(Preprint) AAS 25-412

LINEAR COVARIANCE FIDELITY CHECKS AND MEASURES OF NON-GAUSSIANITY

Jackson Kulik[*], Braden Hastings[†], and Keith A. LeGrand[‡]

Linear covariance (LinCov) techniques have gained widespread traction in the modeling of uncertainty, including in the preliminary study of navigation performance. While LinCov methods offer improved computational efficiency compared to Monte Carlo based uncertainty analysis, they inherently rely on linearization assumptions. As the study of navigation in Cislunar space grows, it is important to assess the degree to which linear covariance techniques perform under higher dynamical nonlinearity and large state uncertainties. We present and test a number of computational techniques for assessing linear covariance performance using tools from higher-order statistics and constrained optimization.

INTRODUCTION

Consider a nonlinear function $\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^m$ that may represent the flow of a dynamical system or a measurement function. Let $\mathbf{x} \in \mathbb{R}^n$ random variable with mean $\boldsymbol{\mu}_x$ and covariance $\mathbf{P}_x$. Linear covariance techniques rely on the following simple principle for uncertainty propagation: for sufficiently small input covariance $\mathbf{P}_x$, the resulting output mean and covariance are approximately

$$\mathrm{E}[\mathbf{z}] \approx \boldsymbol{\mu}_z^{(1)} = \mathbf{g}(\boldsymbol{\mu}_x) \tag{1}$$

$$\mathbf{P}_z \approx \mathbf{P}_z^{(1)} = \mathbf{G}\mathbf{P}_x\mathbf{G}^T \tag{2}$$

where the superscript "(1)" denotes that these quantities arise from linear covariance analysis and

$$\mathbf{G} = \frac{\mathrm{d}\mathbf{g}}{\mathrm{d}\mathbf{x}}\Big|_{\boldsymbol{\mu}_x} \tag{3}$$

is the Jacobian of $\mathbf{g}$ evaluated at $\boldsymbol{\mu}_x$. Furthermore, if $\mathbf{x}$ is Gaussian-distributed, then by the same aforementioned assumptions, $\mathbf{z}$ is also approximately Gaussian distributed. Often, linear covariance techniques are employed for stochastic systems that include process noise or measurement noise. In this work, we will examine only deterministic systems, where the dominating source of output uncertainty is due the initial uncertainty in $\mathbf{x}$.

This paper develops new methods and metrics for analyzing the quality of LinCov-estimated quantities in nonlinear systems. The methods in this work are based primarily on calculating expectations related to linearization error or solving constrained optimization problems related to linearization error. In both cases, the proposed methods leverage higher-order statistics. For multivariate random variables, central moments beyond order two require tensor representations. Thus,

[*]Assistant Professor, Mechanical and Aerospace Engineering, Utah State University, Logan UT.
[†]PhD Student, Mechanical and Aerospace Engineering, Utah State University, Logan UT.
[‡]Assistant Professor, School of Aeronautics and Astronautics, Purdue University, West Lafayette, IN.

our constrained optimization approach leverages the theory of tensor eigenvalues and tensor operator norms.[1,2] One of the most important elements of our approach is output-whitening, which enables each output to be meaningfully interpretable as a Mahalanobis distance and is unit and scale-invariant.

BACKGROUND

Partial Derivative Tensors

In this work, we will heavily rely on higher-order partial derivative tensors of the nonlinear function $\mathbf{g}$ in order to describe the higher-order terms in the Taylor series expansion of $\mathbf{g}$ about the mean of $\boldsymbol{\mu}_x$ as these will approximately describe the error in approximating $\mathbf{g}(\boldsymbol{\mu}_x+\delta\mathbf{x})$ by its linearization $\mathbf{g}(\boldsymbol{\mu}_x) + \mathbf{G}\delta\mathbf{x}$. The Taylor series expansion of $\mathbf{g}$ about the mean $\boldsymbol{\mu}_x$ is given by

$$\mathbf{g}(\boldsymbol{\mu}_x + \delta\mathbf{x}) = \mathbf{g}(\boldsymbol{\mu}_x) + \mathbf{G}\delta\mathbf{x} + \frac{1}{2}\frac{\partial^2\mathbf{g}}{\partial\mathbf{x}^2}\Big|_{\boldsymbol{\mu}_x}\delta\mathbf{x}^2 + \mathcal{O}(\delta\mathbf{x}^3) \tag{4}$$

where we define the shorthand notation above for double contraction of the tensor with two copies of a vector in terms of the following Einstein notation (where there is a sum over any repeated indices in an expression)

$$\left(\frac{\partial^2\mathbf{g}}{\partial\mathbf{x}^2}\Big|_{\boldsymbol{\mu}_x}\delta\mathbf{x}^2\right)^i = \left(\frac{\partial^2\mathbf{g}}{\partial\mathbf{x}^2}\Big|_{\boldsymbol{\mu}_x}\right)^i_{j,k}\delta\mathbf{x}^j\delta\mathbf{x}^k \tag{5}$$

In the following, we adopt the abbreviated notation

$$\mathbf{G}^{(2)} = \frac{\partial^2\mathbf{g}}{\partial\mathbf{x}^2}\Big|_{\boldsymbol{\mu}_x} \tag{6}$$

to denote the second-order partial derivative tensor. We will primarily be concerned with partial derivatives associated with the flow of a dynamical system. The partial derivatives of the flow of a dynamical system are known as state transition tensors and can be computed by deriving and integrating variational equations associated with the dynamical system[3] or by using differential algebra techniques.[4]

Whitening and Mahalanobis Distance

Given a random variable $\mathbf{x}$ and its covariance matrix $\mathrm{var}(\mathbf{x}) = \mathbf{P}$, a whitening transform is defined by any transformation

$$\mathbf{y} = \mathbf{W}\mathbf{x} \tag{7}$$

such that $\mathbf{y}$ is the same dimension of of $\mathbf{x}$ and the "whitened" covariance matrix $\mathrm{var}(\mathbf{y}) = \boldsymbol{I}$. It is straightforward to show that the inverse of any square root factor $\mathbf{W} = \mathbf{P}^{-1/2}$, where $\mathbf{P} = \mathbf{P}^{1/2}(\mathbf{P}^{1/2})^T$ provides a whitening transformation. A convenient property of the the whitened vector $\mathbf{y}$ is that its 2-norm is equal to the Mahalanobis distance in the original space:

$$\sqrt{\mathbf{y}^T\mathbf{y}} = \sqrt{\mathbf{x}^T\mathbf{P}^{-1}\mathbf{x}} \tag{8}$$

Moments

Denote the linearly propagated distribution of $\mathbf{x}$ by

$$\mathbf{z}^{(1)} = \mathbf{g}(\boldsymbol{\mu}_x) + \mathbf{G}(\mathbf{x} - \boldsymbol{\mu}_x) \tag{9}$$

The statistics of $\mathbf{z}^{(1)}$ are described completely by the mean and covariance in Eqs. (1) and (2), respectively. The third-order central moments of this Gaussian distribution along any direction are zero, so

$$(\mathbf{S}_z^{(1)})^{i,j,k} = \mathrm{E}[(\mathbf{z}^{(1)} - \boldsymbol{\mu}_z^{(1)})^i(\mathbf{z}^{(1)} - \boldsymbol{\mu}_z^{(1)})^j(\mathbf{z}^{(1)} - \boldsymbol{\mu}_z^{(1)})^k] = 0 \tag{10}$$

In general, the standardized third-order central moment–known as *skewness*–of a distribution is given by the third-order central moment of the random variable that results from whitening the original random variable

$$\boldsymbol{\Sigma}^{i,j,k} = (\mathbf{P}^{-1/2})^i_{i'}(\mathbf{P}^{-1/2})^j_{j'}(\mathbf{P}^{-1/2})^k_{k'}\mathbf{S}^{i',j',k'} \tag{11}$$

this is also zero in the case of $\mathbf{z}^{(1)}$. The fourth order central moment tensor of $\mathbf{z}^{(1)}$ is defined as

$$\begin{aligned}
(\mathbf{K}_z^{(1)})^{i,j,k,l} &= \mathrm{E}[(\mathbf{z}^{(1)} - \boldsymbol{\mu}_z^{(1)})^i(\mathbf{z}^{(1)} - \boldsymbol{\mu}_z^{(1)})^j(\mathbf{z}^{(1)} - \boldsymbol{\mu}_z^{(1)})^k)(\mathbf{z}^{(1)} - \boldsymbol{\mu}_z^{(1)})^l] && (12)\\
&= (\mathbf{P}_z^{(1)})^{i,j}(\mathbf{P}_z^{(1)})^{k,l} + (\mathbf{P}_z^{(1)})^{i,k}(\mathbf{P}_z^{(1)})^{j,l} + (\mathbf{P}_z^{(1)})^{i,l}(\mathbf{P}_z^{(1)})^{j,k} && (13)\\
&= \mathrm{sym}\left(3(\mathbf{P}_z^{(1)})^{i,j}(\mathbf{P}_z^{(1)})^{k,l}\right) && (14)
\end{aligned}$$

where the second equality known as Isserlis' theorem can be obtained via the joint characteristic function[3,5] or by way of a symmetry argument, inverse whitening transformation, and by counting the number of pairs that can be selected from four indices. The sym operator denotes averaging the tensor over all permutations of the indices to symmetrize the tensor

$$\mathrm{sym}(\mathbf{T})^{i_1,\ldots,i_m} = \frac{1}{m!}\sum_{\sigma\in S_m}\mathbf{T}^{\sigma(i_1,\ldots,i_m)} \tag{15}$$

