(Preprint) AAS 26-775

MULTI-HYPOTHESIS SENSOR FUSION WITH AN EPISTEMIC KERNEL MIXTURE FILTER

Ben Parsonage*, Pietro De Marchi*, Massimiliano Vasile†

Resident Space Objects (RSOs) may be observed using several alternative measurements methods, such as Light Curves, radar etc. Each method is subject to various sources of uncertainty, such as atmosphere attenuation/degradation effects, temporal/systematic sensor noise, optical aberrations and surface defects, to name a few. Furthermore, the processing of raw measurements to obtain state estimations can result in several possible solutions given inherent ambiguities in the received data.

This paper demonstrates that this multiplicity can be treated as an epistemic uncertainty. For cases in which sequential batches of measurements are available from different types of sensor, we present a robust sequential filter that can handle epistemic uncertainty in the received measurements and the model of the dynamics. This filter can represent an uncertain distribution of arbitrary complexity as a Epistemic Kernel Mixture Model (EKMM) and can propagate a full set of distribution functions by treating each kernel individually. The kernel weights of the EKMM are interval valued, allowing the filter to represent multiple non-Gaussian probability distributions simultaneously. Robust Bayesian process is used to update both the weights and their bounds in order to build a family of posterior distributions from each new measurement. Between observations, kernels are automatically adapted in number and size to optimise the propagation of the prior. From indirect inference through propagation, the filter can also be used to estimate unknown model parameters and update the model distribution accounting for epistemic uncertainty.

Results are presented for an attitude estimation problem in which the measurement ambiguities and resulting multiplicity in possible attitude states from several available sensor types is modelled using the proposed filtering approach. It is shown that the continuous propagation of epistemic distributions allows the sequential inference of unobserved model parameters, meaning that the target state can be estimated with minimal assumptions or prior knowledge regarding the object parameters. Results are validated against both a standard Unscented Kalman Filter (UKF) and Monte Carlo sampling.

INTRODUCTION

A key activity in Space Situational Awareness (SSA) is the robust motion estimation of Resident Space Objects (RSOs) from noisy/uncertain observations, thus enabling status monitoring, precise long-term orbit propagation, Conjunction Assessment Risk Analysis (CARA) and Active Debris Removal (ADR) missions, amongst other applications.

In particular, the attitude motion of RSOs may be observed using one or more of several alternative measurements methods, such as spectral light curves, laser ranging and radar. Each method is subject to various sources of uncertainty, including atmosphere attenuation and degradation effects, temporal and systematic sensor noise, optical aberrations and surface defects, to name a few. Observations are generally sparse and can be received from heterogenous sources with unknown quality. Furthermore, the processing of raw observations to obtain state estimations is hindered by ambiguities inherent to the target object itself. That is, many combinations of geometric shape, orientation, surface materials and inertia properties may correspond

*Research Associate, Department of Mechanical and Aerospace Engineering, University of Strathclyde, United Kingdom, G1 1XJ.
†Professor, Department of Mechanical and Aerospace Engineering, University of Strathclyde, United Kingdom, G1 1XJ.

to the same measured data.[1] Last but not least, these physical properties of the target object, and indeed of the sensor(s) themselves, might be only partially known.

These forms of uncertainty are epistemic in nature as they represent a fundamental lack of knowledge, and thus must be correctly treated. Building on previous work by the authors on the treatment of epistemic uncertainty in conjunction analysis and deep space navigation,[2,3] we propose a novel sequential filter that estimates the bounding region of admissible states under epistemic uncertainty. We remove the assumption of Gaussianity in modelling the belief state of the RSO and propagate a family of admissible distributions that represent the future belief state prior to the next observation. We also introduce the fundamental assumption that every observed state is an individual hypothesis taken from an unknown underlying distribution, similarly to that proposed in Sanchez et al.[4] Hence, when multiple hypotheses are possible we include them in a robust inference process that accounts for the multiplicity of possible belief states.

This paper first presents a description of the fundamental construction of the filter, together with a simplified demonstration considering homogeneous measurements. Next, we present a more complex scenario in which multiple hypotheses are received from several heterogenous sensors providing a series of sparse observations of the true motion of a given RSO.

METHOD

The proposed sequential filter represents an uncertain distribution of arbitrary complexity as a weighted combination of smaller probability functions (or *kernels*). Kernel weights are interval valued, allowing the representation of multiple non-Gaussian probability distributions simultaneously. We refer to this approach as an Epistemic Kernel Mixture Filter (EKMF)*. The Probability Density Function (PDF) of an Epistemic Kernel Mixture Model (EKMM) can be written as:

$$p(\mathbf{x}) = \sum_k \mathbf{w}_k \int_X \mathcal{N}(\mathbf{x}, \mathbf{p}_k)\mathrm{dx} \tag{1}$$

where $\mathcal{N}(\mathbf{x}, \mathbf{p})$ represents a chosen multivariate distribution function with parameters $\mathbf{p}$ evaluated at the state vector $\mathbf{x}$, and the kernel weights $\mathbf{w}_k$ are bounded as:

$$\mathbf{l} \leq \mathbf{w} \leq \mathbf{u} \tag{2}$$

where $\mathbf{l} = [l_1, \ldots, l_k]$ and $\mathbf{u} = [u_1, \ldots, u_k]$ are the lower and upper weights bounds respectively for each kernel. The expected value of some function $f(\mathbf{x})$ represented using an EKMM can thus be written as:

$$\mathrm{E}(f) = \sum_k w_k c_k = \mathbf{c}^T \mathbf{w} \tag{3}$$

where $c_k = \int_X f(\mathbf{x})\mathcal{N}(\mathbf{x}, \mathbf{p}_k)\mathrm{dx}$ is defined as the expected value of $f(\mathbf{x})$ if $\mathbf{x}$ follows the probability distribution given by the k-th kernel. Uncertainty is then quantifiable using the lower expectation of $f(\mathbf{x})$, given as:

$$\begin{aligned} \underline{\mathrm{E}}(f) = &\min_{\mathbf{w}} \mathbf{c}^T \mathbf{w} \\ &s.t.\ \mathbf{l} \leq \mathbf{w} \leq \mathbf{u} \\ &\mathbf{w} \geq 0 \\ &\sum_k w_k = 1 \end{aligned} \tag{4}$$

*Where kernels are specified as Gaussian distributions, this approach is similar to navigation filters that employ the more common Gaussian Mixture Model (GMM), though still with the epistemic modifications described herein. Given the ability in this case to specify alternative kernel distributions, we maintain the distinction for clarity.

Propagation

If the kernel distribution functions are chosen as standard Gaussians such that $\mathbf{p}_k = [\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k]$, where $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$ are the means and covariances of the k-th kernel respectively, the EKMF can propagate the complete set of distribution functions using an Unscented Transform (UT) applied consecutively to each kernel.[5] A UT is used to approximate the distribution of a variable $\mathbf{y}$ resulting from transforming a Gaussian variable $\mathbf{x}$ with distribution $\mathcal{N}(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ using a state transition function $\mathbf{x}^{(f)} = f(\mathbf{x}^{(0)}, t_0, t_f)$, where the final state $\mathbf{x}^{(f)}$ at time t_f is obtained by propagating the initial state $\mathbf{x}^{(0)}$ from time t_0. The UT computes the function f at the sigma-points $\mathcal{X}_i$ of each of N kernels representing the initial distribution:

$$\mathcal{X}_0 = \boldsymbol{\mu} \tag{5}$$

$$\mathcal{X}_i = \boldsymbol{\mu} + \gamma \left(\sqrt{\Sigma}\right)_i \text{ for } i \in \{1, \ldots, N\} \tag{6}$$

$$\mathcal{X}_{i+N} = \boldsymbol{\mu} - \gamma \left(\sqrt{\Sigma}\right)_i \text{ for } i \in \{1, \ldots, N\}\,, \tag{7}$$

where $\left(\sqrt{\Sigma}\right)_i$ is the i-th column of a decomposition of $\sqrt{\Sigma}$ (defined such that $\sqrt{\Sigma}\sqrt{\Sigma}^T = \Sigma$). A *square-root* approach is implemented, meaning that the covariances of each kernel are represented by their Cholesky factor, which offers numerical advantages. Where possible, all operations are performed on these Cholesky factors.[6] The scaling parameter γ is defined by the UT parameters α, β and κ.[5] There are multiple commonly used choices for parameters α, β and κ, with none of them allowing the UT to compute the exact mean and covariance matrix for general multi-variate non-linear transformations, even those given as a second order polynomial.[7] In this work, values are taken as $\alpha = 1$, $\beta = 0$ and $\kappa = 0$, corresponding to a cubature filter. These values have the advantage of ensuring positive weights, which prevents matrices from becoming non-positive-definite and also minimises the use of Cholesky downdating (which becomes necessary when updating the uncertainty with further measurements/observations). This then avoids numerical issues with non-positive-definite matrices occurring due to floating point errors.

