Sequential Monte Carlo Bayesian Estimation Using Gram Charlier Series Mixture Model

Syed Amer A. Gilani and P. L. Palmer

***Abstract*—In this paper discrete time nonlinear Bayesian filter using Gram Charlier Series Mixture (GCSM) model has been developed. Optimal nonlinear sequential state estimation can be described in a unified way by recursive Bayes' formula. The most important quantity of interest in Bayesian recursive formulation is state probability distribution of the system conditioned on available measurements. Exact optimal solution to Bayesian filtering problem is intractable as it requires an infinite dimensional process. Bayes' probability distribution can be approximated by orthogonal expansion of probability density function in terms of higher order moments of the distribution. In general, better series approximations to Bayes' distribution can be achieved by using higher order moment terms and Hermite polynomials termed as Gram Charlier Series (GCS). Sequential Monte Carlo (SMC) method has been adopted for approximating state predictive and filtering distributions parameterized as GCSM. GCSM based parametric bootstrap particle filters are derived for flexible use depending on inference problems under sparse measurement environment. Application of these sequential filters for satellite orbit determination using radar measurements is presented. The results have shown better/comparable performances over other SMC filtering methods such as Particle Filter and Gaussian Mixture Particle Filter (GMPF) under sparse measurement availability.**

***Index Terms*— Hermite polynomials; Gram Charlier Series; Gram Charlier Series Mixture Model; Monte Carlo; Bayes' estimation; Satellite orbit determination; Bootstrap filters;**

I. INTRODUCTION

Efficient algorithms for real time filtering of nonlinear dynamic systems based on Gaussian assumption of state prior and posterior Probability Density Function (PDF) includes the Extended Kalman Filter (EKF), Iterated Extended Kalman filter (IEKF) and variations such as H_∞EKF [1][2][3]. These methods may perform suboptimal for estimation of certain nonlinear problems with multi-modal or heavy tailed posterior PDFs [4]. The Gaussian Sum Filter (GSF) [5] considers this issue by approximating the posterior PDF with a Gaussian Mixture Model (GMM) which is essentially interpreted as a parallel bank of EKFs. However, due to the use of the EKF as a cardinal building block, it also suffers from similar shortcomings as the EKF. Significant improvements in approximations of posterior statistics using sigma points as in Sigma Point Filters Family (SPFF) were developed which provides better results and serves as efficient alternatives to EKF [6][7]. These algorithms are based on Gaussian assumption of the nonlinear system equations and commonly known as Unscented Kalman Filter (UKF) [6]. An efficient improvement to UKF based on an adaptive technique to overcome measurement and dynamic model mismatch is presented in [7]. The use of orthogonal series i.e., Hermite polynomials for correction to a Gaussian PDF was first presented by [8][9] known as Gram Charlier Series (GCS) or Edgeworth series. Edgeworth series has slight variations in ordering of Hermite polynomials terms [10]. Culver developed nonlinear filter based on GCS for continuous-discrete filtering paradigm [11]. Minimum Mean Square Estimates (MMSE) which is minimum variance solution of the state is developed by approximating Bayesian posterior PDF as Gaussian density multiplied with multidimensional Hermite polynomials [12]. Closed form analytical MMSE solutions for nonlinear filtering problem using third order GCS were derived. Numerical techniques like Gauss Hermite Quadrature [13] has also been used extensively to solve GCS based PDFs inside Bayesian integrals owing to their convenient form [14][15][16][17]. In method employing GCS developed by Challa [15] higher order moments were propagated using ito differential rule to compute coefficients of the series [18] and Bayes formula was solved using Gauss Hermite Quadrature using EKF or IEKF generated quadrature points. Edgeworth filters were developed by Horwood employing Gauss Hermite Quadrature based solution of Bayesian integrals for space surveillance and tracking [16]. Most of the dynamical systems in various applications such as satellite tracking and navigation systems are nonlinear [18][19]. This poses a substantial challenge to aerospace engineers and scientists to find efficient algorithms for real time estimation and prediction of such dynamical systems from the sequential observations [19][20]. Tightly coupled Global Positioning System (GPS) Pseudo-Range / Inertial Measurement Unit (IMU) and Precise Point Position (PPP)/IMU navigation systems employed in aerospace systems show nonlinearity during large IMU misalignments and GPS outages. To keep the advantages of the nonlinear filtering methods in dealing with such nonlinear systems, a Cubature Kalman Filter (CKF) [21] + EKF hybrid filtering

Manuscript received April 23, 2019. The work has been financially supported by National University of Sciences and Technology, Islamabad, Pakistan.

Syed Amer A. Gilani is with National University of Sciences and Technology (e-mail: gilani.amer@ gmail.com).

P.L. Palmer (Late) was with Surrey Space Centre, University of Surrey, UK.

.

method based on dual estimation framework is proposed by [20]. CKF is a nonlinear filtering method based on the spherical-radial Cubature rule. Being a deterministic sampling filtering method, CKF needs 2n (n=states/parameters of the system) Cubature points to propagate the state and covariance matrix, which shows a relatively smaller computational load than the UKF, as UKF mostly needs 2n+1 sigma points for the nonlinear states' propagation [20]. A class of filtering methods based on the Sequential Monte Carlo (SMC) approach had been surfaced in the literature known as Bootstrap Particle Filtering (PF) [22]. SMC can be approximately defined as a collection of methods that employs a Monte Carlo simulation scheme in order to resolve on line estimation and prediction requirements. The SMC technique achieves filtering by producing ensemble of weighted samples of the state variables or parameters in a recursive manner. Discrete samples are used to represent a complicated probability distribution. Importance sampling and weighted resampling are performed to complete the online filtering. There have also been many efficient modifications and improvements on these methods [23][24][25][26]. In this paper GCS Mixture (GCSM) model [27] is proposed for approximating state prior and posterior PDF to augment and improve the standard PF.

II. Optimal Bayesian Discrete Filtering

In filtering applications the problem is to recursively estimate posterior PDF for the states as one receives the observations. Consider the nonlinear Stochastic Discrete State Space Model (DSSM) [4]:

$$\mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}) + \mathbf{\Gamma}_k \mathbf{w}_k \tag{1}$$

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k, k) + \mathbf{v}_k \tag{2}$$

where, $\mathbf{x}_k \in \mathbb{R}^d$ is the d-dimensional state vector to be estimated, denoted with discrete time subscript *"k"*, $\mathbf{f}(.) \in \mathbb{R}^{d\times 1}$ is a nonlinear function which evolves the state from $(k-1)$ to k discrete instant of time, $\mathbf{\Gamma}_k \in \mathbb{R}^{d\times m}$ is a dispersion matrix, $\mathbf{y}_k \in \mathbb{R}^q$ is a q-dimensional measurement vector, $\mathbf{h}(.) \in \mathbb{R}^{q\times 1}$ is nonlinear measurement function of evolved state, $\mathbf{w}_k \in \mathbb{R}^m$ and $\mathbf{v}_k \in \mathbb{R}^q$ is the m-dimensional and q-dimensional mutually independent additive white Gaussian process and measurement noise variables, respectively. The whiteness of noise variables is equivalent to requiring the state and measurement sequences to be Markov processes (the development of filtering algorithm is restricted here to such processes only). The state variable $\mathbf{x}_k$ is usually considered as hidden variable, being measured only through $\mathbf{y}_k$ at discrete time instants. The estimation problem is termed nonlinear if at least one of the Equations: 1 or 2 is the nonlinear function of the state. In a Bayesian framework posterior PDF of the state $p(\mathbf{x}_k|\mathbf{y}_{1:k})$ given all the observations $\mathbf{y}_{1:k} = \{\mathbf{y_1}, \mathbf{y}_2, \ldots, \mathbf{y}_k\}$ constitutes the complete solution to the probabilistic inference problem and allows to compute any function of the state $g(\mathbf{x}_k)$. The nonlinear state space model given in Equation: 1 specify conditional transition PDF, $p(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{y}_{1:k-1})$ of the current state given the previous state and complete history of previous observations. Equation: 2 specify the likelihood of current observation $p(\mathbf{y}_k|\mathbf{x}_k)$ given the current state. The predictive conditional PDF $p(\mathbf{x}_k|\mathbf{y}_{1:k-1})$ is defined through Chapman-Kolmogorov Equation (CKE) expressed as [1][19]:

$$p(\mathbf{x}_k|\mathbf{y}_{1:k-1}) = \int_{-\infty}^{+\infty} p(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{y}_{1:k-1}) p(\mathbf{x}_{k-1}|\mathbf{y}_{1:k-1}) d\mathbf{x}_{k-1} \tag{3}$$