The standardized fourth order central moment tensor or *kurtosis* tensor is obtained by applying a whitening transformation to $\mathbf{z}^{(1)}$ and taking the fourth central moment of the resulting random variable:

$$(\boldsymbol{\kappa}_z^{(1)})^{i,j,k,l} = ((\mathbf{P}_z^{(1)})^{-1/2})^i_{i'}((\mathbf{P}_z^{(1)})^{-1/2})^j_{j'}((\mathbf{P}_z^{(1)})^{-1/2})^k_{k'}((\mathbf{P}_z^{(1)})^{-1/2})^l_{l'}(\mathbf{K}_z^{(1)})^{i',j',k',l'} = 3(\mathbf{I}_m^{(4)})^{i,j,k,l} \tag{16}$$

where $\mathbf{I}_m^{(4)}$ is the fourth-order, m-dimensional identity tensor[6,7] such that

$$\mathbf{I}_m^{(4)}\boldsymbol{\xi}^3 = \boldsymbol{\xi} \qquad \text{and} \qquad \mathbf{I}_m^{(4)}\boldsymbol{\xi}^4 = 1 \tag{17}$$

for all $\boldsymbol{\xi} \in \mathbb{R}^m$ such that $\|\boldsymbol{\xi}\|_2 = 1$. The relationship between the Gaussian kurtosis and the identity tensor is a novel observation as far as the authors are aware. We defien the *excess kurtosis* tensor as the difference between the kurtosis $\boldsymbol{\kappa}$ of some distribution and the kurtosis of a Gaussian distribution of the same dimension. That is,

$$\delta\boldsymbol{\kappa} = \boldsymbol{\kappa} - 3\mathbf{I}^{(4)} \tag{18}$$

The skewness and kurtosis tensors as defined here are the central moments of the whitened random variable $\mathbf{y}$ such that

$$(\mathbf{y} - \boldsymbol{\mu}_y) = \mathbf{P}_x^{-1/2}(\mathbf{x} - \boldsymbol{\mu}_x) \tag{19}$$

where $\mathbf{x}$ is the original random variable under consideration. Note that an mth-order central moment tensor that is not standardized (such as $\mathbf{S}$) can be used to find the corresponding mth-order central moment of a marginal distribution along any particular direction (such as the skewness of the marginal distribution along $\hat{\mathbf{v}}$) by contraction m times with a unit vector along that direction. On the other hand, repeated contraction along a particular direction of the standardized central moment tensors (skewness $\boldsymbol{\Sigma}$ and kurtosis $\boldsymbol{\kappa}$) does not necessarily result in the standardized skewness and kurtosis of the marginal distribution under contraction. For example,

$$\Sigma^{i,j,k}\hat{\mathbf{v}}^i\hat{\mathbf{v}}^j\hat{\mathbf{v}}^k \neq \frac{S^{i,j,k}\hat{\mathbf{v}}^i\hat{\mathbf{v}}^j\hat{\mathbf{v}}^k}{(\hat{\mathbf{v}}^T\mathbf{P}\hat{\mathbf{v}})^{3/2}} = \frac{\mathrm{E}[(\hat{\mathbf{v}}^T(\mathbf{x}-\boldsymbol{\mu})]^3}{\sigma_{\hat{\mathbf{v}}}^3} \tag{20}$$

where $\sigma_{\hat{\mathbf{v}}}$ is the standard deviation of the marginal distribution along an arbitrary unit vector $\hat{\mathbf{v}}$. Special cases in which the two expressions are equivalent include when the distribution is a homoscedastic multivariate Gaussian and when the distribution is a multivariate Gaussian marginalized such that $\hat{\mathbf{v}}$ is along one of the principal axes of the associated covariance ellipsoid. This is all to say that one must be careful not to conflate the standardized skewness tensor contracted along a particular direction with the marginal skewness along that direction. Instead, the standardized skewness tensor contracted along a particular direction should be interpreted as the skewness of the marginal distribution in the given direction of the whitened random variable.

Tensor Eigenvalues

Given an mth-order supersymmetric (symmetric under any permutation of the indices) covariant tensor (a multilinear functional that operates on m vectors to produce a scalar value) $\mathbf{T}$, consider the following constrained optimization problem:

$$\max_{\|\mathbf{x}\|_2=1} \mathbf{T}_{i_1,\ldots,i_m}\mathbf{x}^{i_1}\ldots\mathbf{x}^{i_m} \tag{21}$$

The method of Lagrange multipliers may be applied to this constrained optimization to yield the following conditions for optimality:

$$\mathbf{T}\mathbf{x}^{m-1} = \lambda\mathbf{x} \tag{22}$$

$$\|\mathbf{x}\|_2 = 1 \tag{23}$$

The above equations define a z-eigenvalue and eigenvector pair. The theory[8,9] behind tensor z-eigenvalues and their computation[6,10] has recently been adopted by the astrodynamics community to study nonlinearity and to quantify the error associated with linear methods for guidance, navigation, and control.[1,2,11] In order to calculate the maximal z-eigenvector pair, an algorithm known as shifted higher-order power iteration may be employed. Shifted higher-order power iteration is a generalization of shifted power iteration for finding eigenvalues of symmetric matrices. The algorithm is described by the iteration

$$\mathbf{x}_{(k+1)} = \frac{\mathbf{T}\mathbf{x}_{(k)}^{m-1} + \alpha\mathbf{x}_{(k)}}{\|\mathbf{T}\mathbf{x}_{(k)}^{m-1} + \alpha\mathbf{x}_{(k)}\|_2} \tag{24}$$

where the subscript (k) indicates the kth iterate of the algorithm and α is chosen to ensure global convergence of the algorithm to an eigenpair of the tensor $\mathbf{T}$. We use the notation $\mathbf{T}\mathbf{x}^{m-1}$ as shorthand for the the following operation that produces a vector

$$(\mathbf{T}\mathbf{x}^{m-1})_{i_1} = T_{i_1,i_2,\ldots,i_m}x^{i_2}\ldots x^{i_m} \tag{25}$$

The shift parameter α must be chosen to be larger than the order of the tensor minus one multiplied with the spectral radius of the matrix resulting from contracting all but 2 of the indices of the tensor with any vector on the unit sphere:

$$\alpha > (m-1) \max_{\|\mathbf{x}\|_2=1} \rho(\mathbf{T}\mathbf{x}^{m-2}) \tag{26}$$

One simple method to compute a sufficient shifting factor α is given by the sum of the absolute values of the entries of the tensor[6]

$$\alpha = (m-1) \sum_{i_1,\ldots,i_m} |T_{i_1,\ldots,i_m}| \tag{27}$$

This algorithm will converge regardless of the initial guess for the eigenvector, and will tend to converge to eigenpairs with the largest eigenvalues, though is not guaranteed to converge to the eigenpair with the largest eigenvalue from every initial guess. As such, if there is not a good method for generating an initial guess at the eigenvector corresponding to the largest eigenvalue, then the algorithm should be run with a number of random initial guesses.

BASELINE MONTE CARLO APPROACHES

First, we will review a number of baseline metrics that we can employ using Monte Carlo based techniques. These will not be the most efficient, and performing a Monte Carlo defeats the purpose of using linear covariance analysis in the first place. However, these metrics will offer a standard for comparison against with the more computationally efficient methods proposed here. By sampling the initial Gaussian distribution $p(\mathbf{x})$ and propagating each sample through the nonlinear function $\mathbf{g}$ we can obtain samples of the distribution of $\mathbf{z} = \mathbf{g}(\mathbf{x})$. We denote samples of the final distribution by

$$\mathbf{z}_{(i)} = \mathbf{g}(\mathbf{x}_{(i)}) \tag{28}$$

for i from 1 to N where $\{\mathbf{x}_{(i)}\}_{i=1}^{N} \sim p(\mathbf{x})$. With samples of the final distribution we can compare against the linearly propagated distribution of

$$\mathbf{z}^{(1)} = \mathbf{g}(\boldsymbol{\mu}_x) + \mathbf{G}(\mathbf{x} - \boldsymbol{\mu}_x) \tag{29}$$

Differences and Ratios of Moments As a simple test, we can examine the moments of the true random variable $\mathbf{z}$ and the linearly propagated random variable $\mathbf{z}^{(1)}$. The absolute difference between the two means can be taken as a metric of interest, or can be normalized to a squared Mahalanobis distance according to

$$(\boldsymbol{\mu}_z^{(\mathrm{MC})} - \boldsymbol{\mu}_z^{(1)})^T (\mathbf{P}_z^{(1)})^{-1} (\boldsymbol{\mu}_z^{(\mathrm{MC})} - \boldsymbol{\mu}_z^{(1)}) \tag{30}$$

This is depicted in Fig. 1. The concept of moment comparison can be extended to the second moments. The minimum and maximum values of the generalized Rayleigh quotient

$$\frac{\boldsymbol{\zeta}^T (\mathbf{P}^{(1)})^{-1} \boldsymbol{\zeta}}{\boldsymbol{\zeta}^T (\mathbf{P}^{(\mathrm{MC})})^{-1} \boldsymbol{\zeta}} \tag{31}$$

can be employed to assess the difference between the two distributions. This generalized Rayleigh quotient can be interpreted as the minimal/maximal linear covariance Mahalanobis distance of the

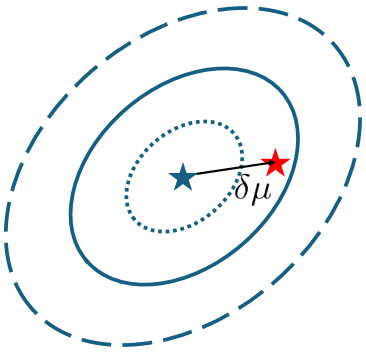

Figure 1. Notional depiction of the contours of equal Mahalanobis distance with a difference in two means superimposed.

