

# Spline Filter for Target Tracking

# SPLINE FILTER FOR TARGET TRACKING

BY

DONNA LYNN KOCHERRY, B.Tech.

A THESIS

SUBMITTED TO THE DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING

AND THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

© Copyright by Donna Lynn Kocherry, September 2010

All Rights Reserved

Master of Applied Science (2010)  
(Electrical & Computer Engineering)

McMaster University  
Hamilton, Ontario, Canada

TITLE: Spline Filter for Target Tracking

AUTHOR: Donna Lynn Kocherry  
B.Tech., (Electrical and Electronics Engineering)  
National Institute of Technology, Calicut, India

SUPERVISOR: Dr. Thiagalingam Kirubarajan

NUMBER OF PAGES: viii, 58

*I would like to dedicate this thesis to my precious family, my parents and my sister.*

# Abstract

In this thesis an efficient approach to nonlinear non-Gaussian state estimation based on spline filtering is presented. The estimation of the conditional probability density of the unknown state can be ideally achieved through Bayes rule. However, the associated computational requirements make it impossible to implement this online filter in practice. In the general particle filtering problem, estimation accuracy increases with the number of particles at the expense of increased computational load. In this thesis, B-Spline interpolation is used to represent the density of the state pdf through a low order continuous polynomial. The motivation is to reduce the computational load and to improve accuracy. The motion of spline control points and corresponding coefficients is achieved through implementation of the Fokker-Planck equation, which describes the propagation of state probability density function between measurement instants. The solution of the Fokker Planck equation is achieved by calculating the state transition probability matrix. The state transition matrix is calculated using Dirac Feynman approximation. This filter is applicable for a general state estimation problem as no assumptions are made about the underlying probability density. Finally, simulation results are presented to demonstrate the effectiveness of the proposed algorithm.

# Acknowledgements

My extreme gratitude to Almighty God for everything.

I express my most sincere gratitude to my supervisor Dr. T. Kirubarajan for giving me an opportunity to be a graduate student under him and for his continual guidance and inspiration from the start to finish of the journey that culminated in this thesis. His admirable patience with the students and ever enthusiastic nature to help are greatly appreciated. I would like to extend my heartfelt gratitude to Dr. Thomas for his prompt guidance and valuable suggestions throughout my work.

The timely help and administrative assistance from Cheryl Gies and Helen Jachna are appreciated. I thank Terrence Greenlay for helping me with solving my computer related problems. I am thankful to my friends Ali, Biruk, Herath, Xiaofan, Chen, Kaibing, Darcy and others in the ETF Laboratory for their companionship and support. Most importantly, I would like to acknowledge the priceless contribution of my parents in teaching me the immense value of education. I am deeply grateful to them, as they have made great sacrifices to make my sister and me what we are now. Finally, but definitely not the least, I thank my sister for being there for me always.

# Contents

<b>Abstract</b>	<b>iv</b>
<b>Acknowledgements</b>	<b>v</b>
<b>1 Introduction and Problem Statement</b>	<b>1</b>
1.1 Nonlinear Filtering Theory . . . . .	1
1.2 Problem Statement . . . . .	6
1.3 Contribution of the Thesis . . . . .	7
<b>2 Nonlinear Filtering Theory</b>	<b>10</b>
2.1 Nonlinear Bayesian Filtering . . . . .	10
2.2 Nonlinear Filter Representation via Spline Functions . . . . .	12
2.3 Nonlinear Estimation using Monospline . . . . .	15
2.4 Spline Filter for Nonlinear/Non-Gaussian Bayesian Tracking . . . . .	17
2.5 Spline Filter for Multidimensional Nonlinear/Non-Gaussian Bayesian Tracking . . . . .	20
2.6 Spline Histogram Method for Reconstruction of PDF . . . . .	21
2.7 Nonlinear Filters with Particle Flow Induced by Log-Homotopy . . . . .	23
2.8 Nonlinear Estimation for a Class of Systems . . . . .	26

2.9	Cubature Kalman Filters . . . . .	28
2.9.1	Square - Root Cubature Kalman Filter . . . . .	30
2.10	Spline Filter . . . . .	31
2.11	Particle filtering . . . . .	32
<b>3</b>	<b>Spline in Estimation</b>	<b>34</b>
3.1	B-Splines . . . . .	34
3.1.1	Positivity Property of B-Splines . . . . .	35
3.1.2	Integral of B-splines . . . . .	36
3.1.3	Construction of B-splines . . . . .	36
3.1.4	Multidimensional Spline . . . . .	36
<b>4</b>	<b>The Theory</b>	<b>39</b>
4.0.5	Current working Approach . . . . .	39
4.0.6	Update Step . . . . .	42
4.0.7	Finding Estimate . . . . .	43
4.0.8	Initialization . . . . .	43
<b>5</b>	<b>Simulations and Results</b>	<b>46</b>
5.1	SPLINE FILTER . . . . .	46
5.1.1	Simulation . . . . .	46
5.2	Results . . . . .	49
<b>6</b>	<b>Summary</b>	<b>52</b>
6.1	Conclusion . . . . .	52
6.2	Future work and improvements possible . . . . .	53



# List of Figures

4.1	Gaussian sample Values . . . . .	44
4.2	The initial probability density function . . . . .	45
5.1	The probability density at the final instant of Spline filter and numerical filter . . . . .	50
5.2	Comparing the truth and the calculated results . . . . .	51

# Chapter 1

## Introduction and Problem Statement

### 1.1 Nonlinear Filtering Theory

Most real life systems are nonlinear in nature. A nonlinear system is one in which the variables to be solved cannot be written as a combination of independent components. Or it can be defined as a system in which the output of the system is not a linear function of the input of the system. Nonlinear systems are of interest to researchers because most natural systems are inherently nonlinear and dynamic. A dynamic system is one where the parameters of the system change with time as is observed in most natural systems. So it is usually very difficult to exactly predict the output of a real system based on knowledge of the input. The variable that is to be determined in any system can be described using its state. The state of a dynamic system completely defines it. It is defined as the minimal set of variables, the knowledge of which at the present time instant, along-with the knowledge of the future

input applied for a time period, enables one to analyze the response of the system for the future time instant. The equations governing the dynamics of the state of a variable in a nonlinear system are usually not known exactly. For decades researchers have been trying to find methods to estimate the state of a variable in a nonlinear system based on measurements of system parameters. Since system parameters cannot be directly measured they have to be estimated from observations made from a distance. Estimation is the process of inferring the value of a quantity of interest from indirect, inaccurate and uncertain observations. Tracking is the estimation of the state of a moving object based on remote measurements. Estimation provides a systematic approach to deal with the discrepancies in different measurements on the same object.

Observability of a system is a concept that is useful to check whether the states of a system can be reconstructed or estimated by observation of the outputs and inputs of the system as described in (K.P.Mohandas, 2006). A linear system is observable in an interval  $[t_0, t_1]$  if for any initial state  $\mathbf{x}(t_0)$ , knowing the input  $u(t)$  and the output  $\mathbf{y}(t)$  over the same interval, it is possible to solve for any state  $\mathbf{x}(t)$ . This essentially means that it would be possible to find  $\mathbf{x}(t_0)$  once  $u(t)$  and  $\mathbf{y}(t)$  are known. If this property holds for any initial time  $t_0$  and any initial state  $\mathbf{x}(t_0)$ , the system is said to be completely observable.

Nonlinear filtering is the process of estimation of the current state of a dynamic nonlinear system. There are a great variety of applications for state estimation in the real world, a few of which are listed below.

- Tracking and Surveillance
- Navigation of ships and aircrafts

- Control Systems
  - Making of Steel
  - Chemical, nuclear, manufacturing and industrial processes
- Power systems to identify faults and leakages
- Signal processing
- Geophysical systems to estimate ground faults, earthquakes etc

These and many other applications in the real world contribute to raise the significance of nonlinear estimation.

The problem of estimating the state of a linear system has been studied in depth by researchers over the years. Gauss made the following (philosophical) observations on the (physical) observations that could be made on the planetary motion as described in (Bar-Shalom *et al.*, 2001):

- If the observations were absolutely correct, the parameters could be determined with perfect accuracy from a minimum number of observations (n observations for n parameters).
- Then subsequent observations would confirm, but not correct, the values obtained for the parameters.
- But, since the observations are only approximations of the truth, we should combine more observations than the minimum to determine more accurately the unknown quantities.

- Thus, starting with approximate knowledge, we can correct it with subsequent observations so as to satisfy all the observations in the most accurate manner possible.

The above (philosophical) observations lead to the following implications:

- A basic description (model) of the system is available with some unknown parameters to be estimated.
- Redundant data are required to reduce the effect of measurement errors.
- To satisfy all the observations in the most accurate manner possible, the residuals (differences between the observed values and the values predicted from the estimates) should be as small as possible.
- Inaccuracy of observations necessitates probabilistic modeling.
- Combination of the initial knowledge and the subsequent observations leads to the recursive algorithm concept.

Estimation theory provides a systematic approach to deal with the discrepancies in different measurements on the same object. Under the Gaussian assumption for the initial state and all the noises entering into the system, the Kalman filter is the optimal Minimum Mean Square Error state estimator. Various measures are used to gauge linear filter algorithms such as consistency, optimality and stability. For non gaussian random variables the Kalman filter algorithm is found to be the best linear estimator or the Linear Minimum Mean Square Error estimator (LMMSE).

The problem of nonlinear estimation however is more complex. The concept of Information state was developed to be a function of the available information set (the

cumulative set of observations from the beginning to the current time instant and the set of known inputs prior to the current time instant) that completely summarizes the past of the system in a probabilistic sense. The conditional pdf  $p(t_k, x|Y(t_{k-1}))$  of the state effectively can be called the information state of the system. The Extended Kalman Filter(EKF) was developed initially to estimate the state in nonlinear systems. It works on a principle of linearization of the nonlinearities of the system and uses the LMMSE estimation technique. The prediction step of filtering is achieved in the EKF using the Taylor series expansion of the nonlinear terms in the state around the latest estimate with first order terms for the first order EKF and the second order terms for the second order EKF respectively. The EKF differs from the Kalman filter in its covariance calculations, which are now no longer decoupled from the state estimate calculations and cannot be done off-line. However, the EKF was found to be unstable as the inherent approximations in it could lead to the divergence of the filtering algorithm (unbounded estimation errors). These errors occurred mainly due to the Taylor series expansions which were used in the EKF. The expansions lead to errors because the higher order terms which went neglected may not always have been negligible. Also the calculations of the covariance matrices are done using the estimated or predicted values of the state rather than the exact values. The Iterated Extended Kalman filter was developed to improve accuracy and later on the Unscented Kalman Filter (UKF) was developed. Researchers have developed various filtering methods such as Particle filtering in an effort to solve the hurdles of nonlinear estimation. Some of them have been described in detail in Chapter 2.

