Group and Extended Target Tracking with the Probability Hypothesis Density Filter

Anthony Jack Swain

Submitted for the degree of Doctor of Philosophy

to Heriot Watt University

on completion of research in the

Department of Electrical, Electronic and Computing Engineering.

November 2013

Abstract

Multiple target tracking concerns the estimation of an unknown and time-varying number of objects (targets) as they dynamically evolve over time from a sequence of measurements obtained from sensors at discrete time intervals. In the Bayesian filtering framework the estimation problem incorporates natural phenomena such as false measurements and target birth/death. Though theoretically optimal, the generally intractable Bayesian filter requires suitable approximations. This thesis is particularly motivated by a first-order moment approximation known as the Probability Hypothesis Density (PHD) filter.

The emphasis in this thesis is on the further development of the PHD filter for handling more advanced target tracking problems, principally involving multiple group and extended targets. A group target is regarded as a collection of targets that share a common motion or characteristic, while an extended target is regarded as a target that potentially generates multiple measurements.

The main contributions are the derivations of the PHD filter for multiple group and extended target tracking problems and their subsequent closed-form solutions. The proposed algorithms are applied in simulated scenarios and their estimate results demonstrate that accurate tracking performance is attainable for certain group/extended target tracking problems. The performance is further analysed with the use of suitable metrics.

Dedication

In memory of Patricia Swain.

Acknowledgements

To start with, I wish to express my gratitude to Dr Daniel Clark for his continual support and guidance as primary supervisor throughout the course of my studies. His knowledge and enthusiasm were a constant source of inspiration. I would also like to thank Neil Cade for the additional perspective that he provided with his wealth of experience in industry. By extension, the School of Engineering and Physical Sciences deserves grateful acknowledgement as the host institution for my studies at Heriot Watt University as well as Selex ES (formerly known as Selex Galileo Ltd) for their financial contribution in the EPSRC Industrial CASE Award Studentship and ultimately for the job opportunity which followed.

In addition, I would like to thank Murat Uney for the constructive feedback he provided on the first couple of chapters I wrote for the thesis and Edmund Brekke for his helpful suggestion concerning the derivations in the appendix.

Finally, but by no means least, I have my partner Kirsty to thank for helping me to see it through to the end. Without your support at home, especially through a particularly difficult period during the write-up stage, who knows whether I would have had the strength and belief to finish. Thank you.

Contents

List of Tables

List of Figures

List of Acronyms

RFS	random finite set
PHD	Probability Hypothesis Density
CPHD	Cardinalized Probability Hypothesis Density
i.i.d.	independent identically distributed
SMC	sequential Monte Carlo
MeMBer	Multi-target Multi-Bernoulli
GM	Gaussian mixture
RMSE	root mean squared error
OSPA	optimal subpattern assignment
JoTT	joint target-detection and tracking
p.g.fl.	probability generating functional
MHT	multiple hypothesis tracking
PDA	probabilistic data association
JPDA	joint probabilistic data association

Chapter 1

Introduction

1.1 Motivation

The research area motivating the work documented in this thesis is that of target tracking, which concerns the estimation of a dynamic target under surveillance based on a noise corrupted sensor measurement. In a Bayesian filtering framework the state representing the target and the measurement are regarded as random variables, so that the target state is estimated recursively via its probability density. The extension to multiple target tracking is of particular interest, whereby the state now represents a multi-target system and is regarded as a set-valued random variable in the optimal Bayesian filtering framework. Specifically, multiple target tracking concerns the joint estimation of an unknown and varying number of dynamic targets and their states. The multi-target estimation problem requires the processing of multiple measurements, some of which may not be target generated and referred to as false alarms or clutter. The origin of each measurement is subsequently unknown and some targets may not even generate a measurement, which gives rise to detection uncertainty.

Sub-optimal filtering approaches to multiple target tracking are more favourable in practice. Section 1.1.1 introduces one such approach, significant with regards to this thesis, namely the Probability Hypothesis Density (PHD) filter, and reviews some of its recent developments. To conclude the motivational section of this chapter, two types of advanced multiple target tracking problems are introduced, group target tracking in Section 1.1.2 and extended target tracking in Section 1.1.3.

1.1.1 Multiple target tracking with the PHD filter

Proposed by Mahler in [1], the PHD filter is an approximation to the multi-target Bayesian filter that recursively propagates the first-order moment density, or intensity, instead of the full probability density. The recursive equations in the PHD filter account for natural phenomena such as target birth and death, as well as false

alarms and missed detections. The expected number of targets are uniquely determined from the propagated intensities in the recursion such that the PHD filter is also first-order in target number. This sub-optimal filter has been successfully demonstrated in a number of applications, such as sonar [2, 3, 4], acoustics [5], radar [6, 7, 8, 9, 10] and computer vision [11, 12].

The PHD filter has been shown to handle maneuvering targets as demonstrated in [13, 14, 15, 16, 17, 18]. Specifically, in [13, 14], Punithakumar, Kirubarajan & Sinha resolve dynamic model uncertainty by assuming multiple models for possible motion modes and then combining the mode-dependent estimates in a similar manner to that used in the interacting multiple model estimator [19]. Alternatively, jump Markov system models have been utilised for tracking maneuvering targets using the PHD filter as proposed by Pasha, Vo, et. al. [16, 17, 18].

The PHD filter is association free, i.e. the state estimates of the individual targets given by the PHD filter have no labels associated with them. This lack of association renders the PHD filter unable to provide individual target track estimates, which is potentially problematic in the event of target trajectories crossing each other. Further development of the PHD filter to address this issue of track labelling has thus been the consideration in [20, 21, 22, 23, 24, 25, 26, 27]. Early attempts at incorporating data association with the PHD filter has been considered independently in [20] and [21]. Schemes presented by Panta, Vo, et. al. in [21], and more recently [22], propose the use of the PHD filter to reduce the size of the measurement sets on which data association techniques are applied. In contrast, a peak-to-track association is proposed by Lin, Bar-Shalom & Kirubarajan [20], whereby associations are established between peaks of the intensity function in the PHD filter over two subsequent time intervals in order to construct tracks. Examples of such track management schemes in practice can be found in [23, 24, 25, 26, 27].

In many tracking problems, the measurements received from a sensor, such as radar or sonar, often contain a signal strength (amplitude) along with position observations. The strength of the signal from a target is typically stronger than those from false alarms and provides a valuable source of information which can be incorporated into the PHD filter, as proposed independently by Clark, Ristic, et. al. [28] and Punithakumar, Kirubarajan & Sinha [29].

Most implementations of the PHD filter (see Chapter 2 for a description of typical implementations) assume that it is known *a priori* where the new targets are likely to appear. However, recent examples of multiple target tracking using the PHD filter can be found in [30, 31, 32], when there is no *a priori* spatial information

available on where new targets can appear. Motivated by the idea that targets are more likely to appear around measurements, the use of a diffuse target birth model is demonstrated by Houssineau & Laneuville [30] and Ristic, Clark & Vo [31, 32]. Specifically, in [31, 32], a new formulation of the PHD filter is presented which distinguishes between persistent and newly appearing targets.

The core results for the recursive equations in the PHD filter [1] address only single sensor scenarios. An approximation was proposed by Mahler [1, p. 1169] to extend the results to multi-sensor scenarios, that involves performing the PHD update iteratively for each set of measurements received from the multiple sensors, and referred to as the *iterated-corrector approximation*. This approximation is known to have a peculiarity, that changing the order each sensor's measurements are iterated, results in different update intensities. This was further investigated by Nagappa & Clark in [33], where a multi-sensor scenario was considered in which a single poor quality sensor (modelled with a low probability of detection) was used in combination with good sensors. Theoretically rigorous formulae for the multi-sensor PHD filter, though computationally intractable, can be derived as shown for the two sensor case by Mahler in [34] and for any number of sensors by Delande, Duflos et. al. in [35]. A simplification was also introduced in [35] that took into account the sensors' field-of-views, some of which may not overlap each other, resulting in tractable solutions.

1.1.2 Group target tracking

A group target is considered to be a collection of individual targets, each with dynamics that are no longer strictly independent of one another. Examples of group targets include: a formation of vehicles/aircraft; a flock of migrating birds; a shoal of fish, etc. Group target tracking, therefore, concerns the estimation of a dynamic system that consists of an unknown number of groups each with an unknown number of constituent targets. The estimation problem can present a substantial theoretical and practical challenge which is due, in part, to the difficulty that arises from processing measurements that are generated from the groups' constituents rather than the groups themselves.

Most existing approaches that address the group target tracking problem assume that the strong interdependencies observed between individual measurement data are due to a group formation pattern, whereby the individual target dynamics are superimposed on a common group effect. Such an approach was proposed by Gordon & Salmond in [36], and later applied in [37], in a Bayesian filtering framework where the state variable incorporates a vector representing the state of

all individual targets in the group and a vector representing the 'bulk' behaviour of the group. The number of targets in this application was considered to be constant and known *a priori*. A further application of this approach is demonstrated by Ristic, Arulampalam & Gordon in [38], where the 'bulk' vector describing the common motion allows for translation, rotation and scaling of the group target.

A more general approach to group target tracking is considered by Mahler in [39, 40], whereby a systematic statistical representation is proposed which provides the foundation for a natural generalisation of the multi-target Bayesian filter to group targets tracking problems. Furthermore, Mahler proposes approximation strategies based on generalisations of the PHD which will be revisited in more detail in this thesis (Chapter 3).

1.1.3 Extended target tracking

In extended target tracking, targets potentially produce more than one measurement per time-step, which may occur in scenarios where targets are close to a sensor, resulting in back-reflections. When back-reflections occur, the observation of a target consists of multiple point detections. The collection of measurements generated by each target can be modelled in such a way that they resemble a cluster of points, an approach proposed by Gilholm, Salmond et. al. [41, 42] (see Chapter 2 for details). However, the consideration of a structured arrangement of target generated measurements allows for the possibility of estimating the extent of targets, i.e. their shape and size (this consideration is formalised in Chapter 3 using the same statistical representation as for multiple group target systems).

Several approaches have been proposed in recent years for target extent estimation. One such approach, introduced by Koch [43], employs the use of random matrices whereby the extent of elliptical shaped targets are modelled by means of a symmetric positive definite matrix and is assumed to follow an inverse Wishart density. Further developed by Feldmann & Fränken [44, 45], regarding an extended target as a special case of a group target, the random matrices approach was demonstrated on a simulated example involving a group formation consisting of 5 individual targets travelling with constant velocity. An alternative approach was proposed by Baum, Hanebeck, et. al. [46] for modelling an elliptic extent using a so-called Random Hypersurface Model, previously introduced in [47] for arbitrary shapes and later developed for multiple extended targets [48] by the same principal authors. A detailed comparison of the random matrices approach and Random Hypersurface Model method is given in [49]. Both these approaches augment the target state into kinematic and extent parameter components. Such an

augmented state is considered in [50], where an extension of the standard Bayesian filter for single target tracking is proposed which bounds the target generated measurements by a circular region. Finally, objects which are line shaped or curved can be modelled using polynomials and tracked as extended targets, where the state estimates consist of the polynomial coefficients, as proposed by Lundquist, Orguner & Gustafsson [51] and applied to a road map estimation problem.

Methods based on the same mathematical foundation for which the multi-target Bayesian filter is formalised and from which the PHD filter is derived, have also been proposed by Granström, Lundquist & Orguner [52, 53] for estimating the extent. In [52], an implementation of the PHD filter for non-standard (extended) targets (see Chapter 2) is proposed, which involves calculating a likelihood function that relates a set of measurements, potentially, to a target whose structure resembles that of a rectangle or an ellipse. In [53], an augmented state representation of an extended target is considered, that consists of linear and non-linear components. The PHD filter is then applied for jointly estimating this target state and a set of measurement generating points defined on the boundary of each target's geometrically structured shape.

1.2 Scope of Thesis

The PHD filter has been the subject of great interest in the tracking community and proven to be applicable in a variety of different multiple target tracking scenarios, as emphasised in Section 1.1.1. The scope of this thesis is in the further development of the PHD filter for handling advanced multiple target systems, specifically those involving multiple group and extended targets as described in Sections 1.1.2 and 1.1.3 respectively. The focus is on developing such extensions of the PHD filter for the single sensor case only. This section provides an outline of the thesis' structure, along with the key contributions and related publications.

1.2.1 Organisation

Background material is provided in Chapter 2 which covers the notion of Bayesian filtering for the estimation problem in the context target tracking and particularly the fundamental theory behind the generalisation to multiple target problems. The same mathematical concepts needed to formulate the multi-target Bayesian filter are used in the derivation of the PHD filter, the resulting recursion for which is presented in Chapter 2.

Chapter 3 focuses on the extension of the PHD filter to group and extended

target tracking problems. Explicit formulae are derived for the recursive equations of these PHD filter extensions and constitute the main results of this thesis. Though the PHD filter results presented in Chapter 3 offer improved tractability over the generalised Bayesian filter for group/extended target problems, there still remains computational challenges in their potential implementations. Chapter 3 therefore concludes with a discussion on such limitations and a review of various methods available to address them.

Chapter 4 proposes a closed-form solution to the PHD filter for group targets and demonstrates its efficiency via simulations, while Chapter 5 considers a special case applicable to the PHD filter for extended targets. A direct application of the subsequent closed-form solution presented in Chapter 5 is demonstrated on a simulated example and its performance is assessed. An implementation of the PHD filter for extended targets is also proposed in Chapter 5 which incorporates target extent estimation with the addition of variable shape parameters.

The final chapter summarises the work presented in this thesis and discusses current related research.

1.2.2 Contributions

This thesis contributes to the development of the PHD filter, specifically in multiple group and extended target applications. The key original contributions are:

- The derivation of explicit formulae for the recursive equations in the PHD filters for group and extended targets.
- The formulation of closed-form solutions for the respective PHD filters.
- The successful implementation of the closed-form solutions, demonstrated via simulations.
- A novel implementation of the PHD filter for extended targets with estimable extent shape parameters.

1.2.3 Publications

The following papers (listed in chronological order) have been published based on the work presented in this thesis:

1. A. Swain and D. E. Clark, "First-moment filters for spatial independent cluster processes", in *Proc. SPIE*, vol. 7697, 2010.

2. A. Swain and D. E. Clark, "Extended object filtering using spatial independent cluster processes", in *13th Int. Conf. on Information Fusion*, (Edinburgh, UK), July 2010.

3. A. Swain and D. E. Clark, "Bayesian estimation of the intensity for independent cluster point processes: an analytic solution", *Procedia Environmental Sciences*, (Special Issue: Spatial Statistics 2011, Mapping Global Change), vol. 7, pp. 56–61, 2011.

4. A. Swain and D. E. Clark, "The single group PHD filter: an analytic solution", in *14th Int. Conf. on Information Fusion*, (Chicago, Illinois, USA), pp. 592–599, July 2011.

5. A. Swain and D. E. Clark, "The PHD filter for extended target tracking with estimable extent shape parameters of varying size", in *15th Int. Conf. on Information Fusion*, (Singapore), pp. 1111–1118, July 2012.

Chapter 2

Background

The notion of Bayesian filtering is essential to the development of this thesis. Standard Bayesian filtering, applicable to single target tracking problems, concerns the estimation of a random vector via its probability density given a sequence of measurement vectors. The non-trivial extension to multiple target tracking problems requires the notion of a random finite set and its probabilistic descriptor.

This chapter presents the random finite set formalism of the multiple target tracking problem in a Bayesian filtering framework and provides the fundamental basis for the focus of this thesis. Section 2.1 introduces the estimation problem in the context of target tracking and the standard Bayesian filter. A brief overview on the mathematical background of random finite sets is given in Section 2.2, including probabilistic descriptions and a few useful instances. Section 2.3 presents the formulation of the Bayesian filter for multiple target tracking in the random finite set framework, and discusses its general intractability as well as common approximations. The Probability Hypothesis Density (PHD) filter, a first-order moment approximation to the generalised Bayesian filter, is presented in Section 2.4, along with the cardinalized variant. Finally, performance evaluation for single and multiple target tracking is reviewed in Section 2.5.

2.1 Estimation and Target Tracking

Estimation is simply defined as the problem of estimating an unknown state $\mathbf{x}$ from a noisy observation $\mathbf{z}$. In the context of target tracking, the state is a vector containing information about a target under surveillance such as its position and velocity, while the observation, also a vector, is a noise corrupted sensor measurement such as a radar return. In the Bayesian filtering framework, the state and measurement are treated as random variables and the state is estimated via its probability density as summarised in Section 2.1.1.

2.1.1 Bayesian filtering

The following summary of the Bayesian filter considers the estimation problem in discrete time. The time evolution of the state vector is defined by the dynamic equation:

$$\mathbf{x}_k = f_k(\mathbf{x}_{k-1}, \mathbf{v}_{k-1}), \tag{2.1}$$

where f_k is a function that transforms any given state vector $\mathbf{x}_{k-1}$ and process noise $\mathbf{v}_{k-1}$ at time $k-1$ into a new state vector $\mathbf{x}$ at time k. The vector $\mathbf{v}_{k-1}$ represents the uncertainty of the target's motion. This dynamic model can also be described, probabilistically, by the Markov transition density $f_{k|k-1}(\mathbf{x}_k \,|\, \mathbf{x}_{k-1})$. The relation between the measurement and the state is defined by the measurement equation:

$$\mathbf{z}_k = h_k(\mathbf{x}_k, \mathbf{w}_k), \tag{2.2}$$

where h_k is a function that transforms any given state vector $\mathbf{x}_k$ and observation noise $\mathbf{w}_k$ at time k into a measurement vector $\mathbf{z}_k$. The vector $\mathbf{w}_k$ represents the observation error. This measurement model can also be described, probabilistically, by the likelihood function $g_k(\mathbf{z}_k \,|\, \mathbf{x}_k)$. Note that the process and observation noise are uncorrelated.

The Bayesian filter estimates the state vector recursively in time via the probability density $p_k(\mathbf{x}_k \,|\, \mathbf{z}_{1:k})$ given a sequence of measurements vectors $\mathbf{z}_{1:k} = \mathbf{z}_1, \ldots, \mathbf{z}_k$. Suppose the prior density $p_{k-1}(\mathbf{x}_{k-1} \,|\, \mathbf{z}_{1:k-1})$ at time $k-1$ is known, then the standard Bayesian filter recursion is given by the predicted and updated densities

$$p_{k|k-1}(\mathbf{x}_k \,|\, \mathbf{z}_{1:k-1}) = \int f_{k|k-1}(\mathbf{x}_k \,|\, \mathbf{x}_{k-1})\, p_{k-1}(\mathbf{x}_{k-1} \,|\, \mathbf{z}_{1:k-1})\, \mathrm{d}\mathbf{x}_{k-1}, \tag{2.3}$$

$$p_k(\mathbf{x}_k \,|\, \mathbf{z}_{1:k}) = \frac{g_k(\mathbf{z}_k \,|\, \mathbf{x}_k)\, p_{k|k-1}(\mathbf{x}_k \,|\, \mathbf{z}_{1:k-1})}{\displaystyle\int g_k(\mathbf{z}_k \,|\, \mathbf{x}_k)\, p_{k|k-1}(\mathbf{x}_k \,|\, \mathbf{z}_{1:k-1})\, \mathrm{d}\mathbf{x}_k}, \tag{2.4}$$

and illustrated by the following diagram:

$$\cdots \rightarrow p_{k-1}(\mathbf{x}_{k-1} \,|\, \mathbf{z}_{1:k-1}) \xrightarrow{\text{prediction}} p_{k|k-1}(\mathbf{x}_k \,|\, \mathbf{z}_{1:k-1}) \xrightarrow{\text{update}} p_k(\mathbf{x}_k \,|\, \mathbf{z}_{1:k}) \rightarrow \cdots$$

Solutions to the Bayesian recursion vary, depending on certain dynamic and measurement model specifications. A closed-form solution exists under linear Gaussian model assumptions as shown in Section 2.1.2, while Section 2.1.3 discusses non-linear approaches.

2.1.2 Linear estimation with the Kalman filter

When the dynamic and measurement models are linear transformations with additive Gaussian noise then the estimation problem given in the Bayesian filter has a closed-form solution known as the Kalman filter [54, 55, 56]. Specifically, the Kalman filter assumes the dynamic and measurement equations are given as

$$\mathbf{x}_k = \mathbf{F}_{k-1}\mathbf{x}_{k-1} + \mathbf{v}_{k-1}, \tag{2.5}$$

$$\mathbf{z}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{w}_k, \tag{2.6}$$

where $\mathbf{F}_{k-1}$ is a transition matrix, $\mathbf{H}_k$ is a projection matrix, while $\mathbf{v}_{k-1}$ and $\mathbf{w}_k$ are independent zero-mean Gaussian noise variables with covariances matrices $\mathbf{Q}_{k-1}$ and $\mathbf{R}_k$ respectively. Hence, the transition density and likelihood function are

$$f_{k|k-1}(\mathbf{x}_k \,|\, \mathbf{x}_{k-1}) = \mathcal{N}\left(\mathbf{x}_k \,;\, \mathbf{F}_{k-1}\,\mathbf{x}_{k-1}, \mathbf{Q}_{k-1}\right), \tag{2.7}$$

$$g_k(\mathbf{z}_k \,|\, \mathbf{x}_k) = \mathcal{N}\left(\mathbf{z}_k \,;\, \mathbf{H}_k\,\mathbf{x}_k, \mathbf{R}_k\right), \tag{2.8}$$

where $\mathcal{N}(\cdot\,;\mathbf{m},\mathbf{P})$ denotes a Gaussian density with mean and covariance $\mathbf{m}$ and $\mathbf{P}$ respectively.

Under these assumptions, the Kalman filter recursion proceeds as follows. Given the prior probability density is a Gaussian function of the form

$$p_{k-1}(\mathbf{x}_{k-1} \,|\, \mathbf{z}_{1:k-1}) = \mathcal{N}\left(\mathbf{x}_{k-1} \,;\, \mathbf{m}_{k-1}, \mathbf{P}_{k-1}\right), \tag{2.9}$$

the predicted probability density is also a Gaussian function of the form

$$p_{k|k-1}(\mathbf{x}_k \,|\, \mathbf{z}_{1:k}) = \mathcal{N}\left(\mathbf{x}_k \,;\, \mathbf{m}_{k|k-1}, \mathbf{P}_{k|k-1}\right), \tag{2.10}$$

where the corresponding mean and covariance are given by

$$\mathbf{m}_{k|k-1} = \mathbf{F}_{k-1}\,\mathbf{m}_{k-1}, \qquad \mathbf{P}_{k|k-1} = \mathbf{F}_{k-1}\mathbf{P}_{k-1}\mathbf{F}_{k-1}^T + \mathbf{Q}_{k-1}, \tag{2.11}$$

and subsequently the updated probability density is a Gaussian of the form

$$p_k(\mathbf{x}_k \,|\, \mathbf{z}_{1:k}) = \mathcal{N}\left(\mathbf{x}_k \,;\, \mathbf{m}_k, \mathbf{P}_k\right), \tag{2.12}$$

where the corresponding mean and covariance are given by

$$\mathbf{m}_k = \mathbf{m}_{k|k-1} + \mathbf{K}_k\left(\mathbf{z}_k - \mathbf{H}_k\,\mathbf{m}_{k|k-1}\right), \qquad \mathbf{P}_k = \left(\mathbf{I} - \mathbf{K}_k\,\mathbf{H}_k\right)\mathbf{P}_{k|k-1}. \tag{2.13}$$

The matrix $\mathbf{K}_k$ is referred to as the Kalman gain and is given by

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{H}_k^T \left(\mathbf{H}_k \, \mathbf{P}_{k|k-1} \, \mathbf{H}_k^T + \mathbf{R}_k\right)^{-1}. \tag{2.14}$$

Essentially, given linear Gaussian dynamic and measurement models, the Kalman filter propagates the probability density as a Gaussian function over time.

2.1.3 Non-linear estimation

The Kalman filter is not applicable when the dynamic and/or measurement models are non-linear. For non-linear estimation problems, approximations are necessary. A selected number of approximate methods are discussed here, further details of which can be found, collectively, in [38].

One such method assumes the process and observation noise are still independent zero-mean Gaussian variables and approximates the non-linear functions in the dynamic and measurement equations by the first term in their Taylor series expansions. This approximate method is known as the *extended Kalman filter* (EKF) [56] and it simply linearises the non-linear functions in the model equations so that the Kalman filter can be applied.

Rather than approximating the non-linear functions in the model equations, an alternative method known as the *unscented Kalman filter* (UKF) [57], approximates the probability density propagated in the Bayesian filter with a Gaussian function which is represented by a set of deterministically chosen sample points. As non-linear approximations of the Kalman filter, the EKF and UKF both operate in the framework of a Gaussian approximation for the probability density $p_k(\mathbf{x}_k \,|\, \mathbf{z}_{1:k})$ so that their implementation is simple.

Other approximate methods worth mentioning for non-linear estimation include Gaussian sum filters and particle filters. Gaussian sum filters [58, 59] approximate the probability density $p_k(\mathbf{x}_k \,|\, \mathbf{z}_{1:k})$ with a Gaussian mixture (a weighted sum of Gaussian functions, such that the weights sum to 1). Approximations of this type are applicable when the probability density is multimodal, such as is applicable for tracking a manoeuvring target. Particle filters [60] on the other hand perform sequential Monte Carlo (SMC) estimation based on an approximate particle representation of the probability density. The particles are random sample points, typically drawn from the transition density. Particle filters are very computationally demanding and in practice are generally used when other methods for non-linear estimation, such as the EKF and UKF, produce unsatisfactory results.

2.1.4 Traditional multiple target tracking methods

Early development of multiple target tracking techniques combined correlation with Kalman filtering theory [61, 62, 63, 64]. Correlation refers the process of determining a one-to-one assignment of targets and observations, based on the fundamental assumption that each target generates at most one observation. The correlation procedure consists of two stages and is performed in parallel with Kalman filtering methods. The first stage in the correlation procedure is gating, used to determine which observations have potentially been generated by each target, given the predicted target state estimates obtained from performing multiple Kalman predictions, as illustrated in Figure 2.1.

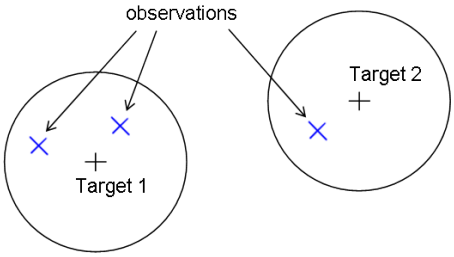


Figure 2.1: An example of gating for two targets where their predicted positions are indicated with black crosses.

The next step is to make final observation-to-target assignments and perform multiple Kalman updates. In the case where a single observation is within the gate of a single target, the assignment is trivial. However for closely spaced targets, as shown in the example illustrated in Figure 2.2, the one-to-one assignment is less trivial. Conflict situations arise when multiple observations fall within the same gate (or gates) and when observations fall within the gates of more than one target. The most common approach to this assignment problem is the *nearest neighbour*

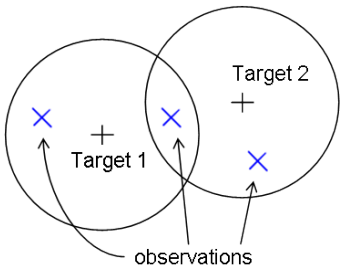


Figure 2.2: An example of gating for two closely spaced targets.

approach which looks for unique pairings that minimises the total summed distance between observations and their assigned targets. Two computationally efficient

suboptimal solutions to this assignment problem based on the nearest neighbour approach were reviewed by Blackman in [65] and shown to often produce differing solutions to one another. *Data association* methods, discussed next, further develop the concept of observation-to-target assignment.

The data association problem concerns the optimal assignment of observations into *tracks*. Multiple hypothesis tracking (MHT) [62, 66] is a data association based method that considers all possible combinations of such assignments (referred to as hypotheses) for all observations recevied up to the current time step in order to determine the most likely tracks. Alternatively the joint probabilistic data association (JPDA) method [64], an extension of probabilistic data association (PDA) proposed by Bar-Shalom and Tse in [67] for tracking a single target in a cluttered environment, uses joint association events and probabilities to avoid conflicting hypotheses in the presence of multiple targets. JPDA can be regarded as a special case of MHT that does not require the storage of past observations nor multiple hypotheses.

Data association methods suffer from the combinatorial growth in the number of hypotheses with the number of observations and targets. For this reason JPDA and MHT methods can not feasibly be considered for tracking scenarios that involve dense target environments and will not be pursued further in this thesis. Instead the remainder of this chapter focuses on association free methods based on the theory of random finite sets.

2.2 Random Finite Sets

Fundamental to the formulation of the Bayesian filter for multiple target tracking, is the idea of regarding states and observations as random finite sets. This section reviews the theory behind random finite sets, which was introduced by Matheron [68], in the context of integral geometry, and exploited by Mahler [69] for multiple target tracking.

2.2.1 Definition and probabilistic descriptor

A random finite set (RFS) is simply a random variable that takes values as (unordered) finite sets, i.e. a finite-set valued random variable. Essentially, a RFS consists of a random number of constituent points (cardinality) and the points themselves are random, distinct and unordered. Note that, since a RFS Ξ is without ordering, an instance $\Xi = X = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$, where the constituents $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are distinct, is equivalently represented by $n!$ random vectors in the augmented

product space $\mathfrak{X}^n = \mathfrak{X} \times \cdots \times \mathfrak{X}$ by arbitrarily permuting $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathfrak{X}$. For example, consider the instance where $n = 2$ and $\mathfrak{X} = \mathbb{R}^d$ for $d \geq 1$, i.e. $X = \{\mathbf{x}_1, \mathbf{x}_2\}$ where $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^d$. This instance can be equivalently represented by the two different random vectors $\begin{bmatrix}\mathbf{x}_1 & \mathbf{x}_2\end{bmatrix}^T$ and $\begin{bmatrix}\mathbf{x}_2 & \mathbf{x}_1\end{bmatrix}^T$. The importance of this property is shown in the following probabilistic description for a RFS.

Let $\mathcal{C}(\mathfrak{X})$ be the class of closed subsets of $\mathfrak{X}$[1]. The probability density of a RFS Ξ is the function $p_\Xi(X)$ such that

$$\int_S p_\Xi(X)\,\delta X = \Pr(\Xi \subseteq S), \tag{2.15}$$

for all $S \in \mathcal{C}(\mathfrak{X})$. Setting $S = \mathfrak{X}$ gives

$$\int_{\mathfrak{X}} p_\Xi(X)\,\delta X = \Pr(\Xi \subseteq \mathfrak{X}) = 1. \tag{2.16}$$

Furthermore, since the number of constituents of X is unknown, the probability $\Pr(\Xi \subseteq S)$ in equation (2.15) becomes

$$\begin{aligned}\Pr(\Xi \subseteq S) &= \sum_{n=0}^{\infty} \Pr(\Xi \subseteq S, |\Xi| = n) \\ &= \sum_{n=0}^{\infty} \int_{S^n} p_\Xi(\mathbf{x}_1, \ldots, \mathbf{x}_n)\,\mathrm{d}\mathbf{x}_1, \ldots, \mathrm{d}\mathbf{x}_n,\end{aligned} \tag{2.17}$$

where $p_\Xi(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ denotes the joint probability density of the random vectors $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathfrak{X}$ and $S^n = S \times \cdots \times S$ denotes the product space for positive integer n. Since the same finite set $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ is obtained for all $n!$ permutations of $\mathbf{x}_1, \ldots, \mathbf{x}_n$, the probability assigned to X must be equally distributed among the $n!$ vector permutations. That is, the probability density of the finite set X is

$$p_\Xi(X) = p_\Xi(\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}) = n!\, p_\Xi(\mathbf{x}_1, \ldots, \mathbf{x}_n). \tag{2.18}$$

The definition of a set integral follows immediately from equations (2.15), (2.17) and (2.18).

2.2.2 Set integral

Given a real-valued function $f(X)$ of a finite-set variable X, its integral over a region $S \subseteq \mathfrak{X}$ is

[1] A closed subset $S \in \mathcal{C}(\mathfrak{X})$ is not strictly finite

$$\begin{aligned}\int_S f(X)\,\delta X &:= \sum_{n=0}^{\infty} \frac{1}{n!} \int_{S^n} f(\{\mathbf{x}_1, \ldots \mathbf{x}_n\})\,\mathrm{d}\mathbf{x}_1, \ldots, \mathrm{d}\mathbf{x}_n \\ &= f(\emptyset) + \int_S f(\{\mathbf{x}\})\,\mathrm{d}\mathbf{x} + \frac{1}{2}\int_{S\times S} f(\{\mathbf{x}_1, \mathbf{x}_2\})\,\mathrm{d}\mathbf{x}_1\,\mathrm{d}\mathbf{x}_2 + \cdots .\end{aligned} \tag{2.19}$$

Note that the set integral in equation (2.19) accounts for the random variability of the number of constituents in X, hence the sum in (2.19) has, in general, an infinite number of terms. Furthermore, the scaling factor $1/n!$ arises as a consequence of the relation given in equation (2.18).

2.2.3 The point process theory equivalence

The key concepts presented so far in this section for random finite sets have an equivalence in point process theory [70]. That is, any point process in which the total number of points is finite with probability 1, is referred to as a finite point process. Intuitively, such a point process is considered as a finite multi-set consisting of random points, with possible repetition, i.e. $Y = \{\mathbf{y}_1, \ldots, \mathbf{y}_n\}$ where the constituent points $\mathbf{y}_1, \ldots, \mathbf{y}_n$ are not distinct. The statistical behaviour of a finite point process is characterised by *Janossy densities* [70, p. 137]. A finite point process is *simple* if and only if its Janossy densities exist (i.e. are finite-value functions) and vanish whenever $\mathbf{y}_i = \mathbf{y}_j$ for some $1 \leq i \neq j \leq n$ (i.e are completely symmetric in all arguments). A simple finite point process is therefore equivalent to a random finite set and the Janossy densities actually coincide with the probability densities $p_\Xi(\{\mathbf{x}_1, \ldots \mathbf{x}_n\})$ given in equation (2.18).

This relation between random finite sets and finite point processes was described by Mahler in the context of multiple target tracking [69], having been theoretically established by Ripley earlier in [71]. The decision whether to adopt one formalism over the other is simply a matter of preference and the RFS formalism will be maintained throughout this thesis hereafter. However, an important concept in point process theory is of particular interest with regards to the scope of this thesis and subject to further investigation in the next chapter.

2.2.4 Useful instances of random finite sets

A selected number of important RFS classes, which are commonly encountered in filtering methods for multiple target tracking and frequently referred to hereafter, are summarised below.

Independent identically distributed clusters

An independent identically distributed (i.i.d.) cluster RFS Ξ on $\mathfrak{X}$, with instance $\Xi = X = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$, is uniquely characterised by its cardinality distribution denoted by $\rho(n) = \Pr(|\Xi| = n)$ and probability density functions $p(\mathbf{x}_i)$ for $i = 1, \ldots, n$. The probability density of an i.i.d. cluster RFS Ξ is then written as

$$p_{\Xi}(X) = n!\,\rho(n)\prod_{i=1}^{n} p(\mathbf{x}_i). \tag{2.20}$$

Poisson

A Poisson RFS Ξ on $\mathfrak{X}$ is a particular type of i.i.d. cluster in which the cardinality distribution is Poisson, i.e. $\rho(n) = \mathrm{e}^{-\lambda}\,\lambda^n/n!$ with Poisson rate λ. The probability density of a Poisson RFS Ξ is then written as

$$p_{\Xi}(X) = \mathrm{e}^{-\lambda}\,\lambda^n \prod_{i=1}^{n} p(\mathbf{x}_i) = \mathrm{e}^{-\lambda}\prod_{i=1}^{n} \lambda\, p(\mathbf{x}_i). \tag{2.21}$$

A Poisson RFS Ξ is uniquely characterised by its intensity function denoted by $v(\mathbf{x}) = \lambda\, p(\mathbf{x})$.

Bernoulli

A Bernoulli RFS Ξ on $\mathfrak{X}$ has probability $1 - \rho$ of being empty and probability ρ of consisting of only one element, which is distributed according to the probability density $p(\cdot)$ defined on $\mathfrak{X}$. The probability density of a Bernoulli RFS Ξ is then

$$p_{\Xi}(X) = \begin{cases} 1 - \rho & \text{if } X = \emptyset, \\ \rho\, p(\mathbf{x}) & \text{if } X = \{\mathbf{x}\}, \\ 0 & \text{otherwise,} \end{cases} \tag{2.22}$$

where $\rho = \Pr(|\Xi| = 1)$ can be interpreted as an existence probability. Thus in this case, the cardinality distribution of RFS Ξ is Bernoulli with parameter ρ.

Multi-Bernoulli

A multi-Bernoulli RFS Ξ on $\mathfrak{X}$ is a union of a fixed number of independent Bernoulli RFS classes $\Xi^{(i)}$ with existence probability $\rho^{(i)}$ and probability density $p^{(i)}$ defined on $\mathfrak{X}$ for $i = 1, \ldots, N$. That is $\Xi = \Xi^{(1)} \cup \cdots \cup \Xi^{(N)}$ and the probability density of a multi-Bernoulli RFS Ξ is

$$p_{\Xi}(X) = \begin{cases} \prod_{i=1}^{N} \left(1 - \rho^{(i)}\right) & \text{if } X = \emptyset, \\ \prod_{i=1}^{N} \left(1 - \rho^{(i)}\right) \sum_{1 \leq i_1 \neq \cdots \neq i_n \leq N} \prod_{j=1}^{n} \frac{\rho^{(i_j)} \, p^{(i_j)}(\mathbf{x}_j)}{1 - \rho^{(i_j)}} & \text{if } |X| = n \leq N, \\ 0 & \text{otherwise.} \end{cases} \tag{2.23}$$

A multi-Bernoulli RFS is therefore completely described by the corresponding multi-Bernoulli parameter set $\{\rho^{(i)}, p^{(i)}\}_{i=1}^{N}$.

2.3 Multiple Target Tracking using Random Finite Sets

The formulation of the Bayesian filter for multiple target tracking presented in this section regards the unknown number of targets and their states as well as the collection of measurements as random finite sets. Let $X_k = \{\mathbf{x}_{1,k}, \ldots, \mathbf{x}_{n,k}\}$ denote the multi-target set at time-step k, whose constituent target states $\mathbf{x}_{1,k}, \ldots, \mathbf{x}_{n,k} \in \mathbb{R}^{d_x}$ are vectors with known dimension d_x. Further, let $Z_k = \{\mathbf{z}_{1,k}, \ldots, \mathbf{z}_{m,k}\}$ denote the measurement set at time-step k, whose constituent observation measurements $\mathbf{z}_{1,k}, \ldots \mathbf{z}_{m,k} \in \mathbb{R}^{d_z}$, received from multiple sensors, are vectors with known dimension d_z.

The objective of multiple target tracking is to estimate the RFS X_k, i.e. its cardinality $|X_k| = n$ and constituent target states $\mathbf{x}_{i,k}$ for $i = 1, \ldots, n$, at each time-step k given a sequence of measurement sets received up to time k, denoted by $Z_{1:k} = Z_1, \ldots, Z_k$. In the Bayesian filtering framework, the multi-target RFS X_k is estimated, as it evolves over time, via the posterior distribution $p_k(X_k \,|\, Z_{1:k})$. The posterior distributions assume the probabilistic description given in Section 2.2, i.e.

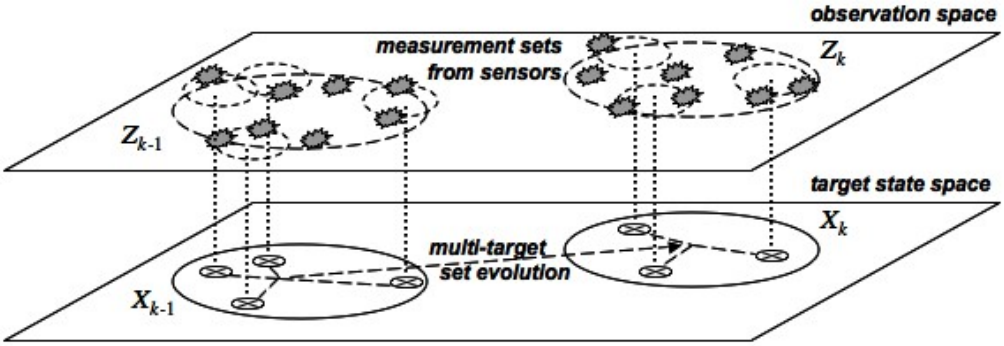


Figure 2.3: Visualisation of the observed multi-target set evolution over two consecutive time-steps

the probability density of a RFS $p(X)$ (noting that the subscript Ξ is dropped for simplicity and brevity). The time evolution of the observed multi-target set is depicted over two consecutive time intervals in Figure 2.3.

The RFS framework naturally allows for the incorporation of target birth and death, false measurements and detection uncertainty into the estimation problem described above. Such occurrences can be incorporated into the dynamic and measurement model as shown in the following two subsections.

2.3.1 Dynamic model

Multi-target dynamics involves a time-varying number of targets due to the appearance of new targets ('birth') and/or the disappearance of existing targets ('death'). The time evolution of the multi-target RFS is described by another RFS which is modelled as follows.

For a given multi-target RFS X_{k-1} at time $k-1$, the survival of each target state $\mathbf{x}_{k-1} \in X_{k-1}$ is modelled by the RFS

$$\Psi_{k|k-1}(\mathbf{x}_{k-1}) = \begin{cases} \{\mathbf{x}_k\} & \text{when the target survives,} \\ \emptyset & \text{otherwise.} \end{cases} \tag{2.24}$$

New targets at time k appear due to spontaneous births, which are modelled by the RFS Γ_k. The time evolution of the multi-target RFS from time $k-1$ to k, depicted in Figure 2.4, is modelled by the RFS X_k which is given by the union of the surviving targets and the spontaneous births, i.e.

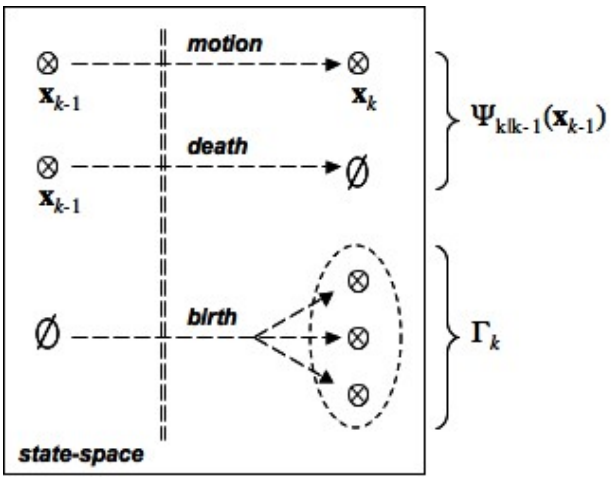


Figure 2.4: Visualisation of the dynamic model

$$X_k = \left(\bigcup_{\mathbf{x} \in X_{k-1}} \Psi_{k|k-1}(\mathbf{x}) \right) \cup \Gamma_k. \tag{2.25}$$

The RFSs constituting the union in equation (2.25) are assumed to be independent of each other and the randomness of this multi-target evolution is captured in the *multi-target transition density* $f_{k|k-1}(X_k \mid X_{k-1})$.

2.3.2 Measurement model

The RFS measurement model accounts for detection uncertainty and false measurements and is described as follows. For a given multi-target RFS X_k at time k, the detection of each target state $\mathbf{x}_k \in X_k$ is modelled by the following RFS

$$\Theta_k(\mathbf{x}_k) = \begin{cases} \{\mathbf{z}_k\} & \text{when a target is detected,} \\ \emptyset & \text{otherwise,} \end{cases} \tag{2.26}$$

which is representative of most standard multiple target tracking problems where each target generates no more than one measurement. Those measurements that are not target generated are considered to be false measurements, or clutter, and modelled by the RFS K_k. The collection of measurements, modelled by the RFS Z_k and depicted in Figure 2.5, is therefore given by the union of target generated measurements and clutter, i.e.

$$Z_k = \left(\bigcup_{\mathbf{x} \in X_k} \Theta_k(\mathbf{x}) \right) \cup K_k. \tag{2.27}$$

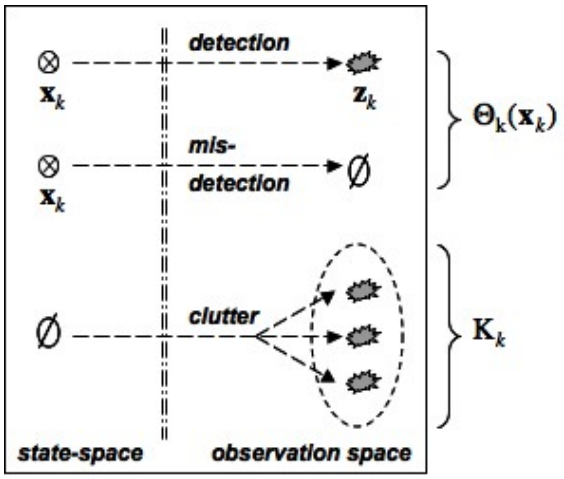


Figure 2.5: Visualisation of the measurement model

The RFSs constituting the union in equation (2.27) are assumed to be independent of each other and the randomness of the observations is captured in the *multi-target likelihood* $g_k(Z_k \mid X_k)$.

2.3.3 Bayesian filter

By modelling the multi-target state and the measurements as RFSs and given the previously defined notions of the probability density of a RFS and the set integral, the Bayesian filter for multiple target tracking is formulated as follows. Suppose that the prior density $p_{k-1}(X_{k-1} \mid Z_{1:k-1})$ is known, then the multi-target Bayesian filter recursion is given by the predicted and updated probability densities

$$p_{k|k-1}(X_k \mid Z_{1:k-1}) = \int f_{k|k-1}(X_k \mid X_{k-1})\, p_{k-1}(X_{k-1} \mid Z_{1:k-1})\, \delta X_{k-1}, \tag{2.28}$$

$$p_k(X_k \mid Z_{1:k}) = \frac{g_k(Z_k \mid X_k)\, p_{k|k-1}(X_k \mid Z_{1:k-1})}{\int g_k(Z_k \mid X_k)\, p_{k|k-1}(X_k \mid Z_{1:k-1})\, \delta X_k}. \tag{2.29}$$

That is, the Bayesian filter for multiple target tracking propagates the probability density of the multi-target RFS over time, given a sequence of measurement sets, as depicted in the following diagram:

$$\cdots \rightarrow p_{k-1}(X_{k-1} \mid Z_{1:k-1}) \xrightarrow{\text{prediction}} p_{k|k-1}(X_k \mid Z_{1:k-1}) \xrightarrow{\text{update}} p_k(X_k \mid Z_{1:k}) \rightarrow \cdots$$

In general, the multi-target Bayesian filter is intractable for a variable number of targets due to the combinatorial nature of set integrals (see Section 2.2.2), the computation of which is required in the recursive equations (2.28)-(2.29). None-theless, there exist several examples of the multi-target Bayesian filter having been applied to a small number of targets [72, 73, 74, 75]. However, for applications involving higher numbers of targets and/or measurements, approximations of the filter are necessary.

2.3.4 Approximations to the Bayesian filter

In principle, the multi-target Bayesian filter recursion can be implemented using sequential Monte Carlo (SMC) techniques [76] for approximating the posterior probability density $p_k(X_k \mid Z_{1:k})$ with a large number of weighted particles. Such SMC methods are predominantly employed for the applications addressed in [72, 73, 74, 75]. Suppose the number of targets, say n, is known, then a particle representation of the multi-target Bayesian filter requires the order of $n!N_p$ oper-

ations where N_p is the number of particles used, which generally needs to be a large number. Consequently, for large n also, SMC implementations via particle representations alone can be impractical for applications of multi-target Bayesian filtering, due to the combinatorial nature of the factor $n!$.

A more tractable approximation to the multi-target Bayesian filter was proposed by Mahler in [69] using the multi-Bernoulli RFS and so called the Multi-target Multi-Bernoulli (MeMBer) filter. Intuitively, the MeMBer filter propagates the posterior probability density in time by propagating a finite but time-varying number of target estimates, each characterised by the probability of existence and the probability density of the current estimated target state, i.e. the parameter set $\{\rho^{(i)}, p^{(i)}\}_{i=1}^{N}$ describing the multi-Bernoulli RFS (see Section 2.2.4). SMC implementations of the MeMBer fiter can be found in [77, 78, 79], and a recent modification of the MeMBer filter was proposed in [80] to incorporate unknown clutter and sensor field-of-view models.

It has been analytically shown in [81], that the MeMBer filter suffers from an over-estimation in the number of targets, a bias caused by the multi-Bernoulli approximation in the update step and resulting in a high number of false tracks. A novel multi-Bernoulli approximation was therefore proposed in [81, 82], called the cardinality-balanced (CB-)MeMBer filter, that alleviates the bias problem. The CB-MeMBer filter has recently been modified for the application of audio-visual multiple target tracking [83].

It is, however, the earlier first-order approximations proposed by Mahler, the Probability Hypothesis Density (PHD) filter [1] and, to a lesser extent, the cardinalized PHD (CPHD) filter [84], that are of particular interest with regards to this thesis. Such approximations to the mulit-target Bayesian filter, introduced in the next section, alleviate the intractability of propagating full posterior probability densities by instead propagating first-order moments which operate on the space occupying a single target.

2.4 Probability Hypothesis Density Filtering

This section summarises the concept of a probability hypothesis density, presents the recursion equations and describes its typical implementations, before briefly introducing the cardinalized PHD filter. The extension of the PHD filter to a non-standard target problem, specifically involving targets that generate multiple measurements (known as extended targets), is also presented.

2.4.1 The PHD filter recursion

The first-order moment density of a RFS Ξ is defined by the following expectation

$$v_{\Xi}(\mathbf{x}) = \mathsf{E}\left[\delta_{\Xi}(\mathbf{x})\right] = \int \delta_X(\mathbf{x})\, p_{\Xi}(X)\, \delta X, \tag{2.30}$$

where $\delta_X(\mathbf{x}) = 0$ if $X = \emptyset$ and, otherwise, $\delta_X(\mathbf{x}) = \sum_{\mathbf{w}\in X} \delta_{\mathbf{x}}(\mathbf{w})$ is a sum of Dirac delta functions $\delta_{\mathbf{x}}(\mathbf{w})$ concentrated at $\mathbf{x} \in \mathfrak{X}$. It was shown by Mahler [69], that equation (2.30) gives rise to the following formula

$$v_{\Xi}(\mathbf{x}) = \int p_{\Xi}(\{\mathbf{x}\} \cup W)\, \delta W = \int_{X \ni \mathbf{x}} p_{\Xi}(X)\, \delta X. \tag{2.31}$$

From the expression given in equations (2.30) and (2.31), the function $v_{\Xi} : \mathfrak{X} \to \mathbb{R}$ can be interpreted as the density of targets at $\mathbf{x}$ and is so called the *probability hypothesis density* (PHD) or *intensity*. The PHD is not a probability density though, since the integral of $v_{\xi}(\mathbf{x})$ in any region $S \subseteq \mathfrak{X}$ gives the expected number of targets of Ξ in that region, i.e.

$$\int_S v_{\Xi}(\mathbf{x})\, \mathrm{d}\mathbf{x} = \mathsf{E}\left[|\Xi \cap S|\right]. \tag{2.32}$$

Further details on Dirac delta functions, and the derivations of the expressions for the first-order moment density of a RFS Ξ and its integral over S, can be found in Appendix A.

The PHD filter [1] approximates the predicted and updated probability densities $p_{k|k-1}(X_k) = p_{k|k-1}(X_k \,|\, Z_{1:k-1})$ and $p_k(X_k) = p_k(X_k \,|\, Z_{1:k})$ with the respective first-order moment densities, denoted by $v_{k|k-1}(\mathbf{x}_k) = v_{k|k-1}(\mathbf{x}_k \,|\, Z_{1:k-1})$ and $v_k(\mathbf{x}_k) = v_k(\mathbf{x}_k \,|\, Z_{1:k})$, which are propagated over time as depicted in the following diagram:

$$\begin{array}{ccccccc}
\cdots & \longrightarrow & p_{k-1}(X_{k-1}) & \xrightarrow{\text{prediction}} & p_{k|k-1}(X_k) & \xrightarrow{\text{update}} & p_k(X_k) \longrightarrow \cdots \\
 & & \downarrow & & \downarrow & & \downarrow \\
\cdots & \longrightarrow & v_{k-1}(\mathbf{x}_{k-1}) & \xrightarrow{\text{prediction}} & v_{k|k-1}(\mathbf{x}_k) & \xrightarrow{\text{update}} & v_k(\mathbf{x}_k) \longrightarrow \cdots
\end{array}$$

Then, based on the following assumptions,

- each target evolves and generates measurements independently of one another, and independent of target births and clutter respectively,

- the RFS $\Psi_{k|k-1}$ modelling the survival uncertainty for each target evolving

from state $\mathbf{x}_{k-1}$ to $\mathbf{x}_k$, is Bernoulli with parameter $p_S(\mathbf{x}_{k-1})$ and transition density $f_{k|k-1}(\mathbf{x}_k \,|\, \mathbf{x}_{k-1})$ describing its evolution,

- the RFS Θ_k modelling the detection uncertainty for each target $\mathbf{x}_k \in X_k$, is also Bernoulli with parameter $p_D(\mathbf{x}_k)$ and a likelihood that $\mathbf{z}_k$ is observed from $\mathbf{x}_k$ given by $g_k(\mathbf{z}_k \,|\, \mathbf{x}_k)$,

- the clutter RFS K_k is Poisson with intensity $\kappa_k(\cdot)$,

- the predicted multi-target RFS X_k is approximated with a Poisson RFS prior to the measurement update,

the *PHD recursion* is given by

$$v_{k|k-1}(\mathbf{x}_k) = \gamma_k(\mathbf{x}_k) + \int p_S(\mathbf{x}_{k-1})\, f_{k|k-1}(\mathbf{x}_k \,|\, \mathbf{x}_{k-1})\, v_{k-1}(\mathbf{x}_{k-1})\, \mathrm{d}\mathbf{x}_{k-1}, \tag{2.33}$$

$$v_k(\mathbf{x}_k) = \big(1 - p_D(\mathbf{x}_k)\big) v_{k|k-1}(\mathbf{x}_k) + \sum_{\mathbf{z} \in Z_k} \frac{p_D(\mathbf{x}_k)\, g_k(\mathbf{z} \,|\, \mathbf{x}_k)\, v_{k|k-1}(\mathbf{x}_k)}{\kappa_k(\mathbf{z}) + v_{k|k-1}[p_D\, g_k(\mathbf{z} \,|\, \cdot)]}. \tag{2.34}$$

The intensity of the birth RFS Γ_k is denoted by $\gamma_k(\mathbf{x}_k)$ in equation (2.33) and $v_{k|k-1}\left[p_D\, g_k(\mathbf{z} \,|\, \cdot)\right] = \int p_D(\mathbf{x})\, g_k(\mathbf{z} \,|\, \mathbf{x})\, v_{k|k-1}(\mathbf{x})\, \mathrm{d}\mathbf{x}$ in equation (2.34). The derivation of the PHD recursion requires the use of probability generating functionals, an important concept in the theory of point processes which applies to random finite sets (see Appendix A). Full details on the derivation can be found in [1].

2.4.2 Implementations of the PHD filter

The PHD filter is typically implemented using a particle representation, i.e. sequential Monte Carlo (SMC) techniques [85, 86, 87, 76, 88], or with a Gaussian mixture formulation [89, 90, 91, 92]. Particle filtering methods have been independently proposed for the PHD filter in [85] and [86], which can be considered as special cases of the generalised SMC implementation of the PHD filter, so called the SMC-PHD filter, proposed in [87, 76]. SMC implementations approximate the PHD filter using a weighted set of particles at each iteration of the recursion. Alternatively, a closed-form solution to the PHD filter was established in [89, 90] under linear Gaussian model assumptions. The resulting filter is a recursion of Gaussian mixtures and so called the GM-PHD filter. Convergence results have been established for the particle PHD filter in [93, 94] and for the GM-PHD filter in [95].

While the GM-PHD filter is typically applied to linear estimation problems, variations have been developed for non-linear dynamic models in [92] and using

bearings-only measurements in [91]. A hybrid implementation of the PHD filter has also been proposed in [96], in which the Gaussian components in the GM-PHD filter are represented with Gaussian particle filters [97] similar to the Gaussian sum particle filter proposed in [98]. This hybrid implementation is applicable for non-linear estimation problems and results in a filter that no longer has a closed form solution and requires Monte Carlo integration for its approximation.

2.4.3 The cardinalized PHD filter

The propagated intensities of the PHD filter recursion have the unique property that their integral over any region of state space equates to the expected number of targets contained in that region, as shown by equation (2.32). That is, the PHD filter is not only first-order in the multiple target states, but also first-order in target number, since it propagates cardinality information with only a single parameter (the mean). As a result, the cardinality distribution is approximated with a Poisson distribution. Consequently, since the mean and variance of a Poisson distribution are equal, when the true cardinality is high the corresponding estimate has high variance, potentially leading to oversensitive cardinality estimates in practice, particularly when the probability of detection is low.

A generalisation of the PHD filter was therefore proposed by Mahler in [84] which remains first-order in the target states, but relaxes the first-order assumption on the number of targets, and is known as the cardinalized PHD filter. Not only does the CPHD filter propagate first-order moment densities over time, but also the cardinality distribution, as depicted in the following diagram:

$$\begin{array}{ccccccc}
\cdots & \longrightarrow & p_{k-1}(X_{k-1}) & \xrightarrow{\text{prediction}} & p_{k|k-1}(X_k) & \xrightarrow{\text{update}} & p_k(X_k) \longrightarrow \cdots \\
 & & \downarrow & & \downarrow & & \downarrow \\
\cdots & \longrightarrow & \begin{cases} v_{k-1}(\mathbf{x}_{k-1}) \\ \rho_{k-1}(n) \end{cases} & \xrightarrow{\text{prediction}} & \begin{cases} v_{k|k-1}(\mathbf{x}_k) \\ \rho_{k|k-1}(n) \end{cases} & \xrightarrow{\text{update}} & \begin{cases} v_k(\mathbf{x}_k) \\ \rho_k(n) \end{cases} \longrightarrow \cdots
\end{array}$$

The Cardinalized Probability Hypothesis Density (CPHD) filter generalises the PHD filter by relaxing the Poisson assumption on the number of targets and instead approximating the multi-target RFS with an i.i.d. cluster RFS (see Section 2.2.4). The prediction and update equations for the CPHD recursion are explicitly given in [84] along with details on their derivations, and a Gaussian mixture implementation can be found in [99].

2.4.4 The PHD filter for non-standard targets

Most standard multiple target tracking applications, including the PHD and CPHD filters, assume that each target generates no more than one measurement. Suppose, however, that targets potentially produce more than one measurement, so that the RFS modelling target generated measurements is now given by

$$\Theta_k(\mathbf{x}_k) = \begin{cases} Z_{\mathbf{x},k} & \text{when a target is detected,} \\ \emptyset & \text{otherwise,} \end{cases} \tag{2.35}$$

where $Z_{\mathbf{x},k}$ is also a RFS. Targets that generate measurement sets in this way are otherwise known as *extended targets*. An extension of the PHD filter that handles extended targets was introduced by Mahler in [100], where the RFS $Z_{\mathbf{x},k}$ is modelled with a Poisson RFS, as was first proposed in [41, 42].

The PHD filter recursion for extended target tracking and the recursion given in Section 2.4.1, only differ in the update equation. Based on the following assumptions,

- each target generates a set of measurements independently of one another, and independent of clutter,
- the clutter RFS K_k is Poisson with intensity $\kappa_k(\cdot)$,
- the RFS Θ_k modelling the detection uncertainty for each target $\mathbf{x}_k \in X_k$, is Bernoulli with parameter $p_D(\mathbf{x}_k)$ and a likelihood that the set $Z = Z_{\mathbf{x},k}$ is observed from $\mathbf{x}_k$ given by $g_k(Z \,|\, \mathbf{x}_k) = \mathrm{e}^{-\alpha(\mathbf{x}_k)} \prod_{\mathbf{z}\in Z} \alpha(\mathbf{x}_k)\, l(\mathbf{z} \,|\, \mathbf{x}_k)$, where $\alpha(\mathbf{x}_k)$ is the expected number of measurements generated by $\mathbf{x}_k$ and $l(\mathbf{z} \,|\, \mathbf{x}_k)$ is the single-measurement likelihood,
- the predicted multi-target RFS X_k is approximated with a Poisson RFS,

the update equation in the RFS filter for extended targets is then

$$\begin{aligned} v_k(\mathbf{x}_k) = v_{k|k-1}(\mathbf{x}_k) \Bigg\{ & 1 - p_D(\mathbf{x}_k) + p_D(\mathbf{x}_k)\, \mathrm{e}^{-\alpha(\mathbf{x}_k)} \\ & + \sum_{\pi\in\Pi_{Z_k}} \varpi_\pi \sum_{\varphi\in\pi} \frac{p_D(\mathbf{x}_k)\, \mathrm{e}^{-\alpha(\mathbf{x}_k)} \left(\prod_{\mathbf{z}\in\varphi} \alpha(\mathbf{x}_k)\, l(\mathbf{z} \,|\, \mathbf{x}_k)/\kappa_k(\mathbf{z}) \right)}{\delta_{|\varphi|,1} + v_{k|k-1} \left[p_D\, \mathrm{e}^{-\alpha} \left(\prod_{\mathbf{z}\in\varphi} \alpha\, l_{\mathbf{z}}\, /\kappa_k(\mathbf{z}) \right) \right]} \Bigg\}, \end{aligned} \tag{2.36}$$

denoting $l_{\mathbf{z}} = l(\mathbf{z} \,|\, \cdot)$. The summation $\sum_{\pi\in\Pi_{Z_k}}$ is taken over all possible partitions of the measurement set Z_k, denoted by Π_{Z_k}. A single partition, denoted by π,

contains subsets $\varphi \in \pi$ whose constituent elements consist of measurements in Z_k such that $\bigcup_{\varphi\in\pi} \varphi = Z_k$. Partitioning of the measurement set will be looked at in more detail in the next chapter. The partition weights are denoted by ϖ_π and given by

$$\varpi_\pi = \frac{\prod\limits_{\varphi\in\pi} \left(\delta_{|\varphi|,1} + v_{k|k-1}\left[p_D\, \mathrm{e}^{-\alpha}\left(\prod_{\mathbf{z}\in\varphi} \alpha\, l_{\mathbf{z}}\, /\kappa_k(\mathbf{z})\right)\right]\right)}{\sum\limits_{\pi'\in\Pi_{Z_k}} \prod\limits_{\varphi'\in\pi'} \left(\delta_{|\varphi'|,1} + v_{k|k-1}\left[p_D\, \mathrm{e}^{-\alpha}\left(\prod_{\mathbf{z}\in\varphi'} \alpha\, l_{\mathbf{z}}\, /\kappa_k(\mathbf{z})\right)\right]\right)}, \tag{2.37}$$

where $\delta_{|\varphi|,1}$ is the Kronecker delta[2]. For details on the derivation refer to [100]. A GM implementation of the PHD filter for extended targets can be found in [101].

A cardinalized variant of the PHD filter for extended target tracking has since been proposed in [102, 103], whereby the Poisson assumption on the number of targets and the number of measurements is relaxed. Like the CPHD filter, the multi-target RFS is modelled as an i.i.d. cluster RFS and in addition the RFS of target generated measurements is also modelled as an i.i.d. cluster. This so called CPHD filter for extended targets has been implemented with a Gaussian mixture formulation in [104].

2.5 Performance Evaluation

Assessing the performance of a filter, whether for single or multiple target estimation/tracking problems, requires the notion of a metric describing the error between the true and estimated target state(s). For single target estimation, such a notion is well established in the form of, for example, the root mean squared error, the details of which are given in Section 2.5.1. On the other hand, for multiple target tracking problems, a suitable metric describing the error between set-valued states must account for errors in the cardinality as well as the state estimation error. One such method is introduced in Section 2.5.2.

2.5.1 The root mean squared error

The following details for the root mean squared error (RMSE) can be found in [64], which provides an appropriate metric for describing the state estimation error, especially in linear estimation problems. Let $\mathbf{x}$ and $\mathbf{y}$ be the true and estimated state vectors respectively, then the root mean squared error is given by

[2]The Kronecker delta is defined to be $\delta_{|W|,1} = 1$ if $|W| = 1$ and $\delta_{|W|,1} = 0$ otherwise

$$\mathrm{RMSE}(\mathbf{y}) = \sqrt{\mathrm{MSE}(\mathbf{y})}, \tag{2.38}$$

where $\mathrm{MSE}(\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|^2$ is the mean squared error defined by the squared norm of the estimation error, so that

$$\mathrm{RMSE}(\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|. \tag{2.39}$$

For linear estimation errors, the Euclidean distance is suitable for evaluating the norm, i.e.

$$\|\mathbf{x}-\mathbf{y}\| = \sqrt{(\mathbf{x}-\mathbf{y})^T(\mathbf{x}-\mathbf{y})}. \tag{2.40}$$

2.5.2 The OSPA metric

The optimal subpattern assignment (OSPA) metric, proposed in [105], accounts for differences in the cardinality and the individual elements between two finite sets in a mathematically consistent way, that allows for a meaningful physical interpretation. Its definition is given as follows.

Let $d^{(c)}(\mathbf{x},\mathbf{y}) := \min(c, \|\mathbf{x}-\mathbf{y}\|)$ denote the distance between $\mathbf{x},\mathbf{y} \in \mathfrak{X}$ with a cut-off at $c > 0$ and Ω_k denote the set of permutations on $\{1,2,\ldots,k\}$ for any positive integer k. Then the OSPA metric, denoted by $\bar{d}_p^{(c)}$, for $p \geq 1$, $c > 0$ and arbitrary finite subsets $X = \{\mathbf{x}_1,\ldots,\mathbf{x}_m\}$ and $Y = \{\mathbf{y}_1,\ldots,\mathbf{y}_n\}$ is given by

$$\bar{d}_p^{(c)}(X,Y) := \left(\frac{1}{n}\left(\min_{\varsigma\in\Omega_n}\sum_{i=1}^{m} d^{(c)}(\mathbf{x}_i,\mathbf{y}_{\varsigma(i)})^p + c^p(n-m)\right)\right)^{1/p}, \tag{2.41}$$

if $m \leq n$, and $\bar{d}_p^{(c)}(X,Y) := \bar{d}_p^{(c)}(Y,X)$ if $m > n$ ($\bar{d}_p^{(c)}(X,Y) = \bar{d}_p^{(c)}(Y,X) = 0$ if $m = n = 0$). The OSPA distance is interpreted as a p^{th}-order per-target error, comprised of a p^{th}-order per-target localisation error and a p^{th}-order per-target cardinality error, defined respectively as follows

$$\begin{aligned} \bar{e}_{\text{loc}}^{(c)}(X,Y) &:= \left(\frac{1}{n}\min_{\varsigma\in\Omega_n}\sum_{i=1}^{m} d^{(c)}(\mathbf{x}_i,\mathbf{y}_{\varsigma(i)})^p\right)^{1/p}, \\ \bar{e}_{\text{card}}^{(c)}(X,Y) &:= \left(\frac{c^p(n-m)}{n}\right)^{1/p}, \end{aligned} \tag{2.42}$$

if $m \leq n$, and $\bar{e}_{\text{loc}}^{(c)}(X,Y) := \bar{e}_{\text{loc}}^{(c)}(Y,X)$, $\bar{e}_{\text{card}}^{(c)}(X,Y) := \bar{e}_{\text{card}}^{(c)}(Y,X)$ if $m > n$. The order parameter p determines the sensitivity of the metric to outliers, and the

cut-off parameter c determines the relative weighting of the penalties assigned to cardinality and localisation errors. When $p = 1$, the OSPA distance is equal to the sum of the 'per-target' localisation and cardinality components, which facilitates a direct interpretation of the metric. The choice of the cut-off parameter c depends on whether the accurate estimation of the the number of objects is more or less important than accurate position estimates and can be decided upon using the following guidelines.

Small values of c which corresponds to the magnitude of a typical localisation error, determined by, for example, the observation noise, has the effect of emphasising cardinality errors. Large values of c which corresponds to the maximal distance between targets has the effect of emphasising localisation errors. Any moderate value of c which is significantly larger than a typical localization error, but significantly less than the maximal distance between targets, maintains a balance between the two components.

Chapter 3

Group and Extended Target Tracking

Random finite sets provide a natural and rigourous framework for multi-target Bayesian filtering and its approximations, including the Probability Hypothesis Density (PHD) filter. The focus of this chapter is on the further development of the PHD filter for handling more advanced target tracking problems. In particular, the advanced target tracking problems of interest are those involving multiple group/extended targets. A group target is regarded as a collection of targets that share a common motion or characteristic, while an extended target is regarded as a target that potentially generates multiple measurements. The PHD filter results presented in this chapter provide a computationally tractable basis for group and extended target tracking, the implementations of which are the subject of subsequent chapters.

Section 3.1 introduces hierarchical RFS formalisms, based on a fundamental concept from point process theory, of group and extended target tracking problems in a Bayesian filtering framework. Section 3.1 concludes with considerations for first-order moment approximations based on generalisations of the PHD. Section 3.2 presents the recursive equations of the PHD filter for group target tracking and subsequent special cases, including the extended target tracking problem. Finally, limitations of the first-order moment filters for group and extended target systems are discussed in Section 3.3.

3.1 Bayesian Filter for Group and Extended Targets

The generalisation of the Bayesian filter to group (and extended) target tracking requires a systematic statistical representation based on an important concept in the theory of point processes, known as the cluster process. This foundation is introduced in the next subsection before specifying the RFS analogue in the representation of a multiple group (or extended) target system whose estimation is the objective of the corresponding Bayesian filter.

3.1.1 Cluster processes

Cluster processes are an essential concept in the theory of point processes [70], stochastic geometry [106] and spatial population processes [107]. A cluster process is a doubly-stochastic process and can be described as a hierarchical point process in which individual points are grouped into independent clusters. An early example of this can be found in [108] where Neyman and Scott proposed the stochastic process of clustering, specifically a Poisson cluster process, as a mathematical model of the Universe. Here the so called Neyman-Scott process model consists of a Poisson distributed number of independent and identically spatially distributed cluster centres, which collectively form an unseen Poisson point process. Each cluster centre has associated with it a random number of points that are independent from one another and identically distributed about the centre, which collectively form a subsidiary i.i.d. cluster process. Figure 3.1 illustrates a realisation of the Thomas cluster process, a special case of the Neyman-Scott process in which the cluster centres are distributed uniformly and each cluster consists of a Poisson distributed number of points whose positions are isotropic Gaussian displacements from the centre, i.e. the subsidiary processes are Poisson.

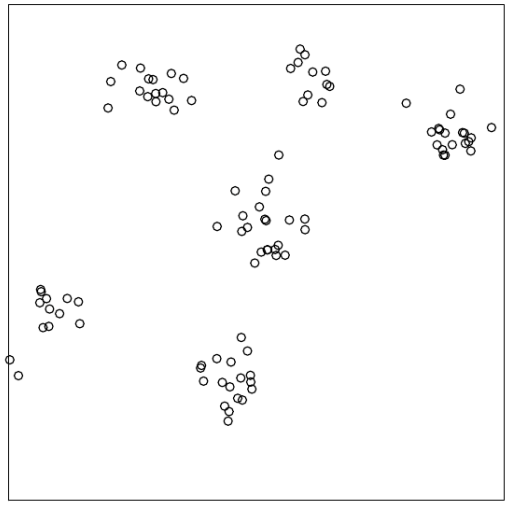


Figure 3.1: A realisation of the Thomas cluster process with Poisson rate 6 for the cluster centre process, Poisson rate 12 and standard deviation 0.04 for each subsidiary process.

There exist numerous applications of statistical inference for cluster processes, some of which were discussed by Neyman and Scott in [109]. For example population dynamics, an application referring to the development of a biological population such as weeds, insects or bacteria. It was noted by Harris in [110] that the process of clustering for the application in population dynamics is closely related

to general branching processes, which concerns the sequence of the number of individuals in generations of a population. A cluster process is regarded as a segment of a branching process, which consists of only two generations, the cluster centres constitute the parent generation and collectively their descendants constitute the next generation. For this reason, the process of cluster centres will be frequently referred to hereafter as the *parent* and the subsidiary processes as *daughters*.

Cluster processes provide an appropriate statistical representation of multiple group/extended target systems, whereby the constituent members of the parent process represent the group states while the constituent members of each daughter process represent the individual target states belonging to a single group. This representation can be written in terms of random finite sets as shown in the following subsection.

3.1.2 Group and extended target cluster models

The estimation problem for a multiple group target system, unlike the multiple target problem in the previous chapter, consists of three levels: a *doubly hidden parent level* (the measurable space in which the unknown group states occupy); a *singularly hidden daughter level* (the measurable space in which the unknown individual target states occupy); and a *visible observation level* (the measurable space in which the known observation measurements occupy). This is illustrated in Figure 3.2. In this problem, the system of unknown quantities can be represented by a cluster process and written in terms of the RFS framework as follows.

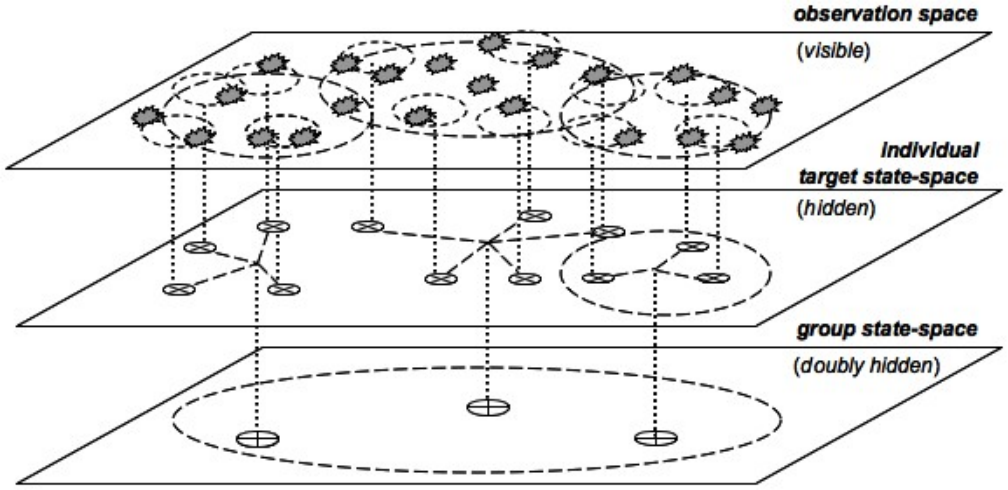


Figure 3.2: Visualisation of the multiple group estimation problem. The known measurements are direct observations of the unknown individual target states and indirect observations of the unknown group states.

Underlying everything is the RFS Υ (parent process), a random variable with instance $\Upsilon = X = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ where X is a finite set containing group state vectors $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{R}^{d_x}$. Each group state $\mathbf{x}_i$ has a dependent RFS $\mathcal{E}_i$ (daughter process), a random variable with instance $\mathcal{E}_i = \Xi_i = \{\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_{n_i}\}$ where Ξ_i is a finite set containing individual target state vectors $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_{n_i} \in \mathbb{R}^{d_\xi}$. The multi-group set X is indirectly observable, since those measurements in the finite set $Z = \{\mathbf{z}_1, \ldots, \mathbf{z}_m\}$ that have been generated by targets are observations of the individual targets $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_{n_i}$ for $i = 1, \ldots, |X|$.

The complete state of a multiple group target system can be specified in one of two ways. The first specifies the complete state as a finite set given by $\mathbb{X} = \{(\mathbf{x}_1, \Xi_1), \ldots, (\mathbf{x}_n, \Xi_n)\}$. The statistical representation of a system with instance $\mathbb{X}$ is therefore the random finite subset $\mathcal{X} = \bigcup_{\mathbf{x}\in\Upsilon}\{(\mathbf{x}, \mathcal{E}^{\mathbf{x}})\}$ on the space of pairs, where $\mathcal{E}^{\mathbf{x}}$ is the RFS on individual target space associated with group $\mathbf{x}$. This is consistent with Moyal's definition of a cluster process found in [107, p. 24].

An alternative specification for the complete state of a multiple group target system is the finite set now given by $\mathbb{X} = \{\Xi_1^{\mathbf{x}}, \ldots, \Xi_n^{\mathbf{x}}\}$, where for $i = 1, \ldots, n$ the set of points $\Xi_i^{\mathbf{x}}$ may or may not include the state vector $\mathbf{x}_i$ with which it is associated. If the state vector $\mathbf{x}_i$ is included in the set $\Xi_i^{\mathbf{x}}$ then it can be considered as the leader of group i, otherwise $\mathbf{x}_i$ can be regarded as the virtual leader of the individual targets $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_{n_i} \in \Xi_i^{\mathbf{x}}$. The statistical representation is now the random finite set $\mathcal{X} = \bigcup_{\mathbf{x}\in\Upsilon} \mathcal{E}^{\mathbf{x}}$, where $\mathcal{E}^{\mathbf{x}}$ is the RFS associated with group $\mathbf{x}$ and defined solely on the individual target space regardless of its inclusion/exclusion of $\mathbf{x}$. This representation is consistent with definitions given by Daley & Vere-Jones in [70] and Stoyan et. al. in [106] for cluster processes.

Either cluster model is applicable to multiple extended target systems. The state vectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$, which collectively form the parent process (RFS Υ), instead represent the targets, while the state vectors $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_{n_i}$, which collectively form the dependent daughter process (RFS $\mathcal{E}_i$) for each target state $\mathbf{x}_i$, instead represent the feature points of a target's extent.

3.1.3 Bayesian recursion

The objective of group/extended target tracking is to estimate the RFS $\mathcal{X}_k$ with instance $\mathbb{X}_k = \{(\mathbf{x}_{1,k}, \Xi_{1,k}), \ldots, (\mathbf{x}_{n,k}, \Xi_{n,k})\}$ (or $\mathbb{X}_k = \{\Xi_{1,k}^{\mathbf{x}}, \ldots, \Xi_{n,k}^{\mathbf{x}}\}$) at each time-step k given a sequence of measurement sets received up to time k ($Z_{1:k} = Z_1, \ldots, Z_k$). In the Bayesian filtering framework the estimation of RFS $\mathcal{X}_k$ is propagated via the posterior distribution $p_k(\mathbb{X}_k \,|\, Z_{1:k})$.

