Adaptive Gaussian Mixture Method for Uncertainty Propagation in Space Surveillance*

Jan Krejčí and Ondřej Straka

Department of Cybernetics, University of West Bohemia
Univerzitní 8, 306 14 Pilsen, Czech Republic
{jkrejci,straka30}@kky.zcu.cz
WWW home page: https://idm.kky.zcu.cz.

Abstract. The orbital uncertainty propagation problem is treated, where the uncertainty concerning an object position and velocity is propagated over a long time interval because observations are scarce. This paper focuses on the Gaussian mixture description of the uncertainty and proposes a method that adaptively changes the number of mixture components to represent the uncertainty efficiently. The proposed method uses the mean square error based measure of nonlinearity to generate a decision whether the mixture components should be split in order to preserve the fidelity of the uncertainty description. Further, the paper analyzes the performance of four local propagators, which propagate individual mixture components in time. The performance analysis is accomplished using a low-earth orbit scenario.

Keywords: Gaussian mixture, measures of nonlinearity, uncertainty propagation, space surveillance, orbital mechanics

1 Introduction

Uncertainty propagation is a domain of interest in many fields of research, including signal processing [1], target tracking [3], fault detection [12], optimal and predictive control problems [4], and space surveillance. The space surveillance refers to the study and monitoring of space objects. An important issue in the space surveillance, is tracking and uncertainty quantification of near-Earth orbiting resident space objects (RSOs)[1] along their orbital trajectories. The need of efficient and accurate methods for long time period uncertainty propagation is essential in many safety-critical functions, such as filtering, conjunction analysis, calculation of collision probability, sensor resource management, and anomaly detection [27]. The need is also crucial due to the number of RSOs, which grows rapidly in time. The number was beyond 17,000 by December 6, 2016 [27]. The space is a data-sparse environment and therefore, the evolution of the uncertainty for long time periods is strictly based on a mathematical dynamic model.

* This work was supported by the Czech Ministry of Education, Youth and Sports, project LO1506.

[1] RSOs are typically defined to be of a size equal or larger than 10 cm [27].

Such problem is commonly known as the orbital uncertainty propagation (OUP), and has recently gained considerable attention.

The OUP problem is commonly given by a nonlinear (stochastic) ordinary differential equation (ODE) with an initial condition given in a probabilistic manner, mostly by a mean and a covariance matrix, representing a (Gaussian) probability density function (PDF). The accurate long time period OUP is always an extremely computationally difficult process, because obtaining the true PDF at each time instant involves a solution to the Fokker-Planck equation (FPE) [13], which is a partial differential equation.

Various uncertainty propagation methods, also called propagators, use different approaches to uncertainty representation and its transformation. The uncertainty represented as a PDF is commonly expressed by a Dirac mixture (DM) PDF, a Gaussian PDF, or a Gaussian mixture (GM) PDF. Monte Carlo (MC) simulation is a simple, though computationally intensive method, expressing the uncertainty using the DM PDF. It involves solutions to multiple ODEs with initial values drawn randomly according to the OUP initial condition. It has a special position among other methods, because its results can be considered reliable to validate other uncertainty propagation methods.

Methods that use a Gaussian PDF provide a transformation of the first two statistical moments (mean and covariance matrix) only. Usually, they are computationally undemanding, but taking the nonlinear setting into account, the Gaussian PDF is unable to fully capture the true PDF, which, even though given as Gaussian at the beginning, quickly becomes non-Gaussian. These methods can be divided into i) methods that use linearization of some kind, such as the linear covariance analysis (LinCov) [17], first order Taylor expansion based propagation (FOTE), and the covariance analysis describing function technique (CADET) [9] and ii) sigma-point based methods such as the unscented transformation (UT) [14] that uses nonlinear transformation of the sigma-points.

The advantage of using the GM PDF for the uncertainty representation is, that it can describe any PDF with arbitrary accuracy [21]. Each GM component (term) is fully described by its mean and covariance matrix, and captures the uncertainty of the RSO within a small area. The nonlinear effect of the propagation diminishes when reducing the area and thus for smaller areas the effect can be considered linear. As a consequence, the GM PDF is better suited to the OUP than a single Gaussian PDF as its support can be thought of as a union of smaller areas. To update each GM term in time, the previously mentioned propagators which use the Gaussian PDF can be used and hence they will be called the local propagators hereafter.

Other methods proposed for the OUP include techniques such as the polynomial chaos expansion (PC) [25], state transition tensors [18, 26], differential algebra technique [22], or numerically solving Fokker-Planck equation [10]. A comprehensive study of all the mentioned propagators is given in [27].

Let us point out a recently developed method, called the adaptive entropy-based Gaussian-mixture information synthesis (AEGIS) [6]. This method uses an adaptive GM approximation in order to capture efficiently the temporal change

of the true PDF from the initial Gaussian, to a non-Gaussian PDF. A powerful technique, that can be injected into the GM framework, is adaptive splitting of the GM components in order to decrease the approximation error of the true PDF caused by the nonlinear transformation. A quantification of such error is crucial in order to determine the time instants for the splitting. In the AEGIS, the measure used to determine the time instants is based on the differential entropy [2]. More precisely, for each GM component, the difference between the differential entropy of linear (FOTE) and nonlinear (UT) propagation is assessed. If the entropy difference exceeds a certain user-defined threshold, the GM term is split into a GM. This allows a more reliable representation of the uncertainty. The disadvantage of the differential entropy is that it is unbounded, which deteriorates its tractability and the ability of adjusting the threshold [15].

The goal of the paper is twofold. First, to avoid the threshold specification problems of the differential entropy, a measure of nonlinearity (MoNL) will be used in the AEGIS, for which the threshold can be specified conveniently. More specifically, the mean square error MoNL will be used, as suggested in [11]. Second, the performance of the designed adaptive uncertainty propagation method will be analyzed for different local propagators in order to select the propagator with the best performance.

The rest of the paper is organized as follows: In Section 2, the OUP problem is stated, following with a presentation of the local propagators and a general description of the AEGIS method. In Section 3, the mean square error based measure of nonlinearity is presented, followed by a description of the proposed adaptive GM method in Section 4. The performance analysis is given in Section 5 and the paper is concluded in Section 6.

2 Problem Formulation and Related Work

In this section, the OUP propagation is presented. The summary of local propagators, which will be analyzed within the proposed method, follows together with a general description of the AEGIS method.

First, the Gaussian PDF and the GM PDF are stated. The PDF of a Gaussian random variable $\mathbf{x} \in \mathbb{R}^n,\ n \in \mathbb{N}$ is given by

$$p(\mathbf{x}) = \mathcal{N}(\mathbf{x}; \mathbf{m}, \mathbf{P}) = \frac{1}{\sqrt{\det(2\pi\mathbf{P})}} \exp\left(-\frac{1}{2}(\mathbf{x}-\mathbf{m})\mathbf{P}^{-1}(\mathbf{x}-\mathbf{m})^T \right), \tag{1}$$

where $\mathbf{m} \in \mathbb{R}^n$ is mean, and $\mathbf{P} = \mathbf{P}^T \in \mathbb{R}^{n\times n},\ \mathbf{P} \succ 0$ is covariance matrix. The GM PDF is given by a sum of weighted Gaussian PDFs,

$$p_{\mathrm{GM}}(\mathbf{x}) = \sum\nolimits_{i=1}^{M} w_i \mathcal{N}(\mathbf{x}; \mathbf{m}_i, \mathbf{P}_i), \tag{2}$$

where M is the number of components, $w_i \geq 0$ is weight of the i-th GM component such that $\sum_{i=1}^{M} w_i = 1$, and $\mathcal{N}(\mathbf{x}; \mathbf{m}_i, \mathbf{P}_i)$ is the GM component itself $\forall i \in \{1, 2, \ldots, M\},\ M \in \mathbb{N}$.

