Issues with Satellite Collision Risk Aggregation

Hunter Morris, Matthew Hejduk
The Aerospace Corporation
Lauri Newman
NASA Headquarters

ABSTRACT

With the observed growth in the orbital population due to the deployment of large satellite constellations, there is a strong need for a credible measure of long-term satellite collision risk. One widespread and intuitive tool for this task is the so-called "aggregate P_c", which is based on the P_c or "probability of collision" metric widely used in conjunction assessment. This approach, while possessing many attractive qualities, has one major drawback: it may not accurately reflect the expected long-term frequency of collision. This is because the derivation of the P_c itself relies on a hidden assumption that is questionable from a frequentist perspective. The P_c may be justifiable on Bayesian grounds, but a Bayesian probability can be potentially misleading in a long-term risk assessment context, where the natural interpretation is in terms of long-term frequencies. In light of this problem, several alternative methods of long-term risk assessment are explored, some of which show promise but none of which is entirely satisfactory for immediate deployment.

1. INTRODUCTION

For the purposes of setting orbital safety policy, as well as for mission planning, estimating the long-term risk of collision faced or posed by a satellite or constellation is essential. The most common method used historically to perform this analysis utilizes flux-based models of collision risk, such as the Orbital Debris Engineering Model (ORDEM) maintained by the NASA Orbital Debris Program Office (ODPO), the Meteoroid and Space Debris Terrestrial Environment Reference (MASTER) model maintained by ESA, or one of their derivatives. These models allow one to input an initial orbit, the spacecraft size, period of time, and minimum conjuncting object size; the model will then calculate the likelihood of at least one collision over that period with an object of that size or larger. Such models are rigorously constructed and yield a frequentist probability as their product, but they cannot account for the effects of active collision avoidance (CA) operations. One could presume that these active CA operations entirely eliminate the collision risk with trackable objects over the modeled period; but in fact modern CA practices do not eliminate this risk, only reduce it. With large constellations performing active CA operations, a small residual risk for individual spacecraft can quickly add up to a large ensemble risk of at least one collision for the entire constellation. In order to account for this reality, many conjunction assessment researchers and practitioners have sought a different method of aggregate collision-risk assessment.

One family of methods for attempting this analysis that has gained traction in recent years is based on the "aggregate collision probability," or "aggregate P_c".[1] This quantity is computed as follows. Let

$$\{P_{c,i}\}_{i=1}^{N}$$

be a sequence of collision probabilities, each corresponding to a separate conjunction. Depending on the application, these probabilities may be obtained by simulation or from historical data, and they may involve one primary satellite or many. Assuming that collisions are mutually independent events, the probability that none of the conjunctions will result in a collision is given by the aggregate survival probability [4],[7]:

$$S_{agg} = \prod_{i=1}^{N} \left(1 - P_{c,i}\right).$$

[1] Some references call this the *cumulative* collision probability.

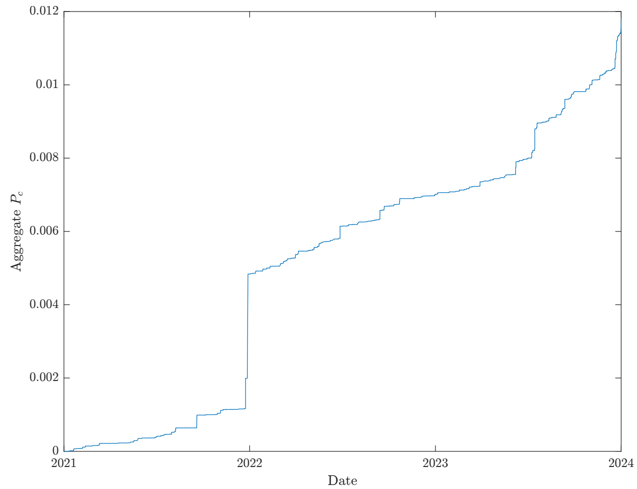

Fig. 1: Typical evolution of aggregate P_c over time, based on historical CDM data from the Hubble Space Telescope.

Thus the probability that *at least one* of these conjunctions will result in a collision is the complement of this probability, or

$$P_{c,agg} = 1 - \prod_{i=1}^{N} \left(1 - P_{c,i}\right). \tag{1}$$

This quantity is called the *aggregate collision probability* or *aggregate P_c*.

For a concrete example of this kind of computation, see Figure 1. This figure was generated using historical conjunction data message (CDM) data from the Hubble Space Telescope, spanning the time period from 2021-2024. Since each conjunction may have multiple CDMs associated with it, it is necessary to choose a representative CDM for each conjunction. For this exercise, the last available CDM before the time of closest approach (TCA) was used to compute the P_c for each conjunction (an appropriate choice for this spacecraft, since it is not maneuverable).

In the general case, this computation can be adjusted to take account of collision avoidance maneuvers. Let T_α be the threshold value for P_c, above which a collision avoidance maneuver will be carried out. In the sequence $\{P_{c,i}\}_{i=1}^{N}$, any value above the maneuver threshold can be replaced with some lower value, $P_{c,i}^{rem}$, that represents the risk remaining after a collision avoidance maneuver has been performed. Then a similar calculation as that in (1) can be used to calculate the *residual risk*:

$$P_{agg}^{rem} = 1 - \prod_{i=1}^{N} \left(1 - P_{c,i}^{rem}\right). \tag{2}$$

From this one can also calculate the *fractional risk reduction* (FRR):

$$FRR = 1 - \frac{P_{agg}^{rem}}{P_{agg}}. \tag{3}$$

The FRR represents the proportion of the aggregate risk of collision that has been eliminated by the chosen maneuver strategy.

An example of the residual (or "mitigated") risk calculation just described is shown in Figure 2. The data here are again taken from the Hubble Space Telescope, over the same time period as in Figure 1. For this exercise we imagine that the Hubble Space Telescope is maneuverable. The latest possible time at which a satellite operator can safely commit to performing a risk-mitigation maneuver, should one be deemed necessary, is known as the maneuver commitment point (MCP). For this analysis an MCP of 1 day before TCA was assumed, and consequently the latest CDM available at 1 day before TCA was chosen to represent each conjunction. For comparison's sake, two aggregate P_c curves have been plotted in Figure 2: an unmitigated version where none of the individual P_c values have been replaced, and a

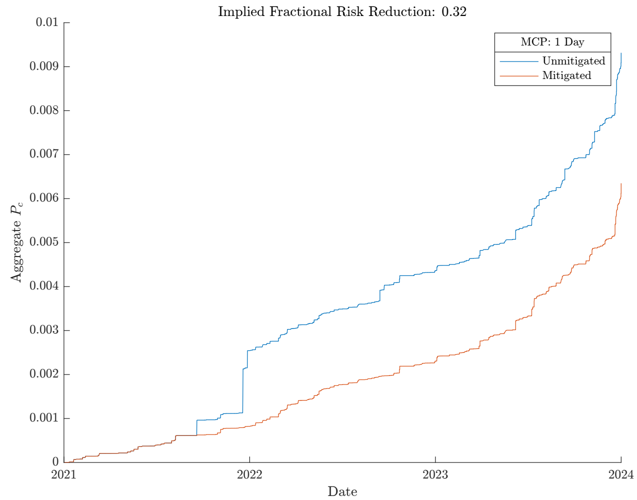

Fig. 2: Comparing the aggregate P_c in the absence of maneuver with the aggregate P_c calculated by replacing all P_c values above 10^{-4} with 3.1×10^{-6}. This replacement simulates the NASA CARA risk mitigation strategy.

mitigated version where all P_c values above 10^{-4} have been replaced with 3.1×10^{-6} (this is intended to model the NASA CARA risk mitigation strategy). The implied fractional risk reduction is also shown.