The posterior distribution at time-step k is either given by the probability

density of RFS $\mathcal{X}_k = \bigcup_{\mathbf{x}\in\Upsilon_k}\{(\mathbf{x}, \mathcal{E}^{\mathbf{x}})\}$, which satisfies

$$\int_{S_x\times S_\xi^n} p_k(\mathbb{X}_k \,|\, Z_{1:k})\,\delta\mathbb{X}_k = \Pr(\Upsilon_k \subseteq S_x, \mathcal{E}_{1,k} \subseteq S_\xi, \ldots, \mathcal{E}_{n,k} \subseteq S_\xi \,|\, Z_{1:k}), \tag{3.1}$$

for $\mathbb{X}_k = \{(\mathbf{x}_{1,k}, \Xi_{1,k}), \ldots, (\mathbf{x}_{n,k}, \Xi_{n,k})\}$, or by the probability density of RFS $\mathcal{X}_k = \bigcup_{\mathbf{x}\in\Upsilon_k}\mathcal{E}^{\mathbf{x}}$, which satisfies

$$\int_{S_\xi^n} p_k(\mathbb{X}_k \,|\, Z_{1:k})\,\delta\mathbb{X}_k = \Pr(\mathcal{E}^{\mathbf{x}}_{1,k} \subseteq S_\xi, \ldots, \mathcal{E}^{\mathbf{x}}_{n,k} \subseteq S_\xi \,|\, Z_{1:k}), \tag{3.2}$$

for $\mathbb{X}_k = \{\Xi^{\mathbf{x}}_{1,k}, \ldots, \Xi^{\mathbf{x}}_{n,k}\}$, and for all $S_x \subseteq \mathcal{C}(\mathbb{R}^{d_x})$, $S_\xi \subseteq \mathcal{C}(\mathbb{R}^{d_\xi})$. The expressions in equations (3.1) and (3.2) are generalisations of that in equation (2.15) such that the integrals are given by

$$\int f(\mathbb{X})\,\delta\mathbb{X} = \sum_{n=0}^{\infty}\frac{1}{n!}\int f(\{(\mathbf{x}_1, \Xi_1), \ldots, (\mathbf{x}_n, \Xi_n)\})\,\mathrm{d}\mathbf{x}_1\cdots\mathrm{d}\mathbf{x}_n\,\delta\Xi_1\cdots\delta\Xi_n, \tag{3.3}$$

$$\int f(\mathbb{X})\,\delta\mathbb{X} = \sum_{n=0}^{\infty}\frac{1}{n!}\int f(\{\Xi^{\mathbf{x}}_1, \ldots, \Xi^{\mathbf{x}}_n\})\,\mathrm{d}\mathbf{x}_1\cdots\mathrm{d}\mathbf{x}_n\,\delta\Xi_1\cdots\delta\Xi_n, \tag{3.4}$$

respectively, where each of the indicated integrals $\int\cdot\,\delta\Xi_i$ are a set integral as given in equation (2.19).

The Bayesian filter recursion for group/extended target tracking is a generalisation of the recursion given in Section 2.3.3 and is formulated as follows. Suppose the multiple group/extended target system has a cluster process representation, written in terms of either RFS model described in Section 3.1.2, and that the prior distribution $p_{k-1}(\mathbb{X}_{k-1} \,|\, Z_{1:k-1})$ is known. Then the multi-group/extended target Bayesian filter recursion is given by the predicted and updated posterior distributions

$$p_{k|k-1}(\mathbb{X}_k \,|\, Z_{1:k-1}) = \int f_{k|k-1}(\mathbb{X}_k \,|\, \mathbb{X}_{k-1})\, p_{k-1}(\mathbb{X}_{k-1} \,|\, Z_{1:k-1})\,\delta\mathbb{X}_{k-1}, \tag{3.5}$$

$$p_k(\mathbb{X}_k \,|\, Z_{1:k}) = \frac{g_k(Z_k \,|\, \mathbb{X}_k)\, p_{k|k-1}(\mathbb{X}_k \,|\, Z_{1:k-1})}{\displaystyle\int g_k(Z_k \,|\, \mathbb{X}_k)\, p_{k|k-1}(\mathbb{X}_k \,|\, Z_{1:k-1})\,\delta\mathbb{X}_k}, \tag{3.6}$$

where $f_{k|k-1}(\mathbb{X}_k \,|\, \mathbb{X}_{k-1})$ denotes the Markov transition density encapsulating the randomness of the cluster model's evolution and $g_k(Z_k \,|\, \mathbb{X}_k)$ denotes the cluster model likelihood encapsulating the randomness of the observations received at time-step k. That is, the Bayesian filter for group/extended target tracking prop-

agates the probability density of the RFS $\mathcal{X}_k$ with instance $\mathbb{X}_k$, given a sequence of measurement sets, as depicted in the following diagram:

$$\cdots \to p_{k-1}(\mathbb{X}_{k-1} \,|\, Z_{1:k-1}) \xrightarrow{\text{prediction}} p_{k|k-1}(\mathbb{X}_k \,|\, Z_{1:k-1}) \xrightarrow{\text{update}} p_k(\mathbb{X}_k \,|\, Z_{1:k}) \to \cdots$$

While the multi-target Bayesian filter is computationally intractable for all but the simplest problems, the practical implementation of its generalisation to group and extended target tracking problems, given by the recursive equations (3.5)–(3.6), is essentially impossible under almost all circumstances. Consequently, suitable approximations are even more crucial for group and extended target tracking problems. This section concludes with a review of possible strategies for practical implementation, based on generalisations of the first-order moment approximation introduced in Section 2.4 to group and extended target tracking problems.

3.1.4 First-order moment approximations

The following strategies are based on generalisations of the PHD, first proposed by Mahler in [39].

The total group target PHD (Mahler [39])
Denoted by $v_k(\mathbf{x}, \Xi \,|\, Z_{1:k})$, this is defined as the first-order moment of the RFS $\mathcal{X} = \bigcup_{\mathbf{x}\in\Upsilon}\{(\mathbf{x}, \mathcal{E}^{\mathbf{x}})\}$ whose probability density is $p_k(\mathbb{X} \,|\, Z_{1:k})$, given the instance $\mathbb{X} = \{(\mathbf{x}_1, \Xi_1), \ldots, (\mathbf{x}_n, \Xi_n)\}$. That is

$$v_k(\mathbf{x}, \Xi \,|\, Z_{1:k}) = \int_{\mathbb{X}\ni(\mathbf{x},\Xi)} p_k(\mathbb{X} \,|\, Z_{1:k})\, \delta\mathbb{X}, \tag{3.7}$$

which operates on the space occupying the single pair $(\mathbf{x}, \Xi)$, where $\mathbf{x}$ denotes the group state whose actual target set is Ξ.

The condensed group target PHD
Denoted by $v_k(\mathbf{x}, \boldsymbol{\xi} \,|\, Z_{1:k})$, this is defined as the first-order moment of the RFS $\mathcal{X} = \bigcup_{\mathbf{x}\in\Upsilon} \mathcal{E}^{\mathbf{x}}$ whose probability density is $p_k(\mathbb{X} \,|\, Z_{1:k})$, given the instance $\mathbb{X} = \{\Xi_1^{\mathbf{x}}, \ldots, \Xi_n^{\mathbf{x}}\}$. That is

$$v_k(\mathbf{x}, \boldsymbol{\xi} \,|\, Z_{1:k}) = \int_{\substack{\mathbb{X}\ni\Xi^{\mathbf{x}}, \\ \Xi^{\mathbf{x}}\ni\boldsymbol{\xi}}} p_k(\mathbb{X} \,|\, Z_{1:k})\, \delta\mathbb{X}, \tag{3.8}$$

which operates on the space occupying the single group and single individual target states jointly.

The marginal group target PHD

Denoted by $v_k(\mathbf{x} \,|\, Z_{1:k})$, this is defined as the marginal density of the group target PHD. That is

$$v_k(\mathbf{x} \,|\, Z_{1:k}) = \int v_k(\mathbf{x}, \Xi \,|\, Z_{1:k}) \, \delta\Xi = \int_{X \ni \mathbf{x}} p_k(X \,|\, Z_{1:k}) \, \delta X, \tag{3.9}$$

where $p_k(X \,|\, Z_{1:k})$ is the probability density of the RFS Υ (parent process) and $v_k(\mathbf{x} \,|\, Z_{1:k})$ is its first-order moment, which operates on the space solely occupying the single group state.

Any recursion that propagates the total group target PHD over time will still be computationally intractable in general. This can be accounted for by the support of the density $v_k(\mathbf{x}, \Xi \,|\, Z_{1:k})$ being the space occupying a single pair $(\mathbf{x}, \Xi)$, which contains the multi-target set Ξ. Essentially the filtering equations of such a recursion will have similar complexity as those for the multi-target Bayesian recursion given in Section 2.3.3. Consequently, it will be recursions that instead propagate the condensed group target PHD (as depicted over time in the diagram below[1]) that provide tractable solutions to the group target tracking problem and such will be the focus of the next section.

$$\begin{array}{ccccccc}
\cdots \to & p_{k-1}(\mathbb{X}_{k-1} \,|\, Z_{1:k-1}) & \xrightarrow{\text{prediction}} & p_{k|k-1}(\mathbb{X}_k \,|\, Z_{1:k-1}) & \xrightarrow{\text{update}} & p_k(\mathbb{X}_k \,|\, Z_{1:k}) & \to \cdots \\
& \downarrow & & \downarrow & & \downarrow & \\
\cdots \to & v_{k-1}(\mathbf{x}_{k-1}, \boldsymbol{\xi}_{k-1}) & \xrightarrow{\text{prediction}} & v_{k|k-1}(\mathbf{x}_k, \boldsymbol{\xi}_k) & \xrightarrow{\text{update}} & v_k(\mathbf{x}_k, \boldsymbol{\xi}_k) & \to \cdots
\end{array}$$

Finally, any recursion that propagates the marginal group target PHD is of interest for estimation problems that only concern the group states and not the individual targets that comprise the groups, which is most applicable to extended target tracking problems. Such a recursion (as depicted over time in the diagram below) will therefore also be considered in the next section for an extended target tracking application.

$$\begin{array}{ccccccc}
\cdots \to & p_{k-1}(\mathbb{X}_{k-1} \,|\, Z_{1:k-1}) & \xrightarrow{\text{prediction}} & p_{k|k-1}(\mathbb{X}_k \,|\, Z_{1:k-1}) & \xrightarrow{\text{update}} & p_k(\mathbb{X}_k \,|\, Z_{1:k}) & \to \cdots \\
& \downarrow & & \downarrow & & \downarrow & \\
\cdots \to & v_{k-1}(\mathbf{x}_{k-1} \,|\, Z_{1:k-1}) & \xrightarrow{\text{prediction}} & v_{k|k-1}(\mathbf{x}_k \,|\, Z_{1:k-1}) & \xrightarrow{\text{update}} & v_k(\mathbf{x}_k \,|\, Z_{1:k}) & \to \cdots
\end{array}$$

[1]Denoting the predicted and updated intensities by $v_{k|k-1}(\mathbf{x}_k, \boldsymbol{\xi}_k) = v_{k|k-1}(\mathbf{x}_k, \boldsymbol{\xi}_k \,|\, Z_{1:k-1})$ and $v_k(\mathbf{x}_k \, \boldsymbol{\xi}_k) = v_k(\mathbf{x}_k \, \boldsymbol{\xi}_k \,|\, Z_{1:k})$ respectively.

3.2 PHD Filtering for Group and Extended Targets

This section presents the filtering equations of the PHD recursions for specific group and extended target tracking applications, the derivations of which are given in Appendix B. The results for the filtering equations in Sections 3.2.3 and 3.2.6 are based on similar results in the author's conference papers [111] and [112] respectively[2], while Sections 3.2.4 and 3.2.5 present special cases of the filter equations in Section 3.2.3.

3.2.1 Dynamic model

The hierarchical RFS representation $\mathcal{X}_{k-1} = \bigcup_{\mathbf{x}\in\Upsilon_{k-1}} \mathcal{E}^{\mathbf{x}}_{k-1}$ allows for the incorporation of group and individual target birth and death, which is encapsulated in the Markov transition density $f_{k|k-1}(\mathbb{X}_k \,|\, \mathbb{X}_{k-1})$ as follows.

For a given parent RFS Υ_{k-1} at time-step $k-1$ with instance X_{k-1}, the survival of each cluster, i.e. each subsidiary RFS $\mathcal{E}^{\mathbf{x}}_{k-1}$, associated with $\mathbf{x}_{k-1} \in X_{k-1}$ whose instance is $\Xi^{\mathbf{x}}_{k-1}$ is modelled by the RFS

$$\Phi_{k|k-1}(\Xi^{\mathbf{x}}_{k-1}) = \begin{cases} \Xi^{\mathbf{x}}_{k} & \text{when } \mathbf{x}_{k-1} \text{ survives,} \\ \emptyset & \text{otherwise.} \end{cases} \tag{3.10}$$

If a group survives, the time evolution of its corresponding multi-target set $\Xi^{\mathbf{x}}_{k-1}$ is modelled by the RFS $\Xi^{\mathbf{x}}_{k}$ given by

$$\Xi^{\mathbf{x}}_{k} = \left(\bigcup\nolimits_{\boldsymbol{\xi}\in\Xi^{\mathbf{x}}_{k-1}} \Psi_{k|k-1}(\boldsymbol{\xi}) \right) \cup \Gamma^{\mathbf{x}}_{\xi,k}, \tag{3.11}$$

where $\Psi_{k|k-1}$ is the same RFS as given in equation (2.24), that now models individual target survival, and $\Gamma^{\mathbf{x}}_{\xi,k}$ is the RFS that models the appearance of new individual targets. New clusters (groups) appear at time k due to spontaneous births, which are modelled by the hierarchical RFS Γ_k. The time evolution of the cluster model $\mathcal{X}_{k-1}$ is then given by

$$\mathbb{X}_k = \left(\bigcup\nolimits_{\mathbf{x}\in\Upsilon_{k-1}} \Phi_{k|k-1}(\Xi^{\mathbf{x}}_{k-1}) \right) \cup \Gamma_k. \tag{3.12}$$

[2]These papers primarily presented results for the predicted and updated intensities in first-order moment filters for group and extended target tracking, each regarded as an extension of the multi-target CPHD filter. In [111], the main updated intensity result assumed a simplistic group tracking scenario in which there were no false alarms and no missed detections, while in [112], the considered extended target scenarios also included the case in which false alarms and missed detections were present.

The RFSs constituting the unions in equations (3.11) and (3.12) are assumed to be statistically independent of each other.

3.2.2 Measurement model

The RFS measurement model accounts for group and individual target detection uncertainty, as well as false measurements, and is specified as follows. For a given parent RFS Υ_k with instance X_k in the cluster model $\mathcal{X}_k = \bigcup_{\mathbf{x}\in\Upsilon_k} \mathcal{E}_k^{\mathbf{x}}$ at time-step k, the detection of each cluster, i.e. each subsidiary RFS $\mathcal{E}_k^{\mathbf{x}}$ with instance $\Xi_k^{\mathbf{x}}$, associated with $\mathbf{x}_k \in X_k$ is modelled by the RFS

$$\Theta_{1,k}(\Xi_k^{\mathbf{x}}) = \begin{cases} Z_{\mathbf{x},k} & \text{when } \mathbf{x}_k \text{ is detected,} \\ \emptyset & \text{otherwise.} \end{cases} \tag{3.13}$$

The collection of measurement generated from a detected group is modelled by the RFS $Z_{\mathbf{x},k}$ given by

$$Z_{\mathbf{x},k} = \bigcup_{\boldsymbol{\xi}\in\Xi_k^{\mathbf{x}}} \Theta_{2,k}(\boldsymbol{\xi} \,|\, \mathbf{x}_k), \tag{3.14}$$

where the RFS $\Theta_{2,k}$ models the detection of each individual target $\boldsymbol{\xi} \in \Xi_k^{\mathbf{x}}$ and is given by

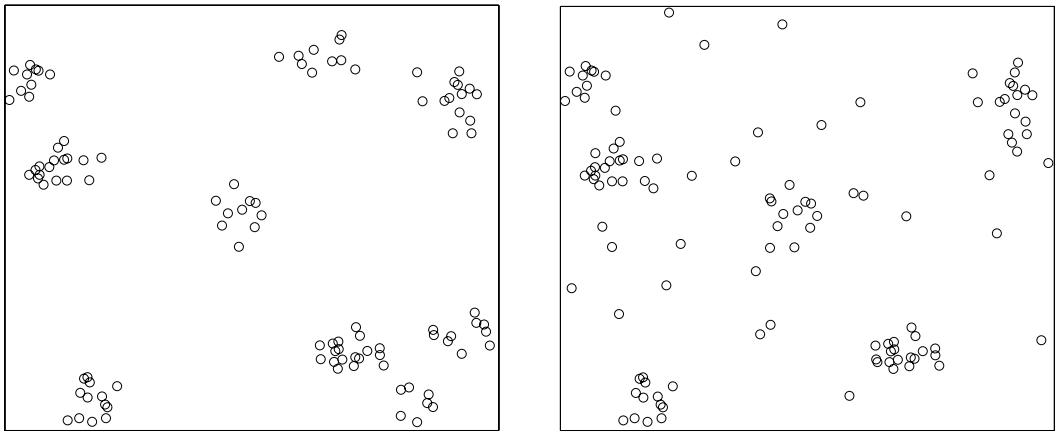
$$\Theta_{2,k}(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k) = \begin{cases} \{\mathbf{z}_k\} & \text{when an individual target } \boldsymbol{\xi}_k \text{ in group } \mathbf{x}_k \text{ is detected,} \\ \emptyset & \text{otherwise.} \end{cases} \tag{3.15}$$

Those measurements that are not target generated are considered to be false alarms, or clutter, and modelled by the RFS K_k. The total collection of measurements, modelled by the RFS Z_k, is then given by

$$Z_k = \left(\bigcup\nolimits_{\mathbf{x}\in X_k} \Theta_{1,k}(\Xi_k^{\mathbf{x}})\right) \cup K_k. \tag{3.16}$$

The clutter and target generated measurements are assumed to be independent and the randomness of the observations is captured in the likelihood $g_k(Z_k \,|\, \mathbb{X}_k)$.

The cluster model for RFS $\mathcal{X}_k$ results in target generated measurements that naturally appear in clusters. It is convenient then to also apply cluster modelling to the clutter. A more general measurement model can be obtained by assuming a cluster model for RFS K_k. To illustrate, consider the two realisations of



(a) Example 1 (b) Example 2

Figure 3.3: Realisations of measurement sets

the measurement set Z_k in Figure 3.3. Example 1 (Figure 3.3a) clearly shows a measurement set that consists of a collection of measurement clusters, which is representative of a measurement model that includes cluster clutter. Example 2 (Figure 3.3b) shows a measurement set where the clutter appears not to be distributed in clusters. However each false measurement in the second example can simply be considered as a cluster that consists of only one component, so that the cluster model for RFS K_k is still valid.

3.2.3 A PHD filter for group targets

The propagation of the predicted intensity in the condensed group target PHD recursion, presented in Theorem 3.1, requires the following assumptions regarding the group and individual target dynamics:

- each group evolves independently of one another and independently of group births;

- the RFS $\Phi_{k|k-1}$ modelling the survival uncertainty for each group evolving from state $\mathbf{x}_{k-1}$ to $\mathbf{x}_k$ is Bernoulli with parameter $p_{S,1}(\mathbf{x}_{k-1})$ and transition density $f_{k|k-1}(\mathbf{x}_k \,|\, \mathbf{x}_{k-1})$ describing its evolution;

- each individual target, though conditioned on a group, evolves independently of other individual targets in that group and independently of births;

- the RFS $\Psi_{k|k-1}$ modelling the survival uncertainty of each individual target evolving from state $\boldsymbol{\xi}_{k-1}$ to $\boldsymbol{\xi}_k$ is also Bernoulli with parameter $p_{S,2}(\boldsymbol{\xi}_{k-1})$ and

transition density $f_{k|k-1}(\boldsymbol{\xi}_k \,|\, \boldsymbol{\xi}_{k-1}, \mathbf{x}_{k-1})$ describing its evolution, which is also dependent on a group state $\mathbf{x}_{k-1}$.

Theorem 3.1 (**Multi-group PHD prediction**). *Given the listed assumptions and a prior intensity which satisfies the following factorisation* $v_{k-1}(\mathbf{x}_{k-1}, \boldsymbol{\xi}_{k-1}) = v_{k-1}(\mathbf{x}_{k-1}) \times v_{k-1}(\boldsymbol{\xi}_{k-1} \,|\, \mathbf{x}_{k-1})$*, then the predicted intensity is*

$$v_{k|k-1}(\mathbf{x}_k, \boldsymbol{\xi}_k) = \gamma_k(\mathbf{x}_k, \boldsymbol{\xi}_k) + \int p_{S,1}(\mathbf{x}_{k-1})\, f_{k|k-1}(\mathbf{x}_k \,|\, \mathbf{x}_{k-1})\, v_{k|k-1}(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k, \mathbf{x}_{k-1})\, v_{k-1}(\mathbf{x}_{k-1})\, \mathrm{d}\mathbf{x}_{k-1}, \tag{3.17}$$

where the conditional predicted intensity of each RFS $\mathcal{E}_k$ *is*

$$v_{k|k-1}(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k, \mathbf{x}_{k-1}) = \gamma_{\xi,k}(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k) + \int p_{S,2}(\boldsymbol{\xi}_{k-1})\, f_{k|k-1}(\boldsymbol{\xi}_k \,|\, \boldsymbol{\xi}_{k-1}, \mathbf{x}_{k-1})\, v_{k-1}(\boldsymbol{\xi}_{k-1} \,|\, \mathbf{x}_{k-1})\, \mathrm{d}\boldsymbol{\xi}_{k-1}. \tag{3.18}$$

Note that $\gamma_k(\mathbf{x}_k, \boldsymbol{\xi}_k)$ *denotes the intensity of the RFS* Γ_k *and* $\gamma_{\xi,k}(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k)$ *denotes the conditional intensity of the RFS* $\Gamma_{\xi,k}$*.*

Proof. Details on the derivation of this result are given in Appendix B (Section B.3.1). □

To complete the condensed group target PHD recursion, the propagation of the updated intensity, presented in Theorem 3.2, requires the following assumptions:

- each group generates a set of measurements independently of one another, and independently of clutter;
- the RFS $\Theta_{1,k}$ modelling the detection uncertainty for each group $\mathbf{x}_k \in X_k$ is Bernoulli with parameter $p_{D,1}(\mathbf{x}_k)$ and a likelihood that the set $Z = Z_{\mathbf{x},k}$ is observed from $\mathbf{x}_k$ given by $g_k(Z \,|\, \mathbf{x}_k, \Xi_k)$;
- each individual target in a group generates measurements independently of one another;
- the RFS $\Theta_{2,k}$ modelling the detection uncertainty for each individual target $\boldsymbol{\xi}_k \in \Xi_k$, where Ξ_k is the multi-target set corresponding to a group $\mathbf{x}_k$, is Bernoulli with parameter $p_{D,2}(\boldsymbol{\xi}_k)$ and a likelihood that $\mathbf{z}_k$ is observed from $\boldsymbol{\xi}_k$ in group $\mathbf{x}_k$ given by $g_k(\mathbf{z}_k \,|\, \boldsymbol{\xi}_k, \mathbf{x}_k)$;

- the RFS Υ_k and each RFS $\mathcal{E}_k$ in the hierarchical RFS representation of the predicted multi-group state are both approximated with a Poisson RFS.

The final assumption concerns the clutter in the measurement model. As alluded to in Section 3.2.2, the clutter is modelled so that they appear in clusters. Specifically, the clutter RFS K_k is modelled with the same hierarchical RFS representation as for the predicted multi-group state. The probability density of the clutter RFS is then given by

$$p_\kappa(K_k) = \exp\left(-\int \kappa_1(\mathbf{u}_{\varphi_\kappa})\,\mathrm{d}\mathbf{u}_{\varphi_\kappa}\right) \sum_{\pi_\kappa \in \Pi_{K_k}} \left(\prod_{\varphi_\kappa \in \pi_\kappa} \Lambda(\varphi_\kappa)\right), \tag{3.19}$$

where the following notation is introduced

$$\Lambda(\varphi_\kappa) = \int \exp\left(-\int \kappa_2(\mathbf{z} \,|\, \mathbf{u}_{\varphi_\kappa})\,\mathrm{d}\mathbf{z}\right) \left(\prod_{\mathbf{z} \in \varphi_\kappa} \kappa_2(\mathbf{z} \,|\, \mathbf{u}_{\varphi_\kappa})\right) \kappa_1(\mathbf{u}_{\varphi_\kappa})\,\mathrm{d}\mathbf{u}_{\varphi_\kappa}. \tag{3.20}$$

In (3.19), Π_{K_k} denotes the set of all partitions of the RFS K_k and π_κ denotes a single partition containing subsets $\varphi_\kappa \in \pi_\kappa$ such that $\bigcup_{\varphi_\kappa \in \pi_\kappa} \varphi_\kappa = K_k$. The underlying Poisson RFS consisting of cluster centres $\mathbf{u}_{\varphi_\kappa}$ for each subset φ_κ in a possible partition π_κ is characterised by the intensity $\kappa_1(\mathbf{u}_{\varphi_\kappa})$, and $\kappa_2(\mathbf{z} \,|\, \mathbf{u}_{\varphi_\kappa})$ denotes the intensity of the corresponding dependent Poisson RFS. The probability density given by equations (3.19)–(3.20), discussed further in Appendix B (Section B.2.2), closely resembles the probability density for the example of a cluster process presented by van Lieshout in [113], where the parent and daughter processes are also Poisson.

Theorem 3.2 (**Multi-group PHD update**). *Suppose a new set of measurements, denoted by $Z = Z_k$, is received at time-step k then, given the preceding assumptions hold, the updated intensity is*

$$\begin{aligned} v_k(\mathbf{x}_k, \boldsymbol{\xi}_k) = v_{k|k-1}(\mathbf{x}_k, \boldsymbol{\xi}_k) \Big\{ & 1 - p_{D,1}(\mathbf{x}_k) \\ & + p_{D,1}(\mathbf{x}_k) \Big(\exp\left(-v_{k|k-1}[p_{D,2} \,|\, \mathbf{x}_k]\right) \left(1 - p_{D,2}(\boldsymbol{\xi}_k)\right) \Big) \Big\} \\ & + \sum_{\pi \in \Pi_Z} \varpi_\pi \sum_{\varphi \in \pi} \frac{p_{D,1}(\mathbf{x}_k)\, L_\varphi(\mathbf{x}_k)\, v_{k|k-1}(\mathbf{x}_k)}{v_{k|k-1}[p_{D,1}\, L_\varphi] + \Lambda(\varphi)} v_k(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k)_\varphi \,, \end{aligned} \tag{3.21}$$

where the conditional updated intensity for each possible RFS $\mathcal{E}_k$ coinciding with

subset $\varphi \in \pi$ *is*

$$v_k(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k)_\varphi = \left(1 - p_{D,2}(\boldsymbol{\xi}_k) + \sum_{\mathbf{z}\in\varphi} \frac{p_{D,2}(\boldsymbol{\xi}_k)\, g_k(\mathbf{z} \,|\, \boldsymbol{\xi}_k, \mathbf{x}_k)}{v_{k|k-1}[p_{D,2}\, l_{\mathbf{z}} \,|\, \mathbf{x}_k]}\right) v_{k|k-1}(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k), \quad (3.22)$$

denoting $l_{\mathbf{z}} = g_k(\mathbf{z} \,|\, \boldsymbol{\xi}_k, \mathbf{x}_k)$. *The summation* $\sum_{\pi\in\Pi_Z}$ *is taken over all possible partitions of the entire measurement set* Z, *denoted by* Π_Z. *A single partition* π *contains subsets* $\varphi \in \pi$ *such that* $\bigcup_{\varphi\in\pi} \varphi = Z$ *The partition weights are given by*

$$\varpi_\pi = \frac{\prod_{\varphi\in\pi}\Big(v_{k|k-1}[p_{D,1}\, L_\varphi] + \Lambda(\varphi)\Big)}{\sum_{\pi'\in\Pi_Z}\prod_{\varphi'\in\pi'}\Big(v_{k|k-1}[p_{D,1}\, L_{\varphi'}] + \Lambda(\varphi')\Big)}, \quad (3.23)$$

and $L_\varphi = L_\varphi(\mathbf{x}_k)$ *denotes the pseudo multiple measurement likelihood for the subset* φ *given by*

$$L_\varphi(\mathbf{x}_k) = \exp\left(-v_{k|k-1}[p_{D,2} \,|\, \mathbf{x}_k]\right) \prod_{\mathbf{z}\in\varphi} v_{k|k-1}[p_{D,2}\, l_{\mathbf{z}} \,|\, \mathbf{x}_k]. \quad (3.24)$$

The clutter term $\Lambda(\varphi)$ *is as defined in equation* (3.20).

Proof. Details on the derivation of this result are given in Appendix B (Section B.3.2). □

Theorems 3.1 and 3.2 are, respectively, the prediction and update step in the condensed group target PHD recursion, which will be referred to hereafter as the *multi-group PHD filter* recursion. The result for the multi-group PHD prediction can simply be interpreted as a doubly-stochastic version of the multi-target PHD prediction, consisting of predicted components for group and individual target states from equations (3.17) and (3.18) conditionally. The result for the multi-group PHD update, however, is more complex. The update equation (3.21) consists of one missed group detection term and $|\Pi_{Z_k}| \times |\pi|$ group detection terms. Each group detection term relates to a subset φ in a possible partition π of the measurement set Z_k, in which a group state is updated with the pseudo multiple measurement likelihood L_φ. A group detection term also includes the update equation for an individual target state, given by equation (3.22), which is similar to the multi-target PHD update and consists of one missed detection term as well as $|\varphi|$ detection terms (relating to the number of detections in the subset φ).

3.2.4 Special case 1: single group target

A single group target is a special case of the multiple group target system and the corresponding state representation is such that the underlying RFS Υ strictly contains a single group state. That is, the complete state specification is given by $\mathbb{X} = \{(\mathbf{x}, \Xi)\}$ with group state vector $\mathbf{x}$ and finite set Ξ of individual target state vectors $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_n$.

The single group RFS representation is equivalent to fixing the group cardinality at $|X| = 1$ in the complete state specification for the multiple group target system. As a consequence of this restriction on the group cardinality, the dynamic and measurement model specifications for a single group target system are simplified.

Firstly, the group cardinality restriction does not allow group births or death in the dynamic model. That is, there is no survival uncertainty for the cluster associated with the group state $\mathbf{x}_{k-1}$ so that $\Phi_{k|k-1}(\Xi^{\mathbf{x}}_{k-1}) = \{\Xi^{\mathbf{x}}_k\}$ with probability $f_{k|k-1}(\mathbf{x}_k \,|\, \mathbf{x}_{k-1})$. The time evolution of the complete state $\mathbb{X}_{k-1}$ for the single group target scenario is therefore given by $\mathbb{X}_k = \{\Xi^{\mathbf{x}}_k\}$ where $\Xi^{\mathbf{x}}_k$ is the RFS that models the time evolution of the multi-target set Ξ_{k-1} conditioned on group state $\mathbf{x}$, which is given by the same union of RFSs as shown in equation (3.11).

Likewise for the measurement model, it is assumed there is no detection uncertainty for the cluster associated with the group state $\mathbf{x}_k$ so that $\Theta_{1,k}(\Xi^{\mathbf{x}}_k) = Z_{\mathbf{x},k}$ with likelihood $g_k(Z \,|\, \mathbf{x}_k, \Xi_k)$, where $Z = Z_{\mathbf{x},k}$ still denotes the collection of measurements generated by the cluster $\Xi^{\mathbf{x}}_k$ given by equation (3.14), and the total collection of measurements, modelled by the RFS Z_k, becomes

$$Z_k = \Theta_{1,k}(\Xi^{\mathbf{x}}_k) \cup K_k. \tag{3.25}$$

Finally, it was natural to assume, for the multiple group target scenario, that the measurements appear in clusters and either a cluster of measurements are target generated or otherwise clutter. This assumption is not necessarily appropriate for a single group scenario and instead the clutter RFS K_k is modelled with a Poisson RFS characterised its intensity $\kappa(\cdot)$.

The remaining assumptions required for the propagation of the filtering equations in the condensed group target PHD recursion for this single group special case, presented in Corollary 3.1 and Theorem 3.3, are as follows:

- each individual target, though conditional on the single group, evolves independently of other individual targets and independently of births;

- the RFS $\Psi_{k|k-1}$ modelling the survival uncertainty of each individual target

evolving from state $\boldsymbol{\xi}_{k-1}$ to $\boldsymbol{\xi}_k$ is Bernoulli with parameter $p_S(\boldsymbol{\xi}_{k-1})$ and transition density $f_{k|k-1}(\boldsymbol{\xi}_k \,|\, \boldsymbol{\xi}_{k-1}, \mathbf{x}_{k-1})$, which is again also dependent on group state $\mathbf{x}_{k-1}$;

- each individual target within the single group generates measurements independently of one another and independently of clutter;

- the RFS $\Theta_{2,k}$ modelling the detection uncertainty of each individual target $\boldsymbol{\xi}_k \in \Xi_k^{\mathbf{x}}$ is Bernoulli with parameter $p_D(\boldsymbol{\xi}_k)$ and and a likelihood that $\mathbf{z}_k$ is observed from $\boldsymbol{\xi}_k$ given by $g_k(\mathbf{z}_k \,|\, \boldsymbol{\xi}_k, \mathbf{x}_k)$;

- the RFS $\mathcal{E}_k$ in the RFS representation of the predicted single group state is approximated with a Poisson RFS.

Corollary 3.1 (**Single group PHD prediction**). *Given the preceding assumptions regarding the group and individual target dynamics and a prior intensity which satisfies the factorisation* $v_{k-1}(\mathbf{x}_{k-1}, \boldsymbol{\xi}_{k-1}) = p_{k-1}(\mathbf{x}_{k-1}) \times v_{k-1}(\boldsymbol{\xi}_{k-1} \,|\, \mathbf{x}_{k-1})$*, then the predicted intensity for the single group special case follows immediately from equation* (3.17)*, though a detailed derivation is given in Appendix B (Section B.4.1). That is*

$$v_{k|k-1}(\mathbf{x}_k, \boldsymbol{\xi}_k) = \int f_{k|k-1}(\mathbf{x}_k \,|\, \mathbf{x}_{k-1})\, v_{k|k-1}(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k, \mathbf{x}_{k-1})\, p_{k-1}(\mathbf{x}_{k-1})\, \mathrm{d}\mathbf{x}_{k-1}, \quad (3.26)$$

and $v_{k|k-1}(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k, \mathbf{x}_{k-1})$ *is the conditional predicted intensity of the RFS* $\mathcal{E}_k$ *as given in equation* (3.18)*, where the probability of individual target survival is instead denoted by* $p_S(\boldsymbol{\xi}_{k-1})$*.*

Theorem 3.3 (**Single group PHD update**). *Suppose a new set of measurements* Z_k *is received at time-step* k *then, given the assumptions regarding the observations and the RFS representation of the predicted state hold, the updated intensity for the single group special case is*

$$v_k(\mathbf{x}_k, \boldsymbol{\xi}_k) = \frac{L_{Z_k}(\mathbf{x}_k)\, p_{k|k-1}(\mathbf{x}_k)}{\int L_{Z_k}(\mathbf{x}_k)\, p_{k|k-1}(\mathbf{x}_k)\, \mathrm{d}\mathbf{x}_k} v_k(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k), \quad (3.27)$$

where $L_{Z_k}(\mathbf{x}_k)$ *is a pseudo multiple measurement likelihood given by*

$$L_{Z_k}(\mathbf{x}_k) = \exp\left(-v_{k|k-1}[p_D \,|\, \mathbf{x}_k]\right) \prod_{\mathbf{z} \in Z_k} \Big(\kappa(\mathbf{z}) + v_{k|k-1}[p_D\, l_{\mathbf{z}} \,|\, \mathbf{x}_k]\Big), \quad (3.28)$$

again denoting $l_{\mathbf{z}} = g_k(\mathbf{z} \,|\, \boldsymbol{\xi}_k, \mathbf{x}_k)$. *The conditional updated intensity for the RFS* $\mathcal{E}_k$ *is given by*

$$v_k(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k) = \left(1 - p_D(\boldsymbol{\xi}_k) + \sum_{\mathbf{z} \in Z_k} \frac{p_D(\boldsymbol{\xi}_k)\, g_k(\mathbf{z} \,|\, \boldsymbol{\xi}_k, \mathbf{x}_k)}{\kappa(\mathbf{z}) + v_{k|k-1}[p_D\, l_{\mathbf{z}} \,|\, \mathbf{x}_k]}\right) v_{k|k-1}(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k). \quad (3.29)$$

Proof. Details on the derivation are given in Appendix B (Section B.4.2). □

3.2.5 Special case 2: single cluster existence model

The scenario considered now is a generalisation of the joint target-detection and tracking (JoTT) filter [69, Section 14.7, pp. 514–528] to group targets, which describes an interesting problem of detecting the existence of a collection of targets evolving over time in a heavily cluttered environment. Referred to hereafter as the *single cluster existence model*, its corresponding state representation is such that the underlying RFS Υ either contains a single group state or is empty. Consequently, the dynamic model specification for the single cluster existence model differs again from those given in Sections 3.2.1 and 3.2.4.

Since the group cardinality restriction described in Section 3.2.4 is relaxed to also include $|X| = 0$, the dynamic model can incorporate group survival uncertainty and birth which are modelled by the RFS $\mathbb{X}_k = \overline{\Phi}_{k|k-1}$ where for a given hierarchical RFS representation $\mathbb{X}_{k-1}$ at time-step $k-1$

$$\overline{\Phi}_{k|k-1}(\mathbb{X}_{k-1}) = \begin{cases} \Phi_{k|k-1}(\Xi^{\mathbf{x}}_{k-1}) & \text{when a group exists at time } k-1, \\ \Gamma_k & \text{when a group does not exist at time } k-1. \end{cases} \quad (3.30)$$

That is, when $\mathbb{X}_{k-1} = \{\Xi^{\mathbf{x}}_{k-1}\}$, $\overline{\Phi}_{k|k-1}(\mathbb{X}_{k-1}) = \Phi_{k|k-1}(\Xi^{\mathbf{x}}_{k-1})$ where $\Phi_{k|k-1}$ is the RFS modelling the survival uncertainty of the cluster associated with the group state $\mathbf{x}_{k-1}$ as given by equation (3.10). When $\mathbb{X}_{k-1} = \emptyset$, $\overline{\Phi}_{k|k-1}(\mathbb{X}_{k-1}) = \Gamma_k$ where Γ_k is the RFS modelling the group birth given by

$$\Gamma_k = \begin{cases} \{\Xi^{\mathbf{x}}_k\} & \text{when a new group appears at time } k, \\ \emptyset & \text{when no new group appears at time } k. \end{cases} \quad (3.31)$$

The remaining assumptions required for the propagation of the prediction equation in the condensed group target PHD recursion for the single cluster existence model, presented in Theorem 3.4, are as follows:

- the RFS $\overline{\Phi}_{k|k-1}$ is modelled with a Bernoulli RFS with parameters ρ_{k-1};

- the RFS $\Phi_{k|k-1}$ is modelled with a Bernoulli RFS with parameter $p_{S,1}(\mathbf{x}_{k-1})$ and transition density $f_{k|k-1}(\mathbf{x}_k \,|\, \mathbf{x}_{k-1})$;

- each individual target within the single group, if it exists at time $k-1$ and survives to the next time-step, evolves independently of one another and independently of births;

- the RFS $\Psi_{k|k-1}$ is modelled with a Bernoulli RFS with parameter $p_{S,1}(\boldsymbol{\xi}_{k-1})$ and transition density $f_{k|k-1}(\boldsymbol{\xi}_k \,|\, \boldsymbol{\xi}_{k-1}, \mathbf{x}_{k-1})$;

- the RFS Γ_k is a cluster model whose parent process is approximated with a Bernoulli RFS with parameter p_B;

Theorem 3.4 (**Single cluster existence prediction**). *Given the listed assumptions and a prior intensity which satisfies the factorisation* $v_{k-1}(\mathbf{x}_{k-1}, \boldsymbol{\xi}_{k-1}) = p_{k-1}(\mathbf{x}_{k-1}) \times v_{k-1}(\boldsymbol{\xi}_{k-1} \,|\, \mathbf{x}_{k-1})$*, then the predicted intensity for the single cluster existence model is*

$$\begin{aligned} v_{k|k-1}(\mathbf{x}_k, \boldsymbol{\xi}_k) &= (1 - \rho_{k-1})\, \gamma_k(\mathbf{x}_k, \boldsymbol{\xi}_k) \\ &\quad + \rho_{k-1} \int p_{S,1}(\mathbf{x}_{k-1})\, f_{k|k-1}(\mathbf{x}_k \,|\, \mathbf{x}_{k-1})\, v_{k|k-1}(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k, \mathbf{x}_{k-1})\, p_{k-1}(\mathbf{x}_{k-1})\, \mathrm{d}\mathbf{x}_{k-1}, \end{aligned} \tag{3.32}$$

where $\gamma_k(\mathbf{x}_k, \boldsymbol{\xi}_k) = p_B\, p_{\gamma,k}(\mathbf{x}_k)\, \gamma_{\xi,k}(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k)$ *is the intensity of the group birth RFS* Γ_k *and* $v_{k|k-1}(\boldsymbol{\xi}_k \,|\mathbf{x}_k, \mathbf{x}_{k-1})$ *is the conditional predicted intensity of the RFS* $\mathcal{E}_k$*, which has precisely the same expression as given in equation* (3.18)*.*

The predicted probability of existence $\rho_{k|k-1}$ *is also propagated and is given by*

$$\rho_{k|k-1} = (1 - \rho_{k-1})\, p_B + \rho_{k-1} \int p_{S,1}(\mathbf{x}_{k-1})\, p_{k-1}(\mathbf{x}_{k-1})\, \mathrm{d}\mathbf{x}_{k-1}. \tag{3.33}$$

Proof. Details on the derivation are given in Appendix B (Section B.5.1). □

To complete the recursion, note that the measurement model specification for the single cluster existence model also differs from those given in Sections 3.2.2 and 3.2.4. The collection of measurements is modelled by the RFS Z_k given by

$$Z_k = \begin{cases} \widetilde{Z}_k & \text{when a group exists at time } k, \\ K_k & \text{when a group does not exist at time } k, \end{cases} \tag{3.34}$$

where $\widetilde{Z}_k$ is given by the same union of RFSs as shown in equation (3.25) and K_k is the clutter RFS.

The remaining assumptions required for the propagation of the update equation in the condensed group target PHD recursion for the single cluster existence model, presented in Theorem 3.5, are as follows:

- the RFS $\Theta_{1,k}$, constituting the union in the RFS $\widetilde{Z}_k$ (along with RFS K_k) and modelling the detection uncertainty of a single group $\mathbf{x}_k$, is Bernoulli with parameter $p_{D,1}(\mathbf{x}_k)$ and a likelihood that the set $Z = Z_{\mathbf{x},k}$ is observed from $\mathbf{x}_k$ given by $g_k(Z \,|\, \mathbf{x}_k, \Xi_k)$;

- each individual target within the single group, if it exists at time k, generates measurements independently of one another;

- the RFS $\Theta_{2,k}$ modelling the detection uncertainty for each individual target $\boldsymbol{\xi}_k \in \Xi_k$, where Ξ_k denotes the multi-target set corresponding to the single group $\mathbf{x}_k$, is Bernoulli with parameter $p_{D,2}(\boldsymbol{\xi}_k)$ and a likelihood that $\mathbf{z}_k$ is observed from $\boldsymbol{\xi}_k$ in the group $\mathbf{x}_k$ given by $g_k(\mathbf{z}_k \,|\, \boldsymbol{\xi}_k, \mathbf{x}_k)$;

- the RFS K_k is modelled with a Poisson RFS characterised by intensity κ;

- the RFS Υ_k in the RFS representation of the predicted state for the single cluster existence model is approximated by a Bernoulli RFS with parameter $\rho_{k|k-1}$ and the subsidiary RFS $\mathcal{E}_k$ is approximated with a Poisson RFS. Likewise for the RFS representation at time $k-1$ with Bernoulli parameter ρ_{k-1}.

Theorem 3.5 (**Single cluster existence update**). *Suppose a new set of measurements Z_k is received at time-step k then, given the listed assumptions hold, the updated intensity for the single cluster existence model is*

$$v_k(\mathbf{x}_k, \boldsymbol{\xi}_k) = \frac{\rho_{k|k-1} \left\{ (1 - p_{D,1}(\mathbf{x}_k))\, v_{k|k-1}(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k) + p_{D,1}(\mathbf{x}_k)\, \tilde{L}_{Z_k}(\mathbf{x}_k)\, v_k(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k) \right\} p_{k|k-1}(\mathbf{x}_k)}{1 - \rho_{k|k-1} + \rho_{k|k-1} \int \left(1 - p_{D,1}(\mathbf{x}_k) + p_{D,1}(\mathbf{x}_k)\, \tilde{L}_{Z_k}(\mathbf{x}_k) \right) p_{k|k-1}(\mathbf{x}_k)\, \mathrm{d}\mathbf{x}_k}, \tag{3.35}$$

where $v_k(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k)$ is given by equation (3.29) *and, given non-empty subsets $W \subseteq Z_k$,*

$$\tilde{L}_{Z_k}(\mathbf{x}_k) = \exp\left(-v_{k|k-1}[p_{D,2} \,|\, \mathbf{x}_k]\right) \sum_{W \subseteq Z_k} \prod_{\mathbf{z} \in W} \frac{v_{k|k-1}[p_{D,2}\, l_{\mathbf{z}} \,|\, \mathbf{x}_k]}{\kappa(\mathbf{z})}, \tag{3.36}$$

is the pseudo multiple measurement likelihood.

The updated probability of existence is also propagated and is given by

$$\rho_k = \frac{\rho_{k|k-1} \int \left(1 - p_{D,1}(\mathbf{x}_k) + p_{D,1}(\mathbf{x}_k)\, \tilde{L}_{Z_k}(\mathbf{x}_k)\right) p_{k|k-1}(\mathbf{x}_k)\, \mathrm{d}\mathbf{x}_k}{1 - \rho_{k|k-1} + \rho_{k|k-1} \int \left(1 - p_{D,1}(\mathbf{x}_k) + p_{D,1}(\mathbf{x}_k) \tilde{L}_{Z_k}(\mathbf{x}_k)\right) p_{k|k-1}(\mathbf{x}_k)\, \mathrm{d}\mathbf{x}_k}. \tag{3.37}$$

Proof. Details on the derivation are given in Appendix B (Section B.5.2). □

3.2.6 A PHD filter for extended targets

As suggested in Section 3.1.4, the marginal group target PHD recursion is now considered for the multiple extended target tracking problem given the hierarchical RFS representation described in Section 3.1.2. The propagation of the filtering equations in the recursion, presented in Theorems 3.6 and 3.7 respectively, require the following assumptions:

- each target evolves independently of one another and independently of target births;
- the RFS $\Phi_{k|k-1}$, now modelling the survival uncertainty for each target evolving from state $\mathbf{x}_{k-1}$ to $\mathbf{x}_k$, is Bernoulli with parameter $p_S(\mathbf{x}_{k-1})$ and transition density $f_{k|k-1}(\mathbf{x}_k \,|\, \mathbf{x}_{k-1})$;
- each target generates a set of measurements independently of one another and independently of clutter;
- the RFS $\Theta_{1,k}$ now modelling the detection uncertainty for each extended target $\mathbf{x}_k \in X_k$ is Bernoulli with parameter $p_{D,1}(\mathbf{x}_k)$ and a likelihood that the set $Z = Z_{\mathbf{x},k}$ is observed from the collection of feature points associated with $\mathbf{x}_k$ given by $g_k(Z \,|\, \mathbf{x}_k, \Xi_k)$;
- each feature point, on the extents of independently observed targets, generates measurements independently of one another, and independently of clutter;
- the RFS $\Theta_{2,k}$ now modelling the detection uncertainty of each feature point $\boldsymbol{\xi}_k \in \Xi_k$ is Bernoulli with parameter $p_{D,2}(\boldsymbol{\xi}_k)$ and a likelihood that $\mathbf{z}_k$ is observed from $\boldsymbol{\xi}_k$ on the extent of target $\mathbf{x}_k$ given by $g_k(\mathbf{z}_k \,|\, \boldsymbol{\xi}_k, \mathbf{x}_k)$;
- the RFS K_k is modelled with a hierarchical RFS representation whose probability density is given by equations (3.19) and (3.20);

- the RFS Υ_k in the hierarchical RFS representation of the predicted state is approximated with a Poisson RFS, and likewise each subsidiary RFS $\mathcal{E}_k$ is also approximated with a Poisson RFS whose intensity is an arbitrary function denoted by $\mu(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k)$.

Note that the assumptions regarding the target dynamics only consider the evolution of members in the RFS Υ_{k-1}, since there is no interest in the evolution of the feature points constituting the subsidiary RFSs $\mathcal{E}_{k-1}$.

Theorem 3.6 (**PHD prediction for extended targets**). *Given the listed assumptions regarding target dynamics and a prior marginal intensity denoted by* $v_{k-1}(\mathbf{x}_{k-1}) = v_{k-1}(\mathbf{x}_{k-1} \,|\, Z_{1:k-1})$*, then the predicted marginal intensity, denoted by* $v_{k|k-1}(\mathbf{x}_k) = v_{k|k-1}(\mathbf{x}_k \,|\, Z_{1:k-1})$*, is*

$$v_{k|k-1}(\mathbf{x}_k) = \gamma_k(\mathbf{x}_k) + \int p_S(\mathbf{x}_{k-1})\, f_{k|k-1}(\mathbf{x}_k \,|\, \mathbf{x}_{k-1})\, v_{k-1}(\mathbf{x}_{k-1})\, \mathrm{d}\mathbf{x}_{k-1}, \tag{3.38}$$

where $\gamma_k(\mathbf{x}_k)$ *is the intensity of target births.*

Proof. Details on the derivation are given in Appendix B (Section B.6.1). □

Theorem 3.7 (**PHD update for extended targets**). *Suppose a new set of measurements, denoted by* $Z = Z_k$*, is received at time-step* k *then, given the assumptions regarding the observations and the RFS representation of the predicted state hold, the updated marginal intensity, denoted by* $v_k(\mathbf{x}_k) = v_k(\mathbf{x}_k \,|\, Z_{1:k})$*, is*

$$\begin{aligned} v_k(\mathbf{x}_k) = v_{k|k-1}(\mathbf{x}_k) \Bigg\{ & 1 - p_{D,1}(\mathbf{x}_k) + p_{D,1}(\mathbf{x}_k) \exp\left(-\mu[p_{D,2} \,|\, \mathbf{x}_k]\right) \\ & + \sum_{\pi \in \Pi_Z} \varpi_\pi \sum_{\varphi \in \pi} \frac{p_{D,1}(\mathbf{x}_k)\, L_\varphi(\mathbf{x}_k)}{v_{k|k-1}[p_{D,1}\, L_\varphi] + \Lambda(\varphi)} \Bigg\}. \end{aligned} \tag{3.39}$$

The summation $\sum_{\pi \in \Pi_Z}$ *is taken over all possible partitions of the measurement set* Z *and the partition weights are given by the same expression in equation* (3.23)*. The pseudo multiple measurement likelihood also has the same expression as given in equation* (3.24) *except with* μ *replacing* $v_{k|k-1}$*, and the clutter term* $\Lambda(\varphi)$ *is as defined in equation* (3.20)*.*

Proof. Details on the derivation are given in Appendix B (Section B.6.2). □

Note that the result in Theorem 3.6 is precisely the multi-target PHD prediction and the result in Theorem 3.7 is closely related to the PHD update for non-standard targets given by equation (2.36).

3.3 Limitations

Despite the first-order moment approximations operating on the joint single group, single individual target space in group target tracking problems (single target space in extended target tracking problems) their implementation still presents a computational challenge due to the combinatorial nature of the update equations in particular. This section provides further explanation as to the origin of this combinatorial complexity and a review of various techniques available in addressing the problem.

3.3.1 Partitioning of the measurement set

An integral part of group and extended target tracking via first-order moment approximations is the partitioning of the measurement set. The update equations (3.21) for multiple group targets and (3.39) for multiple extended targets, involve the summation over all possible partitions Π_Z of the measurement set $Z = Z_k$, where a single partition π contains subsets φ consisting of measurements in Z such that $\bigcup_{\varphi\in\pi}\varphi = Z$. To illustrate, consider the partitions of $Z_1 = \{z_1, z_2\}$ and $Z_2 = \{z_1, z_2, z_3\}$. For the measurement set containing two measurements, Z_1, there are only two possible partitions:

$$\pi_1 = \{z_1, z_2\}; \quad \pi_2 = \{z_1\}, \{z_2\}. \tag{3.40}$$

The partitions of the measurement set Z_2 can be determined from the partitions of $Z_1 \cup \{z_3\}$ as follows. Firstly, consider taking a single partition $\pi \in \Pi_{Z_1}$ and appending the subset $\{z_3\}$ to get $\pi \cup \{z_3\}$. Repeating for all partitions of Z_1 gives

$$\pi_1 = \{z_1, z_2\}, \{z_3\}; \quad \pi_2 = \{z_1\}, \{z_2\}, \{z_3\}. \tag{3.41}$$

The remaining partitions of $Z_2 = Z_1 \cup \{z_3\}$ are found by replacing a subset φ from a single partition $\pi \in \Pi_{Z_1}$ with $\varphi \cup \{z_3\}$, repeating for each subset $\varphi \in \pi$ to get a new set of partitions $\{(\pi \setminus \varphi) \cup (\varphi \cup \{z_3\}) : \forall\, \varphi \in \pi\}$, and then repeating for all partitions of Z_1 to give

$$\pi_3 = \{z_1, z_2, z_3\}; \qquad \pi_4 = \{z_1, z_3\}, \{z_2\}; \qquad \pi_5 = \{z_1\}, \{z_2, z_3\}. \tag{3.42}$$

Hence there are a total of five possible partitions of the measurement set containing three measurements. Note that the approach, demonstrated above, for determining the partitions of a set whose size is $n + 1$ from the known partitions of a set whose

size is n is utilised in the inductive derivation of the PHD update equation for extended targets in [100] and more recently in the CPHD update equation for group targets in [111]. Furthermore, the number of partitions $|\Pi_Z|$ for a measurement set of size $|Z| = n$ is in fact given by the n^{th} Bell number, B_n [114]. Table 3.1 lists the Bell numbers for $n = 1, \ldots, 10$ and shows how combinatorially complex the operation of summing over all possible partitions of Z becomes for a large number of measurements. However, there exist various clustering techniques for addressing the partitioning problem.

Centroid based clustering assumes clusters are represented by a central vector, not necessarily a point in the measurement set. When the number of clusters is fixed to say k, the k-means algorithm can be applied [115], which finds k cluster centres and assigns points to the nearest centre. The resulting partition is equivalent to the Voronoi tessellation [116], which given a set of distinct central vectors in the Euclidean plane, associates all locations in that space with the closet member of the central vector set, so that the plane is partitioned into k regions. The drawbacks of this clustering technique are that it requires the number of clusters k to be specified in advance and prefers clusters to be approximately similar sizes, otherwise the borders between clusters (i.e. the boundaries of the regions in the Voronoi tessellation) are likely to be ill-defined. Such a technique would not be appropriate for partitioning a measurement set that incorporates clutter.

Connectivity based clustering, otherwise referred to as hierarchical clustering, is a technique that involves connecting points to form clusters based on their distance, for example nearest neighbour methods [117]. As such, a cluster is described by the maximum distance needed to connect points in the cluster. Different distance choices give different cluster formations and so unique partitioning of the measurement set is not always achievable. An example of such a technique, is a nearest neighbour approach proposed for determining the partitions of measurement sets in the Gaussian mixture application of the PHD filter for extended targets in [101]. In this proposed approach, the measurement set is divided into N_d sub-partitions

$\lvert Z\rvert$	$\lvert\Pi_Z\rvert = B_n$	$\lvert Z\rvert$	$\lvert\Pi_Z\rvert = B_n$
1	1	6	203
2	2	7	877
3	5	8	4,140
4	15	9	21, 147
5	52	10	115,975

Table 3.1: List of the number of partitions $|\Pi_Z|$, where $|Z| = n$, for $n = 1, \ldots, 10$

with distance thresholds $\{d_i\}_{i=1}^{N_d}$ where $d_i < d_{i+1}$, $\forall\, i$, such that partition i consists of subsets containing measurements that are no more than d_i metres apart from their nearest neighbours.

Density based clustering techniques [118] are often similar to connectivity based clustering, and assume clusters are defined as areas of higher density than the remainder of the measurement set. That is, points are connected to form clusters based on a certain distance threshold and whether they satisfy a density criterion. Measurement points in sparse areas (outlying the areas of high density) are considered to be clutter. The main drawback of density based clustering is that the cluster densities are expected to drop in order to determine distinct clusters, which is a problem when clusters are close or overlapping. However, this is a common problem in cluster analysis/set partitioning, and one which is not so easy to deal with, whatever technique is employed.

Distribution based clustering refers to statistical approaches to clustering, where clusters are defined from point measurements that most likely belong to the same distribution, which often closely resembles the way simulated measurement sets are generated. One such approach is the well known expectation-maximization (EM) clustering method [119, 120], which models the measurement set with a fixed number of Gaussian distributions, employed recently as a mixture reduction method in extended target tracking [121]. In 'hard' clustering (each measurement belongs to a cluster or not), measurements are assigned to the Gaussian distribution they most likely belong to, whereas in 'soft' clustering each measurement is assigned to each Gaussian distribution and weighted according to their likelihood. The appropriateness of distribution-based clustering techniques in general depends on the choice of distribution models to describe the measurement set.

In the following chapters, Gaussian mixture implementations of the PHD filter are presented for group and extended targets, formulated under linear Gaussian assumptions, so the focus is on the latter type of clustering techniques. In particular, a procedure that validates measurements lying within a region defined by a Gaussian ellipse for each estimated cluster centre, the details of which will be provided with the Gaussian mixture (GM) formulations later. Known as validation or gating technique [64], the procedure has previously been applied to the partitioning of the set of measurements for defining groups of targets in [122]. The main attraction of gating is that it results in a single partition $\boldsymbol{\pi}$, hence removing the summation $\sum_{\pi\in\Pi_Z}$ (and weights ϖ_π) in the update equations (3.21) for group targets and (3.39) for extended targets.

3.3.2 Multiple measurement likelihood

The partitioning problem which arises from implementing the first-order moment approximations of the Bayesian filter for multiple group/extended targets, specifically in the update equations (3.21) & (3.39), is not present in the special case of the single group target. However, there is an aspect of the update equations given in Theorem 3.3 (and Theorem 3.5) which does present a similar computational challenge at the implementation stage, namely the multiple measurement likelihood.

Consider the expression for this multiple measurement likelihood $L_Z(\,\cdot\,)$ given by equation (3.28). In order to implement either of the single group target update intensities the product term in $L_Z(\,\cdot\,)$ needs to be re-written as

$$\begin{aligned}\prod_{\mathbf{z}\in Z}\Big(\kappa(\mathbf{z}) + v_{k|k-1}[p_{D,2}\, l_{\mathbf{z}}\,|\,\cdot]\Big) &= \sum_{W\subseteq Z}\left(\prod_{\mathbf{z}\in Z\setminus W}\kappa(\mathbf{z})\right)\left(\prod_{\mathbf{w}\in W} v_{k|k-1}[p_{D,2}\, l_{\mathbf{w}}\,|\,\cdot]\right)\\ &= \left(\prod_{\mathbf{z}\in Z}\kappa(\mathbf{z})\right)\sum_{W\subseteq Z}\prod_{\mathbf{w}\in W}\frac{v_{k|k-1}[p_{D,2}\, l_{\mathbf{w}}\,|\,\cdot]}{\kappa(\mathbf{w})}. \qquad (3.43)\end{aligned}$$

That is, $L_Z(\,\cdot\,)$ can be written in terms of the multiple measurement likelihood associated with the single cluster existence update, i.e. $L_Z(\,\cdot\,) = \Big(\prod_{\mathbf{z}\in Z}\kappa(\mathbf{z})\Big)\tilde{L}_Z(\,\cdot\,)$. Computation of $L_Z(\,\cdot\,)$ written in this form now involves the summation over all possible non-empty subsets W in the entire measurement set Z, the order of which is $\mathcal{O}(2^n)$ for $|Z| = n$, inclusive of the empty subset $W = \emptyset$.

The combinatorial complexity here can be overcome by performing the gating procedure mentioned earlier for addressing the partitioning problem. Gating in this case will filter out the majority of the clutter, leaving measurements (some of which still may be clutter) more closely related to the single group target within the validation region. The validated measurements constitute a single subset, say $\overline{W}$, and consequently the summation $\sum_{W\subseteq Z}$ is removed from the equation (3.43), so that now the multiple measurement likelihood is denoted by $L_{\overline{W}}(\,\cdot\,)$. This useful gating procedure significantly reduces the computational demand for the single group target update and considered for the illustrative example presented in the next chapter. It can also be applied in the single cluster existence update as demonstrated for the closely related Bernoulli filter proposed by Ristic and Sherrah [123, 124] for the joint detection and tracking of an extended target, a generalisation of the JoTT filter.

Chapter 4

GM implementation of the PHD filter for Group Targets

In general, the recursion for the multi-group PHD filter in Section 3.2.3 has no closed-form solutions. However, under linear Gaussian assumptions on the dynamic and observation models, a closed-form solution exists that formulates the posterior intensity at any time-step as a Gaussian mixture. Since the group states and individual target states are jointly estimated via the first-order moments in this recursion, the components in the mixture are formulated as joint Gaussian functions whose means, covariances and weights are analytically propagated in time.

A Gaussian mixture (GM) formulation for the multi-group PHD filter recursion is proposed in Section 4.1, whose implementation requires the careful management of the number of Gaussian components as considered in Section 4.2. For illustration purposes, the implementation of a similar GM formulation for the special case, the single group PHD filter, is shown in Section 4.3 with a simulated example and its performance is analysed.

4.1 The Multi-Group GM-PHD Filter

A closed-form solution exists for the multi-group PHD filter recursion under certain Gaussian assumptions shown in Sections 4.1.1–4.1.4. The resulting multi-group GM-PHD filter recursion is then presented in Sections 4.1.5 and 4.1.6.

4.1.1 GM prior intensity

Suppose that at the beginning of each iteration of the filter, the prior intensity is known and formulated with the following Gaussian mixture

$$v_{k-1}(\mathbf{x}_{k-1}, \boldsymbol{\xi}_{k-1}) = \sum_{i=1}^{J_{k-1}} \nu_{k-1}^{(i)} \sum_{j=1}^{J_{k-1}^{(i)}} \omega_{k-1}^{(j)} \, \mathcal{N}\left(\mathbf{X}_{k-1}\,; \boldsymbol{\mu}_{k-1}^{(i,j)}, \boldsymbol{\Sigma}_{k-1}^{(i,j)}\right), \tag{4.1}$$

where the components are joint Gaussian functions with $\mathbf{X}_{k-1} = \begin{bmatrix}\boldsymbol{\xi}_{k-1} & \mathbf{x}_{k-1}\end{bmatrix}^T$ and

$$\boldsymbol{\mu}_{k-1}^{(i,j)} = \begin{bmatrix}\mathbf{m}_{\boldsymbol{\xi},k-1}^{(j|i)} & \mathbf{m}_{\mathbf{x},k-1}^{(i)}\end{bmatrix}^T, \qquad \boldsymbol{\Sigma}_{k-1}^{(i,j)} = \begin{bmatrix}\mathbf{P}_{11,k-1}^{(j|i)} & \mathbf{P}_{12,k-1}^{(j|i)} \\ \left(\mathbf{P}_{12,k-1}^{(j|i)}\right)^T & \mathbf{P}_{22,k-1}^{(i)}\end{bmatrix}. \tag{4.2}$$

The notation $\mathcal{N}(\,\cdot\,;\boldsymbol{\mu},\boldsymbol{\Sigma})$ is used to denote a Gaussian distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Furthermore, $\mathbf{m}_{\mathbf{x},k-1}^{(i)}$, $\mathbf{P}_{22,k-1}^{(i)}$ are the means and covariances for the Gaussian components corresponding to group states, and $\mathbf{m}_{\boldsymbol{\xi},k-1}^{(j|i)}$, $\mathbf{P}_{11,k-1}^{(j|i)}$ are the means and covariances for the Gaussian components corresponding to individual target states for each group. The cross diagonal covariances $\mathbf{P}_{12,k-1}^{(j|i)}$ relate to the strength of dependence each group has on its constituent targets.

The Gaussian mixture formulation in equation (4.1) satisfies the factorisation of the joint prior intensity given by $v_{k-1}(\mathbf{x}_{k-1},\boldsymbol{\xi}_{k-1}) = v_{k-1}(\mathbf{x}_{k-1}) \times v_{k-1}(\boldsymbol{\xi}_{k-1}\,|\,\mathbf{x}_{k-1})$ so that

$$v_{k-1}(\mathbf{x}_{k-1}) = \sum_{i=1}^{J_{k-1}} \nu_{k-1}^{(i)}\, \mathcal{N}\Big(\mathbf{x}_{k-1}\,;\mathbf{m}_{\mathbf{x},k-1}^{(i)},\mathbf{P}_{22,k-1}^{(i)}\Big), \tag{4.3}$$

$$v_{k-1}(\boldsymbol{\xi}_{k-1}\,|\,\mathbf{x}_{k-1}) = \sum_{j=1}^{J_{k-1}^{(i)}} \omega_{k-1}^{(j)}\, \mathcal{N}\Big(\boldsymbol{\xi}_{k-1}\,;\boldsymbol{\mu}_{\boldsymbol{\xi}|\mathbf{x},k-1}^{(i,j)},\boldsymbol{\Sigma}_{\boldsymbol{\xi}|\mathbf{x},k-1}^{(i,j)}\Big). \tag{4.4}$$

The Gaussian components in (4.4) are conditional Gaussian realisations from the joint Gaussian components in (4.1), with means and covariances given by

$$\begin{aligned}\boldsymbol{\mu}_{\boldsymbol{\xi}|\mathbf{x},k-1}^{(i,j)} &= \mathbf{m}_{\boldsymbol{\xi},k-1}^{(j|i)} + \mathbf{P}_{12,k-1}^{(j|i)}\left(\mathbf{P}_{22,k-1}^{(i)}\right)^{-1}\left(\mathbf{x}_{k-1} - \mathbf{m}_{\mathbf{x},k-1}^{(i)}\right), \\ \boldsymbol{\Sigma}_{\boldsymbol{\xi}|\mathbf{x},k-1}^{(i,j)} &= \mathbf{P}_{11,k-1}^{(j|i)} - \mathbf{P}_{12,k-1}^{(j|i)}\left(\mathbf{P}_{22,k-1}^{(i)}\right)^{-1}\left(\mathbf{P}_{12,k-1}^{(j|i)}\right)^T.\end{aligned} \tag{4.5}$$

Further details on the formulation of conditional Gaussian functions are given in the appendix.

4.1.2 Linear Gaussian dynamics

Given the hierarchical RFS representation of the multiple group target system at time $k-1$, the evolution of each independent group state and a constituent target state to the next time-step is characterised by the joint Markov transition density, denoted by $f_{k|k-1}(\mathbf{x}_k,\boldsymbol{\xi}_k\,|\,\mathbf{x}_{k-1},\boldsymbol{\xi}_{k-1})$. Suppose the joint Markov transition density

is modelled with the following linear Gaussian

$$f_{k|k-1}(\mathbf{x}_k, \boldsymbol{\xi}_k \,|\, \mathbf{x}_{k-1}, \boldsymbol{\xi}_{k-1}) = \mathcal{N}\left(\mathbf{X}_k \,;\, \mathbf{F}_{k-1}\mathbf{X}_{k-1}, \mathbf{Q}_{k-1}\right), \tag{4.6}$$

where $\mathbf{X}_k = \begin{bmatrix}\boldsymbol{\xi}_k & \mathbf{x}_k\end{bmatrix}^T$ and

$$\mathbf{F}_{k-1} = \begin{bmatrix}\mathbf{F}_{11} & \mathbf{F}_{12} \\ \mathbf{0} & \mathbf{F}_{22}\end{bmatrix}, \qquad \mathbf{Q}_{k-1} = \begin{bmatrix}\mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{12}^T & \mathbf{Q}_{22}\end{bmatrix}. \tag{4.7}$$

The transition matrix $\mathbf{F}_{22}$ describes the group target dynamics, $\mathbf{F}_{11}$ describes the individual target dynamics and $\mathbf{F}_{12}$ is the transition matrix for the daughter process due to the group target dynamics.

The linear Gaussian model in equation (4.6) satisfies the following factorisation $f_{k|k-1}(\mathbf{x}_k, \boldsymbol{\xi}_k \,|\, \mathbf{x}_{k-1}, \boldsymbol{\xi}_{k-1}) = f_{k|k-1}(\mathbf{x}_k \,|\, \mathbf{x}_{k-1}) \times f_{k|k-1}(\boldsymbol{\xi}_k \,|\, \boldsymbol{\xi}_{k-1}, \mathbf{x}_{k-1})$ so that

$$f_{k|k-1}(\mathbf{x}_k \,|\, \mathbf{x}_{k-1}) = \mathcal{N}\left(\mathbf{x}_k \,;\, \mathbf{F}_{22}\mathbf{x}_{k-1}, \mathbf{Q}_{22}\right), \tag{4.8}$$

$$f_{k|k-1}(\boldsymbol{\xi}_k \,|\, \boldsymbol{\xi}_{k-1}, \mathbf{x}_{k-1}) = \mathcal{N}\big(\boldsymbol{\xi}_k \,;\, \mathbf{F}_{11}\boldsymbol{\xi}_{k-1} + \mathbf{F}_{12}\mathbf{x}_{k-1} + \mathbf{Q}_{12}\mathbf{Q}_{22}^{-1}(\mathbf{x}_k - \mathbf{F}_{22}\mathbf{x}_{k-1}), \mathbf{Q}_{\text{Schur}}\big), \tag{4.9}$$

where $\mathbf{Q}_{\text{Schur}} = \mathbf{Q}_{11} - \mathbf{Q}_{12}\mathbf{Q}_{22}^{-1}\mathbf{Q}_{12}^T$ given process noise covariance matrices $\mathbf{Q}_{11}$, $\mathbf{Q}_{12}$ & $\mathbf{Q}_{22}$. Note that $f_{k|k-1}(\mathbf{x}_k \,|\, \mathbf{x}_{k-1})$ and $f_{k|k-1}(\boldsymbol{\xi}_k \,|\, \boldsymbol{\xi}_{k-1}, \mathbf{x}_{k-1})$ denote the single-group and single-individual target Markov transition densities given in equations (3.17) and (3.18) respectively. In addition, suppose the probabilities of survival, given by $p_{S,1}(\mathbf{x}_{k-1})$ and $p_{S,2}(\boldsymbol{\xi}_{k-1})$ in equations (3.17) and (3.18) respectively, are state independent, i.e. they are arbitrary constants denoted now by $p_{S,1}$ and $p_{S,2}$ respectively.

4.1.3 GM birth intensities

The intensity of the RFS Γ_k modelling the group births and the conditional intensity of the RFS $\Gamma_{\xi,k}$ modelling the birth of new individual targets are formulated with the following Gaussian mixtures

$$\gamma_k(\mathbf{x}_k, \boldsymbol{\xi}_k) = \sum_{i=1}^{J_{\gamma,k}} \nu_{\gamma,k}^{(i)} \sum_{j=1}^{J_{\gamma,k}^{(i)}} \omega_{\gamma,k}^{(j)} \, \mathcal{N}\Big(\mathbf{X}_k \,;\, \boldsymbol{\mu}_{\gamma,k}^{(i,j)}, \boldsymbol{\Sigma}_{\gamma,k}^{(i,j)}\Big), \tag{4.10}$$

$$\gamma_k(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k) = \sum_{j=1}^{J_{\gamma,k}^{(i)}} \omega_{\gamma,k}^{(j)} \, \mathcal{N}\Big(\boldsymbol{\xi}_k \,;\, \boldsymbol{\mu}_{\gamma,\boldsymbol{\xi},k}^{(i,j)}, \boldsymbol{\Sigma}_{\gamma,\boldsymbol{\xi},k}^{(i,j)}\Big). \tag{4.11}$$

In equation (4.10), the means and covariances of the Gaussian components in the mixture are given by

$$\boldsymbol{\mu}_{\gamma,k}^{(i,j)} = \begin{bmatrix} \mathbf{m}_{\gamma,\boldsymbol{\xi},k}^{(j|i)} & \mathbf{m}_{\gamma,\mathbf{x},k}^{(i)} \end{bmatrix}^T, \qquad \boldsymbol{\Sigma}_{\gamma,k}^{(i,j)} = \begin{bmatrix} \mathbf{P}_{\gamma,11,k}^{(j|i)} & \mathbf{P}_{\gamma,12,k}^{(j|i)} \\ \left(\mathbf{P}_{\gamma,12,k}^{(j|i)}\right)^T & \mathbf{P}_{\gamma,22,k}^{(i)} \end{bmatrix}. \tag{4.12}$$

while in equation (4.11), the means and covariances of the Gaussian components in the mixture are given by

$$\begin{aligned} \boldsymbol{\mu}_{\gamma,\boldsymbol{\xi},k}^{(i,j)} &= \mathbf{m}_{\gamma,\boldsymbol{\xi},k}^{(j|i)} + \mathbf{P}_{\gamma,12,k}^{(j|i)} \left(\mathbf{P}_{22,k|k-1}^{(i)}\right)^{-1} \left(\mathbf{x}_{k-1} - \mathbf{m}_{\mathbf{x},k|k-1}^{(i)}\right), \\ \boldsymbol{\Sigma}_{\gamma,\boldsymbol{\xi},k}^{(i,j)} &= \mathbf{P}_{\gamma,11,k}^{(j|i)} - \mathbf{P}_{\gamma,12,k}^{(j|i)} \left(\mathbf{P}_{22,k|k-1}^{(i)}\right)^{-1} \left(\mathbf{P}_{\gamma,12,k}^{(j|i)}\right)^T. \end{aligned} \tag{4.13}$$

The Gaussian components in (4.11) are further conditional Gaussian realisations, this time from the joint Gaussian components with means and covariances given by

$$\boldsymbol{\mu}_{\gamma,k}^{(i,j)} = \begin{bmatrix} \mathbf{m}_{\gamma,\boldsymbol{\xi},k}^{(j|i)} & \mathbf{m}_{\mathbf{x},k|k-1}^{(i)} \end{bmatrix}^T, \qquad \boldsymbol{\Sigma}_{\gamma,k}^{(i,j)} = \begin{bmatrix} \mathbf{P}_{\gamma,11,k}^{(j|i)} & \mathbf{P}_{\gamma,12,k}^{(j|i)} \\ \left(\mathbf{P}_{\gamma,12,k}^{(j|i)}\right)^T & \mathbf{P}_{22,k|k-1}^{(i)} \end{bmatrix}, \tag{4.14}$$

for $i = 1, \ldots, J_{k-1}$, where $\mathbf{m}_{\mathbf{x},k|k-1}^{(i)}$, $\mathbf{P}_{22,k|k-1}^{(i)}$ are the predicted means and covariance of the Gaussian components corresponding to group states in the mixture to be determined for the predicted intensity.

4.1.4 Linear Gaussian measurement likelihood

The final model assumption concerns the single measurement likelihood, denoted by $g_k(\mathbf{z} \,|\, \boldsymbol{\xi}_k, \mathbf{x}_k)$ in equation (3.22). Suppose this likelihood function is modelled with the following factorisation of linear Gaussian functions

$$g_k(\mathbf{z} \,|\, \boldsymbol{\xi}_k, \mathbf{x}_k) = \mathcal{N}(\mathbf{z}\,; \mathbf{H}\boldsymbol{\xi}_k, \mathbf{R}_1) \times \mathcal{N}(\mathbf{z}\,; \widehat{\mathbf{H}}\mathbf{x}_k, \mathbf{R}_2), \tag{4.15}$$

where $\mathcal{N}(\mathbf{z}\,; \mathbf{H}\boldsymbol{\xi}_k, \mathbf{R}_1)$ is the likelihood that the measurement vector $\mathbf{z}$ is related to individual target $\boldsymbol{\xi}_k$, with projection matrix $\mathbf{H}$ and observation noise covariance matrix $\mathbf{R}_1$, and $\mathcal{N}(\mathbf{z}\,; \widehat{\mathbf{H}}\mathbf{x}_k, \mathbf{R}_2)$ is the likelihood that the measurement $\mathbf{z}$ relates to group target $\mathbf{x}_k$, with projection matrix $\widehat{\mathbf{H}}$ and observation noise covariance matrix $\mathbf{R}_2$. In addition, suppose the probabilities of detection, given by $p_{D,1}(\mathbf{x}_k)$ and $p_{D,2}(\boldsymbol{\xi}_k)$ in equations (3.21) and (3.22) respectively, are state independent, i.e. they are arbitrary constants denoted now by $p_{D,1}$ and $p_{D,2}$ respectively.

4.1.5 The multi-group GM-PHD prediction

Given the linear Gaussian assumptions on the dynamics along with the GM formulations of the birth intensities and prior intensity described in Sections 4.1.1–4.1.3, the predicted intensity from Theorem 3.1 can be formulated as a Gaussian mixture, as proposed in the following.