2.1 Orbital Uncertainty Propagation Problem

The differential equations governing the movement of any RSO can be given in multiple ways. Consider the following dynamic deterministic models commonly used in the space surveillance:

$$\dot{\mathbf{r}} = \mathbf{v}, \qquad \dot{\mathbf{v}} = -\frac{\mu}{r^3}\mathbf{r} + \mathbf{a}_{\text{pert}}(\mathbf{r}, \mathbf{v}), \tag{3}$$

where $\mathbf{r}$ and $\mathbf{v}$ are the position and velocity of the RSO in the Earth-centered inertial (ECI) coordinates, respectively, $r = \|\mathbf{r}\|_2$ is its distance from the center of the Earth, μ is the gravitational constant of the Earth, and the term $\mathbf{a}_{\text{pert}}$ refers to the acceleration perturbations. The initial conditions for (3) are given by random variables, $\mathbf{r}(t_0) \sim p(\mathbf{r}(t_0))$ and $\mathbf{v}(t_0) \sim p(\mathbf{v}(t_0))$, which both can be well described by the Gaussian distribution.

The equations given in (3) can be rewritten in terms of the ODE

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), t), \quad \mathbf{x}(t_0) \sim p(\mathbf{x}(t_0), t_0) \tag{4}$$

commonly known as the nonlinear state-space model with an uncertain initial condition and $\mathbf{x} = \mathbf{x}(t) \in \mathbb{R}^n$ being the state, and with (omitting time indices)

$$\mathbf{x} = \begin{bmatrix} \mathbf{r} \\ \mathbf{v} \end{bmatrix}, \quad \text{and} \quad \mathbf{f}(\mathbf{x}) = \begin{bmatrix} \mathbf{v} \\ -\mu r^{-3}\mathbf{r} + \mathbf{a}_{\text{pert}}(\mathbf{r}, \mathbf{v}) \end{bmatrix}. \tag{5}$$

Solution to (4) with a deterministic initial condition $\mathbf{x}_0$ can be written as

$$\mathbf{x}(t) = \phi(\mathbf{x}(t_0), t), \tag{6}$$

where $\phi(\mathbf{x}(t_0), t)$ is the state solution flow.

Considering for simplicity that the RSO movement is confined in to the equatorial plane, the two-dimensional description of the position, $\mathbf{r}(t) = [x(t)\ y(t)]^T$, and velocity, $\mathbf{v}(t) = [u(t)\ v(t)]^T$ of the RSO, can be used, and therefore, $\mathbf{x}(t) \in \mathbb{R}^4$.

As stated in the introduction, the time evolution of $p(\mathbf{x}(t), t)$ is given by the solution to the Fokker-Planck equation (FPE) as

$$\frac{\partial}{\partial t} p(\mathbf{x}(t), t) = -\sum_{i=1}^{4} \frac{\partial}{\partial x_i(t)} [p(\mathbf{x}(t), t) f_i(\mathbf{x}(t), t)], \tag{7}$$

where f_i is the i-th element of $\mathbf{f}$ in (4).

In the high-earth-orbit (HEO) scenario, the term $\mathbf{a}_{\text{pert}}$ is usually neglected. In the low-earth-orbit (LEO) scenario, the term $\mathbf{a}_{\text{pert}}$ can be modeled using the exponential atmospheric density model [6] as

$$\mathbf{a}_{\text{pert}}(\mathbf{r}(t), \mathbf{v}(t), t) = -\frac{1}{2}\rho_0 \exp\left(-\frac{r(t) - R - h_0}{h_s} \right) \beta\, v_{\text{rel}}(t) \mathbf{v}_{\text{rel}}(t), \tag{8}$$

where ρ_0 is the reference atmospheric mass density, R is the Earth radius, h_0 is the reference height, h_s is the scale height, β is the ballistic coefficient, and

omitting the time index, $\mathbf{v}_{\text{rel}} = [u - \omega y \quad v + \omega x]^T$ with ω being the angular velocity of the Earth, and $v_{\text{rel}} = \|\mathbf{v}_{\text{rel}}\|_2$.

Note that the ballistic coefficient β and the atmospheric mass density modelled by the term $\rho_0 \exp(-\frac{r(t)-R-h_0}{h_s})$ are often uncertain and should be modelled as stochastic variables [5,19,23]. However, their stochastic properties are seldom provided in the literature. For convenience, this paper considers deterministic ballistic coefficient and atmospheric mass density even though some of the propagators can easily take into account such uncertainty.

2.2 Local Propagators

Consider the initial PDF for the OUP given in terms of eq. (4), $\mathbf{x}(t_0)$, being a Gaussian random variable with mean $\mathbf{m}(t_0) = \mathbf{m}_0$ and covariance matrix $\mathbf{P}(t_0) = \mathbf{P}_0$. The local propagators then deal with approximating the mean $\mathbf{m}(t)$ and covariance matrix $\mathbf{P}(t)$, related to $\mathbf{x}(t)$ at any desired time t, as a subject to the nonlinear propagation. The analyzed local propagators are the FOTE based propagation, LinCov, CADET, and the propagation based on the UT.

FOTE: The first order Taylor expansion based propagation is a local linearization method, based on replacing the ODE governing function $\mathbf{f}(\mathbf{x}(t), t)$ with its first order Taylor series expansion, along some reference trajectory. Assuming the reference trajectory is an approximated mean $\mathbf{m}(t)$, obtained as the solution to (4) with the initial condition $\mathbf{x}(t_0) = \mathbf{m}_0$, the new ODE can be written as

$$\dot{\mathbf{x}}(t) \approx \mathbf{f}(\mathbf{m}(t), t) + \mathbf{A}(\mathbf{m}(t), t)[\mathbf{x}(t) - \mathbf{m}(t)], \tag{9}$$

where

$$\mathbf{A}(\mathbf{m}(t), t) = \left.\frac{\partial \mathbf{f}(\mathbf{x}(t), t)}{\partial \mathbf{x}(t)}\right|_{\mathbf{x}(t)=\mathbf{m}(t)}. \tag{10}$$

The propagation of the mean and covariance matrix is then given by

$$\dot{\mathbf{m}}(t) = \mathbf{f}(\mathbf{m}(t), t), \qquad \mathbf{m}(t_0) = \mathbf{m}_0, \tag{11}$$

$$\dot{\mathbf{P}}(t) = \mathbf{A}(\mathbf{m}(t), t)\mathbf{P}(t) + \mathbf{P}(t)\mathbf{A}^T(\mathbf{m}(t), t), \quad \mathbf{P}(t_0) = \mathbf{P}_0. \tag{12}$$

Note that the FOTE based propagation has no user-design parameter.

