APPLICATION OF CHEBECHEV'S INEQUALITY THEOREM IN THE DESIGN OF OPTIMAL NON-LINEAR FILTERS

Subhash Challa *Farhan A. Faruqi*

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

ABSTRACT

Chebechev's inequality theorem from the theory of probability and statistics provides an upperbound for the amount of probability in the "tails" of any given probability density function. This theorem has interesting applications in the numerical solution of the Fokker-Planck-Kolmogorov Equation (FPKE) as shown in this paper. Numerical solution of FPKE is an essential component of the design of optimal nonlinear filters. The solution of the FPKE in conjunction with the Bayes' conditional density lemma provides optimal (minimum variance) state estimates of any general stochastic dynamic system (SDS).

1. INTRODUCTION

The time evolution of the states of SDS is completely described, mathematically, by the FPKE [3]. The solution of the FPKE gives transition probability densities of the states of the associated SDS. These densities can be conditioned on the measurements by using the Bayes' theorem giving rise to conditional densities. Means of these conditional densities give the optimal state estimates of the SDS [4]. It has been well documented in literature that the analytical solution for the FPKE is extremely difficult to obtain except in a few special cases [5]. The special cases for which analytical solutions exist are for those systems that have a stationary solution or for those systems having PDF's in analytically tractable functional forms. This motivates the use of numerical methods for solving the FPKE. This approach is deemed timely because of the growing interest in the numerical solution of FPKE [6, 7, 2]. The generally used methods use either finite difference methods or finite element methods. We used finite difference methods in solving FPKE (numerically) in the examples considered in this paper. While implementing the finite difference methods the state variable domain is discretised and is defined at finite number of equally spaced grid points. Before we define these points we need to define the maximum limits of the state variable domain. This brings us to the central problem of this paper and is described in the next section.

The solution of FPKE is a probability density function. As the probability densities extends over the entire state variable (or RV) domain ($-\infty$ to $+\infty$ in most cases) some form of truncation of this domain is sought when solving the FPKE numerically. Many *ad hoc* methods exist for such truncation of the state variable (or RV) domain but to the best of the authors knowledge this issue did not receive much attention in literature. This situation, according to our conjecture, arose because much of the research work published in this area is motivated by physical sciences dealing with systems having stationary analytical solutions. In many physical sciences' applications, where transition probability densities of a SDS were required, numerical solutions for the FPKE were used. Once again, as these numerical solutions were sought for systems exhibiting stationary solutions they were, to some extent, free from an inherent problem affecting SDS's having non-stationary solutions. In this paper we not only present, formally, the problem of random variable domain truncation in a recursive estimation scheme, but also provide a statistically consistent and optimal method, based on the Chebechev's inequality theorem, to truncate the the same.

2. THE PROBLEM OF MOVING *SIGNIFICANT* DOMAIN IN OPTIMAL RECURSIVE ESTIMATION

In non-linear filtering problems the dynamical systems seldom have stationary solutions thus the associated FPKE's also rarely have stationary solutions. When an FPKE has a stationary solution the evolving probability densities approach a stationary PDF which "settles" about a certain mean with certain functional form, and as time evolves the significant values of the PDF does not move over its domain. However, in nonlinear filtering problems (e.g. target tracking problems) the significant mass of the evolving PDF moves over the state variable domain. Thus care must be taken while truncating this domain. The domain has to be truncated in such a manner that the evolved PDF's significant mass must not go beyond the maximum limits of the state variable domain. To avoid this problem researchers usually set the domain large enough for the evolved PDF's to stay within the maximum limits for the time of propagation. This has a major disadvantage that a great number of computations are carried out on that part of the domain that has very little contribution to the over all solution of FPKE. Sometimes, to reduce computations, the state variable domain is chosen to be too small resulting in the loss of vital information regarding the underlying PDF. Moreover, the domain corresponding to the significant mass of the PDF changes with time as the PDF evolves in time in accordance with FPKE, thus leading to the problem of moving *significant* domain. Thus the three fundamental problems that exist in the numerical solution of FPKE may be summarised as follows:

