# GOOD THINGS COME TO THOSE WHO WEIGHT! IMPROVING GAUSSIAN MIXTURE WEIGHTS FOR CISLUNAR DEBRIS TRACKING

**Dalton Durant**[*], **Uwe D. Hanebeck**[†], **and Renato Zanetti**[‡]

Kessler syndrome is the process of orbital debris colliding and compounding over time. Within several orbits, 100% of satellites are lost. It is then critical to have accurate and efficient tracking methods. In this study, the Gaussian Mixture Probability Hypothesis Density filter is used to track multiple targets generated from a debris cloud clustered around a similar cislunar near rectilinear halo orbit. Traditionally, updating the weights of individual Gaussian mixture components involved linearizing the measurement model about the prior means. However, recent advancements have shown that linearizing about the posteriors significantly improves estimation performance without sacrificing efficiency. This work applies the posterior linearization technique to multi-target filtering and further eliminates the need for linearization altogether by employing techniques from importance sampling. The study proves that each method has exact equivalence under linear measurement models, and demonstrates empirically better performance for the posterior methods under nonlinear measurement models.

## INTRODUCTION

The Kessler syndrome, or Kessler effect, describes a catastrophic scenario where orbital debris continuously collides, leading to an exponential increase in space debris.[1] In just a few orbits, all satellites could be lost——a grim reality that poses a severe threat to future space missions and exploration.

Orbital debris has been extensively studied,[2–5] but only recently has attention shifted toward debris threats in cislunar space. Recent studies explore topics such as spacecraft explosions at Lagrange points,[6–10] re-collision risks involving debris from NASA's Lunar Gateway,[11–13] and tracking methods related to Space Situational Awareness (SSA)[14–21] to name a few. This growing body of work highlights the increasing interest in studying cislunar debris.

It is the interest of this work to advance the area of cislunar debris tracking through accurate and efficient multi-target filtering methods. Comprehensive reviews of multi-target filtering and tracking methods can be found in References 22 and 23. One such method, the Gaussian Mixture Probability Hypothesis Density (GM-PHD) filter,[24] is a popular method for multi-target filtering, with applications in radar tracking,[25,26] computer vision,[27,28] and autonomous navigation.[29,30] The

---
[*]Ph.D. Student, Department of Aerospace Engineering & Engineering Mechanics, The University of Texas at Austin, Austin, TX.
[†]Chaired Professor, Department of Computer Science, Karlsruhe Institute of Technology, Karlsruhe, Germany
[‡]Associate Professor, Department of Aerospace Engineering & Engineering Mechanics, The University of Texas at Austin, Austin, TX.

GM-PHD filter estimates the states and cardinality of multiple targets from noisy measurements without requiring explicit data association between measurements and targets. While it does not label or catalog targets—necessary in some SSA applications—it is well suited for tracking multiple debris fragments clustered around a cislunar near rectilinear halo orbit (NRHO), where labeling is unnecessary, and the focus is on estimating the amount, location, and trajectories of the debris.

The GM-PHD filter is a specific implementation of the Probability Hypothesis Density (PHD) filter,[31] based on Random Finite Set (RFS) theory. Instead of propagating the full multi-target posterior distribution, the PHD filter propagates the first-order moment of the RFS, also known as the intensity function. A defining feature of the GM-PHD filter is its approximation of the intensity function as a weighted sum of Gaussian components, which makes it analytically tractable. However, inaccuracies in computing these weights can lead to filter divergence, erroneous state estimates, and misalignment with the true intensity distribution. The weights are updated by approximating the marginal likelihood of each component, traditionally by linearizing the measurement model around the prior means of the components. While this method is straightforward, it introduces limitations. Recent advances indicate that linearizing around the posterior means of the components significantly improves estimation accuracy without compromising computational efficiency,[21,32] though this approach has only been applied to single-target filters so far.

It is then the purpose of this work to extend the posterior linearization approach to multi-target filters, specifically, the GM-PHD. Furthermore, this work also proposes eliminating the need for linearization altogether by employing importance sampling techniques. This anti-linearization approach is generalized for all Gaussian mixture-type filters and is empirically shown to significantly improve the GM-PHD filter's accuracy, consistency, and computational efficiency.

This work begins by reviewing Gaussian mixture weights and introducing the anti-linearization weighting method. Next, under linear measurement models, this work proves that the proposed method is exactly equivalent to traditional and recent posterior linearization weighting methods. This work then demonstrates improved performance under nonlinear models for general Gaussian mixture-type filters. The proposed anti-linearization method and the recent posterior linearization method are then applied to the GM-PHD filter, showing improved performance in a cislunar debris tracking example compared to traditional methods. The work then concludes with a summary and recommendations for future research.

## GAUSSIAN MIXTURE WEIGHTS

This work models the state dynamics and measurements as

$$\dot{\mathbf{x}} = f(\mathbf{x}, \boldsymbol{\nu}), \tag{1}$$

$$\mathbf{y} = h(\mathbf{x}) + \boldsymbol{\eta}, \tag{2}$$

where $\mathbf{x} \in \mathbb{R}^{n_x}$ is the state with dynamics function $f(\cdot) \colon \mathbb{R}^{n_x} \to \mathbb{R}^{n_x}$ and noise $\boldsymbol{\nu}$. Additionally, $\mathbf{y} \in \mathbb{R}^{n_y}$ is the measurement with mapping function $h(\cdot) \colon \mathbb{R}^{n_x} \to \mathbb{R}^{n_y}$. The measurements are modeled with additive, white, Gaussian measurement noise $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$ uncorrelated to all $\boldsymbol{\nu}$.

In Bayesian inversion, the posterior is usually the function of interest and is represented by Bayes' rule

$$p(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{x})\, p(\mathbf{x})}{p(\mathbf{y})}, \tag{3}$$

where $p(\mathbf{x})$ is the prior, $p(\mathbf{y}|\mathbf{x})$ is the likelihood, and $p(\mathbf{y})$ is the marginal likelihood; also known as the measurement probability density function (PDF) or the Bayesian evidence. In problems like cislunar debris tracking, dynamics and measurements are highly nonlinear causing the prior and the marginal likelihood to exhibit significant non-Gaussian uncertainties making them analytically intractable and therefore, as a result, there are no analytical solutions to the posterior PDF.

Fortunately, any non-Gaussian prior can be approximated as a weighted sum of $n$ Gaussians with associated prior weights $w_i^-$ arbitrarily well:[33]

$$p(\mathbf{x}) \approx \sum_{i=1}^{n} w_i^- \, p_i(\mathbf{x}). \tag{4}$$

Performing Bayesian inversion on the individual components, the posterior distribution is

$$
\begin{aligned}
p(\mathbf{x}|\mathbf{y}) &\propto \sum_{i=1}^{n} w_i^- \, p(\mathbf{y}|\mathbf{x}) \, p_i(\mathbf{x}) \\
&= \sum_{i=1}^{n} \underbrace{w_i^- \, p_i(\mathbf{y})}_{\propto w_i^+} \, p_i(\mathbf{x}|\mathbf{y}),
\end{aligned}
\tag{5}
$$

with updated weights

$$w_i^+ = \frac{w_i^- \, p_i(\mathbf{y})}{p(\mathbf{y})}, \tag{6}$$

and,

$$p(\mathbf{y}) \approx \sum_{j=1}^{n} w_j^- \, p_j(\mathbf{y}), \tag{7}$$

where $p_i(\mathbf{x})$ is the $i$-th *component's* prior, $p_i(\mathbf{y})$ is its marginal likelihood, and $p_i(\mathbf{x}|\mathbf{y})$ is its posterior. Note that if the prior is exactly a Gaussian sum and the measurements are linear, then a closed form solution exists where the posterior is also a Gaussian sum. In this case, it is optimal to update each component with a Kalman filter.[34]

However, for nonlinear measurements, it is generally not possible to compute $p_i(\mathbf{y})$ and $p_i(\mathbf{x}|\mathbf{y})$ exactly, and some approximations need to be made to perform Bayesian inversion on the components (via EKF,[35] UKF,[36,37] CKF,[38] *etc.*). In recent work,[21,32] it is argued that once the $i$-th component's posterior $p_i(\mathbf{x}|\mathbf{y})$ is calculated, it is a more accurate representation of the unknown state than the prior, and hence any approximation made to compute the marginal likelihood $p_i(\mathbf{y})$ should rely on the posterior $p_i(\mathbf{x}|\mathbf{y})$ rather than the prior $p_i(\mathbf{x})$. In-other-words, the updated weights should rely on the posterior rather than the prior because the posterior typically contains more information about the system.

**Traditional Weights**

There is a common assumption—albeit whose implications have not been explored in depth—in which the updated weights in Equation (6) are typically computed by linearizing the measurement model. This linearization is computed using a first-order Taylor series expansion centered around the $i$-th component's prior state estimate $\bar{\mathbf{x}}^{(i)}$ as

$$h(\mathbf{x}^{(i)}) \approx h(\bar{\mathbf{x}}^{(i)}) + \bar{\mathbf{H}}^{(i)}(\mathbf{x} - \bar{\mathbf{x}}^{(i)}), \tag{8}$$

3

where $\bar{\mathbf{H}}^{(i)} = \frac{\partial h(\bar{\mathbf{x}}^{(i)})}{\partial \bar{\mathbf{x}}^{(i)}}$ is the Jacobian matrix, which captures the gradient of the measurement model with respect to the state evaluated at the prior estimate. It then follows that the innovation covariance is given by

$$
\begin{aligned}
\bar{\mathbf{P}}_{\mathbf{yy}}^{(i)} &= \mathbb{E}[(\mathbf{y} - h(\bar{\mathbf{x}}^{(i)}))\,(\mathbf{y} - h(\bar{\mathbf{x}}^{(i)}))'] \\
&= \bar{\mathbf{H}}^{(i)}\bar{\mathbf{P}}_{\mathbf{xx}}^{(i)}\bar{\mathbf{H}}^{(i)'} + \mathbf{R},
\end{aligned}
\tag{9}
$$

where $\bar{\mathbf{P}}_{\mathbf{xx}}^{(i)}$ is the $i$-th component's prior covariance estimate.

