

QUANTIFYING VOLUME CHANGES FROM COVARIANCE PROPAGATION

Jackson Kulik*, and Keith A. LeGrand†

In a Hamiltonian dynamical system, the volume associated with a highest density region containing a given probability mass is exactly conserved. Similarly, the volume of covariance ellipsoids are conserved under linear covariance propagation in these systems. In this work, we quantitatively explore the non-conservation of volume associated with ellipsoids arising from higher-order and unscented propagation of the covariance. Along the way, we compare the volume preservation of the unscented transform based on different whitening transformations and parameters. Finally, we note a novel family of conserved quantities related to the state transition tensors.

INTRODUCTION

A nonlinear transformation $\mathbf{g}(\mathbf{x})$ is considered volume preserving if everywhere

$$\det \left(\frac{\partial \mathbf{g}}{\partial \mathbf{x}} \right) = 1 \quad (1)$$

Note that a volume preserving linear transformation \mathbf{A} preserves the determinant of the covariance of a Gaussian distribution propagated through that linear transformation as a result of the multiplicative property of the determinant. If the random variable $\mathbf{x}_1 \in \mathbb{R}^n$ is distributed as a Gaussian with prescribed mean and covariance $\boldsymbol{\mu}_1, \mathbf{P}_1$

$$\mathbf{x}_1 \sim \mathcal{N}(\boldsymbol{\mu}_1, \mathbf{P}_1) \quad (2)$$

and

$$\mathbf{x}_2 = \mathbf{A}\mathbf{x}_1 \quad (3)$$

then

$$\mathbf{x}_2 \sim \mathcal{N}(\boldsymbol{\mu}_2, \mathbf{P}_2) \quad (4)$$

where

$$\mathbf{P}_2 = \mathbf{A}\mathbf{P}_1\mathbf{A}^T \quad (5)$$

So, when \mathbf{A} is volume preserving, the determinant of the covariance before and after the linear transformation are equal:

$$\det(\mathbf{P}_2) = \det(\mathbf{A}) \det(\mathbf{P}_1) \det(\mathbf{A}^T) = \det(\mathbf{P}_1) \quad (6)$$

*Assistant Professor, Mechanical and Aerospace Engineering, Utah State University, Logan UT

†Assistant Professor, School of Aeronautics and Astronautics, Purdue University, West Lafayette, IN

As such, we consider a transformation between two Gaussian distributions to be volume preserving if the determinant of each covariance is equal to the other

$$\det(\mathbf{P}_1) = \det(\mathbf{P}_2) \quad (7)$$

Another equivalent sense in which two Gaussian distributions can be considered to have the same volume is in the sense of the volume of their highest density regions, the ellipsoidal region which contains all points with probability density above a given threshold or equivalently Mahalanobis distance from the mean below a given threshold. The k - σ ellipsoid is the region defined by

$$S_{k,\mathbf{P}} = \{\mathbf{x} \text{ s.t. } \mathbf{x}^T \mathbf{P}^{-1} \mathbf{x} \leq k^2\} \quad (8)$$

This describes an ellipsoid with semi-axes given by the vectors

$$\mathbf{a}_i = k\sqrt{\lambda_i} \mathbf{v}_i \quad (9)$$

where $(\lambda_i, \mathbf{v}_i)$ is an eigenpair of the covariance \mathbf{P} . The volume of a hyperellipse is given by the volume of a unit hypersphere multiplied by the determinant of a linear transformation that brings that sphere to the chosen hyperellipse. Such a linear transformation has singular values given by the semi-axes of the ellipsoid. This implies that the determinant of the linear transformation is given by the product of the lengths of the semi-axes of the ellipsoid and the overall volume of the k - σ covariance hyperellipse in n -dimensions is

$$\text{Vol}(S_{k,\mathbf{P}}) = \left(k^n \sqrt{\det(\mathbf{P})}\right) \frac{2\pi^{n/2}}{n\Gamma(n/2)} \quad (10)$$

where Γ is the gamma function which generalizes the factorial. Thus, the square root of the determinant of the covariance matrix is proportional to the volume enclosed by a covariance ellipsoid and preservation of the covariance determinant is equivalent to preservation of the volume of k - σ covariance hyperellipsoids.

The flow of a Hamiltonian dynamical system is volume preserving according to Liouville's theorem¹ so that the Jacobian of the flow, also known as the state transition matrix, is always a volume preserving linear transformation for any reference trajectory and time-of-flight. This implies that by using linear covariance analysis in a Hamiltonian dynamical system, the volume of any covariance ellipsoid containing a fixed probability mass is conserved as the linear covariance propagation is carried out.

