Spring 1-1-2018

Orbit Uncertainty Propagation with Separated Representations

Marc Balducci
University of Colorado at Boulder, marc.balducci@colorado.edu

Recommended Citation

Balducci, Marc, "Orbit Uncertainty Propagation with Separated Representations" (2018). *Aerospace Engineering Sciences Graduate Theses & Dissertations*. 236.
https://scholar.colorado.edu/asen_gradetds/236

Orbit Uncertainy Propagation with Separated Representations

by

Marc Balducci

B.S., Rensselaer Polytechnic Institute, 2011

M.S., University of Colorado at Boulder, 2018

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Aerospace Engineering Sciences

2018

This thesis entitled:
Orbit Uncertainy Propagation with Separated Representations
written by Marc Balducci
has been approved for the Department of Aerospace Engineering Sciences

Prof. Brandon A. Jones

Prof. Alireza Doostan

Date ___________________

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Balducci, Marc (Ph.D., Aerospace Engineering Sciences)

Orbit Uncertainy Propagation with Separated Representations

Thesis directed by Prof. Brandon A. Jones

In light of recent collisions and an increasing population of objects in Earth orbit, the space situational awareness community has significant motivation to develop novel and effective methods of predicting the behavior of object states under the presence of uncertainty. Unfortunately, approaches to uncertainty quantification often make simplifying assumptions in order to reduce computation cost. This thesis proposes the method of separated representations (SR) as an efficient and accurate approach to uncertainty quantification. The properties of an orthogonal polynomial basis and a uni-directional least squares regression approach allow for the theoretical computation cost of SR to remain low when compared to Monte Carlo or other surrogate methods. Specifically, SR does not suffer from the curse of dimensionality, where computation cost increases exponentially with respect to input dimension. Benefits of this low computation cost are shown in a series of low Earth orbit test cases, where SR is used to accurately approximate non-Gaussian posterior distribution functions. Here, the dimension of the problem is increased from 6 to 20 without incurring significantly more computation time. Taking advantage of a large input dimension, this research presents a global sensitivity analysis computed via SR, which affords a more nuanced analysis of a previously examined case in the literature. By considering design variables, SR is formulated to perform optimization under uncertainty. A novel method that utilizes a Brent optimizer to create training data at unique times of closest approach is devised and implemented in order to detect low probability collision events. This methodology is leveraged to design an optimal avoidance maneuver, which would be intractable when using traditional Monte Carlo. Lastly, a multi-element algorithm is formulated and presented to estimate solutions that are challenging for unmodified SR. This multi-element SR leads to orders of magnitude in accuracy improvement when considering the ability of unmodified SR to approximate discontinuous, multimodal, or diffuse solutions.

Dedication

To my parents

Acknowledgements

The material for the work by Marc Balducci is provided by the NSTRF fellowship, NASA Grant NNX15AP41H. Additional funding is provided by the GAANN fellowship as well as TA positions in the Department of Aerospace Engineering Sciences of the University of Colorado at Boulder. The work presented in Chapters 4 and 7 utilizes code found in the py-orthpol Python package. This code can be downloaded at https://github.com/PredictiveScienceLab/py-orthpol

With regards to personal acknowledgements, I would be unable to thank anyone if I didn't first recognize my parents, Louis and Tamera, and the efforts that they put into raising and supporting me throughout the years. Although many people have positively influenced my life to this point, there are a certain few with whom our interactions have afforded me the motivation to accomplish such goals as the work contained within this thesis. It would be impossible to ignore my good friend Eduardo Villalba who was always willing to talk about more than just the numbers. It is from discussions with you, among my most valued of friends, that I have learned valuable lessons. Specifically, it is not the amount of achievements or posessions that one strives for or earns, it is the amount of genuine significance that is attached to each of these things. A million suns won't fill you up, if you can't see the wine flowing over your cup.

Contents

Chapter

1 Introduction 1

Appendix

Tables

Table

Figures

Figure

Chapter 1

Introduction

1.1 Motivation

Space situational awareness (SSA) requires that an accurate estimate of a space object's state and uncertainty be known. This estimation is a major component in conjunction assessment and computing the probability of collision, or reacquiring an object to facilitate a state update. Collisions between catalogued objects [2, 71] has spurred investigations into typical methodologies of SSA and increased interest in research for novel approaches to this problem. Research in this field becomes increasingly important as the ratio of objects in space to telescopes on the ground increases, which is expected with improving sensor technologies [68]. To meet demand, SSA analyses must accurately propagate uncertainties over increasingly long time spans. Unfortunately, these involve nonlinear dynamics that cause the position and velocity multivariate probability distribution function (PDF) to possibly become highly non-Gaussian.

While SSA is traditionally viewed as restricted to the realm of Earth orbit, state awareness is required in many environments. As the number of objects orbiting Mars grows, assessments of their current and future states becomes necessary to avoid conflicts. State awareness and spacecraft navigation also faces challenges in complex outer planet and small body systems such as asteroids and comets [95, 93]. In the case of small bodies, the gravity may be highly asymmetric and relatively unknown. Therefore if left uncontrolled for a short period, spacecraft may deviate significantly from their planned trajectories or even impact the surface [95]. Complex planetary systems such as the Jovian or Saturnian moon system also present challenges in maneuver planning and state estimation or prediction. These environments often have many sources of perturbations, which complicate analyses such as planetary protection. In addition to this, time between flybys of interest may be so long that the ability to utilize satellite dynamics in order to minimize fuel use is limited [93].

Calculating the probability of collision between space objects necessitates an accurate assessment of SSA, which in turn provides a representation of the uncertainty of a space object's state. This estimation process becomes increasingly difficult and important as the ratio of objects in space to telescopes on the ground increases, which is expected with improving sensor technologies [83]. In order to identify and therefore avoid potential collisions, methods must be developed such that the state of objects in orbit can be accurately determined and the probability of collision estimated. This thesis seeks to aid the field of conjunction assessment by providing an efficient means of propagating uncertainty without posterior assumptions, from which the probability of collision can be computed. The developed methodology is then applied to an optimization problem focused on avoiding the identified conjunction, thereby tractably performing a design that is robust to uncertainty.

1.2 Uncertainty Quantification Background

Methods of uncertainty quantification (UQ) seek to estimate the variability of a system response due to input and modeling errors. This variability results from, for example, force model truncation and statistical uncertainties in inputs. One way of quantifying the system response is to generate realizations of random variables and propagate each to a final state (e.g. Monte Carlo simulation). Using this method, the statistical characteristics of the PDF can be analyzed but at a large computational cost. However, a recent assessment of the Air Force Space Command's astrodynamic standards indicates a need for such sampling-based methods for UQ in certain scenarios [83]. Therefore, if the large computation cost of standard methods, such as Monte Carlo (MC) simulation, is to be avoided, a more computationally efficient procedure is of interest.

Traditionally, uncertainty mapping using the state transition matrix (STM) is used to estimate a posterior PDF [101]. As an alternative, the statistical method of MC may be used [96, 33].

Unfortunately, each method has disadvantages. The accuracy of Monte Carlo is known to be inversely proportional to the square root of the number of samples used. The result of which is that incremental improvements in the accuracy require significant increases in sample size and computation costs. On the other hand, the STM relies on a linearization scheme, which is undesirable in the nonlinear regime of orbit propagation [70]. More recently, the unscented transform (UT) has been used as an efficient uncertainty propagator due to its ability to do so nonlinearly. The UT, like the STM, still assumes a Gaussian distribution and *a posteriori* Gaussian assumptions have been proven to be inaccurate under common conditions [70]. Cases of high variance, significant time between observations, or both tend to yield non-Gaussian posterior PDFs. Avoiding Gaussian assumptions in high-dimensional stochastic systems is therefore important for proposed uncertainty propagation techniques if the ratio of objects in space to sensors continues to increase.

Methods such as polynomial chaos expansions (PCEs) [50, 122, 68, 67, 66] and Gaussian Mixture Methods (GMM) [62, 30, 32], which do not assume Gaussian distributions, have been proposed as an alternative to current approaches [68]. While these new methods provide an improvement in efficiency, computation time still increases quickly (up to exponential) with respect to the number of uncertain inputs or stochastic dimensions [36, 62]. This effect has been dubbed the *curse of dimensionality*, the mitigation of which requires dimensional reduction or truncation, while inaction leads to increased computation time [20, 36, 62]. State transition tensors (STT) are also being researched as a method to nonlinearly propagate state uncertainty [85, 46, 80]. These STT methods, however, require the derivation of complex partial derivatives or numerical methods for their approximation.

In addition to solution statistics, it is often desired to determine the (global) *sensitivity* of the solution of interest or its statistics with respect to uncertain inputs. A sensitivity analysis quantifies the relative effects that random inputs have on uncertainty in quantities of interest (QOIs) [99]. The result of such an analysis can be represented via a sensitivity index based on the analysis of variance (ANOVA). These indices are helpful in a variety of circumstances, including practical (rather than mandatory) dimension truncation or prioritizing dimension determination. If the goal is to optimize reliability or minimize solution variance, among other desires, it is helpful to identify the input that is most deserving of better determination in order to reduce the output uncertainty by the largest amount possible [99]. The effects of the curse of dimensionality, however, can limit the scope of sensitivity analyses by prematurely truncating stochastic dimension count for the sake of such an analysis. Although PC has been shown to efficiently compute global sensitivity indices [106, 21], the efficient generation of a PC expansion lies in the assumption of low stochastic dimension or low polynomial degree.

1.3 Conjunction Assessment

Modern conjunction assessment (CA) requires the calculation of the probability of collision. The calculation of this value includes a methodology of UQ due to the fact that the risk of collision is a statistical event as long as the conjunction is not considered to be a deterministic process. As research devoted to CA has progressed, methodologies have been introduced and developed over time. Early attempts to determine a conjunction focused on the distance of closets approach between two nominal trajectories. Due to the fact that this approach does not consider uncertainty, the astrodynamics community later developed methods incorporating uncertainty in the form of probability ellipsoids. The probability of collision could then be calculated via overlapping areas or volumes [53]. Evolving from this, methods have been developed to incorporate these uncertainty volumes or areas, as well as the shape of objects, often in the guise of hard body radii [4, 18, 86]. Although such approaches have proven themselves in operational environments, collisions do still occur [87, 2, 71]. Typically, the established methods make assumptions such as posterior Gaussian PDFs or constant velocity near the time of closest approach [89, 87, 88]. Oftentimes, these assumptions are made for the sake of computational efficiency. Therefore, it is desirable to provide a posterior estimation technique that is tractable and lessens the number of assumptions. Both GMM [31, 45] and PC have been applied to CA, with PC currently used in operation on the Magnetospheric Multiscale (MMS) Mission [68, 67, 66, 100]. Results have shown that these methods provide improved efficiency in UQ and estimating the probability of collision. However as previously stated, both GMM and PC suffer from the *curse of dimensionality.*

As an alternative to distance of closest approach and probability of collision, Patera has proposed a quantity known as the *probability of conflict* [87, 88]. Used within the aviation community, the conflict probability is quantified by calculating the probability that the uncertainty of a primary object will enter a conflict volume of a predetermined shape. A cylinder is commonly used in aviation, as altitude rarely changes, while Patera developed a method compatible with spheres and ellipsoids. These shapes are more applicable to astrodynamics, as spacecraft are often in orbits which are not perfectly circular. When applied to publicly available operational conjunction data collected over a calendar year, it was found that the probability of conflict reliably predicted possible collisions while mitigating the chance of a false positive when compared to more common methods such as the probability of collision or maximum probability of collsion [87]. Although a promising quantity, the probability of conflict is still calculated while assuming Gaussian posteriors.

1.4 Optimization Application

Significant work in the astrodynamics community has undertaken the task of designing optimal maneuvers for spacecraft, with a focus on a deterministic optimization problem. The Lambert algorithm is familiar to the field, while recent work has helped address a miriad of scenarios [113]. For example, Izzo et al. [64] uses search space pruning to optimize a multiple gravity assist problem. In Pontani et al. [92] and Pontani and Conway [91], particle swarms are utilized for optimizing multiple burn rendezvous trajectories as well as continuous finite thrust rendezvous trajectories in the rotating Euler-Hill frame, respectively. In order to avoid the high computation cost of a brute force optimization problem, e.g., one that relies on evaluations of a black box propagator, Peng et al. [90] and Gano et al. [47] both use surrogates, with Gano et al. [47] comparing three methodologies. Although these approaches are relevant for the stated deterministic problems, they are not robust to system uncertainty.

Complex systems are often subject to uncertainty or stochastics due to incomplete models or inadequate state knowledge. Once this uncertainty has been quantified, a deterministic optimization scheme does not adequately account for the statistical nature of the system. The methodology of optimization under uncertainty (OUU) accounts for such uncertainty. OUU enables solving an optimization problem while also taking into consideration the uncertainty of random inputs and the resulting variability in a performance or reliability function [73]. In order to optimize a problem, OUU supports the concept of *reliability design* as well as *robust design* [38]. Reliability design focuses on finding an optimal solution that is constrained by a minimum allowed reliability index. This index is based on a cumulative or complementary cumulative distribution function. Robust design, on the other hand, emphasizes a constraint on the maximum allowed variance in a design.

OUU has recently become a topic of interest in the astrodynamics community. Utilizing linear covariance propagation for UQ has seen a number of applications, while more advanced methods of UQ (such as the unscented transform or state transition tensors) have also recently been applied [82, 79, 125, 124, 123]. As an example of a robust maneuver design, Patera and Peterson [89] considers uncertainty in a maneuver design methodology intended to avoid a potential collision. The cited work covers an operational algorithm that is used to identify conjunction events and quantify the risk of collision via probability of collision. The approach employs a two step method for calculating a maneuver which avoids the identified possible collision. This process involves first determining the gradient of the probability of collision with respect to maneuver direction, and then, using a one dimensional root finding scheme, the optimal maneuver magnitude is computed [89]. Although the methodology provides an improvement in computational efficiency, it maintains a Gaussian posterior assumption. In the case of Mueller and Larsson [82], the robust optimization technique is applied to collision avoidance via maneuver design. Robust rendezvous planning is addressed by Louembet et al [78] and Deaconu et al. [28] while respectively considering maneuver error and navigation uncertainty. Robust rendezvous planning is again addressed by Dell'Elce and Kerschen [29], where differential flatness is applied to the optimization approach. Considering the various approaches to UQ, a review of the techniques therefore becomes of interest when applying OUU to astrodynamics.

1.5 Proposed Methodology

As an alternative to the discussed methods, we present the approach of *separated representations* (SR) for propagating uncertainties associated with the initial state of a satellite and other parameters of an orbital system to a future time. SR, also known as canonical decompositions (CANDECOMP) or parallel factor analysis (PARAFAC), provides a surrogate model to efficiently quantify the response of a system to a set of inputs. The main idea behind SR is to decompose a multivariate function of inputs into a sum of products of univariate functions of those inputs. First introduced by Hitchcock in 1927, SR has been applied to several areas including data mining, machine learning [20] and chemistry [6, 58]. Recent work by [19, 40] shows that SR may be an effective algorithm for estimating a function of many variables, while alleviating the curse of dimensionality, i.e., SR shows promise to significantly reduce computation times for high-dimension stochastic systems [35, 34, 84, 72, 36, 24, 54, 110, 94]. SR provides a means to propagate uncertainty with nonlinear models where no assumptions of a Gaussian *a posteriori* distribution are made.

First applied to astrodynamics by [16] as part of this work, SR has also been used to produce a direct solution of the Fokker-Planck equation for perturbed Keplerian mechanics [107]. The expectation is that, similar to studies in the literature, computation costs of orbit UQ will decrease with respect to several other methods that fall prey to the curse of dimensionality [36]. Theoretically, as demonstrated by [20], the computational complexity of SR remains linear with respect to the number of random inputs. With this fact in mind, the available suite of random or stochastic dimensions can be expanded without significantly increasing computation time. This work presents the mathematical tools required to use SR in astrodynamics, provides an initial proof of concept for efficiently propagating orbit uncertainty and uses an SR to efficiently perform sensitivity analyses.

With the ability to efficiently propagate uncertainty, SR is proposed as a method for quantifying the conjunction probability between objects in space. Once a statistically significant conjunction has been identified, a natural next step in the process is to address this concern and attempt to avoid the collision. Designing a maneuver for one (or more) of the involved objects is one such approach. PC has been used to perform OUU, and it has been shown to be both efficient and accurate when compared to previously used methods [73, 39]. Feldehacker et al. [41] applies PC to a mission design problem by utilizing the reduced computational burden of a surrogate within an OUU problem. An assumption when using PC for OUU, however, is that the system is accurately represented by a stochastic dimension that is low enough in order to avoid the *curse of dimensionality.* Therefore, SR has an expected benefit of reducing the computation time of OUU under high stochastic dimension when compared to PC. In this thesis, we consider probability of collision test cases with a dimension up to 18. This would be theoretically expensive for methods that suffer from the curse of dimensionality.

1.6 Multi Element Improvement

The unmodified SR is most effective when estimating smooth, continuous functions. However, this behavior is not guaranteed in many fields such as astrodynamics. Additionally, multimodal posterior PDFs are challenging for polynomial surrogate methods to achieve a converged solution, while diffuse PDFs with large variances also present issues. Multimodal behavior has been observed in Molniya orbits along with expected difficulties in creating surrogate approximations [67]. Flybys also pose estimation issues, and multimodal distributions have been documented while running simulations of the Cassini mission [42]. This poses a risk to mission survival or planetary protection, where it may be required to perform a tractable estimation of a non-Gaussian PDF, but it is possible for a surrogate method to fail to converge. An additional problem is that of resonant dynamics, where a “keyhole” problem subjects a very small number of samples to potential impacts. This is seen in scenarios such as asteroid flybys or discarded rocket stage bodies [119].

Unstable resonant orbits have increasingly been used for tour design around outer planets with moons. An application that has been of interest recently includes Jovian petit grand tours with an emphasis on missions to Europa. Anderson and Lo showed that resonant orbits are integral to understanding the dynamics of flybys for ballistic [11], impulsive [10], and low-thrust trajectories [9]. They have also been used directly in the design of tours and endgame scenarios with their invariant manifolds and heteroclinic connections [12, 7, 115, 116], as components in the overall tour design strategy [23], and with optimization [75]. Resonant orbits have also been shown to be key to understanding elements of the final approach to moons such as Europa [8, 13]. The instability of these perturbed environments, in particular after flybys, can establish a discontinuous or multi-modal relationship between the input and solution space. Therefore, computationally efficient tools that are able to accurately converge onto such solutions are desirable.

As it is a similar methodology, we are able to look to research on PCEs for additional techniques and enhancements in capability. For example, general PCEs have been modified to include compressive sampling, which may reduce the required training sample count by eliminating insignificant coefficients [56]. Although applying this approach to SR is not a topic of this research, we focus on other research aimed at enhacing the performance of PCEs. Specifically, those made to improve upon the reliance on a continuous relationship between random inputs and the solution space. In an effort to accurately estimate a discontinuous solution, the Wiener-Haar method was developed [76, 77]. The approach of this cited work relies on a wavelet basis rather than the orthogonal polynomial method previously discussed. Therefore, in order to develop a method that maintains the simple polynomial formulation of PCEs, Wan and Karniadakis [120] proposes a multi-resolution support. This approach is a significant improvement due to the ability for a multi-element generalized polynomial chaos method (ME-gPC) to remain accurate when considering discontinuous functions or random inputs, while maintaining the same PCE solver methodology with few modifications. These characteristics are equally desirable for an SR approximation, and we seek a similar improvement in performance while remaining within the basic framework of polynomial formulation.

1.7 Contributions and Overview

The contents of this thesis contain the background necessary to provide motivation for the research of SR, its formulation and implementation, as well as various applications to the field of astrodynamics. These applications are run on a number of test cases, and the accuracy of the surrogate technique is verified and compared to reference methodolgies. Taken together, this study presents work comprised of the following contributions.

This work includes the first application of SR to the field of astrodynamics and considers both relatively low and high dimension count. The efficacy of the method is quantified in both computation cost and accuracy. Leveraging the benefits that SR provides to computation time, a surrogate based sensitivity analysis is applied to two test cases. This is the first example of using SR to compute Sobol indices for a global sensitivity analysis. The results of one of these analyses provides additional insight to the relationship between input uncertainty and the variance of the QoIs in a previously examined test case in [62]. A formulation for using SR with OUU is then presented, which is also a new contribution. In a first for the astrodynamic community, the process of using SR with OUU is later applied to a probability of collision case, where a maneuver is designed to avoid two collisions. For the calculation of this probability quantity, a novel method, which incorporates a one dimensional optimizer, is presented. Following the formulation for OUU, a multi-element approach is discussed and formulated. Due to the multi-element formulation being a unique contribution of this thesis, implementation guidelines are given in order to faciliate further development. For the first time, a multi-element SR surrogate is applied to an astrodynamic test case and the results compared to a single element SR, thus highlighting the benefits of this new method.

The remaining chapters are used to introduce the formulation and theory of SR, discuss using SR with OUU, introduce the formulation of multi-element SR, and present results on low Earth orbit (LEO) test cases, an OUU application, as well as cases highlighting the increased performance generated by a multi-element approach. Within Chapter 2, Section 2.5.2 introduces the method of SR, its formulation and an implementation guideline. Following this, Section 2.5.3 presents formulations for solution statistics and discusses an approach to sensitivity analysis. Chapter 3 examines the concept of OUU, and the formulation of SR needed for OUU is established in Section 3.3. Chapter 4 discusses the general weakness of SR to capture multi-modal PDFs, introduces a formulation to overcome this, and includes a rough implementation algorithm.

In order to establish the performance of SR, its applications and any improvements to the general methodology, a series of test cases are evaluated. Chapter 5 presents three test cases that analyze the computation cost of SR, its ability to approxiate joint PDFs, and its capability of performing sensitivity analyses. Following this, Chapter 6 considers an OUU application using the formulation derived in Chapter 3. For this, we apply SR to a case where the probability of collision must be calculated. Therefore, Chapter 6 contains sections on the calculation of this quantity, the case that is considered, and the performance of SR when performing OUU on the problem. Lastly, Chapter 7 takes the theory and implementation guidelines of Chapter 4 and applies this methodology to a series of test cases. The accuracy of the multi-element approach is compared to that of an unmodified SR, thereby proving both the requirement and improvement in estimation of the multi-element method.

Chapter 2

Uncertainty Quantification

This chapter presents a more thorough exploration of the concept of uncertainty quantification (UQ) than dicussed in Chapter 1. Here, we cover some of the more common methods of UQ, their typical applications and defining characteristics. In addition to this, more recent developments in UQ, as applied to astrodynamics, are introduced. Among these, surrogate methods are presented, which include polynomial chaos expansions (PCEs) and separated representations (SR). The latter is the focus of this work.

Therefore, a significant portion of this chapter is devoted to the theoretical description of SR and the formulations needed for implementation. An algorithm for converging on a solution is given, and, in addition to this, analytic expressions for the first two moments are derived, as well as a method for calculating sensitivity indices. The latter is a topic which will be covered more broadly at first, so that the advantages of a surrogate appproach can be appreciated.

2.1 Problem Setup and Objective

Consider $(\Omega, \mathcal{F}, \mathcal{P})$ to be a suitable probability space on which the random inputs for a given system are defined. The sample space Ω contains the set of elementary events, while the probability measure $\mathcal{P}$ exists on the σ-algebra $\mathcal{F}$. The random inputs are represented by the vector quantity $\boldsymbol{y}(\omega) = (y_1(\omega), \ldots, y_d(\omega)) : \Omega \rightarrow \mathbb{R}^d$, $\omega \in \Omega$ and $d \in \mathbb{N}$. Additionally, the stochastic function $\boldsymbol{q}(\boldsymbol{y})$ exists on this probability space. For the purposes of an astrodynamics application of UQ, the state of the satellite is usually the quantity of interest (QoI) of the stochastic function $\boldsymbol{q}(\boldsymbol{y}) \in \mathbb{R}^M$, and it is assumed to depend on d random variables $\boldsymbol{y} \in \mathbb{R}^d$. Here, M is the dimension of the QoI, and it is indexed via $m = 1, \ldots, M$. For the case of Cartesian position and velocity, $M = 6$, for example. These inputs characterize the uncertainty in the initial state and possibly other parameters such as the coefficient of drag.

The state of the satellite at time t is denoted as $\boldsymbol{q}(t, \boldsymbol{y})$ and satisfies a set of ODEs

$$\boldsymbol{\mathcal{X}}\left(t, \boldsymbol{y}; \boldsymbol{q}\right) = \mathbf{0} \tag{2.1}$$

describing the temporal evolution of the satellite state. In this scenario, $t \in [t_0, t_f]$ is the temporal variable and the initial condition

$$\boldsymbol{q}(t_0, \boldsymbol{y}) = \boldsymbol{q}_0(\boldsymbol{y}) \tag{2.2}$$

is considered. For the interest of a cleaner presentation, the temporal dependence of $\boldsymbol{q}(\boldsymbol{y})$ is restricted to a fixed instance of t and only the short notation of $\boldsymbol{q}(\boldsymbol{y})$ is use henceforth.

2.2 Traditional Techniques

Although a number of methodologies are used within the astrodynamics community, we focus on two examples here. That is, we present Monte Carlo, as well as linearization and the state transition matrix. These two techniques are chosen as they represent differing approaches, i.e., the Monte Carlo method utilizes fully non-linear dynamics, makes no posterior assumptions and relies on the law of large numbers to derive statistics from a sampling method, while the linearization approach assumes a Gaussian posterior, with mean and standard deviation being direct properties of this assumption. This choice of either full non-linear simulations or linearized propagation models is representative of contemporary uncertainty quantification in the astrodynamics community.

2.2.1 Monte Carlo

In the case of Monte Carlo (MC) simulation, the statistics of $\boldsymbol{q}(\boldsymbol{y})$ are estimated by calculating multiple realizations of $\boldsymbol{q}(\boldsymbol{y})$, organized in the set $\{\boldsymbol{q}\left(\boldsymbol{y}_j\right)\}_{j=1}^{N_{MC}}$, from the initial state condition $\{(\boldsymbol{y}_j, \boldsymbol{q}_0\left(\boldsymbol{y}_j\right))\}_{j=1}^{N_{MC}}$, where $\boldsymbol{y}_j$ denotes a realized sample of $\boldsymbol{y}$. Similarly, the vector $\boldsymbol{y} \in \mathbb{R}^d$ is indexed by $i = 1, \ldots, d$, which results in the j-th sample of direction i being indicated by $y_{i,j}$. The initial samples are generated by applying various *a priori* statistical distributions to the initial conditions. For example, many astrodynamics applications assume a Gaussian *a priori* distribution. Therefore, samples at t_0 are generated via $\boldsymbol{q}_0 \sim \mathcal{N}(\bar{\boldsymbol{q}}_0, \boldsymbol{\Sigma})$, with a given mean vector $\bar{\boldsymbol{q}}_0$ and covariance matrix $\boldsymbol{\Sigma}$. Both the mean and covariance are intended to be dictated by an informed decision such as measurement data or hardware specifications. Each sample $\boldsymbol{y}_j$ is first mapped to an initial sample $\boldsymbol{q}_0(\boldsymbol{y}_j)$ via a Cholesky decomposition of $\boldsymbol{\Sigma}$. Specifically,

$$\boldsymbol{q}_0\left(\boldsymbol{y}_j\right) = \bar{\boldsymbol{q}}_0 + \boldsymbol{L}\boldsymbol{y}_j, \quad j = 1, \ldots, N, \tag{2.3}$$

where $\boldsymbol{\Sigma} = \boldsymbol{L}\boldsymbol{L}^{\mathrm{T}}$, and y_i has the probability density function

$$\rho_i = \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{(x-\mu_i)^2}{2\sigma_i^2}}, \tag{2.4}$$

with (in the case of this work) $\mu_i = 0$ and $\sigma_i = 1$ for all values of i. Similarly, if the statistical distribution considered is uniform, initial samples are generated as $\boldsymbol{q}_0 \sim \mathcal{U}(\boldsymbol{q}_0^{lower}, \boldsymbol{q}_0^{upper})$, where the inputs are the lower and upper bounds of the desired initial distribution. If we consider direction i and $y_i \sim \mathcal{U}(0, 1)$,

$$q_{0,i}\left(y_{i,j}\right) = q_{0,i}^{lower} + (q_{0,i}^{upper} - q_{0,i}^{lower})y_{i,j}, \quad j = 1, \ldots, N, \tag{2.5}$$

where $q_{0,i}$ is the initial quantity related to the i-th direction, and $y_{i,j}$ is sample j of this input direction. Typically in this work, any uniform inputs $\boldsymbol{y}$ are bounded $[-1, 1]$ in order to be consistent with what is known as the Wiener-Askey scheme. The formulation of Eq. (2.5) is still valid, however, as the inputs may be mapped to $[0, 1]$ before the upper and lower bounds of the initial condition are considered. More information on the Wiener-Askey scheme is presented later in this chapter within Section 2.5.

The elements of $\boldsymbol{q}_0\left(\boldsymbol{y}_j\right)$ are then propagated using a desired integration method as a black box to produce the set $\{\boldsymbol{q}\left(\boldsymbol{y}_j\right)\}$ at some time of interest. By leveraging the law of large numbers, MC methods are able to utilize the full dynamics of the desired system and estimate moments of the QoIs. In addition to moments, the joint PDFs of $\boldsymbol{q}(\boldsymbol{y})$ may be approximated using a sufficiently large number of realizations of $\boldsymbol{q}(\boldsymbol{y})$. However, it is well-known that large values of N_{MC} may be needed for an accurate estimation of these statistics, thus making the method computationally intractable for certain orbit problems. Thusly, although MC methods may provide an estimation of posterior PDFs without an assumption on the final distribution or a simplification of the dynamics, applications are limited by the high computation cost. There is therefore an interest in finding alternative, more efficient methods.

2.2.1.1 Solution Sensitivity

In addition to desiring the statistics or posterior PDF of $\boldsymbol{q}(\boldsymbol{y})$, a sensitivity analysis of the solution may be performed. The results of such an analysis determine the effect that a random input has on the the uncertainty of each QoI. Here, we first introduce the calculation of a sensitivity analysis as performed by a traditional MC method. The structure of the presented algorithm provides motivation for the construction of an approximation such as a surrogate method.

Sobol [104] discusses such an analysis and the difference between local and global sensitivities. Local sensitivities specify the derivative of the solution with respect to a stochastic input at a given realization of the inputs. Global sensitivities, however, consider the entire solution rather than a single solution realization [104, 99]. That is, local sensitivity measures, such as differential analysis, give values relative to single points in a multi-dimensional space, while global sensitivities quantify the effect of an input on an output while considering the distribution of the input in question, as well as the distributions of all other inputs [98].

We propose determining the global *sensitivity indices* using a Sobol approach [97]. Specifically, the presented method determines the variability of each component of the QoI, q_m, with respect to each direction i. In order to compute the indices, we first generate two input matrices

$$\boldsymbol{Y} = \begin{bmatrix} y_{1,1} & \cdots & y_{1,d} \\ \vdots & \ddots & \vdots \\ y_{N_{MC},1} & \cdots & y_{N_{MC},d} \end{bmatrix}, \quad \boldsymbol{Y}' = \begin{bmatrix} y'_{1,1} & \cdots & y'_{1,d} \\ \vdots & \ddots & \vdots \\ y'_{N_{MC},1} & \cdots & y'_{N_{MC},d} \end{bmatrix}, \tag{2.6}$$

where $\boldsymbol{Y}$ and $\boldsymbol{Y}'$ are composed of independent samples of N_{MC} random inputs that are appropriately distributed considering the density functions ρ_i dictated by the probability space of the problem.

Following [97], the (total) sensitivity indices, $\{S_{i,m}\}_{m=1}^{M}$, are given by

$$S_{i,m} = \frac{V\left(E\left(q_m|y_i\right)\right)}{V\left(q_m\right)} = \frac{U_{i,m} - \left(E(q_m)\right)^2}{V\left(q_m\right)}, \tag{2.7}$$

where

$$U_{i,m} = \int \left(E(\hat{q}_m|y_i = \tilde{y}_i)\right)^2 \rho_i(\tilde{y}_i)\mathrm{d}\tilde{y}_i \tag{2.8}$$

and $\rho_i(y_i)$ is the marginal density function of y_i. Eqs. (2.7) and (2.8) utilize the operator $E(\cdot)$, which takes the the expectation value of a given input. Additionally, $V\left(q_m\right)$ here denotes the variance of the quantity q_m, and $\tilde{y}_i$ is a fixed value of y_i. In practice, as detailed in [97], $U_{i,m}$ in Eq. (2.7) is estimated by MC sampling,

$$\hat{U}_{i,m} = \frac{1}{N_{MC} - 1} \sum_{j=1}^{N_{MC}} \hat{q}_m\left(y_{1,j}, \ldots, y_{d,j}\right) \hat{q}_m\left(y'_{1,j}, \ldots, y'_{(i-1),j}, y_{i,j}, y'_{(i+1),j}, \ldots, y'_{d,j}\right). \tag{2.9}$$

It is in Eq. (2.9) where the overall approach of the sensitivity index algorithm is readily apparent. If $\boldsymbol{Y}$ is thought of as the sample matrix and $\boldsymbol{Y}'$ as the *resample* matrix, then we can note how Eq. (2.7) together with Eq. (2.9) is evaluating the variation of the QoI q_m as the direction i also varies. That is, we quantify the relative effect that an input direction i has on the variance of the function $\boldsymbol{q}(\boldsymbol{y})$.

This result is significant, because the quantities $\{S_{i,m}\}$ provide valuable information. That is, although the effects of an input direction on the uncertainty of a QoI is not quantified absolutely, it is still compared to all other inputs. Therefore, operators or researchers are able to prioritize the analysis, inclusion or even observation of specific directions. For example, if the semi-major axis is found to dominate the variance of the QoIs, work can be done to mitigate these effects via opportunistic observations and/or particular observation types. Alternatively, models or approximations may be generated that priorize the input of semi-major axis over all other inputs for the sake of computational efficiency.

2.2.2 Linearization and the State Transition Matrix

Uncertainty mapping using a linearization assumption and the state transition matrix (STM) has enjoyed a long and successful history in the astrodynamics field [101]. Its applications are numerous and include conjunction assessment, maneuver design and predicting close encounters between the Earth and asteroids of interest (while the methodology remains accurate) [53, 4, 18, 86, 89, 52]. This methodology forgoes the sampling approach of MC and the reliance on random variable as seen in Eqs. (2.1) and (2.2). Rather, a linearization of the problem is made by leveraging the method of the Taylor series. We first begin with the definition of the Taylor series,

$$f(x) \approx f(a) + f'(a)(x-a) + \frac{1}{2!}f''(a)(x-a)^2 + \dots, \tag{2.10}$$

where an approximation to the function $f(\cdot)$ evaluated at x is desired. Here, the Taylor series expansion originates at the function evaluated at a. Since linearization and the STM is constructed for use in a filter, we often consider the unknown true solution $\boldsymbol{q}(t)$ alongside a known reference $\boldsymbol{q}^*(t)$. Thusly, the STM is derived from approximating the QoIs at a desired time $\boldsymbol{q}(t)$ by expanding a Taylor series about the QoIs evaluated from a reference solution $\boldsymbol{q}^*(t)$. Therefore, regard the previously defined function $f(\cdot)$ as the *solution flow* of the QoIs. Using this notion and applying Eq. (2.10), we desire to approximate the true solution by constructing a Taylor series about the reference solution, i.e.,

$$\boldsymbol{q}(t) = f(\boldsymbol{q}(t_0)) = f(\boldsymbol{q}^*(t_0)) + \left[\frac{\delta f(\boldsymbol{q}(t_0))}{\delta \boldsymbol{q}(t_0)}\right]^* \left[\boldsymbol{q}(t_0) - \boldsymbol{q}^*(t_0)\right] + O_q\left[\boldsymbol{q}(t_0) - \boldsymbol{q}^*(t_0)\right], \tag{2.11}$$

where $O_q[\cdot]$ are higher order terms in the series. The quantity $\left[\frac{\delta f(\boldsymbol{q}(t_0))}{\delta \boldsymbol{q}(t_0)}\right]^*$ is equivalent to the STM that we desire. However, this value is not typically known. To numerically approximate the STM, we consider the time derivative of Eq. (2.11), i.e.,

$$\dot{\boldsymbol{q}}(t) = \dot{\boldsymbol{q}}^*(t) + \left[\frac{\delta \dot{\boldsymbol{q}}(t)}{\delta \boldsymbol{q}(t)}\right]^* \left[\boldsymbol{q}(t) - \boldsymbol{q}^*(t)\right] + O_q\left[\boldsymbol{q}(t) - \boldsymbol{q}^*(t)\right], \tag{2.12}$$

where the notation $(\cdot)^*$ indicates an evaluation of the QoI taken at a dynamical reference.

By dropping higher order terms, Eq. (2.12) simplifies to

$$\Delta\dot{\boldsymbol{q}}(t) = A(t)\Delta\boldsymbol{q}(t), \tag{2.13}$$

where $A(t) \equiv \left[\frac{\delta\dot{\boldsymbol{q}}(t)}{\delta\boldsymbol{q}(t)}\right]^*$. The introduced quantities $\Delta\boldsymbol{q}$ and $\Delta\dot{\boldsymbol{q}}$ denote the QoI and time derivative QoI deviations, respectively . We now have a linear differential equation which elegantly integrates to

$$\Delta\boldsymbol{q}(t) = \boldsymbol{\Phi}(t, t_0)\Delta\dot{\boldsymbol{q}}(t_0), \tag{2.14}$$

where $\boldsymbol{\Phi}(t, t_0)$ is the STM. The STM linearly propagates the initial deviations $\Delta\boldsymbol{q}(t_0)$ to a desired time t. Levergating the STM's ability to propagate deviations, we now focus on the ability of the STM to linearly propagate the initial covariance matrix $\boldsymbol{\Sigma}_0$ of the system, i.e.,

$$\boldsymbol{\Sigma}_t = \boldsymbol{\Phi}(t, t_0)\boldsymbol{\Sigma}_0\boldsymbol{\Phi}(t, t_0)^{\mathrm{T}}, \tag{2.15}$$

with $\boldsymbol{\Sigma}_t$ being the covariance at a designated time t, which represents the uncertainty of the chosen object. Although the linearization of the problem results in the remarkably computationally efficient STM, this uncertainty is linearly propagated (where accuracy is lost due to truncating the higher order moments) and the posterior is assumed to be Gaussian. Since this has been shown to be inaccurate under certain conditions, there has been motivation for further research [70].

2.3 Recent Advances

Due to the deisire to increase efficiency yet reduce assumptions and simplifications, research on UQ has continued within the astrodynamics community. Methodologies such as the unscented transform (UT) and Gaussian mixture methods (GMMs) leverage a wealth of information, tools and applications that the long history of linearization and the STM has laid down as a foundation. Other methods, such as differential algebra (DA), state transition tensors (STTs) and surrogate methods, vary in approach and, in some cases, attempt to approximate a MC sampling process. The UT has recently gained traction as a relatively simple yet significant upgrade over the STM, due to its property of nonlinear uncertainty propagation, while certain surrogate methods are currently being used in an operational setting [100].

2.3.1 Unscented Transform

The unscented transform (UT) is utilized to provide a non-linear means of uncertainty propagation for a filter known as the unscented Kalman filter (UKF). The UKF relies on the propagation of deterministic *sigma points* to determine the mean and standard deviation of the uncertainty ellipsoid, thereby not requiring any derivative work such as with the STM [69]. Typically, the number of required sigma points is much smaller when compared to that of a MC method, so the UKF is able to remain efficient even when compared to the linearization and STM methodology. These points have been shown to be equivalent to those using a Gauss-Hermite quadrature rule [63]. In practice, the UKF is used with the UT, which replicates a third order Gauss-Hermite quadrature rule [60]. Therefore, the number of required sigma points is $2 * M + 1$, where M is the size of the state space. Although the UKF utilizes a procedure which non-linearly propagates uncertainty, there is still an assumption of a Gaussian posterior. Therefore, other methods are worthy of consideration.

2.3.2 Gaussian Mixture Methods

Gaussian mixture methods (GMMs), or Gaussian sum methods, approximate PDFs (without distribution assumptions) as a weighted sum of Gaussian PDFs [5]. That is,

$$F_{\mathcal{N}}(\boldsymbol{y}) = \sum_{\lambda=1}^{\Lambda_{GMM}} w_\lambda \mathcal{N}(\boldsymbol{y}; \boldsymbol{\mu}_\lambda, \boldsymbol{\Sigma}_\lambda), \tag{2.16}$$

where $F_{\mathcal{N}}(\boldsymbol{y})$ is the approximated PDF, Λ_{GMM} is the total number of Gaussian elements indexed by λ and w_λ are the weights of the Gaussian PDFs with means $\boldsymbol{\mu}_\lambda$ and variance $\boldsymbol{\Sigma}_\lambda$. The weights are constrained by the properties $1 = \sum_{\lambda=1}^{\Lambda_{GMM}} w_\lambda$ and $w_\lambda \geq 0$. This formulation provides advantages such as facilitating parallelization over λ and compatibility with Gaussian filters and associated methods. Unfortunately, the GMMs suffer from the *curse of dimensionality*, where the splitting along input directions becomes a computational cost that increases exponentially. This can lead to intractability and therefore motivates research in other areas.

2.3.3 State Transition Tensors

By leveraging analytic expressions as an approximation for the solution flow of the Fokker-Planck equation, Park and Scheeres [85] utilize state transition tensors (STT) to propagate uncertainty in a series of test cases. The results compare well with that of fully non-linear propagations. STTs rely on analytic high order Taylor series expansions to replace a solution flow function. The method has been shown to be accurate for propagations lasting many (20) orbital periods in two body dynamics and with a J_2 gravity field [46]. Beyond pure uncertainty propagation, STTs have also been applied to estimation filters [80]. As an analytic method, STTs are computationally efficient, and, once derived, do not need to be altered as long as the dynamics are not changed. However, the derivation of high order, analytic Taylor expansions is complex. For example, when considering dynamics such as gravity fields beyond the J_2 order, it is possible that deriving the STT expression would be an overly involved process. Additionally, many operators have legacy code which is treated as a blackbox. Because of this, it is desirable to explore methods that are less intrusive.

2.3.4 Differential Algebra

Differential algebra (DA) is relatively new to the field of astrodynamics, and it relies on a high order expansion to replace a solution flow [114, 81]. That is, DA imitates the analytic Taylor series approximation of some function via an algebraic approach, which utilizes Taylor polynomials. Due to being calculated in a computer environment, DA relieves some of the burden of an analytic Taylor series. The reliance on these high order series affords the method a great deal of computational efficiency and does not rely on large numbers of numerical integrations for the construction of a DA model. Rather, DA propagates the Taylor polynomial over a desired time span in order to intrusively create an efficient analytic model for further propagation purposes. Typically, DA is capable of functioning in place of a traditional blackbox in a MC environment, and therefore provides a sampling based method for determining solution statistics. DA, however, is entirely reliant on the Taylor series-like method and its limitations. Therefore, certain conditions, such as discontinuity, prevent the approximation from converging to an accurate solution.

2.3.5 Kriging

As a non-intrusive method, kriging has received some attention in the astrodynamics literature. Originally conceived in the field of geostatistics, kriging approximates an unknown function using a known trend function in addition to a Gaussian random function [47]. Using correlation values between sample points, estimate values of a function are calculated for a given input. Gano et al [47] uses kriging along with other surrogate techniques in various optimization problems, with results showing that kriging performs relatively well when data is sparse. In addition to this, kriging is compared to other non-intrusive methods of UQ such as PCEs [112]. In this case, generalized kriging performed poorly when considering the accuracy of the various methods.

2.4 Sampling Techniques

Throughout various research topics, integrations are often performed. These calculations may be done analytically, numerically or via approximations. For example, the first moment (mean) of a function is denoted by

$$\mu = \int f(x)\rho(x)dx \tag{2.17}$$

where the function $f(x)$ is reliant on the input x which has the distribution function $\rho(x)$. Often, this moment equation is approximated by the discrete form of the first moment, i.e.,

$$\mu \approx \frac{1}{N}\sum_{n=1}^{N} f(x_n). \tag{2.18}$$

Here, N is considered to be a number of samples sufficiently large for Eq. (2.18) to converge to the result of Eq. (2.17). This is a specific example illustrating the logic behind MC simulations where the law of large numbers is leveraged.

It is possible to approximate Eq. (2.17) and other integrations without solely relying on very large numbers of samples N as seen in the discrete Eq. (2.18). Techniques that utilize "smart" sampling are often able to reduce N to a tractable number. An example of this is collocation method of Gaussian quadrature (GC). GC approximates an integral as a weighted sum, i.e.,

$$\int_{-1}^{1} f(y)\,dy = \sum_{n=1}^{N} w_n f(y_n) + R_N(f), \tag{2.19}$$

where N is the total number of quadrature nodes and w_n are the nodal weights determined via the particular quadrature algorithm that is chosen. The error $R_N(f)$ of the approximation is zero if the function f is a polynomials of degree $2N-1$.

Extending the univariate form of Eq. (2.19) to a mulivariate formulation is useful in the case of astrodynamics, and this is possible via a tensor product of sets of univariate nodes. That is,

$$\int_{-1}^{1} f(\boldsymbol{y})\,d\boldsymbol{y} = \sum_{n_1=1}^{N_1} \ldots \sum_{n_d=1}^{N_d} w_{n_1} f(y_{1,n_1},\ldots,y_{d,n_1}) \ldots w_{n_d} f(y_{1,n_d},\ldots,y_{d,n_d}) + R_N(f). \tag{2.20}$$

Due to the tensor products needed in Eq. (2.20), the computation cost of the multivariate GC increases dramatically with respect to the input dimension d. Therefore, this method suffers from what is known as the *curse of dimensionality*. In an effort to reduce the number of samples needed for higher dimension problems, the Smolyak grid method was developed [103].

Based on the principles of a sparse grid, Smolyak's method reduces the total number of nodes needed for an accurate discretization by ignoring high-order interactions among dimensions. Although a Smolyak grid based quadrature still maintains the curse of dimensionality, the cost at higher dimensions is greatly reduced when compared to the full tensor product method of Eq. (2.20). Jones et al [67] highlights the cost comparison by providing a table of required samples. As an example, when assuming that all directions of a $d = 7$ case require the same number of nodes, if $N = 6$, the full tensor quadrature requires 279,936 samples while a Smolyak approach reduces that number to 8,583.