This method of long-term risk assessment has some attractive features, as described by [4]. For one, it does not rely on a physical model, e.g. a flux model, but relies entirely on empirically observed collision probabilities. Another attractive feature of this method is that it does not assume that all of the risk from a conjunction is eliminated by a maneuver, but rather allows some risk to remain. At the same time, there is a hidden pitfall contained in this method which, while arguably not fatal, seriously curbs its usefulness.

As shown in Figure 1, at the end of the period 2021-2024 the aggregate P_c reaches a value of about .012. But how should we interpret this quantity? Probably the most natural interpretation is in terms of long-run frequencies: if one could observe an infinite number of random experiments, where in each experiment an HST-like object is allowed to orbit in the same orbital environment over the same period, then one should expect that a collision would occur in about 12 out of every 1,000 of these experiments. Unfortunately, as will be shown, this interpretation is at best questionable, and at worst actively misleading. The fundamental issue is with the individual "collision probabilities" that underlie the aggregate P_c. As will be shown in the next section, the standard method for calculating collision probabilities involves the (heretofore implicit) assumption that, prior to analysis, all close-approach vectors are equally likely.[2] In a certain sense, this assumption is extremely sensible; it seems that, before a conjunction is analyzed, we have no principled reason to favor one close approach vector over another. But this reasoning hints at the issue: although our state of *knowledge* may be such that we have no reason to favor one close approach vector over another, that does not mean that all close approach vectors are equally *frequent*. The questionable validity of the assumption of a uniform prior means that, while the standard collision probability may be in some sense justifiable, it may not give an accurate idea of the long-term *frequency* of collision. This does not necessarily imply that P_c should be abandoned operationally, but it does require that long-term risk analysis carried out on the basis of aggregate P_c be handled with care, so that important decisions are not made on the basis of a faulty interpretation.

To lay out in detail the argumentation supporting the above claims, the next section is dedicated to reviewing the standard P_c calculation and showing that it relies on an implicit assumption that brings its interpretation as a long-term frequency into question. After this follows a review of possible alternatives to aggregate P_c for long-term risk analysis, which includes certain recommendations. The final section offers concluding remarks and recommended items for future work.

[2] When specified in Cartesian coordinates.

2. THE COLLISION PROBABILITY

2.1 Review of Collision Probability

If we knew exactly where two satellites were going to be at the time of their closest approach, there would be no need to calculate collision probabilities. As it is, the prediction of a satellite's state (position and velocity) at a future time is affected by many unknown factors, including but not limited to measurement error and unpredictable space weather. In order to take the uncertainty arising from these factors into account, it has become standard to model the state of a satellite at a particular time as a random variable. Specifically, the state of satellite i at time t, denoted X_i^t, is often assumed to be distributed normally, with mean state $\widehat{x}_i^t$, and covariance matrix Σ_i^t, i.e.

$$X_i^t \sim N\left(\widehat{x}_i^t, \Sigma_i^t\right). \tag{4}$$

The mean state, $\widehat{x}_i^t$, represents the "best-guess" prediction of the state at time t. The covariance matrix is intended as a summary of the uncertainties induced by measurement error, space weather forecasting uncertainties, etc., and determines how "spread out" the state of the satellite is around its mean value. The following development presents the typical derivation of the probability of collision, or P_c, taking (4) as given. In a subsequent section, we will attempt use a statistical model of the measurement errors to derive P_c from first principles, and this process will reveal an issue that undermines the interpretation of P_c as a long-term frequency.

For the present, however, assume that a close approach has been detected between two satellites, which we refer to as the primary and the secondary, and which will be denoted with the indices $i = 1,2$ respectively. Their *relative* state at TCA will be given by

$$X^{TCA} = X_2^{TCA} - X_1^{TCA}.$$

If we assume that the state distributions of the primary and secondary are independent of each other, then

$$X^{TCA} \sim N\left(\widehat{x}^{TCA}, \Sigma^{TCA}\right),$$

where $\widehat{x}^{TCA} = \widehat{x}_2^{TCA} - \widehat{x}_1^{TCA}$ and $\Sigma_{TCA} = \Sigma_2^{TCA} + \Sigma_1^{TCA}$.

Because most conjunctions involve high relative velocities and short encounter times, it is often a good approximation to assume that the relative trajectory of two satellites is linear near TCA (i.e. that their relative velocity vector is constant). Furthermore, the components of Σ^{TCA} corresponding to the velocity are often relatively small, so that we can treat the velocity as known. Under these two conditions, the point of closest approach between the two satellites will be constrained to a fixed plane which is perpendicular to the known relative velocity vector; we call this plane the "conjunction plane". To determine the probability of collision, then, we need only consider the distribution of the relative position vector at TCA *projected* onto the conjunction plane [6]. If we denote this projected, two-dimensional relative position or "miss" vector by Y, then we have

$$Y \sim N(\widehat{y}, D), \tag{5}$$

where $\widehat{y}$ and D are projections of $\widehat{x}^{TCA}$ and Σ^{TCA} onto the conjunction plane. After an appropriate rotation of the coordinate system, it can be assumed that $D = \text{diag}\left(d_1^2, d_2^2\right)$. The collision probability is then calculated as the probability that Y will come within a certain radius of the origin, called the hard-body radius (HBR), i.e.

$$\begin{aligned} P_c &= \mathbb{P}(\|Y\| \leq \text{HBR}) \\ &= \frac{1}{2\pi d_1 d_2} \iint_{\|y\| \leq \text{HBR}} e^{-\frac{1}{2}\left(\frac{(y_1-\widehat{y})^2}{d_1^2} + \frac{(y_2-\widehat{y}_2)^2}{d_2^2}\right)} dy_1 dy_2. \end{aligned} \tag{6}$$

There are alternate methods for calculating the collision probability that do not rely on this two-dimensional approximation (see [8]). For the present argument, however, we restrict our attention to the two-dimensional case.

One aspect of the preceding development requires comment. In the above derivation, the "mean" or "estimated" states (e.g. $\widehat{x}_i^{t_0}$, $\widehat{y}$, etc.) are treated as fixed parameters, about which the true states vary randomly. In reality, as will be shown in the next section, these estimated states are themselves realizations of random variables (which will be represented by the corresponding capitalized symbols, $\widehat{X}_i^{t_0}$, $\widehat{Y}$, etc.). Expression (5) is thus a statement about the distribution of Y (the true miss vector in the conjunction plane), *given* that the random variable $\widehat{Y}$ takes on a particular value $\widehat{y}$ (i.e. given that the estimated miss vector is equal to its observed value). It is thus possible to rewrite (5) in a more statistically rigorous way, as

$$Y \,\big|\, \widehat{Y} = \widehat{y} \sim N\left(\widehat{y}, D\right). \tag{7}$$

The notation on the left hand side of the $\sim$ sign is read "*Y conditional on $\widehat{Y} = \widehat{y}$*". It follows that P_c is a *conditional* probability of collision; in particular, it is the probability of collision, *given* a particular value of the *estimated* miss vector.