## 1.2 Problem Statement

The task addressed in this thesis is to estimate the state of a nonlinear system based on measurements of the system output. The two main steps involved in the process are the prediction step and the correction or update step. In the prediction step, the state of the target at a future time instant needs to be forecasted. The optimal state estimator needs to compute the conditional pdf of the state given all the information available at the present time instant  $t_k$ : the prior information about the initial state, the intervening inputs and the measurements through time  $t_k$ . The approach used in this thesis uses Transition probabilities of the state probability density function. Integrating the Chapman-Kolmogorov equation or the prediction equation over the entire domain of the state probability density function is the difficulty faced in determining the predicted density. Continuous integration over the entire domain of the state pdf is practically impossible and numerical approximations are usually used. This raises the computational complexity of the problem exponentially with an increase in the number of dimensions of the state. The state probability density function is a continuous function but it is usually discretized in particle filtering methods in order to evaluate the Chapman-Kolmogorov prediction equation. The approach used in this thesis is to represent it using spline polynomials so as to get an accurate representation using fewer parameters. An attempt is made to strike a satisfactory balance between accuracy of the representation and the resulting computational complexity involved. The approach used for the correction or update step in this thesis involves the Bayes' equation.

## 1.3 Contribution of the Thesis

Many problems in science require estimation of the state of a system that changes over time using a sequence of noisy measurements made on the system. In this thesis, the state-space approach to modeling dynamic systems is used, and the focus will be on the continuous-discrete time formulation of the problem. Thus, a continuous differential equation is used to model the evolution of the states with time and the measurements are assumed to be available at discrete times as governed by a discrete time difference equation. For dynamic state estimation, the continuous-discrete-time approach is widespread and convenient. The state-space approach to time-series modeling focuses attention on the state vector of a system. The state vector contains all relevant information required to describe the system under investigation. For example, in tracking problems, this information could be related to the kinematic characteristics of the target.

This thesis presents a multivariate spline filter as a solution to implementation problem for nonlinear Bayesian estimator. Generally, a good representation of the probabilities and likelihood functions is essential to the success of the filtering algorithm in nonlinear estimation problems. A variety of numerical methods have been developed to address the problem, which include the Extended Kalman filter(EKF) (Bar-Shalom *et al.*, 2001), unscented Kalman filter(UKF) (Julier and Uhlmann, 1997) and various Sequential Monte Carlo filters (Arulampalam *et al.*, 2002) (Gordon *et al.*, 1993) (Daum, 2005). The idea of using discrete splines and spline filtering is discussed in (Ustuner and Ferrari, 1992). A varied implementation of the particle filter using log homotopy is discussed in (Daum and Huang, 2008). In particle filtering, the computational load increases exponentially with number of particles as described



in (Daum and Huang, 2003). A continuous-discrete state estimation problem is presented in this thesis. A multi-dimensional B-spline based approach is used to obtain a continuous representation of the probability density function over the target space.

In the literature, methods to construct splines for representing a probability density are widely analyzed in various application areas as described in (Docenko and Berzins, 2003), (Zhang and Yue, 2004), (Lee *et al.*, 1997), (Lin *et al.*, 2004) and (Kimeldorf and Wahba, 1970). While working with clustered data, spline methods for representing the data can account for cluster correlation and are non-local as described in (Kimeldorf and Wahba, 1970). Since the elements in the state vector are cross correlated, spline methods are used in this thesis to represent the probability density of the state as opposed to kernel methods. An approach using mono-splines to address the problem of nonlinear estimation using quadratures is presented in (Wang and Klein, 1976). This approach assumes that the process and measurement noises are zero mean Gaussian. In (Bucy and Youssef, 1974), another method is presented to obtain the nonlinear estimate of phase demodulation with a two dimensional phase process model as an alternative to the fourier filter. In (Yu and Deng, 2009), a method for solving nonlinear estimation problems using splines is presented in which the nonlinear function is converted to linear in a higher dimensional space. Spline based filtering for target tracking allows easy manipulation of the density functions since they are polynomials of finite order. For instance, it is a straightforward process to find the target probability in any region in the state space by evaluating the integral of the spline representation over that region.

However, to be of better practical use in estimation, it is essential to extend the univariate filter to a multi-dimensional one. This is achieved using tensor product

approach. Although the complexity of the tensor product approach increases exponentially with dimension, complexity can be reduced by exploiting the structure of the linear system that needs to be solved to construct the spline polynomial. Once the target probability density function is represented by an equivalent B-spline representation, the prediction and update steps are implemented. For prediction, a practical implementation of the Fokker-Planck-Kolmogorov equation is used. In solving the Fokker-Planck-Kolmogorov equation no assumptions are made about the probability density of the state. The predicted pdf at each instant is used in the update step to find the final updated density of the state. The resulting multi-dimensional filter can be used in most of the common circumstances and the density could even be multi-modal.

# Chapter 2

## Nonlinear Filtering Theory

The problem of state estimation of nonlinear continuous time stochastic systems with discrete time measurements is of special interest in nonlinear state estimation. The optimal nonlinear filtering estimator or the Bayesian estimator is described below.

### 2.1 Nonlinear Bayesian Filtering

The general target dynamic model, which governs the evolution of the state sequence is assumed to be of a continuous discrete nature. This model is used because the system state actually evolves continuously with time but measurements are only available at discrete time instants due to the physical limitations of sensors. The system state is assumed to evolve in continuous time according to

$$\dot{X}(t) = F[t, X(t)] + G(t, X(t))\eta_t \quad (2.1)$$

where  $X(t) = [x_1, x_2, \dots, x_n]'_t$  represents the state vector of the system at any time  $t$ ,  $F[t, X(t)]$  is possibly a nonlinear function vector valued function of the state  $X(t)$ ,  $G(t, X(t))$  is an  $n \times m$  real matrix and  $\eta_t$  is a zero mean continuous time white process noise with intensity  $\tilde{Q}(t)$  and autocorrelation

$$E[\eta_t \eta_{t'}] = Q_t \delta(t - t') \quad (2.2)$$

The objective of target tracking is to recursively estimate  $X(t)$  from the measurements obtained at discrete times as

$$Z(t_k) = H_k(X(t_k)) + w_k \quad (2.3)$$

where  $Z(t_k)$  is the measurement at time instant  $t_k$ ,  $H_k$  is, in general, a nonlinear function and  $w_k$  is an i.i.d. measurement noise sequence of known statistics.

In filtering, the main purpose is estimating the probability distribution of the state  $X(t)$  conditioned up to the current time step. This leads to the pdf  $p(X(t_k)|Z(1:k))$ , which can be obtained probabilistically in two steps: prediction and update. It is assumed that the initial pdf  $p(X(0)|Z(0))$  of the state vector is available. Suppose the prior pdf  $p(X(t_{k-1})|Z(1:k-1))$  at time  $t = k-1$  is available. Then the pdf associated with the predicted state can be obtained using the continuous system model (2.1) via Kolmogorov forward or Fokker-Planck equation as described in (Bar-Shalom *et al.*, 2001).

$$\frac{\partial}{\partial t} p[X(t)] = - \sum_{i=1}^n \frac{\partial [p[X(t)] F[t, X(t)]_i]}{\partial X(t)_i} + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial [p(GQG^T)_{ij}]}{\partial X(t)_i \partial X(t)_j} \quad t_k > t > t_{k-1} \quad (2.4)$$

where  $p(X(t))$  denotes the state density at time  $t$  and  $Z(1 : i) = Z(1), \dots, Z(i)$  denotes the set of measurements up to time  $i$ .

The pdf associated with the update state is given by Bayes equation

$$p(X(t_k)|Z(1 : k)) = \frac{p(Z(t_k)|X(t_k))p(X(t_k)|Z(1 : k - 1))}{p(Z(t_k)|Z(1 : k - 1))} \quad (2.5)$$

where the denominator

$$p(Z(t_k)|Z(1 : k - 1)) = \int p(Z(t_k)|X(t_k))p(X(t_k)|Z(1 : k - 1))dX(t_k) \quad (2.6)$$

is a normalization term. The normalization term depends on the likelihood function  $p(Z(t_k)|X(t_k))$  defined by the measurement model (2.3). Over the years researchers have developed various practical implementations of the Fokker-Planck equation and some of those are discussed in the following sections. Various filtering methods have been developed by researchers over the years and a review of some of their work is presented here.

## 2.2 Nonlinear Filter Representation via Spline Functions

One of the crucial factors responsible for success of a nonlinear filter is the accurate representation of the conditional probability density of the signal, given the observations. In (Bucy and Youssef, 1974), the one dimensional basis used is spline under tension and multidimensional basis is generated by a tensor product of the one dimensional basis. Under Spline generation, the spline basis is discussed, the compact

support of the spline basis will be 4 intervals which will be chosen equal, each piece satisfying this common differential locally over one interval.

$$\frac{d^4}{dx^4}B_i(x) + \sigma^2 \frac{d^2}{dx^2}B_i(x) = 0$$

The spline coefficients of one dimensional basis  $C_i$  are then uniquely determined.

The 2 dimensional basis is generated by tensor product of each one dimensional basis and spline coefficients  $C_{ij}$  are determined by matrix inversion. The matrix involved is a kronecker product of two tridiagonal matrices each of which can be inverted explicitly. The next section of the paper is devoted to getting a closed form solution of the matrix inversion with generalized boundary conditions. The matrix can be separated and the matrix manipulation becomes a two vector manipulation. The nonlinear filter problem considered is that of two dimensional phase demodulation. The model used is the following

$$x_{n+1}^1 = x_n^1 + \Delta x_n^2 x_{n+1}^2 = x_n^2 + u_n$$

where  $x^1$  is the phase,  $x^2$  the phase rate,  $\Delta$  time step and  $u_n$  a Gaussian noise sequence with zero mean and covariance  $\Delta q$  independent of initial state vector, which is gaussian zero mean with covariance  $r$ . The phase is observed by the following sensors:

$$z_n^1 = \cos x_n^1 + v_n^1 z_n^2 = \sin x_n^1 + v_n^2$$

$v_n^1$  and  $v_n^2$  are i.i.d. Gaussian noise sequences zero mean and covariance  $\frac{r}{\Delta}$ .  $J_n(x, y)$  is the conditional density of  $x^1, x^2$  given all  $z^j$  with  $j = 1 \dots n$ .