Proposition 4.1. *Suppose that the assumptions on the dynamics and the probabilities of survival hold. Then, given the Gaussian mixture formulations of the prior intensity and birth intensities in equations* (4.1), (4.10) *and* (4.11) *respectively, the predicted intensity is also a Gaussian mixture of the form*

$$v_{k|k-1}(\mathbf{x}_k, \boldsymbol{\xi}_k) = \sum_{\bar{\imath}=1}^{J_{k|k-1}} \nu_{k|k-1}^{(\bar{\imath})} \sum_{\bar{\jmath}=1}^{J_{k|k-1}^{(\bar{\imath})}} \omega_{k|k-1}^{(\bar{\jmath})} \mathcal{N}\Big(\mathbf{X}_k \,; \boldsymbol{\mu}_{k|k-1}^{(\bar{\imath},\bar{\jmath})}, \boldsymbol{\Sigma}_{k|k-1}^{(\bar{\imath},\bar{\jmath})}\Big), \tag{4.16}$$

with means and covariances given by

$$\boldsymbol{\mu}_{k|k-1}^{(\bar{\imath},\bar{\jmath})} = \begin{bmatrix} \mathbf{m}_{\boldsymbol{\xi},k|k-1}^{(\bar{\jmath}|\bar{\imath})} & \mathbf{m}_{\mathbf{x},k|k-1}^{(\bar{\imath})} \end{bmatrix}^T, \quad \boldsymbol{\Sigma}_{k|k-1}^{(\bar{\imath},\bar{\jmath})} = \begin{bmatrix} \mathbf{P}_{11,k|k-1}^{(\bar{\jmath}|\bar{\imath})} & \mathbf{P}_{12,k|k-1}^{(\bar{\jmath}|\bar{\imath})} \\ \left(\mathbf{P}_{12,k|k-1}^{(\bar{\jmath}|\bar{\imath})}\right)^T & \mathbf{P}_{22,k|k-1}^{(\bar{\imath})} \end{bmatrix}, \tag{4.17}$$

and $J_{k|k-1} = J_{\gamma,k} + J_{k-1}$. *For new group targets* $\{\mathbf{x}_{k,\bar{\imath}} : \bar{\imath} = i = 1, \ldots, J_{\gamma,k}\}$ *with new individual targets* $\{\boldsymbol{\xi}_{k,\bar{\jmath}} : \bar{\jmath} = j = 1, \ldots, J_{\gamma,k}^{(i)}\}$, *the weights, means and covariances in equation* (4.16) *are given by*

$$\nu_{k|k-1}^{(\bar{\imath})} = \nu_{\gamma,k}^{(i)}, \quad \omega_{k|k-1}^{(\bar{\jmath})} = \omega_{\gamma,k}^{(j)}, \quad \boldsymbol{\mu}_{k|k-1}^{(\bar{\imath},\bar{\jmath})} = \boldsymbol{\mu}_{\gamma,k}^{(i,j)}, \quad \boldsymbol{\Sigma}_{k|k-1}^{(\bar{\imath},\bar{\jmath})} = \boldsymbol{\Sigma}_{\gamma,k}^{(i,j)}, \tag{4.18}$$

and $J_{k|k-1}^{(\bar{\imath})} = J_{\gamma,k}^{(i)}$. *For persistent group targets* $\{\mathbf{x}_{k,\bar{\imath}} : \bar{\imath} = J_{\gamma,k} + i \text{ for } i = 1, \ldots, J_{k-1}\}$ *the weights, means and covariances of the Gaussian components in equation* (4.16) *corresponding to the group states are given by*

$$\begin{aligned} \nu_{k|k-1}^{(\bar{\imath})} &= p_{S,1}\, \nu_{k-1}^{(i)}, \\ \mathbf{m}_{\mathbf{x},k|k-1}^{(\bar{\imath})} &= \mathbf{F}_{22}\mathbf{m}_{\mathbf{x},k-1}^{(i)}, \qquad \mathbf{P}_{22,k|k-1}^{(\bar{\imath})} = \mathbf{F}_{22}\mathbf{P}_{22,k-1}^{(i)}\mathbf{F}_{22}^T + \mathbf{Q}_{22}, \end{aligned} \tag{4.19}$$

with $J_{k|k-1}^{(\bar{\imath})} = J_{\gamma,k}^{(i)} + J_{k-1}^{(i)}$, *while the weights, means and covariances of the Gaussian components corresponding to new individual targets* $\{\boldsymbol{\xi}_{k,\bar{\jmath}} : \bar{\jmath} = j = 1, \ldots, J_{\gamma,k}^{(i)}\}$ *are given by*

$$\omega_{k|k-1}^{(\bar{\jmath})} = \omega_{\gamma,k}^{(j)}, \quad \mathbf{m}_{\boldsymbol{\xi},k|k-1}^{(\bar{\jmath}|\bar{\imath})} = \mathbf{m}_{\gamma,\boldsymbol{\xi},k}^{(j|i)}, \quad \mathbf{P}_{11,k|k-1}^{(\bar{\jmath}|\bar{\imath})} = \mathbf{P}_{\gamma,11,k}^{(j|i)}, \tag{4.20}$$

and those corresponding to persistent individual targets $\{\boldsymbol{\xi}_{k,\bar{\jmath}} : \bar{\jmath} = J_{\gamma,k}^{(i)} + j \text{ for } j = 1, \ldots, J_{k-1}^{(i)}\}$ *are given by*

$$\begin{aligned}
\omega_{k|k-1}^{(\bar{\jmath})} &= p_{S,2}\, \omega_{k-1}^{(j)}, \\
\mathbf{m}_{\boldsymbol{\xi},k|k-1}^{(\bar{\jmath}|\bar{\imath})} &= \mathbf{F}_{11}\mathbf{m}_{\boldsymbol{\xi},k-1}^{(j|i)} + \mathbf{F}_{12}\mathbf{m}_{\mathbf{x},k-1}^{(i)}, \\
\mathbf{P}_{11,k|k-1}^{(\bar{\jmath}|\bar{\imath})} &= \mathbf{F}_{11}\mathbf{P}_{11,k-1}^{(j|i)}\mathbf{F}_{11}^{T} + \Big(\mathbf{F}_{11}\mathbf{P}_{12,k-1}^{(j|i)}\mathbf{F}_{12}^{T}\Big)^{T} \\
&\quad + \mathbf{F}_{11}\mathbf{P}_{12,k-1}^{(i)}\mathbf{F}_{12}^{T} + \mathbf{F}_{12}\mathbf{P}_{22,k-1}^{(j|i)}\mathbf{F}_{12}^{T} + \mathbf{Q}_{11}.
\end{aligned} \tag{4.21}$$

The respective cross diagonal covariances corresponding to new and persistent individual targets in persistent groups are given by $\mathbf{P}_{12,k|k-1}^{(\bar{\jmath}|\bar{\imath})} = \mathbf{P}_{\gamma,12,k}^{(j|i)}$ *and*

$$\mathbf{P}_{12,k|k-1}^{(\bar{\jmath}|\bar{\imath})} = \mathbf{F}_{11}\mathbf{P}_{12,k-1}^{(j|i)}\mathbf{F}_{22}^{T} + \mathbf{F}_{12}\mathbf{P}_{22,k-1}^{(i)}\mathbf{F}_{22}^{T} + \mathbf{Q}_{12}. \tag{4.22}$$

Details on how the GM formulation in Proposition 4.1 can be established are given in Appendix C (Sections C.2.1).

4.1.6 The multi-group GM-PHD update

Given the remaining assumption described in Section 4.1.4 and the GM formulation for the predicted intensity proposed in Section 4.1.5, the updated intensity from Theorem 3.2 can also be formulated as a Gaussian mixture as proposed in the following.

Proposition 4.2. *Suppose that the assumptions on the single-measurement likelihood and probabilities of detections hold. Then, given the GM formulation of the predicted intensity in equation* (4.16)*, the updated intensity is also a Gaussian mixture of the form*

$$\begin{aligned}
&v_k(\mathbf{x}_k, \boldsymbol{\xi}_k) = \\
&\big(1 - p_{D,1} + p_{D,1}\, \mathrm{e}^{-p_{D,2}\,\alpha}\,(1 - p_{D,2})\big)\, v_{k|k-1}(\mathbf{x}_k, \boldsymbol{\xi}_k) + \sum_{\pi \in \Pi_Z} \varpi_\pi \sum_{\varphi \in \pi} v_{D,k}(\mathbf{x}_k, \boldsymbol{\xi}_k)_\varphi,
\end{aligned} \tag{4.23}$$

where $\alpha = \sum_{j=1}^{J_{k|k-1}^{(i)}} \omega_{k|k-1}^{(j)}$*; and*

$$v_{D,k}(\mathbf{x}_k, \boldsymbol{\xi}_k)_\varphi = \sum_{i=1}^{J_{k|k-1}} \sum_{j_{1:|\varphi|}} \nu_{\varphi,k}^{(j_{1:|\varphi|},i)} \sum_{\bar{\jmath}=1}^{J_k^{(1:|\varphi|,i)}} \omega_k^{(\bar{\jmath})}\, \mathcal{N}\Big(\mathbf{X}_k\,;\, \boldsymbol{\mu}_k^{(j_{1:|\varphi|},i,\bar{\jmath})}, \boldsymbol{\Sigma}_k^{(j_{1:|\varphi|},i,\bar{\jmath})}\Big), \tag{4.24}$$

with the means and covariances

$$\boldsymbol{\mu}_k^{(j_{1:|\varphi|},i,\bar{\jmath})} = \begin{bmatrix} \mathbf{m}_{\boldsymbol{\xi},k}^{(j_{1:|\varphi|},i,\bar{\jmath})} & \mathbf{m}_{\mathbf{x},k}^{(j_{1:|\varphi|},i)} \end{bmatrix}^T, \quad \boldsymbol{\Sigma}_k^{(j_{1:|\varphi|},i,\bar{\jmath})} = \begin{bmatrix} \mathbf{P}_{11,k}^{(j_{1:|\varphi|},i,\bar{\jmath})} & \mathbf{P}_{12,k}^{(j_{1:|\varphi|},i,\bar{\jmath})} \\ \left(\mathbf{P}_{12,k}^{(j_{1:|\varphi|},i,\bar{\jmath})}\right)^T & \mathbf{P}_{22,k}^{(j_{1:|\varphi|},i)} \end{bmatrix}, \tag{4.25}$$

and $J_k^{(1:|\varphi|,i)} = J_{k|k-1}^{(i)} + |\varphi| \times J_{k|k-1}^{(i)}$. *In equation* (4.24), *the sum* $\sum_{j_{1:|\varphi|}}$ *is taken over all permutations, denoted by* $j_{1:|\varphi|} = (j_1, \ldots, j_{|\varphi|})$, *of the indices* $j_\ell = 1, \ldots, J_{k|k-1}^{(i)}$ *for* $\ell = 1, \ldots, |\varphi|$.

The weights, means and covariances of the Gaussian components in (4.24) *relating to detected group states are* $\nu_{\varphi,k}^{(j_{1:|\varphi|},i)}$, $\mathbf{m}_{\mathbf{x},k}^{(j_{1:|\varphi|},i)}$ *and* $\mathbf{P}_{22,k}^{(j_{1:|\varphi|},i)}$. *More precisely, for* $1 \leq \ell \leq |\varphi|$,

$$\begin{aligned} \mathbf{m}_{\mathbf{x},k}^{(j_{1:\ell},i)} &= \boldsymbol{\eta}_{\mathbf{x},k}^{(j_{1:\ell},i)} + \boldsymbol{\mathcal{P}}_{22,k}^{(j_{1:\ell},i)} \left(\mathbf{A}_k^{(i,j_\ell)}\right)^T \left(\mathbf{S}_1^{(j_{1:\ell},i)}\right)^{-1} \left(\mathbf{z}_\ell - \mathbf{b}_1^{(j_{1:\ell},i)}\right) \\ &\quad + \mathbf{K}_{\mathbf{x},k}^{(j_{1:\ell},i)} \left(\mathbf{z}_\ell - \mathbf{b}_2^{(j_{1:\ell},i)}\right), \\ \mathbf{P}_{22,k}^{(j_{1:\ell},i)} &= \left(\mathbf{I} - \mathbf{K}_{\mathbf{x},k}^{(j_{1:\ell},i)} \widehat{\mathbf{H}}\right) \left(\mathbf{I} - \boldsymbol{\mathcal{P}}_{22,k}^{(j_{1:\ell},i)} \left(\mathbf{A}_k^{(i,j_\ell)}\right)^T \left(\mathbf{S}_1^{(j_{1:\ell},i)}\right)^{-1} \mathbf{A}_k^{(i,j_\ell)}\right) \boldsymbol{\mathcal{P}}_{22,k}^{(j_{1:\ell},i)}, \\ \mathbf{K}_{\mathbf{x},k}^{(j_{1:\ell},i)} &= \left(\mathbf{I} - \boldsymbol{\mathcal{P}}_{22,k}^{(j_{1:\ell},i)} \left(\mathbf{A}_k^{(i,j_\ell)}\right)^T \left(\mathbf{S}_1^{(j_{1:\ell},i)}\right)^{-1} \mathbf{A}_k^{(i,j_\ell)}\right) \boldsymbol{\mathcal{P}}_{22,k}^{(j_{1:\ell},i)} \widehat{\mathbf{H}}^T \left(\mathbf{S}_2^{(j_{1:\ell},i)}\right)^{-1}, \end{aligned} \tag{4.26}$$

denoting $\mathbf{A}_k^{(i,j_\ell)} = \mathbf{H}\mathbf{P}_{12,k|k-1}^{(j_\ell|i)} \left(\mathbf{P}_{22,k|k-1}^{(i)}\right)^{-1}$, *where*

$$\left.\begin{aligned} \mathbf{b}_1^{(j_{1:\ell},i)} &= \mathbf{H}\mathbf{m}_{\boldsymbol{\xi},k|k-1}^{(j_\ell|i)} + \mathbf{A}_k^{(i,j_\ell)} \left(\boldsymbol{\eta}_{\mathbf{x},k}^{(j_{1:\ell},i)} - \mathbf{m}_{\mathbf{x},k|k-1}^{(i)}\right), \\ \mathbf{S}_1^{(j_{1:\ell},i)} &= \mathbf{A}_k^{(i,j_\ell)} \boldsymbol{\mathcal{P}}_{22,k}^{(j_{1:\ell},i)} \left(\mathbf{A}_k^{(i,j_\ell)}\right)^T + \widetilde{\mathbf{R}}_k^{(i,j_\ell)}, \end{aligned}\right\} \tag{4.27}$$

$$\left.\begin{aligned} \mathbf{b}_2^{(j_{1:\ell},i)} &= \widehat{\mathbf{H}}\boldsymbol{\eta}_{\mathbf{x},k}^{(j_{1:\ell},i)} + \widehat{\mathbf{H}}\boldsymbol{\mathcal{P}}_{22,k}^{(j_{1:\ell},i)} \left(\mathbf{A}_k^{(i,j_\ell)}\right)^T \left(\mathbf{S}_1^{(j_{1:\ell},i)}\right)^{-1} \left(\mathbf{z}_\ell - \mathbf{b}_1^{(j_{1:\ell},i)}\right), \\ \mathbf{S}_2^{(j_{1:\ell},i)} &= \widehat{\mathbf{H}} \left(\mathbf{I} - \boldsymbol{\mathcal{P}}_{22,k}^{(j_{1:\ell},i)} \left(\mathbf{A}_k^{(i,j_\ell)}\right)^T \left(\mathbf{S}_1^{(j_{1:\ell},i)}\right)^{-1} \mathbf{A}_k^{(i,j_\ell)}\right) \boldsymbol{\mathcal{P}}_{22,k}^{(j_{1:\ell},i)} \widehat{\mathbf{H}}^T + \mathbf{R}_2. \end{aligned}\right\} \tag{4.28}$$

In (4.26) *and* (4.27), $\boldsymbol{\eta}_{\mathbf{x},k}^{(j_{1:\ell},i)}$ *and* $\boldsymbol{\mathcal{P}}_{22,k}^{(j_{1:\ell},i)}$ *are iteratively propagated means and covariances relating to group targets, given by*

$$\boldsymbol{\eta}_{\mathbf{x},k}^{(j_{1:\ell},i)} = \begin{cases} \mathbf{m}_{\mathbf{x},k|k-1}^{(i)} & \text{for } \ell = 1, \\ \mathbf{m}_{\mathbf{x},k}^{(j_{1:\ell-1},i)} & \text{for } 1 < \ell \leq |\varphi|, \end{cases} \quad \boldsymbol{\mathcal{P}}_{22,k}^{(j_{1:\ell},i)} = \begin{cases} \mathbf{P}_{22,k|k-1}^{(i)} & \text{for } \ell = 1, \\ \mathbf{P}_{22,k}^{(j_{1:\ell-1},i)} & \text{for } 1 < \ell \leq |\varphi|. \end{cases} \tag{4.29}$$

The weights $\nu_{\varphi,k}^{(j_{1:|\varphi|},i)}$ *are evaluated as follows*

$$\nu_{\varphi,k}^{(j_{1:|\varphi|},i)} = \frac{p_{D,1}\,\nu_{k|k-1}^{(i)}\mathrm{e}^{-p_{D,2}\,\alpha}\,L_{j_{1:|\varphi|}}^{(i)}}{\Lambda(\varphi)+\sum_{i=1}^{J_{k|k-1}}\sum_{j_{1:|\varphi|}} p_{D,1}\,\nu_{k|k-1}^{(i)}\mathrm{e}^{-p_{D,2}\,\alpha}\,L_{j_{1:|\varphi|}}^{(i)}}, \tag{4.30}$$

where $L_{j_{1:|\varphi|}}^{(i)}$ *denotes the multiple measurement likelihood given by*

$$L_{j_{1:|\varphi|}}^{(i)} = \prod_{\ell=1}^{|\varphi|}\omega_{k|k-1}^{(j_\ell)}\,p_{D,2}\,q_1^{(j_{1:\ell},i)}(\mathbf{z}_\ell)\,q_2^{(j_{1:\ell},i)}(\mathbf{z}_\ell), \tag{4.31}$$

and, for $1 \le \ell \le |\varphi|$,

$$\begin{aligned} q_1^{(j_{1:\ell},i)}(\mathbf{z}_\ell) &= \mathcal{N}\Big(\mathbf{z}_\ell\,;\mathbf{b}_1^{(j_{1:\ell},i)},\mathbf{S}_1^{(j_{1:\ell},i)}\Big),\\ q_2^{(j_{1:\ell},i)}(\mathbf{z}_\ell) &= \mathcal{N}\Big(\mathbf{z}_\ell\,;\mathbf{b}_2^{(j_{1:\ell},i)},\mathbf{S}_2^{(j_{1:\ell},i)}\Big). \end{aligned} \tag{4.32}$$

The weights, means and covariances of the Gaussian components in (4.24) *relating to individual targets are* $\omega_k^{(\bar{\jmath})}$, $\mathbf{m}_{\boldsymbol{\xi},k}^{(j_{1:|\varphi|},i,\bar{\jmath})}$ *and* $\mathbf{P}_{11,k}^{(j_{1:|\varphi|},i,\bar{\jmath})}$, *for missed detections, i.e.* $\bar{\jmath}=j=1,\ldots,J_{k|k-1}^{(i)}$,

$$\begin{aligned} \mathbf{m}_{\boldsymbol{\xi},k}^{(j_{1:|\varphi|},i,\bar{\jmath})} &= \mathbf{m}_{\boldsymbol{\xi},k|k-1}^{(j|i)} + \mathbf{P}_{12,k}^{(j_{1:|\varphi|},i,\bar{\jmath})}\Big(\mathbf{P}_{22,k}^{(j_{1:|\varphi|},i)}\Big)^{-1}\Big(\mathbf{m}_{\mathbf{x},k}^{(j_{1:|\varphi|},i)} - \mathbf{m}_{\mathbf{x},k|k-1}^{(i)}\Big),\\ \mathbf{P}_{11,k}^{(j_{1:|\varphi|},i,\bar{\jmath})} &= \mathbf{P}_{11,k|k-1}^{(j|i)} - \mathbf{P}_{12,k}^{(j_{1:|\varphi|},i,\bar{\jmath})}\Big(\mathbf{P}_{22,k}^{(j_{1:|\varphi|},i)}\Big)^{-1}\Big(\mathbf{P}_{12,k|k-1}^{(j|i)} - \mathbf{P}_{12,k}^{(j_{1:|\varphi|},i,\bar{\jmath})}\Big)^{T}, \end{aligned} \tag{4.33}$$

with weights $\omega_k^{(\bar{\jmath})} = (1-p_{D,2})\,\omega_{k|k-1}^{(j)}$ *and cross diagonal covariances given by*

$$\mathbf{P}_{12,k}^{(j_{1:|\varphi|},i,\bar{\jmath})} = \mathbf{P}_{12,k|k-1}^{(j|i)}\Big(\mathbf{P}_{22,k|k-1}^{(i)}\Big)^{-1}\mathbf{P}_{22,k}^{(j_{1:|\varphi|},i)}. \tag{4.34}$$

For detections, i.e. $\bar{\jmath} = J_{k|k-1}^{(i)} + (\ell-1)\times J_{k|k-1}^{(i)} + j_\ell$ *for* $j_\ell = 1,\ldots,J_{k|k-1}^{(i)}$ *and* $\ell = 1,\ldots,|\varphi|$,

$$\begin{aligned} \mathbf{m}_{\boldsymbol{\xi},k}^{(j_{1:|\varphi|},i,\bar{\jmath})} &= \widetilde{\mathbf{m}}_{\boldsymbol{\xi},k}^{(i,j_\ell)} + \mathbf{P}_{12,k}^{(j_{1:|\varphi|},i,\bar{\jmath})}\Big(\mathbf{P}_{22,k}^{(j_{1:|\varphi|},i)}\Big)^{-1}\Big(\mathbf{m}_{\mathbf{x},k}^{(j_{1:|\varphi|},i)} - \mathbf{m}_{\mathbf{x},k|k-1}^{(i)}\Big),\\ \mathbf{P}_{11,k}^{(j_{1:|\varphi|},i,\bar{\jmath})} &= \widetilde{\mathbf{P}}_{11,k}^{(i,j_\ell)} - \mathbf{P}_{12,k}^{(j_{1:|\varphi|},i,\bar{\jmath})}\Big(\mathbf{P}_{22,k}^{(j_{1:|\varphi|},i)}\Big)^{-1}\Big(\widetilde{\mathbf{P}}_{12,k}^{(i,j_\ell)} - \mathbf{P}_{12,k}^{(j_{1:|\varphi|},i,\bar{\jmath})}\Big)^{T}, \end{aligned} \tag{4.35}$$

with cross diagonal covariances given by

$$\mathbf{P}_{12,k}^{(j_{1:|\varphi|},i,\bar{\jmath})} = \widetilde{\mathbf{P}}_{12,k}^{(i,j_\ell)}\Big(\mathbf{P}_{22,k|k-1}^{(i)}\Big)^{-1}\mathbf{P}_{22,k}^{(j_{1:|\varphi|},i)}, \tag{4.36}$$

where

$$\begin{aligned}
\widetilde{\mathbf{m}}_{\boldsymbol{\xi},k}^{(i,j_\ell)} &= \mathbf{m}_{\boldsymbol{\xi},k|k-1}^{(j_\ell|i)} + \mathbf{K}_{\boldsymbol{\xi},k}^{(i,j_\ell)} \left(\mathbf{z}_\ell - \mathbf{H}\mathbf{m}_{\boldsymbol{\xi},k|k-1}^{(j_\ell|i)} \right), \\
\widetilde{\mathbf{P}}_{11,k}^{(i,j_\ell)} &= \left(\mathbf{I} - \mathbf{K}_{\boldsymbol{\xi},k}^{(i,j_\ell)} \mathbf{H} \right) \mathbf{P}_{11,k|k-1}^{(j_\ell|i)} - \widetilde{\mathbf{P}}_{12,k}^{(i,j_\ell)} \left(\mathbf{P}_{22,k|k-1}^{(i)} \right)^{-1} \left(\mathbf{K}_{\boldsymbol{\xi},k}^{(i,j_\ell)} \mathbf{H} \mathbf{P}_{12,k|k-1}^{(j_\ell|i)} \right)^T, \\
\widetilde{\mathbf{P}}_{12,k}^{(i,j_\ell)} &= \left(\mathbf{I} - \mathbf{K}_{\boldsymbol{\xi},k}^{(i,j_\ell)} \mathbf{H} \right) \mathbf{P}_{12,k|k-1}^{(j_\ell|i)}, \\
\mathbf{K}_{\boldsymbol{\xi},k}^{(i,j_\ell)} &= \boldsymbol{\Sigma}_{\boldsymbol{\xi}|\mathbf{x},k|k-1}^{(i,j_\ell)} \mathbf{H}^T \left(\widetilde{\mathbf{R}}_k^{(i,j_\ell)} \right)^{-1},
\end{aligned} \tag{4.37}$$

and weights evaluated as follows

$$\omega_k^{(\bar{\jmath})} = \frac{p_{D,2}\, \omega_{k|k-1}^{(j_\ell)}\, q_1^{(j_{1:\ell},i)}(\mathbf{z}_\ell)\, q_2^{(j_{1:\ell},i)}(\mathbf{z}_\ell)}{\sum_{j_\ell=1}^{J_{k|k-1}^{(i)}} p_{D,2}\, \omega_{k|k-1}^{(j_\ell)}\, q_1^{(j_{1:\ell},i)}(\mathbf{z}_\ell)\, q_2^{(j_{1:\ell},i)}(\mathbf{z}_\ell)}. \tag{4.38}$$

Finally, the partition weights are given by

$$\varpi_\pi = \frac{\prod_{\varphi\in\pi} \left(\Lambda(\varphi) + \sum_{i=1}^{J_{k|k-1}} \sum_{j_{1:|\varphi|}} p_{D,1}\, \nu_{k|k-1}^{(i)}\, \mathrm{e}^{-p_{D,2}\,\alpha}\, L_{j_{1:|\varphi|}}^{(i)} \right)}{\sum_{\pi'\in\Pi_Z} \prod_{\varphi'\in\pi'} \left(\Lambda(\varphi') + \sum_{i=1}^{J_{k|k-1}} \sum_{j_{1:|\varphi'|}} p_{D,1}\, \nu_{k|k-1}^{(i)}\, \mathrm{e}^{-p_{D,2}\,\alpha}\, L_{j_{1:|\varphi'|}}^{(i)} \right)}. \tag{4.39}$$

Details on how the GM formulation in Proposition 4.2 can be established are given in Appendix C (Section C.2.2).

4.1.7 Remarks on the GM-PHD recursion for group targets

Propositions 4.1 and 4.2 present a closed-form solution to the PHD filter recursion for multiple group targets, the resulting algorithm for which is given in Tables 4.1 and 4.2. It follows from Proposition 4.1 and 4.2 that, given the intensity at the initial time-step $k = 0$ is formulated with a Gaussian mixture, subsequent predicted and posterior intensities are also GM formulations.

The expected number of groups $N_{k|k-1}$ and the expected number of individual targets $N_{k|k-1}^{(\bar{\imath})}$, for each group, associated with the predicted intensity $v_{k|k-1}$ are obtained by summing the appropriate mixture weights. That is

$$N_{k|k-1} = p_{S,1}\, N_{k-1} + \sum_{i=1}^{J_{\gamma,k}} \nu_{\gamma,k}^{(i)}, \tag{4.40}$$

where $N_{k-1} = \sum_{i=1}^{J_{k-1}} \nu_{k-1}^{(i)}$, and for $\bar{\imath} = 1, \ldots, J_{k|k-1}$

given $\left\{\nu_{k-1}^{(i)}, \left\{\omega_{k-1}^{(j)}, \boldsymbol{\mu}_{k-1}^{(i,j)}, \boldsymbol{\Sigma}_{k-1}^{(i,j)}\right\}_{j=1}^{J_{k-1}^{(i)}}\right\}_{i=1}^{J_{k-1}}$

prediction

$\bar{\imath} := 0.$

for $i = 1, \ldots, J_{\gamma,k}$ (*New group targets*)

$\bar{\imath} := \bar{\imath} + 1, \quad \nu_{k|k-1}^{(\bar{\imath})} := \nu_{\gamma,k}^{(i)},$

$\bar{\jmath} := 0,$

for $j = 1, \ldots, J_{\gamma,k}^{(i)}$

$\bar{\jmath} := \bar{\jmath} + 1,$

$\omega_{k|k-1}^{(\bar{\jmath})} := \omega_{\gamma,k}^{(j)}, \quad \boldsymbol{\mu}_{k|k-1}^{(\bar{\imath},\bar{\jmath})} := \boldsymbol{\mu}_{\gamma,k}^{(i,j)}, \quad \boldsymbol{\Sigma}_{k|k-1}^{(\bar{\imath},\bar{\jmath})} := \boldsymbol{\Sigma}_{\gamma,k}^{(i,j)}.$

end

$J_{k|k-1}^{(\bar{\imath})} := \bar{\jmath}.$

end

for $i = 1, \ldots J_{k-1}$ (*Persistent group targets*)

$\bar{\imath} := \bar{\imath} + 1,$

$\nu_{k|k-1}^{(\bar{\imath})}, \ \mathbf{m}_{\mathbf{x},k|k-1}^{(\bar{\imath})}, \ \mathbf{P}_{22,k|k-1}^{(\bar{\imath})}$ are formulated as per equation (4.19),

$\bar{\jmath} := 0.$

for $j = 1, \ldots, J_{\gamma,k}^{(i)}$ (*New individual targets*)

$\bar{\jmath} := \bar{\jmath} + 1,$

$\omega_{k|k-1}^{(\bar{\jmath}|\bar{\imath})}, \ \mathbf{m}_{\boldsymbol{\xi},k|k-1}^{(\bar{\jmath}|\bar{\imath})}, \ \mathbf{P}_{11,k|k-1}^{(\bar{\jmath}|\bar{\imath})}$ are formulated as per equation (4.20),

and $\mathbf{P}_{12,k|k-1}^{\bar{\jmath}|\bar{\imath})} := \mathbf{P}_{\gamma,12,k}^{(j|i)}.$

end

for $j = 1, \ldots, J_{k-1}^{(i)}$ (*Persistent individual targets*)

$\bar{\jmath} := \bar{\jmath} + 1,$

$\omega_{k|k-1}^{(\bar{\jmath})}, \ \mathbf{m}_{\boldsymbol{\xi},k|k-1}^{(\bar{\jmath}|\bar{\imath})}, \ \mathbf{P}_{11,k|k-1}^{(\bar{\jmath}|\bar{\imath})}$ and $\mathbf{P}_{12,k|k-1}^{(\bar{\jmath}|\bar{\imath})}$ are formulated

as per equations (4.21)–(4.22).

end

$J_{k|k-1}^{(\bar{\imath})} := \bar{\jmath}.$

end

$J_{k|k-1} := \bar{\imath}.$

output $\left\{\nu_{k|k-1}^{(i)}, \left\{\omega_{k|k-1}^{(j)}, \boldsymbol{\mu}_{k|k-1}^{(i,j)}, \boldsymbol{\Sigma}_{k|k-1}^{(i,j)}\right\}_{j=1}^{J_{k|k-1}^{(i)}}\right\}_{i=1}^{J_{k|k-1}}$

Table 4.1: Pseudo-code for the multi-group GM-PHD prediction

$$N_{k|k-1}^{(\bar{\imath})} = \begin{cases} \displaystyle\sum_{j=1}^{J_{\gamma,k}^{(i)}} \omega_{\gamma,k}^{(j)} & \text{for } \bar{\imath} = 1, \ldots, J_{\gamma,k}, \\ p_{S,2}\, N_{k-1}^{(i)} + \displaystyle\sum_{j=1}^{J_{\gamma,k}^{(i)}} \omega_{\gamma,k}^{(j)} & \text{for } \bar{\imath} = J_{\gamma,k} + i \ \&\ i = 1, \ldots, J_{k-1}, \end{cases} \tag{4.41}$$

where $N_{k-1}^{(i)} = \sum_{j=1}^{J_{k-1}^{(i)}} \omega_{k-1}^{(j)}$ for $i = 1, \ldots, J_{k-1}$. Likewise for the expected number of groups N_k and the expected number of individual targets $N_k^{(\bar{\imath})}$, for each group, associated with the updated intensity v_k.

given $\left\{\nu_{k|k-1}^{(i)}, \left\{\omega_{k|k-1}^{(j)}, \boldsymbol{\mu}_{k|k-1}^{(i,j)}, \boldsymbol{\Sigma}_{k|k-1}^{(i,j)}\right\}_{j=1}^{J_{k|k-1}^{(i)}}\right\}_{i=1}^{J_{k|k-1}}$

update

$\bar{\imath} := 0.$

for $i = 1, \ldots, J_{k|k-1}$ (*Missed detected group targets*)

$\bar{\imath} := \bar{\imath} + 1,$

$\nu_k^{(\bar{\imath})} := (1 - p_{D,1} + p_{D,1}\, \mathrm{e}^{-p_{D,2}\,\alpha}\,(1 - p_{D,2}))\,\nu_{k|k-1}^{(i)}, \quad (\textit{where } \alpha = \sum_{j=1}^{J_{k|k-1}^{(i)}} \omega_{k|k-1}^{(j)})$

for $j = 1, \ldots, J_{k|k-1}^{(i)}$

$\bar{\jmath} := \bar{\jmath} + 1,$

$\omega_k^{(\bar{\jmath})} := \omega_{k|k-1}^{(j)}, \quad \boldsymbol{\mu}_k^{(\bar{\imath},\bar{\jmath})} := \boldsymbol{\mu}_{k|k-1}^{(i,j)}, \quad \boldsymbol{\Sigma}_k^{(\bar{\imath},\bar{\jmath})} := \boldsymbol{\Sigma}_{k|k-1}^{(i,j)}.$

end

$J_k^{(\bar{\imath})} := \bar{\jmath}.$

end

for each partition $\pi \in \Pi_{Z_k}$ (*Detected group targets*)

for each subset $\varphi \in \pi$

for $i = 1, \ldots, J_{k|k-1}$

for each permutation $j_{1:|\varphi|} = (j_1, \ldots, j_{|\varphi|})$ $\quad (\textit{where } j_\ell = 1, \ldots, J_{k|k-1}^{(i)})$

$\bar{\imath} := \bar{\imath} + 1,$

$\nu_k^{(\bar{\imath})} := \nu_{\varphi,k}^{(j_{1:|\varphi|},i)}, \quad \mathbf{m}_{\mathbf{x},k}^{(\bar{\imath})} := \mathbf{m}_{\mathbf{x},k}^{(j_{1:|\varphi|},i)}, \quad \mathbf{P}_{22,k}^{(\bar{\imath})} := \mathbf{P}_{22,k}^{(j_{1:|\varphi|},i)},$

where the formulae for $\nu_{\varphi,k}^{(j_{1:|\varphi|},i)}$, $\mathbf{m}_{\mathbf{x},k}^{(j_{1:|\varphi|},i)}$, $\mathbf{P}_{22,k}^{(j_{1:|\varphi|},i)}$ are given in equations (4.26)–(4.32).

$\bar{\jmath} := 0.$

for $j = 1, \ldots, J_{k|k-1}^{(i)}$ (*Missed detected individual targets*)

$\bar{\jmath} := \bar{\jmath} + 1, \quad \omega_k^{(\bar{\jmath})} := (1 - p_{D,2})\,\omega_{k|k-1}^{(j)},$

$\mathbf{m}_{\boldsymbol{\xi},k}^{(\bar{\jmath}|\bar{\imath})} := \mathbf{m}_{\boldsymbol{\xi},k}^{(j_{1:|\varphi|},i,j)}, \quad \mathbf{P}_{11,k}^{(\bar{\jmath}|\bar{\imath})} := \mathbf{P}_{11,k}^{(j_{1:|\varphi|},i,j)}, \quad \mathbf{P}_{12,k}^{(\bar{\jmath}|\bar{\imath})} := \mathbf{P}_{12,k}^{(j_{1:|\varphi|},i,j)},$

where the formulae for $\mathbf{m}_{\boldsymbol{\xi},k}^{(j_{1:|\varphi|},i,j)}$, $\mathbf{P}_{11,k}^{(j_{1:|\varphi|},i,j)}$, $\mathbf{P}_{12,k}^{(j_{1:|\varphi|},i,j)}$ are given in equations (4.33)–(4.34).

end

for $\ell = 1, \ldots, |\varphi|$ (*Detected individual targets*)

for $j_\ell = 1, \ldots, J_{k|k-1}^{(i)}$

$\bar{\jmath} := \bar{\jmath} + 1,$

$\mathbf{m}_{\boldsymbol{\xi},k}^{(\bar{\jmath}|\bar{\imath})} := \mathbf{m}_{\boldsymbol{\xi},k}^{(j_{1:|\varphi|},i,\bar{\jmath})}, \quad \mathbf{P}_{11,k}^{(\bar{\jmath}|\bar{\imath})} := \mathbf{P}_{11,k}^{(j_{1:|\varphi|},i,\bar{\jmath})}, \quad \mathbf{P}_{12,k}^{(\bar{\jmath}|\bar{\imath})} := \mathbf{P}_{12,k}^{(j_{1:|\varphi|},i,\bar{\jmath})},$

where the formulae for $\mathbf{m}_{\boldsymbol{\xi},k}^{(j_{1:|\varphi|},i,\bar{\jmath})}$, $\mathbf{P}_{11,k}^{(j_{1:|\varphi|},i,\bar{\jmath})}$, $\mathbf{P}_{12,k}^{(j_{1:|\varphi|},i,\bar{\jmath})}$ are given in equations (4.35)–(4.37).

Each weight $\omega_k^{(\bar{\jmath})}$ is formulated as per equation (4.38).

end

end

$J_k^{(\bar{\imath})} := \bar{\jmath}.$

end

end

end

Each partition weight ω_π is formulated as per equation (4.39).

end

$J_k := \bar{\imath}.$

output $\left\{\nu_k^{(i)}, \left\{\omega_k^{(j)}, \boldsymbol{\mu}_k^{(i,j)}, \boldsymbol{\Sigma}_k^{(i,j)}\right\}_{j=1}^{J_k^{(i)}}\right\}_{i=1}^{J_k}$

Table 4.2: Pseudo-code for the multi-group GM-PHD update

That is

$$N_k = \left(1 - p_{D,1} + p_{D,1}\,\mathrm{e}^{-p_{D,2}\,\alpha}\right) N_{k|k-1} + \sum_{\pi\in\Pi_Z} \varpi_\pi \sum_{\varphi\in\pi} \sum_{i=1}^{J_{k|k-1}} \sum_{j_{1:|\varphi|}} \nu_{\varphi,k}^{(j_{1:|\varphi|},i)}, \tag{4.42}$$

and for $\bar{\imath} = 1, \ldots, J_k$

$$N_k^{(\bar{\imath})} = \begin{cases} N_{k|k-1}^{(i)} + (1 - p_{D,2})\, N_{k|k-1}^{(i)} & \text{for } \bar{\imath} = i = 1, \ldots, J_{k|k-1}, \\ (1 - p_{D,2})\, N_{k|k-1}^{(i)} + p_{D,2} \sum\limits_{z\in\varphi_r} \sum\limits_{\bar{\jmath}=1}^{J_{k|k-1}^{(i)}} \omega_k^{(\bar{\jmath})} & \text{for } \bar{\imath} = J_{k|k-1} + J_k^{(t,r,i)} + j, \\ & \;\&\; j = 1, \ldots, |\varphi_r| \times J_{k|k-1}^{(i)}, \end{cases} \tag{4.43}$$

where, for $t = 1, \ldots, |\Pi_Z|$, $r = 1, \ldots, |\pi_t|$ and $i = 1, \ldots, J_{k|k-1}$,

$$J_k^{(t,r,i)} = \left(\left(r + |\pi|_{1:t-1}^{\times} - 1\right) \times J_{k|k-1} \times \prod_{\ell=1}^{r-1} \left(|\varphi_\ell| \times \prod_{i'=1}^{J_{k|k-1}} J_{k|k-1}^{(i')}\right)\right) + \left(|\varphi_r| \times J_{k|k-1}^{(1:i-1)}\right), \tag{4.44}$$

denoting $|\pi|_{1:t-1}^{\times} = |\pi_1| \times \cdots \times |\pi_{t-1}|$ and $J_{k|k-1}^{(1:i-1)} = J_{k|k-1}^{(1)} \times \cdots \times J_{k|k-1}^{(i-1)}$ (such that $|\pi|_{1:t-1}^{\times} = 0$ and $J_{k|k-1}^{(1:i-1)} = 0$ when $t = 1$ and $i = 1$ respectively).

It can be deduced from equations (4.40) and (4.42) that the GM-PHD filter for multiple group targets, at any time-step k, requires

$$J_k = J_{k|k-1} + \prod_{\pi\in\Pi_Z} \prod_{\varphi\in\pi} \prod_{\bar{\imath}=1}^{J_{k|k-1}} |\varphi| \times J_{k|k-1}^{(\bar{\imath})}, \tag{4.45}$$

Gaussian components relating to group states in v_k, where $J_{k|k-1} = J_{\gamma,k} + J_{k-1}$ is the number of components relating to group states in $v_{k|k-1}$ and the number of components relating to individual targets in $v_{k|k-1}$ is given by

$$J_{k|k-1}^{(\bar{\imath})} = \begin{cases} J_{\gamma,k}^{(i)} & \text{for } \bar{\imath} = i, \ldots, J_{\gamma,k}, \\ J_{\gamma,k}^{(i)} + J_{k-1}^{(i)} & \text{for } \bar{\imath} = J_{\gamma,k} + i \text{ and } i = 1, \ldots, J_{k-1}, \end{cases} \tag{4.46}$$

given that the number of components relating to group and individual target states in v_{k-1} are J_{k-1} and $J_{k-1}^{(i)}$ respectively. It is clear from equation (4.45) that the number of Gaussian components (relating to group states alone) in the posterior intensities increases without bound over time. The next section focuses on methods for reducing the number of Gaussian components and presents various procedures with this in mind.

4.2 Mixture Reduction Procedures

This section addresses implementation issues with the multi-group GM-PHD filter, in particular the potential for exponential growth in the number of Gaussian components at the update step.

4.2.1 Reducing the number of partitions of the measurement set

The number of partitions of a measurement set contributes significantly to the increase in the number of Gaussian components in the posterior intensities. As established in the previous chapter (Section 3.3), the operation of partitioning the measurement presents a substantial computational challenge. To address the partitioning problem a gating procedure is considered which reduces the number of partitions to $|\Pi_Z| = 1$ as follows.

Based on the validation or gating technique [64], the pre-update procedure, whose details are given in Table 4.3, assigns measurements to subsets of a single partition $\boldsymbol{\pi}$. Each subset $\varphi \in \boldsymbol{\pi}$ consists of those measurements in Z_k which lie within a

---

given $\left\{\nu_{k|k-1}^{(i)}, \mathbf{m}_{\mathbf{x},k|k-1}^{(i)}, \mathbf{P}_{22,k|k-1}^{(i)}\right\}_{i=1}^{J_{k|k-1}}$, measurement set Z_k, projection matrix $\widehat{\mathbf{H}}$, noise covariance matrix $\mathbf{R}_G \propto \mathbf{R}_2$ and gating threshold T_2.

procedure

Set $\boldsymbol{\pi} = \emptyset$, $n = 0$ and $I = \{i = 1, \ldots, J_{k|k-1}\}$.

repeat

$i_{\max} := \arg\max_{i \in I} \nu_{k|k-1}^{(i)}$

$\boldsymbol{\mu} := \widehat{\mathbf{H}}\mathbf{m}_{\mathbf{x},k|k-1}^{(i_{\max})}, \quad \boldsymbol{\Sigma} = \widehat{\mathbf{H}}\mathbf{P}_{22,k|k-1}^{(i_{\max})}\widehat{\mathbf{H}}^T + \mathbf{R}_G,$

$W := \left\{\mathbf{z} \in Z_k \,\middle|\, (\mathbf{z} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{z} - \boldsymbol{\mu}) \leq T_2 \right\}.$

if $W \neq \emptyset$

$n := n + 1,$

$\varphi_n := W,$

$\boldsymbol{\pi} := \{\boldsymbol{\pi}, \varphi_n\},$

$Z_k := Z_k \setminus \varphi_n.$

end

$I := I \setminus \{i_{\max}\}.$

until $I = \emptyset$.

for each remaining measurement $\mathbf{z} \in Z_k$

$n := n + 1,$

$\varphi_n := \{\mathbf{z}\}.$

end

output $\boldsymbol{\pi} = \{\varphi_i\}_{i=1}^{n}$.

Table 4.3: Gating procedure for partitioning the measurement set

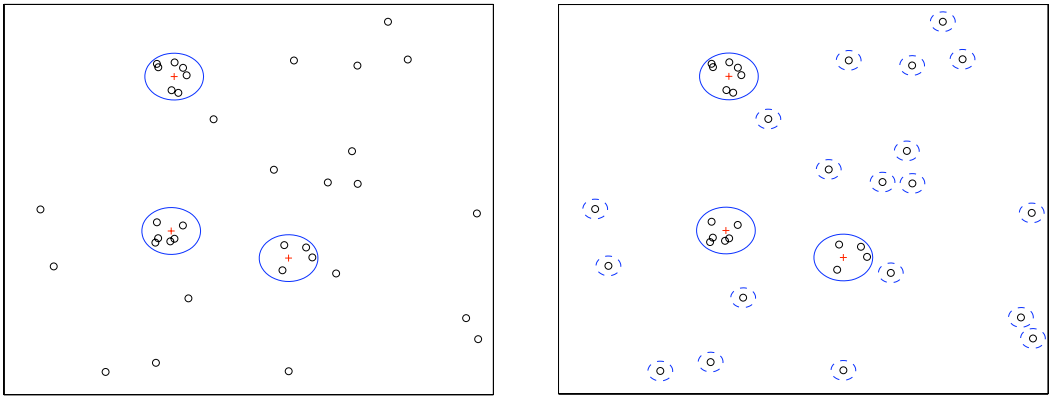


Figure 4.1: A measurement set partitioned according to the gating procedure in Table 4.3. Observations are indicated by black 'circles', predicted group target estimates are indicated by red 'crosses' and gates are indicated by blue 'lines'.

validation region for each predicted group target estimate. The validation or gating regions are defined by the Gaussian ellipses from $\mathcal{N}(\mathbf{z}\,;\widehat{\mathbf{H}}\mathbf{m}^{(i)}_{\mathbf{x},k|k-1},\widehat{\mathbf{H}}\mathbf{P}^{(i)}_{22,k|k-1}\widehat{\mathbf{H}}^T + \mathbf{R}_G)$ for $i = 1,\ldots,J_{k|k-1}$.[1] Any remaining measurements in Z_k are each individually assigned to a subset in $\boldsymbol{\pi}$, as illustrated in Figure 4.1. As a result, this procedure removes the summation over all partitions $\sum_{\pi\in\Pi_Z}$ in equations (4.23) and (4.42), leaving instead the following expressions for the updated intensity

$$v_k(\mathbf{x}_k,\boldsymbol{\xi}_k) = \left(1 - p_{D,1} + p_{D,1}\,\mathrm{e}^{-p_{D,2}\,\alpha}\,(1-p_{D,2})\right) v_{k|k-1}(\mathbf{x}_k,\boldsymbol{\xi}_k) + \sum_{\varphi\in\boldsymbol{\pi}} v_{D,k}(\mathbf{x}_k,\boldsymbol{\xi}_k)_{\varphi}, \tag{4.47}$$

and the expected number of group targets associated with v_k

$$N_k = \left(1 - p_{D,1} + p_{D,1}\,\mathrm{e}^{-p_{D,2}\,\alpha}\right) N_{k|k-1} + \sum_{\varphi\in\boldsymbol{\pi}} \sum_{i=1}^{J_{k|k-1}} \sum_{j_{1:|\varphi|}} \nu_{\varphi,k}^{(j_{1:|\varphi|},i)}. \tag{4.48}$$

4.2.2 Validating measurements

Besides the number of partitions of the measurement set, the other significant contributor of the increase in the number of Gaussian components is the pseudo multiple measurement likelihood L_φ. Another validation technique based procedure, whose details are given in Table 4.4, can be performed (during the update step) for each subset φ of a partition π, that validates measurements lying within a

[1]The Gaussian $\mathcal{N}(\mathbf{z}\,;\widehat{\mathbf{H}}\mathbf{m}^{(i)}_{\mathbf{x},k|k-1},\widehat{\mathbf{H}}\mathbf{P}^{(i)}_{22,k|k-1}\widehat{\mathbf{H}}^T + \mathbf{R}_G)$ is interpreted as the updated single measurement likelihood in relation to group targets, which is obtained by applying the standard result given in Lemma C.2 (Appendix C) to the product $\mathcal{N}(\mathbf{z}\,;\widehat{\mathbf{H}}\mathbf{x}_k,\mathbf{R}_G)\,\mathcal{N}(\mathbf{x}_k\,;\mathbf{m}^{(i)}_{\mathbf{x},k|k-1},\mathbf{P}^{(i)}_{22,k|k-1})$.

given $\left\{\left\{\mathbf{m}_{\boldsymbol{\xi},k|k-1}^{(j|i)}, \mathbf{P}_{11,k|k-1}^{(j|i)}, \mathbf{P}_{12,k|k-1}^{(j|i)}\right\}_{j=1}^{J_{k|k-1}^{(i)}}\right\}_{i=1}^{J_{k|k-1}}$, a partition π,
projection matrix $\mathbf{H}$, observation noise covariance matrix $\mathbf{R}_1$, and
gating threshold T_1.
procedure
$\tilde{\pi} := \pi$.
for each $\varphi \in \pi$
$J_{\varphi,k|k-1} := 0$.
for $i = 1, \ldots, J_{k|k-1}$
$n_\varphi := 0$.
for $\ell = 1, \ldots, |\varphi|$
$L_\ell := \emptyset$,
for $j = 1, \ldots, J_{k|k-1}^{(i)}$
Let $\boldsymbol{\mu} := \mathbf{H}\mathbf{m}_{\boldsymbol{\xi},k|k-1}^{(j|i)}$ and $\boldsymbol{\Sigma} := \mathbf{H}\mathbf{P}_{11,k|k-1}^{(j|i)}\mathbf{H}^T + \mathbf{R}_1$.
if $(\mathbf{z}_\ell - \boldsymbol{\mu})^T (\boldsymbol{\Sigma})^{-1} (\mathbf{z}_\ell - \boldsymbol{\mu}) \leq T_1$
$L_\ell := \{L_\ell, j\}$,
end
end
if $L_\ell \neq \emptyset$
$n_\varphi := n_\varphi + 1$, $\mathbf{z}_{n_\varphi} := \mathbf{z}_\ell$, $J_{n_\varphi}^{(i)} := |L_\ell|$.
for $\bar{\jmath} = 1, \ldots, J_{n_\varphi}^{(i)}$
$\mathbf{m}_{\boldsymbol{\xi}}^{(\bar{\jmath}|i)} = \mathbf{m}_{\boldsymbol{\xi},k|k-1}^{(L_{\ell,\bar{\jmath}}|i)}$,
$\mathbf{P}_{11}^{(\bar{\jmath}|i)} = \mathbf{P}_{11,k|k-1}^{(L_{\ell,\bar{\jmath}}|i)}$, $\mathbf{P}_{12}^{(\bar{\jmath}|i)} = \mathbf{P}_{12,k|k-1}^{(L_{\ell,\bar{\jmath}}|i)}$,
(*where* $L_{\ell,\bar{\jmath}}$ *denotes the* $\bar{\jmath}^{th}$ *entry in the set* L_ℓ)
end
end
end
if $n_\varphi > 0$
$J_{\varphi,k|k-1} := J_{\varphi,k|k-1} + 1$.
end
end
if $J_{\varphi,k|k-1} = 0$
$\tilde{\pi} := \tilde{\pi} \setminus \varphi$.
end
end
output $\left\{\left\{\left\{\mathbf{m}_{\boldsymbol{\xi}}^{(\bar{\jmath}|i)}, \mathbf{P}_{11}^{(\bar{\jmath}|i)}, \mathbf{P}_{12}^{(\bar{\jmath}|i)}\right\}_{\bar{\jmath}=1}^{J_l^{(i)}}\right\}_{l=1}^{n_\varphi}\right\}_{i=1}^{J_{\varphi,k|k-1}}$ for each $\varphi \in \tilde{\pi}$.

Table 4.4: Gating procedure for validating measurement in each subset φ of a partition π

region defined by Gaussian ellipses from $\mathcal{N}(\mathbf{z}\,;\mathbf{H}\mathbf{m}^{(j|i)}_{\boldsymbol{\xi},k|k-1},\mathbf{H}\mathbf{P}^{(j|i)}_{11,k|k-1}\mathbf{H}^T+\mathbf{R}_1)$ for $i=1,\ldots,J_{k|k-1}$ and $j=1,\ldots,J^{(i)}_{k|k-1}$.[2] To start with let $\tilde{\pi}=\pi$, then those subsets $\bar{\varphi}\in\pi$ whose constituent measurements do not lie within any validation region for all $i=1,\ldots,J_{k|k-1}$ and $j=1,\ldots,J^{(i)}_{k|k-1}$ are stricken from the partition $\tilde{\pi}$. For each remaining subset $\varphi\in\tilde{\pi}$, the procedure not only reduces the number of Gaussian components relating to group states in the expression corresponding to detections $v_{D,k}(\mathbf{x}_k,\boldsymbol{\xi}_k)$ in equation (4.24) to $J_{\varphi,k|k-1}\leq J_{k|k-1}$, but also reduces the summation over permutations $j_{1:|\varphi|}$ in equation (4.24) to $\sum_{j_{1:n_\varphi}}$ where $j_{1:n_\varphi}=(j_1,\ldots,j_{n_\varphi})$ denotes permutations of the indices $j_l=1,\ldots,J^{(i)}_l$ for $l=1,\ldots,n_\varphi$, such that $J^{(i)}_l\leq J^{(i)}_{k|k-1}$ and $n_\varphi\leq|\varphi|$. As a result, the expression corresponding to detections in equation (4.47) becomes

$$v_{D,k}(\mathbf{x}_k,\boldsymbol{\xi}_k)_\varphi=\sum_{i=1}^{J_{\varphi,k|k-1}}\sum_{j_{1:n_\varphi}}\nu^{(j_{1:n_\varphi},i)}_{\varphi,k}\sum_{\bar{\jmath}=1}^{\tilde{J}^{(i)}_{n_\varphi,k}}\omega^{(\bar{\jmath})}_k\mathcal{N}\left(\mathbf{X}_k\,;\boldsymbol{\mu}^{(j_{1:n_\varphi},i,\bar{\jmath})}_k,\boldsymbol{\Sigma}^{(j_{1:n_\varphi},i,\bar{\jmath})}_k\right),\quad(4.49)$$

where $\tilde{J}^{(i)}_{n_\varphi,k}=J^{(i)}_{k|k-1}+\prod_{l=1}^{n_\varphi}J^{(i)}_l$. In addition to performing the gating procedure in Table 4.3, the expected number of groups associated with v_k given by equation (4.47) is now

$$N_k=\left(1-p_{D,1}+p_{D,1}\,\mathrm{e}^{-p_{D,2}\,\alpha}\right)N_{k|k-1}+\sum_{\varphi\in\tilde{\boldsymbol{\pi}}}\sum_{i=1}^{J_{\varphi,k|k-1}}\sum_{j_{1:n_\varphi}}\nu^{(j_{1:n_\varphi},i)}_{\varphi,k}.\quad(4.50)$$

Consequently, the multi-group GM-PHD filter, incorporating both gating procedures (Tables 4.3 and 4.4) at every time-step k, now requires

$$\tilde{J}_k=J_{k|k-1}+\prod_{\varphi\in\tilde{\boldsymbol{\pi}}}\prod_{i=1}^{J_{\varphi,k|k-1}}\prod_{l=1}^{n_\varphi}J^{(i)}_l,\quad(4.51)$$

Gaussian components relating to group targets in v_k and for $\bar{\imath}=1,\ldots,\tilde{J}_k$

$$\tilde{J}^{(\bar{\imath})}_k=\begin{cases}J^{(i)}_{k|k-1} & \text{for } \bar{\imath}=i=1,\ldots,J_{k|k-1},\\ \tilde{J}^{(i)}_{n_\varphi,k} & \text{for } \bar{\imath}=J_{k|k-1}+J^{(r,i)}_k+j \text{ and } j=1,\ldots,J^{(i)}_{n_\varphi,k},\end{cases}\quad(4.52)$$

[2]The Gaussian $\mathcal{N}(\mathbf{z}\,;\mathbf{H}\mathbf{m}^{(j|i)}_{\boldsymbol{\xi},k|k-1},\mathbf{H}\mathbf{P}^{(j|i)}_{11,k|k-1}\mathbf{H}^T+\mathbf{R}_1)$ can be interpreted as the updated single measurement obtained by applying the standard result given in Lemma C.2 (Appendix C) to the product $\mathcal{N}(\mathbf{z}\,;\mathbf{H}\boldsymbol{\xi}_k,\mathbf{R}_1)\,\mathcal{N}(\boldsymbol{\xi}_k\,;\mathbf{m}^{(j|i)}_{\boldsymbol{\xi},k|k-1},\mathbf{P}^{(j|i)}_{11,k|k-1})$.

where $J_{n_\varphi,k}^{(i)} = \prod_{l=1}^{n_\varphi} J_l^{(i)}$ and

$$J_k^{(r,i)} = \prod_{\ell=1}^{r-1}\left(\prod_{i'=1}^{J_{\varphi_\ell,k|k-1}}\left(\prod_{l=1}^{n_{\varphi_\ell}} J_l^{(i')}\right)\right) + \prod_{i'=1}^{i}\left(\prod_{l=1}^{n_{\varphi_r}} J_l^{(i')}\right), \tag{4.53}$$

for $r = 1, \ldots, |\tilde{\boldsymbol{\pi}}|$ and $i = 1, \ldots, J_{\varphi_r,k|k-1}$.

4.2.3 State extraction

Having reduced the number of Gaussian components, the next task is to extract the group and individual target state estimates, which is straightforward since the means of the remaining Gaussian components are local maxima of v_k, provided that they are reasonably well separated. Closely spaced Gaussian components can be merged, as shown in Table 4.5, based on the pruning algorithm proposed in [90]. Gaussian components are judged to be closely spaced or not based on the Mahalanobis distance measure [125], given appropriate thresholds[3].

The Mahalanobis distance is usually a suitable distance measure for Gaussian functions whose covariances are identical. However, the covariances of the Gaussian components in v_k, prior to merging, may be very different and the use of the Mahalanobis distance could then lead to the merging of such Gaussian components, resulting in less than accurate representations of the group and individual target states. To avoid this, it is perhaps worth considering an alternative measure of distance which takes into account differing covariances.

The Hellinger distance [126], for example, quantifies the difference between two probability distributions, say f and g, given by the following expression

$$d_H(f,g) = \sqrt{1 - \rho_B(f,g)}, \tag{4.54}$$

where $\rho_B(f,g)$ denotes the Bhatacharyya coefficient [126] given by

$$\rho_B(f,g) = \int \sqrt{f(\mathbf{x})\, g(\mathbf{x})}\, \mathrm{d}\mathbf{x}. \tag{4.55}$$

Unlike the Mahalanobis distance, the Hellinger distance is constrained to $0 \leq d_H \leq 1$, where the maximum distance is achieved when the two distributions do not overlap. For two Gaussian distributions $f(\mathbf{x}) = \mathcal{N}(\mathbf{x}\,; \mathbf{m}_1, \mathbf{P}_1)$ and $g(\mathbf{x}) =$

[3]See [64] for further explanation on threshold values.

given $\left\{\nu_k^{(i)}, \mathbf{m}_{\mathbf{x},k}^{(i)}, \mathbf{P}_{22,k}^{(i)}, \left\{\omega_k^{(j)}, \mathbf{m}_{\boldsymbol{\xi},k}^{(j|i)}, \mathbf{P}_{11,k}^{(j|i)}, \mathbf{P}_{12,k}^{(j|i)}\right\}_{j=1}^{\tilde{J}_k^{(i)}}\right\}_{i=1}^{\tilde{J}_k}$, truncation thresholds τ_1, τ_2, merging thresholds U_1, U_2 and maximum allowable number of Gaussian components. relating to group and individual targets respectively $J_{\max,1}$, $J_{\max,2}$.

Set $\bar{\imath} = 0$, and $I_{\mathbf{x},k} = \left\{ i = 1, \ldots, \tilde{J}_k \middle| \nu_k^{(i)} > \tau_1 \right\}$.

repeat (*Merging of parent components*)

$\bar{\imath} := \bar{\imath} + 1, \quad i_{\max} := \arg\max_{i \in I_{\mathbf{x},k}} \nu_k^{(i)}$,

$\bar{I}_{\mathbf{x},k} := \left\{ i \in I_{\mathbf{x},k} \middle| \left(\mathbf{m}_{\mathbf{x},k}^{(i)} - \mathbf{m}_{\mathbf{x},k}^{(i_{\max})}\right)^T \left(\mathbf{P}_{22,k}^{(i)}\right)^{-1} \left(\mathbf{m}_{\mathbf{x},k}^{(i)} - \mathbf{m}_{\mathbf{x},k}^{(i_{\max})}\right) \leq U_1 \right\}$,

$\tilde{\nu}_k^{(\bar{\imath})} := \sum_{i \in \bar{I}_{\mathbf{x},k}} \nu_k^{(i)}, \quad \widetilde{\mathbf{m}}_{\mathbf{x},k}^{(\bar{\imath})} := \frac{1}{\tilde{\nu}_k^{(\bar{\imath})}} \sum_{i \in \bar{I}_{\mathbf{x},k}} \nu_k^{(i)} \mathbf{m}_{\mathbf{x},k}^{(i)}, \quad \widetilde{\mathbf{P}}_{22,k}^{(\bar{\imath})} := \frac{1}{\tilde{\nu}_k^{(\bar{\imath})}} \sum_{i \in \bar{I}_{\mathbf{x},k}} \nu_k^{(i)} \mathbf{P}_{22,k}^{(i)}$,

$\ell := 0$.

for all $i \in \bar{I}_{\mathbf{x},k}$

for $j = 1, \ldots, \tilde{J}_k^{(i)}$

$\ell := \ell + 1, \quad \omega_k^{(\ell)} := \omega_k^{(j)}$,

$\mathbf{m}_{\boldsymbol{\xi},k}^{(\ell|\bar{\imath})} := \mathbf{m}_{\boldsymbol{\xi},k}^{(j|i)} + \mathbf{P}_{12,k}^{(\ell|\bar{\imath})} \left(\widetilde{\mathbf{P}}_{22,k}^{(\bar{\imath})}\right)^{-1} \left(\widetilde{\mathbf{m}}_{\mathbf{x},k}^{(\ell)} - \mathbf{m}_{\mathbf{x},k}^{(i)}\right)$,

$\mathbf{P}_{11,k}^{(\ell|\bar{\imath})} := \mathbf{P}_{11,k}^{(j|i)} - \mathbf{P}_{12,k}^{(\ell|\bar{\imath})} \left(\widetilde{\mathbf{P}}_{22,k}^{(\bar{\imath})}\right)^{-1} \left(\mathbf{P}_{12,k}^{(j|i)} - \mathbf{P}_{12,k}^{(\ell|\bar{\imath})}\right)^T$,

$\mathbf{P}_{12,k}^{(\ell|\bar{\imath})} := \mathbf{P}_{12,k}^{(j|i)} \left(\mathbf{P}_{22,k}^{(i)}\right)^{-1} \widetilde{\mathbf{P}}_{22,k}^{(\bar{\imath})}$.

end

end

$\tilde{J}_{\text{new},k}^{(\bar{\imath})} := \ell$.

Set $\bar{\jmath} = 0$, and $I_{\boldsymbol{\xi},k} = \left\{ j = 1, \ldots, \tilde{J}_{\text{new},k}^{(\bar{\imath})} \middle| \omega_k^{(j)} > \tau_2 \right\}$.

repeat (*Merging of daughter components*)

$\bar{\jmath} := \bar{\jmath} + 1, \quad j_{\max} := \arg\max_{j \in I_{\boldsymbol{\xi},k}} \omega_k^{(j)}$,

$\bar{I}_{\boldsymbol{\xi},k} := \left\{ j \in I_{\boldsymbol{\xi},k} \middle| \left(\mathbf{m}_{\boldsymbol{\xi},k}^{(j|\bar{\imath})} - \mathbf{m}_{\boldsymbol{\xi},k}^{(j_{\max}|\bar{\imath})}\right)^T \left(\mathbf{P}_{11,k}^{(j|\bar{\imath})}\right)^{-1} \left(\mathbf{m}_{\boldsymbol{\xi},k}^{(j|\bar{\imath})} - \mathbf{m}_{\boldsymbol{\xi},k}^{(j_{\max}|\bar{\imath})}\right) \leq U_2 \right\}$,

$\widetilde{\omega}_k^{(\bar{\jmath})} := \sum_{j \in \bar{I}_{\boldsymbol{\xi},k}} \omega_k^{(j)}, \quad \widetilde{\mathbf{m}}_{\boldsymbol{\xi},k}^{(\bar{\jmath}|\bar{\imath})} := \frac{1}{\widetilde{\omega}_k^{(\ell)}} \sum_{j \in \bar{I}_{\boldsymbol{\xi},k}} \omega_k^{(j)} \mathbf{m}_{\boldsymbol{\xi},k}^{(j|\bar{\imath})}$,

$\widetilde{\mathbf{P}}_{11,k}^{(\bar{\jmath}|\bar{\imath})} := \frac{1}{\widetilde{\omega}_k^{(\ell)}} \sum_{j \in \bar{I}_{\boldsymbol{\xi},k}} \omega_k^{(j)} \mathbf{P}_{11,k}^{(j|\bar{\imath})}, \quad \widetilde{\mathbf{P}}_{12,k}^{(\bar{\jmath}|\bar{\imath})} := \frac{1}{\widetilde{\omega}_k^{(\ell)}} \sum_{j \in \bar{I}_{\boldsymbol{\xi},k}} \omega_k^{(j)} \mathbf{P}_{12,k}^{(j|\bar{\imath})}$,

$I_{\boldsymbol{\xi},k} := I_{\boldsymbol{\xi},k} \setminus \bar{I}_{\boldsymbol{\xi},k}$.

until $I_{\boldsymbol{\xi},k} = \emptyset$.

$\tilde{J}_{\text{new},k}^{(\bar{\imath})} := \bar{\jmath}, \quad I_{\mathbf{x},k} := I_{\mathbf{x},k} \setminus \bar{I}_{\mathbf{x},k}$.

until $I_{\mathbf{x},k} = \emptyset$.

If $\bar{\imath} > J_{\max,1}$ then replace $\left\{\tilde{\nu}_k^{(i)}, \widetilde{\mathbf{m}}_{\mathbf{x},k}^{(i)}, \widetilde{\mathbf{P}}_{22,k}^{(i)}\right\}_{i=1}^{\bar{\imath}}$ by those $J_{\max,1}$ Gaussian components with the largest weights, setting $\tilde{J}_{\text{new},k} = J_{\max,1}$ (otherwise set $\tilde{J}_{\text{new},k} := \bar{\imath}$), and if, for each $i = 1, \ldots, \tilde{J}_{\text{new},k}$, $\tilde{J}_{\text{new},k}^{(i)} > J_{\max,2}$ then replace $\left\{\tilde{\omega}_k^{(j)}, \widetilde{\mathbf{m}}_{\boldsymbol{\xi},k}^{(j)}, \widetilde{\mathbf{P}}_{11,k}^{(j)}, \widetilde{\mathbf{P}}_{12,k}^{(j)}\right\}_{j=1}^{\tilde{J}_{\text{new},k}^{(i)}}$ by those $J_{\max,2}$ Gaussian components with the largest weights, re-setting $\tilde{J}_{\text{new},k}^{(i)} = J_{\max,2}$.

output $\left\{\tilde{\nu}_k^{(i)}, \widetilde{\mathbf{m}}_{\mathbf{x},k}^{(i)}, \widetilde{\mathbf{P}}_{22,k}^{(i)} \left\{\widetilde{\omega}_k^{(j)}, \widetilde{\mathbf{m}}_{\boldsymbol{\xi},k}^{(j|i)}, \widetilde{\mathbf{P}}_{11,k}^{(j|i)}, \widetilde{\mathbf{P}}_{12,k}^{(j|i)}\right\}_{j=1}^{\tilde{J}_{\text{new},k}^{(i)}}\right\}_{i=1}^{\tilde{J}_{\text{new},k}}$.

Table 4.5: Merging closely spaced Gaussian components

$\mathcal{N}(\mathbf{x}\,;\mathbf{m}_2,\mathbf{P}_2)$, the Bhatacharyya coefficient admits a closed form solution given by

$$\rho_B(f,g)=\sqrt{\frac{\sqrt{\det(\mathbf{P}_1\,\mathbf{P}_2)}}{\det\left(\frac{1}{2}(\mathbf{P}_1+\mathbf{P}_2)\right)}}\exp\left\{-\frac{1}{4}(\mathbf{m}_1-\mathbf{m}_2)^T(\mathbf{P}_1+\mathbf{P}_2)^{-1}(\mathbf{m}_1-\mathbf{m}_2)\right\}, \tag{4.56}$$

from which the Hellinger distance immediately follows. Consequently, using the Hellinger distance rather than the Mahalanobis distance, the set of indices $\bar{I}_{\mathbf{x},k}$ in Table 4.5 is now set to be

$$\bar{I}_{\mathbf{x},k}:=\left\{i\in I_{\mathbf{x},k}\;\middle|\;d_H\left(\mathcal{N}\left(\mathbf{x}\,;\mathbf{m}_{\mathbf{x},k}^{(i)},\mathbf{P}_{22,k}^{(i)}\right),f\right)\leq U_1\right\}, \tag{4.57}$$

where $f(\mathbf{x})=\mathcal{N}\left(\mathbf{x}\,;\mathbf{m}_{\mathbf{x},k}^{(i_{\max})},\mathbf{P}_{22,k}^{(i_{\max})}\right)$ and $0<U_1<1$ (likewise for $\bar{I}_{\boldsymbol{\xi},k}$).

Depending on the value set for thresholds U_1, U_2, merging of Gaussian components, using the Hellinger distance, may be minimal, potentially leaving a number of Gaussian components with low weights. Therefore to extract strong state estimates, it is desirable to select the means of those Gaussian components with weights that are greater than some appropriate thresholds $\nu_{\mathrm{Th}},\omega_{\mathrm{Th}}>0$, as shown in Table 4.6, whether or not merging has been performed.

---

given $\left\{\nu_k^{(i)},\mathbf{m}_{\mathbf{x},k}^{(i)},\mathbf{P}_{22,k}^{(i)},\left\{\omega_k^{(j)},\mathbf{m}_{\boldsymbol{\xi},k}^{(j|i)},\mathbf{P}_{11,k}^{(j|i)},\mathbf{P}_{12,k}^{(j|i)}\right\}_{j=1}^{J_k^{(i)}}\right\}_{i=1}^{J_k}$ and thresholds ν_{Th}, ω_{Th}.
Set $\tilde{X}_k=\emptyset$, and $\ell=0$.
for $i=1,\ldots,J_k$
  if $\nu_k^{(i)}>\nu_{\mathrm{Th}}$
    $\ell:=\ell+1,\quad \tilde{X}_k:=\left\{\tilde{X}_k,\mathbf{m}_{\mathbf{x},k}^{(i)}\right\},\quad \tilde{\Xi}_k^{(\ell)}:=\emptyset.$
    for $j=1,\ldots,J_k^{(i)}$
      if $\omega_k^{(j)}>\omega_{\mathrm{Th}}$
        $\tilde{\Xi}_k^{(\ell)}:=\left\{\tilde{\Xi}_k^{(\ell)},\mathbf{m}_{\boldsymbol{\xi},k}^{(j|i)}\right\}.$
      end
    end
  end
end
output $\tilde{X}_k$ as the multiple group target state estimate and
  $\tilde{\Xi}_k^{(\bar{\imath})}$ as the multiple individual target state estimate associated with each state estimate in $\tilde{X}_k$, for $\bar{\imath}=1,\ldots,\ell$.

Table 4.6: Group and individual target state extractions

4.3 The Single Group GM-PHD Filter

This section concerns the implementation of a similar Gaussian mixture formulation for the single group PHD filter. An overview of this formulation is presented in Section 4.3.1, which is applied to the simulated example described in Section 4.3.2. The simulation results are then presented in Section 4.3.3.

4.3.1 The single group GM-PHD recursion

Under linear Gaussian assumptions on the dynamic and measurement model, precisely those given in Sections 4.1.2 and 4.1.4 for the multi-group GM-PHD filter (with $p_{S,1} = 1$ and $p_S = p_{S,2}$), a closed-form solution to the single group PHD filter exists that formulates the predicted and posterior intensities as Gaussian mixture as illustrated in the following diagram:

$$
\begin{array}{lcl}
 & \vdots & \vdots \\
 & \downarrow & \downarrow \\
v_{k-1}(\mathbf{x}_{k-1}, \boldsymbol{\xi}_{k-1}) & = & \displaystyle\sum_{i=1}^{J_{k-1}} \omega_{k-1}^{(i)} \mathcal{N}\left(\mathbf{X}_{k-1}\,; \boldsymbol{\mu}_{k-1}^{(i)}, \boldsymbol{\Sigma}_{k-1}^{(i)}\right) \\
 & \downarrow & \downarrow \\
\text{Prediction:} \quad v_{k|k-1}(\mathbf{x}_k, \boldsymbol{\xi}_k) & = & \displaystyle\sum_{i=1}^{J_{k|k-1}} \omega_{k|k-1}^{(i)} \mathcal{N}\left(\mathbf{X}_{k}\,; \boldsymbol{\mu}_{k|k-1}^{(i)}, \boldsymbol{\Sigma}_{k|k-1}^{(i)}\right) \\
 & \downarrow & \downarrow \\
\text{Update:} \quad v_{k}(\mathbf{x}_k, \boldsymbol{\xi}_k) & = & \displaystyle\sum_{i=1}^{J_{k}} \omega_{k}^{(i)} \mathcal{N}\left(\mathbf{X}_{k}\,; \boldsymbol{\mu}_{k}^{(i)}, \boldsymbol{\Sigma}_{k}^{(i)}\right) \\
 & \downarrow & \downarrow \\
 & \vdots & \vdots
\end{array}
$$

The components in each of the mixtures formulating the predicted and posterior intensities at each iteration are joint Gaussian functions with state variable $\mathbf{X}_k = \begin{bmatrix} \boldsymbol{\xi}_k & \mathbf{x}_k \end{bmatrix}^T$, with respective means given by

$$
\boldsymbol{\mu}_{k|k-1}^{(i)} = \begin{bmatrix} \mathbf{m}_{\boldsymbol{\xi},k|k-1}^{(i)} & \mathbf{m}_{\mathbf{x},k|k-1} \end{bmatrix}^T, \qquad \boldsymbol{\mu}_{k}^{(i)} = \begin{bmatrix} \mathbf{m}_{\boldsymbol{\xi},k}^{(i)} & \mathbf{m}_{\mathbf{x},k} \end{bmatrix}^T, \tag{4.58}
$$

and with respective covariances given by the following block matrices

$$
\boldsymbol{\Sigma}_{k|k-1}^{(i)} = \begin{bmatrix} \mathbf{P}_{11,k|k-1}^{(i)} & \mathbf{P}_{12,k|k-1}^{(i)} \\ \left(\mathbf{P}_{12,k|k-1}^{(i)}\right)^T & \mathbf{P}_{22,k|k-1} \end{bmatrix}, \qquad \boldsymbol{\Sigma}_{k}^{(i)} = \begin{bmatrix} \mathbf{P}_{11,k}^{(i)} & \mathbf{P}_{12,k}^{(i)} \\ \left(\mathbf{P}_{12,k}^{(i)}\right)^T & \mathbf{P}_{22,k} \end{bmatrix}, \tag{4.59}
$$

as denoted for time-step k. Further details on the closed-form expressions for computing these means and covariances, along with the respective weights, in $v_{k|k-1}$ and v_k are given in Appendix D (Section D.1).

4.3.2 Simulation set-up

The single group target scenario considered for demonstrating the implementation of the single group GM-PHD filter, using simulated data, consists of 10 individual targets whose positions do not change relative to the centroid defining the group state. The group target is observed via measurements that are generated by its constituent targets over a surveillance region $[0\ 500]\times[0\ 500]$ and its trajectory is shown in Figure 4.2.

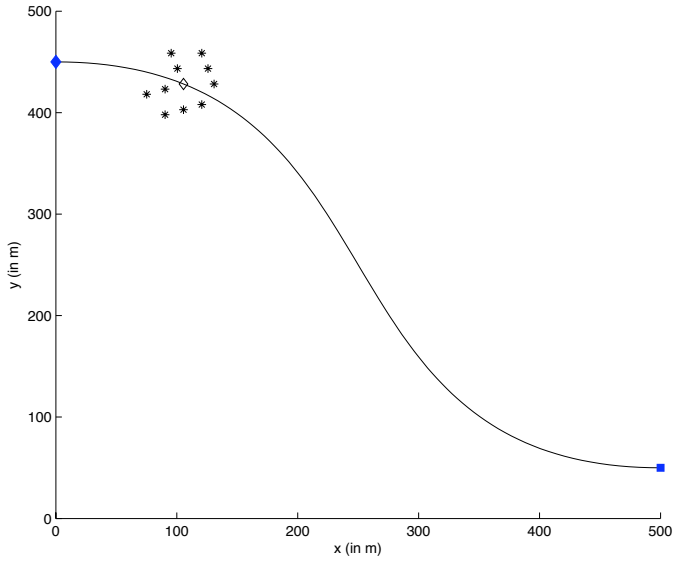


Figure 4.2: The group target trajectory (line) with its starting position at time $k = 1$ indicated by the blue 'filled diamond' and its terminal position at time $k = 100$ indicated by the blue 'filled square'. The group target state position and the individual target state positions at time $k = 15$ are indicated by the 'unfilled diamond' and the 'asterisks' respectively.

The group target state $\mathbf{x}_k = \begin{bmatrix} x_{k,1} & x_{k,2} & x_{k,3} & x_{k,4} \end{bmatrix}^T$ consists of two-dimensional Cartesian co-ordinate position vector $\begin{bmatrix} x_{k,1} & x_{k,3} \end{bmatrix}^T$ and velocity vector $\begin{bmatrix} x_{k,2} & x_{k,4} \end{bmatrix}^T$ at time-step k. It follows the linear Gaussian dynamic model given in equation (4.8) with

$$\mathbf{F}_{22} = \begin{bmatrix} \widehat{\mathbf{F}}_{k-1} & \mathbf{0}_2 \\ \mathbf{0}_2 & \widehat{\mathbf{F}}_{k-1} \end{bmatrix}, \qquad \mathbf{Q}_{22} = \sigma_{\mathbf{x}}^2 \begin{bmatrix} \widehat{\mathbf{Q}}_{k-1} & \mathbf{0}_2 \\ \mathbf{0}_2 & \widehat{\mathbf{Q}}_{k-1} \end{bmatrix}, \tag{4.60}$$

such that, for sampling period $t = 1\,\mathrm{s}$,

$$\widehat{\mathbf{F}}_{k-1} = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}, \qquad \widehat{\mathbf{Q}}_{k-1} = \begin{bmatrix} t^3/3 & t^2/2 \\ t^2/2 & t \end{bmatrix}, \tag{4.61}$$

where $\mathbf{0}_2$ denotes the 2×2 zero matrix and $\sigma_{\mathbf{x}}^2 = \sqrt{0.8}\,\mathrm{m/s^2}$ is the standard deviation of the parent process noise.

Each individual target state $\boldsymbol{\xi}_k = \begin{bmatrix} \xi_{k,1} & \xi_{k,2} \end{bmatrix}^T$ is strictly a two-dimensional Cartesian co-ordinate position vector. They each follow the linear Gaussian dynamic model given in equation (4.9) with $\mathbf{F}_{11} = \mathbf{I}_2$, $\mathbf{F}_{12} = \widehat{\mathbf{H}}(\mathbf{F}_{22} - \mathbf{I}_4)$, where $\mathbf{I}_2$ and $\mathbf{I}_4$ denote the 2×2 and 4×4 identity matrices respectively, and for $t = 1\,\mathrm{s}$,

$$\mathbf{Q}_{11} = \sigma_{\mathbf{x}}^2 \frac{1}{3} t^3\, \mathbf{I}_2, \qquad \mathbf{Q}_{12} = \sigma_{\mathbf{x}}^2 \begin{bmatrix} \begin{bmatrix} t^3/3 & t^2/2 \end{bmatrix} & \mathbf{0}_{2\times1} \\ \mathbf{0}_{2\times1} & \begin{bmatrix} t^3/3 & t^2/2 \end{bmatrix} \end{bmatrix}, \tag{4.62}$$

where $\mathbf{0}_{2\times1}$ denotes the 2×1 zero matrix. (Appendix D.4 gives further details on the expressions for the noise covariance matrices $\mathbf{Q}_{11}$, $\mathbf{Q}_{12}$ and $\mathbf{Q}_{22}$.)

The Cartesian point measurements $\mathbf{z} = \begin{bmatrix} z_1 & z_2 \end{bmatrix}^T \in Z_k$ relate to individual targets $\boldsymbol{\xi}_k$ according to the likelihood $\mathcal{N}(\mathbf{z}\,; \mathbf{H}\boldsymbol{\xi}_k, \mathbf{R}_1)$ in equation (4.15) with $\mathbf{H} = \mathbf{I}_2$ and $\mathbf{R}_1 = \sigma_1^2\, \mathbf{I}_2$ where $\sigma_1 = \sqrt{10}\,\mathrm{m}$. They relate to the single group target $\mathbf{x}_k$ according to the likelihood $\mathcal{N}(\mathbf{z}\,; \widehat{\mathbf{H}}\mathbf{x}_k, \mathbf{R}_2)$ in equation (4.15) with

$$\widehat{\mathbf{H}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \tag{4.63}$$

and $\mathbf{R}_2 = \sigma_2^2\, \mathbf{I}_2$ where $\sigma_2 = \sqrt{44}\,\mathrm{m}$. The measurement set Z_k includes false alarms which are modelled as a Poisson point process with intensity $\kappa(\mathbf{z}) = \lambda\, U(\mathbf{z})$ where $U(\cdot)$ is the uniform density over the surveillance region and λ denotes the average number of false alarms over this area. Figure 4.3 shows simulated measurements plotted with the individual target trajectories against time with $\lambda = 50$. At each iteration, the simulated measurements undergo two stages of validation, similar to the procedures shown in Tables 4.3 (Section 4.2.1) and 4.4 (Section 4.2.2). The precise details of the first of these validation procedures, illustrated in Figure 4.4, are given in Appendix D (Section D.2).

