EDGEWORTH FILTERS FOR SPACE SURVEILLANCE TRACKING

Joshua T. Horwood, Nathan D. Aragon, and Aubrey B. Poore

Numerica Corporation, 4850 Hahns Peak Drive, Suite 200, Loveland CO, 80538

ABSTRACT

The accurate and consistent representation of a space object's uncertainty is essential in the problems of data association (correlation), conjunction analysis, sensor resource management, and anomaly detection. While standard Kalman-based filtering algorithms, Gaussian assumptions, and covariance-weighted metrics are very effective in data-rich tracking environments, their use in the data sparse environment of space surveillance is largely inadequate. It is shown how improved uncertainty consistency can be maintained using the higher fidelity Edgeworth or adaptive Gaussian mixture filters in an orbital element space and how statistics beyond a Gaussian state and covariance can be represented correctly. A simulation scenario which considers the implications of correct uncertainty management to data association (correlation) and anomaly detection is presented.

1. INTRODUCTION

A prerequisite to space situational awareness (SSA) control and reconnaissance is the consistent characterization of the state uncertainty of resident space objects (RSOs) in the space catalog as well as additional objects such as uncorrelated tracks (UCTs). The problem of tracking RSOs and correctly managing their uncertainties presents some unique challenges not found in other tracking environments. In contrast to air or ground tracking, the space surveillance environment is *data-starved*. Typical RSO tracking problems (such as resolving UCTs) require the long-term propagation of state uncertainties, often on the order of several orbital periods, using high fidelity dynamical models in the absence of measurement or track updates. *Uncertainty consistency*, the accurate reflection of the discrepancy from truth exhibited by an estimated state, is critical in the problems of data/track association processing, collision avoidance including probability of collision and probability of reacquisition, sensor cueing and tasking, and anomaly (e.g., maneuver, change) detection. Consequently, inconsistent covariances or uncertainties cause misassociations (correlation errors) that can substantially degrade system level performance.

The state uncertainty of an RSO is commonly assumed to be Gaussian which, although may be appropriate initially, will inevitably become non-Gaussian if propagated for a sufficiently long time period. This departure from "Gaussianity" plagues the standard Kalman-based filtering algorithms such as those of the extended (EKF) and unscented (UKF) variety. More accurate filters which better approximate model nonlinearities are therefore sought. One such example is the adaptive Gaussian mixture filter (AGMF) which has been investigated in SSA applications by the authors [1, 2] and other researchers [3–6].

Rather than using Gaussian mixtures, we propose an alternate representation of RSO state uncertainties defined through an *Edgeworth series*, which is an asymptotic expansion of the probability density function in terms of its cumulants. The corresponding Edgeworth filter, first proposed in a series of papers of Singer [7, 8], propagates the state mean, covariance, and higher-order cumulants (up to some fixed order) using the unscented transform. Thus, the Edgeworth filter is a natural extension of the UKF and the higher-order Gauss-Hermite filters. The primary objective of this work is to compare and contrast the performance of the Edgeworth filter and AGMF and to assess their suitability for use in SSA.

To demonstrate the impact of correct uncertainty characterization and management, we have evaluated the prediction error component of the likelihood ratio used to score the association of a track (or tracklet) to an orbit in a low Earth orbit simulation scenario. Results are compared using the UKF, Edgeworth filter, and AGMF. The traditional (UKF) covariance-weighted prediction error, while sufficient for short term prediction in a data-rich environment, is demonstrated to be inadequate for the long term prediction required in the data-starved space surveillance environment. Further, the impact on correct anomaly detection is also presented. Thus, we conclude that correct uncertainty management can lead to robust tracking, surveillance, and anomaly detection not currently present in operational systems.

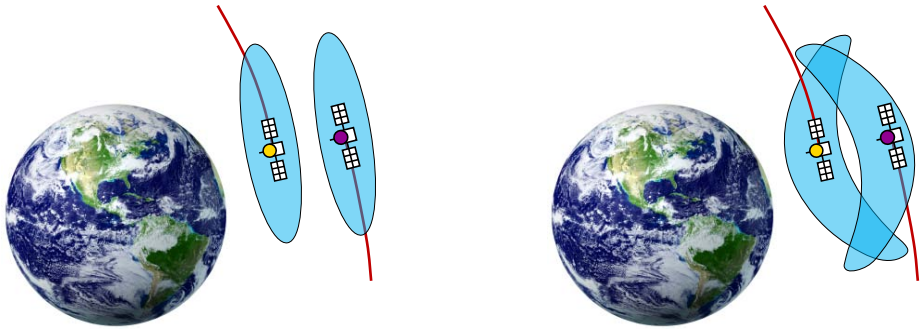

Figure 1. Depiction of the probability of association (POA) or collision between two satellites. A Gaussian assumption in the uncertainties (left) could imply a low POA. The true uncertainties could be non-Gaussian (right) and could imply a high POA.

The plan of the paper is as follows. In Section 2, we give an overview of uncertainty management and reiterate its importance for SSA. In Section 3, the mathematical theory of nonlinear estimation and filtering is developed starting from the Bayesian state estimator equations. This theory is then used in Section 4 to outline the derivation of the Edgeworth filter. In Section 5, an online metric based on the prediction error is derived which provides a means to assess track state uncertainty consistency. Finally, results are presented in Section 6 and conclusions are made in Section 7.

2. OVERVIEW OF UNCERTAINTY MANAGEMENT

Dynamical and measurement modeling errors are ubiquitous in any tracking environment and space surveillance is no exception. Knowledge of the precise state (e.g., position and velocity) of each RSO with 100% confidence cannot be expected and a degree of uncertainty will always be present. Providing an accurate and truthful representation of the state uncertainty is the objective of uncertainty management.