Using Equation (8) and its associated innovation covariance from Equation (9), the marginal likelihood is approximated by

$$
p(\mathbf{y}) \approx \bar{p}(\mathbf{y}) = \sum_{j=1}^{n} w_j^- \, \bar{p}_j(\mathbf{y}),
\tag{10}
$$

$$
\bar{p}_i(\mathbf{y}) = \mathcal{N}(\mathbf{y}; h(\bar{\mathbf{x}}^{(i)}), \bar{\mathbf{P}}_{\mathbf{yy}}^{(i)}),
\tag{11}
$$

resulting in the traditionally computed weights

$$
w_i^+ \approx \bar{w}_i^+ = \frac{w_i^- \, \bar{p}_i(\mathbf{y})}{\bar{p}(\mathbf{y})}.
\tag{12}
$$

Although they are used traditionally, these weights can hinder the performance of Gaussian Sum Filters (GSF),[39–41] Ensemble Gaussian Mixture Filters (EnGMF),[21,42–49] and other Gaussian Mixture Model (GMM)—type filters.[50–52]

**Posterior Linearization Weights**

Reference 32 found that better weights can be computed by linearizing the measurement model about each component's posterior estimate $\hat{\mathbf{x}}^{(i)}$ rather than the prior $\bar{\mathbf{x}}^{(i)}$ as

$$
h(\mathbf{x}^{(i)}) \approx h(\hat{\mathbf{x}}^{(i)}) + \hat{\mathbf{H}}^{(i)}(\mathbf{x} - \hat{\mathbf{x}}^{(i)}),
\tag{13}
$$

where $\hat{\mathbf{H}}^{(i)} = \frac{\partial h(\hat{\mathbf{x}}^{(i)})}{\partial \hat{\mathbf{x}}^{(i)}}$ is the Jacobian matrix, which now captures the gradient of the measurement model with respect to the state evaluated at the posterior estimate. It then follows that

$$
\begin{aligned}
\hat{\mathbf{P}}_{\mathbf{yy}}^{(i)} &= \mathbb{E}[(\mathbf{y} - h(\hat{\mathbf{x}}^{(i)}))\,(\mathbf{y} - h(\hat{\mathbf{x}}^{(i)}))'] \\
&= \hat{\mathbf{H}}^{(i)}\hat{\mathbf{P}}_{\mathbf{xx}}^{(i)}\hat{\mathbf{H}}^{(i)'} + \hat{\mathbf{R}} - \hat{\mathbf{H}}^{(i)}\mathbf{K}^{(i)}\mathbf{R} - (\hat{\mathbf{H}}^{(i)}\mathbf{K}^{(i)}\mathbf{R})' \\
&= (\hat{\mathbf{H}}^{(i)} - \bar{\mathbf{H}}^{(i)})\hat{\mathbf{P}}_{\mathbf{xx}}^{(i)}(\hat{\mathbf{H}}^{(i)} - \bar{\mathbf{H}}^{(i)})' \\
&\quad + (\mathbf{I} - \bar{\mathbf{H}}^{(i)}\mathbf{K}^{(i)})\bar{\mathbf{P}}_{\mathbf{yy}}^{(i)}(\mathbf{I} - \bar{\mathbf{H}}^{(i)}\mathbf{K}^{(i)})',
\end{aligned}
\tag{14}
$$

where $\hat{\mathbf{P}}_{\mathbf{xx}}^{(i)}$ is the $i$-th component's posterior covariance estimate, $\mathbf{I}$ is the $n_y \times n_y$ identity, and $\mathbf{K}^{(i)} = \bar{\mathbf{P}}_{\mathbf{xx}}^{(i)}\bar{\mathbf{H}}^{(i)'}\bar{\mathbf{P}}_{\mathbf{yy}}^{(i)-1}$.

Using Equation (13) and its associated innovation covariance from Equation (14), the marginal likelihood is approximated by

$$
p(\mathbf{y}) \approx \hat{p}(\mathbf{y}) = \sum_{j=1}^{n} w_j^- \, \hat{p}_j(\mathbf{y}),
\tag{15}
$$

$$
\hat{p}_i(\mathbf{y}) = \mathcal{N}(\mathbf{y}; h(\hat{\mathbf{x}}^{(i)}), \hat{\mathbf{P}}_{\mathbf{yy}}^{(i)}),
\tag{16}
$$

4

resulting in the posterior linearization weights

$$w_i^+ \approx \hat{w}_i^+ = \frac{w_i^- \; \hat{p}_i(\mathbf{y})}{\hat{p}(\mathbf{y})}. \tag{17}$$

Reference 32 proved that, for linear measurement models, both approximations $\bar{w}_i^+$ and $\hat{w}_i^+$ yielded exactly equivalent results. This work empirically demonstrates, in nonlinear cases, improved GM-PHD filter performance when linearizing the measurement model about the posterior estimates.

**Anti-Linearization Weights**

A new weight update process is proposed utilizing importance sampling to bypass linearization altogether. This means that this method is Jacobian-free and innovation matrix-free; two sometimes costly computations.

Revisiting Equation (5), the unnormalized contribution of the $i$-th component is

$$w_i^- \; p_i(\mathbf{y}|\mathbf{x}) \; p_i(\mathbf{x}) = w_i^- \; p_i(\mathbf{y}) \; p_i(\mathbf{x}|\mathbf{y}), \tag{18}$$

and has the relationship:

$$w_i^- \; p_i(\mathbf{y}) \; p_i(\mathbf{x}|\mathbf{y}) \propto w_i^+ \; p_i(\mathbf{x}|\mathbf{y}). \tag{19}$$

Traditionally the weights are calculated from the above equation to obtain $w_i^+ \propto w_i^- \; p_i(\mathbf{y})$, which requires some approximation to calculate $p_i(\mathbf{y})$ (such as linearization). This work proposes to instead start from $w_i^- \; p_i(\mathbf{y}|\mathbf{x}) \; p(\mathbf{x}) \propto w_i^+ \; p_i(\mathbf{x}|\mathbf{y})$ to obtain

$$w_i^+ \propto \frac{w_i^- \; p_i(\mathbf{y}|\mathbf{x}) \; p_i(\mathbf{x})}{p_i(\mathbf{x}|\mathbf{y})}, \tag{20}$$

since $p_i(\mathbf{y}|\mathbf{x})$ and $p_i(\mathbf{x})$ are known exactly.

If Bayesian inversion could be performed exactly on each component, evaluating Equation (20) at any value of $\mathbf{x}$ would produce the same value for $w_i^+$. However, when the component's posterior $p_i(\mathbf{x}|\mathbf{y})$ is approximated, Equation (20) returns different values of $w_i^+$ when different values of $\mathbf{x}$ are used. In this work, the weights are calculated by evaluating Equation (20) at the posterior mean:

$$
\begin{aligned}
w_i^+ &\propto \frac{w_i^- \; p_i(\mathbf{y}|\mathbf{x} = \hat{\mathbf{x}}^{(i)}) \; p_i(\mathbf{x} = \hat{\mathbf{x}}^{(i)})}{p_i(\mathbf{x} = \hat{\mathbf{x}}^{(i)}|\mathbf{y})} \\
&= \frac{w_i^- \; \mathcal{N}(\mathbf{y}; \; h(\hat{\mathbf{x}}^{(i)}), \; \mathbf{R}) \, \mathcal{N}(\hat{\mathbf{x}}^{(i)}; \; \bar{\mathbf{x}}^{(i)}, \; \bar{\mathbf{P}}_{\mathbf{xx}}^{(i)})}{\mathcal{N}(\hat{\mathbf{x}}^{(i)}; \; \hat{\mathbf{x}}^{(i)}, \; \hat{\mathbf{P}}_{\mathbf{xx}}^{(i)})}.
\end{aligned}
\tag{21}
$$

The proposed weight update above in Equation (21) bypasses linearization, retaining the full nonlinear measurement model, and avoids computing Jacobians and innovation matrices.

**PROOF OF EQUIVALENCE FOR LINEAR MEASUREMENT MODELS**

This section proves, for linear measurement models, that the proposed anti-linearization weights are equivalent to the posterior linearization weights and the traditional prior linearization weights.

**Theorem 1.** *(Equivalent Weights Under Linear Measurement Models) Consider Gaussian measurement probability distributions $\bar{p}(\mathbf{y})$ and $\bar{p}_i(\mathbf{y})$. These distributions are computed from Equation (10) and Equation (11), respectively, based on prior estimates. Similarly, consider $\hat{p}(\mathbf{y})$ and $\hat{p}_i(\mathbf{y})$, calculated from Equation (15) and Equation (16), respectively, utilizing posterior estimates. In light of these considerations, and given the prior weights $w_i^-$, the traditional weights take the form*

$$\bar{w}_i^+ \;=\; \frac{w_i^- \; \bar{p}_i(\mathbf{y})}{\sum_{j=1}^n w_j^- \; \bar{p}_j(\mathbf{y})}, \tag{22}$$

*the posterior linearization weights take the form*

$$\hat{w}_i^+ \;=\; \frac{w_i^- \; \hat{p}_i(\mathbf{y})}{\sum_{j=1}^n w_j^- \; \hat{p}_j(\mathbf{y})}, \tag{23}$$

*and the anti-linearization weights take the form*

$$\tilde{\hat{w}}_i^+ \;=\; \frac{\frac{w_i^- \; p_i(\mathbf{y}|\mathbf{x}=\hat{\mathbf{x}}^{(i)}) \; p_i(\mathbf{x}=\hat{\mathbf{x}}^{(i)})}{p_i(\mathbf{x}=\hat{\mathbf{x}}^{(i)}|\mathbf{y})}}{\sum_{j=1}^n \frac{w_j^- \; p_j(\mathbf{y}|\mathbf{x}=\hat{\mathbf{x}}^{(j)}) \; p_j(\mathbf{x}=\hat{\mathbf{x}}^{(j)})}{p_j(\mathbf{x}=\hat{\mathbf{x}}^{(j)}|\mathbf{y})}}. \tag{24}$$

*Then the linear case yields $\mathbf{y} \;=\; \mathbf{H}\mathbf{x}^{(i)} \;+\; \eta$, and gives*

$$\bar{w}_i^+ \;=\; \hat{w}_i^+ \;=\; \tilde{\hat{w}}_i^+. \tag{25}$$

*Proof.* If the numerators of Equations (22), (23), and (24) $\forall \; i \;=\; 1, \ldots, n$ are equivalent, then their denominators are also equivalent. Therefore, to prove equivalence, it suffices to prove only equivalence in the numerators and then the denominators follow suite.

