

Problems in nonlinear Bayesian filtering

Author:

Pasha, Syed

Publication Date:

2009

DOI:

<https://doi.org/10.26190/unsworks/14783>

License:

<https://creativecommons.org/licenses/by-nc-nd/3.0/au/>

Link to license to see what you are allowed to do with this resource.

Downloaded from <http://hdl.handle.net/1959.4/43792> in <https://unsworks.unsw.edu.au> on 2024-05-01

UNSW



>014179059

PLEASE TYPE

**THE UNIVERSITY OF NEW SOUTH WALES
Thesis/Dissertation Sheet**

Surname or Family name: PASHA

First name: SYED

Other name/s: AHMED

Abbreviation for degree as given in the University calendar: PhD.

School: ELECTRICAL ENGINEERING

Faculty: ENGINEERING

Title: PROBLEMS IN NONLINEAR BAYESIAN FILTERING

Abstract 350 words maximum: (PLEASE TYPE)

This dissertation presents solutions to two open problems in estimation theory. The first is a tractable analytical solution for problems in multi-target filtering which are too complex to solve using traditional techniques. The second explores a new approach to the nonlinear filtering problem for a general class of models.

The approach to the multi-target filtering problem which involves jointly estimating a random process of the number of targets and their state, developed using the probability hypothesis density (PHD) filter alleviates the intractability of the problem by avoiding explicit data association. Moreover, the notion of linear jump Markov systems is generalized to the multiple target case to accommodate births, deaths and switching dynamics to derive a closed form solution to the PHD recursion for this so-called linear Gaussian jump Markov multi-target model. The proposed solution is general enough to accommodate a broad class of practical problems which are deemed intractable using traditional techniques. Based on this closed form solution, an efficient method is developed for tracking multiple maneuvering targets that switch between multiple models without the need for gating, track initiation and termination, or clustering for extracting state estimates.

The approach to the nonlinear filtering problem explores the framework of the virtual linear fractional transformation (LFT) model which localizes the nonlinearity to the feedback with a simple and sparse structure. The LFT is an exact representation for any differentiable nonlinear mapping and therefore amenable to a general class of problems. An alternative analytical approximation method is presented which avoids linearization of the state space model. The uncorrelated structure of the feedback connection gives better second-order moment approximation of the nonlinearly mapped variables. By arranging the unscented transform in the feedback, the prediction and estimation steps are derived in closed form. The proposed filters for the discrete-time model and continuous-time dynamics with sampled-data measurements respectively are shown to be robust under highly nonlinear and uncertain conditions where standard analytical approximation based filters diverge. Moreover, the LFT based filters are efficient for online implementation. In addition, the LFT framework is applied to extend the closed form solution of the PHD recursion to the nonlinear jump Markov multi-target model.

Declaration relating to disposition of project thesis/dissertation

I hereby grant to the University of New South Wales or its agents the right to archive and to make available my thesis or dissertation in whole or in part in the University libraries in all forms of media, now or here after known, subject to the provisions of the Copyright Act 1968. I retain all property rights, such as patent rights. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

I also authorise University Microfilms to use the 350 word abstract of my thesis in Dissertation Abstracts International (this is applicable to doctoral theses only).

.....
Signature

.....
Witness

.....
Date

The University recognises that there may be exceptional circumstances requiring restrictions on copying or conditions on use. Requests for restriction for a period of up to 2 years must be made in writing. Requests for a longer period of restriction may be considered in exceptional circumstances and require the approval of the Dean of Graduate Research.

FOR OFFICE USE ONLY

Date of completion of requirements for Award:

Declaration

I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, nor material which to a substantial extent has been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis.

I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.

Signature _____ Date _____

COPYRIGHT STATEMENT

'I hereby grant the University of New South Wales or its agents the right to archive and to make available my thesis or dissertation in whole or part in the University libraries in all forms of media, now or here after known, subject to the provisions of the Copyright Act 1968. I retain all proprietary rights, such as patent rights. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

I also authorise University Microfilms to use the 350 word abstract of my thesis in Dissertation Abstract International (this is applicable to doctoral theses only).

I have either used no substantial portions of copyright material in my thesis or I have obtained permission to use copyright material; where permission has not been granted I have applied/will apply for a partial restriction of the digital copy of my thesis or dissertation.'

Signed

Date

AUTHENTICITY STATEMENT

'I certify that the Library deposit digital copy is a direct equivalent of the final officially approved version of my thesis. No emendation of content has occurred and if there are any minor variations in formatting, they are the result of the conversion to digital format.'

Signed

Date

Problems in nonlinear Bayesian filtering

by

Syed Ahmed Pasha

B.E., National University of Sciences and Technology in

Pakistan, 2001



A thesis submitted in fulfillment

of the requirements for the degree of

Doctor of Philosophy

School of Electrical Engineering & Telecommunications

University of New South Wales

2009

This thesis entitled:
Problems in nonlinear Bayesian filtering
written by Syed Ahmed Pasha
has been approved for the School of Electrical Engineering &
Telecommunications

Supervisor: A/Prof. Tuan Duong Hoang

The final copy of this thesis has been examined by the signatory, and I find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Abstract

This dissertation presents solutions to two open problems in estimation theory. The first is a tractable analytical solution for problems in multi-target filtering which are too complex to solve using traditional techniques. The second explores a new approach to the nonlinear filtering problem for a general class of models.

The approach to the multi-target filtering problem which involves jointly estimating a random process of the number of targets and their state, developed using the probability hypothesis density (PHD) filter alleviates the intractability of the problem by avoiding explicit data association. Moreover, the notion of linear jump Markov systems is generalized to the multiple target case to accommodate births, deaths and switching dynamics to derive a closed form solution to the PHD recursion for this so-called linear Gaussian jump Markov multi-target model. The proposed solution is general enough to accommodate a broad class of practical problems which are deemed intractable using traditional techniques. Based on this closed form solution, an efficient method is developed for tracking multiple maneuvering targets that switch between multiple models without the need for gating, track initiation and termination, or clustering for extracting state estimates.

The approach to the nonlinear filtering problem explores the framework of the virtual linear fractional transformation (LFT) model which localizes the nonlinearity to the feedback with a simple and sparse structure. The LFT is an exact representation for any differentiable nonlinear mapping and therefore amenable to a general class of problems. An alternative analytical approximation method is presented which avoids linearization

of the state space model. The uncorrelated structure of the feedback connection gives better second-order moment approximation of the nonlinearly mapped variables. By arranging the unscented transform in the feedback, the prediction and estimation steps are derived in closed form. The proposed filters for the discrete-time model and continuous-time dynamics with sampled-data measurements respectively are shown to be robust under highly nonlinear and uncertain conditions where standard analytical approximation based filters diverge. Moreover, the LFT based filters are efficient for online implementation. In addition, the LFT framework is applied to extend the closed form solution of the PHD recursion to the nonlinear jump Markov multi-target model.

Dedication

To my mother

in appreciation of all your devotion and efforts

Acknowledgements

In the name of *Allah*, Most Merciful, All-Merciful. To Him I extend my profound gratitude. It is because of Him alone that I am able to complete this dissertation. I am also obliged to a number of people for their support and encouragement.

First and foremost, I would like to express my appreciation to my supervisor A/Prof. Tuan Duong Hoang for his careful criticism of my work, insightful discussions and financial support. His guidance has been most valuable.

Next, I wish to thank my co-supervisor A/Prof. Ba-Ngu Vo for introducing me to random sets, who has been instrumental in improving my understanding of the subject, Dr. Wing-Kin Ma for helpful comments. I would also like to acknowledge my friends: members of the Systems & Control Laboratory.

Finally, I am grateful to my mother and sisters for their support and to my wife for her love and patience. May *Allah* prolong the lives of these people and reward them in the best possible way.

This work was partially funded by the Higher Education Commission of Pakistan.

Author's Publications

- Journal papers:

[109] Syed Ahmed Pasha, Ba-Ngu Vo, Hoang Duong Tuan and W-Kin Ma, "A Gaussian mixture PHD filter for jump Markov system models," to appear in *IEEE Transactions on Aerospace and Electronic System*, 2009.

[108] Syed Ahmed Pasha, Hoang Duong Tuan, Ba-Ngu Vo, "Nonlinear Bayesian Filtering using the Unscented Linear Fractional Transformation Model," accepted for publication in *IEEE Transactions on Signal Processing*, 2009.

- Conference papers:

[103] A. Pasha, B. Vo, H.D. Tuan and W.-K. Ma, "Closed form PHD filtering for linear jump Markov models," in *Proceedings of the 9th International Conference on Information Fusion*, (Florence, Italy), pp. 1–8, 2006. **Best Paper award in Information Fusion.**

[135] Ba-Ngu Vo, Ahmed Pasha and Hoang Duong Tuan, "A Gaussian mixture PHD filter for nonlinear jump Markov models," in *Proceedings of the 45th IEEE Conference on Decision & Control*, (San Diego, USA), pp. 3162–3167, 2006.

[105] Syed Ahmed Pasha, Hoang Duong Tuan, "Nonlinear filtering for the linear fractional transformation model," in *Proceedings of the 2nd International Conference on Communications and Electronics*, (Hoi An, Vietnam) pp. 180–185, 2008.

- [104] Syed Ahmed Pasha, Hoang Duong Tuan, “Closed form filtering for linear fractional transformation models,” in *Proceedings of the 17th IFAC World Congress*, (Seoul, Korea), pp. 14510–14515, 2008.
- [106] Syed Ahmed Pasha, Hoang Duong Tuan, “Nonlinear filtering for continuous-time systems using the linear fractional transformation model,” to appear in *Proceedings of the ICASSP*, (Taipei, Taiwan), 2009.
- [107] Syed Ahmed Pasha, Hoang Duong Tuan and Pierre Apkarian, “The LFT based PHD filter for nonlinear jump Markov models in multi-target tracking,” to appear in *Proceedings of the 48th IEEE Conference on Decision & Control*, (Shanghai, China), 2009.

Contents

Chapter	
1	Introduction 3
2	Mathematical foundations 10
2.1	MMSE estimation 11
2.2	Kalman filter 14
2.3	Random sets and counting measures 16
2.4	Random set filtering 23
2.5	Nonlinear filtering 26
3	Gaussian mixture PHD filter for jump Markov system models 31
3.1	Introduction 32
3.2	Problem formulation 34
3.2.1	Jump Markov system (JMS) 34
3.2.2	Random finite sets in multi-target tracking 35
3.2.3	The probability hypothesis density filter 36
3.3	Closed form solution to the PHD recursion for LGJMS multi-target model 39
3.3.1	Linear Gaussian jump Markov system multi-target models 39
3.3.2	Closed form PHD recursion for LGJMS multi-target model 44
3.3.3	General solution to the PHD recursion 48
3.3.4	Simulation results 51

3.4	The PHD filter for nonlinear GJM multi-target models	58
3.4.1	Simulation results	62
4	Nonlinear filtering based on LFT modeling	66
4.1	Introduction	67
4.2	LFT for linear filtering of nonlinear models	69
4.2.1	The linear fractional transformation (LFT) model	69
4.2.2	Recursive Bayes estimation	72
4.3	Simulation results	74
4.3.1	Example I	74
4.3.2	Example II	76
4.3.3	Example III	79
4.4	General solution to Bayes recursion	82
4.4.1	Example IV	84
5	LFT based filtering of the continuous-time model	88
5.1	Introduction	89
5.2	Background: Moment propagation	91
5.3	LFT in filtering nonlinear continuous-time stochastic processes	96
5.3.1	Moment calculations by the EKF and UKF	97
5.3.2	Moment calculation by LFT modeling	100
5.3.3	Recursive Bayes filter by LFT	103
5.4	Simulation results	105
5.4.1	Example I	105
5.4.2	Example II	106
5.4.3	Example III	108
5.4.4	Example IV	111
6	Conclusions	116

A Proofs and Definitions	119
A.1 Proof of Lemma 2.6	119
A.2 Functional derivative	120

Appendix

Bibliography	121
---------------------	-----

List of Figures

Figure

3.1	Aircraft and payload trajectories. ‘o’– locations of start of flight; ‘□’– locations of end of flight (‘×’– location of sensor).	55
3.2	Measurement data and true target positions.	56
3.3	Position estimates of the Gaussian mixture PHD filter.	56
3.4	Mean absolute error of estimated number of targets and probability of track loss.	57
3.5	Aircraft and payload trajectories. ‘o’– locations of start of flight; ‘□’– locations of end of flight (‘×’– location of sensor).	57
3.6	Tracking performance and computational complexity versus clutter rate for $p_{D,k} = 0.98$ and CPEP radius = 50 m.	58
3.7	Tracking performance versus probability of detection for $\lambda_c = 3.47 \times 10^{-3} km^{-2}$ and CPEP radius = 50 m.	59
3.8	Aircraft and payload trajectories. ‘o’– locations of start of flight; ‘□’– locations of end of flight (‘×’– location of sensor).	64
3.9	Position estimates of the Gaussian mixture PHD filter.	65
3.10	Mean absolute error of estimated number of targets and probability of track loss.	65
4.1	Trajectory of the state $x_k = (x_k(1), x_k(2))^T$	75

4.2	True trajectory and the estimate of the state $\mathbf{E}(x_k Z_k)$ given by the proposed filter.	76
4.3	Mean square error (MSE) in the estimates using the proposed filter and the UKF.	76
4.4	Trajectory of the state $x_k = (x_k(1), x_k(2))^T$	78
4.5	True trajectory and the estimate of the state $\mathbf{E}(x_k Z_k)$ given by the proposed filter.	78
4.6	MSE in the estimates using the proposed filter	79
4.7	Trajectory of the vehicle. ‘o’- location of vehicle at $k = 1$; ‘□’- location of vehicle at $k = 100$ (‘×’- location of sensor).	81
4.8	Measurement data (projected on the x and y axis) and true target positions.	82
4.9	Position estimates of the Gaussian mixture PHD filter using the LFT model.	82
4.10	Mean absolute error of estimated number of targets using the LFT based JMS-PHD filter and the unscented JMS-PHD filter.	83
4.11	Trajectories of the cart and the pendulum.	87
4.12	True motion parameters of the cart and estimates using the proposed filter.	87
4.13	MSE in the estimates of motion parameters of the cart using the proposed filter.	87
5.1	Trajectory of the state $x(t) = (x_1(t), x_2(t))^T$	107
5.2	True trajectory and the estimate of the state $\mathbf{E}(x(t_k) Z(t_k))$ given by the proposed filter.	107
5.3	Mean square error (MSE) in the estimates using the proposed filter.	108
5.4	Trajectory of the state $x(t) = (x_1(t), x_2(t))^T$	109
5.5	True trajectory and the estimate of the state $\mathbf{E}(x(t_k) Z(t_k))$ given by the proposed filter.	109
5.6	MSE in the estimates using the proposed filter	110

5.7	Trajectory of the vehicle. ‘o’– location of vehicle at $k = 1$; ‘□’– location of vehicle at $k = 100$ (‘x’– location of sensor).	110
5.8	Measurement data (projected on the x and y axis) and true target positions.	111
5.9	Position estimates of the Gaussian mixture PHD filter using the LFT model.	111
5.10	Mean absolute error of estimated number of targets using the LFT model and the UKF.	112
5.11	Trajectories of the oscillator and actuator.	113
5.12	True oscillator motion parameters and estimates using the proposed filter.	114
5.13	True actuator motion parameters and estimates using the proposed filter.	114
5.14	MSE in the estimates of oscillator motion parameters using the proposed filter and the UKF.	114
5.15	MSE in the estimates of actuator motion parameters using the proposed filter and the UKF.	115

Acronyms and abbreviations

CR-JPDA	Column recursive joint probabilistic data association
EKF	Extended Kalman filter
FISST	Finite set statistics
GJM	Gaussian jump Markov
IMMJPDA	Interacting multiple model joint probabilistic data association
JMS	Jump Markov system
JPDA	Joint probabilistic data association
LCHS	Locally compact, Hausdorff and separable
LFT	Linear fractional transformation
LG	Linear Gaussian
LGJMS	Linear Gaussian jump Markov system
LMMSE	Linear minimum mean square error
LRKF	Linear regression Kalman filter
MHT	Multiple hypotheses tracking
MMSE	Minimum mean square error
MSE	Mean square error
NFT	Nonlinear fractional transformation
NN	Nearest neighbor
p.g.fl.	Probability generating functional
PHD	Probability hypothesis density
RFS	Random finite set

RTAC	Rotational-translational actuator
SMC	Sequential Monte Carlo
UKF	Unscented Kalman filter

Chapter 1

Introduction

The objective of this thesis can be summarized into two main goals. The first is the development of a tractable solution for problems in multi-target filtering where traditional techniques are deemed intractable. The second is the development of a new approach for solution to the nonlinear filtering problem for a general class of nonlinear models. The viability of the results is particularly confirmed by comparison with recognized benchmark techniques.

Multi-target filtering is one of the most critical functions in many civilian and military applications. Some of these applications include air traffic control, financial econometrics, global positioning systems, air defense systems, ocean surveillance systems and ballistic missile defense. A typical scenario of a surveillance system includes time-varying number of targets as new targets appear in the surveillance region while existing targets disappear in the presence of false alarms, noise and uncertainties in the target dynamics, data association and detection. In order to provide accurate and reliable estimates of the state of the objects in the surveillance space, a reliable modeling of interconnected components such as target births, deaths, switching target dynamics, detection uncertainty and clutter is required. In addition, the model must be computationally efficient for real-time implementation. As such, the problem is extremely challenging in both theory and implementation and involves jointly estimating the number of targets and their state at each time step. Traditional multi-target filtering techniques

are computationally intractable under such a general setting.

A target may not follow a predefined trajectory and the uncertainty in the dynamics is modeled by additive noise to compensate for the modeling inaccuracy. If the variance of the noise is small and the target maneuvers sharp turns, the model suffers the loss of track of the target. On the other hand, if the variance of the noise is large, the model is able to track the target but the accuracy of the estimate is degraded. A target that can maneuver exhibits different kinematic behavior from time to time. In such a case, a single model is insufficient to describe completely the behavior of the target at all times. The estimates based on the single model often lead to either poor accuracy of the state estimate or loss of track. The jump Markov system (JMS) or multiple models approach has proven to be an effective tool for single maneuvering target tracking [9, 7, 80, 18, 81]. In this approach, the target can switch between a set of models in a Markovian fashion.

False alarms produced by clutter due to interference between signals and multiple path returns detected by the sensor introduce uncertainty in the origin of the measurement. It is not known whether a measurement is produced by a target or spurious detection. Moreover, noise, occlusion and sensor resolution in case of targets in close proximity may cause the sensor to miss a target. Since the standard filtering theory assumes that the origin of the measurement is known in order to update the target state, traditional tracking approaches solve the data association problem in order to apply standard Bayesian filtering techniques. An error can be made in associating spurious observations to a target. Therefore, the performance and reliability of the traditional methods are highly dependent on the data association. A Bayesian procedure in the tracking problem was introduced in [121] to solve the data association problem which combined with the work [118] led to the development of the nearest neighbor (NN) method [120], applicable in low cluttered environments. The probabilistic data association (PDA) method [8, 11] by weighted averaging the observations gives better

performance than the NN method in environments where clutter rate is high [43].

The multi-target filtering problem is further compounded by the additional uncertainty in data association introduced due to the presence of multiple targets. It is not known which target in the state space produced which measurement in the observation space. The extended probabilistic data association algorithm [5] and its revised version which performs better under track crossing known as the joint probabilistic data association (JPDA) filter [41, 42] were developed to solve the data association problem in the multi-target setting. However, these methods require that the number of targets be known and fixed. A more generalized algorithm known as the multiple hypotheses tracking (MHT) [113, 114] filter forms a hypothesis about the origin of the measurements and retains likely hypotheses based on a pruning criterion for tracking an unknown and time-varying number of targets. JPDA and MHT do not handle maneuvering targets. The JMS approach can be combined with these traditional data association algorithms to track multiple maneuvering targets [6, 19, 130, 29, 60, 38, 73, 111]. However, these data association-based approaches are computationally intensive in general and heuristic techniques are used to reduce the computational load.

Early attempts at tracking an unknown fixed number of targets using point process formulation were made in [140, 99]. Although various approaches have been proposed for simultaneous estimation of the time-varying number of targets and their state in [95, 94, 123, 15, 100], the first systematic treatment of Bayesian multi-target filtering based on random finite sets (RFS) [93] is given in [92, 45, 91, 88]. The RFS approach treats the finite sets of targets and observations at each time step as the multi-target state and multi-target observation respectively and is an elegant generalization of the single target Bayes filter. The works cited above formulate the multi-target filtering problem using finite set statistics (FISST) [88] and fall short of providing a rigorous treatment of the FISST Bayes recursion which involves set derivatives of belief mass functions instead of probability densities. A measure-theoretic formalism of the problem

was provided in [138, 137] which also established the validity of the FISST Bayes recursion by showing the relationship between FISST and conventional probability theory. For a comparison of the RFS approach and traditional multi-target tracking methods see [45].

The Probability Hypothesis Density (PHD) filter [92, 90, 89, 124] which propagates the first moment of the multi-target posterior only called the intensity function circumvents the combinatorial computations that arise from data association while accommodating detection uncertainty, Poisson false alarms, target motion and time-varying number of targets. In [132], the PHD filter was applied to track multiple maneuvering targets using sequential Monte Carlo (SMC) implementations [138, 137]. However, the main drawbacks of the SMC approach are the large number of particles and the unreliability of clustering for extracting multi-target state estimates [138, 131, 133]. Recently, a closed form solution to the PHD recursion has been found for linear Gaussian models that led to the development of the Gaussian mixture PHD filter [133, 131]. Although this approach is efficient and capable of handling nonlinear models, it is not general enough for addressing targets with JMS dynamics. At present there is no tractable analytical technique for tracking multiple targets with JMS dynamics.

Many real world problems do not follow linear models. A JMS comprising of nonlinear models accommodates an even wider range of applications by providing a greater generality for modeling systems that switch between various models. However, for the nonlinear filtering problem there exists no analytical expression for the optimal Bayes solution and in the single target environment a significant challenge is to find an efficient method for on-line, real-time estimation of the state given a nonlinear mapping of the state. Consequently, approximate nonlinear filters have been proposed. These methods are based on either analytical approximations [59, 44, 62, 63, 116, 57, 101], numerical approximations [14, 85, 21, 37, 117, 12, 27, 69, 68] or simulation based approaches [47, 46, 30, 40, 39, 87, 26]. The most widely used analytical approximation method is

the extended Kalman filter (EKF) [59, 44]. The EKF applies a local linearization to the nonlinear mapping around the state estimate. This suggests that the region of stability may be small, significant bias and convergence problems are commonly encountered due to the crude approximation. The unscented Kalman filter (UKF) [62, 63] applies the unscented transform [62] which uses the statistical linear regression technique [77, 78] to approximate the moments of random variables. Similar to the UKF, the divided difference filter [116, 57, 101] also adopts a derivative-free linearization method called the central difference approximation for functional evaluation. The conditional expectation evaluated using the UKF has a higher order accuracy than the estimate given by the EKF. This has been substantiated by empirical studies on the EKF and the UKF showing that in most applications the UKF gives better approximation [62, 63, 77, 33, 115]. Despite the advantage of the UKF over the EKF, the two approaches work reasonably well under mildly nonlinear conditions only.

Over the past few years SMC methods have attracted attention for nonlinear Bayesian filtering applications [40, 39]. These methods approximate the filtering distribution by a set of samples drawn from a proposal distribution. Under the assumption that the proposal distribution includes the region of support of the filtering distribution, SMC methods give more accurate estimates than the analytical approximation based methods. In practice, a sufficiently large number of samples is needed. It is only in the limit that the number of samples approaches infinity that the simulation-based methods guarantee convergence of the estimate to the optimal Bayes solution. There have been many recent modifications and improvements on the SMC methods [39]. However, some of the problems related to the choice of proposal density, optimal sampling from the distribution and computational complexity still persist.

In nonlinear control, two transformation methods are well-known. The exact feedback linearization transforms a nonlinear control system into an equivalent linear one through a variable change [55, 56, 70]. However, it is applicable to a limited class

of nonlinear systems. On the other hand, the linear fractional transformation (LFT) method (see e.g., [143, 4, 3] and the references therein) exists for a broad class of nonlinear systems and is extensively employed in \mathcal{H}_2 and \mathcal{H}_∞ gain-scheduling based control and filtering to represent nonlinear plants, where the uncertainty appears as a LFT (see e.g., [129, 128, 127, 25] and the references therein). The LFT is attractive in that it localizes the nonlinearity to the feedback with a structure that is both simple and sparse. The approximation in the LFT is therefore sufficiently localized to the feedback to linearize a simple nonlinear structure.

This dissertation presents a generalization of the Bayesian multi-target filtering framework using the RFS approach to facilitate tracking of multiple maneuvering targets in the presence of noise, clutter and uncertainties in target dynamics, detection and data association. The proposed multi-target model accommodates target spawning, of interest in military applications to detect missiles spawned from a target for early detection and interception. The nonlinearity of the radar measurement model is addressed by the second main theme of this thesis in developing a nonlinear filtering technique that can accommodate a general class of nonlinear models to provide a perspective that has not been investigated previously in the literature. This research impacts both military and civilian applications. Specifically, the original contributions of the thesis are the following.

- The notion of linear jump Markov systems is generalized to the multiple target case to accommodate births, deaths and switching dynamics. A closed form solution to the PHD recursion is derived for this so-called linear Gaussian jump Markov multi-target model. Based on this closed form solution, an efficient method is developed for tracking multiple maneuvering targets that switch between multiple models. In addition, the proposed approach is extended to nonlinear jump Markov multi-target models by combining the closed form solution and the unscented transform. Further details can be found in Chapter 3.

- A nonlinear Bayesian filtering technique is developed for the discrete-time model based on LFT modeling. A closed form solution to Bayes recursion is derived by arranging the unscented transform in the feedback. In addition, the proposed filtering approach is generalized to handle any smooth nonlinear mapping using the nonlinear fractional transformation (NFT) model. Further details can be found in Chapter 4.
- A nonlinear Bayesian filtering technique is developed for the continuous-time dynamical model based on the efficient LFT modeling approach for simple approximation of the stochastic differential equation of the state prediction and accurate estimation of the state conditional on observations. Further details can be found in Chapter 5.

This thesis is divided into six chapters. The first chapter gives an introduction. The second chapter is devoted to the mathematical foundations. The third presents the linear Gaussian jump Markov multi-target model which is more general than those in standard multi-target tracking algorithms. While traditional multi-target filtering techniques are computationally intractable for a model of such generality it is shown that using the RFS approach, this model is amenable to computationally efficient multi-target filtering techniques. The fourth chapter contains a discussion on the solution of the nonlinear filtering problem using LFT modeling for the discrete-time system. This discussion is continued in the fifth chapter for the continuous-time dynamical model with sampled-data measurements. The conclusion of this thesis in the sixth chapter highlights the importance of developing effective multi-target filtering and nonlinear filtering techniques and shows that the proposed methods provide sound developments in estimation theory.

Chapter 2

Mathematical foundations

Bayesian estimation of the state at time k of a given dynamical stochastic process based on the available information up to time k is known as *filtering*. The objective is to find an estimate that is optimal for a given criterion. Such problems have been dealt with in great detail in the Kalman filter [65] framework for second-order stationary processes and the minimum mean square error (MMSE) criterion [59, 44, 1]. The advantage of the approach can be found in the recursive form of the solution with favorable practical implication in terms of real-time implementation. The estimation problem for random sets is also of theoretical and practical importance. An application that has attracted considerable attention is the target tracking problem for multiple targets where the state and observation are finite sets.

This chapter briefly covers mathematical concepts that will facilitate discussions presented in this thesis. Section 2.1 reviews optimal estimation of a function of a random variable for the MMSE criterion. Section 2.2 presents the Kalman filter as the optimal Bayes filter for the linear state space model. Background on random finite sets (RFS) is presented in Section 2.3. Using the point process interpretation of RFS, a discussion on RFS filtering is given in Section 2.4. Analytical approximation methods in nonlinear filtering are discussed in Section 2.5.

The notation adopted is as follows: $X|Y$ denotes a random variable X restricted by a realization of the conditioning random variable Y , $\mathbf{E}_X(\cdot)$ denotes the expectation

with respect to random variable X , $\langle \cdot, \cdot \rangle$ is the dot product and $\text{tr}(\cdot)$ denotes the trace operator. $X \sim \mathcal{N}(X; \bar{x}, R_X)$ denotes a normally distributed random variable with mean \bar{x} and covariance R_X , while $\mathcal{N}(\cdot; \bar{x}, R_X)$ is its probability density function. R_X^\dagger denotes the pseudo-inverse of R_X .

2.1 MMSE estimation

Consider the state space model given by

$$x_{k+1} = f(x_k) + B_k w_k, \quad (2.1)$$

$$z_k = g(x_k) + D_k v_k, \quad (2.2)$$

where f and g denote arbitrary nonlinear mappings. Here, $x_k \in \mathbb{R}^n$ and $z_k \in \mathbb{R}^m$ are the system state and measurement, $w_k \sim \mathcal{N}(w_k; 0, Q_k)$ and $v_k \sim \mathcal{N}(v_k; 0, R_k)$ are mutually uncorrelated process noise and measurement noise which are also statistically independent of x_k . $B_k \in \mathbb{R}^{n \times p}$ and $D_k \in \mathbb{R}^{m \times q}$ denote noise gain matrices.

The filtering problem involves estimating the state x_k at time k which evolves in a Markovian fashion in the presence of uncertainty w_k conditional on the history of observations $Z_k = (z_1, \dots, z_k)$ corrupted by noise process $\{v_1, \dots, v_k\}$. As the present state x_k also plays an intermediate role as the information carrier for all the past observations (z_1, \dots, z_{k-1}) , filtering is two step process: using observation equation (2.2) to estimate the current state x_k from the current observation z_k and then using state equation (2.1) to predict the future state x_{k+1} . These steps in essence are particular cases of the following estimation problem.

Let X and Y be two random variables with expected values $\mathbf{E}_X(X) = \bar{x}$ and $\mathbf{E}_Y(Y) = \bar{y}$ respectively and with covariance $\text{cov}(Y, X) = R_{YX}$ and $\mathbf{E}_{YX}(\langle Y, X \rangle) = \text{tr}(\text{cov}(Y, X) + \mathbf{E}_Y(Y)\mathbf{E}_X(X)^T)$. The central problem of linear estimation is how to estimate X by an affine function $AY + b$ with deterministic matrix A and vector b ?

Using the minimum mean square error (MMSE) criterion, A and b are found from

$$\min_{A,b} \mathbf{E}_{XY}(\|X - (AY + b)\|^2). \quad (2.3)$$

If X and Y are zero mean i.e., $\bar{x} = 0$ and $\bar{y} = 0$, it follows that $b = 0$. Moreover, if X and Y are also uncorrelated, then $\mathbf{E}_{XY}(\|X - AY\|^2) = \mathbf{tr}(\mathbf{cov}(X, X) + A\mathbf{cov}(Y, Y)A^T)$ is attained minimum at $A = 0$, i.e., $0 = 0 \cdot Y$ is the optimal linear estimator of X conditional on Y . An important result pertinent to estimation and filtering is stated as a theorem below.

Theorem 2.1. *Let X and Y be two random variables with expected values \bar{x} and \bar{y} , auto-covariances R_X and R_Y respectively and cross-covariance R_{YX} . Then*

$$(R_{YX}^T R_Y^\dagger, \bar{x} - R_{YX}^T R_Y^\dagger \bar{y}) = \arg \min_{A,b} \mathbf{E}_{XY}(\|X - (AY + b)\|^2).$$

Consequently, the linear estimator of X based on the observation $Y = y$ for any random variables X and Y is

$$\bar{x} + R_{YX}^T R_Y^\dagger (y - \bar{y}).$$

Moreover, the exact linear statistical relation of X and Y is

$$X = R_{YX}^T R_Y^\dagger (Y - \bar{y}) + \bar{x} + e, \quad (2.4)$$

where the random error

$$e = X - R_{YX}^T R_Y^\dagger (Y - \bar{y}) - \bar{x} \quad (2.5)$$

is uncorrelated to Y ,

$$\mathbf{E}_{XY} \left((X - R_{YX}^T R_Y^\dagger (Y - \bar{y}) - \bar{x}) Y^T \right) = 0, \quad (2.6)$$

and

$$\mathbf{E}_{XY}(\|X - R_{YX}^T R_Y^\dagger (Y - \bar{y}) - \bar{x}\|^2) = \mathbf{tr}(R_X - R_{YX}^T R_Y^\dagger R_{YX}).$$

Proof. The proof is based on well known results [112, 1].

Let $[X^T, Y^T]^T$ be the augmented vector formed by concatenating X and Y and $R_{X,Y}$

be the covariance matrix of the augmented vector, then for

$$R_{X,Y} = \begin{bmatrix} R_X & R_{XY} \\ R_{XY}^T & R_Y \end{bmatrix} \geq 0,$$

any vector orthogonal to columns of R_Y (R_X) must also be orthogonal to the columns of R_{YX} (R_{YX}^T). Let $\mathcal{M}(X)$ denote the linear space spanned by columns of X , then $\mathcal{M}(R_{YX}) \subset \mathcal{M}(R_Y)$, $\mathcal{M}(R_{YX}^T) \subset \mathcal{M}(R_X)$ and there is a matrix B such that $R_{YX}^T = BR_Y$. Using the property $R_Y R_Y^\dagger R_Y = R_Y$,

$$R_{YX}^T R_Y^\dagger R_Y = BR_Y = R_{YX}^T,$$

and hence

$$\begin{aligned} \mathbf{E}^1((X - \bar{x} - R_{YX}^T R_Y^\dagger (Y - \bar{y}))Y^T) &= \mathbf{E}((X - \bar{x})Y^T) - R_{YX}^T R_Y^\dagger \mathbf{E}((Y - \bar{y})Y^T) \\ &= R_{YX}^T - R_{YX}^T R_Y^\dagger R_Y = 0, \end{aligned}$$

i.e., $X - \bar{x} - R_{YX}^T R_Y^\dagger (Y - \bar{y})$ and Y are uncorrelated which implies that 0 is the optimal linear estimator of $X - \bar{x} - R_{YX}^T R_Y^\dagger (Y - \bar{y})$ conditional on Y or equivalently $\bar{x} + R_{YX}^T R_Y^\dagger (y - \bar{y})$ is the optimal linear estimator of X conditional on $Y = y$. Furthermore,

$$\begin{aligned} \mathbf{E}(\|X - \bar{x} - R_{YX}^T R_Y^\dagger (Y - \bar{y})\|^2) &= \text{tr}(\mathbf{E}((X - \bar{x} - R_{YX}^T R_Y^\dagger (Y - \bar{y}))X^T)) \quad (2.7) \\ &= \text{tr}(R_x - R_{YX}^T R_Y^\dagger R_{YX}). \end{aligned}$$

Alternatively, like [1] using Schur's complement

$$R_{X,Y} = \begin{bmatrix} I & R_{YX}^T R_Y^\dagger \\ 0 & I \end{bmatrix} \begin{bmatrix} R_X - R_{YX}^T R_Y^\dagger R_{YX} & 0 \\ 0 & R_Y \end{bmatrix} \begin{bmatrix} I & R_{YX}^T R_Y^\dagger \\ 0 & I \end{bmatrix}^T$$

The covariance of

$$\begin{aligned} \begin{bmatrix} I & R_{YX}^T R_Y^\dagger \\ 0 & I \end{bmatrix}^{-1} \begin{bmatrix} X - \bar{x} \\ Y - \bar{y} \end{bmatrix} &= \begin{bmatrix} I & -R_{YX}^T R_Y^\dagger \\ 0 & I \end{bmatrix} \begin{bmatrix} X - \bar{x} \\ Y - \bar{y} \end{bmatrix} \\ &= \begin{bmatrix} X - \bar{x} - R_{YX}^T R_Y^\dagger (Y - \bar{y}) \\ Y - \bar{y} \end{bmatrix} \end{aligned}$$

¹ The expectation is taken for all X and Y . In the following, the subscript of the expectation operator indicating the variables with respect to which expectation is taken is dropped. There should be no confusion.

is given as

$$\begin{bmatrix} R_X - R_{YX}^T R_Y^\dagger R_{YX} & 0 \\ 0 & R_Y \end{bmatrix}$$

which implies that $X - \bar{x} - R_{YX}^T R_Y^\dagger (Y - \bar{y})$ and $Y - \bar{y}$ are uncorrelated. \square

Theorem 2.1 makes no assumption on the distribution of the random variables and holds true for any X and Y . If X and Y are Gaussian random variables then the random error e defined by (2.5) is Gaussian too, so (2.6) means that e and Y are independent. The linear estimator of X conditional on Y thus coincides with the general minimum mean square error (MMSE) estimator

$$\arg \min_{\tilde{x}} \mathbf{E}(\|X - \tilde{x}\|^2 | Y = y),$$

i.e. it is sufficient to consider linear estimators for Gaussian random variables as non-linear estimators cannot perform better in term of variance of the error. The Kalman filter fully explores this fact as it is concerned with the linear mappings f and g in (2.1)-(2.2). When either X or Y is non-Gaussian, Theorem 2.1 still provides the optimal linear estimator for X conditional on Y .

2.2 Kalman filter

Consider the following linear state space model

$$x_{k+1} = A_k x_k + B_k w_k, \quad (2.8)$$

$$z_k = C_k x_k + D_k v_k, \quad (2.9)$$

where $A_k \in \mathbb{R}^{n \times n}$ and $C_k \in \mathbb{R}^{m \times n}$. When x_k , w_k and v_k are Gaussian, z_k and x_{k+1} must also be Gaussian and thus only Gaussian random variables are concerned in (2.8)-(2.9).

Suppose at the initial time $k = 0$, the estimate of the random variable x_0 is $\bar{x}_0 = m_{0|-1}$ and the covariance is $R_{x,0} = P_{0|-1}$. By (2.9) the random variable z_0

has mean $\eta_0 = C_0 m_{0|-1}$ and covariance $R_{z,0} = C_0 P_{0|-1} C_0^T + D_0 R_0 D_0^T$ with the cross-covariance of z_0 and x_0 given by $R_{zx,0} = C_0 P_{0|-1}$. On arrival of data z_0 , by Theorem 2.1 the conditional mean $\mathbf{E}(x_0|Z_0)$ of the state x_0 given $Z_0 = z_0$ is

$$m_0 = m_{0|-1} + K_0(z_0 - \eta_0),$$

and the covariance of the conditional random variable $\tilde{x}_0 = x_0|Z_0$ is $P_0 = P_{0|-1} - K_0 C_0 P_{0|-1}$, where $K_0 = R_{zx,0}^T R_{z,0}^{-1} = P_{0|-1} C_0^T (C_0 P_{0|-1} C_0^T + D_0 R_0 D_0^T)^{-1}$.

By (2.8), the predicted state $\hat{x}_1 = x_1|Z_0$ at the next time step conditional on data z_0 has expectation $m_{1|0} = \mathbf{E}(x_1|Z_0) = A_0 \mathbf{E}(x_0|Z_0) = A_0 m_0$ and covariance $P_{1|0} = A_0 P_0 A_0^T + B_0 Q_0 B_0^T$. These results are stated more formally as prediction and estimation theorems.

