Spooky Coordinated Tasking and Estimation on Uninformative Priors

Samuel Fedeler
University of Colorado at Boulder
Marcus Holzinger
University of Colorado at Boulder

William Whitacre
The Charles Stark Draper Laboratory, Inc.

August 30, 2021

ABSTRACT

For a variety of mission-critical objectives in Space Domain Awareness (SDA), it is relevant to maintain estimates over large regions of state space using highly non-Gaussian distributions. This problem is broadly applicable in initial orbit determination, maneuver detection, and collision avoidance, and observing targets in this context often necessitates search over the feasible region. Often, this search process results in a large set of null detections. This paper explores methodologies for utilizing negative information in this context, with a special focus on the well-known admissible region. Given a Gaussian mixture representation of the admissible region, a novel methodology for splitting mixands in an arbitrary measurement space is presented. A mixand weight update is derived for the key scenario in which no detection is made at a measurement epoch. Merging methodologies are applied, and the resultant Gaussian sum filter is demonstrated for a representative case in which follow-on tracking of a geostationary object is desired.

1. INTRODUCTION

Whether the goal in consideration is near-Earth asteroid detection, debris mitigation, or collision avoidance, the ability to quickly make follow-up observations on space objects (SOs) is imperative. Often, these goals are further challenged by the prescence of underdetermined or uncertain target states. A newly detected SO may not have been fully observed. A target object may have maneuvered, introducing further uncertainty. If there are collision or operational concerns, urgency is the most critical component in the process of target recovery. This motivates extension of the sensor tasking problem to consider search over feasible regions of state space in an optimal manner, recovering targets as quickly as possible.

Generally, methodologies for driving sensor tasking can be categorized depending on the objective considered. First, one may wish to maintain existing estimates, informing knowledge on a catalog of SOs. A variety of strategies have been proposed assuming a priori knowledge on state estimates and uncertainties. Erwin et al apply linear optimization to form a tasking solution and define useful quantities for interpreting the value of a tasking decision [5]. This work is extended by Williams et al, using Lyapunov exponents to probe the stability of SO estimates [25]. A variety of approaches have also taken inspiration from the machine learning literature, with techniques such as stochastic gradient ascent [23], asynchronous actor-critic methods [18], and Monte Carlo Tree Search [6]. In each of these methods, the driving goal is determination of an optimal policy for decision making given a large set of candidate observations.

Alternately, one may wish to generate new state estimates, expanding the set of SOs studied by searching for natural objects, orbiting satellites, or debris. Wide-ranging techniques for this objective exist in literature. Often, long-period stares over an optical field are performed, acting as a sweep through orbital parameter space [1]. Striping methodologies may also be formed in measurement space, and this strategy accomodates optimization [9].

It is also important to note that detections made with optical sensors generally do not fully observe the object state; as a result of this "Too Short Arc" problem, an admissible region (AR) [19] of unobservable ranges ρ and range rates $\dot{\rho}$ may be formed. This admissible region is a two-dimensional manifold of feasible pairs $(\rho, \dot{\rho})$ that may be projected into the six-dimensional state space. Note that this region may be uniformly distributed or extended to incorporate

measurement uncertainty [26]. Gehly et al. leverage the AR methodology in tandem with Finite Set Statistics to approach the tracking problem, representing the admissible region as a Gaussian mixture to be ingested by a CPHD filter [11]. Methodologies for generating Gaussian mixture representations of admissible regions are introduced by Demars and Jah [4]. AR pairs over longer observation intervals may be used for initial orbit determination [10, 22]. These methodologies are not typically used in an online manner, but rather consider large populations of admissible regions generated from detected tracklets over several observation campaigns.

This literature illustrates the need for follow-up observation to fully observe target states. Tasking in the context of this objective also becomes critical when considering maneuvering targets. Jaunzemis applies the Dempster-Shafer theory of evidence to this problem with success [16]. Decision theoretic approaches have also been applied in tandem with multiple model filters [24]. It is recognized that in either scenario, it becomes challenging to locate a target object, because the projection of the admissible region or reachable set for a target may become quite large relative to the sensor field of view.

The problem of exhausting a feasible set has been explored by Hobson [15] and Murphy[20]. Murphy considers the direct follow-up tasking problem for an admissible region, representing the region as a feasible set that has grown over time in state space. The area of the admissible region is computed over time using high order Taylor series expansions, and the idea of minimizing the search set using the divergence of the set as an observational metric is considered. Simulated annealing in combination with this observational metric is found to achieve some success in minimizing the search space in a time-optimal manner. This research was extended by Fedeler [7], where further search heuristics were developed and leveraged in combination with Monte Carlo Tree Search.

The primary goal of this research is development of an efficient estimation scheme to be utilized in tandem with a prior tasking methodology [7]. Of particular interest is whether the "spooky effect" is apparent, as demonstrated by multi-target filters such as the Gaussian mixture Cardinalized Probability Hypothesis Density (CPHD) filter [8]. This behavior describes the impact of a missed or null detection in one region of measurement space on the probability hypothesis density arbitrarility far away. A logical extension of this effect, then, is determining how a null detection in a subset of the projected feasible set may affect knowledge on other regions within that volume.

With this purpose in mind, the following developments are outlined. First, a brief overview of the process for generating a mixture representation of an admissible region is presented. A mixand weight update is then derived for null detections. Within this update, the expectation propagation [3] methodology for efficient Gaussian integration is utilized. A novel methodology for splitting mixands is then outlined; this methodology was developed to ensure that updates avoid non-Gaussianity. Finally, a merging methodology is introduced to ensure the update process remains computationally efficient. The resultant filter is then fully outlined and applied to a geostationary follow-up tasking problem.

2. METHODS

In order to instantiate an estimation scheme, one must first consider how to represent the probability density function (PDF) over the feasible set within which the target lies. The probability density over this region may be uniform [19] or non-uniform, as in the case of the probabilistic admissible region [26] or a reachable set generated with some a priori knowledge on maneuver probability. It is clear that it is illogical for the probability density to be represented by a univariate Gaussian, and that the density is typically strongly non-Gaussain. Particle representations may be considered, but these methods suffer from a curse of dimensionality and quickly become computationally expensive. Thus, the DeMars method [4] for instantiating a Gaussian mixture representation of probablity density for an admissible region is considered in further detail.

2.1 Gaussian Mixture Approximation for Admissible Regions

Broadly, the AR approximation process may be considered an extension of the problem of approximating a univariate uniform distribution. This approximation may be solved as a root-finding problem with several constraints on the structure of mixands. Assuming equal variance σ^2, equal weights $\omega = \frac{1}{L}$ for L mixands, and evenly distributed mixand means μ, an optimization may be performed to determine optimal variances. Specifically, the L_2 distance [4] between the mixture PDF q and a uniform PDF p with support on the closed interval $[a,b]$ is

$$L_2[p||q] = \frac{1}{b-a} + \frac{\omega^2}{2\sqrt{\pi}\sigma}\sum_{i=1}^{L}\sum_{j=1}^{L}\exp\left(-\frac{1}{4}\left(\frac{\mu_i-\mu_j}{\sigma}\right)^2\right) - \frac{\omega}{b-a}\sum_{k=1}^{L}\left(\text{erf}(B_k)-\text{erf}(A_k)\right) \tag{1}$$

$$A_k = \frac{a-\mu_k}{\sqrt{2}\sigma} \quad B_k = \frac{b-\mu_k}{\sqrt{2}\sigma}. \tag{2}$$

The derivative of the L_2 distance with respect to standard deviation may then be explicitly or numerically computed using any optimizer. In this work, the Levenberg-Marquardt algorithm is utilized with support from the Eigen template library [13].

