State Uncertainty Normality Detection

Introducing an Unscented Transform-Based Test

Sven K. Flegel1 · James C. Bennett2

Published online: 6 December 2019 © The Author(s) 2019

Abstract

Two fundamentally different approaches of determining normality of the probability density function of the state estimation error are compared by application to a range of test cases. The first method is the Henze-Zirkler test, which operates on a random particle sample. The variability of its result is quantified. Using this method, departure from normality has been found to occur in three stages which are detailed. The second test compares the offset in whitened space of the predicted state to the predicted covariance mean obtained from the unscented transform. This test is much more efficient than the random particle based approach and can be applied using any perturbations model. The comparison is performed on the state estimation error in Cartesian space and using two-body motion without process noise. The more efficient, unscented transform based approach shows excellent agreement with the Henze-Zirkler test for constructed test cases. Application to orbit determination results from passive optical observations assessed with a Batch-Least-Squares orbit determination however reveals some discrepancies which have yet to be understood and underline the importance of rigorous testing.

Keywords Multi-variate normality · Astrodynamics · State uncertainty · Unscented transform based normalised offset (UNO) · Henze-Zirkler · Optical data

Introduction

The probability density function (PDF) of the state estimation error (uncertainty) of space objects is utilised in a variety of applications including collision probability estimation, manoeuvre detection, track association, multi-target tracking and sensor
scheduling. Each of these areas has its own accuracy and realism requirements of the uncertainty. Traditional orbit determination methods produce a state fit to observations based on root-mean-square minimisation [28]. Although the transformation from observation to state space is non-linear, the state estimation error PDF corresponding to the estimated state is typically given as a normal distribution (Gaussian) in the state parameters and expressed in the form of the covariance matrix. The point in time for which the state estimate is given is simply referred to as ‘epoch’ or ‘at epoch’ in the following. It seems to be generally accepted that the PDF of the state estimation error may be assumed to remain close to normal for a time after epoch, and many methods have been developed based on this assumption [5, 9, 18, 23, 24]. While some methods require the PDF to be normal on the entire $\pm \infty$ range [23], others operate on a portion of the PDF surrounding the covariance mean [24]. Over time, some features attributed to normality of the uncertainty PDF dissolve, however, so that methods operating on the traditional covariance matrix may become unreliable. As an example, symmetry and the distinct bell-shape are two well-known features of the normal distribution. Changing the weight equally in the two tails of the distribution will lead to a non-normal distribution which retains the feature of symmetry. There is no common agreement on how the departure from one feature should be quantified in relation to others or indeed how it should be determined. This merits the existence of different methods and also helps to better understand where differences in Type I and Type II errors in each may come from. An extensive comparison of methods applied to samples drawn from well known functions is presented in [21]. Some publications that apply these tests to the case of state error estimates in astrodynamics are [10, 27, 29]. The methods for Multivariate (Multi-dimensional) Normality (MVN) detection used in the current paper assess the time frame until a bounded volume surrounding the mean becomes non-normal. The tails of the PDF are not evaluated, so that the results should not be used for methods requiring the PDF to be normal in its entirety.

The Space Environment Research Centre (SERC Limited) is in the process of setting up a catalogue of space objects based on passive and active optical observations. The aim of the work presented here is to augment orbit determination (OD) results with information on the time frame until departure from normality is detected in the uncertainty PDF. Methods for collision likelihood estimation as well as others mentioned before, take the state and uncertainty data supplied in the catalogue as starting points for extrapolation to other points in time. The time frame until MVN departure is of particular importance when selecting appropriate methods that require extrapolation outside the time frame of available observations.

In the current paper, two fundamentally different approaches are used to determine departure from normality in state uncertainty. Commonality between the methods exists in that both operate on a set of states given at a common point in time. State propagation is limited to two-body motion and assumes no process noise. Proper modelling of process noise in itself can easily become quite a complex topic, if sources are to be properly taken into account. Although this topic is important, the current analysis instead concentrates on the assessment of the fundamental capabilities and limitations in the absence of these further complexities. As neither method requires knowledge on how the states were generated, at least formally, they can be
applied regardless of the propagation method. In Section “Multivariate Normality Test Baseline” the first of the two methods, the Henze-Zirkler test [11], is discussed. This test belongs to the family of consistent approaches and operates on a random particle sample. Examples of other tests that have been used in astrodynamics include the Kolmogorov-Smirnov, Royston and Doornik and Hansen tests [10, 29]. The test is applied to a set of constructed test cases and the time evolution is discussed. The second test introduced in Section “Efficient MVN Testing” relies solely on the unscented transform. The offset of the state mean from the covariance mean in whitened space is used to test for normality. Within the section, the two methods are compared based on the previously assessed test cases. The paper concludes by applying both approaches to a select number of orbit determination results from SERC’s own catalogue.

**Multivariate Normality Test Baseline**

The random sample approach was chosen to create a baseline for the assessment of the departure from normality in state uncertainty predictions. The approach involves the approximation of the continuous uncertainty PDF by a random set of discrete states. Most states will be generated in the vicinity of the highest probability density. Each state could be an exact representation of the true state and may be propagated through time independent of all other generated states. In this paper, the collection of states that are generated from the initial covariance in this fashion are called “sample”. The number of states within the sample describe the “sample size” and is denoted as \( n_p \). An MVN test takes into account all states within a sample and evaluates the likelihood that these could have been drawn from a multivariate normal distribution. Although relying on few assumptions, these approaches suffer from two major drawbacks: 1) Different MVN tests exist and each gives a different result for a given a sample, in particular when samples are ‘near’ normally distributed. 2) The outcome depends to some quantifiable degree on the particular random sample drawn from the PDF. The latter issue can be overcome by repeating the test by drawing new samples from the same PDF and looking at all outcomes. This is the basis of Monte-Carlo assessment and the number of samples used are denoted \( n_{mc} \). Monte-Carlo analysis can be computationally expensive. The former issue requires a good understanding of available statistical tests to chose one or several appropriate methods.