Theorem 2.2. *Suppose the estimate of the state x_{k-1} at time $k-1$ given the history of observations Z_{k-1} is m_{k-1} and the covariance is P_{k-1} . Then, the predicted state $\hat{x}_k = x_k|Z_{k-1}$ at time k conditional on the history of observations up to time $k-1$ has the conditional expectation $m_{k|k-1}$ and the covariance of the prediction is $P_{k|k-1}$ where*

$$m_{k|k-1} = A_{k-1} m_{k-1}, \quad (2.10)$$

$$P_{k|k-1} = A_{k-1} P_{k-1} A_{k-1}^T + B_{k-1} Q_{k-1} B_{k-1}^T. \quad (2.11)$$

Proof. Given the estimate of \tilde{x}_{k-1} as m_{k-1} and covariance P_{k-1} , applying the expectation operator in (2.8) and using $\mathbf{E}(w_{k-1}) = 0$ gives the expression for the conditional mean in (2.10). The variance of the error in prediction in (2.11) is determined by obtaining the expression for the covariance of x_{k-1} in (2.8) and using uncorrelation of x_{k-1} and w_{k-1} . \square

Theorem 2.3. *Suppose the predicted state $\hat{x}_k = x_k|Z_{k-1}$ has mean $m_{k|k-1}$ and is distributed with covariance $P_{k|k-1}$. Then, the conditional expectation of \hat{x}_k also conditional on the data z_k at time k is*

$$m_k = m_{k|k-1} + K_k(z_k - \eta_k), \quad (2.12)$$

and the covariance of $\tilde{x}_k = x_k|Z_k$ is

$$P_k = P_{k|k-1} - K_k C_k P_{k|k-1}, \quad (2.13)$$

with

$$\eta_k = C_k m_{k|k-1}, \quad (2.14)$$

$$K_k = P_{k|k-1} C_{1,k}^T (C_k P_{k|k-1} C_k^T + D_k R_k D_k^T)^{-1}. \quad (2.15)$$

Proof. Given the prediction $m_{k|k-1}$ of \hat{x}_k and the covariance of the prediction $P_{k|k-1}$, the conditional mean of the state \tilde{x}_k follows from Theorem 2.1 after substituting expressions for the expectation and covariance of z_k given by η_k in (2.14) and the expression within the parenthesis in (2.15) respectively and the cross-covariance of z_k and $x_{k|k-1}$ given by $\mathbf{E}(z_k x_{k|k-1}^T) - \eta_k m_{k|k-1}^T$ using the uncorrelation of $x_{k|k-1}$ and v_k . \square

2.3 Random sets and counting measures

Random sets can be defined as set-valued random elements in a certain topological space \mathbb{E} . It is assumed that \mathbb{E} is a locally compact Hausdorff second countable (LCHS) space (Euclidean space \mathbb{R}^d is an example of space \mathbb{E}) [97, 67]. Random sets that belong to the collection of closed subsets of \mathbb{E} are called random closed sets. If these sets contain only a finite number of elements then they are known as random finite sets (RFS). The notation $\mathcal{F}_{\mathbb{E}} = \mathcal{F}(\mathbb{E})$ is adopted to denote the space of RFS in \mathbb{E} .

Let the probability space be defined by the triple $(\Omega, \mathcal{A}, \mathbf{P})$, then the distribution of a RFS $X \in \mathcal{A}$ is determined by $\mathbf{P}_X(\mathcal{X}) = \mathbf{P}(X \in \mathcal{X})$ for all $\mathcal{X} \in \mathcal{B}(\mathcal{A})$ where $\mathcal{B}_{(\cdot)} = \mathcal{B}(\cdot)$ is the Borel σ -algebra. While conditional distributions of RFS can also be determined analogously to random elements in a measurable space, this is not true for expectation which alludes to a linear structure on the space of RFS which is nonlinear. The conventional concepts of expectations in linear spaces can however be applied if the RFS is represented by an equivalent random counting measure [93, 34, 67, 125].

Definition 2.4. A measure μ on $\mathcal{B}_{\mathbb{E}}$ in \mathbb{E} is called counting if it takes only non-negative integer values. μ is locally finite if it is finite on bounded subsets of \mathbb{E} .

Let \mathcal{N} be the family of all locally finite measures μ endowed with a σ -algebra generated by $\{\mu \in \mathcal{N} : \mu(B) = n\}$ for $n = 0, 1, \dots$, and $B \in \mathcal{B}_{\mathcal{N}}$, then a random counting measure can be defined as a random element N in \mathcal{N} . For RFS $X \in \mathcal{F}_{\mathbb{E}}$, $\mu(X)$ is a random variable representing the number of points in X and if X is the union of disjoint sets X_1, X_2, \dots , then $\mu(X) = \sum_i \mu(X_i)$. A random counting measure is also called a point process [93, 34, 20, 67, 125]. A counting measure (or corresponding point process) is called simple if $N(\{x\}) = 0$ or 1 for all $x \in \mathcal{X}$. The distribution of a simple point process N is uniquely determined by

$$\begin{aligned} \mathbf{P}_N(\mathcal{X}) &= \mathbf{P}(N \in \mathcal{X}) \\ &= \mathbf{P}(\{\omega \in \Omega : N(\omega) \in \mathcal{X}\}) \quad \text{for } \mathcal{X} \in \mathcal{B}_{\mathcal{N}}. \end{aligned} \quad (2.16)$$

The notation $\mathcal{N}_{\mathbb{E}}^s = \mathcal{N}_s(\mathbb{E})$ is adopted to denote the family of all simple counting measures $N \in \mathcal{N}$ and $\mathcal{B}_{\mathcal{N}}^s = \mathcal{B}(\mathcal{N}^s)$ the σ -field of its Borel sets.

In the univariate and finite multivariate cases the moments (particularly mean and covariance) provide means to describe various distributions. In the point process context it is the moment measures that describe the distributions. While random variables have moments that are real numbers, the moments of point processes are measures. The expectation measure or intensity measure Λ of N is a characteristic analogous to the mean of a real-valued random variable defined by

$$\begin{aligned} \Lambda(B) &= \mathbf{E}[N(B)] \\ &= \int \omega(B) \mathbf{P}(d\omega), \end{aligned}$$

for Borel set B . It is clear from Fubini's theorem that $\Lambda(\cdot)$ inherits countable additivity from $N(\cdot)$ so it defines a measure on $\mathcal{B}_{\mathcal{N}}$ [34]. $\Lambda(B)$ determines the mean number of points in B . The expectation measure is identical to the first factorial moment measure of N . An important example of a point process is the Poisson point process which is

completely characterized by the intensity measure Λ [71, 34, 125, 67, 32, 122]. The Poisson point process is defined as follows.

Definition 2.5. Let Λ be a diffuse Radon measure on \mathbb{R}^d . The Poisson point process X with intensity measure Λ is a random subset of \mathbb{R}^d such that the following properties are satisfied.

1. For each bounded Borel set B , the cardinality $|X \cap B|$ has a Poisson distribution with mean $\Lambda(B)$

$$\mathbf{P}(X(B) = n) = \frac{(\Lambda(B))^n}{n!} e^{-\Lambda(B)} \quad \text{for } n = 0, 1, \dots \quad (2.17)$$

2. The number of points of X in each of disjoint sets B_1, \dots, B_n are independent for every $n \geq 2$ and any collection of disjoint Borel sets.

If Λ is absolutely continuous with respect to the Lebesgue measure, then the corresponding Raydon-Nikodym derivative (or density) λ is called the intensity function [34, 125, 67] given by:

$$\Lambda(B) = \int_B \lambda(x) dx, \quad (2.18)$$

for Borel sets B .

Let \mathcal{M} be a measurable space. A point process in the product space $\mathbb{E} \times \mathcal{M}$ is called a marked point process. The second component being the mark and the first component called the location. Let \mathcal{M} be the space of the marks and $\mathcal{B}_{\mathcal{M}}$ the σ -field of its Borel sets, then for Borel set B and M in $\mathcal{B}_{\mathcal{M}}$ the number of points of X in B with marks in M is denoted by $X(B \times M)$. Let Λ_g denote the intensity measure of the ground point process (i.e., X stripped of the marks) the measure $\Lambda(\cdot \times M)$ for fixed M in $\mathcal{B}_{\mathcal{M}}$ is absolutely continuous with respect to Λ_g and it can be shown that [125].

$$\Lambda(d(x, m)) = \mathbf{P}_x(dm) \Lambda_g(dx), \quad (2.19)$$

where \mathbf{P}_x is a probability measure on $\mathcal{B}_{\mathcal{M}}$ or the distribution of mark m at a point x for $[x; m] \in X$.

Just as factorial moments in the univariate and finite multivariate cases are related to the Taylor series expansion of the probability generating function about unity, the factorial moment measures are related to the expansion of the probability generating functionals (p.g.fl.) about unity [34, 125, 32]. For a point process $X \in \mathcal{N}_{\mathcal{X}}^s$ the p.g.fl. is defined as

$$G_X[\zeta] = \mathbf{E} \left[\prod_{i=1}^N \zeta(x_i) \right], \quad (2.20)$$

for any Borel measurable function ζ satisfying the condition $|\zeta(x)| \leq 1$. The product is zero if $N > 0$ and $\zeta(x_i) = 0$ for any i and is unity if $N = 0$. Let

$$\zeta(x) = \sum_i^m z_i \mathbf{1}_{B_i}(x), \quad (2.21)$$

where $|z_i| < 1$ and $\mathbf{1}_{B_i}(x)$ is the indicator function of set B_i for measurable partition (B_1, \dots, B_m) of a LCHS space \mathcal{X} , then

$$G_X[\zeta] = \mathbf{E} \left[\prod_{i=1}^m z_i^{N(B_i)} \right], \quad (2.22)$$

is the multivariate probability generating function of the number of points in the sets of the partition and the intensity measure of X is given by

$$\int \xi(x) \Lambda_X(dx) = \left(\frac{d}{d\zeta} G_X[\cdot] \right) \Big|_{\zeta=1} [\xi], \quad (2.23)$$

The p.g.fl. of a Poisson point process X with intensity measure Λ_X is

$$G_X[\zeta] = \exp \left(\int (\zeta(x) - 1) \Lambda_X(dx) \right). \quad (2.24)$$

Proof. The proof is trivial and follows in a straightforward manner (see [32, pp. 39], [125, pp. 116]). \square

A useful property of the p.g.fl. is that the p.g.fl. of the union of two independent finite point processes X_1 and X_2 is the product of those corresponding to the individual processes. For point process $X = X_1 \cup X_2$

$$G_X = G_{X_1} G_{X_2}. \quad (2.25)$$

Let $\Pi_{\mathbf{x}}[g] = \prod_{x \in \mathbf{x}} g(x)$. Given \mathbf{x} and \mathbf{y} as the realizations of two point processes $X \in \mathcal{N}_{\mathcal{X}}^s$ and $Y \in \mathcal{N}_{\mathcal{Y}}^s$ on LCHS spaces \mathcal{X} and \mathcal{Y} respectively, the joint p.g.fl. of X and Y can be defined by [34, 136]

$$\begin{aligned} G_{XY}[g, h] &= \mathbf{E} [\Pi_{\mathbf{x}}[g] \Pi_{\mathbf{y}}[h]] \\ &= \int \Pi_{\mathbf{x}}[g] \Pi_{\mathbf{y}}[h] \mathbf{P}_{XY}(d\mathbf{x}, d\mathbf{y}) \\ &= \int \int \Pi_{\mathbf{x}}[g] \Pi_{\mathbf{y}}[h] \mathbf{P}_X(d\mathbf{x}) \mathbf{P}_{Y|X}(d\mathbf{y}|\mathbf{x}) \end{aligned} \quad (2.26)$$

for any g and h satisfying the condition for (2.20). The conditional p.g.fl. of $Y|X$ is defined by

$$\begin{aligned} G_{Y|X}[h] &= \mathbf{E} [\Pi_{\mathbf{y}}[h]] \\ &= \int \Pi_{\mathbf{y}}[h] \mathbf{P}_{Y|X}(d\mathbf{y}|\mathbf{x}). \end{aligned} \quad (2.27)$$

Then the marginal (unconditional) p.g.fl. of Y follows from averaging over all realizations \mathbf{x} ,

$$\begin{aligned} G_Y[h] &= \mathbf{E}_X [G_{Y|X}[h]] \\ &= \int G_{Y|X}[h] \mathbf{P}_X(d\mathbf{x}), \end{aligned} \quad (2.28)$$

and given a conditional p.g.fl., the joint p.g.fl. of X and Y follows the property of conditional expectation,

$$\begin{aligned} G_{XY}[g, h] &= \mathbf{E}_X [\Pi_{\mathbf{x}}[g] G_{Y|X}[h]] = \mathbf{E}_Y [\Pi_{\mathbf{y}}[h] G_{X|Y}[g]] \\ &= \int \Pi_{\mathbf{x}}[g] G_{Y|X}[h] \mathbf{P}_X(d\mathbf{x}) = \int \Pi_{\mathbf{y}}[h] G_{X|Y}[g] \mathbf{P}_Y(d\mathbf{y}). \end{aligned} \quad (2.29)$$

In particular, the following results are useful. Using $\Pi_{\mathbf{x}}[1] = 1$, we obtain

$$G_{XY}[1, h] = \mathbf{E}_X [\Pi_{\mathbf{x}}[1] G_{Y|X}[h]] = G_Y[h]. \quad (2.30)$$

Lemma 2.6. *Let $\mathbf{y} = \{y_1, \dots, y_m\}$ be a realization of a point process $Y \in \mathcal{N}_{\mathcal{Y}}^s$ and $(d^m G_{XY}[g, \cdot])_{h=0} [\delta_{y_1}, \dots, \delta_{y_m}]^2$ be the m -th derivative w.r.t. h of the joint p.g.fl. eval-*

² For a definition of the functional derivative see Appendix (A.2)

uated at the origin where δ_{y_i} is the Dirac delta function at point y_i , then

$$G_{X|Y}[g|\mathbf{y}] = \frac{(d^m G_{XY}[g, \cdot])_{h=0}[\delta_{y_1}, \dots, \delta_{y_m}]}{(d^m G_{XY}[1, \cdot])_{h=0}[\delta_{y_1}, \dots, \delta_{y_m}]} \quad (2.31)$$

Proof. Deferred to the appendix (A.1). □

Proposition 2.7. *Let \mathcal{X} and \mathcal{Y} be two LCHS spaces and $Y \in \mathcal{N}_{\mathcal{Y}}^s$ be a point process related to a realization \mathbf{x} of the Poisson point process $X \in \mathcal{N}_{\mathcal{X}}^s$ with intensity measure Λ_X by*

$$Y(\mathbf{x}) = \bigcup_{x \in \mathbf{x}} S(x),$$

where $S(x)$ is either singleton or empty with distribution

$$\begin{aligned} \mathbf{P}_{S(x)}(\mathcal{S}) &= \mathbf{P}(S(x) \in \mathcal{S}) \\ &= \mathbf{1}_{\emptyset}(\mathcal{S} \cap \emptyset)(1 - p_S(x)) + p_S(x)F(\mathcal{S} \cap \mathcal{Y}, x), \end{aligned}$$

for any \mathcal{S} in the family of all simple counting measures on $\mathcal{B}_{\mathcal{Y}}$, $F: \mathcal{B}_{\mathcal{Y}} \times \mathcal{X} \mapsto \mathbb{R}^+$ is a kernel function (see [34, pp. 641], [136]); for fixed x , $F(\cdot, x)$ is a probability measure on $\mathcal{B}_{\mathcal{Y}}$ with density $f(\cdot, \cdot)$ w.r.t. a reference Lebesgue measure $\lambda_{\mathcal{Y}}$ and p_S is a $\mathcal{B}_{\mathcal{X}}$ -measurable function with $0 \leq p_S(x) \leq 1$ for all $x \in \mathbf{x}$. Then, the conditional intensity measure $\Lambda_{X|Y}$ is given by

$$\zeta \circ \Lambda_{X|Y} = \left(1 - p_S + \sum_{y \in \mathbf{y}} \frac{p_S f(y, \cdot)}{p_S f(y, \cdot) \circ \Lambda_X} \right) \zeta \circ \Lambda_X, \quad (2.32)$$

where $\xi \circ \Psi = \int \xi(x)\Psi(dx)$.

Proof. The p.g.fl. of Y conditional on X is

$$\begin{aligned} G_{Y|X}[h] &= \prod_{x \in \mathbf{x}} G_{S(x)}[h] \\ &= \Pi_{\mathbf{x}}(1 - p_S(x) + p_S(x)h \circ F) \\ &= \Pi_{\mathbf{x}}A[h], \end{aligned} \quad (2.33)$$

where $A[h] = 1 - p_S(x) + p_S(x)h \circ F$. From (2.29) the joint p.g.fl. of X and Y is

$$\begin{aligned} G_{XY}[g, h] &= \mathbf{E}_X [\Pi_x[g]G_{Y|X}[h]] \\ &= \mathbf{E}_X [\Pi_x[g]A[h]] \\ &= G_X[gA[h]] \end{aligned} \quad (2.34)$$

From (2.24) the p.g.fl. of a Poisson point process X takes the form

$$G_{XY}[g, h] = \exp((gA[h] - 1) \circ \Lambda_X) \quad (2.35)$$

The m -th derivative w.r.t. h of the joint p.g.fl. above evaluated at the origin gives

$$(d^m G_{XY}[g, \cdot])_{h=0}[\xi_1, \dots, \xi_m] = e^{B[g]} \prod_{i=1}^m g p_S F \xi_i \circ \Lambda_X, \quad (2.36)$$

where $B[g] = (g(1 - p_S) - 1) \circ \Lambda_X$. Let $\mathbf{y} = \{y_1, \dots, y_m\}$ be the realization of Y . Since density $f(\cdot, \cdot)$ exists w.r.t. λ_Y for the regular kernel $F(\cdot, \cdot)$,

$$\begin{aligned} H[g] &= (d^m G_{XY}[g, \cdot])_{h=0}[\delta_{y_1}, \dots, \delta_{y_m}] \\ &= e^{B[g]} \Pi_{\mathbf{y}} g p_S f(y, \cdot) \circ \Lambda_X. \end{aligned} \quad (2.37)$$

Differentiating w.r.t. g ,

$$(dH)_g[\zeta] = \left(de^{B[\cdot]} \right)_g [\zeta] \Pi_{\mathbf{y}} g p_S f(y, \cdot) \circ \Lambda_X + e^{B[g]} (d\Pi_{\mathbf{y}} (g p_S f(y, \cdot)) \circ \Lambda_X)_g [\zeta] \quad (2.38)$$

where

$$\left(de^{B[\cdot]} \right)_g [\zeta] = e^{B[g]} (1 - p_S) \zeta \circ \Lambda_X \quad (2.39)$$

and

$$(d\Pi_{\mathbf{y}} (g p_S f(y, \cdot)) \circ \Lambda_X)_g [\zeta] = \sum_{y \in \mathbf{y}} p_S f(y, \cdot) \zeta(y) \circ \Lambda_X \prod_{\mathbf{y} \setminus \{y\}} g p_S f(y, \cdot) \circ \Lambda_X \quad (2.40)$$

Substituting in (2.38)

$$(dH)_g[\zeta] = H[1] \left(1 - p_S + \sum_{y \in \mathbf{y}} \frac{p_S f(y, \cdot)}{g p_S f(y, \cdot) \circ \Lambda_X} \right) \zeta \circ \Lambda_X \quad (2.41)$$

Applying Lemma 2.6 gives the required result. \square

2.4 Random set filtering

In this section the filtering problem is stated in the random set formalism by treating the state and observation at each time step k by finite sets. The RFS approach provides a natural representation of a random number of points in the state space \mathcal{X} and the observation space \mathcal{Z} . The meta-state $X_k = \{x_{k,1}, \dots, x_{k,m}\} \in \mathcal{F}_{\mathcal{X}}$ is the RFS of the states at time k . Note that the sequence in which points are arranged is arbitrary with all permutations equally likely, so $x_{k,1}$ does not allude to a point closest to a reference. Similarly, the observation $Z_k = \{z_{k,1}, \dots, z_{k,n}\} \in \mathcal{F}_{\mathcal{Z}}$ is the RFS of measurements at time k where the points are arranged in an unordered configuration.

Let \mathbf{x}_{k-1} be a realization of the corresponding point process X_{k-1} at time $k-1$. Consider the point process X_k constructed as the disjoint union

$$X_k = \left[\bigcup_{x \in \mathbf{x}_{k-1}} S_{k|k-1}(x) \right] \cup \Gamma_k, \quad (2.42)$$

where Γ_k is a Poisson point process with intensity measure $\Lambda_{\Gamma,k}$ independent of $S_{k|k-1}(x)$ which is either singleton or empty with distribution

$$\begin{aligned} \mathbf{P}_{S_{k|k-1}(x)}(\mathcal{S}) &= \mathbf{P}(S_{k|k-1}(x) \in \mathcal{S}) \\ &= \mathbf{1}_{\emptyset}(\mathcal{S} \cap \emptyset)(1 - p_{S,k|k-1}(x)) + p_{S,k|k-1}(x)F_k(\mathcal{S} \cap \mathcal{X}, x), \end{aligned}$$

for any \mathcal{S} in the space of all simple counting measures on $\mathcal{B}_{\mathcal{X}}$, $F_k : \mathcal{B}_{\mathcal{X}} \times \mathcal{X} \mapsto \mathbb{R}^+$ is a transition probability and $p_{S,k|k-1}$ is a $\mathcal{B}_{\mathcal{X}}$ -measurable function with $0 \leq p_{S,k|k-1}(x) \leq 1$ for all $x \in \mathbf{x}_{k-1}$.

Let \mathbf{z}_k be a realization of the point process Z_k which defines the observation model of RFS X_k

$$Z_k = \left[\bigcup_{x \in \mathbf{x}_k} D_k(x) \right] \cup K_k, \quad (2.43)$$

where K_k is a Poisson point process with intensity measure $\Lambda_{\Gamma,k}$ and admits a density

$\lambda_{K,k}$ independent of $D_k(x)$ which is either singleton or empty with distribution

$$\begin{aligned} \mathbf{P}_{D_k(x)}(\mathcal{D}) &= \mathbf{P}(D_k(x) \in \mathcal{D}) \\ &= \mathbf{1}_{\emptyset}(\mathcal{D} \cap \emptyset)(1 - p_{D,k}(x)) + p_{D,k}(x)L_k(\mathcal{D} \cap \mathcal{Z}, x), \end{aligned}$$

for any \mathcal{D} in the family of all simple counting measures on $\mathcal{B}_{\mathcal{Z}}$, $L_k : \mathcal{B}_{\mathcal{Z}} \times \mathcal{X} \mapsto \mathbb{R}^+$ is a transition probability with density $l_k(\cdot, \cdot)$ and $p_{D,k}$ is a $\mathcal{B}_{\mathcal{Z}}$ -measurable function with $0 \leq p_{D,k}(x) \leq 1$ for all $x \in \mathbf{x}_k$.

Theorem 2.8. *Given the intensity measure of the RFS X_{k-1} at time $k-1$ as $\Lambda_{X_{k-1}}$ and a Markov process (2.42), the intensity measure of RFS X_k at time k is given by*

$$\zeta \circ \Lambda_{X_k} = p_{S,k|k-1} F_k \zeta \circ \Lambda_{X_{k-1}} + \zeta \circ \Lambda_{\Gamma,k}. \quad (2.44)$$

Proof. Using the property of the p.g.fl. for independent point processes in (2.25) and the result in (2.33), the conditional p.g.fl. of X_k given X_{k-1} is

$$G_{X_k|X_{k-1}}[h] = [\Pi_{\mathbf{x}_{k-1}} A_k[h]] \cdot G_{\Gamma,k}[h], \quad (2.45)$$

where $A_k[h] = 1 - p_{S,k|k-1}(x) + p_{S,k|k-1}(x)h \circ F_k$. From (2.28), the unconditional p.g.fl. of X_k factors into the p.g.fl. of two independent point processes

$$G_{X_k}[h] = G_{X_{k-1}}[A_k[h]] G_{\Gamma,k}[h]. \quad (2.46)$$

The required intensity measure follows from (2.23) by differentiating w.r.t. h at unity,

$$\begin{aligned} \zeta \circ \Lambda_{X_k} &= d(G_{X_k}[\cdot])_{h=1}[\zeta] \\ &= (dA[\cdot])_{h=1}[\zeta] \circ \Lambda_{X_{k-1}} + \zeta \circ \Lambda_{\Gamma,k}, \end{aligned} \quad (2.47)$$

where

$$(dA[\cdot])_{h=1}[\zeta] = p_{S,k|k-1} F_k \zeta. \quad (2.48)$$

This completes the proof. \square

Theorem 2.9. *Given the unconditional intensity measure of the Poisson RFS X_k as Λ_{X_k} and a realization \mathbf{z}_k of the point process Z_k at time k , the conditional intensity measure of RFS X_k given \mathbf{z}_k is*

$$\zeta \circ \Lambda_{X_k|\mathbf{z}_k} = \left(1 - p_{D,k} + \sum_{z \in \mathbf{z}_k} \frac{p_{D,k} l_k(z, \cdot)}{\lambda_{K,k} + p_{D,k} l_k(z, \cdot) \circ \Lambda_{X_k}} \right) \zeta \circ \Lambda_{X_k}. \quad (2.49)$$

Proof. The proof is similar to that of Proposition 2.7 with $p_S = p_{D,k}$ and $f(\cdot, \cdot) = l_k(\cdot, \cdot)$. Using independence of X_k and K_k , the p.g.fl. of Z_k conditional on a realization \mathbf{x}_k of RFS X_k is given as the product of two p.g.fl.s.,

$$G_{Z_k|X_k}[h] = [\Pi_{\mathbf{x}_k} A_k[h]] \cdot G_{K,k}[h], \quad (2.50)$$

where $A_k[h] = 1 - p_{D,k}(x) + p_{D,k}(x)h \circ L_k$. The joint p.g.fl. of X_k and Z_k is then the product of the p.g.fl.s. of two independent Poisson point processes X_k and K_k .

$$G_{X_k Z_k}[g, h] = \exp(g A_k[h] \circ \Lambda_{X_k} + h \circ \Lambda_{K_k} - 1 \circ (\Lambda_{X_k} + \Lambda_{K_k})). \quad (2.51)$$

The n -th derivative w.r.t. h of the joint p.g.fl. at $h = 0$ gives

$$(d^n G_{X_k Z_k}[g, \cdot])_{h=0}[\xi_1, \dots, \xi_m] = e^{B_k[g]} \prod_{i=1}^n g p_{D,k} L_k \xi_i \circ \Lambda_{X_k} + \xi_i \circ \Lambda_{K_k}, \quad (2.52)$$

where $B_k[g] = g(1 - p_{D,k}) \circ \Lambda_{X_k} - 1 \circ (\Lambda_{X_k} + \Lambda_{K_k})$. Let $\mathbf{z}_k = \{z_1, \dots, z_n\}$. Since density $l_k(\cdot, \cdot)$ exists for the regular kernel $L_k(\cdot, \cdot)$,

$$\begin{aligned} H_k[g] &= (d^n G_{X_k Z_k}[g, \cdot])_{h=0}[\delta_{z_1}, \dots, \delta_{z_m}] \\ &= e^{B_k[g]} \Pi_{\mathbf{z}_k} g p_{D,k} l(z, \cdot) \circ \Lambda_{X_k} + \lambda_{K,k}. \end{aligned} \quad (2.53)$$

Differentiating w.r.t. g ,

$$(dH_k)_g[\zeta] = \left(de^{B_k[\cdot]} \right)_g [\zeta] \Pi_{\mathbf{z}_k} g p_{D,k} l_k(z, \cdot) \circ \Lambda_{X_k} + e^{B_k[g]} (d\Pi_{\mathbf{z}_k} (g p_{D,k} l_k(z, \cdot)) \circ \Lambda_{X_k})_g[\zeta] \quad (2.54)$$

where

$$\left(de^{B_k[\cdot]} \right)_g [\zeta] = e^{B_k[g]} (1 - p_{D,k}) \zeta \circ \Lambda_{X_k} \quad (2.55)$$

and

$$(d\Pi_{\mathbf{z}_k}(g_{PD,k}l_k(z, \cdot)) \circ \Lambda_{X_k})_g[\zeta] = \sum_{z \in \mathbf{z}_k} p_{D,k}l_k(z, \cdot)\zeta \circ \Lambda_{X_k} \prod_{\mathbf{z}_k \setminus \{z\}} g_{PD,k}l_k(z, \cdot) \circ \Lambda_{X_k} \quad (2.56)$$

Substituting in (2.54)

$$(dH_k)_g[\zeta] = H_k[1] \left(1 - p_{D,k} + \sum_{z \in \mathbf{z}_k} \frac{p_{D,k}l_k(z, \cdot)}{\lambda_{K,k} + g_{PD,k}l_k(z, \cdot) \circ \Lambda_{X_k}} \right) \zeta \circ \Lambda_{X_k} \quad (2.57)$$

The required expression of the conditional intensity measure of $X_k|\mathbf{z}_k$ is given by the result (2.23) after applying Lemma 2.6. \square

Theorems 2.8 and 2.9 give the recursion of the probability hypothesis density (PHD) filter [92, 45] for propagating the intensity measure in time and provide the basis of the closed form solution to the PHD filter for the so-called *linear Gaussian jump Markov system multi-target model* discussed in Section 3.3.

2.5 Nonlinear filtering

The class of state space models for which Kalman filter gives the optimal estimate encompasses only a small subset of real systems. For the general class of problems Bayes filter concedes an approximation in order to estimate the state. For the nonlinear state space model (2.1)-(2.2) both state and observation variables x_k and z_k respectively are not necessarily Gaussian. Thus, the central issue with using linear estimation for filtering nonlinear models is the approximation of second-order moments of all concerned state and observation random variables.

In the literature there are two standard approaches to approximate the conditional expectation of the state. Both the extended Kalman filter (EKF) [59, 44] and the unscented Kalman filter (UKF) [62, 63] actually involve only linear estimators (by using Theorem 2.1) and are different only in the way the second-order moments are approximated. The EKF uses the first-order Taylor series approximation of the mappings

f and g around some estimate. The approximation assumes that the estimate lies in the proximity of the global trajectory. With increasing nonlinearity trend of the mappings, the validity of the assumption weakens, causing the estimate to become biased and inconsistent. Furthermore, the convergence of the estimate is not guaranteed. This is demonstrated by a simple example in [75] where the EKF fails to converge. The EKF is better suited for the class of mildly nonlinear problems. On the other hand, the UKF applies the unscented transform [62, 63] based on the statistical linear regression technique to directly compute the first and second-order moments. The expected value of a random variable given by the UKF is correct up to the second-order if the second-order moments can be computed exactly, which is a higher order accuracy compared to the estimate given by the EKF. This is reflected in the empirical studies which demonstrate that the UKF gives better estimates than the EKF [62, 63, 77, 33, 115].

The EKF and the UKF approximations can be derived exactly using the linear regression Kalman filter (LRKF) [77, 78] which approximates the nonlinear mapping using statistical linear regression technique through some regression points. The EKF is derived using a single regression point only while the UKF is derived using $p = 2n$ regression points for an n -dimensional kinematic state.

Suppose $X := (X(1), \dots, X(n))^T \in \mathbb{R}^n$ is a random variable with $\bar{x} = (\bar{x}(1), \dots, \bar{x}(n))^T$ and R_X as the first two moments of its distribution $p_X(x)$. A second random variable Y depends on X through the differentiable nonlinear mapping $Y = f(X)$, where $f(X) = (f_1(X), \dots, f_m(X))^T$. As mentioned above, the central issue for linear estimation of X conditional on $Y = y$ is the approximation of computationally intractable integrals

$$\begin{aligned} \bar{y} := \mathbf{E}(Y) &= \int f(x)p_X(x)dx, \\ R_Y &= \int (f(x) - \bar{y})(f(x) - \bar{y})^T p_X(x)dx, \\ R_{XY} &= \int (x - \bar{x})(f(x) - \bar{y})^T p_X(x)dx. \end{aligned} \tag{2.58}$$

The EKF employs the simplest approximation, which is to linearize the nonlinear map-

ping f around \bar{x} by

$$f(X) \approx A(X - \bar{x}) + f(\bar{x}), \quad A = \left. \frac{\partial f(X)}{\partial X} \right|_{X=\bar{x}}. \quad (2.59)$$

This works well under the assumption that the expected value \bar{x} lies in the proximity of distributed values of X justifying the truncation of higher order terms of Taylor series.

Substituting the approximation of $f(X)$ from (2.59) in the integrals in (2.58) results in

$$\bar{y} \approx A\bar{x} + f(\bar{x}), \quad R_Y \approx AR_XA^T, \quad R_{XY} \approx R_XA^T. \quad (2.60)$$

Since the expected value is correct up to the first order only, as the mapping departs from linear behavior the validity of this assumption weakens. A bias is introduced in the estimate causing the estimates to become inconsistent. According to Theorem 2.1, the exact statistical form of $f(X)$ is actually

$$f(X) = R_{YX}R_X^{-1}(X - \bar{x}) + f(\bar{x}) + e, \quad (2.61)$$

The error $e = Y - R_{YX}R_X^{-1}(X - \bar{x}) - f(\bar{x})$ is a random quantity and is uncorrelated to X . In other words, the random variable Y can be expressed exactly by the statistical linear regression of $f(X)$ around \bar{x} . The quality of the linear approximation

$$f(X) \approx (AX + b) + e, \quad (2.62)$$

depends on how it matches (2.61), i.e. how A and b approximate $R_{YX}R_X^{-1}$ and $f(\bar{x}) - A\bar{x}$, respectively. It is not clear how far away the local approximation (2.59) distances from (2.61).

The UKF aims at the direct approximation of R_{YX} , R_X and $f(\bar{x})$ so its linearized model is more accurate than that of the EKF, as demonstrated in [62, 63]. Regression points x_i , $i = 0, \dots, p$ where $p = 2n$ are selected around \bar{x} in a manner such that the sample mean and covariance of the points are identical to the mean and covariance of X ,

$$\bar{x} = \frac{1}{p+1} \sum_{i=0}^p x_i, \quad R_X = \frac{1}{p+1} \sum_{i=0}^p (x_i - \bar{x})(x_i - \bar{x})^T. \quad (2.63)$$

As $R_X > 0$ and thus admits Cholesky decomposition $R_X = \sum_{i=0}^n q_i q_i^T$, a choice of these regression points is

$$x_0 = \bar{x}, \quad x_i = \bar{x} + \sqrt{(p+1)/2} q_i, \quad x_{n+i} = \bar{x} - \sqrt{(p+1)/2} q_i.$$

Let $y_i = f(x_i)$, $i = 0, \dots, p$, then the mean and covariance of the random variable Y and the cross-covariance of Y and X are approximated by the distribution of the regression points x_i and y_i , $i = 0, \dots, p$ as,

$$\bar{y} = \frac{1}{p+1} \sum_{i=0}^p y_i, \quad R_Y = \frac{1}{p+1} \sum_{i=0}^p (y_i - \bar{y})(y_i - \bar{y})^T, \quad (2.64)$$

$$R_{YX} = \frac{1}{p+1} \sum_{i=0}^p (y_i - \bar{y})(x_i - \bar{x})^T. \quad (2.65)$$

One can see that (2.64)-(2.65) are indeed approximations for the integrals (2.58) with the continuous distribution $p_X(x)$ approximated by the discrete uniform distribution

$$P(X = x_i) = 1/(p+1), \quad i = 0, \dots, p, \quad (2.66)$$

i.e, the distribution $p_X(x)$ is statistically linearized around the regression points x_i , $i = 0, \dots, p$ in the UKF. In summary, for approximation of integrals (2.58), the EKF linearizes the nonlinear deterministic mapping f while the UKF linearizes the random distribution $p_X(x)$.

Based on the above exposition on linear approximation, an intermediary linearization technique can be derived. Note that any differentiable nonlinear mapping f can be represented by

$$f(X) = A_i(X)(X - x_i) + f(x_i), \quad i = 0, \dots, p,$$

with

$$A_i(X) = \int_0^1 \frac{\partial}{\partial X} f(tX + (1-t)x_i) dt, \quad i = 0, \dots, p. \quad (2.67)$$

The approximation (2.59) can be considered as a special case,

$$f(X) \approx A_0(x_0)(X - x_0) + f(x_0), \quad (2.68)$$

with $x_0 = \bar{x}$. Let $y_i = A_i(X)(X - x_i) + f(x_i)$, $i = 0, \dots, p$, then the moments of Y can be computed from (2.64) and the joint moment of Y and X from (2.65). It can be shown that in general this technique works better than the approximation in (2.59).

Chapter 3

Gaussian mixture PHD filter for jump Markov system models

The probability hypothesis density (PHD) filter [92, 45] is an attractive approach to tracking an unknown and time-varying number of targets in the presence of data association uncertainty, clutter, noise, and detection uncertainty. The PHD filter admits a closed form solution for a linear Gaussian multi-target model [131, 133]. However, this model is not general enough to accommodate maneuvering targets that switch between several models. In this chapter, the notion of linear jump Markov systems is generalized to the multiple target case to accommodate births, deaths and switching dynamics. A closed form solution to the PHD recursion (Theorems 2.8 and 2.9) is then derived for the proposed linear Gaussian jump Markov multi-target model. Based on this an efficient method for tracking multiple maneuvering targets that switch between a set of linear Gaussian models is developed. An analytic implementation of the PHD filter using statistical linear regression technique is also proposed for targets that switch between a set of nonlinear models. It is demonstrated through simulations that the proposed PHD filters are effective in tracking multiple maneuvering targets.

The chapter is structured as follows: The problem of tracking multiple maneuvering targets in the presence of uncertainty is discussed in Section 3.1. Section 3.2 presents some background on JMS for modeling a maneuvering target and the PHD

filter. In Section 3.3 the JMS multi-target model for the PHD filter is described, the main result of this chapter, a closed-form solution to the PHD recursion for linear JMS is given and the capability of the proposed algorithm is demonstrated through simulations. In Section 3.4 the approximate solution to the PHD recursion for nonlinear JMS is discussed.

3.1 Introduction

While a non-maneuvering target motion can be described by a fixed model, a combination of motion models that characterize different maneuvers may be needed to describe the motion of a maneuvering target. Tracking a maneuvering target in clutter is a challenging problem and is the subject of numerous works [7, 10, 82, 83]. In the multi-target setting, the number of targets changes due to targets appearing, disappearing, and it is not known which target generated which measurement. Tracking multiple maneuvering targets involves jointly estimating the number of targets and their states at each time step in the presence of noise, clutter, uncertainties in target maneuvers, data association and detection. As such, this problem is extremely challenging in both theory and implementation.

The jump Markov system (JMS) or multiple models approach has proven to be an effective tool for single maneuvering target tracking [18, 10]. In this approach, the target can switch between a set of models in a Markovian fashion. The JMS approach can also be combined with traditional data association techniques such as joint probabilistic data association (JPDA) [6, 19, 130, 60] or multiple hypothesis tracking (MHT) [38, 73] to track multiple maneuvering targets. However, these data association-based approaches are computationally intensive in general and heuristic techniques are used to reduce the computational load.

Mahler's Probability Hypothesis Density (PHD) filter [92, 45] is a multi-target filter that circumvents the combinatorial computations due to data association while

accommodating detection uncertainty, Poisson false alarms, target motion and time-varying number of targets. The generic sequential Monte Carlo implementation of the PHD filter [137, 138] can, in principle, accommodate any Markovian target dynamics including jump Markov systems. However, the drawbacks of the particle approach are the large number of particles, and the unreliability of clustering techniques for extracting state estimates [138, 133]. These problems are alleviated in the *Gaussian mixture PHD filter* implementation, which is developed from a closed form solution to the PHD recursion for linear Gaussian multi-target models [131, 133]. This approach is efficient and is capable of handling certain types non-linear models [133] but is not general enough to accommodate JMS models. At present there is no tractable analytical techniques for tracking multiple targets with JMS dynamics.

In this chapter, the notion of linear jump Markov systems is generalized to the multiple target case to accommodate births, deaths and switching dynamics. A closed form solution to the PHD recursion is then derived for this so-called linear Gaussian jump Markov multi-target model. This solution generalizes the result in [131, 133] to a broader class of practical models. Based on this closed form solution, an efficient method is developed for tracking multiple maneuvering targets that switch between multiple models. The proposed approach can handle problems that are deemed intractable using traditional tracking techniques. Comparison with the classical IMMJPDA filter showed that the proposed approach is computationally much more efficient while exhibiting similar tracking performance, despite the fact that the IMMJPDA filter uses exact knowledge of the fixed number of targets. In addition, the proposed approach is extended to nonlinear jump Markov multi-target models by combining the closed form solution and the unscented transform (see Section 2.5). The proposed multi-target filters sidestep the data association problem and do not require clustering for extracting state estimates. Simulation results are presented to demonstrate the capability of the proposed method.

3.2 Problem formulation

In Section 3.2.1 the JMS is reviewed and in particular the class of linear JMS for modeling maneuvering targets. Using the random finite set (RFS) representations for multi-target states and sensor measurements, the problem is posed as a Bayesian filtering problem in Section 3.2.2. Section 3.2.3 describes the PHD filter.

