Summer 2020

Generalization of polynomial chaos for estimation of angular random variables

Christine Louise Schmid

Department: Mechanical and Aerospace Engineering

GENERALIZATION OF POLYNOMIAL CHAOS FOR ESTIMATION OF ANGULAR RANDOM VARIABLES

by

CHRISTINE LOUISE SCHMID

A DISSERTATION

Presented to the Graduate Faculty of the

MISSOURI UNIVERSITY OF SCIENCE AND TECHNOLOGY

In Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

in

AEROSPACE ENGINEERING

2020

Approved by:

Kyle J. DeMars, Advisor
Jacob E. Darling
Serhat Hosder
Robert Paige
Henry Pernicka

ABSTRACT

The state of a dynamical system will rarely be known perfectly, requiring the variable elements in the state to become random variables. More accurate estimation of the uncertainty in the random variable results in a better understanding of how the random variable will behave at future points in time. Many methods exist for representing a random variable within a system including a polynomial chaos expansion (PCE), which expresses a random variable as a linear combination of basis polynomials.

Polynomial chaos expansions have been studied at length for the joint estimation of states that are purely translational (i.e. described in Cartesian space); however, many dynamical systems also include non-translational states, such as angles. Many methods of quantifying the uncertainty in a random variable are not capable of representing angular random variables on the unit circle and instead rely on projections onto a tangent line. Any element of any space $\mathbb{V}$ can be quantified with a PCE if $\mathbb{V}$ is spanned by the expansion's basis polynomials. This implies that, as long as basis polynomials span the unit circle, an angular random variable (either real or complex) can be quantified using a PCE.

A generalization of the PCE is developed allowing for the representation of complex valued random variables, which includes complex representations of angles. Additionally, it is proposed that real valued polynomials that are orthogonal with respect to measures on the real valued unit circle can be used as basis polynomials in a chaos expansion, which reduces the additional numerical burden imposed by complex valued polynomials. Both complex and real unit circle PCEs are shown to accurately estimate angular random variables in independent and correlated multivariate dynamical systems.

ACKNOWLEDGMENTS

First, I would like to thank my advisor Dr. Kyle J. DeMars for taking me in as an undergraduate and introduce me to the complex* world of state estimation. Without his guidance*, I would not be the researcher, engineer, or person that I am today. I must also thank him for his endless supply of revisions: I think, I can finally say, I've got the hang of, where to correctly place all of my commas.

I would also like to thank my committee members, Drs. Jacob Darling, Serhat Hosder, Robert Paige, and Henry Pernicka. Whether through traditional teaching, or less formal methods of instruction, each of you has been instrumental in my education and in the preparation of this dissertation.

In addition to my committee, I would like to thank all of the other students throughout my graduate career that have kept me sane during this entire process. You helped me work through all of my roadblocks, even when my main contributions to the lab were groan inducing puns. A specific thank you goes to Matt and Kari; it hasn't been the easiest of semesters, some might even say it's been the most challenging of semesters, but we made it through, and everything is fine, because life, uh, finds a way.

Finally, I'd like to thank all my friends and family for supporting me through this entire process. Rob, you and Kodiak have been my rocks; I don't know where I'd be without you two. To my parents, Amy and Charlie Schmid, I have never felt like there was something I couldn't do, even when I decided to challenge the cliché "the sky's the limit".

* pun intended

TABLE OF CONTENTS

Page

LIST OF ILLUSTRATIONS

LIST OF TABLES

1. INTRODUCTION

1.1. MOTIVATION AND LITERATURE SURVEY

Engineering is an imperfect science. Noisy measurements from sensors in state estimation[1,2], a constantly changing environment in guidance[3,4,5], and improperly actuated controls[6] are all major sources of error. The more these sources of error are understood, the better the final product will be.

The question that arises is in what way should the error in a state be represented? Ideally, every variable with some sort of uncertainty associated with it, i.e. a random variable, would be completely and analytically described with its probability density function (pdf). Unfortunately, even if this is feasible for the initialization of the random variable, the pdf's evolution through time rarely yields a pdf with an analytic form when the governing dynamics are nonlinear.

When representing the error using the random variable's pdf is impractical, then an alternative is to use the moments of the pdf to represent the random variable in a dynamical system. The statistical moments of the random variable describe the shape of the pdf, including the average, the concentration, symmetry about the mean, etc.; therefore, providing enough information about the shape of the pdf is seemingly sufficient to describe the pdf. Given a finite number of parameters in the form of moments of the random variable that need to evolve over time, which is surely better than a continuous function without an analytic form. Unfortunately, it isn't the case that describing the random variable using only moments enables a sufficient quantification of the random variable for multiple reasons. Firstly, the complexity associated with calculating a moment and transforming it in time increases with each successive moment. Secondly, unless the pdf has a closed support (i.e. $[a, b]$), or has open support with specific properties (determinateness),

the complete sequence of moments is not sufficient to uniquely describe the pdf. For further discussion on the moment problem the reader is encouraged to reference Sohat and Tamarkin[7], which covers the Hausdorff moment problem for supports over closed intervals[8,9], the Stieltjes moment problem for supports over the semi-infinite interval[10] (i.e. $[0, \infty)$), or the Hamburger moment problem for supports over the infinite interval[11] (i.e. $(-\infty, \infty)$).

Some sort of concession must be made, resulting in an approximation of the random variable or its pdf. Many times, instead of the entire sequence of moments, only the first two moments (mean and variance/covariance) are considered, as with the unscented transform[12] (UT) and the Kalman filter including extended[13], quadrature[14], cubature[15], etc. variants. If the dynamics governing a state are not highly nonlinear, then assuming the pdf maintains a consistent relative shape (e.g. the pdf remains Gaussian, uniform, etc.) is not necessarily a poor assumption, if the shape of the pdf is known, then the moment problem becomes easier. Also, smaller uncertainties in the random variable generally result in smaller deformations of the pdf when compared with larger uncertainties. In these cases, the higher order moments are highly dependent on the first two moments; i.e., there is a minimal amount of unique information in the higher order moments. In contrast, if either the uncertainty is large or the dynamics become highly nonlinear, the higher order moments become less dependent on the first two moments and contain larger amounts of unique information. As a result, the error associated with using only the first two moments becomes significant[16,17].

In terms of computational resources, it is generally much easier to propagate the uncertainty of a random variable through time when the governing dynamics are linear. Even if the dynamics are nonlinear, they can be approximated as linear over very small time intervals by using a linearization method like the Taylor series expansion[18]. The Taylor series expansion, as it is used in the extended Kalman filter[13] for state estimation, utilizes the dynamics Jacobian to propagate the first two moments forward in time. Similar to the

UT, linearization can be very useful (even utilized in crewed missions to the surface of the Moon[19]). However, while the inclusion of higher order moments[20] is possible, and helps capture some higher order effects resulting in more accurate state estimation[21,22], it generally requires an assumption that the pdf is normal*. Additionally, the accuracy of a linearization deteriorates as the dynamics become nonlinear.

One method of quantifying uncertainty that does not require an assumption of the random variable's pdf, but is still able to approximate the shape of the pdf is the polynomial chaos expansion (PCE)[23,24,25,26]. PCE characterizes a random variable as a coordinate in a polynomial vector space. Useful deterministic information about the random variable lies in this coordinate, including the moments of the random variable[27]. The expression of the coordinate depends on the basis with which it is expressed. In the case of PCE, the bases are made up of polynomials that are chosen based on the assumed density of the random variable; however, any random variable can be represented using any basis[28]. It is strongly noted that assuming the density of the random variable simply eases computation; with enough computing power, any random variable can be quantified with any basis[29]. The most common basis polynomials are those that are orthogonal with respect to common pdfs, such as the Hermite-Gaussian and Jacobi-beta polynomial-pdf pairs.

One of the first uses of PCE with different Askey polynomials was to model the uncertainty in boundary conditions in computational fluid dynamics (CFD) simulations[30], showing that PCE can be used with Legendre polynomials to propagate initial uncertainty and that the estimates using PCE are shown to converge to the results of Monte Carlo simulations. Additional studies[31,32,33] into using PCE for CFD simulations have been performed, favoring PCE not only for the accuracy in the propagation of uncertainty, but for the ease of implementation. As opposed to linearization, which requires partial derivatives of the state, PCE can use existing dynamical models without alteration.

*Within this document "normal" and "Gaussian" are used interchangeably referring to a Gaussian pdf.

In addition to the aerodynamics, PCE has been used used to propagate the uncertainty of space based objects. Space presents a very challenging environment, requiring large amounts of autonomy in an extremely isolated environment; as a result, poor knowledge and utilization of uncertainty in the system can cause cascading, potentially catastrophic, errors, especially when considering missions can last years. Although PCE is a relatively new concept within the orbital mechanics field, it has proved to be a useful tool in uncertainty propagation including, the propagation of uncertainty using Cartesian and Poincaré elements[34]. This study observed the performance of PCE compared with linearization and the UT as methods of uncertainty propagation, showing that PCE is the only method (as they have been presented) that is capable of capturing higher order moments and propagation in Cartesian space required more samples than Poincaré elements for the same level of accuracy.

By no means is PCE confined to use in different disciplines of aerospace engineering. The study of epidemiology contains many mathematical models, the inputs of which are often not known deterministically. As opposed to measurable quantities like fevers or the presence of antigens, epidemiologists must include symptomatic factors that are much harder to model. For example, the transmission rate of a disease is of importance when estimating the impact of the disease. Using PCE to model the uncertainties in the human interactions that drive transmission rates not only provides a method of predicting the growth of a disease, but a sensitivity analysis using PCE (without additional computation) indicates how the transmission rate is affected by different types of human interaction.

While polynomial chaos has been well-studied for variables that exist in the n-dimensional space of real numbers ($\mathbb{R}^n$), many variables do not lie in this field. For example, if an angular random variable, such as the true anomaly of a body in orbit or vehicle attitude, were to be estimated using Askey-chaos, the random variable's projection

from $\mathbb{S}^n$ onto a tangent plane in $\mathbb{R}^n$ would be estimated, rather than on its natural manifold $\mathbb{S}^n$. When the uncertainty is very small, this approximation is relatively effective; however, as the uncertainty increases, this approximation becomes invalid.

In addition to increased approximation error as the uncertainty of an angular random variable increases, there are additional properties that must be considered. For example, the support of a 1-dimensional angle is of length 2π, but when considering the common unit circle, the points at $0°$ and $360°$ are the same point. Not only do the pdfs of angular random variables have to have the same value at these points, they must be continuous across them. Approximating an angle as a translational random variable makes no guarantee of this continuity.

Possibly the most beneficial aspect of polynomial chaos is that assumptions about the random variable are not necessary. Approximating an angular random variable as a translational random variable would induce an assumption that is not necessary. Rather than make this type of approximation, directional statistics can be used, which are statistical methods that are specifically formulated to analyze angular random variables on their natural manifold, the n-dimensional sphere $\mathbb{S}^n$. The directional statistics associated with angular random variables has been incorporated into state estimation, providing a method for estimating angular random variables directly on the n-dimensional special orthogonal group $(\mathbb{SO}^n)$[35,36,37].

1.2. CONTRIBUTIONS

This dissertation presents two different methods of representing an angular random variable as a polynomial chaos expansion. The proposed families of polynomials used for angular random variables are orthogonal with respect to measures on either the complex or real unit circle.

The existing framework of PCE uses the orthogonality of the basis polynomials in many ways that are invalid if the polynomials are complex valued. The first contribution of this dissertation is the generalization of PCE for use with complex valued basis polynomials. Recent work[38,39] has made a similar generalization of PCE presenting favorable results in the estimation of angular mean and variance; however, the work presented herein has been explored independently and includes the additional development of correlation between angular random variables. The presented PCE form is usually discussed as a method of representing an angular random variable; however, any random variable with a pdf on a complex manifold can be represented as a result of this generalization.

The second major contribution of this dissertation is the use of real valued polynomials that are orthogonal with respect to measures on the real unit circle in a chaos expansion. With real valued polynomials, the existing PCE framework is appropriate; as a result, the correlation between translational and angular random variables can be estimated directly.

Together, these methods present a way to represent angular (and general complex) random variables using a PCE without having to make any assumptions about the random variable at all. The reduction in assumptions not only produces more accurate results, but also allows produces a more adaptive tool that is not problem specific. Even when a real valued random variable is expanded using the complex expansion is considered, the mathematics result in an expansion that is equivalent to the traditional PCE. The primary benefits of PCE that are of most interest in this work are that PCE:

- does not make an assumption based on the shape of the random variable;
- represents the random variable itself, not just the pdf or the moments of the pdf, meaning higher order effects are not neglected; and
- it can be used with existing "black box" dynamics.

1.3. ORGANIZATION

The remainder of this dissertation is organized as follows:

- Section 2 contains mathematical preliminaries, including general concepts of bases that span different vector spaces. Bases of polynomial vector spaces are discussed in length including the established Askey and Szegő polynomials, as well as polynomials that are orthogonal with respect to arbitrary measures. Additionally, some of the common pdfs on both $\mathbb{R}^n$ and $\mathbb{S}^n$ are given along with the general form of statistical moments.

- Section 3 provides overviews of the UT and PCE as they are used for random variables on the real line. The process for implementation is also given for PCE.

- Section 4 presents the quantification of an angular random variable on either the complex or real unit circle. The generalization of PCE to suit complex valued random variables is developed as well as expressions for the first two moments of a correlated multivariate angular state. Then, the framework and specifications necessary for PCE on the real valued unit circle are given. The numerical complexities of each method are also discussed.

- Section 5 presents two simulations to evaluate the feasibility of PCE on the unit circle and compares results against the UT:

  - First, the two angles describing the state of a double pendulum are jointly estimated. Feasibility of PCE on the unit circle is examined by comparing the estimation of the first two moments against estimates achieved via Monte Carlo simulation.

  - Second, an object in orbit governed by two-body dynamics with initial uncertainty in the mean longitude and semimajor axis is simulated. The first two moments and the Hellinger distance from a Monte Carlo simulation are examined to determine not only feasibility, but also performance against other PCE methods and the UT.

- Section 6 presents conclusions about polynomial chaos on the unit circle based on performance and provides suggestions for continued development.

2. MATHEMATICAL PRELIMINARIES

2.1. MATHEMATICAL BASES

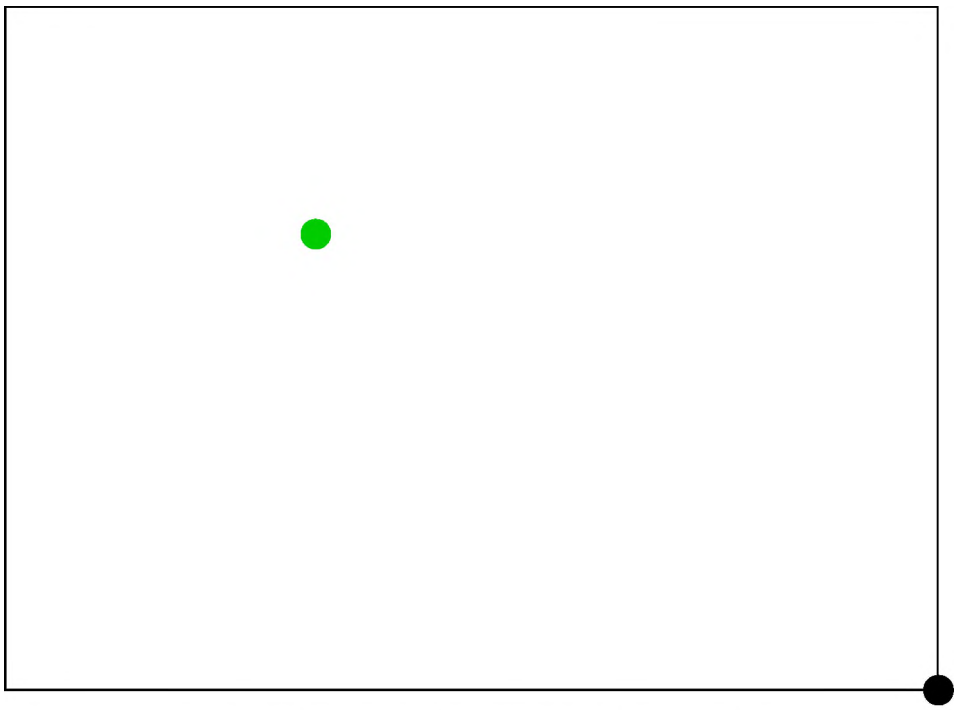
Let $\mathbb{V}^n$ be an n-dimensional vector space. A basis of this vector space is the minimal set of vectors that spans the vector space. An element $\alpha \in \mathbb{V}$ can be expressed in terms of an ordered basis, $\mathcal{B} = \{\beta_1, \beta_2, \ldots, \beta_n\}$, as the linear combination

$$\alpha = a_1\beta_1 + a_2\beta_2 + \cdots + a_n\beta_n, \tag{2.1}$$

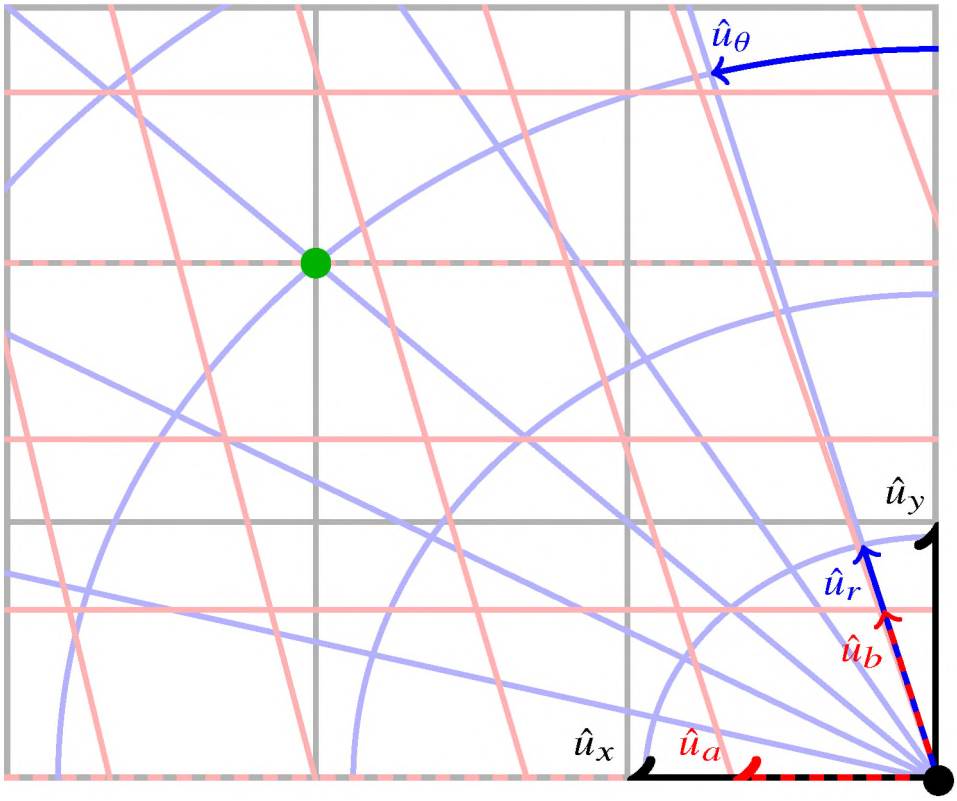
where $[\boldsymbol{a}] = [a_1, a_2, \ldots, a_n]$ is the coordinate of α. While any set of independent vectors can be used as a basis, different bases can prove beneficial – possibly by making the system more intuitive or more mathematically straightforward. When expressing the state of a physical system, the selection of a coordinate frame is effectively choosing a basis for the inhabited vector space. Figures (2.1) show how the comparison between the two different points in Figure (2.1a) can be expressed using the three different sets of bases in Figure (2.1b).

The different bases in Figure (2.1b) are polar (in blue), orthogonal Cartesian (in black), and a nonorthogonal basis in $\mathbb{R}^2$ (in red). The corresponding coordinates are $3\hat{u}_r + 3\hat{u}_\theta$, $2\hat{u}_x + 2\hat{u}_y$, and $7/3\hat{u}_a + 3\hat{u}_b$, respectively. All three expressions describe the relative position of the green dot with respect to the black dot; however, the resulting coordinates are all different.

Consider two types of frames commonly used in orbital mechanics, Earth-fixed frames and body-fixed frames. Earth-fixed frames are fixed at the center of the Earth and rotate about the z-axis along with the Earth, in other words, a fixed point on the Earth will not have any velocity in an Earth-fixed frame. For this reason, if a satellite's ground track is of high importance (such as weather or telecommunications satellites), the translational states (position, velocity, etc) in an Earth-fixed frame (e.g. latitude, longitude, altitude)



(a) Two points in a vector space without a basis.



(b) Two points in a vector space with three different sets of bases.

Figure 2.1. Description of two points within a vector space.

are easiest to visualize. However, in cases where a satellite's actions are dictated by other space-based objects (such as proximity operations), an Earth-fixed frame would not be ideal. In this case, the position of the space-based objects with respect to the satellite gives a more clear understanding of proximity and relative motion. In both cases, any frame *could* be used, but some frames have advantages over others.

While all of the simple coordinate systems mentioned are unique, they are all used to describe points in the vector space $\mathbb{R}^n$, where $n = 2$ in the example from Figures (2.1). This vector space consists of all matrices of size $[n \times 1]$ with real valued entries. In general, the vector space of matrices is $\mathbf{F}^{m\times n}$ where $\mathbf{F}$ could be the field of real ($\mathbb{R}$), imaginary ($\mathbb{I}$), or complex ($\mathbb{C}$) numbers: implying $\mathbf{F}^{m\times n}$ is the vector space of $\mathbf{F}$ valued matrices of size $m \times n$. These matrix spaces are not the only types of vector spaces, even though they are the most common (especially in elementary mathematics).

2.1.1. Inner Product. In general, not all vector spaces are inner product spaces; however, only inner product spaces will be considered herein. An inner product space is a vector space that has a defined inner product. For a vector space $\mathbb{V}$ over the field $\mathbf{F}$, let there be any elements $x, y, z \in \mathbb{V}$ and scalars $\alpha, \beta \in \mathbf{F}$. The inner product performs the following mapping

$$\langle x, y\rangle : \mathbb{V} \times \mathbb{V} \to \mathbf{F},$$

effectively mapping any two elements of $\mathbb{V}$ to a scalar value in $\mathbf{F}$. Additionally, the inner product has the following properties:

- linearity in the first argument:

$$\langle \alpha x, y\rangle = \alpha \langle x, y\rangle \tag{2.2a}$$

$$\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle \tag{2.2b}$$

- conjugate symmetry

$$\langle x, y\rangle = \overline{\langle y, x\rangle} \tag{2.3}$$

- positive definiteness

$$\langle x, x\rangle = \beta > 0 \qquad x \in \mathbb{V}\backslash\mathbf{0}, \tag{2.4}$$

where $\mathbf{0}$ is the zero element of $\mathbb{V}$. Alternatively, the inner product can be thought of as the percentage of x that lies in the y direction, or the projection of x onto y. If x is orthogonal to y, the inner product is 0, if $x = ay$ then the inner product is a. Figure (2.2) shows projection of a point A onto vectors x and y. Using the point A and its projection onto y from Figure (2.2), it is clear that a right triangle is formed between A, y, and the vector orthogonal to y (in this case x). Trigonometric properties can then be used to relate the angle between A and y according to

$$\cos\theta = \frac{\text{Real}\langle A, y\rangle}{||A||||y||}.$$

2.1.2. Orthogonal Bases. An orthogonal basis of an n-dimensional vector space is a subset of bases consisting of exactly n bases

$$\mathcal{B} = \{\beta_0, \ldots, \beta_{n-1}\}, \tag{2.5}$$

such that the inner product between any two basis elements, β_j and β_k, is proportional to the Kronecker delta ($\delta_{jk} = 1$ if $j = k$ and $\delta_{jk} = 0$ otherwise). Given mathematically with angle brackets, this orthogonal inner product takes the following form:

$$\langle \beta_j, \beta_k\rangle = c_\beta\, \delta_{jk} \qquad \forall j, k = 0, 1, \ldots, n\,.$$

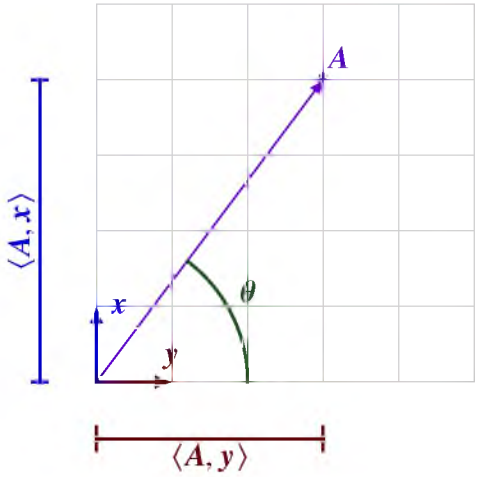


Figure 2.2. Application of inner product to describe projection and angle.

where the normalization factor c_β associated with the basis $\mathcal{B}$. In the event $c = 1$, the set is termed orthonormal. The k^{th} standard basis vector of $\mathbb{V}^n$ is generally the k^{th} vector of the n-dimensional identity; for example, consider the rows of the $n \times n$ identity matrix (I_n) and the unit vectors (u_k) that form the traditional basis for the n-dimensional Euclidean space

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \quad \rightarrow \quad \begin{aligned} u_1 &= [1 \quad 0 \quad \cdots \quad 0] \\ u_2 &= [0 \quad 1 \quad \cdots \quad 0] \\ &\vdots \\ u_n &= [0 \quad 0 \quad \cdots \quad 1] \end{aligned} \; . \tag{2.6}$$

Note that while the general definition of a basis in Eq. (2.5) contains a zeroth element, many geometric spaces (like the Euclidean space in Eq. (2.6)) begin with the first element for practicality. The standard bases are most common; however, there are infinitely many bases for each vector space. While it is possible to have a basis that is not an orthogonal set (as shown in Figure (2.1b)), the use of nonorthogonal bases is practically unheard-of and will not be considered in the remainder of this work.

2.2. POLYNOMIAL BASES

While it is common to restrict the idea of a vector space to matrix spaces, due to their relatively intuitive visualization, it was previously mentioned that polynomial vector spaces also exist. This type of vector space is of particular importance for polynomial chaos.

Let $\mathbb{P}^{n+1}$ be an $(n+1)$-dimensional vector space made up of all polynomials of positive degree n or less with standard basis $\mathcal{B} = \{1, x, \ldots, x^n\}$. The inner product for polynomial vector spaces over $\mathbf{R}$ (the field of real numbers) is defined as the functional inner product with respect to the function $\omega(x)$

$$\langle f(x), g(x) \rangle_{\omega(x)} = \int_{\mathbb{X}} f(x) g(x) d\omega(x), \tag{2.7}$$

where $\omega(x)$ is a non-decreasing function with support $\mathbb{X}$ and f and g are any two polynomials of degree n or less. Let $\mathbf{\Psi}^n(x)$ be a set of $n+1$ polynomials, with monotonically increasing order up to and including n. These polynomials form an orthogonal basis of $\mathbb{P}^{n+1}$ if, $\forall\, j; k = 0, 1, \ldots, n; 0, 1, \ldots, n$,

$$\langle \Psi_j(x), \Psi_k(x) \rangle_{\omega(x)} = \int_{\mathbb{X}} \Psi_j(x) \Psi_k(x) d\omega(x) = c_\psi \delta_{jk}, \tag{2.8}$$

where $\Psi_k(x)$ is the polynomial of order k and c_ψ is the normalization factor for the polynomials $\mathbf{\Psi}$.

2.2.1. Askey Polynomials. The most commonly used orthogonal polynomials are categorized in the Askey scheme, which groups the polynomials based on the generalized hypergeometric function,

$$
\begin{aligned}
{}_pF_q(a_1,\ldots,a_p;b_1,\ldots,b_q;x) &= \sum_{n=0}^{\infty} \frac{(a_1)_n \cdots (a_p)_n}{(b_1)_n \cdots (b_p)_n} \frac{x^n}{n!} \\
(a)_0 &= 1 \\
(a)_n &= \prod_{j=0}^{n\geq 1} (a+j)\,,
\end{aligned}
$$

with which they are generated[40,41,42]. Table 2.1 lists some of the polynomial families, their support, the non-decreasing function they are orthogonal with respect to (commonly referred to as a weight function), and the hypergeometric function they can be written in terms of. Among many other applications, the continuous polynomials are of particular importance when using Gauss quadrature for numerical integration. For an integral approximation of the form

$$
\int_a^b g(x)dx \approx \sum_{k=0}^{m} w_k g(x_k)\,,
$$

where the abscissas (x_k) and weights (w_k) are most commonly generated using the continuous polynomials in Table 2.1. For example, the Legendre polynomials are used in Gauss-Legendre quadrature for integrals over $[-1,1]$, the Laguerre polynomials in Gauss-Laguerre quadrature for integrals over $[0,\infty)$, and the Hermite polynomials in Gauss-Hermite quadrature for integrals over $(-\infty,\infty)$. Additional types of Gaussian quadrature are given by Abramowitz and Stegun[43]. Similarly, the discrete polynomials have been used for Gaussian summation formulas[44].