2.5 Surrogate Methods

The surrogate methods discussed in this work seek to approximate $\boldsymbol{q}(\boldsymbol{y})$ as a function of $\boldsymbol{y}$, i.e., the mapping $\boldsymbol{y} \rightarrow \boldsymbol{q}(\boldsymbol{y})$, which we will then use to estimate the statistics of $\boldsymbol{q}(\boldsymbol{y})$, such as the mean, standard deviation (STD), and possibly marginal and joint PDFs, as well as the sensitivity of the components of $\boldsymbol{q}(\boldsymbol{y})$ to each random input y_i. Explicitly, the approximation

$$\hat{\boldsymbol{q}}(\boldsymbol{y}) = \boldsymbol{q}(\boldsymbol{y}) + \epsilon \tag{2.21}$$

is sought such that the error ϵ is below an acceptable level, where $\hat{\boldsymbol{q}}(\boldsymbol{y})$ is given here as a generic surrogate approximation to the multivariate, vector-valued function $\boldsymbol{q}(\boldsymbol{y})$. To find the left hand side of Eq. (2.21), we train the surrogate on the black box propagationed samples. This collection of samples and realizations is organized into the data set

$$\mathcal{D} = \left\{\left(\boldsymbol{y}_j, \boldsymbol{q}\left(\boldsymbol{y}_j\right)\right)\right\}_{j=1}^{N}, \tag{2.22}$$

which is referred to as the *training data*. Using a regression approach, the training data $\mathcal{D}$ is used to construct the approximation of $\boldsymbol{q}(\boldsymbol{y})$, where the distance between $\boldsymbol{q}(\boldsymbol{y})$ and the approximation $\hat{\boldsymbol{q}}(\boldsymbol{y})$ is minimized at the samples $\{\boldsymbol{y}_j\}$.

The advantage of this approach is that the construction of $\hat{\boldsymbol{q}}(\boldsymbol{y})$ by regression may require far fewer realizations of $\boldsymbol{q}(\boldsymbol{y})$, in the form of $\mathcal{D}$, than a MC approach would need in order to accurately produce statistics or estimations of joint PDFs, that is, $N \ll N_{MC}$. By reducing the number of samples propagated in an ODE solver, the computation cost of calculating desired results is, in turn, reduced. In the case of the two discussed methods of PCEs and SR, the surrogate is often composed of a polynomial basis. This formulations offers many advantages. One of which is a beneficial evaluation cost, where computing a sample $\hat{\boldsymbol{q}}(\boldsymbol{y}_j)$ can be assumed to be significantly more efficient than that of the complex blackbox $\boldsymbol{q}(\boldsymbol{y}_j)$. Therefore, the surrogate $\hat{\boldsymbol{q}}(\boldsymbol{y})$ can be used in lieu of a traditional MC but with a more tractable cost. An additional feature of this polynomial formulation is the ability to compute analytical moments via the surrogate's coefficients. This approach provides an advantage over even the sampling of the surrogate itself, as the analytical expressions do not rely on the law of large numbers for the moment approximation. As a final example, the derivatives of the polynomial basis are generally known. Therefore, values such as local sensitivities with respect to inputs can be computed with relative efficiency.

For conciseness, we define the data-dependent (semi-) inner product of two vectors $\boldsymbol{q}(\boldsymbol{y}), \boldsymbol{q}'(\boldsymbol{y}) \in \mathbb{R}^M$ as

$$\left\langle \boldsymbol{q}, \boldsymbol{q}' \right\rangle_D = \frac{1}{N} \sum_{j=1}^{N} \left\langle \boldsymbol{q}(\boldsymbol{y}_j), \boldsymbol{q}'(\boldsymbol{y}_j) \right\rangle_2 , \tag{2.23}$$

where $\langle \cdot, \cdot \rangle_2$ denotes the standard Euclidean inner product. The (semi-) inner product in Eq. Eq. (2.23) induces the (semi-) norm

$$\|\boldsymbol{q}\|_D = \langle \boldsymbol{q}, \boldsymbol{q} \rangle_D^{1/2} , \tag{2.24}$$

which is used hereafter and is referred to as the *data norm*. It is noted that if the inputs of the data norm are scalar quantities rather than vectors, the result of Eq. (2.24) is the root mean square (RMS) of a data set of size N.

2.5.1 Polynomial Chaos

As a surrogate method, polynomial chaos is a relatively recent UQ method in astrodynamics. However, it has proven to to be efficent in research as well as operations [67, 100]. Leveraging properties of orthogonal polynomials and their relations to random distributions via the Wiener-Askey scheme, PCEs project a stochastic solution onto a polynomial basis. This approximation can be evaluated in various methods as discussed earlier in this section, and the reader is encouraged to explore earlier work on the topic if a large amount of detail and theory is desired [67, 66, 68]. A brief overview of the method is provided, however, as key portions of the theory impact the motivation behind this research.

2.5.1.1 Polynomial Chaos Formulation

For PCEs, the surrogate approximation for a scalar function is given as

$$\hat{q}_{PC}(\boldsymbol{y}) = \sum_{\boldsymbol{\lambda} \in \Lambda_{P,d}} c_{\boldsymbol{\lambda}} \psi_{\boldsymbol{\lambda}}(\boldsymbol{y}), \tag{2.25}$$

where Λ is referred to as the *multi-index*, P is the maximum order of polynomial desired, $\{\psi_{\boldsymbol{\lambda}}\}$ are a set of multidimensional polynomials that are orthonormal with respect to the weighting function of the random inputs $\boldsymbol{y}$, and $c_{\boldsymbol{\lambda}}$ are the polynomial chaos coefficients computed by the projection of $q(\boldsymbol{y})$ onto each basis function $\psi_{\boldsymbol{\lambda}}$.

Remark 1. *It is noted that Eq. (2.25) is an estimate of a multivariate function $q(\boldsymbol{y})$ with a scalar output. It is possible to formulate a PCE to produce the output of a vector function, and interested readers are encouraged to investigate the cited works of Jones, et al., for this vector valued QoI estimation. Here, we present the scalar formulation for the sake of simplicity.*

The orthonormality of the polynomial basis can be expressed as

$$\int \psi_{\boldsymbol{\lambda}}(\boldsymbol{y})\psi_{\boldsymbol{\lambda}'}(\boldsymbol{y})\rho(\boldsymbol{y})d\boldsymbol{y} = \delta_{\boldsymbol{\lambda},\boldsymbol{\lambda}'}, \tag{2.26}$$

where $\rho(\boldsymbol{y})$ is the joint probability measure of the random inputs $\boldsymbol{y}$. In Eq. (2.26), $\delta_{\boldsymbol{\lambda},\boldsymbol{\lambda}'}$ operates as a Dirac delta function over a set, i.e.,

$$\delta_{\boldsymbol{\lambda},\boldsymbol{\lambda}'} = \begin{cases} 1 & \boldsymbol{\lambda} = \boldsymbol{\lambda}' \\ \\ 0 & \text{otherwise.} \end{cases} \tag{2.27}$$

Each basis function $\psi_{\boldsymbol{\lambda}}$ is a tensor product of univariate polynomials of degree $\lambda_i \in \mathbb{N}_0^1$, i.e.,

$$\psi_{\boldsymbol{\lambda}} = \psi_{\lambda_1}(y_1), \ldots, \psi_{\lambda_d}(y_d), \quad \boldsymbol{\lambda} \in \mathbb{N}_0^d. \tag{2.28}$$

As previously discussed, these univariate functions are known *a priori* due to the known relationship between the polynomials and the chosen distributions of the random inputs $\boldsymbol{y}$.

The set of multi-indices is

$$\Lambda_{P,d} := \{\boldsymbol{\lambda} \in \mathbb{N}_0^d \colon \|\boldsymbol{\lambda}\|_1 \leq P, \|\boldsymbol{\lambda}\|_0 \leq d\}, \tag{2.29}$$

where $\|\boldsymbol{\lambda}\|_1 = \sum_{i=1}^{d} \lambda_i$ and $\|\boldsymbol{\lambda}\|_0 = \{i \colon \lambda_i > 0\}$. The cardinality of $\Lambda_{P,d}$, that is, the number of terms, is known via the equation

$$\Lambda_{PC} = \frac{(P+d)!}{P!d!}. \tag{2.30}$$

Eq. (2.30) highlights the exponential cost increase of PCEs with respect to dimension d. This relationship is a product of the *curse of dimensionality*, and, as a result, increasing d can quickly lead to the number of terms in the expansion growing rapidly.

2.5.1.2 Solution Via Least Squares

Since the polynomial basis is known, it is therefore necessary to solve for the unknown coefficients $c_{\boldsymbol{\lambda}}$ for all values of $\boldsymbol{\lambda} \in \Lambda_{P,d}$. Although Jones et al. [67] explores multiple methods for calculating the coefficients, including quadrature methods discussed in Sampling Techniques, this presentation focuses on the least squares approach. Using least squares regression, we are able to solve for $c_{\boldsymbol{\lambda}}$ by considering the cost function

$$J(c_{\boldsymbol{\lambda}}) = \sum_{j=1}^{N_{PC}} (\hat{q}(\boldsymbol{y}_j; c_{\boldsymbol{\lambda}}) - q(\boldsymbol{y}_j))^2. \tag{2.31}$$

Applying Eq. (2.25) to Eq. (2.31), we rewrite the least squares cost function as

$$J(c_{\boldsymbol{\lambda}}) = \|\boldsymbol{HC} - \boldsymbol{Y}\|_F^2, \tag{2.32}$$

where $\|\cdot\|_F$ indicates the Frobenius norm. The matrices found within this norm are defined as

$$\boldsymbol{H} := \begin{bmatrix} \psi_{\lambda_1}(\boldsymbol{y}_1) & \cdots & \psi_{\lambda_{\Lambda_{PC}}}(\boldsymbol{y}_1) \\ \vdots & \ddots & \vdots \\ \psi_{\lambda_1}(\boldsymbol{y}_{N_{PC}}) & \cdots & \psi_{\lambda_{\Lambda_{PC}}}(\boldsymbol{y}_{N_{PC}}) \end{bmatrix}, \quad \boldsymbol{C} := \begin{bmatrix} c_{\boldsymbol{\lambda}_1} \\ \vdots \\ c_{\boldsymbol{\lambda}_{\Lambda_{PC}}} \end{bmatrix} \tag{2.33}$$

$$\boldsymbol{Y} := \begin{bmatrix} q(\boldsymbol{y}_1) \\ \vdots \\ q(\boldsymbol{y}_{N_{PC}}) \end{bmatrix}, \tag{2.34}$$

where $\boldsymbol{H} \in \mathbb{R}^{N_{PC} \times \Lambda_{PC}}$ is referred to as the measurement matric, $\boldsymbol{C} \in \mathbb{R}^{\Lambda_{PC}}$ is the collection of PCE coefficients, and $\boldsymbol{Y} \in \mathbb{R}^{N_{PC}}$ is the training data organized into a vector for computation purposes. Minimizing the cost function Eq. (2.32) is possible by evaluating

$$\hat{\boldsymbol{C}} = (\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H})^{-1}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Y}, \tag{2.35}$$

when $N_{PC} \geq \Lambda_{PC}$. Here, the matrix $\hat{\boldsymbol{C}}$ contains the previously unknown coefficients, thus completing the surrogate approximation seen in Eq. (2.25).

2.5.2 Separated Representations

The separable approach is based on approximating a multivariate scalar function $q\left(\boldsymbol{y}\right)$ with a sum of products of univariate functions. Consider the example function in Section 2.1, which is presented as the state of an orbiting satellite. In the scalar setting, let the QoI $q(\boldsymbol{y})$ be a single element of the satellite state such as the x-position. In the framework of SR, the separated approximation of the satellite state is then a sum of separable products

$$q(\boldsymbol{y}) \approx \hat{q}(\boldsymbol{y}) = \sum_{l=1}^{r} s^l \prod_{i=1}^{d} u_i^l\left(y_i\right), \tag{2.36}$$

where $\hat{q}(\boldsymbol{y})$ is the estimation of $q(\boldsymbol{y})$, and $\{u_i(y_i)\}_{i=1}^{d}$, are unknown univariate functions to be determined so that $\hat{q}(\boldsymbol{y})$ is as close as possible to $q(\boldsymbol{y})$. Additionally, $\{u_i^l(y_i)\}_{l=1}^{r}$, $i = 1, \dots, d$ are referred to as *factors*, and $\{s^l\}_{l=1}^{r}$ are normalization constants such that each $u_i^l(y_i)$ has a unit norm, as elaborated in Section 2.5.2.3. These normalization constants provide numerical stability to the formulation of an SR. In Eq. (2.36), the *a priori* unknown constant r is referred to as the *separation rank* and is ideally the smallest number of separated terms – to be determined in the construction of SR – in order to achieve a desired accuracy in approximating $q(\boldsymbol{y})$. The approximation $\hat{q}(\boldsymbol{y})$ is considered to be *low-rank* if r remains small for the target accuracy. It is worthwhile highlighting the fact that the separated representation Eq. (2.36) is a nonlinear approximation of $q(\boldsymbol{y})$ with a small number of *parameters* in which the expansion basis functions $\prod_{i=1}^{d} u_i^l\left(y_i\right)$ are not predefined; they are sought such that the approximation error is minimized, as discussed in the following.

When $q(\boldsymbol{y})$ admits a low separation rank r, this allows a fast decay of the error with respect to r. In addition, as we shall explain in Section 2.5.2.3, the nonlinear approximation Eq. (2.36) may be computed using multilinear approaches due to its separable form with respect to variables y_i. The combination of these two attributes of SR will allow its construction with a number of samples that is linear in d, which is smaller than that of standard approximation techniques relying on *a priori* fixed bases. For scenarios where the desired accuracy may not be achieved by a small r, e.g., when $q(\boldsymbol{y})$ is discontinuous in $\boldsymbol{y}$ along an arbitrary hyperplane, SR may not lead to an efficient approximation. Therefore, in practice, it is crucial to assess the quality of a constructed

SR – see Remark 2 and Chapter 5 – prior to using it to learn an arbitrary $q(\boldsymbol{y})$ and its statistics. Such an assessment is also a key step in the construction of other surrogate models. We refer the interested reader to the review manuscripts [74, 25] and the references therein for examples of successful application of SR to various types of problems in engineering and sciences.

The construction of the SR in Eq. (2.36) may be posed in the form of a nonlinear optimization involving unknowns $\{u_i^l(y_i)\}$ in order to minimize the distance between $q(\boldsymbol{y})$ and $\hat{q}(\boldsymbol{y})$ at the samples $\{\boldsymbol{y}_j\}$,

$$\min_{\{u_i^l(y_i)\}} \|q - \hat{q}\|_D^2 \,, \tag{2.37}$$

as detailed in Section 2.5.2.3. These unknowns are often approximated in an *a priori* selected basis, thus allowing for a numerical solution to the optimization problem. Here, we consider the approximation of each factor in a basis of Hermite polynomials in y_i,

$$u_i^l\,(y_i) \quad \approx \quad \sum_{p=0}^{P} c_{i,p}^l \psi_p(y_i), \tag{2.38}$$

where $\psi_p(y_i)$ is the Hermite polynomial of degree p with maximum order P. In general, the basis functions $\psi_p(y_i)$ are selected such they are orthogonal with respect to the PDF of y_i, as given by the Askey family of orthogonal polynomials [14]. This allows for selecting a different basis for each direction i, depending on the distribution of y_i. Given the discretization of factors in Eq. (2.38), the optimization problem Eq. (2.37) is now reduced to finding the unknown coefficients $\boldsymbol{c}_i^l = [c_{i,0}^l, \dots, c_{i,P}^l]$ for each direction i and rank l via problem

$$\min_{\{\boldsymbol{c}_i^l\}} \|q - \hat{q}\|_D^2 \,. \tag{2.39}$$

The number of unknowns in the optimization problem Eq. (2.39) is $r \cdot d \cdot (P+1)$, which is linear in the dimension d, as far as the separation rank r is independent of d. As discussed in [20], this linear scaling of the number of unknowns in turn suggests a linear (in d) growth requirement on the number of samples N used to solve Eq. (2.39).

Remark 2. *While increasing the separation rank* r *improves the accuracy of SR, it is not possible to determine a priori if an arbitrary* $q(\boldsymbol{y})$ *lends itself to an accurate separated representation with low separation rank. In practice, the accuracy of a constructed SR,* $\hat{q}(\boldsymbol{y})$*, may be assessed empirically via, for instance, cross-validation techniques commonly used for other types of surrogate models, e.g., based on multivariate polynomial expansions (See [43]). For certain classes of problems, e.g., linear elliptic, [26], semi-linear elliptic, [57], and parabolic, [59], partial differential equations with random inputs, a priori error estimates have been derived, demonstrating the sparsity of the solution when expanded in multivariate polynomial bases. That is, only a small fraction of the basis functions have non-negligible coefficients. When each SR factor* $u_i^l(y_i)$ *is approximated in the same (univariate) polynomial basis, as in Eq. (2.38), such sparse solutions are in principle guaranteed to admit SRs with low separation ranks. This is because the two approaches employ basis functions spanning the same space within which the solution exists or is well approximated.*

2.5.2.1 SR of Vector-Valued Functions

In many cases, it is desirable to approximate a vector-valued QoI via SR. For instance, in the current application to orbit uncertainty propagation, an estimate for the Cartesian components of position and velocity is needed. In such cases, the SR approximation of $\boldsymbol{q}(\boldsymbol{y}) \in \mathbb{R}^M$ is given by

$$\boldsymbol{q}(\boldsymbol{y}) \approx \hat{\boldsymbol{q}}(\boldsymbol{y}) = \sum_{l=1}^{r} s^l \, \boldsymbol{u}_0^l \prod_{i=1}^{d} u_i^l \left(y_i\right), \tag{2.40}$$

where the definitions and approximation of $u_i^l(y_i)$ remains the same as in the scalar-valued SR. A significant difference between Eq. (2.40) and Eq. (2.36) is the addition of the vector of deterministic factors $\boldsymbol{u}_0^l = [u_{0,1}^l, \dots, u_{0,M}^l]^{\mathrm{T}} \in \mathbb{R}^M$, which is used to extend the approximation from the scalar-valued $q(\boldsymbol{y})$ to the vector-valued $\boldsymbol{q}(\boldsymbol{y})$ by solving the optimization problem

$$\min_{\{\boldsymbol{c}_i^l\},\{\boldsymbol{u}_0^l\}} \|\boldsymbol{q} - \hat{\boldsymbol{q}}\|_D^2 \, . \tag{2.41}$$

In the case of Eq. (2.40), the set of scalars s^l are now normalization constants such that both $u_i^l(y_i)$ and $\boldsymbol{u}_0^l$ have unit norm. The method of determining these values is detailed in Section 2.5.2.3.

2.5.2.2 Process of Constructing SR

As an introduction to the implementation of SR, a high level algorithm for the vector output is provided in this section. The sampling approach to constructing SR requires the generation of training data $\mathcal{D}$, i.e., the set $\{(\boldsymbol{y}_j, \boldsymbol{q}(\boldsymbol{y}_j))\}$ of N samples from Eq. (2.22), that represent propagated realizations of $\boldsymbol{q}(\boldsymbol{y})$ at (random) samples of inputs $\boldsymbol{y}$. Such algorithms are dubbed **non-intrusive** as the evaluation of $\boldsymbol{q}(\boldsymbol{y})$ does not require any alteration of the solvers for Eq. (2.1). In the context of orbit state propagation where $\boldsymbol{y}_j$ is Gaussian, the non-intrusive SR process may be summarized by:

(1) Generate a set of independent, random realizations $\{\boldsymbol{y}_j\}_{j=1}^N$, where each $\boldsymbol{y}_j \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_{d\times d})$.

(2) Using the **a priori** state distribution, generate the set of samples $\{\boldsymbol{q}_0(\boldsymbol{y}_j)\}_{j=1}^N$ at the epoch time.

(3) Propagate each of the N samples to the time of interest using a given black box ODE solver to get $\{\boldsymbol{q}(\boldsymbol{y}_j)\}_{j=1}^N$.

(4) Use the training data $\{(\boldsymbol{y}_j, \boldsymbol{q}(\boldsymbol{y}_j))\}_{j=1}^N$ to generate the SR approximation.

Steps 1 through 3 of Section 2.5.2.2 include the generation of $\mathcal{D}$. Some of the data used in the presented numerical results assume an *a priori* Gaussian PDF, i.e., $\boldsymbol{q}_0 \sim \mathcal{N}(\bar{\boldsymbol{q}}_0, \boldsymbol{\Sigma})$, with a given mean vector $\bar{\boldsymbol{q}}_0$ and covariance matrix $\boldsymbol{\Sigma}$. Each sample $\boldsymbol{y}_j$ is first mapped to an initial sample $\boldsymbol{q}_0(\boldsymbol{y}_j)$ via a Cholesky decomposition of $\boldsymbol{\Sigma}$. Discussed later on in Chapters 3 and 6, some inputs used in surrogate creation are uniform, i.e., $y \sim \mathcal{U}(-1, 1)$. The reader is also reminded that the SR methodology presented in this chapter is formulated such that any input-polynomial relationship that follows the Wiener-Askey rule is applicable. The elements of $\boldsymbol{q}_0(\boldsymbol{y}_j)$ are then propagated using a desired integration method as a black box to produce the set $\{\boldsymbol{q}(\boldsymbol{y}_j)\}$ at some time of interest.

Although the number of samples N required in $\mathcal{D}$ for an approximation is not known *a priori*, the growth of N with respect to stochastic dimension d is linear. This relationship is discussed in further detail within Section 2.5.2.4 and can be seen in Eq. (2.57).

2.5.2.3 Solution via Alternating Least Squares

Step 4 in Section 2.5.2.2 is the generation of the SR approximation, which here is done via alternating least squares (ALS) regression using a set of N training samples $\{(\boldsymbol{y}_j, \boldsymbol{q}(\boldsymbol{y}_j))\}$. This method reduces the larger nonlinear optimization process into a series of linear least squares regression problems [20].

Overall approach. Given an initial r, e.g., $r = 1$, pre-selected basis $\{\psi_p(y_i)\}$, and initial coefficient values $\{\boldsymbol{c}_i^l\}$ and $\{\boldsymbol{u}_0^l\}$, the ALS algorithm updates the coefficients $\{\boldsymbol{c}_i^l\}$ by alternating through a sequence of one-dimensional and linear optimization problems. For the results contained within this body of work, initial values of $c_{i,0}^l = 1$, $c_{i,p}^l = 0 \quad p = 1, \ldots, P$ and $u_{0,m}^l = 0$ are used. Let $i' = 1, \ldots, d$ denote one direction of interest for each of these problems. The coefficients $\{\boldsymbol{c}_{i'}^l\}$ are updated by solving the linear least squares regression

$$\{\boldsymbol{c}_{i'}^l\}_{l=1}^r = \underset{\{\boldsymbol{c}_{i'}^l\}_{l=1}^r}{\arg\min} \; \|\boldsymbol{q} - \hat{\boldsymbol{q}}\|_D^2 \,, \tag{2.42}$$

while the coefficients $\{\boldsymbol{c}_i^l\}$ for other directions $i \neq i'$ and $\{\boldsymbol{u}_0^l\}$ are fixed at their current values. The ALS algorithm continues with a sweep through each direction i' in alternation. After reaching the final direction d, the values of $\{\boldsymbol{u}_0^l\}$ are solved for (completing what will be referenced as a full ALS loop) using the linear least squares problem

$$\{\boldsymbol{u}_0^l\}_{l=1}^r = \underset{\{\boldsymbol{u}_0^l\}_{l=1}^r}{\arg\min} \; \|\boldsymbol{q} - \hat{\boldsymbol{q}}\|_D^2 \,, \tag{2.43}$$

while fixing the coefficients $\{\boldsymbol{c}_i^l\}$ for all i at their current values. Full loops of ALS are continued until an *a priori* selected convergence criteria is met. Upon this convergence of the solution with rank r, the SR solution is tested for solution precision when compared to the training data. If this second convergence criteria is not met, then the separation rank is increased ($r = r + 1$) and the ALS procedure is repeated to generate a solution for the larger set of coefficients $\{\boldsymbol{c}_i^l\}$ and $\{\boldsymbol{u}_0^l\}$.

The remainder of this section provides a more detailed description of the ALS method used to generate an SR approximation.

Updating stochastic coefficients $\{\boldsymbol{c}_{i'}^l\}$. Setting the derivative of the cost function $\|\boldsymbol{q} - \hat{\boldsymbol{q}}\|_D^2$ with respect to $\{\boldsymbol{c}_{i'}^l\}$ to zero leads to the normal equation

$$\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)\boldsymbol{z} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{h} \tag{2.44}$$

for the solution $\{\boldsymbol{c}_{i'}^l\}$ to problem Eq. (2.42), organized in

$$\boldsymbol{z} = \begin{bmatrix} \left(\boldsymbol{c}_{i'}^1\right)^{\mathrm{T}} & \ldots & \left(\boldsymbol{c}_{i'}^r\right)^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}. \tag{2.45}$$

In Eq. (2.44),

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_{11} & \cdots & \boldsymbol{A}_{1r} \\ \vdots & \ddots & \vdots \\ \boldsymbol{A}_{N1} & \cdots & \boldsymbol{A}_{Nr} \end{bmatrix}, \tag{2.46}$$

where the (j,l) block of $\boldsymbol{A}$, $\boldsymbol{A}_{jl} \in \mathbb{R}^{M\times P}$, is given by

$$\boldsymbol{A}_{jl} = s^l \begin{bmatrix} \boldsymbol{u}_0^l\, \psi_0(y_{i',j}) & \cdots & \boldsymbol{u}_0^l\, \psi_P(y_{i',j}) \end{bmatrix} \prod_{i\neq i'} u_i^l\left(y_{i,j}\right) \tag{2.47}$$

and the data vector

$$\boldsymbol{h} = \begin{bmatrix} \boldsymbol{q}(\boldsymbol{y}_1)^{\mathrm{T}} & \cdots & \boldsymbol{q}(\boldsymbol{y}_N)^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^{MN}, \tag{2.48}$$

contains the samples of $\boldsymbol{q}(\boldsymbol{y})$. Before continuing to the next direction in the alternation, the values of s^l, and therefore $\boldsymbol{c}^l$, are updated via

$$\boldsymbol{c}_{i'}^l \leftarrow \frac{\boldsymbol{c}_{i'}^l}{\|u_{i'}^l\|_D} \quad \text{and} \quad s^l \leftarrow s^l \|u_{i'}^l\|_D. \tag{2.49}$$

Remark 3. *Notice that, while the total number of unknowns in Eq. (2.42) is* $r\cdot d\cdot P + 1$, *only* $r\cdot P + 1$ *of them appear in the one-dimensional optimization problems Eq. (2.44), which is independent of* d, *as long as* r *does not depend on* d.

Updating deterministic factors $\{\boldsymbol{u}_0^l\}$. After cycling through all directions $i' = 1, \ldots, d$ in the ALS process to compute $\boldsymbol{c}_{i'}^l$, the best estimate of $\{\boldsymbol{u}_0^l\}$ is found by setting the derivative of $\|\boldsymbol{q} - \hat{\boldsymbol{q}}\|_D^2$ with respect to $\boldsymbol{u}_0^l$ to zero. The resulting normal equation associated with problem Eq. (2.43) is

$$\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)\boldsymbol{Z} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{H}, \tag{2.50}$$

where,

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_1 & \cdots & \boldsymbol{A}_N \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^{N \times r} \tag{2.51}$$

and each block $\boldsymbol{A}_j \in \mathbb{R}^{1 \times r}$ is of the form

$$\boldsymbol{A}_j = \begin{bmatrix} s^1 \prod_{i=1}^{d} u_i^1 \left(y_{i,j}\right) & \cdots & s^r \prod_{i=1}^{d} u_i^r \left(y_{i,j}\right) \end{bmatrix}. \tag{2.52}$$

Additionally, the solution and data matrices $\boldsymbol{H}$ and $\boldsymbol{Z}$ are, respectively,

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{q}(\boldsymbol{y}_1) & \cdots & \boldsymbol{q}(\boldsymbol{y}_N) \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^{N \times M} \tag{2.53}$$

and

$$\boldsymbol{Z} = \begin{bmatrix} \boldsymbol{u}_0^1 & \cdots & \boldsymbol{u}_0^r \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^{r \times M}. \tag{2.54}$$

Once $\boldsymbol{Z}$ is solved for using Eq. (2.50), $\boldsymbol{u}_0^l$ is found an normalized using a method similar to Eq. (2.49),

$$\boldsymbol{u}_0^l \leftarrow \frac{\boldsymbol{u}_0^l}{\left\|\boldsymbol{u}_0^l\right\|_2} \quad \text{and} \quad s^l \leftarrow s^l \left\|\boldsymbol{u}_0^l\right\|_2. \tag{2.55}$$

After $\{\boldsymbol{u}_0^l\}$ has been estimated and normalized, and if the solution fails the criteria for convergence, then either another rank is added or the alternation through the directions continues. As discussed in [20], the error $\|\boldsymbol{q} - \hat{\boldsymbol{q}}\|_D$ can never increase throughout these ALS updates.

Rank increase and solution convergence. In order to develop an SR estimate that achieves desired accuracies, a set of tolerances should be defined. One must consider the minimization of error $\|\boldsymbol{q} - \hat{\boldsymbol{q}}\|_D$ between the surrogate and training data as well as the accuracy improvement from one full ALS sweep to the next. In order to enforce a maximum r, the user could set the desired maximum rank or aim for a particular solution precision. The method used in this work sets a desired maximum rank, but a precision based solution achieves convergence when the relative residual

$$\gamma = \frac{\|\boldsymbol{q} - \hat{\boldsymbol{q}}\|_D}{\|\boldsymbol{q}\|_D} < \epsilon, \tag{2.56}$$

in which ϵ is a desired relative tolerance. Theoretically, the approximation developed from SR should decrease γ as the separation rank r is increased. However, there is a limit to the precision that can be reached with a fixed rank [20]. The ALS process is repeated until the difference between the surrogate and training data no longer changes significantly from one iteration to the next. If the ALS process has converged on a solution, but the minimum relative residual has yet to be reached, another rank may be added.

In order to identify such a case, the current implementation uses the difference in relative residuals Eq. (2.56) as an indication of convergence. Specifically, the rank is increased when the solution convergence has not been met and the difference between the relative residual from the most recent iteration and the relative residual from two iterations previous is below some desired relative tolerance δ. This process is able to determine when improvement in the precision of the surrogate stalls or becomes insignificant.

Remark 4. *We highlight that choosing a small* ϵ *(or an unnecessarily large* r*) may result in over-fitting; that is, the difference between* $\boldsymbol{q}(\boldsymbol{y})$ *and its SR approximation* $\hat{\boldsymbol{q}}(\boldsymbol{y})$ *may be large, while* γ *is small. To avoid this issue and as discussed in [36], the least squares problems Eq. (2.49) and Eq. (2.50) may be regularized and appropriate error indicators may be utilized to estimate an optimal value for* ϵ *or* r*.*

2.5.2.4 ALS Algorithm and Computational Cost

In this section, two different approaches to depicting the construction of an SR surrogate are presented. The first, Figure 2.1, illustrates the process used for creating a $d = 3$ surrogate of a scalar output. The scalar output formulation of SR is chosen for the sake of illustrative simplicity. Here, we see the algorithm alternating through the three sets of coefficient vectors represented in a color scheme, i.e., $\{\boldsymbol{c}_1^l\}_{l=1}^r$ in green, $\{\boldsymbol{c}_2^l\}_{l=1}^r$ in blue, and $\{\boldsymbol{c}_3^l\}_{l=1}^r$ in red. These coefficients go directly into computing the most up-to-date approximation, such that the following coefficient is found using all previously calculated sets of coefficients. Algorithm 1 summarizes the SR approximation process in a pseudo-code fashion, unlike the illustration of Figure 2.1. Here, the focus is on the broader algorithm, where processes such as updating s^l are included.

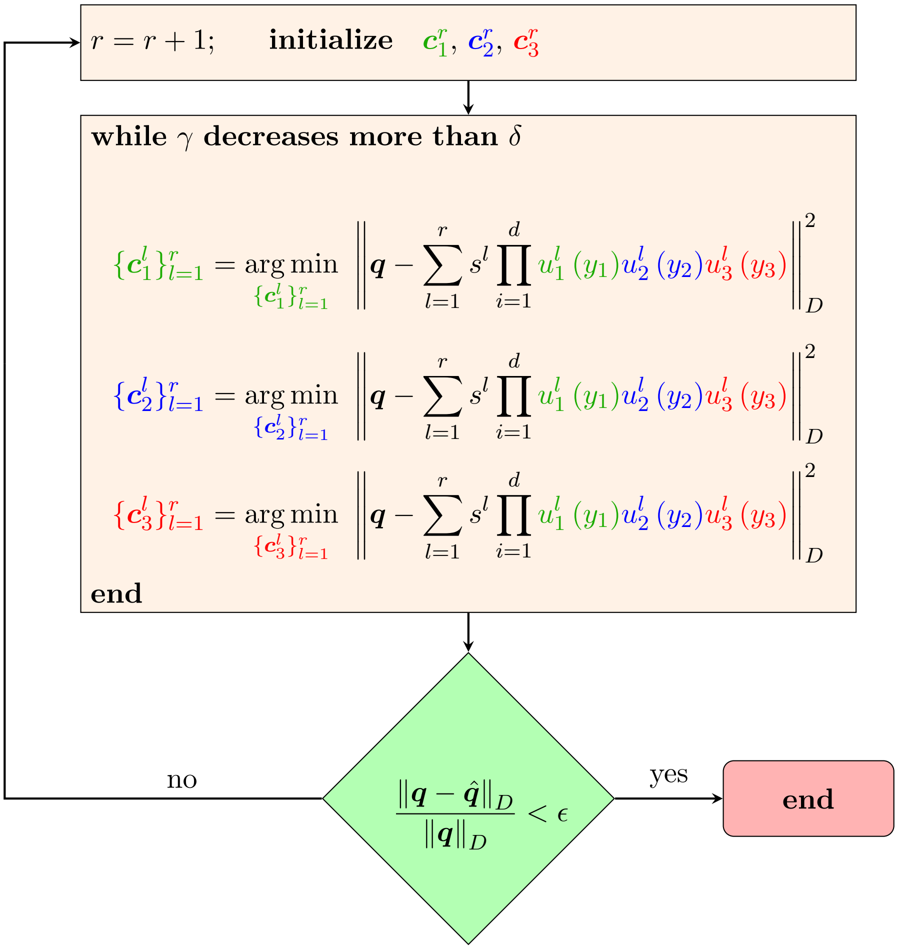


Figure 2.1: Flowchart of the ALS process with $d = 3$

```
$r = 0$;
while $\gamma > \epsilon$ do
    $r = r + 1$;
    Initialize $c_{i'}^r$, $i' = 1, \ldots, d$, $u_0^r$, and $s^r$;
    while $\gamma$ decreases more than $\delta$ (See end of Section 2.5.2.3) do
        for $i' = 1$ to $d$ do
            Solve for $c_{i'}^l$ as elements of $z$ using least squares problem Eq. (2.44);
            Update $s^l$ and $c_{i'}^l$ using Eq. (2.49);
        end
        Solve for $u_0^l$ as columns of $Z$ using least squares problem Eq. (2.50);
        Update $s^l$ and $u_0^l$ using Eq. (2.55);
        Generate $\hat{q}(y)$ using Eq. (2.40);
        Calculate $\gamma$ using Eq. (2.56);
    end
end
```

Algorithm 1: Algorithm for an SR approximation

As discussed in [20] and [36], when $N \gg r(P+1)$, the total cost of generating and solving the least squares problems Eq. (2.44) and Eq. (2.50) for a full ALS sweep is $\mathcal{O}(d{\cdot}r^2{\cdot}(P+1)^2{\cdot}M{\cdot}N)$, which is linear in d. The total number of unknowns in the vector-valued SR Eq. (2.40) is $r{\cdot}(P+1){\cdot}d+r{\cdot}M$, which is linear in d, assuming that the separation rank r is independent of d. This therefore suggests a linear dependence of the number of samples N on d, i.e.,

$$N \sim \mathcal{O}(r \cdot (P+1) \cdot d + r \cdot M) \tag{2.57}$$

for a successful SR computation. Assuming this estimate for N, the total cost of a full ALS sweep grows quadratically in d. We highlight that, for situations where evaluating the QoI is significantly expensive, this cost is reasonable.

Remark 5. *As a reminder to the reader, the computation costs of PCEs are discussed once more. For results discussed in Chapter 5, the PC expansion is computed using a least squares regression. In this case, the number of required samples N_{PC} is given by [55] as*

$$N \sim \mathcal{O}\left(\Lambda_{PC}\right), \tag{2.58}$$

where Λ_{PC} is defined in Eq. (2.30).

As seen in Eq. (2.58) and its reliance on Eq. (2.30), the required number of samples N for a PC expansion increases much faster with respect to dimension d than the largely linear relationship of SR previously discussed in this section, i.e. $N \sim \mathcal{O}(r \cdot (P+1) \cdot d + r \cdot M)$.

2.5.3 Solution Statistics

Once the coefficients $\{\boldsymbol{c}_i^l\}$ and $\{\boldsymbol{u}_0^l\}$ have been calculated, the statistics of $\boldsymbol{q}(\boldsymbol{y})$ may be approximated using the surrogate model $\hat{\boldsymbol{q}}(\boldsymbol{y})$. This can be done by sampling $\hat{\boldsymbol{q}}(\boldsymbol{y})$ using a Monte Carlo method or analytically from the coefficients $\{\boldsymbol{c}_i^l\}$ and the deterministic modes $\{\boldsymbol{u}_0^l\}$. Either of these methods may be applied to reduce the computation time of solution statistics when compared to traditional Monte Carlo methods. Notice that the accuracy of the approximate statistics depends on the accuracy of $\hat{\boldsymbol{q}}(\boldsymbol{y})$, which may be first verified based, for instance, on validation experiments as illustrated in the examples of Chapter 5.

SR relies on a polynomial basis that is orthogonal with respect to the density of the inputs. Following the Wiener-Askey theme, examples include using Hermite polynomials for Gaussian inputs or Legendre polynomials for uniform inputs [14]. When integrated over the support of the probability density function, the polynomials are shown to form an orthonormal basis, i.e.,

$$\int \psi_p(y_i)\psi_{p'}(y_i)\rho(y_i)dy_i = \delta_{p,p'}, \tag{2.59}$$

where $\rho(y_i)$ is the probability density function of direction i, and δ is the Dirac-delta function, indicating that $\psi_p(y_i)$ forms an orthogonal and normalized basis when considering $\rho(y_i)$.

The analytical mean and second moment of an SR solution are derived in [36]. The mean for the mth entry of $\boldsymbol{q}(\boldsymbol{y})$ is

$$E\left(\hat{q}_m\right) = \sum_{l=1}^{r} s^l\, u_{0,m}^l \prod_{i=1}^{d} c_{i,1}^l, \tag{2.60}$$

When considering the $m \times m'$ covariance matrix for $m = 1 \ldots M$ and $m' = 1 \ldots M$, each value $COV_{mm'} = E\left((\hat{q}_m - E(\hat{q}_m))(\hat{q}_{m'} - E(\hat{q}_{m'}))^{\mathrm{T}}\right)$ is computed as

$$COV_{mm'} = \sum_{l=1}^{r}\sum_{l'=1}^{r} s^l\, u_{0,m}^l\, s^{l'}\, u_{0,m'}^{l'} \prod_{i=1}^{d}\left(\sum_{p=1}^{P} c_{i,p}^l c_{i,p}^{l'}\right) - E\left(\hat{q}_m\right) E\left(\hat{q}_{m'}\right). \tag{2.61}$$

Using Eq. (2.61), the variances of $\hat{q}$ are found along the diagonal, or when $m = m'$.

For the low-order moments, these analytical methods are computationally more efficient than a traditional Monte Carlo sampling performed on $\hat{\boldsymbol{q}}(\boldsymbol{y})$. In general, higher order moments of q_m may be estimated using a method similar to Eq. (2.60) and Eq. (2.61), but a statistical sampling method of $\hat{q}_m$ may also be employed [36]. This sampling method may also be used to construct the PDF of $\boldsymbol{q}$. By evaluating the solution $\hat{\boldsymbol{q}}(\boldsymbol{y}_j)$ for large N, an estimate of the PDF of $\boldsymbol{q}$ is found without a large number of ODE solves. The results of this method can be seen in Chapter 5.

2.6 Solution Sensitivity

Unfortunately, in order to calculate $\{S_{i,m}\}$, Eq. (2.9) must be computed $N_{MC} \times d$ times. When considering the law of large numbers and the convergence rate of MC, N_{MC} is often an intractable number of computations. Therefore, repeating this d times further exceeds tolerable amounts of computation time. Therefore, it is possible to significantly increase efficiency by replacing all evaluations of $q_m(\boldsymbol{y}_j)$ in Eq. (2.7-2.9) with a surrogate approximation $\hat{q}_m(\boldsymbol{y}_j)$. In the case of the two presented methods, if d is large, then choosing SR over PCEs saves additional computation time due to lower requirements in the number of training samples N. The advantage of an SR approach lies in its separated structure and the relationship between computation costs and the dimension of $\boldsymbol{y}$. A larger dimension can be analyzed tractably when compared to methods that suffer from the curse of dimensionality. Taking advantage of a large dimension ensures a more complete view of the sensitivities, which leads to more informed decisions on system design and the selection of random inputs for operations. The surrogate approach of PC has been previously shown to reduce computation cost over a Monte Carlo based sensitivity analysis [21]. Chapter 5 provides examples of comparing a PCE evaluated global sensitivity analysis to that of SR, including a discussion of computation cost, as well as insight provided on case studies by the values of $\{S_{i,m}\}$.

Chapter 3

Optimization Under Uncertainty

3.1 Introduction

A common application of UQ is that of a conjunction assessment. Although a simplification of linear motion is appropriate in certain conjunction cases where relative velocities are large, nonlinear motion between objects has been shown to occur in cases of low relative velocity or long term encounters [100, 66, 68, 31]. Therefore, we seek to develop a method of conjunction assessment that does not make such a linear motion assumption. Once this capability is established, it is desirable to physically avoid any identified possible collisions. Therefore, the methodology of OUU is applied to maneuver design.

In order to present the material necessary for OUU with SR, this chapter first introduces OUU in the context of Monte Carlo. We then introduce the SR formulation for OUU applications, as well as the analytic moment equations that are compatible with an OUU problem. Motivation for finding an alternative to MC is also discussed, due to the fact that the methodology of OUU with MC is intractable as applied to this case. The presented alternative is that of separated representation (SR). In Section 3.4, we provide two initial test cases on which the performance of OUU with SR is compared to a MC approach as well as a known solution calculated via a Lambert Solver. These provide a validation background, which Chapter 6 expands upon.

3.2 Optimization Under Uncertainty

The methodology of OUU allows for the optimization of a problem while taking into account the uncertainty or variability of a system [39, 73]. By doing so, a prescribed reliability index or statistical constraint is considered along with the more traditional optimization problem, thus producing a solution that is more robust than an approach that does not take uncertainty into account. The work contained within this chapter utilizes a reliability index χ as a constraint in the optimization problem. Using the presented notation, this reliability design is generally formulated as

$$\begin{aligned} \text{minimize} \quad & f(\boldsymbol{\xi}, \boldsymbol{\theta}) \\ \text{subject to} \quad & \bar{\chi}(\boldsymbol{\xi}, \boldsymbol{\theta}) \geq \chi, \end{aligned} \tag{3.1}$$

where the design variables $\boldsymbol{\theta}$ characterize the *design space* of the problem. In Eq. (3.1), we have some function of the random and design inputs f that we seek to minimize. A deviation from that of a traditional optimization problem lies in the dependence of f not only on the design inputs $\boldsymbol{\theta}$ but also on the random variables $\boldsymbol{\xi}$. A further difference of reliability design when combined with traditional methods comes from the inclusion of the inequality $\bar{\chi}(\boldsymbol{\xi}, \boldsymbol{\theta}) \geq \chi$. The value of $\bar{\chi}(\boldsymbol{\xi}, \boldsymbol{\theta})$ is that of the chosen reliability index dependant on the previously mentioned random variables, as well as with the design inputs. These particular values of the design inputs $\boldsymbol{\theta}$ are those that have been computed in the minimization of f. A solution is found if the design inputs associated with the minimized value of $f(\boldsymbol{\xi}, \boldsymbol{\theta})$ also satisfy the stochastic inequality constraint.

As mentioned in Chapter 1, OUU also supports robust design. For a robust design based OUU, Eq. (3.1) is altered such that the variance of the design is limited, i.e.,

$$\begin{aligned} \text{minimize} \quad & f(\boldsymbol{\xi}, \boldsymbol{\theta}) \\ \text{subject to} \quad & \sigma^2 \leq \bar{\sigma}^2, \end{aligned} \tag{3.2}$$

where $\bar{\sigma}^2$ is a maximum allowable variance of the design. The approach of robust design does not necessarily rely on UQ, as "robustness" may be quantified using local derivatives. We note that robust design, although applicable in many cases, is not utilized in the results of this thesis.

As an example, consider the problem of designing a wrench to minimize material use but also remain within an acceptable rate of mechanical failure [73]. OUU is applied by considering manufacturing tolerance errors and the variability of applied forces, thereby altering the would-be design that only acknowledged a perfectly determined system. These uncertainties introduce statistics to the problem, allowing for a reliability index χ (such as the probability of failure) to be defined and evaluated.

For the sake of clarity, additional notation is provided here. To reiterate, the input variables of $\boldsymbol{\theta}$ are those associated with the *design space* of the problem, while $\boldsymbol{\xi}$ remain those associated with the inherent randomness of the system. These are referred to as the design and random variables, respectively. In order to track the *dimension* of the problem, we introduce the variable d, which is the *total dimension* of the problem. If one wishes to categorize the contributions to the dimension of the problem, d can be split between the dimension of the design space d_θ and that of random variables d_ξ, where $d = d_\xi + d_\theta$.