The update for the density is

$$J_{n+1}(x, y) = K_{n+1} S_{n+1}(x) \int_{-\infty}^{\infty} \exp \frac{-1}{2q\Delta} (y - \xi)^2 J_n(x - \xi\Delta), \xi) d\xi$$

with

$$S_n(x) = \exp\left(\frac{z_n^1 \cos x + z_n^2 \sin x}{\frac{r}{\Delta}}\right)$$

And the optimal estimate is given by

$$x_n^* = \arctan \sin \hat{x}_n / \cos \hat{x}_n$$

where

$$\cos \hat{x}_n + i \sin \hat{x}_n = \int_{-\Pi}^{\Pi} \int_{-\Pi/\Delta}^{\Pi/\Delta} \exp i\xi J_n(\xi, \eta) d\xi d\eta$$

The density  $J_n(x, y)$  is represented using splines by

$$J_n(x, y) = \sum_{i=1}^{M+3} \sum_{j=1}^{N+3} B_i(x) C_{ij}^m B_j(y)$$

The estimate time and accuracy depends on matrix inversion and the convolution matrix. The closed form solution was found to be better than factorization using continuant fraction for matrix inversion. It was concluded in the discussed paper (Bucy and Youssef, 1974) that the spline filter was not as fast as the Fourier filter but was retained more accurate density information and was faster than the cyclic point mass filter.

## 2.3 Nonlinear Estimation using Monospline

Nonlinear estimation of state of a discrete time dynamical system is a recursive algorithm arising from Bayes' Law. Numerical approaches include curve fitting the conditional density and quadrature approximation to the integration. The essential step in obtaining the conditional density of the state given the observations is integration. In (Wang and Klein, 1976) a method using quadrature formulae related to monosplines with optimal knots is used. The system is represented by

$$x_{k+1} = f(x_k) + w_k,$$

where  $x_k, w_k \in R^n$

and the measurement is represented by

$$y_k = h(x_k) + v_k$$

where  $y_k, v_k \in R^m$ .

Noise sequences are assumed independent gaussian random vectors with zero mean. Assuming that initial state pdf is known, the conditional density function of the state given the observations can be computed by



$$\phi(x_k|Y_k) = \frac{u(y_k|x_k)p(x_k|Y_{k-1})}{\int u(y_k|x_k)p(x_k|Y_{k-1})dx_k} \quad (2.7)$$

and

$$p(x_{k+1}|Y_k) = \int \tau(x_{k+1}|x_k)(x_k|Y_k)dx_k \quad (2.8)$$

where  $\phi$  is filtering density,  $u$  is measurement density,  $\tau$  is transition density and  $p$  is prediction density.

A monospline is defined in the paper as follows. Let  $S_x$  be a spline of degree  $2m-1$  with distinct knots  $x_1, x_2, \dots, x_n$  in  $[-1, 1]$  with  $k$  fold knots at  $-1$  and  $1$  where  $0 \leq k \leq m$ . Then if a monomial of degree  $2m$ ,  $x^{2m}/(2m)!$  is interpolated by  $s_x$  in  $[-1, 1]$  the difference between the two functions becomes a monospline

$$H(x) = \frac{x^{2m}}{(2m)!} - s(x) = \frac{x^{2m}}{(2m)!} - \sum_{i=1}^{2m} \alpha_i \frac{x^{i-1}}{(i-1)!} - \sum_{j=1}^n B_j \frac{(x-x_j)^{2m-1}}{(2m-1)!} \quad (2.9)$$

Then quadrature formula is induced and when the knots of the monospline, at which the conditional density is discretized, are allowed to vary, a class of optimal quadrature formulas is obtained. The numerical solution for the monospline with optimal knots is obtained by solving  $2m+2n$  nonlinear algebraic equations and due to symmetry of monospline the number of equations is halved. Then a modified Newton's iteration equation is used to solve for the optimal monospline. To maintain positivity of the discretized density function in the recursive equations only those quadrature formulas with  $k = 0$  or  $1$  and all positive weights are useful for evolution of conditional densities. Increasing  $m$  increases the accuracy of the approximation but makes it more sensitive to truncation factor. The quality of the approximation depends on the degree of monospline, number of knots and truncation factor and

the approximation error is reduced by efficiently representing the density function in digital form. The parameters are computed a priori off line by simplifying the nonlinear equations specifying them. Thus an accurate iteration scheme is developed to obtain monospline with optimal knots.

## 2.4 Spline Filter for Nonlinear/Non-Gaussian Bayesian Tracking

Nonlinear estimation of state of a discrete time dynamical system implements a recursive algorithm arising from Bayes' Law. A good representation of the probability densities and likelihood functions is an essential issue in the success of the filtering algorithm especially in nonlinear case. In this paper (Punithakumar and Kirubaranjan, 2007), a spline based approach is used to give a continuous representation of the probability densities over the target state space.

The state equation is given by

$$x_k = f_k(x_{k-1}) + v_k$$

where  $f_k$  is a possibly nonlinear function of the state  $x_{k-1}$  and  $v_k$  is an i.i.d. process noise sequence of known statistics.

The measurement equation is given by

$$z_k = h_k(x_k) + w_k$$

where  $h_k$  is a possibly nonlinear function and  $w_k$  is an i.i.d. measurement noise

sequence of known statistics. The purpose is to find  $p(x_k|z_{1:k})$ . The first step is prediction where the predicted density  $p(x_k|z_{1:k-1})$  is found, and it is assumed the initial pdf  $p(x_0|z_0)$  is available. The predicted state is obtained from the prior pdf at time  $k-1$  using

Chapman-Kolmogorov equation.

$$p(x_k|z_{1:k-1}) = \int p(x_k|x_{k-1})p(x_{k-1}|z_{1:k-1})dx_{k-1} \quad (2.10)$$

The updated pdf is calculated in the paper (Punithakumar and Kirubarajan, 2007) by

$$p(x_k|z_{1:k}) = \frac{p(z_k|x_k)p(x_k|z_{1:k-1})}{p(z_k|z_{1:k-1})} \quad (2.11)$$

$$\text{And } p(z_k|z_{1:k-1}) = \int p(z_k|x_k)p(x_k|z_{1:k-1})dx_k$$

B-splines are introduced and their properties are discussed namely positivity property, knot insertion property and stable evaluation of B-splines. The integral of B-spline, construction of B-Spline and the Schoenberg-Whitney theorem regarding invertibility is also discussed.

Spline filter is discussed in detail. The spline representation of the prior density is used.  $B_{p,k-1}x_{k-1} \approx p(x_{k-1}|z_{1:k-1})$

The optimal prediction step involving state equation is approximated by assuming the density in the region outside the interval  $[t_1(k-1)t_s(k-1)]$  to be zero and by decomposing the interval  $[t_1(k-1)t_s(k-1)]$  into  $N_s$  cells  $\{x_{k-1}^i : i = 1, \dots, N_s\}$  and

evaluating density at these points. The  $i$  th point is calculated by

$$y_{k-1}^i = B_{k-1|k-1}(x_{k-1}^i)x_k^i = f_{k-1}(x_{k-1}^i) + n_{k-1}^i \quad (2.12)$$

Then spline interpolation for the predicted density is constructed using these points and the new interval is set to  $[\hat{t}_1(k-1) \cdots \hat{t}_s(k-1)]$

To update, the spline representation  $B_{l,k}(x_k)$  of the likelihood function is constructed over the new interval. The same knot sequence is used to reduce computational costs and the updated density is calculated by

$$B_{k|k}(x_k) = \frac{B_{k|k-1}(x_{k-1})B_{l,k}(x_k)}{\int_{\hat{t}_1(k-1)}^{\hat{t}_s(k-1)} B_{k|k-1}(x_{k-1})B_{l,k}(x_k)} \quad (2.13)$$

The insignificant ends of the posterior density are removed and a new interval is selected. The mean of the estimate  $x_k$  at time  $k$  is found as

$$\hat{x}_k = \int_{\hat{t}_1(k)}^{\hat{t}_s(k)} x_k B_{p,k}(x_k) dx_k \quad (2.14)$$

The developed spline filter and Sample Importance Re-sampling (SIR) particle filter are compared in the paper (Punithakumar and Kirubarajan, 2007) by simulation for a one dimensional nonlinear estimation problem and it is found that the spline filter is computationally more expensive than the particle filter for same performance but it retains the probability density function and gives a more natural representation over the discrete approximation. Also finding the target probability over a region in spline filter is straight forward.

## 2.5 Spline Filter for Multidimensional Nonlinear/Non-Gaussian Bayesian Tracking

The univariate spline filter approach presented in (Punithakumar and Kirubarajan, 2007) is extended to multi-dimensional spaces using tensor product approaches in (K.Punithakumar *et al.*, 2008). The nonlinear estimation problem is and B-splines and their properties are the same as described in (Punithakumar and Kirubarajan, 2007). Tensor product of two linear spaces of functions in abstract algebra is defined and bivariate B-spline is obtained as

$$B_{ij}(x, y) = \sum_{i=1}^m \sum_{j=1}^n B_{i,h,s}(x) B_{j,k,t}(y) \alpha_{ij}$$

The construction of spline polynomials involves solving set of equations and the construction of bivariate spline involves solving  $mn$  equations for  $mn$  unknowns. Triangular factorization is used to reduce computational costs from  $(mn)^3$  to  $(m^3 + m^2n + mn^2 + n^3)$ . The paper then describes the prediction and update steps of the optimal recursive Bayesian filtering using univariate spline representation of density and likelihood functions for single dimensional problem (Punithakumar and Kirubarajan, 2007).

The univariate spline filter is extended to multi-dimensional spaces using the tensor product approach. The multi-dimensional spline representation of the density function is constructed using the univariate spline polynomials of each individual elements of the state space  $x_k$ . To reduce computational requirements, the spline is constructed only over the region in which the target density is significant. The prediction and update steps of the multi-dimensional spline filter are generalizations of

those of univariate spline filter.

The multidimensional B-spline filter is illustrated by simulation on a two dimensional example and the results are displayed. The spline filter retains the probability density information in a finite polynomial form and finding the target probability over a region in spline filter is straightforward.

## 2.6 Spline Histogram Method for Reconstruction of PDF

In (Docenko and Berzins, 2003) spline histogram method is used to estimate pdf of a one dimensional data set in automatic and optimal way. The spline histogram method is a non-parametric approach for reconstruction of probability density function underlying statistical selection. Tensioned cubic splines are found to be suitable and the corresponding algorithm has been called TCSplin. From spectroscopic observations, the red-shifts of galaxies in clusters are obtained.

