Uninformative Prior Multiple Target Tracking Using Evidential Particle Filters

Johnny L. Worthy III*, Marcus J. Holzinger†
Georgia Institute of Technology, Atlanta, GA, 30332

Space situational awareness requires the ability to initialize state estimation from short measurements and the reliable association of observations to support the characterization of the space environment. The electro-optical systems used to observe space objects cannot fully characterize the state of an object given a short, unobservable sequence of measurements. Further, it is difficult to associate these short-arc measurements if many such measurements are generated through the observation of a cluster of satellites, debris from a satellite break-up, or from spurious detections of an object. An optimization based, probabilistic short-arc observation association approach coupled with a Dempster-Shafer based evidential particle filter in a multiple target tracking framework is developed and proposed to address these problems. The optimization based approach is shown in literature to be computationally efficient and can produce probabilities of association, state estimates, and covariances while accounting for systemic errors. Rigorous application of Dempster-Shafer theory is shown to be effective at enabling ignorance to be properly accounted for in estimation by augmenting probability with belief and plausibility. The proposed multiple hypothesis framework will use a non-exclusive hypothesis formulation of Dempster-Shafer theory to assign belief mass to candidate association pairs and generate tracks based on the belief to plausibility ratio. The proposed algorithm is demonstrated using simulated observations of a GEO satellite breakup scenario.

I. Introduction

The goal of space situational awareness (SSA) is to characterize as fully as possible the space environment and better understand how this environment can and will change in the near future [1, 2]. The ability to track multiple space objects across a number of possibly spurious, noisy, or otherwise degraded observations is a primary area of concern for SSA. This problem is made more challenging when the objects of interest are closely grouped, for example when in formation or as the result of a debris generating incident. In either case, the observation of these objects can lead to ambiguities in how to instantiate a state estimate and how to associate future observations. The challenge is then to determine a robust way to autonomously initialize state estimates and perform association for any number objects observed over a given time period.

There are many existing approaches which attempt to address some aspect of this issue. Association can be done via multi-target tracking methods that operate directly in the image plane, largely derived from the field of computer vision and tracking. The global nearest neighbor and k-nearest neighbor approaches are two of the simplest implementations of multi-target tracking which do

*Graduate Researcher, School of Aerospace Engineering, and AIAA Student Member.
†Assistant Professor, School of Aerospace Engineering, and AIAA Senior Member.

not require models of the dynamics of the object in the image frame [4–6]. These methods can be improved upon given a model of the dynamics of the target in the frame and the implementation of a sequential filter, such as the EKF or UKF [7]. Recent advances in finite state statistics have lead to the development of robust multi-target tracking schemes in the field of computer vision [8]. These techniques are finding an increasing presence in SSA by demonstrating the capability to track low SNR objects through detectionless frame to frame tracking and the implementation of finite set statistics filtering [9–11]. An advantage of multiple target tracking methods implemented in the image plane is that they can be more computationally efficient due to the reduced dimensionality. However, because the dynamics of the problem are restricted to the image plane, in-plane implementations do not permit the direct computation of state estimates.

Association methods have also been proposed based on the admissible region approach. Attribution penalties can be computed for sampled points in the admissible region to identify potentially associated objects by setting a maximum allowed penalty [12, 13]. Fujimoto et. al. and Maruskin et. al. map a discritized admissible region to either Delunay or Poincaré element space and show that the intersection, if it exists, of two disparate admissible regions is the full state solution [14–16]. Fujimoto and Alfriend implement a boundary value problem approach to gate potential state hypotheses using the angle-rate information [17]. Optimization methods based on the admissible region have been proven to be computationally efficient ways to attempt the association problem for UCTs as well. Siminski et. al. define a loss function optimized over the admissible region to identify potential state solutions given two uncorrelated observations [18]. Worthy and Holzinger present an optimization based approach based on the admissible region which probabilistically determines if two observations are associated [19]. However this class of optimization methodologies alone do not enable one to discriminate fully between the observations between closely grouped objects. In these cases, optimization based method can give overly confident probabilities of association to more than one set of observations, proving problematic without a human-in-the-loop to help interpret the results.

The Constrained Admissible Region Multiple Hypothesis Filter (CAR-MHF) approach is often cited as one of the best methods by which to handle the association of a sequece of UCTs of multiple objects [20–22]. It is also founded in the admissible region method which uses hypothesized constraints to bound the continuum of potential solutions [23]. CAR-MHF takes advantage of the ability to sample this bounded set and instantiate a bank of filters on the resulting set of possible solutions. Each proposed solution is propagated and updated as new measurements are ingested. However, this method can be limited computationally by the discretization level of the admissible region. Furthermore, this method is still limited by the unobservability posed over short measurement periods.

Recently, Worthy and Holzinger proposed Dempster-Shafer Theory (DST) as a method by which unobservability is accounted for through the concept of ignorance [24]. DST augments probability theory by enabling one to account for imprecision in the assignment of probability to a given hypothesis. In DST, belief mass functions are defined which gather evidence supporting propositions under a given hypothesis. These mass functions then enable one to define both the plausibility and belief of a proposition under a given hypothesis rather than just the probability. The idea of imprecision is only necessary when there is ignorance in the system. As more observations are made, as ignorance is driven out of the system, the plausibility and belief values collapse to standard Bayesian probability. This makes DST a powerful tool in estimation. Worthy and Holzinger develop a sequential estimation tool which uses DST to define plausibility and belief surfaces over the admissible region. As expected, these surfaces collapse to a PDF once ignorance

is eliminated. The major drawback of this methodology is that it is very sensitive to the data ingested by the filter. Unlike with association based methods, the method described in Worthy and Holzinger is naive to probabilities of association. In fact, this method is demonstrated to be prone to failure when ingesting measurements from clustered spacecraft without a human-in-the-loop.

This paper proposes the marriage of DST and optimization based methods to determine probability of association to develop a robust multi-target tracking filter for SSA. The motivating idea is to enable the multi-target tracking algorithm to autonomously handle clustered or ambiguous detections of objects as well as unobservable observations. Section II recaps admissible region theory. Section III reviews the optimization based methodology and defines the notation used in this paper. Section IV reviews DST and the evidential particle filter. Section V introduces the evidential particle filter multi-target tracking scheme. Section VI demonstrates the strengths of the method for simulated data. Section VII discusses future directions to further improve the proposed methodology.

II. Admissible Region

The admissible region method bounds the continuum of possible states consistent with a given measurement through hypothesized constraints [23, 25]. Consider the general nonlinear measurement model

$$\mathbf{y} = \mathbf{h}(\mathbf{x}; \mathbf{k}, t) \tag{1}$$

where $\mathbf{x} \in \mathbb{R}^n$ is the state, $\mathbf{k} \in \mathbb{R}^z$ is a set of parameters, and $t \in \mathbb{R}$ is the time. Under the admissible region approach, the state vector can be partitioned into determined states $\mathbf{x}_d \in \mathbb{R}^d$ which directly affect the measurements and undetermined states $\mathbf{x}_u \in \mathbb{R}^u$ which do not affect the measurement [26]. Eqn. (1) then becomes

$$\mathbf{y} = \mathbf{h}(\mathbf{x}_d; \mathbf{k}, t) \tag{2}$$

implying there is a one-to-one mapping only between $\mathbf{x}_d$ and $\mathbf{y}$. By Eqn. (2), all values of $\mathbf{x}_u$ generate the same measurement, yielding a continuum of potential solutions. Let the i^{th} hypothesized constraint be defined as

$$\kappa_i(\mathbf{x}_u, \mathbf{y}, \mathbf{k}, t) \leq 0 \tag{3}$$

The admissible region corresponding to this constraint can then be defined as

$$\mathcal{AR}_i = \{\mathbf{x}_u \in \mathbb{R}^u \mid \kappa_i(\mathbf{x}_u, \mathbf{y}, \mathbf{k}, t) \leq 0\} \tag{4}$$

