Università degli Studi di Firenze

Dottorato in Informatica, Sistemi e Telecomunicazioni
Indirizzo: Ingegneria Informatica e dell'Automazione
Ciclo XXVII
Coordinatore: Prof. Luigi Chisci

Distributed multi-object tracking over sensor networks: a random finite set approach

Dipartimento di Ingegneria dell'Informazione
Settore Scientifico Disciplinare ING-INF/04

Author:	Supervisors:	Coordinator:
Claudio Fantacci	Prof. Luigi Chisci	Prof. Luigi Chisci
	Prof. Giorgio Battistelli	

Years 2012/2014

Contents

List of Theorems and Propositions

List of Figures

List of Tables

List of Acronyms

BF	Bayes Filter
BMF	Belief Mass Function
CBMM	Centralized Bayesian Multiple-Model
CDF	Cumulative Distribution Function
CI	Covariance Intersection
CMδGLMB	Consensus Marginalized δ-Generalized Labelled Multi-Bernoulli
CGM-CPHD	Consensus Gaussian Mixture-Cardinalized Probability Hypothesis Density
CLCP	Consensus on Likelihoods and Priors
CLMB	Consensus Labelled Multi-Bernoulli
COM	Communication
CP	Consensus on Posteriors
CPHD	Cardinalized Probability Hypothesis Density
CT	Coordinated Turn
δ-GLMB	δ-Generalized Labeled Multi-Bernoulli
$DGPB_1$	Distributed First Order Generalized Pseudo-Bayesian
DIMM	Distributed Interacting Multiple Model
DIMM-CL	Distributed Interacting Multiple Model with Consensus on Likelihoods
DMOF	Distributed Multi-Object Filtering
DMOT	Distributed Multi-Object Tracking
DOA	Direction Of Arrival
DSOF	Distributed Single-Object Filtering
DWNA	Discrete White Noise Acceleration
EKF	Extended Kalman Filter
EMD	Exponential Mixture Density
FISST	FInite Set STatistics
GCI	Generalized Covariance Intersection

GGM-CPHD	Global Gaussian Mixture - Cardinalized Probability Hypothesis Density
GLMB	Generalized Labeled Multi-Bernoulli
GM	Gaussian Mixture
GM-CMδGLMB	Gaussian Mixture-Consensus Marginalized δ-Generalized Labelled Multi-Bernoulli
GM-CLMB	Gaussian Mixture-Consensus Labelled Multi-Bernoulli
GM-δGLMB	Gaussian Mixture-δ-Generalized Labelled Multi-Bernoulli
GM-LMB	Gaussian Mixture-Labelled Multi-Bernoulli
GM-MδGLMB	Gaussian Mixture-Marginalized δ-Generalized Labelled Multi-Bernoulli
GPB_1	First Order Generalized Pseudo-Bayesian
I.I.D. / i.i.d. / iid	Independent and Identically Distributed
IMM	Interacting Multiple Model
JPDA	Joint Probabilistic Data Association
KF	Kalman Filter
KLA	Kullback-Leibler Average
KLD	Kullback-Leibler Divergence
LMB	Labeled Multi-Bernoulli
Mδ-GLMB	Marginalized δ-Generalized Labelled Multi-Bernoulli
MHT	Multiple Hypotheses Tracking
MM	Multiple Model
MOF	Multi-Object Filtering
MOT	Multi-Object Tracking
MSKF	Multi-Sensor Kalman Filter
NCV	Nearly-Constant Velocity
NWGM	Normalized Weighted Geometric Mean
OSPA	Optimal SubPattern Assignment
OT	Object Tracking
PF	Particle Filter
PDF	Probability Density Function
PHD	Probability Hypothesis Density
PMF	Probability Mass Function
PRMSE	Position Root Mean Square Error
RFS	Random Finite Set
SEN	Sensor
SMC	Sequential Monte Carlo
SOF	Single-Object Filtering

TOA	Time Of Arrival
UKF	Unscented Kalman Filter
UT	Unscented Transform

List of Symbols

$\mathcal{G}$	Directed graph
$\mathcal{N}$	Set of nodes
$\mathcal{A}$	Set of arcs
$\mathcal{S}$	Set of sensors
$\mathcal{C}$	Set of communication nodes
$\mathcal{N}^j$	Set of in-neighbours (including j itself)
$\mathrm{col}(\,\cdot^i)_{i\in\mathcal{I}}$	Stacking operator
$\mathrm{diag}(\,\cdot^i)_{i\in\mathcal{I}}$	Square diagonal matrix operator
$\langle f, g\rangle \triangleq \int f(x)\, g(x) dx$	Inner product operator for vector valued functions
$\top$	Transpose operator
$\mathbb{R}$	Real number space
$\mathbb{R}^n$	n dimensional Euclidean space
$\mathbb{N}$	Natural number set
k	Time index
x_k	State vector
$\mathbb{X}$	State space
n_x	Dimension of the state vector
$f_k(x)$	State transition function
w_k	Process noise
$\varphi(k\|\zeta)$	Markov transition density
y_k	Measurement vector
$\mathbb{Y}$	Measurement space
n_y	Dimension of the measurement vector
$h_k(x)$	Measurement function
v_k	Measurement noise
$g_k(y\|x)$	Likelihood function
$x_{1:k}$	State history
$y_{1:k}$	Measurement history

$p_k(x)$	Posterior/Filtered probability density function
$p_{k\|k-1}(x)$	Prior/Predicted probability density function
$p_0(x)$	Initial probability density function
$g_k^i(y\|x)$	Likelihood function of sensor i
A_{k-1}	$n_x \times n_x$ state transition matrix
C_k	$n_y \times n_x$ observation matrix
Q_k	Process noise covariance matrix
R_k	Measurement noise covariance matrix
$\mathcal{N}(\cdot;\, \cdot, \cdot)$	Gaussian probability density function
$\hat{x}_k$	Updated mean (state vector)
P_k	Updated covariance matrix (associated to $\hat{x}_k$)
$\hat{x}_{k\|k-1}$	Predicted mean (state vector)
$P_{k\|k-1}$	Predicted covariance matrix (associated to $\hat{x}_{k\|k-1}$)
e_k	Innovation
S_k	Innovation covariance matrix
K_k	Kalman gain
α_σ, β_σ, κ_σ	Unscented transformation weights parameters
c, w_m, w_c, W_c	Unscented transformation weights
X	Finite-valued-set
$h^X \triangleq \prod_{x \in X} h(x)$	Multi-object exponential
$\delta_X(\cdot)$	Generalized Kronecker delta
$1_X(\cdot)$	Generalized indicator function
$\|\cdot\|$	Cardinality operator
$\mathcal{F}(X)$	Space of finite subsets of X
$\mathcal{F}_n(\mathbb{X})$	Space of finite subsets of X with exactly n elements
$f(X)$, $\pi(X)$	Multi-object densities
$\rho(n)$	Cardinality probability mass function
$\beta(X)$	Belief mass function
$d(x)$	Probability hypothesis density function
$E[\cdot]$	Expectation operator
$\bar{n}$, D	Expected number of objects
$s(x)$	Location density (normalized $d(x)$)
r	Bernoulli existence probability
$q \triangleq 1 - r$	Bernoulli non-existence probability
X_k	Multi-object random finite set
Y_k	Random finite set of the observations
$Y_{1:k}$	Observation history random finite set

Y_k^i	Random finite set of the observations of sensor i
B_k	Random finite set of new-born objects
$P_{S,k}$	Survival probability
$\mathcal{C}$, $\mathcal{C}_k$	Random finite sets of clutter
$\mathcal{C}^i$, $\mathcal{C}_k^i$	Random finite sets of clutter of sensor i
$P_{D,k}$	Detection probability
$\pi_k(X)$	Posterior/Filtered multi-object density
$\pi_{k\|k-1}(X)$	Prior/Predicted multi-object density
$\varphi_{k\|k-1}(X\|Z)$	Markov multi-object transition density
$g_k(Y\|X)$	Multi-object likelihood function
$\rho_{k\|k-1}(n)$	Predicted cardinality probability mass function
$d_{k\|k-1}(x)$	Predicted probability hypothesis density function
$\rho_k(n)$	Updated cardinality probability mass function
$d_k(x)$	Updated probability hypothesis density function
$p_{b,k}(\cdot)$	Birth cardinality probability mass function
$\rho_{S,k\|k-1}(\cdot)$	Cardinality probability mass function of survived objects
$d_{b,k}(\cdot)$	Probability hypothesis density function of new-born objects
$\mathcal{G}_k^0(\cdot,\cdot,\cdot)$, $\mathcal{G}_{Y_k}(\cdot)$	Cardinalized probability hypothesis density function generalized likelihood functions
ℓ	Label
$\mathbf{X}$	Labeled finite-valued-set
$\mathbf{x}$	Labeled state vector
$\mathcal{L}(\mathbf{X})$	Label projection operator
$\Delta(\mathbf{X})$	Distinct label indicator
$\mathbb{L}_k$, $\mathbb{B}$	Label set for objects born at time k
$\mathbb{L}_{1:k-1}$, $\mathbb{L}_-$	Label set for objects up to time $k-1$
$\mathbb{L}_{1:k}$, $\mathbb{L}$	Label set for objects up to time k
ξ	Association history
Ξ	Association history set
$\boldsymbol{f}(\mathbf{X})$, $\boldsymbol{\pi}(\mathbf{X})$	Labeled multi-object densities
$w^{(c)}$, $w^{(c)}(\mathbf{X})$	Generalized labeled multi-Bernoulli weights indexed with $c \in \mathbb{C}$
$p^{(c)}$	Generalized labeled multi-Bernoulli location probability density function indexed with $c \in \mathbb{C}$
$(I,\xi) \in \mathcal{F}(\mathbb{L}) \times \Xi$	Labeled multi-object hypothesis
$w^{(I,\xi)}(\mathbf{X})$	δ-Generalized labeled multi-Bernoulli weight of hypothesis (I,ξ)
$p^{(I)}$	δ-Generalized labeled multi-Bernoulli location probability density function with association history ξ

$r^{(\ell)}$	Bernoulli existence probability associated to the object with label ℓ
$q^{(\ell)}$	Bernoulli non-existence probability associated to the object with label ℓ
$\boldsymbol{\varphi}_{k\|k-1}(X\|Z)$	Markov labeled multi-object transition density
$(p \oplus q)(x) \triangleq \frac{p(x)\,q(x)}{\langle p, q\rangle}$	Information fusion operator
$(\alpha \odot p)(x) \triangleq \frac{[p(x)]^{\alpha}}{\langle p^{\alpha}, 1\rangle}$	Information weighting operator
$\overline{p}(x)$, $\overline{q}$, $\overline{f}$, ...	Weighted Kullback-Leibler average
(q, Ω)	Information vector and (inverse covariance) matrix
$\omega^{i,j}$	Consensus weight of node i relative to j
$\omega_l^{i,j}$	Consensus weight of node i relative to j at the consensus step l
Π	Consensus matrix
l	Consensus step
L	Maximum number of consensus steps
ρ^i	Likelihood scalar weight of node i
b^i	Estimate of the fraction $\|\mathcal{S}/\mathcal{N}\|$
$\delta\Omega_k^i \triangleq \left(C_k^i\right)^{\top}\left(R_k^i\right)^{-1} C_k^i$	Information matrix gain
$\delta q_k^i \triangleq \left(C_k^i\right)^{\top}\left(R_k^i\right)^{-1} y_k^i$	Information vector gain
$\mathcal{P}_c$	Sets of probability density functions over a continuous state space
$\mathcal{P}_d$	Sets of probability mass functions over a discrete state space
p_{jt}	Markov transition probability from mode t to j
μ_k^j	Filtered modal probability of mode j
$\mu_{k\|k-1}^j$	Predicted modal probability of mode j
$\mu_k^{j\|t}$	Filtered modal probability of mode j conditioned to mode t
α_i	Gaussian mixture weight of the component i
N_G	Number of components of a Gaussian mixture
α_{ij}	Fused Gaussian mixture weight relative to components i and j
T_s	Sampling interval
β	Fused Gaussian mixture normalizing constant
p_x, p_y	Object planar position coordinates
$\dot{p}_x$, $\dot{p}_y$	Object planar velocity coordinates
λ_c	Poisson clutter rate
N_{mc}	Number of Monte Carlo trials
N_{max}	Maximum number of Gaussian components
γ_m	Merging threshold
γ_t	Truncation threshold
σ_{TOA}	Standard deviation of time of arrival sensor

σ_{DOA}	Standard deviation of direction of arrival sensor
$\boldsymbol{\pi}_k(\mathbf{X})$	Posterior/Filtered labeled multi-object density
$\boldsymbol{\pi}_{k\|k-1}(\mathbf{X})$	Prior/Predicted labeled multi-object density
$\boldsymbol{f}_B(\mathbf{X})$	Labeled multi-object birth density
$w_B(\mathbf{X})$	Labeled multi-object birth weight
$p_B(x,\ell)$	Labeled multi-object birth location probability density function
$r_B^{(\ell)}$	Labeled multi-Bernoulli newborn object weight
$p_B^{(\ell)}(x)$	Labeled multi-Bernoulli newborn object location probability density function
$w_S^{(\xi)}(\mathbf{X})$	Labeled multi-object survival object weight with association history ξ
$p_S(x,\ell)$	Labeled multi-object survival object location probability density function with association history ξ
θ	New association map
$\Theta(I)$	New association map set corresponding to the label subset I
$w_k^{(I,\xi)}$	δ-Generalized labeled multi-Bernoulli posterior/filtered weight of hypothesis (I,ξ)
$p_k^{(\xi)}(x,\ell)$	δ-Generalized labeled multi-Bernoulli posterior/filtered location probability density function with association history ξ
$w_{k\|k-1}^{(I,\xi)}$	δ-Generalized labeled multi-Bernoulli prior/predicted weight of hypothesis (I,ξ)
$p_{k\|k-1}^{(\xi)}(x,\ell)$	δ-Generalized labeled multi-Bernoulli prior/predicted location probability density function with association history ξ and label ℓ
$w_k^{(I)}$	Marginalized δ-Generalized labeled multi-Bernoulli posterior/filtered weight of the set I
$p_k^{(I)}(x,\ell)$	Marginalized δ-Generalized labeled multi-Bernoulli posterior/filtered location probability density function of the set I and label ℓ
$w_{k\|k-1}^{(I)}$	Marginalized δ-Generalized labeled multi-Bernoulli prior/predicted weight of the set I
$p_{k\|k-1}^{(I)}(x,\ell)$	Marginalized δ-Generalized labeled multi-Bernoulli prior/predicted location probability density function of the set I and label ℓ
$r_S^{(\ell)}$	Labeled multi-Bernoulli survival object weight
$p_S^{(\ell)}(x)$	Labeled multi-Bernoulli survival object location probability density function
$r_k^{(\ell)}$	Labeled multi-Bernoulli posterior/filtered existence probability of the object with label ℓ
$p_k^{(\ell)}(x)$	Labeled multi-Bernoulli posterior/filtered location probability density function of the object with label ℓ

$r_{k\|k-1}^{(\ell)}$	Labeled multi-Bernoulli prior/predicted existence probability of the object with label ℓ
$p_{k\|k-1}^{(\ell)}(x)$	Labeled multi-Bernoulli prior/predicted location probability density function of the object with label ℓ
$\langle f, g \rangle \triangleq \int f(X)\, g(X) \delta X$	Inner product operator for finite-valued-set functions
$\langle \mathbf{f}, \mathbf{g} \rangle \triangleq \int \mathbf{f}(\mathbf{X})\, \mathbf{g}(\mathbf{X}) \delta \mathbf{X}$	Inner product operator for labeled finite-valued-set functions

Acknowledgment

First and foremost, I would sincerely like to thank my supervisors Prof. Luigi Chisci and Prof. Giorgio Battistelli for their constant and relentless support and guidance during my 3-year doctorate. Their invaluable teaching and enthusiasm for research have made my Ph.D. education a very challenging yet rewarding experience. I would also like to thank Dr. Alfonso Farina, Prof. Ba-Ngu Vo and Prof. Ba-Tuong Vo. It has been a sincere pleasure to have the opportunity to work with such passionate, stimulating and friendly people. Our collaboration, knowledge sharing and feedback have been of great help and truly appreciated. Last but not least, I would like to thank all my friends and my family who supported, helped and inspired me during my studies.

Foreword

Statistics, mathematics and computer science have always been the favourite subjects in my academic career. The Ph.D. in automation and computer science engineering brought me to address challenging problems involving such disciplines. In particular, multi-object filtering concerns the joint detection and estimation of an unknown and possibly time-varying number of objects, along with their dynamic states, given a sequence of observation sets. Further, its distributed formulation also considers how to efficiently address such a problem over a heterogeneous sensor network in a fully distributed, scalable and computationally efficient way. Distributed multi-object filtering is strongly linked with statistics and mathematics for modeling and tackling the main issues in an elegant and rigorous way, while computer science is fundamental for implementing and testing the resulting algorithms. This topic poses significant challenges and is indeed an interesting area of research which has fascinated me during the whole Ph.D. period.

This thesis is the result of the research work carried out at the University of Florence (Florence, Italy) during the years 2012-2014, of a scientific collaboration for the biennium 2012-2013 with Selex ES (former SELEX SI, Rome, Italy) and of 6 months spent as a visiting Ph.D. scholar at the Curtin University of Technology (Perth, Australia) during the period January-July 2014.

Abstract

The aim of the present dissertation is to address distributed tracking over a network of heterogeneous and geographically dispersed nodes (or agents) with sensing, communication and processing capabilities. Tracking is carried out in the Bayesian framework and its extension to a distributed context is made possible via an information-theoretic approach to data fusion which exploits *consensus* algorithms and the notion of *Kullback–Leibler Average* (KLA) of the Probability Density Functions (PDFs) to be fused.

The first step toward distributed tracking considers a single moving object. Consensus takes place in each agent for spreading information over the network so that each node can track the object. To achieve such a goal, consensus is carried out on the local single-object posterior distribution, which is the result of local data processing, in the Bayesian setting, exploiting the last available measurement about the object. Such an approach is called *Consensus on Posteriors* (CP). The first contribution of the present work [BCF14a] is an improvement to the CP algorithm, namely *Parallel Consensus on Likelihoods and Priors* (CLCP). The idea is to carry out, in parallel, a separate consensus for the novel information (likelihoods) and one for the prior information (priors). This parallel procedure is conceived to avoid underweighting the novel information during the fusion steps. The outcomes of the two consensuses are then combined to provide the fused posterior density. Furthermore, the case of a single highly-maneuvering object is addressed. To this end, the object is modeled as a jump Markovian system and the *multiple model* (MM) filtering approach is adopted for local estimation. Thus, the consensus algorithms needs to be re-designed to cope with this new scenario. The second contribution [BCF+14b] has been to devise two novel consensus MM filters to be used for tracking a maneuvering object. The novel consensus-based MM filters are based on the *First Order Generalized Pseudo-Bayesian* (GPB_1) and *Interacting Multiple Model* (IMM) filters.

The next step is in the direction of distributed estimation of multiple moving objects. In order to model, in a rigorous and elegant way, a possibly time-varying number of objects present in a given area of interest, the *Random Finite Set* (RFS) formulation is adopted since it provides the notion of *probability density for multi-object states* that allows to directly extend existing tools in distributed estimation to multi-object tracking. The multi-object Bayes filter proposed by Mahler is a theoretically grounded solution to recursive Bayesian tracking based on RFSs. However, the multi-object Bayes recursion, unlike the single-object coun-

terpart, is affected by combinatorial complexity and is, therefore, computationally infeasible except for very small-scale problems involving few objects and/or measurements. For this reason, the computationally tractable *Probability Hypothesis Density* (PHD) and *Cardinalized PHD* (CPHD) filtering approaches will be used as a first endeavour to distributed multi-object filtering. The third contribution [BCF$^+$13a] is the generalisation of the single-object KLA to the RFS framework, which is the theoretical fundamental step for developing a novel consensus algorithm based on CPHD filtering, namely the *Consensus CPHD* (CCPHD). Each tracking agent locally updates multi-object CPHD, i.e. the cardinality distribution and the PHD, exploiting the multi-object dynamics and the available local measurements, exchanges such information with communicating agents and then carries out a fusion step to combine the information from all neighboring agents.

The last theoretical step of the present dissertation is toward distributed filtering with the further requirement of unique object identities. To this end the labeled RFS framework is adopted as it provides a tractable approach to the multi-object Bayesian recursion. The δ-GLMB filter is an exact closed-form solution to the multi-object Bayes recursion which jointly yields state and label (or trajectory) estimates in the presence of clutter, misdetections and association uncertainty. Due to the presence of explicit data associations in the δ-GLMB filter, the number of components in the posterior grows without bound in time. The fourth contribution of this thesis is an efficient approximation of the δ-GLMB filter [FVPV15], namely *Marginalized δ-GLMB* (Mδ-GLMB), which preserves key summary statistics (i.e. both the PHD and cardinality distribution) of the full labeled posterior. This approximation also facilitates efficient multi-sensor tracking with detection-based measurements. Simulation results are presented to verify the proposed approach. Finally, distributed labeled multi-object tracking over sensor networks is taken into account. The last contribution [FVV$^+$15] is a further generalization of the KLA to the labeled RFS framework, which enables the development of two novel consensus tracking filters, namely the *Consensus Marginalized δ-Generalized Labeled Multi-Bernoulli* (CM-δGLMB) and the *Consensus Labeled Multi-Bernoulli* (CLMB) tracking filters. The proposed algorithms provide a fully distributed, scalable and computationally efficient solution for multi-object tracking.

Simulation experiments on challenging single-object or multi-object tracking scenarios confirm the effectiveness of the proposed contributions.

1 Introduction

Recent advances in wireless sensor technology have led to the development of large networks consisting of radio-interconnected nodes (or agents) with sensing, communication and processing capabilities. Such a net-centric technology enables the building of a more complete picture of the environment, by combining information from individual nodes (usually with limited observability) in a way that is *scalable* (w.r.t. the number of nodes), *flexible* and *reliable* (i.e. *robust* to failures). Getting these benefits calls for architectures in which individual agents can operate without knowledge of the information flow in the network. Thus, taking into account the above-mentioned considerations, *Multi-Object Tracking* (MOT) in sensor networks requires redesigning the architecture and algorithms to address the following issues:

- lack of a central fusion node;
- scalable processing with respect to the network size;
- each node operates without knowledge of the network topology;
- each node operates without knowledge of the dependence between its own information and the information received from other nodes.

To combine limited information (usually due to low observability) from individual nodes, a suitable *information fusion* procedure is required to reconstruct, from the node information, the state of the objects present in the surrounding environment. The scalability requirement, the lack of a fusion center and knowledge on the network topology call for the adoption of a *consensus* approach to achieve a collective fusion over the network by iterating local fusion steps among neighboring nodes [OSFM07, XBL05, CA09, BC14]. In addition, due to the possible data incest problem in the presence of network loops that can causes *double counting* of information, robust (but suboptimal) fusion rules, such as the *Chernoff fusion* rule [CT12, CCM10] (that includes *Covariance Intersection* (CI) [JU97, Jul08] and its generalization [Mah00]) are required.

The focus of the present dissertation is *distributed estimation*, from the single object to the more challenging multiple object case.

In the context of *Distribute Single-Object Filtering* (DSOF), standard or Extended or Unscented Kalman filters are adopted as local estimators, the consensus involves a single Gaussian component per node, characterized by either the estimate-covariance or the information pair. Whenever multiple models are adopted for better describing the motion of the object in the tracking scenario, multiple Gaussian components per node arise and consensus has to be extended to this multicomponent setting. Clearly the presence of different Gaussian components related to different motion models of the same object or to different objects imply different issues and corresponding solution approaches that will be separately addressed. In this single-object setting, the main contributions in the present work are:

I. the development of a novel consensus algorithm, namely *Parallel Consensus on Likelihoods and Priors* (CLCP), that carries out, in parallel, a separate consensus for the novel information (likelihoods) and one for the prior information (priors);

II. two novel consensus MM filters to be used for tracking a maneuvering object, namely *Distributed First Order Generalized Pseudo-Bayesian* (DGPB_1) and *Distributed Interacting Multiple Model* (DIMM) filters.

Furthermore, *Distribute Multi-Object Filtering* (DMOF) is taken into account. To model a possibly time-varying number of objects present in a given area of interest in the presence of detection uncertainty and clutter, the *Random Finite Set* (RFS) approach is adopted. The RFS formulation provides the useful concept of *probability density* for *multi-object states* that allows to directly extend existing tools in distributed estimation to multi-object tracking. Such a concept is not available in the MHT and JPDA approaches [Rei79, FS85, FS86, BSF88, BSL95, BP99]. However, the multi-object Bayes recursion, unlike the single-object counterpart, is affected by combinatorial complexity and is, therefore, computationally infeasible except for very small-scale problems involving very few objects and/or measurements. For this reason, the computationally tractable *Probability Hypothesis Density* (PHD) and *Cardinalized PHD* (CPHD) filtering approaches will be used to address DMOF. It is recalled that the CPHD filter propagates in time the discrete distribution of the number of objects, called *cardinality distribution*, and the spatial distribution in the state space of such objects, represented by the PHD (or intensity function). It is worth to point out that there have been several interesting contributions [Mah00, CMC90, CJMR10, UJCR10, UCJ11] on multi-object fusion. More specifically, [CMC90] addressed the problem of optimal fusion in the case of known correlations while [Mah00, CJMR10, UJCR10, UCJ11] concentrated on robust fusion for the practically more relevant case of unknown correlations. In particular, [Mah00] first generalized CI in the context of multi-object fusion. Subsequently, [CJMR10] specialized the *Generalized Covariance Intersection* (GCI) of [Mah00] to specific forms of the multi-object densities providing, in particular, GCI fusion of cardinality distributions and PHD functions. In [UJCR10], a Monte Carlo (particle) realization is proposed for the GCI fusion of PHD functions. The two key contributions in this thesis work are:

I. the generalisation of the single-object KLA to the RFS framework;

II. a novel *consensus CPHD* (CCPHD) filter, based on a Gaussian Mixture (GM) implementation.

Multi-object tracking (MOT) involves the on-line estimation of an unknown and time-varying number of objects and their individual trajectories from sensor data [BP99, BSF88, Mah07b]. The key challenges in multi-object tracking also include *data association uncertainty*. Numerous multi-object tracking algorithms have been developed in the literature and most of these fall under the three major paradigms of: *Multiple Hypotheses Tracking* (MHT) [Rei79, BP99], *Joint Probabilistic Data Association* (JPDA) [BSF88] and *Random Finite Set* (RFS) filtering [Mah07b]. The proposed solutions are based on the recently introduced concept of labeled RFS that enables the estimation of multi-object trajectories in a principled manner [VV13]. In addition, labeled RFS-based trackers do not suffer from the so-called "*spooky effect*" [FSU09] that degrades performance in the presence of low detection probability like in the multi-object filters [VVC07, BCF$^+$13a, UCJ13]. Labeled RFS conjugate priors [VV13] have led to the development of a tractable analytic multi-object tracking solution called the *δ-Generalized Labeled Multi-Bernoulli* (δ-GLMB) filter [VVP14]. The computational complexity of the δ-GLMB filter is mainly due to the presence of explicit data associations. For certain applications such as tracking with multiple sensors, partially observable measurements or decentralized estimation, the application of a δ-GLMB filter may not be possible due to limited computational resources. Thus, cheaper approximations to the δ-GLMB filter are of practical significance in MOT. Core contribution of the present work is a new approximation of the δ-GLMB filter. The result is based on the approximation proposed in [PVV$^+$14] where it was shown that the more general *Generalized Labeled Multi-Bernoulli* (GLMB) distribution can be used to construct a principled approximation of an arbitrary labeled RFS density that matches the PHD and the cardinality distribution. The resulting filter is referred to as *Marginalized δ-GLMB* (Mδ-GLMB) since it can be interpreted as a marginalization over the data associations. The proposed filter is, therefore, computationally cheaper than the δ-GLMB filter while preserving key summary statistics of the multi-object posterior. Importantly, the Mδ-GLMB filter facilitates tractable multi-sensor multi-object tracking. Unlike PHD/CPHD and multi-Bernoulli based filters, the proposed approximation accommodates statistical dependence between objects. An alternative derivation of the *Labeled Multi-Bernoulli* (LMB) filter [RVVD14] based on the newly proposed Mδ-GLMB filter is presented.

Finally, *Distributed MOT* (DMOT) is taken into account. The proposed solutions are based on the above-mentioned labeled RFS framework that has led to the development of the δ-GLMB tracking filter [VVP14]. However, it is not known if this filter is amenable to DMOT. Nonetheless, the Mδ-GLMB and the LMB filters are two efficient approximations of the δ-GLMB filter that

- have an appealing mathematical formulation that facilitates an efficient and tractable closed-form fusion rule for DMOT;

- preserve key summary statistics of the full multi-object posterior.

In this setting, the main contributions in the present work are:

I. the development of the first distributed multi-object tracking algorithms based on the labeled RFS framework, generalizing the approach of [BCF+13a] from moment-based filtering to tracking with labels;

II. the development of *Consensus Marginalized δ-Generalized Labeled Multi-Bernoulli* (CMδ-GLMB) and *Consensus Labeled Multi-Bernoulli* (CLMB) tracking filters.

Simulation experiments on challenging tracking scenarios confirm the effectiveness of the proposed contributions.

The rest of the thesis is organized as follows.

Chapter 2 - Background

This chapter introduces notation, provides the necessary background on recursive Bayesian estimation, Random Finite Sets (RFSs), Bayesian multi-object filtering, distributed estimation and the network model.

Chapter 3 - Distributed single-object filtering

This chapter provides novel contributions on distributed nonlinear filtering with applications to nonlinear single-object tracking. In particular: I) a *Parallel Consensus on Likelihoods and Priors* (CLCP) filter is proposed to improve performance with respect to existing consensus approaches for distributed nonlinear estimation; II) a consensus-based multiple model filter for jump Markovian systems is presented and applied to tracking of a highly-maneuvering objcet.

Chapter 4 - Distributed multi-object filtering

This chapter introduces consensus multi-object information fusion according to an information-theoretic interpretation in terms of Kullback-Leibler averaging of multi-object distributions. Moreover, the *Consensus Cardinalized Probability Hypothesis Density* (CCPHD) filter is presented and its performance is evaluated via simulation experiments.

Chapter 5 - Centralized multi-object tracking

In this chapter, two possible approximations of the *δ-Generalized Labeled Multi-Bernoulli* (δ-GLMB) density are presented, namely I) the *Marginalized δ-Generalized Labeled Multi-Bernoulli* (Mδ-GLMB) and II) the *Labeled Multi-Bernoulli* (LMB). Such densities will allow to develop a new centralized tracker and to establish a new theoretical connection to previous

work proposed in the literature. Performance of the new centralized tracker is evaluated via simulation experiments.

Chapter 6 - Distributed multi-object tracking

This chapter introduces the information fusion rules for *Marginalized δ-Generalized Labeled Multi-Bernoulli* (Mδ-GLMB) and *Labeled Multi-Bernoulli* (LMB) densities. An information-theoretic interpretation of such fusions, in terms of Kullback-Leibler averaging of labeled multi-object densities, is also established. Furthermore, the *Consensus Mδ-GLMB* (CMδ-GLMB) and *Consensus LMB* (CLMB) tracking filters are presented as two new labeled distributed multi-object trackers. Finally, the effectiveness of the proposed trackers is discussed via simulation experiments on realistic distributed multi-object tracking scenarios.

Chapter 7 - Conclusions and future work

The thesis ends with concluding remarks and perspectives for future work.

2 Background

Contents

2.1 Network model

Recent advances in wireless sensor technology has led to the development of large networks consisting of radio-interconnected nodes (or agents) with sensing, communication and processing capabilities. Such a net-centric technology enables the building of a more complete picture of the environment, by combining information from individual nodes (usually with limited observability) in a way that is scalable (w.r.t. the number of nodes), flexible and reliable (i.e. robust to failures). Getting these benefits calls for architectures in which individual agents can operate without knowledge of the information flow in the network. Thus, taking into account the above-mentioned considerations, *Object Tracking* (OT) in sensor networks requires redesigning the architecture and algorithms to address the following issues:

- lack of a central fusion node;
- scalable processing with respect to the network size;
- each node operates without knowledge of the network topology;
- each node operates without knowledge of the dependence between its own information and the information received from other nodes.

The network considered in this work (depicted in Fig. 2.1) consists of two types of heterogeneous and geographically dispersed nodes (or agents): *communication* (COM) nodes have only processing and communication capabilities, i.e. they can process local data as well as exchange data with the neighboring nodes, while *sensor* (SEN) nodes have also sensing capabilities, i.e. they can sense data from the environment. Notice that, since COM nodes do not provide any additional information, their presence is needed only to improve network connectivity.

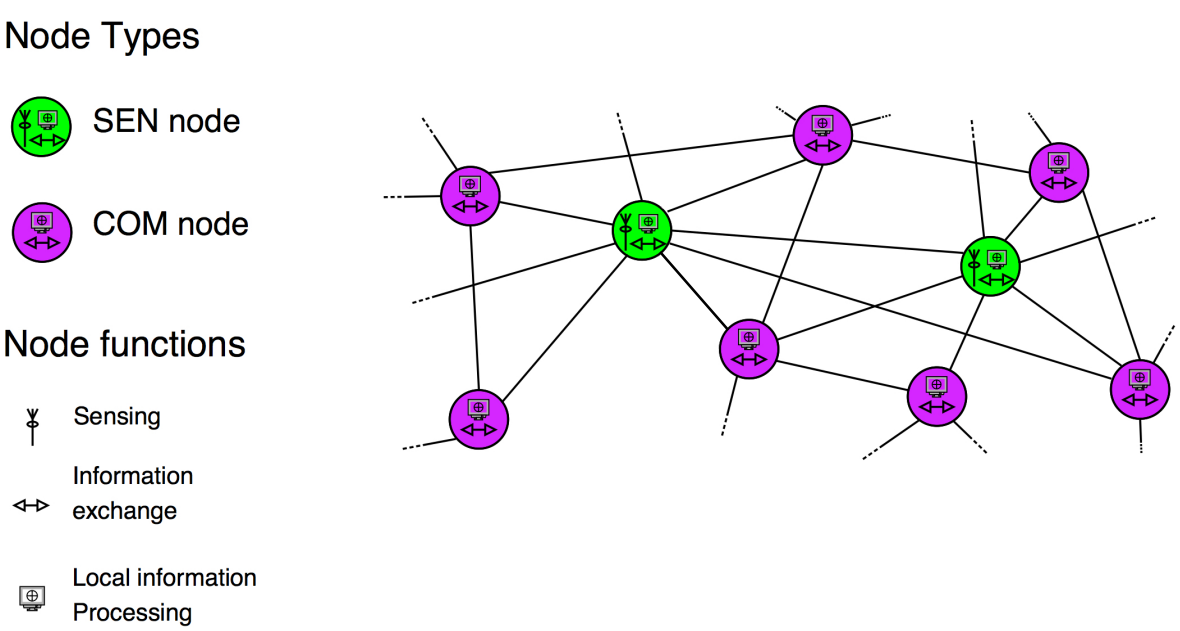


Figure 2.1: Network model

From a mathematical viewpoint, the network is described by a directed graph $\mathcal{G} = (\mathcal{N}, \mathcal{A})$ where $\mathcal{N} = \mathcal{S} \cup \mathcal{C}$ is the set of nodes, $\mathcal{S}$ is the set of sensor and $\mathcal{C}$ the set of communication nodes, and $\mathcal{A} \subseteq \mathcal{N} \times \mathcal{N}$ is the set of arcs, representing links (or connections). In particular, $(i, j) \in \mathcal{A}$ if node j can receive data from node i. For each node $j \in \mathcal{N}$, $\mathcal{N}^j \triangleq \{i \in \mathcal{N} : (i, j) \in \mathcal{A}\}$ denotes the set of in-neighbours (including j itself), i.e. the set of nodes from which node j can receive data.

Each node performs local computation, exchanges data with the neighbors and gathers measurements of kinematic variables (e.g., angles, distances, Doppler shifts, etc.) relative to objects present in the surrounding environment (or surveillance area). The focus of this thesis

will be the development of networked estimation algorithms that are scalable with respect to network size, and to allow each node to operate without knowledge of the dependence between its own information and the information from other nodes.

2.2 Recursive Bayesian estimation

The main interest of the present dissertation is *estimation*, which refers to inferring the values of a set of unknown variables from information provided by a set of noisy measurements whose values depend on such unknown variables. Estimation theory dates back to the work of Gauss [Gau04] on determining the orbit of celestial bodies from their observations. These studies led to the technique known as *Least Squares*. Over centuries, many other techniques have been proposed in the field of estimation theory [Fis12, Kol50, Str60, Tre04, Jaz07, AM12], e.g., the *Maximum Likelihood*, the *Maximum a Posteriori* and the *Minimum Mean Square Error* estimation. The Bayesian approach models the quantities to be estimated as random variables characterized by *Probability Density Functions* (PDFs), and provides an improved estimation of such quantities by conditioning the PDFs on the available noisy measurements. Hereinafter, we refer to the Bayesian approach as to *recursive Bayesian estimation* (or *Bayesian filtering*), a renowned and well-established probabilistic approach for recursively propagating, in a principled way via a two-step procedure, a PDF of a given time-dependent variable of interest. The first key concept of the present work is, indeed, Bayesian filtering. The propagated PDF will be used to describe, in a probabilistic way, the behaviour of a moving object. In the following, a summary of the *Bayes Filter* (BF) is given, as well as a review of a well known closed-form solution of it, the *Kalman Filter* (KF) [Kal60, KB61] obtained in the linear Gaussian case.

2.2.1 Notation

The following notation is adopted throughout the thesis: $\mathrm{col}\big(\,\cdot^i\big)_{i\in\mathcal{I}}$, where I is a finite set, denotes the vector/matrix obtained by stacking the arguments on top of each other; $\mathrm{diag}\big(\,\cdot^i\big)_{i\in\mathcal{I}}$, where I is a finite set, denotes the square diagonal matrix obtained by placing the arguments in the (i,i)-th position of the main diagonal; the standard inner product notation is denoted as

$$\langle f, g\rangle \triangleq \int f(x)\, g(x) dx\,; \tag{2.1}$$

vectors are represented by lowercase letters, e.g. x, $\mathbf{x}$; spaces are represented by blackboard bold letters e.g. $\mathbb{X}$, $\mathbb{Y}$, $\mathbb{L}$, etc. The superscript $\top$ stems for the transpose operator.

2.2.2 Bayes filter

Consider a discrete-time state-space representation for modelling a dynamical system. At each time $k \in \mathbb{N}$, such a system is characterized by a *state vector* $x_k \in \mathbb{X} \subseteq \mathbb{R}^{n_x}$, where n_x is the dimension of the state vector. The state evolves according to the following discrete-time

stochastic model:

$$x_k = f_{k-1}(x_{k-1}, w_{k-1}) \ , \tag{2.2}$$

where f_{k-1} is a, possibly nonlinear, function; w_{k-1} is the process noise modeling uncertainties and disturbances in the object motion model. The time evolution (2.2) is equivalently represented by a Markov transition density

$$\varphi_{k|k-1}(x|\zeta) \ , \tag{2.3}$$

which is the PDF associated to the transition from the state $\zeta = x_{k-1}$ to the new state $x = x_k$.

Likewise, at each time k, the dynamical system described with state vector x_k can be observed via a noisy *measurement vector* $y_k \in \mathbb{Y} \subseteq \mathbb{R}^{n_y}$, where n_y is the dimension of the observation vector. The measurement process can be modelled by the measurement equation

$$y_k = h_k(x_k, v_k) \ , \tag{2.4}$$

which provides an indirect observation of the state x_k affected by the measurement noise v_k. The modeling of the measurement vector is equivalently represented by the *likelihood function*

$$g_k(y|x) \ , \tag{2.5}$$

which is the PDF associated to the generation of the measurement vector $y = y_k$ from the dynamical system with state $x = x_k$.

The aim of recursive state estimation (or filtering) is to sequentially estimate over time x_k given the measurement history $y_{1:k} \triangleq \{y_1, \ldots, y_k\}$. It is assumed that the PDF associated to $y_{1:k}$ given the state history $x_{1:k} \triangleq \{x_1, \ldots, x_k\}$ is

$$g_{1:k}(y_{1:k}|x_{1:k}) = \prod_{\kappa=1}^{k} g_\kappa(y_\kappa|x_\kappa) \, , \tag{2.6}$$

i.e. *the measurements $y_{1:k}$ are conditionally independent on the states $x_{1:k}$*. In the Bayesian framework, the entity of interest is the *posterior density* $p_k(x)$ that contains all the information about the state vector x_k given all the measurements up to time k. Such a PDF can be recursively propagated in time resorting to the well know Chapman-Kolmogorov equation and the Bayes' rule [HL64]

$$p_{k|k-1}(x) = \int \varphi_{k|k-1}(x|\zeta) \, p_{k-1}(\zeta) \, d\zeta \, , \tag{2.7}$$

$$p_k(x) = \frac{g_k(y_k|x) \ p_{k|k-1}(x)}{\displaystyle\int g_k(y_k|\zeta) \ p_{k|k-1}(\zeta) \, d\zeta} \, , \tag{2.8}$$

given an *initial density* $p_0(\cdot)$. The PDF $p_{k|k-1}(\cdot)$ is referred to as the *predicted density*, while

$p_k(\cdot)$ is the *filtered density*.

Let us consider a multi-sensor *centralized* setting in which a sensor network $(\mathcal{N}, \mathcal{A})$ conveys all the measurements to a central fusion node. Assuming that the measurements taken by the sensors are independent, the Bayesian filtering recursion can be naturally extended as follows:

$$p_{k|k-1}(x) = \int \varphi_{k|k-1}(x|\zeta)\, p_{k-1}(\zeta)\, d\zeta\,, \tag{2.9}$$

$$p_k(x) = \frac{\prod\limits_{i\in\mathcal{N}} g_k^i\Big(y_k^i|x\Big)\; p_{k|k-1}(x)}{\int \prod\limits_{i\in\mathcal{N}} g_k^i\left(y_k^i|\zeta\right)\; p_{k|k-1}(\zeta)\, d\zeta}\,. \tag{2.10}$$

2.2.3 Kalman Filter

The KF [Kal60, KB61, HL64] is a closed-form solution of (2.7)-(2.8) in the linear Gaussian case. That is, suppose that (2.2) and (2.4) are linear transformations of the state with additive Gaussian white noise, i.e.

$$x_k = A_{k-1}x_{k-1} + w_{k-1}\,, \tag{2.11}$$

$$y_k = C_k x_k + v_k\,, \tag{2.12}$$

where A_{k-1} is the $n_x \times n_x$ state transition matrix, C_k is the $n_y \times n_x$ observation matrix, w_{k-1} and v_k are mutually independent zero-mean white Gaussian noises with covariances Q_{k-1} and R_k, respectively. Thus, the Markov transition density and the likelihood functions are

$$\varphi_{k|k-1}(x|\zeta) = \mathcal{N}(x;\, A_{k-1}\zeta, Q_{k-1})\;, \tag{2.13}$$

$$g_k(y|x) = \mathcal{N}(y;\, C_k x, R_k)\;, \tag{2.14}$$

where

$$\mathcal{N}(x;\, m, P) \triangleq |2\pi P|^{-\frac{1}{2}}\, e^{-\frac{1}{2}(x-m)^\top P^{-1}(x-m)} \tag{2.15}$$

is a Gaussian PDF. Finally, suppose that the prior density

$$p_{k-1}(x) = \mathcal{N}(x;\, \hat{x}_{k-1}, P_{k-1}) \tag{2.16}$$

is Gaussian with mean $\hat{x}_{k-1}$ and covariance P_{k-1}. Solving (2.7), the predicted density turns out to be

$$p_{k|k-1}(x) = \mathcal{N}\Big(x;\, \hat{x}_{k|k-1}, P_{k|k-1}\Big)\;, \tag{2.17}$$

a Gaussian PDF with mean $\hat{x}_{k|k-1}$ and covariance $P_{k|k-1}$. Moreover, solving (2.8), the posterior density (or *updated density*), turns out to be

$$p_k(x) = \mathcal{N}(x;\, \hat{x}_k, P_k)\;, \tag{2.18}$$

i.e. a Gaussian PDF with mean $\hat{x}_k$ and covariance P_k.

Remark 1. If the posterior distributions are in the same family as the prior probability distribution, the prior and posterior are called *conjugate distributions*, and the prior is called a *conjugate prior* for the likelihood function. The Gaussian distribution is a conjugate prior.

The KF recursion for computing both predicted and updated pairs $(\hat{x}_{k-1}, P_{k-1})$ and $(\hat{x}_k, P_k)$ is reported in Table 2.1.

Table 2.1: The Kalman Filter (KF)

for $k = 1, 2, \dots$ **do**

Prediction

$\hat{x}_{k|k-1} = A_{k-1}\hat{x}_{k-1}$ ▷ Predicted mean

$P_{k|k-1} = A_{k-1}P_{k-1}A_{k-1}^\top + Q_{k-1}$ ▷ Predicted covariance matrix

Correction

$e_k = y_k - C_k\hat{x}_{k|k-1}$ ▷ Innovation

$S_k = R_k + C_kP_{k|k-1}C_k^\top$ ▷ Innovation covariance matrix

$K_k = P_{k|k-1}C_k^\top S_k^{-1}$ ▷ Kalman gain

$\hat{x}_k = \hat{x}_{k|k-1} + K_ke_k$ ▷ Updated mean

$P_k = P_{k|k-1} - K_kS_kK_k^\top$ ▷ Updated covariance matrix

end for

In the centralized setting the network $(\mathcal{N}, \mathcal{A})$ conveys all the measurements

$$y_k^i = C_k^i x_k + v_k^i \,, \tag{2.19}$$

$$v_k^i \sim \mathcal{N}\Big(0, R_k^i\Big) \,, \tag{2.20}$$

$i \in \mathcal{N}$, to a fusion center in order to evaluate (2.10). The result amounts to stack all the information from all nodes $i \in \mathcal{N}$ as follows

$$y_k = \mathrm{col}\Big(y_k^i\Big)_{i\in\mathcal{N}} \tag{2.21}$$

$$C_k = \mathrm{col}\Big(C_k^i\Big)_{i\in\mathcal{N}} \tag{2.22}$$

$$R_k = \mathrm{diag}\Big(R_k^i\Big)_{i\in\mathcal{N}} \tag{2.23}$$

and then to perform the same steps of the KF. A summary of the *Multi-Sensor KF* (MSKF) is reported in Table 2.2.

The KF has the advantage of being Bayesian optimal, but is not directly applicable to nonlinear state-space models. Two well known approximations have proven to be effective in situations where one or both the equations (2.2) and (2.4) are nonlinear: I) Extended KF (EKF) [MSS62] and II) Unscented KF (EKF) [JU97]. The EKF is a first order approximation of the Kalman filter based on local linearization. The UKF uses the sampling principles of

Table 2.2: The Multi-Sensor Kalman Filter (MSKF)

for $k = 1, 2, \ldots$ **do**

Prediction

$\hat{x}_{k|k-1} = A_{k-1}\hat{x}_{k-1}$ ▷ Predicted mean

$P_{k|k-1} = A_{k-1}P_{k-1}A_{k-1}^{\top} + Q_{k-1}$ ▷ Predicted covariance matrix

Stacking

$y_k = \mathrm{col}(y_k^i)_{i\in\mathcal{N}}$

$C_k = \mathrm{col}(C_k^i)_{i\in\mathcal{N}}$

$R_k = \mathrm{diag}(R_k^i)_{i\in\mathcal{N}}$

Correction

$e_k = y_k - C_k\hat{x}_{k|k-1}$ ▷ Innovation

$S_k = R_k + C_kP_{k|k-1}C_k^{\top}$ ▷ Innovation covariance matrix

$K_k = P_{k|k-1}C_k^{\top}S_k^{-1}$ ▷ Kalman gain

$\hat{x}_k = \hat{x}_{k|k-1} + K_ke_k$ ▷ Updated mean

$P_k = P_{k|k-1} - K_kS_kK_k^{\top}$ ▷ Updated covariance matrix

end for

the *Unscented Transform* (UT) [JUDW95] to propagate the first and second order moments of the predicted and updated densities.

2.2.4 The Extended Kalman Filter

This subsection presents the EKF which is basically an extension of the linear KF whenever one or both the equations (2.2) and (2.4) are nonlinear transformations of the state with additive Gaussian white noise [MSS62], i.e.

$$x_k = f_{k-1}(x_{k-1}) + w_{k-1}\,, \tag{2.24}$$

$$y_k = h_k(x_k) + v_k\,. \tag{2.25}$$

The prediction equations of the EKF are of the same form as the KF, with the transition matrix A_{k-1} of (2.11) evaluated via linearization about the updated mean $\hat{x}_{k-1}$, i.e.

$$A_{k-1} = \left.\frac{\partial f_{k-1}(\cdot)}{\partial x}\right|_{x=\hat{x}_{k-1}}\,. \tag{2.26}$$

The correction equations of the EKF are also of the same form as the KF, with the observation matrix C_k of (2.12) evaluated via linearization about the predicted mean $\hat{x}_{k|k-1}$, i.e.

$$C_k = \left.\frac{\partial h_k(\cdot)}{\partial x}\right|_{x=\hat{x}_{k|k-1}}\,. \tag{2.27}$$

The EKF recursion for computing both predicted and updated pairs $(\hat{x}_{k-1}, P_{k-1})$ and $(\hat{x}_k, P_k)$

is reported in Table 2.3.

Table 2.3: The Extended Kalman Filter (EKF)

for $k = 1, 2, \ldots$ **do**

Prediction

$A_{k-1} = \left.\dfrac{\partial f_{k-1}(\cdot)}{\partial x}\right|_{x=\hat{x}_{k-1}}$ ▷ Linearization about the updated mean

$\hat{x}_{k|k-1} = A_{k-1}\hat{x}_{k-1}$ ▷ Predicted mean

$P_{k|k-1} = A_{k-1}P_{k-1}A_{k-1}^{\top} + Q_{k-1}$ ▷ Predicted covariance matrix

Correction

$C_k = \left.\dfrac{\partial h_k(\cdot)}{\partial x}\right|_{x=\hat{x}_{k|k-1}}$ ▷ Linearization about the predicted mean

$e_k = y_k - C_k\hat{x}_{k|k-1}$ ▷ Innovation

$S_k = R_k + C_kP_{k|k-1}C_k^{\top}$ ▷ Innovation covariance matrix

$K_k = P_{k|k-1}C_k^{\top}S_k^{-1}$ ▷ Kalman gain

$\hat{x}_k = \hat{x}_{k|k-1} + K_ke_k$ ▷ Updated mean

$P_k = P_{k|k-1} - K_kS_kK_k^{\top}$ ▷ Updated covariance matrix

end for

2.2.5 The Unscented Kalman Filter

The UKF is based on the *Unscented Transform* (UT), a derivative-free technique capable of providing a more accurate statistical characterization of a random variable undergoing a nonlinear transformation [JU97]. In particular, the UT is a deterministic technique suited to provide an approximation of the mean and covariance matrix of a given random variable subjected to a nonlinear transformation via a minimal set of its samples. Let us consider the mean m and associated covariance matrix P of a generic random variable along with a nonlinear transformation function $g(\cdot)$, the UT proceeds as follows:

- generates $2n_x + 1$ samples $X \in \mathbb{R}^{n_x \times (2n_x+1)}$, the so called σ-points, starting from the mean m with deviation given by the matrix square root Σ of P;
- propagates the σ-points through the nonlinear transformation function $g(\cdot)$ resulting in $G \in \mathbb{R}^{n_x \times (2n_x+1)}$;
- calculates the new transformed mean m' and associated covariance matrix P_{gg} as well as the cross-covariance matrix P_{xg} of the initial and transformed σ-points.

The pseudo-code of the UT is reported in Table 2.4.