LinCov: Linear covariance analysis is also a local linearization method, based on replacing the ODE governing function $\mathbf{f}(\mathbf{x}(t), t)$ with its first order Taylor series expansion, along some reference trajectory. Assuming the reference trajectory is an approximated mean $\mathbf{m}(t)$, satisfying the eq. (11), the evolution of an error, defined as $\mathbf{e}(t) = \mathbf{x}(t) - \mathbf{m}(t)$, is given by

$$\dot{\mathbf{e}}(t) \approx \mathbf{A}(\mathbf{m}(t), t)\mathbf{e}(t). \tag{13}$$

Assuming discrete time instants $t_k,\ k \in \{1, 2, \dots\}$, denote Δt_k the difference between the time instants t_k and t_{k+1}, which can be viewed as a user-design parameter. Then, the propagation of the covariance matrix between the time instants is given by

$$\mathbf{P}(t_{k+1}) = \mathbf{\Phi}(t_k, t_{k+1})\mathbf{P}(t_k)\mathbf{\Phi}^T(t_k, t_{k+1}), \quad \mathbf{P}(t_0) = \mathbf{P}_0, \tag{14}$$

where

$$\mathbf{\Phi}(t_k, t_{k+1}) = e^{\mathbf{A}(\mathbf{m}(t_{k+1}), t_{k+1})\Delta t_k} \tag{15}$$

is called the state transition matrix. Note that the difference between the FOTE based propagation and the LinCov method is a coarser approximation of the covariance matrix propagation. As the LinCov does not need to solve another ODE for finding the covariance matrix evolution, it should be less computationally intensive than the FOTE based propagation, depending on the $\Delta t_k,\ \forall k$, and the ODE solver.

CADET: The covariance analysis describing function technique [9] is a method that uses a statistical linearization as

$$\mathbf{x}(t) \approx \mathbf{m}(t) + \mathbf{e}(t), \tag{16}$$

$$\mathbf{f}(\mathbf{x}(t), t) \approx \mathbf{N}_\mathbf{m}(t)\mathbf{m}(t) + \mathbf{N}_\mathbf{e}(t)\mathbf{r}(t), \tag{17}$$

where $\mathbf{e}(t)$ is a zero-mean independent random process, which covariance matrix is assumed to be $\mathbf{P}(t)$, representing the state deviations, and $\mathbf{N}_\mathbf{m}(t),\ \mathbf{N}_\mathbf{e}(t)$ are the so-called multiple-input describing function gain matrices that are chosen to minimize the mean square error of approximation of the function $\mathbf{f}(\mathbf{x}(t), t)$. The further derivation results in the propagation scheme given by

$$\dot{\mathbf{m}}(t) = E[\mathbf{f}(\mathbf{x}(t), t)], \qquad \mathbf{m}(t_0) = \mathbf{m}_0, \tag{18}$$

$$\dot{\mathbf{P}}(t) = E[\mathbf{f}(\mathbf{x}(t), t)\mathbf{e}^T(t)] + E[\mathbf{e}(t)\mathbf{f}^T(\mathbf{x}(t), t)], \qquad \mathbf{P}(t_0) = \mathbf{P}_0 \tag{19}$$

where $E[.]$ denotes the expectation operator. Calculating their expectations $E[\mathbf{f}(\mathbf{x}(t), t)]$ and $E[\mathbf{f}(\mathbf{x}(t), t)\mathbf{e}^T(t)]$ can be difficult because it involves multivariate integration and evaluation of the state PDF at each time instant t. Note that within the simulations, the UT approximation was used for this purpose.

UT: The unscented transformation [14] is a nonlinear method based on the idea that it may be easier to approximate the state PDF than to approximate the nonlinear transformation. It uses a set of so-called sigma-points $\{\mathcal{X}_j(t_0),\ j = 1, \dots, 2n\}$, which are carefully chosen points from the initial PDF as follows. First, the square root decomposition $\sqrt{\mathbf{P}(t_0)}$ of the covariance matrix $\mathbf{P}(t_0)$ is found such that $\sqrt{\mathbf{P}(t_0)}(\sqrt{\mathbf{P}(t_0)})^T = \mathbf{P}(t_0)$. Then, the sigma-points are given by

$$\begin{aligned} \mathcal{X}_j(t_0) &= \mathbf{m}(t_0) + \sqrt{n}\big(\sqrt{\mathbf{P}(t_0)}\big)_j, & j = 1, 2, \dots, n, \\ \mathcal{X}_j(t_0) &= \mathbf{m}(t_0) - \sqrt{n}\big(\sqrt{\mathbf{P}(t_0)}\big)_{j-n}, & j = n+1, \dots, 2n, \end{aligned} \tag{20}$$

where $\left(\sqrt{\mathbf{P}(t_0)}\right)_j$ denotes the j-th column of $\sqrt{\mathbf{P}(t_0)}$. Each sigma-point is transformed by the nonlinear ODE,

$$\dot{\mathcal{X}}_j(t) = \mathbf{f}(\mathcal{X}_j(t), t), \quad \text{with the initial condition} \quad \mathcal{X}_j(t_0), \ \forall j, \tag{21}$$

forming the transformed set $\{\mathcal{X}_j(t), \ j = 1, \ldots, 2n\}$ of the sigma-points. The mean and covariance matrix at a desired final time $t > t_0$ are then given by

$$\mathbf{m}(t) = \sum\nolimits_{j=1}^{2n} w_j \mathcal{X}_j(t), \tag{22}$$

$$\mathbf{P}(t) = \sum\nolimits_{j=1}^{2n} w_j [\mathcal{X}_j(t) - \mathbf{m}(t)][\mathcal{X}_j(t) - \mathbf{m}(t)]^T, \tag{23}$$

where $w_j = \frac{1}{2n} \forall l$ are the weights of the sigma-points. Note that the above version of the method has no user-design parameter, and other formulations exist [24].

2.3 Brief description of AEGIS

The adaptive entropy-based Gaussian-mixture information synthesis (AEGIS) was first described in [6]. It is a method representing the RSO uncertainty by a GM PDF that increases the number of its components depending on a difference between the differential entropy evaluated for nonlinear propagated state by the UT and linearly evolved differential entropy calculated for each GM term. Note that the (Shannon) differential entropy [20] for the random variable $\mathbf{x} \in \mathbb{R}^n$ described by PDF $p(\mathbf{x})$, is given by

$$H(\mathbf{x}) = -\int p(\mathbf{x}) \log p(\mathbf{x}) d\mathbf{x} = E[-\log p(\mathbf{x})]. \tag{24}$$

The method can be summarized by the following steps.