Eqn. (4) defines a set in $\mathbb{R}^u$ where each member of $\mathcal{AR}_i$ satiesfies the i^{th} constraint. When uncertainties are considered, it is useful to determine the probability of set membership for a given state in $\mathcal{AR}_i$. The probability of set membership is given by

$$P_i[(\mathbf{x}_u \in \mathcal{AR}_i)] \approx \frac{1}{2}\left[1 + \text{erf}\left(\frac{\|\mathbf{x}_\text{u} - \mathbf{x}_{\text{u},\mathcal{B}_\perp}\|}{\sqrt{2\text{tr}\mathbf{P}_{\mathbf{x}_{\text{u},\mathcal{B}_\perp}}}}\right)\right] \tag{5}$$

as defined in [26]. Let $\mathcal{AR} = \mathcal{AR}_i \cap \cdots \mathcal{AR}_c$ denote the joint admissible region for c hypothesized constraints. Then the joint probability of set membership is bounded by

$$P(\mathbf{x}_u \in \mathcal{AR}) = P(\mathbf{x}_u \in \cap_{i=1}^{c} \mathcal{AR}_i) \geq \left[\sum_{i}^{c} P_i(\mathbf{x}_u \in \mathcal{AR}_i)\right] - (c-1) \tag{6}$$

which is a lower bound that is derived from set theory [24].

III. Optimization Based Association

The admissible region approach presents a tool through which unobservable measurements can be used to attempt to extract information about the state of the object. One such use is the association of observations by attempting to find at least one point of intersection between a collection of admissible regions.

A. Admissible Region Intersections

For any admissible region, $\mathcal{AR}$, it is possible to construct a full n dimensional state through

$$\mathbf{x} = \mathbf{g}(\mathbf{x}_u, \mathbf{x}_d, \mathbf{k}) \tag{7}$$

The set of all such states formed from $\mathcal{AR}$ lie on a u-dimensional manifold in $\mathbb{R}^n$. Let this time varying manifold be defined as

$$\mathcal{X}^i(t) = \{\mathbf{x}(t) \ : \ \mathbf{x}_u(t_i) \in \mathcal{AR}\}$$

where i denotes this manifold belongs to the i^{th} admissible region. Given q such admissible regions, and thus q such manifolds $\mathcal{X}^1(t) \cdots \mathcal{X}^q(t)$, it is required for association that there exist at least one state satisfying

$$\mathbf{x}(t) \in \mathcal{X}^i(t), \quad i = 1, \cdots, q \tag{8}$$

For $q = 2$, it is necessary then that

$$\mathcal{X}^1(t) \cap \mathcal{X}^2(t) \neq \emptyset, \forall t$$

which is the foundation of intersection based association methods [14, 16, 18, 27]. Given some distance metric $d(\mathbf{a}, \mathbf{b}) \ : \ \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^+$, the general optimization problem solving this intersection finding method is given by

$$\begin{aligned} \underset{\mathbf{x}_{u,i}, \mathbf{x}_{u,j}}{\text{minimize}} \quad & d(\mathbf{x}^i(t), \mathbf{x}^j(t)) \\ \text{subject to} \quad & \mathbf{x}^i \in \mathcal{X}^i(t) \\ & \mathbf{x}^j \in \mathcal{X}^j(t) \end{aligned}$$

Through the simplifications outlined in [28], this optimization problem can be reduced in order to

$$\begin{aligned} \underset{\mathbf{x}^i_u \in \mathcal{AR}^i}{\text{minimize}} \quad & \sum_{i=1}^{N} \sum_{j=1, j\neq i}^{N} \tilde{d}(\mathbf{x}^i_d(t), \mathbf{x}^j_d(t)) \\ \text{subject to} \quad & \mathbf{x}^i_u \in \mathcal{X}^i_u(t_j) \end{aligned}$$

where $\tilde{d}(\mathbf{a}, \mathbf{b}) : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^+$ is a reduced order distance metric. This reduced order optimization problem attempts to find points of intersection between the determined subsets of $\mathcal{X}^i_d$ and $\mathcal{X}^j_d$ at both times t_i and t_j. Define $\mathcal{M}_i$ and $\mathcal{M}_j$ as the sets containing the local minima resulting from this reduced order optimization problem. If a point of intersection exists between all pairs $(i, j) \in [1, q], i \neq j$, or equivalently $\mathcal{M}_i \cap \mathcal{M}_j \neq \emptyset$, then all q observations are associated.

B. Hypothesis Testing for Association

When uncertainties are considered, $\mathbf{X}_d^i$ and $\mathbf{X}_d^j$ are the random variables representing the resulting distributions in $\mathbf{x}_d^i$ and $\mathbf{x}_d^j$. Assuming Gaussian uncertainties, these random variables cause $\mathcal{X}_d^i(t)$ and $\mathcal{X}_d^j(t)$ to also become probabilistic

$$\mathcal{X}_d^i(t_i) = \{\mathbf{X}_d^i(t_i) : \mathbf{X}_d^i(t_i) \sim \mathcal{N}(\mathbf{x}_d^i(t_i), \mathbf{\Sigma}_i), \mathbf{X}_{u,i}(t_i) \sim \mathcal{AR}^i\} \tag{9}$$

$$\mathcal{X}_d^j(t_j) = \{\mathbf{X}_{d,j}(t_i) : \mathbf{X}_{d,j}(t_j) \sim \mathcal{N}(\mathbf{x}_{d,j}(t_j), \mathbf{\Sigma}_j), \mathbf{X}_{u,j}(t_j) \sim \mathcal{AR}^j\} \tag{10}$$

where $\mathbf{x}_u$ is sampled from the i^{th} joint uncertain admissible region and $\mathbf{\Sigma}_i \in \mathbb{R}^{d\times d}$ is the covariance for the i^{th} observation.

The distribution about the single point $\mathbf{x}_{d,j} \in \mathcal{X}_d^j(t_j)$ becomes an ellipsoid in $\mathbb{R}^d$ centered at $\mathbf{x}_d^j$. The reduced orer distance metric also becomes a random variable defined as

$$D = \tilde{d}(\mathbf{X}_d^i(t), \mathbf{X}_d^j(t)) \tag{11}$$

If an intersection exists, then there should exist a state $\mathbf{x}_d^i(t)$ such that the deterministic version of Eqn. (11) equals zero. Let $l_i = \text{card}(\mathcal{M}_\text{i})$ and $l_j = \text{card}(\mathcal{M}_\text{j})$ and define

$$D_{i\to j}(\mathbf{x}_u(t_i)) = \tilde{d}(\mathbf{X}_d^i(t_j), \mathbf{X}_d^j(t_j)) \tag{12}$$

as the random variable of the distance metric mapping a state $\mathbf{x}_u \in \mathcal{M}_i$ from time t_i to time t_j where $\mathbf{x}_d(t_i) \sim \mathcal{N}(\mathbf{x}_d^i(t_i), \mathbf{\Sigma}_i)$. Likewise, $D_{j\to i}(\mathbf{x}_u(t_j))$ is the random variable of the distance metric mapping a state $\mathbf{x}_u \in \mathcal{M}_j$ from time t_j to time t_i. Let $f_0(s)$ be the distribution defined by imposing the intersection is at at a local minima

$$f_0(s) = \{D_0 \ : \ d_0 = \tilde{d}(\mathbf{x}_d^i(t), \mathbf{X}_d^j(t) + \mathbf{c}), \ \mathbf{X}_d^j(t) \sim \mathcal{X}_d^j(t)\} \tag{13}$$

where $\mathbf{x}_d^i$ is a fixed point in $\mathcal{X}_d^i$ corresponding to a local minima and $\mathbf{c}$ shifts $\mathcal{X}_d^j$ to be centered at $\mathbf{x}_d^i$ and $\mathbf{x}_u \in \mathcal{M}_i$. The true distribution of D is represented by

$$f_1(s) = \{D_0 \ : \ d_0 = \tilde{d}(\mathbf{x}_d^i(t), \mathbf{X}_d^j(t) + \mathbf{c}), \ \mathbf{X}_d^j(t) \sim \mathcal{X}_d^j(t)\} + d_{ij,\mathbf{x}_u(t_i)} \tag{14}$$

$$= f_0(s) + d_{ij,\mathbf{x}_u(t_i)} \tag{15}$$

where since $\mathbf{c}$ is a constant value, $d_{ij,\mathbf{x}_u(t_i)}$ shifts the location of $f_0(s)$. Thus, $d_{ij,\mathbf{x}_u(t_i)}$ is only zero if the local minima is indeed a point of intersection. These distributions help to define a binary hypothesis test to determine from which distribution D is drawn.

