mathbf{x}}$. It is assumed that a single object metric $d(x, \hat{x})$ is given. Further, let

$$d^{(c)}(x, \hat{x}) = \min(c; d(x, \hat{x})) \tag{3.17}$$

denote the distance cut-off at a distance $c > 0$ where c is a parameter.

OSPA

The Optimal Sub-Pattern Assignment (OSPA) multi-object metric is defined as [65], [66]

$$d_p^{(c)}(\mathbf{x}, \hat{\mathbf{x}}) = \begin{cases} \left(\frac{1}{|\hat{\mathbf{x}}|} \left(\min_{\pi \in \Pi_{|\hat{\mathbf{x}}|}} \sum_{i=1}^{|\mathbf{x}|} d^{(c)} \left(x_i, \hat{x}_{\pi(i)} \right)^p + c^p \left(|\hat{\mathbf{x}}| - |\mathbf{x}| \right) \right) \right)^{\frac{1}{p}} & \text{if } |\hat{\mathbf{x}}| \geq |\mathbf{x}| \\ \left(\frac{1}{|\mathbf{x}|} \left(\min_{\pi \in \Pi_{|\mathbf{x}|}} \sum_{i=1}^{|\hat{\mathbf{x}}|} d^{(c)} \left(\hat{x}_i, x_{\pi(i)} \right)^p + c^p \left(|\mathbf{x}| - |\hat{\mathbf{x}}| \right) \right) \right)^{\frac{1}{p}} & \text{if } |\hat{\mathbf{x}}| < |\mathbf{x}|, \end{cases} \tag{3.18}$$

where $c > 0$, $1 \leq p < \infty$, and Π_i is the set of permutations of the set of integers $\{1, \ldots, i\}$ for any $i \in \mathbb{N}$ and for any element $\pi = (\pi(1), \ldots, \pi(i)) \in \Pi_i$. The parameter p determines the severity of penalizing the outliers in the localization component. The OSPA metric (for the case $|\hat{\mathbf{x}}| \geq |\mathbf{x}|$) can be decomposed into two parts:

- Normalized sum of state errors: $\frac{1}{|\hat{\mathbf{x}}|} \sum_{i=1}^{|\mathbf{x}|} d^{(c)} \left(x_i, \hat{x}_{\pi(i)} \right)^p$.
- Normalized cardinality error: $\frac{1}{|\hat{\mathbf{x}}|} c^p (|\hat{\mathbf{x}}| - |\mathbf{x}|)$.

GOSPA

The OSPA metric only accounts for localization errors for the objects in the smallest set and cardinality mismatch; this is not desired in traditional MOT performance evaluation. For example, OSPA does not encourage trackers to

have as few misdetections and false detections as possible. As a solution to this, the generalized OSPA (GOSPA) multi-object metric was proposed in [67], which is able to penalize localization errors for properly detected objects, misdetections and false detection. The GOSPA metric is defined as

$$d_p^{(c,\alpha)}(\mathbf{x},\hat{\mathbf{x}}) = \begin{cases} \left(\min_{\pi\in\Pi_{|\hat{\mathbf{x}}|}} \sum_{i=1}^{|\mathbf{x}|} d^{(c)}\left(x_i,\hat{x}_{\pi(i)}\right)^p + \frac{c^p}{\alpha}(|\hat{\mathbf{x}}|-|\mathbf{x}|)\right)^{\frac{1}{p}} & \text{if } |\hat{\mathbf{x}}| \geq |\mathbf{x}| \\ \left(\min_{\pi\in\Pi_{|\mathbf{x}|}} \sum_{i=1}^{|\hat{\mathbf{x}}|} d^{(c)}\left(\hat{x}_i,x_{\pi(i)}\right)^p + \frac{c^p}{\alpha}(|\mathbf{x}|-|\hat{\mathbf{x}}|)\right)^{\frac{1}{p}} & \text{if } |\hat{\mathbf{x}}| < |\mathbf{x}|, \end{cases} \tag{3.19}$$

where $c > 0$, $1 \leq p < \infty$ and $0 < \alpha \leq 2$.

When the GOSPA metric is used for MOT performance evaluation to measure the "distance" between the true set of object states $\mathbf{x}$ and the estimated set of object states $\hat{\mathbf{x}}$, it is most appropriate to set $\alpha = 2$. In this case, the GOSPA metric can be re-written as

$$d_p^{(c,2)}(\mathbf{x},\hat{\mathbf{x}}) = \left(\min_{\theta\in\Theta^{(|\mathbf{x}|,|\hat{\mathbf{x}}|)}} \sum_{(i,j)\in\theta} d^{(c)}\left(x^i,\hat{x}^j\right)^p + \frac{c^p}{2}(|\mathbf{x}|-|\theta|+|\hat{\mathbf{x}}|-|\theta|)\right)^{\frac{1}{p}}, \tag{3.20}$$

where $\Theta^{(|\mathbf{x}|,|\hat{\mathbf{x}}|)}$ is the set of all possible 2D assignments. An assignment set θ between the sets $\{1,\ldots,|\mathbf{x}|\}$ and $\{1,\ldots,|\hat{\mathbf{x}}|\}$ is a set that has the following properties: $\theta \subseteq \{1,\ldots,|\mathbf{x}|\} \times \{1,\ldots,|\hat{\mathbf{x}}|\}$, $(i,j),(i,j') \in \theta \Rightarrow j = j'$ and $(i',j),(i,j) \in \theta \Rightarrow i = i'$ where the last two properties ensure that every i and j gets at most one assignment. Equation (3.20) facilitates the decomposition of the GOSPA metric into three parts:

- Sum of state errors (to the p-th power): $\sum_{(i,j)\in\theta} d^{(c)}\left(x^i,\hat{x}^j\right)^p$.
- Misdetection error: $\frac{c^p}{2}(|\mathbf{x}|-|\theta|)$.
- False detection error: $\frac{c^p}{2}(|\hat{\mathbf{x}}|-|\theta|)$.

Compared to the OSPA metric in (3.18), there is no normalization factor $\max(|\mathbf{x}|,|\hat{\mathbf{x}}|)$ and an additional parameter α has been introduced. Setting $\alpha = 1$ gives the OSPA metric without normalization. Moreover, it has been shown in [68] that the spooky effect arises in optimal estimation of multiple objects with the OSPA metric; however, this is not the case for the GOSPA metric

with $\alpha = 2$. Furthermore, an analysis in [69] shows that GOSPA ($\alpha = 2$) is more suitable for metric-driven sensor management for usual multiple target estimation tasks than OSPA, where the optimal action should be in principle taken for each sensor independently.

CHAPTER 4

Multi-object modelling

The formulation of an MOT problem needs systematic formal modelling for multi-object dynamics and measurements. This chapter introduces the multi-object dynamic and measurement models used in this thesis.

4.1 Multi-object dynamic model

Multi-object dynamic models should include models for the motion of individual objects as well as models for object appearance and disappearance to handle an unknown and time-varying number of objects. In MOT literature, object appearance and disappearance are often referred to as object birth and death, respectively. The multi-object dynamic model $\pi(\mathbf{x}_{k+1}|\mathbf{x}_k)$ used in this thesis is based on the following assumptions:

- Single object with state x_k at time step k moves to a new state x_{k+1} at time step $k+1$ with a Markov transition density $\pi(x_{k+1}|x_k)$.

- A single object with state x_k at time step k has a probability $P^S(x_k)$ of surviving into time step $k+1$.

- The RFS of objects at time step $k+1$ is the union of the RFS of objects that survive from time step k to time step $k+1$, denoted $\mathbf{x}_{k+1}^{S}$, and the RFS of newborn objects, denoted $\mathbf{x}_{k+1}^{B}$,

$$\mathbf{x}_{k+1} = \mathbf{x}_{k+1}^{S} \cup \mathbf{x}_{k+1}^{B}. \tag{4.1}$$

- Object birth, object death, and object motion are conditionally independent of the previous multi-object state.

Birth models

The RFS models for object birth include a model for the number of objects expected to be born, i.e. a birth cardinality distribution, and a model for the states of the newborn objects, i.e. a birth state density. In many cases, a new object can appear anywhere in the surveillance area, so it is important to use a spatial distribution that covers the entire area. In some cases, objects may only be born in specific areas, and it is then more suitable to use a spatial distribution localized to those areas.

Poisson birth model

To model the number of new objects that appear at a time step as being Poisson distributed is the most common choice in MOT. When using RFSs to model MOT, a Poisson number of new objects means that the object birth is modelled as a Poisson RFS. For a Poisson birth model, at time step k a possibly empty set of newborn objects $\mathbf{x}_k^B$ appears, distributed according to a Poisson RFS with intensity $\lambda_k^B(x_k)$. That the set is possibly empty means that it is possible that no object is born at all in that time step.

The birth cardinality distribution is Poisson, with parameter $\bar{\lambda}_k^B$,

$$\rho(n_k^B = j) = e^{-\bar{\lambda}_k^B} \frac{\left(\bar{\lambda}_k^B\right)^j}{j!}, \tag{4.2}$$

where

$$\bar{\lambda}_k^B = \int \lambda_k^B(x_k) dx_k, \tag{4.3}$$

which is the expected number of births. For a non-empty set of new objects,

their states are independent and identically distributed (i.i.d.) with density

$$\frac{\lambda_k^B(x_k)}{\bar{\lambda}_k^B}. \tag{4.4}$$

The most common representation of the Poisson RFS birth intensity is an unnormalized density mixture,

$$\lambda_k^B(x_k) = \sum_{i=1}^{N_k^B} w_k^{B,i} p_k^{B,i}(x_k), \tag{4.5}$$

where $\sum_{i=1}^{N_k^B} w_k^{B,i} = \bar{\lambda}_k^B$ and $\int p_k^{B,i}(x_k)dx_k = 1$ for $i \in \{1, \ldots, N_k^B\}$. In this case, the birth parameters are the set of weights and densities,

$$\left\{ \left(w_k^{B,i}, p_k^{B,i}(\cdot) \right) \right\}_{i=1}^{N_k^B}. \tag{4.6}$$

For Gaussian implementations, it is suitable to let the Poisson RFS intensity be a Gaussian mixture,

$$\lambda_k^B(x_k) = \sum_{i=1}^{N_k^B} w_k^{B,i} \mathcal{N}\left(x_k; \bar{x}_k^{B,i}, P_k^{B,i} \right), \tag{4.7}$$

with each Gaussian component covering an area that new objects may appear.

