# Anonymous, Extent-Informed Navigation for Map-Based Localization using Random Finite Sets

James D. Brouk
Department of Aerospace Engineering
*Texas A&M University*
College Station, TX, USA
james.brouk@tamu.edu

Kyle J. DeMars
Department of Aerospace Engineering
*Texas A&M University*
College Station, TX, USA
demars@tamu.edu

***Abstract*—This paper presents the anonymous, extent-informed (AEI) update for map-based localization. The presented approach builds upon anonymous feature processing (AFP) by introducing a new prior and landmark likelihood model such that extent-dependencies in the detection process are accounted for directly in the measurement update. The AEI update is applied to a lunar descent scenario, where the simulated vehicle collects optical observations of the lunar surface and compares them to an onboard map. The AEI update is compared to both a Gaussian mixture implementation of AFP and the standard extended Kalman filter (EKF). The results show that the AEI update is able to provide more precise and consistent estimates of the vehicle's position and velocity than the other methods, while also requiring fewer components in the posterior mixture. The AEI update is also shown to be more robust to the presence of clutter and false detection processes.**

***Index Terms*—Multiple object tracking, extended object tracking, random finite sets, random matrix, Gaussian inverse Wishart, Gaussian mixture, Kalman filter**

## I. Introduction

Navigational aids that many terrestrial vehicles benefit from, such as GPS, are often unavailable for planetary descent and exploration. One potential solution is to develop an *a priori* surface model (e.g., a "feature map" that contains landmark position and spatial extent) using a precursor mission. The resulting map can then be used in conjunction with map-based, feature-relative navigation techniques for subsequent missions. For instance, modern crater-based navigation uses observations of lunar impact craters to estimate the vehicle state using an onboard catalog. Unfortunately, standard spacecraft navigation systems (typically based upon the extended Kalman filter (EKF)) rely upon heuristic approaches for outlier detection, require explicit measurement-to-feature associations, and approximate the extended object observation as a point-source measurement. This paper proposes a map-based approach to navigation that leverages a random finite set (RFS)-derived observation model and adopts the random matrix (RM)framework to introduce landmark extent-dependencies and mapping uncertainties via an assumed density approximation.

Recent advances in map-based localization [1], [2] and simultaneous localization and mapping (SLAM) [3] leverage the RFS framework to model the environment and resulting observations – under this construction, measurement-to-landmark association uncertainties and missed/false detection processes (due to clutter or untracked landmarks) can be modeled directly and tackled in a statistically consistent manner. For the anonymous feature processing (AFP) method of [1] and [2], an approximate likelihood is presented that allows these detection processes to be accounted for within a traditional navigation framework. However, by truncating the true multitarget likelihood to only consider the "one-at-a-time" approximation, the update becomes sensitive to high-clutter environments, potentially leading to poor estimation performance. In this paper, a new formulation of the methods presented by [1] and [2] is developed that incorporates extent knowledge (i.e., the shape and orientation of the mapped/observed features) into the measurement update. By this design, the likelihood-based association of the "one-at-a-time" approximation is made less sensitive to clutter and accounts for an extent-dependent error in landmark centroid detections.

Originally introduced to Bayesian estimation by [4], Gaussian mixture (GM) models have been widely adopted for the estimation of nonlinear and non-Gaussian systems. The popularized implementation of [5] uses an EKF-based implementation, where the EKF approximation is made for each mixand. For multitarget tracking, the GM-based implementation of the PHD filter, introduced in [6], is a popular choice for implementation of the PHD filter [7]. By modeling the posterior as a GM, the impact of nonlinearities in both the dynamics and measurement update can be effectively reduced. For this reason, this paper expands upon the original presentation of AFP by retaining the posterior as a GM and showing that the extent-informed update effectively reduces the number of posterior components.

Due to advancements in sensing and computational technologies, extended object tracking has been a popular field of research. While many models have been proposed, approaches based upon the RM model (originally popularized by the Bayesian solution of [8]) have remained popular, with approximate solutions [9] adopted more widely for multiple extended object tracking [10], [11], based upon the extended object likelihood function presented by [12]. Unfortunately, when dealing with optical measurements, it can be difficult

This work was supported by a NASA Space Technologies Graduate Research Opportunity.

to independently detect the extent-distributed measurements. Reference [13] notes that, in the case of images, it is often possible to reduce complexity by "applying a 'blob detector' to a camera image" and selecting "the centroids of the blobs." In this paper, the detected "blob" (i.e., the extent of the object) is used to inform the measurement update, assuming that the centroid detection is extent-dependent. Leveraging the developments in RM-based extended object tracking, this work specifies a Gaussian inverse Wishart (GIW) prior for the landmark states. Assuming the posterior for each landmark is also GIW, landmark states can be marginalized from the posterior to yield a GM in the vehicle state (akin to the Bayesian formulation of the consider Kalman filter [14], [15]).

## II. Dynamics Preliminaries

Suppose that the navigated vehicle has state

$$\boldsymbol{x}_k = \begin{bmatrix} \boldsymbol{p}_k & \boldsymbol{v}_k & \bar{\boldsymbol{q}}^b_{i,k} \end{bmatrix}^T ,$$

composed of position $\boldsymbol{p}_k \in \mathbb{R}^3$, velocity $\boldsymbol{v}_k \in \mathbb{R}^3$, and orientation $\bar{\boldsymbol{q}}^b_{i,k} \in \mathbb{S}^3$, where $\bar{\boldsymbol{q}}^b_{i,k}$ is the unit quaternion that rotates from the inertial to a body-fixed frame. The state evolves according to the discrete-time, nonlinear, stochastic process model

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

where $\boldsymbol{w}_{k-1}$ is taken to be a zero-mean, white noise sequence with covariance $\boldsymbol{P}_{ww,k-1}$. Note that many vehicles of interest to this work operate in a "model-replacement mode" whereby measurements of nongravitational acceleration and angular rotations are integrated to predict the current state of the vehicle (the reader is directed to [16] for further details on inertial navigation).

Let the prior probability density function (pdf) of the vehicle state be approximated by a finite Gaussian mixture of $n_x$ components, defined as

$$p(\boldsymbol{x}_k|\boldsymbol{\zeta}_{1:k-1}) = \sum_{j=1}^{n_x} w_k^{j-} p_g(\boldsymbol{x}_k; \boldsymbol{m}_{x,k}^{j-}, \boldsymbol{P}_{xx,k}^{j-}) ,$$

where each component, defined with weight $w_k^{j-}$, mean $\boldsymbol{m}_{x,k}^{j-}$, and covariance $\boldsymbol{P}_{xx,k}^{j-}$, is conditioned on the measurements $\boldsymbol{\zeta}_{1:k-1} = \{\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_{k-1}\}$. With this construction, the classical form of recursive Bayesian estimation uses the Chapman-Kolmogorov equation to propagate the pdf of $\boldsymbol{x}_k$ via its transition density $\boldsymbol{f}(\boldsymbol{x}_k|\boldsymbol{x}_{k-1})$, i.e.,

$$p(\boldsymbol{x}_k|\boldsymbol{\zeta}_{1:k-1}) = \int f(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}) p(\boldsymbol{x}_{k-1}|\boldsymbol{\zeta}_{1:k-1}) \mathrm{d}\boldsymbol{x}_{k-1} .$$

For the extended GM representation used in this paper, linearization about the component mean is applied to each component of the mixture such that the second moment is propagated via a linearized form of the dynamics. Thus, the propagated moments of the $j^{\text{th}}$ GM component are

$$\begin{aligned} \boldsymbol{m}_{x,k}^{j-} &= \boldsymbol{f}(\boldsymbol{m}_{x,k-1}^{j+}) , \\ \boldsymbol{P}_{xx,k}^{j-} &= \boldsymbol{F}_{x,k-1}^{j} \boldsymbol{P}_{xx,k-1}^{j+} (\boldsymbol{F}_{x,k-1}^{j})^T + \boldsymbol{P}_{ww,k-1} , \end{aligned}$$

where the Jacobian of $\boldsymbol{f}(\boldsymbol{x}_{k-1})$ w.r.t. $\boldsymbol{x}_{k-1}$ is defined to be

$$\boldsymbol{F}_{x,k-1} = \frac{\partial \boldsymbol{f}(\boldsymbol{x}_{k-1})}{\partial \boldsymbol{x}_{k-1}}\Big|_{\boldsymbol{x}_{k-1}=\boldsymbol{m}_{x,k-1}^{j+}} ,$$

evaluated at the $t_{k-1}$ posterior mean $\boldsymbol{m}_{x,k-1}^{j+}$.