The sigma-points $\mathcal{X}_i$ are propagated using the state transition function such that:

$$\mathcal{Y}_i = f(\mathcal{X}_i, t_0, t_f) \tag{8}$$

The mean $\boldsymbol{\mu}^{\mathbf{y}}$ and covariance $\boldsymbol{\Sigma}^{\mathbf{y}}$ of $\mathcal{Y}$ are then estimated as:

$$\boldsymbol{\mu}^{\mathbf{y}} \approx \sum_{i=0}^{2N} W_i^{(m)} \mathcal{Y}_i \tag{9}$$

$$\Sigma^{\mathbf{y}} \approx \sum_{i=0}^{2N} W_i^{(c)} \left(\mathcal{Y}_i - \hat{\boldsymbol{\mu}}_{\mathbf{y}}\right) \left(\mathcal{Y}_i - \hat{\boldsymbol{\mu}}_{\mathbf{y}}\right)^T \tag{10}$$

where the weights $W_i^{(m)}$ and $W_i^{(c)}$ are defined as a function of the tunable parameters α, β, and κ:[5]

$$W_0^{(m)} = \kappa/(N + \kappa) \tag{11}$$

$$W_0^{(c)} = \kappa/(N + \kappa) + (1 - \alpha^2 + \beta) \tag{12}$$

$$W_i^{(m)} = W_i^{(c)} = 1/\left(2(N + \kappa)\right) \tag{13}$$

The above formulae can be used to propagate the state uncertainty between each measurement timepoint. Each individual kernel of the complete distribution is propagated separately as if it were an independent Unscented Kalman Filter (UKF). The individual kernel weights w_k are not changed at this stage.

To represent uncertainty in the dynamics model, and to avoid the issue of what is referred to as a "smug" filter,[8] it is also possible to include noise in the state dynamics (often termed "process noise"). In this case, the state transition function can take an additional parameter $\mathbf{n}$, becoming:

$$\mathbf{x}^{(f)} = f(\mathbf{x}^{(0)}, t_0, t_f, \mathbf{n}) \tag{14}$$

The UT then proceeds in the same way using an augmented input variable that includes the new term $\mathbf{n}$. Note that if $\mathbf{n}$ is also a Mixture Model (with N_n kernels), the combination with $\mathbf{x}^{(0)}$ (a mixture of N_x kernels) results in mixture $\mathbf{x}^{(f)}$ with $N_n \times N_x$ kernels, potentially resulting in quality degradation, which will be discussed later.[9]

Adaptation

When a Gaussian kernel is propagated through a non-linear function using a linear approximation, a loss of accuracy is incurred. This loss can be expressed in terms of the KL-divergence, itself a measure of the difference between a model distribution and the 'true' distribution. The KL-divergence can be represented by evaluating the non-linearity metric D_{KL} for each individual kernel of a Mixture Model using a second order approximation of the non-linear function and of the KL-divergence integral itself.[10]

Splitting The kernels with the highest non-linearity metric $D_{KL}^{(i)}$ are split along the direction with highest $\hat{\eta}_i$, where $\hat{\eta}_i$ is a non-dimensional measure of how much the gradient of the non-linear propagation function f varies within a certain range of the mean of input $\mathbf{x}$, see the work by Faubel et al.[11] Splitting a Gaussian kernel along a specified direction means replacing it with another Gaussian Mixture, in which each kernel has a smaller variance along that direction such that the original kernel is well approximated. The new kernels obtained from splitting a kernel $\mathcal{N}(\mathbf{x}; \boldsymbol{\mu}_k, \Sigma_k)$ are given as:

$$\mathcal{N}(\mathbf{x}; \boldsymbol{\mu}_{k,i}, \Sigma_{k,i}) \tag{15}$$

with weights $w_{k,i}$, means $\boldsymbol{\mu}_{k,i}$ and covariances $\Sigma_{k,i}$ defined by:

$$\begin{aligned} w_{k,i} &= \alpha_i w_k \\ \boldsymbol{\mu}_{k,i} &= \sqrt{\Sigma_k}\beta_i \mathbf{d}_s + \boldsymbol{\mu}_k \\ \Sigma_{k,i} &= \sqrt{\Sigma_k} S_s \sqrt{\Sigma_k}^T \end{aligned} \tag{16}$$

The vector $\mathbf{d}_s$ is zero in every element except the s-th, which represents the splitting direction and is equal to one. The matrix S_s is equal to the identity matrix except at the s-th diagonal element, which is equal to the variance ratio σ_s^2. This represents the fraction of the variance along the splitting direction of the kernels after splitting, relative to the pre-splitting values. The values of α_i and β_i are pre-computed, and obtained from a publicly available library of optimised splits of standard Gaussians.[12] These minimise the L_1 loss between a 1D standard Gaussian and the split with weights and means given by α_i and β_i respectively.

For the split step, as seen in (16), the new weights are given as the product of the previous weights and some pre-computed weights. If $\mathbf{w}_0$ are the weights before splitting kernel i into kernels $N, \ldots, N-1+N_S$, and $\mathbf{w}_f$ are the weights of the new Mixture, then:

$$\mathbf{w}_f = A_{S,i}\mathbf{w}_0 \tag{17}$$

where

$$A_{S,i} = \begin{array}{c} \\ 1,\cdots,i-1 \\ i,\cdots,N-1 \\ N,\cdots,N-1+N_S \end{array} \begin{array}{c} \begin{array}{ccc} 1,\cdots,i-1 & i & i+1,\cdots,N \end{array} \\ \begin{pmatrix} I & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & I \\ \mathbf{0} & \boldsymbol{\alpha} & \mathbf{0} \end{pmatrix} \end{array} \tag{18}$$

Merging If the process/measurement noise is also modelled using a Mixture Model, the number of kernels in the complete Mixture describing the system state can grow exponentially. Furthermore, some of the kernels produced by the splitting process described above can reduce to a practically negligible size. In these cases, it can be more appropriate to merge such kernels with others in their vicinity. A method for merging kernels and subsequently estimating the incurred KL-divergence is proposed by Runnalls[13] and also described in Wills et al.[9]

The merging of the i^{th} and j^{th} kernels results in a single kernel ij with the following parameters:

$$w_{ij} = w_i + w_j \tag{19}$$
$$\mu_{ij} = w_{i|ij}\mu_i + w_{j|ij}\mu_j \tag{20}$$
$$\Sigma_{ij} = w_{i|ij}\Sigma_i + w_{j|ij}\Sigma_j \tag{21}$$
$$+ w_{i|ij}w_{j|ij}\left(\mu_i - \mu_j\right)\left(\mu_i - \mu_j\right)^T \tag{22}$$
$$w_{i|ij} = \frac{w_i}{w_i + w_j} \tag{23}$$
$$w_{j|ij} = \frac{w_j}{w_i + w_j} \tag{24}$$

The cost of merging the i^{th} and j^{th} kernels, B_{ij}, provides an estimate of the KL-divergence between the weighted sum of these kernels and the merged kernel:[13]

$$B_{ij} = \frac{1}{2}\left[w_{ij}\log|\Sigma_{ij}| - w_i\log|\Sigma_i| - w_j\log|\Sigma_j|\right] \tag{25}$$

where $|\Sigma|$ denotes the determinant of Σ. This cost function ignores, however, the increase in non-linearity induced errors resulting from the growth of the covariance ellipsoid. to compensate for the increase in non-linearity which results from the increase in size of the kernel, the cost metric B_{ij} can be augmented with the addition of the following term:

$$C_{ij} = \frac{7}{8}\left(w_{ij}\sum_k \max\left\{\left(\varphi_k^{(i\to ij)}\eta_k^{(i)}\right)^2, \left(\varphi_k^{(j\to ij)}\eta_k^{(j)}\right)^2\right\} - w_i\sum_k\left(\eta_k^{(i)}\right)^2 - w_j\sum_k\left(\eta_k^{(j)}\right)^2\right) \tag{26}$$

From (19), when multiple kernels are merged, the weight of the merged kernel is the sum of the weights of the kernels being merged. As such, the operation remains linear. If kernels i and j are merged into kernel $N+1$, the kernel weight following the merge would be:

$$\mathbf{w}_f = A_{M,i,j}\mathbf{w}_0 \tag{27}$$

where

$$A_{M,i,j} = \begin{array}{c} \\ 1,\cdots,i-1 \\ i,\cdots,j-2 \\ j-1,\cdots,N-2 \\ N-1 \end{array} \begin{array}{c} \begin{array}{ccccc} 1,\cdots,i-1 & i & i+1,\cdots,j-1 & j & j+1,\cdots,N \end{array} \\ \begin{pmatrix} I & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & I & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & I \\ \mathbf{0} & 1 & \mathbf{0} & 1 & \mathbf{0} \end{pmatrix} \end{array} \tag{28}$$

Ideally, the set of kernel splits and merges that minimises the estimated KL-divergence would be chosen. However, this is a difficult optimisation problem to solve, due in part to its combinatorial nature. Instead, a heuristic approach can be taken, which proceeds by propagating the complete Mixture via UT, storing the sigma-points $\mathcal{X}$ and $\mathcal{Y}$ and computing the non-linearity metrics accordingly. The reduction in KL-divergence obtained by splitting the kernel with highest non-linearity is calculated. If the resulting number of kernels is below the maximum number set by the user, the new Mixture is accepted. Otherwise, if N_e is the excess number of kernels, the N_e lowest cost merges are estimated. If their total cost is lower than the reduction in KL-divergence resulting from splitting, the new Mixture is accepted and another iteration of splitting and merging is performed. If, on the contrary, the merging costs have a higher total than the gain from splitting, the algorithm stops and the original Mixture is returned.