Since

$$\left[\sum_{i=1}^{q}\sum_{j=1,j\neq i}^{q} d_{ij,\mathbf{x}_u(t_i)}\right] = 0 \tag{16}$$

must be true for association, the following hypothesis test is used to determine if Eqn. (16) is satisfied.

$$\mathcal{H}_0 : D_{i\to j}(\mathbf{x}_u(t_j)) \sim f_0(s) + d_{ij,\mathbf{x}_u(t_i)} \qquad i,j = 1,\cdots,N; i \neq j \tag{17}$$

$$\mathcal{H}_1 : D_{i\to j}(\mathbf{x}_u(t_j)) \sim f_0(s) + 0 \qquad i,j = 1,\cdots,N; i \neq j \tag{18}$$

The alternative hypothesis $\mathcal{H}_1$ assumes that all q observations are associated, and the null hypothesis $\mathcal{H}_0$ assumes that there is no association between at least one pair of the q observations. These hypotheses can be tested of all $l_i \times l_j$ pairwise combinations of solutions in $\mathcal{M}_i$ and $\mathcal{M}_j$ independently.

The binary hypothesis test is constructed by determining the PDFs of the null and alternative hypotheses and computing the probabilities of false and true association [29]. The probability of false association P_{FA} is given by

$$P_{FA} = \int_r^\infty f_0(s)\mathrm{ds} \tag{19}$$

where r is selected specifically to obtain a desired P_{FA}. The probability of association is then given by

$$P_A = \int_r^\infty f_1(s)\mathrm{ds} \tag{20}$$

Because the hypothesis test is performed for each pairwise combination of solutions, the result from this set of hypothesis tests is a $l_i \times l_j$ set of values of P_A. The overall probability of uncertainty is found by taking the weighted norm as follows

$$\mathbb{P}\left[(\mathbf{x}_{u,i}(t_i), \mathbf{x}_{u,j}(t_j))\right] = \frac{P_{A,(ij)}}{\sum_{i=1}^{l_i}\sum_{j=1}^{l_j} P_{A,(i,j)}} \tag{21}$$

$$\mathcal{P}_A = \sum_{i=1}^{l_i}\sum_{j=1}^{l_j} \mathbb{P}\left[(\mathbf{x}_{u,i}(t_i), \mathbf{x}_{u,j}(t_j))\right] P_{A,(i,j)} \tag{22}$$

providing a single quantity indicating the likelihood that the measurements are associated.

A key feature of this approach is that through the properties of the maximum likelihood estimator, the covariance on the state estimate can be approximated. Let $\hat{\mathbf{x}}_u$ be a local minima found from after solving the optimization problem. By treating the distance metric as a log-likelihood,

$$\mathbb{E}\left[\sum_{i=1}^{N}\sum_{j=1,j\neq i}^{N} \frac{\partial}{\partial \hat{\mathbf{x}}_u}\left(-\tilde{d}(\mathbf{x}_{d,j}(t), \mathbf{x}_{d,i}(t))\right)\right] = \mathbf{0}_{u\times 1} \tag{23}$$

$$\mathbb{E}\left[\sum_{i=1}^{N}\sum_{j=1,j\neq i}^{N} \frac{\partial^2}{\partial \hat{\mathbf{x}}_u^2}\left(\tilde{d}(\mathbf{x}_{d,j}(t), \mathbf{x}_{d,i}(t))\right)\right] = \mathcal{I}(\hat{\mathbf{x}}_u)_{u\times u} \tag{24}$$

where Eqn. (23) and Eqn. (24) indicate that $\hat{\mathbf{x}}_u$ is a MLE and $\mathcal{I}(\hat{\mathbf{x}}_u)_{u\times u}$ is the Fisher information matrix [30]. Let $\mathbf{P}_u$ be the covariance matrix associated with $\hat{\mathbf{x}}_u$, by the Cramer-Rao bound

$$\mathbf{P}_u \geq \mathcal{I}(\hat{\mathbf{x}}_u)^{-1} \tag{25}$$

[31]. This bound is attained for MLEs, and thus a state estimate and covariance can be approximated directly for each local minima. However, this approximation only holds if the time scale is sufficiently large.

$$\mathrm{cond}(\mathcal{I}_\mathrm{k}(\cdot)) < (\mathrm{t_j} - \mathrm{t_i})^{-1}\sqrt{\frac{\|\mathbf{r}\|^3}{3\mu}}, \tag{26}$$

Eqn. 26 is a condition derived by relating the condition number of the Fisher information matrix to the time scale of the dynamics. If this condition is satisfied then the MLE approximation holds, otherwise an alternative method must be used to update the PDF [28].

In summary, association methods generally first prove association and then filters are used to ingest the associated measurements and generate state estimates and covariances. This optimization based approach merges the two steps into one when the problem is observable by assuming approximately Gaussian behavior in this region of the local optimum. However, these assumptions break down over short time periods and the underlying PDF of the admissible region cannot be readily determined. The next section describes how DST can be used to permit the natural evolution from an admissible region uninformative prior to a fully defined PDF.

IV. Dempster Shafer Evidential Filter

Traditional Bayesian probability is limited in that evidence, i.e. a measurement or observation, can only support or refute a hypothesis. In many real world applications there often exists states for which evidence neither supports nor refutes a given hypothesis. DST adds the concepts of plausibility and ignorance to address this particular situation.

A. Dempster Shafer Preliminaries

A full treatment of DST can be found in [32–35], however the primary constructs of DST are briefly introduced as follows. Define the frame of discernment, Ω, as the set which contains all the elements to which belief mass can be assigned. A mass function $m : 2^{\Omega} \rightarrow [0, 1]$ is then defined to allocate belief mass to members of the power set of Ω, that is the set of all possible combinations of the members of Ω. For $A \subset \Omega$, a belief mass function must satisfy

$$\sum_{A \subseteq \Omega} m(A) = 1 \tag{27}$$

From the mass function follows the computation of two useful quantities,

$$\forall A \subseteq \Omega, \mathrm{Bel(A)} = \sum_{\emptyset \neq B \subseteq A} m(B) \tag{28}$$

$$\forall A \subseteq \Omega, \mathrm{Pl(A)} = \sum_{A \cap B \neq \emptyset} m(B) \tag{29}$$

The belief function, $\mathrm{Bel}_i(A)$, gathers evidence to support a given state A. The plausibility function, $\mathrm{Pl}_i(A)$, gathers evidence which simply permits the occurrence of A. Belief and plausibility are related through duality

$$\forall A, \mathrm{Pl(A)} + \mathrm{Bel}(\bar{\mathrm{A}}) = 1 - m(\emptyset) \tag{30}$$

where $\bar{A}$ is the complement of A. If $m_i(\emptyset) = 0$ then the solution must exist in Ω. The definition of either the mass function, belief function, or plausibility function directly enables the other two to be determined. The Ignorance function can quantified through the relationship

$$\mathrm{Ig(A)} = \mathrm{Pl(A)} - \mathrm{Bel(A)} \tag{31}$$

since it is shown that plausibility and belief are upper and lower bounds on probability [36].