## III. Environment and Observation Models

Assume that the vehicle is equipped with a map of distinct landmarks $\boldsymbol{\Lambda}_k$ while simultaneously collecting observations $\mathrm{Z}_k$ of the observed landmarks at each time $t_k$. Let the RFS of *a priori* known and detectable features in the environment of interest $\boldsymbol{E}_k$ be partitioned into two disjoint sets, $\boldsymbol{\Lambda}_k$ and $\boldsymbol{N}_k$, i.e.,

$$\boldsymbol{E}_k = \boldsymbol{\Lambda}_k \cup \boldsymbol{N}_k \quad \text{and} \quad \boldsymbol{\Lambda}_k \cap \boldsymbol{N}_k = \emptyset ,$$

where $\boldsymbol{\Lambda}_k$ corresponds to the set of "map" landmarks (i.e., those selected for localization) and $\boldsymbol{N}_k$ represents the set of remaining features in the environment, deemed the "nuisance" feature set. The set of all landmarks $\boldsymbol{E}_k$ is assumed to be finite, containing $n_E = n_M + n_N$ distinct members, where $n_M = |\boldsymbol{\Lambda}_k|$ and $n_N = |\boldsymbol{N}_k|$. Additionally, the set of all landmarks is assumed to be complete, such that $\boldsymbol{E}_k = \boldsymbol{E}_{k+1}$ for all $k$; as a result, the set of landmarks $\boldsymbol{\Lambda}_k$ and the set of nuisance landmarks $\boldsymbol{N}_k$ are assumed to be static.

Each landmark is taken to be described by a centroid position vector $\boldsymbol{\xi}_k$ and extent matrix $\boldsymbol{\Xi}_k$, such that the $\ell^{\text{th}}$ landmark state is represented by the set $\boldsymbol{\lambda}_k^\ell = \{\boldsymbol{\xi}_k^\ell,\ \boldsymbol{\Xi}_k^\ell\}$. Assuming that the map has been constructed from previous observations, the feature states are taken to be uncertain, with a GIW density model. The landmark centroid position is described by a Gaussian distribution

$$\boldsymbol{\xi}_k^\ell \sim p_g(\boldsymbol{\xi}_k^\ell;\ \boldsymbol{m}_{\xi,k}^\ell, \boldsymbol{P}_{\xi\xi,k}^\ell) ,$$

with mean $\boldsymbol{m}_{\xi,k}^\ell$ and covariance $\boldsymbol{P}_{\xi\xi,k}^\ell$. The extent of the $\ell^{\text{th}}$ landmark is assumed to be distributed according to an inverse Wishart distribution, i.e.,

$$\boldsymbol{\Xi}_k^\ell \sim p_{iw}(\boldsymbol{\Xi}_k^\ell;\ \nu_k^\ell, \boldsymbol{\Psi}_k^\ell) ,$$

where $\nu_k^\ell$ is the degrees of freedom parameter and $\boldsymbol{\Psi}_k^\ell$ is the scale matrix. The extent of the landmark is assumed to be independent of the centroid position, such that the joint distribution of the landmark state is given by

$$p(\boldsymbol{\lambda}_k^\ell) = p(\{\boldsymbol{\xi}_k^\ell, \boldsymbol{\Xi}_k^\ell\}) = p_g(\boldsymbol{\xi}_k^\ell|\boldsymbol{\Xi}_k^\ell) p_{iw}(\boldsymbol{\Xi}_k^\ell) .$$

Let the observation RFS $\boldsymbol{\Theta}_k \in \mathbb{Z}$, for a vehicle of state $\boldsymbol{x}_k$ equipped with the map $\boldsymbol{\Lambda}_k$, be defined such that

$$\boldsymbol{\Theta}_k(\boldsymbol{x}_k, \boldsymbol{\Lambda}_k) = \boldsymbol{K}_k(\boldsymbol{x}_k, \boldsymbol{N}_k) \cup \Big[ \bigcup_{\ell=1}^{n_M} \boldsymbol{\theta}_k(\boldsymbol{x}_k, \boldsymbol{\lambda}_k^\ell) \Big] , \quad (1)$$

where $\boldsymbol{\theta}_k(\boldsymbol{x}_k, \boldsymbol{\lambda}_k^\ell)$ describes the RFS for the observation of the $\ell^{\text{th}}$ feature, and $\boldsymbol{K}_k(\boldsymbol{x}_k, \boldsymbol{N}_k)$ denotes the clutter RFS that simultaneously encapsulates the set of observations derived from the nuisance feature set and false feature detections.

Let the sensor generate $m$ measurements $\boldsymbol{\zeta}_k^i$ at time $t_k$, such that the measurement set $\mathrm{Z}_k \subset \boldsymbol{\Theta}_k$ is expressed as

$$\mathrm{Z}_k = \{\boldsymbol{\zeta}_k^1,\ \boldsymbol{\zeta}_k^2,\ \ldots,\ \boldsymbol{\zeta}_k^m\}\,.$$

The RFS that describes the measurement of the $\ell^{\text{th}}$ landmark is given by

$$\boldsymbol{\theta}_k(\boldsymbol{x}_k, \boldsymbol{\lambda}_k^\ell) = \{\boldsymbol{\zeta}_k^\ell\} \cup \boldsymbol{\emptyset}(p_D(\boldsymbol{x}_k, \boldsymbol{\lambda}_k^\ell))\,,$$

where $\boldsymbol{\emptyset}(p_D(\boldsymbol{x}_k^\ell, \boldsymbol{\lambda}_k^\ell))$ denotes a null detection set, defined such that the measurement for the $\ell^{\text{th}}$ landmark is generated with probability of detection $p_D(\boldsymbol{x}_k, \boldsymbol{\lambda}_k^\ell)$, otherwise stated as

$$\boldsymbol{\theta}_k(\boldsymbol{x}_k, \boldsymbol{\lambda}_k^\ell) = \begin{cases} \boldsymbol{\zeta}_k^\ell & \text{with probability } p_D(\boldsymbol{x}_k, \boldsymbol{\lambda}_k^\ell) \\ \emptyset & \text{with probability } 1 - p_D(\boldsymbol{x}_k, \boldsymbol{\lambda}_k^\ell) \end{cases}. \quad (2)$$

For this work, the detection of the $i^{\text{th}}$ landmark in $\boldsymbol{\Lambda}_k$ produces a set that contains measurements of the centroid position $\boldsymbol{z}_k^i$ and extent $\boldsymbol{Z}_k^i$ according to the measurement likelihood function $g(\boldsymbol{\zeta}_k^i|\boldsymbol{x}_k, \boldsymbol{\lambda}_k^i)$, i.e.,

$$\boldsymbol{\zeta}_k^i = \{\boldsymbol{z}_k^i,\ \boldsymbol{Z}_k^i\}\,, \qquad \boldsymbol{\zeta}_k^i \sim g(\boldsymbol{\zeta}_k^i|\boldsymbol{x}_k, \boldsymbol{\lambda}_k^i)\,.$$

## IV. Filter Development

The posterior is obtained using via Bayes' rule, such that

$$p(\boldsymbol{x}_k|\boldsymbol{\zeta}_{1:k}) \propto g(\boldsymbol{\zeta}_k|\boldsymbol{x}_k)p(\boldsymbol{x}_k|\boldsymbol{\zeta}_{1:k-1})\,.$$

With this construction, statistical inference is performed using the prior state pdf $p(\boldsymbol{x}_k|\boldsymbol{\zeta}_{1:k-1})$ and the observation $\boldsymbol{\zeta}_k$, incorporating new measurement information into the posterior estimate. The extension of Bayesian estimation to multiple and extended objects has been a popular topic of research with many recent developments the result of the RFS architecture and multiobject calculus [13].

