INFORMATION-THEORETIC APPROACHES TO SPACE OBJECT COLLISION

K. DeMars and M. Gualdoni

Missouri University of Science and Technology, Rolla, MO 65409, USA, Email: {demarsk, mjg6y5}@mst.edu

ABSTRACT

In view of the high value of space assets and a growing space debris population, collision analysis is of great significance. Collisions between space objects can, at best, be determined in a probabilistic manner. Collision probability between objects is closely related to close approaches of the objects, where the traditional approach to determining the collision probability is predicated upon finding a time of closest approach and bounding this with a deterministic time interval. This paper investigates the application of information theory to the determination of a conjunction interval. Cases of two and more-than-two objects in close proximity are considered, and methods for determining the interval in which the objects are interacting are developed. Simulations are carried out to compare the developed methods to more conventional techniques.

Key words: Collision assessment, information theory, space object tracking.

1. INTRODUCTION

The population of tracked space objects is constantly increasing due to new launches, increases in sensor tracking capabilities, fragmentation of existing space objects, and the generation of new debris caused by random or intentional collisions. Collisions between space objects (either active spacecraft or space debris) can only be determined in a probabilistic manner due to the lack of perfect knowledge regarding the parameters that characterize the motion of space objects, i.e. the translational and rotational state of the space objects as well as the environment in which the space objects are operating. Fundamentally, collision analysis that is predicated upon nominal orbit parameters without taking into consideration the uncertainty in the parameters is inaccurate.

The probability of collision between two space objects provides a quantitative measure of the likelihood that the space objects will collide with each other. Collision probability between two objects is closely related to close approaches of the two objects. The traditional method for initializing a collision probability assessment is to determine the time of closest approach (TCA) based on predictions of the space object motion that are solely made with estimates of the states of the objects. Determining the TCA can be carried out in several ways. One approach is to numerically propagate the state estimates and find the time at which the propagated estimates come closest, particularly within a combined hard body radius [24]. More sophisticated approaches employ Taylor series expansions [13] and minimization of the relative position vector using surrogate based optimization [9] to improve upon the determination of the TCA. These approaches, however, neglect the influence that the uncertainty of the space object states may have on determining when the objects are in close proximity to one another. While the omission of such effects may be appropriate for regularly tracked objects, it begins to break down as the uncertainties in the estimates of the states grows larger.

This paper investigates the connection between concepts in information theory, such as information divergence and entropic information, and those of collisions between space objects. Such methods naturally enable the inclusion of the uncertainty of the objects and generalize the time of closest approach to a window of time over which the two objects are in close proximity. The developments are first considered for the case where the uncertainties in the space object states are taken to be Gaussian, and then this assumption is relaxed to handle the situations where the uncertainties are non-Gaussian.

Another generalization to the problem of multi-object collisions is also considered. Whereas a minimum approach distance can be employed for the close encounter of two objects, there is no natural extension to the case where more than two objects are in close proximity other than pairwise considerations of two objects at a time. Close proximity of multiple objects can readily occur in multi-spacecraft formations, such as those assembled from small satellite platforms. The information theoretic approaches,

Proc. 7th European Conference on Space Debris, Darmstadt, Germany, 18–21 April 2017, published by the ESA Space Debris Office
Ed. T. Flohrer & F. Schmitz, (http://spacedebris2017.sdo.esoc.esa.int, June 2017)

unlike the traditional methods, can be naturally extended into multi-object domains in order to provide a single framework for handling these problems.

Simulations are carried out to highlight the efficacy of the developed methods using synthetic data for space object collision scenarios. Two-object and three-object collision scenarios are considered, and the results of the information theoretic approaches are compared to more traditional approaches for determining the relevant characteristics involved in space object collisions.

2. PROBLEM STATEMENT

The central focus of this paper is on the determination of an interval during which two or more space objects are in close proximity and exhibit positional interactions, where the uncertainty of the space objects, in addition to their nominal locations, is taken into account. As such, a probabilistic quantification of the states through the use of a probability density function (pdf) is required. Assume that M objects are considered. The states of the objects at time t_k are defined as

$$\boldsymbol{x}_k^{(i)} = \begin{bmatrix} \boldsymbol{r}_k^{(i)} \\ \boldsymbol{v}_k^{(i)} \end{bmatrix},$$

where $i \in \{1, 2, \ldots, M\}$, $\boldsymbol{r}_k^{(i)}$ is the position of the i^{th} object at time t_k, and $\boldsymbol{v}_k^{(i)}$ is the velocity of the i^{th} object at time t_k. Additionally, the uncertainties in the object states are represented by the pdfs $p(\boldsymbol{x}_k^{(i)})$. In this paper, the pdfs of the single-object states are taken to be either Gaussian distributions or Gaussian mixture (GM) distributions. Regardless of the distribution, it is assumed that the pdf is known at a sequence of times, and these pdfs are the inputs to the different methods described herein to determine when the objects are or are not in close proximity. For ease of notation, the time index will be dropped in the following developments, with the understanding that the results can be applied at any time.

3. TWO-OBJECT APPROACHES

Assume that there are two objects with states given by $\boldsymbol{x}^{(i)} = [(\boldsymbol{r}^{(i)})^T \quad (\boldsymbol{v}^{(i)})^T]^T$, where $i \in \{1, 2\}$. Additionally, it is assumed that the uncertainties of the objects are Gaussian, such that the pdf of the i^{th} object is

$$p(\boldsymbol{x}^{(i)}) = p_g(\boldsymbol{x}^{(i)}; \boldsymbol{m}_x^{(i)}, \boldsymbol{P}_x^{(i)}),$$

where $p_g(\boldsymbol{x}; \boldsymbol{a}, \boldsymbol{A})$ represents the Gaussian pdf in $\boldsymbol{x}$ with mean $\boldsymbol{a}$ and covariance $\boldsymbol{A} = \boldsymbol{A}^T > \boldsymbol{0}$, such that

$$p_g(\boldsymbol{x}; \boldsymbol{a}, \boldsymbol{A}) = |2\pi\boldsymbol{A}|^{-1/2} \times \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{a})^T \boldsymbol{A}^{-1}(\boldsymbol{x}-\boldsymbol{a})\right\}.$$

The mean and covariance of the Gaussian distribution can be partitioned into position and velocity elements as

$$\boldsymbol{m}_x^{(i)} = \begin{bmatrix} \boldsymbol{m}_r^{(i)} \\ \boldsymbol{m}_v^{(i)} \end{bmatrix} \quad \text{and} \quad \boldsymbol{P}_x^{(i)} = \begin{bmatrix} \boldsymbol{P}_{rr}^{(i)} & \boldsymbol{P}_{rv}^{(i)} \\ \boldsymbol{P}_{vr}^{(i)} & \boldsymbol{P}_{vv}^{(i)} \end{bmatrix},$$

where, for the i^{th} object, $\boldsymbol{m}_r^{(i)}$ is the position mean, $\boldsymbol{m}_v^{(i)}$ is the velocity mean, $\boldsymbol{P}_{rr}^{(i)}$ is the position covariance, $\boldsymbol{P}_{vv}^{(i)}$ is the velocity covariance, and $\boldsymbol{P}_{rv}^{(i)} = (\boldsymbol{P}_{vr}^{(i)})^T$ is the position-velocity cross-covariance.

3.1. Distribution Sampling

One straightforward approach to determine the interval of interaction between two objects is to generate a number of samples from the pdfs of the objects and determine the number of samples that are within close proximity to one another. This process can be carried out at any time, which enables the determination of the interval in which the objects are interacting. Given the pdfs of the two objects, samples can be drawn from the pdfs as

$$\boldsymbol{x}^{(i,j)} \sim p_g(\boldsymbol{x}^{(i)}; \boldsymbol{m}_x^{(i)}, \boldsymbol{P}_x^{(i)}),$$

where $\boldsymbol{x}^{(i,j)}$ denotes the j^{th} sample of the i^{th} object. These samples can then be partitioned into position and velocity components as $\boldsymbol{x}^{(i,j)} = [(\boldsymbol{r}^{(i,j)})^T \quad (\boldsymbol{v}^{(i,j)})^T]^T$. As only position interactions are of interest, the velocity samples are not required. Given a sample position from each object, the distance between the samples is computed as

$$r^{(j)} = \|\boldsymbol{r}^{(2,j)} - \boldsymbol{r}^{(1,j)}\|,$$

where $\|\cdot\|$ represents the Euclidean norm. If the distance is not greater than a specified cutoff distance, c, i.e. if $r^{(j)} \leq c$, then the j^{th} samples of the objects' positions are deemed to be in close proximity. This process is carried out for a set of N samples taken from the pdfs of both of the objects. The total number of interacting samples, i.e. the number of samples satisfying $r^{(j)} \leq c$, is denoted S, and the percentage of interacting samples is computed as $T = 100(S/N)$. A tolerance, T_0, can then be set on the percentage of interacting samples, and any time at which $T > T_0$, the objects are in close proximity to one another.

There is no restriction of this method to operating on Gaussian distributions. While the preceding discussion has focused on the Gaussian case, it is straightforward to extend it to non-Gaussian distributions, provided that the distribution can be sampled.

3.2. Mahalanobis Distance

While the distribution sampling approach provides a direct method for determining an interval over which two objects are in close proximity, it also requires significant computational effort. Therefore, alternative methods are sought which can also provide a method for determining the relevant interval while accounting for the uncertainties of the objects.

Let $\boldsymbol{y}$ be defined as the joint state of the two objects, i.e.

$$\boldsymbol{y} = \begin{bmatrix} \boldsymbol{x}^{(1)} \\ \boldsymbol{x}^{(2)} \end{bmatrix} . \tag{1}$$

Provided that the two objects are uncorrelated and that their individual distributions are Gaussian, the pdf of $\boldsymbol{y}$ is also Gaussian, or

$$p(\boldsymbol{y}) = p_g(\boldsymbol{y}; \boldsymbol{m}_y, \boldsymbol{P}_y) ,$$

where the mean and covariance of $\boldsymbol{y}$ are

$$\boldsymbol{m}_y = \begin{bmatrix} \boldsymbol{m}_x^{(1)} \\ \boldsymbol{m}_x^{(2)} \end{bmatrix} \quad \text{and} \quad \boldsymbol{P}_y = \begin{bmatrix} \boldsymbol{P}_x^{(1)} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{P}_x^{(2)} \end{bmatrix} .$$

When the two objects are correlated, the pdf of $\boldsymbol{y}$ is still Gaussian; however, in this case, the covariance of $\boldsymbol{y}$ will not be block diagonal.