The assumption of a *Gaussian* or *normal* state uncertainty is often made to simplify the ensuing mathematics and facilitate computations. An n-dimensional random variable $\boldsymbol{x}$, which could represent a six-dimensional position-velocity or orbital element state of an RSO, is Gaussian if its probability density function (PDF) enjoys the form

$$p(\boldsymbol{x}) = \frac{1}{\sqrt{\det(2\pi\mathbf{P})}} \exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T\mathbf{P}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right] \equiv \mathcal{N}(\boldsymbol{x};\boldsymbol{\mu},\mathbf{P}). \tag{1}$$

In (1), the mean (or expected value) of $\boldsymbol{x}$ is $\boldsymbol{\mu}$ and provides an estimate of the state. For a Gaussian PDF, the mean $\boldsymbol{\mu}$ also coincides with the mode which is the point of maximum likelihood. The uncertainty in a Gaussian state $\boldsymbol{x}$ is encoded in the symmetric positive-definite covariance matrix $\mathbf{P}$. The level surfaces (i.e., surfaces of equal likelihood) of a Gaussian are ellipsoids; their orientation and principal axes lengths are determined from the eigenvectors and eigenvalues of $\mathbf{P}$.

The Gaussian assumption works remarkably well in data-rich tracking environments such as air and missile defense, but its use in the data-starved environment of space surveillance can be harder to justify. In particular, the propagation of uncertainties through nonlinear dynamical models over very long time spans (perhaps on the order of months if resolving UCTs) can induce highly non-Gaussian uncertainties. Instead of looking ellipsoidal, the uncertainties typically appear "banana" shaped even when represented in coordinate systems adapted to the dynamics such as orbital elements (see Figure 5 in Section 6). Representation of such non-Gaussian densities present formidable computational challenges since they possess an infinite hierarchy of higher-order statistics (cumulants). Thus, in this context, uncertainty consistency is how to obtain a consistent *finite-dimensional* representation of the non-Gaussian PDF and how to accurately propagate the defining statistics.

The need for non-Gaussian uncertainties is illustrated in Figure 1 which illustrates the significance of correct uncertainty management in the problems of track association and collision analysis. The depicted

scenario shows two closely spaced objects (CSOs) which could associate or collide. The goal is to provide a measure of confidence or probability of association (POA) or collision. Without any knowledge in the uncertainty of the CSOs, it is impossible to quantify the POA; the question of "how close is close?" is purely subjective. If the uncertainties of the CSOs are assumed Gaussian, then the POA is computed using a covariance-weighted norm of the difference between the states (i.e., the Mahalanobis distance) and could be very low as seen in the left-half of the diagram (since the ellipsoidal uncertainties do not overlap). In the right half of Figure 1, the true state uncertainties could be non-Gaussian. If the non-Gaussian statistics are correctly used to evaluate the POA, then one could obtain a higher POA than what one would obtain if the uncertainties were approximated by Gaussians as in the left half of the figure.

3. NONLINEAR ESTIMATION AND FILTERING

The problem of nonlinear filtering requires the definition of dynamical and measurement models. It is assumed that the state evolves deterministically according to the first-order dynamical system

$$\boldsymbol{x}'(t) = \boldsymbol{f}(\boldsymbol{x}(t), t). \tag{2}$$

For maximal generality, one could model the dynamics using a (continuous-time) stochastic model with an additive Gaussian white-noise diffusion term. In space surveillance, the physics encoded by $\boldsymbol{f}$ are high-fidelity and we believe it is advantageous to avoid including Gaussian white process noise since any mismodeling in said noise could contaminate the model. Indeed, the methodology presented in this paper, including the Gaussian sum and Edgeworth filters, are readily extendible to stochastic dynamical systems. A corresponding discrete-time model can be derived from (2) via

$$\boldsymbol{x}_k = \boldsymbol{f}_{k-1}(\boldsymbol{x}_{k-1}), \tag{3}$$

where $\boldsymbol{x}_k = \boldsymbol{x}(t_k)$ and $\boldsymbol{f}_{k-1}$ denotes the solution flow at time t_{k-1}. A sequence of measurements $\boldsymbol{Z}_k \equiv \{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_k\}$ is related to the corresponding kinematic states $\boldsymbol{x}_k$ via measurement functions $\boldsymbol{h}_k$ according to the discrete-time measurement model

$$\boldsymbol{z}_k = \boldsymbol{h}_k(\boldsymbol{x}_k) + \boldsymbol{v}_k, \tag{4}$$

where $\{\boldsymbol{v}_k\}$ is a zero-mean Gaussian white-noise sequence with $\mathrm{E}[\boldsymbol{v}_k\, \boldsymbol{v}_j^T] = \mathbf{R}_k\, \delta_{kj}$.

In the Bayesian approach to dynamic state estimation (see, for example, [9]), one constructs the *posterior* probability density function (PDF) of the state based on information of a *prior* state and received measurements. Encapsulating all available statistical information, the posterior PDF $p(\boldsymbol{x}_k|\boldsymbol{Z}_k)$ may be regarded as the complete solution to the estimation problem and various optimal state estimates can be computed from it.

In the recursive filtering approach, measurement data is processed *sequentially*, rather than as a batch. Given the initial density of the state, $p(\boldsymbol{x}_0) \equiv p(\boldsymbol{x}_0|\boldsymbol{Z}_0)$, the PDF $p(\boldsymbol{x}_k|\boldsymbol{Z}_k)$ is obtained recursively in two stages, namely prediction and correction, as illustrated in the flowchart of Figure 2. For the former, the transitional density $p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1})$ is obtained from (3) in conjunction with the standard change of variables theorem for PDFs. Equivalently, it can be obtained from the continuous-time model (2) by solving the noiseless Fokker-Planck-Kolmogorov (FPK) equation [10]

$$\frac{\partial p}{\partial t} = -\nabla_{\boldsymbol{x}}^T(p\boldsymbol{f}), \tag{5}$$

where $\nabla_{\boldsymbol{x}}$ is the gradient with respect to $\boldsymbol{x}$ viewed as a column operator. In the measurement update stage, also called the fusion step, the density $p(\boldsymbol{z}_k|\boldsymbol{x}_k)$ is evaluated from the measurement model (4). The term $p(\boldsymbol{z}_k|\boldsymbol{Z}_{k-1})$ in the denominator of the correction step is called the *prediction error* and appears in the likelihood ratios for scoring an assignment of a report (i.e., $\boldsymbol{z}_k$) to a track. Thus, its accurate evaluation is critical for correct data association (correlation).

