Multi-target Multi-Bernoulli Tracking and Joint Multi-target Estimator

MULTI-TARGET MULTI-BERNOULLI TRACKING AND JOINT MULTI-TARGET ESTIMATOR

BY

ERKAN BASER, B.Sc., M.Sc.

A THESIS
SUBMITTED TO THE DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING
AND THE SCHOOL OF GRADUATE STUDIES
OF MCMASTER UNIVERSITY
IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Doctor of Philosophy (2016) McMaster University
(Electrical & Computer Engineering) Hamilton, Ontario, Canada

TITLE:	Multi-target Multi-Bernoulli Tracking and Joint Multi-target Estimator
AUTHOR:	Erkan Baser B.Sc., (Electronics Engineering), Hacettepe University, Ankara, Turkey M.Sc., (Electrical & Electronics Engineering), Middle East Technical University University, Ankara, Turkey
SUPERVISOR:	Prof. Thia Kirubarajan
NUMBER OF PAGES:	xvi, 173

To my loved ones

Abstract

This dissertation concerns with the formulation of an improved multi-target multi-Bernoulli (MeMBer) filter and the use of the joint multi-target (JoM) estimator in an effective and efficient manner for a specific implementation of MeMBer filters. After reviewing random finite set (RFS) formalism for multi-target tracking problems and the related Bayes estimators the major contributions of this dissertation are explained in detail.

The second chapter of this dissertation is dedicated to the analysis of the relationship between the multi-Bernoulli RFS distribution and the MeMBer corrector. This analysis leads to the formulation of an unbiased MeMBer filter without making any limiting assumption. Hence, as opposed to the popular cardinality balanced multi-target multi-Bernoulli (CBMeMBer) filter, the proposed MeMBer filter can be employed under the cases when sensor detection probability is moderate to low. Furthermore, a statistical refinement process is introduced to improve the stability of the estimated cardinality of targets obtained from the proposed MeMBer filter. The results from simulations demonstrate the effectiveness of the improved MeMBer filter.

In Chapters III and IV, the Bayesian optimal estimators proposed for the RFS based multi-target tracking filters are examined in detail. First, an optimal solution to the unknown constant in the definition of the JoM estimator is determined by solving a multi-objective optimization problem. Thus, the JoM estimator can be implemented for tracking of a Bernoulli target using the optimal joint target detection and tracking (JoTT) filter. The results from simulations confirm assertions about its performance obtained by theoretical analysis in the literature. Finally, in the third chapter of this dissertation, the proposed JoM estimator is reformulated for RFS multi-Bernoulli distributions. Hence, an effective and efficient implementation of the JoM estimator is proposed for the Gaussian mixture implementations of the MeMBer filters. Simulation results demonstrate the robustness of the proposed JoM estimator under low-observable conditions.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my advisor Dr. Thia Kirubarajan for his support, patience and supervision. His guidance always helped me in all the time of my research. Without his precious support it would not be possible to conduct this research.

Very special thanks to Dr. Murat Efe who helped me a lot to pursue my PhD degree. I will never forget his help and support.

I would also like to thank Dr. Ratnasingham Tharmarasa for taking his time to answer my questions. His insightful discussions and suggestions during the course of my studies are immensely appreciated.

Thanks to my friends in estimation, tracking and fusion research laboratory. In particular, I would like to thank Dr. Ehsan Taghavi for his help in writing my thesis.

Last but not the least, I would like to thank my family: my parents and to my sister for supporting me spiritually throughout writing this thesis and my life in general.

Abbreviations

AEP	Asymptotic Equipartition Property
AR	Autoregressive
CBMeMBer	Cardinality Balanced Multi-Target Multi-Bernoulli
CPHD	Cardinalized Probability Hypothesis Density
EAP	Expected A Posteriori
FISST	Finite Set Statistics
GLMB	Generalized Labeled Multi-Bernoulli
GM	Gaussian Mixture
GMAP	Global Maximum A Posteriori
GNN	Global Nearest Neighbor
IID	Independent Identically Distributed
IMeMBer	Improved Multi-Target Multi-Bernoulli
IPDA	Integrated Probabilistic Data Association
JoM	Joint Multi-Target
JoTT	Joint Target Detection and Tracking
JPDA	Joint Probabilistic Data Association
KL	Kullback-Leibler
KKT	Karush-Kuhn-Tucker

LMB	Labeled Multi-Bernoulli
MaM	Marginal Multi-Target
MAP	Maximum A Posteriori
MeMBer	Multi-Target Multi-Bernoulli
MHT	Multiple Hypothesis Tracking
OSPA	Optimal Subpattern Assignment
p.g.f.	Probability Generating Function
p.g.fl.	Probability Generating Functional
PHD	Probability Hypothesis Density
RFS	Random Finite Set
SMC	Sequential Monte Carlo
SNR	Signal-to-Noise Ratio
SQP	Sequential Quadratic Programming
SVD	Singular Value Decomposition
WSS	Wide Sense Stationary

Contents

List of Tables

List of Figures

Declaration of Academic Achievement

This research presents analytical and computational work carried out solely by Erkan Baser, herein referred to as "the author", with advice and guidance provided by the academic supervisor Prof. T. Kirubarajan. Information that is presented from outside sources which has been used towards analysis or discussion, has been cited when appropriate, all other materials are the sole work of the author.

Chapter 1

Introduction

1.1 Multi-target Tracking

Multi-target tracking is a joint estimation problem of unknown number of targets and their states. The target number varies due to appearance and disappearance of targets while their states evolve over time according to target dynamics. In addition, sensor imperfections (e.g., noise and missed detections), and clutter introduce measurement uncertainties which make the estimation problem too difficult to solve directly using the Bayesian filtering techniques, e.g., Kalman filters and particle filters [1, 14].

The traditional approaches solve multi-target tracking problems in two steps: first, the measurement uncertainty is addressed using a data association method. Then, the Bayesian filtering techniques are used to update individual target states using the associated measurements. Several data association methods have been proposed in the literature. They range from the simple global nearest neighbor (GNN) data association to the more complex ones such as joint probabilistic data association (JPDA) and the multiple hypothesis tracking (MHT) [2, 3]. In the followings, some fundamental data association methods will be briefly discussed.

The GNN method assigns a unique measurement for each target by minimizing a total cost function defined over all combinations of possible assignments of measurements to targets. The possible assignments are determined under the constraint that a measurement can be associated with at most one target and vice versa. Even though this data association method is easy to implement, it is not robust since the selected joint assignment is assumed to be correct. Therefore, tracks can be lost in scenarios with high clutter rate or low probability of detection [2, 3]

The JPDA method uses all possible joint assignments with their association probabilities. That is, instead of a hard joint assignment, soft joint assignments of measurements to targets are performed using their association probabilities. The computational complexity of the JPDA increases exponentially with the number of targets and the number of measurements. Therefore, some approximations are utilized to propose its efficient but suboptimal implementations. [15, 16, 13]. The standard JPDA assumes that the number of targets is fixed and known. This limiting assumption is relaxed by incorporating target existence model into the JPDA framework [11].

The MHT keeps a set of tracks hypothesized for all targets. Each hypothesis has a posterior probability computed using Bayes rule. The basic idea in the MHT is to progressively remove the data-association uncertainty with the help of new measurements. Therefore, these hypotheses are propagated in time and used to generate new hypotheses. However, the number of hypotheses grows exponentially with time. In order to deal with this drawback on its implementation, hypothesis pruning, hypothesis merging and gating techniques are utilized [3]. Consequently, a small set of hypotheses with high probabilities are maintained for each target.

1.2 Random Finite Set Based Multi-target Filters

1.2.1 Random Finite Set

The main challenge with the traditional multi-target tracking filters surfaces while associating measurements to targets. In the random finite set (RFS) framework, the modeling of multi-target tracking systems is completely free of explicit data associations. This is achieved by representing collections of target states and measurements as finite set-valued random variables [8]. That is, an RFS X is a random variable that takes values as finite subsets of a single-object state space, e.g., $X \subseteq \mathbb{R}^n$ where $\mathbb{R}^n$ is Euclidean n-space. Thus, it consists of a random number of unordered elements whose states are also random vectors. Hence, it can be completely characterized by two distribution functions: *i)* a probability mass function of its cardinality variable $|X|$, and *ii)* a symmetric joint probability distribution function of its elements for a given cardinality. For any closed region $S \subseteq \mathbb{R}^n$, the multi-target probability density function $f(X)$ is defined such that [4, 8]

$$\Pr(X \subseteq S) = \int_S f(X)\,\delta X, \tag{1.1}$$

where the set integral of $f(X)$ is given by

$$\int_S f(X)\,\delta X = f(\emptyset) + \sum_{i=1}^{\infty} \frac{1}{i!} \int_{\underbrace{S\ \times \ldots \times\ S}_{i\ times}} f(\{x_1, \ldots, x_i\})\, dx_1 \ldots dx_i, \tag{1.2}$$

where the infinitesimal volume $dx_1...dx_i$ has units of υ^i if υ is the unit of hyper-volume on S. Hence, $f(\{x_1,...,x_i\})$ must have units of υ^{-i} whereas $f(\emptyset)$ is a unitless probability. Each term of the summation in (1.2) is the probability that X is contained in the region S for a given cardinality. Then, the cardinality distribution of X is computed as

$$p_{|X|}(i) = \frac{1}{i!}\int_{|X|=i} f(\{x_1,...,x_i\})\, dx_1...dx_i, \tag{1.3}$$

where the set integral is taken over all S for $|X| = i$. A complete statistical description of X is provided by the multi-target probability density function $f(X)$. Therefore, for any $X = \{x_1,...,x_i\}$ this density is defined as follows [8]:

$$f(\{x_1,...,x_i\}) = i!\, p_{|X|}(i)\, f(x_1,...,x_i), \tag{1.4}$$

where $f(x_1,...,x_i)$ is a symmetric joint probability distribution for all possible permutations of random vectors $(x_1,...,x_i)$.

The probability generating functional (p.g.fl.) is another statistical descriptor of an RFS. It provides a convenient way to obtain statistics of an RFS like the generating functions used in probability theory, e.g., the characteristic function and the probability generating function (p.g.f.). The p.g.fl. of an RFS X is defined by [8]

$$G[h] = \int h^X f(X)\, \delta X, \tag{1.5}$$

where h is any test function on X such that $0 \leq h(x) \leq 1$, and

$$h^X = \begin{cases} 1 \quad \text{if } X = \emptyset, \\ \prod_{x\in X} h(x) \quad \text{otherwise.} \end{cases} \tag{1.6}$$

Substituting the constant function $h(x) = \chi \ \forall x \in X$ into the p.g.fl. in (1.5) yields the p.g.f. of $|X|$ [8]:

$$\begin{aligned} G_{|X|}(\chi) &= G[h]\big|_{h(x)=\chi}, \\ &= \sum_{i=0}^{\infty} \chi^i p_{|X|}(i). \end{aligned} \tag{1.7}$$

The p.g.f. completely characterize $|X|$, i.e., $p_{|X|}(i) = \frac{1}{i!} G^{(i)}_{|X|}(0)$ where $G^{(i)}_{|X|}(\chi)$ is the ith derivative of $G_{|X|}(\chi)$. In addition, the first two moments of $|X|$ are given by $E[|X|] = G^{(1)}_{|X|}(1)$ and $\sigma^2_{|X|} = G^{(2)}_{|X|}(1) - \left(G^{(1)}_{|X|}(1)\right)^2 + G^{(1)}_{|X|}(1)$ [7, 8].

The important RFSs used in multi-target tracking are *i)* Poisson RFS, *ii)* independent identically distributed (IID) cluster RFS, *iii)* Bernoulli RFS, and *iv)* Multi-Bernoulli RFS [8]. Poisson RFS is a special case of the IID cluster RFS when the cardinality distribution is Poisson [7]. Bernoulli RFS and its multi-target extension, i.e., multi-Bernoulli RFS are the parametric RFSs. These two RFSs are the foundation of the multi-target filters studied in this dissertation. Therefore, a brief review of Bernoulli and multi-Bernoulli RFSs are provided in the following paragraphs.

Bernoulli

A Bernoulli RFS is an empty set to refer a nonexistent target with probability of $1-q$ or is a singleton set whose random element x is statistically characterized by a spatial probability density function $f(x)$ on a single target state space. Thus, the probability density function of a Bernoulli RFS X with the parameter pair $\{q, f\}$ is given by

$$f(X) = \begin{cases} 1-q & \text{if } X = \emptyset, \\ qf(x) & \text{if } X = \{x\}. \end{cases} \tag{1.8}$$

The cardinality of a Bernoulli RFS is characterized by a Bernoulli distribution with the parameter q denoting the existence probability of a single target. This can be verified by substituting (1.8) into (1.3). In addition, the p.g.fl. of the Bernoulli RFS X is given by $G[h] = 1 - q + qf[h]$.

Multi-Bernoulli

A multi-Bernoulli RFS is the union of independent Bernoulli RFSs $X^{(i)}$ with the parameter pair $\{q^{(i)}, f^{(i)}\}$, i.e., $X = \bigcup_{i=1}^{I} X^{(i)}$. Therefore, it is described by the ensemble of the parameter pairs $\{(q^{(i)}, f^{(i)})\}_{i=1}^{I}$. Suppose that there exist M targets out of I constituent Bernoulli RFSs. Then, the probability density function of the multi-Bernoulli RFS is given by

$$f(X) = \begin{cases} \left(1 - q^{(1)}\right) ... \left(1 - q^{(I)}\right) & \text{if } X = \emptyset, \\ \sum\limits_{\beta} \prod\limits_{j=1}^{M} \Theta_\beta f^{\beta(j)}(x_j) & \text{if } X = \{x_1, \ldots, x_M\}, \end{cases} \tag{1.9}$$

where

$$\Theta_\beta = \left(1 - q^{(1)}\right) ... \left(1 - q^{(I)}\right) \frac{q^{\beta(1)}}{\left(1 - q^{\beta(1)}\right)} \cdots \frac{q^{\beta(M)}}{\left(1 - q^{\beta(M)}\right)}, \tag{1.10}$$

and the sum is taken over all permutations of the joint association hypotheses of those existent Bernoulli targets that are represented by a one-to-one function $\beta : \{1, ..., M\} \to \{1, ..., I\}$ for $M \le I$ [8]. Using (1.3) the corresponding cardinality distribution of the multi-Bernoulli RFS is obtained as

$$p_{|X|}(M) = \frac{1}{M!} \sum_{\beta} \Theta_\beta. \tag{1.11}$$

Using the independence of Bernoulli RFSs $X^{(i)}$, the p.g.fl. of the multi-Bernoulli RFS

X is given by

$$G[h] = G_1[h] \ldots . G_I[h], \tag{1.12}$$

where $G_i[h] = 1 - q^{(i)} + q^{(i)} f^{(i)}[h]$ for $i = 1, ..., I$ [8].

1.2.2 Multi-Bernoulli Approximations

The Bayesian multi-target filtering is a recursive process performed in two steps [4, 8]:

Prediction: Let $\varphi_{k|k-1}(X|X')$ be the multi-target transition density, which characterizes target motions, target births and target deaths. Then, the posterior multi-target probability density $f_{k-1|k-1}\left(X|Z^{(k-1)}\right)$ at time $k-1$ is propagated according to

$$f_{k|k-1}\left(X|Z^{(k-1)}\right) = \int f_{k-1|k-1}\left(X'|Z^{(k-1)}\right) \varphi_{k|k-1}(X|X')\, \delta X. \tag{1.13}$$

Update: Let $g_k(Z_k|X)$ be the multi-target likelihood of observing the measurement RFS Z_k in the presence of noise, false alarms and missed detections. Then, the predicted multi-target probability density is updated by the Bayes rule:

$$f_{k|k}\left(X|Z^{(k)}\right) = \frac{g_k(Z_k|X)\, f_{k|k}\left(X|Z^{(k-1)}\right)}{\int g_k(Z_k|X)\, f_{k|k}(X|Z^{(k-1)})\, \delta X}. \tag{1.14}$$

Due to the combinatorial nature of $\varphi_{k|k-1}(X|X')$ and $g_k(Z_k|X)$ several integrals must be evaluated over high dimensional product spaces in (1.13) and (1.14). Therefore, the Bayes multi-target filter is computationally intractable [8]. However, if each target is modeled by a Bernoulli RFS, a parametric approximation to the Bayes multi-target filter can be obtained by propagating and updating their Bernoulli parameters, i.e., the multi-Bernoulli RFS X.

The multi-target multi-Bernoulli (MeMBer) filter provides a parametric approximation to the Bayes multi-target filter [8]. However, the posterior cardinality estimate from the MeMBer filter is positively biased due to Bernoulli RFS approximation followed for measurement-updated targets. The cardinality balanced multi-target multi-Bernoulli (CBMeMBer) filter removes the bias [17]. However, the CBMeMBer filter requires a limiting assumption on probability of detection.

1.3 Bayes Estimators for Multi-target Filters

The conventional state estimators, i.e., the expected a posteriori (EAP) and the maximum a posteriori (MAP) estimators cannot be used to estimate RFS [8, 5]. There are two fundamental reasons why these two estimators are unavailable for RFS [8, 6]. First, as indicated above, the multi-target density $f(\{x_1, ...x_i\})$ has units of υ^{-i} where υ is the unit of volume on a single target space, e.g., meter in $\mathbb{R}$, and i is a natural number including zero. Since it is impossible to compare $f(\{x_1, ...x_i\})$ for different values of $|X| = i$, the MAP estimate of X, i.e., $X^{(MAP)}$ cannot be determined. Second, the RFS is a finite-valued random set. Since addition and subtraction operations are not defined on sets properly, the EAP estimate of X, i.e., $X^{(EAP)}$ cannot be determined as well.

Instead of the EAP and MAP estimators, two Bayes multi-target estimators are available to estimate RFS from multi-target filters [8]. These two estimators can be interpreted as MAP-like estimators. This is because they both determine the MAP estimates of individual target states from a given multi-target probability density. The marginal multi-target (MaM) estimator determines this multi-target density by computing the MAP estimate of cardinality, i.e.,

$$\hat{X}^{MaM} = \arg \sup_{x_1,...,x_{\hat{i}^{MAP}}} f_{k|k}\left(\{x_1, ..., x_{\hat{i}^{MAP}}\} \left| Z^{(k)} \right.\right), \tag{1.15}$$

where

$$\hat{i}^{MAP} \triangleq \arg \sup_i p_{|X|}(i), \tag{1.16}$$

On the other hand, the joint multi-target (JoM) estimator determines the MAP estimates of cardinality and target states simultaneously [8], i.e.,

$$\hat{X}^{JoM} = \arg \sup_X f_{k|k}\left(X \left| Z^{(k)} \right.\right) \frac{\varepsilon^{|X|}}{|X|!}, \tag{1.17}$$

where the parameter ε is an unknown constant and must satisfy that $f(\{x_1, ..., x_i\})\, \varepsilon^i \leq 1$ for all $|X| = i$. In addition, there is a trade-off in the selection of ε between the accuracy of multi-target state estimates and the speed of convergence to the true multi-target state [8, 5].

The MaM and JoM estimators are both optimal in the sense that they minimize their cost functions. However, the JoM estimator is more robust than the MaM estimator because of its well-designed cost function [4].

1.4 Pareto Optimization

The multi-objective optimization problems consists of multiple conflicting objective functions that need to be satisfied simultaneously [12, 9]. For example, the problem of determining the most effective portfolio is a multi-objective problem. There are many asset classes like stock, bond, cash and real-estate to invest in with different risk

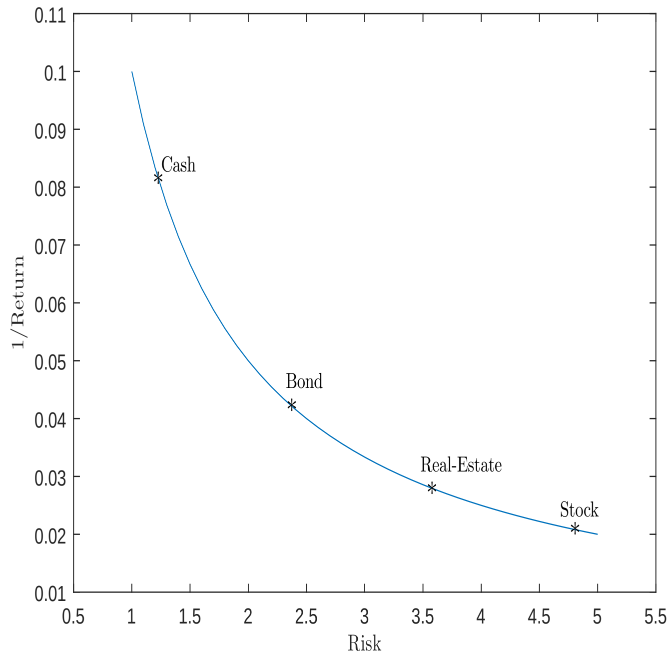

Figure 1.1: Illustration of Pareto front for the bi-objective portfolio problem.

factors. On average, a more risky asset class would have higher returns. Therefore, there is no solution that maximizes return while minimizing risk. Instead, there are many optimal solutions that do not dominate each other [12, 9]. These solutions are called Pareto optimal and form a set of Pareto optimal objective vectors (also known as Pareto front). Consider the dynamics of the economic system observed in a country; the risk-versus-return analysis is illustrated in Fig. 1.1. The curve in Fig. 1.1 is the Pareto front where any improvement to risk incurs a detriment to return.

From optimization point of view, each point on Pareto front is an acceptable solution. However, as in the case of single objective optimization problems, a final solution shall be selected among these solutions. This can be achieved using decision maker's preference on each objective function that specifies the definition of the optimum solution [12, 10, 9]. For example, the decision maker can determine the relative importance of each objective functions in multi-objective optimization by ranking them.

1.5 Theme and Objectives of Dissertation

In compliance with the terms and regulations of McMaster University, this dissertation has been written in sandwich thesis format by assembling three journal articles. These articles are results of the independent research performed by the author of this thesis.

The articles in the dissertation are focused on the multi-Bernoulli process and the use of Bayesian estimators for MeMBer filters. In this scope, the relationship between the multi-Bernoulli RFS distribution and the MeMBer corrector are analyzed in Paper I. This analysis is utilized to formulate an unbiased MeMBer filter like the CBMeMBer filter but without making any limiting assumptions. In addition to addressing an open issue in the CBMeMBer filter, the performance of the proposed MeMBer filter on cardinality estimate of targets was improved. Then, the Bayesian estimators proposed for the RFS based multi-target tracking filters are studied in Paper II. The implementation of the robust JoM estimator was infeasible due to its computational complexity and an unknown constant in its definition. In order to determine an optimal solution to that constant, some fundamental concepts in information theory and multi-objective Pareto optimization are studied. Thus, the JoM estimator is implemented for tracking of a Bernoulli target. The results from simulations confirm assertions about its performance obtained by theoretical analysis in the literature. Finally, in Paper III, the proposed JoM estimator is reformulated for RFS multi-Bernoulli distributions. Thus, the computationally complex JoM estimator can be used for the Gaussian mixture implementations of the MeMBer filters.

1.6 Summary of Enclosed Articles

The papers enclosed in this thesis are listed as follows:

1.6.1 Paper I (Chapter 2)

Erkan Baser, Thia Kirubarajan, Murat Efe, and Bhashyam Balaji

Improved MeMBer Filter with Modeling of Spurious Targets, *IET Radar, Sonar & Navigation*, no. 2, vol. 10, pp. 285–298, February 2016.

Preface: Positive bias is observed on cardinality estimates obtained from the original MeMBer filter. The CBMeMBer filter removes the bias under the assumption that the probability of detection is close to unity. This assumption restricts its use in multi-target tracking scenarios where sensor detection probability is moderate to small. In order to deal with this problem, this paper expresses the relationship between the posterior distribution obtained in the MeMBer data update step and the multi-Bernoulli RFS distribution. Then, this relationship was utilized to introduce spurious targets that account for the bias arising from multi-Bernoulli modeling of targets in the MeMBer data update process. The modeling of spurious targets removes the limiting assumption on the probability of detection. In addition, a statistical method was proposed to refine existence probabilities of Bernoulli targets in the light of measurements. Consequently, the proposed MeMBer filter outperforms the CBMeMBer filter in stability of cardinality estimates of targets. Finally, in addition to simulations, the strength and limitations of the IMeMBer filter was investigated by comparing it with optimal JoTT filter.

1.6.2 Paper II (Chapter 3)

Erkan Baser, Mike McDonald, Thia Kirubarajan, and Murat Efe

A Joint Multitarget Estimator for the Joint Target Detection and Tracking Filter, *IEEE Transactions on Signal Processing*, no. 15, vol. 63, pp. 3857 –3871, May 2015.

Preface: The JoM estimator is a robust optimal Bayesian estimator proposed for the RFS based multi-target tracking filters. However, it could not be put into use due to the unknown JoM estimation constant. This paper provides an optimal solution to that constant by solving a multi-objective optimization problem. The multi-objective optimization problem consists of two conflicting objectives. The first objective function is defined using some fundamental concepts in information theory and aims to maximize the entropy. On the other hand, the second one is obtained from the definition of the JoM estimator and aims to improve the accuracy of JoM estimates. The solution to these two conflicting objective functions is obtained as a Pareto optimal solution to their weighted sum. The weights are adjusted at each time step considering the performance of tracking filter. The track management performance of the proposed JoM estimator is compared with another optimal estimator called MaM estimator in simulations. The results provide a justification to the theoretical analysis performed about the performances of these two estimators in the literature.

1.6.3 Paper III (Chapter 4)

Erkan Baser, Thia Kirubarajan, Murat Efe, and Bhashyam Balaji

A Novel Joint Multitarget Estimator for Multi-Bernoulli Models, *IEEE Transactions* *on Signal Processing*, no. 19, vol. 64, pp. 5038–5051, June 2016.

Preface: The JoM estimator is computationally intractable for RFS multi-target distributions. Therefore, this paper extends the proposed JoM estimator to be used with RFS multi-Bernoulli distributions in an effective and efficient manner. For this purpose, two approximations are followed: first, the MAP estimates from an RFS multi-Bernoulli distribution are approximated for given cardinalities of targets. Then, in order to derive the multi-Bernoulli versions of the two conflicting objective functions in Paper II, the posterior probability densities in the GM form are approximated as Gaussian densities for each Bernoulli target. In addition, the computation of mixing weights for these two conflicting objectives are reformulated considering dynamics of multi-target tracking and characteristics of the JoM estimator. Thus, the Pareto optimal solution to the JoM estimation constant is determined according to the different localization conditions such as occlusion and target birth. Two simulations are provided, one for smaller and one for larger probabilities of detection. The results conclude that the JoM estimator provides better tracking performance than the MaM estimator when the uncertainty on cardinality estimates of targets is high, i.e., under low observable conditions.

Bibliography

[1] Y. Bar-Shalom, X.R. Li, and T. Kirubarajan. *Estimation with Applications to Tracking and Navigation.* J. Whiley and Sons, New York City, NY, 2001.

[2] Y. Bar-Shalom, P. K. Willett, and X. Tian. *Tracking and Data Fusion: A Handbook of Algorithms.* YBS Publishing, Storrs, CT, 2011.

[3] S. S. Blackman and R. Popoli. *Design and Analysis of Modern Tracking Systems.* Artech House, Boston,London, 1999.

[4] I. Goodman, R. Mahler, and H. Nguyen. *Mathematics of Data Fusion.* Kluwer Academic Publisher, Norwell, MA, 1997.

[5] R. Mahler. Why multi-source, multitarget data fusion is tricky. In *Proc. of IRIS National Symposium on Sensor and Data Fusion*, pages 135–153, May 1999.

[6] R. Mahler. 'Statistics 101' for multisensor, multitarget data fusion. *IEEE Aerosp. and Electron. Sys. Magazine Part 2: Tutorials*, 19(1):53–64, Jan. 2004.

[7] R. Mahler. PHD filters of higher order in target number. *IEEE Trans. Aerosp. Electron. Syst.*, 43(4):1523–1542, Oct. 2007.

[8] R. Mahler. *Statistical Multisource-Multitarget Information Fusion.* Artech House, Boston, MA, 2007.

[9] R. T. Marler and J. S. Arora. Survey of multi-objective optimization methods for engineering. *Structural and Multidisciplinary Optimization*, 26(6):369–395, Mar. 2004.

[10] R. T. Marler and J. S. Arora. The weighted sum method for multi-objective optimization. *Structural and Multidisciplinary Optimization*, 41(6):853–862, Dec. 2009.

[11] D. Musicki and R. Evans. Joint integrated probabilistic data association - JIPDA. In *Proc. of 5th Int'l Conf. on Information Fusion*, pages 1120–1125, Maryland, USA, Jul. 2002.

[12] P. Ngatchou, A. Zarei, and M. A. El-Sharkawi. Pareto multi objective optimization. In *Proc. of the 13th Int'l Conf. on Intell. Syst. Application to Power Systems*, pages 84–91, Va, USA, Nov. 2005.

[13] S. Oh and S. Sastry. A polynomial-time approximation algorithm for joint probabilistic data association. In *Proc. of American Control Conference*, pages 1283–1288, Portland, USA, Jul. 2005.

[14] B. Ristic, S. Arulampalam, and N. Gordon. *Beyond the Kalman Filter : particle filters for tracking applications.* Artech House, Boston, London, 2004.

[15] J. Roecker. Suboptimal joint probabilistic data association. *IEEE Trans. Aerosp. and Electron. Syst.*, 29(2):510–517, Apr. 1993.

[16] J. Roecker. A class of near optimal JPDA algorithms. *IEEE Trans. Aerosp. and Electron. Syst.*, 30(2):504–510, Apr. 1994.

[17] B.-T. Vo, B.-N. Vo, and A. Cantoni. The cardinality balanced multitarget multi-Bernoulli filter and its implementations. *IEEE Trans. Signal Process.*, 57(2):409–423, Feb. 2009.

The following chapter is a reproduction of a peer-reviewed article in the Institution of Engineering and Technology (IET) with permission:

Erkan Baser, Thia Kirubarajan, Murat Efe, and Bhashyam Balaji, Improved MeMBer Filter with Modeling of Spurious Targets, *IET Radar, Sonar & Navigation*, no. 2, vol.10, pp. 285–298, February 2016.

Chapter 2

Improved MeMBer Filter with Modeling of Spurious Targets

2.1 Abstract

The cardinality-balanced multi-target multi-Bernoulli (CBMeMBer) filter removes the positive bias from the data-updated cardinality estimate in the MeMBer filter. In this paper, the relationship between the MeMBer corrector and the multi-Bernoulli random finite set (RFS) distribution is analyzed. By utilizing this relationship, a filter that offers a new statistical framework for the MeMBer data update process is proposed. Thus, the multi-Bernoulli RFS distribution is extended to model spurious targets arising from targets under the legacy track set with high probabilities of existence. Unlike the CBMeMBer filter, the proposed filter removes the bias observed in the MeMBer filter by distinguishing spurious targets from actual targets, and while doing this, it does not make any limiting assumption on the probability of target detection. In addition, the modeling of spurious targets allows the refinement of the

existence probabilities of targets in light of measurements. As a result, the stability of the cardinality estimate is improved while removing the bias. The theoretical analysis performed on the joint detection and state estimation problem of a single target reveals the strengths and limitations of the proposed filter. In addition, numerical simulations are performed in a scenario involving targets with crossing trajectories to demonstrate the filter performance.

2.2 Introduction

The objective in multi-target tracking is to estimate both the time-varying number of targets and their random states from the measurements received in the presence of noise, false alarms, and missed detections. The random finite set (RFS) formalism facilitates unified and probabilistic modeling of all presumed uncertainties associated with multi-target tracking problems [12, 15]. Thus, the Bayesian framework for single target tracking can be systematically translated into its multi-target counterpart. For the resulting multi-target Bayes filter, the required Bayesian statistics are computed using the finite set statistics (FISST) [12, 15, 14]. However, the implementation of the multi-target Bayes filter using FISST is computationally intractable. Similar to the single target Bayes filter, the statistical moment-based approximations to the multi-target Bayes filter, known as the probability hypothesis density (PHD) and the cardinalized PHD (CPHD) filters were developed in [11, 13] and implemented in [29, 28, 32]. Recently, significant research has been devoted to improving the performance of the PHD/CPHD filters [23, 24, 3] and to dealing with the problems in their implementations such as extraction of state estimates and track labeling [10, 20, 7]. In addition, further improvements have enabled these two filters to operate when clutter and/or detection profile are not known a priori [17, 18, 6].

The multi-Bernoulli assumption on the RFS of targets represents each target as a randomly and independently switching on/off dynamic system [26]. In other words, for each target an independent Bernoulli RFS provides a unified statistical representation of target existence and target states. Thus, the multi-target multi-Bernoulli (MeMBer) filter was proposed as a tractable solution to multi-target tracking problems [14]. Since the MeMBer filter propagates the multi-target RFS as a multi-Bernoulli RFS, it is a parameterized approximation to the RFS based multi-target Bayes filter. Unlike the moment-based approximations, the MeMBer filter does not suffer from the problem of the extraction of state estimates [14, 33]. That is, as an advantage over the PHD/CPHD filters, its formulation for nonlinear models using sequential Monte Carlo (SMC) methods does not require employing error-prone clustering algorithms.

The multi-Bernoulli RFS formalism implicitly followed in the formulation of the PHD and CPHD filters draws more interest through the development of the MeMBer filter. Subsequently, alternative implementations and improvements to the MeMBer filter were proposed [35, 36, 21]. In addition, the Bernoulli RFS formalism was used in the development of single-target tracking filters. First, the traditional integrated probabilistic data association (IPDA) filter was formulated as RFS based Bayesian recursion [5]. Then, the joint target detection and tracking (JoTT) filter (also known as the Bernoulli filter) was developed [14, 30]. In [9], the Bernoulli filter was further improved to handle imprecise measurements. For comprehensive overview of the theory, implementation and applications, interested readers are referred to the tutorial on Bernoulli filters [26].

In the MeMBer filter, the estimated number of targets (i.e., the cardinality estimate) is positively biased. The reason for the bias was explained as a consequence of the multi-Bernoulli parameters computed for targets under the data-induced track set [33, 30]. The cardinality-balanced MeMBer (CBMeMBer) filter was introduced in [33] to remove the bias by revising the computation of these data-updated parameters. In addition, the SMC implementations of the CBMeMBer and MeMBer filters were proposed as well as their Gaussian mixture (GM) implementations for mildly nonlinear multi-target models. Similar to the improvements in [17, 18] the CBMeMBer filter was then improved further to handle the joint problem of unknown background clutter and detection profile while filtering [34].

Recently, the notion of labeled RFS was introduced as well as the conjugate prior distributions of these new RFSs with respect to the standard multi-target likelihood function [31]. Thus, the analytically tractable and closed-form solutions can be derived for their Bayesian inference. Hence, the generalized labeled multi-Bernoulli (GLMB) filter was proposed along with its relatively efficient version (known as $\delta-$GLMB filter) in terms of computational and memory resources [31]. These two filters propagate the history of data-associations together with track sets. However, the RFS formalism using the FISST benefits from the elimination of the computational problems due to explicit data associations in multi-target tracking [15, 14, 30]. Since this advantage is sacrificed for accuracy and estimation of target tracks, it is computationally more expensive than the CBMeMBer filter [31, 22]. In [22], the LMB filter was proposed as an efficient approximation of the $\delta-$GLMB filter. Thus, it inherits strong sides of the CBMeMBer filter and the $\delta-$GLMB filter. Therefore, it removes the restrictive high signal-to-noise ratio (SNR) and low clutter assumptions while it provides track estimates with better accuracy, compared to the CBMeMBer filter. However, the computational complexity of the LMB filter is still more expensive than that of the CBMeMBer filter since it requires explicit data associations as the $\delta-$GLMB filter.

In this paper, the relationship between the MeMBer corrector and the multi-Bernoulli RFS distribution is analyzed. By utilizing this analysis, a filter that offers a new statistical framework for the MeMBer data update process is proposed. Hence, the multi-Bernoulli RFS distribution is extended by spurious targets arising from targets under the legacy track set with high probabilities of existence. In the MeMBer update, the legacy track set is obtained from the predicted Bernoulli targets, assuming that they are not detected. However, all Bernoulli targets in the MeMBer prediction are also combined for each measurement to introduce targets under the data-induced track set. To resolve this ambiguity, the concept of spurious target is introduced under the data-induced track set against actual targets for the same measurements. Therefore, the proposed filter attempts to remove the positive bias observed in the MeMBer filter by distinguishing spurious targets from actual targets.

The modeling of spurious targets allows the refinement of the existence probabilities of targets in light of measurements. As a result, the stability of cardinality estimate is improved while the positive bias observed in the MeMBer filter is removed. When compared to the CBMeMBer filter, the proposed filter, which is hereinafter referred to as the improved MeMBer (IMeMBer) filter, does not make any limiting assumption on the probability of detection. In addition to simulations, theoretical analysis is performed in order to demonstrate the strengths and limitations of the proposed IMeMBer filter. Preliminary results on this work were published in [4].

This paper is organized as follows: Section 2.3 provides the necessary background on the multi-Bernoulli RFS, the MeMBer filter, and the MeMBer data update process. Section 2.4 analyzes the relationship between the MeMBer corrector and the multi-Bernoulli RFS distribution. By utilizing this analysis, the IMeMBer filter is derived in Section 2.5. The strengths and limitations of the IMeMBer filter are theoretically analyzed in Section 2.6. Simulation results are presented in Section 2.7. Finally, conclusions are drawn in Section 2.8.

2.3 Background

2.3.1 Multi-target Multi-Bernoulli Process

Let the spatial probability density functions of individual targets be denoted as $f^{(1)}(y), ..., f^{(\nu)}(y)$ if they do exist with probabilities $q^{(1)}, ..., q^{(\nu)}$. Since each target evolves independently from one another and follows a Bernoulli distribution with the parameter pair $\{q, f\}$, the multi-target state can be modeled as a multi-Bernoulli RFS, which is the union of the independent Bernoulli RFSs. Thus, the multi-Bernoulli RFS is described by the parameter set $\left\{q^{(i)}, f^{(i)}\right\}_{i=1}^{\nu}$. That is, the multi-target distribution is given by

$$f(Y) = \begin{cases} \left(1-q^{(1)}\right) ... \left(1-q^{(\nu)}\right) & \text{if } Y = \emptyset, \\ \sum\limits_{\beta} \prod\limits_{j=1}^{n} \Theta_{\beta} f^{\beta(j)}(y_j) & \text{if } Y = \{y_1, \ldots, y_n\}, \end{cases} \tag{2.1}$$

where

$$\Theta_{\beta} = \left(1-q^{(1)}\right) ... \left(1-q^{(\nu)}\right) \frac{q^{\beta(1)}}{(1-q^{\beta(1)})} \cdots \frac{q^{\beta(n)}}{(1-q^{\beta(n)})},$$

and the summation is taken over all permutations of the joint association hypotheses that are defined as one-to-one functions $\beta : \{1, ..., n\} \rightarrow \{1, ..., \nu\}$ for $n \leq \nu$ [14].