Starting from Equation (24), it is proportional to its numerator and the following

$$\begin{aligned}
\tilde{\hat{w}}_i^+ &\propto \frac{w_i^- \; p_i(\mathbf{y}|\mathbf{x} = \hat{\mathbf{x}}^{(i)}) \; p_i(\mathbf{x} = \hat{\mathbf{x}}^{(i)})}{p_i(\mathbf{x} = \hat{\mathbf{x}}^{(i)}|\mathbf{y})} \\
&= \frac{w_i^- \; \hat{p}_i(\mathbf{y}|\mathbf{x}) \; \hat{p}_i(\mathbf{x})}{\hat{p}_i(\mathbf{x}|\mathbf{y})} \\
&= \underbrace{w_i^- \; \hat{p}_i(\mathbf{y})}_{\propto \; \hat{w}_i^+} \\
&= \underbrace{w_i^- \; \bar{p}_i(\mathbf{y})}_{\propto \; \bar{w}_i^+}.
\end{aligned} \tag{26}$$

The final two lines of Equation (26) express that the numerator of the posterior linearization weights in Equation (23) are equal to the numerator of the traditional weights in Equation (22). There is no need to prove this because it was already proven for linear measurement models in Reference 32. Inherently, it also serves as a natural validation for the proof of Equation (25) in the contributions presented in this work. □

**A Linear Scalar Example**

As a sanity check, this work ran a numerical test to confirm the presented proof. The truth $x$ is distributed by $\mathcal{N}(10, 9)$. Linear measurements are generated around the truth $x$:

$$y = x + \eta, \tag{27}$$

where $\eta \sim \mathcal{N}(0, 1)$.

The prior is expressed as a Gaussian mixture of $n = 100$ components each having mean $\bar{x}^{(i)}$ and covariance $\bar{P}_{xx}^{(i)} = \beta_S \sigma_x^2$. The component means $\{\bar{x}^{(i)}\}_{i=1}^{100}$ are each sampled from the Gaussian distribution $\mathcal{N}(x, 9)$ centered on the truth $x$. Kernel Density Estimation (KDE) is used to assign the covariances, where $\sigma_x^2$ is the sample covariance of $\{\bar{x}^{(i)}\}_{i=1}^{100}$ and $\beta_S$ is the bandwidth parameter determined using Silverman's Rule of Thumb[43] which reduces computational cost:

$$\beta_S = \left(\frac{4}{n_x + 2}\right)^{\frac{2}{n_x + 4}} n^{-\frac{2}{n_x + 4}}. \tag{28}$$

This method of setting up a Gaussian mixture prior is common and frequently associated with the Ensemble Gaussian Mixture Filter (EnGMF)[42, 44, 49, 53] (*i.e.*, a practical Gaussian mixture-type filter). The filters then perform individual EKF updates for each component and assign weights according to the three different presented methods.

All three methods use the same truth, prior, and measurements for direct comparison purposes. Table 1 compares the three different weighting methods for the scalar linear measurement model $h(x) = x$, averaging results over 10,000 Monte Carlo simulations.

**Table 1**: Comparing the different weighting methods for the scalar linear measurement model example $h(x) = x$. Each component is individually updated with an Extended Kalman Filter (EKF).

|             | Error      | Cov.    | RMSE    | SNEES   |
|-------------|------------|---------|---------|---------|
| Traditional | 0.00088695 | 0.91273 | 0.73165 | 0.92123 |
| Post. Lin.  | 0.00088695 | 0.91273 | 0.73165 | 0.92123 |
| Anti-Lin.   | 0.00088695 | 0.91273 | 0.73165 | 0.92123 |

Using 100 GMM components and averaged over 10,000 Monte Carlo simulations.

The first column in Table 1 is the average Monte Carlo error of the weighted sum of each component's posterior mean estimate $\hat{x}$ with respect to the truth $x$. The second column is the average weighted covariance estimate $\hat{P}_{xx}$. The third column represents the Monte Carlo averaged filter accuracies. This done via the Root Mean Square Error (RMSE) where a lower RMSE indicates a more accurate filter. It is computed by

$$\text{RMSE} = \sqrt{\frac{1}{n_x}(x - \hat{x})'(x - \hat{x})}, \tag{29}$$

where $n_x = 1$ is the size of the state-space for this example. The fourth column represents the Monte Carlo averaged filter consistencies using the Scaled Normalized Estimation Error Squared

(SNEES) and is computed by

$$\text{SNEES} = \frac{1}{n_x}(x - \hat{x})' \hat{P}^{-1}(x - \hat{x}). \tag{30}$$

A SNEES value of 1 means good filter consistency. Anything less than 1 indicates that the filter is too conservative and anything greater than 1 indicates that the filter is too confident. Table 1 indicates that each method computes the exact same statistics, further supporting the justification, numerically, that each method is exactly equivalent.

For linear measurement models, it is now sufficient to conclude that there are three exactly equivalent methods to calculate the weights of the posterior distribution. All three methods are optimal and produce exact Bayesian posteriors for linear, Gaussian mixture systems.

## NONLINEAR MODELS

For the nonlinear case, things are not so clean-cut-and-dried. Instead, there is no analytical general proof that can say one method is better than another. It is model specific. In this Section, this work demonstrates empirically that, for certain measurement models, the presented anti-linearization weights are superior.

Table 2 compares the three different weighting methods for the scalar nonlinear measurement model $h(x) = x^3$, averaging results over 10,000 Monte Carlo simulations. The results were generated using the same truth, prior, and measurement generation as the linear case.

**Table 2**: Comparing the different weighting methods for the scalar nonlinear measurement model example $h(x) = x^3$. Each component is individually updated with an Extended Kalman Filter (EKF).

|  | Error | Cov. | RMSE | SNEES |
|---|---|---|---|---|
| Traditional | 0.1342600 | 0.0535870 | 0.1366000 | 0.52332 |
| Post. Lin. | 0.0401440 | 0.0267740 | 0.0519450 | 1.56720 |
| Anti-Lin. | **0.0020463** | **0.0016690** | **0.0053505** | **1.04800** |

Using 100 GMM components and averaged over 10,000 Monte Carlo simulations.

These results now express that the anti-linearization weights have smaller error and uncertainty, with better consistency than the prior and posterior linearization methods. However, this is only for $n = 100$, which is a specific number of components. So, to further validate its superiority, Figure 1 and Figure 2 plot the Monte Carlo averaged RMSE and SNEES, respectively, as the number of Gaussian mixture components $n$ varies between simulations.

Figure 1 shows the RMSE results and indicates the anti-linearization and posterior linearization weights demonstrate superior accuracy compared to traditional prior linearization weights as the number of components $n$ varies. Although the posterior linearization weights achieve slightly lower RMSE as $n$ increases, the anti-linearization weights outperform them when fewer components are used. This suggests that the anti-linearization weights are more computationally efficient, offering better accuracy with fewer components.

Figure 2 shows the SNEES results. Although both the anti-linearization and posterior linearization weights demonstrate improved consistency over their traditional prior linearization counterpart,
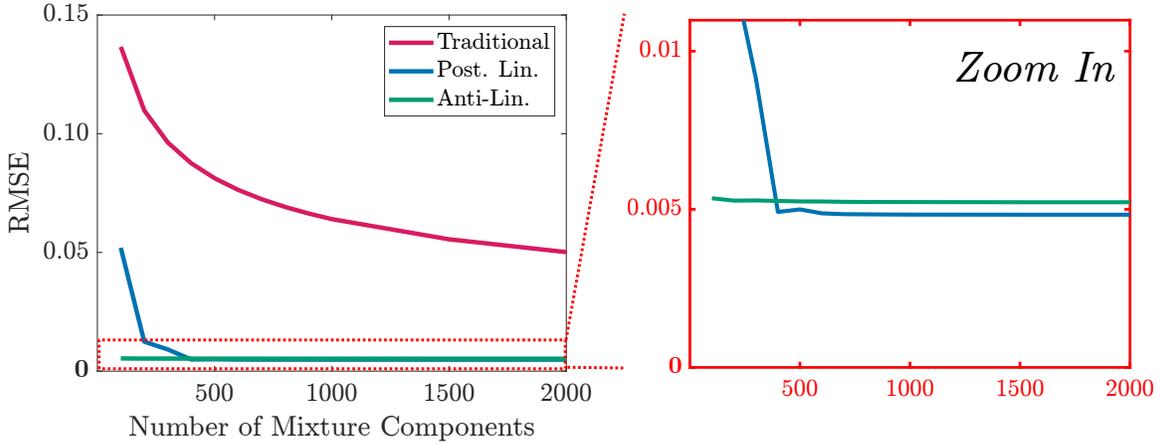
**Figure 1**: Comparing the different weighting methods for the scalar nonlinear measurement model example $h(x) = x^3$. This figure shows the root mean square error (RMSE) vs. the number of mixture components of the compared methods: Traditional weights (red), Posterior Linearization weights (blue), and Anti-Linearization weights (green). Averaged over 10,000 Monte Carlo simulations.
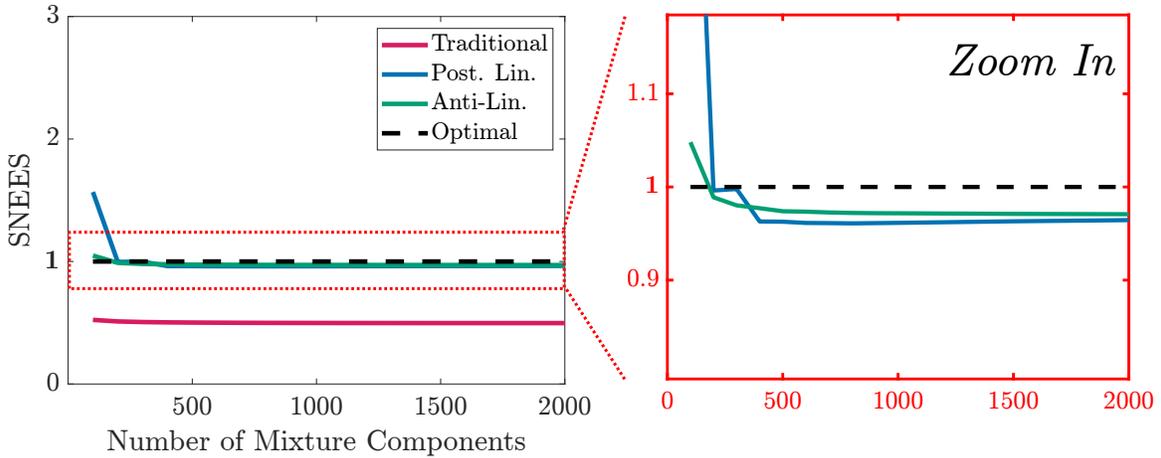


**Figure 2**: Comparing the different weighting methods for the scalar nonlinear measurement model example $h(x) = x^3$. This figure shows the scaled normalized estimation error squared (SNEES) vs. the number of mixture components of the compared methods: Traditional weights (red), Posterior Linearization weights (blue), and Anti-Linearization weights (green). The optimal SNEES of 1 is also shown (black dashed). Averaged over 10,000 Monte Carlo simulations.

the anti-linearization approach has slightly better consistency than the posterior linearization. This is likely due to its avoidance of linearizing the measurement model, thereby preserving more system information. Furthermore, the anti-linearization weights achieve better consistency with fewer components when compared to the other methods.