Here the previous posterior PDF at $k-1$ instant is defined as $p(\mathbf{x}_{k-1}|\mathbf{y}_{1:k-1})$ which has become prior PDF for the k step. The correction or measurement update step generates the posterior PDF function of the form:

$$p(\mathbf{x}_k|\mathbf{y}_{1:k}) = C p(\mathbf{y}_k|\mathbf{x}_k) p(\mathbf{x}_k|\mathbf{y}_{1:k-1}) \tag{4}$$

where, C is the normalization constant given by

$$\begin{aligned} C &\triangleq p(\mathbf{y}_k|\mathbf{y}_{k-1}) \\ &= \left(\int_{-\infty}^{+\infty} p(\mathbf{y}_k|\mathbf{x}_k) p(\mathbf{x}_k|\mathbf{y}_{1:k-1}) d\mathbf{x}_k\right)^{-1} \end{aligned} \tag{5}$$

The filtering problem is to estimate in recursive manner moments of $\mathbf{x}_k$ given $\mathbf{y}_{1:k}$. For any distribution $p(\mathbf{x}_k|\mathbf{y}_{1:k})$ this involves recursive estimation of the expected value of any function of $\mathbf{x}_k$ i.e., $E[g(\mathbf{x}_k)]$ utilizing posterior PDF obtained from Equation: 4, which requires computations of integrals of the form:

$$E[g(\mathbf{x}_k)] = \int_{-\infty}^{+\infty} g(\mathbf{x}_k) p(\mathbf{x}_k|\mathbf{y}_{1:k})\, d\mathbf{x}_k \tag{6}$$

For example, an optimal estimate of the state $g(\mathbf{x}_k) = \mathbf{x}_k$ in terms of Minimum Mean Square Error (MMSE) [1][4] estimation criterion would be:

$$E[g(\mathbf{x}_k)] \triangleq \hat{\mathbf{x}}_k = \int_{-\infty}^{+\infty} \mathbf{x}_k p(\mathbf{x}_k|\mathbf{y}_{1:k}) d\mathbf{x}_k \tag{7}$$

where, $E[.]$ is the expectation operator.

III. Gram Charlier Series And its Mixture Model

A. Univariate Gram Charlier Series

Univariate GCS expansion of the PDF around its best Gaussian estimate $p_g(x_k, \mu, \sigma^2)$ with mean μ and variance σ^2 is given by [27]:

$$\begin{aligned} &p_{gcs}(x_k) \approx p_g(x_k, \mu, \sigma^2) \times \\ &\left[1 + \frac{1}{3!}\kappa_3 h_3(x_k, \mu, \sigma^2) + \frac{1}{4!}\kappa_4 h_4(x_k, \mu, \sigma^2) + \cdots\right] \end{aligned} \tag{8}$$

where,
κ_i is i^{th} standardized cumulant ($\kappa_i = \frac{\kappa_i}{\sigma^i}$) and h_i is the univariate Hermite polynomial of order i. The standard Hermite polynomials obtained by putting $\mu = 0$ and $\sigma^2 = 1$ are defined as [10]:

$$\begin{gathered} h_1(z) = z, h_2(z) = z^2 - 1 \\ h_3(z) = z^3 - 3z, h_4(z) = z^4 - 6z^2 + 3 \end{gathered} \tag{9}$$

where,

$$z = \frac{(x - \mu)}{\sigma}$$

B. Multivariate GCS

If all the moments of a d-dimensional random vector $\mathbf{x}_k$ are finite, then any probability density $p(\mathbf{x}_k)$ can be represented by a Gaussian density $p_g\left(\mathbf{x}_k, \boldsymbol{\mu}_k, \mathbf{P}_k\right)$ with mean $\boldsymbol{\mu}_k$and $\mathbf{P}_k$ covariance matrix multiplied by an infinite series of Hermite polynomials as [10]:

$$\begin{gathered} p_{gcs}(\mathbf{x}_k) \approx p_g(\mathbf{x}_k, \boldsymbol{\mu}_k, \mathbf{P}_k) \\ \times \left[1 + \sum_{i,j,l} \frac{\kappa_{i,j,l}}{3!} h_{ijl}(\mathbf{x}_k, \boldsymbol{\mu}_k, \mathbf{P}_k) + \cdots \right] \end{gathered} \tag{10}$$

where,
function $h_{ijl}(\mathbf{x}_k, \boldsymbol{\mu}_k, \mathbf{P}_k)$ is multidimensional Hermite polynomials, with corresponding input dimensions $i, j, l \in \{1, \dots d\}$, and $\kappa_{i,j,l}$ is the corresponding third cumulant over input dimensions i, j, l, and sum over all input dimensions i, j, l is considered. Hermite polynomials can be obtained by differentiating $p_g\left(\mathbf{x}_k, \boldsymbol{\mu}_k, \mathbf{P}_k\right)$ and defined in [10]:

$$\begin{gathered} h_i\left(\mathbf{x}_k, \boldsymbol{\mu}_k, \mathbf{P}_k\right) = \mathrm{P}_{ij}^{-1}\left(x_j - \mu_j\right) \\ h_{ij}\left(\mathbf{x}_k, \boldsymbol{\mu}_k, \mathbf{P}_k\right) = h_i h_j - \mathrm{P}_{ij}^{-1} \\ h_{ijl}\left(\mathbf{x}_k, \boldsymbol{\mu}_k, \mathbf{P}_k\right) = h_i h_j h_l - h_i \mathrm{P}_{jl}^{-1}\{3\} \end{gathered} \tag{11}$$

where,
P_{ij}^{-1} and similar forms indicate ij^{th} component of inverse of covariance matrix, x_j and μ_j indicate j^{th} variable and its mean respectively. The subscripts implicitly imply summation over indices. The connection between cumulants and multivariate central moments is defined as [10]:

$$\begin{gathered} \kappa_{i,j} = \mathrm{P}_{ij} \\ \kappa_{i,j,l} = \mathrm{P}_{ijl}^{(3)} \end{gathered} \tag{12}$$

where,
$\mathrm{P}_{ijl}^{(3)}$ and similar forms indicate ijl^{th} component of third order (co-skewness) tensor. The bracket notations used in Equation: 11 is sum over combinations of indices. For example:

$$\mathrm{P}_{ij}\mathrm{P}_{lm}\{3\} = \mathrm{P}_{ij}\mathrm{P}_{lm} + \mathrm{P}_{il}\mathrm{P}_{jm} + \mathrm{P}_{im}\mathrm{P}_{jl}$$

C. Gram Charlier Series Mixture Model

GCS expansions do not estimate well near the centroid of the PDF also it does not always result in positive definite approximations [28]. To improve density estimation accuracy one can increase the order of these expansions, but unfortunately it renders the estimate more sensitive to outliers. Rather than increasing the order of the GCS, it was suggested by [27] to use mixtures of GCS expanded Gaussian kernels of moderate order.