2.3.2 MeMBer Filter

The MeMBer filter consists of two steps: *i*) prediction and *ii*) update. These two steps are briefly presented below. Interested readers are referred to [14, 33] for further details.

Step-1: Prediction

Given a multi-Bernoulli RFS distribution parameterized by $\{q_{k-1}^{(i)}, f_{k-1}^{(i)}\}_{i=1}^{M_{k-1}}$ at time step $k-1$, the predicted distribution is also in the form of multi-Bernoulli but parameterized by the union of two independent multi-Bernoulli RFSs: the first RFS includes independent Bernoulli targets surviving from time $k-1$ according to a Markov state transition density $p_{k|k-1}(\cdot|y_{k-1})$ with probability $p_{S,k}(y)$. This RFS is described by the following Bernoulli parameters for each target:

$$q_{k|k-1}^{(i)} = q_{k-1}^{(i)} f_{k-1}^{(i)} \left[p_{S,k}\right], \tag{2.2}$$

$$f_{k|k-1}^{(i)}(y) = \frac{p_{k|k-1}\left[p_{S,k} f_{k-1}^{(i)}\right]}{f_{k-1}^{(i)}\left[p_{S,k}\right]}, \tag{2.3}$$

where for a given test function $h(y)$, a functional of $f(y)$ is defined by $f[h] = \int h(y) f(y)\, dy$.

In addition to surviving targets, another RFS is used as a birth model to explore newborn Bernoulli targets at time step k. This second RFS is independent of the first one and is described by $\{q_{\Gamma,k}^{(i)}, f_{\Gamma,k}^{(i)}\}_{i=1}^{M_{\Gamma,k}}$.

Step-2: Update

Suppose that the predicted multi-Bernoulli RFS distribution at time step k is given by $\{q^{(i)}_{k|k-1}, f^{(i)}_{k|k-1}\}_{i=1}^{M_{k|k-1}}$, where $M_{k|k-1} = M_{k-1} + M_{\Gamma,k}$. Then, the posterior distribution is approximately parameterized by the union of two multi-Bernoulli RFSs: the first RFS includes the parameters given by

$$q^{(i)}_{L,k} = q^{(i)}_{k|k-1} \frac{1 - f^{(i)}_{k|k-1}\left[p_{D,k}\right]}{1 - q^{(i)}_{k|k-1} f^{(i)}_{k|k-1}\left[p_{D,k}\right]}, \tag{2.4}$$

$$f^{(i)}_{L,k}(y) = \frac{1 - p_{D,k}(y)}{1 - f^{(i)}_{k|k-1}\left[p_{D,k}\right]} f^{(i)}_{k|k-1}(y), \tag{2.5}$$

for the predicted Bernoulli targets, assuming that they are not detected with probability $1 - p_{D,k}(y)$.

The second RFS includes the parameters referring to joint updates of all predicted Bernoulli targets. For each measurement z in the set Z_k these parameters are computed as

$$q_{U,k}(z) = \frac{\sum_{i=1}^{M_{k|k-1}} \frac{q^{(i)}_{k|k-1} f^{(i)}_{k|k-1}\left[p_{D,k} g_k(z|\cdot)\right]}{1 - q^{(i)}_{k|k-1} f^{(i)}_{k|k-1}\left[p_{D,k}\right]}}{\kappa(z) + \sum_{i=1}^{M_{k|k-1}} \frac{q^{(i)}_{k|k-1} f^{(i)}_{k|k-1}\left[p_{D,k} g_k(z|\cdot)\right]}{1 - q^{(i)}_{k|k-1} f^{(i)}_{k|k-1}\left[p_{D,k}\right]}}, \tag{2.6}$$

$$f_{U,k}(y; z) = \frac{\sum_{i=1}^{M_{k|k-1}} \frac{q^{(i)}_{k|k-1} p_{D,k}(y) g_k(z|y)}{1 - q^{(i)}_{k|k-1} f^{(i)}_{k|k-1}\left[p_{D,k}\right]} f^{(i)}_{k|k-1}(y)}{\sum_{i=1}^{M_{k|k-1}} \frac{q^{(i)}_{k|k-1} f^{(i)}_{k|k-1}\left[p_{D,k} g_k(z|\cdot)\right]}{1 - q^{(i)}_{k|k-1} f^{(i)}_{k|k-1}\left[p_{D,k}\right]}}, \tag{2.7}$$

where $g_k(z|y)$ is the sensor likelihood function at time step k and $\kappa(z)$ is the intensity function of Poisson distributed clutters.

2.3.3 MeMBer Data Update

To elaborate on the data update process given by (2.6) and (2.7), suppose that the single-sensor multi-target measurements at time step k are modeled by the RFS:

$$\mathrm{Z}_k = \left[\bigcup_{y \in Y_k} \Xi_k\left(y\right) \right] \mathrm{U}\mathrm{C}_k, \tag{2.8}$$

where C_k denotes measurements due to Poisson distributed clutter with intensity $\kappa(z) = \lambda_c \Phi(z)$ and $\Xi_k(y)$ denotes the measurement set produced by a target with state y. More precisely, a specified target is either detected with probability of target detection $p_{D,k}(y)$, thus $\Xi_k(y) = \{z\}$, or missed with probability $1 - p_{D,k}(y)$, thus $\Xi_k(y) = \emptyset$.

In the probability generating functional (p.g.fl.) form, the MeMBer corrector consists of two products given by [14]

$$G_k\left[h\right] \approx \prod_{i=1}^{M_{k|k-1}} G_{L,k}\left[h\right] \prod_{z \in \mathrm{Z}_k} G_{U,k}\left[z; h\right], \tag{2.9}$$

where the first product corresponds to the updated p.g.fl. of targets under the legacy track set, i.e., the individual updates of the $M_{k|k-1}$ predicted Bernoulli targets, assuming that they are not detected, while the second one corresponds to the updated p.g.fl. of targets under the data-induced track set, i.e., the joint updates of the $M_{k|k-1}$ predicted Bernoulli targets for each $z \in Z_k$.

Although the updated p.g.fl. of targets under the legacy track set is of multi-Bernoulli form, the updated p.g.fl. of targets under the data-induced track set is not. To address this problem, two approximations were proposed in the MeMBer filter:

first, the clutter process is assumed sparsely distributed in time. This approximation yields [14]

$$G_{U,k}\left[z;h\right]=\frac{\kappa\left(z\right)+\sum_{i=1}^{M_{k|k-1}}G_{U,k}^{(i)}\left[z;h\right]}{\kappa\left(z\right)+\sum_{i=1}^{M_{k|k-1}}G_{U,k}^{(i)}\left[z;1\right]}, \tag{2.10}$$

where

$$G_{U,k}^{(i)}\left[z;h\right]=\frac{q_{k|k-1}^{(i)}f_{k|k-1}^{(i)}\left[hp_{D,k}g_k\left(z\left|\cdot\right.\right)\right]}{1-q_{k|k-1}^{(i)}+q_{k|k-1}^{(i)}f_{k|k-1}^{(i)}\left[h\left(1-p_{D,k}\right)\right]}.$$

Since the numerator and denominator are both functions of h, $G_{U,k}[z;h]$ cannot be written in Bernoulli form, i.e., $G[h]=1-q+qf[h]$. Therefore, a second approximation is made by setting $h=1$ in the denominator of $G_{U,k}^{(i)}[z;h]$ and thus leading to [14]:

$$G_{U,k}^{(i)}\left[z;h\right]\approx\frac{q_{k|k-1}^{(i)}f_{k|k-1}^{(i)}\left[hp_{D,k}g_k\left(z\left|\cdot\right.\right)\right]}{1-q_{k|k-1}^{(i)}f_{k|k-1}^{(i)}\left[p_{D,k}\right]}. \tag{2.11}$$

In [33], it is asked why (2.11) is a good approximation. Therefore, we first examine this issue by analyzing the relationship between the MeMBer corrector and the multi-Bernoulli RFS distribution. This analysis introduces an alternative derivation for the MeMBer data update process and thus yields a new statistical framework to remove the positive bias from the data-updated cardinality estimate in the MeMBer filter.

2.4 Alternative MeMBer Data Update

The MeMBer filter is a parameterized approximation to the RFS based multi-target Bayes filter. Given this parameterization, our analysis is based on establishing an equivalence between the multi-target posterior distribution of targets under the data-induced track set and the multi-Bernoulli RFS distribution given by (2.1).

Using the product rule for set derivatives [11, 14], the posterior distribution corresponding to the p.g.fl. of the targets under the data-induced track set in (2.9) can be derived as follows:

$$\begin{aligned} f_{U,k}(Y) =& \frac{\delta^n}{\delta Y} \prod_{z \in \mathrm{Z}_k} G_{U,k}[z;h] \bigg|_{h=0}, \\ =& \begin{cases} \sum_{\beta} \prod_{j=1}^{n} \frac{\partial G_{U,k}}{\partial \delta y_j} \left[z_{\beta(j)}; 0 \right] \prod_{\substack{l=1 \\ \ell \notin \mathrm{Im}(\beta)}}^{m} G_{U,k}[z_\ell; 0] \quad \text{if } n \leq m, \\ 0 \quad \text{if } n > m, \\ \prod_{\ell=1}^{m} G_{U,k}[z_\ell; 0] \quad \text{if } n = 0, \end{cases} \end{aligned} \tag{2.12}$$

where $|Y| = n$, $|\mathrm{Z}_k| = m$ and the summation is taken over all permutations of the joint associations defined as one-to-one functions $\beta : \{1, ..., n\} \rightarrow \{1, ..., m\}$, and $\mathrm{Im}(\cdot)$ denotes the image of the function β.

The equivalence of $f_{U,k}(Y)$ and $f(Y)$ results in

$$\prod_{\ell=1}^{m} G_{U,k}[z_\ell; 0] = \left(1 - q^{(1)}\right) ... \left(1 - q^{(m)}\right), \tag{2.13}$$

and

$$\sum_{\beta} \prod_{j=1}^{n} \frac{\partial G_{U,k}}{\partial \delta y_j} \left[z_{\beta(j)}; 0 \right] \prod_{\substack{\ell=1 \\ \ell \notin \mathrm{Im}(\beta)}}^{m} G_{U,k}[z_\ell; 0] = \sum_{\beta} \prod_{j=1}^{n} \left[q^{\beta(j)} f^{\beta(j)}(y_j) \right] \prod_{\substack{\ell=1 \\ \ell \notin \mathrm{Im}(\beta)}}^{m} \left(1 - q^{\ell}\right), \tag{2.14}$$

for $Y = \emptyset$ and $Y = \{y_1, ..., y_n\}$, respectively. The equivalence expressed by (2.13) and (2.14) are analytically valid since the number of targets under the data-induced track set is equal to m in the MeMBer data update process. In (2.13), all of them are assumed to be nonexistent, while in (2.14), only n out of m targets under the data-induced track set are assumed to be existent.

From (2.13), the computation of existence probabilities using $1-q_{U,k}(z) = G_{U,k}[z;0]$ where

$$G_{U,k}\left[z;0\right] = 1 - \frac{\sum_{i=1}^{M_{k|k-1}} G_{U,k}^{(i)}\left[z;1\right]}{\kappa\left(z\right) + \sum_{i=1}^{M_{k|k-1}} G_{U,k}^{(i)}\left[z;1\right]}, \tag{2.15}$$

produces the same results as those computed in the MeMBer data update process.

Using (2.14), the existence probabilities obtained from (2.13) can be justified. In addition, the spatial probability density functions for targets under the data-induced track set can be computed if they do exist. Specifically, a particular concern here is to establish the facts that result in

$$\frac{\partial G_{U,k}}{\partial \delta y_j}\left[z_{\beta(j)};0\right] = q^{\beta(j)} f^{\beta(j)}\left(y_j\right). \tag{2.16}$$

When targets under the data-induced track set are approximated as multi-Bernoulli RFS, each target follows a Bernoulli RFS with a p.g.fl. of the form $G[h] = 1-q+qf[h]$. Suppose that a target under the data-induced track set exists, then using the functional derivative defined in [14] the following identity is obtained:

$$\frac{\partial G_{U,k}}{\partial \delta y}\left[z;0\right] = \frac{\partial G_{U,k}}{\partial \delta y}\left[z;1\right], \tag{2.17}$$

where the left hand side represents the probability density function of a Bernoulli RFS, while the right hand side represents its PHD function. Therefore, the PHD and the probability density function of a Bernoulli RFS are identical if the corresponding target does exist. In the MeMBer filter, the identity in (2.17) cannot be validated unless one of the following prior assumptions is made: $p_{D,k}(y) = 1$ or $h = 1$ in the denominator of $G_{U,k}^{(i)}[z;h]$ (see Appendix A for proof). This fact explains why the approximation setting $h = 1$ was proposed for the derivation of the MeMBer filter in [14].

Using the relationship between the p.g.fl. of an RFS and the probability generating function (p.g.f.) of its discrete cardinality distribution [13] (i.e., $G[h]|_{h(y)=\chi} = G(\chi)_{|Y|}$), the expected value of the cardinality, i.e., the existence probability, for each target under the data-induced track set is computed as

$$\begin{aligned} \frac{\partial G_{U,k}(\chi)_{|Y|}}{\partial \chi}|_{\chi=1} &= \frac{\partial}{\partial \chi}(1 - q + \chi q), \\ &= q. \end{aligned} \tag{2.18}$$

Thus, from (2.10) where $G_{U,k}^{(i)}[z;h]$ is approximated by (2.11), we get the result claimed in (2.15) as

$$q_{U,k}(z) = \frac{\sum_{i=1}^{M_{k|k-1}} G_{U,k}^{(i)}[z;1]}{\kappa(z) + \sum_{i=1}^{M_{k|k-1}} G_{U,k}^{(i)}[z;1]}. \tag{2.19}$$

In order to compute the corresponding spatial probability density function the following two steps are performed: First, the PHD of the Bernoulli RFS is computed from its p.g.fl. as

$$\frac{\partial G_{U,k}}{\partial \delta y}[z;1] = \frac{\sum_{i=1}^{M_{k|k-1}} \frac{q_{k|k-1}^{(i)} f_{k|k-1}^{(i)}(y) p_{D,k}(y) g_k(z|y)}{1 - q_{k|k-1}^{(i)} f_{k|k-1}^{(i)}[p_{D,k}]}}{\kappa(z) + \sum_{i=1}^{M_{k|k-1}} G_{U,k}^{(i)}[z;1]}, \tag{2.20}$$

and then following from (2.16) and (2.17), (2.20) is divided by (2.19). This normalization yields the same spatial probability density function as that computed in [14], i.e.,

$$f_{U,k}\left(y;z\right)=\frac{\sum_{i=1}^{M_{k|k-1}}\frac{q_{k|k-1}^{(i)}f_{k|k-1}^{(i)}(y)p_{D,k}(y)g_k(z|y)}{1-q_{k|k-1}^{(i)}f_{k|k-1}^{(i)}\left[p_{D,k}\right]}}{\sum_{i=1}^{M_{k|k-1}}G_{U,k}^{(i)}\left[z;1\right]}. \tag{2.21}$$

In the MeMBer filter, setting $h=1$ before deriving the multi-Bernoulli parameters produces a bias in the cardinality estimate [33]. The CBMeMBer filter deals with this issue by computing the existence probabilities of targets under the data-induced track set from (2.10) without making the approximation given by (2.11). In other words, the identity required for multi-Bernoulli RFS approximation in (2.17) is not validated for the computation of the existence probabilities in the CBMeMBer filter. Nevertheless, the spatial probability density functions computed in the CBMeMBer filter are valid, provided that $p_{D,k}(y)$ is set at $p_{D,k}(y)\approx 1$. That is, in the CBMeMBer filter, the required identity in (2.17) is validated at the stage where $f_{U,k}(y;z)$ is computed. Although no assumption was made on $p_{D,k}(y)$ in [14], substituting $p_{D,k}(y)\approx 1$ into (2.21) yields the same spatial probability density function as that computed in the CBMeMBer filter.

2.5 Improved MeMBer Filter

2.5.1 Modeling of Spurious Targets

In the MeMBer data update process, with probability $1-q_{U,k}(z)$ a target rising from any $z\in Z_k$ can be considered as non-existent. That is, each measurement can be thought of as having originated from the clutter process with probability $1-q_{U,k}(z)$. On the other hand, due to positive bias a target under the data-induced track set can be modeled as either actual or spurious. Therefore, if a target does exist, the disambiguation between spurious target and actual target requires an augmented Bernoulli RFS. Similar to the extension in [34], the state space of Bernoulli RFS Y is temporarily extended as $\Upsilon' = \Upsilon \times \{0,1\}$ in the data update process to label these targets, i.e., $Y' = \{y, u\} \in \Upsilon'$. Thus, the corresponding RFS distribution is defined as

$$f_{U,k}\left(Y'\right) = \begin{cases} 1 - q_{U,k}\left(z\right) & \text{if } Y' = \emptyset, \\ \tilde{q}_{U,k}(z)\tilde{f}_{U,k}\left(y\right) & \text{if } Y' = \left\{y, 0\right\}, \\ \bar{q}_{U,k}(z)\bar{f}_{U,k}\left(y\right) & \text{if } Y' = \left\{y, 1\right\}, \end{cases} \tag{2.22}$$

where the parameter pairs $\{\tilde{q}_{U,k}, \tilde{f}_{U,k}\}$ and $\{\bar{q}_{U,k}, \bar{f}_{U,k}\}$ for $u = 0$ and $u = 1$ characterize spurious and actual targets, respectively, provided that a target does exist with total probability $\tilde{q}_{U,k}(z) + \bar{q}_{U,k}(z) = q_{U,k}(z)$.

The p.g.fl. of this augmented Bernoulli RFS, if a target does exist, is written as

$$\begin{aligned} \bar{q}_{U,k}(z)\bar{f}_{U,k}\left[h\right] + \tilde{q}_{U,k}(z)\tilde{f}_{U,k}\left[h\right] &= \bar{G}_{U,k}\left[z;h\right] + \tilde{G}_{U,k}\left[z;h\right], \\ &= G_{U,k}\left[h;z\right], \end{aligned} \tag{2.23}$$

where the last equation proves the total probability by (2.18) for the mutually exclusive events $u = 0$ and $u = 1$. Thus, the multi-target posterior distribution given by (2.12) is extended using the product rule as follows:

$$f_{U,k}\left(Y'\right) = \begin{cases} \displaystyle\sum_{\beta}\sum_{\theta}\prod_{\substack{i=1 \\ \theta(i)>0}}^{\rho} \frac{\partial\tilde{G}}{\partial y_i}\left[z_{\beta(\theta(i))};0\right] \prod_{\substack{j=1 \\ j\notin\mathrm{Im}(\theta)}}^{n} \frac{\partial\bar{G}}{\partial y_j}\left[z_{\beta(j)};0\right] \prod_{\substack{\ell=1 \\ \ell\notin\mathrm{Im}(\beta)}}^{m} G_{U,k}\left[z_\ell;0\right] \quad \text{if} \quad n \leq m, \\ 0 \quad \text{if} \quad n > m, \\ \displaystyle\prod_{\ell=1}^{m} G_{U,k}\left[z_\ell;0\right] \quad \text{if} \quad n = 0, \end{cases} \tag{2.24}$$

where spurious targets correspond to the factors of the first product computed for all joint associations defined by a function $\theta : \{1, ..., \rho\} \rightarrow \{0, 1, ..., n\}$ for $\rho \leq n$. Note that $\theta(i) > 0$ denotes a unique association, while $\theta(i) = 0$ denotes a null association (i.e., no spurious target). The actual targets correspond to the factors of the second product computed for all joint associations defined by the one-to-one function β, provided that $j \notin \mathrm{Im}(\theta)$.

For each Bernoulli target i in the MeMBer prediction, there is a corresponding target with probability $q_{L,k}^{(i)}$ under the legacy track set. In the MeMBer data update process, all Bernoulli targets in the MeMBer prediction are also compiled for each $z \in Z_k$ to declare a target under the data-induced track set. Without considering targets under the legacy track set, their contributions to targets under the data-induced track set would give rise to a positive bias in the cardinality estimate like an internal clutter generator. Based on this premise, the p.g.fl. in (2.11) for any z can be separated into two terms:

$$G_{U,k}^{(i)}[z;h] = \frac{q_{k|k-1}^{(i)} q_{L,k}^{(i)} f_{k|k-1}^{(i)}[h p_{D,k} g_k(z|\cdot)]}{1 - q_{k|k-1}^{(i)} f_{k|k-1}^{(i)}[p_{D,k}]} + \frac{q_{k|k-1}^{(i)} \left(1 - q_{L,k}^{(i)}\right) f_{k|k-1}^{(i)}[h p_{D,k} g_k(z|\cdot)]}{1 - q_{k|k-1}^{(i)} f_{k|k-1}^{(i)}[p_{D,k}]}, \tag{2.25}$$

where the first term on the right hand side denotes the contribution to the p.g.fl. of a spurious target, i.e., $\tilde{G}_{U,k}[z;h]$, since the predicted Bernoulli target i is already updated as a target under the legacy track set, while the second term denotes the contribution to the p.g.fl. of an actual target, i.e., $\bar{G}_{U,k}[z;h]$, by satisfying the total probability. Thus, (2.25) shows how to use (2.23) for the multi-Bernoulli RFS.

Using (2.18) with $\bar{G}_{U,k}[z;h]$ and $\tilde{G}_{U,k}[z;h]$, the probability of existence for the actual and spurious targets are computed as

$$\bar{q}_{U,k}(z) = \frac{\sum_{i=1}^{M_{k|k-1}} \frac{q_{k|k-1}^{(i)}\left(1-q_{L,k}^{(i)}\right) f_{k|k-1}^{(i)}\left[p_{D,k} g_k(z|\cdot)\right]}{1-q_{k|k-1}^{(i)} f_{k|k-1}^{(i)}\left[p_{D,k}\right]}}{\kappa\left(z\right) + \sum_{i=1}^{M_{k|k-1}} G_{U,k}^{(i)}\left[z;1\right]}, \tag{2.26}$$

and

$$\tilde{q}_{U,k}\left(z\right) = \frac{\sum_{i=1}^{M_{k|k-1}} \frac{q_{k|k-1}^{(i)} q_{L,k}^{(i)} f_{k|k-1}^{(i)}\left[p_{D,k} g_k(z|\cdot)\right]}{1-q_{k|k-1}^{(i)} f_{k|k-1}^{(i)}\left[p_{D,k}\right]}}{\kappa\left(z\right) + \sum_{i=1}^{M_{k|k-1}} G_{U,k}^{(i)}\left[z;1\right]}, \tag{2.27}$$

respectively.

According to (2.22), if a spurious target does exist, its probability density function has the form of $\tilde{q}_{U,k}(z)\tilde{f}_{U,k}(y;z)$. In addition, (2.17) is still valid for $\tilde{G}_{U,k}\left[z;h\right]$. This is because each contribution from (2.11) to $\tilde{G}_{U,k}\left[z;h\right]$ using (2.25) is proportional to $q_{L,k}^{(i)}$, which is independent of h. Thus, its spatial probability density function can be obtained by dividing the PHD computed from $\tilde{G}_{U,k}\left[z;h\right]$ using (2.17) by (2.27):

$$\tilde{f}_{U,k}\left(y;z\right) = \frac{\sum_{i=1}^{M_{k|k-1}} \tilde{\alpha}^{(i)} f_{k|k-1}^{(i)}\left(y\right) p_{D,k}\left(y\right) g_k\left(z\,|y\right)}{\sum_{i=1}^{M_{k|k-1}} \tilde{\alpha}^{(i)} f_{k|k-1}^{(i)}\left[p_{D,k} g_k\left(z\,|\cdot\right)\right]}, \tag{2.28}$$

where $\tilde{\alpha}^{(i)} = q_{k|k-1}^{(i)} q_{L,k}^{(i)}(1 - q_{k|k-1}^{(i)} f_{k|k-1}^{(i)}[p_{D,k}])^{-1}$. This spatial probability density function is bounded except when $p_{D,k}(y) = 1$ or $p_{D,k}(y) = 0$. The reason is that either the bias in the cardinality estimate is equal to zero when $p_{D,k}(y) = 1$ or there are no targets under the data-induced track set when $p_{D,k}(y) = 0$, in which case there would be no spurious targets generated in the MeMBer filter.

Similarly, the spatial probability density function of each actual target under the data-induced track set can be obtained by dividing the PHD computed from $\bar{G}_{U,k}[z;h]$ using (2.17) by (2.26):

$$\bar{f}_{U,k}(y;z) = \frac{\sum_{i=1}^{M_{k|k-1}} \bar{\alpha}^{(i)} f_{k|k-1}^{(i)}(y)\, p_{D,k}(y)\, g_k(z\,|y)}{\sum_{i=1}^{M_{k|k-1}} \bar{\alpha}^{(i)} f_{k|k-1}^{(i)}\left[p_{D,k} g_k(z\,|\cdot)\right]}, \tag{2.29}$$

where $\bar{\alpha}^{(i)} = q_{k|k-1}^{(i)}(1-q_{L,k}^{(i)})(1-q_{k|k-1}^{(i)} f_{k|k-1}^{(i)}[p_{D,k}])^{-1}$. In contrast to the CBMeMBer filter, no limiting assumption is made on $p_{D,k}(y)$. Nevertheless, substituting $p_{D,k}(y) \approx 1$ into (2.29) yields the spatial probability density function, which is the same as that computed in the CBMeMBer filter.

In [33], the PHD function of the updated p.g.fl. has positive and negative parts. Therefore, a valid spatial probability density function cannot be obtained unless the negative part is neglected by the prior assumption of $p_{D,k}(y) \approx 1$. The existence probability of a spurious target given by (2.27) explains the cardinality distribution from that negative part of the PHD function. See Appendix B for this alternative explanation for modeling spurious targets.

As in the IMeMBer and CBMeMBer filters, the prediction step of the LMB filter corresponds to that of the MeMBer filter. However, the approximation followed in its update step matches its PHD with that of the multi-target posterior density [22]. Hence, it removes the high SNR and low clutter assumptions made for the CBMeMBer filter while it requires computationally expensive exact data associations for this efficient approximation. It is important to note that the proposed IMeMBer filter eliminates the high SNR assumption but it still requires low clutter assumption because of the first approximation in the MeMBer filter. In addition, it benefits from the RFS formalism to eliminate exact data associations. Therefore, the time complexity of the CBMeMBer (or IMeMBer) filter is approximately same as that of the PHD filter, i.e., linear in the number of measurements [16]. On the other hand, the time complexity of the LMB filter is at worst cubic in the number of measurements [22].

2.5.2 Physical Interpretation of Track Sets

The MeMBer data update compiles all Bernoulli targets in the MeMBer prediction for each $z \in Z_k$ to declare a target under the data-induced track set. By modeling of spurious targets their contributions to targets under the data-induced track set decreases, especially if the corresponding targets under the legacy track set have high existence probabilities. However, the existence probabilities of targets under the legacy track set are computed without the current measurement set. In addition, the use of two different mathematical models under the legacy track set and the data-induced track set for the existence probability of the same target would yield poor performance in the cardinality estimate. This is because they refer to the same physical event in the real world. This situation is illustrated with the following example customized to a single Bernoulli target from [1].

Assume that we have an unfair coin with probabilities $P_1\left('head'\right) = 0.6$ and $P_1\left('tail'\right) = 0.4$. The events $'head'$ and $'tail'$ can be considered as mutually exclusive $'target\ existence'$ and $'target\ death'$ events for a Bernoulli target, respectively. Thus, a probability space is constructed as $(\Omega_1, \sigma_1, P_1)$ where $\Omega_1 = \{'head','tail'\}$ is the sample space, and σ_1 is the sigma field of measurable subsets of Ω_1 through the measure P_1.

Table 2.1: Cardinality Distributions for Alternative Probability Spaces

Cardinality distribution \Target number	$n=0$	$n=1$	$n=2$
$p_{\|Y\|,1}(n)$	0.4	0.6	n/a
$p_{\|Y\|,2}(n)$	0.49	0.42	0.09

Alternatively, a second probability space is formed as $(\Omega_2, \sigma_2, P_2)$ where $\Omega_2 = \{'head1', 'head2', 'tail'\}$ with $P_2('head1') = 0.3$, $P_2('head2') = 0.3$, and $P_2('tail') = 0.4$. Similar to the first probability space, the event $'tail'$ can be considered as $'target\ death'$ event for a Bernoulli target. On the other hand, the events $'head1'$ and $'head2'$ can be considered as $'target\ existence'$ event for the same Bernoulli target but under the legacy track set and data-induced track set, respectively.

As indicated in [1], these probability spaces can be considered the equivalent models of randomness, i.e., $P_1('tail') = P_2('tail')$ and $P_1('head') = P_2('head1') + P_2('head2')$. However, the (CB)MeMBer filter determines the number of existing targets with the maximum a posteriori (MAP) estimate from the cardinality distribution of multi-Bernoulli RFS [14, 33]. Using the existence probabilities of the Bernoulli target(s) in these two alternative probability spaces, the cardinality distributions are computed in Table 2.1.

According to Table 2.1, the MAP estimates from the cardinality distributions contradict with each other. The same contradiction would also arise if the existence of a target were determined with the highest probability in both probability spaces [1].

2.5.3 Refinement of Existence Probabilities

To address the problem observed in the physical interpretation of track sets, the refinement of existence probabilities are proposed. Before the multi-target state estimation, the refinement process determines targets detected under the legacy track set using the following test statistics obtained from (2.27):

$$\tilde{q}_{U,k}\left(z,i\right) = \frac{\frac{q^{(i)}_{k|k-1}\, q^{(i)}_{L,k} f^{(i)}_{k|k-1}\left[p_{D,k} g_k(z|\cdot)\right]}{1-q^{(i)}_{k|k-1}\, f^{(i)}_{k|k-1}\left[p_{D,k}\right]}}{\kappa\left(z\right) + \sum_{i=1}^{M_{k|k-1}} G^{(i)}_{U,k}\left[z;1\right]}. \tag{2.30}$$

Then, it purifies the legacy track set by eliminating these targets in light of measurements. Thus, the contributions to actual targets in (2.26) are to be updated, i.e., computed as in the MeMBer data update process.

To determine targets detected under the legacy track set, the following two methods are designed: i-) using the existence probabilities given by (2.27) for each $z \in Z_k$, compute the cardinality distribution of spurious targets [14]. Thus, the MAP estimate of their cardinality, i.e., $\tilde{n}^{MAP}$, is determined. Then, select the $\tilde{n}^{MAP}$ spurious targets with the highest existence probabilities $\tilde{q}_{U,k}(z)$. However, the predicted Bernoulli targets may not be resolved for the measurements $Z^s_k \subset Z_k$, which give rise to the selected spurious targets. Considering one target can generate at most one measurement, targets detected under the legacy track set are determined by maximizing (2.30) for each $z \in Z^s_k$, i.e.,

$$i^* = \arg\max_{i\in I}\left(\tilde{q}_{U,k}\left(z,i\right)\right), \tag{2.31}$$

where $I = \left\{1, ..., M_{k|k-1}\right\}$ is the dynamic list consisting of the indices of targets, which are not detected, under the legacy track set. Therefore, an index j is removed from the list I if $i^* = j$ for any $z \in Z_k^s$. The legacy track set is purified by eliminating the target with index i^*, i.e., $q_{L,k}^{(i^*)} = 0$. Thus, the existence probability of the actual target against the spurious one for the same z is updated as

$$\bar{q}_{U,k}\left(z\right) = \bar{q}_{U,k}\left(z\right) + \tilde{q}_{U,k}\left(z\right). \tag{2.32}$$

This is because no positive bias is observed for the measurement z after the purification. In addition, *ii*-) for the remaining indices in the list I, some targets under the legacy track set may exist with small to moderate probabilities. On the other hand, there may be actual targets with high to moderate existence probabilities to refer to the same physical events. That is, spurious targets with small to moderate existence probabilities may indicate resolved targets, which are detected under the legacy track set. To determine these targets, the remaining measurements in $Z_k - Z_k^s$ are used in (2.31). Thus, the legacy track set is purified by setting $q_{L,k}^{(i^*)} = 0$ conditional on

$$\tilde{q}_{U,k}\left(z, i\right) \begin{cases} \geq Th \text{ if } i = i^*, \\ < Th \text{ otherwise,} \end{cases} \tag{2.33}$$

where $z \in Z_k - Z_k^s$ and Th is the termination threshold for Bernoulli targets. Then, the existence probability of the actual target for the same z is updated as in (2.32). Consequently, all Bernoulli RFSs under the legacy track set and data-induced track set represent one physical event for different targets. Table 2.2 shows the pseudo-code of the refinement process carried out according to these two methods.

Table 2.2: Refinement Process

Input: $\tilde{Q}_k = \left\{\tilde{q}_{U,k}^{(i)}\right\}_{i=1}^{N_m}$ where $N_m = |Z_k|$

- Compute the cardinality distribution from $\tilde{Q}_k$ [14].
- Determine MAP estimate of the cardinality, i.e., $\tilde{n}^{MAP}$.
- Sort the existence probabilities of spurious targets:
  - $[\sim, Id] = sort\left(\tilde{Q}_k, \text{'descend'}\right)$.
- Obtain Z_k^s from the ordered list Id using $\tilde{n}^{MAP}$.
- Initialize the dynamic list $I = \left\{1, ..., M_{k|k-1}\right\}$.
- for $z \in Z_k^s$
  - Evaluate (2.31) to find the index i^*.
  - Set $q_{L,k}^{(i^*)} = 0$ and evaluate (2.32).
  - Update the dynamic list: $I = I - \{i^*\}$.

  end
- for $z \in Z_k - Z_k^s$
  - Evaluate (2.31) to find the index i^*.
  - If the condition in (2.33) is satisfied,
    - Set $q_{L,k}^{(i^*)} = 0$ and evaluate (2.32).
    - Update the dynamic list: $I = I - \{i^*\}$.

    end

  end

2.6 Theoretical Analysis of IMeMBer Filter

The aim of this section is to theoretically analyze the strengths and limitations of the proposed IMeMBer filter. The most convenient way of doing this analysis is to substitute Bernoulli RFS for multi-Bernoulli RFS. Thus, the complex problem at hand simplifies to the joint detection and state estimation problem of a single target with measurements of uncertain origin. In this case, the joint target detection and tracking (JoTT) filter provides an exact and theoretically optimal Bayesian solution [14].

In traditional single target tracking problems, a target of interest is assumed to be always present [2]. The integrated probabilistic data association (IPDA) is an algorithm that removes this assumption by estimating the probability of target existence along with the target's states [19]. The RFS formulation of the IPDA algorithm was derived in [5]. Then, its extended RFS versions referred to as the Bernoulli filter [30] or the JoTT filter [14] was developed.

Suppose that at most one target is present and this is known a priori. In this case, the RFS of target state is modeled as a Bernoulli RFS with the parameter pair $\{q, f\}$ so that

$$f\left(Y\right) = \begin{cases} 1 - q & \text{if } Y = \emptyset \\ qf\left(y\right) & \text{if } Y = \{y\}\,. \end{cases} \tag{2.34}$$

The prediction step of the MeMBer filter is exact, i.e., given the multi-Bernoulli RFS, the prediction for each target is independently modeled by a Bernoulli RFS. However, the MeMBer data update approximates the exact Bayesian multi-target update [14]. Therefore, the theoretical analysis will focus on the comparison between the IMeMBer corrector and the JoTT corrector.

In the original derivation of the JoTT filter, the false alarm process is modeled as an arbitrary RFS. When the Poisson false alarm RFS is substituted for the arbitrary false alarm RFS, the original update equations of the JoTT filter defined in [14] are given by

$$q_{k|k} = \frac{1 - f_{k|k-1}\left[p_{D,k}\right] + \sum_{z \in Z_k} \frac{f_{k|k-1}\left[p_{D,k} g_k(z|\cdot)\right]}{\kappa(z)}}{q_{k|k-1}^{-1} - f_{k|k-1}\left[p_{D,k}\right] + \sum_{z \in Z_k} \frac{f_{k|k-1}\left[p_{D,k} g_k(z|\cdot)\right]}{\kappa(z)}}, \tag{2.35}$$

$$f_{k|k}\left(y\right) = \frac{1 - p_{D,k}\left(y\right) + p_{D,k}\left(y\right) \sum_{z \in Z_k} \frac{g_k(z|y)}{\kappa(z)}}{1 - f_{k|k-1}\left[p_{D,k}\right] + \sum_{z \in Z_k} \frac{f_{k|k-1}\left[p_{D,k} g_k(z|\cdot)\right]}{\kappa(z)}} f_{k|k-1}\left(y\right). \tag{2.36}$$

For a Bernoulli RFS, the legacy part of the updated p.g.fl. in (2.9) is given by [14]

$$G_{L,k}\left[h\right] = \frac{1 - q_{k|k-1} + q_{k|k-1} f\left[h\left(1 - p_{D,k}\right)\right]}{1 - q_{k|k-1} + q_{k|k-1} f\left[\left(1 - p_{D,k}\right)\right]}, \tag{2.37}$$

while its data-induced part given by (2.10) simplifies to

$$\bar{G}_{U,k}\left[z;h\right] = \frac{\kappa\left(z\right) + \bar{G}^{(i)}_{U,k}\left[z;h\right]|_{i=1}}{\kappa\left(z\right) + G_{U,k}\left[z;1\right]}, \tag{2.38}$$

where $\bar{G}^{(i)}_{U,k}\left[z;h\right]|_{i=1}$ denotes the p.g.fl. of an actual target in (2.25). Thus, using (2.18) the probability of existence of a single target is computed as

$$\begin{aligned} q_{k|k} &= \frac{\partial}{\partial \chi}\left[G_{L,k}\left(\chi\right)\prod_{z\in Z_k}\bar{G}_{U,k}\left(z;\chi\right)\right]\Bigg|_{\chi=1}, \\ &= \frac{\partial G_{L,k}}{\partial \chi}\left(\chi\right)\Bigg|_{\chi=1}\prod_{z\in Z_k}\bar{G}^{(0)}_{U,k}\left(z;\chi\right)\Bigg|_{\chi=1} + \\ &\quad G^{(0)}_{L,k}\left(\chi\right)|_{\chi=1}\sum_{z\in Z_k}\frac{\partial \bar{G}_{U,k}}{\partial \chi}\left(z;\chi\right)\Bigg|_{\chi=1}, \end{aligned} \tag{2.39}$$

where the last equation follows from the product rule [11, 14].

Since a Bernoulli target is present as either a legacy or a data-induced track, i.e., these two events are mutually exclusive for a Bernoulli RFS, it follows that

$$\begin{aligned} G^{(0)}_{L,k}\left(\chi\right)|_{\chi=1} &= p_{L,|Y_k|}\left(0\right), \\ &= 1.0, \end{aligned} \tag{2.40}$$

and

$$\begin{aligned}\prod_{z\in Z_k} \bar{G}^{(0)}_{U,k}(z;\chi)\bigg|_{\chi=1} &= \bar{p}_{U,|Y_k|}(0)\,, \\ &= 1.0,\end{aligned} \tag{2.41}$$

where the last equations in (2.40) and (2.41) are results of [13]

$$G(n) = \sum_{n=0}^{\infty} \chi^n p_{|Y|}(n)\,,$$

where $p_{|Y|}(n)$ denotes the cardinality distribution evaluated at nonnegative integer n.