1-sigma contour of the Gaussian fitted from the sample mean and covariance of the distribution. The minimum and maximum values of this generalized Rayleigh quotient can be calculated by solving the generalized eigenvalue problem with

$$(\mathbf{P}^{(1)})^{-1}\boldsymbol{\zeta} = \lambda(\mathbf{P}^{(\mathrm{MC})})^{-1}\boldsymbol{\zeta} \tag{32}$$

and finding the minimum and maximum eigenvalue which will give the minimum and maximum ratios. Finally, the sample skewness $\boldsymbol{\Sigma}^{\mathrm{MC}}$ from Eq. (11) and excess sample kurtosis $\boldsymbol{\kappa}^{\mathrm{MC}}$ from Eq. (18) can be employed as measures of non-Gaussianity. We will present more analysis of the generalized eigenvalues of the precision matrices and on the skewness and kurtosis tensors in the next section where take an analytical approach to approximating the moments of the distribution to second-order.

Expected Squared Mahalanobis Distance of True Distribution The expected squared Mahalanobis distance of a random variable $\boldsymbol{\zeta}$ with respect to the linearly propagated mean and covariance is

$$\mathrm{ESMD}_{\mathbf{g},\mathbf{x}} = \mathrm{E}[(\boldsymbol{\zeta} - \boldsymbol{\mu}_z^{(1)})^T(\mathbf{P}^{(1)})^{-1}(\boldsymbol{\zeta} - \boldsymbol{\mu}_z^{(1)})] \tag{33}$$

Given the uniformly weighted samples $\{\boldsymbol{\zeta}_{(i)}\}_{i=1}^N$, the empirical expectation is found via

$$\mathrm{ESMD}_{\mathbf{g},\mathbf{x}}^{\mathrm{MC}} = \frac{1}{N}\sum_i (\boldsymbol{\zeta}_{(i)} - \boldsymbol{\mu}_z^{(1)})^T(\mathbf{P}^{(1)})^{-1}(\boldsymbol{\zeta}_{(i)} - \boldsymbol{\mu}_z^{(1)}) \tag{34}$$

Using the samples $\mathbf{z}_{(i)}$ gives an approximation of the expectation of the squared Mahalanobis distance of the true distribution. If the true distribution is close to the linearly propagated distribution, then the expected squared Mahalanobis distance should be given by m, the dimension of $\mathbf{z}$, which is the expectation of a Chi-square random variable with m degrees of freedom. Eq. 33 can also be written in tensor notation

$$\mathrm{ESMD}_{\mathbf{g},\mathbf{x}} = \mathrm{E}[(\mathbf{P}^{(1)})_{ij}^{-1}(\boldsymbol{\zeta} - \boldsymbol{\mu}_z^{(1)})^j(\boldsymbol{\zeta} - \boldsymbol{\mu}_z^{(1)})^i] \tag{35}$$

Because $(\mathbf{P}^{(1)})_{ij}^{-1}$ is constant, we can pull it outside the expectation

$$\begin{aligned}\mathrm{ESMD}_{\mathbf{g},\mathbf{x}} &= (\mathbf{P}^{(1)})_{ij}^{-1}\mathrm{E}[(\boldsymbol{\zeta} - \boldsymbol{\mu}_z^{(1)})^j(\boldsymbol{\zeta} - \boldsymbol{\mu}_z^{(1)})^i] \\ &= (\mathbf{P}^{(1)})_{ij}^{-1}\mathrm{E}[(\boldsymbol{\zeta} - \boldsymbol{\mu}_\zeta + \boldsymbol{\mu}_\zeta - \boldsymbol{\mu}_z^{(1)})^i(\boldsymbol{\zeta} - \boldsymbol{\mu}_\zeta + \boldsymbol{\mu}_\zeta - \boldsymbol{\mu}_z^{(1)})^j] \\ &= (\mathbf{P}^{(1)})_{ij}^{-1}(\mathbf{P}_\zeta^{ij} + \delta\boldsymbol{\mu}^j\delta\boldsymbol{\mu}^i)\end{aligned} \tag{36}$$

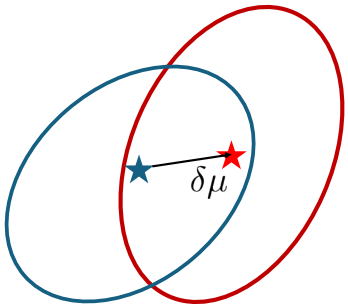

Figure 2. Notional depiction of two distributions with distinct means and covariance ellipsoids.

where $\delta\boldsymbol{\mu} = \boldsymbol{\mu}_\zeta - \boldsymbol{\mu}_z^{(1)}$. The first expression above can be reformulated in the notation of matrix algebra as follows

$$(\mathbf{P}^{(1)})^{-1}_{ij}(\mathbf{P}_\zeta)^{ij} = \mathrm{Tr}((\mathbf{P}^{(1)})^{-1}\mathbf{P}_\zeta) \tag{37}$$

Thus,

$$\mathrm{ESMD}_{\mathbf{g},\mathbf{x}} = \mathrm{Tr}((\mathbf{P}^{(1)})^{-1}\mathbf{P}_\zeta) + \delta\boldsymbol{\mu}^T(\mathbf{P}^{(1)})^{-1}\delta\boldsymbol{\mu} \tag{38}$$

and this can be computed from samples by directly employing the sample covariance and the difference between the sample mean and the propagated mean. The separation of the two terms in the equation above provides insight into the effects of the difference in means on the expected Mahalanobis distance and the effects of the difference in distribution about the respective means. The $\mathrm{Tr}((\mathbf{P}^{(1)})^{-1}\mathbf{P}_\zeta)$ term in the equation above can be interpreted as the expected Mahalanobis distance if the means of the random variable and the linear propagation are the same. The $\delta\boldsymbol{\mu}^T(\mathbf{P}^{(1)})^{-1}\delta\boldsymbol{\mu}$ term is the expected Mahalanobis distance from the difference in means between the true random variable and the linear propagation. In Fig. 2, the two different effects accounted for in expected squared Mahalanobis distance are displayed: a difference in means and a difference in covariances. It is worth noting that the expected squared Mahalanobis distance contains many of the terms present in the Kullback-Leibler divergence of two Gaussian random variables, missing only a constant term subtracting the dimension of the Gaussian random variables and another adding the logarithm of the ratios of the determinants of the two distributions.

Expected Squared Mahalanobis Distance of Linearization Error The linear prediction from any given input data point $\mathbf{x}_{(i)}$ is

$$\mathbf{z}^{(1)}_{(i)} = \mathbf{z}^{(1)}(\mathbf{x}_{(i)}) = \mathbf{g}(\boldsymbol{\mu}_x) + \mathbf{G}(\mathbf{x}_{(i)} - \boldsymbol{\mu}_x) \tag{39}$$

and the linearization error for that sample is

$$\delta\mathbf{z}_{(i)} = \mathbf{z}^{(MC)}_{(i)} - \mathbf{z}^{(1)}_{(i)} \tag{40}$$

so the expectation of the squared Mahalanobis distance of the linearization error is approximated by

$$\mathrm{ESMDoLE}^{\mathrm{MC}}_{\mathbf{g},\mathbf{x}} = \frac{1}{N}\sum_i (\delta\mathbf{z}_{(i)})^T(\mathbf{P}^{(1)})^{-1}\delta\mathbf{z}_{(i)} \tag{41}$$

Expectation-Based Analytical Approaches

Expected Squared Mahalanobis Distance of Linearization Error Let

$$\delta\mathbf{x} = \mathbf{x} - \boldsymbol{\mu}_x \tag{42}$$

The error in the approximation of $\mathbf{g}(\mathbf{x})$ with $\mathbf{g}(\boldsymbol{\mu}_x) + \mathbf{G}\delta\mathbf{x}$ will be dominated by a second-order contribution assuming $\delta\mathbf{x}$ is sufficiently small. This second-order contribution is given by the second term in the Taylor series

$$\mathbf{g}(\boldsymbol{\mu}_x + \mathbf{x}) \approx \mathbf{g}(\boldsymbol{\mu}_x) + \mathbf{G}\delta\mathbf{x} + \frac{1}{2}\mathbf{G}^{(2)}\delta\mathbf{x}^2 \tag{43}$$

where

$$\mathbf{G}^{(2)} = \frac{\partial^2 \mathbf{g}}{\partial \mathbf{x}^2}\Big|_{\boldsymbol{\mu}_x} \tag{44}$$

and the double contraction in shorthand is explicitly in Einstein notation

$$(\mathbf{G}^{(2)}\delta\mathbf{x}^2)^i = (\mathbf{G}^{(2)})^i_{j,k}\delta\mathbf{x}^j\delta\mathbf{x}^k \tag{45}$$