Successive one-dimensional approximations are appliedo extend this methodology to the two-dimensional problem of AR approximation. First, consider that at any range within the AR, the marginalized density over range rate is uniform. However, the range-marginal PDF, computed as the integral over range rate, is generally not uniform. To account for this, a further optimization may be performed on mixand weights once the univariate uniform distribution has been approximated over the support of the range-marginal PDF. In order to perform this optimization, M ranges must be sampled over the support of the range-marginal PDF. One may then compute the $i \times j$ likelihood matrix Λ, where $\Lambda(i,j)$ is the likelihood range i is drawn from mixand j. The vector $\vec{p}$ is then computed by evaluating the range-marginal PDF at each sampled range. A least-squares problem may then be formulated as

$$\min J = ||\vec{p} - \Lambda\vec{\omega}|| \text{ subject to } \vec{\omega} \geq \vec{0} \text{ and } \vec{1}^T\vec{\omega} = 1 \tag{3}$$

where $\vec{\omega}$ is the concatenated set of weights.

Finally, taking the mean range at each mixand in the approximated range-marginal PDF, range rate may be incorporated. Considering range rate as an independent random variable, standard deviation in range rate may be optimized using the same process. Mixands may be augmented with no correlation between range and range rate uncertainties. Note that care is taken to determine the number of mixands L to utilize in each case as a function of minimum desired variance in range and range rate. Further detail on this process may be found in [4]. The resultant mixands, augmented with the measurement utilized to generate the AR $\vec{y}$ with measurement uncertainty R, may then be linearly transformed into state space.

2.2 Measurement updates in the presence and absence of observations

Given the generated set of mixands, one may now consider how mixands are updated when tasking decisions are made and measurements are taken. The simpler scenario is that of a newly made detection. Here, a typical measurement update for a Gaussian sum filter may be performed.

It is important to first note the measurement model utilized in this work. Optical sensors make angular detections on the celestial sphere, and during long exposures, angular rates may also be determined. Right ascension and declination measurements are computed as

$$\alpha = \arctan\left(\frac{\rho_y}{\rho_x}\right) \tag{4}$$

$$\delta = \arcsin\left(\frac{\rho_z}{|\rho|}\right) \tag{5}$$

and

$$\vec{\rho} = \vec{r} - \vec{r_{obs}} \tag{6}$$

is the expected relative position in an inertial frame. Since the angular measurements are explicitly a function of ρ, one may apply the measurement jacobian $H = \frac{d\vec{y}}{d\vec{\rho}}$, where $\vec{y} = [\alpha \;\; \delta]$, to compute angular rates. Then,

$$[\dot{\alpha} \;\; \dot{\delta}] = \frac{d\vec{y}}{d\vec{\rho}}\dot{\vec{\rho}} = H\dot{\vec{\rho}} \tag{7}$$

$$H = \begin{bmatrix} -\frac{\rho_y}{|\rho_{xy}|^2} & \frac{\rho_x}{|\rho_{xy}|^2} & 0 \\ -\frac{\rho_x\rho_z}{|\vec{\rho}|^2|\rho_{xy}|} & -\frac{\rho_y\rho_z}{|\vec{\rho}|^2|\rho_{xy}|} & \frac{|\rho_{xy}|}{|\vec{\rho}|^2} \end{bmatrix} \tag{8}$$

$$|\rho_{xy}| = (\rho_x^2 + \rho_y^2)^{1/2} \tag{9}$$

and the expected relative velocity vector $\dot{\vec{\rho}}$ is easily computed given knowledge of the observer trajectory and the object state estimate. With angular uncertainty R_θ, angular rate uncertainty is explicitly $R_{\dot{\theta}} = \frac{2R}{\Delta t^2}$ for exposure time Δt. This is found trivially with the assumption of two angular measurements of uncertainty R at the beginning and end of an exposure; the expected value of the angular rate may then be determined as

$$E[\dot{\vec{y}}] = \dot{\hat{y}} = \frac{\vec{y}_2 - \vec{y}_1}{\Delta t} \tag{10}$$

and the covariance of this transformed random variable is also trivially computed.

$$cov(\dot{\vec{y}}) = E[(\dot{\vec{y}} - \dot{\hat{y}})(\dot{\vec{y}} - \dot{\hat{y}})^T] \tag{11}$$

$$= \frac{1}{\Delta t^2} E[(\vec{y}_2 - \vec{y}_1 - \hat{y}_2 + \hat{y}_1)(\vec{y}_2 - \vec{y}_1 - \hat{y}_2 + \hat{y}_1)^T] \tag{12}$$

$$= \frac{1}{\Delta t^2}\left(E[(\vec{y}_2 - \hat{y}_2)(\vec{y}_2 - \hat{y}_2)^T] + E[(\vec{y}_1 - \hat{y}_1)(\vec{y}_1 - \hat{y}_1)^T]\right) = \frac{2R}{\Delta t^2}. \tag{13}$$

Given measurement and measurement uncertainty

$$\vec{y} = \begin{bmatrix}\alpha \; \delta \; \dot{\alpha} \; \dot{\delta}\end{bmatrix}^T, \;\; R = \begin{bmatrix} R_\theta & 0 \\ 0 & R_{\dot{\theta}} \end{bmatrix}, \tag{14}$$

the traditional Gaussian sum filter updates may be performed [2]. In this work, unscented measurement updates are utilized [17], with likelihood-based weight updates in addition such that

$$\omega_i = \frac{\omega_i' \mathcal{L}_i(\vec{y}, R)}{\sum_{j=1}^{L} \omega_j' \mathcal{L}_j(\vec{y}, R)} \tag{15}$$

and

$$\mathcal{L}_i(\vec{y}, R) = \mathcal{N}\left(\vec{y} - \vec{h}(\vec{x}_i), HP_iH^T + R\right). \tag{16}$$

Updating the PDF when no detection is made introduces further complexities. First, one must consider the probabiliity that the target SO lies in a given field of view (FOV) of an optical sensor. If mixand uncertainties are sufficiently small in range space relative to the range between the mixand and the optical sensor, one may assume that the probability of detection p_D is uniform over the mixand. Then, the cumulative likelihood of observing the target is

$$P(\vec{y} \neq \emptyset) = \int_{\text{FOV}} p_D(\vec{z}) P(\vec{z}) d\vec{z}. \tag{17}$$

Applying the mixand representation of the PDF, a distinct probability of detection may be assumed for each mixand, and

$$P(\vec{y} \neq \emptyset) = \int_{\text{FOV}} \sum_{i=1}^{L} p_{D,i} \omega_i P(\vec{z} | k = i) d\vec{z}. \tag{18}$$

One must then evaluate the observation likelihood $P(\vec{z}|k=i)$ conditioned on mixand i. Note that the transformation from state space to measurement space is locally linear with the same assumptions on range. Therfore, the mixand density projected into measurement space is still Gaussian, with

$$P(\vec{z}|k=i) \approx \mathcal{N}\left(\vec{h}(\mu_i),\, HP_iH^T\right). \tag{19}$$