This demonstrates one of the primary benefits of including a KL-divergence based metric. That is, to be able to directly compare the decrease in KL-divergence resulting from splitting some kernel with the increase

resulting from merging others. Thus the total number of kernels, and therefore also the cost of the algorithm, can be held constant.

Once all of the merge and split steps have been performed, the final kernel weights $\mathbf{w}_a$ are written as a function of the initial weights $\mathbf{w}$:

$$\mathbf{w}_f = A\mathbf{w}_0 \tag{29}$$

where A is given by the product of all the split and merge matrices. To obtain bounds on the weights of the new kernels, such that the lower expectation can be solved using (4), the following expression must be solved:

$$\begin{aligned} l_{f,k} &= \min_{\mathbf{l}\leq\mathbf{w}\leq\mathbf{u}} (A\mathbf{w})_k \\ u_{f,k} &= \max_{\mathbf{l}\leq\mathbf{w}\leq\mathbf{u}} (A\mathbf{w})_k \end{aligned} \tag{30}$$

Solving (30) gives a conservative estimate of the new bounds, since weight vectors respecting $\mathbf{l}_a \leq \mathbf{w}_a \leq \mathbf{u}_a$ may not correspond to any value of $A\mathbf{w}$ with $\mathbf{l} \leq \mathbf{w} \leq \mathbf{u}$.

Measurement Update

Robust Bayesian process is used to update the weights of the EKMF and their bounds in order to construct a *family* of posterior distributions. When a new measurement is obtained, each kernel of the EKMF (including the kernel weights) is updated the same way as a standard UKF. Let $\mathbf{y} = h(\mathbf{x}, \mathbf{n})$ be the measurement function that obtains a measurement $\mathbf{y}$ of the underlying state $\mathbf{x}$. Measurement noise is represented by $\mathbf{n}$. The state uncertainty is propagated through function h, obtaining a Mixture Model representing the distribution of $\mathbf{y}$. Continuing to assume Gaussian kernels, each kernel of that distribution is written as $\mathcal{N}(\mathbf{y}, \boldsymbol{\mu}_k^y, \Sigma_k^y)$. The covariance matrix between the variables $\mathbf{x}$ and $\mathbf{y}$ is obtained following a similar approach as (10):

$$\Sigma^{\mathbf{x},\mathbf{y}} \approx \sum_{i=0}^{2N} W_i^{(c)} \left(\mathcal{X}_i - \hat{\boldsymbol{\mu}}_{\mathbf{x}}\right) \left(\mathcal{Y}_i - \hat{\boldsymbol{\mu}}_{\mathbf{y}}\right)^T \tag{31}$$

where $\mathcal{X}_i$ are the sigma-points of the state variable at the time of the observation, and $\mathcal{Y}_i$ are the values of the observation $\mathbf{y}$ for those same points. Each kernel in the distribution of the state, $\mathcal{N}(\mathbf{x}, \boldsymbol{\mu}_k, \Sigma_k)$, is updated according to the following formulae for Bayesian estimation:[14]

$$\mathcal{K} = \Sigma^{\mathbf{x},\mathbf{y}} \left(\Sigma_k^y\right)^{-1} \tag{32}$$

$$\hat{\boldsymbol{\mu}}_k = \boldsymbol{\mu}_k + \mathcal{K}\left(\mathbf{y}_m - \boldsymbol{\mu}_k^y\right) \tag{33}$$

$$\hat{\Sigma}_k = \Sigma_k \tag{34}$$

where $\mathbf{y}_m$ is the value of the actual measurement and $\hat{\boldsymbol{\mu}}_k$ and $\hat{\Sigma}_k$ are the updated values of the k-th kernel's mean and covariance. The updated weights are then calculated as:

$$\bar{w}_k^{t+1} = w_k^t \mathcal{N}(\mathbf{y}, \boldsymbol{\mu}_k^y, \Sigma_k^y) = \gamma_k w_k^t \tag{35}$$

followed by a normalisation:

$$w_k^{t+1} = \frac{\bar{w}_k^{t+1}}{\sum_j \bar{w}_j^{t+1}} \tag{36}$$

For interval valued weights $w_k^t \in [\alpha_k^t, \beta_k^t]$, this update affects the weights in a non-linear way. The goal thus becomes the determination of α_k^{t+1} and β_k^{t+1} such that $w_k^{t+1} \in [\alpha_k^{t+1}, \beta_k^{t+1}]$. The measurement update of the weights can be rewritten* as:

$$w_k^{t+1} = \frac{\gamma_k w_k^t}{\sum_j \gamma_j w_j^t} = \frac{\gamma_k w_k^t}{\gamma_k w_k^t + \sum_{j\neq k} \gamma_j w_j^t} = \frac{1}{1 + \frac{1}{\gamma_k w_k^t} \sum_{j\neq k} \gamma_j w_j^t} \tag{37}$$

*Of note, these calculations are implemented using logarithms, where γ_k is replaced with $c \log \gamma_k$, with c chosen to avoid underflow errors. For simplicity in notation, this modification is omitted.

Next, it is assumed that w_k^{t+1} is minimised/maximised by assigning $w_k^t = \alpha_k^t$ or $w_k^t = \beta_k^t$ respectively. Defining:

$$A_k^t = \min \sum_{j \neq k} \gamma_j w_j^t \quad s.t. \sum_{j \neq k} w_j^t = 1 - \beta_k^t, \quad \alpha_j^t \leq w_j^t \leq \beta_j^t \tag{38}$$

$$B_k^t = \max \sum_{j \neq k} \gamma_j w_j^t \quad s.t. \sum_{j \neq k} w_j^t = 1 - \alpha_k^t, \quad \alpha_j^t \leq w_j^t \leq \beta_j^t \tag{39}$$

where A_k^t and B_k^t can both be found using linear programming. Finally, the optimal* new bounds can be calculated by substituting into (37):

$$\alpha_k^{t+1} = \frac{\gamma_k \alpha_k^t}{\gamma_k \alpha_k^t + B_k^t} \tag{40}$$

$$\beta_k^{t+1} = \frac{\gamma_k \beta_k^t}{\gamma_k \beta_k^t + A_k^t} \tag{41}$$

Importantly, if $\gamma_k =$ constant, the bounds are unchanged.

Finally, from indirect inference through propagation, the EKMF can also be used to estimate unknown model parameters in the form of upper and lower confidence bounds. In this work, this is facilitated by including additional state variables in the problem definition. As for the other states, these new variables are also represented by the interval-valued kernel weights of the EKMF, but with a propagation function equal to identity. Mathematically, suppose the state of a system $\mathbf{x}$ depends on some parameters $\mathbf{p}$ through the following state transition function:

$$\mathbf{x}^{(f)} = F(\mathbf{x}^{(0)}, t_0, t_f, \mathbf{p}) \tag{42}$$

An augmented state including parameters $\mathbf{p}$ may therefore be written as:

$$\tilde{\mathbf{x}} = \begin{bmatrix} \mathbf{x} \\ \mathbf{p} \end{bmatrix} \tag{43}$$

with the corresponding state transition function:

$$\tilde{F}(\tilde{\mathbf{x}}^{(0)}, t_0, t_f) = \begin{bmatrix} F(\mathbf{x}^{(0)}, t_0, t_f, \mathbf{p}) \\ \mathbf{p} \end{bmatrix} \tag{44}$$

This allows the observed state variables to become correlated with the uncertain model parameters such that when the filter is updated with measurements of the state, the uncertain parameters are also updated despite not being observed directly themselves.