1) Univariate GCSM Model

Considering Univariate GCS expressed in Equation: 8, a mixture of such PDF (expanded until order four of Hermite polynomials) can be formulated as:

$$\begin{aligned} p_{gcsm}(x_k) \approx & \sum_{g=1}^{G} \alpha_k^{(g)} p_g\left(x_k, \mu_k^{(g)}, \sigma_k^{2(g)}\right) \\ & \times \left[1 + \frac{1}{3!}\kappa_3^{(g)} h_3\left(x_k, \mu_k^{(g)}, \sigma_k^{2(g)}\right) \right. \\ & \left. + \frac{1}{4!}\kappa_4^{(g)} h_4\left(x_k, \mu_k^{(g)}, \sigma_k^{2(g)}\right) + \cdots \right] \end{aligned} \tag{13}$$

where "*G*" are the number of mixands of the mixture model. The parameters of above PDF can be estimated using statistical Expectation Maximization (EM) Algorithm provided in ReBEL Matlab Toolkit [29][30][31]:

$$\begin{gathered} \alpha^{(g)} = \frac{1}{N}\sum_{j}^{N} \tau_j^{(g)}, \mu_1^{(g)} = \frac{1}{N}\sum_{j}^{N} \frac{\tau_j^{(g)} x_j}{\alpha^{(g)}} \\ \mu_2^{(g)} = \frac{1}{N}\sum_{j}^{N} \frac{\tau_j^{(g)}\left(x_j - \mu_1^{(g)}\right)^2}{\alpha^{(g)}} \\ \mu_3^{(g)} = \frac{1}{N}\sum_{j}^{N} \frac{\tau_j^{(g)}\left(x_j - \mu_1^{(g)}\right)^3}{\alpha^{(g)}} \end{gathered} \tag{14}$$

where "*N*" are the number of data particles, $\tau_j^{(g)}$ is Gaussian posterior probabilities and μ_1, μ_2, μ_3 are mean, second, and third order moments respectively. Likewise, higher order moments can be computed. The standardized third cumulant is:

$$\kappa_3^{(g)} = \frac{\mu_3^{(g)}}{\sqrt{\mu_2^{(g)}\mu_2^{(g)}\mu_2^{(g)}}} \tag{15}$$

2) Multivariate GCSM Model

The Multivariate GCSM model expansion up to order three about its Gaussian estimates is given by [10]:

$$p_{gcsm}(\mathbf{x}_k) \approx \sum_{g=1}^{G} \alpha_k^{(g)} p_g\left(\mathbf{x}_k, \boldsymbol{\mu}_k^{(g)}, \mathbf{P}_k^{(g)}\right) \times \left[1 + \sum_{i,j,l} \frac{\kappa_{i,j,l}^{(g)}}{3!} h_{ijl}\left(\mathbf{x}_k, \boldsymbol{\mu}_k^{(g)}, \mathbf{P}_k^{(g)}\right) + \cdots\right] \tag{16}$$

The parameters of above PDF, moments and cumulants can also be estimated using EM Algorithm [29][30][31] using following equations:

$$\alpha^{(g)} = \frac{1}{N}\sum_{j}^{N} \tau_j^{(g)} \tag{17}$$

$$\boldsymbol{\mu}^{(g)} = \frac{1}{N}\sum_{j}^{N} \frac{\tau_j^{(g)}\mathbf{x}_j}{\alpha^{(g)}}$$

$$\mathbf{P}^{(g)} = \frac{1}{N}\sum_{j}^{N} \frac{\tau_j^{(g)}\left(\mathbf{x}_j - \boldsymbol{\mu}^{(g)}\right)\left(\mathbf{x}_j - \boldsymbol{\mu}^{(g)}\right)^{\mathrm{T}}}{\alpha^{(g)}}$$

$$\mathbf{P}^{(\mathbf{3})(g)} = \frac{1}{N}\sum_{j}^{N} \frac{\tau_j^{(g)}\left(\mathbf{x}_j - \boldsymbol{\mu}^{(g)}\right)\left(\mathbf{x}_j - \boldsymbol{\mu}^{(g)}\right)^{\mathrm{T}} \otimes \left(\mathbf{x}_j - \boldsymbol{\mu}^{(g)}\right)^{\mathrm{T}}}{\alpha^{(g)}}$$

where, time subscript *"k"* has been omitted for clarity and replaced with dimension variable *"j"* and $\otimes$ denotes Kronecker product.

IV. Gaussian Copula Random Number Generation

One of the most vital components of GCS based Particle filtering algorithms is the random number generator. Gaussian Copula is used to generate correlated multivariate random numbers with GCS marginal PDFs [32][33]. For example the Bi-variate Gaussian Copula function could be written as:

$$C(\mathrm{u},\mathrm{v}) = \mathcal{N}_\rho\left(\phi^{-1}(\mathrm{u}), \phi^{-1}(\mathrm{v})\right) = \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\phi^{-1}(\mathrm{u})}\int_{-\infty}^{\phi^{-1}(\mathrm{v})} e^{-\left(\frac{s^2 - 2\rho st + t^2}{2(1-\rho^2)}\right)} ds dt \tag{18}$$

where,

u, v and ρ are marginal probability distributions for Bi-variate random numbers and correlation coefficient respectively. We chose GCS as marginal distribution for each dimension in multivariate distributions and select ρ with Gaussian Copula to generate correlations. The following steps are used to generate Gaussian Copula based random numbers with GCS marginals up to order three:

- *Step 1*. Consider $\mathbf{x}_k$ is *d*-dimensional vector random variable. Compute mean μ, variance σ^2, skew μ_3 for each dimension separately.
- *Step 2*. Compute linear correlation ρ or rank correlation τ [33] to construct dependency. For example the multivariate linear correlation is expressed as

  $$\rho_{ij} = \frac{\mathrm{P}_{ij}}{\sqrt{\mathrm{P}_{ii}\mathrm{P}_{jj}}}$$

  where,

  P_{ij}, ij component of covariance matrix $\mathbf{P}$
- *Step 3*. Establish grid for each dimension.
- *Step 4*. Convert statistics from *Step 1* to standardized Cumulants $\kappa_i = \frac{\kappa_i}{\sigma^i}$ where κ_i are cumulants expressed in terms of central moments (only first four are shown):

  $$\kappa_2 = \sigma^2 = \mu_2$$

  $$\kappa_3 = \mu_3$$

  $$\kappa_4 = \mu_4 - 3\mu_2^2$$
- *Step 5*. Compute Cumulative Distribution Function (CDF) for Gram Charlier Marginals for each dimension as:

  $$\mathrm{CDF} = \int_{-\infty}^{\mathrm{x}_i} p_G(t_i, \mu, \sigma)\left[1 + \frac{1}{3!}\kappa_3 h_3(t_i, \mu, \sigma) + \cdots\right] dt_i$$
- *Step 6*. Compute inverse CDF $\phi^{-1}(u_i)$ for each dimension from step (5) by inverting the function.
- *Step 7*. Generate Gaussian Copula Uniform random variables with dependency structure as in *Step 2* by writing a Matlab function `copularnd`.

Generate vector random variables $\mathbf{x}_k$ by table look up method of probabilities from above mentioned *Step 3* and *Step 6* with structure provided by *Step 7*.