Bernoulli and multi-Bernoulli birth models

Object birth models can be also built upon Bernoulli RFSs. For a Bernoulli birth model, at time step k an object is born with probability r_k^B (i.e. no object with probability $1 - r_k^B$), and if an object is born, the initial object state x_k has density $p_k^B(x_k)$. In a Bernoulli birth model, the probability of birth r_k^B gives the birth cardinality distribution, which is a Bernoulli,

$$\rho(n_k^B = j) = \begin{cases} 1 - r_k^B & \text{if } j = 0 \\ r_k^B & \text{if } j = 1 \\ 0 & \text{otherwise,} \end{cases} \tag{4.8}$$

where n_k^B denotes the number of births at time step k.

It is more general to use an MB birth model, which is the union of independent Bernoulli births. The parameters of an MB birth are the set of Bernoulli birth parameters,

$$\left\{\left(r_k^{B,i}, p_k^{B,i}(\cdot)\right)\right\}_{i=1}^{N_k^B}. \tag{4.9}$$

For an MB birth model, the expected number of births is given by the sum of the probabilities of birth $\sum_{i=1}^{N_k^B} r_k^{B,i}$. The MB birth cardinality distribution is given by the convolution of the Bernoulli birth cardinality distributions. In case $r_k^{B,i} = r$ for all i, then the birth cardinality distribution is a Binomial,

$$\rho(n_k^B = j) = \begin{cases} \binom{n_k^B}{j} r^j (1-r)^{n_k^B - j} & \text{if } n_k^B = \{0, 1, \ldots, N_k^B\} \\ 0 & \text{if } n_k^B > N_k^B. \end{cases} \tag{4.10}$$

Compared to the Poisson birth model, the MB birth model bounds a priori the number of objects that can appear in the surveillance area. For linear and Gaussian models, it is suitable to let the Bernoulli birth densities be Gaussian,

$$p_k^{B,i}(x_k) = \mathcal{N}(x_k; \bar{x}_k^{B,i}, P_k^{B,i}). \tag{4.11}$$

Single-object dynamic models

This section introduces two common 2D single-object dynamic models used in this thesis. A survey of dynamic models for object tracking is given in [70].

Nearly constant velocity model

In this model, the single object state at time step k is $x_k = [p_{k,x}, v_{k,x}, p_{k,y}, v_{k,y}]^{\mathrm{T}}$ where $[p_{k,x}, p_{k,y}]^{\mathrm{T}}$ is the position vector and $[v_{k,x}, v_{k,y}]^{\mathrm{T}}$ is the velocity vector. The single object state transition density from the object state x_k at time step k to the object state x_{k+1} at time step $k+1$ is

$$\pi(x_{k+1}|x_k) = \mathcal{N}(x_{k+1}; Fx_k, Q), \tag{4.12}$$

where

$$F = I_2 \otimes \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}, \tag{4.13a}$$

$$Q = \sigma_{\dot{v}}^2 I_2 \otimes \begin{bmatrix} T^3/3 & T^2/2 \\ T^2/2 & T \end{bmatrix}. \tag{4.13b}$$

Here T is the sampling time, $\sigma_{\dot{v}}$ is a parameter controlling the process noise level, I_n is the identity matrix of size n, and $\otimes$ denotes the Kronecker product.

Coordinate turn model with polar velocity

In this model, the single object state at time step k is $x_k = [p_{k,x}, p_{k,y}, v_k, \phi_k, \omega_k]^{\mathrm{T}}$ where $[p_{k,x}, p_{k,y}]^{\mathrm{T}}$ is the position vector, v_k is the polar velocity, ϕ_k is the heading and ω_k is the turn rate. By linearization first and then discretization, the relation between the object state x_k at time step k and the object state x_{k+1} at time step $k+1$ can be written as

$$x_{k+1} = \begin{bmatrix} p_{k,x} + (2v_k/\omega_k)\sin(\omega_k T/2)\cos(\phi_k + \omega_k T/2) \\ p_{k,y} + (2v_k/\omega_k)\sin(\omega_k T/2)\sin(\phi_k + \omega_k T/2) \\ v_k \\ \phi_k + \omega_k T \\ \omega_k \end{bmatrix} + w_k, \tag{4.14}$$

where w_k is white noise with covariance

$$Q = \mathrm{blkdiag}\left(\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, T^2\sigma_{\dot{v}}^2, \begin{bmatrix} T^3/3 & T^2/2 \\ T^2/2 & T \end{bmatrix} \sigma_{\dot{\omega}}^2 \right), \tag{4.15}$$

where blkdiag denotes block diagonalization, and $\sigma_{\dot{v}}^2$ and $\sigma_{\dot{\omega}}^2$ are parameters controlling the process noise level.

4.2 Multi-object measurement model

Multi-object measurement models should include models for the measurements originating from individual objects as well as models for clutter measurements, i.e. measurements not originating from objects. This section introduces the standard (single-sensor) multi-object measurement model $g(\mathbf{z}_k|\mathbf{x}_k)$ with a generalized measurement model [21], where each object generates an independent set of measurements. This measurement model is based on the following common assumptions:

- The measurement set $\mathbf{z}_k$ at time step k is an RFS consisting of the RFS

$\mathbf{z}_k^O(\mathbf{x}_k)$ of measurements generated by the set of objects with states $\mathbf{x}_k$ and the RFS $\mathbf{z}_k^C$ of clutter measurements, i.e.

$$\mathbf{z}_k = \mathbf{z}_k^O(\mathbf{x}_k) \cup \mathbf{z}_k^C. \tag{4.16}$$

- The RFSs $\mathbf{z}_k^O(\mathbf{x}_k)$ and $\mathbf{z}_k^C$ are statistically independent.
- No measurement is generated by more than one object.
- Given a set $\mathbf{x}_k$ of objects, each object $x_k \in \mathbf{x}_k$ is either detected with probability $P^D(x_k)$ and generates a set of measurements $\mathbf{z}_k^O(x_k)$ with conditional density $g(\mathbf{z}_k^O(x_k)|x_k)$, or missed with probability $1-P^D(x_k)$.
- The RFS $\mathbf{z}_k^C$ is a Poisson RFS with intensity $\lambda_k^C(\cdot)$.

Both point and extended object measurement models are special cases of this generalized measurement model.

Point objects

In point object tracking, it is assumed that each object generates at most a single measurement per time step, i.e. a single resolution cell is occupied by an object. In this case, $|\mathbf{z}^O(x)| \leq 1$, and the single-object measurement likelihood for a point object is

$$p(\mathbf{z}^O(x)|x) = \begin{cases} P^D(x)g(z|x) & \mathbf{z}^O(x) = \{z\} \\ 1 - P^D(x) & \mathbf{z}^O(x) = \varnothing \\ 0 & \text{otherwise.} \end{cases} \tag{4.17}$$

Extended objects

In extended object tracking (EOT), each object may generate multiple measurements per time step and the measurements are spatially distributed around the objects, i.e. multiple resolution cells are occupied by an object [71]. In this case, it is commonly assumed that $P^D(x) = 1$ and that the object-generated measurements $\mathbf{z}^O(x)$ are modelled by a Poisson RFS [72]. This models the number of detections as $|\mathbf{z}^O(x)|$ Poisson distributed with a rate $\lambda^O(x)$ that is a function of the object state x. The single-object measurement likelihood for

an extended object is

$$p(\mathbf{z}^O(x)|x) = e^{-\lambda^O(x)}\lambda^O(x)^{|\mathbf{z}^O(x)|}\prod_{z\in\mathbf{z}^O(x)} g(z|x). \tag{4.18}$$

It is also possible to model the object-generated measurements $\mathbf{z}^O(x)$ as a zero-inflated Poisson RFS, and in this case the single-object measurement likelihood is

$$p(\mathbf{z}^O(x)|x) = \begin{cases} P^D(x)e^{-\lambda^O(x)}\lambda^O(x)^{|\mathbf{z}^O(x)|}\prod_{z\in\mathbf{z}^O(x)} g(z|x) & \mathbf{z}^O(x) \neq \varnothing \\ 1 - P^D(x) + P^D(x)e^{-\lambda^O(x)} & \mathbf{z}^O(x) = \varnothing. \end{cases} \tag{4.19}$$

The zero-inflated Poisson RFS may offer more flexibility when modelling object occlusion than the Poisson RFS.

CHAPTER 5

Multi-object conjugate priors for multi-object tracking

Due to the unknown correspondence between measurements and object states, MOT is combinatorial in nature, and thus is highly computationally demanding. MOT conjugate priors are one tool for managing complexity, by exploiting forms which maintain structure through prediction and update steps. This chapter introduces single-object and multi-object conjugacy used in this thesis.

5.1 Conjugate prior

If $\mathscr{L}$ is a class of measurement likelihoods $g(z|x)$, and $\mathscr{F}$ is a class of prior distributions for x, then the class $\mathscr{F}$ is conjugate for $\mathscr{L}$ if

$$p(x|z) \in \mathscr{F},\ \forall\ g(z|x) \in \mathscr{L} \text{ and } \forall\ p(x) \in \mathscr{F}. \tag{5.1}$$

Many examples of conjugate priors can be found from the exponential family of distributions. In each of these cases, the posterior has the same form with the same number of parameters as the prior.

Another common structure is when the prior family contains mixtures of

distributions of a given form. Let $\bar{\mathscr{F}}$ denote the family of mixtures of distributions in $\mathscr{F}$, and let $\bar{\mathscr{L}}$ denote the family of mixtures of likelihoods in $\mathscr{L}$. It can be shown that, if $\mathscr{F}$ is conjugate for the measurement likelihood $\mathscr{L}$, then $\bar{\mathscr{F}}$ is also conjugate for $\mathscr{L}$, and the posterior will contain the same number of mixture components as the prior. Further, $\bar{\mathscr{F}}$ is also conjugate for $\bar{\mathscr{L}}$, but the number of mixture components in the posterior will be the product of the number of components in the prior and in the measurement likelihood. In this case, the posterior is in the same form, but the complexity of the representation grows and eventually approximation becomes necessary [73]. Many examples of this kind can be found in (multiple) object tracking in clutter.