It is also clear that the integral is separable, and thus,

$$P(\vec{y} \neq \emptyset) = \sum_{i=1}^{L}\left(p_{D,i}\omega_i \int_{\text{FOV}} P(\vec{z}|k=i)d\vec{z}\right). \tag{20}$$

To compute the likelihood of observation, one must then evaluate the Gaussian integral in measurement space over the rectangular FOV. Methods such as Genz integration [12] or expectation propagation [3] may be applied for this purpose. Given this result, one must now consider how a null detection may affect existing mixands. It is clear that the probablility a null detection occurs is the complement of Equation 20. Then, one must consider how to apply this result to knowledge on each mixand. Working from first principles, we may apply Bayes' rule.

$$P(\vec{x}|k=i,\vec{y}=\emptyset) = \frac{P(\vec{y}=\emptyset|\vec{x},k=i)P(\vec{x}|k=i)}{P(\vec{y}=\emptyset)} \tag{21}$$

Immediately, challenges arise when considering the term $P(\vec{y}=\emptyset|\vec{x},k=i)$, the probability a null detection is made, conditioned on the SO state captured by mixand i. Consider the PDF for mixand i in further detail, with the temporary assumption that the projection of probability density into measurement space is larger in spread than the sensor FOV. At any point within the support of the projected PDF outside of the FOV, the probability of a null detection must be unity, since that point cannot be captured during the observation. This leads to a scenario in which a subset of the mixand is scaled as a function of the probability of detection, while the remainder is unaffected. The structure of this update is clearly non-Gaussian, and is further illustrated in Figure 1.

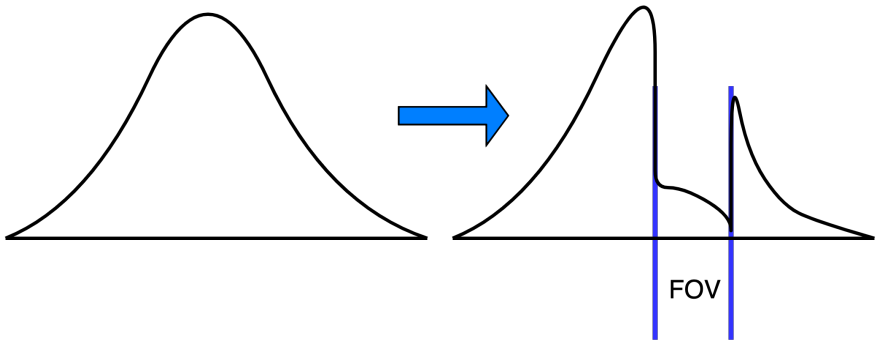


Fig. 1: Non-Gaussianity in a negative information update.

This behavior may be broadly categorized into distinct groups. First, the density captured in the field of view can be relatively small; this commonly occurs when the mixand is either a large normalized distance away from the sensor FOV in measurement space, or quite large in spread in measurement space as compared to the field of view. In this case, a null detection's impact on the PDF is negligible, since the probability $P(\vec{y}=\emptyset|\vec{x},k=i)$ is effectively unity. Alternatively, the probability density may be almost entirely captured by the sensor FOV, in which case the entire mixand is rescaled and the null detection probability for the given mixand nears zero. The third case, in which the sensor field of view overlays a significant portion of the mixand, but not the entirity, requires further consideration. To resolve this case, a splitting method is proposed to reduce the projected spread of mixands in measurement space.

2.3 Oriented Gaussian splitting

Consider a mixand with mean μ and covariance P in state space $\mathbb{S}$. This mixand may be defined relative to an observer $\mathbb{O}$ with measurement function $\vec{h}$. Letting the measurement function be differentiable, the local behavior of $\vec{h}$ may be examined utilizing the gradient. For each scalar measurement, this leads to a tangent that may be normalized and considered as the direction in state space leading to a maximal change in the associated scalar measurement. The resultant set of tangent vectors forms the basis of a tangent space of dimension n, where n is the rank of the measurement Jacobian. Each tangent vector may be explicitly computed as

$$\hat{l}_i = \frac{\frac{\partial h_i}{\partial \vec{x}}\Big|_{\mu}^{T}}{\left|\frac{\partial h_i}{\partial \vec{x}}\Big|_{\mu}^{T}\right|}. \tag{22}$$

It is desired to split the mixand into a set of mixands with unknown means and equivalent covariance P^*, while enforcing that the combined PDF of the resultant mixands captures the same first and second moment of the original mixand. Additionally, it is desired that the observer is also taken into consideration, such that the new mixands are perturbed about the defined measurement bases. Without loss of generality, let the new set contain $N = 2m+1$ mixands, where m is the dimension of the measurement space. Let one mixand be placed at the original mean, with another pair of mixands evenly distanced along the the tangent vector associated with each scalar measurement with distance $a_k P^{\frac{1}{2}}$, where $P^{\frac{1}{2}}$ is the matrix square root of the original mixand covariance. Note that this methodology is typical when considering transformations on Gaussians, exemplified by the work of Havlak and Campbell [14]. The matrix square root is incorporated to ensure that similar mixand structure is in place, and to enforce positive-definiteness. Additionally, let each new mixand have equivalent weight $\omega = \frac{1}{2m+1}$. The mean of the resultant distribution is then

$$\mu_{TOT} = \sum_{k=1}^{2m+1} \omega_k \mu_k \tag{23}$$

$$= \frac{1}{2m+1}\left(\mu + \sum_{k=1}^{m}\left(\mu + a_k P^{\frac{1}{2}}\hat{l}_k\right) + \sum_{k=1}^{m}\left(\mu - a_k P^{\frac{1}{2}}\hat{l}_k\right)\right) = \mu. \tag{24}$$

The covariance of the new distribution must also be determined. For a Gaussian mixture, the total covariance is

$$P_{TOT} = \sum_{k=1}^{N} \omega_i P_i + \sum_{k=1}^{N} \omega_i \left(\mu_i - \mu_{TOT}\right)\left(\mu_i - \mu_{TOT}\right)^T. \tag{25}$$

In this case, then, we find

$$P_{TOT} = \sum_{k=1}^{2m+1} \omega_k P^* + \sum_{k=1}^{2m+1} \omega_k \left(\mu_k - \mu\right)\left(\mu_k - \mu\right)^T \tag{26}$$

$$= P^* + 2\sum_{k=1}^{M} \frac{1}{2m+1} a_k^2 P^{\frac{1}{2}} \hat{l}_k \hat{l}_k^T P^{\frac{T}{2}} = P \tag{27}$$

With this result, one can determine the updated covariance

$$P^* = P - 2\sum_{k=1}^{m} \frac{1}{2m+1} a_k^2 P^{\frac{1}{2}} \hat{l}_k \hat{l}_k^T P^{\frac{T}{2}}. \tag{28}$$

