

Institut für Parallele und Verteilte Systeme

Universität Stuttgart
Universitätsstraße 38
D-70569 Stuttgart

Bachelorarbeit

Polynomial Chaos Expansion mit räumlich adaptiven Sparse Grids

Thomas Albrecht

Studiengang: Data Science
Prüfer/in: Prof. Dr. Dirk Pflüger
Betreuer/in: M. Sc. Michael Rehme

Beginn am: 2. April 2020
Beendet am: 2. November 2020

Kurzfassung

Die Polynomial Chaos Expansion (generalized Polynomial Chaos) ist eine Methode aus der Uncertainty Quantification. Mit ihr können die stochastischen Momente einer Funktion $R(\xi)$, $R : \mathbb{R}^D \rightarrow \mathbb{R}$, deren Parameter $\xi = (\xi_1, \dots, \xi_D)$ gemäß Verteilungsfunktionen verteilt sind, schnell berechnet werden. Außerdem liefert die Methode ein akkurates Surrogat das effizient ausgewertet werden kann. Dazu werden abhängig von den Verteilungsfunktionen in jeder Dimension bestimmte orthogonale Polynombasen verwendet. Die verwendeten Polynome sind jeweils genau bezüglich der Dichtefunktion der jeweiligen Verteilung orthogonal. Die Koeffizienten der Expansion werden mithilfe eines Spectral Projection Ansatzes berechnet. Bei diesem müssen hochdimensionale Integrale gelöst werden. In dieser Arbeit wird die Polynomial Chaos Expansion implementiert. Da traditionelle Integrationsmethoden unter dem sogenannten Fluch der Dimensionalität leiden, werden die Integrale mithilfe von räumlich adaptiven Sparse Grids (dünnen Gittern) gelöst. Die Ergebnisse der Polynomial Chaos Expansion mit räumlich adaptiven Sparse Grids werden mit denen einer Polynomial Chaos Expansion mit regulären Sparse Grids verglichen. Außerdem werden die Ergebnisse noch mit einer weiteren Methode aus der Uncertainty Quantification, der Stochastic Collocation, verglichen.

Inhaltsverzeichnis

1	Einleitung	7
2	Grundlagen	9
2.1	Uncertainty Quantification	9
2.2	Polynomial Chaos Expansion	9
3	Sparse Grids	17
3.1	Grundlagen	17
3.2	Sparse Grids	19
3.3	Räumliche Adaptivität	19
3.4	Ansatzfunktionen	20
3.5	Integration mittels Sparse Grids und Stochastic Collocation	21
4	Ergebnisse	23
4.1	Methoden	23
4.2	Gering-dimensionale Beispiele	24
4.3	Short Column	27
5	Zusammenfassung und Ausblick	33
	Literaturverzeichnis	35

1 Einleitung

Computersimulationen und mathematische Modelle haben sich in vielen Wissenschaftsgebieten etabliert. Sie bieten viele Vorteile gegenüber traditionellen Experimenten und ermöglichen auch neuartige Experimente, die in der Realität nicht durchführbar wären. Da die mathematischen Modelle jedoch die Realität nicht exakt beschreiben können, ist es wichtig, die auftretenden Ungenauigkeiten und Unsicherheiten zu behandeln und analysieren. Die Uncertainty Quantification beschäftigt sich mit diesen und weiteren Fragestellungen, dabei werden die Unsicherheiten stochastisch modelliert.

Die Polynomial Chaos Expansion (PCE) ist eine Methode aus der Uncertainty Quantification. Die hier behandelte Methode ist als „Generalized Polynomial Chaos“ bekannt. Sie erlaubt es, den Einfluss der Unsicherheiten in der Eingabe auf die Ausgabe einer Funktion oder eines Modells zu analysieren. Dazu werden die Parameter der Funktion als Zufallsvariablen modelliert, die die Unsicherheiten der einzelnen Parameter reflektieren. Beispielsweise können stochastische Momente wie Erwartungswert und Varianz berechnet werden, es kann jedoch auch die der Funktion zugrundeliegende Verteilungsfunktion bestimmt werden. Gleichzeitig liefert die Methode eine Approximation der Funktion oder des Modells basierend auf Polynomen. Diese Approximation kann häufig effizienter ausgewertet werden als die ursprüngliche Funktion. Es existieren verschiedene Ansätze um die Koeffizienten α_j aus der Expansion zu berechnen. Der hier vorgestellte Ansatz verwendet keine zusätzlichen Eigenschaften der approximierten Funktion, diese muss lediglich als „Black-Box“ ausgewertet werden können. Zur Berechnung der Koeffizienten wird ein mehrdimensionales Integral aufgestellt. Klassische Quadratur-Verfahren leiden unter dem sogenannten „Fluch der Dimensionalität“: Die Anzahl der erforderlichen Auswertungen der Zielfunktion und damit der Aufwand der Integration steigen exponentiell mit der Zahl der Dimensionen. Ein Verfahren, das diesem Sachverhalt zu einem gewissen Grad entgegenwirkt, ist die numerische Integration mithilfe von Sparse Grids (dünnere Gitter).

