Linköping Studies in Science and Technology
Dissertation No. 1968

Gaussian Processes for Positioning Using Radio Signal Strength Measurements

Yuxin Zhao

LINKÖPING UNIVERSITY

Linköping studies in science and technology. Dissertations.
No. 1968

Gaussian Processes for Positioning Using Radio Signal Strength Measurements

Yuxin Zhao

LINKÖPING UNIVERSITY

Linköping studies in science and technology. Dissertations.
No. 1968

Gaussian Processes for Positioning Using Radio Signal Strength Measurements

Yuxin Zhao

yuxin.zhao@liu.se
www.control.isy.liu.se
Division of Automatic Control
Department of Electrical Engineering
Linköping University
SE–581 83 Linköping
Sweden

ISBN 978-91-7685-162-3 ISSN 0345-7524

Printed by LiU-Tryck, Linköping, Sweden 2019

To my family

Abstract

Estimation of unknown parameters is considered as one of the major research areas in statistical signal processing. In the most recent decades, approaches in estimation theory have become more and more attractive in practical applications. Examples of such applications may include, but are not limited to, positioning using various measurable radio signals in indoor environments, self-navigation for autonomous cars, image processing, radar tracking and so on. One issue that is usually encountered when solving an estimation problem is to identify a good system model, which may have great impacts on the estimation performance. In this thesis, we are interested in studying estimation problems particularly in inferring the unknown positions from noisy radio signal measurements. In addition, the modeling of the system is studied by investigating the relationship between positions and radio signal strength measurements.

One of the main contributions of this thesis is to propose a novel indoor positioning framework based on proximity measurements, which are obtained by quantizing the received signal strength measurements. Sequential Monte Carlo methods, to be more specific particle filter and smoother, are utilized for estimating unknown positions from proximity measurements. The Cramér-Rao bounds for proximity-based positioning are further derived as a benchmark for the positioning accuracy in this framework.

Secondly, to improve the estimation performance, Bayesian non-parametric modeling, namely Gaussian processes, have been adopted to provide more accurate and flexible models for both dynamic motions and radio signal strength measurements. Then, the Cramér-Rao bounds for Gaussian process based system models are derived and evaluated in an indoor positioning scenario.

In addition, we estimate the positions of stationary devices by comparing the individual signal strength measurements with a pre-constructed fingerprinting data base. The positioning accuracy is further compared to the case where a moving device is positioned using a time series of radio signal strength measurements.

Moreover, Gaussian processes have been applied to sports analytics, where trajectory modeling for athletes is studied. The proposed framework can be further utilized to carry out, for instance, performance prediction and analysis, health condition monitoring, etc. Finally, a grey-box modeling is proposed to analyze the forces, particularly in cross-country skiing races, by combining a deterministic kinetic model with Gaussian process.

Populärvetenskaplig sammanfattning

Skattning av okända parametrar anses vara ett av de viktigaste forskningsområdena inom statistisk signalbehandling. Under de senaste årtiondena har metoderna i skattningsteori blivit allt mer attraktiva i praktiska tillämpningar. Några exempel på tillämpningar är positionering baserat på olika mätbara radiosignaler i inomhusmiljöer, navigering för självkörande bilar, bildbehandling, radarspårning och så vidare. Ett delproblem som brukar uppstå är att identifiera en exakt systemmodell, vilket kan ha stor inverkan på skattningsnoggrannheten. I den här avhandlingen är vi intresserade av att studera skattning av okända positioner från brusiga radiosignalmätningar. En del rör modellering av förhållandet mellan position och radiosignalers utbredning.

Ett av de viktigaste bidragen från denna avhandling är att föreslå ett nytt positioneringsramverk baserat på rapporter av närhet till olika kända positioner. Sekventiella Monte Carlo-metoder, nämligen partikelfilter och partikelglättare används för att skatta okända positioner. Vi har studerat praktisk och teoretisk positioneringsnoggrannheten för inomhusmiljöer.

För att förbättra prestandan så har Bayesianska icke-parametriska metoder baserat på Gaussprocesser använts för att modellera både dynamiska rörelser och radioutbredning på ett bra och flexibelt sätt. Vidare har vi härlett teoretiska undre gränser för modeller baserat på Gaussprocesser.

Vi jämför även positionering av rörliga objekt där man har tillgång till tidsserier av mätningar, med positionering av stillastående objekt där vi bara har enskilda mätningar som jämförs med sparade ”fingeravtryck” av radiosignalmätningar.

Dessutom har vi använt Gaussprocesser för sportanalys. Det föreslagna ramverket kan utnyttjas för att exempelvis utföra prestandauppskattningar och analyser, hälsostatusövervakning, osv. Vidare föreslås metoder för modellering av en grupp av atleter, som skulle kunna användas för flödesmönsterigenkänning, strategiplanering, osv. Till sist föreslås en gråboxmodellering för att analysera ansträngning och glid inom längdskidåkning, genom att kombinera en deterministisk kinetisk modell med Gaussprocesser.

Acknowledgments

During these five years, I have experienced a really enjoyable research life, full of new things and adventures. It has been an exciting journey to join the Automatic Control group at Linköping University and Ericsson research Linköping for the first three years. I feel grateful and lucky to collaborate with so many talented and wonderful people.

First of all, I would like to express my great gratitude to my supervisors, Fredrik Gunnarsson, Fredrik Gustafsson and Carsten Fritsche, for your support, encouragement, and guidance during the past years. Thanks to both Fredriks, for your inspiring ideas, in-depth knowledge, and enthusiasm towards research, which have helped me a lot in accomplishing the research work. Thanks to Carsten, for every detailed discussion and great efforts you have put in guiding me through tough situations. I would also like to thank Gustaf Hendeby, for the timely research discussions whenever I got stuck.

Great thanks to Martin Enqvist, Svante Gunnarsson and Ninna Stensgård for providing a very nice and comfortable work environment at Linköping University. Also special thanks to Samuel Axelsson and Mehdi Amirijoo who have helped me a lot with the working conditions at Ericsson Research Linköping. Thanks to TRAX project, which is an European Union FP7 Marie Curie training program, and ELLIT, which is a strategic research project funded by the Swedish government, for funding my research during the last five years.

It is my great pleasure to know all the colleagues at Automatic Control group, Linköping University. Great thanks to all of you (too many names to mention) for the friendly environment here. I really enjoy the information sharing during fika and lunch time. Thanks also to my colleagues at Ericsson research, Linköping, for making such a warm and motivating environment.

Thanks to all my co-authors and special thanks to Fredrik Gunnarsson, Fredrik Gustafsson, Carsten Fristche, Parinaz Kasebzadeh, Kamiar Radnosrati and Angela Fontan who proofread this thesis and provided helpful comments. All the remaining errors are my own. I would like to thanks Gustaf again for providing the LaTeX template that has been used to write this thesis. Also, I would like to express my gratitude for research collaborations with University of Sheffield, especially Lyudmila S. Mihaylova and Chao Liu. Thanks for the support and discussions while I was doing my secondment at Sheffield, UK. Thanks to Feng Yin and Tianshi Chen, for arranging the research visit to Chinese University of Hong Kong, Shenzhen, China.

Thanks to all my friends, for the wonderful time we have during weekends and holidays. The time outside work would be less fantastic without you. Thanks to my family for always being there and supporting me. Special thanks to my parents for your continuous encouragement. Finally, a big thank you to my dear

husband, Kun Qian, for letting me know how wonderful life could be!

Linköping, Sweden, January, 2019
Yuxin Zhao

Contents

Notation **xv**

1 Introduction **1**
1.1 Application Examples . . . 3
1.2 Main Contributions . . . 5
1.3 Thesis outline . . . 6
1.3.1 Outline of Part I . . . 6
1.3.2 Outline of Part II . . . 6

I Background

2 Estimation **15**
2.1 Static Estimators . . . 15
2.1.1 Best Linear Unbiased Estimator . . . 16
2.1.2 Maximum Likelihood Estimator . . . 16
2.1.3 Least Squares . . . 17
2.2 Dynamic Estimators . . . 17
2.3 Sequential Monte Carlo Methods for State Inference . . . 19
2.3.1 Particle Filter . . . 21
2.3.2 Particle Smoother . . . 23

3 Modeling **25**
3.1 Linear Regression for Parametric Modeling . . . 25
3.2 Gaussian Process for Non-parametric Modeling . . . 27

4 The Cramér-Rao Bound **33**
4.1 Cramér-Rao Bound for Static Estimator . . . 34
4.2 Cramér-Rao Bound for Dynamic Estimator . . . 34

5 Concluding Remarks **37**
5.1 Summary of Contribution . . . 37
5.2 Some Insights into Future Work . . . 39

Bibliography **41**

II Publications

A Paper A **55**
1 Introduction . . . 57
1.1 Background . . . 57
1.2 Related Work and Our Contributions . . . 58
1.3 Paper Organization . . . 60
2 Prerequisites . . . 61
2.1 Deployment . . . 61
2.2 RSS Model . . . 62
2.3 Evaluation Set of Sample Positions . . . 66
2.4 Concluding Remarks . . . 66
3 Fundamental Lower Bounds on Position Estimation . . . 66
3.1 Preliminaries . . . 66
3.2 Proximity Report Based Position Estimator . . . 67
3.3 Fundamental Lower Bounds . . . 67
3.4 Computational Complexity . . . 70
3.5 Discussions on Bias . . . 70
4 RSS Threshold Optimization . . . 71
5 Experimental Validation . . . 74
5.1 Measurement Campaign . . . 75
5.2 RSS Threshold Optimization . . . 79
6 Conclusions . . . 84

B Paper B **89**
1 Introduction . . . 91
1.1 Background . . . 91
1.2 Related work and our contribution . . . 92
1.3 Main Motivation and Target Applications . . . 94
1.4 Paper organization . . . 95
2 Proximity Reports . . . 96
3 System Model . . . 97
4 Bayesian Filtering and Smoothing . . . 99
4.1 Particle Filtering . . . 100
4.2 Particle Smoothing . . . 102
5 Parametric Cramér-Rao Lower Bounds . . . 102
5.1 Parametric CRB for Filtering . . . 104
5.2 Parametric CRB for Smoothing . . . 105
6 Results . . . 106
6.1 Experimental Setup . . . 106
6.2 Performance Metrics . . . 107
6.3 Simulated Data . . . 109
6.4 Experimental Data . . . 114

7 Conclusions . . . 119

C Paper C **123**
1 Introduction . . . 125
2 Models . . . 127
2.1 Propagation model . . . 127
2.2 State-Space Model . . . 130
3 Particle Filtering Algorithm Based on GP . . . 130
4 Experimental Results . . . 132
4.1 Setup . . . 132
4.2 Propagation modeling . . . 133
4.3 Performance Evaluation . . . 134
5 Conclusions . . . 136

D Paper D **137**
1 Introduction . . . 139
1.1 Motivations and Background . . . 139
1.2 Related Work and Contributions . . . 140
1.3 Paper Organization . . . 141
2 Brief Introduction of Gaussian Process . . . 142
3 Gaussian Process Based System Models . . . 144
3.1 GP State Transition and GP Measurement Model . . . 145
3.2 GP Measurement Model . . . 150
3.3 GP State Transition Model . . . 150
3.4 Some Implications to State Estimation . . . 151
4 Parametric CRB . . . 153
4.1 Parametric CRB for Case 1 . . . 153
5 Posterior CRB . . . 155
5.1 Posterior CRB for Case 1 . . . 156
5.2 Posterior CRB for Case 2 . . . 157
5.3 Posterior CRB for Case 3 . . . 159
6 Application Example . . . 160
6.1 Parametric CRB for Indoor Positioning . . . 160
6.2 Posterior CRB for Indoor Positioning: Case 1 . . . 162
6.3 Posterior CRB for Indoor Positioning: Case 2 . . . 162
6.4 Posterior CRB for Indoor Positioning: Case 3 . . . 164
7 Results . . . 165
7.1 Data Collection and Simulation Setup . . . 165
7.2 Performance Metrics . . . 166
7.3 Simulation Results . . . 167
8 Conclusions . . . 169

E Paper E **175**
1 Introduction . . . 178
2 Standard Gaussian Process for Fingerprint Construction . . . 179
2.1 Characterizing Spatial Correlation . . . 181

2.2 Estimate Model Parameters . . . 181
2.3 Construct New Fingerprints . . . 181
3 Gaussian Process with Variogram . . . 182
3.1 Characterizing Spatial Correlation . . . 182
3.2 Estimate Parameters and Build New Fingerprints . . . 183
4 Location Fingerprinting Method . . . 184
5 Comparison Between SGP and GPV . . . 185
5.1 Stationarity Assumption . . . 185
5.2 Modeling and Regression Analysis . . . 185
5.3 Estimation . . . 186
5.4 Computational Complexity . . . 186
6 Field Campaign . . . 186
6.1 Data Collection . . . 186
6.2 RSS Map Reconstruction Results . . . 187
6.3 Localization Results . . . 187
7 Conclusions . . . 192
8 Acknowledgment . . . 193
9 Appendix . . . 193

F Paper F **195**
1 Introduction . . . 197
1.1 Background . . . 197
1.2 Contribution . . . 199
1.3 Paper Organization . . . 199
2 Flow Model for a Single Individual . . . 200
2.1 Flow Model . . . 200
2.2 Standard Gaussian Process Regression . . . 200
2.3 Grid Based On-line Gaussian Process Regression . . . 201
2.4 Kernel Selection . . . 203
2.5 Hyperparameters Determination . . . 204
3 Aggregated Flow Model for Multiple Individuals . . . 204
3.1 A Brief Overview of Sequence Clustering . . . 205
3.2 Flow Model for Clusters of Individuals . . . 205
4 Force Analysis for a Single Individual . . . 206
5 Data Description . . . 209
6 Results . . . 211
6.1 Individual Flow Model and Prediction . . . 211
6.2 Aggregated Flow Modeling . . . 213
6.3 Force Analysis . . . 216
7 Conclusions . . . 223
7.1 Hyperparameters for Individual Model . . . 223
7.2 Hyperparameters for Aggregated Model . . . 224
7.3 Hyperparameters for Force Analysis . . . 225

Notation

Operators and Symbols

Notation	Meaning
$[\cdot]^T$	Vector/matrix transpose
$[\cdot]^{-1}$	Inverse of a non-singular square matrix
$\text{tr}(\cdot)$	Trace of a square matrix
$\|\| \cdot \|\|$	Euclidean norm of a vector
$\|\cdot\|$	Cardinality of a set
$\mathbb{E}(\cdot)$	Statistical expectation
$X(\cdot)^T$	Short-hand notation for XX^T
$\ln(\cdot)$	Natural logarithm
$\log_{10}(\cdot)$	Logarithm to base 10
$\otimes$	Kronecker product
$\nabla_{\boldsymbol{\theta}} = \partial/\partial\boldsymbol{\theta}$	The gradient operator
$\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} = \nabla_{\boldsymbol{\theta}}\nabla_{\boldsymbol{\theta}}^T$	The Laplace operator
$\text{erf}(\cdot)$	The standard Gaussian error function.
I	Identity matrix
I_N	Identity matrix of size $N \times N$
$\mathbf{1}$	A vector of all 1s
$\mathbf{0}$	A vector of all 0s
$\boldsymbol{x}_{0:k}$	Stacked state vector from time 0 to time k, $[\boldsymbol{x}_0^T, \ldots, \boldsymbol{x}_k^T]^T$
$\boldsymbol{y}_{0:k}$	Stacked measurement vector from time 0 to time k, $[\boldsymbol{y}_0^T, \ldots, \boldsymbol{y}_k^T]^T$

Distributions

Abbreviation	Meaning
$\mathcal{N}(\mu, \sigma^2)$	Gaussian distribution with mean μ and variance σ^2
$Cat(\cdot)$	Categorical distribution

Abbreviations

Abbreviation	Meaning
BB	Barankin bound
BIM	Bayesian information matrix
BLUE	Best linear unbiased estimator
CDF	Cumulative distribution function
CI	Confidence interval
CMSE	Conditional mean squared error
CRB	Cramér-Rao bound
FFBSi	Forward filtering backward simulation
FIM	Fisher information matrix
GP	Gaussian process
GPR	Gaussian process regression
GPS	Global positioning system
IID	Independent and identically distributed
LS	Least squares
MAP	Maximum a posteriori
ML	Maximum likelihood
MLE	Maximum likelihood estimator
MSE	Mean squared error
PDF	Probability density function
PF	Particle filter
PS	Particle smoother
RCMSE	Root conditional mean squared error
RMSE	Root mean squared error
RSS	Received-signal-strength
SIR	Sequential importance resampling
SMC	Sequential Monte Carlo
SSM	State space model
TDOA	Time difference of arrival
UE	User equipment

1

Introduction

In statistics, two kinds of estimation problems are always formulated. The first problem relies on the estimation of a static system (static estimation as mentioned later in this thesis). In such a case, the target is to estimate a set of parameters that are of interest based on observations/measurements that are potentially corrupted by different kinds of noise. Approaches in estimation theory are usually applied to different applications, for instance, one typical use case is machine learning or system identification. In this thesis, estimation problems for static systems formulated under various applications will be addressed. Coming to the second estimation problem, where the system becomes dynamic, we are aiming at estimate a series of parameters that are changing as time evolves (dynamic estimation as mentioned later in this thesis). There are possibly correlations between parameters over time. To infer the latent parameters from the noisy measurements, a state-space signal model is usually formulated, which describes the time evolution of parameters and the way measurements are generated.