Similarly, the volume of a highest density region² is always conserved under a Hamiltonian flow. The highest density region is defined as the minimum volume region enclosing a given probability mass. This is one generalization of the covariance ellipsoid to non-Gaussian distributions. Because a Hamiltonian flow is a diffeomorphism with determinant of the Jacobian equal to unity everywhere, the probability density function may be mapped using a change of variables approach involving the inverse of the flow.¹

On the other hand, if the first two moments of the distribution are mapped forward under the dynamics using higher-order techniques (beyond linear covariance analysis) and used to fit a Gaussian distribution, the ellipsoids containing a given probability mass associated with that higher-order approximation of the covariance will generally not have the same volume as the initial covariance ellipsoid containing the same probability mass. In this work, we will quantify this effect under Taylor series-based propagation of the moments as well as unscented transform-based propagation of

the moments. Boodram et al. numerically investigated the non-preservation of Gaussian mixtures propagated with third-order expansions of the flow and presented arguments based on Gromov's widths for invariants of the phase-space densities.³ We present additional analysis from a perturbations perspective on the volume of higher-order propagated covariances and develop an additional invariant for the partial derivatives of the flow-map of a volume preserving dynamical system.

The main tool we will use in this analysis is the following expansion of the determinant known as the Jacobi formula to quantify how volume preserving a transformation is:

$$\det(\mathbf{A} + \epsilon\mathbf{B}) = \det(\mathbf{A})(1 + \epsilon \operatorname{tr}(\mathbf{A}^{-1}\mathbf{B})) + \mathcal{O}(\epsilon^2) \quad (11)$$

INVARIANTS IN A VOLUME PRESERVING FLOW

First, we will examine the volume preserving qualities of the second-order Taylor series of the flow of a dynamical system

$$\varphi^i(\mathbf{x}_0 + \delta\mathbf{x}_0; t) = \varphi^i(\mathbf{x}_0; t) + \Phi(\mathbf{x}_0; t)_j^i \delta x_0^j + \frac{1}{2} \Psi(\mathbf{x}_0; t)_{jk}^i \delta x_0^j \delta x_0^k \quad (12)$$

where Φ, Ψ are the first two partial derivative tensors of the flow-map with respect to the initial state variables. These are called the state transition matrix and second-order state transition tensor respectively.

By the Jacobi formula, we have that

$$\det(\Phi + \Psi\delta\mathbf{x}) = \det(\Phi) (1 + \operatorname{tr}(\Phi^{-1}\Psi\delta\mathbf{x})) + \mathcal{O}(\delta\mathbf{x}^2) \quad (13)$$

On the other hand, we know that the determinant is constant for the Jacobian of the true nonlinear flow map φ_t , and the Taylor series for the determinant about any point must have all zero coefficients past the first constant term.

$$1 = \det\left(\frac{\partial\varphi_t}{\partial\mathbf{x}}\Big|_{\mathbf{x}_0+\delta\mathbf{x}_0}\right) = \det(\Phi) (1 + \operatorname{tr}(\Phi^{-1}\Psi\delta\mathbf{x})) + \mathcal{O}(\delta\mathbf{x}^2) \quad (14)$$

$$\implies \operatorname{tr}(\Phi^{-1}\Psi\delta\mathbf{x}) = 0 \quad (15)$$

for all $\delta\mathbf{x}$. Note that this trace expression above is equivalent to a linear function of $\delta\mathbf{x}$

$$\operatorname{tr}(\Phi^{-1}\Psi\delta\mathbf{x}) = (\Phi^{-1})_j^i \Psi_{ik}^j \delta x^k \quad (16)$$

$$\implies \operatorname{tr}(\Phi^{-1}\Psi\delta\mathbf{x}) = \mathbf{c}^T \delta\mathbf{x} \quad (17)$$

where

$$c_k = (\Phi^{-1})_j^i \Psi_{ik}^j = 0 \quad (18)$$

This trace equality can be framed as an approximate volume preservation condition on the first-order Taylor series approximation of the STM. In particular, it requires that for perturbations to the STM, $\Psi\delta\mathbf{x}$, given a perturbed reference trajectory, $\delta\mathbf{x}$, the generalized eigenvalues of that perturbation with respect to the original STM all sum to zero.

Knowing that \mathbf{c} is the zero vector from qualities of the state transition matrix and state transition tensor, we can derive qualities about the partial derivatives of the dynamics by examining the time derivative of this invariant vector \mathbf{c}

$$0 = \dot{c}_k = (\dot{\Phi}^{-1})_j^i \Psi_{ik}^j + (\Phi^{-1})_j^i \dot{\Psi}_{ik}^j \quad (19)$$

We review the two time derivatives that appear in this expression

$$\frac{d}{dt} \left((\Phi(t, t_0))^{-1} \right) = -\Phi^{-1} \dot{\Phi} \Phi^{-1} = -\Phi^{-1} \frac{\partial F}{\partial x} \Phi \Phi^{-1} = -\Phi^{-1} \frac{\partial F}{\partial x} \quad (20)$$

$$\dot{\Psi} = \frac{\partial^2 F}{\partial x^2} \Phi^2 + \frac{\partial F}{\partial x} \Psi \quad (21)$$

which implies that

$$0 = \dot{c}_k = -(\Phi^{-1})_l^i \left(\frac{\partial F}{\partial x} \right)_j^l \Psi_{ik}^j + (\Phi^{-1})_j^i \left(\left(\frac{\partial^2 F}{\partial x^2} \right)_{pq}^j \Phi_i^p \Phi_k^q + \left(\frac{\partial F}{\partial x} \right)_l^j \Psi_{ik}^l \right) \quad (22)$$