The filter is initialised at time-step $k = 1$ with the intensity $v_1(\mathbf{x}, \boldsymbol{\xi})$, which is formulated as the following Gaussian mixture

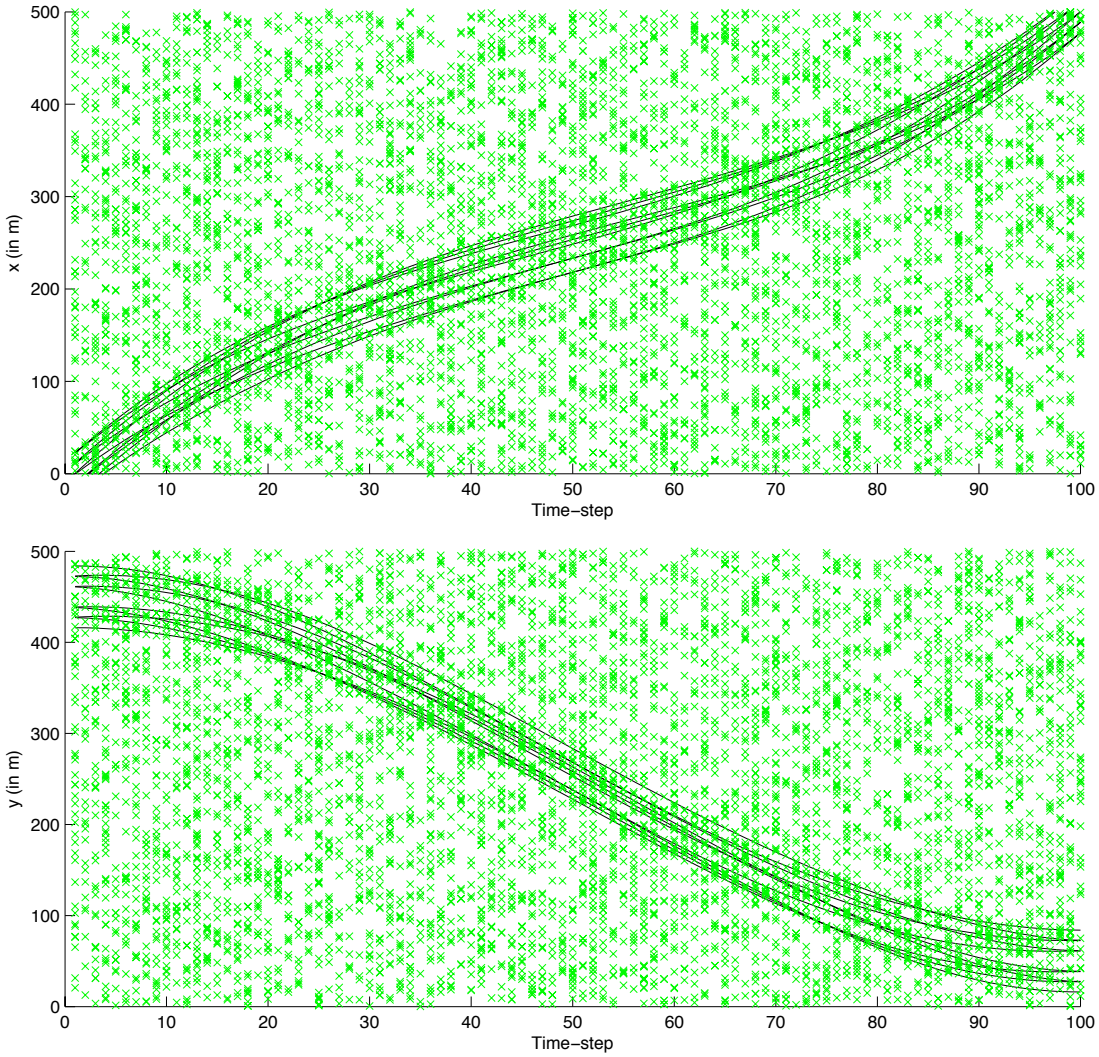


Figure 4.3: Measurements (green '×') and individual target trajectories (lines).

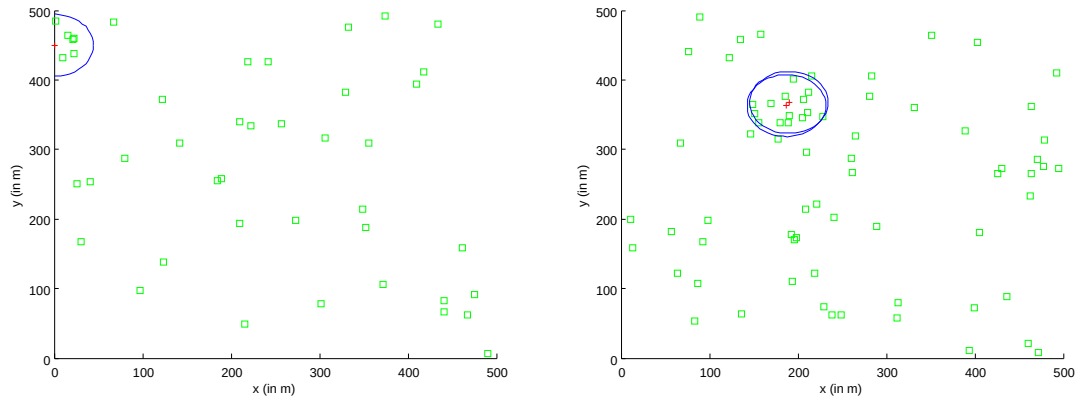


Figure 4.4: Illustrating the use of the validation procedure detailed in Section D.2 on the simulated measurements shown in Figure 4.3 at time-steps $k = 1$ (left) and $k = 31$ (right)

$$v_1(\mathbf{x}, \boldsymbol{\xi}) = \sum_{i=1}^{J_1} \omega^{(i)} \mathcal{N}\left(\begin{bmatrix} \boldsymbol{\xi} \\ \mathbf{x} \end{bmatrix} ; \begin{bmatrix} \mathbf{m}_{\boldsymbol{\xi}}^{(i)} \\ \mathbf{m}_{\mathbf{x}} \end{bmatrix}, \begin{bmatrix} \mathbf{P}_{11}^{(i)} & \mathbf{P}_{12}^{(i)} \\ \left(\mathbf{P}_{12}^{(i)}\right)^T & \mathbf{P}_{12} \end{bmatrix} \right). \tag{4.64}$$

Full details on the initialisation of the weights $\omega^{(i)}$, means $\mathbf{m}_{\boldsymbol{\xi}}^{(i)}$, $\mathbf{m}_{\mathbf{x}}$ and covariances $\mathbf{P}_{11}^{(i)}$, $\mathbf{P}_{12}^{(i)}$, $\mathbf{P}_{22}$ in the above mixture are given in Appendix D (Section D.3).

The birth of newly appearing individual targets in the group at time-steps $k > 1$ are modelled by the following Gaussian mixture

$$\gamma_{\boldsymbol{\xi},k}(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k) = \sum_{i=1}^{J_{\gamma,k}} \omega_{\gamma,k}^{(i)} \mathcal{N}\left(\boldsymbol{\xi}_k \,;\, \boldsymbol{\mu}_{\gamma,\boldsymbol{\xi},k}^{(i)}, \boldsymbol{\Sigma}_{\gamma,\boldsymbol{\xi},k}^{(i)}\right), \tag{4.65}$$

with means and covariances respectively given by

$$\begin{aligned} \boldsymbol{\mu}_{\gamma,\boldsymbol{\xi},k}^{(i)} &= \mathbf{m}_{\gamma,\boldsymbol{\xi},k}^{(i)} + \mathbf{P}_{\gamma,12,k}^{(i)} \mathbf{P}_{22,k|k-1}^{-1} \left(\mathbf{x}_k - \mathbf{m}_{\mathbf{x},k|k-1}\right), \\ \boldsymbol{\Sigma}_{\gamma,\boldsymbol{\xi},k}^{(i)} &= \mathbf{P}_{\gamma,11,k}^{(i)} - \mathbf{P}_{\gamma,12,k}^{(i)} \mathbf{P}_{22,k|k-1}^{-1} \left(\mathbf{P}_{\gamma,12,k}^{(i)}\right)^T. \end{aligned} \tag{4.66}$$

The initialisation of the means $\mathbf{m}_{\gamma,\boldsymbol{\xi},k}^{(i)}$ and covariances $\mathbf{P}_{\gamma,11,k}^{(i)}$, $\mathbf{P}_{\gamma,12,k}^{(i)}$ in the above expressions are dealt with in a similar way to those for the initial intensity, full details of which are given in Appendix D (Section D.3).

4.3.3 Principal results

The single group GM-PHD filter is applied to the scenario described in Section 4.3.2 with probabilities of survival and detection $p_S = p_D = 0.9$ and clutter rate $\lambda = 50$. The mixture reduction procedures detailed in Appendix D (Section D.2) are applied with parameters $T_2 = 9$, $\mathbf{R}_G = 5 \times \mathbf{R}_2$ and $J_{\min} = 5$ (see Tables D.6 & D.7). An additional reduction technique, similar to the procedure shown in Table 4.4, is also applied with parameter $T_1 = 4$. Finally, closely spaced Gaussian components relating to the individual targets are merged, as shown in Table 4.5 for those corresponding components, using the Hellinger distance given by equations (4.54) and (4.56) and with parameters $\tau_2 = 10^{-2}$, $U_2 = 0.4$ (satisfying $0 \leq U_2 \leq 1$) and $J_{\max} = J_{\max,2} = 10$.

From the position estimates shown in Figure 4.5, it can be seen that the single group GM-PHD filter provides accurate estimation of this group formation as it evolves over time. The filter not only accurately estimates the position of the group target state, defined by the centroid of the formation, it successfully detects, on

average, 90% of the individual targets.

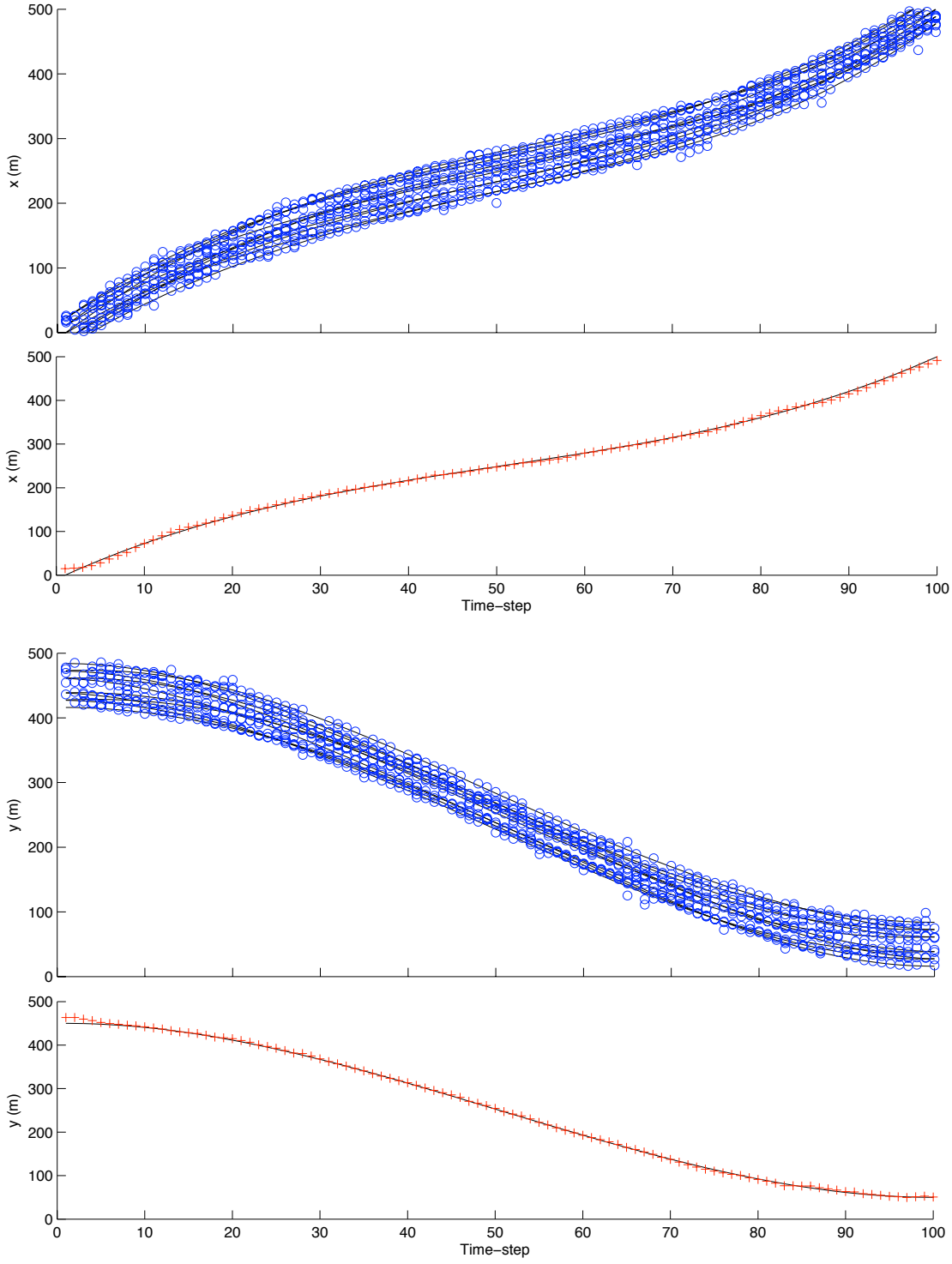


Figure 4.5: Position estimates from the single group GM-PHD filter (with $p_S = p_D = 0.9$) for the individual target states (blue 'circles') and for the group target state (red 'crosses')

The performance of the single group GM-PHD filter is now examined for this scenario, by calculating the RMSE between the estimated and true single group target state over time. It useful, for performance analysis of a filter, to benchmark error results against those of an alternative/existing method that is applicable to the same problem. For this purpose, the GM-PHD filter proposed in [90] for multiple target systems, referred to hereafter as the multi-target GM-PHD filter, will serve as an appropriate benchmark. The multi-target GM-PHD filter can be applied to the multiple constituent targets of the single group, whose state estimate can then be obtained by averaging the means of the Gaussian components in the updated intensity at each iteration.

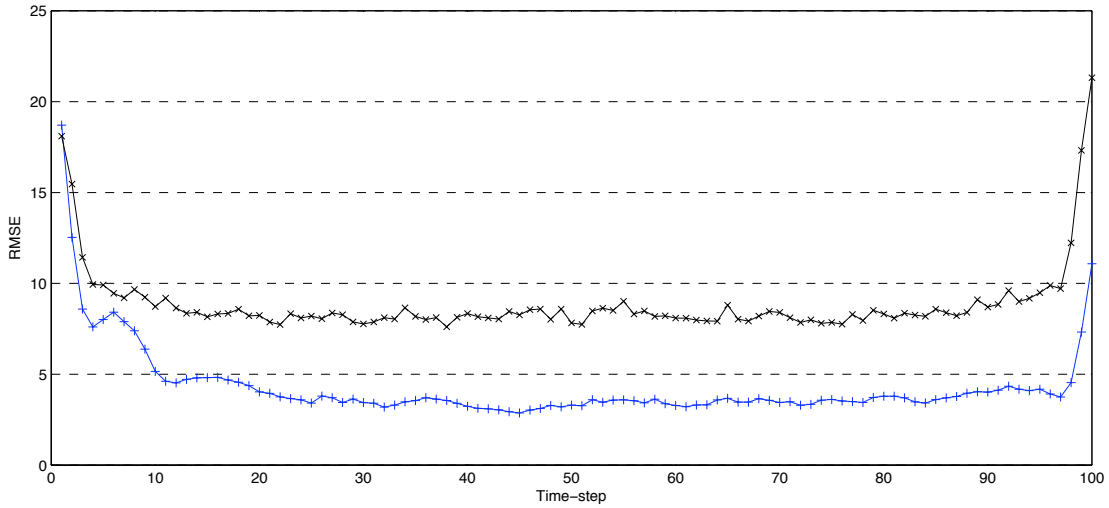


Figure 4.6: The mean errors (RMSE) for the single group target state over 50 MC runs of the multi-target GM-PHD filter (black line '$-\times-$') and the single group GM-PHD filter (blue line '$-+-$')

Figure 4.6 shows the calculated mean errors for 50 Monte Carlo simulations of both single group and multi-target GM-PHD filters. Note that high errors at the beginning and end of the simulations for both filters occur when constituent targets in the single group move beyond the observation region. As expected the results in Figure 4.6 show that the single group GM-PHD filter performs better than the multi-target GM-PHD filter in estimating a group target state, since the single group GM-PHD filter actually accounts for such a state in the cluster model representation of the group target system rather than simply averaging the multiple constituent target estimates.

Furthermore, the peformance of the multi-target GM-PHD filter is known to suffer when the parameter value for the probability of detection p_D is lowered [90]. Even with $p_D = 0.9$ the results in Figure 4.6 show that the mean errors of the group

target state for the mulit-target GM-PHD filter are higher than those for the single group GM-PHD filter, and lowering the probability of detection in this scenario would likely result in an increase of mean errors for both filters. But how much would the performance of the single group GM-PHD filter for this scenario suffer by lowering the probability of detection (or changing some of the other parameters)? This is the subject of further investigation in the following section.

4.3.4 Sensitivity studies

The single group GM-PHD filter is re-applied to the scenario described in Section 4.3.2 with various different parameter settings for the probabilities of survival/detection, clutter rate and validation threshold, in particular. The impact these changes to the parameter settings have on the performance of the filter for this scenario are then analysed based on the estimation results from the simulations.

Lower probabilities of survival and detection

Figure 4.7 shows the position estimate results for the group target state from the simulation run with $p_S = 0.9$ and $p_D = 0.7$. Figure 4.8 shows the position estimate results for the group target state from the simulation run with $p_S = 0.6$ and $p_D = 0.5$. All other parameter settings specified in Section 4.3.3 remain the same.

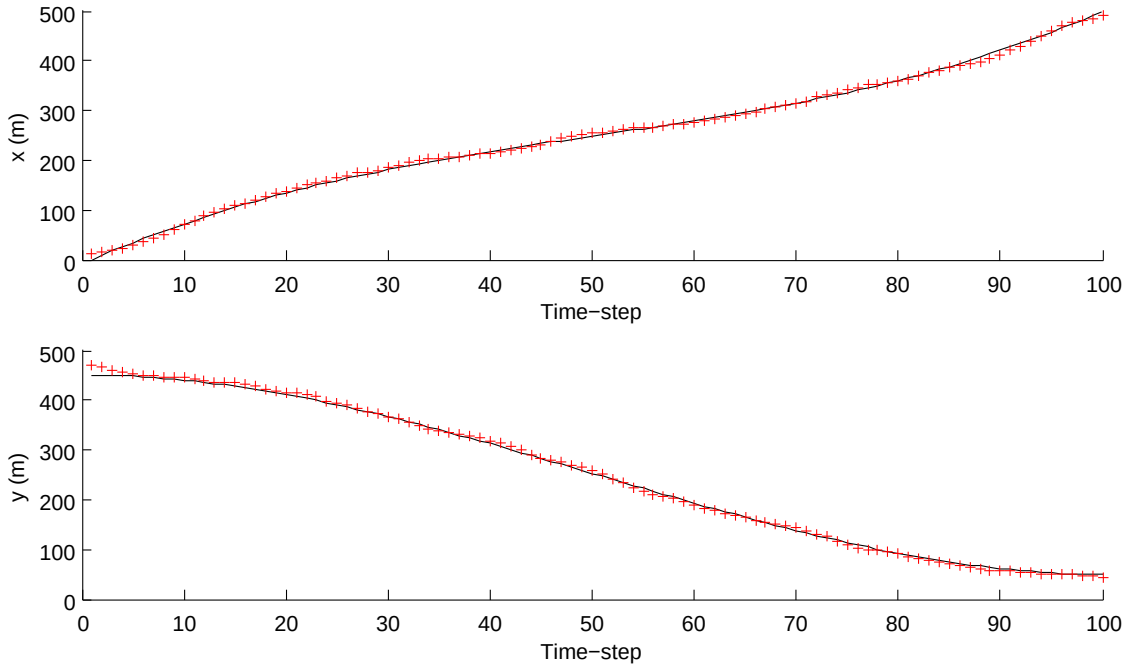


Figure 4.7: Position estimates from the single group GM-PHD filter, with $p_S = 0.9$ and $p_D = 0.7$, for the group target state (red 'crosses')

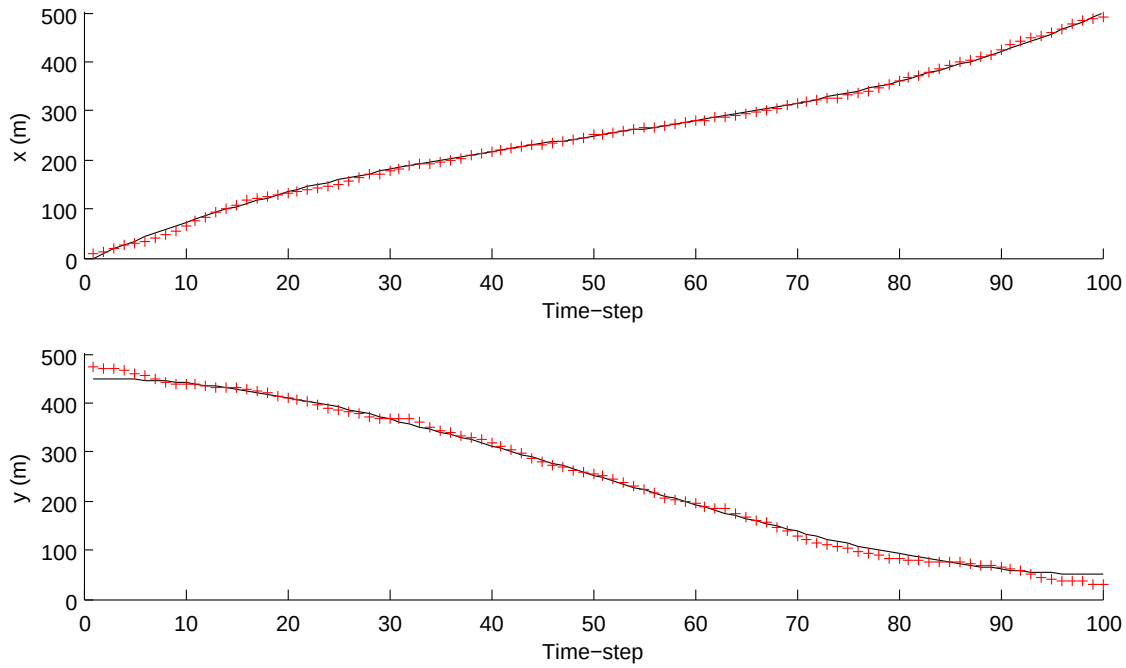


Figure 4.8: Position estimates from the single group GM-PHD filter, with $p_S = 0.6$ and $p_D = 0.5$, for the group target state (red 'crosses')

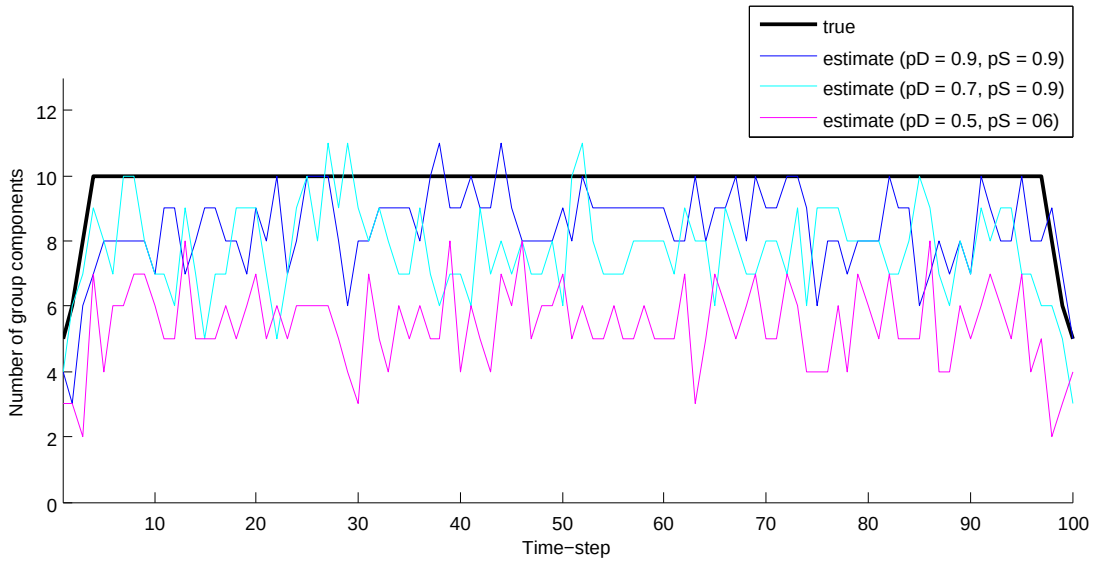


Figure 4.9: Cardinality estimates from each application of the single group GM-PHD filter with the different settings for the probabilities of survival/detection

From the results in Figures 4.7 & 4.8 it can be seen that lowering the probability of detection (and probability of survival) has a nominal affect on the accuracy of the group state estimation. However, it does have an impact on the individual target cardinality estimates as shown in Figure 4.9. The estimation accuracy in

the number of constituent targets in the group deteriorates as the probabilities of survival/detection are lowered. This gives an indication as to how poorly the multi-target GM-PHD filter would perform for this scenario with these parameter changes, since it relies on the multiple constituent target estimates to provide an estimate for the group target state.

Higher clutter rate

Figure 4.10 shows the position estimate results for the group target state from the simulation run with clutter rate $\lambda = 100$. Figure 4.11 shows the position estimate results for the group target state from the simulation run with clutter rate $\lambda = 250$. All other parameter settings specified in Section 4.3.3 remain the same. From the results in Figures 4.10 & 4.11 it can be seen that increasing the clutter rate has more of an impact on the accuracy of the group state estimation than lowering the probabilities of survival/detection.

The affect that increasing the clutter rate has on the individual target cardinality estimates is shown in Figure 4.12. In contrast to the cardinality results from the simulations with lower probability of survival/detection (Figure 4.9), increasing the clutter rate appears to improve the estimation accuracy in the number of constituent targets in the group. However these results are somewhat deceptive since the number of false estimates will increase by virtue of the increase in clut-

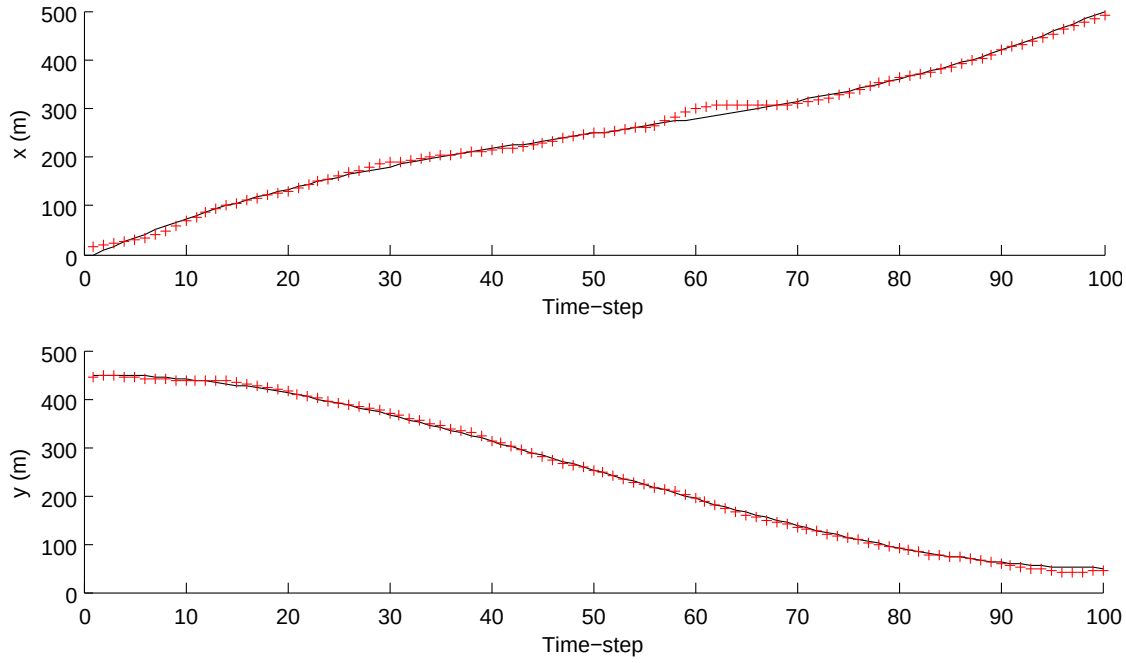


Figure 4.10: Position estimates from the single group GM-PHD filter, with $\lambda = 100$, for the group target state (red 'crosses')

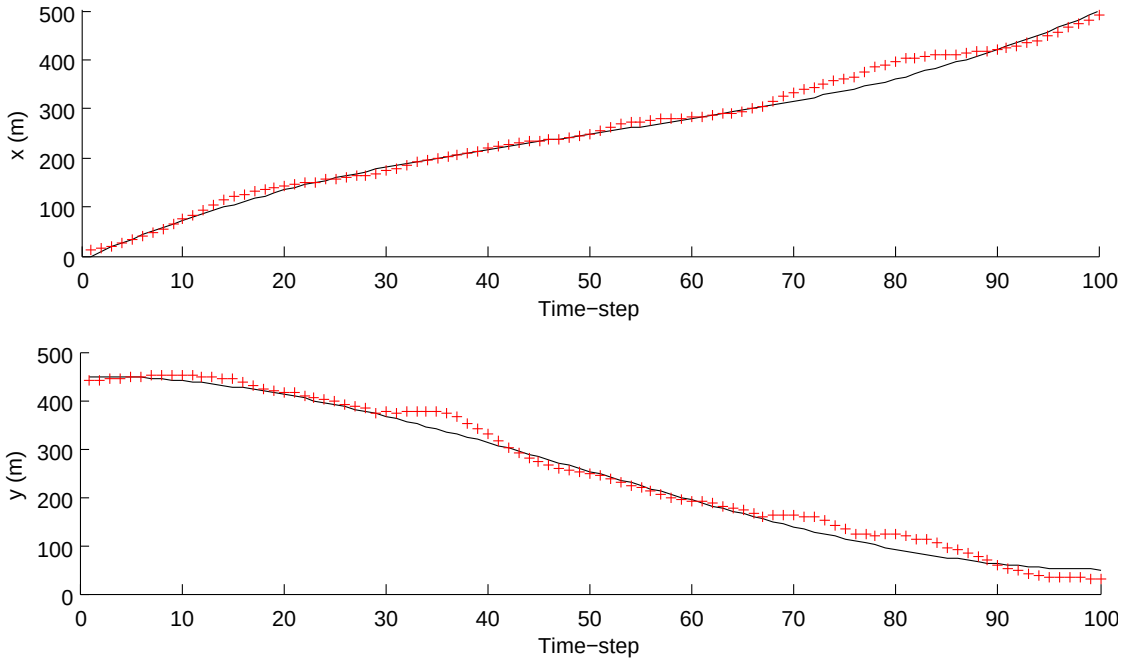


Figure 4.11: Position estimates from the single group GM-PHD filter, with $\lambda = 250$, for the group target state (red 'crosses')

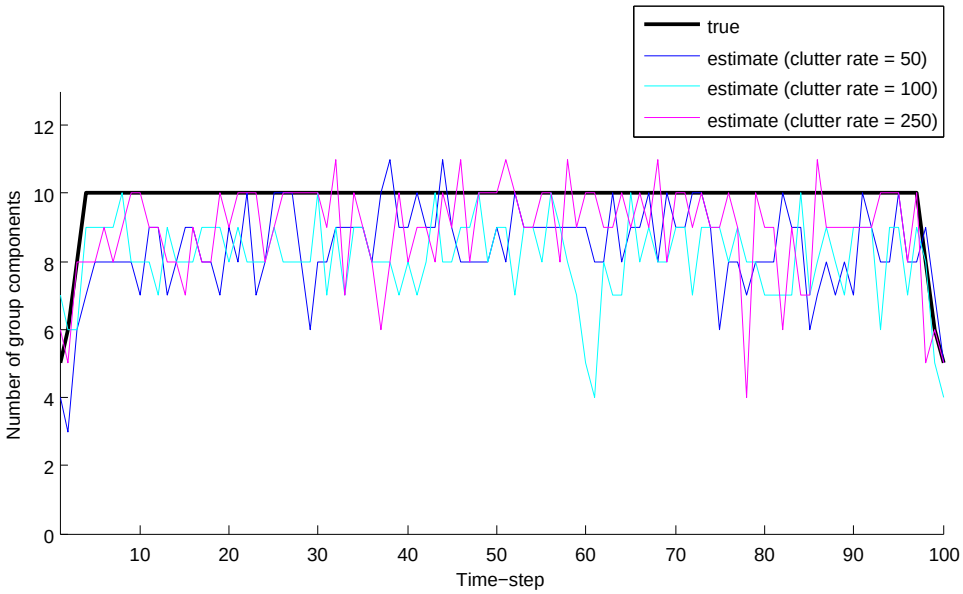


Figure 4.12: Cardinality estimates from each application of the single group GM-PHD filter with the different settings for the clutter rate

ter rate. In other words the apparent improvement in cardinality estimation due to the increase in clutter rate may just be a consequence of extracting more false estimates for the group's constituent targets. For instance, consider the simulated measurements shown in Figure 4.13. The higher the clutter rate the more likely

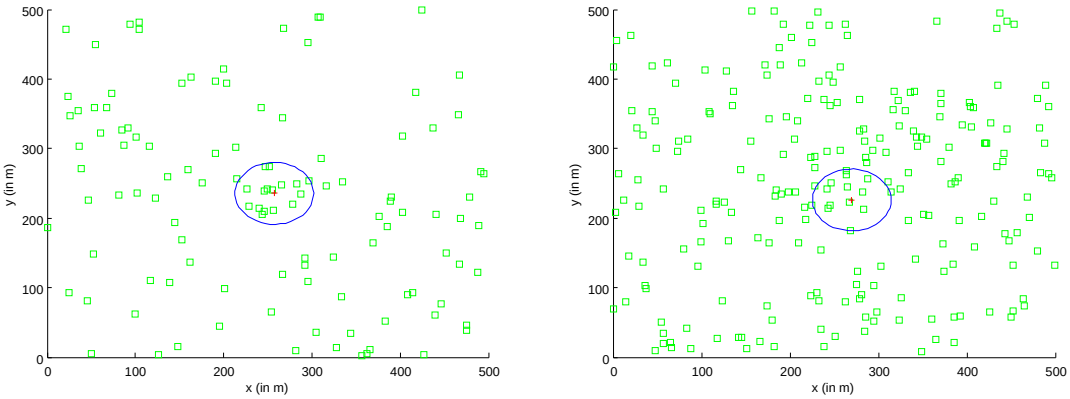


Figure 4.13: Examples of the simulated measurements and validation gates at time-step $k = 53$ with clutter rates $\lambda = 100$ (left) and $\lambda = 250$ (right)

false measurements will be validated in the gating procedure which inevitably leads to false estimates. Hence the multi-target GM-PHD filter approach to this scenario with these parameter changes would again unreliable.

Lower validation threshold

Figures 4.14 & 4.15 show the cardinality and position estimate results from the simulation run with the gating threshold in the validation procedure detailed in Appendix D (Section D.2) reduced to $T_2 = 4$ and with $J_{\min} = 2$. All other parameter settings specified in Section 4.3.3 remain the same.

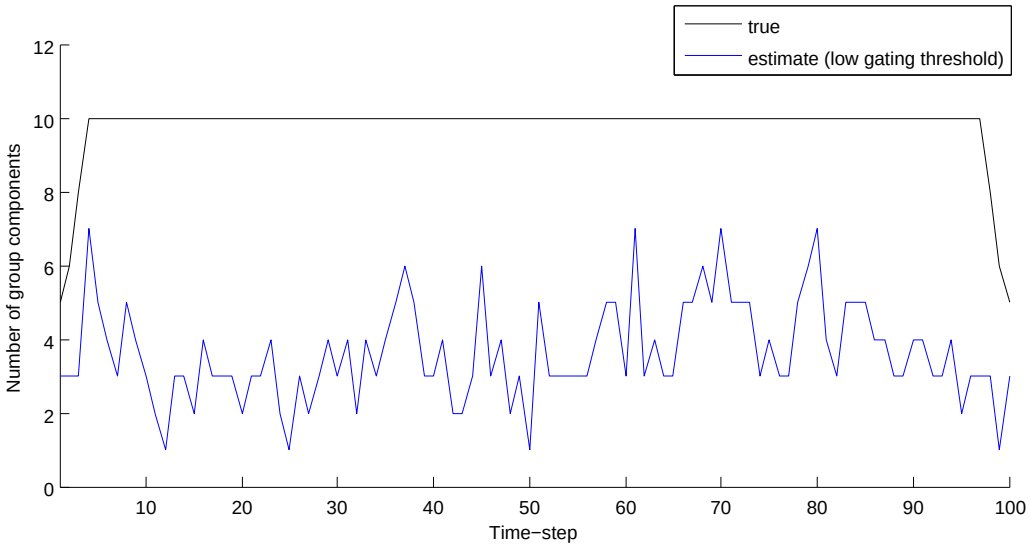


Figure 4.14: Cardinality estimates from the single group GM-PHD filter with lower gating threshold parameters

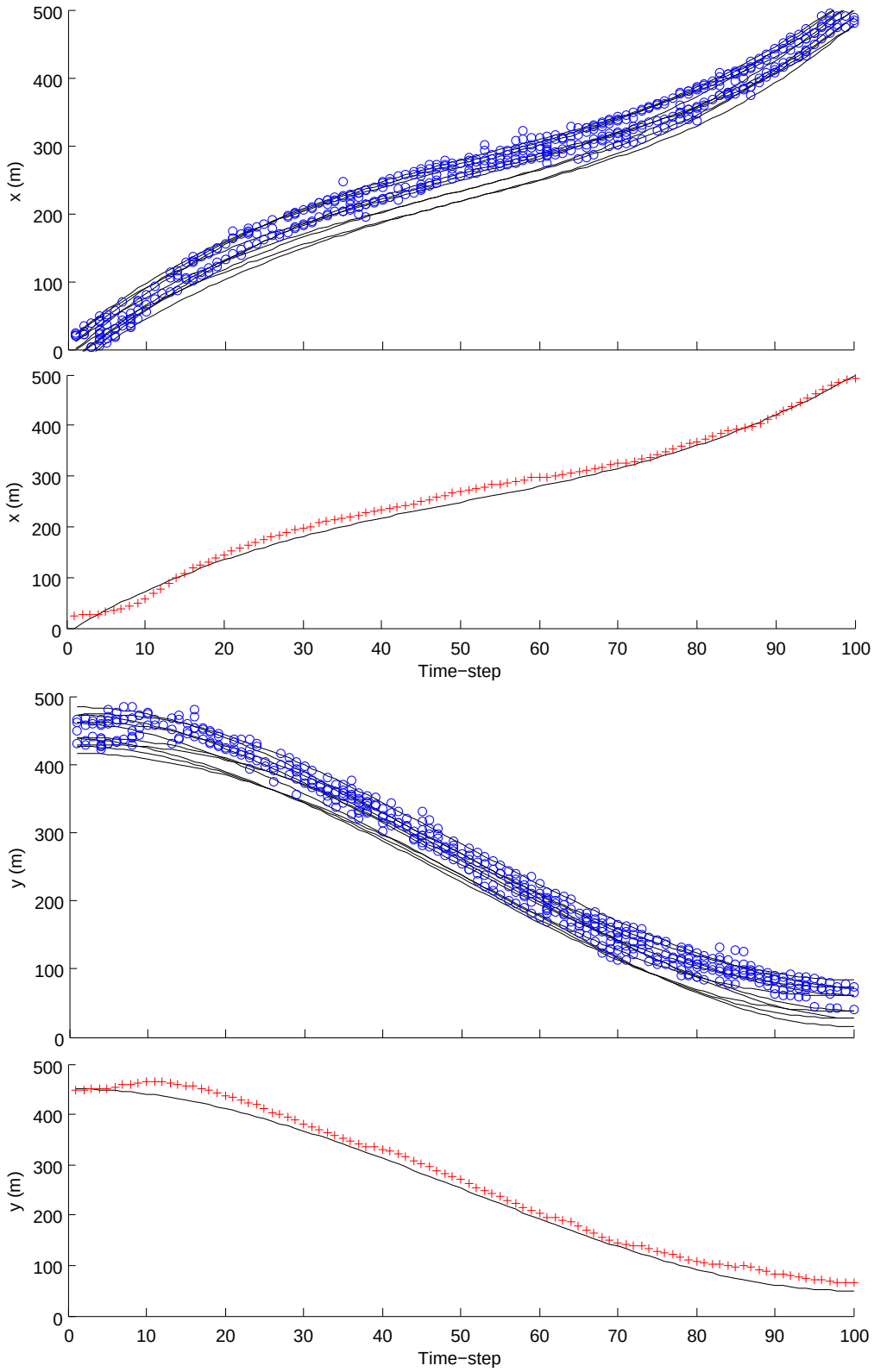


Figure 4.15: Position estimates from the single group GM-PHD filter, with lower gating threshold parameters, for the individual target states (blue 'circles') and for the group target state (red 'crosses')

From the results in Figures 4.14 & 4.15 it can be seen that lowering the gating threshold has a detrimental affect on the accuracy of the group and individual target state estimation as well as on the estimation accuracy in the number of constituent targets. By lowering the gating threshold the validation region is decreased, as can be seen when comparing Figure 4.16 below with Figures 4.4 and 4.13. As a consequence of decreased validation regions fewer potentially target generated measurements are validated, resulting in fewer individual target state estimates being extracted and less accurate group state estimates.

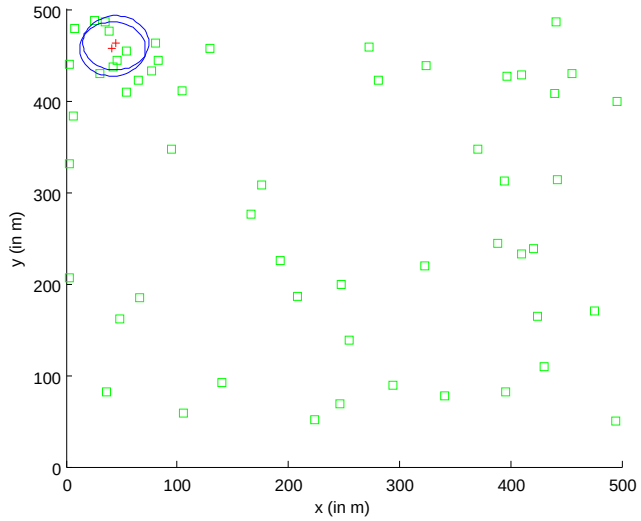


Figure 4.16: Simulated measurements and validation gates at time-step $k = 8$ with threshold $T_2 = 4$ and $J_{\min} = 2$ (see Table D.7)

In an attempt to compensate for decreased validation regions, the validation procedure allows for multiple gates to be considered. The number of validation gates is determined from the ratio of the number of individual target related Gaussian components at the previous time-step against the minimum allowable number of such Gaussian components and rounded to the nearest integer. Only validation gates containing measurements that haven't already been validated are kept for consideration. Figure 4.16 shows two such gates (out of a total possible number of two) from the set of simulated measurements at time-step $k = 8$. This is reason for reducing the parameter setting for the minimum allowable number of individual target related Gaussian components $J_{\min}$ along with lowering the gating threshold for this simulation. It increases the chances of multiple gates being considered in the validation procedure, which potentially expands the region for validation over multiple gates that are reduced in size. Without this compensation in the validation procedure the estimation results would potentially suffer more so than the

results shown in Figures 4.14 & 4.15, with lower gating thresholds. Full details of the algoirthm for this validation procedure are given in Table D.7 (Section D.2, Appendix D).

4.4 Discussion

This chapter presented a closed-form solution to the multi-group PHD filter under linear Gaussian conditions, called the multi-group GM-PHD filter. Adapted from this closed-form solution, the single group GM-PHD filter was applied to a single group formation scenario and tested for robustness against changes to certain parameter settings, in particular for the probabilities of survival/detection, clutter rate and validation threshold.

Now, rather than implement the full multi-group GM-PHD filter and demonstrate its application on a similar scenario, the focus will instead shift to the implementation of the multiple extended target special case, which is covered in the next chapter. For extended target tracking, the concern of the estimation problem is not necessarily on target kinematic states solely, but also on the size and shape of the targets, i.e. the target extents. Extent estimation has been the subject of great interest in tracking literature of late, as discussed in Chapter 1 (Section 1.1.3), and cluster models provide a novel approach to the problem. This aspect of the estimation problem can also apply to group target tracking. The size of group targets may be of interest along with the positions and velocities of groups and their constituent targets. However, to consider this aspect of the estimation problem for group target tracking using cluster models and PHD filtering would require jointly estimating some shape parametric state as well as two levels of kinematic states relating to each group and its constituent targets. Since the PHD filter for extended targets (Section 3.2.6) is a recursion that propagates the marginal group target PHD, there is no interest in the estimation of feature points under the cluster model representation, so the consideration of extent estimation is more feasible for this special case and will be further explored in Chapter 5.

Chapter 5

Implementations of the PHD filter for Extended Targets

It was reasoned in Chapter 3 that extended targets can be mathematically represented in the same way as group targets and that the corresponding PHD filter is a special case of the multi-group PHD filter. Firstly, the hierarchical RFS representation considered in Section 3.1.2, allows for the construction of a geometric shaped extent for each target, as illustrated in Figure 5.1, which is an appropriate consideration in certain scenarios where targets generate multiple point detections that appear in structured arrangements. The additional introduction of variable shape parameters then gives rise to the possibility of estimating the size of each target's extent.

So it is, that the purpose of this chapter is to demonstrate the PHD filter for extended targets presented in Section 3.2.6 with two extended target tracking scenarios. The first, considered in Section 5.2, is a multiple extended target scenario in which the size and shape of each target's extent is fixed, while in Section 5.3 a second multiple extended target scenario is considered with estimable extent shape

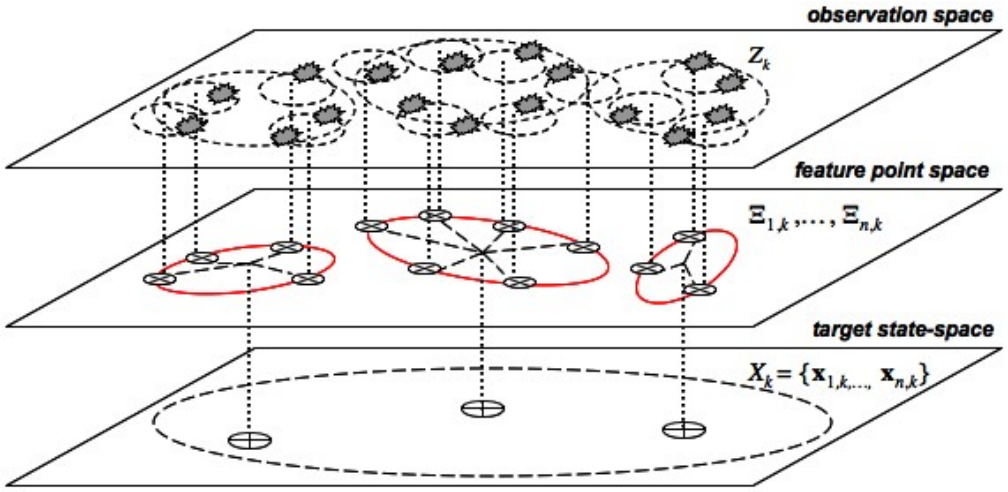


Figure 5.1: Visualisation of a multiple extended target tracking problem under the cluster model representation

parameters of varying size. Simulated examples are included for both scenarios. To start with though, a special case of the multi-group GM-PHD filter proposed in the previous chapter, is introduced in Section 5.1.

5.1 The GM-PHD Filter for Extended Targets

The PHD filter for extended targets presented in Section 3.2.6 is considered a special case of the multi-group PHD filter. It propagates the marginal group target PHD $v_k(\mathbf{x}_k)$ of the hierarchical RFS, which mathematically represents a multiple group/extended target system, while the multi-group PHD filter propagates the condensed group target PHD $v_k(\mathbf{x}_k, \boldsymbol{\xi}_k)$. It follows that a closed-form solution to the PHD filter for extended targets exists that formulates the predicted and posterior intensities as Gaussian mixture and is a special case of the multi-group GM-PHD filter as shown in Sections 5.1.1 and 5.1.2.

5.1.1 The recursion

$$
\begin{array}{ccc}
\vdots & & \vdots \\
\downarrow & & \downarrow \\
v_{k-1}(\mathbf{x}_{k-1}) & = & \displaystyle\sum_{i=1}^{J_{k-1}} \omega_{k-1}^{(i)} \mathcal{N}\left(\mathbf{x}_{k-1}\,; \mathbf{m}_{k-1}^{(i)}, \mathbf{P}_{k-1}^{(i)}\right) \\
| & & | \\
\text{prediction} & & \text{prediction} \\
\downarrow & & \downarrow \\
v_{k|k-1}(\mathbf{x}_k) & = & \displaystyle\sum_{i=1}^{J_{k|k-1}} \omega_{k|k-1}^{(i)} \mathcal{N}\left(\mathbf{x}_k\,; \mathbf{m}_{k|k-1}^{(i)}, \mathbf{P}_{k|k-1}^{(i)}\right) \\
| & & | \\
\text{update} & & \text{update} \\
\downarrow & & \downarrow \\
v_k(\mathbf{x}_k) & = & \displaystyle\sum_{i=1}^{J_k} \omega_k^{(i)} \mathcal{N}\left(\mathbf{x}_k\,; \mathbf{m}_k^{(i)}, \mathbf{P}_k^{(i)}\right) \\
\downarrow & & \downarrow \\
\vdots & & \vdots
\end{array}
$$

At each iteration of the multi-group GM-PHD filter the predicted and posterior intensities satisfy the respective factorisations: $v_{k|k-1}(\mathbf{x}_k, \boldsymbol{\xi}_k) = v_{k|k-1}(\mathbf{x}_k) \times v_{k|k-1}(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k)$ and $v_k(\mathbf{x}_k, \boldsymbol{\xi}_k) = v_k(\mathbf{x}_k) \times v_k(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k)$, given the formulations in Propo-

sitions 4.1 and 4.2. The GM-PHD recursion for extended targets is illustrated in the diagram on the previous page, where the predicted and updated intensities are formulated as mixtures of the Gaussian components relating to the state $\mathbf{x}_k$ in Propositions 4.1 and 4.2 respectively. The closed-form expression for the weights, means and covariances of the Gaussian components in these formulations are subject to slight changes where appropriate for the extended target tracking problem, the precise details of which are given in Section 5.1.2.

5.1.2 The algorithm

Suppose the dynamics relating to the state vector $\mathbf{x}_{k-1}$ is modelled with the linear Gaussian given in equation (4.8) (Section 4.1.2), instead denoting $\mathbf{F}_k = \mathbf{F}_{22}$ & $\mathbf{Q}_k = \mathbf{Q}_{22}$, and the birth model is formulated with the following Gaussian mixture

$$\gamma_k(\mathbf{x}_k) = \sum_{i=1}^{J_{\gamma,k}} \omega_{\gamma,k}^{(i)} \mathcal{N}\left(\mathbf{x}_k \,;\, \mathbf{m}_{\gamma,k}^{(i)}, \mathbf{P}_{\gamma,k}^{(i)}\right). \tag{5.1}$$

Then, given the prior intensity $v_{k-1}(\mathbf{x}_{k-1})$ is a Gaussian mixture, the **predicted intensity** $v_k(\mathbf{x}_k)$ is also a Gaussian mixture, each with the respective forms as illustrated in the diagram on the previous page. The closed-form expressions of the predicted weights, means and covariances are detailed in Table 5.1 and are the same as those relating to the group states in the multi-group GM-PHD prediction

---

given $\left\{\omega_{k-1}^{(i)}, \mathbf{m}_{k-1}^{(i)}, \mathbf{P}_{k-1}^{(i)}\right\}_{i=1}^{J_{k-1}}$, transition matrix $\mathbf{F}_{k-1}$ $(= \mathbf{F}_{22})$,
noise covariance matrix $\mathbf{Q}_{k-1}$ $(= \mathbf{Q}_{22})$, and probability of survival p_S $(= p_{S,1})$.
prediction
$\bar{\imath} := 0$.
for $i = 1, \ldots, J_{\gamma,k}$ (*i.e. for extended target births*)
$\bar{\imath} := \bar{\imath} + 1$,
$\omega_{k|k-1}^{(\bar{\imath})} := \omega_{\gamma,k}^{(i)}, \quad \mathbf{m}_{k|k-1}^{(\bar{\imath})} := \mathbf{m}_{\gamma,k}^{(i)}, \quad \mathbf{P}_{k|k-1}^{(\bar{\imath})} := \mathbf{P}_{\gamma,k}^{(i)}$.
end
for $i = 1, \ldots, J_{k-1}$ (*i.e. for persistent extended targets*)
$\bar{\imath} := \bar{\imath} + 1$,
$\omega_{k|k-1}^{(\bar{\imath})} := p_S\, \omega_{k-1}^{(i)}$,
$\mathbf{m}_{k|k-1}^{(\bar{\imath})} := \mathbf{F}_{k-1}\mathbf{m}_{k-1}^{(i)}, \quad \mathbf{P}_{k|k-1}^{(\bar{\imath})} := \mathbf{F}_{k-1}\mathbf{P}_{k-1}^{(i)}\mathbf{F}_{k-1}^T + \mathbf{Q}_{k-1}$.
end
$J_{k|k-1} := \bar{\imath}$.
output $\left\{\omega_{k|k-1}^{(\bar{\imath})}, \mathbf{m}_{k|k-1}^{(\bar{\imath})}, \mathbf{P}_{k|k-1}^{(\bar{\imath})}\right\}_{\bar{\imath}=1}^{J_{k|k-1}}$.

Table 5.1: Pseudo-code for the GM-PHD prediction for multiple extended targets

(Proposition 4.1). Although note that now $\omega_{k-1}^{(i)} = \nu_{k-1}^{(i)}$, $\mathbf{m}_{k-1}^{(i)} = \mathbf{m}_{\mathbf{x},k-1}^{(i)}$, $\mathbf{P}_{k-1}^{(i)} = \mathbf{P}_{22,k-1}^{(i)}$, $\omega_{k|k-1}^{(i)} = \nu_{k|k-1}^{(i)}$, $\mathbf{m}_{k|k-1}^{(i)} = \mathbf{m}_{\mathbf{x},k|k-1}^{(i)}$ and $\mathbf{P}_{k|k-1}^{(i)} = \mathbf{P}_{22,k|k-1}^{(i)}$.

To complete the recursion, suppose the measurement model is given by factorisation of linear Gaussian functions in equation (4.15) (Section 4.1.4), and the arbitrary intensity of each subsidiary RFS $\mathcal{E}_k$ in the hierarchical RFS representation is formulated with the following Gaussian mixture

$$\mu(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k) = \sum_{j=1}^{J_{\boldsymbol{\xi}}^{(i)}} \mathcal{N}\left(\boldsymbol{\xi}_k \,;\, \mu_{\boldsymbol{\xi}|\mathbf{x},k|k-1}^{(i,j)}, \boldsymbol{\Sigma}_{\boldsymbol{\xi}|\mathbf{x},k|k-1}^{(i,j)}\right), \tag{5.2}$$

where the means and covariances are given by

$$\begin{aligned} \boldsymbol{\mu}_{\boldsymbol{\xi}|\mathbf{x},k|k-1}^{(i,j)} &= \mathbf{m}_{\boldsymbol{\xi}}^{(j|i)} + \mathbf{P}_{\boldsymbol{\xi}|\mathbf{x}}^{(i|j)} \left(\mathbf{P}_{k|k-1}^{(i)}\right)^{-1} \left(\mathbf{x}_k - \mathbf{m}_{k|k-1}^{(i)}\right), \\ \boldsymbol{\Sigma}_{\boldsymbol{\xi}|\mathbf{x},k|k-1}^{(i|j)} &= \mathbf{P}_{\boldsymbol{\xi}}^{(i|j)} - \mathbf{P}_{\boldsymbol{\xi}|\mathbf{x}}^{(i|j)} \left(\mathbf{P}_{k|k-1}^{(i)}\right)^{-1} \left(\mathbf{P}_{\boldsymbol{\xi}|\mathbf{x}}^{(i|j)}\right)^{T}, \end{aligned} \tag{5.3}$$

for $i = 1, \ldots, J_{k|k-1}$. Note that, due to the arbitrariness of the intensity $\mu(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k)$, the number of Gaussian components $J_{\boldsymbol{\xi}}^{(i)}$ in the mixture given by equation (5.2) can be considered as fixed, negating the need to assign a weight to each component. Then, given the GM formulation of the predicted intensity $v_{k|k-1}(\mathbf{x}_k)$, the **updated intensity** $v_k(\mathbf{x}_k)$ is also a Gaussian mixture of the form shown in the illustrative diagram in Section 5.1.1.

The closed-form expressions of the updated weights, means and covariances are detailed in Tables 5.2 and 5.3. The means and covariances given in Tables 5.2 and 5.3 have the same expressions as those relating to the group states in the multi-group GM-PHD update (Proposition 4.2), although now $\mathbf{m}_k^{(i)} = \mathbf{m}_{\mathbf{x},k}^{(i)}$ and $\mathbf{P}_k^{(i)} = \mathbf{P}_{22,k}^{(i)}$. However, the expressions for the weights differ slightly from those relating to the group states in Proposition 4.2 and now denoted by $\omega_k^{(i)} = \nu_k^{(i)}$. In particular, for missed detections (see Table 5.2), the updated weight now does not include the factor $1 - p_{D,2}$, while for detections (see Table 5.3), the computation of the multiple measurement likelihood, which is present in the updated weights, does not now include any weights associated with Gaussian components relating to individual target states (feature points in the case of extended targets).

The GM formulation of the PHD filter for extended targets provides the following expected number of targets

$$N_{k|k-1} = p_S \, N_{k-1} + \sum_{i=1}^{J_{\gamma,k}} \omega_{\gamma,k}^{(i)}, \tag{5.4}$$

$$N_k = \left(1 - p_{D,1} + p_{D,1}\,\mathrm{e}^{-p_{D,2}\,\alpha}\right) N_{k|k-1} + \sum_{\pi \in \Pi_{Z_k}} \varpi_\pi \sum_{\varphi \in \pi} \sum_{i=1}^{J_{k|k-1}} \sum_{j_{1:|\varphi|}} \widetilde{\omega}_{\varphi,k}^{(i)}, \tag{5.5}$$

associated with the predicted and the updated intensities, $v_{k|k-1}$ and v_k respectively, where $j_{1:|\varphi|} = (j_1, \ldots, j_{|\varphi|})$ denotes a permutation of the indices $j_\ell = 1, \ldots, J_{\boldsymbol{\xi}}^{(i)}$ for $\ell = 1, \ldots, |\varphi|$ and

$$\widetilde{\omega}_{\varphi,k}^{(i)} = \frac{p_{D,1}\,\omega_{k|k-1}^{(i)}\,\mathrm{e}^{-p_{D,2}\,\alpha} \prod_{\ell=1}^{|\varphi|} L_\ell^{(i)}}{\Lambda(\varphi) + p_{D,1}\,\mathrm{e}^{-p_{D,2}\,\alpha} \sum_{i=1}^{J_{k|k-1}} \omega_{k|k-1}^{(i)} \sum_{j_{1:|\varphi|}} \prod_{\ell=1}^{|\varphi|} L_\ell^{(i)}}. \tag{5.6}$$

It can be deduced from the expressions for $N_{k|k-1}$ and N_k in (5.4) and (5.5) respectively, that the GM-PHD filter for extended targets requires, at any time-step k,

$$J_k = J_{k|k-1} + \prod_{\pi \in \Pi_{Z_k}} \prod_{\varphi \in \pi} \prod_{i=1}^{J_{k|k-1}} |\varphi| \times J_{\boldsymbol{\xi}}^{(i)}, \tag{5.7}$$

Gaussian components in v_k, where $J_{k|k-1} = J_{\gamma,k} + J_{k-1}$ is the number of components in $v_{k|k-1}$.

---

given $\left\{\omega_{k|k-1}^{(i)}, \mathbf{m}_{k|k-1}^{(i)}, \mathbf{P}_{k|k-1}^{(i)} \left\{\mathbf{m}_{\boldsymbol{\xi}}^{(j|i)}, \mathbf{P}_{\boldsymbol{\xi}}^{(j|i)}, \mathbf{P}_{\boldsymbol{\xi}|\mathbf{x}}^{(j|i)}\right\}_{j=1}^{J_{\boldsymbol{\xi}}^{(i)}}\right\}_{i=1}^{J_{k|k-1}}$.

step 1. (*Update for missed detections*)

$\bar{\imath} := 0$.

for $i = 1, \ldots, J_{k|k-1}$

$\bar{\imath} := \bar{\imath} + 1, \quad \alpha = J_{\boldsymbol{\xi}}^{(i)}$,

$\omega_k^{(\bar{\imath})} := (1 - p_{D,1} + p_{D,1}\,\mathrm{e}^{-p_{D,2}\,\alpha})\,\omega_{k|k-1}^{(i)}, \quad \mathbf{m}_k^{(\bar{\imath})} := \mathbf{m}_{k|k-1}^{(i)}, \quad \mathbf{P}_k^{(\bar{\imath})} := \mathbf{P}_{k|k-1}^{(i)}$.

end

step 2. (*Construction of GM-PHD update components*)

for $i = 1, \ldots, J_{k|k-1}$

for $j = 1, \ldots, J_{\boldsymbol{\xi}}^{(i)}$

$\widetilde{\mathbf{R}}^{(i,j)} := \mathbf{H}\boldsymbol{\Sigma}_{\boldsymbol{\xi}|\mathbf{x},k|k-1}^{(i,j)}\mathbf{H}^T + \mathbf{R}_1, \quad \mathbf{A}_k^{(i,j)} := \mathbf{H}\mathbf{P}_{\boldsymbol{\xi}|\mathbf{x}}^{(j|i)} \left(\mathbf{P}_{k|k-1}^{(i)}\right)^{-1}$,

where $\boldsymbol{\Sigma}_{\boldsymbol{\xi}|\mathbf{x},k|k-1}^{(i,j)}$ is given by the expression in (5.3).

end

end

$J_k := \bar{\imath}$.

output $\left\{\omega_k^{(i)}, \mathbf{m}_k^{(i)}, \mathbf{P}_k^{(i)}\right\}_{i=1}^{J_k}$, $\left\{\left\{\widetilde{\mathbf{R}}^{(i,j)}, \mathbf{A}_k^{(i,j)}\right\}_{j=1}^{J_{\boldsymbol{\xi}}^{(i)}}\right\}_{i=1}^{J_{k|k-1}}$.

Table 5.2: Pseudo-code for steps 1 & 2 of the GM-PHD update for multiple extended targets

given $\left\{\omega_{k|k-1}^{(i)}, \mathbf{m}_{k|k-1}^{(i)}, \mathbf{P}_{k|k-1}^{(i)} \left\{\mathbf{m}_{\boldsymbol{\xi}}^{(j|i)}, \widetilde{\mathbf{R}}^{(i,j)}, \mathbf{A}_k^{(i,j)}\right\}_{j=1}^{J_{\boldsymbol{\xi}}^{(i)}}\right\}_{i=1}^{J_{k|k-1}}$, $\left\{\omega_k^{(i)}, \mathbf{m}_k^{(i)}, \mathbf{P}_k^{(i)}\right\}_{i=1}^{J_k}$,
measurement set Z_k and the set of all possible partitions Π_{Z_k},
step 3. (*Update for detections*)
$\bar{\imath} := J_k$,
for each partition $\pi \in \Pi_{Z_k}$
for $i_\pi = 1, \ldots, |\pi|$ (*i.e. for each subset* $\varphi_{i_\pi} \in \pi$)
$m := |\varphi_{i_\pi}|, \quad n := \bar{\imath}$.
for $i = 1, \ldots, J_{k|k-1}$
for each permutation $j_{1:m} = (j_1, \ldots, j_m)$ (*such that* $j_\ell = 1, \ldots, J_\ell^{(i)}$)
$\bar{\imath} := \bar{\imath} + 1, \quad \boldsymbol{\eta}_k^{(j_1,i)} := \mathbf{m}_{k|k-1}^{(i)}, \quad \boldsymbol{\mathcal{P}}_k^{(j_1,i)} := \mathbf{P}_{k|k-1}^{(i)}$.
for $\ell = 1, \ldots, m$
$\mathbf{b}_1^{(j_{1:\ell},i)} := \mathbf{H}\mathbf{m}_{\boldsymbol{\xi}}^{(j_\ell|i)} + \mathbf{A}_k^{(i,j_\ell)}\left(\boldsymbol{\eta}_k^{(j_{1:\ell},i)} - \mathbf{m}_{k|k-1}^{(i)}\right)$,
$\mathbf{S}_1^{(j_{1:\ell},i)} := \mathbf{A}_k^{(i,j_\ell)}\boldsymbol{\mathcal{P}}_k^{(j_{1:\ell},i)}\left(\mathbf{A}_k^{(i,j_\ell)}\right)^T + \widetilde{\mathbf{R}}_k^{(i,j_\ell)}$,
$\widetilde{\mathbf{K}}_k^{(j_{1:\ell},i)} := \boldsymbol{\mathcal{P}}_k^{(j_{1:\ell},i)}\left(\mathbf{A}_k^{(i,j_\ell)}\right)^T\left(\mathbf{S}_1^{(j_{1:\ell},i)}\right)^{-1}$,
$\widetilde{\mathbf{m}}_k^{(j_{1:\ell},i)} := \boldsymbol{\eta}_k^{(j_{1:\ell},i)} + \widetilde{\mathbf{K}}_k^{(j_{1:\ell},i)}\left(\mathbf{z}_\ell - \mathbf{b}_1^{(j_{1:\ell},i)}\right)$,
$\widetilde{\mathbf{P}}_k^{(j_{1:\ell},i)} := \left(\mathbf{I} - \widetilde{\mathbf{K}}_k^{(j_{1:\ell},i)}\mathbf{A}_k^{(i,j_\ell)}\right)\boldsymbol{\mathcal{P}}_k^{(j_{1:\ell},i)}$,
(*where* $\boldsymbol{\eta}_k^{(j_{1:\ell},i)} = \boldsymbol{\eta}_k^{(j_1,i)}$ *and* $\boldsymbol{\mathcal{P}}_k^{(j_{1:\ell},i)} = \boldsymbol{\mathcal{P}}_k^{(j_1,i)}$ *for* $\ell = 1$)
$\mathbf{S}_2^{(j_{1:\ell},i)} := \widehat{\mathbf{H}}\widetilde{\mathbf{P}}_k^{(j_{1:\ell},i)}\widehat{\mathbf{H}}^T + \mathbf{R}_2, \quad \mathbf{K}_k^{(j_{1:\ell},i)} := \widetilde{\mathbf{P}}_k^{(j_{1:\ell},i)}\widehat{\mathbf{H}}^T\left(\mathbf{S}_2^{(j_{1:\ell},i)}\right)^{-1}$,
$\boldsymbol{\eta}_k^{(j_{1:\ell},i)} := \widetilde{\mathbf{m}}_k^{(j_{1:\ell},i)} + \mathbf{K}_k^{(j_{1:\ell},i)}\left(\mathbf{z}_\ell - \widehat{\mathbf{H}}\widetilde{\mathbf{m}}_k^{(j_{1:\ell},i)}\right)$,
$\boldsymbol{\mathcal{P}}_k^{(j_{1:\ell},i)} := \left(\mathbf{I} - \mathbf{K}_k^{(j_{1:\ell},i)}\widehat{\mathbf{H}}\right)\widetilde{\mathbf{P}}_k^{(j_{1:\ell},i)}$,
$L_\ell^{(i)} := p_{D,2}\,\mathcal{N}\left(\mathbf{z}_\ell\,;\mathbf{b}_1^{(j_{1:\ell},i)}, \mathbf{S}_1^{(j_{1:\ell},i)}\right)\mathcal{N}\left(\mathbf{z}_\ell\,;\widehat{\mathbf{H}}\widetilde{\mathbf{m}}_k^{(j_{1:\ell},i)}, \mathbf{S}_2^{(j_{1:\ell},i)}\right)$.
end
$\mathbf{m}_k^{(\bar{\imath})} := \boldsymbol{\eta}_k^{(j_{1:m},i)}, \quad \mathbf{P}_k^{(\bar{\imath})} := \boldsymbol{\mathcal{P}}_k^{(j_{1:m},i)}$,
$\alpha := J_{\boldsymbol{\xi}}^{(i)}, \quad \tilde{\omega}_k^{(\bar{\imath})} := p_{D,1}\,\omega_{k|k-1}^{(i)}\,\mathrm{e}^{-p_{D,2}\,\alpha}\prod_{\ell=1}^{m} L_\ell^{(i)}$.
end
end
$\omega_\varphi^{(i_\pi)} := \Lambda(\varphi_{i_\pi}) + \sum_{i=n}^{\bar{\imath}} \tilde{\omega}_k^{(i)}$.
for $i = n, \ldots, \bar{\imath}$
$\omega_k^{(i)} := \tilde{\omega}_k^{(i)} \Big/ \omega_\varphi^{(i_\pi)}$.
end
end
$\widetilde{\varpi}_\pi := \prod_{i_\pi=1}^{|\pi|} \omega_\varphi^{(i_\pi)}$.
end
for each $\pi \in \Pi_{Z_k}$
$\varpi_\pi := \widetilde{\varpi}_\pi \Big/ \sum_{\pi \in \Pi_{Z_k}} \widetilde{\varpi}_\pi$.
end
$J_k := \bar{\imath}$.
output $\left\{\omega_k^{(i)}, \mathbf{m}_k^{(i)}, \mathbf{P}_k^{(i)}\right\}_{i=1}^{J_k}$.

Table 5.3: Pseudo-code for step 3 of the GM-PHD update for multiple extended targets

It is clear from equation (5.7) that the number of Gaussian components in the posterior intensity increases without bound over time. In fact, the rate of increase is equivalent to that of the number of Gaussian components relating to group states in the multi-group GM-PHD filter, due to the similarities between the expression given in equations (4.45) and (5.7). Consequently, mixture reduction procedures are once again considered, as discussed in Section 5.1.3.

5.1.3 Mixture reduction procedures

Like the multi-group GM-PHD filter in the previous chapter, the most significant contributor to the unbounded increase in the number of Gaussian components at each iteration of the algorithm shown in Section 5.1.2, is potentially the number of partitions of a measurement set. This accounts for the products $\prod_{\pi\in\Pi_{Z_k}}$ and $\prod_{\varphi\in\pi}$ in equation (5.7). The computational intractability associated with finding all possible partitions of a measurement set, particularly if that set contains a large number of measurements, is addressed by adopting the same gating procedure detailed in Table 4.3, noting the changes in notation for the weights, means and covariances as alluded to in Section 5.1.2.

The other significant contributor to the increase in the number of Gaussian components is the pseudo multiple measurement likelihood $L_{\varphi}(\mathbf{x}_k)$, which accounts for the product $|\varphi| \times J_{\boldsymbol{\xi}}^{(i)}$ in equation (5.7). To alleviate the computational demand associated with updating the state $\mathbf{x}_k$ (by multiplying the predicted intensity $v_{k|k-1}(\mathbf{x}_k)$ with $L_{\varphi}(\mathbf{x}_k)$), the same gating procedure detailed in Table 4.4 can be performed, noting the changes in notation for the means and covariances as now found in (5.3), due to the formulation for $\mu(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k)$ given in equation (5.2). In the case of extended targets, the procedure validates measurements in each subset φ of a partition π for the feature points $\{\boldsymbol{\xi}_{k,j} : j = 1, \ldots, J_{\boldsymbol{\xi}}^{(i)}\}\ \forall\, i = 1, \ldots, J_{k|k-1}$.

As a result of performing both gating procedures, the expected number of targets associated with v_k becomes

$$N_k = \left(1 - p_{D,1} + p_{D,1}\,\mathrm{e}^{-p_{D,2}\,\alpha}\right) N_{k|k-1} + \sum_{\varphi\in\tilde{\boldsymbol{\pi}}} \sum_{i=1}^{J_{k|k-1}} \sum_{j_{1:n_\varphi}} \widetilde{\omega}_{\varphi,k}^{(i)}, \tag{5.8}$$

from which it follows that the number of Gaussian components reduces to

$$\tilde{J}_k = J_{k|k-1} + \prod_{\varphi\in\tilde{\boldsymbol{\pi}}} \prod_{i=1}^{J_{k|k-1}} \prod_{\ell=1}^{n_\varphi} J_{\boldsymbol{\xi},\ell}^{(i)}\,, \tag{5.9}$$

where $j_{1:n_\varphi} = (j_1, \ldots, j_{n_\varphi})$ now denotes the permutation of the indices $j_\ell = 1, \ldots, J^{(i)}_{\boldsymbol{\xi},\ell}$ for $\ell = 1, \ldots, n_\varphi$, such that $J^{(i)}_{\boldsymbol{\xi},\ell} \leq J^{(i)}_{\boldsymbol{\xi}}$ and $n_\varphi \leq |\varphi|$, which alters the expression for $\widetilde{\omega}^{(i)}_{\varphi,k}$ in equation (5.6) accordingly. Those of the remaining $\tilde{J}_k$ Gaussian components which are closely spaced can be merged using a similar method to that shown in Table 4.5, the details of which are given in Table 5.4. Finally, extended target state estimates are extracted from the means of the Gaussian components whose weights are greater than some threshold as shown in Table 5.5.

---

given $\left\{\omega^{(i)}_k, \mathbf{m}^{(i)}_k, \mathbf{P}^{(i)}_k\right\}_{i=1}^{\tilde{J}_k}$, truncation threshold τ, merging threshold U_{Mah} (*or* U_{Hel}), and maximum allowable number of Gaussian components $J_{\max}$.

Set $\bar{\imath} = 0$ and $I = \left\{i = 1, \ldots, \tilde{J}_k \,\middle|\, \omega^{(i)}_k > \tau\right\}$.

repeat

$\bar{\imath} := \bar{\imath} + 1, \quad i_{\max} := \arg\max_{i \in I} \omega^{(i)}_k,$

$\bar{I} := \left\{i \in I \,\middle|\, \left(\mathbf{m}^{(i)}_k - \mathbf{m}^{(i_{\max})}_k\right)^T \left(\mathbf{P}^{(i)}_k\right)^{-1} \left(\mathbf{m}^{(i)}_k - \mathbf{m}^{(i_{\max})}_k\right) \leq U_{\text{Mah}}\right\},$

$\left(or\ \bar{I} := \left\{i \in I \,\middle|\, d_H\left(\mathcal{N}\left(\mathbf{x}_k\,; \mathbf{m}^{(i)}_k, \mathbf{P}^{(i)}_k\right), \mathcal{N}\left(\mathbf{x}_k\,; \mathbf{m}^{(i_{\max})}_k, \mathbf{P}^{(i_{\max})}_k\right)\right) \leq U_{\text{Hel}}\right\}\right)$

$\widetilde{\omega}^{(\bar{\imath})}_k := \sum_{i \in \bar{I}} \omega^{(i)}_k, \quad \widetilde{\mathbf{m}}^{(\bar{\imath})}_k := \frac{1}{\widetilde{\omega}^{(\bar{\imath})}_k} \sum_{i \in \bar{I}} \omega^{(i)}_k \mathbf{m}^{(i)}_k, \quad \widetilde{\mathbf{P}}^{(\bar{\imath})}_k := \frac{1}{\widetilde{\omega}^{(\bar{\imath})}_k} \sum_{i \in \bar{I}} \omega^{(i)}_k \mathbf{P}^{(i)}_k,$

$I := I \setminus \bar{I}.$

until $I = \emptyset$.

If $\bar{\imath} > J_{\max}$ then replace $\left\{\widetilde{\omega}^{(i)}_k, \widetilde{\mathbf{m}}^{(i)}_k, \widetilde{\mathbf{P}}^{(i)}_k\right\}_{i=1}^{\bar{\imath}}$ by those $J_{\max}$ Gaussian components with the largest weights and setting $J_{\text{new},k} = J_{\max}$ (otherwise set $J_{\text{new},k} = \bar{\imath}$).

output $\left\{\widetilde{\omega}^{(i)}_k, \widetilde{\mathbf{m}}^{(i)}_k, \widetilde{\mathbf{P}}^{(i)}_k\right\}_{i=1}^{J_{\text{new},k}}$.

Table 5.4: Merging closely spaced Gaussian components

---

given $\left\{\omega^{(i)}_k, \mathbf{m}^{(i)}_k, \mathbf{P}^{(i)}_k\right\}_{i=1}^{J_k}$ and threshold ω_{Th}.

Set $\tilde{X}_k = \emptyset$.

for $i = 1, \ldots, J_k$

  if $\omega^{(i)}_k > \omega_{\text{Th}}$

    $\tilde{X}_k := \left\{\tilde{X}_k, \mathbf{m}^{(i)}_k\right\},$

  end

end

output $\tilde{X}_k$ as the multiple extended target state estimate.

Table 5.5: Extended target state extractions

5.2 Scenario 1: Extended Target Tracking with Fixed Extent

A direct application of the GM-PHD filter presented in the previous section, is demonstrated in this section on the extended target tracking scenario described in Section 5.2.1 with a simulated example. The simulation results for the estimation problem are then presented in Section 5.2.2, along with results for the filter's performance evaluation.

5.2.1 Simulation set-up

The GM-PHD filter presented in this section applies to extended targets whose extents have any structural shape in general. For illustration purposes, a multiple extended target scenario is considered in which the extents are elliptical in shape and the filter is tested on simulated data. The generated target trajectories and elliptical extents for this scenario are shown in Figure 5.2. The elliptical shaped extents are parameterized by the major and minor axes, which are fixed at values $r_{\text{major}} = 20\,\text{m}$ and $r_{\text{minor}} = 15\,\text{m}$ respectively.

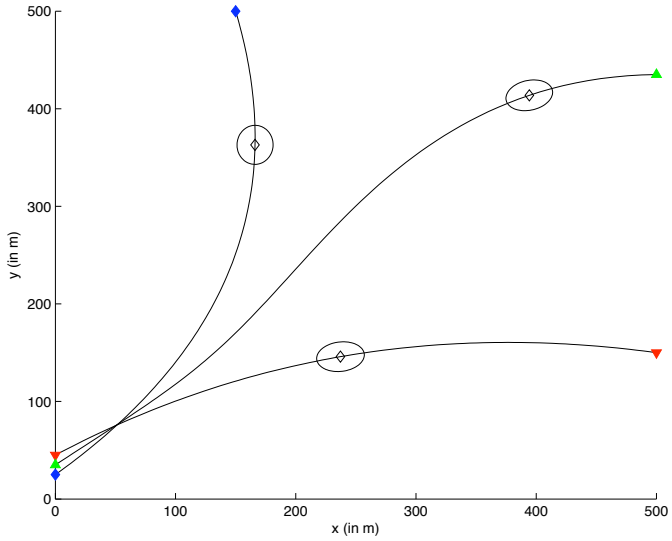


Figure 5.2: The extended target trajectories (line) whose start and end positions are indicated by: blue 'diamonds' for Trajectory 1 at $k = 1$ and $k = 100$; red 'downward triangles' for Trajectory 2 at $k = 26$ and $k = 125$; and green 'upward triangles' for Trajectory 3 at $k = 51$ and $k = 150$.