Now, define $\boldsymbol{z}$ to be the relative position of the two objects, which is to say

$$\boldsymbol{z} = \boldsymbol{r}^{(2)} - \boldsymbol{r}^{(1)} .$$

It is clear that $\boldsymbol{z}$ is a linear transformation of $\boldsymbol{y}$, where the transformation is given by

$$\boldsymbol{z} = \begin{bmatrix} -\boldsymbol{I} & \boldsymbol{0} & \boldsymbol{I} & \boldsymbol{0} \end{bmatrix} \boldsymbol{y} = \boldsymbol{H}\boldsymbol{y} .$$

If the dimension of $\boldsymbol{r}^{(i)}$ is n_r, then the dimension of $\boldsymbol{x}^{(i)}$ is $2n_r$, and the dimension of $\boldsymbol{H}$ is $n_r \times 4n_r$. The individual elements that comprise $\boldsymbol{H}$ are of dimension $n_r \times n_r$. Since $\boldsymbol{z}$ is a linear transformation of a Gaussian random variable, it directly follows that $\boldsymbol{z}$ is also Gaussian. The pdf of $\boldsymbol{z}$ is

$$p(\boldsymbol{z}) = p_g(\boldsymbol{z}; \boldsymbol{m}_z, \boldsymbol{P}_z) ,$$

where the mean and covariance are

$$\begin{aligned} \boldsymbol{m}_z &= \boldsymbol{H}\boldsymbol{m}_y \\ \boldsymbol{P}_z &= \boldsymbol{H}\boldsymbol{P}_y\boldsymbol{H}^T . \end{aligned}$$

Alternatively, by applying the relationship for the linear mapping, the mean and covariance of $\boldsymbol{z}$ can be found to be

$$\begin{aligned} \boldsymbol{m}_z &= \boldsymbol{m}_r^{(2)} - \boldsymbol{m}_r^{(1)} \\ \boldsymbol{P}_z &= \boldsymbol{P}_{rr}^{(1)} + \boldsymbol{P}_{rr}^{(2)} . \end{aligned}$$

Given a sample from the distribution $p(\boldsymbol{z})$, denoted by $\boldsymbol{z}^*$, the squared Mahalanobis distance is [19]

$$d^2 = (\boldsymbol{z}^* - \boldsymbol{m}_z)^T \boldsymbol{P}_z^{-1} (\boldsymbol{z}^* - \boldsymbol{m}_z) .$$

It is well known that the squared Mahalanobis distance, when calculated for samples drawn from Gaussian distributions, is statistically described by the chi-squared distribution, which is characterized by a parameter that is known as the degrees-of-freedom. The degrees-of-freedom employed in the chi-squared distribution is equal to the dimension of the random variable; hence, it follows that the distribution of the squared Mahalanobis distance is [23]

$$p(d^2) = p_{\chi^2}(d^2; n_r) ,$$

where $p_{\chi^2}(a; k)$ represents the chi-squared pdf in the variable a with parameter k (with k degrees-of-freedom). Making use of the chi-squared distribution, a probability gate, or threshold, can be used to associate (or reject association of) a sample. If the probability of accepting the sample is denoted by P, then it is possible to tabulate values of γ such that $d^2 \leq \gamma$ indicates an associated sample. For instance, when the probability gate is set to accept 90% of the samples for a 2 degree-of-freedom case, the chi-squared distribution dictates that $\gamma = 4.6052$. This method is often employed to establish confidence intervals for data association [3, 16].

In the context of determining whether two objects are in close proximity, the objective is to determine if the origin, $\boldsymbol{z}^* = \boldsymbol{0}$, is associated to the relative position distribution, $p_g(\boldsymbol{z}; \boldsymbol{m}_z, \boldsymbol{P}_z)$. The Mahalanobis distance, with respect to the origin and the relative position distribution, is

$$\begin{aligned} d^2 &= \boldsymbol{m}_z \boldsymbol{P}_z^{-1} \boldsymbol{m}_z \\ &= (\boldsymbol{m}_r^{(2)} - \boldsymbol{m}_r^{(1)})^T (\boldsymbol{P}_{rr}^{(1)} + \boldsymbol{P}_{rr}^{(2)})^{-1} (\boldsymbol{m}_r^{(2)} - \boldsymbol{m}_r^{(1)}) . \end{aligned} \tag{2}$$

Then, the interval over which $d^2 \leq \gamma$, where γ is dictated by the probability gate, P, and the degrees-of-freedom, n_r, is the interval over which the objects are in close proximity. Much as the distribution sampling approach relies on selecting the cutoff parameter, the Mahalanobis distance approach relies on selecting the probability gate.

3.3. Kullback-Leibler Divergence

The Mahalanobis distance approach improves upon the distribution sampling approach by removing the computational complexity, but it is restricted to operate on Gaussian distributions in order to establish a probability gate. To begin generalizing the interval determination to the realm of information-theoretic quantities, which also provide the means for handling non-Gaussian pdfs, the information divergence is considered.

Generally speaking, an information divergence is a measure of similarity or dissimilarity between two pdfs. Given a generic information divergence describing the directed distance from $p(\boldsymbol{x})$ to $q(\boldsymbol{x})$ denoted by $D[p||q]$, the "distance" is called a metric if [5]

1. $D[p||q] \geq 0$ with equality if and only if $p(\boldsymbol{x}) = q(\boldsymbol{x})$ (non-negativity and positive definiteness),
2. $D[p||q] = D[q||p]$ (symmetry), and
3. $D[p||r] \leq D[p||q] + D[q||r]$ (sub-additivity/triangle inequality).

Information divergences that only satisfy the first condition are not metrics and are referred to as asymmetric divergences. Satisfaction of the second condition necessarily removes the restriction of referring to the divergence as asymmetric.

The most often used information divergence is the Kullback-Leibler (KL) divergence, which is given by [6, 17]

$$D_{KL}[p||q] = \int_{\mathcal{X}} p(\boldsymbol{x}) \log \frac{p(\boldsymbol{x})}{q(\boldsymbol{x})} \, \mathrm{d}\boldsymbol{x} \,. \tag{3}$$

The KL divergence is defined for arbitrary pdfs, $p(\boldsymbol{x})$ and $q(\boldsymbol{x})$, where $\boldsymbol{x} \in \mathcal{X} \subseteq \mathbb{R}^d$. It is, however, only possible to determine closed-form relationships for certain cases, such as the case where both pdfs are taken to be Gaussian. Let $p(\boldsymbol{x})$ be a Gaussian pdf with mean $\boldsymbol{a}$ and covariance $\boldsymbol{A}$, and let $q(\boldsymbol{x})$ be a Gaussian pdf with mean $\boldsymbol{b}$ and covariance $\boldsymbol{B}$; that is,

$$p(\boldsymbol{x}) = p_g(\boldsymbol{x}; \boldsymbol{a}, \boldsymbol{A}) \quad \text{and} \quad q(\boldsymbol{x}) = p_g(\boldsymbol{x}; \boldsymbol{b}, \boldsymbol{B}) \,. \tag{4}$$

Substituting $p(\boldsymbol{x})$ and $q(\boldsymbol{x})$ from Eq. (4) into the KL divergence of Eq. (3), it can be shown that the KL divergence for Gaussian distributions is

$$D_{KL}[p||q] = \frac{1}{2}\Big[\log\left|\boldsymbol{B}\boldsymbol{A}^{-1}\right| + \mathrm{tr}\left\{\boldsymbol{B}^{-1}\boldsymbol{A}\right\} + (\boldsymbol{a}-\boldsymbol{b})^T\boldsymbol{B}^{-1}(\boldsymbol{a}-\boldsymbol{b}) - d\Big] \,,$$

where $\mathrm{tr}\{\cdot\}$ represents the matrix trace.

One downside to the KL divergence is that it is an asymmetric divergence. While this does not prevent one from forming the directed "distance" between the two distributions $p(\boldsymbol{x})$ and $q(\boldsymbol{x})$, it does mean that the "distance from $p(\boldsymbol{x})$ to $q(\boldsymbol{x})$" is not the same as the "distance from $q(\boldsymbol{x})$ to $p(\boldsymbol{x})$." For instance, the reverse KL divergence, which is given by

$$D_{KL}[q||p] = \frac{1}{2}\Big[\log\left|\boldsymbol{A}\boldsymbol{B}^{-1}\right| + \mathrm{tr}\left\{\boldsymbol{A}^{-1}\boldsymbol{B}\right\} + (\boldsymbol{b}-\boldsymbol{a})^T\boldsymbol{A}^{-1}(\boldsymbol{b}-\boldsymbol{a}) - d\Big] \,,$$

is not equivalent to $D_{KL}[p||q]$. Thus, an interval determination using the KL divergence would differ from an interval determination using the reverse KL divergence. To circumvent this, the symmetric KL divergence is defined as

$$D_S[p||q] = \frac{1}{2}\Big(D_{KL}[p||q] + D_{KL}[q||p]\Big) \,.$$

Substituting for the relationships for the KL and reverse KL divergences, it follows that the symmetric KL divergence can be expressed for Gaussian distributions as

$$D_S[p||q] = \frac{1}{4}\Big[\mathrm{tr}\left\{\boldsymbol{B}^{-1}\boldsymbol{A} + \boldsymbol{A}^{-1}\boldsymbol{B}\right\} + (\boldsymbol{a}-\boldsymbol{b})^T(\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1})(\boldsymbol{a}-\boldsymbol{b}) - 2d\Big] \,. \tag{5}$$

