



Article Angular Correlation Using Rogers-Szegő-Chaos

Christine Schmid ^{1,*} and Kyle J. DeMars ²

- ¹ Department of Mechanical and Aerospace Engineering, Missouri University of Science and Technology, Rolla, MO 65409, USA
- ² Department of Aerospace Engineering, Texas A&M University, College Station, TX 77843, USA; demars@tamu.edu
- * Correspondence: clshg6@mst.edu

Received: 30 December 2019; Accepted: 21 January 2020; Published: 1 February 2020



Abstract: Polynomial chaos expresses a probability density function (pdf) as a linear combination of basis polynomials. If the density and basis polynomials are over the same field, any set of basis polynomials can describe the pdf; however, the most logical choice of polynomials is the family that is orthogonal with respect to the pdf. This problem is well-studied over the field of real numbers and has been shown to be valid for the complex unit circle in one dimension. The current framework for circular polynomial chaos is extended to multiple angular dimensions with the inclusion of correlation terms. Uncertainty propagation of heading angle and angular velocity is investigated using polynomial chaos and compared against Monte Carlo simulation.

Keywords: polynomial chaos; Szegő polynomials; directional statistics; Rogers-Szegő; state estimation

1. Introduction

Engineering is an imperfect science. Noisy measurements from sensors in state estimation [1,2], a constantly changing environment in guidance [3–5], 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. Ideally, every variable with some sort of uncertainty associated with it would be completely and analytically described with its probability density function (pdf). Unfortunately, even if this is feasible for the initialization of a random variable, its evolution through time rarely yields a pdf with an analytic form. If the pdf cannot be given in analytic form, then approximations and assumptions must be made.

In many cases, a random variable is quantified using only its first two moments—as with the unscented transform [7]—and a further assumption is that the distribution is Gaussian. In cases where the variable's uncertainty is relatively small and the dynamics governing its evolution are not highly nonlinear, this is not necessarily a poor assumption. 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. In this case, the error associated with using only the first two moments becomes significant [8,9].

One method of quantifying uncertainty that does not require an assumption of the random variable's pdf is polynomial chaos expansion (PCE) [10–14]. 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 [15]. The expression of the coordinate depends on the basis in 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 [16]. 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 [17]. 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.

Polynomial chaos has been used to quantify and propagate the uncertainty in nonlinear systems including fluid dynamics [18–21], orbital mechanics in multiple element spaces [22], and has been expanded to facilitate Gaussian mixture models [23]. 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. The variables that are of concern in this paper are angular variables. When estimating an angular random variable, such as the true anomaly of a body in orbit or vehicle attitude, a common approach is to assume that the variable is approximately equal to its projection on \mathbb{R}^n and use methods designed for variables on that space. When the uncertainty is very small, this approximation is relatively effective; however, as the uncertainty increases, this approximation becomes invalid. Recently, work on directional statistics has been incorporated into state estimation, providing a method for estimating angular random variables directly on the *n*-dimensional special orthogonal group (\mathbb{S}^n) [24–26]. Recent work [27] has shown that polynomial chaos can be used to estimate the mean and variance of one-dimensional angular random variables and the set of equinoctial orbital elements elements [28]. In neither case was a correlation estimated that included an angular state.

The sections of this paper include a detailed discussion of orthogonal polynomials in Section 2, including those that are orthogonal on the unit circle and two of the more common circular pdfs. In Section 3 an overview of polynomial chaos is given as well as the extension that has been made to include angular random variables, including the correlation between two angles. Finally, in Section 4, numerical results are presented comparing the correlation between two angular random variables calculated using Monte Carlo and PCE.

2. Orthogonal Polynomials

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 orthogonal basis is a subset of bases consisting of exactly *n* basis vectors such that the inner product between any two basis vectors, β_m and β_n , is proportional to the Kronecker delta (δ_{mn}). Given mathematically with angle brackets, this orthogonal inner product takes the following form:

$$\langle eta_m, eta_n
angle = c \, \delta_{mn}$$
 ,

where *m* and *n* are part of the set of positive integers, including zero. In the event c = 1, the set is termed orthonormal. The *i*th standard basis vector of \mathbb{V}^n is generally the *i*th vector of the *n*-dimensional identity; however, there are infinitely many bases for each vector space. It should be noted that it is not a requirement that a basis be orthogonal, merely linearly independent; however, the use of non-orthogonal bases is practically unheard-of.

An element $\alpha \in \mathbb{V}$ can be expressed in terms of an ordered basis $\mathcal{B} = \{\beta_1, \beta_2, ..., \beta_n\}$, as the linear combination

$$\alpha = a_1\beta_1 + a_2\beta_2 + \dots + a_n\beta_n \,, \tag{1}$$

where $[a_1, a_2, ..., a_n]$ 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. Consider a satellite in orbit. If the satellite's ground track is of high importance (such as weather or telecommunications satellites), an Earth-fixed frame would be ideal. However, in cases where a satellite's actions are

dictated by other space-based objects (such as proximity operations), a body-fixed frame would be ideal.

It is common to constrict the term vector space to the spaces that are easiest to visualize, most notably a Cartesian space, where the bases are vectors radiating from the origin at right angles. The term vector space is much more broad than this though. A vector space need only contain the zero element and be closed under both scalar addition and multiplication, which applies to much more than vectors.

Most notable in this work is the idea of a polynomial vector space. 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 $\mathcal{B} = \{1, x, ..., x^n\}$. The inner product with respect to the function ω on the real-valued polynomial space is given by

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

where $\omega(x)$ is a non-decreasing function with support \mathbb{S} and f and g are any two polynomials of degree n or less. A polynomial family $\Phi(x)$ is a set of polynomials with monotonically increasing order that are orthogonal. The orthogonality condition is given mathematically as

$$\langle \phi_m(x), \phi_n(x) \rangle_{\omega(x)} = \int_{\mathbb{S}} \phi_m(x) \phi_n(x) d\omega(x) = 0$$
 (2a)

$$\langle \phi_m^2(x) \rangle_{\omega(x)} = \int_{\mathbb{S}} \phi_m(x) \phi_m(x) d\omega(x) = c$$
, (2b)

where $\phi_k(x)$ is the polynomial of order k, c is a constant, and S is the support of the non-decreasing function $\omega(x)$. Note that while polynomials of negative orders (k < 0), referred to as Laurent polynomials, exist, they are not covered in this work.

The most commonly used polynomial families are categorized in the Askey scheme, which groups the polynomials based on the generalized hypergeometric function ($_pF_q$) from which they are generated [29–31]. Table 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. For completeness, Table 1 lists both continuous and discrete polynomial groups; however, the remainder of this work only considers continuous polynomials.

Туре	Polynomial	Hypergeometric Series	Support	Weight Function/Distribution
Continuous	Legendre	${}_{2}F_{1}$	[-1,1]	Uniform
	Jacobi	${}_{2}F_{1}$	[-1, 1]	Beta
	Laguerre	$_{1}F_{1}$	$[0,\infty)$	Exponential
	Probabilists' Hermite	${}_{2}F_{0}$	$(-\infty,\infty)$	Normal
Discrete	Charlier	$_{2}F_{0}$	$\{0, 1, 2, \dots\}$	Poisson
	Meixner	${}_{2}F_{1}$	$\{0, 1, 2, \dots\}$	Negative Binomial
	Krawtchouk	${}_{2}F_{1}$	$\{0, 1, \ldots, N\}$	Binomial
	Hahn	$_{3}F_{2}$	$\{0,1,\ldots,N\}$	Hypergeometric

Table 1. Common Orthogonal Polynomials.

2.1. Polynomials Orthogonal on the Unit Circle

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. These polynomials are orthogonal with respect to measures on the real line. In the event that a set of polynomials orthogonal with respect a measure on a curved interval (e.g., the unit circle) is desired, the Askey polynomials would be insufficient.

