Finite Point Processes and Their Application to Target Tracking

Dissertation
zur
Erlangung des Doktorgrades (Dr. rer. nat)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von
Dipl.-Math. Christoph Degen
aus
Neuwied

Bonn, 2015

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter: Priv.-Doz. Dr. Wolfgang Koch
2. Gutachter: Prof. Dr. Reinhard Klein

Tag der Promotion: 19.12.2016
Erscheinungsjahr: 2017

Zusammenfassung

Tracking ist ein Teilgebiet der Sensoratenfusion und beschäftigt sich mit der Verfolgung dynamischer Objekte auf Basis unvollständiger und mit Fehler behafteter Messungen. Auf Grund der enormen Nachfrage nach leistungsfähigen Algorithmen wurden in den vergangenen Jahrzehnten eine Vielzahl von Methoden und Verfahren in diesem Teilgebiet der angewandten Informatik entwickelt. Die Vielfalt und Verflechtung der existierenden Konzepte befähigt einerseits zur Lösung von Szenarien mit unterschiedlichsten Randbedingungen, kann andererseits aber nur von den wenigsten Wissenschaftlern vollständig durchschaut werden. Eine Vereinheitlichung der bestehenden Trackingverfahren ist daher für ein tiefgehendes Verständnis dieses Forschungsfeldes von großer Bedeutung.

Die Theorie der Punktprozesse ist ein hochentwickeltes Werkzeug der Wahrscheinlichkeitstheorie und Statistik, das genau wie die Theorie der stochastischen Prozesse eine Vielzahl von Anwendungen in der Finanz- und Wirtschaftsmathematik, Biologie und Physik erfährt. Punktprozesse eignen sich in besonderer Weise zur Modellierung von dynamischen Objekten und Sensormessungen aus Trackinganwendungen, da sowohl die Anzahl als auch die räumliche Verteilung der entsprechenden Elemente nachgebildet werden kann. Das den Punktprozess eindeutig und vollständig charakterisierende wahrscheinlichkeitserzeugende Funktional bietet sich auf Grund seiner kompakten Form als Repräsentant eines Trackingfilters und somit als Grundlage für die einheitliche Darstellung von Trackingverfahren an, da alle notwendigen statistischen Informationen über das Filter auf eine intuitiv verständliche Art und Weise verschlüsselt und zusammengefasst in ihm vorliegen.

In dieser Dissertation wird die Entwicklung, Charakterisierung und Vereinheitlichung von Trackingverfahren mit Hilfe von finiten Punktprozessen untersucht und auf die passive, nicht kooperative Lokalisierung und Verfolgung von elektromagnetischen Emittern im städtischen Gebiet mit Hilfe eines mobilen Antennenarrays angewendet.

Der erste Teil dieser Dissertation erarbeitet ein theoretisches Fundament für den Einsatz von finiten Punktprozessen zur Vereinheitlichung und Herleitung von Trackingfiltern. Es wird gezeigt, dass sich viele bekannte Trackingfilter mit Hilfe von wahr-

scheinlichkeitserzeugenden Funktionalen charakterisieren lassen und die verschiedenen Bausteine eines Trackingfilters werden an Hand seines wahrscheinlichkeiterzeugenden Funktionals erklärt. In besonderer Weise eignet sich die vorgestellte Darstellung dazu, die Gemeinsamkeiten und Unterschiede zwischen einzelnen Filtern hervorzuheben und ein grundlegendes Verständnis für die bestehenden Trackingkonzepte zu entwickeln. Weiterhin wird gezeigt, dass sich individuelle und maßgeschneiderte Trackingfilter leicht mit Hilfe der vorgestellten Theorie modellieren lassen. Damit bietet die entwickelte Zusammenfassung die Chance für jeden praktisch arbeitenden Ingenieur der Sensordatenfusion auf eine intuitive Art und Weise innovative Lösungen für aktuelle Fragestellungen des Trackings zu finden. Den Abschluss des ersten Teils bildet die umfassende mathematische Fundierung der Herleitung von Trackingfiltern aus der vorgestellten Vereinheitlichung.

Der zweite Teil dieser Arbeit behandelt die passive, nicht kooperative Lokalisierung und Verfolgung von elektromagnetischen Emittern im urbanen Umfeld mit Hilfe eines einzelnen mobilen Antennenarrays. Dazu werden die im ersten Teil diskutierten Trackingfilter durch sequentielle Monte–Carlo Verfahren implementiert und auf Basis der Problemstellung weiterentwicklet. Auf Grund des urbanen Umfelds teilt sich das radial emittierte Signal durch Reflektion, Beugung und Streuung in mehrere elektromagnetische Wellen auf, den sogenannten Mehrwegen, die entlang verschiedener Pfade zum Empfänger gelangen. Die Algorithmen der Sensordatenfusion werden durch diese Nebenbedingung vor die Herausforderung gestellt, dass ein Ziel mehrere Messungen erzeugt, die im Messraum räumlich nicht zusammengefasst werden können. Vor diesem Hintergrund werden sogenannte PHD und Intensitätsfilter erstmals auf das vorgestellte Szenario angewendet und umfassend untersucht. Die unterschiedlichen Messmodelle, die dabei zum Einsatz kommen erfordern die Entwicklung verschiedenster komplexer Verfahren bevor eine Anwendung der Filter erfolgen kann. Das Standard-Messmodell trifft die Annahme, dass ein Ziel höchstens eine Messung erzeugt. Aus diesem Grund ist die Entwicklung von leistungsfähigen Extraktionsverfahren für den Zielzustands und neuartigen Likelihood Funktionen zur Bewertung einzelner empfangener Mehrwege erforderlich. Aufwendige numerische Approximationen und Verwerfungsstrategien von Mess-Partitionen sind notwendig, wenn das verallgemeinerte Messmodell angewendet wird. Aus diesem Grund werden verschiedene Kriterien zur Reduktion der numerischen Komplexität unter Verwendung der im ersten Teil vorgestellten Theorie hergeleitet und analysiert. Die Anwendbarkeit der entwickelten Verfahren wird dabei sowohl in Simulations- als auch in Realdatenszenarien demonstriert. Den Abschluss bildet die Entwicklung eines neuartigen Multi-Hypothesen basierten Parameter–Trackingverfahrens für relative Laufzeiten, das Falschmessungen zurückweist, die vor dem ersten Ziel-basierten Mehrweg empfangen werden und auch bei einem hohen Anteil von Falschmessungen, Messausfällen und Sensorrauschen eine zuverlässige Lokalisierung und Verfolgung ermöglicht.

Acknowledgments

First of all, I would like to express my sincere gratitude to Wolfgang Koch for the opportunity to write a thesis as a combination of theory and practice in the department of sensor data and information fusion at the Fraunhofer FKIE in Wachtberg/Bonn and supporting me in all concerns.

I cordially thank Prof. Dr. Reinhard Klein for being the second supervisor and giving me the chance to write this thesis at the Fraunhofer FKIE and the University of Bonn.

I would like to thank Prof. Dr. Peter Martini and Prof. Dr. Patrik Ferrari for their contribution to the supervisory board.

I am deeply grateful to Wolfgang Koch that he introduced me to Roy Streit, who I would like to thank from the bottom of my heart for the many fruitful discussions on the first part of this thesis.

I especially thank Felix Govaers for being my mentor in multitarget tracking and introducing me to the topic of blind mobile localization. His encouragement to follow my ideas, support in programming, countless discussions and an always open door made this thesis possible.

Many thanks go to all my colleagues in the department of sensor data and information fusion at the Fraunhofer FKIE. In particular, I would like to thank Daniel Bender, Julian Hörst, Marek Schikora, Klaus Wild, Marianne Wilms, Martin Michaelis, Hichem El Mokni, Michael Mertens, Alex Charlish and Folker Hoffmann.

I am grateful to AWE Communications from Böblingen/Germany for providing the latest version of their ray tracing simulation "WinProp" to evaluate the data fusion algorithms of the second part of this thesis. In particular, I thank Reiner Hoppe for his support in using the ray–tracer and his help to understand the theory behind.

I would like to thank Stephan Häfner, Martin Käske and Reiner Thomä from the Technical University of Ilmenau for providing a simulation of the antenna array used

within the numerical evaluations of the simulated data and the blind channel estimation algorithms applied to the real world scenario in Chapter 6.

Many thanks go to Saab Medav Technologies GmbH for carrying out the real world experiment presented in Section 6.4. In particular, I thank Alexis Paolo Garcia Ariza and Uwe Trautwein for their support.

The second part of this thesis was supported by the Federal Ministry of Education and Research of Germany (BMBF), within the Project EiLT: “Emitter Lokalisierung unter Mehrwegeausbreitungsbedingungen”, http://eilt.medav.de/.

Contents

CHAPTER I

Introduction

Sensor data fusion, a branch of applied informatics, is the application and automation of well–established methodologies from nature that were present millions of years before the first computer was built. All living creatures perform an intelligent fusion of data, which is produced by their sense organs, by weighting the received information according to lessons learned and the communication with other creatures. It is sensor data fusion that enables living creatures to have situation awareness and reach their goals based on strategic mission planning [Koc14].

One challenge of sensor data and information fusion is the simultaneous tracking of targets and the derivation of related algorithms, called *multitarget tracking filters*. The aim of every multitarget tracking filter is the optimal estimation of statistical information like the spatial distribution and the number of targets present, given a set of incomplete, noisy and even false measurements. The vast growth of challenging questions arising from the applied science of sensor data fusion implied the derivation of an enormous amount of multitarget tracking filters over the past decades. The *diversity* of multitarget tracking filters available nowadays is not only capable to solve a variety of multitarget tracking scenarios, it can also be considered as a richly filled toolbox of concepts whose combination, extension and generalization can be used to solve new problems in multitarget tracking. However, *diversity* also means a burden for the tracking engineer that sees himself confronted with a new tracking scenario. *"Does a solution to my problem already exist?"*, *"Are there closely related concepts that could be used as a starting point for deriving the solution of my problem?"* and *"How do I have to modify an existing approach so that it is applicable to my problem?"* are questions any scientist will be confronted with in his/her working–life.

In mathematics various working fields with a countless number of specialization branches exist. This *diversity* of concepts and solutions implies the aim of *unification*. The *unification* in mathematics can be studied with respect to several topics, e.g. abstraction (in terms of definitions like *function*, *group*, *topology*, etc.), the com-

bination or relation of exiting theories or working areas, etc. [Ott13]. A *unification* in terms of abstraction enables the classification and comparison of existing results, which yields a deeper understanding of the unified concepts.
Analogously to mathematics the working area of sensor data fusion would benefit if a *unification* of multitarget tracking filters via an appropriate abstraction would exist. It would not only help to understand similarities and differences in existing concepts, but a classification via abstraction would also help to concentrate the daily work of tracking engineers on deriving new and customized concepts by exploiting the existing knowledge, instead of wasting resources on the complex re–identification of existing work using textual descriptions.
This work studies the *unification* of multitarget filters by applying a comprehensive and theoretically founded framework. Furthermore, the passive, non–cooperative localization and tracking of electromagnetic emitters in an urban environment using a single antenna array is solved by the application of concepts that are contained in the unifying framework.

1.1 Methodology

The theory of finite point processes is a well–known and highly developed concept from statistics and probability theory, which finds its application in various fields of financial–mathematics, physics and biology. In particular, finite point processes are perfectly suited for modeling multitarget tracking problems. In [Moy62] it is proven that probability generating functionals, which are a generalization of probability generating functions, fully and uniquely characterize a finite point process if the underlying probability distribution is symmetric. Furthermore, it is shown in [Moy62], that probability generating functionals inherently encode the full and complete statistical information about the corresponding point processes and reveal their summary statistics if functional differentiation is applied. The definition of random finite sets, a theory closely related to finite point processes, is due to Mahler [Mah03], [Mah07a], [Mah07b]. Using random finite sets and finite set statistics, he first derived a multitarget tracking filter by the application of probability generating functionals.
The first part of this thesis studies the *unification* of many well–known multitarget tracking filters, which can be modeled by finite point processes, in terms of the corresponding probability generating functionals. According to the fact that the studied filters can be formulated via finite point process theory, the *unification* is called the family of pointillist filters. This framework was first proposed by Streit in [Str14b]. In this thesis many well–known tracking filters are proven to be members of this *unification* and a classification of the filters in terms of the superposition of target states is proposed. Several similarities and differences between tracking filters are easily identified by investigating the respective probability generating functional. A

demonstration on how to use the *unification*–framework is presented to show tracking engineers the benefits of the framework in terms of the customized design of tracking filters. Additionally, the derivation of the summary statistics from the probability generating functional is discussed in a mathematically rigorous manner. Summary statistics are needed to implement and apply the filter to tracking scenarios. It is proven that the summary statistics of *all* members of the family of pointillist filters proposed in this thesis can be derived by applying an appropriate definition of the functional derivative with respect to the Dirac delta. Furthermore, it is shown that these summary statistics can be derived for *all* pointillist filters using ordinary differentiation by the extension of the theory of secular functions [Str14e] to a general class of probability generating functionals. All statements concerning the derivation of summary statistics are proven in this thesis using standard theorems from functional analysis.

An application of a subclass of multitarget tracking filters from the *unification*–framework is presented in terms of the passive and non–cooperative localization and tracking of an electromagnetic emitter in an urban environment using a single mobile antenna array. The first fundamental investigation of the problem is done in [Alg10]. Due to physical propagation effects like scattering, diffraction and reflection of the emitted signal, multiple electromagnetic waves that have travelled along different paths and are therefore referred to as *multipaths* can be received by the antenna array. Thus, a single target generates multiple measurements per sensor scan, which are, in contrast to the well–studied field of extended target tracking [KS05], spatially not related in the measurement space. Probability hypothesis density [Mah03], [Mah07a], [Mah07b], [CM12] and intensity filters [SKSC12], [Deg14] using standard and general target–oriented measurement models are applied. Due to a mismatch in the target–oriented measurement model enhanced target state extraction schemes and likelihood function definitions for single multipaths are needed when standard probability hypothesis density and intensity filters are applied. In contrast, the application of a general target–oriented measurement model within generalized probability hypothesis density and intensity filters is computationally complex. Therefore, approximation schemes for rejecting measurement partitions have to be derived for these filters. The proposed tracking filters are numerically evaluated using simulated and real world data. Furthermore, a multi–hypothesis based parameter tracking algorithm is proposed, which takes false multipaths into account that arrive before the first target–related measurement. The corresponding update equations are derived by a marginalization procedure taking into account so–called clutter hypotheses.

1.2 Structure

This thesis is divided into two parts. The first theoretical part is on the application of finite point processes in target tracking, the second part presents the application of blind mobile localization and tracking, which is solved using the methods derived in the first part.

Finite Point Processes in Target Tracking

In **Chapter 2** the fundamentals of point process theory is established and the connection to multitarget tracking filters is shown using standard literature on point process theory.

Chapter 3 presents a framework for unifying multitarget tracking filters that are modeled using finite point process theory and therefore called the family of pointillist filters. Many well–known multitarget tracking filters are presented in terms of the corresponding probability generating functional and according to their application of target superposition. The differences and similarities between pointillist filters are discussed. An example demonstrates the benefits of the proposed *unification* in a practical way. The process of characterizing multitarget tracking filters using finite point process theory presented in this chapter is called the *Discovery Step*.

The derivation of the summary statistics of pointillist filters is discussed in **Chapter 4**. Using the classic Lebesgue dominated convergence theorem the definition of the functional derivative with respect to the Dirac delta is shown to be mathematically correct for all pointillist filters formulated in Chapter 3. Additionally, an extension of the theory of secular functions [Str14e] to all proposed pointillist filters is carried out. The process of deriving the summary statistics of a pointillist filter presented in this chapter is called the *Analytical Step* of pointillist filters.

An Application to Emitter Tracking under Multipath Propagation

In **Chapter 5** the passive and non–cooperative localization and tracking of an electromagnetic emitter in an urban environment using a single mobile antenna array is introduced. Afterwards, open questions left in the related work are identified and used to formulate the contributions of Chapters 6 and 7.

Chapter 6 extends the work on bind mobile localization and tracking presented in [Alg10] to standard and generalized probability hypothesis density and intensity filters, a subclass of pointillist filters that superpose targets in a single state space. Enhanced target state extraction schemes and likelihood functions that are defined on single multipaths are derived for the probability hypothesis density and intensity filter, which use the standard target–oriented measurement model. Furthermore, approximation criteria for generalized probability hypothesis and intensity filters are derived,

which do not apply any information about the spatial distribution of measurements in the parameter space. Finally, the proposed filters are numerically compared using simulated and real world data.

A novel filter for tracking parameters of blind mobile localization and tracking scenarios in relative time of arrival, based on a track–oriented multi–hypothesis approach, is presented in **Chapter 7**, which is capable of handling falsely detected multipaths that arrive before the first target–related measurement.

The conclusions are drawn and future work is presented in **Chapter 8**.

Part I

Finite Point Processes in Target Tracking

CHAPTER II

Finite Point Processes and Probability Generating Functionals

The multitarget tracking problems studied in this thesis employ finite point process models of multiple target states and measurement sets. Probability generating functionals (PGFLs) are a generalization of probability generating functions (PGFs) for multivariate discrete random variables (RVs). Moyal [Moy62] used PGFLs to characterize and study finite point processes. Of particular relevance here, Moyal showed that PGFLs characterize point processes in terms of their functional derivatives. The statistics which are encoded in the PGFL can be decoded via functional differentiation and yield the factorial moments of a PGFL. The factorial moments of a finite point process are an analogous concept to the moments of an RV. The difference is that factorial moments of a point process are given by the functional derivatives of a PGFL, while moments of an RV are represented by ordinary derivatives of a PGF. Thus, factorial moments of a point process are given by functions, while the moments of a RV are given by scalars. Another interpretation of PGFs is presented in [Wil94], where a generating function is described as *"... a clothesline on which we hang up a sequence of numbers for display"*. Following this description, the factorial moments of a finite point process can be interpreted as an infinite dimensional vector of scalars. This goes along with the fact that PGFLs can be derived as the *small cell* limit of PGFs of histogram counts and, vice versa, histogram PGFs can be derived out of a PGFL by substituting a weighted sum of Dirac deltas [SSCB14, Section 5.4.4].
Finite point processes are well–suited to model multitarget tracking applications. The PGFL of the joint target–measurement point process can be derived for many standard tracking problems directly from the assumptions of the tracking problem, making PGFLs an appropriate tool to design tracking filters based on point process models [SDK15]. Furthermore, PGFLs can be used to find similarities and differences in multitarget tracking filters, which use point process models for the target state and

the measurement set (for details see Chapter 3 or [SDK15]). The fundamentals of point processes, PGFLs and their connection to target tracking are presented in this chapter.
This chapter is structured as follows. In Section 2.1 the fundamentals of point processes are presented and the PGFL of a point process is defined. The Gâteaux derivative of a PGFL is defined in Section 2.3 before the event likelihood function is determined by functional differentiation of the PGFL in Section 2.4. In Section 2.5 factorial moments are defined and interpreted. Poisson point processes (PPPs), which represent an important class of point processes for applications, are given in Section 2.6. The PGF of the canonical, or cardinal, number of points is presented in Section 2.7. Multivariate PGFLs are defined in Section 2.8, before the connection of point processes to target tracking is shown via the PGFL of the Bayes posterior process in Section 2.9. Summary statistics of the Bayes posterior process PGFL are defined Section 2.10 and the branching form of the PGFL is derived in Section 2.11. Finally, point processes are compared to RFS in Section 2.13 .
The content of this chapter is a review of well–known results and can be found (in different notations) in [Str13a], [HDC13], [Mah07b], [SKM95], [Moy62], [DVJ03], [Kar91] and the references cited therein.

2.1 Point Process Fundamentals

Let $\mathcal{X}$ be a complete separable metric space. A typical choice for $\mathcal{X}$, sufficient for most applications appearing in multitarget tracking problems and used throughout this thesis, is $\mathbb{R}^d$, $d > 0$. Then, the space of sets of points or event space in $\mathcal{X}$ is defined by the disjoint union

$$E_{\mathcal{X}} \equiv \emptyset \cup \bigcup_{n \geq 1} \mathcal{X}^{(n)}, \tag{2.1}$$

where $\emptyset$ denotes the empty–set and $\mathcal{X}^{(n)}$ is the space of sets of size $n \in \mathbb{N}$, that is

$$\mathcal{X}^{(n)} \equiv \Big\{ \{x_1, ..., x_n\} | x_i \in \mathcal{X}, i = 1, ..., n \Big\}. \tag{2.2}$$

In general it is not assumed that the elements of $\mathcal{X}^{(n)}$ need to be distinct, that is repetitions of elements are allowed. Thus, the elements of $\mathcal{X}^{(n)}$ are called multisets (a multiset is a set, for which the repetition of elements is allowed [Knu98]). Furthermore, the order of the corresponding indices is not unique. In physics (thermodynamics), $E_{\mathcal{X}}$ is called the grand canonical ensemble, $\mathcal{X}^{(n)}$ is called the nth canonical ensemble and n is called the canonical number.
A stochastic point process in the sense of [SKM95, p. 109], [HDC13] is defined to be a measurable mapping

$$\Phi : (\Omega, \mathcal{F}, \mathbb{P}) \to \left(E_{\mathcal{X}}, \mathcal{B}\left(E_{\mathcal{X}}\right)\right), \tag{2.3}$$

where $(\Omega, \mathcal{F}, \mathbb{P})$ is an arbitrary probability space and $\mathcal{B}(E_{\mathcal{X}})$ denotes the Borel σ–algebra of $E_{\mathcal{X}}$. Since the point process Φ is a measurable mapping, the stochastic model of the point process is defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The associated counting function for an arbitrary $B \in \mathcal{B}(\mathcal{X})$ is defined by

$$\begin{aligned} N_{\cdot}(B) : \big(E_{\mathcal{X}}, \mathcal{B}(E_{\mathcal{X}})\big) &\to \big(\mathbb{N}, \mathcal{B}(\mathbb{N})\big) \\ \varphi &\mapsto N_{\varphi}(B) \equiv |\varphi \cap B|, \end{aligned} \tag{2.4}$$

which counts the number of elements of φ in B and is measurable. Then, the composition

$$N_{\Phi(\cdot)}(B) \equiv N_{\cdot}(B) \circ \Phi : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{N}, \mathcal{N}) \tag{2.5}$$

is measurable, since the composition of measurable functions is measurable again. Here, $\mathcal{N}$ denotes the smallest σ-algebra, such that $N_{\Phi(\cdot)}(B)$ is measurable. Thus, $N_{\Phi(\cdot)}(B)$ is an integer–valued RV for all $B \in \mathcal{B}(\mathcal{X})$.
A class of measures on locally compact Haussdorf spaces, which is of particular interest for the definition of the point process are Radon measures. This family comprises among others the Dirac measure, the counting measure and the Lebesgue measure. Readers, who are not familiar with measure theory might skip the following paragraph and proceed with the definition of point processes studied in this thesis.

Excursion to Radon Measures Let M denote the set of Radon measures [Els09]. The corresponding $\sigma-$algebra of M is generated by the coordinate mapping $\mu \mapsto \mu(f) = \int f d\mu$, where f ranges over the set of continuous functions on $\mathcal{X}$ with compact support (for details see [Kar91]). A Radon measure μ on $\mathcal{B}(\mathcal{X})$ satisfies the following two conditions.

1. For all $x \in \mathcal{X}$ there exists $U \subset \mathcal{X}$ open such that $\mu(U) < \infty$. (locally finite)
2. $\mu(A) = \sup\{\mu(K) | K \subset A,\ K \text{ compact}\}$ for all $A \in \mathcal{B}(\mathcal{X})$. (inner regular)

Then, according to [HDC13, Chapter 4] $N_{\Phi(\cdot)}(\cdot)$ can be interpreted in the following way.

1. $N_{\Phi(\omega)}(B) \in \mathbb{N}$ denotes the number of points of the realization of the point process $\Phi(\omega)$ in $B \in \mathcal{B}(\mathcal{X})$.
2. $N_{\Phi(\cdot)}(B) : (\Omega, \mathcal{F}, \mathbb{P}) \to \big(\mathbb{N}, \mathcal{N}\big)$ is a RV that maps an element of the underlying probability space to the number of points from the realization $\Phi(\omega)$ in $B \in \mathcal{B}(\mathcal{X})$.
3. $N_{\Phi(\omega)}(\cdot) : \mathcal{B}(\mathcal{X}) \to \mathbb{N}$ is a counting measure.
4. $N_{\Phi(\cdot)}(\cdot) : (\Omega, \mathcal{F}, \mathbb{P}) \to M_p \equiv \{\mu \in M : \mu(A) \in \mathbb{N} \text{ for all } A \in B\}$ maps an element of the probability space to the point measure $N_{\Phi(\omega)}(\cdot)$.

In [Kar91] a point process is defined to be a measurable mapping from $(\Omega, \mathcal{F}, \mathbb{P})$ to M_p, that is $N_{\Phi(\cdot)}(\cdot)$ in our notation. Hence, point processes in the sense of [Kar91] are a special kind of random measures. Different classes of point processes are presented in [Kar91], defined in terms of their codomain, that is

1. $M_s \equiv \{\mu \in M_p : \mu(\{x\}) \leq 1 \text{ for all } x \in \mathcal{X}\}$
2. $M_a \equiv \{\mu \in M : \mu \text{ is purely atomic}\}$
3. $M_d \equiv \{\mu \in M : \mu \text{ is diffuse}\}$.

The $\sigma-$algebras of M_p, M_s, M_a and M_d in M are given by the corresponding trace $\sigma-$ algebras. In [Kar91] a point process Φ is called simple if $\mathbb{P}(N_{\Phi(\cdot)}(\cdot) \in M_s) = 1$. Then, the points in $\mathcal{X}^{(n)}$ are $\mathbb{P}$–almost surely (a.s.) distinct. The random measure $N_{\Phi(\cdot)}(\cdot)$ is called atomic if $\mathbb{P}(N_{\Phi(\cdot)}(\cdot) \in M_a) = 1$ and diffuse if $\mathbb{P}(N_{\Phi(\cdot)}(\cdot) \in M_d) = 1$. A measure $\mu \in M$ is diffuse if $\mu(\{x\}) = 0$ for all $x \in \mathcal{X}$. Every Radon measure μ can be decomposed into

$$\mu = \mu_d + \sum_{i=1}^{K} a_i \delta_{x_i}, \tag{2.6}$$

where δ_x denotes the Dirac measure at $x \in \mathcal{X}$ [Kar91, Theorem A.4], $a_i > 0$, $x_i \in \mathcal{X}$ distinct, for all $i = 1, ..., K$, $K \in \mathbb{N}^+$ and μ_d is a diffuse measure. A Radon measure μ is then called purely atomic if its diffuse component μ_d is zero, that is it can be represented by a sum of Dirac measures. An atomic measure is a point measure if and only if for all $i = 1, ..., K$, $a_i \in \mathbb{N}$. If this condition is satisfied, we speak about an atomic point processes instead of an atomic random measure. Note that for a simple point process $a_i = 1$, for all $i = 1, ..., K$. In the following, when speaking about an atomic point process we implicitly assume $a_i \in \mathbb{N}$, for all $i = 1, ..., K$.
Combining the definitions of [SKM95, Chapter 4] and [Kar91, Chaper 1] a point process Φ is called in this thesis

1. *locally finite*, if each bounded subset of $\mathcal{X}$ must only contain a finite number of points of φ, for all $\varphi \in E_{\mathcal{X}} \setminus \emptyset$ $\mathbb{P}$–a.s
2. *simple*, if for all $x_i, x_j \in \varphi$, $\quad x_i = x_j \Rightarrow i = j$ holds $\mathbb{P}$–a.s., for all $\varphi \in E_{\mathcal{X}} \setminus \emptyset$.

In this thesis a simple finite point process is meant, unless otherwise stated, if we speak about a point process. In terms of multitarget tracking the two conditions on the elements of the event space $E_{\mathcal{X}}$ represent the assumptions on the multitarget tracking problem that only finitely many targets can be present in a scenario, where the field of view is bounded (*locally finite*) and that no two targets share the same target state (*simple*).
The intensity measure (first moment measure, mean) of the point process Φ is defined

for an arbitrary $B \in \mathcal{B}(\mathcal{X})$ by the expectation value

$$\mu_\Phi(B) \equiv \mathbb{E}[N_{\Phi(\cdot)}(B)] = \int_\Omega N_{\Phi(\omega)}(B)\mathbb{P}(d\omega) = \int_{E_\mathcal{X}} N_\varphi(B) P_\Phi(d\varphi), \tag{2.7}$$

where P_Φ denotes the push–forward (image) measure of $\mathbb{P}$, which uses the point process Φ as a measurable mapping. Therefore, the moment measure $\mu_\Phi(B)$ yields the expected number of points in B. The expectation (2.7) can be extended further by the nth moment measure

$$\mu_\Phi^{(n)}(B_1, ..., B_n) \equiv \mathbb{E}\left[N_{\Phi(\cdot)}(B_1) \cdots N_{\Phi(\cdot)}(B_n)\right], \tag{2.8}$$

where $B_1, ..., B_n \in \mathcal{B}(\mathcal{X})$. Note that $\mu_\Phi^{(n)}(B^n)$ denotes the nth moment of the RV $N_{\Phi(\cdot)}(B^n)$.
For practical applications like multitarget tracking problems it is of particular interest (e.g. for implementation issues), whether the first moment μ_Φ (2.7) has a density. Not every point process has a moment measure density. For example, if points are arranged on a lattice (see for example [Str14c]) no density corresponding to the first moment exists. However, if the nth moment measure is absolutely continuous with respect to the Lebesgue measure and corresponds to a locally finite, simple point process, the corresponding density exists and is called the nth moment density. For simple, locally finite point processes the first moment density in case of existence is also often referred to as intensity function or (in the tracking community) as probabilistic hypothesis density (PHD) [Mah07b], [Mah03], [Mah07a].

2.2 Probability Generating Functionals

This section follows the considerations of [HDC13], [Str13a] and [Moy62]. For any complex-valued Lebesgue–integrable function

$$h : (\mathcal{X}, \mathcal{B}(\mathcal{X})) \to (\mathbb{R}, \mathcal{B}(\mathbb{R})) \tag{2.9}$$

satisfying $|h(x)| \leq 1$ for all $x \in \mathcal{X}$ and

$$x \in \Phi \equiv \left\{x \in \mathcal{X} : \exists\ \omega \in \Omega \text{ such that } x \in \Phi(w)\right\} \tag{2.10}$$

the PGFL of the locally finite point process Φ is defined by the expectation of $\prod_{x\in\Phi} h(x)$ with respect to the image measure P_Φ of the point process Φ by

$$\Psi_\Phi(h) \equiv \mathbb{E}\left(\prod_{x\in\Phi} h(x)\right) = \sum_{n\geq 0} \int_{\mathcal{X}^{(n)}} \prod_{i=1}^{n} h(x_i) P_\Phi(d\{x_1, ..., x_n\}) \tag{2.11}$$

$$= \sum_{n\geq 0} p_N^\Phi(n) \int_{\mathcal{X}^n} \prod_{i=1}^{n} h(x_i) p^\Phi_{\mathcal{X}|N}(x_1, ..., x_n|n) dx_1 ... dx_n \tag{2.12}$$

$$= \sum_{n\geq 0} \frac{1}{n!} \int_{\mathcal{X}^n} \prod_{i=1}^{n} h(x_i) p_n^\Phi(x_1, ..., x_n) dx_1 ... dx_n, \tag{2.13}$$

where $p_n^\Phi : \mathcal{X}^n \to \mathbb{R}$ is the multi–object density of the corresponding Janossy measure and defined such that

$$\int_B n! P_\Phi(d\{x_1, ..., x_n\}) = \int_{\tilde{B}} p_n^\Phi(x_1, ..., x_n) dx_1 .... dx_n \tag{2.14}$$

holds for all $B \in \mathcal{B}(E_\mathcal{X})$ and $\tilde{B} \in \mathcal{B}(\mathcal{X}^n)$, where the tuples of points contained in B and $\tilde{B}$ are identical. For $n = 0$ in (2.11)–(2.13) it holds that $\prod_{i=1}^{n} h(x_i) \equiv 1$, $p_n^\Phi(\cdot) \equiv 1$, $p^\Phi_{\mathcal{X}|N}(\cdot|n) \equiv 1$ and $P_\Phi(\cdot) \equiv 1$. Equations (2.12) and (2.13) hold if absolute continuity of P_Φ with respect to the Lebesgue measure is assumed due to the application of the Radon–Nikodym theorem.
Note that $P_\Phi(d\{x_1, ..., x_n\})$ considers the ordered event and $p_n^\Phi(x_1, ..., x_n)$ defines the probability density function (PDF) evaluated at the unordered event. The factorization

$$\int_{\mathcal{X}^{(n)}} P_\Phi(d\{x_1, ..., x_n\}) = p_N^\Phi(n) \int_{\mathcal{X}^n} p^\Phi_{\mathcal{X}|N}(x_1, ..., x_n|n) dx_1 ... dx_n \tag{2.15}$$

is due to the definition of the conditional probability and the assumption that P_Φ is absolutely continuous. The PGFL is well–defined for any measurable function h, with $|h(x)| \leq 1$ for all $x \in \mathcal{X}$ due to the assumption that the point process Φ is locally finite [DVJ08, p. 59].
In [Moy62] it is shown that finite point processes are characterized by their PGFL via functional derivatives, which are studied in the following section.

On the absolute continuity of the image measure of a point process We assume from now on that first order moment measure defined in 2.7 of the point process is

absolutely continuous with respect to the Lebesgue measure. Therefore, the corresponding density, that is the Radon–Nikodym derivative exists due to the Radon–Nikodym theorem. However, point processes exists, which do not have an absolute continuous (with respect to the Lebesgue measure) first moment measure and thus no first (and higher) order factorial moment densities exist. In Section 2.12 we study two families of point processes which do not possess an absolute continuous (with respect to the Lebesgue measure) image measure, that are Palm processes and point processes, where the state space has at least one discrete element. For such point processes the results from Section 2.3 to Section 2.11 need to be reformulated using the Radon measure P_Φ instead of the first moment density p_n^Φ.

2.3 The Functional Derivative of a PGFL

This section follows from [Str13a, Section 3.2]. Let Ψ_Φ be the PGFL of a locally finite point process Φ defined as in (2.13). Then, the Gâteaux derivative of Ψ_Φ with respect to the variation

$$\omega : (\mathcal{X}, \mathcal{B}(\mathcal{X})) \to (\mathbb{R}, \mathcal{B}(\mathbb{R})) \tag{2.16}$$

is defined by

$$\frac{\partial}{\partial \omega}\Psi_\Phi(h) \equiv \lim_{\epsilon \searrow 0} \frac{d}{d\epsilon}\Psi_\Phi(h + \epsilon\omega) = \lim_{\epsilon \searrow 0} \frac{\Psi_\Phi(h + \epsilon\omega) - \Psi_\Phi(h)}{\epsilon}, \tag{2.17}$$

where ω is a complex–valued, bounded and Lebesgue–integrable function on $\mathcal{X}$. Note that we are considering only the limit from above, that is, $\lim_{\epsilon \searrow 0}$, since then the PGFLs in (2.17) are well–defined due to [DVJ08]. In [Moy62, Section 4] it is shown that

$$\Psi_\Phi(h + \epsilon\omega) = \sum_{n \geq 0} \frac{1}{n!} \int_{\mathcal{X}^n} \prod_{i=1}^{n} \big(h(x_i) + \epsilon\omega(x_i)\big)\, p_n^\Phi(x_1, ..., x_n) dx_1 ... dx_n \tag{2.18}$$

is an analytic function in the variable $\epsilon \in \mathbb{C}$ in some open region of the complex plane containing $0 \in \mathbb{C}$. Thus, the analyticity of Ψ_Φ justifies an interchange of limit and the (infinite) sum when taking the Gâteaux derivative of (2.18). Therefore, the Gâteaux derivative of (2.18) can be determined and is given by

$$\frac{\partial \Psi_\Phi}{\partial \omega}(h) = \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{k=1}^{n} \int_{\mathcal{X}^n} \omega(x_k) \prod_{i=1, i \neq k}^{n} h(x_i) p_n^\Phi(x_1, ..., x_n) dx_1 ... dx_n. \tag{2.19}$$

Since the summand for $n = 0$ is zero, the outermost sum starts at $n = 1$. The product rule for ordinary differentiation and the assumption that p_n is symmetric yields the inner sum.

The differentiation with respect to multiple real–valued, bounded and integrable variations $\omega_1, ..., \omega_n : (\mathcal{X}, \mathcal{B}(\mathcal{X})) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is defined iteratively, that is, by

$$\frac{\partial^n \Psi_\Phi}{\partial \omega_1 \cdots \partial \omega_n}(h) = \frac{\partial^n \Psi_\Phi}{\partial \epsilon_1 \cdots \partial \epsilon_n} \left(h + \sum_{j=1}^{n} \epsilon_j \omega_j \right) \Bigg|_{\epsilon_1 = \cdots = \epsilon_n = 0}, \tag{2.20}$$

which is called simultaneous perturbation and which is due to [Moy62].
In [Dir27] Dirac defined a function, which satisfies

$$\delta(x) = 0, \text{ when } x \neq 0 \tag{2.21}$$

and

$$\int \delta(x) dx = 1. \tag{2.22}$$

This function δ later became referred to as Dirac delta.

$$\delta^c(x) \equiv \delta(x - c) \tag{2.23}$$

for $c \in \mathcal{X}$. The Gâteaux derivative with respect to the Dirac delta at the point c (or as in [Str13a]: functional derivative with respect to an impulse at c) is informally defined by inserting $\omega(\cdot) = \delta^c(\cdot)$

$$\frac{\partial \Psi_\Phi}{\partial c}(h) = \frac{\partial \Psi_\Phi}{\partial \delta^c}(h) = \frac{\partial \Psi_\Phi}{\partial \omega}(h) \Bigg|_{\omega(\cdot) = \delta^c(\cdot)} \tag{2.24}$$

$$= \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{k=1}^{n} \int_{\mathcal{X}^n} \delta^c(x_k) \prod_{i=1, i \neq k}^{n} h(x_i) p_n^\Phi(x_1, ..., x_n) dx_1 ... dx_n \tag{2.25}$$

$$= \sum_{n=1}^{\infty} \frac{1}{(n-1)!} \int_{\mathcal{X}^{n-1}} \prod_{i=2}^{n} h(x_i) p_n^\Phi(c, ..., x_n) dx_1 ... dx_n. \tag{2.26}$$

Here, (2.26) is due to the application of the sampling property of Dirac delta, the symmetry of p_n^Φ and a relabeling of the arguments. Note that $\frac{\partial \Psi_\Phi}{\partial c}(h)$ is again a functional that explicitly depends on the point $c \in \mathcal{X}$. Analogously, The Gâteaux derivative with respect to several Dirac deltas at $c_1, ..., c_n \in \mathcal{X}$ is given by

$$\frac{\partial^n \Psi_\Phi}{\partial x_1 \cdots \partial x_n}(h) \equiv \frac{\partial^n \Psi_\Phi}{\partial \omega_1 \cdots \partial \omega_n}(h) \Bigg|_{\omega_1 = \delta^{x_1}, \cdots, \omega_n = \delta^{x_n}} \tag{2.27}$$

$$= \sum_{k=n}^{\infty} \frac{1}{(k-n)!} \int_{\mathcal{X}^{k-}} \prod_{i=n+1}^{k} h(x_i) p_k^\Phi(x_1, ..., x_n, x_{n+1}, ..., x_k) dx_{n+1} \cdots dx_k, \tag{2.28}$$

where the product is equal to one for $k = n$. For all non-negative, bounded by one and Lebesgue–integrable test–functions h, the functional derivative of $\Psi_\Phi(h)$ with respect to Dirac delta for $n = 0$ is defined to be $\Psi_\Phi(h)$.
Obviously, Dirac delta cannot be a proper function (and therefore no valid variation for which the Gâteaux derivative in its original form is well–defined) due to the fact, that the value of the Lebesgue integral does not change if the integrand is changed on sets of measure zero. For this chapter we omit the details on the mathematically correct definition of Dirac delta via the limit (in a distributional sense) of an approximate identity, that is, a sequence of test–functions. Furthermore, we omit the study for which PGFLs (2.24) is well–defined. Instead we use (2.24) and (2.25) for the rest of this chapter, keeping in mind, that a mathematically careful treatment is readily provided by using a sequence of test–functions. In Section 4.1 the functional differentiation with respect to the Dirac delta is derived in a mathematically correct way for a large class of PGFLs that is sufficient for almost all multitarget tracking problems.

2.4 Event Likelihood

As mentioned in the introduction of this chapter functional differentiation encrypts the statistical information of the point process, which is encoded in the PGFL analogously to the moments of a RV. In particular, the probability density function (PDF) is contained as a statistical information in the PGFL [Str13a]. A locally finite point process is completely characterized by its PGFL [Moy62], which means that finite point processes can be derived via their PGFLs. It holds, if absolute continuity of P_Φ with respect to the Lebesgue measure is assumed, that

$$\frac{\partial \Psi_\Phi}{\partial c_1}(0) = p_1^\Phi(c_1) = 1! p_N^\Phi(1) p_{\mathcal{X}|N}^\Phi(c_1) \tag{2.29}$$

and

$$\frac{\partial^2 \Psi_\Phi}{\partial c_1 \partial c_2}(0) = \frac{\partial^2 \Psi_\Phi}{\partial c_2 \partial c_1}(0) = p_2^\Phi(c_1, c_2) = 2! p_N^\Phi(2) p_{\mathcal{X}|N}^\Phi(c_1, c_2), \tag{2.30}$$

$c_1,\ c_2 \in \mathcal{X}$. Hence, the Gâteaux derivative with respect to Dirac delta at c evaluated at $h = 0$ is the PDF of the event $\{c\}$ and the Gâteaux derivative with respect to Dirac delta at c_1 and c_2 yields the PDF of the unordered event $\{c_1, c_2\}$ or equivalently 2! times the PDF of the ordered event [Str13a, Section 3.3]. Analogously one derives for $n \geq 1$

$$\frac{\partial^n \Psi_\Phi}{\partial c_1 \cdots \partial c_n}(0) = p_n^\Phi(c_1, ..., c_n) = n! p_N^\Phi(n) p_{\mathcal{X}|N}^\Phi(c_1, ..., c_n). \tag{2.31}$$

2.5 Factorial Moments

The nth–order factorial moment measure $\alpha^{(n)}$ of the point process Φ is defined for a non–negative measurable function $h : (\mathcal{X}, \mathcal{B}(\mathcal{X})) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ in [SKM95] by

$$\int_{\mathcal{X}^n} h(x_1, ..., x_n)\alpha^{(n)}(d(x_1, ..., x_n)) \equiv \int_{\mathcal{X}^{(n)}} \sum_{x_1,...,x_n \in \varphi}^{\neq} h(x_1, ..., x_n)P(d\varphi), \tag{2.32}$$

where the notation $\sum^{\neq}_{x_1,...,x_n \in \varphi}$ means that the summation is over all n–tuples of distinct points in φ including all permutations of given points. Therefore, if $B_1, ..., B_n \in \mathcal{B}(\mathcal{X})$ are pairwise disjoint sets, the nth order factorial moment measure is equal to the nth moment measure, that is

$$\mu^{(n)}(B_1 \times \cdots \times B_n) = \alpha^{(n)}(B_1 \times \cdots \times B_n). \tag{2.33}$$

Note that the nth–order factorial moment measure $\alpha^{(n)}(B^n)$ is the nth–order factorial moment of the RV $N_{\Phi(\cdot)}(B^n)$. If the nth–order factorial moment measure is locally finite and absolute continuous with respect to the Lebesgue measure, due to the Radon–Nikodym theorem, the corresponding density, called nth–order factorial moment density (also called nth–order product density) exists [SKM95].
The first–order moment density, or intensity, of the point process Φ can be derived by the Gâteaux derivative of Ψ_Φ with respect to the Dirac delta at $c \in \mathcal{X}$ by

$$m^{\Phi}_{[1]}(c) = \frac{\partial \Psi_\Phi}{\partial c}(1) = \sum_{n=1}^{\infty} \frac{1}{(n-1)!} \int_{\mathcal{X}^{n-1}} p^{\Phi}_n(c, x_2, ..., x_n)dx_2...dx_n. \tag{2.34}$$

It is equal to the intensity function [SKM95]. The nth–order factorial moment density can be derived by the Gâteaux derivative of Ψ_Φ with respect to the Dirac deltas at $c_1, ..., c_n \in \mathcal{X}$ by

$$\begin{aligned} m^{\Phi}_{[n]}(c_1, ..., c_n) =& \frac{\partial^n \Psi_\Phi}{\partial c_1 \cdots \partial c_n}(1) \\ =& \sum_{k=n}^{\infty} \frac{1}{(k-n)!} \int_{\mathcal{X}^{k-n}} p^{\Phi}_k(c_1, ..., c_n, x_{n+1}, ..., x_k)dx_{n+1} \cdots dx_k. \end{aligned} \tag{2.35}$$

The first moment measure (2.7) can be obtained via functional differentiation of the corresponding PGFL Ψ_Φ, that is for all $B \in \mathcal{B}(\mathcal{X})$

$$\mu^{(1)}_\Phi(B) = \frac{\partial \Psi_\Phi}{\partial 1_B}(h)\bigg|_{h=1}, \quad 1_B(x) = \begin{cases} 0, & \text{if } x \notin B \\ 1, & \text{if } x \in B \end{cases} \tag{2.36}$$

For details see [SKM95] and [HDC13].
Assume that the nth–order factorial moment measure $\alpha^{(n)}$ is absolutely continuous

with respect to the Lebesgue measure and corresponds to a locally finite and simple point process. Then, due to [DVJ03, Lemma 5.4.III] the nth–order factorial moment density is the Radon–Nikodym derivative of the nth–order factorial moment measure defined in (2.7) with respect to the Lebesgue measure and according to [DVJ03, Equation (5.4.12)] the nth–order factorial moment density can be intuitively written as

$$m^{\Phi}_{[n]}(x_1, ..., x_n)dx_1 \cdots dx_n = \Pr\left(\begin{array}{c}\text{one point of the process is}\\ \text{located in each infitesimal subset}\\ [x_i, x_i + dx_i), i = 1, ..., n\end{array}\right), \tag{2.37}$$

which shows that the factorial moment densities can be interpreted as multi–point intensity functions if the points are distinct with probability one. In particular, the intensity function can be intuitively interpreted as

$$m^{\Phi}_{[1]}(x)dx = \Pr\left(\begin{array}{c}\text{one point of the process}\\ \text{is located in } [x, x + dx)\end{array}\right). \tag{2.38}$$

The pair–correlation function is given by $m^{\Phi}_{[2]}(x_1, x_2)$. In tracking applications it characterizes the *spooky action* [FSU09] between two targets in the Bayes posterior process due to assignment interference. Multipoint–correlation functions can also be computed. An application of factorial moment densities to multitarget tracking applications is given in [BES13].

The first moment, or intensity, for multivariate PGFLs is an extension of the definition (2.34). Let the test functions $h_j : (\mathcal{X}_j, \mathcal{B}(\mathcal{X}_j)) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, $j = 1, \ldots, n$ be non–negative, bounded by one and Lebesgue–integrable, and let $x = (x_1, \ldots, x_n)^T \in \mathcal{X}_1 \times \cdots \times \mathcal{X}_n$. The first moment of a multivariate PGFL $\Psi(h_1, \ldots, h_n)$ is defined as the mixed first–order partial derivative

$$m_{[1,\ldots,1]}(x) \equiv \left.\frac{\partial \Psi(h_1, \ldots, h_n)}{\partial x_1 \cdots \partial x_n}\right|_{h_1 = \cdots = h_n = 1}. \tag{2.39}$$

The intensity is seen to be a multivariate function defined on $\mathcal{X}_1 \times \cdots \times \mathcal{X}_n$. The higher order (mixed) factorial moments $m_{[k_1,\ldots,k_n]}(x)$ can be defined analogously, as the mixed partial derivative of order $k_1, \ldots, k_n$ with respect to the test functions $h_1, \ldots, h_n$, respectively.

The functional $\Psi(\cdot)$ is linear if, for all test functions h, g and constants $a, b \in \mathbb{C}$,

$$\Psi(ah + bg) = a\Psi(h) + b\Psi(g). \tag{2.40}$$

It is straightforward to see that for linear functionals, the only events with a *nonzero* probability are realizations of the point process that have exactly one point, and the PDF of this point (in continuous spaces $\mathcal{X}$) is identical to the intensity defined in

(2.34). Therefore, the intensity integrates to one if the functional is linear. Multivariate functionals are multilinear if they are linear in each test function separately. As in the univariate case, it is easily verified that, with probability one, realizations are singleton points (in a Cartesian product space), and the multivariate PDF is identical to the intensity function.

2.6 Poisson Point Processes

According to [Str13a] a locally finite point process Φ on $\mathcal{X}$ is a PPP, if for all $B \in \mathcal{B}(\mathcal{X})$ the RV $N_{\Phi(\cdot)}(B)$ is Poisson distributed with mean $\mu = \int_B f^\Phi(x)dx$, where $f^\Phi(x) \geq 0$ for all $x \in \mathcal{X}$ is the intensity function and the points of Φ are identically and independently distributed with PDF f^Φ/μ in $\mathcal{X}$. The PGFL of a PPP is given by

$$\Psi_\Phi(h) = \exp\left(-\int_{\mathcal{X}} f^\Phi(x)dx + \int_{\mathcal{X}} h(x)f^\Phi(x)dx\right). \tag{2.41}$$

It can be shown that the intensity is the Radon–Nikodym derivative of the first moment measure μ_Φ defined in (2.7), that is, for all $B \in \mathcal{B}(\mathcal{X})$ it holds that

$$\mu_\Phi(B) = \int_B f^\Phi(x)dx = \int_B m^\Phi_{[1]}(x)dx. \tag{2.42}$$

Furthermore, it holds that

$$m^\Phi_{[n]}(x_1, ..., x_n) = f^\Phi(x_1) \cdot ... \cdot f^\Phi(x_n), \tag{2.43}$$

where $x_1, ..., x_n \in \mathcal{X}$ [SKM95].
In [Str10] and [SKM95] many details and references on PPPs can be found. PPPs play an important role in applications (just like the Poisson distribution). Some of them are presented in [Str10].

2.7 Probability Generating Function of the Canonical Number

This section follows [Str13a, Section 3.5]. The PGF of the RV describing the number of points $N_{\Phi(\cdot)}(\mathcal{X})$ of the locally finite point process Φ in $\mathcal{X}$ is denoted by F^Φ. It can be obtained by evaluating the PGFL Ψ_Φ for the constant function $h(x) = x \in \mathbb{R}$, for all $x \in \mathcal{X}$ that is

$$F^\Phi(x) \equiv \Psi_\Phi(h)\Big|_{h=x} = \Psi_\Phi(x) = \sum_{n=0}^{\infty} p^\Phi_N(n)x^n. \tag{2.44}$$

The PGF $F^\Phi(z^{-1})$ is also known as the z–transform of the sequence of probabilities $\{p^\Phi_N(n)\}_{n\in\mathbb{N}}$ in the signal processing literature. The discrete probability distribution

of the number of points p_N^{Φ}, which is defined in (2.12), can be obtained via ordinary differentiation of the PGF of the number of points, that is for all $n \in \mathbb{N}$

$$p_N^{\Phi}(n) = \frac{1}{n!} \frac{d^n F^{\Phi}}{dx^n}(0). \tag{2.45}$$

The expected number of points in one realization of the locally finite point process Φ is given by the first ordinary derivative of F^{Φ} evaluated at one, that is

$$\frac{dF^{\Phi}}{dx}(1) = \sum_{n=0}^{\infty} p_N^{\Phi}(n) x^{n-1} \Bigg|_{x=1} = \mathbb{E}[N_{\Phi(\cdot)}(\mathcal{X})]. \tag{2.46}$$

2.8 Multivariate Probability Generating Functionals

In the previous sections univariate PGFLs are studied. A univariate PGFL corresponds to a single locally finite point process. For multitarget tracking applications a minimal extension has to be made. There, at least two point processes are needed to model the multitarget tracking problem: (at least) one point process for the target and (at least) one point process for the measurement process, which leads to multivariate PGFLs.

A multivariate PGFL is defined on multiple function spaces and corresponds to multiple locally finite point processes. For the definition of the multivariate PGFL let $\mathcal{X}_1, ..., \mathcal{X}_r$ be the spaces of the finite point processes $\Phi_1, ..., \Phi_r$. For almost all practical target tracking applications $\mathcal{X}_i \equiv \mathbb{R}^{d_i}$, $d_i > 0$, $i = 1, ..., r$ is sufficient. Note that the dimensions d_i can mutually differ. Let

$$E_{\mathcal{X}_i} \equiv \emptyset \cup \bigcup_{n \geq 1} \mathcal{X}_i^{(n)} \tag{2.47}$$

denote the event space in $\mathcal{X}_i$, $i = 1, ..., r$ analogously to (2.1). The event space $E_{\mathcal{X}_i}$ models all possible realizations of the locally finite point process Φ_i on $\mathcal{X}_i$. Let $h_1, ..., h_r$ be bounded and Lebesgue–integrable functions with

$$h_i : (\mathcal{X}_i, \mathcal{B}(\mathcal{X}_i)) \to (\mathbb{R}, \mathcal{B}(\mathbb{R})) \tag{2.48}$$

and $1 \geq h_i(x_i) \geq 0$ for all $x_i \in \mathcal{X}_i$ and $i = 1, ..., r$. Consider the joint point process $(\Phi_1, ..., \Phi_r)$ with events in the Cartesian product space $E_{\mathcal{X}_1} \times \cdots \times E_{\mathcal{X}_r}$. Then, the multivariate PGFL of the joint process $(\Phi_1, ..., \Phi_r)$ is defined by the expectation of the product

$$\prod_{x_1 \in \Phi_1} h_1(x_1) \cdot ... \cdot \prod_{x_r \in \Phi_r} h_r(x_r), \tag{2.49}$$

that is

$$\begin{aligned}
&\Psi_{\Phi_1\cdots\Phi_r}(h_1,...,h_r) \\
&\equiv \mathbb{E}\left[\prod_{x_1\in\Phi_1} h_1(x_1)\cdot...\cdot\prod_{x_r\in\Phi_r} h_r(x_r)\right] \qquad (2.50)\\
&= \sum_{n_1=0}^{\infty}\cdots\sum_{n_r=0}^{\infty}\int_{\mathcal{X}_1^{n_1}}\cdots\int_{\mathcal{X}_r^{n_r}}\prod_{i_1=1}^{n_1} h_1(x_{i_1,1})\cdot...\cdot\prod_{i_r=1}^{n_r} h_r(x_{i_r,r}) \\
&\quad P_{\Phi_1\cdots\Phi_r}(d\{x_{1:n_1,1},...,x_{1:n_r,r}\}) \qquad (2.51)\\
&= \sum_{n_1=0}^{\infty}\cdots\sum_{n_r=0}^{\infty} p^{\Phi_1\cdots\Phi_r}_{N_1\cdots N_r}(n_1,...,n_r)\int_{\mathcal{X}_1^{n_1}}\cdots\int_{\mathcal{X}_r^{n_r}}\prod_{i_1=1}^{n_1} h_1(x_{i_1,1})\cdot...\cdot\prod_{i_r=1}^{n_r} h_r(x_{i_r,r}) \\
&\quad p^{\Phi_1\cdots\Phi_r}_{\mathcal{X}_1\cdots\mathcal{X}_r|N_1\cdots N_r}(x_{1:n_1,1},...,x_{1:n_r,r}|n_1,...,n_r)dx_{1:n_1,1}\cdots dx_{1:n_r,r} \qquad (2.52)\\
&= \sum_{n_1=0}^{\infty}\cdots\sum_{n_r=0}^{\infty}\frac{1}{n_1!\cdots n_r!}\int_{\mathcal{X}_1^{n_1}}\cdots\int_{\mathcal{X}_r^{n_r}}\prod_{i_1=1}^{n_1} h_1(x_{i_1,1})\cdot...\cdot\prod_{i_r=1}^{n_r} h_r(x_{i_r,r}) \\
&\quad p^{\Phi_1\cdots\Phi_r}_{n_1+\cdots+n_r}(x_{1:n_1,1},...,x_{1:n_r,r})dx_{1:n_1,1}\cdots dx_{1:n_r,r}, \qquad (2.53)
\end{aligned}$$

where $dx_{1:n_k,k} = dx_{1,k}\cdots dx_{n_k,k}$, $k = 1,...,r$ and $dx_{1:n_k,k} = 1$ for $n_k = 0$. The short–hand notation $x_{1:n_1,1},...,x_{1:n_r,r}$ denotes all points of a realization of the point processes $\Phi_1,...,\Phi_r$ with number of elements $n_1,...,n_r$, respectively. The probability that the processes $\Phi_1,...,\Phi_r$ have simultaneously $n_1,...,n_r$ points is denoted by $p^{\Phi_1\cdots\Phi_r}_{N_1\cdots N_r}(n_1,...,n_r)$, while the conditional probability that the processes $\Phi_1,...,\Phi_r$ have points $x_{1:n_1,1},...,x_{1:n_r,r}$, given that there are $n_1,...,n_r$ points is defined by $p^{\Phi_1\cdots\Phi_r}_{\mathcal{X}_1\cdots\mathcal{X}_r|N_1\cdots N_r}(x_{1:n_1,1},...,x_{1:n_r,r}|n_1,...,n_r)$. If $n_j = 0$, $j = 1,...,r$ the corresponding product $\prod_{i_j=1}^{n_j} h_j(x_{i_j})$ is defined to be one. Equations (2.52) and (2.53) hold if absolute continuity of $P_{\Phi_1\cdots\Phi_r}$ with respect to the Lebesgue measure is assumed. Analogously to Section 2.2 $p^{\mathcal{X}_1\cdots\mathcal{X}_r}_{n_1+\cdots+n_r}(x_{1:n_1,1},...,x_{1:n_r,r})$ denotes PDF evaluated at the unordered event and $P_{\Phi_1\cdots\Phi_r}(d\{x_{1:n_1,1},...,x_{1:n_r,r}\})$ denotes the probability measure that considers ordered events.

The PGFL of the process Φ_j, $j = 1,...,r$ can be obtained by marginalizing over all other processes, that is

$$\Psi_{\Phi_1\cdots\Phi_r}(1,...,1,h_j,1,...,1) = \Psi_{\Phi_j}(h_j). \qquad (2.54)$$

The marginalization (2.54) is obtained by substituting for all test–functions h_i, $i = 1,...,r$, $i \neq j$, $j \in \{1,...,r\}$ the identity function and integrating the corresponding integrals from (2.50) [SDK15].

In multitarget tracking, the processes Φ_i, $i = 1,...,r$ can either be used to model

multiple target state spaces and (at least one) measurement space, multiple measurement spaces and (at least one) target state space or multiple measurement and multiple target state spaces. Multiple target state spaces can be used to have target identification or separation of distinct target types. Multiple measurement spaces can be used to model multiple sensors or multiple scans. Details on the differences and similarities of tracking filters that employ multiple or single state spaces for the target processes and a single measurement space are presented in Chapter 3 of this thesis or in [SDK15].
For the following sections of this chapter we restrict ourselves to bivariate point processes, that is, to a joint point process that consists out of two finite point processes, one modeling the measurement, the other modeling the target process. In Chapter 3 it will be explained how the following considerations can be generalized to multivariate PGFLs.

2.9 PGFL of the Bayes Posterior Point Process

This section follows [Str13a, Section 4.2]. The PGFL of the Bayes posterior process and its intensity function are derived in the following. Let X and Y be the target state and measurement space, respectively. Typically, $X \subset \mathbb{R}^{d_1}$ and $Y \subset \mathbb{R}^{d_2}$, $d_1,\ d_2 > 0$. Let Υ be a finite point process with events in E_Y modeling the measurement process and let Ξ be a finite point process with elements in E_X modeling the target process. Then, under the assumption that $P_{\Xi\Upsilon}$ is absolutely continuous with respect to the Lebesgue measure, the bivariate PGFL is defined analogously to (2.50) on the product space $E_X \times E_Y$ as the expectation of the product

$$\prod_{x\in\Xi} h(x) \prod_{y\in\Upsilon} g(y), \tag{2.55}$$

that is

$$\begin{aligned}\Psi_{\Xi\Upsilon}(h,g) \equiv &\sum_{n=0}^{\infty}\sum_{m=0}^{\infty} \frac{1}{m!n!} \int_{X^n}\int_{Y^m} \prod_{j=1}^{n} h(x_j) \prod_{i=1}^{m} g(y_i) \\ &p_{n+m}^{\Xi\Upsilon}(x_1,...,x_n,y_1,...,y_m)dx_1\cdots dx_n dy_1\cdots dy_m,\end{aligned} \tag{2.56}$$

where

$$h : (X, \mathcal{B}(X)) \to (\mathbb{R}, \mathcal{B}(\mathbb{R})), \tag{2.57}$$

and

$$g : (Y, \mathcal{B}(Y)) \to (\mathbb{R}, \mathcal{B}(\mathbb{R})) \tag{2.58}$$

are bounded by one, non–negative Lebesgue–integrable test–functions. Analogously to (2.54), marginalization yields the PGFL of the remaining process, that is

$$\Psi_{\Xi\Upsilon}(h,1) = \Psi_{\Xi}(h) \text{ and } \Psi_{\Xi\Upsilon}(1,g) = \Psi_{\Upsilon}(g). \tag{2.59}$$

Further details on marginalization in terms of the PGFL are presented in Chapter 3. First, assume we have given the observations $y_1, ..., y_m \in Y$ of the finite point process defined on Y. Then, differentiation of $\Psi_{\Xi\Upsilon}$ with respect to m Dirac deltas at the measurements $y_1, ..., y_m \in Y$, evaluated at the test–function $g = 0$ yields

$$\frac{\partial^m \Psi_{\Xi\Upsilon}}{\partial y_1 \cdots \partial y_m}(h,0) = \sum_{n=0}^{\infty} \frac{1}{n!} \int_{X^n} \prod_{j=1}^{n} h(x_i) p_{n+m}^{\Xi\Upsilon}(x_1, ..., x_n, y_1, ..., y_m) dx_1 \cdots dx_n. \tag{2.60}$$

An additional marginalization with respect to the target process, that is evaluating (2.60) at $h = 1$ (the identity function) gives

$$\frac{\partial^m \Psi_{\Xi\Upsilon}}{\partial y_1 \cdots \partial y_m}(1,0) = p_m^{\Upsilon}(y_1, ..., y_m), \tag{2.61}$$

which is a constant. Therefore, the PGFL $\Psi_{\Xi|\Upsilon}$ of the Bayes posterior process $\Xi|\Upsilon$ is

$$\Psi_{\Xi|\Upsilon}(h|y_1, ..., y_m) = \sum_{n=0}^{\infty} \frac{1}{n!} \int_{X^n} \prod_{j=1}^{n} h(x_j) p_n^{\Xi|\Upsilon}(x_1, ..., x_n|y_1, ..., y_m) dx_1 \cdots dx_n \tag{2.62}$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \int_{X^n} \prod_{j=1}^{n} h(x_j) \frac{p_{n+m}^{\Xi\Upsilon}(x_1, ..., x_n, y_1, ..., y_m)}{p_m^{\Upsilon}(y_1, ..., y_m)} dx_1 \cdots dx_n \tag{2.63}$$

$$= \frac{\frac{\partial^m \Psi_{\Xi\Upsilon}}{\partial y_1 \cdots \partial y_m}(h,0)}{\frac{\partial^m \Psi_{\Xi\Upsilon}}{\partial y_1 \cdots \partial y_m}(1,0)}. \tag{2.64}$$

Here, (2.62) is due to the definition of the PGFL, (2.63) is justified by the application of Bayes Theorem and (2.64) is obtained by inserting (2.60) and (2.61). Note that for determining the PGFL of the Bayes posterior process only the PGFL of the joint process is needed. In Chapter 3 the joint PGFLs of several well–known filters are presented.
Equation (2.64) is presented in [Mah07b, p. 757] as the PGFL *form of the multitarget corrector*.