5.2 Single-object conjugate prior

If $\mathscr{L}$ is a class of measurement likelihoods $g(z_k|x_k)$, $\mathscr{T}$ is a class of transition densities $\pi(x_k|x_{k-1})$, and $\mathscr{F}$ is a class of single-object densities for x, then the class $\mathscr{F}$ is single-object conjugate for $\mathscr{L}$ and $\mathscr{T}$ if

$$p(x_k|z_{1:k-1}) \in \mathscr{F},\ \forall\ \pi(x_k|x_{k-1}) \in \mathscr{T} \text{ and } p(x_{k-1}|z_{1:k-1}) \in \mathscr{F}, \tag{5.2}$$

$$p(x_k|z_{1:k}) \in \mathscr{F},\ \forall\ g(z_k|x_k) \in \mathscr{L} \text{ and } p(x_k|z_{1:k-1}) \in \mathscr{F}. \tag{5.3}$$

Single-object conjugate priors can be understood as a generalization of conjugacy, which originally only concerns the Bayesian update, to the whole filtering recursion [73].

Gaussian conjugate prior

Single point object tracking typically concerns the following Bayesian filtering recursions:

$$p\left(x_k|z_{1:k}\right) = \frac{g\left(z_k|x_k\right) p\left(x_k|z_{1:k-1}\right)}{\int g\left(z_k|y_k\right) p\left(y_k|z_{1:k-1}\right) dy_k}, \tag{5.4}$$

$$p\left(x_k|z_{1:k-1}\right) = \int \pi\left(x_k|x_{k-1}\right) p\left(x_{k-1}|z_{1:k-1}\right) dx_{k-1}. \tag{5.5}$$

The Kalman filter, introduced in Section 2.2, is an example of single-object filters based on single-object conjugate prior where both the predicted density

and the posterior density are Gaussian.

Gamma Gaussian inverse-Wishart conjugate prior

For EOT using the Poisson spatial model and the so-called random matrix approach, the objects are assumed to have elliptic shapes, and the Bayesian filtering recursions are

$$p(\xi_k|\mathbf{z}_{1:k}) = \frac{g(\mathbf{z}_k|\xi_k)p(\xi_k|\mathbf{z}_{1:k-1})}{\int g(\mathbf{z}_k|\varsigma_k)p(\varsigma_k|\mathbf{z}_{1:k-1})d\varsigma_k}, \tag{5.6}$$

$$p(\xi_k|\mathbf{z}_{1:k-1}) = \int \pi_k(\xi_k|\xi_{k-1})p(\xi_{k-1}|\mathbf{z}_{1:k-1})d\xi_{k-1}, \tag{5.7}$$

where $\xi_k = (\gamma_k, x_k, X_k)$ is a combination of a Poisson rate γ_k, a kinematic state vector x_k and an extent matrix X_k. Various random matrix approaches with different prediction and update steps have been proposed in the literature; see [4, Section III.A] for an overview. Here the random matrix approach with improved noise modelling proposed in [74] is taken as an example to motivate the use of a gamma Gaussian inverse-Wishart (GGIW) conjugate prior in EOT.

For the improved noise modelling with the number of measurements being Poisson, the measurement likelihood can be factorized as

$$g(\mathbf{z}_k|\xi_k) = e^{-\gamma_k}\gamma_k^{|\mathbf{z}_k|}\prod_{z_k\in\mathbf{z}_k}\mathcal{N}(z_k; Hx_k, \rho X_k + R), \tag{5.8}$$

where ρ is a scaling factor and R is the measurement noise covariance matrix. The Bayesian conjugate prior for an unknown Poisson rate is the gamma distribution. Also, for Gaussian measurements, the conjugate priors for unknown mean and covariance are the Gaussian and the inverse-Wishart distributions, respectively. This motivates the use of a GGIW representation for the object state density,

$$p(\xi_k|\mathbf{z}_{1:k}) = \mathcal{G}(\gamma_k; \alpha_{k|k}, \beta_{k|k})\mathcal{N}(x_k; \bar{x}_{k|k}, P_{k|k})\mathcal{IW}(X_k; \nu_{k|k}, V_{k|k}), \tag{5.9}$$

where $\mathcal{G}(\gamma_k; \alpha_{k|k}, \beta_{k|k})$ represents a gamma distribution with shape $\alpha_{k|k}$ and rate $\beta_{k|k}$, and $\mathcal{IW}(X_k; \nu_{k|k}, V_{k|k})$ represents an inverse-Wishart distribution with degrees of freedom $\nu_{k|k}$ and scale matrix $V_{k|k}$. For a GGIW prior of

the form (5.9), and a measurement likelihood of the form (5.8), the measurement update (5.6) is not analytically tractable; however, the filtering density $p(\xi_k|\mathbf{z}_{1:k})$ can still be approximated as a GGIW. The approximate update can be based on the approximation of non-linear functions of the extent using either matrix square roots [74] or variational Bayesian inference [75].

The GGIW conjugacy also holds for the prediction step. Early works use a Wishart transition density of the extent matrix, where the prediction step has no closed-form solution, but the predicted density (5.7) can be approximated as a GGIW by either applying a simple heuristic or minimizing the KLD [76]. A recent work in [77] shows that the predicted density is guaranteed to be of inverse Wishart form by modelling the state transition density of the object extent as a non-central inverse Wishart distribution.

5.3 Multi-object conjugate prior

If $\mathscr{L}$ is a class of multi-object measurement likelihoods $g(\mathbf{z}_k|\mathbf{x}_k)$, $\mathscr{T}$ is a class of multi-object transition densities $\pi(\mathbf{x}_k|\mathbf{x}_{k-1})$, and $\mathscr{F}$ is a class of multi-object densities for $\mathbf{x}$, then the class $\mathscr{F}$ is multi-object conjugate for $\mathscr{L}$ and $\mathscr{T}$ if

$$p(\mathbf{x}_k|\mathbf{z}_{1:k-1}) \in \mathscr{F},\ \forall\ \pi(\mathbf{x}_k|\mathbf{x}_{k-1}) \in \mathscr{T} \text{ and } p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1}) \in \mathscr{F}, \tag{5.10}$$

$$p(\mathbf{x}_k|\mathbf{z}_{1:k}) \in \mathscr{F},\ \forall\ g(\mathbf{z}_k|\mathbf{x}_k) \in \mathscr{L} \text{ and } p(\mathbf{x}_k|\mathbf{z}_{1:k-1}) \in \mathscr{F}. \tag{5.11}$$

Multi-object conjugacy is a generalization of the single-object conjugacy to consider the multi-object prediction and update,

$$p(\mathbf{x}_k|\mathbf{z}_{1:k-1}) = \int \pi(\mathbf{x}_k|\mathbf{x}_{k-1}) p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1}) \delta\mathbf{x}_{k-1}, \tag{5.12}$$

$$p\left(\mathbf{x}_k|\mathbf{z}_{1:k}\right) = \frac{g\left(\mathbf{z}_k|\mathbf{x}_k\right) p\left(\mathbf{x}_k|\mathbf{z}_{1:k-1}\right)}{\int g\left(\mathbf{z}_k|\mathbf{y}_k\right) p\left(\mathbf{y}_k|\mathbf{z}_{1:k-1}\right) \delta\mathbf{y}_k}. \tag{5.13}$$

Multi-object conjugacy is important for designing MOT algorithms. With the use of multi-object conjugacy, it is convenient to

- express the theoretically exact density (for the given models),
- describe which parameters are needed to represent the density,
- find computationally tractable approximations, and

- analyze the approximation error.

For the multi-object models introduced in Chapter 4, there are two known classes of multi-object conjugate priors for both point and extended objects: the MBM densities [16], [27], [78] for MB (and more general MBM) birth model and the Poisson MBM (PMBM) densities [17]–[19], [21] for Poisson (and more general PMBM) birth model.

Multi-Bernoulli mixture conjugate prior

For the MB(M) birth, the MBM density $\mathcal{MBM}_{k|k}(\mathbf{x}_k)$ is a multi-object conjugate prior to the multi-object transition density $\pi(\mathbf{x}_k|\mathbf{x}_{k-1})$ and measurement model $g(\mathbf{z}_k|\mathbf{x}_k)$ [27]. In other words, if the posterior density at time step $k-1$ is an MBM, then the predicted density at time step k is an MBM,

$$\mathcal{MBM}_{k|k-1}(\mathbf{x}_k) = \int \pi(\mathbf{x}_k|\mathbf{x}_{k-1})\mathcal{MBM}_{k-1|k-1}(\mathbf{x}_{k-1})\delta\mathbf{x}_{k-1}. \tag{5.14}$$

Further, if the prior density at time step k is an MBM, then the Bayes posterior at time step k is an MBM,

$$\mathcal{MBM}_{k|k}\left(\mathbf{x}_k\right) = \frac{g\left(\mathbf{z}_k|\mathbf{x}_k\right)\mathcal{MBM}_{k|k-1}\left(\mathbf{x}_k\right)}{\int g\left(\mathbf{z}_k|\mathbf{y}_k\right)\mathcal{MBM}_{k|k-1}\left(\mathbf{y}_k\right)\delta\mathbf{y}_k}. \tag{5.15}$$

The MBM density is defined as

$$\mathcal{MBM}_{k|k}(\mathbf{x}_k) = \sum_{h_k=1}^{\mathcal{H}_k} w_{k|k}^{h_k}\mathcal{MB}_{k|k}^{h_k}\left(\mathbf{x}_k\right) \tag{5.16a}$$

$$= \sum_{h_k=1}^{\mathcal{H}_k} w_{k|k}^{h_k} \sum_{\uplus \mathbf{x}_k^i = \mathbf{x}_k} \prod_{i=1}^{N_k^{h_k}} \mathcal{B}_{k|k}^{i,h_k}\left(\mathbf{x}_k^i\right), \tag{5.16b}$$

where $\mathcal{MB}(\cdot)$ and $\mathcal{B}(\cdot)$ denote the multi-Bernoulli density (3.11) and the Bernoulli density (3.10), respectively. Each global hypothesis h_k corresponds to a sequence of data associations, $\mathcal{H}_k$ denotes the number of global hypotheses, and the weight $w_{k|k}^{h_k}$ is the probability of the global hypothesis h_k. The

Figure 5.1: The block diagram of the MBM filter.