For completeness, Table 2.1 lists both continuous and discrete polynomial groups; however, the remainder of this work only considers continuous polynomials. The reader is encouraged to reference Xiu[28] for the generalization of polynomial chaos to use discrete polynomials.

Table 2.1. Common sets of Askey polynomials.

Type	Polynomial	Hypergeometric Series	Support	Weight Function
Continuous	Legendre	${}_2F_1$	$[-1, 1]$	Uniform
	Jacobi	${}_2F_1$	$[-1, 1]^*$	Beta
	Laguerre	${}_1F_1$	$[0, \infty)$	Exponential
	Probabilists' Hermite	${}_2F_0$	$(-\infty, \infty)$	Normal
Discrete	Charlier	${}_2F_0$	$\{0, 1, 2, \dots\}$	Poisson
	Meixner	${}_2F_1$	$\{0, 1, 2, \dots\}$	Negative Binomial
	Krawtchouk	${}_2F_1$	$\{0, 1, \dots, N\}$	Binomial
	Hahn	${}_3F_2$	$\{0, 1, \dots, N\}$	Hypergeometric

2.2.2. Szegő Polynomials. While the Askey polynomials are useful in many applications, their standard forms place them in the polynomial ring $\mathbb{R}[x]$, or all polynomials with real-valued coefficients that are closed under polynomial addition and multiplication. Additionally, these polynomials are orthogonal with respect to measures on the real line. In the event that a set of polynomials orthogonal with respect to a measure on a complex interval (e.g. the complex unit circle) is desired, the Askey polynomials are insufficient.

In his book, Szegő[45] uses the connection between points on the unit circle and points on a finite real interval to develop polynomials that are orthogonal on the complex unit circle, as well as any manifold in complex space. As a result of his work, polynomials orthogonal on the unit circle are now known as Szegő polynomials. Because the unit circle is defined to have unit radius, every point can be described on a real interval of length 2π

*Support of the beta function could either be the closed interval shown or the open interval $(-1, 1)$ based on the specified shape parameters.

(the entirety of this work will use the interval $[-\pi, \pi)$) and can be mapped from the angle $\theta \in [-\pi, \pi]$ to the complex variable $z = e^{i\theta}$, where i is the imaginary unit. All use of the variable z in the following work will correspond to this definition. Recall that the inner product in Eq. (2.7) is defined for polynomial vector spaces over $\mathbf{R}$. Now that the standard polynomial basis is complex, this definition is insufficient. For polynomial vector spaces over $\mathbf{C}$ (the field of complex numbers), the inner product introduces the complex conjugate from Eq. (2.3). When the Szegő polynomials are of concern, the inner product is given as

$$\langle \psi_m(z), \psi_n(z) \rangle_{\omega(\theta)} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \psi_m(z) \overline{\psi_n(z)} \omega(\theta) d\omega(\theta) = \delta_{mn}, \tag{2.9}$$

where the second argument of the inner product attains a complex conjugate (denoted by a bar) and $\omega(\theta)$ is a monotonically increasing weight function over the support. Note that the Kronecker delta is not scaled, implying all polynomials using Szegő's formulation are naturally orthonormal.

One example of a set of Szegő polynomials is the Rogers-Szegő polynomials, which have been well-studied[46,47,48] and were developed by Szegő based on work done by Leonard James Rogers over the q-Hermite polynomials. For a more detailed description of the relationship between the Askey scheme of polynomials and their q-analogs, the reader is encouraged to reference Koekoek and Swarttouw[42].

The popularity of these polynomials is partially due to the fact that they are orthogonal with respect to the mean centered wrapped normal distribution

$$\omega(\theta) = \frac{1}{\sqrt{-2\pi \ln(q)}} \sum_{j=-\infty}^{\infty} \exp\left\{ -\frac{(\theta - 2\pi j)^2}{-2\ln(q)} \right\}, \tag{2.10}$$

where $q \in (0, 1)$ describes the concentration of the pdf. Further discussion of the wrapped normal pdf can be found in Section 2.3.2.1.

The Rogers-Szegő polynomials are generated according to

$$\phi_n(z;q) = \sum_{k=0}^{n} \binom{n}{k}_q z^k \qquad 0 < q < 1\,, \tag{2.11}$$

where $\binom{n}{k}_q$ is the q-binomial coefficient

$$\binom{n}{k}_q = \frac{(q;q)_n}{(q;q)_k (q;q)_{n-k}}\,,$$

and

$$(a;q)_n = \prod_{j=0}^{n-1} (1 - aq^j)\,. \tag{2.12}$$

Because q is not constant, it is interesting to observe the polynomials at the asymptotic limits of q. As q approaches its bounds, the q-binomial, and thus the generating function, approach two different analytic polynomials. The q-binomial can be expanded as the product of k fractions

$$\begin{aligned} \binom{n}{k}_q &= \frac{(1-q^{n-k+1})(1-q^{n-k+2})\cdots(1-q^n)}{(1-q^1)(1-q^2)\cdots(1-q^k)} \\ &= \frac{1-q^{n-k+1}}{1-q}\,\frac{1-q^{n-k+2}}{1-q^2}\cdots\frac{1-q^n}{1-q^k}\,. \end{aligned} \tag{2.13}$$

Written in this way, it is apparent that as q approaches zero, each fraction approaches $1/1$, and thus the limit of the q-binomial is

$$\lim_{q\to 0} \binom{n}{k}_q = 1\,. \tag{2.14}$$

Using the limit behavior in Eq. (2.14) to generate polynomials using Eq. (2.11) results in polynomials whose coefficients all unitary with generating function

$$\phi_n(z) = \sum_{k=0}^{n} z^k \,.$$

As q approaches 1, each of the fractions in Eq. (2.13) approaches 0/0 necessitating L'Hopital's rule. The limit is therefore transformed into

$$\begin{aligned}\lim_{q\to 1}\binom{n}{k}_q &= \lim_{q\to 1}\frac{(n-k+1)q^{n-k}}{q^0}\frac{(n-k+2)q^{n-k+1}}{2q^1}\cdots\frac{nq^{n-1}}{kq^{k-1}} \\ &= \frac{n-k+1}{1}\frac{n-k+2}{2}\cdots\frac{n}{k}\end{aligned}$$

It becomes apparent that the limit can be expressed in terms of factorials as

$$\begin{aligned}\lim_{q\to 1}\binom{n}{k}_q &= \frac{\frac{n!}{(n-k)!}}{k!} \\ &= \frac{n!}{(n-k)!(k!)},\end{aligned}$$

which is the standard binomial coefficient resulting in the generating function

$$\phi_n(z) = \binom{n}{k} z^k \,.$$

Table 2.2 lists the polynomial coefficients in the two limit cases for $n = 0, \ldots, 4$ and $k = 0, \ldots, n$ in a type of "q-Pascal's triangle". In looking at the difference in the limit cases, it is obvious that coefficients increase with increasing q; in fact, the coefficients increase monotonically with increasing q. In effect, for the polynomial

$$\phi_n(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n \,,$$

the limit cases provide bounds for the coefficients such that $a_k \in \left(1, \binom{n}{k}\right)$

Table 2.2. q-binomial coefficients limit behavior as $q \to 0$ and $q \to 1$ for $n = 0$ to 4 and $k = 0$ to n.

	$\lim_{q\to 0} \binom{n}{k}_q$						$\lim_{q\to 1} \binom{n}{k}_q$				
$n = 0$	1					$\to$	1				
$n = 1$	1	1				$\to$	1	1			
$n = 2$	1	1	1			$\to$	1	2	1		
$n = 3$	1	1	1	1		$\to$	1	3	3	1	
$n = 4$	1	1	1	1	1	$\to$	1	4	6	4	1
	$k=0$	$k=1$	$k=2$	$k=3$	$k=4$		$k=0$	$k=1$	$k=2$	$k=3$	$k=4$

In addition to the generating function in Eq. (2.11), a three-step recurrence[49] exists, which is given by

$$\phi_{n+1}(z;q) = (1+z)\phi_n(z;q) - (1-q^n)z\phi_{n-1}(z;q)\,. \tag{2.15}$$

For convenience, the first five Rogers-Szegő polynomials are

$$\phi_0 = 1$$

$$\phi_1 = z + 1$$

$$\phi_2 = z^2 + (q+1)z + 1$$

$$\phi_3 = z^3 + (q^2+q+1)z^2 + (q^2+q+1)z + 1$$

$$\phi_4 = z^4 + (q+1)(q^2+1)z^3 + (q+1)(q^2+q+1)z^2 + (q+1)(q^2+1)z + 1\,.$$

As is apparent, the q-binomial term causes the coefficients to be symmetric, which eases computation, and additionally, the polynomials are naturally monic.

2.2.3. Arbitrary Polynomials. While the more common polynomials have attractive analytic generating functions or recursion expressions, it is sometimes necessary to use polynomials that are orthogonal with respect to different (possibly non-analytic) weight functions. In these cases, methods for defining arbitrary orthogonal polynomials are required.

Two different methods of generating polynomials with respect to specified weight functions are discussed. These methods are the the one outlined by Szegő specifically for polynomials orthogonal on the unit circle, and the Gram-Schmidt[50] orthogonalization.

2.2.3.1. General polynomial formulation. The following is a method of formulating orthonormal polynomials with respect to a given measure. While very similar, the method for measures on the real line and on the unit circle have slight differences. Where appropriate, the differing equations based on measure type are given in parallel.

The process begins with computing the moments of real line measures, and the Fourier coefficients of unit circle measurements, which are

$$m_n = \int_{\mathbb{X}} \omega(x) x^n dx \qquad \text{and} \qquad c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \omega(\theta) e^{-in\theta} d\theta \qquad\qquad \forall\, n \in \mathbb{N}^{+0},$$

respectively, where $\mathbb{X} \subseteq \mathbb{R}$ is the support of x and $\mathbb{N}^{+0}$ is the set of natural numbers including zero. These coefficient are then collected into three different matrices. The first two are standard Hankel matrices $\boldsymbol{H}_{n-1}$ and $\boldsymbol{H}_n$, with the generic forms

$$\boldsymbol{H}_k(\boldsymbol{m}) = \begin{bmatrix} m_0 & m_1 & \cdots & m_k \\ m_1 & m_2 & & \vdots \\ \vdots & & & \vdots \\ \vdots & & m_{2k-2} & m_{2k-1} \\ m_k & \cdots & m_{2k-1} & m_{2k} \end{bmatrix} \quad \text{and} \quad \boldsymbol{H}_k(\boldsymbol{c}) = \begin{bmatrix} c_0 & c_1 & \cdots & c_k \\ c_1 & c_2 & & \vdots \\ \vdots & & & \vdots \\ \vdots & & c_{2k-2} & c_{2k-1} \\ c_k & \cdots & c_{2k-1} & c_{2k} \end{bmatrix},$$

respectively, where $\boldsymbol{m}$ and $\boldsymbol{c}$ are the sets of moments and Fourier coefficients respectively. The third matrix is a slightly altered Hankel matrix, where the bottom row is the standard polynomial basis for real or complex unit circle measures, i.e. x^k or $z^k = e^{i\theta k}, k = 0, 1, \ldots, n$. This gives the results for $\boldsymbol{X}_n(m)$ and $\boldsymbol{X}_n(c)$, respectively to be

$$\boldsymbol{X}_n(\boldsymbol{m}) = \begin{bmatrix} m_0 & m_1 & \cdots & m_n \\ m_1 & m_2 & & \vdots \\ \vdots & & & \vdots \\ \vdots & & m_{2n-2} & m_{2n-1} \\ 1 & \cdots & x^{n-1} & x^n \end{bmatrix} \quad \text{and} \quad \boldsymbol{X}_n(\boldsymbol{c}) = \begin{bmatrix} c_0 & c_1 & \cdots & c_n \\ c_1 & c_2 & & \vdots \\ \vdots & & & \vdots \\ \vdots & & c_{2n-2} & c_{2n-1} \\ 1 & \cdots & z^{n-1} & z^n \end{bmatrix}.$$

With the matrices $\boldsymbol{H}_{n-1}$, $\boldsymbol{H}_n$, and $\boldsymbol{X}_n$ computed, the n^{th} order polynomial can be computed directly according to

$$\phi_n = (|\boldsymbol{H}_{n-1}|\,|\boldsymbol{H}_n|)^{-1/2}|\boldsymbol{X}_n|,$$

where $|\cdot|$ denotes the matrix determinant. The zeroth order polynomial is simply

$$\phi_0 = H_0^{-1/2}, \tag{2.16}$$

which is 1 if the measure ω is a pdf.

Now consider the computational requirements associated with this method. For a polynomial of degree n, $2n + 1$ moments or Fourier coefficients must be calculated to populate the Hankel matrices. Then, the determinants of $\boldsymbol{H}_{n-1}$ $(n \times n)$, $\boldsymbol{H}_n$ $(n+1 \times n+1)$, and $\boldsymbol{X}_n$ $(n+1 \times n+1)$ must be calculated. While this general formulation is cumbersome, it does provide a framework for developing a set of polynomials orthogonal with respect to any conceivable continuous weight function. It is additionally beneficial that this process is not preformed sequentially, which reduces cascading numerical inaccuracies. Beyond the work

by Szegő, Simon[51,52] and Geronimus[53] have written books on polynomials orthogonal on the unit circle, providing in-depth analysis of many aspects of the polynomials including the generation process: specifically the asymptotic behaviors in polynomials with very large orders.

2.2.3.2. Gram-Schmidt. Another method of developing a set of orthogonal polynomials is the Gram-Schmidt method of orthogonalization. The most commonly used application of this process is to orthogonalize a set linearly independent vectors into an orthonormal basis (which can be used for matrix QR decomposition), but the process extends to polynomials. As a result the polynomial Gram-Schmidt orthogonalizes the linearly independent set of polynomials $\{x^n\}$ for $n \in \mathbb{N}^{+0}$ with respect to a weight function $\omega(x)$. The Gram-Schmidt is performed sequentially and is initialized with the zeroth polynomial being 1 just like Eq. (2.16)

$$\phi_0(x) = 1\,.$$

From here the first order polynomial, x, is orthogonalized by removing any amount of x that is not orthogonal to $\phi_0(x)$ with respect to $\omega(x)$

$$\phi_1(x) = x - \frac{\langle x, \phi_0\rangle_{\omega(x)}}{\langle \phi_0, \phi_0\rangle_{\omega(x)}}\phi_0\,,$$

resulting in a first order polynomial orthogonal to $\phi_0(x)$ with respect to $\omega(x)$. The orthogonalization of a vector A given a vector ϕ_0 is shown in Figure (2.3). A is linearly independent of ϕ_0, so when the components of A that line in the ϕ_0 direction are removed, the remaining vector (ϕ_1) is nontrivial and lies orthogonal to ϕ_0. In the figure, geometric notation is used indicating the parallelism of the two horizontal lines and the orthogonality of ϕ_0 and ϕ_1. While this figure shows vectors and not polynomials, the same basic principle applies. In Figure (2.3), the double arrows indicate parallel lines.

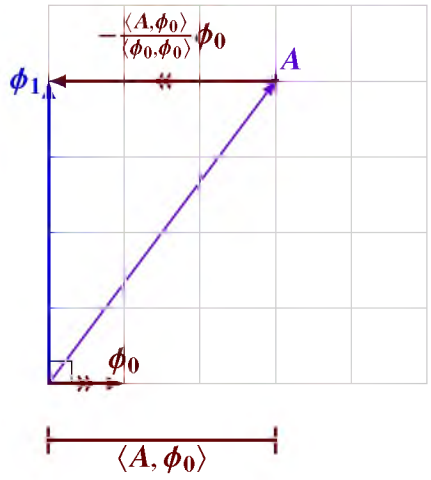


Figure 2.3. Generation of an orthogonal vector given an initial vector ϕ_0 and a linearly independent vector A.

This process continues for the n^{th} element by taking the polynomial x^n and removing any part of that polynomial that is not orthogonal with respect to all of the previous polynomials.

$$\phi_n(x) = x^n - \sum_{k=0}^{n-1} \frac{\langle x^n, \phi_k \rangle_{\omega(x)}}{\langle \phi_k, \phi_k \rangle_{\omega(x)}} \phi_k \,.$$

The computation of inner products is particularly important to this process, and can therefore become numerically unstable. For this reason, polynomials in this work that do not have established generating functions are developed according to Section 2.2.3. If Gram-Schmidt orthogonalization is preferred over the methods in Section 2.2.3, many variants of this process exist including the modified Gram-Schmidt and block Gram-Schmidt algorithm[54].

2.3. PROBABILITY DENSITY FUNCTIONS

The final mathematical preliminary to be covered is the probability density function. The pdf describes a random variable based on the likelihood it has of taking on a certain range of variables. While the probability of a random variable X taking a singular value x is zero, the probability P that $X = x$ can be approximated by integrating the pdf of X (p_X)

$$P(x) \approx \int_x^{x+\delta} p_X(t)dt\,.$$

where δ is an infinitesimally small element. Note that the probability P and pdf p are not the same quantity. As δ approaches 0, the approximation becomes more accurate.

In addition to the pdf, random variables are also described by a function taking on values from 0 to 1 that describes the probability that the random variable takes a value less than or equal to a given value, known as the cumulative distribution function (cdf). If the random variable X has lower bound $\mathbb{X}$, then the probability X is less than or equal to x is given by

$$\begin{aligned} F_X(x) &= P(X \leq x) \\ &= \int_{\mathbb{X}}^{x} p_X(t)dt\,. \end{aligned}$$

Because a pdf can never take a negative value, the cdf is a monotonically increasing function (assuming a continuous pdf) between 0 and 1. These properties are obvious when considering that the pdf is a positive function and that the probability that a random variable can assume any of the values in its domain is 1, implying the random variable exists over its domain, i.e.

$$\int_{\mathbb{X}}^{\mathbb{Y}} p_X(t)dt = 1\,.$$

2.3.1. Densities in $\mathbb{R}$. Most common in real-life applications utilize random variables in $\mathbb{R}$, resulting in univariate pdfs on the real line. A real valued pdf can have any support on real line; e.g., the beta pdf has a closed, bounded support, the exponential pdf has a semi-infinite support, and the normal pdf has infinite support. The most naturally occurring pdf in practical aerospace applications is the normal, or Gaussian, distribution with its distinctive "bell" shaped pdf. For convenience, the univariate normal pdf is given as

$$p_n(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \tag{2.17}$$

and the univariate normal cdf is

$$F_n(x;\mu,\sigma) = \frac{1}{2} + \frac{1}{2}\text{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right), \tag{2.18}$$

where μ and σ are the mean and standard deviation, respectively, and $\text{erf}(\cdot)$ is the error function given by

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt\,.$$

Note that in Eqs. (2.17) and (2.18) the subscript n indicates a random variable that is distinctly random in shape, parameterized by μ and σ, as opposed to the generic X in Section 2.3.

2.3.2. Densities on $\mathbb{S}$. Since the primary focus of this dissertation is angular random variables on both the complex and real unit circles, the most important pdfs are those on the unit circle. The pdfs of circular random variables have the same properties as real line random variables, with the addition that the support of any circular random variable must be 2π or shorter, and the pdf must be continuous across the end points of the support.

2.3.2.1. Wrapped normal density. Any pdf on the real line can be wrapped around the unit circle creating circular analogs to pdfs on $\mathbb{R}$. Let $g(x)$ be a pdf; then, the associated wrapped pdf $p_w(\theta)$ is

$$p_w(\theta) = \sum_{k=-\infty}^{\infty} g(\theta + 2\pi k) \,. \tag{2.19}$$

For example, wrapping a normal distribution takes the pdf

$$p_n(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \qquad -\infty < x < \infty,$$

where μ and σ are the mean and standard deviation, respectively, and wraps it around the unit circle, resulting in

$$p_{wn}(\theta; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \sum_{k=-\infty}^{\infty} \exp\left\{-\frac{(\theta - \mu + 2\pi k)^2}{2\sigma^2}\right\} \qquad -\pi \le \theta < \pi \tag{2.20}$$

which was mentioned briefly in Section 2.2.2 due to its connection with the Rogers-Szegő polynomials.

In many instances, the cdf is required. In general, the cdf of a wrapped distribution can be written in terms of the unwrapped pdf, similar to Eq. (2.19). If $G(x)$ is the cdf of $g(x)$, then the cdf of p_w is

$$F_w(\theta) = \sum_{k=-\infty}^{\infty} G(\theta + 2\pi k) - G(\pi(2k-1)) \,.$$

If the standard normal distribution (i.e. $\mu = 0, \sigma = 1$) is wrapped around the unit circle, the result is the standard wrapped normal pdf. This pdf is of particular interest in this work; therefore, the cdf is given for convenience

$$F_{wn}(\theta) = \frac{1}{2} \sum_{k=-\infty}^{\infty} \operatorname{erf}\left(\frac{\theta + 2\pi k}{\sqrt{2}}\right) - \operatorname{erf}\left(\frac{\pi(2k-1)}{\sqrt{2}}\right) \,. \tag{2.21}$$

The error function is symmetric about zero and it asymptotically approaches ± 1 at a rate such that $\text{erf}(\pm 2\pi) \approx \pm 1$; therefore, the infinite summation can be approximated as

$$F_{wn}(\theta) \approx \frac{1}{2}\text{erf}\left(\frac{\theta}{2}\right) + \sum_{k=-1}^{1} \text{erf}\left(\frac{\theta + 2\pi k}{\sqrt{2}}\right) .$$

This expression becomes particularly useful when drawing samples of the wrapped normal distribution, both randomly and intelligently.

Written according to Eq. (2.20), it is easy to see the similarity between the wrapped normal pdf and the normal pdf; however, the implementation of an infinite summation can be computationally inefficient. It is therefore important to acknowledge another form of the wrapped normal pdf:

$$p_{wn}(\theta; \mu, \sigma^2) = \frac{1}{2\pi}\vartheta_3\left(\frac{\theta - \mu}{2}, \sqrt{e^{-\sigma^2}}\right), \tag{2.22}$$

where ϑ_3 is the Jacobi-theta function given by[43]

$$\vartheta_3(z, q) = 1 + 2\sum_{k=1}^{\infty} q^{k^2}\cos(2kz) .$$

While the Jacobi-theta function contains an infinite summation, many programming languages have built-in functions that provide more accurate results than user coded summations.

Zero-mean normal distributions with varying values of σ are wrapped, with the results shown in Figure (2.4). For relatively small σ values (0.5 and 0.7 in Figure (2.4)), the wrapping effect is not very noticeable; however, as σ increases, the pdf approaches a wrapped uniform distribution. Additionally, with σ values greater than 1, the pdf across the entire support becomes non-zero and the "tails" that are present in normal pdfs are no longer present.

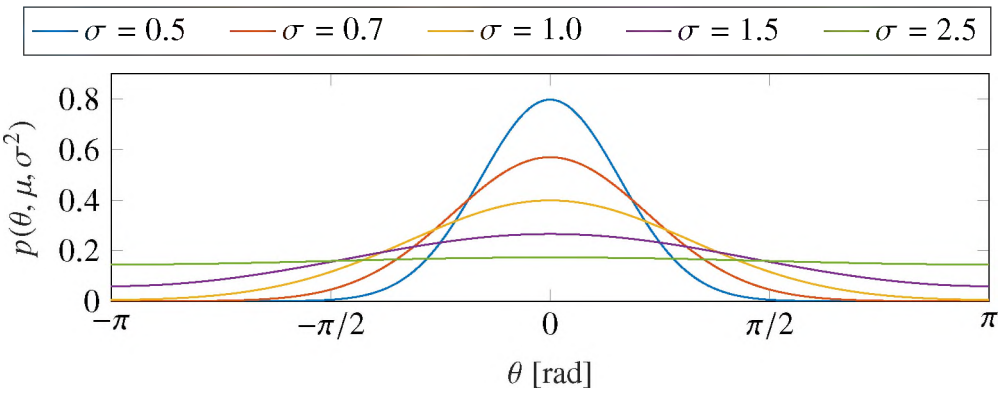


Figure 2.4. Wrapped normal distribution with multiple values of σ.

2.3.2.2. von Mises density. Another common pdf used in directional statistics is the von Mises/von Mises-Fisher distribution. The von Mises distribution lies on $\mathbb{S}^1$ (the subspace of $\mathbb{R}^2$ containing all points that are unit distance from the origin), whereas the von Mises-Fisher has extensions into higher dimensional spheres. The circular von Mises pdf is given as

$$p_{vm}(\theta; \mu, \kappa) = \frac{e^{\kappa \cos(\theta - \mu)}}{2\pi I_0(\kappa)}, \tag{2.23}$$

where μ is the mean angular direction, κ is a concentration parameter, similar to the inverse of standard deviation, and I_0 is the zeroth order modified Bessel function of the first kind. Unfortunately, the von Mises distribution does not have a cdf with an analytic form, presenting a distinct advantage for the wrapped normal pdf. For multiple concentration parameters, the von Mises pdf is shown in Figure (2.5). It is clear to see the similarities between the Gaussian pdf for higher concentration parameters, and similar to the wrapped normal pdf, as the spread of the pdf increases, it approaches a wrapped uniform pdf.

2.3.2.3. Bingham density. Another circular distribution is the Bingham distribution[55]. Like the von Mises distribution, the Bingham distribution is similar to the normal distribution; however, it has the distinct difference of being antipodally symmetric, i.e.

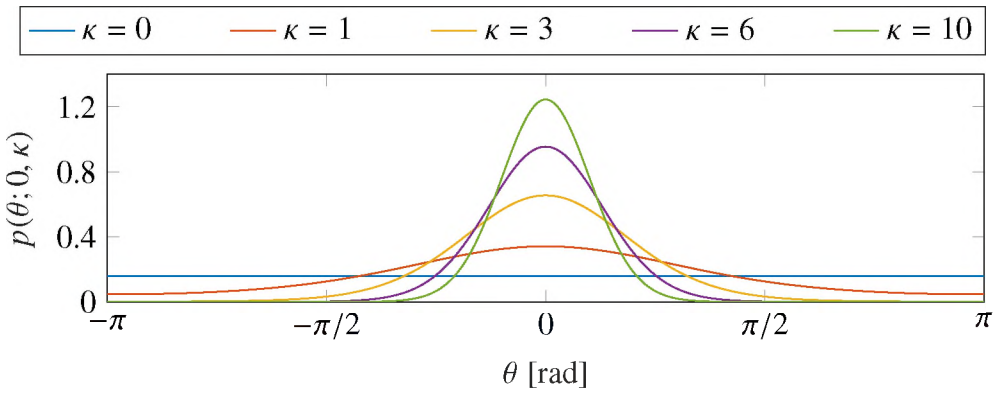


Figure 2.5. von Mises distribution with multiple values of κ.

$p_b(\theta) = p_b(-\theta)$. This property has made it appealing for expressing attitude pdfs when quaternions are used[37]. The pdf of the circular Bingham is given as

$$p_b(\boldsymbol{x}; \boldsymbol{Z}, \boldsymbol{M}) = \left({}_1F_1\left(\frac{1}{2}; \frac{3}{2}; \boldsymbol{Z}\right)\right)^{-1} \exp\left\{\text{tr}(\boldsymbol{Z}\boldsymbol{M}^T\boldsymbol{x}\boldsymbol{x}^T\boldsymbol{M})\right\} \tag{2.24}$$

where $\boldsymbol{x} \in \mathbb{S}^1$, $\boldsymbol{M}$ is a matrix describing the orientation of the mean direction, $\boldsymbol{Z}$ is a diagonal matrix of concentration parameters, and ${}_1F_1$ is a confluent hypergeometric function. For the circular case, the matrices $\boldsymbol{Z}$ and $\boldsymbol{M}$ have dimension 2×2. If the final element of $\boldsymbol{Z} = 0$, then only one value, z, remains to be set. Figure (2.6) shows the result of setting $\boldsymbol{M}$ as the identity matrix, (i.e. no rotation) and varying z. Note that if positive values are used for $\boldsymbol{Z}$, the mean direction axis lies through the trough of the distribution, rather than the apex.

The zeroth order modified Bessel function of the first kind in Eq. (2.23) and the confluent hypergeometric function in Eq. (2.24) are both functions of infinite summations, this along with the Jacobi-theta form of the wrapped normal pdf in Eq. (2.22) helps to show some of the similarities in the three pdfs, as well as challenges in implementation.