$$= (\Phi^{-1})_j^i \left(\frac{\partial^2 F}{\partial x^2} \right)_{pq}^j \Phi_i^p \Phi_k^q \quad (23)$$

$$= \left(\frac{\partial^2 F}{\partial x^2} \right)_{pq}^j \delta_j^p \Phi_k^q \implies \quad (24)$$

$$0 = \left(\frac{\partial^2 F}{\partial x^2} \right)_{pq}^j \delta_j^p \quad (25)$$

where δ is the Kronecker delta, and the final implication comes from the fact that the state transition matrix is always full rank, and thus null vectors of the state transition matrix must be the zero vector. For the circular restricted three-body problem as well as the two-body problem in cartesian coordinates, this fact trivially arises from the fact that all terms of the form $\left(\frac{\partial^2 F}{\partial x^2} \right)_{jq}^j$ are equal to zero.

This first-order analysis of the Jacobi identity applied to the determinant of the state transition matrix has yielded two interesting sets of n identities that relate the first- and second-order partial derivatives of the dynamics and the flow of a Hamiltonian dynamical system. It does not seem that either equation is identical to the second-order symplectic invariant discussed in the literature⁴ that generalizes the well-known symplectic condition associated with the state transition matrix.^{5,6}

HIGHER-ORDER COVARIANCE PROPAGATION

Taylor Series Based Covariance Propagation

We examine the change in volume of a Gaussian distribution propagated by fitting the first two statistical moments propagated using a second-order approximation of the dynamics. We will assume the Taylor series of the flow-map of a dynamical system over some time t is given by (12). By this approximation, the second-order approximation of the covariance matrix is given in terms of the partial derivatives of the flow-map by^{7,8}

$$\mathbf{P}_t^{ij} = \Phi_p^i \Phi_q^j \mathbf{P}_0^{pq} + \frac{1}{4} \Psi_{k_1 k_2}^i \Psi_{k_3 k_4}^j E(\delta x_0^{k_1} \delta x_0^{k_2} \delta x_0^{k_3} \delta x_0^{k_4}) - \frac{1}{4} \Psi_{k_1 k_2}^i \Psi_{k_3 k_4}^j \mathbf{P}_0^{k_1 k_2} \mathbf{P}_0^{k_3 k_4} \quad (26)$$

$$\mathbf{P}_t = \mathbf{P}_t^{(1)} + \delta \mathbf{P}_t^{(2)} \quad (27)$$

where we have split up the first- and second-order contributions and $\delta \mathbf{x}_0 = \mathbf{x}_0 - \boldsymbol{\mu}_0$. The matrix $\mathbf{P}_t^{(1)}$ represents the linear covariance propagation. While linear covariance propagation can be a poor approximation,⁹ it has the property that it preserves volumes exactly.

Unscented Transform Based Covariance Propagation

The unscented transformation may also be used for higher-order covariance propagation. We specifically consider the scaled unscented transform, which prevents undesirable sigma point spreading with increasing state dimension.¹⁰ Given a distribution with mean $\boldsymbol{\mu}_0$ and covariance \mathbf{P}_0 , the $2n + 1$ input sigma points are formed according to

$$\begin{aligned}\mathcal{X}_0^{(0)} &= \boldsymbol{\mu}_0 \\ \mathcal{X}_0^{(i)} &= \boldsymbol{\mu}_0 + \sqrt{n + \lambda} \left[\sqrt{\mathbf{P}_0} \right]_i \quad i = 1, \dots, n \\ \mathcal{X}_0^{(i+n)} &= \boldsymbol{\mu}_0 - \sqrt{n + \lambda} \left[\sqrt{\mathbf{P}_0} \right]_i \quad i = 1, \dots, n\end{aligned}\tag{28}$$

where

$$\lambda = \alpha^2(n + \kappa) - n\tag{29}$$

and α and κ control the size of the sigma point distribution. The matrix $\sqrt{\mathbf{P}_0}$ in (28) can be any (square) matrix square root of \mathbf{P}_0 , and $[\sqrt{\mathbf{P}_0}]_i$ represents the i^{th} column of this matrix. Associated with each sigma point are weights for the mean and covariance propagation, given by

$$\begin{aligned}w_0^{(m)} &= \frac{\lambda}{n + \lambda} \\ w_0^{(c)} &= \frac{\lambda}{n + \lambda} + (1 - \alpha^2 + \beta) \\ w_i^{(m)} &= \frac{1}{2(n + \lambda)} \quad i = 1, \dots, 2n \\ w_i^{(c)} &= \frac{1}{2(n + \lambda)} \quad i = 1, \dots, 2n\end{aligned}\tag{30}$$

where β is yet another tuning parameter which can be adjusted to account for higher order moments of the input distribution. Throughout this work, $\beta = 2$, which is considered optimal for a Gaussian distribution. Each sigma point is mapped through the nonlinear flow as

$$\mathcal{X}_t^{(i)} = \varphi \left(\mathcal{X}^{(i)}; t \right), \quad i = 0, \dots, 2n\tag{31}$$