V. GCSM Particle Filtering

A. Particle Filtering

Particle filtering is based on MC simulations to obtain approximation of PDFs given in Equations: 3 to 7. The main objective is to sequentially sample and resample particles from a particular choice of PDF known as proposal PDF, considered by the filter as approximation of Bayes' posterior

PDF. The choice of proposal PDF is a major issue for the different variants of PF [19][23][24]. PFs employ MC integration scheme to compute integrals. Expectations for functions of states (Equation: 1) are computed from particles drawn from proposal PDF. The conditional transition PDF specified by nonlinear state space model given in Equation: 1, $p(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{y}_{1:k-1})$ is choice of proposal PDF for generic PF also known as Sequential Importance Sampling-Resampling (SIS-R) PF [22].

B. *GCSM Particle Filtering*

Based on improved fidelity of GCSM model demonstrated in ref [27][28], truncated GCS up to order three in a mixture model configuration is used in nonlinear SMC filtering. The Bayes' posterior and noise PDF in this filter are considered as GCSM. However, one may consider additive Gaussian noise also. The compact form of GCSM can be expressed as:

$$p_{gcsm}(\mathbf{x}_k) = \sum_{g=1}^{G} \alpha_k^{(g)} p_{gcs}\left(\mathbf{x}_k, \boldsymbol{\mu}_k^{(g)}, \mathbf{P}_k^{(g)}, \mathbf{P}_k^{(\mathbf{3})(g)}\right) \tag{19}$$

An important point to note is the ability of the GCSMPF to incorporate (additive) highly non-Gaussian process noise expressed compactly as:

$$p_{gcsm}(\mathbf{w}_k) = \sum_{i=1}^{I} \beta_k^{(i)} p_{gcs}\left(\mathbf{w}_k, \boldsymbol{\mu}_{\mathbf{w},k}^{(i)}, \mathbf{Q}_k^{(i)}, \mathbf{Q}_k^{\mathbf{3}(i)}\right) \tag{20}$$

During the time update, firstly the samples from the PDF expressed in Equations: 19 and 20 are drawn as per weights $\alpha_k^{(g)}$ and $\beta_k^{(i)}$. One may use the Sequential Importance Resampling (SIR) or Residual Resample (RR) as explained in Ref [4][31]. These samples are propagated through the nonlinear dynamical system $\mathbf{f}(.)$ (Equation: 1). By approximating the propagated distribution as GCSM one employs EM to obtain time updated *"G"* component state predictive GCSM PDF. The proposal PDF in this filter is considered as state predictive PDF available from the time update. In measurement update the samples are redrawn from state predictive GCSM PDF and the weights for *"M"* particles of each mixand are computed using the observation likelihood $p\left(\mathbf{y}_k|\mathbf{x}_k = \boldsymbol{x}_k^{(i)}\right)$ considered Gaussian as in Ref [23][24][31]. The weighted updates of parameters for each mixand are computed as:

$$\boldsymbol{\mu}_k^{(g)} = \frac{\sum_{j=1}^{M} w_k^{(j)(g)} \boldsymbol{x}_k^{(j)(g)}}{\sum_{j=1}^{M} w_k^{(j)(g)}},$$

$$\mathbf{P}_k^{(g)} = \frac{\sum_{j=1}^{M} w_k^{(j)(g)} (\mathrm{A})(\mathrm{A})^{\mathrm{T}}}{\sum_{j=1}^{M} w_k^{(j)(g)}} \tag{21}$$

$$\mathbf{P}_k^{(\mathbf{3})(g)} = \frac{\sum_{j=1}^{M} w_k^{(j)(g)} (\mathrm{A})(\mathrm{A})^{\mathrm{T}} \otimes (\mathrm{A})^{\mathrm{T}}}{\sum_{j=1}^{M} w_k^{(j)(g)}}$$

where,
$\mathrm{A} = \boldsymbol{x}_k^{(j)(g)} - \boldsymbol{\mu}_k^{(g)}$. The Pseudo-code for the filter is presented in Table: 1.

TABLE I. PSEUDO-CODE GCSM PARTICLE FILTER

Time Update

1. For $g = 1, \dots G$, obtain samples as per the weights $\alpha_{k-1}^{(g)}$

$$\left\{\mathbf{x}_{(k-1)}^{(j)}\right\}_{j=1}^{M} \sim p_{gcsm}(\mathbf{x}_{k-1}|\mathbf{y}_{1:k-1})$$
$$= \sum_{g=1}^{G} \alpha_{k-1}^{(g)} p_{gcs}\left(\mathbf{x}_{k-1}, \boldsymbol{\mu}_{k-1}^{(g)}, \mathbf{P}_{k-1}^{(g)}, \mathbf{P}_{k-1}^{(\mathbf{3})(g)}\right)$$

2. For $i = 1, \dots, I$, obtain samples $\left\{\mathbf{w}_{(k-1)}^{(j)}\right\}_{j=1}^{M}$ from $p_{gcsm}(\mathbf{w}_{k-1})$ as per weights $\beta_{k-1}^{(i)}$

3. Propagate particles $\left\{\mathbf{x}_{(k-1)}^{(j)}\right\}_{j=1}^{M}$ through nonlinear function Equation: 1 from time instants $k-1 \rightarrow k$ to get $\left\{\mathbf{x}_{(k)}^{(j)}\right\}_{j=1}^{M}$

4. Add particles from 2 and 3 above.

5. Perform EM step on propagated on particles from 4 above to extract "*G*" component GCSM time updated predictive PDF:

$$p_{gcsm}(\mathbf{x}_k|\mathbf{y}_{1:k-1}) = \sum_{g=1}^{G} \alpha_k^{(g)} p_{gcs}\left(\mathbf{x}_k, \boldsymbol{\mu}_k^{(g)}, \mathbf{P}_k^{(g)}, \mathbf{P}_k^{(\mathbf{3})(g)}\right)$$

Measurement Update

1. For $g = 1, \dots G$, obtain samples from $p_{gcs}^{(g)}(\mathbf{x}_k)$ and denote them as $\left\{\mathbf{x}_k^{(j)(g)}\right\}_{j=1}^{M}$.

2. For $g = 1, \dots G$ each j = *1,…,M,* compute weights , $w_k^{(j)(g)} = p\left(\mathbf{y}_k|\mathbf{x}_k = \boldsymbol{x}_k^{(j)(g)}\right)$

3. For $g = 1, \dots G$, Compute mean, covariance and tensor components $\mathbf{P}_k^{(\mathbf{3})(g)}$: (see Equation: 21)

$$\boldsymbol{\mu}_k^{(g)} = \frac{\sum_{j=1}^{M} w_k^{(j)(g)} \boldsymbol{x}_k^{(j)(g)}}{\sum_{j=1}^{M} w_k^{(j)(g)}},$$

$$\mathbf{P}_k^{(g)} = \frac{\sum_{j=1}^{M} w_k^{(j)(g)} \left(\boldsymbol{x}_k^{(j)(g)} - \boldsymbol{\mu}_k^{(g)}\right)\left(\boldsymbol{x}_k^{(j)(g)} - \boldsymbol{\mu}_k^{(g)}\right)^{\mathrm{T}}}{\sum_{j=1}^{M} w_k^{(j)(g)}}$$

4. Update weights $\alpha_k^{(g)} = \alpha_{k-1}^{(g)} \frac{\sum_{j=1}^{M} w_k^{(j)g}}{\sum_g^G \sum_{j=1}^{M} w_k^{(j)g}}, \alpha_k^{(g)} = \frac{\alpha_k^{(g)}}{\sum_g^G \alpha_k^{(g)}}$

Inference

The conditional mean state estimate $\hat{\mathbf{x}}_k = E[\mathbf{x}_k|\mathbf{y}_{1:k}]$ and Covariance can be estimated by:

$$\hat{\mathbf{x}}_k = \sum_{g=1}^{G} \alpha_k^{(g)} \boldsymbol{\mu}_k^{(g)},$$

$$\hat{\mathbf{P}}_k = \sum_{g=1}^{G} \alpha_k^{(g)} \left(\mathbf{P}_k^{(g)} + \left(\boldsymbol{\mu}_k^{(g)} - \hat{\mathbf{x}}_k\right)\left(\boldsymbol{\mu}_k^{(g)} - \hat{\mathbf{x}}_k\right)^{\mathrm{T}}\right)$$

Optional Step: Residual Resampling applied on mixture weights to avoid use of insignificant weights in next time step

The pictorial form of time and measurement update for GCSMPF pseudo-code is described in Figure: 1 and 2 respectively. The use of (-) superscript in $\mathbf{x}_k^-$ indicates time updated estimates of state variable.