Given three parameters α_σ, β_σ and κ_σ, the weights c, w_m and W_c are calculated exploiting the algorithm in Table 2.5. Moment matching properties and performance improvements are discussed in [JU97,WvdM01] by resorting to specific values of α_σ, β_σ and κ_σ. It is of common

Table 2.4: The Unscented Transformation (UT)

procedure UT(m, P, g)
c, w_m, $W_c = \text{UTW}(\alpha_\sigma, \beta_\sigma, \kappa_\sigma)$ ▷ Weights are calculated exploiting UTW in Table 2.5
$\Sigma = \sqrt{P}$
$X = \Big[m \dots m\Big] + \sqrt{c}\,\Big[\underline{0}, \Sigma, -\Sigma\Big]$ ▷ $\underline{0}$ is a zero column vector
$G = g(X)$ ▷ $g(\cdot)$ is applied to each column of X
$m' = G w_m$
$P_{gg} = G W_c G^\top$
$P_{yg} = G W_c G^\top$
Return m', P_{gg}, P_{xg}
end procedure

Table 2.5: Unscented Transformation Weights (UTW)

procedure UTW(α_σ, β_σ, κ_σ)
$\varsigma = \alpha_\sigma^2(n_x + \kappa_\sigma) - n_x$
$w_m^{(0)} = \varsigma\,(n_x + \varsigma)^{-1}$
$w_c^{(0)} = \varsigma\,(n_x + \varsigma)^{-1} + (1 - \alpha_\sigma^2 + \beta_\sigma)$
$w_m^{(1,\dots,2n_x)}, w_c^{(1,\dots,2n_x)} = [2(n_x + \varsigma)]^{-1}$
$w_m = \Big[w_m^{(0)}, \dots, w_m^{(2n_x)}\Big]^\top$
$w_c = \Big[w_c^{(0)}, \dots, w_c^{(2n_x)}\Big]^\top$
$W_c = (I - [w_m \dots w_m])\,\text{diag}\Big(w_c^{(0)} \dots w_c^{(2n)}\Big)\,(I - [w_m \dots w_m])^\top$
$c = \alpha_\sigma^2(n_x + \kappa_\sigma)$
Return c, w_m, W_c
end procedure

practice to set these three parameters as constants, thus computing the weights once at the beginning of the estimation process.

The UT can be applied in the KF recursion allowing to obtain a nonlinear recursive estimator known as UKF [JU97]. The pseudo-code of the UKF is shown in Table 2.6.

The main advantages of the UKF approach are the following:

- it does not require the calculation of the Jacobians (2.26) and (2.27). The UKF algorithm is, therefore, very suitable for highly nonlinear problems and represents a good trade-off between accuracy and numerical efficiency;
- being derivative free, it can cope with functions with jumps and discontinuities;
- it is capable of capturing higher order moments of nonlinear transformations [JU97].

Due to the above mentioned benefits, the UKF is herewith adopted as the nonlinear recursive estimator.

The KF represents the basic tool for recursively estimating the state, in a Bayesian framework, of a moving object, i.e. to perform *Single-Object Filtering* (SOF) [FS85, FS86, BSF88,

Table 2.6: The Unscented Kalman Filter (UKF)

for $k = 1, 2, \ldots$ **do**

Prediction

$\hat{x}_{k|k-1}, P_{k|k-1} = \mathrm{UT}\Big(\hat{x}_{k-1|k-1}, P_{k-1|k-1}, f(\cdot)\Big)$

$P_{k|k-1} = P_{k|k-1} + Q$

Correction

$\hat{y}_{k|k-1}, S_k, C_k = \mathrm{UT}\Big(\hat{x}_{k|k-1}, P_{k|k-1}, h(\cdot)\Big)$

$S_k = S_k + R$

$\hat{x}_k = \hat{x}_{k|k-1} + C_k S_k^{-1}\left(y_k - \hat{y}_{k|k-1}\right)$

$P_k = P_{k|k-1} - C_k S_k^{-1} C_k^{\top}$

end for

BSL95, BSLK01, BP99]. A natural evolution of SOF is *Multi-Object Tracking* (MOT), which involves the on-line estimation of an unknown and (possibly) time-varying number of objects and their individual trajectories from sensor data [BP99, BSF88, Mah07b]. The key challenges in multi-object tracking include *detection uncertainty*, *clutter*, and *data association uncertainty*. Numerous multi-object tracking algorithms have been developed in the literature and most of these fall under three major paradigms: *Multiple Hypotheses Tracking* (MHT) [Rei79, BP99]; *Joint Probabilistic Data Association* (JPDA) [BSF88]; and *Random Finite Set* (RFS) filtering [Mah07b]. In this thesis, the focus is on the RFS formulation since it provides the concept of *probability density* for *multi-object state* that allows to directly extend the single-object Bayesian recursion. Such a concept is not available in the MHT and JPDA approaches [Rei79, FS85, FS86, BSF88, BSL95, BP99].

2.3 Random finite set approach

The second key concept of the present dissertation is the RFS [GMN97, Mah07b, Mah04, Mah13] approach. In the case where the need is to recursively estimate the state of a possibly time varying number of multiple dynamical systems, RFSs allow to generalize standard Bayesian filtering to a unified framework. In particular, states and observations will be modelled as RFSs where not only the single state and observation are random, but also their number (set cardinality). The purpose of this section is to cover the aspects of the RFS approach that will be useful for the subsequent chapters. Overviews on RFSs and further advanced topics concerning point process theory, stochastic geometry and measure theory can be found in [Mat75, SKM95, GMN97, Mah03, VSD05, Mah07b].

Remark 2. It is worth pointing out that in multi-object scenarios there is a subtle difference between *filtering* and *tracking*. In particular, the first refers to estimating the state of a possibly time-varying number of objects without, however, uniquely identifying them, i.e. after having estimated a multi-object density a decision-making operation is needed to extract

the objects themselves. On the other hand, the term "tracking" refers to jointly estimating a possibly time-varying number of objects and to uniquely mark them over time so that no decision-making operation has to be carried out and object trajectories are well defined. Finally, it is clear that in single-object scenarios the terms *filtering* and *tracking* can be used interchangeably.

2.3.1 Notation

Throughout the thesis, finite sets are represented by uppercase letters, e.g. X, $\mathbf{X}$. The following multi-object exponential notation is used

$$h^X \triangleq \prod_{x \in X} h(x)\,, \tag{2.28}$$

where h is a real-valued function, with $h^{\varnothing} = 1$ by convention [Mah07b]. The following generalized Kronecker delta [VV13, VVP14] is also adopted

$$\delta_Y(X) \triangleq \begin{cases} 1, \text{ if } X = Y \\ 0, \text{ otherwise} \end{cases}, \tag{2.29}$$

along with the inclusion function, a generalization of the indicator function, defined as

$$1_Y(X) \triangleq \begin{cases} 1, \text{ if } X \subseteq Y \\ 0, \text{ otherwise} \end{cases}. \tag{2.30}$$

The shortand notation $1_Y(x)$ is used in place of $1_Y(\{x\})$ whenever $X = \{x\}$. The cardinality (number of elements) of the finite set X is denoted by $|X|$. The following PDF notation will be also used.

$$\text{Poisson}_{[\lambda]}(n) = \frac{e^{-\lambda}\lambda^n}{n!}\,, \lambda \in \mathbb{N}\,, n \in \mathbb{N}\,, \tag{2.31}$$

$$\text{Uniform}_{[a,b]}(n) = \begin{cases} \dfrac{1}{b-a}\,, & n \in [a,b] \\ 0\,, & n \notin [a,b] \end{cases}, a \in \mathbb{R}\,, b \in \mathbb{R}\,, a < b\,. \tag{2.32}$$

2.3.2 Random finite sets

In a typical multiple object scenario, the number of objects varies with time due to their appearance and disappearance. The sensor observations are affected by misdetection (e.g., occlusions, low radar cross section, etc.) and false alarms (e.g., observations from the environment, clutter, etc.). This is further compounded by association uncertainty, i.e. it is not known which object generated which measurement. The objective of multi-object filtering is to jointly estimate over time the number of objects and their states from the observation

history.

In this thesis we adopt the RFS formulation, as it provides the concept of *probability density of the multi-object state* that allows us to directly generalize (single-object) estimation to the multi-object case. Indeed, from an estimation viewpoint, the multi-object system state is naturally represented as a finite set [VVPS10]. More concisely, suppose that at time k, there are N_k objects with states $x_{k,1}, \ldots, x_{k,N_k}$, each taking values in a state space $\mathbb{X} \subseteq \mathbb{R}^{n_x}$, i.e. the *multi-object state* at time k is the finite set

$$X_k = \{x_{k,1}, \ldots, x_{k,N_k}\} \subset \mathbb{X}. \tag{2.33}$$

Since the multi-object state is a finite set, the concept of RFS is required to model, in a probabilistic way, its uncertainty.

An RFS X on a space $\mathbb{X}$ is a random variable taking values in $\mathcal{F}(\mathbb{X})$, the space of finite subsets of $\mathbb{X}$. The notation $\mathcal{F}_n(\mathbb{X})$ will be also used to refer to the space of finite subsets of $\mathbb{X}$ with exactly n elements.

Definition 1. An RFS X is a random variable that takes values as (unordered) finite sets, i.e. a finite-set-valued random variable.

At the fundamental level, like any other random variable, an RFS is described by its probability distribution or probability density.

Remark 3. What distinguishes an RFS from a random vector is that: I) the number of points is random; II) the points themselves are random and unordered.

The space $\mathcal{F}(\mathbb{X})$ does not inherit the usual Euclidean notion of integration and density. In this thesis, we use the *FInite Set STatistics* (FISST) notion of integration/density to characterize RFSs [Mah03, Mah07b].

From a probabilistic viewpoint, an RFS X is completely characterized by its *multi-object density* $f(X)$. In fact, given $f(X)$, the cardinality *Probability Mass Function* (PMF) $\rho(n)$ that X have $n \geq 0$ elements and the joint conditional PDFs $f(x_1, x_2, \ldots, x_n|n)$ over $\mathbb{X}^n$ given that X have n elements, can be obtained as follows:

$$\rho(n) = \frac{1}{n!} \int_{\mathbb{X}^n} f\left(\{x_1, \ldots, x_n\}\right) \, dx_1 \cdots dx_n \tag{2.34}$$

$$f(x_1, \ldots, x_n|n) = \frac{1}{n!\,\rho(n)} \, f\left(\{x_1, \ldots, x_n\}\right) \tag{2.35}$$

Remark 4. The multi-object density $f(X)$ is nothing but the multi-object counterpart of the state PDF in the single-object case.

In order to measure probability over subsets of $\mathbb{X}$ or compute expectations of random set variables, it is convenient to introduce the following definition of *set integral* for a generic

real-valued function $g(X)$ (not necessarily a multi-object density) of an RFS variable X:

$$\int_{\mathbb{X}} g(X)\,\delta X \triangleq \sum_{n=0}^{\infty} \frac{1}{n!} \int_{\mathbb{X}^n} g(\{x_1,\ldots,x_n\})\,dx_1\cdots dx_n \tag{2.36}$$
$$= g(\varnothing) + \int_{\mathbb{X}} g(\{x\})\,dx + \frac{1}{2}\int_{\mathbb{X}^2} g(\{x_1,x_2\})\,dx_1 dx_2 + \cdots$$

In particular,

$$\beta(X) \triangleq \text{Prob}(X \subset \mathbb{X}) = \int_{\mathbb{X}} f(X)\,\delta X \tag{2.37}$$

measures the probability that the RFS X is included in the subset $\mathbb{X}$ of $\mathbb{R}^{n_x}$. The function $\beta(X)$ is also known as the *Belief-Mass Function* (BMF) [Mah07b].

Remark 5. The BMF $\beta(X)$ is nothing but the multi-object counterpart of the state *Cumulative Distribution Function* (CDF) in the single-object case.

It is also easy to see that, thanks to the set integral definition (2.36), the multi-object density, like the single-object state PDF, satisfies the trivial normalization constraint

$$\int_{\mathbb{X}} f(X)\,\delta X = 1. \tag{2.38}$$

It is worth pointing out that the multi-object density, while completely characterizing an RFS, involves a combinatorial complexity; hence simpler, though incomplete, characterizations are usually adopted in order to keep the *Multi-Object Filtering* (MOF) problem computationally tractable. In this respect, the *first-order moment* of the multi-object density, better known as *Probability Hypothesis Density* (PHD) or *intensity function*, has been found to be a very successful characterization [Mah04, Mah03, Mah07b]. In order to define the PHD function, let us introduce the number of elements of the RFS X which is given by

$$N_X = \int_{\mathbb{X}} \phi_X(x)\,dx\,, \tag{2.39}$$

where

$$\phi_X(x) \triangleq \sum_{\xi\in X} \delta_x(\xi)\ . \tag{2.40}$$

We would like to define the PHD function $d(x)$ of X over the state space $\mathbb{X}$ so that the expected number of elements of X in $\mathbb{X}$ is obtained by integrating $d(\cdot)$ over $\mathbb{X}$, i.e.

$$\text{E}[N_X] = \int_{\mathbb{X}} d(x)\,dx\,. \tag{2.41}$$

Since

$$\text{E}[N_X] = \int N_X f(X)\,\delta X \tag{2.42}$$
$$= \int \left[\int_{\mathbb{X}} \phi_X(x)\,dx\right] f(X)\,\delta X \tag{2.43}$$

$$= \int_{\mathbb{X}} \left[\int \phi_X(x)\, f(X)\, \delta X \right] dx\,, \tag{2.44}$$

comparing (2.41) with (2.44), it turns out that

$$d(x) \triangleq \mathrm{E}[\phi_X(x)] = \int \phi_X(x)\, f(X)\, \delta X \tag{2.45}$$

Without loss of generality, the PHD function can be expressed in the form

$$d(x) = \overline{n}\, s(x) \tag{2.46}$$

where

$$\overline{n} = \mathrm{E}[n] = \mathrm{E}[n(\mathbb{X})] = \sum_{n=0}^{\infty} n\rho(n) \tag{2.47}$$

is the expected number of objects and $s(\cdot)$ is a single-object PDF, called *location density*, such that

$$\int_{\mathbb{X}} s(x)\, dx = 1\,. \tag{2.48}$$

It is worth to highlight that, in general, the PHD function $d(\cdot)$ and the cardinality PMF $\rho(\cdot)$ do not completely characterize the multi-object distribution. However, for specific RFSs defined in the next section, the characterization is complete.

2.3.3 Common classes of RFS

A review of the common RFS densities is provided [Mah07b] hereafter.

Poisson RFS

A *Poisson* RFS X on $\mathbb{X}$ is uniquely characterized by its intensity function $d(\cdot)$. The Poisson RFSs have the unique property that the distribution of the cardinality of X is Poisson with mean

$$D = \int_{\mathbb{X}} d(x)\, dx\,, \tag{2.49}$$

and for a given cardinality the elements of X are i.i.d. with probability density

$$s(x) = \frac{d(x)}{D}\,. \tag{2.50}$$

The probability density of X can be written as

$$\pi(X) = e^{-D} \prod_{x \in X} d(x) \tag{2.51}$$

The Poisson RFS is traditionally described as characterizing no spatial interaction or complete spatial randomness in the following sense. For any collection of disjoint subsets $B_i \in \mathbb{X}$, $i \in \mathbb{N}$, it can be shown that the count functions $|X \cap B_i|$ are independent random variables which

are Poisson distributed with mean

$$D_{B_i} = \int_{B_i} d(x)\, dx\,, \tag{2.52}$$

and for a given number of points occurring in B_i the individual points are i.i.d. according to

$$\frac{d(x)\, 1_{B_i}(x)}{N_{B_i}}\,. \tag{2.53}$$

The procedure in Table 2.7 illustrates how to generate a sample from a Poisson RFS.

Table 2.7: Sampling a Poisson RFS

$X = \varnothing$
Sample $n \sim \text{Poisson}_{[D]}$
for $i = 1, \ldots, n$ **do**
Sample $x_i \sim s(\cdot)$
$X = X \cup \{x\}$
end for

Independent identically distributed cluster RFS

An i.i.d. cluster RFS X on $\mathbb{X}$ is uniquely characterized by its cardinality distribution $\rho(\cdot)$ and matching intensity function $d(\cdot)$. The cardinality distribution must satisfy

$$D = \sum_{n=0}^{\infty} n\rho(n) = \int_{\mathbb{X}} d(x)\, dx\,, \tag{2.54}$$

but can otherwise be arbitrary, and for a given cardinality the elements of X are i.i.d. with probability density

$$s(x) = \frac{d(x)}{D} \tag{2.55}$$

The probability density of an i.i.d. cluster RFS can be written as

$$\pi(X) = |X|!\, \rho(|X|) \prod_{x \in X} s(x) \tag{2.56}$$

Note that an i.i.d. cluster RFS essentially captures the spatial randomness of the Poisson RFS without the restriction of a Poisson cardinality distribution. The procedure in Table 2.8 illustrates how a sample from an i.i.d. RFS is generated.

Bernoulli RFS

A Bernoulli RFS X on $\mathbb{X}$ has probability $q \triangleq 1 - r$ of being empty, and probability r of being a singleton whose only element is distributed according to a probability density p defined

Table 2.8: Sampling an i.i.d. RFS

$X = \varnothing$
Sample $n \sim \rho(\cdot)$
for $i = 1, \ldots, n$ **do**
Sample $x_i \sim s(\cdot)$
$X = X \cup \{x\}$
end for

on $\mathbb{X}$. The cardinality distribution of a Bernoulli RFS is thus a Bernoulli distribution with parameter r. A Bernoulli RFS is completely described by the parameter pair $(r, p(\cdot))$.

Multi-Bernoulli RFS

A multi-Bernoulli RFS X on $\mathbb{X}$ is a union of a fixed number of independent Bernoulli RFSs $X^{(i)}$ with *existence probability* $r^{(i)} \in (0, 1)$ and probability density $p^{(i)}(\cdot)$ defined on $\mathbb{X}$ for $i = 1, \ldots, I$, i.e.

$$X = \bigcup_{i=1}^{I} X^{(i)} . \tag{2.57}$$

It follows that the mean cardinality of a multi-Bernoulli RFS is

$$D = \sum_{i=1}^{I} r^{(i)} \tag{2.58}$$

A multi-Bernoulli RFS is thus completely described by the corresponding multi-Bernoulli parameter set $\left\{\left(r^{(i)}, p^{(i)}(\cdot)\right)\right\}_{i=1}^{I}$. Its probability density is

$$\pi(X) = \prod_{j=1}^{I} \left(1 - r^{(j)}\right) \sum_{1 \leq i_1 \neq \cdots \neq i_{|X|} \leq I} \prod_{j=1}^{|X|} \frac{r^{(i_j)} \, p^{(i_j)}(x_j)}{1 - r^{(i_j)}} \tag{2.59}$$

For convenience, probability densities of the form (2.59) are abbreviated by the form $\pi(X) = \left\{\left(r^{(i)}, p^{(i)}\right)\right\}_{i=1}^{I}$. A multi-Bernoulli RFS jointly characterizes unions of non-interacting points with less than unity probability of occurrence and arbitrary spatial distributions. The procedure in Table 2.9 illustrates how a sample from a multi-Bernoulli RFS is generated.

2.3.4 Bayesian multi-object filtering

Let us now introduce the basic ingredients of the MOF problem [Mah07b, Mah04, Mah13, Mah14], i.e. the object RFS $X_k \subset \mathbb{X}$ at time k and the RFS Y_k^i of measurements gathered by node $i \in \mathcal{N}$ at time k. It is also convenient to define the overall measurement RFSs at time k,

$$Y_k \triangleq \bigcup_{i \in \mathcal{N}} Y_k^i , \tag{2.60}$$

Table 2.9: Sampling a multi-Bernoulli RFS

$X = \varnothing$
for $i = 1, \ldots, I$ **do**
  Sample $u \sim \text{Uniform}_{[0,1]}$
  if $u \leq r^{(i)}$ **then**
    Sample $x \sim p^{(i)}(\cdot)$
    $X = X \cup \{x\}$
  end if
end for

and up to time k,

$$Y_{1:k} \triangleq \bigcup_{\kappa=1}^{k} Y_\kappa \,. \tag{2.61}$$

In the random set framework, the Bayesian approach to MOF consists, therefore, of recursively estimating the object set X_k conditioned to the observations $Y_{1:k}$. The object set is assumed to evolve according to a multi-object dynamics

$$X_k = \Phi_{k-1}(X_{k-1}) \cup B_{k-1} \tag{2.62}$$

where B_{k-1} is the RFS of *new-born* objects at time $k-1$ and

$$\Phi_{k-1}(X) = \bigcup_{x \in X} \phi_{k-1}(x) \;, \tag{2.63}$$

$$\phi_{k-1}(x) = \begin{cases} \{x'\} \,, & \text{with survival probability } P_{S,k-1} \\ \varnothing \,, & \text{otherwise} \end{cases} \tag{2.64}$$

Notice that according to (2.62)-(2.64) each object in the set X_k is either a new-born object from the set B_{k-1} or an object survived from X_{k-1}, with probability $P_{S,k-1}$, and whose state vector x' has evolved according to the single-object dynamics (2.2) with $x = x_{k-1}$ and $x' = x_k$. In a similar way, observations are assumed to be generated, at each node $i \in \mathcal{N}$, according to the measurement model

$$Y_k^i = \Psi_k^i(X_k) \cup \mathcal{C}_k^i \tag{2.65}$$

where $\mathcal{C}_k^i$ is the *clutter* RFS (i.e. the set of measurements not due to objects) at time k and node i, and

$$\Psi_k^i(X) = \bigcup_{x \in X} \psi_k^i(x) \;, \tag{2.66}$$

$$\psi_k^i(x) = \begin{cases} \{y_k\} \,, & \text{with detection probability } P_{D,k} \\ \varnothing \,, & \text{otherwise} \end{cases} \tag{2.67}$$

Notice that according to (2.65)-(2.67) each measurement in the set Y_k^i is either a false one from the clutter set $\mathcal{C}_k^i$ or is related to an object in X_k, with probability $P_{D,k}$, according to the single-sensor measurement equation (2.4). Hence it is clear how one of the major benefits of the random set approach is to directly include in the MOF problem formulation (2.62)-(2.67) fundamental practical issues such as object birth and death, false alarms (clutter) and missed detections. Following a Bayesian approach, the aim is to propagate in time the *posterior* (*filtered*) *multi-object densities* $\pi_k(X)$ of the object set X_k given $Y_{1:k}$, as well as the *prior* (*predicted*) ones $\pi_{k|k-1}(X)$ of X_k given $Y_{1:k-1}$. Exploiting random set theory, it has been found [Mah07b] that the multi-object densities follow recursions that are conceptually analogous to the well-known ones of Bayesian nonlinear filtering, i.e.

$$\pi_{k|k-1}(X) = \int \varphi_{k|k-1}(X|Z)\, \pi_{k-1}(Z)\, \delta Z\,, \tag{2.68}$$

$$\pi_k(X) = \frac{g_k(Y_k|X)\, \pi_{k|k-1}(X)}{\int g_k(Y_k|Z)\, \pi_{k|k-1}(Z)\, \delta Z}\,, \tag{2.69}$$

where $\varphi_{k|k-1}(\,\cdot\,|\,\cdot\,)$ is the *multi-object transition density* to time k, $g_k(\,\cdot\,|\,\cdot\,)$ is the *multi-object likelihood function* at time k [Mah03, Mah07a, Mah07b].

Remark 6. Despite this conceptual resemblance, however, the multi-object Bayes recursions (2.68)-(2.69), unlike their single-object counterparts, are affected by combinatorial complexity and are, therefore, computationally infeasible except for very small-scale MOF problems involving few objects and/or measurements.

For this reason, the computationally tractable PHD and *Cardinalized PHD* (CPHD) filtering approaches will be followed and briefly reviewed.

2.3.5 CPHD filtering

The CPHD filter [Mah07a, VVC07] propagates in time the cardinality PMFs $\rho_{k|k-1}(n)$ and $\rho_k(n)$ as well as the PHD functions $d_{k|k-1}(x)$ and $d_k(x)$ of X_k given $Y_{1:k-1}$ and, respectively, $Y_{1:k}$ assuming that the clutter RFS, the predicted and filtered RFSs are i.i.d. cluster processes (see (2.56)). The resulting CPHD recursions (prediction and correction) are as follows

Prediction

$$\begin{aligned} \rho_{k|k-1}(n) &= \sum_{j=0}^{n} p_b(n-j)\ \rho_{S,k|k-1}(j) \\ d_{k|k-1}(x) &= d_{b,k}(x) + \int \varphi_{k|k-1}(x|\zeta)\ P_{S,k-1}(\zeta)\, d_{k-1}(\zeta)\, d\zeta \end{aligned} \tag{2.70}$$

Correction

$$\rho_k(n) = \frac{\mathcal{G}_k^0\Big(d_{k|k-1}(\cdot)\,, Y_k, n\Big)\ \rho_{k|k-1}(n)}{\sum_{\eta=0}^{\infty} \mathcal{G}_k^0\Big(d_{k|k-1}(\cdot)\,, Y_k, \eta\Big)\ \rho_{k|k-1}(\eta)} \tag{2.71}$$
$$d_k(x) = \mathcal{G}_{Y_k}(x)\ d_{k|k-1}(x)$$

where: $p_{b,k}(\cdot)$ is the assumed birth cardinality PMF; $\rho_{S,k|k-1}(\cdot)$ is the cardinality PMF of survived objects, given by

$$\rho_{S,k|k-1}(j) = \sum_{t=j}^{\infty} \left(\begin{array}{c} t \\ j \end{array} \right) P_{S,k}^{j}\,(1 - P_{S,k})^{h-j}\,\rho_{k-1}(t)\ ; \tag{2.72}$$

$d_{b,k}(\cdot)$ is the assumed PHD function of new-born objects; $\varphi_{k|k-1}(x|\xi)$ is the state transition PDF (2.3) associated to the single object dynamics; the generalized likelihood functions $\mathcal{G}_k^0(\cdot,\cdot,\cdot)$ and $\mathcal{G}_{Y_k}(\cdot)$ have cumbersome expressions which can be found in [VVC07].

Enforcing Poisson cardinality distributions for the clutter and object sets, the update of the cardinality PMF is no longer needed and the CPHD filter reduces to the PHD filter [Mah03,VM06]. The CPHD filter provides enhanced robustness with respect to misdetections, clutter as well as various uncertainties on the multi-object model (i.e. detection and survival probabilities, clutter and birth distributions) at the price of an increased computational load: a CPHD recursion has $O(m^3 n_{max})$ complexity compared to $O(mn_{max})$ of PHD, m being the number of measurements and n_{max} the maximum number of objects. There are essentially two ways of implementing PHD/CPHD filters, namely the *Particle Filter* (PF) or *Sequential Monte Carlo* (SMC) [VSD05] and the *Gaussian Mixture* (GM) [VM06, VVC07] approaches, the latter being cheaper in terms of computation and memory requirements and, hence, by far preferable for distributed implementation on a sensor network wherein nodes have limited processing and communication capabilities.

2.3.6 Labeled RFSs

The aim of MOT is not only to recursively estimate the state of the objects and their time varying number, but also to keep track of their trajectories. To this end, the notion of label is introduced in the RFS approach [VV13, VVP14] so that each estimated object can be uniquely identified and its track reconstructed.

Let $\mathcal{L} : \mathbb{X}\times\mathbb{L} \to \mathbb{L}$ be the projection $\mathcal{L}((x,\ell)) = \ell$, where $\mathbb{X} \subseteq \mathbb{R}^{n_x}$, $\mathbb{L} = \{\alpha_i : i \in \mathbb{N}\}$ and α_i's are distinct. To estimate the trajectories of the objects they need to be uniquely identified by a (unobserved) label drawn from a discrete countable space $\mathbb{L} = \{\alpha_i : i \in \mathbb{N}\}$, where the α_i's are distinct. To incorporate object identity, a label $\ell \in \mathbb{L}$ is appended to the state x of each object and the multi-object state is regarded as a finite set on $\mathbb{X} \times \mathbb{L}$, i.e. $\mathbf{x} \triangleq (x,\ell) \in \mathbb{X}\times\mathbb{L}$. However, this idea alone is not enough since $\mathbb{L}$ is discrete and it is possible (with non-zero probability) that multiple objects have the same identity, i.e. be marked with

the same label. This problem can be alleviated using a special case of RFS called *labeled* RFS [VV13, VVP14], which is, in essence, a marked RFS with distinct labels. Then, a finite subset $\mathbf{X}$ of $\mathbb{X}\times\mathbb{L}$ has distinct labels if and only if $\mathbf{X}$ and its labels $\mathcal{L}(\mathbf{X}) = \{\mathcal{L}(\mathbf{x}) : \mathbf{x} \in \mathbf{X}\}$ have the same cardinality, i.e. $|\mathbf{X}| = |\mathcal{L}(\mathbf{X})|$. The function $\Delta(\mathbf{X}) \triangleq \delta_{|\mathbf{X}|}(|\mathcal{L}(\mathbf{X})|)$ is called the *distinct label indicator* and assumes value 1 if and only if the condition $|\mathbf{X}| = |\mathcal{L}(\mathbf{X})|$ holds.

Definition 2. A labeled RFS with state space $\mathbb{X}$ and (discrete) label space $\mathbb{L}$ is an RFS on $\mathbb{X} \times \mathbb{L}$ such that each realization $\mathbf{X}$ has distinct labels, i.e.

$$|\mathbf{X}| = |\mathcal{L}(X)| \ . \tag{2.73}$$

The unlabeled version of a labeled RFS with density $\boldsymbol{\pi}$ is distributed according to the *marginal density* [VV13]

$$\pi(\{x_1, \ldots, x_n\}) = \sum_{(\ell_1,\ldots,\ell_n)\in\mathbb{L}^n} \boldsymbol{\pi}(\{(x_1, \ell_1), \ldots, (x_n, \ell_n)\}) \ . \tag{2.74}$$

The unlabeled version of a labeled RFS is its projection from $\mathbb{X} \times \mathbb{L}$ into $\mathbb{X}$, and is obtained by simply discarding the labels. The cardinality distribution of a labeled RFS is the same as its unlabeled version [VV13].

Hereinafter, symbols for labeled states and their distributions are bolded to distinguish them from unlabeled ones, e.g. $\mathbf{x}$, $\mathbf{X}$, $\boldsymbol{\pi}$, etc.

2.3.7 Common classes of labeled RFS

A review of the common labeled RFS densities is provided [VV13].

Labeled Poisson RFS

A labeled Poisson RFS $\mathbf{X}$ with state space $\mathbb{X}$ and label space $\mathbb{L} = \{\alpha_i : i \in \mathbb{N}\}$, is a Poisson RFS X on $\mathbb{X}$ with intensity $d(\cdot)$, tagged with labels from $\mathbb{L}$. A sample from such labeled Poisson RFS can be generated by the procedure reported in Table 2.10.

Table 2.10: Sampling a labeled Poisson RFS

$\mathbf{X} = \varnothing$
Sample $n \sim \text{Poisson}_{[D]}$
for $i = 1, \ldots, n$ **do**
Sample $x \sim d(\cdot)/D$
$\mathbf{X} = \mathbf{X} \cup \{(x, \alpha_i)\}$
end for

The probability density of an LMB RFS is given by

$$\boldsymbol{\pi}(\mathbf{X}) = \delta_{\mathbb{L}(|\mathbf{X}|)}(\mathcal{L}(\mathbf{X}))\ \text{Poisson}_{[D]}(|\mathbf{X}|) \prod_{\mathbf{x}\in\mathbf{X}} \frac{d(x)}{D} \,, \tag{2.75}$$

where $\mathbb{L}(n) = \{\alpha_i \in \mathbb{L}\}_{i=1}^{n}$; $\delta_{\mathbb{L}(n)}(\{\ell_1, \ldots, \ell_n\})$ serves to check whether or not labels are distinct.

Remark 7. By a tracking point of view, each label of a labeled Poisson RFS $\mathbf{X}$ cannot directly refer to an object since all the corresponding states are sampled from the same intensity $d(x)$, i.e. there is not a 1-to-1 mapping between labels and location PDFs. Thus, it is not clear how to properly use such a labeled distribution in a MOT problem.

Labeled independent identically distributed cluster RFS

In the same fashion of the unlabeled RFSs, the labeled Poisson RFS can be generalized to the labeled i.i.d. cluster RFS by removing the Poisson assumption on the cardinality and specifying an arbitrary cardinality distribution. A sample from such a labeled i.i.d. RFS can be generated by the procedure reported in Table 2.11.

Table 2.11: Sampling a labeled i.i.d. cluster RFS

$\mathbf{X} = \varnothing$
Sample $n \sim \rho(\cdot)$
for $i = 1, \ldots, n$ **do**
Sample $x \sim d(\cdot)/D$
$\mathbf{X} = \mathbf{X} \cup \{(x, \alpha_i)\}$
end for

The probability density of an LMB RFS is given by

$$\boldsymbol{\pi}(\mathbf{X}) = \delta_{\mathbb{L}(|\mathbf{X}|)}(\mathcal{L}(\mathbf{X}))\ \rho(|\mathbf{X}|) \prod_{\mathbf{x} \in \mathbf{X}} \frac{d(x)}{D}\,, \tag{2.76}$$

The same consideration drawn for the labeled Poisson RFS is inherited by the i.i.d. cluster RFS.

Generalized Labeled Multi-Bernoulli RFS

A *Generalized Labeled Multi-Bernoulli* (GLMB) RFS [VV13] is a labeled RFS with state space $\mathbb{X}$ and (discrete) label space $\mathbb{L}$ distributed according to

$$\boldsymbol{\pi}(\mathbf{X}) = \Delta(\mathbf{X}) \sum_{c \in \mathbb{C}} w^{(c)}(\mathcal{L}(\mathbf{X})) \left[p^{(c)}\right]^{\mathbf{X}} \tag{2.77}$$

where: $\mathbb{C}$ is a discrete index set; $w^{(c)}(L)$ and $p^{(c)}$ satisfy the normalization constraints:

$$\sum_{L \subseteq \mathbb{L}} \sum_{c \in \mathbb{C}} w^{(c)}(L) = 1\,, \tag{2.78}$$

$$\int p^{(c)}(x, \ell)\, dx = 1\,. \tag{2.79}$$

A GLMB can be interpreted as a mixture of $|\mathbb{C}|$ multi-object exponentials

$$w^{(c)}(\mathcal{L}(\mathbf{X}))\left[p^{(c)}\right]^{\mathbf{X}} . \tag{2.80}$$

Each term in the mixture (2.77) consists of the product of two factors:

1. a weight $w^{(c)}(\mathcal{L}(\mathbf{X}))$ that only depends on the labels $\mathcal{L}(\mathbf{X})$ of the multi-object state $\mathbf{X}$;
2. a multi-object exponential $\left[p^{(c)}\right]^{\mathbf{X}}$ that depends on the entire multi-object state.

The cardinality distribution of a GLMB is given by

$$\rho(n) = \sum_{L\in\mathcal{F}_n(\mathbb{L})}\sum_{c\in\mathbb{C}} w^{(c)}(L) , \tag{2.81}$$

from which, in fact, summing over all possible n implies (2.78). The PHD of the unlabeled version of a GLMB is given by

$$d(x) = \sum_{c\in\mathbb{C}}\sum_{\ell\in\mathbb{L}} p^{(c)}(x,\ell)\sum_{L\subseteq\mathbb{L}} 1_L(\ell)\, w^{(c)}(L) . \tag{2.82}$$

Remark 8. The GLMB RFS distribution has the appealing feature of having a 1-to-1 mapping between labels and location PDFs provided by the single-object densities $p^{(c)}(x,\ell)$, for each term of the multi-object exponential mixture indexed with c. However, in the context of MOT, it is not clear how to exploit (in particular how to implement) such a distribution [VV13, VVP14].

δ-Generalized Labeled Multi-Bernoulli RFS

A δ-Generalized Labeled Multi-Bernoulli (δ-GLMB) RFS with state space $\mathbb{X}$ and (discrete) label space $\mathbb{L}$ is a special case of a GLMB with

$$\mathbb{C} = \mathcal{F}(\mathbb{L})\times\Xi , \tag{2.83}$$

$$w^{(c)}(L) = w^{(I,\xi)}(L) = w^{(I,\xi)}\delta_I(L) , \tag{2.84}$$

$$p^{(c)}(\cdot) = p^{(I,\xi)}(\cdot) = p^{(\xi)}(\cdot) , \tag{2.85}$$

i.e. is distributed according to

$$\boldsymbol{\pi}(\mathbf{X}) = \Delta(\mathbf{X})\sum_{(I,\xi)\in\mathcal{F}(\mathbb{L})\times\Xi} w^{(I,\xi)}\delta_I(\mathcal{L}(\mathbf{X}))\left[p^{(\xi)}\right]^{\mathbf{X}} \tag{2.86}$$

$$= \Delta(\mathbf{X})\sum_{I\in\mathcal{F}(\mathbb{L})}\delta_I(\mathcal{L}(\mathbf{X}))\sum_{\xi\in\Xi} w^{(I,\xi)}\left[p^{(\xi)}\right]^{\mathbf{X}} , \tag{2.87}$$

where Ξ is a discrete space. In MOT a δ-GLMB can be used (and implemented) to represent the multi-object densities over time. In particular:

- each finite set $I \in \mathcal{F}(\mathbb{L})$ represents a different configurations of labels whose cardinality provides the number of objects;
- each $\xi \in \Xi$ represents a history of association maps, e.g. $\xi = (\theta_1, \ldots, \theta_k)$, where an association map at time $1 \leq j \leq k$ is a function θ_j which maps track labels at time j to a measurement at time j with the constraint that a track can generate at most one measurement, and a measurement can be assigned to at most one track;
- the pair (I, ξ) is called *hypothesis*.

To clarify the notion of δ-GLMB, let us consider the following two examples.

E1: Suppose the following two possibilities

1. 0.4 chance of 1 object with label ℓ_1, i.e. $I_1 = \{\ell_1\}$, and density $p(x, \ell_1)$;
2. 0.6 chance of 2 objects with, respectively, labels ℓ_1 and ℓ_2, i.e. $I_2 = \{\ell_1, \ell_2\}$, and densities $p(x, \ell_1)$ and $p(x, \ell_2)$.

Then, the δ-GLMB representation is:

$$\boldsymbol{\pi}(\mathbf{X}) = 0.4\, \delta_{\{\ell_1\}}(\mathcal{L}(\mathbf{X}))\; p(x, \ell_1) + 0.6\, \delta_{\{\ell_1, \ell_2\}}(\mathcal{L}(\mathbf{X}))\; p(x, \ell_1)\; p(x, \ell_2)\; . \tag{2.88}$$

Notice that in this example there are no association histories, i.e. $\Xi = \varnothing$. Thus (2.88) has exactly two hypotheses $(I_1, \varnothing)$ and $(I_2, \varnothing)$.

E2: Let us now consider the previous example with $\Xi \neq \varnothing$:

1. 0.4 chance of 1 object with label ℓ_1, i.e. $I_1 = \{\ell_1\}$, association histories ξ_1 and ξ_2, i.e. there are two hypotheses (I_1, ξ_1) and (I_1, ξ_2), weights $w^{(I_1, \xi_1)} = 0.3$ and $w^{(I_1, \xi_2)} = 0.1$, densities $p^{(\xi_1)}(x, \ell_1)$ and $p^{(\xi_2)}(x, \ell_1)$;
2. 0.6 chance of 2 objects with, respectively, label ℓ_1 and ℓ_2, i.e. $I_2 = \{\ell_1, \ell_2\}$, association histories ξ_1 and ξ_2, i.e. there are two hypotheses (I_2, ξ_1) and (I_2, ξ_2), weights $w^{(I_2, \xi_1)} = 0.4$ and $w^{(I_2, \xi_2)} = 0.2$, densities $p^{(\xi_1)}(x, \ell_1)$, $p^{(\xi_1)}(x, \ell_2)$, $p^{(\xi_2)}(x, \ell_1)$ and $p^{(\xi_2)}(x, \ell_2)$.

Then, the δ-GLMB representation is:

$$\begin{aligned}\boldsymbol{\pi}(\mathbf{X}) = {} & \delta_{\{\ell_1\}}(\mathcal{L}(\mathbf{X}))\; \left[0.3\, p^{(\xi_1)}(x, \ell_1) + 0.1\, p^{(\xi_2)}(x, \ell_1)\right] + \\ & \delta_{\{\ell_1, \ell_2\}}(\mathcal{L}(\mathbf{X}))\; \left[0.4\, p^{(\xi_1)}(x, \ell_1)\; p^{(\xi_1)}(x, \ell_2) + 0.2\, p^{(\xi_2)}(x, \ell_1)\; p^{(\xi_2)}(x, \ell_2)\right]\end{aligned} \tag{2.89}$$

Thus (2.89) has exactly four hypotheses (I_1, ξ_1), (I_1, ξ_2), (I_2, ξ_1) and (I_2, ξ_2).

The weight $w^{(I, \xi)}$ represents the probability of hypothesis (I, ξ) and $p^{(\xi)}(x, \ell)$ is the probability density of the kinematic state of track ℓ for the association map history ξ.

Labeled Multi-Bernoulli RFS

A *labeled multi-Bernoulli* (LMB) RFS $\mathbf{X}$ with state space $\mathbb{X}$, label space $\mathbb{L}$ and (finite) parameter set $\{(r^{(\ell)}, p^{(\ell)}) : \ell \in \mathbb{L}\}$, is a multi-Bernoulli RFS on $\mathbb{X}$ augmented with labels corresponding to the successful (non-empty) Bernoulli components. In particular, $r^{(\ell)}$ is the *existence probability* and $p^{(\ell)}$ is the PDF on the state space $\mathbb{X}$ of the *Bernoulli component* $(r^{(\ell)}, p^{(\ell)})$ with unique label $\ell \in \mathbb{L}$. The procedure in Table 2.12 illustrates how a sample from a labeled multi-Bernoulli RFS is generated.

Table 2.12: Sampling a labeled multi-Bernoulli RFS

$\mathbf{X} = \varnothing$
for $\ell \in \mathbb{L}$ **do**
  Sample $u \sim \text{Uniform}_{[0,1]}$
  if $u \leq r^{(\ell)}$ **then**
    Sample $x \sim p^{(\ell)}(\cdot)$
    $\mathbf{X} = \mathbf{X} \cup \{(x, \ell)\}$
  end if
end for

The probability density of an LMB RFS is given by

$$\boldsymbol{\pi}(\mathbf{X}) = \Delta(\mathbf{X})\, w(\mathcal{L}(\mathbf{X}))\, p^{\mathbf{X}} \tag{2.90}$$

where

$$w(L) = \prod_{\ell \in L} 1_{\mathbb{L}}(\ell)\, r^{(\ell)} \prod_{\ell \in \mathbb{L} \setminus L} \left(1 - r^{(\ell)}\right) , \tag{2.91}$$

$$p^{(\ell)}(x) \triangleq p(x, \ell) . \tag{2.92}$$

For convenience, the shorthand notation $\boldsymbol{\pi} = \left\{\left(r^{(\ell)}, p^{(\ell)}\right)\right\}_{\ell \in \mathbb{L}}$ will be adopted for the density of an LMB RFS. The LMB is also a special case of the GLMB [VV13, VVP14, RVVD14] having $\mathbb{C}$ with a single element (thus the superscript is simply avoided) and

$$p^{(c)}(x, \ell) = p(x, \ell) = p^{(\ell)}(x) , \tag{2.93}$$

$$w^{(c)}(L) = w(L) . \tag{2.94}$$

2.3.8 Bayesian multi-object tracking

The labeled RFS paradigm [VV13, VVP14], along with the mathematical tools provided by FISST [Mah07b], allows to formalize in a rigorous and elegant way the multi-object Bayesian recursion for MOT.

For MOT, the object label is an ordered pair of integers $\ell = (k, i)$, where k is the *time*

of birth and $i \in \mathbb{N}$ is a unique index to distinguish objects born at the same time. The label space for objects born at time k is $\mathbb{L}_k = \{k\} \times \mathbb{N}$. An object born at time k has state $\mathbf{x} \in \mathbb{X} \times \mathbb{L}_k$. Hence, the label space for objects at time k (including those born prior to k), denoted as $\mathbb{L}_{0:k}$, is constructed recursively by $\mathbb{L}_{0:k} = \mathbb{L}_{0:k-1} \cup \mathbb{L}_k$ (note that $\mathbb{L}_{0:k-1}$ and $\mathbb{L}_k$ are disjoint). A multi-object state $\mathbf{X}$ at time k is a finite subset of $\mathbb{X} \times \mathbb{L}_{0:k}$.

Suppose that, at time k, there are N_k objects with states $\mathbf{x}_{k,1}, \ldots, \mathbf{x}_{k,N_k}$, each taking values in the (labeled) state space $\mathbb{X} \times \mathbb{L}_{0:k}$, and M_k measurements $y_{k,1}, \ldots, y_{k,M_k}$ each taking values in an observation space $\mathbb{Y}$. The *multi-object state* and *multi-object observation*, at time k, [Mah03, Mah07b] are, respectively, the finite sets

$$\mathbf{X}_k = \{\mathbf{x}_{k,1}, \ldots, \mathbf{x}_{k,N_k}\} , \tag{2.95}$$

$$Y_k = \{y_{k,1}, \ldots, y_{k,M_k}\} . \tag{2.96}$$

Let $\boldsymbol{\pi}_k(\cdot)$ denote the *multi-object filtering density* at time k, and $\boldsymbol{\pi}_{k|k-1}(\cdot)$ the *multi-object prediction density* to time k (formally $\boldsymbol{\pi}_k(\cdot)$ and $\boldsymbol{\pi}_{k|k-1}(\cdot)$ should be written respectively as $\boldsymbol{\pi}_k(\cdot|Y_0, \ldots, Y_{k-1}, Y_k)$, and $\boldsymbol{\pi}_{k|k-1}(\cdot|Y_0, \ldots, Y_{k-1})$, but for simplicity the dependence on past measurements is omitted). Then, the *multi-object Bayes recursion* propagates $\boldsymbol{\pi}_k$ in time [Mah03, Mah07b] according to the following update and prediction

$$\boldsymbol{\pi}_{k|k-1}(\mathbf{X}) = \int \boldsymbol{\varphi}_{k|k-1}(\mathbf{X}|\mathbf{Z})\, \boldsymbol{\pi}_{k-1}(\mathbf{Z})\, \delta\mathbf{Z} , \tag{2.97}$$

$$\boldsymbol{\pi}_k(\mathbf{X}) = \frac{g_k(Y_k|\mathbf{X})\, \boldsymbol{\pi}_{k|k-1}(\mathbf{X})}{\displaystyle\int g_k(Y_k|\mathbf{Z})\, \boldsymbol{\pi}_{k|k-1}(\mathbf{Z})\, \delta\mathbf{Z}} , \tag{2.98}$$

where $\boldsymbol{\varphi}_{k|k-1}(\,\cdot\,|\,\cdot\,)$ is the *labeled multi-object transition density* to time k, $g_k(\,\cdot\,|\,\cdot\,)$ is the *multi-object likelihood function* at time k, and the integral is a *set integral* defined, for any function $\boldsymbol{f} : \mathcal{F}(\mathbb{X} \times \mathbb{L}) \to \mathbb{R}$, by

$$\int \boldsymbol{f}(\mathbf{X})\, \delta\mathbf{X} = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{(\ell_1,\ldots,\ell_n)\in\mathbb{L}} \int_{\mathbf{X}^n} \boldsymbol{f}(\{(x_1,\ell_1), \ldots, (x_n,\ell_n)\})\, dx_1 \cdots dx_n . \tag{2.99}$$

The multi-object posterior density captures all information on the number of objects, and their states [Mah07b]. The multi-object likelihood function encapsulates the underlying models for detections and false alarms while the multi-object transition density embeds the underlying models of motion, birth and death [Mah07b].

Remark 9. Eqs. (2.97)-(2.98) represent the multi-object counterpart of (2.7)-(2.8).

Let us now consider a multi-sensor *centralized* setting in which a sensor network $(\mathcal{N}, \mathcal{A})$ conveys all the measurement to a central fusion node. Assuming that the measurements taken by the sensors are independent, the multi-object Bayesian filtering recursion can be

naturally extended as follows:

$$\boldsymbol{\pi}_{k|k-1}(\mathbf{X}) = \int \boldsymbol{\varphi}_{k|k-1}(\mathbf{X}|\mathbf{Z})\, \boldsymbol{\pi}_{k-1}(\mathbf{Z})\, \delta\mathbf{Z}\,, \tag{2.100}$$

$$\boldsymbol{\pi}_{k}(\mathbf{X}) = \frac{\prod\limits_{i\in\mathcal{N}} g_k^i\left(Y_k^i|\mathbf{X}\right) \boldsymbol{\pi}_{k|k-1}(\mathbf{X})}{\int \prod\limits_{i\in\mathcal{N}} g_k^i\left(Y_k^i|\mathbf{Z}\right) \boldsymbol{\pi}_{k|k-1}(\mathbf{Z})\, \delta\mathbf{Z}}\,, \tag{2.101}$$

Remark 10. Eqs. (2.100)-(2.101) represents the multi-object counterpart of (2.9)-(2.10) which has been made possible thanks to the concept of RFS densities.

In the work [VV13], the full labeled multi-object Bayesian recursion is derived for the general class of GLMB densities, while in [VVP14] the analytical implementation of the multi-object Bayesian recursion for the δ-GLMB density is provided. Moreover, in the work [RVVD14], a multi-object Bayesian recursion for the LMB density, which turns out not to be in a closed form, is discussed and presented.

For convenience, hereinafter, explicit reference to the time index k for label sets will be omitted by denoting $\mathbb{L}_- \triangleq \mathbb{L}_{0:k-1}$, $\mathbb{B} \triangleq \mathbb{L}_k$, $\mathbb{L} \triangleq \mathbb{L}_- \cup \mathbb{B}$.

2.4 Distributed information fusion

The third and fourth core concepts of this dissertation are the *Kullback-Leibler Average* (KLA) and *consensus*, which are two mathematical tools for distributing information over sensor networks.

To combine limited information from individual nodes, a suitable information fusion procedure is required to reconstruct, from the information of the various local nodes, the state of the objects present in the surrounding environment. The scalability requirement, the lack of a fusion center and knowledge on the network topology (see section 2.1) dictate the adoption of a consensus approach to achieve a collective fusion over the network by iterating local fusion steps among neighboring nodes [OSFM07, XBL05, CA09, BC14]. In addition, due to the data incest problem in the presence of network loops that causes *double counting* of information, robust (but suboptimal) fusion rules, such as the *Chernoff fusion* rule [CT12, CCM10] (that includes *Covariance Intersection* [JU97, Jul08] and its generalization [Mah00]) are required.

In this chapter, the KLA and consensus will be separately dealt with, respectively, as a robust suboptimal fusion rule and as a technique to spread information, in a scalable way, throughout sensor networks.

2.4.1 Notation

Given PDFs p, q and a scalar $\alpha > 0$, the information fusion $\oplus$ and weighting operators $\odot$ [BCF+13a, BC14, BCF14a] are defined as follows

$$(p \oplus q)(x) \triangleq \frac{p(x)\, q(x)}{\langle p, q \rangle} \tag{2.102}$$

$$(\alpha \odot p)(x) \triangleq \frac{[p(x)]^{\alpha}}{\langle p^{\alpha}, 1 \rangle}. \tag{2.103}$$

It can be checked that the fusion and weighting operators satisfy the following properties:

P.A $(p \oplus q) \oplus h = p \oplus (q \oplus h) = p \oplus q \oplus h$

P.B $p \oplus q = q \oplus p$

P.C $(\alpha\, \beta) \odot p = \alpha \odot (\beta \odot p)$

P.D $1 \odot p = p$

P.E $\alpha \odot (p \oplus q) = (\alpha \odot p) \oplus (\alpha \odot q)$

P.F $(\alpha + \beta) \odot p = (\alpha \odot p) \oplus (\beta \odot q)$

for any PDFs p and q, positive scalars α and β.