It is fair to wonder whether this level of conceptual and notational refinement is actually necessary. As we turn in the next section to an attempted derivation of P_c from first principles, the need for this seemingly excessive level of precision will become clear. We begin with the orbit determination process, focusing in particular on the effects of sensor measurement error.

2.2 Orbit Determination and the Covariance Matrix

The standard method of orbit determination involves using measurements/observations of a satellite (look-angles, ranges, range-rates, etc.) to estimate the state of that satellite at a particular epoch time. Assume we have a K-dimensional vector of measurements

$$M = \begin{bmatrix} M_1 \\ \vdots \\ M_K \end{bmatrix}.$$

In addition, we need a function $h : \mathbb{R}^6 \to \mathbb{R}^K$, called the *measurement model*, which maps possible epoch states into the measurement vector that would be expected in the absence of measurement error. With the measurement model in hand, we can write

$$M = h(X_i^{t_0}) + \varepsilon,$$

where ε is a K-dimensional vector of sensor measurement errors. Furthermore, it is assumed that

$$\varepsilon \sim N\left(0, S\right),$$

where the covariance matrix of the measurement errors, S, is known. Operationally, S is obtained from the sensor calibration process. In addition, ε is assumed to be independent of $X_i^{t_0}$.[3] Intuitively, the distribution of ε describes the relative frequency with which measurement errors of a certain size can be expected to occur.

If we linearize the measurement model around an appropriate reference state, we can estimate the epoch state of the satellite using weighted-least-squares (WLS), with $W = S^{-1}$ as the weighting matrix; additionally, it can be shown that the WLS estimate of the epoch state, denoted $\hat{X}_i^{t_0}$, is approximately

$$\widehat{X}_i^{t_0} \approx X_i^{t_0} + (H^T W H)^{-1} H^T W \varepsilon,$$

where H is the Jacobian of the measurement model evaluated at the reference state.[4] Notice that $\widehat{X}_i^{t_0}$, which will henceforth be referred to as the *estimated* epoch state of satellite i, is itself a random variable, since it depends on the random sensor measurement errors ε. However, it also depends on the unknown true epoch state, $X_i^{t_0}$, which is itself a random variable with unknown distribution. Despite this, if we *condition* on $X_i^{t_0} = x_i^{t_0}$,[5] then, because of the independence of $X_i^{t_0}$ and ε, we obtain

$$\widehat{X}_i^{t_0} \,\big|\, X_i^{t_0} = x_i^{t_0} \sim N\left(x_i^{t_0}, \Sigma_i^{t_0}\right), \tag{8}$$

[3] This may seem like an unrealistic assumption at first glance. However, in order to acquire a satellite for observation, we have to have a reasonable idea of where it is already. Even if this independence assumption does not hold for *all possible states*, it is probably approximately true in the narrow range of states that the satellite could plausibly occupy. In addition, if S is known, it cannot be dependent on an unknown quantity.

[4] For details see [9].

[5] This can be thought of as "holding the random variable $X_i^{t_0}$ fixed at the value $x_i^{t_0}$".

where $\Sigma_i^{t_0} = (H^T W H)^{-1}$.

To illustrate the significance of the above symbolic expression, imagine that we could "turn back the clock", and perform the same OD an infinite number of times, on the same satellite, following the same trajectory (i.e. conditioning on $X_i^{t_0} = x_i^{t_0}$), making measurements at the same times and with the same instruments, allowing only the measurement errors at the various sensors to vary randomly, according to their distributions. Then the OD process would produce an infinitely large sample of epoch state estimates, whose mean would be the true, unknown epoch state $x_i^{t_0}$, and whose covariance would be $\Sigma_i^{t_0}$. It bears repeating that the covariance matrix $\Sigma_i^{t_0}$ describes the distribution of the epoch state *estimates* and not the distribution of the true epoch state itself! This observation is significant because, in the previous section, we assumed that the covariance matrix did indeed describe the distribution of the true epoch state (see expression (4)), and this assumption was crucial to the derivation of P_c. The consequences of this observation will become clearer as we proceed.

2.3 Deriving P_c with Bayes' Theorem

Returning to our derivation, once the epoch states of two conjuncting satellites have been estimated, these estimated states can be propagated forward to TCA, along with their respective covariance matrices. For the sake of exposition, assume that the propagation from epoch to TCA contains no error; that is, assume the only source of error in the problem is the measurement error at the sensors. We then have the following conditional distribution for the propagated state estimate:

$$\widehat{X}_i^{TCA} \mid X_i^{TCA} = x_i^{TCA} \sim N\left(x_i^{TCA}, \Sigma_i^{TCA}\right), \tag{9}$$

where $\widehat{X}_i^{TCA}$ represents the state estimate of satellite i, propagated from epoch to TCA, Σ_i^{TCA} represents the corresponding propagated covariance, and X_i^{TCA} represents the *true* state of satellite i at TCA, which is fixed at the value x_i^{TCA}. Similarly to expression (8), expression (9) states that the propagated *estimate* of the satellite's state at TCA, given a particular true TCA state, is a normal random variable, with mean at the unknown true state and with a covariance matrix that represents the propagated effect of sensor measurement error.

If we further assume that the sensor measurement errors and epoch states of the primary and secondary are mutually independent, then the estimated *relative* state at TCA will have the following distribution, conditional on a particular true relative state $X^{TCA} = x^{TCA}$:

$$\widehat{X}^{TCA} \mid X^{TCA} = x^{TCA} \sim N\left(x^{TCA}, \Sigma^{TCA}\right)$$

where $\Sigma^{TCA} = \Sigma_i^{TCA} + \Sigma_i^{TCA}$.

Finally, under the assumptions of section 2.1 (linear relative trajectories, known velocities), we can derive the (conditional) distribution of the *estimated* miss vector projected into the conjunction plane:

$$\widehat{Y} \mid Y = y \sim N(y, D). \tag{10}$$

This is deceptively similar to (7), and it may seem that one can simply integrate the PDF corresponding to (10) over the hard-body circle to obtain the probability of collision. There are two difficulties with this, however. The first difficulty is that the distribution in (10) has an unknown mean, since its mean is the true 2D miss vector. Of course, if the true 2D miss vector were known, then there would be no need to perform conjunction risk assessment in the first place. Admittedly, this difficulty could be patched somewhat by replacing y with its estimate, $\widehat{y}$, which is a quantity we do observe. But this leads us to the second, far more serious difficulty.

Recall that (10) describes the distribution of the 2D miss vector *estimates* that would be observed in repeated orbit determinations, given the *true* 2D miss vector. In calculating the probability of collision, however, we are interested in the distribution of *true* 2D miss vectors, given the *estimated* 2D miss vector. If we were to integrate the PDF corresponding to (10) over the hard-body circle, the result would not be the probability of collision, i.e. the probability that the *true* miss vector would pass within the HBR, but rather the probability that the *estimated* 2D miss vector would pass within the HBR (were we to repeat the orbit determination and propagation experiment)[5]. Thus, in our attempt to derive the probability of collision from first principles, we have reached an impasse.