### A. Likelihood Derivation

The belief mass function (BMF) $\beta_X(\boldsymbol{Y})$ is used to describe the probability of a RFS $\boldsymbol{X} \in \mathbb{Z}$, much like the cumulative density function (cdf) is used to describe the probability of a random variable [13]. Here, the BMF is used to derive a likelihood function for RFS observation models presented in Section III, where the multiobject pdf describing the measurement set likelihood is obtained by taking the set derivative of the BMF and evaluating it at the null set. For the RFS $\boldsymbol{\Theta}_k \in \mathbb{Z}$, its BMF is defined such that

$$\beta_\Theta(\boldsymbol{Y}|\boldsymbol{x}_k, \boldsymbol{\Lambda}_k) = \Pr(\boldsymbol{\Theta}_k \subseteq \boldsymbol{Y}|\boldsymbol{x}_k,\ \boldsymbol{\Lambda}_k)\,. \quad (3)$$

Assuming that each observation in the observation set composed according to Eq. (1) is generated independently, Eq. (3) can be expressed as

$$\beta_\Theta(\boldsymbol{Y}|\boldsymbol{x}_k, \boldsymbol{\Lambda}_k) = \beta_K(\boldsymbol{Y}|\boldsymbol{x}_k, \boldsymbol{N}_k) \prod_{i=1}^{n_M} \beta_{\theta,i}(\boldsymbol{Y}|\boldsymbol{x}_k, \boldsymbol{\Lambda}_k)\,.$$

From the definition of the detection set in Eq. (2), the BMF describing the probability of a measurement being generated by the $i^{\text{th}}$ landmark is expressed as

$$\beta_{\theta,i}(\boldsymbol{Y}|\boldsymbol{x}_k, \boldsymbol{\Lambda}_k) = (1 - p_D(\boldsymbol{x}_k, \boldsymbol{\lambda}_k^i)) + p_D(\boldsymbol{x}_k, \boldsymbol{\lambda}_k^i) \int_Y g(\boldsymbol{\zeta}|\boldsymbol{x}_k, \boldsymbol{\lambda}_k^i)\mathrm{d}\boldsymbol{\zeta}\,,$$

Additionally, assuming that clutter detections are temporally Poisson distributed, the BMF describing the probability of a clutter measurement being generated is

$$\beta_K(\boldsymbol{Y}|\boldsymbol{x}_k, \boldsymbol{N}_k) = \exp\left(\lambda_c \int_Y c(\boldsymbol{\zeta})\mathrm{d}\boldsymbol{\zeta} - \lambda_c\right),$$

where $\lambda_c$ and $c(\boldsymbol{\zeta})$ denote the clutter rate and spatial density, respectively. Because this clutter model incorporates both false alarms and nuisance features simultaneously, the clutter rate $\lambda_c$ is assumed to be the sum of the false alarm rate $\lambda_f$ and the nuisance feature rate $\lambda_n$, i.e., $\lambda_c = \lambda_f + \lambda_n$. The spatial density $c(\boldsymbol{\zeta})$ also accounts for the presence of nuisance landmarks in a given measurement scan.

As shown by [13], the set derivative of the BMF becomes, by application of the general product rule, a sum over all partitions of $\mathrm{Z}_k$ into $n_M+1$ disjoint subsets (including empty sets). Because this result is combinatoric and difficult to apply for large-scale problems, this work adopts several approximations of this set derivative to formulate a multiobject likelihood. In [1], the authors approximate the set derivative using a "one-at-a-time" approximation that considers the assignment of a single measurement to each landmark while relegating the remaining measurements to the clutter set. In [2], the author expanded this approximation to also include a term for the "null scan" case, accounting for the case where no measurements are assigned to any landmark. Applying these approximations, the set derivative can be as expressed as

$$\frac{\delta\beta_\Theta(\boldsymbol{Y}|\boldsymbol{x}_k, \boldsymbol{\Lambda}_k)}{\delta \mathrm{Z}_k} = \frac{\delta\beta_K(\boldsymbol{Y}|\boldsymbol{x}_k)}{\delta \mathrm{Z}_k} \prod_{i=1}^{n_M} \frac{\delta\beta_{\theta,i}(\boldsymbol{Y}|\boldsymbol{x}_k, \boldsymbol{\Lambda}_k)}{\delta\boldsymbol{\emptyset}} + \sum_{\substack{\boldsymbol{\zeta}\in\mathrm{Z}_k \\ \boldsymbol{W}=\mathrm{Z}_k\setminus\boldsymbol{\zeta}}} \frac{\delta\beta_K(\boldsymbol{Y}|\boldsymbol{x}_k)}{\delta\boldsymbol{W}} \sum_{i=1}^{n_M} \frac{\delta\beta_{\theta,i}(\boldsymbol{Y}|\boldsymbol{x}_k, \boldsymbol{\Lambda}_k)}{\delta\boldsymbol{\zeta}}\,,$$

where the likelihood function is then obtained by evaluating the set derivative at the null set.

To separate terms, the likelihood function can be written as

$$g(\mathrm{Z}_k|\boldsymbol{x}_k, \boldsymbol{\Lambda}_k) = g_{\text{null}}(\mathrm{Z}_k|\boldsymbol{x}_k, \boldsymbol{\Lambda}_k) + g_{\text{one}}(\mathrm{Z}_k|\boldsymbol{x}_k, \boldsymbol{\Lambda}_k)\,, \quad (4)$$

where the first term is the null scan case, and the second term is the one-at-a-time approximation. By applying the definition of the set derivative and the fundamental theorem of multiobject calculus (see p. 68 and 70, [13]), the derivative of the clutter term, evaluated at the null set, can be reduced to

$$\frac{\delta\beta_K(\boldsymbol{\emptyset}|\boldsymbol{x}_k)}{\delta \mathrm{Z}_k} = \exp(-\lambda_c) \prod_{\boldsymbol{y}\in\mathrm{Z}_k} \lambda_c c(\boldsymbol{y})\,. \quad (5)$$

By the linear rule for set derivatives (see p. 73, [13]),

$$\frac{\delta\beta_{\theta,i}(\boldsymbol{Y}|\boldsymbol{x}_k,\boldsymbol{\Lambda}_k)}{\delta\emptyset} = \beta_{\theta,i}(\emptyset|\boldsymbol{x}_k,\boldsymbol{\Lambda}_k) = \Big(1 - p_D(\boldsymbol{x}_k,\boldsymbol{\lambda}_k^i)\Big)\,.$$

Thus, the likelihood for the null scan is given by

$$g_{\text{null}}(\mathrm{Z}_k|\boldsymbol{x}_k,\boldsymbol{\Lambda}_k) = \phi_k\Big(1 - p_D(\boldsymbol{x}_k,\cdot)\Big)^{\boldsymbol{\Lambda}_k}, \tag{6}$$

where, for convenience, $\phi_k \triangleq \exp(-\lambda_c)\left[\lambda_c c(\cdot)\right]^{\mathrm{Z}_k}$ is defined and the power functional notation of [13] is adopted, i.e., $h^{\boldsymbol{Y}} = \prod_{\boldsymbol{y}\in\boldsymbol{Y}} h(\boldsymbol{y})$. Using Eq. (5), and noting that the set derivative of the landmark term is

$$\frac{\delta\beta_{\theta,i}(\boldsymbol{Y}|\boldsymbol{x}_k,\boldsymbol{\Lambda}_k)}{\delta\boldsymbol{\zeta}} = p_D(\boldsymbol{x}_k,\boldsymbol{\lambda}_k^i)g(\boldsymbol{\zeta}|\boldsymbol{x}_k,\boldsymbol{\lambda}_k^i)\,,$$

the one-at-a-time likelihood is given by

$$g_{\text{one}}(\mathrm{Z}_k|\boldsymbol{x}_k,\boldsymbol{\Lambda}_k) = \phi_k \sum_{\boldsymbol{\zeta}\in\mathrm{Z}_k}\sum_{i=1}^{n_M} \frac{p_D(\boldsymbol{x}_k,\boldsymbol{\lambda}_k^i)g(\boldsymbol{\zeta}|\boldsymbol{x}_k,\boldsymbol{\lambda}_k^i)}{\lambda_c c(\boldsymbol{\zeta})}\,, \tag{7}$$

after evaluating at the empty set. Thus, the likelihood of Eq. (4) is the sum of Eqs. (6) and (7).