The ability to consider optimization as an option is enabled by introducing design variables to the problem. To apply the variables to the problem space, the design inputs are mapped to parameters dictated by the problem design. When used in this application, the design variables $\boldsymbol{\theta} \in \mathcal{R}^{d_\theta}$ are first given equal weight and defined over the domain $[-1, 1]$. The design variables are then mapped to deterministic inputs considered for optimization using the equation

$$\Xi(\theta) = \frac{\Xi_{max} - \Xi_{min}}{2}(\theta + 1) + \Xi_{min} \tag{3.3}$$

where Ξ_{max} and Ξ_{min} indicate the upper and lower bounds chosen for a deterministic input. As a function of θ, the deterministic inputs $\Xi(\theta)$ allow for the problem to explore a larger space of solutions. For instance, consider Ξ to be an input in a black box representing the time of a rocket launch, e.g., the span between 13:00 and 14:00 local time. For OUU purposes, the domain of θ is mapped to the limits of Ξ, thus allowing for the variation of the time of the launch as an input. By choosing a particular value of θ and evaluating the black box propagator M_{MC} times with different $\boldsymbol{\xi}$, the uncertainty of a rocket launch (with user determined QoIs) can be quantified at any point along the launch window. Using the examples that we have discussed, a value of $\theta = 1$ corresponds to a launch time of 14:00, while $\theta = -1$ accounts for a time of 13:00, and $\theta = 0$ represents the time of 13:30. Thus, it becomes possible that, with a chosen value of θ, the system can be optimized with respect to some predetermined scheme, e.g., a launch time that minimizes the total velocity needed from the third stage.

Remark 6. *As can be inferred, the computational cost of OUU is directly related to that of the black box propagator needed to evaluate* M_{MC} *samples. This cost would be repeated every time a new value of* $\boldsymbol{\theta}$ *is considered by the optimization algorithm. Therefore if the optimizer requires numerous iterations in the design space (as is typical), the cost of OUU becomes intractable when considering an* M_{MC} *large enough for converged statistics. Thus, there is motivation to replace a traditional black box propagator with a more computationally efficient method.*

3.3 Surrogate Approach

When considering an OUU formulation of Eq. (2.40), alterations are made such that

$$\boldsymbol{q}(\boldsymbol{\xi}, \boldsymbol{\theta}) \approx \hat{\boldsymbol{q}}(\boldsymbol{\xi}, \boldsymbol{\theta}) = \sum_{l=1}^{r} s^l \, \boldsymbol{u}_0^l \prod_{i'=1}^{d_\theta} u_{i'}^l \left(\theta_{i'}\right) \prod_{i=1}^{d_\xi} u_i^l \left(\xi_i\right), \tag{3.4}$$

where the previous presentation has been altered to include the relevant inputs of $\boldsymbol{\xi}$ and $\boldsymbol{\theta}$. The index of $i' = 1, \ldots, d_\theta$ is provided in order to separate the indexing of the design and random variables.

Remark 7. *It is brought to the reader's attention that although we have provided a distinction between* ξ *and* θ *for problem statement purposes, the product series seen in Eq. (3.4) is mathematically equivalent to the product series found in Eq. (2.40) when considering* $\boldsymbol{y} = \begin{bmatrix} \boldsymbol{\xi} & \boldsymbol{\theta} \end{bmatrix}^T$. *In the context of this example,* $\boldsymbol{\theta}$ *are implemented as uniform distributions while* $\boldsymbol{\xi}$ *are chosen only as Gaussian, but this limitation is not a general rule. Indeed, the distinction of being defined over the domain* $[-1, 1]$ *with equal weight is not unique to the design variables, as random variables are not excluded from being uniformly distributed.*

3.3.1 Analytic Moments

If the first or second moments are desired in the context of OUU, (2.60) may be rewritten to include the separation of design variables from the random inputs

$$\mathbb{E}[\hat{\boldsymbol{q}}(\boldsymbol{\theta}, \boldsymbol{\xi})] = \sum_{l=1}^{r} s_l\, \boldsymbol{u}_0^l \prod_{i'=1}^{d_\theta} u_{i'}^l\left(\theta_{i'}\right) \left\langle \prod_{i=1}^{d_\xi} u_i^l\left(\xi_i\right) \right\rangle, \tag{3.5}$$

where the inner product is defined as

$$\langle a_1(\boldsymbol{\xi}) \cdot a_2(\boldsymbol{\xi}) \cdots a_d(\boldsymbol{\xi}) \rangle = \int a_1(\boldsymbol{\xi}) \cdot a_2(\boldsymbol{\xi}) \cdots a_d(\boldsymbol{\xi}) \rho(\boldsymbol{\xi}) d\boldsymbol{\xi}. \tag{3.6}$$

In Eq. (3.6), $a_d(\cdot)$ is some general function of the inputs $\boldsymbol{\xi}$. These inputs have the density function $\rho(\boldsymbol{\xi})$.

Considering the inner product in Eq. (3.5) is taken over the set of N random variables, we leverage the independence of random inputs $\xi_{i,j}$. The inner product in Eq. (3.5) can then be replaced with,

$$\mathbb{E}[\hat{\boldsymbol{q}}(\boldsymbol{\theta}, \boldsymbol{\xi})] = \sum_{l=1}^{r} s_l\, \boldsymbol{u}_0^l \prod_{i'=1}^{d_\theta} u_{i'}^l\left(\theta_{i'}\right) \prod_{i=1}^{d_\xi} \left(\int u_i^l(\xi_i) \rho(\xi_i) d\xi_i \right). \tag{3.7}$$

As before, the univariate functions $u_i^l\left(\xi_i\right)$ may be approximated as

$$u_i^l\left(\xi_i\right) = \sum_{p=1}^{P} c_{p,i}^l \psi_{p,i}. \tag{3.8}$$

Therefore,

$$\mathbb{E}[\hat{\boldsymbol{q}}(\boldsymbol{\theta}, \boldsymbol{\xi})] = \sum_{l=1}^{r} s_l\, \boldsymbol{u}_0^l \prod_{i'=1}^{d_\theta} u_{i'}^l\left(\theta_{i'}\right) \prod_{i=1}^{d_\xi} \left(\int \left(\sum_{p=1}^{P} c_{p,i}^l \psi_{p,i} \right) \rho(x_i) dx_i \right). \tag{3.9}$$

Due to orthogonality, the general formulation for the analytic expression of an SR mean as a function of design variables simplifies to be

$$\mathbb{E}[\hat{\boldsymbol{q}}(\boldsymbol{\theta}, \boldsymbol{\xi})] = \sum_{l=1}^{r} s_l\, \boldsymbol{u}_0^l \prod_{i'=1}^{d_\theta} u_{i'}^l\left(\theta_{i'}\right) \prod_{i=1}^{d_\xi} c_{0,i}^l \psi_{0,i}, \tag{3.10}$$

where (in the case of an orthogonal polynomial) $\psi_{0,i} = 1$,

$$\therefore \mathbb{E}[\hat{\boldsymbol{q}}(\boldsymbol{\theta}, \boldsymbol{\xi})] = \sum_{l=1}^{r} s_l\, \boldsymbol{u}_0^l \prod_{i'=1}^{d_\theta} u_{i'}^l\left(\theta_{i'}\right) \prod_{i=1}^{d_\xi} c_{0,i}^l. \tag{3.11}$$

Using the same methodology, the analytic expression for the standard deviation σ may be derived. Once again by separating the design and stochastic variables, the analytic expression for the second moment is

$$\mathbb{E}[\hat{\boldsymbol{q}}(\boldsymbol{\theta},\boldsymbol{\xi})^2] = \sum_{l=1}^{r}\sum_{l'=1}^{r'} s_l\, s_{l'}\, \boldsymbol{u}_0^l\, \boldsymbol{u}_0^{l'} \prod_{i'=1}^{d_\theta}\left(u_{i'}^l\,(\theta_{i'})\; u_{i'}^{l'}\,(\theta_{i'})\right)\left\langle \prod_{i=1}^{d_\xi} u_i^l\,(\xi_i)\; u_i^{l'}\,(\xi_i)\right\rangle . \tag{3.12}$$

By replacing the inner product in (3.12) with an integral, the second moment is formulated as

$$\mathbb{E}[\hat{\boldsymbol{q}}(\boldsymbol{\theta},\boldsymbol{\xi})^2] = \sum_{l=1}^{r}\sum_{l'=1}^{r'} s_l\, s_{l'}\, \boldsymbol{u}_0^l\, \boldsymbol{u}_0^{l'} \prod_{i'=1}^{d_\theta}\left(u_{i'}^l\,(\theta_{i'})\; u_{i'}^{l'}\,(\theta_{i'})\right)\prod_{i=1}^{d_\xi}\left\{\int u_i^l\,(\xi_i)\; u_i^{l'}\,(\xi_i)\,\rho(\xi_i)d\xi_i\right\} . \tag{3.13}$$

Substituting (3.8) into (3.13)

$$\begin{aligned}\mathbb{E}[\hat{\boldsymbol{q}}(\boldsymbol{\theta},\boldsymbol{\xi})^2] = &\sum_{l=1}^{r}\sum_{l'=1}^{r'} s_l\, s_{l'}\, \boldsymbol{u}_0^l\, \boldsymbol{u}_0^{l'} \prod_{i'=1}^{d_\theta}\left(u_{i'}^l\,(\theta_{i'})\; u_{i'}^{l'}\,(\theta_{i'})\right)\\ &\prod_{i=1}^{d_\xi}\left\{\int\left(\sum_{p=1}^{P} c_{p,i}^l\psi_{p,i}\right)\left(\sum_{p=1}^{P} c_{p,i}^{l'}\psi_{p,i}\right)\rho_i d\xi_i\right\} .\end{aligned} \tag{3.14}$$

Due to the orthonormality of $\psi_{p,i}$,

$$\int \psi_{p,i}\psi_{p',i} = \delta_{p,p'} . \tag{3.15}$$

(3.14) then simplifies to

$$\mathbb{E}[\hat{\boldsymbol{q}}(\boldsymbol{\theta},\boldsymbol{\xi})^2] = \sum_{l=1}^{r}\sum_{l'=1}^{r'} s_l\, s_{l'}\, \boldsymbol{u}_0^l\, \boldsymbol{u}_0^{l'} \prod_{i'=1}^{d_\theta}\left(u_{i'}^l\,(\theta_{i'})\; u_{i'}^{l'}\,(\theta_{i'})\right)\prod_{i=1}^{d_\xi}\left(\sum_{p=1}^{P} c_{p,i}^l c_{p,i}^{l'}\right) . \tag{3.16}$$

By taking the square root of the difference between the evaluation of (3.16) and the square of the first moment, the standard deviation about the mean is found.

3.4 Initial Test Cases

As a proof of concept, two test cases are presented in this section. The purpose of the first case is to compare the accuracy of an SR OUU application to a result in the literature as well as a MC reference. The second case is contrived in such a way that the performance of optimizing with the surrogate method can be compared to that of a deterministic method; specifically, we compare to a Lambert Solver. For this, the only random inputs of the surrogate are design variables, thus posing a deterministic optimization problem.

3.4.1 Rosenbrock Application

In order to validate SR's performance for OUU, a test case previously used in the existing OUU literature is examined. The results of using PC for OUU by [39] are compared to MC and SR results from this section. By using a relatively simple function to generate samples for a black box optimizer, the optimization potential of SR can be compared with that of MC and established research. In this case, the optimizer is MATLAB's *fmincon*. As in [39], other optimizers, such as DAKOTA [1], may also be used for minimization purposes. In this example, the SR will be utlized inside *fmincon* to find the mean and standard deviation of a function with uncertainty. The function in question is the Rosenbrock function

$$q(\boldsymbol{y}) = 100(\Xi(\theta) - \xi^2)^2 + (1 - \Xi(\theta))^2 \tag{3.17}$$

As discussed in Section 3.2, $\Xi(\theta)$ is the deterministic variable, while ξ is a source of uncertainty. As a design variable, θ is a variable that is sampled within the optimizer. As the source of uncertainty, ξ is a normally distributed random inputs. The optimization goal is to maximize a reliability index β_{CDF}

$$\beta_{CDF} = \frac{\bar{\mu} - \bar{z}}{\bar{\sigma}} \tag{3.18}$$

where $\bar{z} = 10$. In Eq.(3.18), the values of $\bar{\mu}$ and $\bar{\sigma}$ correspond to the evaluations of the first and second moment about the means found via Eqs.(3.11) and (3.16), respectively. Here, the bar notation denotes that a particular value of $\boldsymbol{\theta}$ is chosen by the optimizer.

The work by [39] uses a similar methodology by using means and standard deviations calculated with PC and optimization done using the DAKOTA software suite [1]. In the case of the Rosenbrock function, the design variable Ξ has the bounds $-2 \leq \Xi \leq 2$ and initial value of -0.75, while $\xi \sim \mathcal{N}(0, \mathbb{I})$. As stated previously, the goal is to maximize β_{CDF} of (3.18). Using values of $N = 900$, $r = 3$, $P = 5$ and $\delta = 10^{-6}$, an SR was created. The result of [39] is compared to optimization results using *fmincon* with MC and SR. The MC results use $1,000,000$ values of ξ during the optimization process. Table 3.1 shows values for β_{CDF} and Ξ calculated via each method. It is important to note that *fmincon* tends to avoid Ξ values at either bound. Also included in Table 3.1

are run times for *fmincon* using SR or MC to perform OUU. The elapsed time for using MC is approximately eight times longer than using SR, primarily due to the need to generate $1,000,000$ MC samples for each optimization iteration. Figure 3.1 shows the relative error between expectation values derived from (3.11) and (3.16) and a MC analysis using $1,000,000$ samples of ξ. In this chapter, the relative error is defind as

Table 3.1: β_{CDF} and Ξ of Optimized Rosenbrock Function

Method	β_{CDF}	Ξ	Run Time (s)
[39]	2.0913	-2	N/A
$fmincon_{\mathrm{SR}}$	2.0913	$-2+1.0425e-06$	0.085
$fmincon_{\mathrm{MC}}$	2.0914	$-2+1.0421e-06$	0.691

$$\epsilon_{rel} = \left|\frac{\hat{\lambda}-\lambda}{\lambda}\right|, \tag{3.19}$$

where λ is the quantity being approximated by $\hat{\lambda}$. The abscissae contain the values of θ which are used to generate all values of Ξ since Legendre polynomials are orthogonal on the bounds $-1 \leq \theta \leq 1$. It is possible that this systematic error is due to overfitting.

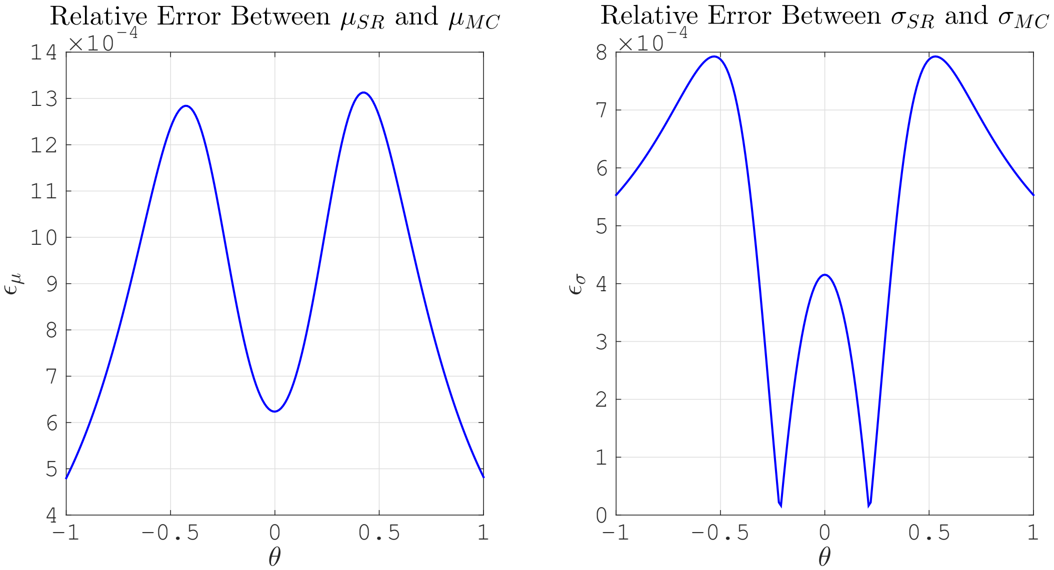


Figure 3.1: Relative error of mean ϵ_μ and standard deviation ϵ_σ.

3.4.2 Maneuver Design

One possible application of OUU in astrodynamics is maneuver design. Optimizing a maneuver is crucial, because on board fuel often makes up a significant portion of a spacecraft's operational weight. This weight dictates launch costs and the scientific payload that can fit on the remaining volume and weight budget. In addition to this, successive maneuver optimizations can lead to a longer lifespan, thereby extending the operation and increasing the value of the craft. As an example, consider the routine use of "stochastic maneuvers". These particular changes in velocity, also known as "clean-up maneuvers", are used to correct errors introduced by previously executed maneuvers. Therefore, it is possible to save fuel is these stochastic maneuvers are not necessary. Due to the UQ required to characterize the system, OUU is a candidate for such a task.

For a brief proof of concept, a scenario of a maneuvering craft seeking to intercept an object at a given location is considered. The spacecraft starts in an equatorial, circular orbit with an altitude of 200 km. Assuming a flight time of $t = 46.755$ minutes, a Lambert solver is used to determine the trajectory and $\Delta\boldsymbol{V}$ necessary to reach the target 179° away at an altitude of 500 km. This $\Delta\boldsymbol{V}_0$ is considered the nominal maneuver and is included in Table 3.2 along with other parameters used in the design of the orbits and maneuver. It is noted that no truly random inputs are used in this particular example. It is formulated such that the focus is on deterministic inputs and the ability of SR to solve for a known optimization solution. In this case, the reference optimal maneuver is calculated via the previously mentioned Lambert solver.

Considering this, the nominal maneuver is used as the mean of the design inputs $\Xi_{\Delta V_x}(\theta_1)$, $\Xi_{\Delta V_y}(\theta_2)$, $\Xi_{\Delta V_z}(\theta_3)$. Using 10% of $\|\Delta\boldsymbol{V}_0\|$ to dictate the bounds of our deterministic inputs $\boldsymbol{\Xi}$, the surrogate is able to consider the following bounds of maneuver components

$$\begin{aligned} 0.0214 \leq \Delta V_x \leq 0.0505 \quad & \text{km/s} \\ 0.1266 \leq \Delta V_y \leq 0.1557 \quad & \text{km/s} \\ -0.0145 \leq \Delta V_z \leq 0.0145 \quad & \text{km/s}, \end{aligned} \tag{3.20}$$

Table 3.2: Parameters of the Maneuver Scenario

	Value	Units
a_1	6571	km
a_2	7071	km
e_1	0	N/A
e_2	0	N/A
ν_1	0	degrees
ν_2	179	degrees
ΔV_x	0.035997	km/s
ΔV_y	0.141187	km/s
ΔV_z	0	km/s
$\|\Delta \boldsymbol{V}_0\|$	0.145704	km/s

Using a two body propagator, training samples for the position and velocity of the craft at time t are created. These training samples are used to create an SR that outputs Cartesian position and velocity of the spacecradt and is a function of the $d = 3$ design inputs $\boldsymbol{\theta}$. Figure 3.2 illustrates the PDF of the spacecraft at time t, as well as the SR estimation evaluated at 50,000 independent input samples. The PDF of each element is highly non-Gaussian, but via a qualitative assessment, it is shown to be estimated well by the SR surrogate.

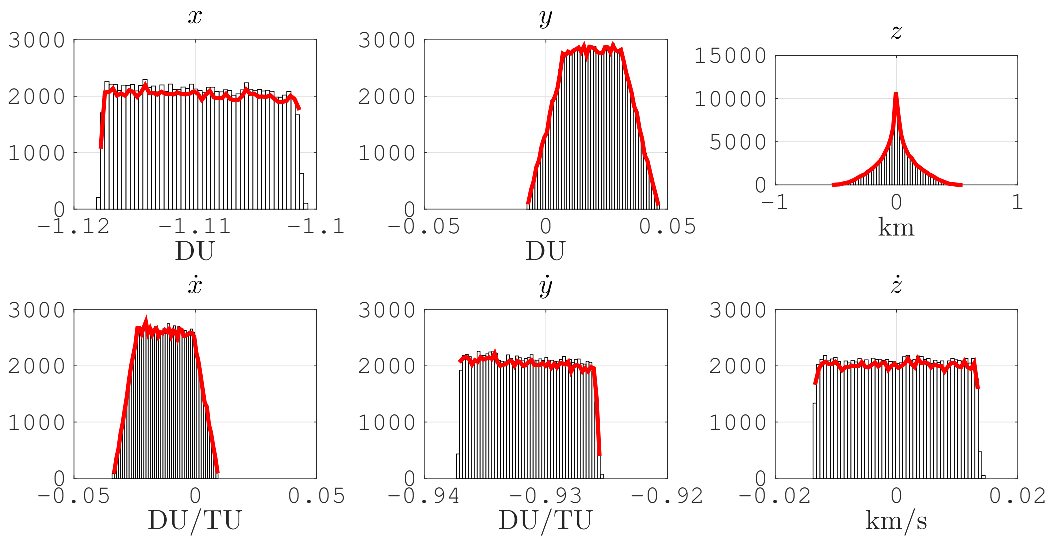


Figure 3.2: PDFs of MC (bins) and SR solution (red line) for the maneuver test case

To test SR's capabilities with OUU as applied to astrodynamics, we consider the simple case of minimizing the distance between the craft and target at time t. Because the only inputs are that of the design variables, the optimized values of the burn should be that of the nominal solution $\Delta\boldsymbol{V}_0$. As expected, on the limits of $-1 \leq \boldsymbol{\theta} \leq 1$, the SR and *fmincon* converged to an accurate estimation of the nominal solution. The SR solution is accurate to within at least five digits, and the results can be seen in Table 3.3.

Table 3.3: Results for Maneuver Optimization

Method	$\Delta V_x(\theta_1)$	$\Delta V_y(\theta_2)$	$\Delta V_z(\theta_3)$
Lambert Solver	0.035998	0.14119	0
$\textit{fmincon}_{\text{SR}}$	0.035972	0.14118	$2.452e-07$
Relative Error	$7e-4$	$1e-05$	N/A

Chapter 4

Multi-Element Separated Representations

4.1 Introduction

First formulated in [120], ME-gPC is presented as a method of adaptively discretizing the random space to achieve K-P_{PC} convergence, where K is the number of elements, or partitions, in random input space and P_{PC} is the polynomial chaos order. By leveraging a method to detect the need for spatial decomposition, splits, or elements, are created along input directions where needed. The approach of partitioning a large, difficult problem into more numerous smaller problems allows ME-gPC to accurately converge to a solution when unmodified PCEs require a prohibitively high order expansion. This adaptive method is presented in [120] with examples using uniform inputs. The work of [121] generalizes the approach by applying an arbitrary probability measure method of ME-gPC to Gaussian and Beta distributions. Due to the decomposition of the input space effectively removing the orthogonal relationship between input space and polynomial basis, care is needed when formulating a PCE for an element. Therefore, [121] demonstrates the numerical construction of orthogonal polynomials (with respect to an element's conditional PDF) for an arbitrary distribution.

As applied to astrodynamics, ME-gPC is shown to provide a performance improvement regarding accuracy when compared to general PCEs [42]. Since PCEs suffer from the curse of dimensionality, high dimension systems quickly become intractable for general PCEs without mitigation. The potential for high computation costs is motivation to research a multi-element formulation of SR (ME-SR). Theoretically, an ME-SR surrogate would be capable of efficiently estimating a high-dimensional, multi-modal stochastic function without making posterior PDF assumptions.

4.1.1 Chapter Overview

Although the multi-element approach has been devised for PCEs, modifications should be made when applying the methodology to SR. The unique formulation of SR (when compared to a PCE) dictates a different approach to choosing an input direction to decompose. Therefore, we present the processes necessary for constructing a ME-SR approximation. In order to do so, we first provide the methodology utilized to decompose the sample space. With this as a foundation, the formulation of SR is altered such that multiple elements of input space are considered. The unique approach to detecting the need and direction of element creation is then introduced. These processes are then organized into an algorithm which adaptively creates a multi-element surrogate for SR.

4.2 Multi-Element Separated Representations

For this chapter, we present the ME-SR surrogate formulation. Beginning with the decomposition on the input space, the properties of a surrogate element are described. Following this, insight is given on the construction of polynomials within these elements, while the last section covers the adaptive procedure that is implemented to identify and create splits in the random input space. It is noted that the notation and mathematics of this chapter are largely based on that found in [120] and [121] with customizations for SR as indicated. In this work, we provide a ME-SR formulation that is valid for scalar QoIs. A vector QoI is equally valid, and the absence of such a formulation is for the sake of simplicity.

4.2.1 Input Space Decomposition

Due to the same relationship between the polynomial base and the random inputs (which follows the Wiener-Askey scheme), both PCEs and SR have identical processes when considering the decomposition of input space. Thusly, we establish the random variables $\boldsymbol{y}$ on $B = \times_{i=1}^{d} [a_i, b_i]$

where a_i and b_i are the limits of the random variables for direction i. These limits can be finite or infinite on $\mathbb{R}$. The decomposition D of B is then defined as

$$D = \begin{cases} B_k = [a_{k,1}, b_{k,1}) \times [a_{k,2}, b_{k,2}) \times \cdots \times [a_{k,d}, b_{k,d}], \\ B = \cup_{k=1}^{K} B_k, \\ B_k \cap B_{k'} = \emptyset \text{ if } k \neq k', \end{cases} \tag{4.1}$$

where K is the total number of elements, and $k = 1, \ldots, K$. In order to evaluate equations over the entire input space B, we define an indicator function I_{B_k}, that is

$$I_{B_k}(\boldsymbol{y}) = \begin{cases} 1 & \text{if } \boldsymbol{y} \in B_k, \\ 0 & \text{otherwise} . \end{cases} \tag{4.2}$$

Therefore, we are able to reconstruct the total sample space Ω via the union of all elements which result in a positive result for Eq. (4.2), i.e.,

$$\Omega = \cup_{k=1}^{K} I_{B_k}^{-1}(1). \tag{4.3}$$

Conversely, a property of these elements, and therefore the indicator function, is that there is no overlap of the probability space

$$I_{B_k}^{-1}(1) \cap I_{B_{k'}}^{-1}(1) = \emptyset \text{ for } k \neq k'. \tag{4.4}$$

Given the properties of the decomposed sample space, it is therefore possible to formulate Eq. (2.36) such that all of Ω is able to be considered via the random inputs $\boldsymbol{y}$ even though it is decomposed into D. First, we consider the SR surrogate of an individual element k. Utilizing Eq. (4.2),

$$\hat{q}_k(\boldsymbol{y}) = I_{B_k}(\boldsymbol{y}) \sum_{l=1}^{r} s_k^l \prod_{i=1}^{d} u_{k,i}^l(y_i), \tag{4.5}$$

with the element index k now included for the normalizing constants s_k^l and univariate functions $u_{k,i}^l$. The approximation of $q(\boldsymbol{y})$ on the k-th element is parsed with respect to the random inputs by the indicator function I_{B_k}, such that only inputs found in the k-th element are non-zero. Using Eq. (4.5), it is possible to now formulate an equation for a global approximation of $q(\boldsymbol{y})$ using a multi-element expression, i.e.,

$$q(\boldsymbol{y}) \approx \hat{q}^K(\boldsymbol{y}) = \sum_{k=1}^{K} \hat{q}_k(\boldsymbol{y}), \tag{4.6}$$

where the composition of the SR surrogate now includes a sum over the number of elements K. It is noted that although Eq. (4.6) is a sum of all elements, only one term is nonzero when considering an input $\boldsymbol{y}$. The indicator function I_{B_k}, found in Eq. (4.5), is only nonzero for the appropriate surrogate $\hat{q}_k$.

4.2.2 Adaptive Element Construction

While the previous sections discuss how a surrogate is defined on an individual element k, the logic behind the creation of an element is to be elaborated upon. As in [120] and [121], we present an algorithm for determining when and where to split the input space. Due to differences in formulation, however, ME-SR differs from ME-gPC in the detection of the direction needing to be split. The performance of this algorithm is later tested in Chapter 7.

Considering Eq. (4.5) with the analytic moments in Eq. (2.60) and Eq. (2.61), we first give the moments of each individual element as

$$\mu_k = E\left[\hat{q}_k\right] = \sum_{l=1}^{r} s_k^l \prod_{i=1}^{d} c_{i,0,k}^l, \tag{4.7}$$

$$\sigma_k^2 = E\left[\hat{q}_k^2\right] = \left\{ \sum_{l=1}^{r} \sum_{l'=1}^{r} s_k^l s_k^{l'} \prod_{i=1}^{d} \left(\sum_{p=0}^{P} c_{i,p,k}^l c_{i,p,k}^{l'} \right) \right\} - \mu_k^2. \tag{4.8}$$

With the equations of the moments for each element, it is then possible to formulate the global mean $\bar{\mu}$ and variance $\bar{\sigma}^2$ of the multi-variante function $q(\boldsymbol{y})$. Since the input spaces for the elements are disjoint, moments for the QoI are given by summing the moments of each element with weights chosen based on the probability of $\boldsymbol{y} \in B_k$, i.e.,

$$\bar{\mu} = \sum_{k=1}^{K} \mu_k \Pr(I_{B_k} = 1), \tag{4.9}$$

$$\bar{\sigma}^2 = \sum_{k=1}^{K} \left[\sigma_k^2 + (\mu_k - \bar{\mu})^2\right] \Pr(I_{B_k} = 1). \tag{4.10}$$

In Eqs. (4.9) and (4.10), the quantity $\Pr(I_{B_k} = 1)$ is introduced. This indicates the cumulative probability of the element B_k within B, with the reliance of I_{B_k} on $\boldsymbol{y}$ dropped for conciseness. As seen in later results, the error of estimating the variance of $q(\boldsymbol{y})$ usually exceeds that of the estimated mean. Therefore, it is of interest to formulate a splitting algorithm which relies on quantities derived from the second moment.

Similar to previous ME-gPC work, this formulation of ME-SR relies on evaluating the relative contribution of the highest order polynomials of each element's variance. That is, we utilize the standard deviation of an individual element, presented in Eqs. (4.7) and (4.8), to define the decay rate parameter

$$\eta_k = \frac{\sum_{l=1}^{r} \sum_{l'=1}^{r} s_k^l \, s_k^{l'} \prod_{i=1}^{d} \; c^l_{i,P,k} c^{l'}_{i,P,k}}{\sigma_k^2}, \tag{4.11}$$

where it is noted that there is no sum over the polynomial order, due to the order of coefficients fixed at the maximum value P. The formulation of Eq. (4.11) differs from that found in [120] and [121] due to the use of the multi-index in ME-gPC. Here, the separable nature of SR does not mimic the combinatorial nature of the highest multi-indices in ME-gPC. Following the previously cited works in ME-gPC, we seek a *K-type refinement*. That is, the element k is decomposed if the probabilistically scaled decay rate parameter exceeds a certain value

$$\eta_k^{\gamma} \Pr(I_{B_k} = 1) \geq \theta_1, \quad 0 < \gamma < 1. \tag{4.12}$$

The values of γ and θ_1 are predetermined constants or tuning parameters. Since η_k is considered the decay rate parameter, this work refers to θ_1 as the *decay check* parameter when necessary.

Once an element has been determined to require a split, it is important to determine which direction in which to decompose the input space. In order to reduce exponential growth in computation cost (due to combinatorial effects of splitting more than one input at a time), it is best to split as few directions as possible. As previously mentioned, PCEs utilize a combinatorial multi index. Thusly in [120] and [121], the detection of which direction(s) to split is calculated via the coefficient associated with the highest order polynomial of direction i being divided by a sum over the highest multi indices over all directions. This sum over a multi-index is not a sum that is separable with respect to dimension as seen in Eq. (2.36). Therefore, SR requires a different approach.

In the current methodology to determine the direction(s) that should be split, we propose the method of constructing uni-directional surrogates. The direction that constructs a surrogate, which has the most significant improvement in the RMS error of a validation data set, is then considered the most sensitive direction. That is, we first construct an individual surrogate $\hat{q}_i(y_i)$ for each dimension which minimizes the error of $\|q(\boldsymbol{y}) - \hat{q}_i(y_i)\|_D$. Explicitly, the formulation of each uni-directional surrogate is given as

$$\hat{q}_i(y_i) = I_{B_k}(y_i) \sum_{l=1}^{r} s^l \, u_i^l \, (y_i) \, . \tag{4.13}$$

This one-dimensional surrogate is then evaluated for some validation data set independent of the training data

$$\mathcal{D}_{val} = \{(y_{i,j}, q(\boldsymbol{y}_j))\}_{j=1}^{N_{val}} \, , \tag{4.14}$$

where N_{val} is the sample size of the validation data set, and the sensitivity parameter is calculated using the RMS value of the uni-directional surrogate's validation error.

Thusly, the directional sensitivity parameter is formulated as

$$r_i = \|q(\boldsymbol{y}) - \hat{q}_i(y_i)\|_D, \qquad i = 1, \ldots, d. \tag{4.15}$$

If Eq. (4.12) results in a split requirement, then the results of Eq. (4.15) are used to determine which input direction(s) require further decomposition. A direction i is split if

$$r_i \leq \theta_2 \cdot \min_{i'=1,\ldots,d} r_{i'}, \qquad \theta_2 \geq 1, \tag{4.16}$$

where θ_2 is a prescribed tuning parameter which dictates the tendency of the algorithm to split more than just the minimum r_i (most sensitive) direction. In the case of more than one dimension satisfying Eq. (4.16), a combinatorial approach is taken to create elements such that the properties of Eq. (4.1) are maintained. Due to this, Eq. (4.16) is significant in maintaining a desired computational cost, as the number of elements increases exponentially with respect to dimension.

Once the directions to split are identified, the bounds of the new elements are created. As an example, we consider a "parent" element k that is split in one direction, resulting in two "children" k' and k''. For this given direction i, the bounds are evenly divided such that

$$\begin{aligned} b_{k',i} &= b_{k,i} \\ a_{k',i} &= \frac{1}{2}(a_{k,i} - b_{k,i}) \\ b_{k'',i} &= \frac{1}{2}(a_{k,i} - b_{k,i}) \\ a_{k'',i} &= a_{k,i} \end{aligned} \tag{4.17}$$

whereafter the element k is destroyed, and the newly created bounds of elements k' and k'' are organized in the decomposition of the input space, i.e. $[b_{k',i}, a_{k',i})$ and $[b_{k'',i}, a_{k'',i})$ are now the bounds of the i-th direction within the input spaces of $B_{k'}$ and $B_{k''}$ respectively. As previously stated, the splitting process results in an exponential increase of elements when considering the number of directions to split. That is, the choice to include the newly created upper or lower bounds found in Eq. (4.17) is a binomial coefficient problem. If d' is the number of directions which satisfy Eq. (4.16), the change in the total number of elements (when splitting a single element) is

$$\Delta K = \sum_{i'=0}^{d'} \binom{d'}{i'} - 1 = 2^{d'} - 1 \tag{4.18}$$

where ΔK is the stated change, and the subtraction accounts for the destruction of the parent element k.

4.3 Implementation Guidelines

Decomposing the input space into a series of elements creates implementation challenges. Therefore, this section is devoted to covering a number of issues to consider. First, numerical matters are presented with an approach to mitigate precision problems. Due to a need for an ME-SR algorithm which is compatible with arbitrary distributions, background is provided for various methods of constructing orthogonal polynomials. Lastly, an implementation algorithm is supplied as a guide for an overall view of the ME-SR methodology.

4.3.1 Numerical considerations

As in [120] and [121], this implementation also utilizes a random inputs normalization technique. As the difference between element bounds, e.g. $b_{k,1} - a_{k,1}$, becomes smaller, the absolute difference between random input samples within these bounds also becomes smaller. This tightening of the bounds may lead to numerical issues in the evaluation of orthogonal polynomial bases, and it results in underflow. To avoid underflow, we begin with the transformation of random inputs with limits of $[-1, 1]$ to some desired limits, i.e.,

$$y_i^k(\bar{y}) \equiv y(\bar{y}; k, i) = \frac{b_{k,i} - a_{k,i}}{2}\bar{y} + \frac{b_{k,i} + a_{k,i}}{2}, \tag{4.19}$$

where $\bar{y}$ is a random variable in $[-1, 1]$. With Eq. (4.19), the density function of a particular direction in an element can then become a function of $\bar{y}$, that is $\rho(y(\bar{y}; k, i))$. To reduce notation clutter, we introduce the shorter $\rho_{k,i}(\bar{y}) \equiv \rho(y(\bar{y}; k, i))$.

The test cases within Chapter 7 contain random inputs which include uniform and Gaussian distributions. Although the uniform distributions require no special treatment, the method of constructing polynomials orthogonal to the infinite bounds of a Gaussian distribution pose certain difficulties in the execution of the adaptive algorithm. To numerically generate polynomials, the implementation of this ME-SR algorithm utilizes truncated Gaussian distributions on the limits $[-6, 6]$, where ∞ and $-\infty$ are not considered for choices of $a_{k,i}$ or $b_{k,i}$. Since the cumulative probability of sampling outside of these bounds is $\sim 1.9 \times 10^{-9}$, the modeling errors imposed by not considering the infinite bounds of a true Gaussian are deemed acceptable.

4.3.2 Constructing Orthogonal Polynomials

When an input space is decomposed into more than one element, the orthonormal properties of Eq. (2.59) are likely lost. For a uniform distribution, it is possible to simply rescale the inputs to the [-1,1] limits required by the Legendre basis in order to maintain orthogonality. When considering a truncated Gaussian distribution for example, this simple rescaling does not result in an orthonormal relationship due to the weighting function of the k-th element $\rho_{k,i}(\bar{\boldsymbol{y}})$ differing from that of the unbroken input space regardless of any scaling parameter. An illustration of decomposed input spaces is seen in Figure 4.1. Here, the relatively simple decomposition of the uniform distribution is contrasted with that of the Gaussian distribution. For the latter, a linear scaling of the weighting function with respect to the element bounds does not maintain orthonormality with the same basis polynomials. Therefore, a new polynomial base must be constructed.

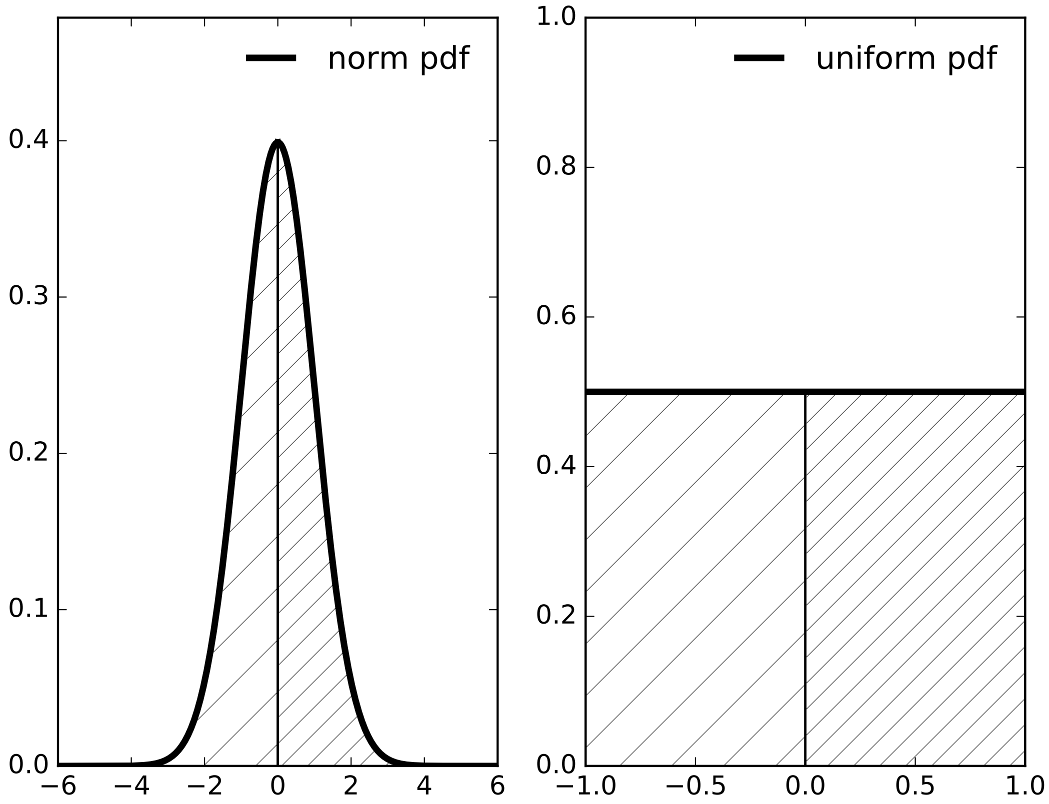


Figure 4.1: Probability density functions of Gaussian (left) and uniform (right) distributions with two elements

For such a construction, we first consider the three-term recurrence relation that is indicative of an orthogonal polynomial system

$$\begin{aligned} \pi_{p+1}(y) &= (y-\alpha_p)\pi_p(y) - \beta_p \pi_{p-1}(y), \quad p = 0, 1, \ldots, \\ \pi_0(y) &= 1, \quad \pi_{-1}(y) = 0, \end{aligned} \tag{4.20}$$

where the coefficients α_p and β_p correspond to the unique weighting function for which the orthogonal polynomial system is defined. This recurrance property is useful for a number of reasons. The first of which is that the reliance on the two parameters α_p and β_p allows for a compact representation of the polynomials. That is, only a linear array of $2(P+1)$ coefficients is needed. Furthermore, the derivatives of the polynomials may also be derived recursively using the derivative of Eq. (4.20) and the linear array of coefficients. Therefore, computing these coefficients is of interest in creating a three-term recurrance formula for an orthogonal polynomial basis. Together with $y_i^k(\bar{y})$ and $\rho_{k,i}(\bar{y})$, the newly constructed polynomials are able to satisfy Eq. (2.59) when considering the input bounds of the numerically stable $\bar{y}$ and the particular probability density function of the element's inputs.

In order to calculate such polynomials, two approaches are typically considered: the Stieltjes procedure and a modified Chebyshev method. In this work, we utilize the Python package Py-Orthpol for computing the coefficients. This package is based on FORTRAN code, which is elaborated upon in [49]. The following sections are devoted to a discussion of properties of orthogonal polynomials and how the two methods can be used to recursively construct a desired set of these functions. The Stieltjes procedure is introduced, followed by the Chebyshev, its modified version and a methodology for discretizing numerically ill-conditioned equations.

First, let $d\lambda(y)$ be a nonnegative measure on the real line $\mathbb{R}$. This measure can have compact or infinite support, and its moments are defined as

$$\mu_r = \int_{\mathbb{R}} y^r \, d\lambda(y), \quad r = 0, 1, 2, \dots, \tag{4.21}$$

where r, in this case, denotes the r-th moment of the measure $d\lambda(y)$. If this measure is absolutely continuous, then Eq. (4.21) simplifies to

$$\mu_r = \int_{\mathbb{R}} y^r \rho(y) \, dy, \quad r = 0, 1, 2, \dots, \tag{4.22}$$

where $\rho(y)$ is a non-negative function and $\int_{\mathbb{R}} \rho(y) \, dy > 0$ [109]. In the literature, $\rho(y)$ is referred to as the weighting function. Considering the application of the orthogonal polynomial basis for SR, we consider the probability density function of the variable y, as seen in Eq. (2.59). Together with Eq. (4.20), Eq. (4.22) allows us to construct both α_p and β_p.

4.3.2.1 Stieltjes Procedure

The construction of the Stieltjes procedure begins with considering Eq. (4.20). A characteristic of polynomials satisfying this formula is that the values of α_p and β_p can be found via

$$\begin{aligned} \alpha_p &= \frac{\int_{\mathbb{R}} y\pi_p^2 \rho(y)\,dy}{\int_{\mathbb{R}} \pi_p^2 \rho(y)\,dy}, \quad p = 0, 1, \ldots, \\ \beta_p &= \frac{\int_{\mathbb{R}} \pi_p^2 \rho(y)\,dy}{\int_{\mathbb{R}} \pi_{p-1}^2 \rho(y)\,dy}, \quad p = 1, 2 \ldots. \end{aligned} \tag{4.23}$$

Although the usefulness of Eq. (4.23) is not immediately apparent, it is known that both $\pi_0(y) = 1$ and $\beta_0 = \int_{\mathbb{R}} \rho(y)\,dy$ [48]. This allows for the computation of α_0 and β_0, which then set off the recursive relationship of Eq. (4.20), as π_{p+1} is a function of α_p, β_p and the previously computed π_p. The formulas of Eq. (4.23) were independtly developed by both Darboux and Stieltjes, from which the name of this method originates [48, 27, 105]

The Stieltjes procedure, although reasonably computable, is not numerically stable. Any rounding errors in the moment calculations needed for Eq. (4.23) will quickly propagate and grow throughout the recursive process. Considering that any SR computations will be done on a computer with finite precision, left unmodified, this method is not tenable.

4.3.2.2 Chebyshev Algorithm

As an alternative to the Stieltjes procedure, the Chebyshev algorithm is also capable of solving for the coefficients α_p and β_p [48]. This is done by first computing

$$\sigma_{p,l} = \int_{\mathbb{R}} \pi_p(y) y^l \rho(y)\,dy, \tag{4.24}$$

and using these terms to then solve for the coefficients

$$\begin{aligned}\alpha_0 &= \frac{\sigma_{0,1}}{\sigma_{0,0}}, \quad \beta_0 = \sigma_{0,0}, \\ \alpha_p &= \frac{\sigma_{p,p+1}}{\sigma_{p,p}} - \frac{\sigma_{p-1,p}}{\sigma_{p-1,p-1}}, \\ \beta_p &= \frac{\sigma_{p,p}}{\sigma_{p-1,p-1}}.\end{aligned} \tag{4.25}$$

A property of the terms $\sigma_{p,l}$ is that they may be generated recursively from the moments μ_l via

$$\begin{aligned}&\sigma_{p,l} = \sigma_{p-1,l+1} - \alpha_{p-1}\sigma_{p-1,l} - \beta_{p-1}\sigma_{p-2,l}, \quad l = p, p+1, \ldots, 2*(P+1) - p - 1, \\ &\sigma{-1,0} = 0, \quad \sigma_{0,l} = \mu_l.\end{aligned} \tag{4.26}$$

Therefore, Eq. (4.22) is used to compute the first $2(P+1)$ moments, such that the first $(P+1)$ coefficients of $\alpha_0, \ldots, \alpha_P$ and $\beta_0, \ldots, \beta_P$ are evaulated via Eqs.(4.26) and (4.25). However, the mapping from moments to coefficients is once again ill-conditioned, and any rounding errors will quickly expand within the recursive algorithm. Therefore, a modification is desired.