From [13] we know that the expected value of the cardinality is given by

$$\begin{aligned}\frac{\partial G}{\partial \chi}(\chi)\bigg|_{\chi=1} &= G^{(1)}(1)\,, \\ &= \textstyle\sum_{n=1}^{\infty} n p_{|Y|}(n)\,.\end{aligned}$$

For a Bernoulli RFS, this simplifies to

$$G^{(1)}(1) = p_{|Y|}(1)\,,$$

where $p_{|Y|}(1)$ denotes the probability that a single target is present, i.e., the probability of target existence. The gradient derivatives in (2.39) are therefore expressed as

$$G^{(1)}_{L,k}(1) = q_{L,k}, \tag{2.42}$$

$$\bar{G}^{(1)}_{U,k}(z,1) = \bar{q}_{U,k}(z)\,. \tag{2.43}$$

Then, substituting (2.40)–(2.43) into (2.39), the posterior probability of target existence is defined as

$$q_{k|k} = q_{L,k} + \sum_{z \in Z_k} \bar{q}_{U,k}(z). \tag{2.44}$$

This probability measure is valid, i.e., $0 \leq q_{k|k} \leq 1.0$ as long as clutter points are sparsely distributed. In other words, it is unlikely that any two measurements are highly related to a Bernoulli target via the sensor likelihood function $g_k(z|y)$. Recall that, for multi-Bernoulli RFS formalism of data-induced part of the MeMBer corrector, the clutter distribution is also approximated not to be too dense [14]. Accordingly, two cases are possible: 1) if no measurement is collected, i.e. $Z_k = \emptyset$, then

$$\begin{aligned} q_{k|k} &= q_{L,k}, \\ &= q_{k|k-1} \frac{1 - f_{k|k-1}[p_{D,k}]}{1 - q_{k|k-1} f_{k|k-1}[p_{D,k}]}, \end{aligned} \tag{2.45}$$

and 2) if there is only one measurement that makes significant contribution to the single target under the data-induced track set while the others are negligible, i.e., implicitly assuming that $Z_k = \{z\}$, then

$$q_{k|k} \approx q_{L,k} + \bar{q}_{U,k}(z). \tag{2.46}$$

After substituting $q_{L,k}$ and $\bar{q}_{U,k}(z)$ into (2.46), some algebraic manipulations yield

$$q_{k|k} \approx \frac{1 - f_{k|k-1}[p_{D,k}] + \frac{f_{k|k-1}[p_{D,k} g_k(z|\cdot)]}{\kappa(z)}}{q_{k|k-1}^{-1} - f_{k|k-1}[p_{D,k}] + \frac{f_{k|k-1}[p_{D,k} g_k(z|\cdot)]}{\kappa(z)}}. \tag{2.47}$$

From the identity in (2.17), the PHD of the posterior density of a Bernoulli RFS is given by $q_{k|k} f_{k|k}(y)$. Thus, the corresponding spatial probability density functions can be obtained by dividing the PHDs computed under these two cases by (2.45) and (2.47), respectively (see Appendix C for proof). That is,

$$f_{L,k}(y) = \frac{1 - p_{D,k}(y)}{1 - f_{k|k-1}\left[p_{D,k}\right]} f_{k|k-1}(y), \tag{2.48}$$

$$f_{U,k}(y) = \frac{1 - p_{D,k}(y) + p_{D,k}(y)\frac{g_k(z|y)}{\kappa(z)}}{1 - f_{k|k-1}\left[p_{D,k}\right] + \frac{f_{k|k-1}\left[p_{D,k}g_k\right]}{\kappa(z)}} f_{k|k-1}(y). \tag{2.49}$$

Likewise, for each of these two cases, JoTT corrector given by (2.35) and (2.36) exactly reduces to the same equations, i.e., (2.45) and (2.48) or (2.47) and (2.49), respectively. Therefore, the IMeMBer corrector agrees with the theoretically optimal JoTT corrector for a Bernoulli RFS as long as the Poisson clutter is sparsely distributed. On the other hand, for each of these two cases, it was demonstrated that the legacy part of the MeMBer corrector resembles the JoTT corrector, while its data-induced part resembles the PHD corrector [14]. Since the PHD is the first order moment approximation to the multi-target posterior density, significant information is lost even in the single target case [8]. This may theoretically explain why the IMeMBer filter outperforms the MeMBer filter. Furthermore, this theoretical analysis indicates that the CPHD filter would outperform the IMeMBer filter in highly cluttered environments. This is because for each of these two cases, it was demonstrated that the CPHD filter reduces to the JoTT filter without making any limiting assumption on the rate of clutter [14].

2.7 Simulation Results

In this section, the performance of the proposed IMeMBer filter is validated by two simulation examples. The performance evaluation is based on the stability and accuracy of cardinality estimate. For this purpose, the Optimal Subpattern Assignment (OSPA) metric is employed [25]. The OSPA metric compares two RFSs by measuring difference in their cardinalities and localization error between associated elements after an optimal assignment algorithm. The sensitivity of the OSPA metric to these two types of errors are adjusted by the cut-off parameter c and the order parameter p. As indicated in [27], for $p = 2$ smooth distance curves are obtained by computing the localization error as in other traditional metrics, e.g., the root mean squared error (RMSE). Therefore, the OSPA metric, which is sensitive to the stability and accuracy in cardinality estimate, is computed by setting $p = 2$ and $c = 25$.

Implementation of the IMeMBer filter differs from the CBMeMBer filter in the data update process, where (2.26) and (2.29) are computed. In addition, the test statistics given by (2.30) are obtained from (2.27) in order to refine the existence probabilities in light of measurements before the multi-target state estimation. Details about both SMC and GM implementations of the CBMeMBer filter can be found in [33].

2.7.1 Nonlinear Multi-target Tracking Example

In this example, the IMeMBer filter is compared with the CBMeMBer and LMB filters through the multi-target tracking scenario shown in Fig. 2.1. The state vector of each individual target comprises positions and velocities in $x - y$ directions, i.e., $y = [p_x,\, p_y,\, v_x,\, v_y]^T$. If target does survive with probability $p_{S,k}(y) = 0.95$, its states evolve according to the discrete white noise acceleration model [2]. The state evolution error is modeled as white Gaussian noise with standard deviations $\sigma_{v,x} = 0.3\mathrm{m/s}^2$ and $\sigma_{v,y} = 0.3\mathrm{m/s}^2$ to cover small accelerations.

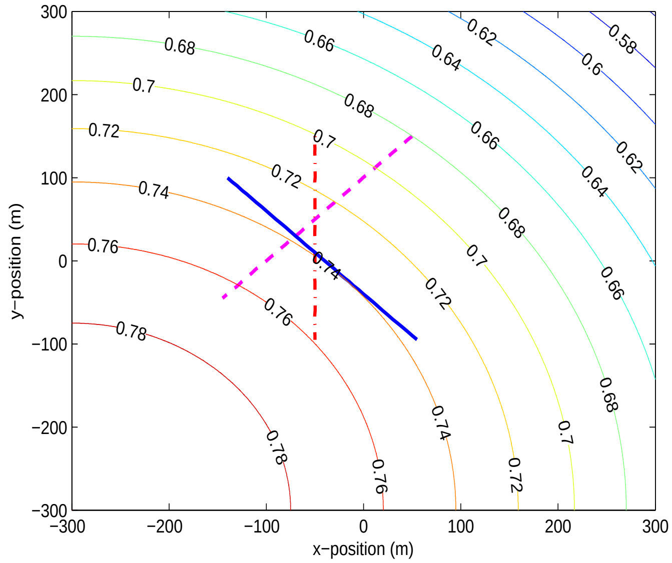

Figure 2.1: Trajectories of three targets along with the change in the state-dependent probability of target detection.

The scene is monitored by a range-bearing sensor located at $[p_{s,x},\ p_{s,y}] = [-300\mathrm{m},\ -300\mathrm{m}]^T$ and the target-originated measurements are given by

$$\Theta_k = \arctan\left((p_x - p_{s,x})/(p_y - p_{s,y})\right) + \omega_{\Theta,\,k},$$

$$r_k = \left(\sqrt{(p_x - p_{s,x})^2 + (p_y - p_{s,y})^2}\right) + \omega_{r,k},$$

where $\omega_{\Theta,\,k}$ and $\omega_{r,k}$ denote mutually independent, zero-mean Gaussian noise sequences for bearing and range measurements with the standard deviations $\sigma_{\Theta,k} = \pi/180\mathrm{rad}$ and $\sigma_{r,k} = 0.1\mathrm{m}$, respectively. As can be seen in Fig. 2.1, the sensor has a state-dependent probability of target detection given by

$$p_{D,k}(y) = \frac{0.80N\left([p_x, p_y]^T; [p_{s,x}, p_{s,y}]^T, 1000^2 I_2\right)}{N\left([p_{s,x}, p_{s,y}]^T; [p_{s,x}, p_{s,y}]^T, 1000^2 I_2\right)}, \tag{2.50}$$

In addition, a Poisson clutter model generates uniformly distributed false alarms over the surveillance region $V = [-300\text{m}, 300\text{m}] \times [-300\text{m}, 300\text{m}]$ with the average rate of $\lambda_c = 5$ per scan. Note that the positive bias observed in the MeMBer filter reduces as $p_{D,k}$ takes values close to unity such that when $p_{D,k} = 1.0$, it vanishes [33]. In Fig. 2.1, the state dependent probability of detection given by (2.50) takes values in the range of [0.58, 0.80] over the surveillance region. Hence, it is expected that the cardinality estimate in the MeMBer filter would be biased significantly. In addition, for the moderate values of $p_{D,k}(y)$ the assumption for the derivation of valid spatial probability density functions in the CBMeMBer filter is not satisfied.

In the considered example, the three different filters explore newborn targets according to the birth model given by $\{q^{(i)}_{\Gamma,k}, f^{(i)}_{\Gamma,k}\}_{i=1}^{3}$ where the existence probabilities are set to $q^{(1)}_{\Gamma,k} = q^{(2)}_{\Gamma,k} = q^{(3)}_{\Gamma,k} = 0.05$, and the spatial density functions are modeled by Gaussian densities $f^{(i)}_{\Gamma,k} = N(y; m^{(i)}_{\Gamma,k}, P_\Gamma)$ with means $m^{(1)}_{\Gamma,k} = [-50, 150, 0, 0]^T$, $m^{(2)}_{\Gamma,k} = [50, 150, 0, 0]^T$, $m^{(3)}_{\Gamma,k} = [-140, 100, 0, 0]^T$, and identical covariance matrices $P_\Gamma = diag([25, 25, 15, 15])$. Fig 2.2 shows the x and y components of the target trajectories, measurements and the position estimates obtained from the IMeMBer filter for one Monte Carlo trial. It can be seen that two targets cross each other at time steps $k = 21$, $k = 25$ and $k = 29$.

The three filters are initialized by using the multi-Bernoulli birth model at time step $k = 1$. At the end of each iteration, Bernoulli targets with existence probabilities

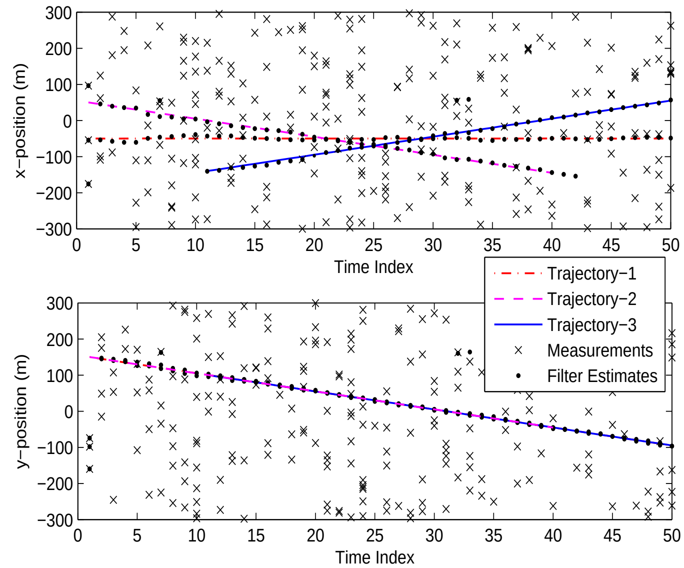

Figure 2.2: x and y components of target trajectories, measurements and IMeMBer filter estimates.

less than $Th = 10^{-2}$ are terminated. Then, the resampling step allocates a number of particles to each Bernoulli target, which is proportional to its probability of existence between $l_{max} = 2000$ and $l_{min} = 1000$, i.e., $l_k^{(i)} \propto q_k^{(i)} l_{\max}$ s.t. $l_k^{(i)} > l_{\min}$ [29].

The performances of the three filters are demonstrated by running the scenario shown in Fig. 2.1 for 500 Monte Carlo runs. In each trial, a random measurement data set is generated for the same target trajectories. Fig. 2.3 shows the average cardinality estimates and their ± 1 standard deviations for the three filters. The average standard deviations measured from the plots of the IMeMBer, CBMeMBer, and LMB filters are 0.59, 0.81, and 0.62, respectively. These values indicate that, in terms of the stability of cardinality estimate, the IMeMBer filter outperforms the LMB filter, which in turn outperforms the CBMeMBer filter even though the cardinality estimates from the three filters converge to the true cardinality. In addition, observe that both the IMeMBer and LMB filters' responses to target death after time step $k = 40$ is slower than that of the CBMeMBer filter. This is because, like the comparison between the PHD and CPHD filters in [32], these filters have higher confidence on their cardinality estimates with smaller standard deviations, compared to the CBMeMBer filter.

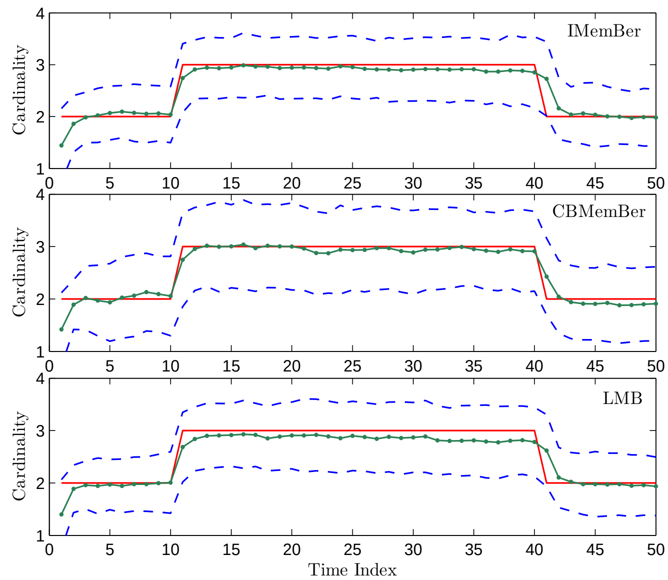

Figure 2.3: The average cardinality estimates over 500 Monte Carlo runs: true cardinality (red solid line), estimated cardinality (green dotted line), and their ±1 standard deviations (blue dashed lines) for the SMC-IMeMBer, SMC-CBMeMBer, and SMC-LMB filters.

In Fig. 2.4, the cardinality and localization performances of the three filters are measured as one using the OSPA metric. As expected, Fig. 2.4 corroborates observations inferred from Fig. 2.3. In other words, the OSPA metric with parameters $p = 2$ and $c = 25$ penalizes the CBMeMBer filter much more than the other two filters due to the highest instability observed in its cardinality estimate. In addition, it can be seen that the the IMeMBer filter outperforms the LMB filter. However, observe the abrupt changes in the OSPA metric for these two filters due to their slow responses to the target death after time step $k = 40$. At this point, it is important to note that both Fig. 2.3 and Fig. 2.4 demonstrate that the IMeMBer filter successfully handles the crossing targets at time steps $k = 21$, $k = 25$ and $k = 29$.

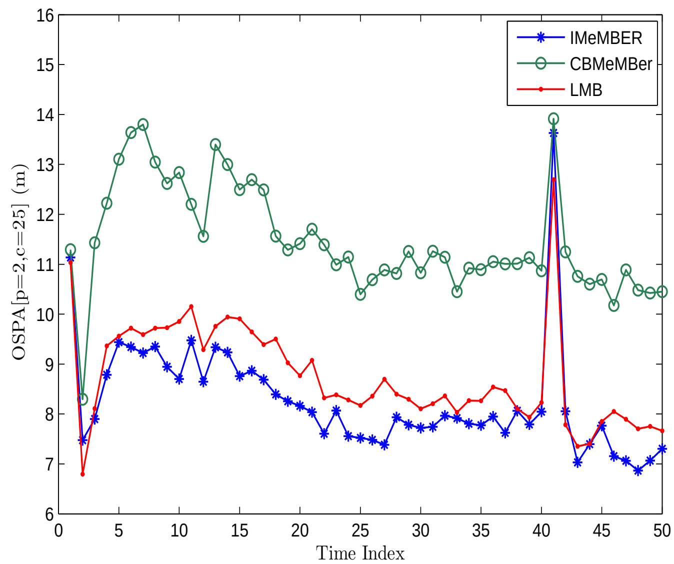

Figure 2.4: The average OSPA metrics over 500 Monte Carlo runs for the SMC-IMeMBer, SMC-CBMeMBer, and SMC-LMB filters.

The spatial probability density function of an actual target in the IMemBer filter, i.e., (2.29) is different from that in the CBMeMBer filter. To evaluate the localization performances of these two filters, the localization component of the OSPA metric can be examined. The contribution of the localization error to the OSPA metric is computed after determining optimal assignment as [25, 27]

$$e_{p,loc}^{(c)}\left(X,Y\right) = \left(\frac{1}{n}\min_{\pi\in\Pi_n}\sum_{i=1}^{m}\min\left(c, d\left(x_i, y_{\pi(i)}\right)\right)^p\right)^{\frac{1}{p}},$$

where Π_n denotes all permutations between the true RFS $X = \{x_1, ..., x_n\}$ and the estimated RFS $Y = \{y_1, ..., y_m\}$ for $m \leq n$, and $d\left(x_i, y_{\pi(i)}\right)$ is the base distance, that corresponds to the Euclidean norm for $p = 2$. In OSPA metric, the base distance is mitigated by the cut-off parameter c if the distance between two points exceeds *c.* In Fig. 2.5, the localization components of the average OSPA metrics and the RMSEs show that the IMeMBer filter produces more accurate state estimates than the CBMeMBer filter. In addition, it can be seen that the IMeMBer and LMB filters have similar localization performance. Based on this result, the performance difference between the IMemBer and LMB filters arises from the stability of their cardinality estimates.

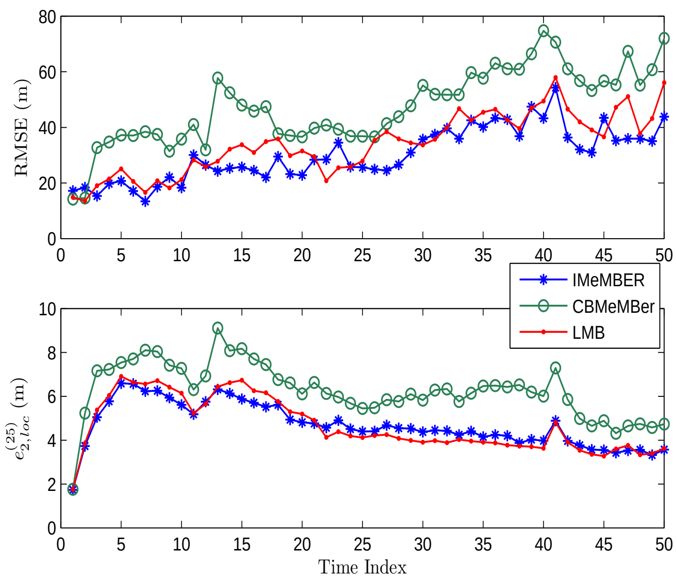

Figure 2.5: The average localization errors over 500 Monte Carlo runs for the SMC-IMeMBer, SMC-CBMeMBer, and SMC-LMB filters.

2.7.2 Linear-Gaussian Multi-target Tracking Example

In the last example, the GM-IMeMBer filter is compared with the GM-CBMeMBer and GM-LMB filters. For the ease of implementation, rather than using (2.50) the detection probability is set to constant values. In addition, the target-originated measurements are linear-Gaussian processes modeled as

$$z_k = Hy_k + \eta_k,$$

where η_k is the zero-mean Gaussian measurement noise with covariance matrix $R_k = diag([1.0,\, 1.0])$m. Thus, the positions of the targets are monitored, i.e., the observation matrix is given by $H = [I_{2\times 2},\, 0_{2\times 2}]$, where $I_{2\times 2}$ and $0_{2\times 2}$ denoting the 2×2 identity and zero matrices, respectively. For computational efficiency of the GM implementation with the maximum of 25 components (i.e., $J_{max} = 25$), the termination, pruning and merging thresholds are set to $Th = 10^{-2}$, $Pr = 10^{-3}$, and $U = 2.5$, respectively [28]. All other model and scenario parameters are the same as those in Section 2.7.1.

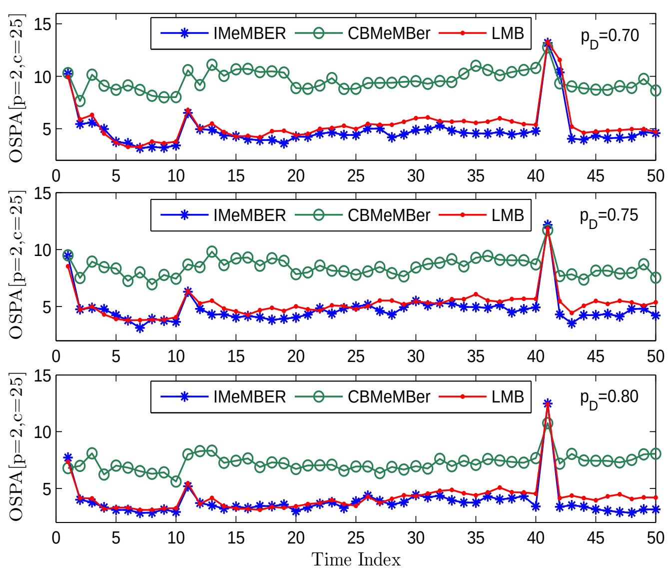

Figure 2.6: The average OSPA metrics over 500 Monte Carlo runs for the GM-IMeMBer, GM-CBMeMBer, and GM-LMB filters at different detection probabilities.

The performances of the three filters were measured by running the same scenario in Fig. 2.1 for 500 Monte Carlo runs. In Fig. 2.6, the average OSPA metrics show the performances of these filters for three different values of the detection probability. It can be seen the GM-IMeMBer and LMB filters have the same performance as p_D takes higher values, and they outperform the GM-CBMeMBer filter. In addition, abrupt changes are observed for the GM-IMeMBer and LMB filters after time step $k = 40$. As explained in Section 2.7.1, this is due to their slow response to the target death. Finally, Fig. 2.6 demonstrates that the GM-IMeMBer filter successfully deals with the crossing targets since no abrupt changes due to missing targets are observed at time steps $k = 21$, $k = 25$, and $k = 29$.

In Fig. 2.7, the localization components of the average OSPA metrics are shown. It can be seen that the IMeMBer and LMB filters have the same localization performance, and they produce more accurate state estimates than the CBMeMBer filter.

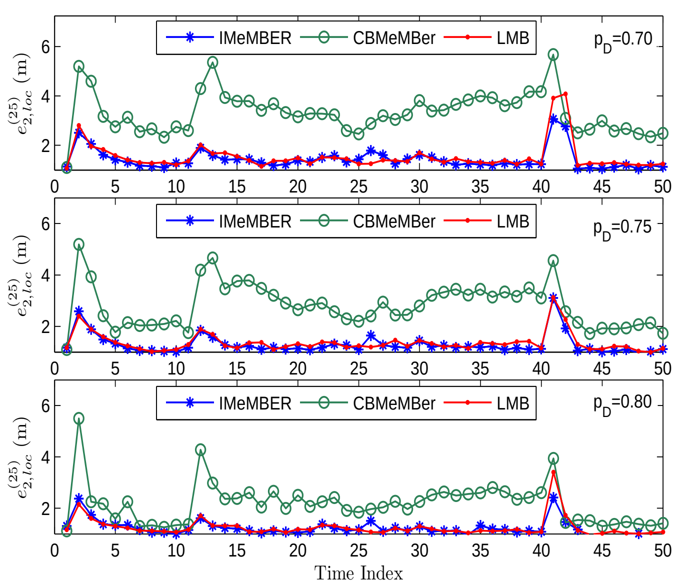

Figure 2.7: The average localization errors over 500 Monte Carlo runs for the GM-IMeMBer, GM-CBMeMBer, and GM-LMB filters.

2.8 Conclusion

The CBMeMBer filter removes the positive bias from the data-updated cardinality estimate in the MeMBer filter. The derivation of the CBMeMBer filter depends on the computation of the cardinality from the exact p.g.fl. without making the second approximation in [14]. In this paper, an alternative derivation of the MeMBer data update process from another perspective is presented. In contrast to the CBMeMBer filter, the new MeMBer data update process follows the two approximations made in the derivation of the MeMBer filter. Then, it is extended to model spurious targets arising from targets under the legacy track set. The modeling of spurious targets yields an unbiased MeMBer filter, referred to as the IMeMBer filter.

In the IMeMBer filter, the spatial probability density function of an actual target differs from that computed for a target under data-induced track set in the CBMeMBer filter. However, its formulation relaxes the limiting assumption on the probability of target detection required for the derivation of valid spatial probability density functions in the CBMeMBer filter. In addition, the modeling of spurious targets is utilized to refine the existence probabilities before multi-target state estimation. Thus, the stability of cardinality estimate is improved. Simulation results were provided to demonstrate the effectiveness of the proposed filter.

2.9 Appendix A

The gradient derivative of the p.g.fl. $G_{U,k}[z;h]$ with respect to a target state y is given by

$$\begin{aligned}\frac{\partial G_{U,k}}{\partial \delta y}\left[z;h\right] =& \Psi^{-1}\sum_{i=1}^{M_{k|k-1}} q_{k|k-1}^{(i)} f_{k|k-1}^{(i)}\left(y\right)\times \\ & \left[\frac{p_{D,k}\left(y\right) g_k\left(z\left|y\right.\right)}{\left(1-q_{k|k-1}^{(i)}+q_{k|k-1}^{(i)} f_{k|k-1}^{(i)}\left[hp'_{D,k}\right]\right)}- \right. \\ & \left. \frac{q_{k|k-1}^{(i)} f_{k|k-1}^{(i)}\left[hp_{D,k}g_k\left(z\left|\cdot\right.\right)\right] p'_{D,k}\left(y\right)}{\left(1-q_{k|k-1}^{(i)}+q_{k|k-1}^{(i)} f_{k|k-1}^{(i)}\left[hp'_{D,k}\right]\right)^2}\right],\end{aligned} \tag{2.9.1}$$

where $p'_{D,k}\left(y\right) = 1 - p_{D,k}\left(y\right)$ and $\Psi = \kappa(z) + \sum_{i=1}^{M_{k|k-1}} G_{U,k}^{(i)}[z;1]$. Note that (2.9.1) does not in general produce the same results when it is evaluated at $h=0$ and $h=1$.

Now, take the gradient derivative under the prior assumption of $p_{D,k}(y) = 1$. Then,

$$\frac{\partial G_{U,k}}{\partial \delta y}\left[z;h\right] = \Psi^{-1}\sum_{i=1}^{M_{k|k-1}} \frac{q_{k|k-1}^{(i)} f_{k|k-1}^{(i)}\left(y\right) g_k\left(z\left|y\right.\right)}{\left(1-q_{k|k-1}^{(i)}\right)}. \tag{2.9.2}$$

Since (2.9.2) is independent of h, it produces the same results when it is evaluated at $h=0$ and $h=1$.

Again, take the gradient derivative but, now under the prior assumption of setting $h=1$ in the denominator of $G_{U,k}^{(i)}[z;h]$, i.e., by following (2.11) as

$$\frac{\partial G_{U,k}}{\partial \delta y}\left[z;h\right] = \Psi^{-1}\sum_{i=1}^{M_{k|k-1}} \frac{q_{k|k-1}^{(i)} f_{k|k-1}^{(i)}\left(y\right) p_{D,k}\left(y\right) g_k\left(z\left|y\right.\right)}{1-q_{k|k-1}^{(i)} f_{k|k-1}^{(i)}\left[p_{D,k}\right]}, \tag{2.9.3}$$

Since (2.9.3) is independent of h, it produces the same results when it is evaluated at $h = 0$ and $h = 1$.

2.10 Appendix B

In [33], the PHD function of the updated p.g.fl. is given by (see (34) and (35) in [33])

$$D_{k|k-1}(y;z) = \frac{\sum_{i=1}^{M_{k|k-1}} D_{U,k}^{(i)}(y;z)}{\kappa(z) + \sum_{i=1}^{M_{k|k-1}} G_{U,k}^{(i)}[1;z]}, \tag{2.10.1}$$

where the negative part of $D_{U,k}^{(i)}(y;z)$ is

$$\tilde{D}_{U,k}^{(i)}(y;z) = \frac{q_{k|k-1}^{(i)2} f_{k|k-1}^{(i)}[p_{D,k} g_k(z|\cdot)](1 - p_{D,k}(y))}{\left(1 - q_{k|k-1}^{(i)} f_{k|k-1}^{(i)}[p_{D,k}]\right)^2} f_{k|k-1}^{(i)}(y). \tag{2.10.2}$$

Using (2.16), the existence probability from (2.10.2) for $i = 1, ..., M_{k|k-1}$ can be computed as

$$\begin{aligned} \tilde{q}_{U,k}(z) &= \int \tilde{D}_{k|k-1}(y;z)\, dy, \\ &= \frac{\sum_{i=1}^{M_{k|k-1}} \frac{q_{k|k-1}^{(i)2} f_{k|k-1}^{(i)}[p_{D,k} g_k(z|\cdot)]\left(1 - f_{k|k-1}^{(i)}[p_{D,k}]\right)}{\left(1 - q_{k|k-1}^{(i)} f_{k|k-1}^{(i)}[p_{D,k}]\right)^2}}{\kappa(z) + \sum_{i=1}^{M_{k|k-1}} G_{U,k}^{(i)}[1;z]}. \end{aligned} \tag{2.10.3}$$

Then, substituting (2.4) into (2.10.3) yields the same existence probability as that computed for a spurious target, i.e., (2.27). Thus, spurious targets appear as targets that are detected under the legacy track set. This is another explanation of the premise that results in (2.25).

2.11 Appendix C

The posterior p.g.fl. of the multi-Bernoulli RFS has the form

$$G_k[h] = G_{L,k}[h] \prod_{z \in Z_k} G_{U,k}[h; z]. \tag{2.11.1}$$

Using the product rule, the PHD of $G_k[h]$ is derived by taking its first order set derivative and then evaluating it at $h = 1$, i.e.,

$$\begin{aligned} D_k(y) &= \frac{\delta G_k}{\delta y}[h]\bigg|_{h=1}, \\ &= \frac{\delta G_{L,k}}{\delta y}[h]\bigg|_{h=1} \prod_{z \in Z_k} G_{U,k}[z; h]\bigg|_{h=1} + \\ &\quad G_{L,k}[h]|_{h=1} \sum_{z \in Z_k} \frac{\delta G_{U,k}}{\delta y}[z; h]\bigg|_{h=1}, \end{aligned} \tag{2.11.2}$$

where $G_{L,k}$ and $G_{U,k}$ are p.g.fl.s of targets under the legacy and data-induced track sets and are defined by (2.37), and (2.38), respectively for a Bernoulli RFS. For the two possible cases defined for a Bernoulli target, i.e., $Z_k = \emptyset$ and $Z_k = \{z\}$, considering the fact that a single target cannot be classified as a target under legacy and data-induced track sets at the same time, (2.11.2) simplifies to

$$\begin{aligned} D_k(y) &= \frac{\delta G_k}{\delta y}[h]\bigg|_{h=1}, \\ &= \frac{\delta G_{L,k}}{\delta y}[h]\bigg|_{h=1} G_{U,k}[z; h]|_{\substack{h=1 \\ q_{L,k=1}}} + \\ &\quad G_{L,k}[h]|_{h=1} \frac{\delta G_{U,k}}{\delta y}[z; h]\bigg|_{\substack{h=1 \\ q_{L,k=0}}}, \end{aligned} \tag{2.11.3}$$

For $Z_k = \emptyset$, the derivation of the PHD can be found in [14, 33]. On the other hand, for $Z_k = \{z\}$, the evaluation of the expression in (2.11.3) using p.g.fl.s given by (2.37) and (2.38) results in

$$\begin{aligned} D_k(y) = q_{k|k-1} \Bigg[& \frac{\frac{1-p_{D,k}(y)}{1-q_{k|k-1} f_{k|k-1}\left[p_{D,k}\right]}}{1+\frac{q_{k|k-1} f_{k|k-1}\left[p_{D,k} g_k(z|\cdot)\right]}{1-q_{k|k-1} f_{k|k-1}\left[p_{D,k}\right]} \times \frac{1}{\kappa(z)}} + \\ & \frac{\frac{p_{D,k}(y) g_k(z|y)}{1-q_{k|k-1} f_{k|k-1}\left[p_{D,k}\right]} \times \frac{1}{\kappa(z)}}{1+\frac{q_{k|k-1} f_{k|k-1}\left[p_{D,k} g_k(z|\cdot)\right]}{1-q_{k|k-1} f_{k|k-1}\left[p_{D,k}\right]} \times \frac{1}{\kappa(z)}} \Bigg] f_{k|k-1}(y) . \end{aligned} \tag{2.11.4}$$

After some algebraic manipulations, (2.11.4) can be rewritten as follows:

$$D_k(y) = \frac{1-p_{D,k}(y)+p_{D,k}(y)\frac{g_k(z|y)}{\kappa(z)}}{q_{k|k-1}^{-1} - f_{k|k-1}\left[p_{D,k}\right] + \frac{f_{k|k-1}\left[p_{D,k} g_k(z|\cdot)\right]}{\kappa(z)}} f_{k|k-1}(y) . \tag{2.11.5}$$

Finally, dividing the PHD by the corresponding existence probability given by (2.47) yields the claimed result, i.e.,

$$f_{U,k}(y) = \frac{1-p_{D,k}(y)+p_{D,k}(y)\frac{g_k(z|y)}{\kappa(z)}}{1 - f_{k|k-1}\left[p_{D,k}\right] + \frac{f_{k|k-1}\left[p_{D,k} g_k\right]}{\kappa(z)}} f_{k|k-1}(y) . \tag{2.11.6}$$

Bibliography

[1] E.H. Aoki, A. Bagchi, P. Mandal, and Y. Boers. A theoretical analysis of Bayes-optimal multi-target tracking and labelling. Technical report, University of Twente, Netherlands, 2011.

[2] Y. Bar-Shalom, X.R. Li, and T. Kirubarajan. *Estimation with Applications to Tracking and Navigation*. J. Whiley and Sons, New York City NY, 2001.

[3] E. Baser and M. Efe. A novel auxiliary particle PHD filter. In *Proc. IEEE 15th Int'l Conf. Information Fusion*, pages 165–172, Istanbul, Turkey, 2012.

[4] E. Baser, T. Kirubarajan, and M. Efe. Improved MeMBer filter with modeling of spurious targets. In *Proc. IEEE 16th Int'l Conf. Information Fusion*, pages 813–819, Istanbul, Turkey, 2013.

[5] S. Challa, B.-N. Vo, and X. Wang. Bayesian approaches to track existence - IPDA and random sets. In *Proc. IEEE 5th Int'l Conf. Information Fusion*, pages 1228–1235, Annapolis, Maryland, Jul. 2002.

[6] X. Chen, R. Tharmarasa, M. Pelletier, and T. Kirubarajan. Integrated clutter estimation and target tracking using Poisson point process. *IEEE Trans. Aerosp. Electron. Syst.*, 48(2):1210–1235, Apr. 2012.

[7] D. Dunne and T. Kirubarajan. Weight partitioned probability hypothesis density filters. In *Proc. IEEE 14th Int'l Conf. Information Fusion*, pages 1–8, Chicago, IL, 2011.

[8] O. Erdinc, P. Willett, and Y. Bar-Shalom. Probability hypothesis density filter for multitarget multisensor tracking. In *Proc. IEEE 8th Int'l Conf. Information Fusion*, pages 146–153, Philadelphia, PA, 2005.

[9] A. Gning, B. Ristic, and L. Mihaylova. Bernoulli particle/box filters for detection and tracking in the presence of triple measurement uncertainty. *IEEE Trans. Signal Process.*, 60(5):2138–2151, Jan. 2012.

[10] W. Liu, C. Han, and H. Zhu. Multitarget state extraction for the PHD filter using MCMC approach. *IEEE Trans. Aerosp. Electron. Syst.*, 46(2):864–883, Apr. 2010.

[11] R. Mahler. Multitarget Bayes filtering via first-order multitarget moments. *IEEE Trans. Aerosp. and Electron. Syst.*, 39(4):1152–1178, Oct. 2003.

[12] R. Mahler. 'Statistics 101' for multisensor, multitarget data fusion. *IEEE Aerosp. and Electron. Sys. Magazine Part 2: Tutorials*, 19(1):53–64, Jan. 2004.

[13] R. Mahler. PHD filters of higher order in target number. *IEEE Trans. Aerosp. Electron. Syst.*, 43(4):1523–1542, Oct. 2007.

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

[15] R. Mahler. 'Statistics 102' for multisource-multitarget detection and tracking. *IEEE J. Sel.Topics Signal Process.*, 7(3):376–388, Jun. 2013.

[16] R. Mahler. *Advances in Statistical Multisource-Multitarget Information Fusion.* Artech House, Norwood MA, 2014.

[17] R. Mahler and A. El-Fallah. CPHD and PHD filters for unknown backgrounds, III: Tracktable multitarget filtering in dynamic clutter. In *Proc. SPIE Signal and Data Processing of Small Targets*, pages 76980F–76980F–12, Apr.

[18] R. Mahler, B.-T. Vo, and B.-N. Vo. CPHD filtering with unknown clutter rate and detection profile. *IEEE Trans. Signal Process.*, 59(8):3497–3513, Mar. 2011.

[19] D. Musicki, R. Evans, and S. Stankovic. Integrated probabilistic data association. *IEEE Trans. Automatic Control*, 39(6):1237–1241, Jun. 1994.

[20] K. Panta, B.-N. Vo, and S. Singh. Novel data association schemes for the probability hypothesis density filter. *IEEE Trans. Aerosp. Electron. Syst.*, 43(2):556–570, Apr. 2007.

[21] V. C. Ravindra, L. Svensson, L. Hammarstrand, and M. Morelande. A cardinality preserving multitarget multi-bernoulli RFS tracker. In *Proc. IEEE 15th Int'l Conf. Information Fusion*, pages 832–839, Singapore, Jul. 2012.