The authors of [5] show that this impasse can be overcome by using Bayes' theorem. In its continuous form, Bayes' theorem states

$$f_{Y \mid \widehat{Y}=\widehat{y}}(y \mid \widehat{y}) = \frac{f_{\widehat{Y} \mid Y=y}(\widehat{y} \mid y) f_Y(y)}{\int f_{\widehat{Y} \mid Y=y}(\widehat{y} \mid y) f_Y(y) dy}, \tag{11}$$

where $f_{Y \mid \hat{Y}=\hat{y}}$ denotes the PDF of the random variable $Y \mid \hat{Y} = \hat{y}$, and so on for the other subscripts. Notice that Bayes' theorem allows us to derive the PDF of $Y \mid \widehat{Y} = \widehat{y}$ (the distribution of the *true* miss, conditional on the *estimated* miss) from the PDF of $\widehat{Y} \mid Y = y$ (which is what we have derived so far). We are still missing a necessary piece of information to apply this theorem, however, and that is the function $f_Y(y)$, the PDF of the "prior" distribution for the true miss vector Y. This prior distribution is the distribution of the true 2D miss vector, *without* taking the estimate $\hat{Y}$ into account, and can be thought of as the distribution of miss vectors that obtains "prior" to the conjunction analysis. Bayes' rule allows us to update this prior distribution in light of the observed data ($\widehat{y}$ and D), and the updated distribution is called the "posterior". The authors of [5] point out that assuming a "uniform" prior, where $f_Y(y) \equiv c$,[6] allows us to recover the typical P_c calculation, since

$$\begin{aligned}
f_{Y \mid \widehat{Y}=\widehat{y}}(y \mid \widehat{y}) &= \frac{f_{\widehat{Y} \mid Y=y}(\widehat{y} \mid y) f_Y(y)}{\int f_{\widehat{Y} \mid Y=y}(\widehat{y} \mid y) f_Y(y) dy} \\
&= \frac{c \cdot f_{\widehat{Y} \mid Y=y}(\widehat{y} \mid y)}{c \cdot \int f_{\widehat{Y} \mid Y=y}(\widehat{y} \mid y) dy} \\
&= \frac{f_{\widehat{Y} \mid Y=y}(\widehat{y} \mid y)}{\int f_{\widehat{Y} \mid Y=y}(\widehat{y} \mid y) dy} \\
&= f_{\widehat{Y} \mid Y=y}(\widehat{y} \mid y) \\
&= \frac{1}{2\pi d_1 d_2} e^{-\frac{1}{2}\left(\frac{(y_1-\widehat{y})^2}{d_1^2} + \frac{(y_2-\widehat{y}_2)^2}{d_2^2}\right)}.
\end{aligned}$$

This implies (7), and from this we can calculate the collision probability in the usual way (see equation (6)).

2.4 The Uniform Prior

We have seen that P_c can indeed be derived from a statistical model of sensor measurement error, but to do so it was necessary to assume a uniform prior distribution on the 2D miss vector in the conjunction plane. Until [5], this assumption remained implicit. Now that it has been laid bare, we must ask ourselves if and how it can be justified.

One way to justify this uniform prior is to argue that, before an estimate of the 2D miss vector is made, we have no reason to favor one miss vector over another. It could be anything, and we do not want to artificially favor some possibilities over others by assigning a non-uniform prior. Another way of putting this is that the uniform prior is "uninformative." Now this is a perfectly reasonable justification, and there is a sizable segment of the statistical literature dedicated studying the properties of such "uninformative priors".[7] The issue is that this justification is fundamentally a *Bayesian* one, and under the Bayesian paradigm, probabilities are seen as describing, not the long-term frequencies of certain events, but our state of "knowledge" or "belief" about these events. This is a problem if we wish to interpret the collision probability as a statement about the long-term *frequency* of collision. For the P_c to accurately represent the long-term frequency of collision, it must be the case that a uniform prior accurately represents the relative *frequency* with which different 2D miss vectors occur over a large number of conjunctions. But it is not obvious that the uniform prior does this.

One questionable aspect of the uniform prior in this respect is the fact that it is an "improper" prior, i.e. it is not technically a probability distribution at all. This is because a "PDF" with an arbitrary non-zero constant value will have an infinite integral over the plane; in particular, it assigns infinite probability to miss vectors *outside* of the HBR

[6] Notice this PDF is uniform in *Cartesian* coordinates.

[7] See [2] for an example.

[5], expressing a prior certainty that no collision will occur. This fact does not cause any issues with the *calculation* of the P_c, since the posterior distribution that results from applying (11) is a proper probability distribution with integral of one. Rather, the fact that a uniform prior over the plane is improper calls into question whether it is a plausible model of the relative frequency of different 2D miss vectors, and thus whether P_c can be interpreted as a long-term frequency. Now this issue of "impropriety" can be remediated by "truncating" the prior distribution outside of an arbitrarily large but bounded region of the plane and renormalizing the uniform PDF to the area of the chosen region; this will cause an arbitrarily small deviation in the P_c calculation if we choose a large enough region of the plane. This procedure does introduce a certain level of arbitrariness to the problem, however, since there are many possible choices for where the truncation should occur.

At the very least, such considerations should lead us to *question* whether the uniform prior accurately portrays the relative long-term frequencies of different miss vectors, and therefore whether P_c gives an accurate idea of the long-term frequency of collision. This applies *a fortiori* to long-term risk assessment in the form of aggregate P_c calculations. But if aggregate P_c cannot be interpreted as a long-term frequency, but must instead be understood as an aggregate "strength of belief", then any policy or mission-planning decisions made on the basis of such calculations is rendered somewhat arbitrary. This is especially the case since, if we are willing to choose a prior simply because it is "uninformative," there are many possible choices, each of which would lead to a different calculation of the probability of collision (see [2]).

It is thus worth asking if there are alternative methods of performing long-term risk assessment that have more plausible connections with long-term frequencies. The next section considers some possible proposals.

3. ALTERNATIVES

3.1 Estimating the Prior from Data

Since the uniform prior is questionable as a description of the long-term frequencies of different 2D miss vectors, it is natural to ask whether the prior can be improved. One possible way of doing this, suggested in [3], is to *estimate* the prior from past conjunction data. Now, the first thing one must do in such an exercise is to decide the functional form of the distribution to be fitted. The authors of [3] choose a gamma distribution in the miss distance (they assume the distribution in the miss angle is uniform, and independent of the miss distance). A gamma distribution in miss distance assigns more probability mass near the center of the conjunction plane then the "uniform" distribution above,[8] giving a higher prior weight to a collision. This would make any collision probabilities based on this prior more conservative than P_c as presently formulated (i.e., with a uniform prior). However, the gamma distribution is chosen in [3] mainly for computational convenience; it is a conjugate prior for this problem, which means that when one employs the Bayes' rule calculation in (11), the posterior distribution, from which we calculate the collision probability, is guaranteed also to be a gamma distribution in the miss distance. Despite this nice property, we still have to wonder how well this distribution can reflect the long-term frequency of 2D miss vectors, even with appropriately chosen parameters, and whether a different distributional form might be more appropriate. One might question why it is necessary to have a convincing *a priori* justification for a particular distributional form if the model can generate an acceptable fit to the data. One reason for desiring such an a priori justification is that it would give us more confidence in the model in ranges of the miss distance where we have little data. When assessing the long-term risk of collision, we are most interested in accurately modeling the left tail of the miss-distance distribution. However, most of the data used to fit the model comes from conjunctions with relatively large (estimated) miss distances. Thus, we are effectively attempting to reconstruct the properties of the tail of the distribution from the properties of the main part of the distribution. If the form of the distribution is not correctly chosen, then the dependence between the tail and main part of the distribution will be misspecified, leading to a model that is inaccurate in the range where it is most necessary for it to be accurate.