4.3.2.3 Modified Chebyshev Algorithm

In order to expand upon the Chebyshev algorithm, *modified moments* are proposed [48]. That is, consider orthogonal polynomials separate from those we are attempting to generate which still obey a recursion formula

$$\begin{aligned}\Pi_{p+1}(y) &= (y - a_p)\Pi_p(y) - b_p\Pi_{p-1}(y), \quad p = 0, 1, \ldots, \\ \Pi_0(y) &= 1, \quad \Pi_{-1}(y) = 0.\end{aligned} \tag{4.27}$$

For our polynomials Π_p, we assume that a_p and b_p are known. Therefore, this is a familiar orthogonal polynomial such as Legendre or Hermite. We then have the *modified moment* equation of

$$\nu_r = \int_{\mathbb{R}} \Pi_r(y)\rho(y)\, dy, \quad r = 0, 1, 2, \ldots. \tag{4.28}$$

The modification continues by inserting the known polynomials into Eq. (4.26), i.e.,

$$\sigma_{p,l} = \int_{\mathbb{R}} \pi_p(y)\Pi_i(y)\rho(y)\, dy. \tag{4.29}$$

Thusly, this alteration to the $\sigma_{p,l}$ equation also changes the calculations in Eq. (4.25). The initialization of the algorithm begins with,

$$
\begin{aligned}
\sigma_{-1,l} &= 0, \quad l = 1, 2, \ldots, 2(P+1) - 2, \\
\sigma_{0,l} &= \nu_l \quad l = 0, 1, \ldots, 2(P+1) - 1, \\
\alpha_0 &= a_0 + \frac{\nu_1}{\nu_0}, \\
\beta_0 &= \nu_0.
\end{aligned}
\tag{4.30}
$$

Using the values of Eq. (4.30), we are then able to construct $\sigma_{p,l}$ and, as a result, α_k and β_k. That is,

$$
\begin{aligned}
\sigma_{p,l} &= \sigma_{p-1,l+1} - (\alpha_{p-1} - a_l)\sigma_{p-1,l} \\
&\qquad - \beta_{p-1}\sigma_{p-2,l} + b_l \sigma_{p-1,l-1}, \quad l = p, p+1, \ldots, 2*(P+1) - p - 1, \\
\alpha_p &= a_p + \frac{\sigma_{p,p+1}}{\sigma_{p,p}} - \frac{\sigma_{p-1,-}}{\sigma_{p-1,p-1}}, \\
\beta_p &= \frac{\sigma_{p,p}}{\sigma_{p-1,p-1}}.
\end{aligned}
\tag{4.31}
$$

As seen in the results of [48], the modified Chebyshev algorithm maintains a stability that surpasses both the unmodified Chebyshev and the Stieltjes procedure.

4.3.2.4 Discretized Algorithms

Taking into account the transition from the measure $d\lambda(y)$ to the weight function $\rho(y)dy$ (as seen in Eqs. (4.21) and (4.22)), we now consider the weight function $\rho(y)$ to exist on $(-1, 1)$. Therefore, it is possible to approximate the newly bounded integrals, e.g., Eq. (4.22), using a quadrature rule, i.e.,

$$
\int_{-1}^{1} \phi(y)\rho(y)\, dy = \sum_{n=1}^{N} w_n \phi_n(y_n)\rho(y_n) + R_N(\phi\rho), \quad N > (P+1),
\tag{4.32}
$$

where $\phi(y)$ is a generic function and N total quadrature points or nodes are taken.

By using this rule in place of continuous integrals throughout the recursive construction of the polynomials and their coefficients, we achieve a numerically stable algorithm [48]. Indeed, if

$\rho \in \mathbb{P}_s$, then $R_N(\phi\rho) = 0$ for all $\phi \in \mathbb{P}_{2(P+1)-1}$ as long as $N \geq 2(P+1) + s$. Here, the term s is dependent on the type of quadrature used, e.g., classical Gaussian $s = 1$. It is noted here that the use of Eq. (4.32) requires that the input space have finite bounds. This choice of finite bounds is previously referred to in Section 4.3.1.

4.3.3 Algorithm

For the purposes of implementation, an ME-SR algorithm summary is provided. It is noted that the following pseudocode serves primarily as an overall review of the ME-SR process, and it does not contain nuances such as ALS convergence criteria that are covered in [17, 20, 36]. It is emphasized here that this algorithm relies on convergence in the K-sense. Therefore, quantities such as element validation error are not considered as stopping criteria, and the outer loop of the algorithm relies on the statement **while any** $\eta_k^\gamma \Pr(I_{B_k} = 1) \geq \theta_1$. In other words, the ME-SR construction process continues until all elements satisfy the limit set by the decay check parameter θ_1.

```
Create single element SR surrogate;
Check for need to split Eq. (4.12);
while any $\eta_k^{\gamma}\,\mathrm{Pr}(I_{B_k} = 1) \geq \theta_1$ do
    $K_{new} = 0$;
    for $k = 1$ to $K$ do
        if $\eta_k^{\gamma}\,\mathbf{Pr}(I_{B_k} = 1) \geq \theta_1$ then
            Compute directional split sensitivities using Eq. (4.13) and Eq. (4.15);
            Determine the $d'$ directions which satisfy Eq. (4.16);
            Create element "children" following the example of Eq. (4.17) and binomial
              coefficient logic;
            Using Eq. (4.18), $K_{new}$ += $\Delta K$;
        end
    end
    Destroy all elements satisfying Eq. (4.12);
    Add element children to input space decomposition $D$;
    $K$ += $K_{new}$;
    for $k = 1$ to $K$ do
        if $\hat{q}_k(\boldsymbol{y})$ is new then
            Convert $\boldsymbol{y}$ to $\bar{\boldsymbol{y}}$;
            for $i = 1$ to $d$ do
                Construct polynomials orthogonal to $\rho_{k,i}$ using algorithm of choice (such as
                  Py-Orthpol);
            end
            Perform ALS to construct $\hat{q}_k(\boldsymbol{y})$;
            Compute $\eta_k$ using Eq. (4.11);
        end
    end
end
```

Algorithm 2: Algorithm for ME-SR

In Algorithm 2, the introduced variable K_{new} is used to keep track of the total number of elements to be added. The exact number of elements required to achieve a solution converged to a desired accuracy is not known a priori. Therefore, it is up to the operator's choice of θ_1 to dictate how flexible or inflexible the adaptive algorithm is.

Chapter 5

Low Earth Orbit Applications

As discussed in the Introduction, there is motivation to provide UQ techniques in low Earth orbit (LEO). Updating the state of an object in orbit via observations requires knowledge of where that object could possibly be. The results of [70] show that it is possible for Gaussian assumptions to become inaccurate. Additionally, the risk of a collision between a spacecraft and another object is unacceptable to an operator, and steps must be taken to mitigate or identify any threat to a mission. In order to be able to perform a conjunction assessment, the operational procedure must first have a method of quantifying the uncertainty of the system. Therefore, we present a series of test cases in LEO, which we have applied SR for the purposes of UQ. The framework on which these test cases are built upon serves as a blueprint for later applications found within this thesis, such as the probability of collision work in Chapter 6.

5.1 Analysis Methodology

In order to visualize the distributions of a test case and characterize the accuracy of an SR-based PDF, MC runs of $100,000$ samples are created for all examples within this section. In addition to this, $100,000$ evaluations of an SR surrogate are created using independent sets of samples of the inputs and the appropriate uncertainties. By computing these data sets side-by-side, this approach is used in all test cases to produce figures on which the PDFs are qualitatively compared. Thus, the efficacy of SR to estimate the PDF of an object in LEO is analyzed. The histograms of the MC results are displayed as a solid line (derived from interpolating the centers of the bins) plotted over the histograms of the SR-generated samples. This method allows for qualitative assessments of SR's ability to capture the third and higher moments. Each test case is also examined quantitatively. These analyses consist of relative residuals calculated for the mean and STD of QoIs. The relative residual provides information for knowing the digits of accuracy in a solution. Previously defined in Chapter 3, we remind the reader that this quantity is calculated by

$$\epsilon_{rel} = \left| \frac{\hat{\lambda} - \lambda}{\lambda} \right|, \tag{5.1}$$

where ϵ_{rel} is the relative residual, λ is a reference value, and $\hat{\lambda}$ is an estimate of this reference. The output of Eq.(5.1) is used throughout the results of this work, and its overall format remains the same. Additionally, the accuracy of each surrogate is estimated by evaluating absolute residuals of a small set of random samples, which are not used in the training of the surrogate. This process computes the difference between the deterministic solution using the block-box propagator and the surrogate-based solution, with each method using the same random inputs. The RMS of the difference is taken over the number of validation samples and provided a quantitative measure for determining the *goodness* of the surrogate and solution convergence. Such a strategy is referred to as *cross-validation* (or *validation* for brevity) and may be extended to multiple constructions of the surrogate and residual evaluation on independent samples of the QoI, see, e.g., [44, Chapter 7]. Validation results are included for each test case in Tables 5.4, 5.5, 5.6 and 5.12, which include the RMS of the difference taken over the number of validation samples as well as the RMS of the MC-based samples for comparison.

5.2 Case Characteristics

When considering Sections 2.1 and 2.5.2, test cases for this section seek the estimation of the function $\boldsymbol{q}(\boldsymbol{y})$. As discussed in Section 2.1, realizations of this function are the position and velocity state of a spacecraft at a considered time. These states are propagated using an ODE integrator and are ultimately derived from an initial condition and Gaussian random variables as seen in Section 2.5.2.2. Therefore, the final states $\{\boldsymbol{q}(\boldsymbol{y}_j)\}$ are treated as training samples and, along with associated random variables $\{\boldsymbol{y}_j\}$, are used to estimate the coefficients $\{\boldsymbol{c}_i^l\}$ and $\{\boldsymbol{u}_0^l\}$. These coefficients are used to construct an SR surrogates from Eq.(2.40). This surrogate evaluates many more sets of samples of the inputs, and the resulting sets of QoIs are then used to create an approximate PDF. As we shall demonstrate, the QoI considered in the following test cases admit small separation ranks r, which range from $r = 3$ to $r = 6$ depending on the case, for the considered accuracies. Such small separation ranks lead to accurate estimation of statistics of the QoIs, its PDFs, and sensitivities with respect to the input variables, using relatively small numbers of samples of QoI. For scenarios where the separation rank is not small, one shall not anticipate similar accuracies as in these experiments.

Each test compares the performance of SR with that of a PC result that is converged with respect to STD. PC is chosen as a reference solution due to its proven nature of converging efficiently and accurately [68]. In the case of high stochastic dimension, PC is still used, albeit with high computation cost due to the *curse of dimensionality.* All test cases incorporate a 50×50 spherical harmonics model of the Earth's gravity perturbations, as determined by GRACE GGM02C gravity model [111], as well as atmospheric drag perturbations based on the exponential cannonball model presented in [113]. All cases are propagated for 36 hours using a Dormand-Prince 5(4) integrator with a tolerance of 10^{-13}.

The first test case considers STDs of 1 km and 1 m/sec in the initial position and velocity, respectively, generating a stochastic dimension of $d = 6$. The second case is similar to the first, with the exception that 14 random inputs (to bring the total to $d = 20$) are added in the form of uncertainty in the low degree Stokes coefficients, drag parameters and the gravitational parameter. The Stokes coefficients used as random inputs are included in Appendix A. Initial values and uncertainties are taken from [111]. Parameters for the *a priori* state PDF for Test Case 2 can be seen in Table 5.1. Test Case 1 uses the first six random inputs in the form of ECI coordinates and velocity. In both test cases, ECI position and velocity as a function of the random inputs are estimated using the SR surrogate for a vector-valued function.

For the third test, a problem presented in [62] is analyzed using SR in place of the original examination with a GMM. In this case, ECI coordinates are replaced with equinoctial elements. This coordinate system is composed of the semimajor axis and five additional elements. The equations for these additional elements as a function of the Keplerian elements are included in Appendix B. Each sample in equinoctial elements is transformed into ECI coordinates and propagated for 36 hrs. The propagated ECI state is then transformed back to equinoctial elements to be used as a training sample. In this case, all six equinoctial elements are estimated using the SR estimation process for a vector-valued function. All random inputs and their relevant uncertainties are presented in Table 5.2.

Table 5.1: Random inputs and associated STDs for the first two test cases

	Mean	STD
x (km)	757.700	1.0
y (km)	5222.607	1.0
z (km)	4851.800	1.0
$\dot{x}$ (m/s)	2213.210	1.0
$\dot{y}$ (m/s)	4678.340	1.0
$\dot{z}$ (m/s)	-5371.300	1.0
μ (km^3/s^2)	3.986×10^5	10^{-3}
C_D	2.0	0.398
A/m (m^2/kg)	0.01	1.7×10^{-3}

Table 5.2: Random inputs and associated STDs for Test Case 3

	Mean	STD
a (km)	6980.0	20.0
h	0.0	10^{-3}
k_e	0.0	10^{-3}
p_e	0.0	10^{-3}
q_e	0.0	10^{-3}
$\lambda_{\mathcal{M}}$ (rad)	0.0	$10^{-2}\frac{\pi}{180}$

In all cases, canonical units are utilized for estimation and plotting. Therefore, all units of distance have been normalized with respect to the Earth's mean radius ($r_\oplus = 6371$ km), resulting in distance units DU, while seconds have been normalized with the time unit TU, where

$$\mathrm{TU} = \sqrt{r_{\oplus}^{3}/\mu}, \tag{5.2}$$

and μ is the gravitational parameter. The result of using canonical units is a data set that is mostly of the same magnitude with respect to quantities of interest and therefore more numerically stable.

5.3 Test Case 1

For an initial example, a case of a satellite in low Earth orbit is examined. The initial conditions for position and velocity in Table 5.1 are used as the initial mean solution, with the respective uncertainties being applied with the methodology described in Section 2.5.2.2. Figure (5.1) depicts the distribution of Radial-Intrack-Crosstrack (RIC) coordinates after 36 hours when considering 100,000 MC samples.All RIC plots (MC- and SR-based) are generated using the same MC propagated position and velocity sample, pulled from the training data set, as the center of the coordinate frame. This sample is unique for each test case, and the particular value was randomly chosen as it puts the origin of the frame at a possible realization of the posterior PDF. As depicted, the *a posteriori* position distribution is non-Gaussian. This is most evident in the Radial-Intrack plot, where a large amount of skewness is evident. Hence, any Gaussian assumptions on the posterior are invalid when attempting to accurately estimate the PDF. Therefore, this case requires a higher-fidelity method.

Using 350 training samples, $r = 5$, $P = 4$ and $\delta = 10^{-7}$, an SR surrogate is created for approximating the ECI state of the satellite. With (2.60) and (2.61), means and STDs for each random input of the satellite's state are compared to those of a converged PC solution. The results can be seen in Table 5.3. Since the reference solution is converged with respect to STD, relative residuals for the third and fourth moments are not provided. Table 5.4 includes the validation results for Test Case 1. The table presents RMS values of the absolute residuals for 70 independent samples of random inputs, as well as the RMS values of the MC-based validation samples. By

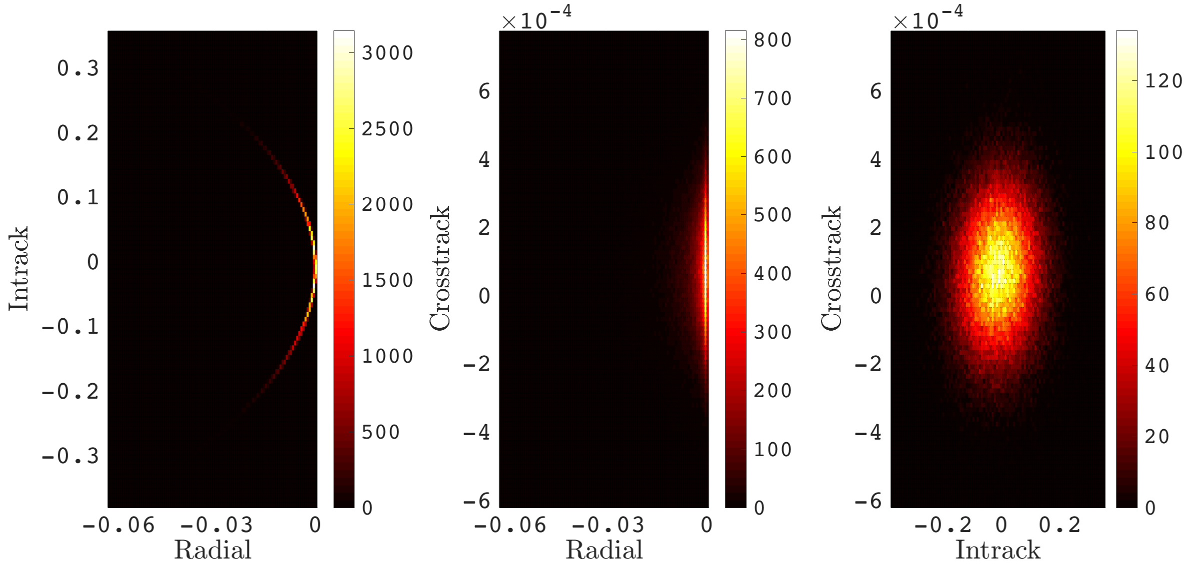


Figure 5.1: MC results for Test Case 1 plotted in RIC coordinates. Note that the colorbars illustrate object count.

comparing the RMS of the residuals with the RMS of the MC-based validation samples, we observe the accuracy of the SR in predicting the states at samples of inputs not included in the training samples.

Qualitatively, it can be seen in Figure 5.2 that the SR solution captures the non-Gaussian distributions well. In the case of $\dot{x}$, the distribution is highly non-Gaussian. The histogram of the SR solution follows the MC distribution well, capturing the overall skewness and kurtosis. The results of the SR evaluation are also transformed to the RIC frame and plotted in Figure 5.3. In addition to these qualitative fits, three digits of precision or more are captured in the first and second moments when comparing the converged PC solution to that of SR. Figure 5.4 depicts the change in STD (computed via (2.61)) as a function of N, which may be used as one metric to illustrate solution convergence. With increasing values of N, change in the STD is quantified via a relative difference, which provides one convergence metric without the use of validation samples. The decreasing difference value indicates convergence of the SR surrogate on its solution for the QoIs. The relative difference is computed via non-consecutive STD approximations to avoid local minima. It is important to note that the plot illustrates solution precision, while results from validation methods (e.g., see Table 5.4) provide information on the surrogate's accuracy. Figure 4 demonstrates the reduced change in the solution with larger N. For this case, the $N = 350$ solution retains a precision of three digits or more, and STD estimates do not vary significantly with more training samples. The convergence tolerance employed to identify a sufficiently large N varies with the application.

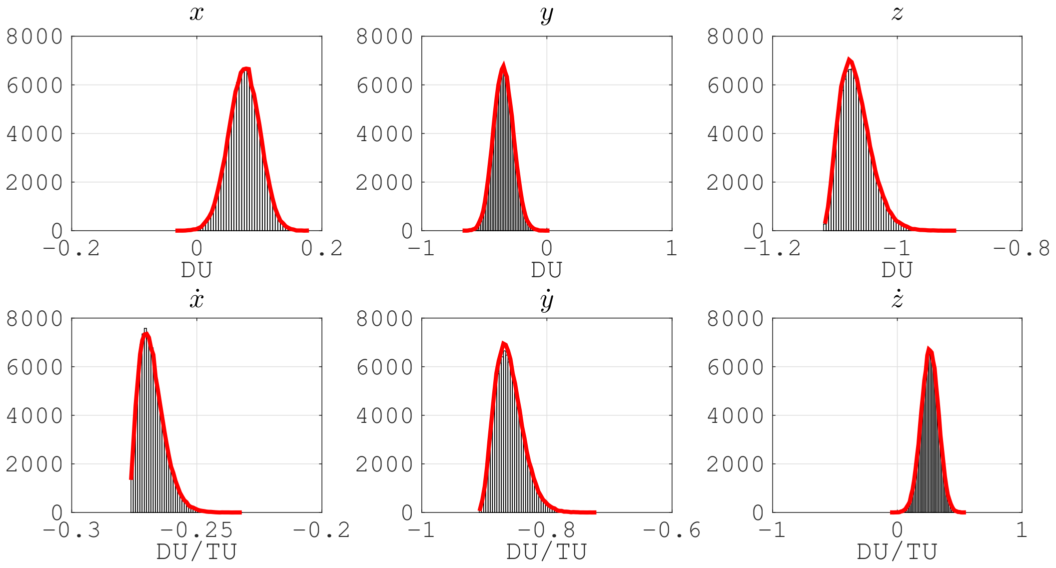


Figure 5.2: Histograms of SR results for quantities of interest in Test Case 1.

Table 5.3: Agreement between SR- and PC-based mean and STD for Test Case 1

	Ref. Mean	Ref. STD	Rel. Mean	Rel. STD
x	0.075765 (DU)	0.02498 (DU)	9.8e-05	2.7e-04
y	-0.34683 (DU)	0.07986 (DU)	6.5e-05	1.8e-04
z	-1.06680 (DU)	0.02534 (DU)	3.4e-06	1.5e-04
$\dot{x}$	-0.26786 (DU/TU)	5.13e-03 (DU/TU)	2.0e-05	1.0e-03
$\dot{y}$	-0.859985 (DU/TU)	0.02238 (DU/TU)	4.3e-06	1.8e-04
$\dot{z}$	0.2603 (DU/TU)	0.06937 (DU/TU)	1.3e-04	3.0e-04

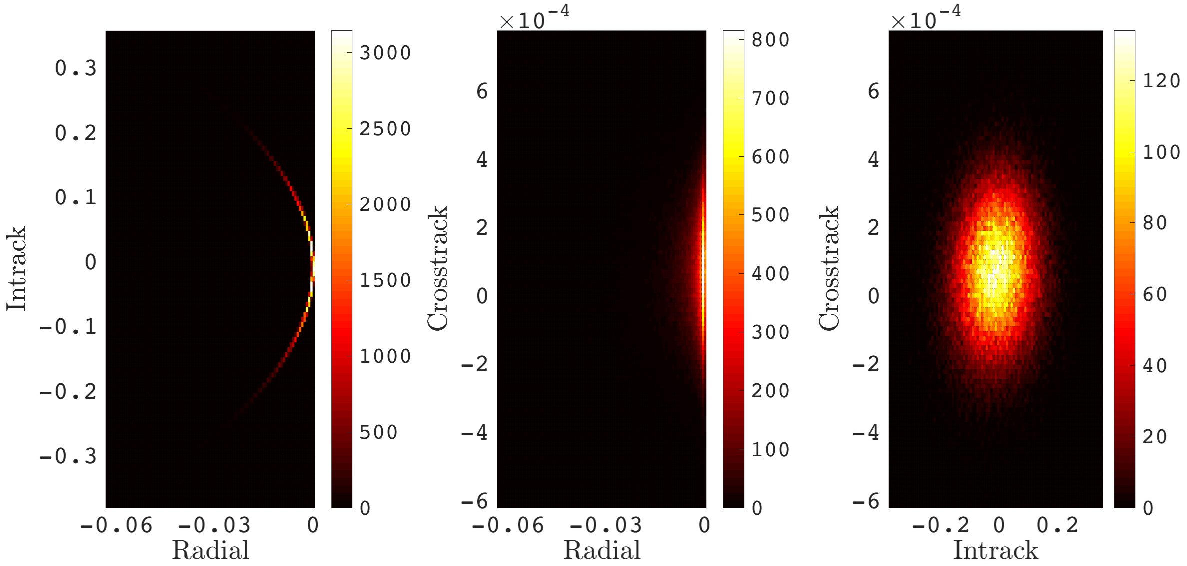


Figure 5.3: SR results for Test Case 1 plotted in the RIC coordinates. Note that the colorbars illustrate object count.

Table 5.4: Residual RMS of 70 SR- and MC-based validation samples for Test Case 1

	MC Sample RMS	Residual RMS	Units
x	0.07679	1.0e-04	(DU)
y	0.3705	3.9e-04	(DU)
z	1.0620	2.9e-05	(DU)
$\dot{x}$	0.2685	1.2e-04	(DU/TU)
$\dot{y}$	0.85600	4.7e-05	(DU/TU)
$\dot{z}$	0.2823	3.0e-04	(DU/TU)

5.4 Test Case 2

For Test Case 2, the previous initial conditions are kept but the stochastic dimension is expanded to include all elements from Table 5.1, in addition to the Stokes coefficients from Table A.1. Theoretically, an estimation method that suffers from the *curse of dimensionality* would require significantly more samples than a method, such as SR, that increases the number of samples linearly with respect to stochastic dimension. This cost comparison is elaborated upon in Section 2.5.2.4 and its associated Remark 5.

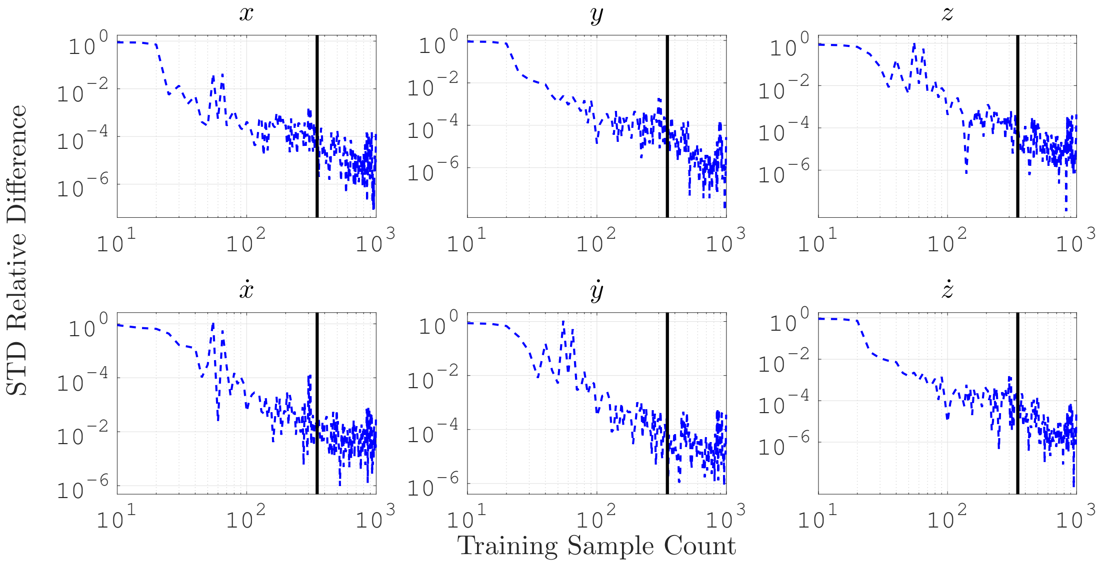


Figure 5.4: Plots of relative differences between approximated STD values for Test Case 1. Note the black line highlighting the choice of N for the surrogate.

Using $N = 750$ training samples, $P = 4$ and $\delta = 10^{-7}$, an SR surrogate is created for estimating the ECI state of the satellite. In this case, results are included for surrogates of rank 3, 4 and 5 with each using the same 750 training samples. By presenting results for these three choices of r, a clearer picture of the convergence of an SR surrogate can be seen. Figure 5.5 presents the 100,000 MC realizations in the RIC frame, which resembles that of Figure 5.1. This figure can then be compared to the results presented in Figures 5.6, 5.7, and 5.8. Each of these figures illustrates realizations generated by the SR surrogates of ranks 3, 4 and 5. Convergence to the MC distributions can been seen as the ranks increase, with $r = 5$ providing a good qualitative fit.

Table 5.5 includes the quantitative validation results for Test Case 2. The table presents RMS values of the residuals for 150 independent random input vectors as evaluated by either a rank 3, 4 or 5 surrogate, as well as RMS values of the MC-based validation samples. Although no improvement in accuracy may be readily apparent in the ECI frame, when the error is transformed to the RIC frame as displayed in Table 5.6, the estimate of the crosstrack state improves with each additional rank. By comparing the validation results of the MC-based RMS with the residual

RMS, the sample RMS in the crosstrack direction indicates a smaller mean compared to the radial and intrack directions. In this case, the $r = 5$ case is the only solution to yield a residual RMS smaller than the sample RMS. As illustrated in Figures 5.7 and 5.8, the crosstrack QoI lies within bounds on the order of 10^{-4}. As seen in Table 5.6, the crosstrack error magnitude for ranks 3 and 4 indicates that one digit of accuracy has not been achieved. The improvement of adding a fifth rank, however, is enough for the crosstrack accuracy of the rank 5 solution to be on the order of 10^{-5} and therefore accurate enough to capture the crosstrack distribution.

As this discussed investigation utilizes the *a priori* knowledge found in Figure 5.5, an operational implementation concerned with crosstrack accuracy would require a different approach. Table 5.7 contains STD values of Test Case 2 approximated by the previously discussed SR solutions and transformed into the RIC frame. By comparing these values to those of Table 5.6, it can be seen that the STDs of the crosstrack direction for the rank 3 and 4 surrogates are smaller than each correspoding residual RMS by an order of magnitude or more. The rank 5 approximation of crosstrack STD, however, is larger than the respective residual RMS. Although this is not proof of solution convergence to STD, it does explain the inability to capture information from the PDFs using the rank 3 and 4 solutions. With a residual RMS larger than the STD, the approximation of the crosstrack direction variance is largely due to the accuracy of the solution and not the actual PDF. It should be noted that although the approximate value of STD for the crosstrack direction, as seen in Table 5.7, increases by an order of magnitude with each increase in rank, a solution of rank 6 yields a crosstrack STD of 1.6e-04 DU. This value compares well with the rank 5 approximation.

Figure 5.9 illustrates the precision of STD estimates in Test Case 2 as a function of N. The methodology for the generation of the figure is the same as the previous case, but with the covariance matrix rotated into the RIC frame to match other results for Test Case 2. This figure also exhibits the previously noted slow convergence in the crosstrack direction when compared to radial and intrack. The N used for the calculation of the presented results retains a precision of three digits or more for the radial and intrack STD. The crosstrack quantity has a precision of

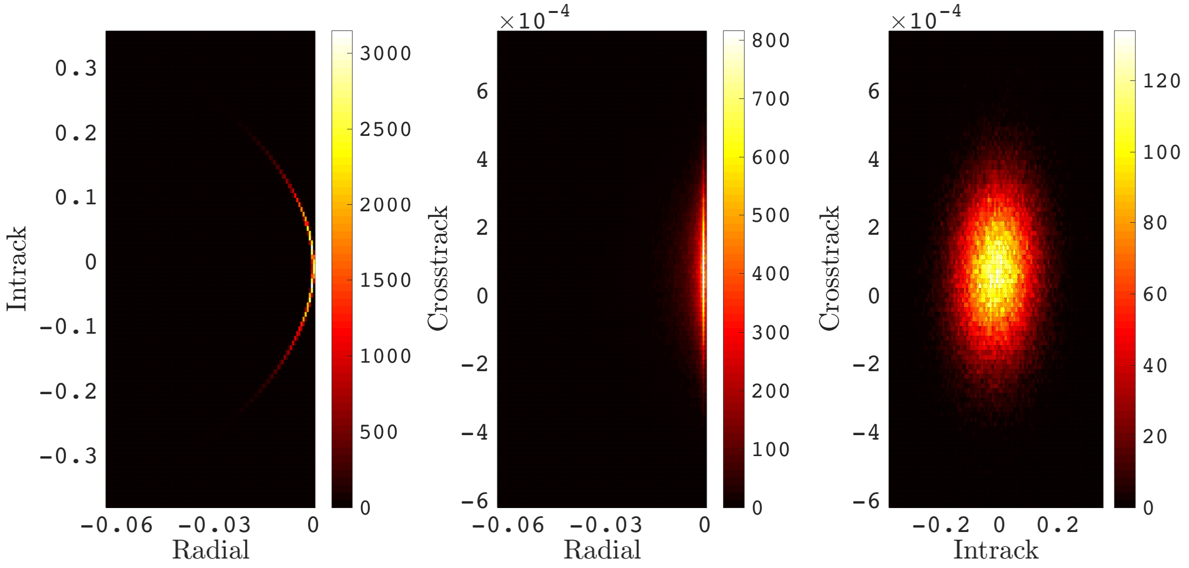


Figure 5.5: MC results for Test Case 2 plotted in RIC coordinates. Note that the colorbars illustrate object count.

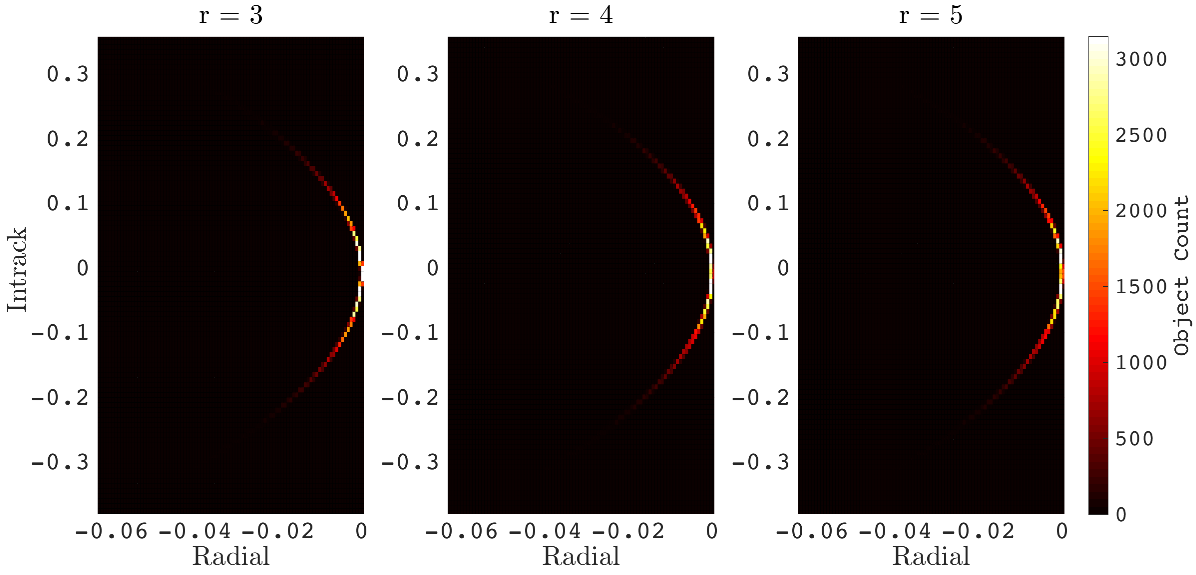


Figure 5.6: SR results for Test Case 2 plotted as radial and intrack.

one digit that continues to improve with N. When considering the implications of this figure, it should be noted that the STD value of the crosstrack direction is smaller than the radial and intrack values by an order of magnitude or more (see Table 5.6). Comparing the surrogate computed for the

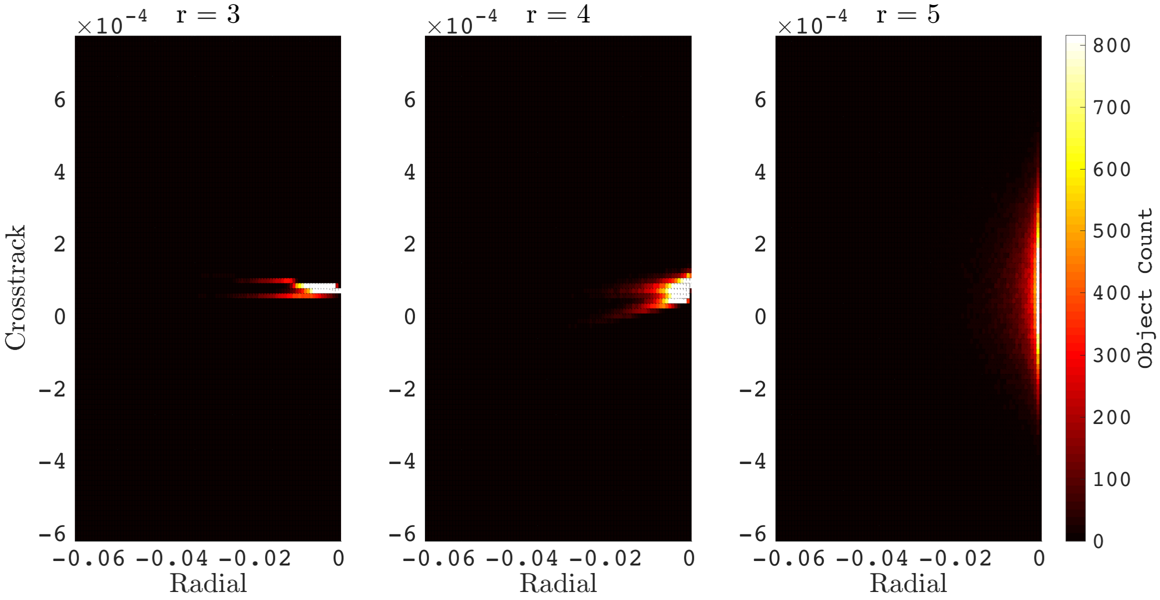


Figure 5.7: SR results for Test Case 2 plotted as radial and crosstrack.

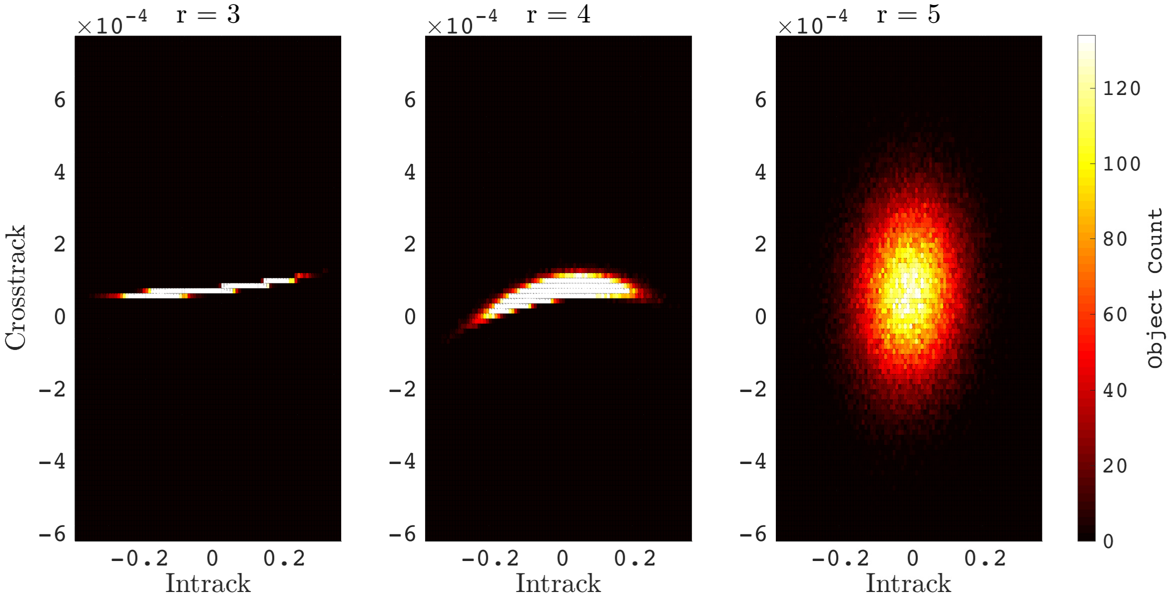


Figure 5.8: SR results for Test Case 2 plotted as intrack and crosstrack.

results of the presented case ($N = 750$) and a surrogate computed in the previous non-consecutive step ($N = 740$), the absolute difference for each estimated RIC STD is on the order of 10^{-6} DU. Hence, relative and absolute differences may be considered to present a more complete picture of

Table 5.5: Residual RMS of 150 SR- and MC-based validation samples for Test Case 2

		Residual RMS			
	MC Sample RMS	$r = 3$	$r = 4$	$r = 5$	Units
x	0.0790	1.7e-04	1.6e-04	1.0e-04	(DU)
y	0.3594	3.8e-04	2.7e-04	2.7e-04	(DU)
z	1.0658	1.1e-04	1.9e-04	1.7e-04	(DU)
$\dot{x}$	0.2680	1.2e-04	1.3e-04	1.1e-04	(DU/TU)
$\dot{y}$	0.8592	1.3e-04	1.6e-04	1.4e-04	(DU/TU)
$\dot{z}$	0.2724	3.2e-04	2.2e-04	2.2e-04	(DU/TU)

Table 5.6: Residual RMS of 150 SR- and MC-based validation samples in RIC frame for Test Case 2

		Residual RMS			
	MC Sample RMS	$r = 3$	$r = 4$	$r = 5$	Units
Radial	1.1239	1.9e-04	2.5e-04	2.3e-04	(DU)
Intrack	0.0896	3.6e-04	2.3e-04	2.2e-04	(DU)
Crosstrack	1.8e-04	1.3e-04	1.3e-04	6.5e-05	(DU)

Table 5.7: STD estimates in RIC frame for Test Case 2

	STD			
	$r = 3$	$r = 4$	$r = 5$	Units
Radial	4.8e-03	4.8e-03	4.8e-03	(DU)
Intrack	8.7e-02	8.7e-02	8.7e-02	(DU)
Crosstrack	8.3e-06	2.7e-05	1.5e-04	(DU)

solution convergence as a function of N.

As $r = 5$ provides the best fit, further results are calculated with a surrogate corresponding to that rank. Figure 5.10 shows the SR result as a histogram plotted alongside a 100,000 MC result. Once again, it can be seen that the SR solution captures the non-Gaussian distributions well. The skewness and tails of the non-Gaussian distributions are represented in both the MC and SR PDFs. In addition to these qualitative fits, Table 5.8 shows that three digits of precision or more are captured in the first and second moments, when compared to the PC result that has

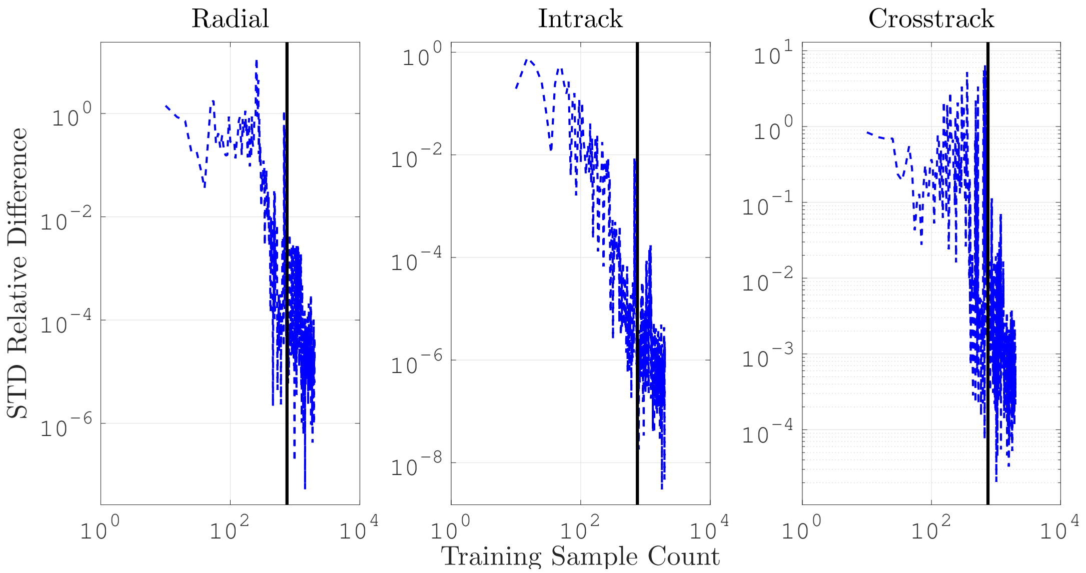


Figure 5.9: Plots of relative differences between approximated STD values for Test Case 2. Note the black line highlighting the choice of N for the surrogate.

been converged with respect to STD.

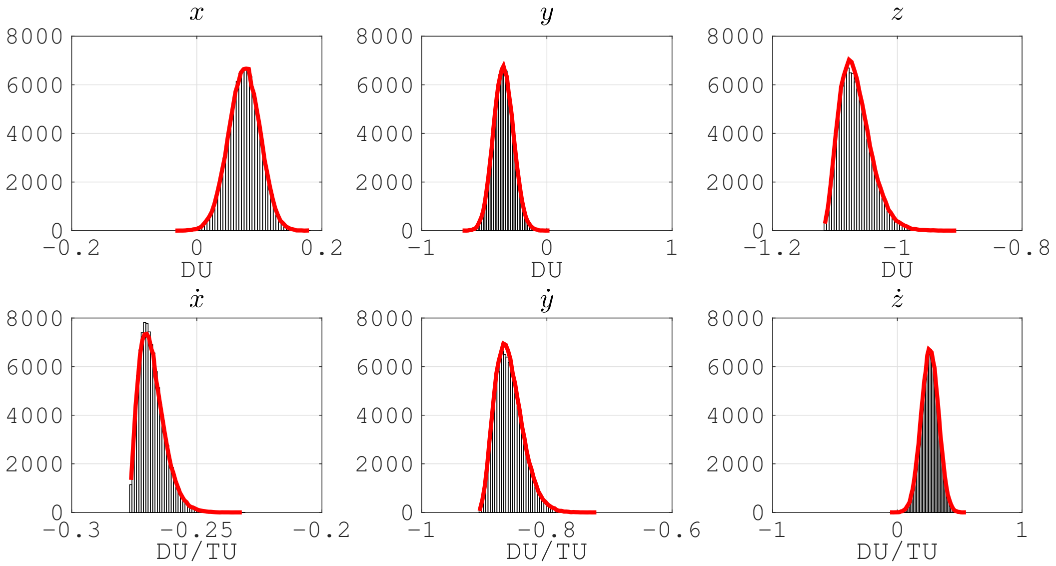


Figure 5.10: Histograms of SR results for quantities of interest in Test Case 2.