2.3.3. Statistics. While the pdf fully describes a random variable, it can rarely be described in full after evolution of the random variable over time. Instead, the state and its uncertainty are commonly defined by the moments of the pdf, which can be approximated much more easily than the full pdf. Most notable of these moments are the mean and

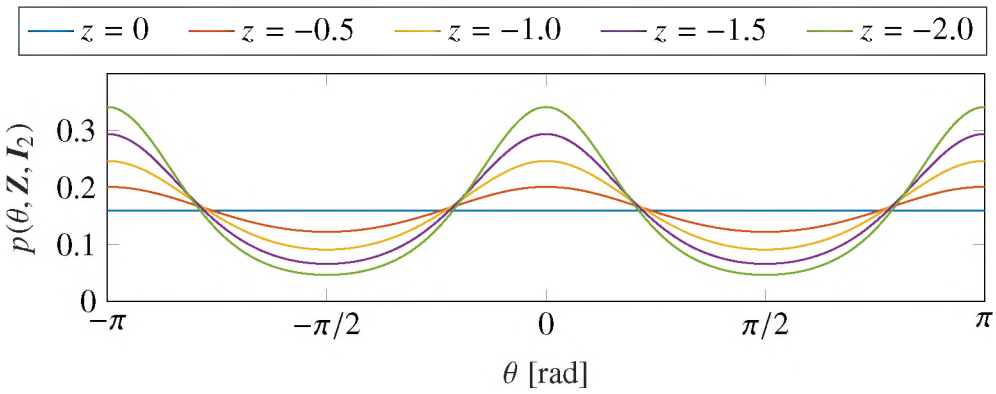


Figure 2.6. Bingham distribution with multiple values of z.

variance, the first two central moments. For pdfs ($p(x)$) on the real line that are continuously integrable, the central moments are given as

$$\begin{aligned} \mu_1 &= \int_S x p(x) dx \\ \mu_k &= \int_S (x - \mu_1)^k p(x) dx \quad k \geq 2 \,. \end{aligned} \tag{2.25}$$

Although less utilized, the raw moments are given as

$$\rho_k = \int_S x^k p(x) dx \,. \tag{2.26}$$

In directional statistics, raw moments are used exclusively and have a slightly different form, which is

$$m_k = \int_S \left(e^{i\theta}\right)^k p(\theta) d\theta \,, \tag{2.27}$$

which is commonly broken up into an angular (θ_k) and a length (R_k) part according to

$$\theta_k = \text{Arg}(m_k) = \tan^{-1} \frac{\text{imag}(m_k)}{\text{real}(m_k)} \qquad \text{and} \qquad R_k = \|m_k\| \,,$$

where $\|\cdot\|$ is the L^2-norm. Similar to pdfs on the real line, the higher order moments give additional information about the shape of the pdf that is not captured by the mean and concentration; however, for circular pdfs the mean direction and the concentration are both derived from the first central moment[35]. Whereas the mean direction comes directly from the angle of the first moment, the "spread" of the pdf comes from the length as either the circular variance (V) calculated by

$$V_1 = 1 - R_1\,, \tag{2.28}$$

or the "circular standard deviation" (σ) calculated by

$$\sigma_1 = \sqrt{-2\ln(R_1)}\,. \tag{2.29}$$

3. UNCERTAINTY QUANTIFICATION

The quantification of a random variable and associated propagation of its uncertainty is a well-studied topic. The true state of an element in all but the simplest of environments is typically unknown, requiring an expression for the uncertainty associated with the state. Not only must the uncertainty in the state be quantified at a given point in time, but as the state evolves, the uncertainty does as well. In cases where the governing dynamics are linear, the entire, continuous, pdf of the state can be propagated directly. However; the types of dynamics considered herein are nonlinear, given by a differential equation of the type

$$\dot{x}(t) = f(x(t)), \tag{3.1}$$

where $x(t)$ is continuous-time state of the system and f is the nonlinear function describing the evolution of x in time, t. In many cases, the differential equation in Eq. (3.1) cannot be solved analytically, and as a result, the pdf describing the state uncertainty becomes non-analytic.

While the full knowledge of the uncertainty in the state provided by the pdf is ideal, its application is generally implausible. This introduces the need for different methods of quantifying the uncertainty in a state that are feasible when the dynamics governing the state are nonlinear.

Choosing a type of uncertainty quantification (UQ) is a trade-off between numerical complexity and accuracy. a Monte Carlo simulation[56] is simple to implement, as only a singular sample is transformed by the dynamics at a time; however, a large number of independent samples are required to achieve an accurate estimate of the propagated state. As a result, Monte Carlo simulations tend to be used as a baseline to evaluate estimates obtained via other, more computationally efficient methods of UQ.

Of particular interest in this work are the unscented transform and polynomial chaos expansion. Chief among reasons for their interest is that when propagating uncertainty, neither require an assumption be made with regards to the shape of the random state's pdf.

3.1. UNSCENTED TRANSFORM

The unscented transform[12] utilizes a number of specifically, and deterministically, chosen sigma points and associated weights to quantify the first two moments of a random variable (or collection of random variables). When drawn, the sigma points and weights fully capture the mean and covariance of an n-dimensional state. At time t, calculation of the sigma points utilizes the mean ($\boldsymbol{\mu}_t$) and matrix square root factor $\sqrt{\boldsymbol{P}_t}$ such that $\boldsymbol{P}_t = \sqrt{\boldsymbol{P}_t}\sqrt{\boldsymbol{P}_t}^T$, which can be attained using methods such as LU decomposition[57] (section 3.5), Schur decomposition[57] (section 2.3), or Cholesky decomposition[58]. The expressions for drawing $2n + 1$ sigma points are give as

$$\boldsymbol{\Sigma}_{t,0} = \boldsymbol{\mu}_t \tag{3.2}$$

$$\boldsymbol{\Sigma}_{t,k=1:n} = \boldsymbol{\mu}_t + \sqrt{n + \lambda}\left[\sqrt{\boldsymbol{P}_t}\right]_k \tag{3.3}$$

$$\boldsymbol{\Sigma}_{t,k=n+1:2n} = \boldsymbol{\mu}_t - \sqrt{n + \lambda}\left[\sqrt{\boldsymbol{P}_t}\right]_k , \tag{3.4}$$

where $[\cdot]_k$ indicates the k^{th} column of the matrix and λ is a parameter defined by[59]

$$\lambda = \alpha^2(n + \kappa) - n .$$

In λ, the value of α is directly related to the distance of all but the centroid, or zeroth, sigma points from the mean, as α approaches zero, the non-centroidal sigma points converge on the mean, and as $|\alpha|$ increases, so do the distances of the non-centroidal sigma points from the mean. The same is true for κ, but large changes in κ more slowly affect the non-centroidal sigma point placement.

The two sets of weights for the mean (w^{μ}) and covariance (w^{Σ}) are calculated according to

$$w_0^{\mu} = \frac{\lambda}{n+\lambda} \qquad w_k^{\mu} = \frac{1}{2(n+\lambda)} \tag{3.5}$$

$$w_0^{\Sigma} = \frac{\lambda}{n+\lambda} + (1-\alpha^2+\beta) \qquad w_k^{\Sigma} = \frac{1}{2(n+\lambda)} \qquad k = 1, 2, \ldots, 2n, \tag{3.6}$$

where β is used when prior information about the covariance is available[59,60]. The choices of α and κ not only affect the placement of the sigma points, but also the weights. In general, increasing $|\alpha|$ and κ decrease the weights.

After being generated, each of the sigma points are transformed according to some nonlinear function

$$\boldsymbol{\Sigma}_{t+1} = \boldsymbol{f}\left(\boldsymbol{\Sigma}_{t+1}\right),$$

which are then used to estimate the mean and covariance of the transformed state according to

$$\begin{aligned} \boldsymbol{\mu}_{t+1} &\approx \sum_{k=0}^{2n} w_k^{\mu} \boldsymbol{\Sigma}_{t+1,k} \\ \boldsymbol{P}_{t+1} &\approx \sum_{k=0}^{2n} w_k^{\Sigma} (\boldsymbol{\Sigma}_{t+1,k} - \boldsymbol{\mu}_{t+1})(\boldsymbol{\Sigma}_{t+1,k} - \boldsymbol{\mu}_{t+1})^T . \end{aligned}$$

To visualize this process, refer to Figure (3.1). The analytic true initial (left) and final (right) states are represented by the mean ("+") and $1-\sigma$ covariance interval (ellipses). The 5 sigma points (using $\alpha = 0.5$, $\beta = 0$, and $\kappa = 1$) that fully capture the mean and covariance of the initial state are shown scaled according to weight. From the initial state, the sigma points are transformed (connected to show initial and final locations, not necessarily the path taken) and used to estimate the final mean ("×", which appears as a $*$ at the final time due overlapping with the true mean) and covariance (dashed ellipse). The true and

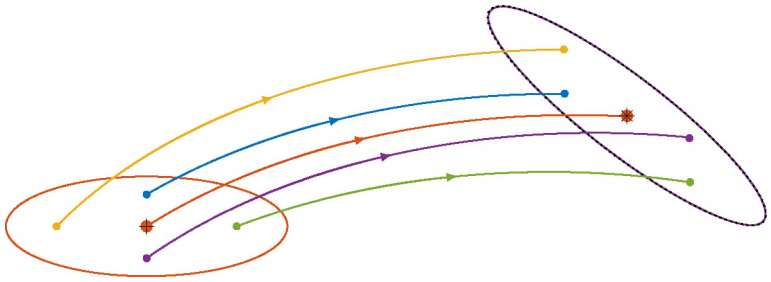


Figure 3.1. Estimation of propagated mean and covariance using the unscented transform with $2n + 1$ sigma points.

estimated mean and covariance are both shown clearly indicating the accuracy of the UT. It should be noted that the error ellipses suggest Gaussian distributions; however, the ellipse is chosen purely for illustrative purposes as the UT makes no assumption on the shape of the pdf.

The accuracy in estimating the mean and covariance of a state in nonlinear dynamics, as well as a relatively low number of sigma points, has made the UT a popular choice not only in uncertainty propagation, but also in filtering based state estimation[61,60]. While higher order moments can be can be calculated using the unscented transform[62], the process to develop expression for these moments is cumbersome requiring additional weights and sigma points and associated rules for calculation. Beyond the estimated moments, all other information about the pdf is lost.

3.2. POLYNOMIAL CHAOS EXPANSION

In many cases, the mean and variance of a random variable are insufficient representations of a random variable. It is therefore necessary to use a more complete representation of a random variable.

The following provides a brief overview of how polynomial chaos was initially conceptualized. There are many instances where specific types of random variables are mentioned; however, generalizations have been made and are covered in this work following the conceptualization.

Every random variable X is defined in some probability space $(\Omega, \mathcal{F}, P)$, where Ω is the sample space of X, $\mathcal{F}$ is the associated σ-algebra, and P is the probability measure. The space of mean centered Gaussian random variables is a Gaussian linear space, which when closed is known as a Gaussian Hilbert space $\mathcal{G}$, meaning any sequence in $\mathcal{G}$ converges to an element in $\mathcal{G}$. This closure becomes useful when examining the moments of an element $\xi \in \mathcal{G}$. Gaussian Hilbert spaces are subspaces of $L^2(\Omega, \mathcal{F}, P)$, which means that the L^2 inner product and norm is defined on these spaces[24]. With the space defined, a way to quantify $\mathcal{G}$ is needed, providing a way of describing a random variable $\xi \in \mathcal{G}$.

The fundamental component of polynomial chaos is that the space L^2 can be expressed as the infinite direct sum of polynomial spaces that have strictly increasing orders, known as the Wiener chaos decomposition[24]. The application of this in homogeneous chaos[23] uses the Wiener chaos decomposition to expresses the element $\varepsilon \in L^2(\Omega, \mathcal{F}, P)$ in terms of a linear combination of elements that span the previously mentioned polynomial spaces. Since $\xi \in \mathcal{G}$ is a Gaussian random variable and the (normalized probabilists') Hermite polynomials have strictly increasing orders and are orthogonal with respect to the standard normal distribution[63], these are the logical choice for the polynomial space spanning elements. In other words, ε can be expressed as a linear combination of Hermite polynomials evaluated at ξ.

Not only does this decomposition apply for Gaussian processes, but the Cameron-Martin theorem[64] says that any process with a finite second moment can be expressed as a linear combination of dense orthogonal polynomial basis functions. In other words, any set of dense orthogonal polynomials and associated standard random variable can be used to express a random variable. In general, the approximation of ε by a linear combination

converges to the true random variable as the number of elements in the linear combination approaches infinity; however, intelligently selecting the orthogonal polynomials can increase the rate of convergence[28]. The different selection of orthogonal polynomials to represent different types of random variables is the generalization from Wiener's homogeneous chaos to polynomial chaos (PC).

Let a stochastic variable, ε, be expressed over the polynomial vector space $\mathbb{P}^{\infty}$, which is spanned by the set of orthogonal polynomials $\mathbf{\Psi} = \{\Psi_0, \Psi_1, \ldots\}^*$, i.e.

$$\varepsilon(x, \xi) = \sum_{k=0}^{\infty} \epsilon_k(x)\Psi_k(\xi), \tag{3.7}$$

where $[\boldsymbol{\epsilon}]$ is the infinite dimensional coordinate (called chaos coefficients) of ε, x is redefined from Section 3.1 as a deterministic component (usually taken to be time), and $\Psi_k(\xi)$ is the k^{th}-order polynomial of Ψ that is evaluated at, and orthogonal with respect to, the standard weight function ξ. This is a polynomial chaos expansion of the random variable ε. The polynomial families listed in Table (2.1 have been shown by Xiu[28] to provide convenient orthogonal bases/weight function pairs based on the approximate pdf of ε.

Since the L^2 inner product can be used, the basis polynomials (which do not *have* to come from Table 2.1) satisfy the general orthogonality condition in Eq. (2.8).

3.2.1. Statistics. Using the coordinate, the useful information about ε including the moments and shape of ε can be estimated. For example, the mean (μ) and variance (σ^2) of the univariate state ε in Eq. (3.7) are

$$\mu = \epsilon_0 \qquad \text{and} \qquad \sigma^2 = \sum_{k=1}^{\infty} \epsilon_k^2 c_\psi,$$

respectively, where c_ψ is the scale factor associated with the orthogonality of the polynomial basis $\mathbf{\Psi}$.

*Discussion of the truncation follows in Eqs. (3.23)–(3.24) as well as in Section 3.2.4.

Now, let some multivariate state include the two elements a and b, which have independent expansions

$$a(x,\xi) = \sum_{j=0}^{\infty} \alpha_j(x)\Psi_j(\xi) \qquad \text{and} \qquad b(z,\zeta) = \sum_{k=0}^{\infty} \beta_k(z)\Phi_k(\zeta), \tag{3.8}$$

where $\mathbf{\Psi}$ and $\mathbf{\Phi}$ are two, possibly different, sets of basis polynomials, with standard random variables ξ and ζ, respectively. For now, x and z are kept independent for completeness. In this case, the expansion of the correlated system uses a set of multivariate basis polynomials that are a combination of $\mathbf{\Psi}$ and $\mathbf{\Phi}$. The orthogonality of the new multivariate basis can easily be proven since ξ and ζ remain independent. Beginning with the integral representation of an inner product

$$\left\langle \Psi_i\Phi_m, \Psi_j\Phi_n \right\rangle = \int_{\mathbb{X}}\int_{\mathbb{Z}} [\Psi_i(\xi)\Phi_m(\zeta)][\Psi_j(\xi)\Phi_n(\zeta)]p(\xi)p(\zeta)d\zeta d\xi\,,$$

it is apparent that each of the elements are with respect to only one of the integration variables; therefore, the double integral can be split into the multiplication of single integrals as

$$\left\langle \Psi_i\Phi_m, \Psi_j\Phi_n \right\rangle = \int_{\mathbb{X}} \Psi_i(z)\Psi_j(z)p(z)dz \int_{\mathbb{Z}} \Phi_m(\zeta)\Phi_n(\zeta)p(\zeta)d\zeta\,. \tag{3.9}$$

The orthogonality of both $\mathbf{\Phi}$ and $\mathbf{\Psi}$ reduce the integrals in Eq. (3.9) to

$$\left\langle \Psi_i\Phi_m, \Psi_j\Phi_n \right\rangle = c_\psi c_\phi \delta_{ij}\delta_{mn}\,, \tag{3.10}$$

where c_ψ and c_ϕ are the normalization factors associated with $\mathbf{\Psi}$ and $\mathbf{\Phi}$, respectively.

Let $\mathbf{\Gamma}$ be the multivariate orthogonal basis polynomials formed from the univariate bases $\mathbf{\Psi}$ and $\mathbf{\Phi}$. If the elements a and b from Eq. (3.8) are expressed using $\mathbf{\Gamma}$ as the basis polynomials, their expansions then become

$$a(x,z,\zeta) = \sum_{j=0}^{\infty} \alpha_j(x)\Gamma_j(z,\zeta) \qquad \text{and} \qquad b(z,z,\zeta) = \sum_{k=0}^{\infty} \beta_k(z)\Gamma_k(z,\zeta)\,. \tag{3.11}$$

In all cases where the weight function is a pdf, the zeroth polynomial of an orthogonal set is 1; therefore, the mean of a random variable expanded with a multivariate basis will still be the corresponding zeroth coefficient, i.e.

$$\begin{aligned} \boldsymbol{\mu} &= \mathrm{E}\left[\begin{bmatrix} a(x,z,\zeta) \\ b(z,z,\zeta) \end{bmatrix}\right] \\ &= \begin{bmatrix} \alpha_0(x) \\ \beta_0(z) \end{bmatrix}, \end{aligned} \tag{3.12}$$

where $\mathrm{E}[\cdot]$ denotes the expected value of the input. For the general n dimensional state, the mean is

$$\boldsymbol{\mu} = \boldsymbol{\epsilon}_0\,. \tag{3.13}$$

On the other hand, expanding the expression for the second moment is not quite as straight-forward. The covariance between a and b ($\sigma_{a,b}$) can be expressed in terms of nested expected values as

$$\sigma_{a,b} = \mathrm{E}[(a - \mathrm{E}[a])(b - \mathrm{E}[b])]p(a,b)\,,$$

where $p(a, b)$ is the joint pdf of a and b. The external expected value can be expressed as a double integral yielding

$$\sigma_{a,b} = \int_{\mathbb{A}} \int_{\mathbb{B}} (a - \mathrm{E}[a])(b - \mathrm{E}[b]) p(a, b) dbda, \tag{3.14}$$

where $\mathbb{A}$ and $\mathbb{B}$ are the supports of a and b, respectively. Expanding Eq. (3.14) as

$$\sigma_{a,b} = \int_{\mathbb{A}} \int_{\mathbb{B}} (ab - a\mathrm{E}[b] - b\mathrm{E}[a] + \mathrm{E}[a]\mathrm{E}[b]) p(a, b) dbda,$$

acknowledging the expected value is the zeroth coefficient, and breaking using the linearity of the integrals gives

$$\begin{aligned}\sigma_{a,b} = &\int_{\mathbb{A}} \int_{\mathbb{B}} abp(a, b) dbda - \int_{\mathbb{A}} \int_{\mathbb{B}} a\beta_0 p(a, b) dbda - \int_{\mathbb{A}} \int_{\mathbb{B}} b\alpha_0 p(a, b) dbda + \\ &+ \int_{\mathbb{A}} \int_{\mathbb{B}} \alpha_0 \beta_0 p(a, b) dbda.\end{aligned}$$

The standard random variables z and ζ are independent by design; therefore, the joint pdf can be broken into the product of two independent pdfs, yielding

$$\begin{aligned}\sigma_{a,b} = &\int_{\mathbb{A}} \int_{\mathbb{B}} abp(a)p(b) dbda - \int_{\mathbb{A}} \int_{\mathbb{B}} a\beta_0 p(a)p(b) dbda - \int_{\mathbb{A}} \int_{\mathbb{B}} b\alpha_0 p(a)p(b) dbda + \\ &+ \int_{\mathbb{A}} \int_{\mathbb{B}} \alpha_0 \beta_0 p(a)p(b) dbda.\end{aligned}$$

Since the zeroth coefficient is a constant value, at least one of the integrals can be removed in each of the terms, except the first resulting in

$$\sigma_{a,b} = \int_{\mathbb{A}} \int_{\mathbb{B}} abp(a)p(b) dbda - \beta_0 \int_{\mathbb{A}} ap(a) dbda - \alpha_0 \int_{\mathbb{B}} bp(b) dbda + \alpha_0 \beta_0.$$

The two single integrals reduce to the product of means, resulting in the expression

$$\sigma_{a,b} = -\alpha_0\beta_0 + \int_{\mathbb{A}}\int_{\mathbb{B}}(ab)p(a,b)dbda\,. \tag{3.15}$$

Substituting the expansions from Eq. (3.11) into Eq. (3.15) gives

$$\sigma_{a,b} = -\alpha_0\beta_0 + \int_{\mathbb{X}}\int_{\mathbb{Z}}\sum_{j=0}^{\infty}\alpha_j(x)\Gamma_j(z,\zeta)\sum_{k=0}^{\infty}\beta_k(z)\Gamma_k(z,\zeta)p(\xi)p(\zeta)d\zeta d\xi\,. \tag{3.16}$$

Recall that the Wiener chaos decomposition is performed on a Gaussian Hilbert space, as opposed to a Gaussian linear space, which means that the convergence of any sequence will be in the same space. It can therefore be concluded that the sequence in Eq. (3.16) converges to a finite number. With a finite covariance, the summation and the integrals can be interchanged[65], leading to

$$\sigma_{a,b} = \sum_{j=1}^{\infty}\sum_{k=1}^{\infty}\alpha_j(x)\beta_k(z)\int_{\mathbb{X}}\int_{\mathbb{Z}}\Gamma_j(\xi,\zeta)\Gamma_k(\xi,\zeta)p(\xi)p(\zeta)d\zeta d\xi\,. \tag{3.17}$$

From Eq. (3.10), the double integral in Eq. (3.17) reduces to a Kronecker delta scaled by c_ψ and c_ϕ, such that the final form of the covariance is

$$\sigma_{a,b} = \sum_{k=1}^{\infty}\alpha_k(x)\beta_k(z)c_\psi c_\phi\,. \tag{3.18}$$

Assuming the variance of a (σ_a^2) is desired instead, it follows that

$$\sigma_a^2 = \sum_{k=1}^{\infty}\alpha_k^2(x)c_\psi c_\phi\,, \tag{3.19}$$

which is the same as the univariate expansion, except for the additional normalization factor c_ϕ. A similar expression is given for σ_b^2 by replacing α_k with β_k resulting in

$$\sigma_b^2 = \sum_{k=1}^{\infty} \beta_k^2(x) c_\psi c_\phi \,. \tag{3.20}$$

Placing Eq. (3.19) in the upper diagonal, Eq. (3.20) in the lower diagonal, and Eq. (3.18) in both off-diagonals the covariance matrix of the 2×2 system of a and b is given as

$$\boldsymbol{P} = \sum_{k=1}^{\infty} \begin{bmatrix} \alpha_k^2 & \alpha_k\beta_k \\ \alpha_k\beta_k & \beta_k^2 \end{bmatrix} c_\psi c_\phi \,.$$

Letting the multivariate state be n-dimensional, and $\boldsymbol{\epsilon}$ be the $n \times \infty$ matrix for the n – theoretically infinite – chaos coefficients, the covariance matrix can be written generally as

$$\boldsymbol{P} = \sum_{k=1}^{\infty} \boldsymbol{\epsilon}_k \boldsymbol{\epsilon}_k^T c_\psi c_\phi \,. \tag{3.21}$$

In cases where orthonormal polynomials are used, the normalization factors disappear completely leaving only the summation of the estimated chaos coefficients

$$\boldsymbol{P} = \sum_{k=1}^{\infty} \boldsymbol{\epsilon}_k \boldsymbol{\epsilon}_k^T \,. \tag{3.22}$$

As is expected, expressions for higher order moments can be developed; however, the complexity of development, and of the expressions themselves, increases. For this reason only first and second order moments are of concern, but if higher order moments are of interest, the reader is strongly encouraged to reference Savin and Faverjon[27], which includes expressions for third and fourth order moments for Jacobi, Hermite, or Laguerre

chaos expansions. This work may only include first and second order moments for PCE like the UT; however, the chaos coefficients can still be used to approximate the *shape* of the pdf, not just the moments.

3.2.2. Coefficient Approximation - Least Squares. In Eqs. (3.7)–(3.22) the random variable ε is expressed as an infinite summation, which is impractical. It is therefore necessary to truncate the summation, yielding

$$\varepsilon(x,\xi) \approx \hat{\varepsilon}(x,\xi) = \sum_{k=0}^{N} \epsilon_k(x)\Psi_k(\xi)\,, \tag{3.23}$$

where the truncation term N, which depends on the dimension of the state n and the total expansion order p, is generally defined as

$$N + 1 = \frac{(n+p)!}{n!p!}, \tag{3.24}$$

where n is the dimension of the state, and p is the total expansion order. Section 3.2.4 provides additional details pertaining to the "total expansion order" as well as how it affects the number of terms in the expansion. For the remainder of Section 3.2.2, it is sufficient to understand that the truncation term N is some strictly positive ($N \geq 1$), finite integer.

In general, the chaos coefficients in Eq. (3.7) are solved for either with sampling-based or projection-based methods. The first, and most common, approach is performed by drawing Q samples of ξ, where $Q > N$, and evaluating Ψ_k and ε at these points, effectively randomly sampling ε. After initial sampling, ε can be transformed in x (commonly x is taken to be time, so this indicates propagating the variable forward in time) resulting in a system of Q equations with $N + 1$ unknown chaos coefficients. This system of equations

takes the form

$$\begin{bmatrix} \varepsilon(x,\xi_1) \\ \varepsilon(x,\xi_2) \\ \vdots \\ \varepsilon(x,\xi_Q) \end{bmatrix} = \begin{bmatrix} \Psi_0(\xi_1) & \Psi_2(\xi_1) & \cdots \Psi_N(\xi_1) \\ \Psi_0(\xi_2) & \Psi_2(\xi_2) & \cdots \Psi_N(\xi_2) \\ \vdots & & \\ \Psi_0(\xi_Q) & \Psi_2(\xi_Q) & \cdots \Psi_N(\xi_Q) \end{bmatrix} \begin{bmatrix} \epsilon_1(x) \\ \epsilon_2(x) \\ \vdots \\ \epsilon_N(x) \end{bmatrix}, \tag{3.25}$$

where $[\epsilon]$ is unknown. This overdetermined system can be solved to produce an estimate of $\boldsymbol{\epsilon}$ using a least-squares solution, i.e.

$$[\epsilon] = \left([\Psi]^*[\Psi]\right)^{-1}[\Psi]^*[\varepsilon], \tag{3.26}$$

where $[\Psi]$ and $[\varepsilon]$ are the matrices in Eq. (3.25) populated by those elements, and $[\cdot]^*$ refers to the conjugate transpose. The coefficients can then be used to approximate the random variable or its moments, e.g. the mean and covariance from Eqs. (3.12) and (3.22), respectively.

3.2.2.1. Sampling methods. The sampling of the standard random variable ξ in Eq. (3.25) can be performed in a number of different ways. The simplest of sampling methods is implemented by numerically generating random[†] numbers, where the probability of a given sample is reliant only on the random variable's pdf. While random samples are generally the easiest to generate, there is no assurance that the samples will adequately represent the pdf.

Alternatively, samples can be drawn deterministically such that some elements of the pdf are fully characterized; for example, the placement of the UT sigma points in Eq. (3.2) and selection of their weights in Eq. (3.5) fully capture the given mean and covariance. While deterministically choosing samples can better represent the pdf, computation can become expensive when many samples or higher order moments are required.

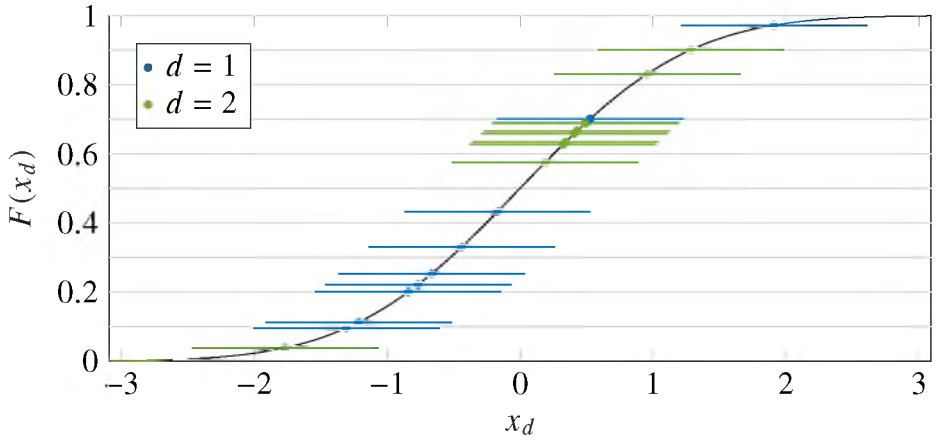
[†]It is noted that computer generated samples are pseudo-random, not truly random; however, for convenience, computer generated random samples are simply referred to as random samples.