The section starts out with an outline of MVN tests presented in literature from which a suitable test is selected (Section “Random Sample MVN Test”). The second (Section “Henze-Zirkler Test for Normality”) contains several sub-topics: First, important aspects of the chosen test statistic are outlined. The evolution of this test statistic for an example case is then detailed to give some insight into its behaviour in describing the transition into non-normality (Section “Example Evolution of Henze-Zirkler MVN Test Values”). This is followed by an assessment of the impact of the sample size as well as the definition of an appropriate one for the given task (Section “Sample Size”). The central limit theorem (Section “Monte-Carlo Runs & Central Limit Theorem”) is used to assess how the variability of sample-by-sample results at each epoch translates into a time window. In the last
Section “Constructed Test Cases” a set of constructed test scenarios is introduced and the transition into non-normality is assessed using the outlined method.

**Random Sample MVN Test**

Results from large scale comparisons of different MVN tests which operate on random samples are published, for instance, in [21] and [6]. A grouping of MVN tests is outlined in [21]:

- Graphical and correlational approaches
- Skewness and kurtosis approaches
- Goodness-of-fit approaches, and
- Consistent approaches.

The well known Kolmogorov-Smirnov, Cramér-von Mises, and Anderson-Darling tests all fall into the “Goodness-of-fit approach” category. Among the selected methods was an extension of the Anderson-Darling test published by [26] which is given the name Romeu–Ozturk test. The Henze–Zirkler [11] test falls into the category of “Consistent approaches”. Consistent is used to indicate that it has been mathematically shown that the test will – at least in theory – consistently reject all non-MVN distributions. An example of tests of the category “Skewness and kurtosis approaches” are approaches based on the work of [20]. Among the criteria which the authors used to assess the tests was the rate of Type I and Type II errors as well as feasibility for implementation and desirable mathematical properties. The two major conclusions of this analysis were: a) No single test for MVN delivered perfect results and it was recommended to employ multiple methods for testing of MVN where possible; and, b) If only one test is used, the Henze–Zirkler was recommended. The Romeu–Ozturk was rejected early on in the study due to high Type I type error rates which in some cases exceeded 10%. These results are supported by a later study by [6]. In these publications, particle sample sizes were varied between 25 and 250 and sample space dimensions of up to 10 were considered. Based on these results, the Henze–Zirkler test was selected as baseline method.

**Henze-Zirkler Test for Normality**

The following equation calculates the test metric:

\[
HZ = \left[ \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} e^{-\frac{\beta^2}{2} D_{ij}} \right] - \left[ 2 \left( 1 + \beta^2 \right)^{-\frac{d}{2}} \sum_{i=1}^{n} e^{-\frac{\beta^2}{2 \left( 1 + \beta^2 \right)} D_i} \right] + \left[ n \left( 1 + 2 \beta^2 \right)^{-\frac{d}{2}} \right].
\]  

\text{(1)}

---

1Parts of this description have been taken with only minor alterations from [8].

2Type I error: An MVN distributed sample is incorrectly identified as being non-MVN distributed.

3Type II error: A non-MVN distributed sample is mistakenly identified as being MVN distributed.
Herein $D_i$ is the squared Mahalanobis distance of the $i^{th}$ observation to the centroid and $D_{ij}$ gives the squared Mahalanobis distance between $i^{th}$ and $j^{th}$ observations. The authors give a relation for $\beta$ which uses the particle sample size $n^4$ and sample space dimensions $d$:

$$\beta = \frac{1}{\sqrt{2}} \left( \frac{n (2d + 1)}{4} \right)^{\frac{1}{d+4}}. \quad (2)$$

This value is used to test the null-hypothesis $H_0$ which is that the sample is indeed MVN distributed. In general, the test statistic $HZ$ is small when the particles are MVN distributed and increases with deviation from MVN. To ease assessment, the $HZ$ value is normalised so that any value $HZ^*$ above 1 rejects the null-hypothesis and the distribution is considered to be non-normal. For a complete description of its implementation, the reader is referred to [8].

Example Evolution of Henze-Zirkler MVN Test Values

Using the fast Mikkola/Halley-4 single step solution for two-body motion as detailed in [22], a common structure is observed in all simulations performed to date. This is not a general proof of the structure being true for all cases. Such an assessment would be extremely valuable but has not yet been attempted. An example of the observed structure is given in Fig. 1. So far, the evolution could always be split roughly into four stages: i) fully normal, ii) transition iii) fully non-normal and iv) deep non-normal.

**Phase I: Fully Normal** During the initial phase following the state epoch, the uncertainty remains normal throughout an entire revolution. During this stage, Type I error causes some samples to still flag non-normality.

**Phase II: Transition** The uncertainty volume then becomes non-normal during parts of one revolution. During this phase, a common structure materialises between the time-evolution of all initial samples.

**Phase III: Fully Non-Normal** In the final phase, all samples exhibit the same time evolution and the uncertainty volume is non-normal throughout the entire revolution.

**Phase IV: Deep Non-Normal** The simple structure observed in the fully non-normal phase gives way to some more complex behaviour.