[22] S. Reuter, B.-T. Vo, B.-N. Vo, and K. Dietmayer. The labeled multi-Bernoulli filter. *IEEE. Trans. Signal Process.*, 62(12):3246–3260, Jun. 2014.

[23] B. Ristic, D. Clark, and B.-N. Vo. Improved SMC implementation of the PHD filter. In *Proc. IEEE 13th Int'l Conf. Information Fusion*, pages 1–8, Edinburgh, UK, 2010.

[24] B. Ristic, D. Clark, B.-N. Vo, and B.-T. Vo. Adaptive target birth intensity in

PHD and CPHD filters. *IEEE Trans. Aerosp. Electron. Syst.*, 48(2):1656–1668, Apr. 2012.

[25] B. Ristic, B.-N. Vo, D. Clark, and B.-T. Vo. A metric for performance evaluation of multi-target tracking algorithms. *IEEE. Trans. Signal Process.*, 59(7):3452–3457, Jul. 2011.

[26] B. Ristic, B.-T. Vo, B.-N. Vo, and A. Farina. A tutorial on Bernoulli filters: Theory, implementation and applications. *IEEE Trans. Signal Process.*, 61(13):3406–3430, Jul. 2013.

[27] D. Schuhmacher, B.-T. Vo, and B.-N. Vo. A consistent metric for performance evaluation of multi-object filters. *IEEE. Trans. Signal Process.*, 56(7):3447–3457, Aug. 2008.

[28] B.-N. Vo and W.-K. Ma. The Gaussian mixture probability hypothesis density filter. *IEEE Trans. Signal Process.*, 54(11):4091–4104, Nov. 2006.

[29] B.-N. Vo, S. Singh, and A. Doucet. Sequential Monte Carlo methods for multi-target filtering with random finite sets. *IEEE Trans. Aerosp. Electron. Syst.*, 41(41):1224–12245, Oct. 2005.

[30] B.-T. Vo. *Random finite sets in multi-object filtering.* PhD thesis, Electrical, Electronic and Computer Engineering, University of Western Australia, Australia, 2008.

[31] B.-T. Vo and B.-N. Vo. Labeled random finite sets and multi-object conjugate priors. *IEEE. Trans. Signal Process.*, 61(13):3460–3475, Jul. 2013.

[32] B.-T. Vo, B.-N. Vo, and A. Cantoni. Analytic implementations of the cardinalized probability hypothesis density filter. *IEEE Trans. Signal Process.*, 55(7):3553–3567, Jul. 2007.

[33] B.-T. Vo, B.-N. Vo, and A. Cantoni. The cardinality balanced multitarget multi-Bernoulli filter and its implementations. *IEEE Trans. Signal Process.*, 57(2):409–423, Feb. 2009.

[34] B.-T. Vo, B.-N. Vo, R. Hoseinnezhadz, and R. Mahler. Robust multi-Bernoulli filtering. *IEEE J. Sel. Topics Signal Process.*, 7(3):399–409, Jun. 2013.

[35] J. L. Williams. Alternative multi-bernoulli filters (extended version). *CoRR*, abs/1203.2995, 2012. `http://arxiv.org/abs/1203.2995`.

[36] J. L. Williams. Hybrid Poisson and multi-Bernoulli filters. In *Proc. IEEE 15th Int'l Conf. Information Fusion*, pages 1103–1110, Singapore, Jul. 2012.

The following chapter is a reproduction of a peer-reviewed article in the Institute of Electrical and Electronics Engineers (IEEE):

Erkan Baser, Mike McDonald, Thia Kirubarajan, and Murat Efe, A Joint Multi-target Estimator for the Joint Target Detection and Tracking Filter, *IEEE Transactions on Signal Processing*, no. 15, vol. 63, pp. 3857–3871, May 2015.

Chapter 3

A Joint Multitarget Estimator for the Joint Target Detection and Tracking Filter

3.1 Abstract

This paper proposes a joint multitarget (JoM) estimator for the joint target detection and tracking (JoTT) filter. An efficient choice to the unknown JoM estimation constant (i.e., hypervolume around target state estimate) is proposed as a Pareto-optimal solution to a multi-objective nonlinear convex optimization problem. The multi-objective function is formulated as two convex objective functions in conflict. The first objective function is the information theoretic part of the problem and aims for entropy maximization, while the second one arises from the constraint in the definition of the JoM estimator and aims to improve the accuracy of the JoM estimates.

The Pareto-optimal solution is obtained using the weighted sum method, where objective weights are determined as linear predictions from autoregressive models. In contrast to the marginal multitarget (MaM) estimator, the "target-present" decision from the JoM estimator depends on the spatial information as well as the cardinality information in the finite-set statistics (FISST) density. The simulation results demonstrate that the JoM estimator achieves better track management performance in terms of track confirmation latency and track maintenance than the MaM estimator for different values of detection probability. However, the proposed JoM estimator suffers from track termination latency more than the MaM estimator since the localization performance of the JoTT filter does deteriorate gradually after target termination.

3.2 Introduction

Target tracking is the process of estimating the state of a dynamic object by filtering noisy measurements in the presence of false alarms and missed detections. The whole process can be divided into track confirmation, track maintenance, and track termination functions. Hence, it is necessary to verify the existence of the target from the received measurements. A number of statistical algorithms have been proposed for the detection and tracking of single (or multiple) target(s) [2]. A recent innovation in the area of target detection and tracking is in the application of the Random Finite Sets (RFS) using the finite-set statistics (FISST) [27, 26].

The RFS formalism of the Bayesian multitarget filter provides a formal mechanism for propagating and updating FISST densities. Using the Almost Parallel Worlds Principle (APWOP) along with the relationship between the FISST probability and the measure theoretic probability, some statistical concepts and techniques in filtering theory and information theory can be established for the RFS formalism [27, 26]. However, the conventional single target state estimators (e.g., the maximum a posteriori (MAP) estimator and the expected a posteriori (EAP) estimator) are undefined for RFS based multitarget filters [27, 25]. Hence, two Bayesian optimal estimators were proposed to obtain the multitarget states from FISST densities. The first multitarget state estimator is called the marginal multitarget (MaM) estimator. This estimator only considers the cardinality information (i.e., the number of elements of a given RFS) in FISST densities. The second multitarget state estimator is called the joint multitarget (JoM) estimator. This estimator, as its name suggests, considers both the cardinality and spatial information related to multitarget states in FISST densities. These two estimators are Bayesian optimal, i.e., they minimize their Bayes risk functions. Recently, the minimum mean optimal sub-pattern assignment (MMOSPA) estimator in [14] was generalized for the probability hypothesis density (PHD) filter [4]. Thus, a theoretical basis also has been established for the commonly used k-means clustering method.

The multi-Bernoulli assumption on the RFS of targets represents each target independently by a parameter pair $\{q, f\}$ [34]. That is, for each target an independent Bernoulli RFS provides a unified statistical representation of target existence via the probability q and target states via the spatial probability density $f(x)$. Using the multi-Bernoulli RFS representation, tractable approximations of the multitarget Bayes filter, generally known as the multi-target multi-Bernoulli (MeMBer) filters, were developed [27, 39, 3]. In addition, the Bernoulli RFS formalism was used in the development of an exact solution to the single-target tracking problem. First, the integrated probabilistic data association (IPDA) filter [30] was formulated as an RFS based Bayes filter [11]. Then, this RFS formulation was extended by making use of a target birth model, state-dependent detection probability and arbitrary false alarm process in its framework. Thus, the joint target detection and tracking (JoTT) filter (also known as the Bernoulli filter) was developed with the objective of estimating the target existence probability along with its state(s)[27, 38]. For more detailed information regarding the theory, implementation and applications of Bernoulli filters, interested readers are referred to [34].

The performance of tracking algorithms and state estimators can be evaluated by metrics defined in terms of cardinality, time, and accuracy [17, 35]. The performance metrics should be determined according to which attributes of the tracking algorithm or the state estimator are selected to be monitored. For example, the mean OSPA (MOSPA) metric is appropriate to reduce jitters and track coalescence [14, 5]. In addition, they should be consistent with the criteria that the tracking algorithm or state estimator is developed to optimize [14, 23]. Based on these facts, it is important to point out that the estimated states from the JoTT filter using the JoM estimator is identical to that using the MaM estimator if "target-present" decision is confirmed by these two estimators. Therefore, the performance metric(s) should be selected so as to monitor the cardinality and time attributes of these two estimators regarding track confirmation, track maintenance quality after the target birth, and track termination. There are numerous metrics defined in terms of cardinality and time. Nevertheless, the OSPA metric is defined as a rigorous and robust performance measure for the (multi)target Bayes filters [36, 33].

Even though the MaM estimator is used in MeMBer type filters, the exact use of the JoM estimator with these filters has not been studied so far. In this paper, we propose a JoM estimator to obtain the estimate of the target RFS from the JoTT filter. The proper choice to the unknwon JoM estimation constant (i.e., hypervolume around target state estimate) is obtained as a Pareto-optimal solution to a multi-objective nonlinear convex optimization problem. The multi-objective function is formulated as two convex objective functions in conflict. The first objective function is the information theoretic part of the problem and aims for entropy maximization, while the second one arises from the constraint in the definition of the JoM estimator and aims to improve the accuracy of the JoM estimates. The Pareto-optimal solution is obtained using the weighted sum method [31, 28, 18, 29]. This method aggregates two or more objective functions into a single objective function using weights selected according to their relative importance. Then, the resulting single-objective optimization problem can be solved using any standard optimization technique [31, 28].

This paper is organized as follows: Section 3.3 provides the necessary background on information theory and multitarget state estimation. In Section 3.4, the Bayesian optimal multitarget estimators (i.e., MaM and JoM estimators) are presented along with their evaluations for estimation of multitarget states. The proper choice to the JoM estimation constant is formulated in Section 3.5. For its Pareto-optimal solution, linear predictions of objective weights are proposed in Section 3.6. The implementation of the JoM estimator for the JoTT filter under Gaussian assumptions is presented in Section 3.7. Simulation results are shown in Section 3.8. Finally, conclusions and future research directions are given in Section 3.9.

3.3 Background

3.3.1 Concepts in Information Theory

In the following, we introduce some of the basic concepts of information theory. For the sake of completeness and clarity, we also summarize how each concept is utilized later.

Entropy: A random variable is statistically characterized by its probability density function (pdf). In traditional statistics, variance of a random variable is used to measure its uncertainty. However, in the information theoretic sense, entropy is a measure of the amount of uncertainty in a random variable [13]. For a discrete random variable x characterized by the probability mass function (pmf) $p(x)$ over its sample space $\mathcal{X}$, the entropy is computed as

$$H(p) = -\sum_{x \in \mathcal{X}} p(x) \log(p(x)), \tag{3.3.1}$$

where $-\log(p(x))$ is called the self-information obtained by the observation of x. For the continuity of entropy, $0 \log(0) = 0$, and thus zero probability does not change the uncertainty in x.

Entropy is a nonnegative measure, i.e., $H(p) \geq 0$ with the properties that $H(p)$ is maximized if $p(x)$ is uniform, and $H(p) = 0$ if there is no uncertainty in x, i.e., $p(x) = 0$ or 1 [13]. Hence, larger entropy means that less information is available for the realization of a random variable through its pmf [22].

Differential Entropy: For continuous random variables, the information theoretic uncertainty analogous to the entropy is called the differential entropy, and is defined as

$$H(f) = -\int_{s} f(x) \log(f(x))\, dx, \tag{3.3.2}$$

where S is the support set of the continuous pdf $f(x)$. Unlike the entropy, the differential entropy has values in the range $[-\infty, \infty]$. Therefore, its standalone value cannot be interpreted as the amount of uncertainty on a continuous time random variable. Besides, it makes sense within the definition of the following concepts.

Entropy and differential entropy will be utilized to analyze uncertainties related to the cardinality and spatial information in a FISST density, respectively. Thus, we can evaluate how appropriate the MaM and JoM estimators are for estimation of the multitarget states.

Asymptotic Equipartition Property: In information theory, the weak law of large numbers corresponds to asymptotic equipartition property (AEP) [13]. That is, given that $\tilde{x}_1, ..., \tilde{x}_n$ are independent and identically distributed (i.i.d.) random samples from $f(x)$, then the normalized self-information of this sequence weakly converges to the (differential) entropy of $f(x)$ with a small positive tolerance, i.e., $\tau > 0$ if n is large enough to satisfy [13]

$$\Pr\left(\left|-\frac{1}{n}\log f(\tilde{x}_1, ..., \tilde{x}_n) \to H(f)\right| < \tau\right) > 1 - \delta, \tag{3.3.3}$$

where $\delta \to 0$ as $n \to \infty$ (proof is given by Chebyshev's inequality). The collection of these sequences forms typical set A_τ^n. Most of the total probability is contained in this set, i.e., $\Pr(A_\tau^n) > 1 - \tau$ and is almost uniformly distributed [13] as

$$2^{-n(H(f)+\tau)} \leq \Pr(\tilde{x}_1, ..., \tilde{x}_n) \leq 2^{-n(H(f)-\tau)}. \tag{3.3.4}$$

Hence, if any statistical conclusion is drawn for a typical set, it would be true in general with high probability [13]. In addition, the volume of typical set is almost given by [13, 24]

$$Vol\left(A_{\tau}^{n}\right) \approx 2^{nH(f)}. \tag{3.3.5}$$

Then, the larger\smaller the differential entropy is, the more $f(x)$ disperses\concentrates over its support set S. Note that the typical set has the smallest volume, compared to all possible sets that contain most of the total probability [13, 22].

Typical set of a standard Gaussian density will lead us to define another important set, where the sequences of mostly likely state estimates exist. Thus, our aim would be the entropy maximization by defining a uniform density over this set.

Quantization: The relationship between the entropy and the differential entropy is established by quantization. To see this, assume that the range of a continuous random variable x is divided into bins of Δ where $f(x)$ is continuous. Then, the entropy of the quantized random variable is given by

$$\begin{aligned} H(p) &= -\sum_{-\infty}^{\infty} p_i \log\left(p_i\right), \\ &= -\sum_{-\infty}^{\infty} f\left(x_i\right)\Delta \log\left(f\left(x_i\right)\Delta\right), \\ &= -\sum_{-\infty}^{\infty} f\left(x_i\right)\Delta \log\left(f\left(x_i\right)\right) - \log\left(\Delta\right), \end{aligned} \tag{3.3.6}$$

where the first term approaches $-\int_S f(x)\log\left(f(x)\right)$ as $\Delta \rightarrow 0$. Thus, for n bit quantization of a continuous random variable, i.e., $\Delta = 2^{-n}$, the entropy increases with n as

$$H(p) = H(f) + n.$$

This means that in order to represent an n-bit quantized information from $x \sim f(x)$ the average number of bits required is $H(f) + n$ [13].

This concept will be utilized to analyze the entropy of a FISST density when the corresponding RFS is quantized. This analysis demonstrates an important fact about the selection of the JoM estimation constant.

Kullback-Leibler Divergence (Relative Entropy): Kullback-Leibler (KL) divergence is a statistical measure of the difference of a model or a theory based pdf $f(x)$ from a true or a reference pdf $f_t(x)$ on the same support set. If $f_t(x)$ is absolutely continuous with respect to $f(x)$ or $+\infty$ otherwise, KL divergence of $f(x)$ from $f_t(x)$ is defined as

$$\begin{aligned} K(f_t \| f) &= \int f_t(x) \log\left(\frac{f_t(x)}{f(x)}\right) dx, \\ &= \int f_t(x) \log(f_t(x)) dx - \int f_t(x) \log(f(x)) dx, \\ &= H(f_t \| f) - H(f_t), \end{aligned} \tag{3.3.7}$$

where the first term measures the uncertainty introduced by using a model or theory based $f(x)$ instead of the true or reference $f_t(x)$ while the second term is the differential entropy of $f_t(x)$. Hence, the more $f(x)$ resembles $f_t(x)$, the less is the information lost due to using $f(x)$. That is, $K(f \| f_t) \geq 0$ gets smaller values with equality if and only if $f(x) = f_t(x)$.

KL divergence is an important concept used in the development of other consistent concepts in information theory. For example, mutual information is a special case of KL divergence [13], and entropy maximization is in general formulated as the minimization of KL divergence instead of Shannon's entropy given by (3.3.1) and (3.3.2) [22, 19].

With the help of other relevant concepts KL divergence will be utilized to define the information theoretic part of the multi-objective optimization problem.

3.3.2 Multitarget State Estimation

In the following, we exemplify the problems of the MAP and EAP estimators when they are generalized for estimation of multitarget states. Then, we define the global MAP estimators, i.e., the GMAP-I and GMAP-II estimators, which were introduced in [16] and also known as the MaM and JoM estimators in [27, 25], respectively.

Consider the scenario in [27, 26], where a Bernoulli target moves in the one dimensional interval $[0, 2]$ with units given in meters. In addition, suppose that the target existence probability is set to 0.5 and if the Bernoulli target does exist, its spatial probability density is uniform over $[0, 2]$. That is, suppose that the FISST density in units of $\mathrm{m}^{-|\mathrm{X}|}$ is

$$f(X) = \begin{cases} 0.5, \text{ if } X = \emptyset \\ 0.25\,\mathrm{m}^{-1}, \text{ if } \begin{cases} X = \{x\} \\ 0 \leq x \leq 2 \end{cases} \\ 0, \text{ otherwise} \end{cases}$$

First, we try to obtain the MAP estimate using $X^{MAP} = \arg\sup_{X} f(X)$. However, the MAP estimator is undefined since $f(\emptyset) = 0.5$ cannot be compared with $f(\{x\}) = 0.25\,\mathrm{m}^{-1}$. This problem would be eliminated by converting $f(X)$ into a unitless quantity by multiplying it with $\mathrm{m}^{|\mathrm{X}|}$. Thus, we obtain the MAP estimate as $X^{MAP} = \emptyset$. However, this conversion results in a paradox. That is, if the Bernoulli target moved in the same interval with units given in kilometer instead of meter, this would result in $f(\{x\}) = 250\,\mathrm{m}^{-1}$. Thus, we would obtain the MAP estimate as $X^{MAP} = \{x\}$ after the conversion. That is, the change in unit of measurements from m to km also changes the MAP estimate [27, 26].

Now, using the set integral we try to obtain the EAP estimate from

$$
\begin{aligned}
X^{EAP} &= \int X f(X)\,\delta X, \\
&= \emptyset f(\emptyset) + \int_{0}^{2} x f(\{x\})\,dx, \\
&= 0.5\,(\emptyset + 1\,\mathrm{m})\,.
\end{aligned}
$$

As indicated in [27, 26], the EAP estimator faces additional problems arising from ill-defined arithmetic operations on sets. Therefore, like the MAP estimator, the EAP estimator is undefined when generalized for estimation of multitarget states.

The GMAP-I and GMAP-II are Bayesian estimators, which are defined according to the minimization of the following cost functions [16]

$$
C_0(X,Y) = \begin{cases} 0, \text{ if } |X| = |Y| \\ 1, \text{ if } |X| \neq |Y| \end{cases} \tag{3.3.8}
$$

and

$$
C(X,Y) = C_0(X,Y) + C_1(X,Y)\,, \tag{3.3.9}
$$

respectively. The second cost function in (3.3.9) takes into account the spatial information in a FISST density, i.e.,

$$C_1(X,Y) = \begin{cases} 0, \text{ if } \begin{cases} s = r, \\ (\ell_1, ..., \ell_s) = (\varphi_{\beta_1}, ..., \varphi_{\beta_r}), \\ (x_1, ..., x_s) = (y_{\beta_1}, ..., y_{\beta_r}) \in K \end{cases} \\ 1, \text{ otherwise} \end{cases}$$

where the hybrid RFSs are defined as $X = \{\xi_1, ..., \xi_s\}$ and $Y = \{\zeta_1, ..., \zeta_r\}$ with their identities $(\ell_1, ..., \ell_s)$ and $(\varphi_1, ..., \varphi_r)$, i.e., $\xi_i = (x_i, \ell_i)$ for $i = 1, ..., s$ and $\zeta_i = (y_i, \varphi_i)$ for $i = 1, ..., r$. The RFSs consisting of $\forall x, y \in \mathbb{R}^n$ are surrounded by a closed ball K in $(\mathbb{R}^n)^r$ and are associated through a one-to-one function given by $\beta : (\ell_1, ..., \ell_s) \rightarrow (\varphi_1, ..., \varphi_r)$. Thus, the cost function in (3.3.8) just weights the cardinality discrepancy, whereas the cost function in (3.3.9) weights both the cardinality and spatial discrepancies. These properties of the GMAP-I and GMAP-II estimators will help us in evaluating the corresponding MaM and JoM estimators for estimation of multitarget states.

3.4 Multitarget Bayes Estimators

For RFSs with different cardinalities, their FISST densities have incommensurable scales (i.e., different physical dimensions). Furthermore, addition and subtraction operations on RFSs are not defined properly. Therefore, the multitarget analogues of the MAP and EAP estimators are undefined [27, 26, 25, 16]. Nevertheless, two MAP like multitarget estimators were proposed for FISST densities. In the following, we show how multitarget states are obtained using these Bayes estimators. In addition, we evaluate how appropriate they are for this purpose based on the results obtained from the analysis of uncertainties related to the cardinality and spatial information in a FISST density.

Marginal Multitarget (MaM) Estimator: The MaM estimate of an RFS is computed in a two-step procedure: first, the MAP estimate of the cardinality is determined:

$$\hat{n}^{MAP} \triangleq \arg\sup_{n} p_{|X|}\left(n\right), \tag{3.4.1}$$

where $|X|$ denotes the cardinality variable for the RFS X and is characterized by its probability mass function. That is, the cardinality distribution of the RFS X, given that $Z^{(k)}$ is the RFS of measurements at time k, is

$$p_{|X|}\left(n\right) \triangleq \frac{1}{n!}\int f_{k|k}\left(\{x_1, ..., x_n\}\left|Z^{(k)}\right.\right) dx_1...dx_n. \tag{3.4.2}$$

Then, the MAP estimate of the multitarget states is determined from the corresponding FISST posterior density for the given cardinality estimate $n = \hat{n}^{MAP}$ as

$$\hat{X}^{MaM} = \arg\sup_{x_1,...,x_{\hat{n}^{MAP}}} f_{k|k}\left(\{x_1, ..., x_{\hat{n}^{MAP}}\}\left|Z^{(k)}\right.\right). \tag{3.4.3}$$

The MaM estimator is Bayesian optimal [27, 25, 16]. However, it does not utilize all the information contained in the multitarget posterior density. Hence, it would be statistically unreliable when the target number is related to the spatial information in the FISST posterior density [27, 25]. That is, using the relationship between the FISST probability and measure theoretic probability, the differential entropy of an

RFS X is given by [32, 15]

$$\begin{aligned} H\left(f_X\right) &= -\int f\left(X\right)\log\left(v^{|X|}f\left(X\right)\right)\delta X, \\ &= -\sum_{n=0}^{\infty}\frac{1}{n!}\int f\left(\{x_1,...,x_n\}\right)\times \\ &\qquad\qquad \log\left(v^n f\left(\{x_1,...,x_n\}\right)\right)dx_1...dx_n, \end{aligned} \tag{3.4.4}$$

where $v^{-|X|}$ is the unit of the FISST density $f\left(X\right)$. Note that the dependence of the FISST posterior density on the RFS $Z^{(k)}$ is dropped here for conciseness.

Substituting $f\left(\{x_1,...,x_n\}\right) = n!p_{|X|}\left(n\right)f\left(x_1,...,x_n\right)$ into (3.4.4) yields

$$\begin{aligned} H\left(f_X\right) = -\sum_{n=0}^{\infty}p_{|X|}\left(n\right)\int f\left(x_1,...,x_n\right)\times \\ \log\left(n!v^n p_{|X|}\left(n\right)f\left(x_1,...,x_n\right)\right)dx_1...dx_n, \end{aligned} \tag{3.4.5}$$

and, after some algebraic manipulations, the differential entropy may be rewritten as the sum of the three terms, i.e.,

$$\begin{aligned} H\left(f_X\right) = &-\sum_{n=0}^{\infty}p_{|X|}\left(n\right)\log\left(p_{|X|}\left(n\right)\right)\int f\left(x_1,...,x_n\right)dx_1...dx_n+ \\ &-\sum_{n=0}^{\infty}p_{|X|}\left(n\right)\int f\left(x_1,...,x_n\right)\log\left(v^n f\left(x_1,...,x_n\right)\right)dx_1...dx_n+ \\ &-\sum_{n=0}^{\infty}p_{|X|}\left(n\right)\log\left(n!\right)\int f\left(x_1,...,x_n\right)dx_1...dx_n, \end{aligned} \tag{3.4.6}$$

where the first term is the entropy of the cardinality distribution:

$$\begin{aligned} H(p) &= \sum_{n=0}^{\infty} p_{|X|}(n) \log \left(p_{|X|}(n) \right) \int f(x_1, ..., x_n)\, dx_1 ... dx_n, \\ &= - \sum_{n=0}^{\infty} p_{|X|}(n) \log \left(p_{|X|}(n) \right), \end{aligned}$$

and the second term is the average differential entropy of the joint pdf of $x_1, ..., x_n$ over $p_{|X|}(n)$:

$$E[H(f_{X,n})] = \sum_{n=0}^{\infty} p_{|X|}(n) H(f_{x,n}).$$

The probability assigned to the FISST density with cardinality n, i.e., $f_{X,n} = f(\{x_1, ..., x_n\})$, is uniformly distributed among joint pdfs $f_{x,n} = f(\{x_1, ..., x_n\})$ of $n!$ possible vectors for all permutations of $\{x_1, ..., x_n\}$, i.e., $f_{x,n}$ are symmetric joint pdfs of $(x_{\sigma 1}, ..., x_{\sigma n})$, where σ indicates the permutation on the numbers $\{1, ..., n\}$ [27, 15]. Hence, the third term indicates the information uncertainty due to change in the representation from RFSs, i.e., $\{x_1, ..., x_n\}$, to vectors of indistinguishable points, i.e., $(x_1, ..., x_n)$ [32, 15]:

$$E[\log(n!)] = \sum_{n=0}^{\infty} p_{|X|}(n) \log(n!),$$

The MaM estimator's cost function only penalizes the cardinality discrepancy between the true RFS and its estimate [16]. Therefore, the MaM estimator determines multitarget states without considering the uncertainty represented by the second and the third terms in the FISST densities.

Joint Multitarget (JoM) Estimator: In contrast to the MaM estimator, the JoM estimator determines the target number and multitarget states simultaneously from the FISST posterior density [27] as

$$\hat{X}^{JoM} = \arg\sup_{X} f_{k|k}\left(X \left| Z^{(k)}\right.\right) \frac{\varepsilon^{|X|}}{|X|!}, \tag{3.4.7}$$

where the parameter ε denotes a small constant (hereinafter called as the JoM estimation constant) and satisfies that $f\left(\{x_1, ..., x_n\}\right)\varepsilon^n \leq 1$ for all integers $n \geq 0$. However, there is a trade-off in the selection of ε. That is, smaller values of ε yield better accuracy in multitarget state estimates, but with slower convergence to the true multitarget states [27, 25]. In Appendix A, information theoretic analysis demonstrates that the uncertainty in multitarget state estimates cannot be improved by selecting too small values for ε.

Alternatively, the JoM estimator can be performed in a two-step procedure [27]. First, for integer values $n \geq 0$ the MAP estimates of the RFSs are computed from the corresponding posterior FISST densities:

$$\hat{X}^{n} = \arg\sup_{x_1,...,x_n} f\left(\{x_1, ..., x_n\} \left| Z^{(k)}\right.\right). \tag{3.4.8}$$

Then, using $\hat{X}^{n}$ for each n, the JoM estimate is determined as $\hat{X}^{JoM} = \hat{X}^{\hat{n}}$, where $\hat{n}$ denotes the solution to the following maximization problem:

$$\hat{n} = \arg\sup_{n} f\left(\{\hat{x}_1, ..., \hat{x}_n\} \left| Z^{(k)}\right.\right) \frac{\varepsilon^n}{n!}. \tag{3.4.9}$$

Like the MaM estimator, the JoM estimator is Bayesian optimal [27, 25, 16]. However, it is naturally more appropriate for the estimation of multitarget states since its cost function penalizes both discrepancies in cardinality and multitarget

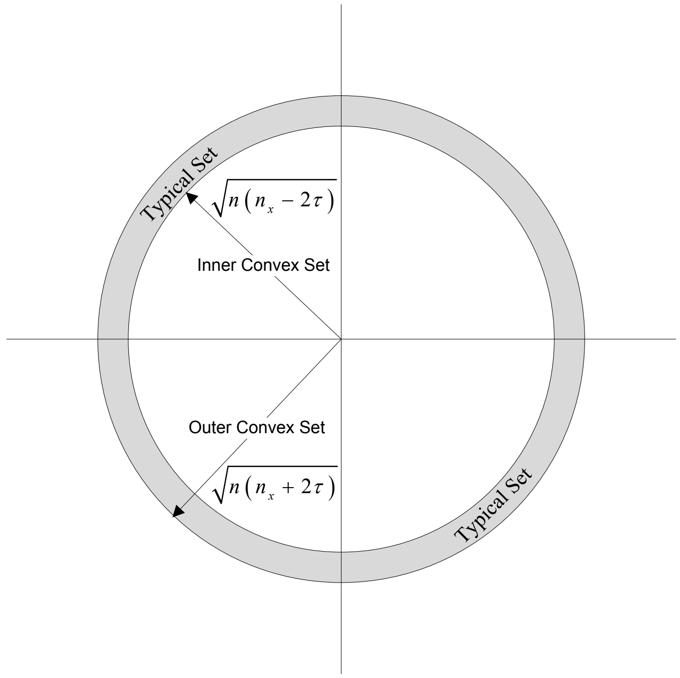

Figure 3.1: The cross section of the typical set of the standard Gaussian density in $\mathbb{R}^{n_x}$.

states [16]. In addition, it is known that the JoM estimator is statistically convergent [27, 25].

3.5 Optimization of the JoM Estimation Constant

The differential entropy of a pdf is roughly represented by a uniform density over its typical set [13, 24]. However, typical sets do not include the sequences of all the most (least) probable state estimates [13, 24]. For example, Fig. 3.1 shows the cross-section of the typical set of a standard Gaussian density around a hypersphere centered at the origin of $\mathbb{R}^{n_x}$ [24, 8]. It can be seen that the typical set is represented by a thin shell bounded by two convex sets (see Appendix B). Instead, for log-concave pdfs (e.g., a Gaussian pdf) superlevel sets can be defined so as to include the sequences of most likely state estimates [8, 10]:

$$S_{\lambda} = \left\{ x \in \mathbb{R}^{n_x} \middle| \, f\left(\tilde{x}_1, ..., \tilde{x}_n\right) \geq \lambda \right\}, \tag{3.5.1}$$

where $\tilde{x}_1, ..., \tilde{x}_n$ are i.i.d. samples drawn from the log-concave pdf $f(x)$, and λ is the supremum value of the uniform probability on the typical set for a small positive constant τ, i.e., $\lambda = e^{-n(H(f)-\tau)}$ [13], where $H(f)$ is in nats. In particular, if x is Gaussian-distributed with mean μ and covariance matrix P in $\mathbb{R}^{n_x}$, i.e., $x \sim N(\mu, P)$, then substituting $H(f) = 0.5 \log\left((2\pi e)^{n_x} |P|\right)$ [13] for λ yields

$$\lambda = \left((2\pi)^{n_x} |P|\right)^{-n/2} e^{-n\left(\frac{n_x}{2} - \tau\right)},$$

and the joint probability distribution of i.i.d. samples are given by

$$\begin{aligned} f\left(\tilde{x}_1, ..., \tilde{x}_n\right) &= \prod_{i=1}^{n} f\left(\tilde{x}_i\right), \\ &= f\left(\hat{x}\right)^n e^{-\frac{1}{2} \sum\limits_{i=1}^{n} (\tilde{x}_i - \mu)^T P^{-1} (\tilde{x}_i - \mu)}, \end{aligned}$$

where $f(\hat{x}) = \left((2\pi)^{n_x} |P|\right)^{-1/2}$.

Thus, the superlevel set given by (3.5.1) can be alternatively defined as

$$S_{\lambda} = \left\{ \tilde{x} \in \mathbb{R}^{n_x} \middle| \, \frac{1}{n} \sum_{i=1}^{n} (\tilde{x}_i - \mu)^T P^{-1} (\tilde{x}_i - \mu) \leq n_x - 2\tau \right\}. \tag{3.5.2}$$

In general, this bounded and closed set includes the sequences of most likely random samples drawn from $f(x)$. However, our aim is to define a confined set that exclusively consists of good state estimates from the JoM estimator. To this end, the superlevel set in (3.5.2), when evaluated at $n = 1$, gives the least upper bound for this special subset as

$$S_{\lambda}^{(1)} = \left\{ x \in \mathbb{R}^{n_x} |\, (x-\mu)^T P^{-1} (x-\mu) \leq n_x - 2\tau \right\}, \tag{3.5.3}$$

where $0 < 2\tau < n_x$. This means that $S_{\lambda}^{(1)}$ is a hyperellipsoid (i.e., a convex set) with the centroid at μ in the region surrounded by the inflection points of the Gaussian density $f(x)$.

The entropy maximization helps ignore spurious details like tail probabilities and side-lobes for which samples from these parts can be hardly ever observed [20]. Over bounded and closed sets, the entropy maximization is achieved by uniform densities [13]. Then, the KL divergence of $f(x)$ from the uniform density defined on $S_{\lambda}^{(1)}$, i.e., $u(x) = \varepsilon_{\lambda}^{-1}$ is given by

$$\begin{aligned} K(u \,\|f) &= \int u(x) \log\left(\frac{u(x)}{f(x)}\right) dx, \\ &= H(u\,\|f) - \log(\varepsilon_{\lambda}), \end{aligned} \tag{3.5.4}$$

where $\log(\varepsilon_{\lambda})$ is the differential entropy of $u(x) = \varepsilon_{\lambda}^{-1}$, i.e., $H(u) = \log(\varepsilon_{\lambda})$, and

$$\begin{aligned} H(u\,\|f) &= -\log(f(\hat{x})) + \frac{1}{2\varepsilon_{\lambda}} \int_{\varepsilon_{\lambda}} (x-\mu)^T P^{-1} (x-\mu) dx, \\ &\leq -\log(f(\hat{x})) + \frac{1}{2}(n_x - 2\tau), \end{aligned}$$

where the last inequality follows from (3.5.3). Thus, the KL divergence in (3.5.4) can be rewritten as

$$K(u\,\|f) \leq -\log(f(\hat{x})\,\varepsilon_{\lambda}) + \frac{1}{2}(n_x - 2\tau), \tag{3.5.5}$$

where the first term on the right hand side is the approximated KL divergence of $f(x)$ from $u(x) = \varepsilon_\lambda^{-1}$ when ε_λ takes so small values, i.e., $n_x - 2\tau \to 0$. Note that the sum on the right hand side of (3.5.5) is always nonnegative since $K(u \| f) \geq 0$ on ε_λ.

The volume of the hyperellipsoid $S_\lambda^{(1)}$ can be expressed in terms of τ as follows [7]:

$$\varepsilon_\lambda = C(n_x)\, |P|^{1/2}\, r^{n_x/2}, \tag{3.5.6}$$

where $r = n_x - 2\tau$ is the critical value for the total probability of $f(x)$ in the hyperellipsoid, and $C(n_x)$ is the volume of the hypersphere with the unit radius in $\mathbb{R}^{n_x}$.

After substituting for ε_λ into (3.5.5), the problem at hand (i.e., determining the optimum volume of the hyperellipsoid) can be formulated as a nonlinear convex optimization problem that determines the optimum value of τ for the least upper bound of the KL divergence. That is,

$$\begin{aligned} &\text{minimize} && f_{o,I}(\tau) = -\log\left(f(\hat{x})\,\varepsilon_\lambda\right) + \frac{1}{2}(n_x - 2\tau), \\ &\text{subject to} && g_1(\tau) = -\tau \leq 0, \\ & && g_2(\tau) = -(n_x - 2\tau) + \gamma_{\min} \leq 0, \end{aligned} \tag{3.5.7}$$

where $\gamma_{\min}$ is a small constant determined according to the chi-square table, considering the degree of freedom (i.e., n_x) and the probability of the confidence level indicating the smallest hyperellipsoid, e.g., $\Pr((n_x - 2\tau) \geq \gamma_{\min}) \geq 95\%$.

The convex optimization problem in (3.5.7) is solely formulated in terms of information theoretic sense. In other words, the objective function $f_{o,I}(\tau)$ in (3.5.7) is minimized as $n_x - 2\tau \to n_x$ (see Appendix C for proof). Thus, the computation of the least upper-bound on the KL divergence through the optimization problem in (3.5.7) corresponds to the minimization of information gain in magnitude measured by

$$K\left(u \,\|\, f\right) \leq H\left(f\right) - \log\left(\varepsilon_{\lambda}\right).$$

In the JoM estimator, the selected hyperellipsoid surrounding the estimated states of targets should satisfy

$$\begin{aligned}\int_{\varepsilon_{\lambda}^{n}} f\left(\{x_1, ..., x_n\}\right) dx_1 ... dx_n &\triangleq \int_{\varepsilon_{\lambda}^{n}} f\left(x_1, ..., x_n\right) dx_1 ... dx_n, \\ &\cong f\left(\hat{x}_1, ..., \hat{x}_n\right) \varepsilon_{\lambda}^{n},\end{aligned} \tag{3.5.8}$$

where the first expression follows from $f\left(\{x_1, ..., x_n\}\right) \triangleq n!\, f\left(x_1, ..., x_n\right)$ and implies that the volume of the hyperellipsoid ε_{λ} for each target should be so small that only one permutation of the RFS is possible in the product space $\varepsilon_{\lambda}^{n}$, i.e., $\{x_1, ..., x_n\} = (x_1, ...x_n)$ [16]. However, as indicated in [27], setting $\varepsilon_{\lambda}^{n}$ to extremely small values would be impractical without considering the information provided by $f\left(x\right)$. In other words, $u\left(x\right)$ would be more informative than $f\left(x\right)$ as $n_x - 2\tau \to 0$. However, this contradicts the information theoretic part of the optimization problem in (3.5.7), which aims for entropy maximization by minimizing information gain obtained using $u\left(x\right)$ instead of $f\left(x\right)$.

In contrast to single-objective optimization, there is usually no unique solution that simultaneously achieves the optimization of more than one objective function. Instead, in multi-objective optimization problems, Pareto-optimal solutions can be computed according to the relative importance of individual objective functions [31, 28].

For a vector of conflicting objective functions given by $F(x) = [f_1(x), ..., f_N(x)]$ a solution x^* is said to be Pareto optimal if there does not exist another solution that dominates it [31]. That is, given that T is the feasible design space, there is no another point, $x \in \mathrm{T}$ satisfying $F(x) \leq F(x^*)$ and $f_i(x) < f_i(x^*)$ for at least one objective function. There are multiple methods for multi-objective optimization problems. However, the conversion of the multi-objective problem into a single-objective problem is the standard way of solving [31, 28].