Red-shift of the  $i$ th galaxy denoted by  $z_i$  and ordered in ascending order i.e. ( $z_i \leq z_{i+1}$ ). Next, a step-like Cumulative Distribution Function(CDF) is constructed, obtained purely from observational data:

$$F_{obs}(cz) = N(z_j < z)/N_{gal} \quad (2.15)$$

where  $N(z_j < z)$  is a number of galaxies with red-shift smaller than  $z$ , and  $N_{gal}$  is a total number of detected galaxies,  $c$  is the speed of light. The PDF  $f(x)$  by definition is the derivative of  $F_{obs}(cz)$  in respect to  $cz$ . If CDF is constructed as shown

before then  $f(x)$  is a sum of Dirac  $\delta$ -functions. The points  $z_i$  with ordinates  $F_{obs}(cz_i)$  are consequently connected by non-decreasing smooth analytical spline  $S(cz)$ . Then analytically differentiated leading to the PDF estimate  $\hat{f}(cz)$ . Continuous PDF thus obtained matches the discrete distribution of the data points. Tensioned cubic splines (TCS) are defined such that the cubic polynomial spline length between two data points is minimal, and only the interpolating function and its first derivative are continuous in data points. When TCS derivative becomes negative non tensioned splines are used by increasing spline length. Close data points give unrealistic high pdf peaks, which are smoothed. Savitzky - Golay filters are used in the paper because noise is in the form of narrow high peaks in pdf.

Integrated Square Error (ISE) function is used to define the optimal width such that noise is reduced but pdf is not over smoothed.

$$ISE(\hat{f}(cz)) = \int_{cz_{min}}^{cz_{max}} (\hat{f}(cz) - f(cz))^2 d(cz) \quad (2.16)$$

is the true pdf and  $\hat{f}(cz)$  is estimated pdf. For discrete filter  $P(h)$  is minimized for smoothing width  $h$  and can be easily calculated

$$P(h) = \sum_{i=1}^N (\hat{f}(cz_i)^2 - 2\hat{f}(cz_i) + 2C_0^{(h)}) \quad (2.17)$$

Statistical noise level is found using bootstrapping technique. Using the obtained spline histogram as a true PDF, the same amount  $N$  of random numbers are generated and another smoothed spline histogram is computed. This is repeated and the average value of the simulated spline histograms and their scattering is calculated.

The simulated distributions were compared statistically with the original values

and the average values agreed till  $1 \sigma$  but the standard deviations are about 10% larger due to the smoothing effect. For Gaussian distributions the asymmetry and excess are significantly different from zero, although in non-Gaussian cases they are rather close to original values.

The optimal smoothing size depends on the selection volume  $N$  in the following way: theoretically  $h_{opt} \propto N^{-1/5}$  and empirically  $h_{opt} \propto N^{-0.195}$ .

The algorithm was implemented on two clusters of galaxies - Abell 2256 and 3526 and clear structures were obtained demonstrating the quality of the algorithm. A direct implementation of the algorithm leads to a good estimate of the PDF of clusters of galaxies. However the dispersions of the group velocities are overestimated due to PDF smoothing and may be calculated from the original data. In principle the spline histograms may be expanded to higher dimensional cases but sampling noise increases. Signal to noise ratio depends on the distribution character.

## 2.7 Nonlinear Filters with Particle Flow Induced by Log-Homotopy

A new concept in particle filter is explored in (Daum and Huang, 2009) making it radically different and vastly superior to the classic particle filter. The computational complexity of the new filter is many orders of magnitude less than the classic particle filter with optimal estimation accuracy for problems of dimension 4 and above. Log-homotopy is used to implement Bayes' rule with particle flow rather than a point-wise multiplication of two functions. Thus the particles migrate smoothly avoiding particle collapse or degeneracy. Particles are well distributed to represent the prior



probability density of the state vector, but with a new measurement, the likelihood of the measurement is not well represented by this set of particles (no particles near peak of the likelihood function). For higher dimensions particle collapse is a serious problem.

Particle flow induced by log-homotopy solves this problem by smoothly transforming the set of particles used to represent the prior density into a new set of particles that is good for representing the product of the prior and the likelihood (Bayes' rule).

The arrows show the direction of flow and the length of arrow represents speed of flow of each particle. Un-normalized conditional probability density of the dimensional state vector  $x$  is computed using Bayes' rule as follows

$$p(x, t_k | Z_k) = p(z_k | x, t_k) p(x, t_k | Z_{k-1})$$

where  $z_k$  is  $k^{th}$  measurement,

$$Z_k = \{z_1, z_2, \dots, z_k\} \tag{2.18}$$

where  $p(z_k | x, t_k) =$  probability density of measurement  $z_k$  at time  $t_k$  conditioned on  $x$ .

In particle filters Bayes' rule is implemented by point-wise multiplication of two functions leading to degeneracy. The fundamental idea used in this paper is to create a differential equation to implement Bayes' rule, and induce a flow of particles analogous to the natural flow in time induced by the Fokker-Planck equation. Because Bayes'

rule operates at discrete points in time a scalar valued parameter  $\lambda$  varying from 0 to 1 is inserted at each discrete measurement time as synthetic time to create homotopy.

$$f = gh \quad (2.19)$$

where  $g$  is the prior density of the  $d$  dimensional state vector and  $h$  is the density of the measurement.

Define homotopy function as

$$\log(f_\lambda) = \log(g) + \lambda \log(h) \quad (2.20)$$

The partial differential equation is created by differentiating with respect to  $\lambda$  as

$$\frac{\partial \log(f_\lambda)}{\partial \lambda} = \log(h) \quad (2.21)$$

Finally the induced particle flow is simplified as

$$\frac{dx}{d\lambda} = -\log(h) \left( \frac{\partial \log(f_\lambda)}{\partial x} \right)^T / \left\| \frac{\partial \log(f_\lambda)}{\partial x} \right\|^2 \quad (2.22)$$

for nonzero gradient and  $dx/d\lambda = 0$  otherwise.

So the induced flow of particles is in the gradient direction of the log-homotopy, with speed proportional to  $\log(h)$  and the flow stops when the gradient is zero. By integrating ODEs we  $\log(f)$  is obtained i.e. a good set of particles to represent

g\*h effectively solving the particle degeneracy issue. Thus a gradient flow in the log-homotopy is obtained, similar to diffusion sampling Monte Carlo methods, which uses a gradient flow of the log-density. But log-homotopy is not a Monte Carlo method; rather it is completely deterministic and does not use any random numbers.

## 2.8 Nonlinear Estimation for a Class of Systems

The paper (Charalambous and Socratous, 2006) contains nonlinear estimation problems for a class of models and employs relative entropy to describe the uncertainty classes. Two types of uncertainty models are considered.

- Uncertainty Models on Conditional Distributions or otherwise known Stochastic Kernels
- Uncertainty Models on Joint Distributions

This approach leads to minimax techniques, in which the worst case estimate of the uncertain measure subject to the uncertainty description is sought. A measurable space  $(\Omega, F)$  is considered on which the unobserved Random Variable (R.V.),  $X$  and the observed  $Y$  are defined:

$$X : (\Omega, F) \rightarrow (X, \sum x), Y : (\Omega, F) \rightarrow (Y, \sum y) \quad (2.23)$$

Thus  $X$  is the space of the unobserved R.V., and  $Y$  is the space of the observed R.V. The relation between the unobserved R.V.  $X$  and the observed R.V.  $Y$  is defined via a probabilistic mapping  $\mu : X \times \sum y \rightarrow [0, 1]$ . which is a probability measure on  $\sum y$  for every  $x \in X$  and is  $X$  measurable for every  $F \in Y$ . This is called a stochastic

kernel or transition probability.  $X$  is estimated as a function of  $Y$  by introducing a pay off and then minimizing it.

$$J(\Phi^*) = \inf_{\Phi \in \chi_{ad}} \int_{X \times Y} l(x, \Phi(y)) dP_{X,Y}(x, y). \quad (2.24)$$

The true kernel is unknown so the estimation problem is formulated as a minimax problem

$$J_1(\Phi^*, \nu^*) = \inf_{\Phi \in \chi_{ad}} \sup_{\nu \in A(\mu)} \int_{X \times Y} l(x, \Phi(y)) \nu(x, dy) dP_X(x) \quad (2.25)$$

and solved using Lagrangians.

Two other alternatives are discussed using different constraints. Uncertainty on the joint distribution is handled in a similar fashion. This pay off function is considered

$$J_4(\Phi) = \int_{X \times Y} l(x, \Phi(y)) dQ_{X,Y}(x, y) \quad (2.26)$$

where  $Q_{X,Y}$  is the uncertain joint measure,  $P_{X,Y}(x, y)$  is the nominal system model and the true model belongs to an uncertainty set described by  $C(P_{X,Y})$  defined as

$$\{C_{X,Y} \in \mathcal{M}(X \times Y); H(Q_{X,Y} | P_{X,Y}) \leq R\} \quad (2.27)$$

where  $R \in (0, \infty)$  and  $H$  is relative entropy between  $P$  and  $Q$ . The estimation problem is formulated as a mini-max problem defined by

$$J_4(\Phi^*, Q^*) = \inf_{\Phi \in \chi_{ad}} \sup_{Q_{X,Y} \in C(P_{X,Y})} \int_{X \times Y} l(x, \Phi(y)) dQ_{X,Y}(x, y) \quad (2.28)$$

and is solved using Lagrangians.

Some properties associated with the estimate are discussed and some applications are illustrated. Thus the concept of entropy is linked successfully with nonlinear estimation.

## 2.9 Cubature Kalman Filters

A new nonlinear filter for high dimensional state estimation is developed called the cubature kalman filter in (Arasaratnam and Haykin, 2009). A nonlinear filtering problem can be solved sub optimally in two ways - locally by assuming a fixed prior posterior distribution and globally using point mass methods etc. Cubature Kalman filter uses a local approach. The predictive density and the filter likelihood density are assumed Gaussian leading to a Gaussian posterior density. So the Bayesian solution reduces to computation of means and covariances in time and measurement update equations. In the time update the Bayesian filter computes the predicted mean and covariance and in measurement update the predicted measurement, the updated state and associated covariances are found using Kalman filter approach. The heart of the Bayesian filter involves computing Gaussian weighted integrals whose integrands are all of the form nonlinear function Gaussian density.

Numerical integration methods to compute multi-dimensional weighted integrals are discussed. An integral of this form  $I(f) = \int_D f(x)w(x)dx$  consisting of function  $f$  and weight  $w$  is solved. A set of points  $x_i$  and weights  $w_i$  are found to approximate the integral as a summation. The gauss-hermite quadrature rule is used to find the gaussian weighted integral for integrals of dimension  $< 5$  . For higher dimension integrals, non product rules like randomized Monte Carlo methods, quasi-Monte Carlo

methods, lattice rules and sparse grids are used, which choose points directly from the domain of integration. A non product cubature rule is derived whose complexity increases linearly with  $n$ . The higher degree cubature rule yields a more accurate solution but the solution is more tedious. The symmetry of the system is made use of to simplify the process.