The relationship for the symmetric KL divergence given in Eq. (5) holds for any Gaussian distributions $p(\boldsymbol{x})$ and $q(\boldsymbol{x})$. The symmetric KL divergence is now specialized to the space object collision case. First, note that "collisions" in the velocity space are not of interest; only collisions that occur in position space are of interest. Therefore, $p(\boldsymbol{x})$ and $q(\boldsymbol{x})$ are chosen to be the position marginal densities

$$p(\boldsymbol{r}_1) = p_g(\boldsymbol{r}_1; \boldsymbol{m}_r^{(1)}, \boldsymbol{P}_{rr}^{(1)})$$
$$q(\boldsymbol{r}_2) = p_g(\boldsymbol{r}_2; \boldsymbol{m}_r^{(2)}, \boldsymbol{P}_{rr}^{(2)}) \,.$$

As the symmetric KL divergence is under consideration, it is not relevant which object is represented by $p(\boldsymbol{x})$ and which object is represented by $q(\boldsymbol{x})$; for either the KL or reverse KL divergence, this is not the case. With this specialization of the pdfs, the symmetric KL divergence is

$$D_S[p||q] = \frac{1}{4}\Big[\mathrm{tr}\left\{\boldsymbol{\Lambda}_{rr}^{(2)}\boldsymbol{P}_{rr}^{(1)} + \boldsymbol{\Lambda}_{rr}^{(1)}\boldsymbol{P}_{rr}^{(2)}\right\} - 2n_r + (\boldsymbol{m}_r^{(1)} - \boldsymbol{m}_r^{(2)})^T(\boldsymbol{\Lambda}_{rr}^{(1)} + \boldsymbol{\Lambda}_{rr}^{(2)})(\boldsymbol{m}_r^{(1)} - \boldsymbol{m}_r^{(2)})\Big] \,, \tag{6}$$

where $\boldsymbol{\Lambda}_{rr}^{(i)} = (\boldsymbol{P}_{rr}^{(i)})^{-1}$.

An interesting connection between the symmetric KL divergence and the Mahalanobis distance can be established in the special case where the position covariances of the two objects are identical. Let $\boldsymbol{P}_{rr} = \boldsymbol{P}_{rr}^{(1)} = \boldsymbol{P}_{rr}^{(2)}$; then, it can be shown that the symmetric KL divergence given in Eq. (6) can be expressed as

$$D_S[p||q] = \frac{1}{2}(\boldsymbol{m}_r^{(2)} - \boldsymbol{m}_r^{(1)})^T\boldsymbol{P}_{rr}^{-1}(\boldsymbol{m}_r^{(2)} - \boldsymbol{m}_r^{(1)}) \,.$$

In a similar fashion, the squared Mahalanobis distance given in Eq. (2) becomes

$$d^2 = \frac{1}{2}(\boldsymbol{m}_r^{(2)} - \boldsymbol{m}_r^{(1)})^T\boldsymbol{P}_{rr}^{-1}(\boldsymbol{m}_r^{(2)} - \boldsymbol{m}_r^{(1)}) \,.$$

Thus, when the two objects possess the same position covariance, the symmetric KL divergence and the squared Mahalanobis distance (computed with respect to the origin) coincide. This is, of course, a rather specialized case. It does, however, provide some insight into the nature of the symmetric KL divergence.

4. MULTI-OBJECT APPROACHES

The methods described thus far are predicated and built upon the assumption that the interval determination for the interaction of the objects is determined for the case of only two objects. As described, the distribution sampling, Mahalanobis distance, and symmetric KL divergence methods necessarily depend upon the restriction to two objects. One way of overcoming this shortcoming is to consider all possible pairwise combinations of the set of $M \geq 2$ objects and apply the methods as described previously. Even when advanced screening methods are employed to reduce the number of objects considered [1, 14], the combinatoric growth associated with pairwise combinations can significantly increase the required computational burden.

4.1. Information Entropy

An alternative approach is to further investigate information-theoretic methods, like the symmetric KL divergence, but within the context of multitarget tracking. A related idea to the information divergence is the information entropy, or simply entropy.

Information entropy originated in the context of quantifying the amount of uncertainty in the generation of a received message or in the amount of "choice" present in the transmission of a message in communication [27], which built upon earlier works that investigated descriptions of the amount of information present in an event space [12]. Essentially, the entropy is a measure of the "size" of the uncertainty of a random variable. The entropy is minimal when there is no uncertainty present, i.e. the random variable becomes deterministic; the entropy is maximal as all possible outcomes of the random variable become equiprobable; and the entropy monotonically increases as the uncertainty increases.

For the communication case, the random variable is discrete in nature, and the uncertainty of the random variable is described in terms of a probability mass function (pmf). The resulting entropy is termed the Shannon entropy, which is [27]

$$H[\rho] = -\sum_{n=-\infty}^{\infty} \rho(n) \log\{\rho(n)\}\,,$$

where n is the random variable and $\rho(n)$ is the pmf of n. Subsequently, this definition was extended to continuous random variables. For a probability density $p(\boldsymbol{x})$ with support over the entire state space $\mathcal{X} \subseteq \mathbb{R}^d$, the Shannon (differential) entropy, $H[p]$, is defined to be

$$H[p] = -\int_{\mathcal{X}} p(\boldsymbol{x}) \log\left\{p(\boldsymbol{x})\right\} \mathrm{d}\boldsymbol{x}\,. \tag{7}$$

The terminology differential entropy is often used to distinguish between the entropy for discrete and continuous random variables, but this distinction will be omitted herein. Whereas the Shannon entropy for discrete random variables is absolute, the entropy of Eq. (7) is relative to the coordinate system. That is, if the coordinates are changed, including the units of $\boldsymbol{x}$, the entropy will change.

While Shannon entropy provides a scalar measure for characterizing the uncertainty of a random variable, it is not the only option available. Shannon's definition of entropy was generalized by Rényi to produce a family of entropy measures [7, 25], which were originally termed "informations of order α," but which are now referred to simply as Rényi entropy. The Rényi entropy is defined as

$$H^{(\alpha)}[p] = \frac{1}{1-\alpha} \log\left\{\int_{\mathcal{X}} p^{\alpha}(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}\right\}, \tag{8}$$

where α is a control parameter. The Rényi entropy is undefined for $\alpha = 1$, but in the limit as $\alpha \to 1$, it is well known that the Rényi entropy reproduces the Shannon entropy. Similar to the Shannon entropy, the Rényi entropy is relative to the coordinate system.

Entropy can be used to describe the relative concentrations of uncertainty for different scenarios. That is, a smaller entropy indicates a more highly concentrated, or more highly localized, uncertainty. As an example, consider the schematic representation of collisions occurring between three objects shown in Fig. 1. In the beginning frame, which is the one denoted t_1, the three objects are distinctly separated. As time continues, Objects #2 and #3 begin to interact, as observed in frame t_2. This interaction continues through frame t_3, at which time Object #1 begins interacting with the other two objects. Finally, in frame t_4, the interaction between three objects has returned to interaction between two objects. This example shows that concentrations of uncertainty are exactly representing interactions between objects.

Shannon entropy, as defined in Eq. (7), cannot be used to describe the level of interaction between multiple objects, as it is the entropy obtained from a pdf representing the uncertainty of a single object. Augmenting the state to include multiple objects, such as is done in Eq. (1), also does not work. When the objects are independent, it is straightforward to

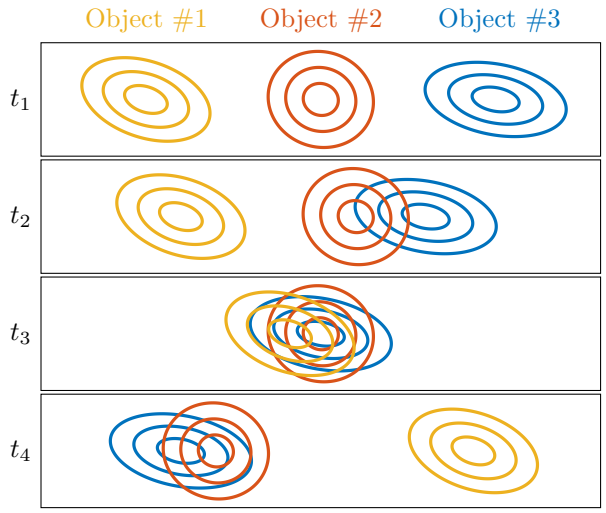

Figure 1. Schematic representation of collisions between three objects.

show that the entropy of the distribution for the augmented state is equal to the sum of the entropies of the distributions of the individual objects. In essence, since the state space is extended to represent the multiple objects, there is no collision between the objects in this space. Instead, the appropriate approach is to work within the realm of multitarget tracking.

4.2. Random Finite Sets

When considering multi-object collisions using multitarget tracking, the use of random vectors is no longer appropriate for representing the state of the system. Instead, random finite sets (RFSs) are used to represent the multitarget state as [11, 22]

$$\boldsymbol{X} = \{\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \ldots, \boldsymbol{x}^{(M)}\},$$

where there are M objects, or targets, in the multitarget state. Each of the distinct elements $\boldsymbol{x}^{(i)} \in \boldsymbol{X}$, $\forall i \in \{1, \ldots, M\}$ is a conventional target state, such as the position and velocity of an object.

An RFS can also be described statistically with a density, albeit a multitarget pdf, in a similar way to describing a random vector using a conventional (single-target) pdf. This is a fundamental element of multitarget filtering using concepts from finite set statistics (FISST) [11, 22]. One common assumption is to take the the RFS, $\boldsymbol{X}$, to be an independent and identically distributed (i.i.d.) cluster process; then, the multitarget pdf can be represented as

$$f(\boldsymbol{X}) = n!\rho(n)\prod_{k=1}^{n} s(\boldsymbol{x}^{(k)}), \tag{9}$$

where n is the cardinality of the RFS, $\rho(n)$ is the cardinality distribution, and $s(\boldsymbol{x}^{(i)})$ is the single target spatial density of the i^{th} target, which is taken to be a valid (single-target) pdf. Further descriptions can be obtained within the i.i.d. cluster process by modeling the cardinality distribution as a Poisson distribution, which is given by

$$\rho(n) = \frac{1}{n!}\lambda^n e^{-\lambda}, \tag{10}$$

where λ is the rate parameter. It is also well known that λ is both the mean and the variance of the Poisson distribution.