3.2.1 Jump Markov system (JMS)

A jump Markov system (JMS) can be described by a set of parameterized state space models whose underlying parameters evolve with time according to a finite state Markov chain. Such a system finds a range of applications in signal processing and provides a natural means to model a maneuvering target whose behavior cannot be characterized at all times by a single model [18, 82, 83].

Let $\xi_k \in \mathbb{R}^n$ and $z_k \in \mathbb{R}^m$ denote the kinematic state (e.g. target coordinates and velocity) and observation, respectively, at time k . Suppose that $r_k \in \mathcal{M}$ is the label of the model in effect at time k , where \mathcal{M} denotes the (discrete) set of all model labels (also called modes). Then, the state evolution and measurement are described by the transition density and measurement likelihood:

$$\tilde{f}_{k|k-1}(\xi_k|\xi_{k-1}, r_k), \quad (3.1)$$

$$g_k(z_k|\xi_k, r_k). \quad (3.2)$$

In addition, the modes follow a discrete Markov chain with transition probability $t_{k|k-1}(r_k|r_{k-1})$ and the transition of the augmented state vector $x_k = [\xi_k^T, r_k]^T \in \mathcal{X} = \mathbb{R}^n \times \mathcal{M}$ is governed by

$$f_{k|k-1}(x_k|x_{k-1}) = \tilde{f}_{k|k-1}(\xi_k|\xi_{k-1}, r_k)t_{k|k-1}(r_k|r_{k-1}). \quad (3.3)$$

A linear Gaussian JMS (LGJMS) is a JMS with linear Gaussian models, i.e. conditioned on mode r_k the state transition density and observation likelihood are given

by

$$\tilde{f}_{k|k-1}(\xi_k|\xi_{k-1}, r_k) = \mathcal{N}(\xi_k; F_{k-1}(r_k)\xi_{k-1}, Q_k(r_k)), \quad (3.4)$$

$$g_k(z_k|\xi_k, r_k) = \mathcal{N}(z_k; H_k(r_k)\xi_k, R_k(r_k)), \quad (3.5)$$

where $\mathcal{N}(\cdot; m, Q)$ denotes a Gaussian density with mean m and covariance Q , $F_{k-1}(r_k)$ and $H_k(r_k)$ denote the transition and observation matrices of model r_k . $Q_k(r_k)$ and $R_k(r_k)$ denote covariance matrices of the process noise and measurement noise.

Tracking a maneuvering target amounts to estimating the kinematic state ξ_k or augmented state x_k at time k , from the sequence of observations $z_{1:k} = (z_1, \dots, z_k)$. The JMS (or multiple models) approach has been shown to be highly effective for maneuvering target tracking [18, 10].

3.2.2 Random finite sets in multi-target tracking

In a multi-target scenario, suppose that $x_{k,1}, \dots, x_{k,N(k)} \in \mathcal{X}$ are the augmented states at time k , where $N(k)$ denotes the number of targets. At the next time step, some of these targets may die, new targets may appear and the surviving targets evolve to their new states. At the sensor, $M(k)$ measurements $z_{k,1}, \dots, z_{k,M(k)} \in \mathbb{R}^m$ are received at time k , some of which are due to targets while the rest are clutter. Note that only some of the existing targets are detected by the sensor, and that the corresponding measurements are indistinguishable from clutter. Hence, the orders in which the states, and the measurements are listed bear no significance. Jointly estimating the time-varying number of states and the values of the states is a fundamentally difficult problem because in addition to the target maneuvers, the number of targets and the number of measurements both vary randomly in time and it is not known which target generated which measurement.

Mahler's finite set statistics (FISST) [92, 45, 88] approach provides an elegant Bayesian formulation of the multi-target filtering problem by treating the finite sets of targets and observations, at time k , as the *multi-target state* and *multi-target observa-*

tion, respectively

$$X_k = \{x_{k,1}, \dots, x_{k,N(k)}\} \subset \mathcal{X}, \quad (3.6)$$

$$Z_k = \{z_{k,1}, \dots, z_{k,M(k)}\} \subset \mathbb{R}^m. \quad (3.7)$$

To model uncertainty in multi-target states and observations, we appeal to the notion of a *random finite set* (RFS) (see Section 2.3). A RFS on a state space \mathcal{X} is simply a random variable taking values in the finite subsets of \mathcal{X} [97, 34]. The *intensity* of an RFS on \mathcal{X} is a non-negative function v on \mathcal{X} such that $v(x)$ is the instantaneous expected number of targets per unit volume at x (2.18). A RFS is *Poisson* if its cardinality distribution is Poisson with mean $N = \int v(x)dx$ and given a cardinality the elements of X are i.i.d. according to v/N (see Definition 2.5). The reader is referred to [138, 133] for overviews on FISST and [45, 92, 88] for comprehensive treatments.

Along the same vein as the single-target filtering problem, a multi-target transition density can be constructed from the RFS model for the time evolution of the multi-target state, which incorporates target motion, spontaneous births, spawnings (off existing targets) and deaths (2.42). Similarly, a multi-target likelihood can be constructed from the RFS measurement model, which accounts for detection uncertainty and clutter (2.43). The posterior distribution of the RFS of targets can be propagated in time by the multi-target Bayes recursion [45, 92, 138]. However, this recursion involves multiple integrals on the space of finite subsets of \mathcal{X} . In addition, the multi-target densities are combinatorial in nature. Hence, the multi-target Bayes filter is computationally intractable in general. Sequential Monte Carlo implementations can be found in [137, 138, 119, 134, 86], although these methods are still computationally intensive, especially when the number of targets is large.

3.2.3 The probability hypothesis density filter

An intelligent approximation to the multi-target Bayes filter, known as the Probability Hypothesis Density (PHD) filter, and which avoids any data association compu-

tations, has been proposed in [92]. The PHD filter propagates the posterior intensity of the RFS of targets in time, based on the following assumptions:

Assumption 3.1. Targets evolve in time and generate measurements independently of one another.

Assumption 3.2. The clutter RFS is Poisson and is independent of the measurements.

Assumption 3.3. The predicted multi-target RFS is Poisson.

Assumptions 3.1 and 3.2 are quite common in many multi-target tracking algorithms [7, 17]. The additional Assumption 3.3 is a reasonable approximation in applications where interactions between targets are negligible [92].

The PHD propagation is a recursion consisting of a prediction step and a data update step. Let $v_{k|k-1}$ and v_k denote the predicted intensity and posterior intensity at time k , respectively. Then the *PHD prediction* is given by

$$v_{k|k-1}(x) = \int [p_{S,k|k-1}(x')f_{k|k-1}(x|x') + \beta_{k|k-1}(x|x')] v_{k-1}(x') dx' + \gamma_k(x), \quad (3.8)$$

where it is understood that an integral with respect to a discrete variable means a sum, and

$f_{k|k-1}(\cdot|x')$ = probability density of a target at time k , given that its previous state is x' ,

$p_{S,k|k-1}(x')$ = probability that a target still exists at time k given that its previous state is x' ,

$\beta_{k|k-1}(\cdot|x')$ = intensity of the RFS of targets spawned at time k by a target with previous state x' ,

$\gamma_k(\cdot)$ = intensity of the birth RFS at time k .

On arrival of a new multi-target measurement, the posterior intensity v_k is computed from the predicted intensity $v_{k|k-1}$ via the *PHD update*:

$$v_k(x) = \left[1 - p_{D,k}(x) + p_{D,k}(x) \sum_{z \in Z_k} \frac{g_k(z|x)}{\kappa_k(z) + \int p_{D,k}(x) g_k(z|x) v_{k|k-1}(x) dx} \right] v_{k|k-1}(x), \quad (3.9)$$

where

- Z_k = multi-target measurement at time k ,
- $g_k(\cdot|x)$ = single-target measurement likelihood at time k ,
- $p_{D,k}(x)$ = probability of detection given a state x at time k ,
- $\kappa_k(\cdot)$ = intensity of clutter RFS at time k .

The PHD recursion is generally intractable due to the ‘curse of dimensionality’ in numerical integration. A generic sequential Monte Carlo (SMC) implementation was proposed in [137, 138] with relevant convergence results (see also [61, 31] for more detailed asymptotic studies). This so-called particle-PHD filter can accommodate targets with JMS dynamics, and has been used to track multiple maneuvering targets in [132, 110]. However, the drawbacks of the particle approach are the large number of particles, and the unreliability of clustering techniques for extracting state estimates [138, 133]. The recently proposed Gaussian mixture PHD filter [131, 133] does not suffer from these drawbacks but is not general enough to handle JMS dynamics. In the following sections, a closed form solution to the PHD recursion is derived for LGJMS dynamics and an efficient and reliable multi-target filter is developed for tracking maneuvering targets.

3.3 Closed form solution to the PHD recursion for LGJMS multi-target model

This section presents a closed-form PHD solution that can accommodate targets that switch between linear Gaussian models. The LGJMS multi-target model is described in Section 3.3.1 and the corresponding closed form PHD recursion is derived in Section 3.3.2. In Section 3.3.3, a general closed form solution to the PHD recursion is derived in the hybrid state space $\mathcal{X} = \mathbb{R}^n \times \mathcal{M}$. Illustrations of the proposed multi-target tracking algorithm on simulated data are given in Section 3.3.4.

For notational convenience, the symbol Θ is used to denote the ordered pair of mean and covariance (m, P) of a Gaussian distribution, i.e

$$\mathcal{N}(x; \Theta) = \mathcal{N}(x; m, P). \quad (3.10)$$

Given a linear Gaussian model $z = Hx + v$, where v is Gaussian noise with mean d and covariance matrix R , the notation Ω is used to denote the ordered triplet of model parameters (H, R, d) , and

$$\mathcal{L}(x, z; \Omega) = \mathcal{N}(z; Hx + d, R) \quad (3.11)$$

to denote the probability density at z . This notation is suggestive of the mapping of x to z via the linear model with parameter Ω . Note that $\mathcal{N}(x; m, P) = \mathcal{L}(m, x; (I, P, 0)) = \mathcal{L}(x, m; (I, P, 0))$.

3.3.1 Linear Gaussian jump Markov system multi-target models

This subsection presents the *linear Gaussian JMS (LGJMS) multi-target model*, which accommodates targets with switching linear dynamics. Campbell's theorem (see (2.19), [125]) is used in the modeling of target births and spawning.

In addition to Assumptions 3.1 - 3.3, the LGJMS multi-target model comprises a LGJMS model for individual targets, kinematic-independent survival and detection

probabilities, and models for target births and spawnings. Like the motion model, birth and spawning models are naturally described in terms of the kinematic state. However, while the distribution of the augmented state can be taken as the product of the mode distribution and the kinematic state distribution conditional on the mode, i.e. $p(\xi, r) = p(r)p(\xi|r)$, this line of reasoning does not extend to birth and spawning intensities. The intensity of the augmented state is not necessarily the product of the intensity of the mode and the intensity of the kinematic state conditioned on the mode.

To specify birth and spawning models for the kinematic state and mode that yield valid birth and spawning intensities in the augmented state, we appeal to a well-known result in point process theory, namely Campbell's theorem (for marked point processes) [125, pp. 106–108]. In particular, Campbell's theorem implies that the intensity of the point process on $\mathbb{R}^n \times \mathcal{M}$ formed by the Cartesian product of a point process on the kinematic state space \mathbb{R}^n , with intensity \tilde{v} , and a point process on the mode space \mathcal{M} , is given by

$$v(\xi, r) = p(r|\xi)\tilde{v}(\xi), \quad (3.12)$$

where $p(\cdot|\xi)$ is the mode distribution given that a point of the product point process has kinematic state ξ . Moreover, if the point process on \mathbb{R}^n is Poisson, then the product point process on $\mathbb{R}^n \times \mathcal{M}$ is also Poisson [71].

3.3.1.1 Birth model 1

In the context of the proposed multi-target birth model, the intensity of augmented state births at time k is given by

$$\gamma_k(\xi, r) = \pi_k(r|\xi)\tilde{\gamma}_k(\xi),$$

where $\tilde{\gamma}_k$ is the intensity of kinematic state births at time k , and $\pi_k(\cdot|\xi)$ is the probability distribution of the modes for a given birth with kinematic state ξ at time k . In line with the standard LGJMS assumption that the mode transition probability $t_{k|k-1}$ is not a function of the kinematic states, the LGJMS multi-target model also assumes that

the mode distribution does not depend on the kinematic state, i.e. $\pi_k(r|\xi) = \pi_k(r)$. Moreover, it is also assumed that the intensity $\tilde{\gamma}_k$ of kinematic state births is a Gaussian mixture

$$\tilde{\gamma}_k(\xi) = \sum_{i=1}^{J_{\gamma,k}} w_{\gamma,k}^{(i)} \mathcal{N}(\xi; \Theta_{\gamma,k}^{(i)}), \quad (3.13)$$

where $J_{\gamma,k}$, $w_{\gamma,k}^{(i)}$, $\Theta_{\gamma,k}^{(i)} = (m_{\gamma,k}^{(i)}, Q_{\gamma,k}^{(i)})$, $i = 1, 2, \dots, J_{\gamma,k}$ are given model parameters. The mean $m_{\gamma,k}^{(i)}$ is a peak of the intensity $\tilde{\gamma}_k$ and has the highest local concentrations of expected number of births, and represents, for example, airbases or airports where targets are most likely to appear. The covariance matrix $P_{\gamma,k}^{(i)}$ determines the spread of $\tilde{\gamma}_k$ in the vicinity of the peak $m_{\gamma,k}^{(i)}$. The weight $w_{\gamma,k}^{(i)}$ gives the expected number target births originating from $m_{\gamma,k}^{(i)}$.

Similarly, the intensity of augmented states spawned, at time k , from a target with augmented state $[\xi'^T, r']^T$, at time $k-1$, is given by

$$\beta_{k|k-1}(\xi, r|\xi', r') = \pi_{k|k-1}(r|\xi, \xi', r') \tilde{\beta}_{k|k-1}(\xi|\xi', r'),$$

where $\tilde{\beta}_{k|k-1}(\cdot|\xi', r')$ is the intensity of kinematic states spawned at time k from $[\xi'^T, r']^T$, and $\pi_{k|k-1}(\cdot|\xi, \xi', r')$ is the probability distribution of the mode for a given kinematic state ξ , spawned at time k from $[\xi'^T, r']^T$. Consistent with standard LGJMS assumption, the LGJMS multi-target model assumes that the mode distribution of a spawned target does not depend on its kinematic state nor its parent's kinematic state, i.e. $\pi_{k|k-1}(r|\xi, \xi', r') = \pi_{k|k-1}(r|r')$, and that the intensity $\tilde{\beta}_{k|k-1}(\cdot|\xi', r')$ of spawned kinematic states is a Gaussian mixture

$$\tilde{\beta}_{k|k-1}(\xi|\xi', r') = \sum_{j=1}^{J_{\beta,k|k-1}(r')} w_{\beta,k|k-1}^{(j)}(r') \mathcal{L}(\xi', \xi; \Omega_{\beta,k|k-1}^{(j)}(r')), \quad (3.14)$$

where $J_{\beta,k|k-1}(r')$, $w_{\beta,k|k-1}^{(j)}(r')$, $\Omega_{\beta,k|k-1}^{(j)}(r') = (F_{\beta,k-1}^{(j)}(r'), Q_{\beta,k-1}^{(j)}(r'), d_{\beta,k-1}^{(j)}(r'))$, $j = 1, 2, \dots, J_{\beta,k-1}(r')$ are given model parameters. A similar interpretation to $\tilde{\gamma}_k$ applies to the intensity $\tilde{\beta}_{k|k-1}$, except that the j th peak, $F_{\beta,k-1}^{(j)}(r')\xi' + d_{\beta,k-1}^{(j)}(r')$, is an affine function of ξ' . Usually, a spawned target is modeled to be in the proximity of its parent.

3.3.1.2 Birth model 2

Alternatively, by interchanging the roles of the kinematic state space and mode space in (3.8), consistent models for births and spawnings can also be derived¹. In this case, the intensity of augmented state births at time k is given by

$$\gamma_k(\xi, r) = \tilde{\gamma}(\xi|r)\pi_k(r),$$

where π_k is now the intensity of mode births and $\tilde{\gamma}_k(\cdot|r)$ is now the distribution of the birth kinematic state given mode r . Note that the intensity of mode births is not a function of kinematic state. It is assumed, in the LGJMS multi-target model, that the distribution $\tilde{\gamma}_k(\cdot|r)$ of kinematic state births is a Gaussian mixture

$$\tilde{\gamma}_k(\xi|r) = \sum_{i=1}^{J_{\gamma,k}(r)} w_{\gamma,k}^{(i)}(r) \mathcal{N}(\xi; \Theta_{\gamma,k}^{(i)}(r)), \quad (3.15)$$

where $J_{\gamma,k}(r)$, $w_{\gamma,k}^{(i)}(r)$, $\Theta_{\gamma,k}^{(i)}(r) = (m_{\gamma,k}^{(i)}(r), Q_{\gamma,k}^{(i)}(r))$, $i = 1, 2, \dots, J_{\gamma,k}(r)$ are given model parameters that depend on the mode r . Similarly, the intensity of augmented states spawned, at time k , spawned from $[\xi'^T, r']^T$ is

$$\beta_{k|k-1}(\xi, r|\xi', r') = \tilde{\beta}_{k|k-1}(\xi|r, \xi', r')\pi_{k|k-1}(r|\xi', r'),$$

where $\pi_{k|k-1}(\cdot|\xi', r')$ is now the intensity of mode spawnings and $\tilde{\beta}_{k|k-1}(\cdot|r, \xi', r')$ is now the distribution of spawned kinematic state given mode r . The LGJMS multi-target model assumes that the intensity of spawned modes does not depend on the kinematic state of its parent, i.e. $\pi_{k|k-1}(r|\xi', r') = \pi_{k|k-1}(r|r')$, and that the distribution $\tilde{\beta}_{k|k-1}(\cdot|r, \xi', r')$ of the spawned kinematic state is a Gaussian mixture

$$\tilde{\beta}_{k|k-1}(\xi|r, \xi', r') = \sum_{j=1}^{J_{\beta,k|k-1}(r, r')} w_{\beta,k|k-1}^{(j)}(r, r') \mathcal{L}(\xi', \xi; \Omega_{\beta,k|k-1}^{(j)}(r, r')), \quad (3.16)$$

¹ One technicality is that we need to restrict the kinematic state space to a compact subset of \mathbb{R}^n . This technicality does not pose any problem in practice since the targets occupy a bounded region of space.

where $J_{\beta,k|k-1}(r, r')$, $w_{\beta,k|k-1}^{(j)}(r, r')$, $\Omega_{\beta,k|k-1}^{(j)}(r, r') = (F_{\beta,k-1}^{(j)}(r, r'), Q_{\beta,k-1}^{(j)}(r, r'), d_{\beta,k-1}^{(j)}(r, r'))$, $j = 1, 2, \dots$, $J_{\beta,k|k-1}(r, r')$ are given model parameters that depend on the current mode r and the parent's previous mode r' .

From a modeling and application point of view, models 1 and 2 are different. However, from an algorithmic or computational viewpoint, the first model can be treated as a special case of the second model with the distribution of the birth kinematic state being independent of mode r , i.e., $\tilde{\gamma}_k(\xi|r) = \tilde{\gamma}_k(\xi)$.

Summarizing, in addition to Assumptions 3.1 - 3.3, the *linear Gaussian JMS (LGJMS) multi-target model*, assumes:

Assumption 3.4. Each target follows a LGJMS model, i.e. the dynamic and measurement models for the augmented state have the form:

$$f_{k|k-1}(\xi, r|\xi', r') = \mathcal{L}(\xi', \xi; \Omega_{f,k|k-1}(r))t_{k|k-1}(r|r'), \quad (3.17)$$

$$g_k(z|\xi, r) = \mathcal{L}(\xi, z; \Omega_{g,k}(r)), \quad (3.18)$$

where $\Omega_{f,k|k-1}(r) = (F_{f,k-1}(r), Q_{f,k}(r), 0)$ is the parameter of the linear target dynamics model conditioned on mode r , $\Omega_{g,k}(r) = (H_k(r), R_k(r), 0)$ is the parameters of the linear observation model conditioned on mode r , and $t_{k|k-1}(r|r')$ is the mode transition probability. In particular, conditional on mode r , $F_{f,k-1}(r)$ is the state transition matrix, $Q_{f,k}(r)$ is the process noise covariance matrix, $H_k(r)$ is the measurement matrix and $R_k(r)$ is the measurement noise covariance matrix.

Assumption 3.5. The probabilities of target survival and target detection are independent of the kinematic state:

$$p_{S,k|k-1}(\xi', r') = p_{S,k|k-1}(r') \quad (3.19)$$

$$p_{D,k}(\xi, r) = p_{D,k}(r). \quad (3.20)$$

Assumptions 3.4 and 3.5 follow from those commonly used in maneuvering target tracking algorithms (see for example [7, 16, 113]),

Assumption 3.6. The intensities of the birth and spawn RFSs can be expressed as Gaussian mixtures of the form:

$$\gamma_k(\xi, r) = \pi_k(r) \sum_{i=1}^{J_{\gamma,k}(r)} w_{\gamma,k}^{(i)}(r) \mathcal{N}(\xi; \Theta_{\gamma,k}^{(i)}(r)), \quad (3.21)$$

$$\beta_{k|k-1}(\xi, r|\xi', r') = \pi_{k|k-1}(r|r') \sum_{j=1}^{J_{\beta,k|k-1}(r,r')} w_{\beta,k|k-1}^{(j)}(r, r') \mathcal{L}(\xi', \xi; \Omega_{\beta,k|k-1}^{(j)}(r, r')), \quad (3.22)$$

where $J_{\gamma,k}(r)$, $w_{\gamma,k}^{(i)}(r)$, $\Theta_{\gamma,k}^{(i)}(r) = (m_{\gamma,k}^{(i)}(r), Q_{\gamma,k}^{(i)}(r))$, $i = 1, 2, \dots, J_{\gamma,k}(r)$ are given parameters of the (Gaussian mixture) density of the kinematic state of a new born target with mode r at time k , and $\pi_k(\cdot)$ is the intensity of mode births at time k . Similarly, $J_{\beta,k|k-1}(r, r')$, $w_{\beta,k|k-1}^{(j)}(r, r')$, $\Omega_{\beta,k|k-1}^{(j)}(r, r') = (F_{\beta,k-1}^{(j)}(r, r'), Q_{\beta,k-1}^{(j)}(r, r'), d_{\beta,k-1}^{(j)}(r, r'))$, $j = 1, 2, \dots, J_{\beta,k-1}(r, r')$ are given parameters of the (Gaussian mixture) density of the kinematic state of a target with mode r , spawned at time k from a target with augmented state $[\xi'^T, r']^T$ at time $k-1$, and $\pi_{k|k-1}(\cdot|r')$ is the intensity of modes spawned at time k from a target with mode r' at time $k-1$.

The LGJMS multi-target model is more general than those in standard multi-target tracking algorithms. While most existing algorithms do not account for births or spawnings, the proposed multi-target model incorporates both. Models for births and spawnings for a given mode r accommodate different intensities of mode births and modes spawned respectively when births and spawnings are likely to vary between different modes. Similarly, the proposed model incorporates models for target death (survival) and target detection for a given mode r . Moreover, traditional multi-target filtering techniques are computationally intractable for a model of such generality. As shown later, using a random finite set approach [92], this model is amenable to computationally efficient multi-target filtering techniques.

3.3.2 Closed form PHD recursion for LGJMS multi-target model

To derive the closed form PHD recursion for the LGJMS multi-target model, Lemmas 1 and 2 in [131, 133] are required, which are stated using the new notation as

follows:

Lemma 3.7. *If $\Omega = (H, R, d)$ and $\Theta = (m, P)$, then*

$$\int \mathcal{L}(x, z; \Omega) \mathcal{N}(x; \Theta) dx = \mathcal{N}(z; \Pi(\Omega, \Theta)) \quad (3.23)$$

where

$$\Pi(\Omega, \Theta) = (Hm + d, R + HPH^T)$$

Lemma 3.8. *If $\Omega = (H, R, d)$ and $\Theta = (m, P)$, then*

$$\mathcal{L}(x, z; \Omega) \mathcal{N}(x; \Theta) = \mathcal{N}(z; \Pi(\Omega, \Theta)) \mathcal{N}(x; \Psi(z, \Omega, \Theta)) \quad (3.24)$$

where

$$\Psi(z, \Omega, \Theta) = (\tilde{m}(z - d), \tilde{P}) \quad (3.25)$$

$$\tilde{m}(z - d) = m + K(z - d - Hm) \quad (3.26)$$

$$\tilde{P} = (I - KH)P \quad (3.27)$$

$$K = PH^T(HPH^T + R)^{-1} \quad (3.28)$$

Proposition 3.9. *For a LGJMS multi-target model, if the posterior intensity v_{k-1} at time $k - 1$ has the form*

$$v_{k-1}(\xi', r') = \sum_{i=1}^{J_{k-1}(r')} w_{k-1}^{(i)}(r') \mathcal{N}(\xi'; \Theta_{k-1}^{(i)}(r')). \quad (3.29)$$

Then the predicted intensity $v_{k|k-1}$ is given by

$$v_{k|k-1}(\xi, r) = \gamma_k(\xi, r) + v_{f,k|k-1}(\xi, r) + v_{\beta,k|k-1}(\xi, r), \quad (3.30)$$

where

$$v_{\beta,k|k-1}(\xi, r) = \sum_{r'} \sum_{i=1}^{J_{k-1}(r')} \sum_{j=1}^{J_{\beta,k|k-1}(r,r')} w_{\beta,k|k-1}^{(i,j)}(r, r') \mathcal{N}(\xi; \Theta_{\beta,k|k-1}^{(i,j)}(r, r')), \quad (3.31)$$

$$w_{\beta,k|k-1}^{(i,j)}(r, r') = \pi_{k|k-1}(r|r') w_{\beta,k|k-1}^{(j)}(r, r') w_{k-1}^{(i)}(r'), \quad (3.32)$$

$$\Theta_{\beta,k|k-1}^{(i,j)}(r, r') = \Pi(\Omega_{\beta,k|k-1}^{(j)}(r, r'), \Theta_{k-1}^{(i)}(r')), \quad (3.33)$$

$$v_{f,k|k-1}(\xi, r) = \sum_{r'} \sum_{i=1}^{J_{k-1}(r')} w_{f,k|k-1}^{(i)}(r, r') \mathcal{N}(\xi; \Theta_{f,k|k-1}^{(i)}(r, r')), \quad (3.34)$$

$$w_{f,k|k-1}^{(i)}(r, r') = p_{S,k|k-1}(r') t_{k|k-1}(r|r') w_{k-1}^{(i)}(r'), \quad (3.35)$$

$$\Theta_{f,k|k-1}^{(i)}(r, r') = \Pi(\Omega_{f,k|k-1}(r), \Theta_{k-1}^{(i)}(r')). \quad (3.36)$$

Proof. From (3.8), the predicted intensity consists of three terms γ_k (already given in the multi-target model), $v_{\beta,k|k-1}$ and $v_{f,k|k-1}$, due to births, spawnings and motion, respectively. For $v_{\beta,k|k-1}$, substituting (3.22), (3.29) into $\int \beta_{k|k-1}(x|x') v_{k-1}(x') dx'$, exchanging the order of sums and integral, and applying Lemma 3.7 to individual terms yields (3.31). For $v_{f,k|k-1}$ substitute (3.17) and (3.29) into $\int f_{k|k-1}(x|x') v_{k-1}(x') dx'$, exchange the order of sums and integral, and apply Lemma 3.7 to individual terms to obtain (3.34). \square

Corollary 3.10. *Under the premises of Proposition 3.9, the expected number of predicted targets is*

$$\hat{N}_{k|k-1} = \hat{N}_{\gamma,k} + \hat{N}_{f,k|k-1} + \hat{N}_{\beta,k|k-1}, \quad (3.37)$$

where

$$\hat{N}_{\gamma,k} = \sum_r \sum_{i=1}^{J_{\gamma,k}(r)} \pi_k(r) w_{\gamma,k}^{(i)}(r), \quad (3.38)$$

$$\hat{N}_{\beta,k|k-1} = \sum_r \sum_{r'} \sum_{i=1}^{J_{k-1}(r')} \sum_{j=1}^{J_{\beta,k|k-1}(r,r')} \pi_{k|k-1}(r|r') w_{\beta,k|k-1}^{(j)}(r, r') w_{k-1}^{(i)}(r'), \quad (3.39)$$

$$\hat{N}_{f,k|k-1} = \sum_r \sum_{r'} \sum_{i=1}^{J_{k-1}(r')} p_{S,k|k-1}(r') t_{k|k-1}(r|r') w_{k-1}^{(i)}(r'), \quad (3.40)$$

Proposition 3.11. *For a LGJMS multi-target model, if the predicted intensity $v_{k|k-1}$ has the form*

$$v_{k|k-1}(\xi, r) = \sum_{i=1}^{J_{k|k-1}(r)} w_{k|k-1}^{(i)}(r) \mathcal{N}(\xi; \Theta_{k|k-1}^{(i)}(r)). \quad (3.41)$$

Then the posterior intensity v_k is given by

$$v_k(\xi, r) = (1 - p_{D,k}(r))v_{k|k-1}(\xi, r) + \sum_{z \in Z_k} v_{g,k}(\xi, r; z), \quad (3.42)$$

where

$$v_{g,k}(\xi, r; z) = \sum_{i=1}^{J_{k|k-1}(r)} w_{g,k}^{(i)}(r; z) \mathcal{N}(\xi; \Theta_{g,k}^{(i)}(r; z)), \quad (3.43)$$

$$w_{g,k}^{(i)}(r; z) = \frac{p_{D,k}(r) w_{k|k-1}^{(i)}(r) q_{g,k}^{(i)}(r; z)}{\kappa_k(z) + \sum_r p_{D,k}(r) \sum_{i=1}^{J_{k|k-1}(r)} w_{k|k-1}^{(i)}(r) q_{g,k}^{(i)}(r; z)}, \quad (3.44)$$

$$q_{g,k}^{(i)}(r; z) = \mathcal{N}(z; \Pi(\Omega_{g,k}(r), \Theta_{k|k-1}^{(i)}(r))), \quad (3.45)$$

$$\Theta_{g,k}^{(i)}(r; z) = \Psi(z, \Omega_{g,k}(r), \Theta_{k|k-1}^{(i)}(r)). \quad (3.46)$$

Proof. From (3.9), the updated intensity consists of three components. The first is the predicted intensity $v_{k|k-1}$ (given), the second is the product $p_{D,k}v_{k|k-1}$ denoted as $v_{D,k}$, and the third is the sum $\sum_{z \in Z_k} v_{g,k}(x; z)$, where

$$v_{g,k}(x; z) = \frac{g_k(z|x)v_{D,k}(x)}{\kappa_k(z) + \int g_k(z|x)v_{D,k}(x)dx}. \quad (3.47)$$

For $v_{g,k}$, first substitute (3.18), (3.41) into the numerator of (3.47) and apply Lemma 3.8 to yield a sum of weighted Gaussians. Second, applying Lemma 3.7 to the integral in the denominator of (3.47) gives the (double) sum in the denominator of (3.44). Combining the results for the numerator and denominator of (3.47) gives (3.43). \square

Corollary 3.12. *Under the premises of Proposition 3.11, the expected number of target is*

$$\hat{N}_k = \sum_r [1 - p_{D,k}(r)] \sum_{i=1}^{J_{k|k-1}(r)} w_{k|k-1}^{(i)}(r) + \sum_{z \in Z_k} \sum_r \sum_{i=1}^{J_{k|k-1}(r)} w_{g,k}^{(i)}(r; z). \quad (3.48)$$

Propositions 3.9 and 3.11 show how the intensities $v_{k|k-1}$ and v_k are analytically propagated in time under linear Gaussian assumption on the JMS multi-target model. The recursions for the means and covariances of $v_{f,k|k-1}$ and $v_{\beta,k|k-1}$ are the Kalman prediction and the recursive computations of the means and covariances of $v_{D,k}$ are the Kalman update. The PHD filter has a complexity of $\mathcal{O}(J_{k-1}|Z_k|)$ where J_{k-1} is the number of Gaussian components representing v_{k-1} for a fixed model r' at time $k-1$ and $|Z_k|$ denotes the number of measurements at time k .

These propositions also indicate that the number of components of the predicted and posterior intensity increases with time, which can be a problem in implementation. However, this problem can be effectively handled by applying some simple pruning procedures [131, 133].

Given the posterior intensity v_k at time k

$$v_k(\xi, r) = \sum_{i=1}^{J_k(r)} w_k^{(i)}(r) \mathcal{N}(\xi; \Theta_k^{(i)}(r)), \quad (3.49)$$

the peaks of the intensity are points of highest local concentration of the expected number of targets. In order to extract the state of the targets from the posterior intensity at time k , an estimate of the number of targets \hat{N}_k is needed. This number is simply $\sum_{i=1}^{J_k(r)} w_k^{(i)}(r)$ rounded to the nearest integer. The estimate of the multi-target state is the set of \hat{N}_k ordered pairs of means and modes $(m_k^{(i)}(r), r)$ with the largest weights $w_k^{(i)}(r)$, $r \in \mathcal{M}, i = 1, \dots, J_k(r)$.

3.3.3 General solution to the PHD recursion

Apart from the LG and LGJMS multi-target models, the PHD recursion also admits closed form solutions under more general settings. In this section a general analytic solution to the PHD recursion is derived in the hybrid state space $\mathcal{X} = \mathbb{R}^n \times \mathcal{M}$. Readers who are interested in the simulation results of the above developed PHD filter may proceed directly to Section 3.3.4.

Proposition 3.13. *Given a multi-target transition model with*

$$p_{S,k|k-1}(\xi', r') = w_{S,k|k-1}^{(0)}(r') + \sum_{l=1}^{J_{S,k|k-1}(r')} w_{S,k|k-1}^{(l)}(r') \mathcal{L}(\xi', m_{S,k|k-1}^{(l)}(r'); \Omega_{S,k|k-1}^{(l)}(r')), \quad (3.50)$$

$$f_{k|k-1}(\xi, r|\xi', r') = \sum_{j=1}^{J_{f,k|k-1}(r, r')} w_{f,k|k-1}^{(j)}(r, r') \mathcal{L}(\xi', \xi; \Omega_{f,k|k-1}^{(j)}(r, r')), \quad (3.51)$$

$$\beta_{k|k-1}(\xi, r|\xi', r') = \sum_{j=1}^{J_{\beta,k|k-1}(r, r')} w_{\beta,k|k-1}^{(j)}(r, r') \mathcal{L}(\xi', \xi; \Omega_{\beta,k|k-1}^{(j)}(r, r')). \quad (3.52)$$

If the posterior intensity v_{k-1} at time $k-1$ has the form

$$v_{k-1}(\xi', r') = \sum_{i=1}^{J_{k-1}(r')} w_{k-1}^{(i)}(r') \mathcal{N}(\xi'; \Theta_{k-1}^{(i)}(r')). \quad (3.53)$$

Then the predicted intensity $v_{k|k-1}$ is given by

$$v_{k|k-1}(\xi, r) = \gamma_k(\xi, r) + v_{f,k|k-1}(\xi, r) + v_{\beta,k|k-1}(\xi, r), \quad (3.54)$$

where

$$v_{\beta,k|k-1}(\xi, r) = \sum_{r'} \sum_{i=1}^{J_{k-1}(r')} \sum_{j=1}^{J_{\beta,k|k-1}(r, r')} w_{\beta,k|k-1}^{(i,j)}(r, r') \mathcal{N}(\xi; \Theta_{\beta,k|k-1}^{(i,j)}(r, r')), \quad (3.55)$$

$$w_{\beta,k|k-1}^{(i,j)}(r, r') = w_{\beta,k|k-1}^{(j)}(r, r') w_{k-1}^{(i)}(r'), \quad (3.56)$$

$$\Theta_{\beta,k|k-1}^{(i,j)}(r, r') = \Pi(\Omega_{\beta,k|k-1}^{(j)}(r, r'), \Theta_{k-1}^{(i)}(r')), \quad (3.57)$$

$$v_{f,k|k-1}(\xi, r) = \sum_{r'} \sum_{i=1}^{J_{k-1}(r')} \sum_{l=0}^{J_{S,k|k-1}(r')} \sum_{j=1}^{J_{f,k|k-1}(r, r')} w_{f,k|k-1}^{(i,j,l)}(r, r') \mathcal{N}(\xi; \Theta_{f,k|k-1}^{(i,j,l)}(r, r')), \quad (3.58)$$

$$w_{f,k|k-1}^{(i,j,l)}(r, r') = w_{f,k|k-1}^{(j)}(r, r') w_{S,k|k-1}^{(l)}(r') w_{k-1}^{(i)}(r') q_{S,k|k-1}^{(i,l)}(r'), \quad (3.59)$$

$$q_{S,k|k-1}^{(i,l)}(r') = \mathcal{N}(m_{S,k|k-1}^{(l)}(r'); \Pi(\Omega_{S,k|k-1}^{(l)}(r'), \Theta_{k-1}^{(i)}(r'))), \quad q_{S,k|k-1}^{(i,0)}(r') = 1, \quad (3.60)$$

$$\Theta_{f,k|k-1}^{(i,j,l)}(r, r') = \Pi(\Omega_{f,k|k-1}^{(j)}(r, r'), \Theta_{S,k|k-1}^{(i,l)}(r')), \quad (3.61)$$

$$\Theta_{S,k|k-1}^{(i,l)}(r') = \Psi(m_{S,k|k-1}^{(l)}(r'), \Omega_{S,k|k-1}^{(l)}(r'), \Theta_{k-1}^{(i)}(r')), \quad \Theta_{S,k|k-1}^{(i,0)}(r') = \Theta_{k-1}^{(i)}(r'). \quad (3.62)$$

Proof. $v_{\beta,k|k-1}$ is obtained as before. For $v_{f,k|k-1}$ first substitute (3.50), (3.53) into $p_{S,k|k-1}(x')v_{k-1}(x')$ and applying Lemma 3.8 to yield a (double) sum of weighted Gaussians. Then substitute the resulting Gaussian mixture and (3.51) into $\int p_{S,k|k-1}(x') \cdot f_{k|k-1}(x|x')v_{k-1}(x')dx'$, exchange the order of sums and integral, and apply Lemma 3.7 to individual terms to obtain (3.58). \square

Proposition 3.14. *Given a multi-target measurement model with*

$$p_{D,k}(\xi, r) = w_{D,k}^{(0)}(r) + \sum_{l=1}^{J_{D,k}(r)} w_{D,k}^{(l)}(r) \mathcal{L}(\xi; m_{D,k}^{(l)}(r), \Omega_{D,k}^{(l)}(r')), \quad (3.63)$$

$$g_k(z|\xi, r) = \sum_{j=1}^{J_{g,k}(r)} w_{g,k}^{(j)}(r) \mathcal{M}(\xi; z, \Omega_{g,k}^{(j)}(r)). \quad (3.64)$$

If the predicted intensity $v_{k|k-1}$ has the form

$$v_{k|k-1}(\xi, r) = \sum_{i=1}^{J_{k|k-1}(r)} w_{k|k-1}^{(i)}(r) \mathcal{N}(\xi; \Theta_{k|k-1}^{(i)}(r)). \quad (3.65)$$

Then the posterior intensity v_k is given by

$$v_k(\xi, r) = v_{k|k-1}(\xi, r) - v_{D,k}(\xi, r) + \sum_{z \in Z_k} v_{g,k}(\xi, r; z), \quad (3.66)$$

where

$$v_{D,k}(\xi, r) = \sum_{i=1}^{J_{k|k-1}(r)} \sum_{l=0}^{J_{D,k}(r)} w_{D,k|k-1}^{(i,l)}(r) \mathcal{N}(\xi; \Theta_{D,k|k-1}^{(i,l)}(r)), \quad (3.67)$$

$$w_{D,k|k-1}^{(i,l)}(r) = w_{D,k}^{(l)}(r) w_{k|k-1}^{(i)}(r) q_{k|k-1}^{(i,l)}(r), \quad (3.68)$$

$$q_{D,k|k-1}^{(i,l)}(r) = \mathcal{N}(m_{D,k}^{(l)}(r); \Pi(\Omega_{D,k}^{(l)}(r), \Theta_{k|k-1}^{(i)}(r))), \quad q_{D,k|k-1}^{(i,0)}(r) = 1, \quad (3.69)$$

$$\Theta_{D,k|k-1}^{(i,l)}(r) = \Psi(m_{D,k}^{(l)}(r), \Omega_{D,k}^{(l)}(r), \Theta_{k-1}^{(i)}(r)), \quad \Theta_{k|k-1}^{(i,0)}(r) = \Theta_{k|k-1}^{(i)}(r), \quad (3.70)$$

$$v_{g,k}(\xi, r; z) = \sum_{i=1}^{J_{k|k-1}(r)} \sum_{l=0}^{J_{D,k}(r)} \sum_{j=1}^{J_{g,k}(r)} w_{g,k}^{(i,j,l)}(r; z) \mathcal{N}(\xi; \Theta_{g,k}^{(i,j,l)}(r; z)), \quad (3.71)$$

$$w_{g,k}^{(i,j,l)}(r; z) = \frac{w_{D,k|k-1}^{(i,l)}(r) w_{g,k}^{(j)}(r) q_{g,k}^{(i,j,l)}(r; z)}{\kappa_k(z) + \sum_{i=1}^{J_{k|k-1}(r)} \sum_{l=0}^{J_{D,k}(r)} \sum_{j=1}^{J_{g,k}(r)} w_{D,k|k-1}^{(i,l)}(r) w_{g,k}^{(j)}(r) q_{g,k}^{(i,j,l)}(r; z)}, \quad (3.72)$$

$$q_{g,k}^{(i,j,l)}(r; z) = \mathcal{N}(z; \Pi(\Omega_{g,k}^{(j)}(r), \Theta_{D,k|k-1}^{(i,l)}(r))), \quad (3.73)$$

$$\Theta_{g,k}^{(i,j,l)}(r; z) = \Psi(z, \Omega_{g,k}^{(j)}(r), \Theta_{D,k|k-1}^{(i,l)}(r)). \quad (3.74)$$

Proof. For $v_{D,k}$, substituting (3.63), (3.65) into $p_{D,k}(x)v_{k|k-1}(x)$, and applying Lemma 3.8 to individual terms yields (3.67). For $v_{g,k}$, first substitute (3.67), (3.64) into the numerator of (3.47) and apply Lemma 3.8 to yield a (triple) sum of weighted Gaussians. Second, applying Lemma 3.7 to the integral in the denominator of (3.47) gives the (triple) sum in the denominator of (3.72). Combining the results for the numerator and denominator of (3.47) gives (3.71). \square

3.3.4 Simulation results

In this subsection simulation results for two examples are presented to demonstrate the performance of the proposed PHD filter for LGJM models. For illustration purposes a two-dimensional scenario is considered where aircraft appear in the surveillance region $[-60, 60] \times [-60, 60] km^2$. A single sensor located at $(0, 0) km$ provides position-only measurements to a controller. The interval between the samples is $T = 5 s$ and the true number of aircraft at each sampling instant is not known.