### *B. Anonymous, Extent-Informed Update*

Following the likelihood derivation, all that remains is the update formulation. Let the prior be defined such that the map is independent of both the vehicle state and the measurements collected, such that the Bayesian posterior is

$$p(\boldsymbol{x}_k,\boldsymbol{\Lambda}_k|\mathrm{Z}_k) = \eta_k^{-1} g(\mathrm{Z}_k|\boldsymbol{x}_k,\boldsymbol{\Lambda}_k)p(\boldsymbol{x}_k|\mathrm{Z}_{1:k-1})p(\boldsymbol{\Lambda}_k)\,,$$

where $\eta_k$ denotes the Bayesian normalization constant.

Adopting the vehicle prior as a GM and assuming that each landmark is independent with prior represented by a GIW pdf, the map pdf is expressed as

$$p(\boldsymbol{\Lambda}_k) = \prod_{i=1}^{n_M} p_g(\boldsymbol{\xi}_k^i;\boldsymbol{m}_{\xi,k}^i,\boldsymbol{P}_{\xi\xi,k}^i)p_{iw}(\boldsymbol{\Xi}_k^i;\nu_k^i,\boldsymbol{\Psi}_k^i)\,. \tag{8}$$

Additionally, let the likelihood for an individual measurement $\boldsymbol{\zeta}_k = \{\boldsymbol{z}_k,\boldsymbol{Z}_k\}$ be Gaussian Wishart, such that

$$\begin{aligned} g(\boldsymbol{\zeta}_k|\boldsymbol{x}_k,\boldsymbol{\lambda}_k^i) = {} & p_g(\boldsymbol{z}_k;\boldsymbol{h}(\boldsymbol{x}_k,\boldsymbol{\xi}_k^i),\boldsymbol{\Sigma}_k^i/n_k) \\ & \times p_w(\boldsymbol{Z}_k;n_k-1,\boldsymbol{\Sigma}_k^i/(n_k-1))\,, \end{aligned} \tag{9}$$

where $\boldsymbol{h}(\boldsymbol{x}_k,\boldsymbol{\xi}_k^i)$ is the nonlinear measurement function for landmark $\boldsymbol{\xi}_k^i$, $\boldsymbol{\Sigma}_k^i = \boldsymbol{P}_{vv,k} + s\boldsymbol{\Xi}_k^i$ is the effective measurement noise covariance for the centroid, with the covariance of a single point detection described by $\boldsymbol{P}_{vv,k}$, and $s$ defined as a scaling factor that matches the covariance of an assumed uniform distribution of point detections. Furthermore, $\boldsymbol{Z}_k$ denotes an 'extent-measurement', and describes the distribution of extent-distributed points in the detection. Notice that $n_k$ is a scalar degree of freedom parameter that describes the precision of the extent measurements – as $n_k$ increases, the extent measurement converges to the true extent of the object. While loosely defined here, $n_k$ could be an additional parameter delivered by the sensing system describing how many points are used in generating $\boldsymbol{Z}_k$. For this work, $n_k$ is assumed to be object-dependent and proportional to the observable extent.

Combining the likelihood function of Eq. (4) with the single-object likelihood of Eq. (9), and adopting the map prior of Eq. (8), the posterior pdf for the vehicle state is

$$\begin{aligned} p(\boldsymbol{x}_k|\mathrm{Z}_k) = \eta_k^{-1}\phi_k \Bigg[ & \int \Big(1 - p_D(\boldsymbol{x}_k,\cdot)\Big)^{\boldsymbol{\Lambda}_k} p(\boldsymbol{\lambda}_k^i)\mathrm{d}\boldsymbol{\lambda}_k^i \\ & + \sum_{\boldsymbol{\zeta}\in\mathrm{Z}_k}\sum_{i=1}^{n_M} \frac{\int p_D(\boldsymbol{x}_k,\boldsymbol{\lambda}_k^i)g(\boldsymbol{\zeta}|\boldsymbol{x}_k,\boldsymbol{\lambda}_k^i)p(\boldsymbol{\lambda}_k^i)\mathrm{d}\boldsymbol{\lambda}_k^i}{\lambda_c c(\boldsymbol{\zeta})} \Bigg] \\ & \times p(\boldsymbol{x}_k|\mathrm{Z}_{1:k-1})\,, \end{aligned} \tag{10}$$

where the map states are marginalized out of the posterior to produce the marginal posterior for the vehicle state. By substituting the vehicle prior and bringing it into the integral, the integral in Eq. (10) that corresponds to the one-at-a-time approximation can also be expressed as

$$\begin{aligned} & p(\boldsymbol{x}_k|\mathrm{Z}_{1:k-1}) \int p_D(\boldsymbol{x}_k,\boldsymbol{\lambda}_k^i)g(\boldsymbol{\zeta}|\boldsymbol{x}_k,\boldsymbol{\lambda}_k^i)p(\boldsymbol{\lambda}_k^i)\delta\boldsymbol{\lambda}_k^i \\ & = \sum_{j=1}^{n_x} w_k^j \int p_D(\boldsymbol{x}_k,\boldsymbol{\lambda}_k^i)p_g(\boldsymbol{z}_k|\boldsymbol{x}_k,\boldsymbol{\lambda}_k^i)p_g^j(\boldsymbol{x}_k|\mathrm{Z}_{1:k-1}) \\ & \qquad \times p_g(\boldsymbol{\xi}_k^i)p_w(\boldsymbol{Z}_k|\boldsymbol{\Xi}_k^i)p_{iw}(\boldsymbol{\Xi}_k^i)\delta\boldsymbol{\xi}_k^i\delta\boldsymbol{\Xi}_k^i\,, \end{aligned} \tag{11}$$

where the shorthand $p_g^j(\cdot)$ denotes the $j^{\text{th}}$ Gaussian component of the vehicle prior. To simplify, the probability of detection model is approximated to zeroth order, i.e.,

$$\begin{aligned} p_D(\boldsymbol{x}_k,\boldsymbol{\lambda}_k^i) &\approx p_D(\boldsymbol{m}_{x,k}^{j-},\boldsymbol{m}_{\lambda,k}^i) \\ \Big(1 - p_D(\boldsymbol{x}_k,\cdot)\Big)^{\boldsymbol{\Lambda}_k} &\approx \Big(1 - p_D(\boldsymbol{m}_{x,k}^{j-},\cdot)\Big)^{\hat{\boldsymbol{\Lambda}}_k}\,, \end{aligned}$$