2.4.2 Kullback-Leibler average of PDFs

A key ingredient for networked estimation is the capability to fuse in a consistent way PDFs of the quantity to be estimated provided by different nodes. In this respect, a sensible information-theoretic definition of fusion among PDFs is the *Kullback-Leibler Average* (KLA) [BCMP11, BC14] relying on the *Kullback-Leibler Divergence* (KLD). Given PDFs $\{p^i(\cdot)\}_{i \in \mathcal{N}}$ and relative weights $\omega^i > 0$ such that $\sum_{i \in \mathcal{N}} \omega^i = 1$, their weighted KLA $\overline{p}(\cdot)$ is defined as

$$\overline{p}(\cdot) = \arg\inf_{p(\cdot)} \sum_{i \in \mathcal{N}} \omega^i D_{KL}\left(p \parallel p^i\right) \tag{2.104}$$

where

$$D_{KL}\left(p \parallel p^i\right) = \int p(x) \log\left(\frac{p(x)}{p^i(x)}\right) dx \tag{2.105}$$

denotes the KLD between the PDFs $p(\cdot)$ and $p^i(\cdot)$. In [BC14] it is shown that the weighted KLA in (2.104) coincides with the *normalized weighted geometric mean* (NWGM) of the PDFs, i.e.

$$\overline{p}(x) = \frac{\prod_{i \in \mathcal{N}} \left[p^i(x)\right]^{\omega^i}}{\int \prod_{i \in \mathcal{N}} \left[p^i(x)\right]^{\omega^i} dx} \triangleq \bigoplus_{i \in \mathcal{N}} \left(\omega^i \odot p^i(x)\right) \tag{2.106}$$

where the latter equality follows from the properties of the operators $\odot$ and $\oplus$. Note that in the unweighted KLA $\omega^i = 1/|\mathcal{N}|$, i.e.

$$\overline{p}(x) = \bigoplus_{i\in\mathcal{N}} \left(\frac{1}{|\mathcal{N}|} \odot p^i\right)(x) \ . \tag{2.107}$$

If all PDFs are Gaussian, i.e. $p^i(\cdot) = \mathcal{N}(\,\cdot\,; m^i, P^i)$, $\overline{p}(\cdot)$ in (2.106) turns out to be Gaussian [BC14], i.e. $\overline{p}(\cdot) = \mathcal{N}\Big(\,\cdot\,; \overline{m}, \overline{P}\Big)$. In particular, defining the *information* (inverse covariance) *matrix*

$$\Omega \triangleq P^{-1} \tag{2.108}$$

and information vector

$$q \triangleq P^{-1}m \tag{2.109}$$

associated to the Gaussian mean m and covariance P, one has

$$\overline{\Omega} = \sum_{i\in\mathcal{N}} \omega^i \Omega^i \,, \tag{2.110}$$

$$\overline{q} = \sum_{i\in\mathcal{N}} \omega^i q^i \,. \tag{2.111}$$

which corresponds to the well known *Covariance Intersection* (CI) fusion rule [JU97]. Hence the KLA of Gaussian PDFs is a Gaussian PDF whose information pair $(\overline{\Omega}, \overline{q})$ is obtained by the weighted arithmetic mean of the information pairs (Ω^i, q^i) of the averaged Gaussian PDFs.

2.4.3 Consensus algorithms

In recent years, *consensus algorithms* have emerged as a powerful tool for distributed computation over networks [OSFM07, XBL05] and have been widely used in distributed parameter/state estimation algorithms [OS07, KT08, CS10, CCSZ08, CA09, SSS09, BCMP11, FFTS10, LJ12, HSH+12, BCF+13a, BC14, OHD14, BCF14a, BCF+14b, FBC+12, BCF+13b, BCF+13c, BCF+13d, BCF+14c]. In its basic form, a consensus algorithm can be seen as a technique for distributed averaging over a network; each agent aims to compute the collective average of a given quantity by iterative regional averages, where the terms "*collective*" and "*regional*" mean "*over all network nodes*" and, respectively, "*over neighboring nodes only*". Consensus can be exploited to develop scalable and reliable distributed fusion techniques.

To this end, let us briefly introduce a prototypal average consensus problem. Let node $i \in \mathcal{N}$ be provided with an estimate $\hat{\theta}^i$ of a given quantity of interest θ. The objective is to develop an algorithm that computes in a distributed way, in each node, the average

$$\bar{\theta} = \frac{1}{|\mathcal{N}|}\sum_{i\in\mathcal{N}} \hat{\theta}^i \,. \tag{2.112}$$

To this end, let $\hat{\theta}_0^i = \hat{\theta}^i$, then a simple consensus algorithm takes the following iterative form:

$$\hat{\theta}_{l+1}^i = \sum_{j \in \mathcal{N}^i} \omega^{i,j}\, \hat{\theta}_l^j \,, \qquad \forall i \in \mathcal{N} \tag{2.113}$$

where the *consensus weights* must satisfy the conditions

$$\sum_{j \in \mathcal{N}^i} \omega^{i,j} = 1 \,, \qquad \forall i \in \mathcal{N} \tag{2.114}$$

$$\omega^{i,j} \geq 0 \,, \qquad \forall i, j \in \mathcal{N} \,. \tag{2.115}$$

Notice from (2.113)-(2.115) that at a given consensus step the estimate in any node is computed as a convex combination of the estimates of the neighbors at the previous consensus step. In other words, the iteration (2.113) is simply a regional average computed in node i, the objective of consensus being convergence of such regional averages to the collective average (2.112). Important convergence properties, depending on the consensus weights, can be found in [OSFM07, XBL05]. For instance, let us denote by Π the consensus matrix whose generic (i, j)-element coincides with the consensus weight $\omega^{i,j}$ (if $j \notin \mathcal{N}^i$ then $\omega^{i,j}$ is taken as 0). Then, if the consensus matrix Π is primitive and doubly stochastic[1], the consensus algorithm (2.119) asymptotically yields the average (2.112) in that

$$\lim_{l \to \infty} \hat{\theta}^i = \bar{\theta} \,, \qquad \forall i \in \mathcal{N} \,. \tag{2.116}$$

A necessary condition for the matrix Π to be primitive is that the graph $\mathcal{G}$ associated with the sensor network be strongly connected [CA09]. Moreover, in the case of an undirected graph $\mathcal{G}$, a possible choice ensuring convergence to the collective average is given by the so-called *Metropolis weights* [XBL05, CA09].

$$\omega^{i,j} = \frac{1}{1 + \max\{|\mathcal{N}^i| \,, |\mathcal{N}^j|\}} \,, \qquad i \in \mathcal{N} \,, j \in \mathcal{N}^i \,, i \neq j \tag{2.117}$$

$$\omega^{i,i} = 1 - \sum_{j \in \mathcal{N}^i,\, j \neq i} \omega^{i,j} \,. \tag{2.118}$$

2.4.4 Consensus on posteriors

The idea is to exploit consensus to reach a *collective agreement* (over the entire network), by having each node iteratively updating and passing its local information to neighbouring nodes. Such repeated local operations provide a mechanism for propagating information throughout the whole network. The idea is to perform, at each time instant and in each node of the network $i \in \mathcal{N}$, a local correction step to find the local posterior PDF followed by consensus on such posterior PDFs to determine the collective unweighted KLA of the posterior densities p_k^i.

[1] A non-negative square matrix Π is doubly stochastic if all its rows and columns sum up to 1. Further, it is primitive if there exists an integer m such that all the elements of Π^m are strictly positive.

Suppose that at time k, each agent i starts with the posterior p_k^i as the initial iterate $p_{k,0}^i$, and computes the l-th consensus iterate by

$$p_{k,l}^i = \bigoplus_{j \in \mathcal{N}^i} \left(\omega^{i,j} \odot p_{k,l-1}^j \right) \tag{2.119}$$

where $\omega^{i,j} \geq 0$, satisfying $\sum_{j \in \mathcal{N}^i} \omega^{i,j} = 1$, are the consensus weights relating agent i to nodes $j \in \mathcal{N}^i$. Then, using the properties of the operators $\oplus$ and $\odot$, it can be shown that [BC14]

$$p_{k,l}^i = \bigoplus_{j \in \mathcal{N}} \left(\omega_l^{i,j} \odot p_k^j \right) \tag{2.120}$$

where $\omega_l^{i,j}$ is the (i,j)-th entry of Π^n, with Π the (square) consensus matrix with (i,j)-th entry $\omega^{i,j} 1_{\mathcal{N}^i}(j)$ (it is understood that p_k^j is omitted from the fusion whenever $\omega_l^{i,j} = 0$). More importantly, it was shown in [OSFM07, XBL05] that if the consensus matrix Π is primitive, (i.e. non-negative, and there exists an integer m such that Π^m is positive) and doubly stochastic (all rows and columns sum to 1), then for any $i, j \in \mathcal{N}$, one has

$$\lim_{l \to \infty} \omega_l^{i,j} = \frac{1}{|\mathcal{N}|}. \tag{2.121}$$

In other words, at time k, if the consensus matrix is primitive then the consensus iterate of each node in the network "tends" to the collective unweighted KLA (2.107) of the posterior densities [CA09, BC14].

The *Consensus on Posteriors* (CP) approach to *Distributed SOF* (DSOF) is summarized by the algorithm of Table 2.13 to be carried out at each sampling interval k in each node $i \in \mathcal{N}$. In the linear Gaussian case, the CP algorithm involves a Kalman filter for local prediction and correction steps and CI for the consensus steps. In this case, it has been been proved [BCMP11, BC14] that, under suitable assumptions, the CP algorithm guarantees mean-squared bounded estimation error in all network nodes for any number $L \geq 1$ of consensus iterations. In the nonlinear and/or non Gaussian case, the PDFs in the various steps can be approximated as Gaussian and the corresponding means-covariances updated via, e.g., EKF [MSS62] or UKF [JU04].

Remark 11. Whenever the number of consensus steps L tends to infinity, consensus on posteriors is unable to recover the solution of the Bayesian filtering problem. In fact, in the CP approach, the novel information undergoing consensus combined with the prior information, is unavoidably underweighted.

Table 2.13: Consensus on Posteriors (CP)

procedure CP(NODE i, TIME k)

Prediction

$$p^i_{k|k-1}(x_k) = \int \varphi_{k|k-1}(x_k|x_{k-1})\, p^i_{k-1}(x_{k-1})\, dx_{k-1}$$

Local Correction

$$p^i_k(x_k|y^i_k) = \begin{cases} \left(g^i_k(y^i_k|\cdot) \oplus p^i_{k|k-1}(\cdot)\right)(x) \ , & i \in \mathcal{S} \\ p^i_{k|k-1}(x_k) \ , & i \in \mathcal{C} \end{cases}$$

Consensus

$p^i_{k,0}(x_k) = p^i_k(x_k|y^i_k)$

for $l = 1, \ldots, L$ **do**

$$p^i_{k,l}(x_k) = \bigoplus_{j \in \mathcal{N}^i} \left(\omega^{i,j} \odot p^j_{k,l-1}(x)\right)$$

end for

$p^i_k(x_k) = p^i_{k,L}(x_k)$

end procedure

3 Distributed single-object filtering

Contents

This chapter presents two applications of distributed information fusion for single-object filtering [BCF14a, FBC+12, BCF+14b]. The first application presents an improvement of the CP approach, described in subsection 2.4.4, and is applied to track a single non-maneuvering object. The second application takes into account the possibility of the object of being highly maneuvering. Thus, a multiple-model filtering approach is used to devise a consensus-based algorithm capable of tracking a single highly-maneuvering object. The effectiveness of the proposed algorithms is demonstrated via simulation experiments on realistic scenarios.

3.1 Consensus-based distributed filtering

The approach proposed in this section is based on the idea of carrying out, in parallel, a separate consensus for the novel information (likelihoods) and one for the prior information (priors). This parallel procedure is conceived as an improvement of the CP approach to avoid underweighting the novel information during the fusion steps. The outcomes of the two consensuses are then combined to provide the fused posterior density.

3.1.1 Parallel consensus on likelihoods and priors

The proposed parallel *Consensus on Likelihoods and Priors* (CLCP) approach to DSOF is summarized by the algorithm of Table 3.1 to be carried out at each sampling interval k in

each node $i \in \mathcal{N}$. Notice that in the correction step, a suitable positive weight $\rho_k^i \geq 1$ is

Table 3.1: Consensus on Likelihoods and Priors (CLCP) pseudo-code

procedure CLCP(NODE i, TIME k)

Prediction

$$p_{k|k-1}^i(x) = \int \varphi_{k|k-1}(x|\zeta)\, p_{k-1}^i(x)\, d\zeta$$

Consensus

$$g_{k,0}^i(x) = \begin{cases} g_k^i(y_k^i|x)\,, & i \in \mathcal{S} \\ 1\,, & i \in \mathcal{C} \end{cases}$$

$$p_{k|k-1,0}^i(x) = p_{k|k-1}^i(x)$$

for $l = 1, \ldots, L$ **do**

$$g_{k,l}^i(x) = \bigoplus_{j \in \mathcal{N}^i} \left(\omega^{i,j} \odot g_{k,l-1}^j(x)\right)$$

$$p_{k|k-1,l}^i(x) = \bigoplus_{j \in \mathcal{N}^i} \left(\omega^{i,j} \odot p_{k|k-1,l-1}^j(x)\right)$$

end for

Correction

$$p_k^i(x) = p_{k|k-1,L}^i(x) \oplus \left(\rho_k^i \odot g_{k,L}^i(x)\right)$$

end procedure

applied to the outcome of the consensus on likelihoods, in order to possibly counteract the underweighting of novel information. To elaborate more on this issue, observe that each local posterior PDF $p_k^i(x)$ resulting from application of the CLCP algorithm turns out to be equal to

$$p_k^i(x) = \left(\bigoplus_{j \in \mathcal{N}} \left(\omega_L^{i,j} \odot p_{k|k-1}^j(x)\right)\right) \oplus \left(\bigoplus_{j \in \mathcal{S}} \left(\rho_k^i \omega_L^{i,j} \odot g_k^i\left(y_k^i|x\right)\right)\right)\,. \tag{3.1}$$

As it can be seen, $p_k^i(x)$ is composed of two parts: I) a weighted geometric mean of the priors $p_{k|k-1}^j(x)$; II) a weighted combination of the likelihoods $g_k^i(y_k^i|x)$. The fact that in the first part the weights $\omega_L^{i,j}$ sum up to one ensures that no double counting of the common information contained in the priors can occur [Jul08, BJA12]. As for the second part, care must be taken in the choice of the scalar weights ρ_k^i. For instance, in order to avoid overweighting some of the likelihoods, it is important that, for any pair i, j, one has $\rho_k^i \omega_L^{i,j} \leq 1$. In this way, each component of independent information is replaced with a conservative approximation, thus ensuring the conservativeness of the overall fusion rule [BJA12]. On the other hand, the product $\rho_k^i \omega_L^{i,j}$ should not be too small in order to avoid excessive underweighting. Based on these considerations and recalling the consensus property (2.121), different strategies for choosing ρ_k^i can be devised.

A reasonable choice would amount to letting

$$\rho_k^i = \min_{j \in \mathcal{S}} \left(\frac{1}{\omega_L^{i,j}} \right) \tag{3.2}$$

whenever at least one of the weights $\omega_L^{i,j}$ is different from zero. Such a choice would assign to ρ_k^i the closest value to the true number of agents $|\mathcal{N}|$ of the network. This strategy is easily applicable when $L = 1$ but, unfortunately, for $L > 1$ requires that each node of the network can compute Π^L. In most settings, this is not possible and alternative solutions must be adopted.

For example, when $L \gg 1$ and the matrix Π is primitive and doubly stochastic, one has that $\omega_L^{i,j} \approx 1/|\mathcal{N}|$ for any i, j. Then, in this case, one can let

$$\rho_k^i = |\mathcal{N}| \,. \tag{3.3}$$

Such a choice has the appealing feature of giving rise to a distributed algorithm converging to the centralized one as L tends to infinity. However, when only a moderate number of consensus steps is performed, it leads to an overweighting of some likelihood components. Furthermore, the number of nodes $|\mathcal{N}|$ might be unknown, in particular for a time-varying network. Notice that when such a choice is adopted, due to the properties of the fusion $\oplus$ and weighting $\odot$ operators, the CLCP algorithm can be shown to be mathematically equivalent (in the sense that they would generate the same $p_k^i(x)$) to the optimal distributed protocol of [OC10].

An alternative solution is to exploit consensus so as to compute, in a distributed way, a normalization factor to improve the filter performance while preserving consistency of each local filter. For example, an estimate of the fraction $|\mathcal{S}|/|\mathcal{N}|$ of sensor nodes in the network can be computed via the consensus algorithm

$$b_{k,l}^i = \sum_{j \in \mathcal{N}^i} \omega^{i,j} \; b_{k,l-1}^j, \qquad l = 1, \ldots, L \tag{3.4}$$

with the initialization $b_{k,0}^i = 1$ if $i \in \mathcal{S}$, and $b_{k,0}^i = 0$ otherwise. Then, the choice

$$\rho_k^i = \begin{cases} 1 \,, & \text{if } b_{k,L}^i = 0 \\ \dfrac{1}{b_{k,L}^i} \,, & \text{otherwise} \end{cases} \tag{3.5}$$

has the desirable property of ensuring that $\rho_k^i \omega_L^{i,j} \leq 1$ for any i, j. Another positive feature of (3.5) is that no a priori knowledge of the network is required.

3.1.2 Approximate CLCP

It must be pointed out that unfortunately the CLCP recursion of Table 3.1 does not admit an exact analytical solution except for the linear Gaussian case. Henceforth it will be assumed that the noises are Gaussian, i.e. $w_k \sim \mathcal{N}(0, Q_k)$ and $v_k^i \sim \mathcal{N}(0, R_k^i)$, but that the system, i.e. $f_k(\cdot)$ and/or $h_k^i(\cdot)$, might be nonlinear. The main issue is how to deal with nonlinear sensors in the consensus on likelihoods as for the other tasks (i.e. consensus on priors, correction, prediction) it is well known how to handle nonlinearities exploiting, e.g., EKF [MSS62] or UKF [JU04] or particle filters [RAG04]. Under the Gaussian assumption, the local likelihoods take the form

$$g_k^i(x) = g_k^i\left(y_k^i|x\right) = \mathcal{N}\left(y_k^i - h_k^i(x); 0, R_k^i\right) \tag{3.6}$$

Whenever sensor i is linear, i.e. $h_k^i(x) = C_k^i x$,

$$g_k^i(x) \propto e^{-\frac{1}{2}\left(x^\top \delta\Omega_k^i x - 2x^\top \delta q_k^i\right)} \tag{3.7}$$

where

$$\begin{aligned} \delta\Omega_k^i &\triangleq \left(C_k^i\right)^\top \left(R_k^i\right)^{-1} C_k^i\,, \\ \delta q_k^i &\triangleq \left(C_k^i\right)^\top \left(R_k^i\right)^{-1} y_k^i\,, \end{aligned} \tag{3.8}$$

Then it is clear that multiplying, or exponentiating by suitable weights, the likelihoods $g_k^i(\cdot)$ is equivalent to adding, or multiplying by such weights, the corresponding $\delta\Omega_k^i$ and δq_k^i defined in (3.8). Then, in the case of linear Gaussian sensors, the consensus on likelihoods reduces to a consensus on the information pairs defined in (3.8). For nonlinear sensors, a sensible approach seems therefore to approximate the nonlinear measurement function $h_k^i(\cdot)$ by a linear affine one, i.e.

$$h_k^i(x) \cong C_k^i\left(x - \hat{x}_{k|k-1}^i\right) + h^i\left(\hat{x}_{k|k-1}\right) \tag{3.9}$$

and then replace the measurement equation (2.4) with $\overline{y}_k^i = C_k^i x_k + v_k^i$ for a suitably defined pseudo-measurement

$$\overline{y}_k^i \;=\; y_k^i - \hat{y}_{k|k-1}^i + C_k^i \hat{x}_{k|k-1}^i \tag{3.10}$$

The sensor linearization (3.9) can be carried out for example by following the EKF paradigm. As an alternative, exploiting the unscented transform and the unscented information filter [Lee08], from $\hat{x}_{k|k-1}^i$ and the relative covariance $P_{k|k-1}^i$, one can get the σ-points $\hat{x}_{k|k-1}^{i,j}$ and relative weights ω_j for $j = 0, 1, \ldots, 2n$ ($n = \dim(x)$), and thus compute

$$\hat{y}_{k|k-1}^i = \sum_{j=0}^{2n} \omega_j\ \hat{y}_{k|k-1}^{i,j},\ \hat{y}_{k|k-1}^{i,j} = h_k^i\left(\hat{x}_{k|k-1}^{i,j}\right) \tag{3.11}$$

$$P_{k|k-1}^{yx,i} = \sum_{j=0}^{2n} \omega_j\ \left(\hat{y}_{k|k-1}^i - \hat{y}_{k|k-1}^{i,j}\right)\left(\hat{x}_{k|k-1}^i - \hat{x}_{k|k-1}^{i,j}\right)^\top \tag{3.12}$$

$$C_k^i = P_{k|k-1}^{yx,i}\ \left(P_{k|k-1}^i\right)^{-1} \tag{3.13}$$

Summarizing the above derivations, the proposed consensus-based distributed nonlinear filter is detailed in Table 3.2.

The reason for performing in parallel consensus on priors and likelihoods is the counteraction of underweighting of the novel information contained in the likelihoods by using the weighting factor ρ_k^i. Clearly, such a choice requires transmitting, at each consensus step, approximately twice the number of floating-point data with respect to other algorithms like CP. Notice that, depending on the packet size, this need not necessarily increase the packet rate.

Table 3.2: Analytical implementation of Consensus on Likelihoods and Priors (CLCP)

procedure CLCP(NODE i, TIME t)

Consensus

If $i \in \mathcal{S}$: $\begin{cases} \delta\Omega_{k,0}^i = \left(C_k^i\right)^\top \left(R_k^i\right)^{-1} C_k^i \\ \delta q_{k,0}^i = \left(C_k^i\right)^\top \left(R_k^i\right)^{-1} \overline{y}_k^i \end{cases}$

If $i \in \mathcal{C}$: $\begin{cases} \delta\Omega_{k,0}^i = 0 \\ \delta q_{k,0}^i = 0 \end{cases}$

$\Omega_{k|k-1,0}^i = \left(P_{k|k-1}^i\right)^{-1}$

$q_{k|k-1,0}^i = \left(P_{k|k-1}^i\right)^{-1} \hat{x}_{k|k-1}^i$

for $l = 1, \ldots, L$ **do**

Likelihood

$\delta\Omega_{k,l}^i = \sum_{j \in \mathcal{N}^i} \omega^{i,j} \delta\Omega_{k,l-1}^j$

$\delta q_{k,l}^i = \sum_{j \in \mathcal{N}^i} \omega^{i,j} \delta q_{k,l-1}^j$

Prior

$\Omega_{k|k-1,l}^i = \sum_{j \in \mathcal{N}^i} \omega^{i,j} \Omega_{k|k-1,l-1}^j$

$q_{k|k-1,l}^i = \sum_{j \in \mathcal{N}^i} \omega^{i,j} q_{k|k-1,l-1}^j$

end for

Correction

$\Omega_k^i = \Omega_{k|k-1,L}^i \; + \; \rho_k^i \; \delta\Omega_{k,L}^i$

$q_k^i = q_{k|k-1,L}^i \; + \; \rho_k^i \; \delta q_{k,L}^i$

Prediction

$P_k^i = \left(\Omega_k^i\right)^{-1}$, $\hat{x}_k^i = \left(\Omega_k^i\right)^{-1} q_k^i$

from $\hat{x}_k^i, P_k^i$ **compute** $\hat{x}_{k+1|k}^i, P_{k+1|k}^i$ **via UKF**

end procedure

3.1.3 A tracking case-study

To evaluate performance of the proposed approach, the networked object tracking scenario of fig. 3.1 has been simulated. The network consists of 50 COM nodes and 10 SEN measuring the

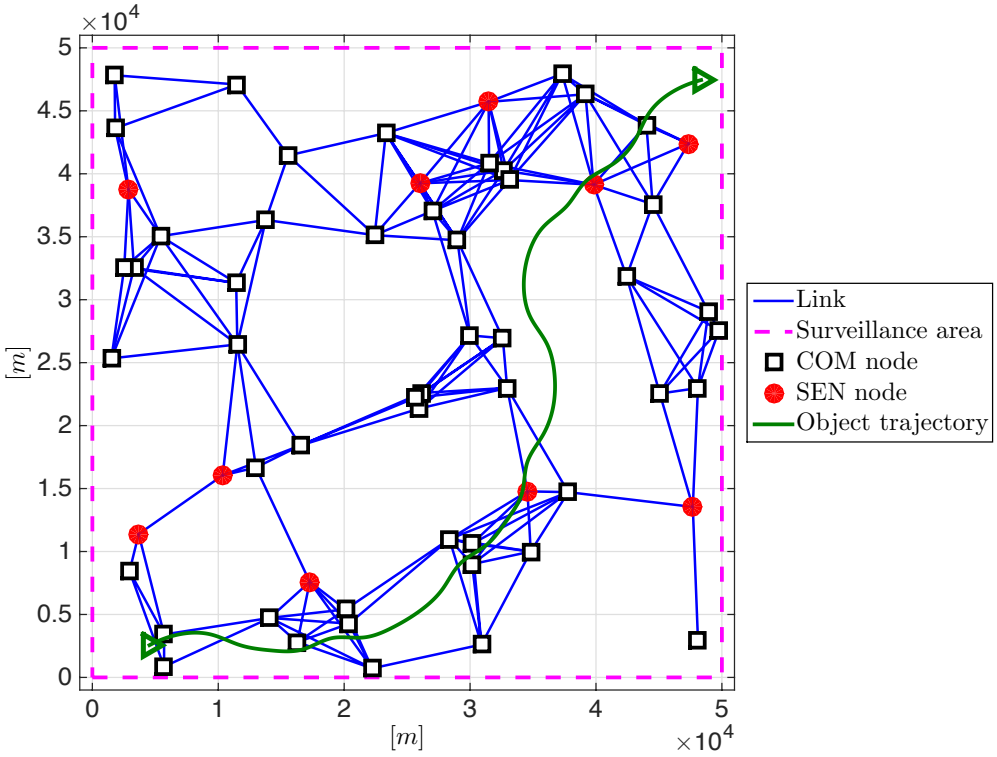


Figure 3.1: Network and object trajectory.

Time Of Arrival (TOA), i.e. the object-to-sensor distance. A linear white-noise acceleration model [BSLK01, p. 270], with 4-dimensional state consisting of positions and velocities along the coordinate axes, is assumed for the object motion; with this choice of the state, sensors are clearly nonlinear. Standard deviation of the measurement noise has been set to $100\,[m]$, the sampling interval to $5\,[s]$, the duration of the simulation to $1500\,[s]$, i.e. 300 samples. In the above scenario, 1000 Monte Carlo trials with independent measurement noise realizations have been performed. The *Position Root Mean Square Error* (PRMSE) averaged over time, Monte Carlo trials and nodes is reported in Table 3.3 for different values of L (number of consensus steps) and different consensus-based distributed nonlinear filters. In all simulations, Metropolis weights [CA09, XBL05] have been adopted for consensus. For each filter and each L two values are reported: the top one refers to the PRMSE computed over whole the simulation horizon, whereas the bottom one refers to the PRMSE computed after the transient period. Three approaches are compared in the UKF variants: the CLCP, CP (Consensus on Posteriors) [BCMP11, BC14] and CL (Consensus on Likelihoods) wherein consensus is carried out on the novel information only[1]. For the CLCP, three different

[1] In order to have a fair comparison, the CL algorithm is implemented as in Table 3.2 with the difference that no consensus on the priors is performed.

choices for the weights ρ_k^i are considered. From Table 3.3, it can be seen that the weight choice $\rho_k^i = |\mathcal{N}|$ provides the best performance only when L is sufficiently large, while for few consensus steps the other choices are preferable. It is worth pointing out that the CL approach, which becomes optimal as $L \to \infty$, guarantees a bounded estimation error only when the number L of consensus steps is sufficiently high (in Table 3.3 "$-$" indicates a divergent PRMSE).

Table 3.3: Performance comparison

	CP	CL	CLCP		
Choice of ρ_k^i		$\|\mathcal{N}\|$	(3.5)	$\|\mathcal{N}\|$	$\min_{j\in\mathcal{S}}(1/\omega_L^{i,j})$
$L=1$	1733	$-$	1909	$-$	1645
	513	$-$	407	$-$	332
$L=2$	1299	$-$	1182	1621	821
	423	$-$	257	545	226
$L=3$	1041	$-$	491	736	480
	389	$-$	231	129	198
$L=4$	936	4269	391	455	355
	370	3238	219	118	182
$L=5$	868	977	341	283	299
	358	134	210	111	170
$L=15$	681	222	230	138	173
	318	94	182	85	122

3.2 Consensus-based distributed multiple-model filtering

The section addresses distributed state estimation of jump Markovian systems and its application to tracking of a maneuvering object by means of a network of heterogeneous sensors and communication nodes. It is well known that a single-model Kalman-like filter is ineffective for tracking a highly maneuvering object and that *Multiple Model* (MM) filters [AF70, BBS88, BSLK01, LJ05] are by far superior for this purpose. The contributions provide novel consensus MM filters to be used for tracking maneuvering objects with sensor networks. Two novel consensus-based MM filters are presented. Simulation experiments in a tracking case-study, involving a strongly maneuvering object and a sensor network characterized by weak connectivity, demonstrate the superiority of the proposed filters with respect to existing solutions.

3.2.1 Notation

The following new notations are adopted throughout the next sections. $\mathcal{P}_c$ and $\mathcal{P}_d$ denote the sets of PDFs over the *continuous* state space $\mathbb{R}^n$ and, respectively, of PMFs over the *discrete* state space $\mathcal{R}$, i.e.

$$\mathcal{P}_c = \left\{ p(\cdot) : \mathbb{R}^n \to \mathbb{R} \;\middle|\; \int_{\mathbb{R}^n} p(x)\,dx = 1 \text{ and } p(x) \geq 0, \forall x \in \mathbb{R}^n \right\}, \tag{3.14}$$

$$\mathcal{P}_d = \left\{ \mu = \mathrm{col}\left(\mu^j\right)_{j\in\mathcal{R}} \in \mathbb{R}^{|\mathcal{R}|} \;\middle|\; \sum_{j\in\mathcal{R}} \mu^j = 1 \text{ and } \mu^j \geq 0, \forall j \in \mathcal{R} \right\}. \tag{3.15}$$

3.2.2 Bayesian multiple-model filtering

Let us now focus on state estimation for the *jump Markovian system*

$$x_k = f(m_k, x_{k-1}) + w_{k-1} \tag{3.16}$$

$$y_k = h(m_k, x_k) + v_k \tag{3.17}$$

where $m_k \in \mathcal{R} \triangleq \{1, 2, \ldots, r\}$ denotes the system *mode* or discrete state; w_k is the process noise with PDF $p_w(m_k, \cdot)$; v_k is the measurement noise, independent of w_k, with PDF $p_v(m_k, \cdot)$. It is assumed that the system can operate in r possible modes, each mode $j \in \mathcal{R}$ being characterized by a mode-matched model with state-transition function $f(j, \cdot)$ and process noise PDF $p_w(j, \cdot)$, measurement function $h(j, \cdot)$ and measurement noise PDF $p_v(j, \cdot)$. Further, mode transitions are modelled by means of a homogeneous Markov chain with suitable transition probabilities

$$p_{jt} \triangleq \mathrm{prob}(m_k = j | m_{k-1} = t), \quad j, t \in \mathcal{R} \tag{3.18}$$

whose possible time-dependence is omitted for notational simplicity. It is well known that a jump Markovian system (3.16)-(3.18) can effectively model the motion of a maneuvering object [AF70, BSLK01, section 11.6]. For instance a *Nearly-Constant Velocity* (NCV) model can be used to describe straight-line object motion while *Coordinated Turn* (CT) models with different angular speeds can describe object maneuvers. In alternative, models with different process noise covariances (small for straight-line motion and larger for object maneuvers) can be used with the same kinematic transition function $f(\cdot)$. For multimodal systems, the classical single-model filtering approach is clearly inadequate. To achieve better state estimation performance, the MM filtering approach [BSLK01, section 11.6] can be adopted. MM filters can provide, in principle, the Bayes-optimal solution for the jump Markov system state estimation problem by running in parallel r mode-matched Bayesian filters (one for each mode) on the same input measurements. Each Bayesian filter provides, at each time k, the conditional PDFs $p^j_{k|k-1}(\cdot)$ and $p^j_k(\cdot)$ of the state vector x_k given the observations $y^k \triangleq \{y_1, \ldots, y_k\}$ and the mode hypothesis $m_k = j$. Further, for each hypothesis $m_k = j$, the

conditional modal probability $\mu^j_{k|\tau} \triangleq \text{prob}(m_k = j|y^\tau)$ is also updated. In summary, Bayesian MM filtering amounts to representing and propagating in time information of the state vector of the jump Markovian system in terms of the r mode-matched PDFs $p^j_\tau(\cdot)\,, p^j_{k|\tau}(\cdot) \in \mathcal{P}_c$, $j \in \mathcal{R}$, along with the PMF $\mu_{k|\tau} = \text{col}\Big(\mu^j_{k|\tau}\Big)_{j\in\mathcal{R}} \in \mathcal{P}_d$. Then, from such distributions the overall state conditional PDF is obtained by means of the PDF mixture

$$p_{k|k-1}(\cdot) = \sum_{j\in\mathcal{R}} \mu^j_{k|k-1}\, p^j_{k|k-1}(\cdot) \;=\; \mu^\top_{k|k-1}\, p_{k|k-1}(\cdot)\;, \tag{3.19}$$

$$p_k(\cdot) = \sum_{j\in\mathcal{R}} \mu^j_{k|k}\, p^j_k(\cdot) \;=\; \mu^\top_{k|k}\, p_k(\cdot)\;, \tag{3.20}$$

where $p_{k|k-1}(\cdot) \triangleq \text{col}\Big(p^j_{k|k-1}(\cdot)\Big)_{j\in\mathcal{R}}$ and $p_k(\cdot) \triangleq \text{col}\Big(p^j_k(\cdot)\Big)_{j\in\mathcal{R}}$.
The Bayesian MM filter is summarized in Table 3.4.

It is worth to highlight, however, that the filter in Table 3.4 is not practically implementable even in the simplest linear Gaussian case, i.e. when I) the functions $f(j,\cdot)$ and $h(j,\cdot)$ are linear for all $j \in \mathcal{R}$; and II) the PDFs $p^j_0(\cdot)$, $p_w(j,\cdot)$ and $p_v(j,\cdot)$ are Gaussian for all $j \in \mathcal{R}$. In fact, both the mode fusion and the mixing steps (see Table 3.4) produce, just at the first time instant $k = 1$, a Gaussian mixture of r components. When time advances, the number of Gaussian components grows exponentially with time k as r^k. To avoid this exponential growth, the two most commonly used MM filter algorithms, i.e. *First Order Generalized Pseudo-Bayesian* (GPB_1) and *Interacting Multiple Model* (IMM), adopt a different strategy to keep all mode-matched PDFs Gaussian throughout the recursions. In particular, the GPB_1 algorithm (see Table 3.5) performs at each time k, before the prediction step, a re-initialization of all mode-matched PDFs with a single Gaussian component having the same mean and covariance of the fused Gaussian mixture. Conversely, the IMM algorithm (see Table 3.6) carries out a mixing procedure to suitably re-initialize each mode-matched PDF with a different Gaussian component. In the linear Gaussian case, the correction and prediction steps of MM filters are carried out, in an exact way, by a bank of mode-matched Kalman filters. Further, the likelihoods g^j_k needed to correct the modal probabilities are evaluated as

$$g^j_k = \frac{1}{\sqrt{\det\Big(2\omega S^j_k\Big)}}\, e^{-\frac{1}{2}\left(e^j_k\right)^\top \left(S^j_k\right)^{-1} e^j_k} \tag{3.21}$$

where e^j_k and S^j_k are the innovation and relative covariance provided by the KF matched to mode j. Whenever the functions $f(j,\cdot)$ and/or $h(j,\cdot)$ are nonlinear, even correction and/or prediction destroy the Gaussian form of the mode-matched PDFs. To preserve Gaussianity, as it is common practice in nonlinear filtering, the posterior PDF can be approximated as Gaussian by making use either of linearization of the system model around the current estimate [MSS62], i.e. EKF, or of the unscented transform [JU04], i.e. UKF.

3.2.3 Networked multiple model estimation via consensus

A fundamental issue for networked estimation is to consistently fuse PDFs of the continuous quantity, or PMFs of the discrete quantity, to be estimated coming from different nodes. A suitable information-theoretic fusion among continuous probability distributions has been introduced in subsection 2.4.2, which can be straightforwardly applied in this context.

Let us now turn the attention to the case of discrete variables. Given PMFs $\mu, \nu \in \mathcal{P}_d$ their KLD is defined as

$$D_{KL}\left(\mu \parallel \nu\right) \triangleq \sum_{j \in \mathcal{R}} \mu^j \, \log\left(\frac{\mu^j}{\nu^j}\right) \tag{3.22}$$

According to (2.104), the weighted KLA of the PMFs $\mu^i \in \mathcal{P}_c$, $i \in \mathcal{N}$, is defined as

$$\overline{\mu} \; = \; \arg \inf_{\mu \in \mathcal{P}_d} \sum_{i \in \mathcal{N}} \omega^i \, D_{KL}\left(\mu \parallel \mu^i\right). \tag{3.23}$$

The following result holds.

Theorem 1 (KLA of PMFs)**.**
The weighted KLA in (3.23) is given by

$$\overline{\mu} = \mathrm{col}\Big(\overline{\mu}^j\Big)_{j \in \mathcal{R}}, \qquad \overline{\mu}^j = \frac{\prod\limits_{i \in \mathcal{N}} \left(\mu^{i,j}\right)^{\omega^i}}{\sum\limits_{t \in \mathcal{R}} \prod\limits_{i \in \mathcal{N}} \left(\mu^{i,t}\right)^{\omega^i}} \tag{3.24}$$

Analogously to the case of continuous probability distributions, fusion and weighting operators can be introduced for PMFs $\mu, \nu \in \mathcal{P}_d$ and $\omega > 0$, as follows:

$$\begin{aligned} &\mu \oplus \nu = \eta = \mathrm{col}\big(\eta^j\big)_{j \in \mathcal{R}} \quad \text{with} \quad \eta^j = \frac{\mu^j \nu^j}{\sum\limits_{t \in \mathcal{R}} \mu^t \nu^t} \triangleq \mu^j \oplus \nu^j, \quad \forall j \in \mathcal{R} \\ &\omega \odot \mu = \eta = \mathrm{col}\big(\eta^j\big)_{j \in \mathcal{R}} \quad \text{with} \quad \eta^j = \frac{\left(\mu^j\right)^{\omega}}{\sum\limits_{t \in \mathcal{R}} \left(\mu^t\right)^{\omega}} \triangleq \omega \odot \mu^j, \quad \forall j \in \mathcal{R} \end{aligned} \tag{3.25}$$

Thanks to the properties P.A-P.F of the operators $\oplus$ and $\odot$ (see subsection 2.4.1), the KLA PMF in (3.23) can be expressed as

$$\overline{\mu} \; = \; \bigoplus_{i \in \mathcal{N}} \left(\omega^i \odot \mu^i\right) \tag{3.26}$$

The collective fusion of PMFs (3.26) can be computed, in a distributed and scalable fashion, via consensus iterations

$$\mu_l^i \; = \; \bigoplus_{j \in \mathcal{N}^i} \left(\omega^{i,j} \odot \mu_{l-1}^j\right) \tag{3.27}$$

initialized from $\mu_0^i = \mu^i$ and with consensus weights satisfying $\omega^{i,j} \geq 0$ and $\sum_{j \in \mathcal{N}^i} \; \omega^{i,j} = 1$.

3.2.4 Distributed multiple-model algorithms

Let us now address the problem of interest, i.e. DSOF for the jump Markovian system (3.16)-(3.17) over a sensor network $(\mathcal{N}, \mathcal{A})$ of the type modeled in section 2.1. In this setting, measurements in (3.17) are provided by sensor nodes, i.e.

$$y_k = \text{col}\Big(y_k^i\Big)_{i\in\mathcal{S}} , \tag{3.28}$$

$$h(\cdot,\cdot) = \text{col}\Big(h^i(\cdot,\cdot)\Big)_{i\in\mathcal{S}} , \tag{3.29}$$

$$v_k = \text{col}\Big(v_k^i\Big)_{i\in\mathcal{S}} , \tag{3.30}$$

where the sensor measurement noise v_k^i, $i \in \mathcal{S}$, is characterized by the PDF $p_{v^i}(\cdot,\cdot)$. The main difficulty, in the distributed context, arises from the possible lack of complete observability from an individual node (e.g. a communication node or a sensor node measuring only angle or range or Doppler-frequency shift of a moving object) which makes impossible for such a node to reliably estimate the system mode as well as the continuous state on the sole grounds of local information. To get around this problem, consensus on both the discrete state PMF and either the mode-matched or the fused continuous state PDFs can be exploited in order to spread information throughout the network and thus guarantee observability in each node, provided that the network is strongly connected (i.e., for any pair of nodes i and j there exists a path from i to j and viceversa) and the jump Markovian system is collectively observable (i.e. observable from the whole set $\mathcal{S}$ of sensors).

Let us assume that at sampling time k, before processing the new measurements y_k, each node $i \in \mathcal{N}$ be provided with the prior mode-matched PDFs $p^i_{k|k-1}(\cdot) = col\left(p^{i,j}_{k|k-1}(\cdot)\right)_{j\in\mathcal{R}}$ along with the mode PMF $\mu^i_{k|k-1} = col\left(\mu^{i,j}_{k|k-1}\right)_{j\in\mathcal{R}}$, which are the outcomes of the previous local and consensus computations up to time $k-1$. Then, sensor nodes $i \in \mathcal{S}$ correct both mode-matched PDFs and mode PMF with the current local measurement y_k^i while communication nodes leave them unchanged; the PDFs and PMF obtained in this way are called local posteriors. At this point, consensus is needed in each node $i \in \mathcal{N}$ to exchange local posteriors with the neighbours and regionally average them over the subnetwork $\mathcal{N}^i$ of in-neighbours. The more consensus iterations are performed, the faster will be convergence of the regional KLA to the collective KLA at the price of higher communication cost and, consequently, higher energy consumption and lower network lifetime. Two possible consensus strategies can be applied:

Consensus on the Fused PDF - first carry out consensus on the mode PMFs, then fuse the local mode-matched posterior PDFs over $\mathcal{R}$ using the modal probabilities resulting from consensus, and finally carry out consensus on the fused PDF;

Consensus on Mode-Matched PDFs - carry out in parallel consensus on the mode PMFs and on the mode-matched PDFs and then perform the mode fusion using the modal probabilities and mode-matched PDFs resulting from consensus.

Notice that the first approach is cheaper in terms of communication as the nodes need to exchange just a single fused PDF, instead of multiple ($r = |\mathcal{R}|$) mode-matched PDFs. Further, it seems the most reasonable choice to be used in combination with the GPB_1 approach, as only the fused PDF is needed in the re-initialization step. Conversely, the latter approach is mandatory for the distributed IMM filter. Recall, in fact, that the IMM filter re-initializes each mode-matched PDF with a suitable mixture of such PDFs, totally disregarding the fused PDF. Hence, to spread information about the continuous state through the network, it is necessary to apply consensus to all mode-matched PDFs.

Summing up, two novel distributed MM filters are proposed: the *Distributed GPB*$_1$ (DGPB_1) algorithm of Table 3.7 which adopts the *Consensus on the Fused PDF* approach to limit data communication costs, and the *Distributed IMM* (DIMM) algorithm of Table 3.8 which, conversely, adopts the *Consensus on Mode-Matched PDFs* approach. Since both DGPB_1 and DIMM filters, like their centralized counterparts, propagate Gaussian PDFs completely characterized by either the estimate-covariance pair $(\hat{x}, P)$ or the information pair $(q = P^{-1}\hat{x}, \Omega = P^{-1})$, the consensus iteration is simply carried as a weighted arithmetic average of the information pairs associated to such PDFs, see (2.110)-(2.111).

3.2.5 Connection with existing approach

It is worth to point out the differences of the proposed DGPB_1 and DIMM algorithms of Tables 3.7 and, respectively, 3.8 with respect to the distributed IMM algorithm of reference [LJ12]. In the latter, each node carries out consensus on the newly acquired information and then updates the local prior with the outcome of such a consensus. Conversely, in the DGPB_1 and DIMM algorithms, each node first updates its local prior with the local new information, and then consensus on the resulting local posteriors is carried out. To be more specific, recall first that, following the Gaussian approximation paradigm and representing Gaussian PDFs with information pairs, the local correction for each node i and mode j, takes the form

$$\Omega_k^{i,j} = \Omega_{k|k-1}^{i,j} + \delta\Omega_k^{i,j}\,, \tag{3.31}$$

$$q_k^{i,j} = q_{k|k-1}^{i,j} + \delta q_k^{i,j}\,, \tag{3.32}$$

where $\delta\Omega_k^{i,j}, \delta q_k^{i,j}$ denote the innovation terms due to the new measurements y_k^i. In particular, for a linear Gaussian sensor model characterized by measurement function $h^i(j,x) = C^{i,j}x$ and measurement noise PDF $p_{v^i}(j,\cdot) = \mathcal{N}(\cdot; 0, R^{i,j})$, the innovation terms in (3.31)-(3.32) are given by

$$\delta\Omega_k^{i,j} = \left(C^{i,j}\right)^\top \left(R^{i,j}\right)^{-1} C^{i,j}\,, \tag{3.33}$$

$$\delta q_k^{i,j} = \left(C^{i,j}\right)^\top \left(R^{i,j}\right)^{-1} y_k^i\,. \tag{3.34}$$

Further, for a nonlinear sensor i, approximate innovation terms $\left(\delta\Omega_k^{i,j}, q_k^{i,j}\right)$ can be obtained making use either of the EKF or of the unscented transform as shown in [LJ12].

Now, the following observations can be made.

- The algorithm in [LJ12], which will be indicated below with the acronym DIMM-CL (*DIMM with Consensus on Likelihoods*), applies consensus to such innovation terms scaled by the number of nodes, i.e. to $|\mathcal{N}|\,\delta\Omega_k^{i,j}$ and $|\mathcal{N}|\,\delta q_{k|k-1}^{i,j}$. Conversely, the DIMM algorithm of Table 3.8 first performs the correction (3.31)-(3.32) and then applies consensus to the local posteriors $\left(\Omega_k^{i,j}, q_k^{i,j}\right)$, and similarly does the GPB_1 algorithm of Table 3.7 with the fused information pairs (Ω_k^i, q_k^i).

- A further difference between DIMM and DIMM-CL concerns the update of the discrete state PMF: the former applies consensus to the posterior mode probabilities $\mu_k^{i,j}$, while the latter applies it to the mode likelihoods $g_k^{i,j}$.

A discussion on the relative merits/demerits of the consensus approaches adopted for DIMM and, respectively, DIMM-CL is in order. The main positive feature of DIMM-CL is that it converges to the centralized IMM as $L \to \infty$. However, it requires a minimum number of consensus iterations per sampling interval in order to avoid divergence of the estimation error. In fact, by looking at the outcome of CL

$$\delta\Omega_{k,L}^{i,j} = \sum_{\imath \in \mathcal{N}} \omega_L^{i,\imath}\, \delta\Omega_k^{\imath,j}\,,$$

it can be seen that a certain number of consensus iterations is needed so that the local information has time to spread through the network and the system becomes observable or, at least detectable, from the subset of sensors $\mathcal{S}_L^i = \left\{ j \in \mathcal{S} : \omega_L^{i,j} \neq 0 \right\}$.

The proposed approach does not suffer from such limitations thanks to the fact that the whole posterior PDFs are combined. Indeed, as shown in [BCMP11, BC14], the single-model consensus Kalman filter based on CP ensures stability, for any $L \geq 1$ under the assumptions of system observability from the whole network and strong network connectivity, in that it provides a mean-square bounded estimation error in each node of the network. Although the stability results in [BCMP11, BC14] are not proved for jump Markov and/or nonlinear systems, the superiority of DIMM over DIMM-CL for tracking a maneuvering object when only a limited number of consensus iterations is performed is confirmed by simulation experiments, as it will be seen in the next section. With this respect, since the number of data transmissions, which primarily affect energy consumption, is clearly proportional to L, in many situations it is preferable for energy efficiency to perform only a few consensus steps, possibly a single one. A further advantage of CP over CL is that the former does not require any prior knowledge of the network topology as well as on the number of nodes in order to work properly. As a by-product, the CI approach can also better cope with time-varying networks where, due to node join/leave and link disconnections, the network graph is changing in time. On the other hand, the proposed approach does not approach the centralized filter as $L \to \infty$ since the adopted fusion rule follows a cautious strategy so as to achieve robustness with respect to data incest [BC14].

3.2.6 A tracking case-study

To assess performance of the proposed distributed multiple-model algorithms described in section 3.2.4, the 2D tracking scenario of Fig. 3.2 is considered. From Fig. 3.2 notice that the object is highly maneuvering with speeds ranging in $[0, 300][m/s]$ and acceleration magnitudes in $[0, 1.5]$ g, g being the gravity acceleration. The object state is denoted by $x = [x,\ \dot{x},\ y,\ \dot{y}]^\top$ where $(x,\ y)$ and $(\dot{x},\ \dot{y})$ represent the object Cartesian position and, respectively, velocity components. The sampling interval is $T_s = 5[s]$ and the total object navigation time is $640[s]$, corresponding to 129 sampling intervals. Specifically, $r = 5$ different *Coordinated-Turn* (CT) models [BSLK01, FS85, FS86] are used in the MM algorithms. All models have the state dynamics

$$x_{k+1} = \begin{bmatrix} 1 & \frac{\sin(\omega T_s)}{\omega} & 0 & -\frac{1-\cos(\omega T_s)}{\omega} \\ 0 & \cos(\omega T_s) & 0 & -\sin(\omega T_s) \\ 0 & \frac{1-\cos(\omega T_s)}{\omega} & 1 & \frac{\sin(\omega T_s)}{\omega} \\ 0 & \sin(\omega T_s) & 0 & \cos(\omega T_s) \end{bmatrix} x_k + w_k\,, \tag{3.35}$$

$$\mathbf{Q} = \mathrm{var}(w_k) = \begin{bmatrix} \frac{1}{4}T_s^4 & \frac{1}{2}T_s^3 & 0 & 0 \\ \frac{1}{2}T_s^3 & T_s^2 & 0 & 0 \\ 0 & 0 & \frac{1}{4}T_s^4 & \frac{1}{2}T_s^3 \\ 0 & 0 & \frac{1}{2}T_s^3 & T_s^2 \end{bmatrix} \sigma_w^2 \tag{3.36}$$

for five different constant angular speeds $\omega \in \{-1, -0.5, 0, 0.5, 1\}$ $[^\circ/s]$. Notice, in particular, that, taking the limit for $\omega \to 0$, the model corresponding to $\omega = 0$ is nothing but the well known *Discrete White Noise Acceleration* (DWNA) model [FS85, BSLK01], The standard deviation of the process noise is taken as $\sigma_w = 0.1\,[m/s^2]$ for the DWNA ($\omega = 0$) model and $\sigma_w = 0.5\,[m/s^2]$ for the other models. Jump probabilities (3.18), for the Markov chain, are chosen as follows:

$$[p_{jk}]_{j,k=1,\ldots,5} = \begin{bmatrix} 0.95 & 0.05 & 0 & 0 & 0 \\ 0.05 & 0.9 & 0.05 & 0 & 0 \\ 0 & 0.05 & 0.9 & 0.05 & 0 \\ 0 & 0 & 0.05 & 0.9 & 0.05 \\ 0 & 0 & 0 & 0.05 & 0.95 \end{bmatrix} \tag{3.37}$$

All the local mode-matched filters exploit the UKF [JU04]. All simulation results have been obtained by averaging over 300 Monte Carlo trials. No prior information on the initial object position is assumed, i.e. in all filters the object state is initialized in the cen-

ter of the surveillance region with zero velocity and associated covariance matrix equal to $\text{diag}\{10^8, 10^4, 10^8, 10^4\}$.