The squared Mahalanobis distance in terms of the final linearized covariance prediction of the linearization error is then

$$\frac{1}{4}((\mathbf{P}_z^{(1)})^{-1})_{i_1,i_2}(\mathbf{G}^{(2)})^{i_1}_{j_1,k_1}\delta\mathbf{x}^{j_1}\delta\mathbf{x}^{k_1}(\mathbf{G}^{(2)})^{i_2}_{j_2,k_2}\delta\mathbf{x}^{j_2}\delta\mathbf{x}^{k_2} \tag{46}$$

which has expectation (Expected Squared Mahalanobis Distance of Linearization Error) given by

$$\mathrm{ESMDoLE}_{\mathbf{g},\mathbf{x}} = \frac{1}{4}((\mathbf{P}_z^{(1)})^{-1})_{i_1,i_2}(\mathbf{G}^{(2)})^{i_1}_{j_1,j_2}(\mathbf{G}^{(2)})^{i_2}_{j_3,j_4}(\mathbf{K}_x)^{j_1,j_2,j_3,j_4} \tag{47}$$

where $\mathbf{K}_x$ is the fourth-order central moment tensor for the Gaussian random variable $\mathbf{x}$. This is given in terms of the covariance of the random variable $\mathbf{x}$ as

$$\begin{aligned}(\mathbf{K}_x)^{j_1,j_2,j_3,j_4} &= \mathrm{E}[\delta\mathbf{x}^{j_1}\delta\mathbf{x}^{j_2}\delta\mathbf{x}^{j_3}\delta\mathbf{x}^{j_4}] &(48)\\ &= \mathrm{sym}\left(3(\mathbf{P}_x)^{j_1,j_2}(\mathbf{P}_x)^{j_3,j_4}\right) &(49)\end{aligned}$$

Maximal Ratio Between First- and Second-Order Covariance Ellipsoids The covariance of the random variable $\mathbf{z}$ as determined up to a linear approximation of the function $\mathbf{g}$ is given by

$$\mathbf{P}_z^{(1)} = \mathbf{G}\mathbf{P}_x\mathbf{G}^T \tag{50}$$

An approximation of the covariance matrix up to second-order is given by[3]

$$(\mathbf{P}_z^{(2)})^{i_1,i_2} = (\mathbf{G}\mathbf{P}_x\mathbf{G}^T)^{i_1,i_2} + \frac{1}{4}(\mathbf{G}^{(2)})^{i_1}_{j_1,j_2}(\mathbf{G}^{(2)})^{i_2}_{j_3,j_4}(\mathbf{K}_x)^{j_1,j_2,j_3,j_4} \tag{51}$$

We may examine the ratio of equal likelihood covariance ellipsoids along a given direction using the generalized Rayleigh quotient

$$\frac{\boldsymbol{\zeta}^T(\mathbf{P}_z^{(1)})^{-1}\boldsymbol{\zeta}}{\boldsymbol{\zeta}^T(\mathbf{P}_z^{(2)})^{-1}\boldsymbol{\zeta}} \tag{52}$$

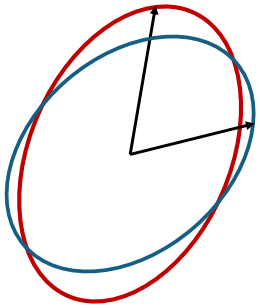

Figure 3. Notional depiction of minimal and maximal ratios of two ellipses. Note that the two directions are not orthogonal.

The generalized Rayleigh quotient is minimized or maximized by solving the generalized eigenvalue problem with

$$(\mathbf{P}_z^{(1)})^{-1}\boldsymbol{\zeta} = \lambda(\mathbf{P}_z^{(2)})^{-1}\boldsymbol{\zeta} \tag{53}$$

and finding the minimum and maximum eigenvalue which correspond to these minimum and maximum ratios. Again, this can be interpreted as giving the maximum and minimum $\mathbf{P}^{(1)}$ based Mahalanobis distance of the 1-sigma ellipsoid based on $\mathbf{P}^{(1)}$. A notional depiction of the minimal and maximal ratio directions of two ellipsoids is shown in Fig. 3. Note that the eigenvalues of the pair $\left((\mathbf{P}_z^{(1)})^{-1}, (\mathbf{P}_z^{(2)})^{-1}\right)$ are the same as the eigenvalues of the pair $\left(\mathbf{P}_z^{(2)}, \mathbf{P}_z^{(1)}\right)$. Thus, the maximum/reciprocal minimum ratios of the precisions are the same as the maximum/reciprocal minimum ratios of the variances along any direction. Computing the generalized eigenvalues of the covariance matrices is more numerically stable and efficient than computing inverses and then computing the generalized eigenvalues of these.

Expected Mahalanobis Distance Between Linear and Higher-Order Models Similar to the analysis done with the expected squared Mahalanobis Distance of the true distribution, the expected squared Mahalanobis Distance can be approximated comparing the linear propagation of the mean and covariance with the second-order propagation of the mean and covariance. The second-order mean is calculated as the expectation of Eq. (43) as

$$\left(\boldsymbol{\mu}_z^{(2)}\right)^i = \mathbf{g}(\boldsymbol{\mu}_x)^i + (\delta\boldsymbol{\mu}_z^{(2)})^i \tag{54}$$

where

$$(\delta\boldsymbol{\mu}_z^{(2)})^i = \frac{1}{2}\left(\mathbf{G}^{(2)}\right)^i_{j,k}(\mathbf{P}_x)^{j,k} \tag{55}$$

and the second-order covariance is calculated using Eq. (51). Once these values have been calculated, the expected squared Mahalanobis Distance can be found using Eq. (38), where the true mean and covariance in the equation, which were found using a Monte Carlo simulation, are replaced by the second-order propagated mean and covariance. This second-order approximation of the mean can also be compared to the first-order approximation of the mean using a linear covariance Mahalanobis distance as described in Eq. (30).

Maximal Normalized Skewness and Kurtosis Analysis Computing a second- or higher-order approximation of the skewness and kurtosis of $\mathbf{z}$ can be accomplished using partial derivative tensors.[12,13] Up to second-order in the approximation of the nonlinear function $\mathbf{g}$, we have that the

central third-order moment tensor is given

$$(\mathbf{S}_z^{(2)})^{i,j,k} = \frac{1}{2}(\mathbf{G}_{l_1}^i \mathbf{G}_{l_2}^j (\mathbf{G}^{(2)})_{l_3,l_4}^k + \mathbf{G}_{l_1}^i \mathbf{G}_{l_2}^k (\mathbf{G}^{(2)})_{l_3,l_4}^j + \mathbf{G}_{l_1}^j \mathbf{G}_{l_2}^k (\mathbf{G}^{(2)})_{l_3,l_4}^i)(\mathbf{K}_x)^{l_1,l_2,l_3,l_4} \quad (56)$$
$$+ \frac{1}{8}\mathbf{G}_{l_1,l_2}^i \mathbf{G}_{l_3,l_4}^j \mathbf{G}_{l_5,l_6}^k (\mathbf{K}_x^{[6]})^{l_1,\ldots,l_6} - 3\mathrm{sym}((\mathbf{P}_z^{(2)})^{i,j}(\delta\boldsymbol{\mu}_z^{(2)})^k) + 2(\delta\boldsymbol{\mu}_z^{(2)})^i(\delta\boldsymbol{\mu}_z^{(2)})^j(\delta\boldsymbol{\mu}_z^{(2)})^k$$

because the third-order central moments of $\mathbf{x}$ are all zero. The notation $\mathbf{K}^{[2M]}$ denotes the even $2M$-th order central moment tensor (generalizing the 4th order moment tensor $\mathbf{K}$). The central fourth-order moment tensor up to second-order is given by

$$(\mathbf{K}_z^{(2)})^{i,j,k,l} = \mathbf{G}_{q_1}^i \mathbf{G}_{q_2}^j \mathbf{G}_{q_3}^k \mathbf{G}_{q_4}^l (\mathbf{K}_x)^{q_1,q_2,q_3,q_4} \quad (57)$$
$$+ \frac{1}{4}\binom{4}{2}\mathrm{sym}((\mathbf{G}^{(2)})_{q_1,q_2}^i (\mathbf{G}^{(2)})_{q_3,q_4}^j \mathbf{G}_{q_5}^k \mathbf{G}_{q_6}^l (\mathbf{K}_x^{[6]})^{q_1,\ldots,q_6})$$
$$+ \frac{1}{16}(\mathbf{G}^{(2)})_{q_1,q_2}^i (\mathbf{G}^{(2)})_{q_3,q_4}^j (\mathbf{G}^{(2)})_{q_5,q_6}^k (\mathbf{G}^{(2)})_{q_7,q_8}^l (\mathbf{K}_x^{[8]})^{q_1,\ldots,q_8} - 4\mathrm{sym}((\mathbf{S}_z^{(2)})^{i,j,k})(\delta\boldsymbol{\mu}_z^{(2)})^l)$$
$$+ 6\mathrm{sym}((\mathbf{P}_z^{(2)})^{i,j}(\delta\boldsymbol{\mu}_z^{(2)})^k(\delta\boldsymbol{\mu}_z^{(2)})^l) - 3(\delta\boldsymbol{\mu}_z^{(2)})^i(\delta\boldsymbol{\mu}_z^{(2)})^j(\delta\boldsymbol{\mu}_z^{(2)})^k(\delta\boldsymbol{\mu}_z^{(2)})^l$$

where the coefficient of the second term is "4 choose 2" which is equal to 6, the symmetrization of a tensor is given in terms of the average over all permutations of the indices (though if there are underlying symmetries, the sum need not be computed over all permutations if the normalization factor is adjusted)

$$\mathrm{sym}(\mathbf{T}^{i_1\ldots i_N}) = \frac{1}{N!}\sum_{\sigma\in S_N} \mathbf{T}^{\sigma(i_1,\ldots,i_N)} \quad (58)$$

and the even higher-order central moments of a Gaussian distribution are given in terms of the covariance as

$$(\mathbf{K}^{[2M]})^{i_1,\ldots,i_{2M}} = \frac{(2M)!}{2^M(M)!}\mathrm{sym}(\mathbf{P}^{i_1,i_2}\ldots\mathbf{P}^{i_{2M-1},i_{2M}}) \quad (59)$$

where the constant out front is the number of ways to partition a set of $2M$ objects into pairs when order does not matter. The normalized central-moment tensors can be calculated by applying the whitening transformation to each index of the central-moment tensor where the whitening transformation is given by the inverse matrix square root of the second-order approximation of the covariance. Once these tensors are calculated, a constrained optimization problem may be solved in order to find the maximal skewness and excess kurtosis of any marginal of the whitened transformation of the random variable representing the second-order approximation $\mathbf{z}^{(2)}$.

$$\max_{\|\mathbf{x}\|_2=1} T^{i_1,\ldots,i_N} x_{i_1}\ldots x_{i_N} \quad (60)$$

This maximal skewness or excess kurtosis of the whitened second-order approximation is a measure of the non-Gaussianity of the distribution. In Fig. 4, we show three distribution which serve as a reference for the meaning in the magnitude of the normalized skewness and excess normalized kurtosis.