B. Dempster Shafer Admissible Region

The DST approach can be applied to the admissible region method by defining an appropriate frame of discernment and mass function. Let,

$$\mathcal{X} = \{(\mathbf{x}_u, \mathbf{x}_d) \ : \ \mathbf{x}_u \in \mathcal{R}\} \tag{32}$$

$$\bar{\mathcal{X}} = \{(\mathbf{x}_u, \mathbf{x}_d) \ : \ \mathbf{x}_u \notin \mathcal{R}\} \tag{33}$$

the set of full n dimensional states satisfying the hypothesized constraints and the set of all inadmissible states respectively. The frame of discernment is then defined as

$$\Omega = \{\mathbf{x} \in \mathcal{X}, \bar{\mathcal{X}}\} \tag{34}$$

Due to the properties of the power set applied to real numbers, it is computationally useful to also define

$$\Theta = \{\mathbf{x} \in \mathcal{X}, \mathcal{X}, \bar{\mathcal{X}}\} \tag{35}$$

$\Theta \subset 2^{\Omega}$ is a special subset of the full power set of Ω that, by construction, must still contain the solution but makes the problem more computationally tractable [34, 37, 38].

The mass function for the admissible region problem is by definition vacuous since given an unobservable sequence of measurements, there is no evidence that either refutes nor supports a particular state [24]. Instead, it is useful to construct a plausibility function and derive a relationship between plausibility and the mass function directly. A pseudo-plausibility function for a singleton hypothesis $\mathbf{x} \in \Theta$ is defined as

$$\widetilde{\mathrm{Pl}}(\mathbf{x}|\mathbf{y}) = P(\mathbf{x}_u \in \cap_{i=1}^{c} \mathcal{R}_i) p(\mathbf{x}_d|\mathbf{y}) \tag{36}$$

Eqn. (36) is not a plausibility function since the plausibility of any state in $\mathcal{X}$ must be lower bounded by the belief of $\mathcal{X}$ itself. The true plausibility function is then given by

$$\mathrm{Pl}(\mathbf{x}|\mathbf{y}) = \widetilde{\mathrm{Pl}}(\mathbf{x}|\mathbf{y}) + m(\mathcal{X}|\mathbf{y}) \tag{37}$$

The mass and plausibility functions for $\mathcal{X}$ and $\tilde{\mathcal{X}}$ are defined as follows

$$\mathrm{m}(\mathcal{X}|\mathbf{y}) = \frac{\int_{\mathcal{X}} \widetilde{\mathrm{Pl}}(\mathbf{x}|\mathbf{y}) \mathrm{d}\mathbf{x}}{\int_{\mathcal{X}} \mathrm{d}\mathbf{x}} \tag{38}$$

$$\mathrm{Pl}(\bar{\mathcal{X}}|\mathbf{y}) = \mathrm{m}(\bar{\mathcal{X}}|\mathbf{y}) = 1 - \max_{\mathbf{x} \in \mathcal{X}} \mathrm{Pl}(\mathbf{x}|\mathbf{y}) \tag{39}$$

The relationship between the plausibility and mass functions is then defined as

$$\begin{bmatrix} \mathrm{Pl}(\mathbf{x}_1|\mathbf{y}_{0:\mathrm{k}-1}) \\ \mathrm{Pl}(\mathbf{x}_2|\mathbf{y}_{0:\mathrm{k}-1}) \\ \vdots \\ \mathrm{m}(\mathcal{X}|\mathbf{y}_{0:\mathrm{k}-1}) \\ \mathrm{m}(\bar{\mathcal{X}}|\mathbf{y}_{0:\mathrm{k}-1}) \end{bmatrix} = \frac{1}{\zeta} \begin{bmatrix} 1 & 0 & \cdots & 1 & 0 \\ 0 & 1 & \cdots & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix} m \begin{pmatrix} \mathbf{x}_1|\mathbf{y}_{0:k-1} \\ \mathbf{x}_2|\mathbf{y}_{0:k-1} \\ \vdots \\ \mathcal{X}|\mathbf{y}_{0:k-1} \\ \bar{\mathcal{X}}|\mathbf{y}_{0:k-1} \end{pmatrix} \tag{40}$$

where ζ is a normalizing factor necessary due to considering only Θ instead of the full power set of Ω.

C. Sequential Belief Filter

Given this treatment of DST for the admissible region problem, the particle filter implementation is described as follows. The plausibility function is initialized with the admissible region probability of set membership and us updated as the product of sequential plausibility values.

$$\mathrm{Pl}_0(\mathbf{x}|\mathbf{y}_0) = P(\mathbf{x}_u \in \cap_{i=1}^{c} \mathcal{R}_i|\mathbf{y}_0)p(\mathbf{x}_d|\mathbf{y}_0) \tag{41}$$

$$\mathrm{Pl}_\mathrm{k}(\mathbf{x}|\mathbf{y}_{0:\mathrm{k}}) = \mathrm{Pl}_\mathrm{k}(\mathbf{x}|\mathbf{y}_\mathrm{k})\mathrm{Pl}_{\mathrm{k}-1}(\mathbf{x}|\mathbf{y}_{0:\mathrm{k}-1}) \tag{42}$$

The mass function is vacuous at initialization

$$m_0(\mathbf{x}|\mathbf{y}_0) = 0 \tag{43}$$

$$m_0(\mathcal{X}|\mathbf{y}_0) = 1 \tag{44}$$

$$m_0(\bar{\mathcal{X}}|\mathbf{y}_0) = 0 \tag{45}$$

and the mass function is updated for the singleton states through Eqn. (40)

$$m_k(\mathbf{x}|\mathbf{y}_{0:k}) = \zeta \mathbf{A}^{-1}\mathrm{Pl}_\mathrm{k}(\mathbf{x}|\mathbf{y}_{0:\mathrm{k}}) \tag{46}$$

whereas Eqns. (38) and (39) are used to update the mass functions for the non-singleton sets $\mathcal{X}$ and $\bar{\mathcal{X}}$. By construction of the problem, the belief function for the singleton states is identically equal to the mass function for those states. Furthermore, as more evidence is gained and the system becomes observable, the belief function and the plausibility functions become equal for at most one of these singleton states and the belief of the non-singleton sets go to zero under the appropriate constraint hypotheses. At this point, a well defined PDF exists for a singleton state in Θ and this state and PDF can be used in traditional estimation schemes. Another novelty of this introduced sequential estimation scheme lies in the ability to test the constraint hypotheses. If measurements are ingested under an invalid set of constraint hypotheses, belief mass is attributed to $\bar{\mathcal{X}}$, the set of all inadmissible states, indicating that there is evidence the truth solution does not lie in the original admissible region as posed [24].

However, the drawback of this methodology is that if an incorrect measurement is ingested, or a sufficiently noisy measurement is ingested, all belief mass can be incorrectly attributed to $\bar{\mathcal{X}}$. The goal of the next section is to introduce a more robust way by which evidential particle filters tracking the state of space objects can avoid this problem by implementing the association methodology from Section III.

V. Multi-target Tracking

DST is well suited to address the limitations posed by both methods discussed in the previous two sections. The ambiguities produced by the uncorrelated tracks of closely grouped space objects fall directly into the concept of ignorance. Until sufficient evidence has been gathered, it is possible that several combinations of measurements from the objects observed are true. Essentially, the multi-target tracking approach proposed uses DST to handle ambiguities in tracking over short time periods and the optimization based approach to continue tracks and produce state estimates.

A. Frame of Discernment for MTT

As stated in the previous section, the primary constructs of DST are the frame of discernment and the belief assignment function. Consider again the measurement model from Eqn. (1) and let $\mathbf{y}_k^i$

denote the i^{th} measurement obtained at time k. Let the frame of discernment at each time then be defined as

$$\Omega(k) \subseteq \left\{ \bigcup_{j=1}^{k} \bigcup_{i=1}^{q_k} \mathbf{y}_j^i \right\} \tag{47}$$

where q_k is the number of measurements gathered at time k. $\Omega(k)$ is by construction equal to, or a subset of, the set of all measurements gathered at all times. The power set of $\Omega(k)$ is then the set of all possible combinations of these measurements, giving a fully exhaustive set over which to search for potential associations and from which to form tracks.