An interesting observation is a half-period periodicity in all simulations. This has some likeness to debris clouds from on-orbit fragmentations. In position space and in early stages after a breakup, the debris accumulates into a ‘pinch-point’ at the point on orbit where the fragmentation occurred, and into a ‘pinch-line’ at the opposite side of the orbit. The debris cloud is smallest at these two points and spreads out in

---

4Later on, the particle sample size $n$ is denoted as $n_p$ to avoid confusion with the number of Monte-Carlo runs $n_{mc}$

5This description expands on the one given in [7] and the companion paper [4]
between. Similar to the observed behaviour in the position uncertainty, two somewhat similar states are thus obtained at half-period intervals. It must be noted however that unlike particles generated from an uncertainty at epoch, fragmentation debris has no – or very little – initial position offset.

**Sample Size**

Figure 2 visualises the importance of proper sample size selection. The left hand plot has been produced by predicting an object’s position using only uncertainty in two dimensions of the position in ECI coordinates. The blue lines are made up of individual particles created at the \( \sigma \)-lines 1 to 10 at the initial epoch where normality is given by definition. Although the technical significance at the outskirts of the chosen scale is questionable, the same relation can be observed on any other scale as well and depends on the propagation time and the magnitude of initial uncertainties. The red lines are made up of the same particles after propagation to \( t_1 = t_0 + \Delta t \). To ease comparison of the two samples, the particle positions are transformed into Mahalanobis space. Within this space, particle distances from the mean are given in multiples of standard deviations. In the plot, particles close to the state mean describe concentric circles at \( t_0 \) as well as \( t_1 \). Particles at the outskirts however exhibit marked deformation in two-dimensional position space at \( t_1 \). In two-dimensional space, this may be taken as one feature required to allow classifying a PDF to be “normal”. Other features such as the area contained in the tails are not well captured in this representation. The amount of ‘bending’ which is observed to increase with distance from the
mean indicates that normality may remain intact for longer periods of time close to the mean state compared to regions that are further removed.

The right hand figure aims at demonstrating by example that the observed qualitative effect is measurable in simulations and must be accounted for appropriately. It displays the evolution of the Henze-Zirkler MVN test statistic over time for samples of different sizes drawn from a 6-d initial position & velocity covariance matrix. Only the 3-d position sub-space is tested for normality. Data points and trend fits are shown for the highest on orbit values only. The figure thus shows the epoch at which normality evolution enters the Transition Regime detailed in Section “Example Evolution of Henze-Zirkler MVN Test Values”. Each data point represents the mean value of \( n_{mc} = 5000 \) Monte-Carlo runs for initial detection of non-normality.

To define a suitable sample size, a relation is first required between the sample size \( n_p \) and the number of particles outside a defined confidence interval \( p \):

\[
n_p (v > v_0) = n_p (1 - p(v_0, k = 1)) \tag{3}
\]

The \( p \)-value is the area under the integrated \( \chi^2(v, k) \) function with degrees of freedom \( k \) that is outside the defined limit variance \( v_0 \). For \( k = 1 \), the \( \chi^2 \) function describes the distribution of residuals of the one-dimensional normal distribution. For the current purpose, the 3-\( \sigma \) bound (99.7% confidence level) is selected as the limit beyond which no significant particles are required. In addition, a significant number

---

Fig. 2 Left: Particles taken at the 1-\( \sigma \) to 10-\( \sigma \) lines before propagation \( (t_0) \) and after propagation \( (t_1) \). Right: Example of effect using different sample sizes from a 6-d position & velocity covariance. Plot shows mean value of \( n_{mc} = 5000 \) Monte-Carlo runs for initial detection of non-normality.

Given the non-linearity of the underlying dynamics, the hypothesis is introduced that there exists a radius in Mahalanobis space outside which the uncertainty PDF has significant non-normal features for any point in time \( t \neq t_0 \).
of particles should still be present between $98.7% \leq p \leq 99.7%$ confidence levels. The significant number of particles is set to 30. The number ‘30’ is borrowed from the central limit theorem (see Section “Monte-Carlo Runs & Central Limit Theorem”) without further investigation and may be revised in future studies. From Table 1 it can be seen that any sample size between 3 000 and 11 000 satisfies the above restrictions. With the understanding visualised in Fig. 2, larger sample sizes lead to earlier detection of non-MVN. For the current purpose it is decided to err on the conservative side and a sample size of $n_p = 10 000$ is chosen. This is the same number as was selected in previous studies.

Monte-Carlo Runs & Central Limit Theorem

It is in the nature of the approach that for any given epoch, not all samples lead to the same MVN-test result. Often, a decision is based on their mean value. A different set of samples would create a slightly different mean. The PDF of many such means follows a normal distribution when each mean used more than 25 - 30 samples. If fewer samples were averaged over, the PDF is expressed by the “Student’s T” distribution [12, pp. 200].

In the following, the decision will be based on $n_{mc} = 50$ so that the normal distribution applies. In this case, the standard deviation of the PDF of means $\sigma_{CLT}$ can be estimated from the standard deviation of the distribution of a single set of 50 results $s(n_{mc} = 50)$. This is the basis of the Central Limit Theorem (CLT) whose general relation is:

$$\sigma_{CLT} (n_{mc}) = \frac{s(n_{mc})}{\sqrt{n_{mc}}}.$$  \hspace{1cm} (4)

