Spacecraft Uncertainty Propagation using

Gaussian Mixture Models and Polynomial Chaos Expansions

Vivek Vittaldev\(^1\) and Ryan P. Russell\(^2\)

\textit{The University of Texas at Austin, Austin, Texas, 78712}

Richard Linares\(^3\)

\textit{University at Minnesota, Minneapolis, Minnesota, 55403}

Polynomial Chaos Expansion (PCE) and Gaussian Mixture Models (GMMs) are combined in a hybrid fashion to propagate state uncertainty for spacecraft with initial Gaussian errors. PCE models uncertainty by performing an expansion using orthogonal polynomials (OPs). The accuracy of PCE for a given problem can be improved by increasing the order of the OP expansion. The number of terms in the OP expansion increases factorially with dimensionality of the problem, thereby reducing the effectiveness of the PCE approach for problems of moderately high dimensionality. This paper shows a combination of GMM and PCE, GMM-PC as an alternative form of the multi-element PCE. GMM-PC reduces the overall order required to reach a desired accuracy. The initial distribution is converted to a GMM, and PCE is used to propagate the state uncertainty represented by each of the elements through the nonlinear dynamics. Splitting the initial distribution into a GMM reduces the size of the covariance associated with each new element thereby reducing the domain of approximation and allowing for lower order polynomials to be used. Several spacecraft uncertainty propagation examples are shown using GMM-PC. The resulting distributions are shown to

\(^1\) Graduate Student, Department of Aerospace Engineering and Engineering Mechanics, v.vittaldev@utexas.edu, Member AIAA.
\(^2\) Associate Professor, Department of Aerospace Engineering and Engineering Mechanics, ryan.russell@utexas.edu, Associate Fellow AIAA.
\(^3\) Assistant Professor, Department of Aerospace Engineering & Mechanics, rлинаres@umn.edu, Member AIAA.
efficiently capture the full shape of the true non-Gaussian distribution.

I. Introduction

Space situational awareness (SSA) is a key area of research for both the civil and military space communities [1]. SSA involves understanding the space environment and its influence on active space assets. The increased population in space of both controllable and uncontrollable Space Objects (SOs) creates difficult challenges for SSA [2]. Many of the components of SSA such as conjunction assessment, orbit determination (OD), and track correlation and prediction require an accurate knowledge of the uncertainty in the states of SOs as a function of time. Due to the nonlinear dynamics of SOs and the latency in observations due to the limitation of the surveillance network, non-Gaussian uncertainty quantification is a key challenge for SSA [3].

Common estimation techniques used for OD such as the Extended Kalman Filter (EKF) or the Batch Least Squares (BLS) work under the assumption that the state uncertainty has a Gaussian distribution [4]. A Gaussian initial distribution for the state uncertainty of SOs at epoch is a frequent assumption because conventional OD techniques only output a state estimate and covariance of the estimated error. OD approaches, therefore, assume Gaussian distributions due to either the simplicity or the validity of this assumption. However, as the state of the SO is propagated in time, the assumption of a Gaussian distribution is no longer valid due to the nonlinearity of the dynamics of orbital motion. For many SSA applications it is desired to know the uncertainty of the SO at a future epoch. Therefore, uncertainty propagation techniques are essential.

A common, computationally intensive method for propagating uncertainty is Monte Carlo (MC) simulations [5, 6]. To achieve statistically valid measures of uncertainty, MC approaches may require on the order of tens or hundreds of millions of propagations for the orbital uncertainty problem. Parallelizing the computations on multiprocessor CPUs or on Graphics Processing Units (GPUs) reduces the runtime of the simulations significantly [7–9] at the cost of increasing the difficulty of implementation [10]. Reducing the number of sample points required for a result with satisfactory confidence bounds is sometimes possible through variance reduction techniques [11]. Although
the computational cost can be prohibitive for most applications due to the slow convergence, the
generality and robustness of MC techniques make them ideal benchmarks to compare other methods.

A spectrum of techniques exists that propagate the state and uncertainty of an initially Gaussian
distribution through a non-linear function, such as orbit propagation [12]. Reduction in computation
cost often comes with a sacrifice in the accuracy of the final probability density function (PDF).
Considering the runtime vs. accuracy spectrum, linear propagation (i.e. the propagation step of an
Extended Kalman Filter (EKF) [4]) lies on the low computation and accuracy extremity whereas
MC lies on the high computation and accuracy extremity. Two techniques that lie in between the
first order Taylor series and MC are Gaussian Mixture Models (GMMs) and Polynomial Chaos
expansions (PCEs).

GMMs can approximate any PDF using a weighted sum of Gaussian distributions and this ap-
proximation converges to any PDF as the number of elements is increased (in a $L_1$-norm sense) [13].
To effectively use GMMs for uncertainty propagation, the initial Gaussian element has to be split
into multiple GMM elements. The splitting of an initial Gaussian distribution into a GMM effec-
tively accomplishes a local piecewise linearization of the nonlinear function. The GMM has spa-
tially distributed means and each element has a smaller uncertainty (i.e. differential entropy). Each
Gaussian element can be propagated through the nonlinear function using State Transition Tensors
(STTs) [14], Fourier-Hermite series [15, 16], sigma-point based methods [17–19], quadrature [20–22],
or cubature [23–26]. The Gaussian assumption remains valid longer for each of the smaller elements
and therefore, the weighted sum of the Gaussian elements after propagation approximates the re-
sulting non-Gaussian PDF. GMMs have been successfully used in orbital mechanics for uncertainty
propagation [27, 28], orbit determination [29, 30], and conjunction assessment [31–33].

PCE [34] quantifies non-Gaussian uncertainty by using a linear combination of orthogonal
polynomials (OP) as a surrogate model for the underlying nonlinear function. The Wiener-Askey
scheme [35] generalizes PCE to other initial distributions and is known as generalized PCE (gPCE).
It is also possible to compute orthogonal polynomials for arbitrary PDFs that are not part of the
Wiener-Askey scheme using arbitrary PCE (aPCE) [36]. After selecting the appropriate OPs, the
PCE approximation is determined by the coefficients of the linear combination of these polynomials.
Once the coefficients are computed, sampling from the PCE generally has a lower computational cost than MC. PCE has been used in many fields for uncertainty quantification of computationally intensive models [37–40]. In orbital mechanics, PCE has been previously used for uncertainty propagation [41] and conjunction assessment [42, 43].

