A NEW APPROACH TO OPTIMAL NONLINEAR FILTERING

Subhash Challa *Farhan A. Faruqi*

Signal Processing Research Centre, QUT
GPO box 2434, Brisbane, Q. 4001, Australia
s.challa@qut.edu.au@qut.edu.au

ABSTRACT

The classical approach to designing filters for systems where system equations are linear and measurement equations are nonlinear is to linearise measurement equations, and apply an Extended Kalman Filter (EKF). This results in suboptimal, biased, and often divergent filters. Many schemes proposed to improve the performance of the EKF concentrated on better linearisation techniques, iterative techniques and adaptive schemes. The improvements achieved were marginal and often reduced the bias and divergence problems but were far from optimal unbiased estimators. In this paper, we present a new approach to Optimal Nonlinear filtering in linear system - nonlinear measurements case. It is based on approximation of evolved probability density functions using quasi-moments followed by numerical evaluation of Bayes' conditional density equation.

1. INTRODUCTION

Nonlinear filtering problems manifest in three different forms; linear system - nonlinear measurements, nonlinear system - linear measurements, nonlinear system - nonlinear measurements. This paper considers the design of Optimal Nonlinear filters in linear system - nonlinear measurements case. The idea presented in this paper gains motivation from the work done by stratonovich[5] and Culver[2] on the significance of quasi-moments for the approximation of probability density functions and by Challa and Faruqi [3] on the investigation of numerical evaluation of Bayes theorem for nonlinear filtering applications. The conventional approaches using EKF are suboptimal due to the linearisation of measurement equations. The linearisation based methods were popular in literature not only because the complete analytical solution for Bayes' conditional density equation is not possible (except in few special cases) but also because the present days computing power was not available when earlier attempts were made to solve this problem. In view of the tremendous progress taking place in the computer industry, it is deemed timely to undertake numerical approaches where it is difficult to obtain analytical solution. Hence the numerical solution of Bayes' conditional density is suggested in this paper.

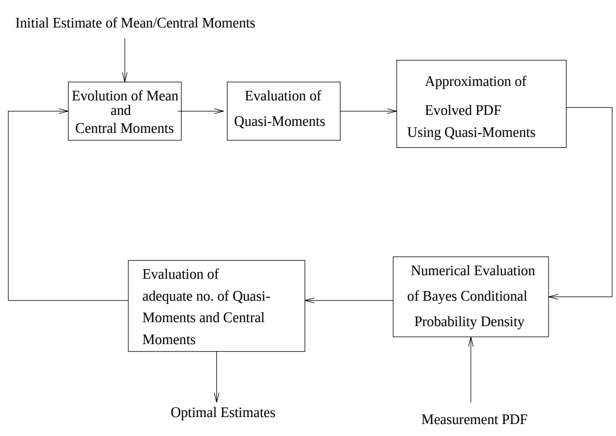

Figure 1. The Block Diagram For The New Method Proposed

In the case of linear system - nonlinear measurements problem, the evolution of central moments are governed by relatively simple equations compared to nonlinear systems but the correction of these moments is extremely difficult. This is because the evolution of lower order moments in linear systems does not depend on higher order moments but the correction of these moments does depend on the higher order moments. The number of higher order moments needed depends on the degree of measurement nonlinearity. This simplification of evolving central moments in linear systems enables one to evolve the probability density function (PDF) using moments as opposed to using Fokker-Planck-Kolmogorov (FPK) forward diffusion equation in nonlinear systems which is extremely difficult to solve. The evolved central moments can be used to evaluate the quasi-moments [5, 4] and find an approximation of the evolved probability density function. The accuracy of this approximation depends on the error bounds tolerated by the particular application in consideration. This method is of great importance in target tracking and navigation problems where the system dynamics, expressed in Cartesian coordinate system, are linear and measurements are nonlinear.

2. NEW FILTERING ALGORITHM

The block diagram in figure 1 summerises the new filtering algorithm. The key point to note is that in any linear system, moment evolution equations of certain order doesn't depend on the moment evolutions of higher order (unlike the moment evolutions in any nonlinear system). This enables one to find a nearly optimal nonlinear filter in a linear system - nonlinear measurements case. The algorithm converges to optimality with the increase in the number of significant quasi-moments considered. This filtering algorithm reduces to a Kalman filtering algorithm in a linear system - linear measurement case as the quasi-moments evaluated would be zero for all orders indicating that the densities remain normal before and after correction.