During a level flight the aircraft dynamics can be modeled by a non-maneuver model and a maneuver model. Motion along a fixed heading at constant speed can be described by a non-maneuver model, for example, a constant velocity model. A level turn can be described by a maneuver model, for example, a coordinated turn model [10, 9]. The kinematic state of an aircraft is defined as $\xi = (p_x, \dot{p}_x, p_y, \dot{p}_y)^T$, where (p_x, p_y) denotes its Cartesian co-ordinates in the horizontal plane and (\dot{p}_x, \dot{p}_y) denotes its velocities. The speed of the aircraft is in the range Mach $[0.9, 1.1]$.

At a turn rate of $0^\circ s^{-1}$ the coordinated turn model reduces to the constant velocity model and the uniform motion of the aircraft can be modeled by the maneuver model. The aircraft motion models are described as follows. Model $r = 1$ is a coordinated turn model with a turn rate of $0^\circ s^{-1}$ with linear Gaussian dynamics (3.17) given by $\Omega_{f,k|k-1}(r = 1) = (F_{k-1}(r = 1), Q_k(r = 1), 0)$, with

$$F_{k-1}(r = 1) = \begin{bmatrix} 1 & \frac{\sin \omega T}{\omega} & 0 & -\frac{1 - \cos \omega T}{\omega} \\ 0 & \cos \omega T & 0 & -\sin \omega T \\ 0 & \frac{1 - \cos \omega T}{\omega} & 1 & \frac{\sin \omega T}{\omega} \\ 0 & \sin \omega T & 0 & \cos \omega T \end{bmatrix}, \quad Q_k(r = 1) = \sigma_{v_1}^2 \begin{bmatrix} \frac{T^4}{4} & \frac{T^3}{2} & 0 & 0 \\ \frac{T^3}{2} & T^2 & 0 & 0 \\ 0 & 0 & \frac{T^4}{4} & \frac{T^3}{2} \\ 0 & 0 & \frac{T^3}{2} & T^2 \end{bmatrix},$$

where ω denotes turn rate. Perturbations in the lift and drag characteristics due to changes in the properties of the atmosphere are modeled as zero-mean Gaussian white noise with a standard deviation, $\sigma_{v_1} = 5 m s^{-2}$.

Model $r = 2$ is a coordinated turn model with a counterclockwise turn rate of

$3^\circ s^{-1}$ with standard deviation of noise, $\sigma_{v_2} = 20 m s^{-2}$ to reflect the different noise characteristics during a level turn. Model $r = 3$ is also a coordinated turn model but with a clockwise turn rate of $3^\circ s^{-1}$. The switching between the motion models is given by Markovian transition probability matrix as

$$[t_{k|k-1}(r|r')] = \begin{bmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.8 & 0.1 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}.$$

The probability of target survival may change from one application to another and between different scenarios of an application. The reason is that in addition to some factors internal to the target, for example, aircraft altitude, fault-tolerance of instrumentation, fuel consumption and length of flight, target survival depends on certain external factors, for example, weather conditions. In general, the probability of target survival in military applications is lower than that in civilian applications where it may additionally depend on the maneuver an aircraft executes and the position of the aircraft relative to the location and type of threat (e.g. radar, anti-aircraft artillery, etc.) in the enemy surveillance region. A realistic model of the probability accounts for all of the above factors. In this chapter, modeling issues are not covered and it is assumed that a model of the probability is given. Furthermore, the probability of target survival may be treated as a random variable and incorporated in the state vector to be estimated. However, for simplicity it is assumed that the probability is known. $p_{S,k|k-1} = 0.99$ is assumed for modes $r' = 1, 2, 3$. Similarly, the probability of target detection may also vary depending on, for example, sensor characteristics, signal interference, weather conditions in civilian applications and in addition, countermeasures in military applications. A realistic model of the probability should consider these issues. Modeling issues are beyond the scope of this chapter. It is assumed that such a model is given. In the examples that follow $p_{D,k} = 0.98$ is assumed for modes $r = 1, 2, 3$.

Measurements follow the observation model (3.18) given by $\Omega_{g,k} = (H_k, R_k, 0)$,

with

$$H_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad R_k = \sigma_\epsilon^2 I_2,$$

where I_n denotes a $n \times n$ identity matrix. The error in the sensor measurements is modeled as zero-mean Gaussian white noise with a standard deviation, $\sigma_\epsilon = 40 m$. Clutter is modeled as a Poisson RFS with intensity $\kappa_k(z) = \lambda_c V \mathcal{U}(z)$, where $\mathcal{U}(\cdot)$ denotes a uniform density over the surveillance region, $V = 1.44 \times 10^4 km^2$ is the *volume* of the surveillance region and $\lambda_c = 3.47 \times 10^{-3} km^{-2}$ denotes the average number of clutter returns per unit volume.

The models for target births and spawnings are described next. Consider a scenario where the surveillance region includes three airport locations at $(40, -50) km$, $(-50, 40) km$ and $(-10, 0) km$. The intensity of the Poisson RFS of spontaneous births is given by

$$\gamma_k(\xi, r) = 0.1 \pi_k(r) \left(\mathcal{N}(\xi; m_\gamma^{(1)}, P_\gamma) + \mathcal{N}(\xi; m_\gamma^{(2)}, P_\gamma) + \mathcal{N}(\xi; m_\gamma^{(3)}, P_\gamma) \right),$$

with

$$\begin{aligned} m_\gamma^{(1)} &= (4 \times 10^4, 0, -5 \times 10^4, 0)^T, \\ m_\gamma^{(2)} &= (-5 \times 10^4, 0, 4 \times 10^4, 0)^T, \\ m_\gamma^{(3)} &= (-1 \times 10^4, 0, 0, 0)^T, \\ P_\gamma &= \text{diag} \left((10^6, 10^4, 10^6, 10^4) \right), \end{aligned}$$

and the distribution of the models at birth is taken as $[\pi_k(r)] = (0.8, 0.1, 0.1)$.

Also consider the case where payloads originating from an aircraft contribute to sensor measurements, the intensity of the Poisson RFS of spawn births is given by

$$\begin{aligned} \beta_{k|k-1}(\xi, r | \xi', r') &= 0.05 \pi_{k|k-1}(r | r') \mathcal{N}(\xi; \xi', Q_\beta), \\ Q_\beta &= \text{diag} \left((10^4, 4 \times 10^2, 10^4, 4 \times 10^2) \right), \end{aligned}$$

and the distribution of the models for a given aircraft state is taken as

$$[\pi_{k|k-1}(r|r')] = \begin{bmatrix} 0.8 & 0.1 & 0.1 \\ 0.8 & 0.1 & 0.1 \\ 0.8 & 0.1 & 0.1 \end{bmatrix}.$$

For simplicity it is assumed that the payload dynamics follow models $r = 1, 2, 3$.

3.3.4.1 Example 1

At time $k = 1$ an aircraft takes-off from $(-41, -51) \text{ km}$ and accelerates northwards. At time $k = 3$ a second aircraft takes-off from $(-51, 39) \text{ km}$ and accelerates towards $N80^\circ E$. A third aircraft takes-off from $(-9, 1) \text{ km}$ at time $k = 11$ and accelerates westwards. As the first aircraft initiates a counterclockwise turn at $k = 31$ a payload separates from the aircraft and continues northwards. At time $k = 44$ a payload separates from the second aircraft as it initiates a clockwise turn and continues along $S70^\circ E$.

Fig. 3.1 shows the true aircraft and payload trajectories in the horizontal plane. A 1-D view of these trajectories along with the sensor measurements is shown in Fig. 3.2. Simulations show that the PHD filter works well even when the simulated data is not generated from the same models used by the filter. The position estimates of the PHD filter in Fig. 3.3 demonstrate that the filter provides accurate tracking performance in clutter. Since at each sampling instant the number of targets is not known the filter occasionally exhibits false estimates. However, as shown these estimates do not propagate with time.

The mean absolute error in the number of targets and the probability of track loss (see [133] for a definition of these measures), estimated from 10^3 Monte Carlo runs, are shown in Fig. 3.4 for a position error radius of 50 m .

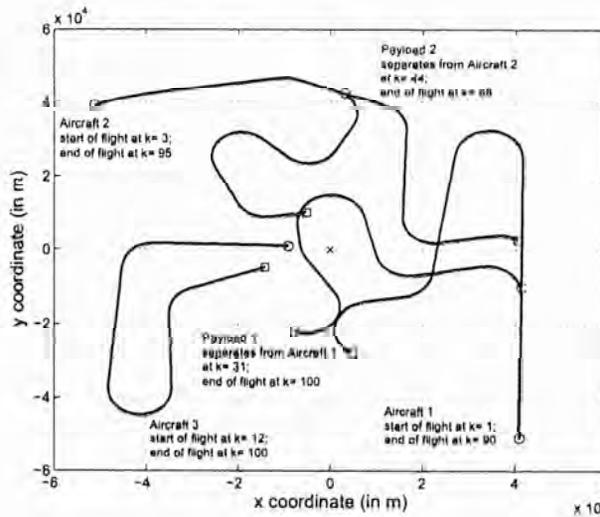


Figure 3.1: Aircraft and payload trajectories. ‘o’– locations of start of flight; ‘□’– locations of end of flight (‘x’– location of sensor).

3.3.4.2 Example 2

At time $k = 1$ three aircraft take-off simultaneously from the three airport locations. Aircraft 1 flies at a bearing of $N45^\circ W$ from $(-41, -51) km$, aircraft 2 flies eastwards from $(-51, 39) km$ and aircraft 3 flies at a bearing of $S45^\circ E$ from $(-9, 1) km$. Assuming all three aircraft exist at each sampling instant and no other targets appear in the surveillance region, the performance of the proposed PHD filter can be compared with that of the well-known IMMJPDA filter which tracks a fixed and known number of targets.

As indicated previously the PHD filter has a complexity of $\mathcal{O}(J_{k-1}|Z_k|)$ where J_{k-1} is the number of Gaussian components representing v_{k-1} for a fixed model r' at time $k - 1$ and $|Z_k|$ denotes the number of measurements at time k . Computationally efficient implementation of data association in JPDA has been the subject of much research. Exploiting parallel implementation, the column-recursive algorithm CR-JPDA [102] has a complexity of $\mathcal{O}(N|Z_k|^2 2^N)$ for N targets.

Fig. 3.5 shows the trajectories of the three aircraft. Fig. 3.6 (a) shows the mean absolute error in the estimate of the number of aircraft by the PHD filter. Fig. 3.6 (b) shows the probability of track loss at various clutter rates while a comparison of the

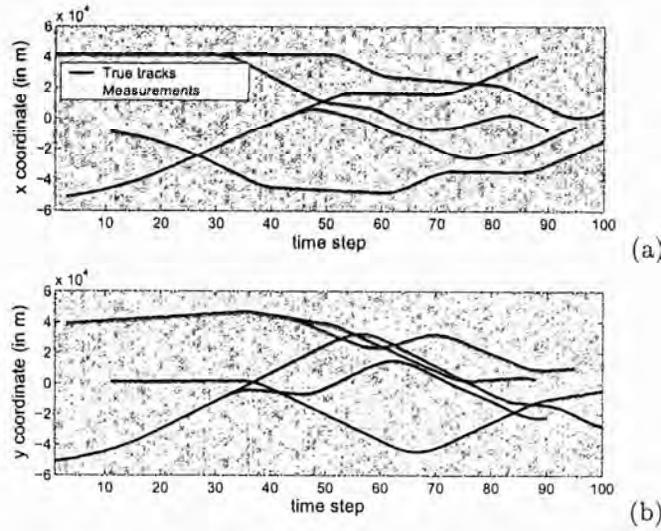


Figure 3.2: Measurement data and true target positions.

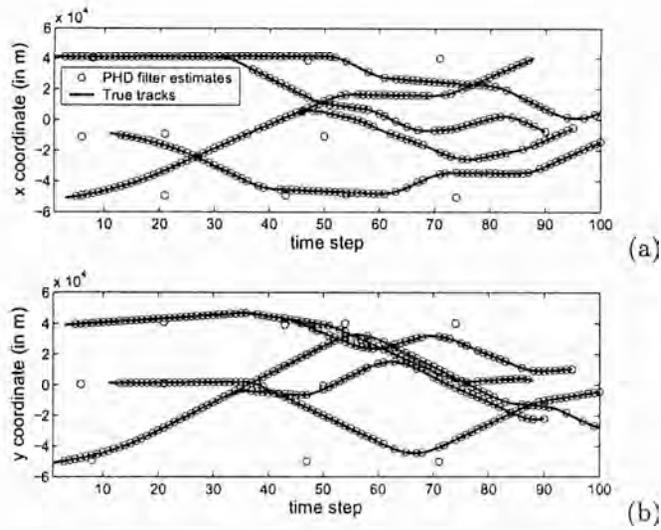


Figure 3.3: Position estimates of the Gaussian mixture PHD filter.

averaged CPU time involved at each step for the two filters is shown in Fig. 3.6 (c). Simulation results obtained from 10^3 Monte Carlo runs indicate that at any given clutter rate the tracking performance of the PHD filter is similar to that of the IMMJPDA filter at lower computational complexity.

Fig. 3.7 shows the tracking performance of the Gaussian mixture PHD filter versus the probability of target detection $p_{D,k}$ in the range $[0.7, 1.0]$ with a fixed clutter rate $\lambda_c = 3.47 \times 10^{-3} km^{-2}$. Fig. 3.7 (a) shows the mean absolute error in the estimate of the number of aircraft by the PHD filter. However, a comparison with the performance

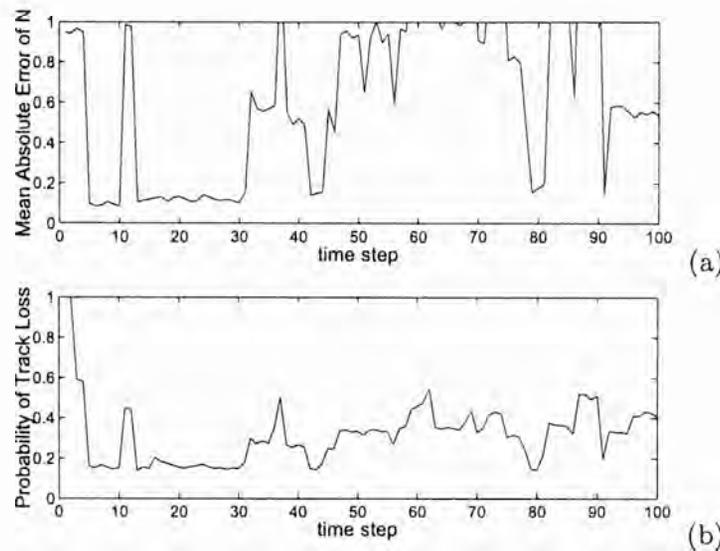


Figure 3.4: Mean absolute error of estimated number of targets and probability of track loss.

of the IMMJPDA filter is more intuitive. This result is remarkable because the PHD filter must resolve detection uncertainty in addition to the uncertainty in the number of targets and therefore is expected to perform poorly with increasing uncertainty in the number of targets due to increasing detection uncertainty. However, as shown in Figs. 3.7 (b) and (c) the tracking performance of the Gaussian mixture PHD filter is very similar to that of the IMMJPDA filter at a much lower computational cost.

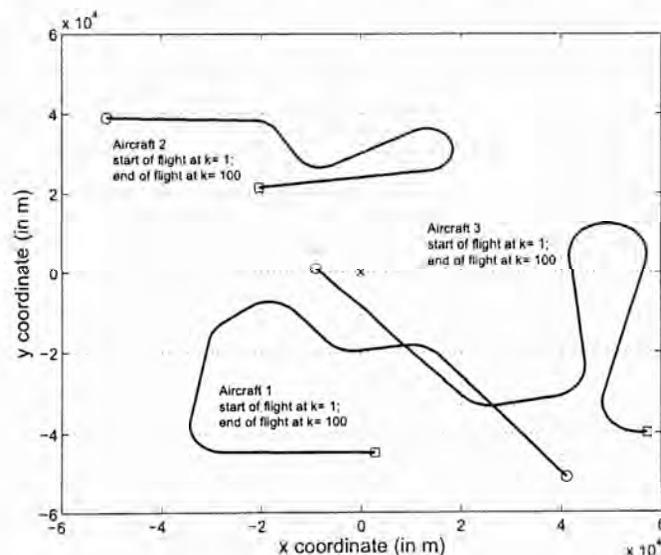


Figure 3.5: Aircraft and payload trajectories. 'o'- locations of start of flight; '□'- locations of end of flight ('x'- location of sensor).

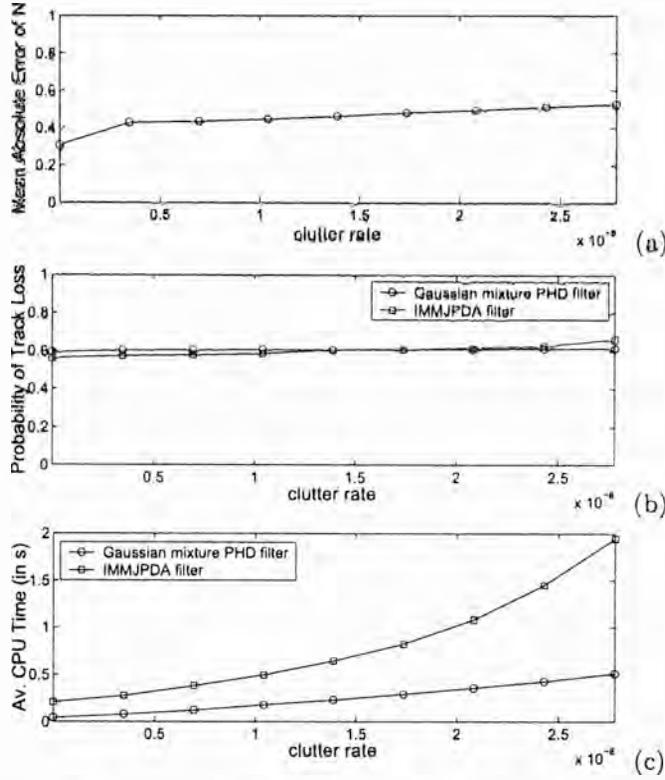


Figure 3.6: Tracking performance and computational complexity versus clutter rate for $p_{D,k} = 0.98$ and CPEP radius = 50 m.

3.4 The PHD filter for nonlinear GJM multi-target models

A JMS comprising of nonlinear models accommodates an even wider range of applications by providing a greater generality for modeling systems that switch between various models. Extension of the PHD filter for nonlinear models relaxes Assumption 3.4 and the state transition density and the observation likelihood take the form

$$f_{k|k-1}(\xi, r|\xi', r') = \mathcal{N}(\xi, F_{k-1}(\xi', r), Q_k(r)) t_{k|k-1}(r|r'), \quad (3.75)$$

$$g_k(z|\xi, r) = \mathcal{N}(z, H_k(\xi, r), R_k(r)), \quad (3.76)$$

where $F_{k-1}(\cdot, r)$ and $H_k(\cdot, r)$ denote nonlinear mappings parameterized by model r . The contribution of the intensity term due to the motion of the targets $v_{f,k|k-1}(\xi, r)$ to the predicted intensity at time k in (3.30) for a given prior intensity v_{k-1} is given by

$$v_{f,k|k-1}(\xi, r) = \sum_{r'} p_{S,k|k-1}(r) t_{k|k-1}(r|r') \int \mathcal{N}(\xi; F_{k-1}(\xi', r), Q_k(r)) v_{k-1}(\xi', r') d\xi'. \quad (3.77)$$

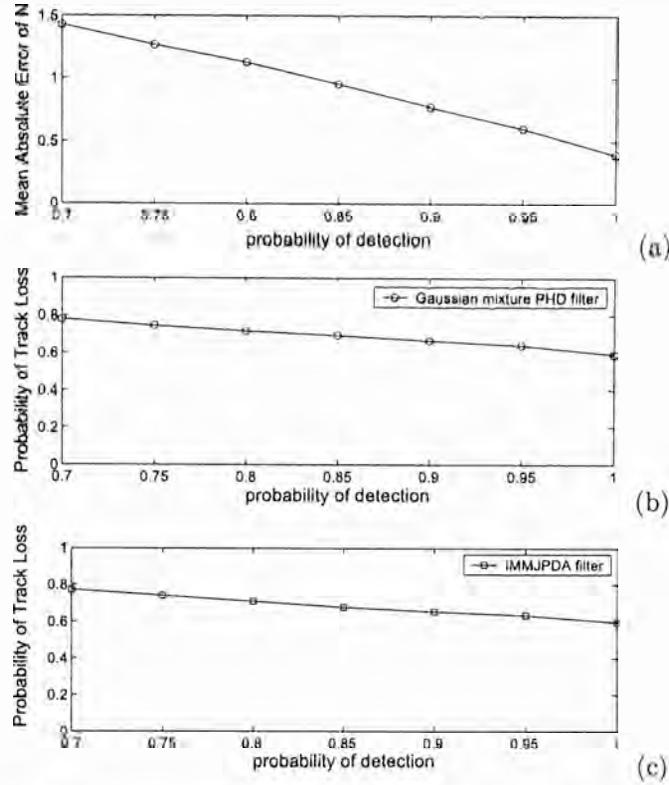


Figure 3.7: Tracking performance versus probability of detection for $\lambda_c = 3.47 \times 10^{-3} \text{ km}^{-2}$ and CPEP radius = 50 m.

Since $F_{k-1}(\cdot, r)$ is a nonlinear mapping, $v_{f,k|k-1}(\xi, r)$ does not admit a closed form. The predicted intensity $v_{k|k-1}(\xi, r)$ at time k is a weighted sum of various functions of ξ , many of which are non-Gaussian due to $v_{f,k|k-1}(\xi, r)$.

Similarly, the contribution of the intensity term due to the detected targets $v_{g,k}(\xi, r)$ to the posterior intensity at time k in (3.42) for a given predicted intensity of Gaussian mixture form is given by

$$v_{g,k}(\xi, r) = \frac{p_{D,k}(r) \mathcal{N}(z; H_k(\xi, r)) v_{k|k-1}(\xi, r)}{\kappa_k(z) + \sum_{r'} p_{D,k}(r') \int \mathcal{N}(z; H_k(\xi', r')) v_{k|k-1}(\xi', r') d\xi'}. \quad (3.78)$$

Since $H_k(\cdot, r)$ is nonlinear, $v_{g,k}(\xi, r)$ does not admit a closed form and the posterior intensity $v_k(\xi, r)$ at time k comprises of non-Gaussian components due to $v_{g,k}(\xi, r)$. At present there exists no tractable analytic method for tracking multiple targets with nonlinear jump Markov dynamics. In this section an analytic approximation of the PHD recursion is presented.

In single target filtering, analytic approximations of the nonlinear Bayes filter in-

clude the extended Kalman filter (EKF) [59, 44] and the unscented Kalman filter (UKF) [62, 63]. The EKF approximates the posterior density by a Gaussian, which is propagated in time by applying Kalman recursions to local linearizations of the (nonlinear) mappings $F_{k-1}(\cdot, r)$ and $H_k(\cdot, r)$. The UKF also approximates the posterior density by a Gaussian, but instead of using the linearized model, it computes the Gaussian approximation of the posterior density at the next time step using the unscented transform (see the discussion on nonlinear filtering in Section 2.5). In the sequel linear approximations of $F_{k-1}(\cdot, r)$ and $H_k(\cdot, r)$ using the UKF are discussed and the performance of the PHD filter is demonstrated for nonlinear models through a simulation example.

Consider the nonlinear mapping $F_{k-1}(\cdot, r)$ evaluated in p points $(\zeta_{k-1,i}(r), y_{k-1,i}(r))$, $i = 0, \dots, p$ around $\bar{m}_{k-1}(r)$ with covariance $P_{k-1}(r)$ where $y_{k-1,i}(r) = F_{k-1}(\zeta_{k-1,i}(r), r)$, $i = 0, \dots, p$ such that

$$\begin{aligned}\bar{m}_{k-1}(r) &= \frac{1}{p+1} \sum_{i=0}^p \zeta_{k-1,i}(r), \\ P_{k-1}(r) &= \frac{1}{p+1} \sum_{i=0}^p (\zeta_{k-1,i}(r) - \bar{m}_{k-1}(r)) (\zeta_{k-1,i}(r) - \bar{m}_{k-1}(r))^T.\end{aligned}\quad (3.79)$$

From (2.62) the statistical linear regression of $F_{k-1}(\cdot, r)$ around $\bar{m}_{k-1}(r)$ is the linear approximation $A_{k-1}(r)m + b_{k-1}(r)$ with

$$A_{k-1}(r) = P_{\zeta y, k-1}^T(r) P_{k-1}^{-1}(r), \quad (3.80)$$

$$b_{k-1}(r) = \bar{y}_{k-1}(r) - A_{k-1}(r) \bar{m}_{k-1}(r), \quad (3.81)$$

where $\bar{y}_{k-1}(r)$ can be computed like (2.64) along with the covariance $P_{y, k-1}(r)$ as

$$\begin{aligned}\bar{y}_{k-1}(r) &= \frac{1}{p+1} \sum_{i=0}^p y_{k-1,i}(r), \\ P_{y, k-1}(r) &= \frac{1}{p+1} \sum_{i=0}^p (y_{k-1,i}(r) - \bar{y}_{k-1}(r)) (y_{k-1,i}(r) - \bar{y}_{k-1}(r))^T,\end{aligned}\quad (3.82)$$

and the cross-covariance term $P_{\zeta y, k-1}(r)$ in (3.80) as (2.65)

$$P_{\zeta y, k-1}(r) = \frac{1}{p+1} \sum_{i=0}^p (\zeta_{k-1,i}(r) - \bar{m}_{k-1}(r)) (y_{k-1,i}(r) - \bar{y}_{k-1}(r))^T. \quad (3.83)$$

Let $e_{k|k-1,i}(r) = y_{k-1,i}(r) - (A_{k-1}(r)\zeta_{k-1,i} + b_{k-1}(r))$ and $e_{k|k-1}(r)$ be the error in the approximation, then the covariance of the error is

$$P_{e,k-1}(r) = \frac{1}{p+1} \sum_{i=0}^p \|e_{k|k-1,i}(r)\|^2 = P_{y,k-1}(r) - A_{k-1}(r)P_{k-1}(r)A_{k-1}^T(r). \quad (3.84)$$

Admitting the following approximation in (3.75)

$$\mathcal{N}(\xi; F_{k-1}(\xi', r), Q_k(r)) \approx \mathcal{N}(\xi, A_{k-1}(r)\xi' + b_{k-1}(r), P_{e,k-1}(r) + Q_k(r)), \quad (3.85)$$

Lemma 3.7 can be applied in (3.77) to obtain $v_{f,k|k-1}$ expressed in Gaussian mixture form.

Similarly, $H_k(\cdot, r)$ can be evaluated in p points $(\zeta_{k|k-1,j}(r), \varphi_{k,j}(r))$, $j = 0, \dots, p$ around $\bar{m}_{k|k-1}(r)$ with covariance $P_{k|k-1}(r)$ where $\varphi_{k,j}(r) = H_k(\zeta_{k|k-1,j}(r), r)$, $j = 0, \dots, p$ such that

$$\begin{aligned} \bar{m}_{k|k-1}(r) &= \frac{1}{p+1} \sum_{j=0}^p \zeta_{k|k-1,j}(r), \\ P_{k|k-1}(r) &= \frac{1}{p+1} \sum_{j=0}^p (\zeta_{k|k-1,j}(r) - \bar{m}_{k|k-1}(r)) (\zeta_{k|k-1,j}(r) - \bar{m}_{k|k-1}(r))^T, \end{aligned} \quad (3.86)$$

and the statistical linear regression of $H_k(\cdot, r)$ around $\bar{m}_{k|k-1}(r)$ is the approximation $C_k(r)m + d_k(r)$. Using (2.62),

$$C_k(r) = P_{\zeta\varphi,k}^T(r)P_{k|k-1}^{-1}(r), \quad (3.87)$$

$$d_k(r) = \bar{\varphi}_k(r) - C_k(r)\bar{m}_{k|k-1}(r), \quad (3.88)$$

where the first and second-order moment terms can be computed like (2.64)-(2.65),

$$\begin{aligned} \bar{\varphi}_k(r) &= \frac{1}{p+1} \sum_{j=0}^p \varphi_{k,j}(r), \\ P_{\varphi,k}(r) &= \frac{1}{p+1} \sum_{j=0}^p (\varphi_{k,j}(r) - \bar{\varphi}_k(r)) (\varphi_{k,j}(r) - \bar{\varphi}_k(r))^T, \\ P_{\zeta\varphi,k}(r) &= \frac{1}{p+1} \sum_{j=0}^p (\zeta_{k|k-1,j}(r) - \bar{m}_{k|k-1}(r)) (\varphi_{k,j}(r) - \bar{\varphi}_k(r))^T. \end{aligned} \quad (3.89)$$

Let $e_{k,j}(r) = \varphi_{k,j}(r) - (C_k(r)\zeta_{k|k-1,i} + d_k(r))$ and $e_k(r)$ be the error in the approximation, then the covariance of the error is

$$P_{e,k}(r) = \frac{1}{p+1} \sum_{j=0}^p \|e_{k,j}(r)\|^2 = P_{\varphi,k}(r) - C_k(r)P_{k|k-1}(r)C_k^T(r). \quad (3.90)$$

Admitting the following approximation in (3.76)

$$\mathcal{N}(z; H_k(\xi, r), R_k(r)) \approx \mathcal{N}(z, C_k(r)\xi + d_k(r), P_{e,k}(r) + R_k(r)), \quad (3.91)$$

Lemmas 3.7 and 3.8 can be applied in (3.78) to obtain $v_{g,k}$ in Gaussian mixture form.

Note that for nonlinear jump Markov spontaneous birth and spawn models each non-Gaussian constituent function of the mixture models can be approximated by a Gaussian using the linear approximation method described above. The expressions for the PHD recursion are notationally cumbersome and therefore omitted.

3.4.1 Simulation results

In this subsection the performance of the proposed PHD filter is demonstrated for nonlinear Gaussian jump Markov models. Assuming the turn rate is not a known constant the maneuver model becomes a nonlinear one. Augmenting the state vector to estimate the turn rate, the kinematic state of the aircraft is defined as $\xi = (p_x, \dot{p}_x, p_y, \dot{p}_y, \omega)^T$.

The motion models are as follows. Model $r = 1$ is a co-ordinated turn model with a known turn rate of $0^\circ s^{-1}$ and standard deviation of process noise, $\sigma_{v_1} = 5m s^{-2}$. Model $r = 2$ is a co-ordinated turn model with an unknown turn rate given by

$$F_{k-1}(\omega, r=2) = \begin{bmatrix} 1 & \frac{\sin \omega T}{\omega} & 0 & -\frac{1-\cos \omega T}{\omega} & 0 \\ 0 & \cos \omega T & 0 & -\sin \omega T & 0 \\ 0 & \frac{1-\cos \omega T}{\omega} & 1 & \frac{\sin \omega T}{\omega} & 0 \\ 0 & \sin \omega T & 0 & \cos \omega T & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad Q_k(r=2) = \sigma_{v_2}^2 \begin{bmatrix} \frac{T^4}{4} & \frac{T^3}{2} & 0 & 0 & 0 \\ \frac{T^3}{2} & T^2 & 0 & 0 & 0 \\ 0 & 0 & \frac{T^4}{4} & \frac{T^3}{2} & 0 \\ 0 & 0 & \frac{T^3}{2} & T^2 & 0 \\ 0 & 0 & 0 & 0 & T^2 \end{bmatrix},$$

and a process noise standard deviation of $10m s^{-2}$ and $0.5^\circ s^{-2}$ for the linear and turn portions respectively. The Markovian transition probability matrix is taken as

$$[t_{k|k-1}(r|r')] = \begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix}$$

Aircraft are observed by a sensor providing bearing and range measurements in the region $[-\pi, \pi] \text{ rad} \times [0, 60] \text{ km}$. The measurements are given by

$$z = \begin{bmatrix} \arctan(p_y/p_x) \\ \sqrt{p_x^2 + p_y^2} \end{bmatrix} + \epsilon_k \quad (3.92)$$

where $\epsilon_k \sim \mathcal{N}(\cdot; 0, R_k)$ with $R_k = \text{diag}([\sigma_\theta^2, \sigma_r^2])$, $\sigma_\theta = (\pi/180) \text{ rad s}^{-1}$ and $\sigma_r = 10 \text{ m}$. The average number of clutter returns per unit volume is $\lambda_c = 1.326 \times 10^{-1} (\text{rad km})^{-1}$.

The models for the births and spawnings are described as follows. The surveillance region includes three airport locations at $(40, -50) \text{ km}$, $(-50, 40) \text{ km}$ and $(-10, -10) \text{ km}$. The intensity of the Poisson RFS of spontaneous births is given by

$$\gamma_k(\xi, r) = 0.1\pi_k(r) \left(\mathcal{N}(\xi; m_\gamma^{(1)}, P_\gamma) + \mathcal{N}(\xi; m_\gamma^{(2)}, P_\gamma) + \mathcal{N}(\xi; m_\gamma^{(3)}, P_\gamma) \right),$$

with

$$\begin{aligned} m_\gamma^{(1)} &= (4 \times 10^4, 0, -5 \times 10^4, 0, 0)^T, \\ m_\gamma^{(2)} &= (-5 \times 10^4, 0, 4 \times 10^4, 0, 0)^T, \\ m_\gamma^{(3)} &= (-1 \times 10^4, 0, -1 \times 10^4, 0, 0)^T, \\ P_\gamma &= \text{diag}((10^6, 10^4, 10^6, 10^4, 10^{-8})), \end{aligned}$$

and the distribution of the models at birth is taken as $[\pi_k(r)] = (0.8, 0.2)$. The intensity of the Poisson RFS of spawn births is given by

$$\begin{aligned} \beta_{k|k-1}(\xi, r|\xi', r') &= 0.05\pi_{k|k-1}(r|r')\mathcal{N}(\xi; \xi', Q_\beta), \\ Q_\beta &= \text{diag}((10^4, 4 \times 10^2, 10^4, 4 \times 10^2, 10^{-8})), \end{aligned}$$

and the distribution of the models for a given aircraft state is taken as

$$[\pi_{k|k-1}(r|r')] = \begin{bmatrix} 0.8 & 0.2 \\ 0.8 & 0.2 \end{bmatrix}.$$

The settings for all other parameters are identical to those in Section 3.3.4.

At time $k = 1$ an aircraft takes-off from $(-41, -51) \text{ km}$ and accelerates northwards. At time $k = 31$ the aircraft executes a clockwise turn through 45° at 1° s^{-1} . 30 s

later the aircraft executes a $2^\circ s^{-1}$ counterclockwise turn. The aircraft then executes a $1^\circ s^{-1}$ clockwise turn at time $k = 70$. At time $k = 3$ a second aircraft takes-off from $(-51, 39) km$ and accelerates at a bearing of $N80^\circ E$. The aircraft executes two clockwise turns at $1^\circ s^{-1}$ and $2^\circ s^{-1}$ and flies at a heading of $S60^\circ W$ for 55 s before executing a 90° counterclockwise turn at $2^\circ s^{-1}$. A third aircraft takes-off from $(-9, -11) km$ at time $k = 12$ and accelerates along the initial heading of $S80^\circ W$. At time $k = 35$ the aircraft performs a 180° counterclockwise maneuver at $1^\circ s^{-1}$ followed by a sequence of clockwise and counterclockwise maneuvers at $2^\circ s^{-1}$. Two payloads separate from Aircraft 1 and Aircraft 2 at time $k = 31$ and $k = 56$ respectively and continue until $k = 100$.

Fig. 3.8 shows the true trajectories in the horizontal plane. As shown in Fig. 3.9 the proposed PHD filter provides reasonably accurate position estimates at most times. Fig. 3.10 shows the mean absolute error in the number of targets and the probability of track loss for a position error radius of 50 m estimated from 10^3 Monte Carlo runs.

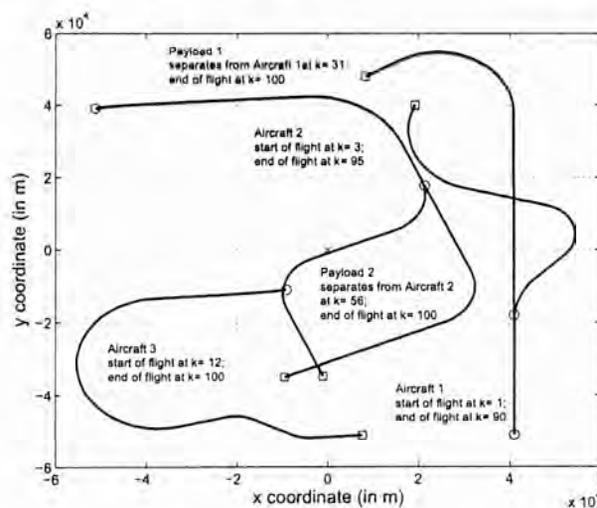


Figure 3.8: Aircraft and payload trajectories. ‘o’– locations of start of flight; ‘□’– locations of end of flight (‘x’– location of sensor).