Sparse Grids sind eine Methode aus der numerischen Mathematik, die es ermöglicht gute Approximationen für Zielfunktion zu bilden, die günstiger ausgewertet werden können als die ursprüngliche Funktion. Für diese Approximation kann dann das ursprüngliche Problem, das für die Zielfunktion zu lösen war, gelöst werden. Dabei ist die Zahl der benötigten Auswertungen der Zielfunktion deutlich geringer als bei traditionellen Lösungsverfahren. Der durch die Approximation entstehende Fehler ist jedoch gering. Auf diese Art können hochdimensionale Probleme, für deren Lösung der Aufwand mit traditionellen Lösungsverfahren zu hoch wäre, trotzdem gelöst werden.

Im Folgenden werden zuerst die mathematischen Grundlagen der Polynomial Chaos Expansion und der Sparse Grids vorgestellt. Danach werden verschiedene Ergebnisse der Implementierung dieser Methode vorgestellt. Zuletzt werden die Ergebnisse noch mit denen einer anderen etablierten Methode aus der Uncertainty Quantification, der Stochastic Collocation, und denen einer anderen PCE Implementierung (DAKOTA) verglichen.

2 Grundlagen

2.1 Uncertainty Quantification

Computersimulationen und wissenschaftliches Rechnen gewinnen in verschiedenen Natur- und Ingenieurwissenschaften, aber auch in der Ökonomie immer mehr an Bedeutung. Die Simulationen ermöglichen beispielsweise Experimente kosteneffizienter, schneller und ohne Folgen in der Realität durchzuführen. Diese Simulationen basieren auf mathematischen Modellen. Diese Modelle können die Realität jedoch nicht exakt beschreiben, da gewisse Faktoren vereinfacht dargestellt oder weggelassen werden. Außerdem lassen sich oft nicht alle Parameter des Systems genau bestimmen oder sind inhärent variabel [LK10].

Uncertainty Quantification (UQ) beschäftigt sich mit diesen und weiteren Unsicherheiten in Systemen und mathematischen Modellen und wie sich diese auf die Ausgaben der Systeme beziehungsweise Modelle auswirken. Dafür werden die unsicheren Parameter des Modells durch Zufallsvariablen modelliert, die gemäß bestimmten Verteilungsfunktionen verteilt sind. Es ist dann möglich mithilfe verschiedener Methoden zu bestimmen, wie die zufällig verteilten Parameter die Ausgabe des Modells beeinflussen. Da das Modell durch die probabilistischen Parameter nun selber als eine Zufallsvariable gesehen werden kann, sind insbesondere der Erwartungswert und die Varianz des Modells von Interesse.

2.2 Polynomial Chaos Expansion

Die Polynomial Chaos Expansion ist eine Methode aus dem Bereich der Uncertainty Quantification, mithilfe welcher der Einfluss der Unsicherheiten in der Eingabe auf die Ausgabe eines Modells analysiert werden kann. Das Modell ist dabei als eine Zielfunktion $R(\xi)$, $R : \mathbb{R}^D \rightarrow \mathbb{R}$ mit $\xi = (\xi_1, \dots, \xi_D)$ einem Vektor unabhängiger Zufallsvariablen gegeben. Diese Zielfunktion kann unter Umständen nur als Black-Box gegeben sein und das Auswerten der Funktion kann ineffizient sein. Das Ziel ist es nun, ein sogenanntes Surrogat für die Zielfunktion zu erhalten, sowie die stochastischen Momente der Zielfunktion zu bestimmen. Unter einem Surrogat ist eine Approximation der Zielfunktion zu verstehen, die die Eigenschaften der Zielfunktion möglichst genau abbildet, jedoch gleichzeitig effizienter ausgewertet werden kann. Dieses wird hier durch eine Summe von Polynomen dargestellt. Dadurch ist das Surrogat selbst ein Polynom und damit oft deutlich effizienter auszuwerten als die Zielfunktion. Die Polynomial Chaos Expansion ermöglicht außerdem eine effiziente Berechnung der stochastischen Momente, diese lassen sich direkt aus den Koeffizienten α_j der Expansion berechnen.