2.10 Summary Statistics of the Bayes Posterior Point Process

This section studies the summary statistics of the Bayes posterior point process analogously to [Str13a, Section 4.3]. The intensity function (or first moment density)

defined in (2.7) and (2.34) of the Bayes posterior process $\Xi|\Upsilon$ is defined at a target state $x \in X$ by the Gâteaux derivative of the PGFL of the Bayes posterior point process with respect to x, due to Section 2.5. Since the Bayes posterior process is given by (2.64) the intensity function is determined by the ratio of functional derivatives

$$f^{\Xi|\Upsilon}(x) = \frac{\frac{\partial^m \Psi_{\Xi\Upsilon}}{\partial x \partial y_1 \cdots \partial y_m}(1,0)}{\frac{\partial^m \Psi_{\Xi\Upsilon}}{\partial y_1 \cdots \partial y_m}(1,0)}. \tag{2.65}$$

The PGF of the RV of the number of targets in X $N_{\Xi|\Upsilon(\cdot)}(X)$ is (due to Section 2.7)

$$F^{\Xi|\Upsilon}(x) = \frac{\frac{\partial^m \Psi_{\Xi\Upsilon}}{\partial y_1 \cdots \partial y_m}(h,0)\Big|_{h(\cdot)=x}}{\frac{\partial^m \Psi_{\Xi\Upsilon}}{\partial y_1 \cdots \partial y_m}(1,0)}. \tag{2.66}$$

The posterior PDF of the canonical number is due to (2.45) given for all $n \in \mathbb{N}$ by

$$p_N^{\Xi|\Upsilon}(n) = \frac{1}{n!}\frac{d^n}{dx^n}F^{\Xi|\Upsilon}(0) = \frac{\frac{1}{n!}\frac{d^n}{dx^n}\left(\frac{\partial^m \Psi_{\Xi\Upsilon}}{\partial y_1 \cdots \partial y_m}(1,0)\right)\Big|_{x=0}}{\frac{\partial^m \Psi_{\Xi\Upsilon}}{\partial y_1 \cdots \partial y_m}(1,0)}, \tag{2.67}$$

and the expected number of targets is given by $\mathbb{E}[N^{\Xi|\Upsilon}] = \frac{d}{dx}F^{\Xi|\Upsilon}(1)$. The definitions (2.65)–(2.67) hold for general locally finite point processes. However, in general a point process does not need to have an absolutely continuous (with respect to the Lebesgue measure) moment measure, so that the intensity function is not necessarily the density of the first moment measure $\mu_{\Xi|\Upsilon}$ defined in (2.7).
Note that if $\Xi|\Upsilon$ is assumed to be a PPP

$$\mathbb{E}[N_{\Xi|\Upsilon}(X)] = \frac{d}{dx}F^{\Xi|\Upsilon}(1) = \int_{\mathcal{X}} f^{\Xi|\Upsilon}(x)dx. \tag{2.68}$$

holds as seen in Section 2.6.

2.11 Branching Process Form of the Bivariate PGFL

Let $A : (\Omega, \mathcal{A}, P) \to (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ be an RV that models the target process and $B : (\Omega, \mathcal{A}, P) \to (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ be an RV that models the measurement process. For simplicity of notation A and B are assumed to have the same domain and co–domain. Furthermore, let $B|A$ and $A|B$ denote the RV that model the conditional processes. Then, the definition of conditional probability yields (denoting the conditional probabilities of the RVs A and B also by P)

$$P(A,B) = P(A|B)P(B) = P(B|A)P(A) \tag{2.69}$$

and thus

$$P(A|B) = \frac{P(A,B)}{P(B)}. \tag{2.70}$$

Bayes Theorem for the conditional target process is obtained by replacing $P(A, B)$ by the right hand side of (2.69), that is

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}. \tag{2.71}$$

In Section 2.9 the PGFL of the Bayes Posterior process is derived using the bivariate PGFL of the joint target–measurement process. The derivation is done analogously to (2.70), replacing the random variables A and B by the finite point processes Ξ and Υ, respectively (see (2.62)–(2.64)). In this derivation the conditional measurement point process has not been used. The conditional measurement process $\Upsilon|\Xi$ is important to make inferences about the target processes by using the measurement information. In the following, according to [Str13a, Section 4.5], the bivariate PGFL is expressed as the composition of two functionals, that is, it is written in branching form. This is done analogously to the right hand side of (2.69). It holds, if $P_{\Upsilon|\Xi}$ is absolutely continuous with respect to the Lebesgue measure, that

$$\begin{aligned}
\Psi_{\Xi\Upsilon}(h, g) = &\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\frac{1}{m!n!}\int_{X^n}\int_{Y^m}\prod_{j=1}^{n}h(x_j)\prod_{i=1}^{m}g(y_i) \\
&p_{n+m}^{\Xi\Upsilon}(x_1, ..., x_n, y_1, ..., y_m)dx_1\cdots dx_n dy_1\cdots dy_m \qquad (2.72)\\
=&\sum_{n=0}^{\infty}\frac{1}{n!}\int_{X^n}\left(\sum_{m=0}^{\infty}\frac{1}{m!}\int_{Y^m}\prod_{i=1}^{m}g(y_i)p_m^{\Upsilon|\Xi}(y_1, ..., y_m|x_1, ..., x_n)dy_1\cdots dy_m\right) \\
&\prod_{j=1}^{n}h(x_j)p_n^{\Xi}(x_1, ..., x_n)dx_1\cdots dx_n \qquad (2.73)\\
=&\sum_{n=0}^{\infty}\frac{1}{n!}\int_{X^n}\Psi_{\Upsilon|\Xi}(g|x_1, ..., x_n)\prod_{j=1}^{n}h(x_j)p_n^{\Xi}(x_1, ..., x_n)dx_1\cdots dx_n, \qquad (2.74)
\end{aligned}$$

where the interchange of the sum and the integral in (2.73) is justified by absolute convergence [Str13a]. The definition of conditional probability for point processes yields analogously to (2.69)

$$p_{n+m}^{\Xi\Upsilon}(x_1, ..., x_n, y_1, ..., y_m) = p_m^{\Upsilon|\Xi}(y_1, ..., y_m|x_1, ..., x_n)p_n^{\Xi}(x_1, ..., x_n), \tag{2.75}$$

$y_1, ..., y_m \in Y$ and $x_1, ..., x_n \in X$ which is also used in (2.73).
According to Section 2.4 the PDF of the conditional measurement process $\Upsilon|\Xi$, in multitarget tracking better known as the *likelihood function*, is given by

$$p_m^{\Upsilon|\Xi}(y_1, ..., y_m|x_1, ..., x_n) = \frac{\partial^m\Psi_{\Upsilon|\Xi}}{\partial y_1\cdots\partial y_m}(0|x_1, ..., x_n). \tag{2.76}$$

If the PGFL of the conditional measurement process is given by the superposition of independent measurement processes, then the PGFL factors, that is

$$\Psi_{\Upsilon|\Xi}(g|x_1,...,x_n) = \prod_{i=1}^{n} \Psi_{\Upsilon|\Xi}(g|x_i), \tag{2.77}$$

where $\Psi_{\Upsilon|\Xi}(g|x_i)$, $x_i \in X$ is the measurement process conditioned on the ith target. Then, inserting (2.77) into (2.74) yields the branching process form of the bivariate PGFL as

$$\Psi_{\Xi\Upsilon}(h,g) = \Psi_{\Xi}\left(h\Psi_{\Upsilon|\Xi}(g|\cdot)\right), \tag{2.78}$$

where Ψ_{Ξ} denotes the PGFL of the target point process Ξ.

2.12 Point Processes With a Measure Comprising Dirac Measures

The Dirac measure [Alt12] is defined by

$$\delta_x(A) \equiv \begin{cases} 0, \text{ if } x \notin A \\ 1, \text{ if } x \in A \end{cases} \tag{2.79}$$

for all $A \in \mathcal{B}(\mathcal{X})$. A measure ν is called absolutely continuous with respect to the Lebesgue measure λ if for all $A \in \mathcal{B}(\mathcal{X})$

$$\lambda(A) = 0 \Rightarrow \nu(A) = 0 \tag{2.80}$$

[Alt12]. The Dirac measure is obviously not absolutely continuous with respect to the Lebesgue measure, since

$$\lambda(\{x\}) = 0, \text{ but } \delta_x(\{x\}) = 1, \text{ for all } x \in \mathcal{X}. \tag{2.81}$$

Furthermore, it holds that the Lebesgue and the Dirac measure are singular, since $\mathcal{X} \setminus \{x\} \cup \{x\} = \mathcal{X}$ with

$$\lambda(\mathcal{X} \setminus \{x\}) = 1 \text{ and } \lambda(\{x\}) = 0 \tag{2.82}$$

and

$$\delta_x(\{x\}) = 1 \text{ and } \delta_x(\mathcal{X} \setminus \{x\}) = 0. \tag{2.83}$$

Hence, there cannot exist a density of the Dirac measure with respect to the Lebesgue measure. Indeed, assume it exists. Then, the density of the Dirac measure would be given by a function $f : \mathcal{X} \to \mathbb{R}$ which satisfies

$$1 = \delta_x(\{x\}) = \int_{\{x\}} f(x)\lambda(dx), \tag{2.84}$$

and

$$0 = \delta_x(A) = \int_A f(x)\lambda(dx), \tag{2.85}$$

for all $A \in \mathcal{B}(\mathcal{X})$, $x \notin A$, that is f is equal to the Dirac delta, which cannot be a proper function due to the definition of the Lebesgue integral [Wal74]. Its behavior as an operator can only be approximated by a sequence of test functions (see [Wal74] or Chapter 4). As mentioned in Section 2.3 the Dirac delta function plays an important role in the derivation of the PGFL of the Bayes posterior process and its factorial moment measures (in case of existence). It will be investigated more closely in the context of functional differentiation in Chapter 4.

One class of point processes with first moment measures that comprise Dirac measures is given by Palm Processes [BES13]. Palm processes formalize conditioning on several points of the corresponding process. The distribution of an unreduced Palm process can be heuristically described as the law of a simple point process conditioned on the presence of points at given locations $x_1, ..., x_n \in \mathcal{X}$. The distribution of the reduced Palm process can informally be described as the law of the respective simple point process conditioning on the presence of points at given locations $x_1, ..., x_n \in \mathcal{X}$, excluding the considered points. In [Kar91, Section 1.7] unreduced Palm processes are developed using Dirac measures. As mentioned above, the density of the first moment measure of a Palm process can (due to the fact that its distribution comprises sums of Dirac measures) only be defined as a limit. In [BES13], [BSE15] reduced Palm distributions and the corresponding Papangelou intensity function are proposed for track extraction in multitarget tracking problems and the Bayes posterior process of the Palm process of the PHD filter is derived. The reduced Palm process removes the atoms in the Radon measure and therefore the measure is absolutely continuous and possesses a density.

Another class of point processes that have first moment measures, which comprises Dirac delta measures are point processes that are defined on discrete or partial discrete state spaces [Str14c]. Therefore, the first moment measure of these point processes comprise Dirac measures and thus the corresponding distribution is not absolutely continuous with respect to the Lebesgue measure. Hence, the corresponding density (intensity, PHD, etc.) can only be considered as a limit of test functions.

2.13 Point Processes versus Random Finite Sets

In [Str13a], [Str14e], [SDK15] finite point processes are used to derive multitarget tracking filters. The original derivation and first application of the closely related concept of random finite sets (RFS) to multitarget tracking is due to Mahler [Mah03], [Mah07a], [Mah07b]. In the following the difference between a point process and a random finite set is discussed.

In [Mah07b, p.349] an RFS Ψ is defined as *a random variable that draws its instantiations $\Psi = Y$ from the hyperspace $\mathfrak{V}$ of all finite subsets Y (the null set $\emptyset$ included) of some underlying space $\mathfrak{V}_0$*. Most of the time, in [Mah07b] it is assumed that $\mathfrak{V}_0 = \mathbb{R}^d$, $d > 0$. In [Mah07b, Appendix E] RFSs are compared with multisets and it is said that all elements in an RFS must be distinct and in [Mah07b, p.364] it is said that an RFS contains unordered events. Thus, we conclude that $\mathfrak{V}$ is given by the equivalence class

$$\mathfrak{V} = \Big\{\{y_1, ..., y_n\} : \ y_i \in \mathfrak{V}_0, \ y_1, ..., y_n \text{ distinct}, \ n \geq 0\Big\} / \sim, \tag{2.86}$$

where the set with $n = 0$ is defined to be the empty set and $\sim$ denotes the equivalence relation given by

$$\begin{aligned}\{y_1, ..., y_n\} \sim \{x_1, ...x_n\} \Leftrightarrow \text{ it exists a permutation } \sigma \text{ such that:} \\ \{y_{\sigma(1)}, ..., y_{\sigma(n)}\} = \{x_1, ..., x_n\}\end{aligned} \tag{2.87}$$

and an RFS is given by the RV

$$\Psi : (\Omega, \mathcal{A}, \mathbb{P}) \to (\mathfrak{V}, \mathcal{B}(\mathfrak{V})), \tag{2.88}$$

where $(\Omega, \mathcal{A}, \mathbb{P})$ is an arbitrary probability space. A locally finite point process defined in Section 2.1 on $(\Omega, \mathcal{A}, \mathbb{P})$ with state space $\mathfrak{V}_0$ is defined by the RV (see Section 2.1)

$$\Phi : (\Omega, \mathcal{A}, \mathbb{P}) \to (E_{\mathfrak{V}_0}, \mathcal{B}(E_{\mathfrak{V}_0})), \tag{2.89}$$

where

$$E_{\mathfrak{V}} \equiv \emptyset \cup \bigcup_{n \geq 1} \mathfrak{V}_0^{(n)} = \Big\{\{x_1, ..., x_n\} : x_i \in \mathfrak{V}_0, \ n \geq 0\Big\}. \tag{2.90}$$

is defined analogously to Equation (2.1) Here, the elements of $\mathfrak{V}^{(n)}$ need not be distinct. Hence they are multisets in the sense of [Knu98]. Assume that all elements φ of $E_{\mathfrak{V}_0}$ are simple and further assume that all elements φ are not only locally finite but globally finite, that is that not only each bounded but the complete set $\mathfrak{V}_0$ contains only finitely many points of φ. Such a process is called simple finite point process. It can be seen that RFSs and simple finite point processes are closely related. However, the two concepts are mathematically not the same.
The event space of a general finite point process contains multisets. A simple finite point process is characterized by

$$x_i, x_j \in \varphi, \quad x_i = x_j \Rightarrow i = j \text{ holds } \mathbb{P}\text{–a.s., for all } \varphi \in E_{\mathfrak{V}_0} \setminus \emptyset. \tag{2.91}$$

An RFS is an (unordered) set where the elements are necessarily distinct. Thus, constructing an RFS and a simple finite point process is not the same. To construct a simple finite point process Φ, first, the RV modeling the number of points is drawn

according to $p_N^{\Phi}(\cdot)$. Denote the realization by N. Afterwards, a RV according to $p_{\Phi|N}(\cdot|N)$ is drawn. This denotes the realization of the simple finite point process. No extra check on the distinction of the points is needed, since the condition on simplicity of a simple finite point process has to hold only $\mathbb{P}$–a.s (see Section 2.1). In contrast, when constructing a RFS, after drawing the realization an additional check on the distinction of the points in the realization is needed, since RFS are not allowed to contain repeated element, that is, the definition of *simple* in Section 2.1 holds with logical certainty. Obviously, this check is only relevant in theory, not in practice.
If $\mathfrak{V}_0$ is a discrete state space as mentioned in Section 2.12 it can be shown that the point process can no longer be simple [SKM95], that is, the event space contains multisets in the sense of [Knu98]. Thus, it cannot be modeled mathematically correctly via RFS–theory. In [Str14c] discrete intensity–filters are derived for the purpose of multitarget tracking.
Thus, we conclude that the classical concept of general finite point processes is a versatile tool for the purpose of multitarget tracking.

2.14 Conclusion

In this chapter well known fundamentals of point processes and PGFLs are presented that are needed to follow the discussions of the following chapters. First, the notion of a point process and the corresponding PGFL are defined. Afterwards, the functional derivative of a PGFL is defined and the likelihood function and various definitions of moments of PGFLs are derived using the functional derivative. As an illustrative example, which is often applied in practice, PPPs are studied. It is shown that multivariate PGFLs can be used to formulate the PGFLs of point processes for multitarget tracking applications and the derivation of summary statistics, needed to derive multitarget tracking filters for practical applications is shown. A special class of point processes defined on discrete spaces, which are of particular interest for practice, is discussed and finally point processes are compared to the closely related definition of an RFS.
Point processes are a versatile tool for modeling target and measurement processes. Their application to target tracking can be divided into two steps (see Figures 3.1 and 3.2). The first is called the *Discovery Step* and models the respective tracking filter by finding and exhibiting the corresponding PGFL. In Chapter 3 and [SDK15] it will be seen that many classic tracking filters can be modeled using point processes and that a formulation of tracking filters within a common framework helps not only to understand the connections and differences of existing filters, but also enables the search to find new tracking filters for demanding problems in target tracking. The second step, called the *Analytical Step*, derives the tracking filter of a PGFL analytically and closes the Bayesian recursion by filter–specific approximations. In Chapter 4 and [DSK15] summary statistics of a PGFL, which are needed for the formulation

of a tracking filter, are derived and justified.

CHAPTER III

The Family of Pointillist Filters

The complete content of this chapter has been published in [SDK15]. In the following a large number of well–known recursive, discrete–time, multitarget tracking filters is analyzed and characterized using point process theory. This classical theory can be applied if the sensor output is modeled by points. There, naturally the assignment problem arises, that is, which measurement belong to which target, which measurements are false measurements (clutter)? In [Rei79] all possible measurement to target assignments are represented by a complete set of assignment hypotheses for the standard assumption that a target generates at most one measurement per sensor scan. As discussed in the previous chapter an alternative is to model tracking filters by a joint target–measurement point process and characterize them via PGFLs in the *Discovery Step*. Then, the PGFL can be used to extract summary statistics of the tracking filter in the *Analytical Step*. Classic methods from analytic combinatorics can be applied [FS09] to derive the update equations of the respective tracking filter. It is shown that tracking filters which employ a joint target–measurement point process can be unified in a common framework, which is called the family of pointillist filters. PGFLs completely characterize finite point processes [Moy62] and different tracking filters can be identified by analyzing their PGFL. The corresponding PGFL can often be derived directly from the assumptions of the given tracking problem. In this chapter the *Discovery Step* for designing pointillist filters is proposed. It will be shown that a large number of well–known (and now classic) tracking filters are members of the family of pointillist filters. Additionally the PGFL description of tracking filters helps to investigate differences and similarities in their definition. Furthermore, the ease of formulating new or adapted filters using the proposed framework is demonstrated.

Three classes of pointillist filters are proposed, distinguishing filters by their application of target superposition. One class does *not* use target superposition, which implies that individual target states are maintained by the filter and each target

has its own state space (and thus a specific label). This class contains many well–known multitarget tracking filters, including Bayes–Markov [Jaz70], multi–hypothesis tracking (MHT) [Rei79], probabilistic data association (PDA) [BSF88], joint PDA (JPDA) [BSL95], probabilistic multi–hypothesis tracking (PMHT) [SL93], integrated PDA (IPDA) [MES94], and the joint IPDA (JIPDA) [ME04] filters.

The second class *does* use target superposition, that is only one target state space is present, which is shared by all targets. In particular, targets do not possess individual labels, which implies that enhanced target state extraction methods need to be applied. Note that target superposition might not be possible if the target state spaces are different. Pointillist filters that use superposition include the PHD intensity [Mah03], cardinalized PHD (CPHD) intensity [Mah07a], generalized PHD intensity [CM12], and the multi–Bernoulli intensity [Mah07b] [VVC09] [VVHM13] filters. Furthermore, originally non–superposing filters like the JPDA, the JIPDA, and the PMHT filters can be made members of this class by superposing the targets in a single target state space by a standard method. A close relation between the multi–Bernoulli filter and the JIPDA filter with superposition is recognized by comparing the respective joint target–measurement process PGFLs of the filters. This fact has already studied (independently) in [Wil14].

A third class of pointillist filters is a hybrid class in between the first two classes of filters. There, specified groups of targets are superposed in the same state space. Modifications of the PHD Intensity and generalized PHD Intensity filters represent members of this class of pointillist filters.

The unification of well–known tracking filters in a common framework illuminates similarities and differences in between them. Furthermore, it enables a data fusion engineer to design customized application–specific tracking filters using different components of an available tool–box of target and measurement models to obtain possibly new and unique filters, that are perfectly suited to solve the given problem. The demonstration on the practical application of the presented mathematical framework to design tracking filters is an important part of this chapter.

It is the completely undisputed and lasting contribution of Mahler to have first derived – among other contributions – a popular multiple target tracker by using PGFLs within the theory of RFSs. For several years, however, there has been an ongoing discussion in the tracking community of how to formulate the underlying mathematical structures and to link them to the classical mathematical theory of locally finite point processes that has been developed over decades.

It is the express intent of this chapter to bring different schools of the tracking community together by demonstrating that PGFLs are very precise and succinct models of the combinatorial probability structures involved in multitarget tracking, whether or not these models arise from a classical or an RFS approach. The family of pointillist filters and the connections to analytic combinatorics were first proposed and discussed

in [Str14b]. The first example of a tracking filter explicitly derived from a PGFL was the PHD intensity filter [Mah03]. The first application of PGFLs and the analytic combinatoric method to filters that do not employ target superposition was given for the PDA filter [Str14a].
The chapter is organized as follows. Section 3.1 discusses superposition and marginalization of target processes in terms of PGFLs. General notation and models are presented in Section 3.2. Section 3.2.2 presents models for target detection, and Section 3.2.3 discusses the clutter modeling. Section 3.3 (Section 3.4) presents the class of pointillist filters that do not (do) employ target superposition. Hybrid pointillist filters are discussed in Section 3.5. Closing the Bayesian recursion and target state estimation are briefly discussed in Sections 3.6 and 3.7, respectively. Section 3.8 shows the applicability of the proposed characterization approach for practical working engineers. In particular a concrete non–standard example involving unresolved targets is presented to show how to model a tracking filter using PGFLs. Section 3.9 gives conclusions and provides a table that overviews many of the filters discussed.
This chapter presents an overview of pointillist filters in terms of superposition of the target space. All presented filters have in common that they use a single measurement space, that is all measurements are superposed in one state space. Multi–scan and multi–sensor versions of pointillist filters can be realized by introducing additional measurement spaces, that is, especially additional test–functions have to be introduced. The approach is analogous to modeling pointillist filters with non–superposed targets (see Section 3.3). Furthermore, hybrid versions with partially superposed measurements can also be realized analogously to Section 3.5. Multi–scan and multi–sensor pointillist filters have already been developed, but are outside the scope of this thesis.

Own publications on this subject: The work presented in this chapter is published in [SDK15]. The generalized intensity filter presented in Section 3.4.4 is first derived in [Deg14]

3.1 Superposition and Marginalization of Finite Point Processes

In the following, target superposition and marginalization are explained in terms of the PGFL, since these concepts are used to separate pointillist filters into different classes. First, the special case of n independent target processes is used to intuitively explain the concept of target superposition and its representation in terms of a given PGFL.
Suppose that the PGFLs of $n \geq 1$ target processes are specified on the target state spaces X_i, $i = 1, \ldots, n$. Note that these spaces *do not* have to be identical, since each target has its own test–function. Denote the PGFL of the ith process by $\Psi_i(h_i)$,

where h_i is a complex-valued test function defined on X_i, that is, $h_i : X_i \to \mathbb{C}$. Then, the joint PGFL of the n processes is defined by

$$\Psi(h_1, \ldots, h_n) \equiv \prod_{i=1}^{n} \Psi_i(h_i). \tag{3.1}$$

The product form of the joint PGFL $\Psi(h_1, \ldots, h_n)$ holds if and only if the point processes corresponding to the test functions $h_1, ..., h_n$ are mutually independent (for a more general case that allows allows target correlation, the PGFL must be specified appropriately). In any event, realizations of the joint process are Cartesian products of finite point sets in the spaces X_i.
The marginalization with respect to one point process (which is in the studied example from above used to model a single target) is done by setting the corresponding test–functions equal to the identity function. Thus, the PGFL of the i^{th} marginal process is given by

$$\Psi_i(h_i) = \Psi(\ldots, 1, h_i, 1, \ldots). \tag{3.2}$$

Realizations of the ith marginal process are finite point sets in the space X_i. For mutually independent processes, the PGFLs of the marginal processes are identical to the factors in the product (3.1).
The n processes can be superposed if the spaces X_i are identical, which is denoted by X and referred to as the ground space. The fact that there is only one target state space implies that there is only one test–function. The PGFL of the superposed process is given by the *diagonal* of the joint PGFL, that is,

$$\Psi_X(h) \equiv \Psi(h, \ldots, h), \tag{3.3}$$

where the test function of the superposed process $h(\cdot)$ is defined on X. Realizations of the superposed process are finite point sets in X.
The filters of Section 3.3 do *not* use superposition, and therefore have as many (possibly different) target state spaces and test functions as there are targets. The pointillist filters of Section 3.4 *do* use superposition, and thus all targets share the same state space and there is only one test function. Finally, the hybrid pointillist filters discussed in 3.5 superpose some targets and not others. There, only the superposed targets share the same state space.

3.2 Notation and Models

In this section a consistent notation for the target and measurement models used for the rest of this chapter is presented, while specific probability distributions are not assumed.

The initial reference time is denoted by t_0. The measurement sample times are denoted by t_k, $k \in \mathbb{N}$. It is assumed that $t_{k-1} < t_k$ for all k. The recursive time index is suppressed throughout for ease of presentation.
Targets are assumed to be represented by states in the space denoted by X, where $X \subset \mathbb{R}^{\dim(X)}$. Measurements are points in a measurement space $Y \subset \mathbb{R}^{\dim(Y)}$.
The general models used for pointillist filters that do not superpose targets are discussed in this section. Point target and extended target detection models are discussed in Section 3.2.2. Clutter modelling is described in Section 3.2.3.

3.2.1 Target Motion and Measurement Models

For each target a prior PDF is specified at the recursion start time $t_0 \equiv t_{k-1}$. Six PDFs are needed:

- $\mu_0(x_0)$, the (prior) target PDF at time t_{k-1},
- $p_0(x|x_0)$, the Markovian target motion (transition) model from $x_0 \in X$ at time t_{k-1} to $x \in X$ at t_k,
- $\mu(x)$, the predicted target PDF at time t_k,
- $p(y|x)$, the likelihood function of a measurement $y \in Y$ at time t_k conditioned on target state $x \in X$ at t_k,
- $p(y)$, the PDF of a measurement at time t_k conditioned on the sequence of all measurements up to and including time t_{k-1},
- $p(x|y)$, the Bayes posterior PDF at time t_k conditioned on measurements up to time and including time t_k.

Three of these PDFs are determined by the others:

$$\mu(x) = \int_X \mu_0(x_0) p_0(x|x_0)\, dx_0 \tag{3.4}$$

$$p(y) = \int_X \mu(x) p(y|x)\, dx \tag{3.5}$$

$$p(x|y) = \frac{\mu(x) p(y|x)}{p(y)}. \tag{3.6}$$

The last expression is Bayes Theorem.

3.2.2 Target Detection Modeling

Different target detection models are proposed in this section. First, the standard assumption that a target generates *at most one measurement* is formulated and afterwards the more general extended target model, which enables a target to generate

multiple measurements that are distributed according to a joint distribution in the measurement space is proposed. Finally, the general target–oriented measurement model is presented in terms of a PGFL.

3.2.2.1 Targets With At Most One Point Measurement

Missed target detections are modeled by assuming that a target that is known to be present at state $x \in X$ at time t_k is detected with probability $P_k^D(x)$, where $0 \leq P_k^D(x) \leq 1$. The probability of missing the target detection is then $1 - P_k^D(x)$. Suppressing the recursion time index, for all $x \in X$ let

$$a(x) \equiv 1 - P_k^D(x) \quad \text{and} \quad b(x) \equiv P_k^D(x). \tag{3.7}$$

The corresponding PGF is defined by

$$G_{M|x}^{\text{BMD}}(z) \equiv a(x) + b(x)z, \tag{3.8}$$

where $z \in \mathbb{C}$. Target detection probabilities may or may be not be the same for all targets, depending on the application.

3.2.2.2 Extended Target Model

An extended target is assumed to have a well defined state $x \in X$, e.g., an appropriately defined centroid. It is assumed that extended targets generate a random number $M \geq 0$ of independent and identically distributed (point) measurements in the space Y. The distribution of a target–originated measurement is taken to be the likelihood function $p(y|x)$. The number of measurements can also depend on the target state. The conditional PGF of the RV M is defined by

$$G_{M|x}(z) \equiv \sum_{m=0}^{\infty} \Pr\{M = m \big| x\}\, z^m, \tag{3.9}$$

where $\Pr\{M = m| x\}$ denotes the probability that m measurements are generated by the extended target with state $x \in X$. The target is said to be detected if $M \geq 1$. For $M \geq 1$, let

$$d_m(x) \equiv \Pr\{M = m \big| \text{target at state } x \text{ is detected}\}, \tag{3.10}$$

so that $\sum_{m=1}^{\infty} d_m(x) = 1$. Using the probabilities (3.7) gives

$$G_{M|x}(z) = a(x) + b(x) \sum_{m=1}^{\infty} d_m(x) z^m \tag{3.11}$$

$$\equiv a(x) + b(x)\, G_{D|x}(z), \tag{3.12}$$

where $G_{D|x}(z)$ is the PGF of the number of measurements generated by a detected target at x. It reduces to the "at most one measurement per target" model for $M \equiv 1$, that is, to $G^{\mathrm{BMD}}_{M|x}(z)$ in (3.8) when $d_1(x) = 1$ and $G_{D|x}(z) = z$.

3.2.2.3 Generalized Target–Oriented Measurement Model

A more general target–oriented measurement model is used to formulated the generalized PHD intensity filters from Sections 3.4.4 and 3.5.2. The generalized PHD filter is originally derived in [CM12], the generalized intensity filter is presented first in (iFilter) [Deg14]. The general target–oriented measurement model used there cannot be formulated using a single PGF due to the fact that in this case measurements that correspond to one target can be correlated (measurements of different targets are assumed to be uncorrelated). Instead the target detection is incorporated into the PGFL of the target–oriented measurement process. The respective PGFL is defined by

$$\Psi^{\mathrm{gen}}_{\mathrm{BMD}}(h, g) \equiv \int_X h(x)\mu(x)\Bigg(p_0(\emptyset|x) + \sum_{n\geq 1} \frac{1}{n!} \int_{Y^n} \prod_{i=1}^{n} g(y_i) p_n(y_1, ..., y_n|x)\, dy_1 ... dy_n\Bigg)\, dx, \tag{3.13}$$

where $p_n(y_1, ..., y_n|x)$ denotes the generalized symmetric likelihood function, which is defined on Y^n. As mentioned in [CM12] the probability of detection is defined more generally than the standard detection model from Section 3.2.2.1. Therefore, $p_0(\emptyset|x)$ denotes the probability of a missed detection and $p_n(y_1, ..., y_n|x)$ includes the probability of detecting $(y_1, ..., y_n)^T \in Y^n$, given a specific target state $x \in X$.

3.2.3 Clutter Modeling

The clutter process (also called the false alarm process) in the measurement space Y is assumed to be either a non–homogeneous time–dependent Poisson point process (PPP), a generalization called a cluster process or an arbitrary general point process. For all cluster processes, including PPPs, given the number of points, the points are independently and identically distributed (i.i.d). In contrast, clutter models given by general point processes can be used to model correlated clutter measurements. The PGF of the probability distribution of the number of points plays a key role. PPPs are flexible, well understood, and widely used in diverse applications [Str10].
PPP clutter is considered first. Denote the intensity function of the clutter process at time t_k by $\lambda_k(y)$. The PPP is homogeneous if $\lambda_k(y) \equiv$ constant on Y; otherwise,

it is non-homogeneous. Time dependence is suppressed in the notation, so that $\lambda_k(y)$ is written $\lambda(y)$. The mean number of clutter points in Y is

$$\Lambda \equiv \int_Y \lambda(y)\, dy. \tag{3.14}$$

To assure that Λ is finite (and simplify the discussion), it is assumed that Y is bounded. Define the clutter PDF $p_\Lambda(y)$ to be the normalized intensity function,

$$p_\Lambda(y) \equiv \lambda(y)/\Lambda. \tag{3.15}$$

In this notation, the PGFL of the PPP clutter process is [Kar91]

$$\Psi_C^{\text{PPP}}(g) \equiv \exp\left(-\Lambda + \Lambda \int_Y g(y) p_\Lambda(y)\, dy \right), \tag{3.16}$$

where the test–function $g : Y \to \mathbb{C}$ is bounded by one, non–negative and Lebesgue–integrable.
The PGFL of clutter when modeled as a cluster process is defined by

$$\Psi_C^{\text{Cluster}}(g) \equiv G_C\left(\int_Y g(y) q(y) dy \right), \tag{3.17}$$

where $q(y)$ is the PDF of a clutter point $y \in Y$ and $G_C(z)$ is the PGF of the number of clutter points, namely,

$$G_C(z) \equiv \sum_{c=0}^{\infty} \Pr\{C = c\} z^c, \tag{3.18}$$

where $\Pr\{C = c\}$ denotes the probability that c clutter measurements are generated. Note that (3.16) is a special case of (3.17) for $G_C(z) = \exp(-\Lambda + \Lambda z)$.
The PGFL of an arbitrary general clutter process is denoted by $\Psi_C^{\text{gen}}(g)$ in the following.

3.3 Pointillist Filters without Superposition

The joint PGFLs of several well–known recursive discrete–time tracking filters that do not superpose targets are exhibited in this section. The general notation of Section 3.2 is used, and the recursive time index is suppressed. Throughout this chapter, PGFLs and PGFs will be denoted, respectively, by $\Psi(\cdot)$ and $G(\cdot)$.

3.3.1 Bayes–Markov Filter

3.3.1.1 Classical Problem

The classical Bayes–Markov filter [Jaz70], [SSCB14] is presented in the following. It assumes exactly one target to be always present. It is always detected, and it generates exactly one measurement. There is no clutter. Denote the target and measurement test functions by $h : X \to \mathbb{C}$ and $g : Y \to \mathbb{C}$, respectively. Both are bounded by one, non–negative and Lebesgue–integrable. The PGFL is

$$\Psi_{\mathrm{BM}}(h, g) \equiv \int_X \int_Y h(x)g(y)\mu(x)p(y|x)\, dy\, dx. \tag{3.19}$$

As a check, note that $\Psi_{\mathrm{BM}}(1, 1) \equiv \Psi_{\mathrm{BM}}(h, g)|_{h(\cdot)=1,g(\cdot)=1} = 1$, due to the fact that $\mu(x)$ and $p(y|x)$ are PDFs.

3.3.1.2 With Missed Detections

Missed target detections are modeled using (3.7) [Jaz70], [SSCB14]. The joint target–originated measurement process is

$$\begin{aligned}\Psi_{\mathrm{BMD}}(h, g) &\equiv \int_X h(x)\mu(x)G^{\mathrm{BMD}}_{M|x}\left(\int_Y g(y)p(y|x)\, dy\right) dx \\ &= \int_X h(x)\mu(x)\Big(a(x) + b(x)\int_Y g(y)p(y|x)\, dy\Big)dx.\end{aligned} \tag{3.20}$$

As a check, note that $\Psi_{\mathrm{BMD}}(1, 1) = 1$. This expression shows up often in the sequel. It reduces to (3.19) when the target detection probability is one.

3.3.1.3 With Missed Detections and Extended Target

The PGF (3.9) of the number of measurements generated by the extended target with state $x \in X$ is assumed to be known. The PGFL for the extended target is [Jaz70], [SSCB14], [GGMS05a]

$$\Psi_{\mathrm{BME}}(h, g) \equiv \int_X h(x)\mu(x)G_{M|x}\Big(\int_Y g(y)p(y|x)\, dy\Big)dx. \tag{3.21}$$

If the target can generate at most one measurement, then $M \in \{0, 1\}$ and $G_{M|x}(z) = a(x) + b(x)\, z,\ \ z \in \mathbb{C}$. Thus, (3.20) is a special case of (3.21).

3.3.2 PDA Filter

3.3.2.1 Without Gating

In the PDA problem [BSF88], [BSL95], as in the classical Bayes–Markov problem, exactly one target is assumed to be present. However, the target may or may not be detected, and clutter may be also present. Realizations of the clutter process and the target measurement process are superposed in the sensor measurement space Y. The target measurement and clutter processes are assumed to be independent. The PGFL of superposed independent processes is the product of the PGFLs of the constituent processes [SR01]. The target–originated measurement PGFL is given by (3.20) and the clutter PGFL is (3.16), so the product

$$\begin{aligned}\Psi_{\mathrm{PDA}}^{\mathrm{noGate}}(h,g) =& \Psi_{C}^{\mathrm{PPP}}(g)\,\Psi_{\mathrm{BMD}}(h,g)\\ =& \exp\left(-\Lambda+\Lambda\int_Y g(y)p_\Lambda(y)\,dy\right)\\ &\times\int_X h(x)\mu(x)\left(a(x)+b(x)\int_Y g(y)p(y|x)\,dy\right)dx. \end{aligned} \tag{3.22}$$

is the PGFL of the superposed processes.

3.3.2.2 With Gating

The PGFL is modified to accommodate gating. The gate, denoted by Γ, is a specified subset of Y. The probability that a target–originated measurement falls within the gate is

$$P_\Gamma \equiv \Pr\{y\in\Gamma\} = \int_\Gamma p(y)\,dy. \tag{3.23}$$

The gate can be chosen arbitrarily at time t_k so long as $\Pr\{y\in\Gamma\}\neq 0$; however, it is typically chosen so $\Pr\{y\in\Gamma\}$ is near one. Gating the PPP clutter process to Γ yields a PPP [Str10] whose PGFL is given by

$$\Psi_{C}^{\mathrm{PPPgated}}(g) = \exp\left(-|\Gamma|+|\Gamma|\int_\Gamma g(y)p_\Gamma(y)\,dy\right), \tag{3.24}$$

where the expected number of clutter points in the gate is

$$|\Gamma| \equiv \int_\Gamma \lambda(y)\,dy \tag{3.25}$$

and the clutter PDF is the intensity normalized by gate volume:

$$p_\Gamma(y) = \begin{cases} \lambda(y)/|\Gamma|, & \text{if } y\in\Gamma\\ 0, & \text{if } y\notin\Gamma. \end{cases} \tag{3.26}$$

Censoring target measurements lying outside Γ gives (cf. (3.20))

$$\Psi^{\text{gated}}_{\text{BMD}}(h,g) = \int_X h(x)\mu(x)\bigg(a_\Gamma(x) + b_\Gamma(x)\int_\Gamma g(y)p_\Gamma(y|x)\,dy\bigg)dx,$$

where the probability that a target at $x \in X$ is detected *within* the gate is

$$b_\Gamma(x) \equiv b(x)\,P_\Gamma, \tag{3.27}$$

so that $a_\Gamma(x) \equiv 1 - b(x)P_\Gamma$ is the probability that it is not detected in the gate, and the conditional PDF of a gated target–originated measurement is

$$p_\Gamma(y|x) = \begin{cases} (P_\Gamma)^{-1}\,p(y|x), & \text{if } y \in \Gamma \\ 0, & \text{if } y \notin \Gamma. \end{cases} \tag{3.28}$$

Using these expressions gives

$$\begin{aligned}\Psi_{\text{PDA}}(h,g) \equiv& \Psi_C^{\text{PPPgated}}(g)\,\Psi^{\text{gated}}_{\text{BMD}}(h,g) \\ =& \exp\bigg(-|\Gamma| + |\Lambda|\int_\Gamma g(y)p_\Lambda(y)\,dy\bigg) \\ &\times \int_X h(x)\mu(x)\bigg(1 - b(x)P_\Gamma + b(x)\int_\Gamma g(y)p(y|x)\,dy\bigg)dx. \end{aligned} \tag{3.29}$$

This expression holds only for $y \in \Gamma$. The gated PGFL reduces to the PGFL for ungated measurements when the gate is the entire measurement space, i.e., $\Gamma = Y$.

3.3.2.3 Extended Target

Given the PGF (3.9) of the number of measurements generated by the extended target, the PGFL for the ungated extended target is the product

$$\begin{aligned}\Psi_{\text{PDAE}}(h,g) =& \Psi_C^{\text{PPP}}(g)\,\Psi_{\text{BME}}(h,g) \\ =& \exp\bigg(-\Lambda + \Lambda\int_Y g(y)p_\Lambda(y)\,dy\bigg) \\ &\times \int_X h(x)\mu(x)G_{M|x}\bigg(\int_Y g(y)p(y|x)\,dy\bigg)dx. \end{aligned} \tag{3.30}$$

Gating the extended target measurements requires counting the number of ways that m target–originated measurements can fall within the gate. It is not considered further here.

3.3.3 JPDA Filter

Exactly $n \geq 1$ targets are assumed to be present. Clutter is assumed to be present. Target measurement and clutter processes are assumed to be mutually independent. Gating can be accommodated in the manner outlined below for the JIPDA filter in Section 3.3.6.2, so it is not done here.

The targets may or may not be detected. If detected, a target generates at most one measurement. All measurements generated are superposed in the same measurement space Y. Which measurements are target–originated and which are clutter is unknown. The PGFL of the ith target measurement process is given by

$$\begin{aligned}\Psi_{\mathrm{BMD}(i)}(h_i, g) &= \int_{X_i} h_i(x)\mu_i(x) G^{\mathrm{BMD}}_{M_i|x}\left(\int_Y g(y)p(y|x)\,dy\right) dx \\ &= \int_{X_i} h_i(x)\mu_i(x)\left(a_i(x) + b_i(x)\int_Y g(y)p_i(y|x)\,dy\right)dx,\end{aligned} \tag{3.31}$$

where the quantities in this expression are the same as in (3.20) but with an index i added to make them specific to the ith target. The state spaces X_i can be different or they can be copies of the same space.

Target processes are not superposed in a common state space. Consequently, a different test function $h_i(x) : X_i \to \mathbb{R}$ is necessary for each target. In contrast, the same measurement test function g is used because all measurements, whether target–originated or clutter, are superposed in the measurement space Y.

Because clutter and target measurement processes are mutually independent, the joint PGFL is the product of PGFLs:

$$\Psi_{\mathrm{JPDA}}(h_1, \ldots, h_n, g) = \Psi_C^{\mathrm{PPP}}(g) \prod_{i=1}^{n} \Psi_{\mathrm{BMD}(i)}(h_i, g), \tag{3.32}$$

where $\Psi_C^{\mathrm{PPP}}(g)$ is given by (3.16).

The PGFL of the marginal process for the jth target, $j = 1, \ldots, n$, is found by setting $h_i(x) = 1$ for $i \neq j$. Explicitly,

$$\Psi_{\mathrm{JPDA}(j)}(h_j, g) = \Psi_C^{\mathrm{PPP}}(g)\Psi_{\mathrm{BMD}(j)}(h_j, g) \prod_{i=1,\, i\neq j}^{n} \Psi_{\mathrm{BMD}(i)}(1, g). \tag{3.33}$$

The marginal PGFLs are used for approximation purposes. They do not characterize the full JPDA distribution. For the now–standard derivation of the JPDA filter see [BSF88], [BSL95].

3.3.4 PMHT Filter

The PMHT filter and enhanced versions were published in [SL93], [SL95], [SL94]. It is assumed that there are exactly $n \geq 1$ targets present. Target–originated measurements are modeled as PPPs that are conditioned on target state $x \in X$. The mean

number of measurements generated by the ith target is denoted by $\Lambda_i(x)$, and the conditional likelihood is $p_i(y|x)$. The PGF of the ith target–originated measurement process conditioned on a target at $x \in X$ is (*cf.* (3.9))

$$G^{\text{PPP}}_{M_i|x}(z) \equiv \exp\big(-\Lambda_i(x) + \Lambda_i(x)z\big). \tag{3.34}$$

Although not formulated in terms of PGFLs, Poisson models for the number of measurements were first incorporated into the PMHT and histogram PMHT filters in [Dav15], [VDFL15].

The PGFL of the ith target is

$$\begin{aligned}\Psi^{\text{PPP}}_{\text{BMD}(i)}(h_i, g) &= \int_X h_i(x)\mu_i(x) G^{\text{PPP}}_{M_i|x}\left(\int_Y g(y)p_i(y|x)\,dy\right)\,dx \\ &= \int_X h_i(x)\mu_i(x)\exp\left(-\Lambda_i(x) + \Lambda_i(x)\int_Y g(y)p_i(y|x)\,dy\right).\end{aligned} \tag{3.35}$$

This expression differs from the JPDA model (3.31) only in the choice of the conditional PGFL of the measurements. The ith target model reduces to the extended target measurement model (3.21) when the number of measurements is Poisson distributed with mean $\Lambda_i(x)$, that is, when $G_{M|x}(z) \equiv \exp\big(-\Lambda_i(x) + \Lambda_i(x)z\big)$.

Targets are mutually independent by assumption, so the joint PGFL of the PMHT filter is the product of the target–specific PGFLs,

$$\Psi_{\text{PMHT}}(h_1, \ldots, h_n, g) = \prod_{i=1}^{n} \Psi^{\text{PPP}}_{\text{BMD}(i)}(h_i, g). \tag{3.36}$$

As with JPDA, different targets have different test functions. PMHT, unlike JPDA, models clutter using a "diffuse" target, that is, as a target with a high measurement variance.

The PGFL (3.36) is, by inspection, a linear functional (see the discussion following (2.40)) in the test functions $h_1, \ldots, h_n$. From the discussion in Section 2.5, it follows that realizations of the joint target point process have, with probability one, exactly one target in each state space, and that the Bayes posterior PDF and the intensity functions are identical.

Denote a realization by $x_i \in X_i$, $i = 1, \ldots, n$. It can be shown that the joint posterior PDF is proportional to the product (over the measurements) of a probabilistic mixture of measurement likelihoods. The n mixing proportions are

$$\pi_i(x_i) = \Lambda_i(x_i) \Big/ \sum_{j=1}^{n} \Lambda_j(x_j), \qquad \text{for } i = 1, \ldots, n. \tag{3.37}$$

This result is reasonable given the well–known relationship between Poisson mixtures and the multinomial distribution [Kin92], [JKB97], [Gri84]. Details are straightforward and are omitted. The original PMHT filter assumes that the spatial rates are

constant, $\Lambda_i(x) \equiv \Lambda_i$, and applies the EM method to find point estimates for the n targets and other parameters.

3.3.5 IPDA Filter

The IPDA presented in [MES94], [MS08], [ME94] [CMME11], [MSS14] shares all the assumptions of the PDA except for one – it assumes that *at most* one target exists. Let $\chi_0,\ 0 \leq \chi_0 \leq 1$, denote the initial probability that the target exists, i.e., that there is $N = 1$ target. Let χ denote the updated probability of target existence at current time (see [MES94] for details). The PGF of the number of targets at the current time is therefore

$$G_N(z) = 1 - \chi + \chi\, z. \tag{3.38}$$

The PGFL for the target–generated measurement process is

$$\Psi_{\mathrm{BM}\chi}(h, g) = G_N(\Psi_{\mathrm{BMD}}(h, g)) = 1 - \chi + \chi\, \Psi_{\mathrm{BMD}}(h, g), \tag{3.39}$$

where $\Psi_{\mathrm{BMD}}(h, g)$ is given by (3.20). Superposing an independent clutter process gives the PGFL for IPDA as

$$\Psi_{\mathrm{IPDA}}(h, g) = \big(1 - \chi + \chi\, \Psi_{\mathrm{BMD}}(h, g)\big)\, \Psi_C^{\mathrm{PPP}}(g), \tag{3.40}$$

where the clutter PGFL $\Psi_C^{\mathrm{PPP}}(g)$ is given by (3.16).
In [CVW02] the IPDA filter is derived using RFSs. In [MLS08] and [ME02] further versions of the IPDA filter are presented. These investigations may represent another class of pointillist filters, but they are outside the scope of this thesis.

3.3.6 JIPDA Filter

3.3.6.1 Without Gating

The JIPDA filter was presented first in [ME04] (see also [CMME11], [MSS14]). The JIPDA process assumes that at most $n \geq 1$ targets exist. Let χ_i denote the existence probability for target i, i.e., the probability that target i is present. The target–originated measurement process for target i is, using the same notation as in (3.31),

$$\Psi_{\mathrm{IPDA}(i)}(h_i, g) = 1 - \chi_i + \chi_i\, \Psi_{\mathrm{BMD}(i)}(h_i, g). \tag{3.41}$$

Different test functions h_i are needed because the each target has its own state space. Assuming the target–originated measurement processes are mutually independent of

each other and of the clutter process gives the PGFL for JIPDA as

$$\begin{aligned}\Psi_{\text{JIPDA}}(h_1,\ldots,h_n,g) =& \Psi_C^{\text{PPP}}(g)\prod_{i=1}^{n}\Psi_{\text{IPDA}(i)}(h_i,g)\\ =& \Psi_C^{\text{PPP}}(g)\prod_{i=1}^{n}\Big(1-\chi_i+\chi_i\,\Psi_{\text{BMD}(i)}(h_i,g)\Big).\end{aligned} \tag{3.42}$$

As checks, note that $\Psi_{\text{JIPDA}}(1,\ldots,1)=1$ and that $\Psi_{\text{JIPDA}}(\cdot)$ reduces to $\Psi_{\text{JPDA}}(\cdot)$ for $\chi_i=1$.
Analogously to the JPDA, the PGFL of the marginal process for the jth target, $j=1,...,n$, can be determined by setting $h_i(x)=1$ for $i\neq j$. It is given by

$$\begin{aligned}\Psi_{\text{JIPDA}(j)}(h_j,g) =& \Psi_C^{\text{PPP}}(g)\Big(1-\chi_j+\chi_j\,\Psi_{\text{BMD}(j)}(h_j,g)\Big)\\ &\times\prod_{i=1,i\neq j}^{n}\Big(1-\chi_i+\chi_i\,\Psi_{\text{BMD}(i)}(1,g)\Big).\end{aligned} \tag{3.43}$$

The marginal defined in (3.43) can be used in various ways to approximate the joint distribution and thereby close the Bayesian recursion.
In [MC03], a closely related version of the JIPDA is presented using RFSs. The versions differ only in the method used to start new tracks.

3.3.6.2 With Gating

Gating is incorporated in a manner analogous to that of Section 3.3.2.2 for the PDA. Denote the measurement gate of the ith target by $\Gamma_i\subset Y$, $i=1,...,n$. Let $\Gamma\equiv\cup_{i=1}^{n}\Gamma_i$. Assuming PPP clutter, the expected number of clutter points, $|\Gamma|$, in Γ is given by (3.25). The corresponding clutter PDF $p_\Gamma(y)$ is given by (3.26). The gated clutter process is a PPP, and its PGFL is

$$\Psi_C^{\text{PPPgated}}(g)=\exp\Big(-|\Gamma|+|\Gamma|\int_\Gamma g(y)p_\Gamma(y)dy\Big). \tag{3.44}$$

Let P_{Γ_i} denote the probability that a target–originated measurement falls within the gate of the ith target, that is

$$P_{\Gamma_i}\equiv\Pr\{y\in\Gamma_i\}=\int_{\Gamma_i}p(y)\,dy. \tag{3.45}$$

Then, the PGFL of the joint target–measurement process for the ith target is given by

$$\Psi_{\text{BMD}(i)}^{\text{gated}}(h_i,g)=\int_X h(x)\mu(x)\Big(a_{\Gamma_i}(x)+b_{\Gamma_i}(x)\int_{\Gamma_i}g(y)p_{\Gamma_i}(y|x)\,dy\Big), \tag{3.46}$$

where the probability that the ith target at $x \in X_i$ is detected *within* the gate is

$$b_{\Gamma_i}(x) = b(x)\, P_{\Gamma_i}, \tag{3.47}$$

so that $a_{\Gamma_i}(x) = 1 - b(x)P_{\Gamma_i}$ is the probability that it is not detected in the gate, and the conditional PDF of a gated target–originated measurement is

$$p_{\Gamma_i}(y|x) = \begin{cases} (P_{\Gamma_i})^{-1}\, p(y|x), & \text{if } y \in \Gamma_i \\ 0, & \text{if } y \notin \Gamma_i, \end{cases} \tag{3.48}$$

which means that each target has its own gate and its likelihood function is normalized, so that it is a PDF within this gate. Therefore, the PGFL of the Bayes posterior process of the gated JIPDA filter is given by

$$\Psi^{\text{gated}}_{\text{JIPDA}}(h_1, \ldots, h_n, g) = \Psi_C^{\text{PPPgated}}(g) \prod_{i=1}^{n} \left(1 - \chi_i + \chi_i\, \Psi^{\text{gated}}_{\text{BMD}(i)}(h_i, g)\right). \tag{3.49}$$

The PGFL of the marginal process for the JIPDA including gating is determined analogously to Section 3.3.6.1.

3.3.7 MHT Filter

In the standard MHT algorithm, presented in [Rei79], the filter *generates a set of data association hypotheses to account for all possible origins of every measurement.* In a specific iteration the hypothesis are generated, that a certain target is a false alarm, belongs to an existing target, or occurs due to a new target. Based on the target states various techniques for reducing the numerical complexity like gating, merging and pruning were proposed by the community. However, there is only one un–pruned and complete set of MHT hypothesis.
Assume that the JIPDA filter derived from (3.42) is extended by a process that allows data induced targets. Then, assuming m measurements were received it is given by

$$\begin{aligned} \Psi_{\text{MHT}}(h_1, \ldots, h_{n+m}, g) =& \Psi_C^{\text{PPP}}(g) \prod_{i=1}^{n} \left(1 - \chi_i + \chi_i\, \Psi_{\text{BMD}(i)}(h_i, g)\right) \\ & \times \prod_{j=1}^{m} \left(1 - \gamma_j + \gamma_j\, \Psi^{\text{Data}}_{\text{BMD}(j)}(h_{n+j}, g)\right), \end{aligned} \tag{3.50}$$

where $0 \leq \gamma_j \leq 1$ denotes the probability that measurement j was generated by a target that is not modeled yet, and

$$\Psi^{\text{Data}}_{\text{BMD}(j)}(h_j, g) \equiv \int_X h_j(x)\xi_j(x) G^{\text{BMD}}_{M|x} \left(\int_Y g(y) p_j(y|x)\, dy \right) dx \tag{3.51}$$

is a data–driven PGFL that assumes that this new target has (specified) *a priori* PDF $\xi_j(x)$ and generated a measurement with conditional PDF $p_j(y|x)$. The test functions $h_{n+1}, ..., h_{n+m}$ correspond to the m data-induced PGFLs.
The PGFL given by the data–driven JIPDA in (3.50) captures the complete un–pruned set of all hypothesis of a single scan MHT algorithm. The correct Bayesian way to derive the target state estimates is to apply marginalization over all except one target state, exactly as it is done in the JIPDA filter.
To keep the numerical complexity feasible, many approximation methods (gating, merging) and decision techniques (pruning) can be applied to close the Bayesian recursion. These considerations are outside the scope of this thesis.
The close relationship between the IPDA and the MHT filters was first noted in [ME05]. However, the connection was not discussed there in terms of PGFLs. The connection between MHT, JIPDA, and multi–Bernoulli filters has been studied using the framework of RFSs in [Wil14].
The study of a multiple scan MHT filter, i.e., an MHT filter that keeps measurement origin hypotheses for several successive scans, has already been studied. It is outside the scope of this thesis and will be presented in a further publication.

3.4 Pointillist Filters with Superposition

In this section the PGFLs of filters with target superposition are presented, that is all targets share the same target state space.
The general notation of Section 3.2 is used throughout, and the recursive time index is suppressed.

3.4.1 Superposition in JPDA and Other Multitarget Filters

The multitarget processes in Section 3.3 are such that each target has its own state space. When the state spaces are copies of the same space, then it is possible to superpose the target processes into one process on the common space. It is seen in this section that it is straightforward to superpose targets once the joint multitarget PGFL is known. In the following it is shown how three originally non–superposed filters can be superposed.
One problem with superposition is the loss of target–specific labels. As noted in Section 3.7, estimating target–specific PDFs from the superposed process requires separate procedures that are replete with their own special difficulties.
Another problem, noted below in Section 3.6, is that it mis–models multiple targets. The PDF $\mu(\cdot)$ plays a more significant role for superposed targets than it does for one target. This is because the target states are (typically) i.i.d. samples from an RV whose PDF is $\mu(\cdot)$. The mis–modeling would disappear only if the samples could somehow be drawn without replacement from different targets.

3.4.1.1 JPDA

The joint probability distribution in JPDA is marginalized over all but one of the n targets to obtain a target–specific PDF. When the target state spaces are all the same, i.e., $X_i \equiv X$ for all i, target point processes can be superposed instead of marginalized. The PGFL of the JPDA with superposition (JPDAS) filter is found by substituting

$$h_i(\cdot) \equiv h(\cdot) \tag{3.52}$$

into the PGFL (3.32) for the JPDA. The resulting PGFL is

$$\Psi_{\text{JPDAS}}(h,g) = \Psi_{\text{JPDA}}(h,\ldots,h,g) = \Psi_C^{\text{PPP}}(g)\prod_{i=1}^{n}\Psi_{\text{BMD}(i)}(h,g). \tag{3.53}$$

In JPDAS, as in JPDA, targets can have different motion models, measurement models, and detection probabilities. Note too that the superposition procedure can be limited so that it includes some targets and not others.
A label–free version of the JPDA filter, called set JPDA (SJPDA) filter has already been proposed in [SSGW11a].

3.4.1.2 JIPDA, PMHT, and Other Filters

Targets can be superposed in other joint multitarget point processes with known PGFLs when identical target state spaces are assumed. The PGFL of the superposed process is found by substituting (3.52). For example, from (3.42) and (3.36),

$$\Psi_{\text{JIPDAS}}(h,g) = \Psi_{\text{JIPDA}}(h,\ldots,h,g) \tag{3.54}$$

$$\Psi_{\text{PMHTS}}(h,g) = \Psi_{\text{PMHT}}(h,\ldots,h,g) \tag{3.55}$$

are, respectively, the PGFLs for JIPDA and PMHT with superposition.

3.4.2 PHD Intensity Filter

The probability hypothesis density (PHD) filter is proposed in [Mah07b] and the papers [Mah03] and [Mah07a]. The intensity filter (iFilter) is proposed in [Str10]. The difference between these filters lies in their state space. While the PHD filter uses the state space X (typically $X = \mathbb{R}^d$, $d > 0$), the iFilter uses an augmented state space $X^+ = X \cup \phi$, where X is equal to the state space of the PHD filter and ϕ is the hypothesis space, used to model clutter by scatterers. By a careful choice of certain parameters the Bayes posterior of the iFilter on X^+ is the same as the Bayes posterior of the PHD filter. For details see [SSCB14]. A detailed comparison of both filters using the PGFL derivation can be found in [Str13a].
Targets are assumed to have the same state space X and are superposed. It is assumed

that the targets constitute a locally finite point process at the recursion start time. This is the prior process at the initial (start up) time, but thereafter it is the Bayes posterior point process. The PGFL of the process, denoted by $\Psi_0(h, g)$, is typically not a PPP, so it is approximated by a PPP to close the Bayes recursion. To this end, the joint conditional PDF is approximated by the product of the marginal PDFs. Details can be found in [Mah07b] and [Str10].

It is important that the predicted target process be a PPP. This is guaranteed to be the case assuming independent probabilistic thinning (target death), and assuming independent Markovian targets having the same motion model $p_0(x|x_0)$ of Section 3.2.1. It also holds if a new–target (birth) process is superposed, provided the birth process is a PPP independent of the target process.

The details of the predicted target PPP are of little concern here – it suffices to assume that it is a (recursively) specified PPP with intensity $\overline{N}\,\mu(x)$, where the predicted number of targets is $\overline{N} \equiv E[N]$, N being the RV modeling the number of targets, and $\mu(x)$ is the predicted target PDF. The mean $\overline{N}$ and PDF $\mu(x)$ are determined by details of the prediction process. (In particular, note that $\mu(x)$ is not given by (3.4) except in special cases.)