To solve the first problem, numerous methods have been studied in literature, for instance, the maximum likelihood estimator (MLE), least squares approach, best linear unbiased estimator (BLUE) and Bayesian estimation approaches such as maximum a posteriori (MAP) estimator and minimum mean square error (MMSE) estimator, which are introduced and thoroughly discussed in Kay (1993). Static estimation approaches are usually applied in the process of machine learning or system identification, where a set of inputs and outputs (also known as observations/measurements) are available, and the aim to determine the system that represents the relationship between inputs and outputs. In such a process, the structure of the system is either explicitly determined or there is no deterministic form of the system, which corresponds to a parametric and non-parametric approach, respectively. For parametric approach, the system model is usually

explicitly given by specifying a finite number of model parameters. For example, the linear regression model in Bishop (2006), autoregressive model and state-space model in Ljung (1999). However, in non-parametric approach, the number of parameters grows as the number of observations increases. In addition, there is no explicit model form such that the parameters in the non-parametric modeling are only determined by the data. One example of non-parametric modeling is Gaussian process (GP) Rasmussen and Williams (2006), which will be introduced in section 3.2.

For the second problem, where the dynamics of the system are usually formulated in a state-space form, estimating the unknown parameters (i.e., defined as states $\boldsymbol{x}_{1:k}$, where k is the time stamp) from the noisy observations (or measurements denoted by $\boldsymbol{y}_{1:\tau}$, where τ is also a time stamp) requires a set of statistical inferring steps, which are generally known as prediction, filtering or smoothing depending on if $\tau < k$, $\tau = k$ or $\tau > k$. A statistical framework is always applied to the dynamic estimation problems, one of which is the Bayesian inference framework Box and Tiao (1973), which recursively update the probability of states as more measurements are available. The most famous approach that fits into the Bayesian framework is the Kalman filter Kalman (1960), which gives a closed-form solution for linear Gaussian state-space models (SSMs). For nonlinear SSMs, there are approximation methods, such as extended Kalman filter (EKF), and unscented Kalman filter (UKF), see for instance Jazwinski (1970); Wan and Merwe (2000). However, for highly non-linear and non-Gaussian SSMs, the performance of extended Kalman filter and UKF are quite limited. In fact, this is a general limitation of Bayesian inference when we encounter the non-linear and non-Gaussian SSMs, where there are many intractable integrals.

In the 1980s, Monte Carlo methods emerged as a good tool for solving integration problems in Bayesian inference. It is known as one of the computational methods which use statistical simulations to approximate the estimates. Monte Carlo methods are mainly used in three distinct problem classes: optimization, numerical integration, and generating draws from a probability distribution Kroese et al. (2014). Sequential Monte Carlo methods emerged more recent in 1990s, which sequentially takes the observation and updates information about the hidden states Doucet et al. (2001). Here in this thesis, we will mainly use sequential Monte Carlo methods to solve the problem of inferring the latent states from observations. Detailed background and some examples of sequential Monte Carlo methods will be presented in following chapters.

In addition, the performance of the estimation problems mentioned above also heavily rely on the modeling of system, for instance, for static estimation, the model between the noisy observations and parameters we are aiming to infer. For dynamic estimation problems, both the state transition (i.e., how the dynamics of states evolve over time) and the relationships between states and measurements (i.e., measurement models) are of great importance. However, it is not always easy to find out the proper models, especially when there is no clear clue on how

the states evolves over time, how measurements are generated, and when the states and measurements are perturbed by noises.

Let's taking the received-signal-strength (RSS) as an example, which measures the power level of a radio signal from a reference node at a specific latent position. Due to the additive noise, shadowing, multi-path fading, and other effects, there is no explicit expression to impute the relationship between the latent positions and the observed RSS values. In such cases, empirical models have been proposed. For instance, the Hata-Okumura model Hata (1980) and the piecewise linear model Goldsmith (2006). While all those are parametric models, nonparametric models such as Gaussian processes can be used Ferris et al. (2006). Also, the model can be assumed to be known beforehand, otherwise it needs to be determined at a separate training (i.e., regression modeling) phase with the help of different estimation methods. The latter case is within the scope of this thesis.

The main contributions of this thesis include using different techniques to solve the two estimation problems described above in different applications. In the following section, some application examples will be introduced to further illustrate the problems that are studied in this thesis.

1.1 Application Examples

So far, a general description of the two problems we are aiming to solve has been given. Those estimation problems can be encountered in various research areas and applications, and signal processing is one of them. To be more specific, the applications of estimation theory are involved in for instance, image processing, radar technology, positioning systems, speech recognition, telecommunications and so on. In what follows, some illustrative examples will be provided to further assist the understanding and to better nail down the scope of this thesis.

For example, in a radar system Skolnik (1980), in order to determine the position of for instance, an aircraft, it is important to estimate the range between the radar and the aircraft. Usually, this is done by measuring the round-trip delay of a signal transmitted from the radar and reflected back by the aircraft. However, the measured round-trip delay is always perturbed by environmental random noise and delays in the receiver. Hence, a statistical estimation method is required to infer the range from the noisy round-trip delay measurements.

Another example of using estimation method is to estimate the parameters in a regression modeling problem. For instance, considering a simple problem of studying the relationship between the pressure and temperature of a given amount of gas. It is well known that there is a linear relationship between the pressure and temperature according to the Gay-Lussac's Law. However, in order to learn the explicit model, we measure both the pressures (denoted as $\boldsymbol{x}$) and temperatures

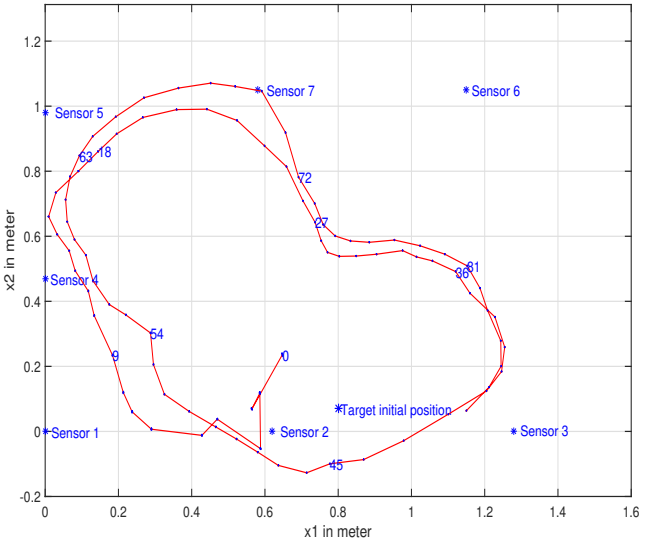


Figure 1.1: *Target tracking in a sensor network Zhao (2015).*

(denoted as $\boldsymbol{y}$), and assume a linear model between them $\boldsymbol{y} = \beta \boldsymbol{x} + \boldsymbol{n}$, where β is the coefficient that we want to find out and $\boldsymbol{n}$ is the noise that may occur in the measurements. Different estimation approaches can be used in this scenario to estimate both the coefficients β and the variance of Gaussian noise $\boldsymbol{n}$.

To infer the latent variables from the observations in a dynamic way, many applications in the target tracking area have been mentioned. One example application is to perform positioning of a moving device in a sensor network. For instance, we would like to track a device which carrying a speaker that can make quite distinctive sounds Zhao (2015). Then, a sensor network of microphones can be built up to receive the sound from the speaker. The observations in this case are the time difference of arrivals (TDOA), which can be computed by correlating the received signals at the microphones. However, the signal is always corrupted by random noises, where a statistical estimation framework can be applied. The latent variables are then the positions of the device carrying the speaker. A deployment of the sensor network with the target we want to track is illustrated in Figure 1.1.

Then, the positioning results are shown in Figure 1.1 by the solid curves. In this example, the aim is to infer the latent positions of the device from the TDOAs, which can be solved, for instance, using Bayesian filtering algorithms.

1.2 Main Contributions

So far we have seen some examples of the estimation problems. In short, the main contributions of this thesis can be summarized as: (I) Use estimation approaches to solve both static and dynamic estimation problems, which can be formulated more specifically as: inferring the unknown positions of a device based on radio signal strength measurements, which are vulnerable to different kinds of noises. (II) Perform both parametric and non-parametric modeling (i.e., Gaussian processes) for different regression applications (e.g., model the relationship between positions and signal strength measurements). The detailed contributions are as follows:

- A novel approach to generate proximity reports from radio received-signal-strength measurements has been proposed. Advanced modeling with Gaussian process has be applied to optimize the thresholds which trigger the proximity reports. Sets of optimized thresholds have been obtained using different criteria, which enables a new positioning fashion. (Paper A)

- Sequential Monte Carlo methods are proposed for the novel proximity based positioning system to estimate the position of a moving device in such a network. In addition, theoretical positioning performance for proximity based system has been derived. This leads to a simpler and more efficient way of positioning devices with less signaling and bandwidth, while maintaining adequate positioning accuracy. (Paper B)

- Modeling of the proximity report measurements, which is based on Gaussian process regression model, has been evaluated. The results well demonstrate that the better modeling of the measurements help to improve the position estimation performance. (Paper C)

- Theoretical lower bounds of generalized position estimation problems using received-signal-strength measurements have been derived, which are validated and compared to the simulated estimation performance. It has also been shown that the derived lower bounds can be used as a benchmark against the positioning accuracy obtained from any estimation algorithm under such scenarios. (Paper D)

- A static position estimation problem has been studied using received-signal-strength measurements modeled by two different Gaussian processes: the standard Gaussian process and Gaussian process with variogram. A statistical fingerprinting algorithm based on the pre-constructed RSS database has been designed for both models. The positioning performance are evaluated with RSS fingerprints constructed in two ways and further compared to the case where dynamic position estimation is performed. (Paper E)

- A novel application of Gaussian processes to analyze positioned data in sports is proposed. More precisely, the flows for both single trajectory and multiple trajectories haven been studied with various Gaussian processes.

A grey-boxing modeling is further proposed to analyze the forces in a typical cross-country skiing race by combing the known kinetic motion model with Gaussian process which accounts for the unknown forces. Evaluations are carried out using real data from a skiing race and the results provide valuable insights into sport performance analysis. (Paper F)

1.3 Thesis outline

The thesis consists of two parts. The first part introduces the general background such as static estimation approaches, sequential Monte Carlo methods for state inference, the parametric/non-parametric modeling and theoretical lower bounds. The second part presents the proposed solutions as a collection of peer-reviewed papers.

1.3.1 Outline of Part I

In the first part of this thesis, the most common estimation methods are introduced, which include but not limited to the best linear unbiased estimator (BLUE), maximum likelihood estimator (MLE), maximum a posterior (MAP) estimator, and Bayesian filtering and smoothing approaches with highlights on particle filtering and smoothing for dynamic state inference. Furthermore, approaches for system modeling are reviewed with two specific examples. Finally, the Cramér-Rao bound is introduced, which provides a theoretical lower limit for the estimator.

1.3.2 Outline of Part II

In the second part of this thesis, a collection of edited papers are presented. The summary is given as below.

Paper A

Paper A of this thesis is an edited version of,

> F. Yin, Y. Zhao, F. Guannarsson, and F. Gustafsson. Received-signal-strength threshold optimization using Gaussian processes. *IEEE Transactions on Signal Processing*, 65(8):2164–2177, 2017.

which is an extension of the two earlier contributions

> F. Yin, Y. Zhao, and F. Gunnarsson. Fundamental bounds on position estimation using proximity reports. In *IEEE 83rd Vehicular Technology Conference: VTC2016-Spring*, pages 1–5, Nanjing, China, 2016.

> F. Yin, Y. Zhao, and F. Gunnarsson. Proximity report triggering threshold optimization for network-based indoor positioning. In *18th International Conference on Information Fusion (FUSION)*, pages 1061–1069, Washington D.C., USA, July 2015.

Summary: This paper presents a generic received-signal-strength (RSS) threshold optimization framework for generating informative proximity reports. The proposed framework contains five main building blocks, namely the deployment information, RSS model, positioning metric selection, optimization process and management. Among others, we focus on Gaussian process regression (GPR) based RSS models and positioning metric computation. The optimal RSS threshold is found by minimizing the best achievable localization root mean squared error (RMSE) formulated with the aid of fundamental lower bound analysis. Computational complexity is compared for different RSS models and different fundamental lower bounds. The resulting optimal RSS threshold enables enhanced performance of new fashioned low-cost and low-complex proximity report based positioning algorithms. The proposed framework is validated with real measurements collected in an office area where bluetooth-low-energy (BLE) beacons are deployed.

The idea of this paper originated from Fredrik Gunnarsson and further refined by discussions among all authors. Both Feng Yin and the author of this thesis have contributions to the majority of writing and experimental results.

Paper B

Paper B of this thesis is an edited version of,

> Y. Zhao, C. Fritsche, F. Yin, F. Gunnarsson, and F. Gustafsson. Sequential Monte Carlo methods and theoretical bounds for proximity report based indoor positioning. *IEEE Transactions on Vehicular Technology*, 67(6):5372 – 5386, 2018a.

which is an extension of the earlier contribution

> Y. Zhao, F. Yin, F. Gunnarsson, M. Amirijoo, E. Özkan, and F. Gustafsson. Partile filtering for positioning based on proximity report. In *18th International Conference on Information Fusion (FUSION)*, pages 1046–1052, Washington D.C., USA, July 2015.

Summary: In paper A, the framework of optimizing thresholds for converting RSS to proximity reports has been developed. In this paper, we further consider positioning of devices based on a time series of proximity reports, which are generated using the optimized thresholds. This corresponds to nonlinear measurements with respect to the device position in relation to the network nodes. Therefore, sequential Monte Carlo methods, namely particle filtering and smoothing, are applicable for positioning. Positioning performance is evaluated in a typical office area with Bluetooth-low-energy beacons deployed for proximity detection and report, and is further compared to parametric Cramér-Rao lower bounds. Accuracy is concluded to vary spatially over the office floor, and in relation to the beacon deployment density.

The idea of this paper originated from Fredrik Gunnarsson and further refined by discussions among all authors. The majority of the paper is written by the author of this thesis. All the experimental results are by the author of this thesis.

Paper C

Paper C of this thesis is an edited version of,

> Y. Zhao, F. Yin, F. Gunnarsson, M. Amirijoo, and G. Hendeby. Gaussian processes for propagation modeling and proximity based indoor positioning. In *IEEE 83rd Vehicular Technology Conference: VTC2016-Spring*, pages 1–5, Nanjing, China, May 2016a.

Summary: In paper B, in order to perform the positioning of device using proximity reports, we assume a linear log-distance model for the RSS observations, which gives linear relationship between the RSS and the distance from the device to the network node in logarithmic scale. Then, the RSS is converted to proximity reports. However, in indoor scenario, the modeling of RSS is more complex. To address this problem, Gaussian process regression (GPR) has been applied for propagation modeling of RSS. This also provides some insights into the spatial correlation of the radio propagation in the considered area. Then, by assuming a dynamic transition of the states, particle filter is combined with GPR to estimate the position of the device. Radio propagation modeling and positioning performance are evaluated in a typical office area with BLE beacons deployed. Results show that the positioning accuracy can be improved by using GPR compared to the case work where linear log-distance model is used.

The idea of this paper originated from discussions among all authors. The majority of the paper is written by the author of this thesis. All the experimental results are by the author of this thesis.

Paper D

Paper D of this thesis is an edited version of,

> Y. Zhao, C. Fritsche, G. Hendeby, F. Yin, T. Chen, and F. Gunnarsson. Cramér-Rao bounds for filtering based on Gaussian process state-space models. *Submitted to IEEE Transactions on Signal Processing*, 2019.

which is an extension of the earlier contribution

> Y. Zhao, C. Fritsche, and F. Gunnarsson. Parametric lower bound for nonlinear filtering based on Gaussian process regression model. In *20th International Conference on Information Fusion (FUSION)*, pages 1–7, Xi'an, China, July 2017a.

Summary: The Cramér-Rao bounds (CRBs) for static estimation problems based on Gaussian processes are derived in Paper A. Hence, considering the dynamic

estimation problem as formulated in Paper C, we are aiming to derive the Cramér-Rao bounds for general state-space models based on Gaussian processes. To be more precise, the posterior CRBs are derived for three different state-space models: parametric state transition with GP measurement model, GP state transition with parametric measurement model, GP state transition with GP measurement model. In addition, the parametric CRB is derived for the first case. The theoretical bounds are evaluated and compared to simulation performance under a positioning scenario. We conclude that the derived CRBs are tight and can be used as benchmarks for the dynamic estimation problems formulated into Gaussian process based state-space models.

The idea of this paper originated from Carsten Fritsche and Feng Yin, which is further refined by discussions among all authors. The majority of the paper is written by the author of this thesis. All simulation results are by the author of this thesis.

Paper E

Paper D of this thesis is an edited version of,

> Y. Zhao, C. Liu, L. S. Mihaylova, and F. Guannarsson. Gaussian processes for RSS fingerprints construction in indoor localization. In *21st International Conference on Information Fusion (FUSION)*, pages 1377 – 1384, Cambridge, UK, July 2018b.

Summary: Paper B and Paper C focus on positioning of a moving device with certain assumed mobility patterns. In practice, there also exist use cases where static positioning is preferred (e.g., the device is not moving or moving without any continuous pattern). In such problems, fingerprinting is usually selected to infer the position of the device. However, the problem associated with fingerprinting method is the collection and maintenance of a relatively large fingerprint database/map. In this work, we propose and compare two algorithms namely, standard Gaussian process and Gaussian process with variogram, to reconstruct the RSS map with incomplete training data. To validate the effectiveness of both algorithms, experiments with BLE infrastructure have been conducted. RSS measurements are collected along predefined tracks. Both algorithms are applied to reconstruct the full RSS map within the whole area of interest. Finally, with the reconstructed complete RSS map, the localization performance using probabilistic fingerprinting is evaluated and compared.