Consider a continuous time linear stochastic dynamic system(SDS) described by

$$\dot{X}(t) = F(t)X(t) + G(t)\eta(t) \;\; t \geq t_0 \tag{1}$$

where $X(t) = [x_1, x_2, \ldots, x_n]$ represents the state vector of the system at any time t. $F(t)$ is a linear vector valued function with real components and $G(t)$ is a $n \times m$ real matrix and $\eta(t)$ is a white Gaussian noise process, $\eta(t) \sim N(0, Q(t))$. Observations $Y(t_k)$ of this system are taken at discrete time instants t_k :

$$\begin{aligned} Y(t_k) = H(X(t_k), t_k) \quad &+ \quad \nu_k, \;\; k = 1, 2, \ldots \\ t_{k+1} \quad &> \quad t_k \geq t_0 \end{aligned} \tag{2}$$

where $H(X(t_k), t_k)$ is a non-linear function of the observable states of the SDS and $\nu_k \sim N(0, R_k)$. The moment evolution equations are given by lemma 6.1 in [1]. When the equations in the referred lemma are expanded for a linear system they result in matrix Riccati type equations which can be solved recursively. In a single dimension case, the evolution equation for the n^{th} order central moment is derived in the next section.

3. DERIVATION OF MOMENT EVOLUTION EQUATIONS

Defining the n^{th} order central moment by

$$C_n = E\{(x - \hat{x})^n\} \tag{3}$$

where $\hat{x}$ implies $E\{x\}$. By expanding this in accordance with binomial series, we have

$$\begin{aligned} C_n \;\; &= \;\; E\{n_{c_0} x^n + n_{c_1} x^{n-1}\hat{x} + n_{c_2} x^{n-2}(\hat{x})^2 + \ldots \\ &+ \;\; n_{c_r} x^{n-r}(\hat{x})^r + \ldots + n_{c_n}\hat{x}^n\} \end{aligned} \tag{4}$$

which implies that

$$\begin{aligned} C_n \;\; &= \;\; n_{c_0}\widehat{x^n} + n_{c_1}\widehat{x^{n-1}}\hat{x} + n_{c_2}\widehat{x^{n-2}}(\hat{x})^2 + \ldots \\ &+ \;\; n_{c_r}\widehat{x^{n-r}}(\hat{x})^r + \ldots + n_{c_n}\hat{x}^n\} \end{aligned} \tag{5}$$

Where

$$n_{c_r} = \frac{n!}{r!(n-r)!} \tag{6}$$

For a general single dimensional linear system represented by

$$\dot{x} = ax + \eta \tag{7}$$

the moment evolution equations are obtained by differentiating the equation 5 and by using lemma 6.1 from [1] one obtains

$$\frac{d\widehat{x^n}}{dt} = na\widehat{x^n} + \frac{q}{2}(n-1)n\widehat{x^{n-2}} \tag{8}$$

and also by using

$$\begin{aligned} \widehat{x^n} \;\; &= \;\; (\hat{x})^n + n_{c_1}(\hat{x})^{n-1}\widehat{x-\hat{x}} + n_{c_2}(\hat{x})^{n-2}\widehat{(x-\hat{x})^2} + \\ \ldots \;\; & \;\; n_{c_r}(\hat{x})^{n-r}\widehat{(x-\hat{x})^r} + \ldots + n_{c_n}\widehat{(x-\hat{x})^n} \end{aligned} \tag{9}$$

being equivalent to

$$\begin{aligned} \widehat{x^n} \;\; &= \;\; (\hat{x})^n + n_{c_1}(\hat{x})^{n-1}C_1 + n_{c_2}(\hat{x})^{n-2}C_2 + \\ \ldots \;\; & \;\; n_{c_r}(\hat{x})^{n-r}C_r + \ldots + n_{c_n}C_n \end{aligned} \tag{10}$$

in 5 one obtains

$$\begin{aligned} \frac{dC_n}{dt} \;\; &= \;\; naC_n + \frac{q}{2}[n(n-1)\widehat{x^{n-2}} + (n-1)(n-2)\widehat{x^{n-3}} + \\ \ldots \;\; & +(n-r)(n-r-1)\widehat{x^{n-r-2}} + \ldots + 6\hat{x} + 2] \end{aligned} \tag{11}$$