It should be noted that any linearly dependent expansion samples provide no additional information to the least squares approximation in Eq. (3.25). Beyond providing no new information, linearly dependent samples increase computation, if there are not at least N linearly independent samples, the matrix inverse in Eq. (3.26) can become singular.

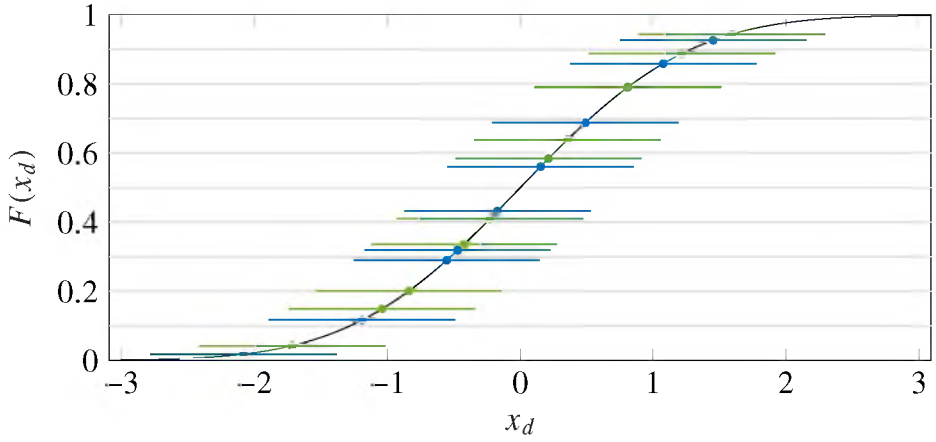
As is expected, there are numerous sampling methods that combine deterministic and random sampling techniques such as Latin hypercube[66] (LHC) and Hammersley[67] sampling. As opposed to random sampling, LHC and Hammersley sampling break the cdf of the random variable into equal segments and create a sample within each segment, which can be seen for 10, two dimensional (each in different colors) samples from the standard normal pdf in Figures (3.2a)–(3.2c). These figures contain the standard normal cdf and the uniform samples on the interval [0, 1] that are used to produce normally distributed samples. The horizontal lines are included to better indicate the positioning of each sample within the cdf segments. It is important to note that these are not whisker plots nor are they meant to show a range of x values for a given $F(x)$.

In Figure (3.2a) it is apparent that the samples from one variable predominately come from the top half of the cdf, and the other variable samples predominately come from the bottom, which produces an incomplete representation of the pdf. Obviously, the clustering in Figure (3.2a) is the result of a singular instance of sample generation, and there is no guarantee of repeatability. The likelihood of drawing all 10 samples exactly uniformly is the same as any other possible spread of samples.

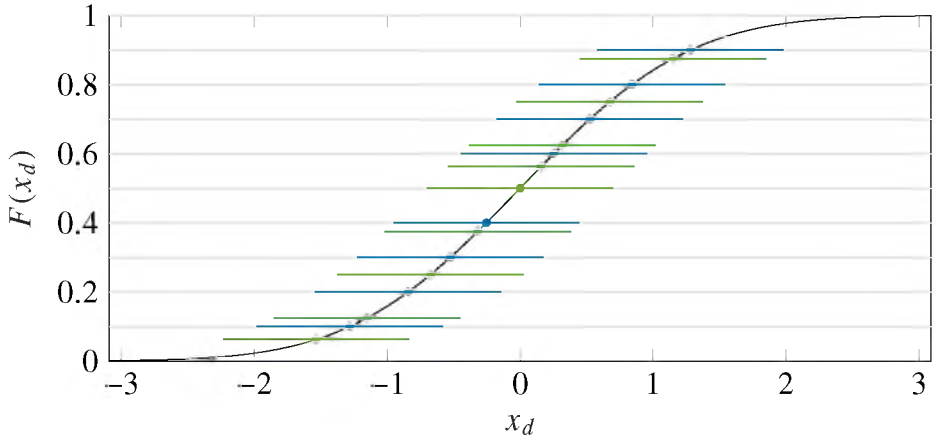
Contrarily, the LHC and Hammersley samples in Figures (3.2b) and (3.2c), respectively, more completely sample the cdf. LHC sampling randomly samples each dimension once in every segment while Hammersley deterministically sets every sample such that the first dimension's samples are the endpoints of each segment, and the remaining dimensions' samples are set based on a van der Corput sequence[68]. The resulting bivariate normal samples for each sampling method are shown in Figures (3.3a)–(3.3c).



(a) Random samples on the cdf.



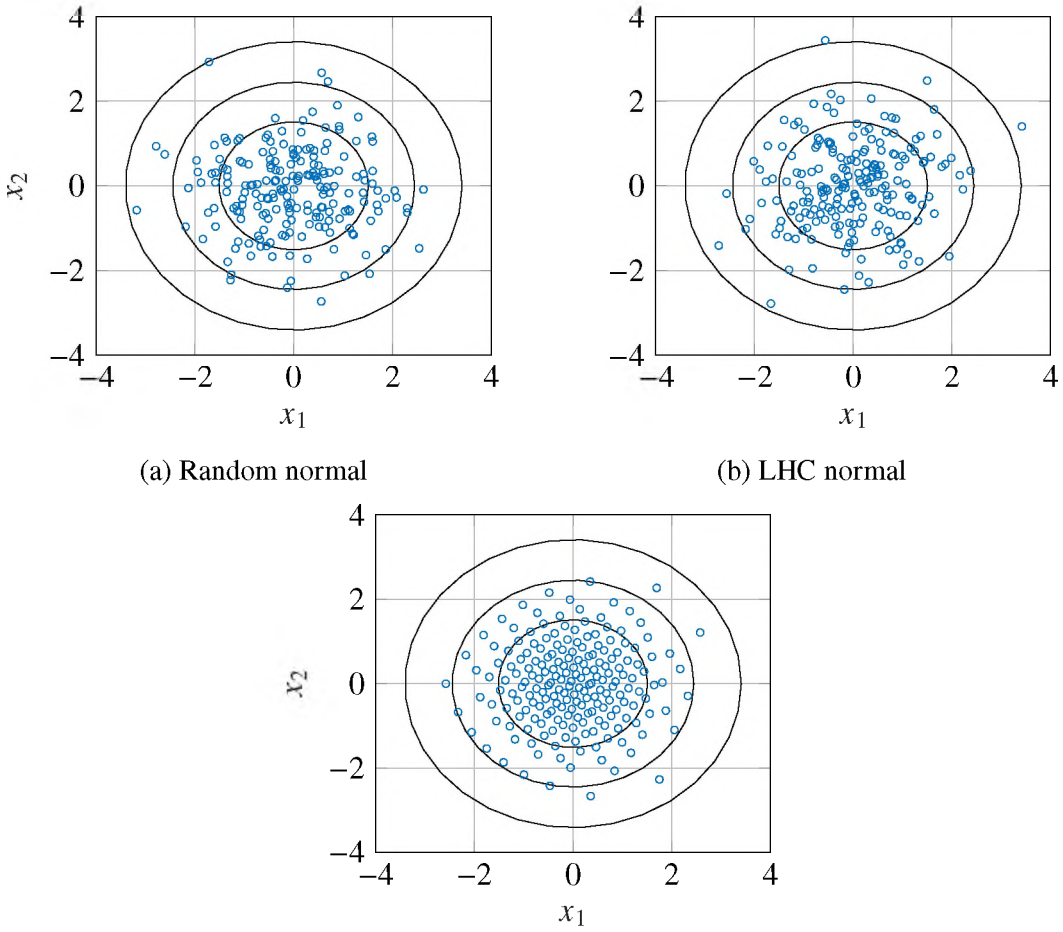
(b) LHC samples on the cdf.



(c) Hammersley samples on the cdf.

Figure 3.2. Uniform cdf sampling.

The random and LHC sample (Figures (3.3a) and (3.3b)) sets both give the appearance of random sets; however, the deterministic quality of the Hammersley technique is immediately apparent in Figure (3.3c). Ultimately, the choice of sampling technique is problem specific depending on how the samples are used.



(a) Random normal

(b) LHC normal

(c) Hammersley normal

Figure 3.3. Bivariate normally distributed samples with Gaussian 1, 2, and 3 – σ intervals.

3.2.2.2. Implementation procedure. For convenience, the procedure for estimating the mean and covariance of a random state using the coefficient approximation method in Section 3.2.2 is given in Algorithm 1. Let $\boldsymbol{\varepsilon}$ be the state of a system with initial uncertainty of any type defined by mean $\boldsymbol{m}_0$ and covariance $\boldsymbol{P}_0$ subject to a set of system dynamics $f(\cdot)$ over the time vector T. The algorithm outlines the steps required to estimate the state (by

chaos coefficients, mean, and covariance) at the times in T. It should be noted that it is possible to represent any uncertainty in the governing dynamics ($\boldsymbol{f}$) with PCE; however, it is assumed the dynamics are perfectly known.

Algorithm 1 Estimation of mean and covariance using a polynomial chaos expansion.

1: **procedure** PCE_EST($\boldsymbol{m}_0, \boldsymbol{P}_0$) ▷ State estimation using PCE with deterministically known system dynamics.
2: **for** $k = 1$ **to** T **do**
3: Draw samples of $\boldsymbol{\xi}$ based on sampling type ▷ Refer to Section 3.2.2.1
4: $[\boldsymbol{\Psi}] \leftarrow \boldsymbol{\Psi}(\boldsymbol{\xi})$
5: $\boldsymbol{\varepsilon}_{k-1} \leftarrow \boldsymbol{m}_{k-1} + \sqrt{\boldsymbol{P}_{k-1}}\xi$
6: $\boldsymbol{\varepsilon}_k \leftarrow \boldsymbol{f}(\boldsymbol{\varepsilon}_{k-1})$
7: $\boldsymbol{\epsilon}_k \leftarrow$Eq. (3.26)
8: $\boldsymbol{m}_k \leftarrow$Eq. (3.13)
9: $\boldsymbol{P}_k \leftarrow$Eq. (3.21) ▷ Summation truncated by user
10: **end for**
11: **return** $\boldsymbol{\epsilon}$, $\boldsymbol{m}$, $\boldsymbol{P}$
12: **end procedure**

This implementation is similar to that of the UT, specifically lines 3, 6, 8, and 9. In both cases, samples are obtained, transformed, and then used to estimate the mean and covariance at some future point in time. In general, the matrix inverse used in the least squares on line 6 is computed using methods that are computationally negligible when compared with the dynamics in line 6. Any additional run time between the UT and PCE is therefore driven by the number of UT sigma points compared with the number PCE samples. It should again be noted that while PCE may be more computationally expensive, estimation of the state is not restricted to the moments.

3.2.3. Coefficient Solution - Inner Product. While the sampling-based method is more practical to apply, the projection based method does not depend on sampling the underlying distribution. Recall the projection of the point $\boldsymbol{A}$ onto the $\boldsymbol{x}$ and $\boldsymbol{y}$ directions in

Figure (2.2). These projections can be used to write the point $\boldsymbol{A}$ as the linear combination of the bases $\boldsymbol{x}$ and $\boldsymbol{y}$ as

$$\boldsymbol{A} = \frac{\langle \boldsymbol{A}, \boldsymbol{x} \rangle}{\langle \boldsymbol{x}, \boldsymbol{x} \rangle} \boldsymbol{x} + \frac{\langle \boldsymbol{A}, \boldsymbol{y} \rangle}{\langle \boldsymbol{y}, \boldsymbol{y} \rangle} \boldsymbol{y} \,,$$

which strongly resembles the Gram-Schmidt orthogonalization in Section 2.2.3.2. In both cases, the normalized projection helps isolate the portion of an element (in this case $\boldsymbol{A}$) that lies in a certain direction. In this way, the an expression for the coordinates of $\boldsymbol{A}$ can be found. The same process can be done to solve for the chaos coefficients in the polynomial chaos expansion.

To find the j^{th} coefficient, project the chaos expansion onto the j^{th} basis yielding

$$\left\langle \varepsilon(x,\xi), \Psi_j(\xi) \right\rangle_{p(\xi)} = \left\langle \sum_{k=0}^{\infty} \epsilon_k(x) \Psi_k(\xi), \Psi_j(\xi) \right\rangle_{p(\xi)} .$$

Utilizing the property of first argument linearity from Eq. (2.2), the summation in the first argument can be removed from the inner product, resulting in

$$\left\langle \varepsilon(x,\xi), \Psi_j(\xi) \right\rangle_{p(\xi)} = \sum_{k=0}^{\infty} \left\langle \epsilon_k(x) \Psi_k(\xi), \Psi_j(\xi) \right\rangle_{p(\xi)} .$$

Additionally, the inner product is with respect to the variable ξ; therefore, the coefficient ϵ acts as a scalar and again by Eq. (2.2), can be removed from the inner product giving

$$\left\langle \varepsilon(x,\xi), \Psi_j(\xi) \right\rangle_{p(\xi)} = \sum_{k=0}^{\infty} \epsilon_k(x) \left\langle \Psi_k(\xi), \Psi_j(\xi) \right\rangle_{p(\xi)} . \tag{3.27}$$

Consider, for the sake of argument, a similar inner product that has the summation in the second argument, i.e.

$$\left\langle X_j(\xi), \sum_{k=0}^{\infty} b_k(x) X_k(\xi) \right\rangle_{p(\xi)} .$$

In this case, the summation can still be brought out of the inner product in the same way, but the order of the inner product has to be flipped before the coefficient can be removed. From the conjugate symmetry property in Eq. (2.3), the flipped inner product takes the form

$$\sum_{k=0}^{\infty} \left\langle \overline{b_k(x) X_k(\xi)}, \overline{X_j(\xi)} \right\rangle_{p(\xi)} .$$

Now that the coefficient is in the first argument it can be removed from the inner product resulting in

$$\sum_{k=0}^{\infty} \overline{b_k(x)} \left\langle \overline{X_k(\xi)}, \overline{X_j(\xi)} \right\rangle_{p(\xi)} .$$

When all the values are real, the complex conjugate has no affect; but, it is important to recognize the underlying mathematical operations.

Returning to Eq. (3.27) and expanding the summation gives

$$\left\langle \varepsilon(x,\xi), \Psi_j(\xi) \right\rangle_{p(\xi)} = \epsilon_0(x) \left\langle \Psi_0(\xi), \Psi_j(\xi) \right\rangle_{p(\xi)} + \cdots \epsilon_j(x) \left\langle \Psi_j(\xi), \Psi_j(\xi) \right\rangle_{p(\xi)} + \cdots .$$

Recall the basis polynomials are specifically chosen to be orthogonal, so the only inner product on the right-hand side that remains non-zero is the j^{th} one, resulting in

$$\left\langle \varepsilon(x,\xi), \Psi_j(\xi) \right\rangle_{p(\xi)} = \epsilon_k(x) \left\langle \Psi_j(\xi), \Psi_j(\xi) \right\rangle_{p(\xi)} ,$$

which can easily be arranged according to

$$\epsilon_j(x) = \frac{\left\langle \varepsilon(x,\xi), \Psi_j(\xi) \right\rangle_{p(\xi)}}{\langle \Psi_j^2(\xi) \rangle_{p(\xi)}}, \tag{3.28}$$

isolating $\epsilon_j(x)$.

Systematically projecting ε onto each orthogonal basis yields similar expressions for each coefficient such that the full coordinate takes the form

$$[\epsilon] = \begin{bmatrix} \frac{\langle \varepsilon(x,\xi), \Psi_0(\xi) \rangle_{p(\xi)}}{\langle \Psi_0^2(\xi) \rangle_{p(\xi)}} \\ \frac{\langle \varepsilon(x,\xi), \Psi_1(\xi) \rangle_{p(\xi)}}{\langle \Psi_1^2(\xi) \rangle_{p(\xi)}} \\ \vdots \\ \frac{\langle \varepsilon(x,\xi), \Psi_j(\xi) \rangle_{p(\xi)}}{\langle \Psi_j^2(\xi) \rangle_{p(\xi)}} \\ \vdots \end{bmatrix} . \tag{3.29}$$

These coefficients can then be used to calculate moments according to Section 3.2.1.

It should be recalled that the inner products in Eq. (3.29) represent integrals of the form

$$\epsilon_j(x) = \frac{\int_{\mathbb{X}} \varepsilon(x,\xi)\Psi_j(\xi)dp(\xi)}{\int_{\mathbb{X}} \Psi_j^2(\xi)dp(\xi)} . \tag{3.30}$$

where $\mathbb{X}$ is the support of ξ. The integral in the denominator is usually straightforward, and is even given analytically for many common families of orthogonal polynomials, but the integral in the numerator is more complicated. The only instances where the integral in the numerator may have an analytic form is when ε is known analytically. For the purposes of state estimation, perfect knowledge of ε would not require a method of uncertainty quantification such as PCE. It is therefore assumed that the integral

$$\int_{\mathbb{Z}} \varepsilon(x,\xi)\Psi_j(\xi)dp(\xi) ,$$

must be approximated using any of the numerous methods numerical integration methods such as quadrature or cubature[69].

3.2.4. Expansion Truncation. In the event the random variable is fully characterized by a PCE with a finite number of terms, the truncation in Eq. (3.23) can impose no approximation error. Consider, for example, the initialization a Gaussian random variable x,

$$x(\xi) = \mu + \sigma\xi, \tag{3.31}$$

where μ is the initial mean, σ is the initial standard deviation, and ξ is the standard normal random variable. Now, consider the Hermite polynomials

$$\begin{aligned} H_0(\xi) &= 1 \\ H_1(\xi) &= \xi \\ H_2(\xi) &= \xi^2 - 1 \\ &\vdots \\ H_n(\xi) &= (-1)^n e^{\xi^2/2} \frac{d^n}{d\xi} e^{-\xi^2/2} \qquad n = 3, 4, \ldots . \end{aligned}$$

Writing x as a linear combination of Hermite polynomials gives

$$x(\xi) = \mu H_0(\xi) + \sigma H_1(\xi) + 0H_2(\xi) + \cdots + 0H_n + \cdots . \tag{3.32}$$

Collecting the coefficients results in a coordinate of the form

$$[\epsilon] = \begin{bmatrix} \mu \\ \sigma \\ 0 \\ \vdots \end{bmatrix}, \tag{3.33}$$

which has infinite dimension, but only two nonzero coefficients. In general, a finite PCE is not possible, and Eq. (3.23) is an approximation with error ($\tilde{\epsilon}$) given by

$$\tilde{\epsilon}(x,\xi) = \sum_{k=N+1}^{\infty} \epsilon_k(x)\Psi_k(\xi),$$

which contains all elements of Eq. (3.7) with orders greater than the truncation.

As long as the orthogonal polynomials $\mathbf{\Psi}$ are dense in L^2, then the summation in Eq. (3.7) converges to the true random variable. For rates of convergence, the reader is encouraged to refer to Xiu[28] and Ernst[70]. In relative terms, convergence rates increase as the standard random variable in the expansion approaches the random variable. For instance, the number of terms required to represent the random variable in Eq. (3.31) using any basis polynomials other than the Hermite polynomials will have more than 2 coefficients.

Recall the truncation term N in Eq. (3.24), and its dependence on the dimension n and is the total expansion order p. The total expansion order refers to the maximum order of all multivariate basis polynomials. The number of combinations of n univariate polynomials, whose orders add to p, or number of n-tuples that add to p including zero is known as the stars and bars combinatorial problem[71]. When 0 is included in the selection set (since polynomials of order 0 are allowed), the number of n-tuples that sum to p (N_p) is

$$N_p = \binom{n+p-1}{n-1}.$$

Combining this with the number of n-tuples that sum to $p-1, p-2, \ldots, 0$ is

$$N_{p,p-1,\ldots,0} = \sum_{k=0}^{p} \binom{n+k-1}{n-1},$$

which has the form of the hockey-stick identity[72],

$$\sum_{k=b}^{c} \binom{a}{b} = \binom{c+1}{b+1},$$

for positive integers and $c > b$. Using this identity produces the truncation term N from Eq. (3.24),

$$N + 1 = \frac{(n+p)!}{n!p!},$$

where the left-hand side is because the polynomial chaos expansion begins with a zeroth term.

Let a 2-dimensional state be expanded with respect to ξ and ζ such that the univariate polynomial bases are $\mathbf{\Psi}$ and $\mathbf{\Phi}$. The resulting multivariate basis for a p^{th} total order expansion is constructed according to the products in Table 3.1, which clearly that indicates no multivariate bases are constructed that have a total order greater than p. The partiality of the tensor product is evident in the empty lower triangle of Table 3.1.

Table 3.1. Traditional construction of multivariate basis polynomials $\mathbf{\Gamma}$ from univariate orthogonal polynomials $\mathbf{\Psi}$ and $\mathbf{\Phi}$ by total expansion order.

	Ψ_0	Ψ_1	$\cdots$	Ψ_{p-1}	Ψ_p
Φ_0	$\Psi_0\Phi_0$	$\Psi_1\Phi_0$	$\cdots$	$\Psi_{p-1}\Phi_0$	$\Psi_p\Phi_0$
Φ_1	$\Psi_0\Phi_1$	$\Psi_1\Phi_1$	$\cdots$	$\Psi_{p-1}\Phi_1$	
$\vdots$	$\vdots$	$\vdots$	$\ddots$		
Φ_{p-1}	$\Psi_0\Phi_{p-1}$	$\Psi_1\Phi_{p-1}$			
Φ_p	$\Psi_0\Phi_p$				

For this type of truncation, all univariate expansions have the same order. As was discussed, the rate of convergence of the PCE increases as the random variable diverges from the standard random variable. In these cases, a higher PCE expansion order is required to produce a more accurate approximation of the random variable. But, if one variable requires a larger expansion, the multivariate basis with terms according to Eq. (3.24) expands *every* element of the state. In other words, if within an n-dimensional state expanded using Hermite-chaos has one state that is very non-Gaussian requiring a p^{th} order expansion and $n - 1$ dimensions requiring k^{th} order expansions such that p is significantly larger

than k, the multivariate basis with Eq. (3.24) terms expands every element with a p^{th} order expansion. The additional expansion in the majority of states is not necessary, and increases the computation burden of PCE.

Instead of representing every element of the state with a p^{th} order expansion, allowing each state to be expanded to a unique expansion order reduces the total number of terms in the multivariate expansion, and ensures computation resources are efficiently applied to the states that require larger expansions. For an n-dimensional state, if no total expansion order is defined and the individual expansion orders are $\{p_1, \ldots, p_n\}$, the "full tensor" truncation term becomes

$$N + 1 = \prod_{j=1}^{n} (p_j + 1) \qquad j = 1, \ldots, n\,. \tag{3.34}$$

In the case $p_j = p \; \forall j$, the number of terms is simply

$$N + 1 = (p + 1)^n\,.$$

For convenience, the multivariate bases resulting from a full tensor expansion in 2 dimensions are collected in Table 3.2.

Table 3.2. Full tensor construction of multivariate basis polynomials $\mathbf{\Gamma}$ from univariate orthogonal polynomials $\mathbf{\Psi}$ and $\mathbf{\Phi}$ by individual expansion orders.

	Ψ_0	Ψ_1	$\cdots$	Ψ_{p_1-1}	Ψ_{p_1}
Φ_0	$\Psi_0\Phi_0$	$\Psi_1\Phi_0$	$\cdots$	$\Psi_{p_1-1}\Phi_0$	$\Psi_{p_1}\Phi_0$
Φ_1	$\Psi_0\Phi_1$	$\Psi_1\Phi_1$	$\cdots$	$\Psi_{p_1-1}\Phi_1$	$\Psi_{p_1}\Phi_1$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
Φ_{p_2-1}	$\Psi_0\Phi_{p_2-1}$	$\Psi_1\Phi_{p_2-1}$	$\cdots$	$\Psi_{p_1-1}\Phi_{p_2-1}$	$\Psi_{p_1}\Phi_{p_2-1}$
Φ_{p_2}	$\Psi_0\Phi_{p_2}$	$\Psi_1\Phi_{p_2}$	$\cdots$	$\Psi_{p_1-1}\Phi_{p_2}$	$\Psi_{p_1}\Phi_{p_2}$

Additionally, each dimension can have a unique expansion order *and* a total expansion order can be defined which would result in yet another truncation factor. For example, the same 2-dimensional basis with individual expansion orders of p_1 and p_2 and a total expansion order of $p < p_1 + p_2$ in Table 3.3.

Expansions of this type are not considered in this work.

Table 3.3. Partial tensor construction of multivariate basis polynomials $\mathbf{\Gamma}$ from univariate orthogonal polynomials $\mathbf{\Psi}$ and $\mathbf{\Phi}$ by individual expansion orders with a total expansion order.

	Ψ_0	Ψ_1	$\cdots$	Ψ_{p_1-1}	Ψ_{p_1}
Φ_0	$\Psi_0\Phi_0$	$\Psi_1\Phi_0$	$\cdots$	$\Psi_{p_1-1}\Phi_0$	$\Psi_{p_1}\Phi_0$
Φ_1	$\Psi_0\Phi_1$	$\Psi_1\Phi_1$	$\cdots$	$\Psi_{p_1-1}\Phi_1$	$\Psi_{p_1}\Phi_1$
$\vdots$	$\vdots$	$\vdots$			$\vdots$
Φ_{p_2-1}	$\Psi_0\Phi_{p_2-1}$	$\Psi_1\Phi_{p_2-1}$			$\Psi_{p_1}\Phi_{p-p_1}$
Φ_{p_2}	$\Psi_0\Phi_{p_2}$	$\Psi_1\Phi_{p_2}$	$\cdots$	$\Psi_{p-p_2}\Phi_{p_2}$	

3.2.5. Coefficient Analysis. Inspection of the chaos coefficients can lead to a better understanding of how the uncertainty of one random variable affects another random variable. Again, let a and b be two random variables in a state with coefficient order according to Table 3.4.

Inspection of the zeroth coefficient does not provide much insight beyond translation of the mean. The remaining coefficients, on the other hand, can be used to estimate higher order moments as well as indicate divergence from the standard random variable in the expansion and interaction between the variables.

Table 3.4. Example coefficient ordering of a bivariate, full tensor, expansion.

$\mathbf{\Psi}(\xi)$	Ψ_0	Ψ_1	$\cdots$	Ψ_0	Ψ_1	$\cdots$	Ψ_0	Ψ_1	$\cdots$	Ψ_{p_1}
$\mathbf{\Phi}(\zeta)$	Φ_0	Φ_0	$\cdots$	Φ_1	Φ_1	$\cdots$	Φ_k	Φ_k	$\cdots$	Φ_{p_2}
$a(x,\xi,\zeta)$	α_0	α_1	$\cdots$	α_{p_1}	α_{p_1+1}	$\cdots$	α_{kp_1}	α_{kp_1+1}	$\cdots$	$\alpha_{p_1\cdot p_2-1}$
$b(y,\xi,\zeta)$	β_0	β_1	$\cdots$	β_{p_1}	β_{p_1+1}	$\cdots$	β_{kp_1}	β_{kp_1+1}	$\cdots$	$\beta_{p_1\cdot p_2-1}$
	$\boldsymbol{\mu}$	$\sigma_{a,b}$								

For example, take the pair of coefficients α_k and β_k. The covariance between a and b is influenced by these two coefficients only if both α_k and β_k are nonzero. If α_k is zero, and β_k is nonzero, β_k contributes to σ_b^2 but provides no impact on a at all, which holds for if β_k is zero and α_k is nonzero.

Of specific consideration are the multivariate polynomials of degree 1. In these cases, only one of the polynomials is a function of its standard random variable. The coefficients α_{kp_1} and β_{kp_1} correspond to the basis $\Psi_0\Phi_k$, which is equivalent to the univariate polynomial Φ_k: implying α_{kp_1} and β_{kp_1} are coordinates with respect to *only* the standard random variable ζ and not ξ. The opposite is true for the multivariate basis $\Psi_k\Phi_0$. As a result, the coefficients corresponding to first order, multivariate bases indicate the portion of the random variable (or its moments beyond the first) that is attributed to the corresponding standard random variable, i.e. α_{kp_1} and β_{kp_1} provide an indication as to the isolated effects of ζ on a and b. Again, this applies to all univariate bases.