Moment approximation methods are often employed in multitarget tracking due to their ability to provide computationally tractable solutions. Two such methods, referred to as the probability hypothesis density (PHD) [20, 30] and cardinalized probability hypothesis density (CPHD) [21, 31, 32] filters, approximate the multitarget Bayes filter by operating on the intensity function, $v(\boldsymbol{x})$, which is the first-order moment of the multitarget pdf. The intensity function is defined such that integrating over the entire support of the intensity resolves to the expected cardinality of the set, or

$$\lambda = \int_{\mathcal{X}} v(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}. \tag{11}$$

For an RFS with a Poisson cardinality distribution, the intensity function is given by

$$v(\boldsymbol{x}) = \lambda s(\boldsymbol{x}). \tag{12}$$

The intensity function is commonly expressed as a weighted sum of Gaussian distributions of the form [22, 30]

$$v(\boldsymbol{x}) = \sum_{\ell=1}^{L} w^{(\ell)} p_g(\boldsymbol{x}; \boldsymbol{m}^{(\ell)}, \boldsymbol{P}^{(\ell)}). \tag{13}$$

It is important to note that the GM of Eq. (13) is slightly different from that of [2, 28] and subsequent works in GM filtering methods. The difference is that the conventional GM representation is constructed to represent a pdf, and the weights must sum to unity. From Eqs. (11) and (13), it directly follows that $\lambda = \sum_{\ell=1}^{L} w^{(\ell)}$. That is, when the GM is used to describe an intensity function, the expected number of objects in the state space is equal to the sum of the weights in the GM representation of the intensity function.

4.3. Shannon Entropy

Just as the random vector is lifted to an RFS, the single-target integral becomes a set integral; thus, the Shannon entropy defined in Eq. (7) may be recast for RFSs via the set integral, such that

$$H[f] = -\int_{S} f(\boldsymbol{X}) \log\{f(\boldsymbol{X})\}\delta\boldsymbol{X}. \tag{14}$$

This definition of the multitarget Shannon entropy depends upon the coordinate system in much the same way that the Shannon entropy of Eq. (7) does. For the multitarget problem, however, the situation is compounded since the units change with the cardinality of the RFS. As such, a more appropriate definition for the multitarget Shannon entropy is [4, 26]

$$H[f] = -\int_S f(\boldsymbol{X}) \log\left\{u^{|\boldsymbol{X}|} f(\boldsymbol{X})\right\} \delta\boldsymbol{X}\,, \qquad (15)$$

where $u^{-|\boldsymbol{X}|}$ is the units of the FISST density, $f(\boldsymbol{X})$, and $|\boldsymbol{X}|$ is the cardinality of the RFS $\boldsymbol{X}$. For simplicity and consistency with existing results, the naive form of the multitarget Shannon entropy will be used in proceeding.

Substituting the i.i.d. cluster RFS distribution of Eq. (9) and applying the definition of a set integral [22] to Eq. (14), the Shannon entropy of an i.i.d. cluster RFS is given by

$$H[f] = -\sum_{n=0}^{\infty} \frac{1}{n!} \int_{\mathcal{X}^n} \left[n!\rho(n) \prod_{k=1}^{n} s(\boldsymbol{x}^{(k)})\right] \times \log\left\{n!\rho(n) \prod_{i=1}^{n} s(\boldsymbol{x}^{(i)})\right\} \mathrm{d}\boldsymbol{x}^{(1)} \cdots \mathrm{d}\boldsymbol{x}^{(n)}\,,$$

where $\mathcal{X}^n$ is the Cartesian product of n copies of the state space. The benefits of the logarithm become apparent quickly, as this allows the logarithm to be decomposed into terms dependent on the single-target spatial densities and terms dependent only on the cardinality n. Thus, the entropy may be expressed as

$$H[f] = -\sum_{n=0}^{\infty} \left\{\rho(n) \int_{\mathcal{X}^n} \left[\prod_{k=1}^{n} s(\boldsymbol{x}^{(k)})\right] \log\left\{n!\rho(n)\right\} \times \mathrm{d}\boldsymbol{x}^{(1)} \cdots \mathrm{d}\boldsymbol{x}^{(n)}\right\} - \sum_{n=0}^{\infty} \left\{\rho(n) \int_{\mathcal{X}^n} \left[\prod_{k=1}^{n} s(\boldsymbol{x}^{(k)})\right] \times \sum_{i=1}^{n} \log\left\{s(\boldsymbol{x}^{(i)})\right\} \mathrm{d}\boldsymbol{x}^{(1)} \cdots \mathrm{d}\boldsymbol{x}^{(n)}\right\}.$$

The product of the single-target spatial densities applies a sifting-like effect; since each logarithm term is only dependent upon $\boldsymbol{x}^{(i)}$, the remaining $n-1$ integrals can be evaluated over each spatial density. The result of each of the $n-1$ evaluations is unity, as the single-target spatial densities are taken to be valid pdfs. Therefore, only a sum of integrals over the target state space, $\mathcal{X}$, remains, and the result is that the entropy is given by

$$H[f] = -\sum_{n=0}^{\infty} \left[\rho(n) \log\left\{n!\rho(n)\right\}\right] - \sum_{n=0}^{\infty} \left[\rho(n) \sum_{i=1}^{n} \int_{\mathcal{X}} s(\boldsymbol{x}^{(i)}) \log\left\{s(\boldsymbol{x}^{(i)})\right\} \mathrm{d}\boldsymbol{x}^{(i)}\right].$$

While the variable $\boldsymbol{x}^{(i)}$ has been maintained up to this point to distinguish between the multiple integration dimensions, it is recognized that the i.i.d. assumption can now be used to simplify the expression by replacing $\boldsymbol{x}^{(i)}$ with a non-indexed $\boldsymbol{x}$ since each n-tuple integral has been reduced to a single integral over the state space $\mathcal{X}$. The result of dropping the index is a sum of n identical integrals, or

$$H[f] = -\sum_{n=0}^{\infty} \left[\rho(n) \log\left\{n!\rho(n)\right\}\right] - \sum_{n=0}^{\infty} \left[n\rho(n)\right] \int_{\mathcal{X}} s(\boldsymbol{x}) \log\left\{s(\boldsymbol{x})\right\} \mathrm{d}\boldsymbol{x}\,.$$

The summation in the second term is simply the definition of the mean of the cardinality distribution. Denoting this mean by μ, it follows that

$$H[f] = -\sum_{n=0}^{\infty} \left[\rho(n) \log\left\{n!\rho(n)\right\}\right] - \mu \int_{\mathcal{X}} s(\boldsymbol{x}) \log\left\{s(\boldsymbol{x})\right\} \mathrm{d}\boldsymbol{x}\,. \qquad (16)$$

Equation (16) is the Shannon entropy for an RFS under the assumption that it is distributed according to an i.i.d. cluster process, which is defined by Eq. (9).

The expression for the Shannon entropy of an i.i.d. cluster process given in Eq. (16) can be specialized by making an assumption on the cardinality distribution. Substituting the Poisson cardinality distribution given by Eq. (10) into the entropy relationship of Eq. (16), while noting that $\mu = \lambda$, it can be shown that the infinite summations are eliminated, resulting in

$$H[f] = \lambda - \lambda \log \lambda - \lambda \int_{\mathcal{X}} s(\boldsymbol{x}) \log\left\{s(\boldsymbol{x})\right\} \mathrm{d}\boldsymbol{x}\,. \qquad (17)$$

Solving for the spatial density in terms of the intensity function from Eq. (12) and substituting the result into Eq. (17), it follows that the Shannon entropy may be expressed solely in terms of the rate parameter, λ, and the intensity function, $v(\boldsymbol{x})$, as

$$H[f] = \lambda - \lambda \log \lambda - \int_{\mathcal{X}} v(\boldsymbol{x}) \log\left\{v(\boldsymbol{x})\right\} \mathrm{d}\boldsymbol{x} + \int_{\mathcal{X}} v(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} \log \lambda\,.$$

The integral in the final term can be replaced by recalling from Eq. (11) that it is simply the rate parameter, λ; therefore,

$$H[f] = \lambda - \int_{\mathcal{X}} v(\boldsymbol{x}) \log\left\{v(\boldsymbol{x})\right\} \mathrm{d}\boldsymbol{x}\,. \qquad (18)$$

The result of Eq. (18) is also given in [8].

Equation (18) is the Shannon entropy for an RFS under the assumption that it is distributed according to an i.i.d. cluster process, with the further stipulation that the cardinality distribution is Poisson. The result shows that the entropy is composed of a cardinality entropy term and a spatial entropy term. It should be noted that the spatial entropy term in Eq. (18) still contains cardinality elements through the representation of the intensity function. It is interesting to note that the spatial entropy term is of the exact form of the single-target entropy given by Eq. (7), but with the multitarget intensity function in place of the single-target pdf. Thus, the spatial term will tend to exhibit the same characteristics observed with the single-target entropy, lending intuition to the analysis of multitarget entropy, and the cardinality term will cause the entropy to rise as the number of targets in the multitarget state increases.

Except in special cases of the intensity, such as an intensity that is Gaussian, the Shannon entropy of Eq. (18) cannot be found in closed-form. For instance, when the intensity is represented as in Eq. (13), no closed-form solution to the entropy of Eq. (18) can be found. In such situations, numerical solutions to the integral, such as those obtained through Monte Carlo integration, must be used.