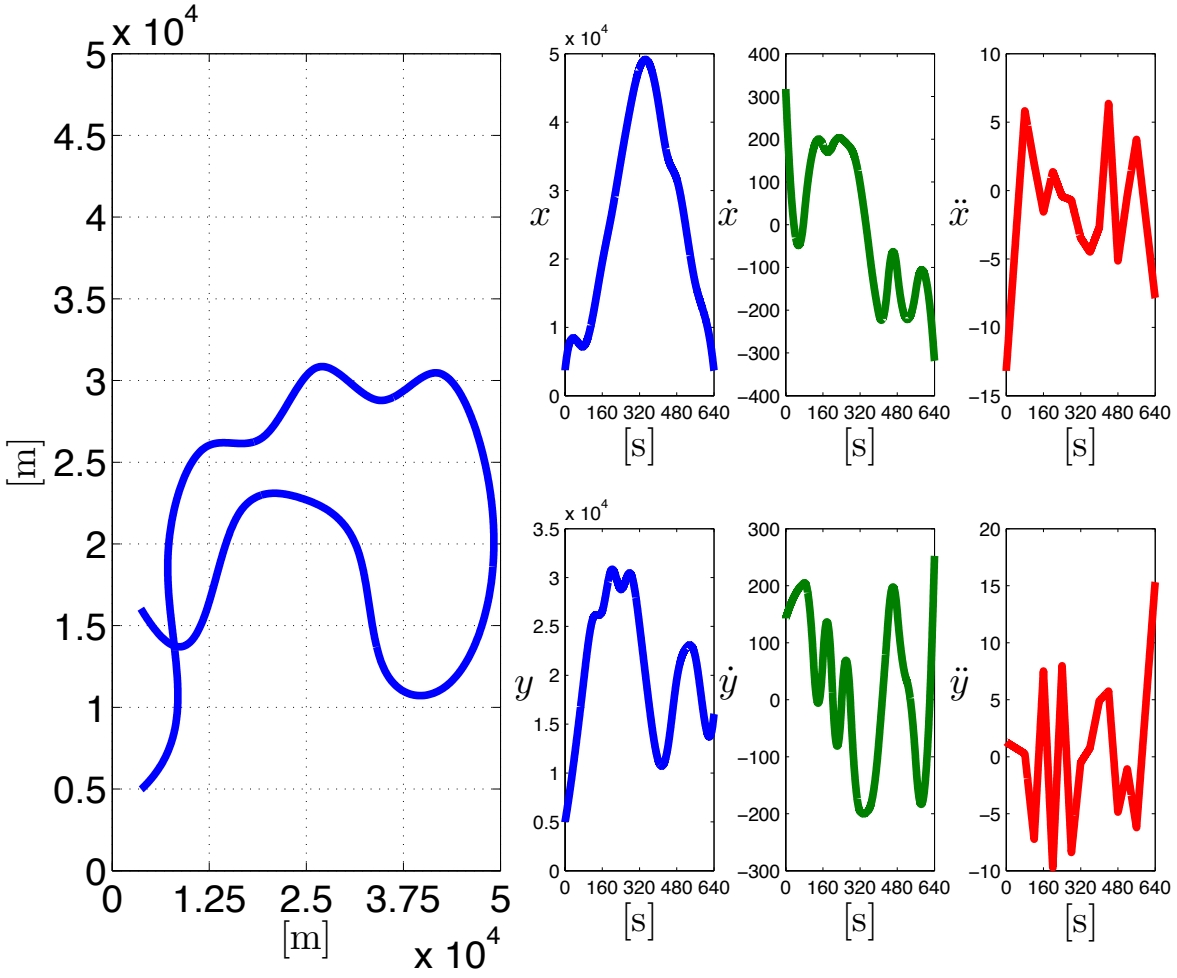


Figure 3.2: Object trajectory considered in the simulation experiments

A surveillance area of $50 \times 50\,[km^2]$ is considered, wherein 4 communication nodes, 4 bearing-only sensors measuring the object *Direction of Arrival* (DOA) and 4 TOA are deployed as shown in Fig. 3.3. Notice that the presence of communication nodes improves energy efficiency in that it allows to reduce average distance among nodes and, hence, the required transmission power of each individual node, but also hampers information diffusion across the network. This type of network represents, therefore, a valid benchmark to test the effectiveness of consensus state estimators. The following measurement functions characterize the DOA and TOA sensors:

$$h^i(x) = \begin{cases} \angle[(x - x^i) + j\,(y - y^i)]\,, & \text{if } i \text{ is a DOA sensor} \\ \sqrt{(x - x^i)^2 + (y - y^i)^2}\,, & \text{if } i \text{ is a TOA sensor} \end{cases} \tag{3.38}$$

where $(x^i,\ y^i)$ represents the known position of sensor i in Cartesian coordinates. The

standard deviation of DOA and TOA measurement noises are taken respectively as $\sigma_{DOA} = 1\,[^\circ]$ and $\sigma_{TOA} = 100\,[m]$. The number of consensus steps used in the simulations ranges from $L = 1$ to $L = 5$, reflecting a special attention towards energy efficiency.

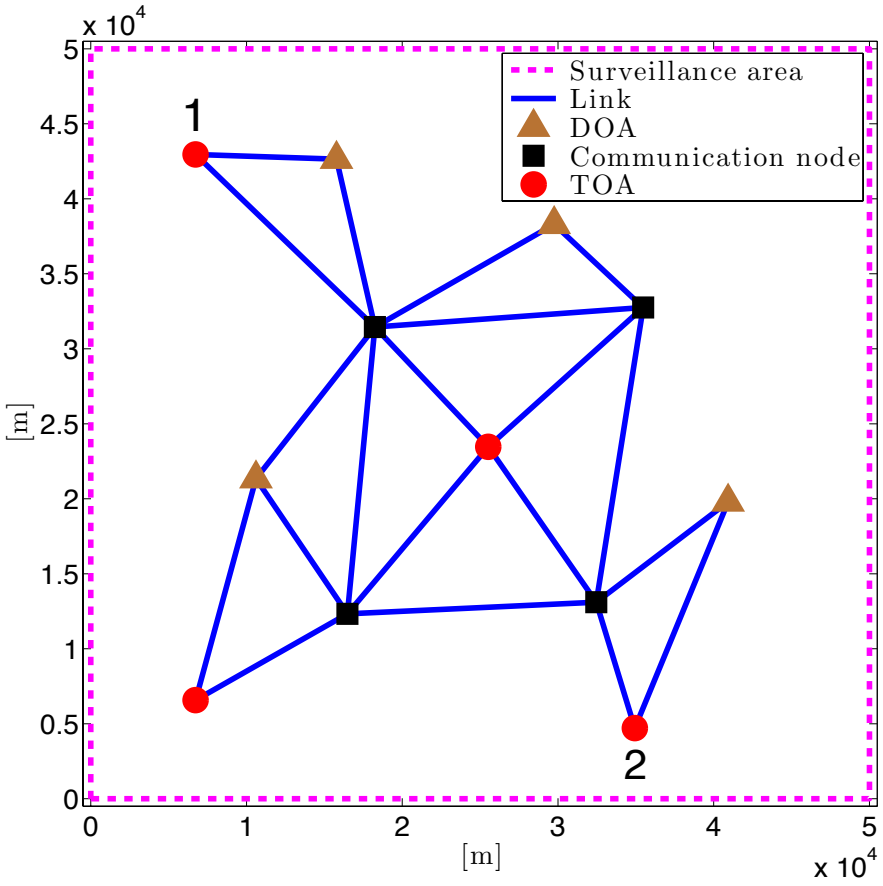


Figure 3.3: Network with 4 DOA, 4 TOA sensors and 4 communication nodes.

The following MM algorithms have been compared: centralized GPB_1 and IMM, distributed GPB_1 and IMM (i.e. DGPB_1 and DIMM) as well as the DIMM-CL algorithm of reference [LJ12]. The performance of the various algorithms is measured by the PRMSE. Notice that averaging is carried out over time and Monte Carlo trials for the centralized algorithms, while further averaging over network nodes is applied for distributed algorithms. Table 3.9 summarizes performance of the various algorithms using the network of Fig. 3.3. Notice that the PRMSEs of DIMM-CL are not included in Table 3.9 since for all the considered values of L such an algorithm exhibited a divergent behaviour. These results are actually consistent with the performance evaluation in [LJ12] where, considering circular networks with $N \geq 4$ nodes, a minimum number of $L = 10$ is required to track the object. Moreover, considering the network in Fig. 3.3 without communication nodes, $L = 80$ consensus steps are required for non-divergent behaviour of the DIMM-CL in [LJ12]. Notice that such a number of consensus steps might still be too large for practical applications, which is also stated in [LJ12]. As it can be seen from table 3.9, DIMM performs significantly

better than DGPB_1 and, due to the presence of communication nodes, the distributed case exhibits significantly worse performance compared to the centralized case. On the other hand it is evident that, even in presence of communication nodes (bringing no information about the object position) and of a highly maneuvering object, distributed MM algorithms still work satisfactorily by distributing the little available information throughout the network by means of consensus. In fact, despite these difficulties, the object can still be tracked with reasonable confidence. Fig. 3.4 shows, for $L = 5$, the differences between two nodes that are placed almost at opposite positions in the network. It is noticeable that the main difficulty in tracking by communication nodes arises when a turn is approaching, especially, in this trajectory, during the last turn. Clearly, one should use a higher number of consensus steps L to allow data to reach two distant points in the network and, hence, to have similar estimated trajectory and mode probabilities.

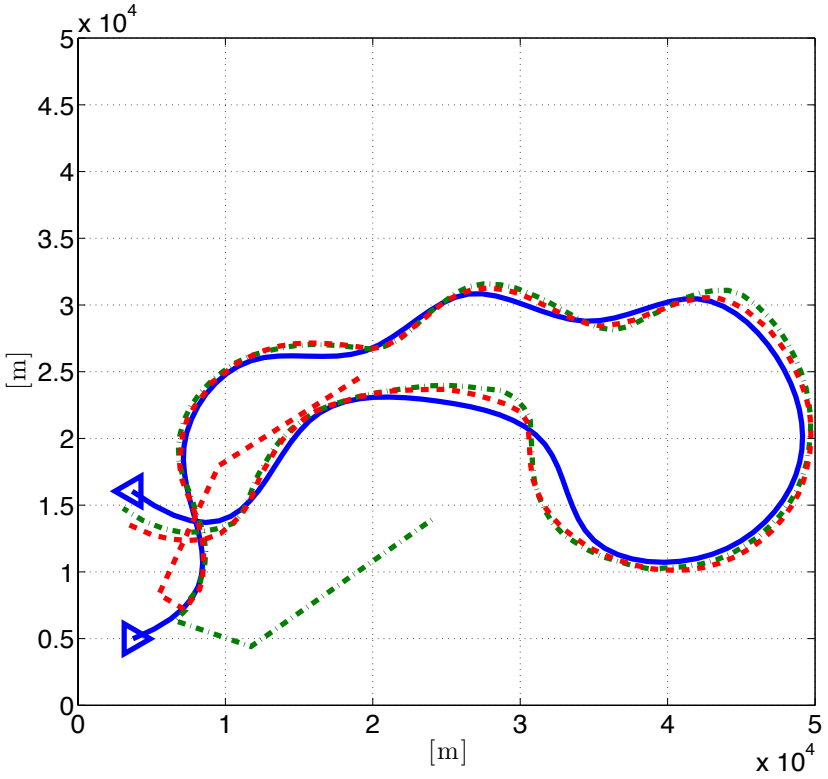


Figure 3.4: Estimated object trajectories of node 1 (TOA, green dash-dotted line) and node 2 (TOA, red dashed line) in the sensor network of fig. 3.3. The real object trajectory is the blue continuous line.

Table 3.4: Centralized Bayesian Multiple-Model (CBMM) filter pseudo-code

procedure CBMM(NODE i, TIME k)

Prediction

for mode $j \in \mathcal{R}$ **do**

$$p^j_{k|k-1}(x) = \int p_w(j, x - f(j,\zeta))\; p^j_{k-1}(\zeta)\; d\zeta$$

$$\mu^j_{k|k-1} = \sum_{t\in\mathcal{R}} p_{jt}\, \mu^t_k$$

end for

Correction

for mode $j \in \mathcal{R}$ **do**

$$p^j_k(x) = \frac{p_v\left(j, y_k - h\left(j,x\right)\right)\; p_{k|k-1}(x)}{\int p_v\left(j, y_k - h(j,\zeta)\right)\; p_{k|k-1}(\zeta)\, d\zeta}$$

$$g^j_k = p\Big(y_k | y^{k-1}, m_k = j\Big)$$

$$\mu^j_k = \frac{g^j_k\, \mu^j_{k|k-1}}{\sum_{t\in\mathcal{R}} g^t_k\, \mu^t_{k|k-1}}$$

end for

Mode fusion

$$p_k(x) = \sum_{j\in\mathcal{R}} \mu^j_k\, p^j_k(x)$$

Mixing

$$\forall j, t \in \mathcal{R}:\; \mu^{t|j}_k = \frac{p_{jt}\, \mu^t_k}{\sum_{\imath\in\mathcal{R}} p_{j\imath}\, \mu^\imath_k}$$

$$\forall j \in \mathcal{R}:\; \overline{p}^j_k(x) = \sum_{t\in\mathcal{R}} \mu^{t|j}_k\; p^t_k(x)$$

Re-initialization

$$\forall j \in \mathcal{R}:\; p^j_k(x) = \overline{p}^j_k(x)$$

end procedure

Table 3.5: Centralized First-Order Generalized Pseudo-Bayesian (CGBP_1) filter pseudo-code

procedure CGBP_1(NODE i, TIME k)

Prediction

for mode $j \in \mathcal{R}$ **do**

from $\hat{x}^j_{k-1}, P^j_{k-1}$ **compute** $\hat{x}^j_{k|k-1}, P^j_{k|k-1}$ **via EKF or UKF**

$$\mu^j_{k|k-1} = \sum_{t \in \mathcal{R}} p_{jt}\, \mu^t_k$$

end for

Correction

for mode $j \in \mathcal{R}$ **do**

from $\hat{x}^j_{k|k-1}, P^j_{k|k-1}$ **compute** $\hat{x}^j_k, P^j_k$, e^j_k, S^j_k **via EKF or UKF**

$$g^j_k = \frac{1}{\sqrt{\det\left(2\omega S^j_k\right)}}\, e^{-\frac{1}{2}\left(e^j_k\right)^\top \left(S^j_k\right)^{-1} e^j_k}$$

$$\mu^j_k = \frac{g^j_k\, \mu^j_{k|k-1}}{\sum_{t \in \mathcal{R}} g^t_k\, \mu^t_{k|k-1}}$$

end for

Mode fusion

$$\hat{x}_k = \sum_{j \in \mathcal{R}} \mu^j_k\, \hat{x}^j_k$$

$$P_k = \sum_{j \in \mathcal{R}} \mu^j_k \left[P^j_k + \left(\hat{x}_k - \hat{x}^j_k\right)\left(\hat{x}_k - \hat{x}^j_k\right)^\top\right]$$

$\forall j \in \mathcal{R} : \hat{x}^j_k = \hat{x}_k,\ P^j_k = P_k$

end procedure

Table 3.6: Centralized Interacting Multiple Model (CIMM) filter pseudo-code

procedure CIMM(NODE i, TIME k)

Prediction

As in the GPB_1 algorithm of Table 3.5

Correction

As in the GPB_1 algorithm of Table 3.5

Mode fusion

As in the GPB_1 algorithm of Table 3.5

Mixing

$$\forall j, t \in \mathcal{R} : \mu_k^{t|j} = \frac{p_{jt}\, \mu_k^t}{\sum_{\imath \in \mathcal{R}} p_{j\imath}\, \mu_k^\imath}$$

for mode $j \in \mathcal{R}$ **do**

$$\overline{x}_k^j = \sum_{t \in \mathcal{R}} \mu_k^{t|j}\, \hat{x}_k^t$$

$$\overline{P}_k^j = \sum_{t \in \mathcal{R}} \mu_k^{t|j} \left[P_k^t + \left(\overline{x}_k^j - \hat{x}_k^t \right) \left(\overline{x}_k^j - \hat{x}_k^t \right)^\top \right]$$

end for

Re-initialization

$$\forall j \in \mathcal{R} : \hat{x}_k^j = \overline{x}_k^j,\ P_k^j = \overline{P}_k^j$$

end procedure

Table 3.7: Distributed GPB_1 (DGPB_1) filter pseudo-code

procedure DGPB_1(NODE i, TIME k)

Prediction

for mode $j \in \mathcal{R}$ **do**

given $\hat{x}^{i,j}_{k-1}$ **and** $P^{i,j}_{k-1}$ **compute** $\hat{x}^{i,j}_{k|k-1}$ **and** $P^{i,j}_{k|k-1}$ **via EKF or UKF**

$\mu^{i,j}_{k|k-1} = \sum_{t \in \mathcal{R}} p_{jt}\, \mu^{i,t}_{k-1}$

end for

Correction

for mode $j \in \mathcal{R}$ **do**

if $i \in \mathcal{S}$ **then**

from $\hat{x}^{i,j}_{k|k-1}, P^{i,j}_{k|k-1}$ **compute** $\hat{x}^{i,j}_{k}, P^{i,j}_{k}$, $e^{i,j}_k$ $S^{i,j}_k$ **via EKF or UKF**

$g^{i,j}_k = \left[\det\left(2\omega S^{i,j}_k\right)\right]^{-\frac{1}{2}} e^{-\frac{1}{2}\left(e^{i,j}_k\right)^\top \left(S^{i,j}_k\right)^{-1} e^{i,j}_k}$

$\mu^{i,j}_{k,0} = g^{i,j}_k \mu^{i,j}_{k|k-1} \left[\sum_{t \in \mathcal{R}} g^{i,t}_k \mu^{i,t}_{k|k-1}\right]^{-1}$

else if $i \in \mathcal{C}$ **then**

$\hat{x}^{i,j}_k = \hat{x}^{i,j}_{k|k-1}$, $P^{i,j}_k = P^{i,j}_{k|k-1}$, $\mu^{i,j}_{k,0} = \mu^{i,j}_{k|k-1}$

end if

end for

Consensus on modal probabilities

for mode $j \in \mathcal{R}$ **do**

for $l = 1, \ldots, L$ **do**

$\mu^{i,j}_{k,l} = \bigoplus_{\imath \in \mathcal{N}^i} \left[\omega^{i,\imath} \odot \mu^{\imath,j}_{k,l-1}\right]$

end for

$\mu^{i,j}_k = \mu^{i,j}_{k,L}$

end for

Mode fusion

$\hat{x}^i_{k,0} = \sum_{j \in \mathcal{R}} \mu^{i,j}_k\, \hat{x}^{i,j}_k$

$P^i_{k,0} = \sum_{j \in \mathcal{R}} \mu^{i,j}_k \left[P^{i,j}_k + \left(\hat{x}^i_{k,0} - \hat{x}^{i,j}_k\right)\left(\hat{x}^i_{k,0} - \hat{x}^{i,j}_k\right)^\top\right]$

Consensus on the fused PDF

$\Omega^i_{k,0} = \left(P^i_{k,0}\right)^{-1}$, $q^i_{k,0} = \Omega^i_{k,0}\hat{x}^i_{k,0}$

for $l = 1, \ldots, L$ **do**

$\Omega^i_{k,l} = \sum_{\imath \in \mathcal{N}^i} \omega^{i,\imath}\, \Omega^\imath_{k,l-1}$

$q^i_{k,l} = \sum_{\imath \in \mathcal{N}^i} \omega^{i,\imath}\, q^\imath_{k,l-1}$

end for

Re-initialization

$\forall j \in \mathcal{R} : \hat{x}^{i,j}_k = \left(\Omega^i_{k,L}\right)^{-1} q^i_{k,L}$, $P^{i,j}_k = \left(\Omega^i_{k,L}\right)^{-1}$

end procedure

Table 3.8: Distributed IMM (DIMM) filter pseudo-code

procedure DIMM(NODE i, TIME k)

Prediction

As in the DGPB_1 algorithm of Table 3.7

Correction

for mode $j \in \mathcal{R}$ **do**

if $i \in \mathcal{S}$ **then**

from $\hat{x}^{i,j}_{k|k-1}, P^{i,j}_{k|k-1}$ **compute** $\hat{x}^{i,j}_{k,0}, P^{i,j}_{k,0}$, $e^{i,j}_k$, $S^{i,j}_k$ **via EKF or UKF**

$$g^{i,j}_k = \left[\det\left(2\omega S^{i,j}_k\right)\right]^{-\frac{1}{2}} e^{-\frac{1}{2}\left(e^{i,j}_k\right)^\top \left(S^{i,j}_k\right)^{-1} e^{i,j}_k}$$

$$\mu^{i,j}_{k,0} = g^{i,j}_k \mu^{i,j}_{k|k-1} \left[\sum_{t\in\mathcal{R}} g^{i,t}_k \mu^{i,t}_{k|k-1}\right]^{-1}$$

end if

if $i \in \mathcal{C}$ **then**

$$\hat{x}^{i,j}_{k,0} = \hat{x}^{i,j}_{k|k-1},\ P^{i,j}_{k,0} = P^{i,j}_{k|k-1},\ \mu^{i,j}_{k,0} = \mu^{i,j}_{k|k-1}$$

end if

end for

Parallel consensus on modal probabilities & mode-matched PDFs

for mode $j \in \mathcal{R}$ **do**

$$\Omega^{i,j}_{k,0} = \left(P^{i,j}_{k,0}\right)^{-1},\ q^{i,j}_{k,0} = \Omega^{i,j}_{k,0}\, \hat{x}^{i,j}_{k,0}$$

for $l = 1, \ldots, L$ **do**

$$\mu^{i,j}_{k,l} = \bigoplus_{\imath\in\mathcal{N}^i} \left[\omega^{i,\imath} \odot \mu^{\imath,j}_{k,l-1}\right]$$

$$\Omega^{i,j}_{k,l} = \sum_{\imath\in\mathcal{N}^i} \omega^{i,\imath}\, \Omega^{\imath,j}_{k,l-1},\ q^{i,j}_{k,l} = \sum_{\imath\in\mathcal{N}^i} \omega^{i,\imath}\, q^{\imath,j}_{k,l-1}$$

end for

$$\mu^{i,j}_k = \mu^{i,j}_{k,L}$$

end for

Mode fusion

$$\forall j \in \mathcal{R} : P^{i,j}_{k,L} = \left(\Omega^{i,j}_{k,L}\right)^{-1},\ \hat{x}^{i,j}_{k,L} = P^{i,j}_{k,L} q^{i,j}_{k,L}$$

$$\hat{x}^i_k = \sum_{j\in\mathcal{R}} \mu^{i,j}_k\, \hat{x}^{i,j}_{k,L}$$

$$P^i_k = \sum_{j\in\mathcal{R}} \mu^{i,j}_k \left[P^{i,j}_{k,L} + \left(\hat{x}^i_k - \hat{x}^{i,j}_{k,L}\right)\left(\hat{x}^i_k - \hat{x}^{i,j}_{k,L}\right)^\top\right]$$

Mixing

for mode $j \in \mathcal{R}$ **do**

$$\forall t \in \mathcal{R} : \mu^{i,t|j}_k = p_{jt}\, \mu^{i,t}_k \left[\sum_{\jmath\in\mathcal{R}} p_{j\jmath}\, \mu^{\jmath,h}_k\right]^{-1}$$

$$\hat{x}^{i,j}_k = \sum_{t\in\mathcal{R}} \mu^{i,t|j}_k\, \hat{x}^{i,j}_{k,L}$$

$$P^{i,j}_k = \sum_{t\in\mathcal{R}} \mu^{i,t|j}_k \left[P^{i,t}_{k,L} + \left(\hat{x}^{i,t}_k - \hat{x}^{i,t}_{k,L}\right)\left(\hat{x}^{i,t}_k - \hat{x}^{i,t}_{k,L}\right)^\top\right]$$

end for

end procedure

Table 3.9: Performance comparison for the sensor network of fig. 3.3

	GPB_1	IMM
PRMSE $[m]$	538	446

PRMSE $[m]$	$DGPB_1$	DIMM
$L = 1$	1603	960
$L = 2$	1312	781
$L = 3$	183	711
$L = 4$	1103	672
$L = 5$	1052	648

4 Distributed multi-object filtering

Contents

The focus of the next sections is on *Distributed MOF* (DMOF) over the network of section 2.1 wherein each node (tracking agent) locally updates multi-object information exploiting the multi-object dynamics and the available local measurements, exchanges such information with communicating agents and then carries out a fusion step in order to combine the information from all neighboring agents. Specifically, a CPHD filtering approach (see subsection 2.3.5) to DMOF will be adopted [BCF+13a]. Hence, the agents will locally update and fuse the cardinality PMF $\rho(\cdot)$ and the location PDF $s(\cdot)$ (see (2.56)) that, for the sake of brevity, will also be referred to as the CPHD.

More formally, the DMOF problem over the network $(\mathcal{N}, \mathcal{A})$ can be stated as follows. Each node $i \in \mathcal{N}$ must estimate at each time $k \in \{1, 2, \dots\}$ the CPHD of the unknown multi-object set X_k in (2.62)-(2.67) given local measurements Y_κ^i for all $\kappa \geq 1$ up to time k and data received from all adjacent nodes $j \in \mathcal{N}^i \backslash \{i\}$ so that the estimated pair $(\rho_k^i(\cdot), s_k^i(\cdot))$ be as close as possible to the one that would be provided by a centralized CPHD filter simultaneously processing information from all nodes.

Clearly local CPHD filtering is the same as the centralized one of (2.70)-(2.71), operating on the local measurement set Y_k^i instead of Y_k. Conversely, CPHD fusion deserves special attention and will be tackled in the next sections.

4.1 Multi-object information fusion via consensus

Suppose that, in each node i of the sensor network, a multi-object density $f^i(X)$ is available which has been computed on the basis of the information (collected locally or propagated

from other nodes) which has become available to node i. Aim of this section is to investigate whether it is possible to devise a suitable distributed algorithm guaranteeing that all the nodes of the network reach an agreement regarding the multi-object density of the unknown multi-object set. An important feature the devised algorithm should enjoy is *scalability*, i.e. the processing load of each node must be independent of the network size. For this reason, approaches based on the optimal (Bayes) fusion rule [CMC90, Mah00] are ruled out. In fact, they would require, for each pair of adjacent nodes (i, j), the knowledge of the multi-object density $f(X|I^i \cap I^j)$ conditioned to the common information $I^i \cap I^j$ and, in a practical network, it is impossible to keep track of such a common information in a scalable way. Hence, some robust suboptimal fusion technique has to be devised.

4.1.1 Multi-object Kullback-Leibler average

The first important issue to be addressed is how to define the average of the local multi-object densities $f^i(X)$. To this end, taking into account the benefit provided by the RFS approach about defining multi-object densities, it is possible to extend the notion of KLA of single-object PDFs introduced in subsection 2.4.2 to multi-object ones. Let us first define the notion of KLD to multi-object densities $f(X)$ and $g(X)$ by

$$D_{KL}\left(f \parallel g\right) \triangleq \int f(X) \log\left(\frac{f(X)}{g(X)}\right) \delta X \tag{4.1}$$

where the integral in (4.1) must be interpreted as a set integral according to the definition (2.36). Then, the weighted KLA $f_{KLA}(X)$ of the multi-object densities $f^i(X)$ is defined as follows

$$f_{KLA} = \arg\inf_{f} \sum_{i \in \mathcal{N}} \omega^i D_{KL}\Big(f \parallel f^i\Big)\,. \tag{4.2}$$

with weights ω^i satisfying

$$\omega^i \geq 0\,, \qquad \sum_{i \in \mathcal{N}} \omega^i = 1. \tag{4.3}$$

Notice from (4.2) that the weighted KLA of the agent densities is the one that minimizes the weighted sum of distances from such densities. In particular, if N is the number of agents and $\omega^i = 1/N$ for $i = 1, \ldots, N$, (4.2) provides the (unweighted) KLA which averages the agent densities giving to all of them the same level of confidence. An interesting interpretation of such a notion can be given recalling that, in Bayesian statistics, the KLD (4.1) can be seen as the information gain achieved when moving from a prior $g(X)$ to a posterior $f(X)$. Thus, according to (4.2), the average PDF is the one that minimizes the sum of the information gains from the initial multi-object densities. Thus, this choice is coherent with the *Principle of Minimum Discrimination Information* (PMDI) according to which the probability density which best represents the current state of knowledge is the one which produces an information gain as small as possible (see [Cam70, Aka98] for a discussion on such a principle and its relation with Gauss' principle and maximum likelihood estimation) or, in other words [Jay03]:

> "the probability assignment which most honestly describes what we know should be the most conservative assignment in the sense that it does not permit one to draw any conclusions not warranted by the data".

The adherence to the PMDI is important in order to counteract the so-called *data incest* phenomenon, i.e. the unaware reuse of the same piece of information due to the presence of loops within the network.

The following result holds.

Theorem 2 (Multi-object KLA)**.**
The weighted KLA defined in (4.2) turns out to be given by

$$f_{KLA}(X) = \frac{\prod\limits_{i\in\mathcal{N}} \left[f^i(X)\right]^{\omega^i}}{\int \prod\limits_{i\in\mathcal{N}} \left[f^i(X)\right]^{\omega^i} \delta X} \tag{4.4}$$

4.1.2 Consensus-based multi-object filtering

The identity (4.4) with $\omega^i = 1/|\mathcal{N}|$ can be rewritten as

$$f_{KLA}(X) = \bigoplus_{i\in\mathcal{N}} \left(\frac{1}{|\mathcal{N}|} \odot f^i(X)\right). \tag{4.5}$$

Then, the global (collective) KLA (4.5), which would require all the local multi-object densities to be available, can be computed in a distributed and scalabale way by iterating regional averages through the consensus algorithm

$$\hat{f}^i_{l+1}(X) = \bigoplus_{j\in\mathcal{N}^i} \left(\omega^{i,j} \odot \hat{f}^j_l(X)\right), \qquad \forall i \in \mathcal{N} \tag{4.6}$$

with $\hat{f}^i_0(X) = f^i(X)$. In fact, thanks to properties P.A - P.F, it can be seen that[1]

$$\hat{f}^i_l(X) = \bigoplus_{j\in\mathcal{N}} \left(\omega^{i,j}_l \odot f^j(X)\right), \qquad \forall i \in \mathcal{N} \tag{4.7}$$

where $\omega^{i,j}_l$ is defined as the element (i,j) of the matrix Ω^l. With this respect, recall that when the consensus weights $\omega^{i,j}$ are chosen so as to ensure that the matrix Ω is doubly stochastic, one has

$$\lim_{l\to+\infty} \omega^{i,j}_l = \frac{1}{|\mathcal{N}|}, \quad \forall i,j \in \mathcal{N}. \tag{4.8}$$

[1]Notice that the scalar multiplication operator is defined only for strictly positive scalars. However, in equation (4.7) it is admitted that some of the scalar weights $\omega^{i,j}_l$ be zero. The understanding is that, whenever $\omega^{i,j}_l$ is equal to zero, the corresponding multi-object density $f^j(X)$ is omitted from the addition. This can always be done, since for each $i \in \mathcal{N}$ and each l, at least one of the weights $\omega^{i,j}_l$ is strictly positive.

Hence, as the number of consensus steps increases, each local multi-object density "tends" to the global KLA (4.5).

4.1.3 Connection with existing approaches

It is important to point out that the fusion rule (4.4), which has been derived as KLA of the local multi-object densities, coincides with the Chernoff fusion [CT12, CCM10] known as *Generalized Covariance Intersection* (GCI) for multi-object fusion, first proposed by Mahler [Mah00]. In particular, it is clear that $f_{KLA}(X)$ in (4.4) is nothing but the NWGM of the agent multi-object densities $f^i(X)$; it is also called *Exponential Mixture Density* (EMD) [UJCR10, UCJ11].

Remark 12. The name GCI stems from the fact that (4.4) is the multi-object counterpart of the analogous fusion rule for (single-object) PDFs [Mah00, Hur02, JBU06, Jul08] which, in turn, is a generalization of CI originally conceived [JU97] for Gaussian PDFs.

Remind that, given estimates $\hat{x}_i$ of the same quantity x from multiple estimators with relative covariances P_i and unknown correlations, their CI fusion is given by

$$\begin{aligned} P &= \left(\sum_{i \in \mathcal{N}} \omega^i P_i^{-1} \right)^{-1} \\ \hat{x} &= P \sum_{i \in \mathcal{N}} \omega^i P_i^{-1} \hat{x}_i \end{aligned} \tag{4.9}$$

The peculiarity of (4.9) is that, for any choice of the weights ω^i satisfying (4.3) and provided that all estimates are consistent in the sense that

$$E\Big[(x - \hat{x}_i) (x - \hat{x}_i)^\top \Big] \leq P_i \,, \qquad \forall i \tag{4.10}$$

then the fused estimate also turns out to be consistent, i.e.

$$E\Big[(x - \hat{x}) (x - \hat{x})^\top \Big] \leq P \,. \tag{4.11}$$

It can easily be shown that, for normally distributed estimates, (4.9) is equivalent to

$$p(x) = \frac{\prod\limits_{i \in \mathcal{N}} \Big[p^i(x) \Big]^{\omega^i}}{\int \prod\limits_{i \in \mathcal{N}} \Big[p^i(x) \Big]^{\omega^i} dx} \tag{4.12}$$

where $p^i(\cdot) \triangleq \mathcal{N}(\cdot; \hat{x}_i, P_i)$ is the Gaussian PDF with mean $\hat{x}_i$ and covariance P_i. This suggested to use (4.12) for arbitrary, possibly non Gaussian, PDFs. As a final remark, notice that the consistency property, which was the primary motivation that led to the development

of the CI fusion rule, is in accordance with the PMDI discussed in Section 4.1.1, which indeed represents one of the main positive features of the considered multi-object fusion rule.

4.1.4 CPHD fusion

Whenever the object set is modelled as an i.i.d. cluster process, the agent multi-object densities to be fused take the form

$$f^i(X) = |X|!\, \rho^i(|X|) \prod_{x \in X} s^i(x) \tag{4.13}$$

where $\left(\rho^i(n), s^i(x)\right)$ is the CPHD of agent i. In [CJMR10] it is shown that in this case the GCI fusion (4.4) yields

$$\overline{f}(X) = |X|!\, \overline{\rho}(|X|) \prod_{x \in X} \overline{s}(x) \tag{4.14}$$

where

$$\overline{s}(x) = \frac{\prod_{i \in \mathcal{N}} \left[s^i(x)\right]^{\omega^i}}{\int \prod_{i \in \mathcal{N}} \left[s^i(x)\right]^{\omega^i} dx}, \tag{4.15}$$

$$\overline{\rho}(n) = \frac{\prod_{i \in \mathcal{N}} \left[\rho^i(n)\right]^{\omega^i} \left\{\int \prod_{i \in \mathcal{N}} \left[s^i(x)\right]^{\omega^i} dx\right\}^n}{\sum_{j=0}^{\infty} \prod_{i \in \mathcal{N}} \left[\rho^i(j)\right]^{\omega^i} \left\{\int \prod_{i \in \mathcal{N}} \left[s^i(x)\right]^{\omega^i} dx\right\}^j}. \tag{4.16}$$

In words, (4.14)-(4.16) amount to state that the fusion of i.i.d. cluster processes provides an i.i.d. cluster process whose location density $\overline{s}(\cdot)$ is the weighted geometric mean of the agent location densities $s^i(\cdot)$, while the fused cardinality $\overline{\rho}(\cdot)$ is obtained by the more complicated expression (4.16) also involving the agent location PDFs besides the agent cardinality PMFs.

Please notice that, in principle, both the cardinality PMF and the location PDF are infinite-dimensional. For implementation purposes, finite-dimensional parametrizations of both need to be adopted. As far as the cardinality PMF $\rho(n)$ is concerned, it is enough to assume a sufficiently large maximum number of objects n_{max} present in the scene and restrict $\rho(\cdot)$ to the finite subset of integers $\{0, 1, \ldots, n_{max}\}$. As for the location PDF, two finitely-parameterized representations based on the SMC or, respectively, GM approaches are most commonly adopted. The SMC approach consists of representing location PDFs as linear combinations of delta Dirac functions, i.e.

$$s(x) = \sum_{j=1}^{N_p} w_j\, \delta(x - x_j) \tag{4.17}$$

Conversely, the GM approach expresses location PDFs as linear combinations of Gaussian

components, i.e.

$$s(x) = \sum_{j=1}^{N_G} \alpha_j \, \mathcal{N}(x; \hat{x}_j, P_j) \tag{4.18}$$

Recent work [UJCR10] has presented an implementation of the distributed multi-object fusion following the SMC approach. For DMOF over a sensor network, typically characterized by limited processing power and energy resources of the individual nodes, it is of paramount importance to reduce as much as possible local (in-node) computations and inter-node data communication. In this respect, the GM approach promises to be more parsimonious (usually the number of Gaussian components involved is orders of magnitude lower than the number of particles required for a reasonable tracking performance) and hence preferable. For this reason, a GM implementation of the fusion (4.15) has been adopted in the present work which, further, exploits consensus in order to carry out the fusion in a fully distributed way. For the sake of simplicity, let us consider only two agents, a and b, with GM location densities

$$s^i(x) = \sum_{j=1}^{N_G^i} \alpha_j^i \, \mathcal{N}\left(x; \hat{x}_j^i, P_j^i\right) , \qquad i = a, b \tag{4.19}$$

A first natural question is whether the fused location PDF

$$\overline{s}(x) = \frac{[s^a(x)]^\omega \left[s^b(x)\right]^{1-\omega}}{\int [s^a(x)]^\omega \left[s^b(x)\right]^{1-\omega} dx} \tag{4.20}$$

is also a GM. Notice that (4.20) involves exponentiation and multiplication of GMs. To this end, it is useful to draw the following observations concerning elementary operations on Gaussian components and mixtures.

- The power of a Gaussian component is a Gaussian component, more precisely

$$[\alpha \, \mathcal{N}(x; \hat{x}, P)]^\omega = \alpha^\omega \, \beta(\omega, P) \, \mathcal{N}\left(x; \hat{x}, \frac{P}{\omega}\right) \tag{4.21}$$

  where

$$\beta(\omega, P) \triangleq \frac{\left[\det(2\pi P \omega^{-1})\right]^{\frac{1}{2}}}{\left[\det(2\pi P)\right]^{\frac{\omega}{2}}} \tag{4.22}$$

- The product of Gaussian components is a Gaussian component, more precisely [WM06, Wil03]

$$\alpha_1 \, \mathcal{N}(x; \hat{x}_1, P_1) \cdot \alpha_2 \, \mathcal{N}(x; \hat{x}_2, P_2) = \alpha_{12} \, \mathcal{N}(x; \hat{x}_{12}, P_{12}) \tag{4.23}$$

  where

$$P_{12} = \left(P_1^{-1} + P_2^{-1}\right)^{-1} \tag{4.24}$$

$$\hat{x}_{12} = P_{12}\left(P_1^{-1}\hat{x}_1 + P_2^{-1}\hat{x}_2\right) \tag{4.25}$$

$$\alpha_{12} \quad = \quad \alpha_1 \, \alpha_2 \, \mathcal{N}(\hat{x}_1 - \hat{x}_2; 0, P_1 + P_2) \tag{4.26}$$

- Due to (4.23) and the distributive property, the product of GMs is a GM. In particular, if $s^a(\cdot)$ and $s^b(\cdot)$ have N_G^a and, respectively, N_G^b Gaussian components, then $s^a(\cdot)\, s^b(\cdot)$ will have $N_G^a N_G^b$ components.

- Exponentiation of a GM does not provide, in general, a GM.

As a consequence of the latter observation, the fusion (4.19)-(4.20) does not provide a GM. Hence, in order to preserve the GM form of the location PDF throughout the computations a suitable approximation of the GM exponentiation has to be devised. In [Jul06] it is suggested to use the following approximation

$$\left[\sum_{j=1}^{N_G} \alpha_j \, \mathcal{N}(x; \hat{x}_j, P_j)\right]^{\omega} \cong \sum_{j=1}^{N_G} \left[\alpha_j \mathcal{N}(x; \hat{x}_j, P_j)\right]^{\omega} = \sum_{j=1}^{N_G} \alpha_j^{\omega} \, \beta(\omega, P_j) \, \mathcal{N}\left(x; \hat{x}_j, \frac{P_j}{\omega}\right) \tag{4.27}$$

As a matter of fact, the above approximation seems reasonable whenever the cross-products of the different terms in the GM are negligible for all x; this, in turn, holds provided that the centers $\hat{x}_i$ and $\hat{x}_j$, $i \neq j$, of the Gaussian components are well separated, as measured by the respective covariances P_i and P_j. In geometrical terms, the more separated are the confidence ellipsoids of the Gaussian components the smaller should be the error involved in the approximation (4.27). In mathematical terms, the conditions for the validity of (4.27) can be expressed in terms of Mahalanobis distance [Mah36] inequalities of the form

$$\begin{aligned} (\hat{x}_i - \hat{x}_j)^{\top} \, P_i^{-1} \, (\hat{x}_i - \hat{x}_j) \quad &\gg \quad 1 \\ (\hat{x}_i - \hat{x}_j)^{\top} \, P_j^{-1} \, (\hat{x}_i - \hat{x}_j) \quad &\gg \quad 1 \end{aligned}$$

Provided that the use of (4.27) is preceded by a suitable merging step that fuses Gaussian components with Mahalanobis (or other type of) distance below a given threshold, the approximation seems reasonable. To this end, the merging algorithm proposed by Salmond [Sal88,Sal90] represents a good tool.

Exploiting (4.27), the fusion (4.20) can be approximated as follows:

$$\overline{s}(x) = \frac{\displaystyle\sum_{i=1}^{N_G^a} \sum_{j=1}^{N_G^b} \alpha_{ij}^{ab} \, \mathcal{N}\left(x; \hat{x}_{ij}^{ab}, P_{ij}^{ab}\right)}{\displaystyle\int \sum_{i=1}^{N_G^a} \sum_{j=1}^{N_G^b} \alpha_{ij}^{ab} \, \mathcal{N}\left(x; \hat{x}_{ij}^{ab}, P_{ij}^{ab}\right) dx} = \frac{\displaystyle\sum_{i=1}^{N_G^a} \sum_{j=1}^{N_G^b} \alpha_{ij}^{ab} \, \mathcal{N}\left(x; \hat{x}_{ij}^{ab}, P_{ij}^{ab}\right)}{\displaystyle\sum_{i=1}^{N_G^a} \sum_{j=1}^{N_G^b} \alpha_{ij}^{ab}} \tag{4.28}$$

where

$$P_{ij}^{ab} = \left[\omega \, (P_i^a)^{-1} + (1 - \omega) \left(P_j^b\right)^{-1}\right]^{-1} \tag{4.29}$$

$$\hat{x}_{ij}^{ab} = P_{ij}^{ab}\left[\omega\left(P_i^a\right)^{-1}\hat{x}_i^a + (1-\omega)\left(P_j^b\right)^{-1}\hat{x}_j^b\right] \tag{4.30}$$

$$\alpha_{ij}^{ab} = \left(\alpha_i^a\right)^{\omega}\left(\alpha_j^b\right)^{1-\omega}\beta(\omega, P_i^a)\,\beta\left(1-\omega, P_j^b\right)\mathcal{N}\left(\hat{x}_i^a - \hat{x}_j^b;\, 0,\, \frac{P_i^a}{\omega} + \frac{P_j^b}{1-\omega}\right) \tag{4.31}$$

Notice that (4.28)-(4.31) amounts to performing a CI fusion on any possible pair formed by a Gaussian component of agent a and a Gaussian component of agent b. Further the coefficient α_{ij}^{ab} of the resulting (fused) component includes a factor $\mathcal{N}\left(\hat{x}_i^a - \hat{x}_j^b;\, 0, P_i^a/\omega + P_j^b/(1-\omega)\right)$ that measures the separation of the two fusing components $(\hat{x}_i^a, P_i^a)$ and $\left(\hat{x}_j^b, P_j^b\right)$. In (4.28) it is reasonable to remove Gaussian components with negligible coefficients α_{ij}^{ab}. This can be done either by fixing a threshold for such coefficients or by checking whether the Mahalanobis distance

$$\sqrt{\left(x_i^a - x_i^b\right)^{\top}\left(\frac{P_a}{1-\omega} + \frac{P_b}{\omega}\right)^{-1}\left(x_i^a - x_i^b\right)} \tag{4.32}$$

falls below a given threshold. The fusion (4.20) can be easily extended to $N \geq 2$ agents by sequentially applying the pairwise fusion (4.29)-(4.30) $N-1$ times. Note that, by the associative and commutative properties of multiplication, the ordering of pairwise fusions is irrelevant.

4.2 Consensus GM-CPHD filter

This section presents the proposed *Consensus Gaussian Mixture-CPHD* (CGM-CPHD) filter algorithm. The sequence of operations carried out at each sampling interval k in each node $i \in \mathcal{N}$ of the network is reported in Table 4.1. All nodes $i \in \mathcal{N}$ operate in parallel at each sampling interval k in the same way, each starting from its own previous estimates of the cardinality PMF and location PDF in GM form

$$\left\{\rho_{k-1}^i(n)\right\}_{n=0}^{n_{max}}, \left\{\left(\alpha_j^i, \hat{x}_j^i, P_j^i\right)_{k-1}\right\}_{j=1}^{\left(N_G^i\right)_{k-1}} \tag{4.33}$$

and producing, at the end of the various steps listed in Table 4.1, its new estimates of the CPHD as well as of the object set, i.e.

$$\left\{\rho_k^i(n)\right\}_{n=0}^{n_{max}}, \left\{\left(\alpha_j^i, \hat{x}_j^i, P_j^i\right)_k\right\}_{j=1}^{\left(N_G^i\right)_k}, \widehat{X}_k^i \tag{4.34}$$

A brief description of the sequence of steps of the CGM-CPHD algorithm is in order.

1. First, each node i performs a local GM-CPHD filter update exploiting the multi-object dynamics and the local measurement set Y_k^i. The details of the GM-CPHD update (prediction and correction) can be found in [VVC07]. A merging step, to be described later, is introduced after the local update and before the consensus phase in order to reduce the number of Gaussian components and, hence, alleviate both the communication and

the computation burden.

2. Then, consensus takes place in each node i involving the subnetwork $\mathcal{N}^i$. Each node exchanges information (i.e., cardinality PMF and GM representation of the location PDF) with the neighbors; more precisely node i transmits its data to nodes j such that $i \in \mathcal{N}^j$ and waits until it receives data from $j \in \mathcal{N}^i \backslash \{i\}$. Next, node i carries out the GM-GCI fusion in (4.16) and (4.28)-(4.31) over $\mathcal{N}^i$. Finally, a merging step is applied to reduce the joint communication-computation burden for the next consensus step. This procedure is repeatedly applied for an appropriately chosen number $L \geq 1$ of consensus steps.

3. After the consensus, the resulting GM is further simplified by means of a pruning step to be described later. Finally, an estimate of the object set is obtained from the cardinality PMF and the pruned location GM via an estimate extraction step to be described later.

Table 4.1: Distributed GM-CPHD (DGM-CPHD) filter pseudo-code

procedure DGM-CPHD(NODE i, TIME k)
  Local Filtering
  LOCAL GM-CPHD PREDICTION ▷ See (2.70) and [VVC07]
  LOCAL GM-CPHD CORRECTION ▷ See (2.71) and [VVC07]
  MERGING ▷ See Table 4.2
  Information Fusion
  for $l = 1, \ldots, L$ **do**
    INFORMATION EXCHANGE
    GM-GCI FUSION ▷ See (4.16) and (4.28)
    MERGING ▷ See Table 4.2
  end for
  Extraction
  PRUNING ▷ See Table 4.3
  ESTIMATE EXTRACTION ▷ See Table 4.4
end procedure

While the local GM-CPHD update (prediction and correction) is thoroughly described in the literature [VVC07] and the GM-GCI fusion has been thoroughly dealt with in the previous section, the remaining steps (merging, pruning and estimate extraction) are outlined in the sequel.

Merging

Recall [VM06, VVC07] that the prediction step (2.70) and the correction step (2.71) of the PHD/CPHD filter make the number of Gaussian components increase. Hence, in order to avoid an unbounded growth with time of such components and make the GM-(C)PHD filter practically implementable, suitable component reduction strategies have to be adopted.

In [VM06, Section III.C, Table II], a reduction procedure based on truncation of components with low weights, merging of similar components and pruning, has been presented. For use in the proposed CGM-CPHD algorithm, it is convenient to deal with merging and pruning in a separate way. A pseudo-code of the merging algorithm is in Table 4.2.

Table 4.2: Merging pseudo-code

procedure MERGING($\{\hat{x}_i, P_i, \alpha_i\}_{i=1}^{N_G}, \gamma_m$)
  $t = 0$
  $\mathcal{I} = \{1, \ldots, N_G\}$
  repeat
    $t = t + 1$
    $j = \underset{i \in \mathcal{I}}{\operatorname{argmax}}\, \alpha_i$
    $\mathcal{MC} = \left\{ i \in \mathcal{I} \middle| (\hat{x}_j - \hat{x}_i)^\top P_i^{-1} (\hat{x}_j - \hat{x}_i) \leq \gamma_m \right\}$
    $\bar{\alpha}_t = \sum_{i \in \mathcal{MC}} \alpha_i$
    $\bar{x}_t = \frac{1}{\bar{\alpha}_k} \sum_{i \in \mathcal{MC}} \alpha_i \hat{x}_i$
    $\bar{P}_t = \frac{1}{\bar{\alpha}_k} \sum_{i \in \mathcal{MC}} \alpha_i \left[P_i + (\bar{x}_k - \hat{x}_i)(\bar{x}_k - \hat{x}_i)^\top \right]$
    $\mathcal{I} = \mathcal{I} \backslash \mathcal{MC}$
  until $\mathcal{I} \neq \emptyset$
  return $\left\{ \bar{x}_i, \bar{P}_i, \bar{\alpha}_i \right\}_{i=1}^{t}$
end procedure

Pruning

To prevent the number of Gaussian components from exceeding a maximum allowable value, say N_{max}, only the first N_{max} Gaussian components are kept while the other are removed from the GM. The pseudo-code of the pruning algorithm is in Table 4.3.

Table 4.3: Pruning pseudo-code

procedure PRUNING($\{\hat{x}_i, P_i, \alpha_i\}_{i=1}^{N_G}, N_{max}$)
  if $N_G > N_{max}$ **then**
    $\mathcal{I} = \{$indices of the N_{max} Gaussian components with highest weights $\alpha_i\}$
    return $\{\hat{x}_i, P_i, \alpha_i\}_{i \in \mathcal{I}}$
  end if
  return $\{\hat{x}_i, P_i, \alpha_i\}_{i=1}^{N_{max}}$
end procedure

Since the CPHD filter directly estimates the cardinality distribution, the estimated num-

ber of objects can be obtained via *Maximum A Posteriori* (MAP) estimation, i.e.

$$\hat{n}_k = \max_n \rho_k(n) \tag{4.35}$$

Given the MAP-estimated number of objects $\hat{n}_k$, estimate extraction is performed via the algorithm in Table 4.4 [VM06, Section III.C, Table III]. The algorithm extracts the $\hat{n}_k$ local maxima (peaks) of the estimated location PDF keeping those for which the corresponding weights are above a preset threshold γ_e.