Optimization-Based Analytical Approaches

In previous work, constrained optimization-based methods were employed to determine the need for and direction of Gaussian mixture splitting in the context of nonlinear uncertainty propagation.[14] We summarize a few of the methods used in that work that provide the most easily interpretable metrics associated with the performance of linear covariance propagation.

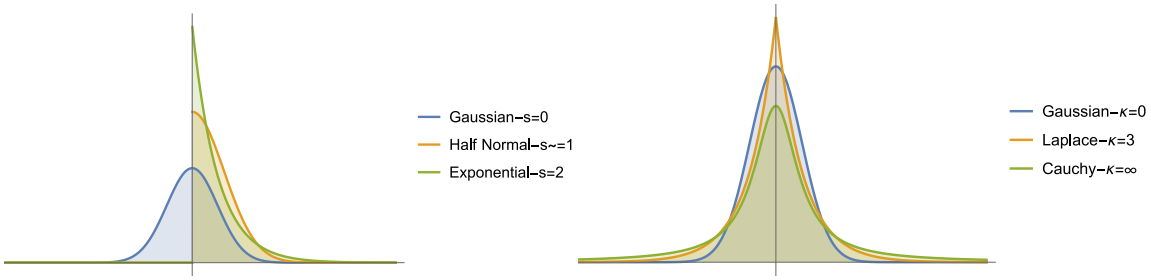

Figure 4. The normalized skewness and kurtosis of three common distributions as reference to interpret skewness values in terms of the asymmetry of the distribution and kurtosis values in terms of the heavy-tailedness of the distribution.

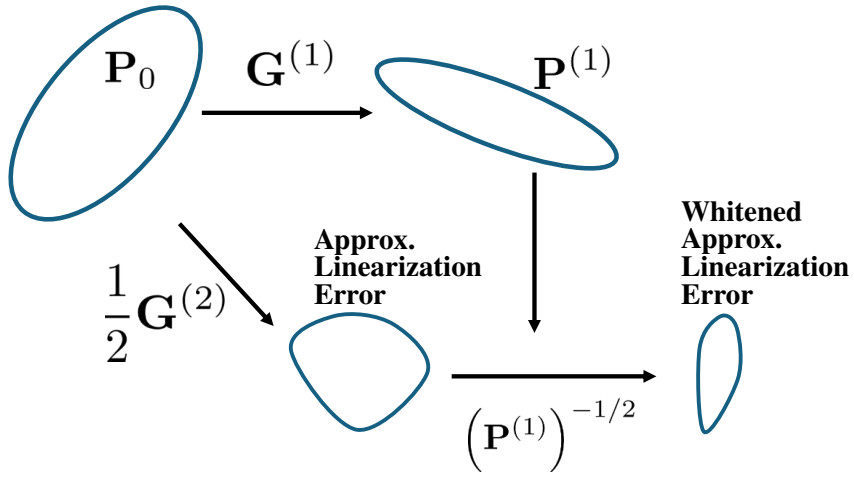

Figure 5. A schematic describing the quantity being maximized in computing WUSSOS.

Maximal Mahalanobis Distance of 1-sigma Linearization Error The whitened uncertainty-scaled second-order stretching (WUSSOS) heuristic

$$\max_{\mathbf{x}^T\mathbf{P}_x^{-1}\mathbf{x}=1} \|\mathbf{G}^{(2)}\mathbf{x}^2\|_{\mathbf{P}_z^{-1}} \tag{61}$$

combines nonlinearity, uncertainty, and output whitening to characterize the effectiveness of linear uncertainty propagation.[14] WUSSOS gives the maximum Mahalanobis distance (induced by the linearly propagated covariance) of the second-order approximation of the linearization error for any initial point on a surface of equal likelihood in the input space. In Fig. 5, we sketch the surface over which the maximum distance from the origin is optimized to arrive at WUSSOS. The WUSSOS optimization problem can be rephrased in a manner computable as a maximal Z-eigenvector with shifted symmetric higher-order power iteration as

$$\max_{\mathbf{y}^T\mathbf{y}=1} (\mathbf{G}^{(2)}(\mathbf{P}_x^{1/2}\mathbf{y})^2)^T\mathbf{P}_z^{-1}(\mathbf{G}^{(2)}(\mathbf{P}_x^{1/2}\mathbf{y})^2) \tag{62}$$

The relevant fourth-order tensor whose square-rooted maximum Z-eigenvalue gives WUSSOS can be written

$$\mathbf{W}_{i_1,i_2,i_3,i_4} = \left(\mathbf{P}_z^{-1}\right)_{j_1,j_2} \left(\mathbf{G}^{(2)}\right)^{j_1}_{k_1,k_2} \left(\mathbf{G}^{(2)}\right)^{j_2}_{k_3,k_4} \left(\mathbf{P}_x^{1/2}\right)^{k_1}_{i_1} \ldots \left(\mathbf{P}_x^{1/2}\right)^{k_4}_{i_4} \tag{63}$$

Relaxation of Maximal Mahalanobis Distance of 1-sigma Linearization Error Tuggle and Zanetti[15,16] first introduced a Gaussian mixture splitting method that employed nonlinear function properties and uncertainty scaling to arrive at the following optimization to select a splitting direction

$$\hat{\mathbf{x}}^* \sim \underset{\mathbf{x}^T\mathbf{P}_x^{-1}\mathbf{x}=1}{\arg\max} \|\mathbf{G}^{(2)}\mathbf{x}\|_F \tag{64}$$

where the subscript F denotes the Frobenius norm and $\mathbf{G}^{(2)}\mathbf{x}$ denotes the matrix with components

$$\left(\mathbf{G}^{(2)}\mathbf{x}\right)_j^i = (\mathbf{G}^{(2)})_{j,k}^i x^k \tag{65}$$

which describes a linear approximation of how the Jacobian of $\mathbf{g}$ changes as it is evaluated at a point $\mathbf{x}$ away from the current point it is being evaluated at. This splitting direction and the associated metric can be calculated using a generalized eigenvalue problem with the two matrices $(\bar{\mathbf{G}}^{(2)}, \mathbf{P}_x^{-1})$ or with the singular value decomposition of $\bar{\mathbf{G}}^{(2)}\mathbf{P}_x^{1/2}$ where the mn by n matricization of the tensor $\mathbf{G}^{(2)}$ is

$$(\bar{\mathbf{G}}^{(2)})_k^{ni+j} = (\mathbf{G}^{(2)})_{j,k}^i \tag{66}$$

In our previous work,[14] we established a scale invariance for this metric and selection criterion by transforming the linear transformation given by the matrix $\mathbf{G}^{(2)}\mathbf{x}$ so that it maps from a whitened input space to a whitened output space. For completeness, we summarize this process here. The input whitening transformation is given by

$$\mathbf{x}' = \mathbf{P}_x^{-1/2}\mathbf{x} \tag{67}$$

and the whitening transformation for the output space is given by

$$\mathbf{z}' = \mathbf{P}_z^{-1/2}\mathbf{z} \tag{68}$$

A generic linear transformation $\mathbf{A}$ from the original input space to the original output space can be transformed to a linear transformation from the whitened input space to the whitened output space according to

$$\mathbf{A}' = \mathbf{P}_z^{-1/2}\mathbf{A}\mathbf{P}_x^{1/2} \tag{69}$$

If we apply this transformation to the linear transformation given by the matrix $\mathbf{G}^{(2)}\mathbf{x}$, we obtain a linear transformation given by the matrix

$$\mathbf{P}_z^{-1/2}(\mathbf{G}^{(2)}\mathbf{x})\mathbf{P}_x^{1/2} \tag{70}$$

whose squared Frobenius norm will characterize the change in the linear approximation of the nonlinear function $\mathbf{g}$ (normalized to map between whitened input and output spaces) at a step $\mathbf{x}$ away from the current point of linearization. In order to assess whether the change in this whitened linearization is significant we should compare against the squared Frobenius norm of the original linearization between whitened spaces. If the original linear transformation is $\mathbf{A} = \mathbf{G}$ and both $\mathbf{P}_x, \mathbf{P}_z$ are nonsingular, then the corresponding linear transformation between whitened spaces $\mathbf{G}'$ is a linear transformation from the unit sphere in n dimensions (the dimension of the domain of $\mathbf{g}$) to the unit sphere in m dimensions (the dimension of the codomain of $\mathbf{g}$). Thus, $\mathbf{G}'$ depicted in Fig. 6 is an orthogonal matrix. Note that since the eigenvalues of an orthogonal matrix only take values

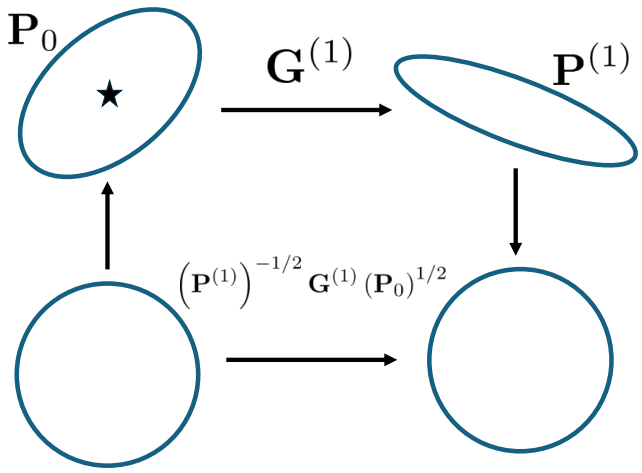

Figure 6. A schematic describing the reference transformation for comparison with the transformations being optimized in WUSSOLC.

