Vol. 5 (Iss. 4) 2024, pp. 302-316, African Journal of Empirical Research https://ajernet.net

Second Order Extended Ensemble Filter for Non-linear Filtering

Kevin Midenyo[1]
Angwenyi David[2]
Oganga Duncan [3]
[1,2,3] Department of Mathematics, Masinde Muliro University of Science and Technology,
P.O Box 190-50100, Kakamega (Kenya).
*kevinmidenyo@gmail.com

Abstract

Whenever the state of a system must be estimated from noisy information, a state estimator is employed to fuse the data with the model to produce an accurate estimate of the state. When the system dynamics and observation models are linear, the Kalman Filter which is optimal, is used. However, in most applications of interest the system dynamics and observations equations are not-linear and suitable extensions of the Kalman Filter have been developed; for example, the Extended Kalman Filter(EKF). The Extended Kalman Filter is based on linearization by the Taylor series expansion about the mean of the state. This filtering process is however computationally expensive especially in high dimensional data. The cause for this is the high cost of integrating the equation of evolution of covariances. Due to this complexity in integration, new methods were sought known as the particle filters. It replaces linearization of non-linearities with Monte Carlo methods. The particle filter formed a basis for Ensemble Kalman Filter (EnKF) an extension of Kalman filter to non-linear filtering. The EnKF reduced the computational cost but its innovation process does not capture information sufficiently hence there is need to improve its performance. This study has developed a new filter, Second order Extended Ensemble Filter (SoEEF).We derived it from stochastic state models by expansion of expected values to the second order by use of Taylor series together with Monte Carlo method and Matlab. We used Lorenz 63 system of ordinary equations and differential Matlab to test the performance of the new filter. Then we compared its performance with four other filters like Bootstrap Particle Filter (BPF),First order Kalman Bucy Filter (FoEKBF),Second order Kalman Bucy Filter (SoKBF) and First order Extended Ensemble Filter (FoEEF). SoEEKF performs much better than the other four filters.

Keywords: Non-linear Filtering; Kalman Filter; Estimate; non-linear; second order extended ensemble Kalman filter
2020 Mathematics Subject Classification: 35B45; 60J27; 93E11

1 Introduction

1.1 Non-linear filtering

The goal of non-linear filtering is to find an estimate of the current state of a given system characterised by stochastic process X. This process is called the signal process [6] [2] [11]. The signal process is measured via a process Y called

Measurement process. These processes can be represented by stochastic differential equations [9]. Non-linear filtering has made extensive use of stochastic calculus, hence we shall start off by reviewing stochastic calculus in time continuous. Consider stochatic differential equation of the form

$$dx_t = f(x_t, t)dt + g(x_t, t)\beta_t; \qquad t_0 \leq t \tag{1}$$

where x_t is $n \times 1$ state vector at time t, $f(x, t)$ is the drift function of dimension $n \times 1$, the uncertainties are captured by $g(x_t, t)$ called diffusion function of dimension $n \times m$ then $\beta_t, t \leq t_0$ is the Brownian motion process [1]. The solution of (1) can be obtained by integrating both sides as follows;

$$x_t = x_{t_0} + \int_{t_0}^{t} f(x_t, \tau)d\tau + \int_{t_0}^{t} g(x_\tau, \tau)d\beta_\tau \tag{2}$$

where $\int_{t_0}^{t} f(x_t, \tau)d\tau$ is the Riemann integral. Now the second integral

$$I(t) = \int_{t_0}^{t} g(x_\tau, \tau)d\beta_\tau \tag{3}$$

is riddled with technicalities which arise from the fact that Brownian motion process is nowhere differentiable. Such integral needs special treatment, due to stochasticity, which is why they are known as stochastic integrals. This stochastic integral can be defined as L^2 -limit of Riemann sum approximation. We set

$$w_k = (1 - \lambda)t_k + \lambda t_{k+1}. \tag{4}$$

We have two choices where $\lambda = 0$ we obtain the Itô stochastic integral

$$\int_a^b g(x_\tau, \tau)d\beta_\tau = \lim_{\delta t \to 0} \sum_{k=1}^{n} g(x_{t_k}, t_k)(\beta_{x_{k+1}} - \beta_{t_k}). \tag{5}$$

Where $\delta t = t_{k+1} - t_k$ and the terms to be summed are evaluated at discretised times $a = t_0 < t_1 < ... < t_k < t_{k+1} < ... < t_n = b$, l.i.m is limit in the mean convergence. The second choice is $\lambda = \frac{1}{2}$ that leads us to Stratonovich stochastic integral

$$\int_a^b g(x_\tau, \tau)d\beta_\tau = \lim \sum_{k=1}^{n} g\{\frac{x_{t_k} + x_{t_{k+1}}}{2}, t_k\}(\beta_{t_{k+1}} - \beta_{t_k}). \tag{6}$$

When the Itô integral is independent of λ then the Itô and Stratonovich integrals coincide [12] , that is

$$\int_a^b g(x_t, t)' d\beta_\tau = \int_a^b g(x_t, t)d\beta_\tau + \frac{1}{2}\int_a^b \partial_x[g(x_\tau, \tau)]d\tau. \tag{7}$$

1.2 Filtering Problem

The filtering problem is to find the best estimate of the state X given the measurement Y. To best solve this problem, the conditional density $\pi_t(x|y)$ of x is considered [2] [11] given the measurement y. The continuous time solution to this problem in special case is given by Fokker-Planck equation while in general case is given by Kushner-Stratonovich equation. We consider the Fokker-Planck equation . Given the density function

$$\pi_t(x) = \pi(x, t|Y, \tau) \tag{8}$$

and its probability distribution function, $\pi_{t|\tau}(x|y)$ the Fokker Plank equation is therefore,

$$\frac{\partial \pi_{t|\tau}(x|y)}{\partial t} = -\frac{\partial(\pi_{t|\tau}(x|y))f(x, t)}{\partial x} + \frac{1}{2}\frac{\partial^2(\pi_{t|\tau})(x|z)g^2(x, t)}{\partial x^2}. \tag{9}$$

This equation describes the evolution of the density function π_x.