In [32], Szegő uses the connection between points on the unit circle and points on a finite real interval to develop polynomials that are orthogonal on the unit circle. Polynomials of this type are now known as Szegő polynomials. Since the unit circle is defined to have unit radius, every point can be described on a real interval of length 2π and mapped to the complex variable $\vartheta = e^{i\theta}$, where *i* is the imaginary unit. All use of the variable ϑ in the following corresponds to this definition. The orthogonality expression for the Szegő polynomials, $\phi_k(\vartheta)$, is

$$\langle \phi_m(\vartheta), \phi_n(\vartheta)
angle_{\omega(heta)} = rac{1}{2\pi} \int_{-\pi}^{\pi} \phi_m(\vartheta) \overline{\phi_n(\vartheta)} \omega(heta) d heta = \delta_{mn}$$
 ,

where $\overline{\phi_n(\vartheta)}$ is the complex conjugate of $\phi_n(\vartheta)$ and $\omega(\theta)$ is the monotonically increasing weight function over the support. Note that, as opposed to Equation (2a), the Kronecker delta is not scaled, implying all polynomials using Szegő's formulation are orthonormal.

While the general formulation outlined by Szegő is cumbersome—requiring the calculation of Fourier coefficients corresponding to the weight function and large matrix determinants—it does provide a framework for developing a set of polynomials orthogonal with respect to any conceivable continuous weight function. In addition to the initial research done by Szegő, further studies have investigated polynomials orthogonal on the unit circle [33–38].

Fortunately, there exist some polynomial families that are given explicitly, such as the Rogers-Szegő polynomials. The Rogers-Szegő polynomials have been well-studied [39–41] and were developed by Szegő based on work done by Rogers over the *q*-Hermite polynomials. For a more detailed description of the relationship between the Askey scheme of polynomials and their *q*-analog, the reader is encouraged to reference [31,42].

The generating function for the Rogers-Szegő polynomials is given as

$$\phi_n(\vartheta;q) = \sum_{k=0}^n \binom{n}{k}_q \vartheta^k, \qquad (3)$$

where $\binom{n}{k}_{q}$ is the *q*-binomial

$$\binom{n}{k}_{q} = \frac{(q;q)_{n}}{(q;q)_{k}(q;q)_{n-k}}$$
 and $(a;q)_{n} = \prod_{j=0}^{n-1} (1 - aq^{j}).$

The weight function that satisfies the orthogonality condition of these polynomials is

$$\omega(\theta) = \frac{1}{\sqrt{-2\pi\ln(q)}} \sum_{j=-\infty}^{\infty} \exp\left\{\frac{(\theta+2\pi j)^2}{2\ln(q)}\right\} \qquad 0 < q < 1.$$

$$\tag{4}$$

In addition to the generating function, a three-step recurrence [43] exists, which is given by

$$\phi_{n+1}(\vartheta;q) = (1+\vartheta)\phi_n(\vartheta;q) - (1-q^n)\vartheta\phi_{n-1}(\vartheta;q).$$
(5)

For convenience, the first five polynomials are:

$$\begin{split} \phi_0 &= 1 \\ \phi_1 &= \vartheta + 1 \\ \phi_2 &= \vartheta^2 + (q+1)\vartheta + 1 \\ \phi_3 &= \vartheta^3 + (q^2 + q + 1)\vartheta^2 + (q^2 + q + 1)\vartheta + 1 \\ \phi_4 &= \vartheta^4 + (q+1)(q^2 + 1)\vartheta^3 + (q+1)(q^2 + q + 1)\vartheta^2 + (q+1)(q^2 + 1)\vartheta + 1 \,. \end{split}$$

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. Distributions on the Unit Circle

With the formulation of polynomials orthogonal on the unit circle, the weight function $\omega(\theta)$ has been continuously mentioned but not specifically addressed. In the general case, the weight function can be any non-decreasing function; however, the most common polynomial families are those that are orthogonal with respect to well-known pdfs, such as the ones listed in Table 1. Because weight functions must exist over the same support as the corresponding polynomials, pdfs over the unit circle are required for polynomial orthogonal on the unit circle.

2.2.1. Von Mises Distribution

One of the most common distributions used in directional statistics is the von Mises/von Mises-Fisher distribution [44–46]. 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 distribution has extensions into higher dimensional spheres. The circular von Mises pdf is given as [24]

$$p_m(heta;\mu,\kappa) = rac{e^{\kappa\cos(\theta-\mu)}}{2\pi I_0(\kappa)}$$
 ,

where μ is the mean angular direction on a 2π interval (usually $[-\pi, \pi]$), $\kappa \ge 0$ is a concentration parameter (similar to the inverse of the standard deviation), and I_0 is the zeroth order modified Bessel function of the first kind. The reason this distribution is so common is its close similarity to the normal distribution. This can be seen in Figure 1a, where von Mises distributions of various concentration parameters are plotted.



Figure 1. Common Circular probability density functions (pdfs). (a) Circular von Mises distribution for multiple values of κ . (b) Wrapped normal distribution for multiple values of σ .

2.2.2. Wrapped Distributions

The easiest to visualize circular distribution, or rather group of distributions, that is discussed is the set of wrapped distributions. The wrapped distributions take a distribution on the real line and wrap it onto the unit circle according to

$$p_w(heta) = \sum_{j=-\infty}^{\infty} p(heta + 2\pi j)$$
 ,

where the support of $p(\cdot)$ is an interval of \mathbb{R} , and the domain of p_w is an interval on \mathbb{R} with length 2π . For example, wrapping a normal distribution takes the pdf

$$p_n(x;\mu,\sigma^2) = rac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-rac{(x-\mu)^2}{2\sigma^2}
ight\},$$

where the domain of *x* is \mathbb{R} , and μ and σ are the mean and standard deviation, respectively, and wraps it, resulting in the wrapped pdf (in this case wrapped normal)

$$p_{wn}(\theta;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \sum_{j=-\infty}^{\infty} \exp\left\{-\frac{(\theta-\mu+2\pi j)^2}{2\sigma^2}\right\},\,$$

where the domain of θ is an interval on \mathbb{R} with length 2π . Zero-mean normal distributions with varying values of σ are wrapped, with the results shown in Figure 1b.

Recall the weight function of the Rogers-Szegő polynomials in Equation (4). As the log function is monotonically increasing, the term $\log(1/q)$ increases monotonically as q decreases. Observing the extremes of q: as q approaches 1, $\log(1/q)$ approaches 0, and as q approaches 0, $\log(1/q)$ approaches ∞ . Letting $\log(1/q) = \sigma^2$ and $\mu = 0$, this becomes a zero-mean wrapped normal distribution.

It is clear from Figure 1 that both distributions described previously have strong similarities to the unwrapped normal distribution. Figure 1 also shows the difference in the standard deviation parameter. Whereas the wrapped normal distribution directly uses the standard deviation of the unwrapped distribution, the von Mises distribution is with respect to concentration parameter that is inversely related to the dispersion of the random variable. This makes the wrapped normal distribution slightly more intuitive when comparing with an unwrapped normal distribution.

2.2.3. Directional Statistics

The estimation of stochastic variables generally relies on calculating the statistics of that variable. Most notable of these statistics are the mean and variance, the first two central moments. For pdfs (p(x)) on the real line that are continuously integrable, the central moments are given as

$$\mu_1 = \int_{\mathbb{S}} x p(x) dx$$

$$\mu_k = \int_{\mathbb{S}} (x - \mu_1)^k p(x) dx \quad k \ge 2,$$
(6)

where S is the support of p(x). Although less utilized in general, raw moments are commonly used in directional statistics and are given as

$$ho_k = \int_{\mathbb{S}} x^k p(x) dx$$
 ,

or

$$ho_k = \int_{\mathbb{S}} artheta^k p(heta) d heta$$
 ,

where the slight distinction is that the integration variable is within the variable $\vartheta = e^{i\theta}$. In addition, mean direction and circular variance are not the first and second central moments [24]. Instead, both are calculated from the first moment's angle (θ_1) and length (R_1):

$$\theta_1 = \arg(\rho_1) = \tan^{-1} \frac{\operatorname{imag}(\rho_1)}{\operatorname{real}(\rho_1)}$$
(7a)

$$R_1 = \|\rho_1\|$$
, (7b)

where $\|\cdot\|$ is the l^2 -norm. From Mardia [24], the length can be used to calculate the circular variance V_1 and circular standard deviation σ_1 according to

$$V_1 = 1 - R_1$$
 (8a)

$$\sigma_1 = \sqrt{-2\ln(R_1)} \,. \tag{8b}$$

Effectively, as the length of the moment decreases, the concentration of the pdf about the mean direction decreases and the unwrapped standard deviation (USTD) increases. Note that while the subscript in Equations (7) and (8) is 1, there are corresponding mean directions and lengths associated with all moments; however, these are rarely used in applications.