The idea of this paper originated from discussions among all authors. Chao Liu writes sections on Gaussian process with variogram and contributes to corresponding experimental results. The author of this thesis writes sections on standard Gaussian process and contributes to corresponding experimental results.

Paper F

Paper E of this thesis is an extension of the earlier contribution

Y. Zhao, F. Yin, F. Gunnarsson, F. Hultkratz, and J. Fagerlind. Gaussian processes for flow modeling and prediction of positioned trajectories evaluated with sports data. In *19th International Conference on Information Fusion (FUSION)*, pages 1461–1468, July 2016b.

Summary: With the help of emerging positioning techniques, more data with position information are available for analysis, which may include but not limited to modeling and making predictions based on the data. Such problems have also been encountered in sports area, which are aiming for analyzing and monitoring athletes' performance and condition, providing guidance for coaches, and attracting interested audience. In this work, we apply Gaussian processes to flow modeling and force analysis in skiing races, but the proposed framework can be generally applied to other use cases with device trajectories of positioned data. Some specific aspects can be addressed when the data is periodic, like in sports where the event is split up over multiple laps along a specific track. Flow models of both the individual skier and a cluster of skiers are derived and analyzed. A grey-box modeling by combining the kinetic model with Gaussian process is also proposed to analyze the forces in skiing races. Performance has been evaluated using data from the Falun Nordic World Ski Championships 2015, in particular the Men's cross-country 4×10 km relay.

The idea of this paper originated from Fredrik Gunnarsson and further refined by discussions among all authors. The majority of the paper is written by the author of this thesis and all experimental results are by the author of this thesis.

Publications

Published work by the author of this thesis are listed below in chronological order. Publications indicated by a ⋆ are included in part II of this thesis.

⋆Y. Zhao, C. Liu, L. S. Mihaylova, and F. Guannarsson. Gaussian processes for RSS fingerprints construction in indoor localization. In *21st International Conference on Information Fusion (FUSION)*, pages 1377 – 1384, Cambridge, UK, July 2018b.

⋆Y. Zhao, C. Fritsche, F. Yin, F. Gunnarsson, and F. Gustafsson. Sequential Monte Carlo methods and theoretical bounds for proximity report based indoor positioning. *IEEE Transactions on Vehicular Technology*, 67(6):5372 – 5386, 2018a.

Y. Zhao, C. Fritsche, and F. Gunnarsson. Parametric lower bound for nonlinear filtering based on Gaussian process regression model. In *20th International Conference on Information Fusion (FUSION)*, pages 1–7, Xi'an, China, July 2017a.

⋆F. Yin, Y. Zhao, F. Guannarsson, and F. Gustafsson. Received-signal-strength threshold optimization using Gaussian processes. *IEEE Transactions on Signal Processing*, 65(8):2164–2177, 2017.

⋆Y. Zhao, F. Yin, F. Gunnarsson, F. Hultkratz, and J. Fagerlind. Gaussian processes for flow modeling and prediction of positioned trajectories evaluated with sports data. In *19th International Conference on Information Fusion (FUSION)*, pages 1461–1468, July 2016b.

⋆Y. Zhao, F. Yin, F. Gunnarsson, M. Amirijoo, and G. Hendeby. Gaussian processes for propagation modeling and proximity based indoor positioning. In *IEEE 83rd Vehicular Technology Conference: VTC2016-Spring*, pages 1–5, Nanjing, China, May 2016a.

F. Yin, Y. Zhao, and F. Gunnarsson. Fundamental bounds on position estimation using proximity reports. In *IEEE 83rd Vehicular Technology Conference: VTC2016-Spring*, pages 1–5, Nanjing, China, 2016.

F. Yin, Y. Zhao, and F. Gunnarsson. Proximity report triggering threshold optimization for network-based indoor positioning. In *18th International Conference on Information Fusion (FUSION)*, pages 1061–1069, Washington D.C., USA, July 2015.

Y. Zhao, F. Yin, F. Gunnarsson, M. Amirijoo, E. Özkan, and F. Gustafsson. Partile filtering for positioning based on proximity report. In *18th International Conference on Information Fusion (FUSION)*, pages 1046–1052, Washington D.C., USA, July 2015.

Part I

Background

2

Estimation

In this chapter, some basic estimation approaches will be introduced based on a probabilistic framework. The subsequent sections first introduce estimation methods for static problems as described in Chapter 1. Then, for the dynamic estimation problems, sequential Monte Carlo methods are discussed with focus on particle filter and smoother.

2.1 Static Estimators

The static estimation problems are usually formulated in a probabilistic framework. There are two philosophies considered in literature, namely, the classical and Bayesian philosophies Kay (1993). In a classical philosophy, the unknown parameters θ we are aiming to estimate are considered as deterministic, and the measurements $\boldsymbol{y}$, from which information about the unknown θ can be extracted, are random. Hence, we have the following statistical model (which relates the measurements with unknown parameters, also called likelihood function),

$$p(\boldsymbol{y}|\boldsymbol{\theta}). \tag{2.1}$$

Alternatively, the Bayesian philosophy considers the parameters θ to be also random, where usually a prior distribution $p(\theta)$ is assumed. Hence, the statistical models becomes $p(\boldsymbol{y}, \theta)$, which can be partitioned as $p(\boldsymbol{y}, \theta) = p(\boldsymbol{y}|\theta)p(\theta)$. In addition, the distribution $p(\boldsymbol{y}|\theta)$ is defined as the likelihood function, and $p(\theta)$ is the prior. In this thesis, we consider the static estimation problems formulated in a classical philosophy, and give some brief introduction of the most commonly used approaches to estimate the unknown but deterministic parameters θ.

2.1.1 Best Linear Unbiased Estimator

The best linear unbiased estimator (BLUE) is an estimator that is based a linear combination of the measurements. If we consider a general linear estimation problem formulated as

$$\boldsymbol{y} = H\theta + \boldsymbol{n}, \tag{2.2}$$

where we are aiming to estimate θ from measurements $\boldsymbol{y}$, and $\boldsymbol{n}$ is noise with zero mean and covariance C (it is not necessary to know the distribution of noise). Hence, the BLUE of $\boldsymbol{\theta}$ is given by

$$\hat{\boldsymbol{\theta}}_{\text{BLUE}} = (H^T C^{-1} H)^{-1} H^T C^{-1} \boldsymbol{y}, \tag{2.3}$$

with covariance

$$C_{\hat{\boldsymbol{\theta}}} = (H^T C^{-1} H)^{-1}. \tag{2.4}$$

Detailed derivations of BLUE can be found in Kay (1993). When the noise is Gaussian distributed, the BLUE is also the optimal estimator in the sense that it is the minimum variance unbiased (MVU) estimator.

2.1.2 Maximum Likelihood Estimator

The maximum likelihood estimator (MLE) is defined as the estimated value of $\boldsymbol{\theta}$ that maximizes the likelihood function $p(\boldsymbol{y}|\boldsymbol{\theta})$, which is usually equivalently formulated as

$$\hat{\boldsymbol{\theta}}_{\text{MLE}} = \arg\max_{\boldsymbol{\theta}} \ln p(\boldsymbol{y}|\boldsymbol{\theta}). \tag{2.5}$$

Assuming the likelihood function is differentiable, the MLE is usually found in the solutions to

$$\frac{\partial \ln p(\boldsymbol{y}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \boldsymbol{0}. \tag{2.6}$$

If there are more than one solution, the value that maximizes the likelihood function is the MLE. For a general linear model given as

$$\boldsymbol{y} = H\boldsymbol{\theta} + \boldsymbol{n}, \tag{2.7}$$

where $\boldsymbol{n}$ is zero mean Gaussian distributed with covariance C, the MLE of $\boldsymbol{\theta}$ is given by

$$\hat{\boldsymbol{\theta}}_{\text{MLE}} = (H^T C^{-1} H)^{-1} H^T C^{-1} \boldsymbol{y}, \tag{2.8}$$

with covariance

$$C_{\hat{\boldsymbol{\theta}}} = (H^T C^{-1} H)^{-1}. \tag{2.9}$$

It should be noted that the expression of MLE for the general linear model is identical to the BLUE. However, this is only the case when the noise is Gaussian distributed. In most of cases, the closed-form solution of MLE is hard to obtain, and numerical methods are always considered. For instance, the expectation-maximization (EM) algorithm as given in Dempster et al. (1977), and Newton-Raphson method as given in Wallis (1685).

2.1.3 Least Squares

In least squares (LS) approaches, we are aiming to find out the parameters $\boldsymbol{\theta}$ such that the sum of residuals (i.e., squared differences between measurements and the signal values generated by the system model which depends on $\boldsymbol{\theta}$) are minimized. Among all least squares problems, it is common to encounter the linear least squares (e.g., in Paper A and Paper B), which can be formulated as

$$\boldsymbol{y} = H\boldsymbol{\theta} + \boldsymbol{n}, \tag{2.10}$$

where the noise $\boldsymbol{n}$ is of zero mean with any arbitrary distribution. Then, the least squares estimator (LSE) becomes

$$\hat{\boldsymbol{\theta}}_{\text{LSE}} = (H^T H)^{-1} H^T \boldsymbol{y}. \tag{2.11}$$

It is noticed that compared with BLUE and MLE, it is not necessary to know the covariance of noise as long as the noise has zero mean. The minimum least squares error can be computed as $(\boldsymbol{y} - H\hat{\boldsymbol{\theta}})^T(\boldsymbol{y} - H\hat{\boldsymbol{\theta}})$.

2.2 Dynamic Estimators

In dynamic estimation problems, we are aiming to estimate parameters in a sequential manner. Hence, it is generally to formulate the dynamic estimation problem in a state-space form, which can be given as

$$\boldsymbol{x}_k = \boldsymbol{f}(\boldsymbol{x}_{k-1}, \boldsymbol{n}_k), \tag{2.12a}$$

$$\boldsymbol{y}_k = \boldsymbol{g}(\boldsymbol{x}_k, \boldsymbol{e}_k), \tag{2.12b}$$

where $\boldsymbol{y}_k$ are the noisy measurements at time k, and we are aiming to estimate the unknown state $\boldsymbol{x}_k$ at each time stamp. $\boldsymbol{f}$ and $\boldsymbol{g}$ are state transition and measurement functions, respectively. In addition, there are usually process noise $\boldsymbol{n}_k$ in the state transition and measurement noise $\boldsymbol{e}_k$. When both the state transition and measurement models are linear and both noises are additive Gaussian distributed, the closed-form solution of the estimator can be found by Kalman filter/smoother, Kalman (1960). However, for most of the dynamic estimation problems, it is hard to find a closed-form solution, where sequential Monte Carlo methods come into play.

Sequential Monte Carlo methods are based on the Bayesian inference framework, where Bayes' theorem is used to update the probability for a hypothesis as more information becomes available. Given a time series of noisy observations $\boldsymbol{y}_{1:T} \triangleq \{\boldsymbol{y}_k\}_{k=1}^T$, we would like to infer the hidden states variables $\boldsymbol{x}_{0:T} \triangleq \{\boldsymbol{x}_k\}_{k=0}^T$. If interpreted in the Bayesian way, we would like to compute the posterior joint distribution of all states given all the observations. A straightforward way to get the posterior distribution by applying Bayes' theorem is given by

$$p(\boldsymbol{x}_{0:T}|\boldsymbol{y}_{1:T}) = \frac{p(\boldsymbol{y}_{1:T}|\boldsymbol{x}_{0:T})p(\boldsymbol{x}_{0:T})}{p(\boldsymbol{y}_{1:T})}, \tag{2.13}$$

where $p(\boldsymbol{x}_{0:T})$ is the prior distribution, $p(\boldsymbol{y}_{1:T}|\boldsymbol{x}_{0:T})$ is the likelihood distribution, and $p(\boldsymbol{y}_{1:T}) = \int p(\boldsymbol{y}_{1:T}|\boldsymbol{x}_{0:T})p(\boldsymbol{x}_{0:T})d\boldsymbol{x}_{0:T}$ is usually a constant normalization factor.

Although it looks very straightforward to compute the posterior distribution, it is sometimes hard. Since when there is a new observation coming in, the posterior distribution for all the states needs to be recomputed. This is a typical problem in the case which is considered in this thesis, where we want to get the best estimate immediately after we have obtained a new observation. As the number of time steps increases, the computation will become intractable, since the computational complexity increases as the dimensionality of the posterior distribution gets larger. This is a very tough problem especially in real-time applications.

Then, if we relax the computation a bit, that is to say, instead of computing the full posterior distribution, we can compute the marginal distributions of the states. In order to get to this marginal distribution, we have to assume the following system models in a probabilistic way Jazwinski (1970).

$$\boldsymbol{x}_0 \sim p(\boldsymbol{x}_0), \tag{2.14a}$$

$$\boldsymbol{x}_k \sim p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}), \tag{2.14b}$$

$$\boldsymbol{y}_k \sim p(\boldsymbol{y}_k|\boldsymbol{x}_k). \tag{2.14c}$$

Here

- $p(\boldsymbol{x}_0)$ is the prior distribution of $\boldsymbol{x}_0$.
- $p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1})$ specifies the dynamic model (indicating how $\boldsymbol{x}_k$ evolves over time) of $\boldsymbol{x}_k$, which also has the assumption that $\boldsymbol{x}_k$ is a Markov process, such that the state at time k only depends on the state at time $k-1$.
- $p(\boldsymbol{y}_k|\boldsymbol{x}_k)$ is the likelihood distribution which is specified by the measurement model as given in (3.1).

The model structure given above is also called the state-space model Ljung (1999). With the model structure defined, we can formulate the following Bayesian filtering and smoothing problems by computing corresponding marginal distributions.

- *Bayesian filtering problem:* we are aiming at computing the marginal distributions of the current state $\boldsymbol{x}_k$ given the current and previous measurements $\boldsymbol{y}_{1:k}$.
$$p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k}), k = 1, \ldots, T. \tag{2.15}$$
- *Bayesian prediction problem:* we are aiming at computing the marginal distributions of the future state $\boldsymbol{x}_{k+m}$, m steps after the current time step k:
$$p(\boldsymbol{x}_{k+m}|\boldsymbol{y}_{1:k}), k = 1, \ldots, T, m = 1, 2, \ldots. \tag{2.16}$$

- *Bayesian smoothing problem:* we are aiming at computing the marginal distributions of the state $\boldsymbol{x}_k$ given a certain interval of measurements $\boldsymbol{y}_{1:T}$ with $k < T$:

$$p(\boldsymbol{x}_k|\boldsymbol{y}_{1:T}), k = 1, \ldots, T. \tag{2.17}$$

In this thesis, we are focusing on the filtering and smoothing problems.

In order to solve the filtering and smoothing problems, many algorithms have been developed for state-space models of various forms. Some example algorithms may include:

- *Kalman filter (KF) Kalman (1960):* it provides a closed form solution to the linear Gaussian filtering problem. This is under the assumption that both the dynamic and measurement model are linear and Gaussian distributed, such that the posterior distribution is exactly Gaussian and no numerical approximations are needed.

- *Rauch-Tung-Striebel* smoother Rauch et al. (1965): this is the corresponding closed form smoother for linear Gaussian state model.

- *Sequential Monte Carlo (SMC) filter and smoother:* namely particle filter and smoother. Since Bayesian optimal filtering and smoothing equations are generally computationally intractable, SMC is one of the numberical methods that represents the posterior distribution as a weighted set of Monte Carlo samples.

In what follows, the SMC methods will be introduced.