The motivation for combining PCE and GMMs (GMM-PC) for uncertainty propagation arises from a comparison with the finite element method (FEM). In this analogy the initial uncertainty is considered to be an object of interest and the nonlinear function is an applied load. In FEM, a mesh grid is generated over the object to discretize it into smaller and simpler geometries. The exact solution is obtained at the nodes, but between the nodes inside each element some polynomial interpolant is used to approximately obtain the functional form of displacements. Increasing the number of elements by reducing the size of each element is known as $h$-refinement and increasing the order of the interpolant is $p$-refinement. Selectively modifying both the size and number of elements, and the order of the polynomial interpolant is $hp$-refinement [44, 45]. Splitting the initial Gaussian distribution into Gaussian distributions with a smaller differential entropy is analogous to $h$-refinement. Using a higher order accurate method of propagating the uncertainty of a single Gaussian element is $p$-refinement. A proof of concept is presented here to demonstrate that a $hp$-refinement method of using PCE for the propagation of a GMM is applicable for orbit uncertainty propagation by increasing accuracy while having a lower computational cost than an MC simulation. GMM-PC is an alternative, but new, form of the Multi-Element generalized Polynomial Chaos (ME-gPC) method [46]. The ME-gPC uses segmented Gaussian distribution for the $h$-refinement, while the GMM-PC uses weighted Gaussian kernels. The GMM-PC trades optimality for ease of implementation, when compared to the tradition ME-gPC. The similarities and differences between the two methods are outlined in Section IV.

The evolution of an initial Gaussian SO state uncertainty distribution into a non-Gaussian distribution is the driving force behind the nonlinear uncertainty propagation techniques for the orbital problem. Changing the coordinate set in which the SO state is expressed can have an effect on the length of the flight time the Gaussian assumption remains approximately valid. Orbit element sets such as Keplerian elements, Equinoctial elements (EE), and Modified Equinoctial elements
(MEE) absorb some of the nonlinearity of orbital motion and therefore, the uncertainty distribution expressed in these states can be assumed to be Gaussian for longer times of flight [47–50]. The Gaussian assumption in all coordinate sets degrades with increasing initial state uncertainty, fidelity of perturbations, degree of mismodeling, and time of flight [48, 51].

The organization of the paper is as follows: first a summarized background is reviewed for PCE and GMMs. Next, the combined GMM-PC is presented. Then the performance of GMM-PC is compared to PCE for three highly nonlinear orbit uncertainty propagation cases. Finally, conclusions are provided from these results.

II. Polynomial Chaos Expansion

For a PCE, orthogonal polynomials provide a basis for the dependence of the quantity of interest on the random inputs. The uncertainty can then be characterized based on this dependence. Let \( u(\xi, t) \) be a second order random process which is a function of time and a random variable \( \xi \in \Omega \) where \( u(\xi, t) : (\Omega \times \mathbb{R}^+) \rightarrow \mathbb{R}^1 \), then \( u(\xi, t) \) can be expressed by a gPCE as:

\[
\begin{align*}
  u(\xi, t) &= \sum_{i=0}^{\infty} c_i(t) \Psi_i(\xi) \quad (1a) \\
  \Psi_i(\xi) &= \sum_{j=0}^{i} a_{i,j} \xi^j \quad (1b)
\end{align*}
\]

Where \( c_i(t) \in \mathbb{R} \) are the \( i^\text{th} \) coefficient associated with the \( i^\text{th} \) orthogonal polynomials \( \Psi_i \). The orthogonal polynomials \( \Psi_i \) and their coefficients \( a_{i,j} \) are defined as being orthogonal w.r.t. the following inner product:

\[
\int_{\Omega} \Psi_m(\xi)\Psi_n(\xi)p(\xi)d\xi = 0, \ m \neq n \quad (2)
\]

where \( p(\xi) \) satisfies the properties of a PDF over the domain \( \Omega \) i.e. \( \xi \in \Omega \). In the Weiner-Askey [35] scheme the orthogonal polynomial type is determined by \( p(\xi) \) for classical PDFs. In general, the aPCE scheme can account for cases where \( p(\xi) \) are not classical but arbitrary PDFs by determining the orthogonal polynomials using the raw moments of the distribution [36]. In this work the initial distribution is assumed to be Gaussian and therefore, probabilists’ Hermite polynomials are chosen according to the Wiener-Askey [35] scheme. A brief description of Hermite polynomials is provided in Appendix A.
The orbit uncertainty propagation problem is a multivariate problem and therefore, requires orthogonal multivariate polynomials. A multivariate Gaussian distribution can be converted to a product of IID standard normal distributions. Multivariate polynomials are generated from a set of independent random variables and created using the multi-index notation. The multi-index vectors, \( \alpha_i \in \mathbb{R}^n \), contain the orders of the \( d \) univariate polynomials that are combined in a tensor product to produce multivariate polynomials. The number of monomials, \( L \), of order \( \ell \) or less in \( d \) variable is computed by

\[
L = \frac{(d + \ell)!}{d! \ell!}
\]  

(3)

Using multi-index notation, the multivariate polynomial is the product:

\[
\Psi_{\alpha i}(\xi) = \prod_{k=1}^{d} \psi_{\alpha_i(k)}(\xi_k)
\]  

(4)

If the output is a vector function of dimension \( m \), \( m \times L \) coefficients \( c_i(t) \) have to be solved for.