With this result in mind, it is still important to consider that whether P^* is positive definite. The key determination is whether the eigenvalues of P^* are all positive. It is possible to left and right multiply P^* by $P^{-\frac{1}{2}}$ and maintain the definiteness of the matrix such that

$$P_{norm} = P^{-\frac{1}{2}} P^* P^{-\frac{T}{2}} \tag{29}$$

$$= I - 2\sum_{k=1}^{m} \frac{1}{2m+1} a_k^2 \hat{l}_k \hat{l}_k^T \tag{30}$$

Now, any eigenvector for the summation must also be an eigenvector of P_{norm}. For each eigenvector $\vec{v}_i$, the assoicated eigenvalue of the summation is λ_i, and the eigenvalue of P_{norm} must be enforced to be strictly greater than zero such that

$$\lambda_P = 1 - \lambda_i > 0. \tag{31}$$

This allows for gains a_k to be chosen as a function of the structure of the cumulative outer product of the tangent space. First, consider the outer product $\hat{l}_k \hat{l}_k^T$. Since the tangent vectors are normalized, this matrix is symmetric and positive semi-definite with a single eigenvalue at unity. This offers an upper bound on the eigenvalues of the summation. If each tangent vector is collinear, the summation will then have a single nonzero eigenvalue

$$\lambda^* = \frac{2ma_k^2}{2m+1} \tag{32}$$

that must be less than unity. Gains must in general then be no greater than

$$a_k < \sqrt{\frac{2m+1}{2m}}. \tag{33}$$

Note that these gains may be increased if there is further knowledge of the measurement space. If two unit vectors are orthogonal, one may infer that the sum of outer products of these vectors has two eigenvalues at unity in addition to zero eigenvalues. This argument may be expanded, considering the full set of tangent vectors in the space. The maximum eigenvalue of the summed matrix must be no greater than, assuming gain is held constant,

$$\lambda^* = \frac{2a_k^2}{2m+1}\left(m+1-\text{rank}(H)\right), \tag{34}$$

where $\text{rank}(H)$ is the rank of the measurement Jacobian, which is equivalent to the rank of the set of tangent vectors.

This result may explicitly be demonstrated for an optical case in which right ascension and declination measurements are taken. This is the critical case for negative information updates, because projected mixands must be split in angular space to ensure the update remains Gaussian. For this case, the dimension of the measurement is $m = 2$. It is also known that the gradients of right ascension and declination are orthonormal in state space. Therefore, we have

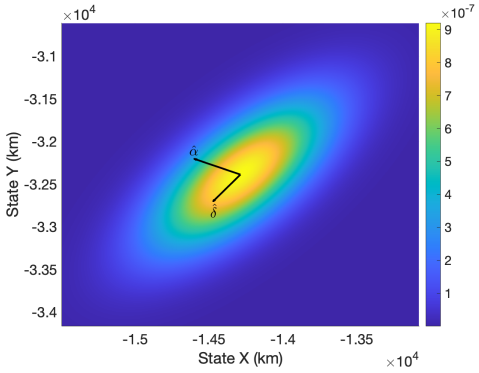
$$\lambda^* = \frac{2a_k^2}{5} < 1 \tag{35}$$

and

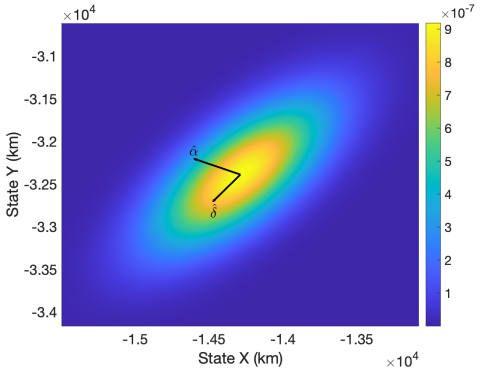
$$a_k < \sqrt{\frac{5}{2}} \tag{36}$$

It is then assured that newly generated mixands continue to have positive definite covariances as splitting occurs. Note that while this gain is the theoretical maximal bound to ensure positive definite covariances, it is not advisible to use. As gain nears this quantity, the projected PDF of the split mixands will become drastically small in the eigendirection associated with the eigenvalues nearing zero. This behavior is visualized in Figure 2, where the PDF is presented for a mixand before a split, with a split using small gains, and a split with gains nearing the theoretical maximum. With this projection, it is most clear that mixand uncertainties become quite small in the $\hat{\alpha}$ direction, and they also become quite small in the $\hat{\delta}$ direction (that is largely out of the plane). More specifically, as the gain approaches the theoretic maximum, the projections of uncertainties into measurement space collapse into discrete points. This behavior is demonstrated in Figure 3.

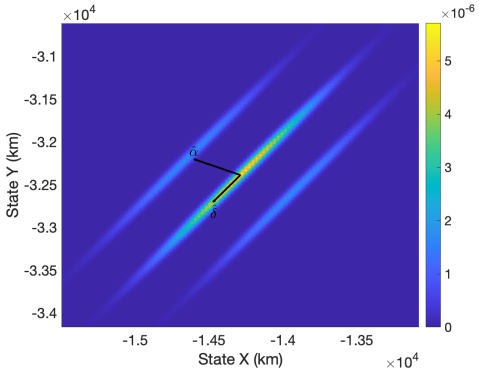
Understanding this behavior helps visualize a trade in gain selection. Increased gain ensures a reduction in spread for the mixands utilized, but also increases relative entropy between the original mixand and resultant mixand. Care must be taken to allow new mixands to become as small as needed while maintaining a distribution sufficiently close to the original.



(a) Original mixand. Measurement tangent vectors $\hat{\alpha}$ and $\hat{\delta}$ are also displayed.



(b) Mixands split with gain $a = 1.0$. Density is approximately equal to a).



(c) Mixands split with gain $a = \sqrt{\frac{5}{2}}$. The first and second moments match those of a), but relative entropy is much higher.

Fig. 2: Mixand density in X-Y space before and after splitting.

2.4 Updating Gaussians

With a formulation for splitting Gaussians such that they will be sufficiently small in measurement space in place, a criterion for determining whether a mixand shall be split must be established. It is logical to incorporate some measure of offset from the center of the sensor FOV in measurement space and the comparative spread of uncertainty to the sensor FOV. A critical Mahalanobis distance may be defined, such that a mixand shall only be split if

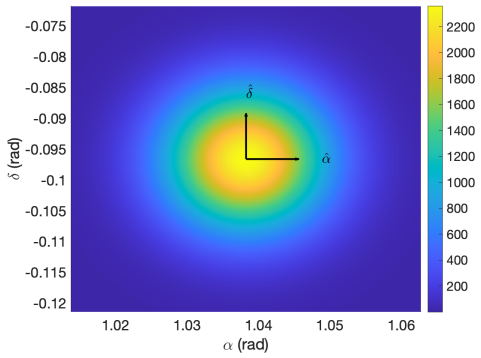
$$D_M(\mu_i, P_i; \mathcal{O}) < d^*. \tag{37}$$

Additionally, the angular spread may be defined as the square root of the maximal diagonal value of the projected covariance trace

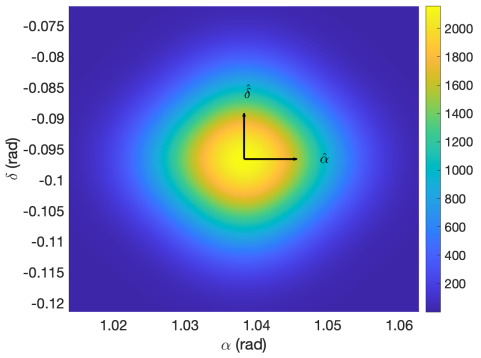
$$s = \sqrt{\max\left(HPH^T\right)} > s^*. \tag{38}$$

This may be compared with a critical value that is a function of the diagonal field of view of the sensor.