3. Polynomial Chaos

At any given instance in time, the deviation of the estimate from the truth can be approximated as a Gaussian distribution centered at the mean of the estimate. The space of these mean-centered Gaussians is known as a Gaussian linear space [11]; when that space is closed (i.e., the distributions have finite second moments), it falls into the Gaussian Hilbert space \mathcal{H} . At this point, what is needed is a way to quantify \mathcal{H} , as this gives the uncertainty between the estimate and the truth. This can be achieved by projecting \mathcal{H} onto a complete set of orthogonal polynomials when those basis functions are evaluated at a random variable $\xi \in \mathcal{H}$. While the distribution at any point in time natively exists in \mathcal{H} , its projection onto the set of orthogonal polynomials provides a way of quantifying it by means of the ordered coordinates, as in Equation (1).

The homogeneous chaos [10] specifies ξ to be normally distributed with zero mean and unit variance (i.e., unit Gaussian), and the orthogonal polynomials to be the Hermite polynomials due to their orthogonality with respect to the standard Gaussian pdf [47]. Not only does this apply for Gaussian processes, but the Cameron-Martin theorem [48] says that this applies for any process with a finite second moment. Although the solution does converge as the number of orthogonal polynomials increases, further development has shown that, for different stochastic processes, certain basis functions cause the solution to converge faster [16], leading to the more general polynomial chaos (PC).

To begin applying this method mathematically for a general stochastic process, let a stochastic variable, ε , be expressed as the linear combination over an infinite-dimensional vector space, i.e.,

$$\varepsilon(x,\xi) = \sum_{k=0}^{\infty} \varepsilon_k(x) \Psi_k(\xi) , \qquad (9)$$

where $\epsilon_i(x)$ is the deterministic component and $\Psi_i(\xi)$ is an *i*th-order orthogonal basis function evaluated at, and orthogonal with respect to, the weight function, ξ . The polynomial families listed in Table 1 have been shown by Xiu [16] to provide convenient types of chaos based on their weight functions.

In general, the elements of the coordinate ($[\epsilon]$) are called the polynomial chaos coefficients. These coefficients hold deterministic information about the distribution of the random variable; for instance, the first and second central moments of ϵ can be calculated easily as

$$\mathbf{E}[\boldsymbol{\varepsilon}] = \boldsymbol{\mu}_1 = \boldsymbol{\varepsilon}_0 \tag{10a}$$

$$\mathbf{E}\left[(\varepsilon - \mathbf{E}[\varepsilon])^2\right] = \mu_2 = \sum_{k=1}^{\infty} \epsilon_k^2 \langle \Psi_k^2 \rangle_{p(\xi)}, \qquad (10b)$$

where E[] denotes expected value.

Now, let ε be an *n*-dimensional vector. Each of the *n* elements in ε are expanded separately; therefore, Equation (9) is effectively identical in vector form

$$\boldsymbol{\varepsilon}(\boldsymbol{x},\boldsymbol{\xi}) = \begin{bmatrix} \varepsilon^{(1)}(x^{(1)},\boldsymbol{\xi}) \\ \varepsilon^{(2)}(x^{(2)},\boldsymbol{\xi}) \\ \vdots \\ \varepsilon^{(n)}(x^{(n)},\boldsymbol{\xi}) \end{bmatrix} = \begin{bmatrix} \sum_{k=0}^{\infty} \varepsilon_{k}^{(1)}(x^{(1)}) \Psi_{k}(\boldsymbol{\xi}) \\ \sum_{k=0}^{\infty} \varepsilon_{k}^{(2)}(x^{(2)}) \Psi_{k}(\boldsymbol{\xi}) \\ \vdots \\ \sum_{k=0}^{\infty} \varepsilon_{k}^{(n)}(x^{(n)}) \Psi_{k}(\boldsymbol{\xi}) \end{bmatrix}$$

Because the central moments are independent, the mean and variance of each their calculations similarly do not change. In addition to mean and variance, the correlation between two random variables is commonly desired. With the chaos coefficients estimated for each random variable and the polynomial basis known, correlation terms such covariance can be estimated.

3.1. Covariance

Let the continuous variables *a* and *b* have chaos expansions

$$a(x,\xi) = \sum_{j=0}^{\infty} \alpha_j(x) \Psi_j(\xi) \qquad \text{and} \qquad b(z,\zeta) = \sum_{k=0}^{\infty} \beta_k(z) \Phi_k(\zeta). \tag{11}$$

The covariance between *a* and *b* can be expressed in terms of two nested expected values

$$\operatorname{cov}(a,b) = \operatorname{E}[(a - \operatorname{E}[a])(b - \operatorname{E}[b])],$$

the external of which can be expressed as a double integral yielding

$$\operatorname{cov}(a,b) = \int_{\mathbb{A}} \int_{\mathbb{B}} (a - \operatorname{E}[a])(b - \operatorname{E}[b]) db da , \qquad (12)$$

where \mathbb{A} and \mathbb{B} are the supports of *a* and *b* respectively. Substituting the expansions from Equation (11) into Equation (12) and acknowledging that the zeroth coefficient is the expected value gives

$$\operatorname{cov}(a,b) = \int_{\mathbb{A}} \int_{\mathbb{B}} (ab - a\mathrm{E}[b] - b\mathrm{E}[a] + \mathrm{E}[a]\mathrm{E}[b])dbda$$
$$= \int_{\mathbb{A}} \int_{\mathbb{B}} (ab - \beta_0 a - \alpha_0 b + \alpha_0 \beta_0)dbda$$
$$= -\alpha_0 \beta_0 + \int_{\mathbb{A}} \int_{\mathbb{B}} (ab)dbda$$
(13a)

$$= -\alpha_0 \beta_0 + \int_{\mathbb{X}} \int_{\mathbb{Z}} \sum_{j=0}^{\infty} \alpha_j(x) \Psi_j(\xi) \sum_{k=0}^{\infty} \beta_k(z) \Phi_k(\zeta) d\zeta d\xi \,. \tag{13b}$$

Note the change of variables between Equations (13a) and (13b). This is possible because the random variable and the weight function (a/ξ and b/ζ in this case) are over the same support. Additionally, the notation of the support variable is changed to be consistent with the integration variable.

As long as the covariance is finite, the summation and the integrals can be interchanged [49], giving a final generalized expression for the covariance to be

$$\operatorname{cov}(a,b) = \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \alpha_j(x) \beta_k(z) \int_{\mathbb{X}} \int_{\mathbb{Z}} \Psi_j(\xi) \Phi_k(\zeta) d\zeta d\xi \,. \tag{14}$$

In general, no further simplifications can be made; however, if the variables *x* and *z* are expanded using the same set of basis polynomials, then integration reduces to

$$\operatorname{cov}(a,b) = \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \alpha_k(x) \beta_j(z) \int_{\mathbb{X}} \Psi_k(\xi) \Psi_j(\xi) p(\xi) d\xi , \qquad (15)$$

containing a single variable with respect to the base pdf. Taking advantage of the basis polynomial orthogonality yields the following simple expression:

$$\operatorname{cov}(a,b) = \sum_{k=1}^{\infty} \alpha_k(x) \beta_k(z) \langle \Psi_k^2 \rangle_{p(\xi)} \,. \tag{16}$$

Combined with the variance, the covariance matrix of the 2 \times 2 system of *x* and *z* just discussed is given as

$$\boldsymbol{P} = \sum_{k=1}^{\infty} \begin{bmatrix} \alpha_k^2 & \alpha_k \beta_k \\ \alpha_k \beta_k & \beta_k^2 \end{bmatrix} \langle \Psi_k^2 \rangle_{p(\xi)} \,.$$