### Table 1  Sample sizes and the number of particles in the outskirts of the distribution

| Sample Size | 98.7% - 99.7% | > 99.7% |
|-------------|--------------|--------|
| 500         | 5.0          | 1.3    |
| 1 000       | 10.0         | 2.7    |
| 2 000       | 20.0         | 5.4    |
| 3 000       | 30.0         | 8.1    |
| 4 000       | 40.0         | 10.8   |
| 5 000       | 50.0         | 13.5   |
| 6 000       | 60.0         | 16.2   |
| 7 000       | 70.0         | 18.9   |
| 8 000       | 80.0         | 21.6   |
| 9 000       | 90.0         | 24.3   |
| 10 000      | 100.0        | 27.0   |
| 11 000      | 110.0        | 29.7   |
We want to know two things at this point. Firstly: What is the true state of the population at a given point in time? As we are dealing with the random particle approach, this information is only useful in combination with its variability. Secondly: How does this variability from epoch to epoch translate into a time window for the transition to non-normality. Such a window is only a further measure of understanding the variability of the results that are inherent to the particle method. They have no relation to the underlying population that we assume represents reality under the simplified physical properties of two-body motion and no process noise. In the following, we derive a continuous function by fitting to the discrete data from the Henze-Zirkler test that allows us to estimate such a time window.

**MVN Departure Onset and Conclusion Mean Epochs** Once MVN starts breaking down, it initially occurs only briefly during a revolution either at one or several locations. Section “Example Evolution of Henze-Zirkler MVN Test Values” has details on this. To identify the state of the uncertainty at any epoch, one must therefore evaluate the entire revolution. To decrease computational burden, departure onset and conclusion epochs are iterated towards. This is possible due to the purely analytic description of orbital motion. For the transition onset, the epoch is recorded which marks the start of the first revolution during which the MVN test has failed at least once. The departure conclusion epoch marks the start of the first revolution during which the MVN test fails at all epochs. Evaluating for MVN at set epochs allows sharp peaks to be missed however. Here, circular orbits are tested at 17 epochs and eccentric orbits at 35 epochs. The higher sampling rate for eccentric orbits is required as the higher dynamic behaviour causes MVN values to spike sharply during departure onset especially. Under-sampling causes peaks and therefore early stages of MVN departure onset to be missed.

**Continuous Function Fits to Discrete Henze-Zirkler Values** The CLT can now be employed to estimate a time frame around the determined mean MVN departure epochs. To achieve this, first, the highest and lowest mean values, $\bar{x}$, as well as sample standard deviation values, $s$, within each assessed revolution are identified. Continuous functions $\bar{x}(t)$ and $s(t)$ are then fit to the mean and sample standard deviation results independently. An example of this is shown in Fig. 3. Values near the MVN decision threshold are given high weights in the fit. This ensures that the variability in the departure epoch is reflected well.

The time dependent sample standard deviation $s(t)$ can be translated into the standard deviation of the PDF of the mean value $\sigma_{CLT}(n_{mc})$ as per Eq. 4. With this information, a continuous function can now be established which describes the time evolution of the Henze-Zirkler test statistic, given a user definable confidence level $R$:

$$0 = \bar{x}(t) + R \left[ \frac{s(t)}{\sqrt{n_{mc}}} \right] - 1 . \quad (5)$$
Confidence level values $R$ according to the cumulative standard normal distribution are:

| $R$ | -3 | -2 | -1 | 0  | 1  | 2  | 3  |
|-----|----|----|----|----|----|----|----|
| $1-\alpha$ [%] | 0.1 | 2.3 | 15.9 | 50.0 | 84.1 | 97.7 | 99.9 |

With $R = 0$, the relation simplifies back to $0 = \bar{x}(t) - 1$ and the threshold epoch as per the mean function fit $\bar{x}(t)$ is obtained. This relates to the 50% confidence level $(1-\alpha)$.

**MVN Departure Onset and Conclusion Windows** The inversion of Eq. 5 to yield $t$ would allow finding the epoch for MVN departure as function of confidence, $R$. For the ensuing analysis, $R$ is fixed and $t$ is solved for numerically.

**Constructed Test Cases**

The assessment of collision probability methods performed by Alfano [1] contains detailed information on states and uncertainties for twelve different close encounter cases. These cases create the basis for the current chapter. Table 2 gives an overview of the simulated scenarios. For more detail, the reader is referred to the original publication. The first number in each case ID is the case number. The second number...
Table 2 Overview of calibration cases based on [1]. Nomenclature is scenario No.-object No.

| Case ID | 1-1 | 3-1 | 4-1 | 1-1a | 3-1a | 4-1a | 5-1 | 6-1 | 7-1 | 8-1 | 9-1 | 10-1 |
|---------|-----|-----|-----|------|------|------|-----|-----|-----|-----|-----|------|
| Regime  | GEO | LEO | MEO | HEO  |
| Period / min | 1396.3 | 94.6 | 675.6 | 717.7 |

references the respective object from the close conjunction case. Cases number 3, 11 and 12 differ from other cases only in parameters not pertaining to the state and covariance and are thus neglected here. Instead the GEO cases have been duplicated with all terms in the covariance matrices multiplied by a factor of 100 (IDs ending with ‘a’).

It should be noted that none of the cases in this section have position-velocity co-variances. This is a crucial difference to the state estimation error covariance matrices contained in SERC’s catalogue at the time of writing and which are used in Section “Departure from Normality for Catalogue OD Results”.

Results for the transition onset and conclusion epochs based on the Henze-Zirkler Method and using 50 Monte-Carlo runs at a sample size of 10 000 are visualised in Fig. 4. The results are characterised in terms of the mean value of each Monte-Carlo analysis and its variability based on the CLT applying the method outlined in Section “Monte-Carlo Runs & Central Limit Theorem”. The variability of results using the CLT is determined for the 0.1 % and 99.9 % (± 3σ) confidence levels. For the GEO cases with enlarged uncertainties, the departure from normality simply advances in time. Although this may seem correct intuitively, it has been observed that not all terms have the same effect; e.g. an appropriate selection of covariance terms for instance have been observed to allow normality to break down later, even though variances may be higher in position as well as velocity.