Observation Multiplicity

Equations (31)-(41) are designed to update the current state approximation using a single observation y_t at time t. However, in many scenarios, there may be several K equally probable state solutions resulting from a single observation, such that $\mathbf{y} = \{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_k\}$. This solution multiplicity may be due to noise (clutter) or measurement ambiguity.[1] Some authors have previously tackled this problem by simply accepting only the most likely observation at a given time, or by combining multiple solutions through the use of a weighted sum approach, such as for the Probabilistic Data Association Filter (PDAF).[15,16]

In Faubel et al,[17] the authors offer an approach that first splits the filter at time t into K filters, then assigns each of the resulting filters to one of the K observations, and finally updates each filter according to a conditioning step. This result in a Gaussian Mixture filtering density, which the authors handle by maintaining

*'Optimal' in the sense that they will include weights that could not have been obtained by the initial set of weights, and that they represent the bounds of the tightest hyperrectangle including all the updated weights.

a set of filters in parallel. To avoid the total number of Gaussian components growing exponentially in time, the authors implement a kernel merging step after each iteration.

This work proposes a similar approach, taking advantage of the unique properties of the EKMF. This approach allows a multiplicity of solutions to be treated as a set of samples drawn from and unknown distribution. By then propagating the resulting set of distributions, a posterior density that covers the solution multiplicity can be easily computed. The process for this is as follows:

1. On receiving one or more observations at time t such that the measured state is composed of multiple hypotheses, $\mathbf{y}_t = \{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_k\}_t$, update the current EKMM using a robust Bayesian process with each of the K hypotheses individually, creating K parallel distributions.
2. Merge the resulting set of K distributions into a single EKMM. The resulting Mixture contains the contribution of each hypothesis to the prior with respect to the non-Gaussian measurement distribution.
3. Apply an adaptive kernel merging step to prevent the exponential growth in total number of Mixture components. This operates by minimising the KL-divergence for a reduction towards a user-specified maximum number of kernels.
4. The single reduced EKMM incorporating all hypotheses can then be propagated forwards in time to the next observation.

This process ensures that non-Gaussian/multi-modal behaviour of multi-hypotheses can be preserved without incurring exponential growth, simply by utilising the epistemic definition of the proposed filtering approach. That is, treating 'families' of distributions. In the case of poor measurements, the confidence region (defined by the upper and lower expectation of the EKMM) is likely to remain wide and a larger number of hypotheses might survive, though this then therefore maintains a correct quantification of uncertainty.

Measurement Inconsistency

A final complication that must be dealt with is inconsistency in the received multiplicity across sequential observations. To explain, we have established in the preceding discussions that a single observation may produce a multiplicity of equally probable states representing some underlying ambiguities. However, in practical scenarios, the extent of this multiplicity may not necessarily be fixed across successive observations. That is, one measurement may produce a certain portion of the ambiguity, the next may produce another, and so on. This issue can be broadly described by two primary high-level behaviours:

1. Multiple hypotheses are received for a single observation, where the ambiguity represented by the complete hypotheses set is variable between observations. This can be the case, for example, where significant noise in the raw measurement data disrupts the state estimation algorithm.
2. A single (or clustered) hypothesis is returned for a single observation, with ambiguity represented across successive observations. This can be the case, for example, when inherent shape ambiguity results in several distinct state solutions matching the received measurement data.

The implication is that in such cases, any received measurement, or set of measurements, can only be considered an 'incomplete' representation of the true uncertainty of the target state. A practical consequence is that for scenarios in which a state estimation filter has been initialised on the first set of received measurements, but receives a new set of measurements representing an as-yet unobserved hypothesis, the new set is thus completely incompatible with the original distribution. The Bayesian inference step is then likely to drastically reduce support given to the original distribution in favour of the new measurements, when in fact both states are representative of the underlying measurement uncertainty.

To address this behaviour, a modification to the Bayesian update process is implemented to allow the representation and progressive collection of incomplete multiplicities. By this approach, the EKMF can

gradually build a stable multi-modal distribution representing ambiguous states compatible with all received observations taking into account both the solution multiplicity and the inherent observational uncertainty. This modification operates as follows:

1. Create a separate EKMM representing the new measurements. Each individual of m hypotheses is represented by a single kernel, with interval-valued weights $w_m \in [0, 1]$. This allows any unique hypothesis to be treated with zero or full support, which maintains the ambiguity.
2. Merge the EKMM from Step 1 with the current prior, creating a new EKMM containing kernels that cover the new measurements. Importantly, this includes the fact that the new measurement weights can be reduced to 0, negating their effect on the prior.
3. Perform the multi-hypothesis Bayesian inference step on the combined EKMM according to the procedure described in the previous sections. This includes expanding the combined EKMM with the measurement uncertainty associated to each hypothesis in turn prior to the update of the kernel means, covariances, weights and weights bounds. The resulting m parallel EKMMs are merged into a single distribution, reduced to a specified maximum number of kernels.
4. Export the final EKMM representing an updated distribution that supports the previous and new measurements.

This modification in essence allows the weights for all new kernels to vary between 0 and 1. This implies that the kernels associated with both the previous and new measurements can be weighted to fully support each ambiguous state separately, but not together. This then by definition maintains both the measurement ambiguity and the inherent measurement uncertainty by virtue of the interval valued weights. The EKMF can therefore adaptively widen, contract and/or translate the confidence bounds to represent a multi-modal distribution containing the new measurement as well as the propagated measurements received up until that point. Whilst constituting a conservative estimate of the true state, the resulting bounds are nonetheless an accurate representation of the uncertainty introduced by ambiguous observations.

DEMONSTRATION

To demonstrate the proposed EKMF, we first present an example in which the received measurements are both consistent and unambiguous, with measurement uncertainty represented by Gaussian noise. This example involves the attitude estimation of an uncontrolled satellite in Low Earth Orbit (LEO). The complete attitude state is described by a set of Modified Rodrigues Parameterss (MRPs) (r_1, r_2, r_3) describing the satellite's orientation with respect to an Earth Centred Inertial (ECI) reference frame, and the three angular velocity components (p, q, r) in the Body Coordinate System (BCS). Figure 1 shows the time history of the confidence bounds associated to each state variable as estimated by the EKMF. From an assumed Gaussian initialisation, the current estimation is refined via sequential noisy measurements processed through the EKMF framework as described throughout the previous sections. Synthetic measurements of the satellites attitude state are assumed available in tracklets; quick successions of measurements taken during the satellite's visibility window, with one tracklet per orbital pass. Each measurement is assumed subject to noise conforming to a normal distribution $\mathbf{y} \in \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\sigma})$, where σ is the measurement variance and the mean value μ corresponds to the ground-truth state.

Comparison is made with a standard UKF, operated effectively as a EKMF with only one Gaussian kernel and using the same predict/update steps (though neglecting the adaptive splitting/merging routines). The UKF is initialised using the same parameters as the EKMF, and applied in parallel (i.e. matching the same predict/update pattern and using the same sets of observations). For additional validation, the posterior distribution of the EKMF at each observation point is sampled with 10,000 Monte-Carlo points, with each then being individually propagated to the next observation time.

In each subfigure of Figure 1, the blue shaded area represents the 99% confidence bounds modelled by the EKMF and the green bounds represent the same but as modelled by the UKF. A subset of Monte-Carlo samples are shown as grey lines, with the ground truth state given by the orange line. Observation

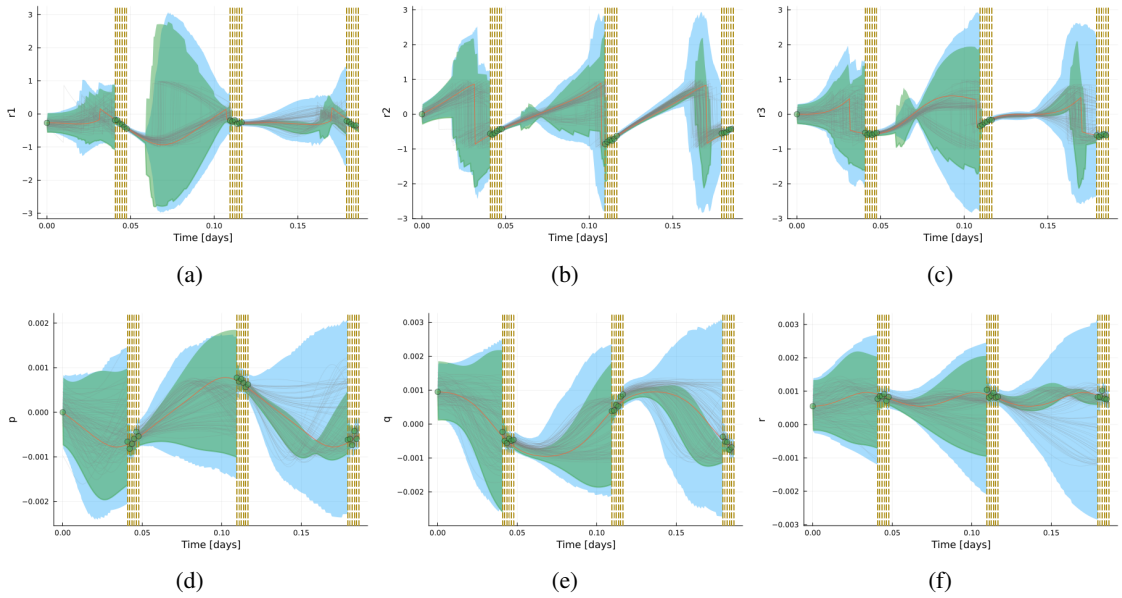


Figure 1: Estimated posterior upper/lower probability bounds of (a-c) the Modified Rodrigues Parameterss (MRPs) (r_1, r_2, r_3) and (d-f) the body-frame rotation rates (p, q, r) for a satellite in LEO.

timepoints are displayed individually as vertical dashed lines, with the actual measurement values shown as green markers.