1.2.1 Kushner-Stratonovich eqaution

The solution to the problem of non-linear filtering in continuous time is given by the Kushner-Stratonovich equation [20]. The equation estimates $\hat{x}_t$ at time t by combining noisy dynamics with noisy measurement and is given as follows.

$$\pi_t(x/z_t) = \pi_{t_0}(x) + \int_{t_0}^{t} \ell(\pi_t(x/z_t))d_\eta + \int_{t_0}^{t} \pi_\eta(x/z_\tau)(h - \hat{h}_\tau)^T R^{-1}(\eta)(dz_\eta - \hat{h}_\eta d_\eta). \tag{10}$$

In differential form we have

$$d\pi_t(x/z_t) = \ell(\pi_t(x/z_t))d\eta + \pi_\eta(x/z_\eta)(h - \hat{h}_\eta)^T R^{-1}(\eta)(dz_\eta - \hat{h}_\eta d\eta) \tag{11}$$

where

$$\ell = f\frac{\partial}{\partial x} + \frac{1}{2}g^2\frac{\partial^2}{\partial x^2}$$

and

$$\hat{h_t} = \int h(x,t)\pi_t(x/z_t)dx\ .$$

The weak form of Kushner-Stratanovich equation is

$$d\pi_t[\phi] = \pi_t[\ell\phi]dt + (\pi_t[\phi(x)h] - m\hat{\phi}\hat{h}_t)^T R^{-1}(t)(dz_t - \hat{h}_t dt)]. \tag{12}$$

We now consider the equation of evolution of the mean and covariance. For a function of x given $\phi(x)$ and conditional density $\pi_t(x/z_t)$ the mean is

$$\pi_t[\phi(x)] = \int \phi(x)\pi_t(x/z_t)dx \tag{13}$$

using the weak form of Kushner-Stratonovich equation(1.12) and taking $\pi_t[\phi(x)] = \hat{\phi}_t(x)$ we get

$$d\hat{\phi}_t(x) = \pi_t[f\frac{\delta\phi}{\delta x}]dt + \frac{1}{2}\pi_t[g^2\frac{\delta^2 x}{\delta x^2}]dt + (\pi_t[\phi h] - \hat{\phi}_t\hat{h}_t)(R)^{-1}(t)(dz_t - \hat{h}_t dt). \tag{14}$$

When we substitute $\phi(x)$ with x, our equation becomes

$$d\hat{x}_t = \hat{f}_t dt + (\pi_t[xh] - \hat{x}\hat{x}_t)R^{-1}(t)(dz_t - \hat{h}_t dt). \tag{15}$$

This is the equation for evolution of the mean.
The conditional variance is derived from

$$p_t = E[(x - \hat{x}_t)]^2|z_t$$

$$= \pi_t[x^2] - \hat{x}_t^2$$

hence

$$dp_t = d\pi[x^2] - d\hat{x}_t. \tag{16}$$

It is now appropriate that we compute an expression $d\pi_t[x^2]$ and $d\hat{x}_t^2$.
Substituting $\phi(x)$ for x^2 in the weak Kushner-Stratonovich equation, we get

$$d\pi_t[x^2] = 2\pi_t[xf]dt + \pi_t[g^2]dt + (\pi_t[x^2h] - \pi_t[x^2]\hat{h}_t)R^{-1}(t)(dz_t - \hat{h}_t dt). \tag{17}$$

Applying Itô formula of measurements

$$dh = \delta_t[h]dt + \nabla[h]^T dx_t + \frac{1}{2}tr[g(x_t,t)g^T(x_t,t)]\nabla[h]dt$$

the equation becomes

$$d\hat{x}_t^2 = 2\hat{x}_t d\hat{x}_t + (\pi[xh] - \hat{x}_t\hat{h_t})^2 R^{-1}(t)dt. \tag{18}$$

Substituting from (16), we get

$$d\hat{x}_t^2 = 2\hat{x}_t + (2\hat{x}_t\pi_t[xh] - 2\hat{x}^2\hat{h}_t)R^{-1}(t)(dy_t - \hat{h}_t dt) + (\pi_t[xh] - \hat{x}_t\hat{h}_t)^2 R^{-1}(t)dt. \tag{19}$$

The evolution of covariance is developed from(7), substituting(17) and (19) in equation 16 we have

$$\begin{aligned} dp_t = (2\pi_t[xf] - 2\hat{x}_t\hat{f}_t)dt + \pi_t(g^2dt - (\pi_t[xh] - \hat{x}_t\hat{h}_t))^2 R^{-1}(t)dt \\ +(\pi_t[x^2h] - 2\hat{x}_t[xh] - \pi_t[x^2]\hat{h}_t + 2\hat{x}_t^2\hat{h}_t)R^{-1}(t)(dy_t - \hat{h}_t dt). \end{aligned} \tag{20}$$

2 Existing filters

We shall look at continuous time filters where there is no predictor- corrector formulation [15]. We begin by looking at evolution of the mean and covariance in linear case then proceed to non-linear filters. Consider the linear model equation :

$$Signal: \quad dx_t = F(t)x_t dt + G(t)d\beta_t; \quad x_{t_0}, \qquad t_0 \le t \tag{21}$$

$$Measurement: \quad dy_t = H(t)x_t dt + R^{\frac{1}{2}}(t)d\eta_t, \qquad t_0 \le t \tag{22}$$

where $F(t)x_t$ is a $n \times n$ continuous time matrix and $G(t)$ is an $n \times m$ continuous time matrix ,$\beta_t, t_0 \le t$ is an m Brownian motion. $H(t)x_t$ is $r \times n$ continuous time matrix,$R(t)$ is an $r \times r$ matrix ,$\eta_t, t_o \le t$ is r dimensional Brownian motion vector. The Kushner-Stratonovich equation (11) can be used to derive the Kalman Bucy filter. Multipying equation(11) by x with and integrating over it we obtain $\hat{x}$

$$d\hat{x} = F(t)\hat{x}_t dt + P_t H^T(t)(dy_t - H(t)\hat{x}_t dt). \tag{23}$$

In the same manner the covariance equation is as follows.