2.2.1 Wiener-Askey Schema

Die Methode der Polynomial Chaos Expansion verwendet orthogonale Polynombasen aus dem Askey-Schema. Diese Polynome bilden für die verschiedenen Wahrscheinlichkeitsverteilungen optimale Basen, da sie orthogonal bezüglich einer Gewichtungsfunktion, die bis auf einen konstanten Faktor der Dichtefunktion der jeweiligen Verteilung entspricht, sind.

Für die Polynomial Chaos Expansion wird jeder Dimension der Zielfunktion, entsprechend der zugrundeliegenden Wahrscheinlichkeitsverteilung der Zufallsvariable ξ_i dieser Dimension, eine Polynombasis zugewiesen. Diese Polynombasen entstammen dem Askey-Schema orthogonaler Polynome.

Eine Menge von Polynomen $\{P_n(x), n \in \mathbb{N}\}$, mit P_n ein Polynom von Grad n , wird als orthogonal bezüglich einer Gewichtungsfunktion $w(x)$ bezeichnet, wenn sie folgende Eigenschaft erfüllt:

$$\int_S P_n(x)P_m(x)w(x)dx = h_n\delta_{nm} \quad \forall n, m \in \mathbb{N} \quad (2.1)$$

Sie sind also orthogonal, falls das Skalarprodukt zweier Polynome genau dann nicht null ist, wenn die Polynome gleich sind. Die Gewichtungsfunktion $w(x)$ ist nicht-negativ und auf dem gesamten Träger S definiert. Gilt außerdem $h_n = 1, \forall n \in \mathbb{N}$ spricht man von orthonormalen Polynomen.

Die Polynome werden so gewählt, dass sie orthogonal bezüglich einer Gewichtungsfunktion, die bis auf einen konstanten Faktor der Dichtefunktion der zugeordneten Wahrscheinlichkeitsverteilung entspricht, sind. Dies ermöglicht eine optimale Konvergenz [XK02].

In Tabelle 2.1 sind die optimalen Polynombasen für einige häufig verwendeten kontinuierlichen Wahrscheinlichkeitsverteilungen gegeben. Die Tabelle entspricht der in [EB09] vorgestellten, welche wiederum auf einer Tabelle aus [XK02] basiert.

Verteilung	Verteilungsfunktion	Polynombasis	Gewichtungsfunktion	Definitionsbereich
Normal	$\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$	Hermiteische $He_n(x)$	$e^{-\frac{x^2}{2}}$	$[-\infty, \infty]$
Uniform	$\frac{1}{2}$	Legendre $P_n(x)$	1	$[-1, 1]$
Beta	$\frac{(1-x)^\alpha(1+x)^\beta}{2^{\alpha+\beta+1}B(\alpha+1, \beta+1)}$	Jacobi $P_n^{(\alpha, \beta)}(x)$	$(1-x)^\alpha(1+x)^\beta$	$[-1, 1]$
Exponential	e^{-x}	Laguerre $L_n(x)$	e^{-x}	$[0, \infty]$
Gamma	$\frac{x^\alpha e^{-x}}{\Gamma(\alpha+1)}$	V. Laguerre $L_n^{(\alpha)}(x)$	$x^\alpha e^{-x}$	$[0, \infty]$

Tabelle 2.1: Wahrscheinlichkeitsverteilungen und die zugewiesenen Polynombasen aus dem Askey Schema [EB09]

Im folgenden werden die orthogonalen Polynome aus Tabelle 2.1 kurz vorgestellt:

Hermitesche Polynome

Die Hermiteschen Polynome (im Englischen als „probabilists’ Hermite Polynomials“ bezeichnet) sind orthogonal bezüglich $w(x) = e^{-\frac{x^2}{2}}$ und sind durch folgende Formel gegeben:

$$He_n(x) = (-1)^n e^{\frac{x^2}{2}} \frac{d^n}{dx^n} e^{-\frac{x^2}{2}}$$

Sie sind durch folgende Rekursionformel darstellbar:

$$He_{n+1}(x) = xHe_n(x) - nHe_{n-1}(x), \quad He_0(x) = 1, \quad He_1(x) = x$$

Ihre Norm hat folgende geschlossene Form:

$$\langle He_n^2 \rangle = \sqrt{2\pi} \cdot n!$$

Legendre Polynome

Die Legendre Polynome sind orthogonal bezüglich $w(x) = 1$ und durch folgende Formel gegeben:

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n$$

Sie sind durch folgende Rekursionformel darstellbar:

$$P_{n+1}(x) = \frac{(2n+1)}{(n+1)} x P_n(x) - \frac{n}{(n+1)} P_{n-1}(x), \quad P_0(x) = 1, \quad P_1(x) = x$$

Ihre Norm hat folgende geschlossene Form:

$$\langle P_n^2 \rangle = \frac{2}{2n+1}$$

Sie sind ein Spezialfall der Jacobi Polynome für $w(x) = \alpha = \beta = 0$.