B. Mass function for MTT

The belief assignment function that readily seems appropriate for $\Omega(k)$ is the probability of association defined in Eqn. (20). Let $T \subseteq \Omega(k)$ and $\mathbf{y}_j^i \in T$ and define

$$\tilde{m}(T) = \begin{cases} \mathcal{P}_A(\mathbf{y}_j^i, \emptyset) = 0 & \text{card(T)} = 1 \\ \mathcal{P}_A(\mathbf{y}_v^1, \mathbf{y}_w^2) & \text{card(T)} = 2, \text{v} \neq \text{w} \\ \prod_{i=1}^{\text{card(T)}} \prod_{j \neq i}^{\text{card(T)}} \mathcal{P}_A(\mathbf{y}_v^i, \mathbf{y}_w^j) & \text{card(T)} > 2, \text{v} \neq \text{w} \\ 0 & \text{otherwise} \end{cases}$$

as the association function. Then

$$m(T) = \xi \tilde{m}(T), \; T \subseteq \Omega(k) \tag{48}$$

is a candidate mass function for the defined frame of discernment, where ξ is a normalization factor to ensure Eqn. (27) is satisfied. Unlike the mass function for the admissible region evidential particle filter, this mass function can be directly defined and is not vacuous. There is no need to attempt to first find a plausibility function from which to map into belief mass. Another difference is the fact that for any non-singleton $T \subseteq \Omega(k)$, the belief mass does not necessarily go to zero. Thus, the belief and plausibility functions can be explicitly defined from Eqn. (48) by Eqns. (28) and (29). Finally, for any set of measurements $T \subset \Omega(k)$,

$$\text{Ig(T)} = \text{Pl(T)} - \text{Bel(T)} \tag{49}$$

When $\text{Ig}(T) = 0$, sufficient evidence has been gathered to confirm the all measurements in the set T are associated. In a broader sense, $\text{Ig}(T) = 0$ implies through Eqns. (28) and (29) that the mass function is only non-zero for T and subsets of T. This signifies that the measurements comprising T cannot be associated with any other measurements from $\Omega(k)$. A more useful quantity to track is given by

$$\mathcal{I}(T) = \text{Bel(T)}/\text{Pl(T)} \tag{50}$$

, where $\text{Ig}(T) = 0 \rightarrow \mathcal{I}(T) = 1$, which essentially gives the confidence in track T.

C. Updating the Frame of Discernment

The motivation of this framework is to provide a robust way to perform association and state estimation. As measurement associations are performed, it follows that $\Omega(k)$ should be modified appropriately to reflect positive associations. That is, once association has been confirmed for a specific $T \subseteq \Omega(k)$, there is no need to continue to retest the measurements constituting T. Let $\bar{T} = \{B \ : \ B \subset T\}$, if $\text{Bel}(T) = \text{Pl}(T)$ then

$$\Omega(k+1) = \left\{\Omega(k) \setminus \bar{T}\right\} \cup \left\{\bigcup_{j=1}^{q_{k+1}} \mathbf{y}_{k+1}^{j}\right\} \tag{51}$$

Equation (51) ensures measurements that have been positively associated are combined into a single track in $\Omega(k)$ while the individual measurements are removed from $\Omega(k)$. This reduces the cardinality of $\Omega(k)$ thus improving computational performance by eliminating the need to retest already associated tracks.

It is then also possible to estimate the number of tracked objects, N, by computing

$$\hat{N} = \sum_{A \subseteq \Omega} \mathcal{I}(A) = \sum_{A \subseteq \Omega} \frac{\text{Bel(A)}}{\text{Pl(A)}}$$

which is the sum of the ratio of belief to plausibility, or the ignorance ratio, of A. By construction $\text{Bel(A)}/\text{Pl(A)} = 1$ only for tracks for which there is sufficient evidence for association. This implies that in general $\hat{N} \leq N$, but if there is any ambiguity about any tracks in $\Omega(k)$ then $\hat{N} < N$.

D. Hybrid Filter for MTT

The previous two sections outline the construction of DST applied to the association problem by defining a frame of discernment containing all possible sets of measurements. Since it is ultimately desired to obtain a state estimate from each track T, the state estimates and probability distributions are updated through a hybrid application of the evidential particle filter introduced in Section IV and the MLE approximation introduced in III. The computation of $\tilde{m}(T)$ requires the solution to the reduced order optimization problem, and for each local minima Eqn. (26) is also computed. If the observability condition is met, then the state estimate and corresponding PDF are produced directly from the solution to the optimization problem. Otherwise, the evidential filter is used to update the bounds on the PDF by computing the updated belief and plausibility surfaces as outlined in Section IV.C.

E. Computational Performance

The full exploration of $\Omega(k)$ requires checking all $2^{\Omega(k)}$ possible combinations of observations for general application of DST. This problem is similar to that posed by MHT methods when ambiguity causes the number of hypotheses to grow exponentially. In terms of computational tractability, it is desired to remove as many members from the full power set as possible without violating the principles of DST. The primary method by which this is achieved is by not considering any observations from the same time for association. Further, the frame of discernment update outlined by Eqn. (51) deals with removing subsets for which the knowledge of superset to which they belong is certain. However, the size of the power set will grow exponentially with the number of ambiguous track associations with the current implementation.

The general computational complexity of this algorithm with the stated modifications is $O(n^2)$ since the pairwise probability of association must be computed for each valid pair of observations. Thus, the computational limiting factor is the optimization based association problem. The run time of the currently implemented DSMTT algorithm, t_{cpu}, can be predicted as follows

$$t_{\text{cpu}} = \omega \times q_{k-1} \times q_k \times t_{\text{opt}} \tag{52}$$

where t_{opt} is the run time of the optimization method, ω is the number of existing measurements in 2^{Ω}, q_{k-1} is number of measurements in the longest associated track, and q_k is the number of new measurements being ingested.

While it may be enticing to eliminate or sub-sample certain ambiguous tracks from $\Omega(k)$, removing propositions which have non-zero belief mass violates DST. Note that none of the modifications to $\Omega(k)$ described remove propositions which have belief mass. Ultimately, the strength of DST is that with enough observations the algorithm can resolve these ambiguities. Improperly deleting items from $\Omega(k)$ can prevent this resolution or even lead indirectly to incorrect resolution of ambiguity.

F. The DSMTT Algorithm

The preceding subsections lay the foundation for the DSMTT framework. This subsection discusses the actual implementation of the algorithm. At a high level, the algorithm ingests measurements from observers at time k and constructs $\Omega(k)$. Note that the observers need not be collocated, these methodologies make no assumptions on the observer parameters. Then belief mass is assigned to every possible combination of measurements in $\Omega(k)$ using Eqn. (48), which is based upon computing the probability of association described in Eqn. (20). The observability of each combination of measurements is tested as a result of computing the probability of association. Lastly, the state estimates of fully observable combinations are updated through the MLE approximation while the unobservable combinations are updated with the evidential particle filter. Algorithm 1 provides more detail into the MTT framework. Note that in implementation, it is more convenient maintain and update the power set, 2^{Ω}, as opposed to Ω.

VI. Simulation and Results

To demonstrate the performance of the proposed methodology, three simulated scenarios are presented. The observations follow the breakup of a GEO satellite generating 50 pieces of debris ($N = 51$) normally distributed with initial changes in velocity, ΔV, of 3 m/s. The dynamics model chosen is simple two body Keplerian motion, however note that the algorithms described are agnostic to the dynamics model used. The observer is assumed to be located in Atlanta, GA (33.755°N, 84.39°W, 300m) with a field of view of 10.6754×8.5528 arcminutes. The instantaneous field of view of the sensor is .2341 arcseconds, any measurement with angular separations below this distance are merged into a single measurement. Angular measurement uncertainty is chosen to be 10 arcseconds.

In the first simulated scenario, the observation system is tasked on the GEO satellite starting 60 minutes post breakup and continues observing for 10 minutes at 60 second intervals generating 500 total measurements of debris. All of the debris objects are contained within the FOV of the sensor and the purpose of this scenario is to demonstrate the in frame multi-target tracking capabilities of this approach.