A general pattern can be seen across each subfigure. That is, the predicted bounds of each state variable will widen over time through propagation due to the modelled uncertainty, then contract towards the true value with additional observations. Convergence is clear within the short time periods represented by each tracklet, with the longer periods of no observations allowing the confidence bounds to widen. An interesting feature are the discontinuities observable in the MRPs. While simply a numerical consequence of the continued rotation of the satellite, the location of each discontinuity is nonetheless uncertain. The filter compensates for this by widening the state bounds around the temporal location of each discontinuity, resulting in a noticeable spike in uncertainty (though successfully containing the possible locations).

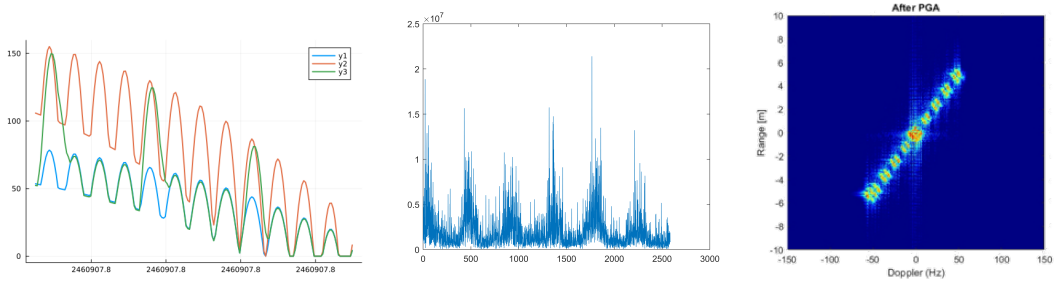
The bounds predicted by the UKF are seen to generally lie within those predicted by the EKMF. This matches expectation, as the UKF essentially predicts one single distribution, whereas the EKMF predicts sets of distributions. In other words, the distribution predicted by the UKF is likely to always physically lie within the set of distributions predicted by the EKMF. It is for this reason that the EKMF bounds are generally wider, successfully containing also the Monte-Carlo samples failed to be captured by the UKF. This suggests the EKMF is successfully representing the uncertainty in the system.

ATTITUDE MEASUREMENT METHODS

To now extend the proposed EKMF and test it's applicability to realistic simulated measurement data from multiple sources, this section briefly introduces and describes three RSO measurement and attitude estimation methods; Multispectral Light Curve Observations, Radar Cross Section (RCS) and Inverse Synthetic Aperture Radar (ISAR), and shows that the uncertainty associated with each can be considered both aleatoric and epistemic, thus are well suited to treatment using the proposed sequential filter.

In each case, we assume the availability of realistic synthetic measurements representing the raw data that can be received via each approach. The following discussion is thus primarily related to the corresponding attitude estimation method used to convert the simulated observation data (see Figure 2 for an example of each) into viable estimates of the attitude state. Discussion of these aspects is relevant to introduce the sources

of ambiguity that can become prevalent also in real data.



(a) (b) (c)

Figure 2: Examples measurements representing (a) Multispectral Light Curves (smoothed), (b) RCS signals and (c) ISAR images.

Multispectral Light-Curves

Multispectral light curve observations are represented as a time series of single-pixel intensity measurements through multiple different colour (or wavelength) channels of the visible spectra, see Figure 2(a). For this work, attitude estimation from multispectral light curves is accomplished via a modified version of the standard Epoch method.[18] Our approach closely follows that of Song et al,[19] and so we refer the reader to this work for the full treatment of this method. By identifying successive maxima in a light curve, the synodic period of the target satellite can be determined. Using this in conjunction with the observational geometry (the Sun-target-observer system), we can then optimise to obtain the sidereal period and spin axis of the target body.

One of the primary disadvantages of the Epoch method is that for light curves of targets with rotational symmetry observed for relatively short durations, the Lomb-Scargle periodogram[20] (used to identify candidate periods) is susceptible to identifying several periodic signals corresponding to successive harmonics of the true spin state. It is therefore to be expected that multiple possible solutions will exist for any given candidate period and target geometry. Specifically due to geometric symmetry, any spin-axis solution defined by right ascension and declination angles (α, δ) will additionally have a second solution corresponding to the axis direction being reversed, $(\alpha + 180, -\delta)$.

In short, given some level of geometric symmetry in the target body but with no assumed prior information available, the Epoch method is often unable to determine which observed spin period is the correct one. The Epoch method may calculate a spin period between any two detectable peaks across all channels. Eliminating duplicates, this can mean that any/all of the quarter period, half period and double period as well as the true period, are detectable with equal levels of probability. The result is a high level of ambiguity in the received measurements, corresponding to several competing, but supported, hypotheses.

The light curve-based measurement method emloyed in this work is assumed to return measurements in the format $\mathbf{y} = [\alpha, \delta, \gamma, T]$ and $\boldsymbol{\sigma}_y = [\sigma_\alpha, \sigma_\delta, \sigma_\gamma, \sigma_T]$, where the satellite inertial orientation is represented by right ascension α and declination δ of an angular momentum vector L about which the body spins in the direction indicated by the right-hand rule with an instantaneous rotational angle of γ and at a rotational velocity defined by the spin period T. The transformation of the measurement vector $\mathbf{y}$ into the state vector $\mathbf{x}$ required by the EKMF is accomplished using an interface function of the form $\tilde{\mathbf{x}} = [r_{IB}, \omega_B] = g(\mathbf{y})$, where

the function g represents a wrapper to the following equations:

$$|\omega| = \frac{2\pi}{T} \tag{45}$$

$$\boldsymbol{\omega}_I = \begin{bmatrix} \cos\alpha\cos\delta \\ \sin\alpha\cos\delta \\ \sin\delta \end{bmatrix} * |\omega| \tag{46}$$

$$\boldsymbol{\omega}_B = A\boldsymbol{\omega}_I \tag{47}$$

$$r_{IB} = f(A) \tag{48}$$

where the attitude tensor A between the ECI reference frame and the embedded rigid BCS is a rotation matrix computed by:

$$A = R_2(-x_p)R_1(-y_p)R_3(\gamma)R_1\left(\frac{\pi}{2} - \delta\right)R_2\left(\frac{\pi}{2} + \alpha\right) \tag{49}$$

Here, R_1, R_2 and R_3 represent the standard rotation matrices about the x, y and z-axes respectively. The pole coordinates x_p and y_p describe the relative position of the momentum vector L with respect to the satellite body axis $+Z_{BCS}$. In the following, the pole coordinates are assumed $x_p = y_p = 0$ (that is, the spin axis is assumed directly aligned with the $+Z_{BCS}$ axis of the target geometry) given the intention to minimise prior knowledge of the system under observation. Given then that by definition the light curve measurement approaches only directly measure the target spin axis and it's corresponding spin rate (with the exact rotational position only able to be inferred from x_p, y_p and γ), the resulting MRPs describing the rotational position of the target body (extracted from the rotation matrix A) are thus to be taken in relativistic terms. For scenarios in which light curve measurements are received in addition to an alternative measurement method, one that does directly measure the rotational position of the target, the pole coordinates can be included as additional unobserved parameters within the state vector processed by the EKMF.

Radar Cross Section (RCS)

Radar Cross Section (RCS) quantifies how detectable a complex 3D object is via radar and provides an accurate representation of how the signal backscatters from its surface. RCS measurements provide a single time-intensity signal as reflected from the target geometry, see Figure 2(b). Similarly to light curves, this signal again exhibits periodic variation due to the rotational motion of the target and its translation across the observable arc relative to the observing Ground Station (GS).

Assuming the availability of raw RCS measurements, attitude estimation can again be performed using the Epoch method as described in the previous section. In this case, the theoretical background is identical besides a change in observational geometry. That is, rather than considering the Sun-object-observer geometry, we replace the illumination source with the observing GS. This geometric reduction means that we can expect to see a maximum in the received signal when a large area of the target is normal to the Line of Sight (LOS). From here, the method proceeds as described for light curves. Accordingly, the state estimations generated by examining RCS observations using this approach remain susceptible to the same ambiguities as described previously for light curve measurements (more so given the lack of multiple data channels).