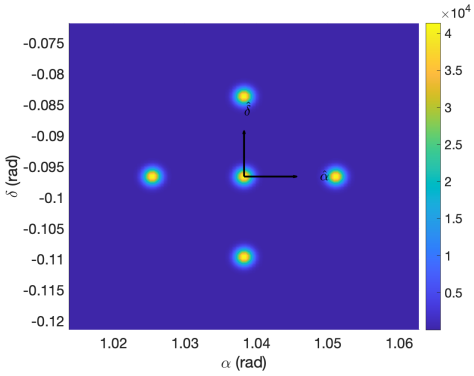
Once mixands are rescaled such that they are either sufficiently distant in measurement space from the sensor FOV or of comparable size to the sensor FOV, the negative information update may be considered in further detail. Revisiting Equation 21, the problematic term $P(\vec{y} = \emptyset | \vec{x}, k = i)$ may now be considered approximately discrete in that the mixand is either fully covered by the sensor FOV or is sufficiently far from the sensor. As such, it is now logical to consider Equation 21 as a weight update on each mixand in much the same manner as a particle filter, using the intermediate



(a) Original mixand projected into measurement space.



(b) Mixands split with gain $a = 1.0$ in measurement space.



(c) Mixands split with gain $a \approx \sqrt{\frac{5}{2}}$. Mixands are reduced into discrete points.

Fig. 3: Mixand density in measurement space before and after splitting.

density

$$g(\vec{y}|k=i) = \frac{P(\vec{y}=\emptyset|\vec{x},k=i)}{P(\vec{y}=\emptyset)} \tag{39}$$

$$= \frac{P(\vec{y}=\emptyset|\vec{x},k=i)}{\sum_{j=1}^{L} P(\vec{y}=\emptyset|\vec{x},k=j)}. \tag{40}$$

The denominator can simply be considered a normalization, while the numerator may be approximately evaluated as

$$P(\vec{y}=\emptyset|\vec{x},k=i) \approx \begin{cases} 1 & i \notin \text{FOV} \\ 1-p_D & i \in \text{FOV} \end{cases} \tag{41}$$

Note that it is still useful to explicitly compute the Gaussian integrals over the FOV because of the stopping criterion on splitting; these results may be scaled by the tail probabilities computed.

2.5 Merging and Filter Outline

With the filter update fully expressed, one now must ensure there is no hypothesis explosion in mixands so that the filter remains computationally efficient. With a goal of minimizing Kullbeck-Liebler divergence during the merging process, the well-known Runnall's method is utilized [21]. A discrimination bound,

$$B(i,j) = \frac{1}{2}\left[(\omega_i+\omega_j)\log|P_{ij}| - \omega_i \log|P_i| - \omega_j|P_j|\right], \tag{42}$$

may be iteratively computed with the merged covariance for mixands i and j P_{ij}. Merging is iteratively performed until a threshold maxima of mixands is reached. Pruning may also be applied if weights are sufficiently small, but care must be taken to ensure that this does not disregard mixands split during the negative information update. With this methodology in place, the full filter is outlined in Figure 4.

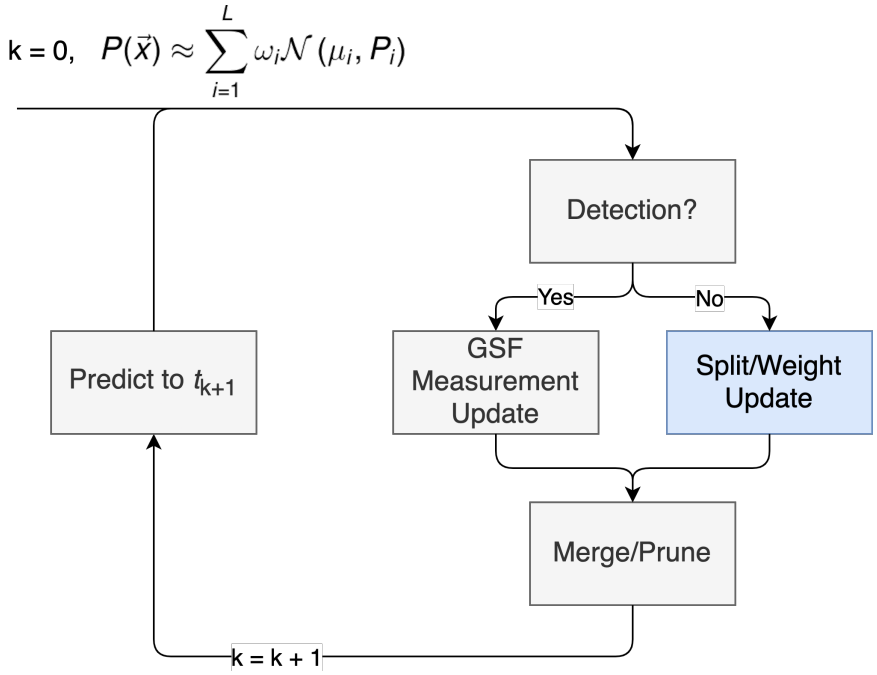


Fig. 4: Gaussian Sum Filter diagram. Major contribution highlighted in blue.

3. RESULTS

The methodology is now presented for a case in which a prior detection is made and follow-up observation is desired. The object tracked has the true initial state (in kilometers and kilometers per second)

$$\vec{x} = \begin{bmatrix} -27100 & -32300 & -100 & 2.36 & -1.98 & 0 \end{bmatrix} \tag{43}$$

in the Earth-centered inertial (ECI) frame. An observation is made by an observer at the initial ECI position (in kilometers)

$$\vec{o} = \begin{bmatrix} 517.859 & -5281.538 & 3526.190 \end{bmatrix}. \tag{44}$$

An admissible region is then formed from knowledge of observer state and the attributable vector (in radians and radians per second) as

$$\vec{y} = \begin{bmatrix} \alpha & \delta & \dot{\alpha} & \dot{\delta} \end{bmatrix} = \begin{bmatrix} -2.36716 & -0.093581 & 7.30762e-05 & -1.53752e-09 \end{bmatrix}. \tag{45}$$

To constrain the admissible region, several assumptions are made on feasible orbits. First, the assumption is made that the target SO is on an ecliptic trajectory following two body dynamics, with energy less than 0. Next, a maximal eccentricity of 0.3 is applied. Finally a minimum radius of periapsis of 6500 km is assumed. The admissible region is approximated using a mixture representation. The resultant admissible region is visualized in Figure 5. Here, the feasible set is presented in the range-range rate halfplane; note that this is the unobservable subspace over which the admissible region is determined. The admissible region is then propagated forward in time until the next possible