Since N is Poisson distributed with mean $\overline{N}$, its PGF is (when there is at most one target, see (3.38))

$$G_N^{\text{PPP}}(z) = e^{-\overline{N}+\overline{N}z}. \tag{3.56}$$

The predicted target states are i.i.d. by PPP assumption and are drawn from the PDF $\mu(x)$. The target measurement functions and detection probabilities are assumed to be the same for all targets. Hence, the PGFL of the predicted target–originated measurement process is

$$\Psi_{\text{PHD}}^{\text{targets}}(h, g) = G_N^{\text{PPP}}\big(\Psi_{\text{BMD}}(h, g)\big), \tag{3.57}$$

where $\Psi_{\text{BMD}}(h, g)$ is given by (3.20). The target–originated measurement process is superposed with the independent clutter process, so the PGFL of the PHD filter is the product

$$\Psi_{\text{PHD}}(h, g) = \Psi_C^{\text{PPP}}(g)\, G_N^{\text{PPP}}\big(\Psi_{\text{BMD}}(h, g)\big). \tag{3.58}$$

Substituting (3.16), (3.56), and (3.20) gives the explicit form

$$\Psi_{\text{PHD}}(h, g) = \exp\left(-\Lambda - \overline{N} + \Lambda \int_Y g(y) p_\Lambda(y)\, dy \right.$$
$$\left. + \overline{N} \int_X h(x)\mu(x)\left(a(x) + b(x)\int_Y g(y)p(y|x)\, dy\right)dx\right). \tag{3.59}$$

It has been noted (see, e.g. [SSCB14] and [Str14c]) that the mathematical form of the PGFL of the Bayes posterior process – before approximation to close the Bayesian

recursion – is the product of a PPP and m Bernoulli target processes, where m denotes the number of measurements. In other words, the Bayes posterior process is the superposition of m Bernoulli processes and a PPP. This form is closely related to that of the multi–Bernoulli filters discussed below in Section 3.4.5.
In [CC09] the connection between a Gaussian Mixture implementation of the PHD intensity filter and the JIPDA filter is investigated. It is shown that under certain conditions (each target has a linear Gaussian dynamical model, target survival and detection probabilities are state independent, no explicit target birth and spawning event) the *composite density* (of the JIPDA) conforms to the definition of probability hypothesis density (see [CC09, Section 4]).

3.4.3 CPHD Intensity Filter

The CPHD intensity filter is proposed in [Mah07a], [Mah07b]. It propagates besides the intensity additionally the cardinality distribution and its PGFL. The assumptions made are essentially the same as in the PHD intensity filter. The differences are that the CPHD intensity filter propagates the cardinality distribution (2.67), that the clutter process is given by an i.i.d. cluster process as defined in (3.17) and that the PGF of the number of present targets is also given by an i.i.d. cluster process, that is

$$G_N^{\text{Cluster}}(z) = \sum_{n=0}^{\infty} p_N^{\text{Cluster}}(n) z^n, \tag{3.60}$$

where $p_N^{\text{Cluster}}(\cdot)$ is the distribution of the number of targets in X. The joint PGFL of the CPHD intensity filter is

$$\Psi_{\text{CPHD}}(g,h) = \Psi_C^{\text{Cluster}}(g)\, G_N^{\text{Cluster}}(\Psi_{\text{BMD}}(g,h)). \tag{3.61}$$

The PGFL of the CPHD intensity filter reduces to the PGFL of the PHD intensity filter if the target and measurement processes are both given by PPPs.
In [MKV12] a CPHD filter is proposed that uses a fixed number of targets. Note that if this CPHD filter employs a clutter process that is given by a PPP it is equivalent to the JPDAS filter proposed in (3.53).
The PGF $F^{\Xi|\Upsilon}(z)$ and the intensity $f^{\Xi|\Upsilon}(s)$ of the Bayes posterior process $\Xi|\Upsilon$ are defined as in (2.66) and (2.65) with respect to the joint PGFL $\Psi_{\text{CPHD}}(g,h)$. To close the Bayesian recursion, the posterior process is approximated by a point process $\widehat{\Xi|\Upsilon}$. The PGF of the number of targets in the approximating process is taken equal to that of the original process:

$$F^{\widehat{\Xi|\Upsilon}}(z) \equiv F^{\Xi|\Upsilon}(z).$$

The probability distribution of $\widehat{\Xi|\Upsilon}$ conditioned on n targets is defined by the product approximation

$$p_n^{\widehat{\Xi|\Upsilon}}(s_1,...,s_n) = \prod_{i=1}^{n} p^{\widehat{\Xi|\Upsilon}}(s_i), \tag{3.62}$$

where the PDF for a single target is defined by

$$p^{\widehat{\Xi|\Upsilon}}(s) = f^{\Xi|\Upsilon}(s) \Big/ \frac{dF^{\Xi|\Upsilon}}{dx}(1). \tag{3.63}$$

In words, the approximate process $\widehat{\Xi|\Upsilon}$ is a cluster process whose PGF of target number is chosen to be equal to that of the posterior process, and whose points are i.i.d. distributed with PDF proportional to the normalized intensity of the posterior process.

3.4.4 Generalized PHD Intensity Filters

The generalized PHD intensity filter was first proposed in [CM12] and shares most of the assumptions of the standard PHD intensity filter presented in Section 3.4.2. The predicted target process is a PPP and the updated target process is approximated by a PPP due to the same arguments as in Section 3.4.2. The differences are the target–originated measurement process and the clutter model. In Section 3.4.2 it is assumed that one target generates at most one measurement per sensor scan. Furthermore, it is assumed that the clutter model is given by (3.16), that is clutter is Poisson–distributed. The generalized PHD intensity filter relaxes these assumptions. First, the clutter model can be chosen arbitrary. Therefore, let $\Psi_C^{\text{gen}}(g)$ be the PGFL of a specified, but arbitrary, clutter process. Furthermore, let $\Psi_{\text{BMD}}^{\text{gen}}(h, g)$ be the PGFL of the target–oriented measurement process defined in Section 3.2.2.3. The PGFL of the generalized PHD intensity filter is therefore given by

$$\Psi_{\text{GenPHD}}(h, g) = \Psi_C^{\text{gen}}(g)\, G_N\big(\Psi_{\text{BMD}}^{\text{gen}}(h, g)\big), \tag{3.64}$$

which was proposed in [CM12].
In [Deg14] the generalization of the target–oriented measurement process is used to formulate the generalized iFilter. Analogously to the standard iFilter its generalized version uses an augmented state space $X^+ = X \cup X_\phi$, where X is equal to the state space of the generalized PHD filter and X_ϕ is the hypothesis space, used to model clutter by scatterers. A close numerical inspection on these two closely related filters is presented in Section 6.3.

3.4.5 Multi–Bernoulli Intensity Filters

A Bernoulli RV (trial) has two outcomes, or events, that are usually labeled “success” and “failure”. Its PGF is $G^{\text{Ber}}(z) \equiv 1 - q + qz,\ z \in \mathbb{C}$, where $q \in [0, 1]$ is the probability of success. A multi–Bernoulli RV is defined to be the number of successes in n mutually independent Bernoulli trials. If $q_i \in [0, 1]$ is the probability of success

in the ith trial, then the PGF of the number of successes is the product

$$G^{\text{multiBer}}(z) \equiv \prod_{i=1}^{n} (1 - q_i + q_i z)\,. \tag{3.65}$$

Bernoulli models were used in Section 3.2 for target detection modeling. They are used here to model target existence.

The multi–Bernoulli intensity filter proposed in [Mah07b] makes the same assumptions as the PHD intensity filter, except that the prior target process is assumed to be given by a multi–Bernoulli process. Enhanced versions, implementations and numerical examples of the multi–Bernoulli intensity filter can be found in [VVC09], [VVHM13]. The predicted target process is again a multi–Bernoulli process. This is guaranteed by assuming the birth process to be a multi–Bernoulli process, which is independent of the target process. Target death is modeled by independent probabilistic thinning. Targets are assumed to be independent Markovian processes having the same motion model $p_0(x|x_0)$ described in Section 3.2.1.

The Bayes posterior process is not a multi–Bernoulli process. Thus, it is approximated by a multi–Bernoulli process consisting of a superposition of two independent multi–Bernoulli processes, one modeling data–induced targets, the other modeling existing targets. This assures the Bayesian recursion to be closed.

The predicted target process is a multi–Bernoulli process with expected number of targets $\overline{N} \equiv E(N)$. Let $n \equiv \lfloor \overline{N} \rfloor$ be the largest integer less than or equal to $\overline{N}$. The PGFL of the multi–Bernoulli target process is given by

$$\Psi_{N\text{MB}}(h, g) = \prod_{i=1}^{n} \big(1 - \chi_i + \chi_i\, \Psi_{\text{BMD}(i)}(h, g)\big)\,. \tag{3.66}$$

Here, $\chi_i \in [0, 1]$ denotes probability that the ith predicted (hypothesized) target is indeed a target, i.e., that it exists. Analogously to (3.50), the PGFL for data–induced targets is given by

$$\Psi_{M\text{MB}}(h, g) = \prod_{j=1}^{m} \Big(1 - \gamma_j + \gamma_j\, \Psi^{\text{Data}}_{\text{BMD}(j)}(h, g)\Big)\,. \tag{3.67}$$

In general a cluster process as defined in (3.17) is proposed here to model clutter. Then, due to the superposition of the target–originated measurement process with the independent clutter process, the PGFL of the multi–Bernoulli filter is

$$\Psi^{\text{Cluster}}_{\text{MB}}(h, g) = \Psi^{\text{Cluster}}_{C}(g)\, \Psi_{N\text{MB}}(h, g)\, \Psi_{M\text{MB}}(h, g). \tag{3.68}$$

Substituting a PPP clutter model for the more general cluster process model in (3.68) gives

$$\Psi_{\text{MB}}(h, g) = \Psi^{\text{PPP}}_{C}(g)\, \Psi_{N\text{MB}}(h, g)\, \Psi_{M\text{MB}}(h, g). \tag{3.69}$$

This is the PGFL of the multi–Bernoulli intensity filter given in [Mah07b].
Comparing the joint PGFL of the data–driven JIDPA filter from (3.50) with the PGFL of the multi–Bernoulli filter in (3.69), it is evident that the PGFLs of both filters differ only in the application of superposition. The data–driven JIPDA has as many target state spaces as there are targets and measurements. In contrast, the targets within the multi–Bernoulli filter all share the same target state space. Therefore, the multi–Bernoulli filter can be described as a data–driven JIDPA filter which employs superposition. The multi–Bernoulli filter is also described as a superposed single–scan MHT algorithm in Section 3.3.7.
The connections between the multi–Bernoulli, the JIPDA, and the MHT filters are studied in terms of RFSs in [Wil14]. The multi–Bernoulli filter derived in [Wil15] is closely related to the set JPDA (SJPDA) [SSGW11b] filter.
Various extensions of the multi–Bernoulli filter are proposed [RVVD14] including labeled versions to keep account of the target identity. The labeled multi–Bernoulli process is described in [RVVD14] as an RFS on the Cartesian product of the state and label spaces, but this is inaccurate, for then the labels would be random, which they are not. As can be seen from Table 3.1 and the discussions above, the labeled version of the multi–Bernoulli filter corresponds to the JIPDA filter since the incorporation of target labels is equivalent to not superposing the targets onto one state space.
Closely related to the multi–Bernoulli filter is the processing of the joint multitarget probability density (JMPD) proposed in [KMKH05]. There, a particle filter is used to implement the JPDA.

3.5 Hybrid Pointillist Filters

The filters presented in this section assume at most $n \geq 1$ target groups to be present. Therefore, each target group has its own state space and, within these groups, targets are superposed.
It is also sometimes possible to superpose some targets and not others. Selective superposition has its uses. For example, targets that are well–separated could be superposed in one target state space, thereby possibly reducing computational complexity without incurring significant information loss, while the remaining targets are not superposed. Other examples involve highly mixed scenarios in which some groups of targets may have the same "within–group" dynamical model (e.g., constant velocity), but with different groups having different dynamical models. Selective superposition procedures are an additional decision step, which is outside the scope of this thesis.

3.5.1 Joint PHD Intensity Filter

The joint PHD intensity filter [Str14d] assumes that exactly $n \geq 1$ target groups are present. The groups are mutually independent. Each group has its own state

space X_i, $i = 1, ..., n$, and the targets in a group generate at most one measurement. Analogously to the PHD intensity filter (see Section 3.4.2) the target groups constitute a locally finite point process at the recursion start time. The PGFL of the updated process is not necessarily a PPP and thus it is approximated by a PPP for closing the Bayesian recursion. The predicted target process is a PPP, since independent probabilistic thinning (target death) is assumed per target group and independent Markovian targets of one target group are assumed to have the same motion model. The clutter process is assumed to be a PPP. Furthermore, the processes of the target groups and clutter are assumed to be mutually independent and hence the PGFL of the joint PHD intensity filter is given by

$$\Psi_{\text{JointPHD}}(h_1, ..., h_n, g) = \Psi_C^{\text{PPP}}(g) \prod_{i=1}^{n} G_{N_i}^{\text{PPP}}\big(\Psi_{\text{BMD}(i)}(h_i, g)\big), \tag{3.70}$$

where $G_{N_i}^{\text{PPP}}(z) = e^{-\overline{N}_i + \overline{N}_i z}$ is the PGF of the target number of the ith target group and $\overline{N}_i$ is the expected number of targets of the predicted target process.

3.5.2 Joint Generalized PHD Intensity Filter

Analogously to the joint PHD intensity filter, the joint generalized PHD filter assumes exactly $n \geq 1$ target groups to be present and these groups are mutually independent. As in the joint PHD intensity filter each group has its own state space X_i, $i = 1, ..., n$. Targets in a group are allowed to possess an arbitrary target–oriented measurement process, that is targets in a group can generate more than one measurement per sensor scan. To the knowledge of the author this filter has not been published yet. The target groups constitute a PPP at the recursion start time. Since the Bayes posterior process is not necessarily a PPP it is approximated by a PPP to close the Bayesian recursion. Due to the same arguments as in Section 3.5.1 the predicted target process is a PPP.
As in Section 3.4.4 the clutter model can be chosen arbitrarily. Thus, the joint PGFL of the joint generalized PHD intensity filter is given by

$$\Psi_{\text{JointGenPHD}}(h_1, ..., h_n, g) = \Psi_C^{\text{gen}}(g) \prod_{i=1}^{n} G_{N_i}^{\text{PPP}}\big(\Psi_{\text{BMD}(i)}^{\text{gen}}(h_i, g)\big), \tag{3.71}$$

where $\Psi_C^{\text{gen}}(g)$ is the PGFL of the arbitrarily specified clutter model used in (3.64) of Section 3.4.4.

3.6 Closing the Bayesian Recursion

As already discussed in Chapter 2, the PGFL of the Bayes posterior point process can be expressed by a ratio of functional derivatives of the joint PGFL (see

also [Mah07b], [Str13a], [CM12]). By definition this is valid for all pointillist filters. The analogous derivative ratio for the PGF of the Bayes conditional distribution has long been known for discrete multivariate distributions, [JKB97, Equ. (34.48), p. 11]. A Bayesian recursion is said to be closed, or exact, if the probability distributions of the prior and posterior processes have the same mathematical forms. (There is a class of filters, outside the scope of this thesis, that is closed in this sense [Dau87], [Dau05].) The pointillist filters studied in this thesis, however, possess prior and posterior distributions with different forms. Therefore, the posterior point process has to be approximated by a process with the same form as the prior to close the Bayesian recursion.

Methods for closing the Bayes recursion for pointillist filters that superpose targets have been discussed above. The resist on using the summary statistics of the Bayes posterior point process.In Section 3.4.1, the approximation of the Bayes posterior process by a simpler point process, e.g., a PPP for closing the Bayesian recursion is discussed and it is pointed out that this approximation implies a problem. The problem is that even if the number of local maxima of the PDF is equal to the number of targets, the target samples can *with nonzero probability* all be drawn from the same local maxima, that is, they are drawn with replacement. Instead, the samples should be drawn *without* replacement so that each target is only sampled once. Therefore, an approximation of the Bayes posterior process via PPPs means a mis–modeling.

In theory it is easier for non–superposed filters to close the Bayesian recursion since each target possesses its own state space. When all targets are assumed to exist, it can be shown from the PGFL that the intensity function (or PDF, if the PGFL is multilinear) is defined over the Cartesian product of target state spaces. Since the high dimensionality makes it impractical to carry an estimate of the joint intensity, a statistical approach to reduce the dimensionality of multivariate problems is to approximate the joint distributions by a product of the n univariate marginal distributions [Gri84], [JKB97]. The product form is reasonable given that targets are mutually independent. It is used for deriving the JPDA filter [BSF88], [BSL95]. The Bayes recursion is closed because the prior is assumed to have the same product form. Non–superposed pointillist filters imply other issues when, as for example in JIPDA, there is at most one target in each target state space, that is, if targets have a probability of existence smaller than one. Then a marginalization over all targets except for the target of interest and conditioning afterwards on the existence of the target enables to find a conditional target state PDF together with the probability of target existence [ME04]. Iterating over all targets for each iteration yields a discrete–continuous marginal distribution for each target. Using the product approximation for the joint distribution closes the Bayes recursion analogously to the JPDA–case.

3.7 Target State Estimation

The minimum Bayes risk estimate is defined for locally finite point processes as the realization of the process that minimizes the expected value of a specified cost function [Str13a]. The minimum Bayes risk estimate yields an estimate of the number of targets and is for many problems equivalent to a maximum *a posteriori* (MAP) estimate for an appropriately defined cost function.

Since the Bayes estimate is impractical, pseudo–Map estimates are found using the summary statistics of the Bayes posterior process. The estimated number of targets, $\hat{n}_{\text{pseudoMAP}}$, is the maximum of the (discrete) posterior distribution of the number of targets, as determined from the PGFL (see (2.67)–(2.66)) and can be obtained from the intensity function of the Bayes posterior process for each iteration. Pseudo–MAP estimates are reasonable estimates if they correspond to stable local peaks of the intensity function. However, the local peaks are often unstable [Mah07b], [EWBS05] and if more than one target falls into one intensity peak, it is not clear how to extract their target states. The problem is, that the intensity does not carry the complete statistical information. Enhanced target state extraction methods are derived to overcome these limitations [RCV10], [LqDfC10].

For tracking filters which do not superpose targets the state extraction is theoretically easy. There, the issue of a large number of targets can be handled by approximating the joint posterior distribution via a product of the marginal distributions.

3.8 How to Design a Tracking Filter: An Engineer's Perspective

The solution of a practical tracking problem using locally finite point processes can be separated into two methodologically different steps. First, the PGFL is constructed using different application–specific ingredients, such as the target, measurement and clutter model, complete, partial or no superposition of target states, the target–oriented measurement model, the sensor resolution model, etc. This process is called the *Discovery Step* and its component steps are depicted in Figure 3.1. Up to this point, this chapter has studied many different filters using this basic procedure. It can be seen from Figure 3.1 that several combinations with various models for the ingredients of a tracking filter are possible, and this leads to an enormous number of different possible tracking filters. The enthusiastic tracking engineer can use Figure 3.1 as an inspiration to design new customized tracking filters that fit the problem of current interest.

If the tracking problem is modeled by constructing the PGFL, the second step is to derive the formulas needed for the implementation of the particular tracking filter, i.e., to compute the summary statistics (first or higher order moments, distribution of the number of targets, etc.) needed to close the Bayesian recursion. These statistics are given by ratios of functional derivatives with respect to the Dirac delta, evaluated at different points of the state and measurement spaces [Mah07b], [CM12], [Str13a]

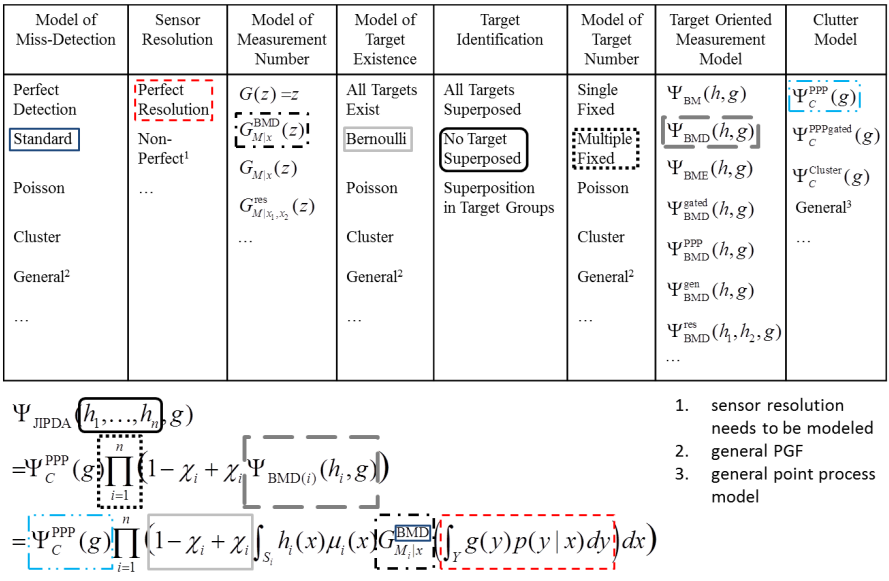

Figure 3.1: *Discovery Step.* A tracking filter is completely characterized by its joint target–measurement PGFL. This figure depicts the palette of available point process models for targets and measurement used in the PGFLs of the filters studied in sections up to and including Section 3.8. It highlights the construction of the PGFL for the JIPDA filter. Ellipses marks indicate that the list is not exhaustive.

(see also Chapter 2). Figure 3.2 visualizes this *Analytical Step* using the PHD filter [Mah07b], [Mah03], [Mah07a], the first filter to be derived from a PGFL.
Alternatives to functional differentiation are discussed in Section 4.3. Several of the methods presented there make it possible to assess particle filter performance on simulated data sets quickly and reliably, that is, they provide a very cost effective way to explore the design space without the need for expensive hand–crafted code.
The framework of PGFLs and point processes used in the previous sections is of more than purely theoretical interest. First, it shows similarities and differences between existing well–known tracking filters. For example, the consideration of how the superposition of targets is modeled within the framework of point processes brings to light the close connection of the multi–Bernoulli, the un–pruned/merged MHT and the JIPDA filter. Furthermore, it enables the reader to understand better the kinds of challenges that arise if new assumptions are made (e.g., on target superposition, Section 3.7) or if old assumptions are altered–seemingly small changes can have disproportionate impact. The proposed design procedure therefore helps the experienced tracking engineer to understand existing filters and their connections.
Second, the framework gives birth to an entirely new class (to the knowledge of the author) of tracking filters, called hybrid pointillist filters, by a straightforward application of the assumption that target states are superposed only within a specific number of target groups. It is evident to ask for detailed numerical evaluations of this new class of filters; however, a close investigation of practical applications is not part of this work and will be presented in future publications.
In addition to the important discoveries mentioned above, a tracking engineer might still ask: What's in it for me? The answer is: Unifying tracking filters in a common framework offers the possibility of customized designs for application–specific tracking problems that engineers are confronted with in practical work. Once understood it offers an easy way of finding out which existing methods can be re–used and which parts of the problem have to be modeled in a different way. Thereby, the ingredients such as clutter model, target–oriented measurement model, etc. of the existing tracking filters can be mixed in a large variety of combinations, as depicted in Figure 3.1. In the following, a guide to how to design a customized tracking filter is presented for a specific tracking scenario by making use of the essential cornerstone–parameters of the problem.

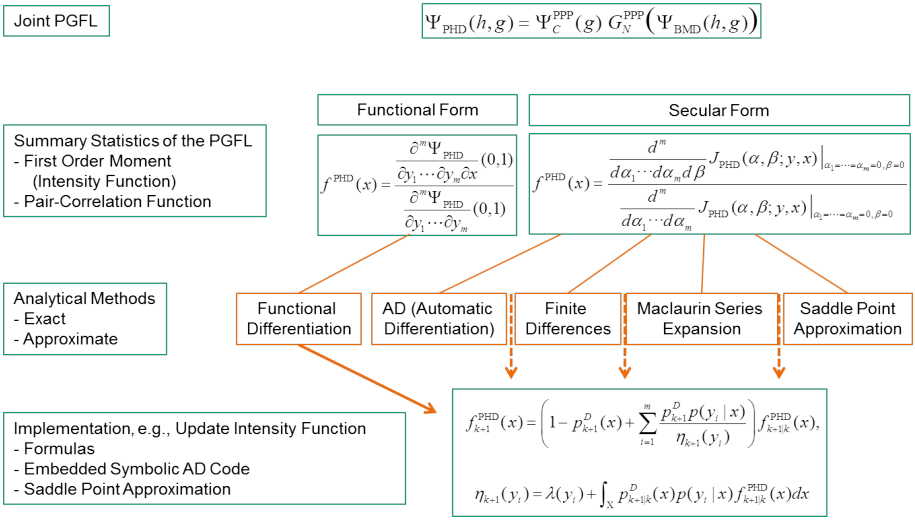

Figure 3.2: *Analytical Step.* Derive summary statistics of the respective tracking filter by differentiating the PGFL of the Bayes posterior point process. The intensity function f in the general case is given by (2.65). This figure depicts the variety of choices available for any PGFL, while indicating those made for the PHD intensity filter. Symbolic functional derivatives of the PGFL must be done by hand, but lead to explicit formulas. The secular form of the PGFL (see (4.80)) can be differentiated symbolically using widely available software. Exact numerical values for particle weights can be found by AD. Derivatives of all orders of the secular PGFL can be written using the Cauchy integral method, which lends itself to saddle point approximation. For the details of secular functions and exact and approximate analytical methods see Chapter 4.

Example Assume exactly two unresolved targets $x_1 \in X_1$, $x_2 \in X_2$ with probability of detection $p_{D,1}(x_1, x_2)$ and $p_{D,2}(x_1, x_2)$, respectively, to be present in a cluttered environment. The target's identity is of particular interest and one target generates at most one measurement. This problem has already been formulated in [SUD10]) (see also [BB06], [CBS84]) and should demonstrate how a tracking filter is formulated in terms of its joint PGFL. The following basic questions for designing a tracking filter are answered for each of the columns of Figure 3.1 for the unresolved target problem. By following these questions, the columns of Figure 3.1 provides a guide for formulating custom filters for other problems.

1. *What is the model of target missed detections?*
   Choose a standard missed detection model. For ease of discussion it is assumed that $p_{D,j}(x_1, x_2) \equiv p_{D,j}(x_j)$, $j = 1, 2$, that is, the probability of detecting target j depends only on target j. Adopting the notation (3.7) gives $a_j(x_j) =$

$1 - p_{D,j}(x_j)$ and $b_j(x_j) = p_{D,j}(x_j)$, $j = 1, 2$.

2. *How is sensor resolution modeled?*
   The resolution of the targets depends on how close they appear in the sensor. For example, the sensor resolution function can be Gaussian, i.e.,

$$f^{\text{res}}(x_1,x_2) = e^{-\left(H_1(x_1)-H_2(x_2)\right)^T \Sigma^{-1}\left(H_1(x_1)-H_2(x_2)\right)/2}, \tag{3.72}$$

   where the matrix Σ is determined by the nature of the sensor and the system functions $H_1 : X_1 \to Y$, $H_2 : X_2 \to Y$ map the target states to points in the measurement space. For this choice, targets at $x_1 \in X_2, x_2 \in X_2$ are poorly resolved by the sensor if $f^{\text{res}}(x_1, x_2) \approx 1$ and well resolved if $f^{\text{res}}(x_1, x_2) \approx 0$. Whether or not unresolved target states are close depends on the system functions H_1 and H_2. The Gaussian function f^{res} can be replace by any reasonable function $f^{\text{res}} : X_1 \times X_2 \to [0, 1]$ provided that $f^{\text{res}}(x_1, x_2) = 0$ for all $x_1 \in X_1$, $x_2 \in X_2$ with $H_1(x_1) = H_2(x_2)$. The resolution function, however defined, is used in the PGF of the number of measurements.

3. *How to model the number of measurements?*
   Let the PGF of the number of measurements be given by

$$G^{\text{res}}_{M|x_1,x_2}(z) \equiv c_0 + c_1 z + c_2 z^2, \tag{3.73}$$

   where

$$c_0 \equiv a_1(x_1)a_2(x_2), \tag{3.74}$$

$$c_1 \equiv a_1(x_1)b_2(x_2) + b_1(x_1)a_2(x_2) + f^{\text{res}}(x_1, x_2)b_1(x_1)b_2(x_2), \tag{3.75}$$

$$c2 \equiv b_1(x_1)b_2(x_2)\left(1 - f^{\text{res}}(x_1, x_2)\right) \tag{3.76}$$

   and $f^{\text{res}}(\cdot, \cdot)$ is defined, for example, by (3.72). The number of target measurements depends on the distance between the two targets in the sensor space. If $x_1 \approx x_2$, then $f^{\text{res}}(x_1, x_2) \approx 1$ and the PGF of the number of targets is

$$G^{\text{res}}_{M|x_1,x_2}(z) \approx c_0 + c_1 z, \quad z \in \mathbb{C}, \tag{3.77}$$

   which means that two poorly resolved targets yield at most one measurement with high probability. On the other hand, the PGF for well resolved targets is quadratic and, moreover, factors into two linear terms:

$$\begin{aligned} G^{\text{res}}_{M|x_1,x_2}(z) &= c_0 + c_1 z + c_2 z^2 \approx \left(a_1(x_1) + b_1(x_1)z\right)\left(a_2(x_2) + b_2(x_2)z\right) \\ &= G^{\text{BMD}}_{M|x_1}(z)\, G^{\text{BMD}}_{M|x_2}(z). \end{aligned} \tag{3.78}$$

   Thus, well resolved detected targets yield two conditionally independent measurements, as expected.

4. *How is target existence model defined?*
   Two targets are assumed present, exactly as in JPDA, so no modeling of the target existence is needed.

5. *Is target identification needed?*
   Yes. Therefore the joint PGFL will have two test functions, one for each target.

6. *How is the target number modeled?*
   A fixed number of targets is assumed.

7. *What is the target–oriented measurement model?*
   It is assumed that a target generates at most one measurement per sensor scan. (This assumption can be relaxed; see (3.20) and the extended version (3.13).) Measurements, however, depend jointly on the states of both targets. The target–oriented measurement model can be shown to be given by a modification of (3.20). It is defined by

$$\Psi^{\text{res}}_{\text{BMD}}(h_1, h_2, g) \equiv \int_{X_1}\int_{X_2} h_1(x_1)\mu_1(x_1)\, h_2(x_2)\mu_2(x_2) \times G^{\text{res}}_{M|x_1,x_2}\left(\int_Y g(y)p(y|x_1,x_2)\,dy\right) dx_1 dx_2, \tag{3.79}$$

   where $p(y|x_1, x_2)$ is an arbitrarily specified likelihood function that is conditioned on both targets. It is worth pointing out that the PGFL (3.79) reduces to the PGFL of the standard JPDA filter for two targets under two assumptions: (*i*) the factorization (3.78) holds, that is, the targets are always well–separated; and (*ii*) the likelihood function $p(y|x_1, x_2)$ depends on only one target state not two. In this case $p(y|x_i)$ is used within $G^{\text{BMD}}_{M|x_i}(\cdot)$, $i = 1, 2$.

8. *Which clutter model is used?*
   In this case the clutter model is an arbitrary locally finite point process. We choose the Poisson clutter model $\Psi^{\text{PPP}}_C(g)$ and assume further that the clutter process is mutually independent of both target–oriented measurement processes.

The target–oriented measurement and clutter processes are mutually independent, so the joint PGFL is given by the product of their PGFLs:

$$\Psi^{\text{res}}_{\text{JPDA}}(h_1, h_2, g) = \Psi^{\text{PPP}}_C(g)\, \Psi^{\text{res}}_{\text{BMD}}(h_1, h_2, g). \tag{3.80}$$

This PGFL fully characterizes the two target example problem for unresolved measurements. This completes the tracking filter *Discovery Step*.
Sensor resolution issues increase the complexity of the tracking problem. In the language of PGFLs, this can be traced to the fact that the double integral involved in the target–oriented measurement model $\Psi^{\text{res}}_{\text{BMD}}$ does not factor into a product of two

integrals as in the perfectly resolved case, since the PGF of the number of measurements is not the product of first degree polynomials (cf. (3.79). This has a significant impact in practice because computing double integrals is computationally much more demanding than computing the product of two lower dimensional integrals.
Secular functions can be defined and derived for PGFLs with general target–oriented measurement models. This can be seen directly for the unresolved target example using a method that is similar to that used in [Str14e]. The proof of the result in the general case is more involved and is presented in Chapter 4.
Another example of multiple integrals that do not factor is given by $\Psi_{\text{GenPHD}}(h, g)$ of the generalized PHD intensity filter (3.64). The reason in this case is that the generalized PHD filter enables a target to create more than one measurement per sensor scan and thus the integrals with respect to the measurement space do not factor.
It is seen from the unresolved target example discussed above that in practice engineers can use the pointillist filter point of view to quickly design and characterize tracking filters for modeling specific problems of interest.

3.9 Conclusion and Future Work

In this chapter the *Discovery Step* for pointillist filters is proposed. Filters that are modeled using locally finite point processes can be characterized in terms of a single functional. This PGFL description has several advantages. First, as seen in Table 3.1, it makes it easy to compare different tracking filters in terms of their boundary conditions. Therefore, existing relations and differences can be identified. Second, it enables data fusion engineers to design filters for their tracking problems at hand.
The pointillist filters are separated into three classes by their use of target superposition. Many well-known filters are in the class of pointillist filters that *do not* use target superposition, including the classical Bayes-Markov filter as well as the PDA, JPDA, IPDA, JIPDA, PMHT, and MHT filters.
The second class of pointillist filters are those that have a single target state space, that is, filters that superpose targets. Examples are the PHD intensity, the CPHD intensity, the generalized PHD intensity, and the multi-Bernoulli filters. Some filters (multi-Bernoulli, CPHD with fixed number of targets) can be formulated and derived from superposed versions of the joint PGFLs of the non-superposed JPDA, JIPDA, and MHT tracking filters (if all target state spaces are identical).
Hybrid pointillist filters, superpose some of the targets and others not. This is a class in between the first two classes. Examples are the joint PHD intensity and the generalized joint PHD intensity filters.
The next step is the *Analytical Step*, which is needed to obtain the explicit formulas for the implementation of the designed tracking filters. In Chapter 4 it is shown how summary statistics such as the first-order moment (intensity function) and higher-

order factorial moments (for pair–correlation) of the Bayes posterior process can be derived from the joint target–measurement PGFL found in the *Discovery Step*.
The study of multi-scan and multi-sensor versions of pointillist filters as well as the extension of family of pointillist filters will be presented in future publications and is not part of this thesis.

Filter Name	**PGFL Notation**	**Joint Target-Measurement PGFL**
Bayes-Markov	$\Psi_{\mathrm{BM}}(h,g)$	$\int_X \int_Y h(x)g(y)\mu(x)p(y\|x)\,dy\,dx$
Missed Detections	$\Psi_{\mathrm{BMD}}(h,g)$	$\int_X h(x)\mu(x)\,G^{\mathrm{BMD}}_{M\|x}\big(\int_Y g(y)p(y\|x)\,dy\big)\,dx$
Missed Detections and Extended Target	$\Psi_{\mathrm{BME}}(h,g)$	$\int_X h(x)\mu(x)\,G_{M\|x}\big(\int_Y g(y)p(y\|x)\,dy\big)\,dx$
PMHT		
Without Superposition	$\Psi_{\mathrm{PMHT}}(h_1,\ldots,h_n,g)$	$\prod_{i=1}^n \Psi^{\mathrm{PPP}}_{\mathrm{BMD}(i)}(h_i,g)$
With Superposition	$\Psi_{\mathrm{PMHTS}}(h,g)$	$\prod_{i=1}^n \Psi^{\mathrm{PPP}}_{\mathrm{BMD}(i)}(h,g)$
PDA		
Without Gating	$\Psi^{\mathrm{noGate}}_{\mathrm{PDA}}(h,g)$	$\Psi_C^{\mathrm{PPP}}(g)\,\Psi_{\mathrm{BMD}}(h,g)$
With Gating	$\Psi_{\mathrm{PDA}}(h,g)$	$\Psi_C^{\mathrm{PPPgated}}(g)\,\Psi^{\mathrm{gated}}_{\mathrm{BMD}}(h,g)$
Extended Target	$\Psi_{\mathrm{PDAE}}(h,g)$	$\Psi_C^{\mathrm{PPP}}(g)\,\Psi_{\mathrm{BME}}(h,g)$
JPDA		
Without Superposition	$\Psi_{\mathrm{JPDA}}(h_1,\ldots,h_n,g)$	$\Psi_C^{\mathrm{PPP}}(g)\,\prod_{i=1}^n \Psi_{\mathrm{BMD}(i)}(h_i,g)$
With Superposition	$\Psi_{\mathrm{JPDAS}}(h,g)$	$\Psi_{\mathrm{JPDA}}(h,\ldots,h,g)$
IPDA	$\Psi_{\mathrm{IPDA}}(h,g)$	$\Psi_C^{\mathrm{PPP}}(g)\left(1-\chi+\chi\,\Psi_{\mathrm{BMD}}(h,g)\right)$
JIPDA		
Without Gating	$\Psi_{\mathrm{JIPDA}}(h_1,\ldots,h_n,g)$	$\Psi_C^{\mathrm{PPP}}(g)\,\prod_{i=1}^n\left(1-\chi_i+\chi_i\,\Psi_{\mathrm{BMD}(i)}(h_i,g)\right)$
With Gating	$\Psi^{\mathrm{gated}}_{\mathrm{JIPDA}}(h_1,\ldots,h_n,g)$	$\Psi_C^{\mathrm{PPPgated}}(g)\,\prod_{i=1}^n\left(1-\chi_i+\chi_i\,\Psi^{\mathrm{gated}}_{\mathrm{BMD}(i)}(h_i,g)\right)$
With Superposition	$\Psi_{\mathrm{JIPDAS}}(h,g)$	$\Psi_{\mathrm{JIPDA}}(h,\ldots,h,g)$
MHT^2	$\Psi_{\mathrm{MHT}}(h_1,\ldots,h_{n+m},g)$	$\Psi_C^{\mathrm{PPP}}(g)\,\prod_{i=1}^n\left(1-\chi_i+\chi_i\,\Psi_{\mathrm{BMD}(i)}(h_i,g)\right)$ $\times\prod_{j=1}^m\left(1-\gamma_j+\gamma_j\,\Psi^{\mathrm{Data}}_{\mathrm{BMD}(j)}(h_{n+j},g)\right)$
Multi-Bernoulli	$\Psi_{\mathrm{MB}}(h,g)$	$\Psi_C^{\mathrm{PPP}}(g)\,\prod_{i=1}^n\left(1-\chi_i+\chi_i\,\Psi_{\mathrm{BMD}(i)}(h,g)\right)$ $\times\prod_{j=1}^m\left(1-\gamma_j+\gamma_j\,\Psi^{\mathrm{Data}}_{\mathrm{BMD}(j)}(h,g)\right)$
PHD	$\Psi_{\mathrm{PHD}}(h,g)$	$\Psi_C^{\mathrm{PPP}}(g)\,G_N^{\mathrm{PPP}}(\Psi_{\mathrm{BMD}}(h,g))$
CPHD	$\Psi_{\mathrm{CPHD}}(h,g)$	$\Psi_C^{\mathrm{Cluster}}(g)\,G_N^{\mathrm{Cluster}}(\Psi_{\mathrm{BMD}}(h,g))$
Generalized PHD	$\Psi_{\mathrm{GenPHD}}(h,g)$	$\Psi_C^{\mathrm{gen}}(g)\,G_N(\Psi^{\mathrm{gen}}_{\mathrm{BMD}}(h,g))$
Joint PHD	$\Psi_{\mathrm{JointPHD}}(h_1,\ldots,h_n,g)$	$\Psi_C^{\mathrm{PPP}}(g)\,\prod_{i=1}^n G^{\mathrm{PPP}}_{N_i}(\Psi_{\mathrm{BMD}}(h_i,g))$
Joint Gen. PHD	$\Psi_{\mathrm{JointGenPHD}}(h_1,\ldots,h_n,g)$	$\Psi_C^{\mathrm{gen}}(g)\,\prod_{i=1}^n G^{\mathrm{PPP}}_{N_i}(\Psi^{\mathrm{gen}}_{\mathrm{BMD}}(h_i,g))$

Ref. Eqn.	Target Model	Clutter	Missed Det.	Super-pos.
(3.19)	= 1 pt. tgt.	No	No	No
(3.20)	= 1 pt. tgt.	No	Std.	No
(3.21)	= 1 ext. tgt.	No	Gen.[5]	No
(3.36)	= n ext. tgts.	No[1]	No[8]	No
(3.55)	= n ext. tgts.	No[1]	No[8]	Yes
(3.22)	= 1 pt. tgt.	PPP	Std.	No
(3.29)	= 1 pt. tgt.	PPP	Std.	No
(3.30)	= 1 ext. tgt.	PPP	Gen.[5]	No
(3.32)	= n pt. tgts.	PPP	Std.	No
(3.53)	= n pt. tgts.	PPP	Std.	Yes
(3.40)	$\leq$ 1 pt. tgt.	PPP	Std.	No
(3.42)	$\leq n$ pt. tgts.	PPP	Std.	No
(3.49)	$\leq n$ pt.. tgts.	PPP	Std.	No
(3.54)	$\leq n$ pt. tgts.	PPP	Std.	Yes
(3.50)	$\leq n + m$ pt. tgts.	PPP	Std.	No
(3.69)	$\leq n + m$ pt. tgts.	PPP	Std.	Yes
(3.58)	PPP	PPP	Std.	Yes
(3.61)	Cluster	Cluster	Std.	Yes
(3.64)	Gen.[7]	Gen.[6]	Gen.[4]	Yes
(3.70)	= n gps.	PPP	Std.[3]	Yes[3]
(3.71)	= n gps.	Gen.[6]	Gen.[3,4]	Yes[3]

Table 3.1: Overview of the pointillist family of multitarget tracking filters
1 Clutter is modeled as target with high variance 2 MHT of un-pruned and complete set of hypothesis 3 Within target groups 4 Arbitrary target-oriented measurement process 5 Extended target measurement model 6 General locally finite point process clutter model 7 General locally finite point process target model 8 Modeled via missing data assignment weight

CHAPTER IV

Factorial Moment Derivation of the Bayes Posterior Point Process

The solution of a tracking problem, modeled by a locally finite point process, can be divided into two steps (see Figures 3.1 and 3.2). First, the *Discovery Step*, which is studied intensively in Chapter 3 and [SDK15], models a tracking filter by constructing the PGFL of the joint target–measurement process using basic target tracking ingredients like the clutter model, the target–generated measurement model, the sensor resolution, etc. In Chapter 3 it is shown that many well–known tracking filters can be situated in this class of PGFL–derivable filters, called the family of pointillist filters. Second, the *Analytical Step*, derives the tracking filter of a PGFL via functional differentiation and closes the Bayesian recursion by filter specific approximations. This chapter studies the theoretical foundations of the *Analytical Step*. Most of the contributions of this chapter are published in [DSK15].

The Bayes posterior target point process is conditioned on the measurement process, and its PGFL is a ratio of functional derivatives of joint PGFLs [Mah07b], [Str13a], [CM12]. The required functional derivatives are Gâteaux derivatives [ED11, p. 406] evaluated with respect to Dirac deltas [Dir27] centered at the locations of the measurements and targets.

In [Str14e] a technique for the derivation of multitarget intensity filters using ordinary derivatives is proposed for the class of PGFLs that can be written as a function of linear functionals. The objections raised in [CM12] are incorrect due to the use of Dirac as a function and not, as accepted in practice, a distribution, which is approximated by a sequence of test–functions. Thus, these objections are readily addressed by taking the limit of a test–function sequence for the Dirac delta with respect to the $L^1(\mu;\mathbb{R}^d)$–norm, where $d > 0$ and μ is the d–dimensional Lebesgue measure. The results in [Str14e] are applicable to many target tracking problems, but they do not extend to problems with correlated measurements because, in this case, the joint

target–measurement PGFL cannot be written as a function of linear functionals. The main purpose of this chapter is to extend the results from [Str14e] to a larger class of functionals which encompasses the family of pointillist filters from Chapter 3. In particular, the general case of correlated measurements is covered. Further extensions might be possible, but the presented extension is very broad and would to cover almost all practical tracking problems. Another purpose is to compare the functional derivative defined in this chapter to the set derivative defined in [Mah07b, p.380-381], [GMN97].
The intention of the work presented is to show that the alternative approach for deriving tracking filters using PGFLs and point process theory proposed in [Str14e] is mathematically exact and can be extended in a mathematically correct way to a larger class of PGFLs that is sufficient for almost all practical tracking applications. This alternative method does not conflict with the methods proposed in [Mah07b], [CM12] therein.
The chapter is organized as follows. In Section 4.1 the definition of the Gâteaux derivative with respect to the Dirac delta from [Str14e] is discussed and extended to a larger class of functionals that encompasses the pointillist family of multitarget tracking filters presented in Chapter 3. This definition is compared to the set derivative defined in [Mah07b, p. 380–381]. In Section 4.2 the technique for deriving multitarget intensity filters using secular functions [Str14e] is described, and it is shown that due to the results in Section 4.1 the technique can be generalized from the class of analytic functions of linear functionals to a larger class of PGFLs. In Section 4.3 a review of methods for computing the ordinary derivatives of secular functions are presented. The conclusions are drawn in Section 4.4.
The theory of secular functions presented in Section 4.2.1 and 4.2.4 on the class of PGFLs that can be written as a function of linear functionals has been derived by Streit in [Str14e]. Furthermore, the application of the well–known techniques for computing the ordinary derivatives of secular functions presented in Section 4.3 are proposed by Streit in [Str15], [Str14a] and [Str14e].

Own publications on this subject: The studies on the mathematical correct definition of the functional derivative with respect to the Dirac delta presented in Section 4.1 and the extension of it to the general family of PGFL $\mathcal{P}_2$ made in Section 4.2 have been published in [DSK15] © 2015 IEEE.

4.1 The Functional Derivative with Respect to the Dirac Delta

4.1.1 Definition and Approximation of Dirac Delta

In [Wal74] Dirac delta is defined as a distribution (see [Str13b] and [FJ99] as alternative references), i.e., δ is a linear and continuous mapping $\delta : \vartheta \to \mathbb{R}$, where the space

of test–functions is given by $\vartheta \equiv C_0^\infty(\mathbb{R}^n)$ (infinitely often continuously–differentiable with compact support). The space of all such distributions is defined by

$$\vartheta' \equiv \{f : C_0^\infty \to \mathbb{R}\}. \tag{4.1}$$

However, other test–function spaces are possible and in the following, we denote by δ^c the functional which satisfies $(\delta^c, \phi) \equiv \phi(c)$, where $c \in \mathbb{R}^n$ and $\phi : \mathbb{R}^n \to \mathbb{R}$.
In [Wal74, pp. 38–39] it is shown that δ^c can be approximated on ϑ by any absolutely integrable function $f : \mathbb{R}^n \to \mathbb{R}$ with $f(x) = 0$ for all $x \notin \{x : \|x - c\| \leq 1\}$ and $\int_{\mathbb{R}^n} f(x)dx = 1$ by $f_\lambda(\cdot) \equiv \lambda^{-n} f\left(\frac{\cdot - c}{\lambda}\right)$ in the sense that $\lim\limits_{\lambda \searrow 0} f_\lambda = \delta^c$ in ϑ, where

$$\lim_{\lambda \searrow 0}(f_\lambda, \phi) \equiv \lim_{\lambda \searrow 0}\left(\int f_\lambda(x)\phi(x)\right) = \int \delta^c(x)\phi(x)dx = (\delta^c, \phi) = \phi(c), \tag{4.2}$$

for all $\phi \in \vartheta$. For the proof, the continuity of the test–functions is sufficient. Note that the limit is not taken into the integral, that is, we are investigating a limit of integrals also known as convergence with respect to the $L^1(\mu; \mathbb{R}^n)$–norm. In the following, we denote by $\{\gamma_\lambda^c\}_{\lambda>0}$ an arbitrary family of test–functions converging in the sense of (4.2) for $\lambda \searrow 0$ to δ^c.
Two families of test-functions are used exemplary in this work for δ^c on $\mathbb{R}^d$. For $x = (x_1, ..., x_d)$, $c = (c_1, ..., c_d) \in \mathbb{R}^d$ the (multivariate) Gaussian test–sequence is defined by

$$\gamma_\lambda^{c,1}(x) \equiv \prod_{i=1}^{d} \frac{1}{\sqrt{2\pi\lambda^2}} \exp\left(-\frac{(x_i - c_i)^2}{2\lambda^2}\right) \tag{4.3}$$

and the step function sequence is given by

$$\gamma_\lambda^{c,2}(x) \equiv \frac{1}{\mu(E_\lambda^c)} 1_{E_\lambda^c}, \tag{4.4}$$

where $E_\lambda^c \subset \mathbb{R}^d$ is a neighbourhood of $c \in \mathbb{R}^d$ with $\mu(E_\lambda^c) = \lambda$, μ being the Lebesgue measure on $\mathbb{R}^d$. Note that $\gamma_\lambda^{c,i} : \mathbb{R}^d \to \mathbb{R}$, $i = 1, 2$ are indeed approximate identities in the sense of [Alt12, p.114] for δ^c. Figure 4.1 and Figure 4.2 visualize $\gamma_\lambda^{c,1}$ and $\gamma_\lambda^{c,2}(x)$ for different values of the parameter λ.
Even though there exists no local integrable function satisfying

$$\int \delta^c(x)\phi(x)dx = (\delta^c, \phi) \equiv \phi(c) \tag{4.5}$$

for $\phi \in C_0^\infty(\mathbb{R}^n)$ [Wal74, p.21], we will use the integral notation from (4.5) in the following, keeping in mind that it is a notational device. However, the reader should note that the right hand side of (4.5), that is the definition of Dirac delta as an operator (or distribution in the sense of [Wal74]), is mathematically well–defined.

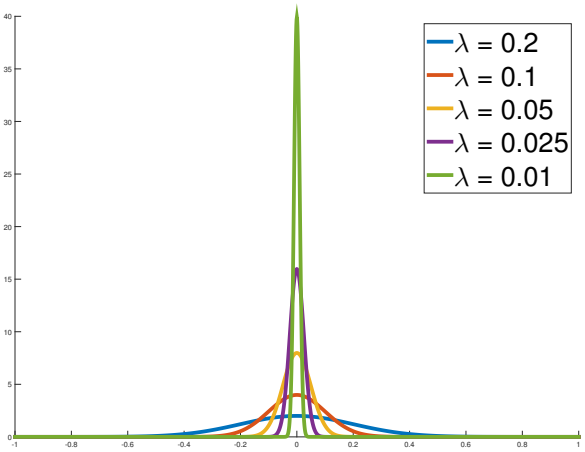

Figure 4.1: Plot of approximate identities $\gamma_\lambda^{0,1}$ for Dirac delta at 0, defined in (4.3) for different values of λ in one–dimension © 2015 IEEE.

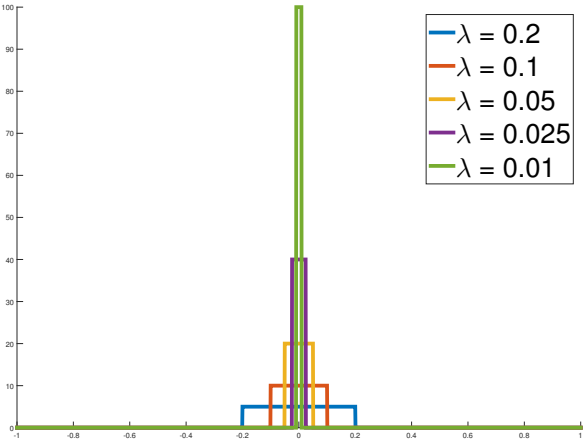

Figure 4.2: Plot of approximate identities $\gamma_\lambda^{0,2}$ for Dirac delta at 0, defined in (4.4) for different values of λ in one–dimension © 2015 IEEE.

4.1.2 Definition of the Gâteaux Derivative with respect to the Dirac Delta

In [Str14e] the functional derivative with respect to an impulse is defined. It can be called Gâteaux derivative with respect to the Dirac delta (and has been done so far in this thesis) to emphasize the fact that it is a directional derivative of a functional with respect to the Dirac delta.
First, the class of PGFLs, for which the Gâteaux derivative with respect to the Dirac delta is defined in [Str14e], is investigated. Let $F : \mathbb{C}^k \to \mathbb{R}$ be a multivariate PGF, that is

$$F(s_1, ..., s_k) \equiv \sum_{i_1,...,i_k \geq 0} p_{i_1,...,i_k} \, s_1^{i_1} \cdot ... \cdot s_k^{i_k}, \tag{4.6}$$

where $(s_1, .., s_k) \in \mathbb{C}^k$ and $p_{i_1,...,i_k} \in [0, 1]$. Let $Y \subset \mathbb{R}^d$ be closed and bounded, and let $q_i : Y \to \mathbb{R}$, $i = 1, ..., k$, be continuously differentiable PDFs except possibly for jump continuities. Furthermore, let

$$g \in \mathcal{G} \equiv \Big\{ g \in L^1(\mu; Y) : \ g \text{ is non–negative, } |g(y)| \leq 1, \text{ for all } y \in Y \}. \tag{4.7}$$

Then, we define the class of analytic functions of linear functionals by

$$\mathcal{P}_1 \equiv \left\{ \Psi : \mathcal{G} \to \mathbb{R} : \ \Psi(g) \equiv F\left(\int_Y g(y) q_1(y) dy, ..., \int_Y g(y) q_k(y) dy \right) \right\}. \tag{4.8}$$

In [Str14e], for the class of PGFLs $\mathcal{P}_1$ and an arbitrary family of test–functions $\{\gamma_\lambda^c\}_{\lambda>0}$ of δ^c, the Gâteaux derivative with respect to the Dirac delta is defined by

$$\frac{\partial \Psi}{\partial c}(g) \equiv \lim_{\lambda \searrow 0} \frac{\partial \Psi}{\partial \gamma_\lambda^c}(g), \tag{4.9}$$

where

$$\frac{\partial \Psi}{\partial g}(g') \equiv \lim_{\epsilon \searrow 0} \frac{\Psi(g + \epsilon g') - \Psi(g)}{\epsilon} \tag{4.10}$$

and $g, g' \in \mathcal{G}$. Since the multivariate PGF is analytic at the origin $\mathbf{0} = (\mathbf{0}, ..., \mathbf{0}) \in \mathbb{C}^{\mathbf{k}}$, the limit (4.9) is well–defined for PGFLs from $\mathcal{P}_1$. Note that the Gâteaux derivative $\frac{\partial \Psi}{\partial \gamma_\lambda^c}(g)$ is well–defined for all $\lambda > 0$.
The Fréchet derivative [Alt12] of a functional requires uniform convergence for all directions (test–functions). Every Fréchet differentiable functional is also Gâteaux differentiable. In general, the converse does not hold. However, for analytic functions Gâteaux differentiability is equivalent to Fréchet differentiability. In the following the Fréchet/Gâteaux derivative is called a functional derivative for simplicity (following the language of [Str14e]). The uniform convergence for all test–functions of the Fréchet derivative is needed for the definition of secular functions in Section 4.2. It is also needed if saddle point methods are employed to find point–wise approximations of the factorial moments of the Bayesian posterior point process [Str15].

4.1.3 Extending the Functional Derivative with respect to the Dirac Delta

This section extends the definition of the functional derivative with respect to the Dirac delta (4.9) to a larger class of PGFLs. Furthermore, it can be seen that the set derivative from [Mah07b] is a special case of the proposed functional derivative. First, we define the class of PGFLs we are interested in by

$$\mathcal{P}_2 \equiv \left\{ \Psi : \mathcal{H} \to \mathbb{R} : \Psi(h) = \sum_{n\geq 0} \frac{a_n}{n!} \int_{\mathcal{X}^n} \prod_{i=1}^{n} h(x_i) f_n(x_1, ..., x_n) dx_1 ... dx_n \right\}, \tag{4.11}$$

where

$$\mathcal{H} \equiv \{ h : \mathcal{X} \to \mathbb{R} : \ h \text{ is bounded by one, non–negative and Lebesgue–integrable} \}, \tag{4.12}$$

$\mathcal{X} = \mathbb{R}^d$, $d \geq 1$, $a_n \in [0,1]$, $\sum_{n\geq 0} a_n = 1$, $n \in \mathbb{N}$ and $f_n : \mathcal{X}^n \to \mathbb{R}$ be a symmetric bounded density, which may depend on $n \geq 1$.
Note that for the univariate case the class of analytic functions of linear functionals $\mathcal{P}_1$ is a subset of $\mathcal{P}_2$. Indeed, consider

$$\Psi(h) \equiv F\left(\int_X h(x) q(x) dx \right) \tag{4.13}$$

with $F(s) = \sum_{n\geq 0} a_n s^n$, $h \in \mathcal{H}$. Then, $\Psi \in \mathcal{P}_1$ and setting

$$f_n(x_1, ..., x_n) \equiv n! \cdot q(x_1) \cdot \ldots \cdot q(x_n), \tag{4.14}$$

$x_i \in \mathcal{X}$, $i = 1, ..., n$ shows that $\Psi \in \mathcal{P}_2$. Thus, $\mathcal{P}_1 \subset \mathcal{P}_2$ for the univariate case. The class of analytic functions of linear functionals $\mathcal{P}_1$ is also a proper subset of $\mathcal{P}_2$, since PGFLs with

$$f_n(x_1, ..., x_n) \equiv (x_1 + ... + x_n) \left(\sqrt{x_1^2 + ... + x_n^2} \right)^{-1}, \tag{4.15}$$

$x_i \in \mathcal{X}$, $i = 1, ..., n$ are in $\mathcal{P}_2$ but not in $\mathcal{P}_1$ for the univariate case. The univariate class of PGFLs $\mathcal{P}_2$ can be extended to the multivariate case by

$$\mathcal{P}_2^{\text{multi}} \equiv \left\{ \Psi : \mathcal{H} \to \mathbb{R} : \ \Psi(h) \equiv \sum_{j=1}^{k} \sum_{n\geq 0} \frac{a_{n,j}}{n!} \int_{\mathcal{X}^n} \prod_{i=1}^{n} h(x_i) f_{n,j}(x_1, ..., x_n) dx_1 ... dx_n \right\}, \tag{4.16}$$

where $a_{n,j} \in [0,1]$, $\sum_{j=1}^{k} \sum_{n\geq 0} a_{n,j} = 1$ and $\sum_{j=1}^{k} f_{n,j} \mathcal{X}^n \to \mathbb{R}$ is assumed to be a symmetric density, which depends on $n \in \mathbb{N}$. However, this class of PGFLs is not investigated further in this work.
Following the idea from definition (4.9) we define the functional derivative of a PGFL

from class $\mathcal{P}_2$ with respect to the Dirac delta δ^c and the two families of test–functions $\{\gamma_\lambda^{c,i}\}_{\lambda>0}$ for $i = 1, 2$ to be

$$\frac{\partial \Psi}{\partial c}(g) \equiv \lim_{\lambda \searrow 0} \frac{\partial \Psi}{\partial \gamma_\lambda^{c,i}}(g). \tag{4.17}$$

Note that the definitions (4.9) and (4.17) are closely related to the considerations on the constructive definition of the Radon–Nikodym derivative from [GMN97, p.145]. There $\{\gamma_\lambda^{c,2}\}_{\lambda>0}$ is used as an approximate identity and the convergence is proven by the Lebesgue density theorem.
The assumptions on f needed to justify the definition (4.17) are now given. To this end, we first consider for $i = 1, 2$ and $\lambda > 0$

$$\frac{\partial \Psi}{\partial \gamma_\lambda^{c,i}}(h) = \lim_{\epsilon \searrow 0} \frac{\Psi(h + \epsilon \gamma_\lambda^{c,i}) - \Psi(h)}{\epsilon} \tag{4.18}$$

$$= \sum_{n \geq 1} \frac{a_n}{n!} \sum_{k=1}^{n} \int_{\mathcal{X}^n} \gamma_\lambda^{c,i}(x_1) \prod_{i=1, i \neq k}^{n} h(x_i) f_n(x_1, ..., x_n) dx_1 ... dx_n, \tag{4.19}$$

$i = 1, 2$. Note that (4.19) holds due to the symmetry of f, and that the sum and the derivative can be interchanged, due to the analyticity of $\Psi(h + \epsilon h')$ in ϵ in some open region containing the origin of the complex plane [Str13a], a fact that is proven in [Moy62, section 4]. The outermost sum starts at $n = 1$, since the derivative of the $n = 0$ term is zero. Now, it has to be investigated under which assumptions

$$\frac{\partial \Psi}{\partial c}(g) \equiv \lim_{\lambda \searrow 0} \frac{\partial \Psi}{\partial \gamma_\lambda^{c,i}}(g) = \lim_{\lambda \searrow 0} \lim_{\epsilon \searrow 0} \frac{\Psi(h + \epsilon \gamma_\lambda^{c,i}) - \Psi(h)}{\epsilon}$$

$$= \sum_{n \geq 1} \frac{a_n}{n!} \sum_{k=1}^{n} \int_{\mathcal{X}^n} \delta^c(x_k) \prod_{i=1, i \neq k}^{n} h(x_i) f_n(x_1, ..., x_n) dx_1 ... dx_n \tag{4.20}$$

$$= \sum_{n \geq 1} \frac{a_n}{n - 1!} \int_{\mathcal{X}^{n-1}} \prod_{i=2}^{n} h(x_i) f_n(c, x_2, ..., x_n) dx_2 ... dx_n \tag{4.21}$$

holds, $i = 1, 2$. Equation (4.21) is due to applying Dirac delta in terms of a distribution, the symmetry of f and the assumption that c is in the domain of $f(\cdot, x_2, ..., x_n) : \mathbb{R} \to \mathbb{R}$, $x_2, ..., x_n \in \mathbb{R}^d$. The interchange of the limit and the outermost sum in (4.20) is not straightforward and has to be investigated carefully. To find out under which conditions on f_n definition (4.20) holds, we apply Lebesgue's dominated convergence theorem (LDC) [Alt12, p. 62] to the space of sequences

$$l^1(\mathbb{R}) \equiv \{x \in \mathbb{R}^{\mathbb{N}} : \sum_{i \in \mathbb{N}} |x_i| < \infty\} \tag{4.22}$$

[Alt12, pp. 28–29] in the following. LDC can be applied to $l^1(\mathbb{R})$, since it is equal to $L^1(m;\mathbb{R})$, where m is the counting measure, defined by

$$m(A) \equiv \begin{cases} |A|, \text{ if } |A| < \infty \\ \infty, \text{ if } |A| = \infty, \end{cases} \tag{4.23}$$

$A \in \mathcal{B}(\mathbb{N})$, where $\mathcal{B}(\mathbb{N})$ denotes the Borel–σ algebra of the natural numbers [Els09, p. 29] and

$$L^p(\mu, \mathbb{R}^n) \equiv \Big\{ f : (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mu) \to (\mathbb{R}, \mathcal{B}(\mathbb{R})) : \\ \text{f is } \mu\text{–measurable and } \left(\int_{\mathbb{R}^n} |f(x)|^p \right)^{\frac{1}{p}} < \infty \Big\}, \tag{4.24}$$

μ being an arbitrary measure defined on $\mathcal{B}(\mathbb{R}^n)$. Define $f_\lambda^{c,i} : \mathbb{N} \to \mathbb{R}$ by

$$f_\lambda^{c,i}(n) \equiv \frac{a_n}{n!} \sum_{k=1}^{n} \int_{\mathcal{X}^n} \gamma_\lambda^{c,i}(x_k) \prod_{i=1, i\neq k}^{n} h(x_i) f_n(x_1, ..., x_n) dx_1 ... dx_n, \tag{4.25}$$

for all $n \in \mathbb{N}$, $\lambda > 0$, $h \in \mathcal{H}$, $c \in \mathcal{X}$ and $i = 1, 2$.
To show that the limit (4.17) is well defined using LDC it suffices to show that first $f_\lambda^{c,i}$ converges for all $c \in \mathcal{X}$ m–almost everywhere (a.e). to

$$f^c(n) \equiv \frac{a_n}{n!} \sum_{k=1}^{n} \int_{\mathcal{X}^n} \delta^c(x_k) \prod_{i=1, i\neq k}^{n} h(x_i) f_n(x_1, ..., x_n) dx_1 ... dx_n \tag{4.26}$$

for $i = 1, 2$ and second

$$|f_\lambda^{c,i}| \leq g, \text{for some } g \in l^1(\mathbb{R}) \ m\text{–a.e.}, \tag{4.27}$$

since then due to LDC $f_\lambda^{c,i}$ converges to f^c in $l^1(\mathbb{R})$, that is (4.20) holds. Note that due to the definition of the counting measure m, a statement holds m–a.e. if it holds for all $n \in \mathbb{N}$.
The next Lemma proves Condition (4.26). From now on μ denotes the n–dimensional Lebesgue measure, defined on $\mathcal{X}^n$, $n \geq 1$.