![Fig. 4](image-url) Beginning and end of transition into non-normality of position uncertainty. Scenario identifiers refer to those detailed in Table 2. Error bars are for the CLT standard deviation ± 3σCLT. Blue is for MVN departure onset epoch. Red is for MVN departure conclusion epoch.
Efficient MVN Testing

The transition start and end epochs described in Section “Example Evolution of Henze-Zirkler MVN Test Values” have different implications for different methods of covariance propagation. An error-state transition matrix approach uses a prior covariance matrix to produce an update. As no information on non-normally distributed uncertainties is contained in this representation, any covariance matrix produced after the transition phase is entered may no longer be reliable. The covariance matrix propagation based on the unscented transform on the other hand allows each particle to move independently through time and space. The covariance matrix generated from these at any epoch requires no prior and so will accurately portray normally distributed uncertainties within the transition region. The key here is to identify at each epoch, whether the underlying uncertainty space is normally distributed or not.

Unscented Transform based Normalised Offset (UNO)

In astrodynamics, non linear two-body motion causes any normal uncertainty volume to bend in two dimensional space after some time. This deformation from normality is often used in the literature to visualise an important effect of non-linear motion in astrodynamics on state uncertainty divergence from normality [2, 13, 14, 17]. For any covariance extrapolation method that retains information on this bending, the deformation should be quantifiable simply by the offset of the propagated state $\mathbf{z}^{(0)}$ from the propagated covariance mean state $\bar{\mathbf{z}}$. Apart from the bending, this offset should also be able to detect non-symmetry in one-dimensional normal distributions (skewing). Kurtosis on the other hand, which describes a symmetrical deviation pertaining only to the tails of the PDF may remain undetected.

Here, the unscented transform is chosen for its efficiency and ease of implementation as covariance extrapolation method. More specifically, the symmetric set extended by the mean as per Equation (12) of [15, 16] is applied. The sample covariance mean state is determined when the covariance is estimated from the given set of sigma points. Reference [15] defines it as “the weighted average of the transformed points”:

$$\bar{\mathbf{z}} = \sum_{i=0}^{p} W^{(i)} \mathbf{z}^{(i)}.$$  \hspace{1cm} (6)

Herein, $p$ is the number of states generated by the Unscented Transform, $\mathbf{z}^{(i)}$ is the $i^{th}$ state and $W^{(i)}$ is the corresponding weight. The offset used to measure the deviation from normality is labelled NED for normalised Euclidean distance in this paper:

$$\text{NED} = \sqrt{\mathbf{D}} = ||\mathbf{L}^{-1} \left( \mathbf{z}^{(0)} - \bar{\mathbf{z}} \right) ||.$$  \hspace{1cm} (7)

$L^{-1}$ is the inverse lower triangular – obtained through the Cholesky decomposition – of the sample covariance matrix. The multiplication of the vector difference with

---

6For the extended set describing position and velocity in 3-D space, $p = 13$. 

\[ \sum_{i=0}^{p} W^{(i)} \mathbf{z}^{(i)} \]
Fig. 5 Sketch of the idea behind the Unscented Transform based Normalised Offset (UNO) MVN test metric

the inverse lower triangular matrix retrieves the relative vector in whitened space. Its magnitude is the Mahalanobis distance. This normalisation is required to decouple the offset from orbit regime and covariance size. Figure 5 visualises the method.

Calibration

Unlike the Henze-Zirkler test, the NED is not subject to a hypothesis test. Thus, we must determine which NED we believe to yield a reasonable estimate for detection of significant non-normality in the propagated PDF of the state estimation error. Rather than choosing an arbitrary value, here we attempt a calibration by comparison to the Henze-Zirkler test results presented in the previous section. If a strong relationship exists between the two methods, then it should be possible to associate each Henze-Zirkler value exactly one NED. Ideally, this should be possible regardless of the initial covariance or orbit type. Some variability in the results is expected due to the established variability in the Henze-Zirkler mean values. More pronounced deviations may however occur due to the fundamental differences between the two tests as well as short comings in the UNO test to detect certain deviations from normality as outlined in the previous section.

The left hand plot in Fig. 6 plots the mean of the Henze-Zirkler test statistic vs. NED for all cases of Section “Constructed Test Cases”. A strong relation between the two can be seen up to a Henze-Zirkler test value of about 2. Beyond, the relation is no longer independent of orbit or covariance. As we are only interested in transition between normal and non-normal state uncertainties, a trend is fit to the region below the Henze-Zirkler Test value of 1.5. Its intersection with the normality boundary given by the Henze-Zirkler test yields the calibrated decision threshold for UNO which is: \( \text{NED} = 0.0501 \). This value is rounded to 0.05 in the ensuing computations.
Comparing UNO and Henze-Zirkler Test for Constructed Cases

Table 3 contains the epochs for transition onset and conclusion calculated with UNO as well as the Henze-Zirkler test. It can be seen that for all orbit types including highly eccentric, the results of the Unscented Transform based test are well within the $\pm 3\sigma_{\text{CLT}}$ window of the Henze-Zirkler test.

Departure from Normality for Catalogue OD Results

Measurements from passive and active optical measurements of objects in all Earth orbits are being processed in a Batch-Least-Squares orbit determination and ingested into a catalogue. An account of the status of this catalogue is made in the papers by Dr. Michael Lachut et al. [19] and James C.S. Bennett et al. [4]. The OD results assessed here are from mid 2018 and include passive observations only.