Jacobi Polynome

Die Jacobi Polynome sind orthogonal bezüglich $w(x) = (1-x)^\alpha (1+x)^\beta$ und durch folgende Formel gegeben:

$$P_n^{\alpha,\beta}(x) = \frac{\Gamma(\alpha+n+1)}{n! \Gamma(\alpha+\beta+n+1)} \sum_{m=0}^n \binom{n}{m} \frac{\Gamma(\alpha+\beta+n+m+1)}{\Gamma(\alpha+m+1)} \left(\frac{x-1}{2}\right)^m$$

Sie sind ebenfalls durch eine Rekursionformel darstellbar. Ihre Norm hat folgende geschlossene Form:

$$\langle P_n^{(\alpha,\beta)^2} \rangle = \frac{2^{\alpha+\beta+1}}{(2n+\alpha+\beta+1)} \frac{\Gamma(n+\alpha+1)\Gamma(n+\beta+1)}{\Gamma(n+\alpha+\beta+1)n!}$$

Verallgemeinerte Laguerre Polynome

Die verallgemeinerten Laguerre Polynome (auch zugeordneten Laguerre Polynome) sind orthogonal bezüglich $x^\alpha e^{-x}$ und durch folgende Formel gegeben:

$$L_n^\alpha(x) = \frac{e^x x^{-\alpha}}{n!} \frac{d^n}{dx^n} (e^{-x} x^{n+\alpha})$$

Sie sind durch folgende Rekursionformel darstellbar:

$$L_{n+1}^\alpha(x) = \frac{(2n + \alpha + 1 - x)}{(n + 1)} L_n^\alpha(x) - \frac{(n + \alpha)}{(n + 1)} L_{n-1}^\alpha(x), \quad L_0^\alpha(x) = 1, \quad L_1^\alpha(x) = -x + \alpha + 1$$

Ihre Norm hat folgende geschlossene Form:

$$\langle L_n^{\alpha 2} \rangle = \frac{\Gamma(n + \alpha + 1)}{n!}$$

Die Laguerre Polynome sind ein Spezialfall für $\alpha = 0$.

2.2.2 Generalized Polynomial Chaos

Die Zielfunktion $R(\xi)$, mit dem Zufallsvektor $\xi = (\xi_1, \dots, \xi_D)$ ist darstellbar als folgende Expansion:

$$R(\xi) = \sum_{j=0}^{\infty} \alpha_j \Psi_j(\xi) \tag{2.2}$$

Dabei sind $\Psi_j(\xi)$ mehrdimensionale Polynome, die folgendermaßen aus den eindimensionalen Polynomen $\psi(\xi)$ aus Tabelle 2.1 gebildet werden können:

$$\Psi_j(\xi) = \prod_{i=1}^n \psi_{m_j^i}(\xi_i)$$

Hierbei ist der Index m_j^i Teil eines Multi-Indexes, gegeben über $m_j = (m_j^1, \dots, m_j^D)$, $m_j^i \in \mathbb{N}$, geordneten Tupeln natürlicher Zahlen, über welche die Kombinationen der eindimensionalen Polynome definiert sind. Eine Möglichkeit der Konstruktion eines solchen Multi-Indexes wird in Abschnitt 2.2.5 vorgestellt.

Die so konstruierten mehrdimensionalen Polynome sind auch orthogonal [Sud15].