MBM density is fully parameterized by the parameters,

$$\left\{ \left(w_{k|k}^{h_k}, \left\{ \left(r_{k|k}^{i,h_k}, p_{k|k}^{i,h_k}(\cdot) \right) \right\}_{i=1}^{N_k^{h_k}} \right) \right\}_{h_k=1}^{\mathcal{H}_k}. \tag{5.17}$$

Predicting (5.14) and updating (5.15) the MBM density then comes down to computing the predicted and updated parameters.

The MBM filter is an MOT algorithm based on the MBM conjugate prior and consists of four building blocks: 1) prediction, 2) update, 3) reduction, and 4) estimation. A block-diagram for the MBM filter is presented in Fig. 5.1. MBM reduction typically consists of 1) pruning MBs with low weights and 2) pruning Bernoullis with low probability of existence. The reader is referred to [27] for explicit equations and implementation details of the MBM filter.

The MBM_{01} filter is based on the MBM_{01} conjugate prior. It has the same filtering recursion as the MBM filter but performs MBM_{01} expansion after the prediction step. This results in an hyper-exponential increase in the number of global hypotheses in the MBM_{01} filter. The δ-generalized labelled multi-Bernoulli (δ-GLMB) filter [28] is equivalent to the labelled MBM_{01} filter where labels are used for sequential track formation.

Poisson multi-Bernoulli mixture conjugate prior

With a Poisson birth, the PMBM density $\mathcal{PMBM}_{k|k}(\mathbf{x}_k)$ is a multi-object conjugate prior to the standard multi-object models [17], [18], [21]. If the posterior at time step $k-1$ is PMBM, then for the standard multi-object transition density $\pi(\mathbf{x}_k|\mathbf{x}_{k-1})$ [14] the predicted multi-object density is also PMBM,

$$\mathcal{PMBM}_{k|k-1}(\mathbf{x}_k) = \int \pi(\mathbf{x}_k|\mathbf{x}_{k-1}) \mathcal{PMBM}_{k-1|k-1}(\mathbf{x}_{k-1}) \delta\mathbf{x}_{k-1}. \tag{5.18}$$

Further, if the predicted multi-object density at time step k is PMBM, then for the standard point and extended multi-object measurement model $g(\mathbf{z}_k|\mathbf{x}_k)$ the Bayes posterior is also PMBM

$$\mathcal{PMBM}_{k|k}(\mathbf{x}_k) = \frac{g(\mathbf{z}_k|\mathbf{x}_k)\,\mathcal{PMBM}_{k|k-1}(\mathbf{x}_k)}{\int g(\mathbf{z}_k|\mathbf{y}_k)\,\mathcal{PMBM}_{k|k-1}(\mathbf{y}_k)\,\delta\mathbf{y}_k}. \tag{5.19}$$

When using the PMBM conjugate prior for MOT, the set of objects $\mathbf{x}_k$ at time step k is the union of detected objects and undetected objects, i.e. $\mathbf{x}_k = \mathbf{x}_k^d \uplus \mathbf{x}_k^u$. The set of detected objects $\mathbf{x}_k^d$ consists of objects that the sensors have detected at least once. The set of undetected objects $\mathbf{x}_k^u$ consists of objects that have never been detected by any of the sensors (but that are within the region in which the objects are modelled).

The PMBM density is defined as

$$\mathcal{PMBM}_{k|k}(\mathbf{x}_k) = \sum_{\mathbf{x}_k^u \uplus \mathbf{x}_k^d = \mathbf{x}_k} \mathcal{P}_{k|k}^u(\mathbf{x}_k^u)\,\mathcal{MBM}_{k|k}^d\left(\mathbf{x}_k^d\right), \tag{5.20}$$

with a PPP density $\mathcal{P}_{k|k}^u(\cdot)$ for undetected objects, and an MBM density $\mathcal{MBM}_{k|k}^d(\cdot)$ for detected objects. The undetected PPP intensity is typically an unnormalized density mixture,

$$\lambda_{k|k}^{\mathrm{u}}(x_k) = \sum_{t=1}^{N_k^{\mathrm{u}}} \tilde{w}_{k|k}^{\mathrm{u},t} p_{k|k}^{\mathrm{u},t}(x_k), \tag{5.21}$$

whose parameters are

$$\left\{\left(\tilde{w}_{k|k}^{\mathrm{u},t}, p_{k|k}^{\mathrm{u},t}(\cdot)\right)\right\}_{t=1}^{N_k^{\mathrm{u}}}. \tag{5.22}$$

The MBM density, defined in (5.16), for detected objects has parameters (5.17). Therefore, the PMBM density is fully parameterized by the parameters (5.22) and (5.17). A PMBM becomes an MBM if the intensity of the PPP is zero.

The PMBM density captures relevant uncertainties in MOT elegantly. In MOT the number of objects is unknown, and in the PMBM density this is captured by both the Bernoulli probabilities of existence, and by the undetected PPP intensity. The uncertainty about the object states are captured by the Bernoulli state densities for detected objects, and by the PPP inten-

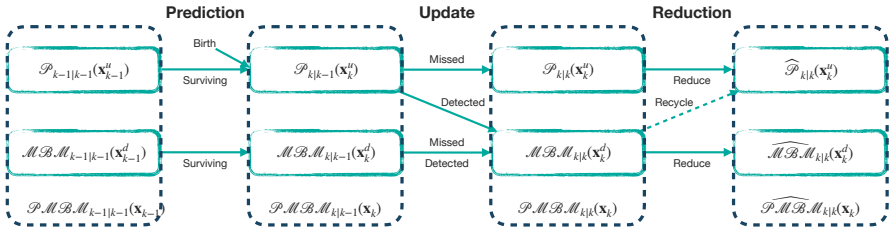

Figure 5.2: The block diagram of the PMBM filter.

sity for undetected objects. Lastly, the unknown data association in MOT is captured by the MB mixture, where each mixture component corresponds to a sequence of (global) data associations, and the component weights are the estimated probability of the corresponding global associations.

The PMBM filter is an MOT algorithm based on the PMBM conjugate prior. The PMBM filter has a hypothesis structure that is similar to MHT, see, e.g. [17], [31] for detailed discussions. Similar to the MBM filter, the PMBM filter also consists of four building blocks: 1) prediction, 2) update, 3) reduction, and 4) estimation. A block-diagram for the PMBM filter is presented in Fig. 5.2. Here recycling means that Bernoullis with low probability of existence are approximated as a PPP and added to the PPP for undetected objects [79]. Compared to Bernoulli pruning, the information contained in the Bernoullis can be approximately retained in the PPP via recycling. The PMBM filter has a more efficient representation of the hypotheses than the MBM filter, thanks to the fact that the initiation of new Bernoullis is measurement-driven. Comparisons of MBM and PMBM filters have shown that, in terms of computational cost and estimation error, PMBM filter have better performance, see, e.g. [19], [27], [37], [80]. The reader is referred to [17]–[19] for explicit equations and implementation details of the PMBM filter.

It is also useful to consider a PMB approximation to the PMBM density (5.20), which offers a trade-off between computational complexity and estimation performance. The PMB filter is a computationally efficient approximation of the PMBM filter by performing the PMB approximation after each update. Various PMB approximation methods exist for both point and extended objects [17], [81], among which the one based on variational approximation [81]

has the best performance when objects move in close proximity.

CHAPTER 6

Multi-object tracking based on sets of trajectories

In RFS-based algorithms for MOT, the main focus has been on the filtering problem. The object state estimates can be easily extracted from the multi-object filtering densities; however, it is not obvious how to build trajectories in a sound manner. A fully Bayesian approach to MOT should characterize the distribution of the trajectories given the measurements, as it contains all information about the trajectories. Performing MOT on sets of trajectories can be attained by considering multi-object density functions in which objects are trajectories [34], referred to as multi-trajectory densities. This chapter briefly reviews basic concepts of sets of trajectories and PMBM conjugate priors for sets of trajectories [36], [38], [82], [83]. In addition, a metric on the space of finite sets of trajectories [84] is introduced for MOT performance evaluation in terms of trajectory estimation error.

6.1 Sets of trajectories

This section reviews the state representation of trajectories and the integration of trajectory densities. In addition, possible problem formulations of MOT based on sets of trajectories used in this thesis are introduced.

Single trajectory

A trajectory consists of a sequence of object states that can start at any time step and end at any time after it starts. Mathematically, a trajectory is represented as a variable $X = (t, x^{1:i})$ where t is the initial time step of the trajectory, i is its length, and $x^{1:i} = (x^1, \ldots, x^i)$ denotes a sequence of length i that contains the object states at consecutive time steps of the trajectory. If k is the current time step, $t + i - 1 < k$ means that the trajectory ended at time $t + i - 1$, and $t + i - 1 = k$ means that the trajectory is ongoing.

As a trajectory $(t, x^{1:i})$ exists from time step t to $t + i - 1$, variable $(t + i)$ belongs to the set,

$$I_{(k')} = \{(t, i) : 0 \leq t \leq k' < \infty \text{ and } 1 \leq i \leq k' - t + 1 < \infty\}. \tag{6.1}$$

A single trajectory X up to a finite time step k' therefore belongs to the space

$$\mathfrak{T}_{k'} = \uplus_{(t,i)\in I_{(k')}} \{t\} \times \mathfrak{X}^i, \tag{6.2}$$

where $\mathfrak{X}^i$ represents i Cartesian products of the single object state space $\mathfrak{X}$. Given a real-valued function $p(\cdot)$ on the single trajectory space $\mathfrak{T}_{(k')}$, its integral is

$$\int p(X)dX = \sum_{(t,i)\in I_{(k')}} \int p\left(t, x^{1:i}\right) dx^{1:i}, \tag{6.3}$$

where single trajectory density $p(t, x^{1:i})$ can be factorized as

$$p(t, x^{1:i}) = p(x^{1:i}|t, i)P(t, i). \tag{6.4}$$

This integral goes through all possible start times, lengths, and object states of the trajectory.