Still, it is worthwhile to see if a good fit to the data can be obtained with a gamma prior in the miss distance. The authors of the present paper attempted to fit a gamma distribution in the miss distance to the HST CDM data that has been mentioned in previous sections, utilizing the method proposed in [3]. This exercise yielded a distribution with such a poor fit to the data that we suspect a mistake in our implementation of the estimation method, and thus we do not display the results here. It is worth noting, however, that the authors of [3] also had difficulties with fitting their model, which they tentatively ascribe to numerical instability in the estimation. Evidently, more work is needed to make estimating the prior a viable improvement to P_c.

[8] Again, uniform in Cartesian coordinates.

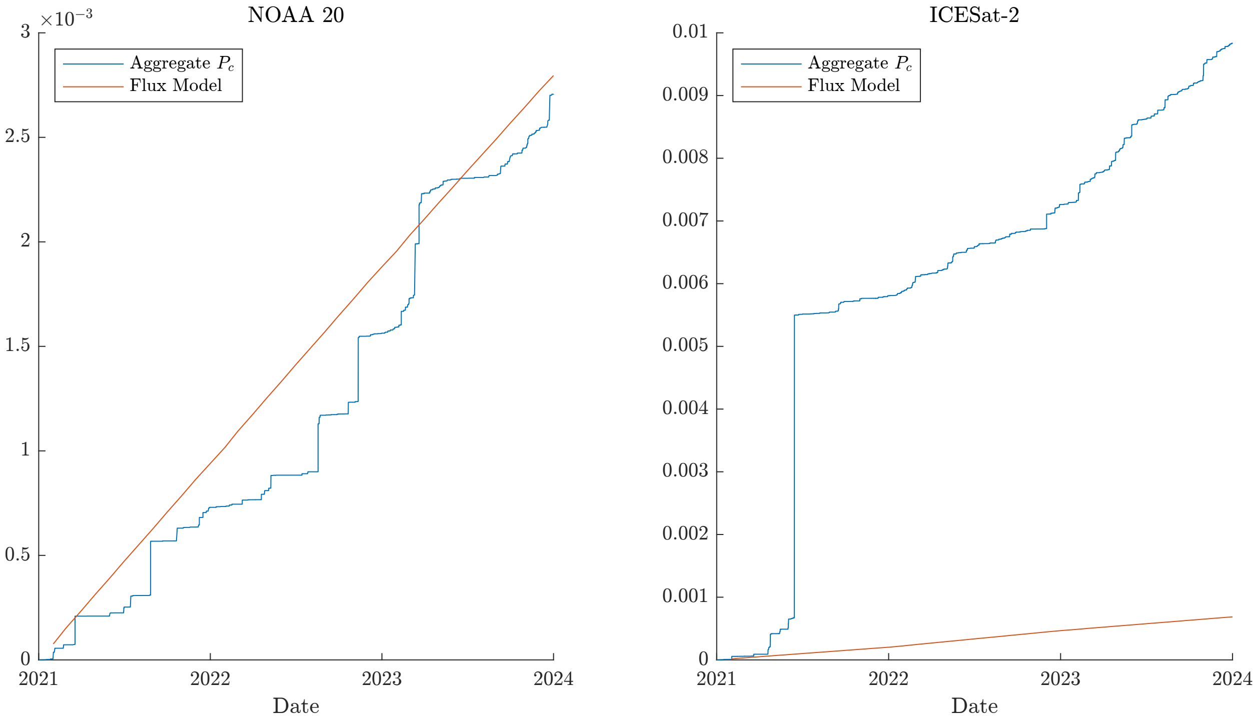

Fig. 3: Comparing the results from the aggregate P_c calculation with the aggregate probability of collision derived from the DAS flux model.

3.2 Flux Models

The classic method for estimating the long-term risk of collision to a satellite in a particular orbit regime is the debris flux model, mentioned briefly in the beginning of this paper. This method employs an analogy to gas dynamics to model the orbital debris environment, and it is explicitly designed to yield an estimate of the long-term *frequency* of collision that an object of a particular size will encounter in a particular orbit regime. Although this is typically phrased in terms of the *number* of collisions expected over a period of time, this can be used to derive the probability of *at least one* collision occurring over a period of time, in a similar manner to aggregate P_c. Still, one might question whether this modeling approach accurately reflects the orbital environment.

Consider Figure 3. This shows the aggregate P_c of the NOAA 20 and ICESat-2 satellites over the period 2021-2024, compared to the output of the ODPO Debris Assessment Software (DAS) flux model for two satellites in similar orbits. Specifically, NOAA 20, at an altitude of 830 km and inclination of 97.8°, was compared against a hypothetical satellite at an altitude of 850 km and inclination of 97°, while ICESat-2, at an altitude of 480 km and inclination of 92°, was compared with a hypothetical satellite at an altitude of 520 km and inclination of 97°. In the first case, as can be seen, the aggregate P_c estimate is about the same as the estimate derived from the flux model, but in the second case the aggregate P_c is much higher than the flux model estimate. The divergence between the two methods in this case suggests that either aggregate P_c or the flux model is an inaccurate measure of the true collision risk, in terms of long-term frequencies. In fact, since the aggregate P_c is based on a uniform prior, which assigns essentially zero prior probability to a collision, one would expect that, if anything, the aggregate P_c should underestimate the true long-term frequency of collision. Thus, if the flux model were accurate, one would expect the aggregate collision probability calculated from the flux model to *exceed* the aggregate P_c, which is the opposite of what we observe.

Another shortcoming of the flux modeling approach is that it is not capable of taking collision avoidance maneuvers into account. Thus, even if one possessed an accurate estimate of the long-term frequency of collision in the absence of maneuver, it would still be necessary to estimate the fraction of collisions that are expected to be avoided by implementing a particular maneuver strategy. One proposal for evaluating this proportion, introduced by [10], is discussed in the following section.