Lemma 1. *Let $f_n \in L^1(\mu; \mathcal{X}^n)$ be continuous and bounded for all $n \in \mathbb{N}$. Denote by $\{\gamma_\lambda^0\}_{\lambda>0}$ an arbitrary approximate identity in the sense of [Alt12, p.114]. Then, the convolution $(\gamma_\lambda^0 * f_n)(x) \equiv \int_{\mathcal{X}} \gamma_\lambda^0(x-y) f_n(y) dy$ converges everywhere for $\lambda \searrow 0$ (for all $x \in \mathcal{X}$) to $f_n(x)$.*

Proof. Since $\{\gamma_\lambda^0\}_{\lambda>0}$ is an approximate identity and therefore integrates to one. Analogously to [Alt12, p.116] it holds that

$$|(\gamma_\lambda^0 * f_n)(x) - f_n(x)| = \left| \int_{\mathcal{X}} \gamma_\lambda^0(x-y)\left(f_n(y) - f_n(x)\right) dy \right| \tag{4.28}$$

$$= \left| \int_{\mathcal{X}} \gamma_\lambda^0(\tilde{y})\left(f_n(x-\tilde{y}) - f_n(x)\right) d\tilde{y} \right| \leq \int_{\mathcal{X}} \left| \gamma_\lambda^0(\tilde{y})\left(f_n(x-\tilde{y}) - f_n(x)\right) \right| d\tilde{y} \tag{4.29}$$

$$= \int_{B_r(0)} \left| \gamma_\lambda^0(\tilde{y})\left(f_n(x-\tilde{y}) - f_n(x)\right) \right| d\tilde{y} + \int_{\mathcal{X}\setminus B_r(0)} \left| \gamma_\lambda^0(\tilde{y})\left(f_n(x-\tilde{y}) - f_n(x)\right) \right| d\tilde{y} \tag{4.30}$$

$$\leq \sup_{y\in B_r(0)} \left| (f_n(x-y) - f_n(x)) \right| \cdot \int_{B_r(0)} \gamma_\lambda^0(\tilde{y}) d\tilde{y} \tag{4.31}$$

$$+ \sup_{y\in \mathcal{X}\setminus B_r(0)} \left| (f_n(x-y) - f_n(x)) \right| \cdot \int_{\mathcal{X}\setminus B_r(0)} \gamma_\lambda^0(\tilde{y}) d\tilde{y}, \tag{4.32}$$

where $B_r(0)$ is a ball with radius $r > 0$ around $0 \in \mathcal{X}$. The first summand (4.31) converges due to the continuity of f and the boundedness of $\int_{B_r(0)} \gamma_\lambda^0(\tilde{y}) d\tilde{y}$ for $r \searrow 0$ to zero. The second summand (4.32) converges for all $r > 0$ and $\lambda \searrow 0$ to zero due to the definition of an approximate identity and the boundedness of f_n. □

Remark 1. *Note that convergence is needed for all $x \in \mathcal{X}$, since Condition (4.26) has to hold for arbitrary $c \in \mathcal{X}$.*

Remark 2. *Lemma 1 holds for arbitrary approximate identity. For the particular case of the Dirac sequence $\{\gamma_\lambda^{c,2}\}$ it seems that the continuity of $f_n \in L^1(\mu; \mathcal{X}^n)$ is not needed. This due to the fact that Lebesgue's density theorem [SG97, p.220–222] can be applied. However, Lemma 1 states convergence not only μ–a.e., but convergence for all $x \in \mathcal{X}$. Hence, continuity of $f_n \in L^1(\mu; \mathcal{X}^n)$ is needed.*

Remark 3. *Due to [Alt12, Theorem 2.15] and [Alt12, A 1.11] it seems that Lemma 1 could have been formulated without the assumption that $f_n \in L^1(\mu; \mathcal{X}^n)$ is continuous, if the approximate identity is replaced by an appropriate subsequence. This is due to the fact that convergence with respect to the $L^1(\mu; \mathbb{R})$–norm implies μ–a.e. convergence for a subsequence [Alt12, A 1.11]. However, analogously to the previous Remark, Lemma 1 states converges not only μ–a.e., but convergence for all $x \in \mathcal{X}$.*

This proves Condition (4.26) for all Dirac sequences, that is especially for the two studied approximate identities $\{\gamma_\lambda^{0,i}\}, i = 1, 2$.

Since for $h \in \mathcal{H}$ it holds that $h \geq 0$ and bounded by some constant $B \in \mathbb{R}$ we obtain

$$|f_\lambda^{c,i}| \leq \frac{B^{n-1}}{n!} \sum_{k=1}^{n} \int_{\mathcal{X}^n} \gamma_\lambda^{c,i}(x_k) f_n(x_1, ..., x_n) dx_1 ... dx_n \tag{4.33}$$

$$= \frac{n \cdot B^{n-1}}{n!} \int_{\mathcal{X}^n} \gamma_\lambda^{c,i}(x_1) f_n(x_1, ..., x_n) dx_1 ... dx_n \tag{4.34}$$

$$= \frac{B^{n-1}}{n-1!} \int_{\mathcal{X}^n} \gamma_\lambda^{c,i}(x_1) f_n(x_1, ..., x_n) dx_1 ... dx_n, \tag{4.35}$$

where (4.35) holds due to the symmetry of $f_n(x_1, ..., x_n)$, $i = 1, 2$. If, it can be shown that

$$\int_{\mathcal{X}^n} \gamma_\lambda^{c,i}(x_1) f_n(x_1, ..., x_n) dx_1 ... dx_n \leq C^{n-1} \cdot A, \tag{4.36}$$

where $A,\ C \in \mathbb{R}$, then $g : \mathbb{N} \to \mathbb{R}$ defined by

$$g(n) \equiv \frac{(B \cdot C)^{n-1}}{n-1!} \cdot A, \tag{4.37}$$

$n \in \mathbb{N}$ is in $l^1(\mathbb{R})$, since

$$\sum_{n=1}^{\infty} |g(n)| = \sum_{n=1}^{\infty} \left| \frac{(B \cdot C)^{n-1}}{n-1!} \cdot A \right| = |A| \cdot \sum_{n=0}^{\infty} \left(\frac{\left(|B| \cdot |C|\right)^n}{n!} \right)$$

$$= |A| \cdot \exp(|B| \cdot |C|) < \infty \tag{4.38}$$

and $|f_\lambda^{c,i}| \leq g$ m–a.e, that is $|f_\lambda^{c,i}(n)| \leq g(n)$ for all $n \in \mathbb{N}$, $i = 1, 2$. If such a function g can be found, Condition (4.27) is fulfilled. Before the function g is constructed the following Lemma needs to be shown.

Lemma 2. *For $f_n : \mathcal{X}^n \to \mathbb{R}$ symmetric, $f_n \geq 0$ and $\int_{\mathcal{X}^n} f_n(x_1, ..., x_n) = K$, $K \in \mathbb{R}$ it holds that there exists a constant $C \in \mathbb{R}$ such that $\int_{\mathcal{X}} f_n(x_1, ..., x_n) dx_i = C < \infty$, for all $i \in \{1, ..., n\}$.*

Proof. First, due to the symmetry of f_n, it holds for all $i, j \in \{1, ..., n\}, i \neq j$ that

$$\int_{\mathcal{X}} f_n(x_1, ..., x_n) dx_i = \int_{\mathcal{X}} f_n(x_1, ..., x_n) dx_j. \tag{4.39}$$

Indeed, assume that there exist $i, j \in \{1, ..., n\}$ such that

$$\int_{\mathcal{X}} f_n(x_1, ..., x_n) dx_i \neq \int_{\mathcal{X}} f_n(x_1, ..., x_n) dx_j. \tag{4.40}$$

Consider the permutation $\sigma : \{1, ..., n\} \to \{1, ...n\}$ with $\sigma(i) = j$, $\sigma(j) = i$ and $\sigma(k) = k$, for all $k \in \{1, ..., n\}$, $k \neq i, j$. Then it holds

$$\int_{\mathcal{X}} f_n(x_1, ..., x_n)dx_i = \int_{\mathcal{X}} f_n(x_{\sigma(1)}, ..., x_{\sigma(n)})dx_i = \int_{\mathcal{X}} f_n(x_1, ..., x_n)dx_j, \tag{4.41}$$

which is a contradiction to the assumption.
Next, we show that $\int_{\mathcal{X}} f_n(x_1, ..., x_n)dx_i < \infty$. Assume that there exists a $j \in \{1, ..., n\}$ such that $\int_{\mathcal{X}} f_n(x_1, ..., x_n)dx_i = \infty$. Then,

$$K = \int_{\mathcal{X}^n} f_n(x_1, ...x_n)dx_1...dx_n = \int_{\mathcal{X}} ... \int_{\mathcal{X}} f_n(x_1, ...x_n)dx_i dx_1....dx_{i-1}dx_{i+1}...dx_n = \infty, \tag{4.42}$$

where (4.42) holds due to Fubini's theorem [Alt12]. This is a contradiction to the assumption and proves the statement. □

Theorem 3. *Let $f_n \in L^1(\mu; \mathcal{X}^n)$ be continuous, bounded and symmetric. Then for all $\lambda > 0$ $|f_\lambda^{c,i}| \leq g$, m–a.e. $i = 1, 2$, where $g : \mathbb{N} \to \mathbb{R}$ is defined by*

$$g(n) \equiv \frac{(B \cdot C)^{n-1}}{n - 1!} \cdot \sup_{x \in \mathcal{X}^n} (f_n(x_1, x_2, ..., x_n)), \tag{4.43}$$

$n \in \mathbb{N}$, $x = (x_1, ..., x_n) \in \mathcal{X}^n$, $B \in \mathbb{R}$ being the bound of the test-function h and $C = \int_{\mathcal{X}} |f_n(x_1, ..., x_n)|dx_i$. Furthermore, it holds that $g \in l^1(\mathbb{R})$.

Proof. Since f_n is bounded the Hölder–inequality [Alt12, p. 54] gives for $p = \infty$, $q = 1$ and $x_2, ..., x_n \in \mathcal{X}$

$$\left| \int_{\mathcal{X}} \gamma_\lambda^{c,i}(x_1) f_n(x_1, ..., x_n)dx_1 \right| \leq \left\| \gamma_\lambda^{c,i} f_n(\cdot, x_2, ..., x_n) \right\|_{L^1(\mu;\mathcal{X})} \tag{4.44}$$

$$\leq \sup_{x \in \mathcal{X}^n} (f_n(x_1, x_2, ..., x_n)) \int_{\mathcal{X}} \gamma_\lambda^{c,i}(x_1)dx_1 = \sup_{x \in \mathcal{X}^n} \left| f_n(x_1, x_2, ..., x_n) \right|, \tag{4.45}$$

$i = 1, 2$, which proves together with Lemma 2 and Fubini's theorem the first part of the statement. Furthermore,

$$\sum_{n=1}^{\infty} |g(n)| = \exp(|B| \cdot |C|) \cdot \sup_{x \in \mathcal{X}} (f_n(x_1, x_2, ..., x_n)) < \infty \tag{4.46}$$

and thus $g \in l^1(\mathbb{R})$. □

This proves together with Lemma 1 and Lemma 2, that if $f_n \in L^1(\mu; \mathbb{R}^n)$ is continuous, bounded and symmetric LDC can be applied to $f_\lambda^{c,i}$ and thus (4.20) holds, which proves that (4.17) is well–defined. It seems that for $\gamma_\lambda^{c,i}$, $i = 1, 2$ boundedness of $f_n(\cdot, x_2..., x_n)$, $x_2, ..., x_n \in \mathcal{X}$ in a neighbourhood of c would have been sufficient. However, if the functional derivative has to be computed at arbitrary peaks, boundedness of $f_n(x_1, ..., x_n)$ on $\mathcal{X}^n$ is needed.
The following proposition shows under which Condition (4.17) can be defined using an arbitrary approximate identity.

Proposition 4. *Let*

$$f_n \in C_0^0(\mathcal{X}^n) \equiv \{f_n : \mathcal{X}^n \to \mathbb{R} : f_n \text{ is continous and the support of } f_n \text{ is compact}\}$$

and h be a continuous test–function. Then, for an absolutely integrable function $\eta : \mathcal{X} \to \mathbb{R}$, that is $\int_{\mathcal{X}} |\eta(x)|dx < \infty$ with $\int_{\mathcal{X}} \eta(x)dx = 1$ and $\eta(x) = 0$ for all $x \notin \{x : |x - c| \leq 1\}$

$$\left| \frac{a_n}{n!} \sum_{k=1}^{n} \int_{\mathcal{X}^n} \lambda^{-1} \eta\left(\frac{x_k}{\lambda}\right) \prod_{i=1, i\neq k}^{n} h(x_i) f_n(x_1, ..., x_n) dx_1 ... dx_n \right|$$
$$\leq \frac{(B \cdot C)^{n-1}}{n-1!} \max_{x \in \mathcal{X}^n} f_n(x_1, x_2, ..., x_n) \tag{4.47}$$

Proof. Due to [Wal74, p.38, p.39] $\lambda^{-1}\eta\left(\frac{\cdot}{\lambda}\right)$ converges as a distribution on the set of all test–functions from $C_0^0(\mathcal{X})$ to δ^c, which holds if h is continuous. Thus, m–a.e. convergence is proven. Furthermore, $f \in C_0^0(\mathcal{X})$ and thus it attends its maximum on its (compact) support. Therefore, applying Lemma 2 and estimating, by using the maximum instead of the supremum in the previous theorem, yields the desired result. □

Let us summarize the results, that is which properties the density f needs to fulfill so that Conditions (4.26) and (4.27) can be verified. First, consider the approximate identities $\{\gamma_\lambda^{c,i}\}_{\lambda>0}$, $i = 1, 2$. Due to Lemma 1 (Condition (4.26)) and Theorem 3 (Condition (4.27)) the assumptions of LDC are fulfilled if the multi–object density $f_n \in L^1(\mu; \mathcal{X}^n)$ is continuous, bounded and symmetric. Second, if $f_n \in C_0^0(\mathcal{X}^n)$ is symmetric, Proposition 4 proves that Conditions (4.26) and (4.27) are valid for continuous test–functions $h \in \mathcal{H}$ if an approximate identity is constructed via [Alt12, 2.14 (2), p.114] using an absolutely integrable function. Note that the continuity of the test–function h is a non–restrictive constraint for the *Analytical Step* of deriving pointillist filters, since $h = 0$ or $h = 1$ are the only needed choices for deriving multitarget tracking filters. Since $C_0^0(\mathcal{X}^n) \subset \{f \in L^1(\mu, \mathcal{X}^n) : f \text{ continuous and bounded}\}$, in the following the multi–object densities are assumed to be in $L^1(\mu, \mathcal{X}^n)$, continuous, bounded and symmetric.

All pointillist tracking filters presented in Chapter 3 have in common that their joint PGFL is contained in $\mathcal{P}_2$, that is for all pointillist filters from Chapter 3 the considerations of Section 4.1 are valid. This can easily be justified by considering Table 3.1 from Chapter 3.

4.1.4 On the Connection of the Set Derivative from [Mah07b] and the Functional Derivative with respect to the Dirac Delta

Note that (4.17) comprises out of two independent limit processes, one for the functional derivative and the other for the approximation of Dirac delta. The set derivative proposed in [Mah07b, pp. 380, 381] can be defined using the family of step test–functions $\left\{\gamma_\lambda^{c,2}\right\}_{\lambda>0}$. Let $h \equiv 1_X$, $X \subset \mathbb{R}^d$, $d > 0$. Then by replacing the variable of the family of test–functions $\lambda > 0$ by the variable $\epsilon > 0$ used to define the derivative, that is using the family of test–functions $\left\{\gamma_\epsilon^{c,2}\right\}_{\epsilon>0}$ to approximate Dirac delta, one obtains

$$\frac{\partial \Psi}{\partial \gamma_\epsilon^{c,2}}(h) = \frac{\Psi(h + \epsilon\gamma_\epsilon^{c,2}) - \Psi(h)}{\epsilon} = \frac{\Psi(1_X + 1_{E_\epsilon^c}(x)) - \Psi(1_X)}{\mu(E_\epsilon^c)}, \tag{4.48}$$

where μ denotes the Lebesgue measure on $\mathbb{R}^d$, $c \in \mathbb{R}^d$, $d > 0$ and E_ϵ^c be a neighborhood of c with $\mu(E_\epsilon^c) = \epsilon$. Thus, the double limit from (4.17) reduces to a single limit due to the fact that the variable ϵ of the derivative and the Lebesgue measure of the support of γ_ϵ^c are chosen to be the same. Under the assumption that X and E_ϵ^c are disjoint

$$\lim_{\epsilon \searrow 0} \frac{\Psi(1_X + 1_{E_\epsilon^c}(x)) - \Psi(1_X)}{\mu(E_\epsilon^c)} \tag{4.49}$$

denotes the set derivative defined in [Mah07b, pp. 380, 381]. Thus, it can be seen that the set derivative can be defined using approximate identities.

4.2 Secular Functions

In this section the technique for deriving multitarget intensity filters using secular functions from [Str14e] is presented. First, the definition of a secular function for the univariate and the multivariate case from [Str14e] is presented. An example for unresolved targets is used to demonstrate that the family of pointillist filter encompasses tracking filters that cannot be represented by a PGFL from $\mathcal{P}_1$. Then, the concept of secular functions is extended to the class of PGFLs $\mathcal{P}_2$ defined in (4.11). Afterwards, secular functions of joint PGFLs are considered and the nth–order factorial moment of the PGFL of the Bayes posterior point process is derived.

4.2.1 Secular Functions on $\mathcal{P}_1$

This section follows [Str14e] to present the theory of secular functions for univariate and multivariate PGFLs.
To begin with, let Ψ is a univariate PGFL given by

$$\Psi(g) \equiv F\left(\int_Y g(y)q(y)dy\right), \tag{4.50}$$

$F(z) \equiv \sum_{n=0}^{\infty} a_n z^n$, $a_n \in [0,1]$, $z \in \mathbb{C}$, as defined in [Str14e, Sec. 4], where $g \in \mathcal{G}$ and $q : Y \to \mathbb{R}$ be a PDF on Y which is continuously differentiable at interior points of Y, except possibly for jump continuities. Here, $Y \subset \mathbb{R}^d$, $d > 1$ be a closed and bounded subset and $c \in Y$. Furthermore, let $\gamma : Y \to \mathbb{R}$ be bounded, non–negative and Lebesgue–integrable function, that is in particular some $B > 0$ exists with $|\gamma(y)| \leq B < \infty$, for all $y \in Y$. Then $\Psi(g + z\gamma) : \mathbb{C} \to \mathbb{C}$, considered a function of $z \in \mathbb{C}$, is analytic in an open neighbourhood of the origin. Let $\{\gamma_\lambda^c\}_{\lambda>0}$ be a family of test–functions.
The secular function corresponding to $\Psi(g)$, $g \in \mathcal{G}$ is defined by

$$J(\alpha; c) \equiv \lim_{\lambda \searrow 0} \Psi(g + \alpha\gamma_\lambda^c) = F\left(\int_Y g(y)q(y)dy + \alpha q(c)\right), \tag{4.51}$$

where the analyticity of F justifies taking the limit into the integral. Then, it can be easily observed, that

$$\frac{\partial \Psi}{\partial c}(g) = \frac{dJ}{d\alpha}(0; c) \equiv \frac{d}{d\alpha} J(\alpha; c)\Big|_{\alpha=0} \tag{4.52}$$

holds, which means that the functional derivative of Ψ is identical to the ordinary derivative of the corresponding secular function J.
Secular functions can also be defined for multiple derivatives. Therefore, according to [Str14e, Sec. 4.1] let $\Psi \in \mathcal{P}_1$ and define the secular function of $\Psi \in \mathcal{P}_1$ by substituting a test–sequence for a weighted train of Dirac deltas by

$$J(\alpha; y) \equiv \lim_{\lambda \searrow 0} \Psi\left(g(y) + \sum_{i=1}^{m} \alpha_i \gamma_\lambda^{y_i}(y)\right) = \Psi\left(h(x) + \sum_{i=1}^{m} \alpha_i \delta^{c_i}(x)\right), \tag{4.53}$$

where $y = \{y_1, ..., y_m\}$, $m > 0$, $\alpha \in \mathbb{C}^m$ and F being not a univariate PGF (as in (4.50)) but a multivariate PGF. Note that the right hand side of (4.53) uses the Dirac delta as an evaluation operator, not as a function. Then, due to [Str14e, Sec. 4]

$$\Psi_y(g) \equiv \frac{\partial^m}{\partial y_1 ... \partial y_m} \Psi(g) = \frac{d^m}{d\alpha_1 \cdots d\alpha_m} J(\alpha; y)\Big|_{\alpha_1=...=\alpha_m=0} \equiv J_\alpha(0; y). \tag{4.54}$$

Thus, (4.54) shows that the functional derivative with respect to multiple Dirac deltas is identical to the first order mixed derivative of the corresponding secular function.

This important fact is needed for example in the derivation of a multivariate intensity function and other summary statistics that are derived by evaluating a simultaneous differentiation with respect to several directions. For example in the derivation of the PHD filter [Mah13], [Str13a] the Bayesian posterior point process is differentiated with respect to impulses at the measurement positions. The notion of the Fréchet derivative is thus important because it justifies that the derivative from (4.54) is commutative, that is the order of the derivatives does not play a role.
Secular functions can also be defined in a multivariate fashion. Let therefore

$$F^{\text{multi}}(z_1, ..., z_k) \tag{4.55}$$

be a multivariate probability generating function (PGF). Then, the multivariate PGFL is given by

$$\Psi(g_1, ..., g_k) = F^{\text{multi}}\left(\int_Y g_1(y)q_1(y)dy, ..., \int_Y g_k(y)q_k(y)dy\right) \tag{4.56}$$

and the corresponding secular functions is defined by

$$J(\alpha_1, ..., \alpha_k; y) \equiv \lim_{\lambda \searrow 0} \Psi\left(g_1 + \alpha_1\gamma_\lambda^{c_i}, ..., g_k + \alpha_k\gamma_\lambda^{c_k}\right). \tag{4.57}$$

4.2.2 $\mathcal{P}_1$ is Not Exhaustive

In Chapter 3 a family of tracking filters is presented that can be derived using point process theory and PGFLs. It is seen that many well–known filters, called the family of pointillist filters, can be modeled using point process theory. The following consideration demonstrates that the family of pointillist filter contains tracking filters which cannot be modeled using the class of PGFLs $\mathcal{P}_1$ and why therefore an extension of the definition of secular functions from $\mathcal{P}_1$ to $\mathcal{P}_2$ is needed.
The definition of a secular function presented in [Str14e] uses PGFLs that can be represented by a function of single integrals (see (4.50)). This definition can be extended to model tracking applications that involve unresolved targets, that is, two or more targets that the sensor interprets as a single measurement. The general problem of processing $r \geq 1$ unresolved targets is modeled by using the PGFL of a multi–cluster process, that is

$$\Psi_r^{\text{multi}}(g_1, ..., g_k) \equiv F\left(\int_Y \cdots \int_Y g_1(y_1) \cdots g_k(y_r)q(y_1, ..., y_r)dy_1...dy_r\right), \tag{4.58}$$

where $F(z) \equiv \sum_{n=0}^{\infty} a_n z^n$. Let the PGFL $\Psi_r(g)$ be defined as the superposition of the multivariate version, that is $\Psi_r(g) \equiv \Psi_r^{\text{multi}}(g, ..., g)$. When writing out (4.58)

particularly for $r = 2$ one obtains

$$\begin{aligned}\Psi_{r=2}(g) =& a_1 \int_{Y^2} \prod_{i=1}^{2} g(y_i) q(y_1, y_2) dy_1 dy_2 \\ &+ a_2 \int_{Y^4} \prod_{i=1}^{4} g(y_i) q(y_1, y_2) q(y_3, y_4) dy_1 ... dy_4 \\ &+ a_3 \int_{Y^6} \prod_{i=1}^{6} g(y_i) q(y_1, y_2) q(y_3, y_4) q(y_5, y_6) dy_1 ... dy_6 \\ &+ ..., \end{aligned} \tag{4.59}$$

which shows that the measurement of each target is always due to exactly $r = 2$, measurements that the sensor could not resolve.

When computing the ordinary derivative of the secular function (4.53) of $\Psi_{r=2}$ one might wonder why these derivatives are well–defined, since products of Dirac deltas arise. However, this is not a problem, since each factor of a product of Dirac deltas possesses a unique argument (within the product), which implies that each integral is evaluated only once by Dirac delta. Note that a product of Dirac deltas can also be interpreted as a Dirac delta operator that is defined on a Cartesian product space.

In practice it is not known a priori how many targets are unresolved by a sensor, since the spatial distribution of targets naturally changes dynamically in enhanced target tracking applications. Therefore, a tracking filter needs to be able to change the number of unresolved targets also dynamically. However, from the definition (4.59) it can be seen that such a tracking filter cannot be constructed using $\Psi_r(g)$, since r needs to be fixed a priori. A more realistic assumption would be that an arbitrary upper bound $1 \leq K < \infty$ of unresolved targets is known, but the exact number of unresolved targets (of a sensor) is unknown. This can be modeled by a PGFL from $\mathcal{P}_2$ by

$$\Psi_{\text{res}}(g) = \sum_{n=1}^{K} a_n \int_{Y^n} \prod_{i=1}^{n} g(y_i) q(y_1, ..., y_n) dy_1 ... dy_n, \tag{4.60}$$

where it should be noted that $\Psi_{\text{res}} \notin \mathcal{P}_1$.

When $K = \infty$ is chosen, the PGFL (4.60) is identical to the PGFL of the generalized PHD intensity filter [CM12] (see also $\Psi_{\text{GenPHD}}(h, g)$ in Chapter 3). It is a special case of a general result given by [Moy62, eqn. (5.3)]. (The special case comes about when the probability measure $Q^{(n)}(\cdot|x_k)$ from [Moy62, eqn. (5.3)] is equivalent to a PDF.)

4.2.3 Extension of Secular Functions to $\mathcal{P}_2$

In the following, according to the considerations of Section 4.2.1, we assume the density f_n of the considered PGFLs to be symmetric, continuous, bounded and in

$L^1(\mu;\mathcal{X}^n)$.
We define for all $\Psi \in \mathcal{P}_2$ and $h \in \mathcal{H}$

$$J(\alpha;x) \equiv \lim_{\lambda \searrow 0} \Psi \left(h(x) + \sum_{i=1}^{m} \alpha_i \gamma_\lambda^{c_i}(x) \right), \tag{4.61}$$

where $x \in \mathcal{X} \subseteq \mathbb{R}^d$, $\alpha = (\alpha_1, ..., \alpha_m)^T \in \mathbb{C}^m$, $c = (c_1, ..., c_m)^T \in \mathbb{R}^m$, $\lambda > 0$ to be the secular function of the PGFL Ψ. Here, $\{\gamma_\lambda^{c_i}\}_{\lambda>0}$ is some family of test–functions for the Dirac delta. First, we are to show that this definition is well–defined, that is we need to show that the limit for $\lambda \searrow 0$ exists. Due to Section 4.1.3 we already know that the functional derivative with respect to the Dirac delta exists if the symmetric density f is bounded, continuous and in $L^1(\mu;\mathcal{X}^n)$. Since

$$\Psi \left(h(x) + \sum_{i=1}^{m} \alpha_i \gamma_\lambda^{c_i}(x) \right) = \sum_{n=0}^{\infty} \tilde{f}_\lambda^{c,i}(n), \tag{4.62}$$

where

$$\tilde{f}_\lambda^{c,i}(n) \equiv \frac{a_n}{n!} \int_{\mathcal{X}^n} \prod_{j=1}^{n} \left(h(x_j) + \sum_{i=1}^{m} \alpha_i \gamma_\lambda^{c_i}(x_j) \right) f_n(x_1, ..., x_n) dx_1 ... dx_n, \tag{4.63}$$

LDC can be applied analogously to the previous section to $\tilde{f}_\lambda^{c,i}$. For the application of LDC, it needs to be checked first that $\tilde{f}_\lambda^{c,i}$ converges for any $c \in \mathcal{X}$ m–a.e. to

$$\tilde{f}^c(n) \equiv \sum_{n \geq 1} \frac{a_n}{n!} \int_{\mathcal{X}^n} \prod_{j=1}^{n} \left(h(x_j) + \sum_{i=1}^{m} \alpha_i \delta^{c_i}(x_j) \right) f_n(x_1, ..., x_n) dx_1 ... dx_n, \tag{4.64}$$

and second that

$$|\tilde{f}_\lambda^c| \leq g, \text{ for some } g \in l^1(\mathbb{C}) \ m\text{–a.e.}, \tag{4.65}$$

where $l^1(\mathbb{C})$ denotes the complexification of $l^1(\mathbb{R})$. Note, that LDC can be applied to complex–valued functions [Rud87]. Due to the fact that (4.63) can be written as the finite sum of integrals Condition (4.64) is proven analogously to Lemma 1 and Proposition 4. The following Lemma proves that Condition (4.65) is valid.

Lemma 5. *For a factor $A_\lambda > 0$, which only depends on λ, a family of test–functions for Dirac delta $\{\gamma_\lambda^{c_i}\}_{\lambda>0}$ and the bounded symmetric density f_n, the function $g : \mathbb{N} \to \mathbb{N}$ defined by*

$$g(n) \equiv \sum_{k=0}^{n} \frac{1}{k!(n-k)!} m^k R^k A^k B^{n-k} C^{n-k}, \tag{4.66}$$

is in $l^1(\mathbb{C})$ *and it holds that* $|\tilde{f}_\lambda^c| \leq g$ *m–a.e., where*

$$\tilde{f}_\lambda^c(n) \equiv \frac{a_n}{n!} \int_{\mathcal{X}^n} \prod_{j=1}^{n} \left(h(x_j) + \sum_{i=1}^{m} \alpha_i \gamma_\lambda^{c_i}(x_j) \right) f_n(x_1, ..., x_n) dx_1 ... dx_n \tag{4.67}$$

and $|h| \leq B$, $B > 0$, $C \equiv \int_{\mathcal{X}} |f(x_1, ..., x_n)| dx_i$, $i = 1, ..., n$, $|\alpha_j| \leq R$, $\alpha_j \in \mathbb{C}$ $j = 1, ..., m$ *and* $A \equiv \sup_{x \in \mathcal{X}^n} (f_n(x_1, x_2, ..., x_n))$

Proof. First, we show that $|\tilde{f}_\lambda^c(n))| \leq g(n)$ for all $n \in \mathbb{N}$. Due to the binomial theorem, the boundedness of h and $a_n \in [0, 1]$ we obtain

$$\left| \tilde{f}_\lambda^c(n) \right| \tag{4.68}$$

$$= \left| \frac{a_n}{n!} \int_{\mathcal{X}^n} \prod_{j=1}^{n} \left(h(x_j) + \sum_{i=1}^{m} \alpha_i \gamma_\lambda^{c_i}(x_j) \right) f_n(x_1, ..., x_n) dx_1 ... dx_n \right| \tag{4.69}$$

$$= \left| \sum_{k=0}^{n} \frac{a_n}{k!(n-k)!} \int_{\mathcal{X}^n} \left(\prod_{j=1}^{n-k} h(x_j) \right) \left(\prod_{j=1}^{k} \sum_{i=1}^{m} \alpha_i \gamma_\lambda^{c_i}(x_j) \right) f_n(x_1, ..., x_n) dx_1 ... dx_n \right| \tag{4.70}$$

$$\leq \sum_{k=0}^{n} \frac{1}{k!(n-k)!} B^{n-k} \int_{\mathcal{X}^n} \left| \left(\prod_{j=1}^{k} \sum_{i=1}^{m} \alpha_i \gamma_\lambda^{c_i}(x_j) \right) f_n(x_1, ..., x_n) \right| dx_1 ... dx_n \tag{4.71}$$

For the evaluation of (4.71) we could use the multinomial theorem. However, we are only interested in finding an upper bound of the integral. Note that

$$\int_{\mathcal{X}^k} \prod_{j=1}^{k} \sum_{i=1}^{m} \alpha_i \gamma_\lambda^{c_i}(x_j) f_n(x_1, ..., x_n) dx_1 ... dx_k \leq m^k R^k A^k \int_{\mathcal{X}^k} \gamma_\lambda^{c_i}(x_j) dx_1 ... dx_k \tag{4.72}$$

where $|\alpha_i| < R$, for some $R > 0$ and $i \in \{1, ..., k\}$. The question how many terms $\prod_{j=1}^{k} \sum_{i=1}^{m} \alpha_i \gamma_\lambda^{c_i}(x_j)$ possesses can be answered by a simple combinatorial consideration. Since for fixed $j \in \{1, ...k\}$ each summand of $\sum_{i=1}^{m} \alpha_i \gamma_\lambda^{c_i}(x_j)$ yields m^{k-1} terms, $\prod_{j=1}^{k} \sum_{i=1}^{m} \alpha_i \gamma_\lambda^{c_i}(x_j)$ possesses m^k terms.
The left $(n-k)$–integrations yield C^{n-k}. Thus, we obtain

$$|\tilde{f}_\lambda^c(n)| \leq g(n) \tag{4.73}$$

for all $n \in \mathbb{N}$, which proves that $|\tilde{f}_\lambda^c| \leq g$ m-a.e.

Second, it holds that

$$g(n) \equiv \sum_{k=0}^{n} \frac{1}{k!(n-k)!} m^k R^k A^k B^{n-k} C^{n-k} \tag{4.74}$$

$$= \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} (m \cdot R \cdot A)^k (B \cdot C)^{n-k} \tag{4.75}$$

$$= \frac{1}{n!} (m \cdot R \cdot A + B \cdot C)^n \tag{4.76}$$

due to the binomial theorem and thus

$$\sum_{n=0}^{\infty} |g(n)| = \exp(m \cdot R \cdot A + B \cdot C) < \infty, \tag{4.77}$$

which proves that $g \in l^1(\mathbb{R})$. □

Hence, the extension of secular functions from $\mathcal{P}_1$ to $\mathcal{P}_2$ is proven. and thus the theory of secular functions can be applied to all members of the family of pointillist filters presented in Chapter 3.

4.2.4 Secular Functions for Joint PGFLs

This section follows [Str14e, Sec. 4.2] and studies secular functions for joint PGFLs. Joint PGFLs are defined in Sections 2.8 and 2.9. Let $\Psi(g, h)$ be a joint PGFL from the extension of $\mathcal{P}_1$ to two arguments, where $h : X \to \mathbb{C}$ and $g : Y \to \mathbb{C}$ are the corresponding non-negative, bounded (by one) and Lebesgue–integrable test–functions. To obtain the secular function of the joint PGFL, the simultaneous perturbation defined in [Moy62, Eqn. (4.11)] is applied by defining

$$g(y) \equiv \sum_{i=1}^{m} \alpha_i \delta^{y_i}(y), \ y \in Y \tag{4.78}$$

$$h(x) \equiv 1 + \sum_{j=1}^{n} \beta_j \delta^{x_j}(x), \ x \in X, \tag{4.79}$$

where $\alpha = (\alpha_1, ..., \alpha_m) \in \mathbb{C}^m$ and $\beta = (\beta_1, ..., \beta_n) \in \mathbb{C}^n$. For $m = 0$ and $n = 0$ the sums are defined to be zero. The secular function corresponding to the joint PGFL is defined by

$$J(\alpha, \beta; y, x) \equiv \Psi \left(\sum_{i=1}^{m} \alpha_i \delta^{y_i}(y), 1 + \sum_{j=1}^{n} \beta_j \delta^{x_j}(x) \right) \tag{4.80}$$

for the PGFLs studied in [Str14e]. By the application of the same arguments used in Section 4.2.3, secular functions of joint PGFLs can be extended to joint PGFLs with a

density f that satisfies the respective conditions ($f \in L^1(\mu; X^n)$ bounded, continuous and symmetric).
According to Section 4.2.3 the ordinary derivatives of J are identical to the functional derivatives of the PGFL Ψ, that is,

$$J_{\alpha\beta}(\alpha, \beta; y, x)|_{\alpha=0,\beta=0} = \Psi_{yx}(g, h)|_{g(\cdot)=0,\ h(\cdot)=1}. \tag{4.81}$$

4.2.5 Example: PHD Filter Update Equation Derivation Using Secular Functions

In the following, the applicability of secular functions is demonstrated exemplary at hand of the PHD filter, a member of the family of pointillist filters presented in Section 3.4.2. To this end, this section follows [Str14e, Section 5.1] and shows how the PHD filter can be derived via ordinary differentiation. According to Section 3.4.2 the joint target–measurement point process has the PGFL

$$\begin{aligned}\Psi_{\mathrm{PHD}}(h, g) = \exp\Bigg(&-\Lambda - \overline{N} + \Lambda \int_Y g(y) p_\Lambda(y)\, dy \\ &+ \overline{N} \int_X h(x)\mu(x)\Big(a(x) + b(x) \int_Y g(y) p(y|x)\, dy\Big) dx\Bigg).\end{aligned} \tag{4.82}$$

Using definition (4.80) together with the test–functions defined in (4.78) and (4.79) the corresponding secular function is given by

$$\begin{aligned}J(\alpha, \beta; y, x) =& K \exp\Bigg(\sum_{i=1}^{m} \alpha_i \Big(\Lambda p_\Lambda(y_i) + \overline{N} \int_X p(y_i|x) b(x) \mu(x) dx\Big) \\ &+ \overline{N}\Bigg(\sum_{j=1}^{n} \beta_j a(x_j)\mu(x) + \sum_{i=1}^{m}\sum_{j=1}^{n} \alpha_i \beta_j p(y_i|x_j) b(x_j) \mu(x_j)\Bigg)\Bigg),\end{aligned} \tag{4.83}$$

where

$$K \equiv \exp\Big(-\Lambda - \overline{N}\Big) \tag{4.84}$$

is a constant. First, the ordinary derivative of (4.83) for $n = 1$ with respect to α yields

$$\begin{aligned}J_\alpha(\mathbf{0}, \beta_1; y, x_1) \equiv& \frac{d^m}{d\alpha_1 \cdots d\alpha_m} J(\alpha, \beta_1; y, x_1)\Big|_{\alpha_1=\cdots=\alpha_m=0} \\ =& J(\mathbf{0}, \beta_1; y, x_1) \\ &\times \prod_{i=1}^{m}\Big(\Lambda p_\Lambda(y_i) + \overline{N} \int_X h(x) p(y_i|x) \mu(x) dx + \beta_1 p(y_i|x_1) b(x_1) \overline{N} \mu(x_1)\Big),\end{aligned} \tag{4.85}$$

where $\mathbf{0} = (0, ..., 0) \in \mathbb{C}^m$. Second, the derivative with respect to β_1 evaluated at $\beta_1 = 0$ gives

$$\begin{aligned} m_{[1]}(x_1|y) =& \frac{1}{J_\alpha(\mathbf{0}, 0; y, x_1)} \frac{d}{d\beta_1} J_\alpha(\mathbf{0}, 0; y, x_1) \\ =& a(x_1)\overline{N}\mu(x_1) + \sum_{i=1}^{m} \frac{p(y_i|x_1)b(x_1)\overline{N}\mu(x_1)}{\Lambda p_\Lambda(y_i) + \overline{N}\int_X p(y_i|x)a(x)\mu(x)dx}, \end{aligned} \tag{4.86}$$

which is the update equation of the PHD intensity filter. The derivation of (4.86) using functional derivatives can be found in [Mah07b] and also in [Str13a].
The (ordinary) differentiation of the secular function defined in (4.83) can be easily done analytically for the PHD filter. However, for PGFL with a complex structure or for a reduction of the numerical complexity other methods for computing the derivatives of secular functions are of interest. The following section gives a short overview of some options to carry out the differentiation.

4.3 Methods for Computing Derivatives of Secular Functions

Due to the previous sections the application of the theory of secular functions from [Str14e] to the family of pointillist filters is mathematically justified. Therefore, the computation of the summary statistics needed for the implementation and application of pointillist filters can be done by ordinary instead of functional differentiation. This section follows [Str14e], [Str14a] and [Str15] and gives an overview of existing, well–known methods that can be used for this *Analytical Step*.
In principal, however, every (exact or approximating) approach for ordinary differentiation can be used for deriving the summary statistics of a PGFL.

4.3.1 Application of Cauchy's Residue Theorem – Saddle Point Methods

In [Str15] saddle point methods for the JPDA and related filters are proposed. This method relies on Cauchy's residue theorem [FS09, p.236] and offers an approach to derive particle implementations of pointillist filters. In the following we study the approach analogously to the considerations made in [Str15].
Consider the complex–valued function

$$\alpha \mapsto J(\alpha; x) \equiv \Psi\left(h(x) + \sum_{i=1}^{m} \alpha_i \delta^{c_i}(x)\right), \tag{4.87}$$

where $x \in \mathbb{R}^n$, $n \geq 0$, $\alpha = (\alpha_1, ..., \alpha_m)^T \in \mathbb{C}^m$, $\Psi \in \mathcal{P}_2$ and $h \in \mathcal{H}$. Then, $J(\cdot; x)$ is analytic in some open region of the complex plane $\Omega \subset \mathbb{C}^m$ containing $(0, ..., 0)^T \in \mathbb{C}^m$ [Moy62, Sec. 4]. Denote by $\Omega_1, \ldots, \Omega_m \subset \mathbb{C}$ the projections of the open region Ω

from $\mathbb{C}^m$ to $\mathbb{C}$. Thus, Cauchy's coefficient formula [FS09, p.236] can be applied to $\alpha \mapsto J(\alpha; x)$, $x \in \mathbb{R}^n$. Hence, the ordinary derivative of $J(\cdot; x)$ is given by

$$\begin{aligned}&\frac{d^m}{d\alpha_1 \cdots d\alpha_m} J(\alpha; x)\Big|_{\alpha_1=\cdots\alpha_m=0} \\ &= \frac{1}{(2\pi i)^m} \int_{C_1} \cdots \int_{C_m} \frac{J(\zeta; x)}{(\alpha_1 - \zeta_1)^2 \ldots (\alpha_m - \zeta_m)^2} d\zeta_1 \ldots d\zeta_m,\end{aligned} \tag{4.88}$$

where $C_1, \ldots, C_m \subset \mathbb{C}$ are simple loops of $\Omega_1, \ldots, \Omega_m$ encircling $0 \in \mathbb{C}$, respectively. Note, that due to the concept of secular functions and Cauchy's coefficient theorem (which is an implication of Cauchy's residue theorem) the first–order factorial moment can be determined by contour–integration. Furthermore, the numerical approximation of contour-integral with saddle point methods has been extensively studied (see [Ble12], [FS09]) and can be used to evaluate (4.88) numerically. A derivation of the JPDA filter using saddle point methods and more details on the approach is presented in [Str15].

4.3.2 Classical Finite Differences

Let M be the total number of derivatives let $U : \mathbb{R}^n \to \mathbb{R}$ be a real–valued function. Then the finite difference approximation of the cross–derivative of U at $(0, ..., 0) \in \mathbb{R}^M$ is given by [Str14e, Section 6.1]

$$\begin{aligned}&\frac{d^M U}{dx_1 \cdots dx_M}(0, ..., 0) \\ &\approx \frac{1}{\epsilon_1 \cdots \epsilon_M 2^M} \sum_{\sigma_1, \ldots, \sigma_M = 0}^{1} (-1)^{\sigma_1 + \ldots + \sigma_M} U((-1)^{\sigma_1}\epsilon_1, ..., (-1)^{\sigma_M}\epsilon_M),\end{aligned} \tag{4.89}$$

where the increments $\epsilon_i > 0$, $i = 1, ..., M$ should small enough to guarantee a good approximation.
This approach might be impractical for large values of M and the problem of underflow can arise due to the alternating signs [Str14e].

4.3.3 Maclaurin Series Expansion

For the examples studied in [Str14e, Sec. 5.] the PPPs are used, which implies that corresponding secular function takes the form

$$J(\alpha, \beta_1; y, x_1) = c_0 \exp(c^T \alpha + \pi(\alpha + \beta_1 \pi(\alpha; x_1))), \tag{4.90}$$

where $\pi : \mathbb{R}^M \times X \to \mathbb{R}$ and $X \subset \mathbb{R}^n$, $n \geq 0$ is the target space. Furthermore, $\pi(a) \equiv \int_X \pi(a; x) dx$, $x \in X$, $a \in \mathbb{R}^M$. Then, expanding the function π in a Maclaurin (Taylor) series at $\mathbf{0} \in \mathbb{R}^M$ yields

$$\pi(a; x) \approx \pi(\mathbf{0}; x) + [\nabla \pi(\mathbf{0}; x)]^T \alpha + \frac{1}{2} \alpha^T [\nabla^2 \pi(\mathbf{0}; x)] \alpha, \tag{4.91}$$

where $\nabla\pi(\mathbf{0}) \equiv (\nabla\pi_l(\mathbf{0};x) : l = 1,...,M) \in \mathbb{R}^M$ and $\nabla^2\pi(\mathbf{0};x) \in \mathbb{R}^M \times \mathbb{R}^M$ are the gradient and the Hessian matrix of $\pi(\alpha;x)$ evaluated at $\mathbf{0} \in \mathbb{R}^M$, respectively. Substituting the first order Maclaurin expansion into the secular function implies that the intensity is given by

$$m_{[1]} = \pi(\mathbf{0};x_1) + \sum_{l=1}^{M} \frac{\nabla\pi_l(\mathbf{0};x_1)}{c_l + \int_X \nabla\pi_l(\mathbf{0};s)ds}. \tag{4.92}$$

The details can be found in [Str13a, Sct. 6.2].

4.3.4 Automatic Differentiation

In [GLLZ14] automatic differentiation (AD) methods are presented, which evaluate the derivatives of a function without finding the symbolic derivative. AD methods are exact and can be used to compute the first–order factorial moment point–wisely, e.g., for the determination of particle–weights [Str14a].

4.4 Conclusion and Future Work

It is shown under general assumptions that a mathematically correct definition of the functional derivative with respect to the Dirac delta exists. The definition is based on a sequence of regular distributions, called approximate identities. It is closely related to the considerations on the constructive definition of the Radon-Nikodym derivative from [GMN97, p.145], but proven in a different way. The close relation to Mahler's definition of the set derivative [Mah07b] is presented. Proofs verify its definition for a large class of PGFLs, which is suitable for almost all practical applications in multitarget tracking. Furthermore, it is shown that the concept of secular functions from [Str14e] can be extended to this larger class of PGFLs. Thus, in particular it is shown in a mathematically correct way, that the theory of secular functions can be applied to the family of pointillist filters presented in Chapter 3. This implies that numerical approximation schemes for ordinary differentiable functions like saddle point methods, Maclaurin series expansion, finite differences, etc. and exact methods like automatic differentiation can be applied to almost all practically relevant multitarget tracking filters.
Future work will focus on the evaluation of numerical approximation methods for the derivation of pointillist filters.

Part II

An Application to Emitter Tracking under Multipath Propagation

CHAPTER V

The Challenge of Blind Mobile Localization

The localization and tracking of mobile electromagnetic emitters, e.g., mobile phones, in an urban environment is an important and ambitious challenge. For a large class of safety, emergency and security applications the accurate localization of the emitter position, also referred to as mobile station (MS), is inevitable. Furthermore, the positioning and tracking of mobile electromagnetic emitters in cellular networks is implied by commercial interests [Alg10]: Fleet operations with delivery, taxi, bus, car sharing [Zha00], location sensitive billing [ZGL02], the design and architecture of wireless systems [HJL04], [ALLH02], [WHBL00], [WSL$^+$04], [Lan99], location based network access, radio resource management, road traffic management [BD00], route guidance and navigation are only a few examples which make the localization of mobile emitters in cellular networks a highly developed research field of current interest [Alg10]. The mentioned approaches for positioning in cellular networks have in common that they make explicitly use of the available cooperation between the electromagnetic emitter, the network (provider) and the observer station (OS). However, if the localization has to be done passively, that is, without emitting a signal by the OS and without any cooperation in between the MS, the OS and the network (provider) the problem is called the blind mobile localization and tracking of mobile electromagnetic emitters (BML) [Alg10]. This implies that the OS has to determine the location of the MS by only inspecting the transmitted electromagnetic waves. Possible fields of application are civilian security, safety and emergency scenarios, where the object of interest is the location of a non–subscribed phone user. Besides the term "blind" BML is hindered by an additional aspect, which complicates the problem. Typically, a BML scenario is located within a densely built–up area. Thus, due to physical propagation effects like reflection, diffraction and scattering the OS receives multiple signals that have traveled along different paths and which are called *multipaths* (see Figure 5.1). This implies that the standard multitarget tracking assumption, that is, *one target generates at most one measurement* is not valid in BML and therefore

complicates the derivation of an appropriate tracking algorithm.
In this chapter the fundamentals of BML are introduced and open questions in existing work are identified to motivate the contribution of Chapters 6 and 7.
This chapter is structured as follows. In Section 5.1 the fundamentals of BML are formulated. First, the boundary conditions of BML are explained in Section 5.1.1. Then, properties of the path propagation and methods for simulating the propagation of the radio channel, that is, ray tracing simulations and their limitations for the application within BML data fusion algorithms are explained in Section 5.1.2. The frameworks of BML for simulated and real world data, consisting out of a mobile antenna array, algorithms for blind channel estimation, a sophisticated ray tracing simulation and the sensor data fusion algorithms, which are studied in this thesis, are presented in Section 5.1.3. In Section 5.2 the related work on BML is discussed. Finally, limitations and open questions of existing work are identified and used to formulate and motivate the main contribution of the second part of this thesis in Section 5.3.

5.1 Fundamentals of Blind Mobile Localization

BML is a challenge that needs to be solved using passive and non–cooperative techniques for localization. In [Alg10, Section 2.1] a detailed taxonomy of localization methods is presented and BML is situated within this description. The boundary conditions of BML are presented next in Section 5.1.1, following the precise study presented in [Alg10, Section 2.2]. The application of ray tracers for the incorporation of context information is discussed in Section 5.1.2. Finally, a complete overview of the BML frameworks studied in this thesis to solve the task of BML is presented in Section 5.1.3.

5.1.1 Boundary Conditions of Blind Mobile Localization

BML is a short, but not exact acronym for the description of the studied challenge. According to [Alg10, Section 2.2] its correct definition is formulated by the statement

"Blind localization/tracking of a mobile station in urban scenarios using a single moving observer station equipped with an antenna array."

The scenario of BML is visualized in Figure 5.1. Four key ingredients are contained in this problem statement, which are used in [Alg10, Section 2.2] to characterize the boundary conditions of BML in the following way.

1. ***"Blind localization/tracking..."*** The OS locates and tracks the MS position in a passive (without emitting a signal) and non–cooperative way (without any cooperation between the MS and the network provider or the OS). This yields to the fact that

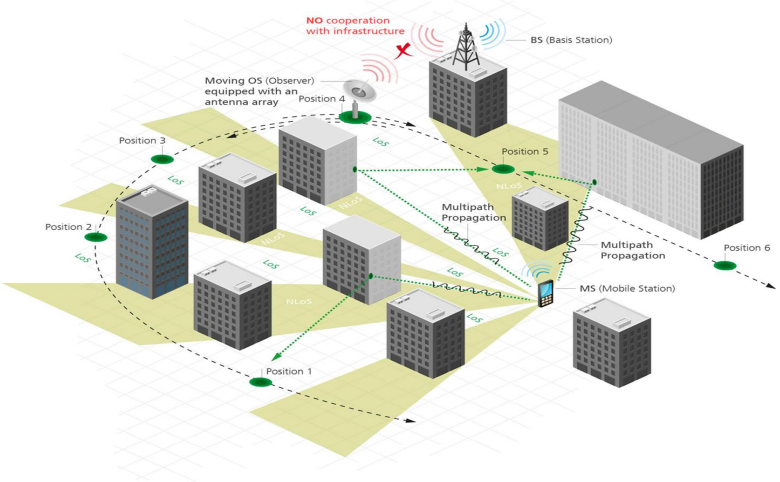

Figure 5.1: Visualization of the BML Scenario. A single OS equipped with an antenna array tries to localize and track an MS using only the received electromagnetic waves (multipaths). Due to an explicit exploitation of the multipaths via context information coming from a ray tracer, a BML tracking algorithm is able to track and localize an MS under LoS as well as NLoS conditions.

only the physically received electromagnetic signal is processed. Hence, blind channel estimation techniques [HKT15] have to be applied for the multipath characterization. Since the signal is not known, the absolute lengths and runtimes of the signal's paths cannot be estimated. Instead, a multipath is characterized (among other quantities like the azimut angle of arrival (AoA) and elevation angle of arrival (EoA)) by its relative time of arrival (RToA).

2. ***"...of a mobile station..."*** The target, which is an electromagnetic emitter and referred to as MS, moves, which implies that not only localization but also tracking algorithms have to be derived. The electromagnetic waves are emitted radially by the MS. According to [Alg10] BML multitarget scenarios reduce without loss of generality to single target scenarios, due to the fact that the multipaths emitted by each target are separable with respect to their signal form. However, in case of non–separability all target tracking algorithms presented in Chapter 6 can easily be applied to multitarget scenarios.

3. ***"...in urban scenarios..."*** The challenge of BML is situated in an urban environment, that is, in a densely built–up area. Since the radially emitted electromagnetic signal of the MS is transformed by reflection, diffraction and scattering during its

propagation, the antenna array of the OS receives multiple signals, which are called multipaths. In conventional localization techniques multipaths deteriorate the estimation process and thus target state estimates are only valid under line of sight (LoS) conditions, which implies that such approaches have a limited applicability in urban environments. In contrast to that, BML explicitly exploits the received multipaths by the incorporation of context information of the urban environment in terms of a ray tracer prediction. Due to the application of assignment techniques BML data fusion algorithms are capable to localize and track the MS either under LoS or non–line of sight (NLoS) conditions [Alg10].

4. ***"...using a single moving observer station equipped with an antenna array."*** A single moving OS implies that the ray tracing prediction needs to be done continuously, that a mission planning (an online estimation of the optimal OS route in terms of target state estimation accuracy) can be performed and that the carried antenna array has to be rather small. An array of antennas has to be used to resolve the multipaths and to characterized them in terms of their AoA, the EoA and the RToA. However, the number of resolvable multipaths is restricted by the number of elements of the antenna array. For example, the antenna array used for recording the real world data presented in Section 6.4 has five elements. Details on antenna arrays can be found in [Bro12] and [Alg10].

In Chapter 6 and Chapter 7 the purpose is to answer open questions with respect to the aspects of target tracking in a BML scenario. Therefore, the processing of the signal, that is in particular, the blind channel estimation of the received electromagnetic signal to characterize the multipaths in terms of their AoA, RToA and the EoA is outside the scope of this thesis. For details on the blind channel estimation the interested reader is referred to the literature. An overview over blind channel estimation techniques is given in [ZT95] and [TP98]. The parameter estimation used within this thesis is presented in [HKT15], which is based on [GS96], [YS94] and [TXK94].

5.1.2 Path Propagation and Ray Tracing

According to the boundary conditions given above, the localization and tracking of an electromagnetic emitter in an urban environment implies that the emitted signal is transformed into multiple signals (each characterized by AoA, EoA and RToA). The multipath propagation is caused by essentially three physical effects, that is, reflection, diffraction and scattering. These effects are described in [Alg10, Section 2.5.2] and standard literature on the propagation of electromagnetic waves. The mathematical model of the split–up signal is usually given by the impulse response, which consists out of a sum of Dirac pulses (see [Alg10, Equation (2.1)]). A detailed discussion on the impulse response can be found in [Alg10] and the references cited therein.

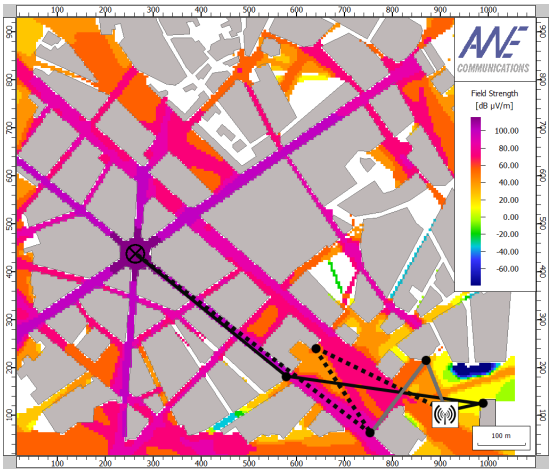

Figure 5.2: Visualization of the field strength prediction of a ray tracing simulation: For a given observer (black cross) – mobile station (antenna) constellation the color at the emitter location indicates the received field strength at the observer. Three multipaths are visualized (black solid, block dotted, gray) and the interaction points are plotted as black dots © 2013 IEEE.

Ray tracers are well established in the field of network planning and the system design of mobile communication systems. Based on a city map, a ray tracer predicts the set of multipaths that are emitted by the MS and received by the OS. The ray tracer used for the evaluation of the methods proposed in Chapters 6 and 7 is computationally efficient and accurate in its prediction [HWLW03]. It is based on [HWLW99], [WHL99], [HWLW03] and the fundamentals given in [Gla89]. Figure 5.2 shows the ray tracing prediction in terms of the received field strength in an urban environment including a subset of multipaths for a fixed MS–OS constellation. Figure 6.8 visualizes a ray tracing prediction in terms of the received field strength for the city of Erlangen.

The ray tracer prediction is based on a city and building map of the investigated scenario. Problems for any BML localization and tracking algorithm arise if these maps do not represent the reality, which might happen due to inaccurate or non–realistic models of the buildings or the city map. But even if the map information of the urban environment is perfect, static and moving obstacles like cars, trucks, pedestrians, trees, etc. cannot be modeled by the ray tracer and deteriorate the localization and tracking result. Additionally, several challenges for a correct ray tracer prediction are given by the receiver side. First, the fact that in a BML scenario a mobile OS is used implies that the antenna array cannot be arbitrary large. Thus, the number of elements and therefore also the number of resolvable multipaths is restricted. Fur-

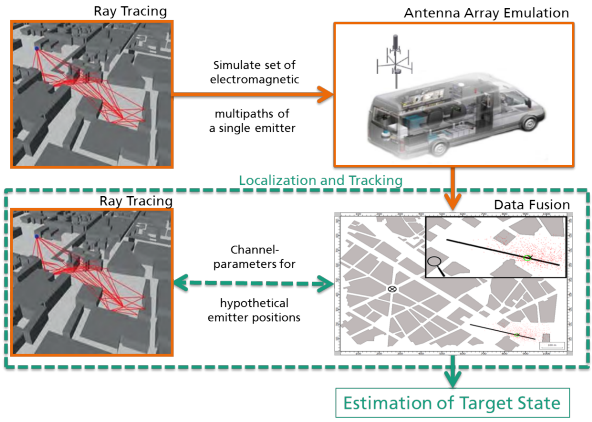

Figure 5.3: BML framework used for simulated (single–target) scenarios. Based on the true trajectory of the MS (ground truth), a fixed OS position and the city and building map of the urban environment the ray tracer predicts for each time instance a set of electromagnetic waves. Then, the set of multipaths is used as input for the emulation of an antenna array, which creates the set of measurements for the data fusion algorithm. Within the data fusion algorithm the ray tracer prediction is used for the evaluation of the assignment–based likelihood function. Ray tracing visualization: © 2015 AWE Communications. OS model: © 2015 Saab Medav Technologies GmbH. Data fusion: © 2013 IEEE.

thermore, due to fading certain predicted multipaths might not be received. All these effects have to be modeled by a ray tracer to guarantee an accurate prediction of the path propagation.