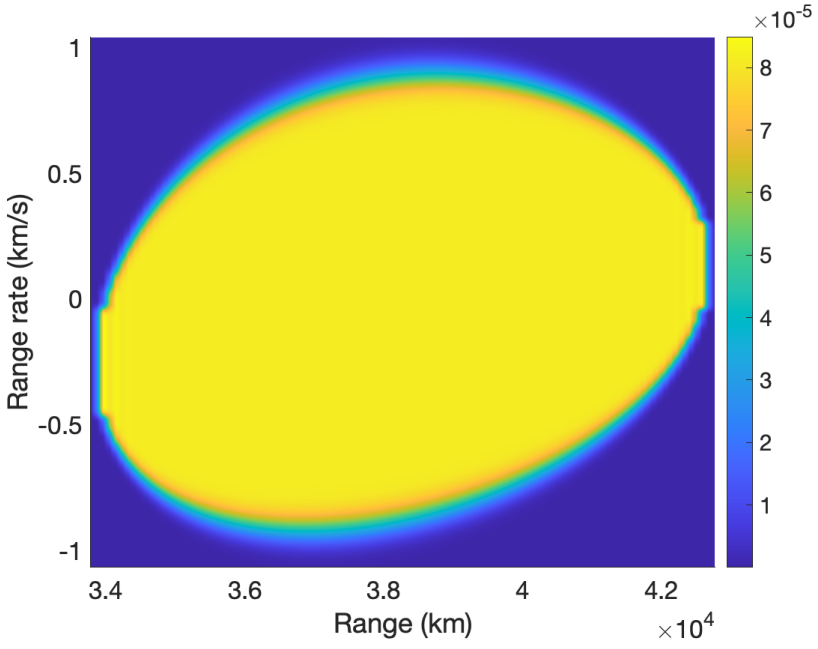


Fig. 5: The initial admissible region in the range range-rate halfplane.

observation period, assumed to be 2 hours. The propagated admissible region is then visualized in measurement space in Figure 6.

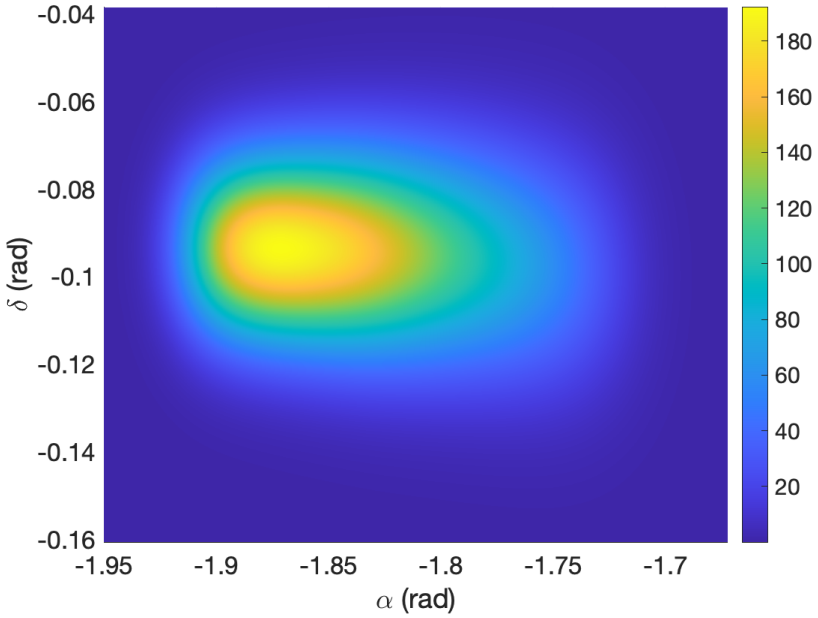


Fig. 6: The initial admissible region propagated and projected into measurement space.

Several characteristics may be noted from these figures. First, probability density is uniform over the unobservable subspace, as no information is available on the SO range or range rate. Therefore, each feasible range and range rate pair has equal probability. However, this is not true when this feasible set is projected into another space. One may

consider that the admissible region initially manifests as a point in measurement space, where it was fully observed. As the admissible region is propagated over time, the feasible region then grows in measurement space through a combination of dynamical uncertainty and rotation of the field of view of the observer, allowing the state to eventually become fully observable. Because different subsets of the admissible region grow at different rates, and because the projection of the admissble region into measurement space may still be concentrated around the mean angular rates over the set, the projection is decidedly not uniform in measurement space. Figure 6 demonstrates this fact, and this behavior is important to note when considering searching for the SO over the region. A successful tasking methodology captures two objectives in approaching this goal. First, it is critical to consider what regions of the feasible set are growing quickly in measurement space so that it is possible to exhaust the admissible region over time. Second, it is critical to search over subregions with high probability density to maximize probability of detection.

The methodology of Fedeler et al. [7] attempts to accomodate these objectives, and is utilized in this demonstration to generate a tasking solution. Observations are taken at a 15 second cadence until the search region is exhausted; the first follow-up observation is performed two hours after the initial dection. The admissible region, with initial area in measurement space of approximately 7 deg^2, is exhausted with a sequence of 66 observations, using a sensor with a square field of view of 0.25 deg^2. Note that this region is growing quickly in measurement space, and at the time of the 66th observation, the projected admissible region has an area of approximately 12 deg^2. Over the course of this tasking solution, a single observation is made at timestep $t_{15} = 7410$ s, when the observer is pointing at the angular coordinates $\alpha = -1.82939$ rad, $\delta = -0.093562$ rad. In each other case, no detection is made, and the negative information is processed. Figure 7 visualizes the effects of processing negative information 7 observations into the tasking scenario. Note that probabiliity density the observed region of measurement space has been greatly reduced.

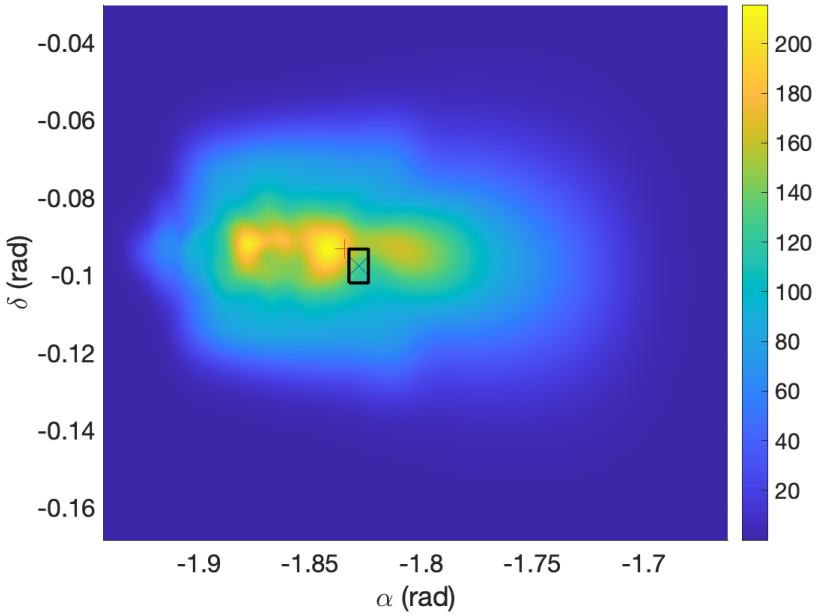


Fig. 7: The admissible region projected into measurement space after 7 null detections.

In the figure, the most recent observation is represented by a shaded rectangle, and the true state, at $\alpha = -1.8346$ rad and $\delta = -0.09318$ rad, is marked by a plus sign. Note that previous observations have greatly reduced probability density in other areas of the projection, most visibly in the region from $\alpha = -1.87$ to -1.85 radians and $\delta = -0.1$ to -0.09 radians. Because of these reductions, unobserved subsets of the projected admissible region are now comparatively more likely. Indeed, the density at the true state is approximately 10 percent higher than prior to the processing of any negative information. This result is quite comparable in essence to the spooky effect described by Franken et al [8]. The negative information in this case describes missed detections on a subset of mixands in the ensemble, increasing the likelihood that each other mixand is "truth" and may be associated with the true state.