2.3 Sequential Monte Carlo Methods for State Inference

The main problem of Bayesian inference is to obtain an estimation of the state $\boldsymbol{x}_k$ at time k ($k \leq T$), which can be reduced to the computation of the expectations over the posterior distribution:

$$\mathbb{E}\{\boldsymbol{x}_k|\boldsymbol{y}_{1:T}\} = \int \boldsymbol{x}_k p(\boldsymbol{x}_k|\boldsymbol{y}_{1:T}) d\boldsymbol{x}. \tag{2.18}$$

However, for most of the time, the computation of such integral is intractable, where numerical approximations are needed. *Monte Carlo methods* are a class of methods for computing the integrals of the form (2.18). They are methods to approximate the closed form expression of statistical terms by drawing samples from the distribution and using sample average to estimate the expectation. For instance, a typical Monte Carlo approximation to (2.18) can be given as: draw

N independent random samples from the distribution $\boldsymbol{x}_k^{(i)} \sim p(\boldsymbol{x}_k|\boldsymbol{y}_{1:T})$, for $i = 1, \ldots, N$. Then the expectation is estimated as

$$\mathbb{E}\{\boldsymbol{x}_k|\boldsymbol{y}_{1:T}\} \approx \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{x}_k^{(i)}. \tag{2.19}$$

It seems to be promising to draw samples from the distribution $p(\boldsymbol{x}|\boldsymbol{y}_{1:T})$. However, it is usually not possible due to the complicated functional form of the distribution. In order to solve this problem, an approximated distribution is used, namely importance distribution $\pi(\boldsymbol{x}|\boldsymbol{y}_{1:T})$, from which we can easily draw samples. With importance distribution, the Monte Carlo approximation to the expectation can be reformulated as

$$\begin{aligned}
\mathbb{E}\{\boldsymbol{x}|\boldsymbol{y}_{1:T}\} &= \int p(\boldsymbol{x}|\boldsymbol{y}_{1:T})\boldsymbol{x}d\boldsymbol{x} \\
&= \frac{\int p(\boldsymbol{y}_{1:T}|\boldsymbol{x})\boldsymbol{x}p(\boldsymbol{x})d\boldsymbol{x}}{\int p(\boldsymbol{y}_{1:T}|\boldsymbol{x})p(\boldsymbol{x})d\boldsymbol{x}} \\
&= \frac{\int \left[\frac{p(\boldsymbol{y}_{1:T}|\boldsymbol{x})\boldsymbol{x}p(\boldsymbol{x})}{\pi(\boldsymbol{x}|\boldsymbol{y}_{1:T})}\right]\pi(\boldsymbol{x}|\boldsymbol{y}_{1:T})d\boldsymbol{x}}{\int \left[\frac{p(\boldsymbol{y}_{1:T}|\boldsymbol{x})p(\boldsymbol{x})}{\pi(\boldsymbol{x}|\boldsymbol{y}_{1:T})}\right]\pi(\boldsymbol{x}|\boldsymbol{y}_{1:T})d\boldsymbol{x}} \\
&\approx \frac{\frac{1}{N}\sum_{i=1}^{N}\frac{p(\boldsymbol{y}_{1:T}|\boldsymbol{x}^{(i)})p(\boldsymbol{x}^{(i)})}{\pi(\boldsymbol{x}^{(i)}|\boldsymbol{y}_{1:T})}\boldsymbol{x}^{(i)}}{\frac{1}{N}\sum_{j=1}^{N}\frac{p(\boldsymbol{y}_{1:T}|\boldsymbol{x}^{(j)})p(\boldsymbol{x}^{(j)})}{\pi(\boldsymbol{x}^{(j)}|\boldsymbol{y}_{1:T})}} \\
&= \sum_{i=1}^{N} w^{(i)}\boldsymbol{x}^{(i)}.
\end{aligned} \tag{2.20}$$

In sequential Monte Carlo method, observations are taken into consideration sequentially and the states are updated sequentially, where we have $\mathbb{E}\{\boldsymbol{x}_k|\boldsymbol{y}_{1:k}\} \approx \sum_{i=1}^{N} w_k^{(i)}\boldsymbol{x}_k^{(i)}$. Then, considering the posterior distribution of states $\boldsymbol{x}_{0:k}$ given the obsearvations $\boldsymbol{y}_{1:k}$, using the Markov properties of the state-space model, we have:

$$\begin{aligned}
p(\boldsymbol{x}_{0:k}|\boldsymbol{y}_{1:k}) &\propto p(\boldsymbol{y}_k|\boldsymbol{x}_{0:k}, \boldsymbol{y}_{1:k-1})p(\boldsymbol{x}_{0:k}|\boldsymbol{y}_{1:k-1}) \\
&= p(\boldsymbol{y}_k|\boldsymbol{x}_k)p(\boldsymbol{x}_k|\boldsymbol{x}_{0:k-1}, \boldsymbol{y}_{1:k-1})p(\boldsymbol{x}_{0:k-1}|\boldsymbol{y}_{1:k-1}) \\
&= p(\boldsymbol{y}_k|\boldsymbol{x}_k)p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1})p(\boldsymbol{x}_{0:k-1}|\boldsymbol{y}_{1:k-1}).
\end{aligned} \tag{2.21}$$

If we draw samples from the the importance distribution $\boldsymbol{x}_{0:k} \sim \pi(\boldsymbol{x}_{0:k}|\boldsymbol{y}_{1:k})$, then, the weights at time k, is given as

$$w_k^{(i)} \propto \frac{p(\boldsymbol{y}_k|\boldsymbol{x}_k^{(i)})p(\boldsymbol{x}_k^{(i)}|\boldsymbol{x}_{k-1}^{(i)})p(\boldsymbol{x}_{0:k-1}^{(i)}|\boldsymbol{y}_{1:k-1})}{\pi(\boldsymbol{x}_{0:k}|\boldsymbol{y}_{1:k})}. \tag{2.22}$$

If assume that the states up to time k depend only on measurements up to time k, the importance distribution can be obtained recursively, as

$$\pi(\boldsymbol{x}_{0:k}|\boldsymbol{y}_{1:k}) = \pi(\boldsymbol{x}_k|\boldsymbol{x}_{0:k-1}, \boldsymbol{y}_{1:k})\pi(\boldsymbol{x}_{0:k-1}|\boldsymbol{y}_{1:k-1}). \tag{2.23}$$

The weights can be updated as

$$w_k^{(i)} \propto \frac{p(\boldsymbol{y}_k|\boldsymbol{x}_k^{(i)})p(\boldsymbol{x}_k^{(i)}|\boldsymbol{x}_{k-1}^{(i)})}{\pi(\boldsymbol{x}_k^{(i)}|\boldsymbol{x}_{0:k-1}^{(i)}, \boldsymbol{y}_{1:k})} \frac{p(\boldsymbol{x}_{0:k-1}^{(i)}|\boldsymbol{y}_{1:k-1})}{\pi(\boldsymbol{x}_{0:k-1}^{(i)}|\boldsymbol{y}_{1:k-1})}. \tag{2.24}$$

Now assume that we have already drawn samples $\boldsymbol{x}_{k-1}^{(i)}$ from the importance distribution $\pi(\boldsymbol{x}_{0:k-1}|\boldsymbol{y}_{1:k-1})$ and computed the corresponding weights $w_{k-1}^{(i)}$, we can draw samples $\boldsymbol{x}_{0:k}^{(i)}$ from $\pi(\boldsymbol{x}_{0:k}|\boldsymbol{y}_{1:k})$ by drawing new state samples at step k as $\boldsymbol{x}_k^{(i)} \sim \pi(\boldsymbol{x}_k|\boldsymbol{x}_{0:k-1}^{(i)}, \boldsymbol{y_{1:k}})$. Then, the weights can be computed recursively as

$$w_k^{(i)} \propto \frac{p(\boldsymbol{y}_k|\boldsymbol{x}_k^{(i)})p(\boldsymbol{x}_k^{(i)}|\boldsymbol{x}_{k-1}^{(i)})}{\pi(\boldsymbol{x}_k^{(i)}|\boldsymbol{x}_{0:k-1}^{(i)}, \boldsymbol{y}_{1:k})} w_{k-1}^{(i)}, \tag{2.25}$$

where $w_{k-1}^{(i)} \propto \frac{p(\boldsymbol{x}_{0:k-1}^{(i)}|\boldsymbol{y}_{1:k-1})}{\pi(\boldsymbol{x}_{0:k-1}^{(i)}|\boldsymbol{y}_{1:k-1})}$. SMC methods are also known as particle filtering and smoothing, which will be introduced in the following subsections.

2.3.1 Particle Filter

With the definition of importance sampling, sequential Monte Carlo approximation, and the state-space model as specified in (2.14), we can have a general particle filtering (also known as sequential importance resampling (SIR) filter) procedure.

- *Initialization:* draw N samples $\boldsymbol{x}_0^{(i)}$ from the prior distribution $p(\boldsymbol{x}_0)$.

$$\boldsymbol{x}_0^{(i)} \sim p(\boldsymbol{x}_0), i = 1, \ldots, N. \tag{2.26}$$

  Assign equal weight to each sample with $w_0^{(i)} = \frac{1}{N}$, for $i = 1, \ldots, N$.

- For each time step $k = 1, \ldots, T$, do the following:

  1. *Time update:* draw samples $\boldsymbol{x}_k^{(i)}$ from the importance distribution

$$\boldsymbol{x}_k^{(i)} \sim \pi(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}^{(i)}, \boldsymbol{y}_{1:k}), i = 1, \ldots, N. \tag{2.27}$$

  2. *Weights update:* compute the new weight at time k as

$$w_k^{(i)} \propto \frac{p(\boldsymbol{y}_k|\boldsymbol{x}_k^{(i)})p(\boldsymbol{x}_k^{(i)}|\boldsymbol{x}_{k-1}^{(i)})}{\pi(\boldsymbol{x}_k^{(i)}|\boldsymbol{x}_{0:k-1}, \boldsymbol{y}_{1:k})} w_{k-1}^{(i)}. \tag{2.28}$$

     Normalize the weights to sum to 1

$$w_k^{(i)} = \frac{w_k^{(i)}}{\sum_{i=1}^N w_k^{(i)}}. \tag{2.29}$$

3. The estimate of state at k is given as

$$\hat{\boldsymbol{x}}_k \approx \sum_{i=1}^{N} w_k^{(i)} \boldsymbol{x}_k^{(i)}. \tag{2.30}$$

4. *Resampling:* if the number of effective particles are too low, i.e., $N_{eff} = \frac{1}{\sum_{i=1}^{N}\left(w_k^{(i)}\right)^2} < N_{th}$, perform resampling as
   (a) Interpret each weight $w_k^{(i)}$ as the probability of obtaining the sample index i in the set $\left\{\boldsymbol{x}_k^{(i)} : i = 1, \ldots, N\right\}$.
   (b) Draw N samples from that discrete distribution and replace the old sample set with this new one.
   (c) Set all weights equal to $\frac{1}{N}$.

Here the resampling step duplicates the particles with large weights and discards the particles with relatively small weights. This has to be done to focus the particle filters to the more relevant part of the state space, i,e,. to states that are with high probabilities. It is also noted that resampling is not performed every step, but only when the number of effective particles becomes too small. The effective number of particles can be estimated from the variance of the particle weights Liu and Chen (1995) as given in the resampling algorithm given above. There are many ways of doing resampling. The resampling method described here is one example which is known as *multinomial resampling*, which will be used later in this thesis. For a comprehensive comparison between different resampling methods, the reader can refer to Douc (2005).

The performance of the particle filter depends on the selection of importance distribution $\pi(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}^{(i)}, \boldsymbol{y}_{1:k})$. The importance distribution should be in such a functional form that we can easily draw samples from it and that it is possible to evaluate the probability densities of the sample points. The optimal importance distribution in terms of variance is given as Doucet et al. (2001)

$$\pi(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}^{(i)}, \boldsymbol{y}_{1:k}) = p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}, \boldsymbol{y}_k). \tag{2.31}$$

However, it is generally hard to sample from this distribution. Here in this thesis, the conditional prior of the state $p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1})$ is used as the importance distribution for simplicity but at the cost of losing useful information in $\boldsymbol{y}_k$. This will result in the *bootstrap filter* given in Gordon et al. (1993), where the corresponding weights update becomes

$$w_k^{(i)} \propto w_{k-1}^{(i)} p(\boldsymbol{y}_k|\boldsymbol{x}_k^{(i)}). \tag{2.32}$$

So far in this thesis, we have introduced a general particle filtering algorithm which uses the observations obtained before and at the current step for computing the best possible estimate of the current state (and possibly future states).

However, sometimes it is also of interest to estimate states for each time step conditional on all the measurements that we have obtained. This problem can be solved with Bayesian smoothing.

2.3.2 Particle Smoother

The purpose of Bayesian smoothing is to compute the marginal posterior distribution of the state $\boldsymbol{x}_k$ at time k, given observations up to time step T, where $T > k$:

$$p(\boldsymbol{x}_k|\boldsymbol{y}_{1:T}). \tag{2.33}$$

As can be see from (2.33), the Bayesian smoother uses also the future observations for computing the estimates. The Bayesian smoothing equations give a general procedure to compute the state estimates given future observations in a backward recursive way. The smoothed distribution of $p(\boldsymbol{x}_k|\boldsymbol{y}_{1:T})$ for any $k < T$ can be computed according to Kitagawa (1994),

$$p(\boldsymbol{x}_{k+1}|\boldsymbol{y}_{1:k}) = \int p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k)p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k})d\boldsymbol{x}_k, \tag{2.34a}$$

$$p(\boldsymbol{x}_k|\boldsymbol{y}_{1:T}) = p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k}) \int \left[\frac{p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k)p(\boldsymbol{x}_{k+1}|\boldsymbol{y}_{1:T})}{p(\boldsymbol{x}_{k+1}|\boldsymbol{y}_{1:k})} \right] d\boldsymbol{x}_{k+1}, \tag{2.34b}$$

where $p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k})$ is the filtering distribution at time k, and $p(\boldsymbol{x}_{k+1}|\boldsymbol{y}_{1:k})$ is the predicted distribution at time $k+1$.

Similarly to the particle filtering introduced before, particle smoothing algorithm is aiming to solve the smoothing problems in non-linear/non-Gaussian state-space models. SIR particle smoother is corresponding to the SIR filtering algorithm in the previous subsection. It is based on direct use of the SIR filtering results for smoothing. In filtering, we approximate the posterior distribution for all states. However, only the current state $\boldsymbol{x}_k^{(i)}$ at time k is kept, while the history states $\boldsymbol{x}_{0:k-1}^{(i)}$ have been discarded.

An approximation to the smoothing distribution can be obtained by keeping the full histories. To get the smoothing solution, we also need to resample the state histories, not only the current states, to prevent the resampling from breaking the state histories. The resulting algorithm is similar to the particle filtering algorithm, except that we store all the particle histories $\left\{\boldsymbol{x}_{0:k}^{(i)}, w_{0:k}^{(i)}\right\}, i = 1, \ldots, N$. The approximation to the smoothed posterior distribution at time step k, given the observations up to the time step $T > k$, can be given as

$$\hat{\boldsymbol{x}}_k|\boldsymbol{y}_{1:T} = \sum_{i=1}^{N} w_T^{(i)} \boldsymbol{x}_k^{(i)}, \tag{2.35}$$

where $\hat{\boldsymbol{x}}_k|\boldsymbol{y}_{1:T}$ denotes the estimated states at time k given measurements up to time T and$\boldsymbol{x}_k^{(i)}$ is the kth component in the particle histories $\boldsymbol{x}_{0:T}^{(i)}$. This SIR

smoothing algorithm is straightforward and simple to understand. However, the estimation performance will be significantly degraded when $T \gg k$ Kitagawa (1996).

Instead of simply storing the particle histories in the SIR smoother, the forward filtering backward-simulation (FFBSi) particle smoother in Godsill et al. (2004) is based on the reuse of filtering results, and propagate individual trajectories backwards from the last time step T to the start time. It can be summarized as

1. From particle filtering, store the set of weighted particles as $\left\{x_k^{(i)}, w_k^{(i)}, i = 1, \ldots, N, k = 1, \ldots, T\right\}$.
2. For time T, choose $\tilde{\boldsymbol{x}}_T = \boldsymbol{x}_T^{(i)}$ with probability $w_T^{(i)}$.
3. For time $k = T-1, \ldots, 0$, do the following:
   (a) Update the weights according to
   $$w_{k|k+1}^{(i)} \propto w_k^{(i)} p(\tilde{\boldsymbol{x}}_{k+1} | \boldsymbol{x}_k^{(i)}). \tag{2.36}$$
   (b) Choose backward state $\tilde{\boldsymbol{x}}_k = \boldsymbol{x}_k^{(i)}$ with probability $w_{k|k+1}^{(i)}$.

Repeat the above procedure for M times, then we get $\tilde{\boldsymbol{x}}_{0:T}^m$ for $m = 1, \ldots, M$. The smoothing estimation can be approximated as:

$$\hat{\boldsymbol{x}}_k | \boldsymbol{y}_{1:T} \approx \frac{1}{M} \sum_{m=1}^{M} \tilde{\boldsymbol{x}}_k^m, k = 0, \ldots, T. \tag{2.37}$$

This algorithm will be adopted to our specific applications in Paper B.

3

Modeling

In this chapter, approaches to model a system will be introduced. The procedure to build a good model for the system is known as machine learning or system identification. In this thesis, we mainly focus on determining a system model which can formulated as

$$\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{e}), \tag{3.1}$$

where $\boldsymbol{x}$ is the input state vector, $\boldsymbol{y}$ is the output observation, $\boldsymbol{e}$ is the noise and function $\boldsymbol{f}$ denotes the model. In a typical statistical modeling procedure, we are aiming to find an appropriate function $\boldsymbol{f}$ which can well represent the relationships between observations and state variables. Modeling problems are usually addressed in two different ways, namely, parametric and non-parametric. For parametric modeling, where the model is parametrized by a finite set of parameters $\boldsymbol{\theta}$, various model structures have been proposed in literature, see for instance, Ljung (1999), Bishop (2006) and Murphy (2012). In the remaining part of this chapter, one of the most commonly used model structure is described, namely, linear regression. For non-parametric modeling, the typical approach, namely, Gaussian process will be described.

3.1 Linear Regression for Parametric Modeling

Linear regression is a series of approaches where the relationships between input and output variables are represented in a linear way. To be more specific, in linear regression (for simplicity, we consider for a scalar observation first), given a data set in the form of $\{y_i, x_{i,1}, \ldots, x_{i,p}\}_{i=1}^{n}$, the model structure can be given in the following general form

$$y_i = w_0 + \sum_{j=1}^{p} w_j x_{i,j} + e_i, \tag{3.2}$$

where y_i is the observation at output, $x_{i,1}, \ldots, x_{i,p}$ are the input variables, w_0 and $w_1, \ldots, w_p$ are linear weights which need to be determined.

In such a linear regression model, the noise factor is typically assumed to be independent and Gaussian distributed with zero mean, i.e., $e_i \sim \mathcal{N}(e_i; 0, \sigma_e^2)$, which assumes that the noise variance is constant for all the observations. The parameters in this model need to be estimated can be defined as $\boldsymbol{\theta} = \{w_{0:p}, \sigma_e\}$. If we stack all the observations into one vector, then we have

$$\boldsymbol{y} = X\boldsymbol{w} + \boldsymbol{e}, \tag{3.3}$$

where $\boldsymbol{y} = [y_1, \ldots, y_n]^T$ ($(\cdot)^T$ denotes transpose operation), $\boldsymbol{w} = [w_0, \ldots, w_p]^T$ and

$$X = \begin{bmatrix} 1 & x_{1,1} & x_{1,2} & \cdots & x_{1,p} \\ 1 & x_{2,1} & x_{2,2} & \cdots & x_{2,p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n,1} & x_{n,2} & \cdots & x_{n,p} \end{bmatrix}. \tag{3.4}$$

To train a linear regression model, it is required to estimate $\boldsymbol{w}$ based on different criteria. If, for instance, the least squares method introduced in Chapter 2 is applied, the parameter can be presented as

$$\hat{\boldsymbol{w}}_{LS} \triangleq \arg\min_{\boldsymbol{w}} \|\boldsymbol{y} - X\boldsymbol{w}\|_2^2, \tag{3.5}$$

where $\|\cdot\|_2$ is the L_2-norm. The solution to (3.5) can be obtained in closed-form by taking the derivative of the squared error function w.r.t parameter $\boldsymbol{w}$, which gives Casella and Berger (2002)

$$\hat{\boldsymbol{w}}_{LS} = (X^T X)^{-1}(X^T \boldsymbol{y}), \tag{3.6}$$

and the noise variance is estimated by

$$\hat{\sigma}_e^2 = \frac{1}{n-p-1} \sum_{i=1}^{n} \left(y_i - X_i \hat{\boldsymbol{w}}_{LS}\right)^2, \tag{3.7}$$

where X_i denotes row i of the matrix X.