4.4. Rényi Entropy

Similar to the concept of extending the Shannon entropy into the multitarget domain, the Rényi entropy is lifted from Eq. (8) through the definition of the set integral to yield

$$H^{(\alpha)}[f] = \frac{1}{1-\alpha} \log \left\{ \int_S f^{\alpha}(\boldsymbol{X}) \delta \boldsymbol{X} \right\},$$

where α is the control parameter. Just the same as $\lim_{\alpha \to 1} H^{(\alpha)}[p] = H[p]$, it can be shown that $\lim_{\alpha \to 1} H^{(\alpha)}[f] = H[f]$; that is, the multitarget Rényi entropy is a generalization of the multitarget Shannon entropy. As with the Shannon entropy, the naive extension of Eq. (8) is considered herein; however, an extension similar to Eq. (15) to consider the units of the FISST density for the multitarget Rényi entropy can be carried out.

The definition of the set integral and the i.i.d. cluster process definition from Eq. (9) can be applied to give the Rényi entropy for i.i.d. cluster processes as

$$H^{(\alpha)}[f] = \frac{1}{1-\alpha} \log \left\{ \sum_{n=0}^{\infty} \frac{1}{n!} \int_{\mathcal{X}^n} [n! \rho(n) \times s(\boldsymbol{x}^{(1)}) \cdots s(\boldsymbol{x}^{(n)})]^{\alpha} \, \mathrm{d}\boldsymbol{x}^{(1)} \cdots \mathrm{d}\boldsymbol{x}^{(n)} \right\}.$$

Distributing the exponent to each term within the integrals and noting that each of the integrals can be separated, it follows that

$$H^{(\alpha)}[f] = \frac{1}{1-\alpha} \log \left\{ \sum_{n=0}^{\infty} \frac{1}{n!} (n!)^{\alpha} \rho^{\alpha}(n) \times \int_{\mathcal{X}} s^{\alpha}(\boldsymbol{x}^{(1)}) \, \mathrm{d}\boldsymbol{x}^{(1)} \cdots \int_{\mathcal{X}} s^{\alpha}(\boldsymbol{x}^{(n)}) \, \mathrm{d}\boldsymbol{x}^{(n)} \right\}.$$

As before, the indices on the $\boldsymbol{x}^{(i)}$ have been used to distinguish between individual integration dimensions; however, now that the integrals involving each of the $\boldsymbol{x}^{(i)}$ terms have been separated into a product of n integrals, this index can be dropped to yield

$$H^{(\alpha)}[f] = \frac{1}{1-\alpha} \log \left\{ \sum_{n=0}^{\infty} \frac{1}{n!} (n!)^{\alpha} \rho^{\alpha}(n) \times \left[\int_{\mathcal{X}} s^{\alpha}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \right]^n \right\}. \tag{19}$$

Equation (19) is the Rényi entropy for an RFS under the assumption that it is distributed according to an i.i.d. cluster process. Unlike the Shannon entropy equivalent given in Eq. (16), there is no clear separation between cardinality-induced entropy and spatial-induced entropy.

The Rényi entropy for an i.i.d. cluster process may be specialized by providing a further restriction on the cardinality distribution. The Poisson cardinality distribution defined by Eq. (10) is now considered. Substituting Eq. (10) for the cardinality distribution in Eq. (19) and reducing yields

$$H^{(\alpha)}[f] = \frac{1}{1-\alpha} \log \left\{ e^{-\alpha\lambda} \sum_{n=0}^{\infty} \frac{1}{n!} \times \left[\int_{\mathcal{X}} \big(\lambda s(\boldsymbol{x})\big)^{\alpha} \, \mathrm{d}\boldsymbol{x} \right]^n \right\}. \tag{20}$$

From the series expansion of the exponential function, $e^z = \sum_{n=0}^{\infty} z^n / n!$, it can be seen that the summation term appearing in Eq. (20) may be written as an exponentional, such that

$$H^{(\alpha)}[f] = \frac{1}{1-\alpha} \log \left\{ e^{-\alpha\lambda + \int_{\mathcal{X}} (\lambda s(\boldsymbol{x}))^{\alpha} \, \mathrm{d}\boldsymbol{x}} \right\},$$

which may be further reduced by taking the logarithm to be base e to yield

$$H^{(\alpha)}[f] = \frac{1}{1-\alpha} \left[-\alpha\lambda + \int_{\mathcal{X}} \big(\lambda s(\boldsymbol{x})\big)^{\alpha} \, \mathrm{d}\boldsymbol{x} \right].$$

From the definition of the intensity function for a Poisson RFS in Eq. (12), the Rényi entropy is given in terms of the rate parameter, λ, and the intensity function, $v(\boldsymbol{x})$, to be

$$H^{(\alpha)}[f] = -\frac{\alpha\lambda}{1-\alpha} + \frac{1}{1-\alpha} \int_{\mathcal{X}} v^{\alpha}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}. \tag{21}$$

Equation (21) is the Rényi entropy of order α for an RFS under the assumption that it is distributed according to an i.i.d. cluster process, with the further stipulation that the cardinality distribution is Poisson. The result, much like the Shannon entropy, shows that the entropy is composed of a cardinality entropy term and a spatial entropy term, where it is worth noting that the spatial entropy term in Eq. (21) contains cardinality elements through the representation of the intensity function. Unlike the Shannon entropy, however, the spatial element of the Rényi entropy does not take on the same form as the single-target Rényi entropy, which can be seen by comparing Eqs. (8) and (21).

In contrast to the Shannon entropy of Eq. (18), the Rényi entropy of Eq. (21) can be found in closed-form for certain choices of the control parameter α when the intensity function is given by Eq. (13). For instance, when $\alpha = 2$, the Rényi entropy is

$$H^{(2)}[f] = 2\lambda - \int_{\mathcal{X}} v^2(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\,. \tag{22}$$

Substituting for the GM representation of the intensity function, it can be shown that

$$H^{(2)}[f] = 2\sum_{\ell=1}^{L} w^{(\ell)} - \sum_{i=1}^{L}\sum_{j=1}^{L}\Big[w^{(i)}w^{(j)} \times \Gamma\big(\boldsymbol{m}^{(i)} - \boldsymbol{m}^{(j)}, \boldsymbol{P}^{(i)} + \boldsymbol{P}^{(j)}\big)\Big]\,, \tag{23}$$

where

$$\Gamma(\boldsymbol{a}, \boldsymbol{A}) = |2\pi\boldsymbol{A}|^{-1/2}\exp\left\{-\frac{1}{2}\boldsymbol{a}^T\boldsymbol{A}^{-1}\boldsymbol{a}\right\}\,.$$

The first term in Eq. (23) is the cardinality entropy and is simply given as the sum of the weights of the GM representation of the intensity; for a constant number of guaranteed existing objects, it is constant. The spatial component of Eq. (23), i.e. the double summation term, is of prime interest when considering the determination of an interval of close approaches. From Eq. (23), the spatial element of the Rényi entropy may be expressed as

$$H_S^{(2)}(\boldsymbol{\Theta}) = -\sum_{i=1}^{L}\sum_{j=1}^{L}\Big[w^{(i)}w^{(j)} \times \Gamma\big(\boldsymbol{m}^{(i)} - \boldsymbol{m}^{(j)}, \boldsymbol{P}^{(i)} + \boldsymbol{P}^{(j)}\big)\Big]\,, \tag{24}$$

where $\boldsymbol{\Theta}$ is the collection of parameters that defines the GM representation of the intensity for which the spatial entropy is computed, i.e. $\boldsymbol{\Theta} = \{w^{(\ell)}, \boldsymbol{m}^{(\ell)}, \boldsymbol{P}^{(\ell)}\}_{\ell=1}^{L}$ when all parameters of the GM representation of the intensity are considered.

5. RESULTS AND DISCUSSION

The methods developed for the determination of an interval during which two or more objects are interacting, or the interval during which the calculation of the probability of collision should be carried out, are applied to three test cases. The first test case examines a two-object collision case, where the uncertainty in the states of the objects is Gaussian, and compares all of the techniques presented in this work. The second test case examines a three-object collision case, again with Gaussian uncertainty, and compares the distribution sampling method to the Rényi entropy method. The final case examines a two-object collision case with non-Gaussian uncertainties and compares the distribution sampling method to the Rényi entropy method.

For simplicity, the dynamic motion of the objects considered in this work is taken to be represented using relative motion dynamics. A fictitious reference object is used to describe the center of a rotating frame, and the motion of nearby objects is taken to be described by the relative motion with respect to this reference object. The reference object is taken to be in a circular orbit, and a local-vertical, local-horizontal frame (with the x-axis representing the cross-track direction and the y-axis representing the along-track direction) is attached to the reference object. Furthermore, only planar motion will be considered in order to simplify the analysis. The Clohessy-Wiltshire (CW) model is used to represent the dynamics of the nearby objects, such that, for a state defined as the relative position and velocity of the form $\boldsymbol{x}_k = [\boldsymbol{r}_k^T \quad \boldsymbol{v}_k^T]^T$, the dynamics of the state are given by the linear, discrete-time, noiseless system

$$\boldsymbol{x}_k = \boldsymbol{F}_{k-1}\boldsymbol{x}_{k-1}\,,$$

where $\boldsymbol{F}_{k-1}$ is the state transition matrix of the CW model, which is taken to be

$$\boldsymbol{F}_{k-1} = \begin{bmatrix} \boldsymbol{\Phi}_{rr} & \boldsymbol{\Phi}_{rv} \\ \boldsymbol{\Phi}_{vr} & \boldsymbol{\Phi}_{vv} \end{bmatrix}\,,$$

where

$$\boldsymbol{\Phi}_{rr} = \begin{bmatrix} 4 - 3\cos\psi & 0 \\ 6(\sin\psi - \psi) & 1 \end{bmatrix}$$
$$\boldsymbol{\Phi}_{rv} = \begin{bmatrix} \frac{1}{n}\sin\psi & \frac{2}{n}(1 - \cos\psi) \\ \frac{2}{n}(\cos\psi - 1) & \frac{1}{n}(4\sin\psi - 3\psi) \end{bmatrix}$$
$$\boldsymbol{\Phi}_{vr} = \begin{bmatrix} 3n\sin\psi & 0 \\ 6n(\cos\psi - 1) & 0 \end{bmatrix}$$
$$\boldsymbol{\Phi}_{vv} = \begin{bmatrix} \cos\psi & 2\sin\psi \\ -2\sin\psi & -3 + 4\cos\psi \end{bmatrix}\,,$$

and $\psi = n(t_k - t_{k-1})$, where n is the mean motion of the reference object. Note that the CW model is only one of many available relative motion models. Other options for relative motion [10, 15, 18, 29] can be used, or non-relative motion, such as inertial, central body motion, can also be employed. The CW model is chosen here because of its linear nature, which allows for simple state propagation.