Figure 5.11 illustrates the convergence of SR as a function of N and compares it to that of MC. The figure shows the relative errors of the estimated STD for the x-position of 100 independent SR solutions. For each chosen N, 100 independent calculations of an SR surrogate and subsequent

Table 5.8: Agreement between SR- and PC-based mean and STD for Test Case 2

	Ref. Mean	Ref. STD	Rel. Mean	Rel. STD
x	0.075765 (DU)	0.024985 (DU)	1.4e-05	4.0e-05
y	-0.346831 (DU)	0.079863 (DU)	4.6e-06	3.0e-05
z	-1.066805 (DU)	0.025340 (DU)	2.9e-07	3.2e-05
$\dot{x}$	-0.26786 (DU/TU)	5.135e-03 (DU/TU)	1.8e-05	5.8e-04
$\dot{y}$	-0.859985 (DU/TU)	0.02238 (DU/TU)	1.5e-06	3.2e-04
$\dot{z}$	0.260341 (DU/TU)	0.069376 (DU/TU)	8.8e-04	3.6e-05

STD estimation are performed. This is done using the Test Case 2 initial conditions with fixed SR parameters excluding the number of training samples. The ordinate axes of the plots are logarithmic, the middle lines of the boxes are the medians, and the top edge of the boxes are the 75th percentiles (upper quartile) with the bottom edges being the 25th (lower quartile). The upper and lower whiskers cover 1.5 times the interquartile range for the upper and lower quartiles, respectively, which is 99.3% of the data if it was normally distributed. The remaining outliers are marked as blue dots. Here we use (2.61) to approximate σ for the SR solutions and sampling based method for the MC results. The fast convergence of SR estimates of STD (as a function of N) can be seen in the left plot of Fig. 5.11, and, when compared to the right plot, the relatively slow convergence of MC can also be seen.

5.4.1 Sensitivity Analysis

Using surrogates to evaluate Eq.(2.7), a sensitivity analysis is applied to the results of Test Case 2. Reference values found using a PC expansion are included in Tables 5.9 and 5.10 along with absolute residuals found when approximating the indices with 10^6 SR realizations. The PC results, which are generated such that they provide six digits of precision, utilize an analytic method for calculating sensitivity indices [21], and serve as a baseline for assessing SR accuracy. Therefore, some values are stated as ~ 0, as the calculated indices are less than or equal to 10^{-6}. Each of these values represents, at most, 10^{-8} percent of the total variance contribution to a QoI due to the fact that total Sobol indices sum to one or greater [102]. The indices in

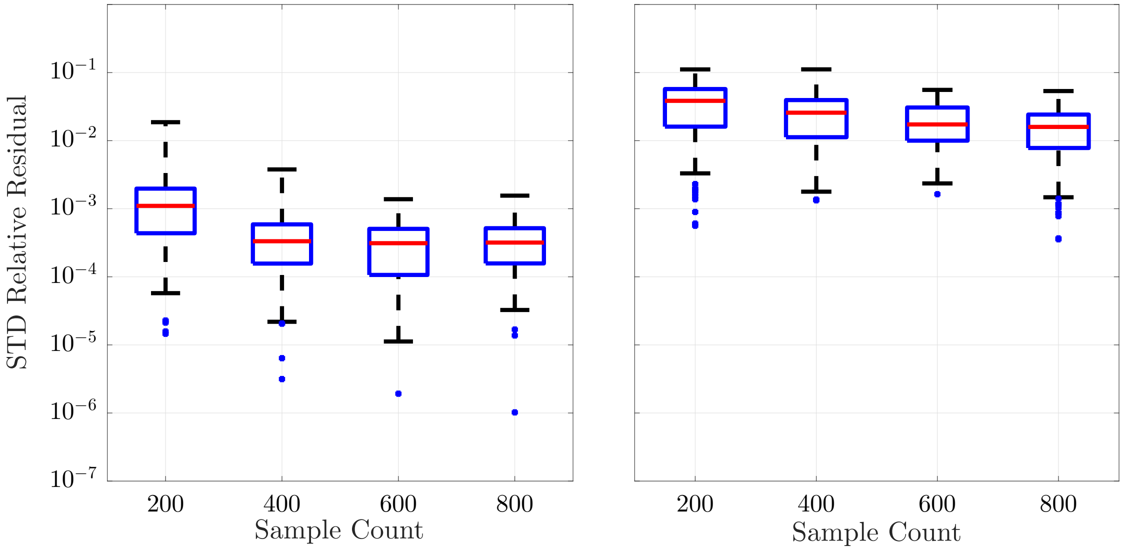


Figure 5.11: STD relative residuals. SR on left and MC on right

Tables 5.9 and 5.10 show that the ECI random inputs contribute variability that is large enough to be quantified, with the x-position having the smallest contribution out of the six. In addition to these six random dimensions, the sensitivity index of μ is included. Using the indices as a guide, it can be concluded that uncertainty in μ produces little variability in the final solution for this test case. Figure 5.12 shows the absolute value of univariate functions $\{u_i^l\}$ for the random inputs provided in Tables 5.9 and 5.10. Constructed using $\{\boldsymbol{c}_i^l\}$, the appropriate polynomial bases and a set of samples of the inputs, the figure illustrates the variability of each element. The behavior of each u_i^l is represented well by respective sensitivity indices. For example, the univariate functions of x and μ exhibit less variability than the other shown random inputs. The results in Tables 5.9 and 5.10 quantitatively reflect this, with x and μ having low index values. In Fig. 5.12, the values of $\{u_i^l\}$ for the other 13 random inputs are omitted due to the lack of variability. The sensitivity indices for the Stokes coefficients, C_D and A/M are all negligible (smaller than 10^{-6}) and are omitted for brevity. The low sensitivity index values for these random inputs, including μ, agree with intuition due to the physics of the high-altitude orbit. Gravitational perturbations as well as atmospheric drag effects are not significant in relation to position uncertainty at an altitude of around 790 km.

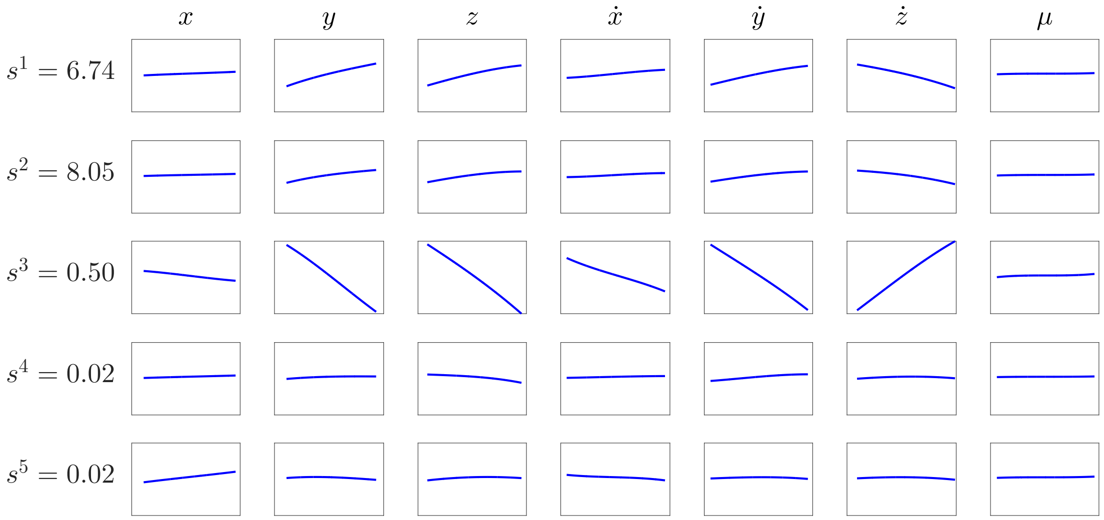


Figure 5.12: Plot of the absolute values of the univariate factors $u_i^l(y_i)$ for Test Case 2.

Table 5.9: Sensitivity indices $S_{i,m}$ and residuals of position QoIs for Test Case 2

Random	Quantities of Interest					
	x		y		z	
Inputs	PCE	Resid.	PCE	Resid.	PCE	Resid.
x	5.28e-03	8e-05	6.10e-03	3e-05	6.43e-03	3e-05
y	0.280	9e-04	0.279	1e-03	0.285	1e-05
z	0.235	1e-03	0.236	1e-03	0.238	8e-04
$\dot{x}$	0.0437	4-e04	0.0423	1-e04	0.0440	2-e04
$\dot{y}$	0.192	1-e04	0.193	2-e04	0.201	6-e04
$\dot{z}$	0.244	2-e04	0.244	8e-03	0.251	1e-05
μ	~ 0	N/A	~ 0	N/A	~ 0	N/A

5.5 Test Case 3

For Test Case 3, a scenario presented by [62] is considered. For this, the initial conditions, random inputs and relevant standard deviations are found in Table 5.2. The distribution of 100,000 MC samples in RIC coordinates and the non-Gaussian distributions can be seen in Fig. 5.13.

Table 5.10: Sensitivity indices $S_{i,m}$ and residuals of velocity QoIs for Test Case 2

Random	Quantities of Interest					
	$\hat{x}$		$\hat{y}$		$\hat{z}$	
Inputs	PCE	Resid.	PCE	Resid.	PCE	Resid.
x	4.30e-03	6e-05	5.93e-03	1e-05	5.99e-03	3e-05
y	0.291	2e-04	0.284	1e-03	0.277	2e-04
z	0.239	1e-04	0.244	1e-03	0.236	7e-04
$\dot{x}$	0.0386	7e-03	0.0428	6e-04	0.0426	1e-04
$\dot{y}$	0.207	5e-03	0.194	6e-04	0.192	4e-04
$\dot{z}$	0.257	5e-03	0.251	2e-04	0.246	1e-03
μ	~ 0	N/A	~ 0	N/A	~ 0	N/A

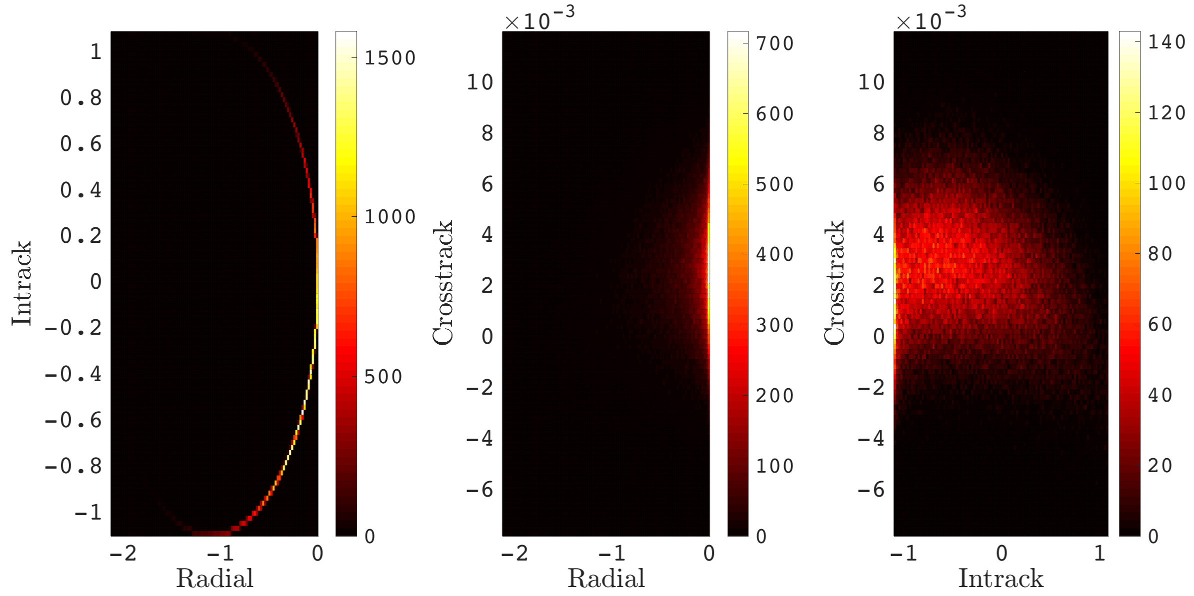


Figure 5.13: MC results for Test Case 3 plotted in the RIC coordinates. Note that the colorbars illustrate object count.

Using $N = 200$ training samples, $r = 6$, $P = 4$ and $\delta = 10^{-6}$, an SR surrogate is estimated for the six equinoctial elements. Figure 5.14 presents the 100,000 SR realizations in the RIC frame, which compares well to Figure 5.13. The results of the moment analysis, and therefore the quantitative fit, can be seen in Table 5.11. Three digits or more of accuracy is shown by the relative residuals for mean and standard deviation. Figure 5.15 illustrates the change in STD for Test Case 3 as a function of N. Generated in the same manner as the previous cases, the figure depicts changes in the estimated STD values for the six equinoctial elements. For this test case, the $N = 200$ value used for the SR solution yields a precision of three digits or more. Table 5.12 includes the validation results for Test Case 3. The table presents RMS values of the relative residuals for 40 independent random input vectors, as well as the RMS values of the MC-based validation samples for comparison. In addition to this, Figure 5.16 shows the SR result as a histogram plotted alongside a 100,000 MC result. Qualitatively, it can be seen that the SR solution captures the distributions well.

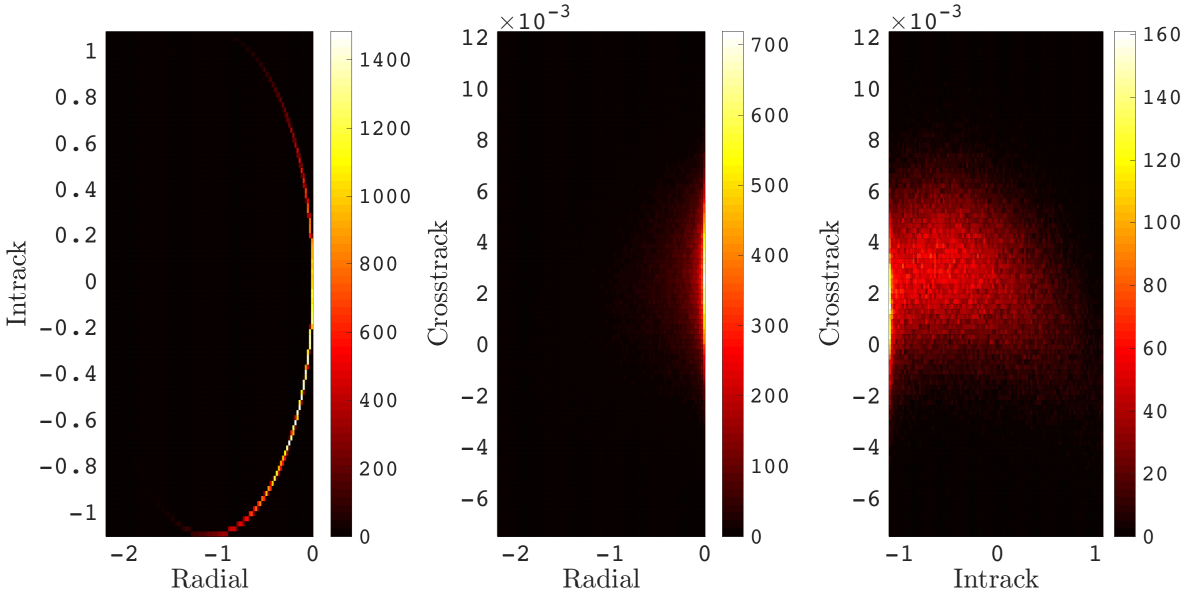


Figure 5.14: SR results for Test Case 3 plotted in the RIC coordinates. Note that the colorbars illustrate object count.

In [62], it is stated that uncertainty in the semimajor axis a has the largest effect on the uncertainty of the final PDF. At a qualitative glance, Figure 5.17 appears to agree with this conclusion. The figure is derived similarly to that of Fig. 5.12, with the appropriate $\boldsymbol{c}_i^l$, $\boldsymbol{u}_0^l$, and polynomial bases being used in lieu of those in Test Case 2. Higher ranks have been omitted from the figure due to the lack of variability with respect to the scale of the image. Depicting the univariate functions for each dimension and rank, the figure shows that the first rank of dimension

Table 5.11: Agreement between SR- and PC-based mean and STD for Test Case 3

	Ref. Mean	Ref. STD	Rel. Mean	Rel. STD
a	1.095586 (DU)	3.1368e-03 (DU)	2.6e-07	7.6e-05
h_e	4.673e-04	1.15e-03	7.8e-04	2.2e-03
k_e	-2.130e-03	1.115e-03	1.5e-04	7.2e-04
p_e	1.85e-05	1.0005e-03	4.2e-03	4.6e-05
q_e	-7.05e-06	9.9948e-04	4.7e-03	1.3e-05
$\lambda_{\mathcal{M}}$	2.46646 (rad)	0.606994 (rad)	1.1e-06	2.2e-06

Table 5.12: Residual RMS of 40 SR- and MC-based validation samples for Test Case 3

	MC Sample RMS	Residual RMS	Units
a	1.09614	4.4e-05	(DU)
h_e	1.3e-03	3.8e-04	(N/A)
k_e	2.3e-03	4.1e-04	(N/A)
p_e	9.1e-03	2.0e-05	(N/A)
q_e	1.02e-03	1.8e-05	(N/A)
$\lambda_{\mathcal{M}}$	2.4063	4.8e-04	(rad)

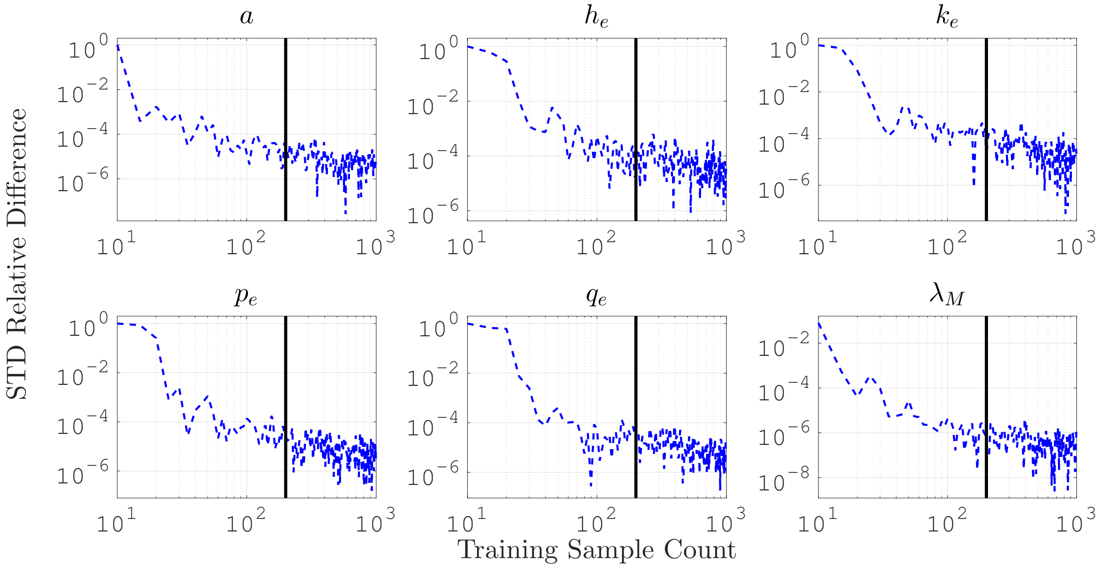


Figure 5.15: Plots of relative differences between approximated STD values for Test Case 3. Note the black line highlighting the choice of N for the surrogate.

a contains significantly more variability than any other univariate function.

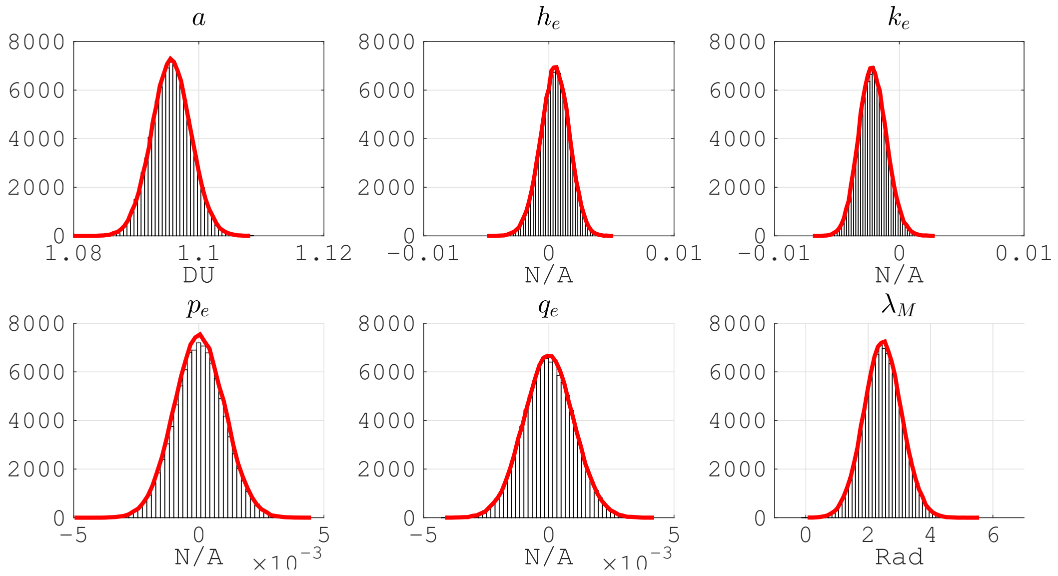


Figure 5.16: Histograms of SR results for quantities of interest in Test Case 3.

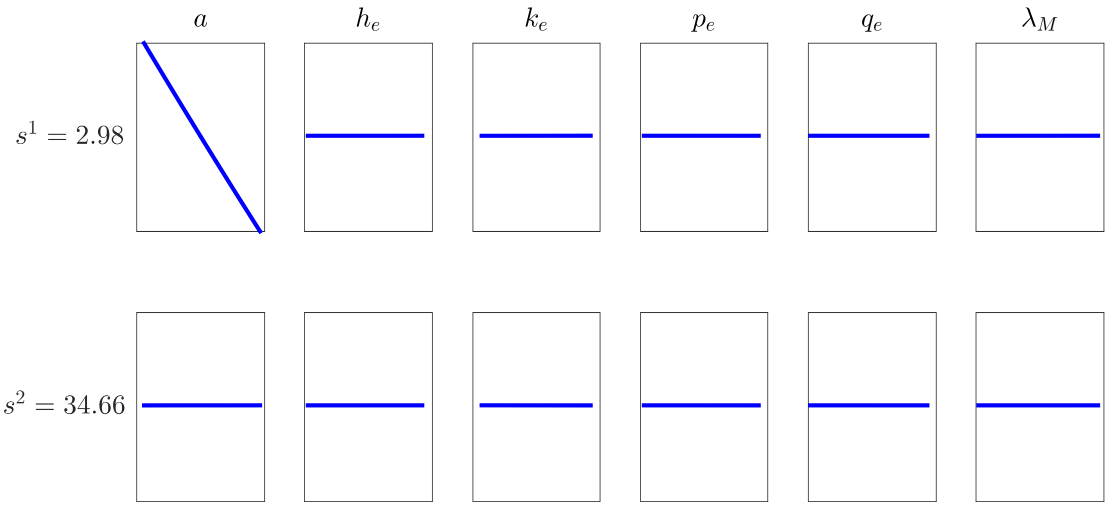


Figure 5.17: Plots of the absolute values of factors $u_i^l(y_i)$ for Test Case 3

5.5.1 Sensitivity Analysis

In order to quantitatively validate these conclusions, a sensitivity analysis is applied to the results of Test Case 3 and compared to a PC baseline. The results of this analysis can be found in Tables 5.13 and 5.14. As in Test Case 2, the PC baseline is converged to six decimal places.

Therefore, some values have been stated to be approximately 0. In [62], it is deduced that uncertainty in a is the most important contribution to equinoctial element variance. In particular, the variance of $\lambda_{\mathcal{M}}$ relies largely on uncertainty in a. Tables 5.13 and 5.14 show that variance in a and $\lambda_{\mathcal{M}}$ does indeed rely mostly on uncertainty in a. However, the Sobol indices present a more detailed analysis. The results are indicative of a relatively independent system, where the variability of each QoI is most affected by the variability of its corresponding random input. Figure 5.18 illustrates the variability of the QoI p_e propagated with uncertainty only in a or uncertainty in p_e. The uncertainty used in this analysis is taken from the respective values in Table 5.2. Therefore, the dependence of the variability of p_e with respect to uncertainty in a and p_e can be compared.

Table 5.13: Sensitivity indices $S_{i,m}$ and residuals of a, h_e and k_e for Test Case 3

	Quantities of Interest					
Random	a		h_e		k_e	
Inputs	PCE	Resid.	PCE	Resid.	PCE	Resid.
a	0.999	1e-03	0.249	5e-03	0.200	3e-03
h_e	~ 0	N/A	0.723	8e-03	0.0297	4e-04
k_e	2e-06	2e-07	0.0267	1e-03	0.769	3e-03
p_e	~ 0	N/A	~ 0	N/A	~ 0	N/A
q_e	~ 0	N/A	~ 0	N/A	~ 0	N/A
$\lambda_{\mathcal{M}}$	~ 0	N/A	~ 0	N/A	~ 0	N/A

Table 5.14: Sensitivity indices $S_{i,m}$ and residuals of p_e, q_e and $\lambda_{\mathcal{M}}$ for Test Case 3

	Quantities of Interest					
Random	p_e		q_e		$\lambda_{\mathcal{M}}$	
Inputs	PCE	Resid.	PCE	Resid.	PCE	Resid.
a	4e-06	3e-06	4e-06	4e-06	0.999	1e-03
h_e	~ 0	N/A	~ 0	N/A	~ 0	N/A
k_e	~ 0	N/A	~ 0	N/A	~ 0	N/A
p_e	0.963	2e-04	0.0364	3e-07	~ 0	N/A
q_e	0.0364	3e-05	0.963	1e-03	~ 0	N/A
$\lambda_{\mathcal{M}}$	~ 0	N/A	~ 0	N/A	~ 0	N/A

As illustrated, the effect of uncertainty in p_e is two orders of magnitude larger than the variability introduced by a. This independent behavior repeats for h_e, k_e and q_e and explains the relatively high rank, $r = 6$, needed for solution convergence.

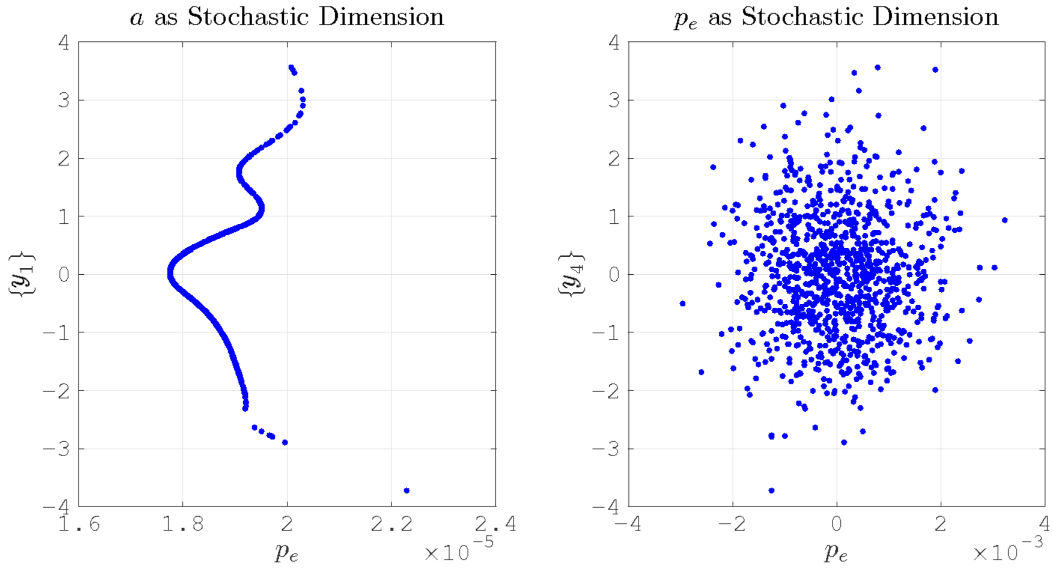


Figure 5.18: Plots of p_e propagated with uncertainty in a or p_e

5.6 Summary of Results

Containing the first applications of SR to the field of astrodynamics, the results of Chapter 5 establish the capability of SR to accurately estimate the PDFs of objects in LEO. Considering large initial uncertainty and long propagation times, the non-Gaussian posterior PDFs of a reference solution are accurately approximated by a surrogate which does not suffer from the curse of dimensionality. The overall linear increase in required training sample count is seen in Sections 5.3 and 5.4. As the input dimension is increased from $d = 6$ to $d = 20$, the required number of samples does not increase exponentially. In addition to this, ANOVA global sensitivity analyses are tractably performed by SR. For the case found in Section 5.5, the sensitivity analysis reveals nuances in the relationships between inputs and QoIs in this equinoctial element example. Specifically, the analysis shows that the system is largely independent, where the growth in uncertainty of each QoI is due to the associated random input.

Chapter 6

Probability of Collision Quantification and Optimization

6.0.1 Introduction

For this chapter, we present a tractable approach to designing a collision avoidance maneuver while considering uncertainty. SR is used for computing the probability of collision between one or more space objects. The capability provided by this surrogate method is then combined with the presented OUU approach such that a reliability design is computationally efficient. This approach is subject to a minimization function and a set of constraints, both of which are functions of uncertainty of the system. We demonstrate the proposed framework to design a single maneuver that avoids collisions in sequence with multiple debris objects. Therefore, we introduce SR as a tractable method for designing a collision avoidance maneuver that is robust with respect to uncertainty.

The tests and numeric results contained in this chapter consider the use of SR methods for the creation of polynomial surrogates that are able to map the relationship between inputs and quantities of interest (QoIs). The resulting surrogates are used for the propagation of orbit state PDFs without a posterior Gaussian assumption. Using a sampling method, the probability of collision is estimated over a time span. The ability to efficiently generate these sample sets with a surrogate is used within an OUU process in order to facilitate a reliability based design of a maneuver in orbit about the Earth.

6.1 Problem Setup

The presented problem, which serves as the focus of this chapter, is that of three objects orbiting the Earth. In this scenario, object 1 is scheduled to perform a maneuver for orbit maintenance. However, it is found that object 1 may collide with object 2 or object 3 depending on the applied ΔV. Thus, a maneuver must be found such that object 1 avoids objects 2 and 3, while desiring to change the prescribed maneuver as little as possible. In this case, we only consider maneuver design for object 1. An illustration of this scenario is included in Figure 6.1. It is noted that the presented picture is a notional diagram and serves as a visual aid for the timeline of events, including the first propagation to the maneuver time $t_{\Delta V}$ and the three maneuvers mentioned and calculated in this chapter. It is also noted that the overall approach taken in this chapter (that of using SR for reliability design) is not limited to designing a maneuver for one object to avoid two sequential collisions. In fact, the problem presented in this section and those following is an expression of the general formulation shown in Eq. (3.1). The following subsections cover the work necessary to perform such an analysis with a sampling approach such as MC.

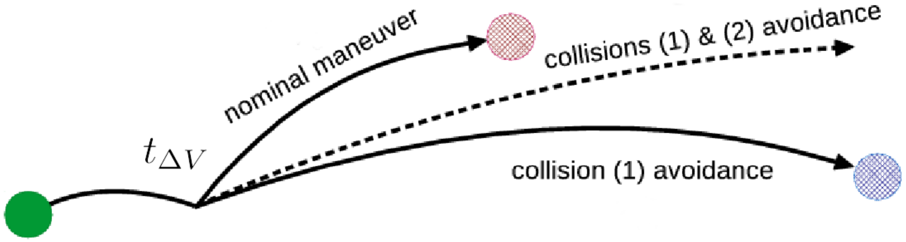


Figure 6.1: Illustration of the possible conjunction events and associated maneuvers.

6.2 Probability of Collision Calculation

As a general case, we present the calculation of the probability of collision via MC by considering the states of two independent objects in orbit about the Earth. Each is dependent on a separate vector of random variables such that $\boldsymbol{r}_1(t, \boldsymbol{\eta})$ and $\dot{\boldsymbol{r}}_1(t, \boldsymbol{\eta})$ are the position and velocity state of object 1 at time t, while $\boldsymbol{r}_2(t,\boldsymbol{\omega})$ and $\dot{\boldsymbol{r}}_2(t,\boldsymbol{\omega})$ are likewise for object 2. Here, $\boldsymbol{\eta}$ and $\boldsymbol{\omega}$ are vectors of independent random variables, where the state of each object is dependent on these separate vectors, and it is not necessary for the dimension of $\boldsymbol{\eta}$ to be the same as $\boldsymbol{\omega}$.

Using the states of the two objects, we may begin to consider the time at which the distance between the two is at a minimum, and, using this time and distance, we seek to compute the probability of collision over a period of time rather than at a single epoch. This parameter is also known as the time integrated probability of collision, but, for the sake of brevity, it will be referred to in this work as the probability of collision, or P_c, hereafter. To enable the calculation of this parameter, a MC approach is presented. As described later, SR allows for rapidly generating the required samples to enable P_c estimation in reliability design.

In order to determine the probability of collision, it is necessary to know the distance between two objects at the time of closest approach. To begin our definition, we present the square of the separation distance to be calculated as

$$s^2(t,\boldsymbol{\eta},\boldsymbol{\omega}) = (\boldsymbol{r}_1(t,\boldsymbol{\eta}) - \boldsymbol{r}_2(t,\boldsymbol{\omega})) \cdot (\boldsymbol{r}_1(t,\boldsymbol{\eta}) - \boldsymbol{r}_2(t,\boldsymbol{\omega})), \tag{6.1}$$

where t is time. Therefore, if we would like to find the minimum separation distance between two objects over a period of time, we evaluate Eq. (6.1) at the time of closest approach. In this work, this epoch is denoted as $\mathcal{T}$, which satisfies

$$\mathcal{T}(\boldsymbol{\eta},\boldsymbol{\omega}) = \arg\min_t s^2(t,\boldsymbol{\eta},\boldsymbol{\omega}). \tag{6.2}$$

Upon combining the initial state PDF for the two spacecraft with realized independent random variables, the realizations of the state at t_0 may be generated. These random samples are propagated to compute a set containing M_{MC} samples of the minimum separation distances (or closest approaches) between the two objects. Evaluating the process of Eqs. (6.1) and (6.2) M_{MC} times results in the set $\{s^2(\mathcal{T}_j,\boldsymbol{\eta}_j,\boldsymbol{\omega}_j)\}_{j=1}^{M_{MC}}$. The process and computation of this methodology are discussed in further detail within Section 6.1. In order for statistics to be calculated to some desired accuracy, it is necessary for M_{MC} to be chosen such that the set is sufficiently large. We also introduce the more compact notation

$$\mathcal{T}_j \equiv \mathcal{T}(\boldsymbol{\eta}_j, \boldsymbol{\omega}_j) \tag{6.3}$$

where $\mathcal{T}_j$ is the time of closest approach as a function of the jth random input.

Remark 8. *As seen in Eqs. (6.1) and (6.2), when applied to the set of data* $\{s^2(\mathcal{T}_j, \boldsymbol{\eta}_j, \boldsymbol{\omega}_j)\}_{j=1}^{M_{MC}}$*, each sample* j *is a function of the two vectors of random inputs* $\boldsymbol{\eta}_j$ *and* $\boldsymbol{\omega}_j$*. In order to reduce notation clutter, the two random input vectors are combined such that*

$$\boldsymbol{\xi}_j = \begin{bmatrix} \boldsymbol{\eta}_j \\ \boldsymbol{\omega}_j \end{bmatrix}, \tag{6.4}$$

where the random inputs are organized into the set $\{\boldsymbol{\xi}_j\}_{j=1}^{M_{MC}}$*, and each random input vector is sized such that* $\boldsymbol{\xi}_j \in \mathcal{R}^{d_\xi}$*. Here, the dimension* d_ξ *is a result of combining the stochastic dimension of* $\boldsymbol{\eta}$ *and* $\boldsymbol{\omega}$*.*

In turn, the set of distance samples is now explicitly dependent on both sets of random inputs $\boldsymbol{\eta}$ and $\boldsymbol{\omega}$, and, with the assistance of compacted notation, is denoted as

$$\{s^2(\mathcal{T}_j, \boldsymbol{\xi}_j)\}_{j=1}^{M_{MC}}. \tag{6.5}$$

The probability of collision is then estimated as

$$P_c = \frac{\text{count}(s^2(\mathcal{T}_j, \boldsymbol{\xi}_j) \leq \mathcal{R}^2)}{M_{MC}}. \tag{6.6}$$

where the operator count() indicates the number of positive tests over the $j = 1, \ldots, M_{MC}$ test points, and $\mathcal{R}$ is defined as the keep out radius.

Due to the fact that the convergence of a MC analysis is known to be inversely proportional to the square root of the number of samples used, it is possible that generating enough samples to compute a converged P_c via Eq. (6.6) would be computationally expensive. Therefore, alternative methods of calculating the P_c have been investigated. Using PC, Jones and Doostan [66] evaluates Eq. (6.2) using a discretized time variable $t_k \in [t_1, \ldots, t_K]$ in order to calculate $\mathcal{T}_j$ and the data set of Eq. (6.5). This method relies on the assumption that the discretized times t_k are dense enough to detect collisions. The work by Jones et al [68] expands on this methodology by introducing an approximation of the system that is continuous with respect to time. To find the minimum square distance between the two objects, an interpolating polynomial, one for each $j = 1, \ldots, M_{MC}$, is used to compute solutions at times between the discrete steps of t_k. The resulting polynomial is used as a surrogate in a given optimization algorithm (such as Brent's method) in order to find the estimated time of closest approach within the continuous problem. For the work presented in this chapter, we consider the creation of a surrogate whose output is at the time of closest approach for any given sample. Therefore, no extra computation work is needed to find the samples at this time of interest for collision analyses.

6.3 Quantities of Interest

Using the initial conditions of three objects in Earth orbit, it is desired to know the uncertainty of the orbital state at a later time in the form of realizations of the state. The set of stochastic ODEs

$$\boldsymbol{\mathcal{F}}\left(t, \boldsymbol{\xi}, \boldsymbol{\theta}; \boldsymbol{r}_1(t_0), \dot{\boldsymbol{r}}_1(t_0), \boldsymbol{\Delta V}, \boldsymbol{r}_2(t_0), \dot{\boldsymbol{r}}_2(t_0), \boldsymbol{r}_3(t_0), \dot{\boldsymbol{r}}_3(t_0)\right) = \boldsymbol{0}, \tag{6.7}$$

describe the temporal evolution of the states of the three objects in Earth orbit, where $\boldsymbol{\mathcal{F}}$ is a stochastic ODE operator, and $t \in [t_0, t_f]$ is time. It is noted that the realized maneuver $\boldsymbol{\Delta V}$ is not only dependent on the random variables $\boldsymbol{\xi}$ but also on the design variables $\boldsymbol{\theta}$, i.e., $\boldsymbol{\Delta V}(\boldsymbol{\xi}, \boldsymbol{\theta})$. The dependence on the random variables is due to the modeling of maneuver execution error, which highlights the need for OUU in this case. The optimization methodology must account for both random variables $\boldsymbol{\xi}$ and determinstic design inputs $\boldsymbol{\theta}$ if a reduction in system information is determined to be undesirable. If $\boldsymbol{\xi}$ were to be excluded, the optimization problem would be deterministic, i.e., non statistical, and unable to consider approaches such as those including reliability indices (as seen in Eq. (3.1)).

By evaluating Eq. (6.7) for the state of each object M_{MC} times, sets of data are collected, e.g.,

$\{\boldsymbol{r}_1(\mathcal{T}_j, \boldsymbol{\xi}_j, \boldsymbol{\theta}_j), \dot{\boldsymbol{r}}_1(\mathcal{T}_j, \boldsymbol{\xi}_j, \boldsymbol{\theta}_j)\}$ and $\{\boldsymbol{r}_2(\mathcal{T}_j, \boldsymbol{\xi}_j, \boldsymbol{\theta}_j), \dot{\boldsymbol{r}}_2(\mathcal{T}_j, \boldsymbol{\xi}_j, \boldsymbol{\theta}_j)\}$, and the uncertainty of the system represented in Eq. (6.7) is quantified. These data sets, with each sample propagated to $\mathcal{T}_j$, are then evaluated in Eqs. (6.1) and (6.2) so that the P_c can be computed in Eq. (6.6). We note that each value of $\mathcal{T}_j$ is not known *a priori*. The process used for finding these values is discussed later in Section 6.4. Because the considered case involves three objects and two potential collisions, notation is introduced to aid in identification. The QoIs discussed in this chapter consist of the position of an object, i.e.,

$$\boldsymbol{q}_{\alpha}^{(\beta)}(\boldsymbol{\xi}, \boldsymbol{\theta}) = \boldsymbol{r}_{\alpha}(\mathcal{T}^{(\beta)}, \boldsymbol{\xi}, \boldsymbol{\theta}) \tag{6.8}$$

where the position components of object $\alpha = 1, 2, 3$ are considered the elements of a single QoI vector. Eq. (6.8) is a generality of the QoIs, i.e., not yet sampled and organized into a set. It is noted that Greek letters α and β are used in lieu of Latin alphabet characters in order to distinguish the indices from subsequent indexing used for SR and other equations. Here, we also introduce the index of $\beta = 1, 2$. This value tracks the considered collision (the first or second) such that confusion between data sets is reduced. Explicitly, we have

$$\boldsymbol{q}_1^{(1)}(\boldsymbol{\xi}, \boldsymbol{\theta}) \neq \boldsymbol{q}_1^{(2)}(\boldsymbol{\xi}, \boldsymbol{\theta}), \tag{6.9}$$

where Eq. (6.9) indicates that the state of object 1 at the first time of closest approach (that with object 2) is not equal to that of object 1 at the second time of closest approach (that with object 3). This work therefore focuses on four distinct states, i.e, $\boldsymbol{q}_1^{(1)}$, $\boldsymbol{q}_2^{(1)}$, $\boldsymbol{q}_1^{(2)}$ and $\boldsymbol{q}_3^{(2)}$.

Remark 9. *Eq. (6.8) denotes the composition of the quantity of interest vector* $\boldsymbol{q}_{\alpha}^{(\beta)}(\boldsymbol{\xi}, \boldsymbol{\theta})$*. It is emphasized here that the dependence of the QoI on time has been dropped due to the fact that* $\boldsymbol{r}_{\alpha}(\mathcal{T}^{(\beta)}, \boldsymbol{\xi}, \boldsymbol{\theta})$ *and* $\dot{\boldsymbol{r}}_{\alpha}(\mathcal{T}^{(\beta)}, \boldsymbol{\xi}, \boldsymbol{\theta})$ *are a function of a fixed time* $\mathcal{T}^{(\beta)}$*, and the QoI is therefore only valid at this single point in time. The vector* $\boldsymbol{q}_{\alpha}^{(\beta)}(\boldsymbol{\xi}, \boldsymbol{\theta})$ *is still dependent on the inputs* $\boldsymbol{\xi}$ *and* $\boldsymbol{\theta}$*.*

6.4 Data Generation

The MC sampling approach of data generation begins by mapping random variables, associated with the object's state, to an initial condition of $\boldsymbol{r}_{\alpha}(t_0)$ and $\dot{\boldsymbol{r}}_{\alpha}(t_0)$. As an aid, we provide the example of adding uncertainty to the state of object 1. For the random inputs associated with object 1's state, samples of Gaussian distributions are generated and each sample $\xi_{i,j} \sim \mathcal{N}(0,1)$ is first mapped to the position and velocities for the object at t_0 via the state covariance $\boldsymbol{\Sigma}$. Therefore, the jth sample of the perturbed state is

$$\begin{bmatrix} \boldsymbol{r}_1(t_0, \boldsymbol{\xi}_j) \\ \dot{\boldsymbol{r}}_1(t_0, \boldsymbol{\xi}_j) \end{bmatrix} = \begin{bmatrix} \boldsymbol{r}_1(t_0) \\ \dot{\boldsymbol{r}}_1(t_0) \end{bmatrix} + \boldsymbol{\Gamma} \begin{bmatrix} \xi_{1,j} \\ \vdots \\ \xi_{6,j} \end{bmatrix}, \tag{6.10}$$

where $\boldsymbol{\Gamma}$ is the lower triangular Cholesky decomposition of $\boldsymbol{\Sigma}$, and we consider the first six entries of $\boldsymbol{\xi}$ to be associated with the position and velocity of object 1.

The state of object 1 is updated at the time of maneuver $t_{\Delta V}$ in order to reflect the change in velocity that the maneuver imparts, here modeled as an instaneous change in velocity. It is at this point that the design variables contribute to the propagated state. In this case, each design input sample is uniformly distributed over the bounds $[-1,1]$, i.e., $\theta_{i,j} \sim \mathcal{U}(-1,1)$ where $i = 1,\ldots,3$ indexes the three components of the maneuver velocity. Using Eq. (3.3), we have

$$\boldsymbol{\Xi}(\boldsymbol{\theta}_j) = \begin{bmatrix} \Xi_1(\theta_{1,j}) \\ \Xi_2(\theta_{2,j}) \\ \Xi_3(\theta_{3,j}) \end{bmatrix}, \tag{6.11}$$

the bounds of which are covered later on in Section 6.5. Eq. (6.11) is used to update the state velocity

$$\dot{\boldsymbol{r}}_1(t_{\Delta V}, \boldsymbol{\xi}_j, \boldsymbol{\theta}_j) = \dot{\boldsymbol{r}}_1(t_{\Delta V}, \boldsymbol{\xi}_j) + \boldsymbol{\Xi}(\boldsymbol{\theta}_j) + \boldsymbol{\Gamma}^{\Delta V} \begin{bmatrix} \xi_{7,j} \\ \xi_{8,j} \\ \xi_{9,j} \end{bmatrix}. \tag{6.12}$$

In Eq. (6.12), we have $\boldsymbol{\Gamma}^{\Delta V}$, which is the Cholesky decomposition of the maneuver error covariance matrix $\boldsymbol{\Sigma}^{\Delta V}$. Here the 7-th, 8-th, and 9-th entries of $\boldsymbol{\xi}$ are associated with the maneuver error of object 1. Although presented as an example centered around object 1, Eqs. (6.10)-(6.12) may be further generalized and applied to other objects.