The linear regression model is one of the most simple model structures. This model will be used later in our Paper A and B. However, in some practical systems, the relationship between the observations and states are highly nonlinear and linear regression model is no longer proper. Hence, more complex models are required. However, the complexity of the parametric models grows at the expense of increasing the number of parameters to be estimated. In such cases, Bayesian non-parametric modeling has been introduced to give more flexibility as the number of observations increase Hjort et al. (2010).

In Bayesian non-parametrics, the degree of freedom increase as the number of observations increases. It is a Bayesian model whose parameter space has infinite dimension. In Bayesian statistics, the parameter is modeled as a random variable: the value of the parameter is unknown, and all forms of such uncertainty in Bayesian statistics are expressed as randomness. In another way of interpreting Bayesian non-parametrics is to consider them as infinite stochastic processes. Roughly speaking, a stochastic process is a generalization of a probability distribution to functions Rasmussen and Williams (2006).

The modeling problem formulated above can be summarized as: given a finite set of training data $\mathcal{D}$, we want to find function f that make predictions for all possible input values. Then, we need to make assumptions about the characteristics of the underlying function. In Bayesian non-parametrics, a prior probability is given to every possible function (infinite numbers), where higher probabilities are given to functions that are considered to be more likely. Gaussian process as one of the Bayesian non-parametrics, provides a solution to this problem.

In what follows, Gaussian process as a typical non-parametric method, will be discussed and further used in Paper C, D, E and F.

3.2 Gaussian Process for Non-parametric Modeling

According to Bayes' theorem, we have the following relationship between the posterior distribution $p(\boldsymbol{x}_i|y_i)$, likelihood function $p(y_i|\boldsymbol{x}_i)$ and prior distribution

$$p(\boldsymbol{x}_i|y_i) = \frac{p(y_i|\boldsymbol{x}_i)p(\boldsymbol{x}_i)}{p(y_i)} = \frac{p(y_i|\boldsymbol{x}_i)p(\boldsymbol{x}_i)}{\int p(y_i|\boldsymbol{x}_i)p(\boldsymbol{x}_i)d\boldsymbol{x}_i}, i = 1, \ldots, n, \tag{3.8}$$

where $p(y_i)$ is the marginal likelihood distribution. From (3.8) we can see that the posterior distribution combines all the information from the likelihood and the prior. Stacking all the observations as a vector $\boldsymbol{y}$ and unknown variables as a matrix X, we can update the beliefs in the unknown variables through the posterior distribution. However, we are usually primarily interested in the prediction of a new observation y^* at a new unknown variable $\boldsymbol{x}^*$

$$y^*|\boldsymbol{x}^*, X, \boldsymbol{y}. \tag{3.9}$$

Gaussian process is a generalization of the Gaussian probability distribution. In a Gaussian process, every point in some continuous input space is associated with a normally distributed random variable. Moreover, every finite collection of those random variables has a multivariate normal distribution. The distribution of a Gaussian process is the joint distribution of all those (infinitely many) random variables.

Generally used as a machine learning method, Gaussian process measures the similarity between different points to predict the observation value for new input

state. Starting with a scalar case for simplicity, let's denote the training data set as $\mathcal{D} = \{(\boldsymbol{x}_i, y_i)\}_{i=1}^{n}$. Then, the observation can be given by a Gaussian process which is written as

$$y_i = f(\boldsymbol{x}_i) + e_i. \tag{3.10}$$

The noise term is assumed to be Gaussian distributed with zero mean and variance σ_e^2 in this work. The function $f(\boldsymbol{x}_i)$ is a Gaussian process, which is characterized by its mean function and covariance function. The mean and covariance (kernal) function can be defined as

$$m(\boldsymbol{x}_i) = \mathbb{E}\{f(\boldsymbol{x}_i)\}, \tag{3.11a}$$

$$k\left(\boldsymbol{x}_i, \boldsymbol{x}_i'\right) = \mathbb{E}\left\{[f(\boldsymbol{x}_i) - m(\boldsymbol{x}_i)]\left[f\left(\boldsymbol{x}_i'\right) - m\left(\boldsymbol{x}_i'\right)\right]\right\}. \tag{3.11b}$$

The Gaussian process $f(\boldsymbol{x}_i)$ can be written as

$$f(\boldsymbol{x}_i) \sim \mathcal{GP}\left(m(\boldsymbol{x}_i), k\left(\boldsymbol{x}_i, \boldsymbol{x}_i'\right)\right). \tag{3.12}$$

There are many choices for the kernal function defined in (3.11b). One of the most commonly used one is called squared exponential (SE) kernal, which is formulated as

$$k\left(\boldsymbol{x}_i, \boldsymbol{x}_i'\right) = \sigma^2 \exp\left(-\frac{\|\boldsymbol{x}_i - \boldsymbol{x}_i'\|^2}{2l^2}\right), \tag{3.13}$$

where σ^2 is the function variance, which determines the variation of function values from their means. l is the length scale, which describes the smoothness of a function. Small value means that function values can change quickly, large values characterize functions that change slowly. In order to distinguish between the free parameters in parametric modeling, here we call these *hyperparameters*.

In order to show how to update the prior distribution conditioned on observations coming in, illustrations of this process are shown in Figure 3.1. In these three subfigures, the posterior mean and 95% confidence interval of the function have been plotted. As we increase the number of observations, the information we get is also richer, which leads to less uncertainty in the model (e.g., for single observation, the model ambiguity is very large. As more observations become available, the ambiguity reduces and the model becomes more distinct).

Usually, we are primarily interested in incorporating the knowledge that the training data set provides about the function. Then we make predictions. It is natural to have the joint distribution of the training observations $\boldsymbol{y}$, and the prediction y^* at new test input value $\boldsymbol{x}^*$. From previous statements about Gaussian process, we know that the joint distribution of a collection of Gaussian distributed variables are still following the Gaussian distribution. Hence, we have

$$\begin{bmatrix} \boldsymbol{y} \\ y^* \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \boldsymbol{m}(X) \\ m(\boldsymbol{x}^*) \end{bmatrix}, \begin{bmatrix} K(X,X) + \sigma_e^2 & k(X, \boldsymbol{x}^*) \\ k(\boldsymbol{x}^*, X) & K(\boldsymbol{x}^*, \boldsymbol{x}^*) \end{bmatrix}\right), \tag{3.14}$$

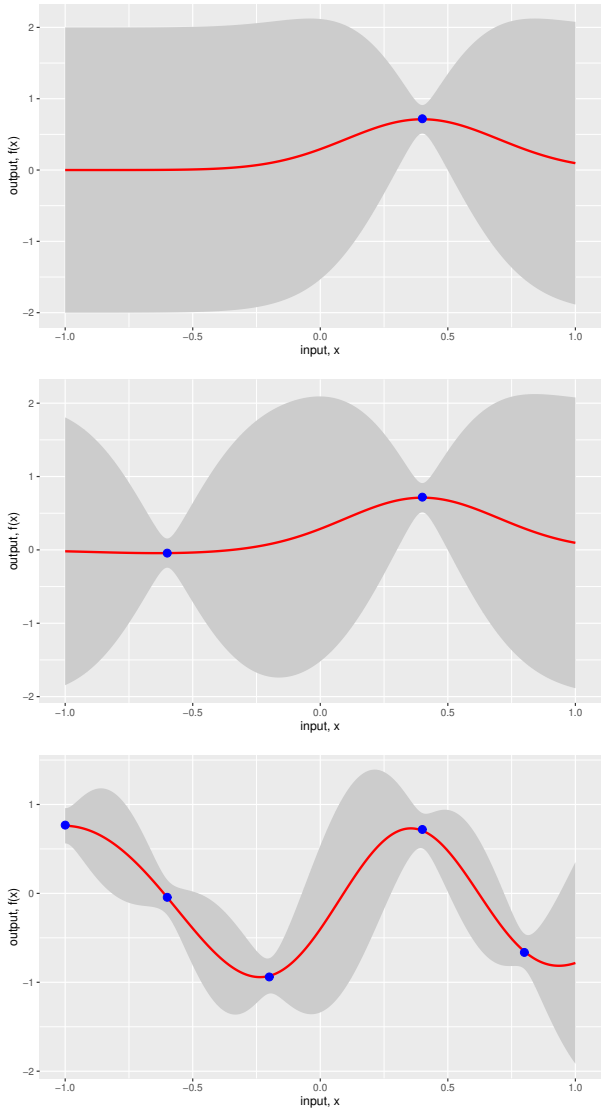


Figure 3.1: *Gaussian process regression with one, two and five observations: observed points are shown by blue dots, red line shows the posterior mean and the shaded areas show the 95% confidence interval.*

where $K(X, X)$ denotes the covariance matrix for the training input data X:

$$K(X, X) = \begin{bmatrix} k(\boldsymbol{x}_1, \boldsymbol{x}_1) & \cdots & k(\boldsymbol{x}_1, \boldsymbol{x}_n) \\ \vdots & \ddots & \vdots \\ k(\boldsymbol{x}_n, \boldsymbol{x}_1) & \cdots & k(\boldsymbol{x}_n, \boldsymbol{x}_n) \end{bmatrix}. \tag{3.15}$$

$k(X, \boldsymbol{x}^*)$ is a symmetric matrix of covariances evaluated at all pairs of training and test data given by $k(X, \boldsymbol{x}^*) = k(\boldsymbol{x}^*, X) = [k(\boldsymbol{x}_1, \boldsymbol{x}^*), k(\boldsymbol{x}_2, \boldsymbol{x}^*), \ldots, k(\boldsymbol{x}_n, \boldsymbol{x}^*)]$. Similarly, $k(\boldsymbol{x}^*, \boldsymbol{x}^*)$ is the covariance of $\boldsymbol{x}^*$. Then, the prediction distribution given all the training data and the test point $\boldsymbol{x}^*$ can be obtained by marginalizing the joint Gaussian distribution on the training data, which gives

$$y^* | \boldsymbol{x}^*, X, \boldsymbol{y} \sim \mathcal{N}\left(\bar{\mu}^*, \bar{k}^*\right), \tag{3.16}$$

where

$$\bar{\mu}^* = m(\boldsymbol{x}^*) + k(\boldsymbol{x}^*, X)[K(X, X) + \sigma_e^2 I]^{-1}(\boldsymbol{y} - \boldsymbol{m}(X)), \tag{3.17a}$$

$$\bar{k}^* = k(\boldsymbol{x}^*, \boldsymbol{x}^*) - k(\boldsymbol{x}^*, X)[K(X, X) + \sigma_e^2 I]^{-1} k(X, \boldsymbol{x}^*) + \sigma_e^2. \tag{3.17b}$$

Detailed derivations can be found in Rasmussen and Williams (2006, Apendix 2). If we take a closer look at the predictive distribution given by (3.17), then the mean of the prediction is a linear combination of observations $\boldsymbol{y}$. It is also noted that the variance in (3.17) does not depend on the observations, which only depends on the input variables. This is a property of the Gaussian distribution.

It is also useful to introduce the term *marginal likelihood*, which is $p(\boldsymbol{y}|X)$ at this point. This can be obtained directly by observing that $\boldsymbol{y} \sim \mathcal{N}(\boldsymbol{m}(X), K(X, X) + \sigma_e^2 I)$. Then the logarithm of the marginal likelihood is given by

$$\begin{aligned} \log p(\boldsymbol{y}|X) = & \\ & -\frac{1}{2}(\boldsymbol{y} - \boldsymbol{m}(X))^T (K(X, X) + \sigma_e^2 I)^{-1} (\boldsymbol{y} - \boldsymbol{m}(X)) \\ & -\frac{1}{2} \log |K(X, X) + \sigma_e^2 I| - \frac{n}{2} \log 2\pi. \end{aligned} \tag{3.18}$$

We see from the kernal definition that there are hyperparameters in the kernal function which are needed to be determined, namely σ and l, which determine the variance and the smoothness of the function, respectively. The effects of varying the two hyperparameters can be significant. This leads to a model selection problem where in this case is to determine the proper hyperparameters for a specific kernal function. Usually, the optimal hyperparameters can be obtained by maximizing the marginal likelihood function given in (3.18). It is noted that the determined hyperparameters for the Gaussian process is closely related to the training data set.

So far we have introduced some general aspects regarding Gaussian process. From what has been stated, we can see that the model is fully determined by the training data set rather than the structure of the model, which can be quite useful

in the cases when we can hardly tell the model structure from the training data. However, if we examine again the prediction equations in (3.17), there the main computational complexity relies on the inverting of the matrix $K(X, X) + \sigma_e^2 I$, which size is proportional to the size of the training data. The computational complexity for inverting n by n matrix is then $\mathcal{O}(n^3)$. In most modeling cases, the more training data, the better the model fit results. However, in this case of GP, the computational complexity has also been increased. The studies of reducing the complexity while still keeping the efficiency of GP have been arise in recent years. One way to avoid inverting the big covariance matrix is to use a grid based on-line Gaussian process, which is derived in Huber (2014). We also present a very practical use case with grid based on-line GP in our Paper E and A.

So far, we have discussed about the modeling aspects. As stated previously, the model can be either assumed to be known or can be trained using parametric/non-parametric modeling approaches. After the model has been selected and well trained, the estimation can be performed as introduced in Chapter 2.

4

The Cramér-Rao Bound

This chapter introduces the Cramér-Rao bound (CRB), which is often used to assess the best achievable performance for a specific estimation problem. The calculation of CRB is beneficial in many aspects. First of all, before we design any estimation algorithm, the CRB can be computed to provide insights into the performance of the targeting estimation problem. Further more, we can compare different estimators by looking into the closeness to the CRB. Then, we aim at improving the estimator if it is far away from the CRB.

Generally speaking, given an estimator of $\boldsymbol{\theta}$, $\hat{\boldsymbol{\theta}} = [\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_p]^T$ of dimension p from an observed vector $\boldsymbol{y}$, the CRB provides a lower bound on the covariance of any estimator under certain regularity conditions. The following relationship can be given for the mean squared error (MSE) matrix of any estimator

$$\mathcal{M}(\hat{\boldsymbol{\theta}}(\boldsymbol{y})) \triangleq \mathbb{E}\left\{\left[\hat{\boldsymbol{\theta}}(\boldsymbol{y}) - \boldsymbol{\theta}\right]\left[\hat{\boldsymbol{\theta}}(\boldsymbol{y}) - \boldsymbol{\theta}\right]^T\right\} \geq P^{CRB}, \tag{4.1}$$

where the operator $\geq$ means that the difference $\mathcal{M}(\hat{\boldsymbol{\theta}}(\boldsymbol{y})) - P^{CRB}$ is positive semidefinite.

As mentioned previously, the estimator can either be obtained in a static way (i.e., there is no dynamic change in the estimator) or a dynamic way (i.e., with filtering and smoothing, where the dynamic model of the estimator is also considered). Correspondingly, we have CRBs for both static estimator and dynamic estimator, which will be introduced in the following sections, respectively.

4.1 Cramér-Rao Bound for Static Estimator

For the static case, as mentioned in Chapter 2 the estimator $\hat{\boldsymbol{\theta}}$ and observation $\boldsymbol{y}$ is related by the probability density function $p(\boldsymbol{y};\boldsymbol{\theta})$. If $p(\boldsymbol{y};\boldsymbol{\theta})$ satisfies the condition that

$$\mathbb{E}_{p(\boldsymbol{y}|\boldsymbol{\theta})}\left[\frac{\partial \ln p(\boldsymbol{y}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right] = \mathbf{0}, \tag{4.2}$$

the covariance of any unbiased estimator satisfies the following according to Kay (1993)

$$\operatorname{cov}(\boldsymbol{\theta}) \geq \left\{-\mathbb{E}_{p(\boldsymbol{y}|\boldsymbol{\theta})}\left[\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \ln p(\boldsymbol{y}|\boldsymbol{\theta})\right]\right\}^{-1}, \tag{4.3}$$

where $J(\boldsymbol{\theta}) = \mathbb{E}_{p(\boldsymbol{y};\boldsymbol{\theta})}\left[-\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \ln p(\boldsymbol{y};\boldsymbol{\theta})\right]$ is called the Fisher information matrix (FIM). The static CRB will be used as a comparison for the static position estimator in Paper A.

4.2 Cramér-Rao Bound for Dynamic Estimator

With an estimation problem formulated in a dynamic way (i.e., given by the state-space model), additional information contained in the state dynamics should also be included when computing the CRB. Hence, for dynamic estimators, we will introduce both the posterior and parametric CRB for Bayesian filtering in what follows.

Usually for the state-space model given in Chapter2, we can generate random models for the state trajectories. The posterior CRB provide a lower bound that is averaged over all possible state trajectories. For a specific trajectory, we have the parametric CRB, which can be used as a lower bound conditioned on this specific state trajectory.