where $\hat{\boldsymbol{\Lambda}}_k = \{\boldsymbol{m}_{\lambda,k}^1,\ \boldsymbol{m}_{\lambda,k}^2,\ \ldots,\ \boldsymbol{m}_{\lambda,k}^{n_M}\}$ is the set containing the estimate for each landmark state and $\boldsymbol{m}_{x,k}^{j-}$ is the prior estimate of the vehicle state. Additionally, by applying the properties of Gaussians [17] and carrying out the landmark centroid state marginalization, the GM of the vehicle state updated with the centroid measurement is obtained as

$$\begin{aligned} & p_g(\boldsymbol{x}_k;\boldsymbol{m}_{x,k}^{i,j+},\boldsymbol{P}_{xx,k}^{i,j+})p_g(\boldsymbol{z}_k;\boldsymbol{m}_{z,k}^{i,j},\boldsymbol{P}_{zz,k}^{i,j}) \\ & \qquad = p_g(\boldsymbol{z}_k;\boldsymbol{h}(\boldsymbol{m}_{x,k}^{j-},\boldsymbol{m}_{\xi,k}^i),n_k^{-1}\boldsymbol{\Sigma}_k^i) \\ & \qquad \times p_g(\boldsymbol{x}_k;\boldsymbol{m}_{x,k}^{j-};\boldsymbol{P}_{xx,k}^{j-})p_g(\boldsymbol{\xi}_k^i;\boldsymbol{m}_{\xi,k}^i,\boldsymbol{P}_{\xi\xi,k}^i) \end{aligned} \tag{12}$$

where the posterior mean $\boldsymbol{m}_{x,k}^{i,j+}$ and covariance $\boldsymbol{P}_{xx,k}^{i,j+}$ for the $j^{\text{th}}$ component updated against the $i^{\text{th}}$ landmark are

$$\boldsymbol{m}_{x,k}^{i,j+} = \boldsymbol{m}_{x,k}^{j-} + \boldsymbol{K}_k^{i,j}(\boldsymbol{z}_k - \boldsymbol{m}_{z,k}^{i,j}) \tag{13a}$$

$$\boldsymbol{P}_{xx,k}^{i,j+} = (\boldsymbol{I}_{\dim(x)} - \boldsymbol{K}_k^{i,j}\boldsymbol{H}_{x,k})\boldsymbol{P}_{xx,k}^{j-} \tag{13b}$$

with the expected measurement, linear gain, and innovation covariance defined to be

$$\boldsymbol{m}_{z,k}^{i,j} = \boldsymbol{h}(\boldsymbol{m}_{x,k}^{j-},\boldsymbol{m}_{\xi,k}^i) \tag{14a}$$

$$\boldsymbol{K}_k^{i,j} = \boldsymbol{P}_{xx,k}^{j-}\boldsymbol{H}_{x,k}^T\big(\boldsymbol{P}_{zz,k}^{i,j}\big)^{-1} \tag{14b}$$

$$\begin{aligned} \boldsymbol{P}_{zz,k}^{i,j} = {} & \boldsymbol{H}_{x,k}\boldsymbol{P}_{xx,k}^{j-}\boldsymbol{H}_{x,k}^T \\ & + \boldsymbol{H}_{\xi,k}\boldsymbol{P}_{\xi\xi,k}^i\boldsymbol{H}_{\xi,k}^T + n_k^{-1}\boldsymbol{M}_{\Sigma,k}^i\,. \end{aligned} \tag{14c}$$

In Eq. (14c), the innovation covariance $\boldsymbol{P}^{i,j}_{zz,k}$ is now dependent upon the *estimate* of the effective noise covariance, i.e.,

$$\boldsymbol{M}^{i}_{\Sigma,k} = \boldsymbol{P}_{vv,k} + s\boldsymbol{M}^{i}_{\Xi,k}\,,$$

where the landmark extent is approximated by the *a priori* estimate, i.e., $\boldsymbol{\Xi}^{i}_{k} \approx \boldsymbol{M}^{i}_{\Xi,k}$. Ultimately, this neglects uncertainty in the extent estimate for the Gaussian state update.

To facilitate a tractable solution, it is assumed that the vehicle-map state density is GIW in prior and posterior. Thus, the remaining components in Eq. (11) can be reorganized to

$$\begin{aligned} L^{i,j}_{\mathrm{Z}_k}(\boldsymbol{\zeta}_k) p_{iw}(\boldsymbol{\Xi}^{i}_{k}; \nu^{i,j+}_{k}, \boldsymbol{\Psi}^{i,j+}_{k}) &= p_g(\boldsymbol{z}_k; \boldsymbol{m}^{i,j}_{z,k}, \boldsymbol{P}^{i,j}_{zz,k}) \\ &\times p_w(\boldsymbol{Z}_k; n_k - 1, \boldsymbol{\Sigma}^{i}_{k}/(n_k-1)) p_{iw}(\boldsymbol{\Xi}^{i}_{k}; \nu^{i}_{k}, \boldsymbol{\Psi}^{i}_{k})\,, \end{aligned} \quad (15)$$

where the pseudolikelihood $L^{i,j}_{\mathrm{Z}_k}(\boldsymbol{\zeta}_k)$ is defined by

$$\begin{aligned} L^{i,j}_{\mathrm{Z}_k}(\boldsymbol{\zeta}_k) = \frac{(n_k-1)^{n_k d/2}}{\pi^{d/2}} &\frac{\Gamma_d\left[(\nu^{i}_{k}+n_k)/2\right]}{\Gamma_d\left[(n_k-1)/2\right]\Gamma_d\left[\nu^{i}_{k}/2\right]} \\ &\times \frac{|\boldsymbol{\Psi}^{i}_{k}|^{\nu^{i}_{k}/2}}{|\boldsymbol{\Psi}^{i}_{k} + \tilde{\boldsymbol{N}}^{i,j}_{k} + (n_k-1)\tilde{\boldsymbol{Z}}^{i}_{k}|^{(\nu^{i}_{k}+n_k)/2}} \\ &\times \frac{|\boldsymbol{Z}_k|^{(n_k-d-2)/2}|\boldsymbol{M}_{\Xi,k}|^{n_k/2}}{|\boldsymbol{P}^{i,j}_{zz,k}|^{1/2}|\boldsymbol{M}^{i}_{\Sigma,k}|^{(n_k-1)/2}}\,, \end{aligned} \quad (16)$$

and selected to preserve the form of the GIW prior, where the scaled innovation matrix and measurement are (using approximations akin to [9], [18])

$$\begin{aligned} \tilde{\boldsymbol{N}}^{i,j}_{k} &= (\boldsymbol{S}^{i,j}_{N,k})^{-T}(\boldsymbol{z}_k - \boldsymbol{m}^{i,j}_{z,k})(\boldsymbol{z}_k - \boldsymbol{m}^{i,j}_{z,k})^{T}(\boldsymbol{S}^{i,j}_{N,k})^{-1} \\ \tilde{\boldsymbol{Z}}^{i}_{k} &= (\boldsymbol{S}^{i}_{Z,k})^{-T}\boldsymbol{Z}_k(\boldsymbol{S}^{i}_{Z,k})^{-1}\,, \end{aligned}$$

respectively, with the scaling matrices

$$\begin{aligned} \boldsymbol{S}^{i,j}_{N,k} &= (\boldsymbol{P}^{i,j}_{zz,k})^{1/2}(\boldsymbol{\Xi}^{i}_{k})^{-1/2} \approx (\boldsymbol{P}^{i,j}_{zz,k})^{1/2}(\boldsymbol{M}^{i}_{\Xi,k})^{-1/2} \\ \boldsymbol{S}^{i}_{Z,k} &= (\boldsymbol{\Sigma}^{i}_{k})^{1/2}(\boldsymbol{\Xi}^{i}_{k})^{-1/2} \approx (\boldsymbol{M}^{i}_{\Sigma,k})^{1/2}(\boldsymbol{M}^{i}_{\Xi,k})^{-1/2} \end{aligned}$$

defined such that

$$\begin{aligned} \mathbb{E}\left\{\boldsymbol{S}^{i,j}_{zz,k}\boldsymbol{\Xi}^{i}_{k}(\boldsymbol{S}^{i,j}_{N,k})^{T}|\boldsymbol{\Xi}^{i}_{k} = \boldsymbol{M}^{i}_{\Xi,k}\right\} &= \boldsymbol{P}^{i,j}_{zz,k} \\ \mathbb{E}\left\{\boldsymbol{S}^{i}_{Z,k}\boldsymbol{\Xi}^{i}_{k}(\boldsymbol{S}^{i}_{Z,k})^{T}|\boldsymbol{\Xi}^{i}_{k} = \boldsymbol{M}^{i}_{\Xi,k}\right\} &= \boldsymbol{M}^{i}_{\Sigma,k}\,. \end{aligned}$$