Each extended target state $\mathbf{x}_k = \begin{bmatrix} x_{k,1} & x_{k,2} & x_{k,3} & x_{k,4} \end{bmatrix}^T$ consists of two-dimensional Cartesian co-ordinate position vector $\begin{bmatrix} x_{k,1} & x_{k,3} \end{bmatrix}^T$ (i.e. the centre of the ellipse) and velocity vector $\begin{bmatrix} x_{k,2} & x_{k,4} \end{bmatrix}^T$ at time-step k. It follows the linear

Gaussian dynamic model given by equation (4.8) with

$$\mathbf{F}_{k-1} = \begin{bmatrix} \widehat{\mathbf{F}}_{k-1} & \mathbf{0}_2 \\ \mathbf{0}_2 & \widehat{\mathbf{F}}_{k-1} \end{bmatrix}, \qquad \mathbf{Q}_{k-1} = \sigma^2 \, \mathbf{\Gamma}_{k-1} \, \mathbf{\Gamma}_{k-1}^T, \tag{5.10}$$

where $\widehat{\mathbf{F}}_{k-1}$ has the same expression as given in (4.61) and

$$\mathbf{\Gamma}_{k-1} = \begin{bmatrix} t^2/2 & t & 0 & 0 \\ 0 & 0 & t^2/2 & t \end{bmatrix}, \tag{5.11}$$

with sampling period $t = 1\,\mathrm{s}$, noting the changes in notation for the transition and process noise covariance matrices, as alluded to in Section 5.1.2.

The Cartesian point measurements $\mathbf{z} = \begin{bmatrix} z_1 & z_2 \end{bmatrix}^T \in Z_k$ relate to feature points $\boldsymbol{\xi}_k$ of a target according to the likelihood $\mathcal{N}(\mathbf{z}\,; \mathbf{H}\boldsymbol{\xi}_k, \mathbf{R}_1)$ in equation (4.15) with $\mathbf{H} = \mathbf{I}_2$ and $\mathbf{R}_1 = \sigma_1^2 \, \mathbf{I}_2$ where $\sigma_1 = \sqrt{10}\,\mathrm{m}$. The probability of detecting these feature points is $p_{D,2} = 0.78$. The measurements relate to the target states $\mathbf{x}_k$ according to the likelihood $\mathcal{N}(\mathbf{z}\,; \widehat{\mathbf{H}}\mathbf{x}_k, \mathbf{R}_2)$ in (4.15) with $\widehat{\mathbf{H}}$ as defined in equation (4.63) and $\mathbf{R}_2 = \sigma_2^2 \, \mathbf{I}_2$ where $\sigma_2 = \sqrt{r_{\text{major}} + \sigma_1^2}\,\mathrm{m}$. As was noted in [100], the effective probability of target detection takes into account both the number of feature points α $(= 9)$ and $p_{D,1}$ $(= 0.9)$. The measurement set Z_k includes false alarms which are modelled with the hierarchical RFS representation whose probability density is given by equations (3.19) and (3.20) with $\kappa_1(\mathrm{d}\mathbf{u}_\varphi) = \lambda_1 \, U(\mathbf{u}_\varphi)$ and $\kappa_2(\mathbf{z} \,|\mathbf{u}_\varphi) = \lambda_2 \, \mathcal{N}(\mathbf{z}\,; \mathbf{u}_\varphi, \mathbf{R}_\kappa)$, where $U(\cdot)$ is the uniform density over the surveillance region $[0\ 500] \times [0\ 500]$, $\lambda_1 = 50$ is the expected number of false measurement clusters, $\lambda_2 = 1$ is the average number of measurement per cluster and $\mathbf{R}_\kappa = 5 \times \mathbf{I}_2$. Figure 5.3 shows the cluttered measurements plotted with the target trajectories against time.

The arbitrary intensity $\mu(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k)$ for each target has the Gaussian mixture form given in equation (5.2) with $J_{\boldsymbol{\xi}}^{(i)} = \alpha$ for $i = 1, \ldots, J_{k|k-1}$ and for $j = 1, \ldots, J_{\boldsymbol{\xi}}^{(i)}$

$$\mathbf{m}_{\boldsymbol{\xi}}^{(j|i)} = \widehat{\mathbf{H}}\mathbf{m}_{k|k-1}^{(i)} + \mathbf{r}_{k|k-1}^{(j|i)}, \qquad \mathbf{P}_{\boldsymbol{\xi}}^{(j|i)} = \mathbf{R}_1 + \mathbf{P}_{\boldsymbol{\xi}|\mathbf{x}}^{(j|i)} \left(\mathbf{P}_{k|k-1}^{(i)} \right)^{-1} \left(\mathbf{P}_{\boldsymbol{\xi}|\mathbf{x}}^{(j|i)} \right)^T, \tag{5.12}$$

where the cross diagonal covariance matrices are given by $\mathbf{P}_{\boldsymbol{\xi}|\mathbf{x}}^{(j|i)} = \widehat{\mathbf{H}}\mathbf{P}_{k|k-1}^{(i)}$. The means $\mathbf{m}_{\boldsymbol{\xi}}^{(j|i)}$ relate to the feature points distributed along the perimeter of the target's elliptical shaped extent at fixed intervals where

$$\mathbf{r}_{k|k-1}^{(j|i)} = \begin{bmatrix} r_{\text{major}} \cos(\theta_j) \cos(\phi_{k|k-1}^{(i)}) + r_{\text{minor}} \sin(\theta_j) \sin(\phi_{k|k-1}^{(i)}) \\ r_{\text{major}} \cos(\theta_j) \sin(\phi_{k|k-1}^{(i)}) + r_{\text{minor}} \sin(\theta_j) \cos(\phi_{k|k-1}^{(i)}) \end{bmatrix}, \tag{5.13}$$

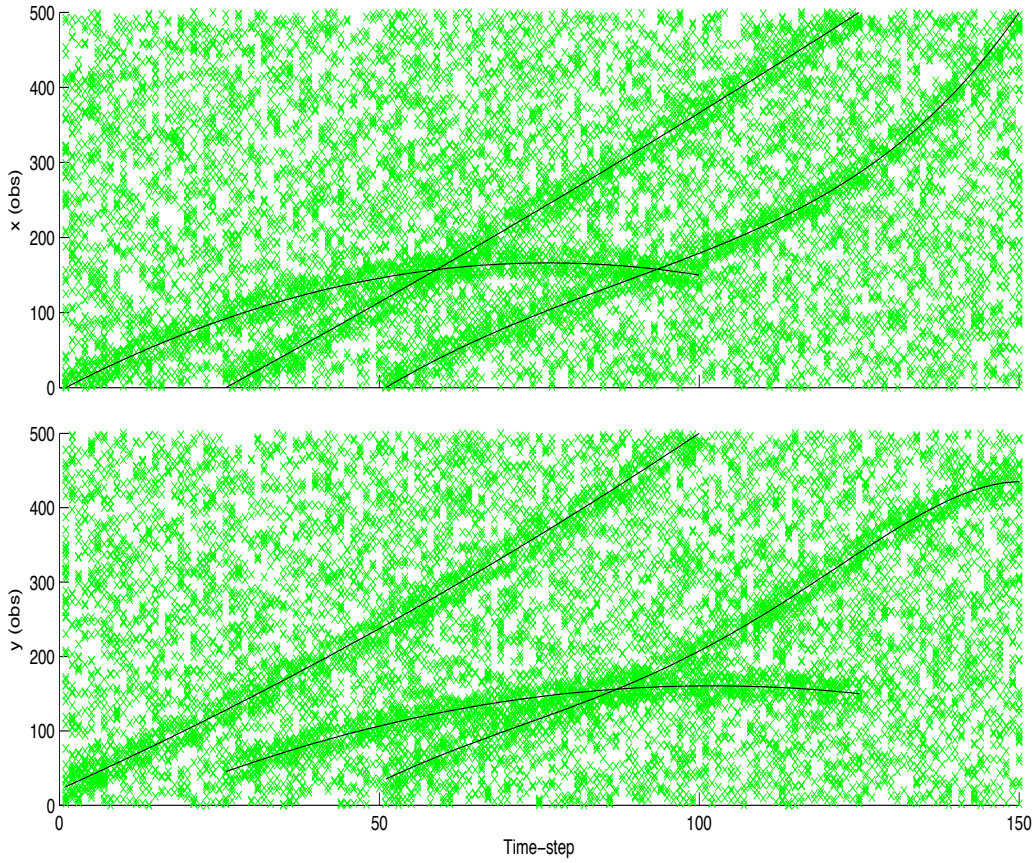


Figure 5.3: Measurements (green 'crosses') and target trajectories (lines)

with $\theta_j = 2\pi j/\alpha$ and $\phi^{(i)}_{k|k-1} = \arctan(m^{(i)}_{k|k-1,4}/m^{(i)}_{k|k-1,2})$, denoting $m^{(i)}_{k|k-1,n}$ as the n^{th} element of the mean vector $\mathbf{m}^{(i)}_{k|k-1}$, which relate to the velocities for $n = 2, 4$. The angles $\phi^{(i)}_{k|k-1}$ relate to the rotation of the elliptical extents which depend on the direction of the targets' trajectories.

The filter is initialised at time-step $k = 1$ with the intensity $\gamma_k(\mathbf{x}_k)$, so that the predicted intensity consists solely of target births, as shown in Table 5.6. For this initialisation procedure, the point of origin for targets and the initial number of Gaussian components are set to be $\mathbf{m}_{\text{source}} = [0 \;\; 35]^T$ and $J_{\text{initial}} = 5$ respectively, while the probability of target birth is $p_B = 0.1$.

At subsequent time-steps $k > 1$, the weights, means and covariances of the Gaussian components relating to persistent targets in $v_{k|k-1}$ are computed first so that, for $i = 1, \ldots, J_{k-1}$, $\omega^{(i)}_{k|k-1}$, $\mathbf{m}^{(i)}_{k|k-1}$ and $\mathbf{P}^{(i)}_{k|k-1}$ have the closed-form expressions shown in Table 5.1. Target births, at time-steps $k > 1$, are measurement driven and determined with respect to the single partition $\boldsymbol{\pi}$ of the measurement set, obtained as a result of performing the gating procedure in Table 4.3 as discussed in Section 5.1.3. Details for such a procedure are given in Table 5.7.

given an initial number of Gaussian components J_{initial},
a point of origin for targets $\mathbf{m}_{\text{source}}$,
an observation region $[a,b] \times [c,d]$ and noise covariance matrix $\mathbf{R}_2 = \sigma_2^2 \mathbf{I}_2$,
and probability of target birth p_B.

procedure At time-step $k = 1$

$n := 0$.

for $i = 1, \ldots, J_{\text{initial}}$

Sample $\widetilde{\mathbf{m}}_{\gamma,k}^{(i)} = \left[\widetilde{m}_{\gamma,k,1}^{(i)} \;\; \widetilde{m}_{\gamma,k,2}^{(i)}\right]^T \sim \mathcal{N}\left(\mathbf{x}_k\,; \mathbf{m}_{\text{source}}, \mathbf{R}_2\right)$, so that $a \leq \widetilde{m}_{\gamma,k,1}^{(i)} \leq b$ and $c \leq \widetilde{m}_{\gamma,k,2}^{(i)} \leq d$,

$n = n + 1, \quad \omega_{k|k-1}^{(n)} := p_B,$

$\mathbf{m}_{k|k-1}^{(n)} := \left[\widetilde{m}_{\gamma,k,1}^{(i)} \;\; 0 \;\; \widetilde{m}_{\gamma,k,2}^{(i)} \;\; 0\right]^T, \quad \mathbf{P}_{k|k-1}^{(n)} := \text{diag}([\sigma_2^2, 2, \sigma_2^2, 2])$.

end

$J_{k|k-1} := n$.

output $\left\{\omega_{k|k-1}^{(i)}, \mathbf{m}_{k|k-1}^{(i)}, \mathbf{P}_{k|k-1}^{(i)}\right\}_{i=1}^{J_{k|k-1}}$.

Table 5.6: Initialisation of the GM-PHD filter for multiple extended targets

given a single partition $\boldsymbol{\pi}$ of the measurement set Z_k,
a point of origin for targets $\mathbf{m}_{\text{source}}$, noise covariance matrix $\mathbf{R}_G \propto \mathbf{R}_2$
and gating threshold T_2.

procedure

Set $W = \emptyset$ and $\tilde{\boldsymbol{\pi}} = \boldsymbol{\pi}$.

for each $\varphi \in \boldsymbol{\pi}$ such that $|\varphi| = 1$

$W := W \cup \varphi, \quad \tilde{\boldsymbol{\pi}} := \tilde{\boldsymbol{\pi}} \setminus \varphi.$

end

$n := |\tilde{\boldsymbol{\pi}}|$,

$\tilde{\varphi} := \left\{\mathbf{z} \in W \,\middle|\, (\mathbf{z} - \mathbf{m}_{\text{source}})^T \mathbf{R}_G^{-1} (\mathbf{z} - \mathbf{m}_{\text{source}}) \leq T_2\right\}$.

if $|\tilde{\varphi}| > 1$

$n := n + 1$,

$\varphi_n := \tilde{\varphi}, \quad \tilde{\boldsymbol{\pi}} := \{\tilde{\boldsymbol{\pi}}, \varphi_n\}$,

$W := W \setminus \varphi_n$.

Create Gaussian components for target births $\left\{\omega_{\gamma,k}^{(i)}, \mathbf{m}_{\gamma,k}^{(i)}, \mathbf{P}_{\gamma,k}^{(i)}\right\}_{i=1}^{J_{\gamma,k}}$ following the initialisation procedure in Table 5.6 with $J_{\gamma,k} = J_{\text{initial}}$.

end

for each remaining measurement $\mathbf{z} \in W$

$n := n + 1$,

$\varphi_n := \{\mathbf{z}\}, \quad \tilde{\boldsymbol{\pi}} := \{\tilde{\boldsymbol{\pi}}, \varphi_n\}$.

end

output $\tilde{\boldsymbol{\pi}} = \{\varphi_j\}_{j=1}^{n}$ and $\left\{\omega_{\gamma,k}^{(i)}, \mathbf{m}_{\gamma,k}^{(i)}, \mathbf{P}_{\gamma,k}^{(i)}\right\}_{i=1}^{J_{\gamma,k}}$.

Table 5.7: Procedure for determining target births at time-steps $k > 1$.

5.2.2 Results

The GM-PHD filter for extended targets, presented in Section 5.1, is applied to the scenario described in Section 5.2.1, with parameters $T_2 = 9$, $\mathbf{R}_G = 5 \times \mathbf{R}_2$ (see Tables 4.3 and 5.7), $T_1 = 4$ (see Table 4.4), $\tau = 10^{-4}$, $U_{\text{Mah}} = 9$, $J_{\max} = 10$ (see Table 5.4) and $\omega_{\text{Th}} = 0.5$ (see Table 5.5). From the position estimates in Figure 5.4, it can be seen that the GM-PHD filter for extended targets provides accurate tracking performance for this scenario.

The filter does generate the occasional spurious estimate at the source or after a target has moved outside the surveillance region. False estimates at the source can be accounted for by the appearance of more than one measurement lying within the gate centred on $\mathbf{m}_{\text{source}}$ (see Table 5.7), thus generating target births. Likewise for false estimates appearing in areas close to where a target has recently moved beyond the surveillance region. These false estimates occur when enough measurements lie within gates centred on the predicted means of Gaussian components relating to targets whose weights may still be large enough for them to continue their estimated existence. However, both types of false estimates soon disappear when there are no further observations in subsequent iterations to maintain the spurious tracks. Increasing the minimum limit on the size of subset $\tilde{\varphi}$ in Table 5.7, hence introducing

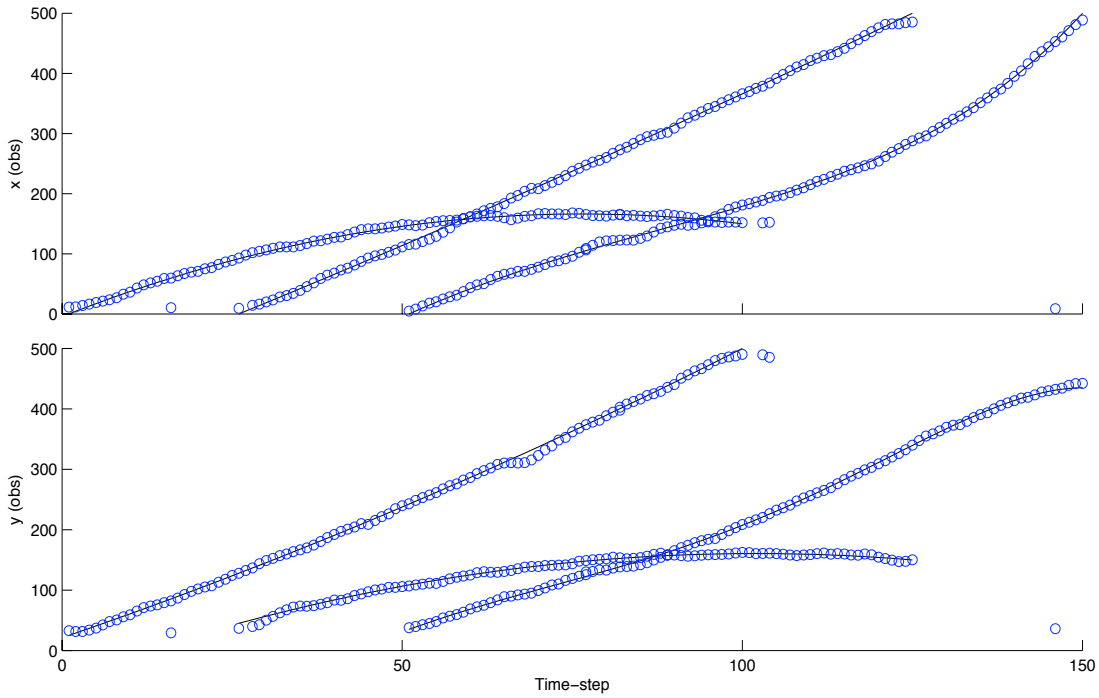


Figure 5.4: Position estimates from the GM-PHD filter for extended targets (blue 'circles') and the trajectories (lines)

a greater restriction on the generation of target births, should eliminate the false estimates appearing at the source altogether.

To assess the performance of the GM-PHD filter for extended targets, an appropriate benchmark is necessary. For this purpose, a Gaussian mixture implementation of a PHD filter similar to that introduced in Section 2.4.4 is considered, a summary of which is given in Appendix E. 50 Monte Carlo (MC) trials were run for each filter with the same trajectories shown in Figure 5.2 but with independently generated measurements for each trial and their performance was evaluated using the OSPA metric with order $p = 1$ and cut-off $c = 75 > \sigma_2^2$ (see Section 2.5.2). The MC average of the OSPA distance, as well as the localisation and cardinality components of the metric, are shown in Figure 5.5.

The OSPA results in Figure 5.5 illustrate how much the benchmark filter is penalised for cardinality errors. It is clear that in a scenario where measurements are generated from feature points which have a geometric structure such as an ellipse, the proposed GM-PHD filter for extended targets in Section 5.1 performs better than the other with regards the estimation of the number of targets. The restriction of the elliptical shaped extent in the cluster process model does however

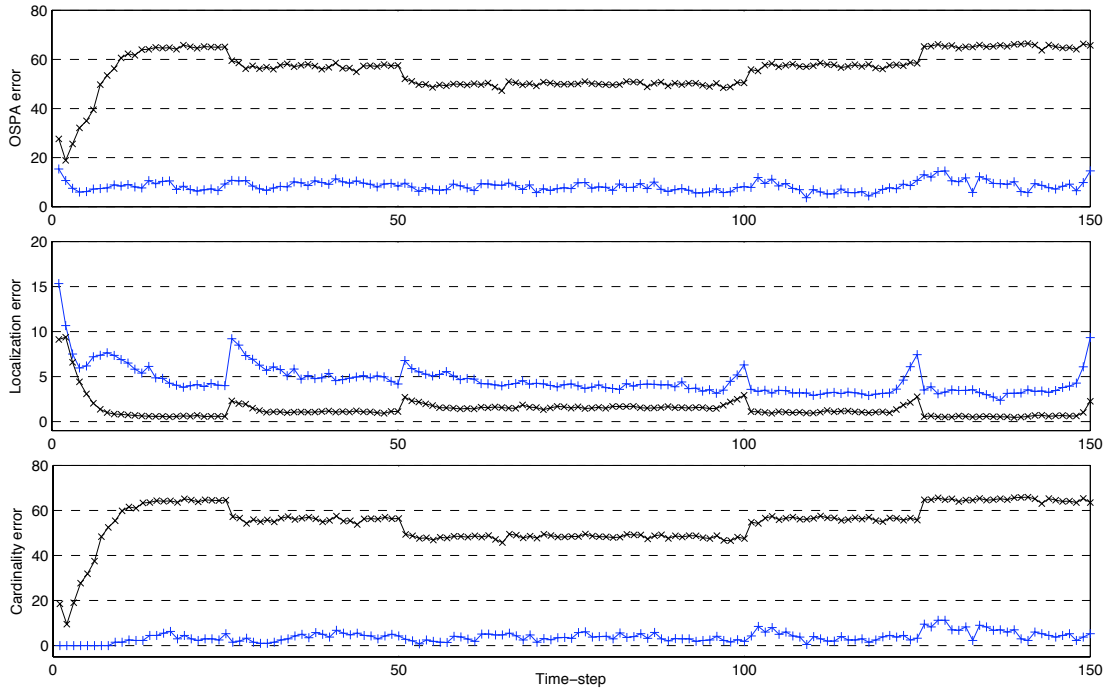


Figure 5.5: The average OSPA error results, with localisation and cardinality components, from 50 MC runs of the GM-PHD filter for extended targets with and without the cluster process representation (blue line '−+−' and black line '−×−' respectively)

result in slightly larger localisation errors for the proposed GM-PHD filter, though the majority of localisation errors for both filters are less the standard deviation of the observation noise (for the parent process). The increases in localisation errors at time intervals $k = 25$ and 50 coincide with the appearance of new targets in the surveillance region and increases leading up to time intervals $k = 100$, 125 and 150 coincide with the approach of existing targets to the edge of the surveillance region.

5.3 Scenario 2: Extended Target Tracking with Variable Extent

The hierarchical RFS approach to extended target tracking not only allows for the construction of geometric shaped extents, improving the estimation capabilities of the PHD filter in scenarios where targets generated measurements that appear in structured arrangements, but also allows for the estimation of variable shape parameters. To do so requires adapting the PHD filter for extended targets as shown in Section 5.3.1, the implementation of which further develops the GM formulation given in Section 5.1 as shown in Section 5.3.2. A simulated example is considered in Section 5.3.3 and the results are presented in Section 5.3.4.

5.3.1 Adapting the PHD filter

Consider the target state $\mathbf{X}_k$, which now consists of the random vector $\mathbf{x}_k$ defining the kinematic state and the random vector $\mathbf{s}_k$ defining the variable shape parameters, i.e. $\mathbf{X}_k = \begin{bmatrix}\mathbf{x}_k & \mathbf{s}_k\end{bmatrix}^T$. The PHD filter for extended targets jointly estimates the kinematic state and variable shape parameters through the propagation of intensities which can be factorised as follows

$$v_k(\mathbf{x}_k, \mathbf{s}_k) = v_k(\mathbf{s}_k) \times v_k(\mathbf{x}_k \,|\, \mathbf{s}_k). \tag{5.14}$$

The factorisation above holds true under the assumption that the estimation of the kinematic state depends on the variable shape parameters and not conversely. This assumption holds for the multiple extended target scenario considered in this section, in which the construction of the geometric shaped extent for each target is partly determined by the variable shape parameters and the measurements generated from the feature points along the extent boundaries update the kinematic states.

Now suppose that the position and shape of each target evolve independently of one another so that the dynamics can be factorised as $f_{k|k-1}(\mathbf{x}_k, \mathbf{s}_k \,|\, \mathbf{x}_{k-1}, \mathbf{s}_{k-1}) =$

$f_{k|k-1}(\mathbf{s}_k \,|\, \mathbf{s}_{k-1}) \times f_{k|k-1}(\mathbf{x}_k \,|\, \mathbf{x}_{k-1})$, where $f_{k|k-1}(\mathbf{s}_k \,|\, \mathbf{s}_{k-1})$ denotes the Markov transition density for the variable shape parameters and $f_{k|k-1}(\mathbf{x}_k \,|\, \mathbf{x}_{k-1})$ denotes the Markov transition density of the kinematic state from one time-step to the next. This is representative of the scenario under consideration, where the variable shape parameters do not include the angle of rotation, which is instead determined by a target's velocities, included in the kinematic state. The **predicted intensity** in the corresponding PHD filter is then given by

$$v_{k|k-1}(\mathbf{x}_k, \mathbf{s}_k) = v_{k|k-1}(\mathbf{s}_k) \times v_{k|k-1}(\mathbf{x}_k \,|\, \mathbf{s}_k), \tag{5.15}$$

where the predicted intensity of the variable shape parameters is

$$v_{k|k-1}(\mathbf{s}_k) = \int v_{k-1}(\mathbf{s}_{k-1})\, f_{k|k-1}(\mathbf{s}_k \,|\, \mathbf{s}_{k-1}) \mathrm{d}\mathbf{s}_{k-1}, \tag{5.16}$$

the predicted intensity of the kinematic state is

$$v_{k|k-1}(\mathbf{x}_k \,|\, \mathbf{s}_k) = \gamma_k(\mathbf{x}_k \,|\, \mathbf{s}_k) + \int p_S(\mathbf{x}_{k-1})\, f_{k|k-1}(\mathbf{x}_k \,|\, \mathbf{x}_{k-1})\, v_{k-1}(\mathbf{x}_{k-1} \,|\, \mathbf{s}_{k-1})\, \mathrm{d}\mathbf{x}_{k-1}, \tag{5.17}$$

which corresponds to the expression given in equation (3.38), with $\gamma_k(\mathbf{x}_k \,|\, \mathbf{s}_k)$ denoting the intensity of birth targets. The probability of target survival $p_S(\mathbf{x}_{k-1})$ is also assumed to be shape independent.

Given a new set of measurements Z_k, the **updated intensity** in the corresponding PHD filter is then given by

$$v_k(\mathbf{x}_k, \mathbf{s}_k) = \frac{L_{Z_k}(\mathbf{x}_k, \mathbf{s}_k)\, v_{k|k-1}(\mathbf{s}_k)\, v_{k|k-1}(\mathbf{x}_k \,|\, \mathbf{s}_k)}{\iint L_{Z_k}(\mathbf{x}_k, \mathbf{s}_k)\, v_{k|k-1}(\mathbf{s}_k)\, v_{k|k-1}(\mathbf{x}_k \,|\, \mathbf{s}_k)\, \mathrm{d}\mathbf{s}_k\, \mathrm{d}\mathbf{x}_k}, \tag{5.18}$$

where the joint pseudo-likelihood is

$$\begin{aligned} L_{Z_k}(\mathbf{x}_k, \mathbf{s}_k) = 1 - p_{D,1}(\mathbf{x}_k, \mathbf{s}_k) + p_{D,1}(\mathbf{x}_k, \mathbf{s}_k) \exp\left(-\mu[p_{D,2} \,|\, \mathbf{x}_k, \mathbf{s}_k]\right) \\ + \sum_{\pi \in \Pi_{Z_k}} \varpi_\pi(\mathbf{s}_k) \sum_{\varphi \in \pi} \frac{p_{D,1}(\mathbf{x}_k, \mathbf{s}_k)\, L_\varphi(\mathbf{x}_k, \mathbf{s}_k)}{\Lambda(\varphi) + v_{k|k-1}\left[p_{D,1}\, L_\varphi \,|\, \mathbf{s}_k\right]}. \end{aligned} \tag{5.19}$$

The joint pseudo-likelihood above corresponds to the expression within the brackets in equation (3.39), except it now updates both the kinematic state and the variable shape parameters. The partition weights, denoted by $\varpi_\pi(\mathbf{s}_k)$ in $L_{Z_k}(\mathbf{x}_k, \mathbf{s}_k)$, are now

given by

$$\varpi_{\pi}(\mathbf{s}_k) = \frac{\prod_{\varphi\in\pi}\Big(\Lambda(\varphi) + v_{k|k-1}\left[p_{D,1}\, L_{\varphi}\,|\,\mathbf{s}_k\right]\Big)}{\sum_{\pi'\in\Pi_{Z_k}}\prod_{\varphi'\in\pi'}\Big(\Lambda(\varphi) + v_{k|k-1}\left[p_{D,1}\, L_{\varphi'}\,|\,\mathbf{s}_k\right]\Big)}, \tag{5.20}$$

and the multiple measurement likelihood is now

$$L_{\varphi}(\mathbf{x}_k, \mathbf{s}_k) = \exp\left(-\mu[p_{D,2}\,|\,\mathbf{x}_k, \mathbf{s}_k]\right)\prod_{\mathbf{z}\in\varphi}\mu\left[p_{D,2}\, l_{\mathbf{z}}\,|\,\mathbf{x}_k, \mathbf{s}_k\right], \tag{5.21}$$

where $\mu[p_{D,2}\,|\,\mathbf{x}_k, \mathbf{s}_k] = \int p_{D,2}(\boldsymbol{\xi})\,\mu(\boldsymbol{\xi}\,|\,\mathbf{x}_k, \mathbf{s}_k)\,\mathrm{d}\boldsymbol{\xi}$ for some arbitrary intensity μ of each subsidiary Poisson RFS $\mathcal{E}_k$ representing the feature points, which is conditioned on both $\mathbf{x}_k$ and $\mathbf{s}_k$. The clutter term $\Lambda(\varphi)$ is as defined in equation (3.20).

If a single partition of the measurement set, denoted by $\boldsymbol{\pi}$, can be determined then the joint pseudo-likelihood reduces to

$$\begin{aligned} L_{Z_k}(\mathbf{x}_k, \mathbf{s}_k) &= 1 - p_{D,1}(\mathbf{x}_k, \mathbf{s}_k) \\ &\quad + p_{D,1}(\mathbf{x}_k, \mathbf{s}_k)\exp\left(-\mu[p_{D,2}\,|\,\mathbf{x}_k, \mathbf{s}_k]\right) + \sum_{\varphi\in\boldsymbol{\pi}}\frac{p_{D,1}(\mathbf{x}_k, \mathbf{s}_k)\, L_{\varphi}(\mathbf{x}_k, \mathbf{s}_k)}{\Lambda(\varphi) + v_{k|k-1}\left[p_{D,1}\, L_{\varphi}\,|\,\mathbf{s}_k\right]}. \end{aligned} \tag{5.22}$$

For simplicity, it will be assumed hereafter that such a partition $\boldsymbol{\pi}$ exists for each measurement set and is known.

5.3.2 Implementation

To implement the PHD filter described in Section 5.3.1, a Dirac mixture model is considered for the intensity of the variable shape parameters and each component of this has associated with a GM formulation for the intensity of the kinematic state conditioned on that particular component. That is, at the beginning of each iteration of the filter, the prior intensities are modelled as follows

$$v_{k-1}(\mathbf{s}_{k-1}) = \sum_{i=1}^{N_{k-1}}\nu_{k-1}^{(i)}\,\delta(\mathbf{s}_{k-1} - \mathbf{s}_{k-1}^{(i)}), \tag{5.23}$$

$$v_{k-1}(\mathbf{x}_{k-1}\,|\,\mathbf{s}_{k-1}^{(i)}) = \sum_{j=1}^{J_{k-1}^{(i)}}\omega_{k-1}^{(j|i)}\,\mathcal{N}\left(\mathbf{x}_{k-1}\,|\,\mathbf{s}_{k-1}^{(i)}; \mathbf{m}_{k-1}^{(j|i)}, \mathbf{P}_{k-1}^{(j|i)}\right). \tag{5.24}$$

The mixture in (5.23) defines a particle representation of $v_{k-1}(\mathbf{s}_{k-1})$ which can be expressed as the set of weighted particles $\{\nu_{k-1}^{(i)}, \mathbf{s}_{k-1}^{(i)}\}_{i=1}^{N_{k-1}}$.

To start with then, N_{k-1} particles are sampled from the Markov transition density describing the change in the shape parameters, i.e. for $i = 1, \ldots, N_{k-1}$, sample $\tilde{\mathbf{s}}_k^{(i)} \sim f_{k|k-1}(\mathbf{s}_k \,|\, s_{k-1}^{(i)})$. For each sampled particle, the predicted and updated intensities for the kinematic state are formulated as GMs of the following forms

$$v_{k|k-1}(\mathbf{x}_k \,|\, \tilde{\mathbf{s}}_k^{(i)}) = \sum_{j=1}^{J_{k|k-1}^{(i)}} \omega_{k|k-1}^{(j|i)} \mathcal{N}\Big(\mathbf{x}_k \,|\, \tilde{\mathbf{s}}_k^{(i)}\,; \mathbf{m}_{k|k-1}^{(j|i)}, \mathbf{P}_{k|k-1}^{(j|i)}\Big)\,, \tag{5.25}$$

$$v_k(\mathbf{x}_k \,|\, \tilde{\mathbf{s}}_k^{(i)}) = \sum_{j=1}^{J_k^{(i)}} \omega_k^{(j|i)} \mathcal{N}\Big(\mathbf{x}_k \,|\, \tilde{\mathbf{s}}_k^{(i)}\,; \mathbf{m}_k^{(j|i)}, \mathbf{P}_k^{(j|i)}\Big)\,, \tag{5.26}$$

where the weights, means and covariances of the respective Gaussian components have the same closed-form expressions as detailed in Tables 5.1 – 5.3, taking into account the slight change in indexing due to the dependency on particle i. Note that the updated intensity for the kinematic state follows from the equation $v_k(\mathbf{x}_k \,|\, \tilde{\mathbf{s}}_k^{(i)}) = L_{Z_k}(\mathbf{x}_k, \tilde{\mathbf{s}}_k^{(i)})\, v_{k|k-1}(\mathbf{x}_k \,|\, \tilde{\mathbf{s}}_k^{(i)})$, where $L_{Z_k}(\mathbf{x}_k, \tilde{\mathbf{s}}_k^{(i)})$ is the joint pseudo-likelihood found in equation (5.18) and given by equation (5.22). The weights in the Dirac mixture are then updated as follows

$$\tilde{\nu}_k^{(i)} = \frac{L_{Z_k}(\tilde{\mathbf{s}}_k^{(i)})}{\sum_{i=1}^{N_{k-1}} L_{Z_k}(\tilde{\mathbf{s}}_k^{(i)})}\, \nu_{k-1}^{(i)}, \tag{5.27}$$

where $L_{Z_k}(\tilde{\mathbf{s}}_k^{(i)})$ is determined by summing the weights in the mixture for the updated intensity $v_k(\mathbf{x}_k \,|\, \tilde{\mathbf{s}}_k^{(i)})$. Finally, a resampling procedure is applied, which makes N_k copies of $\{\tilde{\nu}_k^{(i)}, \tilde{\mathbf{s}}_k^{(i)}\}_{i=1}^{N_{k-1}}$ and eliminates low-weighted particles to give $\{\nu_k^{(i)}, \mathbf{s}_k^{(i)}\}_{i=1}^{N_k}$.

5.3.3 Simulation set-up

The implementation described in Section 5.3.2 applies to extended targets whose extents have any structural shape in general. For illustration purposes, the same multiple extended target scenario from Section 5.2.1 is considered in which the elliptical extents now gradually reduce in size over time, as shown for one of the trajectories in Figure 5.6.

Each target state $\mathbf{X}_k$ consists of the kinematic state $\mathbf{x}_k = \begin{bmatrix} x_{k,1} & x_{k,2} & x_{k,3} & x_{k,4} \end{bmatrix}^T$ and variable shape parameters $\mathbf{s}_k = \begin{bmatrix} s_{k,1} & s_{k,2} & s_{k,3} & s_{k,4} \end{bmatrix}^T$, where $\begin{bmatrix} x_{k,1} & x_{k,3} \end{bmatrix}^T$ is the position vector, $\begin{bmatrix} x_{k,2} & x_{k,4} \end{bmatrix}^T$ is the velocity vector, $s_{k,1}$ & $s_{k,3}$ are the magnitudes of

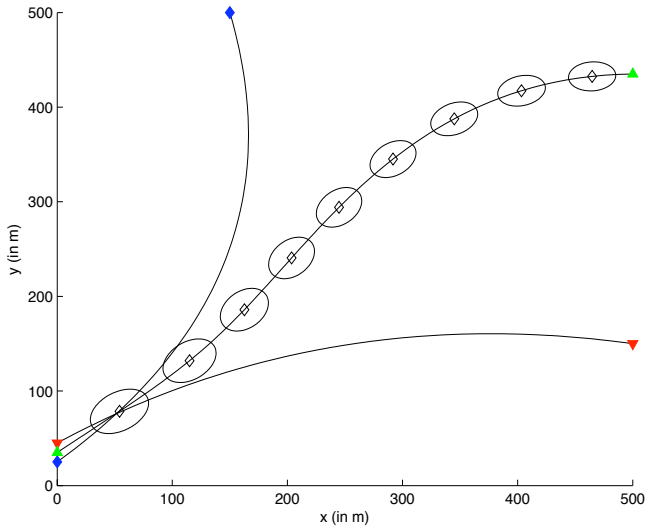


Figure 5.6: The multiple target trajectories ('solid' and 'dashed' lines) and an example of the size varying elliptical extents. Start and end positions are indicated by: blue 'diamonds' for Trajectory 1 at $k = 1$ and $k = 100$; 'red 'downward triangles' for Trajectory 2 at $k = 26$ and $k = 150$; and green 'upward triangles' for Trajectory 3 at $k = 51$ and $k = 125$.

the major and minor axes of the target's elliptical extent respectively, and $s_{k,2}$ & $s_{k,4}$ are the rates of change in the major and minor axes' magnitudes respectively.

Targets follow the same linear Gaussian dynamics given by equation (4.8) with transition matrix $\mathbf{F}_{k-1}$ $(= \mathbf{F}_{22})$ and process noise covariance matrix $\mathbf{Q}_{k-1}$ $(= \mathbf{Q}_{22})$ as defined in equations (5.10)–(5.11). The probability of survival is assumed to be state independent and has the following constant value $p_S = 0.9$.

Point measurements $\mathbf{z} = \begin{bmatrix} z_1 & z_2 \end{bmatrix}^T$ are simulated at each iteration using the single measurement likelihood given in equation (4.15) and from the hierarchical RFS model, whose probability density is given by equations (3.19) and (3.20), for false measurements. The projection matrices $\mathbf{H}$, $\widehat{\mathbf{H}}$ and the observation noise covariance matrices $\mathbf{R}_1$, $\mathbf{R}_2$ have the same definitions as given in Section 5.2.1 and the corresponding standard deviations now have the following values $\sigma_1 = \sqrt{5}\,\text{m}$ and $\sigma_2 = \sqrt{32}\,\text{m}$. The number of feature points for each target is $\alpha = 11$ and the probabilities of detecting targets and their feature points are $p_{D,1} = p_{D,2} = 0.99$. The intensities in the clutter term (3.20), κ_1 and κ_2, are modelled in the same way as given in Section 5.2.1, with the same Poisson rates. Figure 5.7 shows the cluttered measurements plotted with the target trajectories against time.

The filter is initialised with the birth intensity $\gamma_k(\mathbf{x}_k \,|\, \tilde{\mathbf{s}}_k^{(i)})$ at time-step $k = 1$, whose Gaussian components are driven by the measurements received at that time-step and specifically determined from those subsets $\varphi \in \boldsymbol{\pi}$, where $\boldsymbol{\pi}$ is the known

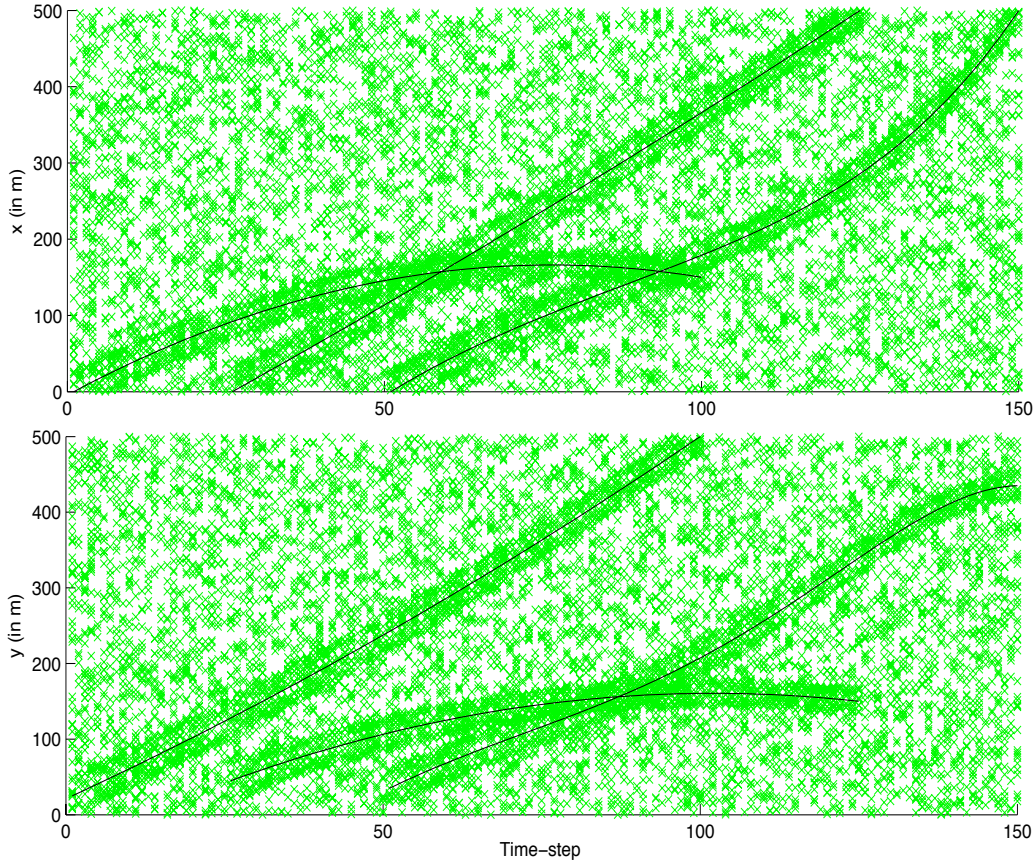


Figure 5.7: Measurements (green 'crosses') and target trajectories (lines)

partition of the measurement set Z_k, that contain more than one measurement. That is, for $j = 1, \ldots, J_{\gamma,k}^{(i)}$, where $J_{\gamma,k}^{(i)} = |\tilde{\boldsymbol{\pi}}|$ and $\tilde{\boldsymbol{\pi}} = \{\varphi \in \boldsymbol{\pi} : |\varphi| > 1\}$, $\omega_{\gamma,k}^{(j|i)} = p_B = 0.1$, $\mathbf{P}_{\gamma,k}^{(j|i)} = \text{diag}([5\sigma_2^2, 2, 5\sigma_2^2, 2])$ and

$$\mathbf{m}_{\gamma,k}^{(j|i)} = \left[\frac{1}{|\tilde{\varphi}_j|} \sum_{\mathbf{z} \in \tilde{\varphi}_j} z_1 \quad 0 \quad \frac{1}{|\tilde{\varphi}_j|} \sum_{\mathbf{z} \in \tilde{\varphi}_j} z_2 \quad 0 \right]^T, \tag{5.28}$$

for each particle i.

The variable shape parameters for new targets are sampled from a truncated linear Gaussian. Specifically, sample $\tilde{\mathbf{c}}_k^{(i)} = \begin{bmatrix} \tilde{c}_{k,1}^{(i)} & \tilde{c}_{k,2}^{(i)} \end{bmatrix}^T \sim \mathcal{N}(\mathbf{c}_k\,; \boldsymbol{\mu}_{\text{initial}}, \boldsymbol{\Sigma}_{\text{initial}})$ such that $\tilde{\mathbf{c}}_k^{(i)}$ lies within the interval $(\mathbf{a}_{\text{initial}}, \mathbf{b}_{\text{initial}})$, given the initial mean $\boldsymbol{\mu}_{\text{initial}} = \mathbf{b}_{\text{initial}}$ and initial covariance $\boldsymbol{\Sigma}_{\text{initial}} = \sigma_{\text{initial}}^2 \mathbf{I}_2$, where $\mathbf{a}_{\text{initial}} = \begin{bmatrix} 19 & 14 \end{bmatrix}^T$, $\mathbf{b}_{\text{initial}} = \begin{bmatrix} 32 & 21 \end{bmatrix}^T$ and $\sigma_{\text{initial}}^2 = \sqrt{2}\,\text{m}$, to give $\tilde{\mathbf{s}}_k^{(i)} = \begin{bmatrix} \tilde{c}_{k,1}^{(i)} & 0 & \tilde{c}_{k,2}^{(i)} & 0 \end{bmatrix}^T$. Similarly for persistent targets, sample $\tilde{\mathbf{s}}_k^{(i)} \sim \mathcal{N}(\mathbf{s}_k\,; \mathbf{F}_{k-1}\,\mathbf{s}_{k-1}^{(i)}, \widetilde{\mathbf{Q}}_{k-1})$ such that $\tilde{\mathbf{s}}_k^{(i)}$ lies within the interval $(\mathbf{a}, \mathbf{b})$ where $\mathbf{a} = \begin{bmatrix} 19 & -0.2 & 14 & -0.1 \end{bmatrix}^T$, $\mathbf{b} = \begin{bmatrix} 32 & 0 & 21 & 0 \end{bmatrix}^T$, $\mathbf{F}_{k-1}$ is

as defined in equation (5.10) and

$$\widetilde{\mathbf{Q}}_{k-1} = \mathbf{\Gamma}_{k-1}\,\mathbf{\Sigma}_{\mathbf{s}}\,\mathbf{\Gamma}_{k-1}^{T}, \qquad \mathbf{\Sigma}_{\mathbf{s}} = \mathrm{diag}([\sigma_{\mathbf{s},1}^{2}, \sigma_{\mathbf{s},2}^{2}, \sigma_{\mathbf{s},3}^{2}, \sigma_{\mathbf{s},4}^{2}]), \tag{5.29}$$

for $\sigma_{\mathbf{s},1} = \sqrt{2}\,\mathrm{m}$, $\sigma_{\mathbf{s},2} = \sqrt{0.2}\,\mathrm{m}$, $\sigma_{\mathbf{s},3} = \sqrt{1}\,\mathrm{m}$, $\sigma_{\mathbf{s},4} = \sqrt{0.1}\,\mathrm{m}$ and $\mathbf{\Gamma}_{k-1}$ as defined in equation (5.11).

Finally, for each particle i and kinematic state j, the arbitrary intensity $\mu(\boldsymbol{\xi}_k, \mathbf{x}_k)$ has a similar Gaussian mixture form to that given in equation (5.2) with $J_{\boldsymbol{\xi}}^{(j|i)} = \alpha$ and for $t = 1, \ldots, J_{\boldsymbol{\xi}}^{(j|i)}$ the means $\mathbf{m}_{\boldsymbol{\xi}}^{(t|i,j)}$ and covariances $\mathbf{P}_{\boldsymbol{\xi}}^{t|i,j)}$ have equivalent expressions to those given in (5.12) in which the cross diagonal covariance matrices are given by $\mathbf{P}_{\boldsymbol{\xi}|\mathbf{x}}^{(t|i,j)} = \widehat{\mathbf{H}}\mathbf{P}_{k|k-1}^{(j|i)}$. Once again, the means $\mathbf{m}_{\boldsymbol{\xi}}^{(t|i,j)}$ relate to the feature points distributed along the perimeter of the target's elliptical shaped extent at fixed intervals now where

$$\mathbf{r}_{k|k-1}^{(t|i,j)} = \begin{bmatrix} \tilde{s}_{k,1}^{(i)} \cos(\theta_t)\cos(\phi_{k|k-1}^{(j|i)}) + \tilde{s}_{k,3}^{(i)} \sin(\theta_t)\sin(\phi_{k|k-1}^{(j|i)}) \\ \tilde{s}_{k,1}^{(i)} \cos(\theta_t)\sin(\phi_{k|k-1}^{(j|i)}) + \tilde{s}_{k,3}^{(i)} \sin(\theta_t)\cos(\phi_{k|k-1}^{(j|i)}) \end{bmatrix}, \tag{5.30}$$

with $\theta_t = 2\pi t/\alpha$ and $\phi_{k|k-1}^{(j)} = \arctan(m_{k|k-1,4}^{(j)}/m_{k|k-1,2}^{(j)})$, which is equivalent to the angle of rotation specified in Section 5.2.1.

5.3.4 Results

The implementation outlined in Section 5.3.2 for the PHD filter presented in Section 5.3.1 is applied to the scenario described in Section 5.3.3 using, in addition, the pruning algorithm given in Table 5.4 for merging closely spaced Gaussian components with parameters $\tau = 10^{-4}$, $U_{\mathrm{Mah}} = 9$ and $J_{\max} = 10$. The estimation results shown in Figure 5.8 are from a simulation run with 500 particles for each target.

At each iteration, the state and error estimates are obtained by applying a similar clustering technique as demonstrated in [31] whereby the weights, means and covariances of the Gaussian components resulting from target detections in v_k are re-indexed for each subset $\varphi \in \boldsymbol{\pi}$, i.e. $\omega_k^{(n|i)}$, $\mathbf{m}_k^{(n|i)}$ and $\mathbf{P}_k^{(n|i)}$ for $n = 1, \ldots, N_{\varphi}^{(i)}$ where $N_{\varphi}^{(i)} = J_{k|k-1}^{(i)} \times |\varphi| \times \alpha$ and $i = 1, \ldots, N_{k-1}$. The corresponding weights for each subset φ are then summed to give

$$W_{\varphi,k} = \sum_{i=1}^{N_{k-1}} \sum_{n=1}^{N_{\varphi}} \omega_k^{(n|i)}, \tag{5.31}$$

and if $W_{\varphi,k} \geq T$, for some threshold T, the position and shape parameter estimates

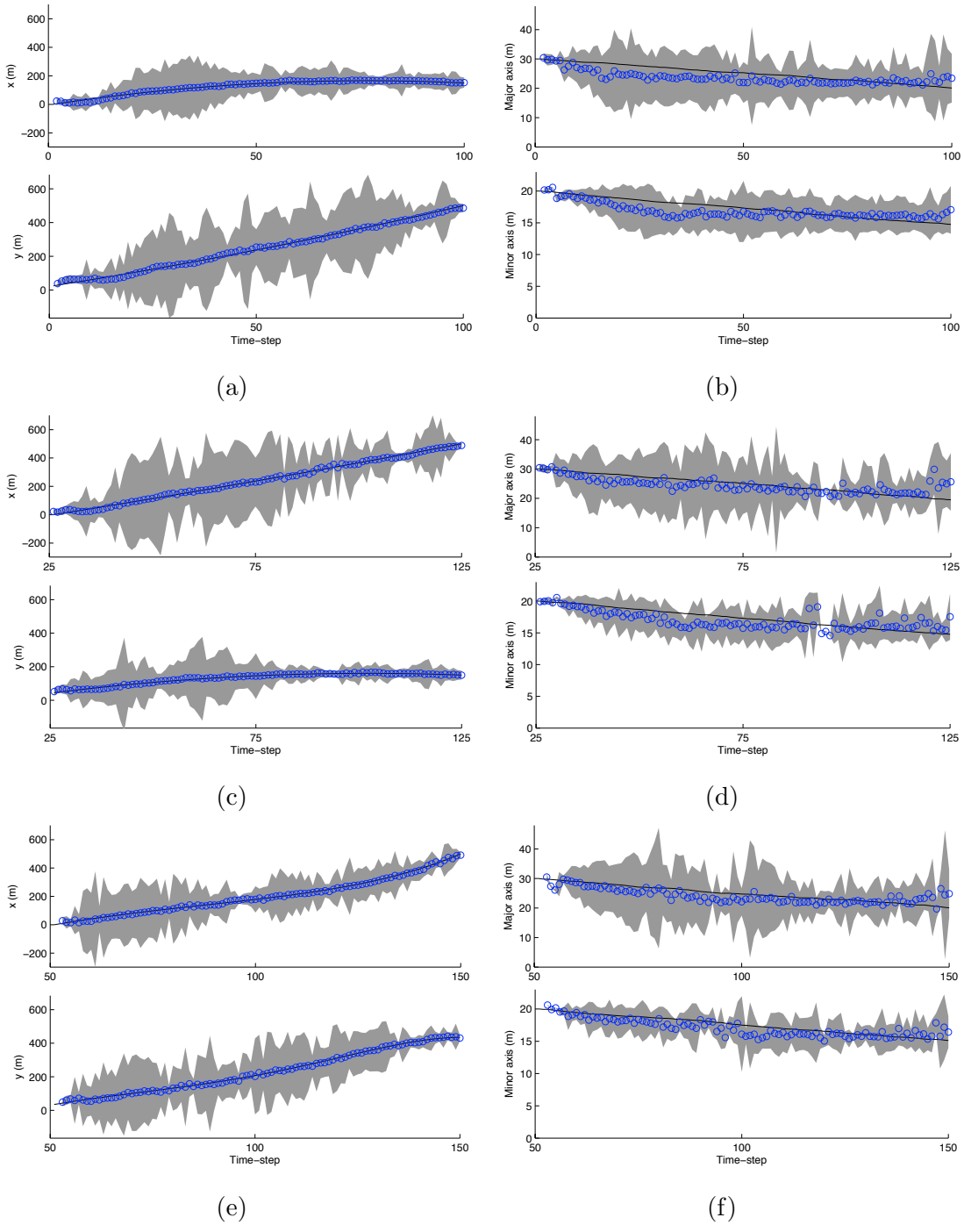


Figure 5.8: Position and shape parameter estimates (blue 'circles'), uncertainty (grey shaded area) and target trajectories (line).

are calculated respectively as

$$\widehat{\mathbf{m}}_{\varphi,k} = \sum_{i=1}^{N_{k-1}} \sum_{n=1}^{N_\varphi} \omega_k^{(n|i)} \, \mathbf{m}_k^{(n|i)}, \qquad \hat{\mathbf{s}}_{\varphi,k} = \sum_{i=1}^{N_{k-1}} \sum_{n=1}^{N_\varphi} \omega_k^{(n|i)} \, \tilde{\mathbf{s}}_k^{(i)}, \tag{5.32}$$

and the corresponding errors are calculated as

$$\begin{aligned} \widehat{\mathbf{P}}_{\varphi,k} &= \sum_{i=1}^{N_{k-1}} \sum_{n=1}^{N_\varphi} \omega_k^{(n|i)} \left(\mathbf{P}_k^{(n|i)} + (\mathbf{m}_k^{(n|i)} - \widehat{\mathbf{m}}_{\varphi,k})(\mathbf{m}_k^{(n|i)} - \widehat{\mathbf{m}}_{\varphi,k})^T \right), \\ \hat{\mathbf{s}}_{\varphi,k}^{\text{err}} &= \sum_{n=1}^{N_{k-1}} \sum_{n=1}^{N_\varphi} \omega_k^{(n|i)} (\tilde{\mathbf{s}}_k^{(i)} - \hat{\mathbf{s}}_{\varphi,k})(\tilde{\mathbf{s}}_k^{(i)} - \hat{\mathbf{s}}_{\varphi,k})^T. \end{aligned} \tag{5.33}$$

Tracks are maintained by labelling each particle set corresponding to a target.

Figure 5.8 shows early results for the position estimates (5.8a, 5.8c & 5.8e) and the shape parameter estimates (5.8b, 5.8d & 5.8f). Where the targets approach the boundary of the observation region towards the end of their existence, so that the extents are only partially visible, the estimation accuracy suffers as a result. Overall the accuracy of this approach, particularly in estimating the shape parameter, is reliant on the choice of sampling distribution and the truncated Gaussian used to sample particles for this simulated example is perhaps too restrictive. Better results could be acheived with the use of random matrices as incorporated in the PHD filtering methods for the extended target tracking problem by Granström, Orguner et. al. [127, 128]. This will discussed further in the following concluding chapter.

Chapter 6

Discussion

6.1 Thesis Summary

This thesis focused on extending the PHD filter to multiple group and extended target tracking problems. A fundamental concept in point process theory, known as cluster processes, was explored in Chapter 3 from which hierarchical RFS formalisms were introduced to represent the multiple group/extended target system. These so called cluster models allow for the generalisation of the Bayesian filter to group and extended target tracking problems and their first-order moments provide suitable approximations for practical implementation when recursively propagated in the extension of the PHD filter to such tracking problems. The results presented in Chapter 3 for the filtering equations in the recursion of the PHD filter for group targets, so called the *multi-group PHD filter*, along with their derivation in Appendix B, constituted the principal contribution of this thesis. The filtering equations of three special cases, including the PHD filter for extended targets, were also explicitly derived.

A closed-form solution to the multi-group PHD filter was proposed in Chapter 4, under certain linear Gaussian conditions, which resulted in a Gaussian mixture formulation. The implementation of a similar GM formulation for the single group special case was demonstrated with a simulated example and shown to perform better than simply using the multi-target GM-PHD filter [90], taking an aggregated estimate of the group state.

Chapter 5 concerned the implementation of the PHD filter for extended targets, a special case of the multi-group PHD filter. The closed-form solution presented in Chapter 5 is subsequently a special case of the GM formulation proposed in Chapter 4. Its implementation was demonstrated with a simulated example, in which the multiple detections for each target have a structured appearance, and shown to perform better than a GM implementation of an alternative PHD filter for non-standard targets (see Appendix E), similar to that proposed in [101]. A novel

implementation of the PHD filter for extended targets was also proposed in Chapter 5 which incorporated variable shape parameters for target extent estimation using a particle representation.

6.2 Current Research

The PHD filters presented in this thesis are not only just applicable to group and extended target tracking problems. In particular, the single group PHD filter has been applied to the Simultaneous Localisation and Mapping (SLAM) problem by Lee, Clark et. al. [129, 130]. The hierarchical RFS representation in this case consists of a set (map) of multiple landmarks conditioned on a single vehicle state. The single group PHD filter, referred to by the authors as the single cluster PHD filter, is implemented using a combination of Dirac mixture and Gaussian mixture models for the vehicle state and landmark estimations respectively. This approach has since been extended by the same authors to include dynamic targets amongst the multiple landmarks in [131].

Similarly, Ristic and Clark et. al. have recently demonstrated a further application of the single group (cluster) PHD filter in a collection of papers [132, 133, 134], namely for the joint estimation of multiple dynamic targets and a multi-sensor bias. The underlying process in the hierarchical RFS representation consists of a single (multi-sensor bias) state vector on which the set of multiple targets are conditioned. The authors adopt a particle implementation of the single cluster PHD filter using a Dirac mixture model for both the multi-sensor bias state and the target state estimations.

Finally, the derivation of the multi-group PHD update requires the use of an induction proof as alluded to to in Appendix B (Proof of Lemma B.2, Section B.3.2). However, recently, a more general approach to PHD filtering has been introduced by Clark and Mahler [135], which removes the need for an induction proof. A general chain rule is introduced in [135], based on the general version of Faà di Bruno's formula for variational calculus proposed by Clark and Houssineau [136, 137], which provides a more direct approach to deriving PHD filter results. This is demonstrated in [135] where general PHD filters are derived using the general chain rule.

6.3 Future Research Directions

This section discusses interesting directions for future work that moves beyond the scope of this thesis. The majority of the suggested research directions involve ex-

tensions and adaptations of existing or newly introduced work from other authors.

6.3.1 Extensions to non-linear models

The closed-form solution to the multi-group PHD filter proposed in this thesis, as well as the single group and extended target special cases, is based on linear Gaussian assumptions for the dynamic and measurement models. This section discusses possible extensions of the multi-group GM-PHD filter to non-linear models, principally, but the same extensions apply to the single group and extended target special cases. Specifically, this would require the state independent assumptions on the probabilities of survival and detection to remain as specified in Sections 4.1.2 and 4.1.4 respectively, as well as the same GM formulation for the birth intensities specified in Section 4.1.3. The extension to non-linearity occurs in the dynamic and measurement model equations which could be expressed as non-linear functions with zero-mean Gaussian process and observation noise whose covariances are still defined as the matrices $\mathbf{Q}_{11}$, $\mathbf{Q}_{22}$ and $\mathbf{R}_{11}$, $\mathbf{R}_{22}$ respectively. This would be straightforward enough for the group dynamics and measurement model, but extra care may need to be taken with such non-linear individual target dynamics due to the dependency on the group dynamics and the additional process noise with covariance $\mathbf{Q}_{12}$. The multi-group GM-PHD filter could then be adapted to accommodate such non-linear Gaussian models following the extended Kalman or unscented Kalman approximations in a similar manner to that proposed for the multi-target GM-PHD filter in [90].

For extensions to non-linear non-Gaussian models, particle representations using SMC techniques would need to be considered. To consider a full particle representation for the multiple group target problem, where the individual target dynamics are dependent on the group dynamics, would be extremely computationally demanding. However, for some applications which consider independent dynamics, a hybrid implementation involving a combination of particle representation and GM formulation is possible, as shown for the SLAM problem in [129].

6.3.2 Incorporating unknown detection profile

In a recent paper by Mahler, Vo & Vo [138] unknown detection probability was incorporated into PHD/CPHD filtering by augmenting the target states to accommodate non-uniform detection profiles. Estimation of the augmented target states via the expanded PHD or CPHD recursions yields information on the number of targets, as well as the individual kinematic states and the unknown detection prob-

ability for each target. Analytic implementations of such PHD/CPHD recursions are acheived by modelling the augmented part with Beta distributions, resulting in so-called Beta Gaussian mixtures.

The same approach can be applied to incorporate unknown detection probability into PHD filtering for group and extended targets, and therefore relaxing the state independent assumptions on the probabilities of detection in the proposed GM implementations. Each group and individual target state would need to be augmented separately, which may intuitively lead to Beta-Beta-Gaussian mixture formulations. A similar approach could also be applied to instead incorporate unknown survival probability with appropriately augmented states.

6.3.3 Shape estimation using random matrices

The implementation of the PHD filter for extended targets which incorporated shape estimation, proposed in Chapter 5 (Section 5.3), had one minor flaw. That was identified as the restriction imposed by the choice of sampling distribution for the particle representation. One method worth investigating, to improve on the results presented in this thesis, involves modelling the shape parametric as a random matrix. With the use of random matrices, the shape parametric can be represented by an inverse Wishart distribution and can be incorporated into an implementation of the PHD filter for extended targets with shape estimation in two possible ways.

Firstly, following the implementation proposed in Section 5.3, the inverse Wishart representation for the shape parametric can be used as the sample distribution. This approach would require checking that the assumptions made in the adaption of the PHD filter still held for a shape paramteric modelled as a random matrix. Secondly, the inverse Wishart distribution can simply be incorporated into the GM-PHD filter for extended targets presented in Section 5.2, following a similar application presented by Granström, Orguner et. al. in [127]. This latter approach allows the assumption that the kinematic and shape parametric dynamics are independent, which was imposed for the implementation presented in Section 5.3, to be relaxed. Either way the incorporation of a random matrix model for the shape parametric would potentially yield better estimation results.

6.3.4 Group classification and identification

Currently the framework for group target tracking presented in this thesis does not consider spawining of targets, either group or individual targets. Since the

PHD filter for multiple target tracking already considers spawning [1], introducing individual target spawning in the multi-group PHD filter provides little in the way novel contribution to tracking research. Considering the spawning of whole groups on the other hand may be of more interest in terms of future research. This is representative of scenario in which members of an existing group separate from the rest and form a new group. The introduction of group spawning leads the concept of group classification, i.e. determining the difference between group targets.

The final interesting research direction to mention is based on the concept of group identification. That is, identifying when a number of individual targets begin to share a common characteristic so that the collection of targets can be regarded as a group target. The main point of interest attached to idea of group identification, specifically with regards to the contents of this thesis, would be how to model the transition from the use of the multi-target PHD filter to the single/multi-group PHD filter or hybrid of both.

Appendix A

Additional Background Material

A.1 Dirac Delta Functions

The Dirac delta concentrated at $\mathbf{x}$ is a function, denoted $\delta_{\mathbf{x}}(\mathbf{y})$, with the following major properties. If $\mathbf{y} \neq \mathbf{x}$, then $\delta_{\mathbf{x}}(\mathbf{y}) = 0$ otherwise

$$\int h(\mathbf{y})\, \delta_{\mathbf{x}}(\mathbf{y})\, \mathrm{d}\mathbf{y} = h(\mathbf{x}), \tag{A.1}$$

for any function $h(\cdot)$ continuous at $\mathbf{y} = \mathbf{x}$. It satisfies

$$\int_S \delta_{\mathbf{x}}(\mathbf{y})\, \mathrm{d}\mathbf{y} = \mathbf{1}_S(\mathbf{x}), \qquad \text{and} \qquad \int \delta_{\mathbf{x}}(\mathbf{y})\, \mathrm{d}\mathbf{y} = 1, \tag{A.2}$$

where $\mathbf{1}_S(\cdot)$ denotes the indicator function which is defined as

$$\mathbf{1}_S(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in S, \\ 0 & \text{if } \mathbf{x} \notin S. \end{cases} \tag{A.3}$$

A.2 Campbell's Theorem [106, page 103]

Let h be a measurable function on $\mathbb{R}^d$. Then for any RFS Ξ with instance $\Xi = Y$, probability density p_Ξ and first-order moment density v_Ξ, the expected value of the sum of $h(\mathbf{y})$ over $\mathbf{y} \in Y$ is given by

$$\mathsf{E}\left[\sum_{\mathbf{y}\in Y} h(\mathbf{y})\right] = \int \left(\sum_{\mathbf{y}\in Y} h(\mathbf{y})\right) p_\Xi(Y)\, \delta Y = \int h(\mathbf{y})\, v_\Xi(\mathbf{y})\, \mathrm{d}\mathbf{y}. \tag{A.4}$$

The expression for the first-order moment density of a RFS given in equation (2.30) follows as direct application of the result in equation (A.4) with $h(\mathbf{y}) = \delta_{\mathbf{x}}(\mathbf{y})$.

A.3 Expressions for the PHD of a RFS and its Integral

The specific formula for the first-order moment density of a RFS Ξ given in equation (2.31) can be established by applying the formula for a set integral (see Section 2.2.2) to the expression in equation (2.30), then expanding the sum of Dirac delta functions and making use of the properties in Section A.1 as follows.

$$\begin{aligned}
v_{\Xi}(\mathbf{x}) &= \int \left(\sum_{\mathbf{w} \in X} \delta_{\mathbf{x}}(\mathbf{w}) \right) p_{\Xi}(X)\, \delta X \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \int \left(\delta_{\mathbf{x}}(\mathbf{x}_1) + \cdots + \delta_{\mathbf{x}}(\mathbf{x}_n) \right) p_{\Xi}(\{\mathbf{x}_1, \ldots, \mathbf{x}_n\})\, \mathrm{d}\mathbf{x}_1 \cdots \mathrm{d}\mathbf{x}_n \\
&= \sum_{n=1}^{\infty} \frac{n}{n!} \int p_{\Xi}(\{\mathbf{x}, \mathbf{w}_1, \ldots, \mathbf{w}_{n-1}\})\, \mathrm{d}\mathbf{w}_1 \cdots \mathrm{d}\mathbf{w}_{n-1} \\
&= \sum_{n=1}^{\infty} \frac{1}{(n-1)!} \int p_{\Xi}(\{\mathbf{x}, \mathbf{w}_1, \ldots, \mathbf{w}_{n-1}\})\, \mathrm{d}\mathbf{w}_1 \cdots \mathrm{d}\mathbf{w}_{n-1} \\
&= \sum_{i=0}^{\infty} \frac{1}{i!} \int p_{\Xi}(\{\mathbf{x}, \mathbf{w}_1, \ldots, \mathbf{w}_i\})\, \mathrm{d}\mathbf{w}_1 \cdots \mathrm{d}\mathbf{w}_i = \int p_{\Xi}(\{\mathbf{x}\} \cup W)\, \delta W. \qquad \text{(A.5)}
\end{aligned}$$

Further, the statement for the integral of v_{Ξ} in any region $S \subseteq \mathfrak{X}$ given in equation (2.32) can be proven as follows. Note that

$$\mathsf{E}\left[|\Xi \cap S|\right] = \int |X \cap S|\, p_{\Xi}(X)\, \delta X = \int \left(\sum_{\mathbf{x} \in X} \mathbf{1}_S(\mathbf{x}) \right) p_{\Xi}(X)\, \delta X, \qquad \text{(A.6)}$$

where $\mathbf{1}_S(\mathbf{x})$ is the indicator function defined in Section A.1. Applying the formula for a set integral to the expression in equation (A.6) and expanding the sum of indicator functions gives

$$\begin{aligned}
\mathsf{E}\left[|\Xi \cap S|\right] &= \sum_{n=0}^{\infty} \frac{1}{n!} \int \left(\mathbf{1}_S(\mathbf{x}_1) + \cdots + \mathbf{1}_S(\mathbf{x}_n) \right) p_{\Xi}(\{\mathbf{x}_1, \ldots, \mathbf{x}_n\})\, \mathrm{d}\mathbf{x}_1 \cdots \mathrm{d}\mathbf{x}_n \\
&= \sum_{n=1}^{\infty} \frac{n}{n!} \int \mathbf{1}_S(\mathbf{x})\, p_{\Xi}(\{\mathbf{x}, \mathbf{w}_1, \ldots, \mathbf{w}_{n-1}\})\, \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{w}_1 \cdots \mathrm{d}\mathbf{w}_{n-1} \\
&= \sum_{n=1}^{\infty} \frac{1}{(n-1)!} \int \mathbf{1}_S(\mathbf{x})\, p_{\Xi}(\{\mathbf{x}, \mathbf{w}_1, \ldots, \mathbf{w}_{n-1}\})\, \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{w}_1 \cdots \mathrm{d}\mathbf{w}_{n-1} \\
&= \int \mathbf{1}_S(\mathbf{x}) \left(\sum_{j=0}^{\infty} \frac{1}{j!} \int p_{\Xi}(\{\mathbf{x}, \mathbf{w}_1, \ldots, \mathbf{w}_j\})\, \mathrm{d}\mathbf{w}_1 \cdots \mathrm{d}\mathbf{w}_j \right) \mathrm{d}\mathbf{x}.
\end{aligned}$$

Then the statement in equation (2.32) immediately follows as a result of the last line in equation (A.5), i.e.

$$\mathsf{E}\left[|\Xi \cap S|\right] = \int \mathbf{1}_S(\mathbf{x}) \left(\int p_\Xi(\{\mathbf{x}\} \cup W)\, \delta W \right) \mathrm{d}\mathbf{x} = \int_S v_\Xi(\mathbf{x})\, \mathrm{d}\mathbf{x}.$$

A.4 Probability Generating Functionals

The probability generating functional (p.g.fl.) plays a fundamental role in the study of point processes, the early use of which was demonstrated by Moyal in [107] for stochastic population processes. The RFS analogue plays an important role in the derivation of the PHD filter recursion, which was presented in Section 2.4.1, and is defined as follows.

Let $X \subseteq \mathfrak{X}$ be any finite set and let $h(\mathbf{x})$ be a test function defined on $\mathfrak{X}$ such that $0 \leq h(\mathbf{x}) \leq 1$. Then the *probability generating functional* of a RFS Ξ whose probability density is $p_\Xi(X)$ is given by the following expectation value

$$G_\Xi[h] = \mathsf{E}\left[\prod_{\mathbf{x} \in X} h(\mathbf{x}) \right] = \int \left(\prod_{\mathbf{x} \in X} h(\mathbf{x}) \right) p_\Xi(X)\, \delta X, \tag{A.7}$$

where the function $\prod_{\mathbf{x} \in X} h(\mathbf{x})$ is symmetric and measurable. The p.g.fl. $G_\Xi[h]$ has the following properties:

$$G_\Xi[h] = \begin{cases} p_\Xi(\emptyset) & \text{for } h(\mathbf{x}) = 0,\ \forall\, \mathbf{x} \in X; \\ \displaystyle\int p_\Xi(X)\, \delta X = 1 & \text{for } h(\mathbf{x}) = 1,\ \forall\, \mathbf{x} \in X; \\ \displaystyle\int_S p_\Xi(X)\, \delta X = \Pr(\Xi \subseteq S) & \text{for } h(\mathbf{x}) = \mathbf{1}_S(\mathbf{x}),\ \forall\, \mathbf{x} \in X. \end{cases} \tag{A.8}$$

Finally, consider the union $\Xi = \Xi_1 \cup \cdots \cup \Xi_n$ where $\Xi_1, \ldots, \Xi_n$ are statistically independent. Then the p.g.fl. G_Ξ satisfies the property, for all h,

$$G_\Xi[h] = G_{\Xi_1}[h] \cdots G_{\Xi_n}[h], \tag{A.9}$$

The probability generating functionals for some of the RFS classes presented in Section 2.2.4 can easily be determined by applying the formula for set integrals to equation (A.7), followed by direct substitution of the respective probability densities given by equation (2.20)–(2.22). For example, the p.g.fl. of an ***i.i.d.***

cluster RFS Ξ is found to be

$$\begin{aligned} G_{\Xi}[h] &= \sum_{n=0}^{\infty} \frac{1}{n!} \int h(\mathbf{x}_1) \cdots h(\mathbf{x}_n)\, p_{\Xi}(\{\mathbf{x}_1, \ldots, \mathbf{x}_n\})\, \mathrm{d}\mathbf{x}_1 \cdots \mathrm{d}\mathbf{x}_n \\ &= \sum_{n=0}^{\infty} \rho(n) \int h(\mathbf{x}_1) \cdots h(\mathbf{x}_n)\, p(\mathbf{x}_1) \cdots p(\mathbf{x}_n)\, \mathrm{d}\mathbf{x}_1 \cdots \mathrm{d}\mathbf{x}_n \\ &= \sum_{n=0}^{\infty} \rho(n) \left(\int h(\mathbf{x})\, p(\mathbf{x})\, \mathrm{d}\mathbf{x} \right)^n . \end{aligned} \tag{A.10}$$

Similarly, the p.g.fl. of a ***Poisson*** RFS Ξ with intensity v is found to be

$$G_{\Xi}[h] = \exp\left(\int (h(\mathbf{x}) - 1)\, v(\mathbf{x})\, \mathrm{d}\mathbf{x} \right). \tag{A.11}$$

and the p.g.fl. of a ***Bernoulli*** RFS Ξ with existence probability ρ is

$$G_{\Xi}[h] = 1 - \rho + \rho \int h(\mathbf{x})\, p(\mathbf{x})\, \mathrm{d}\mathbf{x}. \tag{A.12}$$

A.5 Functional Derivatives

Let $F[h]$ be a functional on $\mathfrak{X}$, whose argument is an ordinary function h. The *functional derivative* of $F[h]$ in the direction η is defined by

$$\delta F[h\,;\eta] = \lim_{\epsilon \to 0^+} \frac{1}{\epsilon} \left(F[h + \epsilon\eta] - F[h] \right). \tag{A.13}$$

The k^{th}-order functional derivative of $F[h]$ in the directions $\eta_1, \ldots, \eta_k$ is defined iteratively by

$$\delta^k F[h\,;\eta_1, \ldots, \eta_k] = \lim_{\epsilon \to 0^+} \frac{1}{\epsilon} \left(\delta^{k-1} F[h + \epsilon\eta_k\,;\eta_1, \ldots, \eta_{k-1}] - \delta^{k-1} F[h\,;\eta_1, \ldots, \eta_{k-1}] \right). \tag{A.14}$$

Importantly, the probability density and first-order moment density of a RFS Ξ can be recovered from the functional derivatives of its p.g.fl. G_{Ξ}. That is, given $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$, the n^{th}-order functional derivative of G_{Ξ} in the directions $\delta_{\mathbf{x}_1}, \ldots \delta_{\mathbf{x}_n}$, evaluated at $h = 0$, gives

$$p_{\Xi}(X) = \delta^n G_{\Xi}[0\,;\delta_{\mathbf{x}_1}, \ldots, \delta_{\mathbf{x}_n}], \tag{A.15}$$

and the functional derivative of G_Ξ in the direction $\delta_{\mathbf{x}}$, evaluated at $h = 1$, gives

$$v_\Xi(\mathbf{x}) = \delta G_\Xi[1\,;\delta_{\mathbf{x}}], \tag{A.16}$$

where $\delta_{\mathbf{x}}, \delta_{\mathbf{x}_1}, \ldots, \delta_{\mathbf{x}_n}$ denote Dirac delta functions. To illustrate the result for the first-order moment of a RFS given in equation (A.16), consider the functional derivative of the p.g.fl. of a Poisson RFS in the direction $\delta_{\mathbf{x}}$, i.e. from equation (A.11),

$$\delta G_\Xi[h\,;\delta_{\mathbf{x}}] = \exp\left(v[h-1]\right)\delta v[h-1\,;\delta_{\mathbf{x}}], \tag{A.17}$$

where $v[h-1] = \int \left(h(\mathbf{x})-1\right) v(\mathbf{x})\,\mathrm{d}\mathbf{x}$. Note that

$$\begin{aligned}
\delta v[h-1\,;\delta_{\mathbf{x}}] &= \lim_{\epsilon\to 0^+} \frac{1}{\epsilon}\left(v[h-1+\epsilon\delta_{\mathbf{x}}] - v[h-1]\right) \\
&= \lim_{\epsilon\to 0^+} \frac{1}{\epsilon}\left(v[h-1] + \epsilon v[\delta_{\mathbf{x}}] - v[h-1]\right) \\
&= v[\delta_{\mathbf{x}}] = v(\mathbf{x}),
\end{aligned}$$

which follows from Dirac delta properties in Section (A.1). Equation (A.17) can then be written as

$$\delta G_\Xi[h\,;\delta_{\mathbf{x}}] = \exp\left(v[h-1]\right) v(\mathbf{x}) = G_\Xi[h]\, v(\mathbf{x}), \tag{A.18}$$

and setting $h = 1$ gives $\delta G_\Xi[1\,;\delta_{\mathbf{x}}] = v(\mathbf{x})$ as expected.

Appendix B

PHD Derivations for Group and Extended Target Tracking

This chapter provides details on the derivations of the filtering equations for the PHD recursions presented in Section 3.2. The derivations are based on the multi-target PHD filter derivation [1] and requires the concept of probability generating functionals for cluster processes, i.e. p.g.fl. of the multiple group/extended target system RFS representations as shown in the next section. The first-order moment approximations of the Bayesian filter for group/extended targets can then be derived from the functional derivatives of such probability generating functionals, as demonstrated in Section B.2.

B.1 Probability Generating Functionals of Cluster Processes

The p.g.fl. of a cluster process which is characterised by the superposition of ordered pairs[1] was proposed by Moyal in [107]. Such a cluster process is equivalent to the multiple group/extended target system RFS representation given by $\mathcal{X} = \bigcup_{\mathbf{x}\in\Upsilon}\{(\mathbf{x}, \mathcal{E}^{\mathbf{x}})\}$, as defined in Section 3.1.2, where Υ is the RFS of group states and $\mathcal{E}^{\mathbf{x}}$ is the subsidiary RFS of individual target states for each $\mathbf{x} \in \Upsilon$. The p.g.fl. of a RFS $\mathcal{X}$ is defined as follows.