- If the maximum limits of the domain of integration for solving FPKE are set too narrow (say Xmin1 and Xmax1 in figure 1)then the evolved PDF may have its significant mass outside these pre-fixed limits.
- If the maximum limits of the domain of integration for solving FPKE are too wide (say Xmin2 and Xmax2 in figure 1) then the algorithm ends up performing a lot of computation on the domain contributing very little to the over all solution of FPKE.
- Even if the maximum limits of the domain are chosen to be optimal for the current time, as the FPKE's solution evolves in time these limits cease to be optimal for all time. Figure 2 provides an illustration of this problem.

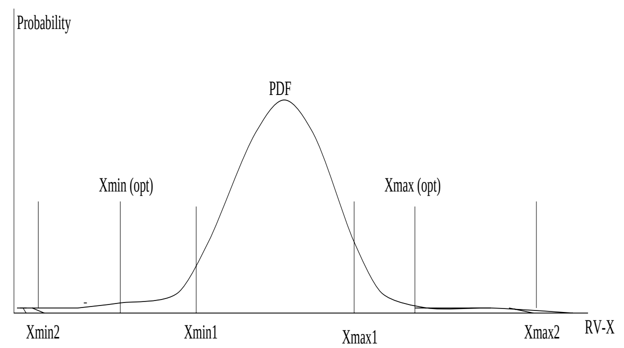

Figure 1. The problem of random variable domain truncation

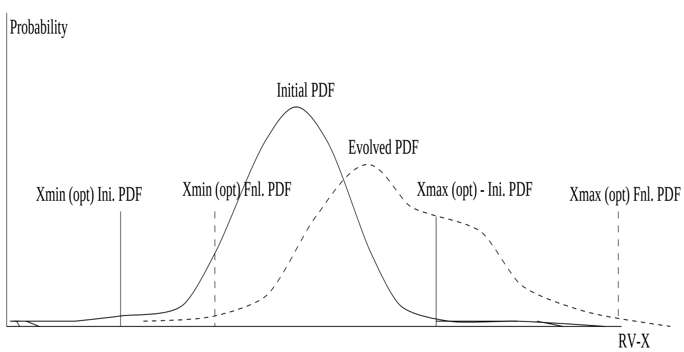

Figure 2. The moving significant domain problem

This necessitates the need to obtain these limits in a recursive manner for each propagation time. In the next section we propose the use of the Chebechev's inequality theorem in conjunction with the moment evolution equations to obtain the limits of the state variable domain in an optimal manner.

3. DETERMINATION OF OPTIMAL SIGNIFICANT DOMAIN IN RECURSIVE ESTIMATION PROBLEMS

In solving the FPKE numerically, with optimal truncation of the state variable domain, one needs to know the limits of both the present and future domains where PDF's significant mass exists. With the knowledge of the present and future second moments these limits can be evaluated using the Chebechev's inequality theorem. Chebechev's inequality theorem from the theory of probability and statistics provides an upperbound for the amount of probability in the "tails" of any given probability density function as a function of its standard deviation. This theorem holds for nearly all types of probability density functions (continuous & discrete). Its application is not limited by any functional form of the PDF's. This generality provides it with the power to deal with the evolving PDF's from the solution of FPKE. The determination of the optimal state variable domain is as shown in the block diagram in figure 3.