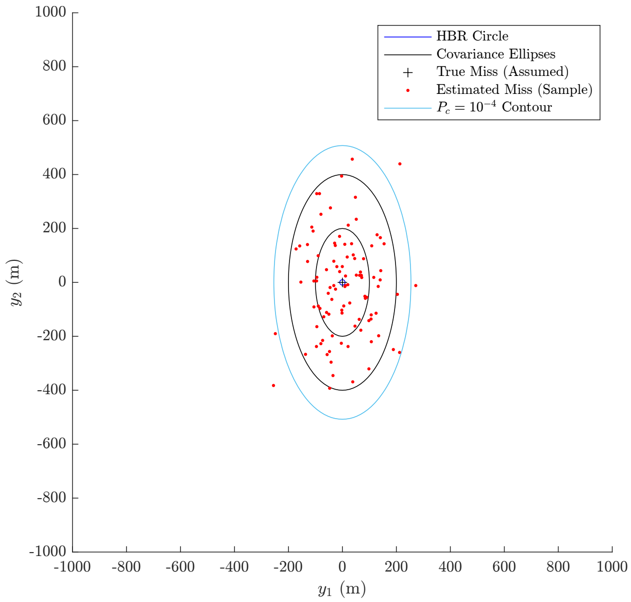

Fig. 4: An illustration of the idea behind P_d. The red dots represent a sample of estimated miss vectors that could have resulted from the orbit determination and propagation process, given the covariance ellipses shown, and given that the true miss vector is 0, i.e. a direct hit. The red dots inside the $P_c = 10^{-4}$ contour represent instances where the impending collision is "detected." Thus the fraction of dots falling inside the contour is an approximation of the "probability of detection".

3.3 Probability of Detection

Recall that, for a particular projected combined covariance matrix D, we know the distribution of possible estimated miss vectors, conditional on the true miss vector:

$$\widehat{Y} \mid Y = y \sim N(y, D). \tag{12}$$

Now, from the perspective of calculating the probability of collision, this is problematic, which is why, in a previous section, we needed to make use of Bayes' rule to arrive at the distribution of the *true* miss vector, conditional on the *estimated* miss vector. However, we can use the distribution given in (12) to calculate the probability that a collision will be *detected*, should the true miss distance be smaller than the HBR. This is done in the following manner. Given a particular 2D covariance and combined HBR, the P_c can be calculated for *any* hypothetical estimated miss vector. From this, we can find a region of the plane such that, if the estimated miss vector were to fall into it, then the P_c would be greater than the maneuver threshold T_α (often 10^{-4}). If we assume that $y = 0$, i.e. that the *true* miss vector is a direct hit, then from (12) we can calculate the probability that the P_c would exceed the maneuver threshold if a direct hit were going to occur. Following the authors of [10], this quantity will be called P_d, or the "probability of detection." The idea is illustrated in Figure 4. It is actually not necessary to assume a *direct* hit; the same calculation could be carried out with y anywhere inside the HBR. However, the P_d so calculated will depend on the particular value of y chosen, and when the HBR is small, this will be well approximated by the direct-hit value anyway.

Now the P_d could be calculated numerically, but the authors of [10] derive a simple analytic approximation:

$$P_d \approx \max\left\{1 - 2T_\alpha \frac{|D|^{\frac{1}{2}}}{\mathrm{HBR}^2}, 0\right\} \tag{13}$$

in which T_α is the chosen maneuver threshold and $|D|$ is the absolute value of the determinant of the 2D covariance matrix (in the conjunction plane). The approximation above relies on the assumption that the HBR is small relative to the covariance matrix, which is often a good assumption.

The significance of P_d to long-term risk assessment is this: if we assume that a collision avoidance maneuver completely eliminates the risk of collision (which, admittedly, is not a fully healthy assumption), and if we exclude potential

Table 1:

Mission	Terra	FGRST	NOAA 20
Average P_d	.71	.39	.70
FRR based on P_c^{agg}	.91	.46	.67

collisions with untracked space debris, then the P_d tells us the expected proportion of collisions avoided over a long period of time, given our chosen maneuver threshold. Notice, however, that P_d depends on the specific covariance and combined HBR used in the calculation; generalizing the P_d to entire histories of encounters with varying covariances and HBRs requires some additional effort. If we wish to estimate the expected proportion of collisions avoided under operational circumstances, we need to *average* the P_d over the empirical distribution of covariances and HBRs encountered in the data. This can be done on a per satellite, per constellation, or per regime basis, as demanded by the application.

Table 1 shows average P_d values calculated using historical CDM data from three NASA missions over the period 2021-2024, using an MCP of 1 day and a maneuver threshold of 10^{-4}. For comparison, Table 1 also contains the fractional risk reduction (FRR) that would be calculated on the same data using the aggregate P_c, according to the procedure described in section 1; figure 5 contains the aggregate P_c plots corresponding to this calculation. The first takeaway from Table 1 is that the aggregate P_c method can give a misleading idea of the risk reduction, if risk reduction is thought of as the average number of collisions avoided. This is shown by column 1 of Table 1, in which the P_d suggests that about 70% of collisions will be avoided on average, whereas the aggregate P_c suggests a much higher fraction of 91%. This discrepancy is not surprising when we remember that P_d is a frequentist probability, whereas the aggregate P_c is based on a Bayesian probability. It is somewhat surprising to note, however, that although the P_d *can* diverge dramatically from the FRR based on aggregate P_c, the two metrics often give similar answers, as shown by the latter two columns of Table 1. This is plausibly explained as follows. In order for the FRR calculation based on the aggregate P_c to yield a high value, there must be several conjunctions in the data that have P_c values above the maneuver threshold. But in order for a conjunction to yield a high P_c value, its covariance must be sufficiently small. This is a corollary of the phenomenon known as "probability dilution", where very large covariances cause the P_c value to be small, regardless of the estimated miss location. But if a conjunction must have a "small" covariance in order to yield a high P_c, this small covariance will cause the P_d to be relatively large, other things being equal. This chain of correlations will cause the FRR estimated with aggregate P_c to be correlated with the average P_d, so that it is less surprising that these two metrics should often be close in value. In spite of this correlation, however, if one is interested in estimating the expected fraction of collisions avoided, P_d will be more reliable.

Once we have a good estimate of the average fraction of collisions that we are actually avoiding, we can multiply this by the probability of collision in the absence of maneuvers, which we may derive from, say, a flux model. This calculation yields the *residual* collision risk.

Notice that to derive the P_d we are only relying on the distribution of the *estimated* miss distance, given by (12). This distribution is derived directly from the distribution of the measurement errors, which is assumed to be known and to accurately reflect their relative *frequencies*. This contrasts with the derivation of (10), which relies on Bayes' rule and the assumption of a uniform prior on the true 2D miss vector. Thus, we can be confident that P_d does, indeed, represent the long-term fraction of collisions detected, as long as our measurement error models are correct.

There is something else about Table 1 that is worthy of comment: namely, the fraction of collisions we can expect the P_c to detect, with a maneuver threshold of 10^{-4}, is surprisingly low, sometimes less than half. What makes this even more surprising is that this is the expected fraction of collisions detected among *tracked* objects. In addition, the average P_d values can vary substantially from satellite to satellite, even with the same maneuver threshold, and thus satellites can receive substantially different amounts of risk mitigation from following the same risk mitigation strategy. Now, the P_d expression above shows that lowering the maneuver threshold will tend to improve the detection performance for individual satellites; but this still leaves the variation of the detection performance from satellite to satellite untouched. This motivates the search for an alternative risk metric which has a fixed, known probability of detection.