In practice, Eq.(1a) is approximated by truncating the infinite series for a fixed number of polynomial functions. The approximate PCE for multivariate polynomials is:

\[
\bar{u}_L(\xi, t) = \sum_{i=0}^{L} c_i(t) \Psi_{\alpha_i}(\xi)
\]  

(5)

It has been shown that for analytic functions the truncated approximate PCE, \( \bar{u}(\xi, t) \), given above converges exponentially to the second order process \( u(\xi, t) \) in the \( L_2 \)-norm sense [35], i.e.,

\[
\mathbb{E} \left[ (\bar{u}_L(\xi, t) - u(\xi, t))^2 \right] \to 0 \text{ as } L \to \infty.
\]

The coefficients \( c_i(t) \) determine the response surface of the surrogate model consisting of the Hermite polynomials. Two methods of finding these coefficients are the intrusive and the non-intrusive methods. The intrusive method requires knowledge of the full nonlinear function that determines the evolution of the random vector of inputs. The truncated polynomial expansions are introduced into the model equations and solved using a Galerkin projection of the equations on the polynomial space. A system of equations is then solved for \( c_i(t) \). The intrusive method cannot be used with black-box dynamics because existing codes have to be rewritten to form additional differential equations for \( c_i(t) \) and therefore, is not considered in this work. The intrusive method is analogous to integrating the analytical Jacobian of the dynamics to compute the first-order STM.
The non-intrusive method does not require any knowledge of the propagation function. Given that the system can be solved for a sample initial condition, the projection property (Galerkin Projection) for approximating the coefficients in Eq.(5) can be used:

\[ c_i(t) = \int_\Omega u(\xi, t) \Psi_{\alpha_i}(\xi) p(\xi) d\xi \]  

where \( p(\xi) \) is the PDF of \( \xi \in \Omega \). The coefficients can be computed by a quadrature numerical approximation of Eq.(6), MC sampling, or with Least Squares (LS) regression. The results in this paper use the LS regression for computing the coefficients. Both LS and MC methods for numerically approximating Eq.(6) are based on random sampling. The quadrature methods on the other hand, require deterministic precomputed nodes and weights. All non-intrusive methods generate sample points from \( \xi \sim \mathcal{N}(0_d, I_d \times d) \). A square-root factor of the covariance matrix linearly converts \( \xi \) from an IID variable to the initial multivariate Gaussian distribution.

Equation 7 is solved by rewriting into the traditional LS form using the following matrix containing the \( L \) multivariate polynomials evaluated at the \( N \) nodes:

\[ \Psi = \begin{bmatrix} \Psi_{\alpha_1}(\xi_1) & \cdots & \Psi_{\alpha_L}(\xi_1) \\ \vdots & \ddots & \vdots \\ \Psi_{\alpha_1}(\xi_N) & \cdots & \Psi_{\alpha_L}(\xi_N) \end{bmatrix} \]  

Arranging the coefficients and function evaluations into vectors results in the following linear system:

\[ \left( \Psi^T \Psi \right) \begin{bmatrix} c_1(t) \\ \vdots \\ c_L(t) \end{bmatrix} = \Psi^T \begin{bmatrix} u(\xi_1, t) \\ \vdots \\ u(\xi_N, t) \end{bmatrix} \]  

The LS method suffers from the curse of dimensionality where the combination of increasing problem dimension and order of the polynomial scale the number of required evaluations in a factorial manner.

For an \( n \)-dimensional input state with univariate polynomials of maximum order \( \ell \), the number of
terms $L$ in the multivariate polynomial is computed using Equation 3. The number of terms in a six-dimensional multivariate polynomial as a function of $\ell$ is shown in Figure 1.

The LS regression method requires approximately $2L$ sample points to solve for the coefficients [52] as a general recommendation, which is not an absolute and can be problem dependent. The exact number of sample points is unknown and dependent on the definition of convergence of the PCE solution for a given problem. However, if many coefficients have a small value and therefore, compressive sensing techniques can generate a sparse representation of $c(t) = [c_1(t), \ldots, c_L(t)]^T$ so that $N < L$ function evaluations suffice. To generate a true sparse representation of $c(t)$, the $L_0$ norm of $c(t)$ has to be minimized subject to the $L_2$ norm conditions between the function evaluations and the PC solution. However, minimizing the $L_1$ norm instead of the $L_0$ norm of $c(t)$ converts the problem into a convex optimization problem that can be solved with common solvers [53, 54].

A full tensor product for the multivariate quadrature requires $N = \ell^d$ function evaluations and therefore, also suffers from the curse of dimensionality. Sparse grids compute quadrature nodes using sparse tensor product and therefore, reduce the number of function evaluations to $N < \ell^d$ for high dimensions. The number of function evaluations required for using a full tensor product quadrature rule and two different sparse grids, Smolyak [55] and Genz-Keister [56], is shown in Figure 1. LS regression is used in the current study, which uses $N = 2L$ function evaluations, where $L$ is computed using Equation (3).

![Graph showing the comparison of full grid, Smolyak, Konrad-Patterson, and Smolyak methods with LS regression for terms required for multivariate polynomials and the number of quadrature points for a 6 dimensional input state. The LS regression requires $2L$ points.](image-url)
III. Gaussian Mixture Models

A GMM is a weighted sum of Gaussian probability distribution functions.

\[ p(x) = \sum_{i=1}^{N} \alpha_i p_g(x; \mu_i, P_i) \quad \sum_{i=1}^{N} \alpha_i = 1 \quad 0 < \alpha_i \leq 1 \]  

(10)

\( N \) is the number of Gaussian probability distribution functions, and \( \alpha_i \) is a weight. For uncertainty propagation, the initial Gaussian distribution is split into a GMM and each element is propagated through the nonlinear function. Standard Gaussian propagation techniques such as STTs [14] or sigma-point methods [17, 18] are commonly used to approximate the Gaussian elements post-propagation. Higher order STTs [14] and sigma-point methods [19] can better approximate a non-Gaussian distribution by computing more moments. Although each element remains Gaussian, the weighted sum forms a non-Gaussian approximation of the true distribution. Modifications of the procedure exist, and are commonly applied when the nonlinear function is the solution of an ODE: the weights can be updated post-propagation [57] or the elements can be further split into more elements or merged mid-propagation [27]. However, these modifications are not considered in the current study.

Instead of forming a GMM approximation of the initial multivariate Gaussian distribution, a univariate GMM library of the standard normal distribution is formed [27, 29, 32, 58]. The univariate splitting library has to be computed only once and is stored in the form of a lookup table. A library where all the standard deviations in the split are the same (homoscedastic) [32, 59]. With increasing \( N \), \( \sigma \) decreases and therefore, the differential entropy of each element decreases.