Analytical solutions to the prediction and correction equations in Figure 2 are generally intractable and are only known in a few restrictive cases. In practice, models are nonlinear and states can be non-Gaussian; one must be content with an approximate or suboptimal algorithm for the Bayesian state estimator. While the EKF and UKF are used extensively in air and missile tracking, state uncertainties are only represented

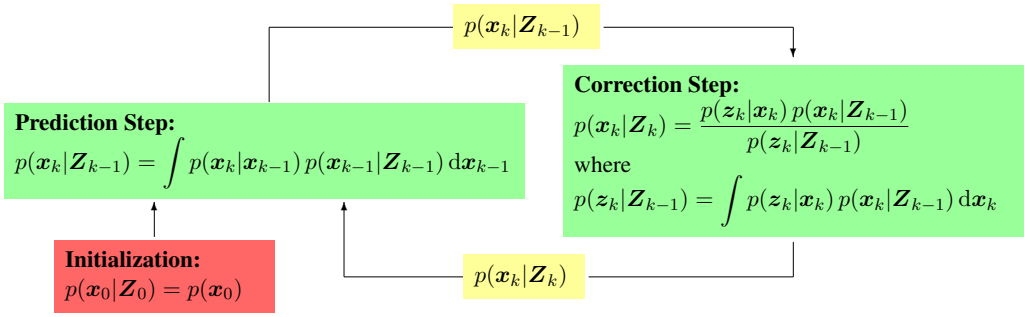

Figure 2. The predictor-corrector step for the recursive Bayesian state estimator.

by a covariance matrix and this assumption need not be valid in the space surveillance environment. Because of the need to propagate uncertainties over extended time intervals in the absence of measurement updates, higher-order cumulants (e.g., skewness, kurtosis) can become non-negligible and must be represented in order to achieve uncertainty consistency. In the next section we present the Edgeworth filters which can correctly correctly capture statistics beyond a state and covariance with the potential to retain uncertainty consistency over long propagations.

4. EDGEWORTH FILTER

In this section, we give a brief introduction to the Edgeworth filter and its derivation using the Bayesian state estimator outlined in the previous section. To begin, we first need to define our notation for multivariate cumulants which form the building blocks of the Edgeworth series representation. We then give the generalized unscented transform used to implement the prediction and correction steps of the Edgeworth filter.

4.1 Multivariate Cumulants

Let $\boldsymbol{x}$ be an n-dimensional random variable with components x^i. The cumulants of order p of $\boldsymbol{x}$, denoted by* $\kappa^{i_1,\ldots,i_p}$, are defined from the cumulant generating function $K_{\boldsymbol{x}} : \mathbb{R}^n \to \mathbb{R}$:†

$$K_{\boldsymbol{x}}(\boldsymbol{\xi}) = \ln \mathrm{E}[\mathrm{e}^{\xi_i x^i}] = \xi_i \kappa^i + \tfrac{1}{2!}\xi_i\xi_j\kappa^{i,j} + \tfrac{1}{3!}\xi_i\xi_j\xi_k\kappa^{i,j,k} + \cdots . \tag{6}$$

Conversions to and from the cumulants, moments, and central moments of $\boldsymbol{x}$ are provided in McCullagh [11]. For example, in terms of the central moments

$$\mathring{\kappa}^{i_1\cdots i_p} = \mathrm{E}[(x^{i_1} - \mu^{i_1})\cdots(x^{i_p} - \mu^{i_p})], \qquad \mu^i = \mathrm{E}[x^i] = \kappa^i,$$

expressions for the cumulants up to fourth order are given by

$$\kappa^{i,j} = \mathring{\kappa}^{ij}, \quad \kappa^{i,j,k} = \mathring{\kappa}^{ijk}, \quad \kappa^{i,j,k,\ell} = \mathring{\kappa}^{ijk\ell} - \mathring{\kappa}^{ij}\mathring{\kappa}^{k\ell} - \mathring{\kappa}^{ik}\mathring{\kappa}^{j\ell} - \mathring{\kappa}^{i\ell}\mathring{\kappa}^{jk}. \tag{7}$$

In particular, note that κ^i and $\mathring{\kappa}^{ij}$ are the components of the mean and covariance of $\boldsymbol{x}$, respectively.

4.2 Edgeworth Representation

Denoting $\boldsymbol{\mu}$ and $\mathbf{P}$ as the mean and covariance of $\boldsymbol{x}$ and $\boldsymbol{\kappa} = (\kappa^{i,j,k}, \kappa^{i,j,k,\ell}, \ldots)$ as the hierarchy of cumulants of $\boldsymbol{x}$ of third-order and higher, the Edgeworth representation of $\boldsymbol{x}$ is

$$p(\boldsymbol{x}) = \mathcal{N}(\boldsymbol{x}; \boldsymbol{\mu}, \mathbf{P})\mathcal{E}(\boldsymbol{x}; \boldsymbol{\mu}, \mathbf{P}, \boldsymbol{\kappa}), \tag{8}$$

*The components of $\boldsymbol{x}$ and its moments and cumulants are indexed by superscripts (rather than subscripts) because they transform like contravariant tensors under linear transformations.