$$dP_t = F(t)P_t dt + P_t F^T(t)dt + G(t)G^T(t)dt - P_t H^T(t)R^{-1}(t)H(t)P_t dt. \tag{24}$$

Equation (3) and (4) are the continuous time linear filtering equations known as Kalman Bucy Filter (KBF)[14]. It gives an optimal estimate of (x) given Y. The Prediction step is given by: The mean estimate $d\hat{x}_t = F(t)\hat{x}_t dt$
and Covariance estimator, $dP_t = F(t)P_t dt + P_t F^T(t)dt + G(t)G^T(t)dt$.
The Additive term Summarises the new measurement in the predicted estimate.
$P_t H^T(t)(dy_t - H(t)\hat{x}_t dt$ the analysis step is made up of the Kalman gain $k_t = P_t H^T(t)R^{-1}(t)$
and the Innovation $dy_t - H(t)\hat{x}_t dt$. Through these equations one can compute an estimate of x [3].

The Kalman Bucy Filter is optimal for linear systems [1], but it is not optimal in non -linear systems. Hence extensions of KBF to non-linear systems have been sought for example EnKF and Extended Kalman Bucy Filter.

2.1 Bootstrap Particle Filter

Particle filters solve filtering problem by using Monte-Carlo methods to numerically approximate estimates of the mean and covariance. Particle Filter approximates the posterior distribution $p(x_t|y_{1:t})$,with a set of particles $N_t = \left(X_t^i, w_t^i\right)$, $i = 1, ..., N$ [17] where x_t^i is the dynamic state. w_t^i is the weight of the particle in time t $y_{1:t}$ measurement

$$\hat{P}(x_t|y_{1:t}) = \sum_{i=1}^{N} w_t^i \delta(x_t - x_t^i). \tag{25}$$

The most applied Particle filter is the Bootstrap particle filter[8]. This is due to their simplicity. It is an applied filtering algorithm which approximate the filtering probability density $p(x_k|y_o....,y_k)$ by a weighted set of M sample (w_k^i, x_k^i) $i\epsilon 1, ..., M$. The importance weights w_k^i are approximations to the relative probabilities. The Bootstrap Particle filter algorithm can be summarised as follows.

Algorithm 1 Bootstrap particle filter

Input: δt,α,M,N,π_{t_0},π_{t_n} and x_0
Output: $\{x_{t_n}^i\}_{n,i=1}^{N,M}$

Obtain $x_{t_0}^i \sim \pi_{t_0}\left(x_{t_0}^i|x_{t_{n-1}}^i\right)$
Assing initial weights $w_{t_0}^i = \frac{1}{M}$
for $n = 1$ **to** $N, \delta t > 0$ **do**
    Obtain $x_{t_n}^i \sim \pi_{t_n}\left(x_{t_n}^i|x_{t_{n-1}}^i\right)$
    Calculate weights $w_{t_n}^i \sim \pi_{t_n}\left(\delta y_{t_n}|x_{t_n}^i\right) w_{t_{n-1}}^i$
    Normalise the weights to obtain $\{\tilde{w}_{t_n}^i\}_{i=1}^M$
    Calculate $ESS_{t_n} = \frac{1}{\sum_{i=1}^M(\tilde{w}_{t_n}^i)^2}$
    if $ESS_{t_n} \leqslant \alpha$ **then**
        Re-sample the particles
        Set the weights to $w_{t_n}^i = \frac{1}{M}$
    end if
end for

2.2 Ensemble Kalman-Bucy Filter

The Ensemble Kalman Bucy Filter is based on particle filters. EnKBF is designed for time continuous models. Estimates the covariance using Monte Carlo techniques. It draws M ensembles members $\left(x_{t_0}^i\right)_{i=1}^M$ which are used to approximate the conditional density [10] . The x^i is propagated according to the equation of evolution of the mean in KBF as follows[1]

$$dx_t^i = F(t)x_t^i dt + G(t)d\beta_t^i + P_t^M H^T(t)R^{-1}(t)(dy_t + R^{\frac{1}{2}}(t)\eta_t^i - H(t)x_t^i dt). \tag{26}$$

The ensemble mean is given by

$$\hat{x_t} = \frac{1}{M}\sum_{i=1}^M x_t^i. \tag{27}$$

The covariance is estimated as

$$P_t^M = \frac{1}{M-1}\sum_{i=1}^M (x_t^i - \hat{x}_t)(x_t^i - \hat{x}_t)^T. \tag{28}$$

2.3 First Order Extended Kalman Bucy Filter (FoEKBF)

The scalar form of 1 and 2 gives us

$$Signal: \quad dx_t = f(x_t)dt + g(t)(t)d\beta_t; \quad x_{t_0}, \qquad t_0 \leq t \tag{29}$$

$$Measurement: \quad dy_t = h(x_t)dt + R^{\frac{1}{2}}(t)d\eta_t, \qquad t_0 \leq t \tag{30}$$

By substituting $g(x,t)$ with $g(t)$ with corresponding derivatives. The first order approximations of equations 16 and 20 in Taylor series expansion about the mean [13] yields,

$$d\hat{x}_t = f(\hat{x}_t)dt + P_t\nabla[h](\hat{x}_t)R^{-1}(t)(dy_t - h(\hat{x}_t)dt) \tag{31}$$

$$dP_t = P_t\nabla[f]^T(\hat{x_t}dt + \nabla[f](\hat{x}_t)P_tdt + g(t)g^T(t)dt - P_t\nabla[h]^T(\hat{x})R^{-1}\nabla[h](\hat{x}_t)P_tdt. \tag{32}$$

Where $\nabla[f] = \frac{\delta f_i}{\delta x_j}$ $\nabla[h] = \frac{\delta h_i}{\delta x_j}$