To apply the univariate splitting library to a multivariate Gaussian distribution with mean \( \mu \) and covariance matrix \( P \), the univariate splitting library is applied along a column of the square-root \( S \) of the covariance matrix:

\[ P = SS^T \]  

(11)

For an \( d \)-dimensional state, the covariance matrix of each element is:

\[ P^i = [s_1 \ldots \sigma s_k \ldots s_d][s_1 \ldots \sigma s_k \ldots s_d]^T \]  

(12)

where \( s_k \) is the desired column of \( S \). The means of the multivariate GMM are:

\[ \mu^i = \mu + \mu^i s_k \]  

(13)
It is possible to apply the univariate splitting direction along any desired direction by generating a square-root matrix with one column parallel to the input direction \([33, 60, 61]\). For extremely non-linear problems, splitting along a single direction may not account for the entire non-linearity of the problem. Therefore, splitting the initial multivariate distribution in multiple directions is required in order to better approximate the non-Gaussian behavior post-propagation \([33, 59]\). In such cases the splitting library can be applied recursively as a tensor product to split along multiple directions.

IV. Gaussian Mixture Model Polynomial Chaos

The Gaussian Mixture Model Polynomial Chaos (GMM-PC) approach is used to approximate the distribution of the following nonlinear function:

$$y = f(x), \quad x \sim \mathcal{N}(\mu, P) \quad (14)$$

where \(x \in \mathbb{R}^d\) and \(y \in \mathbb{R}^{d_2}\). For a pure state propagation problem with only state uncertainty, the nonlinear problem is:

$$x_{k+1} = f_k(x_k, t_k, t_{k+1}), \quad x_k \sim \mathcal{N}(\mu_k, P_k) \quad (15)$$

where \(x_k\) denotes the state of the system at time index \(k\) or time \(t_k\) and \(f_k(\cdot)\) is a nonlinear function. This system is assumed to have Gaussian initial errors with mean and covariance \(\mu_k\) and \(P_k\) respectively. A PCE forms a surrogate model, which is sampled to approximate the distribution of \(x_k\). From Eq. (3) the number of terms \(L\) that are required to achieve a given order \(\ell\) for the PCE approach grows factorially with dimension \(d\), therefore it becomes increasingly difficult to increase the order of accuracy for higher dimensional problems. This provides a strong incentive to use the lowest order PCE possible, subject to an acceptable accuracy level. The GMM-PC reduces the order of the PCE by splitting the initial Gaussian distribution into a GMM. Each weighted Gaussian element is further propagated through the nonlinear function using PCE. The non-Gaussian behavior of the PDF post-propagation depends primarily on the nonlinearity of the function over the domain of the initial uncertainty. Splitting the initial Gaussian distribution into smaller distributions, decreases the size of the initial uncertainty of each element. Splitting reduces the nonlinearity of
the function by reducing the domain of each element, requiring lower order polynomials to achieve desired accuracy. However, increasing the number of splits in the initial distribution increases the required function evaluations linearly, whereas a reduction in order decreases the number of function evaluations factorially.

The Multi-Element generalized Polynomial Chaos (ME-gPC) method [46] also decomposes the PDF into smaller subdomains by partitioning the support of the initial PDF. For a Gaussian initial distribution, the subdomains of the ME-gPC method have a non-classical distribution seen in Figure 2. Therefore, Gram-Schmidt orthogonalization is required to form the polynomials for this non-classical distribution, this is in effect an aPCE technique over the subdomains [36]. Because each element in the GMM-PC technique is itself a Gaussian distribution, standard Hermite polynomials are used which greatly reduces the complexity of the code over the ME-gPC because analytical recursive relations exist for Hermite polynomials. In the ME-gPC method, the initial PDF is exactly represented by the elements, but the numerically computed polynomials are approximations. However, in the GMM-PC method, the initial PDF is approximated by a GMM but the form of the orthogonal polynomials (coefficients $a_i$ from Equation 1b) are available in closed form. It should be noted that the actual PCE coefficients, which determine the linear combination of the orthogonal polynomials (coefficients $c_i$ from Equation 1a), are computed using regression. The classical ME-gPC method handles any discontinuities as a function of random inputs better than the GMM-PC. The GMM-PC uses Hermite polynomials and therefore, assumes a smooth dependence on the input variables over the infinite domain.

Some GMM methods adaptively generate new elements by splitting existing elements in such a manner that the local probability density is close to Gaussian [27] when the function is an ODE. In the current work, functions are considered to be discrete black box functions. Therefore, further splitting and merging mid function evaluation is not treated. For the orbit propagation problem, one function evaluation is the integration of a sample trajectory for the entire time of flight. The failure of the propagated Gaussian components to describe the underlying non-Gaussian distribution is mitigated by converting each element to a PCE. The PCEs employed are a global expansion, i.e., they have infinite domain with no guaranteed exponential decay in the response to the random
inputs. Hence, there are two or more PCEs that can overlap by providing information on the system response to the same random inputs. This overlap causes the GMM-PC technique to be suboptimal compared to traditional ME-gPC, which uses segmented PDFs. However, the benefit is the ease of implementation because an existing PCE implementation that uses Hermite polynomials is easily extended to GMM-PC without having to recompute orthogonal polynomials for arbitrary PDFs.