For nonlinear measurement models, the performance comparison between the different weight-

ing methods becomes more apparent. The anti-linearization weights show improved accuracy and consistency, for certain models, over both the prior and posterior linearization methods. These empirical results demonstrate that, for nonlinear measurement models, the anti-linearization weights have superior accuracy, consistency, and computational efficiency because they achieve better performance using fewer Gaussian mixture components.

## IMPROVED WEIGHTS FOR THE GM-PHD FILTER

The PHD filter, introduced by Reference 31, is based on the Random Finite Set (RFS) theory. It propagates the first factorial moment density (*a.k.a.* the first-order moment or *a.k.a.* the intensity function) of the RFS, rather than the full multi-target posterior distribution. The intensity function is the expected number of points found in a unit volume of the space and is based on Poisson point processes.[26]

A Poisson point process is a type of random process where points are scattered randomly in a space, according to a Poisson distribution. In essence, the PHD provides a density function describing the distribution of multiple targets in a given space, embodying statistical properties akin to those found in Poisson point processes. By doing so, the PHD filter is a computationally feasible alternative to the more complex Multi-Target Bayesian filter, as it avoids the combinatorial complexity of data association.

The intensity function $v(\mathbf{x})$ is defined such that the integral over any region $S$ in the state-space gives the expected number of targets $\hat{N}$ in that region:

$$\hat{N} = \int_S v(\mathbf{x}) \, d\mathbf{x}. \tag{31}$$

The value $v(\mathbf{x}) \, d\mathbf{x}$ is the expected number of targets in an infinitesimally small region $d\mathbf{x}$ of $\mathbf{x}$, *i.e.* $v(\mathbf{x})$ is the expected target density at $\mathbf{x}$.

### Gaussian Mixture Implementation

The Gaussian mixture (GM) implementation of the PHD filter[24] approximates the intensity function as a weighted sum of Gaussian components, allowing for an analytically tractable solution. The GM-PHD filter represents the posterior intensity function at time $k - 1$ as

$$v_{k-1|k-1}(\mathbf{x}) \approx \sum_{i=1}^{J_{k-1|k-1}} w_{k-1|k-1}^{(i)} \mathcal{N}(\mathbf{x}; \hat{\mathbf{x}}_{k-1|k-1}^{(i)}, \hat{\mathbf{P}}_{k-1|k-1}^{(i)}), \tag{32}$$

where $w_{k-1|k-1}^{(i)}$ is the weight, $\hat{\mathbf{x}}_{k-1|k-1}^{(i)}$ is the estimated mean, and $\hat{\mathbf{P}}_{k-1|k-1}^{(i)}$ is the estimated covariance of the $i$-th Gaussian component at time $k-1$, and $J_{k-1|k-1}$ is the number of components.

**Remark 1.** *The posterior intensity from Equation* (32) *is exactly a Gaussian mixture for linear dynamics and measurements.[23, 24] However, this work represents it as an approximation because of nonlinear dynamics and measurements.*

*Prediction Step.* The prediction step of the GM-PHD filter involves predicting the intensity function from time $k - 1$ to time $k$. This includes predicting the birth of new targets, the survival of existing targets, and the motion model of targets [*]:

$$v_{k|k-1}(\mathbf{x}) = v_{B,k}(\mathbf{x}) + v_{S,k|k-1}(\mathbf{x}), \tag{33}$$

---

[*]This work does not touch on spawning. Please refer to References 24 and 23.

where $v_{B,k}(\mathbf{x})$ is the birth intensity accounting for components that get created at $k$ and $v_{S,k|k-1}(\mathbf{x})$ is the survival intensity accounting for components surviving the propagation step from $k-1$ to $k$.

For the nonlinear GM-PHD filter, the prediction step is expressed as:

$$v_{k|k-1}(\mathbf{x}) \approx \sum_{i=1}^{J_{B,k}} w_{B,k}^{(i)} \mathcal{N}(\mathbf{x}; \mathbf{x}_{B,k}^{(i)}, \mathbf{P}_{B,k}^{(i)}) + p_{S,k} \sum_{i=1}^{J_{k-1|k-1}} w_{k|k-1}^{(i)} \mathcal{N}(\mathbf{x}; \hat{\mathbf{x}}_{S,k|k-1}^{(i)}, \hat{\mathbf{P}}_{S,k|k-1}^{(i)}),$$

$$w_{k|k-1}^{(i)} = w_{k-1|k-1}^{(i)},$$

$$\hat{\mathbf{x}}_{S,k|k-1}^{(i)} = f_{k|k-1}(\hat{\mathbf{x}}_{k-1|k-1}^{(i)}),$$

$$\hat{\mathbf{P}}_{S,k|k-1}^{(i)} = \mathbf{F}_{k|k-1} \hat{\mathbf{P}}_{k-1|k-1}^{(i)} \mathbf{F}_{k|k-1}^T + \mathbf{Q},$$

$$(34)$$

where $J_{B,k}$, $w_{B,k}^{(i)}$, $\mathbf{x}_{B,k}^{(i)}$, $\mathbf{P}_{B,k}^{(i)}$, $i = 1, \cdots, J_{B,k}$ are given model parameters that determine the shape of the birth intensity; $J_{k-1|k-1}$, $w_{k|k-1}^{(i)}$, $\hat{\mathbf{x}}_{S,k|k-1}^{(i)}$, $\hat{\mathbf{P}}_{S,k|k-1}^{(i)}$, $i = 1, \cdots, J_{k-1|k-1}$ are the survival intensity components from $k-1$ to $k$. Also, $p_{S,k}$ is the probability of survival, $\mathbf{F}_{k|k-1}$ is the state transition matrix, and $\mathbf{Q}$ is the process noise covariance.

*Update Step.* The update step incorporates the measurements at time $k$ to update the intensity function. The updated intensity function is given by:

$$v_{k|k}(\mathbf{x}) = (1 - p_{D,k})v_{k|k-1}(\mathbf{x}) + \sum_{\mathbf{y} \in \mathcal{Y}_k} v_{D,k|k}(\mathbf{x}|\mathbf{y}),$$

$$v_{D,k|k}(\mathbf{x}|\mathbf{y}) = \sum_{i=1}^{J_{k|k-1}} w_{k|k}^{(i)} \mathcal{N}(\mathbf{x}; \hat{\mathbf{x}}_{k|k}^{(i)}, \hat{\mathbf{P}}_{k|k}^{(i)}),$$

$$w_{k|k}^{(i)} = \frac{p_{D,k} w_{k|k-1}^{(i)} \xi_k^{(i)}}{\lambda_{C,k} p_{C,k} + p_{D,k} \sum_{j=1}^{J_{k|k-1}} w_{k|k-1}^{(j)} \xi_k^{(j)}},$$

$$\hat{\mathbf{x}}_{k|k}^{(i)} = \hat{\mathbf{x}}_{k|k-1}^{(i)} + \mathbf{K}_k^{(i)}(\mathbf{y} - h(\hat{\mathbf{x}}_{k|k-1}^{(i)})),$$

$$\hat{\mathbf{P}}_{k|k}^{(i)} = \hat{\mathbf{P}}_{k|k-1}^{(i)} - \mathbf{K}_k^{(i)} \mathbf{H}_k^{(i)} \hat{\mathbf{P}}_{k|k-1}^{(i)},$$

$$\mathbf{K}_k^{(i)} = \hat{\mathbf{P}}_{k|k-1}^{(i)}(\mathbf{H}_k^{(i)})^T (\mathbf{S}_k^{(i)})^{-1},$$

$$\mathbf{S}_k^{(i)} = \mathbf{H}_k^{(i)} \hat{\mathbf{P}}_{k|k-1}^{(i)}(\mathbf{H}_k^{(i)})^T + \mathbf{R},$$

$$\xi_k^{(i)} = \mathcal{N}\left(\mathbf{y}; h(\hat{\mathbf{x}}_{k|k-1}^{(i)}), \mathbf{S}_k^{(i)}\right),$$

$$(35)$$

where $p_{D,k}$ is the probability of detection, $\mathbf{H}_k^{(i)}$ is the state-measurement mapping matrix, $\mathbf{R}$ is the measurement noise covariance, $\mathcal{Y}_k \subset \mathbb{R}^{n_y}$ is the set of measurements at time $k$, $\lambda_{C,k}$ is the average number of clutter measurements per time step, and $p_{C,k}$ is the spatial distribution of clutter on the surveillance area (clutter density).