The reference object is chosen to have a mean motion of $n = 15.91$ rev/day and the time step is taken to be a constant $\Delta t = t_k - t_{k-1} = 10$ sec. In all three examples, initial pdfs for the objects under consideration are specified by either Gaussian or GM pdfs through the specification of means, covariances, and, for the case of GMs, weights of the GM. All weights are held constant, and all of the means and covariances are propagated using the linear dynamics of the CW model.

5.1. Example #1

The initial means of the two objects considered for the first example are given in Table 1. In addition, the uncertainties of the two objects are Gaussian, with diagonal covariance matrices. The standard deviations are given in Table 2.

Table 1. Initial means for Example #1.

Object	x [m]	y [m]	$\dot{x}$ [mm/s]	$\dot{y}$ [mm/s]
1	5.0	38.3	4.4	−11.2
2	5.4	18.5	−7.2	−7.7

Table 2. Initial standard deviations for Example #1.

Object	x [m]	y [m]	$\dot{x}$ [mm/s]	$\dot{y}$ [mm/s]
1	1.0	0.5	1.0	2.0
2	0.5	1.0	2.0	1.0

The time history of the means for the two objects is computed for 300 time steps, and the resulting trajectories are illustrated in Fig. (2). Similarly, the covariance histories are computed and plotted at 5-step intervals in Fig. (3). Based on these figures, it is observed that there is likely a time interval for which the two objects experience close interaction.

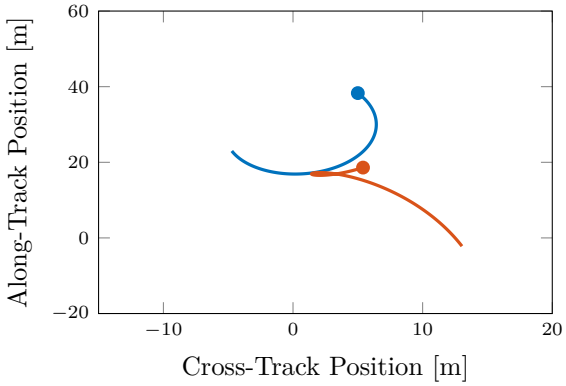

Figure 2. Mean trajectories for Example #1.

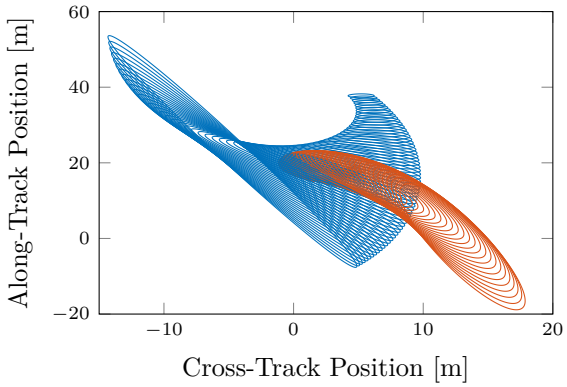

Figure 3. Uncertainty trajectories for Example #1.

To determine when the two objects are in close interaction, the distribution sampling, Mahalanobis distance, symmetric KL divergence, and Rényi entropy methods are applied at each step of the simulation. For each method, the position marginal distributions are obtained from the propagated means and covariances of each object.

First, consider the distribution sampling, Mahalanobis distance, and symmetric KL divergence methods. The position marginal distributions are sampled 1×10^6 times, the relative distance between the samples are computed, and the number of samples with a relative distance less than $c = 1$ m are counted. This provides the percentage of interacting samples shown in Fig. (4). The position marginal distributions are used to compute the squared Mahalanobis distance from Eq. (2). The Mahalanobis distance, along with a 90% confidence interval, which is obtained from the chi-squared distribution, is illustrated in Fig. (5). The position marginal distributions are furthermore used to compute the symmetric KL divergence from Eq. (5). For scaling, the square root of the symmetric KL divergence is shown in Fig. (6). Each of these three methods shows a

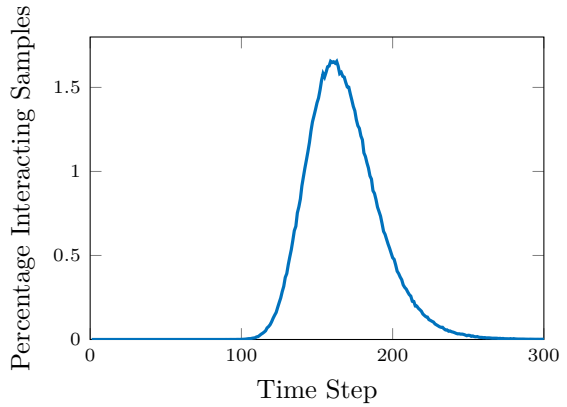

Figure 4. Application of the distribution sampling method to Example #1.

similar time of closest approach for the two objects, where the uncertainty is included in each method. The symmetric KL divergence, however, provides an analysis from which it is difficult to determine the interval over which the two objects are interacting. For this reason, this method will not be considered in the following analyses.

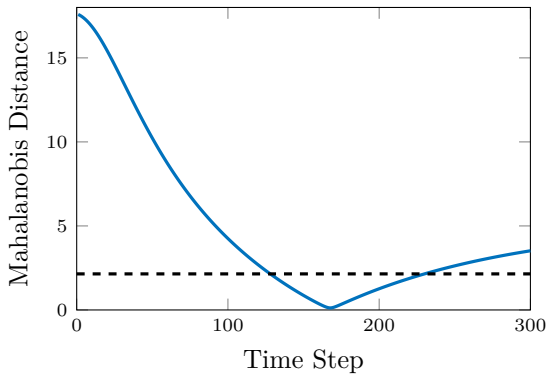

Figure 5. Application of the Mahalanobis distance method to Example #1.

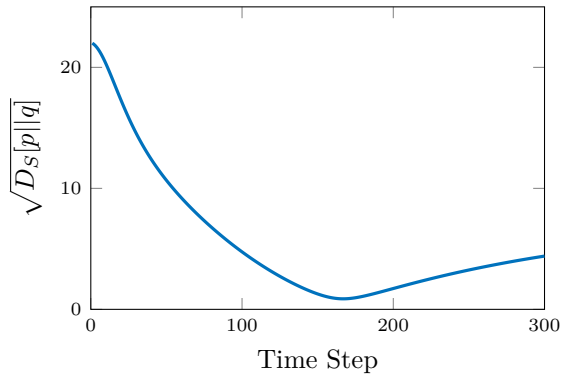

Figure 6. Application of the symmetric KL divergence method to Example #1.

Now, consider the Rényi entropy method. The propagation of the means and covariances for the two objects yields the position marginal means as $\boldsymbol{m}_r^{(i)}$ and the position marginal covariances as $\boldsymbol{P}_{rr}^{(i)}$ for $i \in \{1, 2\}$. To each of these means and covariances, a weight of $w^{(i)} = 1$ is assigned, indicating one target per mean and covariance. These parameters are then used to form $\boldsymbol{\Theta}^{(i)} = \{w^{(i)}, \boldsymbol{m}_r^{(i)}, \boldsymbol{P}_{rr}^{(i)}\}$ as the set of parameters for each object. The total set of parameters is denoted $\boldsymbol{\Theta} = \boldsymbol{\Theta}^{(1)} \cup \boldsymbol{\Theta}^{(2)}$. Using the parameters $\boldsymbol{\Theta}$, the spatial Rényi entropy is computed from Eq. (24), and the result is illustrated in Fig. (7). Inspection of Fig. (7) shows that the interval in which the objects are interacting is obscured by other trends contained in the spatial Rényi entropy. As the entropy accounts for changes in the size of the individual distributions, as well as the translational motion of the individual distributions, the interval of interaction is not immediately apparent. To more clearly see the interval of interaction, the relative spatial Rényi entropy (not to be confused with the concept of relative entropy) is defined as

$$\Delta H_S^{(2)} = H_S^{(2)}(\boldsymbol{\Theta}) - \sum_{i=1}^{M} H_S^{(2)}(\boldsymbol{\Theta}^{(i)}) \,.$$

The relative spatial Rényi entropy can be viewed as the difference between the spatial entropy and the spatial entropy under the assumption of no interaction between the objects. The relative spatial Rényi entropy is illustrated in Fig. (8), from which the interval of interaction is immediately clear. Comparing Figs. (4) and (8), it is seen that the relative spatial Rényi entropy method provides excellent agreement to the more computationally expensive distribution sampling method.

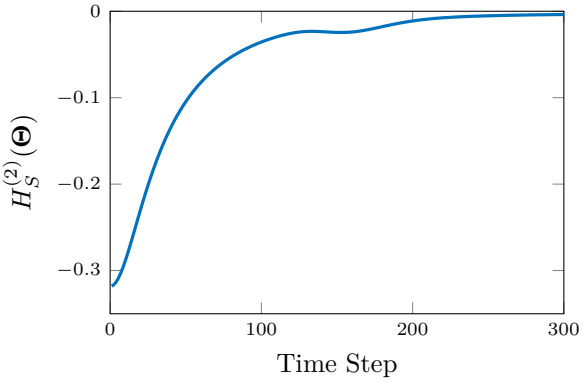

Figure 7. Application of the spatial Rényi entropy method to Example #1.