Let the finite sets X and Ξ be instances of the RFS Υ and a subsidiary RFS $\mathcal{E}$ respectively. Furthermore, let $h(\mathbf{x}, \Xi)$ be a test function defined on the space of all ordered pairs, for $\mathbf{x} \in X$, such that $0 \leq h(\mathbf{x}, \Xi) \leq 1$. Then the p.g.fl. of a RFS $\mathcal{X}$ is given by the following expectation value

$$G_{\mathcal{X}}[h] = G_{\Upsilon}\left[G_{\mathcal{E}}[h\,|\,\cdot]\right] = \mathsf{E}\left[\prod_{\mathbf{x}\in X} G_{\mathcal{E}}[h\,|\,\mathbf{x}]\right] = \int\left(\prod_{\mathbf{x}\in X} G_{\mathcal{E}}[h\,|\,\mathbf{x}]\right) p_{\Upsilon}(X)\,\delta X, \quad \text{(B.1)}$$

where $p_{\Upsilon}(X)$ denotes the probability density of the RFS Υ and

[1] An ordered pair is defined in [107] by (x, y^n), where x is the characteristic of the cluster as a whole and y^n means that the cluster contains n individuals with states $y_1, \ldots, y_n$.

$$G_{\mathcal{E}}[h \,|\, \mathbf{x}] = \int h(\mathbf{x}, \Xi)\, p_{\mathcal{E}}(\Xi \,|\, \mathbf{x})\, \delta\Xi, \tag{B.2}$$

where $p_{\mathcal{E}}(\Xi \,|\, \mathbf{x})$ is the conditional probability density of each subsidiary RFS $\mathcal{E}$.

Consider the alternative characterisation of a cluster process found in [70, Section 6.3] and [106, Section 5.1], which has an equivalent RFS representation given by $\mathcal{X} = \bigcup_{\mathbf{x}\in\Upsilon} \mathcal{E}^{\mathbf{x}}$. Let the finite sets X and Ξ be instances of the RFS Υ and a subsidiary RFS $\mathcal{E}$ respectively. Furthermore, let $h(\boldsymbol{\xi})$ be a test function defined on the individual target space, for $\boldsymbol{\xi} \in \Xi$, such that $0 \leq h(\boldsymbol{\xi}) \leq 1$. Then the p.g.fl. of a RFS $\mathcal{X}$ is then given by the same expectation value defined in equation (B.1), except now $G_{\mathcal{E}}[h \,|\, \mathbf{x}]$ is the p.g.fl. of the RFS $\mathcal{E}$ given by the following expectation value

$$G_{\mathcal{E}}[h \,|\, \mathbf{x}] = \mathsf{E}\left[\prod_{\boldsymbol{\xi}\in\Xi} h(\boldsymbol{\xi}) \,\middle|\, \mathbf{x}\right] = \int \left(\prod_{\boldsymbol{\xi}\in\Xi} h(\boldsymbol{\xi})\right) p_{\mathcal{E}}(\Xi \,|\, \mathbf{x})\, \delta\Xi. \tag{B.3}$$

Suppose the RFS Υ is Poisson with intensity v_{Υ}, then the p.g.fl. of $\mathcal{X}$ is given by

$$G_{\mathcal{X}}[h] = \exp\left(\int (G_{\mathcal{E}}[h \,|\, \mathbf{x}] - 1)\, v_{\Upsilon}(\mathbf{x})\, \mathrm{d}\mathbf{x}\right), \tag{B.4}$$

and the point process equivalence of such a RFS is referred to in point process literature as a *Poisson cluster process*. Furthermore, suppose also that each subsidiary RFS $\mathcal{E}$ is i.i.d., then the RFS $\mathcal{X}$ is equivalent to the *Neyman-Scott process* [108].

For the purpose of deriving the filtering equations for the PHD recursions presented in Section 3.2 for group and extended target tracking applications, it is necessary to introduce an additional test function in the p.g.fl. of RFS $\mathcal{X} = \bigcup_{\mathbf{x}\in\Upsilon} \mathcal{E}^{\mathbf{x}}$. Let $\bar{h}(\mathbf{x})$ be a test function defined on the group state space, for $\mathbf{x} \in X$, such that $0 \leq \bar{h}(\mathbf{x}) \leq 1$. Then the joint p.g.fl. of RFS $\mathcal{X}$ is given the following expectation value

$$\begin{aligned} G_{\mathcal{X}}[\bar{h}, h] = G_{\Upsilon}\left[\bar{h}\, G_{\mathcal{E}}[h \,|\, \cdot]\right] &= \mathsf{E}\left[\prod_{\mathbf{x}\in X} \bar{h}(\mathbf{x})\, G_{\mathcal{E}}[h \,|\, \mathbf{x}]\right] \\ &= \int \left(\prod_{\mathbf{x}\in X} \bar{h}(\mathbf{x})\, G_{\mathcal{E}}[h \,|\, \cdot]\right) p_{\Upsilon}(\mathbf{x})\, \delta X, \end{aligned} \tag{B.5}$$

where $G_{\mathcal{E}}[h \,|\, \mathbf{x}]$ is still the p.g.fl. of RFS $\mathcal{E}$ given by equation (B.3). In addition, the RFS Υ and subsidiary RFSs $\mathcal{E}$ will each be approximated with a Poisson RFS

for the derivations in Sections 3.2-B.6. The joint p.g.fl. of the corresponding RFS $\mathcal{X}$ is then given by

$$\begin{aligned} G_{\mathcal{X}}[\bar{h}, h] &= \exp\left(v_{\Upsilon}\left[\bar{h}\exp\left(v_{\mathcal{E}}[h-1\,|\,\cdot]\right)-1\right]\right) \\ &= \exp\left(\int\left(\bar{h}(\mathbf{x})\exp\left(\int(h(\boldsymbol{\xi})-1)\,v_{\mathcal{E}}(\boldsymbol{\xi}\,|\,\mathbf{x})\,\mathrm{d}\boldsymbol{\xi}\right)-1\right)v_{\Upsilon}(\mathbf{x})\,\mathrm{d}\mathbf{x}\right). \end{aligned} \tag{B.6}$$

B.2 First-Order Moments of Cluster Processes

The specific formula for the first-order moment of the RFS $\mathcal{X} = \bigcup_{\mathbf{x}\in\Upsilon}\{(\mathbf{x}, \mathcal{E}^{\mathbf{x}})\}$, referred to as the total group target PHD and given by equation (3.7) in Section 3.1.4, can be established in a similar way as demonstrated in Section A.3. That is,

$$\begin{aligned} &v_{\mathcal{X}}(\mathbf{x}, \Xi) \\ &= \int\left(\sum\nolimits_{(\mathbf{x}',\Xi')\in\mathbb{X}}\delta_{(\mathbf{x},\Xi)}(\mathbf{x}',\Xi')\right)p_{\mathcal{X}}(\mathbb{X})\,\delta\mathbb{X} \\ &= \sum_{n=0}^{\infty}\frac{1}{n!}\int\left(\delta_{(\mathbf{x},\Xi)}(\mathbf{x}_1,\Xi_1)+\cdots+\delta_{(\mathbf{x},\Xi)}(\mathbf{x}_n,\Xi_n)\right)\times \\ &\qquad\qquad p_{\mathcal{X}}(\{(\mathbf{x}_1,\Xi_1),\ldots,(\mathbf{x}_n,\Xi_n)\})\,\mathrm{d}\mathbf{x}_1\cdots\mathrm{d}\mathbf{x}_n\,\delta\Xi_1\cdots\delta\Xi_n \\ &= \sum_{n=1}^{\infty}\frac{n}{n!}\int p_{\mathcal{X}}(\{(\mathbf{x},\Xi),(\mathbf{x}'_1,\Xi'_1),\ldots,(\mathbf{x}'_{n-1},\Xi'_{n-1})\})\,\mathrm{d}\mathbf{x}'_1\cdots\mathrm{d}\mathbf{x}'_{n-1}\,\delta\Xi'_1\cdots\delta\Xi'_{n-1} \\ &= \sum_{n=1}^{\infty}\frac{1}{(n-1)!}\int p_{\mathcal{X}}(\{(\mathbf{x},\Xi),(\mathbf{x}'_1,\Xi'_1),\ldots,(\mathbf{x}'_{n-1},\Xi'_{n-1})\})\,\mathrm{d}\mathbf{x}'_1\cdots\mathrm{d}\mathbf{x}'_{n-1}\,\delta\Xi'_1\cdots\delta\Xi'_{n-1} \\ &= \int p_{\mathcal{X}}(\{(\mathbf{x},\Xi)\}\cup\mathbb{X}')\,\delta\mathbb{X}' = \int_{\mathbb{X}\ni(\mathbf{x},\Xi)} p_{\mathcal{X}}(\mathbb{X})\,\delta\mathbb{X}. \end{aligned}$$

Similarly, the first-order moment of the RFS $\mathcal{X} = \bigcup_{\mathbf{x}\in\Upsilon}\mathcal{E}^{\mathbf{x}}$, referred to as the condensed group target PHD and given by equation (3.8), can be established as follows:

$$\begin{aligned} &v_{\mathcal{X}}(\mathbf{x}, \boldsymbol{\xi}) \\ &= \int\left(\sum\nolimits_{\Xi^{\mathbf{w}}\in\mathbb{X},\,\boldsymbol{\xi}'\in\Xi^{\mathbf{w}}}\delta_{\mathbf{x}}(\mathbf{w})\,\delta_{\boldsymbol{\xi}}(\boldsymbol{\xi}')\right)p_{\mathcal{X}}(\mathbb{X})\,\delta\mathbb{X} \\ &= \sum_{n=0}^{\infty}\frac{1}{n!}\int\left(\delta_{\mathbf{x}}(\mathbf{x}_1)\sum\nolimits_{\boldsymbol{\xi}'_1\in\Xi_1^{\mathbf{x}}}\delta_{\boldsymbol{\xi}}(\boldsymbol{\xi}'_1)+\cdots+\delta_{\mathbf{x}}(\mathbf{x}_n)\sum\nolimits_{\boldsymbol{\xi}'_n\in\Xi_n^{\mathbf{x}}}\delta_{\boldsymbol{\xi}}(\boldsymbol{\xi}'_n)\right) \\ &\qquad\qquad\times p_{\mathcal{X}}(\{\Xi_1^{\mathbf{x}},\ldots,\Xi_n^{\mathbf{x}}\})\,\mathrm{d}\mathbf{x}_1\cdots\mathrm{d}\mathbf{x}_n\,\delta\Xi_1^{\mathbf{x}}\cdots\delta\Xi_n^{\mathbf{x}} \end{aligned}$$

$$
\begin{aligned}
&= \sum_{n=1}^{\infty} \frac{n}{n!} \int \Big(\sum\nolimits_{\boldsymbol{\xi}' \in \Xi^{\mathbf{x}}} \delta_{\boldsymbol{\xi}}(\boldsymbol{\xi}') \Big) \\
&\qquad\qquad \times p_{\mathcal{X}}(\{\Xi^{\mathbf{x}}, \Xi_1^{\mathbf{w}}, \ldots, \Xi_{n-1}^{\mathbf{w}}\}) \, \mathrm{d}\mathbf{w}_1, \cdots \mathrm{d}\mathbf{w}_{n-1} \, \delta\Xi^{\mathbf{x}} \, \delta\Xi_1^{\mathbf{w}} \cdots \delta\Xi_{n-1}^{\mathbf{w}} \\
&= \int \Big(\sum\nolimits_{\boldsymbol{\xi}' \in \Xi^{\mathbf{x}}} \delta_{\boldsymbol{\xi}}(\boldsymbol{\xi}') \Big) \, p_{\mathcal{X}}(\{\Xi^{\mathbf{x}}\} \cup \mathbb{W}) \, \delta\Xi^{\mathbf{x}} \, \delta\mathbb{W} \\
&= \sum_{m=0}^{\infty} \frac{1}{m!} \int \left(\delta_{\boldsymbol{\xi}}(\boldsymbol{\xi}_1) + \cdots + \delta_{\boldsymbol{\xi}}(\boldsymbol{\xi}_m) \right) p_{\mathcal{X}}(\{\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_m\}^{\mathbf{x}} \cup \mathbb{W}) \, \mathrm{d}\boldsymbol{\xi}_1 \cdots \mathrm{d}\boldsymbol{\xi}_m \, \delta\mathbb{W} \\
&= \sum_{m=1}^{\infty} \frac{m}{m!} \int p_{\mathcal{X}}(\{\boldsymbol{\xi}, \boldsymbol{\xi}_1', \ldots, \boldsymbol{\xi}_{m-1}'\}^{\mathbf{x}} \cup \mathbb{W}) \, \mathrm{d}\boldsymbol{\xi}_1' \cdots \mathrm{d}\boldsymbol{\xi}_{m-1}' \, \delta\mathbb{W} \\
&= \int p_{\mathcal{X}} \left(\{\{\boldsymbol{\xi}\} \cup \Xi'\}^{\mathbf{x}} \cup \mathbb{W} \right) \delta\Xi' \, \delta\mathbb{W} = \int_{\mathbb{X} \ni \Xi^{\mathbf{x}}, \, \Xi^{\mathbf{x}} \ni \boldsymbol{\xi}} p_{\mathcal{X}}(\mathbb{X}) \, \delta\mathbb{X}.
\end{aligned}
$$

As discussed in Section 3.1.4, recursions that propagate the first-order moment of the RFS $\mathcal{X} = \bigcup_{\mathbf{x} \in \Upsilon} \mathcal{E}^{\mathbf{x}}$ will provide tractable solutions to the group target tracking problem. The derivations of the filtering equations for such recursions are detailed in Sections B.3–B.6 based on the property relating the first-order moment of RFS $\mathcal{X}$ with the functional derivative of the joint p.g.fl. $G_{\mathcal{X}}$ as given in Section B.2.2 (and Section B.2.3). Before that, an important result is introduced regarding the functional derivative of a composite functional.

B.2.1 The chain rule for functional derivatives

Since the joint p.g.fl. $G_{\mathcal{X}}[\bar{h}, h]$ is given by the composition $G_{\Upsilon}\left[\bar{h}, G_{\mathcal{E}}[h \,|\, \cdot]\right]$, its functional derivative in the direction η requires the following result.

Lemma B.1 (**Chain rule**). *Let G_{Υ} and $G_{\mathcal{E}}$ be differentiable functionals, then the functional derivative of the composition $G_{\mathcal{X}}[h] = G_{\Upsilon}\left[G_{\mathcal{E}}[h \,|\, \cdot]\right]$ in the direction η is*

$$
\delta G_{\mathcal{X}}[h \,; \eta] = \delta G_{\Upsilon}\left[G_{\mathcal{E}}[h \,|\, \cdot] \,; \eta\right] = \delta G_{\Upsilon}\left[G_{\mathcal{E}}[h \,|\, \cdot] \,; \delta G_{\mathcal{E}}[h \,|\, \cdot \,; \eta]\right]. \tag{B.7}
$$

Proof. Details on how this result is established can be found in [135], where the direction is evaluated at $\eta = \delta_{\mathbf{y}}$. □

Given the Dirac delta function $\delta_{\mathbf{x},\boldsymbol{\xi}}$ concentrated at $\mathbf{x}$ and $\boldsymbol{\xi}$ such that $\delta_{\mathbf{x},\boldsymbol{\xi}} = \delta_{\mathbf{x}}(\mathbf{x}')\delta_{\boldsymbol{\xi}}(\boldsymbol{\xi}')$, it follows from the chain rule that the functional derivative of the joint p.g.fl. $G_{\mathcal{X}}[\bar{h}, h]$ in the direction $\delta_{\mathbf{x},\boldsymbol{\xi}}$ is

$$
\delta G_{\mathcal{X}}[\bar{h}, h \,; \delta_{\mathbf{x},\boldsymbol{\xi}}] = \delta G_{\Upsilon}\left[\bar{h} \, G_{\mathcal{E}}[h \,|\, \cdot] \,; \delta G_{\mathcal{E}}[h \,|\, \cdot \,; \delta_{\boldsymbol{\xi}}] \, \delta_{\mathbf{x}}\right]. \tag{B.8}
$$

B.2.2 The condensed group target PHD

The functional derivative of the joint p.g.fl. $G_{\mathcal{X}}[\bar{h}, h]$ in the direction $\delta_{\mathbf{x},\boldsymbol{\xi}}$, evaluated at $\bar{h} = 1$ and $h = 1$, gives the first-order moment of the RFS $\mathcal{X}$, i.e.

$$v_{\mathcal{X}}(\mathbf{x}, \boldsymbol{\xi}) = \delta G_{\mathcal{X}}[1, 1\,; \delta_{\mathbf{x},\boldsymbol{\xi}}], \tag{B.9}$$

To illustrate this result, consider the functional derivative, in the direction $\delta_{\mathbf{x},\boldsymbol{\xi}}$, of joint the p.g.fl. $G_{\mathcal{X}}$ given in equation (B.6). To start with, since the RFS $\mathcal{E}$ is Poisson, the functional derivative of its p.g.fl. $G_{\mathcal{E}}$ in the direction $\delta_{\boldsymbol{\xi}}$ is given by

$$\delta G_{\mathcal{E}}[h \,|\, \cdot\,; \delta_{\boldsymbol{\xi}}] = \exp\left(v_{\mathcal{E}}[h - 1 \,|\, \cdot]\right) v_{\mathcal{E}}(\boldsymbol{\xi} \,|\, \cdot).$$

It then follows from the chain rule (Lemma B.1) and equation (B.5), that the functional derivative of the joint p.g.fl. in the direction $\delta_{\mathbf{x},\boldsymbol{\xi}}$ is

$$\begin{aligned}\delta G_{\mathcal{X}}[\bar{h}, h\,; \delta_{\mathbf{x},\boldsymbol{\xi}}] &= \delta G_{\Upsilon}\left[\bar{h}\, G_{\mathcal{E}}[h \,|\, \cdot]\,; \exp(v_{\mathcal{E}}[h - 1 \,|\, \cdot])\, v_{\mathcal{E}}(\boldsymbol{\xi} \,|\, \cdot)\, \delta_{\mathbf{x}}\right] \\ &= \exp\left(v_{\Upsilon}\left[\bar{h}\, G_{\mathcal{E}}[h \,|\, \cdot] - 1\right]\right) \delta v_{\Upsilon}\left[\bar{h}\, G_{\mathcal{E}}[h \,|\, \cdot] - 1\,; G_{\mathcal{E}}[h \,|\, \cdot]\, v_{\mathcal{E}}(\boldsymbol{\xi} \,|\, \cdot)\, \delta_{\mathbf{x}}\right],\end{aligned} \tag{B.10}$$

since the RFS Υ is also Poisson. Finally, note that,

$$\begin{aligned}&\delta\, v_{\Upsilon}\left[\bar{h}\, G_{\mathcal{E}}[h \,|\, \cdot] - 1\,; G_{\mathcal{E}}[h \,|\, \cdot]\, v_{\mathcal{E}}(\boldsymbol{\xi} \,|\, \cdot)\, \delta_{\mathbf{x}}\right] \\ &= \lim_{\epsilon \to 0^+} \frac{1}{\epsilon}\Big(v_{\Upsilon}\left[\bar{h}\, G_{\mathcal{E}}[h \,|\, \cdot] - 1 + \epsilon\, G_{\mathcal{E}}[h \,|\, \cdot] v_{\mathcal{E}}(\boldsymbol{\xi} \,|\, \cdot)\, \delta_{\mathbf{x}}\right] - v_{\Upsilon}\left[\bar{h}\, G_{\mathcal{E}}[h \,|\, \cdot] - 1\right]\Big) \\ &= \lim_{\epsilon \to 0^+} \frac{1}{\epsilon}\Big(v_{\Upsilon}\left[\bar{h}\, G_{\mathcal{E}}[h \,|\, \cdot] - 1\right] + \epsilon\, v_{\Upsilon}[G_{\mathcal{E}}[h \,|\, \cdot]\, v_{\mathcal{E}}(\boldsymbol{\xi} \,|\, \cdot)\, \delta_{\mathbf{x}}] - v_{\Upsilon}\left[\bar{h}\, G_{\mathcal{E}}[h \,|\, \cdot] - 1\right]\Big) \\ &= v_{\Upsilon}[G_{\mathcal{E}}[h \,|\, \cdot]\, v_{\mathcal{E}}(\boldsymbol{\xi} \,|\, \cdot)\, \delta_{\mathbf{x}}] = \exp\left(v_{\mathcal{E}}[h - 1 \,|\, \mathbf{x}]\right) v_{\mathcal{E}}(\boldsymbol{\xi} \,|\, \mathbf{x})\, v_{\Upsilon}(\mathbf{x}),\end{aligned}$$

so that equation (B.10) becomes

$$\delta G_{\mathcal{X}}[\bar{h}, h\,; \delta_{\mathbf{x},\boldsymbol{\xi}}] = \exp\left(v_{\Upsilon}\left[\bar{h}\, G_{\mathcal{E}}[h \,|\, \cdot] - 1\right]\right) \exp\left(v_{\mathcal{E}}[h - 1 \,|\, \mathbf{x}]\right) v_{\Upsilon}(\mathbf{x})\, v_{\mathcal{E}}(\boldsymbol{\xi} \,|\, \mathbf{x}).$$

Finally, evaluating at $\bar{h} = 1$ and $h = 1$ gives $\delta G_{\mathcal{X}}[1, 1\,; \delta_{\mathbf{x},\boldsymbol{\xi}}] = v_{\Upsilon}(\mathbf{x})\, v_{\mathcal{E}}(\boldsymbol{\xi} \,|\, \mathbf{x})$. The first-order moment of the RFS $\mathcal{X}$ is a joint density satisfying the factorisation $v_{\mathcal{X}}(\mathbf{x}, \boldsymbol{\xi}) = v_{\Upsilon}(\mathbf{x}) \times v_{\mathcal{E}}(\boldsymbol{\xi} \,|\, \mathbf{x})$ given the dependence of $\boldsymbol{\xi} \in \mathcal{E}$ on $\mathbf{x} \in \Upsilon$. Hence the relation given by equation (B.9) is satisfied for a RFS $\mathcal{X}$ whose joint p.g.fl. is given by equation (B.6).

Note that the probability density of RFS $\mathcal{X}$ can also be recovered from the functional derivative of the joint p.g.fl. That is, given $\mathbb{X} = \{\Xi_1^{\mathbf{x}}, \ldots, \Xi_n^{\mathbf{x}}\}$ where

$\Xi_i^{\mathbf{x}} = \{\boldsymbol{\xi}_{i,1}, \ldots, \boldsymbol{\xi}_{i,n_i}\}$ for $i = 1, \ldots, |X|$ and $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$, the N^{th}-order functional derivative of $G_{\mathcal{X}}[\bar{h}, h]$, where

$$N = \prod_{i=1}^{|X|} n_i,$$

in the directions $\delta_{\Xi_1^{\mathbf{x}}}, \ldots, \delta_{\Xi_n^{\mathbf{x}}}$ (denoting $\delta_{\Xi_i^{\mathbf{x}}} = \delta_{\mathbf{x}_i, \boldsymbol{\xi}_{i,1}}, \ldots, \delta_{\mathbf{x}_i, \boldsymbol{\xi}_{i,n_i}}$ for $i = 1, \ldots, |X|$), evaluated at $\bar{h} = 0$ and $h = 0$, gives

$$p_{\mathcal{X}}(\mathbb{X}) = \delta^N G_{\mathcal{X}}\left[0, 0\,; \delta_{\Xi_1^{\mathbf{x}}}, \ldots, \delta_{\Xi_n^{\mathbf{x}}}\right]. \tag{B.11}$$

The probability density presented in Section 3.2.3 for a RFS whose parent and daughter processes are Poisson RFSs can be derived from equation (B.11) given its p.g.fl. has the same form as that in equation (B.6).

B.2.3 The marginal group target PHD

The functional derivative the joint p.g.fl. $G_{\mathcal{X}}[\bar{h}, h]$ in the direction of $\delta_{\mathbf{x}}$, evaluated at $\bar{h} = 1$ and $h = 1$, gives the first-order moment density of the RFS Υ, i.e.

$$v_{\Upsilon}(\mathbf{x}) = \delta G_{\mathcal{X}}[1, 1\,; \delta_{\mathbf{x}}], \tag{B.12}$$

where $\delta_{\mathbf{x}}$ is the Dirac delta function concentrated at $\mathbf{x}$. To illustrate this result, given by equation (B.12), again consider the functional derivative, this time in the direction $\delta_{\mathbf{x}}$, of the joint p.g.fl. given in equation (B.6). That is

$$\begin{aligned}\delta G_{\mathcal{X}}[\bar{h}, h\,; \delta_{\mathbf{x}}] &= \delta G_{\Upsilon}\left[\bar{h}\, G_{\mathcal{E}}[h\,|\,\cdot]\,; \delta_{\mathbf{x}}\right] \\ &= \exp\left(v_{\Upsilon}\left[\bar{h}\, G_{\mathcal{E}}[h\,|\,\cdot] - 1\right]\right) \delta v_{\Upsilon}\left[\bar{h}\, G_{\mathcal{E}}[h\,|\,\cdot] - 1\,; \delta_{\mathbf{x}}\right], \end{aligned} \tag{B.13}$$

since the direction is strictly parameterised by $\mathbf{x} \in \Upsilon$. Note that

$$\begin{aligned}\delta v_{\Upsilon}\left[\bar{h}\, G_{\mathcal{E}}[h\,|\,\cdot] - 1\,; \delta_{\mathbf{x}}\right] &= \lim_{\epsilon \to 0^+} \frac{1}{\epsilon}\left(v_{\Upsilon}\left[\bar{h}\, G_{\mathcal{E}}[h\,|\,\cdot] - 1 + \epsilon\, \delta_{\mathbf{x}}\right] - v_{\Upsilon}\left[\bar{h}\, G_{\mathcal{E}}[h\,|\,\cdot] - 1\right]\right) \\ &= v_{\Upsilon}[\delta_{\mathbf{x}}] = v_{\Upsilon}(\mathbf{x}).\end{aligned}$$

Equation (B.13) becomes $\delta G_{\mathcal{X}}[\bar{h}, h\,; \delta_{\mathbf{x}}] = \exp\left(v_{\Upsilon}\left[\bar{h}\, G_{\mathcal{E}}[h\,|\,\cdot] - 1\right]\right) v_{\Upsilon}(\mathbf{x})$ then setting $\bar{h} = 1$ and $h = 1$ gives $\delta G_{\mathcal{X}}[1, 1\,; \delta_{\mathbf{x}}] = v_{\Upsilon}(\mathbf{x})$ as expected.

B.3 Derivation of the Multi-Group PHD Filter Recursion

This section summarises the derivation of the multi-group PHD filter recursion presented in Section 3.2.3. Sketch proofs are given in Sections B.3.1 and B.3.2 respectively for the results of the prediction and update equations.

B.3.1 Multi-group PHD prediction

The joint p.g.fl. of the RFS representation of the multiple group target system, i.e. $\mathcal{X}_{k-1} = \bigcup_{\mathbf{x}\in\Upsilon_{k-1}} \mathcal{E}_{k-1}^{\mathbf{x}}$, is given by $G_{k-1}[\bar{h}, h] = G_{\Upsilon,k-1}\big[\bar{h}\, G_{\mathcal{E},k-1}[h\,|\,\cdot]\big]$. Then, given the dynamic model specified in Section 3.2.1, the p.g.fl. corresponding to the predicted posterior in the multi-group Bayesian filter recursion can be written as

$$G_{k|k-1}[\bar{h}, h] = G_{\Gamma,k}[\bar{h}, h] \int \left(\prod_{\mathbf{x}'\in X_{k-1}} G_{\Phi,k|k-1}\big[\bar{h}\, G_{\mathcal{E},k|k-1}[h\,|\,\cdot,:]\big|\,\mathbf{x}'\big] \right) p_{k-1}(X_{k-1})\,\delta X_{k-1},$$

where

$$G_{\mathcal{E},k|k-1}[h\,|\,\cdot,:] = G_{\Gamma,\xi,k}[h\,|\,\cdot\,] \int \left(\prod_{\boldsymbol{\xi}'\in\Xi_{k-1}} G_{\Psi,k|k-1}[h\,|\,\boldsymbol{\xi}',:] \right) p_{k-1}(\Xi_{k-1})\,\delta\Xi_{k-1}.$$

That is, the predicted p.g.fl. is given by

$$G_{k|k-1}[\bar{h}, h] = G_{\Gamma,k}[\bar{h}, h]\, G_{\Upsilon,k-1}\big[\bar{h}\, G_{\Phi,k|k-1}\big[G_{\mathcal{E},k|k-1}[h\,|\,\cdot,:]\big|:\big]\big]\,, \tag{B.14}$$

where $G_{\mathcal{E},k|k-1}[h\,|\,\cdot,:] = G_{\Gamma,\xi,k}[h\,|\,\cdot\,]\, G_{\mathcal{E},k-1}\big[G_{\Psi,k|k-1}[h\,|\,\cdot\cdot,:]\big|:\big]$.

The result for the predicted intensity follows from the relation given in equation (B.9). That is, taking the functional derivative of $G_{k|k-1}$ in the direction $\delta_{\mathbf{x},\boldsymbol{\xi}}$ then evaluating at $\bar{h} = 1$ and $h = 1$, as follows. Given the expression for $G_{k|k-1}$ in equation (B.14), $\delta G_{k|k-1}[\bar{h}, h\,;\delta_{\mathbf{x},\boldsymbol{\xi}}]$ is written as

$$\begin{aligned}\delta G_{k|k-1}[\bar{h}, h\,;\delta_{\mathbf{x},\boldsymbol{\xi}}] = {}& \delta G_{\Gamma,k}[\bar{h}, h\,;\delta_{\mathbf{x},\boldsymbol{\xi}}]\, G_{\Upsilon,k-1}\big[G_{\Phi,k|k-1}\big[\bar{h}\, G_{\mathcal{E},k|k-1}[h\,|\,\cdot,:]\big|:\big]\big] \\ & + G_{\Gamma,k}[\bar{h}, h]\, \delta G_{\Upsilon,k-1}\big[G_{\Phi,k|k-1}\big[\bar{h}\, G_{\mathcal{E},k|k-1}[h\,|\,\cdot,:]\big|:\big]\,;\delta_{\mathbf{x},\boldsymbol{\xi}}\big]\,.\end{aligned} \tag{B.15}$$

By applying the chain rule (Lemma B.1) it follows that

$$\begin{aligned}& \delta G_{\Upsilon,k-1}\big[G_{\Phi,k|k-1}\big[\bar{h}\, G_{\mathcal{E},k|k-1}[h\,|\,\cdot,:]\big|:\big]\,;\delta_{\mathbf{x},\boldsymbol{\xi}}\big] \\ &= \delta G_{\Upsilon,k-1}\big[G_{\Phi,k|k-1}\big[\bar{h}\, G_{\mathcal{E},k|k-1}[h\,|\,\cdot,:]\big|:\big]\,;\delta G_{\Phi,k|k-1}\big[\bar{h}\, G_{\mathcal{E},k|k-1}[h\,|\,\cdot,:]\,;\delta_{\mathbf{x},\boldsymbol{\xi}}\big|:\big]\big]\end{aligned}$$

$$= \int \left(\sum_{\mathbf{x}' \in X_{k-1}} \delta G_{\Phi,k|k-1}\left[\bar{h}\, G_{\mathcal{E},k|k-1}[h \,|\, \cdot, :]\,; \delta_{\mathbf{x},\boldsymbol{\xi}}\right| \mathbf{x}'\right] \times \left(\prod_{i=1}^{|X_{k-1}|-1} G_{\Phi,k|k-1}\left[\bar{h}\, G_{\mathcal{E},k|k-1}[h \,|\, \cdot, :]\right| \mathbf{x}'_i\right] \right) \right) p_{k-1}(X_{k-1})\, \delta X_{k-1}. \tag{B.16}$$

The assumption on the RFS $\Phi_{k|k-1}$ described in Section 3.2.3 implies that the joint p.g.fl. $G_{\Phi,k|k-1}\left[\bar{h}\, G_{\mathcal{E},k|k-1}[h \,|\, \cdot, :]\right| :\right]$ has the following form

$$\begin{aligned} &G_{\Phi,k|k-1}\left[\bar{h}\, G_{\mathcal{E},k|k-1}[h \,|\, \cdot, :]\right| :\right] \\ &= 1 - p_{S,1}(:) + p_{S,1}(:) \int \bar{h}(\mathbf{x}_k)\, G_{\mathcal{E},k|k-1}[h \,|\, \mathbf{x}_k, :]\, f_{k|k-1}(\mathbf{x}_k \,|\, :)\, \mathrm{d}\mathbf{x}_k, \end{aligned} \tag{B.17}$$

so that

$$\begin{aligned} &\delta G_{\Phi,k|k-1}\left[\bar{h}\, G_{\mathcal{E},k|k-1}[h \,|\, \cdot, :]\,; \delta_{\mathbf{x},\boldsymbol{\xi}}\right| :\right] \\ &= \delta G_{\Phi,k|k-1}\left[\bar{h}\, G_{\mathcal{E},k|k-1}[h \,|\, \cdot, :]\,; \delta G_{\mathcal{E},k|k-1}[h\,; \delta_{\boldsymbol{\xi}} \,|\, \cdot, :]\, \delta_{\mathbf{x}}\right| :\right] \\ &= p_{S,1}(:)\, f_{k|k-1}(\mathbf{x} \,|\, :)\, \delta G_{\mathcal{E},k|k-1}[h\,; \delta_{\boldsymbol{\xi}} \,|\, \mathbf{x}, :] \end{aligned} \tag{B.18}$$

Furthermore, given the expression for $G_{\mathcal{E},k|k-1}$, it follows that

$$\begin{aligned} \delta G_{\mathcal{E},k|k-1}[h\,; \delta_{\boldsymbol{\xi}} \,|\, \mathbf{x}, :] = {} & \delta G_{\Gamma,\xi,k}[h\,; \delta_{\boldsymbol{\xi}} \,|\, \mathbf{x}]\, G_{\mathcal{E},k-1}\left[G_{\Psi,k|k-1}[h \,|\, \cdot\cdot, :]\right| :\right] \\ & + G_{\Gamma,\xi,k}[h \,|\, \mathbf{x}]\, \delta G_{\mathcal{E},k-1}\left[G_{\Psi,k|k-1}[h \,|\, \cdot\cdot, :]\,; \delta_{\boldsymbol{\xi}}\right| :\right], \end{aligned}$$

where, applying the chain rule once again,

$$\begin{aligned} &\delta G_{\mathcal{E},k-1}\left[G_{\Psi,k|k-1}[h \,|\, \cdot\cdot, :]\,; \delta_{\boldsymbol{\xi}}\right| :\right] \\ &= \delta G_{\mathcal{E},k-1}\left[G_{\Psi,k|k-1}[h \,|\, \cdot\cdot, :]\,; \delta G_{\Psi,k|k-1}[h\,; \delta_{\boldsymbol{\xi}} \,|\, \cdot\cdot, :]\right| :\right], \end{aligned}$$

which can be written in a similar form as the expression given in equation (B.16). The assumption on the RFS $\Psi_{k|k-1}$ described in Section 3.2.3 implies that its p.g.fl. has a similar form to that given in equation (B.17) so that

$$\delta G_{\Psi,k|k-1}[h\,; \delta_{\boldsymbol{\xi}} \,|\, \cdot\cdot, :] = p_{S,2}(\cdot\cdot)\, f_{k|k-1}(\boldsymbol{\xi} \,|\, \cdot\cdot, :).$$

Hence, by substitution, the functional derivative of $G_{\mathcal{E},k-1}\left[G_{\Psi,k|k-1}[h \,|\, \cdot\cdot, :]\right| :\right]$ in the direction $\delta_{\boldsymbol{\xi}}$, evaluated at $h = 1$, becomes

$$\delta G_{\mathcal{E},k-1}\big[G_{\Psi,k|k-1}[1\,|\,\cdot\cdot,:]\,;p_{S,2}(\cdot\cdot)\,f_{k|k-1}(\boldsymbol{\xi}\,|\,\cdot\cdot)\big]$$
$$= \int p_{S,2}(\boldsymbol{\xi}_{k-1})\,f_{k|k-1}(\boldsymbol{\xi}\,|\,\boldsymbol{\xi}_{k-1},:)\,v_{k-1}(\boldsymbol{\xi}_{k-1}\,|\,:)\,\mathrm{d}\boldsymbol{\xi}_{k-1},$$

which follows from Campbell's theorem (Section A.2). From the relation given in equation (B.9) it also follows that $\delta G_{\Gamma,\xi,k}[1\,;\delta_{\boldsymbol{\xi}}\,|\,\mathbf{x}] = \gamma_{\xi,k}(\boldsymbol{\xi}\,|\,\mathbf{x})$ and so

$$\delta G_{\mathcal{E},k|k-1}[1\,;\delta_{\boldsymbol{\xi}}\,|\,\mathbf{x},:] = v_{k|k-1}(\boldsymbol{\xi}\,|\,\mathbf{x},:), \tag{B.19}$$

where $v_{k|k-1}(\boldsymbol{\xi}\,|\,\mathbf{x},:)$ is the conditional predicted intensity given in equation (3.18). Therefore, it follows from equation (B.19), that setting $\bar{h} = 1$ and $h = 1$ in equation (B.18) gives

$$\delta G_{\Phi,k|k-1}\big[G_{\mathcal{E},k|k-1}[1\,|\,\cdot,:]\,;\delta_{\mathbf{x},\boldsymbol{\xi}}\big|:\big] = p_{S,1}(:)\,f_{k|k-1}(\mathbf{x}\,|\,:)\,v_{k|k-1}(\boldsymbol{\xi}\,|\,\mathbf{x},:).$$

Subsequently setting $\bar{h} = 1$ and $h = 1$ in equation (B.16) gives

$$\delta G_{\Upsilon,k-1}\big[1\,;\delta G_{\Phi,k|k-1}\big[G_{\mathcal{E},k|k-1}[1\,|\,\cdot,:]\,;\delta_{\mathbf{x},\boldsymbol{\xi}}\big|:\big]\big]$$
$$= \int p_{S,1}(\mathbf{x}_{k-1})\,f_{k|k-1}(\mathbf{x}\,|\,\mathbf{x}_{k-1})\,v_{k|k-1}(\boldsymbol{\xi}\,|\,\mathbf{x},\mathbf{x}_{k-1})\,v_{k-1}(\mathbf{x}_{k-1})\,\mathrm{d}\mathbf{x}_{k-1},$$

which again follows from Campbell's theorem (Section A.2). The above expression gives the second term, setting $\bar{h} = 1$ and $h = 1$, in equation (B.15), and the predicted intensity given in equation (3.17) follows by further noting that $\delta G_{\Gamma,k}[1,1\,;\delta_{\mathbf{x},\boldsymbol{\xi}}] = \gamma_k(\mathbf{x},\boldsymbol{\xi})$.

B.3.2 Multi-group PHD update

Under the RFS representation of the multiple group target system given by $\mathcal{X}_k = \bigcup_{\mathbf{x}\in\Upsilon_k}\mathcal{E}_k^{\mathbf{x}}$ with instance $\mathbb{X}_k$, the corresponding joint p.g.fl. for the updated posterior in the multi-group Bayesian filter recursion is given by

$$G_k[\bar{h},h] = \frac{\int\left(\prod_{\mathbf{x}\in X_k}\bar{h}(\mathbf{x})\left(\prod_{\boldsymbol{\xi}\in\Xi_k^{\mathbf{x}}}h(\boldsymbol{\xi})\right)\right)g_k(Z_k\,|\,\mathbb{X}_k)\,p_{k|k-1}(\mathbb{X}_k)\,\delta\mathbb{X}_k}{\int g_k(Z_k\,|\,\mathbb{X}_k)\,p_{k|k-1}(\mathbb{X}_k)\,\delta\mathbb{X}_k}.$$

Note that the likelihood $g_k(Z_k\,|\,\mathbb{X}_k)$ can be recovered from its p.g.fl. That is, given a measurement set $Z_k = \{\mathbf{z}_1,\ldots,\mathbf{z}_m\}$, $g_k(Z_k\,|\,\mathbb{X}_k) = \delta^m G_{Z,k}[0\,;\delta_{\mathbf{z}_1},\ldots,\delta_{\mathbf{z}_m}\,|\,\mathbb{X}_k]$.

Therefore, the following trivariate p.g.fl. is introduced

$$F[g,\bar{h},h]=\int\left(\prod\nolimits_{\mathbf{x}\in X_k}\bar{h}(\mathbf{x})\left(\prod\nolimits_{\boldsymbol{\xi}\in\Xi_k^{\mathbf{x}}}h(\boldsymbol{\xi})\right)\right)G_{Z,k}[g\,|\,\mathbb{X}_k]\,p_{k|k-1}(\mathbb{X}_k)\,\delta\mathbb{X}_k, \quad \text{(B.20)}$$

the m^{th}-order functional derivative of which in the directions $\delta_{\mathbf{z}_1},\ldots,\delta_{\mathbf{z}_m}$ is given by

$$\begin{aligned}&\delta^m F[g,\bar{h},h\,|\,\delta_{\mathbf{z}_1},\ldots,\delta_{\mathbf{z}_m}]\\&=\int\left(\prod\nolimits_{\mathbf{x}\in X_k}\bar{h}(\mathbf{x})\left(\prod\nolimits_{\boldsymbol{\xi}\in\Xi_k^{\mathbf{x}}}h(\boldsymbol{\xi})\right)\right)\delta^m G_{Z,k}[g\,;\delta_{\mathbf{z}_1},\ldots,\delta_{\mathbf{z}_m}\,|\,\mathbb{X}_k]\,p_{k|k-1}(\mathbb{X}_k)\,\delta\mathbb{X}_k,\end{aligned}$$

so that the updated joint p.g.fl. can be written as

$$G_k[\bar{h},h]=\frac{\delta^m F[0,\bar{h},h\,;\delta_{\mathbf{z}_1},\ldots,\delta_{\mathbf{z}_m}]}{\delta^m F[0,1,1\,;\delta_{\mathbf{z}_1},\ldots,\delta_{\mathbf{z}_m}]} \quad \text{(B.21)}$$

The following Lemma shows the specific formulation of the updated joint p.g.fl. given the measurement model specified in Section 3.2.2 and the corresponding assumptions described in Section 3.2.3.

Lemma B.2. *Under the conditions specified in Section 3.2.2 for the measurement model and the corresponding assumptions described in Section 3.2.3, the updated joint p.g.fl. is given by*

$$\begin{aligned}G_k[\bar{h},h]=&\frac{\exp\left(v_{\Upsilon,k|k-1}\left[\bar{h}\,(1-p_{D,1})\exp\left(v_{\mathcal{E},k|k-1}[h-1\,|\,\cdot]\right)\right]\right)}{\exp\left(v_{\Upsilon,k|k-1}[1-p_{D,1}]\right)}\\&\times\frac{\exp\left(v_{\Upsilon,k|k-1}\left[\bar{h}\,p_{D,1}\exp\left(v_{\mathcal{E},k|k-1}[h(1-p_{D,2})-1\,|\,\cdot]\right)\right]\right)}{\exp\left(v_{\Upsilon,k|k-1}\left[p_{D,1}\exp\left(-v_{\mathcal{E},k|k-1}[p_{D,2}\,|\,\cdot]\right)\right]\right)}\\&\times\frac{\sum_{\pi\in\Pi_{Z_k}}\prod_{\varphi\in\pi}\left(\Lambda(\varphi)+v_{\Upsilon,k|k-1}\left[\bar{h}\,p_{D,1}\,\Delta_{\varphi}[h\,|\,\cdot]\right]\right)}{\sum_{\pi'\in\Pi_{Z_k}}\prod_{\varphi'\in\pi'}\left(\Lambda(\varphi')+v_{\Upsilon,k|k-1}[p_{D,1}\,L_{\varphi'}]\right)}.\end{aligned} \quad \text{(B.22)}$$

Proof. Given the measurement model specified in Section 3.2.2, the p.g.fl. corresponding to the likelihood $g_k(Z_k\,|\,\mathbb{X}_k)$ is given by

$$G_{Z,k}[g\,|\,\mathbb{X}_k]=G_{K,k}[g]\prod_{\mathbf{x}\in X_k}G_{\Theta,1,k}[g\,|\,\mathbf{x},\Xi_k],$$

In addition, the predicted state has a RFS representation given by $\mathcal{X}_k=\bigcup_{\mathbf{x}\in\Upsilon_k}\mathcal{E}_k^{\mathbf{x}}$, the predicted joint p.g.fl. has the form $G_{k|k-1}[\bar{h},h]=G_{\Upsilon,k|k-1}\left[\bar{h}\,G_{\mathcal{E},k|k-1}[h\,|\,\cdot]\right]$ which, along with the expression for $G_{Z,k}[g\,|\,\mathbb{X}_k]$ above, results in the following

expression for the trivariate p.g.fl.

$$F[g, \bar{h}, h] = G_{K,k}[g]\, G_{\Upsilon,k|k-1}\big[\bar{h}\, G_{\Theta,1,k}[g \,|\, \cdot, \Xi_k]\, G_{\mathcal{E},k|k-1}[h \,|\, \cdot]\big] .$$

The assumption on the RFS $\Theta_{1,k}$ described in Section 3.2.3 gives

$$\begin{aligned} &G_{\Upsilon,k|k-1}\big[\bar{h}\, G_{\Theta,1,k}[g \,|\, \cdot, \Xi_k]\, G_{\mathcal{E},k|k-1}[h \,|\, \cdot]\big] \\ &= G_{\Upsilon,k|k-1}\big[\bar{h}\,(1 - p_{D,1})\, G_{\mathcal{E},k|k-1}[h \,|\, \cdot] + \bar{h}\, p_{D,1}\, G_{Z,\mathbf{x},k}[g \,|\, \cdot, \Xi_k]\, G_{\mathcal{E},k|k-1}[h \,|\, \cdot]\big] , \end{aligned}$$

where $G_{Z,\mathbf{x},k}[g \,|\, \cdot, \Xi_k] = \prod_{\boldsymbol{\xi} \in \Xi_k} G_{\Theta,2,k}[g \,|\, \boldsymbol{\xi}, \cdot]$, so that

$$\begin{aligned} F[g, \bar{h}, h] = G_{K,k}[g] \times \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (B.23) \\ G_{\Upsilon,k|k-1}\big[\bar{h}\,(1 - p_{D,1})\, G_{\mathcal{E},k|k-1}[h \,|\, \cdot] + \bar{h}\, p_{D,1}\, G_{\mathcal{E},k|k-1}[h\; G_{\Theta,2,k}[g \,|\, :, \cdot]| \cdot]\big] . \end{aligned}$$

The m^{th}-order functional derivative of this in the directions $\delta_{\mathbf{z}_1}, \ldots, \delta_{\mathbf{z}_m}$ is then

$$\begin{aligned} &\delta^m F[g, \bar{h}, h\,;\, \delta_{\mathbf{z}_1}, \ldots, \delta_{\mathbf{z}_m}] = \\ &\quad \sum\nolimits_{W \subseteq Z_k} \delta^{m-|W|} G_{K,k}\big[g\,;\, \delta_{\mathbf{z}_1}, \ldots, \delta_{\mathbf{z}_{m-|W|}}\big] \qquad\qquad\qquad\qquad\qquad\qquad (B.24) \\ &\qquad\qquad \times \delta^{|W|} G_{\Upsilon,k|k-1}\big[\bar{h}\,(1 - p_{D,1})\, G_{\mathcal{E},k|k-1}[h \,|\, \cdot] + \\ &\qquad\qquad\qquad\qquad \bar{h}\, p_{D,1}\, G_{\mathcal{E},k|k-1}[h\; G_{\Theta,2,k}[g \,|\, :, \cdot]| \cdot]\,;\, \delta_{\mathbf{w}_1}, \ldots, \delta_{\mathbf{w}_{|W|}}\big], \end{aligned}$$

which follows from the general product rule for functional derivatives [69, Section 11.6, page 389].

The Poisson assumption on the RFS representation of the predicted state, as described in Section 3.2.3, implies that the predicted joint p.g.fl. in equation (B.23) has the following form

$$\begin{aligned} &G_{\Upsilon,k|k-1}\big[\bar{h}\,(1 - p_{D,1})\, G_{\mathcal{E},k|k-1}[h \,|\, \cdot] + \bar{h}\, p_{D,1}\, G_{\mathcal{E},k|k-1}[h\; G_{\Theta,2,k}[g \,|\, :, \cdot]| \cdot]\big] \\ &= \exp\big(v_{\Upsilon,k|k-1}\big[\bar{h}\,(1 - p_{D,1})\, G_{\mathcal{E},k|k-1}[h \,|\, \cdot] + \bar{h}\, p_{D,1}\, G_{\mathcal{E},k|k-1}[h\; G_{\Theta,2,k}[g \,|\, :, \cdot]| \cdot] - 1\big]\big) , \end{aligned}$$

where $G_{\mathcal{E},k|k-1}[h \,|\, \cdot] = \exp\big(v_{\mathcal{E},k|k-1}[h - 1 \,|\, \cdot]\big)$. Consequently, it can be shown by induction that

$$\begin{aligned} &\delta^{|W|} G_{\Upsilon,k|k-1}\big[\bar{h}\,(1 - p_{D,1})\, G_{\mathcal{E},k|k-1}[h \,|\, \cdot] \\ &\qquad\qquad + \bar{h}\, p_{D,1}\, G_{\mathcal{E},k|k-1}[h\; G_{\Theta,k}[g \,|\, :, \cdot]| \cdot]\,;\, \delta_{\mathbf{w}_1}, \ldots, \delta_{\mathbf{w}_{|W|}}\big] \\ &= \exp\big(v_{\Upsilon,k|k-1}\big[\bar{h}\,(1 - p_{D,1})\, G_{\mathcal{E},k|k-1}[h \,|\, \cdot] + \bar{h}\, p_{D,1}\, G_{\mathcal{E},k|k-1}[h\; G_{\Theta,2,k}[g \,|\, :, \cdot]| \cdot] - 1\big]\big) \end{aligned}$$

$$\times \sum_{\pi \in \Pi_W} \prod_{\varphi \in \pi} v_{\Upsilon,k|k-1}\Big[\bar{h}\, p_{D,1}\, \delta^{|\varphi|} G_{\mathcal{E},k|k-1}\Big[h\; G_{\Theta,2,k}[g \,|\, :,\cdot\,]\,; \delta_{\mathbf{z}_1}, \ldots, \delta_{\mathbf{z}_{|\varphi|}}\Big|\, \cdot\,\Big]\Big].$$

Furthermore, the assumption on the RFS $\Theta_{2,k}$ described in Section 3.2.3 implies that its p.g.fl. has the following form

$$G_{\Theta,2,k}[g \,|\, :,\cdot\,] = 1 - p_{D,2}(\,:\,) + p_{D,2}(\,:\,)\, l_{\mathbf{z}'}[g \,|\, :,\cdot\,],$$

denoting $l_{\mathbf{z}'}[g \,|\, :,\cdot\,] = \int g(\mathbf{z}')\, l_{\mathbf{z}'}(\,:,\cdot\,)\, \mathrm{d}\mathbf{z}'$, where $l_{\mathbf{z}'}(\,:,\cdot\,) = g_k(\mathbf{z}' \,|\, :,\cdot\,)$, so that

$$\begin{aligned}
&\delta^{|\varphi|} G_{\mathcal{E},k|k-1}\Big[h\; (1 - p_{D,2} + p_{D,2}\, l_{\mathbf{z}}[g \,|\, :,\cdot\,])\,; \delta_{\mathbf{z}_1}, \ldots, \delta_{\mathbf{z}_{|\varphi|}}\Big|\, \cdot\,\Big] \\
&= \exp\big(v_{\mathcal{E},k|k-1}[h\; (1 - p_{D,2} + p_{D,2}\, l_{\mathbf{z}}[g \,|\, :,\cdot\,]) - 1|\, \cdot\,]\big) \prod_{\mathbf{z} \in \varphi} v_{\mathcal{E},k|k-1}[h\, p_{D,2}\, l_{\mathbf{z}} \,|\, \cdot\,].
\end{aligned}$$

It follows, given the additional assumption on the clutter RFS K_k as described in Section 3.2.3 and noting that

$$p_\kappa(Z_k \setminus W) = \delta^{m-|W|} G_{K,k}\big[0\,; \delta_{\mathbf{z}_1}, \ldots, \delta_{\mathbf{z}_{m-|W|}}\big], \tag{B.25}$$

that the updated joint p.g.fl. given by equation (B.21) has the form

$$\begin{aligned}
G_k[\bar{h}, h] = {} & \frac{\exp\big(v_{\Upsilon,k|k-1}\big[\bar{h}\,(1 - p_{D,1})\, G_{\mathcal{E},k|k-1}[h \,|\, \cdot\,] + \bar{h}\, p_{D,1}\, G_{\mathcal{E},k|k-1}[h(1 - p_{D,2}) \,|\, \cdot\,]\big]\big)}{\exp\big(v_{\Upsilon,k|k-1}\big[1 - p_{D,1} + p_{D,1} \exp\big(-v_{\mathcal{E},k|k-1}[p_{D,2} \,|\, \cdot\,]\big)\big]\big)} \\
& \times \frac{\displaystyle\sum_{W \subseteq Z_k} \left(\sum_{\pi_\kappa \in \Pi_{Z_k \setminus W}} \prod_{\varphi_\kappa \in \pi_\kappa} \Lambda(\varphi_\kappa) \right) \left(\sum_{\pi \in \Pi_W} \prod_{\varphi \in \pi} v_{\Upsilon,k|k-1}\big[\bar{h}\, p_{D,1}\, \Delta_\varphi[h \,|\, \cdot\,]\big] \right)}{\displaystyle\sum_{W' \subseteq Z_k} \left(\sum_{\pi'_\kappa \in \Pi_{Z_k \setminus W'}} \prod_{\varphi'_\kappa \in \pi'_\kappa} \Lambda(\varphi'_\kappa) \right) \left(\sum_{\pi' \in \Pi'_W} \prod_{\varphi' \in \pi'} v_{\Upsilon,k|k-1}[p_{D,1}\, L_{\varphi'}] \right)},
\end{aligned}$$

where $L_\varphi = L_\varphi(\,\cdot\,)$ is the pseudo multiple measurement likelihood given by equation (3.24) and, for brevity, the following notation is introduced

$$\Delta_\varphi[h \,|\, \cdot\,] = \exp\big(v_{\mathcal{E},k|k-1}[h(1 - p_{D,2}) - 1 \,|\, \cdot\,]\big) \prod_{\mathbf{z} \in \varphi} v_{\mathcal{E},k|k-1}[h\, p_{D,2}\, l_{\mathbf{z}} \,|\, \cdot\,].$$

The summations over all subsets $W \subseteq Z_k$ and all partitions Π_W, $\Pi_{Z_k \setminus W}$ in the numerator of the expression for $G_k[\bar{h}, h]$ above can be re-written as a sum over all partitions Π_{Z_k}, and likewise for the denominator, to give the result in equation (B.22). □

The result for the updated intensity follows from the relation given in equation

(B.9). That is, taking the functional derivative of the expression for $G_k[\bar{h}, h]$ given in equation (B.22) in the direction $\delta_{\mathbf{x},\boldsymbol{\xi}}$ then evaluating at $\bar{h} = 1$ and $h = 1$, as follows. First consider each of the exponential terms in (B.22), the functional derivatives of which, in the direction $\delta_{\mathbf{x},\boldsymbol{\xi}}$, are found to be, respectively,

$$\begin{aligned}&\exp\left(v_{\Upsilon,k|k-1}\left[\bar{h}\,(1-p_{D,1})\exp\left(v_{\mathcal{E},k|k-1}[h-1\,|\,\cdot]\right)\right]\right)\\&\quad\times(1-p_{D,1}(\mathbf{x}))\exp\left(v_{\mathcal{E},k|k-1}[h-1\,|\,\mathbf{x}]\right)v_{\Upsilon,k|k-1}(\mathbf{x})\,v_{\mathcal{E},k|k-1}(\boldsymbol{\xi}\,|\,\mathbf{x}),\end{aligned}$$

and

$$\begin{aligned}&\exp\left(v_{\Upsilon,k|k-1}\left[\bar{h}\,p_{D,1}\exp\left(v_{\mathcal{E},k|k-1}[h\,(1-p_{D,2})-1\,|\,\cdot]\right)\right]\right)\times\\&p_{D,1}(\mathbf{x})\exp\left(v_{\mathcal{E},k|k-1}[h\,(1-p_{D,2})-1\,|\,\mathbf{x}]\right)(1-p_{D,2}(\boldsymbol{\xi}))\,v_{\Upsilon,k|k-1}(\mathbf{x})\,v_{\mathcal{E},k|k-1}(\boldsymbol{\xi}\,|\,\mathbf{x}).\end{aligned}$$

The functional derivative of the product term in equation (B.22), in the direction $\delta_{\mathbf{x},\boldsymbol{\xi}}$, is then

$$\prod_{\varphi\in\pi}\left(\Lambda(\varphi)+v_{\Upsilon,k|k-1}\left[\bar{h}\,p_{D,1}\,\Delta_{\varphi}[h\,|\,\cdot]\right]\right)\sum_{\varphi\in\pi}\frac{\delta v_{\Upsilon,k|k-1}\left[\bar{h}\,p_{D,1}\,\Delta_{\varphi}[h\,|\,\cdot]\,;\delta_{\mathbf{x},\boldsymbol{\xi}}\right]}{\Lambda(\varphi)+v_{\Upsilon,k|k-1}\left[\bar{h}\,p_{D,1}\,\Delta_{\varphi}[h\,|\,\cdot]\right]},\tag{B.26}$$

where the functional derivative of $v_{\Upsilon,k|k-1}\left[\bar{h}\,p_{D,1}\,\Delta_{\varphi}[h\,|\,\cdot]\right]$ in the direction $\delta_{\mathbf{x},\boldsymbol{\xi}}$ is found by applying the chain rule (Lemma B.1). That is

$$\delta v_{\Upsilon,k|k-1}\left[\bar{h}\,p_{D,1}\,\Delta_{\varphi}[h\,|\,\cdot]\,;\delta_{\mathbf{x},\boldsymbol{\xi}}\right]=\delta v_{\Upsilon,k|k-1}\left[\bar{h}\,p_{D,1}\,\Delta_{\varphi}[h\,|\,\cdot]\,;\delta\Delta_{\varphi}[h\,;\delta_{\boldsymbol{\xi}}\,|\,\cdot]\,\delta_{\mathbf{x}}\right],$$

and the functional derivative of $\Delta_{\varphi}[h\,|\,\cdot]$ in the direction $\delta_{\boldsymbol{\xi}}$ is given by

$$\begin{aligned}\delta\Delta_{\varphi}[h\,;\delta_{\boldsymbol{\xi}}\,|\,\cdot]&=\Delta_{\varphi}[h\,|\,\cdot]\,\delta v_{\mathcal{E},k|k-1}[h\,(1-p_{D,2})-1\,;\delta_{\boldsymbol{\xi}}\,|\,\cdot]\\&\quad+\Delta_{\varphi}[h\,|\,\cdot]\sum_{\mathbf{z}\in\varphi}\frac{\delta v_{\mathcal{E},k|k-1}[h\,p_{D,2}\,l_{\mathbf{z}}\,;\delta_{\boldsymbol{\xi}}\,|\,\cdot]}{v_{\mathcal{E},k|k-1}[h\,p_{D,2}\,l_{\mathbf{z}}\,|\,\cdot]}.\end{aligned}\tag{B.27}$$

Introducing the following notation

$$\delta G_k[h\,;\delta_{\boldsymbol{\xi}}\,|\,\cdot]_{\varphi}=\delta v_{\mathcal{E},k|k-1}[h\,(1-p_{D,2})-1\,;\delta_{\boldsymbol{\xi}}\,|\,\cdot]+\sum_{\mathbf{z}\in\varphi}\frac{\delta v_{\mathcal{E},k|k-1}[h\,p_{D,2}\,l_{\mathbf{z}}\,;\delta_{\boldsymbol{\xi}}\,|\,\cdot]}{v_{\mathcal{E},k|k-1}[h\,p_{D,2}\,l_{\mathbf{z}}\,|\,\cdot]},$$

then $\delta\Delta_{\varphi}[h\,;\delta_{\boldsymbol{\xi}}\,|\,\cdot]=\Delta_{\varphi}[h\,|\,\cdot]\,\delta G_k[h\,;\delta_{\boldsymbol{\xi}}\,|\,\cdot]_{\varphi}$, and so

$$\delta v_{\Upsilon,k|k-1}\left[\bar{h}\,p_{D,1}\,\Delta_{\varphi}[h\,|\,\cdot]\,;\delta_{\mathbf{x},\boldsymbol{\xi}}\right]=p_{D,1}(\mathbf{x})\,\Delta_{\varphi}[h\,|\,\mathbf{x}]\,v_{\Upsilon,k|k-1}(\mathbf{x})\,\delta G_k[h\,;\delta_{\boldsymbol{\xi}}\,|\,\mathbf{x}]_{\varphi}.$$

The functional derivative of $G_k[\bar{h}, h]$ in the direction of $\delta_{\mathbf{x},\boldsymbol{\xi}}$ is therefore

$$\begin{aligned}\delta G_k[\bar{h}, h\,; \delta_{\mathbf{x},\boldsymbol{\xi}}] = v_{k|k-1}(\mathbf{x}, \boldsymbol{\xi})\, G_k[\bar{h}, h]\Big\{&(1 - p_{D,1}(\mathbf{x})) \exp\left(v_{\mathcal{E},k|k-1}[h - 1 \,|\, \mathbf{x}]\right) \\ &+ p_{D,1}(\mathbf{x}) \exp\left(v_{\mathcal{E},k|k-1}[h\,(1 - p_{D,2}) - 1 \,|\, \mathbf{x}]\right)(1 - p_{D,2}(\boldsymbol{\xi}))\Big\} \\ &+ \frac{\exp\left(v_{\Upsilon,k|k-1}\left[\bar{h}\,(1 - p_{D,1}) \exp\left(v_{\mathcal{E},k|k-1}[h - 1 \,|\, \cdot]\right)\right]\right)}{\exp\left(v_{\Upsilon,k|k-1}[1 - p_{D,1}]\right)} \\ &\times \frac{\exp\left(v_{\Upsilon,k|k-1}\left[\bar{h}\, p_{D,1} \exp\left(v_{\mathcal{E},k|k-1}[h(1 - p_{D,2}) - 1 \,|\, \cdot]\right)\right]\right)}{\exp\left(v_{\Upsilon,k|k-1}\left[p_{D,1} \exp\left(-v_{\mathcal{E},k|k-1}[p_{D,2} \,|\, \cdot]\right)\right]\right)} \\ &\times \sum_{\pi \in \Pi_{Z_k}} \Omega_\pi(\bar{h}, h) \sum_{\varphi \in \pi} \frac{p_{D,1}(\mathbf{x})\, \Delta_\varphi[h \,|\, \mathbf{x}]\, v_{\Upsilon,k|k-1}(\mathbf{x})\, \delta G_k[h\,; \delta_{\boldsymbol{\xi}} \,|\, \mathbf{x}]_\varphi}{\Lambda(\varphi) + v_{\Upsilon,k|k-1}\left[\bar{h}\, p_{D,1}\, \Delta_\varphi[h \,|\, \cdot]\right]},\end{aligned}$$

where, for further brevity,

$$\Omega_\pi(\bar{h}, h) = \frac{\prod_{\varphi \in \pi}\Big(\Lambda(\varphi) + v_{\Upsilon,k|k-1}\left[\bar{h}, p_{D,1}\, \Delta_\varphi[h \,|\, \cdot]\right]\Big)}{\sum_{\pi' \in \Pi_{Z_k}} \prod_{\varphi' \in \pi'}\Big(\Lambda(\varphi') + v_{\Upsilon,k|k-1}[p_{D,1}\, L_{\varphi'}]\Big)}. \tag{B.28}$$

The result given in equation (3.21) follows by setting $\bar{h} = 1$ and $h = 1$ in the above expression for $\delta G_k[\bar{h}, h\,; \delta_{\mathbf{x},\boldsymbol{\xi}}]$, noting that $L_\varphi(\,\cdot\,) = \Delta_\varphi[1 \,|\, \cdot\,]$, $\varpi_\pi = \Omega_\pi(1, 1)$ and $\delta G_k[1\,; \delta_{\boldsymbol{\xi}} \,|\, \mathbf{x}]_\varphi = v_k(\boldsymbol{\xi} \,|\, \mathbf{x})_\varphi$ is the conditional updated intensity given by equation (3.22).

B.4 Derivation of the Single Group PHD Filter Recursion

This section summarises the derivation of the single group PHD filter recursion presented in Section 3.2.4. First note that, since a single group target is a special case of the multiple group target system as described in Section 3.2.4, the joint p.g.fl. of the corresponding RFS representation has the following form

$$G_{\widetilde{\mathcal{X}}}[\bar{h}, h] = G_\Upsilon\left[\bar{h}, G_\mathcal{E}[h \,|\, \cdot]\right] = p_\Upsilon\left[\bar{h}\, G_\mathcal{E}[h \,|\, \cdot]\right] = \int \bar{h}(\mathbf{x})\, G_\mathcal{E}[h \,|\, \mathbf{x}]\, p_\Upsilon(\mathbf{x})\, \mathrm{d}\mathbf{x}, \tag{B.29}$$

where $G_\mathcal{E}[h \,|\, \cdot\,]$ is the p.g.fl. of the subsidiary RFS $\mathcal{E}$ given by equation (B.3). This p.g.fl. formulation will be exploited in the derivation of the filtering equations for the single group PHD filter recursion as shown in Sections B.4.1 and B.4.2.

B.4.1 Single group PHD prediction

Given a RFS representation such that the underlying RFS Υ strictly contains a single group state and the re-specification of the dynamic model described in

Section 3.2.4, the predicted joint p.g.fl. for the single group target system is

$$\begin{aligned} G_{k|k-1}[\bar{h}, h] &= p_{\Upsilon,k-1}\big[G_{\Phi,k|k-1}\big[\bar{h}\ G_{\mathcal{E},k|k-1}[h\,|\,\cdot,:]\big|:\big]\big] \\ &= \int G_{\Phi,k|k-1}\big[\bar{h}\ G_{\mathcal{E},k|k-1}[h\,|\,\cdot,:]\big|:\big]\ p_{k-1}(\mathbf{x}_{k-1})\,\mathrm{d}\mathbf{x}_{k-1}, \end{aligned}$$

where $G_{\mathcal{E},k|k-1}[h\,|\,\cdot,:] = G_{\Gamma,\xi,k}[h\,|\,\cdot\,]\,G_{\mathcal{E},k-1}\big[G_{\Psi,k|k-1}[h\,|\,\cdot\cdot,:]\big|:\big]$, which is the same expression given in Section B.3.1 for the predicted p.g.fl. of the subsidiary RFS $\mathcal{E}_k$.

The result for the predicted intensity follows from the relation given in equation (B.9). That is, given the expression above for $G_{k|k-1}$, its functional derivative in the direction $\delta_{\mathbf{x},\boldsymbol{\xi}}$ is given by

$$\begin{aligned} \delta G_{k|k-1}[\bar{h}, h\,;\delta_{\mathbf{x},\boldsymbol{\xi}}] &= \delta p_{\Upsilon,k-1}\big[G_{\Phi,k|k-1}\big[\bar{h}\ G_{\mathcal{E},k|k-1}[h\,|\,\cdot,:]\big|:\big]\,;\delta_{\mathbf{x},\boldsymbol{\xi}}\big] \\ &= p_{\Upsilon,k-1}\big[\delta G_{\Phi,k|k-1}\big[\bar{h}\ G_{\mathcal{E},k|k-1}[h\,|\,\cdot,:]\,;\delta_{\mathbf{x},\boldsymbol{\xi}}\,\big|:\big]\big]\,. \end{aligned} \tag{B.30}$$

Give the assumption on the RFS $\Phi_{k|k-1}$ described in Section 3.2.4, the functional derivative of $G_{\Phi,k|k-1}$ in the direction $\delta_{\mathbf{x},\boldsymbol{\xi}}$ is

$$\begin{aligned} &\delta G_{\Phi,k|k-1}\big[\bar{h}\ G_{\mathcal{E},k|-1}[h\,|\,\cdot,:]\,;\delta G_{\mathcal{E},k|k-1}[h\,;\delta_{\boldsymbol{\xi}}\,|\,\cdot,:]\,\delta_{\mathbf{x}}\big|:\big] \\ &= f_{k|k-1}(\mathbf{x}\,|\,:)\,\delta G_{\mathcal{E},k|k-1}[h\,;\delta_{\boldsymbol{\xi}}\,|\,\mathbf{x},:]. \end{aligned}$$

Hence the result for the predicted intensity given in equation (3.26) immediately follows by substitution into equation (B.30) and evaluating at $h = 1$, noting that $\delta G_{\mathcal{E},k|k-1}[1\,;\delta_{\boldsymbol{\xi}}\,|\,\mathbf{x},:]$ results in the same expression as given in equation (B.19).

B.4.2 Single group PHD update

The derivation of this result follows that of the multi-group PHD update given in Section B.3.2. While the group cardinality is fixed at $|X_k| = 1$, the updated joint p.g.fl. can still be written in the form given in equation (B.21). However, given the RFS representation of the predicted single group state and the re-specification of the measurement model described in Section 3.2.4, the trivariate p.g.fl. is now given by

$$\begin{aligned} F[g, \bar{h}, h] &= G_{K,k}[g]\,p_{\Upsilon,k|k-1}\big[\bar{h}\,G_{\Theta,1,k}[g\,|\,\cdot\,,\Xi_k]\ G_{\mathcal{E},k|k-1}[h\,|\,\cdot]\big|\,\cdot\,\big] \\ &= G_{K,k}[g]\,p_{\Upsilon,k|k-1}\big[\bar{h}\,G_{\mathcal{E},k|k-1}[h\ G_{\Theta,2,k}[g\,|\,:,\cdot]|\,\cdot]\big] \\ &= p_{\Upsilon,k|k-1}\big[\bar{h}\,G_{K,k}[g]\,G_{\mathcal{E},k|k-1}[h\ G_{\Theta,2,k}[g\,|\,:,\cdot]|\,\cdot]\big]\,. \end{aligned}$$

The m^{th}-order functional derivative of this trivariate p.g.fl. in the directions $\delta_{\mathbf{z}_1}, \ldots, \delta_{\mathbf{z}_m}$ is therefore

$$\begin{aligned}
&\delta^m F[g, \bar{h}, h\,; \delta_{\mathbf{z}_1}, \ldots, \delta_{\mathbf{z}_m}] \\
&= \delta^m p_{\Upsilon,k|k-1}\big[\bar{h}\, G_{K,k}[g]\, G_{\mathcal{E},k|k-1}[h\, G_{\Theta,2,k}[g\,|\,:,\cdot\,]\,|\,\cdot\,]\,; \delta_{\mathbf{z}_1}, \ldots, \delta_{\mathbf{z}_m}\big] \\
&= p_{\Upsilon,k|k-1}\Big[\bar{h} \sum\nolimits_{W \subseteq Z_k} \delta^{m-|W|} G_{K,k}\big[g\,; \delta_{\mathbf{z}_1}, \ldots, \delta_{\mathbf{z}_{m-|W|}}\big] \\
&\qquad\qquad \times \delta^{|W|} G_{\mathcal{E},k|k-1}\big[h\, G_{\Theta,2,k}[g\,|\,:,\cdot\,]\,; \delta_{\mathbf{w}_1}, \ldots, \delta_{\mathbf{w}_{|W|}}\,\big|\,\cdot\,\big]\Big],
\end{aligned}$$

given the measurement set $Z_k = \{\mathbf{z}_1, \ldots, \mathbf{z}_m\}$.

The Poisson assumption on the RFS $\mathcal{E}_k$ in the RFS representation of the predicted single group state and the assumption on the RFS $\Theta_{2,k}$ described in Section 3.2.4 gives

$$\begin{aligned}
&\delta^{|W|} G_{\mathcal{E},k|k-1}\big[h\, G_{\Theta,2,k}[g\,|\,:,\cdot\,]\,; \delta_{\mathbf{w}_1}, \ldots, \delta_{\mathbf{w}_{|W|}}\,\big|\,\cdot\,\big] \\
&= \exp\big(v_{\mathcal{E},k|k-1}[h\,(1 - p_D + p_D\, l_{\mathbf{z}}[g\,|\,:,\cdot\,]) - 1|\,\cdot\,]\big) \prod_{\mathbf{w} \in W} v_{\mathcal{E},k|k-1}[h\, p_D\, l_{\mathbf{w}}\,|\,\cdot\,],
\end{aligned} \tag{B.31}$$

denoting $l_{\mathbf{z}}[g\,|\,:,\cdot\,] = \int g(\mathbf{z})\, l_{\mathbf{z}}(:,\cdot)\, \mathrm{d}\mathbf{z}$, where $l_{\mathbf{z}}(:,\cdot) = g_k(\mathbf{z}\,|\,:,\cdot)$. It follows, given the final assumption on the clutter RFS K_k as described in Section 3.2.4 and denoting $Z = Z_k$, that the updated joint p.g.fl. given by equation (B.21) has the form

$$G_k[\bar{h}, h] = \frac{p_{\Upsilon,k|k-1}\left[\bar{h} \exp\big(v_{\mathcal{E},k|k-1}[h(1-p_D) - 1\,|\,\cdot\,]\big) \sum_{W \subseteq Z} \left(\prod_{\bar{\mathbf{z}} \in Z \setminus W} \kappa(\bar{\mathbf{z}})\right) \prod_{\mathbf{w} \in W} v_{\mathcal{E},k|k-1}[h\, p_D\, l_{\mathbf{w}}\,|\,\cdot\,]\right]}{p_{\Upsilon,k|k-1}\left[\exp\big(-v_{\mathcal{E},k|k-1}[p_D\,|\,\cdot\,]\big) \sum_{W' \subseteq Z} \left(\prod_{\bar{\mathbf{z}}' \in Z \setminus W'} \kappa(\bar{\mathbf{z}}')\right) \prod_{\mathbf{w}' \in W'} v_{\mathcal{E},k|k-1}[p_D\, l_{\mathbf{w}'}\,|\,\cdot\,]\right]}.$$

This is equivalent to the following formulation

$$G_k[\bar{h}, h] = p_{\Upsilon,k|k-1}\big[\bar{h}\, \Delta_Z[h\,|\,\cdot\,]\big] \big/ p_{\Upsilon,k|k-1}[L_Z], \tag{B.32}$$

where, for brevity, the following notation is introduced

$$\Delta_Z[h\,|\,\cdot\,] = \exp\big(v_{\mathcal{E},k|k-1}[h(1-p_D) - 1\,|\,\cdot\,]\big) \prod\nolimits_{\mathbf{z} \in Z} \Big(\kappa(\mathbf{z}) + v_{\mathcal{E},k|k-1}[h\, p_D\, l_{\mathbf{z}}\,|\,\cdot\,]\Big),$$

and noting that $L_Z(\,\cdot\,) = \Delta_Z[1\,|\,\cdot\,]$ is the pseudo multiple measurement likelihood

given by equation (3.28).

The result for the updated intensity follows from the relation given in equation (B.9). That is, taking the functional derivative of the updated joint p.g.fl. $G_k[h]$ in the direction $\delta_{\mathbf{x},\boldsymbol{\xi}}$, then evaluating at $\bar{h} = 1$ and $h = 1$ as follows. Given the expression for $G_k[\bar{h}, h]$ in equation (B.32),

$$\begin{aligned} \delta G_k[\bar{h}, h\,;\delta_{\mathbf{x},\boldsymbol{\xi}}] &= \delta p_{\Upsilon,k|k-1}\big[\bar{h}\,\Delta_Z[h\,|\,\cdot\,]\,;\delta_{\mathbf{x},\boldsymbol{\xi}}\big] \,/\, p_{\Upsilon,k|k-1}[L_Z] \\ &= \delta p_{\Upsilon,k|k-1}\big[\bar{h}\,\Delta_Z[h\,|\,\cdot\,]\,;\delta\Delta_Z[h\,;\delta_{\boldsymbol{\xi}}\,|\,\cdot\,]\,\delta_{\mathbf{x}}\big] \,/\, p_{\Upsilon,k|k-1}[L_Z], \end{aligned} \tag{B.33}$$

where the functional derivative of $\Delta_Z[h\,|\,\cdot\,]$ in the direction $\delta_{\boldsymbol{\xi}}$ has a similar expression as given in equation (B.27). That is, $\delta\Delta_Z[h\,;\delta_{\boldsymbol{\xi}}\,|\,\cdot\,] = \Delta_Z[h\,|\,\cdot\,]\,\delta G_k[h\,;\delta_{\boldsymbol{\xi}}\,|\,\cdot\,]$ where

$$\delta G_k[h\,;\delta_{\boldsymbol{\xi}}\,|\,\cdot\,] = \delta v_{\mathcal{E},k|k-1}[h(1-p_D)-1\,;\delta_{\boldsymbol{\xi}}\,|\,\cdot] + \sum_{\mathbf{z}\in Z} \frac{\delta v_{\mathcal{E},k|k-1}[h\,p_D\,l_{\mathbf{z}}\,;\delta_{\boldsymbol{\xi}}\,|\,\cdot\,]}{\kappa(\mathbf{z}) + v_{\mathcal{E},k|k-1}[h\,p_D\,l_{\mathbf{z}}\,|\,\cdot\,]} \tag{B.34}$$

Hence, the result given in equation (3.27), immediately follows by substituting the expression for $\delta\Delta_Z[h\,;\delta_{\boldsymbol{\xi}}\,|\,\cdot\,]$ into equation (B.33), completing the functional differentiation in the direction $\delta_{\mathbf{x}}$ and setting $h = 1$, again noting that $L_Z(\,\cdot\,) = \Delta_Z[1\,|\,\cdot\,]$, and that $\delta G_k[1\,;\delta_{\boldsymbol{\xi}}\,|\,\cdot\,] = v_k(\boldsymbol{\xi}\,|\,\cdot\,)$ is the updated conditional intensity given by equation (3.29).

B.5 Derivation of the PHD Filter Recursion for the Single Cluster Existence Model

This section summarises the derivation of the filtering equations for the single cluster existence model presented in Section 3.2.5. The description of the single cluster existence model given in Section 3.2.5 results in the following joint p.g.fl. of the corresponding RFS representation

$$\begin{aligned} G_{\mathcal{X}}[\bar{h}, h] = G_{\Upsilon}\big[\bar{h}\,G_{\mathcal{E}}[h\,|\,\cdot\,]\big] &= 1-\rho+\rho\,p_{\Upsilon}\big[\bar{h}\,G_{\mathcal{E}}[h\,|\,\cdot\,]\big] \\ &= 1-\rho+\rho\int \bar{h}(\mathbf{x})\,G_{\mathcal{E}}[h\,|\,\mathbf{x}]\,p_{\Upsilon}(\mathbf{x})\,\mathrm{d}\mathbf{x}. \end{aligned} \tag{B.35}$$

since the underlying RFS Υ can be modelled with a Bernoulli RFS. The p.g.fl. of the subsidiary RFS $\mathcal{E}$ is given by equation (B.3), if the group exists, and the probability of existence is denoted by ρ. The joint p.g.fl. formulation given by equation (B.35) plays an important role in the derivation of the prediction and update equations

for the single cluster existence model as shown in Sections B.5.1 and B.5.2.