For an *n*-dimensional state, let ϵ be the $n \times \infty$ matrix for the *n*, theoretically infinite, chaos coefficients. Written generally, the covariance matrix in terms of a chaos expansion is

$$\boldsymbol{P} = \sum_{k=1}^{\infty} \boldsymbol{\epsilon}_k \boldsymbol{\epsilon}_k^T \langle \boldsymbol{\Psi}_k^2 \rangle_{p(\boldsymbol{\xi})} \,.$$

In cases where orthonormal polynomials are used, the polynomial inner product disappears completely leaving only the summation of the estimated chaos coefficients

$$\boldsymbol{P} = \sum_{k=1}^{\infty} \boldsymbol{\epsilon}_k \boldsymbol{\epsilon}_k^T \,. \tag{17}$$

3.2. Coefficient Calculation

The two most common methods of solving Equation (9) for the chaos coefficients are sampling-based and projection-based. The first, and most common, approach requires truncating the infinite summation in Equation (9) to yield

$$\varepsilon(x,\xi) = \sum_{k=0}^{N} \epsilon_k(x) \Psi_k(\xi) , \qquad (18)$$

where the truncation term N, which depends on the dimension of the state n and the highest order polynomial p, is

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

Drawing *Q* samples of ξ , where Q > N, and evaluating Ψ_k and ε at these points effectively results in randomly sampling ε directly. 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 unknowns that describe the pdf of ε after the transformation that is given by

$$\begin{aligned} \varepsilon(x,\xi_1) &= \varepsilon_0(x)\Psi_0(\xi_1) + \varepsilon_1(x)\Psi_1(\xi_1) + \dots + \varepsilon_N(x)\Psi_N(\xi_1) \\ \varepsilon(x,\xi_2) &= \varepsilon_0(x)\Psi_0(\xi_2) + \varepsilon_1(x)\Psi_1(\xi_2) + \dots + \varepsilon_N(x)\Psi_N(\xi_2) \\ &\vdots \\ \varepsilon(x,\xi_Q) &= \varepsilon_0(x)\Psi_0(\xi_Q) + \varepsilon_1(x)\Psi_1(\xi_Q) + \dots + \varepsilon_N(x)\Psi_N(\xi_Q) \end{aligned}$$

This overdetermined system can be solved for ϵ using a least-squares approximation. The coefficients can then be used to calculate convenient statistical data about ϵ (e.g., central and raw moments).

While the sampling-based method is more practical to apply, the projection based method is not dependent on sampling the underlying distribution. Projecting the pdf of ε onto the *j*th basis yields

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

The inner product is with respect to the variable ξ ; therefore, the coefficient ϵ acts as a scalar. The inner product is linear in the first argument; therefore, the summation coefficients can be removed from the inner product without alteration, that is

$$\left\langle \varepsilon(x,\xi), \Psi_j(\xi) \right\rangle_{p(\xi)} = \sum_{k=0}^{\infty} \epsilon_k(x) \left\langle \Psi_k(\xi), \Psi_j(\xi) \right\rangle_{p(\xi)} \,. \tag{19}$$

In contrast, if the summation is an element of the second argument, the linearity condition still holds; however, the coefficients incur a complex conjugate. Recall the basis polynomials are generally chosen to be orthogonal, so the right-hand inner product of Equation (19) reduces to the scaled Kronecker delta, resulting in

$$\begin{split} \left\langle \varepsilon(x,\xi), \Psi_j(\xi) \right\rangle_{p(\xi)} &= \sum_{k=0}^{\infty} \epsilon_k(x) \left\langle \Psi_k(\xi), \Psi_j(\xi) \right\rangle_{p(\xi)} \\ &= \sum_{k=0}^{\infty} \epsilon_k(x) c \, \delta_{kj} \,. \end{split}$$

This leaves only the *j*th term (with the constant $c = \langle \Psi_j^2(\xi) \rangle_{p(\xi)}$), and an equation that is easily solvable for ϵ_j is

$$\epsilon_{j}(x) = \frac{\left\langle \epsilon(x,\xi), \Psi_{j}(\xi) \right\rangle_{p(\xi)}}{\langle \Psi_{j}^{2}(\xi) \rangle_{p(\xi)}}$$
(20a)

$$=\frac{\int_{\mathbb{Z}}\varepsilon(x,\xi)\Psi_{j}(\xi)dp(\xi)}{\int_{\mathbb{Z}}\Psi_{i}^{2}(\xi)dp(\xi)},$$
(20b)

which almost always requires numeric approximation.

3.3. Implementation Procedure

For convenience, the procedure for estimating the mean and covariance of a random state is given in Algorithm 1. Let ε be the state of a system with uncertainty defined by mean *m* and covariance *P* subject to a set of system dynamics over the time vector *T*. The algorithm outlines the steps required to estimate the mean and covariance of the state after the amount of time specified by *T*.

Algorithm 1 Estimation of mean and covariance using a polynomial chaos expansion.				
1: procedure PCE_EST(m_0, P_0)	▷ Estimation of <i>m</i> and <i>P</i> using PCE			
2: for $k = 1$ to T do	-			
3: Draw samples of ξ based on chaos type	Either randomly or intelligently			
4: $\varepsilon_{k-1} \leftarrow m_{k-1}, P_{k-1}, \xi$	\triangleright Sample ε based on ξ			
5: ε_k from propagating ε based on state dynamics	_			
6: $\epsilon \leftarrow \text{Equation} (20a)$				
7: $m_k \leftarrow \bar{\text{E}}$ quation (10a)				
8: $P_k \leftarrow \text{Equation (17)}$				
9: end for				
10: return <i>m</i> , <i>P</i>				
11: end procedure				

3.4. Complex Polynomial Chaos

While polynomial chaos has been well-studied and applied to a various number of applications in \mathbb{R}^n , alterations must be made for the restricted space \mathbb{S}^n due to its circular nature. A linear approximation can be made with little error when a circular variable's uncertainty is small; however, as the uncertainty increases, the linearization can impose significant error. Figure 2 shows the effects of projecting two wrapped normal distributions with drastically different standard deviations onto a tangent plane. The two wrapped normal distributions are shown in Figure 2a,b, with USTDs of 0.25 and 0.4 rad, respectively. Clearly, even relatively small USTDs result in approximately uniform wrapped pdfs.

One of the most basic projections is an orthogonal projection from an *n*-dimensional space onto an n - 1 dimensional plane. In this case, the wrapped normal pdf is projected orthogonally onto the plane (1, x, z), which lies tangent to the unit circle at the point (1,0), coinciding with the mean direction of both pdfs. The plane, and the projection of the pdf onto this plane are shown in Figure 2c,d. Approximating the circular pdf as the projected planar pdf comes with an associated loss of information. At the tangent point, there is obviously no information loss; however, when the physical distance from the original point to the projected point is considered, the error associated with the projected point increases. As is the case with many projection methods concerning circular and spherical bodies, all none of the information from the far side of the body is available in the projection. The darkness of the shading in all of Figure 2 comes from the distance of the projection where white is no distance, and black is a distance value of least one (implying the location is on the hemisphere directly opposite the mean direction).

To better indicate the error induced by this type of projection, Figure 2e,f also include a measure that shows how far the pdf has been projected as a percentage of the overall probability at a given point. At the tangent point, there is no projection required, therefore the circular pdf has to be shifted 0% in the *x* direction. As the pdf curves away from the tangent plane, the pdf has to be projected farther. The difference between Figure 2e and Figure 2f is that the probability approaches zero nearing $y = \pm 1$ in Figure 2e; therefore, the effect of the error due to projection is minimal. In cases where the information is closely concentrated about one point, tangent plane projections can be good assumptions. Contrarily, in Figure 2f the pdf does not approach zero, and therefore the approximation begins to become invalid. Accordingly, the red error line approaches the actual pdf, indicating that the majority of the pdf has been significantly altered in the projection.