The propagated statistics are then recovered from the output sigma points as

$$\begin{aligned}\boldsymbol{\mu}_t &= \sum_{i=0}^{2n} w_i^{(m)} \mathcal{X}_t^{(i)} \\ \mathbf{P}_t &= \sum_{i=0}^{2n} w_i^{(c)} \left(\mathcal{X}_t^{(i)} - \boldsymbol{\mu}_t \right) \left(\mathcal{X}_t^{(i)} - \boldsymbol{\mu}_t \right)^T\end{aligned}\tag{32}$$

Perturbation Analysis

Considering (26) and recalling that $\mathbf{P}_t^{(1)}$ is equal in volume to \mathbf{P}_0 , the perturbing effect due to $\delta\mathbf{P}_t^{(2)}$ is of interest. With this, the expansion of the determinant of the covariance up to second-order is given by applying the Jacobi formula as

$$\det(\mathbf{P}_t) = \det\left(\mathbf{P}_t^{(1)}\right) \left(1 + \text{tr} \left(\left(\mathbf{P}_t^{(1)} \right)^{-1} \delta\mathbf{P}_t^{(2)} \right) \right) + \mathcal{O} \left(\left\| \delta\mathbf{P}_t^{(2)} \right\|_F^2 \right)\tag{33}$$

where $\|\cdot\|_F$ represents the Frobenius norm. The perturbing term is given by

$$\text{tr} \left(\left(\mathbf{P}_t^{(1)} \right)^{-1} \delta \mathbf{P}_t^{(2)} \right) = \left(\left(\mathbf{P}_t^{(1)} \right)^{-1} \right)_{ij} \left(\delta \mathbf{P}_t^{(2)} \right)^{ij} \quad (34)$$

where

$$\left(\mathbf{P}_t^{(1)} \right)_{ij}^{-1} = \left(\Phi^{-1} \right)_i^l \left(\mathbf{P}_0^{-1} \right)_{lm} \left(\Phi^{-1} \right)_j^m \quad (35)$$

The trace of a product of an inverse of a symmetric positive definite matrix and another symmetric positive definite matrix is given by the sum of the generalized eigenvalues. In this context, we have the sum of the extrema of the $\mathbf{P}_t^{(1)}$ -Mahalanobis distance of the marginal variance contribution from the second-order perturbation to the covariance. The trace quantity from (34) can be used as an approximation of the percentage of non-volume preservation from the second-order approximation of the covariance preservation and could be used as an indicator of both the inaccuracy of linear covariance propagation and of the inaccuracy of a moment-matched Gaussian approximation based on the second-order or unscented propagation of the moments.

EXAMPLE

This section examines the volume preservation characteristics of uncertainty propagation methods in a cislunar orbit. The equations of motion for the circular restricted three-body problem (CR3BP) are given in the synodic frame as

$$\frac{d}{dt} \mathbf{x} = \mathbf{F}(\mathbf{x}) \quad (36)$$

$$\mathbf{F}(\mathbf{x}) = \left[\dot{x} \quad \dot{y} \quad \dot{z} \quad 2\dot{y} + \frac{\partial \bar{U}}{\partial x} \quad -2\dot{x} + \frac{\partial \bar{U}}{\partial y} \quad \frac{\partial \bar{U}}{\partial z} \right]^T \quad (37)$$

where $\bar{U}(x, y, z) = \frac{1 - \mu^*}{\|\mathbf{r}_1\|} + \frac{\mu^*}{\|\mathbf{r}_2\|} + \frac{x^2 + y^2}{2}$ is the effective potential, the reduced mass is defined as $\mu^* = \frac{m_2}{m_1 + m_2}$ for the two primary bodies with masses m_1, m_2 respectively. The primary body with greater mass is given index 1 so that $m_1 \geq m_2$. Both masses are located along the x -axis at $[-\mu^*, 0, 0]$ and $[1 - \mu^*, 0, 0]$ with respect to their common barycenter at the origin. The position of the satellite of interest with respect to the primary and secondary bodies is given by \mathbf{r}_1 and \mathbf{r}_2 respectively.¹¹ The reference orbit used in the following sections models the proposed NASA Gateway orbit¹² in the CR3BP and is shown in Figure 1. The initial conditions and mass parameter used for the orbit are

$$\begin{aligned} \mu &= 1.0/(81.30059 + 1.0), & x_0 &= 1.022022, \\ z_0 &= -0.182097, & y_0 &= -0.103256 \end{aligned}$$

in nondimensional units with other initial coordinates equal to zero. This initial condition coincides with the apolune of the orbit. The period of the orbit is approximately 1.511111 nondimensional time units where 2π time units correspond to the period of revolution for the Earth-Moon system. In this example, we employ an initial Gaussian distribution with mean equal to the Gateway initial conditions and covariance given by

$$\mathbf{P}_x = 10^{-8} \text{diag}([1, 0, 1, 0, 0, 0]) + 10^{-10} \mathbf{I}_6 \quad (38)$$

in nondimensional canonical units. The 1-sigma distances are on the order of 40 [km] along the x and z directions, 4 [km] along the y direction, and 0.01 [m/s] in each velocity direction. This is the same orbit and same initial uncertainty as employed in our previous work on Gaussian mixture uncertainty propagation¹³ for ease of comparison.