**Remark 2.** *In this work, the GM-PHD uses the EKF[34, 35] for prediction and updates, although it can easily be extended to use other filters.*

The combinatorial explosion of the algorithm is managed by 3 key processes: pruning, merging, and capping. Pruning involves removing components that have low weights, thereby reducing the

computational load. Merging, on the other hand, combines components that are in close proximity to each other, simplifying the overall structure. Capping prevents the number of components from exceeding a predefined set maximum. To estimate target states from the mixture (called state extraction), the components with the highest weights are selected. To estimate the number of targets $\hat{N}$ (cardinality), the posterior weights are summed. For an in-depth explanation of these procedures and more, refer to the detailed discussion and pseudocode provided in.[23,24]

**Traditional GM-PHD Weights**

The traditional weights of the GM-PHD are expressed in Equation (35) where $\xi_k^{(i)}$ represents the marginal likelihood of the $i$-th Gaussian component:

$$\xi_k^{(i)} \;=\; \mathcal{N}\left(\mathbf{y}; h(\hat{\mathbf{x}}_{k|k-1}^{(i)}), \mathbf{S}_k^{(i)}\right), \tag{36}$$

where $\hat{\mathbf{x}}_{k|k-1}^{(i)}$ is the prior associated with this component and $\mathbf{S}_k^{(i)}$ is its innovation covariance matrix:

$$\mathbf{S}_k^{(i)} \;=\; \mathbf{H}_k^{(i)}\hat{\mathbf{P}}_{k|k-1}^{(i)}(\mathbf{H}_k^{(i)})^T \;+\; \mathbf{R}, \tag{37}$$

and $\mathbf{H}_k^{(i)} = \frac{\partial h(\hat{\mathbf{x}}_{k|k-1}^{(i)})}{\partial \hat{\mathbf{x}}_{k|k-1}^{(i)}}$ is the Jacobian of the measurement model evaluated at the prior estimate.

**Posterior Linearization GM-PHD Weights**

The posterior linearization weights have only been shown to be used with single-target tracking filters.[21,32] This work now proposes a way to incorporate them into the GM-PHD framework:

$$\xi_k^{(i)} \;=\; \mathcal{N}\left(\mathbf{y}; h(\hat{\mathbf{x}}_{k|k}^{(i)}), \hat{\mathbf{S}}_k^{(i)}\right), \tag{38}$$

where $\hat{\mathbf{x}}_{k|k}^{(i)}$ is the posterior associated with this component and $\hat{\mathbf{S}}_k^{(i)}$ is its innovation covariance matrix according to Equation (14):

$$\begin{aligned} \hat{\mathbf{S}}_k^{(i)} \;=\;\; & (\hat{\mathbf{H}}_k^{(i)} - \mathbf{H}_k^{(i)})\hat{\mathbf{P}}_{k|k}^{(i)}(\hat{\mathbf{H}}_k^{(i)} - \mathbf{H}_k^{(i)})' \\ & + \; (\mathbf{I} - \mathbf{H}_k^{(i)}\mathbf{K}_k^{(i)})\mathbf{S}_k^{(i)}(\mathbf{I} - \mathbf{H}_k^{(i)}\mathbf{K}_k^{(i)})', \end{aligned} \tag{39}$$

and $\hat{\mathbf{H}}_k^{(i)} = \frac{\partial h(\hat{\mathbf{x}}_{k|k}^{(i)})}{\partial \hat{\mathbf{x}}_{k|k}^{(i)}}$ is the Jacobian of the measurement model evaluated at the posterior estimate. Also, $\hat{\mathbf{P}}_{k|k}^{(i)}$ is the posterior covariance of the $i$-th component, $\mathbf{I}$ is the $n_y \times n_y$ identity, and $\mathbf{K}_k^{(i)} = \hat{\mathbf{P}}_{k|k-1}^{(i)}\mathbf{H}_k^{(i)'}\mathbf{S}_k^{(i)-1}$ is the Kalman gain.

**Anti-Linearization GM-PHD Weights**

This work also proposes incorporating the anti-linearization weights into the GM-PHD:

$$\xi_k^{(i)} \;=\; \frac{p_i(\mathbf{y}|\mathbf{x})\, p_i(\mathbf{x})}{p_i(\mathbf{x}|\mathbf{y})}, \tag{40}$$

and evaluating $\mathbf{x} = \hat{\mathbf{x}}_{k|k}^{(i)}$ gives

$$\xi_k^{(i)} \;=\; \mathcal{N}(\mathbf{y};\, h(\hat{\mathbf{x}}_{k|k}^{(i)}),\, \mathbf{R})\, \mathcal{N}(\hat{\mathbf{x}}_{k|k}^{(i)};\, \hat{\mathbf{x}}_{k|k-1}^{(i)},\, \hat{\mathbf{P}}_{k|k-1}^{(i)}), \tag{41}$$

where the posterior denominator in Equation (40) cancels out when evaluating at the posterior estimate in Equation (41).

## CISLUNAR MULTI-TARGET FILTERING EXAMPLE

Cislunar NRHOs are a class of trajectories that exist in the vicinity of the Moon and are typically positioned close to Lagrange points associated with the Earth-Moon system. NRHOs offer a degree of long-term stability such that, on average, the orbit remains relatively stable over extended periods. This stability is valuable for missions requiring long-duration observations or operations.

### Problem Setup

This work focuses on a scenario where 10 targets are generated from a debris cloud clustered around a 9:2 synodic resonant NRHO. (See Fig. 3 for an illustration.) This specific type of NRHO has a 6.5-day period, a perilune radius of about 3,250 km, and an apolune radius of approximately 71,000 km. It is the lowest-altitude NRHO with a useful resonance, and serves as the baseline orbit for NASA's Lunar Gateway mission.[54]
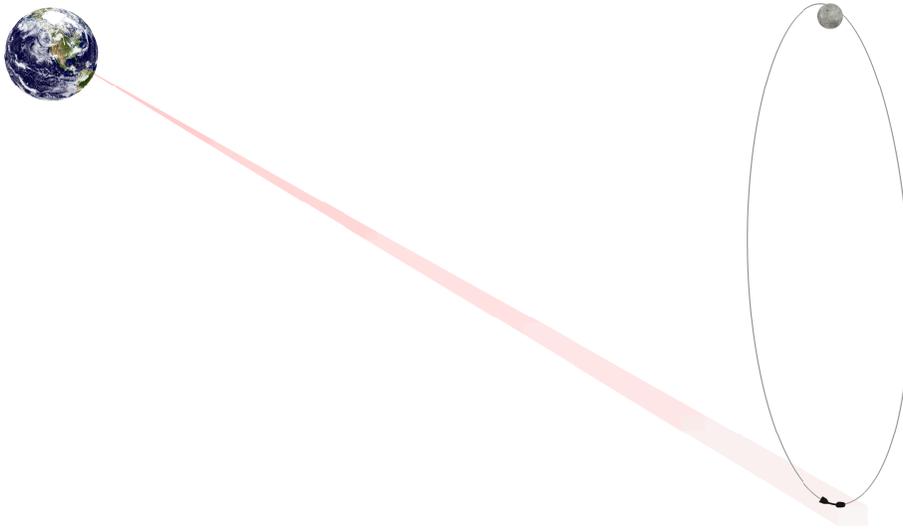


**Figure 3**: An illustration of target trajectories clustered around apolune following a 9:2 synodic resonant NRHO. Measurements are mapped to the Barycenter of the system making it irrelevant whether the measurements come from Earth-based or space-based tracking systems.

Given the high demand for advanced tracking systems like the Deep Space Network (DSN) and the Space Surveillance Network (SSN), the assumption of available range measurements is no longer viable. Consequently, alternative tracking systems are employed and rely solely on angle measurements. These measurements are mapped to the Barycenter of the system making it irrelevant whether the measurements come from Earth-based or space-based tracking systems. In this simulation, targets are assumed trackable such that they have sufficient size and reflectivity. Tracking passes occur close to apolune every orbit for five orbits. During these tracking passes, angle measurements are available every 10 minutes for eight hours.

*Dynamics Model.* This work models cislunar NRHO dynamics using the Circular Restricted Three Body Problem (CR3BP) for the Earth-Moon system with a six dimensional state-space rep-

resented by $\mathbf{x} = [\mathbf{r}(1), \mathbf{r}(2), \mathbf{r}(3), \mathbf{v}(1), \mathbf{v}(2), \mathbf{v}(3)]'$:

$$\dot{\mathbf{r}}(1) = \mathbf{v}(1), \qquad \dot{\mathbf{r}}(2) = \mathbf{v}(2), \qquad \dot{\mathbf{r}}(3) = \mathbf{v}(3),$$

$$\dot{\mathbf{v}}(1) = \mathbf{r}(1) + 2\mathbf{v}(2) - \frac{(1-\mu)(\mathbf{r}(1)+\mu)}{r_\oplus^3} - \frac{\mu(\mathbf{r}(1)-1+\mu)}{r_\mathbb{C}^3},$$

$$\dot{\mathbf{v}}(2) = \mathbf{r}(2) - 2\mathbf{v}(1) - \frac{(1-\mu)\mathbf{r}(2)}{r_\oplus^3} - \frac{\mu\mathbf{r}(2)}{r_\mathbb{C}^3}, \qquad (42)$$

$$\dot{\mathbf{v}}(3) = -\frac{(1-\mu)\mathbf{r}(3)}{r_\oplus^3} - \frac{\mu\mathbf{r}(3)}{r_\mathbb{C}^3},$$

where $\mathbf{r}(1), \mathbf{r}(2), \mathbf{r}(3)$ and $\mathbf{v}(1), \mathbf{v}(2), \mathbf{v}(3)$ represent the scaled Cartesian positions and velocities of the target with-respect-to the Barycenter origin, $\mu$ is the scaled Moon geocentric gravitational constant, and $r_\oplus$ and $r_\mathbb{C}$ are the distances of the target with-respect-to the Earth and Moon in the Barycenter reference frame:

$$\mu = \frac{\mu_\mathbb{C}}{\mu_\oplus + \mu_\mathbb{C}}, \qquad (43)$$

$$r_\oplus = \sqrt{(\mathbf{r}(1)+\mu)^2 + \mathbf{r}(2)^2 + \mathbf{r}(3)^2}, \qquad (44)$$

$$r_\mathbb{C} = \sqrt{(\mathbf{r}(1)-1+\mu)^2 + \mathbf{r}(2)^2 + \mathbf{r}(3)^2}. \qquad (45)$$