The cubature Kalman filter is discussed in detail in this paper. An integral of Cartesian form

$$I(f) = \int_{R^n} f(x) \exp(-x^T x) dx$$

is solved by transforming to a spherical radial integration form and then using a third degree spherical radial rule. The variable is changed from Cartesian vector  $x$  to radius  $r$  and direction  $y$  to get the following integral

$$I(f) = \int_0^\infty \int_{U_n} f(ry) r^{n-1} \exp(-r^2) d\sigma(y) dr$$

The spherical and radial integrals are numerically computed by spherical cubature and gaussian quadrature rules. These are defined and a spherical-radial cubature rule is given

$$\int_{R^n} f(x) \exp(-x^T x) dx \approx \sum_{j=1}^{m_s} \sum_{i=1}^{m_r} a_i b_j f(r_i s_j)$$

Proofs are presented for the above equations in (Arasaratnam and Haykin, 2009). Higher degree cubature rules are avoided as the third degree gives a sufficient approximation and higher degrees may make matters worse due to extremely high computation costs. Also third degree exhibits efficient and robust computation. Use of higher degree rules may sabotage the performance of the CKF

### 2.9.1 Square - Root Cubature Kalman Filter

The square root extension of the CKF is developed while preserving the symmetry and positive definiteness of the error covariance matrix. Positive definiteness may be lost by matrix square-rooting, matrix inversion, matrix squared-form amplifying round off errors and subtraction of the two positive definite matrices present in the covariance update. The square-root version of the CKF or Square-root Cubature Kalman Filter (SCKF) was developed. The SCKF propagates square-root factors of the predictive and posterior error covariances. Hence, matrix square-rooting operations are avoided. In addition, the SCKF preserves symmetry and positive (semi)definiteness of the covariance, improves numerical accuracy and has doubled-order precision.

The SCKF is developed using the least-squares method for the Kalman gain and matrix triangular factorizations or triangularizations (e.g., the QR decomposition) for covariance updates thus avoiding matrix inversion and square-rooting respectively. The Unscented Kalman filter and the CKF are compared. They share a common property, both the UKF and the CKF use a weighted set of symmetric points. But the sigma point set built into UKF has some limitations over CKF namely numerical inaccuracy, unavailability of a square-root solution and filter instability. Overall the CKF is found to be more accurate and principled in mathematical terms than the sigma point approach. The CKF is applied to two nonlinear estimation problems and simulated results are displayed. The general conclusion drawn is that CKF outperforms UKF, CDKF etc. and is a new and improved algorithmic addition to nonlinear filtering.

## 2.10 Spline Filter

As a result of reviewing the papers described in the previous sections it can be concluded that using spline polynomials for representing likelihood functions and density functions is ideal. Spline is preferred because it retains probability and gives a more natural representation than discrete approximations like particle filter. Also finding target probability over a region in spline filter is straight-forward. The optimum nonlinear filter derived using Bayes' theorem and Chapman-Kolmogorov equation is nearly impossible to implement owing to the necessity for storing the entire pdf and iterative integrations, but the spline filter can be a very accurate approximation. The prediction step in paper (Punithakumar and Kirubarajan, 2007) can be used with slight modification. The step of splitting into cells the interval of prior density may perhaps be avoided and some other alternative may be attempted. I think that the knot sequence used to construct the spline interpolation of the likelihood function in the update stage of the spline filter must be updated. After developing the nonlinear filter for single dimension it can then be extended to multiple dimensions by tensor products concept similar to paper (K.Punithakumar *et al.*, 2008).

Another option would be to use the splitting concept in the prediction stage but to use the concept of log homotopy or particle flow as described in paper (Daum and Huang, 2009). By developing a transformation for the spline coefficients so as to represent the product of two functions greater accuracy may perhaps be achieved without an excessive increase in the computation costs. With more research and knowledge of splines, the spline filter can be further optimized and can be the best near optimal nonlinear filter with regard to accuracy, computation costs, accuracy and better representation of density function.



## 2.11 Particle filtering

Particle filters are a novel class of nonlinear filters introduced about a decade before (Gordon *et al.*, 1993). There are a wide variety of particle filters. The basic working principle of most particle filters can be described as following:

- Possible system state sequences (particles) are simulated using importance sampling and resampling.
- At time  $k$ , the particles already include state sequences up to time  $k - 1$ . A proposition of the system state at time  $k$  is added to each of them.
- The particles are weighted based on how well the new states fit the new observation.
- The particles are resampled according to the weights  $\rightarrow$  Particles with high weights are multiplied, particles with low weights may disappear.
- A discrete representation of the filtering distribution is formed using the particles, a state estimate can be obtained from this.

Particle filters approximate the complete non-Gaussian probability density of the state vector conditioned on the measurements. Multidimensional integration is approximated using Monte Carlo sampling, which is the hallmark of particle filters. The key issue in nonlinear filters of any kind is “the curse of dimensionality”, a phrase coined by Richard Bellman over forty years ago to describe the exponential growth of computational complexity as a function of the dimension of the state vector  $X$ . The computational complexity of the Kalman filter grows as the cube of dimension but for the general nonlinear problems using filters that achieve optimal accuracy,

the computational complexity grows exponentially in dimension. A detailed analysis of the computational complexity of particle filters for a given estimation accuracy is given in (Daum, 2005). For low dimensional problems, well designed particle filters achieve optimal estimation accuracy with a computational complexity roughly the same as a EKF or Extended Kalman Filter. But in general for high dimensional problems the computational complexity of particle filters is enormous.

# Chapter 3

## Spline in Estimation

### 3.1 B-Splines

This section briefly introduces B-Splines and their properties. The reader is referred to (Boor, 2001) and (Prautzsch *et al.*, 2002) for a detailed background on splines. The spline interpolation in B-form is given by (Boor, 2001)

$$\mathbf{B}(x) = \sum_{j=1}^n \mathbf{B}_{j,p,t}(x) \alpha_j \quad (3.1)$$

where

$$\mathbf{B}_{j,p,t}(x) \triangleq (t_{j+p} - t_j)[t_j, \dots, t_{j+p}](\cdot - x)_+^{p-1} \quad (3.2)$$

is the  $j$ th B-spline of order  $p$  for the given knot sequence  $t$ . The notation  $[t_j, \dots, t_{j+p}]g$  stands for the divided difference of order  $p$  of  $g$  at the sites  $t_j, \dots, t_{j+p}$  and  $(x)_+ = \max\{x, 0\}$  stands for the truncation function. The placeholder notation is used to

indicate that the  $k$  th divided difference of the function  $(t-x)_+^{p+1}$  of the two variables  $t$  and  $x$  is to be taken by fixing  $x$  and considering it as a function of  $t$  alone.

From definition, the first order B-spline is given by

$$\mathbf{B}_{j,1,t} = \begin{cases} 1 & \text{if } t_j \leq x \leq t_{j+1} \\ 0 & \text{otherwise} \end{cases} \quad (3.3)$$

Starting from the first order, the higher order B-splines can be constructed using the recurrence relation

$$\mathbf{B}_{j,p,t} = w_{j,k} \mathbf{B}_{j,p-1,t} + (1 - w_{j+1,p}) \mathbf{B}_{j+1,p-1,t} \quad (3.4)$$

with

$$w_{j,k}(x) \triangleq \frac{x - t_j}{t_{j+p-1} - t_j} \quad (3.5)$$

### 3.1.1 Positivity Property of B-Splines

The B-spline  $B_{j,p,t}$  is made up of at most  $p$  nontrivial polynomial pieces and vanishes outside the interval  $[t_j, \dots, t_{j+p}]$  and is positive on the interior of that interval (Boor, 2001) i.e.,

$$\mathbf{B}_{j,p,t}(x) > 0, \quad t_j < x < t_{j+p} \quad (3.6)$$

while

$$t_j = t_{j+p} \Rightarrow \mathbf{B}_{j,p,t}(x) = 0 \quad (3.7)$$

With the positivity property of B-splines and having  $\alpha_i \geq 0$  for all  $i$ , the positivity of spline representation of probability and likelihood functions are ensured.

### 3.1.2 Integral of B-splines

The integral of B-spline, for  $t_1 \leq x \leq t_s$ , is given by

$$\int_{t_1}^x \sum_{i=1}^n \alpha_i \mathbf{B}_{ip}(y) dy = \sum_{i=1}^{s-1} \left( \sum_{j=1}^i \alpha_j (t_{j+p} - t_j) / p \right) \mathbf{B}_{i,p+1}(x) \quad (3.8)$$

### 3.1.3 Construction of B-splines

The basis sequence of B-splines  $B_j$  for  $j = 1, \dots, n$  are linearly independent (Boor, 2001). Thus the space defined by the span of the B-spline basis functions  $B_j$ 's is n-dimensional. If the strictly increasing sequence  $t = \tau_1, \tau_2, \dots, \tau_n$  of data sites is given, then for a given function  $g$ , the spline  $f \triangleq \sum_{j=1}^n \alpha_j \mathbf{B}_j$  agrees with  $g$  at  $\tau$ , if

$$\sum_{j=1}^n \alpha_j \mathbf{B}_j(\tau_i) = g(\tau_i) \quad \text{for } i = 1, \dots, n \quad (3.9)$$

This is a linear system of  $n$  equations with  $n$  unknown  $\alpha_i$ 's with coefficient matrix  $(B_j(\tau_i))$ , the spline collocation matrix.

### 3.1.4 Multidimensional Spline

The spline definition given above can be generalized to estimate multidimensional state vectors using the tensor product approach (K.Punithakumar *et al.*, 2008) (Boor, 2001). A tensor product of two linear spaces of functions in abstract algebra is defined as follows:

Let  $U$  and  $V$  be linear space of functions defined on real sets  $X$  and  $Y$ , respectively.

Then for each  $u \in U$  and  $v \in V$ , the function

$$w(x, y) \triangleq u(x)v(y) \quad (x, y) \in X \times Y \quad (3.10)$$

is called the tensor product of  $u$  with  $v$  and denoted by  $u \otimes v$ .

Further, the set of all finite linear combinations of functions on  $X \times Y$  of the form  $u \otimes v$  is called the tensor product of  $U$  with  $V$  and is given by

$$U \otimes V \triangleq \left\{ \sum_{i=1}^n \alpha_i (u_i \otimes v_i) : \alpha_i \in \mathbf{R}, u_i \in U, v_i \in V, \quad i = 1, \dots, n \right\} \quad (3.11)$$

Using this approach, a bivariate B-spline is obtained as

$$B_{ij}(x, y) = \sum_{i=1}^m \sum_{j=1}^n \mathbf{B}_{i,h,s}(x) \mathbf{B}_{j,k,t}(y) \alpha_{ij} \quad (3.12)$$

This concept can be extended to any finite dimensional space and thus provides a mechanism to construct multidimensional B-splines.