†In (6) and in what follows, the *Einstein summation convention* will be implicitly assumed: any repeated upper and lower tensor index implies summation over that index from 1 to n. For example, $\xi_i x^i = \sum_{i=1}^n \xi_i x^i$.

where $\mathcal{N}(\boldsymbol{x};\boldsymbol{\mu},\mathbf{P})$ is the Gaussian PDF defined in (1) and the Edgeworth correction $\mathcal{E}(\boldsymbol{x};\boldsymbol{\mu},\mathbf{P},\boldsymbol{\kappa})$ is[‡]

$$\mathcal{E}(\boldsymbol{x};\boldsymbol{\mu},\mathbf{P},\boldsymbol{\kappa}) = 1 + \tfrac{1}{6}\kappa^{i,j,k}h_{ijk} + \big[\tfrac{1}{24}\kappa^{i,j,k,\ell}h_{ijk\ell} + \tfrac{1}{72}\kappa^{i,j,k}\kappa^{\ell,m,n}h_{ijk\ell mn}\big] + \mathcal{O}(N^{-3/2}). \tag{9}$$

In (9), the $h_{i_1\cdots i_p}$ are the components of the Hermite tensor of order p with respect to $\boldsymbol{\mu}$ and $\mathbf{P}$ which are defined by

$$h_{i_1\cdots i_p} = \frac{(-1)^p}{\mathcal{N}(\boldsymbol{x};\boldsymbol{\mu},\mathbf{P})}\frac{\partial^p}{\partial x^{i_1}\cdots\partial x^{i_p}}\mathcal{N}(\boldsymbol{x};\boldsymbol{\mu},\mathbf{P}). \tag{10}$$

For computational purposes, the series (9) is truncated at some finite order $p \geqslant 3$ resulting in the approximation $\mathcal{E}(\boldsymbol{x};\boldsymbol{\mu},\mathbf{P},\boldsymbol{\kappa}) \approx \mathcal{E}(\boldsymbol{x};\boldsymbol{\mu},\mathbf{P},\boldsymbol{\kappa}_{3:p})$, where $\boldsymbol{\kappa}_{3:p}$ denotes the hierarchy of cumulants of order three to p.

4.3 Generalized Unscented Transform

The UKF and its higher-order generalizations, the Gauss-Hermite (GH) filters, are based on the *unscented transform* [12,13]. Starting with an n-dimensional Gaussian random variable $\boldsymbol{x} \sim N(\boldsymbol{\mu},\mathbf{P})$ and a nonlinear transformation $\boldsymbol{f} : \mathbb{R}^n \to \mathbb{R}^m$, the unscented transform approximates the mean and covariance of the transformed variable $\boldsymbol{f}(\boldsymbol{x})$. Equivalently, the unscented transform provides a means of integrating the function $\boldsymbol{f}$ multiplied by a Gaussian weight function. Gauss-Hermite quadrature effectively approximates the integral by evaluating the function $\boldsymbol{f}$ at strategically chosen sigma points (or nodes) and then taking a weighted average.

Multidimensional Gauss-Hermite quadrature schemes which have a minimal number of nodes required to achieve a given accuracy are developed by Genz and Keister (GK) [14]. The resulting quadrature rules[§] enjoy the property that the number of nodes grows only polynomially with dimension when the order is fixed. For example, the number of nodes (sigma points) for third- and fifth-order GH-GK quadrature methods is $2n+1$ and $2n^2+1$, respectively. In contrast, these numbers become 2^n and 3^n, respectively, when the naive Cartesian product node set is assumed. The GH-GK nodes also have the advantage of being *nested*; nodes from a low order method are embedded in all quadrature methods of higher order. This feature enables one to compute higher-order estimates while re-using the function evaluations from a lower-order method. The difference between the high and low order approximations can be used to estimate the error of the integration.

In the implementation of the Edgeworth filter using Gauss-Hermite quadrature, we will need to formulate a generalization of the unscented transform which permits the evaluation of the mean, covariance, and higher-order cumulants of a non-linear transformation of a Gaussian or any variable represented as an Edgeworth series. If a random variable $\boldsymbol{x}$ is represented by the truncated Edgeworth series (8) and $\boldsymbol{f} : \mathbb{R}^n \to \mathbb{R}^m$, then

$$\mathrm{E}[\boldsymbol{f}(\boldsymbol{x})] = \int \mathcal{N}(\boldsymbol{x};\boldsymbol{\mu},\mathbf{P})\mathcal{E}(\boldsymbol{x};\boldsymbol{\mu},\mathbf{P},\boldsymbol{\kappa})\boldsymbol{f}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

This expectation can be approximated using Gauss-Hermite quadrature with nodes generated from the mean $\boldsymbol{\mu}$ and covariance $\mathbf{P}$ of $\boldsymbol{x}$. The precise algorithm is given below.

Algorithm 1. [Generalized unscented transform] Let $p(\boldsymbol{x}) = \mathcal{N}(\boldsymbol{x};\boldsymbol{\mu},\mathbf{P})\mathcal{E}(\boldsymbol{x};\boldsymbol{\mu},\mathbf{P},\boldsymbol{\kappa}_{3:p})$, and $\boldsymbol{f} : \mathbb{R}^n \to \mathbb{R}^m$. The mean, covariance, and central moments up to order q,

$$\begin{aligned}
\boldsymbol{\mu}_f &= \mathrm{E}[\boldsymbol{f}(\boldsymbol{x})],\\
\mathbf{P}_f &= \mathrm{Cov}[\boldsymbol{f}(\boldsymbol{x})] = \mathrm{E}[\mathring{\boldsymbol{f}}\mathring{\boldsymbol{f}}^T], \quad \mathring{\boldsymbol{f}} = \boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{\mu}_f,\\
\mathring{\kappa}_f^{i_1\cdots i_r} &= \mathrm{E}[\mathring{f}^{i_1}\cdots\mathring{f}^{i_r}], \quad r = 3,\ldots q,
\end{aligned}$$

are computed as follows.