To determine the optimum value of τ, two objective functions $f_{o,I}(\tau)$ and $f_{o,J}(\tau)$, which quantify entropy maximization and the accuracy of the JoM estimator, respectively, are in conflict with one another. An optimization problem with a single convex objective function can be defined by aggregating them with appropriately selected weights. However, a consistent Pareto-optimal solution to this optimization problem requires the normalization of these conflicting objective functions in different magnitudes [28, 18]. To this end, their extreme values are calculated at the vertex points of the Pareto-optimal set [28]. Specifically, for the problem at hand, first set $\tau = 0$ to obtain the minimum of $f_{o,I}(\tau)$, i.e., $F_{o,I}^{Min}$ while setting $f_{o,J}(\tau)$ to its maximum value, i.e., $F_{o,J}^{Max}$. Then, set $\tau = 0.5\,(n_x - \gamma_{\min})$ to obtain $F_{o,I}^{Max}$ and $F_{o,J}^{Min}$ for $f_{o,I}(\tau)$ and $f_{o,J}(\tau)$, respectively. Finally, the following robust normalization is performed for these conflicting objective functions [28, 18]:

$$f_{o,\xi}^{Trans}(\tau) = \frac{f_{o,\xi}(\tau) - F_{o,\xi}^{Min}}{F_{o,\xi}^{Max} - F_{o,\xi}^{Min}}, \forall \xi \in \{I, J\}.$$

Thus, an optimization problem with a single convex objective function can be obtained as follows:

$$\begin{aligned} \text{minimize} \quad & f_m(\tau) = w_I f_{o,I}^{Trans}(\tau) + w_J f_{o,J}^{Trans}(\tau), \\ \text{subject to} \quad & g_1(\tau) = -\tau \le 0, \\ & g_2(\tau) = -(n_x - 2\tau) + \gamma_{\min} \le 0, \end{aligned} \tag{3.5.9}$$

where $f_{o,J}^{Trans}$ is the normalization of the objective function defined as

$$f_{o,J}(\tau) = \begin{cases} (n_x - 2\tau)^2 & \text{if } (n_x - 2\tau) > \gamma_{\min} \\ 0 & \text{otherwise,} \end{cases}$$

considering the accuracy of the JoM estimator.

In this paper, the weights of the conflicting objectives are determined as linear predictions from autoregressive (AR) models. The next section presents details about this process. However, the weights can also be chosen depending on the application and preference of decision maker(s) [31, 28].

The nonlinear convex optimization problem in (3.5.9) can be solved using any standard nonlinear optimization technique [31]. In addition, the solution is strictly Pareto optimal for the positive weights of the convex objective functions [28, 18] . In this paper, the sequential quadratic programming (SQP) is employed to find a Pareto-optimal solution to (3.5.9). The SQP iteratively solves a quadratic approximation to the Lagrangian function, in the sense that the sequence of solutions approaches to optimal solution satisfying the necessary Karush-Kuhn-Tucker (KKT) conditions [9, 6]. Note that there are many other ways to solve the above multi-objective optimization problem. The contribution of this paper is not in optimization, but in multitarget detection and state estimation. Thus, we have used a standard optimization approach that guarantees a Pareto-optimal solution without exhaustive comparison with other approaches.

In order to illustrate the geometrical interpretation of the weighted sum method, let us examine the nonlinear convex optimization problem in (3.5.9) with the following parameters: $P = \text{diag}\left([50, 50, 10, 10]'\right)$, $n_x = 4$ and $\gamma_{\min} = 0.297$ with the confidence probability of 99.9%. Considering the inequality constraints in (3.5.9) the feasible design space of τ, i.e., $\mathrm{T} = \{\tau \,|g_i(\tau) \leq 0,\ i = 1, 2\}$ is obtained as $\mathrm{T} = [0,\ 1.8515]$ [28, 29]. Thus, the feasible criterion space of the vector of the normalized objective functions, i.e., $F = \left[f_{o,I}^{Trans}(\tau), f_{o,J}^{Trans}(\tau)\right]$ is defined as $\Omega = \{F\,|\tau \in \mathrm{T}\}$ [28, 29]. Fig. 3.2 shows the relationship between the Pareto front and the normalized objective functions in the feasible criterion space. The Pareto front is the set of the non-dominated points, i.e., Pareto-optimal points in the criterion space [28]. As can be seen in Fig. 3.2, the Pareto front is a convex curve. Thus, a Pareto-optimal point can always be obtained depending on the weights of the conflicting objective functions [29, 40]. This is because for a given set of weights, the weighted sum method approximates the Pareto front as a line [40]:

$$f_{o,I}^{Trans}(\tau) = -\frac{w_J}{w_I} f_{o,J}^{Trans}(\tau) + \frac{1}{w_I} f_m(\tau^*),$$

where τ^* denotes a Pareto-optimal solution. For example, the SQP finds the Pareto-optimal solution as $\tau^* = 1.1674$ if the conflicting objective functions are considered equally important, i.e., $w_I = w_J = 0.5$. Thus, the Pareto-optimal point in the feasible design space is computed as $F = [0.1747, 0.1687]$. As expected, the normalized objective functions in conflict are penalized almost equally. In Fig. 3.2, the line with the slope -1 is tangent to the Pareto front at $F = [0.1747, 0.1687]$ and locally approximates the convex Pareto front.

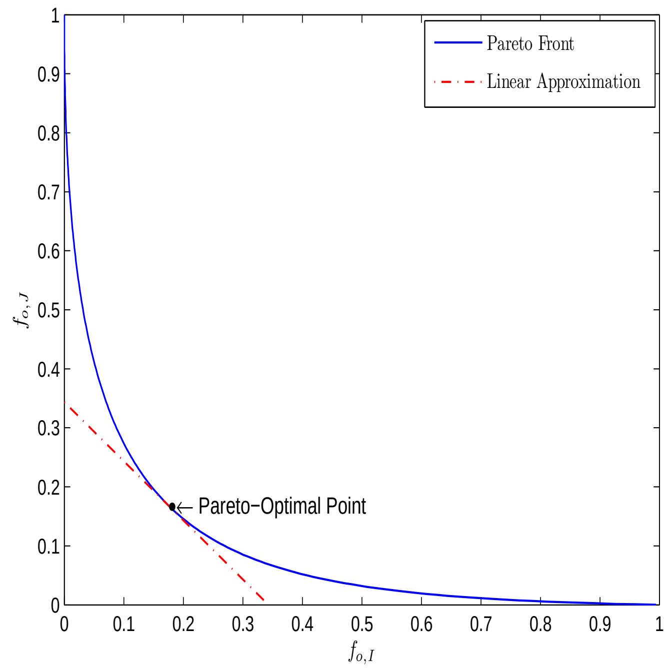

Figure 3.2: Geometrical interpretation of the weighted sum method in the feasible criterion space.

3.6 Linear predictions of objective weights

AR models predict the current output of a stochastic process based on its previous outputs. The AR model of order N, denoted as $AR(N)$, is in general defined by [21]

$$x_k = c + \sum\nolimits_{i=1}^{N} \alpha_i x_{k-i} + \vartheta_k,$$

where c denotes a constant for a non-zero mean value of x_k, $\{\alpha_i\}_{i=1}^{M}$ are predictor coefficients and ϑ_k is a white noise representing prediction error with zero mean and variance σ_{ϑ}^2. For linear predictions of the objective weights, we use the following $AR\,(1)$ model:

$$w_k = c + \alpha w_{k-1} + \vartheta_k, \tag{3.6.1}$$

where the predictor coefficient indicates linear relationship in this time series. For a wide sense stationary (WSS) process, the condition $|\alpha| < 1$ must be satisfied. In this case, the $AR\,(1)$ model is statistically characterized by [21]

$$\begin{aligned} E\left[w_k\right] &= \mu_w = \frac{c}{1-\alpha}, \\ \mathrm{var}\left(w_k\right) &= \sigma_w^2 = \frac{\sigma_{\vartheta}^2}{1-\alpha^2}, \\ \mathrm{cov}\left(w_k, w_{k-i}\right) &= \sigma_w^2 \alpha^i. \end{aligned}$$

Thus, the autocorrelation function between w_k and w_{k-i} decays to zero by α^i as $i \to \infty$. This means that the $AR\,(1)$ model is also stable, i.e., represents a predictable process.

The objective function $f_{o,J}\left(\tau\right)$ in (3.5.9) only considers the degree of freedom, i.e., n_x because of the definition of the hyperellipsoid in (3.5.3). Thus, substituting (3.5.6) into (3.5.8) for a Bernoulli target with parameter pair $\{q_k, f_k\}$ over the volume ε_λ results in

$$\begin{aligned} \int_{\varepsilon_\lambda} f_k\left(\{x\}\right) dx \cong&\ q_k f_k\left(\hat{x}\right) \varepsilon_\lambda, \\ &= q_k \frac{1}{2^{n_x/2}\Gamma\left(\frac{n_x}{2}+1\right)} \left(n_x - 2\tau\right)^{n_x/2}, \end{aligned}$$

where $f_k\left(x\right)$ is a Gaussian pdf and $\Gamma\left(\cdot\right)$ denotes the gamma function. Notice that the approximation is independent of P at time k, denoted as P_k. To consider the covariance of $f_k\left(x\right)$ implicitly in this approximation we determine the degree of correlation between $w_{J,k}$ and $w_{J,k-1}$ as

$$\beta_k = \frac{|P_{k-1}|^{1/2}}{|P_k|^{1/2}} \mathbf{1}_A (q_k),$$

where the first term is the ratio of infinitesimal volumes to locate a Bernoulli target with the same spatial probability at time k and $k-1$, respectively and $\mathbf{1}_A$ denotes an indicator function defined on the set $A = [q_{\min}, 1]$ [27]. The indicator function neglects changes in P_k before confirming a Bernoulli target with the threshold $q_{\min}$. Thus, we keep the weights at their initial states until a probable Bernoulli target is confirmed. In addition, for a stable process the correlation must decay to zero as time lag increases. For this purpose, we set $\alpha = \beta_k$ in (3.6.1) within its control limits as shown in Fig. 3.3.

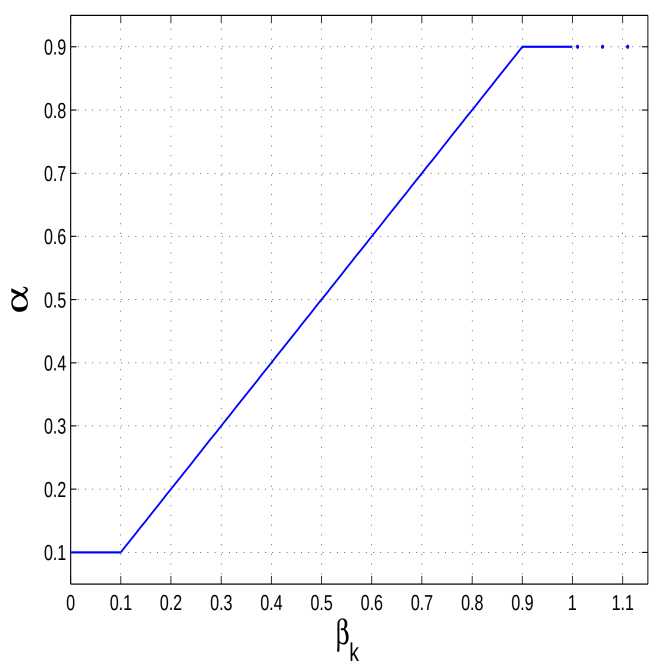

Figure 3.3: Predictor coefficient of AR(1) model versus the degree of correlation between successive weights.

At this point, it is important to note that our $AR\left(1\right)$ model with the predictor coefficient evolving in time does not represent a WSS process. However, it would turn into a WSS process after the optimal JoTT filter converges to its steady-state with detections. Then, the predictor coefficient is set to $\alpha = 0.9$ according to Fig. 3.3 since successive changes in P_k would be small. Thus, the linear predictions monotonically approach to $\mu_{w,J} = 10c_J$, where $0.1 \leq \mu_{w,J} \leq 0.9$ in order to prevent that one objective completely dominates another in the multi-objective optimization. Since $f_k\left(x\right)$ is very peaky after the convergence, $f_{o,J}\left(\tau\right)$ becomes more important than $f_{o,I}\left(\tau\right)$ in (3.5.9). Hence, $\mu_{w,J}$ is set to its maximum value, i.e., $\mu_{w,J} = 0.9$ by $c_J = 0.09$.

Using $w_{I,k} + w_{J,k} = 1$, the the $AR\left(1\right)$ model for $w_{I,k}$ is defined by

$$w_{I,k} = 0.01 + \alpha w_{I,k-1} + \nu_k,$$

where ν_k is a white noise with zero mean and variance $\sigma^2_{w,I} = \sigma^2_{w,J}$ since $\nu_{I,k} = -\vartheta_{J,k}$. Similarly, after the convergence its linear predictions monotonically approach to $\mu_{w,I} = 0.1$.

On the other hand, the optimal JoTT filter gradually deteriorates after target death. Therefore, β_k takes values close to zero and with $\alpha = 0.1$ the linear predictions for $w_{J,k}$ and $w_{I,k}$ monotonically approach to their opposite means, i.e., $\mu_{w,J} = 0.1$ and $\mu_{w,I} = 0.9$, respectively. Consequently, $f_{o,I}\left(\tau\right)$ becomes more important than $f_{o,J}\left(\tau\right)$ in (3.5.9) as $f_k\left(x\right)$ disperses over ε_λ.

3.7 Implementation of the JoM Estimator for the JoTT Filter

Suppose that at most one target is present. In this case, the RFS of a single target can be modeled as a Bernoulli RFS with the parameter pair (q_{k-1}, f_{k-1}). Thus, its FISST density is parameterized as

$$f_{k-1}(X) = \begin{cases} 1 - q_{k-1} & \text{if } X = \emptyset \\ q_{k-1} f_{k-1}(x) & \text{if } X = \{x\}, \end{cases} \tag{3.7.1}$$

where q_{k-1} is the existence probability of the target, and $f_{k-1}(x)$ is its spatial pdf if the target is present.

In the prediction step of the JoTT filter, the FISST density $f_{k-1}(X)$ propagated to time k is parameterized as follows [27, 38]:

$$q_{k|k-1} = p_B (1 - q_{k-1}) + q_{k-1} \int p_{S,k-1}(x) f_{k-1}(x) dx_{k-1}, \tag{3.7.2}$$

$$f_{k|k-1}(x) = \frac{1}{q_{k|k-1}} \left[(1 - q_{k-1}) p_B b_k(x) + q_{k-1} \langle f, p_S \psi \rangle \right], \tag{3.7.3}$$

where a newborn target is declared with probability p_B according to a birth density $b_k(x)$, i.e., the Bernoulli parameter pair (p_B, b_k), and

$$\langle f, p_S \psi \rangle = \int f_{k-1}(x) p_{S,k-1}(x) \psi_{k|k-1}(\cdot | x) dx_{k-1},$$

where $p_{S,k-1}(x)$ is the state-dependent target survival probability and if the target survives, its states evolve according to the Markov state transition density $\psi_{k|k-1}(\cdot | x)$.

Suppose that the single-sensor multitarget measurements at time k are modeled as

$$\mathrm{Z}_k = \Gamma_k\,(x)\mathrm{UC_k},$$

where C_k is the RFS of i.i.d. false alarms and $\Gamma_k\,(x)$ is the Bernoulli RFS of target-originated measurement with the parameter pair $(p_D(x), g_k(z|x))$, where $p_D(x)$ is the detection probability, and $g_k(z|x)$ is the measurement likelihood function.

In the original derivation of the JoTT filter, the false alarm process is modeled as an arbitrary RFS. If the Poisson false alarm RFS with mean rate λ_c and spatial pdf $c\,(z)$ is substituted for the arbitrary false alarm RFS, the original data update equations of the JoTT filter defined in [27, 38] have the form of

$$q_{k|k} = \frac{1 - f_{k|k-1}\,[p_D] + \sum\limits_{z\in Z_k} \frac{f_{k|k-1}\,[p_D g_k(z|\cdot)]}{\kappa(z)}}{q_{k|k-1}^{-1} - f_{k|k-1}\,[p_D] + \sum\limits_{z\in Z_k} \frac{f_{k|k-1}\,[p_D g_k(z|\cdot)]}{\kappa(z)}}, \tag{3.7.4}$$

$$f_{k|k}\,(x) = \frac{1 - p_D\,(x) + p_D\,(x) \sum\limits_{z\in Z_k} \frac{g_k(z|x)}{\kappa(z)}}{1 - f_{k|k-1}\,[p_D] + \sum\limits_{z\in Z_k} \frac{f_{k|k-1}\,[p_D g_k(z|\cdot)]}{\kappa(z)}} f_{k|k-1}\,(x)\,. \tag{3.7.5}$$

where, in general, $f_{k|k-1}\,[x] = \int x\, f_{k|k-1}\,(x)dx$ and $\kappa\,(z) = \lambda_c\, c\,(z)$ is the intensity function of the Poisson false alarm RFS.

For the JoM estimator, the Bayesian risk function to be minimized is given by [16]

$$\int C\,(X, J\,(Z))\, f\,(X)\,\delta X \approx 2 - p_{|X|}\,(|J|) - \frac{f\,(X)\,\varepsilon^{|J|}}{|J|!}, \tag{3.7.6}$$

where J denotes the JoM estimator, C is the cost function that penalizes both discrepancies in cardinality and multitarget states, and $p_{|X|}(|J|)$ is the cardinality distribution evaluated at the target number $|J|$.

Then, using the updated Bernoulli parameters from the JoTT filter, the JoM estimator confirms the presence of a single target if

$$2 - (1 - q_{k|k}) > 2 - q_{k|k} - q_{k|k} f_{k|k}(\hat{x})\,\varepsilon, \tag{3.7.7}$$

where the left hand side is the Bayes risk function evaluated for the "no-target" case, i.e., $X = \emptyset$ and the right hand side is the Bayes risk function evaluated for the "target-present" case, i.e., $X = \{x\}$. Solving this inequality for $q_{k|k}$ yields the following test for "target-present" decision:

$$q_{k|k} > \frac{1}{2 + f_{k|k}(\hat{x})\,\varepsilon}. \tag{3.7.8}$$

As in the original JoM estimator, first, the MAP estimate of $X = \{x\}$ is computed from the parameterized FISST density, i.e., $\left(q_{k|k}, f_{k|k}\right)$ where the spatial pdf $f_{k|k}$ has the Gaussian mixture form, i.e., $f_{k|k}(x) = \sum_{i=1}^{N_k} w_{k|k}^{(i)} f_{k|k}^{(i)}$, with the mixing weights satisfying that $\sum_{i=1}^{N_k} w_{k|k}^{(i)} = 1.0$. Before state estimation, pruning and merging of the Gaussian components are performed. Thus, the state estimation is obtained using the well-separated and significant Gaussian density components according to (3.4.8). For the selected Gaussian density component, its Pareto-optimal volume given by $T_{P,opt} = q_{k|k}\varepsilon_{P,opt}$ is computed. Then, the test for "target-present" decision in (3.7.8) is checked using $\varepsilon_{P,opt}$. That is, $f_{k|k}(\hat{x})\,\varepsilon_{P,opt}$ is set to $\min\left(f_{k|k}(\hat{x})\,\varepsilon_{P,opt}, 1/q_{k|k}\right)$. Consequently, if target is progressively better-localized, all of its probability mass would be almost located in $\varepsilon_{P,opt}$, i.e., $q_{k|k} f_{k|k}(\hat{x}) \varepsilon_{P,opt} \approx 1$ [27].

3.8 Simulation Results

In this section, the proposed JoM estimator is compared with the MaM estimator. To do this, their track management performance using outputs of the JoTT filter is evaluated through the OSPA metric [36, 33]. The OSPA metric compares two finite sets X, and Y, considering the difference in their cardinalities (i.e., cardinality error) and the positional distance between their associated points (i.e., localization error) after an optimal assignment. The sensitivity of the OSPA metric to these two errors are controlled by the cut-off parameter c and the order parameter p. However, for a Bernoulli RFS the OSPA metric reduces to [12]

$$d_p^{(c)}(X,Y) = \begin{cases} 0 \text{ if } X = \emptyset,\ Y = \emptyset \\ c \text{ if } X = \emptyset,\ Y = \{y\} \\ c \text{ if } X = \{x\},\ Y = \emptyset \\ d^{(c)}(x,y) \text{ if } X = \{x\},\ Y = \{y\}, \end{cases}$$

where $d^{(c)}(x,y) = \min(c, d(x,y))$ is the cut-off distance between the points in two non-empty Bernoulli RFSs. Thus, in this case, the OSPA metric is independent of the order parameter p. In addition, the major performance difference between the two estimators is expected to occur in the accuracy of their decisions on track confirmation, track maintenance, and track termination. Then, the cut-off parameter c must be set to a high value in order to make the OSPA metric sensitive to cardinality errors due to false and missing point estimates. In simulations, the OSPA metric is therefore computed with the parameters $p = 1$, and $c = 25$.

The target state vector comprises position and velocities in $\mathrm{x} - \mathrm{y}$ directions, i.e., $x_k = [p_{x,k}, p_{y,k}, v_{x,k}, v_{y,k}]'$. If the target does survive with probability $p_S = 0.90$, its states evolve according to the coordinated turn model with the known turn rate Ω [7, 1], i.e., the state transition model is

$$x_k = F(\Omega)\, x_{k-1} + G\omega_{k-1},$$

where $\omega_{k-1} \sim N(0, Q_{k-1})$ is the zero-mean Gaussian process noise with covariance matrix $Q_{k-1} = diag([0.1, 0.1]')\, \mathrm{m/s}^2$, and the system matrices are

$$F(\Omega) = \begin{bmatrix} 1 & 0 & \frac{\sin(\Omega \mathrm{T})}{\Omega} & -\frac{1-\cos(\Omega \mathrm{T})}{\Omega} \\ 0 & 1 & \frac{1-\cos(\Omega \mathrm{T})}{\Omega} & \frac{\sin(\Omega \mathrm{T})}{\Omega} \\ 0 & 0 & \cos(\Omega \mathrm{T}) & -\sin(\Omega \mathrm{T}) \\ 0 & 0 & \sin(\Omega \mathrm{T}) & \cos(\Omega \mathrm{T}) \end{bmatrix}, \qquad G = \begin{bmatrix} \frac{T^2}{2} & 0 \\ 0 & \frac{T^2}{2} \\ T & 0 \\ 0 & T \end{bmatrix},$$

where T is the sampling interval and set at $\mathrm{T} = 1\mathrm{s}$ in simulations.

The single target tracking scenario runs for 40s. The target appears at time $k = 6$ and moves along a straight line with a constant speed of $|v| = 5\,\mathrm{m/s}$ in the $\mathrm{x} - \mathrm{y}$ directions until time $k = 20$. Then, it starts maneuvering at a constant turn rate of $|\Omega| = 2\,\mathrm{deg/s}$ and is terminated at time $k = 35$. The target birth is modeled as a Bernoulli RFS given by $\{q_b, f_b(x)\}$, where the birth existence probability is set at $q_b = 0.01$, and the spatial pdf is defined as $f_b(x) = N(\hat{x}_b, P_b)$ with mean $\hat{x}_b = [-70, 70, 0, 0]'$ and covariance matrix $P_b = diag\left([50, 50, 10, 10]'\right)$.

The target is detected by a sensor with state-independent detection probability

p_D and the sensor has a linear Gaussian measurement model given by

$$z_k = Hx_k + \eta_k,$$

where $\eta_k \sim N(0, R_k)$ is the zero-mean Gaussian measurement noise with covariance matrix $R_k = diag\left([1,1]'\right)$ m. With $I_{2\times2}$ and $0_{2\times2}$ denoting the $n \times n$ identity and zero matrices, respectively, the observation matrix is given by $H = [I_{2\times2}, 0_{2\times2}]$. In addition to noisy target-originated measurement, the received measurement set includes clutter points. In simulations, clutter is modeled as a Poisson RFS with the mean rate of $\lambda_c = 10$ per scan and uniform spatial distribution over the surveillance region $V = [-300\text{m}, 300\text{m}] \times [-300\text{m}, 300\text{m}]$, i.e., $c(z) = V^{-1}$. The performance of the two estimators is evaluated by running the same scenario for 500 Monte Carlo runs. In each trial, target-originated measurement, detected with p_D, and independent random clutters are generated. Fig. 3.4 shows the x and y components of the target trajectory, measurements and the position estimates obtained from the JoTT filter with $p_D = 0.80$ for one Monte Carlo trial.

In the JoTT filter, the Bernoulli RFS is represented as a Gaussian mixture. The maximum number of Gaussian components is set at $J_{\max} = 100$. They are pruned and merged at each time step with thresholds $T_{prune} = 10^{-3}$ and $T_{merge} = 4.0$, respectively according to the algorithm proposed in [37].

The track management performance of the proposed JoM estimator and the MaM estimator are shown in Fig. 3.5–3.7 for different values of the detection probability, ranging from high to moderately small values, i.e., $p_D = 0.95, 0.90, ..., 0.70$. The MaM estimator confirms "target-present" decision by comparing the existence probability $q_{k|k}$ with the hard threshold 0.5. However, the proposed JoM estimator confirms

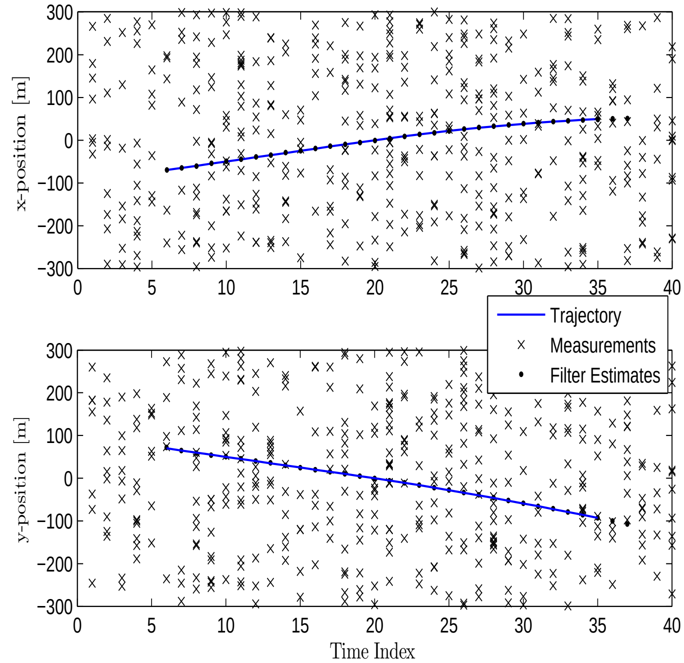

Figure 3.4: x and y components of target trajectory, measurements and JoTT filter estimates.

"target-present" decision by setting a lower margin than this hard threshold considering how well the JoTT filter localizes the target, i.e., the term $f_{k|k}(\hat{x})\,\varepsilon$ in (3.7.8). However, the maximum value of $f_{k|k}(\hat{x})\,\varepsilon$ is set by a confirmation threshold q_{min}. In simulations, q_{min} is set to 0.20. Thus, the track, for which $q_{k|k} > q_{min}$, is confirmed by the JoM estimator. In particular, the use of this threshold helps to prevent false point estimates before the target birth and after the target death.

In Fig. 3.5, it can be seen that the two estimators demonstrate almost the same track management performance in terms of track confirmation before the target birth at time $k = 6$. In addition, the initial track maintenance quality of the proposed JoM estimator with insignificant values of the lower margin is nearly the same as that of

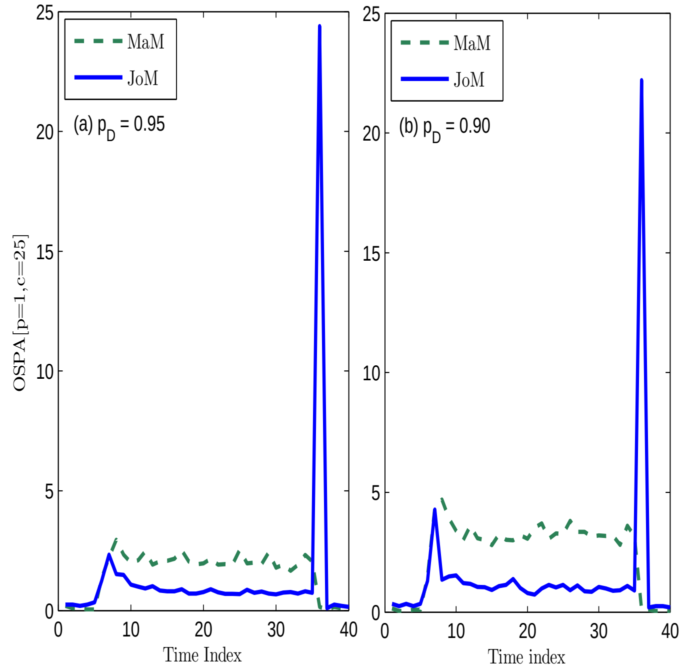

Figure 3.5: 500 Monte Carlo run averages of the OSPA metric computed for the track management performance of the JoM and MaM estimators.

the MaM estimator. However, the JoTT filter localizes the target more accurately using target-originated measurements detected with high probability as time proceeds. Therefore, the lower margin than the hard threshold 0.5 becomes significant, so that the proposed JoM estimator does not prematurely declare track termination if the target is miss-detected due to sensor imperfection. On the other hand, large values of the lower margin than the hard threshold 0.5 result in latency on track termination. That is, after the target is terminated at time $k = 35$, the localization performance of the JoTT filter does deteriorate gradually due to missed detections. Hence, the track termination decision is delayed in the proposed JoM estimator.

In Fig. 3.6(a), it can be seen that the track management performances of the two estimators are nearly the same during the tracking scenario. These results indicates

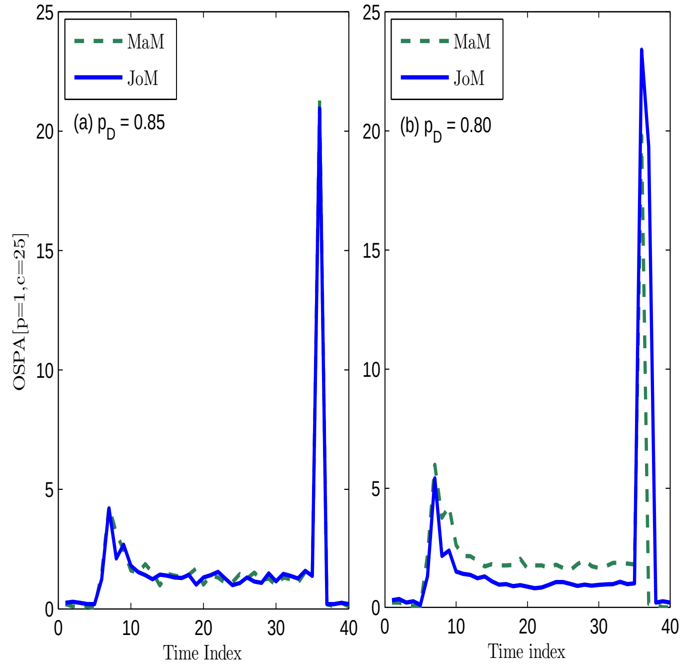

Figure 3.6: 500 Monte Carlo run averages of the OSPA metric computed for the track management performance of the JoM and MaM estimators.

that the decrease in the existence probability of target ($q_{k|k}$) cannot be compensated by the value of the lower margin computed in the proposed JoM estimator when the target is miss-detected. However, Fig. 3.6(b) shows that the track maintenance quality of the proposed JoM estimator is better than that of the MaM estimator after the target birth. That is, the value of the lower margin can compensate the decrease in $q_{k|k}$ due to target being miss-detected. Nevertheless, the proposed JoM estimator suffers from track termination latency more than the MaM estimator due to the statistics indicating a well-localized target obtained from the JoTT filter after time $k = 35$.

Finally, Fig. 3.7 shows the track management performances of the two estimators under moderately small detection probabilities. It can be seen that the initial track management performance of the proposed JoM estimator is better than that of the MaM estimator. More explicitly, the MaM estimator suffers much more from the track confirmation latency using the hard threshold 0.5 than the JoM estimator with insignificant values of the lower margin. In addition, the track maintenance quality of the proposed JoM estimator is better than that of the MaM estimator after a small period of time from the target birth. However, as in Fig. 3.6(b), the proposed JoM estimator confirms track termination with larger time delay after time $k = 35$, compared to the MaM estimator.

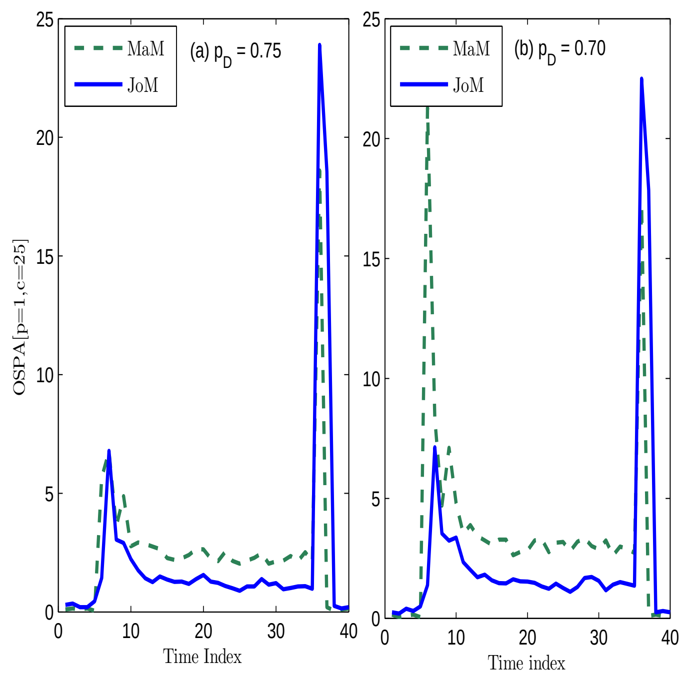

Figure 3.7: 500 Monte Carlo run averages of the OSPA metric computed for the track management performance of the JoM and MaM estimators.

According to the $AR(1)$ models in Section 3.6, time evolution of the weights in (3.5.9) for different values of detection probability is shown in Fig. 3.8 and Fig. 3.9. For considerably high detection probabilities, e.g., $p_D = 0.95$ and $p_D = 0.90$, the weights are adjusted as indicated in Section 3.6, i.e., they monotonically approach to their means after the optimal JoTT filter converges to its steady-state with detections. However, if the detection probability is not so high or close to moderately small values, the weights are predicted based on the estimation error analysis in the optimal JoTT filter. Consequently, the linear predictions can be considered to be adaptive to the JoTT filter's performance.

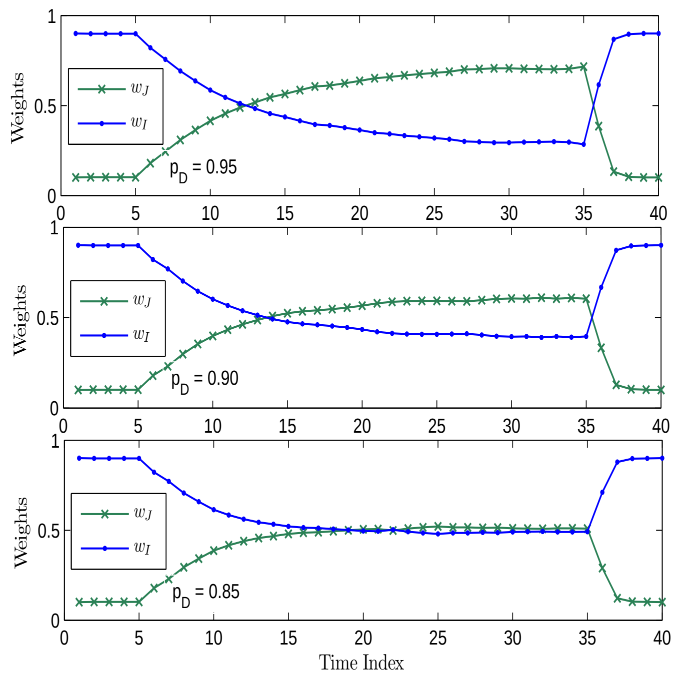

Figure 3.8: 500 Monte Carlo run averages of the weights for high detection probabilities.

3.9 Conclusions

In this paper, we have proposed an optimization algorithm to compute the optimal value of the unknown estimation constant in the JoM estimator. The optimization problem is defined in terms of two conflicting objective functions. The first objective

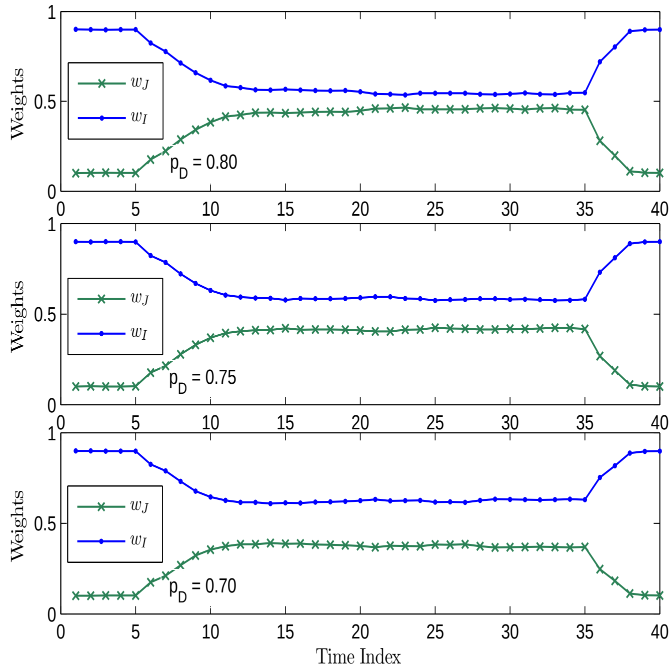

Figure 3.9: 500 Monte Carlo run averages of the weights for moderately small detection probabilities.

function is defined in terms of the information theoretic sense and aims for entropy maximization by setting the estimation constant to its maximum permissible value. In contrast, the second one arises from the constraint in the definition of the JoM estimator and aims to improve the accuracy of the JoM estimates by setting the estimation constant to its minimum value determined by the probability of user's confidence level. We used a standard optimization approach that guarantees a Pareto-optimal solution.

The proposed JoM estimator is used in the JoTT filter and compared to the other MAP type multitarget estimator-called the MaM estimator. The simulation results demonstrate that the track management performance of the proposed JoM estimator in terms of track confirmation latency, and track maintenance quality after target birth is better than that of the MaM estimator for different values of the detection probability, ranging from high to moderately small values. However, the proposed JoM estimator suffers from track termination latency more than the MaM estimator as the localization performance of the JoTT filter does deteriorate gradually after target termination.