To construct an $N$-element GMM-PC for the nonlinear problem from Eq. (14), the initial multivariate Gaussian distribution is first approximated by a GMM.

$$p_g(x; \mu, P) \approx \sum_{i=1}^{N} \alpha_i p_g(x_i; \mu_i, P_i)$$

Note that the subscript $i$ denotes the GMM element number. Samples can be drawn from the GMM in Eq. (16) and PCE is used to propagate the uncertainty of each element through the nonlinear function. For each element, the square-root matrix $S_i$ is a transformation from a reference frame where the variables are IID.

$$x_i = S_i \xi_i + \mu_i, \quad P_i = S_i S_i^T, \quad \xi_i \sim \mathcal{N}(0_d, I_{d \times d})$$

where $i$ represents the GMM element and $d$ is the dimension of the input parameters $x$. A sample from the GMM-PC is generated randomly by the $i^{th}$ element, where $i$ is determined by a random
variable \( X \) with a uniform distribution, \( X \sim U(0,1) \). Given a random sample \( X \), \( i \) is determined such that:

\[
\sum_{k=0}^{i-1} \alpha_k < X \leq \sum_{k=1}^{i} \alpha_k, \quad \alpha_0 = 0
\]

To sample from the GMM-PC, a set of samples \( X \) are drawn from \( U(0,1) \) and a set of \( i \) are found from these samples using Eq. (18). Then the number of samples that are drawn from each element in the GMM is determined by the set of \( i \). Finally, with a desired univariate order \( \ell_i \) for each element, the surrogate model given by the GMM-PC approximation for Eq. (14) is:

\[
y_i = C_i \Psi^*_i(\xi_i)
\]

where

\[
\Psi^*_i = \left[ \psi_{\alpha_1}(\xi_i) \ldots \psi_{\alpha_{L_i}}(\xi_i) \right]^T
\]

\[
C_i \in \mathbb{R}^{d \times L_i}
\]

The matrix of coefficients \( C_i \) are computed for each element using Least Squares.

A univariate example is now considered where a 13th-order normalized Hermite polynomial expansion is used as the nonlinear function \( y = f(x) \). The polynomials are computed using Equations (28) and (29).

\[
y = \sum_{k=0}^{13} c_k \psi_k(x), \quad x \sim \mathcal{N}(0,1)
\]

\[
c_k = (-1)^k \times e^{-\beta \times k}
\]

The assigned value of \( \beta \) controls the influence of the higher order coefficients and a larger value indicates greater nonlinearity. The result of the uncertainty propagation through Eq. (21) is approximated by PCE (1 element GMM-PC) and GMM-PC of orders 3 - 10. The two-sample univariate Cramér-von Mises parameter is used to compare the performance of the PCE and GMM-PC (3, 5, and 7 elements) with respect to an MC simulation [62]:

\[
CvM = \frac{N_1N_2}{N_1 + N_2} \int_{-\infty}^{\infty} [F_{N_1}(x) - G_{N_2}(x)]^2 dH_{N_1+N_2}(x)
\]
where $F_{N_1}(x)$ is the empirical distribution of the surrogate model with $N_1$ samples, $G_{N_2}(x)$ is the empirical distribution of the MC simulation with $N_2$ samples, and $H_{N_1+N_2}(x)$ is the empirical distribution function of two samples together. A lower value of $CvM$ implies higher accuracy. The $CvM$ values for $\beta = 1$ and $\beta = 0.1$ are computed for increasing number of elements and PCE order for GMM-PC in Figures 3a and 3b, respectively. When $\beta = 1$, a one element GMM-PC (standard PCE) is able to accurately capture the uncertainty of the 13th-order PCE for orders greater than 3 and there is no benefit of using more elements. The GMM-PC with 3 elements does not provide accurate results ($CvM < 0.3$) for any order because of the difference between the original Gaussian distribution and the GMM approximation. The behavior of the $CvM$ values resembles a random walk in Figure 3a because the test cannot discriminate between two distributions when they are virtually identical. When $\beta = 0.1$, the benefit of increasing the number of elements is clearly seen and a GMM-PC with $N > 1$ outperforms the PCE for $\ell < 8$. The nonlinearity of the function is large enough that the difference between the GMM distribution and the original Gaussian distribution is not noticeable after propagation.

**Fig. 3 Performance of approximating a 13th-order Hermite PCE using lower order (3 $\leq \ell \leq 10$) GMM-PC using 1,000,000 samples**

Approximating the uncertainty propagation through an arbitrary function using PCE is analogous to approximating a higher order PCE with a lower order PCE. In case the function is not highly nonlinear, a lower order PCE is able to approximate the resulting uncertainty with high
accuracy and there is no benefit of using GMM-PC. However, when the nonlinearity is strong, the benefit of GMM-PC is clearly seen. The computational benefit is greater as the dimension of the problem increases because the rate of growth of the number of multivariate polynomials $L$ is factorial with respect to the dimension. The computational load of GMM-PC can be further reduced if the PCE order of each element in the GMM-PC is varied proportionally to the element weight. If the quadrature or cubature methods are used to compute the PCE coefficients in GMM-PC, the black box function has to be run for the deterministic nodes for each element. In this case, the computational load increases linearly with the number of elements if each GMM-PC element uses the same order of polynomials $\ell$.

Another strategy using Weighted Least Squares (WLS) [63, 64] is also tested in the Results, Section V E, when using the LS collocation method for GMM-PC. The computational budget for the number of function evaluations is first fixed. For the orbit uncertainty propagation problem, a function evaluation is the integration of the trajectory for the desired time of flight. Using the same function evaluations, the coefficients for each GMM-PC are found using WLS. For each element, the initial conditions for the function calls are converted to the uncorrelated reference frame of the element’s covariance matrix using the square-root.

$$\xi = S^{-1}_i(x - \mu_i), \quad P_i = S_iS_i^T$$

The coefficients are then computed using the standard normal equations for WLS with an $n \times n$ diagonal weight matrix $W$ [63, 64]:

$$W(j,j) = (2\pi)^{-\frac{d}{2}}e^{-\frac{1}{2}\xi_j^T\xi_j}$$

The diagonal elements are the PDF value of normalized $d$-dimensional input points with respect to a Gaussian distribution with $0_d$ mean and a covariance of identity $I_{d \times d}$. To keep the total number of function evaluations fixed, the order of the PCE for all the elements is no longer the same. The multivariate order of each element $L_i$ is chosen such that $L_i = \text{floor}(M/2)$ and $M$ is the number of points $\xi_i$ that are within a $3\sigma$ bound for the current GMM element.