For a general dynamic estimation problem formulated in a state-space form as given in (2.12), the aim is to estimate the hidden states from time 0 to k. Hence, the Bayesian information matrix for posterior CRB is defined as

$$J_{0:k}^{\text{pos}} = \mathbb{E}_{p(\boldsymbol{y}_{0:k},\boldsymbol{x}_{0:k})}\left[-\Delta_{\boldsymbol{x}_{0:k}}^{\boldsymbol{x}_{0:k}} \ln p(\boldsymbol{y}_{0:k},\boldsymbol{x}_{0:k})\right]. \tag{4.4}$$

If additive Gaussian noises have been assumed in the state-space model, where

$$\boldsymbol{x}_k = f(\boldsymbol{x}_{k-1}) + \boldsymbol{n}_k, \tag{4.5a}$$

$$\boldsymbol{y}_k = g(\boldsymbol{x}_k) + \boldsymbol{e}_k, \tag{4.5b}$$

and both $\boldsymbol{n}_k$ and $\boldsymbol{e}_k$ are zero mean Gaussian distributed with covariance Q and R, respectively, the Bayesian information submatrix at time k (the inverse of the $p \times p$ lower-right partition of $\left(J_{0:k}^{\text{pos}}\right)^{-1}$) can be iteratively computed from the Bayesian information submatrix at $k-1$ Bergman (1999); Tichavsky et al. (1998), given by

$$J_k^{\text{pos}} = D_k^{22} - D_k^{21}(D_k^{11} + J_{k-1}^{\text{pos}})^{-1} D_k^{12}, \tag{4.6}$$

where

$$D_k^{11} = \mathbb{E}_{\boldsymbol{y}_{0:k},\boldsymbol{x}_{0:k}}\left\{-\Delta_{\boldsymbol{x}_{k-1}}^{\boldsymbol{x}_{k-1}} \log p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1})\right\}$$
$$= \mathbb{E}_{\boldsymbol{y}_{0:k},\boldsymbol{x}_{0:k}}\left\{\nabla_{\boldsymbol{x}_{k-1}} f^T(\boldsymbol{x}_{k-1}) Q^{-1}\left[\nabla_{\boldsymbol{x}_{k-1}} f^T(\boldsymbol{x}_{k-1})\right]^T\right\}, \tag{4.7a}$$
$$D_k^{12} = [D_k^{21}]^T = \mathbb{E}_{\boldsymbol{y}_{0:k},\boldsymbol{x}_{0:k}}\left\{-\Delta_{\boldsymbol{x}_{k-1}}^{\boldsymbol{x}_k} \log p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1})\right\}$$
$$= -\mathbb{E}_{\boldsymbol{y}_{0:k},\boldsymbol{x}_{0:k}}\left\{\nabla_{\boldsymbol{x}_{k-1}} f^T(\boldsymbol{x}_{k-1})\right\} Q^{-1}, \tag{4.7b}$$
$$D_k^{22} = \mathbb{E}_{\boldsymbol{y}_{0:k},\boldsymbol{x}_{0:k}}\left\{-\Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}_k} \log p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1})\right\}$$
$$+ \mathbb{E}_{\boldsymbol{y}_{0:k},\boldsymbol{x}_{0:k}}\left\{-\Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}_k} \log p(\boldsymbol{y}_k|\boldsymbol{x}_k)\right\}$$
$$= Q^{-1} + \mathbb{E}_{\boldsymbol{y}_{0:k},\boldsymbol{x}_{0:k}}\left\{\nabla_{\boldsymbol{x}_k} g^T(\boldsymbol{x}_k) R^{-1}\left[\nabla_{\boldsymbol{x}_k} g^T(\boldsymbol{x}_k)\right]^T\right\}. \tag{4.7c}$$

However, in certain scenarios, to compute the bound it is required to form Monte Carlo averages by simulating state trajectories from the state-space model (in the non-linear case as shown in Paper B). In such cases, the generated state trajectories may not be practical. Or sometimes it is impossible to generate state trajectories with certain constraints. Thus, we can resort to the parametric CRB, which is conditioned on a deterministic state trajectory $\left\{\boldsymbol{x}_{0:k}^*\right\}$, defined as

$$J_{0:k}^{\text{par}}(\boldsymbol{x}_{0:k}^*) = \mathbb{E}_{p(\boldsymbol{y}_{0:k}|\boldsymbol{x}_{0:k})}\left[-\Delta_{\boldsymbol{x}_{0:k}}^{\boldsymbol{x}_{0:k}} \ln p(\boldsymbol{y}_{0:k}|\boldsymbol{x}_{0:k})\right]|_{\boldsymbol{x}_{0:k}^*}. \tag{4.8}$$

Correspondingly, the parametric CRB for filtering is derived in Fritsche et al. (2016). If assume $\boldsymbol{n}_k$ and $\boldsymbol{e}_k$ are additive zero mean Gaussian distributed noises with covariance Q and R, respectively, the Fisher information submatrix at time k (computed as the inverse of the $p \times p$ lower-right partition of $\left(J_{0:k}^{\text{par}}\right)^{-1}(\boldsymbol{x}_{0:k}^*)$) can be updated iteratively from $k-1$:

$$J_k^{\text{par}}(\boldsymbol{x}_{0:k}^*) = \left(F_{k-1}(\boldsymbol{x}_{k-1}^*)\left[J_{k-1}^{\text{par}}(\boldsymbol{x}_{0:k-1}^*)\right]^{-1} F_{k-1}^T(\boldsymbol{x}_{k-1}^*)\right)^{-1} + [G_k(\boldsymbol{x}_k^*)]^T R^{-1} G_k(\boldsymbol{x}_k^*), \tag{4.9}$$

where

$$F_{k-1}(\boldsymbol{x}_{k-1}^*) = \nabla_{\boldsymbol{x}_{k-1}} f^T(\boldsymbol{x}_{k-1})|_{\boldsymbol{x}_{0:k-1}^*}, \tag{4.10a}$$
$$G_k(\boldsymbol{x}_k^*) = \nabla_{\boldsymbol{x}_k} g^T(\boldsymbol{x}_k)|_{\boldsymbol{x}_{0:k}^*}. \tag{4.10b}$$
$$\tag{4.10c}$$

It should be noted that for linear Gaussian state-space models, the derivation of CRBs is straightforward and simple. However, for cases investigated in this thesis, the system models maybe highly non-linear and the expectations are required to be approximated by Monte Carlo simulations. For instance, in Paper B where proximity is obtained by quantizing RSS measurements, the noise term in the measurement model is no longer additive and the function g is highly nonlinear. Hence, the computation of CRB becomes more complicated. In Paper D, the functions f and g may become random. As a consequence, the derivatives, for instance in (4.7), will become more complicated.

5

Concluding Remarks

In this thesis, we are aiming to solve estimation problems in different applications. To be more specific, estimation problems are studied with emphasis on inferring unknown parameters from either a static or dynamic system, and applying both parametric and non-parametric approaches to system modeling.

For static estimation problems, the most commonly used methods, such as best linear unbiased estimator, maximum likelihood estimator, and least squares estimator have been briefly introduced. For dynamic estimation problems, where a state-space model is formulated, the Bayesian inference framework is introduced. Sequential Monte Carlo methods, concretely particle filtering, are emphasized as numerical approximations to the integrals in Bayesian inference where there is no analytical solutions. Approaches for modeling different static and dynamic systems are further introduced, among which Gaussian process is considered as one powerful tool for non-parametric modeling. In addition, Cramér-Rao bounds are briefly introduced as the theoretical lower limits for the estimator.

5.1 Summary of Contribution

The main contribution of this thesis is to apply different estimation methods and modeling tools in a novel positioning framework. Gaussian process is combined with sequential Monte Carlo methods as well as static fingerprinting algorithms to improve position accuracy. Cramér-Rao bounds are derived as benchmarks for the estimation performance of the proposed positioning framework. In addition, we apply Gaussian processes for flow modeling and force analysis in sports, which may lead to new fashions in sport data analytics.

Paper A presents a generic received-signal-strength threshold optimization frame-

work for generating informative proximity reports. In particular, we focus on Gaussian process regression based RSS models and positioning metric computation. The optimal RSS threshold is found through minimizing the best achievable localization performance. The resulting optimal RSS threshold enables enhanced performance of new fashioned low-cost and low-complex proximity report based positioning algorithms.

Based on Paper A, in Paper B, we further consider positioning of devices based on a time series of proximity reports, which are generated using the optimized thresholds. Due to the nonlinearity of proximity measurements, sequential Monte Carlo methods, namely particle filtering and smoothing, are applied for position estimation. The positioning accuracy is concluded to be close to the parametric CRBs and 3 to 4 meter accuracy is achievable for most cases in this indoor scenario.

Paper C proposes an improved model for RSS measurements by applying Gaussian process. Then, the RSS measurements are further converted to proximity for positioning. Then, particle filter is combined with GP to infer the position of the device. Experimental results show that the positioning accuracy can be improved by using GP.

Since Gaussian process has been proved to improve the position estimation performance in Paper C, we derived the theoretical lower limits for a general estimation problem formulated into a state-space form where Gaussian process is used for modeling in Paper D. Parametric and posterior CRBs are derived for state-space models formulated by Gaussian processes. Evaluations are carried out under the positioning framework using RSS measurements. It is concluded that the derived CRBs based on Gaussian process state-space model are tight and can be used as benchmarks for evaluating the estimation performance under such scenarios.

Paper E considers a static position estimation, where a fingerprinting algorithm based on Gaussian process model is proposed. Both standard Gaussian process and Gaussian process with variogram are applied to reconstruct the full RSS map when the fingerprints are incomplete. With the proposed algorithms, the constructed RSS maps provide satisfactory positioning accuracy.

In Paper F, we apply Gaussian processes to flow modeling and force analysis of athletes in skiing races, in particular, the Men's cross-country 4×10 km relay in Falun Nordic World Ski Championships 2015. The results show that the flow models vary spatially for different skiers and clusters. By doing force analysis, we conclude that skiers apply different strategies over multiple laps, and skiers having better performance are good at maintaining propulsive force at inclining area, while the declining performance is mainly determined by the friction on ice and air resistance.

5.2 Some Insights into Future Work

Gaussian processes are proved to be a powerful tool in modeling and prediction in various applications, and we have so far derived theoretical bounds for estimation problems formulated by Gaussian process state-space models. However, the bounds that have been derived are under the assumption that all hyperparameters in Gaussian process and parameters in the parametric model are known. It would be natural to investigate into the cases where the parameters/hyperparameters in the model are unknown, and will be estimated together with the unknown states. The corresponding theoretical lower bounds are known as hybrid bounds.

In addition, we consider the simplified state-space model, where the Markov properties between states are maintained, while the correlation between measurements over time are ignored. However, this leads to a loss of information, which is contained in the off-diagonal elements of the GP posterior covariance matrix. Hence, it would be natural to extend the special case studied in this thesis to a more general state-space formulation based on Gaussian process, where the correlation between states and measurements over time are also considered.

Bibliography

European Union FP7 Tracking in Complex Sensor Systems (TRAX) project. https://www.trax.utwente.nl/, 2013-2017.

3GPP. Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Resource Control (RRC); Protocol specification. TS 36.331, 3rd Generation Partnership Project (3GPP), 2017.

3GPP. Radio Resource Control (RRC); Protocol specification. TS 25.331, 3rd Generation Partnership Project (3GPP), 2017.

J. S. Abel. A bound on mean-square-estimate error. *IEEE Transactions on Information Theory*, 39(5):1675–1680, Sep. 1993.

C. Andrieu, A. Doucet, and V. B. Tadic. On-line parameter estimation in general state-space models. In *IEEE Conference on Decision and Control*, Jan 2006.

L. N. Atallah, J.-P. Barbot, and P. Larzabal. From Chapman-Robbins bound towards Barankin bound in threshold behaviour prediction. *Electronics Letters*, 40(4):279–280, Feb. 2004.

P. Bahl and V. N. Padmanabhan. RADAR: An in-building RF-based user location and tracking system. In *IEEE INFOCOM. The 19th Annual Conference on Computer Communications*, 2000.

Y. Bar-Shalom, T. Kirubarajan, and X.-Rong Li. *Estimation with Applications to Tracking and Navigation*. John Wiley & Sons, Inc., New York, NY, USA, 2002. ISBN 0471221279.

R. Bartlett. *Introduction to Sports Biomechanics: Analysing Human Movement Patterns*. London: Routledge, 2007.

N. Bergman. *Recursive Bayesian Estimation Navigation and Tracking Applications*. PhD thesis, Linköping University, 1999.

C. M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006. ISBN 0387310738.

G. C. Bohling. Introduction to geostatistics and Variogram analysis. Lecture Notes, 2005.

G. E. P. Box and G. C. Tiao. *Bayesian Inference in Statistical Analysis*. John Wiley & Sons, New York, USA, 1973.

M. Briers, A. Doucet, and S. Maskell. Smoothing algorithms for state–space models. *Annals of the Institute of Statistical Mathematics*, 62(1):61–89, 2009. ISSN 1572-9052.

W. L. Brogan. *Modern Control Theory (3rd Ed.)*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1991. ISBN 0-13-589763-7.

L. Bruno, M. Khider, and P. Robertson. On-line training of the path-loss model in Bayesian WLAN indoor positioning. In *International Conference on Indoor Positioning and Indoor Navigation*, pages 1–9, Oct 2013.

M. Bshara, U. Orguner, F. Gustafsson, and L. Van Biesen. Fingerprinting localization in wireless networks based on received-signal-strength measurements: A case study on WiMAX networks. *IEEE Transactions on Vehicular Technology*, 59(1):283–294, Jan. 2010.

N. Bulusu, J. Heidemann, and D. Estrin. GPS-less low-cost outdoor localization for very small devices. *IEEE Personal Communications*, 7(5):28–34, Oct. 2000.

T. Camp, J. Boleng, and V. Davies. A survey of mobility models for ad hoc network research. *Wireless Communications & Mobile Computing (WCMC): Special Issue on Mobile Ad Hoc Networking: Research, Trends and Applications*, 2: 483–502, 2002.

G. Casella and R. L. Berger. *Statistical inference*. Duxbury advanced series. B. Cole, Pacific Grove (Calif.), 2002. ISBN 0-495-39187-5.

C. B. Chang and K. P. Dunn. *Applied State Estimation and Association*. MIT press, 2016.

N. A. C. Cressie. *Statistics for Spatial Data*. John Wiley, 1993.

M. P. Deisenroth and J. W. Ng. Distributed Gaussian processes. *arXiv preprint arXiv:1502.02843*, 2015.

A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal statistical Society, Series B*, 39(1):1–38, 1977.

R. Di Taranto, S. Muppirisetty, R. Raulefs, D. Slock, T. Svensson, and H. Wymeersch. Location-aware communications for 5G networks: How location information can improve scalability, latency, and robustness of 5G. *IEEE Signal Processing Magazine*, 31(6):102–112, Nov. 2014.

R. Douc. Comparison of resampling schemes for particle filtering. In *In 4th International Symposium on Image and Signal Processing and Analysis (ISPA*, pages 64–69, 2005.

A. Doucet and A. M. Johansen. A tutorial on particle filtering and smoothing: Fifteen years later. In *Nonlinear Filtering Handbook*, pages 656–704. Oxford University Press, 2011.

A. Doucet, N. de Freitas, and N. Gordon. *An Introduction to Sequential Monte Carlo Methods*, pages 3–14. Springer New York, New York, NY, 2001.

Z. Duan, V. P. Jilkov, and X. R. Li. State estimation with quantized measurements: Approximate MMSE approach. In *11th International Conference on Information Fusion*, pages 1–6, June 2008.

D. Duvenaud, J. R. Lloyd, R. Grosse, J. B. Tenenbaum, and Z. Ghahramani. Structure discovery in nonparametric regression through compositional kernel search. *arXiv:1302.4922v4*, 2013.

D. Ellis, E. Sommerlade, and I. Reid. Modelling pedestrian trajectory patterns with Gaussian processes. In *Proc. of IEEE 12th International Conference on Computer Vision Workshops (ICCV Workshops)*, pages 1229–1234, Sep. 2009. doi: 10.1109/ICCVW.2009.5457470.

R. Faragher and R Harle. An analysis of the accuracy of Bluetooth low energy for indoor positioning applications. In *27th International Technical Meeting of The Satellite Division of the Institute of Navigation*, pages 201–210, Tampa, Florida, USA, September 2014.

B. Ferris, D. Hähnel, and D. Fox. Gaussian processes for signal strength-based location estimation. In *Robotics: Science and Systems (RSS)*, pages 1–8, Philadelphia, USA, August 2006.

S. Foong and Z. Sun. High accuracy passive magnetic field-based localization for feedback control using principal component analysis. *Sensors (Basel, Switzerland)*, 16 (8)(1280), August 2016.

R. Frigola. *Bayesian Time Series Learning with Gaussian Processes*. PhD thesis, University of Cambridge, 2015.

R. Frigola, F. Lindsten, T. Schön, and C. E. Rasmussen. Identification of Gaussian process state-space models with particle stochastic approximation EM. In *IFAC Proceedings Volumes (IFAC)*, volume 19, pages 4097–4102, 08 2014.

C. Fritsche and U. Orguner. Supplementary material for On parametric lower bounds for discrete-time filtering. Tech. Rep. LiTH-ISY-R, 3090, Linköping University, Department of Electrical Engineering, Linköping, Sweden, 2016.

C. Fritsche and U. Orguner. On parametric smoothing Cramér-Rao bounds. Tech. Rep. LiTH-ISY-R, 3097, Linköping University, Department of Electrical Engineering, Linköping, Sweden, 2017.