Multiple trajectories

A set of trajectories up to a finite time step k' is denoted by $\mathbf{X} \in \mathcal{F}(\mathfrak{T}_{(k')})$ where $\mathcal{F}(\mathfrak{T}_{(k')})$ is the set of all finite subsets of $\mathfrak{T}_{(k')}$. Given a trajectory

$X = (t, x^{1:i})$, the set of the object state at time step k is

$$\tau^k(X) = \begin{cases} \{x^{k+1-t}\} & \text{if } t \leq k \leq t+i-1 \\ \varnothing & \text{otherwise.} \end{cases} \tag{6.5}$$

A trajectory X is present at time step k if and only if $|\tau^k(X)| = 1$. Given a set $\mathbf{X}$ of trajectories, the set $\tau^k(\mathbf{X})$ of object states at time step k is

$$\tau^k(\mathbf{X}) = \bigcup_{X \in \mathbf{X}} \tau^k(X). \tag{6.6}$$

The number of trajectories present at time k is given by $|\tau^k(\mathbf{X})|$.

Given a real-valued function $f(\cdot)$ on the space $\mathcal{F}(\mathfrak{T}_{(k')})$ of sets of trajectories, its set integral is

$$\int f(\mathbf{X}) \delta\mathbf{X} = \sum_{n=0}^{\infty} \frac{1}{n!} \int f\left(\{X_1, \ldots, X_n\}\right) dX_{1:n}, \tag{6.7}$$

where $X_{1:n} = (X_1, \ldots, X_n)$. If $f(\cdot)$ is a multi-trajectory density, then $f(\cdot) \geq 0$ and its set integral is one.

Multi-trajectory processes are analogous to multi-object processes for sets of object. A trajectory Poisson RFS has density

$$f(\mathbf{X}) = e^{-\int \lambda(X) dX} \prod_{X \in \mathbf{X}} \lambda(X), \tag{6.8}$$

where the trajectory Poisson intensity $\lambda(\cdot)$ is defined on the trajectory state space, i.e. realizations of the Poisson RFS are trajectories with a birth time, a length, and a state sequence. A trajectory Bernoulli RFS has density

$$f(\mathbf{X}) = \begin{cases} 1 - r & \text{if } \mathbf{X} = \varnothing \\ rp(X) & \text{if } \mathbf{X} = \{X\} \\ 0 & \text{otherwise.} \end{cases} \tag{6.9}$$

where $p(\cdot)$ is a trajectory state density and r is the Bernoulli probability of existence. A trajectory MB is the disjoint union of multiple trajectory Bernoulli RFSs; trajectory MBM RFS is an RFS whose density is a mixture

of trajectory MB densities.

Problem formulation

There are many ways in which an MOT problem can be formulated depending on the application at hand. In this thesis, the following two variants are considered:

- The set of current trajectories: the objective is to estimate the trajectories of the objects that are present in the surveillance area at the current time.

- The set of all trajectories: the objective is to estimate the trajectories of all objects that have passed through the surveillance area at some point between time step 0 and the current time step, i.e. both the objects that are present in the surveillance area at the current time, and the objects that have left the surveillance area (but were in the surveillance area at least one previous time).

Depending on the problem formulation, the variable that we are interested in is different. For the set of current trajectories, $\mathbf{X}_k$ is the set of trajectories for which $t + i - 1 = k$. For the set of all trajectories, $\mathbf{X}_k$ is the set of trajectories for which $t + i - 1 \leq k$. It is also possible to consider other problem formulations, e.g. where we are interested in trajectories in one time interval while conditioning on measurements in a different time interval [82].

With the set of trajectories $\mathbf{X}_k$ as variables of interest in MOT, the Bayesian filtering recursions are

$$p\left(\mathbf{X}_k|\mathbf{z}_{1:k-1}\right) = \int \pi\left(\mathbf{X}_k|\mathbf{X}_{k-1}\right) p\left(\mathbf{X}_{k-1}|\mathbf{z}_{1:k-1}\right) \delta\mathbf{X}_{k-1}, \tag{6.10}$$

$$p\left(\mathbf{X}_k|\mathbf{z}_{1:k}\right) = \frac{g\left(\mathbf{z}_k|\mathbf{X}_k\right) p\left(\mathbf{X}_k|\mathbf{z}_{1:k-1}\right)}{\int g\left(\mathbf{z}_k|\mathbf{Y}_k\right) p\left(\mathbf{Y}_k|\mathbf{z}_{1:k-1}\right) \delta\mathbf{Y}_k}, \tag{6.11}$$

where $\pi(\mathbf{X}_k|\mathbf{X}_{k-1})$ and $g(\mathbf{z}_k|\mathbf{X}_k)$ are multi-object dynamic and measurement models for sets of trajectories, defined analogously to their counterparts for sets of objects in Chapter 4. See [34] for details.

6.2 PMBMs for sets of trajectories

This section introduces the PMBM conjugate prior for sets of trajectories [36], [38], [82], [83]. The PMBM multi-trajectory density $\mathcal{PMBM}_{k|k}(\mathbf{X}_k)$ is a conjugate prior to the standard multi-object models with Poisson birth:

$$\mathcal{PMBM}_{k|k-1}(\mathbf{X}_k) = \int \pi(\mathbf{X}_k|\mathbf{X}_{k-1})\,\mathcal{PMBM}_{k-1|k-1}(\mathbf{X}_{k-1})\,\delta\mathbf{X}_{k-1}, \tag{6.12}$$

$$\mathcal{PMBM}_{k|k}(\mathbf{X}_k) = \frac{g(\mathbf{z}_k|\mathbf{X}_k)\,\mathcal{PMBM}_{k|k-1}(\mathbf{X}_k)}{\int g(\mathbf{z}_k|\mathbf{Y}_k)\,\mathcal{PMBM}_{k|k-1}(\mathbf{Y}_k)\,\delta\mathbf{Y}_k}. \tag{6.13}$$

In a PMBM multi-trajectory density, the trajectory Poisson RFS models trajectories of the set of undetected objects, whereas the trajectory MBM RFS models trajectories of objects that have been detected (in one of the observed measurement sets). In a trajectory MBM, each Bernoulli RFS models a single potential trajectory given a sequence of associations, an MB RFS models a set of detected trajectories for a global hypothesis, and the weights of MBs are the estimated probability of the corresponding associations.

A trajectory Poisson/Bernoulli RFS density can be marginalized to the current time step to obtain an object Poisson/Bernoulli RFS density. Therefore, marginalizing a PMBM set of trajectories density gives a PMBM set of object states density. The trajectory PMBM filter is an MOT algorithm based on the PMBM conjugate prior for sets of trajectories; see [36], [83] for explicit equations and implementation details. The relation between a trajectory PMBM filter and an object PMBM filter is illustrated in Figure 6.1. Two key insights are: 1) in an object PMBM filter the object state history is marginalized out in every prediction step, and 2) in an trajectory PMBM filter information about the past object states is maintained such that trajectories can be estimated directly from the posterior.

6.3 Metric on the space of sets of trajectories

A metric for sets of trajectories based on multidimensional assignments was proposed in [84]. This metric penalizes localization costs for properly detected objects, misdetections, false detections and track switches. Let $\Pi_{\mathbf{X},\mathbf{Y}}$ be the set of all possible assignment vectors between the index sets $\{1\ldots,|\mathbf{X}|\}$ and

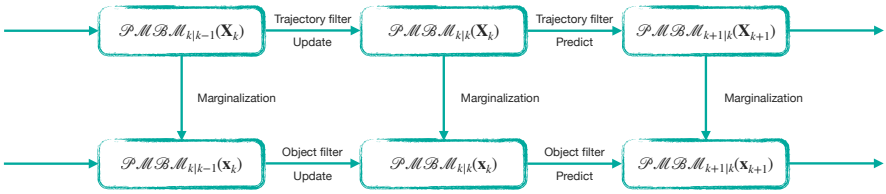

Figure 6.1: Trajectory PMBM filter versus object PMBM filter.

$\{0, \ldots, |\mathbf{Y}|\}$. An assignment vector $\pi^k = [\pi_1^k, \ldots, \pi_{|\mathbf{X}|}^k]^{\mathrm{T}}$ at time step k is a vector $\pi^k \in \{0, \ldots, |\mathbf{Y}|\}^{n_{\mathbf{X}}}$ such that its ith component $\pi_i^k = \pi_{i'}^k = j > 0$ implies that $i = i'$. Here $\pi_i^k = j \neq 0$ implies that trajectory i in $\mathbf{X}$ is assigned to trajectory j in $\mathbf{Y}$ at time step k and $\pi_i^k = 0$ implies that trajectory i in $\mathbf{X}$ is unassigned at time step k.