2.4 Second Order Extended Kalman Bucy Filter(SoEKBF)

In similar manner of finding FoEEKF we apply Taylor series expansion of the mean estimate to second order[13] . This gives us SoEKBF,

$$d\hat{x}_t = f(\hat{x}_tdt + \frac{1}{2}\triangle[f](\hat{x}_tP_tdt + P_t\nabla[h]^T(\hat{x}_t)R^{-1}(t)(dy_t - (h(\hat{x}_t) + \frac{1}{2}\triangle[h](\hat{x}_tP_t)dt) \tag{33}$$

$$\begin{aligned} dP_t = P_t\nabla[f]^T(\hat{x}_t)dt + \nabla[f](\hat{x}_t)P_tdt + g(t)g^T(t)dt + -P_t\nabla[h]^T(\hat{x})R^{-1}\nabla[h](\hat{x}_t)P_tdt \\ +\frac{1}{2}P_t\triangle[h]^T(\hat{x}_t)R^{-1}(t)\big(dy_t - (h(\hat{x}_t) + \frac{1}{2}\triangle[h](\hat{x}_tP_t)dt\big)P_t \end{aligned} \tag{34}$$

where $\triangle[f] = \frac{\delta^2 f}{\delta x_i \delta x_j}$ and $\triangle[h] = \frac{\delta^2 h}{\delta x_i}\delta x_j$

3 Derivation of the filter

We shall make a few assumptions as we develop our filter. We strict ourselves to linear and non-linear models with an assumption that probability densities and conditional probability densities are Gaussian. We shall make approximations based on Taylor series expansion and Monte-Carlo methods. Approximations are necessary due to the nature of high dimensional data [18]. We therefore review the derivation of the second and first order direct approximations to the equation of evolution of the mean and covariance using Taylor series expansion about the mean [[13];[4]]. What is more, we derive sequential Monte-Carlo approximation to the solution to the non-linear filtering problem. The imperfection of the models can be dealt with by quantifying the uncertainty. The uncertainties in the model are mostly represented as random variables, and are referred to as noise in the model. The setting is that of stochastic differential equations.
Suppose it is required of us to obtain an estimate of the state, x_t, given the measurements history, $Y_t; t \leq t_0$. This calls for the use of conditional density functions. Happily, these functions have been established,together with their evolution of the first and second moments.
Consider the following continuous state space models that can be expressed, in $It\hat{o}$ form, as follows.

$$Signal: dx_t = f(x_t,\theta)dt + G(x_t)Q^{\frac{1}{2}}(t)d\beta_t; \qquad x_{t_0} = x(0), \qquad t_0 \leq t \tag{35}$$

$$Measurement: dy_t = h(x_t)dt + R^{\frac{1}{2}}(t)d\eta_t, \qquad t_0 \leq t \tag{36}$$

where: .

The state vector x_t, at a given time t, evolves in time according to a function f, of the state vector t. The uncertainties and imperfections of the model are as much as possible captured in additive term, $G(x_t)Q^{\frac{1}{2}}(t)d\beta_t$, where $\beta_t, t \leq t_0$ is Brownian motion process. The model equation (3.1) is an expression of a non-linear continuous state space model. Given

Term	Name	Dimension
x_t	State vector	$n \times 1$
$f(x_t, \theta)$	drift term	$n \times 1$
θ	vector of parameters	$d \times 1$
$G(x_t)$	matrix function	$n \times m$
$Q(t)$	time - function matrix	$m \times m$
$\beta, t > t_0$	Brownian motion process	$m \times 1$
y_t	output vector	$r \times 1$
$h(x_t)$	measurement function	$r \times 1$
$R(t)$	time- function matrix	$r \times r$
$\eta_t, t > t_0$	Brownian motion process	$r \times 1$

the initial value of the state x_{t0}, the subsequent states in the model can be obtained by formal integration. This involves solving the model each time. The state at a given time depends on the state at the previous time and nothing else about the past states, this model is hence known as the Markov model.

The model is said to be deterministic if it does not include randomness in any of it's terms;otherwise it is said to be Stochastic.

3.1 The mean and covariance

Mean and covariance correspond to the first and second moments, respectively [13]. A moment, by definition, is the expectation of a function, say, $\varphi(x)$, with respect to a given probability density. For the system (35) to (36), the mean, $\hat{x}_t$, and the covariance, P_t satisfy the following equations.

$$d\hat{x}_t = \hat{f}dt + (\widehat{x_t h^T} - \hat{x}_t \hat{h}^T)R^{-1}(t)(dy_t - \hat{h}dt), \qquad \hat{x}_{t_0} = x(0), \tag{37}$$

$$\begin{aligned}(dP_t)_{ij} &= (\widehat{x_i f_i} - \hat{x}_i \hat{f}_i)dt + (\widehat{f_i x_j} - \hat{f}_i \hat{x}_j)dt + (\widehat{GQG^T})_{ij}dt - (\widehat{x_i h} - \hat{x}_i \hat{h})^T R^{-1}(\widehat{h x_j} - \hat{h}\hat{x}_j)dt \\ &+ (\widehat{x_i x_j h} - \widehat{x_i x_j}\hat{h} - \hat{x}_i \widehat{x_j h} - \hat{x}_j \widehat{x_i h} + 2\hat{x}_i \hat{x}_j \hat{h})^T R^{-1}(t)(dy_t - \hat{h}dt), \qquad P_{t0} = P(0)\end{aligned} \tag{38}$$

The equations 37 and 38 are the exact equations of the evolution of the mean and covariance[13] .