[‡]The big-O notation in (9) has the following interpretation. If $\boldsymbol{x}$ is a standardized sum of N independent identically distributed random variables, then $\kappa^{i,j,k} = \mathcal{O}(N^{-1/2})$, $\kappa^{i,j,k,\ell} = \mathcal{O}(N^{-1})$, …, $\kappa^{i_1,\ldots,i_p} = \mathcal{O}(N^{-(p-2)/2})$.

[§]The Genz-Keister nodes and weights can be downloaded at http://sparse-grids.de.

Table 1. The minimum order of the Gauss-Hermite quadrature rule (left) and the corresponding number of Gauss-Hermite-Genz-Keister quadrature nodes in $n = 6$ dimensions (right) required to compute the cumulants up to order q of the random variable $\boldsymbol{f}(\boldsymbol{x})$ if $\boldsymbol{x}$ is represented by an Edgeworth series of order p.

$p \backslash q$	1	2	3	4
2	1	2	3	4
3	4	5	6	7
4	7	8	9	10

Minimum quadrature order

$p \backslash q$	1	2	3	4
2	13	13	13	73
3	73	73	257	257
4	257	749	749	2021

Number of nodes ($n = 6$)

1. Obtain a set of n-dimensional GH-GK nodes and weights of order of at least $3p + q - 6$ (see Table 1) for the standardized Gaussian $N(\mathbf{0}, \mathbf{I})$ and denote them as $\boldsymbol{z}_\sigma$ and w_σ, respectively, for $\sigma = 1, \ldots, N$.
2. For $\sigma = 1, \ldots, N$, compute $\boldsymbol{x}_\sigma = \boldsymbol{\mu} + \mathbf{A}\boldsymbol{z}_\sigma$ (where $\mathbf{P} = \mathbf{A}\mathbf{A}^T$), $\boldsymbol{f}_\sigma = \boldsymbol{f}(\boldsymbol{x}_\sigma)$, and $\tilde{w}_\sigma = w_\sigma \mathcal{E}(\boldsymbol{x}_\sigma; \boldsymbol{\mu}, \mathbf{P}, \boldsymbol{\kappa}_{3:p})$.
3. Compute $\boldsymbol{\mu}_f = \sum_{\sigma=1}^N \tilde{w}_\sigma \boldsymbol{f}_\sigma$.
4. For $i = 1, \ldots, N$, compute $\mathring{\boldsymbol{f}} = \boldsymbol{f}_\sigma - \boldsymbol{\mu}_f$.
5. Compute $\mathbf{P}_f = \sum_{\sigma=1}^N \tilde{w}_\sigma \mathring{\boldsymbol{f}} \mathring{\boldsymbol{f}}^T$.
6. For $r = 3, \ldots, q$, compute $\mathring{\kappa}_f^{i_1 \cdots i_r} = \sum_{\sigma=1}^N \tilde{w}_\sigma \mathring{f}_\sigma^{i_1} \cdots \mathring{f}_\sigma^{i_r}$.
7. Optionally, compute the cumulants $\kappa_f^{i_1 \cdots i_r}$, for $r = 3, \ldots q$, using (7) or generalizations in [11].

Table 1 gives the minimum order of the Gauss-Hermite quadrature rule required to compute the cumulants of $\boldsymbol{f}(\boldsymbol{x})$ up to order q given that $\boldsymbol{x}$ is represented by an Edgeworth expansion of order p. The corresponding number of six-dimensional GH-GK nodes (sigma points) for these orders are tabulated on the right The minimum quadrature order is derived on the assumption that the function $\boldsymbol{f}$ in Algorithm 1 is linear. In practice, it is advisable to choose an order slightly higher than the minimum to account for the nonlinearities of $\boldsymbol{f}$.

The prediction and correction steps of the Edgeworth filter are applications of Algorithm 1. For example, starting with the dynamical model (3) and an Edgeworth representation of the prior state $\boldsymbol{x}_{k-1}$, the cumulants of the predicted state $\boldsymbol{x}_k$ can be obtained by applying Algorithm 1 with the solution flow $\boldsymbol{f}_{k-1}$. Full details will be presented in a future publication.

5. AN ONLINE METRIC FOR UNCERTAINTY CONSISTENCY

The metric we derive in this section is motivated from the prediction error term in the correction step of the Bayesian state estimator. Referring to Figure 2, in the implementation of the correction step in space surveillance, the posterior PDF $p(\boldsymbol{x}_k|\boldsymbol{Z}_k)$ is often computed by solving a *track-to-track fusion problem* by updating the prior density $p(\boldsymbol{x}_k|\boldsymbol{Z}_{k-1})$ with another track density $\tilde{p}(\boldsymbol{x}_k|\tilde{\boldsymbol{Z}}_{k-1})$. If the measurement histories $\boldsymbol{Z}_{k-1}$ and $\tilde{\boldsymbol{Z}}_{k-1}$ are assumed independent, then Bayes' rule in conjunction with the diffuse prior assumption [15] imply that

$$p(\boldsymbol{x}_k|\boldsymbol{Z}_k) = \frac{1}{\mathcal{PE}} p(\boldsymbol{x}_k|\boldsymbol{Z}_{k-1})\, \tilde{p}(\boldsymbol{x}_k|\tilde{\boldsymbol{Z}}_{k-1}), \quad \mathcal{PE} = \int p(\boldsymbol{x}_k|\boldsymbol{Z}_{k-1})\, \tilde{p}(\boldsymbol{x}_k|\tilde{\boldsymbol{Z}}_{k-1})\, \mathrm{d}\boldsymbol{x}_k, \tag{11}$$

where $\boldsymbol{Z}_k \equiv (\boldsymbol{Z}_{k-1}, \tilde{\boldsymbol{Z}}_{k-1})$.