For illustration purposes the evolution equations of first seven central moments are provided.

$$\frac{dC_1}{dt} = aC_1 \tag{12}$$

$$\frac{dC_2}{dt} = 2aC_2 + q \tag{13}$$

$$\frac{dC_3}{dt} = 3aC_3 + q(3C_1 + 1) \tag{14}$$

$$\frac{dC_4}{dt} = 4aC_4 + q(6C_1^2 + 3C_1 + 1) \tag{15}$$

$$\frac{dC_5}{dt} = 5aC_5 + q(10C_1^3 + 30C_1C_2 + 10C_3 + 6C_1^2 + 3C_1 + 1) \tag{16}$$

$$\begin{aligned} \frac{dC_6}{dt} \;\; &= \;\; 6aC_6 + q(15C_1^4 + 60C_1C_3 + 90C_1^2C_2 + 15C_4 \\ &+ \;\; 10C_1^3 + 30C_1C_2 + 10C_3 + 6C_1^2 + 3C_1 + 1) \end{aligned} \tag{17}$$

$$\begin{aligned} \frac{dC_6}{dt} \;\; &= \;\; 6aC_6 + q(21(c_1^5 + 1C_1^2C_3 + 10C_1^3C_2 + 5C_1C_4 + C_5 \\ &+ \;\; 15C_1^4 + 60C_1C_3 + 90C_1^2C_2 + 15C_4 \\ &+ \;\; 10C_1^3 + 30C_1C_2 + 10C_3 + 6C_1^2 + 3C_1 + 1) \end{aligned} \tag{18}$$

The evolved central moments are then used to find the quasi-moments which in turn are used to find the approximate PDF. The evaluation of quasi-moments from central moments and the approximation of evolved density is considered in the next section.

4. APPROXIMATION OF PDF USING QUASI-MOMENTS

The suggested approximation to probability density function is based on the work done by Stratonovich [5, 4]. The interesting thing that is stated in their work is that any probability density can be represented by a Gaussian density times a sum of k_{th} order n dimensional Hermite polynomials. Since we are dealing with single dimension systems we require only k_{th} order single dimensional Hermite polynomials.

The derivation of relation between quasi-moments and central moments involves the expansion of the ratio of characteristic functions in Maclauren's series. The details are given in [2]. The approximation of PDF is simply a product of Gaussian density with same mean and variance as that of the true density with the Hermite polynomials.

$$p(X,t_k|Y_{t_k}^-) = p_g(X,t_k|Y_{t_k}^-)[1+\sum_{k=3}^{\infty}\frac{(-1)^k}{k!}q_{k_n}h_{k_n}] \quad (19)$$

where q_{k_n} is the k_{th} order n dimensional quasi-moment and h_{k_n} is the k_{th} order n dimensional Hermite polynomial. The Hermite polynomials are given by

$$h_k = \frac{(-1)^k}{p_g}\frac{\partial^k p_g}{\partial x^k} \quad (20)$$

where p_g is the Gaussian density having the same mean and variance as that of p the true density. The details of evaluation of these are well elucidated in [2, 5, 4]. The Hermite polynomials upto order seven are given in this paper for illustration purposes. Defining

$$p_g = \frac{1}{\sqrt{2\pi(\sigma)^2}}\exp\frac{-1}{2}\frac{(x-E\{x\})^2}{(\sigma)^2} \quad (21)$$