C. Fritsche, U. Orguner, and F. Gustafsson. On parametric lower bounds for discrete-time filtering. In *2015 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 4338–4342, 2016.

G. Giorgetti, S. K. S. Gupta, and G. Manes. Optimal RSS threshold selection in connectivity-based localization schemes. In *International Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, pages 220–228, New York, NY, USA, 2008. ACM.

S. J. Godsill, A. Doucet, and M. West. Monte Carlo smoothing for nonlinear time series. *Journal of the American Statistical Association*, 99(465):156–168, 2004.

A. Goldsmith. *Wireless Communications*. Cambridge University Press, New York, NY, USA, 2006.

N. J. Gordon, D. J. Salmond, and A. F. M. Smith. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proceedings F - Radar and Signal Processing*, 140(2):107–113, April 1993.

M. Gudmundson. Correlation model for shadow fading in mobile radio systems. *Electronics Letters*, 27(23):2145–2146, Nov. 1991. ISSN 0013-5194. doi: 10.1049/el:19911328.

F. Gustafsson. *Statistical Sensor Fusion*. Studentlitteratur, Lund, Sweden, 2nd edition, 2012.

F. Gustafsson and R. Karlsson. Generating dithering noise for maximum likelihood estimation from quantized data. *Automatica*, 49(2):554–560, Feb. 2013.

F. Gustafsson, F. Gunnarsson, Niclas Bergman, U. Forssell, J. Jansson, R. Karlsson, and P.-J. Nordlund. Particle filters for positioning, navigation, and tracking. *IEEE Transactions on Signal Processing*, 50(2):425–437, Feb. 2002.

I. Guvenc and C. C. Chong. A survey on TOA based wireless localization and NLOS mitigation techniques. *IEEE Communications Surveys Tutorials*, 11(3):107–124, 2009.

M. Hata. Empirical formula for propagation loss in land mobile radio services. *IEEE Transactions on Vehicular Technology*, 29(3):317–325, Aug 1980.

T. He, C. Huang, B. M. Blum, J. A. Stankovic, and T. Abdelzaher. Range-free localization schemes for large scale sensor networks. In *Annual International Conference on Mobile Computing and Networking*, pages 81–95, New York, NY, USA, 2003. ACM.

R. J. Herring. *Real-Time Traffic Modeling and Estimation with Streaming Probe Data using Machine Learning*. PhD thesis, UC Berkeley: Industrial Engineering & Operations Research, 2010.

N. Hjort, C. Holmes, P. Mueller, and S. Walker. *Bayesian Nonparametrics: Principles and Practice*. Cambridge University Press, Cambridge, UK, 2010.

J. D. Hol, T. B. Schön, and F. Gustafsson. On resampling algorithms for particle filters. In *Proc. IEEE Nonlinear Statistical Signal Processing Workshop*, pages 79–82, Sept. 2006.

M. F. Huber. Recursive Gaussian process regression. In *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 3362–3366, May 2013.

M. F. Huber. Recursive Gaussian process: On-line regression and learning. *Pattern Recognition Letters*, 45:85–91, 2014.

S. S. Jan, S. J. Yeh, and Y. Liu. Received signal strength database interpolation by kriging for a Wi-Fi indoor positioning system. *Sensors(Basel, Switzerland)*, 15 (9):21377–21393, 2015.

A. H. Jazwinski. *Stochastic processes and filtering theory*. Mathematics in science and engineering. Academic press, New York, 1970. UKM.

D. Jin, F. Yin, C. Fritsche, A. M. Zoubir, and F. Gustafsson. Dithering in quantized RSS based localization. In *International Conference on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, pages 245–248, December 2015.

J. Joseph, F. Doshi-Velez, S. Huang, and N. Roy. A Bayesian nonparametric approach to modeling motion patterns. *Auton. Robots*, 31(4):383–400, Nov. 2011. ISSN 0929-5593.

R. E. Kalman. A new approach to linear filtering and prediction problems. *ASME Journal of Basic Engineering*, 1960.

R. Karlsson and F. Gustafsson. Filtering and estimation for quantized sensor information. In *Proc. 13th European Signal Processing Conference (EUSIPCO)*, pages 1–4, Antalya, Turkey, Sept. 2005.

S. M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall, Inc., Englewood Cliffs, NJ, 1993. ISBN 0-13-345711-7.

R. Khan, F. Sottile, and M. Spirito. Hybrid positioning through extended Kalman filter with inertial data fusion. *International Journal of Information and Electronics Engineering*, 3(1):127–131, Jan 2013.

H. Kim, W. Seo, and K. Baek. Indoor positioning system using magnetic field map navigation and an encoder system. *Sensors (Basel, Switzerland)*, 17 (3) (651), 2017.

G. Kitagawa. The two-filter formula for smoothing and an implementation of the Gaussian-sum smoother. *Annals of the Institute of Statistical Mathematics*, 46 (4):605–623, 1994.

G. Kitagawa. Monte Carlo filter and smoother for non-Gaussian nonlinear state space models. *Journal of Computational and Graphical Statistics*, 5(1):1–25, 1996.

E. D. Klenske, M. N. Zeilinger, B. Scholkopf, and P. Hennig. Gaussian process-based predictive control for periodic error correction. *IEEE Transactions on Control Systems Technology*, 24(1):110–121, Jan 2016. ISSN 1063-6536. doi: 10.1109/TCST.2015.2420629.

J. Ko and D. Fox. GP-BayesFilters: Bayesian filtering using Gaussian process prediction and observation models. In *International Conference on Intelligent Robots and Systems (IROS)*, pages 3471–3476, Nice, France, Sep. 2008.

A. Konak. Estimating path loss in wireless local area networks using ordinary Kriging. In *Winter Simulation Conference*, pages 2888–2896, Baltimore, MD, 2010.

D. G. Krige. *A Statistical Approach to Some Basic Mine Valuation Problems on the Witwatersrand*. PhD thesis, the University of Witwatersrand, Johannesburg, South Africa, 1951.

D. P. Kroese, T. Brereton, T. Taimre, and Z. I. Botev. Why the Monte Carlo method is so important today. *Wiley Interdisciplinary Reviews: Computational Statistics*, 6(6):386–392, 2014.

T. Lang, C. Plagemann, and W. Burgard. Adaptive non-stationary kernel regression for terrain modelling. In *In Proc. of the Robotics: Science and Systems Conference (RSS)*, 2007.

J. Lategahn, M. Muller, and C. Rohrig. TDoA and RSS based extended Kalman filter for indoor person localization. In *IEEE 78th Vehicular Technology Conference (VTC Fall)*, pages 1–5, Sept 2013.

B. Li, Y. Wang, H. K. Lee, A. Dempster, and C. Rizos. Method for yielding a database of location fingerprints in WLAN. *IEE Proceedings- Communication*, 152(5):580–586, 2005.

Y. Li, X. Chen, M. Coates, and B. Yang. Sequential Monte Carlo radio-frequency tomographic tracking. In *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 3976–3979, May 2011.

J. Lin, M. Vlachos, E. Keogh, and D. Gunopulos. *Advances in Database Technology: 9th International Conference on Extending Database Technology*, chapter Iterative Incremental Clustering of Time Series, pages 106–122. Springer Berlin Heidelberg, Berlin, Heidelberg, 2004. ISBN 978-3-540-24741-8.

F. Lindsten, P. Bunch, S. Särkkä, T. B. Schön, and S. J. Godsill. Rao-Blackwellized particle smoothers for conditionally linear Gaussian models. *IEEE Journal of Selected Topics in Signal Processing*, 10(2):353–365, March 2016.

C. Liu, A. Kiring, N. Salman, and L. Mihaylova. A Kriging algorithm for location fingerprinting based on received signal strength. In *Sensor Data Fusion: Trends, Solutions, Applications (SDF)*, Bonn, Germany, 2015.

J. S. Liu and R. Chen. Blind deconvolution via sequential imputations. *Journal of the American Statistical Association*, 90(430):567–576, 1995.

L. Ljung. *System Identification (2Nd Ed.): Theory for the User*. Prentice Hall PTR, Upper Saddle River, NJ, USA, 1999. ISBN 0-13-656695-2.

E. S. Lohan, J. Talvitie, P. Figueiredo e Silva, H. Nurminen, S. Ali-Löytty, and R. Piché. Received signal strength models for WLAN and BLE-based indoor positioning in multi-floor buildings. In *Proc. International Conference on Location and GNSS (ICL-GNSS)*, pages 1–6, June 2015.

D. J. C. MacKay. Gaussian processes – a replacement for supervised neural networks?, 1997.

L. Mainetti, L. Patrono, and I. Sergi. A survey on indoor positioning systems. In *22nd International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, pages 111–120, Sept 2014.

A. Manirabona and L. C. Fourati. A Kriged fingerprinting for wireless body area network indoor localization. *Wireless Personal Communications*, 80(4):1501–1515, 2015.

T. L. Marzetta. Computing the Barankin bound, by solving an unconstrained quadratic optimization problem. In *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 5, pages 3829–3832, Apr. 1997.

N. A. Maung Maung and M. Kawai. Performance improvement of connectivity-based localization using iterative learning. In *International Conference on Localization and GNSS (ICL-GNSS)*, pages 1–6, Helsinki, Finland, June 2014.

A. Mousa, Y. Dama, M. Najjar, and B. Alsayeh. Optimizing outdoor propagation model based on measurements for multiple RF cell. *International Journal of Computer Applications*, 60(5):5–10, December 2012.

L. S. Muppirisetty, T. Svensson, and H. Wymeersch. Spatial wireless channel prediction under location uncertainty. *IEEE Transactions on Wireless Communications*, 15(2):1031–1044, Feb 2016.

K. P. Murphy. *Machine Learning: A Probabilistic Perspective*. The MIT Press, 2012. ISBN 0262018020, 9780262018029.

M. Nilsson, J. Rantakokko, M. A. Skoglund, and G. Hendeby. Indoor positioning using multi-frequency RSS with foot-mounted INS. In *International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pages 177–186, Oct 2014.

N. Patwari and A. O. Hero, III. Using proximity and quantized RSS for sensor localization in wireless networks. In *ACM International Conference on Wireless Sensor Networks and Applications*, pages 20–29, New York, NY, USA, 2003. ACM.

J. Pelant, Z. Tlamsa, V. Benes, L. Polak, O. Kaller, L. Bolecek, J. Kufa, J. Sebesta, and T. Kratochvil. BLE device indoor localization based on RSS fingerprinting mapped by propagation modes. In *27th International Conference Radioelektronika*, pages 1–5, April 2017.

M. Pohjola. Analysing effectiveness of force application in ski skating using force and motion capture data: a model to support cross-country skiing research and coaching. Master's thesis, University of Jyväskylä, 2014.

J. Quiñonero Candela and C. E. Rasmussen. A unifying view of sparse approximate Gaussian process regression. *Journal of Machine Learning Research*, 6: 1939–1959, Dec. 2005. ISSN 1532-4435.

C. E. Rasmussen and H. Nickisch. Gaussian processes for machine learning (GPML) toolbox. *Journal of Machine Learning Research*, 11:3011–3015, Dec. 2010. ISSN 1532-4435.

C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. MIT Press, Cambridge, MA, USA, 2006.

H. E. Rauch, C. T. Striebel, and F. Tung. Maximum likelihood estimates of linear dynamic systems. *AIAA Journal*, 3:1445–1450, August 1965.

S. Raza, P. Misra, T. He, and T. Voigt. Bluetooth smart: An enabling technology for the Internet of Things. In *IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pages 155–162, Abu Dhabi, UAE, Oct. 2015.

S. Reece, R. Mann, I. Rezek, and S. Roberts. Gaussian process segmentation of co-moving animals. *Proc. AIP Conference*, 1305(1):430–437, 2011.

S. Remes, M. Heinonen, and S. Kaski. Non-Stationary Spectral Kernels. In *31st Conference on Neural Information Processing Systems*, Long Beach, CA, USA., 2017.

P. Richter and M. Toledano-Ayala. Revisiting Gaussian process regression modeling for localization in wireless sensor networks. *Sensors (Basel)*, 15(9):22587–22615, Sep. 2015.

N. Salman, N. Alsindi, L. Mihaylova, and A. H. Kemp. Super resolution WiFi indoor localization and tracking. In *Sensor Data Fusion: Trends, Solutions, Applications (SDF)*, pages 1–5, Oct 2014.

S. Särkkä. *Bayesian Filtering and Smoothing*. Cambridge University Press., 2013.

S. Särkkä, A. Solin, and J. Hartikainen. Spatiotemporal learning via infinite-dimensional Bayesian filtering and smoothing: A look at Gaussian process regression through Kalman filtering. *IEEE Signal Processing Magazine*, 30(4): 51–61, July 2013.

A. Schwaighofer, M. Grigoras, V. Tresp, and C. Hoffmann. *GPPS: A Gaussian process positioning system for cellular networks*. Number 16 in Advances in Neural Information Processing Systems. MIT Press, 2004.

Senion. White paper: ISP 101, an introduction to indoor positioning systems. URL `https://senion.com/indoor-positioning-101/`.

Y. Shen, A. Y. Ng, and M. Seeger. Fast Gaussian process regression using KD-trees. In *Proc. Advances in Neural Information Processing Systems*. MIT Press, 2006.

X. Shi, Y. H. Chew, C. Yuen, and Z. Yang. A RSS-EKF localization method using HMM-based LOS/NLOS channel identification. In *IEEE International Conference on Communications (ICC)*, pages 160–165, June 2014.

M. I. Skolnik. *Introduction to Radar Systems*. McGraw Hill Book Co., New York, 2 edition, 1980.

B. L. Smith and M. J. Demetsky. Short-term traffic flow prediction: neural network approach. *Transportation Research Record*, (1453), 1994.

E. Snelson and Z. Ghahramani. Sparse Gaussian processes using pseudo-inputs. In *International Conference on Advances in Neural Information Processing Systems*, pages 1257–1264. MIT press, 2006.

J. Talvitie, M. Renfors, and E. S. Lohan. Distance-based interpolation and extrapolation methods for RSS-based localization with indoor wireless signals. *IEEE Transactions on Vehicular Technology*, 64(4):1340–1353, April 2015.

P. Tichavsky, C. H. Muravchik, and A. Nehorai. Posterior Cramer-Rao bounds for discrete-time nonlinear filtering. *IEEE Transactions on Signal Processing*, 46 (5):1386–1396, May 1998.

M. K. Titsias. Variational learning of inducing variables in sparse Gaussian processes. In *Proc. Artificial Intelligence and Statistics*, pages 567–574, 2009.

R. Turner, M. Deisenroth, and C. Rasmussen. State-space inference and learning with Gaussian processes. In Yee Whye Teh and Mike Titterington, editors, *The Thirteenth International Conference on Artificial Intelligence and Statistics*, volume 9 of *Proceedings of Machine Learning Research*, pages 868–875, Chia Laguna Resort, Sardinia, Italy, 13–15 May 2010. PMLR.

J. Wagberg, D. Zachariah, T. Schön, and P. Stoica. Prediction Performance After Learning in Gaussian Process Regression. In Aarti Singh and Jerry Zhu, editors, *The 20th International Conference on Artificial Intelligence and Statistics*, volume 54 of *Proceedings of Machine Learning Research*, pages 1264–1272, Fort Lauderdale, FL, USA, 20–22 Apr 2017. PMLR.

J. Wallis. *A treatise of algebra, both historical and practical microform : shewing the original, progress, and advancement thereof, from time to time, and by what steps it hath attained to the heighth at which it now is : with some additional treatises*. John Playford, for Richard Davis London, 1685.

E. A. Wan and R. Van Der Merwe. The unscented kalman filter for nonlinear estimation. In *IEEE Adaptive Systems for Signal Processing, Communications, and Control Symposium*, pages 153–158, Alberta, Canada, Oct 2000.

R. Want, A. Hopper, V. Falcão, and J. Gibbons. The active badge location system. *ACM Transactions on Information Systems*, 10(1):91–102, Jan. 1992.

S. Weisberg. *Applied Linear Regression*. John Wiley & Sons, 3rd edition, 2013.

D. Wu and L. Xia. Reference device-assisted adaptive location fingerprinting. *Sensors*, 16(6):802, 2016.

H. Wymeersch, S. Marano, W. M. Gifford, and M. Z. Win. A machine learning approach to ranging error mitigation for UWB localization. *IEEE Transactions on Communications*, 60(6):1719–1728, June 2012. ISSN 0090-6778. doi: 10.1109/TCOMM.2012.042712.110035.

Y. Xie, K. Zhao, Y. Sun, and D. Chen. Gaussian processes for short-term traffic volume forecasting. *Transportation Research Record: Journal of the Transportation Research Board*, 2165:69–78, 2010.

Z. Yang, Z. Zhou, and Y. Liu. From RSSI to CSI: Indoor localization via channel response. *ACM Computing Surveys*, 46(2):25, 2013.

K. Yedavalli, B. Krishnamachari, S. Ravula, and B. Srinivasan. Ecolocation: a sequence based technique for RF localization in wireless sensor networks. In *Information Processing in Sensor Networks (IPSN)*, pages 285–292, Los Angeles, California, USA, April 2005.

F. Yin. *Robust Wireless Localization in Harsh Mixed Line-of-Sight/Non-Line-of-Sight Environments*. PhD thesis, Technische Universitaet Darmstadt, Darmstadt, Germany, Aug. 2014.

F. Yin, C. Fritsche, F. Gustafsson, and A. M. Zoubir. TOA-based robust wireless geolocation and Cramér-Rao lower bound analysis in harsh LOS/NLOS environments. *IEEE Transactions on Signal Processing*, 61(9):2243–2255, May 2013.