The association problem provides a natural application of the prediction error $\mathcal{PE}$ in (11) as an online metric for uncertainty consistency. Referring to Figure 3, suppose we have two satellite states at times t_1 and t_2 (with $t_1 \leqslant t_2$ without loss of generality) represented by the PDFs $p_1(\boldsymbol{x}, t_1)$ and $p_2(\boldsymbol{x}, t_2)$, respectively,

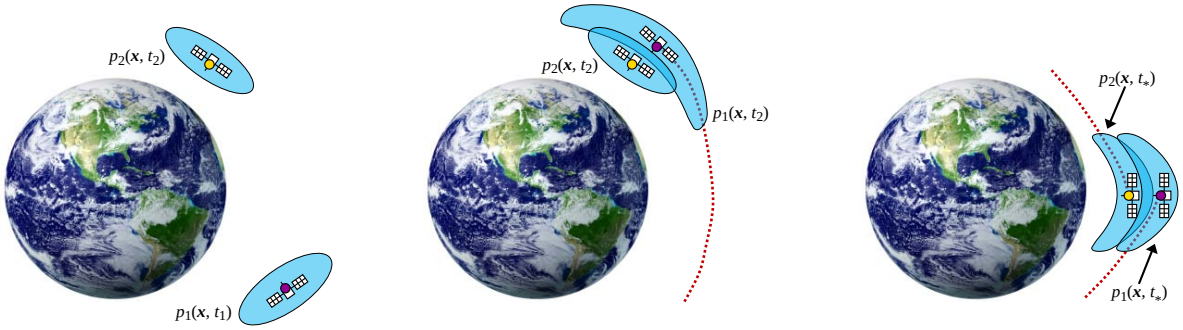

Figure 3. The evaluation of the prediction error (12) used to compute the probability of association between two states at times t_1 and t_2 (left) requires representations of their PDFs at a common time t_* (e.g., $t_* = t_2$ (middle) or $t_1 < t_* < t_2$ (right)). The prediction error is independent of the choice of t_*.

and we wish to assess whether the two objects associate. The likelihood score for computing the probability of association depends on the prediction error (11) which for this scenario is given by

$$\mathcal{PE}(t) = \int p_1(\boldsymbol{x}, t)\, p_2(\boldsymbol{x}, t)\, \mathrm{d}\boldsymbol{x}. \tag{12}$$

Evaluation of the integral (12) requires representations of both PDFs p_1 and p_2 at some common time t. Thus, one could either (i) propagate p_1 to time t_2 and evaluate $\mathcal{PE}(t_2)$ or (ii) propagate p_1 from time t_1 to some intermediate time t_*, then back propagate p_2 to t_*, and finally evaluate $\mathcal{PE}(t_*)$. In fact, the value of the prediction error is *independent of the choice of* t_* under the assumption that the dynamics are governed by a conservative deterministic model and the PDFs are propagated consistently (i.e., using the noiseless FPK equation (5)). This result is formally stated and proved below.

Proposition 1. *Let $p_1(\boldsymbol{x}, t)$ and $p_2(\boldsymbol{x}, t)$ denote the PDFs of two independent states at time t. Suppose the state $\boldsymbol{x}$ is governed by the conservative dynamical model*

$$\boldsymbol{x}'(t) = \boldsymbol{f}(\boldsymbol{x}(t), t), \quad \nabla_x^T \boldsymbol{f} = \boldsymbol{0},$$

and the time evolution of each PDFs satisfies the FPK equation (5). Then, the prediction error (12) is time-independent.

Proof. If p_1 and p_2 both satisfy the FPK equation, then the same can be said for their product. Indeed, by the product rule,

$$\frac{\partial}{\partial t}(p_1 p_2) = p_1 \frac{\partial p_2}{\partial t} + p_2 \frac{\partial p_1}{\partial t} = -p_1 \nabla_x^T (p_2 \boldsymbol{f}) - p_2 \nabla_x^T (p_1 \boldsymbol{f}) = -\nabla_x^T (p_1 p_2 \boldsymbol{f}).$$

Therefore,

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{PE} = \frac{\mathrm{d}}{\mathrm{d}t}\int p_1 p_2\, \mathrm{d}\boldsymbol{x} = \int \frac{\partial}{\partial t}(p_1 p_2)\, \mathrm{d}\boldsymbol{x} = -\int \nabla_x^T (p_1 p_2 \boldsymbol{f})\, \mathrm{d}\boldsymbol{x} = 0,$$

as follows from the Gauss Divergence theorem. □

The metric (12) provides a necessary condition for uncertainty consistency. If the evolution of the prediction error is not constant over time, then the propagated uncertainties of the two states are not consistent. The use of this metric for quantifying uncertainty consistency over long time gaps is illustrated in Section 6.

In closing, we remark that Proposition 1 holds for other "distance metrics" such as the Kullback-Leibler divergence (KLD),

$$\mathcal{KL}(t) = \int p_1(\boldsymbol{x}, t)\, \ln\left[\frac{p_1(\boldsymbol{x}, t)}{p_2(\boldsymbol{x}, t)}\right] \mathrm{d}\boldsymbol{x},$$