The construction of spline polynomials involve solving set of linear equations (Boor, 2001) and, the construction of bivariate spline given in (3.12) require solving  $mn$  equations for  $mn$  unknowns. The straight-forward solution by gaussian elimination thus requires order of  $(mn)^3$  operations. However by constructing a triangular factorization of the linear system, the computational cost would be reduced to the order of  $(m^3 + m^2n + mn^2 + n^3)$ . The multidimensional spline representation of the state density function is constructed using tensor product of univariate spline polynomials of each individual elements of the state space  $\mathbf{x}_k$ . In order to reduce the computational requirements, the spline is constructed only over the region in which

the target density is significant.

# Chapter 4

## The Theory

This section describes the prediction and update steps of the optimal recursive Bayesian filtering using spline representation of density and likelihood functions as described in (Punithakumar and Kirubarajan, 2007) and (K.Punithakumar *et al.*, 2008).

### 4.0.5 Current working Approach

Initially the measurements of the actual data or the truth are available. The system being considered is a single dimensional nonlinear system with continuous state equation as

$$\dot{x}(t) = f(x(t)) + v(t) \quad (4.1)$$

where  $x(t), f(x(t)), v(t) \in R^n$  with covariance  $Q_{ij} = \hbar\nu\delta_{ij}$  and a discrete measurement equation given by

$$y(t_k) = h(x(t_k), w(t_k), t_k) \quad (4.2)$$

where  $y(t) \in R^{m \times 1}, h \in R^{m \times 1}$  and the noise process  $w(t)$  is assumed to be a white noise process. It is assumed that  $p(y(t_k)|x(t_k))$  is known. To define the transition



probability matrix a wide range of possible values of the state variable  $X$  is selected. The probabilities of state transition between the various possible states are calculated using Dirac Feynman approximation as discussed in (Balaji, 2010b) and (Balaji, 2010a). Transition probabilities are represented as a Probability transition matrix  $P$ , where  $P = [p_{ij}]$  and

$$p_{ij} \triangleq Pr[X(t+1) = j | X(t) = i] \quad (4.3)$$

each row sum must be 1.

Suppose the initial probability density is represented by the prior  $\sigma_0(x)$ . Then the evolution of the states is represented by the Fokker Planck- Kolmogorov forward equation.

State estimation consists of mainly two parts - Prediction step and Correction or Update step

### **Prediction step**

The probability of state  $X$  taking value  $x'$  at time  $t$  is represented as  $p(t, x')$ . Let  $t'' > t'$  and let the prior density be represented by  $p(t', x')$ . The predicted density at time  $t''$  is obtained as

$$p(t'', x') = \int P(t'', x'' | t', x') p(t', x') d^n x' \quad (4.4)$$

The equation described above is realized by representing the probability density and the Transition probability matrix as spline polynomials. In particle filter methodology

it can be approximated numerically as follows:

$$p(t_k, x|Y(t_{k-1})) = \sum_{x'} P(t_k, x|t_{k-1}, x')p(t_{k-1}, x'|Y(t_{k-1})) \quad (4.5)$$

where  $Y(t_{k-1}) =$  All the measurements till time instant  $t_{k-1}$

This product can be achieved using matrix multiplication in cases where the range of values that the state variable  $X$  can take are predefined and limited.

However, in reality the range cannot be predefined so easily. A good representation of the probability densities is an essential issue in the success of the filtering algorithm especially in the case of nonlinear filtering since the probability densities could be multi-modal. For an accurate and efficient representation, the state density is represented using B-Splines. In the existing particle filter approach, the accuracy of the prediction depends on the number of particles being used. The probability distribution is approximated using a large number of particles and a point-wise matrix multiplication is used to achieve equation (4.4) as shown in (Balaji, 2010a). By using spline polynomials it is hoped that the state density can be represented accurately using a minimal number of spline control points. Hence to implement equation (4.4), it is required to multiply a two dimensional spline representing Transition Probability matrix with a One- dimensional spline representing the prior state density and the result needs to be a one dimensional spline representing predicted state density.

### Imposing pdf property to spline

When state density is represented using spline polynomials, the values of the probability density must be kept positive. When the derivative of the density spline  $\frac{\partial}{\partial t}p[X(t)]$  is found, it has certain negative values corresponding to the points where the slope

is negative. However, when the prior density is added to it, the negative part of the density should be canceled to leave a positive predicted state density. The predicted pdf must always be positive for every value of  $X$ . Thus, a condition is imposed to ensure that the spline representation predicted in the problem must be positive for every value of  $X$ .

### Moving Spline Points

The domain of the state space having a significant mass of the pdf should be identified for implementing the Prediction step effectively. However, if the limits of the domain are set too narrow then the evolved pdf may have significant mass outside the selected domain. If the limits of the domain are set too wide then the associated computational complexity increases without a significant improvement in accuracy. The limits of the domain have to be optimally chosen keeping aforementioned problems in mind. However, the state pdf evolves in time and hence the limits of the state variable domain need to be recursively evaluated for each propagation time instant to represent the state density more accurately.

#### 4.0.6 Update Step

The spline representation of predicted density  $B_{k|k-1}(X(t_k))$  is available and the update step involves solving (2.5). This can be done by constructing the spline representation  $B_{l,k}(X(t_k))$  of the likelihood function  $p(Z(t_k)|X(t_k))$  over the interval  $[t'_1(k-1), \dots, t'_s(k-1)]$ . The same knot sequence is used to construct the spline interpolation of the likelihood function in order to reduce the computational cost of multiplying  $B_{k|k-1}(X_k)$  and  $B_{l,k}(X_k)$ . The spline interpolation of the update density

is then calculated as

$$B_{k|k}X(t_k) = \frac{B_{k|k-1}X(t_k)B_{l,k}X(t_k)}{\int_{t_1(k-1)}^{t_s(k-1)} B_{k|k-1}X(t_{k-1})B_{l,k}X(t_k)} \quad (4.6)$$

The spline function for the updated posterior density  $B_{k|k}(X(t_k))$  is now available. However, in many cases it is possible to remove the insignificant end portion of the posterior density with lower target probability. This is done by finding the integral value of  $B_{k|k}(X(t_k))$  over interval  $[t'_{1+i}(k-1) \ t'_{1+i+1}(k-1)]$  for each  $i$  starting from 1 until the first value exceeds a threshold value. The procedure is repeated for  $[t'_{s-i}(k-1) \ t'_{s-i+1}(k-1)]$  for each  $i$  starting from 1 and these intervals are removed. The new interval is set to  $[t_1(k) \dots t_s(k)]$ .

#### 4.0.7 Finding Estimate

The mean of the state estimate  $\hat{X}(t_k)$  at time step  $t = k$  is calculated by

$$\int_{t_1(k)}^{t_s(k)} X(t_k)B_{p,k}(X(t_k))dX(t_k) \quad (4.7)$$

Finding the higher moments is also straightforward with spline filter approach.

#### 4.0.8 Initialization

The spline filter is initialized by constructing the spline interpolation  $B_{0|0}$  of the pdf  $p(X(0) | Z(0))$  of initial target state. The initial state probability density function is available as a set of data points, which consist of value of the state variable and corresponding probability density values. Suppose the initial probability density of the state vector is available as a set of data points.

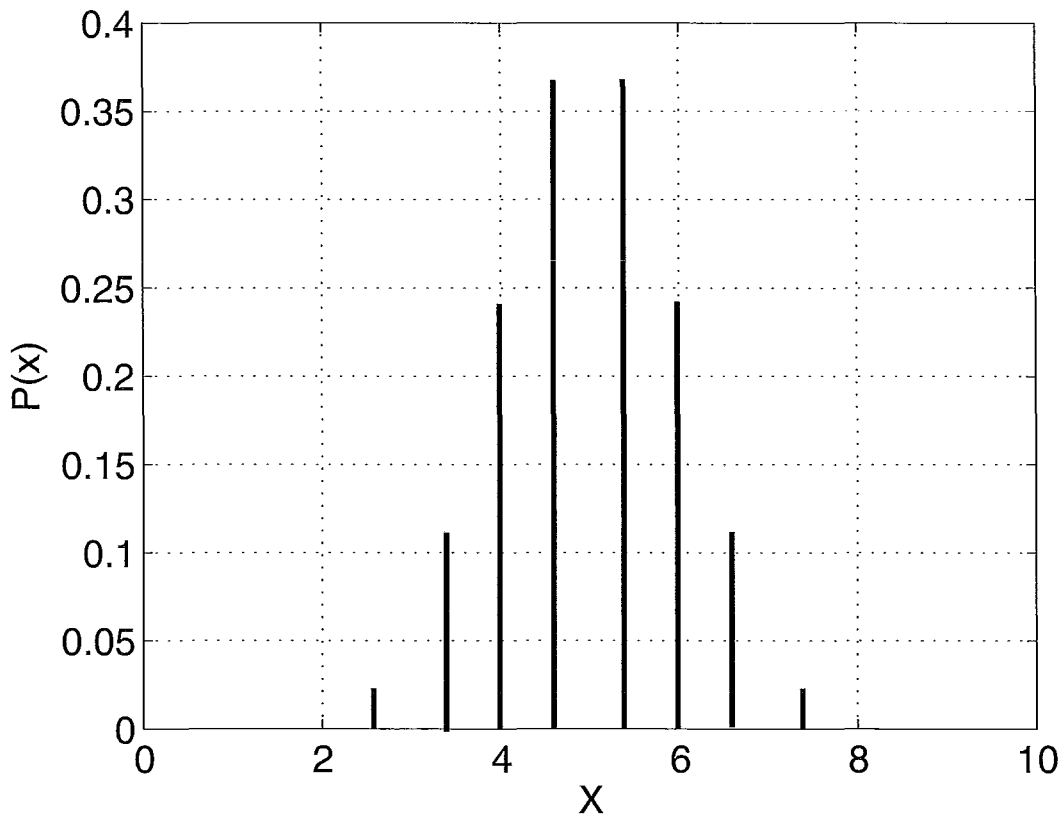


Figure 4.1: Gaussian sample Values

A B-spline interpolation can be fitted smoothly to the available data as shown in the figure below.

The values of the Spline polynomials (denoted as + signs on the graph), evaluated at the knots (the  $X$  coordinates corresponding to the + signs) are the same as the values of the probability density function of the state vector at those knot points.