and

$$T = \frac{(x-E\{x\})}{(\sigma)^2} \quad (22)$$

then we note that

$$\begin{aligned}
\frac{\partial p_g}{\partial x} &= -p_g T && (23)\\
\frac{\partial^2 p_g}{\partial x^2} &= p_g(T^2 - 1\sigma^2) && (24)\\
\frac{\partial^3 p_g}{\partial x^3} &= p_g(-T^3 + \frac{3T}{\sigma^2}) && (25)\\
\frac{\partial^4 p_g}{\partial x^4} &= p_g(T^4 - 6\frac{T^2}{\sigma^2} + \frac{3}{\sigma^4}) && (26)\\
\frac{\partial^5 p_g}{\partial x^5} &= p_g(-T^5 + \frac{10T^3}{\sigma^2} - \frac{15T}{\sigma^4}) && (27)\\
\frac{\partial^6 p_g}{\partial x^6} &= p_g(T^6 - \frac{15T^4}{\sigma^2} + \frac{45T^2}{\sigma^4} - \frac{15}{\sigma^6}) && (28)\\
\frac{\partial^7 p_g}{\partial x^7} &= p_g(-T^7 + \frac{21T^5}{\sigma^2} - \frac{105T^3}{\sigma^4} + \frac{105T}{\sigma^6}) && (29)
\end{aligned}$$

These terms involving the partial derivatives of Gaussian density are useful in getting the Hermite polynomials. It remains to find out the quasi-moments in terms of the central moments. The derivation of these are detailed in [2]. The first seven quasi moments in terms of the first seven central moments are given for illustration purposes.

The first two quasi-moments are always zero. The third quasi-moment is equal to third central moments. The quasi-moments above three are

$$\begin{aligned}
q_4 &= C_4 - 3C_2^2 && (30)\\
q_5 &= C_5 - 10C_2C_3 && (31)\\
q_6 &= C_6 - 15C_2q_4 - 15C_2^3 - 10C_3^2 && (32)\\
q_7 &= C_7 - 21C_2q_5 - 105C_2^2C_3 - 35C_3q_4 && (33)
\end{aligned}$$

We derive here the actual equations that are used while considering upto seventh order quasi- moment, as this was found adequate in the problems considered in this paper. Interested researcher may follow similar steps for obtaining approximations incorporating higher order quasi-moments.

This approximated PDF is used in the corrector in numerical evaluation of the Bayes' conditional density. At an observation (at t_k) the conditional density satisfies the difference equation

$$p(X,t_k|Y_{t_k}) = \frac{p(Y_{t_k}|X)p(X,t_k|Y_{t_k}^-)}{\int p(Y_{t_k}|X)p(X,t_k|Y_{t_k}^-)} \quad (34)$$

respectively, where $p(Y_k|X)$ is given by

$$p(Y_k|X) = \frac{1}{(2\pi)^{\frac{m}{2}}}\times e^{-\frac{1}{2}[Y_k - H(X,t_k)]^T R_k^{-1}[Y_k - H(X,t_k)]} \quad (35)$$

As the functional form of the evolved PDF varies with time, the analytical solution for the Bayes' conditional density equation is extremely difficult to obtain in a recursive estimation scheme. This motivates the use of numerical methods to solve this equation. The conditional PDF obtained is used to obtain the moments and quasi-moments. The quasi-moments of increasing orders are evaluated and the difference between the true PDF and the approximated PDF is compared with an arbitrarily small value ϵ. In the single dimensional case, if the quasi-moments of, say, order n leads to an error less than ϵ, the evaluation of moments of orders higher than n is stopped. The evaluated moments are fed back to the predictor for further prediction. The filter approaches optimality with the increase in the number quasi-moments considered. A typical approximation process is given figures 2, 3 and 4.

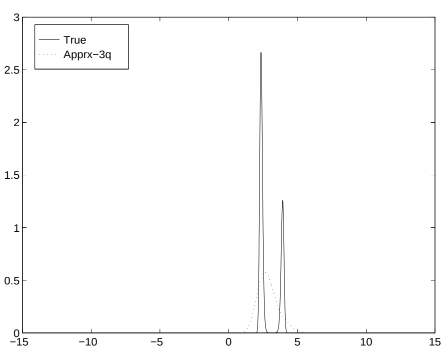

Figure 2. The approximate probability using 3 quasi-moments

The true PDFs and their associated approximations clearly show how the approximate PDFs improve depending on the number of quasi-moments considered. In the course of implementation of this method, some numerical problems were encountered. Some times the approximated density attained negative values. We forced the values to become zero when it was becoming negative and renormalised the density to maintain the area under the density unity.