The final issue relates to optimally constructing the GMM approximation of the initial Gaussian distribution prior to converting each element into a PCE. In this study, univariate splitting libraries
with an odd number of elements are used so that the locations of the means all lie along a single line in the $d$-dimensional space [32]. An initial, optionally lower order, PCE is fit to find the direction with the highest nonlinearity i.e., the direction in which the value of the PCE coefficients decays the slowest. If the absolute values of the coefficients decay rapidly in all directions, there is no reason to switch to a GMM-PC; otherwise. If the GMM-PC is required due to the nonlinearity, the univariate splitting library is applied along the identified direction. The WLS technique can be used to possibly reuse the function evaluations for at least the central element, which is the most similar to the original Gaussian distribution. The WLS approach can also be an effective way of choosing appropriate orders and component numbers. Using a set of simulations of the system (i.e., a computational budget), the WLS can be run for different PCE orders and component numbers until convergence is achieved in the mean and variance estimates. A priori knowledge about the problem can also be used to identify the directions with highest nonlinearity (e.g. the velocity uncertainty for orbit propagation).

V. Results

Four orbital dynamics test cases that benefit from using GMM-PC are presented in this section. The first two, and the fourth cases are subject to two-body dynamics: 1) an eccentric Medium Earth Orbit (MEO), 2) a Molniya orbit, and 4) a circular LEO. Two-body dynamics are used to make the results easily reproducible. The third test case is an object in a Geosynchronous Transfer Orbit (GTO) under the influence of additional perturbations. Cartesian coordinates in the ECI frame are used to express the uncertainties for the first three cases and EE are used for the last case. The univariate splitting library used to generate the initial GMMs is from Vittaldev et. al. [32].

The univariate polynomials that constitute the multivariate polynomial are tested for orders 3 to 10 (i.e., $\ell$ in Eq. (3) is changed from 3 to 10). For the GMM-PC with $N$ elements, the total number of coefficients is computed using Equation (3)

$$L = N \times \frac{(d + \ell)!}{d!\ell!}$$

(26)

The PCE coefficients are computed using LS where the number of function evaluations is twice the number of multivariate coefficients $L$ (i.e., total computational load is $2L$). The univariate splitting
libraries used are easily accessible 1.

The final subsection uses the WLS GMM-PC technique to reuse function evaluations. For the orbit uncertainty propagation examples shown in this section, a function evaluation is the integration of the trajectory of a sample from epoch for the entire flight time.

A. Medium Earth Orbit

The first test case is a MEO with two-body dynamics propagated for 3 days. The initial state and uncertainty at epoch are found in Table 1. The highest nonlinearity and therefore, the non-convergence is found along the initial velocity direction for the MEO case.

<table>
<thead>
<tr>
<th>a</th>
<th>e</th>
<th>i</th>
<th>( \Omega )</th>
<th>( \omega )</th>
<th>( \nu )</th>
<th>( \sigma_{rR} )</th>
<th>( \sigma_{rI} )</th>
<th>( \sigma_{rC} )</th>
<th>( \sigma_{vR} )</th>
<th>( \sigma_{vI} )</th>
<th>( \sigma_{vC} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>24,475 km</td>
<td>0.5</td>
<td>25°</td>
<td>0°</td>
<td>0°</td>
<td>1 km</td>
<td>1 m</td>
<td>1 m</td>
<td>0.1 m/s</td>
<td>0.1 m/s</td>
<td>0.1 m/s</td>
<td></td>
</tr>
</tbody>
</table>

Table 1 Initial osculating orbit elements and uncertainty expressed in the RIC coordinate frame for a MEO object

The coefficient values corresponding to the order of univariate polynomials within the multivariate polynomials from Eq.(4) for a one element GMM-PC are plotted in Figure 4. The square-root of the initial covariance is constructed such that \( \xi_1 \) is along the velocity vector \( \mathbf{v}_v \). Figure 4a shows that this case is highly non-linear along the velocity direction because increasing the order of the PCE does not result in a fast reduction in the values of the coefficients. The expansion along the other directions converges, as shown by the reducing coefficient magnitude in Figure 4b. Applying a univariate split along \( \xi_1 \) reduces the size of the uncertainty and therefore, polynomials of a lower order can be used to achieve convergence.

The initial Gaussian distribution is split into a GMM with up to 9 elements along the \( \mathbf{v}_v \) direction. The sum of \( L \) of all the elements is used as a proxy for the computation cost. The GMM-PC is propagated for the desired flight time of 3 days and is sampled. The \( CuM \) values for the samples in the Radial-Intrack-Crosstrack (RIC) frame with respect to a MC simulation is

1 http://russell.ae.utexas.edu/code/GMMsplittingLibrary.txt
Fig. 4 Coefficient values for the corresponding univariate polynomial orders for the directions of the IID state for a one element GMM-PC used to propagate uncertainty of an object in MEO. The trend of the maximum values of the coefficients as the order increases is indicated by the arrow.

(a). $\xi_1$: Velocity direction  
(b). $\xi_2 - \xi_6$: Other directions

Fig. 5 Two-sample univariate Cramer-von-Mises metric in the RIC frame for GMM-PC, split along velocity, with respect to an MC simulation of 1,000,000 samples of an object in MEO. $L$ is the total number of coefficients required per direction for the multivariate polynomial and is analogous to the compute cost $(2L)$.

(a). PCE (1 element GMM-PC)  
(b). 9 element GMM-PC

shown in Figure 5 for a one element GMM-PC and a nine element GMM-PC. Using only PCE, which is a one element GMM-PC, does not result in a converged solution with low error for all the univariate directions except for $v_I$. It should be noted that there can be significant variation of the test statistic over an orbital period, especially for highly eccentric orbits. As the number of splits
Fig. 6 Two-sample univariate Cramer-von-Mises metric in the RIC frame for GMM-PC, split along velocity, with respect to an MC simulation of 1,000,000 samples of an object in MEO. $L$ is the total number of coefficients required per direction for the multivariate polynomial and is analogous to the compute cost ($2L$).

along the velocity direction $N$ is increased, a lower computation cost results in a more accurate representation of the uncertainty. The norm of the $CvM$ values in all the position and velocity directions of the RIC frame is found in Figure 6. The $CvM$ value along the different directions can then be used to accept the hypothesis that the MC samples and the samples generated by the GMM-PC originate from the same distribution with a certain confidence. The $CvM$ has recently been used as an uncertainty realism metric for orbit propagation [65] and lookup tables can be used for hypothesis testing [65, 66]. Using the $CvM$ provides a holistic measure of accuracy for the distribution rather than computing selected moments.