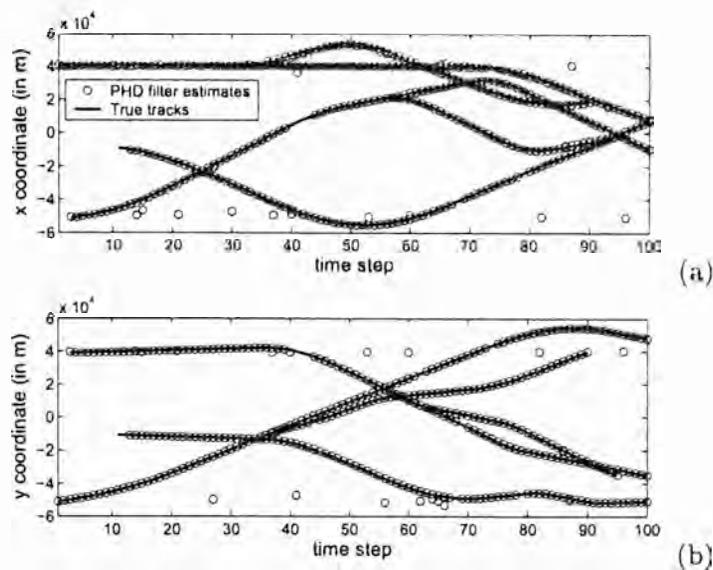


Figure 3.9: Position estimates of the Gaussian mixture PHD filter.

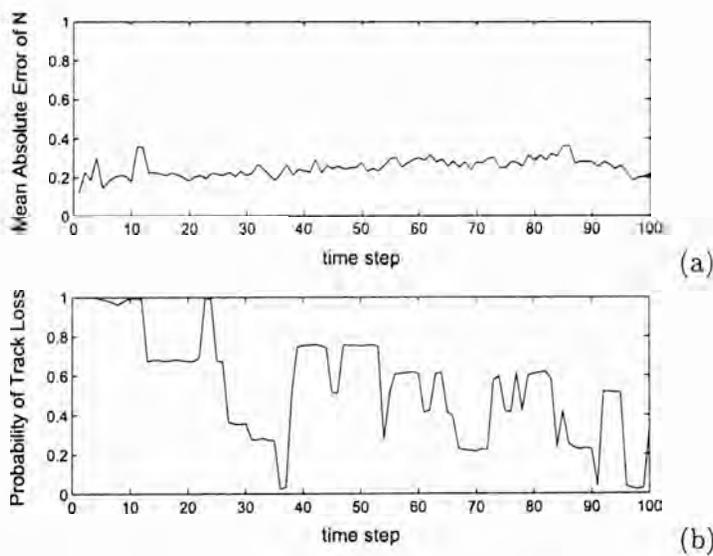


Figure 3.10: Mean absolute error of estimated number of targets and probability of track loss.

Chapter 4

Nonlinear filtering based on LFT modeling

For nonlinear state space model involving random variables with arbitrary probability distributions, the state estimation given a sequence of observations is based on an appropriate criterion such as the minimum mean square error (MMSE) (see Section 2.1). This leads to linear approximation in the state space of the extended Kalman filter (EKF) [59, 44] and the unscented Kalman filter (UKF) [62, 63], which work reasonably well only for mildly nonlinear systems. A Bayesian filtering technique is proposed based on the MMSE criterion in the framework of the virtual linear fractional transformation (LFT) model [143], which is characterized by a linear part and a simple nonlinear structure in the feedback loop. The LFT is an equivalent representation for smooth mappings (differentiable in any order), so the virtual LFT model is amenable to a wide range of nonlinear systems. Simulation results demonstrate that the proposed filtering technique gives better approximation and tracking performance than standard methods like the UKF. Furthermore, for highly nonlinear systems where UKF diverges, the LFT model estimates the conditional mean with reasonable accuracy.

This chapter is organized as follows: Section 4.1 gives some background on analytical approximation techniques for nonlinear filtering and transformation models applied in nonlinear control. In Section 4.2 the LFT model is discussed. An analytical solution to the Bayes recursion is then derived based on this model. Simulation results are given in Section 4.3 to analyze the performance of the proposed approach with the UKF. The

general solution to the Bayes recursion is then given in Section 4.4 using the nonlinear fractional transformation (NFT) model which accommodates the most general class of nonlinear mapping. This is demonstrated through a simulation example.

4.1 Introduction

The linear minimum mean square error (LMMSE) estimate of a random variable conditional on an event can be determined if the unconditional means of the random variable and the conditioning random variable, the covariance of the conditioning random variable and the joint second-order moment or cross-covariance of the two random variables are computable. Based on this, the Bayes filter can be applied in a recursive fashion to estimate the state of a system conditional on a sequence of observations as it evolves with time. Under linearity assumption on the state space model, the estimate is unbiased and admits a closed form solution. However, the class of linear state space models encompasses only a small subset of real systems. For the general class of real systems modeled by nonlinear space representation, the Bayes filter concedes an approximation in order to estimate the state.

In the literature there are two standard approaches to approximate the conditional expectation of the state. The extended Kalman filter (EKF) applies a local linearization to the nonlinear mapping around the state estimate. This method is predicated on the weak premise that the estimate lies in the neighborhood of the global trajectory. As a result, stability of the filter and convergence of the estimate are not guaranteed. This is demonstrated by a simple example in [75] where the EKF fails to converge. The unscented Kalman filter (UKF) [62, 63] on the other hand, applies the unscented transformation [62] which uses the statistical linear regression technique [77, 78] to approximate the moments of random variables. The conditional mean obtained using the UKF has a higher order accuracy than the estimate given by the EKF. This has been substantiated by empirical studies on the EKF and UKF showing that in most

applications the UKF gives better approximation [62, 63, 77, 33, 115]. Despite the advantage of UKF over EKF, the two approaches work reasonably well under mildly nonlinear conditions only.

Over the past few years sequential Monte Carlo (SMC) methods have attracted attention for nonlinear Bayesian filtering applications [40, 39]. These methods approximate the filtering distribution by a set of samples drawn from a proposal distribution. Under the assumption that the proposal distribution includes the region of support of the filtering distribution, SMC methods give better approximation than the linear approximation techniques mentioned above. In practice, a sufficiently large number of samples is needed. It is only in the limit that the number of samples approaches infinity that the simulation-based methods guarantee convergence of the estimate to the optimal Bayes solution.

In nonlinear control, exact feedback linearization is to transform a nonlinear control system into an equivalent linear one through a variable change [55, 56, 70]. Its validity is highly restrictive of local nature and its applicability is for a limited class of nonlinear systems. On the other hand, the *linear fractional transformation* (LFT) method (see e.g., [143, 4, 3] and the references therein) is extensively employed in \mathcal{H}_2 and \mathcal{H}_∞ gain-scheduling based control and filtering to represent nonlinear plants, whereas the uncertainty appears as a LFT (see e.g., [129, 128, 127, 25] and the references therein). Unlike feedback linearization, the LFT approach gives an equivalent representation for a very wide class of nonlinear systems including smooth mappings (differentiable in any order) and those involving complex fractional terms [143, 129, 128]. The LFT representation comprises of a linear model and a simple nonlinear structure in the feedback loop with sparse representation. This structure offers two advantages: firstly, any approximation involved is localized to the feedback loop only. In [141] static nonlinearities appearing in the feedback loop are approximated using a local linearization about the input trajectory yielding a linear-time varying system. Secondly, the highly uncorre-

lated nature of the nonlinear structure gives better approximation of the second-order moments. Moreover, the flexible representation of the LFT system is amenable to different approximation strategies using the linear regression technique which is not obvious using the state space model representation. Based on this, a Bayesian filtering technique is presented for the most general class of nonlinear systems by transforming the state space model into an exact equivalent LFT model. By applying the unscented transformation in the feedback loop only a closed form solution is derived to estimate the conditional mean of the state. The simulation results show that the proposed filtering approach gives a better tracking performance than the UKF in terms of tracking error. There is a case when the proposed approach can track an object which cannot be tracked using the UKF. For the class of nonlinear mappings that cannot be expressed in the fractional form, the *nonlinear fractional transformation* (NFT) model has been proposed [129, 52, 51, 53]. The proposed filtering technique generalizes naturally to the NFT case. Since the LFT model is widely accepted as a tool to express nonlinear systems, the discussion in this chapter in the context of the LFT model and then the results are presented under more general settings for the NFT system.

4.2 LFT for linear filtering of nonlinear models

This section presents a solution to the nonlinear filtering problem which averts the linearization of the state space model by using the LFT model. The LFT model is discussed in Section 4.2.1. The estimation and prediction steps of the recursive Bayes filter are then derived in Section 4.2.2.

4.2.1 The linear fractional transformation (LFT) model

From robust control theory it is known that any nonlinear mapping f in $y = f(x)$, differentiable at any order admits an equivalent representation known as the LFT model

[143, 120, 128],

$$\begin{bmatrix} y \\ y_\Delta \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x \\ w_\Delta \end{bmatrix}, \quad (4.1)$$

$$w_\Delta = \Delta(x)y_\Delta, \quad (4.2)$$

where $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{m \times n_\Delta}$, $C \in \mathbb{R}^{n_\Delta \times n}$ and $D \in \mathbb{R}^{n_\Delta \times n_\Delta}$. The auxiliary variables $w_\Delta \in \mathbb{R}^{n_\Delta}$ and $y_\Delta \in \mathbb{R}^{n_\Delta}$ introduced are related via the feedback connection $\Delta(x)$ which takes the form $\Delta(x) = \sum_{i=1}^n \Delta_i x(i)$ with $x(i)$ as the i -th element of vector x .

The LFT system of (4.1)-(4.2) can be easily seen by its compact expression

$$y = (A + B\Delta(x)(I - D\Delta(x))^{-1}C)x, \quad (4.3)$$

where $\Delta(x)$ enters the relation in a highly nonlinear fashion. Using either the local linearization technique (2.59) or the statistical linear regression method (2.62) to approximate y in this manner gives an approximation that is equivalent to linearizing the nonlinear mapping f . On the other hand, the representation (4.1)-(4.2) is nonlinear in the feedback path only. Under this representation, an approximation is localized to the feedback path for estimation of the auxiliary random variable w_Δ in (4.2). Given the first two moments of x as \bar{x} and R_x , the regression points $x_i, i = 0, \dots, p$ are chosen by (2.63). Define the regression points $w_{\Delta i} = \Delta(x_i)y_{\Delta i}$, where

$$y_{\Delta i} = Cx_i + D\bar{w}_\Delta, \quad (4.4)$$

and $\bar{w}_\Delta \approx \mathbf{E}(w_\Delta)$ is

$$\bar{w}_\Delta = (I - \bar{\Delta}D)^{-1} \left(\frac{1}{p+1} \sum_{i=0}^p \Delta(x_i)Cx_i \right), \quad (4.5)$$

with

$$\bar{\Delta} = \frac{1}{p+1} \sum_{i=0}^p \Delta(x_i). \quad (4.6)$$

Accordingly, the covariance of w_Δ and the cross-covariance with x are computed like (2.64)-(2.65):

$$R_\Delta = \frac{1}{p+1} \sum_{i=0}^p (w_{\Delta i} - \bar{w}_\Delta)(w_{\Delta i} - \bar{w}_\Delta)^T, \quad (4.7)$$

$$R_{\Delta x} = \frac{1}{p+1} \sum_{i=0}^p (w_{\Delta i} - \bar{w}_\Delta)(x_i - \bar{x})^T. \quad (4.8)$$

One can see that the approximation of the first and second-order moments of w_Δ in (4.5)-(4.8) averts the linearization of (4.3) which gives poor approximation for highly nonlinear models. The proposed approximation is expected to work well even in the case of highly nonlinear systems due to the simpler nonlinear structure in the LFT model.

Now, for a random variable $y = f(x) + \tilde{B}w$ which depends on x with mean \bar{x} and covariance R_x and $w \sim \mathcal{N}(\cdot; 0, R_w)$, independent of x , the equivalent LFT representation takes the form

$$y = Ax + \tilde{B}w + Bw_\Delta, \quad y_\Delta = Cx + Dw_\Delta, \quad w_\Delta = \Delta(x)y_\Delta, \quad (4.9)$$

where $\tilde{B} \in \mathbb{R}^{m \times n_w}$. The expectation of y is still

$$\bar{y} = A\bar{x} + B\bar{w}_\Delta,$$

where \bar{w}_Δ is defined from (4.5). The covariance of y and its the cross-covariance with x are

$$\begin{aligned} R_y &= AR_x A^T + \tilde{B}R_w \tilde{B}^T + BR_\Delta B^T + \\ &\quad AR_{\Delta x}^T B^T + BR_{\Delta x} A^T, \end{aligned} \quad (4.10)$$

$$R_{yx} = AR_x + BR_{\Delta x}, \quad (4.11)$$

respectively, with R_Δ and $R_{\Delta x}$ defined from (4.7)-(4.8).

Based on the LFT representation of the state space model (2.1)-(2.2) and the procedure outlined for the approximation of the statistical moments of w_Δ , the estimation and prediction steps of the recursive Bayes filter are derived in the following subsection.

4.2.2 Recursive Bayes estimation

The nonlinear state space model (2.1)-(2.2) can alternatively be expressed in the LFT format

$$x_{k+1} = A_k x_k + B_{1,k} w_k + B_{2,k} w_{\Delta k}, \quad (4.12)$$

$$z_k = C_{1,k} x_k + D_{11,k} v_k + D_{12,k} w_{\Delta k}, \quad (4.13)$$

$$z_{\Delta k} = C_{2,k} x_k + D_{22,k} w_{\Delta k}, \quad (4.14)$$

$$w_{\Delta k} = \Delta(x_k) z_{\Delta k}. \quad (4.15)$$

Here $A_k \in \mathbb{R}^{n \times n}$, $B_{1,k} \in \mathbb{R}^{n \times n_w}$, $B_{2,k} \in \mathbb{R}^{n \times n_{\Delta}}$, $C_{1,k} \in \mathbb{R}^{m \times n}$, $D_{11,k} \in \mathbb{R}^{m \times n_v}$, $D_{12,k} \in \mathbb{R}^{m \times n_{\Delta}}$, $C_{2,k} \in \mathbb{R}^{n_{\Delta} \times n}$, and $D_{22,k} \in \mathbb{R}^{n_{\Delta} \times n_{\Delta}}$. $w_{\Delta k} \in \mathbb{R}^{n_{\Delta}}$ and $z_{\Delta k} \in \mathbb{R}^{n_{\Delta}}$ denote auxiliary variables introduced to model the feedback connection which takes the form indicated above. The noise sequences $\{w_k\}$ and $\{v_k\}$ are assumed mutually uncorrelated and uncorrelated to the state x_k (and thus uncorrelated to the auxiliary variable $w_{\Delta k}$ too).

Denote by $x_{k,i}$, $i = 0, \dots, p$ the i -th regression point of x_k . Suppose the estimate of the state x_{k-1} at time $k-1$ conditional on the history of observations Z_{k-1} is m_{k-1} with covariance P_{k-1} . Then, by applying the technical results of the previous subsection, the prediction and estimation steps of the Bayes recursion in the proposed approach are given by the following.

Proposition 4.1. *The expectation of the predicted state at time k conditional on the data up to time $k-1$ is accepted as $m_{k|k-1}$ with covariance $P_{k|k-1}$ where*

$$m_{k|k-1} = A_{k-1} m_{k-1} + B_{2,k-1} \bar{w}_{\Delta k-1}, \quad (4.16)$$

$$P_{k|k-1} = A_{k-1} P_{k-1} A_{k-1}^T + B_{1,k-1} Q_{k-1} B_{1,k-1}^T + B_{2,k-1} R_{\Delta k-1} B_{2,k-1}^T + A_{k-1} R_{\Delta x, k-1}^T B_{2,k-1}^T + B_{2,k-1} R_{\Delta x, k-1} A_{k-1}^T, \quad (4.17)$$

with

$$\bar{\Delta}_{k-1} = \frac{1}{p+1} \sum_{i=0}^p \Delta(x_{k-1,i}), \quad (4.18)$$

$$\bar{w}_{\Delta k-1} = (I - \bar{\Delta}_{k-1} D_{22,k-1})^{-1} \left(\frac{1}{p+1} \sum_{i=0}^p \Delta(x_{k-1,i}) C_{2,k-1} x_{k-1,i} \right) \quad (4.19)$$

$$w_{\Delta k-1,i} = \Delta(x_{k-1,i}) (C_{2,k-1} x_{k-1,i} + D_{22,k-1} \bar{w}_{\Delta k-1}), \quad i = 0, \dots, p, \quad (4.20)$$

$$R_{\Delta k-1} = \frac{1}{p+1} \sum_{i=0}^p (w_{\Delta k-1,i} - \bar{w}_{\Delta k-1}) (w_{\Delta k-1,i} - \bar{w}_{\Delta k-1})^T, \quad (4.21)$$

$$R_{\Delta x,k-1} = \frac{1}{p+1} \sum_{i=0}^p (w_{\Delta k-1,i} - \bar{w}_{\Delta k-1}) (x_{k-1,i} - m_{k-1})^T. \quad (4.22)$$

Proposition 4.2. *Given the conditional mean and covariance of the predicted state as above, the required estimate of the state x_k conditional on Z_k is m_k with covariance P_k given by*

$$m_k = m_{k|k-1} + K_k(z_k - \eta_k), \quad (4.23)$$

$$P_k = P_{k|k-1} - K_k(C_{1,k}P_{k|k-1} + D_{12,k}R_{\Delta x,k|k-1}), \quad (4.24)$$

where

$$\eta_k = C_{1,k}m_{k|k-1} + D_{12,k}\bar{w}_{\Delta k|k-1}, \quad (4.25)$$

$$K_k = (P_{k|k-1}C_{1,k}^T + R_{\Delta x,k|k-1}^T D_{12,k}^T) (C_{1,k}P_{k|k-1}C_{1,k}^T + D_{11,k}R_k D_{11,k}^T + D_{12,k}R_{\Delta k|k-1}D_{12,k}^T + C_{1,k}R_{\Delta x,k|k-1}^T D_{12,k}^T + D_{12,k}R_{\Delta x,k|k-1}C_{1,k}^T)^{-1}, \quad (4.26)$$

with

$$\bar{\Delta}_{k|k-1} = \frac{1}{p+1} \sum_{i=0}^p \Delta(x_{k|k-1,i}), \quad (4.27)$$

$$\bar{w}_{\Delta k|k-1} = (I - \bar{\Delta}_{k|k-1} D_{22,k})^{-1} \left(\frac{1}{p+1} \sum_{i=0}^p \Delta(x_{k|k-1,i}) C_{2,k} x_{k|k-1,i} \right) \quad (4.28)$$

$$w_{\Delta k|k-1,i} = \Delta(x_{k|k-1,i}) (C_{2,k} x_{k|k-1,i} + D_{22,k} \bar{w}_{\Delta k|k-1}), \quad i = 0, \dots, p, \quad (4.29)$$

$$R_{\Delta k|k-1} = \frac{1}{p+1} \sum_{i=0}^p (w_{\Delta k|k-1,i} - \bar{w}_{\Delta k|k-1}) (w_{\Delta k|k-1,i} - \bar{w}_{\Delta k|k-1})^T, \quad (4.30)$$

$$R_{\Delta x,k|k-1} = \frac{1}{p+1} \sum_{i=0}^p (w_{\Delta k|k-1,i} - \bar{w}_{\Delta k|k-1}) (x_{k|k-1,i} - m_{k|k-1})^T. \quad (4.31)$$

One can see that Proposition 4.1 is similar to the Kalman prediction step with the addition of the terms involving moments of $w_{\Delta k-1}$. Similarly, Proposition 4.2 is the Kalman data update step with the addition of terms involving moments of $w_{\Delta k|k-1}$.

4.3 Simulation results

In this section simulation results are presented to demonstrate the performance of the proposed filtering technique using the LFT model. Example I considers a nonlinear state transition model of the third order and gives a comparison of the tracking performance with the UKF. The effectiveness of the proposed filtering method becomes evident in Example II which considers a highly nonlinear problem for which standard filtering approaches for nonlinear systems such as the EKF and the UKF fail to converge. In Example III a scenario from a multi-target tracking application is considered to estimate the unknown number of targets in the surveillance space and their state based on bearings and range information.

4.3.1 Example I

Consider a typical nonlinear autoregressive (AR) equation $q_{k+2} = -0.1q_{k+1} - q_k^3 + w_k$ with the noisy measurement $z_k = q_k + v_k$, which admits the following state-space equation formulation with the state $x_k = (x_k(1), x_k(2))^T = (q_k, q_{k+1})^T \in \mathbb{R}^2$

$$x_{k+1} = \begin{bmatrix} 0 & 1 \\ -x_k^2(1) & -0.1 \end{bmatrix} x_k + \begin{bmatrix} 0 \\ 1 \end{bmatrix} w_k, \quad (4.32)$$

$$z_k = x_k(1) + v_k, \quad (4.33)$$

where $w_k \sim \mathcal{N}(\cdot; 0, Q)$ with $Q = 0.04$ and $v_k \sim \mathcal{N}(\cdot; 0, R)$ with $R = 0.5$. The nonlinear state space model (4.32)-(4.33) can be represented in the LFT form (4.12)-(4.15) with

$$\begin{aligned} A_k &= \begin{bmatrix} 0 & 1 \\ 0 & -0.1 \end{bmatrix}, & B_{1,k} &= \begin{bmatrix} 0 \\ 1 \end{bmatrix}, & B_{2,k} &= \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}, \\ C_{1,k} &= \begin{bmatrix} 1 & 0 \end{bmatrix}, & D_{11,k} &= 1, & D_{12,k} &= 0_{1,2}, \\ C_{2,k} &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, & D_{22,k} &= \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \end{aligned} \quad (4.34)$$

where $0_{a,b}$ is the $a \times b$ zero matrix. The feedback connection has the simple structure $\Delta(x_k) = x_k(1)I_2$ with I_a as the identity matrix of dimension $a \times a$. The true trajectory of the state x_k for 50 time steps is shown in Fig. 4.1. Using $\bar{x}_0 = (0, 0)^T$ as the initial estimate of the state with covariance $R_{x,0} = I_2$, the estimates given by the proposed filter at each time step are shown in Fig. 4.2 along with the true states. In Fig. 4.3 the mean square error (MSE) in the estimates obtained from 10^2 Monte Carlo runs using the proposed filter and the UKF is shown. Fig. 4.3 (a) depicts that the error using the proposed method drops to 0.1 at $k = 5$ and remains below that obtained using the UKF. A similar trend observed in Fig. 4.3 (b) suggests that the proposed filter gives better tracking performance than the UKF.

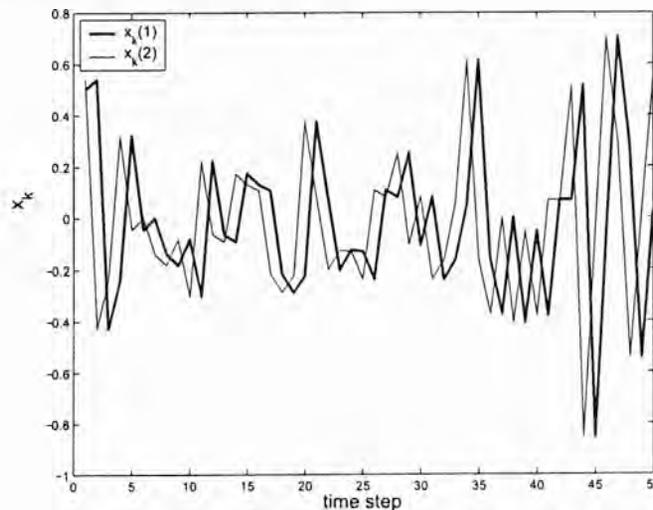


Figure 4.1: Trajectory of the state $x_k = (x_k(1), x_k(2))^T$.

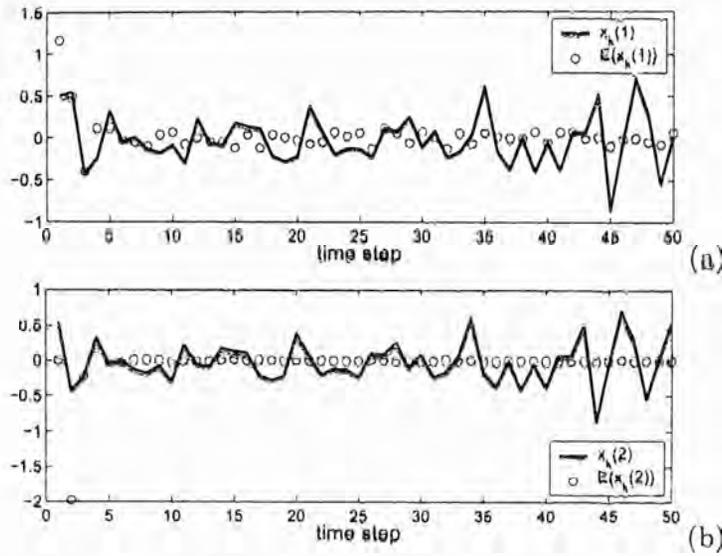


Figure 4.2: True trajectory and the estimate of the state $\mathbf{E}(x_k|Z_k)$ given by the proposed filter.

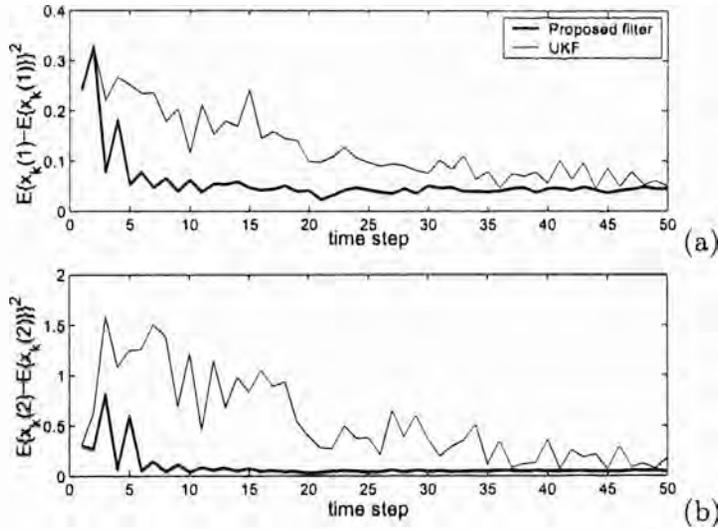


Figure 4.3: Mean square error (MSE) in the estimates using the proposed filter and the UKF.

4.3.2 Example II

Consider the highly nonlinear system [129]

$$x_{k+1} = (Q_0 + Q_1 x_k^3(1) + Q_2 x_k^3(2) + Q_3 x_k(1)x_k^2(2) + Q_4 x_k(1) + Q_5 x_k(2))x_k + B_k w_k, \tag{4.35}$$

$$z_k = 100 \begin{bmatrix} -1 & 1 \end{bmatrix} x_k + v_k, \tag{4.36}$$

where

$$\begin{aligned} Q_0 &= \begin{bmatrix} -0.7 & -1.0 \\ 0.1 & -0.5 \end{bmatrix}, \quad Q_1 = \begin{bmatrix} 0.3 & 0.2 \\ 0.1 & 0.2 \end{bmatrix}, \quad Q_2 = \begin{bmatrix} 0.2 & 0.1 \\ 0.2 & 0.3 \end{bmatrix}, \\ Q_3 &= \begin{bmatrix} 0.4 & 0.1 \\ 0.15 & 0.1 \end{bmatrix}, \quad Q_4 = \begin{bmatrix} 0.25 & 0.25 \\ 0.1 & 0.25 \end{bmatrix}, \quad Q_5 = \begin{bmatrix} 0.25 & 0 \\ 0.1 & 0.25 \end{bmatrix}, \end{aligned} \quad (4.37)$$

$B_k = [-2 \ 1]^T$, $w_k \sim \mathcal{N}(\cdot; 0, Q)$ with $Q = 0.01$ and $v_k \sim \mathcal{N}(\cdot; 0, R)$ with $R = 100$. The exact LFT representation can be constructed with

$$\begin{aligned} A_k &= \begin{bmatrix} -0.7 & -1.0 \\ 0.1 & -0.5 \end{bmatrix}, \quad B_{1,k} = B_k, \\ B_{2,k} &= \begin{bmatrix} 0.25 & 0 & 0.3 & 0 & 0 & 0.1 & 0.2 & 0.3 & 0 & 0.4 & 0.5 \\ 0.1 & 0 & 0.1 & 0.25 & 0 & 0.3 & 0.2 & 0.3 & 0 & 0.15 & 0.35 \end{bmatrix}, \\ C_{1,k} &= 100 \begin{bmatrix} -1 & 1 \end{bmatrix}, \quad D_{11,k} = 1, \quad D_{12,k} = 0_{1,11}, \\ C_{2,k}^T &= \begin{bmatrix} 1 & 0_{1,2} & 0 & 0_{1,6} & 1 \\ 0 & 0_{1,2} & 1 & 0_{1,6} & 0 \end{bmatrix}, \quad D_{22,k} = \begin{bmatrix} 0_{1,5} & 0_{1,5} & 0 \\ L & 0_5 & 0_{5,1} \\ M & N & 0_{5,1} \end{bmatrix}, \end{aligned} \quad (4.38)$$

where

$$L = \begin{bmatrix} I_2 & 0_{2,1} & 0_2 \\ 0_{1,2} & 0 & 0_{1,2} \\ 0_2 & 0_{2,1} & I_2 \end{bmatrix}, \quad M = \begin{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} & 0_{2,1} & \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \\ \begin{bmatrix} 1 & 0 \end{bmatrix} & 0 & \begin{bmatrix} 0 & 0 \end{bmatrix} \\ 0_2 & 0_{2,1} & 0_2 \end{bmatrix}, \quad N = \begin{bmatrix} 0_2 & 0_{2,1} & 0_2 \\ 0_{1,2} & 0 & 0_{2,1} \\ 0_2 & 0_{2,1} & \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \end{bmatrix},$$

and 0_a is the zero matrix of dimension $a \times a$. Using $1_{a,b}$ to denote the $a \times b$ matrix with entries one, the feedback connection is given by

$$\Delta(x_k) = \text{diag}(\{x_k(1)1_{1,3} \quad x_k(2)1_{1,4} \quad x_k(1) \quad x_k(2)1_{1,3}\}). \quad (4.39)$$

The trajectory of the state x_k as it evolves with time for 50 time steps is shown in Fig. 4.4. Using $\bar{x}_0 = (0, 0)^T$ as the estimate of x_0 at time $k = 0$ with covariance

$R_{x,0} = 0.25I_2$, Fig. 4.5 shows the estimate of x_k given by the proposed filter at each time step along with the true trajectory. In this example the EKF and the UKF break down so a comparison of the results with the standard approximation methods is not possible. The MSE obtained from 10^2 Monte Carlo runs using the proposed filter shown in Fig. 4.6 indicates that the proposed filter works reasonably well.

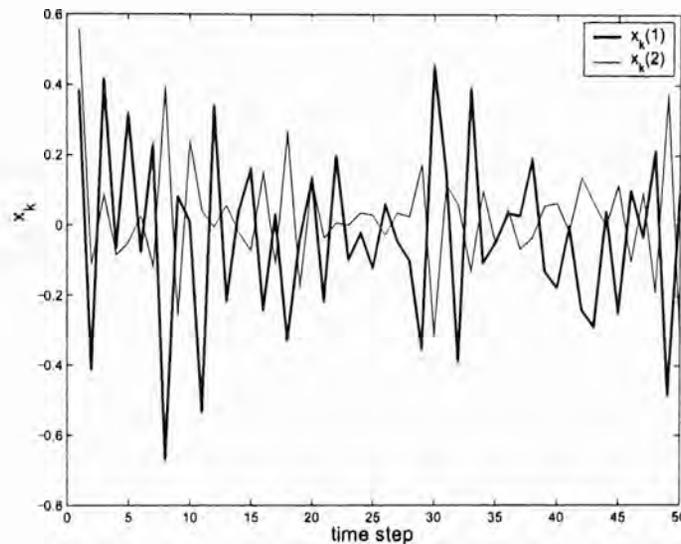


Figure 4.4: Trajectory of the state $x_k = (x_k(1), x_k(2))^T$.

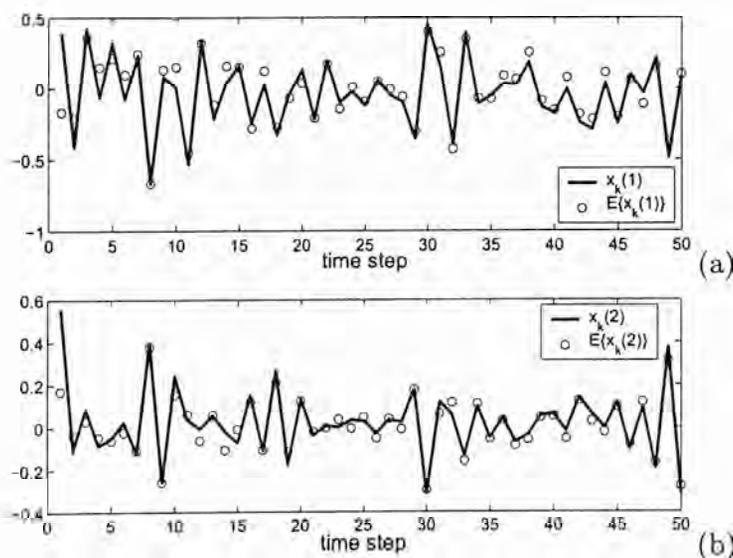


Figure 4.5: True trajectory and the estimate of the state $\mathbf{E}(x_k|Z_k)$ given by the proposed filter.

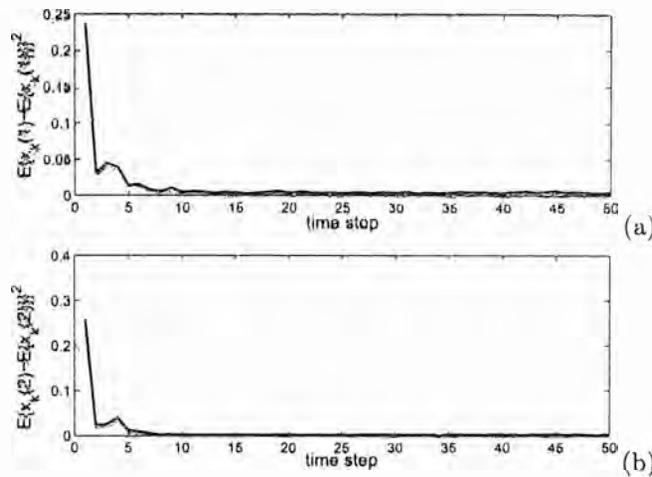


Figure 4.6: MSE in the estimates using the proposed filter

4.3.3 Example III

In this example the problem of multi-target filtering of random number of maneuvering targets in clutter presented in Section 3.4 using the proposed nonlinear JMS-PHD filter is revisited.

The aircraft are observed by a sensor located at $(-120, -40)$ km which provides bearing and range information in the region $[-\pi/4, \pi/4]$ rad \times $[0, 200]$ km. The measurements are given by

$$z_k = \begin{bmatrix} \arctan((p_{y,k} - p_{s,y}) / (p_{x,k} - p_{s,x})) \\ \sqrt{(p_{x,k} - p_{s,x})^2 + (p_{y,k} - p_{s,y})^2} \end{bmatrix} + \epsilon_k \quad (4.40)$$

where $(p_{s,x}, p_{s,y})$ denotes the sensor coordinates, $\epsilon_k = (\epsilon_{1,k}, \epsilon_{2,k})^T \sim \mathcal{N}(\cdot; 0, R_k)$ with $R_k = \text{diag}([\sigma_\theta^2, \sigma_r^2])$, $\sigma_\theta = (\pi/180)$ rad s $^{-1}$ and $\sigma_r = 20$ m. Let $z_k = (z_{1,k}, z_{2,k})^T$, then

$$\tilde{z}_{1,k} = \tan(z_{1,k}) = \frac{p_{y,k} - p_{s,y}}{p_{x,k} - p_{s,x}} + \tilde{v}_k. \quad (4.41)$$

Using the approximation $\tilde{v}_k \sim \mathcal{N}(\cdot; 0, \sigma_\theta^2)$, $\tilde{z}_{1,k}$ can alternatively be expressed in the LFT format

$$\tilde{z}_{1,k} = C_{1,k}\xi_k + D_{11,k}d_k + D_{12,k}w_{\Delta k}, \quad (4.42)$$

$$z_{\Delta k} = C_{2,k}\xi_k + D_{21,k}d_k + D_{22,k}w_{\Delta k}, \quad (4.43)$$

$$w_{\Delta k} = \Delta(x_k)z_{\Delta k}, \quad (4.44)$$

where $d_k = (\epsilon_{1,k}, 1)^T$,

$$\begin{aligned} C_{1,k} &= 0_{1,4}, & D_{11,k} &= \begin{bmatrix} 1 & p_{s,y}/p_{s,x} \end{bmatrix}, & D_{12,k} &= \frac{1}{p_{s,x}^2} \begin{bmatrix} p_{s,y} & -p_{s,x} \end{bmatrix}, \\ C_{2,k} &= 0_{1,4}, & D_{21,k} &= \begin{bmatrix} 0 & 1 \end{bmatrix}, & D_{22,k} &= \begin{bmatrix} 1/p_{s,x} & 0 \end{bmatrix}, \end{aligned} \quad (4.45)$$

and the feedback connection $\Delta(x_k)$ is the vector of vehicle coordinates $(p_{x,k}, p_{y,k})^T$.

For a given mode r , the single target dynamical and range measurement models are approximated by the linear Gaussian models (3.91) and (3.85) respectively as before.

The average number of clutter returns per unit volume is taken as $\lambda_c = \pi^{-1} (\text{rad km})^{-1}$.

For simplicity target spawning is not considered. Consider a scenario where the surveillance region includes the five airport locations at $(-20, -20) \text{ km}$, $(10, 20) \text{ km}$, $(30, -10) \text{ km}$, $(-30, 20) \text{ km}$ and $(-20, 40) \text{ km}$. The spontaneous birth random finite set (RFS) is Poisson with intensity

$$\begin{aligned} \gamma_k(\xi_k, r_k) &= 0.1\pi_k(r_k)(\mathcal{N}(\xi_k; m_\gamma^{(1)}, P_\gamma) + \mathcal{N}(\xi_k; m_\gamma^{(2)}, P_\gamma) + \\ &\quad \mathcal{N}(\xi_k; m_\gamma^{(3)}, P_\gamma) + \mathcal{N}(\xi_k; m_\gamma^{(4)}, P_\gamma) + \mathcal{N}(\xi_k; m_\gamma^{(5)}, P_\gamma)), \end{aligned}$$

with

$$\begin{aligned} m_\gamma^{(1)} &= (-2 \times 10^4, 0, -2 \times 10^4, 0, 0)^T, \\ m_\gamma^{(2)} &= (1 \times 10^4, 0, 2 \times 10^4, 0, 0)^T, \\ m_\gamma^{(3)} &= (3 \times 10^4, 0, -1 \times 10^4, 0, 0)^T, \\ m_\gamma^{(4)} &= (-3 \times 10^4, 0, 2 \times 10^4, 0, 0)^T, \\ m_\gamma^{(5)} &= (-2 \times 10^4, 0, 4 \times 10^4, 0, 0)^T, \\ P_\gamma &= \text{diag}([10^3, 200, 10^3, 200, 0]). \end{aligned}$$

The settings for all other parameters are identical to those in Section 3.4.1.

Fig. 4.7 shows the true trajectories of five aircraft in the horizontal plane that appear in the surveillance region and disappear at different times and locations. The aircraft perform a sequence of maneuvers at a turn rate in the interval $[-2, 2]^\circ \text{ s}^{-1}$.