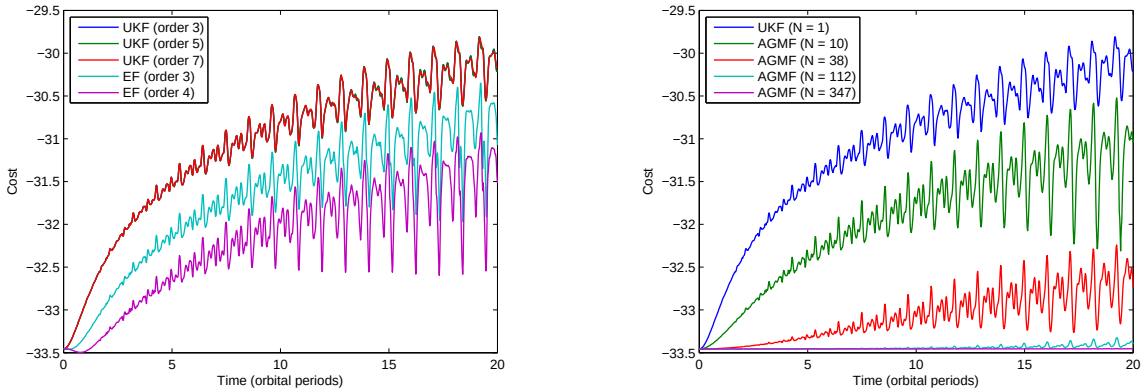

Figure 4. Evolution of the cost metric (negative log of the prediction error (12)) computed using various unscented Kalman filters, Edgeworth filters, and adaptive Gaussian mixture filters.

and can thus serve as an alternative metric for uncertainty consistency. While the KLD is used in sensor resource managment as a measure of mutual information gain [16], it is not what is commonly used in tracking for scoring likelihoods of association. A rigorous derivation of the likelihood ratio used to score the association of a track or measurement to an orbit is provided in [17]; it is precisely the prediction error $p(\boldsymbol{z}_k|\boldsymbol{Z}_{k-1})$ or (11) which appears in the ensuing expression and not the KLD.

6. RESULTS

In this section, we present results demonstrating how the Edgeworth filter (EF) and adaptive Gaussian mixture filter (AGMF) give improved uncertainty consistency over the standard unscented Kalman filter (UKF) in the problem of long term propagation of uncertainties. We then illustrate the impact of correct uncertainty management on the problems of association and anomaly detection. The simulation scenario considers two closely spaced objects (CSOs) in the low Earth regime whose initial uncertainties are Gaussian with respect to equinoctial orbital elements [18]. The initial states of the objects at time $t = 0$ are:

$$\begin{aligned} a_1 &= 6980\,\text{km}, \quad h_1 = k_1 = p_1 = q_1 = \ell_1 = 0 \quad \text{(Object 1)}, \\ a_2 &= 7020\,\text{km}, \quad h_2 = k_2 = p_2 = q_2 = \ell_2 = 0 \quad \text{(Object 2)}. \end{aligned}$$

The initial covariances of the objects are $\mathbf{P}_1 = \mathbf{P}_2 = \mathbf{A}\mathbf{A}^T$, where

$$\mathbf{A} = \text{diag}(20\,\text{km}, 10^{-3}, 10^{-3}, 10^{-3}, 10^{-3}, 0.01^\circ).$$

From (12), it follows that the prediction error between the two objects at $t = 0$ is $\mathcal{PE} = 3.3865 \cdot 10^{14}$ with corresponding "cost"

$$c = -\ln \mathcal{PE} = -33.456. \tag{13}$$

Fixing a particular filter, we propagate each of the two Gaussians for a total time of 32.38 hours (which is about 20 complete orbital periods) and evaluate the prediction error (and cost) at intermediate times. By Proposition 1, the cost must retain the constant value given by (13) in order for us to be able to assert that the state uncertainties are propagated consistently. Thus, any departure from a constant cost signals a degradation in the computed uncertainty.

In the left half of Figure 4, the evolution of the cost function is plotted using the UKF of order three, five, and seven. We observe that there is negligible difference in the computed costs. Therefore, the standard (third-order) UKF achieves *covariance* consistency. But, covariance consistency is only a prerequisite to the more general *uncertainty* consistency. Although the UKF correctly resolves the mean and covariance (first two cumulants) of the probability density, it fails to capture information about the cumulants of third-order and higher even in orbital element space. Because the cost diverges from the initial value, the true state uncertainties of the CSOs are evidently non-Gaussian and contain non-negligible higher-order cumulants which impact the probability of association (POA). After twenty orbital periods, this divergence is nearly

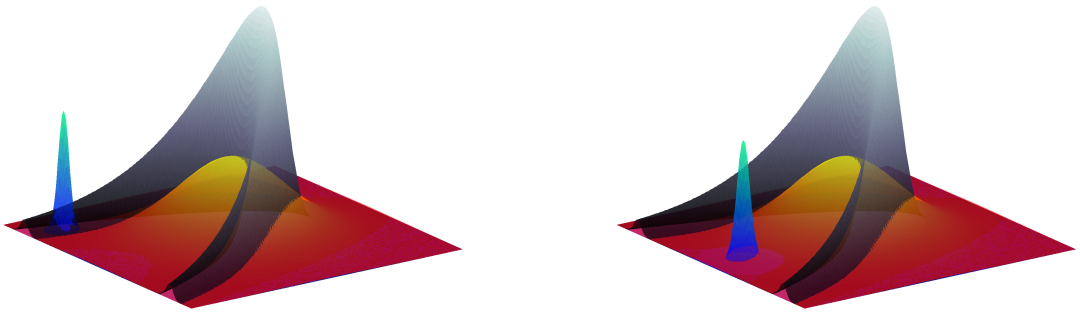

Figure 5. The PDF of Object 1 in the semi-major axis and mean longitude coordinates after four orbital periods computed using the UKF (reddish colormap) and an AGMF with $N = 347$ (grayscale colormap). The blueish Gaussians represent track updates which are subsequently scored against Object 1's PDF.

3.5 units of cost or about a factor of 30 in the prediction error. Consequently, the resulting POA could have an error upwards of a full order of magnitude.

The left half of Figure 4 also shows the improvement in uncertainty consistency when cumulants beyond the mean and covariance are represented and propagated using the EF. The middle curve is the cost obtained by propagating the mean, covariance, and third-order cumulants of each object using the EF formalism. The bottom curve takes into account all cumulants up to order four. By accounting for higher-order cumulants in the uncertainties, improved uncertainty consistency is achieved as exhibited by the mitigated divergence in the time evolution of the cost function.