2.2.3 Trunkierung

Die Expansion, die durch Gleichung (2.2) gegeben ist, stellt die Zielfunktion genau dar. In der Praxis ist es jedoch in der Regel nicht möglich die unendlich vielen Koeffizienten zu berechnen. Deshalb muss die Expansion trunziert werden. Hierzu wird die Summe nach endlich vielen Gliedern abgeschnitten. Die so trunzierte Expansion stellt die Zielfunktion nicht mehr genau dar, sie wird nur noch approximiert. Diese Approximation wird im Folgenden mit \hat{R} bezeichnet.

$$R(\xi) \cong \sum_{j=0}^P \alpha_j \Psi_j(\xi) =: \hat{R}(\xi) \tag{2.3}$$

Es gibt mehrere verschiedene Möglichkeiten zur Trunkierung der Expansion. Eine Möglichkeit ist als „total-order expansion“ bekannt (vgl. [EB09]). Hierfür wird der maximale Polynomgrad p der mehrdimensionalen Polynome aus Gleichung (2.2) beschränkt.

Im Folgenden bezeichne N die Anzahl der Summanden aus Gleichung (2.3), D sei weiterhin die Dimensionalität der Zielfunktion. Die Anzahl der Summanden ist dann folgendermaßen gegeben:

$$N = 1 + P = \frac{(D + p)!}{D! p!}$$

Eine weitere Möglichkeit ist das Beschränken des Polynomgrads jeder einzelnen Dimension. Sei p_i der maximale Polynomgrad der i -ten Dimension, dann gilt für die Anzahl der Summanden:

$$N = 1 + P = \prod_{i=1}^D (p_i + 1)$$

In der Theorie wurden auch weitere Möglichkeiten zur Trunkierung erarbeitet, unter anderem auch adaptive Verfahren [EB09], die die Eigenschaften der Zielfunktion verwenden, um a-priori eine bessere Trunkierung zu ermöglichen. Diese Verfahren werden in dieser Arbeit nicht weiter behandelt.

Bezüglich der Wahl eines „guten“ Wertes für p lässt sich noch folgendes anmerken: Selbst für kleine Werte p explodiert die Anzahl der Summanden mit steigender Zahl von Dimensionen. Für hochdimensionale Probleme sind deshalb Werte $p \in \{2, 3, 4, 5\}$ typisch.

2.2.4 Transformation der Input-Zufallsvariablen

Wendet man die Methode der Polynomial Chaos Expansion auf Probleme aus der Praxis an, kommt es selten vor, dass die modellierten Zufallsvariablen den standardisierten Verteilungen folgen. Entsprechen die den Input-Zufallsvariablen zugrundeliegenden Wahrscheinlichkeitsverteilungen nicht denen aus Tabelle 2.1, so ist eine Transformation der Zufallsvariablen in die gewünschte Form nötig. Sei $\xi = (\xi_1, \dots, \xi_D)$ der Zufallsvektor der Zielfunktion. Um die Methode der Polynomial Chaos Expansion anzuwenden, muss jedes ξ_i mittels einer der Wahrscheinlichkeitsverteilungen aus Tabelle 2.1 beschrieben werden. Die Polynomial Chaos Expansion wird dann für diese transformierten Zufallsvariablen berechnet.

Sei beispielsweise $R(\xi)$ eine eindimensionale Zielfunktion, $\xi \sim \mathcal{N}(\mu, \sigma)$. Hier lässt sich ξ auch folgendermaßen darstellen:

$$\xi = \mu + \sigma X$$

mit $X \sim \mathcal{N}(0, 1)$, einer Standard-Normalverteilten Zufallsvariablen [Sud15].

Wie bereits erwähnt werden im Folgenden nur unabhängige Zufallsvariablen betrachtet, jedoch ist es auch möglich in diesem Transformationsschritt korrelierte Zufallsvariablen derart zu transformieren, dass sie verwendet werden können. Dies wird beispielsweise in [EB09] näher behandelt.

2.2.5 Beispiel

Es sei eine Zielfunktion $R(\xi)$ gegeben, $\xi = (\xi_1, \xi_2)$, mit $\xi_1 \sim \mathcal{U}(a, b)$ und $\xi_2 \sim \mathcal{N}(\mu, \sigma)$. Wendet man nun als ersten Schritt die oben erklärten Transformationen an, um die Zufallsvariablen durch die in Tabelle 2.1 definierten Verteilungen zu beschreiben, erhält man folgende Zufallsvariablen:

$$\begin{aligned} \xi_1 \sim \mathcal{U}(a, b) \quad \xi_1 &= \frac{(a+b)}{2} + \frac{(b-a)}{2}x_1 \quad \text{mit} \quad x_1 \sim \mathcal{U}(-1, 1) \\ \xi_2 \sim \mathcal{N}(\mu, \sigma) \quad \xi_2 &= \mu + \sigma x_2 \quad \text{mit} \quad x_2 \sim \mathcal{N}(0, 1) \end{aligned}$$

Im Folgenden ist beispielhaft die Berechnung der „total-order expansion“ für Polynomgrad $p = 3$ gegeben. In Tabelle 2.2 sind der Index j , der Multi-Indexeintrag m_j und das mehrdimensionale Polynom für die gewählte Expansions-Ordnung und die transformierten Zufallsvariablen dargestellt.