Initially, the position elements of each object are solved for at a series of discrete times t_k which straddle the nominal $\mathcal{T}^{(\beta)}$ at the times $t_k \in [t_1, t_2, \ldots, t_K]$. Using this set of position vectors $\{\boldsymbol{r}_\alpha(t_k, \boldsymbol{\xi}_j, \boldsymbol{\theta}_j)\}$ the distances between the objects can be computed. Searching through the set of realizations at all times t_k, the minimum distance is found over the discrete set. The solution $t_{\hat{k}}$ to the minimum distance is initially a rough estimate for $\mathcal{T}_j^{(\beta)}$. It is further refined using the Brent optimization algorithm among the bounds $[t_{\hat{k}-1}, t_{\hat{k}+1}]$. A basic illustration of this process is included in Figure 6.2. The states at time steps $t_k \in [t_1, t_2, t_3, t_4]$ are included for each object and superimposed over the respective continuous orbit. Although the overall geometry is exaggerated, the basic behavior is representative of that found in the methodology. Even though $t_{\hat{k}} = t_2$ may appear to be the time at which the objects are at a minimum distance when only considering the discrete samples, the actual $\mathcal{T}_j^{(\beta)}$ is at a point between t_2 and t_3. This distance and the associated $\mathcal{T}_j^{(\beta)}$ are then found using a Brent optimizer and the black box propagator [22]. For the results presented in this chapter, the Brent optimization function found in the Python Scipy library is used with a relative tolerance of 10^{-10}. By using a method such as Brent's, the full dynamics of the problem are utilized.

Once a value of $\mathcal{T}_j^{(\beta)}$ has been found, the corresponding sample j can be propagated to the time of closest approach in order to yield $\boldsymbol{q}_\alpha^{(\beta)}(\boldsymbol{\xi}_j, \boldsymbol{\theta}_j)$. Thus, the MC data are calculated and subsequently organized into the data set of $\{(\boldsymbol{\xi}_j, \boldsymbol{\theta}_j, \boldsymbol{q}_\alpha^{(\beta)}(\boldsymbol{\xi}_j, \boldsymbol{\theta}_j))\}_{j=1}^{M_{MC}}$. This data set is then used to solve for various statistics or to analyze the PDF for methodologies such as determining the probability of collision. In order to quantify the risk of collision, the probability of collision, as seen in Eq. (6.6), is utilized.

6.5 Optimization and constraints

The goal of this analysis is to constrain the probability of collision to be below an acceptable value while also minimizing the changes to the previously scheduled maneuver. Because P_c is a value fundamentally reliant upon a probability density function (that of $\boldsymbol{\xi}$), including it in the constraints of an optimization algorithm requires the approach of OUU presented in Eq. (3.1).

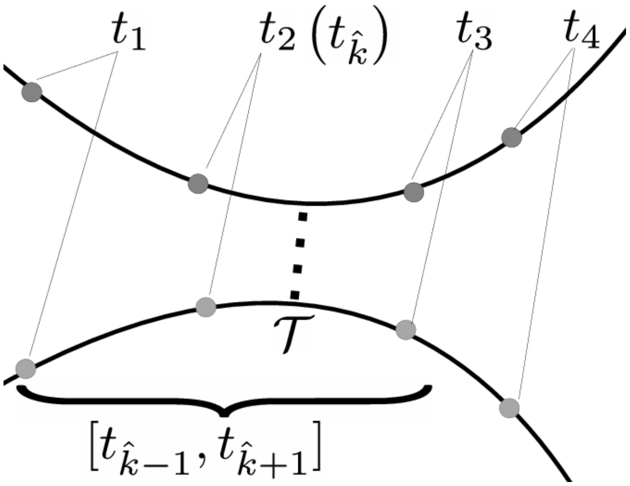


Figure 6.2: Illustration of the data $\{\boldsymbol{r}_\alpha\,(t_k, \boldsymbol{\xi}_j, \boldsymbol{\theta}_j)\}$ and the location of the time of closest approach $\mathcal{T}$.

Therefore the resulting optimization approach is

$$\begin{aligned} \text{minimize} \quad & f = \|\boldsymbol{\Delta V}(\boldsymbol{\xi}, \boldsymbol{\theta}) - \boldsymbol{\Delta V}_0\| \\ \text{subject to} \quad & P_c^{(1)}(\boldsymbol{\xi}, \boldsymbol{\theta}) \leq \kappa \\ & P_c^{(2)}(\boldsymbol{\xi}, \boldsymbol{\theta}) \leq \kappa, \end{aligned} \tag{6.13}$$

where f is the function to be minimized, $\boldsymbol{\Delta V}(\boldsymbol{\xi}, \boldsymbol{\theta})$ is the maneuver to be designed, $\boldsymbol{\Delta V}_0$ is the nominal maneuver, and κ is the acceptable threshold for the probability of collision. The particular value of $\boldsymbol{\theta}$ that satisfies the contraints and minimizes f is given the notation $\bar{\boldsymbol{\theta}}^{(1,2)}$, where the bar notifies that this particular vector of design variables is the argument which satisfies Eq. (6.13), and the superscript indicates compliance with avoiding both collision (1) and (2). Therefore, the maneuver which avoids collisions (1) and (2) is denoted as $\boldsymbol{\Delta V}(\boldsymbol{\xi}, \bar{\boldsymbol{\theta}}^{(1,2)})$. From here on, the notation of the designed maneuver is simplified for conciseness

$$\overline{\boldsymbol{\Delta V}}^{(1,2)} \equiv \boldsymbol{\Delta V}(\boldsymbol{\xi}, \bar{\boldsymbol{\theta}}^{(1,2)}). \tag{6.14}$$

In Eq. (6.13), the reliability index β introduced in Eq. (3.1) is replaced by κ. The switch from the $\geq$ inequality to that of $\leq$ is due to the fact that the time integrated probability of collision considered in this chapter is the compliment of the probability of survival, i.e., $P_c = 1 - P_s$. Rather than present the case of constraining the problem to a P_s greater than some desired survival index, we maintain our use of the probability of collision which must remain below or equal to a threshold value.

The designed maneuver $\overline{\boldsymbol{\Delta V}}^{(1,2)}$ is computed by dictating a vector of design variables $\bar{\boldsymbol{\theta}}^{(1,2)}$ from the given design space. This design space of the maneuver is created by mapping the design variables to the nominal maneuver $\boldsymbol{\Delta V}_0$ in the NTW frame. This coordinate frame has the $\hat{\boldsymbol{T}}$ or transverse axis pointed in the direction of the velocity vector. The $\hat{\boldsymbol{W}}$ axis is then parallel to the angular velocity, while the $\hat{\boldsymbol{N}}$ direction completes the orthogonal coordinate frame. The respective upper and lower limits, as depicted in Eq. (3.3), are user determined, e.g., based on absolute or relative bounds. In the case of this work, the following variables are bounded,

$$\begin{aligned} \Delta V_N^{lower} &\leq \Xi_1(\theta_1) \leq \Delta V_N^{upper} \\ \Delta V_T^{lower} &\leq \Xi_2(\theta_2) \leq \Delta V_T^{upper} \\ \Delta V_W^{lower} &\leq \Xi_3(\theta_3) \leq \Delta V_W^{upper}. \end{aligned} \tag{6.15}$$

Therefore, a solution to the function and constraints of Eq. (6.13) is searched for by applying an optimization algorithm to explore the design space of $\boldsymbol{\Delta V}(\boldsymbol{\xi}, \boldsymbol{\theta})$ represented by the values of $\Xi(\boldsymbol{\theta})$ found in Eq. (6.15). As mentioned in Note 6 however, the computational cost of repeatedly evaluating M_{MC} samples for each candidate vector of design variables can be prohibitive. Therefore, there is motivation to evaluate Eq. (6.13) with a different methodology.

6.6 Surrogate Notation

For the purposes of this application, Eq. (3.4) is adopted with the following notation

$$\boldsymbol{q}_\alpha^{(\beta)}(\boldsymbol{\xi}, \boldsymbol{\theta}) \approx \hat{\boldsymbol{q}}_\alpha^{(\beta)}(\boldsymbol{\xi}, \boldsymbol{\theta}) = \sum_{l=1}^{r} s^l \, \boldsymbol{u}_0^l \prod_{i'=1}^{d_\theta} u_{i'}^l \left(\theta_{i'}\right) \prod_{i=1}^{d_\xi} u_i^l \left(\xi_i\right), \tag{6.16}$$

where the previous presentation has been altered to reflect the introduced notation of object and collision indexing. The surrogates $\hat{\boldsymbol{q}}_\alpha^{(\beta)}(\boldsymbol{\xi}, \boldsymbol{\theta})$ are constructed using the training data $\mathcal{D} = \{(\boldsymbol{\xi}_j, \boldsymbol{\theta}_j, \boldsymbol{q}_\alpha^{(\beta)}(\boldsymbol{\xi}_j, \boldsymbol{\theta}_j))\}_{j=1}^{N}$, and the resulting separable polynomial functions are evaluated M_{MC}

times to compute the sets $\{\boldsymbol{q}_\alpha^{(\beta)}(\boldsymbol{\xi}_j, \boldsymbol{\theta}_j)\}_j^{M_{MC}}$. From these sets, statistics and the probability of collision $P_c^{(\beta)}$ are calculated.

6.7 Results

As discussed in Section 6.1, this chapter considers the case of an object (object 1) attempting to avoid sequential collisions with objects 2 and 3. Each object state is computed with a Runge-Kutta 7(8) propagator that considers solar radiation pressure, gravitational perturbation from the Moon and Sun, as well as a 50×50 spherical harmonic representation of the Earth's gravity field [37]. Object 1 undergoes a small maneuver 24 hours after the case's epoch time (September 12, 2017, 23:58:51.817 UTC), modeled in the NTW frame. Noting a potential collision with object 2 three hours after the maneuver time, an alteration to this nominal maneuver is calculated using the optimization problem presented in Eq. (6.13). Since object 3 is not yet a concern, design inputs $\bar{\boldsymbol{\theta}}^{(1)}$ are solved for such that the cost function is minimized with only the first constraint considered, i.e., $P_c^{(1)}(\boldsymbol{\xi}, \bar{\boldsymbol{\theta}}^{(1)}) \leq \kappa$. This resulting maneuver $\overline{\boldsymbol{\Delta V}}^{(1)} \equiv \boldsymbol{\Delta V}(\boldsymbol{\xi}, \bar{\boldsymbol{\theta}}^{(1)})$ is then found to send object 1 onto a collision course with object 3. By introducing a possible collision between objects 1 and 3, this change artificially complicates the collision avoidance procedure necessary for the object 1 and 2 conjunction.

Therefore, the following results present a potential collision over two encounters, where a maneuver, designed to avoid a statistically significant collision with object 2, is not viable due to another potential collision with object 3. As an aid, we refer the reader to Figure 6.1, which illustrates the described scenario. The nominal maneuver $\boldsymbol{\Delta V}$ is seen as resulting in a conjunction with object 2 in red, while the first optimized maneuver $\overline{\boldsymbol{\Delta V}}^{(1)}$ is labeled as "collision (1) avoidance" and results in a conjunction with object 3 in blue. The desired optimized maneuver $\overline{\boldsymbol{\Delta V}}^{(1,2)}$, which is the ultimate goal of the optimization problem, avoids collisions (1) and (2). We note that this test case is based on "Case 3" from Alfano [3]. Alterations have been made for the sake of increased complexity, including a maneuver, and a third object.

6.7.1 Case Conditions

The initial conditions of the three objects are considered as random inputs for the stochastic problem, and the values of these states are located in Table 6.1, along with additional random inputs in the form of the nominal maneuver components $\boldsymbol{\Delta V}_0 = [\Delta V_{0,N}\, \Delta V_{0,T}\, \Delta V_{0,W}]^{\mathrm{T}}$. These initial conditions are used as the means of the initial density functions, upon which uncertainty is applied. As a reference aid, Table 6.2 contains a list of the major events that occur during the studied case, along with the time of the events relative to the epoch, and the objects for which a surrogate is created. The times of the first and second collision are calculated via the mean states of each object. For the case of object 1, the initial condition found in Table 6.1 occurs at the epoch time t_0. Objects 2 and 3, however, have initial PDFs originating at the time of maneuver $t_{\Delta V}$. Therefore, the state of object 1 is propagated for 24 hours longer than either object 2 or 3.

Table 6.1: Random inputs for objects 1, 2, and 3

	Mean		
	Object 1	Object 2	Object 3
x (m)	-27891608.7	-33021744.7	-31062467.4
y (m)	-30077310.3	-24238863.3	18517034.2
z (m)	196242.2	-643450.3	3874262.4
$\dot{x}$ (m/s)	-2296.4	-1855.7	777.8
$\dot{y}$ (m/s)	2092.3	2498.6	3149.9
$\dot{z}$ (m/s)	-384.1	-374.4	194.7
$\Delta V_{0,N}$ (m/s)	0.1	N/A	N/A
$\Delta V_{0,T}$ (m/s)	0.5	N/A	N/A
$\Delta V_{0,W}$ (m/s)	0.1	N/A	N/A

Table 6.2: Timeline of Nominal Events

Event	Time (s)	Surrogates
Begin Time (t_0)	0	N/A
Maneuver ($t_{\Delta V}$)	86400	N/A
First Collision $\mathcal{T}^{(1)}$	367200	Object 1 ($\hat{\boldsymbol{q}}_1^{(1)}$), Object 2 ($\hat{\boldsymbol{q}}_2^{(1)}$)
Second Collision $\mathcal{T}^{(2)}$	378000	Object 1 ($\hat{\boldsymbol{q}}_1^{(2)}$), Object 3 ($\hat{\boldsymbol{q}}_2^{(2)}$)

For the given case, object 2 is 1000 times more uncertain (with respect to variance) than object 3. This represents a lack of knowledge of object 2's state when compared to object 3. For this chapter, the uncertainty of the object at the epoch time is quantifiably represented by covariance matrices, i.e.,

$$\boldsymbol{\Sigma}_\alpha = \begin{bmatrix} \boldsymbol{\Sigma}_\alpha^{\text{pos}} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_\alpha^{\text{vel}} \end{bmatrix}. \tag{6.17}$$

Equation (6.17) depicts the uncertainty of the random inputs in the form of a covariance matrix for each object, which is applied to the initial conditions by way of Eq. (6.10). It is noted that the covariance matrices used for objects 1 and 3 (and subsequently object 2) are taken from Alfano [3] and rotated to reflect a rotation of the object states from that presented in the referenced material.

Each matrix segment is of size 3 × 3. The sections associated with the position uncertainty of an object $\boldsymbol{\Sigma}_\alpha^{\text{pos}}$ are,

$$\boldsymbol{\Sigma}_1^{\text{pos}} = \begin{bmatrix} 0.0569 & -0.0234 & 0.0035 \\ -0.0234 & 0.0722 & -0.0048 \\ 0.0035 & -0.0048 & 0.0407 \end{bmatrix} \quad \text{m}^2, \tag{6.18}$$

$$\boldsymbol{\Sigma}_3^{\text{pos}} = \begin{bmatrix} 0.0574 & -0.0235 & -0.0035 \\ -0.0235 & 0.0718 & -0.0047 \\ 0.0035 & -0.0047 & 0.0407 \end{bmatrix} \quad \text{m}^2, \tag{6.19}$$

while the values associated with the velocity $\boldsymbol{\Sigma}_\alpha^{\text{vel}}$ are

$$\boldsymbol{\Sigma}_1^{\text{vel}} = \boldsymbol{\Sigma}_3^{\text{vel}} = 10^{-8} \cdot \mathbb{I}_3 \quad \text{m}^2/\text{s}^2. \tag{6.20}$$

The covariance of object 2 is based on that of object 2, where

$$\boldsymbol{\Sigma}_2 = 1000 \times \boldsymbol{\Sigma}_3. \tag{6.21}$$

For this particular case, the covariance of the maneuver random inputs $\boldsymbol{\Sigma}_1^{\Delta V}$ is in the NTW coordinate frame

$$\boldsymbol{\Sigma}_1^{\Delta V} = 4 \times 10^{-8} \cdot \mathbb{I}_3 \quad \text{m}^2/\text{s}^2. \tag{6.22}$$

In the case of object 1, uncertainty is applied to the initial conditions contained in Table 6.1 using appropriate random variables and $\mathbf{\Sigma}_1$. Each sample j is forward propagated the 24 hours until the maneuver time $t_{\Delta V}$. At this point, the maneuver (with errors and design inputs) is added to object 1's velocity as seen in Eq. (6.12), and the state is further propagated until $\mathcal{T}_j$. For objects 2 and 3, there is no initial propagation of 24 hours. Thus, the states of object 2 and 3, presented in Table 6.1, are perturbed using random inputs and the respective covariances and propagated from $t_{\Delta V}$ until $\mathcal{T}_j^{(\beta)}$.

In addition to quantifying the uncertainty, this case also considers a design space which is applied to the OUU approach presented in Eq. (6.13). The parameters needed to complete the OUU analysis are included in Table 6.3. This table contains values for the upper limits of the allowable probability of collision κ, the designated keep out radius $\mathcal{R}$ used in Eq. (6.6), and the lower and upper bounds on the maneuver design space. In this case, the design space is dictated to be ±10% of the nominal maneuver components. It is also noted that the value of $\kappa = 10^{-4}$ is implemented due to its use in the astrodynamics community (e.g., Schilling et al [100]).

Table 6.3: OUU Parameters

	Value	Units
κ	10^{-4}	N/A
$\mathcal{R}$	70	m
ΔV_N^{lower}	0.09	m/s
ΔV_N^{upper}	0.11	m/s
ΔV_T^{lower}	0.45	m/s
ΔV_T^{upper}	0.55	m/s
ΔV_W^{lower}	0.09	m/s
ΔV_W^{upper}	0.11	m/s

Remark 10. *Because there are 3 objects each with a minimum of 6 random inputs associated with respective initial conditions, the addition of* $\mathbf{\Sigma}_1^{\Delta V}$ *and* $\boldsymbol{\theta}$ *could, at first glance, result in a dimension greater than 18. The methodology of creating the training data* $\mathcal{D}$*, however, limits this dimension to* $d = 18$*. Due to the fact that each value of* $\mathcal{T}_j^{(\beta)}$ *is only reliant upon the two objects corresponding*

to the conjunction β*, the inputs attributed with the irrelevant object are not included in the training data. Therefore, they are not included as inputs in the creation of the relevant surrogates.*

6.7.2 Probability of Collision: OUU

Once a converged solution is found for the system response at times $\mathcal{T}_j^{(\beta)}$ as a function of the input variables, it then becomes possible to consider an OUU application. For this analysis, the scenario presented in Section 6.5 is the problem of interest: with the available bounds of the maneuver design space, the goal is to avoid two possible collisions while minimizing the change to the nominally chosen maneuver. Here, we utilize the computational efficiency of the surrogates $\hat{\boldsymbol{q}}_\alpha^{(\beta)}(\boldsymbol{\xi}, \boldsymbol{\theta})$ in order to add tractability to the optimization process.

6.7.2.1 Surrogate Creation and Testing

Using $N = 2000$, $r = 10$ and $P = 5$, an SR solution for each object's Cartesian position is calculated from a single set of training data $\mathcal{D}$. It is here that we note that, for the generation of each surrogate, the total dimension $d = 18$. The results of samples generated with an SR solution and those generated by the reference propagator are contained in Figures 6.3 and 6.4 for an examination of the fit. For these figures, an independent set of $M_{MC} = 10^5$ MC samples is used as a reference. These values are displayed as the black bins of the histograms. The trained surrogates $\hat{\boldsymbol{q}}_\alpha^{(\beta)}(\boldsymbol{\xi}, \boldsymbol{\theta})$ are then evaluated at the associated random inputs, and the results are overlayed on the MC histograms as red lines. Qualitatively, the samples produced with the SR surrogate appear to approximate the distribution of the reference solution well. It is noted that the histograms of Figures 6.3 and 6.4 do not display PDFs due to the variation of the deterministic design inputs. When considering the Wiener-Askey scheme and the orthogonal relationship between a polynomial base and its associated distribution, however, it is helpful to think of the design inputs as random. In this case, the distributions of both figures emphasize the evolution from the prior Gaussian distributions to the posterior non-Gaussian distributions. In particular, the PDFs estimated in Figure 6.3 are neither representative of a Gaussian or uniform distribution, thus they are not a simple fit to either the Hermite or Legendre bases of which the SR surrogates are composed of.

In order to provide quantitative proof of convergence, validation samples are calculated. For this, 1000 samples of the QoIs are computed with the same methodology used for the creation of the training data $\mathcal{D}$. The associated random inputs are then used as inputs to the constructed surrogates, and the output is compared to the black box reference values. By comparing absolute residuals of these small sets of data, which are independent from the training of the surrogate, this validation process evaluates the accuracy of the SR solutions better than comparing solutions of data used during the fitting process. Due to the relatively small number of samples required, this validation approach is more indicative of that used in an operational environment when compared to a process that utilizes large numbers of samples (such as that presented later in Table 6.9). The RMS for each QoI is taken over the number of validation samples and provided as a quantitative measure for identifying solution covergence. The RMS of the differences is presented in Table 6.4. As seen in the table, the validation test indicates that an error of 17.5 m is achieved for any single position component.

Table 6.4: Residual RMS of 1000 SR- and MC-based validation samples for the test case

	MC Sample RMS	Residual RMS	Units
$x_1^{(1)}$	0.01466e04	0.0175	km
$y_1^{(1)}$	4.17156e04	0.0105	km
$z_1^{(1)}$	0.36377e04	0.0044	km
$x_2^{(1)}$	0.01499e04	0.0109	km
$y_2^{(1)}$	4.17156e04	0.0038	km
$z_2^{(1)}$	0.36376e04	0.0012	km
$x_1^{(2)}$	2.98324e04	5.586e-04	km
$y_1^{(2)}$	2.91666e04	5.375e-04	km
$z_1^{(2)}$	1.00684e04	3.941e-04	km
$x_3^{(2)}$	2.98358e04	1.456e-04	km
$y_3^{(2)}$	2.91588e04	1.678e-04	km
$z_3^{(2)}$	1.00670e04	3.876e-04	km

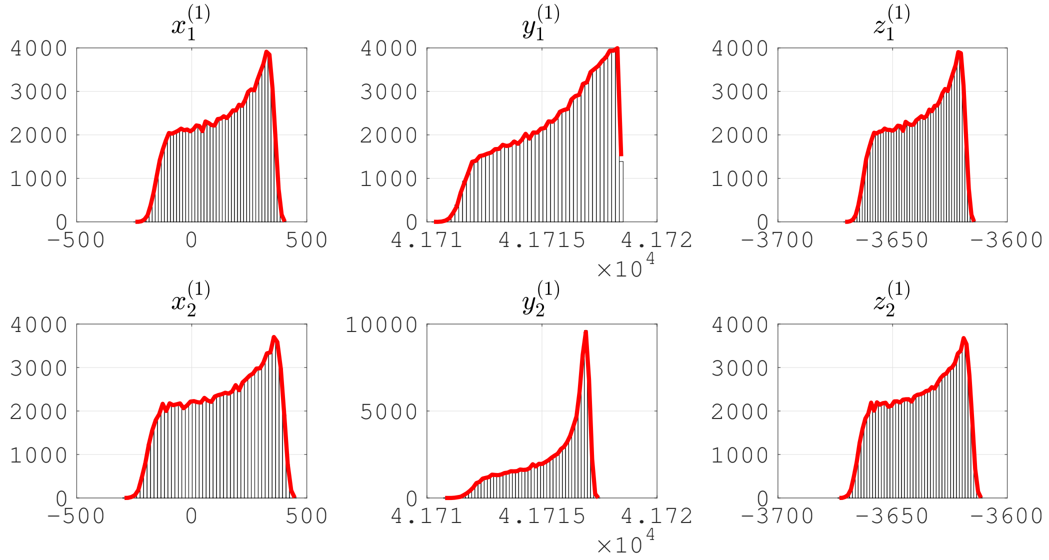


Figure 6.3: Histograms of collision (1) QoIs for the OUU case. Note, all units are in meters.

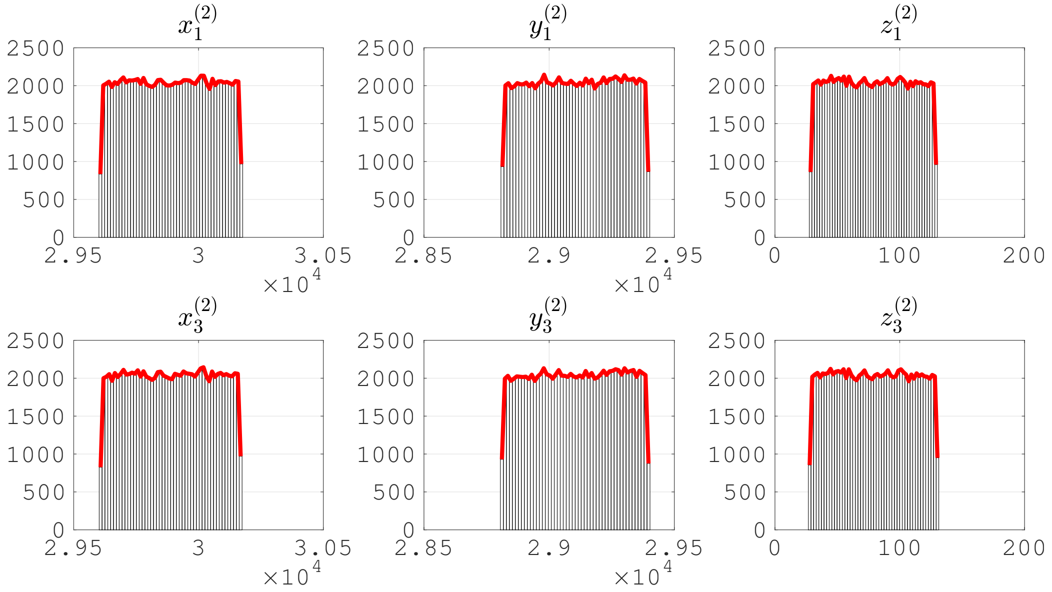


Figure 6.4: Histograms of collision (2) QoIs for the OUU case. Note, all units are in meters.

6.7.2.2 OUU Results

Following the optimization scheme presented in Eq. (6.13), the SR surrogates are leveraged in MATLAB's *fmincon* in order to produce results for the design variables chosen by the optimizer at each step. In this process, default option values of the MATLAB version 2017a are utilized with the exception of the minimum change in design inputs for calculating finite-difference gradients. A value of 10^{-1} is used for this test case due to the need for the probability of collision to differ between choices of the design inputs. This issue arises because values of $P_c^{(\beta)}$ used within the constraint functions are calculated by counting the number of hits between objects, and these discrete values may not change with small variations in the design inputs even though the overall distributions of the QoIs have been altered by such a change[1] . This change in settings emphasizes the care that must be taken when approaching an optimization problem for a given test case.

For the purpose of numerical accuracy, the surrogates are used to produce a set of $M_{MC} = 10^6$ samples for each object at each design input candidate vector. These sets are then computed to calculate the value of $P_c^{(\beta)}$ to be used in Eq. (6.13), and the associated value of $\overline{\boldsymbol{\Delta V}}^{(1,2)}$ is subsequently checked for compliance with respect to the minimization requirement. This process is repeated until the optimization algorithm determines that the optimal value of $\overline{\boldsymbol{\Delta V}}^{(1,2)}$ is found. For the $\overline{\boldsymbol{\Delta V}}^{(1,2)}$ result presented in this chapter, *fmincon* took 12 iterations. As discussed in Note 6, the computation of propagating $M_{MC} = 10^6$ MC samples would be repeated for each of these 12 iteration using a black box propagator if the SR surrogates are not leveraged.

Table 6.5 contains the optimization results relevant to the designed maneuver. In it, the values of the design variables that comply with the optimization requirements $\bar{\boldsymbol{\theta}}^{(1,2)}$ are included, along with the corresponding change in the maneuver $\delta\boldsymbol{\Delta V} = \overline{\boldsymbol{\Delta V}}^{(1,2)} - \boldsymbol{\Delta V}_0$. For comparison, the results of $\overline{\boldsymbol{\Delta V}}^{(1)}$ are included in Table 6.6. Only avoiding the collision between objects 1 and 2, $\overline{\boldsymbol{\Delta V}}^{(1)}$ is inadequate when constraining both $P_c^{(1)}$ and $P_c^{(2)}$. The N and W components of each optimization are orders of magnitude less than the T component. It is suspected that this behavior is due to relative insensitivity between the PDF and the directions in question. A change in velocity in the transverse (T) direction directly correlates with a change in semimajor axis and orbital energy. Therefore, there is a potential application for an optimization alogrithm focused on

---

[1] For this particular application, the setting change for the minimum difference used to calculate finite-difference gradients is the *DiffMinChange* variable within *fmincon*'s *optimoptions* function

sparseness in future work.

Unoptimized (i.e., based on the nominal maneuver) and optimized probability of collision $P_c^{(\beta)}$ values are included in Table 6.7. This table contains values calculated with maneuvers corresponding to the nominal $\boldsymbol{\Delta V}_0$, as well as the first optimization attempt, i.e., $\overline{\boldsymbol{\Delta V}}^{(1)}$, and the OUU case result $\overline{\boldsymbol{\Delta V}}^{(1,2)}$. For each of these maneuvers, there is a reference MC result computed via $M_{MC} = 10^5$ black box calculations as well as those computed by evaluating the $\hat{\boldsymbol{q}}_\alpha^{(\beta)}(\boldsymbol{\xi}, \boldsymbol{\theta})$ surrogates with a set of $M_{MC} = 10^6$ random inputs (SR). The choice of inputs $\boldsymbol{\theta}$ corresponds with the particular maneuver, e.g., $\boldsymbol{\theta} = \mathbf{0}$ for $\boldsymbol{\Delta V}_0$.

An addditional column SR$_{val}$ is included in Table 6.7 as a further check of the optimization results. For these data, the same SR parameters from the OUU analysis, e.g., r and N, are utilized in a surrogate that is fitted to training data that has been generated by fixing $\boldsymbol{\theta}_j = \bar{\boldsymbol{\theta}}^{(1,2)}$. When constructing this surrogate, the dimension is reduced to $d = 15$, due to the lack of variation among the design space. By evaluating 10^6 random samples via this representation $\hat{\boldsymbol{q}}_\alpha^{(\beta)}(\boldsymbol{\xi})$, we obtain a validation on the $P_c^{(\beta)}$ values for the optimized case. A check of this contribution indicates that the inputs dictated via the optimization scheme do indeed avoid collisions with objects 2 and 3.

A quick check of Table 6.7 shows values of $P_c^{(\beta)}$ that are similar to varying degrees between MC and SR results. As presented in Schilling et al [100], confidence intervals are a valuable tool for determining the possible error of a statistical estimation. Specifically, the exact Clopper-Pearson (CP) interval provides bounds that, although conservative, contain the true probability with a high rate of accuracy. The following bounds are given for the CP interval

$$CP = \left[B^{-1}(\alpha/2; n_H, M_{MC} - n_H + 1), B^{-1}(1 - \alpha/2; n_H + 1, N - n_H)\right], \tag{6.23}$$

where n_H is the number of observed hits, and $B^{-1}()$ returns the value producing a probability from the cumulative Beta distribution. When evaluating Eq. (6.23) using the $P_c^{(\beta)}$ values of the MC results (and the associated $M_{MC} = 10^5$), bounds are produced given a 95% confidence interval. That is, the true probability is contained within the calculated bounds under 95% of independent analyses. These bounds are included in Table 6.8, and the results suggest that both the SR and

MC results are within the expected error of each other when considering a M_{MC} of 10^5. These results are further supported by the histograms of Figure 6.5. The histograms of this figure are produced by evaluating the surrogates using appropriate design variables and 100 independent sets of random variables where $M_{MC} = 10^6$. Therefore, the variability between independent estimations of $P_c^{(\beta)}$ using $\hat{\boldsymbol{q}}_\alpha^{(\beta)}(\boldsymbol{\xi}, \boldsymbol{\theta})$ can be seen. Due to a complete lack of variation from the values presented in Table 6.7, results for $P_c^{(2)}$ are omitted. From this, we observe that the estimated values of $P_c^{(1)}$ remain within their respective CI bounds of Table 6.8.

Table 6.5: Reliability design of $\overline{\boldsymbol{\Delta V}}^{(1,2)}$ for the OUU case

	Value	Component	$\delta\boldsymbol{\Delta V}$
$\bar{\theta}_1$	-1.632e-03	N	-1.632e-03 cm/s
$\bar{\theta}_2$	0.2074	T	1.037 cm/s
$\bar{\theta}_3$	5.655e-03	W	5.655e-03 cm/s

Table 6.6: Reliability design of $\overline{\boldsymbol{\Delta V}}^{(1)}$

	Value	Component	$\delta\boldsymbol{\Delta V}$
$\bar{\theta}_1$	-1.051e-03	N	-1.051e-03 cm/s
$\bar{\theta}_2$	-0.2140	T	-1.070 cm/s
$\bar{\theta}_3$	3.603e-04	W	3.603e-04 cm/s

As an additional validation check, the 10^5 nominal MC samples used for the creation of the data in Table 6.7 are compared by evaluating the same random inputs in the surrogate ($\hat{\boldsymbol{q}}_\alpha^{(\beta)}(\boldsymbol{\xi}, \mathbf{0})$). This test is expanded by computing 10^5 additional MC samples, this time with the design inputs resulting from the OUU analysis found in Table 6.5, i.e., $\boldsymbol{\theta} = \bar{\boldsymbol{\theta}}^{(1,2)}$. These samples are then compared to surrogate outputs ($\hat{\boldsymbol{q}}_\alpha^{(\beta)}(\boldsymbol{\xi}, \bar{\boldsymbol{\theta}}^{(1,2)})$) using the same inputs as the MC samples. The results of these two analyses are included in Table 6.9. When considering the results of Table 6.4 to Table 6.9, the errors of the validation samples compare well those computed with the larger set of MC samples. This table also contains a third column dedicated to the post-optimization

Table 6.7: $P_c^{(\beta)}$ results for the OUU case

	$\boldsymbol{\Delta V}_0$		$\overline{\boldsymbol{\Delta V}}^{(1)}$		$\overline{\boldsymbol{\Delta V}}^{(1,2)}$		
	MC	SR	MC	SR	MC	SR	SR$_{val}$
$P_c^{(1)}$	1.89e-02	1.87e-02	1.50e-04	7.50e-05	1.30e-04	1.00e-04	1.23e-04
$P_c^{(2)}$	0.0	0.0	1.0	1.0	0.0	0.0	0.0

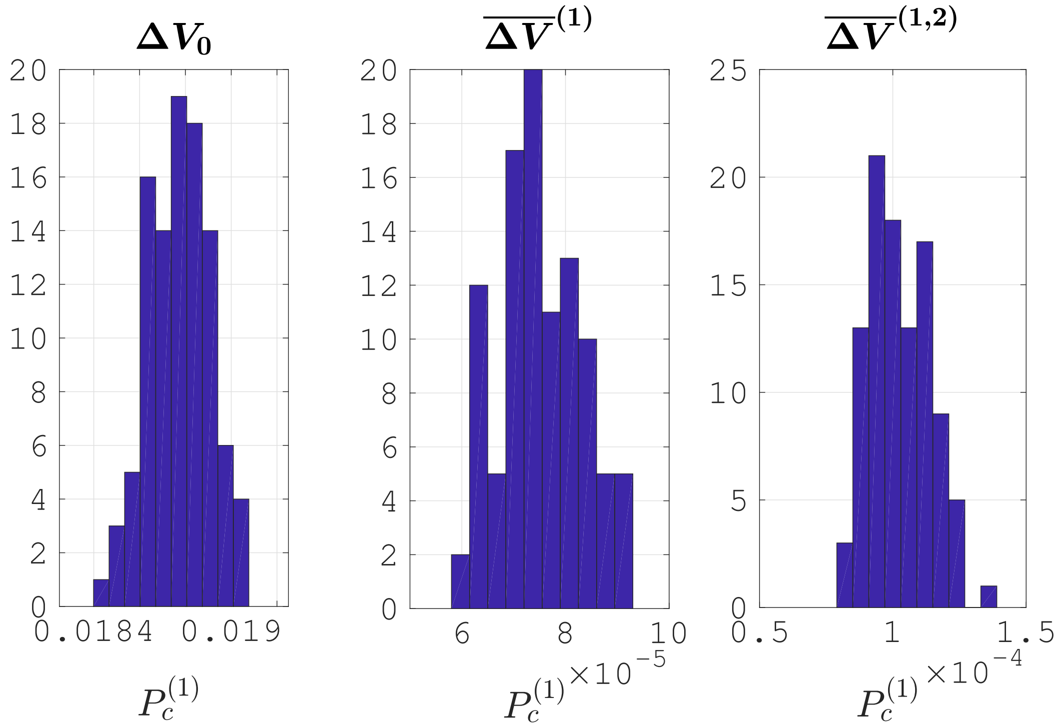


Figure 6.5: Histograms of 100 estimations of $P_c^{(1)}$ using surrogates and independent sets of random variables. Values for $P_c^{(2)}$ are omitted due to lack of variation.

surrogates $\hat{\boldsymbol{q}}_\alpha^{(\beta)}(\boldsymbol{\xi})$ that have been fitted to training data reliant only upon random inputs and not

Table 6.8: CP bounds for the MC results of Table 6.7

	$\boldsymbol{\Delta V}_0$	$\overline{\boldsymbol{\Delta V}}^{(1)}$	$\overline{\boldsymbol{\Delta V}}^{(1,2)}$
$P_c^{(1)}$	[1.81e-02, 1.98e-02]	[8.39e-05, 2.47e-04]	[6.92e-05, 2.22e-04]
$P_c^{(2)}$	[N/A, 3.68e-05]	[0.999, N/A]	[N/A, 3.68e-05]

design inputs. Due to the computation time needed to propagate 10^5 MC samples, it is noted that this sort of validation check is not typical under operational conditions. It does, however, provide a more complete evaluation of the surrogate accuracy. The quantitative results found in Table 6.9 show that both surrogates that include design space have reached an accuracy of at least 10 m in the estimation of the position QoIs. The surrogates $\hat{\boldsymbol{q}}_\alpha^{(\beta)}(\boldsymbol{\xi})$, trained on the post-optimization data that does not include design inputs, show a similar accuracy.

Table 6.9: Residual RMS of 10^5 SR- and MC-based validation samples for the nominal and optimized case

	$\hat{\boldsymbol{q}}_\alpha^{(\beta)}(\boldsymbol{\xi}, \boldsymbol{0})$	$\hat{\boldsymbol{q}}_\alpha^{(\beta)}(\boldsymbol{\xi}, \bar{\boldsymbol{\theta}}^{(1,2)})$	$\hat{\boldsymbol{q}}_\alpha^{(\beta)}(\boldsymbol{\xi})$	Units
$x_1^{(1)}$	9.14e-03	8.16e-03	6.19e-03	km
$y_1^{(1)}$	7.69e-03	7.76e-03	1.08e-03	km
$z_1^{(1)}$	3.13e-03	3.40e-03	6.13e-04	km
$x_2^{(1)}$	7.68e-03	6.62e-03	6.10e-03	km
$y_2^{(1)}$	1.46e-03	1.27e-03	1.53e-03	km
$z_2^{(1)}$	8.72e-04	8.30e-04	1.03e-03	km
$x_1^{(2)}$	1.04e-04	1.05e-04	1.35e-04	km
$y_1^{(2)}$	1.16e-04	1.02e-04	1.46e-04	km
$z_1^{(2)}$	3.98e-05	2.51e-05	3.04e-05	km
$x_3^{(2)}$	1.64e-04	1.57e-04	9.59e-05	km
$y_3^{(2)}$	1.69e-04	1.49e-04	9.47e-05	km
$z_3^{(2)}$	1.46e-04	1.32e-04	2.64e-05	km

6.8 Summary of Results

Building on the results of Chapter 5, the work contained within Chapter 6 seeks to apply the established UQ capabilities of SR to the field of conjunction assessment. Specifically, SR is used to quantify the probability of collision between objects in Earth orbit. This capability is then expanded to include design inputs, thus allowing the surrogate method of SR to be utlilized in an OUU problem. Although computationally prohibitive using a MC method, the OUU application of SR is both efficient and accurate. By considering variables that account for initial uncertainty as well as deterministic adjustments to a collision avoidance maneuver, a methodology is developed to optimally design an maneuver that includes state uncertainty. In this particular case, the change in maneuver is minimized while multiple probability of collisions are constrained.

Chapter 7

Multi-Element Applications

7.1 Test Case Introduction

In order to analyze the performance of the heuristic ME-SR algorithm, a series of test cases are chosen. First, we consider a test case from the ME-gPC literature. Wan et al. [120, 121] applies ME-gPC to a formulation of the Kraichnan-Orszag problem in order to test the ability of the multi-element method to account for disconinuities near the origin. Here, we apply the method of ME-SR. After this, a test case examines an Earth orbiting object. For this orbital scenario, a spacecraft is in a Molniya orbit with initial uncertainty, and its state is propagated for approximately 10.5 days, resulting in a posterior PDF that is difficult to approximate using SR due to large variances and, potentially, a bimodal nature.

In order to highlight the computational efficiency of SR with respect to input dimension, the number of directions within the Molniya case input space is then increased from $d = 6$ to $d = 10$. The former case utilizes a Keplerian element propagator, which results in posterior distributions with large variances. Due to the long propagation time and large standard deviations, unmodified SR has difficulty converging to an accurate solution. Therefore, this example stands to benefit from splitting the complex problem into smaller, more managable problems to approximate with ME-SR. The latter case of $d = 10$ results in bimodal distributions due to gravitational perturbations. Here, ME-SR is once again expected to improve in accuracy when compared to unmodified SR. Next, we apply ME-SR to an orbit case in the circular restricted three body problem (CRTBP) where the PDF of the spacecraft state is considered during a flyby. For all three cases, we compare the results of ME-SR to a single element SR surrogate, as well as a MC reference. From these comparisons, we are able to make conclusions about the performance of the ME-SR algorithm.

The authors note that the provided value of N for each test case is for the number of training samples used in each element, regardless of the value of $\Pr(I_{B_k} = 1)$. Therefore, the total number of samples used in a ME-SR surrogate is $N \times K$. Due to the current implementation of the ME-SR algorithm, however, more than $N \times K$ are utilized in the full training process of a ME-SR surrogate. This is due to the fact that the training data $\mathcal{D}$ is not preserved from one iteration of decomposition to the next. The properties of orthogonal polynomials displayed in Eq. (2.59) dictate that simply transferring inputs from one iteration to the next is mathematically incorrect. Here, the relationship between orthogonal polynomials and the probability density function of the related input direction is seen. As K increases, the bounds of the constructed orthogonal polynomials and associated inputs change. Therefore, the probability density functions are altered as well. As an example, consider the tail of a Gaussian distribution. The density function of an element bounded from 5σ to 6σ is relatively uniform when compared to randomly sampling a normal Gaussian density function until a desired number of samples within the bounds are collected. That is, samples pulled from the full Gaussian are more skewed towards the mean when compared to samples generated directly within the stated tail. This phenomenon is most prominent in areas of low probability (such as Gaussian tails), and a biased sample distribution results in inaccurate surrogate solutions. Reusing samples is possible by utilizing a methodology including weighted least squares, such as that found in [55]. Although more than $N \times K$ samples are needed for the implementation presented in this work, results include the theoretical sample count, as each surrogate is effectively trained on this number of samples. Achieving this theoretical cost is the subject of further research.

For the results presented in this section, an unmodified SR surrogate is constructed in order to compare with the accuracy of ME-SR. This single element SR is trained on the same total number of samples used in the comparitive ME-SR ALS process, i.e., $N \times K$ samples, in order to provide a level of "fairness" with regards to the training of the surrogates. For case results presented with multiple examples of K or N, e.g., some results are evaluated with $K = 4$ and some with $K = 8$, the number of training samples used for an unmodified SR surrogate is the total used for the comparitive ME-SR surrogate.

For the quantitative results presented in this section, a MC data set of $M_{MC} = 10^5$ samples (independent of surrogate training) is utilized as a reference comparison. As an indicator of individual sample accuracy, an RMS of relative error ϵ_{rel} is taken, where the data norm

$$\epsilon_{rel} = \frac{\|q(\boldsymbol{y}_j) - \hat{q}^K(\boldsymbol{y}_j)\|_D}{\|q(\boldsymbol{y}_j)\|_D} \tag{7.1}$$

is taken over the reference MC sample size of M_{MC}. Additionally, empirical values of the STD are taken over the output of each method and compared as a relative error, i.e.,

$$\sigma_{rel} = \frac{|\sigma_{MC} - \hat{\sigma}|}{\sigma_{MC}}. \tag{7.2}$$

In Eq. (7.2), the values of σ_{MC} and $\hat{\sigma}$ are the empirical standard deviations computed from data sets of the MC reference and surrogate output, respectively.

7.1.1 Kraichnan-Orszag Problem

Analyzed in previous ME-gPC literature, the Kraichnan-Orszag (KO) problem is known to be discontinuous near $y_1 = 0$, which unmodified PCEs fail to converge on [120, 121]. Although the KO problem can be formulated with one, two or three random inputs, this work focuses on the $d = 2$ case, as it provides a multidimensional case to exercise the ME-SR algorithm on, while remaining easier to illustrate than the $d = 3$ case. The KO ODE

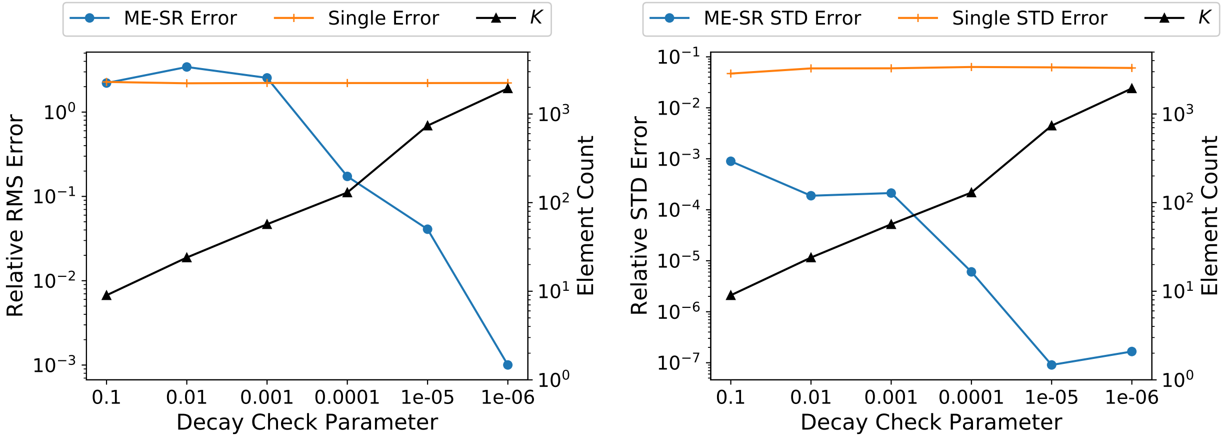
$$\begin{aligned} \frac{dY_1}{dt} &= Y_1 Y_3, \\ \frac{dY_2}{dt} &= -Y_2 Y_3, \\ \frac{dY_3}{dt} &= -Y_1^2 + Y_2^2, \end{aligned} \tag{7.3}$$

are utilized in the generation of test samples, with initial conditions

$$Y_1(0) = 1, \;\; Y_2(0) = 0.1y_1, \;\; \text{and} \;\; Y_3(0) = 0.1y_2. \tag{7.4}$$

Uncertainty is added to these initial conditions in the form of truncated Gaussian variables on the $[-6, 6]$ limits discussed in Section 4.3, and the resulting set of initial samples is propagated from $t = 0$ to $t = 10$. The QoI taken to be estimated is Y_2.