The right half of Figure 4 shows that the AGMF can in principle maintain uncertainty consistency even over many orbital periods. Specifics of the AGMF used to produce these results will be provided in a future publication. We remark that the fourth-order EF performed comparably to the AGMF with $N = 10$ Gaussians in this scenario. However, in terms of computational cost, the former is significantly more expensive requiring a Gauss-Hermite quadrature method of order 11 entailing the propagation of 2021 sigma points (see Table 1). In contrast, each component covariance of the $N = 10$ Gaussian mixture was propagated using a third-order UKF with 13 sigma points and thus the mixture required only 130 total propagations.

Figure 5 depicts the non-Gaussianity of the uncertainty in Object 1 after only four orbital periods. A two-dimensional slice of its PDF along the semi-major axis and mean longitude coordinates is plotted. The true uncertainty (grayscale colormap) was computed using a high fidelity AGMF with $N = 1144$ and is contrasted with the Gaussian uncertainty (reddish colormap) obtained from the UKF.

The use of an inconsistent Gaussian uncertainty computed from a traditional Kalman-based filter can have dire consequences when doing track association or anomaly detection. For example, suppose we wish to compute the prediction error (or cost) associated with a new Gaussian track state (blueish colormap) located along the tail of Object 1's PDF, as illustrated in the left half of Figure 5. In this particular example, we obtain a cost of -30.406 from the UKF approximation (reddish colormap) versus a cost of -33.920 from the $N = 347$ AGMF (grayscale colormap). Therefore, the UKF estimates a larger cost leading to a lower POA and *possibly a failure to associate* and *possibly a misdetection of an anomaly.* Referring now to the right half of Figure 5, suppose the new track is located outside the support of Object 1's true uncertainty. In this particular case, the UKF produces a cost of -32.675 while the $N = 347$ AGMF yields a cost of $+3.5592$. Therefore, the UKF estimates a smaller cost leading to a higher POA and *possibly a misassociation* and *possibly a failure to detect an anomaly.*

7. CONCLUSIONS

In this paper, we have shown that the traditional covariance-weighted approach to track association and anomaly detection by way of the unscented Kalman filter (UKF) is inadequate in the sparse data environment of space surveillance. Instead the full probability density must be represented and propagated using a higher fidelity filter which effectively captures the higher order cumulants of the state uncertainty. Since covariances

propagate consistently in orbital element space with a standard UKF, an Edgeworth representation or a Gaussian sum approximation to the full density in orbital elements is a viable solution.

ACKNOWLEDGMENTS

This work was funded, in part, by the Air Force Office of Scientific Research through a research grant (FA9550-08-1-0419) and a Phase I STTR (FA9550-10-C-0087).

REFERENCES

[1] J. T. Horwood and A. B. Poore, "Adaptive Gaussian sum filters for space surveillance," *IEEE Transactions on Automatic Control*, 2010. (In Press).

[2] J. T. Horwood, N. D. Aragon, and A. B. Poore, "Adaptive Gaussian sum filters for space surveillance tracking," in *Proceedings of the AAS Kyle T. Alfriend Astrodynamics Symposium*, (Monterey, CA), May 2010. (In Press).

[3] D. Alspach and H. Sorenson, "Nonlinear Bayesian estimation using Gaussian sum approximations," *IEEE Transactions on Automatic Control*, vol. 17, no. 4, pp. 439–448, 1972.

[4] G. Terejanu, P. Singla, T. Singh, and P. D. Scott, "Uncertainty propagation for nonlinear dynamic systems using Gaussian mixture models," *Journal of Guidance, Control and Dynamics*, vol. 31, no. 6, pp. 1623–1633, 2008.

[5] K. DeMars, M. Jah, D. Giza, and T. Kelecy, "Orbit determination performance improvements for high area-to-mass ratio space object tracking using an adaptive Gaussian mixtures estimation algorithm," in *21st International Symposium on Space Flight Dynamics*, (Toulouse, France), 2009.

[6] D. R. Giza, P. Singla, J. L. Crassidis, R. Linares, P. J. Cefola, and K. Hill, "Entropy-based space object data association using an adaptive Gaussian sum filter," in *AIAA Guidance, Navigation, and Control Conference*, (Toronto, Canada), August 2010.

[7] H. Singer, "Generalized Gauss-Hermite filtering." (Preprint) FernUniversität in Hagen, 2006.

[8] H. Singer, "Generalized Gauss-Hermite filtering for multivariate diffusion processes." (Preprint) FernUniversität in Hagen, 2006.

[9] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Boston: Artech House, 2004.

[10] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. New York: Dover, 1970.

[11] P. McCullagh, *Tensor Methods in Statistics*. London: Chapman and Hall, 1987.

[12] S. J. Julier, J. K. Uhlmann, and H. F. Durant-Whyte, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Transactions on Automatic Control*, vol. 55, pp. 477–482, 2000.

[13] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol. 92, pp. 401–422, 2004.

[14] A. Genz and B. D. Keister, "Fully symmetric interpolatory rules for multiple integrals over infinite regions with Gaussian weight," *Journal of Computational and Applied Mathematics*, vol. 71, pp. 299–309, 1996.

[15] Y. Bar-Shalom, X.-R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*. New York: John Wiley & Sons, 2001.

[16] A. O. Hero, D. Castañón, D. Cochran, and K. Kastella, *Foundations and Applications of Sensor Management*. New York: Springer, 2008.

[17] A. B. Poore, "Multidimensional assignment formulation of data association problems arising from multitarget tracking and multisensor data fusion," *Computational Optimization and Applications*, vol. 3, pp. 27–57, 1994.

[18] R. A. Broucke and P. J. Cefola, "On the equinoctial orbit elements," *Celestial Mechanics*, vol. 5, pp. 303–310, 1972.