For $1 \leq p < \infty$, cut-off parameter $c > 0$, switching penalty $\gamma > 0$ and a base metric $d_b(\cdot, \cdot)$ in the single object space $\mathfrak{X}$, the multidimensional assignment metric $d_p^{(c,\gamma)}(\mathbf{X}, \mathbf{Y})$ between two sets $\mathbf{X}$ and $\mathbf{Y}$ of trajectories in time interval $1, \ldots, T$ is

$$d_p^{(c,\gamma)}(\mathbf{X}, \mathbf{Y}) = \min_{\pi^k \in \Pi_{\mathbf{X},\mathbf{Y}}} \left(\sum_{k=1}^{T} d_{\mathbf{X},\mathbf{Y}}^k \left(\mathbf{X}, \mathbf{Y}, \pi^k\right)^p + \sum_{k=1}^{T-1} s_{\mathbf{X},\mathbf{Y}} \left(\pi^k, \pi^{k+1}\right)^p \right)^{\frac{1}{p}}, \tag{6.14}$$

where the costs (to the p-th power) for properly detected objects, misdetections and false detections at time step k are

$$\begin{aligned} d_{\mathbf{X},\mathbf{Y}}^k \left(\mathbf{X}, \mathbf{Y}, \pi^k\right)^p = & \sum_{(i,j) \in \theta^k(\pi^k)} d\left(\tau^k(X_i), \tau^k(Y_j)\right)^p \\ & + \frac{c^p}{2} \left(\left|\tau^k(\mathbf{X})\right| + \left|\tau^k(\mathbf{Y})\right| - 2\left|\theta^k\left(\pi^k\right)\right| \right), \end{aligned} \tag{6.15}$$

with

$$\begin{aligned} \theta^k\left(\pi^k\right) = \Big\{ & \left(i, \pi_i^k\right) : i \in \{1, \ldots, n_{\mathbf{X}}\}, \\ & |\tau^k(X_i)| = \left|\tau^k(Y_{\pi_i^k})\right| = 1, d\left(\tau^k(X_i), \tau^k(Y_{\pi_i^k})\right) < c \Big\}, \end{aligned} \tag{6.16}$$

and the switching cost (to the p-th power) from time step k to $k+1$ is given by

$$s_{\mathbf{X},\mathbf{Y}}(\pi^k, \pi^{k+1})^p = \gamma^p \sum_{i=1}^{|\mathbf{X}|} s\left(\pi_i^k, \pi_i^{k+1}\right), \tag{6.17}$$

$$s\left(\pi_i^k, \pi_i^{k+1}\right) = \begin{cases} 0 & \text{if } \pi_i^k = \pi_i^{k+1} \\ 1 & \text{if } \pi_i^k \neq \pi_i^{k+1}, \pi_i^k \neq 0, \pi_i^{k+1} \neq 0 \\ \frac{1}{2} & \text{otherwise.} \end{cases} \tag{6.18}$$

It should be noted that, for $(i,j) \in \theta^k$, $\tau^k(X_i)$ and $\tau^k(Y_{\pi_i^k})$ contain precisely one element and their distance is smaller than c, so $d\left(\tau^k(X_i), \tau^k(Y_j)\right)$ coincides with $d_b(\cdot,\cdot)$ evaluated at the corresponding single object states, which corresponds to the localization error. Therefore, (6.15) represents the sum of the costs (to the pth power) that correspond to localization error for properly detected objects (indicated by the assignments in $\theta^k(\pi^k)$), number of misdetections ($|\tau^k(\mathbf{X})| - |\theta^k(\pi^k)|$) and false detections ($|\tau^k(\mathbf{Y})| - |\theta^k(\pi^k)|$) at time step k.

The metric (6.14) can be computed by solving a multidimensional assignment problem, which may be computationally heavy for large T. It was further shown in [84] that an accurate lower bound on the metric (6.14) can be obtained using linear programming, which can be computed in polynomial time. Note that this lower bound is also a metric. In addition, the metric (6.14) can be extended by including weights to the costs associated to different time steps [85].

CHAPTER 7

Summary of included papers

This chapter provides a summary of the included papers.

7.1 Paper A

Yuxuan Xia, Karl Granström, Lennart Svensson,
Ángel F. García-Fernández, Jason L. Williams
Multi-scan implementation of the trajectory Poisson multi-Bernoulli mixture filter
Published in Journal of Advances in Information Fusion,
vol. 14, no. 2, pp. 213–235, December 2019,
 ISSN: 1557-6418 .

The Poisson multi-Bernoulli mixture (PMBM) and the multi-Bernoulli mixture (MBM) are two multi-target distributions for which closed-form filtering recursions exist. The PMBM has a Poisson birth process, whereas the MBM has a multi-Bernoulli birth process. This paper considers a recently developed formulation of the multi-target tracking problem using a random finite set of trajectories, through which the track continuity is explicitly established. A

multi-scan trajectory PMBM filter and a multi-scan trajectory MBM filter, with the ability to correct past data association decisions to improve current decisions, are presented. In addition, a multi-scan trajectory MBM_{01} filter, in which the existence probabilities of all Bernoulli components are either 0 or 1, is given. This paper proposes an efficient implementation that performs track-oriented N-scan pruning to limit the computational complexity, and uses dual decomposition to solve the involved multi-frame assignment problem. The performance of the presented multi-target trackers, applied with an efficient fixed-lag smoothing method, is evaluated in a simulation study.

7.2 Paper B

Yuxuan Xia, Lennart Svensson, Ángel F. García-Fernández, Jason L. Williams, Daniel Svensson, and Karl Granström
Multiple object trajectory estimation using backward simulation
Published in IEEE Transactions on Signal Processing, vol. 70, pp. 3249–3263, Jun. 2022.
 ISSN: 1941-0476 .

This paper presents a general solution for computing the multi-object posterior for sets of trajectories from a sequence of multi-object (unlabelled) filtering densities and a multi-object dynamic model. Importantly, the proposed solution opens an avenue of trajectory estimation possibilities for multi-object filters that do not explicitly estimate trajectories. In this paper, we first derive a general multi-trajectory backward smoothing equation based on random finite sets of trajectories. Then we show how to sample sets of trajectories using backward simulation for Poisson multi-Bernoulli filtering densities and develop a tractable implementation based on ranked assignment. The performance of the resulting multi-trajectory particle smoothers is evaluated in a simulation study, and the results demonstrate that they have superior performance in comparison to several state-of-the-art multi-object filters and smoothers.

7.3 Paper C

Yuxuan Xia, Karl Granström, Lennart Svensson,
Maryam Fatemi, Ángel F. García-Fernández, Jason L. Williams

Poisson multi-Bernoulli approximations for multiple extended object filtering
Published in IEEE Transactions on Aerospace and Electronic Systems, vol. 58, no. 2, pp. 890–906, Apr. 2022.
 ISSN: 1557-9603 .

The Poisson multi-Bernoulli mixture (PMBM) is a multi-object conjugate prior for the closed-form Bayes random finite set filter. The extended object PMBM filter provides a closed-form solution for multiple extended object filtering with standard models. This article considers computationally lighter alternatives to the extended object PMBM filter by propagating a Poisson multi-Bernoulli (PMB) density through the filtering recursion. A new local hypothesis representation is presented, where each measurement creates a new Bernoulli component. This facilitates the development of methods for efficiently approximating the PMBM posterior density after the update step as a PMB. Based on the new hypothesis representation, two approximation methods are presented: one is based on the track-oriented multi-Bernoulli (MB) approximation, and the other is based on the variational MB approximation via Kullback-Leibler divergence minimization. The performance of the proposed PMB filters with gamma Gaussian inverse-Wishart implementations are evaluated in a simulation study.

7.4 Paper D

Yuxuan Xia, Ángel F. García-Fernández, Florian Meyer, Jason L. Williams, Karl Granström, and Lennart Svensson
Trajectory PMB filters for extended object tracking using belief propagation
Pre-print is available at https://arxiv.org/abs/2207.10164. .

In this paper, we propose a Poisson multi-Bernoulli (PMB) filter for extended object tracking (EOT), which directly estimates the set of object trajectories, using belief propagation (BP). The proposed filter propagates a PMB density on the posterior of sets of trajectories through the filtering recursions over time, where the PMB mixture (PMBM) posterior after the update step is approximated as a PMB. The efficient PMB approximation relies on several important theoretical contributions. First, we present a PMBM conjugate

prior on the posterior of the sets of trajectories for a generalized measurement model, in which each object generates an independent set of measurements. The PMBM density is a conjugate prior in the sense that both the prediction and the update steps preserve the PMBM form of the density. Second, we present a factor graph representation of the joint posterior of the PMBM set of trajectories and association variables for the Poisson spatial measurement model. Importantly, leveraging the PMBM conjugacy and the factor graph formulation enables an elegant treatment of undetected objects via a Poisson point process and efficient inference on sets of trajectories using BP, where the approximate marginal densities in the PMB approximation can be obtained without enumeration of different data association hypotheses. To achieve this, we present a particle-based implementation of the proposed filter, where smoothed trajectory estimates, if desired, can be obtained via single-object particle smoothing methods, and its performance for EOT with ellipsoidal shapes is evaluated in a simulation study.

CHAPTER 8

Concluding Remarks and Future Work

This thesis studies Bayesian object tracking problems for both point and extended objects with a focus on MOT based on sets of trajectories. The concluding remarks and possible future work directions of the included papers are given as follows:

- **Paper A: "Multi-scan implementations of the trajectory Poisson multi-Bernoulli mixture filter"**
  This paper shows that multi-scan data association algorithms used in classical track-oriented MHT can be utilized in trajectory filters based on multi-object conjugate priors, resulting in the multi-scan implementations of the trajectory filters. Interesting future work is to benchmark the multi-scan trajectory filters against efficient track-oriented MHT algorithms in the literature. It would also be interesting to continue the development of more efficient optimization algorithms for solving the multi-scan data association problem.

- **Paper B: "Multiple object trajectory estimation using backward simulation"**
  This paper presents a general forward-backward smoothing equation for

sets of trajectories and proposed a multi-trajectory particle smoother using backward simulation for PMB filtering densities along with a tractable implementation based on ranked assignment. Interesting future work is to consider a one-time-step lagged implementation of the multi-trajectory smoother, such that trajectories can be built upon the computation of multi-object filtering densities. It is also valuable to develop an implementation that works for forward densities with particle representation.

- **Paper C: "Poisson multi-Bernoulli approximations for multiple extended object filtering"**
  This paper presents an extended object PMB filter, which propagates a PMB density through the filtering recursion, as a computationally lighter alternative to the extended object PMBM filter. Two different PMB approximation methods are presented: one is based on the track-oriented MB approximation, and the other is based on variational MB approximation via KLD minimization. It would be valuable to develop different extended object PMB implementations with different object spatial models and verify their performance on real data.

- **Paper D: "Trajectory PMB filters for extended object tracking using belief propagation "**
  This paper presents a PMBM conjugate prior on the posterior of sets of trajectories for a generalized measurement model, and a factor graph representation of the joint posterior of the PMBM set of trajectories and association variables for the Poisson spatial measurement model. Based on these theoretical contributions, two TPMB filters for multiple EOT implementation using particle BP are presented: one estimates the set of alive trajectories, and the other estimates the set of all trajectories. For future work, it would be interesting to study how to make use of variational MB approximation to facilitate track initialization. It is also worth investigating how to extend BP for EOT to consider multi-scan data associations and tracking co-existing point and extended objects.

