Evaluating Probability of Collision (Pc) Uncertainty

M.D. Hejduk
L.C. Johnson
April 2016
Agenda

- Conjunction Assessment Basics
- Probability of Collision (Pc) calculation outline
- Pc uncertainty overview
- Pc uncertainty component: covariance uncertainty
  - Covariance realism assessment
  - Covariance realism PDF generation
- Pc uncertainty component: hard-body radius uncertainty
  - Primary objects using projected-area sampling
  - Secondary objects using radar cross-section values
- Pc uncertainty component: natural variation in Pc calculation
- Example output
- Conclusions and future work
How are Satellite Collision Risks Determined/Mitigated?

- Certain spacecraft are determined to be “defended assets”
  - Will be evaluated for collision risk with other objects
- For seven days into the future, the expected positions of the defended asset and the rest of the objects in the space catalogue are determined
- “Keep-out volume” box drawn around the defended asset at each time-step
  - Typically 5km x 5km x 25km in size, with the longer dimension along the orbit path
- Any satellite that penetrates the keep-out volume during the 7-day analysis is considered a possible “conjunction”
- Particulars of the close approach analyzed to determine actual conjunction risk
“Fly By” Ephemeris Comparison

- Ephemerides generated for primary and secondaries that are possible threats
- Screening volume box (or ellipsoid) constructed about primary
- Box “flown” along the primary’s ephemeris
- Any penetrations of box constitute possible conjunctions
- For these conjunctions, Conjunction Data Message (CDM) generated
  - State estimates and covariances at TCA
  - Relative encounter information
  - OD information
- CDM data used to calculate probability of collision (Pc)
Calculating Probability of Collision (Pc): 3D Situation at Time of Closest Approach (TCA)

Figure taken from Chan (2008)
• Assumptions
  – Error volumes (position random variables about the mean) are uncorrelated

• Result
  – All of the relative position error can be centered at one of the two satellite positions
    • Secondary satellite is typically used
  – Relative position error can be expressed as the additive combination of the two satellite position covariances (proof given in Chan 2008)
    • $C_a + C_b = C_c$
  – Must be transformed into a common coordinate system, combined, and then transformed back
Calculating Pc: 2-D Approximation (2 of 3)

Projection to Conjunction Plane

- Combined covariance centered at position of secondary at TCA
- Primary path shown as “soda straw”
- If conjunction duration is very short
  - Motion can be considered to be rectilinear—soda straw is straight
  - Conjunction will take place in 2-d plane normal to the relative velocity vector and containing the secondary position
  - Problem can thus be reduced in dimensionality from 3 to 2
- Need to project covariance and primary path into “conjunction plane”
Calculating $P_c$: 2-D Approximation (3 of 3)

Conjunction Plane Construction

- Combined covariance projected into plane normal to the relative velocity vector and placed at origin
- Primary placed on x-axis at (miss distance, 0) and represented by circle of radius equal to sum of both spacecraft circumscribing radii ("hard-body radius" or HBR)
- Z-axis perpendicular to x-axis in conjunction plane

Figure taken from Chan (2008)
2-D Probability of Collision Computation

- Rotate axes until they align with principal axes of projected covariance ellipse
- $P_c$ is then the portion of the density that falls within the HBR circle
  - $r$ is $[x \ z]$ and $C^*$ is the projected covariance

$$P_c = \frac{1}{\sqrt{(2\pi)^2|C^*|}} \int_A \exp \left( -\frac{1}{2} r^T C^{-1} r \right) dX dZ$$
2-D vs. 3-D Conjunction Geometry
• 2-D approximation not valid
• Can attempt 3-D integral
  – Messy, but Coppola (2012) outlines methodology with Lebedev quadrature
• Can use Monte Carlo
  – From TCA
    • Propagate both satellites’ states and covariances to nominal TCA
    • Take position (and maybe velocity) perturbations from each covariance to define new states for primary and secondary
    • Find new TCA and record miss distance
    • Tabulate all miss distances; percent that are smaller than HBR is Pc
  – From epoch
    • Similar procedure to above, but perturbations performed at epoch
    • Perturbed states propagated forward to new TCA with full non-linear dynamics
Conjunction Event Canonical Progression