For clarity of exposition, we use the scalar case of the system (3.1) - (3.2) - which we now write as follows.

$$Signal : dx_t = f(x_{t,\theta})dt + g(x_t)q^{\frac{1}{2}}d\beta_t; x_{t0} = x(0), \qquad t_0 \leq t, \tag{39}$$

$$Measurement : dy_t = h(x_t)dt + r^{\frac{1}{2}}d\eta_t, \qquad t_0 \leq t. \tag{40}$$

The corresponding equations for evolution of the conditional mean and variance, for the scalar case, are :

$$d\hat{x}_t = \hat{f}dt + (\widehat{x_t h} - \hat{x}_t \hat{h})r^{-1}(t)(dy_t - \hat{h}dt), \hat{x}_{to} = x(0), \tag{41}$$

$$(dp_t) = 2(\widehat{xf} - \hat{x}\hat{f})dt + (\widehat{qg^2}dt - (\widehat{xh} - \hat{x}\hat{h})^2 r^{-1} + (\widehat{x^2 h} - \widehat{x^2}\hat{h} - 2\hat{x}\widehat{xh} + 2\hat{x}^2\hat{h})r^{-1}(dy^t - \hat{h}dt), 5.5cmdp_0 = p(0), \tag{42}$$

where $\hat{f} = \int f(x)\rho(x|Y_t)dx$.
Solving the system (39) - (40) gives an exact filter. The solution,however,involves calculating conditional expected values, which can be forbiddingly infeasible, seeing that they involve integration over non-linear functions. Which is why we resort to approximation of expected values. In the following, a second order approximation of the exact filter is formulated by negating terms of third- and higher- order.

3.2 Second-order approximate filter

. We summarise the result below and proceed to show how we arrived at it.

Theorem 1. *(Second - order approximate filter). Suppose $f(x)$ and $h(x)$ are continuous functions, whose first and second order derivatives, f_x, f_{xx}, h_x, h_{xx}, respectively, exist. Then the second-order approximation equations for the exact filter equations, (40) - (41), neglecting the third- and higher- order derivatives and moments, are:*

$$d\hat{x}_t = f(\hat{x})dt + \frac{1}{2}p_t f_{xx}(\hat{x}dt + ph_x(\hat{x})r^{-1}(t)(dy_t - (h\hat{x}) + \frac{1}{2}p_t h_{xx}(\hat{x})), \tag{43}$$

$$\begin{aligned} dp_t = 2pf_x(\hat{x})dt + (q(t)g^2(\hat{x}) + pd(t)g_x^2(\hat{x}))dt - (p_t h_x(\hat{x}))^2 r^{-1} \\ + \frac{1}{2}p_t^2 h_{xx}(\hat{x})r^{-1}(t)(dy_t - (h(\hat{x}) + \frac{1}{2}p_t h_{xx}(\hat{x}))dt). \end{aligned} \tag{44}$$

proof. By Taylor expansion about $\hat{x}$, we have

$$f(x) \approx f(\hat{x}) + (x-\hat{x})f_x(\hat{x}) + \frac{1}{2}(x-\hat{x})^2 f_{xx}(\hat{x}), \tag{45}$$

$$h(x) \approx h(\hat{x}) + (x-\hat{x})h_x + \frac{1}{2}(x-\hat{x})^2 h_{xx}(\hat{x}), \tag{46}$$

$$q(t)g^2(x) \approx q(t)(g(\hat{x}) + (x-\hat{x})g_x(\hat{x} + \frac{1}{2}(x-\hat{x})^2 g_{xx}(\hat{x}))^2 \tag{47}$$

$$xf(x) \approx xf(\hat{x}) + x(x-\hat{x})f_x(\hat{x}) + \frac{1}{2}x(x-\hat{x})^2 f_{xx}(\hat{x}). \tag{48}$$

Taking expectations and noticing that $p = \widehat{(x-\hat{x})^2} = x\widehat{(x-\hat{x})}$, and that $(x-\hat{x}) = 0$ leads to

$$\hat{f} \approx f(\hat{x}) + \frac{1}{2}pf_{xx}(\hat{x}) \tag{49}$$

,

$$\hat{h} \approx h(\hat{x}) + \frac{1}{2}ph_{xx}(\hat{x}), \tag{50}$$

$$\widehat{q(t)g^2} \approx q(t)g^2(\hat{x}), \tag{51}$$

$$\widehat{xf} - \hat{x}\hat{f} \approx \hat{x}f(\hat{x}) + pf_x(\hat{x}) + \frac{1}{2}\hat{x}pf_{xx}(\hat{x}) - \hat{x}f(\hat{x} - \frac{1}{2}pf^{xx}(\hat{x}) = pf_x(\hat{x}). \tag{52}$$

It now remains to obtain the approximation of the term $\widehat{x^2h} - \widehat{x^2}\hat{h} - 2\hat{x}\widehat{xh} + 2\hat{x}^2\hat{h}$. By Taylor expansion

$$\widehat{x^2} \approx \hat{x}^2 + (x-\hat{x})^2. \tag{53}$$

Using (3.12) and (3.19)

$$\begin{aligned} \widehat{x^2h} \approx [(\hat{x}^2 + (x-\hat{x})2\hat{x} + (x-\hat{x})^2)(h(\hat{x} + (x-\hat{x}h_x(\hat{x}) + \frac{1}{2}(x-\hat{x})^2 h_{xx}(\hat{x}))]' \\ = \hat{x}^2 h(\hat{x}) + \frac{1}{2}\hat{x}^2 ph_{xx}(\hat{x}) + 2\hat{x}ph_x(\hat{x} + ph(\hat{x}) \end{aligned} \tag{54}$$

Similarly

$$\begin{aligned}\widehat{x^2h} &\approx [(\hat{x}^2 + (x-\hat{x})^2]\hat{}[(h(\hat{x}) + (x-\hat{x})h_x(\hat{x}) + \frac{1}{2}(x-\hat{x})^2 h_{xx}(\hat{x})]\hat{},\\ &= (\hat{x}^2 + P)(h(\hat{x}) + (x-\hat{x}) + \frac{1}{2}(x-\hat{x})^2 h_{xx}(\hat{x})),\\ &= \hat{x}^2 h(\hat{x}) + \frac{1}{2}\hat{x}^2 p h_{xx}(\hat{x}) + ph(\hat{x}) + \frac{1}{2}p^2 h_{xx}(\hat{x}).\end{aligned} \tag{55}$$