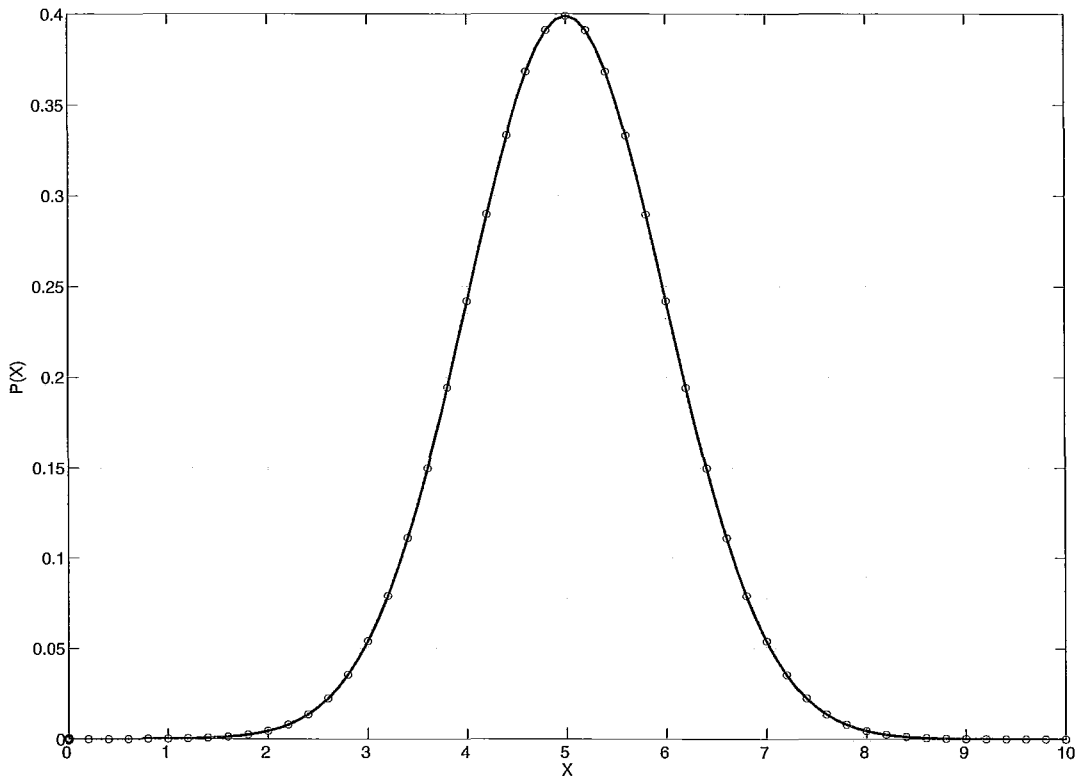


Figure 4.2: The initial probability density function

# Chapter 5

## Simulations and Results

### 5.1 SPLINE FILTER

#### 5.1.1 Simulation

The simulation is explained in detail here. In this simulation, single dimensional tracking problem with the following state model is considered.

$$\dot{x}(t) = 1.2 \cos(3x(t)) + v(t) \quad (5.1)$$

where  $x(t), f(x(t)), v(t) \in \mathbb{R}^n$  with covariance  $Q_{ij} = \hbar \nu \delta_{ij}$ . The prediction step requires the solution of the Fokker Planck forward equation and is achieved through numerical approximation as described in (Balaji, 2010b) as follows:

$$P(t_k, x | Y(t_{k-1})) = \sum_{x'} P(t_k, x | t_{k-1}, x') p(t_{k-1}, x' | Y(t_{k-1})) \quad (5.2)$$

where  $Y(t_{k-1}) =$  All the measurements till time instant  $t_{k-1}$ .

To proceed with the prediction step the transition probability matrix,  $P$  has to be calculated. The Transition Probability matrix can be found using the Dirac Feynman algorithm described in detail in (Balaji, 2010b).

Initially a range is fixed for possible  $X$  values as  $-1.5 \leq X \leq 1.5$  with the values being uniformly distributed and having 0.01 increments between successive values. Next the Transition Probability matrix is calculated. It is assumed that the times  $t'$  and  $t''$  are such that  $t'' > t'$ . The One step pre-point Dirac Feynman approximation as described in (Balaji, 2010a) for additive noise is

$$P(t'', x'' | t', x') = \frac{1}{\sqrt{(2\Pi(t'' - t'))^n \det g(t')}} \exp \left( -\frac{t'' - t'}{2} \sum_{i,j=1}^n \left[ \frac{(x''_i - x'_i)}{(t'' - t')} f_i(x', t') \right] \right. \\ \left. g_{ij}^{-1}(t') \left[ \frac{(x''_j - x'_j)}{(t'' - t')} - f_j(x', t') \right] \right). \quad (5.3)$$

For multiplicative noise it is

$$P(t'', x'' | t', x') = \frac{1}{\sqrt{(2\Pi(t'' - t'))^n \det g(t')}} \exp \left[ -(t'' - t') L^{(r,0)} \left( t, x', x'', \frac{(x'' - x')}{(t'' - t')} \right) \right] \quad (5.4)$$



where the Lagrangian  $L^{(r,0)}(t, x', x'', \frac{(x''-x')}{(t''-t')})$  is given by

$$\begin{aligned}
L^{(r,0)}(t, x', x'', \frac{(x''-x')}{(t''-t')}) &= \frac{1}{2} \sum_{i,j=1}^n \left( \frac{(x''_i - x'_i)}{(t'' - t')} - f_i(x' + r(x'' - x'), t') \right) \\
&\quad \left( \sum_{a,b=1}^{p_e} e_{ia}(x', t') Q_{ab}(t') e_{jb}(x', t') \right)^{-1} \\
&\quad \left( \frac{(x''_j - x'_j)}{(t'' - t')} - f_j(x' + r(x'' - x'), t') \right) \\
&\quad + r \sum_{i=1}^n \frac{\partial f_i}{\partial x_i}(x' + r(x'' - x'), t) \tag{5.5}
\end{aligned}$$

$g_{ij}(t) = \sum_{a,b=1}^{p_e} p_e e_{ia}(t) Q_{ab}(t) e_{bj}(t)$  and  $r \in [0,1]$  specifies the discretization of the stochastic system as  $x^{(r)}(t) = x(t - \Delta t) + r(x(t) - x(t - \Delta t))$ . For this problem the noise is additive and equation(5.3) is used to find the transition probability matrix. A symmetric discretization is chosen and the value of r is fixed at 0.5. The transition probability values are calculated for every combination of states is found. The probability of state transition from  $Y = x(t = jj)$  to  $X = x(t = ii)$  for a  $x^{(r)}(t) = X1r$ ,  $\sigma_x = 0.3$  and  $dt = t'' - t' = 0.01$  is found to be

$$\begin{aligned}
P(jj, ii) &= \left( \frac{1}{(\sqrt{(2 * \pi * dt)} * \sigma_x)} \right) * \exp(-0.5 * dt * (\frac{1}{\sigma_x^2}) * \\
&\quad \left( \frac{(X - Y)}{dt - 1.2 * \cos(3 * X1r)} \right)^2 + r * dt * 3.6 * \sin(3 * X1r)) \tag{5.6}
\end{aligned}$$

The Transition probability matrix does not change with time and so it is calculated only once in the beginning prior to processing the measurements. The transition probability matrix for this system is found to be a  $301 \times 301$  matrix. The probability density vector is represented using a one-dimensional B-spline. This is done using the 'spaps' command in Matlab. In order to execute the product in (5.2), the Transition

probability matrix is represented as a set of One dimensional Splines. This is necessary as it is not possible to directly multiply the transition probability matrix, represented by a 2-dimensional spline, with the probability density vector, represented by a 1-dimensional spline. This product is required to implement equation(5.2).

The  $i$  th column of the Transition probability matrix represents the probability of evolving from all the various states to the  $i$  th state. The prior distribution spline is multiplied by each column of the Transition Probability matrix to obtain a spline representing the probability of the state variable assuming the value corresponding to the respective row of the transition probability matrix. Each resulting spline is integrated to obtain a single value for each probability. The resulting probabilities are used to recreate the density spline to obtain the predicted density spline. Essentially a sort of numerical approximation has been used and the product in (4.4) is achieved using (5.2). The probability of the state variable assuming each value among the range of possible values is calculated as a number and the density spline is recreated for each iteration of the prediction and update loop. A continuous representation for the density vector and a discretized version of the Transition probability matrix is obtained. Recreating the spline at each prediction step is not optimal and an alternative must be found.

## 5.2 Results

The results of this simulation are compared with the results of numerical algorithm as discussed in (Balaji, 2010a) and are presented below. The spline filtering method provides a result having the slightly better accuracy than the numerical method. The RMSE values are calculated for multiple runs of the code and are found to be 0.0224

for the spline filter and 0.0583 for the numerical method. However, the time taken is slightly more for the Spline filter. This is mainly due to the necessity to recreate the spline at each prediction instant.

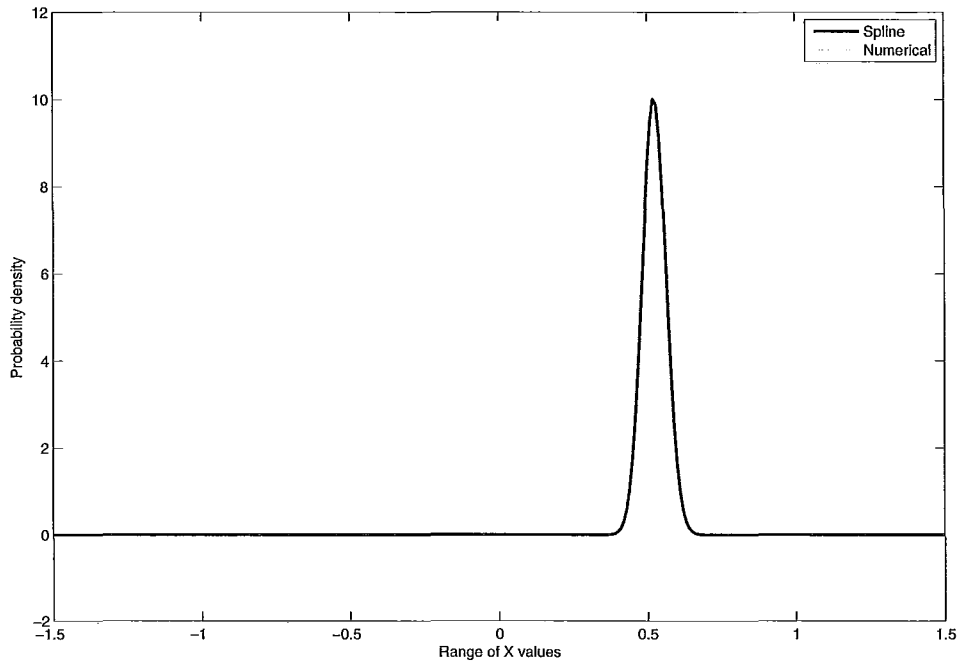


Figure 5.1: The probability density at the final instant of Spline filter and numerical filter