- Conjunction typically first discovered 7 days before TCA
  - Covariances large, so typically \( P_c \) below maximum
- As event tracked and updated, changes to state estimate are relatively small, but covariance shrinks
  - Because closer to TCA, less uncertainty in projecting positions to TCA
- Theoretical maximum \( P_c \) encountered when 1-sigma covariance size to miss distance ratio is \( 1/\sqrt{2} \)
  - After this, \( P_c \) usually decreases rapidly
- Behavior shown in graph at right
  - X-axis is covariance / miss distance
  - Y-axis is \( \log_{10} \left( P_c / \max(P_c) \right) \)
  - Order of magnitude change in \( P_c \) considered significant, thus log-space more appropriate
• **Pc is only a nominal solution for the conjunction**
  – Derived from estimates of the mean
    • If underlying distributions not symmetric, then this is not an expression of central tendency
  – Does not include uncertainties on the inputs
    • “Uncertainty of uncertainty volumes” or uncertainty in HBR

• **Thus, while representing the risk, nominal Pc is just a point estimate**

• **Want to know how much variation or uncertainty in the Pc calculated for any given conjunction**
  – Determine uncertainty PDFs for the Pc calculation inputs
  – Through Monte Carlo trials, vary above inputs to the Pc calculation
  – Include a resampling technique to determine natural variation in the calculation
  – Generate a probability density of resultant Pc values
  – Characterize this distribution empirically
- Generate a $P_c$ distribution, using Monte Carlo (MC) trials of the underlying uncertainties
  - Determine uncertainty for each of the $P_c$ parameters

**Underlying Uncertainties**

- **HBR Uncertainty**
  - Draw from projected area distribution (primary) and RCS PDF (secondary)

- **Covariance Uncertainty**
  - Draw from scale factor distributions for both objects

- **Natural Sampling Variability**
  - Draw from 2D scaled Gaussian covariance

**Generate $P_c$ distribution**
COVARIANCE REALISM AND SCALE FACTORS
Covariance Realism

• **Ways a typical covariance can be unrealistic**
  – Much larger or smaller than the “real” error volume
  – Differently oriented from the “real” error volume
  – Representing a different distribution from the “real” error distribution

• **This last item not addressed in present study**
  – Current form of covariance promotes Gaussian assumption
  – *A priori* arguments for presuming component error distributions close to Gaussian
  – *A posteriori* evidence for component errors following a symmetric distribution
  – Study indicates large-Pc events not affected by “bending” covariances*

• **Large covariances not inherently problematic**
  – Rather, quite appropriate if errors themselves are large

• **Covariance realism assessment approach is combined evaluation of size and orientation, presuming error volume is Gaussian ellipsoid**

JSpOC State and Covariance Accuracy Utility

• Truth ephemeris produced for every satellite
  – Similar to methodology used for generating precision satellite laser ranging orbits
  – “Stitched together” pieces of ephemeris from a “judiciously chosen” portion of the fit-spans of subsequent batch ODs
    • Methodology to minimize overlap of portions drawing from same observation base
  – Covariance for reference orbit also preserved (epoch covariances from generating ODs)

• Each produced precision vector for each object compared to its reference orbit at propagation states of interest
  – Position comparisons at 6, 12, 18, 24, 36, 48, 72, 120, and 168 hrs
  – Propagated position covariance also calculated and retained at each comparison point

• Raw materials for covariance realism investigations thus available:
  – State errors
  – Propagated covariance at point of comparison and reference orbit covariance
Reference Orbit Formation Approaches: Previous and Present

Old Approach

New Approach

Fit Span

~ abutment

0 % 25 % 50 %

0 % 100 %

b.

a.

~ abutment
• Let \( q \) and \( r \) be vectors of values that conform to a Gaussian distribution
  – Commonly called normal deviates

• A normal deviate set can be transformed to a standard normal deviate by subtracting the mean and dividing by the standard deviation
  – This produces the so-called Z-variables

\[
Z_q = \frac{q - \mu_q}{\sigma_q}, \quad Z_r = \frac{q - \mu_r}{\sigma_r}
\]

• The sum of the squares of a series of standard normal deviates produces a chi-squared distribution, with the number of degrees of freedom equal to the number of series combined

\[
Z_q^2 + Z_r^2 = \chi^2_{2 \, \text{dof}}
\]
Covariance Realism: Normal Deviates in State Estimation