In Fig. 2, we plot the logarithm of the determinant of the covariance as propagated by the first-order method—which remains fixed over time, along with the covariance propagated by the second-order covariance propagation from (26), its approximation from (33) and (34), and finally from the unscented transform. The determinant of the covariance grows by around ten orders of magnitude around perilune as propagated by the higher-order methods, indicating that the volume grows by five orders of magnitude at this time, representing a serious dilution in the maximum likelihood of any one point as modeled by a Gaussian distribution using the higher-order propagated covariance as opposed to the linearly propagated covariance. This is further resolved in Fig. 3 where the logarithm of the ratio between the second-order covariance determinant and its approximation as well as the original covariance are presented. We can see that generally the approximation of the second-order covariance determinant is fairly accurate except when the second-order covariance determinant blows up, in which case the second-order approximation still qualitatively demonstrated the blow up despite missing the mark by two orders of magnitude.

The unscented transform employed in the study above used scaling parameters $\alpha = 10^{-3}$, $\kappa = 0$, and $\beta = 2$. In order to demonstrate that the particular square root matrix (representing an inverse whitening transform) employed by the unscented transform does not significantly affect the volume of the resulting Gaussian distribution, we uniformly sample from the space of all valid square root matrices, which can be parameterized by

$$\sqrt{\mathbf{P}_0} \in \{\mathbf{S}_0 \mathbf{Q} : \mathbf{Q} \in O(n)\} \quad (39)$$

where $O(n)$ is the orthogonal group in dimension n , and \mathbf{S}_0 is the lower-diagonal Cholesky factor of \mathbf{P}_0 . Random square root matrices are then obtained by drawing Haar-distributed orthogonal matrix samples on $O(n)$.¹⁴ Fig. 4 shows for one hundred such randomly sampled whitening transformations for the unscented transformation, what the resulting ratio between the UT predicted covariance determinant and the original covariance determinant is after one period of propagation with the unscented transform based on each whitening transformation. The distribution has minor spread, indicating that the degree of non-volume preservation of the unscented transform is largely independent of the particular choice of unscented transform. While not plotted here, the invariants in (18) and (22) are conserved nearly up to machine precision in this example.

To get a visual understanding of the effects of volume non-preservation of uncertainty in practice, we performed volume-preserving linear covariance propagation as well as non-volume-preserving unscented uncertainty propagation and compared the results against a Monte Carlo uncertainty propagation with 10,000 samples at perilune of the orbit. The resulting probability density function and scatter plots from the Monte Carlo are displayed in the $x - y$ marginal distribution in Fig. 5. Even though changes in volume do not necessarily carry over from the full 6-dimensional distribution to a lower dimensional distribution, we see that the marginal distribution from the unscented transform is much more spread out than the linear covariance, and that the mean does not necessarily coincide as well as the linear covariance solution does with the mode of the distribution propagated nonlinearly in a Monte Carlo. While the unscented propagated distribution covers more of the support of the Monte Carlo distribution, it also has regions of high likelihood that are outside the support of the true distribution. The opposite is true of the linear covariance propagation, which

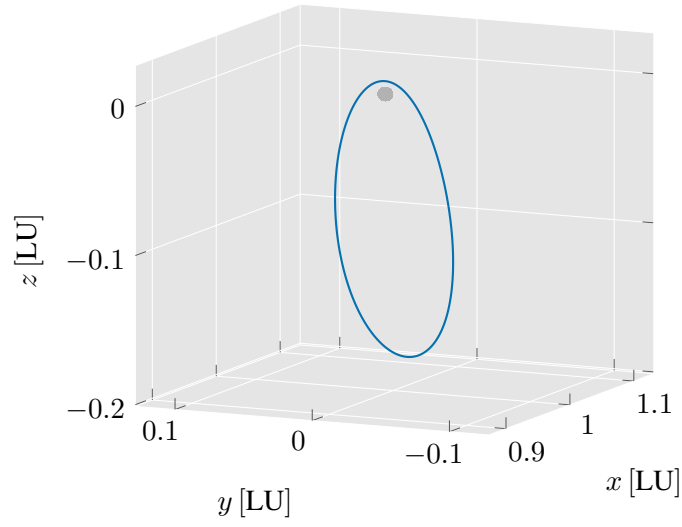


Figure 1. Full NRHO considered in three-body uncertainty propagation application.