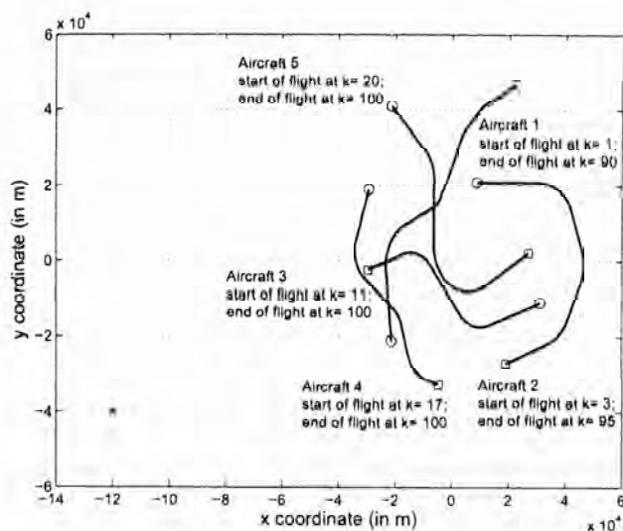


Figure 4.7: Trajectory of the vehicle. 'o'– location of vehicle at $k = 1$; '□'– location of vehicle at $k = 100$ ('x'– location of sensor).

A 1-D view of these trajectories along each axis with cluttered measurements plotted against time is shown in Fig. 4.8. The position estimates of the PHD filter in Fig. 4.9 show that the filter successfully tracks the targets in clutter. Occasionally, the filter underestimates the number of aircraft in the surveillance region and momentarily loses track. Similarly, an overestimate of the number of aircraft produces false estimates which as shown do not propagate with time. There are two causes of this mis-estimation of the number of targets. The predicted multi-target RFS is assumed to be Poisson (see Section 3.2.3), and so the number of targets is Poisson distributed. The mean of a Poisson distribution is the same as its variance so the variance of the estimated number of targets is high. This is compounded by errors in the prediction and update steps which cannot be computed exactly for nonlinear models. However, the mean absolute error in the estimated number of targets averaged over 10^3 Monte Carlo runs shown in Fig. 4.10 suggests that the LFT based JMS-PHD filter gives more reliable estimates than the unscented JMS-PHD filter presented in Section 3.4. The reason for this is that the estimates of the state of the targets distance closer to the true states using the LFT model. The existence of the targets is discriminated from noise and is accounted for in computing the effective number of targets for extracting the state of the targets.

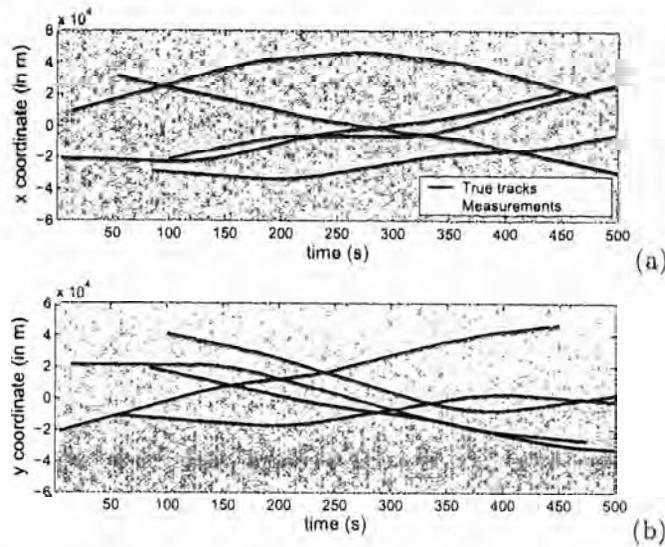


Figure 4.8: Measurement data (projected on the x and y axis) and true target positions.

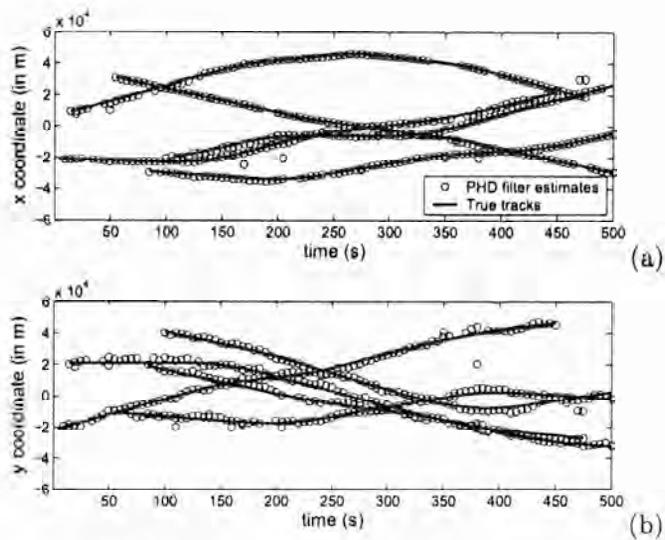


Figure 4.9: Position estimates of the Gaussian mixture PHD filter using the LFT model.

4.4 General solution to Bayes recursion

Nonlinear systems that do not admit an exact equivalent LFT representation can be expressed as a nonlinear fractional transformation (NFT) system [129, 52, 51] which is exact. In this section the solution to the Bayes recursion in Section 4.2.2 is generalized

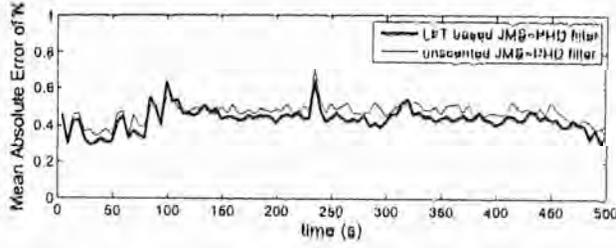


Figure 4.10: Mean absolute error of estimated number of targets using the LFT based JMS-PHD filter and the unscented JMS-PHD filter.

to accommodate any nonlinear structure using the NFT model given below

$$x_{k+1} = A_k x_k + B_{1,k} w_k + B_{2,k} w_{\Delta k}, \quad (4.46)$$

$$z_k = C_{1,k} x_k + D_{11,k} v_k + D_{12,k} w_{\Delta k}, \quad (4.47)$$

$$z_{\Delta k} = C_{2,k} x_k + D_{21,k} u_k + D_{22,k} w_{\Delta k}, \quad (4.48)$$

$$w_{\Delta k} = \Delta(x_k) z_{\Delta k}. \quad (4.49)$$

where $u_k = 1$ is a fixed input, $D_{21,k} \in \mathbb{R}^{n_{\Delta}}$ and the feedback connection matrix is given by $\Delta(x_k) = \sum_{i=0}^n \Delta_i(x_k(i)) x_k(i)$. For rational nonlinearities, the NFT model can be transformed into the LFT representation (4.1)-(4.2) by introducing additional auxiliary variables.

Under the standard assumption on the noise processes in Section 4.2.2, suppose the conditional expectation of x_{k-1} given Z_{k-1} is m_{k-1} with covariance P_{k-1} . Then, Propositions 4.1 and 4.2 generalize to the following results.

Proposition 4.3. *The expectation of the predicted state at time k conditional on the data up to time $k-1$ is accepted as $m_{k|k-1}$ with covariance $P_{k|k-1}$ where $m_{k|k-1}$ is taken as given in (4.16) and $P_{k|k-1}$ is taken as (4.17) with $\bar{w}_{\Delta k-1}$ and $w_{\Delta k-1,i}$ in (4.19) and (4.20) respectively given by*

$$\bar{w}_{\Delta k-1} = (I - \bar{\Delta}_{k-1} D_{22,k-1})^{-1} \left(\frac{1}{p+1} \sum_{i=0}^p \Delta(x_{k-1,i}) (C_{2,k-1} x_{k-1,i} + D_{21,k} u_k) \right) \quad (4.50)$$

$$w_{\Delta k-1,i} = \Delta(x_{k-1,i}) (C_{2,k-1} x_{k-1,i} + D_{21,k} u_k + D_{22,k-1} \bar{w}_{\Delta k-1}), \quad i = 0, \dots, p. \quad (4.51)$$

Proposition 4.4. *Given the conditional mean and covariance of the predicted state as above, the required estimate of the state conditional on Z_k is m_k with covariance P_k*

given by (4.23) and (4.24) respectively with $\bar{w}_{\Delta k|k-1}$ and $w_{\Delta k|k-1,i}$ in (4.28) and (4.29) respectively as

$$\bar{w}_{\Delta k|k-1} = (I - \bar{\Delta}_{k|k-1} D_{22,k})^{-1} \left(\frac{1}{p+1} \sum_{i=0}^p \Delta(x_{k|k-1,i}) (C_{2,k} x_{k|k-1,i} + D_{21,k} u_k) \right) \quad (4.52)$$

$$w_{\Delta k|k-1,i} = \Delta(x_{k|k-1,i}) (C_{2,k} x_{k|k-1,i} + D_{21,k} u_k + D_{22,k} \bar{w}_{\Delta k|k-1}), \quad i = 0, \dots, p. \quad (4.53)$$

4.4.1 Example IV

The inverted pendulum on a cart problem, a well-known unstable nonlinear system [50] is considered. The aim is to accurately track the motion parameters of the cart based on its position measurements. Using Euler's method the system can be discretized into a discrete-time nonlinear system [139],

$$x_{k+1} = \begin{bmatrix} x_{1,k} + T x_{2,k} \\ x_{2,k} + \frac{T(mlx_{4,k}^2 \sin(x_{3,k}) - bx_{2,k} - mg \cos(x_{3,k}) \sin(x_{3,k}))}{M + m \sin^2(x_{3,k})} \\ x_{3,k} + T x_{4,k} \\ x_{4,k} + \frac{T((M+m)g \sin(x_{3,k}) + bx_{2,k} \cos(x_{3,k}) - mlx_{4,k}^2 \sin(x_{3,k}) \cos(x_{3,k}))}{l(M + m \sin^2(x_{3,k}))} \end{bmatrix} + B_k w_k, \quad (4.54)$$

$$z_k = C_k x_k + D_k v_k, \quad (4.55)$$

where $w_k \sim \mathcal{N}(\cdot; 0, Q)$ with $Q = \text{diag}([0.04, 10^{-4}])$ and $v_k \sim \mathcal{N}(\cdot; 0, R)$ with $R = 10^{-6}$. The constant $M = 1378 \text{ g}$ is the mass of the cart and $m = 51 \text{ g}$ the mass of the block on the pendulum of length $l = 0.325 \text{ m}$. The acceleration due to gravity is taken as $g = 9.81 \text{ m s}^{-2}$ and the coefficient of friction due to motion of the cart $b = 12980 \text{ g s}^{-1}$. The sampling period $T = 0.1 \text{ s}$ is taken. Let (p_k, v_k) be the translational motion parameters (position and velocity respectively) of the cart and (θ_k, ω_k) be the angular motion parameters (position and velocity respectively) of the pendulum at time k , the state variables are taken as $x_{1,k} = p_k$, $x_{2,k} = v_k$, $x_{3,k} = \theta_k$ and $x_{4,k} = \omega_k$.

The nonlinear system (4.54)-(4.55) can alternatively be expressed in the NFT

form with

$$A_k = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 - Tb/M & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix}, B_{1,k} = B_k = \begin{bmatrix} 0.005 & 0 \\ 0.1 & 0 \\ 0 & 0.015 \\ 0 & 0.3 \end{bmatrix}, B_{2,k} = \begin{bmatrix} 0_{1,17} \\ [J \ 0_{1,9}] \\ 0_{1,17} \\ [0_{1,9} \ L] \end{bmatrix},$$

$$C_{1,k} = C_k = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}, \quad D_{11,k} = D_k = 1, \quad D_{12,k} = 0_{1,17},$$

$$C_{2,k} = \begin{bmatrix} 0_{1,4} \\ \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix} \\ 0_{8,4} \\ \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\ 0_{5,4} \end{bmatrix}, \quad D_{21,k}^T = \begin{bmatrix} 1 & 0_{1,2} & 1 & 0_{1,4} & 1_{1,2} & 0_{1,7} \end{bmatrix},$$

$$D_{22,k} = \begin{bmatrix} \begin{bmatrix} -m/M & 0_{1,16} \end{bmatrix} \\ 0_{1,17} \\ \begin{bmatrix} 0 & 1 & 0_{1,15} \end{bmatrix} \\ 0_{1,17} \\ \begin{bmatrix} 0_{1,3} & 1 & 0_{1,13} \end{bmatrix} \\ \begin{bmatrix} 0_{1,2} & 1 & 0_{1,14} \end{bmatrix} \\ \begin{bmatrix} 1 & 0_{1,16} \end{bmatrix} \\ \begin{bmatrix} 0_{1,4} & 1 & 0_{1,12} \end{bmatrix} \\ \begin{bmatrix} 0_{1,8} & -m/M & 0_{1,8} \end{bmatrix} \\ 0_{3,17} \\ N \end{bmatrix}, \quad N = \begin{bmatrix} 0_{1,11} & 1 & 0_{1,5} \\ 0_{1,12} & 1 & 0_{1,4} \\ 0_{1,9} & 1 & 0_{1,7} \\ 0_{1,10} & 1 & 0_{1,6} \\ 0_{1,13} & 1 & 0_{1,3} \end{bmatrix}, \quad (4.56)$$

with

$$J = \frac{T}{M} m \begin{bmatrix} 0_{1,2} & 1 & 0 & -g & -m/M \begin{bmatrix} 1 & -b/m & -g \end{bmatrix} \end{bmatrix},$$

$$L = \frac{T}{lM} \begin{bmatrix} (M+m)g & b & 0_{1,2} & -ml & -m/M \begin{bmatrix} (M+m)g & b & -ml \end{bmatrix} \end{bmatrix}.$$

The feedback connection matrix is

$$\Delta(x_k) = \text{diag}([\delta_{1,k}, x_{4,k}, \tilde{\delta}_k, \sin(x_{3,k}), \delta_{2,k}, x_{2,k}, \delta_{2,k}, \delta_{1,k}, \tilde{\delta}_k, x_{4,k}, \tilde{\delta}_k, \delta_{1,k} 1_{1,3}]), \quad (4.57)$$

where

$$\begin{aligned} \delta_{1,k} &= \sin^2(x_{3,k}), \\ \delta_{2,k} &= \left(1 + \delta_{1,k} \frac{m}{M}\right)^{-1} \delta_{1,k}, \\ \tilde{\delta}_k &= [\sin(x_{3,k}) \quad \cos(x_{3,k})]. \end{aligned}$$

Fig. 4.11 shows the simulated translational and angular motion parameters of the cart and the pendulum respectively before the mass m strikes the cart. With the initial estimate of the state taken as $\bar{x}_0 = (0.5, 0.01, 0.01, 0)^T$ and covariance $R_{x,0} = 0.5I_4$, Fig. 4.12 shows the true position and velocity of the cart and the estimates given by the proposed filter. The MSE in the estimates is shown in Fig. 4.13. The results shown in Fig. 4.13 are averaged over 10^3 Monte Carlo runs and indicate that the proposed filter works better than the UKF for most of the time. This example demonstrates that the NFT for filtering of discrete-time models can be applied to a general class of nonlinear problems.

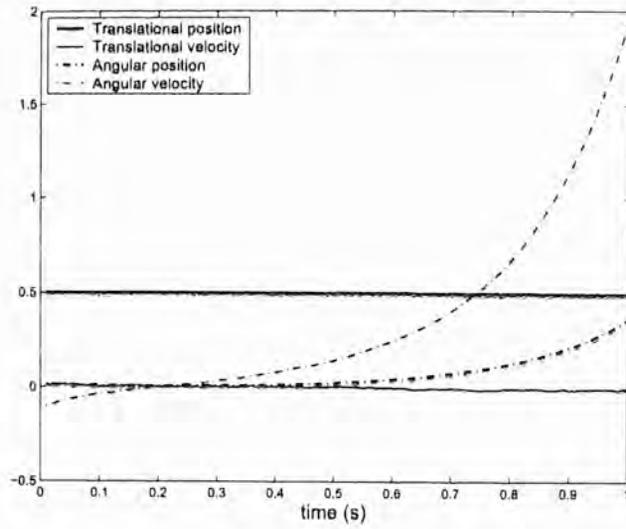


Figure 4.11: Trajectories of the cart and the pendulum.

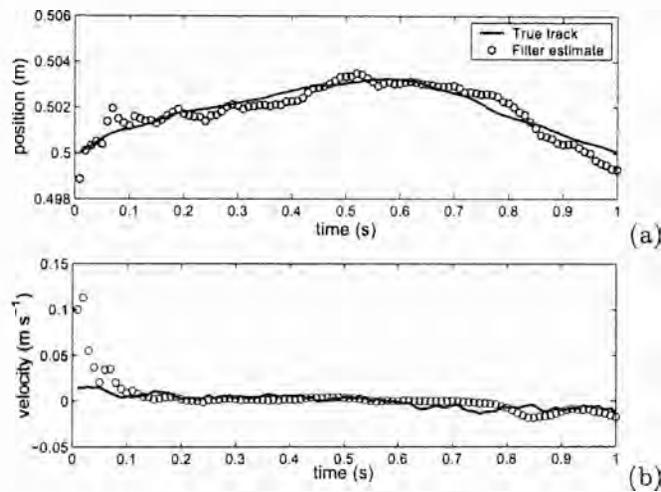


Figure 4.12: True motion parameters of the cart and estimates using the proposed filter.

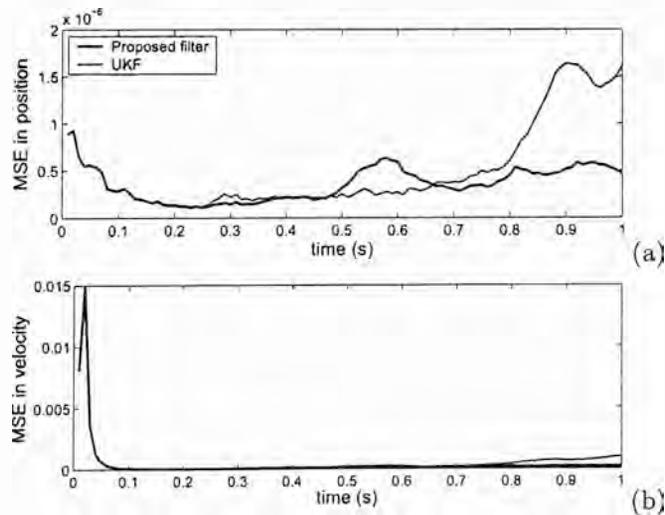


Figure 4.13: MSE in the estimates of motion parameters of the cart using the proposed filter.

Chapter 5

LFT based filtering of the continuous-time model

The linear fractional transformation (LFT) based filter for the discrete-time model presented in Chapter 4 is an attractive approach to filtering for a broad class of nonlinear problems. A closed form solution to Bayes recursion was derived based on the unscented transform localized to a simple nonlinear structure in the feedback. However, the discrete-time state equation is unnatural for many real processes evolving in continuous-time where the stochastic dynamical differential equation gives a realistic model. In this chapter, an analytic solution to the LFT based filter is presented for the continuous-time dynamical model with sampled-data measurements. Simulation results demonstrate that the proposed filtering approach is efficient for online implementation and is applicable to problems where standard analytical approximation based methods fail. Moreover, a comparison with the unscented Kalman filter (UKF) shows that the proposed filter outperforms the UKF in terms of accuracy of estimation as well computational efficiency.

The chapter is structured as follows: Section 5.1 gives the motivation for LFT-based filtering. Section 5.2 states the continuous-discrete filtering problem and the challenge of nonlinear filtering. In Section 5.2 the procedure for the calculation of the moments using the standard analytical approximation methods is examined in some detail before presenting the procedure for the proposed approach. An analytical solution to Bayes recursion using the LFT model for the continuous-discrete problem is then

presented. Simulation results are given in Section 5.4 to demonstrate the performance of the proposed filter and the UKF.

5.1 Introduction

The estimation theory gives elegant solution to nonlinear Bayesian filtering problems in continuous-time dynamics in terms of Kushner-Stratonovich [76, 126, 22] and Zakai [142, 64] partial differential equation for the conditional density. Apart from the linear Gaussian case [66] however, a tractably realizable solution is available only under certain conditions [13] (see also [35, 79] and the references therein).

The stochastic dynamic differential equation is a realistic model of most natural continuous-time processes. For instance, the kinematic state comprising of position and velocity of a moving vehicle. However, observations are often available and processed at discrete time base by digital hardware. For instance, the periodic reports from a radar when an aircraft is illuminated by the rotating radar antenna. From Shannon-Nyquist theory, a sampled signal with high enough sampling rate gives complete information of the band-limited signal, i.e. an analog signal can be exactly recovered from its sampled data. Thus the continuous-time measurement does not necessarily give more information while it may cause additional complications. The estimation problem when the dynamical state model is continuous-time but observations are sampled-data measurements is referred to as *continuous-discrete filtering* [58, 59].

Numerical and grid based methods for the Fokker-Planck equation in continuous-time filtering can still be applied to the continuous-discrete filtering problem for propagating the distribution of the state at the next sampling instant [24, 96]. However, these methods involve quite computationally intensive algorithms. The introduction of heuristics for computational efficiency tailor the algorithms to only a particular problem. The measure transformation based approaches using particle filters [39, 40] can be found in [72, 54, 98]. The particle filter based techniques in practice require a sufficiently

large number of samples and guarantee convergence of the estimate to the optimal Bayes solution only in the limit that the number of samples approaches infinity. The nonlinear projection filter [49] uses the Galerkin method (see e.g., [48]) to solve the Fokker-Planck equation. In [28] the solution is approximated using generalized Edgeworth series and Gauss-Hermite quadrature. Again, the computational load and then the practicability for real time application remain the greatest challenge and unsolved issues within these later methods.

Under certain constraints it is possible to obtain exact filters which can be implemented in real-time by a finite number of ordinary differential equations [36, 74]. For general nonlinear problems, the extended Kalman filter (EKF) [59, 44] is a tractably realizable approximation using the local linearization technique. The resulting linearized dynamical equation can be transferred to the discrete-time equivalent whereby the problem becomes that of discrete-time filtering. Although the solution is obtained in analytical fashion, the accuracy of the discrete-time model is contingent on the quality of the local linearization technique employed by the EKF. For discrete-time systems, the unscented Kalman filter (UKF) [62, 63] has been shown to be computationally efficient and to give estimates with a higher order accuracy than the EKF (see Section 2.5 for further explanation). The UKF directly computes and propagates the first and second-order moments of the state. However, for continuous-discrete filtering this entails solving the differential equations of the first and second-order moments of the predicted state which requires iterative methods to approximate the equivalent discrete-time filter which are often prohibitive for real-time implementation.

The linear fractional transformation (LFT) method [143, 4, 3] gives an exact model for a broad class of nonlinear systems characterized by a linear part and a nonlinear structure in the feedback with sparse representation. By arranging the unscented transformation [62, 63] in the feedback loop the approximation in the LFT model is sufficiently localized to the feedback to linearize a simple nonlinear structure (see Section

4.2). In this chapter the LFT framework is explored for problems in nonlinear filtering of the continuous-discrete system. The objective of this chapter is essentially four fold:

- To present the LFT framework as an efficient approach for simple approximation of the stochastic differential equation of the state prediction.
- To propose the LFT model for accurate estimation of the state conditional on observations.
- To show that the LFT is a powerful tool for online nonlinear filtering which avoids recalculation of system matrices at each sampling instant for processing of the observations.
- To show that the proposed filtering approach economizes on signal processing operations. This is demonstrated through a simulation example where even at a higher sampling rate the UKF does not perform better than the proposed method.

In summary, the aim is to show that the LFT is a new powerful framework for online nonlinear tracking.

5.2 Background: Moment propagation

Consider the dynamical equation in the Itô differential equation form

$$dx(t) = f(x(t))dt + d\beta(t), \quad (5.1)$$

where f denotes an arbitrary nonlinear drift, $x(t) \in \mathbb{R}^n$ is the state of the system at time t , $\beta(\cdot)$ is the Brownian process with diffusion $\tilde{Q}(\cdot)$ assumed independent of the state $x(\cdot)$ which is a reasonable simplification for modeling in applications of target tracking, control and communications¹.

¹ Under such an assumption the Itô and Stratonovich interpretations of the stochastic differential equations are equivalent.

Although in principle the observation flow is continuous time

$$z(t)dt = g(x(t))dt + d\tilde{v}(t), \quad (5.2)$$

it is important to realize by the foundation of digital signal processing that only sampled-data measurements given by

$$z(t_k) = g(x(t_k)) + v(t_k), \quad (5.3)$$

are processed in real-time by digital hardware in most cases for estimation and interpolation purposes. As mentioned above, by Shannon-Nyquist theory, in most cases there is no loss of information by frequently sampling (5.2) to get (5.3). Here g is an arbitrary nonlinear mapping, $z(t_k) \in \mathbb{R}^m$ is the observation at time $t_k = kT$ with T as the sampling period, $v(t_k)$ is the measurement noise with zero mean and covariance $R(t_k)$, statistically independent of the state $x(t_k)$.

Suppose $Z(t_k) = (z(t_1), \dots, z(t_k))$ is the sequence of observations up to time t_k . The continuous-discrete filtering problem is to estimate $x(t_k)|Z(t_k)$, the state $x(t_k)$ at time t_k conditional on $Z(t_k)$. As the state itself already carries all information of the past observations, the filtering recursion constitutes the two steps: (i) based on $x(t_{k-1})|Z(t_{k-1})$ and state equation (5.1) on $[t_{k-1}, t_k]$ to predict (or interpolate) $x(t_k)|Z(t_{k-1})$; (ii) Using $x(t_k)|Z(t_{k-1})$ and observation (5.3) at time t_k to estimate $x(t_k)|Z(t_k)$. We will see that both steps are based on the estimation problem of Theorem 2.1.

With linear mappings f and g in (5.1), (5.3), the continuous-discrete linear state space model is

$$dx(t) = \tilde{A}(t)x(t)dt + d\beta(t), \quad (5.4)$$

$$z(t_k) = C(t_k)x(t_k) + v(t_k), \quad (5.5)$$

where $\tilde{A}(\cdot) \in \mathbb{R}^{n \times n}$ and $C(\cdot) \in \mathbb{R}^{m \times n}$. Suppose at time t_0 , the estimate of the random variable $x(t_0)$ is $\bar{x}(t_0) = m(t_0|_{-1})$ and the covariance is $R_x(t_0) = P(t_0|_{-1})$. The

expectation of $z(t_0)$ in (5.5) is thus $\eta(t_0) = C(t_0)m(t_{0|t_0})$ with covariance $R_z(t_0) = C(t_0)P(t_{0|t_0})C^T(t_0) + R(t_0)$ and the cross-covariance of $z(t_0)$ and $x(t_0)$ is given by $R_{zx}(t_0) = C(t_0)P(t_{0|t_0})$. On arrival of data $z(t_0)$, by Theorem 2.1 $x(t_0)|Z(t_0) = x(t_0)|z(t_0)$ has expectation

$$m(t_0) = m(t_{0|t_0}) + K(t_0)(z(t_0) - \eta(t_0)),$$

and covariance $P(t_0) = P(t_{0|t_0}) - K(t_0)C(t_0)P(t_{0|t_0})$, where $K(t_0) = R_{zx}^T(t_0)R_z^{-1}(t_0) = P(t_{0|t_0})C^T(t_0)(C(t_0)P(t_{0|t_0})C^T(t_0) + R(t_0))^{-1}$.

The estimate $m(t_1|t_0)$ of the predicted state for $x(t)$ at time $t = t_1$ based on $z(t_0)$ is the solution at $t = t_1$ of the differential equation

$$\frac{d}{dt}m(t) = \tilde{A}(t)m(t), \quad (5.6)$$

on $[t_0, t_1]$, or the corresponding integral equation

$$m(t_1|t_0) - m(t_0) = \int_{t_0}^{t_1} \tilde{A}(t)m(t)dt. \quad (5.7)$$

Similarly, the covariance $P(t_1|t_0)$ is the solution at $t = t_1$ of the differential equation

$$\frac{d}{dt}P(t) = \tilde{A}(t)P(t) + P(t)\tilde{A}^T(t) + \tilde{Q}(t), \quad (5.8)$$

on $[t_0, t_1]$. A similar realization for $k \geq 1$ is the following version of the Kalman filter.

Theorem 5.1. *Suppose the estimate of $\tilde{x}(t_{k-1}) = x(t_{k-1})|Z(t_{k-1})$, the state $x(t_{k-1})$ based on the history of observations $Z(t_{k-1})$ at time t_{k-1} is $m(t_{k-1})$ and the covariance is $P(t_{k-1})$. Then, the conditional expectation $m(t_k|t_{k-1})$ and covariance $P(t_k|t_{k-1})$ of the predicted state $\hat{x}(t_k) = x(t_k)|Z(t_{k-1})$ at time t_k are defined by the solutions at $t = t_k$ of the following differential equations*

$$\frac{d}{dt}m(t) = \tilde{A}(t)m(t), \quad (5.9)$$

$$\frac{d}{dt}P(t) = \tilde{A}(t)P(t) + P(t)\tilde{A}^T(t) + \tilde{Q}(t) \quad (5.10)$$

on $[t_{k-1}, t_k]$.

Theorem 5.2. *The conditional expectation $m(t_k)$ and covariance $P(t_k)$ of state estimate $\tilde{x}(t_k) = x(t_k)|Z(t_k)$ are defined by*

$$m(t_k) = m(t_k|t_{k-1}) + K(t_k)(z(t_k) - \eta(t_k)), \quad (5.11)$$

$$P(t_k) = P(t_k|t_{k-1}) - K(t_k)C(t_k)P(t_k|t_{k-1}), \quad (5.12)$$

with

$$\eta(t_k) = C(t_k)m(t_k|t_{k-1}), \quad (5.13)$$

$$K(t_k) = P(t_k|t_{k-1})C^T(t_k)(C(t_k)P(t_k|t_{k-1})C^T(t_k) + R(t_k))^{-1}. \quad (5.14)$$

Note that in case of continuous-time measurements

$$z(t)dt = C(t)x(t)dt + d\tilde{v}(t), \quad (5.15)$$

with $\mathbf{E}(d\tilde{v}(t)d\tilde{v}^T(t)) = \tilde{R}(t)dt$. By setting $t = t_k = t_{k-1}$, $P(t) = P(t_k|t_{k-1})$, $m(t) = m(t_k|t_{k-1})$ and $\tilde{P}(t) = P(t_k)$, $\tilde{m}(t) = m(t_k)$ in (5.11)-(5.12) with (5.14) replaced by

$$K(t) = P(t)C^T(t)\tilde{R}^{-1}(t)C(t)P(t) \quad (5.16)$$

and then replacing $m(t)$ and $P(t)$ by $\tilde{m}(t)$ and $\tilde{P}(t)$ on the right hand sides of (5.9)-(5.10) leads to the well known Kalman-Bucy filter

$$\frac{d}{dt}m(t) = \tilde{A}(t)m(t) + K(t)(z(t) - C(t)m(t)), \quad (5.17)$$

$$\frac{d}{dt}P(t) = \tilde{A}(t)P(t) + P(t)\tilde{A}^T(t) + \tilde{Q}(t) - P(t)C^T(t)\tilde{R}^{-1}(t)C(t)P(t). \quad (5.18)$$

Like [84, Theorem 1], Gaussian *a priori* distribution is not required. Comparing the two sets of equations (5.9)-(5.12) and (5.17)-(5.18), one can see that the computational load for the latter is much heavier due to the differential Riccati equation (5.18). For moderately large dimension, the real-time solution for (5.17)-(5.18) is still a challenging task. On the other hand, intuitively the equations (5.9)-(5.12) are preferable for estimation and filtering purpose. To see further motivation of the later development, consider the nonlinear versions of (5.9)-(5.12) and (5.17)-(5.18). The continuous-discrete equations

(5.1),(5.3) are considered first, where the state dynamics (5.1) is crudely approximated by the interpolation equation of Euler (rectangular) approximation

$$x(t_k) = x(t_{k-1}) + Tf(x(t_{k-1})) + d\beta(t_{k-1}) \quad (5.19)$$

- Suppose the estimate $\tilde{x}(t_{k-1}) := x(t_{k-1})|Z(t_{k-1})$ of the state $x(t_{k-1})$ based on the history of observations $Z(t_{k-1})$ at time t_{k-1} has mean $m(t_{k-1})$ and covariance $P(t_{k-1})$. Then, by Theorem 2.1 the conditional expectation $m(t_k|t_{k-1})$ and covariance $P(t_k|t_{k-1})$ of the predicted state $\hat{x}(t_k) := x(t_k)|Z(t_{k-1})$ at time t_k are defined by

$$m(t_k|t_{k-1}) = m(t_{k-1}) + Tf(\hat{x}(t_{k-1})), \quad (5.20)$$

$$P(t_k|t_{k-1}) = P(t_{k-1}) + T(R_{f_x}(k) + R_{f_x}^T(k)) + T^2 R_f(k) + T\tilde{Q}(t_{k-1}), \quad (5.21)$$

where

$$R_{f_x}(k) = R_{f(\tilde{x}(t_{k-1}))\tilde{x}(t_{k-1})}, \quad R_f(k) = R_{f(\tilde{x}(t_{k-1}))}, \quad \hat{f}(\tilde{x}(t_{k-1})) = \mathbf{E}(f(\tilde{x}(t_{k-1}))). \quad (5.22)$$

- Again, by Theorem 2.1, the conditional expectation $m(t_k)$ and covariance $P(t_k)$ of the state estimate $\tilde{x}(t_k) := x(t_k)|Z(t_k)$ are defined by

$$m(t_k) = m(t_k|t_{k-1}) + K(t_k)(z(t_k) - \eta(t_k)), \quad (5.23)$$

$$P(t_k) = P(t_k|t_{k-1}) - K(t_k)R_{g_x}(k), \quad (5.24)$$

with

$$\eta(t_k) = \mathbf{E}(g(\hat{x}(t_k))), \quad (5.25)$$

$$K(t_k) = R_{g(\hat{x}(t_k))\hat{x}(t_k)}^T (R_{g(\hat{x}(t_k))} + R(t_k)/T)^\dagger, \quad R_{g_x}(k) = R_{g(\hat{x}(t_k))\hat{x}(t_k)}. \quad (5.26)$$

Combining the pair (5.20) and (5.23) and the pair (5.21) and (5.24) to form two equa-

tions leads to

$$\frac{m(t_k) - m(t_{k-1})}{T} = \hat{f}(\tilde{x}(t_{k-1})) + R_{g(\hat{x}(t_k))\hat{x}(t_k)}^T (TR_{g(\hat{x}(t_k))} + R(t_k))^\dagger \cdot (z(t_k) - \mathbf{E}(\hat{x}(t_k))), \quad (5.27)$$

$$\frac{P(t_k) - P(t_{k-1})}{T} = R_{fx}(k) + R_{fx}^T(k) + \tilde{Q}(t_{k-1}) - R_{g(\hat{x}(t_k))\hat{x}(t_k)}^T (TR_{g(\hat{x}(t_k))} + R(t_k))^\dagger R_{g(\hat{x}(t_k))\hat{x}(t_k)}. \quad (5.28)$$

Now, letting $T \rightarrow 0$, then $t_k = t_{k-1} = t$ and $\tilde{x}(t_k) = \hat{x}(t_k) = \hat{x}(t)$, and (5.27)-(5.28) become the following equations

$$\frac{d}{dt}m(t) = \hat{f}(\hat{x}(t)) + R_{g(\hat{x}(t))\hat{x}(t)}^T R^\dagger(t)(z(t) - \mathbf{E}(g(\hat{x}(t)))), \quad (5.29)$$

$$\frac{d}{dt}P(t) = R_{f(\hat{x}(t))\hat{x}(t)} + R_{f(\hat{x}(t))\hat{x}(t)}^T(t) + \tilde{Q}(t) - R_{g(\hat{x}(t))\hat{x}(t)}^T \tilde{R}^\dagger(t) R_{g(\hat{x}(t))\hat{x}(t)}, \quad (5.30)$$

for propagation of the mean $m(t)$ and covariance $P(t)$ of the state estimate $\hat{x}(t) := x(t)|Z(t)$ in the case of using continuous observation (5.2). Of course, the above equations (5.29)-(5.30) are infinite-dimensional in general, which are reduced to the finite dimensional equations (5.17)-(5.18) when both mappings f and g are linear. By examining equations (5.20)-(5.26) and (5.29)-(5.30), one can see the central issue with using linear estimation for filtering nonlinear models is the approximation of the second-order moments of all concerned state and observation random variables. Analytical approximation methods such as the EKF and the UKF actually involve only linear estimators (by using Theorem 2.1) and are different only in the way the second-order moments are approximated (see Section 2.5).

5.3 LFT in filtering nonlinear continuous-time stochastic processes

As can be seen from the analysis in the previous section, the main issues of online filtering nonlinear continuous-time processes are connected with the efficiency of moment prediction of the state process $x(t)$ with distribution $p_x(\cdot)$ satisfying the

stochastic differential equation

$$dx(t) = f(x(t))dt, \quad \mathbf{E}(x(t_{k-1})) = m(t_{k-1}), R_{x(t_{k-1})} = P(t_{k-1}), \quad (5.31)$$

and the conditional moment of

$$x(t_k)|g(x(t_k)) \quad (5.32)$$

This section presents such a calculation. The calculations by the EKF and UKF are analyzed in Section 5.3.1. Then the proposed calculation development which is based on the LFT modeling is presented in Section 5.3.2.

5.3.1 Moment calculations by the EKF and UKF

The challenge in the prediction of $x(t)$ satisfying equation (5.31) is the computation of the intractable integral equations

$$\begin{aligned} d\bar{x}(t) := \mathbf{E}(dx(t)) &= \int (f(x'(t))dt) p_x(x')dx', \\ d\tilde{R}_x(t) &= \int ((f(x'(t))dt) - d\bar{x}(t)) (x'(t) - \bar{x}(t))^T + \\ &\quad (x'(t) - \bar{x}(t)) (f(x'(t))dt - d\bar{x}(t))^T) p_x(x')dx'. \end{aligned} \quad (5.33)$$

which are avoidable in the linear mapping case.

Proposition 5.3. *Regarding the linear stochastic differential equation (SDE) (5.4) on $[t_{k-1}, t_k]$, suppose that $\Psi_k \equiv \Psi(t_k)$ and $\Upsilon_k \equiv \Upsilon(t_k)$ are the solution at $t = t_k$ of the following time-varying differential equations*

$$\frac{d}{dt}\Psi(t) = \tilde{A}(t)\Psi(t), \quad \Psi(t_{k-1}) = I_n, \quad (5.34)$$

$$\frac{d}{dt}\Upsilon(t) = \tilde{A}(t)\Upsilon(t) + \Upsilon(t)\tilde{A}^T(t) + \tilde{Q}(t), \quad \Upsilon(t_{k-1}) = 0_n. \quad (5.35)$$

Suppose that m_{k-1} and P_{k-1} are the mean and auto-covariance of $x(t_{k-1})$. Then, the estimate of the mean and auto-covariance of $x(t_k)$ is

$$m_{k|k-1} = \Psi_k m_{k-1}, \quad (5.36)$$

$$P_{k|k-1} = \Psi_k P_{k-1} \Psi_k^T + \Upsilon_k. \quad (5.37)$$

One can see that (5.36)-(5.37) is the prediction step of the discrete-time Kalman filter. It should be noted that for moderately large dimension the real-time update for Ψ_k and Υ_k at sampling instants is still prohibitive. However, for the particular case of the linear time invariant SDE, i.e. $\tilde{A}(t) \equiv \tilde{A}$ and $\tilde{B}(t) \equiv \tilde{B}$, Ψ_k and Υ_k admit the explicit expressions

$$\Psi_k = e^{\tilde{A}T} \quad (5.38)$$

$$\Upsilon_k = \int_0^T e^{\tilde{A}\tau} \tilde{Q} e^{\tilde{A}^T\tau} d\tau, \quad (5.39)$$

which can be computed off-line and do not require refreshment during the online process. Returning to (5.31), the EKF linearizes the nonlinear mapping f around the expectation $\bar{x}(t)$ of $x(t)$ by

$$f(x(t)) \approx \tilde{A}(t)(x(t) - \bar{x}(t)) + f(\bar{x}(t)), \quad \tilde{A}(t) = \left. \frac{\partial f(x(t))}{\partial x(t)} \right|_{x(t)=\bar{x}(t)}, \quad (5.40)$$

which works well in the case that the expected value $\bar{x}(t)$ lies in the proximity of distributed values of $x(t)$. Substituting the approximation of $f(x(t))$ from (5.40) in the integral equations (2.58) results in

$$\frac{d}{dt} \bar{x}(t) = \tilde{A}(t) \bar{x}(t), \quad (5.41)$$

$$\frac{d}{dt} \tilde{R}_x(t) = \tilde{A}(t) \tilde{R}_x(t) + \tilde{R}_x(t) \tilde{A}^T(t). \quad (5.42)$$

By Proposition 5.3, (5.41) can be transferred to an equivalent discrete-time model $x(t_{k+1}) = \Psi(t_k)x(t_k)$ where $\Psi(t)$ satisfies the differential equation (5.34). The expectation $m_{k|k-1}$ and the covariance $P_{k|k-1}$ of the predicted state $x(t_k)|Z(t_{k-1})$ are then given by

$$m_{k|k-1} = \Psi_k m_{k-1}, \quad (5.43)$$

$$P_{k|k-1} = \Psi_k P_{k-1} \Psi_k^T, \quad (5.44)$$

where $\Psi_k = \Psi(t_k)$ and m_{k-1} and P_{k-1} are the first two moments of $x(t_{k-1})$. Besides unpredictable validity of the approximation, as mentioned earlier, the online implemen-

tation of the EKF in moderately large dimensional cases is indeed a challenging task due to the time-varying characteristic of the equations (5.41)-(5.42).