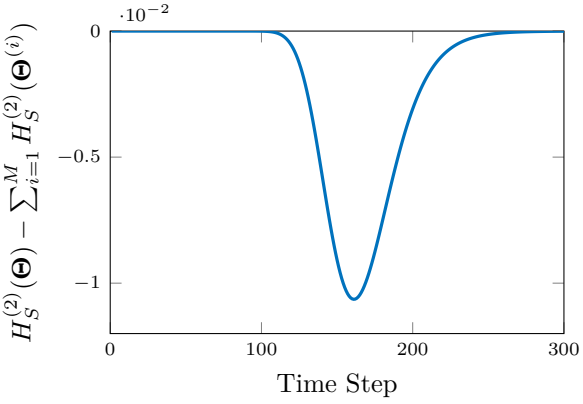

Figure 8. Application of the relative spatial Rényi entropy method to Example #1.

5.2. Example #2

To further investigate the methods developed in this paper, a three-object case is considered. The initial means of the three objects are given in Table 3, and the initial standard deviations are given in Table 4. As with the first example, the initial covariances are taken to be diagonal.

Table 3. Initial means for Example #2.

Object	x [m]	y [m]	$\dot{x}$ [mm/s]	$\dot{y}$ [mm/s]
1	0.0	50.0	0.0	0.0
2	0.0	−30.0	−68.9	52.8
3	30.0	0.0	−87.7	−17.3

Table 4. Initial standard deviations for Example #2.

Object	x [m]	y [m]	$\dot{x}$ [mm/s]	$\dot{y}$ [mm/s]
1	1.0	0.5	1.0	1.0
2	0.8	1.2	1.0	1.0
3	0.6	0.5	1.0	1.0

The time histories of the means and covariances for the three objects are computed for 200 time steps, and the resulting trajectories are illustrated in Fig. (9). The resulting covariances histories are plotted at 3-step intervals in Fig. (10).

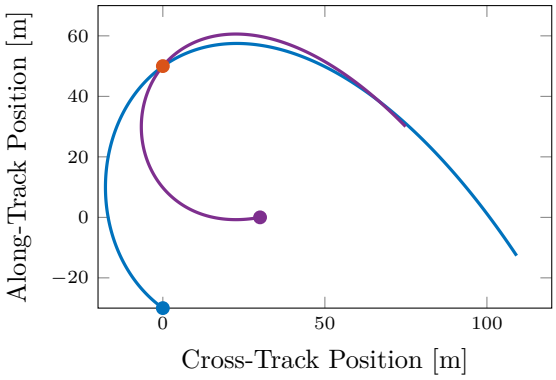

Figure 9. Mean trajectories for Example #2.

Based on the analysis of Example #1, the distribution sampling and relative spatial Rényi entropy methods are implemented to determine when the three objects are interacting. For each method, the position marginal distributions are obtained from the propagated means and covariances of each object.

At each time step, the position marginal distributions are sampled 1×10^6 times, the relative distance between the samples are computed, and the number of samples with a relative distance less than $c = 1$ m

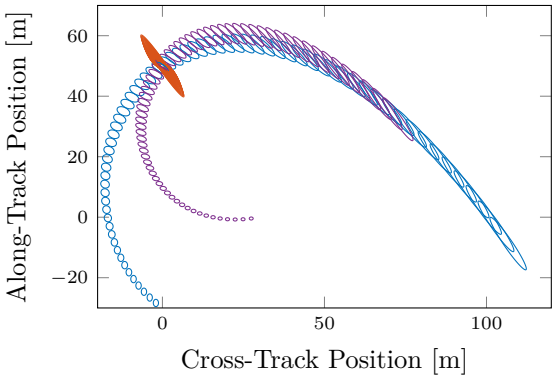

Figure 10. Uncertainty trajectories for Example #2.

are counted. The percentage of interacting samples is then determined, and this is shown in Fig. (11). Figure (11) illustrates the pairwise interactions between the three objects based on the pairwise percentage of interacting samples, as well as the total percentage of interacting samples obtained from considering all three objects. From Fig. (11), it is observed that the second and third objects begin interacting first, then the first and second objects interact, and finally the first and third objects interact.

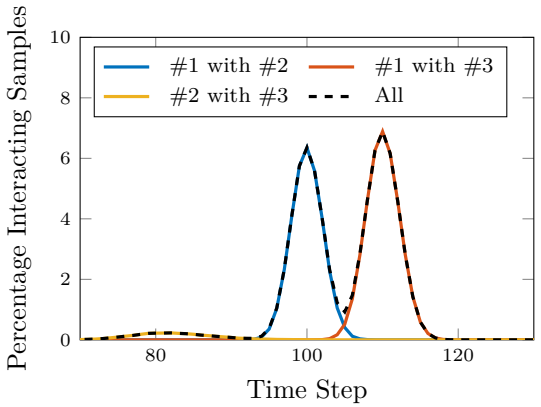

Figure 11. Application of the distribution sampling method to Example #2.

The position marginal means and covariances are obtained from the uncertainty propagation for each object, and these are denoted by $\boldsymbol{m}_r^{(i)}$ and $\boldsymbol{P}_{rr}^{(i)}$, respectively, for $i \in \{1, 2, 3\}$. To each position marginal distribution, a weight of $w^{(i)} = 1$ is assigned, indicating one target per mean and covariance. These parameters are the set of parameters for each object: $\boldsymbol{\Theta}^{(i)} = \{w^{(i)}, \boldsymbol{m}_r^{(i)}, \boldsymbol{P}_{rr}^{(i)}\}$. The total set of parameters is denoted $\boldsymbol{\Theta} = \boldsymbol{\Theta}^{(1)} \cup \boldsymbol{\Theta}^{(2)} \cup \boldsymbol{\Theta}^{(3)}$. The relative spatial Rényi entropy, as defined previously, is then computed from the individual and collective parameter sets, and it is illustrated in Fig. (12). The separate intervals of interaction are clear from Fig. (12), and remarkable agreement with

the distribution sampling method is again observed. One downside to the relative spatial Rényi entropy method is that the information on the pairwise interactions is lost. This could be recovered by considering pairwise combinations of the objects in a similar manner to that undertaken in the distribution sampling method, but this would come at an increase in computational burden.

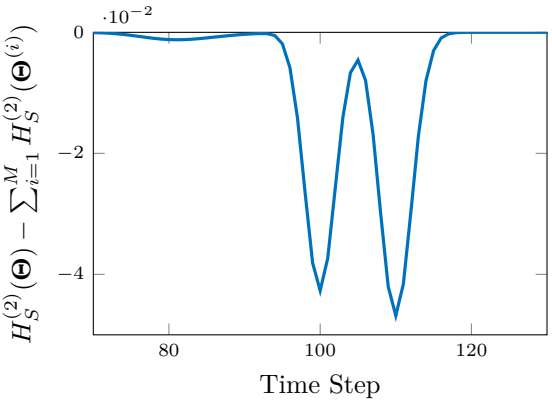

Figure 12. Application of the relative spatial Rényi entropy method to Example #2.

5.3. Example #3

The final example considered applies and compares the distribution sampling and relative spatial Rényi entropy methods to a non-Gaussian problem. The first object in this problem is taken to be described by a GM pdf, where the pdf is constructed for demonstration purposes only, and is not intended to represent a real-world scenario. It is assumed that the uncertainty of the first object is uniform in angular space, centered at zero in the cross-track direction, with a $\pm 20^\circ$ extent in the angular direction. Additionally, the range is taken to be 50.0 m from the origin, with a Gaussian uncertainty (1σ) of 2.0 m. The initial velocity is taken to be Gaussian with zero mean and standard deviations of 1.0 mm/s in both directions. Each component has equal weight in the GM representation of the pdf, and there are L components in the mixture.

The second object is taken to be described by a Gaussian distribution. The mean of the Gaussian distribution is $\boldsymbol{m} = [-25.0 \quad 25.0 \quad 11.6 \quad 50.0]$, with position units of meters and velocity units of mm/s. The covariance of the Gaussian distribution is diagonal with position standard deviations of 1.0 m and velocity standard deviations of 1.0 mm/s. Contours of the initial position distributions are illustrated in Fig. (13). The pdfs of the two objects are propagated for 350 time steps, and the trajectory of the mean of the distribution for the second object is shown in Fig. (13). As the mean of the first object has zero initial velocity and is in only the along-track direction, it remains stationary. The distribution, however, moves in time, but this is not illustrated.

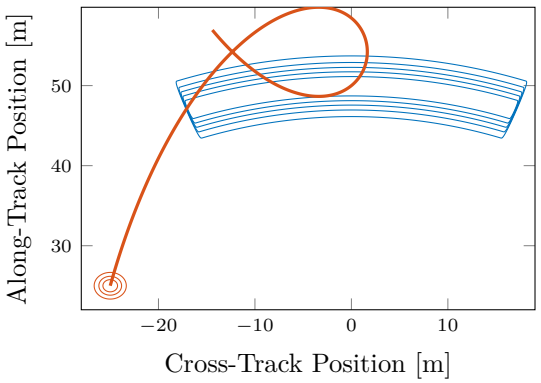

Figure 13. Trajectories for Example #3.

As with the preceding examples, the position marginal distributions are obtained for each object at each time step. These distributions are sampled 1×10^6 times, the relative distance between the samples are computed, and the number of samples with a relative distance less than $c = 1$ m are counted. The percentage of interacting samples is shown in Fig. (14). It is interesting to note that this example of a two-object collision exhibits two separate collisions between the objects. Examination of Fig. (13) shows that the second object "loops around" and passes nearby the first object twice.

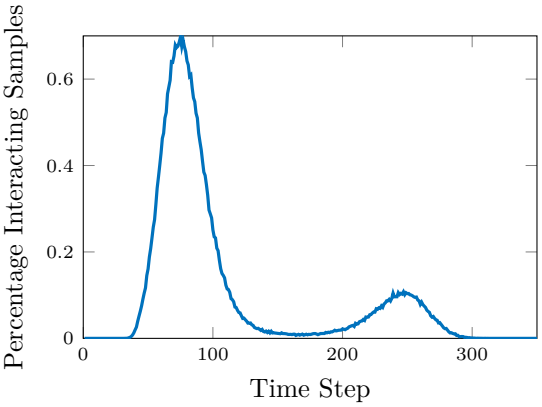

Figure 14. Application of the distribution sampling method to Example #3.