Additionally, observing the chaos coefficient rate of convergence can indicate the relative divergence or convergence of the random variable from the standard random variable. Consider, for example, the Gaussian random variable expanded using Hermite-chaos resulting in the chaos coefficients in Eq. (3.33). If, after a transformation, the first two chaos coefficients remain the only two coefficients that are nonzero, it can be concluded the random variable is still Gaussian, without additional analysis or computation. Conversely, as the other chaos coefficients diverge from zero, the random variable becomes less Gaussian. It should be noted that this is a very high-level analysis, and while it may be possible to define a divergence metric, such as the Kullback-Leibler[73] or Hellinger[74] distances.

In many instances, the dynamics governing a state are not known. Analyzing the chaos coefficients may not provide the user with the means of reverse engineering the dynamics, but it can provide additional insight into how the different states interact beyond the covariance or implementation factors including if certain variables that require higher levels of accuracy.

4. CHAOS ON THE UNIT CIRCLE

4.1. COMPLEX EXPANSION

A PCE is effectively the projection of the $L^2(\Omega, \mathcal{F}, P)$ vector space onto an infinite set of orthogonal polynomials. When a complex space is projected onto a set of complex polynomials, the basis can either be the original set of polynomials or their complex conjugates. To maintain a distinction between PCE for real polynomials (which is denoted as $\text{PCE}_{\mathbb{R}}$) and PCE for complex polynomials (which is denoted as $\text{PCE}_{\mathbb{C}}$), the complex conjugate of the basis polynomials is used in the chaos expansion. Obviously, this notation holds for $\text{PCE}_{\mathbb{R}}$ since $\overline{\mathbb{R}} = \mathbb{R}$. Using the complex conjugate of the set of polynomials, the familiar expansion in Eq. (3.7) becomes

$$\varepsilon(x, z) = \sum_{k=0}^{\infty} \epsilon_k(x)\overline{\Psi_k(z)}, \tag{4.1}$$

where, once again, $z = e^{i\theta}$, ε is the random variable, ϵ_k is the k^{th} chaos coefficient, Ψ_k is the k^{th} order basis polynomial, and $\overline{\Psi_k(z)}$ indicates a complex conjugate. Because the inner product naturally conjugates the second element, the inner product representation of each chaos coefficient from Eq. (3.28) does not change, but a complex conjugate is introduced into the integral form in Eq. (3.30). The resulting expressions for the chaos coefficients

when the basis polynomials are complex are given as

$$\epsilon_j(x) = \frac{\left\langle \varepsilon(x,z), \Psi_j(z) \right\rangle_{p(z)}}{\langle \Psi_j^2(z) \rangle_{p(z)}} \tag{4.2a}$$

$$= \frac{\int_{\mathbb{Z}} \varepsilon(x,z)\overline{\Psi_j(z)}dp(z)}{\int_{\mathbb{Z}} \Psi_j(z)\overline{\Psi_j(z)}dp(z)} \,. \tag{4.2b}$$

4.1.1. Statistics. Taking an approach similar to the one taken in Section 3.2.1, expressions for the first two moments can be developed for a multivariate state that includes the two elements α and β, which have independent expansions

$$a(x,z) = \sum_{j=0}^{\infty} \alpha_j(x)\overline{\Psi_j(z)} \qquad \text{and} \qquad b(y,\varsigma) = \sum_{k=0}^{\infty} \beta_k(y)\overline{\Phi_k(\varsigma)}, \tag{4.3}$$

where $\mathbf{\Psi}$ and $\mathbf{\Phi}$ are two, possibly different, sets of complex basis polynomials, with standard random variables z and ς, respectively. Note the change in notation from Eq. (3.8) to Eq. (4.3). This is done to ensure the variable z is consistent with the $e^{i\theta}$ notation.

The integral expression for the mean of a takes the form

$$\mathrm{E}[a] = \int_{\mathbb{X}} a(x,z)p(z)dz \,,$$

where a can be replaced with its expansion giving

$$\mathrm{E}[a] = \int_{\mathbb{X}} \sum_{k=0}^{\infty} \alpha_k(x)\overline{\Psi_j(z)}p(z)dz \,. \tag{4.4}$$

The zeroth polynomial of the Szegő polynomials is 1, just like the Askey polynomials, so an equivalent statement to Eq. (4.4) can be written that includes the zeroth polynomial. Including the zeroth polynomial and moving the summation/coefficient pair outside of the

integral gives the expression

$$\mathrm{E}[a] = \sum_{k=0}^{\infty} \alpha_k(x) \int_{\mathbb{X}} \overline{\Psi_k(z)} \Psi_0(z) p(z) dz \,,$$

which can be written as the inner product

$$\mathrm{E}[a] = \sum_{k=0}^{\infty} \alpha_k(x) \left\langle \overline{\Psi_k(z)}, \overline{\Psi_0(z)} \right\rangle \,. \tag{4.5}$$

The inner product in Eq. (4.5) represents the orthogonality condition of the polynomials $\mathbf{\Psi}$, and as a result, when $k \neq 0$, the inner products evaluates to 0, and when $k = 0$, the inner product evaluates to 1 (because the pdf of z evaluates to 1). Therefore,

$$\mathrm{E}[a] = \sum_{k=0}^{\infty} \alpha_k(x) \left\langle \overline{\Psi_k(z)}, \overline{\Psi_0(z)} \right\rangle \,, \tag{4.6}$$

when the polynomials are complex.

Even if the mean is written according to

$$\mathrm{E}[a] = \sum_{k=0}^{\infty} \alpha_k(x) \left\langle \overline{\Psi_k(z)}, \Psi_0(z) \right\rangle \,, \tag{4.7}$$

the result is the same because

$$\Psi_0(z) \equiv \overline{\Psi_0(z)} \,.$$

The multivariate first moment is the same for complex random variables as it is for real random variables in Eq. (4.8). The first moment of the multivariate system including complex random variables a and b is therefore,

$$\mathrm{E}\left[\begin{bmatrix} a(x, y\xi) \\ b(x, y, \varsigma) \end{bmatrix}\right] = \begin{bmatrix} \alpha_0(x) \\ \beta_0(z) \end{bmatrix} \,. \tag{4.8}$$

Unfortunately, the orthogonality of the polynomials utilized in the simplification from Eq. (3.17) to Eq. (3.18) is not valid for complex valued polynomials, resulting in more complicated expressions for the raw second moments as

$$\mathrm{E}[a^2] = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \alpha_j \alpha_k \int_{\mathbb{X}} \overline{\Psi_j(z)\Psi_k(z)} p(z) dz \tag{4.9}$$

$$\mathrm{E}[ab] = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \alpha_j \beta_k \int_{\mathbb{X}}\int_{\mathbb{Y}} \overline{\Psi_j(\varsigma)\Phi_k(z)} p(z,\varsigma) d\varsigma dz\,, \tag{4.10}$$

where the expression in Eq. (4.9) is similar for b when the appropriate chaos coefficients (β) and basis polynomials ($\mathbf{\Phi}$) are used. If z and ς are the same standard random variable, then one of the integrals in Eq. (4.10) is eliminated yielding

$$\begin{aligned}\mathrm{E}[ab] &= \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \alpha_j \beta_k \int_{\mathbb{X}} \overline{\Psi_j(z)\Psi_k(z)} p(z) dz \\ &= \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \alpha_j \beta_k \langle \overline{\Psi_j(z)}, \Psi_k(z) \rangle_{p(z)}\,.\end{aligned} \tag{4.11}$$

Consider using a standard random variable that does not have the form $e^{i\theta}$. This results in a real valued standard random variable and real valued orthogonal basis polynomials. In this case Eqs. (4.9) and (4.10) become

$$\mathrm{E}[a^2] = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \alpha_j \alpha_k \int_{\mathbb{X}} \Psi_j(z)\Psi_k(z) p(z) dz \tag{4.12}$$

$$\mathrm{E}[ab] = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \alpha_j \beta_k \int_{\mathbb{X}}\int_{\mathbb{Y}} \Psi_j(\varsigma)\Phi_k(z) p(z,\varsigma) d\varsigma dz\,. \tag{4.13}$$

Now, the orthogonality of real valued polynomials can be utilized. Additionally, beginning the summations at one instead of zero (i.e. the difference between a central moment and a raw moment) results in the equations for the variance and covariance in Eqs. (3.18) and (3.19)

from Section (3.2.1). It is implied that Eqs. (4.9) and (4.10) are the general form of the moment equations, because Eqs. (3.18) and (3.19) can be derived from Eqs. (4.9) and (4.10) after making the specification that real valued polynomials and random variables are used.

4.1.2. Rogers-Szegő-Chaos. The Rogers-Szegő polynomials and the wrapped normal distribution provide a convenient basis and random variable pairing for the linear combination in Eq. (4.1). The Rogers-Szegő polynomials in Eq. (2.11) can be rewritten according to[46]

$$\phi_n\left(-\frac{z}{\sqrt{q}},q\right)=\sum_{k=0}^{n}(-1)^{n-k}\binom{n}{k}_q q^{\frac{n-k}{2}}z^k, \tag{4.14}$$

where q is calculated based on the standard deviation of the unwrapped normal distribution as $q=e^{-\sigma^2}$. These polynomials satisfy the orthogonality condition

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\phi_m\left(-\frac{z}{\sqrt{q}},q\right)\overline{\phi_n\left(-\frac{z}{\sqrt{q}},q\right)}\vartheta_3\left(\frac{\theta}{2},\sqrt{q}\right)d\theta=\frac{(q;q)_n}{q^n}\delta_{mn}, \tag{4.15}$$

where $\vartheta_3(\alpha,\beta)$ is the theta function

$$\vartheta_3(\alpha,\beta)=\sum_{k=-\infty}^{\infty}\beta^{k^2}e^{2ik\alpha}, \tag{4.16}$$

which is another form of the wrapped normal distribution. The number of wrappings in Eq. (4.16) significantly affects the results. For reference, the results presented in this work truncate the summation to ± 1000.

For convenience, the inverse of the theta function is

$$\vartheta_3^{-1}(\alpha,\beta)=2\alpha+\pi+2\sum_{k=1}^{\infty}\frac{\beta^{k^2}\sin(2k\alpha)}{k}.$$

The inverse of the theta function is particularly useful if the cdf is required to draw random samples, because recalling from Section 3.2.2.1, randomly sampling a pdf is performed by uniformly sampling the corresponding cdf.

Written out, the first five orders of the form of the Rogers-Szegő polynomials given in Eq. (4.14) are

$$\phi_0 = 1$$

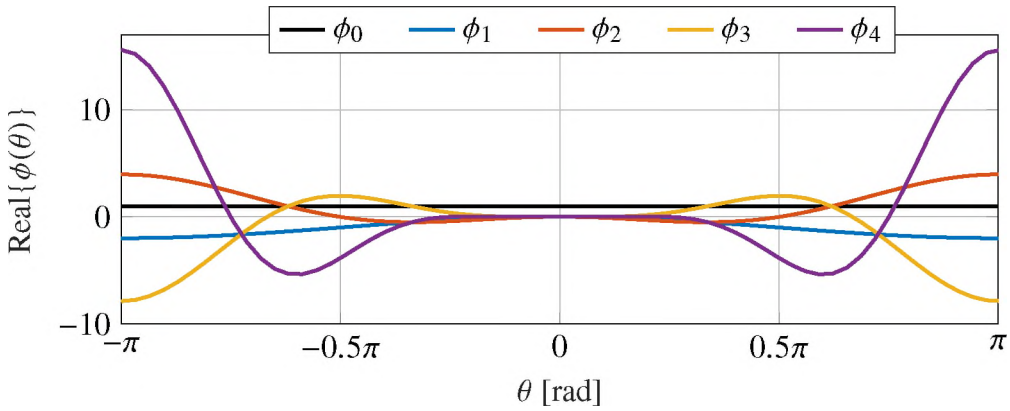
$$\phi_1 = \vartheta - q^{1/2}$$

$$\phi_2 = \vartheta^2 - q^{1/2}(q+1)\vartheta + q$$

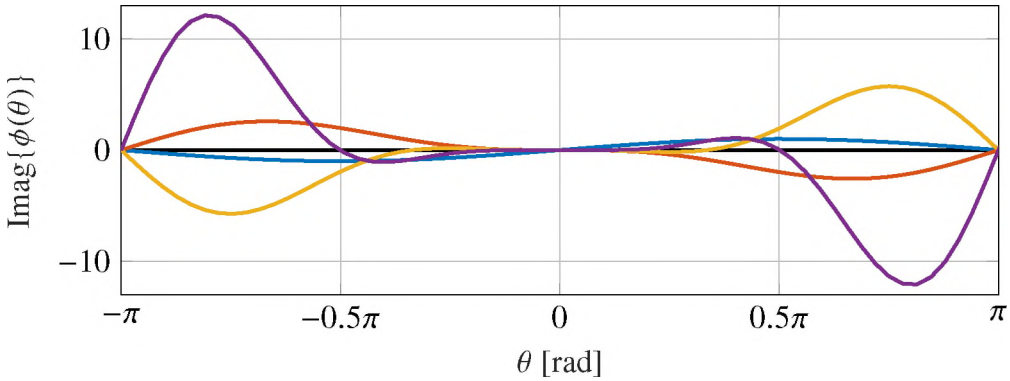
$$\phi_3 = \vartheta^3 - q^{1/2}(q^2+q+1)\vartheta^2 + q(q^2+q+1)\vartheta - q^{3/2}$$

$$\phi_4 = \vartheta^4 - q^{1/2}(q+1)(q^2+1)\vartheta^3 + q(q^2+1)(q^2+q+1)\vartheta^2 - q^{3/2}(q+1)(q^2+1)\vartheta + q^2,$$

and are shown graphically in Figure (4.1). Because the polynomials are complex valued, the real and imaginary components are shown separately. In both cases, the polynomials are oscillatory, with the real component being symmetric about $\theta = 0$ (even), and the imaginary component being antisymmetric about $\theta = 0$ (odd). Additionally, the amplitude of the oscillations increase both with increasing order and distance from $\theta = 0$, similar in behavior to the Askey polynomials. The zeroth polynomial is one, as is standard; therefore, the difference between the two generating functions given in Eqs. (4.14) and (4.16) is only of consequence in the calculation of moments beyond the first.



(a) Real Component



(b) Imaginary Component

Figure 4.1. The zeroth through fourth Rogers-Szegő polynomials with an unwrapped standard deviation of $\sigma = 0.1$.

4.1.3. Function Evaluations and Numerical Complexity. As is to be expected, the computational complexity associated with estimating a random variable using $\text{PCE}_{\mathbb{C}}$ increases with increasing state dimension. It is therefore of interest to develop an expression that bounds the additional number of function evaluations $\text{PCE}_{\mathbb{C}}$ requires beyond $\text{PCE}_{\mathbb{R}}$ as a function of number of states and expansion order that. Due to the many different methods of calculating inner products, all with different computational requirements, the number of functional inner products is what will be enumerated.

Let $\boldsymbol{x} \in \mathbb{S}^P$ be a P-dimensional state vector consisting of angular variables, and let $\boldsymbol{q} \in \mathbb{N}^P$ be the expansion order of each element in $\boldsymbol{x}$, where $\mathbb{N}$ is the set of natural numbers, including zero. The number of inner products required to calculate the chaos coefficients in Eq. (3.30) for element x_i is $2(q_i + 1)$, where $\{i \in \mathbb{N} : k \leq P\}$ and q_i is the i^{th} element of $\boldsymbol{q}$.

Assume that the mean, variance, and covariance are desired for/between each element. The mean does not require any extra inner products, since the mean is simply the zeroth coefficient. The variance from Eq. (4.9) requires an additional $(q_i+1)^2$ inner products for a raw moment, or q_i^2 inner products for a central moment. Similarly, the covariance from Eq. (4.10) between the i^{th} and j^{th} elements requires $(q_i + 1)(q_j + 1)$ additional evaluations for a raw moment and $q_i q_j$ for a central moment. Combining these into one expression, the generalized number of inner product evaluations for raw moments with $P \geq 2$ is

$$2(q_1 + 1) + (q_1 + 1)^2 + \sum_{i=2}^{P} \left(2(q_i + 1) + (q_i + 1)^2 + \sum_{j=1}^{i-1} (q_i + 1)(q_j + 1) \right)$$

and for central moments is

$$2(q_1 + 1) + q_1^2 + \sum_{i=2}^{P} \left(2(q_i + 1) + q_i^2 + \sum_{j=1}^{i-1} q_i q_j \right).$$

It should be noted that this is the absolute maximum number of evaluations that is required for an entirely angular state. In many cases inner products can be precomputed, which reduces the number of computations that must be performed at every time step, if orthonormal polynomials are used, the inner product in the denominator of the coefficient calculation equation (Eq. (4.2a)) is exactly one, which reduces the coefficient calculation inner products by two.

4.2. POLYNOMIAL CHAOS ON THE REAL UNIT CIRCLE

In the Szegő polynomials, the angular random variable is transformed into its complex form; it is for this reason that the Szegő polynomials are complex valued, and the complex conjugate in the inner product (Eq. (4.1)) is important. The complex form of a single angle is inherently two dimensional (one real dimension and one complex dimension). If the radius is fixed, then any point on a circle can be uniquely defined by its angular coordinate. It is proposed that an angular random variable can be treated as a one dimensional variable with the conditions that it be confined to a 2π support (again, using the support $[-\pi, \pi)$) and that the pdf of the random variable be continuous across the endpoints.

Instead of the Szegő polynomials, or any other complex valued polynomial, let a new set of polynomials be developed according to Section 2.2.3 that are orthogonal with respect to a measure on the real unit circle. These polynomials provide a sufficient basis for the PCE of an angular random variable like the Szegő polynomials, but the additional computational burden of the complex conjugate (e.g. Eq. (3.21)) is avoided. A real unit circle PCE can use all of the existing framework utilized by traditional $\text{PCE}_{\mathbb{R}}$ methods and requires none of the additional inner product computations in Section 4.1.3.

For convenience, the real valued polynomials used in this work are orthogonal with respect to the standard wrapped normal pdf (i.e. zero mean and unit variance) and are referred to as wrapped normal polynomials (WNP). The first five orders of the WNP are illustrated in Figure (4.3). For comparison, the normalized Hermite polynomials are also shown. With a standard deviation of 1, the standard normal and the wrapped normal pdfs are fairly similar, with the wrapped normal having larger values when evaluated at $\pm\pi$ than the normalized Hermite polynomials, as can be seen in Figure (4.2).

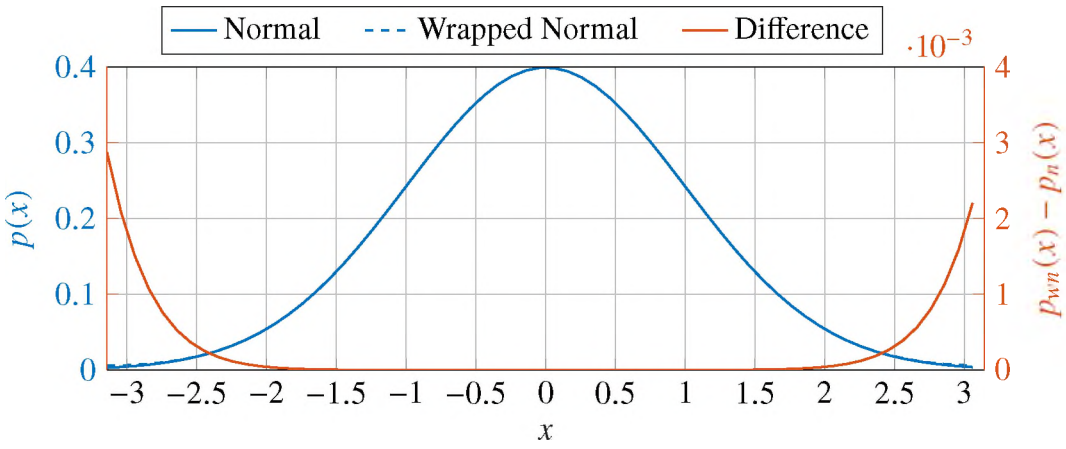


Figure 4.2. Standard normal and wrapped normal pdfs.

Because of the similarity between the standard normal and standard wrapped normal pdfs, the difference between the normalized Hermite and the wrapped normal polynomials (WNP) is not very large for the lower ordered polynomials (see Figure (4.3)); however, as the order increases, the deviation becomes more apparent. As a result, the difference between Hermite-chaos and WNP-chaos will be most noticeable in higher order expansions.

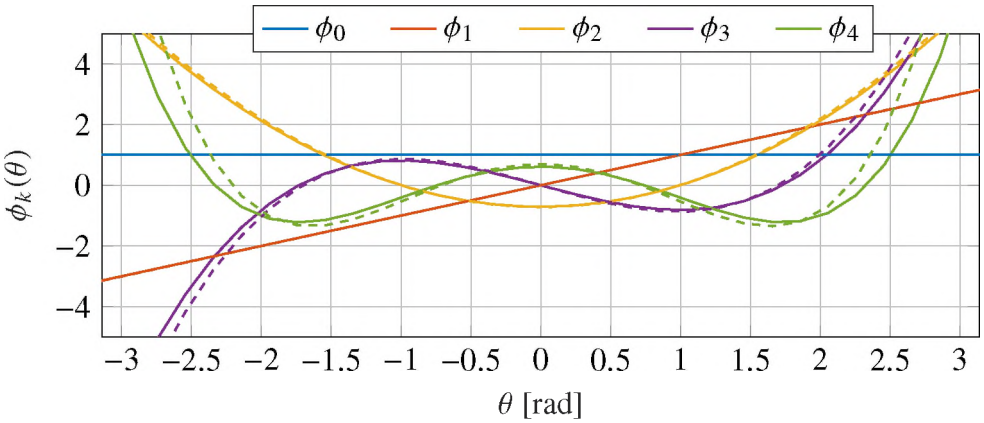


Figure 4.3. First five orders of the normalized Hermite (solid) and WNP (dashed) polynomials on the interval $[-\pi, \pi)$.

A contributing factor of using the standard wrapped normal pdf to develop the WNP is the small values of $p_{wn}(\pm\pi)$, ensuring the general normal shape is maintained, but the tails are not overly large, which would make a wrapped uniform pdf harder to approximate:

recall the PCE converges quicker when the standard random variable in the expansion has a similar shape to the random variable being expanded. In addition, unit variance helps simplify the wrapped normal expression to

$$p_{wn}(\theta; 0, 1^2) = \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} \exp\left\{-\frac{(\theta + 2\pi k)^2}{2}\right\} \qquad -\pi \leq \theta < \pi\,. \tag{4.17}$$

It is important to note that the reduction of dimension can be performed because angular random variables are expressed on the circle with fixed unitary radius; implying the two dimensions are linearly dependent. Consider, for example, this equation of a circle:

$$r^2 = x^2 + y^2 \qquad -r \leq x, y \leq r\,.$$

When the radius is fixed, only one independent variable remains. Instead of using a subset Cartesian coordinates, a fixed radius circle can be described entirely by the angle $-\pi \leq \theta < \pi$, effectively reducing the dimension of the problem from 2 (x and y) to 1 (only θ). Figure (4.4) shows how the points on a unit circle are mapped onto a tangent line of length 2π. Here, the distance between points on the unit circle is preserved so that the arc length between points on the circle is the same as the linear distance between the points on the tangent line.

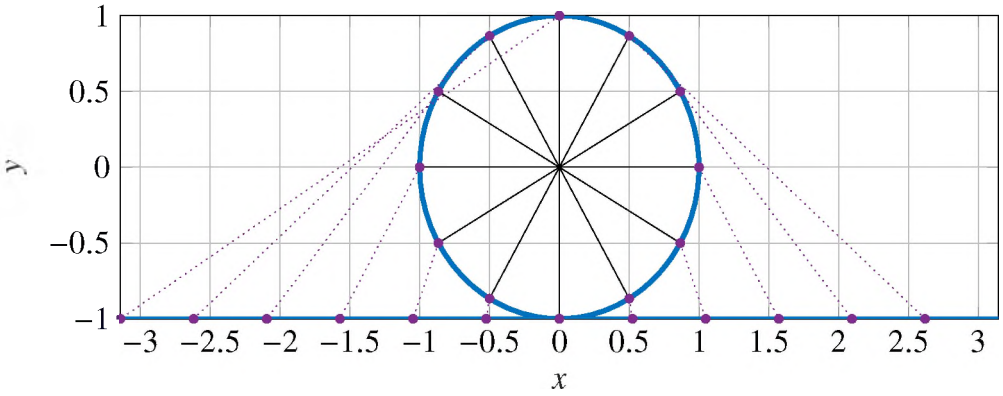


Figure 4.4. Projection of the unit circle onto the line $[-\pi, \pi)$ such that the arc length is preserved as linear length.

It should be noted that this can be performed because the only quantity of interest is the angle, which is not deformed by a mapping of this kind. This is because the unit circle is a curved line, as opposed to the unit disk, which contains all points enclosed* within the unit circle. For random variables on complex manifolds with linearly independent dimensions, representation by $\text{PCE}_{\mathbb{R}}$ would be inappropriate.

4.3. POLYNOMIAL DENSITY

The Rogers-Szegő polynomials and the proposed wrapped normal polynomials are both infinite orthogonal bases; however, the density of the polynomials in $L^2(\Omega, \mathcal{F}, P)$ should also be considered to ensure that the chaos expansion approaches the random variable as the summation approaches infinity. To prove polynomial density, it is sufficient to show that the measure that generates the polynomials is uniquely solved by its moments[75,70]. Carleman's condition[7] says that if the moments $\{m_0, m_1, \ldots\}$ of the measure satisfy

$$\sum_{k=0}^{\infty} \frac{1}{\sqrt[2n]{m_{2n}}} = \infty, \tag{4.18}$$

then no other pdf can have the same moment sequence. For the class of wrapped pdfs, the moment sequence is defined by the characteristic equation of the unwrapped pdf when it is evaluated at integer values. The Gaussian distribution with mean μ and standard deviation σ has the characteristic function (also called the moment generating function)

$$\phi(t) = e^{it\mu - t^2\sigma^2/2},$$

*The open unit disk includes all points strictly enclosed by the unit circle, whereas the closed unit disk includes the unit circle as well.

which asymptotically approaches zero with increasing t. The wrapped normal moment sequence is therefore

$$m_n = e^{in\mu - n^2\sigma^2/2} \qquad \forall n \in \mathbb{Z},$$

where $\mathbb{Z}$ is the set of integers. The limit of the summand within Carleman's condition then becomes

$$\lim_{n\in\mathbb{N}^{+0}\to\infty} \frac{1}{e^{i\mu - n\sigma^2}} = \infty,$$

when the mean is finite. The divergence of the summand satisfies Eq. (4.18), implying that the WNP polynomials used in this work are dense in $L^2(\Omega, \mathcal{F}, P)$. If these polynomials had not been dense in L^2, even the infinite chaos expansion in Eq. (3.7) would be an approximation.

As a counterexample, consider the lognormal pdf with moment sequence

$$m_n = e^{n\mu + n^2\sigma^2/2} \qquad \forall n \in \mathbb{Z}.$$

Again evaluating the limit of the summand within Carleman's condition yields

$$\lim_{n\in\mathbb{N}^{+0}\to\infty} \frac{1}{e^{i\mu + n\sigma^2}} = 0.$$

The Stieltjes-Wigert polynomials[10,76] polynomials are orthogonal with respect to the lognormal pdf; however, because they are not dense, the Stieltjes-Wigert polynomials are not *uniquely* orthogonal with respect to the lognormal pdf. As a result, the Stieltjes-Wigert polynomials *could* be used as the basis for polynomial chaos with a lognormal standard random variable; unfortunately, there is no guarantee that the expansion converges to the random variable: a guarantee provided when dense orthogonal basis polynomials are used.

5. RESULTS AND DISCUSSION

Section 4 develops two different methods of representing and propagating the uncertainty in angular random variables. While the two methods seek to accomplish the same goal, the developmental processes required for each method are significantly different. A complete generalization of the expansion itself, as well as new derivations of moment equations are presented in Section 4.1. Because $\text{PCE}_{\mathbb{C}}$ is entirely novel, its feasibility must be determined before its accuracy can be compared with existing methods.

By comparison, $\text{PCE}_{\mathbb{R}}$ (in Section 4.2) has been well established as a viable method of representing and propagating uncertainty. Because WNP-chaos does not require any changes to existing $\text{PCE}_{\mathbb{R}}$ methodologies, its accuracy can be tested against existing methods that do not treat the random variable as an angle.

Because the two methods in Section 4 are at different levels, they are not tested against each other. In Section 5.1, $\text{PCE}_{\mathbb{C}}$ is used to jointly propagate two correlated angles in a double pendulum setup. The estimates of the first two raw moments (including correlation) are estimated and compared with estimates obtained from Monte Carlo simulation.