Inverse Synthetic Aperture Radar (ISAR)

Inverse Synthetic Aperture Radar (ISAR) imaging allows the reconstruction of two-dimensional radar reflectivity maps of rotating and/or translating objects. When an object rotates with respect to the radar LOS, the relative motion induces Doppler frequency shifts that are used to form high-resolution images, such as that shown in Figure 2(c). Methods to synthetically generate ISAR images can be found in the works by Ozdemir[21] and Chen & Martorella,[22] with the reader being directed towards these works for further details.

Assuming the availability of a time-sequence of ISAR images over a fixed observation window, the attitude (spatial orientation and rotational velocity) of a 3D space object of unknown shape can be estimated using only the corresponding image sequence and the direct radar returns data. An example shape-agnostic 'shape-matching' optimisation process can be summarised as follows:

1. The range residuals from the radar returns across the full observation window are used to estimate the primary dimensions of a bounding cuboid centred on the target geometry.

2. The approximating cuboid is given a rotational velocity, initial orientation and spin axis with respect to the radar LOS frame and propagated in time across the duration of the ISAR image sequence.

3. At the time instance of each ISAR image, a binary mask of the approximating cuboid is created and overlaid onto the ISAR image plane, where the estimated rotational velocity has been used to resolve the ISAR images into range-crossrange scale.

4. The sequence of combined images is evaluated in terms of image similarity and a single numeric score is produced corresponding to the quality of the matching.

5. A numerical optimiser is employed to cycle steps 2-4 for a population of multiple candidate solutions to produce a distribution of possible target states.

This process as described is inherently limited given the strong geometric symmetries of the approximating cuboid used to perform the shape matching. Accordingly, Step (5) may produce several valid and well-matched attitude state estimations representing alternative orientations simultaneously. This in essence describes a strong ambiguity in the attitude measurements received via this approach, thus again is well suited for examination via the EKMF.

NUMERICAL EXPERIMENT

This section presents a more complex attitude estimation problem in which measurement ambiguities and a resulting multiplicity in possible attitude states from several available sensor types are modelled using the proposed filtering approach.

The following example represents a Starlink-3333 satellite in LEO, with orbital conditions defined by the Keplerian parameters listed in Table 1. The satellite is assumed spinning at a constant rate about a single axis aligned with one of the primary body-axes components. To represent an inherent lack of knowledge of the target geometry, an epistemic uncertainty (in the form of an initial confidence bound) is applied to the principle Moments of Inertia $I = [A, B, C]$ of the target. This uncertainty is based on assuming each element conforms to a uniform distribution between prescribed upper and lower bounds such that: $A \sim U(319.3449 \pm 10\%)$kg/m^2, $B \sim U(1099.8847 \pm 10\%)$kg/m^2 and $C \sim U(1288.1368 \pm 10\%)$kg/m^2. A torque-free dynamic model is employed to evaluate the target's ground-truth position and velocity and to propagate the current estimated distributions from one observation point to the next.

The initial attitude state is defined by Euler angles (ϕ, θ, ψ) representing the body rotation with respect to the inertial ECI reference-frame (later converted into MRPs to avoid singularities within the Mixture Model), and body-frame rotational rates (p, q, r), with values given in Table 2. To allow measurements of the satellite attitude and rotational rate to lead to an update of the Inertia moments estimates, the state vector is augmented as described earlier.

Table 1: Orbital elements of the target body.

Parameter	Value
Epoch [UTC]	2025-08-20 00:00:00
a [km]	6917
e [-]	0.0001541
i [deg]	55.3
Ω [deg]	355.2318
ω [deg]	81.9820
M [deg]	0

Table 2: Initial attitude state.

Parameter	Value at t_0
ϕ [deg]	0
θ [deg]	90.0
ψ [deg]	0
p [deg/s]	0
q [deg/s]	6.29
r [deg/s]	0

To provide realistic measurements of the spacecrafts attitude, three measurement types are assumed available: Multispectral Light Curves, RCS and ISAR images. It is further assumed that an appropriate attitude estimation method is available for each measurement type such that processing the raw data from any valid observation produces one or more valid hypotheses of the full attitude state $\mathbf{x} = [\mathbf{r}_{MRP}, \boldsymbol{\omega}_B]$, where the MRPs are $\mathbf{r}_{MRP} = [r_1, r_2, r_3]$ and the angular velocity components are $\boldsymbol{\omega}_B = [p, q, r]$, although the proposed approach remains applicable with only partial state measurements.

Each of the three selected measurement types are assumed available separately from one of three respective ESA-operated GSs spaced around the Earth. These stations include Redu (Belgium), Korou Station, Cayenne (South America) and Santa Maria Station (Portugal) with geographical parameters listed in Table 3, alongside their assigned measurement type. Table 3 additionally lists the operational constraints associated to each measurement method, in terms of the minimum required length of an observation window and the specification of the frequency with which measurements can be obtained from that particular station. Each observation constitutes a data collection period conforming to the visibility window defined by assuming a 30^o horizon. A further constraint is included by recognising that light-curve measurements in particular can only be received when the satellite is illuminated by the Sun (i.e. the target does not lie in the Earth's shadow).

Table 3: Measurement method and Ground Station specifications.

Measurement Type	Ground station	Geodetic Coordinates [deg, deg, m]	Observation Window	Window Frequency	Field of View [deg]
Multispectral Light Curves (Epoch Method)	Redu, Belgium (ESA)	[50.0005, 5.1453, 330.68]	2-6 rotations	1 per day	120
RCS (Epoch Method)	Kourou station, Cayenne, South America (ESA)	[5.2514, -52.8047, 10.32]	2-6 rotations	1 per 2 days	120
ISAR	Santa Maria Station, Portugal (ESA)	[36.9977, -25.1359, 198.0]	1 rotation	1 per 3 days	120

To validate the efficacy of the EKMF, the posterior distribution at each observation point is sampled with 10,000 Monte-Carlo points, with each then being propagated to the next observation point. The resulting Monte-Carlo distribution can then be directly compared with both the prior and posterior distributions output by the EKMF. As before, results are additionally validated against a standard UKF initialised in the same manner as the EKMF and propagated/updated in parallel.

Results

Figures 3 shows the estimated upper/lower bounds for the MRPs describing the satellites orientation (r_1, r_2, r_3) and the body-frame rotation rates (p, q, r) respectively. In each figure, the blue shaded area represents the 99% confidence bounds as estimated by the EKMF, whereas the green shaded area represents that of the UKF. Measurement timepoints are depicted with vertical dashed lines, where the green, red and blue markers correspond to the light-curve, RCS and ISAR measurements respectively. The ground-truth state is shown with yellow markers. For simplicity, only the posterior bounds are shown.

As expected, the received measurements represent a significant level of uncertainty in the target state. The MRPs are typically bounded between $[-1, 1]$ for both the UKF and the EKMF, suggesting the satellite can viably exist at any orientation at a particular measurement time. Slightly more nuance is visible from examining the estimated rotational rates in Figure 3(d-f). Here, we can see measurement ambiguities in several different forms. Firstly, the light curve measurements in general correctly identify the spin axis, which can be seen given the larger magnitudes of the spin attributed to the q component relative to p and r. However, the light curve measurements also consistently return the half-spin rate as well as the true spin rate (noise notwithstanding), likely given the observability of non-consecutive peaks in the light intensity data. Secondly, the RCS measurement are seen less likely to associate the correct spin magnitude to the appropriate body axis, and additionally return an ambiguity in the direction of spin, seen by the 'mirroring' of measurements between positive and negative values. The RCS-based method also returns in one instance