Initialization is given by establishing the terms

$$p(\mathbf{x}(t_0)) = \alpha_1 \mathcal{N}(\mathbf{x}(t_0); \mathbf{m}_1(t_o^1), \mathbf{P}_1(t_o^1)), \tag{25}$$

$$H_1^N(t_o^1) = \frac{1}{2} \log |2\pi e \mathbf{P}_1(t_o^1)|, \tag{26}$$

where $p(\mathbf{x}(t_0))$ is the initial GM containing solely the initial Gaussian PDF with $\alpha_1 = 1$, $\mathbf{m}_1(t_o^1) = \mathbf{m}_0$, $\mathbf{P}_1(t_o^1) = \mathbf{P}_0$ where $t_o^1 = t_0$ is the time instant of the *origination* of the initial GM component, and $H_1^N(t_o^1)$ is the differential entropy for the Gaussian distribution. Moreover, the time instants, in which the GM components will be tested to satisfy the reliability criterion, should be established. Note that the superscript N in $H_1^N(t_0)$ stands for "nonlinear".

Propagation is a process pertaining to each GM component for time $t \in [t_o^i, t_s^i]$, where t_o^i stands for the time instant of origination of the i-th GM component,

and t_s^i is yet unknown time instant of splitting the i-th component. Consider the following GM at the time t,

$$p(\mathbf{x}(t)) = \sum\nolimits_{i=1}^{M} \alpha_i \mathcal{N}(\mathbf{x}(t), \mathbf{m}_i(t), \mathbf{P}_i(t)), \tag{27}$$

where $M \in \mathbb{N}$ is the number of the GM components. Each component is being propagated using the UT propagation scheme as discussed above, yielding the mean $\mathbf{m}_i(t)$ and covariance matrix $\mathbf{P}_i(t)$. During the time span $[t_o^i, t_s^i]$, the "nonlinear" differential entropy $H_i^N(t)$, and "linear" differential entropy $H_i^L(t)$ for the i-th component defined as

$$H_i^N(t) = \frac{1}{2} \log |2\pi e \mathbf{P}_i(t)|, \tag{28}$$

$$\dot{H}_i^L(t) = \operatorname{tr}\big(\mathbf{A}(\mathbf{m}_i(t), t)\big), \quad H_i^L(t_0) = H_i^N(t_o^i), \tag{29}$$

where $\mathbf{A}(\mathbf{m}_i(t), t)$ is given by (10), are monitored and compared. Note that eq. (29) describes the evolution of the entropy for the Gaussian PDF of the state governed by (9), which indirectly indicates the FOTE based propagation.

Testing takes place at each time instant, predefined in the initialization step. For each GM component, the reliability condition is defined by the difference between the nonlinear and linear entropy $|H_i^L(t) - H_i^N(t)|$. If the difference exceeds a given threshold, the GM term is no longer considered to reliably represent the corresponding portion of uncertainty due to the nonlinearity, and it should be split.

Splitting takes place if the i-th GM component was found to satisfy the splitting condition defined within the testing process. This time instant is denoted as t_s^i. The propagation process for the GM component is halted and the GM component is then split along the dominant eigenvector, yielding ΔM new Gaussian PDFs, that are directly injected into the current GM to replace the split component. For details of the splitting process, see [6]. The GM after the splitting can be written as

$$p(\mathbf{x}(t)) = \sum\nolimits_{i=1}^{M+\Delta M-1} \alpha_i \mathcal{N}(\mathbf{x}(t), \mathbf{m}_i(t), \mathbf{P}_i(t)). \tag{30}$$

For each new GM component, the corresponding sigma-points are generated and the propagation process continues with setting $t_o^j := t_s^i$, for each index j corresponding to the new components, and $M := M + \Delta M - 1$. The algorithm is illustrated in Fig. 1. The threshold for the splitting of GM components $\mathcal{T}$, can be given as a fraction of the initial differential entropy (26).

Notes:

Note 1: There are two local propagators involved within the AEGIS. The UT is used for the propagation of the GM components, and the FOTE is indirectly used only for the calculation of the linear entropy $H_i^L(t)$.

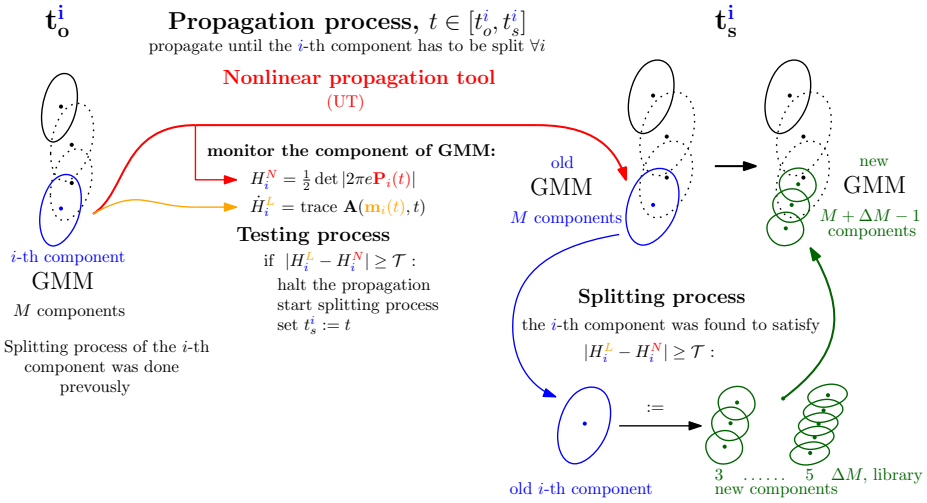


Fig. 1: The AEGIS method illustration

Note 2: The disadvantage of the differential entropy lies in the specification of the threshold $\mathcal{T}$. If the initial covariance matrix is "small" enough to yield a negative entropy, the threshold is negative and, therefore, the reliability condition can never be satisfied. If the threshold is not determined using the initial nonlinear entropy, the choice is ambiguous as the differential entropy is unbounded. The difference between the nonlinear and linear differential entropy can be seen as a measure of nonlinearity of the transformation ϕ in (6) in the region given by the non-negligible support of $\mathcal{N}(\mathbf{x}(t); \mathbf{m}_j(t), \mathbf{P}_j(t))$. Hence, it may be convenient to use a measure of nonlinearity [8] with more becoming properties to govern the splitting of the term $\mathcal{N}(\mathbf{x}(t); \mathbf{m}_j(t), \mathbf{P}_j(t))$

3 Measure of Nonlinearity

As stated above, the differential entropy is unbounded, which deteriorates its tractability and the ability of adjusting the threshold [15]. To tackle this problem, measures of nonlinearity (MoNL) or non-Gaussianity (MoNG) may be involved [11]. An overview of such measures can be found in [8].

As argued in [11], the MoNL which could be conveniently used within the OUP problem is the mean-square error (MSE) based MoNL, defined as follows. Consider a nonlinear function

$$\mathbf{y} = \gamma(\mathbf{u}), \tag{31}$$

where $\mathbf{u} \in \mathbb{R}^{n_u}$, and $\mathbf{y} \in \mathbb{R}^{n_y}$ are both random variables. For the joint distribution of $\begin{bmatrix}\mathbf{u}\\ \mathbf{y}\end{bmatrix}$, let the covariance matrix be denoted as $\begin{bmatrix}\mathbf{P_{uu}} & \mathbf{P_{yu}^T}\\ \mathbf{P_{yu}} & \mathbf{P_{yy}}\end{bmatrix}$. The MSE-MoNL endeavors to measure the L_2-distance between the nonlinear function γ

and its best linear approximation [16]. The measure in the normalized form can be calculated by

$$\nu_{\mathrm{MSE}} = \sqrt{\frac{\mathrm{tr}(\mathbf{P_{yy}} - \mathbf{P_{yu}}\mathbf{P_{uu}}^{-1}\mathbf{P_{yu}}^{T})}{\mathrm{tr}(\mathbf{P_{yy}})}}, \quad \nu_{\mathrm{MSE}} \in [0,1]. \tag{32}$$

The value $\nu_{\mathrm{MSE}} = 0$ indicates that the function γ is linear almost everywhere.