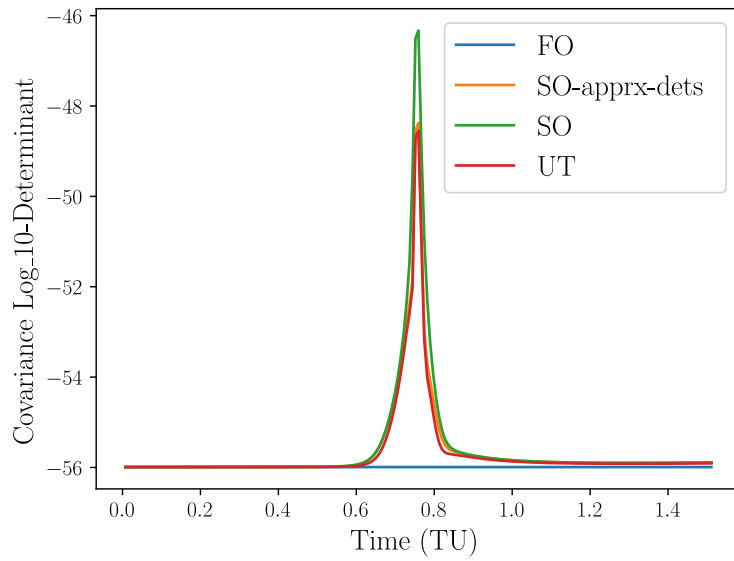


Figure 2. The logarithms of the determinant of the covariance associated with first-order, second-order, trace-based second-order approximate, and unscented propagation of the covariance.

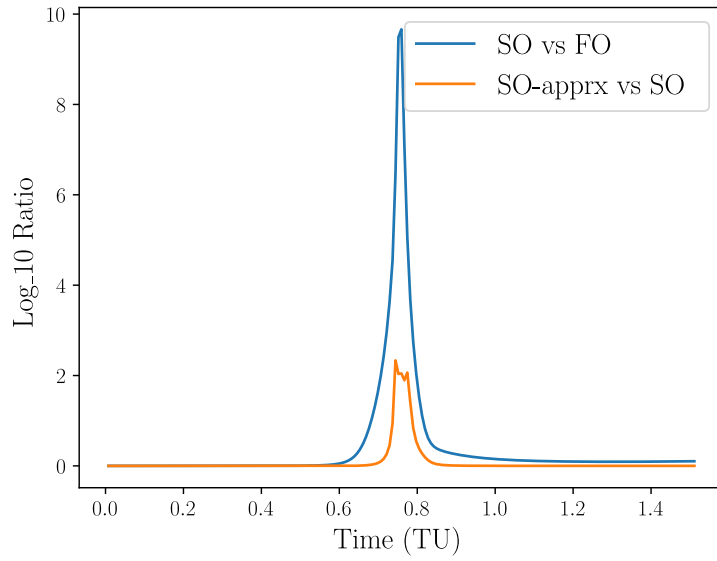


Figure 3. The logarithm of the ratio between the second-order covariance determinant and the first-order covariance determinant (original covariance determinant), as well as the approximate covariance determinant

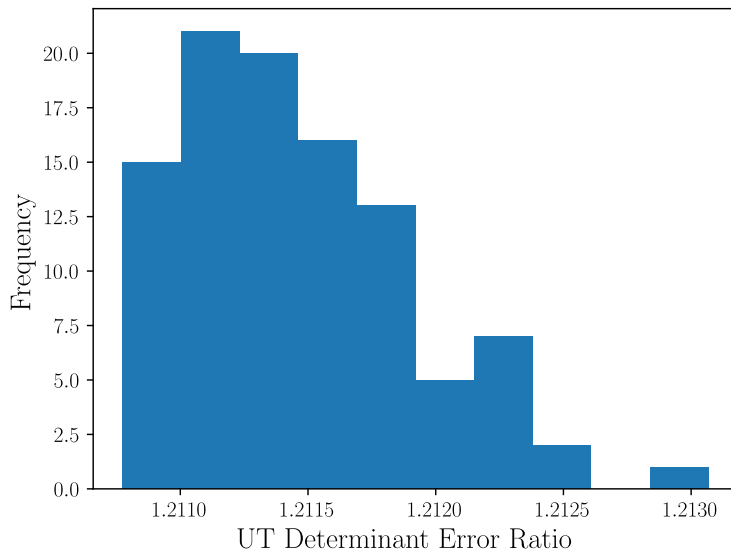


Figure 4. The distribution of covariance determinant growth ratio at the final time under the unscented transforms generated from the Haar measure on orthogonal matrices.