Figure 7 clearly shows the accurate representation of the MC simulation by the GMM-PC. Point clouds have a tendency to exaggerate the visual effect amount of outliers because point density is not easily represented. There are many outliers present in the pure PCE and GMM solutions, especially in the radial direction. The GMM approximation is jagged because it is similar to plotting a circle with only 9 points. Each element still must approximate a curve with a line, so the semi minor axis of the post ellipse gets inflated to accommodate for curvature of the true sub-distribution each element is fitting.
Fig. 7 100,000 samples for velocity in the Radial-Intrack plane from the MC, a PCE, and GMM-PC and GMM simulations with an initial split applied along $v_e$ for the uncertainty propagation of a MEO object

B. Molniya Orbit

The second test case is an orbit from Jones et. al. [41]. The initial state and uncertainty for the Molniya orbit are found in Table 2. The orbit is propagated for 10 days to compare the final distributions. This case is known to not converge when PCE and Cartesian coordinates are used to represent the uncertainty [41]. Although only using GMM-PC expressed in Cartesian coordinates is shown to converge here, the traditional ME-gPC methods will also converge. Additionally, Jones et. al. [41] include uncertainty in the gravitation parameter, whereas the current implementation does not, because it is not required to illustrate the convergence issues.

This case is again highly non-linear along velocity $\xi_1$, direction because increasing the order of
Table 2 Initial osculating orbit elements [41] and uncertainty expressed in Cartesian coordinates in the ECI frame for an object in a Molniya orbit

<table>
<thead>
<tr>
<th>$a$</th>
<th>$e$</th>
<th>$i$</th>
<th>$\Omega$</th>
<th>$\omega$</th>
<th>$\nu$</th>
<th>$\sigma_x$</th>
<th>$\sigma_y$</th>
<th>$\sigma_z$</th>
<th>$\sigma_{v_x}$</th>
<th>$\sigma_{v_y}$</th>
<th>$\sigma_{v_z}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>26,562km</td>
<td>0.741</td>
<td>63.4°</td>
<td>90°</td>
<td>–90°</td>
<td>0°</td>
<td>10 m</td>
<td>10 m</td>
<td>10 m</td>
<td>1 m/s</td>
<td>1 m/s</td>
<td>1 m/s</td>
</tr>
</tbody>
</table>

the PCE does not result in reducing values of the coefficients and the behavior is similar to the behavior in Figure 4. The expansion along the other directions does converge.

The initial Gaussian distribution is split into a GMM with up to 9 elements along the $v_y$ direction. The GMM-PC is propagated for the desired flight time of 10 days and is sampled. The $C_vM$ values for the samples in the RIC frame with respect to a MC simulation is shown in Figure 8. Using only PCE does not result in a converged solution with low error for all the univariate directions except for $r_C$. As the number of splits along the velocity direction is increased, a lower computation cost results in an increased accuracy.

Because only two-body dynamics are used and the initial uncertainty in the position coordinates is only 10 m, the final distribution is thinly spread along the orbit. Therefore, the resulting distributions of the MC, PCE, GMM, and GMM-PC are shown as histograms in the RIC frame of the MC mean in Figure 9. The samples from the 9 element 6$^{th}$-order GMM-PC clearly form a good approximation of the MC histograms in all the univariate directions of the RIC frame. PCE and the pure GMM fail to approximate the tails of the distributions.

C. Geosynchronous Transfer Orbit

The third test case is a GTO with perturbations due to atmospheric drag, non-spherical Earth of degree and order 8, Solar Radiation Pressure (SRP), and third-body attraction of the Sun and the Moon. The initial state and uncertainty for the GTO are found in Table 3. The highest nonlinearity and therefore, the non-convergent source for the PCE is again the initial velocity uncertainty and the behavior is similar to Figure 4.

The $C_vM$ values for the samples in the RIC frame with respect to the MC simulation is shown in Figure 10. Using only a one element GMM-PC, does not result in a converged solution in the $v_x$
Fig. 8 Two-sample univariate Cramer-von-Mises metric in the RIC frame for GMM-PC, split along velocity, with respect to an MC simulation of 1,000,000 samples of an object in a Molniya orbit. $L$ is the total number of coefficients required per direction for the multivariate polynomial and is analogous to the compute cost.

<table>
<thead>
<tr>
<th>$\Omega$</th>
<th>$\omega$</th>
<th>$\nu$</th>
<th>$\sigma_x$</th>
<th>$\sigma_y$</th>
<th>$\sigma_z$</th>
<th>$\sigma_{vx}$</th>
<th>$\sigma_{vy}$</th>
<th>$\sigma_{vz}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>24.475km</td>
<td>0.731</td>
<td>7°</td>
<td>250°</td>
<td>8°</td>
<td>0°</td>
<td>5 m</td>
<td>5 m</td>
<td>5 m</td>
</tr>
</tbody>
</table>

Table 3 Initial osculating orbit elements and uncertainty expressed in Cartesian coordinates in the ECI frame for a GTO object

direction. The behavior of the point clouds and the histograms are similar to the MEO and Molniya cases, respectively and are therefore not presented.

In the three test cases presented so far, the ratio of the velocity and position uncertainties are relatively large compared to typical cases. Velocity uncertainties are usually a thousandth of the position uncertainty. Since the relative scale of the velocity uncertainties (up to one tenth) are larger here, this likely results in the need to apply the split along the velocity direction. Splitting in the velocity direction may not always be the optimal choice for all scenarios.

D. Low Earth Orbit

The final test case uses orbit elements to express the state uncertainty. The initial state and state uncertainty in EE are taken from Horwood et al. [29]. The time of flight, however, is increased to 65.6 hours. The increased flight time makes the uncertainty distribution of the SO wrap around the
Fig. 9 Histograms in the RIC frame of 1,000,000 samples from the MC, a PCE, and a GMM-PC and GMM with an initial split applied along $v_r$ for the uncertainty propagation of an object in a Molniya orbit whole orbit. The major source of the final uncertainty is the initial semi-major axis uncertainty \[29\] and the plot of the coefficients is similar the velocity direction of the previous cases as seen in Figure 4. Splitting the initial distribution along the semi-major axis direction ensures that the
Fig. 10 Two-sample univariate Cramer-von-Mises metric for PCE and GMM-PC, split along velocity, with respect to an MC simulation of 1,000,000 samples for the GTO case. \( L \) is the total number of coefficients required per direction for the multivariate polynomial and is analogous to the compute cost.