Fig. 1. Time update GCSMPF.

Fig. 2. Measurement Update GCSMPF.

VI. TRACKING OF A SATELLITE USING RADAR MEASUREMENTS

In this section algorithms discussed for GCSMPF would now be implemented for tracking of a satellite using ground based radars. We shall compare GSCMPF with SIS-R PF and GSPF. The equations of motion for true model used in this experiment are given as:

$$\dot{\boldsymbol{r}} = \mathbf{v} \tag{22}$$

$$\dot{\mathbf{v}} = -\frac{\mu_E}{r^3}\boldsymbol{r} + \boldsymbol{a}_G$$

where, $\boldsymbol{r} = [X, Y, Z]^{\mathrm{T}}$, $\mathbf{v} = [\dot{X}, \dot{Y}, \dot{Z}]^{\mathrm{T}}$ are position and velocity of a space object in ECI coordinates, $\mathrm{v} = |\mathbf{v}|$, $r = |\boldsymbol{r}|$, $\boldsymbol{a}_G$ = perturbation acceleration due to zonal gravitational harmonic up to J_4.

Given some specific initial conditions $\mathbf{x}_0 = [\boldsymbol{r}_0\ \mathbf{v}_0]^{\mathrm{T}}$ these equations (Equation: 22) are integrated using numerical method such as Runge-Kutta (RK-4) to get time history of position and velocity in ECI reference frame termed here as true trajectory. The true trajectory is being measured by a radar system fixed at some location on Earth. Now we describe the radar measurement system (see Figure: 3). The ECI position vector of satellite is related to radar range vector and radar site vector through following equation [19][34]:

$$\boldsymbol{r} = \boldsymbol{R}_s + \boldsymbol{\rho} \tag{23}$$

where,

$\boldsymbol{r}$ = ECI coordinates of satellite,
$\boldsymbol{R}_s$ = ECI coordinates of radar site, and
$\boldsymbol{\rho}$ = range vector from radar site to satellite.

The range vector $\boldsymbol{\rho}$ from the radar site to satellite is described in Topocentric coordinate system (see Figure: 3 for illustration) in terms of the "zenith", "east" and "north" as:

$$\boldsymbol{\rho} = \rho_u \hat{\boldsymbol{u}} + \rho_e \hat{\boldsymbol{e}} + \rho_n \hat{\boldsymbol{n}} \tag{24}$$

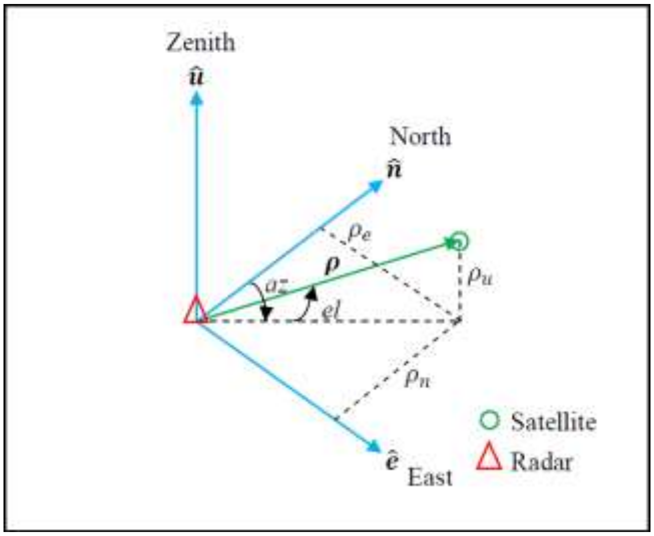

Fig. 3. Topocentric Coordinate System for RADAR Observations

The range can be obtained as:

$$\rho = \sqrt{\rho_u^2 + \rho_e^2 + \rho_n^2} \tag{25}$$

The azimuth (*az*) and elevation (*el*) angles of a radar antenna are expressed by:

$$az = \tan^{-1}\left(\frac{\rho_e}{\rho_n}\right) \tag{26}$$

$$el = \tan^{-1}\left(\frac{\rho_u}{\sqrt{\rho_e^2 + \rho_n^2}}\right)$$

The east, north, and zenith unit vectors in Topocentric coordinate system is given by [34]:

$$\hat{\boldsymbol{e}} = \begin{pmatrix} -\sin\lambda \\ \cos\lambda \\ 0 \end{pmatrix}, \hat{\boldsymbol{n}} = \begin{pmatrix} -\sin\varphi\cos\lambda \\ -\sin\varphi\sin\lambda \\ \cos\varphi \end{pmatrix}, \hat{\boldsymbol{u}} = \begin{pmatrix} \cos\varphi\cos\lambda \\ \cos\varphi\sin\lambda \\ \sin\varphi \end{pmatrix} \tag{27}$$

where,

φ and λ are geographical latitude and longitude of radar site respectively. By defining the orthogonal transformation as:

$$\mathfrak{N} = (\hat{\boldsymbol{e}}\, \hat{\boldsymbol{n}}\, \hat{\boldsymbol{u}}\,)^{\mathrm{T}} \tag{28}$$

The satellite's Topocentric coordinates in terms of radar site latitude and longitude may be obtained through following transformation [34]:

$$\boldsymbol{\rho} = \begin{pmatrix} \rho_e \\ \rho_n \\ \rho_u \end{pmatrix} = \mathfrak{N}(\mathcal{R}_{\mathbf{z}}(\Theta)\boldsymbol{r} - \boldsymbol{R}_s) \tag{29}$$

where, $\mathcal{R}_{\mathbf{z}}(.)$ stands for rotation about z-axis and Θ = Greenwich Mean Sidereal Time (GMST) [34][35]. GMST is also termed as Greenwich hour angle which denotes the angle between the mean vernal equinox of date and the Greenwich meridian. It is a direct measure of Earth's rotation and expressed in angular units as well as time. For example 360 degrees (2π) correspond to 24 hours. Time calculations for satellite orbit predictions and determination are usually carried out in Julian Date (JD) [34] due to its continuous nature. A Julian Date (JD) is the number of days since noon 1 January, 4713 BC including the fraction of day. Presently, the JD numbers are already quite large therefore a Modified Julian Date is defined as: $\mathrm{MJD} = \mathrm{JD} - 2400000.5$. Eglin US Air Force Base (AFB) is selected as radar site with $\varphi = 30.2316$ deg and $\lambda = 86.2147$ deg W. Each measurement consists of range, azimuth and elevation angles and the measurement errors were considered to be Gaussian distributed with following variances (adapted from reference [19]):

$$\sigma_{\mathrm{range}} = 25\ \mathrm{m}, \sigma_{\mathrm{azimuth}} = 0.015\ \mathrm{deg}$$
$$\sigma_{\mathrm{elevation}} = 0.015\ \mathrm{deg}$$

Initial conditions in terms of classical orbital elements [34] of a satellite to generate true trajectory are:

$$\begin{aligned} a &= 6981.0425\ \mathrm{km} \\ e &= 7.5629 \times 10^{-4} \\ I &= 51.6041\ \mathrm{deg} \\ \Omega &= 25.0038\ \mathrm{deg} \\ \omega &= 339.4915\ \mathrm{deg} \\ M &= 56.8164\ \mathrm{deg} \end{aligned} \tag{30}$$

where,

a = semi-major axis, *e* = eccentricity, *I* = inclination, Ω = right ascension of ascending node, ω = argument of perigee and *M* = mean anomaly. The orbital period of this satellite is 96 min 45 sec (approx).