In terms of the nonlinear transformation ϕ given in (6), the MSE-MoNL can be used within the OUP considering $\mathbf{u} = \mathbf{x}(t_0)$ and $\mathbf{y} = \mathbf{x}(t)$. Recalling the UT, assume having the sigma-point sets $\{\mathcal{X}_j(t_0),\ j = 1,\dots,2n\}$, and $\{\mathcal{X}_j(t),\ j = 1,\dots,2n\}$, before and after the nonlinear transformation ϕ (6), respectively. Now, define the cross-covariance matrix $\mathbf{P}_c(t,t_0)$ as

$$\mathbf{P}_c(t,t_0) = \sum\nolimits_{j=1}^{2n} w_j[\mathcal{X}_j(t) - \mathbf{m}(t)][\mathcal{X}_j(t_0) - \mathbf{m}_0]^T. \tag{33}$$

For the joint distribution of $\begin{bmatrix}\mathbf{x}(t_0)\\ \mathbf{x}(t)\end{bmatrix}$, the covariance matrix can be approximated using the UT sigma-points as $\begin{bmatrix}\mathbf{P}(t_0) & \mathbf{P}_c(t,t_0)^T\\ \mathbf{P}_c(t,t_0) & \mathbf{P}(t)\end{bmatrix}$, where $\mathbf{P}(t)$ is also given by the UT using eq. (23). Then the mean-square error based MoNL is computed by

$$\nu_{\mathrm{MSE}}(t) = \sqrt{\frac{\mathrm{tr}\big(\mathbf{P}(t) - \mathbf{P}_c(t,t_0)\mathbf{P}(t_0)^{-1}\mathbf{P}_c(t,t_0)^T\big)}{\mathrm{tr}\big(\mathbf{P}(t)\big)}}. \tag{34}$$

Assume monitoring $\nu_{\mathrm{MSE}}(t)$ for each GM component during the propagation process. The splitting of any GM component can be initiated when $\nu_{\mathrm{MSE}}(t)$ exceeds a given threshold for the particular component. The threshold can be established conveniently, as the measure is bounded between zero and one.

4 The Adaptive GM Algorithm

The proposed algorithm is based on the AEGIS, aiming to tackle the problems with the differential entropy stated above, using the MSE-MoNL instead. Let us distinguish the proposed method from the AEGIS calling it the adaptive MSE-based Gaussian-mixture information synthesis (AMGIS). The method can be summarized by the following steps.

Initialization is given by establishing the terms

$$p(\mathbf{x}(t_0)) = \alpha_1\,\mathcal{N}(\mathbf{x}(t_0);\mathbf{m}_1(t_o^1),\mathbf{P}_1(t_o^1)), \tag{35}$$

$$\mathcal{T} = \in (0,1), \tag{36}$$

where $p(\mathbf{x}(t_0))$ is the initial GM containing solely the initial Gaussian PDF with $\alpha_1 = 1$, $\mathbf{m}_1(t_o^1) = \mathbf{m}_0$, $\mathbf{P}_1(t_o^1) = \mathbf{P}_0$ where $t_o^1 = t_0$ is the time instant

of the origination of the initial GM component, and $\mathcal{T}$ is the threshold representing the maximum L_2-distance between the state solution flow (6) and its best linear approximation, through the propagation. Moreover, the sigma-point set $\{\mathcal{X}_j^1(t_o^1), j = 1, \ldots, 2n\}$ is drawn using (20) for the initial GM term. Again, the time instants, at which the GM terms will be tested to satisfy the splitting condition, should be established.

Propagation is again a process pertaining to each GM term for time $t \in [t_o^i, t_s^i]$, where t_o^i stands for the time instant of the origination of the i-th GM term, and t_s^i is a yet unknown time instant of splitting the i-th term. Consider having the following GM at the time t,

$$p(\mathbf{x}(t)) = \sum\nolimits_{i=1}^{M} \alpha_i \mathcal{N}(\mathbf{x}(t), \mathbf{m}_i(t), \mathbf{P}_i(t)), \tag{37}$$

where $M \in \mathbb{N}$ is the number of the GM components. Each component is being propagated in time, yielding the mean $\mathbf{m}_i(t)$ and covariance matrix $\mathbf{P}_i(t)$. The propagation can be performed by any of the local propagators discussed in Section 2.2. During the time span, the sigma-point set $\{\mathcal{X}_j^i(t_o^i),\ j = 1, \ldots, 2n\}$ *for the i-th GM term* is also being propagated yielding the covariance matrix $\mathbf{P}_i(t)$, and the cross-covariance matrix $\mathbf{P}_{c,i}(t, t_o^i)$ computed using (33) in order to monitor the MSE-MoNL (34) for each GM term as

$$\nu_{\text{MSE}}^i(t) = \sqrt{\frac{\text{tr}\big(\mathbf{P}_i(t) - \mathbf{P}_{c,i}(t, t_o^i)\mathbf{P}_i(t_o^i)^{-1}\mathbf{P}_{c,i}(t, t_o^i)^T\big)}{\text{tr}\big(\mathbf{P}_i(t)\big)}}. \tag{38}$$

Testing takes place at each time instant, predefined in the initialization step. Each GM component is monitored by computing its $\nu_{\text{MSE}}^i(t)$ and its splitting is invoked if the value of MSE-MoNL exceeds the threshold $\mathcal{T}$, as

$$\nu_{\text{MSE}}^i(t) \geq \mathcal{T}. \tag{39}$$

Splitting of the i-th GM term consists of performing the same actions as in the AEGIS discussed above. The AMGIS algorithm is illustrated in Figure 2.

Notes:

As well as in the case of the AEGIS, there are two local propagators involved in the AMGIS. The UT is used for computing the $\nu_{\text{MSE}}^i(t)$, and any of the local propagators (FOTE, LinCov, CADET, UT) can be used for the propagation of the GM terms. The UT can be used for both computing the $\nu_{\text{MSE}}^i(t)$ and the propagation of the GM terms at the same time, in order to reduce the computational costs.