5.1.3 Blind Mobile Localization Framework

The BML frameworks studied in this thesis are closely related to the frameworks proposed in [Alg10, Figure 2.18] and consist out of four components. First, an antenna array with five–elements, which is carried by the mobile OS together with a receiver and a direction finding software is used. Obviously, this component is only needed if real world experiments are carried out. Within the simulation framework the antenna array output is emulated. Second, a blind channel estimation algorithm, presented in [HKT15], is applied to extract a set of multipaths out of the received signal. Due to the fact that no cooperation between the MS and the OS is available, the estimation characterizes a multipath by its AoA, RToA and the EoA. This component is used within the real world data framework to extract the multipaths. The third component is the ray tracer, which predicts for each OS–MS constellation the set of multipaths based upon the city and building map of the investigated environment. It finds its application within real world scenarios as a database used by the data fusion algorithm

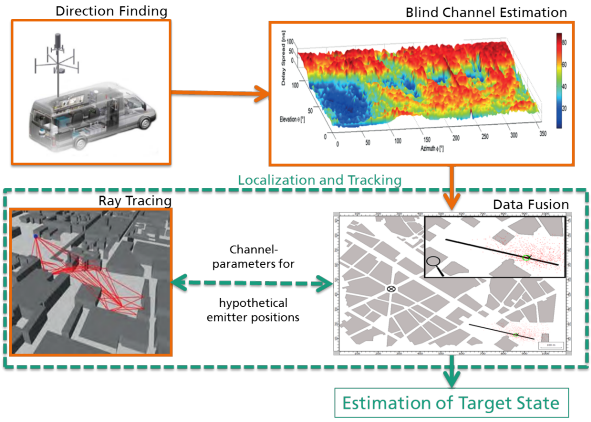

Figure 5.4: Processing chain of a real world BML scenario. After the electromagnetic signal is received by the antenna array, that is carried by the mobile OS, the received signal is processed by the blind channel parameter estimation proposed in [HKT15], which then outputs the set of received multipaths to the data fusion algorithm. The localization and tracking is done analogously to the BML framework used for simulated scenarios. Ray tracing visualization: © 2015 AWE Communications. OS model, parameter estimation visualization: © 2015 Saab Medav Technologies GmbH. Data fusion: © 2013 IEEE.

to model the likelihood function, that is, the assessment of hypothetical MS locations. For the simulation BML framework the ray tracer is additionally used to generate multipaths. For each time instance the ray tracer predicts the set of multipaths for the true position of the MS and the OS location. This set of true multipaths is then used as an input for the antenna emulation. Finally, the fourth component of the BML framework is the tracking algorithm studied in Chapter 6. It fuses the information gathered by the other three components and yields an estimate of the target position (plus some additional information like the number of false multipaths, the probability of the target's existence and an estimation of the covariance matrix, see Chapter 6). The data fusion is essentially based upon the idea that the hypothetical MS position which produces the set of predicted multipaths that fits best to the set of received multipaths is the most likely MS location [Alg10]. This incorporation of context information in terms of the ray tracer prediction distinguishes this approach from classical direction finding methods that suffer from multipath propagation and do not work under NLoS conditions. The approach from [Alg10] followed in this thesis exploits the multipath propagation and works either under LoS or NLoS conditions. In the following, the part of data fusion is studied. A detailed investigation of the remaining components of the BML framework is outside the scope of this thesis. First, related work is discussed and used to formulate the open questions that are answered

in the Chapters 6 and 7.

5.2 Related Work

Since conventional direction finding approaches use a single AoA, multipaths deteriorate the localization result or make it even impossible to estimate the position of the MS. In contrast to such classical localization approaches, several branches of the multitarget tracking community study how multipaths can explicitly be incorporated into appropriate data fusion algorithms. In the following, existing work of emitter localization and tracking under the influence of multipath propagation is presented. It is separated into non–cooperative and cooperative approaches.

5.2.1 Non–Cooperative Methods for Tracking and Localization Under Multipath Propagation

In [Alg10] and the related papers [ADKT08b], [ART04], [ADKT08a], [ADKT06a], [ADKT06b] the challenge of BML is studied fundamentally. The problem of BML is approached by incorporating higher–level information about the urban environment. For a given MS–OS constellation the fusion algorithm possesses on the one hand the measured multipaths and on the other the predicted multipaths of the ray tracing simulation. After computing the cost (or distance) matrix of normalized distances between the path–parameters for the two sets of predicted and measured multipaths, a standard assignment algorithm like the Munkres algorithm [BL71] determines the best match between the two sets of multipaths using the cost matrix. Based on this assignment, likelihood functions for essentially three localization and tracking algorithms are presented in [Alg10]: First, a simple point estimation algorithm maximizes a specific likelihood function for a set of multipaths. Second, the sampling importance resampling (SIR) filter is applied to the problem: Each particle is considered as a hypothetical MS location and the emitted multipaths are determined by the ray tracer for each particle and a fixed OS. Then, a likelihood function which takes a probability of detection and a clutter density for multipaths into account, is used for updating the particle weights. Finally, an MHT–based estimation approach is proposed. By considering all possible subsets of the measurement set the subsets with the largest probability are selected and the hypotheses which pass a threshold test are evaluated. The hypothesis with the largest likelihood value is then used as an estimate of the target location in the current iteration. To keep the computational load moderate several pruning and merging strategies are discussed. The derived tracking approaches are verified using synthetic and real world data of single target scenarios. The idea of the assessing hypothetical emitter positions proposed in [Alg10] and the corresponding papers is closely related to the idea of fingerprinting, which is first carried out for BML in [KST06]. There, AoA fingerprints, which are pre–calculated

by a ray tracer for a single OS, are used to perform the state estimation of the target. Furthermore, in [Alg10] a standard MHT filter is proposed for performing a parameter tracking, which is closely related to the ideas of track before detect (TBD), that is, the proposed algorithm increases the probability of detection and decreases the false alarm probability.

In [VGL+14], [VGLL12], [VGLL13] the authors present localization algorithms for BML–scenarios. In [VGL+14] and [VGLL12] a localization using a multipath component distance (MCD) based fingerprinting technique is described. The authors therein assume that the position of the MS could have already been roughly estimated by the cellular network. The fingerprinting techniques used in these works concentrate on a small field of view and a high accuracy of the localization. Analogously to [Alg10] a ray tracer is used to predict the path propagation between the MS and multiple base stations (BS). Given the estimated paths, which are obtained by a high resolution algorithm (HRA) and the database of the predicted paths, the MCD [CCS+06] is used to determine the most likely cell, that is, the most likely MS position. Then, a pairing procedure, where the MCD is computed for all predicted and estimated pairs and synchronization are performed. The time synchronization is needed since a comparison of the RToAs is not recommended by the authors. This is due to the fact that some of the predicted paths may be not received. In [VGLL13] the polarization diversity is incorporated as additional factor into the MCD [CCS+06] to improve the localization. However, no tracking is performed and an estimation of the position of the MS needs to be available for initialization purposes.

In [OSD15] a time difference of arrival (TDoA) approach is proposed to solve the problem. There, the last scatterers before the emitter are modeled as virtual receivers. The measurements of these virtual receivers are then processed by a TDoA approach. However, this innovative work has some drawbacks. First, at least five interactions, that is, five multipaths are needed to perform a localization in three dimensions. This might not always be the case (small antenna arrays might not produce such a large number of measurements). Furthermore, only reflections can be used to model the virtual receivers, that is, the effect of scattering or diffraction cannot be considered.

In [SD13b], [SD13a], [SNDE14], [SSAA12] the exploitation of multipaths arising when using an active radar to localize a target is studied. Therefore, the environment is assumed to be known exactly. Several performance bounds for different scenarios are derived therein.

Multipath measurements also occur in over Over–The–Horizon radar (OTHR) scenarios. There, ionosphere layers act as reflectors and enable the radar to localize and track targets that are behind the horizon. In [RBSW15a] and [RBSW15b] two ionosphere layers are assumed and the maximum likelihood probabilistic multi hypothesis tracker (ML–PMHT) is proposed to solve the problem. In these works a simple ionosphere model is used to model the path propagation. However, also ray tracers can be used

to model and predict the multipaths for an OTHR scenario with an enhanced ionosphere model. Other methods for solving OTHR scenarios are cited in [RBSW15a]. They are given by the multipath PDA (MPDA) algorithm [PE98], the multiple detection multi hypothesis tracker (MD–MHT) [SCAS13], the PMHT [CDC03] and the multiple model unified PDA (MM–UPDAF) [CD03].
In multistatic active sonar scenarios multiple transmitter (Tx) and receiver (Rx) are used to localize and track the target underwater. The multistatic setup enables to use different aspect angles for the detection of the target and to improve the estimation quality by a fusion of the single measurements [Bro12]. Thereby, reflections of objects underwater imply multipath propagation. The incorporation of these multipaths in multistatic sonar applications is considered in several publications. In [Mic15] a single or multiple known Rx locations are assumed and the Tx locations and the points of reflection are estimated. The estimation of the Tx locations, given known Rx locations and reflection points is studied in [MBE14]. The estimation of a single Rx location given known points of reflection and a known single Tx position is investigated in [DBE11a] and [DBE11b].

5.2.2 Cooperative Methods for Tracking and Localizing Targets Under Multipath Propagation

The study of methods in cooperative scenarios under multipath propagation is highly developed and their study is outside the scope of this work. However, the following work on cooperative localization and tracking is remarkable and of particular interest. In [Li14] several multipaths are processed that are characterized by the time of arrival (ToA) and the AoA (scenario using a monostatic multiple–input multiple–output (MIMO) radar) or the AoA and the exact path length (multistatic range–based scenario). This information is available, since the waveform emitted by the MS is known exactly. The authors propose a combined algorithm that performs simultaneously target tracking, data association and multipath modeling in a Bayesian estimation framework and is called simultaneous target and multipath positioning (STAMP) [LK14]. Within the STAMP algorithm the multipath parameters are modeled using a single–cluster PHD by describing a multipath parameter as an RFS. The target process is then defined as a parent process, that is updated via a particle or Kalman filter. The multipath channel parameter is formulated as daughter process and applied by a Gaussian Mixture implementation of the PHD filter. The innovation of this work on cooperative methods is to perform the target tracking and the signal processing simultaneously.

5.3 Limitations and Open Questions of Existing Work

As mentioned in Section 5.2.1 the localization and tracking under the influence of multipath propagation is of interest in various fields of research. By identifying limitations

and open questions in the existing work this section motivates the main contributions of Chapters 6 and 7.

In [Alg10] standard multitarget tracking approaches like the sampling importance resampling SIR and the MHT filter are combined with an appropriate probabilistic likelihood function to track electromagnetic emitters under the influence of multipath propagation. However, the application of other enhanced pointillist filters from Chapter 3 to BML has not been done so far. Especially filters that do not use target labels, that is, pointillist filters performing target superposition (see Section 3.4) are of particular interest. Furthermore, the derivation and definition of additional likelihood functions promises an improvement in solving the task of BML. Therefore, in Chapter 6 pointillist filters applying target superposition (see Section 3.4) and novel likelihood functions are proposed to solve the challenge of BML. The methodologies are compared using synthetic and real world data.

In [Alg10, Section 9.2] the author mentions that future research should investigate the verification of data fusion algorithms for BML using real world scenarios with a smaller bandwidth than 120 MHz. An evaluation of the proposed tracking filters from this work is carried out using real world data, where the transmitted signal has a bandwidth that is smaller than 5 MHz, in Section 6.4.

In [Alg10, Section 7.3] a standard track–oriented MHT approach to perform a parameter tracking in the measurement space is given using a standard handling of hypothesis in all parameters. The author mentions that *the application of multipath tracking is optional and ... not recommended in situations with high clutter and dense multipath components. Since then the multipath tracks can not be resolved or too much false tracks arise, which might produce even deterioration of the final localization result.* New parameter tracking algorithms need to be enhanced considering these factors. Furthermore, the authors from [VGL$^+$14], [VGLL12], [VGLL13] mention that a processing of the RToA characterization of multipaths is not recommended. Instead, they perform a time synchronization. Thus, in Chapter 7 the issue of clutter multipaths, that are received before the first target related multipath is identified and enhanced MHT filtering equations are derived via a marginalization over clutter hypothesis. The proposed methods are compared to a standard track–oriented MHT algorithm similar to the one applied in [Alg10, Section 7.3] using synthetic data.

CHAPTER VI

Blind Mobile Localization Using PHD Intensity Filters

In the following, standard (see Section 3.4.2) and generalized (see Section 3.4.4) PHD intensity filters are applied to the challenge of BML. As mentioned in the previous chapter the members of this subclass of pointillist filters, which superpose targets, have not yet been applied (to the knowledge of the author) to track electromagnetic emitters passively and without any cooperation under multipath propagation. The fact that the standard target–oriented measurement model, that is, *one target generates at most one measurement per sensor scan* is not valid for the challenge of BML implies that either an adaption of multitarget tracking filters, which use a standard target–oriented measurement model to the given boundary conditions is needed or filters have to be apply which employ a generalized target–oriented measurement model, that is enabling a target to generate more than one measurement per sensor scan.
As mentioned in the previous chapter the challenge of BML bears several challenging tasks. This chapter focuses on achieving sequential Monte Carlo (SMC)–implementations of different versions of PHD intensity filters. This comprises the definition of the observation space, the adaptation of existing target extraction schemes, different definitions of appropriate likelihood functions, approximations of numerically complex update equations, solutions to specific implementation issues and numerical comparisons of the different approaches using simulated and real world data.
This chapter is structured as follows. In Section 6.1 SMC–implementations of PHD intensity filters which use a standard target–measurement model (see Section 3.4.2) are applied to the task of BML by introducing a global assignment–based likelihood function and a generalized target state extraction scheme based on the particle grouping approach presented in [RCV10], which takes into account the mismatch in the target–oriented measurement model. An adaption of the decomposition of a likelihood function presented in [SW09] which is linear in target space to BML is presented

in Section 6.2. Furthermore, both likelihood functions are compared within an SMC-implementation of the adapted standard ifilter. Afterwards, Section 6.3 presents the generalized intensity filter and a numerical comparison to the generalized PHD filter [CM12] is shown. Then, approximations of the update equation of generalized PHD intensity filters are derived which reduce the numerical complexity of the filter update without assuming a joint distribution of the measurements in the parameter space. Then, the generalized PHD filter is compared in a numerical example to the adaption of the standard iFilter in Section 6.3.4.3. The same numerical comparison is done using real world data in Section 6.4. Finally, the conclusions are drawn and future work is discussed in Section 6.5.

Own publications on this subject: The adaption of the PHD intensity filter to BML using a generalized target state extraction scheme from Section 6.1 is published in [DGK13a] © 2013 IEEE. In [DGK14a] © 2014 IEEE the application of a decomposition of a likelihood function that is linear in target space to BML from Section 6.1 is published. The generalized intensity filter presented in Section 6.3 and Section 3.4.4 is derived for the first time in [Deg14] © 2014 IEEE. In [DGK14b] © 2014 IEEE and [DGK15] the derivation of the numerical approximation for generalized PHD intensity filters is given. The application of the generalized PHD filter to BML and a numerical comparison to the adaption of the standard intensity filter from Section 6.1 is published in [DGK15].

6.1 Standard SMC–PHD Intensity Filter Using a Generalized Extraction Scheme

In this section SMC–implementations of filters that propagate the intensity (or PHD, first order moment, etc.) like the iFilter [SKSC12] or the PHD–filter [RCV10] are applied to BML. Such filters are referred to as PHD intensity filters according to Section 3.4.2 and the fact that under certain parameterizations both filters yield the same update of the first order moment. Different definitions of the observation space of a PHD intensity filter within the BML framework are presented. Afterwards, a target state extraction scheme, called grouping of particles, which is originally presented in [RCV10], is generalized for processing multipath measurements in an SMC–implementation of the PHD intensity filter. Finally, the generalized grouping of particles approach is compared in a numerical evaluation to a standard target state extraction scheme, using a single target scenario and a ray tracing simulation.
This Section is organized as follows. Section 6.1.1 describes the problem formulation. A definition of the probabilistic likelihood function from [Alg10] is given in Section 6.1.1.2. In Section 6.1.1.3 different observation space definitions are presented. Issues concerning the distinct definitions, the problem of target state estimation and the

connection between observation and target space are discussed in Section 6.1.1.4. In Section 6.1.2 a generalization of the grouping of particles approach from [RCV10] is developed. Section 6.1.3 studies different methods for the target state extraction. It divides into a simple mean computation for a single–target scenario (see Section 6.1.3.1) and two newly developed generalized grouping of particles approaches (see Sections 6.1.3.2 and 6.1.3.3). A numerical evaluation using an SMC–implementation of the iFilter can be found in Section 6.1.4.1 before the conclusions are drawn in Section 6.1.5.

6.1.1 Formulation of the Problem

6.1.1.1 Measurement Space of BML

In the following the measurement or parameter space of BML used within Chapter 6 and Chapter 7 is defined. Due to the urban environment physical propagation effects like reflection, diffraction and scattering affect the emitted electromagnetic signal. This leads to the fact that the OS receives multiple signals which have traveled along different multipaths. Each received multipath is characterized by its AoA and its RToA. The sequence of measurement sets up to and including time $k \in \mathbb{N}$ is defined by

$$Z^k \equiv \{Z_k, Z^{k-1}\}, \tag{6.1}$$

where

$$Z_k \equiv \{z_k^j\}_{j=1}^{m_k} = \{\eta_k^j + \nu_k^j\}_{j=1}^{m_k} \tag{6.2}$$

denotes the measurement set of the current iteration k. Here, η_k^j is the true jth measured multipath and $\nu_k^j \equiv [w_{k,\varphi}^j w_{k,\tau}^j]^T$ is the jth measurement noise, $w_{k,\varphi}^j \sim \mathcal{N}(0, \sigma_{j,k,\varphi}^2)$, $w_{k,\tau}^j \sim \mathcal{N}(0, \sigma_{j,k,\tau}^2)$. $\sigma_{j,k,\varphi}^2$ and $\sigma_{j,k,\tau}^2$ be the noise variances and $C_k^j \equiv \text{diag}[\sigma_{j,k,\varphi}^2 \sigma_{j,k,\tau}^2]$ defines the covariance matrix.
An element of Z_k is given by

$$z_k^j \equiv \begin{pmatrix} \varphi_k^j \\ \tau_k^j \end{pmatrix}, \tag{6.3}$$

where $\varphi_k^j \in [-\pi, \pi]$, denotes the AoA in radians and $\tau_k^j \in \mathbb{R}^+$ is the RToA of multipath j in iteration k, $j = 1, ..., m_k$. In some numerical evaluations of this chapter the EoA is used additionally. Then, an element of Z_k is given by

$$z_k^j \equiv \begin{pmatrix} \varphi_k^j \\ \vartheta_k^j \\ \tau_k^j \end{pmatrix}, \tag{6.4}$$

where $\vartheta_k^j \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ denotes the EoA of multipath j (in radians) in iteration k. The covariance matrix is then given by $C_k^j \equiv \text{diag}[\sigma_{j,k,\varphi}^2 \sigma_{j,k,\vartheta}^2 \sigma_{j,k,\tau}^2]$.
Without loss of generality it is assumed that the elements of Z_k are stored in ascending order with respect to the RToA, that is, $0 \equiv \tau_k^1 < \tau_k^2 < ... < \tau_k^{m_k}$.
Furthermore, the set of measurements without the first $(t-1)$-incoming multipaths, which is needed in Chapter 7 to formulate the clutter hypothesis, $t \in \{2, ..., m_k\}$, is defined by

$$Z_{k,t} \equiv \{z_{k,t}^j\}_{j=t}^{m_k}, \tag{6.5}$$

where

$$z_{k,t}^j \equiv \begin{pmatrix} \varphi_k^j \\ \tau_k^j - \tau_k^t \end{pmatrix}, \tag{6.6}$$

$j = t, ..., m_k$.
For the rest of this chapter the time index is suppressed for the ease of presentation.

6.1.1.2 Likelihood Function

The probabilistic and assignment based likelihood function defined in [ADKT08b] and [Alg10, chapter 4.4] is presented in this section.
Let $\xi \in X \subseteq \mathbb{R}^2$ be the hypothetical emitter position and let $h_\xi^i,\ i \in \{1, ..., p\}$ be a predicted multipath for a fixed OS location given by the ray tracer, where p is the total number of predicted multipaths. Then, the set of predicted multipaths with respect to an hypothetical emitter position ξ and a fixed OS location is defined by

$$h_\xi \equiv \{h_\xi^i\}_{i=1}^p. \tag{6.7}$$

Possessing a set of measured and predicted multipaths, the possible data interpretations are denoted by $E_{a_1,...,a_p}^m$, where

$$a_i \equiv \begin{cases} 0, & \text{no association, measured multipath} \\ & \text{is not detected} \\ j \in \{1, ..., m\}, & \text{ith predicted multipath is associated} \\ & \text{with measured multipath j} \end{cases} \tag{6.8}$$

is an index for each predicted multipath, which defines its assignment. Let the probability of detection of a specific multipath be denoted by $p_D \in [0, 1]$ and the clutter density by λ_Φ. Let $n \in \{0, ..., \min(m, p)\}$ be the number of assigned multipaths of a specific data interpretaion $E_{a_1,...,a_p}^m$. Then, the probabilistic likelihood function evaluated for a set of measurements Z is defined by

$$p(Z|\xi) \equiv \sum_{E_{a_1,...,a_p}^m} p_D^n \cdot (1 - p_D)^{m-n} \cdot \lambda_\Phi^{p-n} \cdot \prod_{j \in I} \mathcal{N}(h_\xi^j; z^{a_i}, C^{a_i}), \tag{6.9}$$

where $I \equiv \{i \in \{1,...,p\} | a_i \neq 0\}$, that is, the index set of the assigned predicted multipaths and ξ being some hypothetical emitter position in X. Depending on the number of measured and predicted multipaths the computational effort to sum over all data interpretations can be enormous and grows exponentially. Therefore, the likelihood function has to be approximated according to [Alg10, chapter 4.5.2]. First, an optimal assignment for $\min(m,p)$ multipath–assignments is computed. The optimal solution for this assignment problem is given by the Munkres algorithm [BL71]. Then the assigned pairs are sorted with respect to their normalized distances. The corresponding values of the Gaussian distribution are stored in descending order in $\nu \in \mathbb{R}^{\min(m,p)\times 1}$. Then the approximated likelihood function is given by

$$p(Z|\xi) \approx \sum_{n=1}^{\min(m,p)} p_D^n \cdot (1-p_D)^{p-n} \cdot \lambda_\Phi^{m-n} \cdot \prod_{j=1}^{n} \nu_j. \tag{6.10}$$

Note that the likelihood function defined in (6.9) an its approximation (6.10) are defined on sets of measurements. An adaption of the probabilistic likelihood function defined in (6.9) and (6.10) is applied within the generalized PHD Intensity filter in Section 6.3.4.3. However, if the standard PHD intensity filters defined in Section 3.4.2 is applied to BML scenarios, appropriate likelihood functions have to be defined on single multipaths. This is implied by the standard target–oriented measurement model used in standard PHD intensity filters. To this end appropriate observation spaces are studied in the following.

6.1.1.3 Definition of the Observation Space

In the following, two definitions of observation spaces for standard PHD intensity filters are proposed.

Observation Space 1 The first observation space is defined as the subset $Z \subset \mathbb{R}^{T\times 2}$ (in case of processing additionally EoA measurements it is given by $Z \subset \mathbb{R}^{T\times 3}$), where $T \in \mathbb{N}$ represents the number of targets present. Thus, the first observation space is defined by

$$Z \equiv \left\{ \left\{ z^{1,1},...,z^{1,m_1} \right\},...,\left\{ z^{T,1},...,z^{T,m_T} \right\} \right\} \tag{6.11}$$

where the subset $\{z^{t,1},...,z^{m_t}\}$, $t \in \{1,...,T\}$ is the set of multipaths which represents a specific target. The number of received multipaths which belong to target t is specified by $m_t > 0$ (including clutter measurements) .

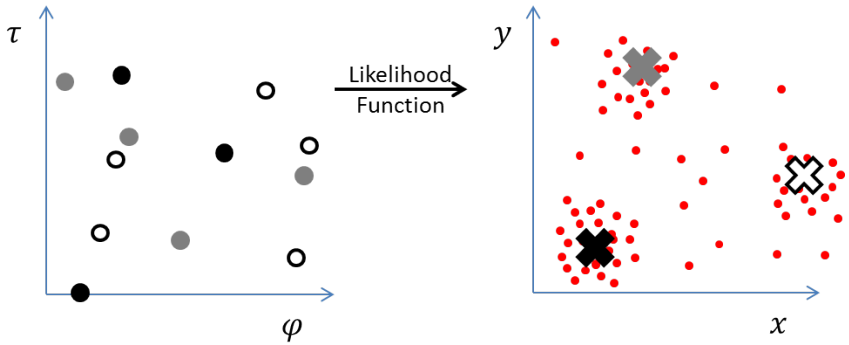

Figure 6.1: Visualization of the first observation space: Since measured multipaths (dots) are assigned to target classes, the ordinary grouping of particles approach yields one target state estimate (crosses) for each target © 2013 IEEE.

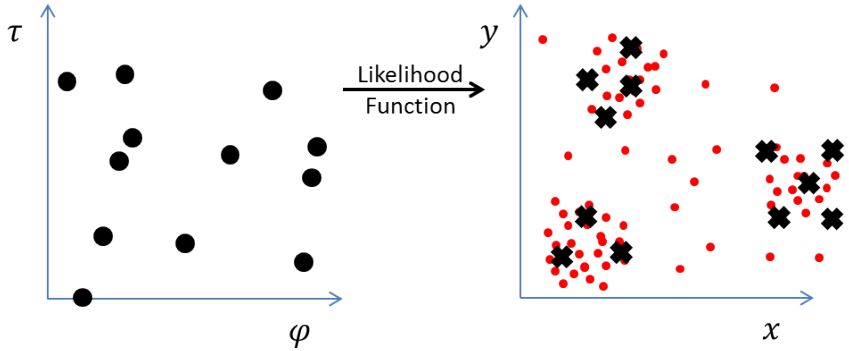

Figure 6.2: Visualization of the second observation space: Each measured multipath is defined as an element of the observation space. The ordinary grouping of particles approach yields for each measurement one estimate if the threshold–test is passed. Therefore, several (suboptimal) target state estimates can be obtained for a single target © 2013 IEEE.

Observation Space 2 The second observation space is defined as the set of received multipaths, that is, $Z \subset \mathbb{R}^{m\times 2}$ ($\mathbb{R}^{m\times 2}$ in case of an additional processing of EoA measurements), where m is the total number of measured multipaths. Therefore, Z is given by

$$Z \equiv \left\{ z^1, ..., z^m \right\}, \tag{6.12}$$

where z^j, $j \in \{1, ..., m\}$ denotes a measured multipath.

6.1.1.4 Discussion of Observation Space Definition and the Problem of Target State Estimation

The main difference between the two observation space definitions is given by the assignment of multipaths to target–classes. For Definition (6.11) an additional pre–processing scheme is required to assign multipaths to targets. Assume such a pre–processing algorithm is available. Then, the derived measurement set contains a one–to–one correspondence between the observation and the target space (see Figure 6.1) as it is assumed in the SMC–implementations from [RCV10] and [SKSC12]. Therefore, a target state estimation using the conventional grouping of particles approach is convenient and yields a target state estimate for a set of multipaths which belongs to a target. In this case, the approximation of the probabilistic likelihood function (6.10) can be used for the update. Hence, the only adaption to be made for applying standard PHD intensity filters using the first observation space to BML is to derive an appropriate pre–processing algorithm for the multipath to target association. However, multipaths which are generated by the same target are not spatially related in the measurement space (as for example in extended target tracking), making such an assignment impossible. Due to this fact the first observation space seems to be inappropriate for using it within a standard PHD intensity filter. In Section 3.4.4 a closely related observation space definition will be discussed and it will be seen that such a generalization of the standard observation space is numerically expensive. However, in a single–target scenario with perfect detection and no false alarms, observation space one yields a possibility to easily apply a standard PHD intensity filter to BML and can thus be considered as a benchmark for the proposed estimation approaches from Section 6.1.3.

The second observation space does not presume an additional assignment and defines each measured multipath as a single measurement. This definition can be seen as an iterative approach of information processing. Since an MS location is represented by a set of multipaths, a single multipath contributes only a part of the complete measurement information to the filter by adjusting the weights via an appropriate likelihood function (see Figure 6.2). Due to the need of updating each multipath separately when choosing the second observation space the following definition of the likelihood function is used for a single multipath.

$$p(z^j|\xi) \equiv \begin{cases} \mathcal{N}(h_\xi^i; z^{a_i}, C^{a_i}), & \text{if } \exists i \in \{1, ..., p\} \\ & \text{such that } a_i = j \\ 0, & \text{otherwise} \end{cases} \tag{6.13}$$

for each $z^j \in Z, \quad j \in \{1, ..., m\}$. The (global) assignment is done via Munkres algorithm [BL71] between the set of measured multipaths Z and the set of predicted multipaths h_ξ of $\xi \in \mathbb{R}^2$. Therefore, the index a_i denotes the assigned measured multipath of the ith predicted multipath from h_ξ analogously to Definition (6.8). Thus, iteratively updating using (6.13) with all elements from Z yields the unlabeled particle set. Due to processing the unsorted multipaths there is no information available for a PHD intensity filter about the relation between the measurements and therefore the ordinary grouping of particles approach yields a target state estimate for each measured multipath. Hence, for a single target several target state estimates might exist. Furthermore, these estimates are suboptimal, since an MS location is represented by a set of multipaths and thus one element of this set does not provide the complete available information about a specific target. To estimate the target state optimally a generalization of the grouping of particles approach is developed in the following section.
Note that the definition of the observation space also influences quantities of a PHD intensity filter. The mean number of targets, which is given by the integral of the intensity over the target state space, corresponds for the first observation space to the mean number of existing targets, where for the second observation space it is given by the mean number of multipaths that correspond to existing targets.

6.1.2 Target State Estimation Using Generalized Particle Grouping

Since the second observation space is used from now on to adapt the standard PHD intensity filter to the challenge of BML, a post–processing of the particle set for the target state extraction is inevitable to determine a single target state estimate based on a subset of multipaths. Therefore, the generalization of the grouping of particles methodology is presented. First, the grouping of particles weights are considered exemplary in terms of the SMC–iFilter presented in [SKSC12]. The corresponding weights using the SMC–PHD filter can be found in [RCV10].
Let $Z \equiv \{z^j\}_{j=1}^m$ be a set of multipath measurements and $\{x_i, w_i\}_{i=1}^N$ be the particle set of the current iteration. Then,

$$w_{j,i} \equiv \frac{p(z^j|x_i)p^D(x_i)}{\lambda(z^j)} \cdot w_i \tag{6.14}$$

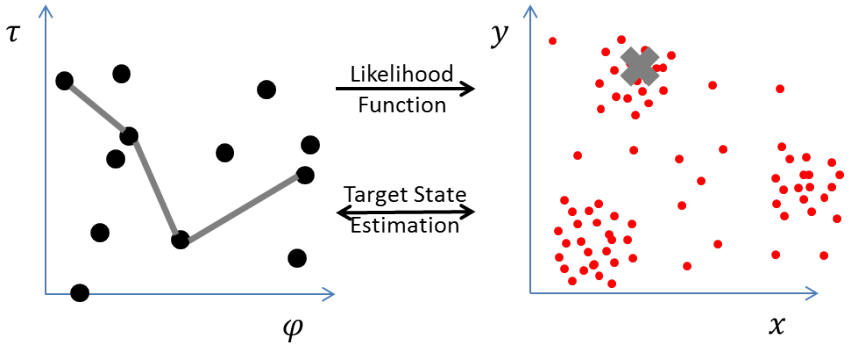

Figure 6.3: Visualization of the generalized grouping of particles approach: By investigating subsets of multipath measurements (connected points) and evaluating the respective grouping weights a probability of target existence can be computed according to [RCV10]. If this exceeds a specific threshold a target estimate (gray cross) is computed for the subset © 2013 IEEE.

is called the particle grouping weight of ith particle and the jth measurement. Here, $p(z^j|x_i)$ is given by the single multipath likelihood function defined in (6.13). $\lambda(z^j)$ denotes the predicted measurement intensity of z^j and weights the respective measurement. $p^D(x_i)$ denotes the detection probability of target x_i. Afterwards,

$$W_j \equiv \sum_{i=1}^{N} w_{j,i} \tag{6.15}$$

is defined. Due to the definition of the predicted measurement intensity, W_j is bounded by one. For measurement j this can be interpreted as a probability of existence. After a threshold–test is passed an estimate for measurement j is determined by

$$\frac{1}{\sum_{i=1}^{N} w_{j,i}} \sum_{i}^{N} w_{j,i} x_i. \tag{6.16}$$

Within an SMC–version of an arbitrary PHD intensity filter that uses the second observation space, this estimation procedure is done for each multipath and therefore might yield several sub–optimal target state estimates for one target (depending on the number of passed threshold–tests). Therefore a generalization of this scheme to sets of measurements for an arbitrary SMC–PHD intensity filter is derived in the following.

Since in BML a target is represented by a set of measurements the aim is not only to determine the probability that particle x_i occurred due to measurement z^j as it

is done in the original grouping of particles presented in [RCV10], but to compute the probability that particle x_i is created due to a specific subset of measurements $\{z^1, ..., z^L\} \subset Z,\ 0 < L \leq m$. This idea is similar to the approach proposed in [Alg10, chapter 7.2.2], where within an MHT algorithm subsets of measurements are assessed and used to estimate the target state. This is visualized in Figure 6.3, where a subset of multipaths yields one estimate for a target. After knowing how to compute this probability it is easy to assess the different subsets and develop procedures for the BML task that offer one state estimate for one target. First consider $p(x_i|z^1, ..., z^L)$: If Bayes Theorem is applied twice and it is assumed that the elements of $\{z^j\}_{j=1}^{L}$, $0 < L \leq m$, are statistically independent the following holds

$$p(x_i|z^1, ..., z^L) = \frac{p(z^1|x_i) \cdot p(x_i|z^2, ..., z^L)}{p(z^1, z^2, ..., z^L)} = \cdots = \frac{\prod_{j=1}^{L} p(z^j|x_i) \cdot p(x_i)}{p(z^1, z^2, ..., z^L)}. \tag{6.17}$$

Having (6.17), the generalized grouping of particles weights can be computed for each subset of the measurement set. But before considering two approaches that utilize (6.17), the generalized grouping of particles weight for a particle x_i and a set of measurements is defined exemplary in terms of the iFilter. First, let Z be the current measurement set and consider analogously to [Alg10, chapter 7.2.2] the set of subsets of the set of measured multipaths, that is,

$$\mathcal{J} \equiv \{J \subset Z\}. \tag{6.18}$$

Due to the binomial series the number of subsets is given by

$$\sum_{i=0}^{|Z|} \binom{|Z|}{i} = 2^{|Z|}, \tag{6.19}$$

where $|\cdot| : \mathcal{J} \to \mathbb{N}$ determines the cardinality of an element from $\mathcal{J}$. Hence the number of subsets of Z grows exponentially with the number of elements of Z.
For an arbitrary subset $J \in \mathcal{J}$ and particle x_i, $i \in \{1, ..., N\}$, where N is the number of persistent particles (present in the previous iteration, that is, not newborn in the current iteration, see [SKSC12]), the generalized grouping of particles weight is defined, following (6.17), by

$$w_{J,i} \equiv \frac{\prod_{j=1}^{|J|} p(z^j|x_i) \cdot p^D(x_i) \cdot w_i}{\lambda(z^J)}, \tag{6.20}$$

where $\lambda(z^J)$ is the predicted measurement intensity of the subset $J \in \mathcal{J}$, that is,

$$\lambda(z^J) = \sum_{i=1}^{N} \prod_{j=1}^{|J|} p(z^j|x_i) \cdot p^D(x_i) \cdot w_i + p(z^J|\phi) \cdot p^D(\phi) \cdot f(\phi). \tag{6.21}$$

Here, $p^D(x_i)$ is the detection probability and w_i the weight of particle x_i, $p^D(\phi)$ is the detection probability and $f(\phi)$ is the intensity in X_ϕ, which is the hypothesis space of the iFilter (X_ϕ is assumed to contain only a constant analogously to [SKSC12]). Note, that the proposed scheme can also be used within an SMC–PHD implementation (compare to [RCV10, Equation (25)]) for all $j \in \{1, ..., |J|\}$.
Obviously, the term *weight* for $w_{J,i}$, defined in (6.20), makes sense since

$$0 \leq \sum_{i=1}^{N} w_{J,i} \leq 1. \tag{6.22}$$

The first inequality of (6.22) is due to the definition of $w_{J,i}$. The second inequality of (6.22) is justified by the fact that $p(z^J|\phi) \cdot p^D(\phi) \cdot f(\phi) \geq 0$ and thus

$$\sum_{i=1}^{N} w_{J,i} \equiv \sum_{i=1}^{N} \frac{\prod_{j=1}^{|J|} p(z^j|x_i) \cdot p^D(x_i) \cdot w_i}{\lambda(z^J)} \leq \frac{\sum_{i=1}^{N} \prod_{j=1}^{|J|} p(z^j|x_i) \cdot p^D(x_i) \cdot w_i}{\sum_{i=1}^{N} \prod_{j=1}^{|J|} p(z^j|x_i) \cdot p^D(x_i) \cdot w_i} = 1. \tag{6.23}$$

In the following, two approaches for the target state estimation in a single target scenario are presented that utilize the generalized grouping of particles equations in a similar way as it is done in [SKSC12] and [RCV10].

6.1.3 Methods for State Extraction

In this section, approaches for the target state estimation in a single–target scenario are presented. Additionally to a standard weighted mean computation two enhanced algorithms, that utilize the generalized grouping of particles equations from the previous section are presented. These state extraction algorithms work in a similar way as the algorithms from [SKSC12] and [RCV10] which use the ordinary grouping of particle approach.

6.1.3.1 Mean Computation

Consider a single–target scenario. The easiest method to do a state extraction is to omit the grouping of particles step from [SKSC12] and to perform a simple mean value computation of the particle set instead. To compute the desired estimate, the normalized weights after the update of the intensity function are used to determine

$$\frac{1}{W} \sum_{i=1}^{N} w_i \cdot x_i, \tag{6.24}$$

where N is the number of particles from the previous iteration (without the newborn particles from the current iteration), $W \equiv \sum_{i=1}^{N} w_i$ and w_i are the weights of the particles $x_i,\ i \in 1, ..., N$.

6.1.3.2 Generalized Mean Computation

When considering a single–target scenario it is implicitly assumed that the obtained set of measured multipaths does not bear any confusion in target assignment, which means that only one subset of multipaths represents the present target. An approach, which has a low computational complexity is to consider only one subset of the measurement set, that is, the set itself. Therefore $w_{Z,i}$ is investigated for each particle x_i. After computing each grouping of particles weight using (6.20)

$$W_Z \equiv \sum_{i=1}^{N} w_{Z,i} \tag{6.25}$$

is determined, which can be interpreted as a probability of existence for the target based on the set of received multipaths (see also [SKSC12]). If $W_Z > \tau$ holds, where $\tau \in [0, 1]$ is a threshold for target existence, the target state estimate is given by

$$\frac{1}{\sum_{i=1}^{N} w_{Z,i}} \sum_{i}^{N} w_{Z,i} \cdot x_i. \tag{6.26}$$

6.1.3.3 Generalized Ranking Estimation

The second generalized grouping of particles approach presented restricts itself not only to the whole measurement set but evaluates the grouping of particles weights from (6.20) for several subsets. Note that due to the exponential growth of the number of subsets with the number of measurements, a consideration of all subsets is computational feasible only for a restricted number of multipaths. To filter out measurements that occur due to clutter the original grouping of particles scheme from [RCV10] is evaluated first. For each measured multipath a probability of existence is determined according to (6.14) and (6.15). Then,

$$\tilde{Z} \equiv \{z^j \in Z | W_j > \tau\}, \tag{6.27}$$

is defined, where $\tau \in [0, 1]$. Note, that a determination of a target state estimate could now be done for each measured multipath without any extra computational burden. Afterwards, the measurements are sorted with respect to to their probability of existence and saved in descending order in $\tilde{Z}_{\text{sorted}}$. The r–best, $0 < r \leq m$, are taken out and all possible subsets except for the empty set and the subsets which only contain one element (the corresponding estimates would be the estimates generated by the ordinary grouping of particles method) are created , that is, for all

$$J \in \mathcal{J} \equiv \left\{ J \subset \left\{ \tilde{z}^1, ..., \tilde{z}^r \right\} \middle| |J| \neq 0, 1 \right\} \tag{6.28}$$

$w_{J,i}$ is computed. Then,

$$\mathcal{J}_{\text{fused}} \equiv \left\{ J \subset \left\{ \tilde{z}^1, ..., \tilde{z}^r \right\} \middle| W_J = \sum_{i=1}^{N} w_{J,i} > \tau_{\text{fused}} \right\} \tag{6.29}$$

is defined, where $\tau_{\text{fused}} \in [0, 1]$. Finally, the element of $\mathcal{J}_{\text{fused}}$ with the highest probability of existence represents the target state estimate for the single–target scenario. Note that the proposed approach is an approximation of the best representing subset, since only the r–best (in terms of the probability of existence for the single multipaths) measurements are used for the estimation. An alternative approximation would be to restrict the cardinality of the considered subsets (see [Alg10, chapter 7.2.3]). For both approximations better results are expected in a single–target scenario without clutter and measurement failure when applying the generalized mean computation from Section 6.1.3.2 compared to the generalized ranking. This is due to the correct measurement being not known a priori (no clutter, no mis detection). However, the generalized ranking approach bears several advantages compared to the generalized mean methodology. First, it is obvious that in a real world scenario clutter and measurement failure are inevitable. Thus, the generalized mean approach incorporates false measurements into its estimation. The generalized ranking on the contrary does not assume to "know" the correct measurement subset, but it determines the best measurement representation for the target by computing the probabilities of existence for a specific number of subsets (the considered number depends on the computational capacity and determines how accurate the approximation of the correct subset is). Hence, it is expected that the generalized ranking approach is more stable and more accurate under the influence of clutter than the generalized mean approach. Furthermore, the generalized ranking approach offers the possibility to extend the proposed PHD intensity filter adaption to a multitarget scenario (see Section 6.1.5).

6.1.4 Numerical Evaluation

6.1.4.1 Comparison of State Extraction Methods

To verify the applicability of the proposed methods a single–target scenario is considered in an urban environment. An SMC–implementation of the iFilter from [SKSC12] is used as a representative of the class of PHD intensity filters. First, a database for a fixed OS position and a grid of mobile station locations is generated, using a ray tracing simulation for a specific city–map (see Figure 5.2). Afterwards, a linear ground truth for a target with constant velocity is created. For each time step the lower–left grid point of the box, in which the target is located in, is determined and the corresponding multipaths from the database are stored as the measurements of the respective iteration. For this evaluation neither clutter nor measurement noise is added to the measurement set and perfect detection of the single multipaths is assumed.

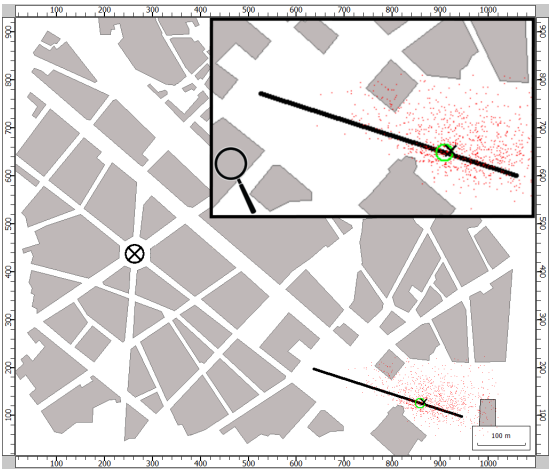

Figure 6.4: Visualization of the simulation scenario: The fixed observer (black circled cross) receives multipath measurements of a single target (ground truth visualized as black solid line) under perfect measurement conditions (no clutter, no measurement noise, perfect detection). The particles are visualized as red dots, the current exact position of the target as green circle and the estimate of the filter as black cross © 2013 IEEE.

In the following the target is moving linearly with constant velocity $v = 2.4\frac{m}{s}$ (see the ground truth in Figure 6.4). The measurement noise in AoA is set to $\sigma_\varphi = 10^\circ$ and in RToA to $\sigma_\tau = \frac{10\text{m}}{c_{\text{light}}}$, where c_{light} denotes electromagnetic propagation. For each iteration and each multipath the same covariance matrix $C \equiv \text{diag}\begin{bmatrix}\sigma_\varphi^2 & \sigma_\tau^2\end{bmatrix}$ is used, which justifies that the index of the standard deviations from Section 6.1.1.1 is omitted. For the SMC–implementation of the iFilter the following detection probabilities are defined. The detection probability in the target state space $X \subseteq \mathbb{R}^2$ is set to $p^D(x_i) = 0.75$ for each particle $x_i \in X$, and the detection probability in X_ϕ is defined by $p^D(\phi) \equiv 0.3$. The transition probability from X_ϕ to X is set to $\Psi(x|\phi) = 0.2$, the transition probability in X_ϕ is defined as $\Psi(\phi|\phi) = 0.2$ and the transition probability from X to X_ϕ is given by $\Psi(\phi|x) = 0.1$. The thresholds for target existence of the ordinary and the generalized grouping of particles are set to 0. This is done to make the generalized grouping of particles approaches comparable to the simple mean computation which is implemented without any threshold criterion for target existence. Furthermore, the maximal number of particles is restricted to 1500 and new particles are created uniformly over the considered city–map, where the number of newborn particles is determined according to [SKSC12].
To assess the different approaches with respect to accuracy 100 Monte–Carlo (MC) runs of the presented scenario are performed. The results are shown in Figures 6.5

and 6.6 in terms of the position and velocity root mean squared error (RMSE). For the investigated scenario, the ordinary mean computation can be seen as a benchmark for the proposed generalized grouping of particles methodologies, since the measurement conditions are perfect, that is, no measurement noise, no clutter and no mis detection imply that the received set of measurements perfectly represents the target. Therefore, the functionality of the proposed algorithms for state extraction can be verified by comparing the RMSE values. For most of the times the best result in terms of accuracy is given by the generalized mean computation approach, which computes the generalized grouping of particles weights for the whole measurement set. Analogously to the ordinary mean computation the generalized mean computation profits from the a priori knowledge that the received set of multipaths is correct. Any approximation of this set by leaving out specific multipaths deteriorates the result. Due to the approach of only taking the r–best measurements into account (in this simulation $r = 5$, the maximal number of received multipaths is 20) the set of measurements for the target state estimation is permanently weakened in terms of missing true multipaths. Therefore, it is remarkable that even though the set of measurements is reduced, the approach of generalized ranking is capable to yield almost the same (and for some iterations an even better) accuracy as the ordinary mean computation, that does not suffer from crossing out true measurements. This capability is mainly due to the individual weighting of measurements and sets of multipaths by the predicted measurement intensity in the procedure. This is also the reason why the generalized mean computation yields the best result of the compared methodologies. However, it is expectable that the influence of clutter and mis detection yields a completely different result. Due to the assessment of measurements in the ordinary grouping of particles step, clutter measurements should be provided with lower probabilities of existence than the true measurements. Therefore, the generalized ranking approach favors measurements which belong to true targets towards clutter measurements. This improves the performance of the generalized ranking methodology compared to the generalized mean computation and the ordinary mean computation in a cluttered single–target scenario with measurement failure and measurement noise.

6.1.5 Conclusion

In this section the application of PHD intensity filters to the problem of BML using an SMC–implementation is presented. First, two observation space definitions are discussed. Due to the absence of an appropriate pre–processing procedure to assign each multipath to the correct target, the second observation space is used, which defines a measurement to be a multipath. The definition of a measurement as a single multipath violates the *at most one measurement per target*–assumption of standard PHD intensity filters. Therefore, an application of the ordinary grouping of particles scheme presented in [SKSC12] and [RCV10] does not yield a single estimate

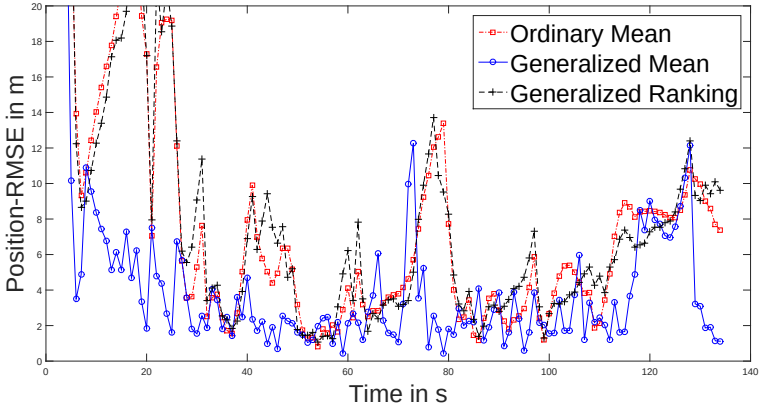

Figure 6.5: Comparison of the Position–RMSE © 2013 IEEE.

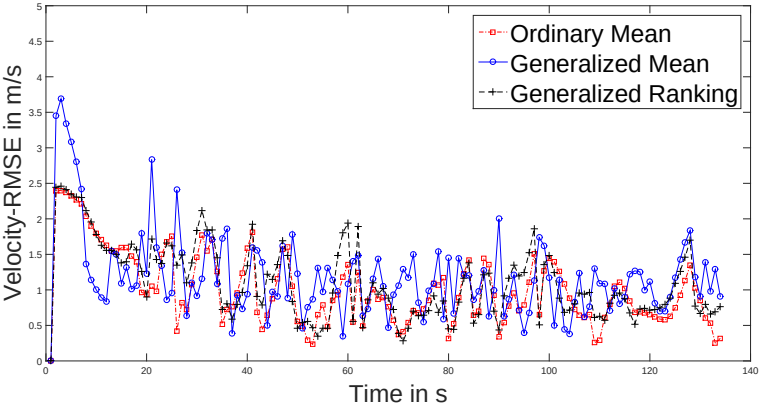

Figure 6.6: Comparison of the Velocity–RMSE © 2013 IEEE.

for the target, but one estimate for each multipath. Thus, a generalization of the grouping of particles scheme for sets of measurements is presented, resulting in the generalized grouping of particles weights. Possessing the definition of these weights two approaches for a single–target scenario are formulated and compared in a clutter– and noise–free scenario with perfect detection to a standard mean computation approach. The results of the numerical evaluation show that the proposed approaches yields comparable (for the generalized ranking) and most of the time even more accurate results (for the generalized mean computation) which motivates the application of standard PHD intensity filters to more complex BML–scenarios.

6.2 SMC–Intensity Filter Using a Decomposition of a Likelihood Function

In [SW09] a general likelihood decomposition is presented. By applying a multivariate version of the chain rule, Maclaurin (Taylor)–series approximation and marginalization, the conventional non–linear likelihood function is approximated by an integral over a manifold of Gaussian kernels, which are linear in target space. Furthermore, several applications are shown. For this work the application of the general likelihood decomposition to the range–bearing scenario, presented in [SW09, chapter 3], is of interest. Since BML considers RToAs only, the range is not observable and thus a marginalization over the range as in [SW09, Equation (20)] represents an appropriate approach to define a decomposed likelihood function for BML and an alternative to the assignment–based likelihood function given in (6.13).
In this section the technique of likelihood decomposition is applied to BML by modifying the likelihood function presented in [SW09, Equation (20)]. To this end, multipaths are divided into straights in between the points of deflection of the signal, which are called interaction points. Afterwards, the decomposed likelihood function for BML is defined by marginalization over the set of multipaths which belong to a specific measurement and summing over all linear connections of one multipath. Furthermore, implementation issues like the drawing of particles along the multipaths, the approximation of the decomposed likelihood function and the determination of the correct multipaths which belong to a specific AoA measurement are discussed. Finally, a numerical evaluation is carried out, using a simulation of the estimation process of a real antenna array and a ray tracer to generate measurements. Therefore, the proposed approach is compared to the likelihood function, which is based on the assignment approach presented in the previous section and defined in (6.13), within an SMC–implementation of the iFilter.
This section is structured as follows. Section 6.2.1 describes the problem formulation. The decomposition of a likelihood function from [SW09] is described in Section 6.2.1.1. Section 6.2.1.2 defines the approximation of the likelihood function for the

range–bearing scenario, given only the bearing measurement, analogously to [SW09, chapter 3] and Section 6.2.1.3 formulates the problem of defining the measurement function for a BML scenario. In Section 6.2.2 the concept of likelihood decomposition is applied to BML. First, Section 6.2.2.1 yields the adapted decomposition for a BML scenario. Implementation issues are discussed in Section 6.2.2.2. A numerical evaluation comparing the assignment–based likelihood function defined in (6.13) and the proposed decomposed likelihood function is presented in Section 6.2.3 before the conclusions are drawn in Section 6.2.4.

6.2.1 Formulation of the Problem

6.2.1.1 Likelihood Function Decomposition

In this section a likelihood function decomposition is derived following the considerations of [SW09]. Therefore, let $g : \mathbb{R}^n \to \mathbb{R}^n$, $n > 0$ denote a continuously differentiable bijective function which maps the target to the measurement space and let $f : \mathbb{R}^n \to \mathbb{R}^n$ be its inverse. Measurement errors are assumed to be Gaussian distributed. If the measurement received $z \in \mathbb{R}^n$ has n dimensions, that is, if the sensor provides information of all measurement coordinates, the likelihood function for the target state $x \in \mathbb{R}^n$ is given by

$$p(z|x) = \mathcal{N}(z; g(x), \Sigma), \tag{6.30}$$

where $\mathcal{N}(z; g(x), \Sigma)$ denotes the multivariate Gaussian PDF on $\mathbb{R}^n$ evaluated at z with mean $g(x)$ and covariance matrix Σ.
In the following the approximation of the likelihood function is derived analogously to [SW09]. To this end, let $h : \mathbb{R}^n \to \mathbb{R}^n$ be a continuously differentiable function. The Jacobian of h evaluated at $x \in \mathbb{R}^n$ is denoted by

$$J[h](x) \equiv \begin{bmatrix} \frac{\partial h_1(x)}{\partial x_1} & \frac{\partial h_1(x)}{\partial x_2} & \cdots & \frac{\partial h_1(x)}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_n(x)}{\partial x_1} & \frac{\partial h_n(x)}{\partial x_2} & \cdots & \frac{\partial h_n(x)}{\partial x_n} \end{bmatrix}. \tag{6.31}$$

Then, the Taylor series expansion of h up to the first term around $a \in \mathbb{R}^n$ is given by

$$h(x) = h(a) + J[h](a)(x - a) + \mathcal{O}((x - a)^2). \tag{6.32}$$

Let $x \in \mathbb{R}^n$ be an element of the target and $z \in \mathbb{R}^n$ be an element of the measurement space. Then, using that f is the inverse of g (and vice versa) and inserting the Taylor series expansion of g yields

$$z - g(x) = g(f(z)) - g(x) \approx \{J[g](f(z))\}(f(z) - x) \tag{6.33}$$

$$= \{J[f](z)\}^{-1}(f(z) - x). \tag{6.34}$$

Here, the approximation of the right hand side of (6.33) is due to the truncation of the Taylor series after the linear term and the generalized chain rule implies (6.34). Therefore, the desired approximation of the likelihood function (which is linear in the target space) is given by

$$\begin{aligned} p(z|x) =& N(z|g(x), \Sigma) && (6.35)\\ =& (2\pi)^{-\frac{n}{2}} |\Sigma|^{-\frac{1}{2}} \exp(-\frac{1}{2}(z-g(x))^T \Sigma^{-1}(z-g(x))) && (6.36)\\ \approx& (2\pi)^{-\frac{n}{2}} |\Sigma|^{-\frac{1}{2}} \exp\big(-\frac{1}{2}(f(z)-x)^T \\ & (\{J[f](z)\}^{-1})^T \, \Sigma^{-1} \{J[f](z)\}^{-1}(f(z)-x)\big) && (6.37)\\ =& |J[f](z)| \, \mathcal{N}(f(z); x, \tilde{\Sigma}(z)), && (6.38) \end{aligned}$$

where $\tilde{\Sigma} \equiv J[f](z)\Sigma J[f](z)^T$, since for $A, B \in \mathbb{R}^{m\times m}$, $m \in \mathbb{N}$ it holds $(A^{-1})^T = (A^T)^{-1}$ and $(AB)^{-1} = B^{-1}A^{-1}$. The approximation from (6.34) implies (6.37) and (6.38) holds due to $|\tilde{\Sigma}(z)| = |\Sigma||J[f](z)|^2$.

6.2.1.2 Likelihood Function Decomposition for Bearing–Only Measurements

Possessing the likelihood approximation (6.38) an application to bearing only measurements can be done (see [SW09]). Let the target and measurement space be given by $\mathbb{R}^2$. Furthermore, define $g, f : \mathbb{R}^2 \to \mathbb{R}^2$ by

$$f(r,\varphi) \equiv \big(r\cos(\varphi), r\sin(\varphi)\big) \tag{6.39}$$

$$g(x,y) \equiv \Big(\sqrt{(x^2+y^2)}, \arctan(y/x)\Big). \tag{6.40}$$

If the sensor provides only a bearing measurement and the range is not observable (bearing–only scenario) a marginalization with respect to the range–coordinate has to be applied. Then, the decomposed likelihood function for bearing–only measurements is given by

$$\mathcal{N}(\varphi; \arctan(y/x), \sigma_\varphi^2) \approx \int_0^\infty r\mathcal{N}\left(\begin{pmatrix} r\cos(\varphi) \\ r\sin(\varphi)\end{pmatrix}; \begin{pmatrix} x \\ y\end{pmatrix}, J[f](r,\varphi)\tilde{\Sigma}J^T[f](r,\varphi)\right) dr, \tag{6.41}$$

where

$$J[f](r,\varphi) \equiv R_f(r,\varphi)D(r,\varphi), \tag{6.42}$$

$$R_f(r,\varphi) \equiv \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi)\end{pmatrix} \tag{6.43}$$

and $D_f(r,\varphi) \equiv \mathrm{Diag}[1, r]$.

6.2.1.3 Problem of Emitter Localization assessing Multipath Measurement with a generalized Likelihood Function Decomposition

The measurement space of BML defined in Section 6.1.1.1 contains AoA and RToA measurements (in case of an additional EoA measurement the following considerations can be easily generalized). The fact that the absolute path length of a multipath is not observable in BML scenarios motivates the application of the likelihood function decomposition from Section 6.2.1.2. A key difference between the standard bearing–only scenario and BML is that the formulation of the propagation of the electromagnetic wave between OS and MS, that is, the definition of the measurement function f from Equation (6.39). In the assignment–based approach for BML [DGK13a], [Alg10] a ray tracer is used to evaluate f pointwisely with respect to the urban environment. This is due to the fact that the ray tracing prediction is highly complex (physical propagation effects like scattering, diffraction and fading are modeled additionally to the reflection) and thus it might be hard or even impossible to invert f. Furthermore, when using a black box for ray tracing f is unknown. Therefore, we assume in this thesis that the inverse of f is in general not given. In the following section, a likelihood function for BML is defined using the considerations from Section 6.2.1.2, Section 6.2.1.1 and an additional approximation.

6.2.2 Likelihood decomposition for multipath measurements

6.2.2.1 Derivation of the Decomposed Likelihood Function for BML

In the following, the output of the ray tracer for a fixed OS location, based on the database of the urban environment and specific radio channel parameters is considered. The ray tracing prediction is modeled according to [HWLW03], which is based on [HWLW99], [WHL99], [HWLW03] and the fundamentals given in [Gla89] For a given AoA $\varphi \in [-\pi, \pi]$ the set of multipaths predicted by the ray tracer is defined by

$$\mathcal{M}_\varphi \equiv \{f_j : [0, \infty) \to \mathbb{R}^2 | j \in \{1, ..., p\}\}, \tag{6.44}$$

where p denotes the number of predicted multipaths for the particular AoA φ and f_j defines a specific multipath, given as a function of the range. Several multipaths which correspond to one AoA can occur due to resolution conflicts and the possibility that an interaction point is modeled as a radially emitting point source. In the following, an arbitrary $j \in \{1, ..., p\}$ is chosen and one multipath f_j is considered. It is assumed that f_j is continuously differentiable and bijective on its image. Let therefore denote g_j the inverse of f_j, that is , $g_j \equiv f_j^{-1}$. Thus, if f_j and g_j are given, Equation (6.38) can be used to define the respective likelihood function.
Unfortunately as mentioned before, an explicit definition of f_j cannot be assumed in general. Therefore, neither its Jacobian nor its inverse are known and hence (6.38) cannot be evaluated as in (6.41) for the bearings–only case. To obtain a general

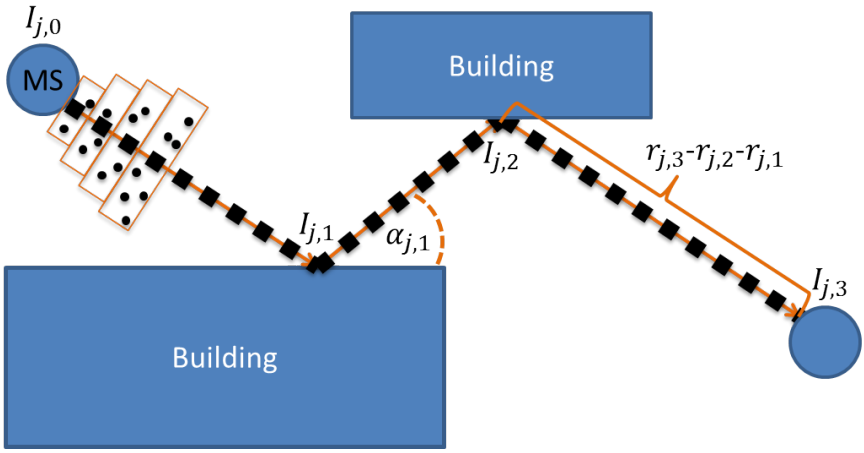

Figure 6.7: Each multipath between the OS and the MS is assumed to propagate linearly in between its interaction points. Boxes around grid points, which are proportional to the covariance matrix of the decomposed likelihood function limit the allowed region of particles (black dots) for each multipath © 2014 IEEE.

decomposed likelihood function, which is linear in target space according to [SW09], an additional approximation has to be made. Therefore, instead of considering the explicit form of f_j the output of the ray tracer is considered. Additionally to the AoA and RToA for a given MS–OS combination, the ray tracer provides the points of interaction for each multipath.