Angular spread in the true anomaly of each element is less than \( 2\pi \). The orbit uncertainty occupying the full range from 0 to \( 2\pi \) in the true anomaly is a common occurrence for debris field propagation cases. A rigorous treatment of uncertainty in a periodic coordinate such as the true anomaly in the context of space surveillance has been carried out by Horwood and Poore. \[67\] The GMM-PC fitting is carried out in EE but the \( CvM \) metric in Figure 11 is computed in Cartesian coordinates by converting the EE GMM-PC samples to Cartesian coordinates. There may be a small loss of uncertainty realism when converting the EE samples to Cartesian coordinates. With number of elements \( N > 1 \), the GMM-PC solution converges because the uncertainty in the true anomaly corresponding to each element is contained within the \([0, 2\pi]\) range.

<table>
<thead>
<tr>
<th>( a )</th>
<th>( h )</th>
<th>( k )</th>
<th>( p )</th>
<th>( q )</th>
<th>( l )</th>
<th>( \sigma_a )</th>
<th>( \sigma_h )</th>
<th>( \sigma_k )</th>
<th>( \sigma_p )</th>
<th>( \sigma_q )</th>
<th>( \sigma_l )</th>
</tr>
</thead>
<tbody>
<tr>
<td>6,980 km</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0(^\circ)</td>
<td>20 km</td>
<td>( 10^{-3} )</td>
<td>( 10^{-3} )</td>
<td>( 10^{-3} )</td>
<td>( 10^{-3} )</td>
<td>( 0.1\circ)</td>
</tr>
</tbody>
</table>

Table 4: Initial osculating orbit and uncertainty expressed in Equinoctial Elements \[29\]
Fig. 11 Two-sample univariate Cramer-von-Mises metric for PCE and GMM-PC, split along the semi-major axis, with respect to an MC simulation of 1,000,000 samples.

E. Weighted Least Squares Technique

Apart from the naive GMM-PC approach, the WLS approach is used for the MEO and GTO orbit uncertainty propagation cases, which use Cartesian Coordinates. The number of function evaluations are held constant and the number of elements in the GMM-PC is increased and the PCE order of each GMM-PC element is adapted based on the number of evaluation nodes available in its $3\sigma$ range. The WLS approach is suboptimal compared to the standard LS method because the function is not reevaluated for each element and the order decreases further from the central element. However, the accuracy is better than the standard PCE approach as seen in Figure 12.

(a). MEO CvMs

(b). GTO CvMs

Fig. 12 Two-sample univariate Cramer-von-Mises metric for PCE and Weighted Least Squares GMM-PC, split along velocity, with respect to an MC simulation of 1,000,000 samples.
VI. Conclusion

A Polynomial Chaos Expansion forms a surrogate model for uncertainty propagation through a non-linear function, which is computationally more efficient to sample when compared to a full Monte Carlo simulation. The performance of the PCE method depends on the nonlinearity of the function and the size of the initial uncertainty. For the four cases of uncertainty propagation shown here, increasing the order of the PCE results in slow convergence and a large computation cost. Splitting the initial Gaussian distribution into a GMM with the means of the weighted Gaussian distributions along the most influential direction (initial velocity vector for the Cartesian Coordinates cases and semi-major axis for the EE case presented here) reduces the uncertainty along that direction. Therefore, a more accurate description of the propagated uncertainty is found at a lower computation cost, compared to simply increasing the order of the PCE. However, the GMM-PC technique is only recommended for scenarios where a PCE does not converge, as is the case for orbit problems with large propagation times, large initial uncertainties, and highly eccentric orbits. The weighted least squares method for computing the coefficients of the PCE is capable of reusing the function evaluations so that accuracy of the GMM-PC method can be increased without increasing the computational load.

Comparing uncertainty propagation to FEM makes the GMM-PC method, which is similar to the ME-gPC, an $hp$-refinement method where the order of the PCE and the number of splits are independently adapted. When compared to a ME-gPC method, the GMM-PC technique is easier to implement because the PDF of each element remains Gaussian. Therefore, the analytical Hermite polynomials are used instead of developing a gPC framework to find orthogonal multivariate polynomials with respect to arbitrary PDFs. In the ME-gPC method, errors may creep into the computation of the orthogonal polynomials, while the error in the GMM-PC technique arises due to the approximation of the initial PDF with a finite number of GMM elements. Therefore, the best possible performance of the GMM-PC has an upper bound which is equivalent to the approximation accuracy of the GMM splitting library. The GMM-PC is also slightly suboptimal compared to the traditional ME-gPC because of the overlap of the Gaussian elements. However, the optimality is traded for ease of implementation. For extending a PCE implementation to the GMM-PC, only
the weights, means, and standard deviations of a univariate splitting library are required, which are previously archived and stored in a tabular manner.

Appendix

A. Hermite Polynomials

The weighting function for the probabilists’ Hermite polynomials is proportional to the standard normal distribution:

\[ w(x) = e^{-x^2/2} \] (27)

For Hermite polynomials, the three term relation is as follows [68]:

\[ \psi^*_{n+1}(\xi) = \xi \psi^*_n(\xi) - n \psi^*_{n-1}(\xi) \] (28)

Equation 28 can be used derive \( n^{\text{th}} \)-order orthogonal polynomials given \( \psi^*_0 = 1 \) and \( \psi^*_1 = \xi \). The polynomials are further made orthonormal:

\[ \psi_n(\xi) = \frac{\psi^*_n(\xi)}{\sqrt{n!}} \] (29)

References


[38] Li, X., Nair, P. B., Zhang, Z., Gao, L., and Gao, C., “Aircraft Robust Trajectory Optimization Using