Algorithm 1 Dempter Shafer Multi Target Tracker

procedure DEMPSTERSHAFERTRACKING($\left\{\bigcup_{i}^{q_k} \mathbf{y}_k^{q_k}\right\}, 2^{\Omega_{k-1}}$)
  $2^{\Omega_k} = \emptyset$ ▷ Initialize the new power set
  $\omega = \text{card}(2^{\Omega_{k-1}})$ ▷ Current size of power set
  for $i \leftarrow 1, q_k$ **do** ▷ Loop over new observations
    for $j \leftarrow 1, \omega$ **do** ▷ Loop over power set
      $T_j \in 2^{\Omega_{k-1}}$ ▷ Select a member from $2^{\Omega_{k-1}}$
      $T_{ij} = T_j \cup \mathbf{y}_k^i$ ▷ Form new track
      $m_{ij} = m(T_{ij})$ ▷ Compute belief mass
      $\text{Pl}(\mathcal{X}|\text{T}_{\text{ij}}), \text{Bel}(\mathcal{X}|\text{T}_{\text{ij}}) = \text{evidentialParticleFilter}(\text{T}_{\text{ij}}, \text{Pl}(\mathcal{X}|\text{T}_{\text{j}}), \text{Bel}(\mathcal{X}|\text{T}_{\text{j}})$ ▷ Update state estimate distributions
      $2^{\Omega_k} = 2^{\Omega_k} \cup T_{ij}$ ▷ Append new track to 2^{Ω_k}
  $\omega^+ = \text{card}(2^{\Omega_k})$ ▷ Updated size of power set
  for $i \leftarrow 1, \omega^+$ **do**
    for $j \leftarrow 1, \omega^+$ **do**
      $T_i, T_j \in 2^{\Omega_k}$ ▷ Select pairs from 2^{Ω_k}
      $\text{Pl}(T_i) = \text{Bel}(T_i) = 0$
      if $T_i \cap T_j \neq \emptyset$ **then**
        $\text{Pl}(T_i)$ += m_j ▷ Eqn. (29)
      if $T_j \subseteq T_i \neq \emptyset$ **then**
        $\text{Bel}(T_i)$ += m_j ▷ Eqn. (28)
  $2^{\Omega_k} = \text{removeZeroPlausibilityTracks}(2^{\Omega_k})$
  $2^{\Omega_k} = \text{pruneZeroIgnoranceTracks}(2^{\Omega_k})$ ▷ Eqn. (51)
  return 2^{Ω_k}

Figure 1 shows the results from processing the observations sequentially using Algorithm 1. The black dots are the center points of the observations, the black lines represent tracks formed with $\text{Ig}(T) = 0$, and the black stars represent the start and end computed for track T. The red lines indicate tracks for which $\text{Ig}(T) \neq 0$, or equivalently ambiguous tracks. As can be seen, there are three pairs of objects whose observations are very closely spaced, producing these ambiguous cases. As a result, Figure 2 shows how the estimated number of objects being tracked gets smaller over time. This is a direct result of these observations spawning spurious tracks in 2^{Ω} since they cannot be fully resolved. After processing, $\text{card}(2^{\Omega}) = 983$ and since $\hat{N} = 48.2$, the remaining 941 members of 2^{Ω} result from ambiguity in closely spaced, non-resolvable associations. This exponential increase in the size of the power set causes compute time to rise significantly and a topic for future work is how to better handle management of the powerset members without violating DST.

In the second scenario, the observation system is tasked by sweeping through the range of right ascensions predicted for the debris cloud location 9 hours post breakup over a 60 minute period. It is assumed for illustration purposes that the observations from the first scenario are unavailable. The debris cloud is much larger than the field of view and observations are taken every 60 seconds giving 1653 total measurements of the debris. Of these 1653 measurements, only 44 out of the 50 total objects have been observed and each observed debris object has between 1 and 34 measurements. The purpose of this scenario is to demonstrate the multi-target tracking

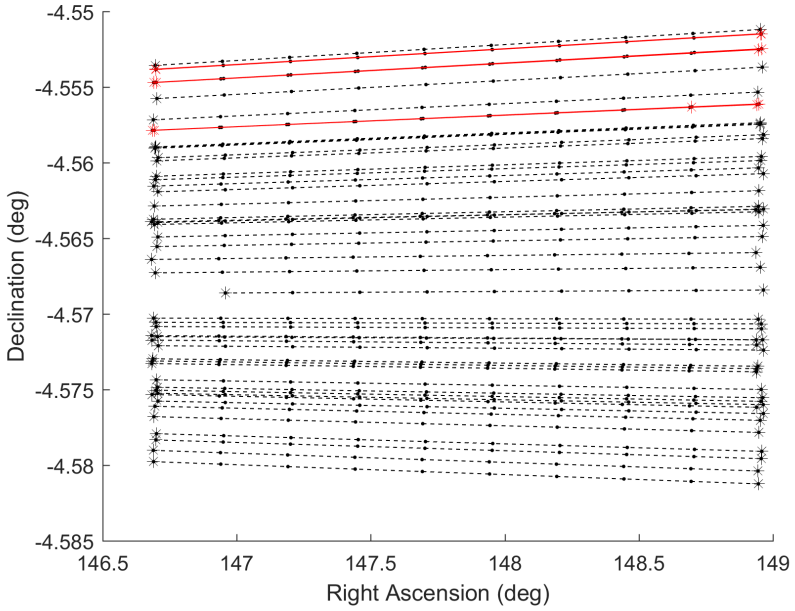

Figure 1. Results of DSTMTT Algorithm

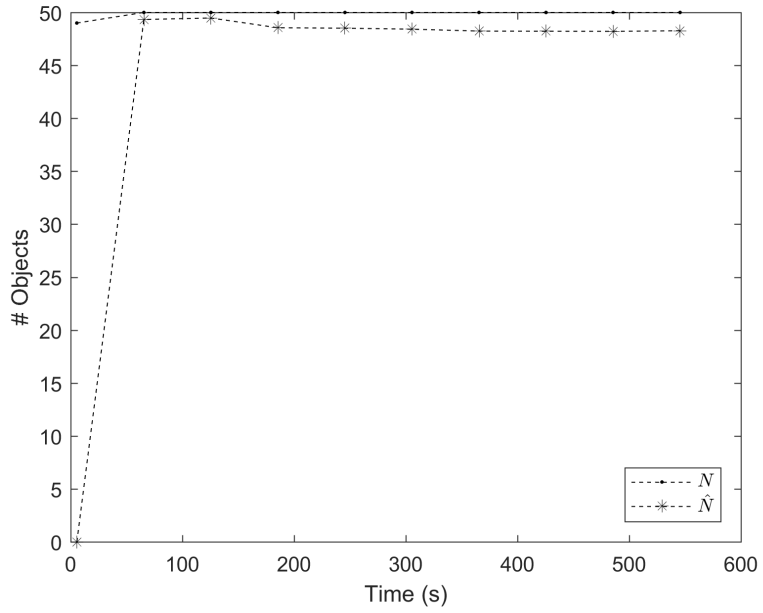

Figure 2. Difference between true and estimated number of tracked objects

capabilities of this algorithm when the targets do not all reside in the frame of the image and when targets drift in and out of the field of view over time. This is more closely related to the problem CAR-MHF attempts to solve. Figure 3 shows the tracks formed as a result of the algorithm. Note the x axis spans nearly 15° since the field of view of the observer is spanning through the debris cloud. Figure 4 shows the difference between how many objects the algorithm is tracking compared with how many objects have truly passed through the field of view. Table 1 shows the association performance for this scenario. The interesting columns to examine are the ignorance ratio and the number of unique labels. The unique labels indicates how many of the true objects are included in this track. Only tracks with unique labels = 1 can be correctly identified tracks, and it is interesting to look at the tracks with unique labels = 1 but ignorance ratio < 1. This indicates tracks for which there is association of at least one measurement which associates with another measurement not in the track. In these cases, the ignorance ratio is maximized when the measurements actually only come from a single object. As a point of future work, it would be interesting to use the ignorance ratio as a means to sample from the ambiguous tracks, keeping only those with high confidence, rather than keeping all ambiguous tracks in the power set.