The remaining term, $-2\hat{x}\widehat{xh} + 2\hat{x}^2\hat{h}$, can be written in the form (3.18) , with h replacing f, which simplifies it a great deal. That is,

$$\begin{aligned}-2\hat{x}\widehat{xh} + 2\hat{x}^2\hat{h} &= -2\hat{x}(\widehat{xh} - \hat{x}\hat{h}),\\ &\approx -2\hat{x}ph_x(\hat{x}).\end{aligned} \tag{56}$$

Now collecting like terms together, we get:

$$\widehat{x^2h} - \widehat{x^2\hat{x}} - 2\hat{x}\widehat{xh} + 2\hat{x}^2\hat{h} = \frac{1}{2}p^2 h_{xx}(\hat{x}). \tag{57}$$

The approximate filter for the exact filter (41) - (42) is obtained by replacing the expected values of the terms with their approximations, which have been developed above. Second-order approximate filter for the vector case have been developed in[[4],[13]]

3.3 First-order approximate filter

The equations for the first-order approximation of the exact filter are summarised in the following Theorem.

Theorem 2. *(First-order approximate filter). Suppose $f(x)$ and $h(x)$ are continuous functions whose first order derivatives, f_x and h_x, respectively, exist. Then the first-order approximation equations for the exact filter equations, (3.7)-(3.8), omitting the second and higher central moments, are:*

$$d\hat{x}_t = f(\hat{x})dt + ph_x(\hat{x})r^{-1}(t)(dy_t - h(\hat{x})dt, \tag{58}$$

$$dp_t = 2pf_x(\hat{x})dt + (q(t)g^2(\hat{x}) + pq(t)g_x^2(\hat{x}))dt - (p_t h_x(\hat{x}))^2 r^{-1}. \tag{59}$$

Proof. Set to zero the second order derivatives of $f(x)$ and $h(x)$ in the equations for second-order approximate filter above. □

3.4 Second-order ensemble filter

Drawing inspiration from the formulation of ensemble Kalman Bucy Filter [[5],[7]], we devise an ensemble filter for non-linear models based on the second-order approximate equations developed above. In fact, it is possible to come up with a first-order and higher-order ensemble filters corresponding to the relevant approximate filter equations. The idea behind ensemble filters is to have, say M, hypotheses of the state, which we denote as $X_t := x_{t\,i=1}^{i\,M}$ at a given time,t. In the prediction step,the hypotheses are propagated according to the model signal. An update is achieved using a relevant minimum variance equation-or its approximation. The empirical mean $\hat{x}_t$ and covariance p_t, are respectively, obtained as follows.

$$\hat{x}_t = \frac{1}{M}\sum_{i=1}^{M} x_t^i; \qquad t_0 \le t, \tag{60}$$

$$p_t = \frac{1}{M-1}\sum_{i=1}^{m}(x_t^i - \hat{x}_t)^2; \qquad t_0 \leq t. \tag{61}$$

The propagation of each particle x_t^i,in the second-order ensemble filter formulation for scalar model system, (3.5) - (3.6), is given by,

$$dx_t^i = f(x_t^i,\theta)dt + g(x_t^i)q^{1/2}(t)d\beta_t^i + ph_x(x_t^i)r^{-1}(t)(dy_t - (h(x_t^i) + \frac{1}{2}p_t h_{xx}(x_t^i))dt); \qquad t_0 \leq t. \tag{62}$$

This is the Second order Extended Ensemble Filter (SoEEF), our new filter.

4 SoEEF Performance

The performance of the new filter SoEEF was tested using Lorenz 63 system of differential equations. Lorenz 63 is one of the most well-known chaotic systems, and it was originally derived by Edward Lorenz [20] [16] . The stochastic Lorenz 63 model,is given by

$$\begin{aligned} \frac{dx}{dt} &= \sigma(x_2 - x_1) \\ \frac{dy}{dt} &= x_1(\rho - x_3) - x_2 \\ \frac{dz}{dt} &= x_1x_2 - bx_3 \end{aligned} \tag{63}$$

with $\sigma = 10$, $r = \frac{8}{3}$ and $b = 28$ these constants determine the behaviour of the system. The three variables (x_1, x_2, x_3) evolve overtime and are sensitive to initial conditions leading to chaotic behaviour. This chaotic behaviour is similar to noise. Lorenz 63 are used to test data assimilation methods due to their sensitivity to the initial conditions. This means that a slight change in initial conditions of the variables leads to a significant change in the outcome. This was repeated with four other filters namely;Bootstrap Particle Filter (BPF),First order Extended Ensemble Filter (FoEEKF),First order Extended Kalman Bucy Filter (FoKBF),Second order Extended Kalman Bucy Filter (SoEKBF).

4.1 The Results

Fig. 1 represents the plot of RMSE (y-axis) against the reciprocal of ensemble of size M (x-axis). The figure has four different types of filters, the Bootstrap Particle Filter (BPF) represented by blue line, First order Extended Ensemble Filter (FoEEF) represented by green line, First order Extended Kalman Bucy Filter (FoEKBF) represented by orange line and the Second order Extended Kalman Bucy Filter (SoEKBF) represented by purple line.

The aim of the fig 1. is to compare the performance of the SoEEF against the other four filters. We can observe that as the number of ensemble size increases, the RMSE of BPF increases while that of FoEKBF,SoEKBF, FoEEF and the SoEEF decreases. However, SoEEF has the lowest RMSE as the ensemble size increases. This indicates that our filter, the SoEEF, performs better as it has a low RMSE.