Substituting Eqs. (12) and (15) and applying the probability of detection approximation, Eq. (11) becomes

$$\begin{aligned} &\int p_D(\boldsymbol{x}_k, \boldsymbol{\lambda}^{i}_{k}) g(\boldsymbol{\zeta}|\boldsymbol{x}_k, \boldsymbol{\lambda}^{i}_{k}) p(\boldsymbol{x}_k, \boldsymbol{\lambda}^{i}_{k}|\mathrm{Z}_{1:k-1}) \delta\boldsymbol{\lambda}^{i}_{k} \\ &= \sum_{j=1}^{n_x} w^{j}_{k} L^{i,j}_{\mathrm{Z}_k}(\boldsymbol{\zeta}) p_D(\boldsymbol{m}^{j-}_{x,k}, \boldsymbol{m}^{i}_{\lambda,k}) p_g(\boldsymbol{x}_k; \boldsymbol{m}^{i,j+}_{x,k}, \boldsymbol{P}^{i,j+}_{xx,k})\,. \end{aligned}$$

Thus, the posterior for the vehicle state can be expressed as

$$\begin{aligned} p(\boldsymbol{x}_k|\mathrm{Z}_k) = &\sum_{j=1}^{n_x} w^{j}_{\mathrm{null},k} p_g(\boldsymbol{x}_k; \boldsymbol{m}^{j-}_{x,k}, \boldsymbol{P}^{j-}_{xx,k}) \\ &+ \sum_{\boldsymbol{\zeta}\in\mathrm{Z}_k}\sum_{i=1}^{n_M}\sum_{j=1}^{n_x} w^{i,j}_{\mathrm{one},k}(\boldsymbol{\zeta}) p_g(\boldsymbol{x}_k; \boldsymbol{m}^{i,j+}_{x,k}, \boldsymbol{P}^{i,j+}_{xx,k})\,, \end{aligned} \quad (17)$$

where the weights are

$$\begin{aligned} w^{j}_{\mathrm{null},k} &= \frac{1}{\tilde{\eta}_k}\left(1 - p_D(\boldsymbol{m}^{j-}_{x,k}, \cdot)\right)^{\hat{\boldsymbol{\Lambda}}_k} w^{j}_{k} \\ w^{i,j}_{\mathrm{one},k}(\boldsymbol{\zeta}) &= \frac{1}{\tilde{\eta}_k}\frac{p_D(\boldsymbol{m}^{j-}_{x,k}, \boldsymbol{m}^{i}_{\xi,k}) L^{i,j}_{\mathrm{Z}_k}(\boldsymbol{\zeta})}{\lambda_c c(\boldsymbol{\zeta})} w^{j}_{k}\,, \end{aligned}$$

the normalization by $\tilde{\eta}_k$ forces the sum of weights to unity.

### C. Implementation Details and GM Reduction

If calculated according to the update of Eq. (17), the posterior GM contains $n_x(n_M n_{\mathrm{Z}}+1)$ components. Left unchecked, this component growth can make further calculations infeasible. Fortunately, the posterior can be accurately represented using only a few of the components. By applying GM reduction techniques, growth can be mitigated such that the solution tractability is preserved. Techniques for component merging, pruning, and capping can be used to reduce the posterior complexity. Component merging can be used to combine sufficiently similar elements (such as by the method of moments). Pruning and capping can be used to eliminate components that contribute minimally to the mixture and to set an artificial cap on the number of components, respectively.

While mixture reduction techniques are important to maintain the tractability of the solution, the update complexity can also be reduced to minimize the number of posterior components. Notably, if the probability of detection model is assumed to trend toward zero for landmarks that are not expected to be measured, then map components that are unlikely to contribute to the posterior can be preemptively eliminated from consideration. This is particularly useful in the context of the one-at-a-time update, where the sum over *all* landmarks is described – eliminating objects not contained in the expected field of view significantly reduces the number of components. Applying these approximations, the posterior GM can be sufficiently reduced.

## V. Results

The developed AEI update is applied to crater-based navigation in a lunar descent scenario, where precise and robust localization techniques are vital to ensure safe touchdown. For brevity, a brief description of the simulation environment is given, though a more detailed discussion of the trajectory, vehicle, and specific sensor models can be found in [19]. The modeled vehicle is equipped with a terrain camera (TC), star camera (SC), and inertial measurement unit (IMU) – specifications for each sensor are described in Table I. The forces acting on the vehicle are assumed to be thrust and two-body lunar gravitational forces, where the IMU characterizes angular rotations and nongravitational accelerations. The vehicle begins at an altitude of 50 km, descending to the lunar surface at a targeted landing site near Beer crater, following the altitude profile shown in Fig. (1). For this study, the first 10 minutes following TC activation are of interest – this period begins approximately 9.8 minutes after descent is initialized, when the vehicle is at an altitude of approximately 30 km. The groundtrack for the descent profile is shown in Fig. (2), with

the inset indicating the observed landmarks (based upon the crater catalog of [20]). The SC provides consistent attitude information, producing quaternion measurements describing the current vehicle orientation. For the first 10 minutes of descent, the vehicle integrates IMU measurements to estimate the vehicle's position and velocity, while the SC provides attitude measurements. At the beginning of the simulation, the filter covariance is uncorrelated with position, velocity, and attitude standard deviations of 1 km, 0.1 m/s, and 100 arcsec, respectively. At TC activation, position standard deviations have grown nearly 20%, while velocity uncertainties have grown by a factor of 10. Since the SC is active throughout this period, attitude uncertainty remains low.

TABLE I
SIMULATED SENSOR SPECIFICATIONS

| Sensor | Rate | Noise ($1\sigma$) | Active Time |
|---|---|---|---|
| Star Camera | 40 Hz | 10 arcsec | 0 to 23.8 min |
| Terrain Camera | 1 Hz | 1/3 pix | 9.8 to 30.6 min |
| Accelerometer | 40 Hz | 35 $\mu g/\sqrt{\text{Hz}}$ | Always |
| Gyroscope | 40 Hz | 0.07 deg$/\sqrt{\text{hr}}$ | Always |

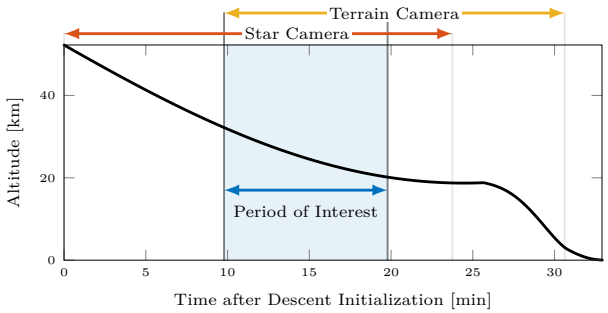


Fig. 1. Altitude profile for lunar descent-to-landing profile with sensor activation windows indicated. The period of interest is highlighted in blue.