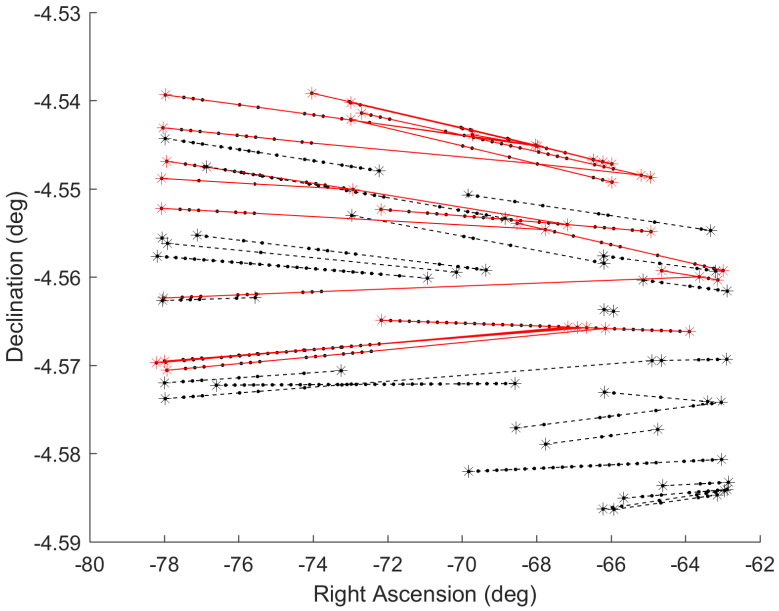

Figure 3. Results of DSTMTT Algorithm

To briefly discuss computational performance, the run times for Scenario 1 and Scenario 2 were 89202.9 seconds and 12652.4 seconds respectively. Computationally scenario 2 is preferable since, while there are more total observations, on average the number of observations being ingested at a given time is substantially smaller. The simulations were performed on a Intel Core i7-6700k CPU with 32GB memory in matlab. The average run time of the observation to observation association method was 1.21 seconds which reinforces the order of magnitude estimation of the run time given in Eqn (52). While the CPU time performance is not ideal, there are many improvements that can be made to this algorithm, for instance, by implementation in a compiled language as opposed to a scripting language.

ID	# Associated	Ignorance Ratio	Unique Labels	ID	# Associated	Ignorance Ratio	Unique Labels
10001	26	1.000000	1	10032	10	0.066206	3
10002	9	1.000000	1	10033	8	0.072031	3
10003	8	1.000000	1	10034	8	0.070169	3
10004	5	1.000000	1	10035	8	0.064325	3
10005	5	1.000000	1	10036	8	0.355032	1
10006	12	1.000000	1	10037	16	0.019664	2
10007	10	1.000000	1	10038	6	1.000000	1
10008	8	1.000000	1	10039	2	1.000000	1
10009	7	1.000000	1	10040	10	0.030980	2
10010	5	1.000000	1	10041	10	0.028430	3
10011	4	0.992179	1	10042	10	0.026622	3
10012	9	1.000000	1	10043	10	0.026113	3
10013	4	1.000000	1	10044	7	0.022761	3
10014	34	0.999999	1	10045	13	0.133369	3
10015	11	0.251810	2	10046	13	0.091231	3
10016	23	1.000000	1	10047	12	0.084184	3
10017	21	0.861727	1	10048	24	0.033964	2
10018	4	1.000000	1	10049	12	0.178488	2
10019	10	1.000000	2	10050	9	0.048686	2
10020	13	0.373452	3	10051	13	0.019786	2
10021	11	0.058676	1	10052	2	0.020190	2
10022	11	0.065351	2	10053	13	0.043982	2
10023	11	0.060098	2	10054	32	1.000000	1
10024	11	0.067875	2	10055	22	1.000000	1
10025	9	0.064370	2	10056	14	1.000000	1
10026	9	0.059507	2	10057	3	1.000000	1
10027	9	0.064184	2	10058	27	1.000000	1
10028	9	0.064735	2	10059	18	1.000000	1
10029	10	0.069590	2	10060	4	0.319874	2
10030	10	0.068980	3	10061	8	1.000000	1
10031	10	0.072673	3	10062	6	1.000000	1
10063	1	0	3				

Table 1. Scenario 2 Association Performance

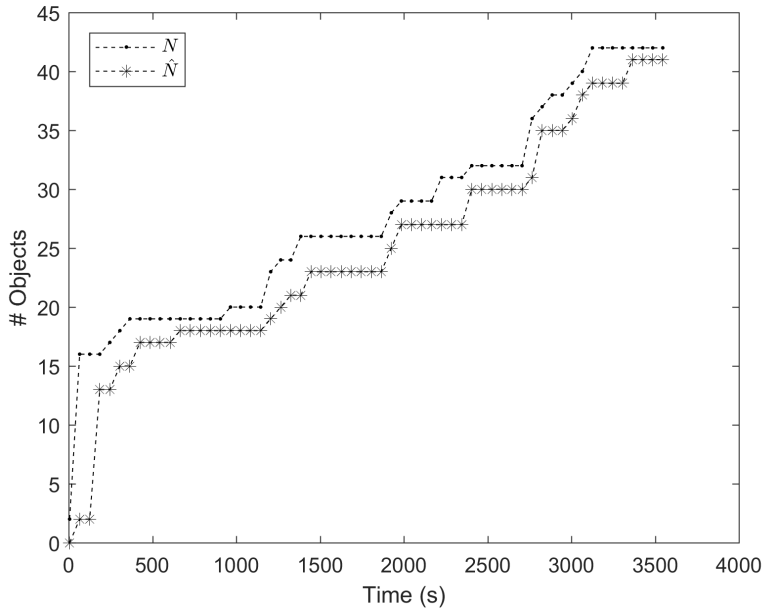

Figure 4. Difference between true and estimated number of tracked objects

VII. Conclusions and Future Work

To conclude, a Dempster-Shafer framework is presented combining an optimization based association metric with an evidential particle filter for updating state estimate distributions. This framework is based on collecting evidence for track formation and following DST only forming tracks when ignorance is driven from the system. A simulated scenario following the breakup of a GEO satellite demonstrates the utility of the algorithm to reconcile object tracks and update state estimates given limited observation availability and short time between observations.

A primary topic of future work is how to improve the computational performance of the algorithm. In particular, how to prevent ambiguities from growing the size of the power set unnecessarily. Another area of future work is to test the algorithm with more complicated dynamics models and with empirical observation data.

VIII. Acknowledgements

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1148903. Shoutout to MITLL for sending me to AMOS.

References

[1] "Joint Publication 3-14, Space Operations," US Joint Chiefs of Staff, Washington, D.C., 2013.

[2] C. D. Wickens, "Situation awareness: Review of Mica Endsley's 1995 articles on situation awareness theory and measurement," *Human factors*, Vol. 50, No. 3, 2008, pp. 397–403.

[3] J. L. Worthy and M. J. Holzinger, "Use of Uninformative Priors to Initialize State Estimation for Dynamical Systems," *Advances in Space Research*, 2017, http://dx.doi.org/10.1016/j.asr.2017.06.040.

[4] P. Konstantinova, A. Udvarev, and T. Semerdjiev, "A study of a target tracking algorithm using global nearest neighbor approach," *Proceedings of the International Conference on Computer Systems and Technologies (CompSysTech03)*, 2003, pp. 290–5.

[5] P. M. Vaidya, "AnO(n logn) algorithm for the all-nearest-neighbors Problem," *Discrete & Computational Geometry*, Vol. 4, Mar 1989, pp. 101–115, 10.1007/BF02187718.

[6] D. Kostishack, B. Burke, and G. Mayer, "Continuous-scan charge-coupled device (CCD) sensor system with moving target indicator (MTI) for satellite surveillance," *QUANTUM*, Vol. 10, 1980, p. 50.