The sizes of the ensemble used are $10, 15, 22, 26, 29, 34, 41, 46, 49$. Other settings are as follows: $dt = 0.001$, $R = 0.17$, and $g = 0.2$

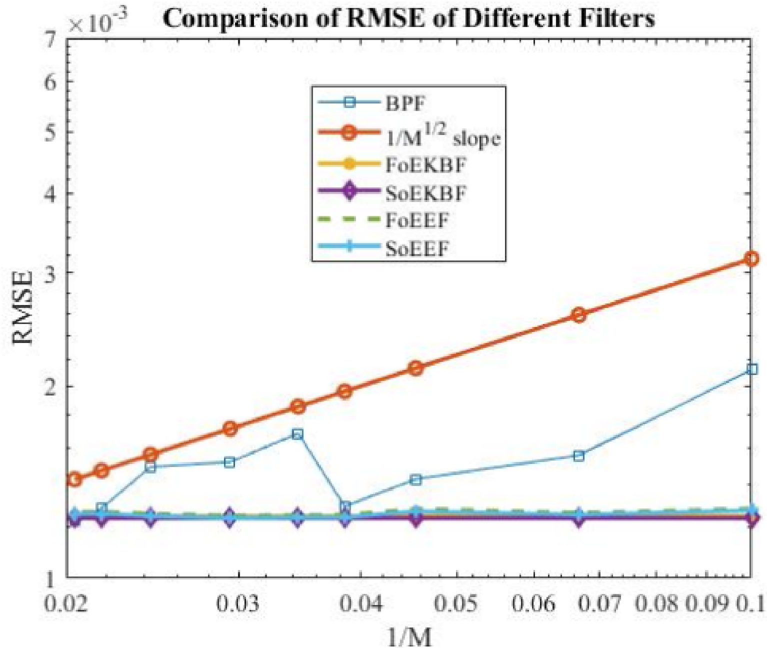

Figure 1: Root mean square error for the reciprocal of Ensemble.

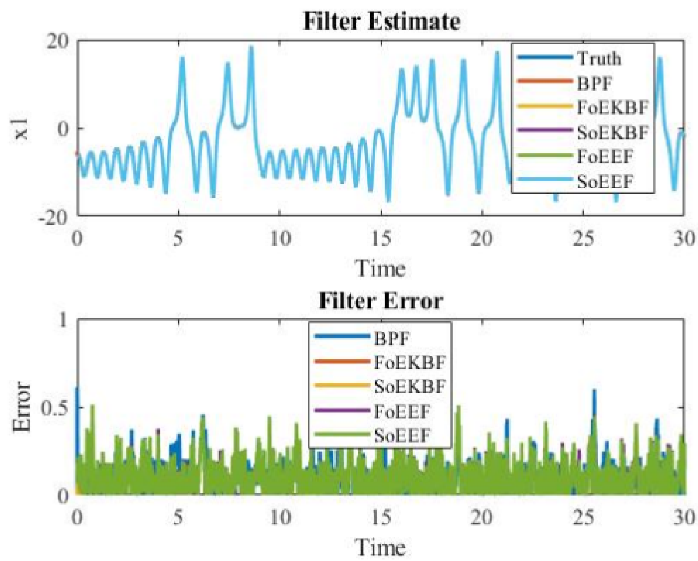

Figure 2: Filter estimate and filter error for $x(1)$ in Lorenz 63 model

Fig. 2 has two plots. The first plot is the comparison of the trajectory of true state generated from x_1 the first variable of Lorenz 63 and the trajectories of evolution of new SoEEF (light blue) with the trajectories of FoEEF(green) SoEKBF(purple),FoEKBF(orange),BPF(red).The paths are analysed on filter estimates using nine ensemble number that are measured between times $T = 0$ and $T = 30$, with a time step of 0.001. The second graph is a measure of filter error against times $T = 0$ and $T = 30$ for SoEEF(blue) model, FoEEF(green) model, SoEKBF (purple) model, FoEKBF (orange) model, BPF (red) model in the first variable.

In Fig. 2, we observe that the filter estimate illustrates that there is no noticeable deviations in trajectory of evolution between SoEEF and the trajectory of the true state generated from Lorenz 63 model . It can also be seen that the four other models, that is FoEEF SoEKBF, FoEKBF and BPF , also shows unvaried trajectories of evolution with both the true state trajectory and SoEKBF model trajectory.

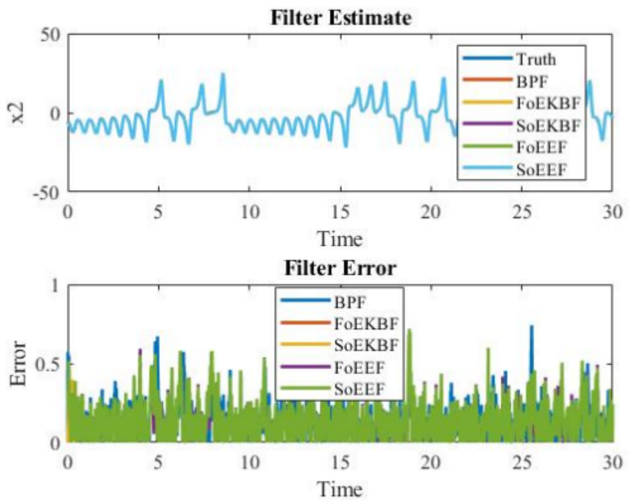

Figure 3: Filter estimate and filter error for $x(2)$ in Lorenz 63 model

From Fig. 3, we observe that there is no significant deviation that occurs between the output of SoEEF model and the output from other FOUR models as displayed in fig 3. SoEEF performs better than the other four filters.

Fig. 4 has two graphical displays with the first graph comparing the trajectories of evolution of new SoEEnKF,FoEEF, SoEKBF, FoEKBFn, BPF and the trajectory of true state for the third variable in Lorenz-63 model. The trajectories are conducted for Filter Estimates using nine ensemble number which are measured between times $T = 0$ and $T = 30$. The second graph measures the output from filter error against times $T = [0, 30]$ for SoEEF,FoEEnKF, SoEKBF, FoEKBF, BPF.