3.10 Appendix A

To understand why selection of too small values for the JoM estimation constant does not ameliorate multitarget state estimates, quantize the FISST density $f(\{x_1, ..., x_n\})$ for all n into small and disjoint hyperspaces Δ^n with volume ε^n. Then using the relation $f(\{x_1, ..., x_n\}) \triangleq n! f(x_1, ..., x_n)$ the probability over a small hyperspace indexed by variable i, i.e., Δ_i^n [27] is computed as:

$$\begin{aligned} p_i(n) &= \frac{1}{n!} \int_{\Delta_i^n} f(\{x_1, ..., x_n\})\, dx_1 ... dx_n, \\ &= \int_{\Delta_i^n} f(x_1, ..., x_n)\, dx_1 ... dx_n, \\ &\approx f(\hat{x}_{1_i}, ..., \hat{x}_{n_i})\, \varepsilon^n. \end{aligned} \tag{3.10.1}$$

where $\hat{x}_{1_i}, ..., \hat{x}_{n_i}$ denotes the multitarget state estimates obtained from $f(\{x_1, ..., x_n\})$ in Δ_i^n. Note that if $f(\{x_1, ..., x_n\})$ is peaky over Δ_i^n, ε must be set to a small value to satisfy the following condition:

$$\sum_{n=0}^{\infty} \sum_{i:\Delta_i^n \in \mathcal{X}^n} p_i(n) \leq 1. \tag{3.10.2}$$

Thus, similar to the quantization of a continuous random variable, the entropy of the quantized RFS is defined as

$$H\left(X^{\Delta}\right) = -\sum_{n=0}^{\infty} \sum_{i:\Delta_i^n \in \mathcal{X}^n} p_i(n) \log\left(p_i(n)\right). \quad (3.10.3)$$

Upon substitution of $p_i(n) = f\left(\hat{x}_{1_i}, ..., \hat{x}_{n_i}\right) \varepsilon^n$ into the logarithmic function in (3.10.3), the entropy of the quantized RFS can be rewritten as

$$\begin{aligned} H\left(X^{\Delta}\right) &= -\sum_{n=0}^{\infty} \sum_{i:\Delta_i^n \in \mathcal{X}^n} p_i(n) \log\left(f\left(\hat{x}_{1_i}, ..., \hat{x}_{n_i}\right) \varepsilon^n\right), \\ &= -\sum_{n=0}^{\infty} \sum_{i:\Delta_i^n \in \mathcal{X}^n} p_i(n) \log\left(f\left(\hat{x}_{1_i}, ..., \hat{x}_{n_i}\right)\right) - \\ &\quad \sum_{n=0}^{\infty} \sum_{i:\Delta_i^n \in \mathcal{X}^n} p_i(n) \log\left(\varepsilon^n\right), \end{aligned} \quad (3.10.4)$$

where the first term is the average self-information of the joint symmetric pdfs (i.e., $f(x_1, ..., x_n)$) over $\Delta^{(n)}$ and the second term is the average self-information of the uniform pdfs (i.e.,$\mathcal{U}(x_1, ..., x_n) = \varepsilon^{-n}$) over $\Delta^{(n)}$.

For simplicity of analysis, assume that $f\left(\hat{x}_{1_i}, ..., \hat{x}_{n_i}\right) \approx 1.0$ over some hyperspaces Δ_i^n indexed by i^*. For the rest, $f\left(\hat{x}_{1_i}, ..., \hat{x}_{n_i}\right) \approx 0$ and thus from (3.10.1) the probability over those regions is $p_i(n) \approx 0$. In this case, using the convention $0 \log 0 = 0$ and $\log 1 = 0$, the first term in (3.10.4) is canceled and the entropy of the quantized RFS simplifies to

$$H\left(X^{\Delta}\right) \approx -\sum_{n=0}^{\infty} \sum_{i^*:\Delta_{i^*}^n \in \mathcal{X}^n} p_{i^*}(n) \log\left(\varepsilon^n\right), \quad (3.10.5)$$

Note that ε is small enough to satisfy the condition given by (3.10.2). Similar to typical sequences with equal probabilities in a typical set, most of the total probability is almost equally divided on some hyperspaces $\Delta_{i^*}^n$ indexed by i^*. Therefore, selecting too small values for ε will not ameliorate the accuracy in multitarget state estimates. On the contrary, the entropy will get larger values due to the uncertainty regarding what multitarget state estimate is true.

3.11 Appendix B

For the standard Gaussian density $f(x)$ defined in $\mathbb{R}^{n_x}$, it follows from (3.3.3) that the amount of self-information associated with the outcome $(\tilde{x}_1, ..., \tilde{x}_n) \in A_n^\tau$ is

$$H(f) - \tau < -\frac{1}{n} \log f(\tilde{x}_1, ..., \tilde{x}_n) < H(f) + \tau, \tag{3.11.1}$$

where $\tilde{x}_1, ..., \tilde{x}_n$ are i.i.d. samples from $f(x)$, and $H(f) = 0.5 \log (2\pi e)^{n_x}$ [13].

Substituting for $-\log f(\tilde{x}_1, ..., \tilde{x}_n) = 0.5n \log (2\pi)^{n_x} + 0.5 \sum_{i=1}^n \tilde{x}_i^T \tilde{x}_i$ into (3.11.1) and making some algebraic manipulations yield

$$n(n_x - 2\tau) < \sum_{i=1}^n \tilde{x}_i^T \tilde{x}_i < n(n_x + 2\tau), \tag{3.11.2}$$

where $\sum_{i=1}^n \tilde{x}_i^T \tilde{x}_i$ represents a thin shell around a hypersphere centered at the origin of $\mathbb{R}^{n_x}$ as claimed.

3.12 Appendix C

The nonlinear convex optimization problem in (3.5.7) is referred to as the primal problem [10]. The Lagrangian of the primal problem is written as

$$L(\tau, \lambda) = f_{o,I}(\tau) + \lambda_1 g_1(\tau) + \lambda_2 g(\tau), \tag{3.12.1}$$

where τ and $\lambda = (\lambda_1, \lambda_2)$ are called primal and dual variables, respectively.

According to the duality theorem, the dual problem has the same optimal solution with the primal problem if Slater's condition holds [6]. Associated with the primal problem, the dual function is defined as

$$\begin{aligned} g(\lambda) &= \min_{\tau} L(\tau, \lambda), \\ &= L(\tau^*, \lambda), \end{aligned} \tag{3.12.2}$$

where τ^* is the primal solution and the dual solutions to $g(\lambda)$, i.e., $\lambda^* = (\lambda_1^*, \lambda_2^*)$ are the Lagrange multipliers of the primal problem.

For any convex optimization problem with differentiable objective and constraint functions, the necessary and sufficient conditions to analyze the optimality of τ^*, and $\lambda^* = (\lambda_1^*, \lambda_2^*)$, are called the Karush-Kuhn-Tucker (KKT) conditions [10, 6]. That is, τ^*, and $\lambda^* = (\lambda_1^*, \lambda_2^*)$ must satisfy the following conditions

$$g_i(\tau^*) \leq 0, \quad \text{for } i = 1, 2 \tag{3.12.3}$$

$$\lambda_i^* \geq 0, \quad \text{for } i = 1, 2 \tag{3.12.4}$$

$$\lambda_i^* g_i(\tau^*) = 0, \quad \text{for } i = 1, 2 \tag{3.12.5}$$

and

$$\nabla_{\tau} L\left(\tau^{*}, \lambda^{*}\right)=\nabla_{\tau} f_{o,I}\left(\tau^{*}\right)+\sum\nolimits_{i=1}^{2} \lambda_{1}^{*} \nabla_{\tau}\ g_{i}\left(\tau^{*}\right)=0, \tag{3.12.6}$$

where (3.12.3) is called primal feasibility conditions of τ^*, (3.12.4) is called the dual feasibility conditions of $\lambda^* = (\lambda_1^*, \lambda_2^*)$, and (3.12.5) is called complementary slackness conditions. Thus, the last KKT condition verifies that τ^* is the global minimum point of $L(\tau, \lambda^*)$.

Based on the KKT conditions three possible cases are distinguished for optimality of τ^* , and $\lambda^* = (\lambda_1^*, \lambda_2^*)$:

1. The constraints are both inactive: this means that $\lambda_i^* = 0$, for $i = 1, 2$. Then, the optimal value of the primal variable is set to $\tau^* = 0$ to satisfy the last KKT condition as

$$\nabla_{\tau} L\left(\tau^{*}, \lambda^{*}\right)=\frac{n_x}{n_x-2\tau^{*}}-1=0. \tag{3.12.7}$$

2. The constraints are both active: this means that $\lambda_i^* > 0$ for $i = 1, 2$. Then, the complementary slackness conditions contradicts for optimality of τ. That is, (3.12.5) for $i = 1$ requires that $\tau^* = 0$, whereas (3.12.5) for $i = 2$ requires that $\tau^* = 0.5\,(n_x - \gamma_{\min})$ where $\gamma_{\min} \ll n_x$. Nevertheless, the optimal value of the primal variable becomes $\tau^* = 0$ if the probability of confidence is excessively set to $\gamma_{\min} = n_x$. Thus, the last KKT condition will have the form

$$\begin{aligned}\nabla_{\tau} L\left(\tau^{*}, \lambda^{*}\right)&=\frac{n_x}{n_x-2\tau^{*}}-1-\lambda_1^*+2\lambda_2^*,\\&=-\lambda_1^*+2\lambda_2^*,\end{aligned} \tag{3.12.8}$$

in which case, $\lambda_1^* = 2\lambda_2^*$. That is, the inequality constraint $g_2(\tau)$ turns into $g_2(\tau) : \tau \leq 0$. Then, the constraints $g_1(\tau)$ and $g_2(\tau)$ contradict each other unless they both turn into the equality constraint given by $\tau = 0$.

3. One active and one inactive constraint: this means that either $\lambda_1^* > 0$ and $\lambda_2^* = 0$ or $\lambda_1^* = 0$ and $\lambda_2^* > 0$. If $\lambda_1^* > 0$ and $\lambda_2^* = 0$, then the complementary slackness condition for $i = 1$ requires that $\tau^* = 0$ but the last KKT condition cannot be satisfied for $\tau^* = 0$. On the other hand, if $\lambda_1^* = 0$ and $\lambda_2^* > 0$, then the complementary slackness condition for $i = 2$ requires that $\tau^* = 0.5\left(n_x - \gamma_{\min}\right)$ and again, the last KKT condition cannot be satisfied for $\lambda_2^* > 0$.

Consequently, the inequality constraints for the nonlinear convex problem are both inactive unless $n_x = \gamma_{\min}$. In addition, $\tau^* = 0$ is the optimal solution for the primal problem. That is, the convex objective function $f_{o,I}\left(\tau\right)$ given by (3.5.7) has a global minimum at $n_x - 2\tau^* = n_x$. Note that the inequality constraints $g_i\left(\tau\right),$ for $i = 1, 2$ are affine in addition to the convexity of $f_{o,I}\left(\tau\right)$, then Slater condition for the strong duality holds. Therefore, the strong duality indicates that the optimal solution to the primal problem $f_{o,I}\left(\tau\right)$ can be attained from the dual problem [10].

Bibliography

[1] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan. *Estimation with Applications to Tracking and Navigation*. J. Whiley and Sons, New York NY, 2001.

[2] Y. Bar-Shalom, P. K. Willett, and X. Tian. *Tracking and Data Fusion: A Handbook of Algorithms*. YBS Publishing, Storrs CT, 2011.

[3] E. Baser, T. Kirubarajan, and M. Efe. Improved MeMBer filter with modeling of spurious targets. In *Proc. of 16th Int'l Conf. on Information Fusion*, pages 813–819, Istanbul, Turkey, Jul. 2013.

[4] M. Baum, P. K. Willett, and U. D. Hanebeck. MMOSPA-based track extraction in the PHD filter - a justification for k-means clustering. In *Proc. of the 53rd IEEE Conf. on Decision and Control*, pages 1816–1821, Ca, USA, Dec. 2014.

[5] M. Baum, P. K. Willett, and U. D. Hanebeck. Polynomial-time algorithms for the exact MMOSPA estimate of a multi-object probability density represented by particles. *IEEE Trans. on Signal Processing*, 63(10):2476–2484, May 2015.

[6] D. P. Bertsakes. *Nonlinear Programming*. Athena Scientific, Belmont MA, 2nd edition, 1999.

[7] S. S. Blackman and R. F. Popoli. *Design and Analysis of Modern Tracking Systems*. Artech House, Norwood MA, 1999.

[8] S. Bobkov and M. Madiman. An equipartition property for high-dimensional log-concave distributions. In *Proc. of the 50th Annual Allerton Conf. on Commun., Control, and Comput*, pages 482–488, Il, USA, Oct. 2012.

[9] P. T. Boggs and J. W. Tolle. Sequential quadratic programming. *Acta Numerica*, (4):1–51, January 1995.

[10] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge UK, 2009.

[11] S. Challa, B.-N. Vo, and X. Wang. Bayesian approaches to track existence - IPDA and random sets. In *Proc. 5th Int'l Conf. on Information Fusion*, pages 1228–1235, Annapolis, Maryland, 2002.

[12] D. Clark, B.-T. Vo, and B.-N. Vo. Forward-backward sequential monte carlo smoothing for joint target detection and tracking. In *Proc. of 16th Int'l Conf. on Information Fusion*, pages 899–906, Seattle, USA, Jul. 2009.

[13] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. J. Wiley and Sons, New York NY, 2nd edition, 2005.

[14] D. F. Crouse, P. K. Willett, and Y. Bar-Shalom. Developing a real-time track display that operators do not hate. *IEEE Trans. on Signal Processing*, 59(7):3441–3447, Jul. 2011.

[15] D. J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Process*

Volume II: General Theory and Structure. Springer, New York NY, 2nd edition, 2008.

[16] I. Goodman, R. Mahler, and H. Nguyen. *Mathematics of Data Fusion*. Kluwer Academic Publisher, Norwell MA, 1997.

[17] A. A. Gorji, R. Tharmarasa, and T. Kirubarajan. Performance measures for multiple target tracking problems. In *Proc. of 14th Int'l Conf. on Information Fusion*, pages 1–8, Illinois, USA, Jul. 2011.

[18] O. Grodzevich and O. Romanko. Discussions on normalization and other topics in multi-objective optimization. In *Proc. of Fields-MITACS Industrial Problem-Solving Workshop (FMIPW)*, pages 89–101, Toronto, Canada, Aug. 2006.

[19] A. Hobson and B.-K. Cheng. A comparison of the Shannon and Kullback information measures. *Journal of Statistical Physics*, 7(6):301–310, Apr. 1973.

[20] E. T. Jaynes. On the rationale of maximum entropy methods. *Proc. IEEE*, 70(9):939–952, Sept. 1982.

[21] G. Kirchgassner, J. Wolters, and U. Hassler. *Introduction to Modern Time Series Analysis*. Springer, New York NY, 2nd edition, 2013.

[22] A. Lesne. Shannon entropy: A rigorous mathematical notion at the crossroads between probability, information theory, dynamical systems and statistical physics. *Math. Struct. in Comp. Science*, 24, 2014.

[23] X. R. Li and Z. Zhao. Evaluation of estimation algorithms part i: Incomprehensive measures of performance. *IEEE Trans. Aerosp. and Electron. Syst.*, 42(4):1340–1358, Jan. 2006.

[24] D. J. C. MacKay. *Information Theory, Inference, and Learning Algorithms.* Cambridge University Press, Cambridge UK, 2003.

[25] R. Mahler. Why multi-source, multitarget data fusion is tricky. In *Proc. of IRIS National Symposium on Sensor and Data Fusion*, pages 135–153, May 1999.

[26] R. Mahler. 'Statistics 101' for multisensor, multitarget data fusion. *IEEE Aerosp. and Electron. Sys. Magazine Part 2: Tutorials*, 19(1):53–64, Jan. 2004.

[27] R. Mahler. *Statistical Multisource-Multitarget Information Fusion.* Artech House, Norwood MA, 2007.

[28] R. T. Marler and J. S. Arora. Survey of multi-objective optimization methods for engineering. *Structural and Multidisciplinary Optimization*, 26(6):369–395, Mar. 2004.

[29] R. T. Marler and J. S. Arora. The weighted sum method for multi-objective optimization. *Structural and Multidisciplinary Optimization*, 41(6):853–862, Dec. 2009.

[30] D. Musicki, R. Evans, and S. Stankovic. Integrated probabilistic data association. *IEEE Trans. Automatic Control*, 39(6):1237–1241, Jun. 1994.

[31] P. Ngatchou, A. Zarei, and M. A. El-Sharkawi. Pareto multi objective optimization. In *Proc. of the 13th Int'l Conf. on Intell. Syst. Application to Power Systems*, pages 84–91, Va, USA, Nov. 2005.

[32] M. Rezaeian and B.-N. Vo. The entropy of random finite sets. In *Proc. of Int'l Symp. on Information Theory*, Seoul, Korea, Jun. 2009.

[33] B. Ristic, B.-N. Vo, D. Clark, and B.-T. Vo. A metric for performance evaluation of tracking algorithms. *IEEE. Trans. Signal Processing*, 59(7):3452–3457, Jul. 2011.

[34] B. Ristic, B.-T. Vo, B.-N. Vo, and A. Farina. A tutorial on Bernoulli filters: Theory, implementation and applications. *IEEE Trans. Signal Processing*, 61(13):3406–3430, Jul. 2013.

[35] R. L. Rothrock and O. E. Drummond. Performance metrics for multiple-sensor, multiple-target tracking. In *Proc. of SPIE Conf. on Signal and Data Processing of Small Targets*, pages 521–531, Orlando, USA, Apr. 2000.

[36] D. Schuhmacher, B.-T. Vo, and B.-N. Vo. A consistent metric for performance evaluation of multi-object filters. *IEEE. Trans. Signal Processing*, 59(7):3447–3457, Aug. 2008.

[37] B.-N. Vo and W.-K. Ma. The Gaussian mixture probability hypothesis density filter. *IEEE Trans. Signal Processing*, 54(11):4091–4104, November 2006.

[38] B.-T. Vo. *Random finite sets in multi-object filtering*. PhD thesis, School of Electrical, Electronic and Computer Engineering, University of Western Australia, Australia, Oct. 2008.

[39] B.-T. Vo, B.-N Vo, and A. Cantoni. The cardinality balanced multi-target multi-bernoulli filter and its implementations. *IEEE Trans. Signal Processing*, 57(2):409–423, Feb. 2009.

[40] X.-S. Yang. *Nature-Inspired Optimization Algorithms*. Elsevier, London UK, 2014.

The following chapter is a reproduction of a peer-reviewed article in the Institute of Electrical and Electronics Engineers (IEEE):

Erkan Baser, Thia Kirubarajan, Murat Efe, and Bhashyam Balaji, A Novel Joint Multitarget Estimator for Multi-Bernoulli Models, *IEEE Transactions on Signal Processing*, no. 19, vol. 64, pp. 5038–5051, June 2016.

Chapter 4

A Novel Joint Multitarget Estimator for Multi-Bernoulli Models

4.1 Abstract

In this paper, the joint multitarget (JoM) estimator proposed for the joint target detection and tracking (JoTT) filter is reformulated for the Gaussian mixture (GM) implementations of the multitarget multi-Bernoulli (MeMBer) filters. For this purpose, a mode-finding algorithm is employed to search for the most significant mode of a GM density. Thus, the maximum a posterior (MAP) estimates of Bernoulli targets are determined. In addition, the multi-Bernoulli versions of the two conflicting objective functions for the Pareto-optimal value of the unknown JoM estimation constant are derived. Simulations compare the performance of the proposed JoM estimator

with that of the marginal multitarget (MaM) estimator in a multitarget tracking scenario, where the probability of target detection is a function of target states. The simulation results demonstrate that the proposed JoM estimator outperforms the MaM estimator under moderately low-observable conditions. This is because the incomplete cost function of the MaM estimator is not adequate to obtain accurate cardinality estimates of targets without considering how well targets are localized. Nevertheless, the proposed JoM estimator may suffer from track termination latency more than the MaM estimator due to the definition of its cost function.

4.2 Introduction

The multitarget tracking (MTT) problem is one of jointly estimating the time-varying number of targets and their random states using noisy sensor measurements in the presence of missed detections and false alarms [2]. As opposed to the classical MTT algorithms, e.g., the multiple hypothesis tracking (MHT) filter [2], the random finite set (RFS) theory and its calculus known as the finite set statistics (FISST) provide a new approach to obtain the multitarget (FISST) posterior probability density without making explicit data association [21, 24, 23]. The probability hypothesis density (PHD) filter was developed as the first-order moment approximation to the FISST posterior probability density [23, 20]. The physical interpretation of the PHD can be explained as target occupancy probabilities of infinitesimal bins in the surveillance region [14, 15]. However, the cardinality estimate of targets from the PHD filter can be inaccurate and unstable due to the PHD's linearization and Poisson RFS approximation [13, 22]. Therefore, the cardinalized PHD (CPHD) filter was developed to estimate the cardinality distribution as well the PHD [22]. Hence, it provides more accurate cardinality estimates compared to the PHD filter [37]. Nevertheless, both the PHD and CPHD filters can suffer from the so-called spooky effect in local cardinality estimates [16].

The Bernoulli RFS formalism models the track of a target by a parameter pair (q, f), where q denotes the existence probability of the target, and f denotes its probability density function (pdf) if it does exist [23, 31]. The integrated probabilistic data-association (IPDA) filter estimates the existence probability of a single-target and its states simultaneously [26]. This optimal Bayesian solution to the complete single-target detection and tracking problem was re-derived using the Bernoulli RFS in [12] and then was extended as the Bernoulli filter, also known as the joint target detection and tracking (JoTT) filter in [23, 35]. In addition, using the multi-Bernoulli RFS, the multitarget multi-Bernoulli (MeMBer) filter was proposed as the parametric approximation to the FISST posterior probability density [23].

The MeMBer filter in [23] (hereinafter called the original MeMBer filter) is positively biased in the cardinality estimate [38]. The cardinality balanced MeMBer (CBMeMBer) filter removes the bias by computing the updated existence probabilities from the exact probability generating functionals (p.g.fl.) [38]. However, this results in a restrictive assumption on the probability of target detection in order to compute valid spatial probability density functions. In [4], the improved MeMBer (IMeMBer) filter removes the bias by introducing spurious Bernoulli targets without any restrictive assumptions. In addition, it refines the existence probabilities in light of measurements using the statistics of these targets. Thus, it improves the cardinality estimate, compared to the CBMeMBer filter, which in turn provides better cardinality estimate than that obtained from the PHD filter [38, 4].

Recently, the notion of labeled RFS was introduced along with their conjugate priors for the standard multitarget likelihood functions [36]. Hence, the generalized labeled multi-Bernoulli (GLMB) filter and its more efficient version known as δ-GLMB filter in terms of computational complexity and memory requirements were proposed. These two filters are analytically tractable and have closed-form solutions [34]. However, they are computationally more expensive than the IMeMBer and CBMeMBer filters [36]. Therefore, the LMB filter was developed as an efficient approximation of the δ-GLMB filter by inheriting the benefits of the CBMeMBer and δ-GLMB filters [29]. Subsequently, a tractable approximation of any labeled RFS density using GLMB densities was proposed in [28]. On the other hand, a conjugate prior distribution for unlabeled RFSs in a special hybrid form was proposed in [40] to obtain their full multitarget posterior. Thus, two tractable multitarget filters were developed using two different approximations of marginal association distributions. Then, a robust alternative to these two filters was proposed by finding the best-fitting multi-Bernoulli RFS to the FISST posterior probability density [41]. The performances of the IMeMBer, CBMeMBer and LMB filters were compared with simulations in [4].

In the literature, two Bayes optimal estimators for RFSs were proposed and applied to multitarget tracking [23, 17, 19]. These estimators are called the marginal multitarget (MaM) and the joint multitarget (JoM) estimators. According to their Bayes cost functions in [17], the JoM estimator minimizes both the cardinality and spatial discrepancies between the true RFS and its estimate simultaneously to determine the MAP estimate, whereas the MaM estimator first minimizes the cardinality discrepancy and then the MAP estimate is extracted from the associated FISST posterior probability density. Therefore, the JoM estimator is more appropriate than the MaM estimator to obtain the estimates of multitarget states, especially when the cardinality estimate is related to the spatial information in the FISST probability density, i.e., targets' states under low-observable conditions [19, 5]. In addition, another Bayes optimal estimator for RFSs was proposed in [1, 7]. This specific estimator was applied to multiuser detection and channel tracking in wireless communications.

In [5], the exact use of the JoM estimator was proposed to obtain the estimate of target RFS from the JoTT filter. For the proposed estimator, the unknown JoM estimation constant is computed as a Pareto-optimal solution to two conflicting objective functions. The first objective is defined in terms of the information theoretic sense, whereas the second one is obtained from the constraint in the definition of the JoM estimator. Track management performance analysis in [5] demonstrates that the JoM estimator outperforms the MaM estimator in terms of track confirmation latency and track maintenance quality, but the track termination latency can be worse in the JoM estimator than that in the MaM estimator.

In this paper, the JoM estimator in [5] is reformulated for the Gaussian Mixture (GM) implementations of the MeMBer filters. For this purpose, a mode-finding algorithm is employed. This allows searching for the most significant mode of a GM density if it has more than one significant component close to one another. Thus, the MAP estimates of Bernoulli targets are obtained by maximizing their FISST posterior probability densities. In addition, GM densities are approximated around their most significant modes as Gaussian densities according to the number of their modes [9]. This local approximation facilitates the derivation of the multi-Bernoulli RFS versions of the two conflicting objective functions in [5]. Under moderately low-observable conditions, the simulation results demonstrate that the JoM estimator provides more reliable cardinality estimates compared to the MaM estimator. As illustrated by examples in [23, 19], the reason is that the cardinality estimates obtained from the MaM estimator are determined by the existence probabilities of targets without considering how well they are localized.

The rest of the paper is organized as follows: Section 4.3 outlines the IMeMBer and CBMeMBer filters. Section 4.4 provides an overview of the multitarget Bayes estimators along with the exact use of the JoM estimator proposed for the JoTT filter. In Section 4.5, the JoM estimator proposed for the JoTT filter is reformulated for the GM implementations of the MeMBer filters. Simulation results are shown in Section 4.6. Finally, conclusions are drawn in Section 4.7.

4.3 Background

4.3.1 Multitarget Multi-Bernoulli RFS Modeling

Let X_k denote a Bernoulli RFS of a mobile target at time k. Then, using the parameter pair (q_k, f_k) for its existence and the detection of target states, the FISST probability density of this Bernoulli RFS is defined as [23, 31]

$$f_k(X) = \begin{cases} 1 - q_k & \text{if } X_k = \emptyset, \\ q_k f_k(x) & \text{if } X_k = \{x\}. \end{cases} \tag{4.3.1}$$

A multi-Bernoulli RFS is a union of independent Bernoulli RFSs. That is, for independent Bernoulli targets $X_k^{(1)}, ..., X_k^{(m)}$ with pdfs $f_k^{(1)}(x), ..., f_k^{(m)}(x)$ and existence probabilities $q_k^{(1)}, ..., q_k^{(m)}$, the multitarget state is described by the parameter set $\{(q_k^{(i)}, f_k^{(i)})\}_{i=1}^{m}$. Thus, its FISST probability density is given by

$$f_k(X) = \begin{cases} \left(1 - q_k^{(1)}\right) \dots \left(1 - q_k^{(m)}\right) & \text{if } X_k = \emptyset, \\ \sum\limits_{\zeta} \prod\limits_{j=1}^{n} \Theta_\zeta f_k^{(\zeta(j))}(x_j) & \text{if } X_k = \{x_1, \dots, x_n\}, \end{cases} \tag{4.3.2}$$

where

$$\Theta_\zeta = \left(1 - q^{(1)}\right) \dots \left(1 - q^{(m)}\right) \frac{q^{(\zeta(1))}}{\left(1 - q^{(\zeta(1))}\right)} \cdots \frac{q^{(\zeta(n))}}{\left(1 - q^{(\zeta(n))}\right)},$$

and the summation is taken over all permutations of the joint association hypotheses that are defined as one-to-one functions $\zeta : \{1, ..., n\} \to \{1, ..., m\}$ for $n \leq m$ [23].

4.3.2 Multitarget Multi-Bernoulli Filters

Using the multi-Bernoulli RFS, the original MeMBer filter was derived under two approximations in [23]. These approximations allow the data-induced tracks (i.e., tracks of detected targets) to be modeled as multi-Bernoulli RFS. Then, the CBMeMBer filter was proposed to remove the positive bias observed in the cardinality estimate from the original MeMBer filter [38]. However, the CBMeMBer filter makes a restrictive assumption on the probability of target detection, namely that it is close to unity. In [4], the Bernoulli RFS given by (4.3.1) was augmented to introduce spurious targets in addition to actual ones. Thus, the IMeMBer filter removes the bias without any assumption on the probability of target detection. In addition, it refines the existence probabilities of Bernoulli targets. This prevents propagating two different hypotheses for the same target, one in the legacy track set (tracks of undetected targets) and another one in the data-induced track set. Hence, the IMeMBer outperforms the CBMeMBer filter in terms of stability of cardinality estimation [4].

In this subsection, we outline the IMeMBer and CBMeMBer filters by presenting their recursive steps since the JoM estimator proposed in [5] will be reformulated for their GM implementations. For detailed information about these two filters we encourage the interested readers to refer to [38, 4].

Prediction Step: This step is exact and identical in the IMeMBer and CBMeMBer filters. Suppose that at time $k-1$, the FISST posterior probability density of multi-Bernoulli RFS is parameterized by $\{(q_{k-1}^{(i)}, f_{k-1}^{(i)})\}_{i=1}^{M_{k-1}}$. The Bernoulli targets from time $k-1$ independently evolve according to a Markov state transition density $p_{k|k-1}(\cdot|x)$ with the probability of target survival $p_{S,k}(x)$. Thus, the parameters for surviving Bernoulli targets are independently computed as [38, 4]

$$q_{k|k-1}^{(i)} = q_{k-1}^{(i)} f_{k-1}^{(i)} \left[p_{S,k}\right], \tag{4.3.3}$$

$$f_{k|k-1}^{(i)}(x) = \frac{p_{k|k-1}\left[p_{S,k} f_{k-1}^{(i)}\right]}{f_{k-1}^{(i)}\left[p_{S,k}\right]}, \tag{4.3.4}$$

where for a given test function $h(x)$, the functional of $f(x)$ is defined by $f[h] = \int h(x) f(x)\, dx$.

In addition to the surviving Bernoulli targets, a new multi-Bernoulli RFS is introduced for target births using the parameter set given by $\{(q_{\Gamma,k}^{(i)}, f_{\Gamma,k}^{(i)})\}_{i=1}^{M_{\Gamma,k}}$. Therefore, the predicted FISST probability density is parameterized by the union of two parameter sets for two types of predicted Bernoulli targets, i.e., $\{(q_{k|k-1}^{(i)}, f_{k|k-1}^{(i)})\}_{i=1}^{M_{k|k-1}}$, where $M_{k|k-1} = M_{k-1} + M_{\Gamma,k}$ [38, 4].

Update Step: This step is divided into two parts for the two types of Bernoulli track sets. The first set is called the legacy track set. The computations of parameters for Bernoulli targets under this set are exact and identical in the IMeMBer and

CBMeMBer filters. That is, using the predicted parameter set $\{(q_{k|k-1}^{(i)}, f_{k|k-1}^{(i)})\}_{i=1}^{M_{k|k-1}}$ each Bernoulli target under the legacy track set is described by the parameters [38, 4]

$$q_{L,k}^{(i)} = q_{k|k-1}^{(i)} \frac{1 - f_{k|k-1}^{(i)}\left[p_{D,k}\right]}{1 - q_{k|k-1}^{(i)} f_{k|k-1}^{(i)}\left[p_{D,k}\right]}, \tag{4.3.5}$$

$$f_{L,k}^{(i)}\left(x\right) = \frac{1 - p_{D,k}\left(x\right)}{1 - f_{k|k-1}^{(i)}\left[p_{D,k}\right]} f_{k|k-1}^{(i)}\left(x\right), \tag{4.3.6}$$

assuming that it is not detected with probability $1 - p_{D,k}(x)$.

The second track set is called the data-induced track set and consists of detected Bernoulli targets. The computations of their parameters require two approximations in the original MeMBer filter. However, in the CBMeMBer filter, the second approximation (setting $h = 1$) is skipped. Instead, using the exact p.g.fl.s, the existence probabilities of these Bernoulli targets are computed as [38]

$$q_{U,k}(z) = \frac{\sum_{i=1}^{M_{k|k-1}} \frac{q_{k|k-1}^{(i)}\left(1-q_{k|k-1}^{(i)}\right) f_{k|k-1}^{(i)}\left[p_{D,k} g_k(z|\cdot)\right]}{\left(1-q_{k|k-1}^{(i)} f_{k|k-1}^{(i)}\left[p_{D,k}\right]\right)^2}}{\kappa\left(z\right) + \sum_{i=1}^{M_{k|k-1}} G_{U,k}^{(i)}\left[z;1\right]}, \tag{4.3.7}$$

where $g_k(z|x)$ is the sensor likelihood function for a given measurement $z \in Z_k$, $\kappa(z)$ is the intensity function of the Poisson distributed clutter, and

$$G_{U,k}^{(i)}\left[z;1\right] \approx \frac{q_{k|k-1}^{(i)} f_{k|k-1}^{(i)}\left[p_{D,k} g_k\left(z\,|\cdot\right)\right]}{1 - q_{k|k-1}^{(i)} f_{k|k-1}^{(i)}\left[p_{D,k}\right]}.$$

The violation of the multi-Bernoulli RFS modeling results in invalid pdfs. Hence, a restrictive assumption on the probability of target detection is made by setting

$p_{D,k}(x) \approx 1$. Thus, the pdfs of these Bernoulli targets are computed as [38]

$$f_{U,k}\left(x;z\right) = \frac{\sum_{i=1}^{M_{k|k-1}} \frac{q_{k|k-1}^{(i)}}{1-q_{k|k-1}^{(i)}} f_{k|k-1}^{(i)}\left(x\right) p_{D,k}\left(x\right) g_k\left(z\left|x\right.\right)}{\sum_{i=1}^{M_{k|k-1}} \frac{q_{k|k-1}^{(i)}}{1-q_{k|k-1}^{(i)}} f_{k|k-1}^{(i)}\left[p_{D,k} g_k\left(z\left|\cdot\right.\right)\right]}. \tag{4.3.8}$$

The IMeMBer filter follows the two approximations in the original MeMBer filter. On the other hand, it augments Bernoulli RFS to remove the positive bias by modeling spurious targets[1]. More explicitly, similar to the extension in [39], the state space of Bernoulli RFS X is temporarily extended as $\Upsilon' = \Upsilon \times \{0,1\}$. Thus, actual and spurious targets under the data-induced track set can be labeled by an augmented variable u. Hence, the FISST probability density of the augmented Bernoulli RFS, i.e., $X'_k = \{x,u\} \in \Upsilon'$, is defined as [4]

$$f_{U,k}\left(X'\right) = \begin{cases} 1 - q_{U,k}\left(z\right) & \text{if } X'_k = \emptyset, \\ \tilde{q}_{U,k}(z)\tilde{f}_{U,k}\left(x\right) & \text{if } X'_k = \left\{x, 0\right\}, \\ \bar{q}_{U,k}(z)\bar{f}_{U,k}\left(x\right) & \text{if } X'_k = \left\{x, 1\right\}, \end{cases} \tag{4.3.9}$$

where the parameter pairs $(\tilde{q}_{U,k}, \tilde{f}_{U,k})$ and $(\bar{q}_{U,k}, \bar{f}_{U,k})$ for $u = 0$ and $u = 1$ represent spurious and actual targets, respectively. Thus, the p.g.fl. of the data-induced track set in [23] is partitioned for actual and spurious targets (see, (25) in [4]). Hence, the

[1]In the update step of the original MeMBer filter, the legacy track set is obtained from the Bernoulli RFSs introduced in the prediction step, assuming that they are not detected. Nevertheless, the predicted Bernoulli RFSs are also used with each measurement to introduce targets under the data-induced track set. To address this ambiguity, i.e., the use of predicted Bernoulli RFSs to generate two contradicting hypotheses for each target, the idea of spurious target is introduced in [4].

parameters of these targets are computed as follows [4]:

$$\tilde{q}_{U,k}(z) = \frac{\sum_{i=1}^{M_{k|k-1}} \frac{q_{k|k-1}^{(i)} q_{L,k}^{(i)} f_{k|k-1}^{(i)}\left[p_{D,k} g_k(z|\cdot)\right]}{1-q_{k|k-1}^{(i)} f_{k|k-1}^{(i)}\left[p_{D,k}\right]}}{\kappa(z) + \sum_{i=1}^{M_{k|k-1}} G_{U,k}^{(i)}[z;1]}, \tag{4.3.10}$$

$$\tilde{f}_{U,k}(x;z) = \frac{\sum_{i=1}^{M_{k|k-1}} \tilde{\alpha}^{(i)} f_{k|k-1}^{(i)}(x)\, p_{D,k}(x)\, g_k(z\,|x)}{\sum_{i=1}^{M_{k|k-1}} \tilde{\alpha}^{(i)} f_{k|k-1}^{(i)}\left[p_{D,k} g_k(z\,|\cdot)\right]}, \tag{4.3.11}$$

where $\tilde{\alpha}^{(i)} = q_{k|k-1}^{(i)} q_{L,k}^{(i)} (1 - q_{k|k-1}^{(i)} f_{k|k-1}^{(i)}[p_{D,k}])^{-1}$ and

$$\bar{q}_{U,k}(z) = \frac{\sum_{i=1}^{M_{k|k-1}} \frac{q_{k|k-1}^{(i)} \left(1-q_{L,k}^{(i)}\right) f_{k|k-1}^{(i)}\left[p_{D,k} g_k(z|\cdot)\right]}{1-q_{k|k-1}^{(i)} f_{k|k-1}^{(i)}\left[p_{D,k}\right]}}{\kappa(z) + \sum_{i=1}^{M_{k|k-1}} G_{U,k}^{(i)}[z;1]}, \tag{4.3.12}$$

$$\bar{f}_{U,k}(x;z) = \frac{\sum_{i=1}^{M_{k|k-1}} \bar{\alpha}^{(i)} f_{k|k-1}^{(i)}(x)\, p_{D,k}(x)\, g_k(z\,|x)}{\sum_{i=1}^{M_{k|k-1}} \bar{\alpha}^{(i)} f_{k|k-1}^{(i)}\left[p_{D,k} g_k(z\,|\cdot)\right]}, \tag{4.3.13}$$

where $\bar{\alpha}^{(i)} = q_{k|k-1}^{(i)} (1 - q_{L,k}^{(i)})(1 - q_{k|k-1}^{(i)} f_{k|k-1}^{(i)}[p_{D,k}])^{-1}$. It is important to note that the existence probabilities given by (4.3.7) and (4.3.12) are identical. This means that the IMeMBer filter removes the positive bias as the CBMeMBer filter. On the other hand, if $p_{D,k}(x) \approx 1$, (4.3.13) will reduce to (4.3.8). As a result, these two filters can be considered identical if $p_{D,k}(x) \approx 1$. However, it is important to note that the aim of this paper is not to compare these two filters with each other.