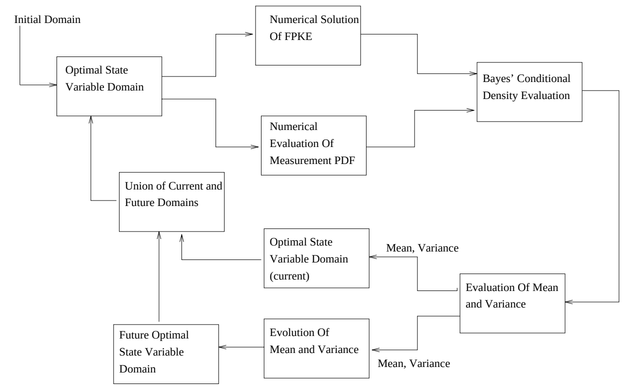

Figure 3. The Block Diagram of the Optimal State Variable Domain Determination

3.1. Chebechev's Inequality Theorem

If x is a random variable having mean μ and variance $\sigma^2 > 0$, then

$$P(|x-\mu| \geq \lambda\sigma) \leq \frac{1}{\lambda^2} \tag{1}$$

where λ is any positive constant. In other words, the probability that a realization of the random variable x lies within the limits $\mu - \lambda\sigma$ and $\mu + \lambda\sigma$ is given by $1 - \frac{1}{\lambda^2}$.

3.2. Application of Chebechev's inequality theorem

In section 2 we presented the moving *significant* domain problem. Here we present a statistically consistent method to obtain an optimal state variable domain. The state variable domain is optimal if the PDF's significant mass lies within that domain. The significant mass of a PDF can be specified as a significance level of the random variable under consideration. For example, if the PDF's significant mass is specified as .99 then the probability that the realization of the random variable will lie within the limits set by the optimal state variable domain is .99. The associated λ is given by $\lambda = \sqrt{\frac{1}{1-0.99}}$. The limits can then be obtained as shown in section 3.1 Since the true PDF is known numerically it is a straight forward matter to evaluate its current second moment but as one doesn't know the actual future PDF, unless one solves the FPKE, an alternate method is suggested for its future second moment thus its future significant domain. This is based on the evolution of second moment using moment evolution methods. We can optimally predict the future second moments for linear SDS's and for certain nonlinear systems in this scheme. Thus in nonlinear filtering problems characterised by linear SDS's and nonlinear measurements and certain nonlinear SDS's and nonlinear measurements, the FPKE can be solved numerically with optimal truncation of state variable domain.

3.3. Moment Evolution Equations

In section 3.2 we have shown that for optimal truncation of the state variable domain we need to know the current PDF's mean and variance as well as the evolved PDF's mean and variance. The later has to be known without finding the actual evolved PDF. A simple way of finding the evolved moments are obtained by applying lemma 6.1 [4]. As we are interested in only first two moments, we can use theorem 6.2 [4] to obtain the evolved mean and variance. Consider a continuous time SDS described by

$$\frac{dX_t}{dt} = F(X_t) + G(X_t)\eta_t \quad t \geq t_0 \tag{2}$$

where $X_t = [x_1, x_2, \ldots, x_n]$ represents the state vector of the system at any time t. $F(X_t)$ is a linear/non-linear vector valued function with real components and $G(X_t)$ is a $n \times m$ real matrix and η_t is a white Gaussian noise process, $\eta_t \sim N(0, Q_t$. Given this SDS the theorem 6.2 [4] states that between measurements the conditional mean ($E[X_t]$) and conditional covariance matrix (P_t) satisfies:

$$\frac{dE[X_t]}{dt} = E[F(X_t)] \tag{3}$$

$$\begin{aligned}\frac{dP_t}{dt} &= (E[X_t F^T(X_t)] - E[X_t]E[F(X_t)]) \\ &+(E[F(X_t)X_t^T] - E[F(X_t)]E[X_t]^T) \\ &+E[G(X_t)QG(X_t)^T]\end{aligned} \tag{4}$$

The $E[(.)]$ refers to the expected value of (.). The expected values appearing in the above equations can be evaluated numerically using the available evolved PDF from the earlier iteration of the FPKE based nonlinear filtering algorithm. This algorithm and the implementation issues are presented in the next section.

The current values of the mean and variance enable us to find the current optimal state-variable domain by the use of Chebechev's inequality theorem. Similarly, by the knowledge of the future moments (from moment evolution equations) we can find the optimal domain for the evolved PDF. Thus the optimal state variable domain for the complete time of PDF evolution, from the current time to the future time, is the union of these two domains.