In this work, $\mu_\oplus = G \cdot m_\oplus$ and $\mu_\mathbb{C} = G \cdot m_\mathbb{C}$. The gravitational constant is $G = 6.6743 \times 10^{-11}$ m3/s2/kg, the mass of the Earth is $m_\oplus = 5.972 \times 10^{24}$ kg, and the mass of the Moon is $m_\mathbb{C} = 7.342 \times 10^{22}$ kg. The units for distance and time are non-dimensionalized by length units LU $= 384400 \times 10^3$ m and time units TU $= \sqrt{\text{LU}^3/(\mu_\oplus + \mu_\mathbb{C})}$ s. The non-dimensional (denoted [-]) period of each target is roughly 1.3632096570 [-] from Reference 55. The system dynamic equations are numerically integrated with an embedded Runge-Kutta 8(7) method.[56]

*Measurement Model.* The measurement vector $\mathbf{y} = [\alpha, \varepsilon]'$ contains azimuth $\alpha$ and elevation $\varepsilon$ of the observed target mapped to the Barycenter origin:

$$\alpha = \tan^{-1}\left(\frac{\mathbf{r}(2) - \mathbf{r}_\oplus(2)}{\mathbf{r}(1) - \mathbf{r}_\oplus(1)}\right), \qquad (46)$$

$$\varepsilon = \sin^{-1}\left(\frac{\mathbf{r}(3) - \mathbf{r}_\oplus(3)}{\|\mathbf{r} - \mathbf{r}_\oplus\|}\right), \qquad (47)$$

where $\|\cdot\|$ is the Euclidean 2-norm, and $\mathbf{r}_\oplus = [\mathbf{r}_\oplus(1), \mathbf{r}_\oplus(2), \mathbf{r}_\oplus(3)]'$ is the position of the tracking system, which in this work is at the Barycenter defined to be the origin: $\mathbf{r}_\oplus = \mathbf{0}$. The measurements are corrupted by additive zero-mean Gaussian white noise with $1\sigma$ azimuth and elevation uncertainties of 16.1 arc-seconds for both. Light travel time delay and measurement biases are not considered.

*Parameters.* Each target's truth and each multi-target filtering method are initialized by the same distribution centered at the non-dimensional coordinates[55] and covariance

$$\hat{\mathbf{x}}_0 = [1.0110350588, \ 0, \ -0.1731500000, \qquad (48)$$
$$\phantom{\hat{\mathbf{x}}_0 =} 0, \ -0.0780141199, \ 0]^\mathrm{T}$$

$$\hat{\mathbf{P}}_0 = \text{diag}([1 \times 10^{-4}, \ 1 \times 10^{-4}, \ 1 \times 10^{-4}, \qquad (49)$$
$$\phantom{\hat{\mathbf{P}}_0 =} 1 \times 10^{-6}, \ 1 \times 10^{-6}, \ 1 \times 10^{-6}]^2).$$

The GM-PHD filter is initialized with a single PHD component of extremely small weight $w_{0|0}^{(1)} = 2.2204 \times 10^{-16}$. The GM-PHD uses the same pruning, merging, capping, and state extraction techniques documented in References 24 and 23. For this scenario, the filter uses a pruning threshold of $1 \times 10^{-5}$, a merging threshold of 4, and a capping limit to the number of PHD components $J_{\max} = 100$.

Clutter exists and is modeled on a Poisson RFS: the average rate of uniform clutter is $\lambda_{C,k} = 1.0$, generated in the surveillance region $\{(\mathbf{r}(1), \mathbf{r}(2), \mathbf{r}(3), \mathbf{v}(1), \mathbf{v}(2), \mathbf{v}(3)) \mid 1.0082 \leqslant \mathbf{r}(1) \leqslant 1.0137,\ -0.0086 \leqslant \mathbf{r}(2) \leqslant 0.0025,\ -0.1758 \leqslant \mathbf{r}(3) \leqslant -0.1691,\ -0.0174 \leqslant \mathbf{v}(1) \leqslant 0.0098,\ -0.0878 \leqslant \mathbf{v}(2) \leqslant -0.0672,\ -0.0098 \leqslant \mathbf{v}(3) \leqslant 0.0510\}$. The uniform clutter density is then $p_{C,k} = 2.4437 \times 10^6$. After being generated, clutter is mapped into measurements using the measurement model from Equations (46) and (47).

The probability of detection is $p_{D,k} = 0.9$ and the probability of survival is $p_{S,k} = 0.99999$. This work does not perform measurement gating. The filter assumes some process noise (although the modeled truths do not): $\mathbf{Q} = 1 \times 10^{-8}[\text{-}] \cdot \mathbf{I}_{n_x}$. The birth model is a single Gaussian with $1\sigma$ position and velocity uncertainties of 1000 km ($2.6 \times 10^{-3}$ [-]) and 0.01 km/s ($9.8 \times 10^{-3}$ [-]), respectively, centered at the initial mean in Equation (48):

$$\mathbf{x}_{B,k} = \hat{\mathbf{x}}_0, \tag{50}$$

$$\mathbf{P}_{B,k} = \mathrm{diag}([2.6 \times 10^{-3},\ 2.6 \times 10^{-3},\ 2.6 \times 10^{-3}, \tag{51}$$
$$9.8 \times 10^{-3},\ 9.8 \times 10^{-3},\ 9.8 \times 10^{-3}]^2),$$

where at every filter update step, after propagation, $J_{B,k} = 10$ components are sampled from this birth model growing the PHD Gaussian mixture. The weight of each new birthed component is $w_{B,k} = 1/100$.

*Multi-Target Filter Accuracy Criterion.* It is common practice to use Optimal Subpattern Assignment (OSPA)[57] to evaluate multi-target filter performance. In this work, the OSPA computes the filter error of the state estimates with-respect-to the truth. It is a combination of the localization and cardinality error components of the finite subsets $\hat{X} = \{\hat{\mathbf{x}}^{(1)}, \cdots, \hat{\mathbf{x}}^{(m)}\}$ (the PHD extracted target estimates) and $X = \{\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(n)}\}$ (the target truths):

$$\bar{d}_p^{(c)} := \left( \frac{1}{\max(m,n)} \left( \min_{\pi \in \Pi_n} \sum_{i=1}^{m} d^{(c)}(\hat{\mathbf{x}}^{(i)}, \mathbf{x}^{\pi(i)})^p + c^p(n-m) \right) \right)^{\frac{1}{p}}, \tag{52}$$

where $\Pi_n$ represents the set of permutations on $\{1, 2, \cdots, k\}$, where $k \in \mathbb{N} = \{1, 2, \cdots\}$ and the notation $d^{(c)}(\cdot, \cdot) \triangleq \min(c, d(\cdot, \cdot))$. The metric $d(\cdot, \cdot)$ can be any metric as defined by Reference 57. This work chooses $d(\cdot, \cdot)$ to be the Euclidean 2-norm of the error:

$$d(\hat{\mathbf{x}}^{(i)}, \mathbf{x}^{\pi(i)}) = \|\hat{\mathbf{x}}^{(i)} - \mathbf{x}^{\pi(i)}\|, \quad \text{with values in } [0, c]. \tag{53}$$

In this work, the OSPA parameter $p = 2$ which is common practice and a cut-off value $c = 1 \times 10^4$ km. A low OSPA value means good filter and tracking accuracy. Assignment of estimate $\hat{\mathbf{x}}$ with truth $\mathbf{x}$ is performed by the Hungarian (*a.k.a.* Munkres) algorithm[58,59] to minimize the summation in Equation (52). Similar procedures and more details for computing the OSPA can be found in References 57, 24, and 23.

**Results**

Figures 4 and 5 visualize the filter accuracy and cardinality accuracy, respectively. Figure 4 shows the 3-dimensional extracted position state estimates and Figure 5 shows the extracted cardinality estimates of the compared weighting methods. Both figures show the results of all 100 Monte Carlo simulations as transparent data points with a single run without transparency for emphasis.
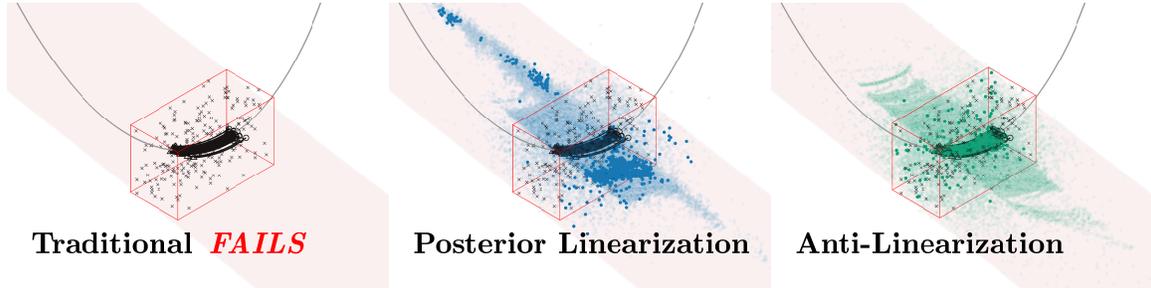


**Figure 4**: For the cislunar NRHO example, this figure shows the 3D true trajectories of the 10 targets (black) and the extracted position state estimates of the compared methods: Traditional weights (red), Posterior Linearization weights (blue), and Anti-Linearization weights (green). The clutter is represented by the gray crosses which are Poisson distributed in the red rectangular surveillance region. These are the results of 100 Monte Carlo simulations with a single run emphasized to illustrate.
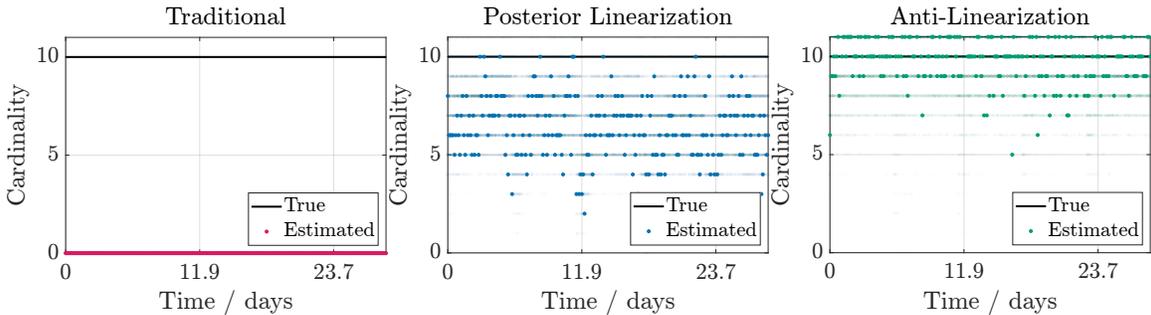


**Figure 5**: For the cislunar NRHO example, this figure shows the true cardinality (black) and the extracted cardinality estimates of the compared methods: Traditional weights (red), Posterior Linearization weights (blue), and Anti-Linearization weights (green). These are the results of 100 Monte Carlo simulations with a single run emphasized to illustrate.