Index j	Multi-Index Eintrag m_j	Polynom Ψ_j
0	(0,0)	$\Psi_0 = P_0(x_1) \cdot He_0(x_2) = 1$
1	(1,0)	$\Psi_1 = P_1(x_1) \cdot He_0(x_2) = x_1$
2	(0,1)	$\Psi_2 = P_0(x_1) \cdot He_1(x_2) = x_2$
3	(2,0)	$\Psi_3 = P_2(x_1) \cdot He_0(x_2) = \frac{1}{2}(3x_1^2 - 1)$
4	(1,1)	$\Psi_4 = P_1(x_1) \cdot He_1(x_2) = x_1x_2$
5	(0,2)	$\Psi_5 = P_0(x_1) \cdot He_2(x_2) = x_2^2 - 1$
6	(3,0)	$\Psi_6 = P_3(x_1) \cdot He_0(x_2) = \frac{1}{2}(5x_1^3 - 3x_1)$
7	(2,1)	$\Psi_7 = P_2(x_1) \cdot He_1(x_2) = \frac{1}{2}(3x_1^2 - 1)x_2$
8	(1,2)	$\Psi_8 = P_1(x_1) \cdot He_2(x_2) = x_1(x_2^2 - 1)$
9	(0,3)	$\Psi_9 = P_0(x_1) \cdot He_3(x_2) = x_2^3 - 3x_2$

Tabelle 2.2: Index, Multiindex und Polynom für das Beispiel

Die Expansion für die gewählte Trunkierung ist also durch folgende Gleichung gegeben:

$$\begin{aligned} R(\xi) \cong \sum_{j=0}^P a_j \Psi_j(\xi) &= \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 \frac{1}{2}(3x_1^2 - 1) + \alpha_4 x_1 x_2 + \alpha_5 (x_2^2 - 1) \\ &+ \alpha_6 \frac{1}{2}(5x_1^3 - 3x_1) + \alpha_7 \frac{1}{2}(3x_1^2 - 1)x_2 + \alpha_8 x_1(x_2^2 - 1) + \alpha_9 (x_2^3 - 3x_2) \end{aligned}$$

2.2.6 Lösen der Koeffizienten mithilfe eines Spectral Projection Ansatzes

Bisher wurden für die Expansion, die in Gleichung (2.3) gegeben ist, nur die mehrdimensionalen Polynome definiert. In der Methode der Polynomial Chaos Expansion sind diese Polynome auch fest gegeben, es müssen also lediglich noch die Koeffizienten α_j bestimmt werden. Für die Berechnung der Koeffizienten gibt es verschiedene Möglichkeiten, die sich in zwei Kategorien einordnen lassen: Intrusive Lösungsverfahren und nicht-intrusive Verfahren. Intrusive Lösungsverfahren verwenden die Eigenschaften oder Formulierung der Zielfunktion. Ein Vorteil nicht-intrusiver Verfahren ist, dass die Eigenschaften der Zielfunktion nicht für die Berechnung der Koeffizienten benötigt

werden, die Zielfunktion muss also nur als „Black-Box“ ausgewertet werden können. Im Folgenden wird der nicht-intrusive Ansatz der Spectral Projection vorgestellt. Dabei wird ein Koeffizient α_j bestimmt, indem das Skalarprodukt der Zielfunktion R mit der zugehörigen Basisfunktion Ψ_j gebildet und berechnet wird. Dies ist eine Orthogonalprojektion im Funktionenraum, der durch die Basisfunktionen Ψ_j gegeben ist.

$$\alpha_j = \frac{\langle R, \Psi_j \rangle}{\langle \Psi_j^2 \rangle} = \frac{1}{\langle \Psi_j^2 \rangle} \int_{\Omega} R(\xi) \Psi_j(\xi) \varrho(\xi) d\xi \quad (2.4)$$

Hier ist $\varrho(\xi) = \prod_{i=1}^n \varrho_i(\xi_i)$ das Tensorprodukt über die standardisierten Dichtefunktionen aus Tabelle 2.1, entsprechend der transformierten Input-Zufallsvariablen. Ω ist das kartesische Produkt über die Definitionsbereiche Ω_n aus Tabelle 2.1.