Figure 2. Error induced by approximating a circular distribution as linear, tangent at the mean. (a) Wrapped normal distribution with pdf $p_{wn}(\theta; 0, 0.25^2)$. (b) Wrapped normal distribution approaching wrapped uniform distribution with pdf $p_{wn}(\theta; 0, 2.4^2)$. (c) Projection of $p_{wn}(\theta; 0, 0.25^2)$ onto a plane tangent to the unit circle. (d) Projection of $p_{wn}(\theta; 0, 2.4^2)$ onto a plane tangent to the unit circle. (e) Error associated with projecting $p_{wn}(\theta; 0, 0.25^2)$ onto a plane tangent to the unit circle. (f) Error associated with projecting $p_{wn}(\theta; 0, 2.4^2)$ onto a plane tangent to the unit circle.

In addition to restricting the space to the unit circle, most calculations required when dealing with angles take place in the complex field. In truth, the bulk of expanding polynomial chaos to be suitable for angular random variables is generalizing it to complex vector spaces. Previous work by the authors [27] has shown that a stochastic angular random variable can be expressed using a polynomial

chaos expansion. Specifically, the chaos expansion is one that uses polynomials that are orthogonal with respect to probability measures on the complex unit circle as opposed to the real line.

3.5. Szegő-Chaos

For the complex angular case, the chaos expansion is transformed slightly, such that

$$\varepsilon(x,\vartheta) = \sum_{k=0}^{\infty} \varepsilon_k(x) \overline{\Psi_k(\vartheta)}, \qquad (21)$$

where, once again, $\vartheta = e^{i\theta}$. The complex conjugate is not required in Equation (21), but it must be remembered that the expansion must be projected onto the *conjugate* of the expansion basis in Equation (20b). While ultimately a matter of choice, it is more convenient to express the expansion in terms of the conjugate basis, rather than the original basis.

Unfortunately, while the first moment is calculated the same for real and complex valued polynomials, the real valued process does not extend to complex valued polynomials. This is because of the slightly different orthogonality condition between real and complex valued polynomials. While the inner product given in Equation (2a) is not incorrect, it is only valid for real valued polynomials. The true inner product of two functions contains a complex conjugate, that is

$$\langle \Psi_m, \Psi_n \rangle_{p(x)} = \int_{\mathbb{X}} \Psi_m(x) \overline{\Psi_n(x)} p(x) dx = c \, \delta_{mn} \, .$$

The difference between $\mathbb{R}[x]$ and $\mathbb{C}[x]$ is that the complex conjugate has no effect on $\mathbb{R}[x]$. Fortunately, the zeroth polynomial of the Szegő polynomials is unitary just like the Askey polynomials. The complex conjugate has no effect; therefore the zeroth polynomial has no imaginary component and is calculated the same for complex and purely real valued random variables.

The complex conjugate of a real valued function has no effect; therefore, the first moment takes the form;

$$\mu_1 = \sum_{k=0}^{\infty} \epsilon_k(x) \int_{\mathbb{X}} \Psi_k(\xi) \Psi_0(\xi) p(\xi) d\xi = \sum_{k=0}^{\infty} \epsilon_k(x) \int_{\mathbb{X}} \Psi_k(\xi) \overline{\Psi_0(\xi)} p(\xi) d\xi.$$
(22)

In general, calculation of the second raw moment and the covariance cannot be simplified beyond

$$\mu_{2} = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \epsilon_{j} \epsilon_{k} \int_{\mathbb{X}} \overline{\Psi_{j}(\xi) \Psi_{k}(\xi)} p(\xi) d\xi$$
$$= \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \epsilon_{j} \epsilon_{k} \left\langle \overline{\Psi_{j}(\xi)}, \Psi_{k}(\xi) \right\rangle_{p(\xi)}$$
(23)

$$\operatorname{cov}(x,z) = \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \alpha_j \beta_k \int_{\mathbb{X}} \int_{\mathbb{Z}} \overline{\Psi_j(\zeta) \Phi_k(\zeta)} p(\zeta,\zeta) d\zeta d\zeta \,.$$
(24)

The simplification from Equation (14) to Equation (15) as a result of shared bases can similarly be applied to Equation (24). This simplifies Equation (24) to a double summation but only a singular inner product (i.e., integral), i.e.,

$$\operatorname{cov}(x,z) = \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \alpha_j \beta_k \int_{\mathbb{X}} \overline{\Psi_j(\xi)} \Psi_k(\xi) p(\xi) d\xi$$
$$= \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \alpha_j \beta_k \langle \overline{\Psi_j(\xi)}, \Psi_k(\xi) \rangle_{p(\xi)}.$$

The familiar expressions for the second raw moment given in Equation (10b) and the covariance given in Equation (16) are special cases for $\mathbb{R}[x]$ rather than general expressions.

3.6. 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 Equation (21). The Rogers-Szegő polynomials in Equation (3) can be rewritten according to [39]

$$\phi_n\left(-\frac{\vartheta}{\sqrt{q}},q\right) = \sum_{k=0}^n (-1)^{n-k} \binom{n}{k}_q q^{\frac{n-k}{2}} \vartheta^k, \qquad (25)$$

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

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

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

$$\vartheta_3(\alpha,\beta) = \sum_{k=-\infty}^{\infty} \beta^{k^2} e^{2ik\alpha} \,, \tag{26}$$

which is another form of the wrapped normal distribution. Note the distinction between the variables ϑ_3 and $\vartheta = e$.

For convenience, the inverse to the given 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 cumulative distribution function (cdf) is required to draw random samples. The number of wrappings in Equation (26) significantly affects the results. For reference, the results presented in this work truncate the summation to ± 1000 .

Written out, the first five orders of this form of the Rogers-Szegő polynomials are

$$\begin{split} \phi_0 &= 1 \\ \phi_1 &= \vartheta - q^{1/2} \\ \phi_2 &= \vartheta^2 - q^{1/2}(q+1)\vartheta + q \\ \phi_3 &= \vartheta^3 - q^{1/2}(q^2+q+1)\vartheta^2 + q(q^2+q+1)\vartheta - q^{3/2} \\ \phi_4 &= \vartheta^4 - q^{1/2}(q+1)(q^2+1)\vartheta^3 + q(q^2+1)(q^2+q+1)\vartheta^2 - q^{3/2}(q+1)(q^2+1)\vartheta + q^2 \,, \end{split}$$

and are shown graphically in Figure 3. Because the polynomials are complex valued, the real and imaginary components are show, separately. In both cases, the polynomials are oscillatory, with the real component being symmetric about $\theta = 0$, and the imaginary component being antisymmetric about $\theta = 0$. Additionally, the amplitude of the oscillations increase both with increasing order and distance from $\theta = 0$.

20

10





 ϕ_1

 ϕ_0

(b)

Figure 3. The zeroth through fourth Rogers-Szegő polynomials with an unwrapped standard deviation of 0.1. (**a**) Real Component. (**b**) Imaginary Component.

The zeroth polynomial is one, as is standard; therefore, the difference between the two generating functions given in Equations (25) and (26) will only be apparent in the calculation of moments beyond the first.

3.7. Function Complexity

As is to be expected, the computational complexity increases with increasing state dimension. It is therefore of interest to develop an expression that bounds the required number of function evaluations as a function of number of states and expansion order. 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 $x \in \mathbb{S}^{P}$ be a *P*-dimensional state vector consisting of angular variables, and let $q \in \mathbb{N}^{P}$ be the expansion order of each element in x, where \mathbb{N} is the set of natural numbers, including zero. The number of inner products required to calculate the chaos coefficients in Equation (20b) 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 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 Equation (23) 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 Equation (24) between the *i*th and *j*th elements requires $(q_i + 1)(q_j + 1)$ additional evaluations for a raw moment and q_iq_j for a central moment. Combining these into one expression, the generalized number of inner product evaluations for raw moments with $P \ge 2$ 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, the use of orthonormal polynomials reduces the coefficient calculation inner products by two, and expansions using real valued polynomials do not require these inner product calculations at all.

4. Numerical Verification and Discussion

To test the estimation methods outlined in Section 3.5, a system with two angular degrees of freedom is considered. The correlated, nonlinear dynamics governing this body, the initial mean directions ϕ/θ , initial USTDs, and constant angular velocities $\dot{\phi}/\dot{\theta}$ are given in Table 2.