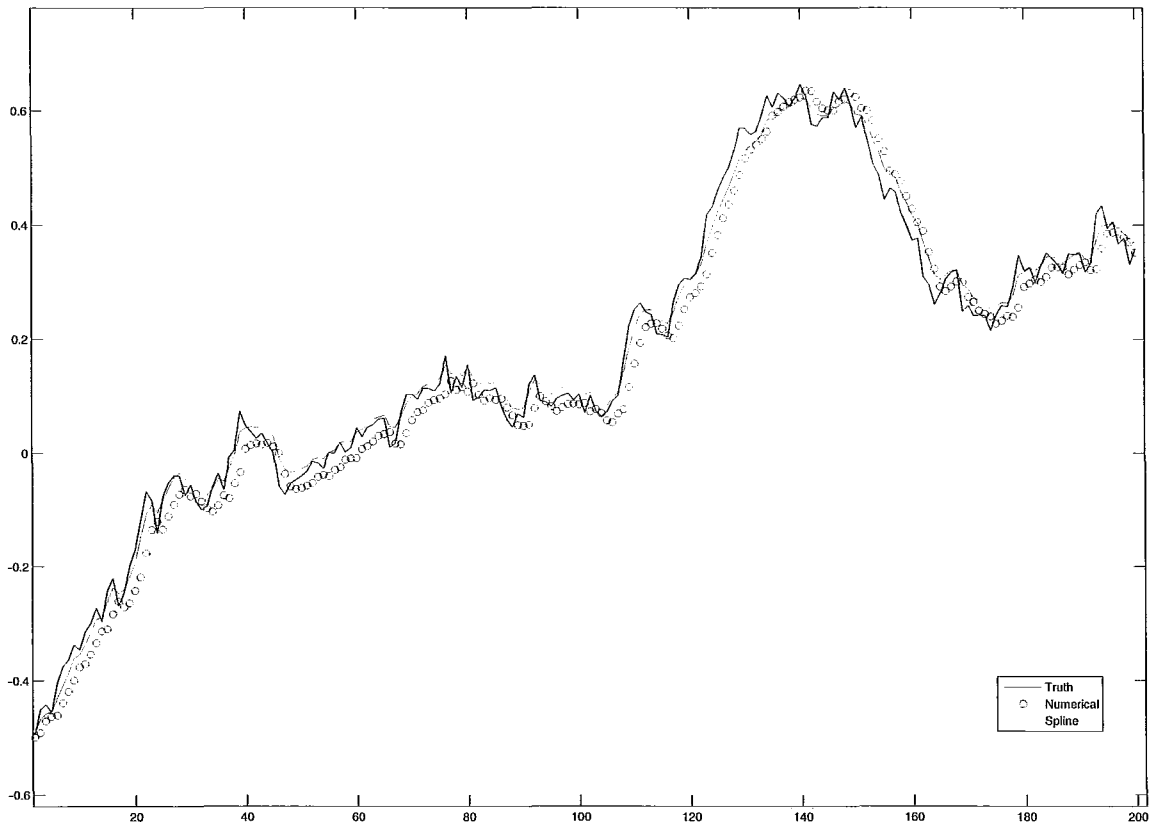


Figure 5.2: Comparing the truth and the calculated results

# Chapter 6

## Summary

### 6.1 Conclusion

The problem being considered in this thesis is the estimation of the state of a nonlinear system based on measurements of the system output. The optimal Bayesian estimator cannot be implemented in practice because the future state of the system can be predicted only by considering the entire domain of the state while implementing the Chapman-Kolmogorov equation. The task of integrating over the continuous state domain is virtually impossible due to huge space requirements. So the problem of this huge integration is handled in this thesis by representing the state pdf using spline polynomials. The initial state density or ‘the prior’ is represented using B splines. The transition probabilities to be used in the Chapman-Kolmogorov equation are calculated from the information available from the system model. Feynman Path integral inspired methods are used for the transition probability calculations. The Chapman Kolmogorov equation was implemented using a spline based method by representing the transition probability matrix using splines. This method simplifies

the representation of the state density through the use of splines. The state density is a continuous function and it is represented as a continuous function itself using splines. This increases the accuracy and reduces the residuals. Also calculation of the moments of the final state density is easy and simple with splines. This new method can be used as an alternative to particle filtering methods. It can be used in various fields of nonlinear estimation such as tracking, navigation, estimation of faults in fields related to power systems and many more. It is a generalized method and can be used for non gaussian filtering since it does not assume that the state density is gaussian. Spline based filtering for target tracking allows easy manipulation of the density functions since they are polynomials of finite order. For instance, it is a straightforward process to find the target probability in any region in the state space by evaluating the integral of the spline representation over that region.

## 6.2 Future work and improvements possible

In this thesis nonlinear state estimation using splines is considered. Currently, the implementation of the Chapman-Kolmogorov equation involves a reconstruction of the spline polynomial representing the state density. Note that this reconstruction issue is not a theoretical limitation but merely an implementation related issue. Matlab® currently does not support multiplication of multivariate splines. Further research can be extended in that area to resolve this issue. Additionally, without further improvements to the theory of the tracking algorithm, there is valid scope for extension to multi-variable nonlinear state estimation. Extension to multi-variable state estimation can be done using multivariate splines, which can be constructed using the tensor product approach. Although the complexity of the tensor product

approach increases exponentially with dimension, complexity can be reduced by exploiting the structure of the linear system that needs to be solved to construct the spline polynomial. The resulting multi-dimensional filter can be used in most of the common circumstances and the density could even be multi-modal.

# Bibliography

- Arasaratnam, I. and Haykin, S. (2009). Cubature Kalman Filters. *IEEE Transactions on Automatic Control*, **54**, 1254–1269.
- Arulampalam, M. S., Maskell, S., Gordon, N., and Clapp, T. (2002). A tutorial on particle filters for online nonlinear/non- gaussian Bayesian tracking . *IEEE Transactions on Signal Processing*, **50**(2), 174–188.
- Balaji, B. (2010a). Feynman Path Integral Inspired Computational Methods for Nonlinear Filtering. *Proceedings of SPIE on Signal Processing, Sensor Fusion and Target Recognition*, **7697**.
- Balaji, B. (2010b). Feynman Path Integrals and Continuous Nonlinear Filtering. *Proceedings of SPIE on Signal Processing, Sensor Fusion and Target Recognition*, **7697**.
- Bar-Shalom, Y., Li, X. R., and Kirubarajan, T. (2001). *Estimation with Application to Tracking and Navigation*. Wiley.
- Boor, C. D. (2001). *A practical guide to Splines*. Springer-verlag, New York.
- Bucy, R. S. and Youssef, H. (1974). Nonlinear filter representation via spline functions. pages 51–60.



- Charalambous, C. D. and Socratous, Y. (2006). Non Linear Estimation for a Class of Systems. *ISIT, Seattle, USA*, 14.
- Daum, F. (2005). Nonlinear filters: Beyond the Kalman filter . *Special tutorial issue of the IEEE Aerospace and Electronic Systems Magazine*.
- Daum, F. and Huang, J. (2003). Curse of Dimensionality and Particle Filters . *Proceedings of IEEE Aerospace Conference, Big Sky Montana*.
- Daum, F. and Huang, J. (2008). Particle flow for nonlinear filters with log-homotopy . *Proceedings of SPIE Conference on signal and data processing, Orlando Florida*.
- Daum, F. and Huang, J. (2009). Nonlinear filters with particle flow induced by log-homotopy. *SPIE Multisensor Fusion, Multitarget Tracking, and resource Management*, **7336**.
- Docenko, D. and Berzins, K. (2003). Spline histogram method for reconstruction of probability density functions of clusters of galaxies. *Lecture Notes on Physics*, **626**, 294–301.
- Gordon, N. J., Salmond, D., and Smith, A. (1993). Novel approach to nonlinear/non-Gaussian Bayesian state estimation . *IEEE Proceedings, Pt. F, Radar and Signal Processing*, **140**(2), 107–113.
- Julier, S. J. and Uhlmann, J. K. (1997). A new extension of the Kalman filter to nonlinear systems.
- Kimeldorf, G. S. and Wahba, G. (1970). A Correspondence between Bayesian Estimation on Stochastic Processes and Smoothing by Splines . *The annals of Mathematical Statistics*, **41**(2), 495–502.

- K.P.Mohandas (2006). *Modern Control Engineering*. Sanguine Technical Publishers, Bangalore, India.
- K.Punithakumar, McDonald, M., and Kirubarajan, T. (2008). Spline Filter for Multidimensional Nonlinear/Non-Gaussian Bayesian Tracking. *Signal and Data Processing of Small Targets 2008, Proceedings of SPIE*, **6969**, 69690K–7.
- Lee, S., Wolberg, G., and Shin, S. Y. (1997). Scattered Data Interpolation with Multilevel B-Spline. *IEEE Transactions on Visualization and Computer Graphics*, **3**(3).
- Lin, X., Wang, N., Welsh, A. H., and Carroll, R. J. (2004). Equivalent Kernels of Smoothing Splines in Nonparametric Regression for Clustered/Longitudinal Data. *Biometrika*, **91**, 177–193.
- Prautzsch, H., Boehm, W., and Paluszny, M. (2002). *Bezier- and B-spline techniques*. Springer-Verlag Berlin Heidelberg New York.
- Punithakumar, K. and Kirubarajan, T. (2007). Spline filter for nonlinear/non-Gaussian Bayesian tracking. *Proceedings of SPIE on Signal and Data Processing of Small Targets*, **6699**.
- Ustuner, K. F. and Ferrari, L. A. (1992). Discrete Splines and Spline Filters. *Ieee Transactions on Circuits and Systems-11: Analog and Digital Signal Processing*, **39**(7), 417–422.
- Wang, A. H. and Klein, R. L. (1976). Implementation of nonlinear estimation using monospline. pages 1305–1307.

Yu, D. and Deng, L. (2009). Solving Nonlinear Estimation Problems using Spline .  
*IEEE Signal Processing Magazine*, **Lecture Notes**, 86–90.

Zhang, J. and Yue, H. (2004). Improved identification algorithm for B-spline modelling of output probability density functions .