$-1, 0$, and 1, the rank of the matrix gives the number of nonzero eiegnvalues, and the Frobenius norm of $\mathbf{G}'$ is given by the sum of squares of the eigenvalues

$$\|\mathbf{G}'\|_F^2 = \min(n, m) \tag{71}$$

As such, the squared Frobenius norm of the whitened change in the linearization given by $\mathbf{G}^{(2)}\mathbf{x}$ can naturally be compared to $\min(n, m)$. If this value is small relative to $\min(n, m)$, then the linearization at a point $\mathbf{x}$ away from the mean would propagate the initial covariance ellipsoid to a final covariance ellipsoid that is very similar. The resulting whitened uncertainty-scaled second-order linearization change (WUSSOLC) metric

$$\max_{\mathbf{x}^T\mathbf{P}_x^{-1}\mathbf{x}=1} \|\mathbf{P}_z^{-1/2}(\mathbf{G}^{(2)}\mathbf{x})\mathbf{P}_x^{1/2}\|_F^2 = \max_{\mathbf{y}^T\mathbf{y}=1} \|\mathbf{P}_z^{-1/2}(\mathbf{G}^{(2)}(\mathbf{P}_x^{1/2}\mathbf{y})))\mathbf{P}_x^{1/2}\|_F^2 \tag{72}$$

can be compared to $\min(n, m)$ as a measure of the maximal normalized change in the linearization over the 1-σ covariance ellipsoid. Note that this metric corresponds to the induced (Frobenius, 2)-norm of the same tensor of which WUSSOS is the induced 2-norm. See previous work on tensor norms for details on the relationship between the induced 2-norm of a tensor and the induced (Frobenius, 2)-norm of a tensor and the inequality between the two.[2] In order to compute WUSSOLC most simply, compute the tensor

$$(\mathbf{G}_w^{(2)})_{j,k}^i = (\mathbf{P}_z^{-1/2})_l^i(\mathbf{G}^{(2)})_{p,q}^l(\mathbf{P}_x^{1/2})_j^p(\mathbf{P}_x^{1/2})_k^q \tag{73}$$

and its mn by n matricization

$$(\bar{\mathbf{G}}_w^{(2)})_k^{ni+j} = (\mathbf{G}_w^{(2)})_{j,k}^i \tag{74}$$

from which we can write the WUSSOLC metric as

$$\max_{\mathbf{y}^T\mathbf{y}=1} \|\bar{\mathbf{G}}_w^{(2)}\mathbf{y}\|_2^2 \tag{75}$$

which is simply the maximal singular value of the matrix $\bar{\mathbf{G}}_w^{(2)}$.

APPLICATIONS

In this section, we test the aforementioned methods in an analysis involving the proposed NASA Gateway Near Rectilinear Halo Orbit (NRHO). An initial Gaussian state uncertainty is assumed, and the distribution is propagated using linear covariance techniques over the course of an orbit. Over that period, we compare the proposed analytic metrics with Monte Carlo based metrics to assess their accuracy and ultimate ability to identify divergent LinCov results.

THREE-BODY MOTION

For simplicity, we adopt the circular three-body problem assumptions. The corresponding equations of motion expressed in the synodic frame are

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x} = \mathbf{F}(\mathbf{x}) \tag{76}$$

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} \dot{x} & \dot{y} & \dot{z} & 2\dot{y} + \frac{\partial \overline{U}}{\partial x} & -2\dot{x} + \frac{\partial \overline{U}}{\partial y} & \frac{\partial \overline{U}}{\partial z} \end{bmatrix}^T \tag{77}$$

where $\overline{U}(x,y,z) = \dfrac{1-\mu^*}{||\mathbf{r}_1||} + \dfrac{\mu^*}{||\mathbf{r}_2||} + \dfrac{x^2+y^2}{2}$ is the effective potential and the mass ratio is defined as $\mu^* = \dfrac{m_2}{m_1+m_2}$ for the two primary bodies with masses m_1, m_2 respectively. We adopt the convention that the primary body with greater mass is assigned the index 1 so that $m_1 \geq m_2$. Both masses lie on the x-axis at $[-\mu^*, 0, 0]$ and $[1-\mu^*, 0, 0]$ with respect to their common barycenter at the origin. The positions of the satellite of interest with respect to the primary and secondary bodies are denoted by $\mathbf{r}_1$ and $\mathbf{r}_2$ respectively.[17]

The reference orbit used in the following sections is a 9:2 resonant Southern L_2 halo orbit like that proposed for the NASA Gateway[18] and shown in Figure 8. These initial conditions were obtained as a canonical unit conversion of the initial conditions used in the QIST model of Gateway.[19] The initial conditions (coinciding with apolune) and mass parameter used for the orbit are

$$\mu = 1.0/(81.30059 + 1.0), \quad x_0 = 1.022022,$$
$$z_0 = -0.182097, \quad \dot{y}_0 = -0.103256$$

in nondimensional units with other initial coordinates equal to zero.

The period of the orbit is approximately $1.511111\,[\mathrm{TU}]$ where

$$2\pi[\mathrm{TU}] = 2.361 \cdot 10^6\,[\mathrm{sec}] \tag{78}$$

In this example, the nonlinear function $\mathbf{g}$ in question is the flow of the circular restricted three-body dynamics φ_t for some fixed value of the time-of-flight t. The flow map associated with the dynamical system in (77) is defined such that

$$\left(\frac{\mathrm{d}}{\mathrm{d}\tau}\varphi_\tau(\mathbf{x}_0)\right)\Bigg|_{\tau=t} = \mathbf{F}(\varphi_t(\mathbf{x}_0)), \quad \varphi_0(\mathbf{x}_0) = \mathbf{x}_0 \tag{79}$$

The Jacobian and the second-order partial derivative tensor of the flow map around some reference trajectory are the state transition matrix and the second-order state transition tensor, respectively. These quantities are obtained by integrating the variational equations,[3] or by employing techniques

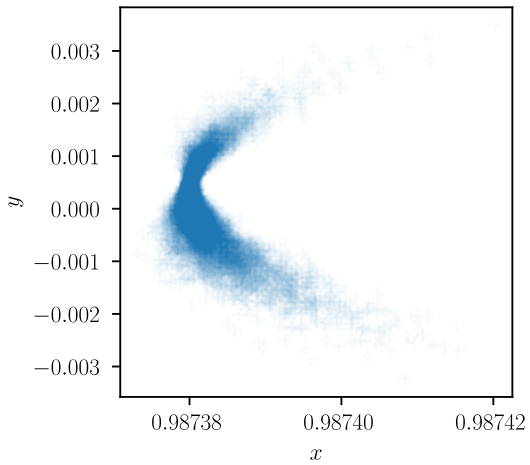

Figure 7. Scatter plot of Monte Carlo samples at reference orbit perilune in the x-y plane.

from differential algebra.[4] While repeated computation of state transition tensors can be costly, techniques exist for precomputing them offline along a known reference trajectory and then efficiently interpolating the state transition tensors online.[19–21] Other efficient approaches are available that leverage the potential low-rank qualities of the state transition tensors.[22–24]

In this example, we employ an initial Gaussian distribution with mean equal to the Gateway-like initial conditions and covariance given by

$$\mathbf{P}_x = 10^{-8}\text{diag}([1, 0, 1, 0, 0, 0]) + 10^{-10}\mathbf{I}_6 \tag{80}$$

in nondimensional canonical units. The 1-sigma distances are on the order of $40\,[\text{km}]$ along the x and z directions, $4\,[\text{km}]$ along the y direction, and $0.01\,[\text{m/s}]$ in each velocity direction. This covariance is designed to be non-isotropic and to highlight the significant non-Gaussianity of the propagated distribution. This non-Gaussianity is evident in Fig. 7, which shows the x-y marginal distribution at perilune, one half period after the initial distribution epoch. All of the Monte Carlo-based metrics in this paper are computed using 10,000 samples propagated with the Circular Restricted Three-Body dynamics.