Der Nenner in Gleichung (2.4) ist für die verwendeten Polynombasen leicht analytisch lösbar und durch eine geschlossene Lösung gegeben. Der Aufwand der Berechnung der Koeffizienten α_j liegt also in der Lösung der mehrdimensionalen Integrale.

In der obigen Definition ist der Definitionsbereich Ω unter Umständen in einer oder mehreren Dimensionen unbeschränkt, zur numerischen Lösung des Integrals muss die Verteilung, und damit der Definitionsbereich in dieser Dimension, auf einen endlichen Bereich begrenzt werden. Hier wird das so realisiert, dass die Streuintervalle der betroffenen Verteilungen betrachtet werden und der neue Definitionsbereich so gewählt wird, dass der Großteil der Werte, die von der Verteilung angenommen werden, innerhalb des Intervalls liegen. Der Anteil der Werte außerhalb des Intervalls ist dann kleiner als die Maschinengenauigkeit und vernachlässigbar.

Die numerische Integration des Zählers in Gleichung (2.4) unterliegt dem sogenannten Fluch der Dimensionalität, der Aufwand der Integration mit klassischen Quadratur-Verfahren steigt exponentiell mit der Anzahl der Dimensionen. Die Zahl der Auswertungen der Zielfunktion steigt also exponentiell in der Anzahl der Dimensionen an, dabei können die einzelnen Auswertungen unter Umständen sehr ineffizient sein. Dementsprechend sind diese Verfahren nur bedingt anwendbar. Bereits ab fünf Dimensionen [Pfl10] kann der geschätzte Aufwand zu groß sein, da die Zahl der benötigten Auswertungen für akkurate Ergebnisse zu hoch ist. Eine Möglichkeit mit diesem Fluch der Dimensionalität umzugehen ist die numerische Quadratur mittels regulärer oder räumlich adaptiver Sparse Grids. Sie wird in Kapitel 3 genauer vorgestellt.

2.2.7 Momentanalyse

Wie bereits in der Einleitung angeführt ist ein großer Vorteil der Polynomial Chaos Expansion die Berechnung der stochastischen Momente der Zielfunktion direkt aus den Koeffizienten der Expansion. Dies ist aufgrund der Konstruktion der Polynombasis möglich. Wegen der Orthogonalität der Basis lässt sich das erste Moment, der Erwartungswert E , folgendermaßen berechnen:

$$E(\hat{R}) = E\left(\sum_{j=0}^P \alpha_j \Psi_j(\xi)\right) = \alpha_0$$

Das zweite Moment, die Varianz Var , lässt sich analog als Summe der Quadrate aller Koeffizienten außer α_0 berechnen:

$$Var(\hat{R}) = E\left((\hat{R} - \alpha_0)^2\right) = \sum_{j=1}^P \alpha_j^2$$

Bei der Varianz ist zu beachten, dass sie bei der Berechnung über die Polynomial Chaos Expansion unterschätzt wird, da hier \hat{R} verwendet wird, also die trunkeerte Expansion. Dadurch wird die Varianz nur über einen Bruchteil der Koeffizienten berechnet.

Auch das dritte und vierte Moment, also die Schiefe und Wölbung der Zielfunktion lassen sich aus den Koeffizienten berechnen, dafür ist jedoch noch der Erwartungswert über das Produkt mehrerer mehrdimensionaler Polynombasen nötig. Dieser Erwartungswert hat jedoch nicht für alle Basisfunktionen aus dem Askey-Schema eine geschlossene Lösung und muss unter Umständen erneut numerisch bestimmt werden [Sud15].

3 Sparse Grids

Sparse Grids (dünne Gitter) stellen für viele Anwendungen eine Alternative zu den meist verwendeten uniformen isotropischen Tensorprodukt-Gittern (vollen Gittern) dar. Sie ermöglichen, zumindest zu einem bestimmten Grad, dem Fluch der Dimensionalität entgegenzuwirken. Das liegt daran, dass die Zahl der Gitterpunkte mit steigender Dimension bei regulären Sparse Grids viel langsamer wächst als die eines vollen Gitters. Gleichzeitig ist der Fehler, der durch die geringere Zahl an Punkten entsteht, sehr gering. Die Methode der räumlich adaptiven Sparse Grids ermöglicht unter Umständen eine noch bessere Konvergenz des Fehlers als reguläre Sparse Grids. Im folgenden werden reguläre und adaptive Sparse Grids nur kurz vorgestellt, eine ausführlichere Beschreibung kann in [Pfl10], [Gar12] oder [Val19] gefunden werden.