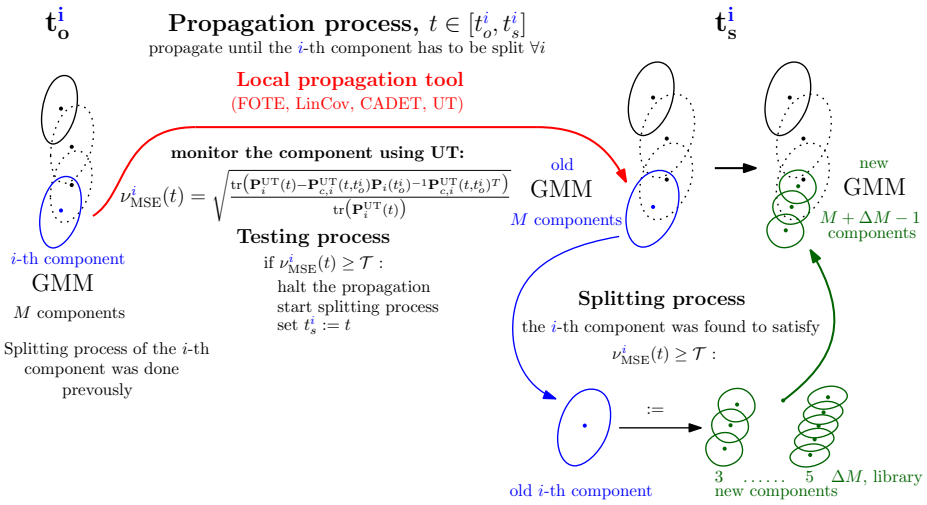


Fig. 2: The AMGIS method illustration

5 Performance Analysis

The performance of the proposed algorithm is tested within the OUP problem for the LEO scenario defined in Section 2.1 with the following initial conditions.

$$\mathbf{m}_0 = \left[2.6 \times 10^3\,\mathrm{km},\, 0\,\mathrm{km},\, 0\,\mathrm{km\,s^{-1}},\, 7.8\,\mathrm{km\,s^{-1}}\right]^T, \tag{40}$$

$$\mathbf{P}_0 = \mathrm{diag}(1.96, 0.25, 6.26 \times 10^{-6}, 2.5 \times 10^{-5}), \tag{41}$$

and the parameters of $\mathbf{a}_{\mathrm{pert}}$ are taken to be $\rho_0 = 3.614 \times 10^{-13}\,\mathrm{kg\,m^{-3}}$, $h_0 = 700\,\mathrm{km}$, $r_s = 88.667\,\mathrm{km}$, $\beta = 1.4$.

The local propagators defined in Section 2.2 were used in the AMGIS algorithm, and the resulting algorithms are denoted as AMGIS–FOTE, AMGIS–LinCov, AMGIS–CADET, and AMGIS–UT. The results are compared using the normalized L_2^2-norm as follows.

5.1 Performance Measure

The square of the L_2-norm measure between the functions $p_1(\mathbf{x})$ and $p_2(\mathbf{x})$ representing distributions of the random variable $\mathbf{x} \in \mathbb{R}^n$ is given by

$$L_2^2(p_1, p_2) = \int_{\mathbb{R}^n} \big(p_1(\mathbf{x}) - p_2(\mathbf{x})\big)^2 d\mathbf{x} \geq 0. \tag{42}$$

The L_2-norm measures the L_2 distance between the distributions and is zero, if $p_1(\mathbf{x}) = p_2(\mathbf{x})$ almost everywhere. Consider its normalized value given by

$$L_{\mathrm{N},2}^2(p_1, p_2) = 1 - \frac{\int_{\mathbb{R}^n} 2\, p_1(\mathbf{x}) p_2(\mathbf{x}) d\mathbf{x}}{\int_{\mathbb{R}^n} p_1^2(\mathbf{x}) d\mathbf{x} + \int_{\mathbb{R}^n} p_2^2(\mathbf{x}) d\mathbf{x}}, \quad L_{\mathrm{N},2}^2(p_1, p_2) \in [0, 1]. \tag{43}$$

For general distributions $p_1(\mathbf{x})$ and $p_2(\mathbf{x})$, the value (43) cannot be calculated analytically, but the example of $p_1(\mathbf{x})$ and $p_2(\mathbf{x})$ being both GM PDFs yields analytical calculation of (43), considering the following relation

$$\int_{\mathbb{R}^n} p_k(\mathbf{x}) p_l(\mathbf{x}) d\mathbf{x} = \sum\nolimits_{i=1}^{M} \sum\nolimits_{j=1}^{N} w_{k,i}\, w_{l,j}\, \mathcal{N}\big(\mathbf{m}_{k,i}; \mathbf{m}_{l,j}, \mathbf{P}_{k,i} + \mathbf{P}_{l,j}\big), \quad (44)$$

for $k, l \in \{1, 2\}$, where $w_{k,i}$, $\mathbf{m}_{k,i}$, and $\mathbf{P}_{k,i}$ are the i-th weight, mean and covariance matrix, respectively, of the GM PDF $p_k(\mathbf{x})$. Within the test simulations, the distributions $p_1(\mathbf{x})$ and $p_2(\mathbf{x})$ will be the resulting GM PDF from the proposed algorithm, and the so-called ground truth GM PDF derived from the Monte Carlo simulation, respectively.

5.2 Calculation of Ground Truth

As stated in the introduction, the MC simulation can be used to validate other uncertainty propagation methods. For the MC simulation, the set of points is drawn randomly from the initial PDF, to be used as a deterministic initial conditions for the ODE (4) and subsequently transformed into a desired time t. The MC simulation results can be described with a DM PDF as

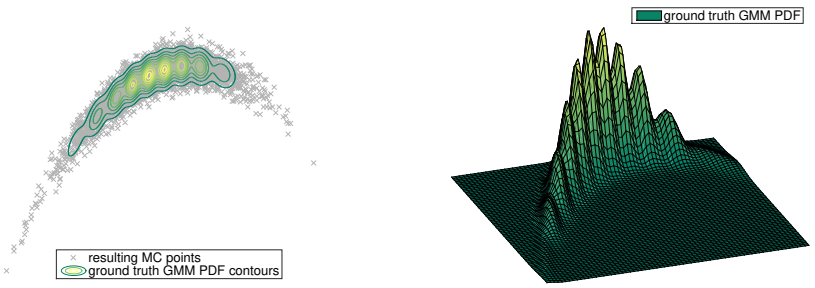
$$p_{\mathrm{MC}}(\mathbf{x}(t)) = \frac{1}{L} \sum_{l=1}^{L} \delta(\mathbf{x}(t) - \mathbf{x}_l(t)), \quad (45)$$

where $\mathbf{x}_l(t) \in \mathbb{R}^n \ \forall l \in \{1, 2, \ldots, L\}$, $L \in \mathbb{N}$ are the locations of the MC points at time t, and δ is the Dirac delta function. Such PDF cannot be used directly to calculate (42). Therefore, w.r.t. the analytical tractability of (42), the DM PDF $p_{\mathrm{MC}}(\mathbf{x}(t))$ will be approximated by a GM PDF approximation $p_{\mathrm{EM}}(\mathbf{x}(t))$, resulting from the EM algorithm [7] fitting at any desired time instant t. Such resulting GM PFD will then be called the ground truth. The normalized L_2-norm can then be readily computed with use of (44). An example of ground truth at a single time instant is depicted in Figs. 3a and 3b.

5.3 Comparison of Local Propagators

To analyze the performance of the algorithms, the uncertainty propagation for a single nominal orbit was considered. The MC method for OUP was performed with 10^5 points, which were then used by the EM to obtain the GM PDF representation of the ground truth.