4. FPKE AND NON-LINEAR FILTER DESIGN

Consider a continuous time SDS presented in section 3.3. Observations $Y(t_k)$ of this system are taken at discrete time instants t_k :

$$\begin{aligned} Y_{t_k} = H(X_{t_k}) \quad &+ \quad \nu_k, \;\; k = 1, 2, \ldots \\ t_{k+1} \quad &> \quad t_k \geq t_0 \end{aligned} \tag{5}$$

where $H(X_{t_k})$ is a linear/non-linear function of the observable states of the SDS and $\nu_k \sim N(0, R_k)$.

Under the assumption that the prior density for the above system exists and is once continuously differentiable with respect to t and twice continuously differentiable with respect to X_t it can be shown that, between observations, the conditional density $p(X_t|Y_t)$ satisfies the FPK forward diffusion equation

$$\frac{\partial p}{\partial t} = -\sum_{i=1}^{n} \frac{\partial [pF_i]}{\partial x_i} + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial [p(GQG^T)_{ij}]}{\partial x_i \partial x_j} \tag{6}$$

where $p(X_t|Y_t)$, $F_i(X_t)$, $G(X_t)$, Q_t are replaced by p, F_i, G, and Q respectively for the sake of simplicity.

At an observation (at t_k) the conditional density satisfies the difference equation

$$p(X_{t_k}|Y_{t_k}) = \frac{p(Y_{t_k}|X)p(X_{t_k}|Y_{t_k}^-)}{\int p(Y_{t_k}|X)p(X_{t_k}|Y_{t_k}^-)} \tag{7}$$

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

$$p(Y_k|X) = \frac{1}{(2\pi)^{\frac{m}{2}}} \times e^{-\frac{1}{2}[Y_k - H(X_{t_k})]^T R_k^{-1}[Y_k - H(X_{t_k})]} \tag{8}$$

where m gives the number of components in measurements vector obtained. The equations (6) and (7) form the predictor and corrector equations of the density evolution method [4]. The means of the conditional density obtained by evaluating equation (7) gives the optimal state estimates.

Since there exists no analytical solution for equation (6), a simple *explicit finite difference method* is used to solve it numerically. In the present implementation the following finite difference approximations are used for the partial derivatives appearing in (1) and the choice of $\triangle t$, h are made in accordance with the condition of stability:

$$\left(\frac{\partial f(X,t)}{\partial t}\right)_{t_i} = \frac{f(X,t_i) - f(X,t_{i-1})}{\triangle t} \tag{9}$$
$$\text{where} \quad \triangle t = t_i - t_{i-1} \quad i = 1, 2, \ldots$$

$$\left(\frac{\partial f(X,t)}{\partial X}\right)_{t_{i-1}} = \frac{f(X_{k+1},t_{i-1}) - f(X_{k-1},t_{i-1})}{2h} \tag{10}$$

$$\left(\frac{\partial^2 f(X,t)}{\partial X^2}\right)_{t_{i-1}} = \frac{f(X_{k+1},t_{i-1}) - 2f(X_k,t_{i-1}) + f(X_{k-1},t_{i-1})}{4h^2} \tag{11}$$
$$\text{where} \quad h = x_k - x_{k-1} \quad k = 1, 2, \ldots$$

Finite-difference methods are approximate in the sense that derivatives at a point are approximated by difference quotients over a small interval but the solutions are not approximate in the sense of being crude estimates.

The condition for stability of the *explicit finite difference method* is that the grid spacing in the time domain, $\triangle t$, and grid spacing in the state variable domain, h, must satisfy the following inequality [1]:

$$0 < \frac{\triangle t}{h^2} \leq \frac{1}{2}$$

The solution obtained by solving equation (6) using the approximations defined in equations (9),(10),and (11) is used in Bayes' conditional lemma given by equation (7) to obtain the conditional density of the evolved PDF conditioned on the measurements obtained through equation (5).

Since the solution was obtained by solving equation (6) numerically we solved the equation (7) also numerically. In this paper all the examples considered used trapezoidal integration in evaluating the integral present in the denominator of equation (7).