Before the multitarget state estimation in the IMeMBer filter, the detection of any Bernoulli target under the legacy track set is checked using the following test statistic obtained from (4.3.10):

$$\tilde{q}_{U,k}(z,i) = \frac{\frac{q^{(i)}_{k|k-1} q^{(i)}_{L,k} f^{(i)}_{k|k-1}\left[p_{D,k} g_k(z|\cdot)\right]}{1-q^{(i)}_{k|k-1} f^{(i)}_{k|k-1}\left[p_{D,k}\right]}}{\kappa(z) + \sum_{i=1}^{M_{k|k-1}} G^{(i)}_{U,k}[z;1]}. \tag{4.3.14}$$

The proposed refinement process in [4] can be summarized as follows: first, it determines the MAP estimate of the cardinality of spurious targets using (4.3.10). Then, it removes Bernoulli targets under the legacy track set, which maximize (4.3.14) for the measurements giving rise to these spurious targets. In addition, if (4.3.14) is small but has a significant contribution from only one Bernoulli target under the legacy track set, this may indicate a resolved and detected Bernoulli target under the legacy track set. Hence, the proposed refinement process removes these Bernoulli targets as well. Finally, it updates existence probabilities of actual targets for the same measurements. Consequently, each Bernoulli RFS under the legacy track set and the data-induced track set represents a particular target hypothesis with a significant probability.

4.4 Multitarget Bayes Estimators

For RFSs with different cardinalities, their FISST probability densities are incomparable because of their different units of measurement. Furthermore, addition and subtraction operations on RFSs are not defined properly. Therefore, the multitarget analogues of the standard MAP and expected a posteriori (EAP) estimators are undefined [23, 17, 19]. Nevertheless, two MAP-like multitarget Bayes estimators were proposed to obtain estimates of an RFS from its FISST probability density.

Marginal Multitarget (MaM) Estimator: The MaM estimate of an RFS is computed in two steps: first, the MAP estimate of the cardinality is determined as

$$\hat{n}^{(MAP)} \triangleq \arg \sup_{n} p_{|X|}(n), \tag{4.4.1}$$

where $|X|$ denotes the cardinality variable for the RFS X and is characterized by its posterior probability mass function (pmf) as follows:

$$p_{|X|}(n) \triangleq \frac{1}{n!} \int f_k\left(\{x_1, ..., x_n\} \middle| Z^{(k)}\right) dx_1 ... dx_n. \tag{4.4.2}$$

Then, assuming that $\hat{n}^{(MAP)}$ is the true cardinality, the MaM estimate of the multi-target RFS is determined as

$$\hat{X}^{(MaM)} = \arg \sup_{x_1, ..., x_{\hat{n}^{(MAP)}}} f_k\left(\{x_1, ..., x_{\hat{n}^{(MAP)}}\} \middle| Z^{(k)}\right). \tag{4.4.3}$$

The MaM estimator is Bayes optimal [23, 17, 19]. However, it does not consider all uncertainties contained in the FISST posterior probability density [17, 5]. That is, using the cost function of the MaM estimator the Bayesian risk function to be minimized is computed as $1-p_{|X|}(n)$ [17]. However, as explained in [5], the differential entropy of an RFS can be expressed as the sum of three terms: i) entropy of the cardinality distribution, ii) average differential entropy of the joint pdf, and iii) uncertainty due to change in representation from RFS to vectors of indistinguishable points. Hence, its estimate given by (4.4.3) can be unreliable, when the cardinality estimate is related to targets' states [19, 5].

The use of the MaM estimator for the MeMBer filters can be illustrated on the parameter set $\{(q_k^{(i)}, f_k^{(i)})\}_{i=1}^{M_k}$ as follows [23]: first, the posterior pmf of the cardinality is computed for $n = 1, ..., M_k$, i.e.,

$$p_{|X|}(n) = \prod_{i=1}^{M_k} \left(1 - q_k^{(i)}\right) \sigma_{M_k,n}(\Theta), \tag{4.4.4}$$

where $\sigma_{M_k,n}$ is the elementary symmetric function of degree n in the set

$$\Theta = \left\{ \frac{q_k^{(1)}}{1 - q_k^{(1)}}, ..., \frac{q_k^{(M_k)}}{1 - q_k^{(M_k)}} \right\},$$

and is computed according to [37]. Thus, $\hat{n}^{(MAP)}$ is obtained from (4.4.1). On the other hand, there is no straightforward solution to (4.4.3). Instead, one approach is to determine the individual state estimates from the PHD as $\hat{X}^{(MaM)} = \arg\sup_{x_1,...,x_{\hat{n}^{(MAP)}}} D_k(x)$, where $D_k(x) = \sum_{i=1}^{M_k} q_k^{(i)} f_k^{(i)}(x)$ [23]. However, as indicated in [33], this may result in selecting peaks from the PHD with small weights instead of significant ones, especially when newborn Bernoulli targets appear with large covariance matrices. This would contradict the first step of the MaM estimator since these selected targets may not make significant contributions to $\hat{n}^{(MAP)}$. Alternatively, a favorable approach is to extract individual state estimates as the MAP estimates from pdfs of Bernoulli targets, considering their existence probabilities [23, 33].

Joint Multitarget (JoM) Estimator: This estimator determines the cardinality and multitarget states from the FISST posterior probability density simultaneously, as opposed to the MaM estimator [23], i.e.,

$$\hat{X}^{(JoM)} = \arg\sup_{X} f_k\left(X \,\middle|\, Z^{(k)}\right) \frac{\varepsilon^{|X|}}{|X|!}, \tag{4.4.5}$$

where the parameter ε denotes a small constant (hereinafter called the JoM estimation constant) and satisfies $f\left(\{x_1, ..., x_n\}\right)\varepsilon^n \leq 1$ for all integers $n \geq 0$. However, the selected value of ε results in a trade-off between the accuracy of the estimated multitarget states and the convergence rate to the true multitarget states. That is, the smaller the value of ε, the more accurate the estimates from the JoM estimator but the slower the convergence speed to that accuracy and vice versa [23]. In addition, note that the value of ε should be no smaller than the best possible localization accuracy, given the limitations of the sensor.

Alternatively, the JoM estimator can be implemented by following these two steps [23]: first, for integer values $n \geq 0$ the MAP estimates of the RFSs are computed as

$$\hat{X}_n^{(MAP)} = \arg\sup_{x_1,...,x_n} f_k\left(\{x_1, ..., x_n\} \,\middle|\, Z^{(k)}\right). \tag{4.4.6}$$

Then, the JoM estimate of the multitarget RFS is determined as $\hat{X}^{(JoM)} = \hat{X}_{\hat{n}}^{(MAP)}$, where $\hat{n}$ is the solution to the following optimization problem:

$$\hat{n} = \arg\sup_{n} f_k\left(\{\hat{x}_1, ..., \hat{x}_n\} \,\middle|\, Z^{(k)}\right) \frac{\varepsilon^n}{n!}. \tag{4.4.7}$$

Like the MaM estimator, the JoM estimator is Bayes optimal [23, 17, 19]. However, it is more appropriate than the MaM estimator to obtain the estimates of multitarget states since its cost function penalizes both discrepancies in cardinality and multitarget states simultaneously [17]. In addition, the JoM estimator is statistically consistent [17]. Therefore, its estimates would be more reliable than those obtained from the MaM estimator [19, 5].

The exact use of the JoM estimator has just been proposed for the JoTT filter in [5]. The proper choice of the unknown JoM estimation constant (ε) is computed as a Pareto-optimal solution to a multi-objective nonlinear convex optimization problem. The multi-objective optimization problem consists of multiple objective functions in conflict. The aggregation of these objective functions with appropriate weights is the standard way of solving this problem [27, 25]. However, the objective functions usually take values in different ranges. Therefore, they should be first normalized to be comparable in magnitudes [18].

In [5], the multi-objective function is defined by two conflicting objective functions. The first objective function is formulated in terms of information theoretic sense as follows:

$$\begin{aligned} \text{minimize} \quad & f_{o,I}(\tau) = -\log\left(f(\hat{x})\,\varepsilon\right) + \frac{1}{2}\left(n_x - 2\tau\right), \\ \text{subject to} \quad & g_1(\tau) = -\tau \le 0, \\ & g_2(\tau) = -\left(n_x - 2\tau\right) + \gamma_{\min} \le 0, \end{aligned} \tag{4.4.8}$$

where $f(x)$ is a Gaussian pdf in $\mathbb{R}^{n_x}$, and $\gamma_{\min}$ is a small constant obtained from the chi-square table, considering the degree of freedom (i.e., n_x) and determines the probability of the confidence level for the smallest hyperellipsoid, i.e, $\Pr\left(\left(n_x - 2\tau\right) \ge \gamma_{\min}\right) = \mathcal{Q}$, where $\mathcal{Q}$ denotes the upper-tail probability. This nonlinear convex objective function aims to minimize the Kullback-Leibler (KL) divergence of the Gaussian pdf $f(x)$ from the uniform density over the hyperellipsoid with volume $\varepsilon = C\left(n_x\right)|P|^{\frac{1}{2}}\left(n_x - 2\right)^{\frac{n_x}{2}}$

is the volume of the hypersphere with unit radius in $\mathbb{R}^{n_x}$. On the other hand, the second objective function is defined as

$$f_{o,J}(\tau) = \begin{cases} (n_x - 2\tau)^2 & \text{if } (n_x - 2\tau) > \gamma_{\min}, \\ 0 & \text{otherwise}, \end{cases} \tag{4.4.9}$$

considering the volume of the hyperellipsoid given by ε, and the constraint given by $g_2(\tau)$ in (4.4.8). The aim of this objective function is to improve the accuracy of the JoM estimator by increasing the resolution of the estimate [17]. This is because the first objective function in (4.4.8) takes smaller values as ε increases, i.e., $n_x - 2\tau \to n_x$ [5]. After the aggregation of these conflicting objective functions, the optimization problem has the form [5]

$$\begin{aligned} \text{minimize} \quad & f_m(\tau) = w_I f_{o,I}^{Trans}(\tau) + w_J f_{o,J}^{Trans}(\tau), \\ \text{subject to} \quad & g_1(\tau) = -\tau \leq 0, \\ & g_2(\tau) = -(n_x - 2\tau) + \gamma_{\min} \leq 0, \end{aligned} \tag{4.4.10}$$

where w_I and w_J are the relative weights of the normalized objective functions given by

$$f_{o,\xi}^{Trans}(\tau) = \frac{f_{o,\xi}(\tau) - F_{o,\xi}^{Min}}{F_{o,\xi}^{Max} - F_{o,\xi}^{Min}}, \forall \xi \in \{I, J\},$$

where $F_{o,\xi}^{Min}$ and $F_{o,\xi}^{Max}$ are the minimum and maximum values of the objective functions, respectively. The weights are adjusted according to the localization performance of the optimal JoTT filter while tracking a single Bernoulli target [5]. The Pareto-optimal solution to (4.4.10) can be obtained using the traditional sequential quadratic programming (SQP) [6] technique developed for nonlinear optimization problems. Details about the conflicting objective functions and the implementation of the JoM estimator proposed for the JoTT filter can be found in [5].

4.5 Reformulation of the JoM Estimator

In this section, we first propose an approximation to (4.4.6) using a mode-finding algorithm. Second, using Gaussian approximations to GM densities we generalize conflicting complex objective functions given by (4.4.8) and (4.4.9) for multi-Bernoulli RFS. As shown in (4.4.10), the normalized objective functions are aggregated with adaptive weights to determine the Pareto-optimal solution to τ. Then, we analyze the operation of the proposed JoM estimator. Considering this analysis and dynamics of multitarget tracking, i.e., target births and missed detections, the predictor coefficient of the autoregressive models which were employed to predict the weights in [5] are redefined. In addition, we define confidence levels for the smallest hyperellipsoid, i.e., γ_{min} according to the predictor coefficient.

4.5.1 Mode Finding and MAP Estimate

The Bayesian risk function[2] corresponding to the JoM estimator $J\left(Z\right)$ is defined as [17, p. 192]

$$\begin{aligned} R\left(J\right) =& E\left[C\left(X, J\left(Z\right)\right)\right], \\ =& \int C\left(X, J\left(Z\right)\right) f\left(X\right) \delta X, \end{aligned} \tag{4.5.1}$$

where C is the cost function that penalizes both discrepancies in cardinality and multitarget states. In the following, the argument Z is dropped from $J\left(Z\right)$ when the

[2]The dependence of the FISST probability density on Z is dropped for conciseness.

latter itself is an argument of another function. The Bayesian risk function is zero in a small neighborhood of $J(Z) = \hat{X}^{(JoM)}$ with volume $\varepsilon^{|J|}$. Therefore, (4.5.1) can be evaluated approximately using the definition of the set integral as follows [17, p. 193]:

$$R(J) \approx 2 - p_{|X|}\left(|\hat{X}^{(JoM)}|\right) - \frac{f\left(\hat{X}^{(JoM)}\right)\varepsilon^{|\hat{X}^{(JoM)}|}}{|\hat{X}^{(JoM)}|!}. \tag{4.5.2}$$

where the last two terms correspond to the total probability subtracted from the maximum cost for the JoM estimator. Thus, the JoM estimator aims to maximize

$$p_{|X|}(|J|) + \left.\frac{f(X)\,\varepsilon^{|J|}}{|J|!}\right|_{X=J(Z)}, \tag{4.5.3}$$

where the first term is always greater than the second one according to (4.4.2). The difference between them is significant, especially when $M_k > 1$ and $n \neq M_k$. The reason is that there are $B(M_k, n) = M_k!/(n!\,(M_k - n)!)$ different multi-Bernoulli RFSs with cardinality n, and all of them contribute to the pmf of their cardinality, i.e., $p_{|X|}(n)$ given by (4.4.2), whereas the probability distribution is computed exclusively around their MAP estimate, i.e., $\hat{X}_n^{(MAP)}$ given by (4.4.6). Therefore, the JoM estimate of a multitarget RFS is formulated as the optimization of the second term, i.e., by (4.4.5). In [17], this fact is explained as follows: the Bayesian risk function is minimized if $f(X)$ is maximized over all finite subsets $X \subseteq \mathbb{R}^{n_x}$ with $|X| = |J|$.

Suppose that at time k, the multi-Bernoulli RFS Ξ_k is described by the parameter set $\{(q_k^{(i)}, f_k^{(i)})\}_{i=1}^{M_k}$. Then, each term in (4.5.3) can be computed using (4.3.2) and (4.4.2). That is, the evaluation of (4.4.2) for the finite set $X_k \subset \{x_1, ..., x_{M_k}\}$ with cardinality n is given by (4.4.4). In addition, the evaluation of (4.3.2) is straightforward if $X_k = \emptyset$. Otherwise, (4.3.2) can be rewritten in a more appropriate form using the relationship $f(\{x_1, ..., x_n\}) \triangleq n! f(x_1, ..., x_n)$ as follows [23]:

$$f_k(X) = n! \prod_{i=1}^{M_k} \left(1 - q_k^{(i)}\right) \sum_{|X_k|=n} \prod_{\substack{j=1 \\ x_j \in X_k}}^{n} \frac{q_k^{(j)}}{1 - q_k^{(j)}} f_k^{(j)}(x_j), \tag{4.5.4}$$

where the summation is taken over all finite sets of existing targets with cardinality n and can be evaluated using the elementary symmetric function of the set given by

$$\Theta_{f,k} = \left\{ \frac{q_k^{(i)}}{1 - q_k^{(i)}} f_k^{(i)}(x_i) \right\}_{i=1}^{M_k}. \tag{4.5.5}$$

Thus, similar to (4.4.4), (4.3.2) can be expressed in a compact notation as

$$f_k(X) = n! \prod_{i=1}^{M_k} \left(1 - q_k^{(i)}\right) \sigma_{M_k,n}(\Theta_{f,k}), \tag{4.5.6}$$

where $\sigma_{M_k,n}(\Theta_{f,k}) = 1$ for $n = 0$ by convention [22].

Using (4.5.6), the solution to (4.4.6) determines the MAP estimate as the maximum contribution to $\sigma_{M_k,n}(\Theta_{f,k})$. However, the exact solution to (4.4.6) is still computationally complex. Instead, GM PHD of $f_k(X)$ can provide an approximate solution to (4.4.6) [23], but cannot adequately distinguish closely-spaced targets [20].

A suboptimal but tractable solution to (4.4.6) can be determined under the following constraint: the most-likely state estimates of each independent Bernoulli RFS are contained within a hyperellipsoid $S^{(i)}$ with volume $\varepsilon^{(i)}$ centered at the most significant mode of its GM density, i.e., $\hat{x}_i^{(MAP)}$. This constraint limits the search space but is reasonable, especially for those MeMBer filters categorized as measurement-oriented type MeMBer filters (e.g., IMeMBer and CBMeMBer filters). This is because they combine all Bernoulli RFSs updated with same measurement into one Bernoulli RFS and thus pdfs of Bernoulli RFSs tend to be localized [40, 41]. Note that as a limiting case of the MAP estimator, the JoM estimator assumes that $\varepsilon^{(i)}$ are so small that $f_k^{(i)}(x_i)$ can approximated as $f_k^{(i)}\left(\hat{x}_i^{(MAP)}\right)$ [23, 17].

The p.g.fl. provides an alternative representation of all statistics represented by a FISST probability density [24, 23]. For the multi-Bernoulli RFS Ξ_k, the p.g.fl. is given by [23]

$$G_k[h] = \prod_{i=1}^{M_k} \left(1 - q_k^{(i)} + q_k^{(i)} f_k^{(i)}[h]\right), \tag{4.5.7}$$

where $h(x)$ can be any real-valued function selected according to the quantity of interest [23, 20]. Considering our constraint on the search space, we define $h(x)$ as an indicator function of the event that x is a member of any measurable search space S, i.e.,

$$\mathbf{1}_S(x) = \begin{cases} 1 \text{ if } x \in S, \\ 0 \text{ otherwise.} \end{cases} \tag{4.5.8}$$

This choice of $h(x)$ defines a special version of the p.g.fl. known as the belief-mass function (bmf) as $\beta_k(S) = G_k[\mathbf{1}_S]$, where $S \subseteq \mathbb{R}^{n_x}$. Thus, $\beta_k(S)$ denotes the probability that X_k is completely contained in the hyperspace S [23, 20].

The FISST probability density can be constructed from the corresponding p.g.fl. using the functional derivative with respect to the finite set $X_k = \{x_1, ..., x_n\}$ as [23]

$$f_k(X) = \left.\frac{\delta G_k}{\delta X_k}[\mathbf{1}_S]\right|_{S=\emptyset} = \frac{\delta^n G_k}{\delta x_1 \cdots \delta x_n}[0]. \tag{4.5.9}$$

Thus, using the product rule [23], (4.4.6) can be rewritten as follows:

$$\hat{X}_n^{(MAP)} = \arg \sup_{x_1,\ldots,x_n} \sum_{|X_k|=n} \prod_{\substack{i=1 \\ X_k^{(i)} \in X_k}}^{n} \frac{1}{1-q_k^{(i)}} \frac{\delta G_k^{(i)}}{\delta X_k^{(i)}} [\mathbf{1}_{S^{(i)}}] \Bigg|_{S^{(i)}=\emptyset}, \tag{4.5.10}$$

where $G_k^{(i)}$ corresponds to the ith factor on the right hand side of (4.5.7). Therefore, its functional derivative is nonzero for $X_k^{(i)} = \{x_i\}$. Then, using the constraint defined on the hyperellipsoid $S^{(i)}$, it can be approximated as

$$\frac{\delta G_k^{(i)}}{\delta X_k^{(i)}} [\mathbf{1}_{S^{(i)}}] \Bigg|_{S^{(i)}=\emptyset} \cong q_k^{(i)} f_k^{(i)} \left(\hat{x}_i^{(MAP)}\right).$$

Consequently, a suboptimal but reasonable solution to (4.4.6) is obtained by solving

$$\begin{aligned} \hat{X}_n^{(MAP)} &\cong \arg \sup_{\hat{x}_1,\ldots,\hat{x}_n} \sum_{|X_k|=n} \prod_{\substack{i=1 \\ X_k^{(i)} \in X_k}}^{n} \frac{q_k^{(i)}}{1-q_k^{(i)}} f_k^{(i)} \left(\hat{x}_i^{(MAP)}\right), \\ &= \arg \sup_{\hat{x}_1,\ldots,\hat{x}_n} \sigma \left(\hat{\Theta}_{f,k}^{(MAP)}\right), \end{aligned} \tag{4.5.11}$$

where $\hat{\Theta}_{f,k}^{(MAP)}$ is the set given by (4.5.5) evaluated at $x_i = \hat{x}_i^{(MAP)}$ for $i = 1, \ldots, M_k$.

In order to solve (4.5.11), we can utilize the following recursive formulation of the elementary symmetric function [22]:

$$\sigma_{M_k,i}\left(\Theta_{f,k}\right) = \sigma_{M_k-1,i}\left(\Theta_{f,k} \backslash \theta_{f,k}^{(j)}\right) + \theta_{f,k}^{(j)} \sigma_{M_k-1,i-1}\left(\Theta_{f,k} \backslash \theta_{f,k}^{(j)}\right), \tag{4.5.12}$$

where $\Theta_{f,k} \backslash \theta_{f,k}^{(j)} = \{q_k^{(i)} (1-q_k^{(i)})^{-1} f_k^{(i)}(x_i) \Big| i = 1, \ldots, M_k \ \& \ i \neq j\}$, and thus the second term includes the jth Bernoulli target in all multi-Bernoulli RFSs of existing targets, whereas the first one does not. Hence, the maximum contribution to $\sigma_{M_k,n}\left(\hat{\Theta}_{f,k}\right)$ is determined by maximizing the second term in (4.5.12) for $\hat{\theta}_{f,k}^{(j)}$ recursively, i.e.,

$$\begin{aligned}\arg \sup_{\substack{\hat{\theta}_{f,k}\subseteq\hat{\Theta}_{f,k}\\ \left|\hat{\theta}_{f,k}\right|=n}} \sigma_{M_k,n}\left(\hat{\Theta}_{f,k}\right) &= \sigma_{M_k-1,n}\left(\hat{\Theta}_{f,k}\backslash\hat{\theta}_{f,k}^{(j),*}\right)\\ &+ \arg \sup_{\hat{\theta}_{f,k}\backslash\hat{\theta}_{f,k}^{(j)}} \sigma_{M_k-1,n-1}\left(\hat{\Theta}_{f,k}\backslash\hat{\theta}_{f,k}^{(j),*}\right)\sup_j \hat{\theta}_{f,k}^{(j)},\end{aligned} \tag{4.5.13}$$

where $\hat{\theta}_{f,k}^{(j),*} = \sup_j \hat{\theta}_{f,k}^{(j)}$ and $n \leq M_k - 1$. The first term on the right hand side of (4.5.13) is a constant since arg sup $_{\hat{\theta}_{f,k}}$ cannot return the maximum contribution to $\sigma_{M_k,n}\left(\hat{\Theta}_{f,k}\right)$ without $\hat{\theta}_{f,k}^{(j),*}$. This recursive process corresponds to determining the n largest elements of $\hat{\Theta}_{f,k}$ and can be performed by finding the indices of its elements sorted in descending order, i.e.,

$$\ell = sort\left(\hat{\Theta}_{f,k}, \text{`descend'}\right), \tag{4.5.14}$$

where ℓ denotes the index vector and its first n elements are the indices of Bernoulli targets in the multi-Bernoulli RFS that make the maximum contribution to $\sigma_{M_k,n}\left(\hat{\Theta}_{f,k}\right)$.

Since the JoM estimator is a MAP-like estimator [23, 17, 19], a mode-finding algorithm can be employed to extract the MAP estimates of Bernoulli targets from their GM densities. To ensure the clarity and integrity of this section, we assume that the MAP estimates of Bernoulli targets are available, i.e., $\left\{\hat{x}_i^{(\mathrm{MAP})}\right\}_{i=1}^{M_k}$, and thus $\Theta_{f,k}$ is evaluated at these estimates to obtain ℓ using (4.5.14). Nevertheless, the details of the mode-finding algorithm can be found in the Appendix.

4.5.2 Objective Functions for Multi-Bernoulli RFSs

In this subsection, we will generalize the conflicting objective functions given by (4.4.8) and (4.4.9) for multi-Bernoulli RFS. To do this, the mode-finding algorithm is utilized again to approximate GM densities as Gaussian densities. Thus, we can evaluate the multitarget KL divergence to derive the multi-Bernoulli version of (4.4.8). Finally, the generalization of (4.4.9) for multi-Bernoulli RFS is performed considering the constraint defined for the first objective function and the total volume for MAP estimates.

In [5], the derivation of (4.4.8) starts with the definition of a special convex set, where the most-likely state estimates from a Gaussian pdf are exclusively found, i.e.,

$$S = \left\{ x \in \mathbb{R}^{n_x} | \left(x - \mu\right)^T P^{-1} \left(x - \mu\right) \leq n_x - 2\tau \right\}, \tag{4.5.15}$$

where $0 < 2\tau < n_x$. The aim in (4.4.8) is to minimize the KL divergence around the mode of a Gaussian pdf. Therefore, GM densities are approximated by Gaussian densities around their most significant modes. For a given GM density, this approximation can be performed according to the number of its modes as follows [9]: *i*) if the GM density is unimodal, the covariance matrix of the approximated Gaussian density around the mixture's mode, i.e., $\mu_k^{(mode)}$ is computed as

$$P_k^{(mode)} = \sum_j^{l_k} \omega_k^{(j)} \left[P_k^{(j)} + \left(\mu_k^{(mode)} - \mu_k^{(j)}\right) \left(\mu_k^{(mode)} - \mu_k^{(j)}\right)^T \right]. \tag{4.5.16}$$

Otherwise, *ii*) if the GM is multimodal, its Hessian $\mathcal{H}_k$ at a mixture's mode contains information about its local concentration. In addition, the Hessian of a Gaussian density at its mode is given by

$$\mathcal{H}_k{}^{(mode)} = -|2\pi P_k|^{-\frac{1}{2}} P_k^{-1}. \tag{4.5.17}$$

Thus, setting the singular value decomposition (SVD) of the negative definite $\mathcal{H}_k{}^{(mode)}$, i.e., $\mathcal{H}_k{}^{(mode)} = -U\Lambda U^T$ to the known $\mathcal{H}_k$ the covariance matrix of the approximated Gaussian density around $\mu_k^{(mode)}$ is computed as

$$P_k^{(mode)} = \left|2\pi(-\mathcal{H}_k)^{-1}\right|^{-\frac{1}{n_x+2}} (-\mathcal{H}_k)^{-1}. \tag{4.5.18}$$

Recall from the Appendix that the Hessians of a GM density are computed while searching for its modes.

After this approximation, we can define (4.5.15) for each component of the multi-Bernoulli RFS X_k. Thus, the information theoretic part of the multi-objective function, i.e., (4.4.8) can be generalized for multi-Bernoulli RFS using the multitarget KL divergence defined as [23, 17]

$$K_k\left(U \| f\right) = \int U_k\left(X\right) \log\left(\frac{U_k\left(X\right)}{f_k\left(X\right)}\right) \delta X. \tag{4.5.19}$$

where $U_k\left(X\right)$ is the reference (or ground truth) FISST probability density, which $f_k\left(X\right)$ is compared to, and is defined over some bounded space, where both of them are continuous.

The convex set in (4.5.15) denotes a hyperellipsoid around the mode of a Gaussian pdf [5]. The volume of this hyperellipsoid represents the degree of accuracy to which most-likely state estimates can resolve [23, 17]. Suppose that there are n ground truth targets. Each of them is characterized by a uniform density over one of those hyperellipsoids centered at the MAP estimates, i.e., $\left\{\hat{x}_{\ell(j)}^{(\mathrm{MAP})}\right\}_{j=1}^{n}$. Thus, the ground truth FISST probability density in (4.5.19) is defined as [23, 17]

$$U_k(X) = \sum_{\varsigma} u_k^{(\varsigma(1))}(x_1)\ldots u_k^{(\varsigma(n))}(x_n) \qquad \text{if } |X_k| = n \text{ and } n > 0, \tag{4.5.20}$$

where ς denotes all permutations on the integers $1, \ldots, n$. For the JoM estimator given by (4.4.5), all hyperellipsoids have the same volume, i.e., $u_k^{(\varsigma(j))}(x_j) = \varepsilon_k^{-1}$ for $j = 1, \ldots, n$. Therefore, (4.5.20) can be simplified to [23, 17]

$$U_k(X) = n!\varepsilon_k^{-n} \text{ if } |X_k| = n \text{ and } n > 0, \tag{4.5.21}$$

and thus using (4.5.6), provided that $|X_k| = n$, (4.5.19) can be written as

$$K_k(U\| f) = \int_{S\times\ldots\times S} {\varepsilon_k}^{-n} \log \frac{{\varepsilon_k}^{-n}}{\prod_{i=1}^{M_k}\left(1 - q_k^{(i)}\right)\sigma_{M_k,n}(\Theta_{f,k})} dx_1 \ldots dx_n. \tag{4.5.22}$$

Since the volumes of each hyperellipsoid centered at $\hat{X}_{n,k}^{(MAP)} = \left\{\hat{x}_{\ell(j)}^{(MAP)}\right\}_{j=1}^{n}$, i.e., ε_k, are so small that (4.5.22) can be approximated for those n Bernoulli targets selected by ℓ as follows:

$$K_k(U\| f) \cong -\log\left(p_{|\hat{X}_k|}(n)\,{\varepsilon_k}^{n}\right) + \sum_{j=1}^{n} H_k\left(u\|\, \tilde{f}^{(\ell(j))}\right), \tag{4.5.23}$$

where $\tilde{f}_k^{(\ell(j))}(x)$ denotes Gaussian approximation to the GM density of the Bernoulli target indexed by $\ell(j)$,

$$H_k\left(u \| \tilde{f}^{(\ell(j))}\right) = -\int_S \varepsilon_k{}^{-1} \log\left(\tilde{f}_k^{(\ell(j))}(x)\right) dx, \tag{4.5.24}$$

is the measure of uncertainty introduced by using $\tilde{f}_k^{(\ell(j))}(x)$ instead of $u_k^{(\ell(j))}(x)$, and $p_{\left|\hat{X}_k\right|}(n)$ is the pmf of the cardinality, i.e., (4.4.2) evaluated for $\hat{X}_{n,k}^{(MAP)}$.

Using (4.5.15), we can determine the least upper bound for the KL divergence from (4.5.24) as [5]

$$H_k\left(u \| \tilde{f}\right) \leq -\log\left(\tilde{f}_k\left(\hat{x}^{(MAP)}\right)\right) + \frac{1}{2}(n_x - 2\tau). \tag{4.5.25}$$

Thus, the multi-Bernoulli RFS version of (4.4.8) is derived from (4.5.23) as follows:

$$\begin{aligned} \text{minimize} \quad & f_{o,I}(\tau) = -\log\left(p_{\left|\hat{X}_k\right|}(n)\right) + \\ & \sum_{j=1}^{n}\left[-\log\left(\tilde{f}_k^{(\ell(j))}\left(\hat{x}_{\ell(j)}^{(MAP)}\right)\varepsilon_k\right) + \frac{1}{2}(n_x - 2\tau)\right], \\ \text{subject to} \quad & g_1(\tau) = -\tau \leq 0, \\ & g_2(\tau) = -(n_x - 2\tau) + \gamma_{\min} \leq 0. \end{aligned} \tag{4.5.26}$$

It is important to note that $-\log\left(p_{\left|\hat{X}_k\right|}(n)\right)$ is constant if there exists at most one Bernoulli target. That is, $p_{\left|\hat{X}_k\right|}(1) = q_k$ if $M_k = 1$. Since $f_{o,I}(\tau)$ is a nonlinear convex objective function [5] this constant can be ignored. Therefore, (4.5.26) will reduce to (4.4.8) if $M_k = 1$.

According to (4.5.15), the volume of the individual hyperellipsoid around each

MAP estimate in $\hat{X}_{n,k}^{(MAP)}$ is given by [5]

$$\varepsilon_k^{(\ell(j))} = C\left(n_x\right)\left|\tilde{P}_k^{(\ell(j))}\right|^{\frac{1}{2}}\left(n_x - 2\tau\right)^{\frac{n_x}{2}}, \tag{4.5.27}$$

where $\tilde{P}_k^{(\ell(j))}$ is the covariance matrix of $\tilde{f}_k^{(\ell(j))}(x)$ and given by either (4.5.16) or (4.5.18). However, the uniform densities in (4.5.26) are identical, see (4.5.21). Therefore, a common hyperellipsoid must be defined around each MAP estimate in $\hat{X}_{n,k}^{(MAP)}$. This hyperellipsoid must maintain the total volume given by

$$\varepsilon_k^n = \prod_{j=1}^{n} \varepsilon_k^{(\ell(j))}, \tag{4.5.28}$$

where $\varepsilon_k^n = C(n_x)^n \left|\tilde{P}_k\right|^{\frac{n}{2}} (n_x - 2\tau)^{n\frac{n_x}{2}}$. Thus, the generalized variance of the common hyperellipsoid is obtained as the geometric average of individual ones:

$$\left|\tilde{P}_k\right| = \sqrt[n]{\prod_{j=1}^{n} \left|\tilde{P}_k^{(\ell(j))}\right|}.$$

Note that the common hyperellipsoid is utilized to evaluate (4.5.26). Nevertheless, the volume of interest for each Bernoulli target is given by (4.5.27) while satisfying (4.5.28). Hence, substituting $\varepsilon_k = \varepsilon_k^{(\ell(j))}$ for $j = 1, ..., n$ into (4.5.26) does not change the objective function in magnitude.

Considering the constraint $g_2(\tau)$ in (4.5.26) and the total volume given by (4.5.28), the multi-Bernoulli version of the second objective function, i.e., (4.4.9), which is in conflict with (4.5.26), is defined by

$$f_{o,J}(\tau) = \begin{cases} (n_x - 2\tau)^{2n} \ \text{ if } \ (n_x - 2\tau) > \gamma_{\min}, \\ 0 \ \text{ otherwise}, \end{cases} \tag{4.5.29}$$

where the exponent keeps the objective function in quadratic form if $n = 1$. Thus, the problems arising from linear objective functions are averted when the weighted sum method is used to model the multi-objective optimization [18].

4.5.3 Operation of the Proposed JoM Estimator

The proposed JoM estimator declares the $(n+1)$th Bernoulli target in addition to n Bernoulli targets with their MAP estimates $\hat{X}_{n,k}^{(MAP)}$ extracted using the mode-finding algorithm if the additional target satisfies

$$\frac{1}{(n+1)!} f_k\left(\hat{X}_{n+1}^{(MAP)}\right) \prod_{j=1}^{n+1} \varepsilon_{k,n+1}^{(\ell(j))} > \frac{1}{n!} f_k\left(\hat{X}_{n}^{(MAP)}\right) \prod_{j=1}^{n} \varepsilon_{k,n}^{(\ell(j))}, \tag{4.5.30}$$

where $\varepsilon_{k,n}^{(\ell(j))}$ and $\varepsilon_{k,n+1}^{(\ell(j))}$ are the volumes of hyperellipsoids computed for the Bernoulli target indexed by $\ell(j)$ when the multi-Bernoulli RFS version of (4.4.10) comprised of (4.5.26) and (4.5.29) is solved for $\hat{X}_{n,k}^{(MAP)}$ and $\hat{X}_{n+1,k}^{(MAP)}$, respectively.

Substituting $f_k(X)$ in (4.5.6) evaluated for $X = \hat{X}_{n,k}^{(MAP)}$ and $X = \hat{X}_{n+1,k}^{(MAP)}$ into (4.5.30) results in

$$\prod_{j=1}^{n+1} \tilde{f}_k^{(\ell(j))}\left(\hat{x}_{\ell(j)}^{(MAP)}\right) \varepsilon_{k,n+1}^{(\ell(j))} \frac{q_k^{(\ell(j))}}{1 - q_k^{(\ell(j))}} > \prod_{j=1}^{n} \tilde{f}_k^{(\ell(j))}\left(\hat{x}_{\ell(j)}^{(MAP)}\right) \varepsilon_{k,n+1}^{(\ell(j))} \frac{q_k^{(\ell(j))}}{1 - q_k^{(\ell(j))}}, \tag{4.5.31}$$

where $\tilde{f}_k^{(\ell(j))}(x)$ are the Gaussian densities defined by following the extraction of MAP estimates in subsection 4.5.2. The common terms on both sides of (4.5.31) cancel each other out and thus it can be reduced to

$$\tilde{f}_k^{(\ell(n+1))}\left(\hat{x}_{\ell(n+1)}^{(MAP)}\right)\varepsilon_{k,n+1}^{(\ell(n+1))}\frac{q_k^{(\ell(n+1))}}{1-q_k^{(\ell(n+1))}} > \rho_{n,n+1}, \tag{4.5.32}$$

where $\rho_{n,n+1}$ denotes the product of the ratios $\varepsilon_{k,n}/\varepsilon_{k,n+1}$ for the first n Bernoulli targets. Consequently, the $(n+1)$th Bernoulli target is declared, if its existence probability satisfies

$$q_k^{(\ell(n+1))} > \frac{\rho_{n,n+1}}{\rho_{n,n+1}+\tilde{f}_k^{(\ell(n+1))}\left(\hat{x}_{\ell(n+1)}^{(MAP)}\right)\varepsilon_{k,n+1}^{(\ell(n+1))}}, \tag{4.5.33}$$

According to (4.5.33), the dynamics of the JoM estimator depend on both the existence probability of the $(n+1)$th Bernoulli target and how well it is localized, considering the first n Bernoulli targets. Therefore, "no-target" decision with the probability of $1-q_k^{(\ell(n+1))}$ for the $(n+1)$th Bernoulli target can be revoked if its localization performance is comparable with those of the first n Bernoulli targets [23].

4.5.4 Adaptive Objective Weights and Confidence Levels

In [5], the weights of the conflicting objective functions in (4.4.10) are adjusted while tracking a single Bernoulli target by the optimal JoTT filter. For this purpose, two autoregressive (AR) models of order 1 are employed to predict these weights, whose sum is unity i.e., $\omega_{J,k} + \omega_{I,k} = 1$. The predictor coefficient of these weights is determined according to [5][3]

$$\mathrm{B}_k = \frac{|P_{k-1}|^{1/2}}{|P_k|^{1/2}} \mathbf{1}_A (q_k), \tag{4.5.34}$$

where the first term is the ratio of volumes of the hyperellipsoids around the MAP estimates of a Bernoulli target with the same chi-square quantile at times $k-1$ and k, and the indicator function keeps the weights at their initial states before declaring a real target, i.e., when $q_k \notin A$, where $A = [q_{\min}, 1]$.