The primary sensor of interest within this study is the nadir-pointing TC that produces optical observations of the lunar surface. Images captured by camera are then fed into an image-processing algorithm that extracts the centroid and extent measurement for each of the detected landmarks in terms of pixel coordinates/distribution. For this work, the result of image processing is modeled by sampling from a Wishart density with the landmark-specific scale matrix defined such that the expected value is the extent projected into the camera frame, while the degree of freedom is assumed proportional to the perimeter of the landmark's extent projection. Centroid measurement errors are modeled such they contain both a white-noise and extent-dependent error source (as outlined in Section IV). To model limitations in detection and classification, the TC is assumed to detect landmarks in the field of view with a probability ($p_D$) of 0.9 – however, any objects in the field of view that are too large (beyond 20% of the frame) or small (less than 5 pixels in the camera projection) are excluded from detection. Figure (2) displays the lunar groundtrack of the vehicle and indicates the landmarks observable by the camera.

The lunar surface feature map used for this study is based upon the crater catalog of [20], though it is assumed that each feature has a centroid mapping error covariance of 10 m ($1\sigma$) in each direction with a degree of freedom equal to the number of points used in the ellipse fit. Uniform clutter measurements are generated over the sensing area with a Poisson rate ($\lambda_c$) of 20, with Wishart distributed extent measurements specified by a scale $\boldsymbol{V}_{\text{c}} = \text{diag}(250, 250)$ and degree of freedom $\nu_{\text{c}} = 6$, generating mostly small to medium-sized clutter detections having large variability in orientation and eccentricity.

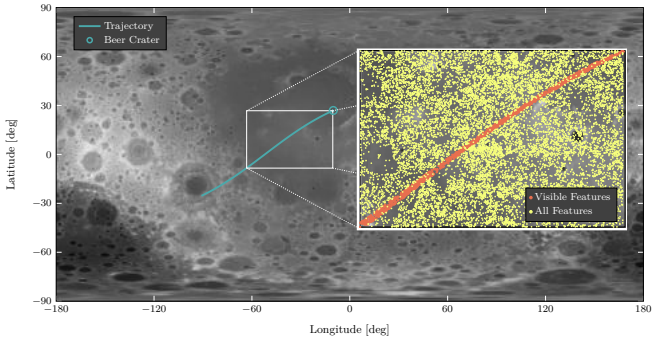


Fig. 2. Groundtrack for descent profile. The inset details a subset of landmarks contained within the lunar crater catalog of [20] with all observed features indicated.

To examine the effectiveness of the AEI update, a comparison is made with both a GM implementation of AFP (GM-AFP) and a standard EKF implementation. The AFP implementation includes the null detection case introduced by [2], and the EKF implementation assumes that the image processing can flawlessly identify the corresponding map feature for each of the observed features and eliminate clutter. If the identification fails, the measurement is omitted from the update. The EKF implementation also uses sequential updates for the set of landmark observations, relinearizing after each measurement is processed. As is common practice with EKF-based navigation, measurement underweighting and residual editing are used to prevent over-convergence – for implementation, an underweighting factor of 5/7 [19] and a 99.5% residual editing gate (based upon a $\chi^2$-test) are used. Both the EKF and GM-AFP implementations lack any knowledge of extent-dependent errors in the feature centroid. For both mixture-based approaches, GM reduction techniques are used: (1) merging is performed for any components that fall within a 99.95% probability test (based upon the Mahalanobis distance), (2) any components with a weight below $1 \times 10^{-5}$ are eliminated from the mixture, and (3) the posterior mixture is capped to 20 components (following merging and pruning).

To compare estimation performance for each of the aforementioned, a 1000 trial Monte Carlo simulation is performed for each the EKF, AFP, and the AEI update. Because the simulation is defined based upon a nominal trajectory, each trial begins with a different set of initial conditions but follows the same true vehicle trajectory in position, velocity, and attitude. With a fixed underlying trajectory, the field of view

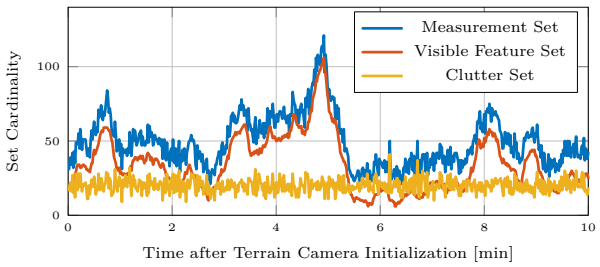


Fig. 3. Cardinality of the measurement set, the set of visible features (without missed detections), and the clutter set for a single trial.

for each terrain camera scan is the same for each optical scan, where it should be noted that the number of detected landmarks and clutter detections can vary from trial to trial. An example of the number of visible landmarks and clutter detections for a single trial is shown in Fig. (3). Notice that the number of clutter returns is, at times, larger than the number of visible features. Initially, many of the observed objects are clutter though there are nearly the same number of visible landmarks. This is a potential stress case for many methods because, without considering multiple objects at once, many observations could be equally likely such that it is harder to eliminate clutter measurements from consideration.

The results of the Monte Carlo simulations are summarized in Fig. (4), where the position and velocity errors are shown for each method – final estimation results for the time interval are provided in Table II. It is worth noting, before delving into the results, that many EKF and GM-AFP trials experience convergence issues due to the presence of unaccounted-for extent-dependent errors in the landmark centroid measurements. As such, any trial that exceeded the $9\sigma$ interval in a single axis is omitted from further calculations. Examining Fig. (4), the AEI update is shown to provide more consistent estimates of the vehicle's position and velocity than both GM-AFP and the EKF. While the AEI update tends to have slower convergence rates, AFP continues to converge, despite the large discrepancy between the predicted and true position statistics, causing poor statistical consistency as the vehicle continues. Because of this over-convergence during the first minute, approximately 75% of GM-AFP trials experience errors $\geq 6\sigma$. The EKF implementation is shown to have the fastest convergence rate, though it is also the most prone to divergence – as reported in Table II, more than 98% of trials experienced divergence. Notice also that, if the EKF's error statistics were shown on many of the figures, it would be difficult to resolve behavior in the other configurations. While each configuration appears to approach the same level of convergence in velocity, it's important to note that the AEI update maintains a conservative estimate while the GM-AFP continues to converge. Additionally, the conservatism of AEI reduces slightly around 5.5 minutes, which aligns with the clutter volume overtaking the number of visible landmarks.

To examine the impact of the cap placed on the number of posterior GM components, a second set of shorter Monte Carlo simulations is briefly examined here, where the cap on the number of posterior GM components is raised to 500 for both the AEI and GM-AFP implementations. The mean number of posterior GM components for each method is presented in Fig. (5), where the shaded region indicates interval containing 99.7% of all trials. The AEI update is shown to maintain fewer components in the posterior mixture than GM-AFP, with the AEI update requiring approximately 30% of the components of GM-AFP. This is likely due to the AEI update's ability to account for extent-dependent errors in the landmark centroid measurements, which allows for a more consistent estimate of the vehicle's position and velocity, while also being better suited to clutter rejection. Note that, during this shortened simulation, which only examines the first minute of the descent profile, the GM-AFP implementation experiences 35 trials that exceeded the $6\sigma$ interval, while the AEI update only has 1 trial that exceeds the $6\sigma$ interval.

## VI. Conclusions

This paper presents a new formulation of anonymous feature processing that incorporates landmark extent to simultaneously improve identification/clutter rejection and introduce previously unmodeled extent-dependencies in the landmark's centroid measurement. The AEI update is shown to provide more consistent estimates of the vehicle's position and velocity than other methods in an extremely cluttered environment while also requiring fewer components in the posterior mixture. The new update is also shown to be more robust to the presence of clutter and false detection processes, reducing the number of failure cases by (at least) a factor of 10. The AEI update is a promising development for space vehicles performing map-based localization, where the environment is often poorly understood.

## References

[1] J. S. McCabe and K. J. DeMars, "Anonymous feature-based terrain relative navigation," *Journal of Guidance, Control, and Dynamics*, vol. 43, no. 3, pp. 410–421, 2020.