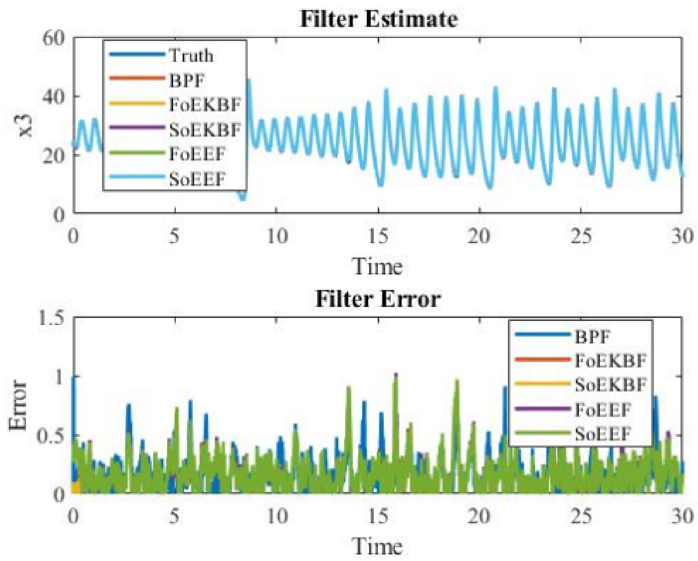

Figure 4: Filter estimate and filter error for $x(3)$ in lorenz63 model 94

5 Conclusion

The experiment of SoEEF, FoEEF, FoEKBF and SoEKBF conducted on a three dimensional Stochastic Lorenz 63 has shown that the performance of SoEEF, as also shown in RMSE plot, does not vary much from that of FoEEF,FoEKBF and SoEKBF. The graphical trajectories of SoEEF in the same three dimensional Lorenz 63 also show resemblance with trajectories of FoEEF, FoEKBF and SoEKBF for the three variables, $x1$, $x2$ and $x3$. As it is known that, as the dimensions increases, solutions to filtering problems become computationally expensive since it requires integration of the equation of evolution of covariances. This cost is reduced in SoEEF, which registers a remarkable performance even at low ensemble sizes (fig.1). In SoEEF, instead of integrating the equation of evolution of covariances, an empirically estimate of covariances is preferred. This greatly reduces the computational cost as compared to FoKBF and SoKBF, where integration of an equation of evolution of covariances is involved. SoEEF is favourable for parallel computing as it does not need re-sampling as in particle filters.

Competing Interests

Authors have declared that no competing interests exist.

References

[1] Angwenyi, D. (2019). Time-Continuous State and Parameter Estimation with application to SPDEs (Doctoral dissertation, PhD Thesis, Institute Fur Mathematik AG Numerische Mathematik, University of Potsdam).

[2] Armstrong, J., Brigo, D., and Hanzon, B. (2024). Optimal projection filters with information geometry. *Information Geometry*, 7(Suppl 1), 525-540.

[3] Baltieri, M., and Buckley, C. L. (2020). On kalman-bucy filters, linear quadratic control and active inference. *arXiv preprint arXiv:2005.06269.*

[4] Bass, R. W., Norum, V. D., and Schwartz, L. (2012). Optimal multichannel nonlinear filtering. Journal of Mathematical Analysis and Applications, 16(1), 152-164.

[5] Bergemann, K., and Reich, S. (2012). An ensemble Kalman-Bucy filter for continuous data assimilation. Meteorologische Zeitschrift, 21(3), 213.

[6] Crisan, D. (2014). The stochastic filtering problem: a brief historical account. Journal of Applied Probability, 51(A), 13-22.

[7] De Wiljes, J., Reich, S., and Stannat, W. (2018). Long-time stability and accuracy of the ensemble Kalman–Bucy filter for fully observed processes and small measurement noise. SIAM Journal on Applied Dynamical Systems, 17(2), 1152-1181.

[8] Doucet, A. (2001). Sequential Monte Carlo Methods in Practice.

[9] Dridi, N., Drumetz, L., and Fablet, R. (2021, August). Learning stochastic dynamical systems with neural networks mimicking the Euler-Maruyama scheme. *In 2021 29th European Signal Processing Conference (EUSIPCO)* (pp. 1990-1994). IEEE.

[10] Ertel, S. W., and Stannat, W. (2024). Analysis of the ensemble Kalman–Bucy filter for correlated observation noise. *The Annals of Applied Probability*, 34(1B), 1072-1107.

[11] Fontenele Magalhães, J. A., Emzir, M. F., and Corona, F. (2023). Approximations to the Solution of the Kushner-Stratonovich Equation for the Stochastic Chemostat. *bioRxiv*, 2023-01.

[12] Hornik, G. and Pavliotis, A. (2015). Stochastic Proceses and applications. Neural Imperial college London.

[13] Jazwinski, A. H. (2007). Stochastic processes and filtering theory. *Courier Corporation.*

[14] Kalman, R. E., and Bucy, R. S. (1961). New results in linear filtering and prediction theory.

[15] Lewis, F. L., Xie, L., and Popa, D. (2017). Optimal and robust estimation: with an introduction to stochastic control theory. CRC press.

[16] Lorenz, E. N. (1963). Deterministic nonperiodic flow. *Journal of atmospheric sciences*, 20(2), 130-141.

[17] Lee, S., Ko, H., and Oh, S. (Eds.). (2018). Multisensor Fusion and Integration in the Wake of Big Data, Deep Learning and Cyber Physical System: *An Edition of the Selected Papers from the 2017 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2017)* (Vol. 501). Springer.

[18] Rajawat, R. S., Singh, K. K., and Mishra, V. N. (2024). Approximation by modified Bernstein polynomials based on real parameters. Mathematical Foundations of Computing, 7(3), 297-309.

[19] Reich, S., and Cotter, C. (2015). Probabilistic forecasting and Bayesian data assimilation. *Cambridge University Press.*

[20] Sparre, P. (1982). The Lorenz Equations: Bifurcations, Chaos, and Strange Attractors. Springer-Verlag.

© *2024 Midenyo et al.; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribu-tion, and reproduction in any medium, provided the original work is properly cited.*