Section 5.2 uses the two-body equinoctial dynamics to test WNP-chaos as a more accurate method of uncertainty propagation than the UT and Hermite-chaos. Again, the first two moments, including correlation, are estimated, and the errors (by comparing with Monte Carlo simulation) are compared. In addition, the Hellinger distance of the pdf obtained from the different methods from the Monte Carlo simulation is analyzed.

5.1. ANGLES ONLY UNCERTAINTY PROPAGATION

To test PCE$_{\mathbb{C}}$, consider a double pendulum, as illustrated in Figure (5.1). In this case, the angles α and β describe the angle each rod (of lengths l_α and l_β, respectively) makes with the gravity vector. The rods themselves are massless, but at the end of each rod are point masses m_α and m_β.

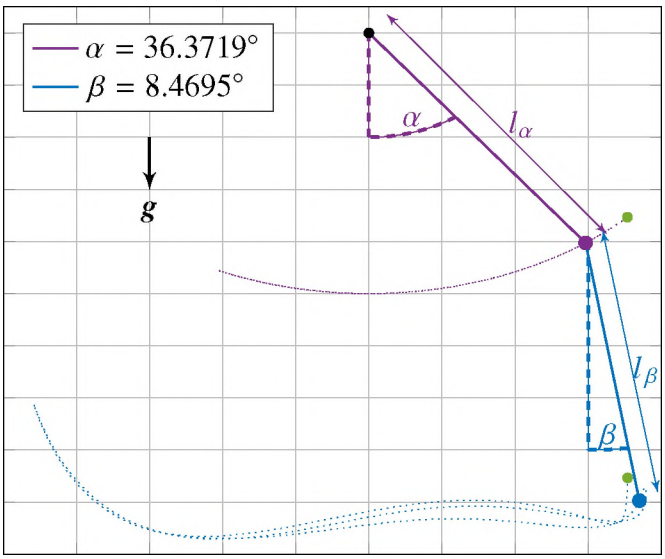


Figure 5.1. Graphical state of a double pendulum after approximately 4 seconds indicating the angles α and β as well as the progression of each point mass throughout the simulation (dashed) beginning at the green dots.

The double pendulum presents convenient dynamical behaviors to test the propagation of uncertainty using PCE$_{\mathbb{C}}$ due to the nonlinear nature of the double pendulum and nonlinear evolution of the uncertainty. The equations of motion can be derived by taking the inertial positions of the masses, i.e.

$$x_\alpha = l_\alpha \sin\alpha \qquad y_\alpha = -l_\alpha \cos\alpha$$

$$x_\beta = x\alpha + l_\beta \sin\beta \qquad y_\beta = y_\beta - l_\beta \cos\beta\,,$$

and differentiating twice with respect to time, acknowledging that the only outside force acting on the system is gravity. The resulting equations of motion are

$$\ddot{\alpha} = \frac{-g[(2m_\alpha + m_\beta)\sin\alpha + m_\beta \sin(\alpha - 2\beta)] - 2\sin(\alpha-\beta)m_\beta(\dot{\beta}^2 l_\beta + \dot{\alpha}^2 l_\alpha \cos(\alpha-\beta))}{l_\alpha(2m_\alpha + m_\beta - m_\beta\cos(2\alpha - 2\beta))} \tag{5.1a}$$

$$\ddot{\beta} = \frac{2\sin(\alpha-\beta)(\dot{\alpha}^2 l_\alpha(m_\alpha + m_\beta) + g(m_\alpha + m_\beta)\cos\alpha + \dot{\beta}^2 l_\beta m_\beta \cos(\alpha-\beta))}{l_\beta(2m_\alpha + m_\beta - m_\beta\cos(2\alpha - 2\beta))}, \tag{5.1b}$$

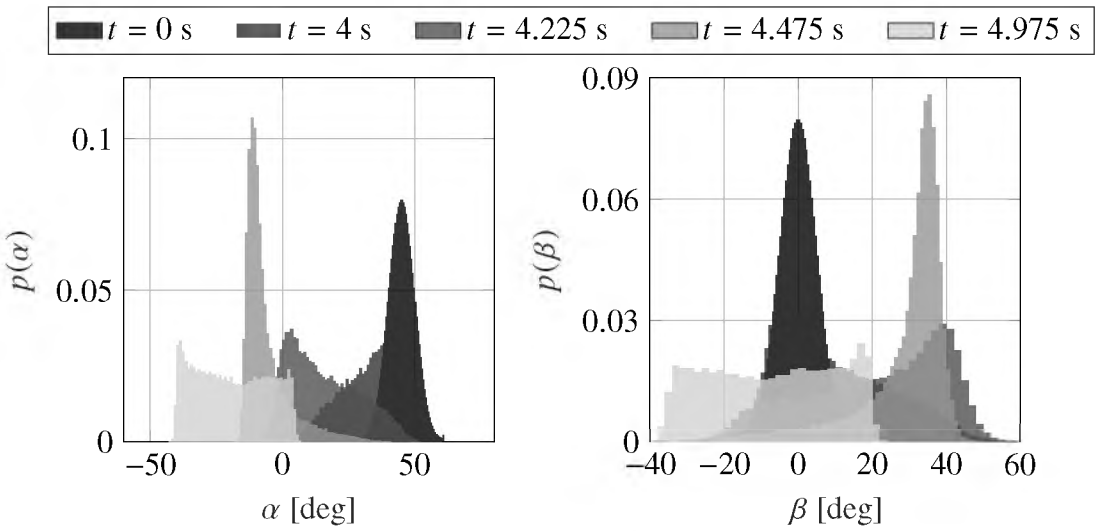
where the dot notation indicates a time derivative (the number of dots indicates the number of time derivatives). Let the state of the system $\boldsymbol{x}$ be comprised of the angles α and β. The initial state of the simulation is given by

$$\boldsymbol{x}_0 \sim \mathcal{WN}\left(\begin{bmatrix} 45° \\ 0° \end{bmatrix}, \operatorname{diag}\begin{bmatrix} (5°)^2 \\ (5°)^2 \end{bmatrix}\right), \tag{5.2}$$

where $\mathcal{WN}$ indicates a wrapped normal pdf with arguments of mean and unwrapped covariance and the notation $\sim$ indicates a distribution, i.e. $x_0 \sim \mathcal{WN}(\mu, \sigma^2)$ implies a wrapped normal pdf with mean μ and variance σ^2. For this scenario, the angular velocities $\dot{\alpha}$ and $\dot{\beta}$ are deterministically known and initially zero, the lengths are both 1 m, and the masses are both 2 kg.

Recalling Figure (5.1), depicted is the state of the pendulum after approximately 4 seconds, chosen so both α and β are nonzero for graphical purposes. In addition to the instantaneous state of the system, the initial locations of m_1 and m_2 are shown as green dots, and the progression each point mass is traced as dotted lines. This figure represents the progression of the mean state obtained using a Monte Carlo simulation of 50,000 samples integrated using MATLAB's `ode45`[77] differential equation solver at a rate of 40 Hz for a total simulation time of 5 seconds.

A primary benefit of PCE is the ability to represent a random variable even when it diverges from a standard random variable e.g. when a random variable becomes non-Gaussian but is expanded using Hermite polynomials. The pdfs of α and β become non-wrapped-Gaussian, as α and β are subjected to the dynamics in Eqs. (5.1), especially after the 4^{th} second of the simulation, which can be seen in Figures (5.2), which are obtained from Monte Carlo simulation.



(a) Progression of $p(\alpha)$ (b) Progression of $p(\beta)$

Figure 5.2. Selected histograms from Monte Carlo simulation of α and β showing the progression from initial wrapped Gaussianity (at $t = 0$ s) to non-wrapped-Gaussian pdfs, particularly after $t = 4$ s.

Both α and β are clearly wrapped normal at the initial time $t = 0$ sec, and in both cases the pdf evolves into an approximate uniform pdf with supports less than 2π. Until $t = 4$ s, the double pendulum behaves approximately like a single pendulum. This provides an opportunity to observe a system that begins uncorrelated and remains minimally correlated until $t = 4$ s, at which point the correlations increase, and the distributions of the states distort exhibiting the range of pdf shapes seen in Figures (5.2).

To test the feasibility of $\text{PCE}_{\mathbb{C}}$, the angular portion of the means ($\sphericalangle\text{E}[\alpha]$, $\sphericalangle\text{E}[\beta]$), length of the means ($||\text{E}[\alpha]||$, $||\text{E}[\beta]||$), second raw moments ($\text{E}[\alpha^2]$, $\text{E}[\beta^2]$), and correlation ($\sphericalangle\text{E}[\alpha\beta]$) are estimated at 40 Hz for 5 seconds. This duration ensures that the pdfs of both states evolve into non-wrappped-Gaussian distributions including pdfs resembling the log normal pdf and uniform pdf on an interval less than 360°. Both α and β are expanded using the Rogers-Szegő polynomials orthogonal with respect to the wrapped normal pdf, and the estimated moment angles are compared against the truth via Monte Carlo simulation.

5.1.1. Assumptions. From the discussion of the generating function for the Rogers-Szegő polynomials (Eq. (4.14)), it is clear that these polynomials depend on the USTD of the wrapped normal pdf. Unfortunately, this means that the polynomials are unique to any given problem, and while they can still be computed ahead of time and looked up, it is not as convenient as problems that use polynomials that are fixed (e.g. Hermite polynomials).

Additionally, the inner product in Eq. (4.9), which describes the calculation of the covariance, requires the knowledge of the joint pdf between the two random variables. This presents a cyclic problem where the joint pdf must be known in order to estimate the state. In practice, there is no reasonable way of obtaining this pdf; and if there is, then the two variables are already so well known that costly estimation methods are irrelevant.

Assumptions about the USTD and the joint pdf must be made. The basis polynomials are evaluated when solving for the chaos coefficients (Eq. (4.2b)) and when estimating the statistical moments (Eqs. (4.6)–(4.11)) at every time step. If no assumption is made about the USTD, then the generating function in Eq. (4.14) (or other method of polynomial generation) must be evaluated at every time step as well. In either case, the computational burden can be greatly reduced if the basis polynomials remain fixed, requiring only an initial evaluation. Additionally, if the same USTD is used for both variables, than the simplification from two to one integrals in Eq. (4.11) can be made. For the results herein, a fixed USTD of 1 is used for both α and β, effectively resulting in the complex analog of the WNP in Section 4.2.

While not required for the mean estimate, a simplification of the joint pdf also significantly reduces the computational burden and increases the feasibility of the problem. The most drastic of simplifications is to use a fixed, uncorrelated joint pdf throughout the entire problem. Note that the pdf used in the inner product is mean centered at zero (even for Askey chaos schemes); therefore, the validity of the estimation is not affected by any movement of the mean.

Not only are the Rogers-Szegő polynomials dependent on the USTD, they utilize a form of the circular standard deviation, which takes values $0 < q < 1$ that are inversely proportional to the USTD. Because the Rogers-Szegő polynomials are orthogonal, the chaos coefficients – beyond the zeroth – are functions of the polynomial normalization constant from (Eq. (4.15)), such that

$$\epsilon_n(x) \propto \frac{1}{\langle \Psi_n(x), \Psi_n(x) \rangle} = \frac{q^n}{\prod_{k=0}^{n-1}(1-q^{k+1})}.$$

As q approaches zero (the wrapped normal pdf approaches a wrapped uniform pdf), the chaos coefficients also approach zero, and as q approaches 1 (the wrapped normal pdf approaches the Dirac delta), the coefficients diverge. In addition, as the polynomial chaos expansion order increases, the chaos coefficients become even more sensitive to the extremes of q. This is not to say that extreme values of q are not viable for PCE, but rather that it is important that numerical precision is maintained for both very small ($q \to 0$) and very large ($q \to 1$) chaos coefficients.

5.1.2. Convergence Behavior. Before discussing the accuracy of each of the seven quantities that is estimated, it is important to examine the convergence behavior of the polynomial chaos expansion. From Section 4.3, the expansion analytically converges to the random variable, but the rate of convergence directly impacts the feasibility of implementation. Additionally, increasing expansion orders can cause numerical instabilities in the decay/divergence rates of the chaos coefficients discussed in Section 5.1.1.

To investigate the convergence behavior of the double pendulum, the error in the estimated moments via $\text{PCE}_{\mathbb{C}}$ with orders ranging from two to 100 is shown in Figure (5.3). The error is defined as the absolute difference in angular components of each moment between the Monte Carlo simulation and Szegő-chaos estimate averaged over the simulation. For example, the error in the moment $\text{E}[\alpha\beta]$ is

$$\frac{1}{201}\sum_{t\in T}\left|\sphericalangle\text{E}[\alpha(t)\beta(t)]_{\text{mc}} - \sphericalangle\text{E}[\alpha(t)\beta(t)]_{\text{sz}}\right| \qquad T = \{0, 0.025, \ldots, 5\},$$

where $\text{E}[\cdot]_{\text{mc}}$ denotes the expected value obtained via Monte Carlo simulation, $\text{E}[\cdot]_{\text{sz}}$ denotes the expected value estimated from Rogers-Szegő-chaos, and the time vector has 201 components.

Note that the chaos coefficients of both angles are calculated from independent expansions, not from a multivariate expansion like the ones in Section 3.2.4.

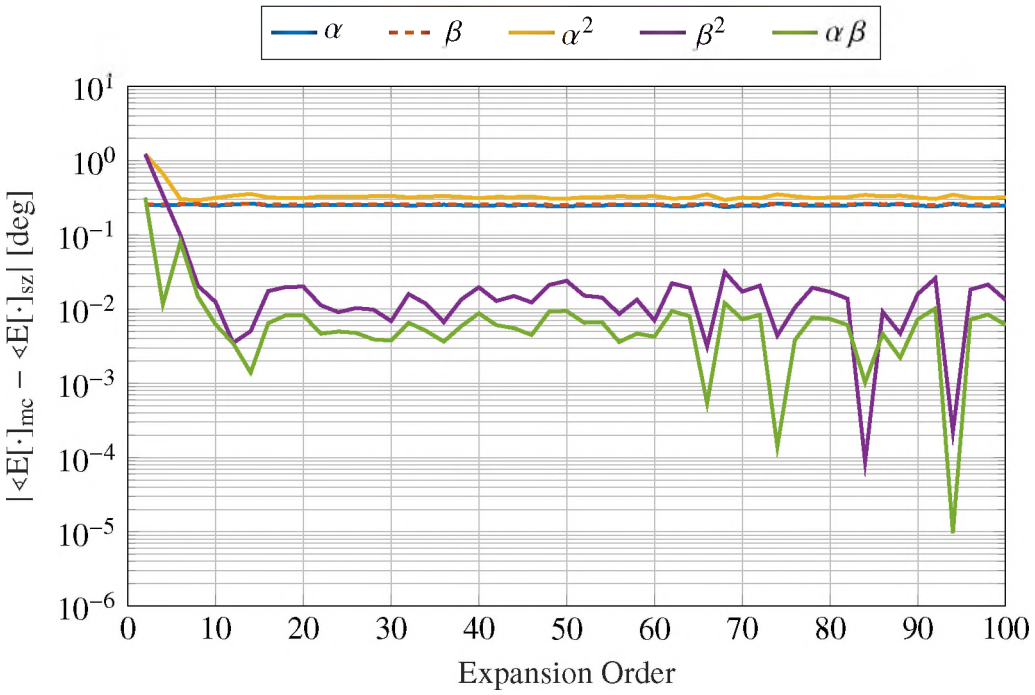


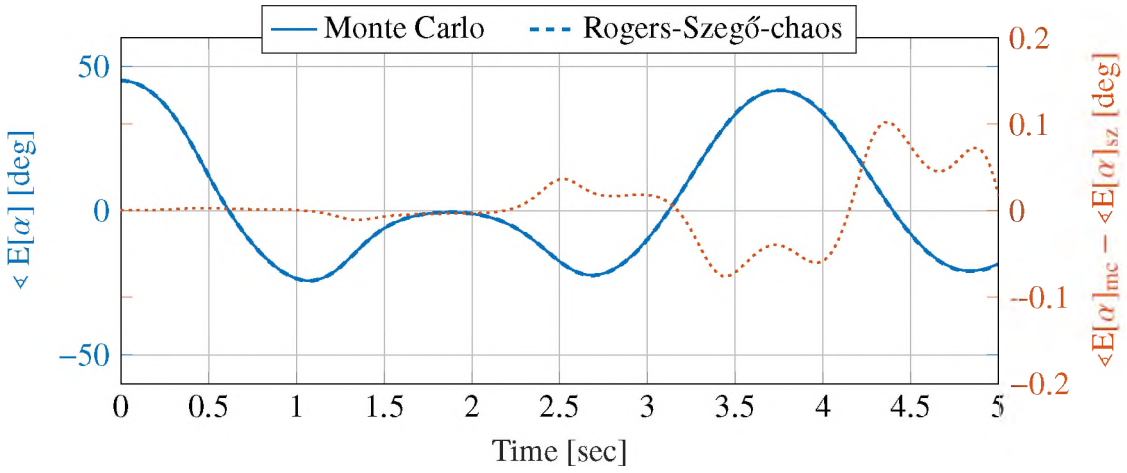
Figure 5.3. Convergence behavior of average (over the simulation) angular difference between Monte Carlo (50, 000 samples per dimension) and $\text{PCE}_{\mathbb{C}}$ due to increasing expansion.

From Figure (5.3), the mean estimates and the second moment of α ($\mu_{2,\alpha}$) reach steady state errors of approximately 0.3° for expansion orders of six and above. While the correlation ($\mu_{\alpha,\beta}$) and second moment of β ($\mu_{2,\beta}$) do not approach steady state values, the angular difference is consistently at least an order of magnitude less than the errors in the means and $\mu_{2,\alpha}$. The increased variability in the errors with increasing orders, which is present in the means and $\mu_{2,\alpha}$, although not as apparent in $\mu_{2,\beta}$ and the correlation, are attributed to numerical inaccuracies in generating the Rogers-Szegő polynomials (given the required approximation of the ϑ_3 function in Eq. (4.14)) as well as in approximating the required polynomial inner products (performed herein using Monte Carlo integration with 12567 points). Additionally, the $\mu_{2,\beta}$ and correlation variation with the large expansion orders is exacerbated due to the logarithmic scaling of the plot.

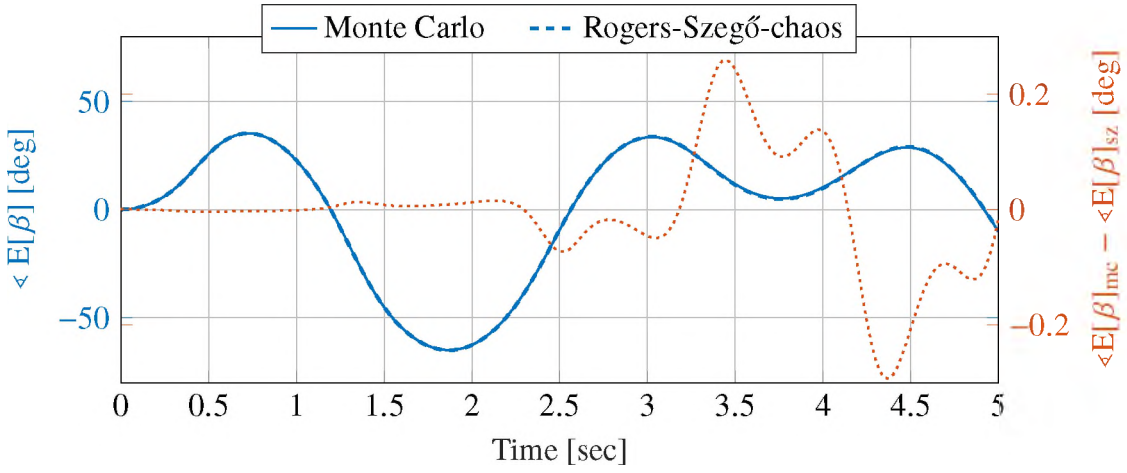
5.1.3. Moment Estimation. Based on the steady state behavior in Figure (5.3), the moment estimates across the five second simulation are investigated for individual expansion orders of 15. Again, the purpose of this investigation is to observe the behavior of the moment estimates, not necessarily the accuracy. In each of the figures in this section (Figures (5.4a)–(5.6b)), the angular portion of the moments obtained via Monte Carlo simulation (e.g. $\sphericalangle \mathrm{E}[\alpha]_{\mathrm{mc}}$) and estimated using Rogers-Szegő-chaos (e.g. $\sphericalangle \mathrm{E}[\alpha]_{\mathrm{sz}}$) are plotted against the left hand axis. The plots from the right hand axis give the difference between the moment via Monte Carlo simulation and the estimate of the same moment from Rogers-Szegő-chaos (e.g. $\sphericalangle \mathrm{E}[\cdot]_{\mathrm{mc}} - \sphericalangle \mathrm{E}[\cdot]_{\mathrm{sz}}$).

Recall, from Eq. (4.8), that the mean is simply the zeroth chaos coefficient. Additionally, the zeroth polynomial is analytic, constant, and normalized, implying all error in the mean is directly attributed to the inaccuracies in numerically approximating the expected value in integral form, i.e., no polynomial self inner products $\left(\left\langle \Psi_j, \overline{\Psi_k} \right\rangle\right)$ are required. As a result, the error in the mean gives an indication of implementation inaccuracies rather than fundamental inaccuracies within $\mathrm{PCE}_{\mathbb{C}}$.

Observing the difference in estimates across the simulation in Figures (5.4), the largest errors are apparent at the peaks and troughs of each state. Dynamically, the largest angles achieved in the simulation occur when the individual pendulum motion changes from clockwise to counterclockwise (or vice versa) and are coincidently the most sensitive to initial conditions. Throughout the simulation, the errors grow cyclically as a result of increasing overestimation of maximum swing angle; however, the general motion of the system is captured.



(a) Angle of α's first moment.



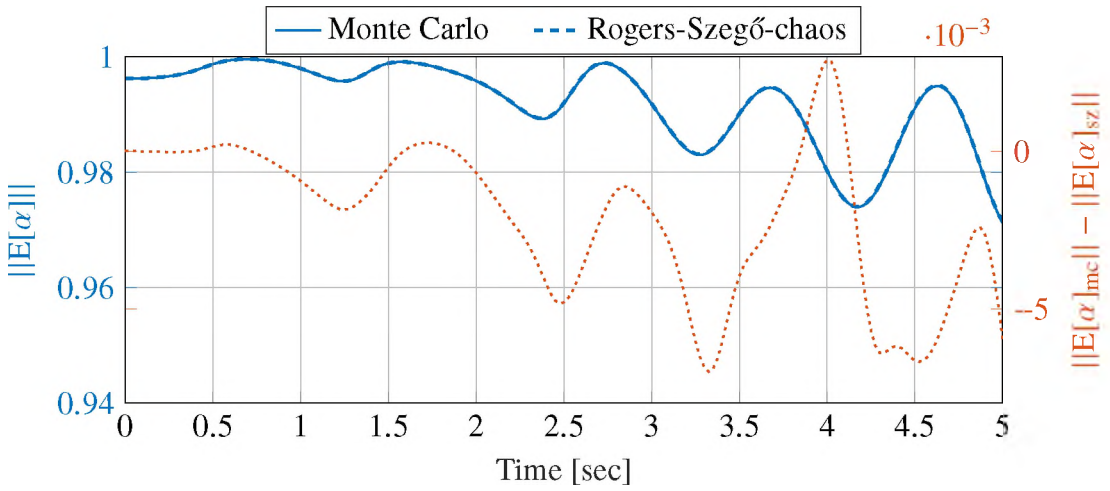
(b) Angle of β's first moment.

Figure 5.4. Angular component of the mean estimated from Monte Carlo and Rogers-Szegő-chaos on the left axis and difference between estimates on the right axis.

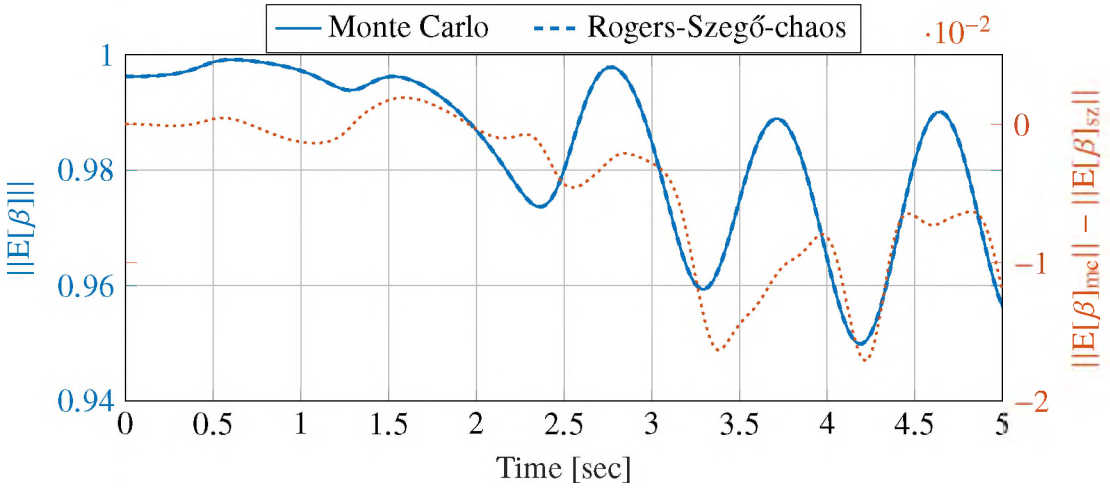
From Section 2.3.3, recall that the concentration of the pdf is found from the length (denoted by $||\cdot||$) of the first moment, which is inversely related to the uncertainty in the random variable. The lengths of the first moments of α and β are depicted in Figures (5.5). Rather than the circular variance or standard deviation, the length is shown, providing an unaltered form of the estimate and a more direct indication of the feasibility of Rogers-Szegő-chaos. If the peaks and troughs of the length are compared with the corresponding angle at the same time, it is apparent that the highest concentrations appear at the switching of swing directions, and the lowest concentration occurs in the middle. Considering the dynamics of the problem, the points where the pendulums switch directions should have the highest concentration, because the random variable "doubles back" on itself, effectively halving the support of the pdfs.

Like the error in the angle, the error in the length grows cyclically, but unlike the angular component, the Rogers-Szegő-chaos estimate almost always under estimates the length of both angles, effectively over estimating the uncertainty in the angle. Rather than an effect of the dynamics, this is most likely a result of numerical approximation. To substantiate this claim, the error in the angular part does not exhibit a similar drift. Between Figures (5.5a) and (5.5b), the error is larger for β, which is to be expected considering the additional range of motion associated with the second pendulum. Whereas the first pendulum is inertially fixed at one end, the second pendulum is not, resulting in larger uncertainty associated with β, which can be seen through the overall downward trend in Figure (5.5b).

In addition to the first moment, it is important to examine the second moments, which use more than just the zeroth chaos coefficient. The estimates of the second raw moments of α and β are given in Figures (5.6), where similar to the angles in Figure (5.4), the error grows cyclically with peaks corresponding to dynamical changes in rotation direction.



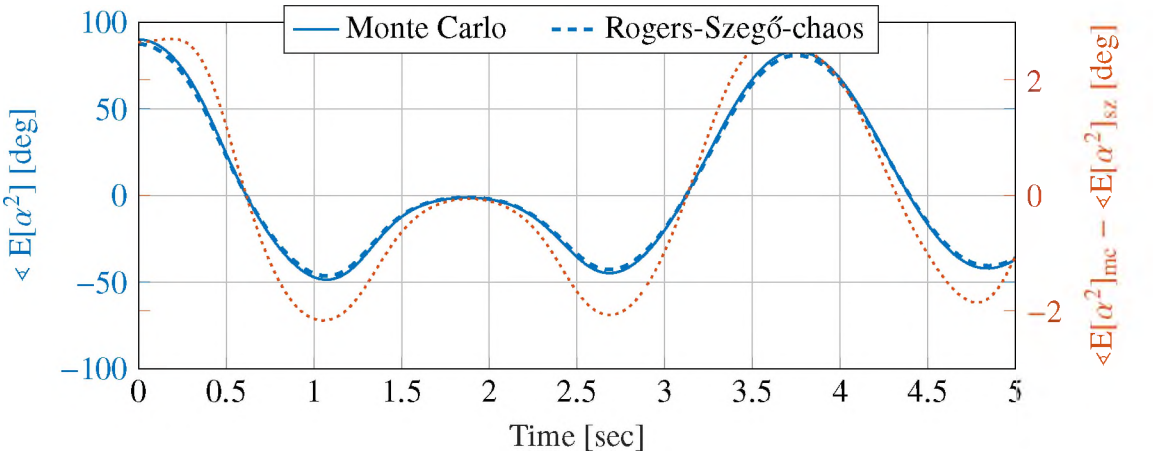
(a) Length of α's first moment.