The GM-PHD filter using the traditional weights fails to estimate any targets in the cluttered surveillance region. A visualization of this is in Figure 4 by no apparent estimates appearing at all. This is additionally supported by Figure 5, which shows this method also failing to estimate any targets resulting in an extracted cardinality estimate of zero. The GM-PHD filter using the traditional weights fails because the weights crash below the state extraction threshold. There is not enough information about the system to keep them afloat, thus the apparent missed detections.

Evidently, by simply switching the weight computation, the GM-PHD filters using the posterior linearization and anti-linearization weights, save the GM-PHD filter from failing and in fact can

give rather good tracking performance. The weights no longer crash and do not get discarded by the prune, merge, capping scheme. The weights essentially stay afloat.

Although better than the traditional prior linearization weights, the GM-PHD filter using the posterior linearization approach does not perform as well as the proposed GM-PHD filter using the anti-linearization weights, which has better state estimation and cardinality estimation. Figure 4, shows visually that the proposed anti-linearization approach produces more precise state estimates of the 10 targets inside the surveillance region. Additionally, it provides more precise cardinality estimates, shown in Figure 5, often estimating correctly that 10 targets actually exist.

To further support the findings of Figures 4 and 5, OSPA is used as a metric to quantify the effectiveness of both filter accuracy and cardinality accuracy. Figure 6 presents the OSPA accuracy of the extracted state estimates in comparison to the truth.
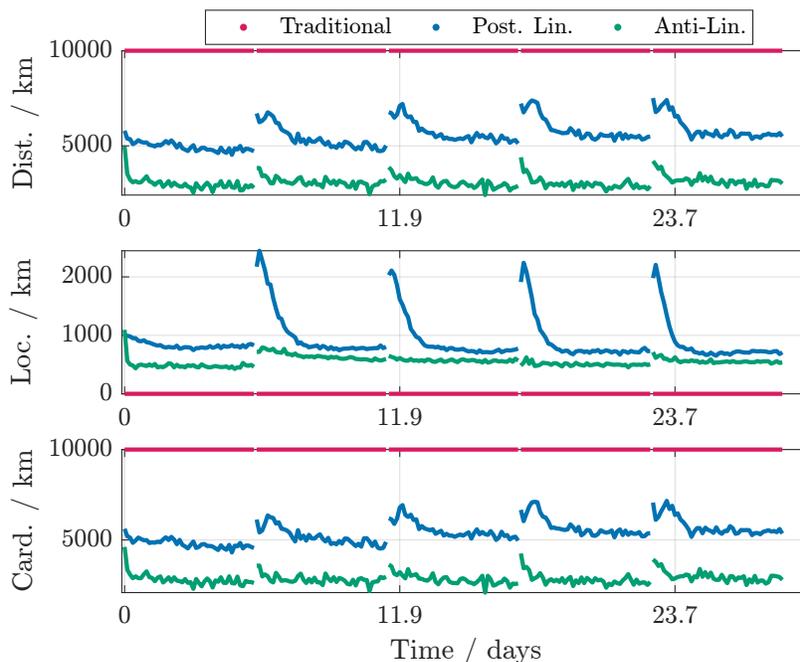


**Figure 6**: For the cislunar NRHO example, this figure shows the Optimal Subpattern Assignment (OSPA) between the extracted position estimates of the compared methods and the truths. It shows the position errors of the cardinality OSPA component versus time (bottom), the localization OSPA component versus time (middle), and the combined cardinality–localization OSPA distance versus time (top). Breaks indicate different orbit passes. Averaged over 100 Monte Carlo simulations.

Because the GM-PHD filter using the traditional weights outright fails, its OSPA is maximized to the cut-off value $c = 1 \times 10^4$ km. Again, by simply switching the weight computation, the GM-PHD filter using the posterior linearization weights provides reasonable results. Despite having a reasonable OSPA, it does show large spikes in the localization component at the beginning of each orbit pass. This is likely due to the linearization about the posterior. Although the localization component struggles early, the cardinality component does not, and shows stable performance over each orbit pass.

17

The GM-PHD filter using the anti-linearization weights shows the best OSPA performance. It has sustained low localization and cardinality components, indicating stable estimation over several orbit passes. This proposed method outperforms the prior and posterior linearization methods resulting in the best overall OSPA, indicating the best multi-target filtering accuracy.

The last piece of evidence for this scenario is the computational speed of each weighting method presented in Figure 7. This figure shows the average wall-clock times it took to run each Monte Carlo simulation. The GM-PHD filter using the traditional weights takes the least time because it carried the least number of PHD components, but this is because it fails to estimate any targets, therefore its time is disregarded and grayed-out in the figure.

The GM-PHD filter using the anti-linearization weights is faster than the filter using the posterior linearization weights. This is because it eliminates the need for linearization and does not compute Jacobians and innovation matrices, which take up more computational resources. Additionally, the anti-linearization approach has better cardinality estimates which means it carried less PHD components to characterize all the hypotheses. It is then sufficient to say that the GM-PHD filter using the proposed anti-linearization weights provides the most accurate state and cardinality estimates while also being the fastest of the compared methods.
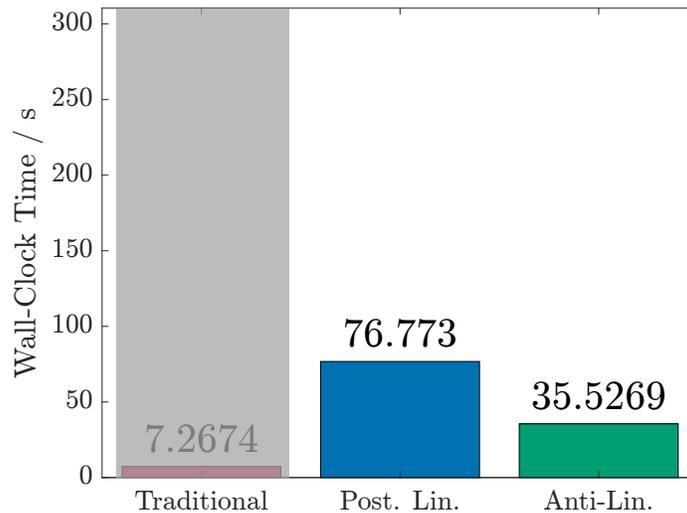


**Figure 7**: For the cislunar NRHO example, these are the wall-clock times of the compared methods. Averaged over 100 Monte Carlo simulations. Since the Traditional weights method fails, its time is disregarded. Using an AMD Ryzen 5 5500U CPU and 8 GB of RAM.

**CONCLUSION**

The Gaussian Mixture Probability Hypothesis Density (GM-PHD) filter is used to accurately estimate the amount, location, and trajectory of simulated space debris. The key driving factor of the GM-PHD is that it approximates the intensity function as a weighted sum of Gaussian mixture components to provide an analytically tractable solution. However, issues arise when the weights are computed inaccurately. This then causes filter divergence, degeneracy, and ultimately, erroneous state estimates.

Traditionally, weights were updated by linearizing the measurement model around prior means. However, recent advancements have improved this by linearizing around posteriors, enhancing estimation performance without sacrificing efficiency——though only for single-target tracking. This work extends the posterior linearization technique to multi-target filtering within the GM-PHD filter.

This work also proposes weights that eliminate the need for linearization entirely. The findings demonstrate their exact equivalence to traditional and recent methods under linear measurement models and superior performance under nonlinear models. The proposed weights, applicable to all Gaussian mixture-type filters, are also integrated into the GM-PHD filter. Empirical results show that they significantly improve the GM-PHD filter's accuracy, consistency, and computational efficiency.

Future work could explore using different samples from the posterior rather than just the mean for the proposed weights. Different sampling methods could be explored such as random sampling or deterministic sampling like quasi-Monte Carlo methods or Fibonacci grids. By continuing the research of Gaussian mixture weight updates it ensures the long-term sustainability of space exploration and satellite operations amid the growing threat of space debris.

## ACKNOWLEDGMENTS

## REFERENCES

[1] D. J. Kessler and B. G. Cour-Palais, "Collision frequency of artificial satellites: The creation of a debris belt," *Journal of Geophysical Research: Space Physics*, Vol. 83, No. A6, 1978, pp. 2637–2646.

[2] D. J. Kessler and J. Loftus Jr, "Orbital debris as an energy management problem," *Advances in Space Research*, Vol. 16, No. 11, 1995, pp. 139–144.

[3] V. A. Chobotov, D. Spencer, D. Schmitt, R. Gupta, R. Hopkins, and D. Knapp, "Dynamics of debris motion and the collision hazard to spacecraft resulting from an orbital breakup," *Aerospace Corp. TR SD-TR-88-96, El Segundo, CA*, 1988.

[4] G. D. Badhwar and P. D. Anz-Meador, "Determination of the area and mass distribution of orbital debris fragments," *Earth, Moon, and Planets*, Vol. 45, No. 1, 1989, pp. 29–51.

[5] A. Tan and R. C. Reynolds, *Theory of satellite fragmentation in orbit*. World Scientific, 2019.

[6] M. Landgraf and R. Jehn, "Space debris hazards from explosions in the collinear Sun-Earth Lagrange points," *arXiv preprint astro-ph/0104277*, 2001.

[7] P. Bandyopadhyay, R. K. Sharma, and A. Tewari, "Space debris hazards from fragmentations in collinear Earth-Moon points," *Moon*, Vol. 1, 2009, p. L2.

[8] N. Boone and R. A. Bettinger, "Debris propogation following a spacecraft mishap at the collinear Earth-Moon Lagrange points," *2021 IEEE Aerospace Conference*, IEEE, 2021, pp. 1–15.

[9] N. R. Boone and R. A. Bettinger, "Artificial debris collision risk following a catastrophic spacecraft mishap in lunar orbit," *2021 Advanced Maui Optical Surveillance Conference (AMOS) Proceedings*, 2021.