VII. Comparison of Non Linear Filters

Satellite initial estimates could be extremely uncertain especially in case of a sparsely tracked object. Therefore, we would observe filtering performance with uncertainty in position variances as $10^5\ \mathrm{m}^2$ and velocity variances as $500\ \mathrm{m}^2.\mathrm{s}^{-2}$:

$$\widehat{\mathbf{P}}_0 = \mathrm{diag}([10^5\ 10^5\ 10^5\ 500\ 500\ 500])$$

$$\widehat{\mathbf{P}}_0^{(3)} = 0$$

$$\hat{\mathbf{x}}_0 = \mathbf{x}_0 + \left[\sqrt{10^5}\ \sqrt{10^5}\ \sqrt{10^5}\ \sqrt{500}\ \sqrt{500}\ \sqrt{500}\right]^{\mathrm{T}}$$

where,

$\mathbf{x}_0$ = true initial conditions in ECI coordinates. ($\hat{\mathbf{x}}_0$) = Initial estimates in ECI coordinates. The position and velocity deviation of our initial estimate ($\hat{\mathbf{x}}_0$) from true initial state ($\mathbf{x}_0$) is $316.22\ \mathrm{m}$ and $22.36\ \mathrm{m.s}^{-1}$ respectively. The filter trajectory for estimation is obtained using two body gravitational acceleration term perturbed by zonal harmonic J_2. Firstly, we would consider frequency of measurements by the radar as 0.1 Hz for 2 min. The time history of Root Mean Square Error (RMSE) for position and velocity are shown in Figures: 4 to 9. These errors have been obtained by averaging over 100 runs for each filter. Two mixand mixture models are considered for GCSMPF and GSPF filters respectively. These figures indicate comparable filtering performance. After the first set of observations the second set of observations is taken once the satellite completes one orbital period. Again the observations are taken as 0.1 Hz for 2 min. The time history of Root Mean Square Error (RMSE) for position and velocity are shown in Figures: 10 to 15. These errors have been obtained after averaging 50 runs for each filter. Gaussian process noise is considered for these simulations.

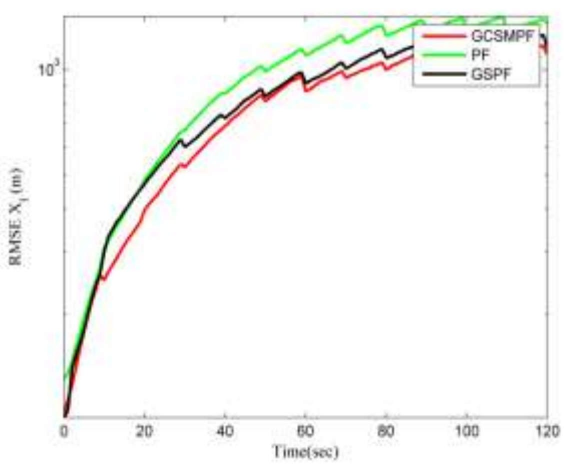

Fig. 4. Time history of position RMSE in ECI (X-Axis).

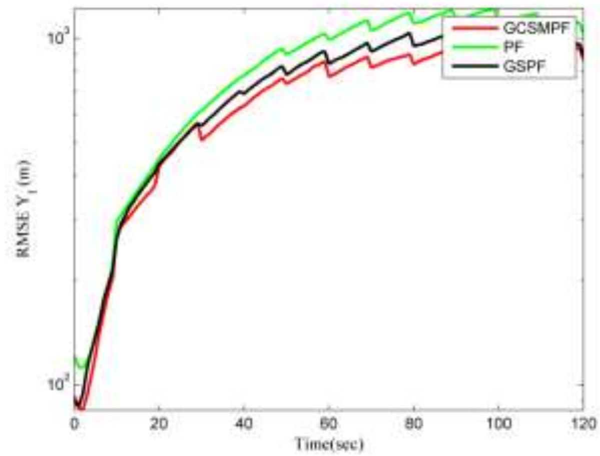

Fig. 5. Time history of position RMSE in ECI (Y-Axis).

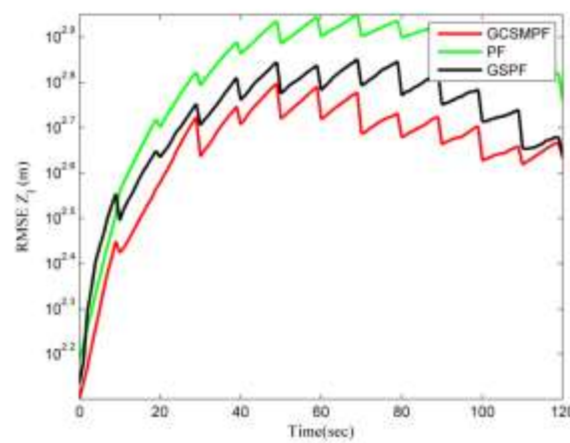

Fig. 6. Time history of velocity RMSE in ECI (Z-Axis).

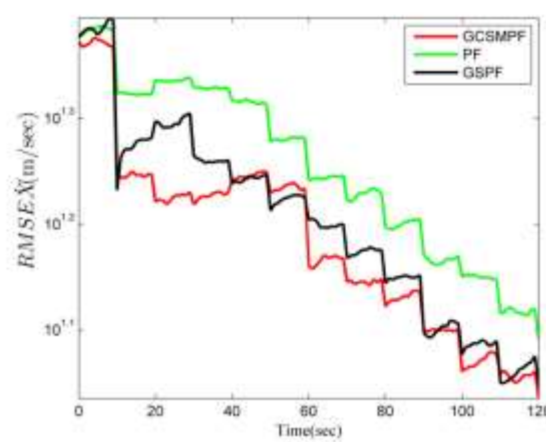

Fig. 7. Time history of velocity RMSE in ECI (X-Axis)

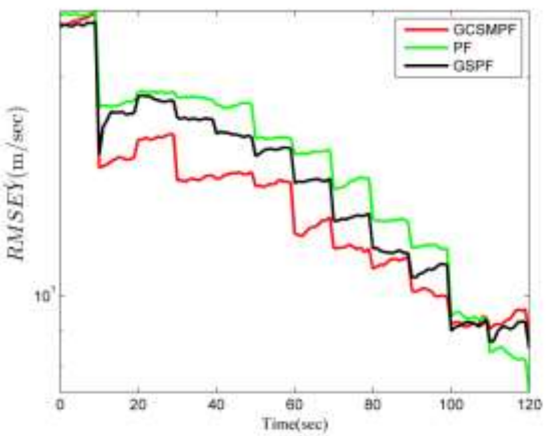

Fig. 8. Time history of velocity RMSE in ECI (Y-Axis)