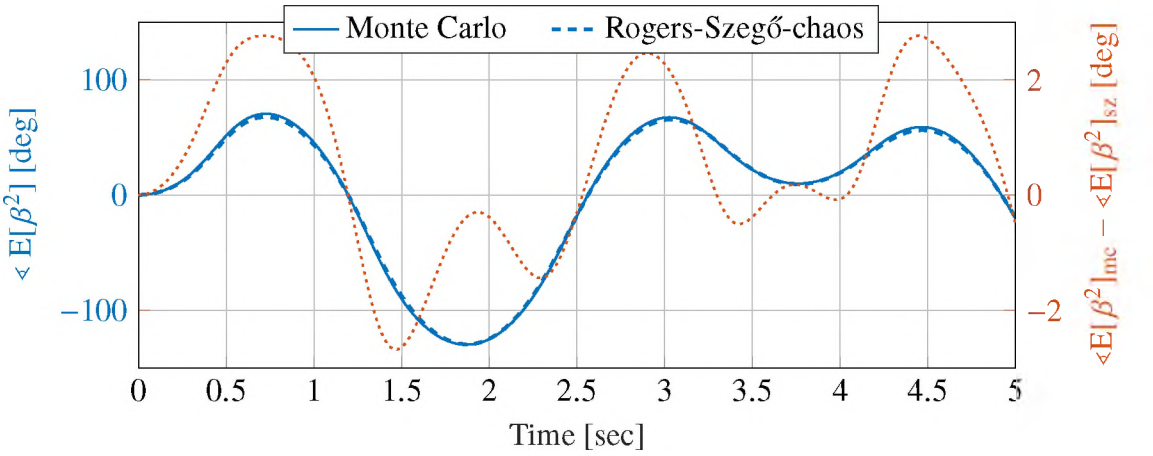
(b) Length of β's first moment.

Figure 5.5. Length of the mean estimated from Monte Carlo and Rogers-Szegő-chaos on the left axis and difference between estimates on the right axis.

Comparing the error in the mean against the corresponding second moment does not reveal any new behavior, only somewhat increased local maximums. Increased error in the peaks is to be expected when considering how the second raw moment is calculated. Recall from Eq. (4.9), the second raw moment requires the inner product between nonorthogonal polynomials, which for this case are approximated numerically (again with Monte Carlo



(a) Angular component of α's second raw moment.



(b) Angular component of β's second raw moment.

Figure 5.6. Angular component of the second raw moment estimated from Monte Carlo and Rogers-Szegő-chaos on the left axis and difference between estimates on the right axis.

integration). It is therefore expected that the errors in higher moments will be larger than the first moment due to the increased number of integrals that must be numerically approximated.

The difference in computational complexity between the mean and higher moments is primarily seen in the polynomial orthogonality factor in the chaos coefficient and the self inner product in the moment equation. While not numerically unstable for an assumed USTD of 1, it is an assumption nonetheless; additionally, the self inner products are numerically approximated in the same way as the chaos coefficients (i.e. via Monte Carlo

integration). It can, therefore, again be assumed that the error in the estimate is due to numerical approximation methods, backed up by the fact that the error in the second raw moment behaves similarly to mean in Figures (5.4).

Based on the mean and second moment estimations, it is not surprising that the correlation estimate (illustrated in Figure (5.7)) tracks the Monte Carlo estimate throughout the simulation with some cyclic growth. What is most apparent is that the errors in the correlation are smaller than those in the mean and second moment estimates. Interestingly, the errors in estimating α and β are similar in value, but opposite in sign. These opposing errors have a negating effect on the correlation error.

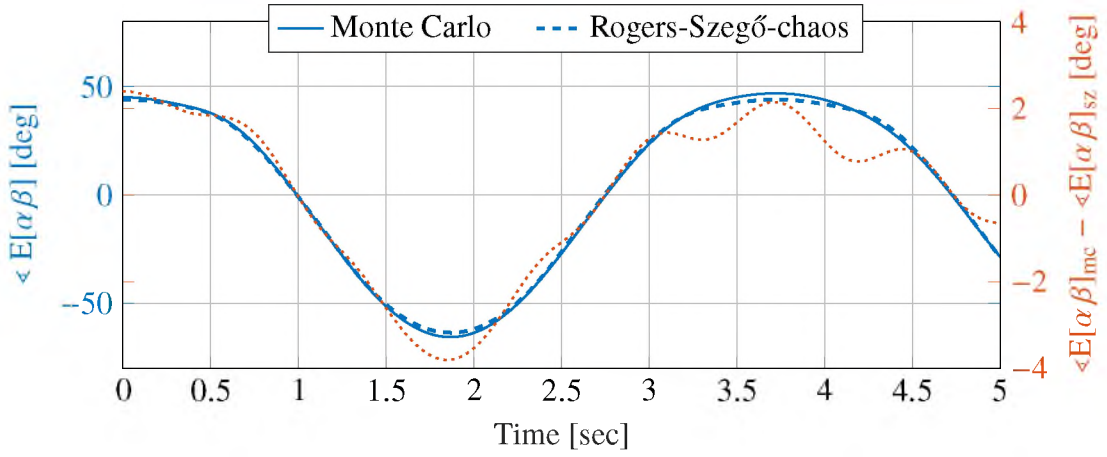


Figure 5.7. Correlation angle estimated from Monte Carlo and Rogers-Szegő-chaos on the left axis and difference between estimates on the right axis.

Given the pattern of error in the estimation of the mean, second raw moment, and correlation, it is clear $\text{PCE}_{\mathbb{C}}$ is a valid method of estimating uncertainty in complex random variables, specifically angular random variables on the complex unit circle. Based on the assumptions made and numerical integration methods, increased accuracy in the estimates can be achieved through the use of basis polynomials that are not problem specific (i.e. dependent on the USTD) and more accurate numerical methods (i.e. Szegő quadrature).

While these results indicate $\text{PCE}_{\mathbb{C}}$ is feasible, it is reasonable to believe the additional computation associated with $\text{PCE}_{\mathbb{C}}$ is too cumbersome for some applications that either require speed (i.e. on-board computing), or that are already very computationally burdensome when using Askey-chaos (i.e. cfd simulation). In cases where random variables must be complex valued, $\text{PCE}_{\mathbb{C}}$ is more than sufficient; however, being able to represent a random variable with real valued polynomials is more efficient.

5.2. TWO-BODY EQUINOCTIAL ELEMENTS

Recently, the equinoctial element set[78] has been gaining popularity in astrodynamics with respect to the Keplerian elements, partially due to the fact that the singularities within the Keplerian elements are not present in the equinoctial elements[79,38]. Similar to the Keplerian elements, the equinoctial elements, given by

$$\boldsymbol{x}_{\text{equ}} = \begin{bmatrix} \text{semimajor axis} & a & \in \mathbb{R} & [\text{km}] \\ \text{eccentricity vector compoments} & \begin{matrix} h \\ k \end{matrix} & \in \mathbb{R}^2 & [\text{N/A}] \\ \text{RAAN vector components} & \begin{matrix} p \\ q \end{matrix} & \in \mathbb{R}^2 & [\text{N/A}] \\ \text{mean longitude} & \lambda & \in \mathbb{S} & [\text{rad}] \end{bmatrix}$$

(where RAAN is the right ascension of the ascending node) are a mixture of translational and angular states. Also like the Keplerian elements, five parameters (a, h, k, p, and q) are used to describe the shape of the orbit and one element (λ) describes the position of the satellite in its orbit. Under the assumption of two-body dynamics, only the mean longitude

varies with time according to

$$\lambda_k = \lambda_0 + \sqrt{\frac{\mu}{a^3}} T_k \,, \tag{5.3}$$

where T_k is the time of flight. Uncertainty in both the mean longitude and the semimajor axis presents a convenient set of dynamics to test the ability of polynomial chaos to estimate the correlation between a translational and angular state.

As is apparent from Eq. (5.3), the mean longitude is dependent on the semimajor axis, which presents an opportunity to test the effectiveness of using real valued polynomials that are orthogonal with respect to measures on the unit circle as the expansion basis for an angular random variable, as well as for testing the ability of mixed expansion $\text{PCE}_{\mathbb{R}}$ to estimate the correlation between an angular random variable and a translational random variable. Because the existing framework for PCE assumes real valued polynomials, the accuracy of the mixed expansion is compared against two-dimensional Hermite-chaos.

In both cases, the chaos coefficients are found using the least squares approach in Section 3.2.2, which is effectively a sampling based method of uncertainty quantification, like the UT. Due to the similarities in numerical methods, and the popularity of the UT in state estimation and uncertainty quantification, the estimates from PCE are compared against the UT. In total, three different methods of estimation are compared: 1) mixed WNP/Hermite-chaos, 2) two state Hermite-chaos, and 3) unscented transform. Within the two types of polynomial chaos, three sampling algorithms are used: random (denoted rand), Latin hypercube (denoted LHC), and Hammersley (denoted Ham). This yields a total of seven approaches that are evaluated.

In all cases, the initial uncertainty is fully captured by the mean and standard deviation of each state, which are taken to be

$$\mu_\lambda = 0^\circ \qquad \text{and} \qquad \mu_a = 7444\ \text{km} \tag{5.4}$$

$$\sigma_\lambda = 1^\circ \qquad\qquad \sigma_a = 100\ \text{km}\,. \tag{5.5}$$

It should be noted that the uncertainties in this example are very large; this is because the primary goal of this work is to test the accuracy of different PCE applications in highly uncertain systems. The unscented transform in this example uses 5 sigma points (corresponding to the $2n + 1$ rule in Section 3.1), and the PCE methods expand λ and a according to

$$\lambda(t,\xi) = \sum_{k=0}^{\infty} \epsilon_{\lambda,k}(t)\Psi_k(\xi) \qquad \text{and} \qquad a(t,\zeta) = \sum_{k=0}^{\infty} \epsilon_{a,k}(t)\Phi_k(\zeta)\,.$$

For the mixed PCE, the polynomials $\mathbf{\Psi}$ are the WNP and ξ is the standard wrapped normal random variable, while the pure Hermite-chaos uses Hermite polynomials and a standard normal random variable. In both expansions, $\mathbf{\Phi}$ are the Hermite polynomials, and ζ is the standard normal random variable.

For both PCE cases, 500 samples are drawn for each state, and the total expansion is 4^{th} order. A Monte Carlo simulation of $100,000$ samples in each state is considered the truth and used to calculate percent errors of the mean, variance, and covariance estimates using all seven methods. In addition to the first two moments, the Monte Carlo simulation provides a full characterization of the random variables; therefore, the accuracy of the full pdf via PCE can be analyzed. If the shape of the pdf is assumed, then the same analysis can be performed with the UT.

The squared Hellinger distance between two univariate pdfs, $p(x)$ and $q(x)$, is given by

$$H^2(p(x), q(x)) = \frac{1}{2}\int_{\mathbb{X}} \left(\sqrt{p(x)} - \sqrt{q(x)}\right)^2 dx, \tag{5.6}$$

where $\mathbb{X}$ is the support of x. The Hellinger distance (and squared Hellinger distance) exists on the closed interval $[0, 1]$, where a Hellinger distance of 0 corresponds to two identical pdfs, and a Hellinger distance of 1 occurs when $p(x)$ and $q(x)$ are mutually singular. As opposed to the Kullback-Leibler (KL) divergence from p to q, given by

$$D_{KL}(p(x))||q(x)) = \int_{\mathbb{X}} p(x) \log\left(\frac{p(x)}{q(x)}\right) dx,$$

if $q(x) = 0$ and $p(x) \neq 0$, the Hellinger distance does not diverge. This is particularly important when considering the fact that PCE can estimate a pdf that does not have an analytic form, resulting in numerical approximations that can reasonably lead to $q(x) = 0$ and $p(x) \neq 0$.

Two different dynamical scenarios are of interest when estimating an angular random variable, and each of these scenarios is examined in the following numerical examples. The first scenario is examined in Section 5.2.1, which looks at the mean longitude at period intervals, beginning after the first orbit and ending after the 50^{th} orbit (i.e. $\mu_\lambda = 2\pi, 4\pi, 6\pi, \ldots, 100\pi$). The second scenario, which is presented in Section 5.2.2, looks at estimates across a single orbit, which highlights the need for continuity across the support bounds at $\pm\pi$.

5.2.1. Period Intervals. The first scenario provides an instance when λ progresses from an approximate normal distribution on the interval $[-\pi, \pi)$ to an approximate wrapped uniform distribution. As time progresses, the pdf flattens causing the probability density function at $\pm\pi$ to become non-zero, and the approximation of a wrapped normal pdf as a normal pdf breaks down. The histograms of λ from the Monte Carlo simulation at different

period intervals are provided in Figure (5.8), where it is apparent that, as the number of orbits increases from 1 to 50, the standard deviation of λ increases and the pdf flattens.

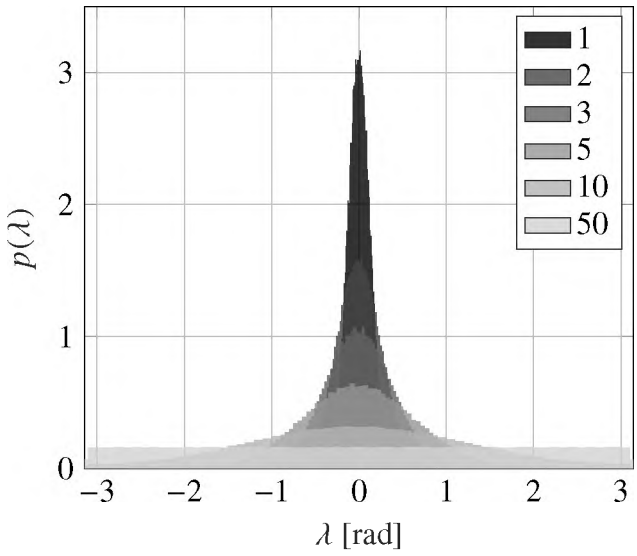


Figure 5.8. Selected histograms indicating the continuous flattening of λ at the same point after multiple revolutions.

5.2.1.1. Moment estimation. Using the Monte Carlo simulation as truth, the percent error in the estimated first and second moments from each of the seven different implementations is compared. Equation (5.3) describes only the propagation of λ, while a remains constant; because of this, the estimates of a never change; for brevity, these results are omitted. The percent error is given by

$$\%E = \left| \frac{\mathrm{E}[\cdot]_{\mathrm{mc}} - \mathrm{E}[\cdot]_{\mathrm{est}}}{\mathrm{E}[\cdot]_{\mathrm{mc}}} \right| \times 100\%$$

where $\mathrm{E}[\cdot]_{\mathrm{mc}}$ is the moment (mean, variance, or covariance) obtained via Monte Carlo simulation, and $\mathrm{E}[\cdot]_{\mathrm{est}}$ is the same moment estimated using one of the seven methods. Table 5.1 presents the percent error of the estimated mean and variance of the mean longitude as

well as the covariance between mean longitude and semimajor axis. Between the endpoints presented in Table 5.1, the percent error in the estimates monotonically decreases from the first orbit (t_1) to the final estimate after 50 orbits (t_{50}) for all seven scenarios.

Table 5.1. Percent error in the initial and final estimates of the first two moments from all seven methods.

Method	mean t_1 [rad]	mean t_{50} [rad]	var t_1 [rad^2]	var t_{50} [rad^2]	cov t_1 [rad·km]	cov t_{50} [rad·km]
H - LHC	1.445e-5	2.542e-7	6.401e-5	1.131e-5	3.650e-4	5.144e-5
H - Ham	1.445e-5	2.543e-7	6.404e-5	1.132e-5	3.650e-4	5.148e-5
H - Rand	1.445e-5	2.539e-7	6.404e-5	1.129e-5	3.650e-4	5.134e-5
WNP - LHC	1.430e-5	2.505e-7	7.457e-4	1.120e-5	3.649e-4	5.071e-5
WNP - Ham	1.430e-5	2.507e-7	7.459e-4	1.139e-5	3.650e-4	5.164e-5
WNP - Rand	1.429e-5	2.505e-7	7.459e-4	1.142e-5	3.650e-4	5.182e-5
UT	1.464e-5	4.345e-7	1.665e-3	1.056e-3	9.572e-4	5.976e-4

The one method that stands apart from the others is the UT. For all six statistical estimates, the UT is outperformed by the PCE methods, especially in the variance and covariance. This is to be expected since the UT is designed to use a specific number of sigma points, whereas the number of polynomial chaos samples is defined by the user and is generally much larger than the number of UT sigma points.

For the sake of argument, let each of the polynomial chaos methods be expanded to a first order total expansion with five total samples. With this type of expansion, only the first two moments can be estimated, mimicking the unscented transform. In this very extreme case, the PCE methods are slightly outperformed by the UT; however, this is not the manner in which PCE should be used at all. It is the ease of increasing expansion order and sample sizes that makes PCE a convenient, and accurate, method of quantifying a random variable.

While the UT is outperformed in all of the estimates, it should be noted that the mean estimate is, relatively speaking, significantly better than the variance and covariance estimates. This occurs because, for both PCE and the UT, the mean is estimated using a weighted average.

Now, a comparison of the six chaos/sampling method combinations is considered. For all three moments, the estimates produced by each method at t_1 are similar. The WNP mean is slightly better than the Hermite, this flips for the variance, and the estimates are nearly identical for the covariance. More interesting are the estimates at t_{50}. Again, there is not a lot of difference between the WNP and Hermite, but for all estimates, the WNP with Latin hypercube sampling produces the smallest error, especially for the variance and covariance.

The underlying assumption of PCE is that the n standard random variables $\boldsymbol{\xi}$ are independent; when the samples are drawn using random sampling, there is no guarantee at all of independence. Latin hypercube intelligently uses random sampling in all dimensions, and Hammersley sampling deterministically sets the first dimension such that the points are evenly dispersed, then intelligently uses random sampling for the remainder of the dimensions. As a result, the random sampling produces reasonable mean estimates, but inconsistent second moments. The even dispersion of the first moment ensures the full support of the random variable is sampled, but the dependence of the samples on each other can lead to a higher correlation between dimensions in lower dimensional problems. As the dimension of the problem increases, the dependence of the samples in the first dimension is less impactful. Because the problem is only two dimensional, Hammersley sampling is not ideal, and as a result, Latin hypercube sampling produces the best moment estimates for this problem.

5.2.1.2. Chaos coefficient examination. While looking at the moment estimates is usually the main purpose of PCE, much can be learned about the system by looking at the individual chaos coefficients. Recall that, for sampling based PCE, the total expansion order describes the multivariate polynomial degree, rather than the univariate polynomial degrees for each dimension. For example, the ordering used in this work produces the coefficient-

multivariate basis pairs given in Table 5.2. The polynomials $\mathbf{\Psi}$ and $\mathbf{\Phi}$ correspond to the expansions in Eq. (5.4). In this case, the state dimension is 2 and the total expansion order is 4; therefore, the total number of terms in the expansion is 15.

Table 5.2. The univariate basis polynomials that are combined to create the multivariate basis polynomial associated with each chaos coefficient.

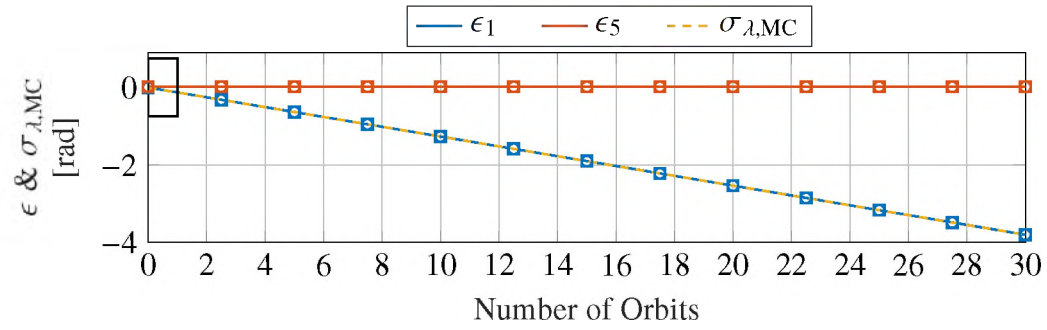
	ϵ_0	ϵ_1	ϵ_2	ϵ_3	ϵ_4	ϵ_5	ϵ_6	ϵ_7	ϵ_8	ϵ_9	ϵ_{10}	ϵ_{11}	ϵ_{12}	ϵ_{13}	ϵ_{14}
ψ_n	0	0	0	0	0	1	1	1	1	2	2	2	3	3	4
ϕ_m	0	1	2	3	4	0	1	2	3	0	1	2	0	1	0

The mean estimates obtained from the WNP- and Hermite-chaos do not differ greatly, mainly due to the fact that the mean relies only on the zeroth coefficients, which has a corresponding polynomial of 1. In effect, the mean becomes a weighted average of the sampled states. By contrast, the second moment – in fact each of the moments beyond the first – uses all of the rest of the coefficients. From Eq. (3.21), the second moment is calculated by the sum of the coefficient products. Because all polynomials used in PCE are dense, the chaos coefficients decay; therefore, the second moment converges to the true second moment as the number of terms in the summation increases. This convergence of the coefficients ensures that truncating the summation still produces accurate estimates, but raises the question: at what point should the truncation be made? Unfortunately, the answer is problem dependent, based on how the random variables are propagated forward in time. The two-body propagation in Eq. (5.3) does not significantly deform the pdf of λ into a non-wrapped-normal pdf (the USTD increases, but the random variable retains an approximately wrapped normal pdf); therefore, a relatively small 4^{th} order total expansion is sufficient. In problems where the system dynamics cause a more distinct deformation of the states' pdf from their initializations, a larger total expansion is appropriate.

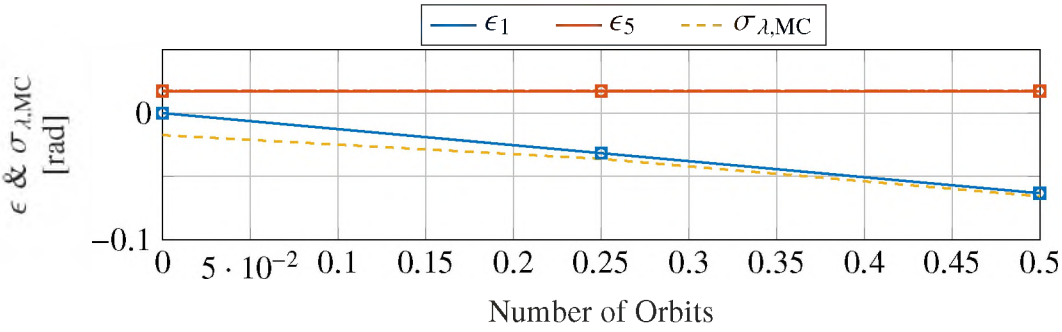
Some of the more interesting coefficients for this example are shown in Figures (5.9). The first and fifth coefficients, shown in Figures (5.9a) and (5.9b), are the largest throughout the simulation and are, therefore, the major contributers to second moment. Even though the fifth coefficient is small compared to the first in Figure (5.9a), a closer inspection of the

first half of an orbit shows that the variance is initially defined by the fifth coefficient. For this problem, and based on the coefficient numbering in Table 5.2, it is apparent that these coefficients coincide with the two multivariate polynomials of combined degree 1. This is expected based on the discussion in Section 3.2.4. When the univariate basis expansions in Eq. (5.4) are expanded using the multivariate basis in Table 5.2, the fifteen element expansions of both elements only contain two nonzero coefficients. For both elements, the first nonzero coefficient is the zeroth, which indicates the mean; the other coefficient is the fifth for λ, and the first for the semimajor axis. From Table 5.2, the fifth and first coefficients correspond to a multivariate polynomial comprised of first order polynomials from the element's univariate expansion and the zeroth order polynomial from the other element's expansion. The initial state is uncorrelated; therefore, the standard deviation of λ in Figure (5.9b) (shown at quarter period intervals to better show convergence) is dominated by the fifth coefficient. As the simulation progresses and the mean longitude is continuously altered based on the semimajor axis, the first coefficient grows, and the contribution from the fifth coefficient becomes overshadowed. It should be noted that the coefficients in Figure (5.9c) are estimated at one period intervals, but the figure marks are shown less frequently to reduce clutter. It should also be noted that all six methods are included in Figures (5.9), even though they are not distinguishable at the shown scales.

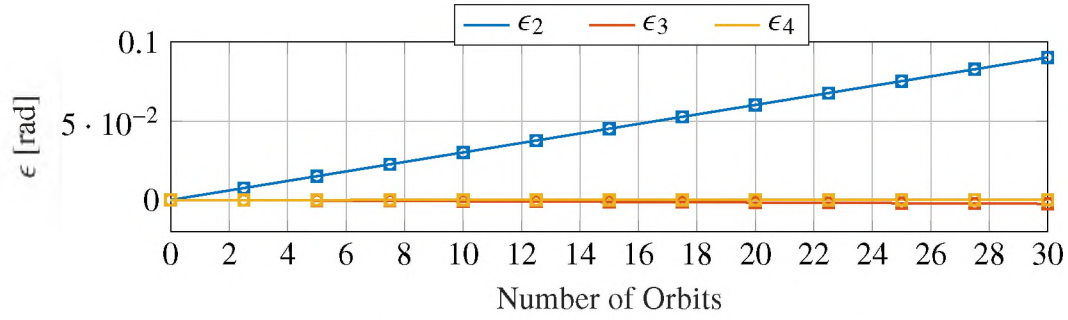
Now, consider the difference between how λ_0 and a are used in Eq. (5.3). The mean longitude is a function of the initial mean longitude without alteration, which is why the fifth coefficient remains stationary throughout the simulation. On the other hand, the semimajor axis is transformed by its $-3/2$ power and scaled by the time of flight, which obviously changes as the simulation progresses. Shown in Figure (5.9c), the coefficients associated with the continuously transformed semimajor axis only (i.e. the second, third, and fourth coefficients) linearly increase with time, similar to the first. Additionally, this plot serves as an illustration that the coefficients converge to zero with increasing polynomial order. The remaining coefficients have similar, approximately linear, trends away from 0.



(a) Standard deviation computed using Monte Carlo as well as first and fifth chaos coefficients at one period intervals.



(b) Standard deviation computed using Monte Carlo as well as first and fifth chaos coefficients over the first orbit.



(c) Second through fourth coefficients at one period intervals.

Figure 5.9. Coefficient evolution at period intervals. The six different methods are shown for each coefficient with line markings Hermite-random(–), Hermite-Hammersley (- -), Hermite-LHC (- -□), WNP-random(- -○), WNP-Hammersley (-·), WNP-LHC (··).

These coefficients are not shown because, while they are not numerically zero, they are insignificant compared with the coefficients shown in Figures (5.9), continuing to illustrate the mentioned coefficient convergence to zero.

5.2.1.3. Hellinger distance. While it is often sufficient to represent a random variable using only the first and second moments, the necessity sometimes arises to have a full understanding of the random variable. The UT provides very good estimates of the first two moments, especially when the underlying pdf is not dominated by higher order moments. Unfortunately, the first two moments alone do not characterize a unique, and arbitrary, pdf. In contrast, a PCE representation characterizes the full random variable. Expressing a random variable as a PCE, and using the chaos coefficients solely to estimate the first two moments, is a drastic underutilization of its capabilities.

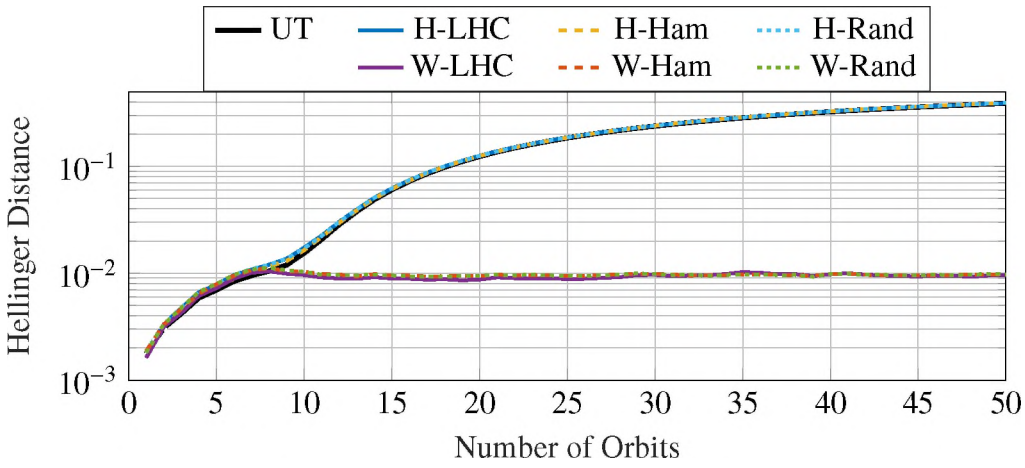
To gain a better understanding of how well the PCE methods are approximating the random variable, the squared Hellinger distance of each estimation type ($\hat{p}(x)$) from the Monte Carlo result ($p(x)$) given by Eq. (5.6) is calculated and shown in Figure (5.10a) at period intervals. Due to the nature of the problem, the moments from the UT are assumed to describe a Gaussian pdf. Note that the legend abbreviation "H" indicates PCE using Hermite polynomials for the expansion of the mean longitude, and "WNP" indicates PCE using wrapped normal polynomials for the longitude. For each of the methods, the semimajor axis is expanded using the Hermite polynomials. Once again within the legend, "LHC" indicates Latin hypercube sampling, "Ham" is Hammersley sampling, and "Rand"is random sampling. From Figure (5.10a), it is apparent that, at the beginning of the simulation, the differences between the methods are small, so much so that the UT and one of the Hermite-chaos methods performs slightly better than WNP-chaos. In these early orbits, the uncertainty in the mean longitude is still relatively small, resulting in very small probabilities that $\lambda = \pm\pi$; therefore, the random variable is reasonably approximated as Gaussian. It is unsurprising that the UT produces such good results at the beginning of the

simulation, since the density obtained from the is assumed to be Gaussian. Similarly, the Hermite-chaos provides very good estimates of translational random variables; therefore, the initially small Hellinger distance is logical.