[10] A. Wilmer, N. Boone, and R. Bettinger, "Artificial debris propagation in cislunar periodic orbits," *2021 Astrodynamics Specialist Conference, AAS Paper*, 2021, pp. 21–525.

[11] K. K. Boudad, D. C. Davis, and K. C. Howell, "Disposal trajectories from near rectilinear halo orbits," *AAS/AIAA Astrodynamics Specialists Conference*, 2018, pp. 18–289.

[12] D. C. Davis, K. K. Boudad, S. M. Phillips, and K. C. Howell, "Disposal, deployment, and debris in near rectilinear halo orbits," *AAS/AIAA Space Flight Mechanics Meeting*, 2019, pp. 19–466.

[13] D. C. Davis, R. J. Power, K. C. Howell, and J. P. Gutkowski, "Lunar impact probability for spacecraft in near rectilinear halo orbits," *AAS/AIAA Space Flight Mechanics Meeting, Virtual*, 2021, pp. 21–452.

[14] C. Frueh, K. Howell, K. J. DeMars, and S. Bhadauria, "Cislunar space situational awareness," *31st AIAA/AAS Space Flight Mechanics Meeting*, 2021, pp. 21–290.

[15] T. Wolf, E. M. Zucchelli, and B. A. Jones, "Multi-fidelity uncertainty propagation for objects in cislunar space," *AIAA SCITECH 2022 Forum*, 2022, p. 1774.

[16] K. A. LeGrand, A. V. Khilnani, and J. L. Iannamorelli, "Bayesian angles-only cislunar space object tracking," *33rd AAS/AIAA Space Flight Mechanics Meeting*, 2023, pp. 23–386.

[17] T. N. Wolf and B. A. Jones, "Information-based trajectory planning for autonomous absolute tracking in cislunar space," *arXiv preprint arXiv:2408.17435*, 2024.

[18] B. A. Jones, "Comparison of sensitivity metrics for orbit uncertainty propagation in cislunar space," *2024 IEEE Aerospace Conference*, IEEE, 2024, pp. 1–13.

[19] B. L. Reifler and B. A. Jones, "Tracking spawning events in cislunar space using a label partitioned GLMB filter," *Proceedings of the 2022 AAS/AIAA Astrodynamics Specialist Conference*, 2022, pp. 22–540.

[20] D. Durant, A. A. Popov, and R. Zanetti, "Processing angles-only tracklets for cislunar multi-target tracking," *AAS/AIAA Astrodynamics Specialist Conference*, 2024, pp. 24–339.

[21] D. Durant, A. A. Popov, and R. Zanetti, "What are you weighting for? Improved weights for Gaussian mixture filtering With application to cislunar orbit determination," *arXiv preprint arXiv:2405.11081*, 2024.

[22] B.-N. Vo, M. Mallick, Y. Bar-Shalom, S. Coraluppi, R. Osborne, R. Mahler, and B.-t. Vo, "Multitarget tracking," *Wiley Encyclopedia of Electrical and Electronics Engineering*, No. 2015, 2015.

[23] W. Wu, H. Sun, M. Zheng, and W. Huang, *Target Tracking with Random Finite Sets*. Springer, 2023.

[24] B.-N. Vo and W.-K. Ma, "The Gaussian mixture probability hypothesis density filter," *IEEE Transactions on Signal Processing*, Vol. 54, No. 11, 2006, pp. 4091–4104.

[25] Y. Qin, H. Ma, J. Chen, and L. Cheng, "Gaussian mixture probability hypothesis density filter for multipath multitarget tracking in over-the-horizon radar," *EURASIP Journal on Advances in Signal Processing*, Vol. 2015, 2015, pp. 1–18.

[26] M. Tobias and A. D. Lanterman, "Probability hypothesis density-based multitarget tracking with bistatic range and Doppler observations," *IEE Proceedings-Radar, Sonar and Navigation*, Vol. 152, No. 3, 2005, pp. 195–205.

[27] V. Eiselein, D. Arp, M. Pätzold, and T. Sikora, "Real-time multi-human tracking using a probability hypothesis density filter and multiple detectors," *2012 IEEE 9th International Conference on Advanced Video and Signal-Based Surveillance*, IEEE, 2012, pp. 325–330.

[28] N. T. Pham, W. Huang, and S. Ong, "Probability hypothesis density approach for multi-camera multi-object tracking," *Computer Vision–ACCV 2007: 8th Asian Conference on Computer Vision, Tokyo, Japan, November 18-22, 2007, Proceedings, Part I 8*, Springer, 2007, pp. 875–884.

[29] C. Lundquist, L. Hammarstrand, and F. Gustafsson, "Road intensity based mapping using radar measurements with a probability hypothesis density filter," *IEEE Transactions on Signal Processing*, Vol. 59, No. 4, 2010, pp. 1397–1408.

[30] K. Y. Leung, F. Inostroza, and M. Adams, "Relating random vector and random finite set estimation in navigation, mapping, and tracking," *IEEE Transactions on Signal Processing*, Vol. 65, No. 17, 2017, pp. 4609–4623.

[31] R. P. Mahler, "Multitarget Bayes filtering via first-order multitarget moments," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 39, No. 4, 2003, pp. 1152–1178.

[32] D. Durant, A. A. Popov, and R. Zanetti, "What are you weighting for? Improved weights for Gaussian mixture filtering," *2024 27th International Conference on Information Fusion (FUSION)*, IEEE, 2024, pp. 1–8.

[33] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with applications to tracking and navigation: theory algorithms and software*. John Wiley & Sons, 2001.

[34] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering*, Vol. 82, No. 1, 1960, pp. 35–45.

[35] A. Gelb *et al.*, *Applied optimal estimation*. MIT press, 1974.

[36] S. J. Julier and J. K. Uhlmann, "New extension of the Kalman filter to nonlinear systems," *Signal processing, sensor fusion, and target recognition VI*, Vol. 3068, Spie, 1997, pp. 182–193.

[37] R. v. d. Merwe and E. Wan, *Sigma-point Kalman filters for probabilistic inference in dynamic state-space models*. PhD thesis, The faculty of the OGI School of Science & Engineering at Oregon, 2004.

[38] I. Arasaratnam and S. Haykin, "Cubature Kalman filters," *IEEE Transactions on Automatic Control*, Vol. 54, No. 6, 2009, pp. 1254–1269.

[39] H. W. Sorenson and D. L. Alspach, "Recursive Bayesian estimation using Gaussian sums," *Automatica*, Vol. 7, No. 4, 1971, pp. 465–479.

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

[41] S. Yun and R. Zanetti, "Sequential Monte Carlo filtering with Gaussian mixture sampling," *Journal of Guidance, Control, and Dynamics*, Vol. 42, No. 9, 2019, pp. 2069–2077.

[42] J. L. Anderson and S. L. Anderson, "A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts," *Monthly Weather Review*, Vol. 127, No. 12, 1999, pp. 2741–2758.

[43] B. W. Silverman, *Density estimation for statistics and data analysis*. Routledge, 2018.

[44] S. Yun, R. Zanetti, and B. A. Jones, "Kernel-based ensemble Gaussian mixture filtering for orbit determination with sparse data," *Advances in Space Research*, Vol. 69, No. 12, 2022, pp. 4179–4197.

[45] A. A. Popov and R. Zanetti, "Ensemble Gaussian mixture filtering with particle-localized covariances," *2023 26th International Conference on Information Fusion (FUSION)*, IEEE, 2023, pp. 1–7.

[46] A. A. Popov and R. Zanetti, "An adaptive covariance parameterization technique for the ensemble Gaussian mixture filter," *arXiv preprint arXiv:2212.10323*, 2022.

[47] A. A. Popov and R. Zanetti, "Ensemble-localized kernel density estimation with applications to the ensemble Gaussian mixture filter," *arXiv preprint arXiv:2308.14143*, 2023.

[48] B. L. Reifler, A. A. Popov, B. A. Jones, and R. Zanetti, "Large-Scale Space Object Tracking in a Proliferated LEO Scenario," *2023 26th International Conference on Information Fusion (FUSION)*, IEEE, 2023, pp. 1–8.

[49] D. Durant, A. A. Popov, and R. Zanetti, "MCMC EnGMF for sparse data orbit determination," *AAS/AIAA Astrodynamics Specialist Conference*, 2023, pp. 23–356.

[50] A. S. Stordal, H. A. Karlsen, G. Nævdal, H. J. Skaug, and B. Vallès, "Bridging the ensemble Kalman filter and particle filters: the adaptive Gaussian mixture filter," *Computational Geosciences*, Vol. 15, 2011, pp. 293–305.

[51] D. Raihan and S. Chakravorty, "Particle Gaussian mixture filters-I," *Automatica*, Vol. 98, 2018, pp. 331–340.

[52] D. Raihan and S. Chakravorty, "Particle Gaussian mixture filters-II," *Automatica*, Vol. 98, 2018, pp. 341–349.

[53] B. L. Reifler, S. Yun, B. A. Jones, and R. Zanetti, "Multi-target ensemble Gaussian mixture tracking with sparse observations," *AMOS Conf. Proc*, 2021.

[54] D. E. Lee, "White Paper: Gateway destination orbit model: a continuous 15 year NRHO reference trajectory," Tech. Rep. Document ID: 20190030294, NASA Johnson Space Center, Houston, TX, 2019.

[55] E. M. Z. Spreen, *Trajectory design and targeting for applications to the exploration program in cislunar space*. PhD thesis, Purdue University, 2021.

[56] J. R. Dormand and P. J. Prince, "A family of embedded Runge-Kutta formulae," *Journal of Computational and Applied Mathematics*, Vol. 6, No. 1, 1980, pp. 19–26.

[57] D. Schuhmacher, B.-T. Vo, and B.-N. Vo, "A consistent metric for performance evaluation of multi-object filters," *IEEE Transactions on Signal Processing*, Vol. 56, No. 8, 2008, pp. 3447–3457.

[58] H. W. Kuhn, "The Hungarian method for the assignment problem," *Naval Research Logistics Quarterly*, Vol. 2, No. 1-2, 1955, pp. 83–97.

[59] J. Munkres, "Algorithms for the assignment and transportation problems," *Society for Industrial and Applied Mathematics*, Vol. 15, 1962, pp. 196–210.