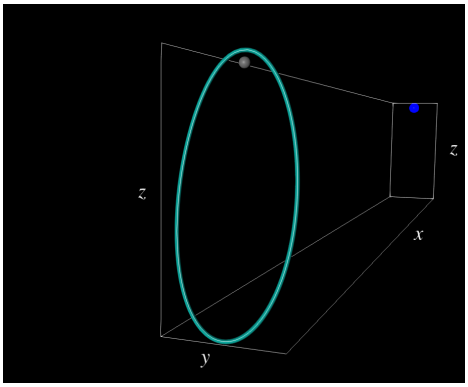

Figure 8. Full NRHO considered in three-body uncertainty propagation application.

A number of orbits in the Earth-Moon Circular Restricted Three-Body problem were consid-

ered in a similar study using the Tensor Eigenpair Measure of Nonlinearity (TEMoN) as a flag for when uncertainty propagation leads to non-Gaussianity of uncertainty.[25] Our work has a similar purpose but each method directly incorporates initial uncertainty and function nonlinearity to provide wholistic metrics rather than correlating distribution independent measures of nonlinearity with measures of non-Gaussianity to try to use one as a proxy for the other.

Results

In all of the following results for each metric, we see that each measure of the quality of the linear covariance propagation degrades extremely for propagation from apolune to near perilune. This matches results on the related nonlinearity index.[1,2] In Fig. 9, the expected squared Mahalanobis

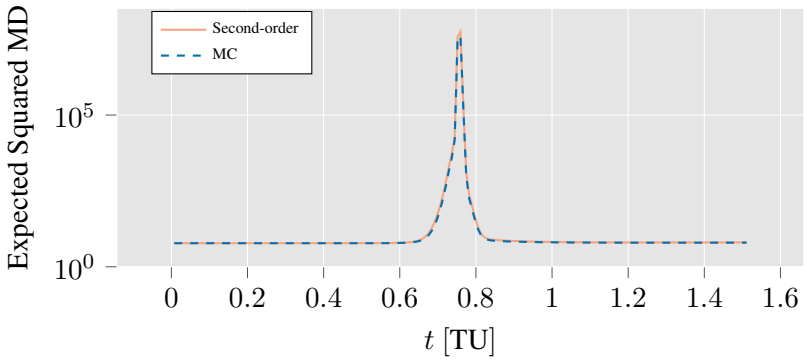

Figure 9. The expected squared Mahalanobis distance relative to the nonlinearly propagated mean and linear covariance.

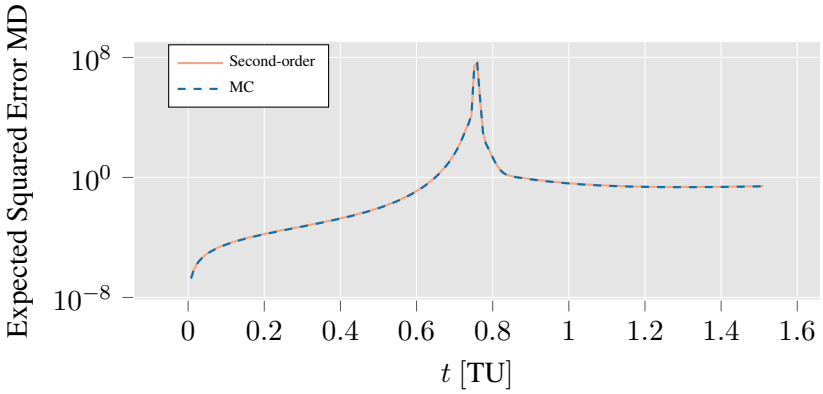

Figure 10. The expected squared Mahalanobis distance of the linearization error with respect to the linear covariance.

distance begins at a value of six. Six is the mean of a chi-square distribution with six degrees of freedom and the squared Mahalanobis distance of the initial distribution follows a chi-square distribution with six degrees of freedom. Except between 0.6 and 0.9 nondimensional time units, the expected squared Mahalanobis distance stays below a value of seven. Over the same time interval, the expected square Mahalanobis distance of the linearization error in Fig. 10 stays below unity, though it begins at zero instead of six. The squared Mahalanobis distance of the difference

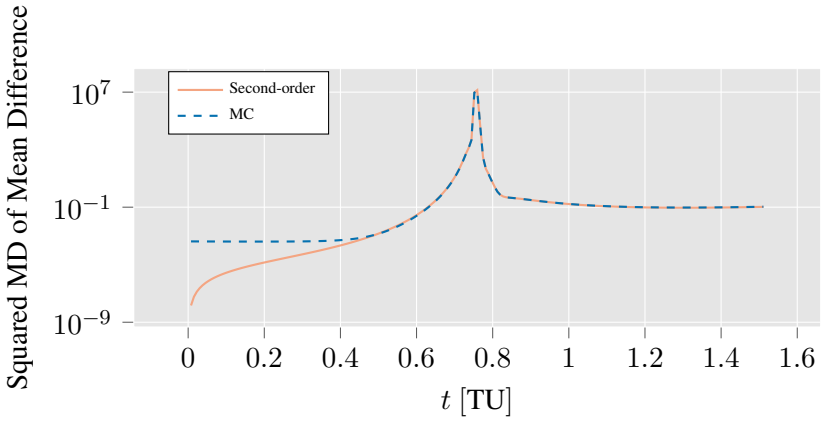

Figure 11. The squared Mahalanobis distance of the nonlinearly propagated mean and another higher-fidelity approximation of the mean.

in the mean propagated versus the mean of the propagated distribution is given in Fig. 11. Here, the Monte Carlo based approach does not start at zero like the analytical approach. This is because the mean of the initial samples from the original distribution will not be exactly the mean of the distribution. By the end of the period, the value settles to around 0.1 which is lower than the value around 0.25 that is approached by the expected squared Mahalanobis distance of the linearization error. This is expected since Jenson's inequality guarantees an inequality between the two metrics.

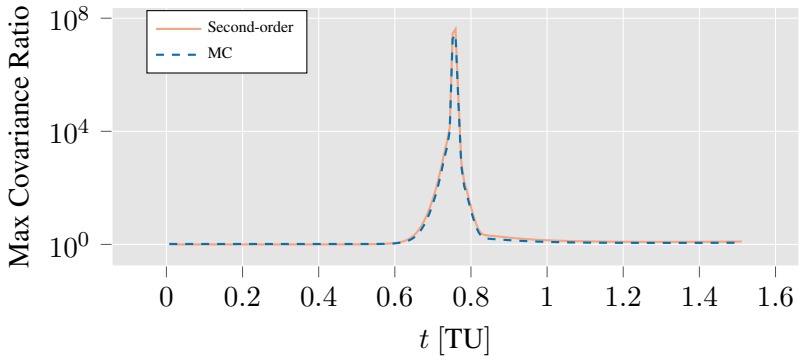

Figure 12. The maximum ratio between the linear variance and a higher-fidelity approximation of the variance in any direction.

The maximum covariance ratio depicted in Fig. 12 begins at a value of one, settles at the end of the period to around 1.25 and 1.15 for the second-order and Monte Carlo estimates respectively. This metric shows more disagreement between the two methods than the other expectation-based metrics. It is also noteworthy that the shape of the graph is very similar to that of the expected squared Mahalanobis distance in Fig. 9. This is because the trace expression in Eq. 38 is equal to the sum of the generalized eigenvalues from Eq. 32 of which the maximum (or the reciprocal of the minimum) is employed as the maximum covariance ratio.

In Fig. 13, both WUSSOS and WUSSOLC begin at zero and settle to a value around 0.5 by the end of a single orbit. As noted in,[2,14] the WUSSOLC metric will always be greater than the WUSSOS metric, however in this case, the two stay consistently within a difference of a few percent

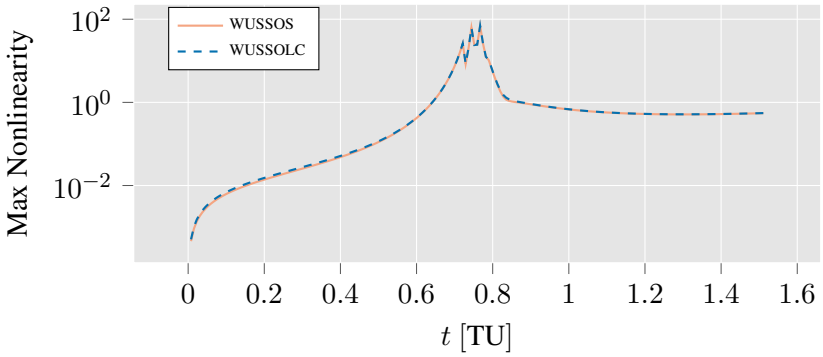

Figure 13. The maximum uncertainty scaled nonlinearity.

between one another. Finally, in Fig. 14 and Fig. 15, the Monte Carlo results and the second-order

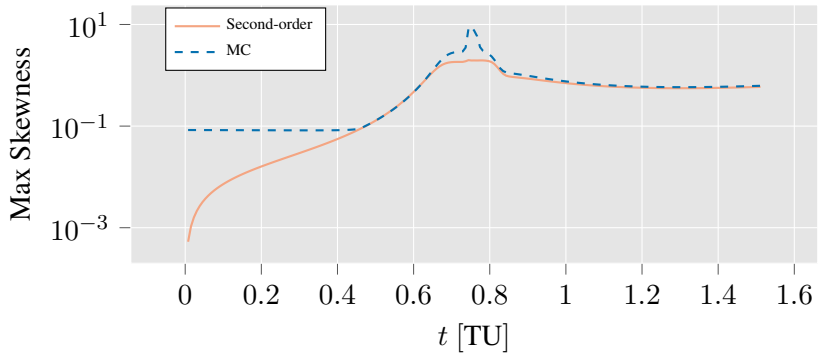

Figure 14. The maximal normalized skewness.