F. Yin, Y. Zhao, and F. Gunnarsson. Proximity report triggering threshold optimization for network-based indoor positioning. In *18th International Conference on Information Fusion (FUSION)*, pages 1061–1069, Washington D.C., USA, July 2015.

F. Yin, Y. Zhao, and F. Gunnarsson. Fundamental bounds on position estimation using proximity reports. In *IEEE 83rd Vehicular Technology Conference: VTC2016-Spring*, pages 1–5, Nanjing, China, 2016.

F. Yin, Y. Zhao, F. Guannarsson, and F. Gustafsson. Received-signal-strength threshold optimization using Gaussian processes. *IEEE Transactions on Signal Processing*, 65(8):2164–2177, 2017.

Y. Zhang and Y. Xie. Forecasting of short-term freeway volume with v-support vector machines. *Transportation Research Record: Journal of the Transportation Research Board*, 2024:92–99, 2008.

Y. Zhao. *Lab Report: Localization Using a Microphone Network, TSRT14 Sensor Fusion*, 2015.

Y. Zhao, F. Yin, F. Gunnarsson, M. Amirijoo, E. Özkan, and F. Gustafsson. Partile filtering for positioning based on proximity report. In *18th International Conference on Information Fusion (FUSION)*, pages 1046–1052, Washington D.C., USA, July 2015.

Y. Zhao, F. Yin, F. Gunnarsson, M. Amirijoo, and G. Hendeby. Gaussian processes for propagation modeling and proximity based indoor positioning. In *IEEE 83rd Vehicular Technology Conference: VTC2016-Spring*, pages 1–5, Nanjing, China, May 2016a.

Y. Zhao, F. Yin, F. Gunnarsson, F. Hultkratz, and J. Fagerlind. Gaussian processes for flow modeling and prediction of positioned trajectories evaluated with sports data. In *19th International Conference on Information Fusion (FUSION)*, pages 1461–1468, July 2016b.

Y. Zhao, C. Fritsche, and F. Gunnarsson. Parametric lower bound for nonlinear filtering based on Gaussian process regression model. In *20th International Conference on Information Fusion (FUSION)*, pages 1–7, Xi'an, China, July 2017a.

Y. Zhao, C. Fritsche, F. Yin, F. Gunnarsson, and F. Gustafsson. Supplementary material for "Sequential Monte Carlo methods and theoretical bounds for proximity report based indoor positioning". Tech. Rep. LiTH-ISY-R, 3101, Linköping University, Department of Electrical Engineering, Linköping, Sweden, 2017b. URL `http://www.control.isy.liu.se/publications/`.

Y. Zhao, C. Fritsche, F. Yin, F. Gunnarsson, and F. Gustafsson. Sequential Monte Carlo methods and theoretical bounds for proximity report based indoor positioning. *IEEE Transactions on Vehicular Technology*, 67(6):5372 – 5386, 2018a.

Y. Zhao, C. Liu, L. S. Mihaylova, and F. Guannarsson. Gaussian processes for RSS fingerprints construction in indoor localization. In *21st International Conference on Information Fusion (FUSION)*, pages 1377 – 1384, Cambridge, UK, July 2018b.

Y. Zhao, C. Fritsche, G. Hendeby, F. Yin, T. Chen, and F. Gunnarsson. Cramér-Rao bounds for filtering based on Gaussian process state-space models. *Submitted to IEEE Transactions on Signal Processing*, 2019.

L. Zuo, R. Niu, and P. K. Varshney. A sensor selection approach for target tracking in sensor networks with quantized measurements. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 2521–2524, March 2008.

Part II

Publications

The papers associated with this thesis have been removed for copyright reasons. For more details about these see:

http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-153944

PhD Dissertations
Division of Automatic Control
Linköping University

M. Millnert: Identification and control of systems subject to abrupt changes. Thesis No. 82, 1982. ISBN 91-7372-542-0.
A. J. M. van Overbeek: On-line structure selection for the identification of multivariable systems. Thesis No. 86, 1982. ISBN 91-7372-586-2.
B. Bengtsson: On some control problems for queues. Thesis No. 87, 1982. ISBN 91-7372-593-5.
S. Ljung: Fast algorithms for integral equations and least squares identification problems. Thesis No. 93, 1983. ISBN 91-7372-641-9.
H. Jonson: A Newton method for solving non-linear optimal control problems with general constraints. Thesis No. 104, 1983. ISBN 91-7372-718-0.
E. Trulsson: Adaptive control based on explicit criterion minimization. Thesis No. 106, 1983. ISBN 91-7372-728-8.
K. NordstrÃűm: Uncertainty, robustness and sensitivity reduction in the design of single input control systems. Thesis No. 162, 1987. ISBN 91-7870-170-8.
B. Wahlberg: On the identification and approximation of linear systems. Thesis No. 163, 1987. ISBN 91-7870-175-9.
S. Gunnarsson: Frequency domain aspects of modeling and control in adaptive systems. Thesis No. 194, 1988. ISBN 91-7870-380-8.
A. Isaksson: On system identification in one and two dimensions with signal processing applications. Thesis No. 196, 1988. ISBN 91-7870-383-2.
M. Viberg: Subspace fitting concepts in sensor array processing. Thesis No. 217, 1989. ISBN 91-7870-529-0.
K. Forsman: Constructive commutative algebra in nonlinear control theory. Thesis No. 261, 1991. ISBN 91-7870-827-3.
F. Gustafsson: Estimation of discrete parameters in linear systems. Thesis No. 271, 1992. ISBN 91-7870-876-1.
P. Nagy: Tools for knowledge-based signal processing with applications to system identification. Thesis No. 280, 1992. ISBN 91-7870-962-8.
T. Svensson: Mathematical tools and software for analysis and design of nonlinear control systems. Thesis No. 285, 1992. ISBN 91-7870-989-X.
S. Andersson: On dimension reduction in sensor array signal processing. Thesis No. 290, 1992. ISBN 91-7871-015-4.
H. Hjalmarsson: Aspects on incomplete modeling in system identification. Thesis No. 298, 1993. ISBN 91-7871-070-7.
I. Klein: Automatic synthesis of sequential control schemes. Thesis No. 305, 1993. ISBN 91-7871-090-1.
J.-E. StrÃűmberg: A mode switching modelling philosophy. Thesis No. 353, 1994. ISBN 91-7871-430-3.
K. Wang Chen: Transformation and symbolic calculations in filtering and control. Thesis No. 361, 1994. ISBN 91-7871-467-2.
T. McKelvey: Identification of state-space models from time and frequency data. Thesis No. 380, 1995. ISBN 91-7871-531-8.
J. SjÃűberg: Non-linear system identification with neural networks. Thesis No. 381, 1995. ISBN 91-7871-534-2.
R. Germundsson: Symbolic systems – theory, computation and applications. Thesis No. 389, 1995. ISBN 91-7871-578-4.

P. Pucar: Modeling and segmentation using multiple models. Thesis No. 405, 1995. ISBN 91-7871-627-6.
H. Fortell: Algebraic approaches to normal forms and zero dynamics. Thesis No. 407, 1995. ISBN 91-7871-629-2.
A. Helmersson: Methods for robust gain scheduling. Thesis No. 406, 1995. ISBN 91-7871-628-4.
P. Lindskog: Methods, algorithms and tools for system identification based on prior knowledge. Thesis No. 436, 1996. ISBN 91-7871-424-8.
J. Gunnarsson: Symbolic methods and tools for discrete event dynamic systems. Thesis No. 477, 1997. ISBN 91-7871-917-8.
M. Jirstrand: Constructive methods for inequality constraints in control. Thesis No. 527, 1998. ISBN 91-7219-187-2.
U. Forssell: Closed-loop identification: Methods, theory, and applications. Thesis No. 566, 1999. ISBN 91-7219-432-4.
A. Stenman: Model on demand: Algorithms, analysis and applications. Thesis No. 571, 1999. ISBN 91-7219-450-2.
N. Bergman: Recursive Bayesian estimation: Navigation and tracking applications. Thesis No. 579, 1999. ISBN 91-7219-473-1.
K. EdstrÃűm: Switched bond graphs: Simulation and analysis. Thesis No. 586, 1999. ISBN 91-7219-493-6.
M. Larsson: Behavioral and structural model based approaches to discrete diagnosis. Thesis No. 608, 1999. ISBN 91-7219-615-5.
F. Gunnarsson: Power control in cellular radio systems: Analysis, design and estimation. Thesis No. 623, 2000. ISBN 91-7219-689-0.
V. Einarsson: Model checking methods for mode switching systems. Thesis No. 652, 2000. ISBN 91-7219-836-2.
M. NorrlÃűf: Iterative learning control: Analysis, design, and experiments. Thesis No. 653, 2000. ISBN 91-7219-837-0.
F. TjÃďrnstrÃűm: Variance expressions and model reduction in system identification. Thesis No. 730, 2002. ISBN 91-7373-253-2.
J. LÃűfberg: Minimax approaches to robust model predictive control. Thesis No. 812, 2003. ISBN 91-7373-622-8.
J. Roll: Local and piecewise affine approaches to system identification. Thesis No. 802, 2003. ISBN 91-7373-608-2.
J. Elbornsson: Analysis, estimation and compensation of mismatch effects in A/D converters. Thesis No. 811, 2003. ISBN 91-7373-621-X.
O. HÃďrkegÃĕrd: Backstepping and control allocation with applications to flight control. Thesis No. 820, 2003. ISBN 91-7373-647-3.
R. Wallin: Optimization algorithms for system analysis and identification. Thesis No. 919, 2004. ISBN 91-85297-19-4.
D. Lindgren: Projection methods for classification and identification. Thesis No. 915, 2005. ISBN 91-85297-06-2.
R. Karlsson: Particle Filtering for Positioning and Tracking Applications. Thesis No. 924, 2005. ISBN 91-85297-34-8.
J. Jansson: Collision Avoidance Theory with Applications to Automotive Collision Mitigation. Thesis No. 950, 2005. ISBN 91-85299-45-6.
E. Geijer Lundin: Uplink Load in CDMA Cellular Radio Systems. Thesis No. 977, 2005. ISBN 91-85457-49-3.
M. Enqvist: Linear Models of Nonlinear Systems. Thesis No. 985, 2005. ISBN 91-85457-64-7.

T. B. SchÃűn: Estimation of Nonlinear Dynamic Systems — Theory and Applications. Thesis No. 998, 2006. ISBN 91-85497-03-7.

I. Lind: Regressor and Structure Selection — Uses of ANOVA in System Identification. Thesis No. 1012, 2006. ISBN 91-85523-98-4.

J. Gillberg: Frequency Domain Identification of Continuous-Time Systems Reconstruction and Robustness. Thesis No. 1031, 2006. ISBN 91-85523-34-8.

M. Gerdin: Identification and Estimation for Models Described by Differential-Algebraic Equations. Thesis No. 1046, 2006. ISBN 91-85643-87-4.

C. GrÃűnwall: Ground Object Recognition using Laser Radar Data – Geometric Fitting, Performance Analysis, and Applications. Thesis No. 1055, 2006. ISBN 91-85643-53-X.

A. Eidehall: Tracking and threat assessment for automotive collision avoidance. Thesis No. 1066, 2007. ISBN 91-85643-10-6.

F. Eng: Non-Uniform Sampling in Statistical Signal Processing. Thesis No. 1082, 2007. ISBN 978-91-85715-49-7.

E. Wernholt: Multivariable Frequency-Domain Identification of Industrial Robots. Thesis No. 1138, 2007. ISBN 978-91-85895-72-4.

D. Axehill: Integer Quadratic Programming for Control and Communication. Thesis No. 1158, 2008. ISBN 978-91-85523-03-0.

G. Hendeby: Performance and Implementation Aspects of Nonlinear Filtering. Thesis No. 1161, 2008. ISBN 978-91-7393-979-9.

J. SjÃűberg: Optimal Control and Model Reduction of Nonlinear DAE Models. Thesis No. 1166, 2008. ISBN 978-91-7393-964-5.

D. TÃűrnqvist: Estimation and Detection with Applications to Navigation. Thesis No. 1216, 2008. ISBN 978-91-7393-785-6.

P-J. Nordlund: Efficient Estimation and Detection Methods for Airborne Applications. Thesis No. 1231, 2008. ISBN 978-91-7393-720-7.

H. Tidefelt: Differential-algebraic equations and matrix-valued singular perturbation. Thesis No. 1292, 2009. ISBN 978-91-7393-479-4.

H. Ohlsson: Regularization for Sparseness and Smoothness — Applications in System Identification and Signal Processing. Thesis No. 1351, 2010. ISBN 978-91-7393-287-5.

S. Moberg: Modeling and Control of Flexible Manipulators. Thesis No. 1349, 2010. ISBN 978-91-7393-289-9.

J. Wallén: Estimation-based iterative learning control. Thesis No. 1358, 2011. ISBN 978-91-7393-255-4.

J. D. Hol: Sensor Fusion and Calibration of Inertial Sensors, Vision, Ultra-Wideband and GPS. Thesis No. 1368, 2011. ISBN 978-91-7393-197-7.

D. Ankelhed: On the Design of Low Order H-infinity Controllers. Thesis No. 1371, 2011. ISBN 978-91-7393-157-1.

C. Lundquist: Sensor Fusion for Automotive Applications. Thesis No. 1409, 2011. ISBN 978-91-7393-023-9.

P. Skoglar: Tracking and Planning for Surveillance Applications. Thesis No. 1432, 2012. ISBN 978-91-7519-941-2.

K. Granström: Extended target tracking using PHD filters. Thesis No. 1476, 2012. ISBN 978-91-7519-796-8.

C. Lyzell: Structural Reformulations in System Identification. Thesis No. 1475, 2012. ISBN 978-91-7519-800-2.

J. Callmer: Autonomous Localization in Unknown Environments. Thesis No. 1520, 2013. ISBN 978-91-7519-620-6.

D. Petersson: A Nonlinear Optimization Approach to H2-Optimal Modeling and Control. Thesis No. 1528, 2013. ISBN 978-91-7519-567-4.

Z. Sjanic: Navigation and Mapping for Aerial Vehicles Based on Inertial and Imaging Sensors. Thesis No. 1533, 2013. ISBN 978-91-7519-553-7.

F. Lindsten: Particle Filters and Markov Chains for Learning of Dynamical Systems. Thesis No. 1530, 2013. ISBN 978-91-7519-559-9.

P. Axelsson: Sensor Fusion and Control Applied to Industrial Manipulators. Thesis No. 1585, 2014. ISBN 978-91-7519-368-7.

A. Carvalho Bittencourt: Modeling and Diagnosis of Friction and Wear in Industrial Robots. Thesis No. 1617, 2014. ISBN 978-91-7519-251-2.

M. Skoglund: Inertial Navigation and Mapping for Autonomous Vehicles. Thesis No. 1623, 2014. ISBN 978-91-7519-233-8.

S. Khoshfetrat Pakazad: Divide and Conquer: Distributed Optimization and Robustness Analysis. Thesis No. 1676, 2015. ISBN 978-91-7519-050-1.

T. Ardeshiri: Analytical Approximations for Bayesian Inference. Thesis No. 1710, 2015. ISBN 978-91-7685-930-8.

N. WahlstrÃűm: Modeling of Magnetic Fields and Extended Objects for Localization Applications. Thesis No. 1723, 2015. ISBN 978-91-7685-903-2.

J. Dahlin: Accelerating Monte Carlo methods for Bayesian inference in dynamical models. Thesis No. 1754, 2016. ISBN 978-91-7685-797-7.

M. Kok: Probabilistic modeling for sensor fusion with inertial measurements. Thesis No. 1814, 2016. ISBN 978-91-7685-621-5.

J. Linder: Indirect System Identification for Unknown Input Problems: With Applications to Ships. Thesis No. 1829, 2017. ISBN 978-91-7685-588-1.

M. Roth: Advanced Kalman Filtering Approaches to Bayesian State Estimation. Thesis No. 1832, 2017. ISBN 978-91-7685-578-2.

I. Nielsen: Structure-Exploiting Numerical Algorithms for Optimal Control. Thesis No. 1848, 2017. ISBN 978-91-7685-528-7.

D. Simon: Fighter Aircraft Maneuver Limiting Using MPC: Theory and Application. Thesis No. 1881, 2017. ISBN 978-91-7685-450-1.

C. Veibäck: Tracking the Wanders of Nature. Thesis No. 1958, 2018. ISBN 978-91-7685-200-2.

C. Andersson Naesseth: Machine learning using approximate inference: Variational and sequential Monte Carlo methods. Thesis No. 1969, 2018. ISBN 978-91-7685-161-6.

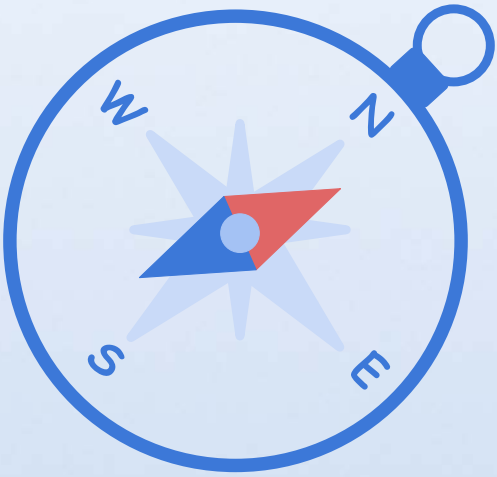
Y. Jung: Inverse system identification with applications in predistortion. Thesis No. 1966, 2018. ISBN 978-91-7685-171-5.

FACULTY OF SCIENCE AND ENGINEERING

Linköping Studies in Science and Technology, Dissertation No. 1968, 2019
Department of Electrical Engineering

Linköping University
SE-581 83 Linköping, Sweden

www.liu.se



LINKÖPING UNIVERSITY