Table 4.4: Estimate extraction pseudo-code

procedure ESTIMATE EXTRACTION($\{\hat{x}_i, P_i, \alpha_i\}_{i=1}^{N_G}, \hat{n}_k, \gamma_e$)
  $\mathcal{I} = \{$indices of the $\hat{n}_k$ Gaussian components with highest weights $\alpha_i\}$
  $\hat{X}_t = \emptyset$
  for $i \in \mathcal{I}$ **do**
    if $\alpha_i \hat{n}_k > \gamma_e$ **then**
      $\hat{X}_t = \hat{X}_t \cup \hat{x}_i$
    end if
  end for
  return $\hat{X}_t$
end procedure

4.3 Performance evaluation

To assess performance of the proposed CGM-CPHD algorithm described in section 4.2, a 2-dimensional (planar) multi-object tracking scenario is considered over a surveillance area of $50 \times 50\,[km^2]$, wherein the sensor network of Fig. 4.1 is deployed. The scenario consists of 6 objects as depicted in Fig. 4.2.

The object state is denoted by $x = [p_x,\, \dot{p}_x,\, p_y,\, \dot{p}_y]^\top$ where (p_x, p_y) and $(\dot{p}_x, \dot{p}_y)$ represent the object Cartesian position and, respectively, velocity components. The motion of objects is modeled by the filters according to the nearly-constant velocity model:

$$x_{k+1} = \begin{bmatrix} 1 & T_s & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T_s \\ 0 & 0 & 0 & 1 \end{bmatrix} x_k + w_k\,, \qquad Q = \sigma_w^2 \begin{bmatrix} \frac{1}{4}T_s^4 & \frac{1}{2}T_s^3 & 0 & 0 \\ \frac{1}{2}T_s^3 & T_s^2 & 0 & 0 \\ 0 & 0 & \frac{1}{4}T_s^4 & \frac{1}{2}T_s^3 \\ 0 & 0 & \frac{1}{2}T_s^3 & T_s^2 \end{bmatrix} \tag{4.36}$$

where $\sigma_w = 2\,[m/s^2]$ and the sampling interval is $T_s = 5\,[s]$.

As it can be seen from Fig. 4.1, the sensor network considered in the simulation consists of 4range-only (*Time Of Arrival*, TOA) and 3 bearing-only (*Direction Of Arrival*, DOA)

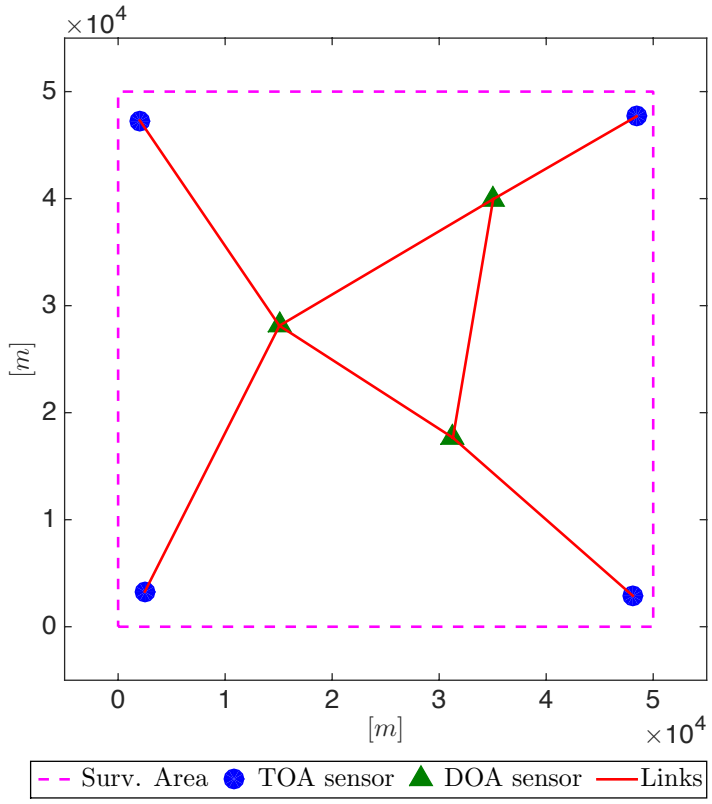


Figure 4.1: Network with 7 sensors: 4 TOA and 3 DOA.

sensors characterized by the following measurement functions:

$$h^i(x) = \begin{cases} \angle[(p_x - x^i) + j\,(p_y - y^i)], & \text{DOA} \\ \sqrt{(p_x - x^i)^2 + (p_y - y^i)^2}, & \text{TOA} \end{cases} \tag{4.37}$$

where (x^i, y^i) represents the known position of sensor i. The standard deviation of DOA and TOA measurement noises are taken respectively as $\sigma_{DOA} = 1\,[^\circ]$ and $\sigma_{TOA} = 100\,[m]$. Because of the non linearity of the aforementioned sensors, the UKF [JU04] is exploited in each sensor in order to update means and covariances of the Gaussian components.

Clutter is modeled as a Poisson process with parameter $\lambda_c = 5$ and uniform spatial distribution over the surveillance area; the probability of object detection is $P_D = 0.99$.

In the considered scenario, objects pass through the surveillance area with no prior information for object birth locations. Accordingly, a 40-component GM

$$d_b(x) = \sum_{j=1}^{40} \alpha_j \, \mathcal{N}(x; \hat{x}_j, P_j) \tag{4.38}$$

has been hypothesized for the birth intensity. Fig. 4.3 gives a pictorial view of $d_b(x)$; notice that the center of the Gaussian components are regularly placed along the border of the

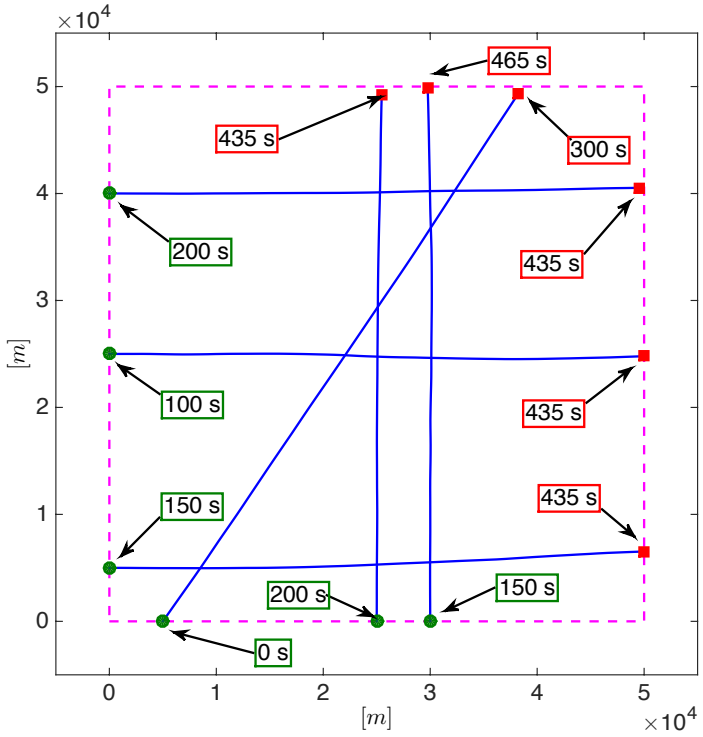


Figure 4.2: object trajectories considered in the simulation experiment. The start/end point for each trajectory is denoted, respectively, by •\■.

surveillance region. For all components, the same covariance $P_j = \text{diag}(10^6, 10^4, 10^6, 10^4)$ and the same coefficient $\alpha_j = \alpha$, such that $n_b = 40\alpha$ is the expected number of new-born objects, have been assumed.

The proposed consensus CGM-CPHD filter is compared to an analogous filter, called *Global GM-CPHD* (GGM-CPHD), that at each sampling interval performs a global fusion among all network nodes. Multi-object tracking performance is evaluated in terms of the *Optimal SubPattern Assignment* (OSPA) metric [SVV08]. The reported metric is averaged over $N_{mc} = 200$ Monte Carlo trials for the same object trajectories but different, independently generated, clutter and measurement noise realizations. The duration of each simulation trial is fixed to $500\,[s]$ (100 samples).

The parameters of the CGM-CPHD and GGM-CPHD filters have been chosen as follows: the survival probability is $P_S = 0.99$; the maximum number of Gaussian components is $N_{max} = 25$; the merging threshold is $\gamma_m = 4$; the truncation threshold is $\gamma_t = 10^{-4}$; the extraction threshold is $\gamma_e = 0.5$; the weight of each Gaussian component of the birth PHD function is chosen as $\alpha_j = 1.5 \cdot 10^{-3}$.

Figs. 4.4, 4.5, 4.6 and 4.7 display the statistics (mean and standard deviation) of the estimated number of objects obtained with CGM-CPHD (Fig. 4.4) and CGM-CPHD with $L = 1$ (Fig. 4.5), $L = 2$ (Fig. 4.6) and $L = 3$ (Fig. 4.7) consensus steps. Fig. 4.8 reports the OSPA metric (with Euclidean distance, $p = 2$, and cutoff parameter $c = 600$) for the

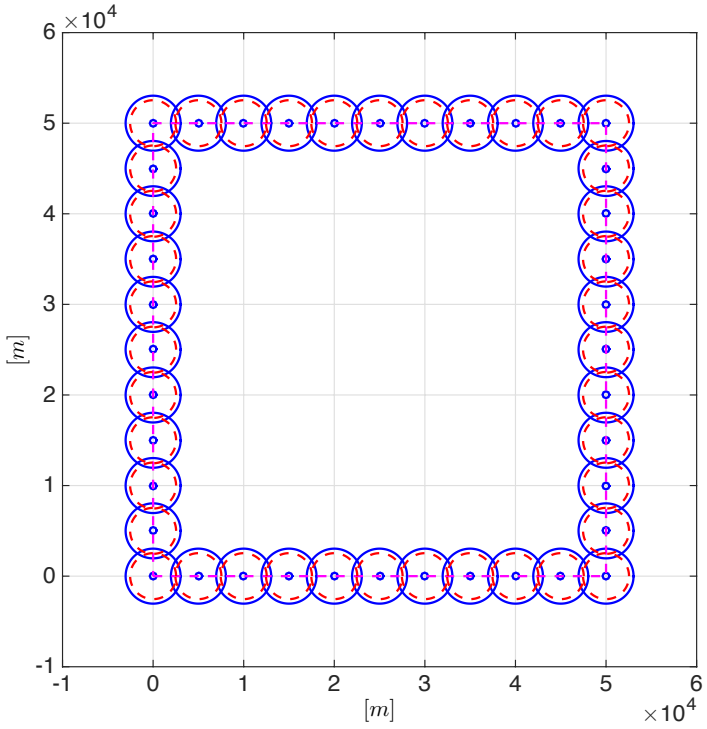


Figure 4.3: Borderline position GM initialization. The symbol ∘ denotes the component mean while the blue solid line and the red dashed line are, respectively, their 3σ and 2σ confidence regions.

same filters. As it can be seen from Fig. 4.8, the performance obtained with three consensus steps is significantly better than with a single one, and comparable with the one given by the non scalable GGM-CPHD filter which performs a global fusion over all network nodes. Similar considerations hold for cardinality estimation. These results show that by applying consensus, for a suitable number of steps, performance of distributed scalable algorithms is comparable to the one provided by the non scalable GGM-CPHD.

To provide an insightful view of the consensus process, Fig. 4.9 shows the GM representation within a particular sensor node (TOA sensor 1 in Fig. 4.1) before consensus and after $L = 1, 2, 3$ consensus steps. It can be noticed that consensus steps provide a progressive refinement of multi-object information. In fact, before consensus the CPHD reveals the presence of several false objects (see Fig. 4.9a); this is clearly due to the fact that the considered TOA sensor does not guarantee observability. Then, in the subsequent consensus steps (see Figs. 4.9b-4.9d), the weights of the Gaussian components corresponding to false objects as well as the covariance of the component relative to the true object are progressively reduced. Summing up, just one consensus step leads to acceptable performance but, in this case, the distributed algorithm tends to be less responsive and accurate in estimating object number and locations when compared to the centralized one (GGM-CPHD). However, by performing additional consensus steps, it is possible to retrieve performance comparable to the one of GGM-CPHD.

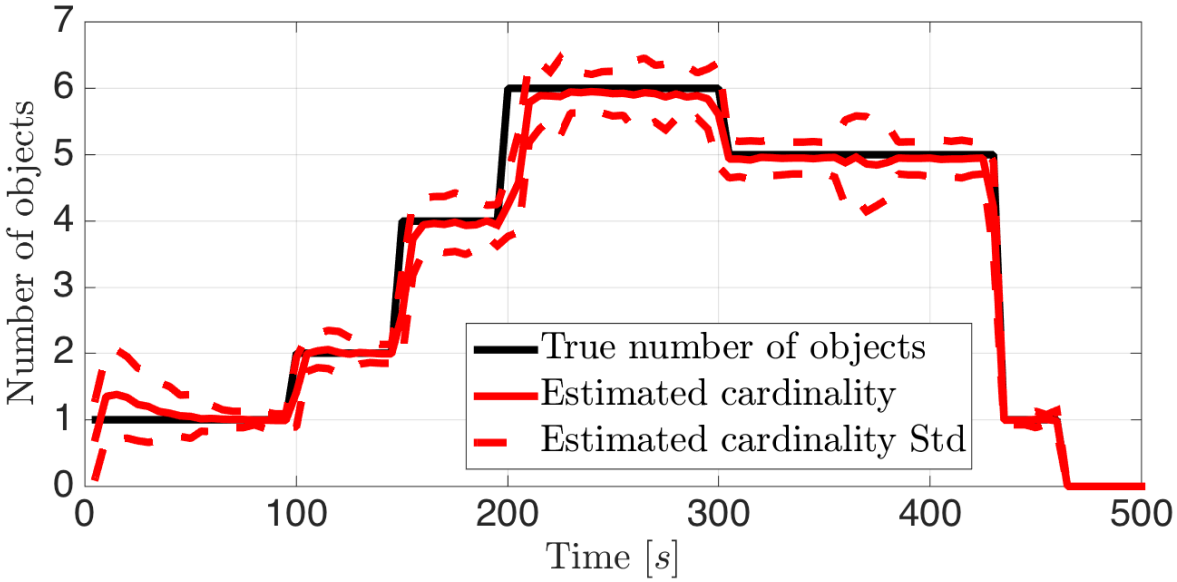


Figure 4.4: Cardinality statistics for GGM-CPHD.

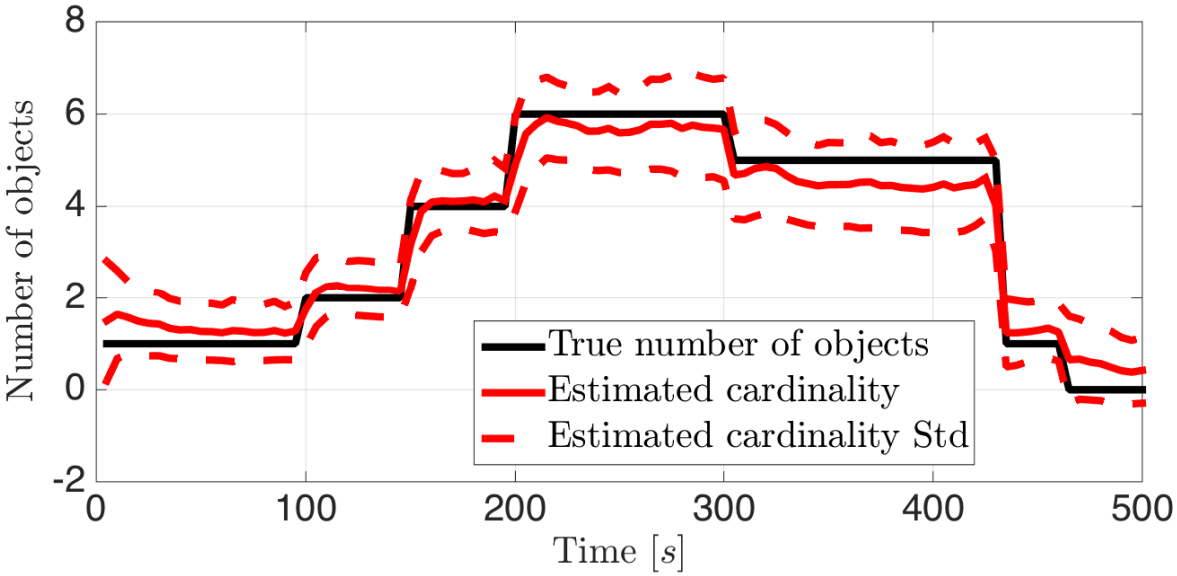


Figure 4.5: Cardinality statistics for CGM-CPHD with $L = 1$ consensus step.

In order to have an idea about the data communication requirements of the proposed algorithm, it is useful to examine the statistics of the number of Gaussian components of each node just before carrying out a consensus step; this determines, in fact, the number of data to be exchanged between nodes. To this end, Table 4.5 reports the average, standard deviation, minimum and maximum number of components in each node and at each consensus step during the simulation. Notice that, in the considered case-study, the average number of Gaussian components to be transmitted is about $20 \div 30$. Since each component involves 15 real numbers (4 for the mean, 10 for the covariance and 1 for the weight), the overall average

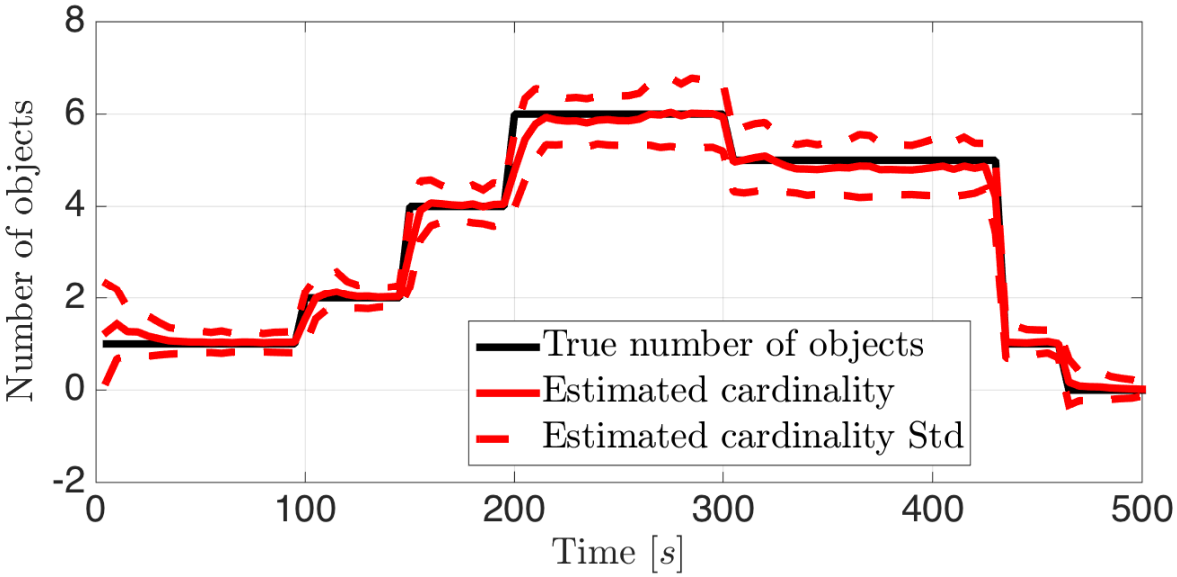


Figure 4.6: Cardinality statistics for CGM-CPHD with $L = 2$ consensus steps.

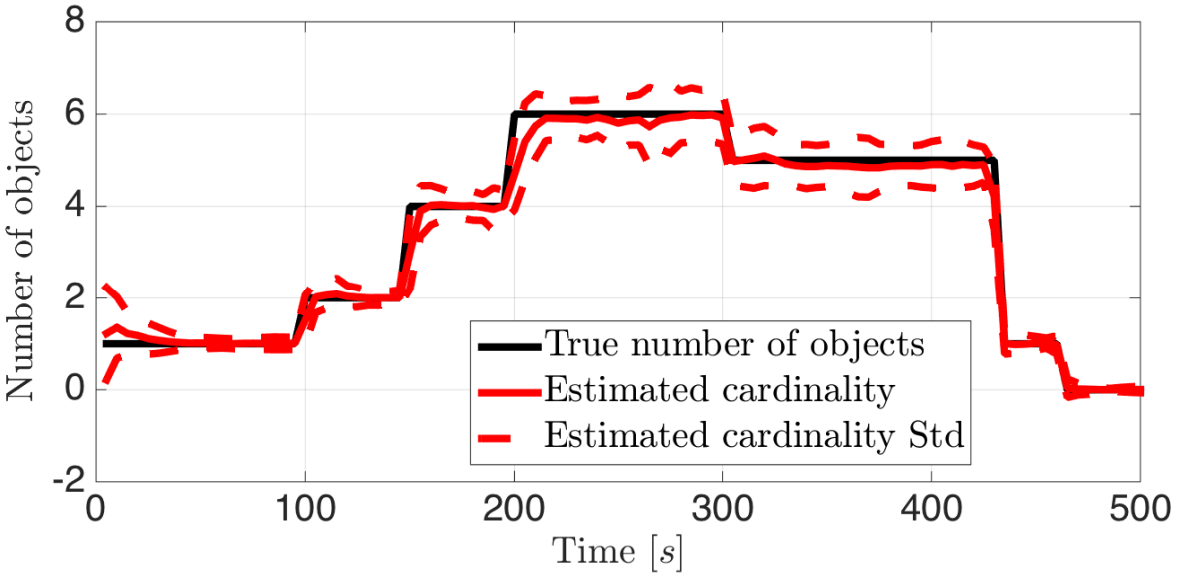


Figure 4.7: Cardinality statistics for CGM-CPHD with $L = 3$ consensus steps.

communication load per consensus step is about $1.2 \div 1.8$ kbytes if a 4-byte single-precision floating-point representation is adopted for each real number.

To give a rough idea of the computation time, our MATLAB implementation on a Personal Computer with 4.2 GHz clock exhibited an average processing time per consensus step in the range of $45 \div 55$ ms. Clearly, this time can be significantly reduced with a different, e.g. C/C++, implementation.

The CCPHD filter has also been successfully applied in scenarios with range and/or Doppler sensors [BCF+13b] and with passive multi-receiver radar systems [BCF+13c].

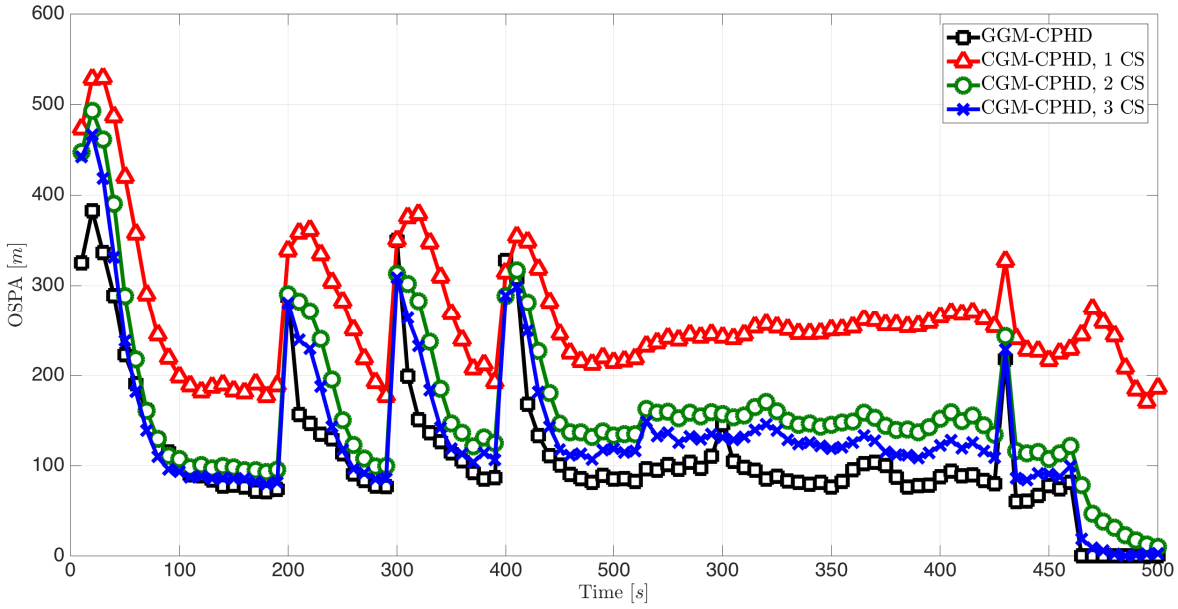


Figure 4.8: Performance comparison, using OSPA, between GGM-CPHD and CGM-CPHD respectively with $L = 1$, $L = 2$ and $L = 3$ consensus steps.

Table 4.5: Number of Gaussian components before consensus for CGM-CPHD

$\mathbf{TOA_1}$	AVG	STD	MAX
$L = 1$	29.39	12.98	63
$L = 2$	10.63	6.97	48
$L = 3$	5.18	4.58	48

$\mathbf{TOA_2}$	AVG	STD	MAX
$L = 1$	29.34	13.17	66
$L = 2$	10.33	7.11	58
$L = 3$	4.61	4.49	58

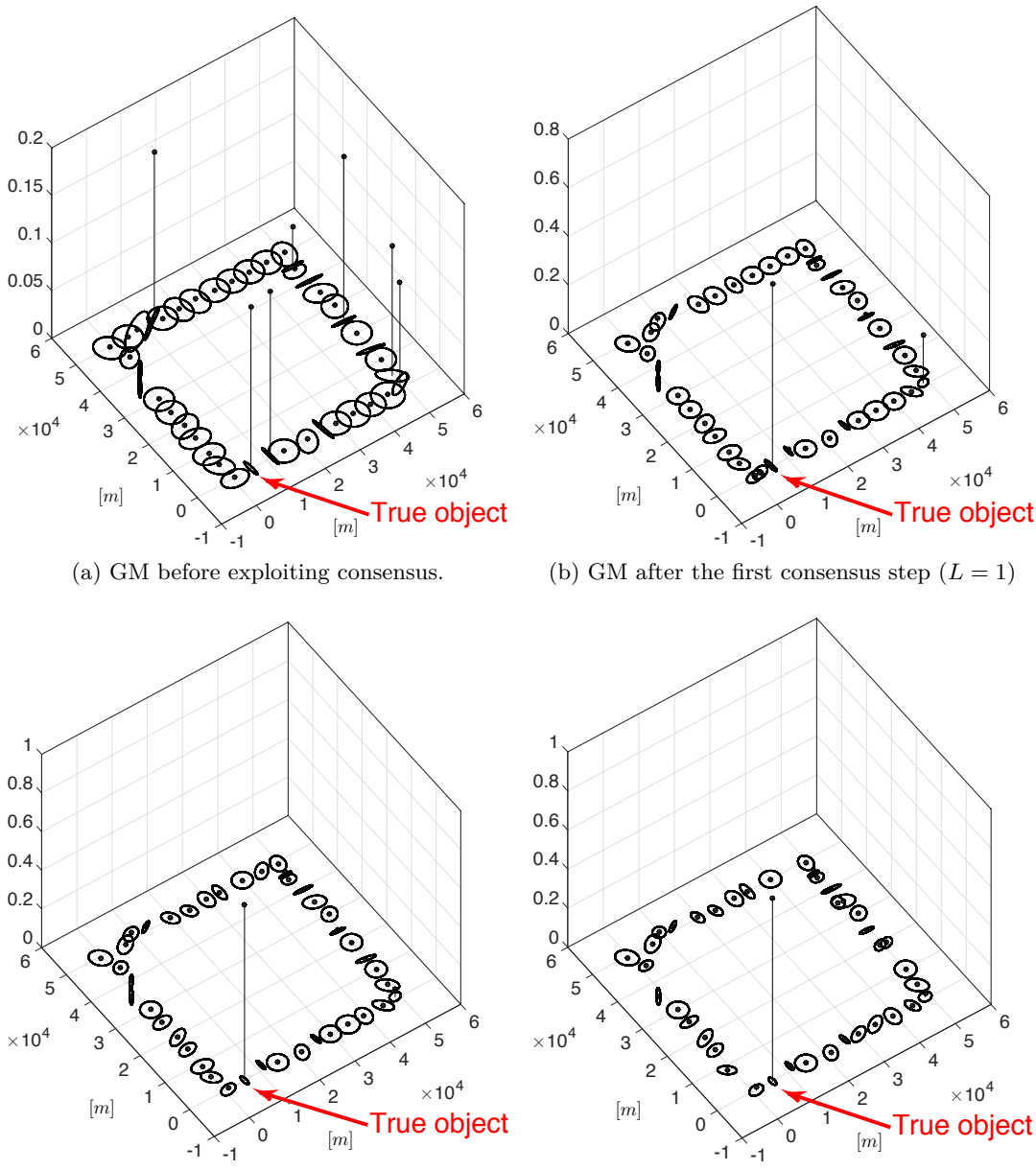
$\mathbf{TOA_3}$	AVG	STD	MAX
$L = 1$	29.08	13.13	63
$L = 2$	10.42	7.06	58
$L = 3$	4.7	4.49	57

$\mathbf{TOA_4}$	AVG	STD	MAX
$L = 1$	29.72	13.06	67
$L = 2$	10.79	6.97	52
$L = 3$	5.25	4.6	52

$\mathbf{DOA_1}$	AVG	STD	MAX
$L = 1$	24.12	15	52
$L = 2$	5.65	4.96	53
$L = 3$	4.22	4.25	51

$\mathbf{DOA_2}$	AVG	STD	MAX
$L = 1$	23.61	15.32	53
$L = 2$	5.6	4.94	50
$L = 3$	4.19	4.24	48

$\mathbf{DOA_3}$	AVG	STD	MAX
$L = 1$	23.45	15.23	52
$L = 2$	4.92	4.78	58
$L = 3$	4.22	4.35	60



(a) GM before exploiting consensus.

(b) GM after the first consensus step ($L = 1$)

(c) GM after the second consensus step ($L = 2$)

(d) GM after the third consensus step ($L = 3$)

Figure 4.9: GM representation within TOA sensor 1 (see Fig. 2) at the initial time instant throughout the consensus process. Means and 99.7% confidence ellipses of the Gaussian components of the location PDF are displayed in the horizontal plane while weights of such components are on the vertical axis.

5 Centralized multi-object tracking

Contents

In this chapter a new approximation of the δ-GLMB filter is presented [FVPV15]. It will be shown that the GLMB distribution can be used to construct a principled approximation to an arbitrary labeled RFS density that matches the PHD and the cardinality distribution [PVV$^+$14]. The resulting filter is referred to as the Marginalized δ-GLMB (Mδ-GLMB) tracking filter since it can be interpreted as a *marginalization over the data associations*. The proposed filter is consequently computationally cheaper than the δ-GLMB filter while still preserving key summary statistics of the multi-object posterior. Importantly, the Mδ-GLMB filter facilitates tractable multi-sensor multi-object tracking. Unlike PHD [VM06], CPHD [VVC07] and Multi-Bernoulli based filters [VVC09], the proposed approximation accommodates statistical dependence between objects. An alternative derivation of the LMB filter proposed in [RVVD14] based on the newly proposed Mδ-GLMB filter is presented [FVPV15].

5.1 The δ-GLMB filter

An efficient approach to multi-object tracking was presented in [VV13] using the δ-GLMB distribution (2.86) (or equivalently (2.87)), i.e.

$$\boldsymbol{\pi}(\mathbf{X}) = \Delta(\mathbf{X}) \sum_{(I,\xi)\in\mathcal{F}(\mathbb{L})\times\Xi} w^{(I,\xi)} \delta_I\left(\mathcal{L}\left(\mathbf{X}\right)\right) \left[p^{(\xi)}\right]^{\mathbf{X}}, \tag{5.1}$$

$$= \Delta(\mathbf{X}) \sum_{I\in\mathcal{F}(\mathbb{L})} \delta_I\left(\mathcal{L}\left(\mathbf{X}\right)\right) \sum_{\xi\in\Xi} w^{(I,\xi)} \left[p^{(\xi)}\right]^{\mathbf{X}}. \tag{5.2}$$

The δ-GLMB density naturally arises in multi-object tracking problems when using the standard detection based measurement model . In the following, the prediction and update steps of the δ-GLMB filter is briefly recalled; details can be found in [VV13, VVP14].

5.1.1 δ-GLMB prediction

The standard multi-object dynamic model is described as follows. Given the multi-object state $\mathbf{Z}$, each state $(\zeta,\ell_-)\in\mathbf{Z}$ either continues to exist at the next time step with probability $P_S(\zeta,\ell_-)$ and evolves to a new state (x,ℓ) with probability density $\varphi_{k|k-1}(x|\zeta,\ell_-)$, or disappears with probability $1-P_S(\zeta,\ell_-)$. Note that the label of the objects is preserved in the transition, only the kinematic part of state changes. Assuming that $\mathbf{Z}$ has distinct labels and that conditional on $\mathbf{Z}$, the transition of the kinematic states are mutually independent, then the set $\mathbf{W}$ of surviving objects at the next time is a labeled multi-Bernoulli RFS [VV13]

$$\boldsymbol{f}_S(\mathbf{W}|\mathbf{Z}) = \Delta(\mathbf{W})\,\Delta(\mathbf{Z})\; 1_{\mathcal{L}(\mathbf{Z})}\left(\mathcal{L}(\mathbf{W})\right)\left[\Phi\left(\mathbf{W};\cdot\right)\right]^{\mathbf{Z}}, \tag{5.3}$$

where

$$\Phi(\mathbf{W};\zeta,\ell_-) = \sum_{(x,\ell)\in\mathbf{W}} \delta_{\ell_-}\left(\ell\right) P_S\left(\zeta,\ell_-\right) f\left(x|\zeta,\ell_-\right) + \left[1-1_{\mathcal{L}(\mathbf{W})}\left(\ell_-\right)\right]\left(1-P_S\left(\zeta,\ell_-\right)\right). \tag{5.4}$$

The $\Delta(\mathbf{Z})$ in (5.3) ensures that only $\mathbf{Z}$ with distinct labels are considered.

The set of new objects born at the next time step is distributed according to

$$\boldsymbol{f}_B(\mathbf{Y}) = \Delta(\mathbf{Y})\, w_B(\mathcal{L}(\mathbf{Y}))\left[p_B\right]^{\mathbf{Y}} \tag{5.5}$$

The birth density $\boldsymbol{f}_B(\cdot)$ is defined on $\mathbb{X}\times\mathbb{L}$ and $\boldsymbol{f}_B(\mathbf{Y})=0$ if $\mathbf{Y}$ contains any element $\mathbf{x}$ with $\mathcal{L}(\mathbf{x})\notin\mathbb{L}$. The birth model (5.5) includes both labeled Poisson and labeled multi-Bernoulli densities. The multi-object state at the next time $\mathbf{X}$ is the superposition of surviving objects and new born objects, i.e. $\mathbf{X}=\mathbf{W}\cup\mathbf{Y}$. Since the label spaces $\mathbb{L}$ and $\mathbb{B}$ are disjoint, the labeled birth objects and surviving objects are independent. Thus the multi-object transition density turns out to be the product of the transition density (5.3) and the density of new

objects (5.5)

$$\boldsymbol{f}(\mathbf{X}|\mathbf{Z}) = \boldsymbol{f}_S(\mathbf{X} \cap (\mathbb{X} \times \mathbb{L})\,|\mathbf{Z})\ \boldsymbol{f}_B\left(\mathbf{X} - (\mathbb{X} \times \mathbb{L})\right). \tag{5.6}$$

Additional details can be found in [VV13, Subsection IV.D].

If the current multi-object prior density is a δ-GLMB of the form (2.86), then the multi-object prediction density is also a δ-GLMB given by

$$\boldsymbol{\pi}_{k|k-1}(\mathbf{X}) = \Delta(\mathbf{X}) \sum_{(I,\xi)\in\mathcal{F}(\mathbb{L})\times\Xi} w_{k|k-1}^{(I,\xi)} \delta_I\left(\mathcal{L}(\mathbf{X})\right) \left[p_{k|k-1}^{(\xi)}\right]^{\mathbf{X}} \tag{5.7}$$

where

$$w_{k|k-1}^{(I,\xi)} = w_B(I\backslash\mathbb{L}_-) w_S^{(\xi)}(I \cap \mathbb{L}_-)\,, \tag{5.8}$$

$$p_{k|k-1}^{(\xi)}(x,\ell) = 1_{\mathbb{L}_-}(\ell) p_S^{(\xi)}(x,\ell) + 1_{\mathbb{B}}(\ell) p_B(x,\ell)\,, \tag{5.9}$$

$$p_S^{(\xi)}(x,\ell) = \frac{\left\langle P_S(\cdot,\ell)\,\varphi_{k|k-1}(x|\cdot,\ell), p_{k-1}^{(\xi)}(\cdot,\ell)\right\rangle}{\eta_S^{(\xi)}(\ell)}\,, \tag{5.10}$$

$$\eta_S^{(\xi)}(\ell) = \left\langle P_S(\cdot,\ell), p_{k-1}^{(\xi)}(\cdot,\ell)\right\rangle\,, \tag{5.11}$$

$$w_S^{(\xi)}(L) = [\eta_S^{(\xi)}]^L \sum_{J\subseteq\mathbb{L}_-} 1_J(L)[1-\eta_S^{(\xi)}]^{J-L} w_k^{(J,\xi)}\,. \tag{5.12}$$

5.1.2 δ-GLMB update

The standard multi-object observation model is described as follows. For a given multi-object state $\mathbf{X}$, each state $\mathbf{x} \in \mathbf{X}$ is either detected with probability $P_D(\mathbf{x})$ and generates a point y with likelihood $g(y|\mathbf{x})$, or missed with probability $1 - P_D(\mathbf{x})$, i.e. $\mathbf{x}$ generates a Bernoulli RFS with parameter $(P_D(\mathbf{x}), g(\cdot|\mathbf{x}))$. Assuming that conditional on $\mathbf{X}$ these Bernoulli RFSs are independent, then the set $W \subset \mathbb{Y}$ of detected points (non-clutter measurements) is a multi-Bernoulli RFS with parameter set $\{(P_D(\mathbf{x}), g(\cdot|\mathbf{x})) : \mathbf{x} \in \mathbf{X}\}$. The set $\mathcal{C} \subset \mathbb{Y}$ of false observations (or clutter), assumed independent of the detected points, is modeled by a Poisson RFS with intensity function $\kappa(\cdot)$. The multi-object observation Y is the superposition of the detected points and false observations, i.e. $Y = W \cup \mathcal{C}$. Assuming that, conditional on $\mathbf{X}$, detections are independent, and that clutter is independent of the detections, the multi-object likelihood is given by

$$g_k(Y|\mathbf{X}) = e^{-\langle\kappa,1\rangle}\kappa^Y \sum_{\theta\in\Theta(\mathcal{L}(\mathbf{X}))} [\psi_Y(\cdot;\theta)]^{\mathbf{X}}\,, \tag{5.13}$$

where $\Theta(I)$ is the set of mappings $\theta : I \to \{0,1,...,M\}$, such that $\theta(i) = \theta(i') > 0$ implies $i = i'$, and

$$\psi_Y(x,\ell;\theta) = \begin{cases} \dfrac{P_D(x,\ell)\, g_k(y_{\theta(\ell)}|x,\ell)}{\kappa(y_{\theta(\ell)})}\,, & \text{if } \theta(\ell) > 0 \\ 1 - P_D(x,\ell)\,, & \text{if } \theta(\ell) = 0 \end{cases}. \tag{5.14}$$

Note that an association map θ specifies which tracks generated which measurements, i.e. track ℓ generates measurement $y_{\theta(\ell)} \in Y$, with undetected tracks assigned to 0. The condition "$\theta(i) = \theta(i') > 0$ implies $i = i'$", means that a track can generate at most one measurement, and a measurement can be assigned to at most one track, at one time instant. Additional details can be found in [VV13, Subsection IV.C].

If the current multi-object prediction density is also a δ-GLMB of the form (2.86), then the multi-object posterior density is a δ-GLMB given by

$$\boldsymbol{\pi}_k(\mathbf{X}) = \Delta(\mathbf{X}) \sum_{(I,\xi)\in\mathcal{F}(\mathbb{L})\times\Xi} \sum_{\theta\in\Theta(I)} w_k^{(I,\xi,\theta)}(Y_k)\delta_I\left(\mathcal{L}(\mathbf{X})\right)\left[p_k^{(\xi,\theta)}\right]^{\mathbf{X}} \tag{5.15}$$

where $\Theta(I)$ denotes the subset of the current maps with domain I, and

$$w_k^{(I,\xi,\theta)}(Y_k) \propto w_{k|k-1}^{(I,\xi)}\left[\eta_{Y_k}^{(\xi,\theta)}(\ell)\right]^I , \tag{5.16}$$

$$\eta_{Y_k}^{(\xi,\theta)}(\ell) = \left\langle p_{k|k-1}^{(\xi)}(\cdot,\ell), \psi_{Y_k}(\cdot,\ell;\theta)\right\rangle , \tag{5.17}$$

$$p_k^{(\xi,\theta)}(x,\ell) = \frac{p_{k|k-1}^{(\xi)}(x,\ell)\psi_{Y_k}(x,\ell;\theta)}{\eta_{Y_k}^{(\xi,\theta)}(\ell)} , \tag{5.18}$$

$$\psi_{Y_k}(x,\ell;\theta) = \begin{cases} \dfrac{P_D(x,\ell)\, g(y_{\theta(\ell)}|x,\ell)}{\kappa(y_{\theta(\ell)})}, & \text{if } \theta(\ell) > 0 \\ 1 - P_D(x,\ell), & \text{if } \theta(\ell) = 0 \end{cases} . \tag{5.19}$$

Notice that the new association maps θ can be added (stacked) to their respective association histories ξ in order to have again the more compact form (2.86) for the updated δ-GLMB (5.15).

5.2 GLMB approximation of multi-object densities

In this section a GLMB approximation of the multi-object density with statistically dependent objects is proposed. In particular a GLMB density that matches the multi-object density of interest in both the PHD and cardinality distribution is derived. The strategy is inspired by Mahler's iid cluster approximation in the CPHD filter [Mah07a], which has proven to be very effective in practical multi-object filtering problems [Mah14, GSW09, SWS09]. Moreover, the GLMB approximation captures object dependencies whereas the iid cluster approximation assumes that individual objects are iid from a common single-object density. Proof of the result is given in the appendix A.

Our result follows from the observation that any labeled RFS density $\boldsymbol{\pi}(\cdot)$ on $\mathcal{F}(\mathbb{X}\times\mathbb{L})$ can be written as

$$\boldsymbol{\pi}(\mathbf{X}) = w(\mathcal{L}(\mathbf{X}))p(\mathbf{X}) \tag{5.20}$$

where

$$w(\{\ell_1,\ldots,\ell_n\}) \triangleq \int \boldsymbol{\pi}(\{(x_1,\ell_1),\ldots,(x_n,\ell_n)\})dx_1\cdots dx_n \tag{5.21}$$

$$p(\{(x_1,\ell_1),\ldots,(x_n,\ell_n)\}) \triangleq \frac{\boldsymbol{\pi}(\{(x_1,\ell_1),\ldots,(x_n,\ell_n)\})}{w(\{\ell_1,\ldots,\ell_n\})} \tag{5.22}$$

It is implicitly assumed that $p(\mathbf{X})$ is defined to be zero whenever $w(\mathcal{L}(\mathbf{X}))$ is zero.

Note that since $\boldsymbol{\pi}(\cdot)$ is symmetric in its arguments, the integral in (5.21) and w is symmetric in $\ell_1,\ldots,\ell_n$. Moreover, since $\sum_{L\in\mathcal{F}(\mathbb{L})} w(L) = 1$, w is indeed a probability distribution on $\mathcal{F}(\mathbb{L})$ and can be interpreted as the probability that the labeled RFS has label set $\{\ell_1,\ldots,\ell_n\}$. For $w(\{\ell_1,\ldots,\ell_n\}) > 0$, $p(\{(x_1,\ell_1),\ldots,(x_n,\ell_n)\})$ is the joint probability density (on $\mathbb{X}^n$) of the kinematic states $x_1,\ldots,x_n$ given that their corresponding labels are $\ell_1,\ldots,\ell_n$.

Proposition 1 (GLMB approximation)**.**
The GLMB that matches the cardinality distribution and the PHD of a labeled RFS with density $\boldsymbol{\pi}(\cdot)$ is given by:

$$\hat{\boldsymbol{\pi}}(\mathbf{X}) = \sum_{I\in\mathcal{F}(\mathbb{L})} w^{(I)}(\mathcal{L}(\mathbf{X}))\left[p^{(I)}\right]^{\mathbf{X}} \tag{5.23}$$

where,

$$w^{(I)}(L) = \delta_I(L)w(I) \tag{5.24}$$

$$p^{(I)}(x,\ell) = 1_I(\ell)p_{I\setminus\{\ell\}}(x,\ell) \tag{5.25}$$

$$p_{\{\ell_1,\ldots,\ell_n\}}(x,\ell) = \int p(\{(x,\ell),(x_1,\ell_1),\ldots,(x_n,\ell_n)\})d(x_1,\ldots,x_n) \tag{5.26}$$

Remark 13. Note that in [BVV14, Section V] a GLMB was proposed to approximate a particular family of labeled RFS densities that arises from multi-object filtering with merged measurements. By applying Proposition 1, it can be shown that the approximation used in [BVV14, Section V] preserves the cardinality distribution and first moment.

Remark 14. In multi-object tracking, the matching of the labeled PHDs $\hat{d}(\cdot,\ell)$ and $d(\cdot,\ell)$ in Proposition 1 is a stronger result than simply matching the unlabeled PHDs alone.

5.3 Marginalizations of the δ-GLMB density

One of the main factors increasing the computational burden of the δ-GLMB filter [VVP14] is the exponential growth of the number of hypotheses in the update of the prior (5.15), which gives rise to the explicit sum over an association history variable ξ. Further, note that the number of association histories is considerably increased in a multi-sensor context. The idea behind the Mδ-GLMB filter is to construct a principled GLMB approximation $\hat{\boldsymbol{\pi}}(\cdot)$ to the posterior density $\boldsymbol{\pi}(\cdot)$ by means of a marginalization with respect to the association histories ξ. In this section, the time index is dropped for the sake of simplicity and the attention is devoted to the Mδ-GLMB approximation corresponding to any δ-GLMB density.

Proposition 2 (The Mδ-GLMB density)**.**
A Marginalized δ-GLMB (Mδ-GLMB) density $\hat{\boldsymbol{\pi}}(\cdot)$ corresponding to the δ-GLMB density $\boldsymbol{\pi}(\cdot)$ in (2.86) is a probability density of the form

$$\hat{\boldsymbol{\pi}}(\mathbf{X}) = \Delta(\mathbf{X}) \sum_{I \in \mathcal{F}(\mathbb{L})} \delta_I(\mathcal{L}(\mathbf{X})) w^{(I)} \left[p^{(I)}\right]^{\mathbf{X}}, \tag{5.27}$$

$$w^{(I)} = \sum_{\xi \in \Xi} w^{(I,\xi)}, \tag{5.28}$$

$$p^{(I)}(x,\ell) = 1_I(\ell) \frac{1}{w^{(I)}} \sum_{\xi \in \Xi} w^{(I,\xi)} p^{(\xi)}(x,\ell). \tag{5.29}$$

The Mδ-GLMB (5.27) preserves both (labeled) PHD and cardinality distribution of the original δ-GLMB density $\boldsymbol{\pi}(\cdot)$ in (2.86).

Note now that the Mδ-GLMB density provided by Proposition 2 is a δ-GLMB which aggregates the association histories ξ relative to the label set I via (5.28)-(5.29). In fact, the marginalized density (5.27) has no association history set Ξ.

The Mδ-GLMB density can be exploited to construct an efficient recursive multi-object tracking filter by calculating the Mδ-GLMB approximation step after the δ-GLMB update, and predicting forward in time using the δ-GLMB prediction.

5.4 The Mδ-GLMB filter

The Mδ-GLMB filter propagates a Mδ-GLMB multi-object posterior density forward in time via the multi-object Bayesian recursion (2.97)-(2.98). Since the Mδ-GLMB is still a δ-GLMB, the equations derived for the δ-GLMB filter [VVP14] as a closed-form solution for the eqs. (2.97)-(2.98) still hold. In the following, the prediction and update steps for the δ-GLMB filter are briefly recalled; additional details can be found in [VVP14, FVPV15].

5.4.1 Mδ-GLMB prediction

Given the multi-object state $\mathbf{Z}$, each state $(\zeta, \ell_-) \in \mathbf{Z}$ either continues to exist at the next time step with probability $P_S(\zeta, \ell_-)$ and evolves to a new state (x, ℓ) with probability density $\varphi_{k|k-1}(x|\zeta, \ell_-)$, or dies with probability $1 - P_S(\zeta, \ell_-)$. The set of newborn objects at the next time step is distributed according to the LMB

$$\boldsymbol{f}_B = \left\{ \left(r_B^{(\ell)}, p_B^{(\ell)} \right) \right\}_{\ell \in \mathbb{B}} \tag{5.30}$$

with $\mathbb{L}_- \cap \mathbb{B} = \varnothing$ and $w_B(\mathbf{X})$ defined in (2.91). The multi-object state at the current time $\mathbf{X}_k$ is the superposition of surviving objects and new born objects.

If the current multi-object prior density is a Mδ-GLMB of the form (5.27), then the

multi-object prediction density is also a Mδ-GLMB given by

$$\boldsymbol{\pi}_{k|k-1}(\mathbf{X}) = \Delta(\mathbf{X}) \sum_{I \in \mathcal{F}(\mathbb{L})} \delta_I\left(\mathcal{L}\left(\mathbf{X}\right)\right) w_{k|k-1}^{(I)} \left[p_{k|k-1}^{(I)}\right]^{\mathbf{X}}, \tag{5.31}$$

where

$$w_{k|k-1}^{(I)} = w_B(I \backslash \mathbb{L}_-)\, w_S^{(I)}(I \cap \mathbb{L}_-)\,, \tag{5.32}$$

$$p_{k|k-1}^{(I)}(x,\ell) = 1_{\mathbb{L}_-}(\ell) p_S^{(I)}(x,\ell) + 1_{\mathbb{B}}(\ell) p_B(x,\ell)\,, \tag{5.33}$$

$$p_S^{(I)}(x,\ell) = \frac{\left\langle P_S(\cdot,\ell)\varphi_{k|k-1}(x|\cdot,\ell), p_{k-1}^{(I)}(\cdot,\ell)\right\rangle}{\eta_S^{(I)}(\ell)}\,, \tag{5.34}$$

$$\eta_S^{(I)}(\ell) = \left\langle P_S(\cdot,\ell), p_{k-1}^{(I)}(\cdot,\ell)\right\rangle\,, \tag{5.35}$$

$$w_S^{(I)}(L) = [\eta_S^{(I)}]^L \sum_{J \subseteq \mathbb{L}_-} 1_J(L)[1-\eta_S^{(I)}]^{J-L} w_k^{(J)}\,, \tag{5.36}$$

which is exactly the δ-GLMB prediction step (5.7) with no association histories from previous time step, i.e. $\Xi = \varnothing$, and with the convention of having the superscript (I) instead of (ξ) due to the marginalization (5.28)-(5.29).

Remark 15. The number of components $\left(w_{k+1|k}^{(I)}, p_{k+1|k}^{(I)}\right)$ computed after the Mδ-GLMB prediction step (5.31) is $|\mathcal{F}(\mathbb{L}_{0:k+1})|$. On the other hand, the number of components $\left(w_{k+1|k}^{(I,\xi)}, p_{k+1|k}^{(\xi)}\right)$ after the δ-GLMB prediction (5.7) is $|\mathcal{F}(\mathbb{L}_{0:k+1}) \times \Xi|$ for $w_{k+1|k}^{(I,\xi)}$ and $|\Xi|$ for $p_{k+1|k}^{(\xi)}$. Notice that the number of weights $w_{k+1|k}^{(I)}$ of the Mδ-GLMB is significantly lower than the $w_{k+1|k}^{(I,\xi)}$ of δ-GLMB. As for the number of location PDFs $p_{k+1|k}^{(\cdot)}$, it is worth noticing that the growth rate of the association histories $\xi \in \Xi$ is super-exponential with time [VV13, VVP14], while the growth rate of the cardinality of $\mathcal{F}(\mathbb{L}_{0:k+1})$ is exponential.