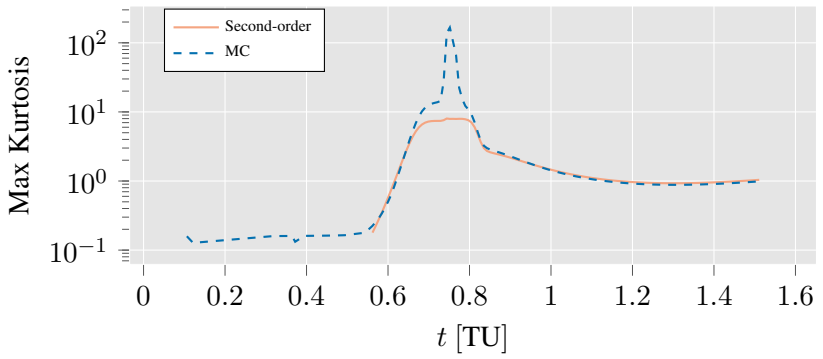

Figure 15. The maximum excess normalized kurtosis

analytical results are significantly different, though they both flag non-Gaussianity around the same time. In both instances of the Monte Carlo, we see that the sample skewness and kurtosis of the initial distribution is nonzero. This effect dominates the skewness and kurtosis metrics for the Monte Carlo until around a quarter of the orbit. Computation of the skewness and kurtosis tensors is generally more costly than the WUSSOS and WUSSOLC measures. Thus, the WUSSOS and

WUSSOLC measures offer an accurate albeit indirect indication of non-Gaussianity at a fraction of the computational cost of the direct skewness and kurtosis measures. The additional computational penalty of the second-order kurtosis stems from additional required shifted symmetric higher-order power iterations, and, in some cases, a maximal eigenvector was not found within one thousand iterations and ten random initializations for time-of-flight below one quarter of the period. We speculate that this is as a result of the small gap between eigenvalues of the nearly zero kurtosis tensor during that time. Convergence of power iteration type algorithms tends to depend on the size of the difference between the largest and next largest eigenvalue. All this goes to show that looking directly at the higher-order moments of the distribution may be the least consistent, interpretable, and efficient method for quantifying non-Gaussianity of those metrics proposed in this paper.

CONCLUSION

We have presented a number of interpretable metrics based on second-order partial derivatives for assessing the error in using linear covariance propagation through nonlinear functions. All metrics presented similar trends when employed in the context of Cislunar astrodynamics uncertainty propagation problem. Those methods which have a Monte Carlo/sampling based equivalent (besides the maximum skewness and kurtosis characterizations), all matched the sampling based method well in the example considered. As the scale of initial uncertainty increases, second-order models of the nonlinear function may not accurately capture the true sampling based equivalents, but should still function as an appropriate warning when linear covariance analysis is failing to accurately describe the true distribution. These second-order validation methods are light-weight when compared with Monte Carlo-based approaches especially if the second-order partial derivatives can be calculated efficiently or precomputed and accessed efficiently. Finally, we have seen that of the methods discussed, the higher-order moment-based approaches have tended to be orders of magnitude slower to compute as compared with the other methods discussed. Additionally, the second-order approximation is not sufficient to match the Monte Carlo approach well, and third-order partial derivatives of the nonlinear function may be necessary for skewness analysis while fourth-order partial derivatives may be necessary for kurtosis analysis. While we initially considered higher-order moment analysis as one of the most natural directions to determine non-Gaussianity and thus ineffectiveness of linear covariance analysis, without significant advances, we believe these to be less effective methods than those that more directly consider uncertainty-weighted measures of nonlinearity.

In future work we plan to more precisely characterize the timing results for these algorithms, as well as present sigma-point based variations on these ideas. We also intend to use some of these metrics to understand the transition to non-Gaussianity and ineffectiveness of linear covariance techniques across an NRHO more broadly rather than always beginning the propagation at apolune.

REFERENCES

[1] E. L. Jenson and D. J. Scheeres, "Semianalytical measures of nonlinearity based on tensor eigenpairs," *Journal of Guidance, Control, and Dynamics*, Vol. 46, No. 4, 2023, pp. 638–653.

[2] J. Kulik, C. Orton-Urbina, M. Ruth, and D. Savransky, "Applications of Induced Tensor Norms to Guidance Navigation and Control," *arXiv preprint arXiv:2408.15362*, 2024.

[3] R. S. Park and D. J. Scheeres, "Nonlinear mapping of Gaussian statistics: theory and applications to spacecraft trajectory design," *Journal of guidance, Control, and Dynamics*, Vol. 29, No. 6, 2006, pp. 1367–1375.

[4] M. Rasotto, A. Morselli, A. Wittig, M. Massari, P. Di Lizia, R. Armellin, C. Valles, and G. Ortega, "Differential algebra space toolbox for nonlinear uncertainty propagation in space dynamics," 2016.

[5] P. McCullagh, *Tensor methods in statistics: Monographs on statistics and applied probability*. Chapman and Hall/CRC, 2018.

[6] T. G. Kolda and J. R. Mayo, "Shifted power method for computing tensor eigenpairs," *SIAM Journal on Matrix Analysis and Applications*, Vol. 32, No. 4, 2011, pp. 1095–1124.

[7] L. Qi, "Eigenvalues of a real supersymmetric tensor," *Journal of symbolic computation*, Vol. 40, No. 6, 2005, pp. 1302–1324.

[8] L. De Lathauwer, B. De Moor, and J. Vandewalle, "On the best rank-1 and rank-(r 1, r 2,..., rn) approximation of higher-order tensors," *SIAM journal on Matrix Analysis and Applications*, Vol. 21, No. 4, 2000, pp. 1324–1342.

[9] L.-H. Lim, "Singular values and eigenvalues of tensors: a variational approach," *1st IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing, 2005.*, IEEE, 2005, pp. 129–132.

[10] T. G. Kolda and B. W. Bader, "MATLAB tensor toolbox," tech. rep., Sandia National Laboratories (SNL), Albuquerque, NM, and Livermore, CA . . . , 2006.

[11] E. L. Jenson and D. J. Scheeres, "Semianalytical Measures of Nonlinearity Based on Tensor Eigenpairs," *Journal of Guidance, Control, and Dynamics*, Vol. 46, Apr. 2023, p. 638–653, 10.2514/1.g006760.

[12] M. Majji, J. Junkins, and J. Turner, "Jth moment extended Kalman filtering for estimation of nonlinear dynamic systems," *AIAA Guidance, Navigation and Control Conference and Exhibit*, 2008, p. 7386.

[13] M. Majji, J. L. Junkins, and J. D. Turner, "A perturbation method for estimation of dynamic systems," *Nonlinear Dynamics*, Vol. 60, 2010, pp. 303–325.

[14] J. Kulik and K. A. LeGrand, "Nonlinearity and Uncertainty Informed Moment-Matching Gaussian Mixture Splitting," *arXiv preprint arXiv:2412.00343*, 2024.

[15] K. Tuggle and R. Zanetti, "Automated splitting Gaussian mixture nonlinear measurement update," *Journal of Guidance, Control, and Dynamics*, Vol. 41, No. 3, 2018, pp. 725–734.

[16] K. E. Tuggle, *Model selection for Gaussian mixture model filtering and sensor scheduling*. PhD thesis, 2020.

[17] W. S. Koon, M. W. Lo, J. E. Marsden, and S. D. Ross, "Dynamical systems, the three-body problem and space mission design," *Equadiff 99: (In 2 Volumes)*, pp. 1167–1181, World Scientific, 9 2000, 10.1142/9789812792617_0222.

[18] "National Aeronautics and Space Administration (NASA) White Paper: Gateway Destination Orbit Model: A Continuous 15 Year NRHO Reference Trajectory," 2019.

[19] D. Cunningham and R. P. Russell, "An Interpolated Second-Order Relative Motion Model for Gateway," *The Journal of the Astronautical Sciences*, Vol. 70, No. 4, 2023, p. 26.

[20] J. Kulik, W. Clark, and D. Savransky, "State Transition Tensors for Continuous-Thrust Control of Three-Body Relative Motion," *Journal of Guidance, Control, and Dynamics*, 5 2023, pp. 1–10, 10.2514/1.g007311.

[21] D. A. Cunningham and R. P. Russell, "Relative motion solutions around an arbitrary SPICE Kernel trajectory," *Astrodynamics Specialist Conference, 2024*, 2024, pp. AAS 24–457.

[22] S. Boone and J. McMahon, "Directional State Transition Tensors for Capturing Dominant Nonlinear Effects in Orbital Dynamics," *Journal of Guidance, Control, and Dynamics*, Vol. 46, No. 3, 2023, pp. 431–442.

[23] S. Boone and J. McMahon, "An Efficient Approximation of the Second-Order Extended Kalman Filter for a Class of Nonlinear Systems," *2024 European Control Conference (ECC)*, IEEE, 2024, pp. 3533–3538.

[24] X. Zhou, R. Armellin, D. Qiao, and X. Li, "Time-Varying Directional State Transition Tensor for Orbit Uncertainty Propagation," *arXiv preprint arXiv:2412.07060*, 2024.

[25] J. Gutierrez, K. Hill, E. L. Jenson, D. J. Scheeres, J. C. Bruer, and R. D. Coder, "Classifying State Uncertainty for Earth-Moon Trajectories," *The Journal of the Astronautical Sciences*, Vol. 71, No. 3, 2024, p. 29.