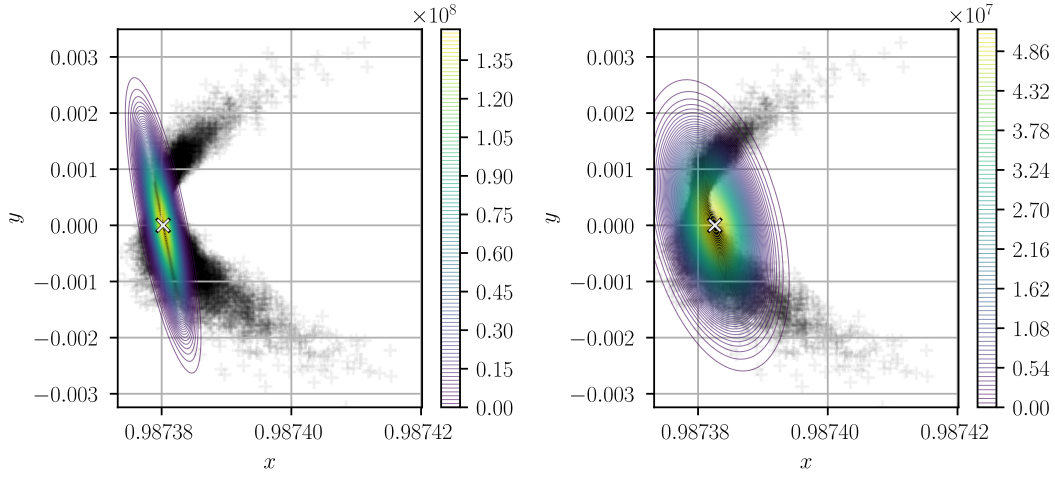


Figure 5. $x - y$ marginal densities from linear covariance (left) versus unscented (right) uncertainty propagation against a Monte Carlo distribution at reference per-illune

covers less of the support of the distribution but coincides better with the most dense parts of the true distribution. This is reflected in a number of statistical metrics of agreement.

To quantify the differences between the linear covariance propagated distribution and the unscented transform propagated distribution to the reference Monte Carlo empirical distribution, we employ the Cramér von Mises (CvM) norm metric. The CvM metric between a univariate marginal distribution and its empirical approximation is given by

$$\omega_j^2 = \int_{-\infty}^{\infty} (F^*(x^j) - F_N(x^j))^2 dF^*(x^j) \quad (40)$$

where $F^*(x^j)$ is the cumulative marginal distribution and $F_N(x^j)$ is the N -sample empirical distribution.¹⁵ To assess the goodness-of-fit of multivariate distributions, we simply consider the 2-norm of the marginal metrics arranged as a vector as

$$\|\boldsymbol{\omega}^2\|_2 = \|\left[\omega_1^2 \quad \dots \quad \omega_m^2\right]\|_2 \quad (41)$$

In order to obtain the marginal cumulative distributions $F_*(x^j)$, we approximate the propagated distribution as Gaussian with moments matching those obtained via the linear covariance and unscented transformation methods.

The expected likelihood kernel (ELK)¹⁶ is also used to compare the moment-matched Gaussian approximations against the reference empirical distribution. If \mathbf{x} has probability density p and \mathbf{x}' has probability density p' then the ELK is

$$l^2(p, p') = E_{p'}[p(\mathbf{x})] = \int_{\Omega} p(\mathbf{z})p'(\mathbf{x})d\mathbf{x} \approx \frac{1}{N} \sum_{i=1}^N p(\mathbf{x}'_{(i)}) \quad (42)$$

where Ω is the support of the probability density functions. The ELK is also known as the likelihood agreement measure (LAM) in other work (see, e.g.,¹⁷).

Table 1 shows the ELK and CvM metrics for the two approaches to uncertainty propagation when compared with the true distribution as reflected by the Monte Carlo study. One can see that

the linear covariance propagation has a higher ELK value which indicates better agreement of the distributions on locations of high density. In fact, the ELK is a measure of modal agreement more so than a measure of overall distribution goodness-of-fit, despite its popular usage for the latter. This property of ELK is discussed in the following subsection.

On the other hand, there is better overall agreement between the unscented propagation distribution and the true distribution on average over the entire domain as measured by the CvM metric (which attains its minimum of zero when the two distributions are identical, though the multi-dimensional CvM metric can also become zero for distributions which merely have identical marginal distributions). Additionally, Fig. 5 shows that the highest likelihood of either marginal distribution differs by nearly an order of magnitude in support of the idea that even the marginal distribution is less dense and higher volume in the case of unscented propagation. In some ways, this shows that higher-order Gaussian propagation (such as the unscented transform or second-order covariance propagation) tends to be more conservative than linear covariance propagation, though at the loss of agreement of the distribution near the mode of the true distribution. Overall, both central moment based approximations fail to capture the true density structure, motivating adaptive Gaussian mixture approaches.^{18,19}

	ELK	CvM Norm
Linear	3.789×10^{22}	1348.363
Unscented	2.516×10^{20}	231.256

Table 1. Distribution error metrics at reference perilune passage.

Limitation of ELK for Measuring Distribution Goodness-of-Fit

The behavior of ELK can be illustrated via the Holder inequality with exponents 1 and ∞ respectively:

$$\int_{\Omega} p(\mathbf{z})p'(\mathbf{z})d\mathbf{z} \leq \max_{\Omega} p(\mathbf{z}) \int_{\Omega} p'(\mathbf{z})d\mathbf{z} = \max_{\Omega} p(\mathbf{z}) \quad (43)$$

where the Dirac distribution located at the mode of the original distribution attains the highest ELK of any distribution, including the original distribution. For example, ELK for a standard univariate Gaussian with itself is $1/2\sqrt{\pi}$ while the ELK between a standard univariate Gaussian and the Dirac distribution at zero is $1/\sqrt{2\pi}$ which is greater by a factor of $\sqrt{2}$. This indicates that ELK is a better metric for modal agreement than for overall agreement of two distributions.