The use of the Mδ-GLMB approximation further reduces the number of hypotheses in the posterior density while preserving the PHD and cardinality distribution. Moreover, the Mδ-GLMB is in a form that is suitable for efficient and tractable information fusion (i.e. multi-sensor processing) as will be shown in the next subsection.

5.4.2 Mδ-GLMB update

Given a multi-object state $\mathbf{X}$, each state $(x,\ell) \in \mathbf{X}$ is either detected with probability $P_D(x,\ell)$ and generates a measurement y with likelihood $g_k(y|x,\ell)$, or missed with probability $1 - P_D(x,\ell)$. The multi-object observation $Y = \{y_1, \ldots, y_M\}$ is the superposition of the detected points and Poisson clutter with intensity function $\kappa(\cdot)$.

If the current multi-object prediction density is a Mδ-GLMB of the form (5.27), then the multi-object posterior density is a δ-GLMB given by

$$\boldsymbol{\pi}_k(\mathbf{X}) = \Delta(\mathbf{X}) \sum_{I \in \mathcal{F}(\mathbb{L})} \sum_{\theta \in \Theta(I)} \delta_I\left(\mathcal{L}\left(\mathbf{X}\right)\right) w_k^{(I,\theta)}(Y_k) \left[p_k^{(I,\theta)}\right]^{\mathbf{X}}, \tag{5.37}$$

where $\Theta(I)$ denotes the subset of the current maps with domain I, and

$$w_k^{(I,\theta)}(Y_k) \propto w_{k|k-1}^{(I)} \left[\eta_{Y_k}^{(\theta)}(\ell)\right]^I , \tag{5.38}$$

$$\eta_{Y_k}^{(I,\theta)}(\ell) = \left\langle p_{k|k-1}^{(I)}(\cdot,\ell), \psi_{Y_k}(\cdot,\ell;\theta) \right\rangle , \tag{5.39}$$

$$p_k^{(I,\theta)}(x) = \frac{p_{k|k-1}^{(I)}(x,\ell)\psi_{Y_k}(x,\ell;\theta)}{\eta_{Y_k}^{(I,\theta)}(\ell)} , \tag{5.40}$$

$$\psi_{Y_k}(x,\ell;\theta) = \begin{cases} \dfrac{P_D(x,\ell)g(y_{\theta(\ell)}|x,\ell)}{\kappa(y_{\theta(\ell)})}, & \text{if } \theta(\ell) > 0 \\ 1 - P_D(x,\ell), & \text{if } \theta(\ell) = 0 \end{cases} . \tag{5.41}$$

Using (5.28)-(5.29), the Mδ-GLMB density corresponding to the δ-GLMB density in (5.37) is a probability density of the form (5.27) with

$$w_k^{(I)} = \sum_{\theta\in\Theta(I)} w_k^{(I,\theta)} , \tag{5.42}$$

$$p_k^{(I)}(x,\ell) = 1_I(\ell)\frac{1}{w_k^{(I)}} \sum_{\theta\in\Theta(I)} w_k^{(I,\theta)} p_k^{(I,\theta)}(x,\ell) . \tag{5.43}$$

The Mδ-GLMB density provided by (5.42)-(5.43) preserves both PHD and cardinality distribution of the original δ-GLMB density.

Remark 16. Each hypothesis $I \in \mathcal{F}(\mathbb{L})$ generates a set of $|\Theta(I)|$ new measurement-to-track association maps for the δ-GLMB posterior. The number of components $\left(w_{k+1|k}^{(I,\theta)}, p_{k+1|k}^{(I,\theta)}\right)$ stored/computed after the Mδ-GLMB update step (5.37) is $|\mathcal{F}(\mathbb{L})| \cdot \sum_{I\in\mathcal{F}(\mathbb{L})} |\Theta(I)|$. On the other hand, the number of hypotheses $\left(w_{k+1|k}^{(I,\xi,\theta)}, p_{k+1|k}^{(\xi,\theta)}\right)$ after the δ-GLMB update (5.15) is $|\mathcal{F}(\mathbb{L}) \times \Xi| \cdot \sum_{I\in\mathcal{F}(\mathbb{L})} |\Theta(I)|$ for $w_{k+1|k}^{(I,\xi,\theta)}$ and $|\Xi| \cdot \sum_{I\in\mathcal{F}(\mathbb{L})} |\Theta(I)|$ for $p_{k+1|k}^{(\xi,\theta)}$. The same conclusions along the lines of Remark 15 hold.

Remark 17. After the marginalization procedure (5.42)-(5.43) only $|\mathcal{F}(\mathbb{L})|$ hypotheses are retained, as all the new contributions provided by the association maps $|\Theta(I)|$ are aggregated in a single component. Notice that $|\mathcal{F}(\mathbb{L})|$ is the same number of hypotheses produced during the prediction step (5.31) (see Remark 15). Thus, the prediction step (5.37) sets the upper bound of the hypotheses that will be retained after each full Mδ-GLMB step.

In the multi-sensor setting, for each time instant k, the update step (5.37) (or (5.15) for the δ-GLMB) is repeatedly evaluated using different measurement sets Y^i provided by the sensors $i \in \mathcal{N}$ (see eq. (2.101)). From Remarks 16 and 17, Mδ-GLMB is preferable to δ-GLMB in terms of memory and computational requirements, since the number of remaining hypotheses after each sensor update step in (2.101) is always set to $|\mathcal{F}(\mathbb{L})|$. Note that this does not apply to δ-GLMB due to the super-exponential memory/computational growth as reported in Remark 15. This is an important property of Mδ-GLMB since it yields a principled approximation which greatly decreases the need of pruning hypotheses with respect

to δ-GLMB [VVP14]. In fact, pruning in the δ-GLMB might lead to poor performance in multi-sensor scenarios with low SNR (e.g. high clutter intensity, low detection probability, etc.) and limited storage/computational capabilities. For instance, this may happen if a subset of the sensors do not detect one or multiple objects and hypotheses associated to the true tracks are removed due to pruning. Furthermore, from a mathematical viewpoint, pruning between corrections generally produces a less informative and order-independent approximation of the posterior distribution in (2.101).

5.5 The LMB filter

The LMB filter introduced in [RVVD14] is a single component approximation of a GLMB density (2.77) that matches the unlabeled PHD. An alternative derivation of the LMB approximation (2.90) first proposed in [RVVD14] through a connection with the Mδ-GLMB approximation is provided hereafter.

5.5.1 Alternative derivation of the LMB Filter

Recall that a LMB density is uniquely parameterized by a set of existence probabilities $r^{(\ell)}$ and corresponding track densities $p^{(\ell)}(\cdot)$:

$$\boldsymbol{\pi}(\mathbf{X}) = \Delta(\mathbf{X})\ w(\mathcal{L}(\mathbf{X}))p^{\mathbf{X}} \tag{5.44}$$

where

$$w(L) = \prod_{\imath\in\mathbb{L}} \left(1 - r^{(\imath)}\right) \prod_{\ell\in L} \frac{1_{\mathbb{L}}(\ell) r^{(\ell)}}{1 - r^{(\ell)}} \tag{5.45}$$

$$p(x,\ell) = p^{(\ell)}(x) \tag{5.46}$$

By extracting individual tracks from the Mδ-GLMB approximation $\hat{\boldsymbol{\pi}}(\cdot)$ (5.27) it is possible to prove that the same expressions for the existence probabilities and state densities originally proposed for the LMB filter in [RVVD14] can be obtained. In fact

$$r^{(\ell)} \triangleq \Pr_{\hat{\boldsymbol{\pi}}}(\ell \in \mathcal{L}(\mathbf{X})) \tag{5.47}$$

$$= \sum_{\ell\in L} w(L) \tag{5.48}$$

$$= \sum_{I\in\mathcal{F}(\mathbb{L})} 1_I(\ell) w(L) \tag{5.49}$$

$$= \sum_{I\in\mathcal{F}(\mathbb{L})} 1_I(\ell) w^{(I)}\,, \tag{5.50}$$

$$= \sum_{(I,\xi)\in\mathcal{F}(\mathbb{L})\times\Xi} 1_I(\ell) w^{(I,\xi)}\,, \tag{5.51}$$

$$p^{(\ell)}(x) \triangleq \frac{\Pr_{\hat{\boldsymbol{\pi}}}((x,\ell) \in \mathbf{X})}{\Pr_{\hat{\boldsymbol{\pi}}}(\ell \in \mathcal{L}(\mathbf{X}))} \tag{5.52}$$

$$= \frac{\int \hat{\boldsymbol{\pi}}((x,\ell) \cup \mathbf{X}) \delta\mathbf{X}}{r^{(\ell)}} \tag{5.53}$$

$$= \frac{\hat{d}(x,\ell)}{r^{(\ell)}}, \tag{5.54}$$

where the notation for the numerator in (5.52) is defined as per [Mah07b, eq. 11.111], while the numerator of (5.54) follows from [Mah07b, eq. 11.112]. Notice that the numerator is precisely the PHD $\hat{d}(\cdot)$ corresponding to $\hat{\boldsymbol{\pi}}(\cdot)$, which by Proposition 1 exactly matches the PHD $d(\cdot)$ corresponding to $\boldsymbol{\pi}(\cdot)$. It can be easily verified that

$$\hat{d}(x,\ell) = \sum_{I \in \mathcal{F}(\mathbb{L})} 1_I(\ell) w^{(I)} p^{(I)}(x,\ell) \tag{5.55}$$

and consequently

$$p^{(\ell)}(x) = \frac{1}{r^{(\ell)}} \sum_{I \in \mathcal{F}(\mathbb{L})} 1_I(\ell) w^{(I)} p^{(I)}(x,\ell) \tag{5.56}$$

$$= \frac{1}{r^{(\ell)}} \sum_{(I,\xi) \in \mathcal{F}(\mathbb{L}) \times \Xi} 1_I(\ell) w^{(I,\xi)} p^{(\xi)}(x,\ell) \tag{5.57}$$

Remark 18. The property of matching the labeled PHD of the δ-GLMB does not hold for the LMB tracking filter, as shown in [RVVD14, Section III], due to the imposed multi-Bernoulli structure for the cardinality distribution. This can be also proved by combining (5.44) with Proposition 1, thus showing again that only the unlabeled PHD of the LMB matches exactly the unlabeled PHD of the original δ-GLMB density.

As suggested by its name, the LMB filter propagates an LMB multi-object posterior density forward in time [RVVD14]. Its recursion will be briefly reviewed in the next subsections. Additional details on its implementation can be found in [RVVD14].

5.5.2 LMB prediction

Given the previous multi-object state $\mathbf{Z}$, each state $(\zeta, \ell_-) \in \mathbf{Z}$ either continues to exist at the next time step with probability $P_S(\zeta, \ell_-)$ and evolves to a new state (x,ℓ) with probability density $\varphi_{k|k-1}(x|\zeta,\ell_-)$, or dies with probability $1 - P_S(\zeta,\ell_-)$. The set of new objects born at the next time step is distributed according to

$$\boldsymbol{f}_B = \left\{ \left(r_B^{(\ell)}, p_B^{(\ell)} \right) \right\}_{\ell \in \mathbb{B}} \tag{5.58}$$

with $\mathbb{L}_- \cap \mathbb{B} = \varnothing$. The multi-object state at the current time $\mathbf{X}_k$ is the superposition of surviving objects and new born objects.

Suppose that the multi-object posterior density at time $k-1$ is LMB with parameter

set $\boldsymbol{\pi}_{k-1} = \left\{\left(r^{(\ell)}, p^{(\ell)}\right)\right\}_{\ell\in\mathbb{L}_-}$. Then, it has been shown in [RVVD14] that the multi-object predicted density is also LMB with parameter set

$$\boldsymbol{\pi}_{k|k-1} = \left\{\left(r_S^{(\ell)}, p_S^{(\ell)}\right)\right\}_{\ell\in\mathbb{L}_-} \cup \left\{\left(r_B^{(\ell)}, p_B^{(\ell)}\right)\right\}_{\ell\in\mathbb{B}}, \tag{5.59}$$

where

$$r_S^{(\ell)} = \eta_S(\ell)\, r^{(\ell)}, \tag{5.60}$$

$$p_S^{(\ell)}(x) = \frac{\left\langle P_S(\cdot,\ell)\varphi_{k|k-1}(x|\cdot,\ell), p(\cdot,\ell)\right\rangle}{\eta_S(\ell)}, \tag{5.61}$$

$$\eta_S(\ell) = \left\langle P_S(\cdot,\ell), p(\cdot,\ell)\right\rangle. \tag{5.62}$$

5.5.3 LMB update

Given a multi-object state $\mathbf{X}$, each state $(x,\ell)\in\mathbf{X}$ is either detected with probability $P_D(x,\ell)$ and generates a measurement y with likelihood $g_k(y|x,\ell)$, or missed with probability $1-P_D(x,\ell)$. The multi-object observation $Y=\{y_1,...,y_M\}$ is the superposition of the detected points and Poisson clutter with intensity function $\kappa(\cdot)$. Assuming that, conditional on $\mathbf{X}$, detections are independent, and that clutter is independent of the detections, the multi-object likelihood is the same of the Mδ-GLMB Update in (5.14).

Suppose that the multi-object predicted density is LMB with parameter set

$$\boldsymbol{\pi}_{k|k-1} = \left\{\left(r_{k|k-1}^{(\ell)}, p_{k|k-1}^{(\ell)}\right)\right\}_{\ell\in\mathbb{L}}. \tag{5.63}$$

The multi-object posterior density cannot be evaluated in a closed form [RVVD14]. Specifically, the LMB $\boldsymbol{\pi}_{k|k-1}$ has to be converted to a δ-GLMB density by (possibly) generating all possible combinations of label sets $I\in\mathcal{F}(\mathcal{L}(\mathbf{X}))$ from the constituent Bernoulli components [RVVD14]. Further, the δ-GLMB $\boldsymbol{\pi}_{k|k-1}(\cdot)$ resulting from the conversion process undergoes the δ-GLMB update step (5.15). Finally, the GLMB posterior $\boldsymbol{\pi}_k(\cdot)$ is approximated by an LMB that matches exactly the first moment of the unlabeled posterior intensity. Thus, the updated LMB is given by

$$\boldsymbol{\pi}_k(\mathbf{X}) = \left\{\left(r_k^{(\ell)}, p_k^{(\ell)}\right)\right\}_{\ell\in\mathbb{L}}, \tag{5.64}$$

where

$$r_k^{(\ell)} = \sum_{(I,\theta)\in\mathcal{F}(\mathbb{L})\times\Theta(I)} 1_I(\ell)\, w_k^{(I,\theta)}(Y_k), \tag{5.65}$$

$$p_k^{(\ell)}(x) = \frac{1}{r_k^{(\ell)}} \sum_{(I,\theta)\in\mathcal{F}(\mathbb{L})\times\Theta(I)} 1_I(\ell)\, w_k^{(I,\theta)}(Y_k)\, p_k^{(\theta)}(x,\ell), \tag{5.66}$$

$$w^{(I,\theta)}(Y_k) \propto \left[\eta_{Y_k}^{(\theta)}\right]^I \prod_{\ell\in\mathbb{L}\setminus I}\left(1-r_{k|k-1}^{(\ell)}\right)\prod_{i\in I} 1_{\mathbb{L}}(i) r_{k|k-1}^{(i)}, \tag{5.67}$$

$$p_k^{(\theta)}(x,\ell) = \frac{p_{k|k-1}(x,\ell)\,\psi_{Y_k}(x,\ell;\theta)}{\eta_{Y_k}^{(\theta)}(\ell)}\,, \tag{5.68}$$

$$\eta_{Y_k}^{(\theta)}(\ell) = \left\langle p_{k|k-1}(\cdot,\ell), \psi_{Y_k}(\cdot,\ell;\theta)\right\rangle\,, \tag{5.69}$$

$$\psi_{Y_k}(x,\ell;\theta) = \begin{cases} \dfrac{P_D(x,\ell)g(y_{\theta(\ell)}|x,\ell)}{\kappa(y_{\theta(\ell)})}, & \text{if } \theta(\ell) > 0 \\ 1 - P_D(x,\ell), & \text{if } \theta(\ell) = 0 \end{cases}. \tag{5.70}$$

Additional details for an efficient implementation of the LMB filter can be found in [RVVD14].

5.6 Performance evaluation

To assess performance of the proposed Marginalized δ-GLMB (Mδ-GLMB) tracker, a 2-dimensional multi-object tracking scenario is considered over a surveillance area of $50 \times 50\,[km^2]$. Two sensor sets are used to represent scenarios with different observability capabilities. In particular:

I. a single radar in the middle of the surveillance region is used as it guarantee observability;

II. a set of 3 *range-only* (Time Of Arrival, TOA) sensors, deployed as shown in Fig. 5.1, are used as they do not guarantee observability individually, but information from different sensors need to be combined to achieve it.

The scenario consists of 5 objects as depicted in Fig. 5.2. For the sake of comparison, the Mδ-GLMB is also compared with the δ-GLMB (δ-GLMB) [VV13, VVP14] and LMB (LMB) [RVVD14] filters. The three tracking filters are implemented using Gaussian Mixtures to represent their predicted and updated densities [VVP14,RVVD14]. Due to the non linearity of the sensors, the *Unscented Kalman Filter* (UKF) [JU04] is exploited to update means and covariances of the Gaussian components.

The kinematic object state is denoted by $x = [p_x,\, \dot{p}_x,\, p_y,\, \dot{p}_y]^\top$, i.e. the planar position and velocity. The object motion is modeled according to the Nearly-Constant Velocity (NCV) model [FS85, FS86, BSF88, BSL95]:

$$x_{k+1} = \begin{bmatrix} 1 & T_s & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T_s \\ 0 & 0 & 0 & 1 \end{bmatrix} x_k + w_k\,, \qquad Q = \sigma_w^2 \begin{bmatrix} \frac{1}{4}T_s^4 & \frac{1}{2}T_s^3 & 0 & 0 \\ \frac{1}{2}T_s^3 & T_s^2 & 0 & 0 \\ 0 & 0 & \frac{1}{4}T_s^4 & \frac{1}{2}T_s^3 \\ 0 & 0 & \frac{1}{2}T_s^3 & T_s^2 \end{bmatrix} \tag{5.71}$$

where: w_k is a white noise with zero mean and covariance matrix Q, $\sigma_w = 5\,[m/s^2]$ and the sampling interval is $T_s = 5\,[s]$.

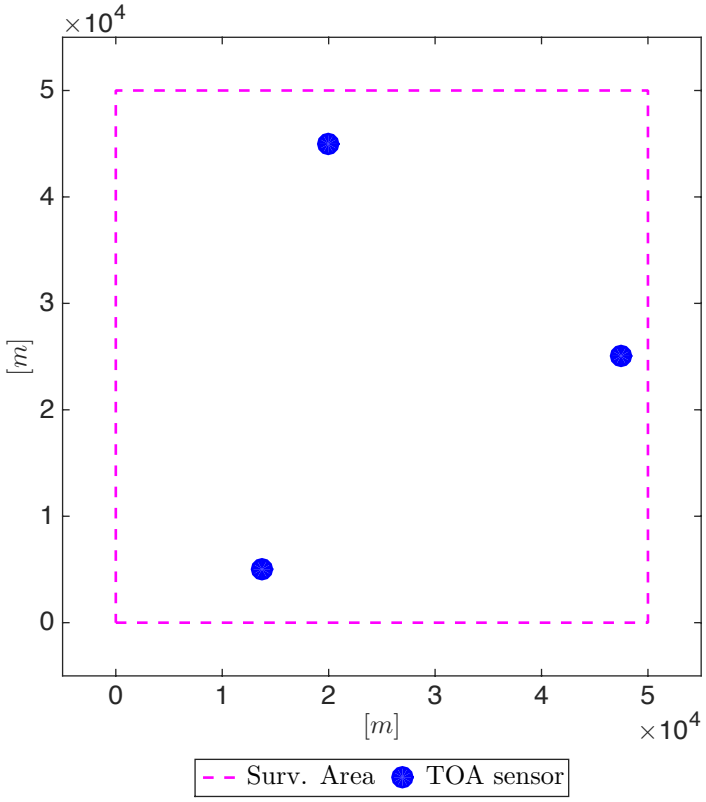


Figure 5.1: Network with 3 TOA sensors.

The radar has the following measurement function:

$$h(x) = \begin{bmatrix} \angle[(p_x - x^r) + j\,(p_y - y^r)] \\ \sqrt{(p_x - x^r)^2 + (p_y - y^r)^2} \end{bmatrix} \tag{5.72}$$

where (x^r, y^r) represents the known position of the radar and the standard deviation of the measurement noise is $\sigma_\theta = 1\,[^\circ]$ for the azimuth and $\sigma_r = 100\,[m]$ for the range. The measurement functions of the 3 TOA of Fig. 5.1 are:

$$h(x) = \sqrt{(p_x - x^s)^2 + (p_y - y^s)^2}\,, \tag{5.73}$$

where (x^s, y^s) represents the known position of sensor (indexed with) s. The standard deviation of the TOA measurement noise is taken as $\sigma_{TOA} = 100\,[m]$.

The clutter is characterized by a Poisson process with parameter $\lambda_c = 15$. The probability of object detection is $P_D = 0.85$.

In the considered scenario, objects pass through the surveillance area with partial prior information for object birth locations. Accordingly, a 10-component LMB RFS $\boldsymbol{\pi}_B = \left\{\left(r_B^{(\ell)}, p_B^{(\ell)}\right)\right\}_{\ell\in\mathbb{B}}$ has been hypothesized for the birth process. Table 5.1 gives a detailed summary of such components. Due to the partial prior information on the object birth locations,

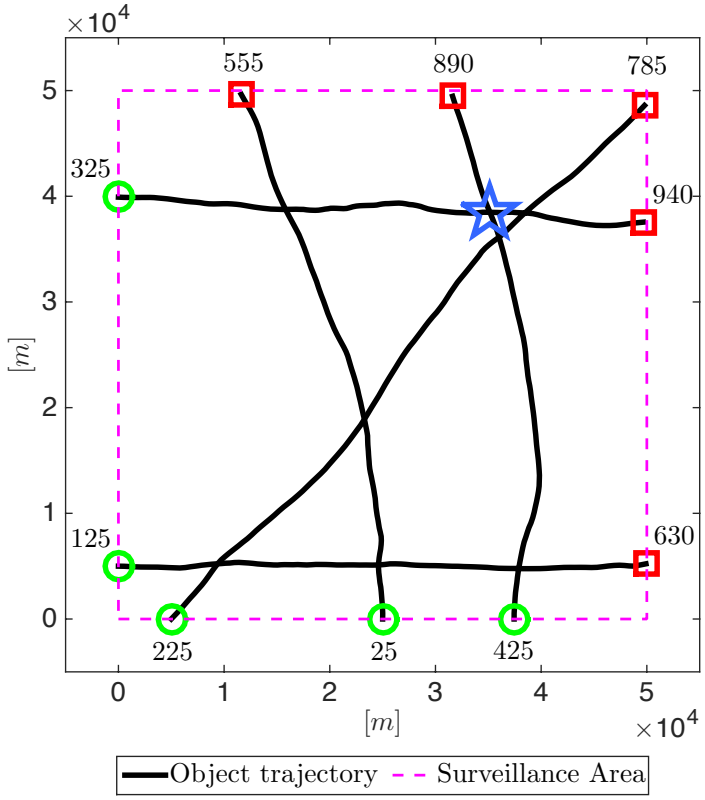


Figure 5.2: Object trajectories considered in the simulation experiment. The start/end point for each trajectory is denoted, respectively, by •\■. The ⋆ indicates a rendezvous point.

Table 5.1: Components of the LMB RFS birth process at a given time k.

$$r^{(\ell)} = 0.09$$
$$p_B^{(\ell)}(x) = \mathcal{N}\left(x;\, m_B^{(\ell)}, P_B\right)$$
$$P_B = \mathrm{diag}(10^6, 10^4, 10^6, 10^4)$$

Label	$(k,\, 1)$	$(k,\, 2)$	$(k,\, 3)$
$m_B^{(\ell)}$	$[0,\, 0,\, 40000,\, 0]^\top$	$[0,\, 0,\, 25000,\, 0]^\top$	$[0,\, 0,\, 5000,\, 0]^\top$

Label	$(k,\, 4)$	$(k,\, 5)$	$(k,\, 6)$
$m_B^{(\ell)}$	$[5000,\, 0,\, 0,\, 0]^\top$	$[25000,\, 0,\, 0,\, 0]^\top$	$[36000,\, 0,\, 0,\, 0]^\top$

Label	$(k,\, 7)$	$(k,\, 8)$
$m_B^{(\ell)}$	$[50000,\, 0,\, 15000,\, 0]^\top$	$[50000,\, 0,\, 40000,\, 0]^\top$

Label	$(k,\, 9)$	$(k,\, 10)$
$m_B^{(\ell)}$	$[40000,\, 0,\, 50000,\, 0]^\top$	$[10000,\, 0,\, 50000,\, 0]^\top$

some of the LMB components cover a state space region where there is no birth. Clutter measurements are, therefore, more prone to generate false objects.

Multi-object tracking performance is evaluated in terms of the *Optimal SubPattern Analysis* (OSPA) metric [SVV08] with Euclidean distance $p = 2$ and cutoff $c = 600\,[m]$. The reported metric is averaged over 100 Monte Carlo trials for the same object trajectories but different, independently generated, clutter and measurement noise realizations. The duration of each simulation trial is fixed to $1000\,[s]$ (200 samples).

The three tracking filters are coupled with the *parallel CPHD look ahead strategy* described in [VV13, VVP14]. The CPHD [VVC07] filter.

5.6.1 Scenario 1: radar

Figs. 5.3, 5.4 and 5.5 display the statistics (mean and standard deviation) of the estimated number of objects obtained, respectively, with Mδ-GLMB, δ-GLMB and LMB tracking filters. As it can be seen, all the algorithms estimate the object cardinality accurately, with no substantial differences. This result implies that, in the presence of a single sensor guaranteeing observability, the approximations made by both Mδ-GLMB and LMB trackers are not critical in that they provide performance comparable to δ-GLMB with the advantage of a cheaper computational burden and reduced storage requirements. Note that the problems introduced by the rendezvous point (e.g. merged or lost tracks) are correctly tackled by all the algorithms.

Fig. 5.6 shows the OSPA distance of the algorithms. Note again that, in agreement with the estimated cardinality distributions, the OSPA distances are nearly identical.

5.6.2 Scenario 2: 3 TOA sensors

Figs. 5.7, 5.8 and 5.9 display the statistics (mean and standard deviation) of the estimated number of objects obtained, respectively, with Mδ-GLMB, δ-GLMB and LMB trackers. The Mδ-GLMB and the δ-GLMB tracking filters estimate the object cardinality accurately, while the LMB exhibits poor performance and higher standard deviation due to losing some tracks when 4 or 5 objects are jointly present in the surveillance area. It is worth noticing that the Mδ-GLMB tracker exhibits performance very close to that of δ-GLMB and that the problems introduced by the rendezvous point are again succesfully tackled.

Fig. 5.10 shows the OSPA distance. Note that the OSPA of Mδ-GLMB is close to the one of δ-GLMB, while LMB shows an overall higher error in agreement with the cardinality error due to losing tracks.

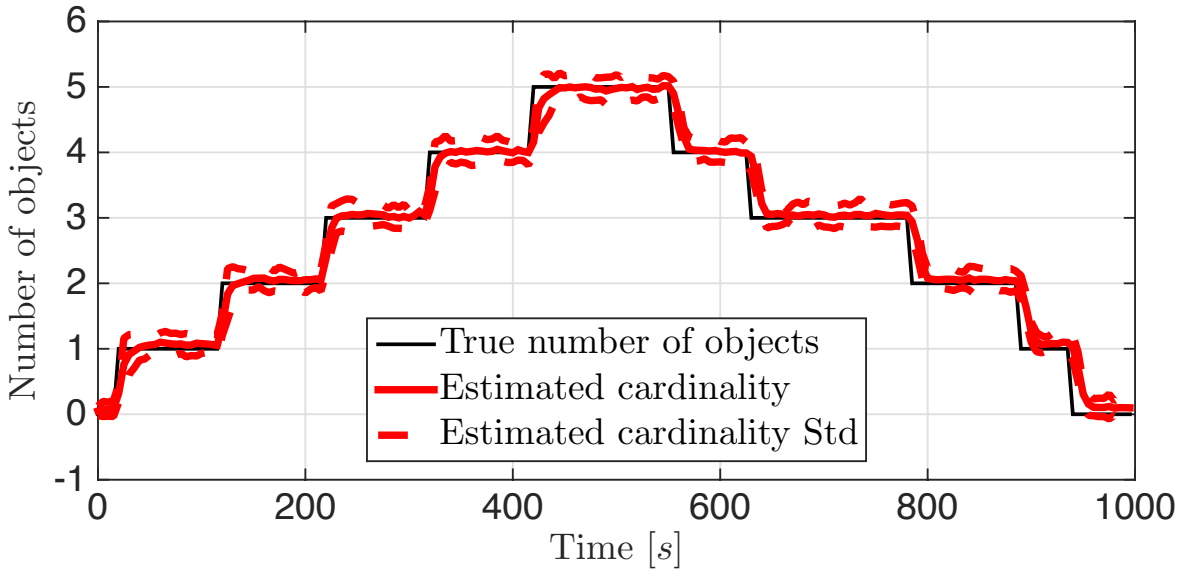


Figure 5.3: Cardinality statistics for Mδ-GLMB tracking filter using 1 radar.

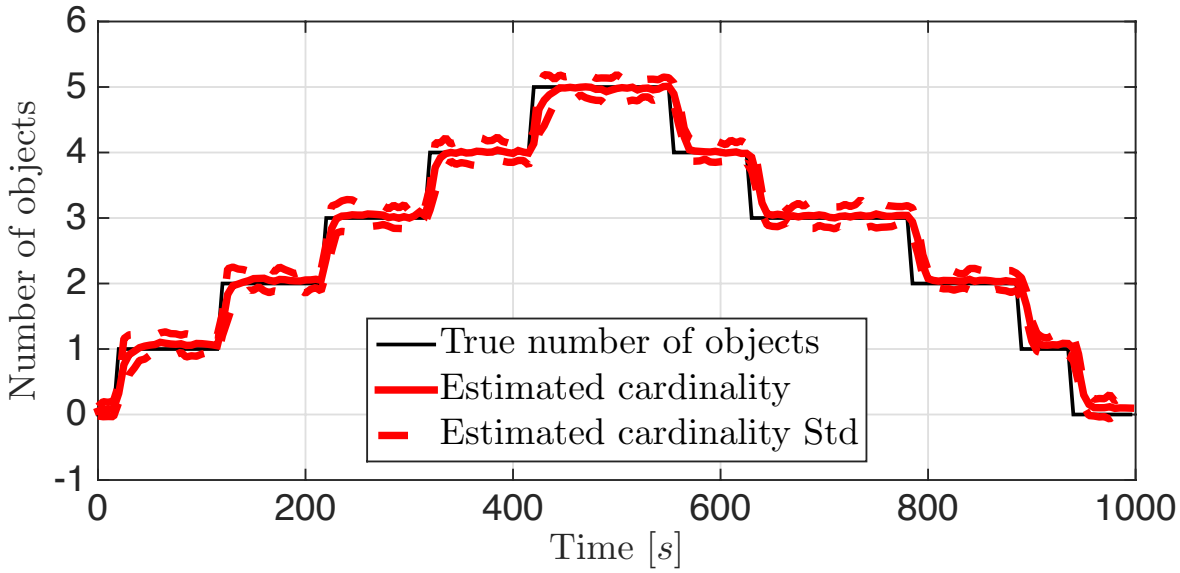


Figure 5.4: Cardinality statistics for δ-GLMB tracking filter using 1 radar.

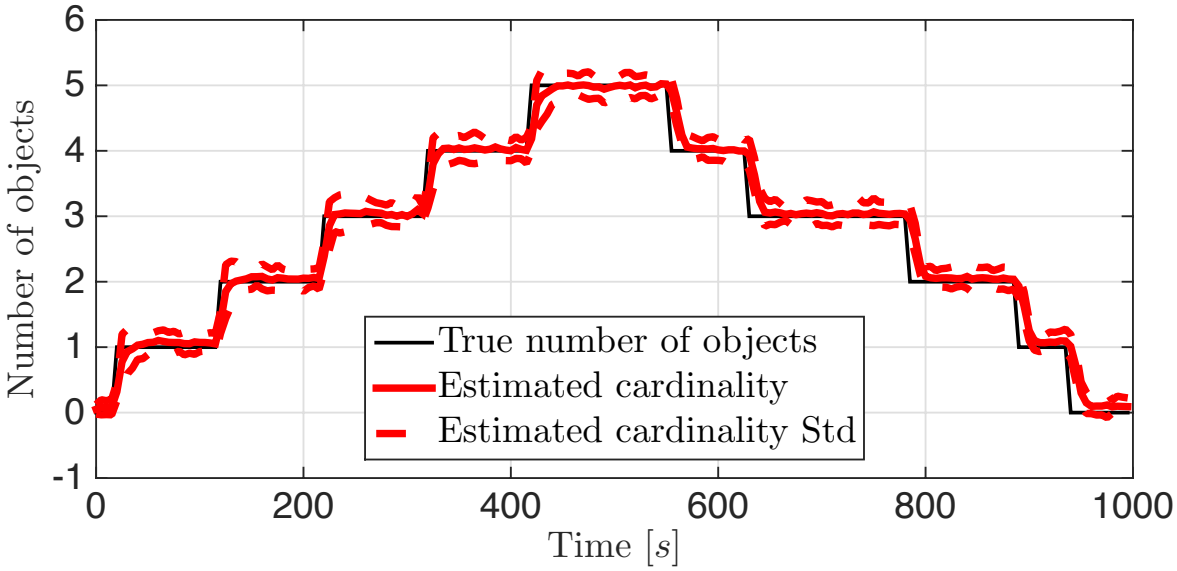


Figure 5.5: Cardinality statistics for LMB tracking filter using 1 radar.

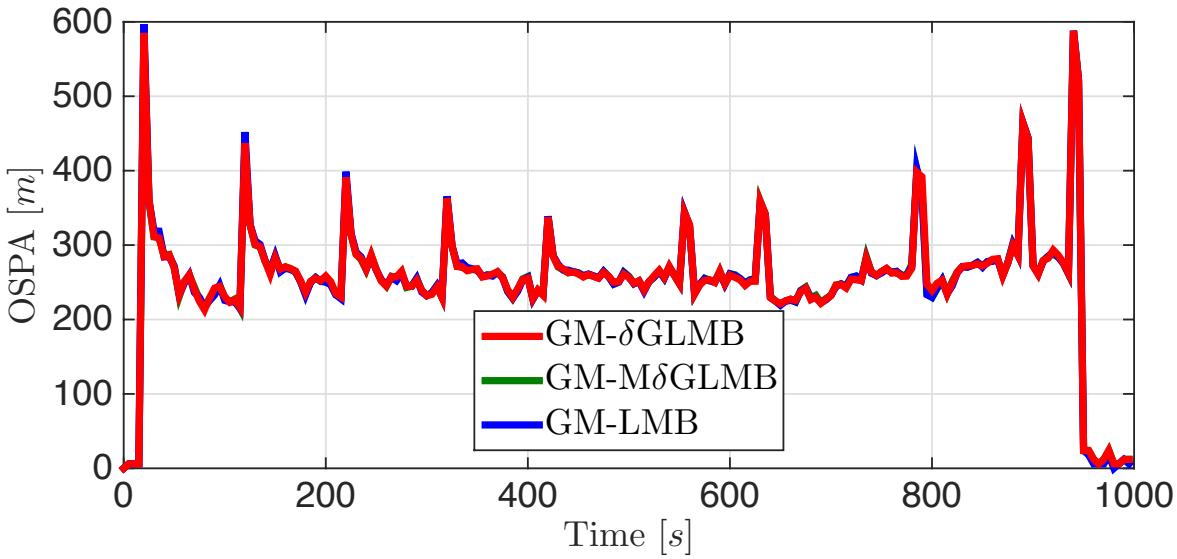


Figure 5.6: OSPA distance ($c = 600\,[m]$, $p = 2$) using 1 radar.

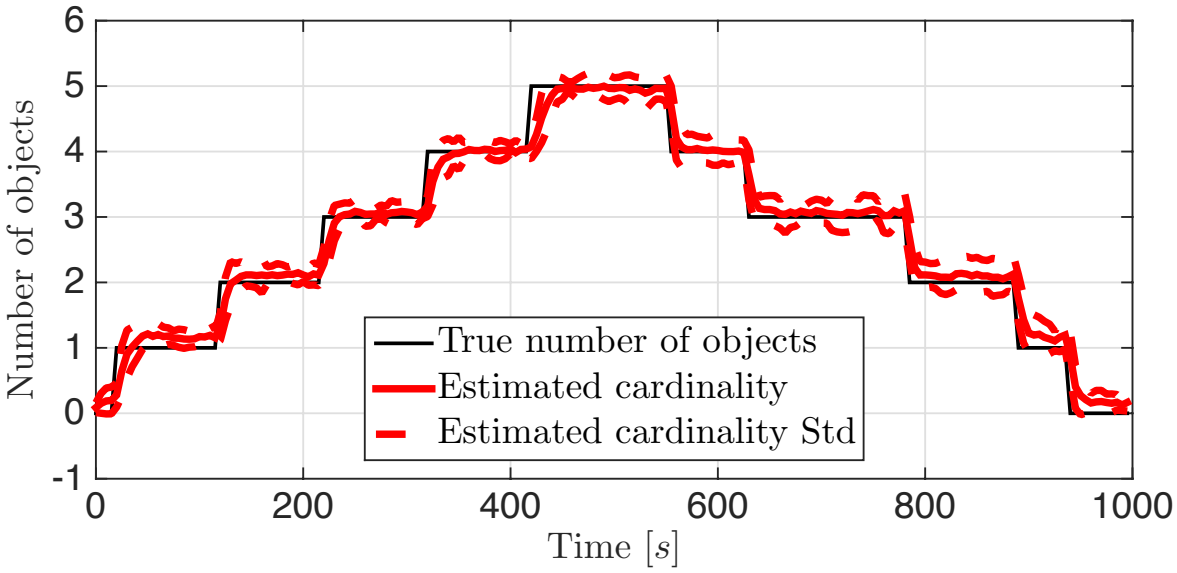


Figure 5.7: Cardinality statistics for Mδ-GLMB tracking filter using 3 TOA.

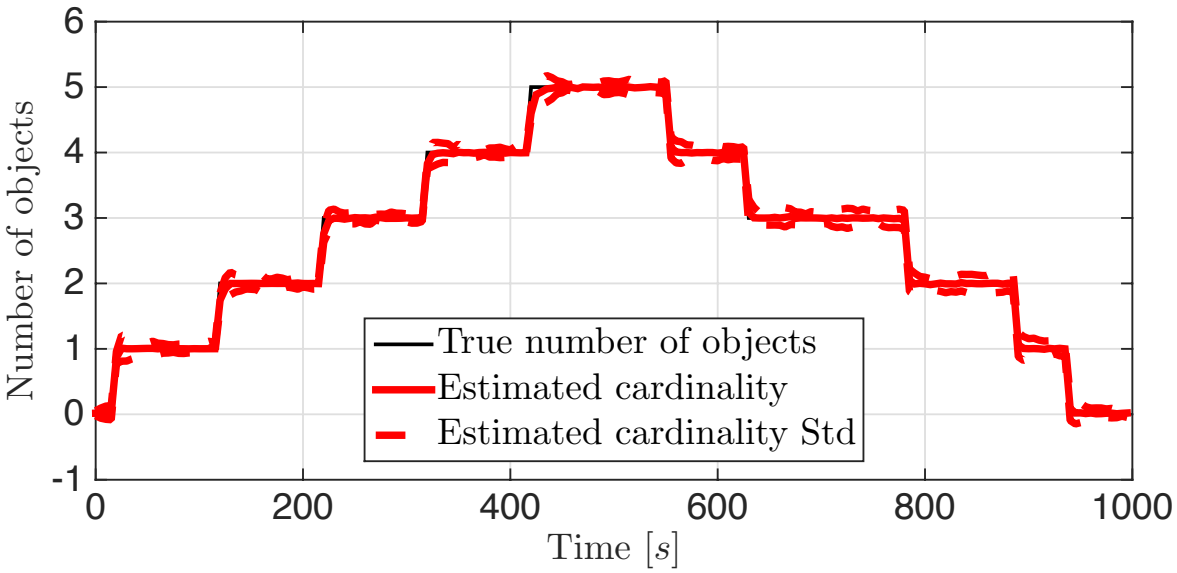


Figure 5.8: Cardinality statistics for δ-GLMB tracking filter using 3 TOA.

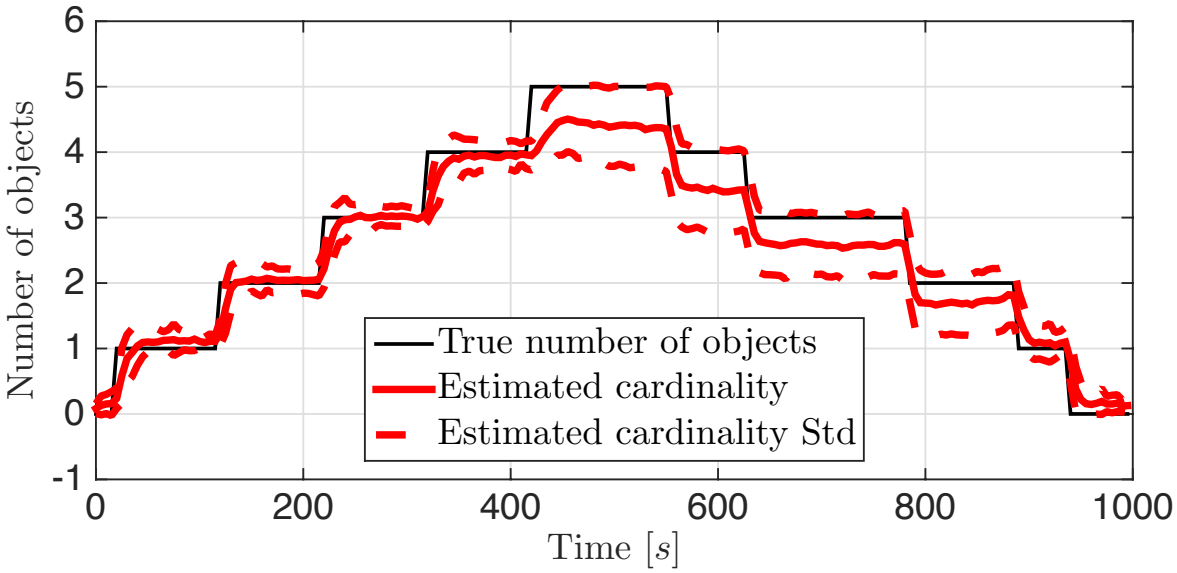


Figure 5.9: Cardinality statistics for LMB tracking filter using 3 TOA.

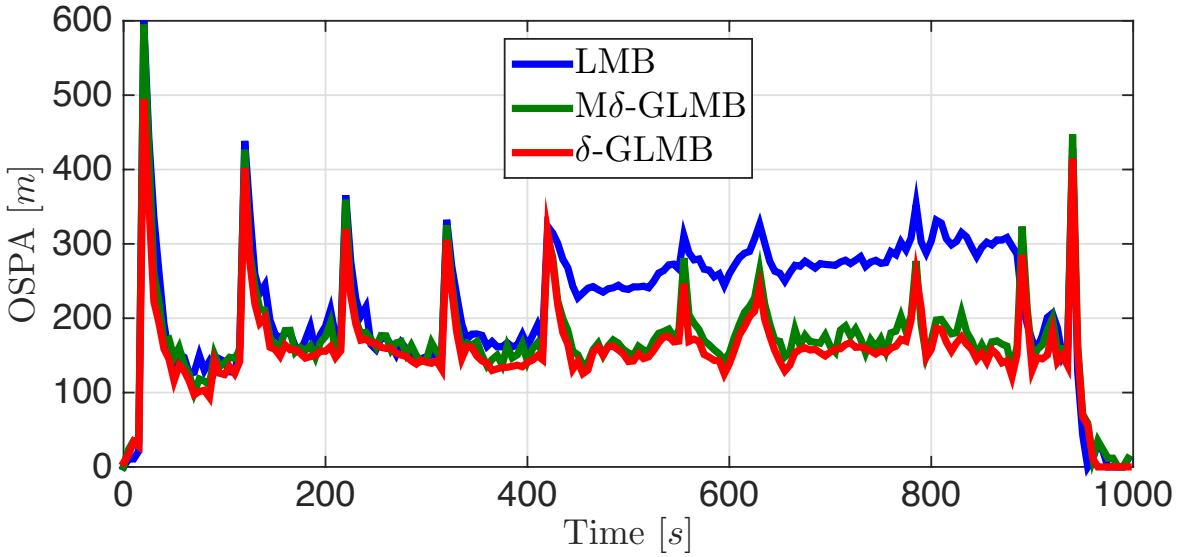


Figure 5.10: OSPA distance ($c = 600\,[m]$, $p = 2$) using 3 TOA.

6 Distributed multi-object tracking

Contents

The focus of this chapter is on *Distributed Multi-Object Tracking* (DMOT). We adopt the RFS formulation since it provides the concept of *probability density* for *multi-object state* that allows us to directly extend existing tools in distributed estimation to the multi-object case. Such a concept is not available in the *Multiple Hypotheses Tracking* (MHT) and *Joint Probabilistic Data Association* (JPDA) approaches [Rei79, FS85, FS86, BSF88, BSL95, BP99].

DMOF based on the RFS paradigm has previously been described in chapter 4 [Mah12, BCF+13a, UCJ13, BCF+14c]. Specifically, an information-theoretic approach to robust distributed multi-object estimation based on the CCPHD filter has been proposed. However, this formulation is not able to provide estimates of the object's trajectories or tracks.

Taking into account the above-mentioned considerations, the main contributions in the present chapter are [FVV+15]:

- the generalization of the robust distributed multi-object filtering approach of subsection 4.1.1 to multi-object tracking;
- the development of the *Consensus Marginalized δ-Generalized Labeled Multi-Bernoulli* (CMδGLMB) and *Consensus Labeled Multi-Bernoulli* (CLMB) filters as the first distributed multi-object trackers in the RFS framework.

The proposed solutions are based on the concept of labeled RFS introduced in subsection

2.3.6, that enables the estimation of multi-object trajectories in a principled manner [VV13]. Moreover, labeled RFS conjugate priors [VV13] have lead to the development of a tractable analytic multi-object tracking solution called the *δ-Generalized Labeled Multi-Bernoulli* (δ-GLMB) filter [VVP14]. However, it is not known if this filter is amenable to DMOT. Nonetheless, the Mδ-GLMB and the LMB filters [FVPV15, RVVD14] are two efficient approximations of the δ-GLMB filter that

I. have an appealing mathematical formulation that facilitates an efficient and tractable closed-form fusion rule for DMOT;

II. preserve key summary statistics of the full multi-object posterior.

In addition, labeled RFS-based trackers do not suffer from the so-called "*spooky effect*" [FSU09] that degrades the performance in the presence of low detection probability like the multi-object filters [VVC07, BCF$^+$13a, UCJ13].

6.1 Information fusion with labeled RFS

In this section, the notion of KLA (4.2) introduced in subsection 4.1.1 will be extended to densities of labeled RFSs and will be shown to be equivalent to the *normalized weighted geometric mean*. Proofs of the results are provided in the appendix A.

6.1.1 Labeled multi-object Kullback-Leibler average

The benefit of using the RFS framework which provides the concept of probability density of the multi-object state, allows us to directly extend the notion of multi-object KLA originally devised in [BCF$^+$13a] to labeled multi-object densities [FVV$^+$15]. Herewith, we adopt the measure theoretic notion of multi-object density given in [VSD05], which does not suffer of the unit compatibility problem in the integrals involving the product of powers of multi-object densities. Note that this density is equivalent to the FISST density as shown in [VSD05]. Thus, the standard inner product notation is extended to multi-object densities

$$\langle \boldsymbol{f}, \boldsymbol{g} \rangle \triangleq \int \boldsymbol{f}(\mathbf{X})\, \boldsymbol{g}(\mathbf{X}) \delta\mathbf{X}\,. \tag{6.1}$$

The weighted KLA $\boldsymbol{f}_{KLA}$ of the labeled multi-object densities $\boldsymbol{f}^i$ is defined by

$$\boldsymbol{f}_{KLA} \triangleq \arg\inf_{\boldsymbol{f}} \sum_{i\in\mathcal{N}} \omega^i D_{KL}\left(\boldsymbol{f} \parallel \boldsymbol{f}^i\right), \tag{6.2}$$

$$\omega^i \geq 0, \qquad \sum_{i\in\mathcal{N}} \omega^i = 1. \tag{6.3}$$

where

$$D_{KL}\left(\boldsymbol{f} \parallel \boldsymbol{g}\right) \triangleq \int \boldsymbol{f}(\mathbf{X}) \log \frac{\boldsymbol{f}(\mathbf{X})}{\boldsymbol{g}(\mathbf{X})} \delta\mathbf{X} \tag{6.4}$$

is the KLD [Mah03, Mah07b] between two multi-object densities $\boldsymbol{f}(\cdot)$ and $\boldsymbol{g}(\cdot)$.

Theorem 3 (KLA of labeled RFSs)**.**
The weighted KLA defined in (6.2) is the normalized weighted geometric mean of the multi-object densities $\boldsymbol{f}^i(\cdot)$, $i \in \mathcal{N}$, i.e.

$$\boldsymbol{f}_{KLA} = \bigoplus_{i\in\mathcal{N}} \left(\omega^i \odot \boldsymbol{f}^i\right). \tag{6.5}$$

Theorem 3 follows from Theorem 1 of [BCF+13a, Section III.B].

Remark 19. The fusion rule (6.5) is the labeled multi-object version of the multi-object Chernoff fusion first proposed by Mahler [Mah00].

Subsequently, it will be shown that the KLAs of Mδ-GLMB and LMB densities are also, respectively, a Mδ-GLMB and an LMB densities. In particular closed-form solutions for the normalized weighted geometric means are derived. These results are necessary to implement the fusion rule (6.5) for Mδ-GLMB and LMB tracking filters.

6.1.2 Normalized weighted geometric mean of Mδ-GLMB densities

The following result holds.