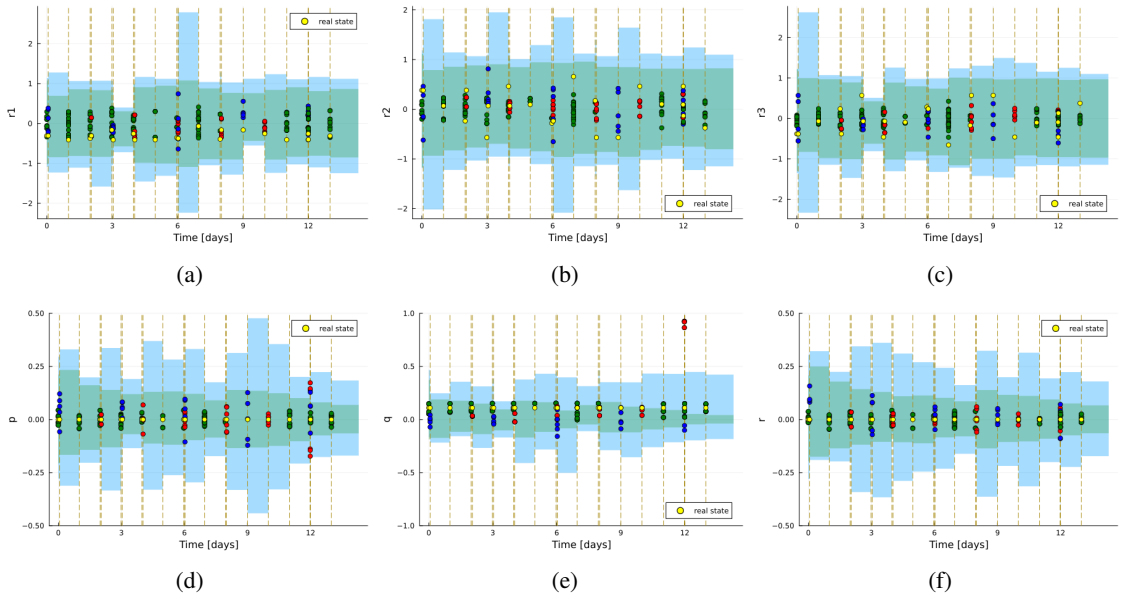


Figure 3: Estimated posterior upper/lower probability bounds of (a-c) the Modified Rodrigues Parameters (r_1, r_2, r_3) and (d-f) the body-frame rotation rates (p, q, r) for a Starlink-3333 satellite in LEO.

a significant harmonic of the true spin, seen as an excessively high spin rate measurement in the q component at the day-12 observation. Finally, the ISAR measurements also fail to consistently attribute the single axis spin to the correct body-axis, and also return an ambiguity in the spin direction. A further complication is that the batched measurements returned by the ISAR method can maintain these inconsistencies within the batch itself, meaning that several possible orientations are available at a single observation instance.

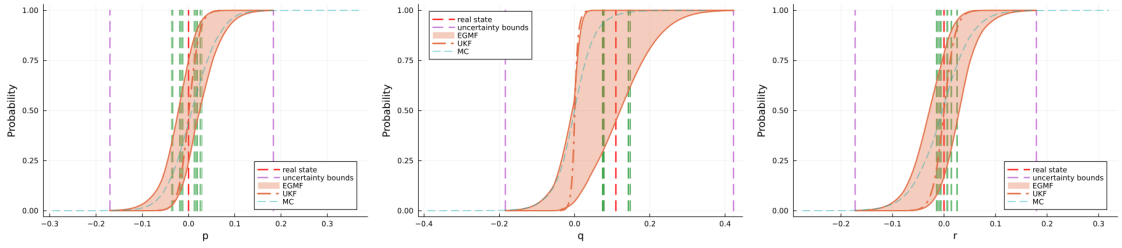
Despite these issues, the EKMF is seen to first maintain the received uncertainty, and second begin to stabilise the estimation as more measurements are received. Where a certain observation period returns high uncertainty/ambiguity in the measurements (such as when ISAR measurements are received), the EKMF widens the confidence bounds*, and where measurements agree (such as around $T \sim 7:8$ days) the bounds begin to contract. With more measurements, the variation in the bounds between successive measurements begin to stabilise (e.g. for $T \geq 11$ days).

The inconsistencies in the received hypotheses can be seen to cause premature convergence in the UKF. More specifically, given that the observations are collectively centred on zero (due to the prevalence of positive/negative ambiguity), this then constitutes the area of greatest support and thus causes the UKF to converge upon zero in all three axes, despite this result being practically improbable. The inconsistency in the received measurements therefore is seen to prevent the UKF from stabilising around the 'correct' component values. However, the EKMF, having been designed to accommodate scenarios in which the measurements themselves are only an incomplete representation of reality (i.e. giving rise to inconsistent and/or ambiguous measurements), correctly maintains the observed uncertainties within a wider set of confidence bounds. Important in a practical sense, the EKMF confidence bounds tend not to contract within the positive and negative values of the true spin magnitude across all three axis components. This reflects the ambiguity received from the measurements and in essence tells the operator that the target object is spinning with a bounded rate around an uncertain axis, the latter most likely due to significant object symmetries.

Figure 4 shows the estimated Cumulative Distribution Functions (CDFs) for the angular velocity components after the final observation update. In these figures, the ground-truth values are shown as a red vertical

*The measurements at $T \sim 12$ days are received in quick succession, with the EKMF bounds first widening to capture the anomalous RCS measurements, before quickly contracting again as the light curve and ISAR-based measurements are received.

dashed line, received measurements are green dashed lines and the upper/lower 99% confidence bounds are purple dashed lines. The orange area represents the epistemic probability distribution as estimated by the EKMF, whereas the orange dashed curve is that of the UKF. The Monte-Carlo sample distribution is shown as the blue dashed curve. Two important features are visible from Figure 4. Firstly, the probability distributions for the p and r components are centred around zero, with upper/lower bounds comfortably bounding a positive and negative value of the true spin state (0.1097 rad/s) accounting for measurement noise. Each of these distributions shows an epistemic component and contains both the Monte-Carlo and UKF distributions, suggesting that the EKMF is both correctly representing the true propagated distribution and a wider range of uncertainty relative to the UKF, which matches expectation. Secondly, the q component (i.e. the component containing the true magnitude of the target spin rate), the epistemic probability distribution is comprised of two distinct functional forms at the extremes. The upper CDF (i.e. the right hand side curve bounding the orange area) is a standard S-curve representing a Gaussian distribution, whereas the lower CDF (the left hand side edge) is much tighter, almost conforming to a vertical epistemic boundary, matching the Monte-Carlo distribution below the 50% probability mark and the UKF above the 50% probability mark. This suggests that while the spin direction ambiguity received from the measurements is still maintained, the positive value is attributed an overall higher level of confidence, though with larger uncertainty in the specific value, corresponding to the harmonics ambiguity given by the Epoch method across the most recent measurements at that time.



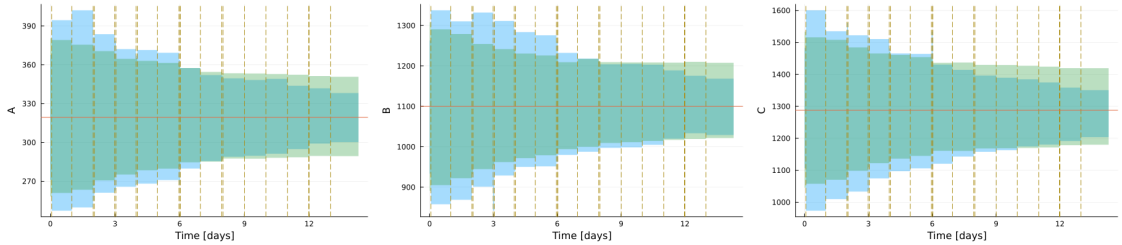
(a) (b) (c)

Figure 4: Cumulative Distribution Functions (CDFs) for the body-frame rotation rates pqr of the Starlink-3333 satellite at the final observation timepoint.

Considering the ability of the EKMF to estimate the values of unobserved model parameters, Figures 5 and 6 show the upper/lower confidence bound and corresponding CDFs for the target inertia parameters. In Figure 5, convergence can be seen in both the EKMF and UKF confidence bounds. This shows that despite the ambiguities and inconsistencies in the received measurements, the EKMF is successfully reducing the uncertainty associated to the target geometry. Interestingly, the UKF appears to plateau at a certain level while the EKMF continues to converge further towards the true inertia values. This can possibly be accounted for by examining the CDFs in Figure 6. The distributions associated to the UKF (orange dotted curves) conform to the standard S-curve of aleatoric uncertainty, as expected for a single Gaussian kernel. The EKMF distributions are however significantly more upright, approaching the vertical form of a purely epistemic interval. This suggests then that the epistemic nature of the uncertainty associated to the inertia parameters, while successfully modelled by the EKMF, cannot be represented by the UKF. This is to be expected given the natural difficulty in fitting a single Gaussian distribution across an interval, and thus would explain the plateauing at wider bounds than the EKMF.

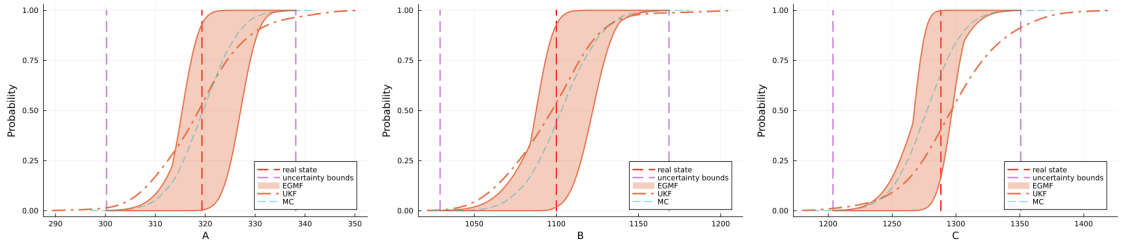
Computational cost

Finally, this section briefly addresses the computational load of the proposed framework in terms of incurred function evaluations relative to both the UKF and the Monte-Carlo based approaches. Figure 7 shows the accumulation of calls to the dynamics propagator for each approach during the course of the example presented in the previous section.