For this case, an exploration of the decay check parameter θ_1 is performed. As θ_1 approaches zero, it is expected that the ME-SR algorithm will produce more element splits than if $\theta_1 \approx 1$. Figure 7.1 illustrates the accuracy of the ME-SR surrogate and a single element SR as the value of the decay check parameter is altered. For this analysis, the values of $N = 250$, $P = 4$, $r = 3$, $\gamma = 0.5$, and $\theta_2 = 1.75$ are fixed, while a range of $\theta_1 = 10^{-1}, \ldots, 10^{-6}$ is used. Displaying values of ϵ_{rel} as relative error in Fig. 7.1(a) and relative standard deviation error in Fig. 7.1(b), the plots show the comparative inability of the unmodified SR to converge on an accurate result. However, the accuracy of the ME-SR surrogate improves as the value of θ_1 approaches zero, while the number of elements increases.



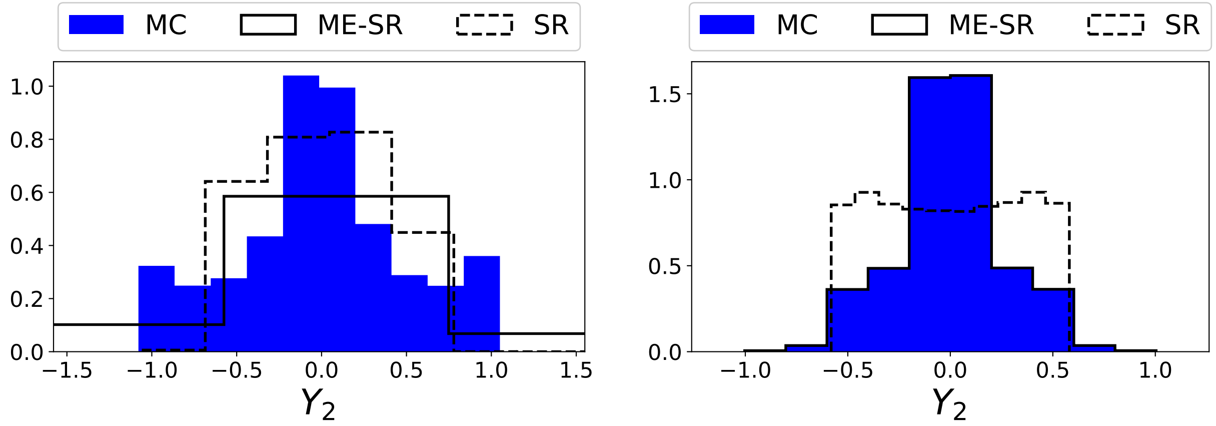
(a) Sample relative error RMS ϵ_{rel}

(b) Empirical STD relative error σ_{rel}

Figure 7.1: Plots of relative errors and element count for the K-O case as a function of θ_1.

Using ME-SR and SR surrogates from the decay check parameter analysis, the lower and upper limits of θ_1 are chosen to illustrate the fitness of the surrogates. Displaying the MC reference as a filled histogram, the solid line represents the fit of the ME-SR surrogate, while the dashed line is that of the single element. We note that the abscissa limits of the plots are truncated so that the

PDFs are easily discernible. It is shown that the ME-SR surrogate improves as θ_1 becomes smaller, and the methodology is able to accurately approximate the posterior distribution of Y_2. The single element SR, however, fails to converge even as the total number of training samples $N \times K$ becomes large.



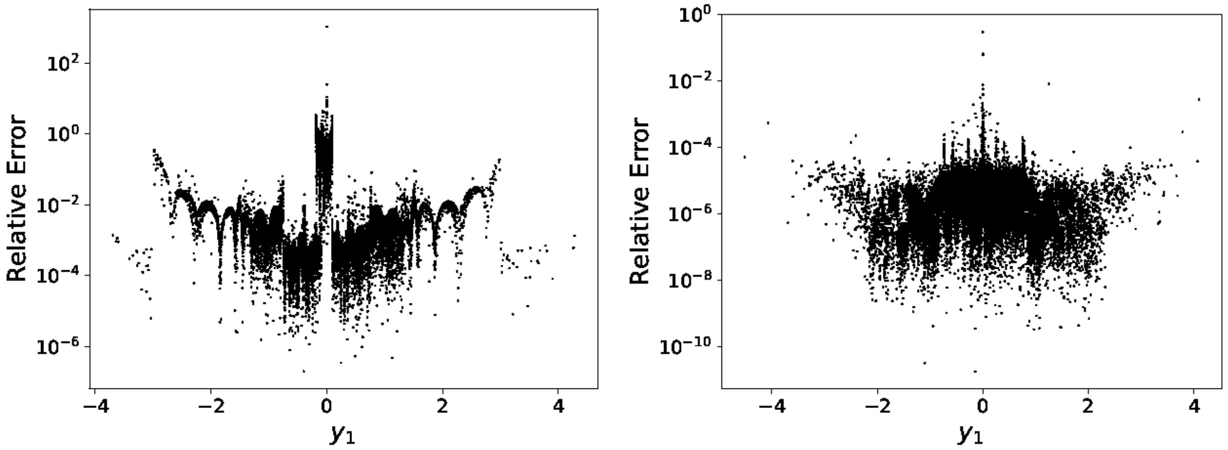
(a) $\theta_1 = 10^{-1}$

(b) $\theta_1 = 10^{-6}$

Figure 7.2: Normalized histograms of the MC reference, as well as the ME-SR and single SR surrogates for the KO problem.

In order to illustrate the error of the surrogates with respect to the input space, Figure 7.3 displays the values of ϵ_{rel} on the ordinate axis with the associated input of y_1 on the abscissa. The inputs and QoIs of the reference MC data are used for this analysis. In this figure, two plots are provided. The first plot contains results for the $\theta_1 = 10^{-2}$ example, while the second presents the $\theta_1 = 10^{-6}$ case. Here, $\theta_1 = 10^{-2}$ is chosen in order to investigate the relatively inaccurate ME-SR solution. As in Figure 7.1, the accuracy of the ME-SR surrogate improves as the magnitude of θ_1 decreases. The inaccuracy of the ME-SR solution in the $\theta_1 = 10^{-2}$ case is discernible near the origin of Figure 7.3(a). Here, a series of larger values of ϵ_{rel} are surrounded by samples of relatively low error. This behavior is responsible for the higher RMS error value seen in Figure 7.1(a), and it is reduced as θ_1 decreases in magnitude. This reduction is seen in Figure 7.3(b).

As visualization aids, the plots of Figure 7.4 depict two versions of the input space of the KO ME-SR surrogate. The first occurs when the decay check parameter is relatively large, producing



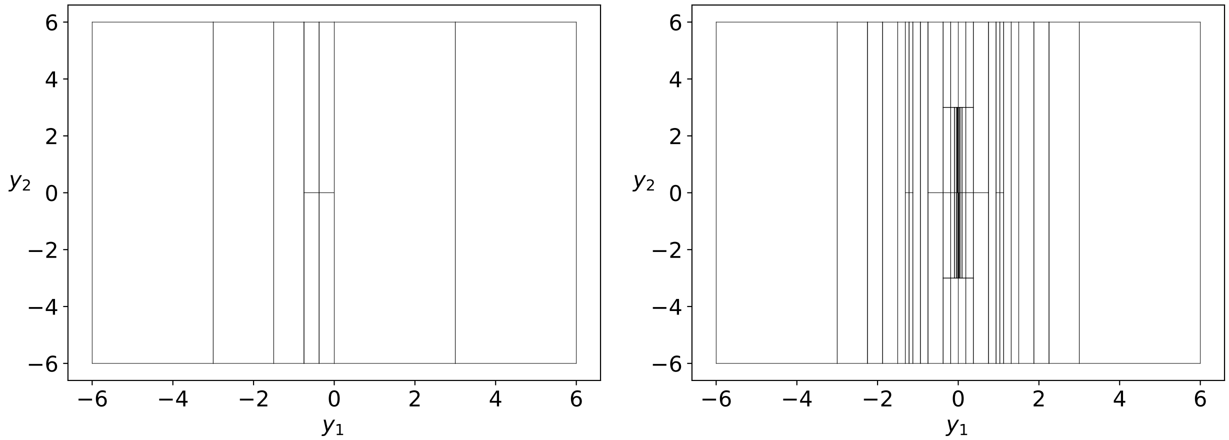
(a) $\theta_1 = 10^{-2}$ (b) $\theta_1 = 10^{-6}$

Figure 7.3: Plots of relative errors and associated inputs for the ME-SR KO case.

$K = 9$ elements, while the second choice of θ_1 is smaller and contains $K = 57$ elements. It is noted that the splitting of the input space appears to occur primarily around the origin, and, in particular, the splits are along the y_1 direction. As previously discussed, the KO problem is discontinuous near $y_1 = 0$. Therefore, the splitting behavior seen in Figure 7.4 supports the deliberate and adaptive behavior of the ME-SR algorithm. The choice of $\theta_1 = 10^{-3}$ for Figure 7.4(b), rather than a smaller magnitude value, is dictated by the number of splits created by smaller values in the analysis displayed in Figure 7.1. The large values of K create plots which contain indiscernible input bounds.

7.1.2 Molniya Orbit - $d = 6$

Previously examined in [67, 42, 117], a Molniya orbit case is considered as an introductory orbit test case for this chapter. First presented in [67], long propagation times and large initial uncertainties lead to a diffuse distribution of the true anomaly. The characteristics of this posterior PDF result in an un-converged PCE. Both [117] and [42] utilize different multi-element methodologies to improve on the performance of a single element PCE. The work of [117] leverages a Gaussian mixture approach with overlapping elements, while [42] follows the implementation



(a) Input space for $\theta_1 = 10^{-1}$

(b) Input space for $\theta_1 = 10^{-3}$

Figure 7.4: Plots of the decomposed input space for the KO case.

guidelines of [120, 121]. Therefore, it is expected for a single element SR surrogate to not converge to an accurate solution, while a ME-SR should improve on the accuracy of the surrogate approach. Although the posterior distributions of this example are not multi-modal, ME-SR proves to be effective due to its ability to pose the previously diffuse and complex estimation problem as a number of smaller and simpler problems.

Table 7.1 contains the initial conditions of the orbit as Keplerian elements, while Table 7.2 presents the initial conditions in Cartesian coordinates along with the standard deviations used as uncertainties in the inputs. The orbit is propagated for 868,105 s (∼10.5 days) using a simple two-body Keplerian propagator.

Table 7.1: Initial conditions in orbital elements of the Molniya orbit

Orbital Element	Mean
semi-major axis	26562 km
eccentricity	0.741
inclination	63.4°
argument of perigee	90°
RAAN	-90°
true anomaly	0°

Table 7.2: Molniya orbit initial means as Cartesian elements with associated random input standard deviations

Input	Mean	STD
x	3.08038e06	10 m
y	0.0	10 m
z	-6.15138e06	10 m
$\dot{x}$	0.0	1 m/s
$\dot{y}$	1.004e04	1 m/s
$\dot{z}$	0.0	1 m/s

Using $r = 3$, $P = 4$, $\theta_1 = 10^{-8}$, $\gamma = 0.5$, and $\theta_2 = 1.25$, the ME-SR algorithm is utilized to construct surrogates for each Cartesian QoI. In [68], the variance of the posterior PDF is considered most sensitive to the input associated with the $\dot{y}$ direction. Therefore, it is expected for the ME-SR algorithm to prioritize splitting this direction. When run using the previously mentioned parameters, only this direction is split via the adaptive algorithm. The number of elements required via the choice of θ_1 varies among each QoI, and, additionally, varies in relation to the number of training samples used.

Figures 7.5, 7.6, and 7.7 illustrate an exploration of this behavior. By increasing the number of training samples per element N, an ME-SR is constructed based on the criteria imposed by θ_1. Therefore, these figures depict information gathered via the K-type convergence of ME-SR as a function of N. For each QoI presented, the top plot contains relative error ϵ_{rel} for both ME-SR as well as an unmodified SR. As previously stated, the unmodified SR utilizes all $N \times K$ training samples of the comparable ME-SR in order to maintain consistency in blackbox computation cost. Below each error plot, there is a graph containing the number of elements split K (represented by the triangle markers and solid line) for the ME-SR, as well as the total number of training samples required, i.e., $N \times K$ (depicted as the plus-sign markers and a dashed line).

Figures 7.5, 7.6, and 7.7 highlight the ability of ME-SR to converge more rapidly to an accurate solution when compared to an unmodified SR. Indeed, the accuracy of the SR solution does not improve markedly as the total number of training samples increases. Each figure starts

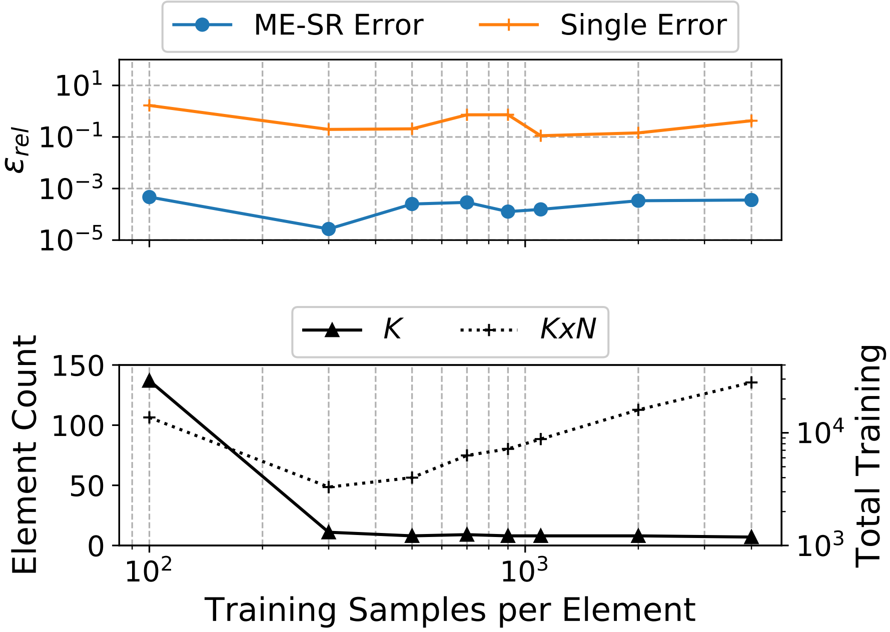


Figure 7.5: Plots of error and sample count for the x-position QoI of the Molniya $d = 6$ case.

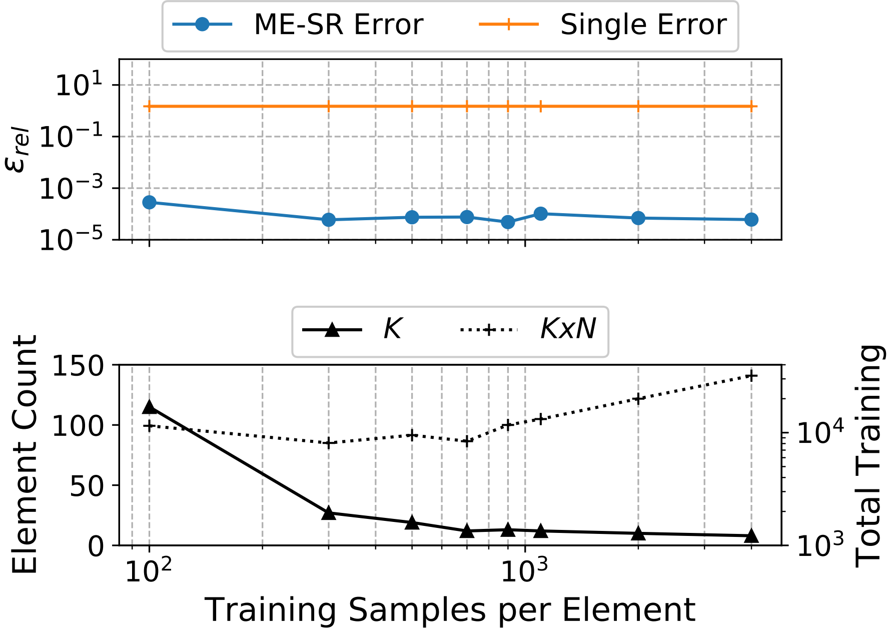


Figure 7.6: Plots of error and sample count for the y-position QoI of the Molniya $d = 6$ case.

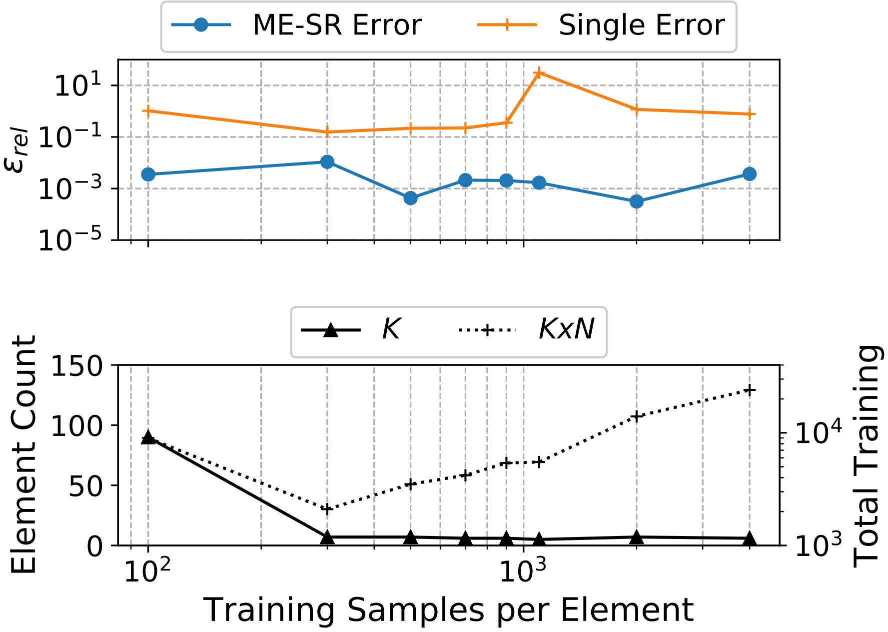


Figure 7.7: Plots of error and sample count for the z-position QoI of the Molniya $d = 6$ case.

with $N = 100$ and increases this value until $N = 4000$. At low values of N, it can be seen that the ME-SR algorithm creates many elements in order to converge in a K-sense. Figure 7.5 illustrates this behavior and the resulting large number of total training samples needed at $N = 100$. However, as the number of training samples increases, the number of elements created by the algorithm begins to decrease. Around $N = 300$, the x-position ME-SR surrogate requires significantly fewer elements than at $N = 100$, and the total number of samples is ultimately lower as well. This toal number of required samples $N \times K$ eventually increases. Although K slowly decreases after the initial drop from $N = 100$ to $N = 500$, the increasing value of N creates an overall trend of increasing the total number of training samples. An interesting behavior seen in these plots is that although the number of training samples per element N as well as the total number of elements K vary, the accuracy of the ME-SR surrogate does not fluctuate signficantly. Therefore, the accuracy of the ME-SR surrogate may be more sensitive to other parameters such as P, r, or the decay check parameter θ_1.

It is noted that the QoIs of the velocity components are left out in order to provide conciseness to this results section. The same behavior is displayed for the velocity QoIs as in the position components, i.e., an initial large K at $N = 100$ followed by an overall decline in K as N increases. Although not displayed in a figure due to required computation times, the number of elements K for each choice of N varies as a function of the training data set. That is, due to the stochastic nature of the generation of training data, it is possible to experience a difference in the fit from one independent construction of ME-SR to the next. Therefore, validation techniques used for statistical model fitting, such as cross- or k-fold validation are recommended. Here, k denotes the number of equal sized training subsamples rather than the index for input space decomposition.

Figure 7.8 provides an illustration of the qualitative improvement in estimating the posterior distribution of the Molniya x-direction QoI with $N = 300$ training samples per element as compared to a single element SR. The shaded region of the histogram represents the bins of the MC reference, while the solid line (which closely follows this region) demarks the bins of the ME-SR posterior PDF histogram. The dashed line depicts the histogram of the unmodified SR surrogate, and its performance is readily compared to that of the ME-SR result. For this particular result, the algorithm converged to $K = 11$ elements. Therefore, a total of 3300 training samples are used.

Table 7.3 provides a quantitative analysis of the ME-SR approximation for the Molniya case. Here, the performance of the unmodified surrogate approach is compared to that of the ME-SR by utilizing relative differences in sample and empirical STD estimation. As previously discussed, the reference solution is that of a 10^5 MC sample set. With respect to the single element SR results, significant digits of accuracy are gained when the ME-SR algorithm splits the random input space. In addition to Table 7.3, Table 7.4 is provided as an aid for quantifying the computation cost of the $N = 300$ ME-SR surrogates by counting the total blackbox propagations.

As a brief exploration into the solution space of a QoI in this test case, the z-position component is considered. For the chosen parameters of this case, the ME-SR fit, while improved in relation to the unmodified SR, gains approximately a digit of precision over the single element solution. Because increasing the amount of training data does not significantly improve the accuracy of

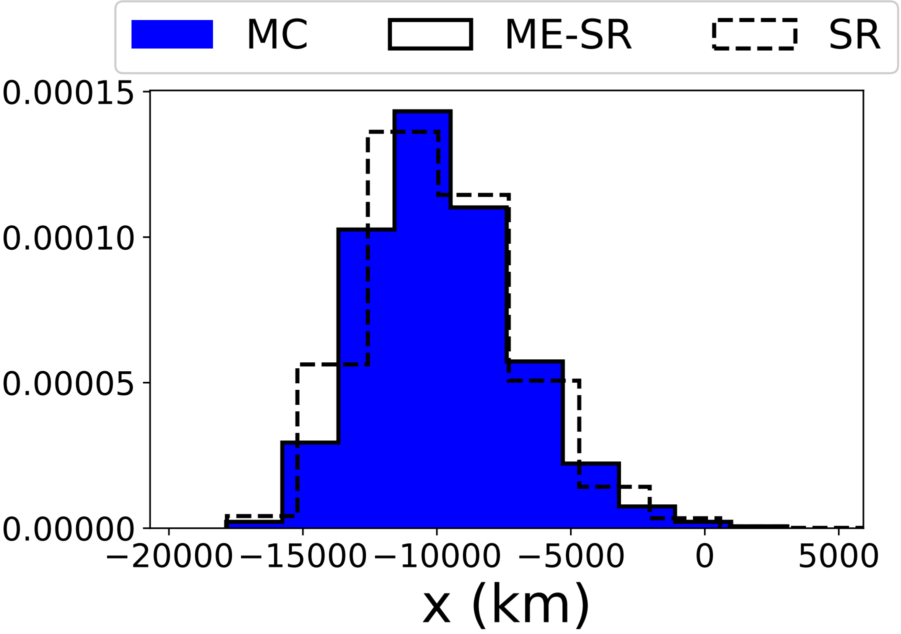


Figure 7.8: Normalized histograms of the x-position Molniya QoI with the MC reference, unmodified SR, and improved ME-SR results

Table 7.3: Performance analysis of $N = 300$ surrogates for the QoIs of the $d = 6$ Molniya problem

Value	Method	x	y	z	$\dot{x}$	$\dot{y}$	$\dot{z}$
σ_{rel}	SR	2.1e-03	1.2e-02	1.8e-03	2.1e-02	4.7e-02	1.8e-02
	ME-SR	2.7e-06	4.5e-05	1.4e-05	1.2e-03	1.8e-05	9.9e-06
σ	Reference MC	2840.2 km	908.5 km	5671.9 km	0.30577 km/s	0.75150 km/s	0.610610 km/s
ϵ_{rel}	SR	1.9e-01	1.4	1.5e-01	1.4e-01	2.9	1.3e-01
	ME-SR	2.7e-05	5.9e-05	1.0e-02	2.7e-03	8.9e-02	4.3e-04
RMS(·)	Reference MC	10199.5 km	17251.6 km	20368.1 km	1.6569 km/s	0.75648 km/s	3.3089 km/s

Table 7.4: Element and total sample count for $N = 300$ ME-SR of the $d = 6$ Molniya QoIs

	x	y	z	$\dot{x}$	$\dot{y}$	$\dot{z}$
K	11	27	7	13	10	14
$N \times K$	3300	8100	2100	3900	3000	4200

the ME-SR surrogate, the decay check parameter is tightened from $\theta_1 = 10^{-8}$ to 10^{-11}. Additionally, the maximum order is lowered to $P = 3$ in order to mitigate potential overfitting. With this change, the relative RMS error of ϵ_{rel} drops to $3.9e-05$. The change in these two parameters is followed by a growth in K. That is, given the surrogate parameters and θ_1, the ME-SR algorithm creates 20 splits instead of 7. Therefore, if greater accuracy is desired for the z QoI, a smaller value of the decay check parameter is encouraged. The behavior of increasing accuracy with decreasing magnitude of θ_1 is seen in the analysis of the KO problem in Section 7.1.1. Although not included for the sake of brevity, applying the same procedure to the other QoIs results in an increase in surrogate accuracy. For example, the error of the ME-SR fit of the x-velocity component decreases from $2.7e-03$ to $5.4e-05$.

7.1.3 Molniya Orbit - $d = 10$

The previously presented Molniya case is extended to include four more random inputs. These uncertain directions take the form of the gravitational parameter of the earth μ, the area to mass (A/M) ratio of the modeled spacecraft, an uncertain coefficient of drag C_d, and the coefficient of reflectivity C_r. Rather than a Keplerian element propagator, the initial conditions are run using a more complex numerical propagator which contains a dynamic model of the atmosphere, solar radiation pressure effects (SRP) and a 10×10 spherical harmonics gravity model, as determined by the GRACE GGM02C gravity model [111]. It is noted that for both drag and SRP, a simple canonball model is utilized. Table 7.5 contains the initial means and uncertainties (in the form of STD) for the additional parameters of this case.

Theoretically, a PCE for a $d = 6$ and $p = 4$ case would require $N = 210$ training samples, while a $d = 10$ and $p = 4$ requires $N = 1001$ samples for the training data set. This relationship is dictated via the equation

$$\Lambda_{PC} = \frac{(P_{PC} + d)!}{P_{PC}!\, d!}, \tag{7.5}$$

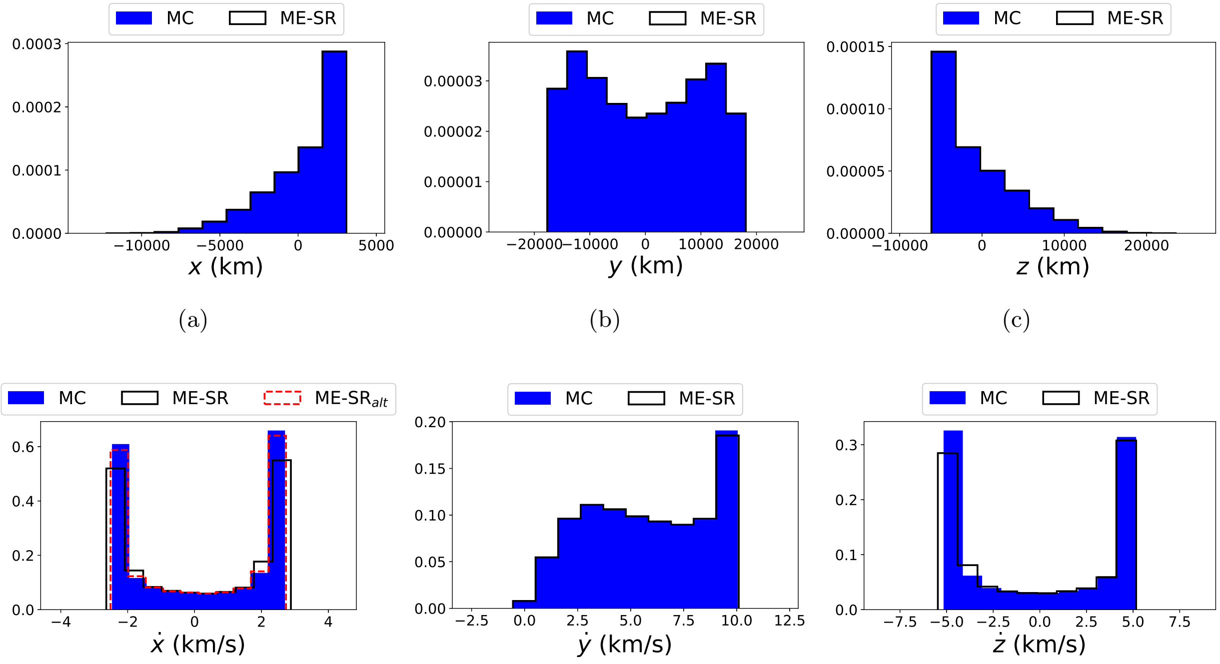
where P_{PC} is the total order of the PCE expansion, and Λ_{PC} are the number of terms required to solve for. This exponential in the number of required training samples with respect to the input dimension is indicative of the *curse of dimensionality.* Due to the ALS format of the solving for SR, this rapid increase in required training sample count is avoidable.

Table 7.5: Additional initial means and standard deviations for uncertain parameters of the $d = 10$ Molniya case

	Mean	STD
$\mu_{GP}(\text{km}^3/\text{s}^2)$	3.986e05	10e-03
$A/m(\text{m}^2/\text{kg})$	9.09e-02	1.0e-02
C_d	2.2	5e-02
C_r	1.2	5e-02

Using $N = 500$, $P = 4$, $\theta_1 = 10^{-8}$, $\gamma = 0.5$, and $\theta_2 = 1.25$, an ME-SR surrogate is constructed for each QoI in the $d = 10$ molniya case. A qualitative analysis of ME-SR is included for each QoI in Figure 7.9. Here, we see the highly non-Gaussian PDFs of the position and velocity components, and the fit via ME-SR (solid line). The authors note that single-element SR results are excluded from these plots due to a lack of convergence. In the case of the position QoIs and $\dot{y}$, the multi-modal distributions are represented well by the ME-SR surrogates. This multi-modal behavior, as opposed to that from the $d = 6$ case, is due to the gravitational perturbations, as well as the more complex gravity model, creating variations in the mean motion of the propagated states. In order to quantify this accuracy, Table 7.6 presents relative error for both individual samples and STD as in the previous case. In this table, multiple digits of precision are gained when comparing ME-SR to SR. In particular, the STD estimate is improved by approximately four digits. As in the previous $d = 6$ Molniya case, Table 7.7 includes the number of elements and total training surrogates required by the chosen ME-SR parameters.

As in the previous case, we continue to explore the fit of one of the QoIs. Specifically, the $\dot{x}$ QoI is considered. The histogram of the PDF in question is displayed in Figure 7.9(d). Here, a relative RMS error of 0.24 is achieved via ME-SR. Searching for a more accurate solution, two approaches are taken: increasing N or using a smaller value of θ_1. For the former, the number of training samples is doubled to $N = 1000$, and the rest of the surrogate parameters are kept the same. Doubling the number of training samples results in a similar relative RMS error, i.e., $\epsilon_{rel} = 0.24$, suggesting that the accuracy is insensitive to increases in training data. This relationship is broadly seen in Figures 7.5, 7.6 and 7.7. On the other hand, changing θ_1 from 10^{-8} to 10^{-11} and P from 4 to 3 (in order to mitigate overfitting) increases the accuracy to $2.9e-03$. This improved fit can be seen as the dashed red line in Figure 7.9(d) where it is labeled as ME-SR$_{alt}$. Therefore, it is again concluded that an operator, seeking accuracies greater than those found in Table 7.6, is recommended to explore altering the decay check parameter and order of an ME-SR surrogate.



(a) (b) (c)

(d) (e) (f)

Figure 7.9: Normalized histograms of the $d = 10$ Molniya case with the MC reference, unmodified SR, and improved ME-SR results

Table 7.6: Performance analysis of $N = 500$ surrogates for the QoIs of the $d = 10$ Molniya problem

Value	Method	x	y	z	$\dot{x}$	$\dot{y}$	$\dot{z}$
σ_{rel}	SR	1.7	8.8e-01	3.2e-01	1.0	3.6	1.8e-01
	ME-SR	2.5e-06	1.1e-06	1.5e-06	8.6e-05	1.1e-05	1.1e-05
σ	Reference MC	2449.1 km	10487.3 km	4851.9 km	2.132 km/s	2.853 km/s	4.266 km/s
ϵ_{rel}	SR	34.7	11.3	53.8	6.9	246.0	14.7
	ME-SR	6.6e-03	4.9e-03	2.2e-02	2.4e-01	6.1e-03	3.1e-03
RMS($\cdot$)	Reference MC	2497.7 km	10489.2 km	4949.7 km	2.141 km/s	6.486 km/s	4.267 km/s

Table 7.7: Element and total sample count for $N = 500$ ME-SR of the $d = 10$ Molniya QoIs

	x	y	z	$\dot{x}$	$\dot{y}$	$\dot{z}$
K	12	11	10	10	12	14
$N \times K$	6000	5500	5000	5000	6000	7000

7.1.4 Europa-Jupiter CRTBP Case

Many of the unstable resonant orbits that are of particular interest for tour design are very unstable and have close flybys of the secondary. One such orbit that is used in this work is the 5:6 resonant orbit shown in Figure 7.10. The circular restricted three-body problem (CRTBP) is the model used in this study for the flyby orbit. See Szebehely [108] for a detailed discussion of the CRTBP, but a brief overview of the problem will be given here. In this model, two bodies orbit in circular orbits around their barycenter, and the motion of a third infinitesimal mass (the spacecraft) is modeled. Using this formulation, the x axis of a rotating coordinate system is chosen so that it is aligned with the two bodies with the origin at the barycenter. The relevant quantities are typically made dimensionless so that the larger body (the primary) has mass $1 - \mu$, while the smaller body (the secondary) has mass μ. The distance between the two bodies, the mean motion, and the gravitational constant are all one in this model, while the period of the rotating system is 2π. The primary is located on the x axis at $x_1 = -\mu$ and the secondary is located at $x_2 = 1 - \mu$. Given this model, the equations of motion for the spacecraft in the rotating frame are

$$\begin{aligned} \ddot{x} - 2\dot{y} &= \frac{\partial \Omega}{\partial x} \\ \ddot{y} + 2\dot{x} &= \frac{\partial \Omega}{\partial y} \\ \ddot{z} &= \frac{\partial \Omega}{\partial z} \end{aligned} \tag{7.6}$$

where

$$\Omega = \frac{x^2 + y^2}{2} + \frac{(1 - \mu)}{r_1} + \frac{\mu}{r_2} \tag{7.7}$$

and

$$r_1 = \sqrt{(x+\mu)^2 + y^2 + z^2},\ r_2 \quad = \sqrt{(x-1+\mu)^2 + y^2 + z^2}. \tag{7.8}$$

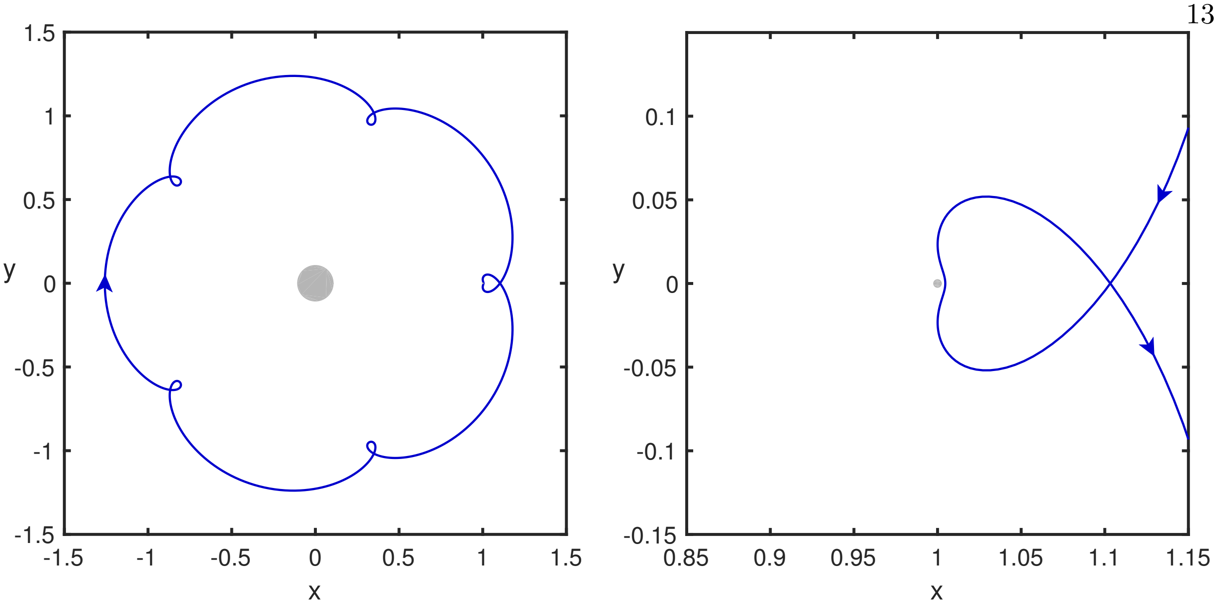
This orbit has a period of approximately 38.2 dimensionless units and a Jacobi constant of C = 2.99736448718112. The close approach results in a highly non-linear propagation of uncertainty, which is of interest when navigating these types of trajectories. An illustration of the orbit examined in this case is found in Figure 7.10. For the case found in this chapter, an uncertain initial condition of an orbiting spacecraft is propagated to various times near the nominal flyby time t_f. Utilizing a uniformly distributed random input y_5, QoIs are propagated from half a period away from the nominal until ±10% of this nominal flyby time. Here, y_5 functions as a *deterministic variable*. That is, although y_5 is randomly sampled for surrogate training, we purposefully select it as an input in the constructed surrogate, thus allowing for the PDF of the spacecraft at a desired time. The mapping of y_5 to the time of propagation is done via

$$t(y_5) = \frac{t_{max} - t_{min}}{2}(y_5 + 1) + t_{min}, \tag{7.9}$$

where t_{max} and t_{min} are the respective lower and upper bounds of our flyby "window" and $t(y_5)$ is the chosen time for the spacecraft.

Therefore, $d = 5$, and the surrogate is able to produce posterior distributions of the QoIs at any time along this flyby "window". This is done post surrogate construction by deterministically choosing y_5 and leaving all other inputs to be randomly selected. This approach of constructing a surrogate, such that one of the inputs is deterministic, is not unlike a previous application of surrogate methods to optimization under uncertainty for conjunction assessment. For further details on this methodology, the works of [41, 15] are recommended. The initial conditions of this orbit and the uncertainty (or bounds) of the inputs are given in Table 7.8.

As in the K-O example, the value of θ_1 is varied in the exploration of this case. Specifically, a range of $\theta_1 = 10^{-1}, \ldots, 10^{-9}$ is used. While the decay check parameter changes for the construction of each surrogate, the values of $N = 250$, $P = 4$, $r = 3$, $\gamma = 0.5$, and $\theta_2 = 1.25$ are fixed. Figure 7.11 illustrates the evolution of ϵ_{rel} as well as σ_{rel} over the range of θ_1 for the $\dot{x}$ QoI. Predictably, we



(a) Overview

(b) Close View Near Secondary

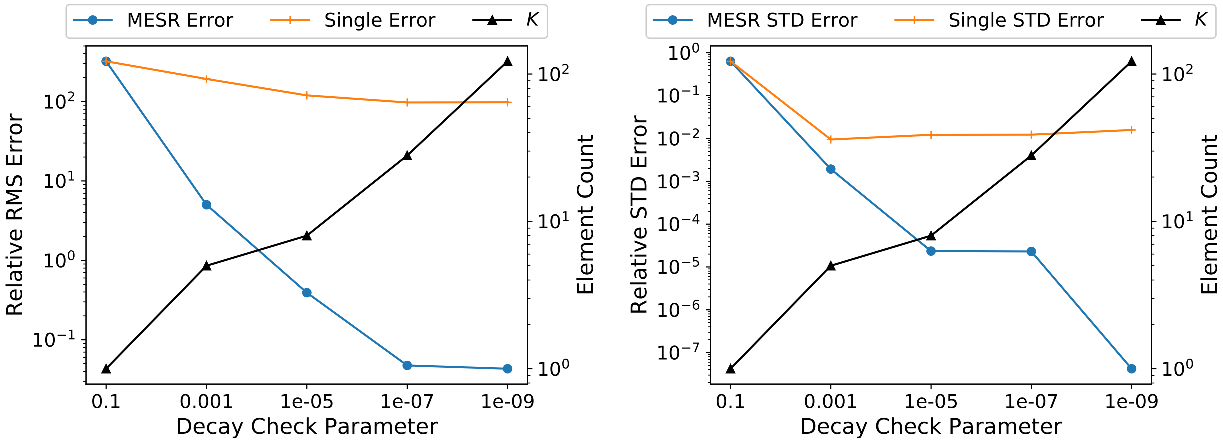
Figure 7.10: 5:6 resonant orbit in the Jupiter-Europa rotating frame.

Table 7.8: CRTBP orbit initial conditions with associated random input standard deviations and bounds

	Mean	STD or Bounds
x	1.00462	10e-05
y	0.0	10e-05
$\dot{x}$	0.0	10e-07
$\dot{y}$	0.11599	10e-07
Flyby Window	N/A	(15.27810,22.91715)

see the number of elements K increase as θ_1 approaches zero. Additionally, as more elements are split, there is an overall trend of accuracy improvement with ME-SR. However, as seen before, the accuracy of the unmodified SR does not significantly improve even as the total number of training samples $N \times K$ increases.

Regarding the splitting of the sample space, it is noted that the algorithm emphasizes decomposing the y_5 direction. This choice is intuitive, as the object state distribution is sensitive to the choice of propagation time. Figure 7.12 displays the splitting of this direction. Unlike the K-O example, the ME-SR algorithm split only one dimension, precluding the need for a "mesh-like"



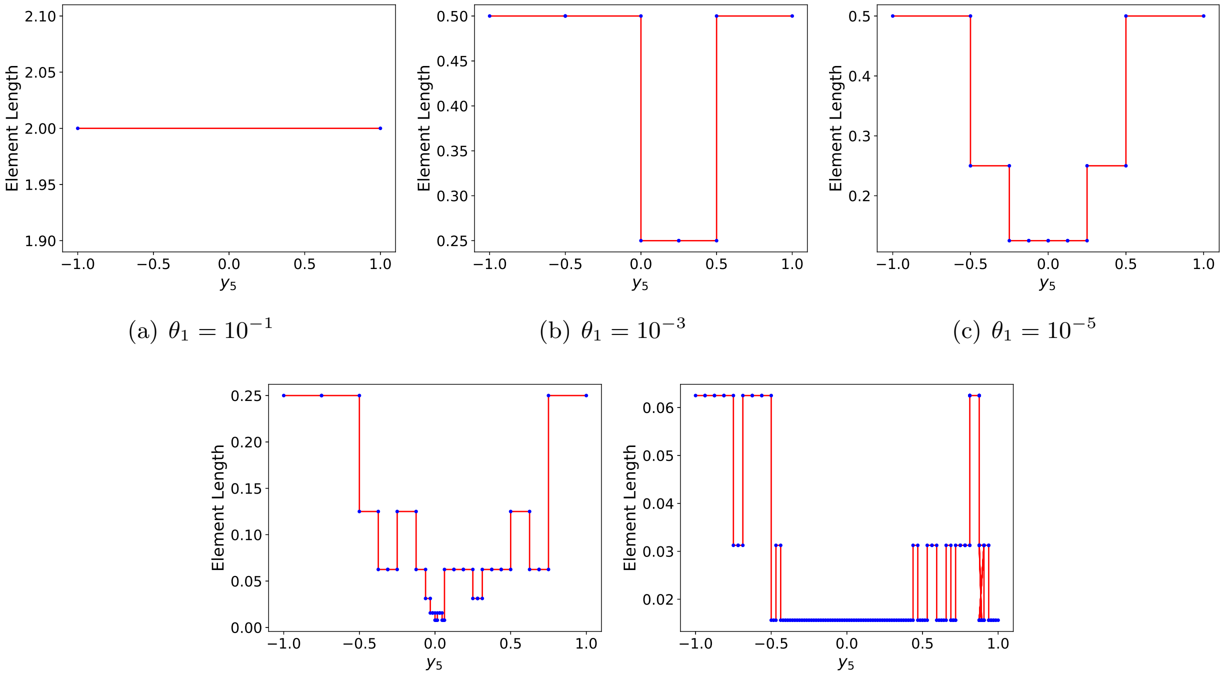
(a) Sample relative error RMS ϵ_{rel}

(b) Empirical STD relative error σ_{rel}

Figure 7.11: Plots of relative errors and element count for the $\dot{x}$ QoI of the CRTBP case as a function of θ_1.

illustration. Rather, the value of y_5 is found on the abcissa, while the ordinate axis represents the overall length of the element, i.e., the distance between bounds of element k. Figure 7.12(a) shows the space of this direction as a single element. As the magnitude of the exponent of θ_1 grows, increased refinement of the input space is seen. Figure 7.12(b) illustrates an input space that has been split into five nearly equal length pieces. However, Figure 7.12(c) shows that, by $\theta_1 = 10^{-5}$, the splits are clustered around $y_5 = 0$ or $t = t_f$. A qualitative assessment of this behavior notes that at $y_5 = 0$ the object is passing by Europa at the nominal time of closest approach. This behavior remains largely intact in Figure 7.12(d). Finally, Figure 7.12(e) presents a highly refined input space (indicated by the ordinate axis) that exhibits relatively little priority to particular locations on the bounds of y_5.