Although other methods like *implicit finite difference methods* and *finite element methods* exist, we used the *explicit finite difference* for the ease of implementation and for quick demonstration of the application of the Chebechev's inequality theorem. We emphasise the fact that the Chebechev's inequality theorem is not limited to any particular numerical method used for solving the FPKE and that it can be used with any numerical method where truncation of the RV's domain is required.

5. SIMULATION RESULTS

In order to illustrate the effectiveness of the application of the Chebechev's inequality theorem in the design of optimal nonlinear filters using the FPKE, we implemented the optimal filters with and without the optimal truncation of the state variable domain on a single dimensional nonlinear filtering problems. We take the example of a linear SDS and nonlinear measurements filtering problem. The initial values of the state variable was chosen as 3 while the initial information given to the filters were in the form of a Gaussian density with mean 2 and variance 0.16. The system noise variance was chosen to be 0.01 and the measurement noise variance was chosen to be 0.2. The integration step length (or equivalently discretisation step in time domain) was chosen as 0.001 sec and the continuous state variable domain was approximated by a set of grid points separated by 0.08 units. When the density evolution method is applied to obtain optimal estimates for a system with following system dynamics,

$$\dot{X}(t) = -1.5X + \eta_t \tag{12}$$

The measurements are nonlinear and are given by

$$Z(t) = X^3(t) + \nu_t \tag{13}$$

where η_t and ν_t are the system and measurement noises respectively and $X(t)$ represents system state at time t.

It is clearly evident from the figure above that the error between the implementations of of FPKE with and without optimal truncation of the state variable domain are negligible while the computations have come down drastically. The FPKE implementation without optimal truncation was working on 1500 points on the state variable domain while the FPKE implementation with optimal truncation was working on progressively lesser number of points.

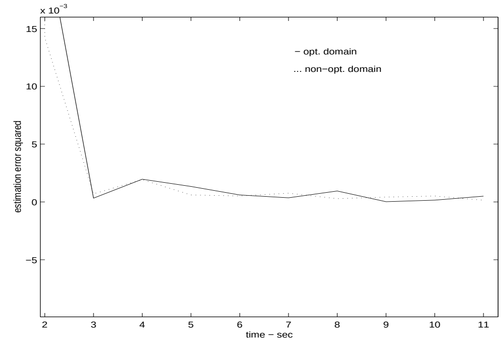

Figure 4. The estimation accuracy with and without optimal truncation of the state variable domain

The number of points stabilised after reaching 129 which is significantly less than 1500 points. The λ was chosen to be 5.

6. CONCLUSIONS

An interesting and extremely useful application of the Chebechev's inequality theorem in the design of the optimal nonlinear filters is presented. It is demonstrated that the proposed method, while reducing the computations involved considerably, does not sacrifice the estimation accuracy. The proposed method using the Chebechev's inequality theorem in arriving at the optimal limits of the state variable domain is extremely powerful because it is not only independent of the numerical methods used in obtaining the solution of the FPKE, but also independent of the form of the PDF's evolving out of the FPKE. Finally, it is shown that in a recursive estimation scheme the computations involved will progressively decrease as the variance of the estimated state variables decrease.

REFERENCES

[1] William F. Ames. *Numerical methods for partial differential equations.* Academic Press, Boston, 1992.

[2] Mohamed A. EL-Gelbeily and Hosam E. Emara Shabaik. Approximate Solution Of The Fokker-Planck-Kolmogorov Equation By Finite Elements. *Communications In Numerical Methods In Engineering*, 10:763–771, 1994.

[3] G.W. Gardiner. *Handbook Of Stochastic Methods.* Springer-Verlag, Berlin, 1990.

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

[5] H. Risken. *The Fokker-Planck Equation.* Springer-Verlag, Berlin, 1989.

[6] V.K. Saul'ev and A.A. Chernikov. Solution Of The Fokker-Planck-Kolmogorov Equation By The Method Of Finite Differences. *Avtomatika i Telemekhanika*, 3:98–102, 1990.

[7] V. Vanaja. Numerical Solution Of A Simple Fokker-Planck Equation . *Applied Numerical Mathematics*, 9:533–540, 1992.