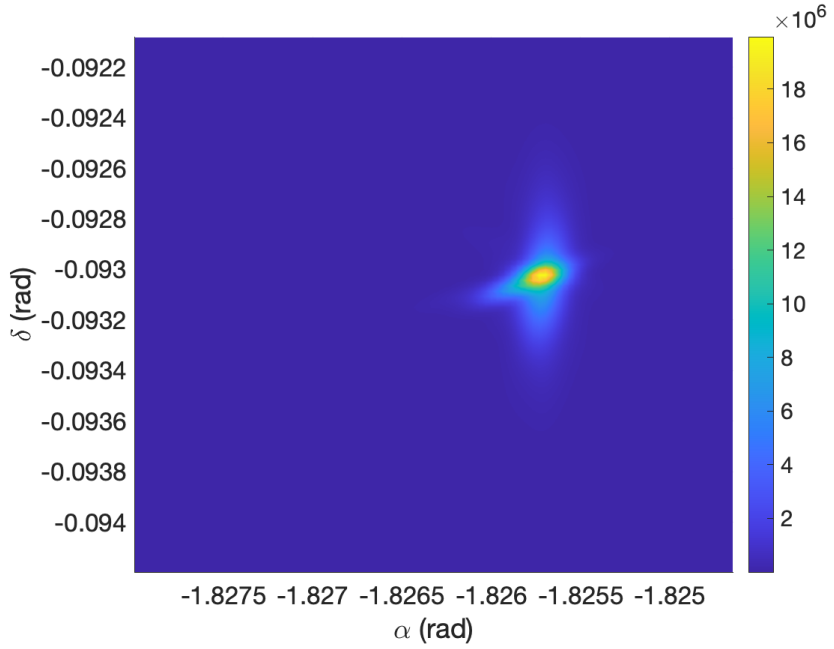


Fig. 8: The admissible region projected into measurement space after a follow up detection is made. Uncertainty in angular space is equivalent to measurement uncertainty.

It is also important to consider the behavior of the filter when a measurement is received. Figure 8 demonstrates the reduction of the mixture when this occurs at time t_{15}. Here, the projected area of the mixture is now reduced to that of measurement uncertainty, assumed to be on the order of 5 arcseconds in this simulation. After this observation is made, there is negligible effect on the state estimate through further processing of negative information, but this is still critical to do in real scenarios, when the likelihood of false alarm measurements is non-negligible. Finally, Figures 9 and 10 visualize estimation error over the course of the simulation. Note that the estimation error remains within the covariance bounds throughout the simulation, and is greatly reduced when the follow-up observation is received.

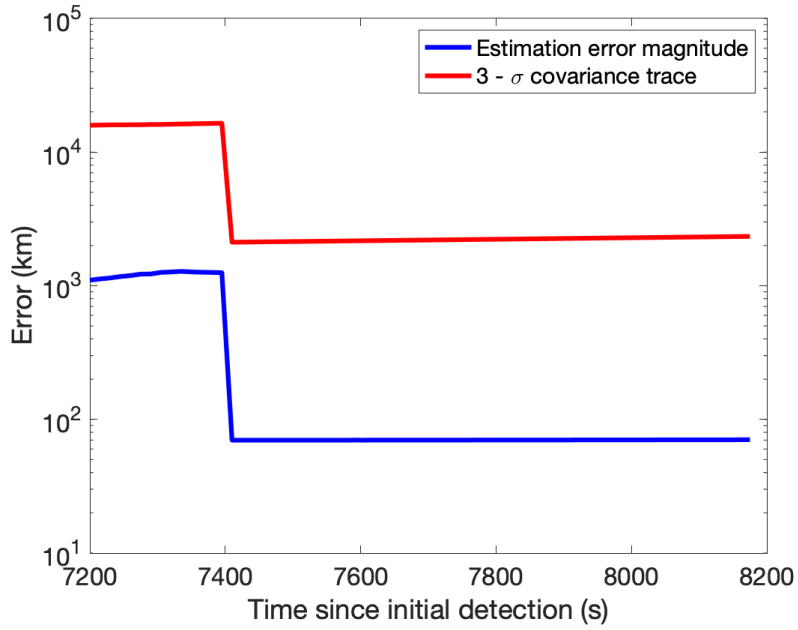


Fig. 9: Positional estimation error over the course of the simulated observation campaign.

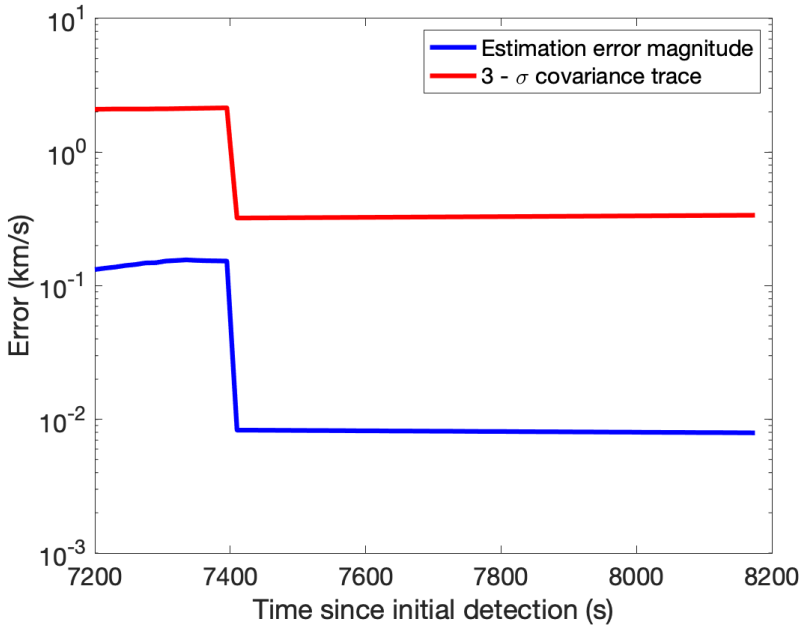


Fig. 10: Velocity estimation error over the course of the simulated observation campaign.

The covariance bounds are a bit conservative in this context because of the time between observations. Interestingly, the full state is only weakly observable, and there is still significant uncertainty in the sensor line of sight direction after the measurement, on the order of 2000 km $3-\sigma$. However, uncertainty is sufficiently reduced such that only a small number of mixands are needed to sufficiently represent probability density after the measurement. There is little impact on the processing of negative information on estimation error, but this is as expected. Before a detection is made, the pdf is still quite large in state space. While some mixands are eliminated through processing negative

information, the ensemble mean is not very useful for such a large pdf; in this case, it just happens to occur that the original mean is somewhat close to the true state. It would be interesting to determine whether the ensemble mean becomes more useful in cases where most of the feasible region is exhausted before a follow-up detection is made.

4. CONCLUSION

Methodologies for utilizing negative information have been developed and discussed. As a primary contribution, a methodology for splitting mixands in an arbitrary measurement space is presented. While the logical focus for this methodology in this paper is the context of optical observations, this methodology could also be extended to consider radar measurements and a variety of alternative observational techniques. The key use case for the developed work is any situation in which the spread of the projected state estimate exceeds the sensor field of regard in measurement space.