Therefore, let

$$I_{j,0}, ..., I_{j,N} \in \mathbb{R}^2 \tag{6.45}$$

be the interaction points of the multipath f_j, where $I_{j,0}$ denotes the position of the OS and $I_{j,N}$ denotes the position of the MS. Then, there exist

$$r_{j,0}, ..., r_{j,N} > 0 \tag{6.46}$$

such that

$$f_j(r_{j,i}) = I_{j,i} \tag{6.47}$$

for all $i \in \{0, ..., N\}$ (see Figure 6.7). Note, that $r_{j,0} = 0$ by definition. Furthermore, let

$$\alpha_{j,0} \equiv \varphi \tag{6.48}$$

$$\alpha_{j,i} \equiv \arctan\left(\frac{I_{j,i+1,y} - I_{j,i,y}}{I_{j,i+1,x} - I_{j,i,x}}\right), \tag{6.49}$$

$i = 0, ..., N-1$ be the angles defined by the multipath at the interaction points. Then for $i \in \{0, ...N-1\}$ the function $f_{j,i} : [0, r_{j,i+1} - r_{j,i}] \to \mathbb{R}^2$ is defined by

$$f_{j,i}(r) \equiv I_{j,i} + \begin{pmatrix} r\cos(\alpha_{j,i}) \\ r\sin(\alpha_{j,i}) \end{pmatrix} \tag{6.50}$$

and

$$J[f_{j,i}](r) \equiv \begin{pmatrix} \cos(\alpha_{j,i}) & -r\ \sin(\alpha_{j,i}) \\ \sin(\alpha_{j,i}) & r\ \cos(\alpha_{j,i}) \end{pmatrix}, \tag{6.51}$$

denotes the corresponding Jacobian. Note that the set $\{f_{j,1}, ..., f_{j,N-1}\}$ approximates the ray tracing prediction function by assuming a linear propagation of the radio channel in between the interaction points. Hence, the likelihood function corresponding to $f_j \in \mathcal{M}_\varphi$ and a target state $(x, y)^T \in \mathbb{R}^2$ can be approximated by

$$p(\varphi|j, (x, y)^T) \approx \sum_{i=0}^{N-1} \int_0^{r_{i+1}-r_i} rN\left(I_{j,i} + \begin{pmatrix} r\cos(\alpha_{j,i}) \\ r\sin(\alpha_{j,i}) \end{pmatrix} \middle| \begin{pmatrix} x \\ y \end{pmatrix}, J[f_{j,i}](r)\Sigma J^T[f_{j,i}](r) \right) dr, \tag{6.52}$$

where the event j denotes that the multipath defined by f_j was chosen. Due to marginalization with respect to the different multipaths belonging to a specific AoA $\varphi \in [-\pi, \pi]$ the likelihood function of the bearing measurement φ given the hypothetical position of the MS $(x, y)^T \in \mathbb{R}^2$ is given by

$$p(\varphi|(x, y)^T) \equiv \sum_{j=1}^{p} p(\varphi|j, (x, y)^T) \approx \sum_{j=1}^{p} \sum_{i=0}^{N-1} \int_0^{r_{j,i+1}-r_{j,i}} r\mathcal{N}\left(I_{j,i} + \begin{pmatrix} r\cos(\alpha_{j,i}) \\ r\sin(\alpha_{j,i}) \end{pmatrix} \middle| \begin{pmatrix} x \\ y \end{pmatrix}, J[f_{j,i}](r)\Sigma J^T[f_{j,i}](r) \right) dr. \tag{6.53}$$

6.2.2.2 Implementation Issues

Specific issues concerning the implementation of the proposed likelihood decomposition are explained in the following.

Determination of Multipaths To evaluate the likelihood function defined in (6.53), multipaths have to be determined for a particular AoA measurement $\varphi \in [-\pi, \pi]$. Therefore, the ray tracer prediction is used. First, by computing the received multipaths for all possible MS positions (on a grid–cell) and a fixed OS, the database of all possible multipaths in the field of view is generated and sorted with respect to the

AoA. Each entry of the database represents a multipath. Additionally to the AoA and the RToA (with respect to its MS position) also the interaction points characterize a specific multipath.
Given the database of all predicted multipaths the paths corresponding to a specific AoA measurement φ need to be determined for the evaluation of the likelihood function defined in (6.53). Since each interaction point emits radially an electromagnetic wave, it is possible that several multipaths belong to the same AoA measurement φ. Depending, on the geometry this fact implies that it might become impossible to obtain a reasonable tracking result, due to a large number of falsely chosen multipaths. Furthermore, the computational effort increases dramatically if a large number of predicted multipaths belongs to the same AoA measurement, since for each path the integral from Equation (6.53) has to be evaluated. Thus, the RToA is utilized to reduce the number of predicted multipaths for a specific AoA measurement. Therefore, the likelihood function from (6.53) is approximated by

$$p((\varphi,\tau)^T|(x,y)^T) \approx \sum_{j=1}^{\tilde{p}} \sum_{i=0}^{N-1} \int_0^{r_{j,i+1}-r_{j,i}} rN\left(I_{j,i} + \begin{pmatrix} r\cos(\alpha_{j,i}) \\ r\sin(\alpha_{j,i}) \end{pmatrix} \middle| \begin{pmatrix} x \\ y \end{pmatrix}, J[f_{j,i}](r)\Sigma J^T[f_{j,i}](r)\right) dr, \tag{6.54}$$

where $\tilde{p}$ denotes the predicted multipaths that are chosen according to the measured AoA φ and RToA τ. Several rules for choosing multipaths out of the database for a given measurement $(\varphi,\tau)^T \in [-\pi,\pi]\times\mathbb{R}^+$ are possible. Note that the assignment problem, which is inherently present when studying BML scenarios, shows up here. In this work, first the paths with the best match for the AoA φ are chosen. Afterwards, out of these the multipath with the closest RToA to the measurement τ is selected. However, a problem for the multipaths with RToA $= 0$ arises. Due to the fact that the ray tracer computes the wave propagation pointwisely per hypothetical emitter position, the RToA cannot be used to distinguish multipaths when it is equal to 0. Due to this reason, measurements that possess an RToA of 0 are not used for the localization.

Drawing of Particles New particles have to be created in the initialization step and for the detection of newborn targets in each iteration. The easiest way of distributing particles is to use a uniform distribution over the FOV. In [RCV10] the authors propose to create newborn particles around the measurements from the previous iteration to avoid a high number of additional particles in scenarios with a high probability of detection. In terms of BML this means that particles are distributed along the received multipaths from the previous iteration. Furthermore, context information from the ray tracer database can be used to draw particles only in regions, where

targets are observable.
The database of the ray tracer includes context information about the urban environment in terms of a city and building map. It consists out of a finite set of grid points, since the ray tracer prediction is done pointwisely per OS–MS combination. If there are no multipaths received by the OS for a specific MS position, the respective grid–point of the ray tracer databases possesses the empty set. This information can be used to find out, where possible targets can be observed by the localization algorithm. Thus, the allowed particle positions with respect to the context information are the grid points of the ray tracer database which possess at least a set of one multipath. In the following, these positions are referred to as valid street points. The implementations of the standard and generalized PHD intensity filters studied in the following sections always use only valid street points for the prediction and initialization of new particles.
By applying for each received measurement (except the one with RToA $= 0$) the approach defined in Section 6.2.2.2, a predicted multipath is chosen out of the ray tracer database. Afterwards, the OS, the interaction points and the MS are linearly connected. The connection is discretized, where the distance between two grid points is given by $d > 0$. Possessing a set of grid points for each multipath, boxes with side length d into the tangential direction and a side length proportional to the covariance of the likelihood function defined in (6.54) into the normal direction are generated. The center of these boxes are the respective grid points (see Figure 6.7). The union of all boxes for one measurement then limits the region of allowed particle positions per measured multipath. Hence, the set of allowed particle positions for a specific measurement is given by the street points which fall into the union of boxes of the chosen multipaths.

Approximation of the Likelihood Function Since the absolute path length is not observable in BML scenarios, a marginalization with respect to the range is needed for the evaluation of the likelihood function defined in (6.54). Therefore, an integral of Gaussian kernels (which are linear in target space) over the distance between two interaction points has to be computed for all interaction points of the chosen multipaths. Since the integral defined (6.53) cannot be evaluated analytically [SW09] it has to be approximated. In this work the integral is approximated by step functions centered at the grid points of the discretized multipath, with support length d.

6.2.3 Numerical Evaluation

To verify that the proposed approach can be used to solve the task of BML a single–target scenario is considered in an urban environment. First, a ray tracer predicts for each possible MS position the set of multipaths that are received by

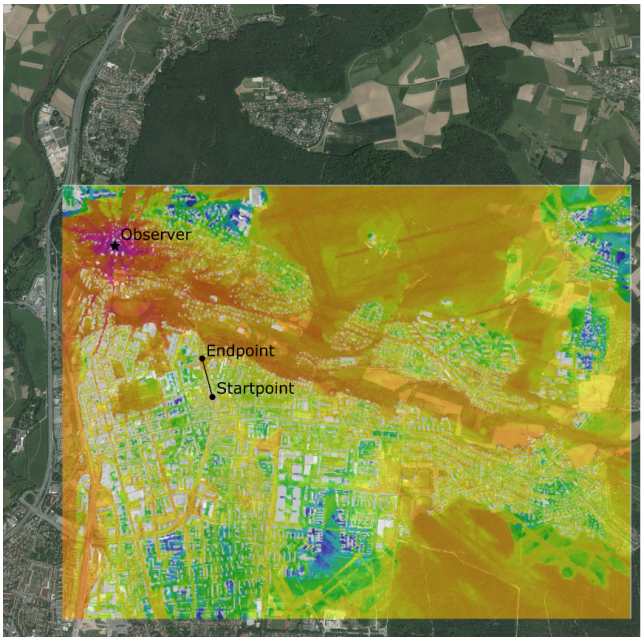

Figure 6.8: Visualization of the scenario used for the evaluation (city of Erlangen/Germany). Map Data: ©GeoBasis–DE/BKG 2015. Ray–Tracer Visualization: © 2015 AWE Communications.

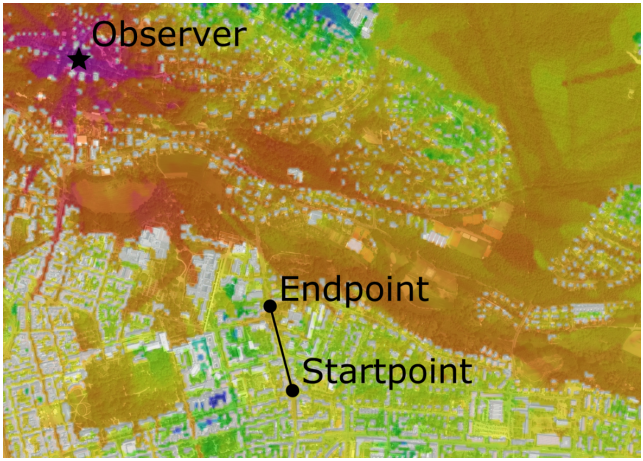

Figure 6.9: Zoom of the investigated scenario. Map Data: ©GeoBasis–DE/BKG 2015. Ray–Tracer Visualization: © 2015 AWE Communications.

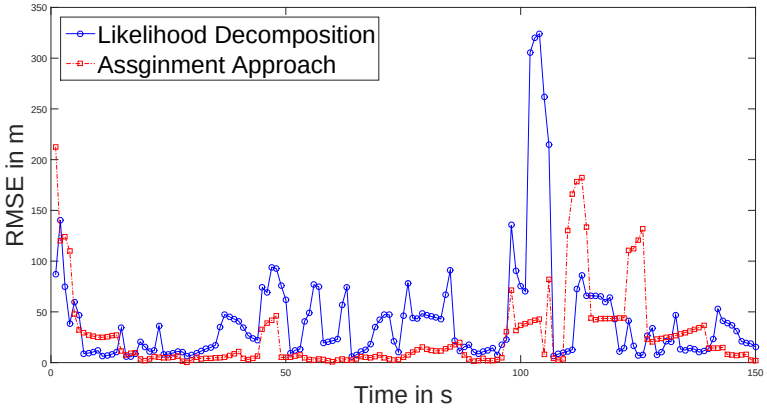

Figure 6.10: Result of the numerical evaluation. The likelihood function which is defined via an assignment approach (6.13) (red) is compared to the approximation of the decomposed likelihood function defined in (6.54) (blue) © 2014 IEEE.

the (fixed) OS. Figures 6.8 and 6.9 depict the trajectory of the target and the position of the OS. The prediction of the ray tracer is also visualized in Figures 6.8 and 6.9, where the colors imply the field strength that is received by the OS for the respective MS position. Due to numerical reasons, a restriction of the database to $[644550, 646550] \times [5495250, 5497250]$ in UTM–zone 32U is used, where the grid–size is 10m in both directions. Afterwards, a linear ground truth for a target that is moving with a constant velocity of $2.4\frac{\mathrm{m}}{\mathrm{s}}$ is created. Then for each time step the lower–left grid point of the box, in which the target is located, is determined and the corresponding multipaths from the database are processed by a simulation of the estimation process of an antenna array with 5 elements. Based on the center frequency of the emitter, the bandwidth of MS and OS, the power of the signal, the number of samples and several other parameters the set of resolved multipaths is determined and to each multipath parameter noise is added. For this evaluation the bandwidth is set to 6 MHz, the center frequency is 470 MHz, the number of samples is 4096 and the power of the emitted signal is set to 1 W. Neither clutter nor measurement failure (in terms of non–detections in the target space) is considered. However, due to the applied emulation of the antenna array a difference between the predicted and the measured multipaths for a particular MS location exists in the number of resolved, false and deteriorated multipaths.

The proposed decomposed likelihood function is compared to the likelihood function based on a (global) assignment approach. For the numerical evaluation of this section both likelihood functions are integrated into an SMC–implementation of the

iFilter, using the generalized mean state extraction scheme from Section 6.1.3.2, that is, assuming implicitly a single–target scenario. Note that according to the previous section also an SMC–implementation of the PHD filter could be used. However, due to the close relation of both filters mentioned in Section 3.4.2 the iFilter is chosen as a representative of this class of pointillist filters. The first filter utilizes the proposed decomposed likelihood function according to (6.54), while the second filter applies the assignment–based likelihood function given in (6.13).

For the implementation of the decomposed likelihood function the measurement noise in AoA is set to $\sigma_\varphi = 10^{-6^\circ}$ and in range to $\sigma_r = 100\text{m}$. Due to numerical reasons the decomposed likelihood function is bounded below by 10^{-10}. The assignment–based likelihood function uses a measurement noise in AoA of $\sigma_\varphi = 0.01^\circ$ and in RToA of $\sigma_\tau = \frac{5\text{m}}{c_{\text{light}}}$, where c_{light} denotes the speed of light. For each iteration and each multipath the same covariance matrix $\text{diag}\begin{bmatrix}\sigma_\varphi^2 & \sigma_r^2\end{bmatrix}$ in the first filter employing the likelihood–decomposition, respectively $\text{diag}\begin{bmatrix}\sigma_\varphi^2 & \sigma_\tau^2\end{bmatrix}$ in the second filter applying the assignment–based likelihood function is used.

The compared approaches share the same parameterization of the SMC–iFilter. The following detection probabilities are defined. The detection probability in the target space $X \subset \mathbb{R}^2$ is set to $p^D(x_i) = 0.75$ for each particle $x_i \in X$ and the detection probability in the space of ϕ hypothesis X_ϕ is defined as $p^D(\phi) \equiv 0.4$. The transition probability from X_ϕ to X is set to $\Psi(x|\phi) = 0.2$, the transition probability in X_ϕ is defined by $\Psi(\phi|\phi) = 0.01$ and the transition probability from X to X_ϕ is given by $\Psi(\phi|x) = 0.1$. The thresholds for target existence of the generalized grouping of particles are set to 0. This is done to obtain an estimate of the SMC–iFilter in each iteration and to compare the two approaches in terms of the RMSE. Furthermore, for both versions the maximal and minimal number of particles is restricted to 1500 and 500, respectively. New particles are created in the second version uniformly over the set of possible MS positions, where the number of newborn particles is determined according to [SKSC12]. For the first version particles are generated along the multipaths according to Section 6.2.2.2 and the particle positions are restricted to the MS positions which posses a distance $r \in [1000\text{m}, 1500\text{m}]$ to the OS position. Both versions draw their particles only on valid street points, that is, possible MS positions that possess at least one multipath.

For the numerical comparison of the two approaches 100 MC runs of the presented scenario are performed for a dataset. The results are shown in Figure 6.10 in terms of the position RMSE. Obviously, the assignment approach provides a more accurate and stable result for the first 100 time steps. The decomposed likelihood shows for the iteration in this period a localization result that is clearly below 100m and which is comparable to the result of the assignment approach until iteration 30. However, it is not as stable and accurate for the time period $[50s, 100s]$ as the assignment approach. This is essentially due to the influence to the fusion algorithm of the paths

with RToA= 0 for these iterations. Due to the reasons mentioned in Section 6.2.2.2 these paths are not considered in the filter using the decomposed likelihood function, since they are not distinguishable. This is a true benefit of the assignment approach, which can take into account these paths. Around iteration 100 the filter which uses the decomposed likelihood function is not able to generate a correct estimate. The wrong result is due to setting the threshold to zero and forcing the filter to estimate in each iteration. However, for most of the iterations after iteration 110 the decomposed likelihood provides a comparable and sometimes even better result than the assignment approach.
One reason for the better result of the assignment approach is that the paths with RToA= 0 are not used by the decomposed likelihood approach. Therefore, important information is not used by the proposed approach due to the fact that the respective paths are indistinguishable. Additionally, the simple approach of selecting the correct multipaths for the evaluation of Equation (6.54) is very sensitive to noise in the multipath parameters and it might happen, that non–correct multipaths are chosen. An enhanced method for determining the predicted multipaths from the ray tracer according to the set of measured multipaths should improve this property. Furthermore, so–called ghosts, that are targets which do not exist and arise due to intersections of selected multipaths are a challenge for the proposed method. Due to the intersection of the selected multipaths from the ray tracer prediction, the weight of the intensity is concentrated not only around the true target position but also around these intersection points. Hence, it appears that the generalized mean target state extraction scheme generates an estimate for an intersection point. This is also the reason why the allowed region of particles is restricted for the SMC–iFilter using the decomposed likelihood function. Since the extraction scheme is not well–suited for this circumstance, an improved extraction approach could provide a better result and make a restriction of the allowed region of particles for the first version superfluous.

6.2.4 Conclusion

In this section a general likelihood decomposition of the bearings–only likelihood function presented in [SW09] is generalized to the problem of BML. First, a decomposition of the likelihood function is presented. To this end, a ray tracer is used to replace the (in general not known) measurement function and its inverse. It determines the multipaths for a given MS–OS combination in terms of the AoA, RToA and the set of interaction points. Assuming a linear propagation of the electromagnetic wave in between interaction points, the decomposed likelihood can be formulated for the BML–scenario.
This approach provides a completely different point of view for the definition of a likelihood function of BML. Up to now the respective likelihood functions were defined via an assignment approach (see Section 6.1, [Alg10]), where the ray tracer prediction

for a hypothetical emitter position is compared to the set of measured multipaths. By applying the decomposition approach, presented in [SW09] the solution is received by weighting particles according to their distance to the different received multipaths. This alternative approach bears several challenges. It turns out that finding the correct multipath to a specific measurement and the appearance of ghosts are challenges of the proposed approach that need to be worked on. However, it is also shown that the proposed approach can be used for localization and tracking purposes in BML.
Due to the fact that the likelihood function defined via the assignment approach clearly outperforms the decomposed likelihood function, the numerical comparisons of the generalized and standard PHD intensity filters in the following section will be done using likelihood functions that are defined using an assignment approach.

6.3 Generalized PHD Intensity Filters Applied to BML

Due to the assumption that targets generate conditionally independent observations with at most one observation per target, the standard PHD intensity filters defined in Section 3.4.2 are not suited a priori for applications where a target might generate multiple measurements in one sensor scan. The application of a PHD intensity filter using a standard target–oriented measurement model implies that enhanced post–processing algorithms for the target state extraction are needed. This approach is proposed in Section 6.1 and promises great applicability to the challenge of BML. However, choosing the mathematically correct target–oriented measurement model of a standard PHD intensity filter has to be considered when studying the applicability of PHD intensity filters to the challenge of BML.
In Section 3.2.2 the extended target–oriented measurement model is proposed. It assumes that a target generates a random number of identically distributed (point) measurements. According to the strategy of designing a pointillist filter from Section 3.8 this target–oriented measurement model can be integrated into a large class of multitarget tracking filters. For the problem of extended object tracking several modifications of the standard PHD filter are available (an excellent overview about existing methods is given in [MCF$^+$14]). In [Mah07b] the target extent is modeled by a set of point scatterers, where each scatterer generates an individual measurement. In [Mah09] an approximation is presented, based on the approximate Poisson model of Gilholm, Godsill, Maskell and Salmon [GGMS05b], where the target extent is modeled by a spatial probability distribution. Furthermore, the set of measurements is preprocessed into associated groups which represent the individual targets. In [KS05] the target extent is modeled by random matrices and in [Gra12], [GLO12], [GO12a], [GO13], [GO12b] the approach is combined using PHD and cardinalized PHD filters. In [RBGD15] the approach is combined using multi–Bernoulli filters. Furthermore, techniques for reducing the number of measurement set partitions, which are essentially based on clustering measurements, are presented

in these references. All methods mentioned have in common that they make explicit use of the target extent. However, in a BML–scenario measurements corresponding to one target are not identically distributed (in fact the measurement distribution in the parameter space is hard to model and depends on multiple factors like the urban environment, the frequency, weather conditions, external factors like pedestrians and cars, etc.) and thus measurements corresponding to the same target are note spatially related in the measurement space. This makes the application of cluster–based extended target models and approximations to a BML–scenario impossible.
To this end, the target–oriented measurement model cannot be designed using a PGF as described in Section 3.2.2.2, but an appropriate PGFL as defined in (3.13) has to be used, which enables a target to generated multiple correlated measurements per sensor scan. Generalized PHD intensity filters presented in Section 3.4.4 are the result of generalizing the target–oriented measurement and the clutter model of a standard PHD intensity filter from Section 3.4.2. The purpose of this section is to study the applicability of generalized PHD intensity filters for tracking targets that generate multiple measurements per sensor scan, which are not drawn in the measurement space according to a joint distribution. In particular, the derivation of non-cluster based approximation schemes for reducing the computational complexity is in the focus of this section. Furthermore, the comparison of the performance of generalized PHD intensity filters to the adapted standard PHD intensity filter proposed in Section 6.1 is also part of the analysis carried out.
This section is structured as follows. In Section 6.3.1 the problem is formulated. In Section 6.3.2 approximations of the update equation are proposed. These approximations are based on restricting the cardinality of the processed partitions (see Section 6.3.2.1) and removing partitions based on the value of the likelihood function value (see Section 6.3.2.2). Furthermore, the generalization of the probability of detection is studied in Section 6.3.3. Three numerical evaluations are carried out to assess the proposed methods. First, in Section 6.3.4.1 the generalized PHD filter is compared to the generalized iFilter. Second, the approximation methods are investigated closely in Section 6.3.4.2 and finally the concepts of generalized and standard PHD intensity filters are compared at hand of a BML scenario in Section 6.3.4.3. Finally, conclusions are drawn in Section 6.3.5.

6.3.1 Formulation of the Problem

In this section the update equations of the generalized PHD intensity filters presented in Section 3.4.4 are given and issues concerning the probability of detection and the computational complexity are identified. The derivation of both filters is not presented here. As mentioned in Chapters 3 and 4 the update equation can be derived straightforwardly in various ways. The generalized PHD is first derived in [CM12] using the general chain rule (GCR) for functional differentiation. The generalized

iFilter is first derived in [Deg14]. The details on the derivation of the filters an be found there.

6.3.1.1 The Generalized PHD Filter

In [CM12] the authors present the GCR, which is a generalization of the fourth chain rule for functional derivatives from [Mah07b]. This method can be used to differentiate complex PGFLs, such as the PGFL of the generalized PHD Intensity filters defined in (3.64). Furthermore, the generalized PHD filter is developed using the GCR. Since targets can generate multiple measurements per sensor scan the likelihood function is defined on sets of measurements. If the clutter is assumed to be Poisson with intensity $\lambda c(\cdot)$ and $\mu s(\cdot)$ denotes the intensity of the PPP modeling the target intensity (λ and μ denote the mean number of clutter and targets, respectively, while $c(\cdot)$ and $s(\cdot)$ denote the distribution of one clutter measurement and one target state in their spaces, respectively), the update equation of the generalized PHD filter for an arbitrary number of measurements per target (and sensor scan) is given for $x \in X \subset \mathbb{R}^{d_1}$ and $z^1, ..., z^m \in Y \subset \mathbb{R}^{d_2}$, $d_1,\ d_2 > 0$ by

$$\mu_{X|Y}(x|z^1, ..., z^m) = \mu s(x) \left(p_0(\emptyset|x) + \frac{\sum\limits_{\pi \in \Pi_{(1:m)}} \sum\limits_{j=1}^{|\pi|} p_{|\pi_j|}(i(\pi_j)|x) \prod\limits_{k=1, k\neq j}^{|\pi|} \eta_{\pi,k}}{\sum\limits_{\pi \in \Pi_{(1:m)}} \prod\limits_{j=1}^{|\pi|} \eta_{\pi,j}} \right), \tag{6.55}$$

where

$$\eta_{\pi,j} \equiv 1_{\{a:|a|=1\}}(i(\pi_j))\lambda c(i(\pi_{j,1})) + \mu \int_X s(x) p_{|\pi_j|}(i(\pi_j)|x) dx \tag{6.56}$$

and $\Pi_{(1:m)}$ denotes the set of all partitions of

$$\{\delta^{z^1}, ..., \delta^{z^m}\}, \tag{6.57}$$

e.g.,

$$\Pi_{(1:2)} = \left\{ \left\{ \{\delta^{z^1}\}, \{\delta^{z^2}\} \right\}, \left\{ \{\delta^{z^1}, \delta^{z^2}\} \right\} \right\}. \tag{6.58}$$

The probability of detection is incorporated in the likelihood function $p_{|\cdot|}(i(\cdot)|x)$, which itself is defined on sets of measurements. In (6.57) δ^c, $c \in Y$ denotes the Dirac delta defined in Section 4.1.1 generalized to d_2 dimensions. The function

$$i : \Pi_{(1:m)} \to \mathcal{P}(Y) \tag{6.59}$$

is defined by

$$i(\{\delta^{z^1},...,\delta^{z^m}\}) \equiv (z^1,...,z^m), \tag{6.60}$$

for all $j \in \{1,...,m\}$ and

$$1_{\{a:|a|=1\}}(\pi) = \begin{cases} 1, \text{if } |\pi| = 1 \\ 0, \text{otherwise} \end{cases} \tag{6.61}$$

defines the indicator function. In (6.59) $\mathcal{P}(Y)$ denotes the power set of the set of received measurements.
According to Chapter 4 an alternative approach to the GCR for deriving the update equation of the generalized PHD filter via ordinary differentiation is given by the theory of secular functions [Str14e].
Note that (6.55) can handle correlated measurements originating from a specific target, since only the assumption that the measurement process is the superposition of $M > 0$ mutually independent (conditioned on $\{x_1,...,x_M\}$) target–oriented measurement processes is needed for the derivation of the update equation (for details see the derivation of the generalized iFilter presented in [Deg14]). In particular, the measurements are not assumed to be independent, conditioned on a specific target state. Measurements originating from different targets cannot be correlated since in the derivation of the generalized PHD filter in [CM12] the corresponding measurement processes need to be mutually independent.

6.3.1.2 The Generalized Intensity Filter

The update equation of the generalized iFilter is given for $z^1,...,z^m \in Y \subset \mathbb{R}^{d_2}$, $d_2 > 0$ by

$$\mu_{X|Y}(x|z^1,...,z^m) = \mu s(x) \left(p_0(\emptyset|x) + \frac{\sum\limits_{\pi\in\Pi_{(1:m)}} \sum\limits_{j=1}^{|\pi|} p_{|\pi_j|}(i(\pi_j)|x) \prod\limits_{k=1,k\neq j}^{|\pi|} \eta_{\pi,k}}{\sum\limits_{\pi\in\Pi_{(1:m)}} \prod\limits_{j=1}^{|\pi|} \eta_{\pi,j}} \right), \tag{6.62}$$

where $x \in X^+ \equiv X \cup X_\phi$, $X \subset \mathbb{R}^{d_1}$ and

$$\eta_{\pi,j} \equiv \int\limits_{X^+} \mu s(x) p_{|\pi_j|}(i(\pi_j)|x)dx = \mu s(\phi) p_{|\pi_j|}(i(\pi_j)|\phi) + \mu \int\limits_{X} s(x) p_{|\pi_j|}(i(\pi_j)|x)dx. \tag{6.63}$$

Equation (6.63) is due to the definition of the integral in $X^+ \equiv X \cup X_\phi$, see [Str13a, Definition (56)]. Analogously to the standard iFilter [SKSC12] the state space of the

generalized iFilter consists of the union of the target space X, which is equal to the target space of the generalized PHD filter, and the hypothesis space X_ϕ, which is used to model clutter by scatterers.
Note that the generalization of the target–oriented measurement model implies that a scatterer with a state in X_ϕ can also generate multiple clutter measurements per scan. This is equivalent to the arbitrary clutter model in the generalized PHD filter. The definition of the false–measurement model for elements from X_ϕ, that is, the clutter model, of the generalized iFilter can be incorporated into the generalized iFilter by the definition of the likelihood function $p_{|\cdot|}(\cdot|x)$ on X_ϕ for $x \in X$.

6.3.1.3 Computational Complexity and Probability Of Detection

The update equations (6.55) and (6.62) of both generalized PHD Intensity filters are numerically highly complex due to the sum over all partitions of the set of received measurements. The number of partitions is growing exponentially with the number m of received measurements and is given by the Bell number B_m. The exponential growth of the Bell number is visualized in Figure 6.11 and it is obvious that for an application of (6.55) and (6.62) approximations are inevitable. In [RBGD15], [Gra12], [GLO12], [GO12a], [GO13], [GO12b] and [SC10] clustering approaches, which are essentially based on the spatial relation of measurements, are used to reduce the number of partitions. These approximations are possible, if measurements that are generated by the same target are spatially related in the measurement space. However, in BML a target generates multiple measurements per sensor scan which are not spatially related in the measurement space. Therefore, the partitions in equation (6.55) need to be reduced without using any information about the spatial distribution of measurements. Instead, other criteria have to be identified to successfully approximate the update equations (6.55) and (6.62) with a feasible number (in terms of the numerical complexity) of partitions.
Furthermore, a modeling of the probability of detection via a simple Bernoulli trial (as it is done for the standard target–oriented measurement model from Section 3.2.2.1) is inappropriate if a target generates multiple measurements per sensor scan, since multiple measurement partitions might be processed which are subsets of each other. Therefore, the definition of the probability of detection in (6.55) and (6.62) (as part of the likelihood function) has to be modeled in a generalized fashion (as already mentioned in [CM12]).

6.3.2 Approximation of the Update Equation

As discussed in the previous section an evaluation of all measurement set partitions within generalized PHD intensity filters is numerically not feasible due to the exponential growth of the number of partitions with increasing set size. Furthermore, a

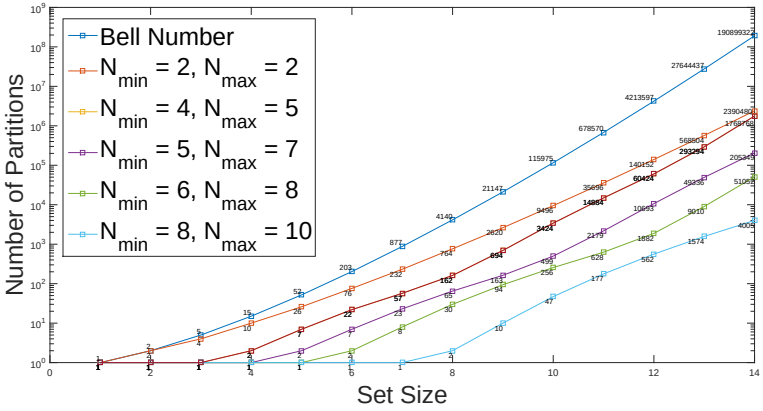

Figure 6.11: Comparison of the Bell number and the number of partitions due to approximation (6.85) © 2014 IEEE.

reduction of the number of partitions by the application of clustering methods is not applicable if measurements that belong to a specific target are not spatially related in the measurement space. To this end, two novel approaches are presented in the following section, which approximate the update equation of generalized PHD intensity filters by reducing the number of investigated partitions without assuming an underlying spatial distribution of the measurements which belong to a specific target. Finally, a generalized definition of the probability of detection is presented.

6.3.2.1 Incorporation of a Priori Information

The first proposed approximation of Equation (6.55) considers available a priori information about the number of generated measurements per target and sensor scan. The idea is to restrict the possible number of generated measurements, that is to assume that a target generates at least $N_{\min} \in \mathbb{N}$ and at most $N_{\max} \in \mathbb{N}$ measurements per sensor scan. This idea is straightforward and closely related to the approach for reducing partition hypothesis presented in [Alg10, Section 7.2.3]. Even though it might seem obvious how a restriction of the number of measurements per target will influence the update equation of the generalized PHD filter and iFilter, a detailed derivation is carried out in the following for the generalized PHD filter to demonstrate how a priori information and specific assumptions can be incorporated via a mathematically correct approach into an existing pointillist filter from Chapter 3.
To derive the respective PHD update equation the general higher order chain rule, presented in [CM12] is used to carry out the functional differentiation. Note that this

Analytical Step can also by done by the application of secular functions. Let

$$g : (Y, \mathcal{B}(Y)) \to (\mathbb{R}, \mathcal{B}(\mathbb{R})) \tag{6.64}$$

and

$$h : (X, \mathcal{B}(X)) \to (\mathbb{R}, \mathcal{B}(\mathbb{R})) \tag{6.65}$$

be bounded (by one), non–negative and Lebesgue–integrable test–functions, where $\mathcal{B}(\cdot)$ denotes the Borel–σ algebra of the respective space. First, the PGFL of the joint state is given analogously to Equation (3.64) by

$$\Psi_{\text{GenPHD}}(h, g) = \Psi_C^{\text{gen}}(g)\, G_N\big(\Psi_{\text{BMD}}^{\text{gen}}(h, g)\big) = (\exp \circ f)(h, g), \tag{6.66}$$

where

$$f(h, g) \equiv \lambda \left(\int_Y c(z) g(z) dz - 1 \right) + \mu \left(\int_X s(x) h(x) \Psi_{\text{obs}}(g|x) dx - 1 \right)$$

and the approximated PGFL of the likelihood function $\Psi_{\text{obs}}(g|\cdot) : X \to \mathbb{R}$, which incorporates the a priori knowledge on the number of measurements per target, is defined by

$$\Psi_{\text{obs}}(g|x) \equiv p_0(\emptyset|x) + \sum_{n=N_{\min}}^{N_{\max}} \frac{1}{n!} \int_{Y^n} \prod_{j=1}^{n} g(z^j) p_n(z^1, ..., z^n|x) dz^1 ... dz^n. \tag{6.67}$$

It holds by definition that

$$\Psi_{\text{BMD}}^{\text{gen}}(h, g) = \int_X h(x) \mu(x) \Psi_{\text{obs}}(g|x) dx \tag{6.68}$$

Note that the for changing the target–oriented measurement model, only $\Psi_{\text{obs}}(g|\cdot)$ has to be adapted. Applying the general higher order chain rule to determine the functional derivative of (6.66) with respect to impulses yields

$$\begin{aligned} \frac{\partial^m \Psi_{\text{GenPHD}}}{\partial \delta^{z^1} \cdots \partial \delta^{z^m}}(h, g) &= \frac{\partial^m (\exp \circ f)}{\partial \delta^{z^1} \cdots \partial \delta^{z^m}}(h, g) = \sum_{\pi \in \Pi_{(1:m)}} \frac{\partial^{|\pi|} \exp(f(h, g))}{\partial \xi_{\pi_1} \cdots \partial \xi_{\pi_{|\pi|}}} \\ &= \sum_{\pi \in \Pi_{(1:m)}} \exp(f(h, g)) \prod_{j=1}^{|\pi|} \xi_{\pi_j}[g, h], \end{aligned} \tag{6.69}$$

where

$$\xi_\omega(h, g) \equiv \frac{\partial^{|\omega|} f(h, g)}{\partial \omega_1 \cdots \partial \omega_{|\omega|}} = \mu \int_X s(x) h(x) \frac{\partial^{|\omega|} \Psi_{\text{obs}}(g|x)}{\partial \omega_1 \cdots \partial \omega_{|\omega|}} dx \tag{6.70}$$

and the functional derivatives are taken with respect to the function g. For the evaluation of (6.70) the functional derivative of definition (6.67) has to be considered. Therefore, let ω be an arbitrary element of a partition from $\Pi_{(1:m)}$. Then, the Gâteaux derivative of the functional is given by

$$\frac{\partial^{|\omega|}\Psi_{\text{obs}}(g)}{\partial\omega_1\cdots\partial\omega_{|\omega|}} = \sum_{n=N_{\min}}^{N_{\max}} \frac{1}{n!}\cdot n\cdot(n-1)\cdot\ldots\cdot(n-|\omega|+1) \times \int_{Y^{n-|\omega|}} \prod_{j=1}^{n-|\omega|} g(z^{j'})p_n(i(\omega),z^{1'},...,z^{(n-|\omega|)'}|x)dz^{1'}...dz^{(n-|\omega|)'} \quad (6.71)$$

if $|\omega| < N_{\min}$. If $|\omega| \in \{N_{\min},...,N_{\max}-1\}$ it is given by

$$\frac{\partial^{|\omega|}\Psi_{\text{obs}}(g)}{\partial\omega_1\cdots\partial\omega_{|\omega|}} = p_{|\omega|}(i(\omega)|x) + \sum_{n=N_{\min}}^{N_{\max}} \frac{1}{n!}\cdot n\cdot(n-1)\cdot\ldots\cdot(n-|\omega|+1) \times \int_{Y^{n-|\omega|}} \prod_{j=1}^{n-|\omega|} g(z^{j'})p_n(i(\omega),z^{1'},...,z^{(n-|\omega|)'}|x)dz^{1'}...dz^{(n-|\omega|)'} \quad (6.72)$$

and if $|\omega| = N_{\max}$ it is equal to

$$\frac{\partial^{|\omega|}\Psi_{\text{obs}}(g)}{\partial\omega_1\cdots\partial\omega_{|\omega|}} = p_{|\omega|}(i(\omega)|x). \quad (6.73)$$

If $|\omega| > N_{\max}$ the derivative is

$$\frac{\partial^{|\omega|}\Psi_{\text{obs}}(g)}{\partial\omega_1\cdots\partial\omega_{|\omega|}} = 0. \quad (6.74)$$

Thus,

$$\frac{\partial^{|\omega|}\Psi_{\text{obs}}(g)}{\partial\omega_1\cdots\partial\omega_{|\omega|}} = 1_A(\omega)p_{|\omega|}(i(\omega)|x) = \begin{cases} p_{|\omega|}(i(\omega)|x), \text{if } |\omega| \in \{N_{\min},...,N_{\max}\} \\ 0, \text{ otherwise,} \end{cases} \quad (6.75)$$

(6.76)

where

$$A \equiv \{a : |a| \in \{N_{\min},...,N_{\max}\}\}. \quad (6.77)$$

In the following, the short–hand notation from (6.75) is used. Given the functional derivative of the PGFL of the joint state with respect to impulses the update equation

of the corresponding PHD filter can be determined. It is given by

$$\mu_{X|Y}(x|z^1,...,z^m) = \left(\frac{\partial^m \Psi_{\text{GenPHD}}(0,1)}{\partial\delta^{z^1}\cdots\partial\delta^{z^m}}\right)^{-1}\left(\frac{\partial^{m+1}\Psi_{\text{GenPHD}}(0,1)}{\partial\delta^{z^1}\cdots\partial\delta^{z^m}\partial\delta^{x}}\right) \tag{6.78}$$

$$= \left(\sum_{\pi\in\Pi_{(1:m)}}\prod_{j=1}^{|\pi|}\xi_{\pi_j}(0,1)\right)^{-1}\left(\sum_{\pi\in\Pi_{(1:m)}}\frac{\partial B_\pi(0,1)}{\partial\delta^x}\right), \tag{6.79}$$

where

$$B_\pi(h,g) \equiv f(h,g)\cdot\prod_{j=1}^{|\pi|}\xi_{\pi_j}(h,g) \tag{6.80}$$

and

$$\begin{aligned}\frac{\partial B_\pi(h,g)}{\partial\delta^x} =& \mu s(x)\Psi_{\text{obs}}[g|x]\prod_{j=1}^{|\pi|}\xi_{\pi_j}(h,g)\\ &+\sum_{j=1}^{|\pi|}\mu s(x)\frac{\partial^{|\pi_j|}\Psi_{\text{obs}}(g|x)}{\partial\pi_{j,1}\cdots\partial\pi_{j,|\pi_j|}}\prod_{k=1,k\neq j}^{\pi}\xi_{\pi_k}(h,g)\end{aligned} \tag{6.81}$$

The evaluation of (6.79) yields the update equation of the approximated generalized PHD filter with Poisson clutter. It is given by

$$\mu_{X|Y}(x|z^1,...,z^m) = \mu s(x)\left(p_0(\emptyset|x) + \frac{\sum_{\pi\in\Pi_{(1:m)}}\sum_{j=1}^{|\pi|}1_A(\pi_j)p_{|\pi_j|}(i(\pi_j)|x)\prod_{k=1,k\neq j}^{|\pi|}\eta_{\pi,k}^{\text{PHD}}}{\sum_{\pi\in\Pi_{(1:m)}}\prod_{j=1}^{|\pi|}\eta_{\pi,j}^{\text{PHD}}}\right), \tag{6.82}$$

where

$$\eta_{\pi,j}^{\text{PHD}} \equiv \xi_{\pi_j}(0,1) = 1_{\{a:|a|=1\}}(\pi_j)\lambda c(i(\pi_{j,1})) + \mu\int_X s(x)1_A(\pi_j)p_{|\pi_j|}(i(\pi_j)|x)dx. \tag{6.83}$$

The update equation of the generalized iFilter is derived analogously. It is given by (6.82), but with $\eta_{\pi,j}^{\text{PHD}}$ replaced by

$$\eta_{\pi,j}^{\text{iFilter}} \equiv 1_{\{a:|a|=1\}}(\pi_j)\mu s(\phi)p_{|\pi_j|}(i(\pi_j)|\phi) + \mu\int_X s(x)p_{|\pi_j|}(i(\pi_j)|x)dx. \tag{6.84}$$

Due to the fact that some summands of Equation (6.82) are zero, computational effort can easily be saved. A summand of the sum over all partitions in (6.82) is zero if for the respective partition $\pi \in \Pi_{(1:m)}$ holds

$$\exists j \in \{1, \ldots, |\pi|\} : |\pi_j| \notin \{1, N_{\min}, ..., N_{\max}\}, \tag{6.85}$$

since then either $1_{\{a:|a|=1\}}(\pi_j) = 0$ or $1_A(\pi_j) = 0$. Therefore, the computational effort can be reduced by rejecting the partitions which fulfill condition (6.85). After rejecting the partitions, Equation (6.55) can be evaluated, since except for the appearance of $1_A(\cdot) = 0$ it is identical to Equation (6.82).
Note that partitions are not rejected, if they have a subset which is of cardinality one. This is independent of the choice of $N_{\min}$ and $N_{\max}$ and holds if a Poisson clutter model is chosen (for the generalized PHD) or if clutter scatterers are allowed to generate only single measurements (generalized iFilter)). For the generalized iFilter this has to be modeled via the likelihood function on X_ϕ. However, more enhanced clutter models could be included. For example, in a BML–scenario the context information, which is available due to a ray tracer, does not consider cars and other road users. Therefore, typical clutter sources in a BML–scenario can be road users, which reflect the signal emitted by the mobile station and act as new point sources of the reflected electromagnetic wave(s). Thus, multipaths which are received due to the same clutter source are not independent and hence clutter models which enable multiple measurements per clutter source could enhance the proposed data fusion algorithms. Obviously, condition (6.85) then needs to be adapted.

6.3.2.2 Evaluation of Significant Summands

In practical applications, the likelihood function is close to zero or might even be represented by zero for unlikely events due to the numerical resolution of the computer. Therefore, another practical approach for reducing the number of partitions which have to be considered in equation (6.82) is to evaluate only the terms for which the likelihood function value is above a specific significance–threshold. To this end, a criterion based on the cardinality of the partition elements is developed to determine these partitions. Let $\pi \in \Pi_{(1:m)}$ be an arbitrary partition which does not fulfill criterion (6.85) and $x \in X$ be an arbitrary target state. Then, if

$$\exists j \in \{1, \ldots, |\pi|\} : |\pi_j| > 1 \text{ and } p_{|\pi_j|}(i(\pi_j)|x) \leq \tau \tag{6.86}$$

is fulfilled

$$\sum_{j=1}^{|\pi|} 1_A(\pi_j) p_{|\pi_j|}(i(\pi_j)|x) \prod_{k=1,k\neq j}^{|\pi|} \eta_{\pi,k}^{\text{PHD/iFilter}} \approx 0 \tag{6.87}$$

approximately holds, where $\tau > 0$ is a chosen suitable small threshold for the significance of a partition. Note that $|\pi_j| > 1$ in (6.86) has to be fulfilled due to the first

summand in $\eta_{\pi,k}^{\mathrm{PHD/iFilter}}$, respectively, since otherwise it might happen that only the jth summand of (6.87) is approximately zero, while the other summands are significantly larger than zero. Hence, condition (6.86) can be used to reduce the number of the considered partitions. If $\tau = 0$ in (6.86), "$\approx$" can be replaced by "$=$" in (6.87). Note that for the application of this condition the likelihood function has to be evaluated for all possible subsets and all target states. The number of all possible subsets is given by the binomial series, e.g., for a set of m measurements,

$$N_{\mathrm{Subsets}} = \sum_{k=0}^{m} \binom{m}{k} = 2^m \tag{6.88}$$

subsets have to be evaluated. However, depending on $N_{\min}$ and $N_{\max}$ the application of condition (6.85) already reduces the number of subsets which have to be considered significantly, that is for $m > 0$ and $1 < N_{\min} \leq N_{\max} \leq m$ an application of condition (6.85) reduces the number of subsets of the measurement set, which have to be considered to

$$N_{\mathrm{Subsets}} = 2^{(N_{\max}+1)-N_{\min}+1}. \tag{6.89}$$

6.3.3 Generalization of the Probability of Detection

In [CM12] the authors emphasize that the probability of detection in [CM12, Equation (27)] is defined more generally than in the standard PHD filter, where the detection is modeled by a single Bernoulli process. If targets generate multiple measurements per scan, the detection process can be modeled by a discrete probability distribution over the number of measurements. This generalization of the probability of detection has also been proposed in [Alg10] by using the binomial distribution for the detection process within the definition of the probabilistic likelihood function for BML. For (6.55), (6.62) and (6.82) the single–target likelihood function can be formulated by

$$p_{|\pi_j|}(i(\pi_j)|x) = p(|\pi_j|, x) \cdot \hat{p}_{|\pi_j|}(i(\pi_j)|x), \tag{6.90}$$

where the sensor likelihood function $\hat{p}_{|\pi_j|}(i(\pi_j)|x)$ is given by

$$\hat{p}_{|\pi_j|}(i(\pi_j)|x) = \tilde{p}_{|\pi_j|}(i(\pi_j)|x) \tag{6.91}$$

and

$$\hat{p}_0(\emptyset|x) = 1. \tag{6.92}$$

Here and in the following, $x \in X$ for the generalized PHD filter (and $x \in X^+$ for the generalized iFilter) and $\pi_j \in \pi$, where $\pi \in \Pi_{(1:m)} \cup \emptyset$ denotes the set of partitions for a set of measurements of cardinality m, defined analogously to Section 6.3.1.1.
For the definition of the generalized probability of detection $p(\cdot, x) : \mathbb{N} \to [0, 1]$ the

detection process of the measurements, which are generated by the same target, needs to be investigated. If the detections (each considered as a random variable) of the single measurements are conditionally (conditioned on a specific target state) independent and have the same distribution (same detection probability), the detection process can be modeled by a series of Bernoulli–trials. If this assumption is fulfilled, a possible choice for $p(\cdot, x)$ is the Poisson distribution, that is

$$p(n, x) \equiv \frac{\lambda^n}{n!} e^{-\lambda}, \tag{6.93}$$

for all $n \in \mathbb{N}$, where the parameter $\lambda \in \mathbb{R}_{>0}$ is the expected number of measurements per target. If the number of measurements which are generated by a single target can be restricted to $N_{\max} \in \mathbb{N}$ the Binomial distribution can be used to model the detection process (see also [Alg10, Section 4.4]). It is given by

$$p(n, x) \equiv \binom{N_{\max}}{n} q^n (1-q)^{N_{\max}-n}, \tag{6.94}$$

for all $n \in \mathbb{N}$, where $q \in [0, 1]$ denotes the detection probability of an individual measurement.
Note that the Binomial distribution can be considered as a special case of the Poisson distribution. To this end, let

$$q \equiv \frac{\lambda}{N_{\max}}. \tag{6.95}$$

Then

$$\lim_{\substack{N_{\max}\to\infty,\\ q\to 0,\\ N_{\max} q\to\lambda}} \binom{N_{\max}}{n} q^n (1-q)^{N_{\max}-n} = \frac{\lambda^n}{n!} e^{-\lambda}, \tag{6.96}$$

for all $n \in \mathbb{N}$ [MS06, p.79].
Thus the Poisson distribution can be used to model the detection process for small detection probabilities of the individual measurements and a large number of trials, that is in scenarios where the target may generate a large number of measurements. The Binomial distribution can be used if the maximal number of measurements per target $N_{\max}$ is known. Note that the two proposed definitions for $p(\cdot, x)$ are only valid if the detections of the single measurements belonging to the same target are conditionally independent and identically distributed. If this assumption is not valid, other distributions need to be considered.

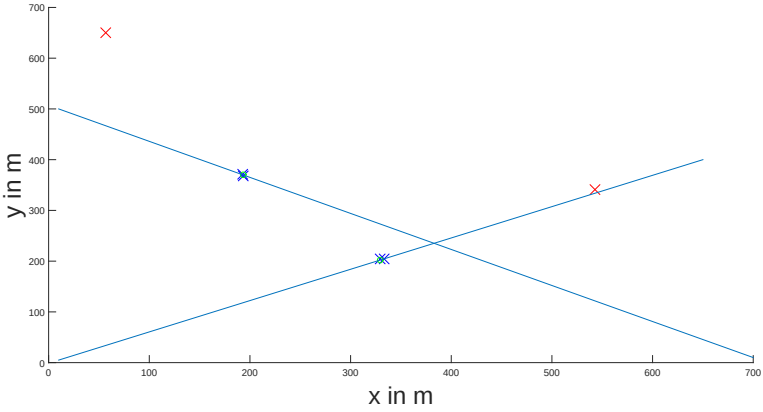

Figure 6.12: Visualization of the two–target scenario used for the comparison of the generalized PHD and the generalized iFilter. Two targets (green circle) are linearly moving with a constant velocity on their trajectory (blue line). In one iteration a target generates two correlated measurements (blue crosses), each with probability of detection $p^D = 0.8$. The measurements are drawn around the targets true position according to a Gauss distribution with covariance matrix Σ. Furthermore, two clutter measurements (red crosses) are generated uniformly over the FOV in each iteration.

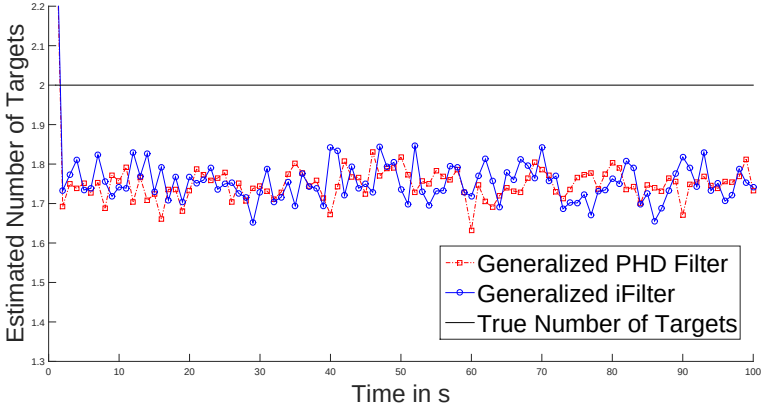

Figure 6.13: Estimation of the number of present targets.

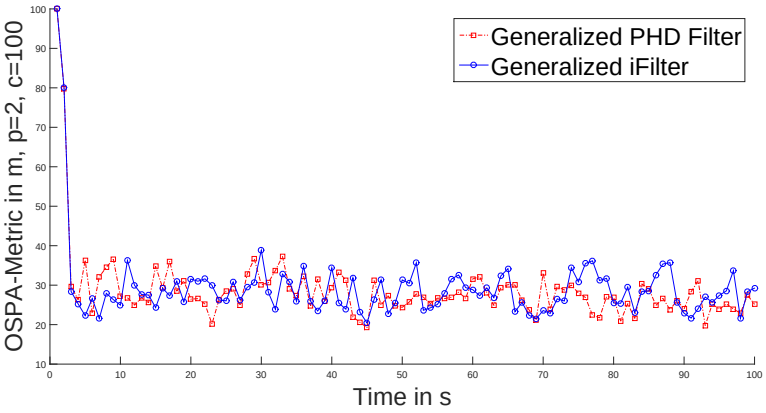

Figure 6.14: Mean of the OSPA–values with order $p = 2$ and cut–off value $c = 100$.

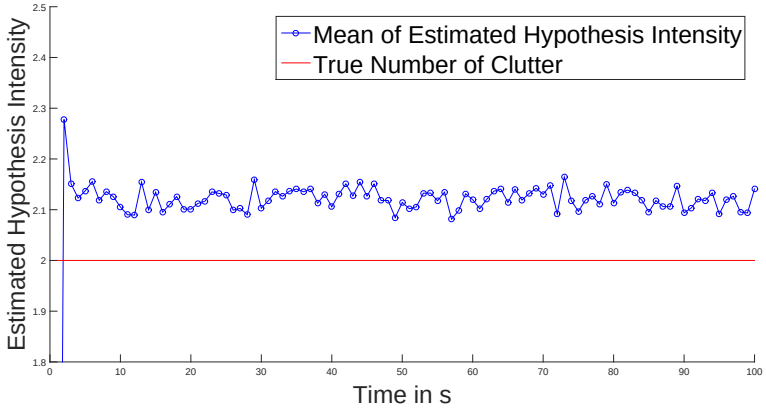

Figure 6.15: Estimated number of clutter of the iFilter.

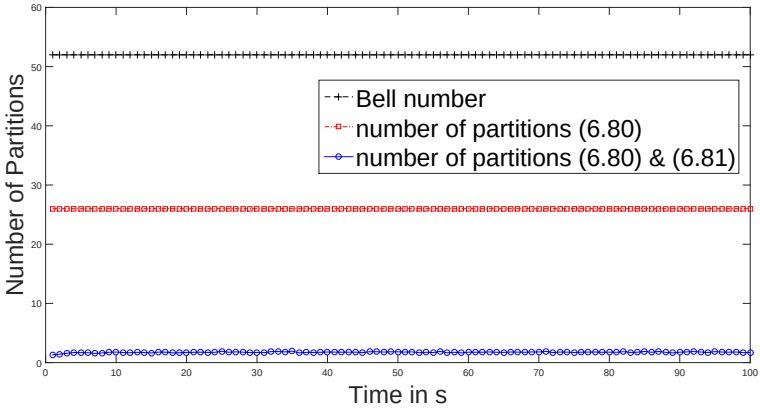

Figure 6.16: Number of partitions used due to the proposed approximation criteria. Here, $N_{\min} = N_{\max} = 2$ and the significance threshold $\tau = 0$ © 2014 IEEE.

6.3.4 Numerical Evaluation

6.3.4.1 Comparison of the Generalized PHD and Intensity Filter in a Multitarget Scenario with Correlated Measurements

To compare the generalized PHD and iFilter a numerical evaluation is carried out. To this end, a two–target scenario is considered (see Figure 6.12), where the targets are moving linearly with a constant speed of $5\frac{\mathrm{m}}{\mathrm{s}}$ (target 1) and $3\frac{\mathrm{m}}{\mathrm{s}}$ (target 2). The trajectory of the first target starts at $(10, 5)^T$ and is directed to $(650, 400)^T$. The second target starts at $(10, 500)^T$ and is directed to $(700, 10)^T$. In each iteration (the scan time is 1s) two correlated measurements are drawn per target according to a Gaussian distribution with the following parameters. The mean is given by the position of the target and the covariance matrix is given by

$$\Sigma = \begin{pmatrix} R & C \\ C^T & R \end{pmatrix}, \tag{6.97}$$

where $R = \mathrm{diag}[10\ 10]$ and $C = \mathrm{diag}[5\ 5]$. Since the probability of detection per measurement is $p^D(x) = 0.8$, for all $x \in \mathbb{R}^2$ the covariance matrix Σ is restricted according to the size of available measurements per target. Furthermore, two clutter measurements are drawn uniformly per iteration in the FOV, which is defined by $\mathrm{FOV} \equiv [0, 700] \times [0, 700]$. For the comparison of the generalized PHD and iFilter SMC–implementations are used. The prediction of both filters is the same as for the standard PHD filter, where the probability of survival is set to $p_s = 1.0$ for all particles and the single–object transition density is defined by the continuous white–noise

acceleration model from [BSLK01] with $\tilde{q} = 1.5$. To reduce the computational complexity, in each iteration 50 newborn particles are generated around the measurements of the previous iteration. For the initialization 100 particles are uniformly drawn in the FOV. A standard resampling algorithm (see [SKSC12]) is carried out and the maximal number of particles is restricted to 150. For the generalized iFilter the prediction of the intensity on X_ϕ is done analogously to [SKSC12]. For the update of the filters the conditions for reducing the number of partitions given in (6.85) and (6.86) are applied to both filters, setting $N_{\min} = N_{\max} = 2$. The likelihood function is given by $p_{|\cdot|}(i(\cdot)|\cdot) : E_Y \setminus \emptyset \times \mathbb{R}^4 \to \mathbb{R}$. Let in iteration $k \in \{1, ..., 100\}$ be $m \in \{2, ..., 6\}$ the number of all received measurements. Let $Y \subseteq \mathbb{R}^2$ be the measurement space. Then,

$$E_Y \equiv \emptyset \cup \bigcup_{n \geq 1} Y^{(n)} \tag{6.98}$$

is defined analogously to Section 2.1 by the space of sets of points in Y. For a subset $z = \{z^1, ..., z^n\} \in E_Y \setminus \emptyset$, $n \in \{1, ..., m\}$ and $x \in X \equiv \mathbb{R}^4$ arbitrary the likelihood function is defined by

$$p_{|z|}(i(z)|x) = p(|z|, x) \cdot \mathcal{N}\left(\begin{pmatrix} z^1 \\ \vdots \\ z^n \end{pmatrix}, \begin{pmatrix} Hx \\ \vdots \\ Hx \end{pmatrix}, \begin{pmatrix} R & C & \dots & C \\ C^T & R & \dots & C \\ \vdots & \vdots & \ddots & \vdots \\ C^T & C^T & \dots & R \end{pmatrix} \right), \tag{6.99}$$

where

$$H \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \tag{6.100}$$

and $p(\cdot, x) : \mathbb{N} \to [0, 1]$ defines the probability of detecting a set of measurements with the respective number of elements for a target state $x \in X$. For the evaluation it is defined by the Binomial distribution

$$p(n) \equiv \begin{cases} 1 - P_D, & \text{if } n = 0 \\ P_D \cdot \binom{2}{n} q^n (1-q)^{2-n}, & \text{otherwise,} \end{cases} \tag{6.101}$$

where $q \equiv p^D = 0.8$ and $x \in X$. Furthermore, the detection probability in X_ϕ in iteration k for the generalized iFilter is defined by

$$p_k(n, x) \equiv \begin{cases} \frac{\left(f_{k-1|k-1}(\phi)\right)^m}{m!} \exp(-f_{k-1|k-1}(\phi)), & \text{if } n = 1 \\ 0, & \text{otherwise,} \end{cases} \tag{6.102}$$

where

$$f_{k-1|k-1} : X^{+} \to \mathbb{R} \tag{6.103}$$

denotes the updated intensity of the iFilter from iteration $k-1$.
The generalized PHD filter uses a Poisson clutter model, which allows clutter to appear as single elements in the measurement space. The mean number of clutter of the generalized PHD filter $\lambda > 0$ is set to two (the correct number of clutter measurements per sensor scan) and $c > 0$ is uniform in the FOV. To extract a target state estimate in each iteration, the k–means algorithm is applied, where the number of clusters k is given by the rounded number of estimated target states from the filters. The transition probabilities of the generalized iFilter are set to $\psi(\phi|x) = 0.1$, $\psi(\phi|\phi) = 0.01$ and $\psi(x|\phi) = 0.2$.
To numerically assess both filters 100 MC runs are performed. Figure 6.13 visualizes the number of estimated targets and Figure 6.14 shows the tracking result in terms of the mean of the optimal subpattern assignment (OSPA)–values (order p= 2 and cut-off c= 100) over all MC runs. Figure 6.15 visualizes the value of the intensity of the generalized iFilter on X_{ϕ}, which represents due to (6.102) the number of present clutter measurements. From Figures 6.13 and 6.14 it can be seen that the generalized iFilter and the generalized PHD filter have a similar performance in terms of the estimation of the number of targets and the OSPA–metric. This result was expected, since both filters only differ in terms of their clutter modulation. However, it is remarkable that the generalized iFilter has no information about the number of clutter and estimates this information by the hypothesis intensity and its general scatterer measurement model. In contrast, the generalized PHD filter possesses in each iteration the correct information that two single clutter measurements are present. It has to be noted, that the clutter models of both filters are defined for single clutter measurements. The PHD filter uses a Poisson clutter model with fixed intensity, the iFilter defines the clutter model via the likelihood function over X_{ϕ}, which is unequal to zero for single measurements only (see (6.102)). In scenarios where clutter scatterers occur that generate multiple clutter measurements per scan (like in BML), these models could be changed. In such cases not the number of clutter, but the number of clutter scatterers is estimated by the generalized iFilter. Analogously, that clutter model for the generalized PHD filter could be generalized for such scenarios.
In the following, the generalized PHD filter as a representative of the class of generalized PHD intensity filters is used for a close inspection of the performance of the approximation criteria defined in (6.85) and (6.86). Afterwards, a possible application of the concept to a BML–scenario in comparison to the adapted standard iFilter is studied using simulated data.

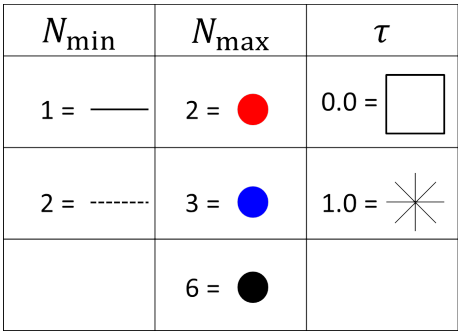

Figure 6.17: Legend for Figures 6.18 – 6.22.

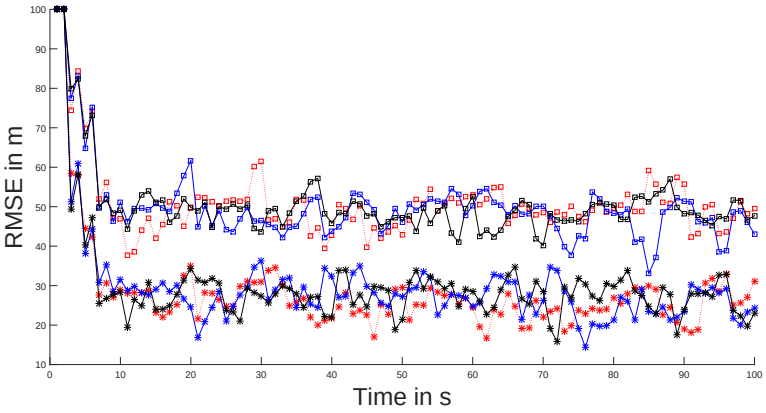

Figure 6.18: RMSE with respect to the first target. Due to the fact that the generalized PHD filter over–estimates the number of present targets if the significance threshold τ is set to 1.0 (see Figure 6.21) and the fact that the target state extraction is based on the rounded number of estimated targets, the parametrization using $\tau = 1.0$ performs better in terms of the RMSE than the parameterizations using $\tau = 0.0$.