On the other hand, using Euler (rectangular) approximation for the differential equation $dx = f(x(t))dt$,

$$x(t) - x(t_{k-1}) = f(x(t_{k-1}))\nu_{k-1}, \quad (5.45)$$

where ν_{k-1} is the length of the interval $[t_{k-1}, t]$, the UKF aims at the direct approximation of the moments of $x(t)$ using the statistical linear regression of $f(x(t))$ around $\bar{x}(t)$. Regression points $x^{(i)}(t_{k-1})$, $i = 0, \dots, p$ where $p = 2n$ are selected for n -dimensional $x(\cdot)$ around $\bar{x}(t_{k-1})$ in a manner such that the sample mean and covariance of the points are identical to the mean and covariance of $x(t_{k-1})$ (see Section 2.5),

$$\begin{aligned} \bar{x}(t_{k-1}) &= \frac{1}{p+1} \sum_{i=0}^p x^{(i)}(t_{k-1}), \\ \tilde{R}_x(t_{k-1}) &= \frac{1}{p+1} \sum_{i=1}^p (x^{(i)}(t_{k-1}) - \bar{x}(t_{k-1}))(x^{(i)}(t_{k-1}) - \bar{x}(t_{k-1}))^T. \end{aligned} \quad (5.46)$$

As $\tilde{R}_x(t_{k-1}) > 0$ and thus admits Cholesky decomposition $\tilde{R}_x(t_{k-1}) = \sum_{i=1}^n q_i q_i^T$, a choice of these regression points is

$$x^{(0)}(t_{k-1}) = \bar{x}(t_{k-1}), \quad x^{(i)}(t_{k-1}) = \bar{x}(t_{k-1}) + \sqrt{\frac{p+1}{2}} q_i, \quad x^{(n+i)}(t_{k-1}) = \bar{x}(t_{k-1}) - \sqrt{\frac{p+1}{2}} q_i.$$

Let $\varphi^{(i)} = x^{(i)}(t_{k-1}) + f(x^{(i)}(t_{k-1}))\nu_{k-1}$, $i = 0, \dots, p$, then the mean and covariance of $x(t)$ can be computed as follows,

$$\bar{x}(t) = \frac{1}{p+1} \sum_{i=0}^p \varphi^{(i)}, \quad \tilde{R}_x(t) = \frac{1}{p+1} \sum_{i=0}^p (\varphi^{(i)} - \bar{x}(t))(\varphi^{(i)} - \bar{x}(t))^T, \quad (5.47)$$

Note that unlike the EKF which approximates the integral equations (5.33), the UKF approximates the solution of (5.33) by utilizing iterative methods which burden the computations for real-time implementation. This is the overhead of using the UKF for the continuous-time dynamical model where accuracy is traded-off for computational efficiency.

In summary, for the approximation of integral equations (5.33), the EKF linearizes the nonlinear deterministic mapping f while the UKF linearizes the distribution $p_x(\cdot)$. For the computation of the conditional moments of (5.32) see Section 2.5.

5.3.2 Moment calculation by LFT modeling

It is known that a broad class of nonlinear mappings including fractional mappings, differentiable at any order, admit an equivalent LFT model [143, 129, 128], i.e. (5.31) can be exactly represented by

$$\begin{bmatrix} dx(t)/dt \\ y_\Delta(t) \end{bmatrix} = \begin{bmatrix} \tilde{A} & \tilde{B} \\ \tilde{C} & \tilde{D} \end{bmatrix} \begin{bmatrix} x(t) \\ w_\Delta(t) \end{bmatrix}, \quad (5.48)$$

$$w_\Delta(t) = \Delta(x(t))y_\Delta(t), \quad (5.49)$$

where $\tilde{A} \in \mathbb{R}^{n \times n}$, $\tilde{B} \in \mathbb{R}^{n \times n_\Delta}$, $\tilde{C} \in \mathbb{R}^{n_\Delta \times n}$ and $\tilde{D} \in \mathbb{R}^{n_\Delta \times n_\Delta}$ are deterministic fixed matrices. The introduced auxiliary variables $w_\Delta(t) \in \mathbb{R}^{n_\Delta}$ and $y_\Delta(t) \in \mathbb{R}^{n_\Delta}$ are related via the feedback channel $\Delta(x(t))$, which admits the structure $\Delta(x(t)) = \sum_{i=1}^n \Delta_i x_i(t)$ where $x_i(\cdot)$ is the i -th element of the vector $x(\cdot)$. In many cases, the LFT representation is straightforward as shown in Section 5.4.

The LFT system of (5.48)-(5.49) is the compact expression

$$dx(t)/dt = (\tilde{A} + \tilde{B}\Delta(x(t))(I - \tilde{D}\Delta(x(t)))^{-1}\tilde{C})x(t), \quad (5.50)$$

where $\Delta(x(t))$ appears in a highly nonlinear fashion. Using either (5.40) to approximate the mapping or (5.45) to directly approximate the moments of $x(t)$ in (5.50) is tantamount to the affine approximations discussed in the previous subsection. On the other hand, a simple nonlinearity appears in the model (5.48)-(5.49) in the feedback only. Under this representation, an alternatively approximated discrete-time solution is the following result.

Proposition 5.4. *The discrete-time approximated solution of the LFT model (5.48)-*

(5.49) given by Proposition 5.3 is

$$\begin{bmatrix} x_{k+1} \\ y_{\Delta k} \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x_k \\ w_{\Delta k} \end{bmatrix}, \quad (5.51)$$

$$w_{\Delta k} = \Delta(x_k)y_{\Delta k}, \quad (5.52)$$

where $A = e^{\tilde{A}T}$, $B = \left(\int_0^T e^{\tilde{A}\tau} d\tau\right) \tilde{B}$, $C = \tilde{C}$, $D = \tilde{D}$ and the feedback connection $\Delta(x_k) = \sum_{i=1}^n \Delta_i x_{k,i}$.

Proof. Consider the differential equation

$$\frac{d}{dt}x(t) = \tilde{A}x(t) + \tilde{B}w_{\Delta}(t). \quad (5.53)$$

By integration

$$x(t) = e^{\tilde{A}t}x(0) + \int_0^t e^{\tilde{A}(t-\tau)} \tilde{B}w_{\Delta}(\tau) d\tau.$$

At time $t_k = kT$

$$x(t_k) = e^{\tilde{A}kT}x(0) + \int_0^{kT} e^{\tilde{A}(kT-\tau)} \tilde{B}w_{\Delta}(\tau) d\tau$$

and

$$\begin{aligned} x(t_{k+1}) &= e^{\tilde{A}(k+1)T}x(0) + \int_0^{(k+1)T} e^{\tilde{A}((k+1)T-\tau)} \tilde{B}w_{\Delta}(\tau) d\tau, \\ &\approx e^{\tilde{A}T}x(t_k) + \left(\int_0^T e^{\tilde{A}\gamma} d\gamma\right) \tilde{B}w_{\Delta}(t_k), \end{aligned} \quad (5.54)$$

where the last step follows from the effective zero-order hold based approximation for $w_{\Delta}(\cdot)$ and introducing $\gamma = (k+1)T - \tau$. \square

From the numerical computation perspective, $B = \tilde{A}^{-1}(A-I)\tilde{B}$, if \tilde{A} is invertible. Moreover, for small T , $A \approx I + \tilde{A}T$ and $B = (IT + \frac{1}{2}\tilde{A}T^2)\tilde{B}$ hold. Note that from the information theory viewpoint, Proposition 5.4 is rather exact discretized solution with moderate sampling rate $1/T$, whereas the UKF approximates numerically this solution at much higher sampling rate (see e.g. [2, Table 1] where Euler and second-order Runge-Kutta discretization require rates higher than 2000 Hz and 1000 Hz respectively for reasonable accuracy and thus are not suitable for real-time update whereas the

approximation like the above Proposition 5.4 works extremely well at 45 Hz and thus is online practicable). For the discrete-time LFT model (5.51)-(5.52), the approximation is localized to the feedback path for estimation of the auxiliary random variable $w_{\Delta k}$ in (5.52). The efficiency of the unscented transformation in the discrete-time case is now exploited. Given \bar{x}_{k-1} and P_{k-1} as the first two central moments of x_{k-1} , the regression points $x_{k-1}^{(i)}$, $i = 0, \dots, p$ are chosen by around \bar{x}_{k-1} as shown in (2.63). Define the regression points $w_{\Delta k-1}^{(i)} = \Delta(x_{k-1}^{(i)})y_{\Delta k-1}^{(i)}$, where

$$y_{\Delta k-1}^{(i)} = Cx_{k-1}^{(i)} + D\bar{w}_{\Delta k-1}, \quad (5.55)$$

and $\bar{w}_{\Delta k-1} \approx \mathbf{E}(w_{\Delta k-1})$ is

$$\bar{w}_{\Delta k-1} = (I - \bar{\Delta}_{k-1}D)^{-1} \left(\frac{1}{p+1} \sum_{i=0}^p \Delta(x_{k-1}^{(i)})Cx_{k-1}^{(i)} \right), \quad (5.56)$$

with

$$\bar{\Delta}_{k-1} = \frac{1}{p+1} \sum_{i=0}^p \Delta(x_{k-1}^{(i)}). \quad (5.57)$$

Accordingly, the second-order central moment of $w_{\Delta k-1}$ and the joint second-order moment with x_{k-1} are computed like (2.64)-(2.65):

$$R_{\Delta k-1} = \frac{1}{p+1} \sum_{i=0}^p (w_{\Delta k-1}^{(i)} - \bar{w}_{\Delta k-1})(w_{\Delta k-1}^{(i)} - \bar{w}_{\Delta k-1})^T, \quad (5.58)$$

$$R_{\Delta x, k-1} = \frac{1}{p+1} \sum_{i=0}^p (w_{\Delta k-1}^{(i)} - \bar{w}_{\Delta k-1})(x_{k-1}^{(i)} - \bar{x}_{k-1})^T. \quad (5.59)$$

As discusses in Chapter 4, the approximation of the moments of $w_{\Delta k-1}$ in (5.56), (5.58)-(5.59) averts the linearization of (5.50) which gives poor approximation for highly nonlinear models. Next, the nonlinear mapping g in (5.32) can be analogously expressed in the LFT format is

$$y_k = Ax_k + Bw_{\Delta k}, \quad y_{\Delta k} = Cx_k + Dw_{\Delta k}, \quad w_{\Delta k} = \Delta(x_k)y_{\Delta k}, \quad (5.60)$$

where $y_k = g(x_k) \in \mathbb{R}^m$ is expressed as a linear mapping of x_k and $w_{\Delta k}$ and the nonlinearity appears in the form of the feedback connection $\Delta(x_k)$ which takes the form

as above. The fixed and known system matrices are of appropriate dimensions. The expectation of y_k is

$$\bar{y}_k = A\bar{x}_k + B\bar{w}_{\Delta k},$$

where $\bar{w}_{\Delta k}$ is defined in a similar manner to (5.56)

$$\bar{w}_{\Delta k} = (I - \bar{\Delta}_k D)^{-1} \left(\frac{1}{p+1} \sum_{i=0}^p \Delta(x_k^{(i)}) C x_k^{(i)} \right), \quad (5.61)$$

with

$$\bar{\Delta}_k = \frac{1}{p+1} \sum_{i=0}^p \Delta(x_k^{(i)}). \quad (5.62)$$

The second-order central moments of $w_{\Delta k}$ are likewise computed as (5.58)-(5.59)

$$R_{\Delta k} = \frac{1}{p+1} \sum_{i=0}^p (w_{\Delta k}^{(i)} - \bar{w}_{\Delta k})(w_{\Delta k}^{(i)} - \bar{w}_{\Delta k})^T, \quad (5.63)$$

$$R_{\Delta x, k} = \frac{1}{p+1} \sum_{i=0}^p (w_{\Delta k}^{(i)} - \bar{w}_{\Delta k})(x_k^{(i)} - \bar{x}_k)^T. \quad (5.64)$$

The calculation of the moments in such fashion alleviates the computational intractability of (5.33) for the computation of the first and second-order moments of $g(x_k)$. In the sequel, we explore the LFT framework for the continuous-discrete state space model.

5.3.3 Recursive Bayes filter by LFT

Consider the state space model (5.1), (5.3). For the differentiable nonlinear mapping f , (5.1) can be expressed in the LFT format

$$dx(t) = (\tilde{A}x(t) + \tilde{B}w_{\Delta}(t))dt + d\beta(t), \quad (5.65)$$

$$z_{\Delta}(t) = \tilde{C}_2 x(t) + \tilde{D}_{22} w_{\Delta}(t), \quad (5.66)$$

$$w_{\Delta}(t) = \Delta(x(t))z_{\Delta}(t), \quad (5.67)$$

where $w_{\Delta}(t)$, $z_{\Delta}(t)$ denote auxiliary variables and the system matrices are of appropriate dimensions. By Proposition 5.4 the discrete-time equivalent of (5.65)-(5.67) is

$$x_{k+1} = Ax_k + w_k + Bw_{\Delta k}, \quad z_{\Delta k} = C_2 x_k + D_{22} w_{\Delta k}, \quad w_{\Delta k} = \Delta(x_k)z_{\Delta k}, \quad (5.68)$$

where $A = e^{\bar{A}T}$, $B = \left(\int_0^T e^{\bar{A}\tau} d\tau\right) \tilde{B}$, $C_2 = \tilde{C}_2$ and $D_{22} = \tilde{D}_{22}$. $\{w_k\}$ is a random process with $\mathbf{E}(w_k) = 0$ and $\mathbf{E}(w_k w_k^T) = Q$ where $Q = \int_0^T e^{\bar{A}\tau} \tilde{Q} e^{\bar{A}^T \tau} d\tau$. Similarly, for the differentiable nonlinear mapping g in (5.3) there exists the LFT system

$$z_k = C_1 x_k + v_k + D_{12} w_{\Delta k}, \quad z_{\Delta k} = C_2 x_k + D_{22} w_{\Delta k}, \quad w_{\Delta k} = \Delta(x_k) z_{\Delta k}, \quad (5.69)$$

In the sequel no distinction is made between the LFT model and the nonlinear fractional transformation (NFT) model as was done in Chapter 4, by adopting a general structure of the feedback connection $\Delta(x_k) = \sum_{i=1}^n \Delta_i(x_{k,i}) x_{k,i}$. The LFT now exists for any smooth nonlinear mappings f, g and (5.68)-(5.69) generalize to the following.

$$x_{k+1} = Ax_k + B_1 w_k + B_2 w_{\Delta k}, \quad (5.70)$$

$$z_k = C_1 x_k + v_k + D_{12} w_{\Delta k}, \quad (5.71)$$

$$z_{\Delta k} = C_2 x_k + D_{21} u_k + D_{22} w_{\Delta k}, \quad (5.72)$$

$$w_{\Delta k} = \Delta(x_k) z_{\Delta k}, \quad (5.73)$$

where $u_k \in \mathbb{R}$ is a fixed input, B_1, B_2 and D_{21} are of appropriate dimensions.

This transfers the filtering problem to that of discrete-time filtering where the prediction and estimation steps of the Bayes recursion for the LFT model (5.70)-(5.73) are given by Propositions 4.3 and 4.4 respectively. The advantage of the LFT model (5.70)-(5.73) lies in the efficiency of its linear structure. The deterministic system matrices admit explicit matrix exponential representation and are not required to be refreshed in real-time so it potentially works even for large dimensional problems. Although the EKF may be computationally more efficient than the UKF for low dimensional problems in the continuous-time case, it still requires the matrices to be refreshed in the linearization step as apparent from (5.40). Moreover, as mentioned above, the quality of the approximation involved in the linearization is poor in general which gives unreliable estimates. For better approximation it may require unrealistic observation sampling rate. A serious drawback of the UKF for continuous-discrete filtering is the manner in which it applies the approximation. While the UKF has been shown to be efficient for the

discrete-time model, it employs iterative methods in the continuous-time case which are time consuming. All these drawbacks are absent in the proposed approach. In Chapter 4, the LFT for filtering of discrete-time nonlinear models was shown to perform better than the UKF at similar complexity. Since the proposed approach still applies the unscented transform in discrete-time, the gain in performance is achieved at considerably lower computational complexity.

5.4 Simulation results

In this section simulation results are given to demonstrate the performance of the proposed nonlinear filtering approach. The potential of the proposed filtering method for the continuous-discrete system is realized immediately in Examples I and II which consider the continuous-discrete versions of the models from Sections (4.3.1) and (4.3.2). While it was possible to consider the UKF for filtering in (4.3.1), it fails to converge for the continuous-discrete model. In Example III the problem of multi-target filtering from the previous chapter is considered. Example IV is based on the nonlinear benchmark model [23] of rotational-translational actuator (RTAC).

The simulated data were generated using 100 steps of Euler-Maruyama method [72] between successive measurements while the UKF was implemented in Examples III and IV for a comparison of the performance using 10 steps of the fourth-order Runge-Kutta integration between successive measurements.

5.4.1 Example I

The continuous-time version of the nonlinear autoregressive (AR) equation considered in 4.3.1 is $\ddot{q}(t) = -0.1\dot{q}(t) - q^3(t) + w(t)$ with the noisy measurement $z(t_k) = q(t_k) + v(t_k)$, which admits the following state-space equation formulation with the state

$$x(t) = (x_1(t), x_2(t))^T = (q(t), \dot{q}(t))^T \in \mathbb{R}^2$$

$$\dot{x}(t) = \begin{bmatrix} 0 & 1 \\ -x_1^2(t) & -0.1 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} w(t), \quad (5.74)$$

$$z(t_k) = x_1(t_k) + v(t_k), \quad (5.75)$$

where $w(t)$ is a stochastic process with $\mathbf{E}(w(t)) = 0$ and $\mathbf{E}(w(t)w^T(\tau)) = \tilde{Q}\delta(t-\tau)$ with $\tilde{Q} = 0.04$. $v(t_k) \sim \mathcal{N}(\cdot; 0, R)$ with $R = 0.5$.

The continuous-discrete model (5.74)-(5.75) can be represented in the LFT form (4.12)-(4.15) with

$$A = \begin{bmatrix} 1 & 0.0995 \\ 0 & 0.99 \end{bmatrix}, B_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, B_2 = \begin{bmatrix} 0 & -0.005 \\ 0 & -0.0995 \end{bmatrix},$$

$$C_1 = \begin{bmatrix} 1 & 0 \end{bmatrix}, D_{12} = 0_{1,2}, C_2 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, D_{21} = 0_{2,1}, D_{22} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad (5.76)$$

where $0_{a,b}$ is the $a \times b$ zero matrix. The feedback connection has the simple structure $\Delta(x(t_k)) = x_1(t_k)I_2$ with I_a as the identity matrix of dimension $a \times a$. The true trajectory of the state $x(t)$ for 20 s is shown in Fig. 5.1. Using $\bar{x}(t_0) = (0, 0)^T$ as the initial estimate of the state with covariance $R_x(t_0) = I_2$, the estimates given by the proposed filter at each sampling instant are shown in Fig. 5.2 along with the true states. In Fig. 5.3 the mean square error (MSE) in the estimates obtained from 10^2 Monte Carlo runs is shown indicating that the proposed filter successfully tracks the true trajectory.

5.4.2 Example II

Consider the continuous-discrete version of the problem from Section (4.3.2)

$$\dot{x}(t) = (Q_0 + Q_1x_1^3(t) + Q_2x_2^3(t) + Q_3x_1(t)x_2^2(t) + Q_4x_1(t) + Q_5x_2(t))x(t) + \tilde{B}w(t), \quad (5.77)$$

$$z(t_k) = Cx(t_k) + v(t_k), \quad (5.78)$$

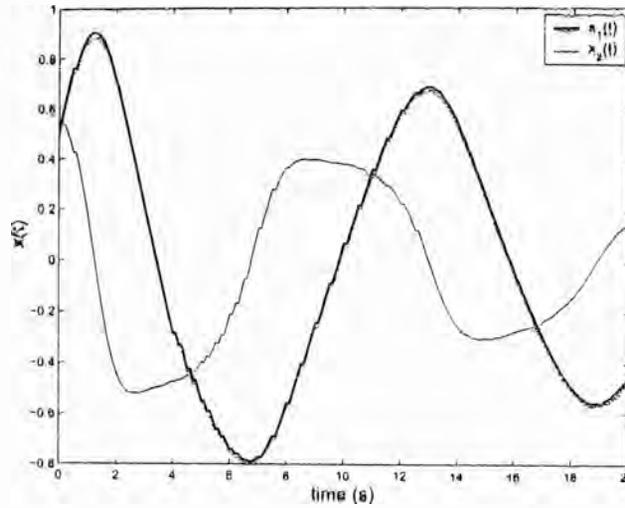


Figure 5.1: Trajectory of the state $x(t) = (x_1(t), x_2(t))^T$.

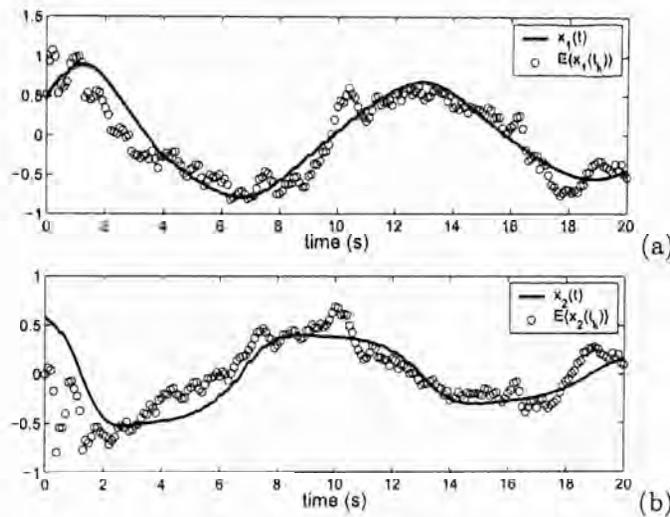


Figure 5.2: True trajectory and the estimate of the state $\mathbf{E}(x(t_k)|Z(t_k))$ given by the proposed filter.

where Q_i , $i = 0, \dots, 5$, \tilde{B} and C are defined in Section (4.3.2), $w(t)$ is zero mean with covariance $\mathbf{E}(w(t)w^T(\tau)) = \tilde{Q}\delta(t-\tau)$ where $\tilde{Q} = 0.01$. $v(t_k) \sim \mathcal{N}(\cdot; 0, R)$ with $R = 100$.

The equivalent discrete-time LFT representation can be constructed with

$$A = \begin{bmatrix} 0.9319 & -0.0942 \\ 0.0094 & 0.9508 \end{bmatrix}, \quad B_1 = \tilde{B},$$

$$B_2 = \begin{bmatrix} 0.0237 & 0 & 0.0285 & -0.0012 & 0 & 0.0082 & 0.0184 & 0.0275 & 0 & 0.0379 & 0.0466 \\ 0.0099 & 0 & 0.0099 & 0.0244 & 0 & 0.0293 & 0.0196 & 0.0294 & 0 & 0.0148 & 0.0344 \end{bmatrix}, \quad (5.79)$$

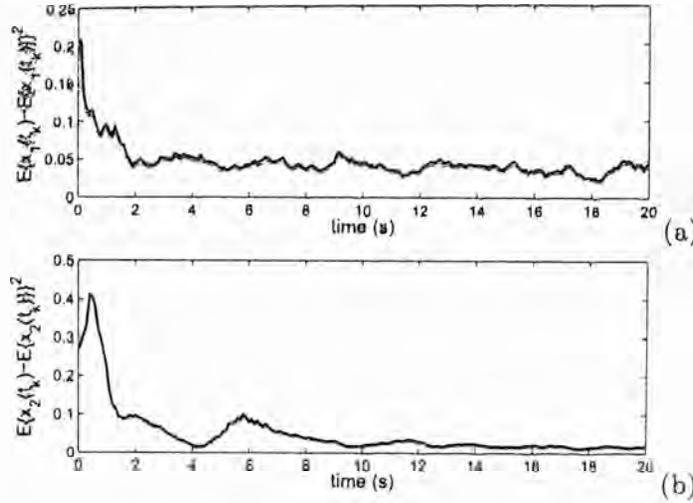


Figure 5.3: Mean square error (MSE) in the estimates using the proposed filter.

and $C_1, C_2, D_{12}, C_2^T, D_{21}, D_{22}$ defined in (4.38). Using $1_{a,b}$ to denote the $a \times b$ matrix with entries one, the feedback connection is given by

$$\Delta(x_k) = \text{diag}([x_{1,k}1_{1,3} \quad x_{2,k}1_{1,4} \quad x_{1,k} \quad x_{2,k}1_{1,3}]). \quad (5.80)$$

The trajectory of the state $x(t)$ as it evolves with time for 10 s is shown in Fig. 5.4. Using $\bar{x}(t_0) = (0, 0)^T$ as the estimate of $x(t_0)$ at time $t = 0$ with covariance $R_x(t_0) = 0.25I_2$, Fig. 5.5 shows the estimate of $x(t)$ given by the proposed filter at each sampling time along with the true trajectory. The MSE obtained from 10^2 Monte Carlo runs using the proposed filter shown in Fig. 5.6 indicates that the proposed filter successfully tracks the true trajectory.

5.4.3 Example III

Returning to the problem of multi-target filtering presented in Section 4.3.3, the aircraft dynamics are modeled by the continuous-time coordinated turn model [10, 9, 33] at different turn rates to describe the maneuver as well as the non-maneuver motion. The motion models are as follows. Model $r(t) = 1$ has a known turn rate of $0^\circ s^{-1}$. The standard deviation of the process noise for model $r(t) = 1$ is 10 m s^{-2} . Model $r(t) = 2$

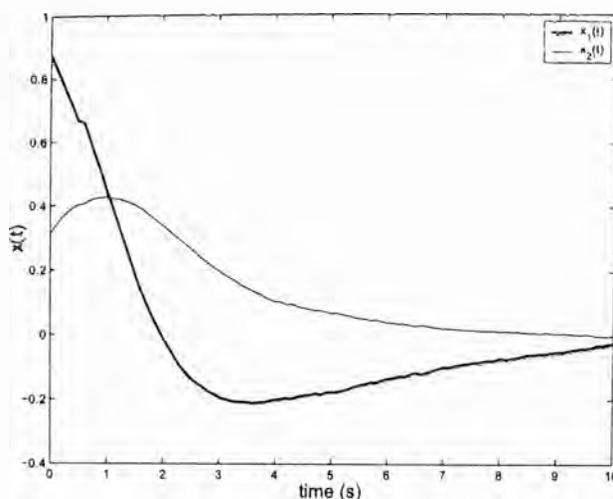


Figure 5.4: Trajectory of the state $x(t) = (x_1(t), x_2(t))^T$.

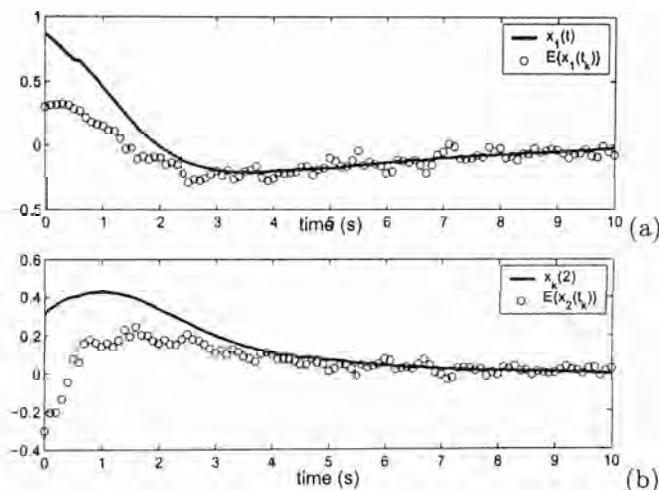


Figure 5.5: True trajectory and the estimate of the state $\mathbf{E}(x(t_k)|Z(t_k))$ given by the proposed filter.

is the nonlinear model with an unknown turn rate $\Omega(t)$ given by

$$\tilde{\mathbf{A}}(r(t)) = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -\Omega(t) & 0 \\ 0 & 0 & \Omega(t) & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \tilde{\mathbf{B}}(r(t)) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The standard deviation of the process noise is 10 m s^{-2} and 0.5° s^{-2} for the the linear and turn portions respectively of the kinematic state during the level turn in $r(t) = 2$.

The sensor model is given in Section 4.3.3. For a given mode r , the single target

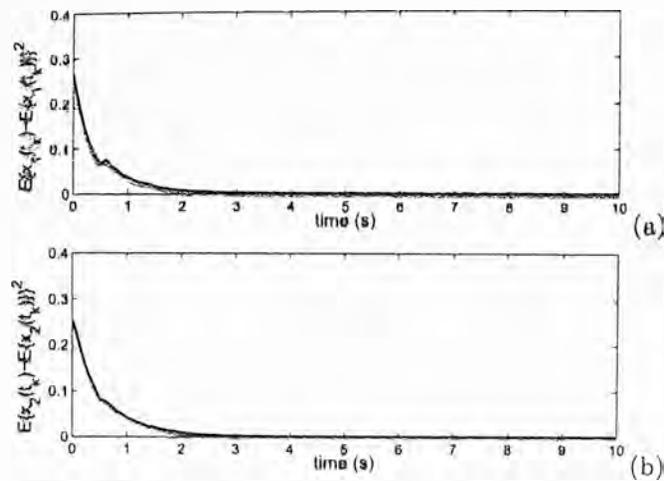


Figure 5.6: MSE in the estimates using the proposed filter

dynamical model is approximated by the linear Gaussian model using the UKF, implemented with 10 steps of the fourth-order Runge-Kutta integration during each sampling period.

Fig. 5.7 shows the true aircraft trajectories. A 1-D view of these trajectories along

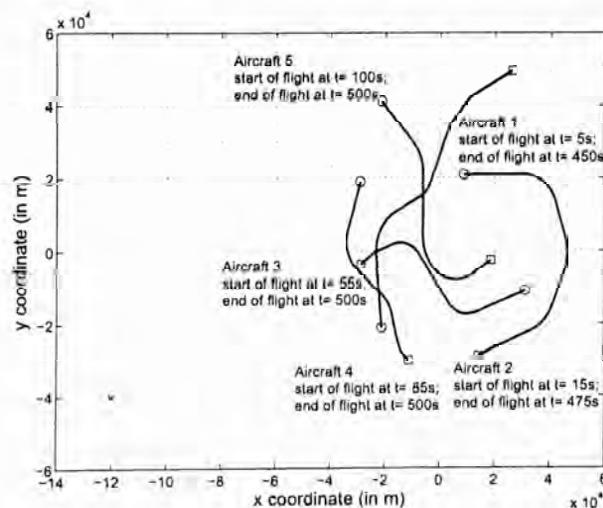


Figure 5.7: Trajectory of the vehicle. ‘o’- location of vehicle at $k = 1$; ‘□’- location of vehicle at $k = 100$ (‘x’- location of sensor).

each axis with cluttered measurements plotted against time is shown in Fig. 5.8. The position estimates of the PHD filter in Fig. 5.9 show that the filter successfully tracks the targets in clutter. The results for the mean absolute error in the estimated number of targets averaged over 10^3 Monte Carlo runs shown in Fig. 5.10 concur with the

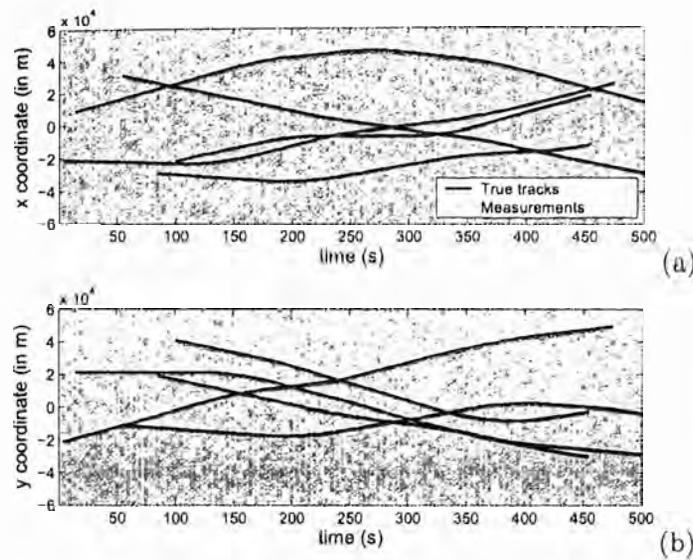


Figure 5.8: Measurement data (projected on the x and y axis) and true target positions.

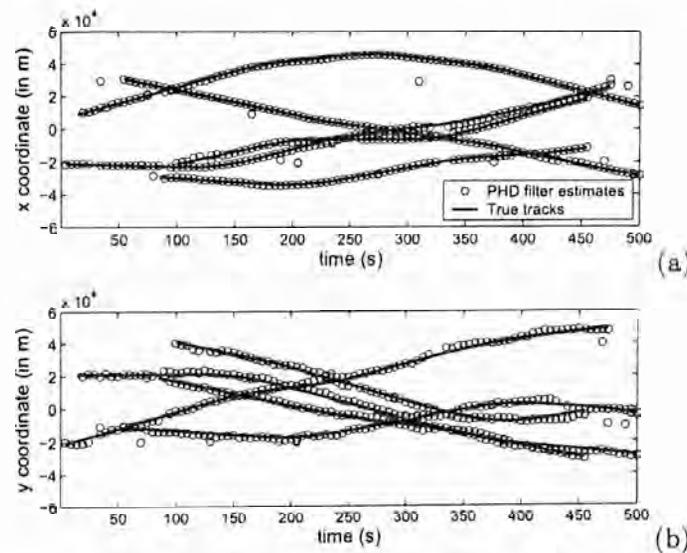


Figure 5.9: Position estimates of the Gaussian mixture PHD filter using the LFT model.

finding in 4.3.3 suggesting that the LFT based JMS-PHD filter is more robust than the unscented JMS-PHD filter.

5.4.4 Example IV

In this example the unstable nonlinear system of the RTAC model [128] is considered. The kinematic state $x(t) = (x_1(t), x_2(t), x_3(t), x_4(t))^T = (\xi(t), \dot{\xi}(t), \theta(t), \dot{\theta}(t))^T$ where $(\xi(t), \dot{\xi}(t))$ denotes the parameters of the oscillator and $(\theta(t), \dot{\theta}(t))$ denotes the

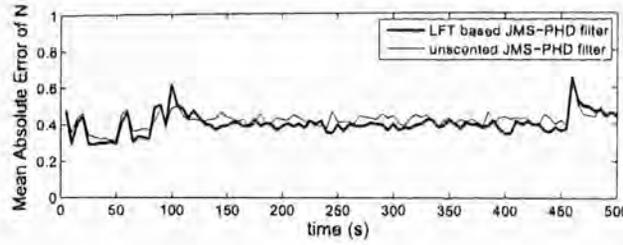


Figure 5.10: Mean absolute error of estimated number of targets using the LFT model and the UKF.

parameters of the actuator. The nonlinear state-space model is given by

$$\dot{x}(t) = \begin{bmatrix} x_2(t) \\ \frac{-x_1(t) + \epsilon x_4^2(t) \sin x_3(t)}{1 - \epsilon^2 \cos^2 x_3(t)} \\ x_4(t) \\ \frac{\epsilon \cos x_3(t)(x_1(t) - \epsilon x_4^2(t) \sin x_3(t))}{1 - \epsilon^2 \cos^2 x_3(t)} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0.7 & 0 \\ 0 & 3 \\ 0 & -1 \end{bmatrix} w(t), \quad (5.81)$$

where $w(t)$ is the random process with $\mathbf{E}(w(t)) = 0$ and $\mathbf{E}(w(t)w^T(\tau)) = \tilde{Q}\delta(t - \tau)$ with $\tilde{Q} = \text{diag}([0.04, 0.001])$. $(\xi(t), \theta(t))$ are available for measurement at sampling interval $T = 10 \text{ ms}$ and $v(t_k) \sim \mathcal{N}(\cdot; 0, R)$ with $R = \text{diag}([0.1, \pi/180])$. The LFT model for (5.81) is

$$\begin{aligned} A &= \begin{bmatrix} 0.9948 & 0.0998 & 0 & 0 \\ -0.1035 & 0.9948 & 0 & 0 \\ 0.0010 & 0 & 1 & 0.1 \\ 0.0194 & 0.0010 & 0 & 1 \end{bmatrix}, \quad B_1 = \begin{bmatrix} 1 & 0 \\ 0.7 & 0 \\ 0 & 3 \\ 0 & -1 \end{bmatrix}, \\ B_2 &= \begin{bmatrix} -0.0002 & 0.0002 & 0.0005 \\ -0.0525 & 0.0525 & 0.1088 \\ 0.0002 & 0.0003 & -0.0001 \\ 0.0525 & 0.0525 & -0.0204 \end{bmatrix}, \quad C_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad D_{12} = 0_{2,3}, \\ C_2 &= \begin{bmatrix} \frac{1}{a_3^2} & 0 & 0 & 0 \\ \frac{1}{a_4^2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad D_{21} = 0_{3,1}, \quad D_{22} = \begin{bmatrix} \frac{1}{a_3} & 0 & -\frac{1}{a_3^2} \\ 0 & -\frac{1}{a_4} & -\frac{1}{a_4^2} \\ 0 & 0 & 0 \end{bmatrix}, \end{aligned} \quad (5.82)$$

$$w_{\Delta k} = [(\epsilon \cos(x_{3,k}) - a_1)\tilde{\Delta}_1 + \epsilon x_{4,k} \sin(x_{3,k})\tilde{\Delta}_2]z_{\Delta k} \quad (5.83)$$

with the same definition of the constants a_1 , a_3 and a_4 as in [128] and $\tilde{\Delta}_1 = \text{diag}([1, 1, 0])$ and $\tilde{\Delta}_2 = \text{diag}([0, 0, 1])$. Under the conditions $\epsilon = 0.2$, $\bar{x}(t_0) = (0.5, 0, 0, 0)^T$ and $R_x(t_0) = \text{diag}([3, 0.3, \pi/60, \pi/60])$ the true trajectories of the oscillator and actuator for 20 s are shown in Fig. 5.11. Fig. 5.12 shows the true position and velocity of the oscillator and the filter estimates while Fig. 5.13 shows the true motion parameters of the actuator and the estimated parameters. In Fig. 5.14 the MSE in estimates of the oscillator parameters is shown obtained using the proposed filter and the UKF. Similarly, the MSE in the estimated parameters of the actuator obtained using both methods is shown in 5.15. The simulation results indicate that the proposed approach gives better performance than the UKF. Also shown is the performance of the UKF processing twice the information available by sampling at intervals of 5 ms. It can be inferred from the results that processing more information does not guarantee an improvement in the estimation.