Theoretic bounds on the size of splits are outlined. It is noted that further perturbation along measurement axes comes with an increase in Kullbeck-Liebler divergence from the original mixture. Future work may aim to formally pose this splitting methodology as a multi-objective optimization, in which Kullbeck-Liebler divergence is minimized with an additional cost objective comparing the spread of resultant mixands to a target value (for example, a sensor FOV).

This splitting methodology is utilized to ensure that the weight update remains Gaussian. Here, it is important to consider that particle representations of the state estimate may also be utilized, with likelihood updates on particles within a sensor FOV when no detection is made. The Gaussian mixture representation may be considered advantageous in this context for several reasons, though. First, it allows for a smooth representation of probability density over the support of the state estimate. This is quite useful in regions of measurement space where particles may be quite sparse. Additionally, there are computational advantages to utilizing Gaussian mixture representations, in that a large set of particles need not be sampled; this is especially critical when particles are sampled from a high-dimensional state space. In the context of admissible regions, this is not the case, but a curse of dimensionality becomes rather important when considering search over reachable sets in Cartesian position-velocity space.

Finally, the methodology is presented for a geostationary follow-up tasking case. Results demonstrate advantages in utilization of negative information before follow-up detections are made. Extensions of these results will be made considering a variety of target orbits, utilization of more interesting dynamics, and follow-up tracking of maneuvering objects.

5. ACKNOWLEDGEMENTS

Research on this project has been supported by the National Science Foundation Graduate Research Fellows Program and the Draper Scholars.

REFERENCES

[1] John Africano, Thomas Schildknecht, Mark Matney, Paul Kervin, Eugene Stansbery, and Walter Flury. A Geosynchronous Orbit Search Strategy. *Journal of Allergy and Clinical Immunology*, 1(2):357–369, 2004.

[2] Daniel L. Alspach and Harold W. Sorenson. Nonlinear bayesian estimation using gaussian sum approximations. *IEEE Transactions on Automatic Control*, 17(4):439–448, 1972.

[3] John P. Cunningham, Philipp Hennig, and Simon Lacoste-Julien. Gaussian Probabilities and Expectation Propagation. 2:1–56, 2011.

[4] Kyle J. DeMars and Moriba K. Jah. Probabilistic initial orbit determination using Gaussian mixture models. *Journal of Guidance, Control, and Dynamics*, 36(5):1324–1335, 2013.

[5] R. Scott Erwin, Paul Albuquerque, Sudharman K. Jayaweera, and Islam Hussein. Dynamic sensor tasking for space situational awareness. *Proceedings of the 2010 American Control Conference, ACC 2010*, pages 1153–1158, 2010.

[6] Samuel J Fedeler, Marcus J Holzinger, and William Whitacre. Optimality and Application of Tree Search Methods for POMDP-based Sensor Tasking. *Advanced Maui Optical and Space Surveillance (AMOS) Technologies Conference*, XX(X), 2020.

[7] Samuel J Fedeler, Marcus J Holzinger, and William Whitacre. Optimally Convergent Minimum-Time Space Object Search and Recovery. In *AAS Space Flight Mechanics Meeting*, pages 1–18, 2021.
[8] D. Fränken, M. Schmidt, and M. Ulmke. "Spooky action at a distance" in the cardinalized probability hypothesis density filter. *IEEE Transactions on Aerospace and Electronic Systems*, 45(4):1657–1664, 2009.
[9] Carolin Frueh, Hauke Fielder, and Johannes Herzog. Heuristic and optimized sensor tasking observation strategies with exemplification for geosynchronous objects. *Journal of Guidance, Control, and Dynamics*, 41(5):1036–1048, 2018.
[10] K. Fujimoto and K. T. Alfriend. Optical short-arc association hypothesis gating via angle-rate information. *Journal of Guidance, Control, and Dynamics*, 38(9):1602–1613, 2015.
[11] Steven Gehly, Brandon A. Jones, and Penina Axelrad. Search-Detect-Track Sensor Allocation for Geosynchronous Space Objects. *IEEE Transactions on Aerospace and Electronic Systems*, 54(6):2788–2808, 2018.
[12] Alan Genz and Giang Trinh. Numerical computation of multivariate normal probabilities using bivariate conditioning. *Journal of Computational and Graphical Statistics*, (1):141–149, 1992.
[13] Gaël Guennebaud, Benoît Jacob, et al. Eigen v3. http://eigen.tuxfamily.org, 2010.
[14] Frank Havlak and Mark Campbell. Discrete and continuous, probabilistic anticipation for autonomous robots in Urban environments. *IEEE Transactions on Robotics*, 30(2):461–474, 2014.
[15] Tyler A. Hobson and I. Vaughan L. Clarkson. A particle-based search strategy for improved Space Situational Awareness. *Conference Record - Asilomar Conference on Signals, Systems and Computers*, pages 898–902, 2013.
[16] Andris D. Jaunzemis, Marcus J. Holzinger, and K. Kim Luu. Sensor tasking for spacecraft custody maintenance and anomaly detection using evidential reasoning. *Journal of Aerospace Information Systems*, 15(3):131–156, 2018.
[17] Simon J Julier and Jeffrey K Uhlmann. New extension of the Kalman filter to nonlinear systems. *Proceedings of SPIE*, (7), 1997.
[18] Richard Linares and Roberto Furfaro. An Autonomous Sensor Tasking Approach for Large Scale Space Object Cataloging. *Advanced Maui Optical and Space Surveillance Technologies Conference*, pages 1–17, 2017.
[19] A. Milani, M. E. Sansaturio, G. Tommei, O. Arratia, and S. R. Chesley. Multiple solutions for asteroid orbits: Computational procedure and applications. *Astronomy & Astrophysics*, 431(2):729–746, 2005.
[20] Timothy S. Murphy and Marcus J. Holzinger. Generalized Minimum-Time Follow-up Approaches Applied to Tasking Electro-Optical Sensor Tasking. *Advanced Maui Optical and Space Surveillance (AMOS) Technologies Conference*, 2017.
[21] Andrew R. Runnalls. Kullback-Leibler approach to Gaussian mixture reduction. *IEEE Transactions on Aerospace and Electronic Systems*, 43(3):989–999, 2007.
[22] Paul W. Schumacher, John A. Gaebler, Christopher W.T. Roscoe, Matthew P. Wilkins, and Penina Axelrad. Parallel initial orbit determination using angles-only observation pairs. *Celestial Mechanics and Dynamical Astronomy*, 130(9):1–20, 2018.
[23] Zachary Sunberg, Suman Chakravorty, Richard Scott Erwin, and Senior Member. Information Space Receding Horizon Control for Multisensor Tasking Problems. *IEEE Transactions on Cybernetics*, 46(6):1325–1336, 2016.
[24] Dylan J. Thomas, Kevin M. Nastasi, Kevin Schroeder, and Jonathan T. Black. Autonomous Multi-Phenomenology Space Domain Sensor Tasking and Adaptive Estimation. *2018 21st International Conference on Information Fusion, FUSION 2018*, (1):1331–1338, 2018.
[25] Patrick S. Williams, David B. Spencer, and Richard S. Erwin. Coupling of estimation and sensor tasking applied to satellite tracking. *Journal of Guidance, Control, and Dynamics*, 36(4):993–1007, 2013.
[26] Johnny L. Worthy and Marcus J. Holzinger. Incorporating Uncertainty in Admissible Regions for Uncorrelated Detections. *Journal of Guidance, Control, and Dynamics*, 38(9):1673–1689, 2015.