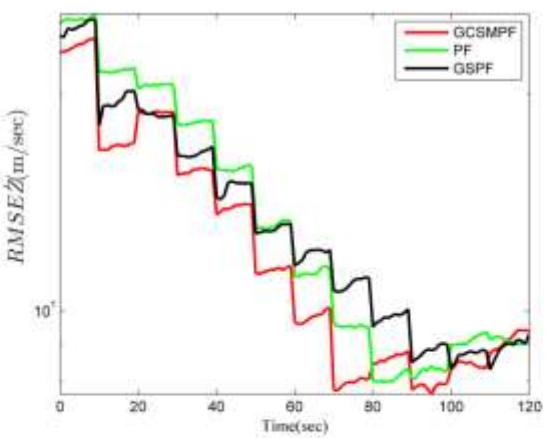

Fig. 9. Time history of velocity RMSE in ECI (Z-Axis)

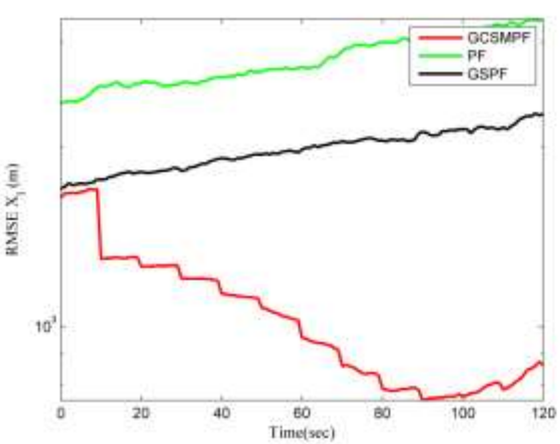

Fig. 10. Time history of position RMSE in ECI (X-Axis) after one orbital period.

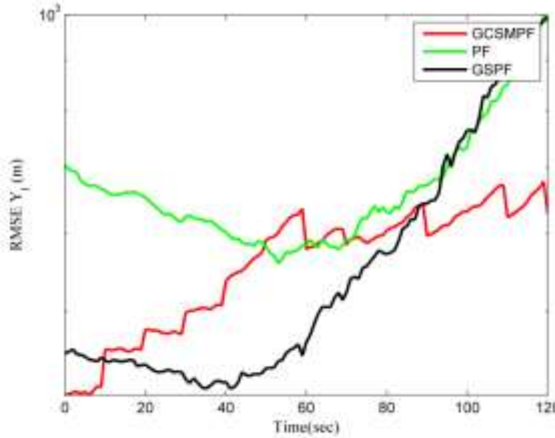

Fig. 11. Time history of position RMSE in ECI (Y-Axis) after one orbital period.

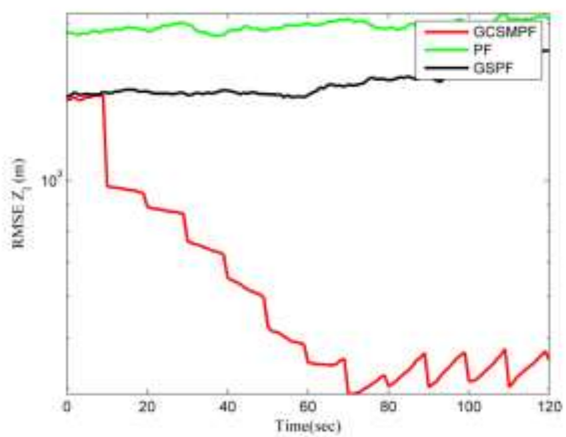

Fig. 12. Time history of position RMSE in ECI (Z-Axis) after one orbital period.

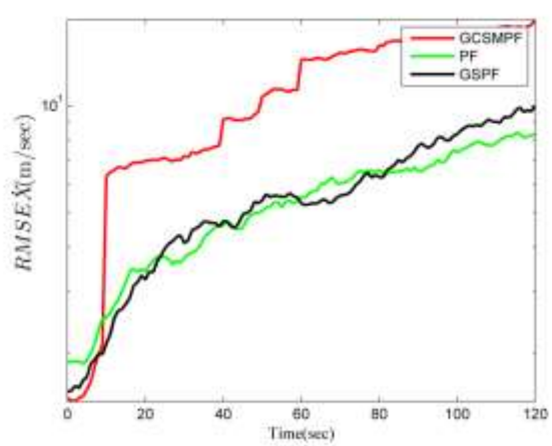

Fig. 13. Time history of velocity RMSE in ECI (X-Axis) after one orbital period

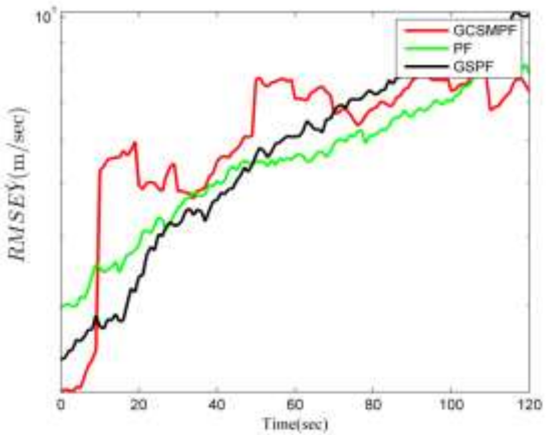

Fig. 14. Time history of velocity RMSE in ECI (Y-Axis) after one orbital period.

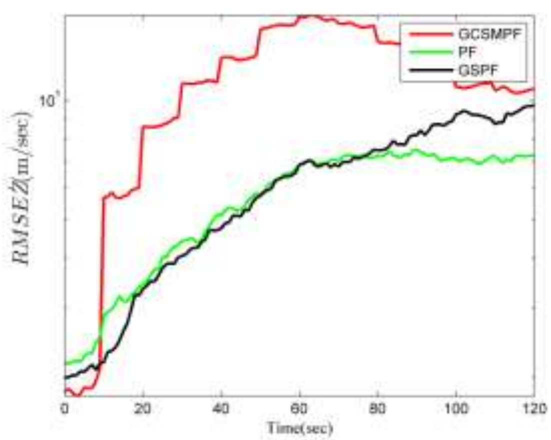

Fig. 15. Time history of velocity RMSE in ECI (Z-Axis) after one orbital period.

VIII. CONCLUSION

In this paper filtering algorithm based on SMC method using GCSM has been described. The algorithm has been compared with PF, GSPF for nonlinear orbit determination through radar measurements. The results show improvements/comparable performance in RMSE for ECI coordinates. GCS and its mixtures can be considered as better choice for replacement of Gaussian PDF in nonlinear filtering applications especially for improvement in particle filtering. An important aspect of filters based on higher order GCS and its mixture is computational complexity associated with generation of random numbers. Gaussian copula based methods are used which may not be always optimal. Therefore, there is a need for development of better random number generator for GCS. In order to implement discrete-time filtering the continuous-time nonlinear dynamical systems used in the experiments are discretized using a fixed time step of numerical integration method (RK-4). In general high fidelity numerical solution is obtained by keeping a very short time step (order of millisecond). This significantly affects the speed of execution in real time particle filtering for satellite which owes to high dimensionality and more number of particles used for such problems. However, GCSMPF can be implemented in parallel processing using high speed Very Large Scale Integrated Circuit (VLSI) based implementation for real time filtering.

ACKNOWLEDGMENT

The author acknowledges Rudolph van der Merwe and Eric A. Wan for provision of open source ReBEL Matlab Toolkit [4][31] for educational purposes, which facilitated development of GCSMPF and comparative analysis with other bootstrap PFs.

REFERENCES

[1] Shalom,Y,B., Li,X,R., Kirubarajan,T., "Estimation with Application to Tracking and Navigation," Wiley-Interscience Publication 2001.