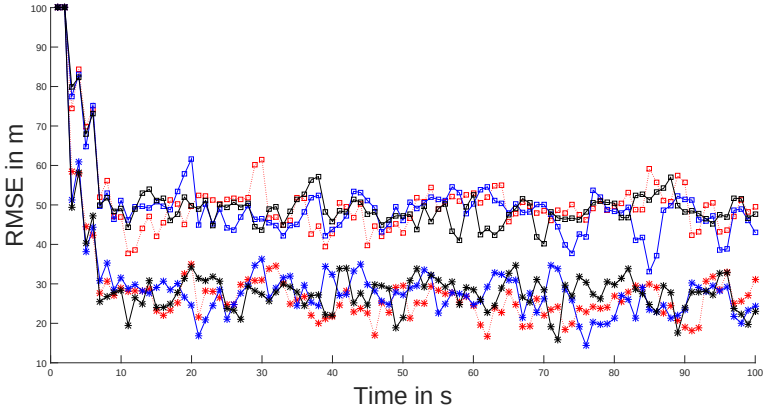

Figure 6.19: RMSE with respect to the second target.

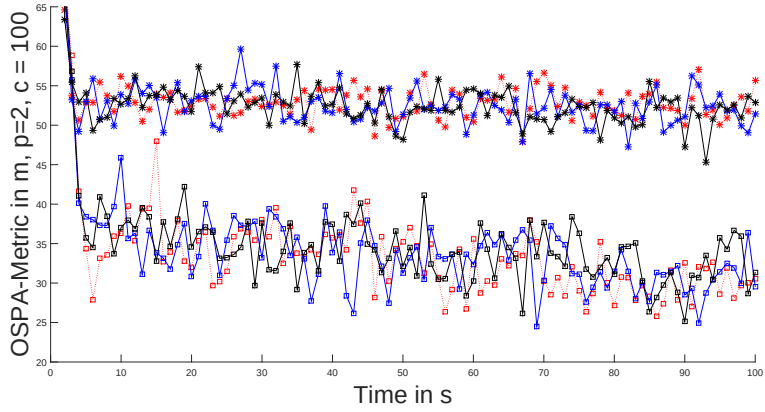

Figure 6.20: Mean of the OSPA–values with order $p = 2$ and cut–off value $c = 100$. In terms of the OSPA–metric the parametrizations using $\tau = 0.0$ perform better compared to those that use $\tau = 1.0$, since the over–estimation of the number of targets (see Figure 6.21) is penalized by the OSPA–metric. The change of the number of investigated partitions does not yield a significant alteration of the results.

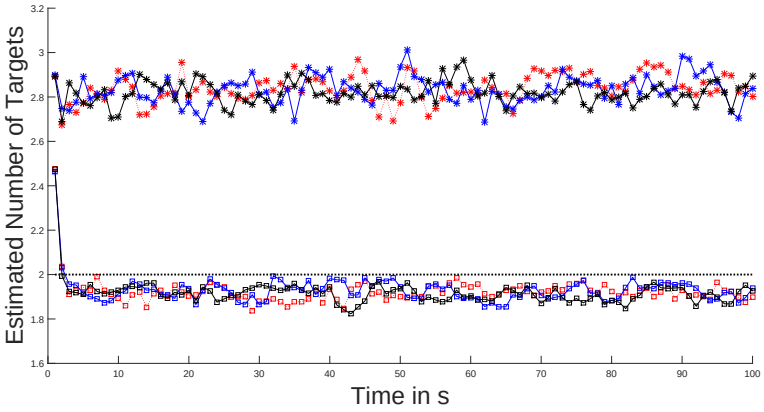

Figure 6.21: Estimated number of targets, where the dashed black line shows the true number of present targets.

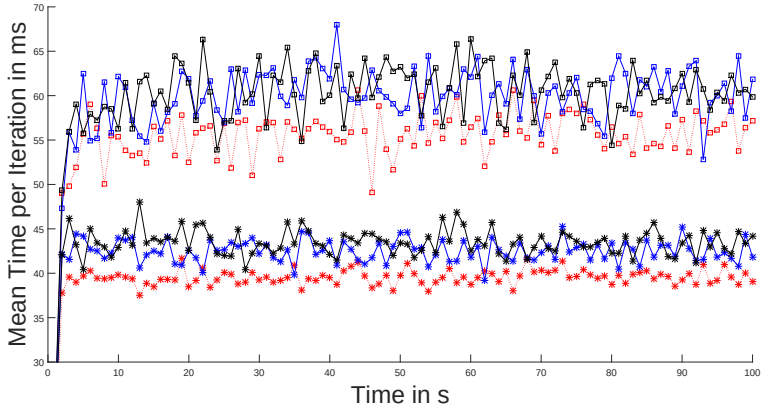

Figure 6.22: Comparison of the mean time for updating the generalized PHD filter of different parameterizations. It can be seen that the more partitions the filter processes and the smaller the significance threshold τ is chosen the longer the update takes.

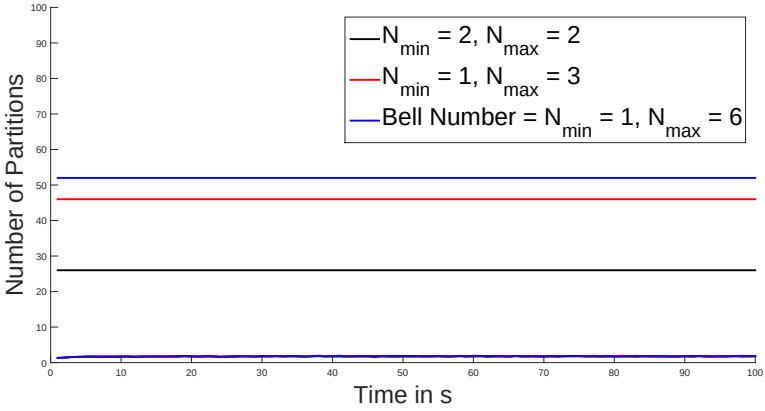

Figure 6.23: Mean number of partitions resulting from condition (6.85) and mean number of partitions due to condition (6.86) for two parameterizations ($N_{\min} = 2/N_{\max} = 2$ (black), $N_{\min} = 1/N_{\max} = 3$ (red)). The number of all partitions is given by the Bell number (blue). For the computation of the mean number of significant partitions in each MC run the mean number of significant partitions is computed for each time step over all particles. Afterwards, the mean of the number of significant partitions is computed over all MC runs. The number of significant partitions is almost the same for the two parameterizations.

6.3.4.2 Close Inspection on the Approximation Criteria in a Multitarget Scenario with Correlated Measurements

For a close inspection on the approximation criteria defined in (6.85) and (6.86), the generalized PHD filter is used as a representative of the family of generalized PHD intensity filters. In the following 100 MC runs of the scenario presented in Section 6.3.4.1 are performed for different parameterizations of the approximation conditions (6.86) and (6.85). Figure 6.21 visualizes the results of the different parameterizations in terms of the estimated number of targets. It can be seen that the estimated number of targets does not depend on the chosen partition sizes, that is, it does not depend on approximation criterion (6.86). This is due to the fact that the number of significant partitions with significance threshold $\tau = 0.0$ (and $\tau = 1.0$) is more or less equal for all three investigated parameterizations $N_{\min} = 2/N_{\max} = 2$, $N_{\min} = 1/N_{\max} = 3$ and $N_{\min} = 1/N_{\max} = 6$. Figure 6.23 visualizes exemplary for one parametrization the mean number of partitions, where in each MC run the mean number of significant partitions is computed for each time step over all particles. Furthermore, it can be seen from Figure 6.21 that the parameterizations with significance threshold $\tau = 1.0$ have a larger deviation in terms of the estimated number of targets than the parameterizations using the significance threshold $\tau = 0.0$ and over–estimate the true number of present targets. This yields to a better performance of the parameterizations using

the significance threshold $\tau = 1.0$ in terms of the RMSE with respect to the two true target states, which can be seen in Figures 6.18 and 6.19. This is due to the fact that the target state extraction is done using a k–means clustering algorithm, where k is given by the rounded estimated number of targets. Therefore, the over–estimation of the number of targets by the parameterizations with significance threshold $\tau = 1.0$ implies a clustering that always estimates at least two clusters. In contrast to that, the parameterizations with significance threshold $\tau = 0.0$ under–estimate the true target number for some iterations and thus the k–means clustering algorithm estimates only one cluster for these iterations. In iterations, where no estimate for a specific target is produced by the generalized PHD filter, the squared error is set to 100^2 m. However, the over–estimation yields to a worse performance of the parameterizations using a significance threshold $\tau = 1.0$ compared to the parameterizations, which use a significance threshold $\tau = 0.0$, since each over–estimation is penalized by the OSPA–metric. The result in terms of the mean time consumption per iteration is shown in Figure 6.22. It can be seen that the parameterizations using the significance threshold $\tau = 1.0$ are faster compared to the parameterizations using a threshold of $\tau = 0.0$. Furthermore, the parameterizations with $N_{\min} = 2/N_{\max} = 2$ perform better in terms of time consumption than $N_{\min} = 1/N_{\max} = 3$ and $N_{\min} = 1/N_{\max} = 6$. In summary: the less partitions and the larger the significance–threshold is, the faster and worser the algorithm performs.

Also a parameterization without using the two approximation conditions has been investigated in terms of the processed time per iteration. Since for this non– approximated SMC generalized PHD filter one iteration took up to $5.30 \cdot 10^3$ s, only one MC run has been performed. Thereby, the mean computation time was $2.28 \cdot 10^3$ s, which shows, that even if $N_{\min} = 1$ and $N_{\max} = 6$ (no approximation in terms of (6.85) has been made) and $\tau = 0.0$ is chosen, the respective generalized PHD filter parameterization ($N_{\min} = 1/N_{\max} = 3, \tau = 0.0$) performs about 45 times faster than the standard version, which does not use any approximation condition at all.

In summary, it is numerically shown that the proposed methods of approximation for generalized PHD intensity filters can be applied to scenarios where targets generate multiple measurements. It should be noted that the definition of the likelihood function does depend on the considered scenario. Furthermore, the following should be kept in mind. The integral of the clutter intensity $\lambda c(\cdot)$ yields the number of false measurements (not false targets). Thus λ denotes the mean number of false measurements per iteration. Hence, clutter is defined in terms of elements of the measurement space, not as clutter targets in the target space when applying a Poisson clutter model.

6.3.4.3 Comparison of the Generalized PHD Filter and the Standard Intensity Filter in a BML Scenario

To demonstrate the connection of generalized PHD intensity filters to the challenge of BML a single–target scenario in a simulated urban environment is presented. For generating multipath measurements, a database for a fixed OS and a grid of MS locations is generated analogously to Section 6.2.3 by using the ray tracing simulation, which is based on [HWLW03]. The distance between two grid points is set to 10 m. The number of received multipaths is restricted to six and each multipath is characterized by its AoA, its EoA and its RToA with respect to the first received multipath. Therefore,

$$Y \equiv [0, 2\pi] \times [-\pi, \pi] \times \mathbb{R}_{>0} \tag{6.104}$$

and

$$E_Y \equiv \emptyset \cup \bigcup_{n \geq 1} \mathcal{Y}^{(n)} \tag{6.105}$$

are defined analogously to Section 2.1.
Afterwards, a linear ground truth for the MS, which is moving with a constant velocity of 2.4 m/s, is simulated (see Figures 6.9 and 6.10). Then, in each time step the lower left grid point of the box, in which the target is located in, is determined. The multipaths which correspond to the chosen grid point are taken to generate the multipath measurements, referred to as the true multipaths in the following. First, in each iteration Gaussian distributed noise is added (this is different to Section 6.2, where a simulation for the antenna array and the receiver setup is used) to the true multipaths, where the standard deviations are set to $\sigma_\varphi = \sigma_\vartheta = 0.001$ rad for the azimuth and elevation of arrival and $\sigma_\tau = \frac{1.0}{c}$ s. Here, $c \equiv 299792458$ m/s defines the speed of light. Furthermore, a binomial detection process with probability of detection of $p^D = 0.95$ is used to model measurement failure. No clutter is added to the set of measurements.
The generalized PHD filter is implemented including the approximations proposed in (6.85) and (6.86), where $N_{\min} \equiv 3$, $N_{\max} \equiv 6$ and the threshold for significance of a partition $\tau \equiv 1.0 \cdot 10^{10}$. The FOV of the considered scenario is given by

$$\text{FOV} \equiv [645259.0, 645999.0] \times [5495257.0, 5496747]. \tag{6.106}$$

Furthermore, the probability of detection is independent of the target's state space, that is, $p(n, x) = p(n)$ for all $x \in X = \text{FOV} \times \mathbb{R}^2$, $n \in \{1, ..., 6\}$. It is modeled by (6.94), where $q \equiv p^D \equiv 0.95$. Then, it is incorporated into the likelihood function,

which is defined for a hypothetical emitter position $\xi \in X$ and a set of multipath measurements $\{z^j\}_{j=1}^m$, where $z^j \in Y$ (according to the ideas presented in [Alg10]) by

$$p\left(\{z^j\}_{j=1}^m|\xi\right) \equiv p(n) \cdot \lambda_\Phi^{m-n} \cdot \prod_{i \in I} \mathcal{N}\left(h_\xi^i; z^{a_i}, C^{a_i}\right), \tag{6.107}$$

where

$$h_\xi \equiv \{h_\xi^i\}_{i=1}^p \tag{6.108}$$

denotes the set of predicted multipaths with respect to ξ and the fixed OS coming from the ray tracer. Note the connection of (6.107) to the probabilistic likelihood function (6.9). The occurrence of clutter in a set of multipaths is modeled by

$$\lambda_\Phi \equiv \frac{0.1}{\text{FOV}}, \tag{6.109}$$

which is equal to the clutter density of the generalized PHD filter. The possible data interpretation is denoted by $E_{a_1,\ldots,a_p}^m$, where a_i is defined as in (6.8) and denotes the association of predicted to measured multipaths. However, due to the computational effort, we only use the best data association analogously to (6.10), which is determined by applying the Munkres algorithm [BL71] to the set of measured and predicted multipaths, using the Mahalanobis distance with the covariance matrix

$$C^{a_i} = C = \text{diag}[\sigma_\varphi^2 \ \sigma_\vartheta^2 \ \sigma_\tau^2] \tag{6.110}$$

for the construction of the cost matrix. Thus, the index a_i in (6.107) denotes the best (global) association for the specific predicted path. The generalized PHD filter is realized by an SMC–implementation, since the likelihood function can be computed only pointwisely. The maximal number of particles of the generalized PHD filter is restricted to 700 and particles are only drawn and predicted to grid points, where at least one multipath can be received due to the available database. In each iteration 200 newborn targets are uniformly drawn over the FOV. The single–object transition density is defined by the continuous white–noise acceleration model from [BSLK01], with $\tilde{q} = 1.5$ and the probability of survival $p_S = 1.0$. To extract the target states the weighted mean of all particles is computed.

For a numerically comparing the result of the proposed generalized PHD filter, the adaption of the standard PHD intensity filter to BML given in Section 6.1 is used. In the following, the generalized mean computation from Section 6.1.3.2 together with an SMC–implementation of the iFilter [SKSC12] is applied. The likelihood function $p(z^j|\xi)$ of a hypothetical emitter position $x_i \in X$ and one multipath $z^j \in Y$, $j \in \{1, ..., m\}$ is defined by (6.13). The assignment is done by an global association via Munkres algorithm between the set of measured multipaths Y and the set of predicted multipaths h_ξ of $\xi \in X$. Therefore, the index a_i denotes the assigned measured multipath z^j of the ith predicted multipath from h_ξ. The probability of detection

is set to $p^D(x) = 0.9$ for all $x \in X$ and the detection probability in the space of hypothesis X_ϕ is defined by $p^D(\phi) = 0.4$. The transition probability from X_ϕ to X is set to $\Psi(x|\phi) = 0.2$, the transition probability in X_ϕ is defined as $\Psi(\phi|\phi) = 0.01$ and the transition probability from X to X_ϕ is given by $\Psi(\phi|x) = 0.1$. The number of particles is restricted to 1500. The thresholds for target existence of the standard iFilter and the generalized PHD filter are set to 0 to make the filters comparable in terms of their RMSE–performance.

To assess both filters with respect to accuracy 100 MC runs of the presented scenario are performed. The result in terms of the RMSE is shown in Figure 6.24. It can be seen that both filters perform more or less equivalent after iteration 30 (the generalized PHD filter is slightly better in terms of its RMSE–performance). However, it also can be seen that until iteration 20 the generalized PHD filter performs worse than the standard iFilter. First, it is obvious that the initialization of the generalized PHD filter is not as good as the initialization of the standard iFilter. This is essentially due to the fact that the likelihood function of the generalized PHD filter is much more restrictive than the likelihood function of the standard iFilter. This is visualized in Figures 6.26 and 6.28, which shows the sum of the likelihood functions given in (6.107) and (6.13), that is,

$$\sum_{\pi \in \Pi^{Y^m}_{(1,3:6)}} p\left(\pi|\xi\right), \tag{6.111}$$

where $\Pi^{Y^m}_{(1,3:6)}$ denotes the set of all partitions of Y^m which have a cardinality $c \in \{1, 3, ..., 6\}$, that is,

$$\Pi^{Y^m}_{(1,3:6)} \equiv \left\{ \{\pi_{z^1}, ...\pi_{z^m}\} : \pi_{z^i} \subseteq Y^m, |\pi_{z^i}| \in \{1, 3, ..., 6\} \right\} \tag{6.112}$$

and

$$\sum_{j=1}^{m} p(z^j|\xi) \tag{6.113}$$

for the standard iFilter and $\xi \in X$ respectively. The number of investigated partitions is restricted due to the approximation condition (6.85), where $N_{\min} = 3$ and $N_{\max} = 6$. For a clearer visualization only values of the likelihood function, which are larger than 10^{10} are plotted. It is obvious that the shape of (6.111) is sharper and therefore more restrictive than (6.113). This is due to the fact that the likelihood function in (6.107) is given by a product of Gaussians. Therefore, partitions with at least one unlikely subset of multipaths (with respect to a hypothetical emitter position) possess a small likelihood function value. This contrasts the likelihood function of the standard iFilter, which only assesses single multipaths (see (6.13)). Due to the fact that the number of particles used by the generalized PHD filter is not sufficient to cover the FOV and due to the small width of the likelihood function

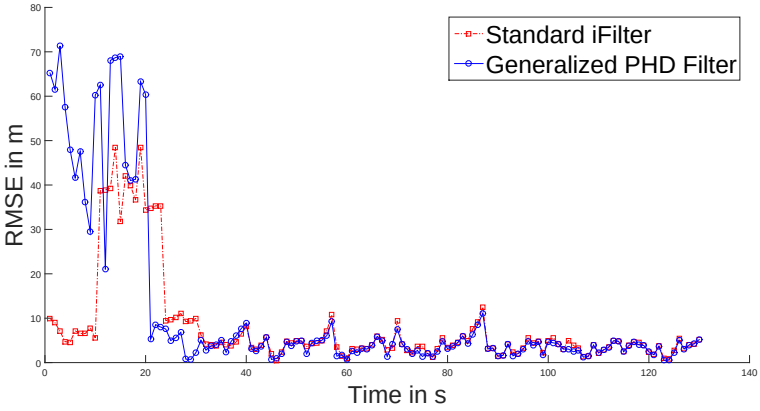

Figure 6.24: RMSE of the standard iFilter, which uses an enhanced post–processing scheme for target state extraction and the generalized PHD filter applying the proposed approximation conditions with $N_{\min} = 3$, $N_{\max} = 6$ and $\tau = 1.0 \cdot 10^{10}$.

(6.107), the mean time of convergence (until iteration 10–11) of the generalized PHD filter is larger than the time of convergence of the standard iFilter. Furthermore, it can be seen from Figure 6.24 that the localization of both filters around iteration 15 gets worse, while the generalized PHD filter performs worse than the standard iFilter. This is due to the fact, that the likelihood functions produce ambiguities in terms of the most likely hypothetical emitter position, which is shown in Figures 6.27 and 6.29 for iteration 15. Since in some MC runs a correct initialization of the generalized PHD filter is not performed until the occurrence of these ambiguities, the generalized PHD filter performs worse than the standard iFilter. However, the generalized PHD filter performs better in iteration 20–25, since the restriction cancels out the ambiguities earlier than the likelihood function of the standard iFilter.
The comparison of both filters in terms of estimated number of targets, that is, the integral of the intensity function over the FOV is visualized in Figure 6.25. Due to the assumption that one target generates at most one measurement per iteration the standard iFilter estimates the number of multipaths which belong to a target. In contrast to this the generalized PHD filter estimates after a few iterations the correct number of present targets.
The standard iFilter clearly outperforms in terms of the time consumption the generalized PHD filter: For one MC run the standard iFilter needs 82614 ms, while the generalized PHD filter takes 20250085 ms, which shows that it is of factor 245 slower than the standard iFilter.

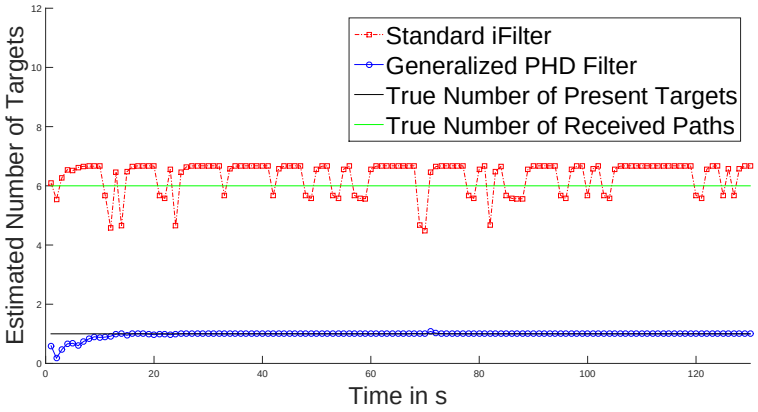

Figure 6.25: Estimated number of targets, that is the sum of all particle weights before resampling of the standard iFilter and the generalized PHD filter. Due to the assumption that one target generates at most one measurement per iteration, which is violated in the considered scenario, the standard iFilter estimates the number of multipaths, which belong to target. The generalized PHD filter is able to estimate the correct number of present targets.

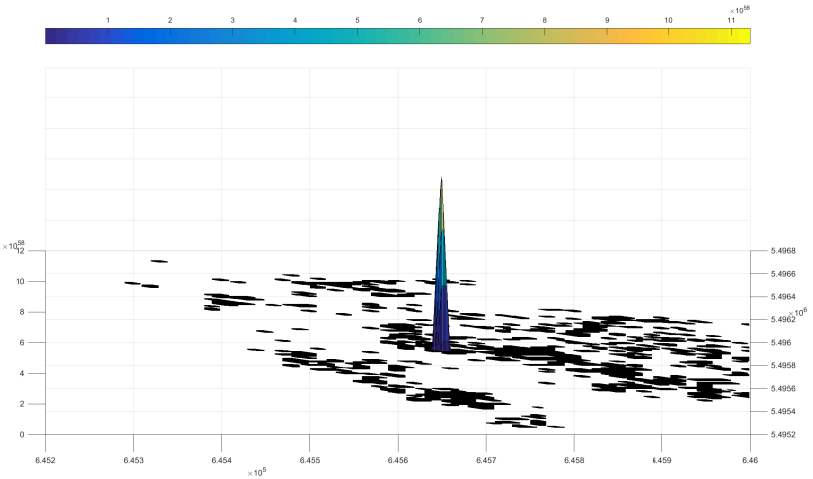

Figure 6.26: Visualization of the likelihood function defined in (6.111) at iteration 1.

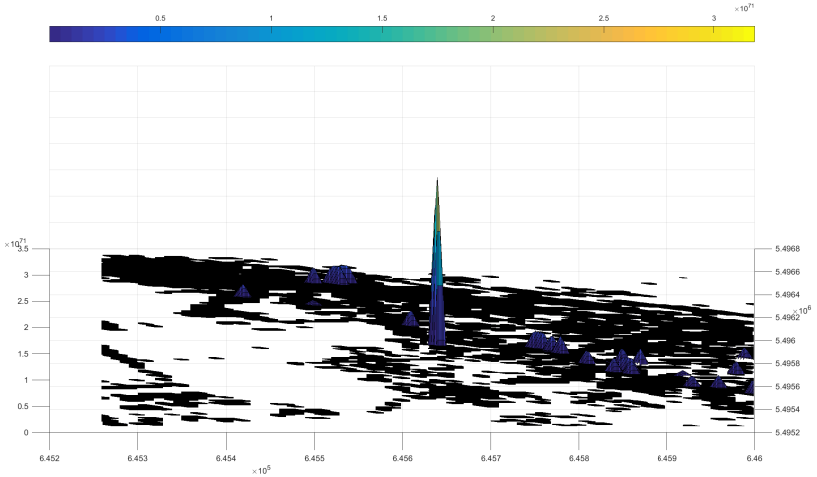

Figure 6.27: Visualization of the likelihood function defined in (6.111) at iteration 15.

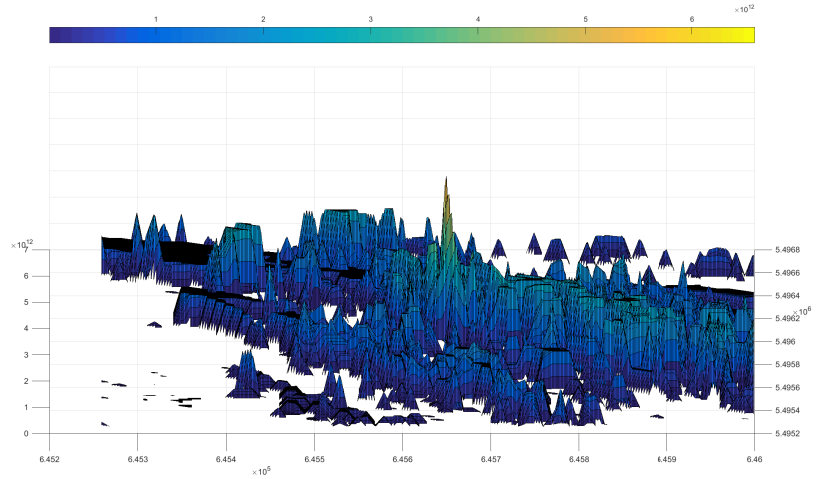

Figure 6.28: Visualization of the likelihood function defined in (6.113) at iteration 1.

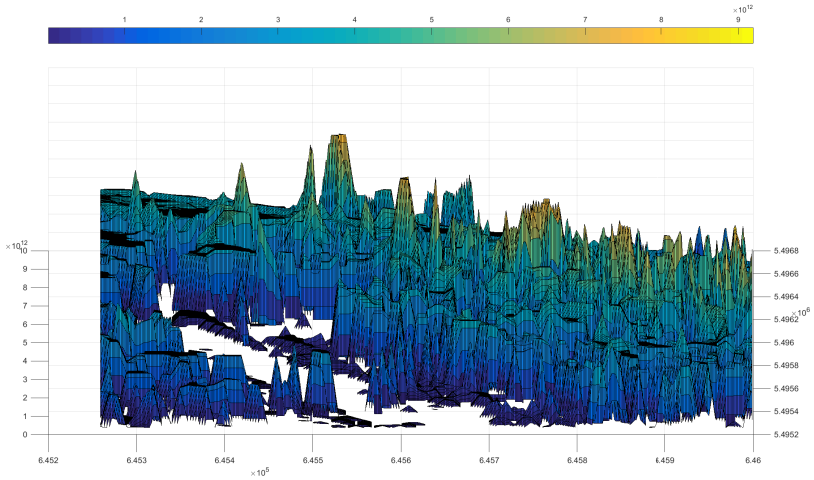

Figure 6.29: Visualization of the likelihood function defined in (6.113) at iteration 15.

6.3.5 Conclusion

In this section generalized PHD intensity filters are studied in terms of their applicability. First, the update equation of the generalized PHD and the generalized iFilter are presented. Thereby, it becomes obvious that the issue of numerical complexity needs to be solved before generalized PHD intensity filters can be applied to concrete scenarios. Furthermore, the issue of the need of a generalization of the probability of detection for such filters is identified. Then, two different ways of approximating the update equation of the generalized PHD and iFilter are proposed. In contrast to approximations for extended object and group tracking, the spatial relation of the measurements in the measurement space is not used. The approximations are based on incorporating the a priori knowledge on the number of measurements per target and the significance of a partition in terms of the likelihood function value. Therefore, the proposed approximations can be applied to the task of BML and other scenarios, where a spatial distribution of the measurements is not available. Furthermore, the detection process is modeled as a function of the target state and number of measurements. Then, the usage of the Binomial and the Poisson distribution for conditionally independent and identical distributed detection processes of the single measurements is motivated. Three numerical examples using simulated data for assessing the proposed methods are presented. First, a two–target scenario with multiple correlated measurements is carried out to compare the generalized PHD to the generalized iFilter. It is obvious that both filters show a comparable performance, which could have been expected, since both versions only vary in the way of modeling clutter. The

only remarkable difference between both filter versions is, that the generalized iFilter is capable to estimate the number of clutter measurements correctly, while the generalized PHD filter possesses the correct number of clutter measurements in every iteration. Second, the same scenario is used to show the applicability of the proposed approximation methods and to discuss the number of partitions that can be successfully reduced without decreasing the tracking performance. Several parameterizations are investigated and compared to each other. Third, a single–target BML scenario is investigated and the generalized PHD filter, using the proposed approximations and the generalization of the probability of detection, is compared to an adaption of the standard intensity filter in terms of runtime, the estimated number of targets and the RMSE performance.
It is demonstrated that generalized PHD intensity filters can be applied to BML only by applying certain approximations of reducing the number of partitions that have to be processed by the update equation. Furthermore, it is seen that the tracking performance of the generalized PHD filter is more robust than the adapted standard iFilter due to the fact that the information of multiple measurements is processed simultaneously in the generalized version. On the contrary, the computational complexity of generalized PHD intensity filters is much higher than the computational complexity of standard PHD intensity filters. In the concrete studied examples this results in a longer initialization phase, due to the fact that the number of particles that can be processed by the generalized PHD filter in an acceptable time are not sufficient to cover the FOV of the investigated scenario. However, when studying multitarget scenarios, partitions arise also for the adapted standard PHD intensity filters due to the enhanced target state extraction schemes needed, which increase the computational burden dramatically.
It is demonstrated that generalized PHD intensity filters are a versatile tool for solving BML scenarios, which need a numerically efficient implementation (e.g. parallel computing) to be applicable for tasks that need to be processed online.

6.4 Evaluation with Real World Data

In the previous sections adapted standard and generalized PHD Intensity filters are compared using the BML framework for simulated data. The corresponding BML framework is presented in Section 5.1.3 and visualized in Figure 5.3. There, the measurement data is created in the following way. First, a ray tracer produces a set of predicted multipaths based on the OS and the ground truth position of the MS for all time instances. Afterwards, these multipaths are deteriorated by an antenna emulation and then used as measurement data by the tracking algorithms.
In this section, the second proposed BML framework from Figure 5.4 for the processing of real world data is used to compare the generalized PHD filter (see Section 6.3) to the adaption of the standard iFilter (see Section 6.1). The data was collected in the

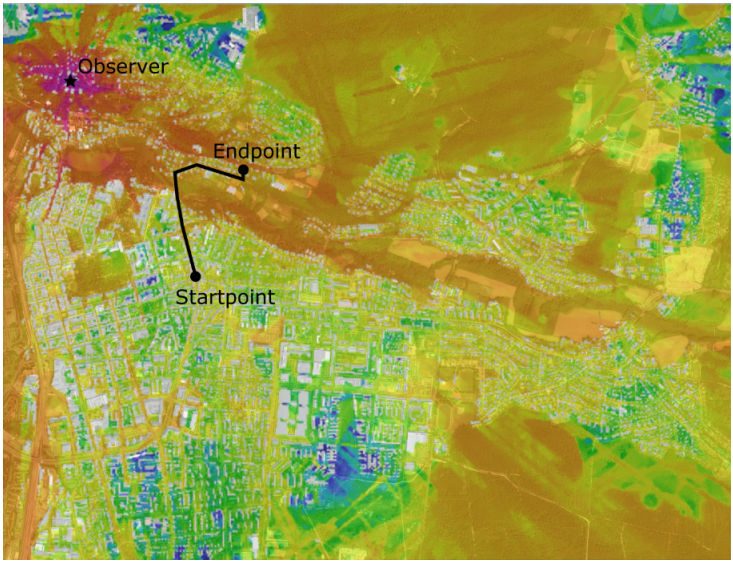

Figure 6.30: Visualization of the MS trajectory for the real world experiment (city of Erlangen/Germany). The colors indicate the received field strength that is received by the observer for the hypothetical emitter positions. Map Data: ©GeoBasis–DE/BKG 2015. Ray–Tracer Visualization: © 2015 AWE Communications.

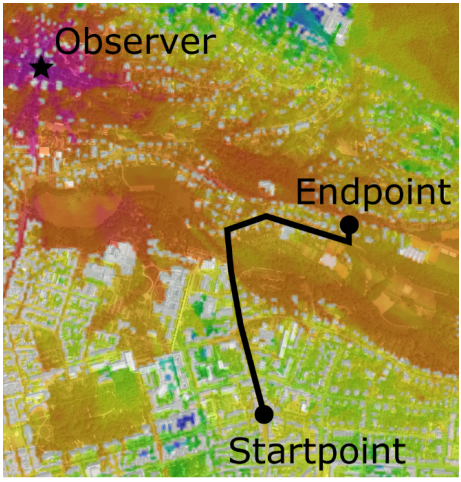

Figure 6.31: Zoom of the studied scenario. Map Data: ©GeoBasis–DE/BKG 2015. Ray–Tracer Visualization: © 2015 AWE Communications.

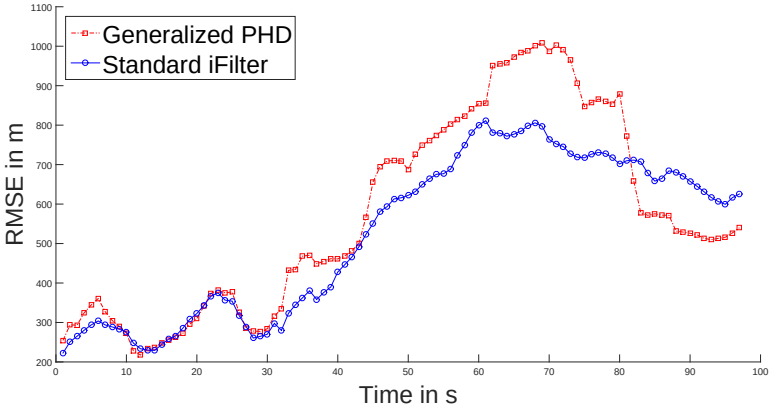

Figure 6.32: Comparison of the generalized PHD and the adapted standard iFilter in terms of the RMSE.

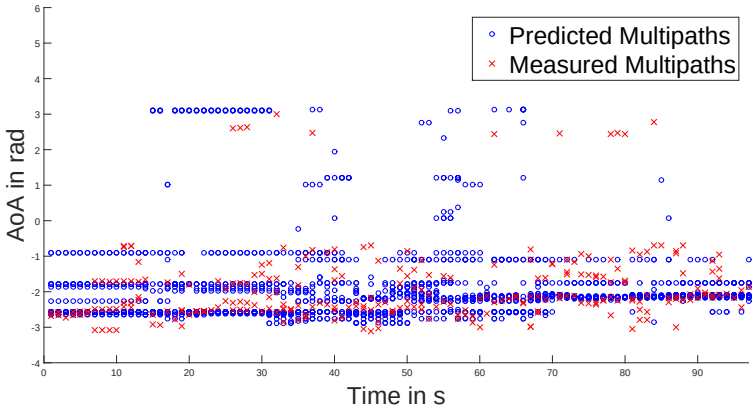

Figure 6.33: Comparison of the received multipaths to the predicted multipaths coming from the ray tracer for the ground truth position in terms of the AoA.

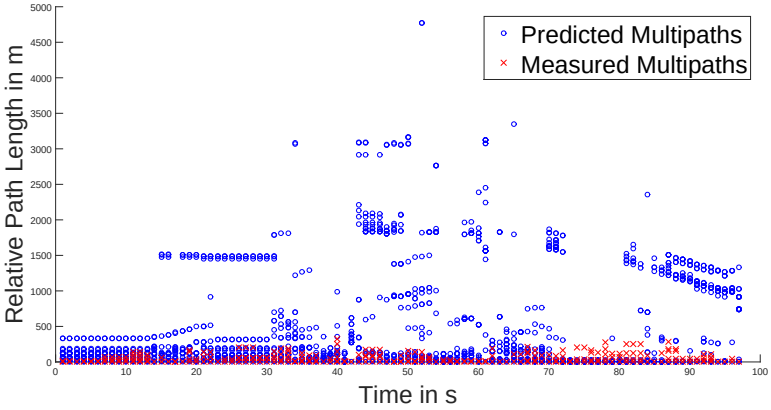

Figure 6.34: Comparison of the received multipaths to the predicted multipaths coming from the ray tracer for the ground truth position in terms of the relative path length.

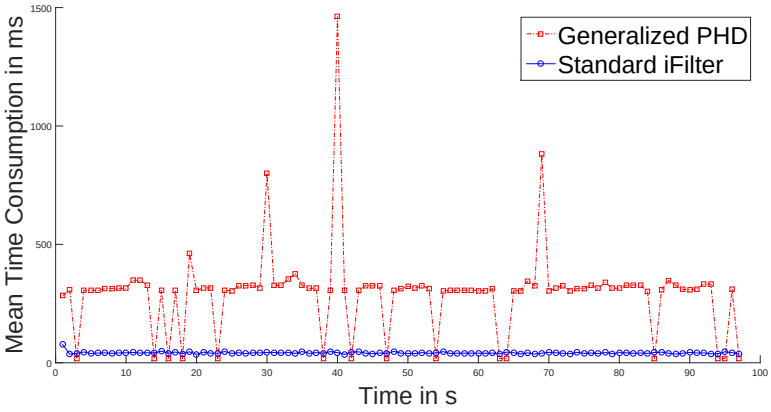

Figure 6.35: Comparison of the generalized PHD and the adapted standard iFilter in terms of the mean computation time per iteration.

Figure 6.36: Picture of the OS © 2013 Saab Medav Technologies GmbH. The antenna array is carried by a lifting platform with a height of 10m.

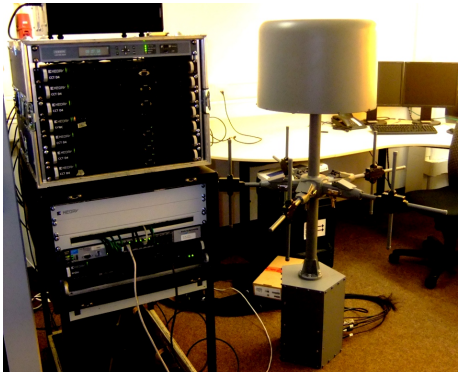

Figure 6.37: Picture of the antenna array (SYST-A0006 from Alaris Antennas) and the receiver setup © 2013 Saab Medav Technologies GmbH.

city of Erlangen. The trajectory of the MS and the location of the OS are depicted in the Figures 6.30 and 6.31.
The MS is given by a car, which is equipped with a signal generator, a power amplifier and an ommidirectional antenna to generate the emitted electromagnetic signal. Furthermore, the target is equipped with an GPS-recorder for evaluating the target state estimations of the data fusion algorithms with respect to accuracy. The emitted signal has a frequency of 475.560 MHz and a bandwidth of 5 MHz. Note that the bandwidth is small compared to the real world data evaluation carried out in [Alg10]. The power of the emitted signal is 10 Watt.
The five–element antenna array carried by the OS is a prototype from Alaris Antennas (SYST-A0006, see http://www.alarisantennas.com/ and Figures 6.36, 6.37), which is mounted on a lifting platform in a height of 10 m (see Figure 6.36). Due to this setup the OS is fixed for the considered scenario. However, in general a moving OS and therefore an online processing of the ray tracing prediction is possible for the proposed data fusion algorithms. Furthermore, the OS carries signal processing equipment including a direction finding system (see Figure 6.37). The received signal is then processed by the blind channel estimation algorithm proposed in [HKT15], which outputs a set of multipaths characterized by the AoA and the RToA.
In Figures 6.33 and 6.34 the received multipaths are compared in terms of their AoA and their relative path length (RPL) (which is equal to the RToA multiplied by the speed of light), respectively, to the predicted multipaths generated by the MS. Using this real world data the two SMC-implementations of the generalized PHD and the adapted standard iFilter from Section 6.3.4.3 are numerically compared. The only difference is given by the parameterization of the filters and the fact that the EoA is not processed in this evaluation. Therefore, only the differences to Section 6.3.4.3 are explained in the following. The standard deviations are set to $\sigma_\varphi = 1.0$ rad for the AoA and $\sigma_\tau = \frac{100.0}{\mathrm{c}}$ s for the RToA, where c denotes the speed of light.
The generalized PHD filter is implemented including the approximations proposed in (6.85) and (6.86), where $N_{\min} \equiv \min(m, 3)$, $N_{\max} \equiv 6$ and the threshold for significance of a partition $\tau \equiv 1.0 \cdot 10^{10}$. Here, m denotes the number of received multipaths. The FOV, that is, the area of possible MS locations which is predicted by the ray tracer is given by

$$\mathrm{FOV} \equiv [644489.0, 646999.0] \times [5494207.0, 5497557.0]. \tag{6.114}$$

The probability of detection of the adaption of the iFilter is given by $p^D(x) \equiv 0.75$ for all $x \in X$, where $X \equiv \mathrm{FOV} \times \mathbb{R}^2$. For the generalized PHD filter the probability of detection is modeled by (6.94), where $q \equiv p^D$. The occurrence of clutter in a set of multipaths is modeled by $\lambda_\Phi \equiv 0.01$, which is equal to the clutter density of the generalized PHD filter. The adaption of the standard iFilter estimates this quantity. The probability of detection in the hypothesis space X_ϕ is set to $p^D(\phi) = 0.3$. The

adaption of the standard iFilter uses the following transition probabilities. The transition probability from X_ϕ to the target state space X is set to $\Psi(x|\phi) = 0.3$, the transition probability in X_ϕ is defined as $\Psi(\phi|\phi) = 0.1$ and the transition probability from X to X_ϕ is given by $\Psi(\phi|x) = 0.3$. For the adaption of the standard iFilter the maximal number of particles is set to 1500 and for the generalized PHD filter the maximal number of particles is restricted by 700.
It can be seen from the RMSE–values after 50 MC runs (using the same data set) (see Figure 6.32) that both filters show a comparable performance until iteration 30. This is caused by the information quality of the received measurements (see Figure 6.33 and Figure 6.34), which is sufficient, such that the small number of particles used by the generalized PHD filter and the approximation criteria (6.85) and (6.86) do not decrease the performance of the generalized PHD filter compared to the standard iFilter adaption. In some iterations the criteria (6.85) and (6.86) imply that no partition is significant, which is caused by the small number of particles and the small width of the likelihood function defined in (6.107) and thus the computation time of the generalized PHD filter is smaller for these iterations than the computation time of the adaption of the standard iFilter. Furthermore, it can be seen from the RMSE visualization in Figure 6.32 that the RMSE of the generalized PHD filter is larger than the RMSE of the adapted iFilter for the iterations 30–80. For some MC runs of the evaluation the generalized PHD filter yields for this time period a smaller (around 200m) RMSE–value than the mean–RMSE of the adapted standard iFilter. However, due to product form of its likelihood function the generalized PHD filter is more restrictive than the adapted standard iFilter, which implies in combination with the small number of particles an instable behavior for these iterations. Additionally, in this time period the number of clutter in the received multipaths is rather high and due to non–adaptive clutter rate of the generalized PHD filter the adaption of the standard iFilter yields a more accurate estimate and is more stable, since it is able to estimate the clutter online. The situation changes drastically after iteration 85, when the generalized PHD filter is able to estimate the target state more accurately than the adapted standard iFilter. This is implied by the fact that the estimated multipaths better match to the predicted multipaths of the ground truth position coming from the ray tracer (see Figures 6.33, 6.34). The restrictive likelihood function of the generalized PHD filter then filters out rigorously hypothetical emitter positions with non–matching multipath predictions. This is the reason why the generalized PHD filter outperforms the standard iFilter adaption for the last ten iterations, even though the number of particles used by the generalized PHD filter is smaller than half the number of particles used by the adapted standard iFilter.
In summary it is shown that the proposed generalized and standard PHD intensity filter can be successfully applied to real world BML data, which is collected under demanding boundary conditions (small number of antenna arrays and small band-

width). The generalized PHD filter would clearly benefit from enlarging the number of particles and decreasing the approximation significance–threshold defined in (6.86). However, due to the high numerical complexity these approximating conditions are needed and imply an instable behavior of the generalized PHD filter under real world conditions. Future work therefore should concentrate on reducing the numerical complexity without approximating the update equation. One approach is parallel computation of the update equation of the generalized PHD filter for different particles and/or partitions.

6.5 Conclusion and Future Work

This chapter is an extension and generalization of the work on passively tracking an electromagnetic emitter under multipath propagation presented in [Alg10] to the class of pointillist filters which superpose targets (see Section 3.4). As representatives of this class of pointillist filters standard and generalized PHD intensity filters are applied to BML. Due to the fact that BML implies a non–standard target–oriented measurement model, that is, *a target generates more than one measurement per sensor scan* standard PHD intensity filters and likelihood functions have to be generalized appropriately. Furthermore, generalized PHD intensity filters (see Section 3.4.4), which apply a generalized target–oriented measurement model, are in their original form numerically highly complex. An application of these filters to BML is only possible if the filter update is approximated by a sophisticated reduction of the partitions. The main challenge thereby is to reject partitions without using a distribution information of the measurements, since multipaths which are created by the same MS are typically not spatially related in the measurement space.
By processing each multipath as an individual measurement, formulating a likelihood function based on an assignment approach for a single multipath and deriving an enhanced target state extraction scheme, an SMC–implementation of standard PHD intensity filters is successfully applied to the task of BML. Furthermore, an alternative likelihood function for BML, which is defined on single multipaths, is formulated using an adaption of the decomposition of a bearing–only likelihood function. Approximation schemes which do not make use of the spatial distribution of measurements and a generalization of the probability of detection enable the numerically highly complex PHD intensity filters to solve the task of BML. Thus, it is shown that either standard or generalized versions of PHD intensity filters can be applied to BML and yield satisfactory results using either simulated or real world data. In particular it is shown, that the derived data fusion algorithms can be applied to real world data which is received with a small bandwidth, an open question formulated in [Alg10].
In future the application of other pointillist filter might imply an improvement of the tracking results. Especially the ease of designing and deriving new pointillist filters for the non–standard challenge of tracking mobile electromagnetic emitters under

multipath propagation using the framework from Chapters 3 and 4 is a promising approach. Also the application of hybrid pointillist filters (see Section 3.5) should be considered due to the partial resolvability of targets in cluttered urban environments. Even though it is assumed in [Alg10] that signals of different MS can be separated due to their signal form, scenarios might exist, where this assumption is not valid. Thus, future work will also investigate the application of the proposed methods in a multitarget BML scenario. In such scenarios an additional level of the assignment problem arises.

Furthermore, the development of more general clutter models for pointillist filters that apply a generalized target–oriented measurement model should be studied in future. In this chapter clutter is assumed to be generated as single elements in the measurement space. In BML however, typical origins of clutter are objects, which are not stored in the database of the urban environment of the ray tracer. Such objects can be pedestrians, cars, trees, new buildings, etc. Since such objects create correlated clutter measurements, clutter should also occur due to clutter scatterers. For the studied filters such enhanced clutter models can be realized by either adapting the likelihood function on the hypothesis space (generalized iFilter) or by introducing a generalized clutter model (generalized PHD filter).

The derivation of further approximations of the update equations of generalized pointillist filters for BML will also be part of future work.

In this work, the blind channel parameter estimation [HKT15] is done independently of the tracking algorithm. However, BML data fusion algorithms would also benefit from a simultaneous target tracking with combined parameter estimation using the received raw signal analogously to the publications on the STAMP algorithm proposed in [Li14] and [LK14].

Mission planning for scenarios with movable OS is a further topic for future work. This challenging task ends up with an optimization problem and should imply a significant improvement of the tracking results in scenarios with a moving OS.

CHAPTER VII

Parameter Tracking for BML

The performance of data fusion algorithms for BML is deteriorated by missing and false multipaths. The BML measurement set consists out of several multipaths which are usually parametrized by their AoA (, sometimes the EoA) and the RToA. Measurement failure in BML occurs for instance if the number of elements of the antenna array is smaller than the number of emitted multipaths, due to fading or resolution conflicts of the antenna array. False multipaths can occur due to moving obstacles like pedestrians, cars, trucks, etc. or due to static objects which are not contained in the building map like buildings, trees, etc. In general one can say that every object which is not contained in the context information of the urban environment used by the ray tracer is a potential origin of false multipaths.
The likelihood functions defined in (6.13) and (6.107) might yield in presence of a high number of clutter or false multipaths ambiguous or completely false information on the position(s) of specific MS(s). In order to improve the information about the received multipaths, this chapter studies the application of algorithms for a tracking of the multipaths in the parameter space, that is, a tracking in the domain of AoA and RToA (all presented algorithms can easily be extended additionally to EoA as well). Since parameter tracking enlarges the probability of detection and decreases the false alarm probability, the quality of the multipaths processed by the data fusion algorithm is increased [Alg10].
This chapter is structured as follows. In Section 7.1 related work for parameter tracking in the context of BML is presented and potential issues in existing approaches are discussed. An MHT algorithm for parameter tracking, which is based on marginalizing over a set of clutter hypothesis, is proposed in Section 7.2. Finally, the conclusions are drawn in Section 7.3

Own work on this subject The proposed parameter tracking from Section 7.2 using an MHT with additional clutter hypothesis is first published in [DGK13b]

7.1 Existing Work

In [SRE+06] and [Sal09] the extended Kalman filter (EKF) is used for the parameter tracking and in [CTW+07] the tracking of clustered multipath parameters is proposed. In order to improve the results of BML–suited data fusion algorithms a parameter tracking in the measurement space, that is, in the AoA and RToA domain using an MHT is proposed in [ADKT08a], where a bank of standard Kalman filter are applied to solve the parameter tracking. Furthermore, in [ART04] an alternative approach using a maximum likelihood batch estimator is derived. An extension and improvement of both works is given in [Alg10]. There, a standard approach from multitarget tracking, that is, an MHT algorithm [Koc10] is applied to the multipath data. The approach is verified using synthetic multipath measurements and the results seem to be satisfactory unless the scenario does not have a too high number of clutter multipaths or too dense multipath components [Alg10].
The measurement function applied in the proposed standard track–oriented MHT from [Alg10] is given by the identity function. Furthermore, the way of creating and processing hypothesis is standard compared to multitarget tracking MHT algorithms. In particular the RToA information is used without any additional processing. However, a problem with the application of this standard approach occurs if clutter multipaths are received that possess a negative RToA, that is, the respective clutter measurement is received before all target–related measurements by the OS. A clutter multipath with negative RToA implies that all target–related RToAs are increased (with the difference between the RToA of the first target–related and the first received multipath, see Figure 7.1). Thus, all likelihood functions defined in this thesis and from [Alg10] are deteriorated due to an error in the assignment between the received and the predicted multipaths.
Since the parameter tracking is applied before the target detection it is closely related to TBD approaches [Alg10]. An overview of existing TBD approaches can be found in [Gov12] and [DRC08]. TBD approaches can essentially be separated into dynamic programming algorithms (DPA) [ASP93], particle filter approaches [RRG05], Hough space transformation [Ric96] and subspace data fusion fusion [DOR08].
In the following, an extension of an MHT algorithm for parameter tracking that uses an additional marginalization over clutter hypothesis is proposed to solve the problem of false associations in between the predicted and measured multipaths.

7.2 MHT–Parameter Tracking Using Clutter Hypothesis

If a clutter multipath is received before the first measurement of a target, that is, if it possesses a negative RToA compared to the target–related measurements, the computation of an assignment–based likelihood function is deteriorated and thus the accuracy of any BML data fusion algorithm decreases.
In this section, a pre–processing of the measurement set by an application of an MHT in the parameter space is proposed. Therefore, two extensions of the MHT, processing additional global clutter hypothesis are derived. Finally, a ray tracing simulation is used to numerically assess the proposed methods for different clutter levels in terms of the OSPA metric. In this section parameter tracking via MHT in AoA and RToA is considered in a completely blind scenario, that is, the possibility exists that a clutter multipath might possess a smaller RToA than all target–related multipaths. To handle this deterioration of the RToA of the target measurements the incorporation of global clutter hypothesis into a track–oriented MHT is derived by applying marginalization to the posterior. Based on the derived filtering equation two approaches are proposed for determining the clutter hypothesis weights. Finally, the proposed MHT–extensions are numerically compared against the ordinary MHT algorithm, using a ray tracing simulation and the OSPA–metric.
This section is organized as follows. Section 7.2.1 describes the problem formulation. The problem of clutter multipaths which are received before the first target–related measurement is described in Section 7.2.1.1. In Section 7.2.2 the incorporation of global clutter hypothesis into the MHT algorithm is presented. The derivation of the respective MHT filtering equations is given in Section 7.2.2.1. The following two subsections study the definition of the clutter hypothesis weights. In Section 7.2.2.2 the weights are defined via a combinatorial approach, in Section 7.2.2.3 the weighting is done by an assignment procedure, using the OPSA–metric. A numerical evaluation of the different MHT versions can be found in Section 7.2.3.

7.2.1 Formulation of the Problem

7.2.1.1 Problem of Clutter Multipaths

The underlying principle of existing data fusion algorithms presented in Chapter 6 and [Alg10] is to assess hypothetical positions of electromagnetic emitters with respect to the current set of measurements by the computation of the likelihood function. The value of the likelihood function is computed by assigning the predicted multipaths of the ray tracer for a hypothetical emitter to the received measurements. A correct assignment between the predicted and measured multipaths is deteriorated or becomes even impossible if clutter measurements (with arbitrary AoA and EoA) have a smaller RToA than every target–related measurement (see Figure 7.1), which are called measurements with negative RToA. This is due to the fact that clutter cannot

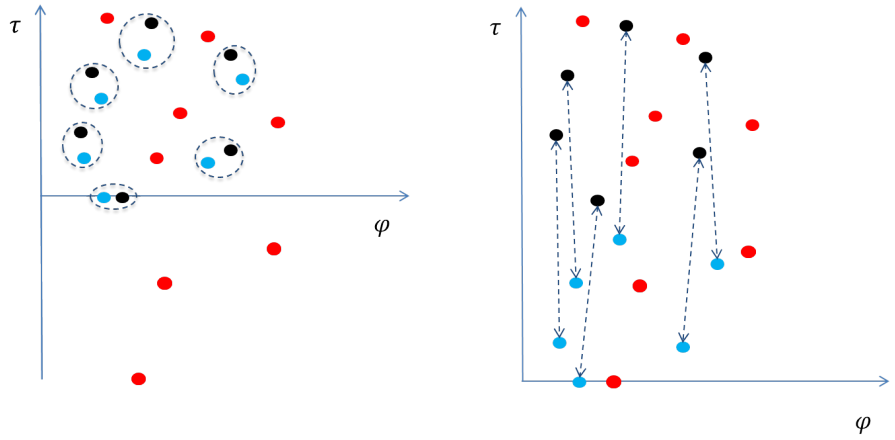

Figure 7.1: Left: Without clutter measurements (red) that possess a negative RToA a correct association between the ray tracing prediction (blue) and measurements (black) is possible (even under the influence of clutter with a positive RToA). Right: Clutter with negative RToA yields to a shift between the measurements and the ray tracing prediction and prevents a correct association © 2013 IEEE.

be modeled by the ray tracer, which implies that the RToAs of the measurements are shifted in RToA compared to the predicted multipaths (see Figure 7.1). Note that clutter which possesses a positive RToA compared to the first target–related measurement might also cause non–correct assignments between the measured and predicted multipaths (a clutter measurement might be closer to a predicted multipath then the correct target–related measurement). However, an application of a standard assignment algorithm [BL71] is still possible in the presence of clutter with positive RToAs (see Figure 7.1, left). Thus, the focus of this work is on clutter measurements with negative RToA.
The main idea to guarantee a correct assignment under the influence of clutter with negative RToA is to consider additionally the hypothesis that the first n received measurements, $n \in \{1, ..., m_k\}$ are false measurements. In the following section an extension of the track–oriented MHT algorithm for tracking multipaths in the AoA–RToA domain is proposed using global clutter hypothesis.

7.2.2 MHT–Parameter Tracking in AoA and RToA

In this section the problem of clutter with negative RToA is solved by introducing clutter–hypothesis and incorporating them into the filtering equation by marginalization. In Section 7.2.2.1 the clutter–hypothesis are introduced and it is explained why

their definition is exhaustive. Furthermore, the filtering equation is derived, where additional clutter–hypothesis weights can be found in each summand. In Section 7.2.2.2 and Section 7.2.2.3 the clutter–hypothesis weights are defined using a combinatorial approach and the OSPA–metric, respectively.

7.2.2.1 Derivation of the MHT–Filtering Equations

In the following the MHT filtering equations are proposed to perform tracking of multipaths with parameters in AoA and RToA. Without loss of generality the elements of Z_k are assumed to be sorted with respect to their RToA. To handle clutter in the AoA–RToA domain, i_C denotes the event that the first i elements of the current measurement set Z_k are clutter in a global sense, $i \in \{1, ..., m_k\}$. Global means that the event i_C implies that the first i elements of Z_k are clutter without stating any relation to a specific estimated target state $x_k \in X \subset \mathbb{R}^d, \quad d > 0$ from time step k. In addition, $\neg i_C$ denotes the corresponding negated event, that is, the event that at least one of the first i measurements is not clutter. This implies that the joint event $\{(i-1)_C, \neg i_C\}$ indicates that the first $(i-1)$ elements of Z_k are false alarms and the ith incoming element of Z_k is a target–related measurement. In particular, this does not imply that the ith element of the set of measurements is originated due to x_k. Note that only clutter measurements with negative RToA are problematic for the assignment–based data fusion algorithms presented in Chapter 6 and [Alg10]. Therefore, it is only important to know whether clutter measurements are received before the first target–related measurement. The clutter hypothesis i_C, $i \in \{1, ..., m_k\}$ cover these events.
In the following the posterior density is investigated in detail for a specific target state x_k. It is given by

$$p(x_k|Z^k) = p(x_k, \neg 1_C|Z^k) + p(x_k, 1_C|Z^k) \tag{7.1}$$

$$= p(x_k, \neg 1_C|Z^k) + p(x_k, \neg 2_C, 1_C|Z^k) + p(x_k, 2_C|Z^k) \tag{7.2}$$

$$= p(x_k, \neg 1_C|Z^k) + \sum_{i=2}^{m_k} p(x_k, \neg i_C, (i-1)_C|Z^k) + p(x_k, m_{kC}|Z^k), \tag{7.3}$$

where (7.1) and (7.2) are the results of applying marginalization to the events x_k and $\{x_k, 1_C\}$, respectively and (7.3) is given due to iterative marginalization on the events $\{x_k, i_C\}$ for $i \in \{2, ..., m_k\}$. Note, that (7.2) holds, since the event 2_C can only occur if the first and the second multipath are clutter. Next, Bayes theorem can be applied

to (7.3), which implies

$$
\begin{aligned}
p(x_k|Z^k) =& p(x_k, \neg 1_C|Z^k) + \sum_{i=2}^{m_k} p(x_k, \neg i_C, (i-1)_C|Z^k) + p(x_k, m_{kC}|Z^k) && (7.4)\\
\propto& p(Z_k|x_k, \neg 1_C) \cdot p(x_k, \neg 1_C|Z^{k-1})\\
&+ \sum_{i=2}^{m_k} p(Z_k|x_k, \neg i_C, (i-1)_C) \cdot p(x_k, \neg i_C, (i-1)_C|Z^{k-1})\\
&+ p(Z_k|x_k, m_{kC}) \cdot p(x_k, m_{kC}|Z^{k-1}) && (7.5)\\
=& p(Z_k|x_k, \neg 1_C) \cdot p(x_k|\neg 1_C, Z^{k-1}) \cdot p(\neg 1_C|Z^{k-1})\\
&+ \sum_{i=2}^{m_k} p(Z_k|x_k, \neg i_C, (i-1)_C) \cdot p(x_k|\neg i_C, (i-1)_C, Z^{k-1})\\
&\times p(\neg i_C, (i-1)_C|Z^{k-1})\\
&+ p(Z_k|x_k, m_{kC}) \cdot p(x_k|m_{kC}, Z^{k-1}) \cdot p(m_{kC}|Z^{k-1}) && (7.6)\\
\approx& p(Z_k|x_k) \cdot p(x_k|\neg 1_C, Z^{k-1}) \cdot p(\neg 1_C|Z^{k-1})\\
&+ \sum_{i=2}^{m_k} p(Z_{k,i}|x_k) \cdot p(x_k|\neg i_C, (i-1)_C, Z^{k-1})\\
&\times p(\neg i_C, (i-1)_C|Z^{k-1})\\
&+ p(\emptyset|x_k) \cdot p(x_k|m_{kC}, Z^{k-1}) \cdot p(m_{kC}|Z^{k-1}) && (7.7)\\
=& p(Z_k|x_k) \cdot p(x_k|Z^{k-1}) \cdot p(\neg 1_C|Z^{k-1})\\
&+ \sum_{i=2}^{m_k} p(Z_{k,i}|x_k) \cdot p(x_k|Z^{k-1}) \cdot p(\neg i_C, (i-1)_C|Z^{k-1})\\
&+ p(\emptyset|x_k) \cdot p(x_k|Z^{k-1}) \cdot p(m_{kC}|Z^{k-1}), && (7.8)
\end{aligned}
$$

where (7.6) holds due to the definition of conditional probability. The approximation of the likelihood function in (7.7) is justified by the following idea. Given the event that the first $(i-1)$ elements of Z_k are clutter and the i-th element is a target–related measurement it approximately suffices to consider $Z_{k,i}$, where the first $(i-1)$-measurements are removed (see Definition (6.5)). Equation (7.8) is valid since the target state x_k of the current iteration is independent of the event that a specific measurement is clutter. This is valid since the event i_C is a global assumption on the measurement set Z_k and does not provide any information about the relation of state x_k to the measurement set. Note that the pairwise statistical independence of the events $\{\neg i_C\}$ and $\{m_{kC}\}$ with $\{x_k\}$ given Z^{k-1} is due to the global definition of the clutter events $i_C, \ \ i \in \{1, ..., m_k\}$.