In recent years, deep learning has been increasingly used in MOT for improving tracking performance using camera [86], lidar [87] and radar [88]. In addition, recent works [89], [90] show that a deep learning based MOT method can match or outperform the performance of the state-of-the-art model-based

Bayesian MOT methods in the model-based setting for point object tracking, in terms of GOSPA and an uncertainty aware performance measure based on negative log-likelihood [91]. Furthermore, a hybrid method for model-based and deep learning based MOT has been recently proposed in [92] using neural enhanced BP. Therefore, it would be interesting to explore how to improve the performance of model-based MOT methods using deep learning techniques as well as how to leverage techniques in model-based MOT methods in designing deep neural network architectures for MOT.

References

[1] Y. Bar-Shalom, P. K. Willett, and X. Tian, *Tracking and data fusion*. YBS publishing Storrs, 2011, vol. 11.

[2] S. Challa, M. R. Morelande, D. Mušicki, and R. J. Evans, *Fundamentals of object tracking*. Cambridge University Press, 2011.

[3] S. Särkkä, *Bayesian filtering and smoothing*. Cambridge University Press, 2013, vol. 3.

[4] K. Granström, M. Baum, and S. Reuter, "Extended object tracking: Introduction, overview, and applications," *Journal of Advances in Information Fusion*, vol. 12, no. 2, 2017.

[5] K. Granström, L. Svensson, S. Reuter, Y. Xia, and M. Fatemi, "Likelihood-based data association for extended object tracking using sampling methods," *IEEE Transactions on Intelligent Vehicles*, vol. 3, no. 1, pp. 30–45, 2017.

[6] Y. Xia, P. Wang, K. Berntorp, *et al.*, "Extended object tracking using hierarchical truncation measurement model with automotive radar," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2020, pp. 4900–4904.

[7] Y. Xia, P. Wang, K. Berntorp, H. Mansour, P. Boufounos, and P. V. Orlik, "Extended object tracking using hierarchical truncation model with partial-view measurements," in *11th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, IEEE, 2020, pp. 1–5.

[8] Y. Xia, P. Wang, K. Berntorp, *et al.*, "Extended object tracking with automotive radar using learned structural measurement model," in *Radar Conference (RadarConf)*, IEEE, 2020, pp. 1–6.

[9] Y. Xia, P. Wang, K. Berntorp, *et al.*, "Learning-based extended object tracking using hierarchical truncation measurement model with automotive radar," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 4, pp. 1013–1029, 2021.

[10] L. D. Stone, R. L. Streit, T. L. Corwin, and K. L. Bell, *Bayesian multiple target tracking*. Artech House, 2013.

[11] Y. Bar-Shalom, F. Daum, and J. Huang, "The probabilistic data association filter," *IEEE Control Systems Magazine*, vol. 29, no. 6, pp. 82–100, 2009.

[12] S. S. Blackman, "Multiple hypothesis tracking for multiple target tracking," *IEEE Aerospace and Electronic Systems Magazine*, vol. 19, no. 1, pp. 5–18, 2004.

[13] I. R. Goodman, R. P. Mahler, and H. T. Nguyen, *Mathematics of data fusion*. Springer Science & Business Media, 2013, vol. 37.

[14] R. P. Mahler, *Statistical Multisource-Multitarget Information Fusion*. Artech House, 2007.

[15] ——, *Advances in statistical multisource-multitarget information fusion*. Artech House, 2014.

[16] B.-T. Vo and B.-N. Vo, "Labeled random finite sets and multi-object conjugate priors," *IEEE Transactions on Signal Processing*, vol. 61, no. 13, pp. 3460–3475, 2013.

[17] J. L. Williams, "Marginal multi-Bernoulli filters: RFS derivation of MHT, JIPDA, and association-based MeMBer," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 51, no. 3, pp. 1664–1687, 2015.

[18] Á. F. García-Fernández, J. L. Williams, K. Granström, and L. Svensson, "Poisson multi-Bernoulli mixture filter: Direct derivation and implementation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 4, pp. 1883–1901, 2018.

[19] K. Granström, M. Fatemi, and L. Svensson, "Poisson multi-Bernoulli mixture conjugate prior for multiple extended target filtering," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 1, pp. 208–225, 2020.

[20] Y. Xia, K. Granström, L. Svensson, M. Fatemi, Á. F. García-Fernández, and J. L. Williams, "Poisson multi-Bernoulli approximations for multiple extended object filtering," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 58, no. 2, pp. 890–906, 2022.

[21] Á. F. García-Fernández, J. L. Williams, L. Svensson, and Y. Xia, "A Poisson multi-Bernoulli mixture filter for coexisting point and extended targets," *IEEE Transactions on Signal Processing*, vol. 69, pp. 2600–2610, 2021.

[22] M. Fatemi, K. Granström, L. Svensson, F. J. Ruiz, and L. Hammarstrand, "Poisson multi-Bernoulli mapping using Gibbs sampling," *IEEE Transactions on Signal Processing*, vol. 65, no. 11, pp. 2814–2827, 2017.

[23] M. Fröhle, C. Lindberg, K. Granström, and H. Wymeersch, "Multisensor Poisson multi-Bernoulli filter for joint target–sensor state tracking," *IEEE Transactions on Intelligent Vehicles*, vol. 4, no. 4, pp. 609–621, 2019.

[24] M. Fröhle, K. Granström, and H. Wymeersch, "Decentralized Poisson multi-Bernoulli filtering for vehicle tracking," *IEEE Access*, vol. 8, pp. 126 414–126 427, 2020.

[25] P. Boström-Rost, D. Axehill, and G. Hendeby, "Sensor management for search and track using the Poisson multi-Bernoulli mixture filter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 5, pp. 2771–2783, 2021.

[26] Y. Ge, O. Kaltiokallio, H. Kim, *et al.*, "A computationally efficient EK-PMBM filter for bistatic mmwave radio SLAM," *IEEE Journal on Selected Areas in Communications*, 2022.

[27] Á. F. García-Fernández, Y. Xia, K. Granström, L. Svensson, and J. L. Williams, "Gaussian implementation of the multi-Bernoulli mixture filter," in *22th International Conference on Information Fusion (FUSION)*, IEEE, 2019, pp. 1–8.

[28] 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*, vol. 62, no. 24, pp. 6554–6567, 2014.

[29] R. P. Mahler, "Multitarget Bayes filtering via first-order multitarget moments," *IEEE Transactions on Aerospace and Electronic systems*, vol. 39, no. 4, pp. 1152–1178, 2003.

[30] R. Mahler, "PHD filters of higher order in target number," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 4, pp. 1523–1543, 2007.

[31] E. Brekke and M. Chitre, "Relationship between finite set statistics and the multiple hypothesis tracker," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 4, pp. 1902–1917, 2018.

[32] Á. F. García-Fernández, J. Grajal, and M. R. Morelande, "Two-layer particle filter for multiple target detection and tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 49, no. 3, pp. 1569–1588, 2013.

[33] E. H. Aoki, P. K. Mandal, L. Svensson, Y. Boers, and A. Bagchi, "Labeling uncertainty in multitarget tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 3, pp. 1006–1020, 2016.

[34] Á. F. García-Fernández, L. Svensson, and M. R. Morelande, "Multiple target tracking based on sets of trajectories," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 3, pp. 1685–1707, 2020.

[35] Á. F. García-Fernández and L. Svensson, "Trajectory PHD and CPHD filters," *IEEE Transactions on Signal Processing*, vol. 67, no. 22, pp. 5702–5714, 2019.

[36] K. Granström, L. Svensson, Y. Xia, J. L. Williams, and Á. F. García-Fernández, "Poisson multi-Bernoulli mixture trackers: Continuity through random finite sets of trajectories," in *21st International Conference on Information Fusion (FUSION)*, IEEE, 2018, pp. 1–8.

[37] Y. Xia, K. Granström, L. Svensson, and Á. F. García-Fernández, "An implementation of the Poisson multi-Bernoulli mixture trajectory filter via dual decomposition," in *21st International Conference on Information Fusion (FUSION)*, 2018, pp. 1–8.

[38] Y. Xia, K. Granström, L. Svensson, Á. F. García-Fernández, and J. L. Williams, "Extended target Poisson multi-Bernoulli mixture trackers based on sets of trajectories," in *22th International Conference on Information Fusion (FUSION)*, IEEE, 2019, pp. 1–8.

[39] Á. F. García-Fernández and L. Svensson, "Tracking multiple spawning targets using poisson multi-bernoulli mixtures on sets of tree trajectories," *IEEE Transactions on Signal Processing*, 2022.

[40] Á. F. García-Fernández, L. Svensson, J. L. Williams, Y. Xia, and K. Granström, "Trajectory Poisson multi-Bernoulli filters," *IEEE Transactions on Signal Processing*, vol. 68, pp. 4933–4945, 2020.

[41] Y. Xia, K. Granström, L. Svensson, Á. F. García-Fernández, and J. L. Williams, "Multi-scan implementation of the trajectory Poisson multi-Bernoulli mixture filter," *Journal of Advances in Information Fusion*, vol. 14, no. 2, pp. 213–235, 2019.

[42] Á. F. García-Fernández, L. Svensson, J. L. Williams, Y. Xia, and K. Granström, "Trajectory multi-Bernoulli filters for multi-target tracking based on sets of trajectories," in *23rd International Conference on Information Fusion (FUSION)*, IEEE, 2020, pp. 1–8.

[43] W. Koch and F. Govaers, "On accumulated state densities with applications to out-of-sequence measurement processing," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 4, pp. 2766–2778, 2011.

[44] C. M. Bishop and N. M. Nasrabadi, *Pattern recognition and machine learning*, 4. Springer, 2006, vol. 4.

[45] F. Lindsten and T. B. Schön, "Backward simulation methods for monte carlo statistical inference," *Foundations and Trends® in Machine Learning*, vol. 6, no. 1, pp. 1–143, 2013.