B.5.1 Single cluster existence prediction

Given the RFS representation and the re-specification of the dynamic model described in Section 3.2.5, the predicted joint p.g.fl. for the single cluster existence model is

$$G_{k|k-1}[\bar{h}, h] = (1 - \rho_{k-1})\, G_{\Gamma,k}[\bar{h}, h] + \rho_{k-1} \int G_{\Phi,k|k-1}\big[\bar{h}\; G_{\mathcal{E},k|k-1}[h \,|\, \cdot, :]\big|\, \mathbf{x}_{k-1}\big]\, p_{\Upsilon,k-1}(\mathbf{x}_{k-1})\, \mathrm{d}\mathbf{x}_{k-1},$$

where $G_{\mathcal{E},k|k-1}[h \,|\, \cdot, :]$ has the same formulation as given in Section B.3.1. That is, the predicted joint p.g.fl. is written as

$$\begin{aligned} G_{k|k-1}[\bar{h}, h] = (1 - \rho_{k-1})\, G_{\Gamma,k}[\bar{h}, h] + \\ \rho_{k-1}\, p_{\Upsilon,k-1}\big[G_{\Phi,k|k-1}\big[\bar{h}\; G_{\mathcal{E},k|k-1}[h \,|\, \cdot, :]\big|\, :\big]\big]\,, \end{aligned} \tag{B.36}$$

and its functional derivative in the direction $\delta_{\mathbf{x},\boldsymbol{\xi}}$ is given by

$$\begin{aligned} \delta G_{k|k-1}[\bar{h}, h\,; \delta_{\mathbf{x},\boldsymbol{\xi}}] = (1 - \rho_{k-1})\, \delta G_{\Gamma,k}[\bar{h}, h\,; \delta_{\mathbf{x},\boldsymbol{\xi}}] + \\ \rho_{k-1}\, p_{\Upsilon,k-1}\big[\delta G_{\Phi,k|k-1}\big[\bar{h},\, G_{\mathcal{E},k|k-1}[h \,|\, \cdot, :]\,; \delta_{\mathbf{x},\boldsymbol{\xi}}\,\big|\, :\big]\big]\,. \end{aligned} \tag{B.37}$$

Given the assumption on the RFS Γ_k described in Section 3.2.5, the joint p.g.fl. $G_{\Gamma,k}[\bar{h}, h]$ is then

$$G_{\Gamma,k}[\bar{h}, h] = 1 - p_B + p_B \int \bar{h}(\mathbf{x}_k)\, G_{\Gamma,\xi,k}[h \,|\, \mathbf{x}_k]\, p_{\gamma,k}(\mathbf{x}_k)\, \mathrm{d}\mathbf{x}_k, \tag{B.38}$$

and its functional derivative in the direction $\delta_{\mathbf{x},\boldsymbol{\xi}}$ is

$$\begin{aligned} \delta G_{\Gamma,k}[\bar{h}, h\,; \delta_{\mathbf{x},\boldsymbol{\xi}}] &= p_B\, \delta p_{\gamma,k}\big[\bar{h}\, G_{\Gamma,\xi,k}[h \,|\, \cdot]\,; \delta_{\mathbf{x},\boldsymbol{\xi}}\big] \\ &= p_B\, \delta p_{\gamma,k}\big[\bar{h}\, G_{\Gamma,\xi,k}[h \,|\, \cdot]\,; \delta G_{\Gamma,\xi,k}[h\,; \delta_{\boldsymbol{\xi}} \,|\, \cdot]\, \delta_{\mathbf{x}}\big]\,. \end{aligned}$$

Evaluating at $\bar{h} = 1$ and $h = 1$ gives $\delta G_{\Gamma,k}[1, 1\,; \delta_{\mathbf{x},\boldsymbol{\xi}}] = \gamma_k(\mathbf{x}, \boldsymbol{\xi}) = p_B\, p_{\gamma,k}(\mathbf{x})\, \gamma_{\xi,k}(\boldsymbol{\xi} \,|\, \mathbf{x})$. Furthermore, it follows from the assumption on the RFS $\Phi_{k|k-1}$ described in Section 3.2.5, that the functional derivative of $G_{\Phi,k|k-1}\big[\bar{h}\; G_{\mathcal{E},k|k-1}[h \,|\, \cdot, :]\big|\, :\big]$ in the direction $\delta_{\mathbf{x},\boldsymbol{\xi}}$ has the same expression as given in equation (B.18). The result for the predicted intensity given in equation (3.32) therefore follows by substitution, setting $\bar{h} = 1$ and $h = 1$ in equation (B.37) and further noting that $\delta G_{\mathcal{E},k|k-1}[1\,; \delta_{\mathbf{x},\boldsymbol{\xi}} \,|\, \cdot, :]$

results in the same expression as given in (B.19).

The predicted probability of existence, denoted by $\rho_{k|k-1}$, can be found by evaluating the predicted joint p.g.fl. $G_{k|k-1}$ at $\bar{h} = 0$ and $h = 0$, since $G_{k|k-1}[0,0] = \Pr(\mathbb{X}_k = \emptyset) = 1 - \rho_{k|k-1}$. That is, from equation (B.36),

$$G_{k|k-1}[0,0] = (1-\rho_{k-1})\, G_{\Gamma,k}[0,0] + \rho_{k-1}\, p_{\Upsilon,k-1}\big[G_{\Phi,k|k-1}[0\,|\,\cdot]\big]\,,$$

where setting $\bar{h} = 0$ anf $h = 0$ in equations (B.38) and (B.17) gives $G_{\Gamma,k}[0,0] = 1 - p_B$ and $G_{\Phi,k|k-1}[0\,|\,\cdot] = 1 - p_{S,1}(\,\cdot\,)$. The predicted probability of existence given in equation (3.33), follows immediately as a result of $\rho_{k|k-1} = 1 - G_{k|k-1}[0,0]$.

B.5.2 Single cluster existence update

The derivation of the results for the single cluster existence update, follows those given in Sections B.3.2 and B.4.2 for the multi-group and single group PHD updates respectively. As such, the updated joint p.g.fl. can still be written in the form given in equation (B.21). However, given the RFS representation of the predicted state and the re-specification of the measurement model described in Section 3.2.5, the trivariate p.g.fl. is now given by

$$\begin{aligned} F[g,\bar{h},h] = {} & (1-\rho_{k|k-1})\, G_{K,k}[g] \\ & + \rho_{k|k-1}\, p_{\Upsilon,k|k-1}\big[\bar{h}\, G_{K,k}[g]\, G_{\Theta,1,k}[g\,|\,\cdot\,;\Xi_k]\; G_{\mathcal{E},k|k-1}[h\,|\,\cdot\,]\big|\,\cdot\,\big]\,. \end{aligned}$$

It follows, from the assumption on the RFS $\Theta_{1,k}$ described in Section 3.2.5, that

$$G_{\Theta,1,k}[g\,|\,\cdot\,;\Xi_k] = 1 - p_{D,1}(\,\cdot\,) + p_{D,1}(\,\cdot\,)\prod\nolimits_{\boldsymbol{\xi}\in\Xi_k} G_{\Theta,2,k}[g\,|\,\boldsymbol{\xi},\,\cdot\,],$$

and so the trivariate p.g.fl. becomes

$$\begin{aligned} F[g,\bar{h},h] = {} & (1-\rho_{k|k-1})\, G_{K,k}[g] + \\ & \rho_{k|k-1}\, p_{\Upsilon,k|k-1}\big[\bar{h}\, G_{K,k}[g]\, \big((1-p_{D,1})\, G_{\mathcal{E},k|k-1}[h\,|\,\cdot\,] \\ & \qquad + p_{D,1}\, G_{\mathcal{E},k|k-1}[h\, G_{\Theta,2,k}[g\,|\,:,\cdot\,]|\,\cdot\,]\big)\big]. \end{aligned}$$

The m^{th}-order functional derivative this in the directions $\delta_{\mathbf{z}_1},\ldots,\delta_{\mathbf{z}_m}$ is then

$$\delta^m F[g,\bar{h},h\,;\delta_{\mathbf{z}_1},\ldots,\delta_{\mathbf{z}_m}] =$$

$$
\begin{aligned}
&(1-\rho_{k|k-1})\,\delta^m G_{K,k}[g\,;\delta_{\mathbf{z}_1},\ldots,\delta_{\mathbf{z}_m}] \\
&+\rho_{k|k-1}\,\delta^m p_{\Upsilon,k|k-1}\Big[\bar{h}\,G_{K,k}[g]\big((1-p_{D,1})\,G_{\mathcal{E},k|k-1}[h\,|\,\cdot\,] \\
&\qquad\qquad + p_{D,1}\,G_{\mathcal{E},k|k-1}[h\,G_{\Theta,2,k}[g\,|\,:,\cdot\,]|\,\cdot\,]\big);\delta_{\mathbf{z}_1},\ldots,\delta_{\mathbf{z}_m}\Big],
\end{aligned}
$$

given the measurement set $Z_k=\{\mathbf{z}_1,\ldots,\mathbf{z}_m\}$. The second of the two m^{th}-order functional derivative in the expression above can be written as

$$
\begin{aligned}
p_{\Upsilon,k|k-1}\Big[&\delta^m G_{K,k}[g\,;\delta_{\mathbf{z}_1},\ldots,\delta_{\mathbf{z}_m}]\,\bar{h}\,(1-p_{D,1})\,G_{\mathcal{E},k|k-1}[h\,|\,\cdot\,]+ \\
&\sum\nolimits_{W\subseteq Z_k}\bar{h}\,p_{D,1}\,\delta^{m-|W|}G_{K,k}\big[g\,;\delta_{\mathbf{z}_1},\ldots,\delta_{\mathbf{z}_{m-|W|}}\big] \\
&\qquad\times\delta^{|W|}G_{\mathcal{E},k|k-1}\big[h\,G_{\Theta,2,k}[g\,|\,:,\cdot\,]\,;\delta_{\mathbf{w}_1},\ldots,\delta_{\mathbf{w}_{|W|}}\,|\,\cdot\,\big]\Big],
\end{aligned}
$$

where $\delta^{|W|}G_{\mathcal{E},k|k-1}\big[h\,G_{\Theta,2,k}[g\,|\,:,\cdot\,]\,;\delta_{\mathbf{w}_1},\ldots,\delta_{\mathbf{w}_{|W|}}\,|\,\cdot\,\big]$ has the same expression as given in equation (B.31) (except with $p_{D,2}$ replacing p_D), as a result of the Poisson assumption on the RFS $\mathcal{E}_k$ in the RFS representation of the predicted state and the assumption on the RFS $\Theta_{2,k}$ described in Section 3.2.5.

It follows, given the final assumption on the clutter RFS K_k as described in Section 3.2.5, that the updated joint p.g.fl. given by equation (B.21) has the form

$$
\begin{aligned}
&G_k[\bar{h},h]= \\
&\frac{1-\rho_{k|k-1}+\rho_{k|k-1}\,p_{\Upsilon,k|k-1}\Big[\bar{h}\,\Big((1-p_{D,1})\,G_{\mathcal{E},k|k-1}[h\,|\,\cdot\,]+p_{D,1}\,\Delta_{Z_k}[h\,|\,\cdot\,]\Big/\prod\nolimits_{\mathbf{z}\in Z_k}\kappa(\mathbf{z})\Big)\Big]}{1-\rho_{k|k-1}+\rho_{k|k-1}\,p_{\Upsilon,k|k-1}\Big[(1-p_{D,1})+p_{D,1}\,L_{Z_k}\Big/\prod\nolimits_{\mathbf{z}\in Z_k}\kappa(\mathbf{z})\Big]},
\end{aligned}
\tag{B.39}
$$

where $\Delta_{Z_k}[h\,|\,\cdot\,]$ is the same notation as introduced in Section B.4.2 (except with $p_{D,2}$ replacing p_D) such that $L_{Z_k}(\,\cdot\,)=\Delta_{Z_k}[1\,|\,\cdot\,]$ is the pseudo multiple measurement likelihood given by equation (3.28). The functional derivative of $G_k[\bar{h},h]$ in the direction $\delta_{\mathbf{x},\boldsymbol{\xi}}$, given the expression above, is therefore

$$
\begin{aligned}
&\delta G_k[\bar{h},h\,;\delta_{\mathbf{x},\boldsymbol{\xi}}]= \\
&\frac{\rho_{k|k-1}\,\delta p_{\Upsilon,k|k-1}\Big[\bar{h}\,\Big((1-p_{D,1})\,G_{\mathcal{E},k|k-1}[h\,|\,\cdot\,]+p_{D,1}\,\Delta_{Z_k}[h\,|\,\cdot\,]\Big/\prod\nolimits_{\mathbf{z}\in Z_k}\kappa(\mathbf{z})\Big)\,;\delta_{\mathbf{x},\boldsymbol{\xi}}\Big]}{1-\rho_{k|k-1}+\rho_{k|k-1}\,p_{\Upsilon,k|k-1}\Big[1-p_{D,1}+p_{D,1}\,L_{Z_k}\Big/\prod\nolimits_{\mathbf{z}\in Z_k}\kappa(\mathbf{z})\Big]},
\end{aligned}
$$

where

$$
\begin{aligned}
&\delta p_{\Upsilon,k|k-1}\Big[\bar{h}\left((1-p_{D,1})\,G_{\mathcal{E},k|k-1}[h\,|\,\cdot\,]+p_{D,1}\,\Delta_{Z_k}[h\,|\,\cdot\,]\Big/\prod\nolimits_{\mathbf{z}\in Z_k}\kappa(\mathbf{z})\right);\delta_{\mathbf{x},\boldsymbol{\xi}}\Big]\\
&=\delta p_{\Upsilon,k|k-1}\big[\bar{h}\,(1-p_{D,1})\,G_{\mathcal{E},k|k-1}[h\,|\,\cdot\,]\,;\delta G_{\mathcal{E},k|k-1}[h\,;\delta_{\boldsymbol{\xi}}\,|\,\cdot\,]\,\delta_{\mathbf{x}}\big]\\
&\quad+\delta p_{\Upsilon,k|k-1}\Big[\bar{h}\,p_{D,1}\,\Delta_{Z_k}[h\,|\,\cdot\,]\Big/\prod\nolimits_{\mathbf{z}\in Z_k}\kappa(\mathbf{z})\;;\delta\Delta_{Z_k}[h\,;\delta_{\boldsymbol{\xi}}\,|\,\cdot\,]\,\delta_{\mathbf{x}}\Big]\,.
\end{aligned}
$$

Note that $\delta\Delta_{Z_k}[h\,;\delta_{\boldsymbol{\xi}}\,|\,\cdot\,]=\Delta_{Z_k}[h\,|\,\cdot\,]\,\delta G_k[h\,;\delta_{\boldsymbol{\xi}}\,|\,\cdot\,]$, where $\delta G_k[h\,;\delta_{\boldsymbol{\xi}}\,|\,\cdot\,]$ has the same expression as given in equation (B.34). The result for the updated intensity given in equation (3.35) immediately follows by substitution into the expression for $\delta G_k[\bar{h},h\,;\delta_{\mathbf{x},\boldsymbol{\xi}}]$, completing the functional differentiation in the direction $\delta_{\mathbf{x}}$ and setting $h=1$, also noting that $\delta G_{\mathcal{E},k|k-1}[1\,;\delta_{\boldsymbol{\xi}}\,|\,\cdot\,]=v_{k|k-1}(\boldsymbol{\xi}\,|\,\cdot\,)$,

$$
\tilde{L}_{Z_k}(\,\cdot\,)=L_{Z_k}(\,\cdot\,)\Big/\prod\nolimits_{\mathbf{z}\in Z_k}\kappa(\mathbf{z})\,,
$$

and $\delta G_k[1\,;\delta_{\boldsymbol{\xi}}\,|\,\cdot\,]=v_k(\boldsymbol{\xi}\,|\,\cdot\,)$ is the conditional updated intensity given in equation (3.29).

Finally, the updated probability of existence given in equation (3.37) simply follows as a result of $\rho_k=1-G_k[0]$, setting $\bar{h}=0$ and $h=0$ in equation (B.39).

B.6 Derivation of the PHD Filter Recursion for Extended Targets

This section summarises the derivation of the PHD filter recursion for extended targets, i.e. the marginal group target PHD recursion, presented in Section 3.2.6. Sketch proofs are given in Sections B.6.1 and B.6.2 respectively for the results of the prediction and update equations.

B.6.1 PHD prediction for extended targets

The derivation of this results follows that of the multi-group PHD prediction in Section B.3.1, although now the relation given in equation (B.12) is utilised instead. That is, taking the functional derivative of the predicted joint p.g.fl. $G_{k|k-1}[\bar{h},h]$ in the direction $\delta_{\mathbf{x}}$, then evaluating at $\bar{h}=1$ and $h=1$. The predicted joint p.g.fl. has the same form as given by equation (B.14), although since there is no interest in the evolution of the subsidiary RFSs $\mathcal{E}_{k-1}$ it is not necessary to specify a formulation for the p.g.fl. $G_{\mathcal{E},k|k-1}$. Hence

$$
\begin{aligned}
\delta G_{k|k-1}[\bar{h},h\,;\delta_{\mathbf{x}}]&=\delta G_{\Gamma,k}[\bar{h},h\,;\delta_{\mathbf{x}}]\,G_{\Upsilon,k-1}\big[G_{\Phi,k|k-1}\big[\bar{h}\;G_{\mathcal{E},k|k-1}[h\,|\,\cdot,:]\big|:\big]\big]\\
&\quad+G_{\Gamma,k}[\bar{h},h]\,\delta G_{\Upsilon,k-1}\big[G_{\Phi,k|k-1}\big[\bar{h}\;G_{\mathcal{E},k|k-1}[h\,|\,\cdot,:]\big|:\big]\,;\delta_{\mathbf{x}}\big]\,.
\end{aligned}
$$

The chain rule (Lemma B.1) can again be applied to the functional derivative of $G_{\Upsilon,k-1}\big[G_{\Phi,k|k-1}\big[G_{\mathcal{E},k|k-1}[h\,|\,\cdot,:]\big|:\big]\big]$ in the direction $\delta_{\mathbf{x}}$ as follows

$$\begin{aligned}&\delta G_{\Upsilon,k-1}\big[G_{\Phi,k|k-1}\big[\bar{h}\;G_{\mathcal{E},k|k-1}[h\,|\,\cdot,:]\big|:\big]\,;\delta_{\mathbf{x}}\big]\\ &=\delta G_{\Upsilon,k-1}\big[G_{\Phi,k|k-1}\big[\bar{h}\;G_{\mathcal{E},k|k-1}[h\,|\,\cdot,:]\big|:\big]\,;\delta G_{\Phi,k|k-1}\big[\bar{h}\;G_{\mathcal{E},k|k-1}[h\,|\,\cdot,:]\,;\delta_{\mathbf{x}}\big|:\big]\big]\,,\end{aligned}$$

which can be written in a similar form as the expression given in equation (B.16).

The assumption on the RFS $\Phi_{k|k-1}$ described in Section 3.2.6 implies that the p.g.fl. $G_{\Phi,k|k-1}\big[G_{\mathcal{E},k|k-1}[h\,|\,\cdot,:]\big|:\big]$ has the same form as given in equation (B.17) (except with p_S replacing $p_{S,1}$) so that

$$\delta G_{\Phi,k|k-1}\big[\bar{h}\;G_{\mathcal{E},k|k-1}[h\,|\,\cdot,:]\,;\delta_{\mathbf{x}}\big|:\big]=p_S(:)\,f_{k|k-1}(\mathbf{x}\,|:)\,G_{\mathcal{E},k|k-1}[h\,|\,\mathbf{x},:].$$

It subsequently follows from Campbell's theorem (Section A.2) that

$$\begin{aligned}&\delta G_{\Upsilon,k-1}\big[1\,;\delta G_{\Phi,k|k-1}\big[G_{\mathcal{E},k|k-1}[1\,|\,\cdot,:]\,;\delta_{\mathbf{x}}\,|:\big]\big]\\ &=\int p_S(\mathbf{x})\,f_{k|k-1}(\mathbf{x}\,|\,\mathbf{x}_{k-1})\,v_{k-1}(\mathbf{x}_{k-1})\,\mathrm{d}\mathbf{x}_{k-1},\end{aligned}$$

since $G_{\mathcal{E},k|k-1}[1\,|\,\cdot,:]=1$. The result for the predicted intensity given in equation (3.38) therefore follows by further noting that $\delta G_{\Gamma,k}[1,1\,;\delta_{\mathbf{x}}]=\gamma_k(\mathbf{x})$.

B.6.2 PHD update for extended targets

The derivation of this result follows that of the multi-group PHD update in Section B.3.2, although, like in the previous subsection, the relation given in equation (B.12) is now utilised instead. That is, taking the functional derivative of the updated joint p.g.fl. $G_k[\bar{h},h]$ in the direction $\delta_{\mathbf{x}}$, then evaluating at $\bar{h}=1$ and $h=1$. Given the assumption regarding the measurement model described in Section 3.2.6, the updated joint p.g.fl. has the same form as the expression given in equation (B.22) (except with μ replacing $v_{\mathcal{E},k|k-1}$).

First consider each of the exponential terms in (B.22), the functional derivatives of which, in the direction $\delta_{\mathbf{x}}$, are respectively found to be

$$\begin{aligned}&\exp\big(v_{\Upsilon,k|k-1}\big[\bar{h}\,(1-p_{D,1})\exp\left(\mu[h-1\,|\,\cdot\,]\right)\big]\big)\\ &\quad\times(1-p_{D,1}(\mathbf{x}))\exp\left(\mu[h-1\,|\,\cdot\,]\right)\;v_{\Upsilon,k|k-1}(\mathbf{x}),\end{aligned}$$

and

$$\exp\big(v_{\Upsilon,k|k-1}\big[\bar{h}\,p_{D,1}\exp\left(\mu[h(1-p_{D,2})-1\,|\,\cdot\,]\right)\big]\big)$$
$$\times\, p_{D,1}(\mathbf{x})\exp\left(\mu[h(1-p_{D,2})-1\,|\,\mathbf{x}\,]\right)\, v_{\Upsilon,k|k-1}(\mathbf{x}).$$

The functional derivative of the product terms in equation (B.22), in the direction $\delta_{\mathbf{x}}$, has a similar form as the expression given in equation (B.26) where the numerator in the summation over subsets $\varphi \in \pi$ is now

$$\delta v_{\Upsilon,k|k-1}[\bar{h}\,p_{D,1}\,\Delta_{\varphi}[h\,|\,\cdot\,]\,;\delta_{\mathbf{x}}] = p_{D,1}(\mathbf{x})\,\Delta_{\varphi}[h\,|\,\mathbf{x}]\,v_{\Upsilon,k|k-1}(\mathbf{x}).$$

The functional derivative of $G_k[\bar{h},h]$ in the direction $\delta_{\mathbf{x}}$ is therefore

$$\begin{aligned}\delta G_k[\bar{h},h\,;\delta_{\mathbf{x}}] = v_{k|k-1}(\mathbf{x})\,G_k[\bar{h},h]\big\{1-p_{D,1}(\mathbf{x})+p_{D,1}(\mathbf{x})\exp\left(\mu[h(1-p_{D,2})-1\,|\,\mathbf{x}\,]\right)\big\}\\ +\frac{\exp\big(v_{k|k-1}\big[\bar{h}\,(1-p_{D,1})\exp\left(\mu[h-1\,|\,\cdot\,]\right)\big]\big)}{\exp\big(v_{k|k-1}[1-p_{D,1}]\big)}\\ \times\frac{\exp\big(v_{k|k-1}\big[\bar{h}\,p_{D,1}\exp\left(\mu[h(1-p_{D,2})-1\,|\,\cdot\,]\right)\big]\big)}{\exp\big(v_{k|k-1}[p_{D,1}\exp\left(-\mu[p_{D,2}\,|\,\cdot\,]\right)]\big)}\\ \times\sum_{\pi\in\Pi_{Z_k}}\Omega_{\pi}(\bar{h},h)\sum_{\varphi\in\pi}\frac{p_{D,1}(\mathbf{x})\,\Delta_{\varphi}[h\,|\,\mathbf{x}]\,v_{k|k-1}(\mathbf{x})}{\Lambda(\varphi)+v_{k|k-1}\big[\bar{h}\,p_{D,1}\,\Delta_{\varphi}[h\,|\,\cdot\,]\big]},\end{aligned}$$

where $\Omega_{\pi}(\bar{h},h)$ has the same form as the expression given in equation (B.28) and denoting $v_{k|k-1}(\mathbf{x}) = v_{\Upsilon,k|k-1}(\mathbf{x})$. The result for the updated intensity given in equation (3.39) follows by setting $\bar{h}=1$ and $h=1$ in the above expression for $\delta G_k[\bar{h},h\,;\delta_{\mathbf{x}}]$, noting that $L_{\varphi}(\,\cdot\,)=\Delta_{\varphi}[1\,|\,\cdot\,]$ and $\varpi_{\pi}=\Omega_{\pi}(1,1)$.

Appendix C

Selected Identities of Multivariate Gaussian Distributions

This chapter summarises some fundamental identities of multivariate Gaussian (normal) distributions. Details are given on joint, marginal and conditional multivariate Gaussian distributions in Section C.1, while Section C.2 presents two fundamental identities for the Gaussian products.

The density of a multivariate Gaussian, denoted as $\mathcal{N}(\mathbf{X}\,;\boldsymbol{\mu},\boldsymbol{\Sigma})$ with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$, is given by

$$p(\mathbf{X}) = \mathcal{N}(\mathbf{X}\,;\boldsymbol{\mu},\boldsymbol{\Sigma}) \propto \exp\left(-\frac{1}{2}(\mathbf{X}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{X}-\boldsymbol{\mu})\right) \tag{C.1}$$

for $n \times 1$ vectors $\mathbf{X}$, $\boldsymbol{\mu}$ and $n \times n$ matrix $\boldsymbol{\Sigma}$.

C.1 Conditional Multivariate Gaussian Distributions

This section includes details on formulating conditional multivariate Gaussian distributions, firstly from joint multivariate Gaussian distributions in Section C.1.1 and then finally using linear regression in Section C.1.2.

C.1.1 Deriving a conditional Gaussian from a joint Gaussian

Suppose $\mathbf{X}$ denotes the state variable consisting of two components $\begin{bmatrix}\boldsymbol{\xi} & \mathbf{x}\end{bmatrix}^T$ and is distributed according to a multivariate Gaussian with mean $\boldsymbol{\mu} = \begin{bmatrix}\mathbf{m}_{\boldsymbol{\xi}} & \mathbf{m}_{\mathbf{x}}\end{bmatrix}^T$ and covariance matrix $\boldsymbol{\Sigma}$ given by a block matrix of the following form

$$\boldsymbol{\Sigma} = \begin{bmatrix}\mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{12}^T & \mathbf{P}_{22}\end{bmatrix}, \tag{C.2}$$

for $q \times 1$ vectors $\boldsymbol{\xi}$, $\mathbf{m}_{\boldsymbol{\xi}}$, $q \times q$ matrix $\mathbf{P}_{11}$, $(n-q)\times 1$ vectors $\mathbf{x}$, $\mathbf{m}_{\mathbf{x}}$, $(n-q)\times(n-q)$ matrix $\mathbf{P}_{22}$ and $q \times (n-q)$ matrix $\mathbf{P}_{12}$. The density can then be written as a joint

multivariate Gaussian denoted as $p(\boldsymbol{\xi}, \mathbf{x}) = p(\mathbf{X})$ and given by

$$p(\boldsymbol{\xi}, \mathbf{x}) \propto \exp\left(-\frac{1}{2}\begin{bmatrix}\boldsymbol{\xi} - \mathbf{m}_{\boldsymbol{\xi}} \\ \mathbf{x} - \mathbf{m}_{\mathbf{x}}\end{bmatrix}^T \begin{bmatrix}\mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{12}^T & \mathbf{P}_{22}\end{bmatrix}^{-1} \begin{bmatrix}\boldsymbol{\xi} - \mathbf{m}_{\boldsymbol{\xi}} \\ \mathbf{x} - \mathbf{m}_{\mathbf{x}}\end{bmatrix}\right), \tag{C.3}$$

which follows directly from equation (C.1). The marginal distributions can be easily found to be

$$p(\boldsymbol{\xi}) = \int p(\boldsymbol{\xi}, \mathbf{x}) \mathrm{d}\mathbf{x} = \mathcal{N}(\boldsymbol{\xi}\,; \mathbf{m}_{\boldsymbol{\xi}}, \mathbf{P}_{11}), \tag{C.4}$$

$$p(\mathbf{x}) = \int p(\boldsymbol{\xi}, \mathbf{x}) \mathrm{d}\boldsymbol{\xi} = \mathcal{N}(\mathbf{x}\,; \mathbf{m}_{\mathbf{x}}, \mathbf{P}_{22}). \tag{C.5}$$

Now a conditional probability function is defined by $p(\boldsymbol{\xi} \,|\, \mathbf{x}) = p(\boldsymbol{\xi}, \mathbf{x}) / \, p(\mathbf{x})$ and this result is used, along with a decomposition result for the covariance matrix $\boldsymbol{\Sigma}$, to recover the conditional distribution $p(\boldsymbol{\xi} \,|\, \mathbf{x})$, given in the following Theorem, from the joint Gaussian in (C.3).

Theorem C.1. *Given a joint multivariate Gaussian* $p(\boldsymbol{\xi}, \mathbf{x})$ *defined in equation* (C.3) *and marginal distribution* $p(\mathbf{x})$ *defined in equation* (C.5)*, the conditional multivariate Gaussian is*

$$p(\boldsymbol{\xi} \,|\, \mathbf{x}) = \mathcal{N}\left(\boldsymbol{\xi}\,; \mathbf{m}_{\boldsymbol{\xi}|\mathbf{x}}, \mathbf{P}_{\text{Schur}}\right) \propto \exp\left(-\frac{1}{2}\left(\boldsymbol{\xi} - \mathbf{m}_{\boldsymbol{\xi}|\mathbf{x}}\right)^T \mathbf{P}_{\text{Schur}}^{-1}\left(\boldsymbol{\xi} - \mathbf{m}_{\boldsymbol{\xi}|\mathbf{x}}\right)\right), \tag{C.6}$$

where $\mathbf{m}_{\boldsymbol{\xi}|\mathbf{x}} = \mathbf{m}_{\boldsymbol{\xi}} - \mathbf{P}_{12}\mathbf{P}_{22}^{-1}(\mathbf{x} - \mathbf{m}_{\mathbf{x}})$ *and* $\mathbf{P}_{\text{Schur}} = \mathbf{P}_{11} - \mathbf{P}_{12}\mathbf{P}_{22}^{-1}\mathbf{P}_{12}^T$.

Proof. To realise this, arising as the result of performing block Gaussian elimination on $\boldsymbol{\Sigma}$, the Schur complement [139] of $\mathbf{P}_{22}$, given by $\mathbf{P}_{11} - \mathbf{P}_{12}\mathbf{P}_{22}^{-1}\mathbf{P}_{12}^T$, is utilised. That is multiplying $\boldsymbol{\Sigma}$ from the right with the block lower triangular matrix

$$\mathbf{L} = \begin{bmatrix}\mathbf{I} & \mathbf{0} \\ -\mathbf{P}_{22}^{-1}\mathbf{P}_{12}^T & \mathbf{I}\end{bmatrix}, \tag{C.7}$$

gives the following

$$\begin{aligned}\boldsymbol{\Sigma}\mathbf{L} = \begin{bmatrix}\mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{12}^T & \mathbf{P}_{22}\end{bmatrix}\begin{bmatrix}\mathbf{I} & \mathbf{0} \\ -\mathbf{P}_{22}^{-1}\mathbf{P}_{12}^T & \mathbf{I}\end{bmatrix} &= \begin{bmatrix}\mathbf{P}_{11} - \mathbf{P}_{12}\mathbf{P}_{22}^{-1}\mathbf{P}_{12}^T & \mathbf{P}_{12} \\ \mathbf{0} & \mathbf{P}_{22}\end{bmatrix} \\ &= \begin{bmatrix}\mathbf{I} & \mathbf{P}_{12}\mathbf{P}_{22}^{-1} \\ \mathbf{0} & \mathbf{I}\end{bmatrix}\begin{bmatrix}\mathbf{P}_{11} - \mathbf{P}_{12}\mathbf{P}_{22}^{-1}\mathbf{P}_{12}^T & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_{22}\end{bmatrix}.\end{aligned} \tag{C.8}$$

Therefore

$$\boldsymbol{\Sigma} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{12}^T & \mathbf{P}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{P}_{12}\mathbf{P}_{22}^{-1} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{P}_{11} - \mathbf{P}_{12}\mathbf{P}_{22}^{-1}\mathbf{P}_{12}^T & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{P}_{22}^{-1}\mathbf{P}_{12}^T & \mathbf{I} \end{bmatrix}, \quad \text{(C.9)}$$

and the inverse of this is

$$\begin{aligned} \boldsymbol{\Sigma}^{-1} &= \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{12}^T & \mathbf{P}_{22} \end{bmatrix}^{-1} \\ &= \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{P}_{22}^{-1}\mathbf{P}_{12}^T & \mathbf{I} \end{bmatrix} \begin{bmatrix} \left(\mathbf{P}_{11} - \mathbf{P}_{12}\mathbf{P}_{22}^{-1}\mathbf{P}_{12}^T\right)^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_{22}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{I} & -\mathbf{P}_{12}\mathbf{P}_{22}^{-1} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}. \end{aligned} \quad \text{(C.10)}$$

Substituting this into (C.3) gives

$$\begin{aligned} p(\boldsymbol{\xi}, \mathbf{x}) \propto \exp\Bigg(-\frac{1}{2} & \begin{bmatrix} \boldsymbol{\xi} - \mathbf{m}_{\boldsymbol{\xi}} - \mathbf{P}_{12}\mathbf{P}_{22}^{-1}(\mathbf{x} - \mathbf{m}_{\mathbf{x}}) \\ \mathbf{x} - \mathbf{m}_{\mathbf{x}} \end{bmatrix}^T \\ & \times \begin{bmatrix} \left(\mathbf{P}_{11} - \mathbf{P}_{12}\mathbf{P}_{22}^{-1}\mathbf{P}_{12}^T\right)^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_{22}^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{\xi} - \mathbf{m}_{\boldsymbol{\xi}} - \mathbf{P}_{12}\mathbf{P}_{22}^{-1}(\mathbf{x} - \mathbf{m}_{\mathbf{x}}) \\ \mathbf{x} - \mathbf{m}_{\mathbf{x}} \end{bmatrix} \Bigg). \end{aligned} \quad \text{(C.11)}$$

Denoting $\mathbf{m}_{\boldsymbol{\xi}|\mathbf{x}} = \mathbf{m}_{\boldsymbol{\xi}} - \mathbf{P}_{12}\mathbf{P}_{22}^{-1}(\mathbf{x} - \mathbf{m}_{\mathbf{x}})$ and $\mathbf{P}_{\text{Schur}} = \mathbf{P}_{11} - \mathbf{P}_{12}\mathbf{P}_{22}^{-1}\mathbf{P}_{12}^T$, then multiplying out these matrices within the exponential in the formulation above gives

$$\begin{aligned} & p(\boldsymbol{\xi}, \mathbf{x}) \propto \\ & \exp\left(-\frac{1}{2} \left(\boldsymbol{\xi} - \mathbf{m}_{\boldsymbol{\xi}|\mathbf{x}}\right)^T \mathbf{P}_{\text{Schur}}^{-1} \left(\boldsymbol{\xi} - \mathbf{m}_{\boldsymbol{\xi}|\mathbf{x}}\right) \right) \exp\left(-\frac{1}{2} (\mathbf{x} - \mathbf{m}_{\mathbf{x}})^T \mathbf{P}_{22}^{-1} (\mathbf{x} - \mathbf{m}_{\mathbf{x}}) \right). \end{aligned} \quad \text{(C.12)}$$

The desired result follows immediately, by applying the result given by $p(\boldsymbol{\xi} \,|\, \mathbf{x}) = p(\boldsymbol{\xi}, \mathbf{x}) / \, p(\mathbf{x})$ to the formulation above. □

The conditional mean in (C.6) $\mathbf{m}_{\boldsymbol{\xi}} + \mathbf{P}_{12}\mathbf{P}_{22}^{-1}(\mathbf{x} - \mathbf{m}_{\mathbf{x}})$ is interpreted as the linear regression equation where $\mathbf{P}_{12}\mathbf{P}_{22}^{-1}$ is the regression coefficient. In fact, there is an alternative way of formulating the conditional Gaussian in (C.6) via linear regression, as described in the next section.

C.1.2 Formulating a conditional Gaussian using linear regression

Consider two variable states $\mathbf{x}$ and $\boldsymbol{\xi}$, where $\mathbf{x}$ is distributed according to a multivariate Gaussian of the form $\mathcal{N}(\mathbf{x}\,; \mathbf{m}_{\mathbf{x}}, \mathbf{P}_{\mathbf{x}})$ and $\boldsymbol{\xi}$ is dependent on $\mathbf{x}$. The depen-

dency of $\boldsymbol{\xi}$ on $\mathbf{x}$ can be represented by the following linear regression equation

$$\boldsymbol{\xi} = \mathbf{a} + \mathbf{B}\mathbf{x}, \tag{C.13}$$

where $\mathbf{B}$ is the regression coefficient and $\mathbf{a}$ is the intercept. The conditional distribution of $\boldsymbol{\xi}$ given $\mathbf{x}$ can also be represented as a multivariate Gaussian of the form

$$p(\boldsymbol{\xi} \,|\, \mathbf{x}) = \mathcal{N}(\boldsymbol{\xi}\,; \mathbf{a} + \mathbf{B}\mathbf{x}, \mathbf{P}_{\boldsymbol{\xi}|\mathbf{x}}). \tag{C.14}$$

First note that $\mathsf{E}(\boldsymbol{\xi} \,|\, \mathbf{x}) = \mathbf{a} + \mathbf{B}\mathbf{x}$. Consequently,

$$\mathsf{E}\left(\mathsf{E}(\boldsymbol{\xi} \,|\, \mathbf{x})\right) = \mathsf{E}\left(\mathbf{a} + \mathbf{B}\mathbf{x}\right) = \mathbf{a} + \mathbf{B}\mathsf{E}(\mathbf{x}). \tag{C.15}$$

Trivially $\mathsf{E}\left(\mathsf{E}(\boldsymbol{\xi} \,|\, \mathbf{x})\right) = \mathsf{E}(\boldsymbol{\xi})$, and so, given $\mathsf{E}(\mathbf{x}) = \mathbf{m}_{\mathbf{x}}$ and denoting $\mathsf{E}(\boldsymbol{\xi}) = \mathbf{m}_{\boldsymbol{\xi}}$, the expression in equation (C.15) becomes $\mathbf{m}_{\boldsymbol{\xi}} = \mathbf{a} + \mathbf{B}\mathbf{m}_{\mathbf{x}}$, from which an expression for $\mathbf{a}$ can be determined. By substitution of this expression for $\mathbf{a}$ into (C.14), the conditional multivariate Gaussian is re-written as

$$p(\boldsymbol{\xi} \,|\, \mathbf{x}) = \mathcal{N}(\boldsymbol{\xi}\,; \mathbf{m}_{\boldsymbol{\xi}} - \mathbf{B}\mathbf{m}_{\mathbf{x}} + \mathbf{B}\mathbf{x}, \mathbf{P}_{\boldsymbol{\xi}|\mathbf{x}}) = \mathcal{N}(\boldsymbol{\xi}\,; \mathbf{m}_{\boldsymbol{\xi}} + \mathbf{B}(\mathbf{x} - \mathbf{m}_{\mathbf{x}}), \mathbf{P}_{\boldsymbol{\xi}|\mathbf{x}}). \tag{C.16}$$

It remains to determine the formulae for $\mathbf{B}$ and $\mathbf{P}_{\boldsymbol{\xi}|\mathbf{x}}$. This is achieved by forming the joint density $p(\boldsymbol{\xi}, \mathbf{x}) = p(\mathbf{x})p(\boldsymbol{\xi} \,|\, \mathbf{x})$ as follows

$$\begin{aligned} p(\boldsymbol{\xi}, \mathbf{x}) &= \mathcal{N}(\mathbf{x}\,; \mathbf{m}_{\mathbf{x}}, \mathbf{P}_{\mathbf{x}}) \times \mathcal{N}(\boldsymbol{\xi}\,; \mathbf{m}_{\boldsymbol{\xi}} + \mathbf{B}(\mathbf{x} - \mathbf{m}_{\mathbf{x}}), \mathbf{P}_{\boldsymbol{\xi}|\mathbf{x}}) \\ &\propto \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{m}_{\mathbf{x}})^T \mathbf{R}_{\mathbf{x}} (\mathbf{x} - \mathbf{m}_{\mathbf{x}})\right) \\ &\quad \times \exp\left(-\frac{1}{2}(\boldsymbol{\xi} - \mathbf{m}_{\boldsymbol{\xi}} - \mathbf{B}(\mathbf{x} - \mathbf{m}_{\mathbf{x}}))^T \mathbf{R}_{\boldsymbol{\xi}|\mathbf{x}} (\boldsymbol{\xi} - \mathbf{m}_{\boldsymbol{\xi}} - \mathbf{B}(\mathbf{x} - \mathbf{m}_{\mathbf{x}}))\right), \end{aligned} \tag{C.17}$$

where $\mathbf{R}_{\mathbf{x}} = \mathbf{P}_{\mathbf{x}}^{-1}$ and $\mathbf{R}_{\boldsymbol{\xi}|\mathbf{x}} = \mathbf{P}_{\boldsymbol{\xi}|\mathbf{x}}^{-1}$. The density $p(\boldsymbol{\xi}, \mathbf{x})$ can be written in the joint multivariate Gaussian format given in (C.3) as follows

$$\begin{aligned} p(\boldsymbol{\xi} \,|\, \mathbf{x}) &\propto \exp\left(-\frac{1}{2}\begin{bmatrix} \boldsymbol{\xi} - \mathbf{m}_{\boldsymbol{\xi}} - \mathbf{B}(\mathbf{x} - \mathbf{m}_{\mathbf{x}}) \\ \mathbf{x} - \mathbf{m}_{\mathbf{x}} \end{bmatrix}^T \begin{bmatrix} \mathbf{R}_{\boldsymbol{\xi}|\mathbf{x}} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{\mathbf{x}} \end{bmatrix} \begin{bmatrix} \boldsymbol{\xi} - \mathbf{m}_{\boldsymbol{\xi}} - \mathbf{B}(\mathbf{x} - \mathbf{m}_{\mathbf{x}}) \\ \mathbf{x} - \mathbf{m}_{\mathbf{x}} \end{bmatrix}\right) \\ &= \exp\left(-\frac{1}{2}\begin{bmatrix} \boldsymbol{\xi} - \mathbf{m}_{\boldsymbol{\xi}} \\ \mathbf{x} - \mathbf{m}_{\mathbf{x}} \end{bmatrix}^T \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{B}^T & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{R}_{\boldsymbol{\xi}|\mathbf{x}} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{\mathbf{x}} \end{bmatrix} \begin{bmatrix} \mathbf{I} & -\mathbf{B} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{\xi} - \mathbf{m}_{\boldsymbol{\xi}} \\ \mathbf{x} - \mathbf{m}_{\mathbf{x}} \end{bmatrix}\right), \end{aligned} \tag{C.18}$$

where the block matrix Σ given in (C.2) equates to

$$\left(\begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{B}^T & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{R}_{\boldsymbol{\xi}|\mathbf{x}} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{\mathbf{x}} \end{bmatrix} \begin{bmatrix} \mathbf{I} & -\mathbf{B} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}\right)^{-1} = \begin{bmatrix} \mathbf{I} & \mathbf{B} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{R}_{\boldsymbol{\xi}|\mathbf{x}}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{\mathbf{x}}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{B}^T & \mathbf{I} \end{bmatrix}. \tag{C.19}$$

That is, since $\mathbf{R}_{\mathbf{x}}^{-1} = \mathbf{P}_{\mathbf{x}}$ & $\mathbf{R}_{\boldsymbol{\xi}|\mathbf{x}}^{-1} = \mathbf{P}_{\boldsymbol{\xi}|\mathbf{x}}$,

$$\begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{12}^T & \mathbf{P}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{P}_{\boldsymbol{\xi}|\mathbf{x}} + \mathbf{B}\mathbf{P}_{\mathbf{x}}\mathbf{B}^T & \mathbf{B}\mathbf{P}_{\mathbf{x}} \\ \mathbf{P}_{\mathbf{x}}\mathbf{B}^T & \mathbf{P}_{\mathbf{x}} \end{bmatrix}. \tag{C.20}$$

Finally, this gives $\mathbf{P}_{\mathbf{x}} = \mathbf{P}_{22}$, $\mathbf{B} = \mathbf{P}_{12}\mathbf{P}_{22}^{-1}$ and $\mathbf{P}_{\boldsymbol{\xi}|\mathbf{x}} = \mathbf{P}_{11} - \mathbf{P}_{12}\mathbf{P}_{22}^{-1}\mathbf{P}_{12}^T$, so that

$$p(\boldsymbol{\xi} \,|\, \mathbf{x}) = \mathcal{N}\big(\boldsymbol{\xi}\,; \mathbf{m}_{\boldsymbol{\xi}} + \mathbf{P}_{12}\mathbf{P}_{22}^{-1}(\mathbf{x} - \mathbf{m}_{\mathbf{x}}), \mathbf{P}_{11} - \mathbf{P}_{12}\mathbf{P}_{22}^{-1}\mathbf{P}_{12}^T\big)\,, \tag{C.21}$$

as desired.

C.2 Multivariate Gaussian product identities

The following standard results for Gaussian functions[1], found in [90], are essential for establishing the Gaussian mixture formulations of the multi-group PHD filter recursion in Chapter 4 (Section 4.1).

Lemma C.1. *Given matrices* $\mathbf{F}$, $\mathbf{Q}$ *and* $\mathbf{P}$ *and vectors* $\mathbf{m}$ *and* $\mathbf{d}$ *of appropriate dimensions, where* $\mathbf{Q}$ *and* $\mathbf{P}$ *are positive definite,*

$$\int \mathcal{N}(\mathbf{X}\,; \mathbf{F}\mathbf{X}' + \mathbf{d}, \mathbf{Q})\, \mathcal{N}(\mathbf{X}'\,; \mathbf{m}, \mathbf{P})\, \mathrm{d}\mathbf{X}' = \mathcal{N}(\mathbf{X}\,; \mathbf{F}\mathbf{m} + \mathbf{d}, \mathbf{F}\mathbf{P}\mathbf{F}^T + \mathbf{Q}). \tag{C.22}$$

Lemma C.2. *Given matrices* $\mathbf{H}$, $\mathbf{R}$ *and* $\mathbf{P}$ *and vector* $\mathbf{m}$ *of appropriate dimensions, where* $\mathbf{R}$ *and* $\mathbf{P}$ *are positive definite,*

$$\mathcal{N}(\mathbf{z}\,; \mathbf{H}\mathbf{X}, \mathbf{R})\, \mathcal{N}(\mathbf{X}\,; \mathbf{m}, \mathbf{P}) = q(\mathbf{z})\, \mathcal{N}(\mathbf{X}\,; \tilde{\mathbf{m}}, \tilde{\mathbf{P}}), \tag{C.23}$$

where $q(\mathbf{z}) = \mathcal{N}(\mathbf{z}\,; \mathbf{H}\mathbf{m}, \mathbf{H}\mathbf{P}\mathbf{H}^T + \mathbf{R})$ *and*

$$\begin{aligned} \tilde{\mathbf{m}} &= \mathbf{m} + \mathbf{K}(\mathbf{z} - \mathbf{H}\mathbf{m}), \\ \tilde{\mathbf{P}} &= (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}, \\ \mathbf{K} &= \mathbf{P}\mathbf{H}^T(\mathbf{H}\mathbf{P}\mathbf{H}^T + \mathbf{R})^{-1}. \end{aligned} \tag{C.24}$$

[1] For a list of useful multivariate Gaussian identities, refer to [140] by Roweis

C.2.1 Formulating the multi-group GM-PHD prediction

To establish the Gaussian mixture (GM) formulation given in Proposition 4.1, first substitute the conditional predicted intensity given by equation (3.18) into equation (3.17) and rewrite as

$$\begin{aligned} v_{k|k-1}(\mathbf{x}_k, \boldsymbol{\xi}_k) = \gamma_k(\mathbf{x}_k, \boldsymbol{\xi}_k) &+ p_{S,1} \int f_{k|k-1}(\mathbf{x}_k \,|\, \mathbf{x}_{k-1})\, v_{k-1}(\mathbf{x}_{k-1})\, \gamma_k(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k)\, \mathrm{d}\mathbf{x}_{k-1} \\ &+ p_{S,1}\, p_{S,2} \int f_{k|k-1}(\mathbf{x}_k, \boldsymbol{\xi}_k \,|\, \mathbf{x}_{k-1}, \boldsymbol{\xi}_{k-1})\, v_{k-1}(\mathbf{x}_{k-1}, \boldsymbol{\xi}_{k-1})\, \mathrm{d}\mathbf{x}_{k-1}\, \mathrm{d}\boldsymbol{\xi}_{k-1}, \end{aligned} \tag{C.25}$$

noting that the probabilities of survival are constant (see Section 4.1.2). Then substituting the GM formulations given in (4.1), (4.3), (4.10) and (4.11), as well as the linear Gaussian models for the joint and single-group Markov transition densities given in (4.6) and (4.8) respectively, into equation (C.25) gives the desired result after applying the standard results for Gaussian functions given in Lemma C.1 and C.2 where appropriate.

The resulting GM formulation given in equation (4.16) satisfies the following factorisation $v_{k|k-1}(\mathbf{x}_k, \boldsymbol{\xi}_k) = v_{k|k-1}(\mathbf{x}_k) \times v_{k|k-1}(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k)$ so that

$$v_{k|k-1}(\mathbf{x}_k) = \sum_{\bar{\imath}=1}^{J_{k|k-1}} \nu_{k|k-1}^{(\bar{\imath})}\, \mathcal{N}\Big(\mathbf{x}_k\,; \mathbf{m}_{\mathbf{x},k|k-1}^{(\bar{\imath})}, \mathbf{P}_{22,k|k-1}^{(\bar{\imath})}\Big), \tag{C.26}$$

$$v_{k|k-1}(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k) = \sum_{\bar{\jmath}=1}^{J_{k|k-1}^{(\bar{\imath})}} \omega_{k|k-1}^{(\bar{\jmath})}\, \mathcal{N}\Big(\boldsymbol{\xi}_k\,; \boldsymbol{\mu}_{\boldsymbol{\xi}|\mathbf{x},k|k-1}^{(\bar{\imath},\bar{\jmath})}, \boldsymbol{\Sigma}_{\boldsymbol{\xi}|\mathbf{x},k|k-1}^{(\bar{\imath},\bar{\jmath})}\Big), \tag{C.27}$$

where the Gaussian components in (C.27) are conditional Gaussian realisations from the joint Gaussian components in (4.16), with means and covariances given by

$$\begin{aligned} \boldsymbol{\mu}_{\boldsymbol{\xi}|\mathbf{x},k|k-1}^{(\bar{\imath},\bar{\jmath})} &= \mathbf{m}_{\boldsymbol{\xi},k|k-1}^{(\bar{\jmath}|\bar{\imath})} + \mathbf{P}_{12,k|k-1}^{(\bar{\jmath}|\bar{\imath})} \Big(\mathbf{P}_{22,k|k-1}^{(\bar{\imath})}\Big)^{-1} \Big(\mathbf{x}_k - \mathbf{m}_{\mathbf{x},k|k-1}^{(\bar{\imath})}\Big), \\ \boldsymbol{\Sigma}_{\boldsymbol{\xi}|\mathbf{x},k|k-1}^{(\bar{\imath},\bar{\jmath})} &= \mathbf{P}_{11,k|k-1}^{(\bar{\jmath}|\bar{\imath})} - \mathbf{P}_{12,k|k-1}^{(\bar{\jmath}|\bar{\imath})} \Big(\mathbf{P}_{22,k|k-1}^{(\bar{\imath})}\Big)^{-1} \Big(\mathbf{P}_{12,k|k-1}^{(\bar{\jmath}|\bar{\imath})}\Big)^{T}. \end{aligned} \tag{C.28}$$

C.2.2 Formulating the multi-group GM-PHD update

To establish the GM formulation given in Proposition 4.2, first note that the assumption regarding the state independent probability of individual target detection (see Section 4.1.4) implies that the exponential term in equation (3.21) and the pseudo multiple measurement likelihood can be written as $\exp\big(-v_{k|k-1}[p_{D,2} \,|\, \cdot]\big) =$

$\mathrm{e}^{-p_{D,2}\,\alpha}$ where

$$\alpha = \sum_{j=1}^{J_{k|k-1}^{(i)}} \omega_{k|k-1}^{(j)}, \tag{C.29}$$

denotes the expected number of individual targets for each group. Then the updated intensity given by equation (3.21) can be re-written as

$$\begin{aligned} v_k(\mathbf{x}_k, \boldsymbol{\xi}_k) = &\left(1 - p_{D,1} + p_{D,1}\,\mathrm{e}^{-p_{D,2}\,\alpha}\,(1 - p_{D,2})\right) v_{k|k-1}(\mathbf{x}_k, \boldsymbol{\xi}_k) \\ &+ \sum_{\pi \in \Pi_Z} \varpi_\pi \sum_{\varphi \in \pi} \frac{p_{D,1}\, L_\varphi(\mathbf{x}_k)\, v_{k|k-1}(\mathbf{x}_k)\, v_k(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k)}{\Lambda(\varphi) + p_{D,1} \int L_\varphi(\mathbf{x}_k)\, v_{k|k-1}(\mathbf{x}_k)\, \mathrm{d}\mathbf{x}_k}, \end{aligned} \tag{C.30}$$

and the conditional updated intensity given by equation (3.22) can be re-written as

$$v_k(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k) = \left(1 - p_{D,2} + \sum_{\mathbf{z} \in \varphi} \frac{p_{D,2}\, g_k(\mathbf{z} \,|\, \boldsymbol{\xi}_k, \mathbf{x}_k)}{p_{D,2} \int g_k(\mathbf{z} \,|\, \boldsymbol{\xi}_k \,|\, \mathbf{x}_k)\, v_{k|k-1}(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k)\, \mathrm{d}\boldsymbol{\xi}_k}\right) v_{k|k-1}(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k), \tag{C.31}$$

noting that the probabilities of detection are constant

Furthermore, the pseudo multiple measurement likelihood, denoted by $L_\varphi(\mathbf{x}_k)$ in equation (C.30) and given by equation (3.24), can be re-written as

$$L_\varphi(\mathbf{x}_k) = \mathrm{e}^{-p_{D,2}\,\alpha} \prod_{\mathbf{z} \in \varphi} p_{D,2} \int g_k(\mathbf{z} \,|\, \boldsymbol{\xi}_k, \mathbf{x}_k)\, v_{k|k-1}(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k)\, \mathrm{d}\boldsymbol{\xi}_k. \tag{C.32}$$

Substituting the GM formulation for $v_{k|k-1}(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k)$ given by equation (C.27) and the linear Gaussian model for the single measurement likelihood given by equation (4.15) into the above expression for L_φ, then applying the standard result for Gaussian functions given in Lemma C.1 gives

$$L_\varphi(\mathbf{x}_k) = \mathrm{e}^{-p_{D,2}\,\alpha} \prod_{\mathbf{z} \in \varphi} \left(p_{D,2}\, \mathcal{N}(\mathbf{z}\,; \widehat{\mathbf{H}}\mathbf{x}_k, \mathbf{R}_2) \sum_{j=1}^{J_{k|k-1}^{(i)}} \omega_{k|k-1}^{(j)}\, \mathcal{N}\left(\mathbf{z}\,; \mathbf{H}\boldsymbol{\mu}_{\boldsymbol{\xi}|\mathbf{x},k|k-1}^{(i,j)}, \widetilde{\mathbf{R}}_k^{(i,j)}\right)\right), \tag{C.33}$$

where, for brevity, $\widetilde{\mathbf{R}}_k^{(i,j)} = \mathbf{H}\boldsymbol{\Sigma}_{\boldsymbol{\xi}|\mathbf{x},k|k-1}^{(i,j)}\mathbf{H}^T + \mathbf{R}_1$, with $\boldsymbol{\mu}_{\boldsymbol{\xi}|\mathbf{x},k|k-1}^{(i,j)}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\xi}|\mathbf{x},k|k-1}^{(i,j)}$ as given by the expressions in (C.28). The product of Gaussian mixtures in equation (C.33) can be expanded and re-written as a sum of Gaussian products for $j_\ell = 1, \ldots, J_{k|k-1}^{(i)}$ and $\ell = 1, \ldots, |\varphi|$ so that the pseudo multiple measurement likelihood

becomes

$$L_{\varphi}(\mathbf{x}_k) = \mathrm{e}^{-p_{D,2}\,\alpha} \sum_{j_{1:|\varphi|}} \left(\prod_{\ell=1}^{|\varphi|} p_{D,2}\, \omega_{k|k-1}^{(j_\ell)}\, \mathcal{N}(\mathbf{z}_\ell\,; \widehat{\mathbf{H}}\mathbf{x}_k, \mathbf{R}_2)\, \mathcal{N}\Big(\mathbf{z}_\ell\,; \mathbf{H}\boldsymbol{\mu}_{\boldsymbol{\xi}|\mathbf{x},k|k-1}^{(i,j_\ell)}, \widetilde{\mathbf{R}}_k^{(i,j_\ell)}\Big) \right), \tag{C.34}$$

where $j_{1:|\varphi|} = (j_1, \ldots, j_{|\varphi|})$ denotes the permutations of indices $j_\ell = 1, \ldots, J_{k|k-1}^{(i)}$ for $\ell = 1, \ldots, |\varphi|$. Then substituting the above expression for $L_{\varphi}(\mathbf{x}_k)$, along with the GM formulation for $v_{k|k-1}(\mathbf{x}_k)$ given in equation (C.26), into equation (C.30) and iteratively applying the standard result for Gaussian functions in Lemma C.2 gives the desired results for the updated weights, means and covariances of the Gaussian components in equation (4.24) relating to detected group states as given by the set of equations 4.26–4.32.

To establish the results for the updated weights, means and covariances of the Gaussian components in equation (4.24) relating to individual targets, first substitute the GM formulation for $v_{k|k-1}(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k)$ given in equation (C.27), along with the linear Gaussian model for $g_k(\mathbf{z} \,|\, \boldsymbol{\xi}_k, \mathbf{x}_k)$ given in equation (4.15), into equation (C.31) and apply the standard result for Gaussian functions in Lemma C.2, resulting in the expressions for the means and covariances given in (4.37). Some algebraic manipulation is necessary when substituting the resulting Gaussian mixture into equation (C.30) so that the updated intensity can be formulated as the Gaussian mixture given in equation (4.24), with updated means and covariances relating to detected individual targets given in (4.35). Likewise with the updated means and covariances relating to missed detected individual targets given in (4.33). Finally, the updated weights of the Gaussian components relating to detected individual targets given in equation (4.38) are determined directly from the multiple measurement likelihood $L_{j_{1:|\varphi|}}^{(i)}$, which is computed as given by the expression in equation (4.32). Similarly for the partition weights ϖ_π given by the expression in equation (4.39).

Appendix D

The Single Group GM-PHD Filter Algorithm

This chapter presents full details on the closed-form solution to the single group PHD filter recursion. In particular, pseudo codes are presented in Section D.1 for the single group GM-PHD prediction and update step, which provide closed-form expressions for computing the weights, means and covariances in the respective Gaussian mixture formulations, summarised in Section 4.3.1.

D.1 Pseudo Codes for the Single Group GM-PHD Recursion

The algorithm for the computation of the weights, means and covariances in the GM formulation of the single group PHD prediction is shown in Table D.1. The

---

given $\mathbf{m}_{\mathbf{x},k-1}$, $\mathbf{P}_{22,k-1}$ and $\left\{\omega_{k-1}^{(i)}, \mathbf{m}_{\boldsymbol{\xi},k-1}^{(i)}, \mathbf{P}_{11,k-1}^{(i)}, \mathbf{P}_{12,k-1}^{(i)}\right\}_{i=1}^{J_{k-1}}$,

prediction

$\mathbf{m}_{\mathbf{x},k|k-1} := \mathbf{F}_{22}\mathbf{m}_{\mathbf{x},k-1}$,

$\mathbf{P}_{22,k|k-1} := \mathbf{F}_{22}\mathbf{P}_{22,k-1}\mathbf{F}_{22}^T + \mathbf{Q}_{22}$, (*for the single group target*)

$\bar{\imath} := 0$,

for $i = 1, \ldots, J_{\gamma,k}$ (*i.e. for individual target births*)

$\bar{\imath} := \bar{\imath} + 1$,

$\omega_{k|k-1}^{(\bar{\imath})} := \omega_{\gamma,k}^{(i)}$, $\mathbf{m}_{\boldsymbol{\xi},k|k-1}^{(\bar{\imath})} := \mathbf{m}_{\gamma,\boldsymbol{\xi},k}^{(i)}$,

$\mathbf{P}_{11,k|k-1}^{(\bar{\imath})} := \mathbf{P}_{\gamma,11,k}^{(i)}$, $\mathbf{P}_{12,k|k-1}^{(\bar{\imath})} := \mathbf{P}_{\gamma,12,k}^{(i)}$.

end

for $i = 1, \ldots, J_{k|k-1}$ (*i.e. for existing individual targets*)

$\bar{\imath} := \bar{\imath} + 1$,

$\omega_{k|k-1}^{(\bar{\imath})} := p_S\, \omega_{k-1}^{(i)}$, $\mathbf{m}_{\boldsymbol{\xi},k|k-1}^{(\bar{\imath})} := \mathbf{F}_{11}\mathbf{m}_{\boldsymbol{\xi},k-1}^{(i)} + \mathbf{F}_{12}\mathbf{m}_{\mathbf{x},k-1}^{(i)}$,

$\mathbf{P}_{11,k|k-1}^{(\bar{\imath})} := \mathbf{F}_{11}\mathbf{P}_{11,k-1}^{(i)}\mathbf{F}_{11}^T + \left(\mathbf{F}_{11}\mathbf{P}_{12,k-1}^{(i)}\mathbf{F}_{12}^T\right)^T + \mathbf{F}_{11}\mathbf{P}_{12,k-1}^{(i)}\mathbf{F}_{12}^T + \mathbf{F}_{12}\mathbf{P}_{22,k-1}\mathbf{F}_{12}^T + \mathbf{Q}_{11}$,

$\mathbf{P}_{12,k|k-1}^{(\bar{\imath})} := \mathbf{F}_{11}\mathbf{P}_{12,k-1}^{(i)}\mathbf{F}_{22}^T + \mathbf{F}_{12}\mathbf{P}_{22,k-1}\mathbf{F}_{22}^T + \mathbf{Q}_{12}$,

end

$J_{k|k-1} := \bar{\imath}$.

output $\mathbf{m}_{\mathbf{x},k|k-1}$, $\mathbf{P}_{22,k|k-1}$ and $\left\{\omega_{k|k-1}^{(i)}, \mathbf{m}_{\boldsymbol{\xi},k|k-1}^{(i)}, \mathbf{P}_{11,k|k-1}^{(i)}, \mathbf{P}_{12,k|k-1}^{(i)}\right\}_{i=1}^{J_{k|k-1}}$.

Table D.1: Pseudo-code for the single group GM-PHD prediction

key steps of the single group GM-PHD update are then shown in Tables D.2–D.5.

The closed-form expressions for the predicted weights, means and covariances in Table D.1 follow directly from the multi-group GM-PHD prediction given in Proposition 4.1 (Section 4.1.5). Likewise, the closed-form expressions for the updated means and covariances in Tables D.2 and D.3 follow directly from the multi-group GM-PHD update given in Proposition 4.2 (Section 4.1.6).

However, the same can not be said for the updated weights due to the difference in multiple measurement likelihoods present in the update equations for the respective PHD filters. The multiple measurement likelihood $L_{Z_k}(\mathbf{x}_k)$ contributing to the update of the single group state in equation (3.27) can be written as

$$L_{Z_k}(\mathbf{x}_k) = \mathrm{e}^{-p_D\,\alpha}\left(\prod_{\mathbf{z}\in Z_k}\kappa(\mathbf{z})\right)\sum_{W\subseteq Z_k}\left(\prod_{\mathbf{z}\in W}\frac{p_D}{\kappa(\mathbf{z})}\int g_k(\mathbf{z}\,|\,\boldsymbol{\xi}_k,\mathbf{z}_k)\,v_{k|k-1}(\boldsymbol{\xi}_k\,|\,\mathbf{x}_k)\mathrm{d}\boldsymbol{\xi}_k\right), \tag{D.1}$$

which follows as a result of the expression given in equation (3.43) and the assumption regarding the state independent probability of individual target detection in Section 4.1.4, now denoted by $p_D = p_{D,2}$. The clutter intensity is denoted by $\kappa(\mathbf{z})$ and $\alpha = \sum_{i=1}^{J_{k|k-1}}\omega_{k|k-1}^{(i)}$ denotes the expected number of individual targets.

It follows that, given the GM formulation of the predicted intensity which satisfies the factorisation $v_{k|k-1}(\mathbf{x}_k,\boldsymbol{\xi}_k) = p_{k|k-1}(\mathbf{x}_k)\times v_{k|k-1}(\boldsymbol{\xi}_k\,|\,\mathbf{x}_k)$, where $p_{k|k-1}(\mathbf{x}_k) = \mathcal{N}\left(\mathbf{x}_k\,;\mathbf{m}_{\mathbf{x},k|k-1},\mathbf{P}_{22,k|k-1}\right)$ and

$$v_{k|k-1}(\boldsymbol{\xi}_k\,|\,\mathbf{x}_k) = \sum_{i=1}^{J_{k|k-1}}\omega_{k|k-1}^{(i)}\,\mathcal{N}\Big(\boldsymbol{\xi}_k\,;\boldsymbol{\mu}_{\boldsymbol{\xi}|\mathbf{x},k|k-1}^{(i)},\boldsymbol{\Sigma}_{\boldsymbol{\xi}|\mathbf{x},k|k-1}^{(i)}\Big), \tag{D.2}$$

with means $\boldsymbol{\mu}_{\boldsymbol{\xi}|\mathbf{x},k|k-1}^{(i)} = \mathbf{m}_{\boldsymbol{\xi},k|k-1}^{(i)} + \mathbf{P}_{12,k|k-1}^{(i)}\mathbf{P}_{22,k|k-1}^{-1}\left(\mathbf{x}_k - \mathbf{m}_{\mathbf{x},k|k-1}\right)$ and covariances given by

$$\boldsymbol{\Sigma}_{\boldsymbol{\xi}|\mathbf{x},k|k-1}^{(i)} = \mathbf{P}_{11,k|k-1}^{(i)} - \mathbf{P}_{12,k|k-1}^{(i)}\mathbf{P}_{22,k|k-1}^{-1}\left(\mathbf{P}_{12,k|k-1}^{(i)}\right)^T, \tag{D.3}$$

the multiple measurement likelihood given in equation (D.1) becomes

$$\begin{aligned} L_{Z_k}(\mathbf{x}_k) = \mathrm{e}^{-p_D\,\alpha}\left(\prod_{\mathbf{z}\in Z_k}\kappa(\mathbf{z})\right)\sum_{W\subseteq Z_k}\Bigg(&\prod_{\mathbf{z}\in W}\frac{p_D}{\kappa(\mathbf{z})}\mathcal{N}\Big(\mathbf{z}\,;\widehat{\mathbf{H}}\mathbf{x}_k,\mathbf{R}_2\Big)\times \\ &\sum_{i=1}^{J_{k|k-1}}\omega_{k|k-1}^{(i)}\,\mathcal{N}\Big(\mathbf{z}\,;\mathbf{H}\boldsymbol{\mu}_{\boldsymbol{\xi}|\mathbf{x},k|k-1}^{(i)},\mathbf{H}\boldsymbol{\Sigma}_{\boldsymbol{\xi}|\mathbf{x},k|k-1}^{(i)}\mathbf{H}^T+\mathbf{R}_1\Big)\Bigg). \end{aligned} \tag{D.4}$$

given $\mathbf{m}_{\mathbf{x},k|k-1}$, $\mathbf{P}_{22,k|k-1}$, $\left\{\omega_{k|k-1}^{(i)}, \mathbf{m}_{\boldsymbol{\xi},k|k-1}^{(i)}, \mathbf{P}_{11,k|k-1}^{(i)}, \mathbf{P}_{12,k|k-1}^{(i)}\right\}_{i=1}^{J_{k|k-1}}$,
measurement set Z_k and constant probability of detection p_D.

step 1. (*Construction of GM-PHD update components*)

for $i = 1, \ldots, J_{k|k-1}$

$\widetilde{\mathbf{R}}_k^{(i)} := \mathbf{H}\boldsymbol{\Sigma}_{\boldsymbol{\xi}|\mathbf{x},k|k-1}^{(i)}\mathbf{H}^T + \mathbf{R}_1, \quad \mathbf{A}_k^{(i)} := \mathbf{H}\mathbf{P}_{12,k|k-1}^{(i)}\mathbf{P}_{22,k|k-1}^{-1}$,
where $\boldsymbol{\Sigma}_{\boldsymbol{\xi}|\mathbf{x},k|k-1}^{(i)}$ is given by the expression in equation (D.3).

end

step 2. (*Update for missed detections*)

$n := 0, \quad \bar{\imath} := 0.$

for each subset $W \subseteq Z_k$ (*Update of the parent process*)

for each permutation $i_{1:|W|} = (i_1, \ldots, i_{|W|})$
(*where* $i_\ell = 1, \ldots, J_{k|k-1}$ *for* $\ell = 1, \ldots, |W|$)

$n := n + 1,$

$\boldsymbol{\eta}_k^{(i_1)} := \mathbf{m}_{\mathbf{x},k|k-1}, \quad \boldsymbol{\mathcal{P}}_k^{(i_1)} := \mathbf{P}_{22,k|k-1}.$

for $\ell = 1, \ldots, |W|$

$\mathbf{b}_1^{(i_{1:\ell})} := \mathbf{H}\mathbf{m}_{\boldsymbol{\xi},k|k-1}^{(i_\ell)} + \mathbf{A}_k^{(i_\ell)}\left(\boldsymbol{\eta}_k^{(i_{1:\ell})} - \mathbf{m}_{\mathbf{x},k|k-1}\right),$

$\mathbf{S}_1^{(i_{1:\ell})} := \mathbf{A}_k^{(i_\ell)}\boldsymbol{\mathcal{P}}_k^{(i_{1:\ell})}\left(\mathbf{A}_k^{(i_\ell)}\right)^T + \widetilde{\mathbf{R}}_k^{(i_\ell)},$

(*where* $\boldsymbol{\eta}_k^{(i_{1:\ell})} = \boldsymbol{\eta}_k^{(i_1)}$ and $\boldsymbol{\mathcal{P}}_k^{(i_{1:\ell})} = \boldsymbol{\mathcal{P}}_k^{(i_1)}$ *for* $\ell = 1$)

$\widetilde{\mathbf{K}}_{\mathbf{x},k}^{(i_{1:\ell})} := \boldsymbol{\mathcal{P}}_k^{(i_{1:\ell})}\left(\mathbf{A}_k^{(i_\ell)}\right)^T\left(\mathbf{S}_1^{(i_{1:\ell})}\right)^{-1},$

$\widetilde{\mathbf{m}}_{\mathbf{x},k}^{(i_{1:\ell})} := \boldsymbol{\eta}_k^{(i_{1:\ell})} + \widetilde{\mathbf{K}}_{\mathbf{x},k}^{(i_{1:\ell})}\left(\mathbf{z}_\ell - \mathbf{b}_1^{(i_{1:\ell})}\right),$

$\widetilde{\mathbf{P}}_{22,k}^{(i_{1:\ell})} := \left(\mathbf{I} - \widetilde{\mathbf{K}}_{\mathbf{x},k}^{(i_{1:\ell})}\mathbf{A}_k^{(i_\ell)}\right)\boldsymbol{\mathcal{P}}_k^{(i_{1:\ell})},$

$\mathbf{S}_2^{(i_{1:\ell})} := \widehat{\mathbf{H}}\widetilde{\mathbf{P}}_{22,k}^{(i_{1:\ell})}\widehat{\mathbf{H}}^T + \mathbf{R}_2, \quad \mathbf{K}_{\mathbf{x},k}^{(i_{1:\ell})} := \widetilde{\mathbf{P}}_{22,k}^{(i_{1:\ell})}\widehat{\mathbf{H}}^T\left(\mathbf{S}_2^{(i_{1:\ell})}\right)^{-1},$

$\boldsymbol{\eta}_k^{(i_{1:\ell})} := \widetilde{\mathbf{m}}_{\mathbf{x},k}^{(i_{1:\ell})} + \mathbf{K}_{\mathbf{x},k}^{(i_{1:\ell})}\left(\mathbf{z}_\ell - \widehat{\mathbf{H}}\widetilde{\mathbf{m}}_{\mathbf{x},k}^{(i_{1:\ell})}\right),$

$\boldsymbol{\mathcal{P}}_k^{(i_{1:\ell})} := \left(\mathbf{I} - \mathbf{K}_{\mathbf{x},k}^{(i_{1:\ell})}\widehat{\mathbf{H}}\right)\widetilde{\mathbf{P}}_{22,k}^{(i_{1:\ell})},$

$L_{i_\ell}^{(n)} := p_D\,\omega_{k|k-1}^{(i_\ell)}\mathcal{N}\left(\mathbf{z}_\ell\,;\mathbf{b}_1^{(i_{1:\ell})}, \mathbf{S}_1^{(i_{1:\ell})}\right)\mathcal{N}\left(\mathbf{z}_\ell\,;\widehat{\mathbf{H}}\widetilde{\mathbf{m}}_{\mathbf{x},k}^{(i_{1:\ell})}, \mathbf{S}_2^{(i_{1:\ell})}\right).$

end

$\mathbf{m}_{\mathbf{x},k}^{(n)} := \boldsymbol{\eta}_k^{(i_{1:|W|})}, \quad \mathbf{P}_{22,k}^{(n)} := \boldsymbol{\mathcal{P}}_k^{(i_{1:|W|})}, \quad \tilde{\nu}_k^{(n)} := \prod_{\ell=1}^{|W|}\left(L_{i_\ell}^{(n)} \Big/ \kappa(\mathbf{z}_\ell)\right).$

end

end

for $j = 1, \ldots, n$

$\nu_k^{(j)} := \tilde{\nu}_k^{(j)} \Big/ \sum_{j'=1}^{n}\tilde{\nu}_k^{(j')}.$

end

for $i = 1, \ldots, J_{k|k-1}$ (*Update of the daughter process*)

$\bar{\imath} := \bar{\imath} + 1, \quad \omega_k^{(\bar{\imath})} := (1 - p_D)\,\omega_{k|k-1}^{(i)},$

$\widetilde{\mathbf{m}}_{\boldsymbol{\xi},k}^{(\bar{\imath})} := \mathbf{m}_{\boldsymbol{\xi},k|k-1}^{(i)}, \quad \widetilde{\mathbf{P}}_{11,k}^{(\bar{\imath})} := \mathbf{P}_{11,k|k-1}^{(i)}, \quad \widetilde{\mathbf{P}}_{12,k}^{(\bar{\imath})} := \mathbf{P}_{12,k|k-1}^{(i)}.$

end

$J_k := \bar{\imath}.$

output $\left\{\omega_k^{(i)}, \widetilde{\mathbf{m}}_{\boldsymbol{\xi},k}^{(i)}, \widetilde{\mathbf{P}}_{11,k}^{(i)}, \widetilde{\mathbf{P}}_{12,k}^{(i)}\right\}_{i=1}^{J_k}$, $\left\{\nu_k^{(j)}, \mathbf{m}_{\mathbf{x},k}^{(j)}, \mathbf{P}_{22,k}^{(j)}\right\}_{j=1}^{n}$ and $\left\{\widetilde{\mathbf{R}}_k^{(i)}, \mathbf{A}_k^{(i)}\right\}_{i=1}^{J_{k|k-1}}$.