The current section should be seen as a proof of concept as at the time of writing the SERC catalogue is still in its build-up phase and the presented efficient method for MVN testing is still in its early testing stage.

Overview of Cases

An overview of example orbit determination results which are assessed in the following is given in Table 4.

Comparing UNO and Henze-Zirkler for OD Cases

Table 5 compares the MVN validity period of OD cases from SERC’s catalogue given in Table 4. The Henze-Zirkler test statistic is averaged from 50 propagated samples.

The results for the transition start show good agreement. The highest deviations between the two methods are observed for cases with short time frames until onset.
Table 3  Beginning and end of transition into non-normality of position uncertainty

| Id       | Transition starts | Transition ends |
|----------|-------------------|-----------------|
|          | UNO               | Henze-Zirkler   | UNO               | Henze-Zirkler   |
|          | Mean              | $\pm 3\sigma_{CLT}$ | Mean              | $\pm 3\sigma_{CLT}$ |
|          | [revs]            | [%]             | [revs]            | [%]             |

**Geosynchronous Orbits**

|               |                  |                  |
|----------------|------------------|------------------|
| 1-1            | 36.0             | 36.0             |
|                | $-12.0/9.4$      |                  |
| 1-2            | 36.0             | 35.3             |
|                | $-14.4/10.4$     |                  |
| 3-1            | 36.0             | 36.6             |
|                | $-14.1/10.9$     |                  |
| 3-2            | 36.0             | 34.6             |
|                | $-16.4/12.6$     |                  |
| 4-1            | 36.0             | 33.7             |
|                | $-16.9/12.9$     |                  |
| 4-2            | 36.0             | 34.7             |
|                | $-15.6/11.7$     |                  |

**Geosynchronous Orbits, Larger Uncertainties**

|               |                  |                  |
|----------------|------------------|------------------|
| 1-1a           | 11.5             | 11.6             |
|                | $-13.3/9.8$      |                  |
| 1-2a           | 11.5             | 11.4             |
|                | $-14.0/11.2$     |                  |
| 3-1a           | 11.5             | 11.5             |
|                | $-14.6/11.1$     |                  |
| 3-2a           | 11.5             | 11.1             |
|                | $-12.9/10.2$     |                  |
| 4-1a           | 11.5             | 11.7             |
|                | $-16.3/13.4$     |                  |
| 4-2a           | 11.5             | 11.2             |
|                | $-15.1/11.5$     |                  |

**Low Earth Orbits**

|               |                  |                  |
|----------------|------------------|------------------|
| 5-1            | 58.0             | 57.4             |
|                | $-14.6/10.7$     |                  |
| 5-2            | 58.0             | 57.1             |
|                | $-21.5/14.5$     |                  |
| 6-1            | 18.5             | 18.6             |
|                | $-15.9/11.0$     |                  |
| 6-2            | 18.5             | 18.2             |
|                | $-15.7/9.4$      |                  |
| 7-1            | 2.4              | 2.1              |
|                | $-17.0/11.1$     |                  |
| 7-2            | 2.4              | 2.1              |
|                | $-40.6/20.0$     |                  |

**Medium Earth Orbits, Near Circular**

|               |                  |                  |
|----------------|------------------|------------------|
| 8-1            | 55.5             | 54.1             |
|                | $-15.1/10.6$     |                  |
| 8-2            | 55.0             | 54.5             |
|                | $-15.7/11.4$     |                  |

**Medium Earth Orbits, Highly Eccentric**

|               |                  |                  |
|----------------|------------------|------------------|
| 9-1            | 47.4             | 47.5             |
|                | $-33.6/22.2$     |                  |
| 9-2            | 47.4             | 56.5             |
|                | $-17.1/16.2$     |                  |
| 10-1           | 47.4             | 52.4             |
|                | $-22.7/20.0$     |                  |
| 10-2           | 47.4             | 55.2             |
|                | $-21.1/22.4$     |                  |

Scenario identifiers refer to those detailed in Table 2. The unit [revs] refers to time after state epoch in revolutions. The Henze-Zirkler test was evaluated at 17 epochs during a revolution except for the highly eccentric cases (9-1 through 10-1), which were evaluated at 100 epochs during a revolution.
Table 4  Overview of orbit determination example cases

| Orbit | GEO | LEO | MEO | HEO |
|-------|-----|-----|-----|-----|
| OD ID | 3715| 3721| 3778| 3783|
| NORAD | 27399| 36287| 17973| 19751|
| Name  | JCSAT 8 | BEIDOU G1 | COSMOS 1844 | ETALON 1 |
|       |       |       |       |       |

of MVN departure. As departure initially occurs only once or twice a revolution, a difference of one revolution between methods yields a higher relative deviation than for cases where the same difference is obtained but at a later departure onset epoch.

The predicted transition end epoch in contrast shows very high deviations in at least three of the nine test cases. The highest deviation is observed for OD ID 3778, but also IDs 3783 and 3769 show deviations which are unlikely to be due to variability in the Henze-Zirkler results, given the markedly better agreement across the other cases. Important differences of the current cases compared to cases of the previous section used for calibration include i) different ratios of position-velocity variances as well as ii) non-zero co-variance terms in the initial covariance matrix. In