Since we assume that $P(\emptyset|x_k) \equiv 0$, (7.8) reduces to

$$p(x_k|Z_k) \approx p(Z_k|x_k) \cdot p(x_k|Z^{k-1}) \cdot p(\neg 1_C|Z^{k-1}) + \sum_{i=2}^{m_k} p(Z_{k,i}|x_k) \cdot p(x_k|Z^{k-1}) \cdot p(\neg i_C, (i-1)_C|Z^{k-1}) \tag{7.9}$$

To evaluate (7.9) it remains to compute

$$p(\neg 1_C|Z^{k-1}) = p(\neg 1_C) \tag{7.10}$$

and

$$p(\neg i_C, (i-1)_C|Z^{k-1}) = p(\neg i_C, (i-1)_C). \tag{7.11}$$

If it is assumed that clutter in the current iteration is independent of measurements from the previous iterations (7.10) and (7.11) hold.

7.2.2.2 Combinatoric Clutter–Hypothesis Weighting

Motivated by the fact that clutter occurs randomly in the measurement set one approach to define the desired quantities from Equation (7.10) and (7.11) is to apply combinatorics. Let $n \in \{0, ..., m_k\}$ denote the number of clutter measurements in iteration k. Then marginalization with respect to n yields

$$p(x_k|Z^k) = \sum_{n=0}^{m_k} p(x_k|Z^k, n) \cdot p(n). \tag{7.12}$$

From now on let $n \in \{0, ..., m_k - 1\}$ be fixed. For $n = m_k$ the measurement set contains only clutter and thus $P(i_C) = 0$ for all $i \in \{1, ..., m_k\}$.
The occurrence of clutter can be modeled by drawing n times out of a set of m_k distinguishable elements without repetition. The m_k distinguishable elements represent the measurements of iteration k sorted and numbered with respect to their RToAs. A clutter measurement is then represented by a drawn element. This can be motivated by the following consideration.
In equation (7.12) marginalization is applied with respect to the number of clutter. Therefore, equations (7.10) and (7.11) only have to be computed given the number of clutter. Hence, each measurement has the same probability to be clutter. Thus, (7.10) and (7.11) can be defined by counting the combinations which represent the events $\{1_C\}$ and $\{\neg i_C, (i-1)_C\}$.
First of all the total number of combinations is given by

$$\binom{m_k}{n}. \tag{7.13}$$

Hence the probability of one combination is

$$\binom{m_k}{n.}^{-1}. \tag{7.14}$$

Therefore, the probability of a specific event can be determined by counting all combinations that represent the desired event. The event that the first received measurement is clutter is thus given by

$$P(1_C|n) = \frac{\binom{m_k}{n} - \binom{m_k - 1}{n}}{\binom{m_k}{n}} \tag{7.15}$$

and hence

$$P(\neg 1_C|n) = 1 - \frac{\binom{m_k}{n} - \binom{m_k - 1}{n}}{\binom{m_k}{n}}. \tag{7.16}$$

Equation (7.15) holds since

$$\binom{m_k - 1}{n} \tag{7.17}$$

denotes the number of combinations without the first element and thus

$$\binom{m_k}{n} - \binom{m_k - 1}{n} \tag{7.18}$$

counts the number of combinations containing the first element. In the following $m_k \geq 2$ is assumed. It holds that

$$P(2_C|n) = P(1_C|n) \cdot \frac{\binom{m_k - 1}{n - 1} - \binom{m_k - 2}{n - 1}}{\binom{m_k - 1}{n - 1}}, \tag{7.19}$$

since $\{2_C\} = \{2_C, 1_C\}$ by definition.
The fact that the event of having clutter in the second measurement is independent of having clutter in the first element justifies the multiplication. Only the number of drawn elements and the total number of measurements have to be subtracted by

one. Analogously to the previous considerations the probability of having clutter in the first but not in the second measurement is given by

$$P(\neg 2_C, 1_C|n) = P(1_C|n) \cdot \left(1 - \frac{\binom{m_k - 1}{n-1} - \binom{m_k - 2}{n-1}}{\binom{m_k - 1}{n-1}} \right). \tag{7.20}$$

Thus, the probability of the event $\{i_C\},\ \ i \in \{2, .., n\}$ can be expressed by

$$\begin{aligned} P(i_C|n) =& P((i-1)_C|n) \cdot \frac{\binom{m_k - (i-1)}{n-(i-1)} - \binom{m_k - i}{n-(i-1)}}{\binom{m_k - (i-1)}{n-(i-1)}} \\ =& \left(\prod_{l=1}^{(i-1)} P(l_C|n) \right) \cdot \frac{\binom{m_k - (i-1)}{n-(i-1)} - \binom{m_k - i}{n-(i-1)}}{\binom{m_k - (i-1)}{n-(i-1)}} \end{aligned} \tag{7.21}$$

and hence

$$P((i-1)_C, \neg i_C|n) = \left(\prod_{l=1}^{(i-1)} P(l_C|n) \right) \cdot \left(1 - \frac{\binom{m_k - (i-1)}{n-(i-1)} - \binom{m_k - i}{n-(i-1)}}{\binom{m_k - (i-1)}{n-(i-1)}} \right). \tag{7.22}$$

Therefore, $p(x_k|Z^k, n)$ from (7.12) is given by

$$\begin{aligned} p(x_k|Z^k, n) =& p(Z_k|x_k, n) \cdot p(x_k|Z^{k-1}) \cdot \left(1 - \frac{\binom{m_k}{n} - \binom{m_k - 1}{n}}{\binom{m_k}{n}} \right) \\ &+ \sum_{i=2}^{m_k} p(Z_{k,i}|x_k, n) \cdot p(x_k|Z^{k-1}) \cdot \left(\prod_{l=1}^{(i-1)} P(l_C) \right) \\ &\times \left(1 - \frac{\binom{m_k - (i-1)}{n - (i-1)} - \binom{m_k - i}{n - (i-1)}}{\binom{m_k - (i-1)}{n - (i-1)}} \right), \end{aligned} \tag{7.23}$$

since $p(x_k|Z^{k-1}) = p(x_k|Z^{k-1}, n)$.

7.2.2.3 OSPA Clutter–Hypothesis Weighting

An alternative approach to determine $p(\neg 1_C)$ and $p(\neg i_C, (i-1)_C)$ is to apply an assignment algorithm. To assess the events $\{(i-1)_C, \neg i_C\}$, $i \in \{2, .., n\}$, the probability that the first $(i-1)$ elements of Z_k are clutter but the ith measurement belongs to a target is expressed by using the spatial relation between the established target state estimates $\{x_1, ..., x_N\}$ of the MHT from the previous iteration and Z_k. Therefore, an appropriate metric for point sets is needed.
The OSPA–metric [SVV08] is originally developed for assessing multitarget tracking filters. It essentially consists out of two summands. One determines the localization and the other describes the cardinality error. For its definition let $X \equiv \{x_1, ..., x_m\}$ and $Y \equiv \{y_1, ..., y_n\}$ be two finite subsets and $m, n \in \mathbb{N}_0$. Denote by $\Pi_{(1:k)}$ the set of permutations on $\{1, 2, ..., k\}$ and let $1 \leq p < \infty$ and $c > 0$. Finally, $d^{(c)}(x, y) \equiv \min(c, d(x, y))$ defines the distance between $x, y \in \mathbb{R}^2$ cut–off at c. Then, the OSPA–metric is defined by

$$\overline{d}_p^{(c)}(X, Y) \equiv \left(\frac{1}{n} \left(\min_{\pi \in \Pi_{(1:n)}} \sum_{i=1}^{m} d^{(c)}(x, y_{\pi(i)})^p + c^p(n - m) \right) \right)^{\frac{1}{p}} \tag{7.24}$$

if $m \leq n$ and $d_p^{(c)}(X, Y) \equiv d_p^{(c)}(Y, X)$ if $m > n$.
Now, (7.10) and (7.11) are defined using the OSPA–metric. Since the MHT filter rejects clutter, established tracks can be used to find out which clutter hypothesis is correct. The idea is thus to compare the established tracks with $Z_{k,i}$, $i \in \{2, ..., m_k\}$

and Z_k using the OSPA–metric. Since a small OSPA–value indicates a good match between two point sets, the probabilities $p(\neg 1_C)$ and $p(\neg i_C, (i-1)_C)$ are defined by

$$p(\neg 1_C) \equiv \frac{\overline{d}_p^{(c)^{-1}}(Z_k, \{x_1, ...x_N\})}{\sum_{j=1}^{m_k} \overline{d}_p^{(c)^{-1}}(Z_{k,j}, \{x_1, ..., x_N\})}, \tag{7.25}$$

$$p((i-1)_C, \neg i_C) \equiv \frac{\overline{d}_p^{(c)^{-1}}(Z_{k,i}, \{x_1, ...x_N\})}{\sum_{j=1}^{m_k} \overline{d}_p^{(c)^{-1}}(Z_{k,j}, \{x_1, ..., x_N\})}, \tag{7.26}$$

$$\tag{7.27}$$

$i \in \{0, ..., m_k\}$. If no established target states are available the probabilities of the clutter hypothesis are equally weighted and normed.

7.2.3 Numerical Evaluation

To assess the proposed approach for parameter tracking with respect to accuracy a numerical evaluation is carried out. An implementation of a track–oriented MHT [Koc10] based on the Kalman filter including merging and pruning for tracks and hypothesis is used to realize the extensions from the previous section by incorporating clutter hypothesis and the two proposed approaches for computing their weights. To create the test data a ray tracing simulation is used: A database is generated for a given OS, a fixed grid of MS locations, a city vector–database and a radio channel parametrization using a ray-tracer model analogously to [HWLW03]. Thus, each MS possesses a specific number of multipaths, consisting out of AoA and RToA, which are received by the OS. Afterwards, a linear ground truth for a target with constant velocity is created. For each time–step the lower–left grid point of the box, in which the target is located, is determined and the corresponding multipaths from the database are stored as measurements of the respective iteration into the ground truth of the parameter space. Finally, measurement failure, clutter and measurement noise is added to the ground truth in the parameter space and the result represents the measurement set. Without clutter and measurement failure the maximal number of received multipaths is 20, the minimal number is 11 and the average number of paths is 17.63. The measurement noise is set to $\sigma_\varphi = 0.0087$ rad and $\sigma_\tau = \frac{5.0m}{c_{\text{light}}}$, where c_{light} denotes the speed of light and the probability of detection is set to $p^D(x) = 1$, for all $x \in X \subseteq \mathbb{R}^2$, where X denotes the target state space. The number of clutter is Poisson distributed with mean $\lambda > 0$ in $[-\pi, \pi] \times [-0.5 \cdot 10^{-6}s, 0.5 \cdot 10^{-6}s] \subset \mathbb{R}^2$. A negative RToA for a specific clutter measurement implies that it is received by the OS before all target–related measurements. After adding clutter to the measurement set the target–related multipath parameters have to be adapted by adding the smallest clutter–RToA to the RToA of each target–related path. To investigate the influence of false measurements, different clutter levels are considered, that is, $\lambda \in \{0, 1, 3, 5, 7\}$.

Each MHT approach is parametrized in the same way: The detection probability is set to $p^D(x) = 0.99$, for all $x \in X$, the measurement noise is set to $\sigma_\varphi = 0.0087$ rad and $\sigma_\tau = \frac{5m}{c_{\text{light}}}$, the false measurement density is set to $\rho_F = 10^{-6}$, the values for the likelihood ratio (LR) tests are set to $A = 0.5$, $B = 10^{20}$. The covariance matrix is constant for all iterations and set to $C \equiv \text{diag}[\sigma_\varphi^2 \ \sigma_\tau^2]$. Furthermore, merging and pruning of hypothesis and tracks is done for each approach using the same parameters.
For the numerical evaluation 250 MC runs of the presented scenario are performed. To compare the results with respect to accuracy the OSPA–metric [SVV08] is computed for the ground truth and the established tracks from the filter, that is, tracks with an LR larger than B. The cut–off value of the OSPA–metric is set to $c = 50$ and the order is set to $p = 2$. The results are shown in Figures 7.2 – 7.6. Obviously, the approach using the combinatoric clutter hypothesis weighting yields for all clutter levels a better result than the ordinary MHT approach. Furthermore, the assignment clutter hypothesis weighting approach yields for $\lambda = 1$ performance comparable to the ordinary MHT algorithm. For higher clutter levels it performs better than the ordinary MHT approach and comparable to the combinatoric clutter hypothesis approach. Note that even if no clutter is present (Figure 7.2) the performance of the combinatoric approach is comparable to the performance of the ordinary MHT. For a clutter level of $\lambda = 7$ the OSPA clutter–hypothesis weighting yields a better result than the combinatoric clutter–hypothesis weighting. This is due to the fact that the OSPA–approach takes knowledge about the past (established tracks) into account and does not ignore the estimated information about the established tracks like the combinatoric approach.

7.3 Conclusion

As mentioned in Chapters 5 and 6 an inherent problem of BML is that no synchronization between the MS and OS can be assumed. Therefore, only RToAs can be measured which yields to the fact that clutter which arrives the OS before the first target–related measurement deteriorates the RToA of each measurement. Related work on the topic of parameter tracking in BML does not cover this fact (see Section 7.1). Thus, in this chapter a clutter hypothesis extension of a track–oriented MHT for parameter tracking in AoA and RToA is presented. First, the extended filtering equation is derived via marginalization over global clutter hypothesis. This extension yields a posterior which takes the possibility into account that clutter measurements are received by the OS before any target–related measurement. To weight each clutter hypothesis the probability that specific elements of the measurement set are clutter has to be determined. Two approaches for defining these probabilities are proposed. First, a combinatorial approach is presented, which models the occurrence of clutter in the measurement set by drawing the expected number of clutter out

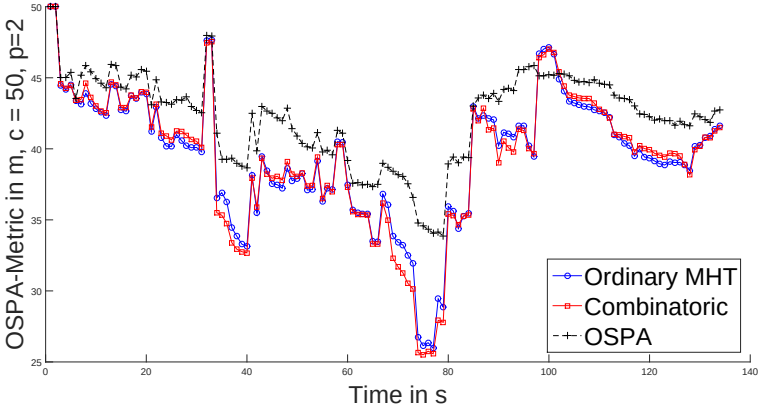

Figure 7.2: Even in the situation where no clutter is present ($\lambda = 0$) the combinatoric hypothesis–weighting approach performances comparable to the ordinary MHT approach © 2013 IEEE.

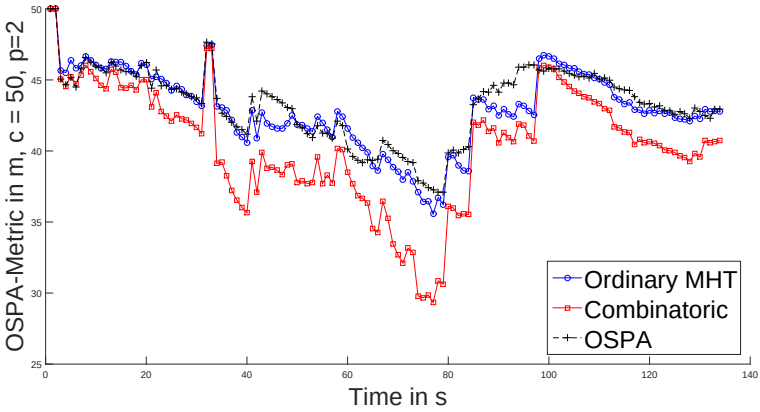

Figure 7.3: OSPA–value for $\lambda = 1$ © 2013 IEEE.

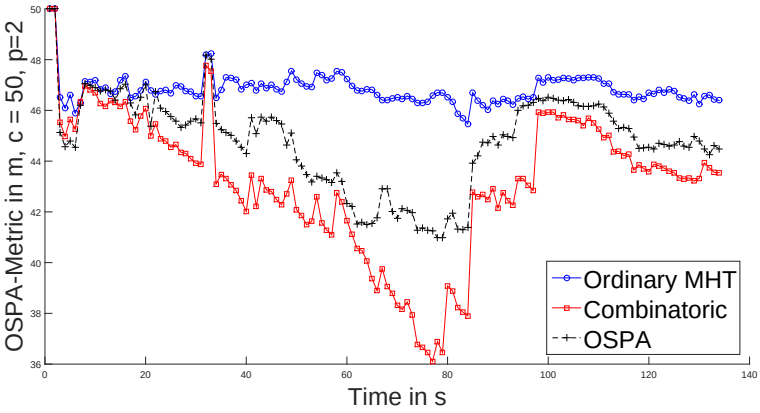

Figure 7.4: OSPA–value for $\lambda = 3$ © 2013 IEEE.

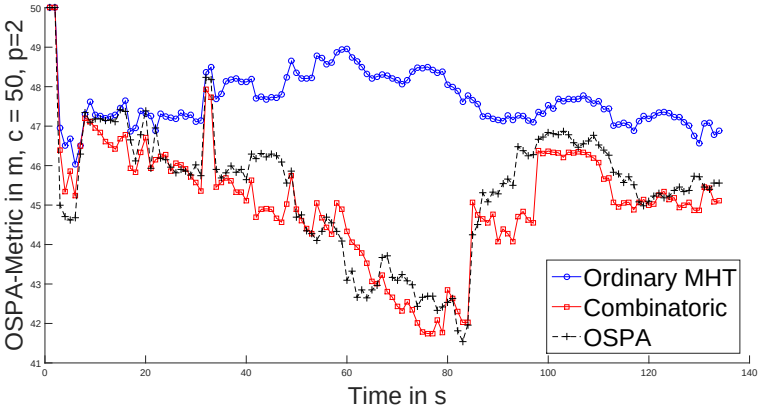

Figure 7.5: OSPA–value for $\lambda = 5$ © 2013 IEEE.

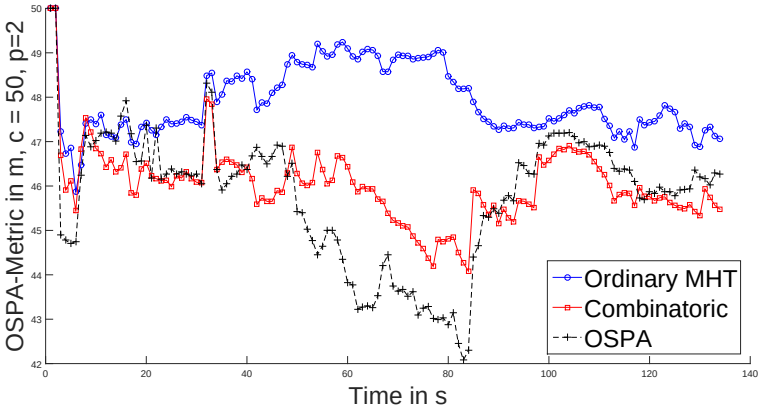

Figure 7.6: OSPA–value for $\lambda = 7$ © 2013 IEEE.

of a set of m_k distinguishable elements, where m_k is the number of measurements. This approach presumes that clutter occurs randomly. Second, an assignment–based procedure for computing the desired probabilities is described. The computation of the OSPA–metric between the established tracks of the ordinary MHT filter and the measurement set is the basis for the clutter hypothesis weighting. This approach takes the knowledge about the previous states into account. Finally, a numerical evaluation for different clutter levels shows that the combinatorial clutter hypothesis weighting yields always a better result than the ordinary MHT if clutter is present. Furthermore, the assignment–based approach yields comparable results (if $\lambda \geq 5$) to the combinatorial method for higher clutter levels and is therefore also a better choice than the ordinary MHT for cluttered scenarios. If no clutter is present the combinatoric approach shows a comparable result to the ordinary MHT.

CHAPTER VIII

Conclusions and Future Work

This thesis is divided into two parts. The first part studies mathematical aspects of the unifying formulation and derivation of multitarget tracking filters by employing finite point process theory. The highly complex challenge of passive and non-cooperative localization and tracking of an electromagnetic emitter in an urban environment using a single antenna array is discussed in the second part. Both parts separate the thesis into theory and practice. However, the multitarget tracking filters applied are a straightforward consequence of the theoretical considerations, since they are themselves members of the proposed unification. This brings to mind James Clerk Maxwell's famous words: "*There is nothing more practical than a good theory.*"

8.1 Conclusions

Finite Point Processes in Target Tracking

Chapter 2 studies well–known fundamentals of point process theory and establishes the connection to multitarget tracking filters in a comprehensive and intuitive way. Furthermore, the benefits of point process theory towards the closely related approach of RFS are discussed and it is noted that the theory of finite point processes is able to model filters for almost all practical relevant target tracking scenarios.
Chapter 3 presents the unification of several well–known multitarget tracking filters by a joint formulation using a single functional, called PGFL. The characterization is based on the finite point process theory proposed in Chapter 2. The multitarget tracking filters are presented according to their use of superposing targets. It is shown that well–known filters that *do not* superpose targets like the Bayes–Markov, MHT, PDA, JPDA, IPDA, JIPDA, PMHT and filters that *do* superpose targets in one state space like the PHD/iFilter, CPHD, multi–Bernoulli and generalized PHD/iFilter are contained in this unification, called the family of pointillist filters. Furthermore, a

family of filters in between the known superposition–classes, that partially superpose targets, are presented and called *hybrid pointillist filters*. Members are the joint PHD/iFilter and the joint generalized PHD/iFilter. The benefit of this simple and unique representation of multitarget tracking filters via a single PGFL is demonstrated by identifying existing relations between filters. The key benefit of using the PGFL–characterization of a pointillist filter is that instead of describing methodologies by a lot of (maybe inaccurate) words, the unique and exact description brings to light the key ingredients of a filter at a glance. An explicit demonstration for tracking engineers on how non–standard and demanding scenarios can be solved by designing customized pointillist filters via PGFLs is also presented using a non–standard example.

Chapter 4 illuminates the derivation of summary statistics of pointillist filters that are needed for the implementation and application of the particular filters to real world tracking scenarios. The mathematically correct definition of the *functional derivative with respect to Dirac delta* is the focus of this chapter. By the application of the classic LDC theorem the definition of the investigated functional derivative is shown to be valid for all pointillist filters formulated in Chapter 3. In particular, the theory of secular functions, which essentially transforms the functional derivative of a PGFL into an ordinary derivative of a secular function, is extended such that all of the pointillist filters from Chapter 3 can be derived in a mathematically correct way. This result implies that highly developed numerical and exact methods for differentiation can be applied to the derivation of the information update equations of pointillist filters. This enables the theoretical results presented in the first part to be applied in practice.

An Application to Emitter Tracking under Multipath Propagation

Chapter 5 introduces the challenge of blindly and passively localizing and tracking an electromagnetic emitter in an urban environment using a single antenna array. Furthermore, existing work is presented and open questions are identified to motivate the contributions of Chapters 6 and 7.

Chapter 6 extends the work on BML presented in [Alg10] to standard and generalized PHD intensity filters, that is, to filters from the class of pointillist filters that superpose targets. For standard PHD intensity filters the target–oriented measurement model does not model that targets in a BML scenario generate multiple measurements per sensor scan. Therefore, standard PHD intensity filters are successfully adapted to BML by formulating enhanced target state extraction schemes (see Section 6.1) and deriving and evaluating likelihood functions that are defined on single multipaths (see Definitions 6.13 and 6.53). Furthermore, the application of generalized PHD intensity filters to BML is studied intensively. In contrast to extended target tracking the spatial relation between measurements cannot be used for reducing measurement partitions. Therefore, approximation criteria, which do

not apply any information about the distribution of measurements in the parameter space, are derived for generalized PHD intensity filters and numerically evaluated. Finally, the proposed standard and generalized PHD intensity filters are numerically compared using simulated and demanding real world data and it is shown that the derived methods promise great applicability to solve the task of BML in further scenarios.

Chapter 7 presents an approach for tracking parameters in AoA and RToA based on a track–oriented MHT. In the existing work on parameter tracking for BML the problem of clutter multipaths which are received before the first target–related multipath is not considered. However, the performance of assignment–based data fusion algorithm decreases if clutter multipaths are received before the first target–related multipath. In this thesis MHT filtering equations that incorporate this effect are derived via a marginalization with respect to clutter hypotheses. A numerical evaluation demonstrates that the derived methodology outperforms the existing state–of–the–art approach from [Alg10].

Summary

The following conclusions summarize the main contributions of this thesis:

1. A unification of multitarget–tracking filters modeled by finite point processes, called the family of pointillist filters, is proposed and the applicability of the concept is demonstrated.

2. The derivation of summary statistics of pointillist filters via functional derivatives with respect to Dirac delta is mathematically justified and the theory of secular functions is extended to the family of pointillist filters.

3. Standard and generalized PHD intensity filters are adapted to the challenge of BML and evaluated using simulated and real world data.

4. A novel parameter tracking algorithm in the AoA–RToA domain using clutter hypotheses is derived.

8.2 Future Work

According to the physicist and mathematician Carl Friedrich Gauss "*Nothing is done as long as there is something left to do*" [1] The challenges studied and methodologies formulated in this thesis imply new questions, which have to be answered in future work.

[1] The original words are: "*Nichts ist getan, wenn noch etwas zu tun übrig ist.*"

Finite Point Processes in Target Tracking

The separation of pointillist filters is done with respect to target superposition and for a single measurement space. Multi–scan and multi–sensor versions of existing pointillist filters can be designed analogously by introducing additional test–functions defined on the measurement space. Future publications will study this extension.
Table 3.1 shows the family of pointillist filters, comprising many well–known tracking filters. However, the list is not exhaustive. The extension of the family of pointillist filters is an ongoing process left for the future.
The theory of secular functions and its extension to the family of pointillist filters discussed in Chapter 4 implies that the summary statistics of every pointillist filter can be derived using ordinary differentiation. Therefore, existing exact and numerical methods for computing these statistics can be applied to save numerical complexity. The evaluation of different methodologies for distinct pointillist filters is an open question left for further scientific work.

An Application to Emitter Tracking under Multipath Propagation

The application of further members of the family of pointillist filters to the task of BML, the design of possibly complete new well–suited pointillist filters for solving the challenge of BML and their evaluation in simulated and real world scenarios is left for future work. In particular the restricted resolvability of targets in BML scenarios motivates the application of *hybrid pointillist filters*, which partially superpose targets. Additionally, finite point process theory could be applied to derive data fusion algorithms which jointly perform blind channel estimation and target tracking in an analogous fashion as it is proposed in [Li14], [LK14].
Enhanced clutter models should be considered, since typically origins of clutter multipaths are objects that create several multipaths by themselves. Therefore, the Poisson model, which models clutter as single multipaths, could be replaced by more general models which enable the consideration of clutter scatterers. In combination with an estimation of the clutter intensity, e.g. by the intensity filter, schemes for the identification of clutter origins could be derived. Such clutter estimates could be used to derive a clutter–map and reject false multipaths or incorporate origins of clutter online into the context information of the ray–tracer.
Mission planning [Alg10] is a topic for improving BML scenarios with a moving OS. Solving the mission planning problem is equivalent to solving an optimization problem. To the knowledge of the author mission planning in terms of a BML scenario has not been done so far. The performance of data fusion algorithms would definitely benefit from an incorporation of mission planning methodologies.
An efficient implementation of pointillist filters using a generalized target–oriented measurement model like the generalized PHD intensity filters proposed in Section 6.3

is necessary to apply such filters in real time to real world scenarios. Since SMC–implementations of the filters are used, a parallel computation of the filter update for each particle and/or each partition could be an approach to save computation time. Furthermore, the numerical schemes for computing the summary statistics presented in Chapter 4 could be integrated to save computational effort. Additionally, further schemes for the reduction of the number of measurement partitions are needed.

List of Figures

Chapter 5

Chapter 6

Chapter 7

List of Tables

Chapter 3

Own Publications

[Deg14] C. Degen. The generalized intensity filter. In *9th Workshop on Sensor Data Fusion: Trends, Solutions, Applications (SDF)*, Bonn, Germany, 2014.

[DGK12] C. Degen, F. Govaers, and W. Koch. Track Maintenance Using the SMC–Intensity Filter. In *7th Workshop on Sensor Data Fusion: Trends, Solutions, Applications (SDF)*, Bonn, Germany, 2012.

[DMG12] C. Degen, H. El Mokni, and F. Govaers. Evaluation of a coupled laser inertial navigation system for pedestrian tracking. In *Proceedings of the 15th International Conference on Information Fusion*, Singapore, 2012.

[DGK13a] C. Degen, F. Govaers, and W. Koch. Emitter localization under multipath propagation using SMC-intensity filters. In *Proceedings of the 16th International Conference on Information Fusion*, Istanbul, Turkey, 2013.

[DGK13b] C. Degen, F. Govaers, and W. Koch. Multi hypothesis parameter tracking in relative time of arrival. In *8th Workshop on Sensor Data Fusion: Trends, Solutions, Applications (SDF)*, Bonn, Germany, 2013.

[DGK14a] C. Degen, F. Govaers, and W. Koch. Emitter localization under multipath propagation using a likelihood function decomposition that is linear in target space. In *Proceedings of the 17th International Conference on Information Fusion*, Salamanca, Spain, 2014.

[DGK14b] C. Degen, F. Govaers, and W. Koch. Tracking targets with multiple measurements per scan. In *Proceedings of the 17th International Conference on Information Fusion*, Salamanca, Spain, 2014.

[DGK15] C. Degen, F. Govaers, and W. Koch. Tracking targets with multiple measurements per scan using the generalized PHD filter. *Journal of Advances in Information Fusion*, 10(2):125–143, Dec. 2015.

[DSK15] C. Degen, R. Streit, and W. Koch. On the functional derivative with respect to the Dirac delta. In *10th Workshop on Sensor Data Fusion: Trends, Solutions, Applications (SDF)*, Bonn, Germany, 2015.

[GDW15] F. Govaers, C. Degen, K. Wild, P. Garcia–Ariza, U. Trautwein, M. Käske, S. Häfner, R. Thomä. Emitter localization in urban scenarios. In *8th Future Security*, Berlin, Germany, 2013.

[SDK15] R. Streit, C. Degen, and W. Koch. The pointillist family of multitarget tracking filters. Submitted to IEEE Transaction on Aerospace and Electronic Systems, May 2015. Online available at http://arxiv.org/abs/1505.08000; Website retrieved at the 18.06.2015.

Bibliography

[ADKT06a] V. Algeier, B. Demissie, W. Koch, and R. Thomä. Blind localization of 3G mobile terminals in multipath scenarios. In *3rd Workshop on Positioning, Navigation and Communication 2006 (WPNC)*, Mar. 2006.

[ADKT06b] V. Algeier, B. Demissie, W. Koch, and R. Thomä. Mobile terminal tracking in urban scenarios using multipath propagation. In *2nd Workshop on Multiple Sensor Data Fusion (MSDF)*, Oct. 2006.

[ADKT08a] V. Algeier, B. Demissie, W. Koch, and R. Thomä. State space initiation for blind mobile terminal position tracking. *EURASIP Journal on Advances in Signal Processing*, DOI: 10.1155/2008/394219, 2008.

[ADKT08b] V. Algeier, B. Demissie, W. Koch, and R. Thomä. Track initiation for blind mobile terminal position tracking using multipath propagation. In *Proceedings of the 11th International Conference on Information Fusion*, Cologne, Germany, 2008.

[Alg10] V. Algeier. *Blind Localization of Mobile Terminals in Urban Scenarios.* PhD thesis, TU Ilmenau, 2010.

[ALLH02] S. Ahonen, J. Lähteenmäki, H. Laitinen, and S. Horsmanheimo. Usage of mobile location techniques for UMTS network planning in urban environment. In *Proceedings of the IST Mobile And Wireless Telecommunications Summit*, pages 823–827, Thessaloniki, Greece, Jun. 2002.

[Alt12] H. W. Alt. *Lineare Funktionalanalysis.* Springer, 2012.

[ART04] V. Algeier, A. Richter, and R. Thomä. A gradient based algorithm for path parameter tracking in sounding. In *COST 273 Meeting TD(04)124*, Jun. 2004.

[ASP93] J. Arnold, S. W. Shaw, and H. Pasternack. Efficient target tracking using dynamic programming. *IEEE Transactions on Aerospace and Electronic Systems*, 29(1):44–56, Jan. 1993.

[BB06] H. A. P. Blom and E. A. Bloem. Bayesian tracking of two possible unresolved maneuvering targets. Technical report, National Aerospace Laboratory (NLR), 2006.

[BD00] R. Bolla and F. Davoli. Road traffic estimation from location tracking data in the mobile cellular network. In *Proceedings Wireless Communications and Networking Conference (WCNC) IEEE*, pages 1107–1112, Chicago, Il, USA, 2000.

[BES13] A. O. Bozdogan, M. Efe, and R. Streit. Reduced Palm intensity for track extraction. In *Proceedings of the 16th International Conference on Information Fusion*, Istanbul, Turkey, 2013.

[BL71] F. Bourgeois and J.-C. Lassalle. An extension of the Munkres algorithm for the assignment problem to rectangular matrices. *Communications of the ACM*, 14(12):802–804, Dec. 1971.

[Ble12] N. Bleistein. Saddle point contribution for an n-fold complex-valued integral–Unpublished Paper. Online, 2012. Available at http://www.cwp.mines.edu/~norm/Papers/steepdesc.pdf; Website retrieved at the 19.05.2015.

[Bro12] M. Broetje. *Multistatic Multihypothesis Tracking Techniques for Underwater- and Air Surveillance Applications.* PhD thesis, University of Siegen, 2012.

[BSE15] A. O. Bozdogan, R. Streit, and M. Efe. Reduced Palm intensity for track extraction. Submitted to IEEE Transaction on Aerospace and Electronic Systems, Oct. 2015. Online available at http://arxiv.org/abs/1510.06732; Website retrieved at the 03.11.2015.

[BSF88] Y. Bar-Shalom and T. E. Fortmann. *Tracking and Data Association.* Boston: Academic Press, 1988.

[BSL95] Y. Bar-Shalom and X. Li. *Multitarget-Multisensor Tracking: Principles and Techniques.* Storrs, Connecticut: YBS Press, 3rd printing edition, 1995.

[BSLK01] Y. Bar-Shalom, X. Li, and T. Kirubarajan. *Estimation with Applications to Tracking and Navigation.* Wiley-Intersience, 2001.

[CBS84] Kuo-Chu Chang and Y. Bar-Shalom. Joint probabilistic data association for multitarget tracking with possibly unresolved measurements and maneuvers. *IEEE Transactions on Automatic Control*, 29(7):585–594, Jul. 1984.

[CC09] R. Chakravorty and S. Challa. Multitarget tracking algorithm - Joint IPDA and Gaussian mixture PHD filter. In *Proceedings of the 12th International Conference on Information Fusion*, pages 316–323, Jul. 2009.

[CCS+06] N. Czink, P. Cera, J. Salo, E. Bonek, J.-P. Nuutinen, and J. Ylitalo. Improving clustering performance using multipath component distance. *Electronics Letters*, 42(1):33–5–, Jan. 2006.

[CD03] S. B. Colegrove and S. J. Davey. PDAF with multiple clutter regions and target models. *IEEE Transactions on Aerospace and Electronic Systems*, 39(1):110–124, Jan. 2003.

[CDC03] S. B. Colegrove, S. J. Davey, and B. Cheung. PDAF versus PMHT performance on OTHR data. In *Proceedings of the International Radar Conference 2003*, pages 560–565, Sep. 2003.

[CM12] D. Clark and R. Mahler. Generalized PHD filter via a general chain rule. In *Proceedings of the 15th International Conference on Information Fusion*, Singapore, 2012.

[CMME11] S. Challa, M. R. Morelande, D. Musicki, and R. J. Evans. *Fundamentals of object tracking.* Cambridge University Press, 2011.

[CTW+07] N. Czink, R. Tian, S. Wyne, F. Tufvesson, J.-P. Nuutinen, J. Ylitalo, E. Bonek, and A. Molisch. Tracking time-variant cluster parameters in MIMO channel measurements. In *Second International Conference on Communications and Networking in China (CHINACOM 07).*, volume 6, pages 1147–1151, Aug. 2007.

[CVW02] S. Challa, B.-N. Vo, and X. Wang. Bayesian approaches to track existence IPDA and random sets. In *Proceedings of the 5th International Conference on Information Fusion*, Annapolis, USA, 2002.

[Dau87] F. E. Daum. Solution of the Zakai equation by separation of variables. *IEEE Transactions on Automatic Control*, 32(10):941–943, Oct. 1987.

[Dau05] F. Daum. Nonlinear filters: beyond the Kalman filter. *Aerospace and Electronic Systems Magazine, IEEE*, 20(8):57–69, Aug. 2005.

[Dav15] S. Davey. Probabilistic multihypothesis tracker with an evolving Poisson prior. *IEEE Transactions on Aerospace and Electronic Systems*, 51(1), Jan. 2015. to appear.

[DBE11a] M. Daun, L. Broetje, and F. Ehlers. Aiding autonomous underwater vehicles navigation using multistatic sonar data. In *Digital Signal Processing (DSP) Workshop for In-Vehicle Systems*, Kiel, Germany, 2011.

[DBE11b] M. Daun, L. Broetje, and F. Ehlers. Simultaneous localisation and tracking onboard AUVs with multistatic sonar data. In *Proceedings of 4th. Underwater Acoustic Conference 2011 (UAM)*, Kos, Greece, 2011.

[Dir27] P. A. M. Dirac. The physical interpretation of the quantum dynamics. *Proceedings Royal Society*, 113:621–641, 1927.

[DOR08] B. Demissie, M. Oispuu, and E. Ruthotto. Localization of multiple sources with a moving array using subspace data fusion. In *Proceedings of the 11th International Conference on Information Fusion*, Cologne, Germany, 2008.

[DRC08] S. J. Davey, M.G. Rutten, and B. Cheung. A comparison of detection performance for several track-before-detect algorithms. In *Proceedings of the 11th International Conference on Information Fusion*, Cologne, Germany, 2008.

[DVJ03] D.J. Daley and D. Vere-Jones. *An introduction to the theory of point processes, Volume 1: Elementary Theory and Methods.*, volume 1. Springer, 2nd edition, 2003.

[DVJ08] D.J. Daley and D. Vere-Jones. *An introduction to the theory of point processes, Volume 2: General Theory and Structure.*, volume 2. Springer, 2nd edition, 2008.

[ED11] E. Engel and R. M. Dreizler. *Density Functional Theory.* Springer, 2011.

[Els09] J. Elstrodt. *Maß- und Integrationstheorie.* Springer, 2009.

[EWBS05] O. Erdinc, P. Willett, and Y. Bar-Shalom. Probability hypothesis density filter for multitarget multisensor tracking. In *Proceedings of the 8th International Conference on Information Fusion*, Jul. 2005.

[FJ99] F. G. Friedländer and M. Joshi. *Introduction to the Theory of Distributions.* Cambridge University Press, 2nd edition, 1999.

[FS09] P. Flajolet and R. Sedgewick. *Analytic Combinatorics.* Cambridge University Press, 1st edition, 2009.

[FSU09] D. Fränken, M. Schmidt, and M. Ulmke. Spooky action at a distance in the cardinalized probability hypothesis density filter. *IEEE Transactions on Aerospace and Electronic Systems*, 45(4):1657–1664, 2009.

[GGMS05a] K. Gilholm, S. Godsill, S. Maskell, and D. Salmond. Poisson models for extended target and group tracking. In *Proceedings SPIE*, volume 5913, pages 230–241, 2005.

[GGMS05b] K. Gilholm, S. Godsill, S. Maskell, and D. Salmond. Poisson models for extended target and group tracking. 5913:230–241, 2005.

[Gla89] A. Glassner. *An Introduction to Ray Tracing.* Academic Press, San Diego, USA, Aug. 1989.

[GLLZ14] A. Griewank, L. Lehmann, H. Leovey, and M. Zilberman. Automatic evaluations of cross-derivatives. *Mathematics of Computation*, 83(285):251–274, 2014.

[GLO12] K. Granström, C. Lundquist, and O. Orguner. Extended target tracking using a gaussian-mixture PHD filter. *IEEE Transactions On Aerospace And Electronic Systems*, 48(4):3268–3286, Oct. 2012.

[GMN97] I. R. Goodman, R. P. S. Mahler, and H. T. Nguyen. *Mathematics of Data Fusion.* Springer, 1997.

[GO12a] K. Granström and O. Orguner. On the reduction of Gaussian inverse Wishart mixtures. In *Proceedings of the 15th International Conference on Information Fusion*, Singapore, 2012.

[GO12b] K. Granström and O. Orguner. A PHD filter for tracking multiple extended targets using random matrices. *IEEE Transactions On Aerospace And Electronic Systems*, 60(11):5675–5671, Nov. 2012.

[GO13] K. Granström and O. Orguner. On spawning and combination of extended/group targets modeled with random matrices. *IEEE Transactions On Signal Processing*, 61(3):678–692, Feb. 2013.

[Gov12] F. Govaers. *Enhanced data fusion in communication constrained multi sensor applications.* PhD thesis, University of Bonn, 2012.

[Gra12] K. Grandström. *Extended Object Tracking using PHD filters.* PhD thesis, Linköping University, Sweden, 2012.

[Gri84] J. Grim. On structural approximating multivariate discrete probability distributions. *Kybernetika*, 20(1):1–17, 1984.

[GS96] J. Gunther and A.L. Swindlehurst. Algorithms for blind equalization with multiple antennas based on frequency domain subspaces. In *Conference Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP-96.*, volume 5, pages 2419–2422 vol. 5, 1996.

[HDC13] J. Houssineau, E. Delande, and D. Clark. Notes of the summer school on finite set statistics. *arXiv preprint arXiv:1308.2586*, Aug. 2013.

[HJL04] S. Horsmanheimo, H. Jormakka, and J. Lähteenmäki. Location-aided planning in mobile network - trial results. *Wireless Personal Communications*, 30(2-4):207–216, Sept. 2004.

[HKT15] S. Häfner, M. Käske, and R. S. Thomä. Polarisation-angle-delay estimation for blind localization approaches under multipath propagation. In *9th European Conference on Antennas and Propagation (EuCAP)*, pages 1–6, Apr. 2015.

[HWLW99] R. Hoppe, P. Wertz, F. M. Landstorfer, and G. Wölfle. Fast 3D ray tracing for the planning of microcells by intelligent preprocessing of the database. In *3rd European Personal and Mobile Communications Conference (EPMCC)*, Paris, France, Mar. 1999.

[HWLW03] R. Hoppe, P. Wertz, F. M. Landstorfer, and G. Wölfle. Advanced ray optical wave propagation modelling for urban and indoor scenarios including wideband properties. *European Transactions on Telecommunications (ETT)*, (1), Jan. 2003.

[Jaz70] A. H. Jazwinski. *Stochastic Processes and Filtering Theory.* Academic Press, 1970.

[JKB97] N. L. Johnson, S. Kotz, and N. Balakrishnan. *Discrete Multivariate Distributions.* New York: Wiley, 1997.

[Kar91] A. F. Karr. *Point Processes and Their Statistical Inference.* CRC Press, 2nd edition, 1991.

[Kin92] J. F. C. Kingman. *Poisson Processes.* Oxford Studies in Probability 3, 1992.

[KMKH05] C. Kreucher, M. Morelande, K. Kastella, and A.O. Hero. Particle filtering for multitarget detection and tracking. In *Aerospace Conference, 2005 IEEE*, pages 2101–2116, Mar. 2005.

[Knu98] D. E. Knuth. *The Art of Computer Programming.* Addison Wesley, 3rd edition, 1998.

[Koc10] W. Koch. On Bayesian tracking and data fusion: A tutorial introduction with examples. *Aerospace and Electronic Systems Magazine, IEEE*, 25(7):29–52, 2010.

[Koc14] W. Koch. *Tracking and Sensor Data Fusion – Methodological Framework and Selected Applications.* Springer, 2014.

[KS05] W. Koch and R. Saul. A Bayesian approach to extended object tracking and tracking of loosely structured target groups. In *Proceedings of the 8th International Conference on Information Fusion*, Philadelphia, PA, USA, 2005.

[KST06] S. Kikuchi, A. Sano, and H. Tsuji. Blind mobile positioning in urban environment based on ray-tracing analysis. *EURASIP Journal on Applied Signal Processing*, 2006.

[Lan99] F. M. Landstorfer. Wave propagation models for the planning of mobile communication networks. In *Proceedings of the 29th European Microwave Conference*, pages 1–6, Munich, Germany, 1999.

[Li14] Li Li. *Simultaneous Target and Multipath Positioning.* PhD thesis, Departement of Electrical and Computer Engineering Duke University, 2014.

[LK14] Li Li and J.L. Krolik. Simultaneous target and multipath positioning. *Selected Topics in Signal Processing, IEEE Journal of*, 8(1):153–165, Feb. 2014.

[LqDfC10] Zhao Li-quan, Zhao Dan-feng, and Xu Cong. An improved peak extraction algorithm in PHD based multi-target tracking algorithm applied to passive radar. In *6th International Conference on Wireless Communications Networking and Mobile Computing (WiCOM)*, Sept. 2010.

[Mah03] R. Mahler. Multitarget filtering via first-order multitarget moments. *IEEE Transactions On Aerospace And Electronic Systems*, 39(4):1152–1178, 2003.

[Mah07a] R. Mahler. PHD filters of higher order in target number. *IEEE Transactions On Aerospace And Electronic Systems*, 43(4):1523–1543, 2007.

[Mah07b] R. Mahler. *Statistical Multisource-Multitarget Information Fusion.* Norwood MA: Archtech House, 2007.

[Mah09] R. Mahler. PHD filters for nonstandard targets, i: Extended targets. In *Proceedings of the 12th International Conference on Information Fusion*, Seattle, WA, USA, 2009.

[Mah13] R. Mahler. Statistics 102 for multisensor-multitarget tracking. *IEEE Journal Of Selected Topics In Signal Processing*, 7(3):376–389, 2013.

[MBE14] Martin Michaelis, Martina Broetje, and Frank Ehlers. Parameter estimation for non-cooperative multistatic sonar. In *Proceedings of the Underwater Acoustics Conference*, Rhodes, Greece, 2014.

[MC03] M. Morelande and S. Challa. A multitarget tracking algorithm based on random sets. In *Proceedings of the 6th International Conference on Information Fusion*, Cairns, Australia, 2003.

[MCF+14] L. Mihaylova, A.Y. Carmi, F.Septier, A. Gning, S.K. Pang, and S. Godsill. Overview of Bayesian sequential Monte Carlo methods for group and extended object tracking. *Digital Signal Processing*, 25:1–16, Feb. 2014.

[ME94] D. Musicki and R. Evans. IPDA-PMM algorithm for tracking maneuvering targets in clutter. In *SPIE International Symposium*, Lake Buena Vista, Florida, Dec. 1994. Proceedings of the 33rd Conference on Decision and Control.

[ME02] D. Musicki and R. Evans. Track decoupling: linear joint IPDA (LJIPDA) and multi-target linear IPDA (MLIPDA). In *Information, Decision and Control, 2002. Final Program and Abstracts*, pages 335–340, Feb. 2002.

[ME04] D. Musicki and R. Evans. Joint integrated probabilistic data association: JIPDA. *IEEE Transactions on Aerospace and Electronic Systems*, 40(3):1093–1099, Jul. 2004.

[ME05] D. Musicki and R.J. Evans. Target existence based MHT. In *44th IEEE Conference on Decision and Control, 2005 and 2005 European Control Conference (CDC-ECC '05)*, pages 1228–1233, Dec. 2005.

[MES94] D. Musicki, R. Evans, and S. Stankovic. Integrated probabilistic data association. *IEEE Transactions on Automatic Control*, 39(6):1237–1241, Jun. 1994.

[Mic15] M. Michaelis. Bistatic simultaneous transmitter localization and mapping. In *10th Workshop Sensor Data Fusion: Trends, Solutions, Applications (SDF)*, Bonn, Germany, 2015.

[MKV12] M. Mallick, V. Krishnamurthy, and B.-N. Vo. *Integrated Tracking, Classification, and Sensor Management: Theory and Applications.* Wiley-IEEE Press, Dec. 2012.

[MLS08] D. Musicki and B. La Scala. Multi-target tracking in clutter without measurement assignment. *Aerospace and Electronic Systems, IEEE Transactions on*, 44(3):877–896, Jul. 2008.

[Moy62] J. E. Moyal. The general theory of stochastic population processes. *Acta Mathematica*, 108:1–31, 1962.

[MS06] K. Mosler and F. Schmid. *Wahrscheinlichkeitsrechung und schließende Statistik.* Springer, 2nd edition, 2006.

[MS08] D. Musicki and S. Suvorova. Tracking in clutter using IMM-IPDA-based algorithms. *IEEE Transactions on Aerospace and Electronic Systems*, 44(1):111–126, Jan. 2008.

[MSS14] D. Musicki, Taek Lyul Song, and R. Streit. Generating function derivation of the IPDA filter. In *9th Workshop on Sensor Data Fusion: Trends, Solutions, Applications (SDF)*, pages 1–6, Oct. 2014.

[OSD15] A. O'Connor, P. Setlur, and N. Devroye. Single-sensor RF emitter localization based on multipath exploitation. *IEEE Transactions on Aerospace and Electronic Systems*, 51(3):1635–1651, Jul. 2015.

[Ott13] M. Otte. *Mathematiker über Mathematik.* Springer, 2013.

[PE98] G. W. Pulford and R. J. Evans. A multipath data association tracker for Over-The-Horizon radar. *IEEE Transactions on Aerospace and Electronic Systems*, 34(4):1165–1183, Oct. 1998.

[RBGD15] S. Reuter, M. Beard, K. Granström, and K. Dietmayer. Tracking extended targets in high clutter using a GGIW–LMB filter. In *10th Workshop on Sensor Data Fusion: Trends, Solutions, Applications (SDF)*, Bonn, Germany, 2015.

[RBSW15a] K. Romeo, Y. Bar-Shalom, and P. Willett. Data fusion with ML-PMHT for very low SNR track detection in an OTHR. In *Proceedings of the 18th International Conference on Information Fusion*, Washington, DC, USA, 2015.

[RBSW15b] K. Romeo, Y. Bar-Shalom, and P. Willett. Fusion of multipath data with ML-PMHT for very low SNR track detection in an OTHR. *submitted to Journal of Advances in Information Fusion*, 2015.

[RCV10] B. Ristic, D. Clark, and B.-N. Vo. Improved SMC implementation of the PHD-filter. In *Proceedings of the 13th International Conference on Information Fusion*, Edinburgh, UK, 2010.

[Rei79] D. B. Reid. An algorithm for tracking multiple targets. *Automatic Control, IEEE Transactions on*, 24(6):843–854, Dec. 1979.

[Ric96] G. Richards. Application of the Hough transformation as a track-before-detect method. In *IEEE Colloquium on Target Tracking and Data Fusion*, 1996.

[RRG05] M. Rutten, B. Ristic, and N. Gordon. A comparison of particle filters for recursive track-before-detect. In *Proceedings of the 8th International Conference on Information Fusion*, Philadelphia, PA, USA, 2005.

[Rud87] W. Rudin. *Real and Complex Analysis*. McGraw-Hill, 3rd edition, 1987.

[RVVD14] S. Reuter, B.-T. Vo, B.-N. Vo, and K. Dietmayer. The labeled multi-Bernoulli filter. *Signal Processing, IEEE Transactions on*, 62(12):3246–3260, Jun. 2014.

[Sal09] J. Salmi. *Contributions to measurement-based dynamic MIMO channel modeling and propagation parameter estimation.* PhD thesis, Helsinki University of Technology, 2009.

[SC10] A. Swain and D. Clark. Extended object filtering using spatial independent cluster processes. In *Proceedings of the 13th Conference on Information Fusion*, pages 1–8, Jul. 2010.

[SCAS13] T. Sathyan, T.-J. Chin, S. Arulampalam, and D. Suter. A multiple hypothesis tracker for multitarget tracking with multiple simultaneous measurements. *IEEE Journal of Selected Topics in Signal Processing*, 7(3):448–460, Jun. 2013.

[SD13a] P. Setlur and N. Devroye. Bayesian and Cramer-Rao bounds for single sensor target localization via multipath exploitation. In *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, pages 5845–5849, May 2013.

[SD13b] P. Setlur and N. Devroye. Multipath exploited Bayesian and Cramer-Rao bounds for single-sensor target localization. *EURASIP Journal on Advances in Signal Processing*, 2013:53, Mar. 2013.

[SG97] G. E. Shilov and B. L. Gurevich. *Integral, Measure and Derivative: A Unified Approach.* Dover Pubn Inc, 2nd edition, 1997.

[SKM95] D. Stoyan, W.S. Kendall, and J. Mecke. *Stochastic geometry and its applications.* Wiley, 2nd edition, Sep. 1995.

[SKSC12] M. Schikora, W. Koch, R. Streit, and D. Cremers. *Sequential Monte Carlo Method for Multi-Target Tracking with the Intensity Filter.*, volume 410. Springer, 2012.

[SL93] R. Streit and T. E. Luginbuhl. A probabilistic multi-hypothesis tracking algorithm without enumeration and pruning. In *Proceedings of the Sixth Joint Service data Fusion symposium*, Laurel, Maryland, USA, 1993.

[SL94] R. Streit and T. E. Luginbuhl. Maximum likelihood method for probabilistic multi-hypothesis tracking. In *SPIE International Symposium*, volume 2235, Orlando, Florida, 1994. Signal and Data Processing of Small Targets.

[SL95] R. Streit and T. E. Luginbuhl. Probabilistic multi-hypothesis tracking. Technical Report 10428, Naval Undersea Warfare Center, Newport, Rhode Island, 1995.

[SNDE14] P. Setlur, T. Negishi, N. Devroye, and D. Erricolo. Multipath exploitation in Non-LOS urban synthetic aperture radar. *IEEE Journal of Selected Topics in Signal Processing*, 8(1):137–152, Feb. 2014.

[SR01] D. N. Shanbhag and C. R. Rao. *Handbook of Statistics 19: Stochastic Processes: Theory and Methods.* North-Holland, 1st edition, 2001.

[SRE+06] J. Salmi, A. Richter, M. Enescu, P. Vainikainen, and V. Koivunen. Propagation parameter tracking using variable state dimension Kalman filter. In *VTC 2006-63rd Spring Vehicular Technology using variable state dimension Kalman filter.*, volume 6, pages 2757–2761, 2006.

[SSAA12] P. Setlur, G.E. Smith, F. Ahmad, and M.G. Amin. Target localization with a single sensor via multipath exploitation. *Aerospace and Electronic Systems, IEEE Transactions on*, 48(3):1996–2014, Jul. 2012.

[SSCB14] L. D. Stone, R. L. Streit, T. L. Corwin, and K. L. Bell. *Bayesian Multi Target Tracking.* Artech House, 2nd edition, 2014.

[SSGW11a] L. Svensson, D. Svensson, M. Guerriero, and P. Willett. Set jpda filter for multitarget tracking. *IEEE Transactions on Signal Processing*, 59(10):4677–4691, Oct 2011.

[SSGW11b] L. Svensson, D. Svensson, M. Guerriero, and P. Willett. Set JPDA filter for multitarget tracking. *Signal Processing, IEEE Transactions on*, 59(10):4677–4691, Oct. 2011.

[Str10] R. Streit. *Poisson Point Processes - Imaging, Tracking, and Sensing.* Springer, 2010.

[Str13a] R. Streit. The probability generating functional for finite point processes, and its application to the comparison of PHD and intensity filters. *Journal of Advances in Information Fusion*, 8(2):119–132, Dec. 2013.

[Str13b] R. S. Strichartz. *Guide to Distribution Theory and Fourier.* World Scientific Publishing Company, 2013.

[Str14a] R. Streit. Generating function derivation of the PDA filter. In *Proceedings of the 17th International Conference on Information Fusion*, Salamanca, Spain, 2014.

[Str14b] R. Streit. Guest appearances of generating functions in multitarget tracking. Plenary Lecture at the 17th International Conference on Information Fusion, Salamanca, Spain, Jul. 2014. Online available at http:/www.fusion2014.org/sites/default/files/Roy-Streit_Gest-Appearances-Generating-Functions-Multitarget- Tracking.pdf; Website retrieved at the 19.05.2015.

[Str14c] R. Streit. Intensity filters on discrete spaces. *IEEE Transactions on Aerospace and Electronic Systems*, 50(2):1590–1599, Apr. 2014.

[Str14d] R. Streit. Joint intensity filter (JiFi) for tracking multiple groups– Unpublished Note., Jan. 2014.

[Str14e] R. Streit. A technique for deriving multitarget intensity filters using ordinary derivatives. *Journal of Advances in Information Fusion*, 9:3–12, Jun. 2014.

[Str15] R. Streit. Saddle point method for JPDA and related filters. In *Proceedings of the 18th International Conference on Information Fusion*, Washington D.C., USA, 2015.

[SUD10] D. Svensson, M. Ulmke, and L. Danielsson. Multi-target tracking with partially unresolved measurements. In *Proceedings of the 40th Annual Conf. of the German Informatics Society (GI): Service Science - New Perspectives in Computer Science INFORMATIK*, volume 2, pages 913–918, Leipzig, Germany, 2010.

[SVV08] D. Schumacher, B.-T. Vo, and B.-N. Vo. A consistent metric for performance evaluation of multi-object filters. *IEEE Transactions on Signal Processing*, 56(8):3447–3457, Aug. 2008.

[SW09] R. Streit and R. Wojtowicz. A general likelihood function decomposition that is linear in target state. In *Aerospace conference, 2009 IEEE*, pages 1–8, 2009.

[TP98] Lang Tong and S. Perreau. Multichannel blind identification: from subspace to maximum likelihood methods. *Proceedings of the IEEE*, 86(10):1951–1968, Oct 1998.

[TXK94] Lang Tong, Guanghan Xu, and Thomas Kailath. Blind identification and equalization based on second-order statistics: A time domain approach. *IEEE Transactions on Information Theory*, 40:340–349, 1994.

[VDFL15] H. X. Vu, S. J. Davey, F. K. Fletcher, and C.-C. Lim. Histogram-PMHT with an evolving Poisson prior. In *Proceedings of the 40th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).*, Brisbane, Australia, 2015.

[VGL+14] I. Vin, D. P. Gaillot, P. Laly, M. Lienard, and P. Degauque. Multipath component distance-based fingerprinting technique for non-cooperative outdoor localization in NLOS scenarios. *IEEE Transactions on Antennas and Propagation*, 62(9):4794–4798, Sept. 2014.

[VGLL12] I. Vin, D. P. Gaillot, P. Laly, and M. Lienard. Non cooperative mobile localization using a 3D data base correlation technique. In *EURO-COST TD(12)05010*, Sept. 2012.

[VGLL13] I. Vin, D. P. Gaillot, P. Laly, and M. Lienard. Non-Line-of-Sight localization algorithms with polarization diversity. In *EURO-COST IC1004 TD(13)06029*, Feb. 2013.

[VVC09] B.-T. Vo, B.-N. Vo, and A. Cantoni. The cardinality balanced multi-target multi-Bernoulli filter and its implementations. *IEEE Transactions on Signal Processing*, 57(2):409–423, 2009.

[VVHM13] B.-T. Vo, B.-N. Vo, R. Hoseinnezhad, and R. P. S. Mahler. Robust multi-Bernoulli filtering. *IEEE Journal of Selected Topics in Signal Processing*, 7(3):399–409, 2013.

[Wal74] W. Walter. *Einführung in die Theorie der Distributionen.* Bibliographisches Institut -Wissenschaftsverlag, Mannheim, Germany, 1974.

[WHBL00] G. Wölfle, R. Hoppe, T. Binzer, and F. Landstorfer. Radio network planning with ray optical propagation models for urban and indoor wireless communication networks. In *Millenium Conference on Antennas And Propagation*, Davos, Switzerland, Apr. 2000.

[WHL99] G. Wölfle, R. Hoppe, and F.M. Landstorfer. A fast and enhanced ray optical propagation model for indoor and urban scenarios, based on an intelligent preprocessing of the database. In *10th IEEE Internat. Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Osaka, Japan F5-3, 1999.

[Wil94] H. S. Wilf. *generatingfunctionology.* Academic Press, 1994.

[Wil14] J. Williams. Marginal multi-Bernoulli filters: RFS derivation of MHT, JIPDA and association-based MeMBer. Website, 2014. Online available at http://arxiv.org/pdf/1203.2995.pdf; Website retrieved at the 27.05.2015.

[Wil15] J. Williams. An efficient, variational approximation of the best fitting multi-Bernoulli filter. *Signal Processing, IEEE Transactions on*, 63(1):258–273, Jan. 2015.

[WSL+04] P. Wertz, M. Sauter, F. Landstorfer, G. Wölfle, and R. Hoppe. Automatic optimization algorithms for the planning of wireless local area networks. In *Proc. VTC2004-Fall Vehicular Technology Conference 2004 IEEE 60th*, pages 3010–3014, Los Angeles, CA, USA, Sept. 2004.

[YS94] J. Yang and A. L. Swindlehurst. DF directed multipath equalization. In *Conference Record of the 28th Asilomar Conference on Signals, Systems and Computers.*, volume 2, pages 1418–1422 vol.2, 1994.

[ZGL02] V. Zeimpekis, G. M. Giaglis, and G. Lekakos. A taxonomy of indoor and outdoor positioning techniques for mobile location services. *SIGecom Exch.*, 3(4):19–27, Dec. 2002.

[Zha00] Y. Zhao. Mobile phone location determination and its impact on intelligent transportation systems. *IEEE Transactions on Intelligent Transportation System*, 1(1):55–64, 2000.

[ZT95] H. Zeng and L. Tong. Some new results on blind channel estimation: performance and algorithms. In *29th Annual Conference on Information Sciences and Systems*, Mar. 1995.