Table D.2: Pseudo-code for steps 1 & 2 of the single group GM-PHD update

given $\mathbf{m}_{\mathbf{x},k|k-1}$, $\mathbf{P}_{22,k|k-1}$, $\left\{\omega_{k|k-1}^{(i)}, \mathbf{m}_{\boldsymbol{\xi},k|k-1}^{(i)}, \mathbf{P}_{11,k|k-1}^{(i)}, \mathbf{P}_{12,k|k-1}^{(i)}\right\}_{i=1}^{J_{k|k-1}}$, Z_k, p_D,
$\left\{\omega_k^{(i)}, \widetilde{\mathbf{m}}_{\boldsymbol{\xi},k}^{(i)}, \widetilde{\mathbf{P}}_{11,k}^{(i)}, \widetilde{\mathbf{P}}_{12,k}^{(i)}\right\}_{i=1}^{J_k}$, $\left\{\nu_k^{(j)}, \mathbf{m}_{\mathbf{x},k}^{(j)}, \mathbf{P}_{22,k}^{(j)}\right\}_{j=1}^{n}$ and $\left\{\widetilde{\mathbf{R}}_k^{(i)}, \mathbf{A}_k^{(i)}, \boldsymbol{\Sigma}_{\boldsymbol{\xi}|\mathbf{x},k|k-1}^{(i)}\right\}_{i=1}^{J_{k|k-1}}$.

step 3. (*Update for detections*)

$\bar{\imath} := J_k, \quad \bar{n} := n, \quad \bar{I} := \emptyset, \quad \bar{J} := \emptyset.$

for each $\mathbf{z} \in Z_k$

$\bar{I} := \{\bar{I}, \bar{\imath}+1\}.$

for $i = 1, \ldots, J_{k|k-1}$

$\bar{\imath} := \bar{\imath}+1, \quad \bar{J} := \{\bar{J}, \bar{n}+1\},$

$\mathbf{S}_1^{(i)} := \mathbf{A}_k^{(i)}\mathbf{P}_{22,k|k-1}\left(\mathbf{A}_k^{(i)}\right)^T + \widetilde{\mathbf{R}}_k^{(i)}, \quad \widetilde{\mathbf{K}}_{\mathbf{x},k}^{(i)} := \mathbf{P}_{22,k|k-1}\left(\mathbf{A}_k^{(i)}\right)^T\left(\mathbf{S}_1^{(i)}\right)^{-1},$

$\widetilde{\mathbf{m}}_{\mathbf{x},k}^{(i)} := \mathbf{m}_{\mathbf{x},k|k-1} + \widetilde{\mathbf{K}}_{\mathbf{x},k}^{(i)}\left(\mathbf{z} - \mathbf{H}\mathbf{m}_{\boldsymbol{\xi},k|k-1}^{(i)}\right), \quad \widetilde{\mathbf{P}}_{22,k}^{(i)} := \left(\mathbf{I} - \widetilde{\mathbf{K}}_{\mathbf{x},k}^{(i)}\mathbf{A}_k^{(i)}\right)\mathbf{P}_{22,k|k-1},$

$\mathbf{S}_2^{(i)} := \widehat{\mathbf{H}}\widetilde{\mathbf{P}}_{22,k}^{(i)}\widehat{\mathbf{H}}^T + \mathbf{R}_2, \quad \mathbf{K}_{\mathbf{x},k}^{(i)} := \widetilde{\mathbf{P}}_{22,k}^{(i)}\widehat{\mathbf{H}}^T\left(\mathbf{S}_2^{(i)}\right)^{-1},$

$\boldsymbol{\eta}_k^{(i)} := \widetilde{\mathbf{m}}_{\mathbf{x},k}^{(i)} + \mathbf{K}_{\mathbf{x},k}^{(i)}\left(\mathbf{z} - \widehat{\mathbf{H}}\widetilde{\mathbf{m}}_{\mathbf{x},k}^{(i)}\right), \quad \boldsymbol{\mathcal{P}}_k^{(i)} := \left(\mathbf{I} - \mathbf{K}_{\mathbf{x},k}^{(i)}\widehat{\mathbf{H}}\right)\widetilde{\mathbf{P}}_{22,k}^{(i)},$

$\tilde{\omega}_k^{(\bar{\imath})} := p_D\,\omega_{k|k-1}^{(i)}\mathcal{N}\left(\mathbf{z}; \mathbf{H}\mathbf{m}_{\boldsymbol{\xi},k}^{(i)}, \mathbf{S}_1^{(i)}\right)\mathcal{N}\left(\mathbf{z}; \widehat{\mathbf{H}}\widetilde{\mathbf{m}}_{\mathbf{x},k}^{(i)}, \mathbf{S}_2^{(i)}\right)\Big/\kappa(\mathbf{z}),$

$\widetilde{\mathbf{m}}_{\boldsymbol{\xi},k}^{(\bar{\imath})} := \mathbf{m}_{\boldsymbol{\xi},k|k-1}^{(i)} + \mathbf{K}_{\boldsymbol{\xi},k}^{(i)}\left(\mathbf{z} - \mathbf{H}\mathbf{m}_{\boldsymbol{\xi},k|k-1}^{(i)}\right),$

$\widetilde{\mathbf{P}}_{11,k}^{(\bar{\imath})} := \left(\mathbf{I} - \mathbf{K}_{\boldsymbol{\xi},k}^{(i)}\mathbf{H}\right)\left(\mathbf{P}_{11,k|k-1}^{(i)} - \mathbf{P}_{12,k|k-1}^{(i)}\mathbf{P}_{22,k|k-1}^{-1}\left(\mathbf{K}_{\boldsymbol{\xi},k}^{(i)}\mathbf{H}\mathbf{P}_{12,k|k-1}^{(i)}\right)^T\right),$

$\mathbf{K}_{\boldsymbol{\xi},k}^{(i)} = \boldsymbol{\Sigma}_{\boldsymbol{\xi}|\mathbf{x},k|k-1}^{(i)}\mathbf{H}^T\left(\widetilde{\mathbf{R}}_k^{(i)}\right)^{-1}, \qquad \left(\widetilde{\mathbf{P}}_{12,k}^{(\bar{\imath})} = \left(\mathbf{I} - \mathbf{K}_{\boldsymbol{\xi},k}^{(i)}\mathbf{H}\right)\mathbf{P}_{12,k|k-1}^{(i)}\right).$

for each subset $V \subseteq Z_k \setminus \{\mathbf{z}\}$

for each permutation $i_{1:|V|} = (i_1, \ldots, i_{|V|})$

$\bar{n} := \bar{n}+1,$

for $\ell = 1, \ldots, |V|$

$\mathbf{b}_1^{(i,i_{1:\ell})} := \mathbf{H}\mathbf{m}_{\boldsymbol{\xi},k|k-1}^{(i_\ell)} + \mathbf{A}_k^{(i_\ell)}\left(\boldsymbol{\eta}_k^{(i,i_{1:\ell-1})} - \mathbf{m}_{\mathbf{x},k|k-1}\right),$

$\mathbf{S}_1^{(i,i_{1:\ell})} := \mathbf{A}_k^{(i_\ell)}\boldsymbol{\mathcal{P}}_k^{(i,i_{1:\ell-1})}\left(\mathbf{A}_k^{(i_\ell)}\right)^T + \widetilde{\mathbf{R}}_k^{(i_\ell)},$

$\widetilde{\mathbf{m}}_{\mathbf{x},k}^{(i,i_{1:\ell})} := \boldsymbol{\eta}_k^{(i,i_{1:\ell-1})} + \boldsymbol{\mathcal{P}}_k^{(i,i_{1:\ell-1})}\left(\mathbf{A}_k^{(i_\ell)}\right)^T\left(\mathbf{S}_1^{(i,i_{1:\ell})}\right)^{-1}\left(\mathbf{z}_\ell - \mathbf{b}_1^{(i,i_{1:\ell})}\right),$

$\widetilde{\mathbf{P}}_{22,k}^{(i,i_{1:\ell})} := \left(\mathbf{I} - \boldsymbol{\mathcal{P}}_k^{(i,i_{1:\ell-1})}\left(\mathbf{A}_k^{(i_\ell)}\right)^T\left(\mathbf{S}_1^{(i,i_{1:\ell})}\right)^{-1}\mathbf{A}_k^{(i_\ell)}\right)\boldsymbol{\mathcal{P}}_k^{(i,i_{1:\ell-1})},$

$\mathbf{S}_2^{(i,i_{1:\ell})} := \widehat{\mathbf{H}}\widetilde{\mathbf{P}}_{22,k}^{(i,i_{1:\ell})}\widehat{\mathbf{H}}^T + \mathbf{R}_2, \quad \mathbf{K}_{\mathbf{x},k}^{(i,i_{1:\ell})} := \widetilde{\mathbf{P}}_{22,k}^{(i,i_{1:\ell})}\widehat{\mathbf{H}}^T\left(\mathbf{S}_2^{(i,i_{1:\ell})}\right)^{-1},$

$\boldsymbol{\eta}_k^{(i,i_{1:\ell})} := \widetilde{\mathbf{m}}_{\mathbf{x},k}^{(i,i_{1:\ell})} + \mathbf{K}_{\mathbf{x},k}^{(i,i_{1:\ell})}\left(\mathbf{z}_\ell - \widehat{\mathbf{H}}\widetilde{\mathbf{m}}_{\mathbf{x},k}^{(i,i_{1:\ell})}\right),$

$\boldsymbol{\mathcal{P}}_{22,k}^{(i,i_{1:\ell})} := \left(\mathbf{I} - \mathbf{K}_{\mathbf{x},k}^{(i,i_{1:\ell})}\widehat{\mathbf{H}}\right)\widetilde{\mathbf{P}}_{22,k}^{(i,i_{1:\ell})},$

$L_{i_\ell}^{(\bar{n})} := \dfrac{p_D}{\kappa(\mathbf{z}_\ell)}\,\omega_{k|k-1}^{(i_\ell)}\mathcal{N}\left(\mathbf{z}_\ell; \mathbf{b}_1^{(i,i_{1:\ell})}, \mathbf{S}_1^{(i,i_{1:\ell})}\right)\mathcal{N}\left(\mathbf{z}_\ell; \widehat{\mathbf{H}}\widetilde{\mathbf{m}}_{\mathbf{x},k}^{(i,i_{1:\ell})}, \mathbf{S}_2^{(i,i_{1:\ell})}\right).$

end

$\mathbf{m}_{\mathbf{x},k}^{(\bar{n})} := \boldsymbol{\eta}_k^{(i,i_{1:|V|})}, \quad \mathbf{P}_{22,k}^{(\bar{n})} := \boldsymbol{\mathcal{P}}_k^{(i,i_{1:|V|})}, \quad \tilde{\nu}_k^{(\bar{n})} := \tilde{\omega}_k^{(\bar{\imath})}\prod_{\ell=1}^{|V|}L_{i_\ell}^{(\bar{n})},$

end

end

end

end

$J_k := \bar{\imath}, \quad \bar{J} := \{\bar{J}, \bar{n}+1\}.$

output $\left\{\widetilde{\mathbf{m}}_{\boldsymbol{\xi},k}^{(i)}, \widetilde{\mathbf{P}}_{11,k}^{(i)}, \widetilde{\mathbf{P}}_{12,k}^{(i)}\right\}_{i=1}^{J_k}$, $\left\{\mathbf{m}_{\mathbf{x},k}^{(j)}, \mathbf{P}_{22,k}^{(j)}\right\}_{j=1}^{\bar{n}}$,
$\left\{\omega_k^{(i)}\right\}_{i=1}^{\bar{I}_1-1}$, $\left\{\tilde{\omega}_k^{(i)}\right\}_{i=\bar{I}_1}^{J_k}$, $\left\{\nu_k^{(j)}\right\}_{j=1}^{\bar{J}_1-1}$, $\left\{\tilde{\nu}_k^{(j)}\right\}_{j=\bar{J}_1}^{\bar{n}}$, with $\bar{I}$ and $\bar{J}$.

Table D.3: Pseudo-code for step 3 of the single group GM-PHD update

given $\left\{\omega_k^{(i)}\right\}_{i=1}^{\bar{I}_1-1}$, $\left\{\nu_k^{(j)}\right\}_{j=1}^{\bar{J}_1-1}$, $\left\{\tilde{\omega}_k^{(i)}\right\}_{i=\bar{I}_1}^{J_k}$, $\left\{\tilde{\nu}_k^{(j)}\right\}_{j=\bar{J}_1}^{\bar{n}}$, with $\bar{I}$ and $\bar{J}$.
step 4. (*Normalisation of the weights for detections*)
  for $j = 1, \ldots, |\bar{I}|$
    if $j = |\bar{I}|$
      for $i = \bar{I}_j, \ldots, J_k$ (*where* $\bar{I}_j$ *denotes the* j^{th} *entry of the set* $\bar{I}$)
        $\omega_k^{(i)} := \tilde{\omega}_k^{(i)} \Big/ \sum_{i'=\bar{I}_j}^{J_k} \tilde{\omega}_k^{(i')}$.
      end
    else
      for $i = \bar{I}_j, \ldots, \bar{I}_{j+1} - 1$
        $\omega_k^{(i)} := \tilde{\omega}_k^{(i)} \Big/ \sum_{i'=\bar{I}_j}^{\bar{I}_{j+1}-1} \tilde{\omega}_k^{(i')}$.
      end
    end
  end
  for $i = 1, \ldots, |\bar{J}| - 1$
    for $j = \bar{J}_i, \ldots, \bar{J}_{i+1} - 1$ (*where* $\bar{J}_i$ *denotes the* i^{th} *entry of the set* $\bar{J}$)
      $\nu_k^{(j)} := \tilde{\nu}_k^{(j)} \Big/ \sum_{j'=\bar{J}_i}^{\bar{J}_{i+1}-1} \tilde{\nu}_k^{(j')}$.
    end
  end
output $\left\{\omega_k^{(i)}\right\}_{i=1}^{J_k}$ and $\left\{\nu_k^{(j)}\right\}_{j=1}^{\bar{n}}$.

Table D.4: Pseudo-code for step 4 of the single group GM-PHD update

given $\mathbf{m}_{\mathbf{x},k|k-1}$, $\mathbf{P}_{22,k|k-1}$, $\left\{\nu_k^{(j)}, \mathbf{m}_{\mathbf{x},k}^{(j)}, \mathbf{P}_{22,k}^{(i)}\right\}_{j=1}^{\bar{n}}$, $\left\{\omega_k^{(i)}, \widetilde{\mathbf{m}}_{\boldsymbol{\xi},k}^{(i)}, \widetilde{\mathbf{P}}_{11,k}^{(i)}, \widetilde{\mathbf{P}}_{12,k}^{(i)}\right\}_{i=1}^{J_k}$,
  and truncation threshold τ_1.
set
  $I_{\mathbf{x},k} := \left\{j = 1, \ldots, \bar{n} \,\middle|\, \nu_k^{(j)} > \tau_1\right\}$,
  $\tilde{\nu}_k := \sum_{j \in I_{\mathbf{x},k}} \nu_k^{(j)}$, $\mathbf{m}_{\mathbf{x},k} := \frac{1}{\tilde{\nu}_k} \sum_{j \in I_{\mathbf{x},k}} \nu_k^{(j)} \mathbf{m}_{\mathbf{x},k}^{(j)}$, $\mathbf{P}_{22,k} := \frac{1}{\tilde{\nu}_k} \sum_{j \in I_{\mathbf{x},k}} \nu_k^{(j)} \mathbf{P}_{22,k}^{(j)}$.
  for $i = 1, \ldots, J_k$
    $\mathbf{m}_{\boldsymbol{\xi},k}^{(i)} := \widetilde{\mathbf{m}}_{\boldsymbol{\xi},k}^{(i)} + \mathbf{P}_{12,k}^{(i)} \mathbf{P}_{22,k}^{-1} \left(\mathbf{m}_{\mathbf{x},k} - \mathbf{m}_{\mathbf{x},k|k-1}\right)$,
    $\mathbf{P}_{11,k}^{(i)} := \widetilde{\mathbf{P}}_{11,k}^{(i)} - \mathbf{P}_{12,k}^{(i)} \mathbf{P}_{22,k}^{-1} \left(\widetilde{\mathbf{P}}_{12,k}^{(i)} - \mathbf{P}_{12,k}^{(i)}\right)^T$, $\mathbf{P}_{12,k}^{(i)} := \widetilde{\mathbf{P}}_{12,k}^{(i)} \mathbf{P}_{22,k|k-1}^{-1} \mathbf{P}_{22,k}$.
  end
output $\mathbf{m}_{\mathbf{x},k}$, $\mathbf{P}_{22,k}$, $\left\{\omega_k^{(i)}, \mathbf{m}_{\boldsymbol{\xi},k}^{(i)}, \mathbf{P}_{11,k}^{(i)}, \mathbf{P}_{12,k}^{(i)}\right\}_{i=1}^{J_k}$.

Table D.5: Merging the Gaussian components relating to the single group target

The updated weights, as evaluated in Tables D.2 and D.4 are therefore representative of the expression for the multiple measurement likelihood $L_{Z_k}(\mathbf{x}_k)$ given in equation (D.4).

Finally, in order for the GM formulation to present a closed-form solution to the single group PHD recursion, it is necessary to merge the Gaussian components relating to the updated group state. The details of this procedure are given in Table D.5. The simulation results presented in Section 4.3.3 are obtained by applying the single group GM-PHD filter to the scenario described in Section 4.3.2 with the parameter $\tau_1 = 10^{-4}$.

D.2 Validating Measurements

Various mixture reduction techniques, similar to those identified in Section 4.2, can also be applied to reduce the computation at each iteration of the single group GM-PHD recursion. In particular, although partitioning the measurement set is not necessary, as it is for the multi-group PHD update, the update equation given by equation (3.27) does involve summing over all possible subsets $W \subseteq Z_k$, as a consequence of expression for the multiple measurement likelihood given in equation (D.1). The computational challenge this present was discussed in Section 3.3.2 and in terms of the GM formulation of the single group PHD update it means having to compute a large number of Gaussian components. This implementation issue can be addressed by performing a validation/gating technique, similar to that shown in Section 4.2.1, as follows.

For the initial time-step $k = 1$, the gating procedure shown in Table D.6 replaces the measurement set Z_k with a single subset W and removes the sum $\sum_{W \subseteq Z_k}$ in equation (D.1), as well as replacing the sum over all measurements in Z_k in equation (3.29) with $\sum_{\mathbf{z} \in W}$. For subsequent time-steps $k > 1$, the gating procedure shown in Table D.7 reduces the number of possible subsets of Z_k to n_g, giving $\{W_j\}_{j=1}^{n_g}$, and replaces the sum $\sum_{W \subseteq Z_k}$ in equation (D.1) with $\sum_{j=1}^{n_g}$, as well as replacing the sum over all measurements in Z_k in equation (3.29) with $\sum_{\mathbf{z} \in W^{\cup}}$ where $W^{\cup} = \bigcup_{j=1}^{n_g} W_j$.

---

given a source point where the group target is expected to originate from $\mathbf{m}_{\text{source}}$, noise covariance matrix $\mathbf{R}_G \propto \mathbf{R}_2$, measurement set Z_k, and gating threshold T_2.

set $W = \left\{ \mathbf{z} \in Z_k \,\middle|\, (\mathbf{z} - \mathbf{m}_{\text{source}})^T \mathbf{R}_G^{-1} (\mathbf{z} - \mathbf{m}_{\text{source}}) \leq T_2 \right\}$.

output W.

Table D.6: Initial gating procedure for determining a validated subset of measurements at time $k = 1$

given $\mathbf{m}_{\mathbf{x},k|k-1}$, $\mathbf{P}_{22,k|k-1}$, J_{k-1}, the minimum allowable number of Gaussian components relating to individual targets $J_{\min}$, measurement set Z_k, noise covariance matrix $\mathbf{R}_G \propto \mathbf{R}_2$ and gating threshold T_2

procedure

If $J_{k-1} > J_{\min}$, then set the number of gates as $N = \lfloor J_{k-1}\,/J_{\min}\,\rceil$, otherwise set $N = 1$.

$n_g := 0, \quad \boldsymbol{\Sigma} := \widehat{\mathbf{H}}\mathbf{P}_{22,k|k-1}\widehat{\mathbf{H}}^T + \mathbf{R}_G$.

for $i = 1, \ldots, N$

Sample $\widetilde{\mathbf{m}} \sim \mathcal{N}(\mathbf{x}_k\,; \mathbf{m}_{\mathbf{x},k|k-1}, \mathbf{P}_{22,k|k-1})$,

$$V := \left\{ \mathbf{z} \in Z_k \,\middle|\, \left(\mathbf{z} - \widehat{\mathbf{H}}\,\widetilde{\mathbf{m}}\right)^T \boldsymbol{\Sigma}^{-1} \left(\mathbf{z} - \widehat{\mathbf{H}}\,\widetilde{\mathbf{m}}\right) \leq T_2 \right\}.$$

if $i = 1$

$n_g := n_g + 1, \quad W_{n_g} := V$.

else

if V is not a duplicate of any of the subsets W_j determined so far, for $j = 1, \ldots, n_g$, then set $n_g = n_g + 1$ and $W_{n_g} = V$.

end

output $\{W_j\}_{j=1}^{n_g}$.

Table D.7: Gating procedure for determining validated subsets of measurements for subsequent time-steps $k > 1$

The consequence of performing either procedure is the significant reduction in the number of Gaussian components for the GM formulation of the update equation.

D.3 Initialisation

For the simulated example described in Section 4.3.2, the single group GM-PHD filter is initialised with the intensity $v_1(\mathbf{x}, \boldsymbol{\xi})$ given by equation (4.64), which is driven by the measurements received at time $k = 1$, a concept that was proposed in [76] and later developed as a technique for the PHD filter in [31, 32]. Specifically, the initial intensity is driven by the measurements from the validated subset W, determined as shown in Table D.6 given the source point $\mathbf{m}_{\text{source}} = [0\ 450]^T$, so that the means of the components relating to the individual targets are sampled from the measurements $\mathbf{z} \in W$ and the mean relating to the group state is found from the averaged sum of the samples. That is, $\mathbf{m}_{\boldsymbol{\xi}}^{(i)} = \left[m_{\boldsymbol{\xi},1}^{(i)} \;\; m_{\boldsymbol{\xi},2}^{(i)}\right]^T \sim \mathcal{N}(\boldsymbol{\xi}\,; \mathbf{z}_i, \mathbf{R}_1)$ for $i = 1, \ldots, |W|$, and

$$\mathbf{m}_{\mathbf{x}} = \left[\frac{1}{|W|} \sum_{i=1}^{|W|} m_{\boldsymbol{\xi},1}^{(i)} \quad 0 \quad \frac{1}{|W|} \sum_{i=1}^{|W|} m_{\boldsymbol{\xi},2}^{(i)} \quad 0 \right]^T, \tag{D.5}$$

The weights and covariances of the components relating to individual targets are then set to be $\omega^{(i)} = 0.1$ and $\mathbf{P}_{11}^{(i)} = \mathbf{R}_1 + \mathbf{P}_{12}^{(i)}\mathbf{P}_{22}^{-1}\left(\mathbf{P}_{12}^{(i)}\right)^T$ for $i = 1, \ldots, |W|$, where the covariance relating to the group state and the cross diagonal covariances are given by $\mathbf{P}_{22} = \text{diag}([5\sigma_2^2, 2, 5\sigma_2^2, 2])$ and $\mathbf{P}_{12}^{(i)} = \widehat{\mathbf{H}}\mathbf{P}_{22}$ respectively. The values of $\mathbf{R}_1$ and σ_2 are as given in Section 4.3.2.

At subsequent time-steps $k > 1$, the individual target births in the group are given by the intensity $\gamma_{\boldsymbol{\xi},k}(\boldsymbol{\xi}_k \,|\, \mathbf{x}_k)$ in equation (4.65) which is driven by measurements from the validated subsets $\{W_j\}_{j=1}^{n_g}$, determined as shown in Table D.7. That is, the means of the birth components are sampled from the measurements $\mathbf{z} \in W^{\cup}$, i.e. $\mathbf{m}_{\gamma,\boldsymbol{\xi},k}^{(i)} \sim \mathcal{N}(\boldsymbol{\xi}_k\,; \mathbf{z}_i, \mathbf{R}_1)$ for $i = 1, \ldots, J_{\gamma,k}$ so that $J_{\gamma,k} = |W^{\cup}|$. The associated weights and covariances are set to be $\omega_{\gamma,k}^{(i)} = 0.1$ and $\mathbf{P}_{\gamma,11,k}^{(i)} = \mathbf{R}_1 + \mathbf{P}_{\gamma,12,k}^{(i)}\mathbf{P}_{22,k|k-1}^{-1}\left(\mathbf{P}_{\gamma,12,k}^{(i)}\right)^T$, where the cross diagonal covariances $\mathbf{P}_{\gamma,12,k}^{(i)}$ are set the same way as for $\mathbf{P}_{12}^{(i)}$ at the initial time-step.

D.4 Derivation of the Process Noise Covariance Matrices

The evolution of the single group target scenario follows the linear Gaussian model given in equation (4.6) (Section 4.1.2), with the noise covariance defined by the block matrix $\mathbf{Q}_k$ given in (4.7), consisting of the covariance matrices $\mathbf{Q}_{11}$, $\mathbf{Q}_{12}$ and $\mathbf{Q}_{22}$. This section provides details on the derivation of the expressions for $\mathbf{Q}_{11}$ and $\mathbf{Q}_{12}$ given in (4.62), and the expression for $\mathbf{Q}_{22}$ given in (4.60) and (4.61), as proposed in Section 4.3.2.

Firstly, the expression for the noise covariance matrix $\mathbf{Q}_{22}$, follows immediately from the constant linear velocity kinematic model described in [64]. Consider the group state vector $\mathbf{x}_k = \begin{bmatrix} x_{k,\mathsf{x}} & \dot{x}_{k,\mathsf{x}} & x_{k,\mathsf{y}} & \dot{x}_{k,\mathsf{y}} \end{bmatrix}^T$, where $\dot{x}_{k,\mathsf{x}}$, $\dot{x}_{k,\mathsf{y}}$ denote the velocities in the x and y directions. Given a sampling interval t, the discrete-time state equation is given by

$$\mathbf{x}_k = \mathbf{F}_{22}\mathbf{x}_{k-1} + \mathbf{v}_{\mathsf{x},k-1} + \mathbf{v}_{\mathsf{y},k-1}, \tag{D.6}$$

where $\mathbf{F}_{22}$ is as defined in (4.60) and (4.61). The discrete-time process noise relating to the change in which the velocity undergoes in the x direction, which is modelled by the white noise $\tilde{v}_{\mathsf{x}}$, is given by

$$\mathbf{v}_{\mathsf{x},k-1} = \begin{bmatrix} \mathbf{I}_2 \\ \mathbf{0}_2 \end{bmatrix} \int_0^t \begin{bmatrix} 1 & t-\tau \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tilde{v}_{\mathsf{x}}\, \mathrm{d}\tau = \begin{bmatrix} \mathbf{I}_2 \\ \mathbf{0}_2 \end{bmatrix} \int_0^t \begin{bmatrix} t-\tau \\ 1 \end{bmatrix} \tilde{v}_{\mathsf{x}}\, \mathrm{d}\tau. \tag{D.7}$$

Similarly, the discrete-time process noise relating to the change the velocity under-

goes in the y direction is given by

$$\mathbf{v}_{\mathsf{y},k-1} = \begin{bmatrix} \mathbf{0}_2 \\ \mathbf{I}_2 \end{bmatrix} \int_0^t \begin{bmatrix} t-\tau \\ 1 \end{bmatrix} \tilde{v}_{\mathsf{y}} \, \mathrm{d}\tau, \tag{D.8}$$

where $\tilde{v}_{\mathsf{y}}$ denotes white noise modelling this change. Suppose that $\mathsf{E}(\tilde{v}_{\mathsf{x}}^2) = \mathsf{E}(\tilde{v}_{\mathsf{y}}^2) = \sigma_{\mathbf{x}}^2$ ($\mathsf{E}(\tilde{v}_{\mathsf{x}}\tilde{v}_{\mathsf{y}}) = 0$), then the process noise covariance $\mathbf{Q}_{22}$ is given by

$$\mathbf{Q}_{22} = \mathsf{E}(\mathbf{v}_{\mathsf{x},k-1}\mathbf{v}_{\mathsf{x},k-1}^T) + \mathsf{E}(\mathbf{v}_{\mathsf{y},k-1}\mathbf{v}_{\mathsf{y},k-1}^T), \tag{D.9}$$

where

$$\begin{aligned} \mathsf{E}(\mathbf{v}_{\mathsf{x},k-1}\mathbf{v}_{\mathsf{x},k-1}^T) &= \sigma_{\mathbf{x}}^2 \begin{bmatrix} \mathbf{I}_2 \\ \mathbf{0}_2 \end{bmatrix} \left(\int_0^t \begin{bmatrix} t-\tau \\ 1 \end{bmatrix} \begin{bmatrix} t-\tau & 1 \end{bmatrix} \mathrm{d}\tau \right) \begin{bmatrix} \mathbf{I}_2 & \mathbf{0}_2 \end{bmatrix} \\ &= \sigma_{\mathbf{x}}^2 \begin{bmatrix} \mathbf{I}_2 \\ \mathbf{0}_2 \end{bmatrix} \begin{bmatrix} t^3/3 & t^2/2 \\ t^2/2 & t \end{bmatrix} \begin{bmatrix} \mathbf{I}_2 & \mathbf{0}_2 \end{bmatrix} = \sigma_{\mathbf{x}}^2 \begin{bmatrix} \mathbf{Q}_k & \mathbf{0}_2 \\ \mathbf{0}_2 & \mathbf{0}_2 \end{bmatrix}, \end{aligned} \tag{D.10}$$

and similarly

$$\mathsf{E}(\mathbf{v}_{\mathsf{y},k-1}\mathbf{v}_{\mathsf{y},k-1}^T) = \sigma_{\mathbf{x}}^2 \begin{bmatrix} \mathbf{0}_2 & \mathbf{0}_2 \\ \mathbf{0}_2 & \mathbf{Q}_k \end{bmatrix}. \tag{D.11}$$

Hence the expression for $\mathbf{Q}_{22}$ given in (4.60) follows immediately with $\mathbf{Q}_k$ given by the expression in (4.61).

Now consider an individual target state vector $\boldsymbol{\xi}_k = \begin{bmatrix} \xi_{k,\mathsf{x}} & \xi_{k,\mathsf{y}} \end{bmatrix}^T$. Given the description of the single group target scenario in Section 4.3.2 and sampling interval t, the discrete-time state equation for individual targets is given by

$$\boldsymbol{\xi}_k = \boldsymbol{\xi}_{k-1} + \mathbf{F}_{12}\mathbf{x}_{k-1} + \bar{\mathbf{v}}_{\mathsf{x},k-1} + \bar{\mathbf{v}}_{\mathsf{y},k-1}, \tag{D.12}$$

where $\mathbf{F}_{12} = \widehat{\mathbf{H}}\,(\mathbf{F}_{22} - \mathbf{I}_4)$, for projection matrix $\widehat{\mathbf{H}}$ given by equation (4.63). The discrete-time process noise relating to the change in which the velocity of the group state undergoes in the x direction is given by

$$\bar{\mathbf{v}}_{\mathsf{x},k-1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \int_0^t \begin{bmatrix} 1 & 0 \end{bmatrix} \left(\begin{bmatrix} 1 & t-\tau \\ 0 & 1 \end{bmatrix} - \mathbf{I}_2 \right) \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tilde{v}_{\mathsf{x}} \, \mathrm{d}\tau = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \int_0^t (t-\tau)\, \tilde{v}_{\mathsf{x}} \, \mathrm{d}\tau. \tag{D.13}$$

Similarly, the discrete-time process noise relating to the change in which the velocity

of the group state undergoes in the y direction is given by

$$\bar{\mathbf{v}}_{\mathsf{y},k-1} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \int_0^t (t-\tau)\, \tilde{v}_{\mathsf{y}}\, \mathrm{d}\tau. \tag{D.14}$$

The inclusion of the process noise vectors $\bar{\mathbf{v}}_{\mathsf{x},k-1}$ and $\bar{\mathbf{v}}_{\mathsf{y},k-1}$ in the state equation (D.12) can be account for by the conditionality of the individual target states on the group state. It follows that the process noise covariance $\mathbf{Q}_{11}$ is given by

$$\mathbf{Q}_{11} = \mathsf{E}(\bar{\mathbf{v}}_{\mathsf{x},k-1}\bar{\mathbf{v}}_{\mathsf{x},k-1}^T) + \mathsf{E}(\bar{\mathbf{v}}_{\mathsf{y},k-1}\bar{\mathbf{v}}_{\mathsf{y},k-1}^T), \tag{D.15}$$

where

$$\mathsf{E}(\bar{\mathbf{v}}_{\mathsf{x},k-1}\bar{\mathbf{v}}_{\mathsf{x},k-1}^T) = \sigma_{\mathbf{x}}^2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} \left(\int_0^t (t-\tau)^2\, \mathrm{d}\tau \right) \begin{bmatrix} 1 & 0 \end{bmatrix} = \sigma_{\mathbf{x}}^2 \begin{bmatrix} t^3/3 & 0 \\ 0 & 0 \end{bmatrix}, \tag{D.16}$$

and similarly

$$\mathsf{E}(\bar{\mathbf{v}}_{\mathsf{y},k-1}\bar{\mathbf{v}}_{\mathsf{y},k-1}^T) = \sigma_{\mathbf{x}}^2 \begin{bmatrix} 0 & 0 \\ 0 & t^3/3 \end{bmatrix}. \tag{D.17}$$

Hence the expression for $\mathbf{Q}_{11}$ given in (4.62) follows by substituting the expressions for $\mathsf{E}(\bar{\mathbf{v}}_{\mathsf{x},k-1}\bar{\mathbf{v}}_{\mathsf{x},k-1}^T)$ and $\mathsf{E}(\bar{\mathbf{v}}_{\mathsf{y},k-1}\bar{\mathbf{v}}_{\mathsf{y},k-1}^T)$ in (D.16) and (D.17) respectively into equation (D.15). Finally, the process noise covariance $\mathbf{Q}_{12}$ can be written as

$$\mathbf{Q}_{12} = \mathsf{E}(\bar{\mathbf{v}}_{\mathsf{x},k-1}\mathbf{v}_{\mathsf{x},k-1}^T) + \mathsf{E}(\bar{\mathbf{v}}_{\mathsf{y},k-1}\mathbf{v}_{\mathsf{y},k-1}^T), \tag{D.18}$$

where

$$\begin{aligned} \mathsf{E}(\bar{\mathbf{v}}_{\mathsf{x},k-1}\mathbf{v}_{\mathsf{x},k-1}^T) &= \sigma_{\mathbf{x}}^2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} \left(\int_0^t (t-\tau) \begin{bmatrix} t-\tau & 1 \end{bmatrix} \mathrm{d}\tau \right) \begin{bmatrix} \mathbf{I}_2 & \mathbf{0}_2 \end{bmatrix} \\ &= \sigma_{\mathbf{x}}^2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} t^3/3 & t^2/2 \end{bmatrix} \begin{bmatrix} \mathbf{I}_2 & \mathbf{0}_2 \end{bmatrix} = \sigma_{\mathbf{x}}^2 \begin{bmatrix} \begin{bmatrix} t^3/3 & t^2/2 \end{bmatrix} & \mathbf{0}_{2\times 1} \\ \mathbf{0}_{2\times 1} & \mathbf{0}_{2\times 1} \end{bmatrix}, \end{aligned} \tag{D.19}$$

and similarly

$$\mathsf{E}(\bar{\mathbf{v}}_{\mathsf{y},k-1}\mathbf{v}_{\mathsf{y},k-1}^T) = \sigma_{\mathbf{x}}^2 \begin{bmatrix} \mathbf{0}_{2\times 1} & \mathbf{0}_{2\times 1} \\ \mathbf{0}_{2\times 1} & \begin{bmatrix} t^3/3 & t^2/2 \end{bmatrix} \end{bmatrix}. \tag{D.20}$$

The expression for $\mathbf{Q}_{12}$ given in (4.62) therefore follows again by substitution.

Appendix E

The Benchmark GM-PHD Filter for Extended Targets

This chapter summarises the benchmark filter used in the performance analysis of the GM-PHD filter for extended targets proposed in Section 5.1, the results of which, for the scenario described in Section 5.2.1, are given in Section 5.2.2. It is based on a Gaussian mixture formulation of a PHD filter similar to that introduced in Section 2.4.4 for non-standard targets.

Section E.1 introduces the alternative PHD filter for non-standard targets and Section E.2 summarises its closed-form solution, paying particular attention to the GM formulation of the update equation. It is then discussed in Section E.3 how this benchmark filter relates to the proposed GM-PHD filter for extended targets from Section 5.1.

E.1 An Alternative PHD Filter for Non-Standard Targets

The PHD filter for non-standard targets [100] introduced in Section 2.4.4 handles targets that potentially produce more than one measurement per time-step, otherwise referred to as *extended targets*. The recursion only differs from that of the multi-target PHD filter [1] in the update equation. One of the assumptions the update equation in (2.36) was based on, concerns the clutter RFS K_k present in the measurement model, that it is Poisson with intensity $\kappa_k(\cdot)$.

Instead consider the hierarchical RFS representation of the clutter RFS K_k whose probability density is given by equation (3.19) and (3.20), so that the RFS Z_k describes a collection of measurement clusters, which are either target generated or false alarms. Based on this alternative assumption concerning the clutter RFS, along with the remaining assumption listed in Section 2.4.4, the update equation

is now

$$v_k(\mathbf{x}_k) = v_{k|k-1}(\mathbf{x}_k) \left\{ 1 - p_D(\mathbf{x}_k) + p_D(\mathbf{x}_k)\, \mathrm{e}^{-\alpha(\mathbf{x}_k)} + \sum_{\pi \in \Pi_{Z_k}} \varpi_\pi \sum_{\varphi \in \pi} \frac{p_D(\mathbf{x}_k)\, \mathrm{e}^{-\alpha(\mathbf{x}_k)} \left(\prod_{\mathbf{z} \in \varphi} \alpha(\mathbf{x}_k)\, l(\mathbf{z} \,|\, \mathbf{x}_k) \right)}{\Lambda(\varphi) + v_{k|k-1} \left[p_D\, \mathrm{e}^{-\alpha} \left(\prod_{\mathbf{z} \in \varphi} \alpha\, l_{\mathbf{z}} \right) \right]} \right\}, \tag{E.1}$$

where the partition weights are now given by

$$\varpi_\pi = \frac{\prod_{\varphi \in \pi} \left(\Lambda(\varphi) + v_{k|k-1} \left[p_D\, \mathrm{e}^{-\alpha} \left(\prod_{\mathbf{z} \in \varphi} \alpha\, l_{\mathbf{z}} \right) \right] \right)}{\sum_{\pi' \in \Pi_{Z_k}} \prod_{\varphi' \in \pi'} \left(\Lambda(\varphi') + v_{k|k-1} \left[p_D\, \mathrm{e}^{-\alpha} \left(\prod_{\mathbf{z} \in \varphi'} \alpha\, l_{\mathbf{z}} \right) \right] \right)}, \tag{E.2}$$

and $\Lambda(\varphi)$ denotes the clutter term given in equation (3.20).

E.1.1 Derivation of the update equation

The derivation of the result for the update equation given in equation (E.1) mostly follows that of the PHD update for non-standard targets [100]. It is recovered from the functional derivative of the updated p.g.fl., corresponding to the updated posterior in the multi-target Bayesian filter, according to the relation given by equation (A.16). The updated p.g.fl. G_k can be written in the form given by equation (B.21) which follows as a result of introducing a bivariate p.g.fl. defined as

$$F[g, h] = \int \left(\prod_{\mathbf{x} \in X_k} h(\mathbf{x}) \right) G_{Z,k}[g \,|\, X_k]\, p_{k|k-1}(X_k)\, \mathrm{d}X_k, \tag{E.3}$$

where $G_{Z,k}[g \,|\, X_k]$ is the p.g.fl. of the RFS Z_k. The measurement model is defined such that Z_k is given by the union target generated measurements and clutter in equation (2.27), resulting in the following p.g.fl.

$$G_{Z,k}[g \,|\, X_k] = G_{K,k}[g] \prod_{\mathbf{x} \in X_k} G_{\Theta,k}[g \,|\, \mathbf{x}], \tag{E.4}$$

The bivariate p.g.fl. is then written as $F[g, h] = G_{K,k}[g]\, G_{k|k-1}[h\, G_{\Theta,k}[g \,|\, \cdot\,]]$, and given the measurement set $Z_k = \{\mathbf{z}_1, \ldots, \mathbf{z}_m\}$, its m^{th}-order functional derivative

in the directions $\delta_{\mathbf{z}_1}, \ldots, \delta_{\mathbf{z}_m}$ is

$$\delta^m F[g, h\,; \delta_{\mathbf{z}_1}, \ldots, \delta_{\mathbf{z}_m}] = \sum_{W \subseteq Z_k} \delta^{m-|W|} G_{K,k}\big[g\,; \delta_{\mathbf{z}_1}, \ldots, \delta_{\mathbf{z}_{m-|W|}}\big]\, \delta^{|W|} G_{k|k-1}\big[h\, G_{\Theta,k}[g\,|\,\cdot\,]\,; \delta_{\mathbf{w}_1}, \ldots, \delta_{\mathbf{w}_{|W|}}\big],$$

which follows from the general product rule for functional derivatives [69, Section 11.6, page 389].

The assumption on the RFS Θ_k described in Section 2.4.4 implies that its p.g.fl. has the following form

$$G_{\Theta,k}[g\,|\,\cdot\,] = 1 - p_D(\,\cdot\,) + p_D(\,\cdot\,) \exp\left(\alpha(\,\cdot\,) \int g(\mathbf{z})\, l(\mathbf{z}\,|\,\cdot\,)\, \mathrm{d}\mathbf{z} - \alpha(\,\cdot\,)\right), \tag{E.5}$$

so that $G_{k|k-1}[h\, G_{\Theta,k}[g\,|\,\cdot\,]] = G_{k|k-1}\big[h\,\big(1 - p_D + p_D\, \mathrm{e}^{\alpha\, l[g\,|\,\cdot\,] - \alpha}\big)\big]$, denoting $l[g\,|\,\cdot\,] = \int g(\mathbf{z})\, l(\mathbf{z}\,|\,\cdot\,)\, \mathrm{d}\mathbf{z}$. Then, given the Poisson assumption on the predicted multi-target RFS X_k, it can be shown by induction that

$$\delta^{|W|} G_{k|k-1}\big[h\,\big(1 - p_D + p_D\, \mathrm{e}^{\alpha\, l[g\,|\,\cdot\,] - \alpha}\big)\,; \delta_{\mathbf{w}_1}, \ldots, \delta_{\mathbf{w}_{|W|}}\big] = \exp\Big(v_{k|k-1}\big[h\,\big(1 - p_D + p_D\, \mathrm{e}^{\alpha\, l[g\,|\,\cdot\,] - \alpha}\big) - 1\big]\Big) \sum_{\pi \in \Pi_W} \prod_{\varphi \in \pi} v_{k|k-1}\Big[h\, p_D\, \mathrm{e}^{\alpha\, l[g\,|\,\cdot\,] - \alpha} \prod_{\mathbf{z} \in \varphi} \alpha\, l_{\mathbf{z}}\Big],$$

denoting $l_{\mathbf{z}} = l(\mathbf{z}\,|\,\cdot\,)$. It follows, given the alternative assumption on the clutter RFS K_k and noting that $p_\kappa(Z_k \setminus W)$ can be recovered as shown in equation (B.25), that the updated p.g.fl. given by equation (B.21) has the form

$$\begin{aligned} G_k[h] = {} & \frac{\exp\big(v_{k|k-1}[h\,(1 - p_D + p_D\, \mathrm{e}^{-\alpha})]\big)}{\exp\big(v_{k|k-1}[1 - p_D + p_D\, \mathrm{e}^{-\alpha}]\big)} \\ & \times \frac{\displaystyle\sum_{\pi \in \Pi_Z} \prod_{\varphi \in \pi} \Big(\Lambda(\varphi) + v_{k|k-1}\Big[h\, p_D\, \mathrm{e}^{-\alpha}\Big(\textstyle\prod_{\mathbf{z} \in \varphi} \alpha\, l_{\mathbf{z}}\Big)\Big]\Big)}{\displaystyle\sum_{\pi' \in \Pi_Z} \prod_{\varphi' \in \pi'} \Big(\Lambda(\varphi') + v_{k|k-1}\Big[p_D\, \mathrm{e}^{-\alpha}\Big(\textstyle\prod_{\mathbf{z} \in \varphi'} \alpha\, l_{\mathbf{z}}\Big)\Big]\Big)}, \end{aligned} \tag{E.6}$$

denoting $Z = Z_k$. The result for the update equation given in equation (E.1) follows by taking the functional derivative of the expression in equation (E.6) for G_k in the direction $\delta_{\mathbf{x}}$ and evaluating at $h = 1$.

E.2 A Summary of the GM Formulation

The Gaussian mixture formulation for this benchmark filter is the same as shown in the illustrative diagram in Section 5.1.1. In fact, the weights, means and covariances of the Gaussian components constituting the predicted intensity mixture have the same closed-form expression as detailed in Table 5.1 (Section 5.1.2). However, the closed-form expression for the updated weights, means and covariances differ from those detailed in Table 5.2 and 5.3, due to the alternative update equation (E.1), as follows.

Suppose that the function denoted by $l(\mathbf{z} \,|\, \mathbf{x}_k)$ in equation (E.1) defines the likelihood that the measurement $\mathbf{z}$ is generated by the target $\mathbf{x}_k$ and that it is modelled with the linear Gaussian function $\mathcal{N}(\mathbf{z}\,; \mathbf{H}\mathbf{x}_k, \mathbf{R})$, with projection matrix $\mathbf{H}$ and noise covariance matrix $\mathbf{R}$. In addition, suppose that the probability of detection $p_D(\mathbf{x}_k)$ and the expected number of measurements generated by the target $\mathbf{x}_k$, denoted by $\alpha(\mathbf{x}_k)$, are state independent, i.e. they are arbitrary constants denoted by p_D and α respectively. Then, given the GM formulation of the predicted intensity $v_{k|k-1}(\mathbf{x}_k)$, the updated intensity $v_k(\mathbf{x}_k)$ is also a Gaussian mixture, where the closed-form expressions for the weights, means and covariances of the constituent Gaussian components are detailed in Table E.1.

This GM formulation provides the following expected number of targets

$$N_k = (1 - p_D + p_D\, \mathrm{e}^{-\alpha})\, N_{k|k-1} + \sum_{\pi \in \Pi_{Z_k}} \varpi_\pi \sum_{\varphi \in \pi} \sum_{i=1}^{J_{k|k-1}} \widetilde{\omega}_{\varphi,k}^{(i)}, \tag{E.7}$$

associated with the updated intensity v_k, where $N_{k|k-1}$ is the expected number of target associated with the predicted intensity given by equation (5.4) and

$$\widetilde{\omega}_{\varphi,k}^{(i)} = \frac{p_D\, \omega_{k|k-1}^{(i)}\, \mathrm{e}^{-\alpha} \prod_{\ell=1}^{m} L_\ell^{(i)}}{\Lambda(\varphi) + p_D\, \mathrm{e}^{-\alpha} \sum_{i'=1}^{J_{k|k-1}} \omega_{k|k-1}^{(i')} \prod_{\ell=1}^{m} L_\ell^{(i')}}. \tag{E.8}$$

From equation (E.7) it can be deduced that this GM formulation of the benchmark filter for extended targets requires, at any time-step k,

$$J_k = J_{k|k-1} + \prod_{\pi \in \Pi_{Z_k}} \prod_{\varphi \in \pi} J_{k|k-1} \times |\varphi|, \tag{E.9}$$

Gaussian components in v_k, where $J_{k|k-1}$ denotes the number of components in $v_{k|k-1}$.

given $\left\{\omega_{k|k-1}^{(i)}, \mathbf{m}_{k|k-1}^{(i)}, \mathbf{P}_{k|k-1}^{(i)}\right\}_{i=1}^{J_{k|k-1}}$, measurement set Z_k,
and the set of all possible partitions Π_{Z_k},
step 1. (*Update for missed detections*)
$\bar{\imath} := 0$,
for $i = 1, \ldots, J_{k|k-1}$
$\bar{\imath} := \bar{\imath} + 1$,
$\omega_k^{(\bar{\imath})} := (1 - p_D + p_D\, \mathrm{e}^{-\alpha})\, \omega_{k|k-1}^{(i)}, \quad \mathbf{m}_k^{(\bar{\imath})} := \mathbf{m}_{k|k-1}^{(i)}, \quad \mathbf{P}_k^{(\bar{\imath})} := \mathbf{P}_{k|k-1}^{(i)}$.
end
step 2. (*Update for detections*)
for each partition $\pi \in \Pi_{Z_k}$
for $i_\pi = 1, \ldots, |\pi|$ (*i.e. for each subset* $\varphi_{i_\pi} \in \pi$)
$m := |\varphi_{i_\pi}|, \quad n := \bar{\imath}$.
for $i = 1, \ldots, J_{k|k-1}$
$\bar{\imath} := \bar{\imath} + 1$,
for $\ell = 1, \ldots, m$
$\mathbf{b}_{1:\ell}^{(i)} := \mathbf{H}\boldsymbol{\eta}_{1:\ell}^{(i)}, \quad \mathbf{S}_1^{(j_{1:\ell},i)} := \mathbf{H}\boldsymbol{\mathcal{P}}_{1:\ell}^{(i)}\mathbf{H}^T + \mathbf{R}$,
$\mathbf{K}_{1:\ell}^{(i)} := \boldsymbol{\mathcal{P}}_{1:\ell}^{(i)}\mathbf{H}^T \left(\mathbf{S}_{1:\ell}^{(i)}\right)^{-1}$,
$\widetilde{\mathbf{m}}_{1:\ell}^{(i)} := \boldsymbol{\eta}_{1:\ell}^{(i)} + \mathbf{K}_{1:\ell}^{(i)} \left(\mathbf{z}_\ell - \mathbf{b}_{1:\ell}^{(i)}\right)$,
$\widetilde{\mathbf{P}}_{1:\ell}^{(i)} := \left(\mathbf{I} - \mathbf{K}_{1:\ell}^{(i)}\mathbf{H}\right) \boldsymbol{\mathcal{P}}_{1:\ell}^{(i)}$,
(*where* $\boldsymbol{\eta}_{1:\ell}^{(i)} = \mathbf{m}_{k|k-1}^{(i)}$ *and* $\boldsymbol{\mathcal{P}}_{1:\ell}^{(i)} = \mathbf{P}_{k|k-1}^{(i)}$ *for* $\ell = 1$)
$\boldsymbol{\eta}_{1:\ell}^{(i)} := \widetilde{\mathbf{m}}_{1:\ell}^{(i)}, \quad \boldsymbol{\mathcal{P}}_{1:\ell}^{(i)} := \widetilde{\mathbf{P}}_{1:\ell}^{(i)}$,
$L_\ell^{(i)} := \alpha\, \mathcal{N}\left(\mathbf{z}_\ell\,; \mathbf{b}_{1:\ell}^{(i)}, \mathbf{S}_{1:\ell}^{(i)}\right)$.
end
$\mathbf{m}_k^{(\bar{\imath})} := \boldsymbol{\eta}_{1:m}^{(i)}, \quad \mathbf{P}_k^{(\bar{\imath})} := \boldsymbol{\mathcal{P}}_{1:m}^{(i)}$,
$\tilde{\omega}_k^{(\bar{\imath})} := p_D\, \omega_{k|k-1}^{(i)}\, \mathrm{e}^{-\alpha} \prod_{\ell=1}^{m} L_\ell^{(i)}$.
end
$\omega_\varphi^{(i_\pi)} := \Lambda(\varphi_{i_\pi}) + \sum_{j=n}^{\bar{\imath}} \tilde{\omega}_k^{(j)}$.
for $j = n, \ldots, \bar{\imath}$
$\omega_k^{(j)} := \tilde{\omega}_k^{(j)} \Big/ \omega_\varphi^{(i_\pi)}$.
end
end
$\widetilde{\varpi}_\pi := \prod_{i_\pi=1}^{|\pi|} \omega_\varphi^{(i_\pi)}$.
end
for each $\pi \in \Pi_{Z_k}$
$\varpi_\pi := \widetilde{\varpi}_\pi \Big/ \sum_{\pi \in \Pi_{Z_k}} \widetilde{\varpi}_\pi$.
end
$J_k := \bar{\imath}$.
output $\left\{\omega_k^{(i)}, \mathbf{m}_k^{(i)}, \mathbf{P}_k^{(i)}\right\}_{i=1}^{J_k}$.

Table E.1: Pseudo-code for update step of the benchmark GM-PHD filter for extended targets

When comparing the expressions for J_k in equations (E.9) and (5.7) for the two filters, it can be seen that the complexity of the proposed GM-PHD filter for extended targets in Section 2.4.4 is marginally greater than that of the benchmark filter summarised here. However, the partitioning of the measurement set is still an integral part of the benchmark filter such that the number of partitions, accounting for the product $\prod_{\pi \in \Pi_{Z_k}}$ in equation (E.9), significantly contributes to the increase in the number of Gaussian components at each iteration. The partitioning problem is once again address by performing the gating procedure detailed in Table 4.3 (Section 4.2.1), while closely spaced Gaussian components are merged and the extended target states are extracted according to the additional mixture reduction procedures given in Tables 5.4 and 5.5 (Section 5.1.3) respectively.

E.3 How the Two Filters Relate

The GM formulation in which the weights, means and covariances have the closed-form expression given in Table E.1 can also be obtained for the update equation in the PHD filter for extended targets presented in Section 3.2.6, under a different linear Gaussian model to that given in Section 4.1.4 for the single measurement likelihood. Rather than the factorisation given in equation (4.15), consider the following linear Gaussian likelihood.

Suppose that the measurement $\mathbf{z}$ originates from the feature point $\boldsymbol{\xi}_k$ so that the measurement equation is given by

$$\mathbf{z} = \mathbf{H}\boldsymbol{\xi}_k + \mathbf{v}_{\mathbf{z}}, \tag{E.10}$$

where $\boldsymbol{\xi}_k$ is a linear transformation of the target state $\mathbf{x}_k$ given by

$$\boldsymbol{\xi}_k = \widehat{\mathbf{H}}\mathbf{x}_k + \mathbf{v}_{\boldsymbol{\xi}}. \tag{E.11}$$

Further suppose $\mathbf{v}_{\mathbf{z}}$ and $\mathbf{v}_{\boldsymbol{\xi}}$ are zero mean (white) Gaussian noise with respective covariances $\mathbf{R}_1$ and $\mathbf{R}_2$, so that

$$\begin{aligned} \mathsf{E}[\mathbf{z}] &= \mathsf{E}[\mathbf{H}\boldsymbol{\xi}_k + \mathbf{v}_{\mathbf{z}}] = \mathbf{H}\mathsf{E}[\boldsymbol{\xi}_k] = \mathbf{H}\widehat{\mathbf{H}}\mathbf{x}_k, \\ \mathsf{Cov}[\mathbf{z}] &= \mathsf{Cov}[\mathbf{H}\boldsymbol{\xi}_k + \mathbf{v}_{\mathbf{z}}] = \mathbf{H}\mathsf{Cov}[\boldsymbol{\xi}_k]\mathbf{H}^T + \mathbf{R}_1 = \mathbf{H}\mathbf{R}_2\mathbf{H}^T + \mathbf{R}_1. \end{aligned} \tag{E.12}$$

It follows that the single measurement likelihood is then given by

$$g_k(\mathbf{z} \,|\, \boldsymbol{\xi}_k, \mathbf{x}_k) = \mathcal{N}(\mathbf{z}\,; \mathbf{H}\widehat{\mathbf{H}}\mathbf{x}_k, \mathbf{H}\mathbf{R}_2\mathbf{H}^T + \mathbf{R}_1). \tag{E.13}$$

Then, given the GM formulation of the predicted intensity $v_{k|k-1}(\mathbf{x}_k)$ (Section 5.1.2), the updated intensity $v_k(\mathbf{x}_k)$ is also a Gaussian mixture and is effectively the same formulation as given in Section E.2 with the closed-form expressions for the updated weights, means and covariances as detailed in Table E.1. Note that the expected number of measurements per target is equivalent to the expected number of feature points and they're both denoted by α in respective filters.

References

[1] R. P. S. Mahler, "Multi-target Bayes filtering via first-order multi-target moments," *IEEE Trans. on AES*, vol. 39, no. 4, pp. 1152–1178, 2003.

[2] D. E. Clark and J. Bell, "Bayesian multiple target tracking in forward scan sonar images using the PHD filter," *Radar, Sonar and Navigation, IEE Proceedings*, vol. 152, pp. 327–334, 2005.

[3] D. E. Clark, B. N. Vo, and J. Bell, "The GM-PHD filter multitarget tracking in sonar images," in *Proc. SPIE*, vol. 6235, 2006.

[4] D. E. Clark, J. Bell, Y. de Saint-Pern, and Y. Pettilot, "PHD filter multi-target tracking in 3D-sonar," in *Proc. of IEEE OCEANS05-Europe*, pp. 265–270, 2005.

[5] D. E. Clark, A. T. Cemgil, P. Peeling, and S. Godsill, "Multi-object tracking of sinusoidal components in audio with the Gaussian mixture probability hypothesis density filter," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pp. 339–342, October 2007.

[6] T. Zajic, B. Ravichandran, R. P. S. Mahler, R. Mehra, and M. Noviskey, "Joint target and identification with robustness against unmodeled targets," in *Proc. SPIE*, vol. 5096, pp. 279–290, 2003.

[7] M. Tobias and A. D. Lanterman, "Probability hypothesis density-based multitarget tracker using multiple bistatic range and velocity measurements," in *Proc. of 36th Southeast Symposium on System Theory*, (Atlanta, GA, USA), pp. 205–209, March 2004.

[8] M. Tobias and A. D. Lanterman, "Multitarget tracking using multiple bistatic range measurements with probability hypothesis densities," in *Proc. SPIE*, vol. 5429, 2004.

[9] M. Tobias and A. D. Lanterman, "Probability hypothesis density-based multitarget tracking with bistatic range and Doppler observations," *Radar, Sonar and Navigation, IEE Proceedings*, vol. 152, no. 3, pp. 195–205, 2005.

[10] S. Jovanoska and R. Thomä, "Multiple target tracking by a distributed UWB sensor network based on the PHD filter," in *Proc. 15th Int'l Conf. on Information Fusion*, (Singapore), pp. 1095–1102, July 2012.

[11] N. Ikoma, T. Uchiro, and T. Maeda, "Tracking of feature points in image sequence by SMC implementations of the PHD filter," in *Proc. of ICE Annual Conference*, vol. 2, (Sapporo), pp. 1696–1701, 2004.

[12] C. P. Haworth, Y. de Saint-Pern, D. E. Clark, E. Trucco, and Y. Petillot, "Detection and tracking of multiple metallic objects in millimetre-wave images," *Int. Journal of Computer Vision*, vol. 71, no. 2, pp. 183–196, 2007.

[13] K. Punithakumar, T. Kirubarajan, and A. Sinha, "A multiple model probability hypothesis density filter for tracking maneuvering targets," in *Proc. SPIE*, vol. 5428, pp. 113–121, 2004.

[14] K. Punithakumar, T. Kirubarajan, and A. Sinha, "Multiple-model probability hypothesis density filter for tracking maneuvering targets," *IEEE Trans. on AES*, vol. 44, no. 1, pp. 87–98, 2008.

[15] B. N. Vo and W. K. Ma, "Joint detection and tracking of multiple maneuvering targets in clutter using random sets," in *Proc. of 8th IEEE Conf. on Control, Automation, Robotics and Vision (ICARCV)*, vol. 2, (Kumming, China), pp. 1485–1490, December 2004.

[16] A. Pasha, B. N. Vo, H. D. Tuan, and W. K. Ma, "Closed form PHD filtering for linear jump Markov models," in *9th Int. Conf. on Information Fusion*, (Florence, Italy), pp. 1–8, July 2006.

[17] B. N. Vo, A. Pasha, and H. D. Tuan, "A Gaussian mixture PHD filter for nonlinear jump Markov models," in *45th IEEE Conf. on Decision and Control*, pp. 3162–3167, December 2006.

[18] A. Pasha, B. N. Vo, H. D. Tuan, and W. K. Ma, "A Gaussian mixture PHD filter for jump Markov system models," *IEEE Trans. on AES*, vol. 45, no. 3, pp. 919–936, 2009.

[19] Y. Bar-Shalom, X. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*. New York: Wiley, 2001.

[20] L. Lin, Y. Bar-Shalom, and T. Kirubarajan, "Data association combined with the probability hypothesis density filter for multitarget tracking," in *Proc. SPIE*, vol. 5428, pp. 464–475, 2004.

[21] K. Panta, B. N. Vo, S. Singh, and A. Doucet, "Probability hypothesis density filter versus multiple hypothesis tracking," in *Proc. SPIE*, vol. 5429, pp. 284–295, 2004.

[22] K. Panta, B. N. Vo, and S. Singh, "Novel data association schemes for the probability hypothesis density filter," *IEEE Trans. on AES*, vol. 43, no. 2, pp. 556–570, 2007.

[23] D. E. Clark and J. Bell, "Data association for the PHD filter," in *Proc. 2nd Int'l Conf. on Intelligent Sensor, Sensor Networks and Information Prcessing*, (Melbourne, Australia), pp. 217–222, December 2005.

[24] K. Panta, B. N. Vo, and S. Singh, "Improved probability hypothesis density (PHD) filter for multi-target tracking," in *Proc. 3rd Int'l Conf. on Intelligent Sensing and Information Processing*, (Bangalore, India), pp. 213–218, December 2005.

[25] D. E. Clark, K. Panta, and B. N. Vo, "The GM-PHD filter multiple target tracker," in *Proc. 9th Int'l Conf. on Information Fusion*, (Florence, Italy), July 2006.

[26] K. Panta, B. N. Vo, and D. E. Clark, "An efficient track management scheme for the Gaussian mixture probability hypothesis density tracker," in *Proc. IEEE Fourth Int'l Conf. on Intelligent Sensing and Information Processing*, (Bangalore, India), December 2006.

[27] K. Panta, D. E. Clark, and B. N. Vo, "Data association and track management for the Gaussian mixture probability hypothesis density filter," *IEEE Trans. on AES*, vol. 45, no. 3, pp. 1003–1016, 2009.

[28] D. E. Clark, B. Ristic, B. N. Vo, and B. T. Vo, "Bayesian multi-object filtering with amplitude feature likelihood for unknown object SNR," *IEEE Trans. on Signal Processing*, vol. 58, no. 1, pp. 26–37, 2010.

[29] K. Punithakumar, T. Kirubarajan, and A. Sinha, "A sequential Monte Carlo probability hypothesis density algorithm for multitarget track-before-detect," in *Proc. SPIE*, vol. 5913, 2005.

[30] J. Houssineau and D. Laneuville, "PHD filter with diffuse spatial prior on the birth process with applications to the GM-PHD filter," in *Proc. 13th Int'l Conf. on Information Fusion*, (Edinburgh, UK), July 2010.

[31] B. Ristic, D. E. Clark, and B. N. Vo, "Improved SMC implementations of the PHD filter," in *Proc. 13th Int'l Conf. on Information Fusion*, (Edinburgh, UK), July 2010.

[32] B. Ristic, D. E. Clark, and B. N. Vo, "Adaptive target birth intensity for PHD and CPHD filters," *IEEE Trans. on AES*, vol. 48, no. 2, pp. 1656–1668, 2012.

[33] S. Nagappa and D. E. Clark, "On the ordering of the sensors in the iterated-corrector probability hypothesis density (PHD) filter," in *Proc. SPIE*, vol. 8050, 2011.

[34] R. P. S. Mahler, "The multisensor PHD filter, I: General solution via multi-target calculus," in *Proc. SPIE*, vol. 7336, 2009.

[35] E. Delande, E. Duflos, D. Heurguier, and P. Vanheeghe, "Multi-target PHD filtering: proposition of extensions to the multi-sensor case," tech. rep., INRIA, 2010.

[36] N. J. Gordon, D. J. Salmond, and D. Fisher, "Bayesian target tracking after group pattern distortion," in *Proc. SPIE*, vol. 3163, pp. 238–248, 1997.

[37] D. J. Salmond and N. J. Gordon, "Group tracking with limited sensor resolution and finite field of view," in *Proc. SPIE*, vol. 4048, pp. 532–540, 2000.

[38] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Artech House, 2004.

[39] R. P. S. Mahler, "Detecting, tracking, and classifying group targets: a unified approach," in *Proc. SPIE*, vol. 4380, pp. 217–228, 2001.

[40] R. P. S. Mahler, "An extended first-order Bayes filter for force aggregation," in *Proc. SPIE*, vol. 4728, pp. 196–207, 2002.

[41] K. Gilholm, S. Godsill, S. Maskell, and D. Salmond, "Poisson models for extended targets and group tracking," in *SPIE Conf., Signal and Data Processing of Small Targets* (O. E. Drummond, ed.), vol. 5913, pp. 230–214, August 2005.

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

[43] W. Koch, "On Bayesian tracking of extended objects," in *IEEE Int. Conf. Multisensor Fusion and Integration for Intelligent Systems*, (Heidelberg, Germany), pp. 209–216, Sept. 2006.

[44] M. Feldmann and D. Fränken, "Tracking of extended objects and group targets using random matrices - a new approach," in *11th Int. Conf. on Information Fusion*, (Cologne, Germany), pp. 242–249, June/July 2008.

[45] M. Feldmann and D. Fränken, "Advances on tracking of extended objects and group targets using random matrices," in *12th Int. Conf. on Information Fusion*, (Seattle, WA, USA), pp. 1029–1036, July 2009.

[46] M. Baum, B. Noack, and U. D. Hanebeck, "Extended object and group tracking with elliptical random hypersurface models," in *13th Int. Conf. on Information Fusion*, (Edinburgh, UK), July 2010.

[47] M. Baum and U. D. Hanebeck, "Random hypersurface models for extended object tracking," in *9th IEEE Int. Symposium on Signal Processing and Information Technology*, (Ajman, UAE), 2009.

[48] M. Baum and U. D. Hanebeck, "Extended object tracking based on combined set-theoretic and stochastic fusion," in *12th Int. Conf. on Information Fusion*, (Seattle, WA, USA), pp. 1288–1295, July 2009.

[49] M. Baum, M. Feldmann, D. Fränken, and U. D. Hanebeck, "Extended object and group tracking: a comparison of random matrices and random hypersurface models," in *IEEE ISIF Workshop on Sensor Data Fusion: Trends, Solutions, Applications*, (Leipzig, Germany), Oct. 2010.

[50] N. Petrov, L. Mihayova, A. Gning, and D. Angelova, "A novel sequential Monte Carlo approach for extended object tracking based on border parameterisation," in *14th Int. Conf. on Information Fusion*, (Chicago, Illinois, USA), pp. 306–313, July 2011.

[51] C. Lundquist, U. Orguner, and F. Gustafsson, "Estimating polynomial structures from radar data," in *13th Int. Conf. on Information Fusion*, (Edinburgh, UK), July 2010.

[52] K. Granström, C. Lundquist, and U. Orguner, "Tracking rectangular and elliptical extended targets using laser measurements," in *14th Int. Conf. on Information Fusion*, (Chicago, Illinois, USA), pp. 592–599, July 2011.

[53] K. Granström, C. Lundquist, and U. Orguner, "Estimating the shape of targets with a PHD filter," in *14th Int. Conf. on Information Fusion*, (Chicago, Illinois, USA), pp. 49–56, July 2011.

[54] R. Kalman, "A new approach to linear filtering and prediction problemns," *Trans. of the AMSE-Journal of Basic Engineering*, vol. 82 (Series D), pp. 35–45, 1960.

[55] Y. C. Ho and R. C. K. Lee, "A Bayesian approach to problems in stochastic estimation and control," *IEEE Transactions on Automatic Control*, vol. AC-9, pp. 333–339, 1964.

[56] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Information and System Sciences, Prentice-Hall, 1979.

[57] S. Julier, J. Uhlmann, and H. F. Durrant-White, "A new method for nonlinear transformation of means and covariances in filters and estimators," *IEEE Trans. on Automatic Control*, vol. 45, no. 3, pp. 477–482, 2000.

[58] H. W. Sorenson and D. L. Alspach, "Recursive Bayesian estimation using Gaussian sums," *Automatica*, vol. 7, pp. 465–479, 1971.

[59] D. L. Alspach and H. W. Sorenson, "Nonlinear Bayesian estimation using Gaussian sum approximations," *IEEE Trans. on Automatic Control*, vol. 17, no. 4, pp. 439–448, 1972.

[60] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statistics and Computing*, vol. 10, pp. 197–208, 2000.

[61] R. A. Singer and J. J. Stein, "An optimal tracking filter for processing sensor data of imprecisely determined origin in surveillance systems," in *IEEE Conf. on Decision and Control*, (Miami Beach, FL), pp. 171–175, Dec 1971.

[62] R. A. Singer, R. G. Sea, and K. B. Housewright, "Derivation and evaluation of improved tracking filters for use in dense multi-target environments," *IEEE Trans. on Information Theory*, vol. 20, pp. 423–432, 1974.

[63] A. J. Jaffer and Y. Bar-Shalom, "On optimal tracking in multiple target environments," in *3rd Symposium on Non-Linear Estimation Theory and Its Applications*, (San Diego, CA), pp. 112–117, Sept 11–13.

[64] Y. Bar-Shalom and T. E. Fortmann, *Tracking and Data Association*, vol. 179 of *Mathematics in Science and Engineering*. Academic Press, Inc., 1988.

[65] S. S. Blackman, *Multiple-Target Tracking with Radar Applications*. Artech House, 1986.

[66] D. B. Reid, "An algorithm for tracking multiple targets," *IEEE Trans. on Automatic Control*, vol. 20, pp. 843–854, 1979.

[67] Y. Bar-Shalom and E. Tse, "Tracking in a cluttered environment with probabilistic data association," *Automatica*, vol. 11, pp. 451–460, 1975.

[68] G. Mathéron, *Random Sets and Integral Geometry*. J. Wiley, 1975.

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

[70] D. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes, Volume 1: Elementary Theory and Methods*. Springer, second ed., 2003.

[71] B. D. Ripley, "Locally finite random sets: foundations for point process theory," *Annals of Probability*, vol. 4, no. 6, pp. 983–994, 1976.

[72] H. Sidenbladh and S. L. Wirkander, "Tracking random sets of vehicles in terrain," in *Proc. of 2003 Conference on Computer Vision and Pattern Recognition*, p. 98, 2003.

[73] W. K. Ma, B. N. Vo, S. S. Singh, and A. Baddeley, "Tracking an unknown time-varying number of speakers using TDOA measurements: a random finite set approach," *IEEE Trans. on Signal Processing*, vol. 54, no. 9, pp. 3291–3304, 2006.

[74] M. Vihola, "Random set particle filter for bearings only multitarget tracking," in *Proc. SPIE*, vol. 5809, pp. 301–312, 2005.

[75] M. Vihola, "Rao-Blackwellised particle filtering in random set multitarget tracking," *IEEE Trans. on AES*, vol. 43, no. 2, pp. 689–705, 2007.

[76] B. N. Vo, S. Singh, and A. Doucet, "Sequential Monte Carlo methods for multi-target filtering with random sets," *IEEE Trans. on AES*, vol. 41, no. 4, pp. 1224–1245, 2005.

[77] B. N. Vo, B. T. Vo, N. T. Pham, and D. Suter, "Bayesian multi-object estimation from image observations," in *12th Int. Conf. on Information Fusion*, pp. 890–898, 2009.

[78] B. N. Vo, B. T. Vo, N. T. Pham, and D. Suter, "Joint detection and estimation of multiple objects from image observations," *IEEE Trans. on Signal Processing*, vol. 58, no. 10, pp. 5129–5241, 2010.

[79] J. Wong, B. T. Vo, B. N. Vo, and R. Hoseinnezhad, "Multi-Bernoulli based track-before-detect with road constraints," in *15th Int'l Conf. on Information Fusion*, (Singapore), pp. 840–846, July 2012.

[80] B. T. Vo, B. N. Vo, R. Hoseinnezhad, and R. P. S. Mahler, "Multi-Bernoulli filtering with unknown clutter intensity and sensor field-of-view," in *Proc. of Conf. on Information Sciences and Systems*, (Baltimore, USA), 2011.

[81] B. T. Vo, B. N. Vo, and A. Cantoni, "On multi-Bernoulli approximations to the Bayes multi-target filter," in *Proc. Int'l Colloquium on Information Fusion*, (Xian, China), 2007.

[82] B. T. Vo, B. N. Vo, and A. Cantoni, "The cardinality balanced multi-target multi-Bernoulli filter and its implementations," *IEEE Trans. on Signal Processing*, vol. 57, no. 2, pp. 409–423, 2009.

[83] R. Hoseinnezhad, B. N. Vo, B. T. Vo, and D. Suter, "Bayesian integration of audio and visual information for multi-target tracking using a CB-MeMBer filter," in *IEEE. Int'l Conf. on Acoustic, Speech and Signal Processing*, (Prague, Czech), pp. 2300–2303, 2011.

[84] R. P. S. Mahler, "PHD filters of higher order in target number," *IEEE Trans. on AES*, vol. 43, no. 4, pp. 1523–1543, 2007.

[85] H. Sidenbladh, "Multi-target particle filtering for the probability hypothesis density," in *6th Int. Conf. on Information Fusion*, (Cairns, Australia), pp. 800–806, July 2003.

[86] T. Zajic and R. P. S. Mahler, "A particle-systems implementation of the PHD multitarget tracking filter," in *Proc. SPIE*, vol. 5096, pp. 291–299, 2003.

[87] B. N. Vo, S. Singh, and A. Doucet, "Sequential Monte Carlo implementations of the PHD filter for multi-target tracking," in *6th Int. Conf. on Information Fusion*, (Cairns, Australia), pp. 792–799, July 2003.

[88] K. Panta and B. N. Vo, "Convolution kernels based sequential Monte Carlo implementation of the probability hypothesis density filter," in *Proc. of Int. Conf. Information, Decision and Control*, (Adelaide, Australia), 2007.

[89] B. N. Vo and W. K. Ma, "A closed-form solution for the probability hypothesis density filter," in *7th Int. Conf. on Information Fusion*, (Philadelhia, USA), pp. 856–863, July 2005.

[90] B. N. Vo and W. K. Ma, "The Gaussian mixture probability hypothesis density filter," *IEEE Trans. on Signal Processing*, vol. 54, no. 11, pp. 4091–4104, 2006.

[91] H. Zhang, Z. Jing, and S. Hu, "Bearings-only multi-target location based on Gaussian mixture PHD filter," in *10th Int. Conf. on Information Fusion*, (Québec, Canada), pp. 1–5, July 2007.

[92] D. E. Clark, B. T. Vo, B. N. Vo, and S. Godsill, "Gaussian mixture implementations of PHD filters for non-linear dynamic models," in *IET Seminar on Target Tracking and Data Fusion: Algorithms and Applications*, (Birmingham), pp. 21–28, April 2008.

[93] D. E. Clark and J. Bell, "Convergence results for the particle PHD filter," *IEEE Trans. on Signal Processing*, vol. 54, no. 7, pp. 2652–2661, 2006.

[94] A. Johansen, S. Singh, A. Doucet, and B. N. Vo, "Convergence of the sequential Monte Carlo implementations of the PHD filter," *Methodology and Computing in Applied Probability*, vol. 8, no. 2, pp. 265–291, 2006.

[95] D. E. Clark and B. N. Vo, "Convergence analysis of the Gaussian mixture PHD filter," *IEEE Trans. on Signal Processing*, vol. 55, no. 4, pp. 1204–1212, 2007.

[96] D. E. Clark, B. T. Vo, and B. N. Vo, "Gaussian particle implementations of probability hypothesis density filters," in *Aerospace Conference, 2007 IEEE*, pp. 1–11, March 2007.

[97] J. Kotecha and P. Djuric, "Gaussian particle filtering," *IEEE Trans. on Signal Processing*, vol. 51, no. 10, pp. 2592–2602, 2003.

[98] J. Kotecha and P. Djuric, "Gaussian sum particle filtering," *IEEE Trans. on Signal Processing*, vol. 51, no. 10, pp. 2603–2613, 2003.

[99] B. T. Vo, B. N. Vo, and A. Cantoni, "Analytic implementations of the cardinalized probability hypothesis density filter," *IEEE Trans. on Signal Processing*, vol. 55, no. 7, pp. 3553–3567, 2007.

[100] R. P. S. Mahler, "PHD filters for nonstandard targets, I: Extended targets," in *12th Int. Conf. on Information Fusion*, (Seattle, WA, USA), pp. 915–921, July 2009.

[101] K. Granström, C. Lundquist, and U. Orguner, "A Gaussian mixture PHD filter for extended target tracking," in *13th Int. Conf. on Information Fusion*, (Edinburgh, UK), July 2010.

[102] U. Orguner, C. Lundquist, and K. Granström, "Extended target tracking with a cardinalized probability hypothesis density filter," in *14th Int. Conf. on Information Fusion*, (Chicago, Illinois, USA), pp. 65–72, July 2011.

[103] U. Orguner, "CPHD filter derivation for extended targets." arXiv: 1011.1512v2, 2010.

[104] U. Orguner, C. Lundquist, and K. Granström, "Extended target tracking with a cardinalized probability hypothesis density filter," in *14th Int. Conf. on Information Fusion*, (Chicago, Illinois, USA), pp. 65–72, July 2011.

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

[106] D. Stoyan, W. Kendall, and J. Mecke, *Stochastic Geometry and its Applications*. Wiley, second ed., 2008.

[107] J. E. Moyal, "The General Theory of Stochastic Population Processes," *Acta Mathematica*, vol. 108, pp. 1–31, 1962.

[108] J. Neyman and E. L. Scott, "Statistical Approach to Problems of Cosmology," *Journal of Royal Statistical Society*, vol. 20, no. 1, pp. 1–29, 1958.

[109] J. Neyman and E. L. Scott, "Processes of Clustering and Applications," in *Stochastic Point Processes: Statistical Analysis, Theory and Application* (P. A. W. Lewis, ed.), pp. 646–681, Wiley, 1972.

[110] T. E. Harris, *The Theory of Branching Processes*. Berlin: Springer, 1963.

[111] A. Swain and D. E. Clark, "First-moment filters for spatial independent cluster processes," in *Proc. SPIE*, vol. 7697, 2010.

[112] A. Swain and D. E. Clark, "Extended object filtering using spatial independent cluster processes," in *13th Int. Conf. on Information Fusion*, (Edinburgh, UK), July 2010.

[113] M. N. M. van Lieshout, *Markov point processes and their applications*. Imperial College Press, 2000.

[114] G. C. Rota, "The number of partitions of a set," *American Mathematical Monthly*, vol. 71, no. 5, pp. 498–504, 1964.

[115] S. P. Lloyd, "Least square quantization in PCM," *IEEE Trans. on Information Theory*, vol. 28, no. 2, pp. 129–137, 1982.

[116] A. Okabe, B. Boots, K. Sugihara, and S. N. Chiu, *Spatial Tessellations. Concept and Applications of Voronoi Diagrams*. Wiley, 2nd ed., 2000.

[117] A. Baddeley and J. Møller, "Nearest-neighbour Markov point processes and random finite sets," *International Statistical Review*, vol. 57, no. 2, pp. 89–121, 1989.

[118] M. Ester, H. P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *2nd Int. Conf. on Knowledge Discovery and Data Mining*, pp. 226–231, AAAI Press, 1996.

[119] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood estimation from incomplete data via the EM algorithm," *J. Royal Statistical Soc.*, vol. B, no. 39, pp. 1–38, 1977.

[120] R. A. REdner and H. F. Walker, "Mixture densities, maximum likelihood and the EM algorithm," *SIAM REv.*, vol. 26, no. 2, pp. 195–239, 1984.

[121] D. J. Salmond, "Mixture reduction algorithms for point and extended object tracking in clutter," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 45, no. 2, pp. 667–686, 2009.

[122] G. Gennari and G. D. Hager, "Probabilistic data association methods in visual tracking of groups," *IEEE Computer Science Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 876–881, 2004.

[123] B. Ristic and J. Sherrah, "Particle filter for sequential detection and tracking of an extended object in clutter," in *20th European Signal Processing Conference (EUSIPCO)*, (Bucharest, Romania), pp. 1199–1203, August 27–31 2012.

[124] B. Ristic and J. Sherrah, "Bernoulli filter for joint detection and tracking of an extended object in clutter," *Radar, Sonar and Navigation, IET*, vol. 7, pp. 26–35, Jan 2013.

[125] P. C. Mahalanobis, "On the generalised distance in statistics," *Proc. of the Nat. Institute of Sciences of India*, vol. 2, no. 1, pp. 49–55, 1936.

[126] K. T. Abou-Moustafa, F. D. L. Torre, and F. P. Ferrie, "Designing a metric for the difference between Gaussian densities," in *Proc. of an Int. Symposium on the Occasion of the 25th Anniversary of the McGill University Centre for Intelligent Machines, Brain, Body and Machine*, vol. 83 of *Advances in Intelligent and Soft Computing*, pp. 57–70, 2012.

[127] K. Granström and U. Orguner, "A PHD filter for tracking multiple extended targets using random matrices," *IEEE Trans. on Signal Processing*, vol. 60, no. 11.

[128] C. Lundquist, K. Granström, and U. Orguner, "An extended target CPHD filter and a Gamma Gaussian inverse Wishart implementation," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 3, pp. 472–483.

[129] C. S. Lee, D. E. Clark, and J. Salvi, "SLAM with single cluster PHD filters," in *IEEE Int. Conf. on Robotics and Automation*, pp. 2096–2101, 14–18 May 2012.

[130] C. S. Lee, D. E. Clark, and S. Nagappa, "Single-cluster PHD filtering and smoothing for SLAM applications," in *Stochastic geometry workshop, IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2012.

[131] C. S. Lee, D. E. Clark, and J. Salvi, "SLAM with dynamic targets via single cluster PHD filtering," *IEEE Journal of Selected Topics in Signal Processing: Special Issue on Multi-Target Tracking*, 2013. accepted for publication.

[132] B. Ristic and D. E. Clark, "Particle filter for joint estimation of multi-object dynamic state and multi-sensor bias," in *IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, pp. 3877–3880, 25–30 March 2012.

[133] B. Ristic, D. E. Clark, and N. Gordon, "Calibration of tracking systems using detections from non-cooperative targets," in *IEEE Workshop on Sensor Data Fusion: Trends, Solutions, Applications*, pp. 25–30, 4–6 September 2012.

[134] B. Ristic, D. E. Clark, and N. Gordon, "Calibration of multi-target tracking algorithms using non-cooperative targets," *IEEE Journal for Selected Topics in Signal Processing: Special Issue on Multi-Target Tracking*, 2013. accepted for publication.

[135] D. E. Clark and R. P. S. Mahler, "Generalized PHD filters via a general chain rule," in *15th Int. Conf. on Information Fusion*, (Singapore), pp. 157–164, July 2012.

[136] D. E. Clark and J. Houssineau, "Faà di Bruno's formula for variational calculus," *Elsevier Nonlinear Analysis: Theory, Methods and Applications*, 2013. accepted for publication.

[137] D. E. Clark and J. Houssineau, "Faà di Bruno's formula for Gâteaux differentials and interacting stochastic population processes." arXiv: 1202.0264, 2012.

[138] R. P. S. Mahler, B.-T. Vo, and B.-N. Vo, "CPHD filtering with unknown cluter rate and detection profile," *IEEE Trans. on Signal Processing*, vol. 59, no. 8, pp. 3497–3513.

[139] F. Zhang, *The Schur complement and its applications*. Springer, 2005.

[140] S. Roweis, "Gaussian identities," tech. rep., NYU Computer Science Depart., July 1999.