[2] J. S. McCabe, "On a higher order method for anonymous feature processing," in *2022 AIAA SCITECH Forum*, 2022.

[3] J. Mullane, B.-N. Vo, M. D. Adams, and B.-T. Vo, "A random-finite-set approach to Bayesian SLAM," *IEEE Transactions on Robotics*, vol. 27, no. 2, pp. 268–282, 2011.

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

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

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

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

[8] J. W. Koch, "Bayesian approach to extended object and cluster tracking using random matrices," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 44, no. 3, pp. 1042–1059, 2008.

[9] M. Feldmann, D. Fränken, and W. Koch, "Tracking of extended objects and group targets using random matrices," *IEEE Transactions on Signal Processing*, vol. 59, no. 4, pp. 1409–1420, 2011.

[10] K. Granström and U. Orguner, "A PHD filter for tracking multiple extended targets using random matrices," *IEEE Transactions on Signal Processing*, vol. 60, no. 11, pp. 5657–5671, 2012.

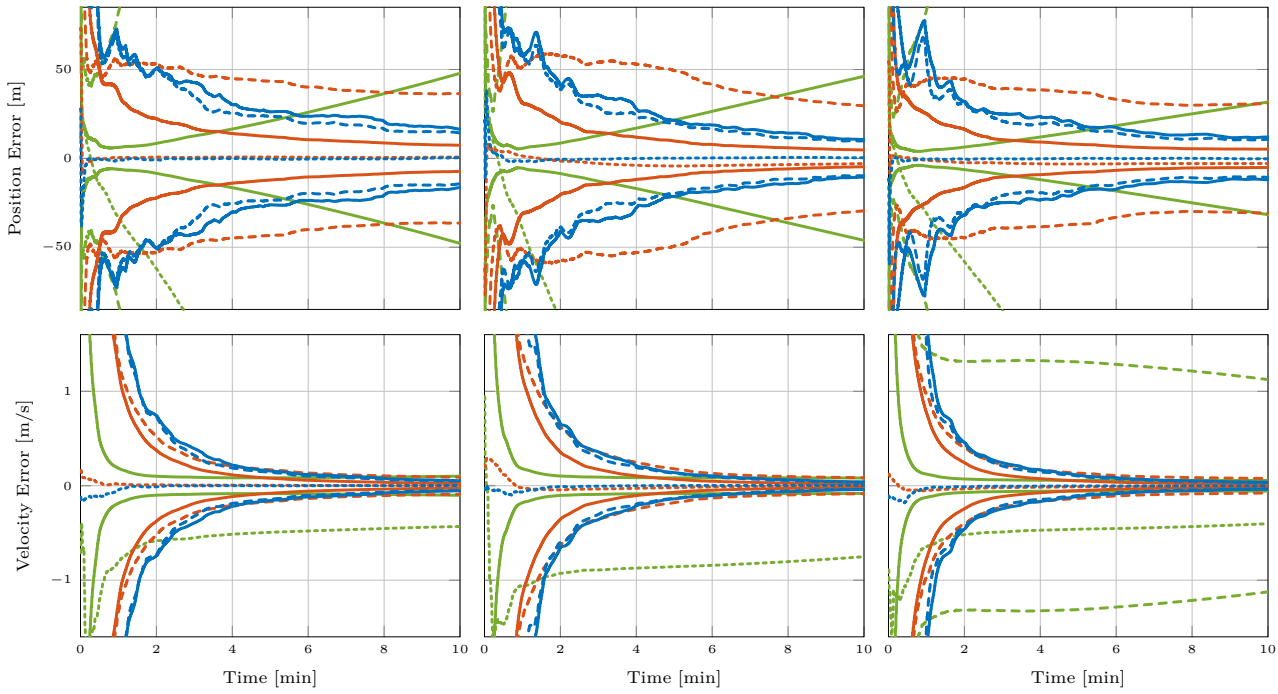


Fig. 4. Estimation results for the vehicle's inertial position and velocity, resulting from 1000 trial Monte Carlo simulations for the EKF (green), GM-AFP (orange), and AEI update (blue). Monte Carlo error statistics ($3\sigma$) are shown as dashed lines, while solid lines represent the averaged filter predicted error covariance ($3\sigma$) and the dotted line is the mean error across the 1000 trials. Note that for each configuration, any trial exceeding $9\sigma$ is excluded from analysis.

TABLE II
MONTE CARLO STATISTICS FROM 1000 TRIALS AFTER 10 MINUTES SIMULATION TIME FOR BOTH POSITION AND VELOCITY.

| Method | Mean Error | Average Filter ($1\sigma$, RSS) | Sample Covariance ($1\sigma$, RSS) | Trials Exceeding $6\sigma$ |
|---|---|---|---|---|
| AEI | 0.54 m, 5.01 cm/s | 7.66 m, 25.45 cm/s | 6.81 m, 23.36 cm/s | 1.5% |
| AFP | 4.16 m, 4.80 cm/s | 3.38 m, 11.73 cm/s | 18.70 m, 48.43 cm/s | 75.8% |
| EKF | 625.31 m, 955.45 cm/s | 25.54 m, 47.45 cm/s | 665.36 m, 1095.63 cm/s | 98.6% |

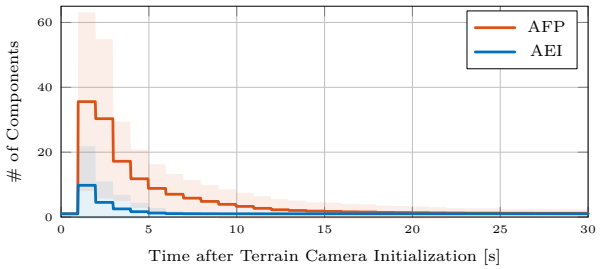


Fig. 5. Mean number of posterior GM components for each method over 1000 Monte Carlo trials – 99.7% sample intervals shown by shaded region.


[11] K. Granström, A. Natale, P. Braca, G. Ludeno, and F. Serafino, "Gamma Gaussian inverse Wishart probability hypothesis density for extended target tracking using X-band marine radar data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 12, pp. 6617–6631, 2015.

[12] R. Mahler, "PHD filters for nonstandard targets, I: Extended targets," in *2009 12th International Conference on Information Fusion*, 2009, pp. 915–921.

[13] R. P. Mahler, *Advances in statistical multisource-multitarget information fusion*. Artech House, 2014.

[14] S. F. Schmidt, "Application of state-space methods to navigation problems," in *Advances in Control Systems*. Elsevier, 1966, vol. 3, pp. 293–340.

[15] J. S. McCabe and K. J. DeMars, "Considering uncertain parameters in non-Gaussian estimation for single-target and multitarget tracking," *Journal of Guidance, Control, and Dynamics*, vol. 40, no. 9, pp. 2138–2150, 2017.

[16] P. G. Savage *et al.*, *Strapdown analytics*. Strapdown Associates Maple Plain, MN, 2000, vol. 2.

[17] Y. Ho and R. Lee, "A Bayesian approach to problems in stochastic estimation and control," *IEEE Transactions on Automatic Control*, vol. 9, no. 4, pp. 333–339, 1964.

[18] J. Lan and X. R. Li, "Tracking of extended object or target group using random matrix: New model and approach," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 6, pp. 2973–2989, 2016.

[19] K. C. Ward, G. S. Fritsch, J. C. Helmuth, K. J. DeMars, and J. S. McCabe, "Design and analysis of descent-to-landing navigation incorporating terrain effects," *Journal of Spacecraft and Rockets*, vol. 57, no. 2, pp. 261–277, 2020.

[20] S. J. Robbins, "A new global database of lunar impact craters > 1–2 km: 1. crater locations and sizes, comparisons with published databases, and global analysis," *Journal of Geophysical Research: Planets*, vol. 124, no. 4, pp. 871–892, 2019.