- In a state estimate, the errors in each component (u, v, and w here) are expected to follow a Gaussian distribution
  - If all systematic errors have been solved for, only random error should remain
- These errors can be standardized to the Z-formulation
  - Mean presumed to be zero (OD should produce unbiased results), so no need for explicit subtraction of mean

\[
Z_u = \frac{u}{\sigma_u}, \quad Z_v = \frac{v}{\sigma_v}, \quad Z_w = \frac{w}{\sigma_w}
\]

- Sum of squares of these standardized errors should follow a chi-squared distribution with three degrees of freedom

\[
Z_u^2 + Z_v^2 + Z_w^2 = \chi^2_{3 \text{dof}}
\]
Covariance Realism:  
State Estimation Example Calculation

- Let us presume we have a precision ephemeris, state estimate, and covariance about the state estimate
  - For the present, further presume covariance aligns perfectly with uvw frame (no off-diagonal terms)
- Error vector \( \varepsilon \) is position difference between state estimate and precision ephemeris, and covariance consists only of variances along the diagonal
  - Inverse of covariance matrix is straightforward

\[
\varepsilon = \begin{bmatrix} \varepsilon_u \\ \varepsilon_v \\ \varepsilon_w \end{bmatrix}, \quad C = \begin{bmatrix} \sigma^2_u & 0 & 0 \\ 0 & \sigma^2_v & 0 \\ 0 & 0 & \sigma^2_w \end{bmatrix}, \quad C^{-1} = \begin{bmatrix} 1/\sigma^2_u & 0 & 0 \\ 0 & 1/\sigma^2_v & 0 \\ 0 & 0 & 1/\sigma^2_w \end{bmatrix}
\]

- Resultant simple formula for chi-squared variables

\[
\varepsilon C^{-1} \varepsilon^T = \frac{\varepsilon_u^2}{\sigma^2_u} + \frac{\varepsilon_v^2}{\sigma^2_v} + \frac{\varepsilon_w^2}{\sigma^2_w} = \chi^2_{dof}
\]
• Mahalanobis distance formulary naturally accounts for correlation terms
• Two-dimensional example:

\[ \varepsilon C^{-1} \varepsilon^T = \frac{1}{1 - \rho^2} \left( \frac{\varepsilon_x^2}{\sigma_x^2} + \frac{\varepsilon_y^2}{\sigma_y^2} - \frac{2\rho \varepsilon_x \varepsilon_y}{\sigma_x \sigma_y} \right) \]

• Conforms to intuition
  – As \( \rho \) approaches zero, diagonal case recovered
Covariance Realism: Testing for Realism

- Mahalanobis distance set should conform to 3-DoF $\chi^2$ distribution
- Expected value for each calculation is DoF, 3 in this case
- Each Mahalanobis point in principle produces a scale factor
  - $mCm$ sizes covariance such that $\varepsilon C^{-1}\varepsilon^T$ will have a value of 3
  - $m^2$ thus the proper factor by which to scale the covariance in order to produce
    the expected value
- However, not every Mahalanobis calculation expected to equal
  expected value
  - Instead, a chi-squared distribution with expected value of 3
- To set scale factor(s), choose factor that brings entire Mahalanobis
distance set into conformity with expected distribution
Empirical Distribution Function (EDF) GOF: Exquisite Solution

- Sum of vertical differences between “ideal” and “real” behavior
  - Hypothetical graph at left
- Cramér – von Mises formulation the most appropriate for current situation
  - Equations at right
  - Weighting function ($\psi$) set to unity a better choice for outlier-infused situations
- $Q$ used to consult tables of $p$-values to determine likelihood of match between ideal and real distribution
  - Best approach is to be able to use $p$-value, as this has a clear statistical meaning

$$Q = n \int_{-\infty}^{\infty} [F_n(x) - F(x)]^2 \psi(x) dx$$
$$\psi(x) = 1$$

- But what if we want a distribution of scale factors?
Covariance Realism: Distribution of Scale Factors

- Rank-ordering of results can give reasonable PDF of scale factors
  - Presume 100 squared Mahalanobis distance values (M2)
    - Derived from JSpOC covariance realism data
  - Rank order list
  - Align each entry with the 3-DoF $\chi^2$ value that corresponds to that percentile
  - Quotient of two terms is (square of) scale factor that produces the $\chi^2$ value expected for that particular percentile

- Examples:

<table>
<thead>
<tr>
<th>Percentile</th>
<th>$x^2$</th>
<th>Square of M Distance</th>
<th>Quotient</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (0.01)</td>
<td>0.115</td>
<td>0.183</td>
<td>1.594</td>
</tr>
<tr>
<td>2 (0.02)</td>
<td>0.185</td>
<td>0.245</td>
<td>1.326</td>
</tr>
<tr>
<td>3 (0.03)</td>
<td>0.245</td>
<td>0.353</td>
<td>1.440</td>
</tr>
<tr>
<td>4 (0.04)</td>
<td>0.300</td>
<td>0.418</td>
<td>1.393</td>
</tr>
</tbody>
</table>
HARD-BODY RADIUS
• HBR is typically found by circumscribing both objects in spheres and combining the objects into one bounding sphere
  – Size of the secondary is typically not known, so added as a large estimate of debris object dimensions

• HBR uncertainties that follow represent a more realistic estimate of the area in the conjunction plane
  – The combined uncertainties are much smaller than the bounding sphere
Hard-Body Radius: Min and Max Using Approximation Equations

Could presume uniform distribution between these values as first-order approximation of PDF, but seems rather arbitrary.

- $HBR_{\text{max}} = \frac{\sqrt{L^2 + (2w + d)^2}}{2} \approx 8.1 \text{m}$
- $L = 13.2 \text{m}$
- $d = 4.2 \text{m}$
- $w = 2.6 \text{m}$

- $HBR_{\text{min}} = \frac{2w + d}{2} \approx 4.7 \text{m}$
• Randomized orientation of primary satellite to capture the average area
  – Ball-and-stick model to be created for each primary asset
  – Includes rotating solar panels

• Projected radius
  – Actual hit area of the satellite expressed as a circular radius
  \[ r = \sqrt{\frac{A}{\pi}} \]
Hard-Body Radius: Projected Area Approach Performance

- NASA/JSC Orbital Debris Program Office (ODPO) has sophisticated satellite model and full Euler angle rotation software to generate projected area PDFs
- Comparison of results for Hubble Space Telescope between ODPO software and ball-and-stick model:

<table>
<thead>
<tr>
<th></th>
<th>Average Area [m²]</th>
<th>Average Effective HBR [m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crude HST model (corresponding to chart)</td>
<td>60.3</td>
<td>4.3</td>
</tr>
<tr>
<td>Sophisticated HST model (Matney*)</td>
<td>63.7</td>
<td>4.5</td>
</tr>
</tbody>
</table>

Good agreement

Hard-Body Radius: Projected Area Approach Implementation

• Assemble ball-and-stick model of primary satellite
• Rotate through all Euler angles and project into plane
• Create empirical PDF of projected areas
• Express as PDF of radii of circles of equivalent area

• If satellite orientation is known at TCA, then area can be projected directly into conjunction plane
  – Can then perform integration by means of a contour integral
  – Lingering problem of how to incorporate area for secondary object
For intact spacecraft, possible to use published dimensions
  – For payloads, these are often not precise enough to be useful, and at least some canonical models would have to be imposed
    • Error in all of this great enough that approach is questionable
  – For rocket bodies, published dimensions are probably adequate
    • But many booster types lack published dimensions

Most common secondaries are debris objects, for which no size information is available

Thus, forced to estimate size from radar cross-section (RCS) value
  – Objects do not have single RCS value but PDF of values, depending on radar response and object aspect function
  – PDFs of individual objects’ RCS values not available, only averaged values
  – As proxy could use canonical distribution
    • Swerling III distribution is most common for debris, and also most conservative in terms of size*

Hard-Body Radius: Swerling Distribution Family

- Swerling distributions derive from the gamma distribution family
  - Location parameter ($\gamma$) = 0
  - Shape parameter ($m$) fixed
  - Scale parameter ($\beta$) estimated from sample (MLE)
- Swerling I/II is gamma with $m=1$
  - Exponential distribution
  - Presumes Rayleigh scattering
- Swerling III/IV is gamma with $m=2$
  - Erlang distribution
  - Presumes correlation with object orientation; more correct assumption
- S-notation is gamma with $m$ given
  - $S_{1.5} =$ gamma with $m=1.5$ &c.