Theorem 4 (NWGM of Mδ-GLMB RFSs)**.**
Let $\boldsymbol{\pi}^{\imath}(\cdot)$, $\imath = 1, \ldots, \mathcal{I}$, be M$\delta$-GLMB densities on $\mathcal{F}(\mathbb{X}\times\mathbb{L})$ and $\omega^{\imath} \in (0,1)$, $\imath = 1, \ldots, \mathcal{I}$, such that $\sum_{\imath=1}^{\mathcal{I}} \omega^{\imath} = 1$. Then the normalized weighted geometric mean is given by

$$\bigoplus_{\imath=1}^{\mathcal{I}} \left(\omega^{\imath} \odot \boldsymbol{\pi}^{\imath}\right)(\mathbf{X}) = \Delta(\mathbf{X}) \sum_{L\in\mathcal{F}(\mathbb{L})} \delta_L(\mathcal{L}(\mathbf{X}))\ \overline{w}^{(L)} \left[\overline{p}^{(L)}\right]^{\mathbf{X}} \tag{6.6}$$

where

$$\overline{p}^{(L)} = \frac{\prod_{\imath=1}^{\mathcal{I}} \left(p_{\imath}^{(L)}\right)^{\omega^{\imath}}}{\int \prod_{\imath=1}^{\mathcal{I}} \left(p_{\imath}^{(L)}\right)^{\omega^{\imath}} dx} \tag{6.7}$$

$$\overline{w}^{(L)} = \frac{\prod_{\imath=1}^{\mathcal{I}} \left(w_{\imath}^{(L)}\right)^{\omega^{\imath}} \left[\int \prod_{\imath=1}^{\mathcal{I}} \left(p_{\imath}^{(L)}(x,\cdot)\right)^{\omega^{\imath}} dx\right]^{L}}{\sum_{F\subseteq\mathbb{L}} \prod_{\imath=1}^{\mathcal{I}} \left(w_{\imath}^{(F)}\right)^{\omega^{\imath}} \left[\int \prod_{\imath=1}^{\mathcal{I}} \left(p_{\imath}^{(F)}(x,\cdot)\right)^{\omega^{\imath}} dx\right]^{F}} \tag{6.8}$$

The fusion rule for Mδ-GLMBs follows by applying Theorem 4 to find the KLA (6.5) of the Mδ-GLMB's $\boldsymbol{\pi}^i(\cdot)$, $i \in \mathcal{N}$. This is summarized in the following Proposition.

Proposition 3 (KLA of Mδ-GLMB RFSs)**.**
Suppose that each agent $i \in \mathcal{N}$ is provided with an Mδ-GLMB $\boldsymbol{\pi}^i(\cdot)$ and that all agents share

the same label space for the birth process, then the Mδ-GLMB components $\overline{p}^{(L)}(\cdot)$ *and* $\overline{w}^{(L)}$ *of the KLA are given by*

$$\overline{p}^{(L)}(x,\ell) = \bigoplus_{i\in\mathcal{N}} \left(\omega^i \odot p^{(L)}\right)(x,\ell) \,, \tag{6.9}$$

$$\overline{w}^{(L)} = \frac{\widetilde{w}^{(L)}}{\sum_{F\subseteq\mathbb{L}} \widetilde{w}^{(F)}} \,, \tag{6.10}$$

where

$$\widetilde{w}^{(L)} = \prod_{i\in\mathcal{N}} \left(w_i^{(L)}\right)^{\omega^i} \left[\int \widetilde{p}^{(L)}(x,\cdot)\,dx\right]^L \,, \tag{6.11}$$

$$\widetilde{p}^{(L)} = \prod_{i\in\mathcal{N}} \left(p_i^{(L)}\right)^{\omega^i} \,. \tag{6.12}$$

Remark 20. Note that the quantities $\widetilde{w}^{(L)}$ and $\overline{p}^{(L)}$ can be independently determined using (6.9) and (6.10). Thus, the overall fusion procedure is fully parallelizable.

Remark 21. Notice that (6.9) is indeed the CI fusion rule [JU97] for the single-object PDFs.

6.1.3 Normalized weighted geometric mean of LMB densities

The following result holds.

Theorem 5 (NWGM of LMB RFSs)**.**
Let $\boldsymbol{\pi}^\imath = \left\{\left(r_i^{(\ell)}, p_i^{(\ell)}\right)\right\}_{\ell\in\mathbb{L}}$, $\imath = 1,\ldots,\mathcal{I}$, *be LMB densities on* $\mathcal{F}(\mathbb{X}\times\mathbb{L})$ *and* $\omega^\imath \in (0,1)$, $\imath = 1,\ldots,\mathcal{I}$, *such that* $\sum_{\imath=1}^{\mathcal{I}} \omega^\imath = 1$. *Then the normalized weighted geometric mean is given by*

$$\bigoplus_{\imath=1}^{\mathcal{I}} \left(\omega^\imath \odot \boldsymbol{\pi}^\imath\right) = \left\{\left(\overline{r}^{(\ell)}, \overline{p}^{(\ell)}\right)\right\}_{\ell\in\mathbb{L}} \tag{6.13}$$

where

$$\overline{r}^{(\ell)} = \frac{\int \prod_{\imath=1}^{\mathcal{I}} \left(r_i^{(\ell)} p_i^{(\ell)}(x)\right)^{\omega^\imath} dx}{\prod_{\imath=1}^{\mathcal{I}} \left(1 - r_i^{(\ell)}\right)^{\omega^\imath} + \int \prod_{\imath=1}^{\mathcal{I}} \left(r_i^{(\ell)} p_i^{(\ell)}(x)\right)^{\omega^\imath} dx} \tag{6.14}$$

$$\overline{p}^{(\ell)} = \bigoplus_{\imath=1}^{\mathcal{I}} \left(\omega^\imath \odot p_i^{(\ell)}\right) \tag{6.15}$$

The fusion rule for LMBs follows by applying Theorem 5 to find the KLA (6.5) of the LMB's $\left\{\left(r_i^{(\ell)}, p_i^{(\ell)}\right)\right\}_{\ell\in\mathbb{L}}$, $i \in \mathcal{N}$. This is summarized in the following Proposition.

Proposition 4 (KLA of LMB RFSs)**.**
Suppose that each agent $i \in \mathcal{N}$ *is provided with an LMB* $\boldsymbol{\pi}^i(\cdot)$ *and that all agents share the*

same label space for the birth process, then the LMB components $\left\{\left(\overline{r}^{(\ell)}, \overline{p}^{(\ell)}\right)\right\}_{\ell\in\mathbb{L}}$ *of the KLA are given by*

$$\overline{r}^{(\ell)} = \frac{\widetilde{r}^{(\ell)}}{\widetilde{q}^{(\ell)} + \widetilde{r}^{(\ell)}}\,, \tag{6.16}$$

$$\overline{p}^{(\ell)}(x) = \bigoplus_{i\in\mathcal{N}} \left(\omega^i \odot p_i^{(\ell)}\right)(x)\,, \tag{6.17}$$

where

$$\widetilde{r}^{(\ell)} = \int \prod_{i\in\mathcal{N}} \left(r_i^{(\ell)} p_i^{(\ell)}(x)\right)^{\omega^i} dx\,, \tag{6.18}$$

$$\widetilde{q}^{(\ell)} = \prod_{i\in\mathcal{N}} \left(1 - r_i^{(\ell)}\right)^{\omega^i}\,. \tag{6.19}$$

Remark 22. Note that each Bernoulli component $\left(\overline{r}^{(\ell)}, \overline{p}^{(\ell)}\right)$ can be independently determined using eqs. (6.16) and (6.17). Thus, the overall fusion procedure is fully parallelizable.

Remark 23. Notice that also in this case eq. (6.17) is the CI fusion rule [JU97] for the single-object PDFs.

6.1.4 Distributed Bayesian multi-object tracking via consensus

At time k, the global KLA (6.5) which would require all the local multi-object densities $\boldsymbol{\pi}_k^i(\cdot)$, $i \in \mathcal{N}$, to be available, can be computed in a distributed and scalable way by iterating regional averages via the consensus algorithm [BCF+13a, Section III.A] [BC14] described in subsection 2.4.3. Thus, each agent $i \in \mathcal{N}$ iterates the consensus steps

$$\boldsymbol{\pi}_{k,l}^i = \bigoplus_{j\in\mathcal{N}^i} \left(\omega^{i,j} \odot \boldsymbol{\pi}_{k,l-1}^j\right)\,, \tag{6.20}$$

with $\boldsymbol{\pi}_{k,0}^i(\cdot) = \boldsymbol{\pi}_k^i(\cdot)$; $\omega^{i,j} \geq 0$, satisfying $\sum_{j\in\mathcal{N}^i}\ \omega^{i,j} = 1$, are the consensus weights relating agent i to nodes $j \in \mathcal{N}^i$. Using the same arguments as the single-object case (subsection 2.4.3), it follows that, at time k, if the consensus matrix is primitive and doubly stochastic, the consensus iterate of each node in the network converges to the global unweighted KLA (2.107) of the multi-object posterior densities as the number of consensus iterations l tends to infinity. Convergence follows along the same line as in [CA09, BC14] since $\mathcal{F}(\mathbb{X}\times\mathbb{L})$ is a metric space [Mah07b]. In practice, the iteration is stopped at some finite l.

Remark 24. The consensus iteration (6.20) is the multi-object counterpart of equation (2.119) reviewed in Subsection 2.4.3.

For Mδ-GLMB multi-object densities, (6.20) can be computed via (6.9) and (6.10), while for LMB densities by means of (6.16) and (6.17). In the present work, each single-object

density p_i is represented by a *Gaussian Mixture* (GM) of the form

$$p(x) = \sum_{j=1}^{N_G} \alpha_j \, \mathcal{N}(x; x_j, P_j) \, . \tag{6.21}$$

Note that the fusion rules (6.9) and (6.17) involve exponentiation and multiplication of GM in the form of (6.21) which, in general, does not provide a GM. Hence, the same solution conceived in subsection 4.1.4 for fusing GMs is also adopted here.

The other common approach for representing a single object location PDF p is using *particles*. Information fusion involving convex combinations of Dirac delta functions requires, at the best of our knowledge, the exploitation of additional techniques like kernel density estimation [UCJ13], least square estimation [HSH$^+$12, OHD14] or parametric model approaches [Coa04] which increase the in-node computational burden. Moreover, the local filtering steps are also more resource demanding with respect to a GM implementation. At this stage, it is preferred to follow the approach devised in [BCF$^+$13a] by adopting a GM representation.

6.2 Consensus labeled RFS information fusion

In this section, two novel fully distributed and scalable multi-object tracking algorithms are described by exploiting Propositions 3 and 4 along with consensus [OSFM07, XBL05, CA09, BC14] to propagate information throughout the network [FVV$^+$15]. Pseudo-codes of the algorithms are also provided.

6.2.1 Consensus Mδ-GLMB filter

This subsection describes the novel *Gaussian Mixture - Consensus Marginalized δ-Generalized Labelled Multi-Bernoulli* (GM-CMδGLMB) filter algorithm. The operations reported in Table 6.1 are sequentially carried out locally by each agent $i \in \mathcal{N}$ of the network. Each node operates autonomously at each sampling interval k, starting from its own previous estimates of the multi-object distribution $\boldsymbol{\pi}^i(\cdot)$, having location PDFs $p^{(I)}(x, \ell)$, $\forall\, \ell \in I$, $I \in \mathcal{F}(\mathbb{L})$, represented with a GM, and producing, at the end of the sequence of operations, its new multi-object distribution $\boldsymbol{\pi}^i(\cdot) = \hat{\boldsymbol{\pi}}^i_N(\cdot)$ as an outcome of the Consensus procedure. The steps of the GM-CδGLMB algorithm follow hereafter.

1. Each agent $i \in \mathcal{N}$ locally performs a GM-δGLMB prediction (5.31) and update (5.37). Additional details of the two procedures can be found in [VVP14, Section IV.B].

2. Consensus takes place in each node i involving the subnetwork $\mathcal{N}^i$. More precisely, at each consensus step, node i transmits its data to nodes j such that $i \in \mathcal{N}^j$ and waits until it receives data from $j \in \mathcal{N}^i \backslash \{i\}$. Next, node i carries out the fusion rule of Proposition 3 over $\mathcal{N}^i$, i.e. performs (6.20) using local information and information

received by $\mathcal{N}^i$. Finally, a merging step for each location PDF is applied to reduce the joint communication-computation burden for the next consensus step. This procedure is repeatedly applied for a chosen number $N \geq 1$ of consensus steps.

3. After consensus, an estimate of the object set is obtained from the cardinality PMF and the location PDFs via an estimate extraction described in Table 6.2.

Table 6.1: Gaussian Mixture - Consensus Marginalized δ-Generalized Labeled Multi-Bernoulli (GM-CMδGLMB) filter

procedure GM-CMδGLMB(Node i, Time k)
- Local Prediction ▷ See (5.31) and [VVP14, Table 2, Section V]
- Local Update ▷ See (5.37) and [VVP14, Table 1, Section IV]
- Marginalization ▷ See (5.42) and (5.43)
- **for** $n = 1, \ldots, N$ **do**
  - Information Exchange
  - GM-MδDGLMB Fusion ▷ See eqs. (6.9) and (6.10)
  - GM Merging ▷ See [VM06, Table II, Section III.C]
- **end for**
- Estimate Extraction ▷ See algorithm in Table 6.2

end procedure

Table 6.2: GM-CMδGLMB estimate extraction

Input: $\boldsymbol{\pi}$, N_{max}
Output: $\hat{\mathbf{X}}$

for $c = 1, \ldots, C_{max}$ **do**
$$\rho(c) = \sum_{I \in \mathcal{F}_c(\mathbb{L})} w^{(I)}$$
end for
$$\hat{C} = \arg\max_c \rho(c)$$
$$\hat{I} = \arg \max_{I \in \mathcal{F}_{\hat{C}}(\mathbb{L})} w^{(I)}$$
$$\hat{\mathbf{X}} = \left\{ \left(\hat{x}, \hat{\ell} \right) : \hat{\ell} \in \hat{I},\ \hat{x} = \arg\max_x p^{(\hat{I})}(x, \hat{\ell}) \right\}$$

6.2.2 Consensus LMB filter

This subsection describes the novel *Gaussian Mixture - Consensus Labelled Multi-Bernoulli* (GM-CLMB) filter algorithm. The operations reported in Table 6.3 will be sequentially carried out locally by each agent $i \in \mathcal{N}$ of the network. Each node operates autonomously at each sampling interval k, starting from its own previous estimates of the multi-object distribution $\boldsymbol{\pi}^i(\cdot)$, having location PDFs $p^{(\ell)}$, $\forall\, \ell \in \mathbb{L}$, represented with a GM, and producing,

at the end of the sequence of operations, its new multi-object distribution $\boldsymbol{\pi}^i(\cdot) = \hat{\boldsymbol{\pi}}^i_N(\cdot)$ as an outcome of the consensus procedure. A summary description of the steps of the GM-CLMB algorithm follows.

1. Each agent $i \in \mathcal{N}$ locally performs a GM-LMB prediction (5.59) and update (5.64). The update procedure involves two more steps: I) a GM δ-GLMB distribution is created from the predicted GM-LMB and updated using the local measurement set Y_k; II) the updated GM δ-GLMB is converted back to a GM-LMB distribution. The details of the GM-LMB prediction and update can be found in [RVVD14, VVP14].

2. Consensus takes place in each node i involving the subnetwork $\mathcal{N}^i$. More precisely, at each consensus step, node i transmits its data to nodes j such that $i \in \mathcal{N}^j$ and waits until it receives data from $j \in \mathcal{N}^i \backslash \{i\}$. Next, node i carries out the fusion rule of Proposition 4 over $\mathcal{N}^i$, i.e. performs (6.20) using local information and information received by $\mathcal{N}^i$. Finally, a merging step for each location PDF is applied to reduce the joint communication-computation burden for the next consensus step. This procedure is repeatedly applied for a chosen number $N \geq 1$ of consensus steps.

3. After consensus, an estimate of the object set is obtained from the cardinality PMF and the location PDFs via an estimate extraction described in Table 6.4.

Table 6.3: Gaussian Mixture - Consensus Labeled Multi-Bernoulli (GM-CLMB) filter

procedure GM-CLMB(NODE i, TIME k)
  LOCAL PREDICTION ▷ See eq. (5.59) and [RVVD14, Proposition 2, Section III.A]
  GM-LMB $\rightarrow$ GM δ-GLMB ▷ See [RVVD14, Section IV.C.1]
  LOCAL UPDATE ▷ See eq. (5.64) and [VVP14, Table 1, Section IV]
  GM δ-GLMB $\rightarrow$ GM-LMB ▷ See [RVVD14, Proposition 4, Section III.B]
  for $n = 1, \ldots, N$ **do**
    INFORMATION EXCHANGE
    GM-LMB FUSION ▷ See eqs. (6.16) and (6.17)
    GM MERGING ▷ See [VM06, Table II, Section III.C]
  end for
  ESTIMATE EXTRACTION ▷ See algorithm in Table 6.4
end procedure

6.3 Performance evaluation

To assess performance of the proposed Gaussian Mixture Consensus Marginalized δ-GLMB (GM-CMδGLMB) and LMB (GM-CLMB) described in section 6.2, a 2-dimensional multi-object tracking scenario is considered over a surveillance area of $50 \times 50\,[km^2]$, wherein the

Table 6.4: GM-CLMB estimate extraction

Input: $\boldsymbol{\pi} = \left\{r^{(\ell)}, p^{(\ell)}\right\}_{\ell\in\mathbb{L}}$, N_{max}

Output: $\hat{\mathbf{X}}$

for $c = 1, \ldots, C_{max}$ **do**

$\rho(c) = \sum_{L\in\mathcal{F}_c(\mathbb{L})} w(L)$

end for

$\hat{C} = \arg\max_c \rho(c)$

$\hat{\mathbb{L}} = \varnothing$

for $\hat{c} = 1, \ldots, \hat{C}$ **do**

$\hat{\mathbb{L}} = \hat{\mathbb{L}} \cup \arg\max_{\ell\in\mathbb{L}\setminus\hat{\mathbb{L}}} r^{(\ell)}$

end for

$\hat{\mathbf{X}} = \left\{\left(\hat{x}, \hat{\ell}\right) : \hat{\ell} \in \hat{\mathbb{L}},\ \hat{x} = \arg\max_x p^{(\hat{\ell})}(x)\right\}$

sensor network of Fig. 6.1 is deployed. The scenario consists of 5 objects as depicted in Fig. 6.2. For the sake of comparison, the trackers are also compared with the Gaussian Mixture Consensus CPHD (GM-CCPHD) filter of [BCF+13a] which, however, does not provide labeled tracks.

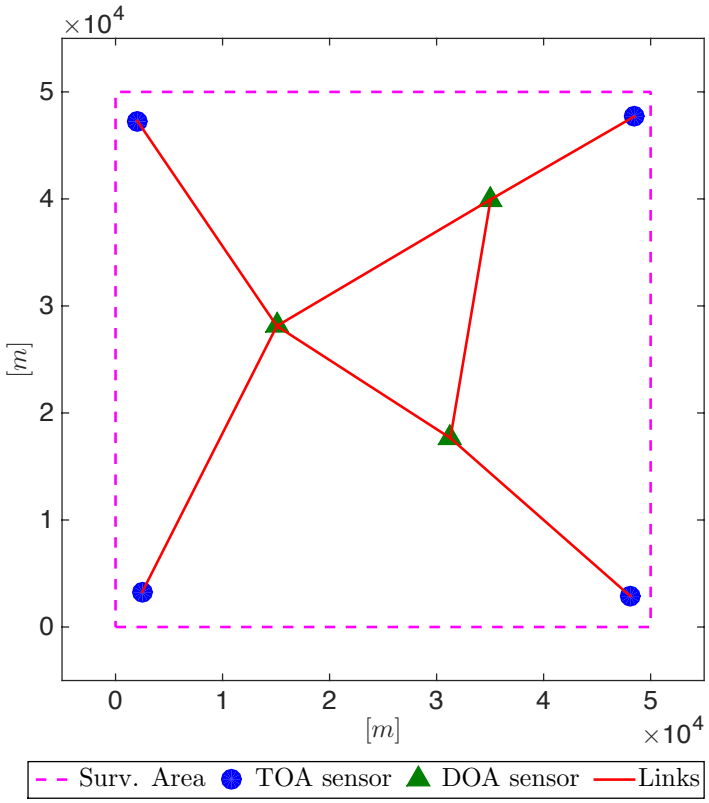


Figure 6.1: Network with 7 sensors: 4 TOA and 3 DOA.

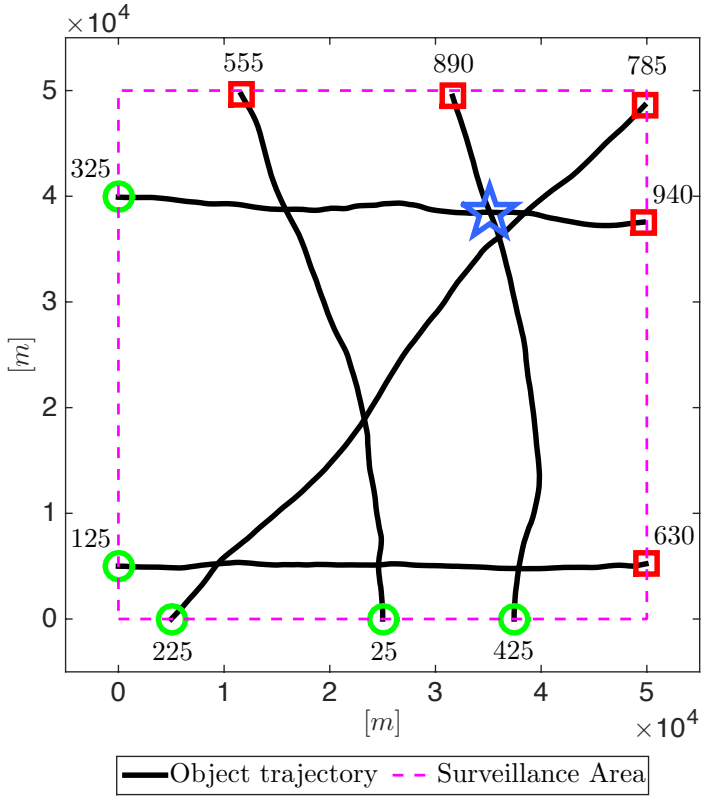


Figure 6.2: Tbject trajectories considered in the simulation experiment. The start/end point for each trajectory is denoted, respectively, by •\■. The ⋆ indicates a rendezvous point.

The kinematic object state is denoted by $x = [p_x,\, \dot{p}_x,\, p_y,\, \dot{p}_y]^\top$, i.e. the planar position and velocity. The motion of objects is modeled by the filters according to the Nearly-Constant Velocity (NCV) model [FS85, FS86, BSF88, BSL95]:

$$x_{k+1} = \begin{bmatrix} 1 & T_s & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T_s \\ 0 & 0 & 0 & 1 \end{bmatrix} x_k + w_k\,, \qquad Q = \sigma_w^2 \begin{bmatrix} \frac{1}{4}T_s^4 & \frac{1}{2}T_s^3 & 0 & 0 \\ \frac{1}{2}T_s^3 & T_s^2 & 0 & 0 \\ 0 & 0 & \frac{1}{4}T_s^4 & \frac{1}{2}T_s^3 \\ 0 & 0 & \frac{1}{2}T_s^3 & T_s^2 \end{bmatrix}, \tag{6.22}$$

where: w_k is a white noise with zero mean and covariance matrix Q, $\sigma_w = 5\,[m/s^2]$ and the sampling interval is $T_s = 5\,[s]$.

The sensor network considered in the simulation (see Fig. 6.1) consists of 4 *range-only* (Time Of Arrival, TOA) and 3 *bearing-only* (Direction Of Arrival, DOA) sensors characterized by the following measurement functions:

$$h^i(x) = \begin{cases} \angle[(p_x - x^i) + j\,(p_y - y^i)], & \text{if } i \text{ is a DOA sensor} \\ \sqrt{(p_x - x^i)^2 + (p_y - y^i)^2}, & \text{if } i \text{ is a TOA sensor} \end{cases} \tag{6.23}$$

where (x^i, y^i) represents the known position of sensor i. The standard deviation of DOA and TOA measurement noises are taken respectively as $\sigma_{DOA} = 1\,[^\circ]$ and $\sigma_{TOA} = 100\,[m]$. Because of the non linearity of the aforementioned sensors, the *Unscented Kalman Filter* (UKF) [JU04] is exploited in each sensor in order to update means and covariances of the Gaussian components.

Three different scenarios will be considered, each of which has a different clutter Poisson process with parameter λ_c and a probability of object detection P_D.

- **High SNR**: $\lambda_c = 5$, $P_D = 0.99$. These parameters were used in the work [BCF+13a] and, therefore, will be used as a first comparison test.
- **Low SNR**: $\lambda_c = 15$, $P_D = 0.99$. These parameters try to describe, in a realistic way, a scenario characterized by high clutter rate λ_c.
- **Low $\mathbf{P_D}$**: $\lambda_c = 5$, $P_D = 0.7$. These parameters test the distributed algorithms in the presence of severe misdetection.

All the above-mentioned case studies have, for each sensor, a uniform clutter spatial distribution over the surveillance area.

In the considered scenario, objects pass through the surveillance area with partial prior information for object birth locations. Accordingly, a 10-component LMB RFS $\boldsymbol{\pi}_B = \left\{\left(r_B^{(\ell)}, p_B^{(\ell)}\right)\right\}_{\ell\in\mathbb{B}}$ has been hypothesized for the birth process. Table 6.5 gives a detailed summary of such components. Due to the partial prior information on the object birth lo-

Table 6.5: Components of the LMB RFS birth process at a given time k.

$$r^{(\ell)} = 0.09$$
$$p_B^{(\ell)}(x) = \mathcal{N}\left(x;\, m_B^{(\ell)}, P_B\right)$$
$$P_B = \mathrm{diag}(10^6, 10^4, 10^6, 10^4)$$

Label	$(k,\,1)$	$(k,\,2)$	$(k,\,3)$
$m_B^{(\ell)}$	$[0,\,0,\,40000,\,0]^\top$	$[0,\,0,\,25000,\,0]^\top$	$[0,\,0,\,5000,\,0]^\top$

Label	$(k,\,4)$	$(k,\,5)$	$(k,\,6)$
$m_B^{(\ell)}$	$[5000,\,0,\,0,\,0]^\top$	$[25000,\,0,\,0,\,0]^\top$	$[36000,\,0,\,0,\,0]^\top$

Label	$(k,\,7)$	$(k,\,8)$
$m_B^{(\ell)}$	$[50000,\,0,\,15000,\,0]^\top$	$[50000,\,0,\,40000,\,0]^\top$

Label	$(k,\,9)$	$(k,\,10)$
$m_B^{(\ell)}$	$[40000,\,0,\,50000,\,0]^\top$	$[10000,\,0,\,50000,\,0]^\top$

cations, some of the LMB components cover a state space region where there is no birth. Therefore, clutter measurements are more prone to generate false objects.

Multi-object tracking performance is evaluated in terms of the *Optimal SubPattern Assignment* (OSPA) metric [SVV08] with Euclidean distance $p = 2$ and cutoff $c = 600\,[m]$. The

reported metric is averaged over 100 Monte Carlo trials for the same object trajectories but different, independently generated, clutter and measurement noise realizations. The duration of each simulation trial is fixed to 1000 $[s]$ (200 samples).

The GM-CMδGLMB and the GM-CLMB are capped to 20000, 8000 and 3000 hypotheses [VV13, VVP14], respectively, for the High SNR, Low SNR and Low P_D scenario, and are coupled with the *parallel CPHD look ahead strategy* described in [VV13, VVP14]. The CPHD filter is similarly capped, for each case study, to the same number of components through pruning and merging of mixture components.

The parameters of the GM-CCPHD filter have been chosen as follows: the survival probability is $P_s = 0.99$; the maximum number of Gaussian components is $N_{max} = 25$; the merging threshold is $\gamma_m = 4$; the truncation threshold is $\gamma_t = 10^{-4}$; the extraction threshold is $\gamma_e = 0.5$; the birth intensity function is the PHD of the LMB RFS of Table 6.5. A single consensus step $L = 1$ is employed for all the simulations.

6.3.1 High SNR

Figs. 6.3, 6.4 and 6.5 display the statistics (mean and standard deviation) of the estimated number of objects obtained, respectively, with GM-CCPHD, GM-CLMB and GM-CMδGLMB. As it can be seen, all the distributed algorithms estimate the object cardinality accurately, with the GM-CMδGLMB exhibiting better estimated cardinality variance. Note that the difficulties introduced by the rendezvous point (e.g. merged or lost tracks) are correctly tackled by all the distributed algorithms.

Fig. 6.6 shows the OSPA distance for the three algorithms. The improved localization performance of GM-CLMB and GM-CMδGLMB is attributed to two factors: (a) the "spooky effect" [FSU09] causes GM-CCPHD filter to temporarily drop objects which are subjected to missed detections and to declare multiple estimates for existing tracks in place of the dropped objects, and (b) the two trackers are generally able to better localize objects due to a more accurate propagation of the posterior density. Note that GM-CLMB and GM-CMδGLMB have similar performance since the approximations introduced by the LMB tracker (see (5.65)-(5.66)) are negligible in the case of high SNR.

6.3.2 Low SNR

Figs. 6.7 and 6.8 display the statistics (mean and standard deviation) of the estimated number of objects obtained, respectively, with the GM-CCPHD and the GM-CMδGLMB. As it can be seen, the two distributed algorithms estimate the object cardinality accurately, with the GM-CMδGLMB exhibiting again better estimated cardinality variance.

Note that the GM-CLMB fails to track the objects. The problem is due to the approximation (5.65)-(5.66) made to convert a δ-GLMB to an LMB, becoming significant with low SNR. In particular, each local tracker fails to properly set the existence probability of the tracks for three main factors: (a) no local observability, (b) high clutter rate and (c) loss of

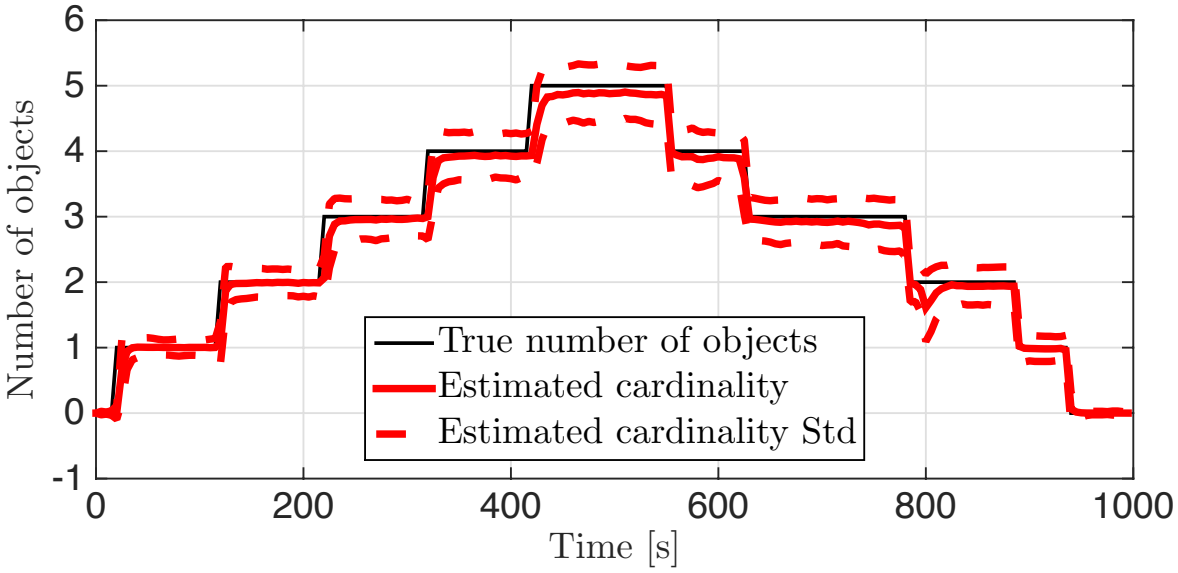


Figure 6.3: Cardinality statistics for the GM-CCPHD filter under high SNR.

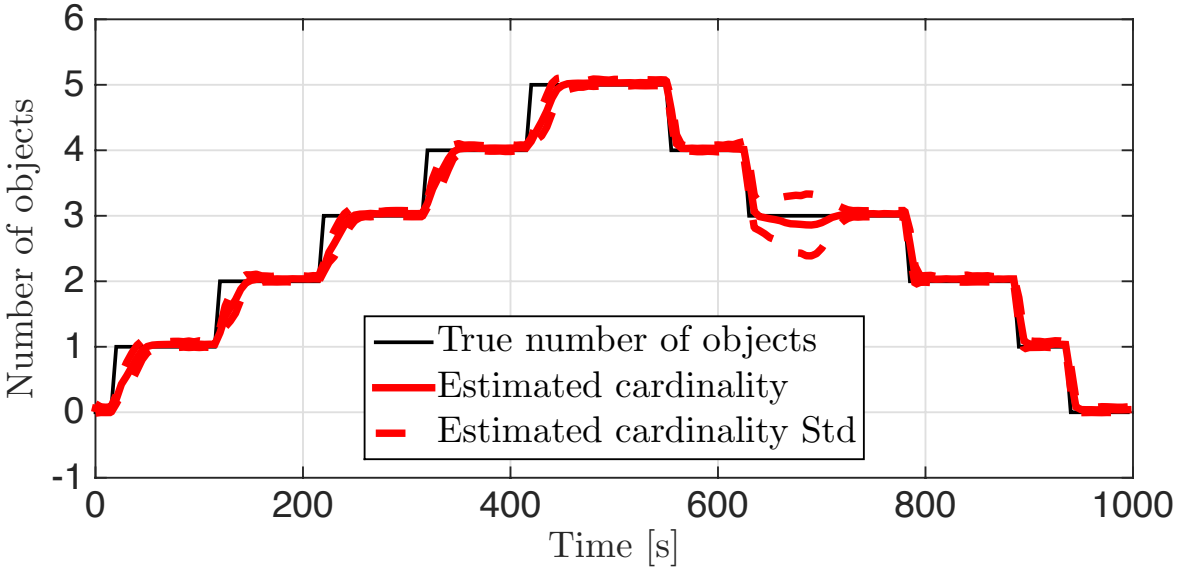


Figure 6.4: Cardinality statistics for the GM-CLMB tracker under high SNR.

the full posterior cardinality distribution after the probability density conversion. By having low existence probabilities, the extraction of the tracks fails even if the single object densities are correctly propagated in time.

Fig. 6.9 shows the OSPA distance for the current scenario. As for the previous case study, GM-CMδGLMB outperforms GM-CCPHD.

6.3.3 Low P_D

Figs. 6.10 displays the statistics (mean and standard deviation) of the estimated number of objects obtained with the GM-CMδGLMB. It is worth pointing out that the only working

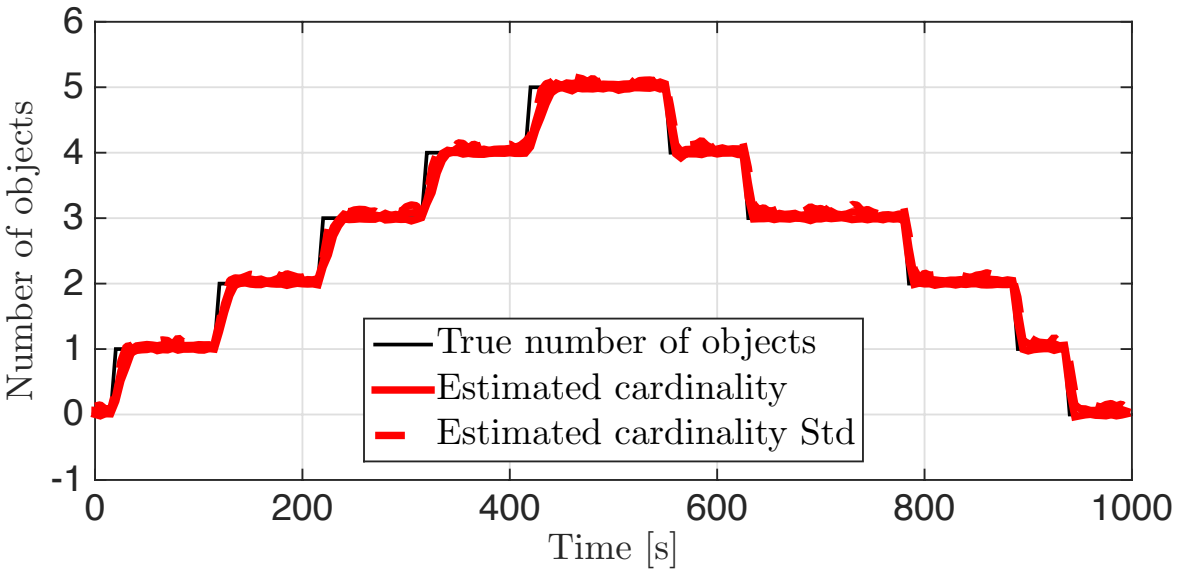


Figure 6.5: Cardinality statistics for the GM-CMδGLMB tracker under high SNR.

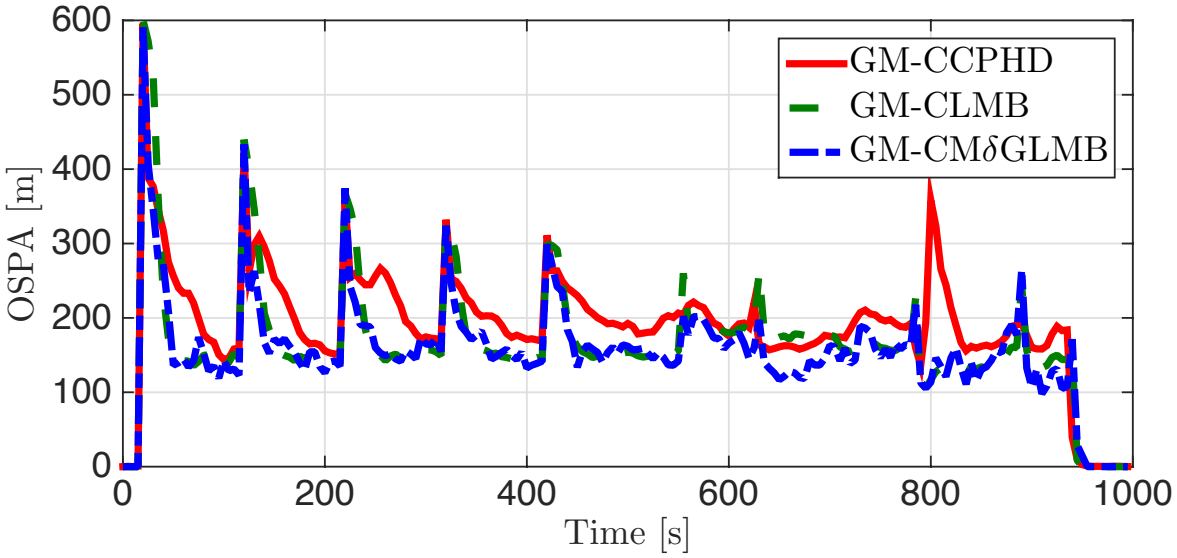


Figure 6.6: OSPA distance ($c = 600\,[m]$, $p = 2$) under high SNR.

distributed algorithm is, indeed, the GM-CMδGLMB and that it exhibits good performance in terms of average number of estimated objects in a very tough scenario with $P_D = 0.7$.

Fig. 6.11 shows the OSPA distance for the current scenario.

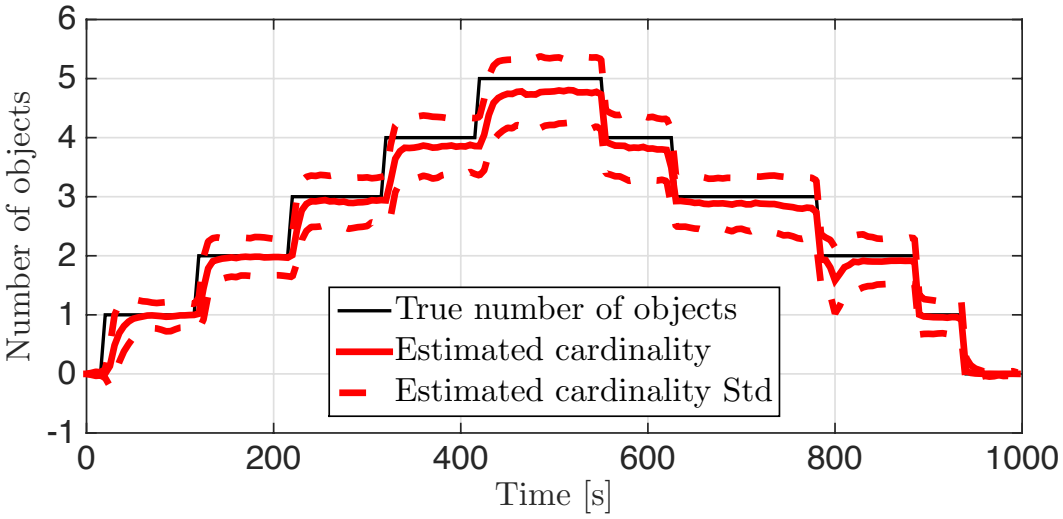


Figure 6.7: Cardinality statistics for the GM-CCPHD filter under low SNR.

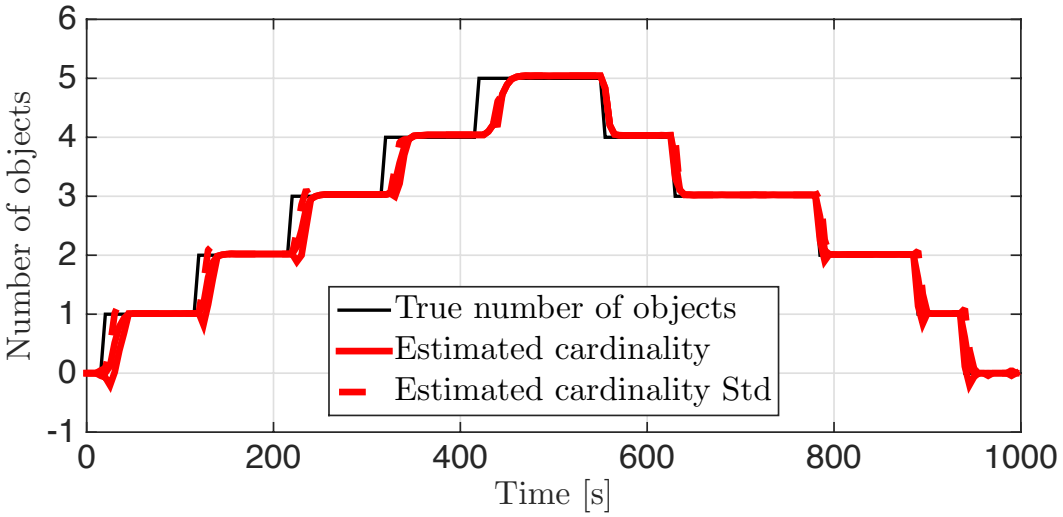


Figure 6.8: Cardinality statistics for the GM-CMδGLMB tracker under low SNR.

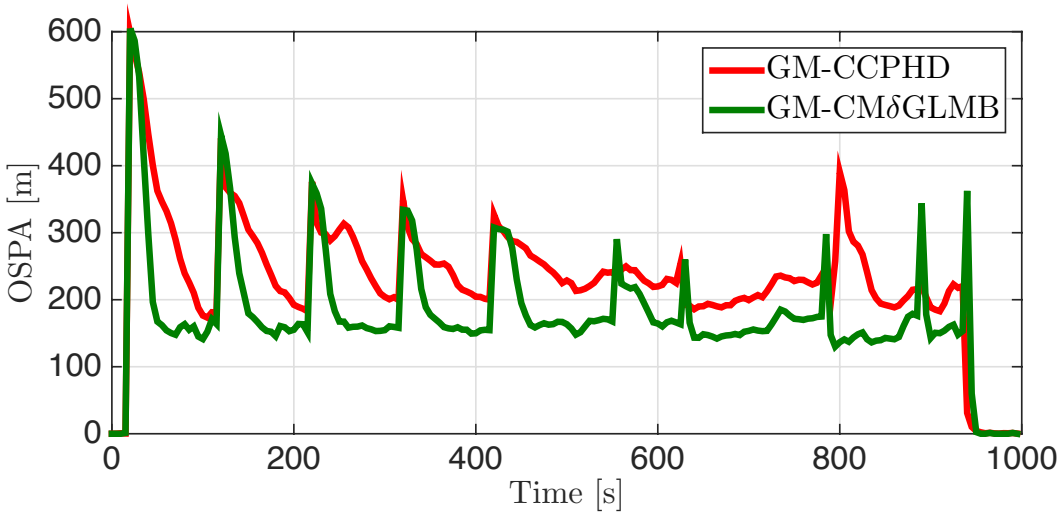


Figure 6.9: OSPA distance ($c = 600\,[m]$, $p = 2$) under low SNR.

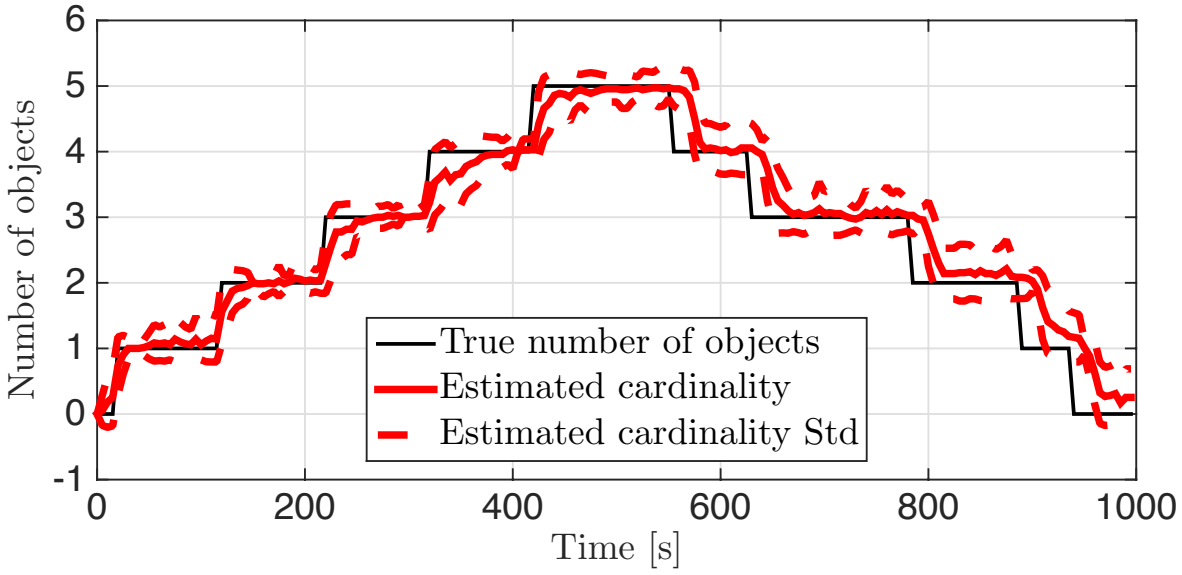


Figure 6.10: Cardinality statistics for the GM-CMδGLMB tracker under low P_D.

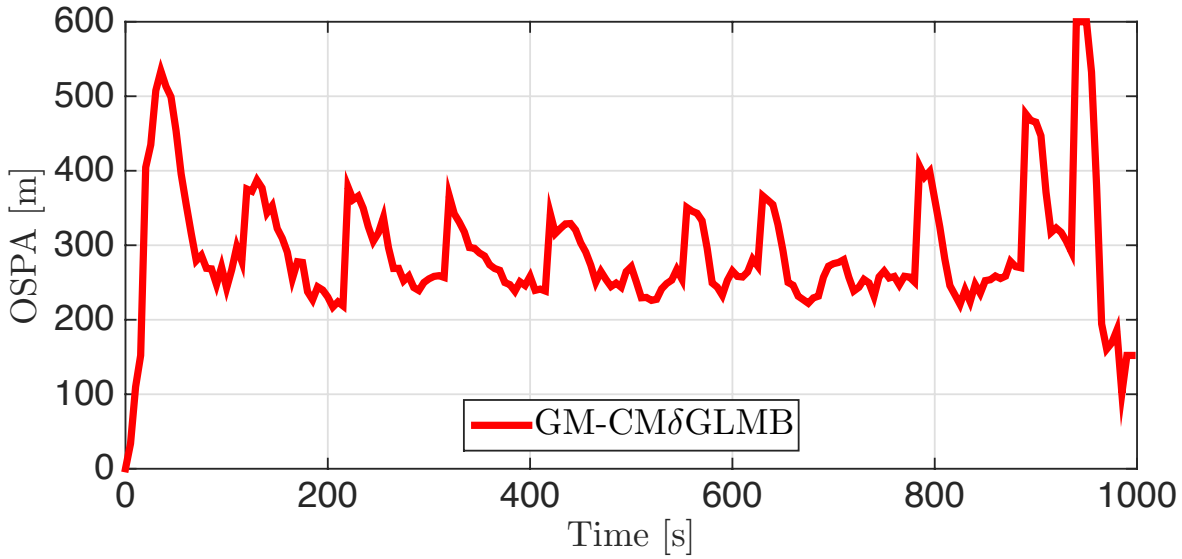


Figure 6.11: OSPA distance ($c = 600\,[m]$, $p = 2$) under low P_D.

7 Conclusions and future work

The present dissertation has summarized the work carried out during my 3-year Ph.D. research work in the field of distributed filtering and tracking. Contributions have spanned from single-object to multi-object scenarios, in both centralized and distributed settings. In particular, distributed tracking over sensor networks requires redesigning the architecture and algorithms to address the following issues:

1. lack of a central fusion node;
2. scalable processing with respect to the network size;
3. each node operates without knowledge of the network network topology;
4. each node operates without knowledge of the dependence between its own information and the information received from other nodes.

Four main ingredients have been considered in order to provide novel research contributions.

I. The Bayesian framework for modeling the quantities to be estimated as random variables characterized by Probability Density Functions (PDFs), and for providing an improved estimation of such quantities by conditioning the PDF on the available noisy measurements.

II. The Random Finite Set (RFS) formulation as it provides the concept of probability density of the multi-object state that allows us to directly generalize single-object estimation to the multi-object case.

III. The Kullback-Leibler Average (KLA) as a sensible information-theoretic definition for consistently fusing PDFs of the quantity to be estimated provided by different nodes (or agents).

IV. Consensus as a tool for developing scalable and reliable distributed fusion techniques over a network where each agent aims to compute the collective average of a given quantity by iterative regional averages, where the terms “collective” and “regional” mean “over all network nodes” and, respectively, “over neighboring nodes only”.

The main contributions of the present thesis are herewith summarized.

- A novel distributed single-object filter, namely *Parallel Consensus on Likelihoods and Priors* (CLCP). The proposed algorithm is based on the idea of carrying out, in parallel, a separate consensus for the novel information (likelihoods) and one for the prior information (priors). This parallel procedure is conceived as an improvement of the *Consensus on Posteriors* (CP) approach to avoid underweighting the novel information during the fusion steps. The outcomes of the two consensuses are then combined to provide the fused posterior density.

- Two novel consensus *Multiple Model* (MM) filters to be used for tracking a maneuvering object, namely *Distributed First Order Generalized Pseudo-Bayesian* (DGPB_1) and *Distributed Interacting Multiple Model* (DIMM) filters.

- A novel result on *Distributed Multi-Object Filtering* (DMOF) over a sensor network using RFSs. Specifically, a generalisation of the single-object *Kullback-Leibler Average* (KLA) to the RFS framework has been devised.

- A novel consensus multi-object filter, the *Consensus Cardinalized Probability Hypothesis Density* (CCPHD) filter. Each tracking agent locally updates multi-object CPHD, i.e. the cardinality distribution and the PHD, exploiting the multi-object dynamics and the available local measurements, exchanges such information with communicating agents and then carries out a fusion step to combine the information from all neighboring agents.

- A novel approximation of the δ-GLMB filter, namely the *Marginalized δ-Generalised Labeled Multi-Bernoulli* (Mδ-GLMB). The result is based on a principled *Generalised Labeled Multi-Bernoulli* (GLMB) approximation of the labeled RFS posterior that matches exactly the posterior PHD and cardinality distribution. The proposed approximation can be interpreted as performing a marginalization with respect to the association histories arising from the δ-GLMB filter. The key advantage of the new filter lies in the reduced growth rate of the number of new components generated at each filtering step. In particular, the approximation (or marginalization) step performed after each update is guaranteed to reduce the number of generated components which normally arise from multiple measurement-to-track association maps. Typically, the proposed Mδ-GLMB filter requires much less computation and storage especially in multi-sensor scenarios compared to the δ-GLMB filter. Furthermore the proposed Mδ-GLMB filter inherits the same implementation strategies and parallelizability of the δ-GLMB filter.