The algorithms AMGIS–FOTE, AMGIS–LinCov, AMGIS–CADET, and AMGIS–UT were run for the threshold values $\mathcal{T} = 5 \times 10^{-3}, 3 \times 10^{-3}, 2 \times 10^{-3}, 1 \times 10^{-3}, 8 \times 10^{-4}$ and 6×10^{-4}. The comparison of the algorithms in terms of the criterion (43) and the number of terms for $\mathcal{T} = 2 \times 10^{-3}$ is illustrated in Fig. 4 and for $\mathcal{T} = 6 \times 10^{-4}$ in Fig. 5. Note that for each threshold the number of components is almost same for all algorithms. From the figures it follows that the AMGIS–LinCov and AMGIS–FOTE algorithms result in almost same value



(a) the MC points and contours of GM PDF obtained by the EM algorithm

(b) the GM PDF

Fig. 3: An illustration of the position marginal of ground truth with 11 GM components.

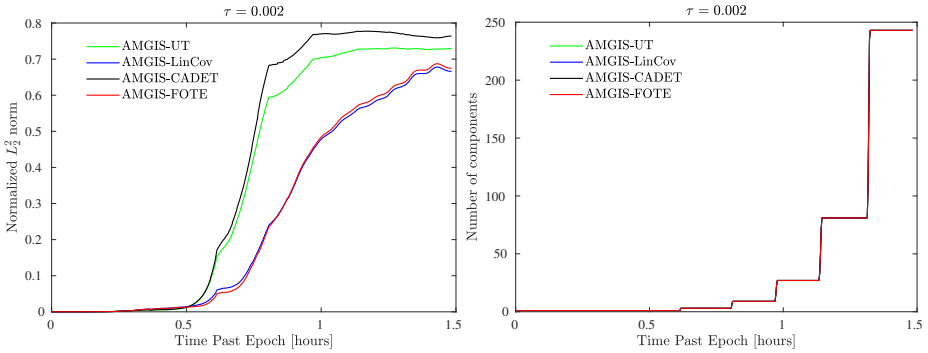
of the criterion (43) and perform better than the AMGIS–CADET and AMGIS–UT algorithms for all values of the threshold.

Performance of the AMGIS–LinCov algorithm for different values of threshold is depicted in Fig. 6 (again in terms of the criterion (43) and the number of components) and similarly performance of the AMGIS-UT algorithm, is depicted in Fig. 7. In both cases, it can be seen that using a lower threshold $\mathcal{T}$ results in GM PDF that is closer to the ground truth but at the cost of high number of GM PDF components.

5.4 Computational time costs

In Table 1, computational costs of the AMGIS algorithm for different local propagators are given. The costs are expressed in terms of duration of an AMGIS run when propagating the uncertainty over the whole orbit. All the numerical simulations in the paper were performed using the R2017b version of Matlab® software running on the PC equipped with Intel® CoreTM i7-4790 CPU (3.60 [GHz]). From the table it is clear that the AMGIS–UT algorithm has the lowest costs, the AMGIS-LinCov and AMGIS-FOTE algorithms have slightly higher costs but the AMGIS-CADET algorithm shows costs almost twice as high as the other methods. The high costs of the AMGIS–CADET follows from the fact that it uses the UT for the moments (18) and (19) computation.

Note that comparison of the AEGIS and the AMGIS algorithms would not make sense as they use different methods to calculate the measure of nonlinearity governing the splitting. Hence, each algorithm can provide uncertainty representation with arbitrary accuracy depending on the threshold used. The advantage of the AMGIS algorithm lies within the favourable properties of the measure used.



(a) (b)

Fig. 4: LEO: Performance analysis of the AMGIS with different local propagators for $\mathcal{T} = 2 \times 10^{-3}$.

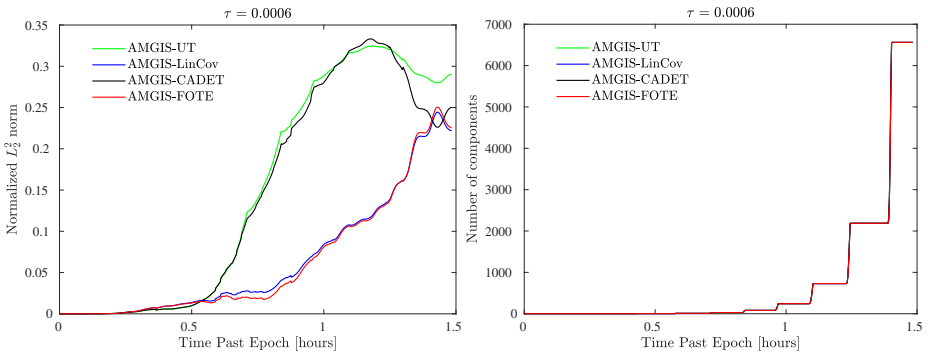
Table 1: Computational costs of an AMGIS run (1 orbit).

	Time (s)			
$\mathcal{T}$	**AMGIS–UT**	**AMGIS–LinCov**	**AMGIS–CADET**	**AMGIS–FOTE**
5×10^{-3}	2.0×10^{1}	2.3×10^{1}	4.4×10^{1}	2.7×10^{1}
3×10^{-3}	6.3×10^{1}	7.4×10^{1}	1.4×10^{2}	8.1×10^{1}
2×10^{-3}	1.9×10^{2}	2.3×10^{2}	4.1×10^{2}	2.4×10^{2}
1×10^{-3}	3.7×10^{3}	4.6×10^{3}	5.6×10^{3}	4.4×10^{3}
8×10^{-4}	7.9×10^{3}	8.3×10^{3}	9.6×10^{3}	8.6×10^{3}
6×10^{-4}	3.8×10^{4}	2.9×10^{4}	3.1×10^{4}	2.9×10^{4}

6 Conclusion and Future Work

The paper dealt with the uncertainty propagation problem, which can be viewed as a time evolution of an uncertainty associated with a residential space object. The adaptive Gaussian mixture method was proposed to efficiently represent the uncertainty using a Gaussian mixture PDF with a variable number of components. The proposed method was built on the adaptive entropy-based Gaussian-mixture information synthesis designed in [6]. The method assesses the nonlinearity of the dynamics during the propagation and the Gaussian mixture components are split if strong non-linearity is detected, i.e. the measure exceeds the threshold. The assessment utilizes the mean square error based measure of nonlinearity.