(a) (b) (c)

Figure 5: Estimated posterior upper/lower probability bounds of the unobserved principle Moments of Inertia of the Starlink-3333 satellite at the final observation timepoint.



(a) (b) (c)

Figure 6: Cumulative Distribution Functions (CDFs) for the unobserved principle Moments of Inertia of the Starlink-3333 satellite at the final observation timepoint.

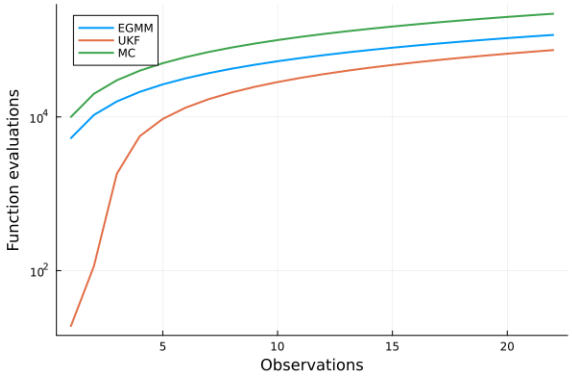


Figure 7: Number of function evaluations incurred by the EKMF, UKF and Monte-Carlo approaches for the Starlink-3333 example, relative to the total number of received observations.

Whilst the UKF represents the least expensive approach computationally, the EKMF can be seen to incur significantly less burden than the traditionally high-expense Monte-Carlo approach. This is important given the observed alignment between the EKMF and the Monte-Carlo predictions thus far.

CONCLUSION

This paper has presented a robust sequential filter for state estimation capable of processing measurements from multiple sensors and representing an ambiguous multiplicity in the target states. We demonstrated that this multiplicity of solutions can be treated as an epistemic uncertainty, and as such modelled using an Epistemic Kernel Mixture Model (EKMM) with interval valued weights. Robust Bayesian process was used

to update the weights and their bounds in order to build a family of posterior distributions from each new measurement. Between observations, kernels are automatically adapted in number and size to optimise the propagation of a family of admissible distributions. From indirect inference through propagation, the filter has also been shown capable of estimating unknown model parameters and updating the current distribution accounting for epistemic uncertainty. We also included the fundamental assumption that every observed state is an individual hypothesis taken from an unknown underlying distribution, thus the robust inference process is modified to incorporate a multiplicity of possible belief states.

We first presented a technical description of the proposed approach and included a demonstrative state estimation example using consistent and unambiguous measurements subject to Gaussian noise. The filter was seen to successfully represent the uncertainty included in the measurements and progressively refine the resulting state estimation accounting for this uncertainty.

In a more complex example, employing several different sensor types with different uncertainties/ambiguities associated to each, the proposed filtering framework was shown to initially maintain the received multiplicity suggested by the measurements before beginning to contract the confidence bounds with successive observations. A particular strength of the EKMF was found in it's ability to represent epistemic distributions (i.e. interval valued estimates) by virtue of the kernel based structure. It was further shown that the continuous propagation of epistemic distributions allowed the sequential inference of unobserved model parameters, meaning that the filter can be applied to state estimation scenarios with minimal assumptions or prior knowledge regarding the target object.

Both the presented examples were validated against a standard UKF, as well as a complete Monte-Carlo based sampling approach. In each case, the EKMF was seen to contain both the UKF solution and the Monte-Carlo samples within the estimated confidence bounds, again suggesting a successful representation of the system uncertainty. The UKF was generally seen to be incapable of representing the complete treatment of uncertainty, particularly when the uncertainty was primarily epistemic in nature.

Future work will evaluate the EKMF in cases with highly non-linear dynamics, or over larger periods of time, with a particular focus on the representation of uncertainty under sparse ambiguous measurements.

REFERENCES

[1] Y. Oualil, F. Faubel, and D. Klakow, "A multiple hypothesis Gaussian mixture filter for acoustic source localization and tracking," *IWAENC 2012; International Workshop on Acoustic Signal Enhancement*, VDE, 2012, pp. 1–4.

[2] C. Greco and M. Vasile, "Robust Bayesian Particle Filter for Space Object Tracking Under Severe Uncertainty," *Journal of Guidance, Control, and Dynamics*, Vol. 45, No. 3, 2022, pp. 481–498, 10.2514/1.G006157.

[3] C. Greco, S. Campagnola, and M. Vasile, "Robust Space Trajectory Design Using Belief Optimal Control," *Journal of Guidance, Control, and Dynamics*, Vol. 45, No. 6, 2022, pp. 1060–1077, 10.2514/1.G005704.

[4] L. Sánchez, M. Vasile, S. Sanvido, K. Merz, and C. Taillan, "Treatment of epistemic uncertainty in conjunction analysis with Dempster-Shafer theory," *Advances in Space Research*, Vol. 74, No. 11, 2024, pp. 5639–5686, https://doi.org/10.1016/j.asr.2024.09.014.

[5] S. J. Julier and J. K. Uhlmann, "A New Extension of the Kalman Filter to Nonlinear Systems," 1997, p. 12.

[6] S. Holmes, G. Klein, and D. Murray, "An O(N²) Square Root Unscented Kalman Filter for Visual Simultaneous Localization and Mapping," Vol. 31, No. 7, 2009, pp. 1251–1263, 10.1109/TPAMI.2008.189.

[7] F. Gustafsson and G. Hendeby, "Some Relations Between Extended and Unscented Kalman Filters," Vol. 60, No. 2, 2012, pp. 545–555, 10.1109/TSP.2011.2172431.

[8] J. R. Carpenter and C. N. D'souza, "Navigation filter best practices," tech. rep., 2018.

[9] A. G. Wills, J. Hendriks, C. Renton, and B. Ninness, "A Bayesian Filtering Algorithm for Gaussian Mixture Models," 2017.

[10] J. R. Hershey and P. A. Olsen, "Approximating the Kullback Leibler divergence between Gaussian mixture models," *2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07*, Vol. 4, IEEE, 2007, pp. IV–317.

[11] F. Faubel, J. McDonough, and D. Klakow, "The Split and Merge Unscented Gaussian Mixture Filter," Vol. 16, No. 9, 2009, pp. 786–789. Conference Name: IEEE Signal Processing Letters, 10.1109/LSP.2009.2024859.

[12] V. Vittaldev and R. P. Russell, "Multidirectional Gaussian mixture models for nonlinear uncertainty propagation," *Computer Modeling in Engineering & Sciences*, Vol. 111, No. 1, 2016, p. 83.

[13] A. Runnalls, "Kullback-Leibler Approach to Gaussian Mixture Reduction," Vol. 43, No. 3, 2007, pp. 989–999. Conference Name: IEEE Transactions on Aerospace and Electronic Systems, 10.1109/TAES.2007.4383588.

[14] E. Wan and R. Van Der Merwe, "The unscented Kalman filter for nonlinear estimation," *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No.00EX373)*, Oct. 2000, pp. 153–158, 10.1109/ASSPCC.2000.882463.

[15] T. Gehrig, U. Klee, J. W. McDonough, S. Ikbal, M. Wölfel, and C. Fügen, "Tracking and beamforming for multiple simultaneous speakers with probabilistic data association filters.," *INTERSPEECH*, 2006.

[16] Y. Bar-Shalom, T. E. Fortmann, and P. G. Cable, "Tracking and data association," 1990.

[17] F. Faubel, M. Georges, B. Fu, and D. Klakow, "Robust Gaussian mixture filter based mouth tracking in a real environment," *Proc. Visual Computing Research Conference (IVCI, Saarbrucken)*, 2009.

[18] P. Magnusson, "Distribution of spin axes and senses of rotation for 20 large asteroids," *Icarus*, Vol. 68, No. 1, 1986, pp. 1–39.

[19] C. Song, H.-Y. Lin, and C.-Y. Zhao, "Analysis of Envisat's rotation state using epoch method," *Advances in Space Research*, Vol. 66, No. 11, 2020, pp. 2681–2688.

[20] J. T. VanderPlas, "Understanding the lomb–scargle periodogram," *The Astrophysical Journal Supplement Series*, Vol. 236, No. 1, 2018, p. 16.

[21] C. Ozdemir, *Inverse synthetic aperture radar imaging with MATLAB algorithms*. John Wiley & Sons, 2021.

[22] V. C. Chen and M. Martorella, "Inverse synthetic aperture radar imaging," *Sci. Tech. Publishing*, Vol. 304, 2014.