CONCLUSION

In this work, we have derived an approximation for the increase in volume of covariance ellipsoids under second-order approximations of the flow of a dynamical system. Additionally, we developed a novel conserved quantity in the second-order approximation of the flow of a volume-preserving dynamical system. In a cislunar example, we demonstrated that higher-order covariance propagation can lead to increases in covariance ellipsoids volume increases of many orders of magnitude despite the system being volume-preserving. Further, we experimentally demonstrated that the particular choice of unscented transform does not significantly affect the level of volume preservation that the unscented transform exhibits when employed with cislunar uncertainty propagation. Linear

covariance propagation, which is volume-preserving, is shown to underestimate the true density in key regions of the effective support, whereas higher-order propagation inflates volume to better encapsulate the effective support, consequently overestimating the density in low probability regions.

ACKNOWLEDGMENT

We would like to thank Oliver Boodram for helpful discussions. Part of this research was sponsored by the United States Air Force Research Laboratory and the United States AFRL Regional Hub and was accomplished under Cooperative Agreement Number FA8750-22-2-0501. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the United States Air Force or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

REFERENCES

- [1] R. Weisman, M. Majji, and K. T. Alfriend, "Solution of Liouville's Equation for Uncertainty Characterization of the Main Problem in Satellite," *Tech Science Press CMES*, Vol. 111, No. 3, 2016, pp. 269–304.
- [2] R. J. Hyndman, "Computing and graphing highest density regions," *The American Statistician*, Vol. 50, No. 2, 1996, pp. 120–126.
- [3] O. Boodram, Y. Khatri, and D. Scheeres, "Investigating the Presence of Symplectic Structure in Gaussian-Mixture Orbit Updates," *AAS Spaceflight Mechanics Meeting*, 2025.
- [4] A. Bani Younes, "Exact computation of high-order state transition tensors for perturbed orbital motion," *Journal of Guidance, Control, and Dynamics*, Vol. 42, No. 6, 2019, pp. 1365–1371.
- [5] R. Broucke, H. Lass, and D. Boggs, "A note on the solution of the variational equations of a class of dynamical systems," *Celestial mechanics*, Vol. 14, No. 3, 1976, pp. 383–392.
- [6] J. T. Peterson, M. Majji, and J. Junkins, "A Generalized Condition for Symplectic Dynamics and the State Transition Matrix,"
- [7] 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.
- [8] M. Valli, R. Armellin, P. Di Lizia, and M. R. Lavagna, "Nonlinear mapping of uncertainties in celestial mechanics," *Journal of Guidance, Control, and Dynamics*, Vol. 36, No. 1, 2013, pp. 48–63.
- [9] J. Kulik, B. Hasting, and K. A. LeGrand, "Linear Covariance Fidelity Checks and Measures of Non-Gaussianity," *2025 AAS/AIAA Spaceflight Mechanics Meeting*, 2025.
- [10] R. V. D. Merwe, *Sigma-Point Kalman Filters for Probabilistic Inference in Dynamic State-Space Models*. PhD thesis, 2004.
- [11] 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.
- [12] "National Aeronautics and Space Administration (NASA) White Paper: Gateway Destination Orbit Model: A Continuous 15 Year NRHO Reference Trajectory," 2019.
- [13] J. Kulik and K. A. LeGrand, "Nonlinearity and Uncertainty Informed Moment-Matching Gaussian Mixture Splitting," *arXiv preprint arXiv:2412.00343*, 2024.
- [14] A. Edelman and N. R. Rao, "Random Matrix Theory," *Acta Numerica*, Vol. 14, May 2005, pp. 233–297, 10.1017/s0962492904000236.
- [15] D. A. Darling, "The Kolmogorov-Smirnov, Cramer-von Mises Tests," *The Annals of Mathematical Statistics*, Vol. 28, Dec. 1957, pp. 823–838, 10.1214/aoms/1177706788.
- [16] T. Jebara, R. Kondor, and A. Howard, "Probability Product Kernels," *Journal of Machine Learning Research*, Vol. 5, Dec. 2004, pp. 819–844.
- [17] K. J. DeMars, R. H. Bishop, and M. K. Jah, "Entropy-based approach for uncertainty propagation of nonlinear dynamical systems," *Journal of Guidance, Control, and Dynamics*, Vol. 36, No. 4, 2013, pp. 1047–1057.
- [18] J. L. Iannamorelli and K. A. LeGrand, "Adaptive Gaussian Mixture Filtering for Multi-sensor Maneuvering Cislunar Space Object Tracking," *The Journal of the Astronautical Sciences*, Vol. 72, Jan. 2025, p. 2, 10.1007/s40295-024-00478-z.

- [19] G. A. Siciliano, K. A. LeGrand, and J. Kulik, “Higher-Order Tensor-Based Deferral of Gaussian Splitting for Orbit Uncertainty Propagation,” July 2025, 10.48550/arXiv.2507.01771.