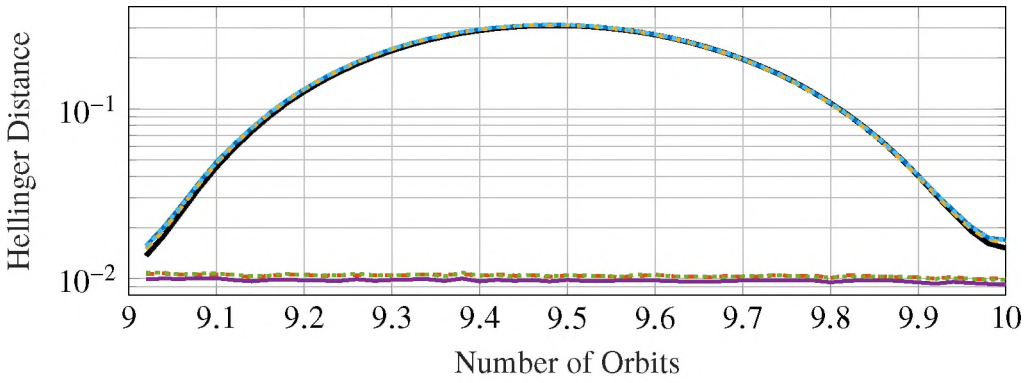
The later orbits are of more interest in this work, since, as the simulation progresses, the real line approximation of λ breaks down. It is between the fifth and tenth orbits that the probabilities at $\pm\pi$ become non-zero, and at this point the WNP-chaos Hellinger distance begins to level off, while the Hermite-chaos and UT distances grow significantly larger. Because the Hermite-chaos and UT do not treat the random variable as an angular random variable, any of the pdf outside the interval $[-\pi, \pi)$ is neglected, which clearly adversely impacts the estimation of the full random variable.

Between the different sampling methods within each type of chaos, there is only a slight difference in Hellinger distances. For both types of chaos, LHC sampling produces slightly more accurate estimates (similar to Table 5.1), but the results are too similar to decisively state which sampling method is most accurate.

5.2.2. Single Complete Orbit. Now that the different estimation types have been examined at period intervals, consider the scenario where the mean is not centered at zero, but instead moves around the unit circle across $\pm\pi$, which causes a discontinuity for a translational random variable. Take, for example, the ninth orbit shown in Figure (5.11). Because no measurements are taken, the uncertainty in the mean longitude grows throughout the orbit; but in general, the points symmetric about the line of apsides should be very similar. This is the reason only selected histograms from the first half of the orbit are shown along with the histogram at the beginning of the next orbit (included to better indicate how much the uncertainty increases over the orbit). Over this orbit, there is not a significant amount of increase in the uncertainty, and the general shapes of the pdfs stay fairly consistent. It can therefore be concluded that any change in accuracy over this orbit is due primarily to the movement of the mean direction around the unit circle.



(a) λ at period intervals.



(b) λ across a single period.

Figure 5.10. Hellinger distances from each estimation type to the Monte Carlo estimate.

Similarly to the trends seen in Table 5.1, the UT is found to produce the least accurate results, and all of the PCE methods perform similarly, with the WNP-chaos with LHC sampling producing the best results. What is much more interesting is the Hellinger distance in Figure (5.10b). In this measure, the wrapped normal-chaos methods outperform both the UT and the Hermite-chaos methods. It is even apparent that the arc present in the Hellinger distance for the Hermite-chaos and UT methods is approximately symmetric with the maximum point occurring at 9.5 periods, exactly where the discontinuity occurs if the random variable is not assumed to be angular. Had the Hellinger distance been calculated more frequently than period intervals, the behavior in Figure (5.10b) would be evident throughout Figure (5.10a).

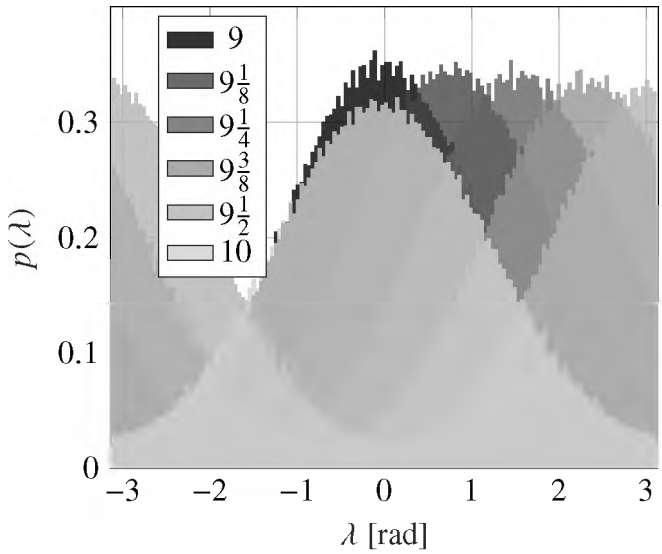


Figure 5.11. Selected histograms indicating the simultaneous shifting and flattening of λ across the period of a single orbit.

Again, like the Hellinger distance in Figure (5.10a), all three sampling methods preform similarly, but for this period, it is apparent that WNP-chaos with Latin hypercube sampling consistently produces the best estimates. It is interesting to note that while there is some variation in the sampling methods for Hermite-chaos at the beginning and end of the orbit, all three sampling methods appear to produce nearly identical results at the half period mark. This is a result of the reduced sensitivity from logarithmic scaling as the Hellinger distance increases.

From this example, it is clear that an angular random variable can not only be propagated using WNP-chaos, but that it can be done so more accurately than the widely used Hermite-chaos and UT. This is especially evident when the approximation of λ's pdf as a Gaussian pdf is invalid: i.e. when the wrapped normal pdf approaches a wrapped uniform pdf, and when the mean direction approaches (and crosses) the support bounds, $\pm\pi$. Not only is λ estimated, but its correlation with the translational random variable a is also estimated to a high degree of accuracy. In addition, these estimates are produced without

using the more costly $PCE_{\mathbb{C}}$ approach. With respect to the different sampling methods presented, Latin hypercube consistently produces the most accurate estimates (when the different methods are distinguishable).

6. CONCLUSIONS

Quantifying the uncertainty of random variables has been the focus of many prior works, including, as one approach, polynomial chaos expansions. The most common types of polynomial chaos expansions are performed on random variables in Euclidean space; however, this work has developed methods that permit using a polynomial chaos expansion to quantify an angular random variable using polynomials that are orthogonal with respect to measures on the complex and real unit circles. Additionally, the covariance between angular random variables and translational random variables can be estimated using the methods developed, even when the expansions of each variable do not use the same basis polynomials.

A generalized set of expressions for the mean and covariance of multi-dimensional systems for both real and complex systems has been presented that does not make the assumption that each variable has been expanded with the same set of basis polynomials. A double pendulum with uncertainty in the two angles defining the state of the system is simulated, presenting an opportunity to jointly estimate two angular random variables under the influence of nonlinear dynamics. When comparing angular portions of the first two circular moments – including correlation – complex polynomial chaos using Rogers-Szegő basis polynomials produce estimates that are, on average, within a degree of Monte Carlo estimates across the simulation time. Over the course of the simulation, the error grows cyclically, corresponding to the largest swings in the pendulums; however, this is to be expected due to the nonlinear nature of the double pendulum and the relatively large amount of initial uncertainty in the state. Even with the observed cyclic divergence, it is clear that the uncertainty in the state governed by nonlinear dynamics can be propagated, including the correlation between angular random variables.

The additional computation required in the estimation of moments, as well as a need for more numerical precision than is needed for expansions using real valued polynomials, makes complex polynomial chaos unappealing for some problems. As an alternative, polynomials that are orthogonal with respect to measures on the real unit circle have been presented and tested in a orbital mechanics problem to estimate the mean longitude of an orbiting body, which is dependent on the initial mean longitude (an angular random variable) and the semimajor axis of the orbit (a random variable on the real line). In addition to the unscented transform, two different types of expansions are considered: pure Hermite-chaos, where both the mean longitude and semimajor major axis are expanded using Hermite polynomials, and a mixed chaos, where the mean longitude is expanded using real valued polynomials that are orthogonal with respect to the wrapped normal probability density function (pdf) and the semimajor axis is expanded using the Hermite polynomials. In comparison with Monte Carlo results, the polynomial chaos expansions outperform the unscented transform and provide similar estimates of the first and second moments, indicating that mixing the bases of expansions yields accurate results, even for covariance estimates. The wrapped normal polynomial chaos expansion significantly outperforms the Hermite-chaos and assumed density unscented transform approximations in the Hellinger distance from the Monte Carlo pdf.

The proposed polynomial chaos expansions for complex random variables and real valued angular random variables are found to be valid methods of representing angular random variables. Additionally, the use of real valued polynomials that are orthogonal with respect to the real valued wrapped normal pdf produce more accurate results than when an angle is expanded using Hermite-chaos.

Future work with the complex chaos expansion includes investigating polynomials orthogonal on the complex unit circle that are not dependent on the standard deviation of the measure with respect to which they are orthogonal, which will reduce the assumptions necessary for implementation. Additionally, more accurate numerical integration are

expected to produce more accurate moment estimation. Once more accurate numerical methods are implemented in complex polynomial chaos, the uncertainty quantification of random variables that are complex with respect to measures on other complex manifolds can be investigated, and the complex expansion can be compared with polynomial chaos using wrapped normal polynomials. Because the wrapped normal polynomials have already proven to be an accurate method of representing an angular random variable, future work also includes simulating more complex dynamical environments with higher dimensions and additional error sources.

REFERENCES

[1] Haberberger, S. J., *An IMU-Based Spacecraft Navigation Architecture Using a Robust Multi-Sensor Fault Detection Scheme*, Master's thesis, Missouri University of Science and Technology, Rolla, MO, USA, 2016.

[2] Galante, J., Eepoel, J. V., Strube, M., Gill, N., Gonzalez, M., Hyslop, A., and Patrick, B., 'Pose measurement performance of the Argon relative navigation sensor suite in simulated-flight conditions,' in 'AIAA Guidance, Navigation, and Control Conference,' Minneapolis, MN, 2012 pp. 1–25, doi:10.2514/6.2012-4927.

[3] Latella, C., Lorenzini, M., Lazzaroni, M., Romano, F., Traversaro, S., Akhras, M. A., Pucci, D., and Nori, F., 'Towards real-time whole-body human dynamics estimation through probabilistic sensor fusion algorithms,' Autonomous Robots, 2019, **43**(6), pp. 1591–1603, ISSN 1573-7527, doi:10.1007/s10514-018-9808-4.

[4] Lubey, D. P. and Scheeres, D. J., 'Supplementing state and dynamics estimation with information from optimal control policies,' 17th International Conference on Information Fusion (FUSION), 2014, pp. 1–7.

[5] Imani, M., Ghoreishi, S. F., and Braga-Neto, U. M., 'Bayesian control of large MDPs with unknown dynamics in data-poor environments,' in 'Advances in Neural Information Processing Systems 31,' pp. 8146–8156, Curran Associates, Inc., 2018.

[6] Hughes, D. L., 'Comparison of three thrust calculation methods using in-flight thrust data,' NASA Technical Memorandum 81360, 1981, pp. 1–32.

[7] Shohat, J. and Tamarkin, J., *The Problem of Moments*, Mathematical surveys and monographs, American Mathematical Society, 1950.

[8] Hausdorff, F., 'Summationsmethoden und momentfolgen I.' Mathematische Zeitschrift, 1921, **9**, pp. 74–109.

[9] Hausdorff, F., 'Summationsmethoden und momentfolgen II.' Mathematische Zeitschrift, 1921, **9**, pp. 280–299.

[10] Stieltjes, T. J., 'Recherches sur les fractions continues,' Annales de la Faculté des sciences de Toulouse : Mathématiques, 1995, **6e série, 4**(1), pp. J1–J35.

[11] Hamburger, H. L., 'Über eine erweiterung des Stieltjesschen momentenproblems,' Mathematische Annalen, 1921, **82**, pp. 168–187, doi:10.1007/BF01498663.

[12] Julier, S., Uhlmann, J., and Durrant-Whyte, H. F., 'A new method for the nonlinear transformation of means and covariances in filters and estimators,' IEEE Transactions on Automatic Control, 2000, **45**(3), pp. 477–482, ISSN 0018-9286, doi: 10.1109/9.847726.

[13] Smith, G., Schmidt, S., and McGee, L., *Application of Statistical Filter Theory to the Optimal Estimation of Position and Velocity on Board a Circumlunar Vehicle*, NASA TR R-135, National Aeronautics and Space Administration, 1962.

[14] Arasaratnam, I., Haykin, S., and Elliott, R. J., 'Discrete-time nonlinear filtering algorithms using GaussâĂŞHermite quadrature,' Proceedings of the IEEE, 2007, **95**(5), pp. 953–977, doi:10.1109/JPROC.2007.894705.

[15] Arasaratnam, I. and Haykin, S., 'Cubature Kalman filters,' IEEE Transactions on Automatic Control, 2009, **54**(6), pp. 1254–1269, doi:10.1109/TAC.2009.2019800.

[16] Wu, C. C., Bossaerts, P., and Knutson, B., 'The affective impact of financial skewness on neural activity and choice,' PLOS ONE, 2011, **6**(2), pp. 1–7, doi: 10.1371/journal.pone.0016838.

[17] Anderson, T. and Mattson, C., 'Propagating skewness and kurtosis through engineering models for low-cost, meaningful, nondeterministic design,' Journal of Mechanical Design, 2012, **134**, p. 100911, doi:10.1115/1.4007389.

[18] Taylor, B., *New Principles of Linear Perspective: Or the Art of Designing on a Plane the Representations of All Sorts of Objects, in a More General and Simple Method Than Has Been Done Before*, Creative Media Partners, LLC, 1719, ISBN 9781379795148.

[19] Mcgee, L. A. and Scmidt, S. F., *Discovery of the Kalman filter as a practical tool for aerospace and industry*, NASA TM-86847, National Aeronautics and Space Administration, 1985.

[20] Athans, M., Wishner, R., and Bertolini, A., 'Suboptimal state estimation for continuous-time nonlinear systems from discrete noisy measurements,' IEEE Transactions on Automatic Control, 1968, **13**(5), pp. 504–514.

[21] Woodburn, J. and Tanygin, S., 'Detection of non-linearity effects during orbit estimation,' Advances in the Astronautical Sciences. Proceedings of the Space Flight Mechanics Meeting, 2010, **136**(10-239).

[22] Huxel, P. J. and Bishop, R. H., 'Navigation algorithms and observability analysis for formation flying missions,' Journal of Guidance, Control, and Dynamics, 2009, **32**(4), pp. 1218–1231, doi:10.2514/1.41288.

[23] Wiener, N., 'The homogeneous chaos,' American Journal of Mathematics, 1938, **60**(4), pp. 897–936, doi:10.2307/2371268.

[24] Janson, S., *Gaussian Hilbert Spaces*, Cambridge Tracts in Mathematics, Cambridge, U.K.; New York: Cambridge University Press, 1997, ISBN 9780521561280.

[25] Eldred, M., 'Recent advances in non-intrusive polynomial chaos and stochastic collocation methods for uncertainty analysis and design,' 50th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, 2009, **4**, pp. 2078–2114, doi:10.2514/6.2009-2274.

[26] Ng, L. and Eldred, M., 'Multifidelity uncertainty quantification using non-intrusive polynomial chaos and stochastic collocation,' 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference, 2012, **9**, pp. 7669–7685, doi:10.2514/6.2012-1852.

[27] Savin, E. and Faverjon, B., 'Higher-order moments of generalized polynomial chaos expansions for intrusive and non-intrusive uncertainty quantification,' 19th AIAA Non-Deterministic Approaches Conference, 2017, pp. 215–223, doi:10.2514/6.2017-0597.

[28] Xiu, D. and Karniadakis, G. E., 'The Wiener-Askey polynomial chaos for stochastic differential equations,' SIAM Journal of Scientific Computing, 2003, **24**(2), pp. 619–644, ISSN 1064-8275, doi:10.1137/S1064827501387826.

[29] Schmid, C. L. and DeMars, K. J., 'Minimum divergence filtering using a polynomial chaos expansion,' Advances in the Astronautical Sciences. Proceedings of the AAS/AIAA Astrodynamics Specialist Conference, 2017, **162**(17-815).

[30] Xiu, D. and Karniadakis, G. E., 'Modeling uncertainty in flow simulations via generalized polynomial chaos,' Journal of Computational Physics, 2003, **187**, pp. 137–167, doi:10.1016/S0021-9991(03)00092-5.

[31] Hosder, S., Walters, R. W., and Perez, R., 'A non-intrusive polynomial chaos method for uncertainty propagation in CFD simulations,' 44th AIAA Aerospace Sciences Meeting and Exhibit, 2006, pp. 1–19.

[32] Hosder, S., Walters, R. W., and Balch, M., 'Point-collocation non-intrusive polynomial chaos method for stochastic computational fluid dynamics,' AIAA Journal, 2010, **48**(12), pp. 2721–2730.

[33] Hosder, S. and Walters, R. W., 'Non-intrusive polynomial chaos methods for uncertainty quantification in fluid dynamics,' 48th AIAA Aerospace Sciences Meeting, 2010, pp. 4–7.

[34] Jones, B. A., Doostan, A., and Born, G. H., 'Nonlinear propagation of orbit uncertainty using non-intrusive polynomial chaos,' Journal of Guidance, Control, and Dynamics, 2013, **36**(2), pp. 430–444, doi:10.2514/1.57599.

[35] Mardia, K. and Jupp, P., *Directional Statistics*, Wiley Series in Probability and Statistics, Wiley, 2009, ISBN 9780470317815.

[36] Markley, F. L., 'Attitude error representations for Kalman filtering,' Journal of Guidance, Control, and Dynamics, 2003, **26**(2), pp. 311–317, doi:10.2514/2.5048.

[37] Darling, J. E., *Bayesian Inference for Dynamic Pose Estimation using Directional Statistics*, Ph.D. thesis, Missouri University of Science and Technology, 2016.

[38] Jones, B. A. and Balducci, M., 'Uncertainty propagation of equinoctial elements via stochastic expansions,' John L. Junkins Dynamical Systems Symposium, 2018, pp. 1–21.

[39] Jones, B. A. and Balducci, M., 'Stochastic expansions including random inputs on the unit circle,' arXiv: Computation, 2018.

[40] Andrews, G. E. and Askey, R., 'Classical orthogonal polynomials,' in C. Brezinski, A. Draux, A. P. Magnus, P. Maroni, and A. Ronveaux, editors, 'Polynômes Orthogonaux et Applications,' Springer Berlin Heidelberg, Berlin, Heidelberg, ISBN 978-3-540-39743-4, 1985 pp. 36–62.

[41] Askey, R. and Wilson, J., *Some Basic Hypergeometric Orthogonal Polynomials that Generalize Jacobi Polynomials*, number 319 in American Mathematical Society: Memoirs of the American Mathematical Society, American Mathematical Society, Providence, RI, USA, 1985, ISBN 9780821823217.

[42] Koekoek, R., Lesky, P. A., and Swarttouw, R. F., *Hypergeometric Orthogonal Polynomials and Their q-Analogues*, Springer Monographs in Mathematics, Springer Berlin Heidelberg, 2010, ISBN 9783642050145.

[43] Hochstrasser, U. W., 'Handbook of mathematical functions, with formulas, graphs, and mathematical tables,,' chapter 22, pp. 771–802, Dover Publications, Inc., USA, ISBN 0486612724, 1974.

[44] Engblom, S., 'Gaussian quadratures with respect to discrete measures,' Technical Report, Department of Information Technology, Upsala University, 2006, (2006-007), p. 17, ISSN 1404-3203.

[45] Szegő, G., *Orthogonal Polynomials*, volume 23 of *American Mathematical Society colloquium publications*, American Mathematical Society, Providence, RI, USA, 1959, ISBN 9780821889527.

[46] Atakishiyev, N. M. and Nagiyev, S. M., 'On the Rogers-Szegő polynomials,' Journal of Physics A: Mathematical and General, 1994, **27**(17), p. L611.

[47] Hou, Q., Lascoux, A., and Mu, Y., 'Continued fractions for Rogers-Szegő polynomials,' Numerical Algorithms, 2004, **35**(1), pp. 81–90.

[48] Vinroot, C. R., 'Multivariate Rogers-Szegő polynomials and flags in finite vector spaces,' ArXiv Mathematics e-prints, 2010, doi:1011.0984.

[49] Andrews, G., *The Theory of Partitions*, Cambridge Mathematical Library, Cambridge University Press, 1998, ISBN 9780521637664.

[50] Wong, Y. K., 'An application of orthogonalization process to the theory of least squares,' The Annals of Mathematical Statistics, 1935, **6**(2), pp. 53–75, ISSN 00034851.

[51] Simon, B., *Orthogonal Polynomials on the Unit Circle*, volume 54 of *Colloquium Publications - American Mathematical Society*, American Mathematical Society, Providence, RI, USA, 2005, ISBN 9780821874769.

[52] Simon, B., 'Orthogonal polynomials on the unit circle: New results,' ArXiv Mathematics e-prints, 2004.

[53] Geronimus, Y. L., 'Orthogonal polynomials: Estimates, asymptotic formulas, and series of polynomials orthogonal on the unit circle and on an interval,' The Mathematical Gazette, 1962, **46**(358), pp. 354–355, doi:10.2307/3611822.

[54] Vanderstraeten, D., 'A stable and efficient parallel block Gram-Schmidt algorithm,' in 'European Conference on Parallel Processing,' Springer, 1999 pp. 1128–1135.

[55] Bingham, C., 'An antipodally symmetric distribution on the sphere,' Annals of Statistics, 1974, **2**(6), pp. 1201–1225, doi:10.1214/aos/1176342874.

[56] Sobol, I. M., *A Primer for the Monte Carlo Method*, CRC Press, 1994, ISBN 9781351469579.

[57] Horn, R. and Johnson, C., *Matrix Analysis*, Cambridge University Press, 1990, ISBN 9780521386326.

[58] Trefethen, L. and Bau, D., *Numerical Linear Algebra*, Other Titles in Applied Mathematics, Society for Industrial and Applied Mathematics, 1997, ISBN 9780898713619.

[59] Särkkä, S., *Bayesian Filtering and Smoothing*, volume 3, Cambridge University Press, 2013.

[60] Wan, E. and van der Merwe, R., 'The unscented Kalman filter for nonlinear estimation,' The Unscented Kalman Filter for Nonlinear Estimation, 2000, **153-158**, pp. 153–158, doi:10.1109/ASSPCC.2000.882463.

[61] Julier, S. J. and Uhlmann, J. K., 'New extension of the Kalman filter to nonlinear systems,' in I. Kadar, editor, 'Signal Processing, Sensor Fusion, and Target Recognition VI,' volume 3068, International Society for Optics and Photonics, SPIE, 1997 pp. 182–193, doi:10.1117/12.280797.

[62] Julier, S. J., 'Skewed approach to filtering,' in O. E. Drummond, editor, 'Signal and Data Processing of Small Targets 1998,' volume 3373, International Society for Optics and Photonics, SPIE, 1998 pp. 271–282, doi:10.1117/12.324626.

[63] Bogachev, V., *Gaussian Measures*, Mathematical Surveys and Monographs, American Mathematical Society, Providence, RI, USA, 2015, ISBN 9781470418694.

[64] Cameron, R. H. and Martin, W. T., 'The orthogonal development of non-linear functionals in series of Fourier-Hermite functionals,' Annals of Mathematics, 1947, **48**(2), pp. 385–392, ISSN 0003486X.

[65] Fubini, G., *Sugli Integrali Multipli: Nota*, Tipografia della R. Accademia dei Lincei, 1907.

[66] McKay, M. D., Beckman, R. J., and Conover, W. J., 'A comparison of three methods for selecting values of input variables in the analysis of output from a computer code,' Technometrics, 1979, **21**(2), pp. 239–245, ISSN 00401706.

[67] Hammersley, J. and Handscomb, D., *Monte Carlo Methods*, Monographs on Applied Probability and Statistics Series, John Wiley & Sons, Incorporated, 1964, ISBN 9780470347249.

[68] del Junco, A. and Steele, J. M., 'Hammersley's law for the van der Corput sequence: An instance of probability theory for pseudorandom numbers,' Annals of Probability, 1979, **7**(2), pp. 267–275, doi:0.1214/aop/1176995087.

[69] Engels, H., *Numerical quadrature and cubature*, Computational mathematics and applications, Academic Press, 1980, ISBN 9780122388507.

[70] Ernst, O. G., Mugler, A., Starkloff, H.-J., and Ullmann, E., 'On the convergence of generalized polynomial chaos expansions,' ESAIM: Mathematical Modelling and Numerical Analysis, 2012, **46**(2), pp. 317–339.

[71] Feller, W., *An introduction to probability theory and its applications*, number v. 2 in Wiley series in probability and mathematical statistics. Probability and mathematical statistics, Wiley, 1966.

[72] Jones, C. H., 'Generalized hockey stick identities and n-dimensional blockwalking,' The Fibonacci Quarterly, 1996, **34**(3), pp. 280–288.

[73] Kullback, S. and Leibler, R. A., 'On information and sufficiency,' Annals Mathematical Statistics, 1951, **22**(1), pp. 79–86, doi:10.1214/aoms/1177729694.

[74] Hellinger, E., 'Neue begründung der theorie quadratischer formen von unendlichvielen veränderlichen,' Journal für die reine und angewandte Mathematik, 1909, **1909**(136), pp. 210–271.

[75] Riesz, M., 'Sur le problème des moments et le théorème de Parseval correspondant,' Scandinavian Actuarial Journal, 1924, **1924**(1), pp. 54–74, doi: 10.1080/03461238.1924.10405368.

[76] Wigert, S., *Sur les polynomes orthogonaux et l'approximation des fonctions continues*, Arkiv för matematik, astronomi och fysik, Almqvist & Wiksell, 1923.

[77] Shampine, L. F. and Reichelt, M. W., 'The MATLAB ode suite,' SIAM Journal on Scientific Computing, 1997, **18**(1), pp. 1–22, ISSN 1095-7197, doi: 10.1137/S1064827594276424.

[78] Broucke, R. A. and Cefola, P. J., 'On the equinoctial orbit elements,' Celestial Mechanics, 1972, **5**(3), pp. 303–310, ISSN 1572-9478, 0008-8714, doi:10.1007/BF01228432.

[79] Sabol, C., Sukut, T., Hill, K., Alfriend, K. T., Wright, B. F., Li, Y., and Schumacher, P. W., 'Linearized orbit covariance generation and propagation analysis via simple Monte Carlo simulations,' Advances in the Astronautical Sciences. Proceedings of the Space Flight Mechanics Meeting, 2010, **136**(10-134).

VITA

Christine L. Schmid was raised in Portland, MO to parents Amy and Charlie Schmid. She graduated from South Callaway High School in May of 2010 and began pursuing a B.S. in Aerospace Engineering in August of 2010 at Missouri University of Science and Technology. After graduating in May of 2014, she remained at Missouri University of Science and Technology as a graduate student under Dr. DeMars. During this time she received the university's Chancellor's fellowship, as well as the Graduate Assistance in Areas of National Need (GAANN) fellowship. Beginning in 2016, she was hired as a Pathways intern at NASA Glenn Research Center in the Mission Architecture and Analysis branch working on multiple elements of the Lunar Orbital Platform-Gateway project. In August, 2020, she received her Ph.D. in Aerospace Engineering from Missouri University of Science and Technology.