Table 5  Beginning and end of transition into non-normality of position uncertainty using Henze-Zirkler test with $n_{mc} = 50$ and UNO with MVN departure decision threshold of 0.05. MVN is evaluated at 35 epochs within each revolution

| OD ID | Transition starts | Transition ends |
|-------|------------------|-----------------|
|       | Henze-Zirkler | UNO | Deviation | Henze-Zirkler | UNO | Deviation |
|       | [days] | [days] | [days] | [%] | [days] | [days] | [days] | [%] |
| Geo synchronous Orbits |
| 3715  | 83.0 | 86.5 | 3.5 | 4.2 | 158.8 | 168.6 | 9.8 | 6.2 |
| 3721  | 126.3 | 129.4 | 3.1 | 2.5 | 220.0 | 213.9 | −6.1 | −2.8 |
| 3778  | 2.4 | 1.6 | −0.8 | −33.3 | 64.3 | 19.1 | −45.2 | −70.3 |
| 3783  | 0.0 | 0.5 | 0.5 | 0.0 | 254.3 | 343.5 | 89.2 | 35.1 |
| Low Earth Orbits |
| 3769  | 29.8 | 25.7 | −4.1 | −13.8 | 68.4 | 89.8 | 21.4 | 31.3 |
| 3753  | 13.6 | 12.5 | −1.1 | −8.1 | 45.8 | 50.3 | 4.5 | 9.8 |
| Medium Earth Orbits, Near Circular |
| 3739  | 97.1 | 96.5 | −0.6 | −0.6 | 126.7 | 127.2 | 0.5 | 0.4 |
| Medium Earth Orbits, Eccentric |
| 3073  | 85.9 | 79.9 | −6.0 | −7.0 | 529.9 | 489.7 | −40.2 | −7.6 |
| 3062  | 0.0 | 0.1 | 0.1 | 0.0 | 0.3 | 0.1 | −0.2 | −66.7 |
Section “Unscented Transform based Normalised Offset (UNO)”, the lacking capability in the current version of UNO in detecting kurtosis has already been identified as a short-coming. Whether the observed differences are due to such an effect has yet to be determined.

Looking at how the results of the two methods differ between constructed test cases (Section “Comparing UNO and Henze-Zirkler Test for Constructed Cases”) and actual OD result cases (this Section), the importance of gaining a sound understanding of capabilities and limitations of any method is highlighted.

A Note on Implementation and Computational Efficiency

The software is implemented in C++ with multithreading support via OpenMP. Compilation is performed with the GNU Compiler Collection release (gcc-9-1.5) for openSUSE Tumbleweed. The majority of underlying mathematical functions required for the MVN tests are implemented based on [25]. The NOVAS library Version C3.1 [3] is used for reference frame conversions. SERC’s internal library libSERC introduced and maintained by Dr. Marek Möckel (SERC) provides additional functionality for reference frame conversions, mathematical functions as well as I/O handling. The hardware used for the following comparison relies on an Intel(R) Core(TM) i7-8700T CPU. The given execution times are the total program run time which includes input file reading and output file writing activities. The comparison is based on evaluation of the state estimation error at a single epoch for the 3D position only. The Henze-Zirkler test is performed using \( n_{mc} = 50 \) Monte-Carlo runs at a sample size of \( n_p = 10\,000 \).

Under these conditions, the Henze-Zirkler test requires just under four minutes in single-thread mode and just over two minutes using 10 threads; this suggests, that the code has room for optimization. The more efficient UNO test takes about seven milliseconds on a single thread; no multithreading support is implemented for this test.

Conclusions

The Henze-Zirkler test for Multivariate Normality is used to test random samples drawn from the covariance matrix representing the state estimation error and propagated through time using two-body motion. A structure is observed in the time evolution of this parameter when Monte-Carlo analysis is employed which underlies all tested LEO, MEO, HEO and GEO cases. After a period of remaining normally distributed over one entire revolution, the state uncertainty enters a phase of transition wherein it will become non-normally distributed for increasing parts throughout a revolution. Finally, the state uncertainty remains non-normally distributed throughout the entire revolution. During this phase, results between samples are most aligned and show a clear structure. After some time, the structure starts to dissolve. Non-normality however remains over the entire revolution.
A more efficient method which operates solely on the states computed using an Unscented Transform has been introduced. Both methods have been applied to state and covariances detailed in a Monte-Carlo study of probability of collision methods by Alfano in 2009 [1] as a baseline as well as orbit determination results from SERC’s own catalogue.

A strong relation between this method and the Henze-Zirkler test has been demonstrated in the relevant region surrounding the state normality decision boundary for all baseline cases and some, but not all, orbit determination result cases. Significant deviations in three out of nine cases in particular in the detection of the MVN departure conclusion exist. Given the fundamental difference of the two methods, deviations are to be expected. The source of these deviations is not yet fully understood however and requires further investigation.