Further, the paper focused on a performance analysis of four local propagators, which propagate individual Gaussian mixture components in time. The



(a) (b)

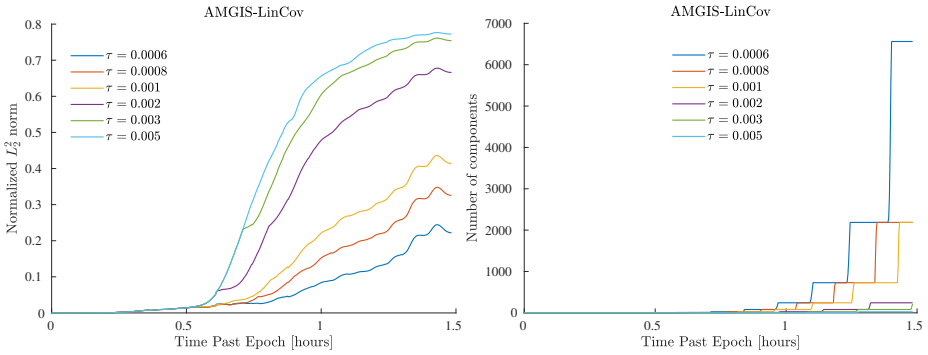
Fig. 5: LEO: Performance analysis of the AMGIS with different local propagators for $\mathcal{T} = 6 \times 10^{-4}$.

analysis was performed using the circular low-earth-orbit scenario. The results indicate that decreasing the threshold leads to more precise approximation of the uncertainty, but with considerably high computational costs given by a rapid increase of number of Gaussian mixture components. The analysis revealed that the linear covariance analysis (LinCov) and the first order Taylor-expansion (FOTE) provide significantly more accurate uncertainty representations than the unscented transform (UT) and the covariance analysis describing function technique (CADET) for any nonlinearity measure threshold.

Future work may include further study of a threshold specification, which may vary over the Gaussian mixture components, or the effect of the direction in which the components are to be split Also, merging of somehow similar components, or propagation in the orbital elements could be involved.

References

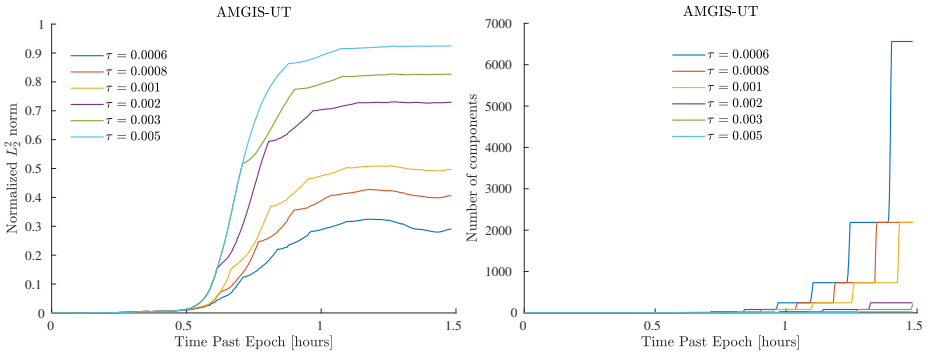
1. Adali, T., Haykin, S.: Adaptive Signal Processing: Next Generation Solutions. Adaptive and Cognitive Dynamic Systems: Signal Processing, Learning, Communications and Control. Wiley (2010)
2. Arndt, C.: Information measures. Springer (2004)
3. Bar-Shalom, Y., Li, X.R., Kirubarajan, T.: Estimation with Applications to Tracking and Navigation: Theory Algorithms and Software. John Wiley & Sons (2001)
4. Bertsekas, D.P.: Dynamic Programming and Optimal Control. 2nd edn. Athena Scientific (2000)
5. Dell'Elce, L., Arnst, M., Kerschen, G.: Probabilistic assessment of the lifetime of low-earth-orbit spacecraft: Uncertainty characterization. Journal of Guidance, Control, and Dynamics **38**(5) (2015) 900–912
6. DeMars, K.J., Bishop, R.H., Jah, M.K.: Entropy-based approach for uncertainty propagation of nonlinear dynamical systems. Journal of Guidance, Control, and Dynamics **36**(4) (2013) 1047–1057



(a) (b)

Fig. 6: LEO: Performance analysis of the AMGIS–LinCov algorithm with different thresholds.

7. Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society, B **39**(1) (1977) 1–38
8. Duník, J., Straka, O., Mallick, M., Blasch, E.: Survey of nonlinearity and non-gaussianity measures for state estimation. In: 2016 19th International Conference on Information Fusion (FUSION). (July 2016) 1845–1852
9. Gelb, A., Warren, R.: Direct statistical analysis of nonlinear systems: Cadet. AIAA Journal **11**(5) (1973) 689–694
10. Giza, D., Singla, P., Jah, M.: An approach for nonlinear uncertainty propagation: Application to orbital mechanics. In: AIAA Guidance, Navigation, and Control Conference. (2012)
11. Havlík, J., Straka, O.: Measures of nonlinearity ans non-gaussianity in orbital uncertainty propagation. In: 2019 22th International Conference on Information Fusion (FUSION). (2019)
12. Isermann, R.: Fault-Diagnosis Systems: An Introduction from Fault Detection to Fault Tolerance. Springer Berlin Heidelberg (2005)
13. Jazwinski, A.: Stochastic Processes and Filtering Theory. Dover Books on Electrical Engineering Series. Dover Publications (2007)
14. Julier, S., Uhlmann, J., Durrant-Whyte, H.F.: A new method for the nonlinear transformation of means and covariances in filters and estimators. IEEE Transactions on Automatic Control **45**(3) (March 2000) 477–482
15. Krejčí, J.: Uncertainty propagation for tracking of moving objects. Bachelor thesis, University of West Bohemia (2018)
16. Li, X.R.: Measure of nonlinearity for stochastic systems. In: 2012 15th International Conference on Information Fusion. (July 2012) 1073–1080
17. Markley, F.L., Carpenter, J.R.: Generalized linear covariance analysis. The Journal of the Astronautical Sciences **57**(1) (Jan 2009) 233–260
18. Park, I.: Dynamical realism and uncertainty propagation. Aerospace Engineering Sciences Graduate Theses & Dissertations. **130** (2015)



(a) (b)

Fig. 7: LEO: Performance analysis of the AMGIS–UT algorithm with different thresholds.

19. Sagnieres, L., Sharf, I.: Uncertainty characterization of atmospheric density models for orbit prediction of space debris. In: 7th European Conference on Space Debris. Volume 7. (2017) N/A
20. Shannon, C.E.: A mathematical theory of communication. Bell System Technical Journal **27**(3) (1948) 379–423
21. Sorenson, H., Alspach, D.: Recursive Bayesian estimation using Gaussian sums. Automatica **7**(4) (1971) 465 – 479
22. Sun, Z.J., Luo, Y.Z., di Lizia, P., Zazzera, F.B.: Nonlinear orbital uncertainty propagation with differential algebra and Gaussian mixture model. Science China Physics, Mechanics & Astronomy **62**(3) (Nov 2018) 34511
23. Titov, E., Burt, J., Josyula, E.: Satellite drag uncertainties associated with atmospheric parameter variations at low earth orbits. Journal of Spacecraft and Rockets **51**(3) (2014) 884–892
24. Wan, E.A., Van Der Merwe, R.: The unscented kalman filter for nonlinear estimation. In: Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No.00EX373). (Oct 2000) 153–158
25. Xiong, F., Chen, S., Xiong, Y.: Dynamic system uncertainty propagation using polynomial chaos. Chinese Journal of Aeronautics **27**(5) (2014) 1156 – 1170
26. Yang, Z., Luo, Y.Z., Zhang, J.: Nonlinear semi-analytical uncertainty propagation of trajectory under impulsive maneuvers. Astrodynamics **3**(1) (Mar 2019) 61–77
27. zhong Luo, Y., Yang, Z.: A review of uncertainty propagation in orbital mechanics. Progress in Aerospace Sciences **89** (2017) 23 – 39