[7] A. Amditis, G. Thomaidis, P. Maroudis, P. Lytrivis, and G. Karaseitanidis, "Multiple hypothesis tracking implementation," *Laser Scanner Technology*, InTech, 2012.

[8] R. Hoseinnezhad, B.-N. Vo, B.-T. Vo, and D. Suter, "Visual tracking of numerous targets via multi-Bernoulli filtering of image data," *Pattern Recognition*, Vol. 45, No. 10, 2012, pp. 3625–3635.

[9] I. I. Hussein, K. J. DeMars, C. Früh, R. S. Erwin, and M. K. Jah, "An AEGIS-FISST integrated detection and tracking approach to Space Situational Awareness," *Information Fusion (FUSION), 2012 15th International Conference on*, IEEE, 2012, pp. 2065–2072.

[10] K. Fujimoto, M. Uetsuhara, and T. Yanagisawa, "Statistical track-before-detect methods applied to faint optical observations of resident space objects," *Advanced Maui Optical and Space Surveillance Technical Conference*, 2015.

[11] T. S. Murphy, M. J. Holzinger, and B. Flewelling, "Visual Tracking Methods for Improved Sequential Image-Based Object Detection," *Journal of Guidance, Control, and Dynamics*, 2017.

[12] A. Milani and G. Gronchi, *Theory of orbit determination*. Cambridge University Press, 2010.

[13] A. Milani, G. Tommei, D. Farnocchia, A. Rossi, T. Schildknecht, and R. Jehn, "Correlation and orbit determination of space objects based on sparse optical data," *Monthly Notices of the Royal Astronomical Society*, Vol. 417, No. 3, 2011, pp. 2094–2103, 10.1111/j.1365-2966.2011.19392.x.

[14] J. M. Maruskin, D. J. Scheeres, and K. T. Alfriend, "Correlation of Optical Observations of Objects in Earth Orbit," *Journal of Guidance, Control, and Dynamics*, Vol. 32, No. 1, 2009, pp. 194–209, doi:10.2514/1.36398.

[15] K. Fujimoto, J. Maruskin, and D. Scheeres, "Circular and zero-inclination solutions for optical observations of Earth-orbiting objects," *Celestial Mechanics and Dynamical Astronomy*, Vol. 106, No. 2, 2010, pp. 157–182, doi:10.1007/s10569-009-9245-y.

[16] K. Fujimoto and D. Scheeres, "Correlation of optical observations of earth-orbiting objects and initial orbit determination," *Journal of guidance, control, and dynamics*, Vol. 35, No. 1, 2012, pp. 208–221.

[17] K. Fujimoto and K. T. Alfriend, "Optical Short-Arc Association Hypothesis Gating via Angle-Rate Information," *Journal of Guidance, Control, and Dynamics*, 2015/06/08 2015, pp. 1–12, 10.2514/1.G000927.

[18] J. Siminski, O. Montenbruck, H. Fiedler, and T. Schildknecht, "Short-arc tracklet association for geostationary objects," *Advances in Space Research*, Vol. 53, No. 8, 2014, pp. 1184 – 1194, doi:10.1016/j.asr.2014.01.017.

[19] J. L. Worthy III and M. J. Holzinger, "AN OPTIMIZATION BASED APPROACH TO CORRELATION OF OBSERVATIONS WITH UNCERTAINTY," *AAS/AIAA Spaceflight Mechanics Meeting, Napa, CA*, 2016.

[20] K. DeMars, M. Jah, and P. Schumacher, "Initial Orbit Determination using Short-Arc Angle and Angle Rate Data," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 48, No. 3, 2012, pp. 2628–2637, doi:10.1109/TAES.2012.6237613.

[21] K. DeMars and M. K. Jah, "Probabilistic Initial Orbit Determination Using Gaussian Mixture Models," *Journal of Guidance Control, and Dynamics*, Vol. 36, No. 5, 2013, pp. 1324–1335, doi:10.2514/1.59844.

[22] J. Stauch, M. Jah, J. Baldwin, T. Kelecy, and K. A. Hill, ch. Mutual Application of Joint Probabilistic Data Association, Filtering, and Smoothing Techniques for Robust Multiple Space Object Tracking (Invited). AIAA SPACE Forum, American Institute of Aeronautics and Astronautics, Aug 2014. 0, 10.2514/6.2014-4365.

[23] A. Milani, G. F. Gronchi, M. d. Vitturi, and Z. Knežević, "Orbit determination with very short arcs. I admissible regions," *Celestial Mechanics and Dynamical Astronomy*, Vol. 90, No. 1-2, 2004, pp. 57–85, doi:10.1007/s10569-004-6593-5.

[24] J. L. W. III and M. J. Holzinger, "Sequential Estimation from Uninformative Priors Using Dempster-Shafer Theory," *IEEE Transactions in Aerospace And Electronic Systems (submitted)*, 2017.

[25] A. Milani, G. F. Gronchi, Z. Knežević, M. E. Sansaturio, and O. Arratia, "Orbit determination with very short arcs: II. Identifications," *Icarus*, Vol. 179, No. 2, 2005, pp. 350 – 374, doi:10.1016/j.icarus.2005.07.004.

[26] J. L. Worthy and M. J. Holzinger, "Incorporating Uncertainty in Admissible Regions for Uncorrelated Detections," *Journal of Guidance, Control, and Dynamics*, Vol. 38, No. 9, 2015, pp. 1673–1689, 10.2514/1.G000890.

[27] K. Fujimoto, D. J. Scheeres, J. Herzog, and T. Schildknecht, "Association of optical tracklets from a geosynchronous belt survey via the direct Bayesian admissible region approach," *Advances in Space Research*, Vol. 53, No. 2, 2014, pp. 295 – 308, doi:10.1016/j.asr.2013.11.021.

[28] J. L. W. III and M. J. Holzinger, "An Optimization Approach for Observation Association with Systemic Uncertainty Applied to Electro-Optical Systems," *Advances in Space Research (submitted)*, 2016.

[29] E. L. Lehmann and J. P. Romano, *Testing statistical hypotheses*. Springer Science & Business Media, 2006.

[30] H. Van Trees, *Detection, Estimation, and Modulation Theory, Optimum Array Processing*. Detection, Estimation, and Modulation Theory, Wiley, 2004.

[31] H. Cramér, *Mathematical methods of statistics*, Vol. 9. Princeton university press, 1945.

[32] A. Dempster, "The Dempster–Shafer calculus for statisticians," *International Journal of Approximate Reasoning*, Vol. 48, No. 2, 2008, pp. 365 – 377. In Memory of Philippe Smets (1938–2005), http://dx.doi.org/10.1016/j.ijar.2007.03.004.

[33] G. Shafer *et al.*, *A mathematical theory of evidence*, Vol. 1. Princeton university press Princeton, 1976.

[34] D. Dubois and H. Prade, *A Set-Theoretic View of Belief Functions*, pp. 375–410. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, 10.1007/978-3-540-44792-4-14.

[35] P. Smets and R. Kennes, "The transferable belief model," *Artificial Intelligence*, Vol. 66, No. 2, 1994, pp. 191 – 234, http://dx.doi.org/10.1016/0004-3702(94)90026-4.

[36] A. D. Jaunzemis and M. J. Holzinger, "EVIDENTIAL REASONING APPLIED TO SINGLE-OBJECT LOSS-OF-CUSTODY SCENARIOS FOR TELESCOPE TASKING," *AAS/AIAA Spaceflight Mechanics Meeting, Napa, CA*, 2016.

[37] J. Rogers and M. Costello, "Smart Projectile State Estimation Using Evidence Theory," *Journal of Guidance, Control, and Dynamics*, Vol. 35, 2017/01/19 2012, pp. 824–833, 10.2514/1.55652.

[38] P. Smets and B. Ristic, "Kalman filter and joint tracking and classification based on belief functions in the TBM framework," *Information fusion*, Vol. 8, No. 1, 2007, pp. 16–27.