In the single target case, the covariance matrix obtained from the optimal JoTT filter would either converge to its steady-state with correct detections or fluctuate due to missed detections after the target is declared. However, the value of B_k abruptly changes after target births while tracking multiple targets. Therefore, the missed detections of the converged targets cannot be monitored through B_k until the covariance matrices of newborn targets converge to their steady-states. To address this issue, (4.5.34) is redefined for $\hat{X}^{(MAP)}_{n,k}$ as

$$\mathrm{B}_k = \frac{\left|\tilde{P}_{\min}\right|^{1/2}}{\left|\tilde{P}_k\right|^{1/2}} \text{ if } q_k^{(\ell(i))} \geq q_{\min} \text{ for } i = 1, ..., n, \tag{4.5.35}$$

where $\left|\tilde{P}_{\min}\right| = \min\left\{\left|\tilde{P}_t\right|\right\}_{t=1}^{k-1}$. Thus, the weights at time $k-1$ are adjusted according to how well multi-Bernoulli targets are localized at time k, compared to the best localization performance achieved until that time.

Considering B_k in (4.5.35), the probability of confidence level given by $\gamma_{\min}$ is also

[3]Here, to avoid confusion, the predictor coefficient in [5] is denoted by B_k.

adjusted. Thus, its value indicates the smallest hyperellipsoid for multi-Bernoulli targets under different localization conditions named as “improvement”, “fluctuation”, “occlusion”, and “birth”. Table 4.1 shows how $\gamma_{\min}$ is adjusted under these conditions according to $\Pr\left((n_x - 2\tau) \geq \gamma_{\min}\right) = \mathcal{Q}$ using the chi-square table.

Table 4.1: Adaptive Confidence Levels

Condition	B_k	$\mathcal{Q}$
“improvement”	$[1, \infty)$	%97.5
“fluctuation”	$[\mathrm{B}_{\max}, 1)$	%95
“occlusion”	$[\mathrm{B}_{\min}, \mathrm{B}_{\max})$	%90
“birth”	$[0, \mathrm{B}_{\min})$	%85

In Table 4.1, the parameters $\mathrm{B}_{\min}$ and $\mathrm{B}_{\max}$ are the control limits for the predictor coefficient and are set at $\mathrm{B}_{\min} = 0.1$ and $\mathrm{B}_{\max} = 0.9$ [5]. Therefore, none of the conflicting objective functions can completely dominate the other. As can be seen from Table 4.1, the value of $\gamma_{\min}$ gradually increases while the upper-tail probability of the chi-square variable in (4.5.15) decreases from “improvement” condition to “birth” condition. Hence, the value of $\rho_{n,n+1}$ in (4.5.33) decreases, especially when the $(n+1)$th Bernoulli target in addition to $\hat{X}_{n,k}^{(MAP)}$ either undergoes an occlusion or appears as a newborn target.

In summary, using γ_{min} obtained from Table 4.1 the conflicting objective functions in (4.5.26) and (4.5.29) are determined and normalized. Then, their aggregation is performed with the adaptive weights computed using (4.5.35). Details on the computation of adaptive weights can be found in [5].

4.6 Simulation Results

In this section, the proposed JoM estimator is compared to the MaM estimator. For this purpose, the optimal subpattern assignment (OSPA) metric is used as the performance criterion [32, 30]. The OSPA metric compares two finite sets, considering the difference in their cardinalities (i.e., cardinality error) and the positional distances between their corresponding elements (i.e., localization error) after an optimal assignment. The sensitivity of this metric to these two errors is adjusted by tuning the cut-off parameter c and the order parameter p. In addition, the type of the localization error computed in the OSPA metric is determined by the order parameter. For example, the average Euclidean distance and the root mean square (RMS) error are computed by setting $p = 1$ and $p = 2$, respectively.

The major performance difference between these two estimators are expected to arise from their cardinality errors due to false alarm or missed detections. Therefore, the OSPA parameters are set at $p = 2$ and $c = 300$. Hence, the cardinality error is penalized more significantly than the localization error. Otherwise, OSPA metric would be biased in favor of the MaM estimator with smaller localization errors, compared to those computed for the proposed JoM estimator when the cardinality estimate obtained from the MaM estimator is not accurate [32].

In simulations, the GM-IMeMBer and GM-CBMeMBer filters are used for the same MTT scenario, but with different state-dependent detection probabilities in the exponential form as in (4.6.1) and (4.6.2). The exponential detection probability is a function of the range between the sensor at $[-300, -300]$ and targets. Therefore, especially in low-observable conditions, missed detections depend on the positions of targets. Thus, the MaM estimator would be expected to suffer from this because of its cost function. Details on the GM implementations of these two filters with exponential detection probability can be found in the supplementary material accompanying this paper [3]. The state vector of each individual target consists of positions and velocities in x-y directions, i.e., $x = [p_x,\, p_y,\, v_x,\, v_y]^T$. If targets survive between two consecutive time scans with the constant probability $p_{S,k}(x) = 0.95$, their states propagate independently according to the linear white noise acceleration model [2]:

$$x_k = \mathcal{F}_k x_{k-1} + \mathcal{G} w_{k-1},$$

where w_{k-1} is a white Gaussian noise with standard deviations $\sigma_{v,x} = 0.3$ m/s^2 and $\sigma_{v,y} = 0.3$ m/s^2, and the system matrices are given by

$$\mathcal{F}_k = \begin{bmatrix} I_{2\times 2} & \Delta I_{2\times 2} \\ 0_{2\times 2} & I_{2\times 2} \end{bmatrix}, \mathcal{G} = \begin{bmatrix} \frac{\Delta^2}{2} I_{2\times 2} \\ \Delta I_{2\times 2} \end{bmatrix},$$

where $I_{2\times 2}$ and $0_{2\times 2}$ denote 2×2 identity and zero matrices, respectively, and Δ denotes the time interval between two consecutive time scans and is set at $\Delta = 1$s in simulations.

The GM-IMeMBer and GM-CBMeMBer filters explore target births according to the multi-Bernoulli RFS described by the parameter set $\{(q_{\Gamma,k}^{(i)}, f_{\Gamma,k}^{(i)})\}_{i=1}^{5}$, where the existence probabilities are set to $q_{\Gamma,k}^{(i)} = 0.05$, and the pdfs are given by Gaussian densities $f_{\Gamma,k}^{(i)}(x) = \mathrm{N}\left(x\,; \mu_{\Gamma,k}^{(i)}, P_\Gamma\right)$ with means $\mu_{\Gamma,k}^{(1)} = [-45,\, 150,\, 0,\, 0]^T$, $\mu_{\Gamma,k}^{(2)} = [45,\, 150,\, 0,\, 0]^T$, $\mu_{\Gamma,k}^{(3)} = [125,\, 60,\, 0,\, 0]^T$, $\mu_{\Gamma,k}^{(4)} = [75,\, 105,\, 0,\, 0]^T$, $\mu_{\Gamma,k}^{(5)} = [-170,\, 185,\, 0,\, 0]^T$ and common covariance matrices $P_{\Gamma,k} = diag([25, 25, 15, 15])$.

Measurements originating from targets are noisy spatial components of their states.

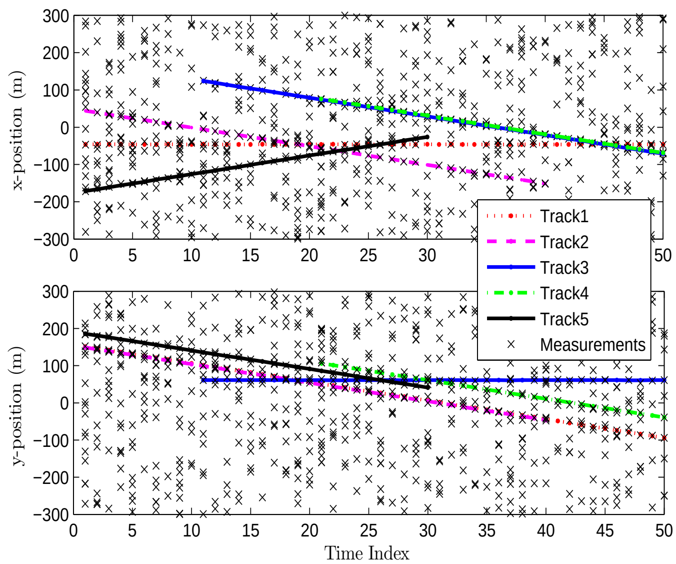

Figure 4.1: x and y components of the true target tracks and their measurements observed in clutter.

That is, they are modeled as a linear-Gaussian process:

$$z_k = Hx_k + \eta_k,$$

where the observation matrix is given by $H = [I_{2\times2}, 0_{2\times2}]$, and η_k is a white Gaussian noise with standard deviations $\sigma_{p,x} = 1\text{m}$ and $\sigma_{p,y} = 1\text{m}$. In addition to noisy target-originated measurements, the measurement set includes clutters. In simulations, clutter is modeled as a Poisson RFS with the mean rate of 10 per scan, i.e., $\lambda_c = 10$ and uniform spatial distribution over the surveillance region $V = [-300\text{m}, 300\text{m}] \times [-300\text{m}, 300\text{m}]$. Fig. 4.1 shows the x and y components of the true target tracks in the MTT scenario together with their measurements in clutter. It can be seen that targets with tracks 1 and 2 at time step $k = 19$ and tracks 3 and 4 at time step $k = 30$ cross each other. The performances of the two estimators are evaluated by running this scenario with random measurement sets for 5000 Monte Carlo runs.

In GM implementations, the maximum number of Gaussian components is set to $J_{max} = 25$. After terminating Bernoulli targets with existence probabilities less than $Th = 0.02$, their components are pruned and merged with thresholds $Pr = 0.01$ and $U = 2.5$ according to the algorithm in [33]. In addition, the control parameters of the fixed-point search to find the modes of a given GM density are set as follows: max_it = 1000, min_step = $10^{-4} \times \Lambda_{min}^{\frac{1}{2}}$, where Λ_{min} is the smallest positive eigenvalue of the covariance matrices, and max_eig = 0. Finally, the parameter $q_{\min}$ in (4.5.35) is determined subject to observable conditions. Thus, any Bernoulli target with $q_k < q_{\min}$ cannot be declared by the JoM estimator.

4.6.1 Example 1: Low-Observable Conditions

In this example, the performances of the GM-IMeMBer filters with the proposed JoM estimator and the MaM estimator are compared under moderately low-observable conditions. As indicated in Section 4.3.2, no restrictive assumption on the probability of target detection is required for the IMeMBer filter. Therefore, the state-dependent probability of target detection is defined as

$$p_D(x) = \frac{0.85 N\Big(Hx; [-300, -300]^T, 1000^2 I_{2\times 2}\Big)}{N\Big([-300, -300]^T; [-300, -300]^T, 1000^2 I_{2\times 2}\Big)}. \tag{4.6.1}$$

As can be seen from Fig. 4.2, the target detection probability given by (4.6.1) takes values in the range of 0.61 to 0.85. According to the trajectories of the five targets, their detection probabilities change approximately in the interval [0.72, 0.81].

Considering low-observable conditions, the minimum existence probability to declare a Bernoulli target is set at $q_{\min} = 1/3$. Since cardinality errors due to missed detections depend on positions of targets, the proposed JoM estimator is expected to outperform the MaM estimator by yielding more reliable estimates. In the following, the performances of these two estimators are compared by evaluating the accuracy and stability of their cardinality estimates.

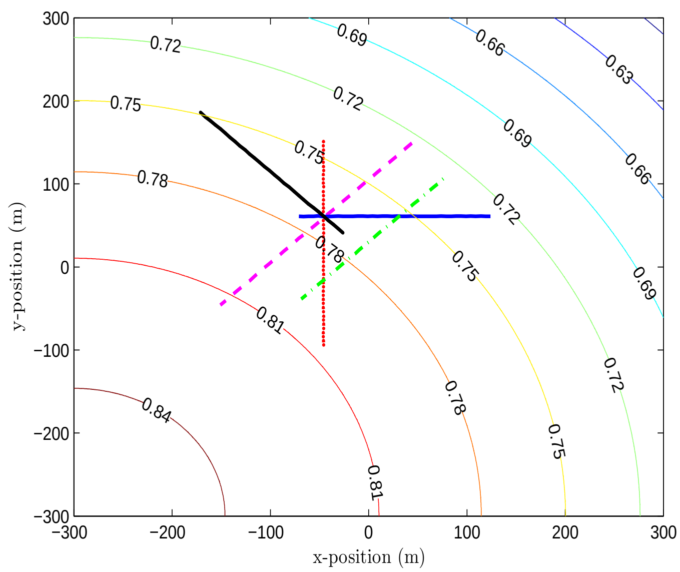

Figure 4.2: Trajectories of five targets and the state-dependent probability of target detection in the surveillance region.

In Fig. 4.3, the average cardinality estimates over 5000 Monte Carlo runs and their $\pm 1\sigma$ standard deviations for the two estimators are shown. The RMS errors measured from the plots of the JoM and MaM estimators are approximately 0.25 and 0.34, respectively, whereas their standard deviations are almost the same. As can be seen from Fig. 4.3, the cardinality estimate from the MaM estimator is negatively biased, compared to that from the JoM estimator. Therefore, these values indicate that the proposed JoM estimator produces more accurate cardinality estimates than the MaM estimator. Another important result obtained from Fig. 4.3 is that the response times of the JoM estimator to track terminations at times $k = 31$ and $k = 41$ are slower than those of the MaM estimator.

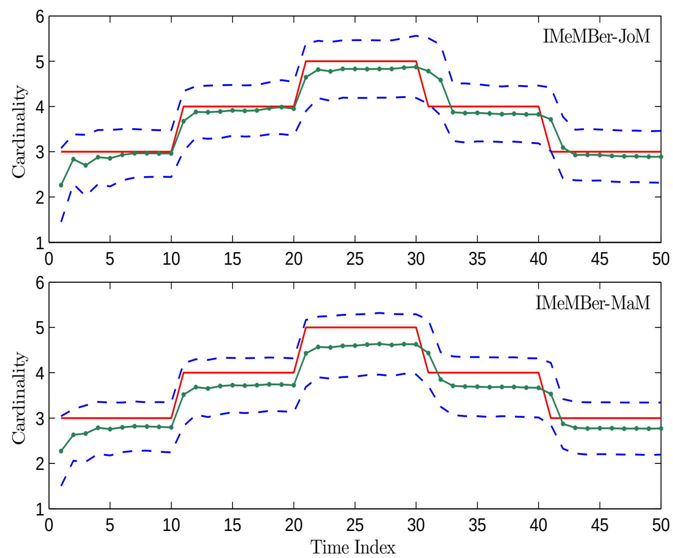

Figure 4.3: The average cardinality estimates over 5000 Monte Carlo runs: true cardinality (red solid line), estimated cardinality (green dotted line), and its ± 1 standard deviations (blue dashed lines) for the IMeMBer filters with the JoM and MaM estimators.

In Fig. 4.4, the OSPA metric shows the cardinality performances of the two estimators. As expected, it corroborates the results inferred from Fig. 4.3. That is, the track maintenance quality of the proposed JoM estimator is better than that of the MaM estimator. However, the instantaneous peaks are observed at times $k = 31$ and $k = 41$ due to the corresponding track termination latencies in Fig. 4.3. Notice that the JoM estimator suffers from these track termination latencies more than the MaM estimator under this condition. The reason is that the MaM estimator responds to target deaths by considering only existence probabilities, whereas the JoM estimator considers existence probabilities together with localization performances. Since the localization performances deteriorate gradually after target deaths at times $k = 30$ and $k = 40$, the track termination decisions in the JoM estimator is slower, compared to those in the MaM estimator.

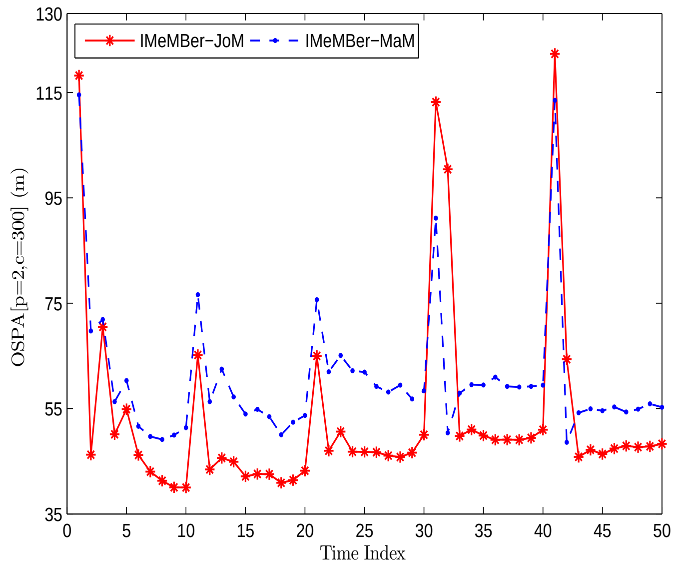

Figure 4.4: The average OSPA over 5000 Monte Carlo runs for the IMeMBer filters with the JoM and MaM estimators.

4.6.2 Example 2: High-Observable Conditions

In this example, the performances of the CBMeMBer filters with the proposed JoM estimator and the MaM estimator are compared. Since the CBMeMBer filter is derived under the assumption of high-observable conditions, i.e., $p_{D,k}(x) \approx 1$, the state-dependent probability of target detection is defined as

$$p_D(x) = \frac{0.98N\left(Hx; [-300, -300]^T, 3000^2 I_{2\times 2}\right)}{N\left([-300, -300]^T; [-300, -300]^T, 3000^2 I_{2\times 2}\right)}. \tag{4.6.2}$$

Note that the aim of our comparison is to demonstrate that there is a performance difference between the two estimators when estimates based on the cardinality distribution are not accurate, especially under low-observable conditions [23, 19]. Otherwise, if the cardinality estimates from the MaM estimator are as accurate as those from the JoM estimator, these two MAP-like estimators perform similarly.

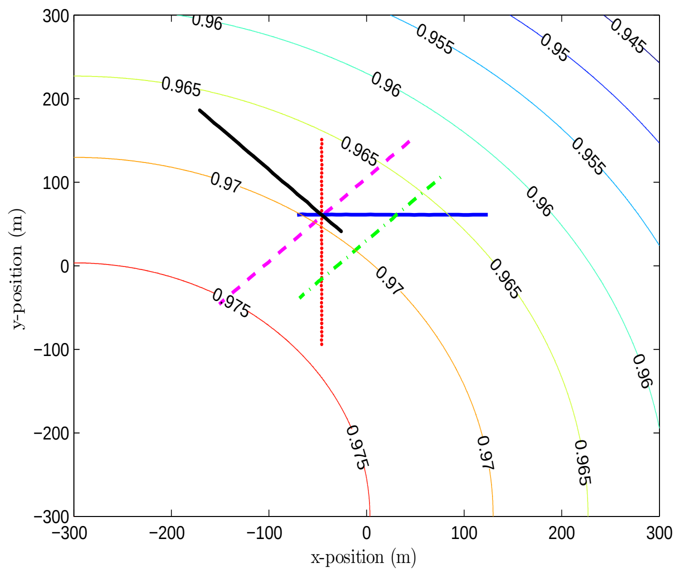

Figure 4.5: Trajectories of five targets and the state-dependent probability of target detection in the surveillance region.

As can be seen from Fig. 4.5, the target detection probability given by (4.6.2) takes values in the range of 0.945 to 0.98. According to the trajectories of the five targets, their detection probabilities change approximately in the interval [0.963, 0.975]. Since cardinality errors due to missed detections rarely occur, the proposed JoM estimator is not expected to outperform the MaM estimator. Accordingly, the minimum existence probability to declare a Bernoulli target is set at $q_{\min} = 0.45$. On the other hand, all other model and scenario parameters are the same as those in subsection 4.6.1.

In Fig. 4.6, the average cardinality estimates over 5000 Monte Carlo runs and their $\pm 1\sigma$ standard deviations for the two estimators are shown. It can be seen that their cardinality estimates converge to the true cardinality. In addition, the RMS errors measured from the plots are both 0.09 approximately. These results demonstrate that the performances of the proposed JoM and MaM estimators are similar under high-observable conditions.

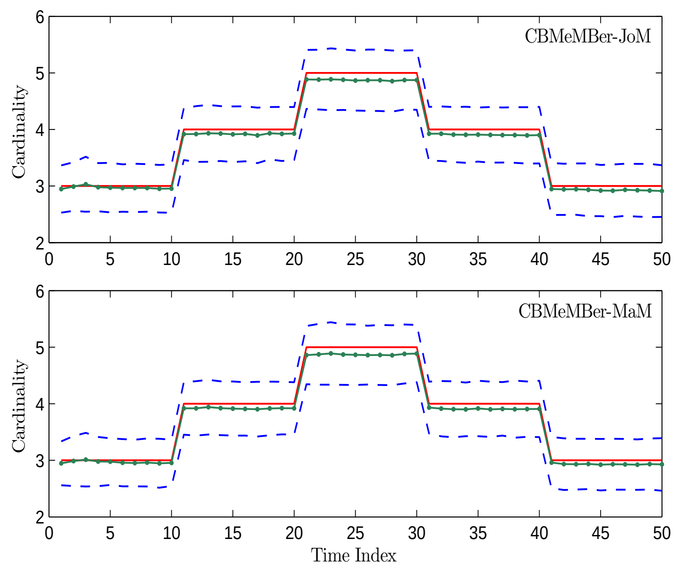

Figure 4.6: The average cardinality estimates over 5000 Monte Carlo runs: true cardinality (red solid line), estimated cardinality (green dotted line), and its ± 1 standard deviations (blue dashed lines) for the CBMeMBer filters with the JoM and MaM estimators.

In Fig. 4.7, the OSPA metric shows the cardinality performances of the two estimators. As expected, it corroborates the results inferred from Fig. 4.6. That is, it can be seen that the performances of the two estimators are almost the same even though the OSPA metric with $c = 300$ is sensitive to errors in cardinality estimates.

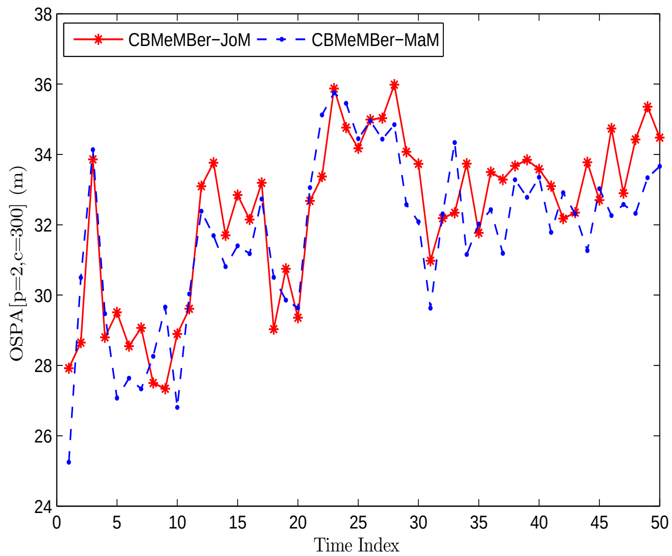

Figure 4.7: The average OSPA over 5000 Monte Carlo runs for the CBMeMBer filters with the JoM and MaM estimators.

4.7 Conclusion

In this paper, the JoM estimator proposed for the JoTT filter was reformulated to obtain the estimate of the multi-Bernoulli RFS. For this purpose, a mode-finding algorithm was employed to obtain the MAP estimates of Bernoulli targets from their GM pdfs. Thus, the conflicting objective functions designed to determine the Pareto-optimal value of the unknown JoM estimation constant were generalized using local statistics of GM densities. In addition, the weights of the conflicting objectives and the confidence level were adjusted considering the characteristics of the multitarget tracking and the proposed JoM estimator.

Simulations demonstrate that the proposed JoM estimator is more reliable in the cardinality estimate than the MaM estimator under moderately low-observable conditions. However, it suffers from track termination latency more than the MaM estimator because of gradually deteriorating localization performances after target deaths. Finally, the performance comparison under high-observable conditions demonstrates that the two estimators achieve the same accuracy and stability in cardinality estimates. This is the expected since the cardinality error due to missed detections is low under high-observable conditions.

4.8 Appendix

There is no analytical solution to find the number of modes of a GM density and their positions over the state space [8, 11]. However, for one-dimensional GM densities and GM densities with either homoscedastic or isotropic covariance matrices, the number of modes is not greater than the number of components. In addition, they are positioned inside the convex hull of GM centroids [8, 11]. It is important to note that the mode of interest in the JoM estimator is the MAP estimate, i.e., the mode with the highest probability among all modes of a GM density. Hence, finding all minor modes are not necessary.

In [8], the fixed-point search was proposed to find modes of GM densities with arbitrary covariance matrices. This algorithm is described as an iterative hill-climbing algorithm since it starts searching modes from each component centroid and stops if a local maximum of the GM density is found. The local maximum, minimum and saddle points are stationary points of multivariate GM densities. Therefore, the fixed-point search first takes the gradient of a given GM density, i.e.,

$$\nabla f_k\left(x\right) = \sum_{j=1}^{l_k} \omega_k^{(j)} N\left(x; \mu_k^{(j)}, P_k^{(j)}\right) \left(P_k^{(j)}\right)^{-1} \left(x - \mu_k^{(j)}\right), \tag{4.8.1}$$

where $f_k(x) = \sum_{j=1}^{l_k} \omega_k^{(j)} N(x; \mu_k^{(j)}, P_k^{(j)})$. Then, (4.8.1) is solved for x, and thus the fixed-point iteration is obtained as

$$x_{n+1} = \left(\sum_{j=1}^{l_k} f_{\omega,k}^{(j)}\left(x_n\right) \left(P_k^{(j)}\right)^{-1}\right)^{-1} \sum_{j=1}^{l_k} f_{\omega,k}^{(j)}\left(x_n\right) \left(P_k^{(j)}\right)^{-1} \mu_k^{(j)}, \tag{4.8.2}$$

where $f_{\omega,k}^{(j)}(x) = \omega_{k-1}^{(j)} N(x; \mu_k^{(j)}, P_k^{(j)})$.

The fixed-point search continues until one of the following convergence conditions is satisfied: *i*) the maximum number of iterations (max_it) is reached or *ii*) the distance between successive iterations is less than a priori threshold value (min_step). To confirm if the stationary point is a local maximum of a given GM density, the fixed-point search checks its Hessian given by

$$\mathcal{H}_k = \sum_{j=1}^{l_k} \omega_k^{(j)} N\left(x; \mu_k^{(j)}, P_k^{(j)}\right)\left(P_k^{(j)}\right)^{-1} \left[\left(x - \mu_k^{(j)}\right)\left(x - \mu_k^{(j)}\right)^T - P_k^{(j)}\right]\left(P_k^{(j)}\right)^{-1}. \tag{4.8.3}$$

More explicitly, if (4.8.3) evaluated at a stationary point is negative definite, the point is characterized as a local maximum, i.e., $\Lambda_{max,\mathcal{H}_k} <$ max_eig, where $\Lambda_{max,\mathcal{H}_k}$ is the maximum eigenvalue of the Hessian, and max_eig is a nonnegative parameter used to eliminate minima without missing any local maximum [8].

The fixed-point search proposed in [8] corresponds to the mean-shift algorithm for the Gaussian kernel, which can be derived as an expectation maximization (EM) algorithm [11, 10]. Thus, using the convergence properties of the EM algorithm, some conclusions can be drawn about the fixed-point search [11, 10]: *i*) its practical convergence is a local maximum from any given starting point and *ii*) the rate of its convergence is usually linear (i.e., slow) unless the components of a GM density are well-separated. Note that closely-spaced components of each GM density are merged after pruning its negligible components in the GM implementations of the MeMBer filters. Purification of GM densities would also accelerate the fixed-point search without missing any significant mode [8].

Bibliography

[1] D. Angelosante, E. Biglieri, and M. Lops. Multiuser detection in a dynamic environment–Part II: Joint user identification and parameter estimation. *IEEE Trans. Info Theory*, 55(9):2365–2374, Sept. 2009.

[2] Y. Bar-Shalom, P. Willtet, and X. Tian. *Tracking and Data Fusion: A Handbook of Algorithms*. YBS Publishing, Storrs CT, 2011.

[3] E. Baser and T. Kirubarajan. The GM implementations of the MeMBer filters with exponential detection probability. `http://ieeexplore.ieee.org`.

[4] E. Baser, T. Kirubarajan, M. Efe, and B. Balaji. Improved MeMBer filter with modeling of spurious targets. *IET Radar, Sonar and Navigation*, 10(2):285–298, Feb. 2016.

[5] E. Baser, M. McDonald, T. Kirubarajan, and M. Efe. A joint multitarget estimator for the joint target detection and tracking filter. *IEEE Trans. Signal Process.*, 63(15):3897–3911, May 2015.

[6] D.P. Bertsekas. *Nonlinear Programming*. Athena Scientific, Nashua, New Hampshire, 1999.

[7] E. Biglieri, E. Grossi, and M. Lops. Random-set theory and wireless communications. *Foundations and Trends in Communications and Information Theory*, 7(4):317–462, Aug. 2012.

[8] M. C.-Perpinan. Mode-finding for mixtures of Gaussian distributions. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22(11):1318–1323, Nov. 2000.

[9] M. C.-Perpinan. *Continuous latent variable models for dimensionality reduction and sequential data reconstruction*. PhD thesis, Department of Computer Science, University of Sheffield, UK, 2001.

[10] M. C.-Perpinan. Gaussian mean-shift is an EM algorithm. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 29(5):767–776, May. 2007.

[11] M. C.-Perpinan and K. Williams. *Scale Space Methods in Computer Vision*, chapter On the number of modes of a Gaussian mixture, pages 625–640.

[12] S. Challa, B.-N. Vo, and X. Wang. Bayesian approaches to track existence-IPDA and random sets. In *Proc. 5th Int'l Conf. Information Fusion*, pages 1228–1235. IEEE, 2002.

[13] O. Erdinc, P. Willett, and Y. Bar-Shalom. Probability hypothesis density filter for multitarget multisensor tracking. In *Proc. 8th Int'l Conf. Information Fusion*, pages 146–153. IEEE, Jul. 2005.

[14] O. Erdinc, P. Willett, and Y. Bar-Shalom. A physical-space approach for the probability hypothesis density and cardinalized probability hypothesis density filters. In *Proc. Signal and Data Processing of Small Targets*, pages 623619–623619–12. SPIE, May 2006.

[15] O. Erdinc, P. Willett, and Y. Bar-Shalom. The bin occupancy filter and its connection to the PHD filter. *IEEE Trans. Signal Process.*, 57(11):4232–4246, Jun. 2009.

[16] D. Franken, M. Schmidt, and M. Ulmke. 'Spooky action at a distance' in the cardinalized probability hypothesis density filter. *IEEE Trans. Aerosp. and Electron. Syst.*, 45(4):1657–1664, Oct. 2009.

[17] I. Goodman, R. Mahler, and H. Nguyen. *Mathematics of Data Fusion.* Kluwer Academic Publisher, Netherlands, 1997.

[18] O. Grodzevich and O. Romanko. Discussions on normalization and other topics in multi-objective optimization. In *Proc. of Fields-MITACS Industrial Problem-Solving Workshop (FMIPW)*, Toronto, Canada, Aug. 2006.

[19] R. Mahler. Why multi-source, multitarget data fusion is tricky. In *Proc. of IRIS National Symposium on Sensor and Data Fusion*, pages 135–153, May 1999.

[20] R. Mahler. Multitarget Bayes filtering via first-order multitarget moments. *IEEE Trans. Aerosp. and Electron. Syst.*, 39(4):1152–1178, Oct. 2003.

[21] R. Mahler. 'Statistics 101' for multisensor, multitarget data fusion. *IEEE Aerosp. and Electron. Sys. Magazine Part 2: Tutorials*, 19(1):53–64, Jan. 2004.

[22] R. Mahler. PHD filters of higher order in target number. *IEEE Trans. Aerosp. Electron. Syst.*, 43(4):1523–1542, Oct. 2007.

[23] R. Mahler. *Statistical Multisource-Multitarget Information Fusion.* Artech House, Boston MA, 2007.

[24] R. Mahler. 'Statistics 102' for multisource-multitarget detection and tracking. *IEEE J. Sel.Topics Signal Process.*, 7(3):376–388, Jun. 2013.

[25] R.T. Marler and J.S. Arora. Survey of multi-objective optimization methods for engineering. *Structural and Multidisciplinary Optimization*, 26(6):369–395, Mar. 2004.

[26] D. Musicki, R. Evans, and S. Stankovic. Integrated probabilistic data association. *IEEE Trans. Automatic Control*, 39(6):1237–1241, Jun. 1994.

[27] P. Ngatchou, A. Zarei, and M.A. El-Sharkawi. Pareto multi objective optimization. In *Proc. of the 13th Int'l Conf. on Intell. Syst. Application to Power Systems*, pages 84–91, Va, USA, Nov. 2005.

[28] F. Papi, B.-N. Vo, B.-T. Vo, C. Fantacci, and M. Beard. Generalized labeled multi-Bernoulli approximation of multi-object densities. *IEEE Trans. Signal. Process.*, 63(20):5487–5497, Jul. 2015.

[29] S. Reuter, B.-T. Vo, B.-N. Vo, and K. Dietmayer. The labeled multi-Bernoulli filter. *IEEE. Trans. Signal Process.*, 62(12):3246–3260, Jun. 2014.

[30] B. Ristic, B.-N. Vo, D. Clark, and B.-T. Vo. A metric for performance evaluation of multi-target tracking algorithms. *IEEE. Trans. Signal Process.*, 59(7):3452–3457, Jul. 2011.

[31] B. Ristic, B.-T. Vo, B.-N. Vo, and A. Farina. A tutorial on Bernoulli filters: Theory, implementation and applications. *IEEE Trans. Signal Process.*, 61(13):3406–3430, Jul. 2013.

[32] D. Schuhmacher, B.-T. Vo, and B.-N. Vo. A consistent metric for performance evaluation of multi-object filters. *IEEE. Trans. Signal Process.*, 56(7):3447–3457, Aug. 2008.

[33] B.-N. Vo and W.-K. Ma. The Gaussian mixture probability hypothesis density filter. *IEEE Trans. Signal Process.*, 54(11):4091–4104, Nov. 2006.

[34] B.-N. Vo, B.-T. Vo, and D. Phung. Labeled random finite sets and the Bayes multi-target tracking filter. *IEEE Trans. Signal. Process.*, 62(24):6554–6567, Dec. 2014.

[35] B.-T. Vo. *Random finite sets in multi-object filtering*. PhD thesis, Electrical, Electronic and Computer Engineering, University of Western Australia, Australia, 2008.

[36] B.-T. Vo and B.-N. Vo. Labeled random finite sets and multi-object conjugate priors. *IEEE. Trans. Signal Process.*, 61(13):3460–3475, Jul. 2013.

[37] B.-T. Vo, B.-N. Vo, and A. Cantoni. Analytic implementations of the cardinalized probability hypothesis density filter. *IEEE Trans. Signal Process.*, 55(7):3553–3567, Jul. 2007.

[38] B.-T. Vo, B.-N. Vo, and A. Cantoni. The cardinality balanced multitarget multi-Bernoulli filter and its implementations. *IEEE Trans. Signal Process.*, 57(2):409–423, Feb. 2009.

[39] B.-T. Vo, B.-N. Vo, R. Hoseinnezhad, and R. Mahler. Robust multi-Bernoulli filtering. *IEEE J. Sel. Topics Signal Process.*, 7(3):399–409, Jun. 2013.

[40] J. L. Williams. Marginal multi-Bernoulli filters: RFS derivation of MHT, JIPDA and association-based MeMBer. 2014. http://arxiv.org/abs/1203.2995.

[41] J. L. Williams. An efficient, variational approximation of the best fitting multi-Bernoulli filter. *IEEE. Trans. Signal Process.*, 63(1):258–273, Jan. 2015.

Chapter 5

Conclusions and Future Research

The work in this thesis studied the problem of cardinality bias observed in the MeMBer filter and proposed a practical and efficient use of the JoM estimator for RFS (multi-)Bernoulli distributions.

To address the positive cardinality bias in the MeMBer filter, the equivalence between the multi-target posterior distribution of data-induced targets and the multi-Bernoulli RFS distribution was established. This provided an alternative derivation of the MeMBer data update process. Then, this alternative derivation was extended to introduce spurious targets that described the cardinality bias. Actually, the spurious targets removed the ambiguity arising from modeling of two contradicting hypothesis for predicted Bernoulli targets: one under the legacy track set and another under the data-induced track set. In contrast to the popular CBMeMBer filter, the modeling of spurious targets removed the bias without making any limiting assumption on sensor probability of detection. This allowed the proposed IMeMBer filter to be employed in scenarios where sensor probability of detection had moderate to small values. At this point, it is important to note that the CBMeMBer filter can be obtained from the IMeMBer filter when sensor probability of detection is so close to unity. The improvement to the proposed IMeMBer filter was achieved by refinement of existing probabilities. Hence, the stability of the cardinality estimate was improved, compared to that in the CBMeMBer filter. In addition to simulations, a theoretical analysis like those presented for the PHD and CPHD filters was performed to demonstrate strength and limitations of the IMeMBer filter. One possible research direction obtained from this analysis would be to remove sparsely-distributed clutter assumption made in the data update process. Thus, like the CPHD filter, the IMeMBer filter would have a satisfactory performance in high cluttered environments.

The another main contribution of this thesis allowed the robust JoM estimator to be used in RFS multi-Bernoulli filters. For this purpose, an optimal solution to the unknown JoM estimation constant was computed by solving a multi-objective optimization problem. For this optimization problem, two convex nonlinear objective functions in conflict were defined: the first objective function aimed to reduce information gain by minimizing the KL divergence of actual spatial probability density function from its uniform approximation. On the other hand, the second objective function aimed to improve the accuracy of estimated states. Since there was no a utopian solution that simultaneously optimized these conflicting objectives, the weighted sum method was used to obtain a unique Pareto optimal solution to the unknown JoM estimation constant. The decision maker's preference on each objective function was determined by adaptive weights. The proposed JoM estimator was applied to the optimal JoTT filter. The simulation results demonstrated its superior track management performance in terms of track confirmation latency and track maintenance against the MaM estimator. However, it suffered from the track termination delay more than the MaM estimator.

Finally, the proposed JoM estimator was reformulated to be used with the multi-Bernoulli distribution. For this purpose, those conflicting objective functions were generalized using local statistics of GM densities. In addition, a new calculation method for adaptive weights were proposed to adjust them according to dynamics of multi-target tracking and characteristics of the JoM estimator. The proposed JoM estimator provided more reliable cardinality estimates than the MaM estimator under moderately low-observable conditions. This was the expected result since the JoM estimator was more robust than the MaM estimator. However, as in the single target tracking case, it suffered from track termination delay more than the MaM estimator. The reason behind this behavior was due to gradually deteriorating localization performance of the multi-target filter after target deaths. Furthermore, the performance comparison under high-observable conditions demonstrated that the performances of these two estimators were same in terms of accuracy and stability of cardinality estimates.

Results obtained from simulations of the proposed JoM estimators in both single and multi-target tracking scenarios indicate that one possible research direction would be to reduce the track termination latency.