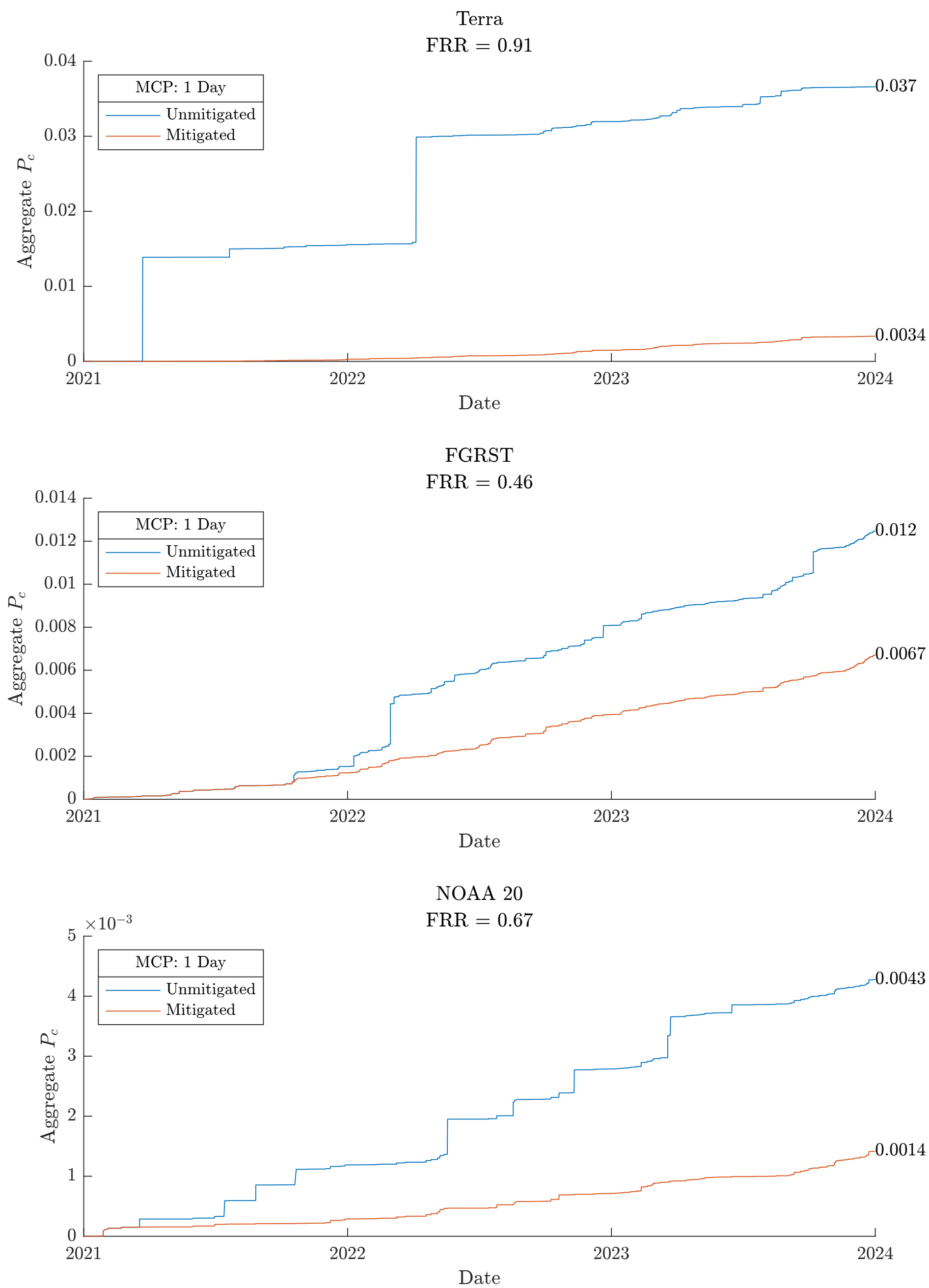

Fig. 5: Illustrates the fractional risk reduction calculated with aggregate P_c.

3.4 Alternative Risk Metrics

The authors of [5] propose a method of collision risk assessment based on the idea of "null hypothesis testing". The proposed method begins with the assumption, or "null hypothesis", that a collision *will occur*. This may seem like a strange way to begin the analysis, but it has useful consequences. To simplify, let us assume not only that a collision will occur, but that it will be a direct hit, i.e. $y = 0$.[9] It follows from (10) that, under this hypothesis

$$\widehat{Y} \sim N(0, D). \tag{14}$$

From this it follows that the squared Mahalanobis distance of the estimated miss vector from the origin,

$$t = \left(\frac{\widehat{Y}_1}{d_1}\right)^2 + \left(\frac{\widehat{Y}_2}{d_2}\right)^2, \tag{15}$$

is distributed chi-square with two degrees of freedom. We call t the "test statistic," because it will allow us to test the null hypothesis. Recall that $\widehat{Y}$ is a random variable that describes the distribution of estimated miss vectors over *many* trials. Let $\widehat{y}$ be the estimated miss vector that is *actually observed* in the particular conjunction under consideration. We can then calculate the observed value of the test statistic,

$$t_{obs} = \left(\frac{\widehat{y}_1}{d_1}\right)^2 + \left(\frac{\widehat{y}_2}{d_2}\right)^2.$$

Now, given the null hypothesis, we know what distribution t_{obs} was drawn from, namely a chi-square distribution with two degrees of freedom. This means that we can compute the probability of drawing a value equal to or larger than t_{obs},

$$p_{obs} = \mathbb{P}(t \geq t_{obs}). \tag{16}$$

This is the probability that we *would have* obtained a larger value for the test statistic than we actually did, if the null hypothesis of a direct hit were true. This probability is also called a "p-value". If this value is "small enough," we reject the null hypothesis in favor of the alternative hypothesis that no collision will occur. Because of the particular form of the test statistic we have derived for our problem, this amounts to concluding that the conjunction is safe if the estimated miss vector is "far enough" away from the origin (relative to the size of the covariance).

The reasoning behind this procedure is analogous to proof by contradiction [5]. We begin by assuming the conclusion we wish to disprove. In our case, this is the proposition that a collision is imminent. We then attempt to show that this assumption leads to an "absurdity." This is the role of the test statistic and p-value. If the p-value is "too small", then we have observed a value of the test statistic that is very unlikely to occur if the null hypothesis were true, and this "absurdity" convinces us to abandon the null hypothesis.

But how do we decide if a given p-value is "too small"? The standard way is to consider the rate at which different possible p-value thresholds will cause us to incorrectly reject the null hypothesis. For the conjunction assessment problem, this corresponds to the expected fraction of actual collisions that will be incorrectly classified as non-collisions. Let us say we are only willing to falsely reject the null at a rate of 10^{-4}. Then we simply reject the null hypothesis when p_{obs} is less than 10^{-4}, and fail to reject otherwise. By construction, if the null-hypothesis is actually true, p_{obs} will only drop below this threshold 1 in 10,000 times, and thus we will only falsely reject the null hypothesis 1 in 10,000 times.

It should be noted that p_{obs} does not even attempt to represent the probability of collision itself; it assumes a collision is going to occur! Rather, p_{obs} represents the consistency of the data with the hypothesis of a collision.