Table 2. Initial conditions and governing equations of the dynamical system in Section 4.

Angle	Mean Direction	USTD	Angular Velocity (const.)	Dynamics
$egin{array}{c} \phi \ heta \end{array} eta \end{array}$	$\pi/30$ [rad] $\pi/4$ [rad]	0.1 [rad] 0.1 [rad]	0.2 [rad/s] 0.2 [rad/s]	$\begin{array}{l} \Delta \phi = \sqrt{ \theta_{k-1} } \dot{\phi} \Delta T \\ \Delta \theta = \theta_{k-1} \dot{\theta} \Delta T \end{array}$

For every simulation, the run time is 4 s with ΔT being 0.05 s; this equates to 81 time steps in each simulation. In Figure 4, the joint pdf propagates from the initial conditions (bottom left) to the final state (top right). The initial joint pdf clearly reflects an uncorrelated bivariate wrapped normal distribution. After being transformed by the dynamics, the final joint pdf exhibits not just translation and scaling, but also rotation: indicating a non-zero correlation between the two angles, which is desired.



Figure 4. Evolution of the joint pdf from uncorrelated to correlated bivarite wrapped normal pdf.

For a practical application, the mean and standard deviation/variance of each dimension, as well as the covariance between dimensions is desired. When dealing with angles, the mean direction and the standard deviation can be obtained from the first moment, omitting the second moment. Therefore, only the first moment and the covariance will be discussed. Recall the equations for the first moment and covariance are in terms of chaos coefficients and are given generally in Equations (22) and (24). Because two angles are being estimated, the supports of the integrals in Equations (22) and (24) are set as $[-\pi, \pi)$: but it should be noted that the support is not rigidly defined this way, the only requirement is that the support length be 2π .

Rather than exploit the computational efficiency of methods such as quadrature integral approximations on the unit circle [50–52], the integrals are computed as Riemann sums. Therefore, it is necessary to determine an appropriate number of elements that provides an adequate numerical approximation, while remaining computationally feasible.

Figure 5 show the settling effect that decreasing the elemental angle has on the estimation of the covariance. Note that this figure is used to show the sensitivity of the simulation to the integration variable rather than the actual estimation of the covariance, which will be discussed later in this section. Both plots show the relative error of each estimate with respect to a Monte Carlo simulation of the joint system described in Table 2. Clearly, as the number of elements increases, the estimates begin to converge until the difference between $d\theta = 0.01$ rad (629 elements) and $d\theta = 0.005$ rad (1257 elements) is perceivable only near the beginning of the simulation. Because of this, it can reasonably be assumed that any error in the polynomial chaos estimate with $d\theta = 0.005$ is not attributed to numerical estimation of the integrals in Equations (22) and (24). Additionally, these results should also indicate the sensitivity of the estimate to the integration variable. Even though the dynamics used in this work's examples result in a joint pdf that somewhat resembles wrapped normal pdfs, the number of elements used in the integration must still be quite large. The final numerical element that must be covered is the Monte Carlo. For these examples, 5×10^7 samples are used in each dimension.



Figure 5. Effects of integration variable on estimated covariance relative error. Each line represents a different elemental angle ranging from 0.04 to 0.005 rad. (a) Angle Component. (b) Length Component.

In each of the examples in Section 4.1, the polynomial chaos estimate first evaluates the Rogers-Szegő polynomials at each of the 1257 uniformly distributed angles (ξ), solves Equation (20b) for the chaos coefficients, and uses Equations (22) and (24) to estimate the mean and covariance. After this, the 1257 realizations of the state ($\varepsilon(\xi)$) are propagated forward in time according to the system dynamics. At each time step the system is expanded using polynomial chaos to estimate the statistics.

The estimations of the first moment and covariance of the system described by the simulation parameters in Table 2 are shown in Figures 6 and 7. In both cases, the angle and length of the estimate are presented, rather than the exponential form used in the polynomial chaos expansion. This representation much more effectively expresses the mean and concentration components of the estimate, which are directly of interest when examining statistical moments.



Figure 6. Estimation of the first moment of both angles using a fifth order polynomial chaos expansion compared against Monte Carlo simulation. (**a**) Angle Component. (**b**) Length Component.

Beginning with the mean estimate from a fifth order polynomial chaos expansion in Figure 6, the mean direction of the angle θ is nearly identical to the Monte Carlo estimate, while the estimate of ϕ begins to drift slightly as the simulation progresses. Of the two angles, this makes the most since; recalling Table 2, only the dynamics governing ϕ are correlated with θ , the dynamics governing θ are only dependent on θ . In comparison, the estimates of the lengths are both much closer to the Monte Carlo result. Looking closely at the end of the simulation, it can be seen that, again, θ is practically identical, and there is some small drift in ϕ downwards, indicating that the estimate reflects a smaller concentration. Effectively, the estimation of the mean reflects some inaccuracy; however, this inaccuracy is partly reflected in the larger dispersion of the pdf.

Similarly to the mean, a small drift can be seen in the estimate of the covariance in Figure 7. In both cases the initial estimate is nearly identical to the Monte Carlo result; however, throughout the simulation a small amount of drift becomes noticeable. While this drift is undesirable, the general tracking of the polynomial chaos estimate to the Monte Carlo clearly shows that the correlation between two angles can be approximated using a polynomial chaos expansion.



Figure 7. Estimation of the covariance between the two angles using a fifth order polynomial chaos expansion compared against Monte Carlo simulation. (**a**) Angle Component. (**b**) Length Component.

4.1.1. Unwrapped Standard Deviation and Joint PDF Assumptions

From the discussion of the generating function for the Rogers-Szegő polynomials Equation (25), it is clear that these polynomials are dependent on the USTD. 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 Equation (23), which describes the calculation of the covariance, requires the knowledge of the joint pdf between the two random variables. In practice, there is no reasonable way of obtaining this pdf; and if there is, then the two variables are already so well know, that costly estimation methods are irrelevant.

It is therefore of interest to investigate what effects making assumptions about the USTD and the joint pdf have on the estimates. The basis polynomials are evaluated when solving for the chaos coefficients Equation (20b) and when estimating the statistical moments Equations (22)–(25) at every time step. If no assumption is made about the USTD, then the generating function in Equation (25) or the three step recursion in Equation (5) 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 Equation (25) can be made.

While only used in the estimation of the covariance, a simplification of the joint pdf will also significantly reduce computation and increase the feasibility of the problem. The most drastic of simplifications is to use the initial, uncorrelated joint pdf. 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 will not be effected by any movement of the mean.

Assuming the USTD to be fixed at 0.1 for both random variables and the joint pdf to be stationary throughout the simulation led to estimates that are within machine precision of the unsimplified results in Figures 6 and 7. This is to be expected when analyzing Askey-chaos schemes (like Hermite-chaos) that are problem invariant. In instances where the USTD of the wrapped normal distribution is

low enough that probabilities at $\pm \pi$ are approximately zero, the wrapped normal distribution is effectively a segment of the unwrapped normal distribution, because the probabilities beyond $\pm \pi$ are approximately equal to zero. However, in problems where the USTD increases, the wrapped normal distribution quickly approaches a wrapped uniform distribution, this makes the time-invariant USTD a poor assumption. While a stationary USTD assumption may not hold as well for large variations in USTD, highly correlated, or poorly modeled, dynamical systems, it shows that some assumptions and simplifications can be made to ensure circular polynomial chaos is a practical method of estimation.

4.1.2. Chaos Coefficient Response

The individual chaos coefficients are not always inspected for problems using Askey-chaos simply due to the commonality of Askey-chaos. The adaptation of polynomial chaos to use the Szegő polynomials, and thus expanding from real to complex valued estimates presents a case that warrants inspection of the chaos coefficients. Figure 8 show the time evolution of the first 13 chaos coefficients (including the zeroth coefficient) that describe the random variable ϕ .



Figure 8. Time evolution of the first 13 chaos coefficients describing the random variable ϕ . (a) Real Coefficient Evolution. (b) Imaginary Coefficient Evolution. (c) Imaginary and Real Coefficient Evolution. (d) Complex Coefficient Evolution.