For the purposes of a qualitative analysis, normalized histograms are once again provided in Figure 7.13. As in previous examples, the MC reference is displayed as a solid histogram, the solid line represents the fit of the ME-SR surrogate, and the dashed line is that of the single element. Unlike in previous analyses however, the results displayed in the histograms contain a *deterministic input*. Therefore, these are not the PDFs of $\dot{x}$ at a single time but over all times that are originally sampled in the reference data set. Although the single-element SR fit improves marginally from



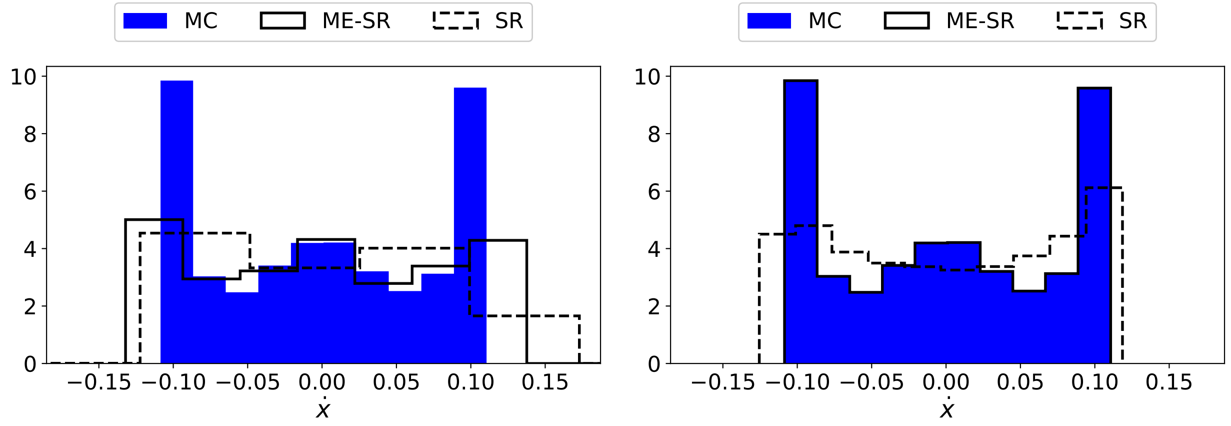
(a) $\theta_1 = 10^{-1}$ (b) $\theta_1 = 10^{-3}$ (c) $\theta_1 = 10^{-5}$

(d) $\theta_1 = 10^{-7}$ (e) $\theta_1 = 10^{-9}$

Figure 7.12: Plots of sample space for the y_5 input and $\dot{x}$ QoI.

Figure 7.13(a) to Figure 7.13(b), the ME-SR fit of the latter case is almost indistinguishable with the MC reference when compared to the dashed line of the single element surrogate. In this case, $\theta_1 = 10^{-3}$ and $\theta_1 = 10^{-9}$ are chosen to represented in Figure 7.13 due to the lack of spltting at $\theta_1 = 10^{-1}$ as seen in Figure 7.11. In addition to this visual comparison, Table 7.9 provides a quantitative analysis for the results of both SR and ME-SR under the $\theta_1 = 10^{-9}$ conditions. These results are promising for future applications, such as planetary protection schemes which seek to quantify the probability of colliding with a planet or moon at multiple times during a flyby.

To illustrate the behavior of a surrogate with a deterministic input, Figure 7.14 displays evaluations of $\hat{q}^K(\boldsymbol{y})$ using 101 equally spaced values of y_5 spanning the bounds $[-1, 1]$. For this figure, fixed values of all other inputs are used, thereby quantifying one possible trajectory, i.e, an analysis using a single realized value of $\boldsymbol{y}$. Here, the orbit seen in Figure 7.10(b) can be compared. It is noted that due to the continuous uniform distribution of y_5 on which the ME-SR surrogate is trained, any possible value of time bounded by the initial condition is a valid input, not just



(a) $\theta_1 = 10^{-3}$

(b) $\theta_1 = 10^{-9}$

Figure 7.13: Normalized histograms of the reference MC, unmodified SR and ME-SR PDFs for the $\dot{x}$ QoI of the CRTBP case

Table 7.9: Performance analysis of $\theta_1 = 10^{-9}$ surrogate methods for the QoIs of the CRTBP case

Value	Method	x	y	$\dot{x}$	$\dot{y}$
σ_{rel}	SR	1.7e-03	5.1e-04	1.5e-02	2.1e-03
	ME-SR	5.9e-07	1.6e-06	1.6e-07	5.6e-07
σ	Reference MC	3.88e-02	3.60e-02	7.66e-02	8.27e-02
ϵ_{rel}	SR	2.5e-03	4.6e-01	99.2	1.6
	ME-SR	1.1e-05	9.6e-03	1.7e-02	1.1e-03
RMS($\cdot$)	Reference MC	1.04	3.60e-02	7.66e-02	8.43e-02

the 101 chosen here. The results presented in Table 7.9 depict a promising methodology for an application such as planetary protection, and the ability for an ME-SR to create Figure 7.14 is one such illustration of the ME-SR surrogate's capability. However, future work is needed in order to accurately estimate a quantity such as the probability of collision with Europa.

7.2 Summary of Results

Chapter 7 applies the formulation of ME-SR presented in Chapter 4 to a series of test cases in order to evaluate the behavior and efficacy of the methodology. For this, a 2D iteration of the

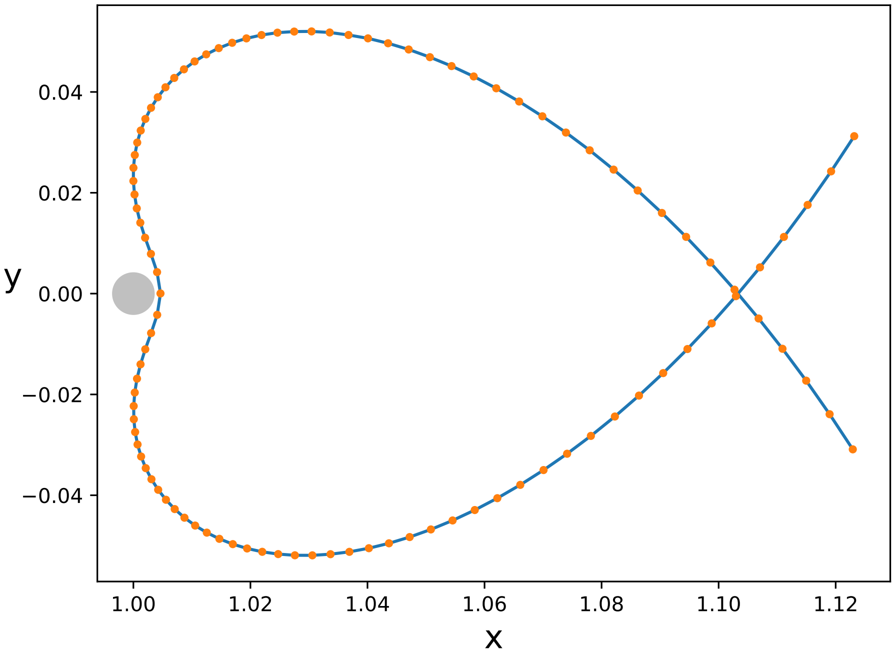


Figure 7.14: ME-SR estimation of one possible trajectory in the CRTBP case

Kraichnan-Orzsag (KO) problem is examined. A need to split the input space in two directions is seen, while the convergence properties of ME-SR, with respect to the decay check parameter, are analyzed and plotted. The accuracy of ME-SR in this case compares favorably to an attempt to construct an unmodified SR solution. Following this, ME-SR is applied to a Molniya case of dimension 6 and 10. Due to the fact that SR does not suffer from the curse of dimensionality, ME-SR is capable of fitting to the 10 dimension case without exponentially increasing in cost when compared to the 6 dimension version of the orbit propagation. For the Molniya case, the convergence of ME-SR with respect to the size of training samples is considered. This analysis shows that increasing the amount of training data does not necessarily improve the surrogate accuracy, although the number of element splits may decrease. Leveraging the conclusions of both the KO and Molniya convergence analyses, the accuracy of one of the Molniya ME-SR surrogates is improved by reducing the magnitude of the decay check parameter. Lastly, ME-SR surrogates are created for a CRTBP orbit. For this example, a flyby is considered over a bounded timespan, and the state of the spacecraft is estimated using ME-SR. The algorithm is shown to create splits primarily in the direction of the deterministic input associated with the flyby time, and the accuracy of the approximation improves as more splits are made in this direction. In all three test cases, ME-SR affords multiple orders of magnitude improvements in the accuracy of STD and sample estimation when compared to SR.

Chapter 8

Conclusions and Future Work

8.1 Summary of Contributions

The work presented in this thesis began with the history and background of UQ in the field of astrodynamics (Chapter 1), covered various UQ methods in use and currently being researched while presenting the method of SR (Chapter 2), discussed OUU and the formulation needed to apply SR (Chapter 3), and introduced a multi-element methodology for SR that relies on the decomposition of input space (Chapter 4). This body of work provided the foundation necessary for the results contained within the subsequent chapters. Chapter 5 explored three basic test cases in LEO so that the accuracy, convergence properties, computation cost and sensitivity analysis performance could be analyzed. After confirming the relative efficiency and accuracy of SR, the work contained in Chapter 6 applied SR to OUU. This optimization problem introduced a novel method of quantifying the probability of collision while designing an avoidance maneuver. Lastly, Chapter 7 illustrated the effectiveness of ME-SR to adaptively fit to a solution that is discontinuous with respect to the input space.

8.2 Conclusions

Altogether, the content of this thesis presents a surrogate method that is novel to the field of astrodynamics, yet is shown to be tractable, accurate and adaptable to complex dynamical systems. As identified in Chapters 1 and 2, there is a need in the astrodynamics community for further research in UQ due to cases of unidentified collisions [2, 71] and increasing numbers of objects in space [83]. Previous work on the surrogate method of PCEs has shown that surrogate methods can work in the field of astrodynamics [67, 41, 100]. It was therefore hypothesized that the promising surrogate method of SR would be found to be applicable as well. After presenting the formulation and implementation algorithm in Chapter 2, the SR methodology was shown to perform well in a variety of test cases. Although no method is universally applicable to UQ in astrodynamics, SR is now known to provide valuable analyses in LEO, especially when considering conjunction assessment with no posterior assumptions. Future applications are possible, and the following section discusses some possibilities.

With regards to quantitative conclusions, SR was shown to accurately reproduce results to four or more digits of precision with respect to a MC reference. Considering a case in LEO, uncertainties of 1 km/s and 1 m/s in position and velocity, respectively, were propagated for 36 hours to produce noticeably non-Gaussian posterior distribution, thereby necessitating a need for a method that does not assume Gaussian posteriors. By increasing the input dimension from $d = 6$ to $d = 20$, the computational tractably of SR was tested. Since PCEs suffer from the curse of dimensionality, the theoretical computation cost for this example would increase from $N = 210$ to $N = 10,626$ training samples. Due to SR's linear relationship between required training samples and input dimension, the resulting training samples count was $N = 150$ and $N = 750$ for the $d = 6$ to $d = 20$ cases, respectively. Leveraging the computational efficiency of SR to compute large data sets, an SR based ANOVA sensitivity analysis shed additional light on a test case previously examined in the literature [62]. With this analysis, it was determined that the equinoctial system in question had a largely independent relationship between inputs and QoIs. This discovery highlights the benefits made possible by an efficient means of UQ such as SR.

In addition to proving itself as a valuable UQ tool, SR has also been shown to work well in conjunction with optimization algorithms that consider uncertainty. Along with the succesful multi-element formulation of SR (ME-SR), the surrogate method is proving to follow relatively closely in the footsteps of PCEs. That is not to say that the methodologies are interchangeable. However, it does highlight that certain surrogate approaches to UQ are both effective and adaptable in the regime of astrodynamics. With regards to applications with OUU, this work focused on the regime of conjunction assessment. In particular, it was desired to quantify the probability of collision while also designing a maneuver to lower this value to an acceptable level. Previous work by Jones et al. had used discretized time "snap shots", as well as interpolation over time, to quantify the probability of collision over a given time span [66, 68]. Rather than generate a large number of surrogates to cover this span, we presented a method which only considers the QoIs at the time of closest approach. By training an SR surrogate using this approach, the probability of collision was computed. This novel method of computing the probability of collision is both efficient and accurate, and other scenario applications are possible.

As previously stated, SR has also been shown to be compatible with the concept of sample space decomposition. That is, in order to increase surrogate accuracy when fitting to a solution that is discontinuous with respect to input direction(s), the methodology of ME-SR splits the input space into element spaces and fits a surrogate to each individual element. While process has been applied to PCEs in the form of ME-gPC, some customization is needed in order to create a functioning multi-element algorithm for SR. The novel algorithm introduced in Chapter 4 relies on a *decay check parameter* for K-type refinement. This parameter quantifies the relative contribution of the highest order polynomials to the standard deviation of the QoI. Once a split is deemed necessary, the direction(s) to split is determined via an error analysis of unidimensional surrogates. Applied to test cases of various systems, the ME-SR algorithm shows orders of magnitude improvement over the unmodified SR approach when solving for a solution that is discontinuous with respect to one or more input directions.

8.3 Future Work

Work is still to be done, however, and a number of future goals have been identified. The first two topics relate to the method of OUU. Although, SR has been shown to perform well with OUU, the application remains limited to working with MATLAB's *fmincon*. The performance of SR with other optimizers, with or without established presence in the astrodynamics community, should be explored. An example is that of the SNOPT optimizer [51]. SNOPT is used within several trajectory software packages such as Copernicus, AeroSpace Trajectory Optimization and Software, General Mission Analysis Tool, and Optimal Trajectories by Implicit Simulation. A feature of SNOPT, and of other optimizers, is the ability to utilize a function gradient in order to aid the optimization algorithm. Since an SR surrogate is a linear combination of known polynomials, the gradient of the surrogate with respect to input directions is trivial to compute. Once the coefficients of s^l, $c^l_{i,p}$, and $u^l_{0,m}$ are calculated via ALS, it is possible to compute the gradient given a table of pre-derived polynomial derivatives.

Leveraging local minima to find an optimal global solution has also been shown to be effective [118]. It is expected that this technique, often called “basin hopping”, could be utilized in an OUU application with SR. Just as the polynomial formulation of SR aids in gradient calculation, local minima are also efficent and trivial to calculate. Although gradient based optimization methods should be explored, some optimization functions are not smooth or continuous. Therefore, there is motivation to explore the application and formulation of solvers for non-smooth functions. The optimization problem presented in Chapter 6 contains such a scenario. Specifically, the constraint functions of Eq. (6.13) are not smooth. An investigation of the effects that this has on the current implementation is desirable, as is an application of an optimizer that specializes in such functions.

In addition to gradient computations and optimization package exploration, the application of ME-SR to OUU is of interest. Presented in Chapter 7, ME-SR shows potential to increase the accuracy of SR when presented with a discontinuous relationship between input directions and the QoIs. In particular, a test case is formulated such that one of the input directions is effectively a design input. By including a uniform input as a function of the time of flight, an ME-SR surrogate is fitted to QoIs representing the state of a spacecraft as it endures a flyby of Europa. A potential application of this surrogate is to utilize it in conjunction with an optimization algorithm to determine the maximum risk of collision with Europa. Indeed, this particular formulation may be altered to include the probability of collision method found in Chapter 6. By considering a maneuver with design space, a procedure could be developed to minimize the risk of planetary contamination by a spacecraft with initial uncertainty.

Previously mentioned in Chapter 7, the efficiency of the presented ME-SR algorithm sampling method has the potential to be improved. The inefficiency results from the fact that the current implementation discards previous training samples when new splits in the input space are made. A new set of training samples are generated for the ME-SR surrogate construction of each new input space B_k. This is because the set of probability density functions $\{\rho_{k,i}\}_{i=1}^{d}$ are unique to each B_k, and inputs generated for $B_{k'}$, where $k \neq k'$, will not maintain orthogonality with polynomials constructed for $\{\rho_{k,i}\}_{i=1}^{d}$. Therefore, without modification, the previously generated inputs do not satisfy the orthogonal relationship depicted in Eq. (2.59). Because of this, the computation time needed to generate an ME-SR surrogate is larger than that of the theoretical computation cost associated with $N \times K$ required training samples.

To rectify this disparity and mitigate an increase in computation time, a weighting scheme similar to that of the work presented in Hampton et al. can be used [55]. Consider the weights,

$$w_{k,i} = \sqrt{\frac{\rho_{k,i}}{\rho_i}}, \tag{8.1}$$

where $w_{k,i}$ is the weight of direction i in the k-th element, and ρ_i is the probability density function of the original, un-split input variable. This orthogonal polynomial generated for direction i in the k-th element is then multiplied by the appropriate weight, satisfying the orthogonal relationship. If we consider y_i to be some input variable from the previously unsplit distribution and $\psi_{k,p}(\cdot)$ to be a polynomial orthogonal to the element k, then

$$\int \psi_{k,p}(y_i)\sqrt{\frac{\rho_{k,i}}{\rho_i}}\psi_{k,p'}(y_i)\sqrt{\frac{\rho_{k,i}}{\rho_i}}\rho_i \, dy_i = \delta_{p,p'}, \tag{8.2}$$

$$\int \psi_{k,p}(y_i)\psi_{k,p'}(y_i)\rho_{k,i} \, dy_i = \delta_{p,p'}, \tag{8.3}$$

thereby allowing inputs y_i and polynomials $\psi_{k,p}(\cdot)$ to maintain an orthogonal basis on element k. If $w_{k,i}^{j}$ is the determined weight of $\xi_{i,j}$, then the combined weight of input vector $\boldsymbol{\xi}_j$ is

$$w_k^j(\boldsymbol{\xi}_j) = \prod_{i=1}^{d} \sqrt{\frac{\rho_{k,i}}{\rho_i}}. \tag{8.4}$$

These combined weights are organized into an $N \times N$ diagonal matrix W, such that

$$\boldsymbol{W} = \begin{bmatrix} w_k^1(\boldsymbol{\xi}_1) & & \\ & \ddots & \\ & & w_k^N(\boldsymbol{\xi}_N) \end{bmatrix}. \tag{8.5}$$

When considering a scalar output ME-SR, Eq. (2.48) is replaced by

$$\boldsymbol{h}' = \boldsymbol{W}\boldsymbol{h}, \tag{8.6}$$

thereby weighting the data vector in the ALS process and accounting for the difference between sample probability density functions. By implementing such a process, the computation cost of the ME-SR algorithm would improve.

Lastly, the field of directional statistics shows promise as a mathematically rigorous means of UQ with surrogates when considering an input direction associated with an angle quantity. This development is significant in the field of astrodynamics, because many coordinate systems or hardware models rely on an angle quantity. Examples include Keplerian and equinoctial elements, attitude, and maneuver pointing. It is possible for any of these quantities to be a relevant and uncertain input. As a PDF for circular data, the Von Mises (VM) distribution has been proposed for astrodynamics applications [61]. Another possible distribution for circular data is the wrapped normal distribution (WND). For the WND, Roger-Szegő polynomials form a basis that is orthogonal to the density function. For VM, however, one of the analytic processes described in Section 4.3.2 of Chapter 4 is currently needed to construct an orthogonal basis. Preliminary results suggest that surrogate methods such as PCEs and SR are capable of accurately accounting for circular data when given proper orthogonal bases [65]. However, there is work to be done. Currently, the accurate construction of an orthogonal basis for the VM distribution is only possible for diffuse inputs. Additionally, SR has exhibited *over fitting* issues when adding excessive ranks to a surrogate model. Therefore, it is desirable to further investigate polynomial bases for the VM and other directional inputs (such as Gaussian-like distributions on a sphere), as well as the convergence properties of SR when considering directional inputs.

Bibliography

[1] Adams B, Bauman L, Bohnhoff W, Dalbey K, Ebeida M, Eddy J, Eldred M, Hough P, Hu K, Jakeman J, Swiler L, Vigil D (2009) Dakota, a multilevel parallel object-oriented framework for design optimization, parameter estimation, uncertainty quantification, and sensitivity analysis: Version 5.4 user's manual. Tech. Rep. SAND2010-2183, Sandia, updated April 2013

[2] Alby F, Lansard E, Michal T (1997) Collision of cerise with space debris. In: Second European Conference on Space Debris, vol 393, p 589

[3] Alfano S (2009) Satellite conjunction monte carlo analysis. AAS Spaceflight Mechanics Mtg, Pittsburgh, PA, Paper pp 09–233

[4] Alfriend KT, Akella MR, Frisbee J, Foster JL, Lee DJ, Wilkins M (1999) Probability of collision error analysis. Space Debris 1(1):21–35

[5] Alspach D, Sorenson H (1972) Nonlinear bayesian estimation using gaussian sum approximations. IEEE transactions on automatic control 17(4):439–448

[6] Ammar A, Chinesta F, Joyot P (2008) The nanometric and micrometric scales of the structure and mechanics of materials revisited: An introduction to the challenges of fully deterministic numerical descriptions. International Journal for Multiscale Computational Engineering 6(3):191–213, DOI 10.1615/IntJMultCompEng.v6.i3.20

[7] Anderson RL (2013) Tour design using resonant orbit heteroclinic connections in patched circular restricted three-body problems. In: 23rd AAS/AIAA Space Flight Mechanics Meeting, Kauai, Hawaii, AAS 13-493

[8] Anderson RL (2015) Approaching moons from resonance via invariant manifolds. Journal of Guidance, Control, and Dynamics 38(6):1097–1109, DOI 10.2514/1.G000286

[9] Anderson RL, Lo MW (2009) Role of invariant manifolds in low-thrust trajectory design. Journal of Guidance, Control, and Dynamics 32(6):1921–1930, DOI 10.2514/1.37516

[10] Anderson RL, Lo MW (2010) Dynamical systems analysis of planetary flybys and approach: Planar europa orbiter. Journal of Guidance, Control, and Dynamics 33(6):1899–1912, DOI 10.2514/1.45060

[11] Anderson RL, Lo MW (2011) A dynamical systems analysis of planetary flybys and approach: Ballistic case. The Journal of the Astronautical Sciences 58(2):167–194, DOI 10.2514/1.45060

[12] Anderson RL, Lo MW (2011) Flyby design using heteroclinic and homoclinic connections of unstable resonant orbits. In: Jah MK, Guo Y, Bowes AL, Lai PC (eds) Spaceflight Mechanics: Proceedings of the 21st AAS/AIAA Space Flight Mechanics Meeting held February 13-17, 2011, New Orleans, Louisiana, American Astronautical Society, Univelt Inc., San Diego, California, Advances in the Astronautical Sciences, vol 140, pp 321–340

[13] Anderson RL, Lo MW (2014) Spatial approaches to moons from resonance relative to invariant manifolds. Acta Astronautica 105:355–372, DOI 10.1016/j.actaastro.2014.09.015

[14] Askey RA, Arthur WJ (1985) Some basic hypergeometric orthogonal polynomials that generalize Jacobi polynomials, vol 319. AMS, Providence RI

[15] Balducci M, Jones BA (2017) Probability of collision estimation and optimization under uncertainty utilizing separated representations. AAS/AIAA Astrodynamics Specialist Conference Stevenson, WA, August 20-24

[16] Balducci M, Jones BA, Doostan A (2013) Orbit uncertainty propagation with separated representations. AAS/AIAA Astrodynamics Specialist Conference Hilton Head, SC, August 11-15

[17] Balducci M, Jones B, Doostan A (2017) Orbit uncertainty propagation and sensitivity analysis with separated representations. Celestial Mechanics and Dynamical Astronomy 129(1-2):105–136, DOI 10.1007/s10569-017-9767-7

[18] Berend N (1999) Estimation of the probability of collision between two catalogued orbiting objects. Advances in Space Research 23(1):243–247

[19] Beylkin G, Mohlenkamp MJ (2005) Algorithms for numerical analysis in high dimensions. SIAM Journal on Scientific Computing 26(6):2133–2159, DOI 10.1137/040604959

[20] Beylkin G, Garcke J, Mohlenkamp MJ (2009) Multivariate regression and machine learning with sums of separable functions. SIAM Journal on Scientific Computing 31(3):1840–1857, DOI 10.1137/070710524

[21] Blatman G, Sudret B (2010) Efficient computation of global sensitivity indices using sparse polynomial chaos expansions. Reliability Engineering and System Safety 95:1216–1229, DOI 10.1016/j.ress.2010.06.015

[22] Brent RP (2013) Algorithms for minimization without derivatives. Courier Corporation

[23] Campagnola S, Russell R (2010) Endgame problem part 2: Multibody technique and the tisserand-poincaré graph. Journal of Guidance, Control, and Dynamics 33(2):476–486, DOI 10.2514/1.44290

[24] Chevreuil M, Lebrun R, Nouy A, Rai P (2015) A least-squares method for sparse low rank approximation of multivariate functions. SIAM/ASA Journal on Uncertainty Quantification 3(1):897–921, DOI 10.1137/13091899X

[25] Chinesta F, Ladeveze P, Cueto E (2011) A short review on model order reduction based on proper generalized decomposition. Archives of Computational Methods in Engineering 18(4):395–404

[26] Cohen A, DeVore R, Schwab C (2010) Convergence rates of best n-term galerkin approximations for a class of elliptic spdes. Foundations of Computational Mathematics 10(6):615–646

[27] Darboux G (1878) Mémoire sur l'approximation des fonctions de très-grands nombres, et sur une classe étendue de développements en série. Journal de Mathématiques pures et appliquées pp 5–56

[28] Deaconu G, Louembet C, Théron A (2014) Minimizing the effects of navigation uncertainties on the spacecraft rendezvous precision. Journal of Guidance, Control, and Dynamics 37(2):695–700

[29] Dell'Elce L, Kerschen G (2014) Robust rendez-vous planning using the scenario approach and differential flatness. In: Proceedings of the 2nd IAA Conference on Dynamics and Control of Space System, Univelt San Diego, CA, vol 153, pp 1–14

[30] DeMars KJ, Bishop RH, Jah MK (2013) Entropy-based approach for uncertainty propagation of nonlinear dynamical systems. Journal of Guidance, Control, and Dynamics 36(4):1047–1057, DOI 10.2514/1.58987

[31] DeMars KJ, Cheng Y, Jah MK (2014) Collision probability with gaussian mixture orbit uncertainty. Journal of Guidance, Control, and Dynamics

[32] DeMars KJ, Cheng Y, Jah MK (2014) Making best use of model evaluations to compute sensitivity indices. Journal of Guidance, Control, and Dynamics 37(3):979–984, DOI 10.2514/1.62308

[33] Dolado J, Legendre P, Garmier R, Revelin B, Pena X (2011) Satellite collision probability computation for long term encounters. Adv Astronaut Sci 142

[34] Doostan A, Iaccarino G (2009) A least-squares approximation of partial differential equations with high-dimensional random inputs. Journal of Computational Physics 228(12):4332–4345, DOI 10.1016/j.jcp.2009.03.006

[35] Doostan A, Iaccarino G, Etemadi N (2007) A least-squares approximation of high-dimensional uncertain systems. Tech. Rep. Annual Research Brief, Center for Turbulence Research, Stanford University

[36] Doostan A, Validi A, Iaccarino G (2013) Non-intrusive low-rank separated approximation of high-dimensional stochastic models. Computation Methods in Applied Mechanical Engineering 263:42–55, DOI 10.1016/j.cma.2013.04.003

[37] Dormand JR, Prince PJ (1980) A family of embedded runge-kutta formulae. Journal of computational and applied mathematics 6(1):19–26

[38] Eldred M, Giunta A, Wojtkiewicz S, Trucano T (2002) Formulations for surrogate-based optimization under uncertainty. In: 9th AIAA/ISSMO symposium on multidisciplinary analysis and optimization, p 5585

[39] Eldred MS, Elman HC (2011) Design under uncertainty employing stochastic expansion methods. International Journal for Uncertainty Quantification 1(2)

[40] Fann G, Beylkin G, Harrison R, Jordan K (2004) Singular operators in multiwavelet bases. IBM Journal of Research and Development 48(2):161–171, DOI 10.1147/rd.482.0161

[41] Feldhacker JD, Jones BA, Doostan A, Hampton J (2016) Reduced cost mission design using surrogate models. Advances in Space Research 57(2):588–603, DOI 10.1016/j.asr.2015.11.002

[42] Feldhacker JD, Smith JJ, Jones BA, Doostan A (2016) Multi-element trajectory models for satellite tour missions. In: AIAA/AAS Astrodynamics Specialist Conference, p 5263

[43] Forrester A, Sobester A, Keane A (2008) Engineering design via surrogate modelling: a practical guide. John Wiley & Sons

[44] Friedman J, Hastie T, Tibshirani R (2001) The elements of statistical learning, vol 1. Springer series in statistics Springer, Berlin

[45] Fujimoto K, Scheeres D (2015) Tractable expressions for nonlinearly propagated uncertainties. Journal of Guidance, Control, and Dynamics

[46] Fujimoto K, Scheeres DJ, Alfriend KT (2012) Analytical nonlinear propagation of uncertainty in the two-body problem. Journal of Guidance, Control, and Dynamics 35(2):497–509, DOI 10.2514/1.54385

[47] Gano S, Kim H, Brown D (2006) Comparison of three surrogate modeling techniques: Datascape, kriging, and second order regression. In: 11th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, p 7048

[48] Gautschi W (1982) On generating orthogonal polynomials. SIAM Journal on Scientific and Statistical Computing 3(3):289–317, DOI 10.1145/174603.174605

[49] Gautschi W (1994) Algorithm 726: Orthpol–a package of routines for generating orthogonal polynomials and gauss-type quadrature rules. ACM Transactions on Mathematical Software (TOMS) 20(1):21–62

[50] Ghanem R, Spanos P (1991) Stochastic Finite Elements: A Spectral Approach. Springer-Verlag, New York

[51] Gill PE, Murray W, Saunders MA (2005) Snopt: An sqp algorithm for large-scale constrained optimization. SIAM review 47(1):99–131

[52] Giorgini JD, Benner LA, Ostro SJ, Nolan MC, Busch MW (2008) Predicting the earth encounters of (99942) apophis. Icarus 193(1):1–19

[53] Gist R, Oltrogge D (1999) Collision vision: Covariance modeling and intersection detection for spacecraft situational awareness

[54] Hadigol M, Doostan A, Matthies HG, Niekamp R (2014) Partitioned treatment of uncertainty in coupled domain problems: A separated representation approach. Computer Methods in Applied Mechanics and Engineering 274:103–124, DOI 10.1016/j.cma.2014.02.004

[55] Hampton J, Doostan A (2015) Coherence motivated sampling and convergence analysis of least squares polynomial chaos regression. Computer Methods in Applied Mechanics and Engineering pp 73–97, DOI doi:10.1016/j.cma.2015.02.006

[56] Hampton J, Doostan A (2015) Compressive sampling of polynomial chaos expansions: Convergence analysis and sampling strategies. Journal of Computational Physics 280:363–386, DOI 10.1016/j.jcp.2014.09.019

[57] Hansen M, Schwab C (2012) Analytic regularity and nonlinear approximation of a class of parametric semilinear elliptic pdes. Mathematische Nachrichten

[58] Harrison RJ, Fann GI, Yanai T, Gan Z, Beylkin G (2004) Multiresolution quantum chemistry: Basic theory and initial applications. Journal of Chemical Physics 121(23), DOI 10.1063/1.1791051

[59] Hoang VH, Schwab C (2013) Sparse tensor galerkin discretization of parametric and random parabolic pdes—analytic regularity and generalized polynomial chaos approximation. SIAM Journal on Mathematical Analysis 45(5):3050–3083

[60] Horwood JT, Poore AB (2011) Adaptive gaussian sum filters for space surveillance. IEEE Transactions on Automatic Control 56(8):1777–1790

[61] Horwood JT, Poore AB (2014) Gauss von mises distribution for improved uncertainty realism in space situational awareness. SIAM/ASA Journal on Uncertainty Quantification 2(1):276–304

[62] Horwood JT, Aragon ND, Poore AB (2011) Gaussian sum filters for space surveillance: Theory and simulations. Journal of Guidance, Control, and Dynamics 34(6):1839–1851, DOI 10.2514/1.53793

[63] Ito K (2000) Gaussian filter for nonlinear filtering problems. In: Decision and Control, 2000. Proceedings of the 39th IEEE Conference on, IEEE, vol 2, pp 1218–1223

[64] Izzo D, Becerra VM, Myatt DR, Nasuto SJ, Bishop JM (2007) Search space pruning and global optimisation of multiple gravity assist spacecraft trajectories. Journal of Global Optimization 38(2):283–296

[65] Jones BA, Balducci M (2018) Stochastic expansions including random inputs on the unit circle. arXiv preprint arXiv:180801052

[66] Jones BA, Doostan A (2013) Satellite collision probability estimation using polynomial chaos expansions. Advances in Space Research 52(11):1860–1875, DOI 10.1016/j.asr.2013.08.027

[67] Jones BA, Doostan A, Born GH (2013) Nonlinear propagation of orbit uncertainty using non-intrusive polynomial chaos. Journal of Guidance, Control, and Dynamics 36(2):430–444, DOI 10.2514/1.57599

[68] Jones BA, Parrish N, Doostan A (2015) Post-maneuver collision probability estimation using sparse polynomial chaos expansions. Journal of Guidance, Control, and Dynamics 36(2):430–444, DOI 10.2514/1.G000595

[69] Julier S, Uhlmann J, Durrant-Whyte HF (2000) A new method for the nonlinear transformation of means and covariances in filters and estimators. IEEE Transactions on automatic control 45(3):477–482

[70] Junkins JL, Akella MR, Alfriend KT (1996) Non-gaussian error propagation in orbital mechanics. Journal of Astronautical Sciences 44(4):541–563

[71] Kelso T, et al (2009) Analysis of the iridium 33-cosmos 2251 collision. Advances in the Astronautical Sciences 135(2):1099–1112

[72] Khoromskij BN, Schwab C (2010) Tensor-structured galerkin approximation of parametric and stochastic elliptic pdes. SIAM Journal on Scientific Computing 33:364–385, DOI 10.1137/100785715

[73] Kim NH, Wang H, Queipo NV (2006) Efficient shape optimization under uncertainty using polynomial chaos expansions and local sensitivities. AIAA journal 44(5):1112

[74] Kolda TG, Bader BW (2009) Tensor decompositions and applications. SIAM Review 51(3):455–500

[75] Lantoine G, Russell RP, Campagnola S (2011) Optimization of low-energy resonant hopping transfers between planetary moons. Acta Astronautica 68(7-8):1361–1378, DOI 10.1016/j.actaastro.2010.09.021

[76] Le Maıtre O, Knio O, Najm H, Ghanem R (2004) Uncertainty propagation using wiener–haar expansions. Journal of computational Physics 197(1):28–57

[77] Le Maıtre O, Najm HN, Ghanem R, Knio O (2004) Multi-resolution analysis of wiener-type uncertainty propagation schemes. Journal of Computational Physics 197(2):502–531

[78] Louembet C, Arzelier D, Deaconu G (2014) Robust rendezvous planning under maneuver execution errors. Journal of Guidance, Control, and Dynamics 38(1):76–93

[79] Luo Y, Yang Z, Li H (2014) Robust optimization of nonlinear impulsive rendezvous with uncertainty. Science China Physics, Mechanics and Astronomy 57(4):731–740

[80] Majji M, Junkins J, Turner J (2008) A high order method for estimation of dynamic systems. The Journal of the Astronautical Sciences 56(3):401–440, DOI 10.1007/BF03256560

[81] Morselli A, Armellin R, Di Lizia P, Zazzera FB (2014) A high order method for orbital conjunctions analysis: sensitivity to initial uncertainties. Advances in Space Research 53(3):490–508

[82] Mueller JB, Larsson R (2008) Collision avoidance maneuver planning with robust optimization. In: International ESA Conference on Guidance, Navigation and Control Systems, Tralee, County Kerry, Ireland

[83] Nielsen P, Alfriend K, Bloomfield M, Emmert J, Miller J, Guo Y, et al (2012) Continuing Kepler's Quest: Assessing Air Force Space Command's Astrodynamic Standards. The National Academies Press, Washington DC

[84] Nouy A (2010) Proper generalized decompositions and separated representations for the numerical solution of high dimensional stochastic problems. Archives of Computational Methods in Engineering 17:403–434, DOI 10.1007/s11831-010-9054-1

[85] Park RS, Scheeres DJ (2006) Nonlinear mapping of gaussian statistics: Theory and applications to spacecraft trajectory design. Journal of Guidance, Control, and Dynamics 29(6):1367–1375, DOI 10.2514/1.20177

[86] Patera RP (2001) General method for calculating satellite collision probability. Journal of Guidance, Control, and Dynamics 24(4):716–722

[87] Patera RP (2007) Space vehicle conflict-avoidance analysis. Journal of guidance, control, and dynamics 30(2):492–498

[88] Patera RP (2007) Space vehicle conflict probability for ellipsoidal conflict volumes. Journal of Guidance, Control, and Dynamics 30(6):1819–1822

[89] Patera RP, Peterson GE (2003) Space vehicle maneuver method to lower collision risk to an acceptable level. Journal of guidance, control, and dynamics 26(2):233–237

[90] Peng H, Yang C, Li Y, Zhang S, Chen B (2013) Surrogate-based parameter optimization and optimal control for optimal trajectory of halo orbit rendezvous. Aerospace Science and Technology 26(1):176–184

[91] Pontani M, Conway BA (2013) Optimal finite-thrust rendezvous trajectories found via particle swarm algorithm. Journal of Spacecraft and Rockets 50(6):1222–1234

[92] Pontani M, Ghosh P, Conway BA (2012) Particle swarm optimization of multiple-burn rendezvous trajectories. Journal of Guidance, Control, and Dynamics 35(4):1192–1207

[93] Quadrelli MB, Wood LJ, Riedel JE, McHenry MC, Aung M, Cangahuala LA, Volpe RA, Beauchamp PM, Cutts JA (2015) Guidance, navigation, and control technology assessment for future planetary science missions. Journal of Guidance, Control, and Dynamics 38(7):1165–1186, DOI 10.2514/1.G000525

[94] Reynolds MJ, Doostan A, Beylkin G (2016) Randomized alternating least squares for canonical tensor decompositions: Application to a pde with random data. SIAM Journal on Scientific Computing 38(5):A2634–A2664, DOI 10.1137/15M1042802

[95] Russell R (2012) Survey of spacecraft trajectory design in strongly perturbed environments. Journal of Guidance, Control, and Dynamics 35(3):705–720, DOI 10.2514/1.56813

[96] Sabol C, Binz C, Segerman A, Roe K, Schumacher Jr PW (2011) Probability of collision with special perturbations dynamics using the monte carlo method. In: AAS/AIAA Astrodynamics Specialist Conference, Girdwood, AK

[97] Saltelli A (2002) Making best use of model evaluations to compute sensitivity indices. Computer Physics Communications 145():280–297, DOI 10.1016/S0010-4655(02)00280-1

[98] Saltelli A, Andres T, Homma T (1993) Sensitivity analysis of model output: an investigation of new techniques. Computational statistics & data analysis 15(2):211–238

[99] Saltelli A, Tarantola S, Campolongo F, Ratto M (2004) Sensitivity Analysis in Practice: A guide to assessing scientific models. Wiley, Hoboken, NJ

[100] Schilling B, Taleb Y, Carpenter JR, Balducci M, Williams TW (2016) Operational experience with the wald sequential probability ratio test for conjunction assessment from the magnetospheric multiscale mission. In: AIAA/AAS Astrodynamics Specialist Conference, p 5424

[101] Schutz B, Tapley B, Born GH (2004) Statistical orbit determination. Academic Press

[102] Smith R (2013) Uncertainty Quantification: Theory, Implementation, and Applications. SIAM-Society for Industrial and Applied Mathematics

[103] Smolyak SA (1963) Quadrature and interpolation formulas for tensor products of certain classes of functions. In: Doklady Akademii Nauk, Russian Academy of Sciences, vol 148, pp 1042–1045

[104] Sobol I (2001) Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. Mathematics and Computers in Simulation 55:271–280

[105] Stieltjes TJ (1884) Quelques recherches sur la théorie des quadratures dites mécaniques. In: Annales scientifiques de l'École Normale Supérieure, Elsevier, vol 1, pp 409–426

[106] Sudret B (2008) Global sensitivity analysis using polynomial chaos expansions. Reliability Engineering & System Safety 93(7):964–979

[107] Sun Y, Kumar M (2015) Uncertainty propagation in orbital mechanics via tensor decomposition. Celestial Mechanics and Dynamical Astronomy pp 1–26, DOI 10.1007/s10569-015-9662-z

[108] Szebehely V (1967) Theory of Orbits: The Restricted Problem of Three Bodies. Academic Press, New York

[109] Szeg G (1939) Orthogonal polynomials, vol 23. American Mathematical Soc.

[110] Tamellini L, Le Maitre O, Nouy A (2014) Model reduction based on proper generalized decomposition for the stochastic steady incompressible navier-stokes equations. SIAM Journal on Scientific Computing 36(3):A1089–A1117, DOI 10.1137/120878999

[111] Tapley B, Ries J, Bettadpur S, Chambers D, Cheng M, Condi F, Gunter B, Kang Z, PNagel, Pastor R, Pekker T, SPoole, Wang F (2005) Ggm02 - an improved earth gravity field model from grace. Journal of Geodesy DOI 10.1007/s00190-005-0480-z

[112] Tardioli C, Kubicek M, Vasile M, Minisci E, Riccardi A (2015) Comparison of non-intrusive approaches to uncertainty propagation in orbital mechanics. In: Proceedings of the AAS/AIAA Astrodynamics Specialist Conference, American Astronautical Society, vol 156, pp 3979–3992

[113] Vallado D (2007) Fundamentals of Astrodynamics and Applications, 3rd edn, Microcosm Press, Hawthorne, CA, chap 8.6, p 562

[114] Valli M, Armellin R, Di Lizia P, Lavagna M (2012) Nonlinear mapping of uncertainties in celestial mechanics. Journal of Guidance, Control, and Dynamics 36(1):48–63

[115] Vaquero M, Howell KC (2011) Poincaré maps and resonant orbits in the circular restricted three-body problem. In: Schaub H, Gunter BC, Russell RP, Cerven WT (eds) Astrodynamics: Proceedings of the AAS/AIAA Astrodynamics Specialist Conference held July 31 to August 4, 2011, in Girdwood, Alaska, American Astronautical Society, Univelt Inc., San Diego, California, Advances in the Astronautical Sciences, vol 142, pp 433–451

[116] Vaquero M, Howell KC (2012) Design of transfer trajectories between resonant orbits in the restricted three-body problem with application to the earth-moon system. In: 1st IAA/AAS Conference on the Dynamics and Control of Space Systems, Porto, Portugal

[117] Vittaldev V, Russell RP, Linares R (2016) Spacecraft uncertainty propagation using gaussian mixture models and polynomial chaos expansions. Journal of Guidance, Control, and Dynamics pp 2615–2626, DOI 10.2514/1.G001571

[118] Wales DJ, Doye JP (1997) Global optimization by basin-hopping and the lowest energy structures of lennard-jones clusters containing up to 110 atoms. The Journal of Physical Chemistry A 101(28):5111–5116

[119] Wallace MS (2015) A massively parallel bayesian approach to planetary protection trajectory analysis and design. Advances in the Astronautical Sciences Astrodynamics 156, pasadena, CA: Jet Propulsion Laboratory, National Aeronautics and Space Administration, 2015

[120] Wan X, Karniadakis GE (2005) An adaptive multi-element generalized polynomial chaos method for stochastic differential equations. Journal of Computational Physics 209(2):617–642, DOI 10.1016/j.jcp.2005.03.023

[121] Wan X, Karniadakis GE (2006) Multi-element generalized polynomial chaos for arbitrary probability measures. SIAM Journal on Scientific Computing 28(3):901–928, DOI 10.1137/050627630

[122] Xiu D (2010) Numerical Methods for Stochastic Computations: A Spectral Method Approach. Princeton University Press

[123] Yang Z, Luo YZ, Zhang J, Tang GJ (2016) Uncertainty quantification for short rendezvous missions using a nonlinear covariance propagation method. Journal of Guidance, Control, and Dynamics pp 2170–2178

[124] Yang Z, Luo Yz, Zhang J (2017) Robust planning of nonlinear rendezvous with uncertainty. Journal of Guidance, Control, and Dynamics pp 1–14

[125] Zhang J, Parks G (2013) Multi-objective optimization for multiphase orbital rendezvous missions. Journal of Guidance, Control, and Dynamics 36(2):622–629

Appendix A

Stokes Coefficients

Table A.1: Low degree Stokes coefficients

	Value	STD
$C_{2,0}$	-4.8416e-04	6.1e-11
$C_{2,2}$	2.4393e-06	3.1e-11
$C_{3,0}$	9.5721e-07	1.1e-11
$C_{3,1}$	2.0304e-06	1.6e-11
$C_{3,2}$	9.0479e-07	2.2e-11
$C_{3,3}$	7.2127e-07	2.6e-11
$C_{4,0}$	5.3999e-07	8.2e-12
$S_{2,2}$	-1.4002e-06	3.1e-11
$S_{3,1}$	2.4820e-07	1.6e-11
$S_{3,2}$	-6.1898e-07	2.2e-11
$S_{3,3}$	1.4143e-06	2.6e-11

Appendix B

Equinoctial Elements

Note that f_r is a *retrograde factor*, where it is $+1$ for all direct orbits and -1 for nearly retrograde orbits [113].

$$h_e = e\sin\omega + f_r\Omega \tag{B.1}$$

$$k_e = e\cos\omega + f_r\Omega \tag{B.2}$$

$$p_e = \frac{\sin i \sin\Omega}{1+\cos^{f_r} i} \tag{B.3}$$

$$q_e = \frac{\sin i \cos\Omega}{1+\cos^{f_r} i} \tag{B.4}$$

$$\lambda_{\mathcal{M}} = \mathcal{M} + \omega + f_r\Omega \tag{B.5}$$