\[
f(x; \gamma, \beta, m) = \frac{1}{\beta^m \Gamma(m)} (x - \gamma)^{m-1} \exp\left(\frac{-(x - \gamma)}{\beta}\right)
\]

\[
f(x; \beta) = \frac{1}{\beta} \exp\left(\frac{-x}{\beta}\right)
\]

\[
f(x; \beta) = \frac{1}{\beta^2} x \exp\left(\frac{-x}{\beta}\right)
\]
• Simulated hyperkinetic destruction of satellite in vacuum chamber
• Collected pieces and subjected them to individual analysis
  – “Observed” each piece with radar in anechoic chamber
  – Articulated full range of aspect angles and full range of radar frequencies
  – Recorded resultant RCS of each aspect/frequency configuration
• Collected results and plotted in dimensionless format
  – RCS / $\lambda^2$; size / $\lambda$
  – Results follow basic theory of Rayleigh, Mie, and optical regions
Hard-Body Radius: Full Process for Secondary Object

- Begin with average RCS
- Produce RCS PDF using Swerling III distribution
  - Scale parameter estimated by mean RCS divided by shape parameter
- Send distribution through ODPO size estimator to generate size PDF
  - Certified only for objects smaller than 20cm, but this is most debris
PC CALCULATION
RESAMPLING
Pc Calculation Resampling

- Resampling/bootstrap methods often used to generate confidence intervals when calculation final distribution unknown
- Early attempts at this with Pc used resampling with invariant covariances
  - Take position draw on primary and secondary covariance at TCA
  - Find new TCA; this defines new nominal miss vector
  - Recompute Pc with this new miss vector and unaltered covariances
  - Problem: covariance is clearly correlated with conjunction geometry
    - Cannot produce new miss distance from covariance-based sampling and then recompute Pc using those same covariances
- Need approach that considers miss distance / covariance linkage
J.H. Frisbee proposed a resampling technique that would also address the correlation problem
– Choose samples from the combined covariance to generate \( m \) miss vectors
– Take mean of \( m \) miss vectors—this is new nominal miss
– Take sample covariance of \( m \) miss vectors—this is new combined covariance
– Compute \( P_c \) using this mean miss distance and sample combined covariance
– Repeat procedure \( n \) times—this produces bootstrap dataset

Resampling Approach Issues

• In this framework, covariances are considered representatives of parent distributions, here further characterized by resampling

• Issue: what should be the value of $m$?
  – In bootstrapping, want the bootstrap sample size to equal the single-sample size that would have been used (or was used) to estimate the parameter
  – Thus, want the number of samples (DoF) of the bootstrap resampling ($m$) to equal the DoF that produced the covariance in the first place
  • That is, the DoF of the generating OD
• DoF is usually calculated as the number of data points minus the number of estimated parameters
  – JSpOC ODs calculated with SSN obs (usually have range, azimuth, and elevation—three observables)
  – Obs provided in “tracks”—group of obs taken during one tracking session

• Thus, tabulation issues arise
  – Each ob provides 3 DoF, minus the estimated parameters
  – However, rather little information content in interior obs of a track
    • JSpOC “track weighting” confirms this—all tracks weighted the same in the OD, regardless of length
  – Better tabulation to count each track as equivalent of one state estimate
    • Longish track about enough data to execute a single state estimate, to first order
    • Total estimated parameters in OD would thus be only one—one state estimated
  – DoF calculation is thus “# of tracks – 1”
    • Would need to be amended for DS, where obs report only two parameters, and needs more work in general
Resampling Approach Schematic

- Repeated thousands of times to calculate distribution of Pc values
- Benefits
  - Correlation of the miss vector and the covariance
  - Maintains an equivalent sampling level to the original OD
    - Naturally responds to variations in tracking density
PROCESS RESULTS
Example #1
Example #2
Conclusions and Future Work

• Proposed method
  – Characterizes the PDF that can represent the Pc from a particular conjunction, given the uncertainties in covariances, HBR and natural variation in the Pc calculation
  – Gives a sense of the dynamic range of the Pc and allow maneuver decisions to be based on percentile points of this range rather than the nominal value alone
  – Provides a mechanism for obtaining a better expression of the calculation’s central tendency (here the median)

• Future Work
  – Refine DoF calculation and generate expansion for angles-only cases
  – Survey results from runs of large datasets
    • Stability studies of simplifying assumptions for faster processing
  – Examine potential as a Pc forecaster