What becomes immediately apparent is that the coefficients are roughly anti-symmetrically paired until the length of the coefficient begins to approach zero. In this specific case, the eighth coefficient in Figure 8 initiates this trend. This is the first coefficient that does not have an initial estimate coinciding with lower ordered coefficients. All coefficients following this one show very little response to the system. This is to be expected. Recall the calculation of the chaos coefficient includes the product of polynomial and pdf as well as a division by the self inner product of each polynomial order (i.e., $\langle \Psi_k^2(\zeta) \rangle_{p(\zeta)}$). The polynomial and pdf product have opposing behaviors when approaching $\pm \pi$ from 0. Whereas the polynomial oscillation amplitude increases, the tails of the pdf approach probability values of zero. This ensures the growth in the higher order polynomials is mitigated.

For brevity, only the coefficients from the variable ϕ are shown. These have a much more interesting response than θ due to the nature of the dynamics. The most notable part of the coefficients from θ is that none of the coefficients ever move beyond the complex unit circle, which from Figure 8c, is clearly not the case for ϕ . In fact, the coefficients describing θ stay close to the complex unit circle and just move clockwise about it. Similarly, the eighth and higher coefficient lengths begin collapse to zero rad. For this problem (and presumably most problems) almost all of the information is coming from the first two coefficients. Comparing the estimates using two, three, and ten coefficients yields the same results to within machine precision. This is not surprising when considering the inner products (Table 3) that are required to estimate the covariance; each of the inner products are effectively zero when compared with $\langle \phi_0, \phi_0 \rangle_{p_{wn}}$ and $\langle \phi_1, \phi_1 \rangle_{p_{wn}}$. While having to only compute two significant chaos coefficients makes computation easier, it also limits the amount of information that is used in the estimate; however, for simple problems such as this one, two significant coefficients are satisfactory.

$\left\langle \psi_{i},\psi_{j} ight angle _{p_{wn}}$	j = 0	j = 1	<i>j</i> = 2	j = 3
i = 0	1.00	-4.77e-16-3.15e-15i	-3.35e-15-2.00e-16i	-1.21e-16-1.65e-16i
i = 1	-4.77e-16+3.15e-15i	-0.01 - 1.94e - 16i	1.94e-4+3.06e-16i	-5.71e-06-3.27e-17i
i = 2	-3.35e-15+2.00e-16i	1.94e-4-3.06e-16i	1.84e-4+4.61e-17i	-1.09e-05-3.70e-17i
i = 3	-1.21e-16+1.65e-16i	-5.71e-06+3.27e-17i	-1.09e-05+3.70e-17i	-4.53e-06-2.45e-16i

Table 3. Rogers-Szegő Inner Products.

5. Conclusions and Future Work

One method of quantifying the uncertainty of a random variable is a polynomial chaos expansion. For variables that exist only on the real line, this type of expansion has been well studied. This work developed the alterations that must be made for a polynomial chaos expansion to be valid for random variables that exist on the unit circle, specifically the complex unit circle (where the *y* coordinate becomes imaginary).

Previous work has shown that polynomial chaos can be used with the Rogers-Szegő polynomials to estimate the raw moments of a random variable with a wrapped normal distribution. A generalized set of expressions for the mean and covariance of multi-dimensional systems for both real and complex systems has been presented that do not make the assumption that each variable has been expanded with the same set of basis polynomials. An example of two angular random variables—one with correlated dynamics, and one without—has been presented. The mean of each random variable as well as the covariance between them is estimated and compared with Monte Carlo estimates. In the case of the uncorrelated random variable, the mean estimates are highly accurate. For the correlated random variable, the estimate is found to slowly diverge from the Monte Carlo result. A similar small divergence is observed in the covariance estimate; however, the general trend is similar enough to indicate the error is not in the formulation of the complex polynomial chaos. Additionally, an approximation to the basis polynomials and time-varying joint probability density function (pdf) is made, without loss of accuracy in the estimate. From the estimates of the mean and covariance, it is clear that the Rogers-Szegő polynomials can be used as an effective basis for angular random

variable estimation. However, for more complex problems, different polynomials should be considered, specifically polynomials with an appropriate number of non-negligible self inner products.

Author Contributions: Conceptualization, C.S.; Methodology, C.S.; Formal Analysis, C.S.; Writing—Original Draft Preparation C.S.; Writing—Review & Editing, K.J.D.; Supervision, K.J.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Graduate Assistance in Areas of National Need fellowship.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

- CDF Cumulative distribution function
- PCE Polynomial chaos expansion
- PDF Probability density function
- USTD Unwrapped standard deviation

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.; Patrick, B. Pose Measurement Performance of the Argon Relative Navigation Sensor Suite in Simulated-Flight Conditions. In Proceedings of the AIAA Guidance, Navigation, and Control Conference, Minneapolis, MN, USA, 13–16 August 2012.
- 3. Latella, C.; Lorenzini, M.; Lazzaroni, M.; Romano, F.; Traversaro, S.; Akhras, M.A.; Pucci, D.; Nori, F. Towards Real-Time Whole-Body Human Dynamics Estimation Through Probabilistic Sensor Fusion Algorithms. *Auton. Robot.* **2019**, *43*, 1591–1603. [CrossRef]
- 4. Lubey, D.P.; Scheeres, D.J. Supplementing state and dynamics estimation with information from optimal control policies. In Proceedings of the 17th International Conference on Information Fusion (FUSION), Salamanca, Spain, 7–10 July 2014; pp. 1–7.
- Imani, M.; Ghoreishi, S.F.; Braga-Neto, U.M. Bayesian Control of Large MDPs with Unknown Dynamics in Data-Poor Environments. In *Advances in Neural Information Processing Systems 31*; Curran Associates, Inc.: Red Hook, NY, USA, 2018; pp. 8146–8156.
- 6. Hughes, D.L. *Comparison of Three Thrust Calculation Methods Using In-Flight Thrust Data;* NASA TM-81360; NASA: Washington, DC, USA, 1981; Unpublished.
- 7. Julier, S.; Uhlmann, J.; Durrant-Whyte, H.F. A New Method for the Nonlinear Transformation of Means and Covariances in Filters and Estimators. *IEEE Trans. Autom. Control* **2000**, *45*, 477–482. [CrossRef]
- 8. Wu, C.C.; Bossaerts, P.; Knutson, B. The Affective Impact of Financial Skewness on Neural Activity and Choice. *PLoS ONE* **2011**, *6*, e16838. [CrossRef] [PubMed]
- 9. Anderson, T.; Mattson, C. Propagating Skewness and Kurtosis Through Engineering Models for Low-Cost, Meaningful, Nondeterministic Design. *J. Mech. Des.* **2012**, *134*, 100911. [CrossRef]
- 10. Wiener, N. The Homogeneous Chaos. Am. J. Math. 1938, 60, 897–936. [CrossRef]
- 11. Janson, S. *Gaussian Hilbert Spaces;* Cambridge Tracts in Mathematics; Cambridge University Press: New York, NY, USA, 1997; Chapter 1.2.
- 12. Alexanderian, A. Gaussian Hilbert Spaces and Homogeneous Chaos: From theory to applications. 2013; Unpublished.
- Eldred, M. Recent Advances in Non-Intrusive Polynomial Chaos and Stochastic Collocation Methods for Uncertainty Analysis and Design. In Proceedings of the 50th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, Palm Springs, CA, USA, 4–7 May 2009; Volume 4, pp. 2078–2114.
- 14. Ng, L.; Eldred, M. Multifidelity Uncertainty Quantification Using Non-Intrusive Polynomial Chaos and Stochastic Collocation. In Proceedings of the 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference, Honolulu, HI, USA, 23–26 April 2012; Volume 9, pp. 7669–7685.