A visual illustration of this method is given in Figure 6. This figure is strikingly similar to Figure 4, and this is no coincidence. It turns out that, if the chosen p_{obs} maneuver threshold is α, then the "probability of detection" using p_{obs} will simply be $1-\alpha$. Thus a threshold of 10^{-4} corresponds to a .9999 probability of detection. This is much higher than the P_d values we observed in the previous section for representative satellites. This performance comes at a cost, however. Preliminary profiling by the NASA CARA analysis team suggests that adopting p_{obs} as the operational

[9] The method proposed in [5] does not require this assumption, but it is made here so that the broad outlines of the method are not obscured by technical detail. The test statistic derived in [5] is also slightly different.

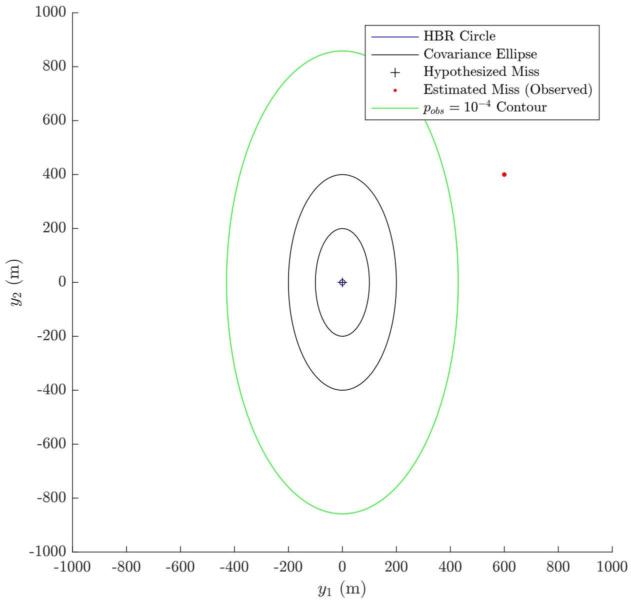

Fig. 6: Illustration of the use of p_{obs}. The estimated miss for this conjunction is far enough away from the origin that we reject the null hypothesis of a collision.

risk metric, with a maneuver threshold of 10^{-4}, would increase the number of mitigation maneuvers by a factor of ~ 8. Even if one were to set the p_{obs} threshold significantly higher, say to .01, the increased maneuver burden would still be unacceptably high for many operators. Even still, the results of the average P_d calculations above show that the p_{obs} threshold could be brought up to .1 or higher and still match or exceed the performance of P_c in terms of the expected fraction of collisions avoided. Now it is true that the P_d can be improved by adjusting the maneuver threshold T_α downward. But this brings us to another attractive feature of p_{obs}. Recall that P_d depends on the size of the covariance and the HBR, and thus to get an estimate of the overall detection performance of P_c for a particular satellite/regime, it is necessary to average the P_d over the distribution of observed covariances and combined HBRs. However, the detection performance of p_{obs} is independent of the covariance, and can be practically read off from the chosen maneuver threshold.[10]

Thus, if we assume, as in the previous section, that all maneuvers are perfectly effective, and if we ignore collisions with untracked space debris, then the fraction of collisions avoided will simply be 1 minus the chosen p_{obs} threshold α. This can be combined with a separate method for calculating the expected frequency of collision in the absence of maneuver, like a flux model, to give a measure of the residual risk in terms of the long-term frequency of collision. This makes p_{obs} an interesting alternative risk metric to P_c.

Another conjunction risk assessment method that enjoys a similar connection with the fraction of collisions avoided is the ellipsoid-overlap method of [1]. This method consists in "flying" a $(1-\alpha)$ confidence ellipsoid along the predicted trajectory of both the primary and secondary and maneuvering if they intersect. Because of the coverage properties of confidence regions, this will ensure that only a fixed fraction of actual collisions will be missed, and thus this method has the same desirable properties for long-term risk assessment as the null-hypothesis testing approach. However, this method suffers from the same practical difficulties in the form of a highly elevated maneuver frequency.

4. CONCLUSION

Aggregate P_c, although an intuitive metric and a natural extension of the current risk assessment framework based on P_c, is most plausibly interpreted as a Bayesian "degree of belief", rather than as the long-term frequency of collision. Thus, if aggregate P_c continues to be utilized in long-term risk analysis, it should be interpreted with care, so as not to draw misleading conclusions.

[10] The detection performance of p_{obs} is still somewhat dependent on the HBR, but the dependence is weak if the HBR is small.

Further research is needed to determine how accurately flux models can model the expected long-term frequency of collision in the absence of maneuvers, and to determine whether there are any viable alternative methods for determining this frequency. However, even in the absence of such a method, it is still possible to estimate the fractional risk *reduction* that results from adopting a particular maneuver strategy, either by utilizing post-hoc analysis methods like the P_d, or by adopting an alternative risk metric that has its collision detection performance built in. One possible future research endeavor would be to carry out a thorough profiling exercise to evaluate the average collision detection performance of the P_c as currently implemented (using the average P_d described [10]), as well as the increased maneuver burden that would result from the adoption of alternative risk assessment methods like p_{obs} or the ellipsoid overlap method. Even in the face of a substantial increase in maneuvers, the added assurance of a fixed, known collision detection rate given by these alternative risk assessment methods could be worth the cost, especially for operators of large constellations of autonomously maneuvering satellites.

REFERENCES

[1] M.S. Balch, R. Martin, and S. Ferson. Satellite Conjunction Analysis and the False Confidence Theorem. In *Proccedings of the Royal Society A*, 475:20180565, 2019.

[2] J.O. Berger, A. Philippe, and C. P. Robert. Estimation of Quadratic Functions: Noninformative Priors for Non-Centrality Parameters, *Statistica Sinica*, 8:359-375, 1998.

[3] J.R. Carpenter, A.C. Davison, S. Elkantassi, and M.D. Hejduk. Inference on the Miss Distance in a Conjunction. In *Proceedings of the 18th Annual Conference on Space Operations*, 2025.

[4] P. Diaz, P.M. Ripoll, M. Duncan, M. Lindsay, T. Harris, and H.G. Lewis. Data-Driven Lifetime Risk Assessment and Mitigation Planning for Large-Scale Satellite Constellations. In Proceedings of the AMOS Technical Conference, 2022.

[5] S. Elkantassi and A.C. Davison. Space Oddity? A Statistical Formulation of Conjunction Assessment, *Journal of Guidance, Control, and Dynamics*, 45(12): 2258-2274, 2022.

[6] J.L. Foster and H.S. Estes. 'A Parametric Analysis of Orbital Debris Collision Probability and Maneuver Rate for Space Vehicles, NASA/JSC-25898, 1992.

[7] Doyle T. Hall. Determining appropriate risk remediation thresholds from empirical conjunction data using survival probability methods, *Advances In The Astronautical Sciences*, 171:3439–3456, 2019.

[8] D. T. Hall, Expected Collision Rates for Tracked Satellites, *Journal of Spacecraft and Rockets*, 28(3):715-728, 2021.

[9] O. Montenbruck and E. Gill. *Satellite Orbits*, pp. 257-265, Springer, 2000.

[10] T.H. Sweetser, B.M. Braun, M. Acocella, and M. A. Vincent. Quantitative Assessment of a Threshold for Risk Mitigation Actions, *Journal of Space Safety Engineering*, 7(3): 318-324, 2020.