[46] R. M. Eustice, H. Singh, and J. J. Leonard, "Exactly sparse delayed-state filters for view-based slam," *IEEE Transactions on Robotics*, vol. 22, no. 6, pp. 1100–1114, 2006.

[47] Y. Xia, L. Svensson, Á. F. García-Fernández, K. Granström, and J. L. Williams, "Backward simulation for sets of trajectories," in *2020 IEEE 23rd International Conference on Information Fusion (FUSION)*, IEEE, 2020, pp. 1–8.

[48] Y. Xia, L. Svensson, Á. F. García-Fernández, J. L. Williams, D. Svensson, and K. Granström, "Multiple object trajectory estimation using backward simulation," *IEEE Transactions on Signal Processing*, vol. 70, pp. 3249–3263, 2022.

[49] Á. F. García-Fernández, L. Svensson, M. R. Morelande, and S. Särkkä, "Posterior linearization filter: Principles and implementation using sigma points," *IEEE Transactions on Signal Processing*, vol. 63, no. 20, pp. 5561–5573, 2015.

[50] E. A. Wan and R. Van Der Merwe, "The unscented Kalman filter for nonlinear estimation," in *Adaptive Systems for Signal Processing, Communications, and Control Symposium*, IEEE, 2000, pp. 153–158.

[51] I. Arasaratnam and S. Haykin, "Cubature Kalman filters," *IEEE Transactions on Automatic Control*, vol. 54, no. 6, pp. 1254–1269, 2009.

[52] H. E. Rauch, F. Tung, and C. T. Striebel, "Maximum likelihood estimates of linear dynamic systems," *AIAA journal*, vol. 3, no. 8, pp. 1445–1450, 1965.

[53] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 498–519, 2001.

[54] F. Meyer, P. Braca, P. Willett, and F. Hlawatsch, "A scalable algorithm for tracking an unknown number of targets using multiple sensors," *IEEE Transactions on Signal Processing*, vol. 65, no. 13, pp. 3478–3493, 2017.

[55] F. Meyer, T. Kropfreiter, J. L. Williams, *et al.*, "Message passing algorithms for scalable multitarget tracking," *Proceedings of the IEEE*, vol. 106, no. 2, pp. 221–259, 2018.

[56] F. Meyer, Z. Liu, and M. Z. Win, "Scalable probabilistic data association with extended objects," in *International Conference on Communications Workshops (ICC Workshops)*, IEEE, 2019, pp. 1–6.

[57] F. Meyer and J. L. Williams, "Scalable detection and tracking of geometric extended objects," *IEEE Transactions on Signal Processing*, vol. 69, pp. 6283–6298, 2021.

[58] Y. Xia, Á. F. García-Fernández, M. Florian, J. Williams, K. Granström, and L. Svensson, "Trajectory PMB filters for extended object tracking using belief propagation," *arXiv preprint arXiv:2207.10164*, 2022.

[59] F. Meyer, O. Hlinka, H. Wymeersch, E. Riegler, and F. Hlawatsch, "Distributed localization and tracking of mobile networks including noncooperative objects," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 2, no. 1, pp. 57–71, 2015.

[60] M. Z. Win, F. Meyer, Z. Liu, W. Dai, S. Bartoletti, and A. Conti, "Efficient multisensor localization for the internet of things: Exploring a new class of scalable localization algorithms," *IEEE Signal Processing Magazine*, vol. 35, no. 5, pp. 153–167, 2018.

[61] E. Leitinger, F. Meyer, F. Hlawatsch, K. Witrisal, F. Tufvesson, and M. Z. Win, "A belief propagation algorithm for multipath-based SLAM," *IEEE Transactions on Wireless Communications*, vol. 18, no. 12, pp. 5613–5629, 2019.

[62] J. Williams and R. Lau, "Approximate evaluation of marginal association probabilities with belief propagation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, no. 4, pp. 2942–2959, 2014.

[63] S. Yang, M. Baum, and K. Granström, "Metrics for performance evaluation of elliptic extended object tracking methods," in *International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, IEEE, 2016, pp. 523–528.

[64] C. R. Givens, R. M. Shortt, *et al.*, "A class of Wasserstein metrics for probability distributions.," *The Michigan Mathematical Journal*, vol. 31, no. 2, pp. 231–240, 1984.

[65] D. Schuhmacher and A. Xia, "A new metric between distributions of point processes," *Advances in Applied Probability*, vol. 40, no. 3, pp. 651–672, 2008.

[66] D. Schuhmacher, B.-T. Vo, and B.-N. Vo, "A consistent metric for performance evaluation of multi-object filters," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 3447–3457, 2008.

[67] A. S. Rahmathullah, Á. F. García-Fernández, and L. Svensson, "Generalized optimal sub-pattern assignment metric," in *20th International Conference on Information Fusion (Fusion)*, IEEE, 2017, pp. 1–8.

[68] Á. F. García-Femández and L. Svensson, "Spooky effect in optimal OSPA estimation and how GOSPA solves it," in *22th International Conference on Information Fusion (FUSION)*, IEEE, 2019, pp. 1–8.

[69] Á. F. García-Fernández, M. Hernandez, and S. Maskell, "An analysis on metric-driven multi-target sensor management: GOSPA versus OSPA," in *24th International Conference on Information Fusion (FUSION)*, IEEE, 2021, pp. 1–8.

[70] X. R. Li and V. P. Jilkov, "Survey of maneuvering target tracking. Part I. Dynamic models," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 39, no. 4, pp. 1333–1364, 2003.

[71] J. Liu, W. Xiong, L. Bai, *et al.*, "Deep instance segmentation with automotive radar detection points," *IEEE Transactions on Intelligent Vehicles*, 2022.

[72] K. Gilholm and D. Salmond, "Spatial distribution model for tracking extended objects," *IEE Proceedings-Radar, Sonar and Navigation*, vol. 152, no. 5, pp. 364–371, 2005.

[73] K. Granström and L. Svensson, *Multi-object tracking for automotive systems*, `https://www.edx.org/course/multi-object-tracking-for-automotive-systems`.

[74] M. Feldmann, D. Franken, and W. Koch, "Tracking of extended objects and group targets using random matrices," *IEEE Transactions on Signal Processing*, vol. 59, no. 4, pp. 1409–1420, 2010.

[75] U. Orguner, "A variational measurement update for extended target tracking with random matrices," *IEEE Transactions on Signal Processing*, vol. 60, no. 7, pp. 3827–3834, 2012.

[76] K. Granström and U. Orguner, "New prediction for extended targets with random matrices," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, no. 2, pp. 1577–1589, 2014.

[77] N. J. Bartlett, C. Renton, and A. G. Wills, "A closed-form prediction update for extended target tracking using random matrices," *IEEE Transactions on Signal Processing*, vol. 68, pp. 2404–2418, 2020.

[78] M. Beard, S. Reuter, K. Granström, B.-T. Vo, B.-N. Vo, and A. Scheel, "Multiple extended target tracking with labeled random finite sets," *IEEE Transactions on Signal Processing*, vol. 64, no. 7, pp. 1638–1653, 2015.

[79] J. L. Williams, "Hybrid Poisson and multi-Bernoulli filters," in *15th International Conference on Information Fusion*, IEEE, 2012, pp. 1103–1110.

[80] Y. Xia, K. Granström, L. Svensson, and Á. F. García-Fernández, "Performance evaluation of multi-Bernoulli conjugate priors for multi-target filtering," in *20th International Conference on Information Fusion (Fusion)*, IEEE, 2017, pp. 1–8.

[81] J. L. Williams, "An efficient, variational approximation of the best fitting multi-Bernoulli filter," *IEEE Transactions on Signal Processing*, vol. 63, no. 1, pp. 258–273, 2014.

[82] K. Granström, L. Svensson, Y. Xia, Á. F. García-Fernández, and J. L. Williams, "Spatiotemporal constraints for sets of trajectories with applications to PMBM densities," in *23rd International Conference on Information Fusion (FUSION)*, IEEE, 2020, pp. 1–8.

[83] K. Granström, L. Svensson, Y. Xia, J. Williams, and Á. F. García-Fernández, "Poisson multi-Bernoulli mixtures for sets of trajectories," *arXiv preprint arXiv:1912.08718*, 2019.

[84] Á. F. García-Fernández, A. S. Rahmathullah, and L. Svensson, "A metric on the space of finite sets of trajectories for evaluation of multi-target tracking algorithms," *IEEE Transactions on Signal Processing*, vol. 68, pp. 3917–3928, 2020.

[85] ——, "A time-weighted metric for sets of trajectories to assess multi-object tracking algorithms," in *24th International Conference on Information Fusion (FUSION)*, IEEE, 2021, pp. 1–8.

[86] J. Cai, M. Xu, W. Li, *et al.*, "MeMOT: Multi-object tracking with memory," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022.

[87] X. Bai, Z. Hu, X. Zhu, *et al.*, "TransFusion: Robust LiDAR-camera fusion for 3D object detection with transformers," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022.

[88] P. Li, P. Wang, K. Berntorp, and H. Liu, "Exploiting temporal relations on radar perception for autonomous driving," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022.

[89] J. Pinto, G. Hess, W. Ljungbergh, Y. Xia, L. Svensson, and H. Wymeersch, "Next generation multitarget trackers: Random finite set methods vs transformer-based deep learning," in *24th International Conference on Information Fusion (FUSION)*, IEEE, 2021, pp. 1–8.

[90] J. Pinto, G. Hess, W. Ljungbergh, Y. Xia, H. Wymeersch, and L. Svensson, "Can deep learning be applied to model-based multi-object tracking?" *arXiv preprint arXiv:2202.07909*, 2022.

[91] J. Pinto, Y. Xia, L. Svensson, and H. Wymeersch, "An uncertainty-aware performance measure for multi-object tracking," *IEEE Signal Processing Letters*, vol. 28, pp. 1689–1693, 2021.

[92] M. Liang and F. Meyer, "Neural enhanced belief propagation for data assocation in multiobject tracking," in *25th International Conference on Information Fusion (FUSION)*, IEEE, 2022, pp. 1–8.