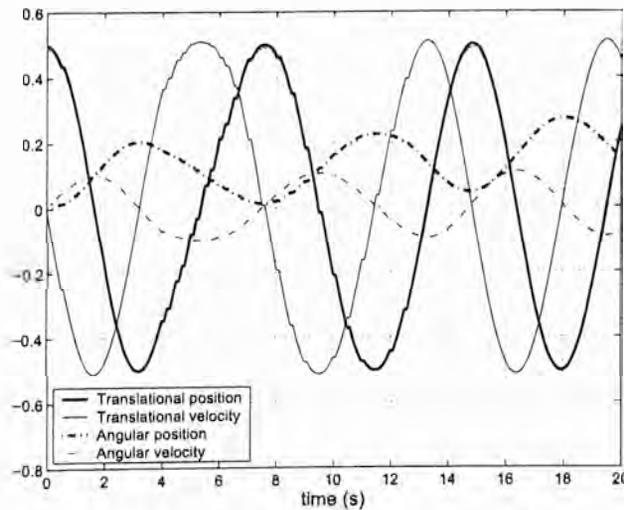


Figure 5.11: Trajectories of the oscillator and actuator.

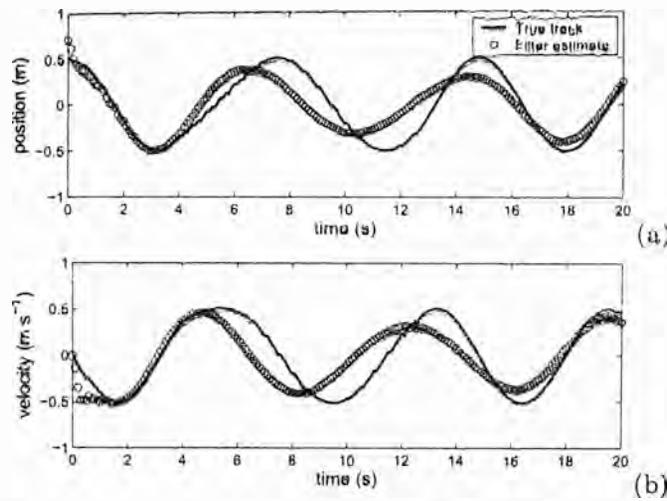


Figure 5.12: True oscillator motion parameters and estimates using the proposed filter.

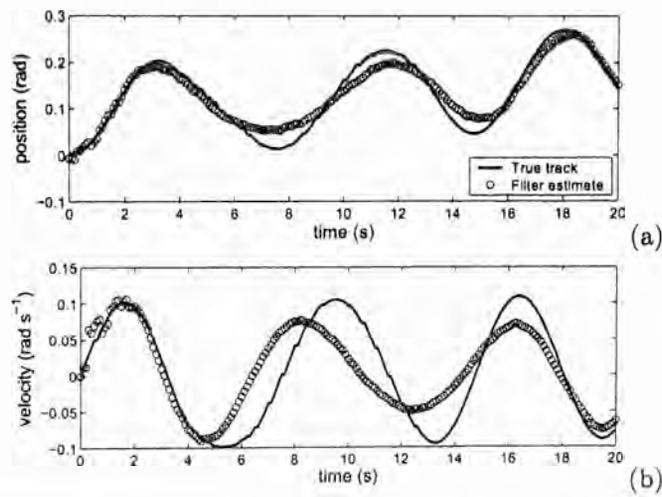


Figure 5.13: True actuator motion parameters and estimates using the proposed filter.

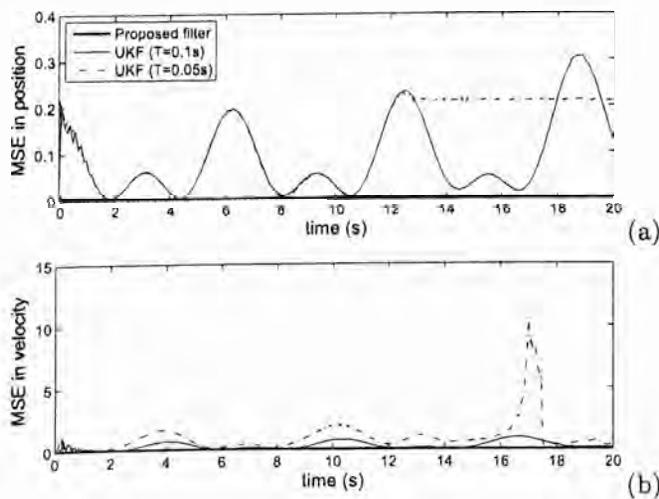


Figure 5.14: MSE in the estimates of oscillator motion parameters using the proposed filter and the UKF.

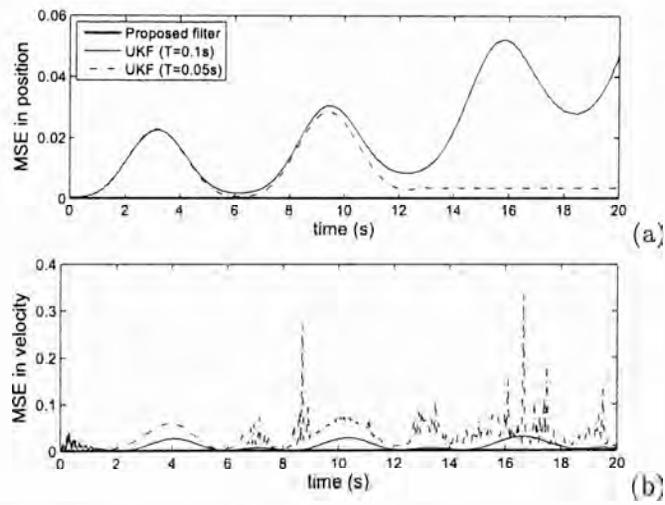


Figure 5.15: MSE in the estimates of actuator motion parameters using the proposed filter and the UKF.

Chapter 6

Conclusions

This dissertation addresses two open problems in estimation theory. First, the problem of multi-target filtering which involves jointly estimating the random number of targets and their state in the presence of noise, clutter, uncertainties in target maneuvers, data association and detection. At present there is no tractable analytical technique for tracking multiple targets under such general settings. Second, the problem of nonlinear filtering for a general class of systems motivated by the inadequacies of the existing analytic approximation based methods.

In Chapter 3, a multi-target model that accommodates births, deaths and switching linear Gaussian dynamics has been proposed based on random finite sets (RFS). For this so-called linear Gaussian jump Markov system (LGJMS) multi-target model, a closed form solution to the probability hypothesis density (PHD) recursion has been derived. The proposed algorithm eliminates the need to perform data association gating, track initiation and termination. Based on this solution, an efficient algorithm that can track an unknown, time-varying number of maneuvering targets in clutter has been developed. Extension of this algorithm to track maneuvering targets with non-linear jump Markov dynamics has also been proposed. In particular, approximate recursions for the weights, means and covariances of the Gaussian components that approximate the multi-target posterior intensity are given. The proposed approach is applicable to a general class of models expedient for a range of practical applications in multi-

target tracking that are deemed intractable using conventional techniques. Simulations have demonstrated the effectiveness of the proposed multi-target filters for tracking an unknown and time-varying number of maneuvering targets in clutter and detection uncertainty. In comparison with the well-known IMMJPDA filter, the proposed approach exhibits an unprecedented combination of good tracking performance and high computational efficiency.

The nonlinear JMS-PHD filter based on the unscented transform proposed in Chapter 3 gives estimates with reasonable accuracy at most times and occasionally misestimates the number of targets partially due to errors in the prediction and update steps which cannot be computed exactly for nonlinear models. In Chapter 4, an alternative analytic approximation of the nonlinear Bayes filter has been proposed based on the linear fractional transformation (LFT) model. The LFT system comprises of a linear part and a simple nonlinear structure in the feedback path. By applying the unscented transformation in the feedback loop a closed form solution to Bayes recursion has been derived. Simulation results have demonstrated that the proposed approach works better than standard filtering techniques such as the unscented Kalman filter (UKF) which linearizes the state space model. For highly nonlinear problems where the UKF breaks down, the proposed filtering technique performs reasonably well. The performance of the JMS-PHD filter for the nonlinear sensor model expressed in the LFT format is shown to be more robust than that using the UKF. The LFT system gives an equivalent representation for nonlinear state space models which is exact for a wide range of nonlinear systems. Extension of the proposed filtering approach for a general class of nonlinear problems has also been proposed using the nonlinear fractional transformation (NFT) model.

The potential of the LFT framework for nonlinear Bayesian filtering is fully recognized when the discussion is continued in Chapter 5 for the continuous-time stochastic process with sampled-data observations. The filtering problem is transferred to that of

discrete-time filtering and the moment propagation is based on the closed form solution to Bayes recursion presented in Chapter 4. Simulation results show a marked difference in the performance of the proposed method and the UKF in the continuous-discrete setting in terms of tracking error and computational complexity.

Following the discussion in this dissertation two interesting problems for further research are proposed. The motivations concerning for which are then presented.

In a multi-sensor scenario, suppose that $s_{k,1}, \dots, s_{k,L(k)} \in \mathcal{S}$ are the active sensors at time k , where L denotes a random process of the number of active sensors. At the sensor $s_{k,i}$, $M_i(k)$ measurements $Z_{k,i} = z_{k,1}^{(i)}, \dots, z_{k,M(k)}^{(i)} \in \mathbb{R}^m$ are received at time k . Find the estimate of the multi-target state X_k based on the measurements $Z_k = \{Z_{k,1}, \dots, Z_{k,L}\}$.

Suppose the LFT model (5.70)-(5.73) exists, find a scaling P_k in the feedback (5.73), $w_{\Delta k} = \Delta(x_k)P_k z_{\Delta k}$ such that $\mathbf{E}(w_{\Delta k} w_k^T) = 0$ and $\mathbf{E}(w_{\Delta k} v_k^T) = 0$, where $\{w_k\}$ and $\{v_k\}$ denote the noise processes of the plant and observation respectively.

The existence of the conditional intensity in Propositions 3.11 and 3.14 is based on the simplicity of the multi-target observation random finite set (RFS), that with probability 1 there are no coincidences among the observations. In the multi-sensor environment, single targets generate multiple observations. Propositions 3.11 and 3.14 do not generalize to such problems. Even in the particular case that at any time for a given target, only one sensor is active, under Poisson assumption on the predicted multi-target RFS, the joint probability generating functional of X_k and Z_k involves product of s functionals in the exponent and is computationally complex in general.

Linear estimation for filtering nonlinear models involves approximation of auto-covariance and cross-covariance of all concerned state and observation random variables. The existence of a scaling P_k satisfying the conditions above implies that the expressions for the covariance (4.17) and (4.24) are exact.

Appendix A

Proofs and Definitions

A.1 Proof of Lemma 2.6

Proof. Let G_X be the probability generating functional (p.g.fl.) of a point process $X \in \mathcal{N}_\lambda^s$ for any Borel measurable function g satisfying the condition in (2.20). Then the m -th functional derivative of G_X w.r.t. g , evaluated about the origin gives the distribution $P_X(\mathbf{x})$ of $\mathbf{x} = \{x_1, \dots, x_m\}$,

$$P_X(\mathbf{x}) = (d^m G_X)_0[\zeta_1, \dots, \zeta_m].$$

If $P_X(\cdot)$ admits a density w.r.t. Lebesgue measure μ , then

$$P_X(d\mathbf{x}) = (d^m G_X)_0[\delta_{x_1}, \dots, \delta_{x_m}], \quad (\text{A.1})$$

where δ_{x_i} is the Dirac delta function at point x_i . Let $\mathbf{y} = \{y_1, \dots, y_n\}$ be a realization of a point process $Y \in \mathcal{N}_\lambda^s$. Using $G_{XY}[g, h] = \mathbf{E}_X(G_{Y|X}[h|\mathbf{x}]\Pi_{\mathbf{x}}[g])$ from (2.27) and differentiating

$$(d^n G_{XY}[g, \cdot])_0[\delta_{y_1}, \dots, \delta_{y_n}] = \mathbf{E}_X((d^n G_{Y|X})_0[\delta_{y_1}, \dots, \delta_{y_n}]\Pi_{\mathbf{x}}[g]). \quad (\text{A.2})$$

Applying the definition of expectation and the conditional density for the differentiation as in (A.1),

$$(d^n G_{XY}[g, \cdot])_0[\delta_{y_1}, \dots, \delta_{y_n}] = \int \int P_{Y|X}(dy|\mathbf{x})P_X(d\mathbf{x})\Pi_{\mathbf{x}}[g]. \quad (\text{A.3})$$

From (2.30),

$$\begin{aligned} (d^n G_{XY}[1, \cdot])_0[\delta_{y_1}, \dots, \delta_{y_n}] &= (d^n G_Y)_0[\delta_{y_1}, \dots, \delta_{y_n}] \\ &= P_Y(dy) \end{aligned} \tag{A.4}$$

Substituting (A.3) and (A.4) on the right hand side of (2.31) and applying Bayes rule, $P_{X|Y}(d\mathbf{x}|\mathbf{y}) = \frac{P_{Y|X}(d\mathbf{y}|\mathbf{x})P_X(d\mathbf{x})}{P_Y(dy)}$, yields $\int P_{X|Y}(d\mathbf{x}|\mathbf{y})\Pi_{\mathbf{x}}[g]$ which is the definition of the conditional p.g.fl. of $X|Y$. This completes the proof. \square

A.2 Functional derivative

Given a test function ϕ and a functional $F : \phi \mapsto F[\phi]$, the functional derivative of F , denoted $(dF)_\phi$ is a distribution $dF[\phi]$,

$$(dF)_\phi[f] = \lim_{\epsilon \rightarrow 0} \frac{F[\phi + \epsilon \cdot f] - F[\phi]}{\epsilon}.$$

Bibliography

- [1] B. Anderson and J. Moore. *Optimal Filtering*. Prentice-Hall, 1979.
- [2] P. Apkarian. On the discretization of LMI-synthesized linear parameter-varying controllers. *Automatica*, 33(4):655–661, 1997.
- [3] P. Apkarian and R.J. Adams. Advanced gain-scheduling technique for uncertain systems. *IEEE Trans. Control System Tech.*, 6(1):21–32, 1997.
- [4] P. Apkarian and P. Gahinet. A convex characterization of gain-scheduled \mathcal{H}_∞ controllers. *IEEE Trans. Automatic Control*, 40(5):853–864, 1995.
- [5] Y. Bar-Shalom. Extension of the probabilistic data association filter to multitarget environment. In *Proc. 5th Symp. Nonlinear Estimation*, pages 16–21, September 1974.
- [6] Y. Bar-Shalom, K.C. Chang, and H.A.P. Blom. Tracking splitting targets in clutter by using an interacting multiple model joint probabilistic data association filter. In Y. Bar-Shalom, editor, *Multitarget Multisensor Tracking: Applications and Advances*, Vol. II, part 3, pages 93–110. Artech House, 1992.
- [7] Y. Bar-Shalom and T. E. Fortmann. *Tracking and Data Association*. Academic Press, San Diego, 1988.
- [8] Y. Bar-Shalom and A.J. Jaffer. Adaptive nonlinear filtering for tracking with measurements of uncertain origin. In *Proc. IEEE Conf. Decision and Control*, pages 243–247, December 1972.
- [9] Y. Bar-Shalom and X.-R. Li. *Multitarget-Multisensor Tracking: Principles and Techniques*. Storrs, CT: YBS Publishing, 1995.
- [10] Y. Bar-Shalom, X.-R. Li, and T. Kirubarajan. *Estimation with Application to Tracking and Navigation*. Wiley, 2001.
- [11] Y. Bar-Shalom and E. Tse. Tracking in a cluttered environment with probabilistic data association. *Automatica*, 11(4):451–460, September 1975.
- [12] R. Beard, J. Kenney, J. Gunther, J. Lawton, and W. Stirling. Nonlinear projection filter based on Galerkin approximation. *Journal of Guidance, Control, and Dynamics*, 22(2):258–266, March-April 1999.

- [13] V. E. Beneš. Exact finite-dimensional filters for certain diffusions with nonlinear drift. *Stochastics*, 5:65–92, 1981.
- [14] V.E. Beneš. New exact nonlinear filters with large Lie algebras. *Systems and Control Letters*, 5(4):217–221, February 1985.
- [15] R.E. Bethel and G.J. Paras. A PDF multitarget-tracker. *IEEE Trans. Aerospace and Electronic Systems*, 30(2):386–403, 1994.
- [16] S.S. Blackman. *Multiple Target Tracking with Radar Applications*. Norwood, MA: Artech House, 1986.
- [17] S.S. Blackman and R. Popoli. *Design and analysis of modern tracking systems*. Norwood, MA: Artech House, 1999.
- [18] H. Blom and Y. Bar-Shalom. The interacting multiple model algorithm for systems with Markovian switching coefficients. *IEEE Trans. Automatic Control*, AC-33:780–783, August 1988.
- [19] H.A.P. Blom and E.A. Bloem. Combining IMM and JPDA for tracking multiple maneuvering targets in clutter. In *Proc. 5th International Conf. Information Fusion*, pages 705–712, Cairns, Australia, 2002.
- [20] Pierre Brémaud. *Point processes and queues*. Springer-Verlag, 1981.
- [21] D. Brigo, B. Hanzon, and F. LeGland. A differential geometric approach to nonlinear filtering: The projection filter. *IEEE Trans. Automatic Control*, 43(2):247–252, February 1998.
- [22] R.S. Bucy and P.D. Joseph. *Filtering for Stochastic Processes with Applications to Guidance*. John Wiley and Sons, 1968.
- [23] R.T. Bupp, D.S. Bernstein, and V.T. Coppola. A benchmark problem for nonlinear control design: Problem, experimental testbed, and passive nonlinear compensation. In *Proc. American Control Conf.*, pages 4363–4367, 1995.
- [24] Z. Cai, F. Le Gland, and H. Zhang. An adaptive local grid refinement method for nonlinear filtering. Internal Report 954, IRISA Campus de Beaulieu, 1995.
- [25] G. Calafiore and L. El Ghaoui. Robust maximum likelihood estimation in the linear model. *Automatica*, 37(4):573–580, 2001.
- [26] J. Carpenter, P. Clifford, and P. Fearnhead. Improved particle filter for nonlinear problems. In *IEE Proc. Radar, Sonar, and Navigation*, number 1 in Vol. 146, pages 2–7, February 1999.
- [27] S. Challa and Y. Bar-Shalom. Nonlinear filtering design using Fokker-Planck-Kolmogorov probability density evolutions. *IEEE Trans. Aerospace and Electronic Systems*, 36(1):309–315, January 2000.
- [28] S. Challa, Y. Bar-Shalom, and V. Krishnamurthy. Nonlinear filtering via generalized Edgeworth series and Gauss-Hermite quadrature. *IEEE Trans. Signal Processing*, 48(6):4697–4706, 2000.

- [29] B. Chen and J.K. Tugnait. Tracking of multiple maneuvering targets in clutter using IMM/JPDA filtering and fixed-lag smoothing. *Automatica*, 37:239–249, February 2001.
- [30] R. Chen and J.S. Liu. Mixture Kalman filters. *Journal of Royal Statistical Society*, 62(3):493–508, 2000.
- [31] D. Clark and J. Bell. Bayesian multiple target tracking in forward scan sonar images using the PHD filter. in *IEE Radar, Sonar and Navigation*, 152(5):327–334, October 2005.
- [32] D.R. Cox and V. Isham. *Point processes*. Chapman and Hall, 1980.
- [33] N. Cui, L. Hong, and J.R. Layne. A comparison of nonlinear filtering approaches with an application to ground target tracking. *Signal Processing*, 85(8):1469–1492, 2005.
- [34] D.J. Daley and D. Vere-Jones. *An introduction to the theory of point processes*. Springer Verlag, 1988.
- [35] F. Daum. New exact nonlinear filters. In J.C. Spall, editor, *Bayesian Analysis of Time Series and Dynamic Models*, pages 199–226. Marcel Dekker, New York, 1988.
- [36] F.E. Daum. Exact finite-dimensional nonlinear filters for continuous time processes with discrete time measurements. In *Proc. IEEE Conf. Decision and Control*, pages 16–22, 1984.
- [37] F.E. Daum. Exact finite-dimensional nonlinear filters. *IEEE Trans. Automatic Control*, 37(7):616–622, July 1986.
- [38] R.J. Dempster, S.S. Blackman, and T. Nichols. Combining IMM filtering and MHT data association for multitarget tracking. In *Proc. 29th Southeastern Symposium*, 1997.
- [39] A. Doucet, C. Andrieu, and S. Godsill. On sequential Monte Carlo sampling methods for Bayesian filtering. *Statistics and Computing*, 10(3):197–208, 2000.
- [40] A. Doucet, N. de Freitas, , and N. Gordon. *Sequential Monte Carlo Methods in Practice*. Springer Verlag, 2001.
- [41] T.E. Fortmann, Y. Bar-Shalom, and M. Scheffe. Multi-target tracking using joint probabalistic data association. In *Proc. IEEE Conf. Decision and Control*, pages 1–6, December 1980.
- [42] T.E. Fortmann, Y. Bar-Shalom, and M. Scheffe. Sonar tracking of multiple targets using joint probabilistic data association. *IEEE Journal of Oceanic Engineering*, OE-8:174–184, July 1983.
- [43] T.E. Fortmann and S. Baron. Problems in multi-target sonar tracking. In *Proc. IEEE Conf. Decision and Control*, pages 1182–1188, December 1979.
- [44] A. Gelb. *Applied optimal estimation*. The MIT Press, 1974.

- [45] I. Goodman, R. Mahler, and H. Nguyen. *Mathematics of Data Fusion*. Kluwer Academic Publishers, 1997.
- [46] N. Gordon, D. Salmond, and A.F.M. Smith. Novel approach to nonlinear/non-gaussian bayesian state estimation. In *Proc. IEE*, number 2 in Vol. 140, pages 107–113, 1993.
- [47] N. Gordon, D. Salmond, and A.F.M. Smith. Bayesian state estimation for tracking and guidance using the bootstrap filter. *Journal of Guidance, Control and Dynamics*, 18(6):1434–1443, 1995.
- [48] R. B. Guenther and J. W. Lee. *Partial Differential Equations of Mathematical Physics and Integral Equations*. Dover Publications, Inc., 1988.
- [49] J. Gunther, R. Beard, J. Wilson, T. Oliphant, and W. Stirling. Fast nonlinear filtering via Galerkins method. In *Proc. American Control Conf.*, pages 2088–2092, 1997.
- [50] R. Gurumoorthy and S.R. Sanders. Controlling nonminimum phase nonlinear systems-The inverted pendulum on a cart example. In *Proc. American Control Conf.*, pages 680–685, 1993.
- [51] N.T. Hoang, H.D. Tuan, P. Apkarian, and S. Hosoe. Gain-scheduled filtering for time-varying discrete systems. *IEEE Trans. Signal Processing*, 52(9):2464–2476, 2004.
- [52] N.T. Hoang, H.D. Tuan, P. Apkarian, and S. Hosoe. Robust filtering for discrete nonlinear fractional transformation systems. *IEEE Trans. Circuits and Systems II*, 51(11):587–592, 2004.
- [53] N.T. Hoang, H.D. Tuan, T.Q. Nguyen, and S. Hosoe. Robust mixed generalized $\mathcal{H}_2/\mathcal{H}_\infty$ filtering of 2-D nonlinear fractional transformation systems. *IEEE Trans. Signal Processing*, 53(12):4697–4706, 2005.
- [54] E.L. Ionides. Inference and filtering for partially observed diffusion processes via sequential Monte Carlo. Tech. Rep. 405, Univ. of Michigan Statistics Dept., 2004.
- [55] A. Isidori. *Nonlinear Control Systems. An Introduction*. Springer Verlag, 1989.
- [56] A. Isidori. *Nonlinear Control Systems*. Springer Verlag, 2001.
- [57] K. Ito and K. Xiong. Gaussian filters for nonlinear filtering problems. *IEEE Trans. Automatic Control*, 45(5):910–927, May 2000.
- [58] A.H. Jazwinski. Filtering for nonlinear dynamical systems. *IEEE Trans. Automatic Control*, AC-11(4):765–766, 1966.
- [59] A.H. Jazwinski. *Stochastic processes and filtering theory*. Academic Press, 1970.
- [60] S. Jeong and J. K. Tugnait. Tracking of multiple maneuvering targets in clutter with possibly unresolved measurements using IMM and JPDAM coupled filtering. In *Proc. American Control Conf.*, pages 1257–1262, 2005.

- [61] A. Johansen, S. Singh, A. Doucet, and B. Vo. Convergence of the sequential Monte Carlo implementation of the PHD filter. to appear in *Methodology and Computing in Applied Probability*, 2006.
- [62] S. J. Julier and J. K. Uhlmann. A new extension of the Kalman filter to nonlinear systems. In *Proc. AeroSense: 11th International Symp. on Aerospace/Defence Sensing, Simulation and Controls*, Orlando, Florida, 1997.
- [63] S. J. Julier and J. K. Uhlmann. Unscented filtering and nonlinear estimation. In *Proc. IEEE*, number 3, pages 401–422, 2004.
- [64] G. Kallianpur. *Stochastic Filtering Theory*. Springer-Verlag, 1980.
- [65] R.E. Kalman. A new approach to linear filtering and prediction problems. *Transactions on ASME, Journal of Basic Engineering*, 82:34–45, March 1960.
- [66] R.E. Kalman and R.S. Bucy. New results in linear filtering and prediction theory. *Transactions on ASME, Journal of Basic Engineering*, 83:95–108, 1961.
- [67] A.F. Karr. *Point processes and their statistical inference*. Marcel Dekker Inc., 1991.
- [68] K. Kastella. Finite difference methods for nonlinear filtering and automatic target recognition. In Y. Bar-Shalom and W.D. Blair, editors, *Multitarget-Multisensor Tracking: Applications and Advances*, Vol. III, pages 233–258. Artech House, Norwood, MA, 2000.
- [69] K. Kastella and A. Zatezalo. Nonlinear filtering for low elevation targets in the presence of multipath propagation. In *Proc. SPIE Conf. Signal and Data Processing of Small Targets*, number 3, pages 452–459, 1998.
- [70] H. Khalil. *Nonlinear Systems*. Prentice Hall, 2002.
- [71] J.F.C. Kingman. *Poisson processes*. Clarendon Press, 1993.
- [72] P.E. Kloeden and E. Platen. *Numerical Solution to Stochastic Differential Equations*. Springer, 1999.
- [73] W. Koch. Fixed-interval retrodiction approach to Bayesian IMM-MHT for maneuvering multiple targets. *IEEE Trans. Aerospace and Electronic Systems*, 36(1):2–14, 2000.
- [74] M.A. Kouritzin. On exact filters for continuous signals with discrete observations. *IEEE Trans. Automatic Control*, AC-43(5):709–715, 1998.
- [75] A.J. Krener and A. Duarte. A hybrid computational approach to nonlinear estimation. In *Proc. IEEE Conf. Decision and Control*, pages 1815–1819, 1996.
- [76] H.J. Kushner. On the differential equations satisfied by conditional probability densities of Markov processes, with applications. *Journal of SIAM Control Series A*, 2(1), 1964.
- [77] T. Lefebvre, H. Bruyninckx, and J. D. Schutter. Kalman filters for nonlinear systems: A comparison of performance. Internal Report 01R033, ME Dept. Katholieke Universiteit Leuven, Belgium, 2001.

- [78] T. Lefebvre, H. Bruyninckx, and J. D. Schutter. Comment on "A new method for the nonlinear transformation of means and covariances in the filters and estimators". *IEEE Trans. Automatic Control*, AC-47(8):1406–1409, 2002.
- [79] C. W. Leung and S. S. T. Yau. Recent results on classification of finite dimensional rank estimation algebras with state space dimension 3. In *Proc. IEEE Conf. Decision and Control*, pages 2247–2250, 1992.
- [80] X.-R. Li and Y. Bar-Shalom. Design of an interacting multiple algorithm for air traffic control tracking. *IEEE Transactions on Control System Technology*, pages 186–194, September 1993.
- [81] X.-R. Li and Y. Bar-Shalom. Multiple model estimation with variable structure. *IEEE Trans. Automatic Control*, AC-41(4):478–493, April 1996.
- [82] X.-R. Li and V. Jilkov. Survey of maneuvering target tracking. Part 1: Dynamic models. *IEEE Trans. Aerospace and Electronic Systems*, 39(4):1333–1364, 2003.
- [83] X.-R. Li and V. Jilkov. Survey of maneuvering target tracking. Part V: Multiple-model methods. *IEEE Trans. Aerospace and Electronic Systems*, 41(4):1255–1321, 2005.
- [84] J. Lo. Finite dimensional sensor orbits and optimal nonlinear filtering. *IEEE Trans. Information Theory*, IT-18(5):583–588, 1972.
- [85] S. Lototsky, R. Milulevicius, and B.E. Rozovskii. Nonlinear filtering revisited: A spectral approach. *SIAM Journal on Control and Optimization*, 35(2):435–461, March 1997.
- [86] W.K. Ma, B. Vo, S. Singh, and A. Baddeley. Tracking an unknown and time varying number of speakers using TDOA measurements: A random finite set approach. *IEEE Trans. Signal Processing*, 54(9), 2006.
- [87] J. MacCormick and A. Blake. A probabilistic exclusion for tracking multiple objects. *International Journal of Computer Vision*, 39(1):57–71, 2000.
- [88] R. Mahler. An introduction to multisource-multitarget statistics and applications. Monograph, Lockheed Martin Technical, 2000.
- [89] R. Mahler. A theoretical foundation for the Stein-Winter Probability Hypothesis Density (PHD) multi-target tracking approach. In *Proc. MSS National Symp. on Sensor and Data Fusion*, Vol. I (Unclassified), San Antonio, TX, June 2000.
- [90] R. Mahler. Multi-target moments and their application to multi-target tracking. In *Proc. Workshop on Estimation, Tracking and Fusion: A tribute to Yaakov Bar-Shalom*, pages 134–166, Monterey, 2001.
- [91] R. Mahler. Random set theory for target tracking and identification. In D. Hall and J. Llinas, editors, *Data Fusion Hand Book*, pages 14/1–14/33. CRC press Boca Raton, 2001.
- [92] R. Mahler. Multi-target Bayes filtering via first-order multi-target moments. *IEEE Trans. Aerospace and Electronic Systems*, 39(4):1152–1178, 2003.

- [93] G. Mathéron. *Random Sets and Integral Geometry*. J. Wiley, 1975.
- [94] M.I. Miller, A. Srivastava, and U. Grenander. Conditional-mean estimation via jump-diffusion processes in multiple target tracking/recognition. *IEEE Trans. Signal Processing*, 43(11):2678–2690, 1995.
- [95] M.I. Miller, R.S. Teichman, A. Srivastava, J.A. O’Sullivan, and D.L. Snyder. Jump-diffusion processes for automated tracking-target recognition. In *Proc. Conf. on Information Science and Systems*, Baltimore, Maryland, March 1993.
- [96] D. Mirković. N-dimensional finite element package. Dept. Math., Iowa State Univ.
- [97] I. Molchanov. *Theory of random sets*. Springer-Verlag, 2005.
- [98] P.D. Moral and L. Miclo. *Branching and interacting particle systems. Approximations of Feynman-Kac formulae with applications to non-linear filtering*. Séminaire de Probabilités XXXIV. Springer-Verlag, 2000.
- [99] S. Mori, C. Chong, E. Tse, and R. Wishner. Tracking and identifying multiple targets without apriori identifications. *IEEE Trans. Automatic Control*, AC-21:401–409, 1986.
- [100] H. Salehi N. Portenko and A. Skorokhod. On optimal filtering of multitarget systems based on point process observations. *Random Operators and Stochastic Equations*, 5(1):1–34, 1997.
- [101] M. Nørgaard, N.K. Poulsen, and O. Ravn. New developments in state estimation for nonlinear systems. *Automatica*, 36(11):1627–1638, November 2000.
- [102] L. Nyland, J. Prins, A. Goldberg, P. Mills, J. Reif, and R. Wagner. A refinement methodology for developing data-parallel applications. In *Proc. EuroPar’96, Lecture Notes in Computer Science*, Springer Verlag, 1996.
- [103] A. Pasha, B. Vo, H.D. Tuan, and W.K. Ma. Closed form PHD filtering for linear jump Markov models. In *Proc. 9th International Conf. Information Fusion*, pages 1–8, Florence, Italy, 2006.
- [104] S.A. Pasha and H.D. Tuan. Closed form filtering for linear fractional transformation models. In *Proc. 17th IFAC World Congress*, pages 14510–14515, Seoul, Korea, 2008.
- [105] S.A. Pasha and H.D. Tuan. Nonlinear filtering for the linear fractional transformation model. In *Proc. 2nd International Conf. Communications and Electronics*, pages 180–185, Hoi An, Vietnam, 2008.
- [106] S.A. Pasha and H.D. Tuan. Nonlinear filtering for continuous-time systems using the linear fractional transformation model. In *Proc. ICASSP*, Taipei, Taiwan, 2009. To appear.
- [107] S.A. Pasha, H.D. Tuan, and P. Apkarian. The LFT based PHD filter for nonlinear jump Markov models in multi-target tracking. In *Proc. IEEE Conf. Decision and Control*, Shanghai, China, 2009. To appear.

- [108] S.A. Pasha, H.D. Tuan, and B. Vo. Nonlinear Bayesian filtering using the unscented linear fractional transformation model. Accepted for publication in *IEEE Trans. Signal Processing*, 2009.
- [109] S.A. Pasha, B. Vo, H.D. Tuan, and W.K. Ma. A Gaussian mixture PHD filter for jump Markov system models. To appear in *IEEE Trans. Aerospace and Electronic Systems*, 2009.
- [110] K. Punithakumar, T. Kirubarajan, and A. Sinha. A multiple model Probability Hypothesis Density filter for tracking maneuvering targets. In *O. E. Drummond (ed.) Signal and Data Processing of Small Targets, Proc. SPIE*, 5428, pages 113–121, 2004.
- [111] S.P. Puranik and J.K. Tugnait. Tracking of multiple maneuvering targets using multiscan JPDA and IMM filtering. *IEEE Trans. Aerospace and Electronic Systems*, 43:23–35, January 2007.
- [112] C.R. Rao. *Linear Statistical Inference and its Applications*. John Wiley and Sons, 1973.
- [113] D.B. Reid. A multiple hypothesis filter for tracking multiple targets in a cluttered environment. Report LMSC, D-560254, Lockheed Missiles and Space Company, September 1977.
- [114] D.B. Reid. An algorithm for tracking multiple targets. *IEEE Trans. Automatic Control*, AC-24:843–854, December 1979.
- [115] A. Romanenko and J.A.A.M. Castro. The unscented filter as an alternative to the EKF for nonlinear state estimation: A simulation case study. *Computers and Chemical Engineering*, 28(8):347–355, 2004.
- [116] T.S. Schei. A finite-difference method for linearization in nonlinear estimation algorithms. *Automatica*, 33(11):2053–2058, November 1997.
- [117] G.C. Schmidt. Designing nonlinear filters based on Daums theory. *Journal of Guidance, Control, and Dynamics*, 16(2):371–376, March-April 1993.
- [118] R.G. Sea. An efficient suboptimal decision procedure for associating sensor data with stored tracks in real-time surveillance systems. In *Proc. IEEE Conf. Decision and Control*, pages 33–37, December 1971.
- [119] H. Sidenbladh and S.L. Wirkander. Tracking random sets of vehicles in terrain. In *Proc. 2003 IEEE Workshop on Multi-Object Tracking*, Madison Wisconsin, 2003.
- [120] R.A. Singer and R.G. Sea. New results in optimizing surveillance system tracking and data correlation performance in dense multitarget environments. *IEEE Trans. Automatic Control*, AC-18(6):571–581, December 1973.
- [121] R.A. Singer and J.J. Stein. An optimal tracking filter for processing sensor data of imprecisely determined origin in surveillance system. *Proc. IEEE Conf. Decision and Control*, pages 171–175, December 1971.
- [122] D.L. Snyder. *Random point processes*. John Wiley & Sons, 1975.

- [123] A. Srivastava, M.I. Miller, and U. Grenander. Jump-diffusion processes for tracking and direction finding. In *Proc. 29th Allerton Conf. on Communication, Control, and Computing*, pages 563–570, University of Illinois Urbana, 1991.
- [124] M.C. Stein and C.L. Winter. An adaptive theory of probabilistic evidence accrual. Report LA-UR-93-3336, Los Alamos National Laboratories, 1993.
- [125] D. Stoyan, W.D. Kendall, and J. Mecke. *Stochastic Geometry and its Applications*. John Wiley & Sons, 1995.
- [126] R.L. Stratonovich. *Conditional Markov Processes and Their Application to the Theory of Optimal Control*. American Elsevier Publishing Company, Inc., 1968.
- [127] K. Sun and A. Packard. Robust \mathcal{H}_2 and \mathcal{H}_∞ filters for uncertain LFT systems. *IEEE Trans. Automatic Control*, AC-50(5):715–720, 2005.
- [128] H.D. Tuan, P. Apkarian, T. Narikiyo, and M. Kanota. New fuzzy control model and dynamic output feedback parallel distributed compensation. *IEEE Trans. Fuzzy Systems*, 12(1):13–21, 2004.
- [129] H.D. Tuan, P. Apkarian, and T.Q. Nguyen. Robust filtering for uncertain nonlinearly parameterized plants. *IEEE Trans. Signal Processing*, 51(7), 2003.
- [130] J.K. Tugnait. Tracking of multiple maneuvering targets in clutter using multiple sensors, IMM and JPDA coupled filtering. *IEEE Trans. Aerospace and Electronic Systems*, AES-40:320–330, January 2004.
- [131] B. Vo and W.K. Ma. A closed-form solution to the Probability Hypothesis Density filter. In *Proc. 8th International Conf. Information Fusion*, pages 792–799, Philadelphia, 2003.
- [132] B. Vo and W.K. Ma. Joint detection and tracking of multiple maneuvering targets using random finite sets. in *Proc. ICARCV*, Kunming, China, 2004.
- [133] B. Vo and W.K. Ma. The Gaussian mixture Probability Hypothesis Density filter. *IEEE Trans. Signal Processing*, 54(11):4091–4104, 2006. <http://www.ee.unimelb.edu.au/staff/bv/index.html>.
- [134] B. Vo, W.K. Ma, and S. Singh. Locating an unknown time-varying number of speakers: A Bayesian random finite set approach. In *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, 4, pages 1073–1076, Philadelphia, 2005.
- [135] B. Vo, A. Pasha, and H.D. Tuan. A gaussian mixture PHD filter for nonlinear jump markov models. In *Proc. IEEE Conf. Decision and Control*, pages 3162–3167, San Diego, USA, 2006.
- [136] B. Vo and S. Singh. Technical aspects of the Probability Hypothesis Density recursion. Tech. Rep. TR05-006, EEE Dept. The University of Melbourne, Australia, 2005.
- [137] B. Vo, S. Singh, and A. Doucet. Sequential Monte Carlo implementation of the PHD filter for multi-target tracking. In *Proc. 6th International Conf. Information Fusion*, pages 792–799, Cairns, Australia, 2003.

- [138] B. Vo, S. Singh, and A. Doucet. Sequential Monte Carlo methods for multi-target filtering with random finite sets. *IEEE Trans. Aerospace and Electronic Systems*, 41(4):1224–1245, 2005. <http://www.ee.unimelb.edu.au/staff/bv/index.html>.
- [139] D. Wang and J. Huang. A neural network based method for solving discrete-time nonlinear output regulation problem in sampled-data systems. In *Advances in Neural Networks - ISNN 2004*, number 3174 in Lecture Notes in Computer Science, pages 59–64. Springer, 2004.
- [140] R. Washburn. A random point process approach to multi-object tracking. In *Proc. American Control Conf.*, vol. 3, pages 1846–1852, March 1987.
- [141] G. Wolodkin, S. Rangan, and K. Poolla. An LFT approach to parameter estimation. In *Proc. American Control Conf.*, pages 2088–2092, 1997.
- [142] M. Zakai. On the optimal filtering of diffusion processes. *Zeit. Wahrsch.*, 11:230–243, 1969.
- [143] K. Zhou, J.C. Doyle, and K. Glover. *Robust and Optimal Control*. Prentice Hall, 1996.

FN 4B