Given that the first object is represented by a GM pdf, the position marginal pdf is obtained by collecting the position elements of the mean and covariance while preserving the weights of the GM pdf. These parameters are denoted by $\boldsymbol{\Theta}^{(1)} = \{w^{(1,\ell)}, \boldsymbol{m}_r^{(1,\ell)}, \boldsymbol{P}_{rr}^{(1,\ell)}\}_{\ell=1}^{L}$, where $\sum_{\ell=1}^{L} w^{(1,\ell)} = 1$. The second object is represented by a Gaussian pdf, meaning that its parameters are simply given by

$\boldsymbol{\Theta}^{(2)} = \{w^{(2)}, \boldsymbol{m}_r^{(2)}, \boldsymbol{P}_{rr}^{(2)}\}$, where $w^{(2)} = 1$. The total set of parameters is denoted $\boldsymbol{\Theta} = \boldsymbol{\Theta}^{(1)} \cup \boldsymbol{\Theta}^{(2)}$. The relative spatial Rényi entropy is computed from the parameters and is shown in Fig. (12).

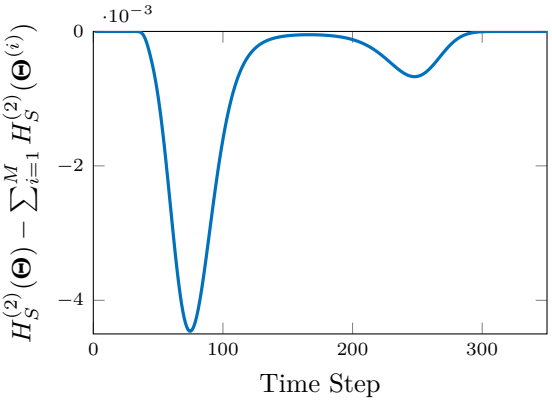

Figure 15. Application of the relative spatial Rényi entropy method to Example #3.

To provide a more direct and visceral comparison of the relative spatial Rényi entropy and the distribution sampling methods, the results from Figs. (14) and (15) are plotted together in Fig. (16), except that the relative spatial Rényi entropy is negated to possess the same sign as the percentage of interacting samples. From Fig. (16), it is seen that the two methods perform identically, modulo scaling, in determining the intervals during which the two objects are in close proximity and the distributions are interacting.

6. CONCLUSIONS

This paper addresses the problem of determining an interval during which two or more objects are in close proximity to one another, including the effects of uncertainty in the states of the objects, with the motivation of better informing the determination of the probability of collision between the objects. Conventional methods are discussed, and new approaches based on the information-theoretic concepts of information divergence and information entropy are developed. The methods are implemented and compared for two-object and three-object close approach scenarios, assuming both Gaussian and non-Gaussian state uncertainties. It is found that the information entropy approach that makes use of the multitarget Rényi entropy produces intervals of close proximity that are nearly identical to a more standard distribution sampling technique while naturally supporting close approaches between more than two objects and uncertainties that are represented by Gaussian mixture probability density functions.

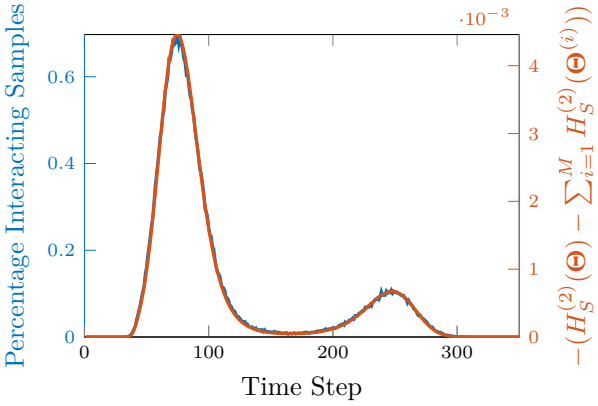

Figure 16. Comparison of the distribution sampling and relative spatial Rényi entropy methods for Example #3.

ACKNOWLEDGMENTS

This work was supported by a US Department of Education GAANN Fellowship (P200A150309). The authors would also like to acknowledge the many fruitful discussions with Mr. James McCabe of Missouri University of Science and Technology.

REFERENCES

1. Alfano, S. (2012). Toroidal path filter for orbital conjunction screening, *Celestial Mechanics and Dynamical Astronomy* **113**(3): 321–334.
2. Alspach, D. L. and Sorenson, H. W. (1972). Nonlinear Bayesian estimation using Gaussian sum approximations, *IEEE Transactions in Automatic Control* **AC-17**(4): 439–448.
3. Bar-Shalom, Y. and Fortmann, T. E. (1988). *Tracking and Data Association*, Vol. 179, Academic Press, chapter Tracking a Single Target in Clutter, pp. 157–190.
4. Baser, E. (2016). *Multi-target Multi-Bernoulli Tracking and Joint Multi-target Estimator*, Ph.D. thesis, McMaster University, Hamilton, Ontario, Canada.
5. Cichocki, A. and Amari, S. (2010). Families of alpha- beta- and gamma- divergences: Flexible and robust measures of similarities, *Entropy* **12**: 1532–1568.
6. Cover, T. M. and Thomas, J. A. (2006). *Elements of Information Theory*, 2nd edn, John Wiley & Sons.
7. Csiszár, I. (2006). *Stochastics: Information Theory*, Springer.
8. Dames, P. M. (2015). *Multi-Robot Active Information Gathering using Random Finite Sets*, Ph.D. thesis, University of Pennsylvania, Philadelphia, Pennsylvania.

9. Denenberg, E. and Gurfil, P. (2016). Improvements to time of closest approach calculation, *Journal of Guidance, Control, and Dynamics* **39**(9): 1967–1979.
10. Gim, D.-W. and Alfriend, K. T. (2003). State transition matrix of relative motion for the perturbed noncircular reference orbit, *Journal of Guidance, Control, and Dynamics* **26**(6): 956–971.
11. Goodman, I. R., Mahler, R. P. S. and Nguyen, H. T. (1997). *Mathematics of Data Fusion*, Kluwer Academic Publishers, Norwell, MA, USA.
12. Hartley, R. V. L. (1928). Transmission of information, *Bell Labs Technical Journal* **7**(3): 535–563.
13. Healy, L. M. (1995). Close conjunction detection on parallel computer, *Journal of Guidance, Control, and Dynamics* **18**(4): 824–829.
14. Hoots, F. R., Crawford, L. L. and Roehrich, R. L. (1984). An analytic method to determine future close approaches between satellites, *Celestial Mechanics* **33**(2): 143–158.
15. Inalhan, G., Tillerson, M. and How, J. P. (2002). Relative dynamics and control of spacecraft formations in eccentric orbits, *Journal of Guidance, Control, and Dynamics* **25**(1): 48–59.
16. Kirubarajan, T. and Bar-Shalom, Y. (2004). Probabilistic data association techniques for target tracking in clutter, *Proceedings of the IEEE* **92.3**: 536–557.
17. Kullback, S. and Leibler, R. A. (1951). On information and sufficiency, *The Annals of Mathematical Statistics* **22**(1): 79–86.
18. LeGrand, K. A. and DeMars, K. J. (2015). Bearings-only initial relative orbit determination, *Journal of Guidance, Control, and Dynamics* **38**(9): 1699–1713.
19. Mahalanobis, P. C. (1936). On the generalised distance in statistics, *Proceedings of the National Institute of Sciences of India* **2**(1): 49–55.
20. Mahler, R. (2003). Multitarget Bayes filtering via first-order multitarget moments, *Aerospace and Electronic Systems, IEEE Transactions on* **39**(4): 1152–1178.
21. Mahler, R. (2007a). PHD filters of higher order in target number, *IEEE Transactions on Aerospace and Electronic Systems* **43**(4): 1523–1543.
22. Mahler, R. P. S. (2007b). *Statistical Multisource-Multitarget Information Fusion*, Artech House, Inc.
23. Papoulis, A. (1984). *Probability, Random Variables, and Stochastic Processes*, 2nd edn, McGraw-Hill.
24. Phillips, M. R. (2012). *Spacecraft collision probability estimation for rendezvous and proximity operations*, M.S. thesis, Utah State University, Logan, Utah.
25. Rényi, A. (1961). On measures of entropy and information, *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability*, Vol. 1, pp. 547–561.
26. Rezaeian, M. and Vo, B.-N. (2009). The entropy of random finite sets, *Proceedings of the International Symposium on Information Theory*.
27. Shannon, C. E. and Weaver, W. (2002). *The Mathematical Theory of Communication*, University of Illinois Press.
28. Sorenson, H. W. and Alspach, D. L. (1971). Recursive Bayesian estimation using Gaussian sums, *Automatica* **7**(4): 465–479.
29. Stringer, M. T., Newman, B., Lovell, T. A. and Omran, A. (2013). Second order nonlinear initial value solution for relative motion using Volterra theory, *Proceedings of the 23rd AAS/AIAA Space Flight Mechanics Conference*, number AAS 13-469, pp. 2133–2149.
30. Vo, B.-N. and Ma, W.-K. (2006). The Gaussian mixture probability hypothesis density filter, *IEEE Transactions on Signal Processing* pp. 4091–4104.
31. Vo, B.-T., Vo, B.-N. and Canton, A. (2006). The cardinalized probability hypothesis density filter for linear Gaussian multi-target models, *40th Annual Conference on Information Sciences and Systems*, pp. 681–686.
32. Vo, B.-T., Vo, B.-N. and Cantoni, A. (2007). Analytic implementations of the cardinalized probability hypothesis density filter, *IEEE Transactions on Signal Processing* **55**(7): 3553–3567.