[2] Havlik, J., Straka, O., "Performance evaluation of iterated extended Kalman filter with variable step-length," 12th European Workshop on Advanced Control and Diagnosis (ACD 2015), Journal of Physics: Conference Series 659 (2015) 012022, doi:10.1088/1742-6596/659/1/012022.

[3] Zhao, J., "Dynamic State Estimation with Model Uncertainties Using $H\infty$ Extended Kalman Filter," IEEE Transaction on Power Systems, Volume: 33 , Issue: 1 , Jan. 2018.

[4] Merwe, R.V.D., "Sigma-Point Kalman Filters for Probabilistic Inference in Dynamic State-Space Models," PhD Dissertation OGI School of Science and Engineering at Oregon Health and Science University, Apr 2004.

[5] Alspach, D, L., Sorenson, H, W., "Nonlinear Bayesian Estimation using Gaussian Sum Approximation," *IEEE Transaction on Automatic Control*, Vol. AC-17, No.4, pages. 439-448, Aug 1972.

[6] Julier, S, J., and Uhlmann, J, K., "A New Extension for Kalman Filter to Nonlinear Systems," In The Proceedings of AeroSense: The 11th International Symposium on Aerospace/Defense Sensing, Simulation and Controls, Orlando, FL, USA, 1997. SPIE. Multi Sensor Fusion, Tracking and Resource Management II.

[7] Zhou, W., Hou, J. "A New Adaptive Robust Unscented Kalman Filter for Improving the Accuracy of Target Tracking," IEEE Access, date of publication 10 Jun 2019, D.O.I: 10.1109/Access.2019.2921794.

[8] Charlier, C,V,L., "Uber Dir Darstelling Willkurlicher Funktionen," Ark.Mat Astr.och Fys.2 No.20 1-35, 1906.

[9] Edgeworth, F, Y., "The Law of Error," *Cambridge Philos. Soc.* 20, 36-66 and 113-141, 1905.

[10] McCullagh, P., "Tensor Methods in Statistics," Publisher Chapman and Hall, 1987.

[11] Culver, C, O., "Optimal Estimation for Nonlinear Stochastic Systems," PhD Dissertation Massachusetts Institute of Technology, Mar 13, 1969.

[12] http://mathworld.wolfram.com/HermitePolynomial.html [Viewed on: 21 Oct 2019].

[13] Abramowitz, M., Stegun, I.A., "Handbook of Mathematical Functions With Formulas, Graph and Mathematical Tables", National Bureau of Standards and Applied Mathematics, 1964.

[14] Bucy, R,S., Hecht, C., Captain Senne, K,D., "An Engineer's Guide to Building Nonlinear Filters," USAF Final Project Report No. 7904-00-17, May 1972.

[15] Challa, S., Shalom, Y, B., Krishnamurty, V., "Nonlinear Filtering via Generalized Edgeworth Series and Gauss-Hermite Quadrature," *IEEE Transaction on Signal Processing*, Vol.48, No.6, pages. 1816-1820, Jun 2000.

[16] Horwood, J., Poore, A., "Adaptive Gaussian Sum Filters for Space Surveillance", IEEE Transactions on Automatic Control, 2010.

[17] Kizner, W., "Optimal Nonlinear Estimation Based on Orthogonal Expansion", Technical Report 32-1366, JPL California Institute of Technology, NASA, 1969.

[18] Jazwinski, A, H., "Stochastic Process and Filtering Theory," Published by New York Academic, 1970.

[19] Lee, D.J., "Nonlinear Bayesian Filtering with Applications to Estimation and Navigation," PhD Dissertation, Texas AandM University May 2005.

[20] Zhao, Y., "Cubature+Extended Hybrid Kalman Filtering Method and its Application in PPP/IMU Tightly-Coupled Navigation Systems," IEEE Sensors Journals D.O.I: 10.1109/JSEN.2015.2469105, IEEE Sensors Journal, 2015.

[21] Pesonen.H., Piche.R., "Cubature-based Kalman Filters for Positioning," 7th Workshop on Positioning, Navigation and Commuincation WPNC 2010.

[22] Gordon, N., Salmond, D., Ewing, C., "Bayesian State Estimation for Tracking and Guidance using the Bootstrap Filter," *Journal of Guidance Control and Dynamics*, Vol. 18, No. 6, pages. 1434-1443, Nov-Dec 1995.

[23] Kotecha, J, H., Djuric, P, M., "Gaussian Particle Filtering," *IEEE Transactions on Signal Processing*, Vol. 51, No. 10, pages. 2592-2601, Oct 2003.

[24] Kotecha, J, H., Djuric, P, M., "Gaussian Sum Particle Filtering for Dynamic State Space Models," Proceedings of ICASSP-2001, Salt Lake City, Utah, May 2001.

[25] Wang, F., Zhang, J., Lin, B., and Li, X., "Two Stage Particle Filter for Nonlinear Bayesian Estimation" IEEE Access-Special Section on Multimedia Analysis for Internet of Things (IoT) D.O.I: 10.1109/ACCESS.2018.2808922, Date of Publication: Feb 23, 2018.

[26] Doucet, A., Godsill, S., and Andrieu, C., "On Sequential Monte Carlo Sampling Methods for Bayesian Filtering," Statistical Computing Vol. 10, No.3, pp. 197-208, 2000.

[27] Van Hulle, M, M., "Edgeworth-Expanded Gaussian Mixture Density Modeling," *Journal of Neural Computation*, Vol. 17, No. 8, Pages. 1706-1714, Aug 2005.

[28] Gilani, S.A.A., Palmer, P., "Nonlinear Bayesian Estimation Based on Mixture of Gram Charlier Series", Proceedings of IEEE Aerospace Conference, Mar 2012.

[29] Dempster. A.P., Liard. N.M., and Rubin. D.B., "Maximum Likelihood for Incomplete Data via the EM Algorithm", Journal of Royal Statistical Society B, 39, 1-38.

[30] Redner. R.A., and Walker. H.F., "Mixture Densities, Maximum Likelihood and the EM algorithm", SIAM Rev., 26, 195-239.

[31] https://github.com/SaeedKeshavarzi/ReBEL (Acccessed on 27 Apr 2019).

[32] Nelson. R.B., "An Introduction to Copulas", Published by Springer, 2006.

[33] Stuart, A., Ord, J, K., "Kendall's Advanced Theory of Statistics," Vol. 1 Distribution Theory, 6th Edition, Copublished in Americas by Halsted Press an Imprint of John Wiley and Sons Inc, 1994.

[34] Montenbruck, O., Gill, E., "Satellite Orbits Models, Methods and Applications," Published by Springer, 2005.

[35] Vallado, D, A., "Fundamentals of Astrodynamics and Applications (Second Edition)," The Space Technology Library, Published by Microcosm Press and Kluwer Academic Publisher 2004.

Syed Amer Ahsan Gilani is Assistant Professor at National University of Sciences and Technology, Islamabad, Pakistan. He did his MSc in Satellite Engineering in 2007 and PhD in Electronics Engineering in 2012, both from University of Surrey, UK. His major interests are Satellite State Estimation and Orbit Determination.

Phil Palmer (Late) was a Professor in Astrodynamics at Surrey Space Centre. Phil was appointed Deputy Director of the Surrey Space Centre (2000-2005) and remained as head of the Astrodynamics team. He had established a research group in Spacecraft Autonomy and also supported the development of the Space Propulsion group. He had interests in computer vision and its application for autonomous navigation of a robot and relative satellite motion. Phil had published widely in the areas of Astronomy, Computer Vision, Pattern Recognition, Remote Sensing and Astronautics. He was a member of the AIAA and a fellow of the IET.