- Savin, E.; Faverjon, B. Higher-Order Moments of Generalized Polynomial Chaos Expansions for Intrusive and Non-Intrusive Uncertainty Quantification. In Proceedings of the 19th AIAA Non-Deterministic Approaches Conference, Grapevine, TX, USA, 9–13 January 2017; pp. 215–223.
- 16. Xiu, D.; Karniadakis, G.E. The Wiener-Askey Polynomial Chaos for Stochastic Differential Equations. *SIAM J. Sci. Comput.* **2003**, *24*, 619–644. [CrossRef]
- Schmid, C.L.; DeMars, K.J. Minimum Divergence Filtering Using a Polynomial Chaos Expansion. In Proceedings of the AAS/AIAA Astrodynamics Specialist Conference, Stevenson, WA, USA, 20–24 August 2017.
- Xiu, D.; Karniadakis, G.E. Modeling Uncertainty in Flow Simulations via Generalized Polynomial Chaos. J. Comput. Phys. 2003, 187, 137–167. [CrossRef]
- Hosder, S.; Walters, R.W.; Perez, R. A Non-Intrusive Polynomial Chaos Method for Uncertainty Propagation in CFD Simulations. In Proceedings of the 44th AIAA Aerospace Sciences Meeting and Exhibit, Reno, NV, USA, 9–12 January 2006.
- 20. Hosder, S.; Walters, R.W.; Balch, M. Point-Collocation Non-Intrusive Polynomial Chaos Method for Stochastic Computational Fluid Dynamics. *AIAA J.* **2010**, *48*, 2721–2730. [CrossRef]
- Hosder, S.; Walters, R.W. Non-Intrusive Polynomial Chaos Methods for Uncertainty Quantification in Fluid Dynamics. In Proceedings of the 48th AIAA Aerospace Sciences Meeting, Orlando, FL, USA, 4–7 January 2010; pp. 1580–1595.
- 22. Jones, B.A.; Doostan, A.; Born, G.H. Nonlinear Propagation of Orbit Uncertainty Using Non-Intrusive Polynomial Chaos. J. Guid. Control. Dyn. 2013, 36, 430–444. [CrossRef]
- 23. Vittaldev, V.; Russell, R.P.; Linares, R. Spacecraft Uncertainty Propagation Using Gaussian Mixture Models and Polynomial Chaos Expansions. *J. Guid. Control. Dyn.* **2016**, *39*, 2615–2626. [CrossRef]
- 24. Mardia, K.; Jupp, P. *Directional Statistics*; Wiley Series in Probability and Statistics; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2009.
- 25. Markley, F.L. Attitude Error Representations for Kalman Filtering. J. Guid. Control. Dyn. 2003, 26, 311–317. [CrossRef]
- 26. Darling, J.E. Bayesian Inference for Dynamic Pose Estimation Using Directional Statistics. Ph.D. Thesis, Missouri University of Science and Technology, Rolla, MO, USA, 2016.
- 27. Schmid, C.L.; DeMars, K.J. Polynomial Chaos Confined to the Unit Circle. In Proceedings of the AAS/AIAA Astrodynamics Specialist Conference, Snowbird, UT, USA, 19–23 August 2018; pp. 2239–2256.
- Jones, B.A.; Balducci, M. Uncertainty Propagation of Equinoctial Elements Via Stochastic Expansions. In Proceedings of the John L. Junkins Dynamical Systems Symposium, College Station, TX, USA, 20–21 May 2018.
- Andrews, G.E.; Askey, R. Classical Orthogonal Polynomials. Polynômes Orthogonaux et Applications; Brezinski, C., Draux, A., Magnus, A.P., Maroni, P., Ronveaux, A., Eds.; Springer: Berlin/Heidelberg, Germany, 1985; pp. 36–62.
- Askey, R.; 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; Chapter 1.
- 31. Koekoek, R.; Lesky, P.A.; Swarttouw, R.F. *Hypergeometric Orthogonal Polynomials and Their q-Analogues*; Springer Monographs in Mathematics; Springer: Berlin/Heidelberg, Germany, 2010.
- 32. Szegő, G. *Orthogonal Polynomials*; American Mathematical Society Colloquium Publications, American Mathematical Society: Providence, RI, USA, 1959; Volume 23, Chapter 11, pp. 287–295.
- 33. Simon, B. *Orthogonal Polynomials on the Unit Circle (Colloquium Publications);* American Mathematical Society: Providence, RI, USA, 2005; Volume 54.
- 34. Simon, B. Orthogonal Polynomials on the Unit Circle: New Results. *arXiv* 2004, arXiv:math.SP/math/0405111.
- 35. Geronimus, Y.L. Orthogonal Polynomials: Estimates, Asymptotic Formulas, and Series of Polynomials Orthogonal on the Unit Circle and on an Interval. *Math. Gaz.* **1962**, *46*, 354–355. [CrossRef]
- 36. Ismail, M.E.; Ruedemann, R.W. Relation Between Polynomials Orthogonal on the Unit Circle with Respect to Different Weights. *J. Approx. Theory* **1992**, *71*, 39–60. [CrossRef]
- 37. Jagels, C.; Reichel, L. On the Construction of Szegő Polynomials. *J. Comput. Appl. Math.* **1993**, *46*, 241–254. [CrossRef]

- 38. Jones, W.B.; Njåstad, O.; Thron, W.J. Moment Theory, Orthogonal Polynomials, Quadrature, and Continued Fractions Associated with the Unit Circle. *Bull. Lond. Math. Soc.* **1989**, *21*, 113–152. [CrossRef]
- 39. Atakishiyev, N.M.; Nagiyev, S.M. On the Rogers-Szegő Polynomials. J. Phys. Math. Gen. 1994, 27, L611. [CrossRef]
- 40. Hou, Q.; Lascoux, A.; Mu, Y. Continued Fractions for Rogers-Szegő Polynomials. *Numer. Algorithms* **2004**, 35, 81–90. [CrossRef]
- 41. Vinroot, C.R. Multivariate Rogers-Szegő Polynomials and Flags in Finite Vector Spaces. *arXiv* 2010, arXiv:math.CO/1011.0984.
- 42. Szabłowski, P.J. On *q*-Hermite Polynomials and Their Relationship With Some Other Families of Orthogonal Polynomials. *Demonstr. Math.* **2011**, *46*, 679–708. [CrossRef]
- 43. Andrews, G. *The Theory of Partitions*; Cambridge Mathematical Library, Cambridge University Press: Cambridge, UK, 1998.
- 44. Fisher, R. Dispersion on a Sphere. *Proc. R. Soc. London. Ser. Math. Phys. Eng. Sci.* **1953**, 217, 295–305. [CrossRef]
- 45. Watson, G.S.; Williams, E.J. On the Construction of Significance Tests on the Circle and the Sphere. *Biometrika* **1956**, *43*, 344–352. [CrossRef]
- 46. Mardia, K.V. Distribution Theory for the von Mises-Fisher Distribution and Its Application. In *A Modern Course on Statistical Distributions in Scientific Work;* Patil, G.P., Kotz, S., Ord, J.K., Eds.; Springer: Dordrecht, The Netherlands, 1975; pp. 113–130.
- 47. Bogachev, V. *Gaussian Measures;* Mathematical Surveys and Monographs, American Mathematical Society: Providence, RI, USA, 2015.
- 48. Cameron, R.H.; Martin, W.T. The Orthogonal Development of Non-Linear Functionals in Series of Fourier-Hermite Functionals. *Ann. Math.* **1947**, *48*, 385–392. [CrossRef]
- 49. Fubini, G. Sugli Integrali Multipli: Nota; Reale Accademia dei Lincei: Rome, Italy, 1907.
- Cruz-Barroso, R.; Mendoza, C.D.; Perdomo-Pío, F. Szegő-Type Quadrature Formulas. J. Math. Anal. Appl. 2017, 455, 592–605. [CrossRef]
- 51. Bultheel, A.; González-Vera, P.; Hendriksen, E.; Njåstad, O. Orthogonal Rational Functions and Quadrature on the Unit Circle. *Numer. Algorithms* **1992**, *3*, 105–116. [CrossRef]
- Daruis, L.; González-Vera, P. Szegő Polynomials and Quadrature Formulas on the Unit Circle. *Appl. Numer. Math.* 2001, 36, 79–112.. [CrossRef]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).