Acknowledgments The authors would like to acknowledge the support of the Cooperative Research Centre for Space Environment Management (SERC Limited) through the Australian Government’s Cooperative Research Centre Programme. Additional thanks go out to Dr. Michael Lachut of Electro Optic Systems Pty Ltd. for his input and efforts in all things relating to orbit determination. Furthermore, thanks go out to PhD student Hansani Kaushalya Perera of University of South Australia for offering her insight in the field of statistics. Thanks also go out to Carlos Yanez of CNES for his independent review of the original conference publication. Finally, the authors would like to thank the journal reviewers for their keen insight and constructive criticism.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Alfano, S.: Satellite conjunction monte carlo analysis. AAS Spaceflight Mechanics Mtg, Pittsburgh, PA., Paper, 09–233 (2009)
2. Alfriend, K.T., Park, I.: When does the uncertainty become non-Gaussian. Adv. Maui Opt. Sp. Survell. Technol. Conf. (2016)
3. Bangert, J., Puatua, W., Kaplan, G., Bartlett, J., Harris, W., Fredericks, A., Monet, A.: User’s guide to NOVAS version C3.1 - Naval observatory vector astrometry software. User’s guide, U.S. Nautical almanac office (2011)
4. Bennett, J.C., Lachut, M., Kucharski, D., Flegel, S.K., Möckel, M., Allworth, J., Kooymans, D., Pollard, A., Smith, C., O’Leary, J., Kankanamalage, H.T., Samuel, R.: Progress in a new conjunction and threat warning service for space situational awareness. In: Proceedings of the 19th Advanced Maui Optical Space Surveillance Technologies Conference, Maui, Hawaii (2018)
5. Coppola, V.T.: Including velocity uncertainty in the probability of collision between space objects. Adv. Astronaut. Sci. - AAS 12-247 143(i), 2159–2178 (2012)
6. Farrell, P.J., Salibian-Barrera, M., Naczk, K.: On tests for multivariate normality and associated simulation studies. J. Stat. Comput. Simul. 77(12), 1065–1080 (2007). https://doi.org/10.1080/10629360600878449
7. Flegel, S.K., Bennett, J.C.: Normality in state uncertainties from orbit determination results fitting optical measurements. In: Proceedings of the 19th Advanced Maui Optical Space Surveillance Technologies Conference, Maui, Hawaii (2018)
8. Flegel, S.K., Möckel, M., Bennett, J.: Covariance size and the breakdown of Gaussianity in GEO uncertainty predictions. Proceedings of the 7th European Space Debris Conference (2017)
9. Foster, J.L.: The analytic basis for debris avoidance operations for the international space station. Eur. Sp. Agency, (Special Publ. ESA SP) 1(473), 441–445 (2001)

10. Gupta, M.: Analysis of the representation of orbital errors and improvement of their modelling. Lulea University of Technology, Department of Computer Science, Electrical and Space Engineering. Master’s thesis (2018)

11. Henze, N., Zirkler, B.: A class of invariant consistent tests for multivariate normality. Commun. Stat. - Theory Methods 19(10), 3595–3617 (1990). https://doi.org/10.1080/03610929008830400. http://www.tandfonline.com/doi/abs/10.1080/03610929008830400

12. Hogg, R.V., Tanis, E.A., Zimmerman, D.L. Probability and Statistical Inference, 9th edn. Pearson Education, London (2015). ISBN-13: 978-0-321-92327-1

13. Horwood, J.T., Poore, A.B.: Orbital state uncertainty realism. Adv. Maui Opt. Sp. Survell. Technol. Conf. (2012)

14. Jones, B.A., Doostan, A., Born, G.H.: Nonlinear propagation of orbit uncertainty using non-intrusive polynomial chaos. J. Guid. Control. Dyn. 36(2), 430–444 (2013). https://doi.org/10.2514/1.57599. http://arc.aiaa.org/doi/abs/10.2514/1.57599

15. Julier, S., Uhlmann, J.: Unscented filtering and non linear estimation. Proc. IEEE 92(3), 401–422 (2004). https://doi.org/10.1109/JPROC.2003.823141

16. Julier, S.J., Uhlmann, J.K.: Correction to “Unscented filtering and nonliner estimation” Proc. IEEE 92(3), 401–422 (2004)

17. Junkins, J.L., Akella, M.R., Alfriend, K.T.: Non-gaussian error propagation in orbital mechanics. J. Astronaut. Sci. 44(4), 541–563 (1996)

18. Krier, G.: Satellite collision probability for Long-Term encounters and arbitrary primary satellite shape. In: Proceedings of the Seventh European Conference on Space Debris held in Darmstadt, Germany, April (2017)

19. Lachut, M., Bennett, J.C.: Automating the assessment of orbit predictions and estimations for building and maintaining a new catalogue. In: Proceedings of the 19th Advanced Maui Optical Space Surveillance Technologies Conference, Maui, Hawaii (2018)

20. Mardia, K.: Measures of multivariate skewness and kurtosis with applications. Biometrika 57, 519–530 (1970)

21. Mecklin, C.J., Mundfrom, D.J.: A Monte Carlo comparison of the Type I and Type II error rates of tests of multivariate normality. J. Stat. Comput. Simul. 75(2), 93–107 (2005). https://doi.org/10.1080/0094965042000193233. http://www.tandfonline.com/doi/abs/10.1080/0094965042000193233

22. Oltrogge, D.L.: Efficient solutions of Kepler’s equation via hybrid and digital approaches. AIAA/AAS Astrodyn. Spec. Conf. - AIAA, pp. 401–420 (2014)

23. Patera, R.P.: General method for calculating satellite collision probability. Journal of Guidance Conrol and Dynamics 24(4) (2001)

24. Patera, R.P.: Satellite collision probability for nonlinear relative motion. J. Guid. Control. Dyn. 26(5), 728–733 (2003). https://doi.org/10.2514/2.5127. http://doi.aiaa.org/10.2514/2.5127

25. Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P. Numerical Recipes - The Art of Scientific Computing, 3rd edn. Cambridge University Press, Cambridge (2007)

26. Romeu, J., Oztur, A.: A comparative study of goodness-of-fit tests for multivariate normality. J. Multivar. Anal. 46, 309–334 (1993)

27. Siminski, J.: Object correlation and orbit determination for geostationary satellites using optical measurements. Dissertation, Universität der Bundeswehr München. https://elib.dlr.de/109770/1/siminski-dissertation_online.pdf (2016)

28. Tapley, B.D., Schutz, B.E., Born, G.H.: Statistical Orbit Determination. Academic Press Inc. ASIN B017N09F9B (2004)

29. Yanez, C., Gupta, M., Morand, V., Dolado, J.: On the Gaussianity validity time for orbital uncertainty propagation. ESA NEO and debris detection conference, Darmstadt, Germany, January 2019 (2019)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.