- The generalisation of the KLA to labeled RFSs.

- Two novel consensus tracking filters, namely a *Consensus Marginalized δ-Generalized Labeled Multi-Bernoulli* (CMδGLMB) and *Consensus Labeled Multi-Bernoulli* (CLMB)

tracking filter. The proposed algorithms provide fully distributed, scalable and computationally efficient solutions for multi-object tracking.

From the above-mentioned contributions, the following conclusions can be drawn.

- ▷ The KLA turns out to be an effective information-theoretic approach for distributing information over a sensor network.
- ▷ Consensus algorithms can be exploited so as to fuse, in a fully distributed and scalable fashion, the information collected from multiple sensors which are supposed to be heterogeneous and geographically dispersed.
- ▷ The CLCP has proven to be a reliable distributed filter. Its performance is crucially dependant on the choice of a suitable weight of the innovation (likelihood) term combined with the prior. In particular, the choice of such a weight equal to the (usually unknown) number of nodes provides the best performance when the number of consensus steps is sufficiently large, but for very few consensus steps can be improved resorting to other simple choices, proposed in this thesis work, which do not require knowledge of the network size.
- ▷ The DGPB_1 and DIMM have been proven to be capable of tracking a highly-maneuvering object with a network consisting of range-only and angle-only sensors along with communication nodes. DIMM performs significantly better than DGPB_1 when a turn is approaching.
- ▷ The multi-object KLA in the RFS framework turns out to be a significant contribution in the field of distributed multi-object filtering. The proposed consensus algorithm admits an intuitive interpretation in terms of distributed averaging of the local multi-object densities. Its formulation has led to the development of the CCPHD filter which proved its effectiveness in a realistic simulation scenario.
- ▷ The labeled RFS approach turns out to be a strong mathematical tool for modeling the probability density of a multi-object state. Its key features allow to directly generalise, in a rigorous and elegant way, the multi-object KLA to the broader labeled multi-object case.
- ▷ The Mδ-GLMB provides an excellent tradeoff between computational/storage requirements and performance for multi-object tracking. Moreover, it has an appealing mathematical formulation which 1) facilitates an efficient and tractable closed-form solution for distributed multi-object tracking and 2) preserves key summary statistics of the full multi-object posterior.
- ▷ The CLMB fails to track objects in low SNR scenarios. The problem is due to the approximation made to convert a δ-GLMB into an LMB, becoming significant with

low SNR. In particular, each local tracking filter fails to properly set the existence probability of the tracks in the case of having: 1) no local observability, 2) high clutter rate and 3) loss of the full posterior cardinality distribution after the probability density conversion. By having low existence probabilities, the extraction of the tracks fails even if the single object densities are correctly propagated in time.

Based on the above achievements, possible future work will concern the following issues.

- To consider sensors with different field-of-view to cope with real-world sensor capabilities. So far, it has been assumed that the area subjected to surveillance is exactly the intersection of the sensor fields-of-view, or that each sensor field-of-view is exactly the surveillance region. By a practical viewpoint this issue needs to be addressed. Thus, the proposed algorithms will have to take into account a state-dependent probability of detection instead of a constant one as assumed in the present work. Moreover, the fusion techniques will have to be redesigned to consider the above-mentioned context.

- To investigate distributed measurement-driven object initialization. The object birth process of the RFS filtering approach is capable of accounting for a large variety of scenarios in which object birth locations could also be unknown. However, a measurement-driven approach would be much more effective compared to a static birth process. The information provided by the sensors could be used to dynamically accommodate each sensor birth process. Thus, a more complete picture of newborn objects could be achieved by fusing local information with the one provided by the other nodes.

A Proofs

Proof of Theorem 1:
The cost to be minimized by the KLA is

$$\begin{aligned} J(\mu) &= \sum_{i\in\mathcal{N}} \pi^i D_{KL}\left(\mu \parallel \mu^i\right) = \sum_{i\in\mathcal{N}} \pi^i \sum_{j\in\mathcal{R}} \mu^j \log\left(\frac{\mu^j}{\mu^{i,j}}\right) \\ &= \sum_{j\in\mathcal{R}} \mu^j \sum_{i\in\mathcal{N}} \pi^i \log\left(\frac{\mu^i}{\mu^{i,j}}\right) \\ &= \sum_{j\in\mathcal{R}} \mu^j \sum_{i\in\mathcal{N}} \log\left(\frac{\mu^j}{\mu^{i,j}}\right)^{\pi^i} \\ &= \sum_{j\in\mathcal{R}} \mu^j \log\left(\prod_{i\in\mathcal{N}} \frac{(\mu^j)^{\pi^i}}{(\mu^{i,j})^{\pi^i}}\right) \\ &= \sum_{j\in\mathcal{R}} \mu^j \log\left(\frac{\mu^j}{\prod_{i\in\mathcal{N}} \left(\mu^{i,j}\right)^{\pi^i}}\right) \end{aligned}$$

where the relationship $\sum_{i\in\mathcal{N}} \pi^i = 1$ has been exploited. Let $\overline{\mu}$ be defined as in (3.24), then

$$J(\mu) = \sum_{j\in\mathcal{R}} \mu^j \log\left(\frac{\mu^j}{c\,\overline{\mu}^j}\right) \tag{A.1}$$

where $c \triangleq \sum_{h\in\mathcal{R}} \prod_{i\in\mathcal{N}} \left(\mu^{i,h}\right)^{\pi^i}$. Hence

$$J(\mu) = D_{KL}(\mu \parallel \overline{\mu}) - \log c \tag{A.2}$$

where the relationship $\sum_{j\in\mathcal{R}} \mu^j = 1$ has been exploited. Then the infimum of $J(\mu)$, i.e. the weighted KLA of (3.23), is just provided by $\mu = \overline{\mu}$ as in (3.24). □

Proof of Theorem 2:
Let

$$\widetilde{f}(X) = \prod_i \left[f^i(X)\right]^{\omega_i} \tag{A.3}$$

$$c = \int \widetilde{f}(X)\delta X\,, \tag{A.4}$$

so that the fused multi-object density in (4.4) can be expressed as $\overline{f}(X) = \widetilde{f}(X)/c$. Then, the cost to be minimized by the KLA is

$$\begin{aligned}
J(f) &= \sum_i \omega_i D_{KL}\left(f \parallel f^i\right) \\
&= \sum_i \omega_i \int f(X) \log\left(\frac{f(X)}{f^i(X)}\right) \delta X \\
&= \sum_i \omega_i \sum_{n=0}^{\infty} \frac{1}{n!} \int f(\{x_1,\ldots,x_n\}) \log\left(\frac{f(\{x_1,\ldots,x_n\})}{f^i(\{x_1,\ldots,x_n\})}\right) dx_1 \cdots dx_n \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \int f(\{x_1,\ldots,x_n\}) \sum_i \omega_i \log\left(\frac{f(\{x_1,\ldots,x_n\})}{f^i(\{x_1,\ldots,x_n\})}\right) dx_1 \cdots dx_n \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \int f(\{x_1,\ldots,x_n\}) \log\left(\frac{f(\{x_1,\ldots,x_n\})}{\prod_i \left[f^i(\{x_1,\ldots,x_n\})\right]^{\omega_i}}\right) dx_1 \cdots dx_n \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \int f(\{x_1,\ldots,x_n\}) \log\left(\frac{f(\{x_1,\ldots,x_n\})}{\widetilde{f}(\{x_1,\ldots,x_n\})}\right) dx_1 \cdots dx_n \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \int f(\{x_1,\ldots,x_n\}) \log\left(\frac{f(\{x_1,\ldots,x_n\})}{c\cdot\overline{f}(\{x_1,\ldots,x_n\})}\right) dx_1 \cdots dx_n \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \int f(\{x_1,\ldots,x_n\}) \log\left(\frac{f(\{x_1,\ldots,x_n\})}{\overline{f}(\{x_1,\ldots,x_n\})}\right) dx_1 \cdots dx_n + \\
&\quad -\sum_{n=0}^{\infty} \frac{1}{n!} \int f(\{x_1,\ldots,x_n\}) \log(c)\, dx_1 \cdots dx_n \\
&= \int f(X)\, \log\left(\frac{f(X)}{\overline{f}(X)}\right) \delta X - \log(c) \int f(X) \delta X \\
&= D_{KL}\Big(f \parallel \overline{f}\Big) - \log(c)
\end{aligned}$$

Since the KLD is always nonnegative and is zero if and only if its two arguments coincide almost everywhere, the above $J(f)$ is trivially minimized by $f_{KLA}(X)$ defined as in (4.4). □

Proof of Proposition 1:
First, the cardinality distributions of $\boldsymbol{\pi}(\cdot)$ and $\hat{\boldsymbol{\pi}}(\cdot)$ are proven to be the same. Following

(2.81), the cardinality distribution of the GLMB in (5.23) is given by:

$$\hat{\rho}(n) = \sum_{L\in\mathcal{F}_n(\mathbb{L})} \sum_{I\in\mathcal{F}(\mathbb{L})} w^{(I)}(L) \tag{A.5}$$

$$= \sum_{L\in\mathcal{F}_n(\mathbb{L})} w(L) \tag{A.6}$$

$$= \frac{1}{n!} \sum_{(\ell_1,\ldots\ell_n)\in\mathbb{L}^n} w(\{\ell_1,\ldots,\ell_n\}) \tag{A.7}$$

$$= \frac{1}{n!} \sum_{(\ell_1,\ldots\ell_n)\in\mathbb{L}^n} \int \pi\left(\{(x_1,\ell_1),\ldots,(x_n,\ell_n)\}\right) dx_1\cdots dx_n \tag{A.8}$$

$$= \frac{1}{n!} \int \pi\left(\{\mathbf{x}_1,\ldots,\mathbf{x}_n)\}\right) d\mathbf{x}_1\cdots d\mathbf{x}_n \tag{A.9}$$

$$= \rho(n) \tag{A.10}$$

Second, the PHD corresponding to $\boldsymbol{\pi}(\cdot)$ and $\hat{\boldsymbol{\pi}}(\cdot)$ are proven to be the same. Note from the definition of $p^{(I)}(x,\ell)$ that

$$p^{(\{\ell,\ell_1,\ldots,\ell_n\})}(x,\ell) = 1_{\{\ell,\ell_1,\ldots,\ell_n\}}(\ell) p_{\{\ell_1,\ldots,\ell_n\}}(x,\ell) \tag{A.11}$$

$$= \int p(\,\{(x,\ell),(x_1,\ell_1),\ldots,(x_n,\ell_n)\}) dx_1\cdots dx_n \tag{A.12}$$

Substituting the GLMB density (5.23) into the PHD of the object with label (or track) ℓ yields

$$\hat{d}(x,\ell) = \int \boldsymbol{\pi}(\{(x,\ell)\cup\mathbf{X}\})\,\delta\mathbf{X} \tag{A.13}$$

from which

$$\hat{d}(x,\ell) = \int \sum_{I\in\mathcal{F}(\mathbb{L})} w^{(I)}\left(\mathcal{L}(\{(x,\ell)\}\cup\mathbf{X})\right)[p^{(I)}]^{\{(x,\ell)\}\cup\mathbf{X}}\delta\mathbf{X} \tag{A.14}$$

$$= \sum_{I\in\mathcal{F}(\mathbb{L})} \left[\int w^{(I)}\left(\{\ell\}\cup\mathcal{L}(\mathbf{X})\right)[p^{(I)}]^{\mathbf{X}}\delta\mathbf{X}\right] p^{(I)}(x,\ell) \tag{A.15}$$

$$= \sum_{I\in\mathcal{F}(\mathbb{L})} \sum_{L\in\mathcal{F}(\mathbb{L})} w^{(I)}\left(\{\ell\}\cup L\right)\left[p^{(I)}(\cdot,\ell)dx\right]^{L} p^{(I)}(x,\ell) \tag{A.16}$$

where the last step follows from Lemma 3 in [VV13, Section III.B]. Noting that $p^{(I)}(\cdot,\ell)$ is a probability density, and using (5.24) gives

$$\hat{d}(x,\ell) = \sum_{L\in\mathcal{F}(\mathbb{L})} \sum_{I\in\mathcal{F}(\mathbb{L})} \delta_I\left(\{\ell\}\cup L\right) w(I) p^{(I)}(x,\ell) \tag{A.17}$$

$$= \sum_{L\in\mathcal{F}(\mathbb{L})} w\left(\{\ell\}\cup L\right) p^{(\{\ell\}\cup L)}(x,\ell) \tag{A.18}$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{(\ell_1,\ldots,\ell_n)\in\mathbb{L}^n} w\left(\{\ell,\ell_1,\ldots\ell_n\}\right) p^{(\{\ell,\ell_1,\ldots\ell_n\})}(x,\ell) \tag{A.19}$$

from which by applying (A.12), we obtain:

$$\hat{d}(x,\ell) = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{(\ell_1,\ldots,\ell_n)\in\mathbb{L}^n} w\left(\{\ell,\ell_1,\ldots\ell_n\}\right) \int p\left(\{(x,\ell),(x_1,\ell_1),\ldots,(x_n,\ell_n)\}\right) dx_1\cdots dx_n \tag{A.20}$$

$$= \int \boldsymbol{\pi}(\{(x,\ell)\}\cup\mathbf{X})\delta\mathbf{X} \tag{A.21}$$

$$= d(x,\ell)\,. \tag{A.22}$$

□

Proof of Proposition 2:

We apply the result in Proposition 1 which can be used to calculate the parameters of the marginalized δ-GLMB density. Notice that such a result applies to any labeled RFS density and our first step is to rewrite the δ-GLMB density (2.86) in the general form for a labeled RFS density [PVV+14], i.e.

$$\boldsymbol{\pi}(\mathbf{X}) = w(\mathcal{L}(\mathbf{X}))p(\mathbf{X}) \tag{A.23}$$

where

$$w(\{\ell_1,\ldots,\ell_n\}) \triangleq \int_{\mathbb{X}^n} \boldsymbol{\pi}(\{(x_1,\ell_1),\ldots,(x_n,\ell_n)\})d(x_1,\ldots,x_n) \tag{A.24}$$

$$= \sum_{I\in\mathcal{F}(\mathbb{L})} \delta_I(\{\ell_1,\ldots,\ell_n\}) \sum_{\xi\in\Xi} w^{(I,\xi)} \int_{\mathbb{X}^n} p^{(\xi)}(x_1,\ell_1)\cdots p^{(\xi)}(x_n,\ell_n)dx_1\cdots dx_n \tag{A.25}$$

$$= \sum_{\xi\in\Xi} w^{(\{\ell_1,\ldots,\ell_n\},\xi)} \sum_{I\in\mathcal{F}(\mathbb{L})} \delta_I(\{\ell_1,\ldots,\ell_n\}) \tag{A.26}$$

$$= \sum_{\xi\in\Xi} w^{(\{\ell_1,\ldots,\ell_n\},\xi)} \tag{A.27}$$

and

$$p(\{(x_1,\ell_1),\ldots,(x_n,\ell_n)\}) \triangleq \frac{\boldsymbol{\pi}(\{(x_1,\ell_1),\ldots,(x_n,\ell_n)\})}{w(\{\ell_1,\ldots,\ell_n\})} \tag{A.28}$$

$$= \Delta(\{(x_1,\ell_1),\ldots,(x_n,\ell_n)\})\frac{1}{w(\{\ell_1,\ldots,\ell_n\})}\cdot \cdot \sum_{I\in\mathcal{F}(\mathbb{L})} \delta_I\left(\{\ell_1,\ldots,\ell_n\}\right) \sum_{\xi\in\Xi} w^{(I,\xi)} \left[p^{(\xi)}\right]^{\{(x_1,\ell_1),\ldots,(x_n,\ell_n)\}} \tag{A.29}$$

Applying Proposition 1, the parameters $w^{(I)}$ and $p^{(I)}(\cdot)$ for the Mδ-GLMB approximation that match the cardinality and PHD are

$$w^{(I)}(L) = \delta_I(L)w(I) = \delta_I(L)\sum_{\xi\in\Xi} w^{(I,\xi)} \tag{A.30}$$

and

$$p^{(I)}(x,\ell) = 1_I(\ell) p_{I-\{\ell\}}(x,\ell) \tag{A.31}$$

$$= 1_I(\ell) \int p(\{(x,\ell),(x_1,\ell_1),\ldots,(x_j,\ell_j)\}) d(x_1,\ldots,x_j) \tag{A.32}$$

where $I\backslash\{\ell\} = \{\ell_1,\ldots,\ell_j\}$. Substituting the expression (A.29) in (A.32), we have

$$p^{(I)}(x,\ell) = 1_I(\ell)\, \Delta(\{(x,\ell),(x_1,\ell_1),\ldots,(x_j,\ell_j)\})\, \frac{1}{w(\{\ell,\ell_1,\ldots,\ell_j\})} \cdot \cdot \sum_{J\in\mathcal{F}(\mathbb{L})} \delta_J(\{\ell,\ell_1,\ldots,\ell_j\}) \sum_{\xi\in\Xi} w^{(J,\xi)} \int \left[p^{(\xi)}\right]^{\{(x,\ell),(x_1,\ell_1),\ldots,(x_j,\ell_j)\}} d(x_1,\ldots,x_j) \tag{A.33}$$

$$= 1_I(\ell)\, \Delta(\{(x,\ell),(x_1,\ell_1),\ldots,(x_j,\ell_j)\})\, \frac{1}{w(\{\ell,\ell_1,\ldots,\ell_j\})} \cdot \sum_{J\in\mathcal{F}(\mathbb{L})} \delta_J(\{\ell,\ell_1,\ldots,\ell_j\}) \sum_{\xi\in\Xi} w^{(J,\xi)} p^{(\xi)}(x,\ell) \tag{A.34}$$

and, noting that $I = \{\ell,\ell_1,\ldots,\ell_j\}$, it follows that only one term in the sum over J is non-zero thus giving

$$p^{(I)}(x,\ell) = 1_I(\ell)\Delta(\{(x,\ell),(x_1,\ell_1),\ldots,(x_j,\ell_j)\}) \frac{1}{\sum_{\xi\in\Xi} w^{(I,\xi)}} \sum_{\xi\in\Xi} w^{(I,\xi)} p^{(\xi)}(x,\ell) \tag{A.35}$$

Consequently, the Mδ-GLMB approximation is given by

$$\hat{\boldsymbol{\pi}}(\mathbf{X}) = \sum_{I\in\mathcal{F}(\mathbb{L})} w^{(I)}(\mathcal{L}(\mathbf{X})) \left[p^{(I)}\right]^{\mathbf{X}} \tag{A.36}$$

$$= \Delta(\mathbf{X}) \sum_{I\in\mathcal{F}(\mathbb{L})} \delta_I(\mathcal{L}(\mathbf{X})) \sum_{\xi\in\Xi} w^{(I,\xi)} \left[1_I(\cdot) \frac{1}{\sum_{\xi\in\Xi} w^{(I,\xi)}} \sum_{\xi\in\Xi} w^{(I,\xi)} p^{(\xi)}(\cdot,\cdot)\right]^{\mathbf{X}} \tag{A.37}$$

$$= \Delta(\mathbf{X}) \sum_{I\in\mathcal{F}(\mathbb{L})} \delta_I(\mathcal{L}(\mathbf{X})) w^{(I)} \left[p^{(I)}\right]^{\mathbf{X}} \tag{A.38}$$

where

$$w^{(I)} = \sum_{\xi\in\Xi} w^{(I,\xi)} \tag{A.39}$$

$$p^{(I)}(x,\ell) = 1_I(\ell) \frac{1}{\sum_{\xi\in\Xi} w^{(I,\xi)}} \sum_{\xi\in\Xi} w^{(I,\xi)} p^{(\xi)}(x,\ell) \tag{A.40}$$

$$= 1_I(\ell) \frac{1}{w^{(I)}} \sum_{\xi\in\Xi} w^{(I,\xi)} p^{(\xi)}(x,\ell) \tag{A.41}$$

□

Proof of Theorem 4:

For the sake of simplicity, let us consider only two Mδ-GLMB densities $\imath \in \{1,2\}$. From (6.6) one gets

$$\overline{\boldsymbol{\pi}}(\mathbf{X}) = \frac{1}{K}\left[\Delta(\mathbf{X}) \sum_{L\in\mathcal{F}(\mathbb{L})} \delta_L(\mathcal{L}(\mathbf{X}))\; w_1^{(L)} \left[p_1^{(L)}\right]^{\mathbf{X}}\right]^{\omega} \left[\Delta(\mathbf{X}) \sum_{L\in\mathcal{F}(\mathbb{L})} \delta_L(\mathcal{L}(\mathbf{X}))\; w_2^{(L)} \left[p_2^{(L)}\right]^{\mathbf{X}}\right]^{1-\omega}. \tag{A.42}$$

Notice that the exponentiation of a sum of delta functions is a sum of the exponentiated delta function terms, i.e

$$\overline{\boldsymbol{\pi}}(\mathbf{X})^{\omega} = \begin{cases} \Delta(\mathbf{X}) \left(w^{(L_1)}\right)^{\omega} \left[\left(p^{(L_1)}\right)^{\omega}\right]^{\mathbf{X}} & \text{if } \mathcal{L}(\mathbf{X}) = L_1 \\ \vdots & \vdots \\ \Delta(\mathbf{X}) \left(w^{(L_n)}\right)^{\omega} \left[\left(p^{(L_n)}\right)^{\omega}\right]^{\mathbf{X}} & \text{if } \mathcal{L}(\mathbf{X}) = L_n \end{cases}$$

Thus, (A.42) yields:

$$\overline{\boldsymbol{\pi}}(\mathbf{X}) = \frac{\Delta(\mathbf{X})}{K} \sum_{L\in\mathcal{F}(\mathbb{L})} \delta_L(\mathcal{L}(\mathbf{X}))\; \left(w_1^{(L)}\right)^{\omega} \left(w_2^{(L)}\right)^{1-\omega} \left[\left(p_1^{(L)}\right)^{\omega} \left(p_2^{(L)}\right)^{1-\omega}\right]^{\mathbf{X}} \tag{A.43}$$

$$= \frac{\Delta(\mathbf{X})}{K} \sum_{L\in\mathcal{F}(\mathbb{L})} \delta_L(\mathcal{L}(\mathbf{X}))\; \left(w_1^{(L)}\right)^{\omega} \left(w_2^{(L)}\right)^{1-\omega} \left[\int \left(p_1^{(L)}(x,\cdot)\right)^{\omega} \left(p_2^{(L)}(x,\cdot)\right)^{1-\omega}\right]^{L} \cdot \cdot \left[\left(\omega \odot p_1^{(L)}\right) \oplus \left((1-\omega) \odot p_2^{(L)}\right)\right]^{\mathbf{X}}. \tag{A.44}$$

The normalization constant K can be easily evaluated exploiting Lemma 3 of [VV13, Section III.B], i.e.

$$K = \int \Delta(\mathbf{X}) \sum_{L\in\mathcal{F}(\mathbb{L})} \delta_L(\mathcal{L}(\mathbf{X}))\; \left(w_1^{(L)}\right)^{\omega} \left(w_2^{(L)}\right)^{1-\omega} \left[\left(p_1^{(L)}\right)^{\omega} \left(p_2^{(L)}\right)^{1-\omega}\right]^{\mathbf{X}} \delta X \tag{A.45}$$

$$= \sum_{L\subseteq\mathbb{L}} \left(w_1^{(L)}\right)^{\omega} \left(w_2^{(L)}\right)^{1-\omega} \left[\int \left(p_1^{(L)}\right)^{\omega} \left(p_2^{(L)}(x,\cdot)\right)^{1-\omega} dx\right]^{L}. \tag{A.46}$$

Applying (A.46) in (A.44) one has

$$\overline{p}^{(L)} = \left[\left(\omega \odot p_1^{(L)}\right) \oplus \left((1-\omega) \odot p_2^{(L)}\right)\right]^{\mathbf{X}} \tag{A.47}$$

$$\overline{w}^{(L)} = \frac{\left(w_1^{(L)}\right)^{\omega} \left(w_2^{(L)}\right)^{1-\omega} \left[\int \left(p_1^{(L)}(x,\cdot)\right)^{\omega} \left(p_2^{(L)}(x,\cdot)\right)^{1-\omega}\right]^{L}}{\sum_{F\subseteq\mathbb{L}} \left(w_1^{(F)}\right)^{\omega} \left(w_2^{(F)}\right)^{1-\omega} \left[\int \left(p_1^{(F)}(x,\cdot)\right)^{\omega} \left(p_2^{(F)}(x,\cdot)\right)^{1-\omega}\right]^{F}} \tag{A.48}$$

It can be proved by induction that Theorem 4 holds considering I Mδ-GLMB densities instead

of 2. □

Proof of Proposition 3:
The proof readily follows by noting that the KLA (6.5) can be evaluated using (6.6) of Theorem 4. □

The following Lemma is useful to prove Theorem 5.

Lemma 1 (Normalization constant of LMB KLA)**.**
Let $\boldsymbol{\pi}^1(\mathbf{X}) = \left\{\left(r_1^{(\ell)}, p_1^{(\ell)}\right)\right\}_{\ell\in\mathbb{L}}$ *and* $\boldsymbol{\pi}^2(\mathbf{X}) = \left\{\left(r_2^{(\ell)}, p_2^{(\ell)}\right)\right\}_{\ell\in\mathbb{L}}$ *be two LMB densities on* $\mathbb{X}\times\mathbb{L}$ *and* $\omega \in (0,1)$ *then*

$$K \triangleq \int \boldsymbol{\pi}^1(\mathbf{X})^{\omega}\boldsymbol{\pi}^2(\mathbf{X})^{1-\omega}\delta\mathbf{X} \tag{A.49}$$

$$= \left\langle \omega \odot \boldsymbol{\pi}^1, (1-\omega)\odot\boldsymbol{\pi}^2 \right\rangle \tag{A.50}$$

$$= \left(\widetilde{q}^{(\cdot)} + \widetilde{r}^{(\cdot)}\right)^{\mathbb{L}}. \tag{A.51}$$

Proof of Lemma 1:
We make use of the Binomial Theorem [AS64] which states

$$\sum_{L\subseteq\mathbb{L}} f^L = (1+f)^{\mathbb{L}} \tag{A.52}$$

Applying Lemma 3 of [VV13, Section III.B] to the definition (A.49) of K gives

$$K = \left(\widetilde{q}^{(\cdot)}\right)^{\mathbb{L}} \sum_{L\subseteq\mathbb{L}} \left(\frac{\widetilde{r}^{(\cdot)}}{\widetilde{q}^{(\cdot)}}\right)^{L}. \tag{A.53}$$

Applying (A.52) to (A.53) gives

$$K = \left(\widetilde{q}^{(\cdot)}\right)^{\mathbb{L}} \left(1 + \frac{\widetilde{r}^{(\cdot)}}{\widetilde{q}^{(\cdot)}}\right)^{\mathbb{L}} \tag{A.54}$$

$$= \left(\widetilde{q}^{(\cdot)} + \widetilde{r}^{(\cdot)}\right)^{\mathbb{L}} \tag{A.55}$$

having defined

$$\widetilde{r}^{(\ell)} \triangleq \int \left(r_1^{(\ell)} p_1^{(\ell)}(x)\right)^{\omega} \left(r_2^{(\ell)} p_2^{(\ell)}(x)\right)^{1-\omega} dx, \tag{A.56}$$

$$\widetilde{q}^{(\ell)} \triangleq \left(1 - r_1^{(\ell)}\right)^{\omega} \left(1 - r_2^{(\ell)}\right)^{1-\omega}. \tag{A.57}$$

□

Proof of Theorem 5:
For the sake of simplicity, let us consider only two LMB densities $\imath \in \{1,2\}$. From (6.13) one

gets

$$\overline{\boldsymbol{\pi}}(\mathbf{X}) = \frac{1}{K}\left[\Delta(\mathbf{X})\, w_1(\mathcal{L}(\mathbf{X}))p_1^{\mathbf{X}}\right]^{\omega}\left[\Delta(\mathbf{X})\, w_2(\mathcal{L}(\mathbf{X}))p_2^{\mathbf{X}}\right]^{1-\omega} \tag{A.58}$$

$$= \frac{\Delta(\mathbf{X})}{K}\left(\widetilde{q}^{(\cdot)}\right)^{\mathbb{L}}\left[1_{\mathbb{L}}(\cdot)\left(\frac{r_1^{(\cdot)}}{1-r_1^{(\cdot)}}\right)^{\omega}\left(\frac{r_2^{(\cdot)}}{1-r_2^{(\cdot)}}\right)^{1-\omega}\right]^{\mathcal{L}(\mathbf{X})}\left(p_1^{\omega}\, p_2^{1-\omega}\right)^{\mathbf{X}} \tag{A.59}$$

$$= \frac{\Delta(\mathbf{X})}{K}\left(\widetilde{q}^{(\cdot)}\right)^{\mathbb{L}}\left[1_{\mathbb{L}}(\cdot)\left(\frac{r_1^{(\cdot)}}{1-r_1^{(\cdot)}}\right)^{\omega}\left(\frac{r_2^{(\cdot)}}{1-r_2^{(\cdot)}}\right)^{1-\omega}\int p_1^{\omega}\, p_2^{1-\omega}dx\right]^{\mathcal{L}(\mathbf{X})}\cdot$$
$$\cdot\left[(\omega\odot p_1)\oplus((1-\omega)\odot p_2)\right]^{\mathbf{X}} \tag{A.60}$$

Hence, recalling definitions (A.56) and (A.57), one has

$$\overline{w}(L) = \frac{\left(\widetilde{q}^{(\cdot)}\right)^{\mathbb{L}}\left(1_{\mathbb{L}}(\cdot)\,\dfrac{\widetilde{r}^{(\cdot)}}{\widetilde{q}^{(\cdot)}}\right)^{L}}{\left(\widetilde{q}^{(\cdot)}+\widetilde{r}^{(\cdot)}\right)^{\mathbb{L}}} \tag{A.61}$$

$$= \left(\frac{\widetilde{q}^{(\cdot)}}{\widetilde{q}^{(\cdot)}+\widetilde{r}^{(\cdot)}}\right)^{\mathbb{L}\setminus L}\left(\frac{1_{\mathbb{L}}(\cdot)\,\widetilde{r}^{(\cdot)}}{\widetilde{q}^{(\cdot)}+\widetilde{r}^{(\cdot)}}\right)^{L} \tag{A.62}$$

$$= \left(\frac{\widetilde{q}^{(\cdot)}}{\widetilde{q}^{(\cdot)}+\widetilde{r}^{(\cdot)}}\right)^{\mathbb{L}}\left(1_{\mathbb{L}}(\cdot)\,\frac{\widetilde{r}^{(\cdot)}}{\widetilde{q}^{(\cdot)}}\right)^{L}, \tag{A.63}$$

$$\overline{p}^{(\ell)}(x) = \left[\left(\omega\odot p_1^{(\ell)}\right)\oplus\left((1-\omega)\odot p_2^{(\ell)}\right)\right]^{\mathbf{X}}. \tag{A.64}$$

It follows from (A.62) that $\overline{q}^{(\ell)}+\overline{r}^{(\ell)}=1$, $\forall\,\ell\in\mathbb{L}$, where

$$\overline{r}^{(\ell)} = \frac{\widetilde{r}^{(\ell)}}{\widetilde{q}^{(\ell)}+\widetilde{r}^{(\ell)}} \tag{A.65}$$

$$\overline{q}^{(\ell)} = \frac{\widetilde{q}^{(\ell)}}{\widetilde{q}^{(\ell)}+\widetilde{r}^{(\ell)}}. \tag{A.66}$$

It can be proved by induction that Theorem 5 holds considering I LMB densities instead of 2. □

Proof of Proposition 4:

The proof readily follows by noting that the KLA (6.5) can be evaluated using (6.13) of Theorem 5. □

Bibliography

[AF70] G. Ackerson and K. S. Fu, "On state estimation in switching environments," *IEEE Trans. on Automatic Control*, vol. 15, no. 1, pp. 10–17, 1970.

[Aka98] H. Akaike, "Information theory and the extension of the maximum likelihood principle," in *Selected Papers of Hirotugu Akaike*. Springer New York, 1998, pp. 199–213.

[AM12] B. Anderson and J. B. Moore, *Optimal filtering*. Courier Dover Publications, 2012.

[AS64] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*. Washington, DC: National Bureau of Standards, 1964. [Online]. Available: http://people.math.sfu.ca/~cbm/aands/abramowitz_and_stegun.pdf

[BBS88] H. Blom and Y. Bar-Shalom, "The interacting multiple model algorithm for systems with Markovian switching coefficients," *IEEE Trans. on Automatic Control*, vol. 33, no. 8, pp. 780–783, 1988.

[BC14] G. Battistelli and L. Chisci, "Kullback-Leibler average, consensus on probability densities, and distributed state estimation with guaranteed stability," *Automatica*, vol. 50, no. 3, pp. 707–718, 2014.

[BCF+13a] G. Battistelli, L. Chisci, C. Fantacci, A. Farina, and A. Graziano, "Consensus CPHD filter for distributed multitarget tracking," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 3, pp. 508–520, 2013.

[BCF+13b] ——, "Distributed multitarget tracking with range-Doppler sensors," in *Proc. 21^{st} European Signal Processing Conference (EUSIPCO)*, pp. 1-5, Marrakech, Morocco, 2013.

[BCF+13c] ——, "Distributed multitarget tracking for passive multireceiver radar systems," in *Proc. 14^{th} International Radar Symposium (IRS)*, pp. 337-342, Dresden, Germany, 2013.

[BCF+13d] ——, "Distributed tracking with Doppler sensors," in *Proc. 52^{nd} IEEE Conf. on Decision and Control (CDC)*, pp. 4760-4765, Florence, Italy, 2013.

[BCF14a] G. Battistelli, L. Chisci, and C. Fantacci, "Parallel consensus on likelihoods and priors for networked nonlinear filtering," *IEEE Signal Processing Letters*, vol. 21, no. 7, pp. 787–791, 2014.

[BCF+14b] G. Battistelli, L. Chisci, C. Fantacci, A. Farina, and A. Graziano, "Consensus-based multiple-model Bayesian filtering for distributed tracking," *IET Radar, Sonar & Navigation*, DOI: 10.1049/iet-rsn.2014.0071, 2014.

[BCF+14c] G. Battistelli, L. Chisci, C. Fantacci, N. Forti, A. Farina, and A. Graziano, "Distributed peer-to-peer multitarget tracking with association-based track fusion," in *Proc. 17th Int. Conf. on Information Fusion (FUSION)*, Salamanca, Spain, 2014.

[BCMP11] G. Battistelli, L. Chisci, S. Morrocchi, and F. Papi, "An information-theoretic approach to distributed state estimation," in *Proc. 18th IFAC World Congress*, pp. 12477-12482, Milan, Italy, 2011.

[BJA12] T. Bailey, S. Julier, and G. Agamennoni, "On conservative fusion of information with unknown non-Gaussian dependence," in *Proc. 15th Int. Conf. on Information Fusion (FUSION)*, pp. 1876-1883, Singapore , 2012.

[BP99] S. Blackman and R. Popoli, *Design and analysis of modern tracking systems.* Artech House, Norwood, MA, USA, 1999.

[BSF88] Y. Bar-Shalom and T. Fortmann, *Tracking and data association.* Academic Press, San Diego, CA, USA, 1988.

[BSL95] Y. Bar-Shalom and X. Li, *Multitarget-multisensor tracking: principles and techniques.* YBS Publishing, Storrs, CT, USA, 1995.

[BSLK01] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with applications to tracking and navigation.* John Wiley & Sons, Hoboken, NJ, USA, 2001.

[BVV14] M. Beard, B.-T. Vo, and B.-N. Vo, "Multi-target tracking with merged measurements using labelled random finite sets," in *Proc. 17th Int. Conf. on Information Fusion (FUSION)*, Salamanca, Spain, 2014.

[CA09] G. C. Calafiore and F. Abrate, "Distributed linear estimation over sensor networks," *International Journal of Control*, vol. 82, no. 5, pp. 868–882, 2009.

[Cam70] L. Campbell, "Equivalence of Gauss's principle and minimum discrimination information estimation of probabilities," *The Annals of Mathematical Statistics*, vol. 41, no. 3, pp. 1011–1015, 1970.

[CCM10] K. C. Chang, C.-Y. Chong, and S. Mori, "Analytical and computational evaluation of scalable distributed fusion algorithms," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 46, no. 4, pp. 2022–2034, 2010.

[CCSZ08] R. Carli, A. Chiuso, L. Schenato, and S. Zampieri, "Distributed Kalman filtering based on consensus strategies," *IEEE Journal on Selected Areas in Communications*, vol. 26, pp. 622–633, 2008.

[CJMR10] D. Clark, S. Julier, R. Mahler, and B. Ristić, "Robust multi-object fusion with unknown correlation," in *Proc. 2010 Sensor Signal Processing for Defense Conference (SSPD 2010)*, London, UK, 2010.

[CMC90] C.-Y. Chong, S. Mori, and K.-C. Chang, "Distributed multitarget multisensor tracking," in *Multitarget-Multisensor Tracking: Advanced Applications*, Y. Bar-Shalom, Ed. Artech House, Norwood, MA, USA, 1990, ch. 8.

[Coa04] M. Coates, "Distributed particle filters for sensor networks," in *Proc. 3rd International Symposium on Information Processing in Sensor Networks (IPSN'04)*, pp. 99-107, New York, NY, USA, 2004.

[CS10] F. S. Cattivelli and A. Sayed, "Diffusion strategies for distributed Kalman filtering and smoothing," *IEEE Trans. on Automatic Control*, vol. 55, pp. 2069–2084, 2010.

[CT12] T. M. Cover and J. A. Thomas, *Elements of information theory.* John Wiley & Sons, Hoboken, NJ, USA, 2012.

[FBC+12] C. Fantacci, G. Battistelli, L. Chisci, A. Farina, and A. Graziano, "Multiple-model algorithms for distributed tracking of a maneuvering target," in *Proc. 15th Int. Conf. on Information Fusion (FUSION)*, Singapore, 2012, pp. 1028–1035.

[FFTS10] M. Farina, G. Ferrari-Trecate, and R. Scattolini, "Distributed moving horizon estimation for linear constrained systems," *IEEE Trans. on Automatic Control*, vol. 55, no. 11, pp. 2462–2475, 2010.

[Fis12] R. A. Fisher, "On an absolute criterion for fitting frequency curves," *Messenger of Math.*, pp. 41–155, 1912.

[FS85] A. Farina and F. A. Studer, *Radar data processing, vol. I: introduction and tracking.* Research Studies Press, Letchworth, Hertfordshire, England, 1985.

[FS86] ——, *Radar data processing, vol. II: advanced topics and applications.* Research Studies Press, Letchworth, Hertfordshire, England, 1986.

[FSU09] D. Franken, M. Schmidt, and M. Ulmke, ""Spooky action at a distance" in the cardinalized probability hypothesis density filter," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 45, no. 4, pp. 1657–1664, 2009.

[FVPV15] C. Fantacci, B.-T. Vo, F. Papi, and B.-N. Vo, "The marginalized δ-GLMB filter," *preprint available online at arXiv:1501.00926*, 2015.

[FVV+15] C. Fantacci, B.-N. Vo, B.-T. Vo, G. Battistelli, and L. Chisci, "Consensus labeled random finite set filtering for distributed multi-object tracking," *preprint available online at arXiv:1501.01579*, 2015.

[Gau04] F. C. Gauss, *Theoria Motus Corporum Coelestium in Sectionibus Conicis Solem Ambientium.* Hamburg, 1809. (Translation of Theoria Motus: *Theory of the Motion of the Heavenly Bodies Moving About the Sun in Conic Sections.* Dover Phoenix ed., 2004).

[GMN97] I. Goodman, R. P. S. Mahler, and H. Nguyen, *Mathematics of Data Fusion.* Kluwer Academic Publishers, Norwell, MA, USA, 1997.

[GSW09] R. Georgescu, S. Schoenecker, and P. Willett, "GM-CPHD and MLPDA applied to the SEABAR07 and TNO-blind multi-static sonar data," in *Proc. 12th Int. Conf. on Information Fusion (FUSION)*, pp. 1851-1858, Seattle, WA, USA, 2009.

[HL64] Y. Ho and R. Lee, "A Bayesian approach to problems in stochastic estimation and control," *IEEE Trans. on Automatic Control*, vol. 9, no. 4, pp. 333–339, 1964.

[HSH+12] O. Hlinka, O. Slučiak, F. Hlawatsch, P. M. Djurić, and M. Rapp, "Likelihood consensus and its application to particle filtering," *IEEE Trans. on Signal Processing*, vol. 60, no. 8, pp. 4334–4349, 2012.

[Hur02] M. Hurley, "An information theoretic justification for covariance intersection and its generalization," in *Proc. 5th Int. Conf. on Information Fusion (FUSION)*, pp. 505-511, Annapolis, MD, USA, 2002.

[Jay03] E. T. Jaynes, *Probability Theory: The Logic of Science.* Cambridge University Press, 2003.

[Jaz07] A. H. Jazwinski, *Stochastic processes and filtering theory.* Courier Dover Publications, 2007.

[JBU06] S. J. Julier, T. Bailey, and J. K. Uhlmann, "Using exponential mixture models for suboptimal distributed data fusion," in *Proc. 2006 IEEE Nonlinear Stat. Signal Proc. Workshop (NSSPW'06)*, pp. 160-163, Cambridge, UK, 2006.

[JU97] S. J. Julier and J. K. Uhlmann, "A non-divergent estimation algorithm in the presence of unknown correlations," in *Proc. IEEE American Control Conference (ACC 1997)*, pp. 2369-2373, Albuquerque, NM, 1997.

[JU04] ——, "Unscented filtering and nonlinear estimation," *Proc. of the IEEE*, vol. 92, no. 3, pp. 401–422, 2004.

[JUDW95] S. J. Julier, J. K. Uhlmann, and H. F. Durrant-Whyte, "A new approach for filtering nonlinear systems," in *Proc. 1995 American Control Conference (ACC)*, 1995, pp. 1628–1632, Seattle, WA, USA, 1995.

[Jul06] S. J. Julier, "An empirical study into the use of Chernoff information for robust, distributed fusion of Gaussian mixture models," in *Proc. 9th Int. Conf. on Information Fusion (FUSION)*, Florence, Italy, 2006.

[Jul08] ——, "Fusion without independence," *2008 IET Seminar on Target Tracking and Data Fusion: Algorithms and Applications*, 2008.

[Kal60] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Trans. ASME - Journal of Basic Engineering*, vol. 82 (Series D), no. 1, pp. 35–45, 1960.

[KB61] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *Trans. ASME - Journal of Basic Engineering*, vol. 83 (Series D), no. 1, pp. 95–108, 1961.

[Kol50] A. N. Kolmogorov, *Foundations of the Theory of Probability.* New York, Chelsea Pub. Co., 1950.

[KT08] M. Kamgarpour and C. Tomlin, "Convergence properties of a decentralized Kalman filter," in *Proc. 47th IEEE Conf. on Decision and Control (CDC)*, pp. 3205-3210, Cancun, Mexico, 2008.

[Lee08] D. J. Lee, "Nonlinear estimation and multiple sensor fusion using unscented information filtering," *IEEE Signal Processing Letters*, vol. 15, no. 8, pp. 861–864, 2008.

[LJ05] X. R. Li and V. Jilkov, "Survey of maneuvering target tracking. Part V: Multiple-model methods," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 41, no. 4, pp. 1255–1321, 2005.

[LJ12] W. Li and Y. Jia, "Consensus-based distributed multiple model UKF for jump Markov nonlinear systems," *IEEE Trans. on Automatic Control*, vol. 57, no. 1, pp. 227–233, 2012.

[Mah36] P. C. Mahalanobis, "On the generalised distance in statistics," *Proc. National Institute of Sciences of India*, vol. 2, no. 1, pp. 49–55, 1936.

[Mah00] R. P. S. Mahler, "Optimal/robust distributed data fusion: a unified approach," *Proc. SPIE Defense and Security Symposium 2000*, 2000.

[Mah03] ——, "Multitarget Bayes filtering via first-order multitarget moments," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 39, no. 4, pp. 1152–1178, 2003.

[Mah04] ——, ""Statistics 101" for multisensor, multitarget data fusion," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 19, no. 1, pp. 53–64, 2004.

[Mah07a] ——, "PHD filters of higher order in target number," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 43, no. 4, pp. 1523–1543, 2007.

[Mah07b] ——, *Statistical multisource-multitarget information fusion.* Artech House, Norwood, MA, 2007.

[Mah12] ——, "Toward a theoretical foundation for distributed fusion," in *Distributed Data Fusion for Network-Centric Operations*, D. Hall, M. L. II, C.-Y. Chong, and J. Llinas, Eds. CRC Press, Boca Raton, FL, USA, 2012, ch. 8.

[Mah13] ——, ""Statistics 102" for multisource-multitarget detection and tracking," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 3, pp. 376–389, 2013.

[Mah14] ——, *Advances in Statistical Multisource-Multitarget Information Fusion.* Artech House, Norwood, MA, 2014.

[Mat75] G. Matheron, *Random sets and integral geometry.* John Wiley & Sons, Hoboken, NJ, USA, 1975.

[MSS62] L. A. McGee, S. F. Schmidt, and G. L. Smith, "Applications of statistical filter theory to the optimal estimation of position and velocity on board a circumlunar vehicle," NASA Technical Report R-135, Tech. Rep., 1962.

[OC10] B. N. Oreshkin and M. J. Coates, "Asynchronous distributed particle filter via decentralized evaluation of Gaussian products," in *Proc. 13th Int. Conf. on Information Fusion (FUSION)*, Edinburgh, UK, 2010.

[OHD14] F. H. O. Hlinka and P. M. Djuric, "Consensus-based distributed particle filtering with distributed proposal adaptation," *IEEE Trans. on Signal Processing*, vol. 62, no. 12, pp. 3029–3041, 2014.

[OS07] R. Olfati-Saber, "Distributed Kalman filtering for sensor networks," in *Proc. 46th IEEE Conf. on Decision and Control (CDC)*, pp. 5492-5498, New Orleans, LA, USA, 2007.

[OSFM07] R. Olfati-Saber, J. Fax, and R. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 49–54, 2007.

[PVV+14] F. Papi, B.-N. Vo, B.-T. Vo, C. Fantacci, and M. Beard, "Generalized labeled multi-Bernoulli approximation of multi-object densities," *preprint available online at arXiv:1412.5294*, 2014.

[RAG04] B. Ristić, S. Arulampalm, and N. J. Gordon., *Beyond the Kalman filter: Particle filters for tracking applications.* Artech House Publishers, 2004.

[Rei79] D. Reid, "An algorithm for tracking multiple targets," *IEEE Trans. on Automatic Control*, vol. 24, no. 6, pp. 843–854, 1979.

[RVVD14] S. Reuter, B.-T. Vo, B.-N. Vo, and K. Dietrmayer, "The labeled multi-Bernoulli filter," *IEEE Trans. on Signal Processing*, vol. 62, no. 12, pp. 3246–3260, 2014.

[Sal88] D. J. Salmond, "Mixture reduction algorithms for uncertain tracking," Tech. Rep. 88004, Royal Aerospace Establishment, Farnborough, UK, DTIC Number ADA197641, Tech. Rep., 1988.

[Sal90] ——, "Mixture reduction algorithms for target tracking in clutter," in *SPIE Signal and Data Processing of Small Targets*, pp. 434-445, Los Angeles, CA, USA, 1990.

[SKM95] D. Stoyan, D. Kendall, and J. Mecke, *Stochastic Geometry and its Applications.* John Wiley & Sons, Hoboken, NJ, USA, 1995.

[SSS09] S. S. Stankovic, M. Stankovic, and D. Stipanovic, "Consensus based overlapping decentralized estimation with missing observations and communication faults," *Automatica*, vol. 45, no. 6, pp. 1397–1406, 2009.

[Str60] R. L. Stratonovich, "Conditional markov processes," *Theory of Probability & Its Applications*, vol. 5, no. 2, pp. 156–178, 1960.

[SVV08] D. Schuhmacher, B.-T. Vo, and B.-N. Vo, "A consistent metric for performance evaluation of multi-object filters," *IEEE Trans. on Signal Processing*, vol. 56, no. 8, pp. 3447–3457, 2008.

[SWS09] D. Svensson, J. Wintenby, and L. Svensson, "Performance evaluation of MHT and GM-CPHD in a ground target tracking scenario," in *Proc. 12th Int. Conf. on Information Fusion (FUSION)*, pp. 300-307, Seattle, WA, USA, 2009.

[Tre04] H. L. V. Trees, *Detection, estimation, and modulation theory.* John Wiley & Sons, Hoboken, NJ, USA, 2004.

[UCJ11] M. Uney, D. Clark, and S. Julier, "Information measures in distributed multitarget tracking," in *Proc. 14th Int. Conf. on Information Fusion (FUSION)*, Chicago, IL, USA, 2011.

[UCJ13] ——, "Distributed fusion of phd filters via exponential mixture densities," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 3, pp. 521–531, 2013.

[UJCR10] M. Uney, S. Julier, D. Clark, and B. Ristić, "Monte Carlo realisation of a distributed multi-object fusion algorithm," in *Proc. 2010 Sensor Signal Processing for Defense Conference (SSPD 2010)*, London, UK, 2010.

[VM06] B.-N. Vo and W. Ma, "The Gaussian mixture probability hypothesis density filter," *IEEE Trans. on Signal Processing*, vol. 54, no. 11, pp. 4091–4104, 2006.

[VSD05] B.-N. Vo, S. Singh, and A. Doucet, "Sequential Monte Carlo methods for multitarget filtering with random finite sets," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 41, no. 4, pp. 1224–1245, 2005.

[VV13] B.-T. Vo and B.-N. Vo, "Labeled random finite sets and multi-object conjugate priors," *IEEE Trans. on Signal Processing*, vol. 61, no. 13, pp. 3460–3475, 2013.

[VVC07] B.-T. Vo, B.-N. Vo, and A. Cantoni, "Analytic implementations of the cardinalized probability hypothesis density filter," *IEEE Trans. on Signal Processing*, vol. 55, no. 7, pp. 3553–3567, 2007.

[VVC09] ——, "The cardinality balanced multi-target multi-Bernoulli filter and its implementations," *IEEE Trans. on Signal Processing*, vol. 57, no. 2, pp. 409–423, 2009.

[VVP14] B.-N. Vo, B.-T. Vo, and D. Phung, "Labeled random finite sets and the Bayes multi-target tracking filter," *IEEE Trans. on Signal Processing*, vol. 62, no. 24, pp. 6554–6567, 2014.

[VVPS10] B.-N. Vo, B.-T. Vo, N.-T. Pham, and D. Suter, "Joint detection and estimation of multiple objects from image observations," *IEEE Trans. on Signal Processing*, vol. 58, no. 10, pp. 5129–5241, 2010.

[Wil03] J. L. Williams, "Gaussian mixture reduction for tracking multiple maneuvering targets in clutter," Master's thesis, Air Force Institute of Technology, Wright-Patterson Air Force Base, OH, 2003. [Online]. Available: http://handle.dtic.mil/100.2/ADA415317

[WM06] J. L. Williams and P. S. Maybeck, "Cost function-based hypothesis control techniques for multiple hypothesis tracking," *Mathematical and Computer Modelling*, vol. 43, pp. 976–989, 2006.

[WvdM01] E. A. Wan and R. van der Merwe, "The unscented Kalman filter," in *Kalman filtering and neural networks*, S. Haykin, Ed. John Wiley & Sons, Hoboken, NJ, USA, 2001, pp. 221–280.

[XBL05] L. Xiao, S. Boyd, and S. Lall, "A scheme for robust distributed sensor fusion based on average consensus," in *Proc. 4^{th} Int. Symposium on Information Processing in Sensor Networks (IPSN)*, pp. 63-70, Los Angeles, CA, USA, 2005.