5. SIMULATION AND RESULTS

The proposed method and EKF are applied on a single dimensional non-linear system with the system dynamics de-

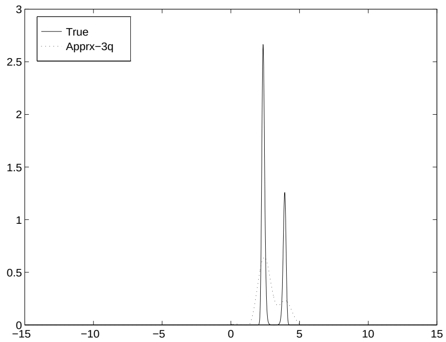

Figure 3. The approximate probability using 5 quasi-moments

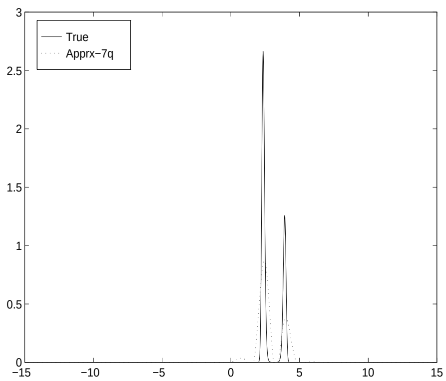

Figure 4. The approximate probability using 7 quasi-moments

scribed by

$$\dot{X}(t) = -0.5X(t) + \eta(t) \tag{36}$$

and the measurements given by $Z(t) = X^3(t) + \nu(t)$ in one simulation and $Z(t) = cos(X(t)) + \nu(t)$ in the other. In

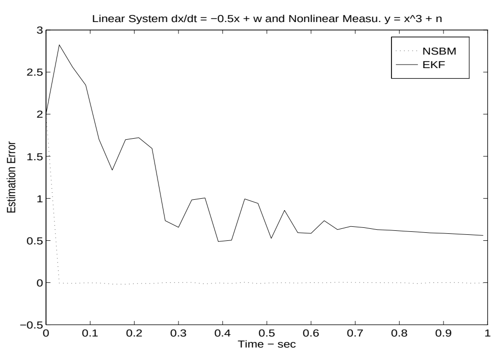

Figure 5. The error plot from simulations with polynomial measurement nonlinearity

the current implementation the system noise variance Q and system measurement variance R are set to 0.001 and 0.005 respectively. The figures 5 and 6 clearly show the superiority of the proposed method over EKF in terms of faster convergence and bias removal in linear system - nonlinear measurements case. It has been observed that as the system noise increased the performance of both these filters become similar. This is due to the fact that the presence of

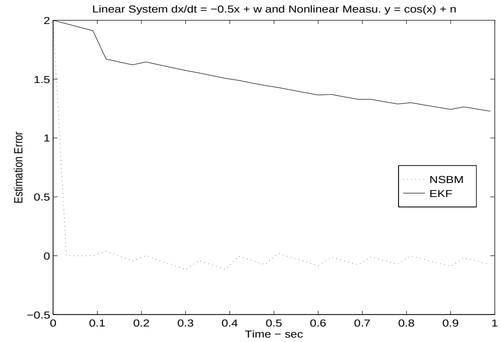

Figure 6. The error plot from simulations with polynomial measurement nonlinearity

system noise occludes the nonlinearities of the system.

6. CONCLUSIONS

A new approach to nonlinear filtering without resorting to linearisation of measurement equations is presented in this paper. The evolution equations of n_{th} order central moments along with their relationship to quasi-moments are presented. The potential of the method is demonstrated on a single dimensional system.

REFERENCES

[1] A. H. Jazwinski. *Stochastic Processes And Filtering Theory*. Academic Press, Inc., 1970.

[2] Culver C. O. *Optimal Estimation For Nonlinear Stochastic Systems, Ph.D Thesis*. MIT, 1969.

[3] Challa S. and Farhan A. F. *Passive Position Location Using Nonlinear Filters*. ISBN 1 86435 249 3 Queensland University Of Technology, 1996.

[4] R. L. Stratonovich. Conditional Markov Processes. *Theory Of Probability And Its Applications*, 5:156–178, 1960.

[5] R. L. Stratonovich. A New representation for Stochastic Integrals And Equations. *Journal Of SIAM Control*, 4:362–371, 1966.