3.1 Grundlagen

Im Folgenden werden entsprechend der Konvention alle Gitter über $[0, 1]^D$ definiert. Dies ist keine Einschränkung, da, wie in Abschnitt 2.2.6 erklärt, trunkierte Wahrscheinlichkeitsverteilungen verwendet werden. Damit ist das in Gleichung (2.4) gegebene Integral über einen endlichen Raum definiert und kann auf $[0, 1]^D$ normiert werden.

Ein eindimensionales Gitter des Levels $l \in \mathbb{N}$ ist durch $2^l + 1$ viele Punkte $x_{l,i}$ gegeben:

$$x_{l,i} := i \cdot h_l, \quad i = 0, \dots, 2^l$$

Das Intervall $[0,1]$ wird dabei in 2^l viele, gleichgroße Intervalle der Schrittweite $h_l = 2^{-l}$ geteilt. Die Menge der Gitterpunkte des Gitters des Levels l wird im Folgenden als Ω_l bezeichnet. Jedem Gitterpunkt wird eine Basisfunktion $\varphi_{l,i}$ zugeordnet. Über diese Basisfunktionen lässt sich nun ein Vektorraum V_l definieren:

$$V_l := \text{span}\{\varphi_{l,i} \mid i = 0, \dots, 2^l\}$$

Betrachtet man Gitter mit unterschiedlichen Leveln l , stellt man fest, dass die Menge der Punkte Ω_l eines Gitters von Level l eine Teilmenge der Punkte des Gitters von Level $l + 1$ bilden. Dabei sind für jedes Gitter Ω_l nur die Punkte $x_{l,i}$ mit ungeradem Index i nicht in der Punktmenge Ω_{l-1} enthalten. Die Indexmenge I_l beinhaltet den Index der Punkte $x_{l,i}$, die in keinem Gitter $\Omega_{\tilde{l}}, \forall \tilde{l} < l$, enthalten sind. Das sind, bis auf den Fall $I_0 = \{0, 1\}$, immer die ungeraden Indizes i . Das Gitter Ω_l lässt sich also als die Vereinigung der Gitter $\tilde{\Omega}_{l'}$, die jeweils nur die neuen Punkte eines Levels l' enthalten, darstellen:

$$\Omega_l = \tilde{\Omega}_1 \cup \dots \cup \tilde{\Omega}_l, \quad \tilde{\Omega}_{l'} := \{x_{l',i} \mid i \in I_{l'}\}$$

Diese hierarchische Darstellung gilt analog für die Basisfunktionen, da diese wie oben gezeigt einzelnen Gitterpunkten zugeordnet wurden. Der Vektorraum, der die Lineare Hülle der Basisfunktionen, die zu einer Indexmenge I_l gehören, aufspannt, wird als hierarchischer Unterraum W_l bezeichnet:

$$W_l := \text{span}\{\varphi_{l,i} \mid i \in I_l\}$$

Die Kombination der hierarchischen Unterräume ermöglicht folgende Konstruktion von V_l :

$$V_l = \bigoplus_{i=0}^l W_i$$

Diese Relation gilt nur wenn die verwendeten Basisfunktionen $\varphi_{l,i}$ bestimmte Glattheitsbedingungen erfüllen[[Val19]]. Funktionen $f(x) \in V_l$ lassen sich dann als Linearkombination dieser Basisfunktionen darstellen:

$$f(x) = \sum_{l'=0}^l \sum_{i \in I_{l'}} \alpha_{l',i} \varphi_{l',i}$$

Die Koeffizienten $\alpha_{l,i}$ werden als hierarchische Überschüße (hierarchical surplus) bezeichnet. Die Berechnung dieser ist abhängig von den verwendeten Basisfunktionen $\varphi_{l,i}$.

Im D -dimensionalen Fall werden die oben beschriebenen Definitionen mittels eines Multi-Indexes $\mathbf{l} = (l_1, \dots, l_D)$ erweitert. Das Gitter $\Omega_{\mathbf{l}}$ des Levels \mathbf{l} ist durch folgende Gitterpunkte gegeben:

$$x_{\mathbf{l},\mathbf{i}} := (i_1 h_{l_1}, \dots, i_D h_{l_D}), \quad i_k = 0, \dots, 2^{l_k}, k \in \{1, \dots, D\}$$

Diesen Gitterpunkten sind erneut Basisfunktionen zugeordnet, die nun durch ein Tensorprodukt gegeben sind:

$$\varphi_{\mathbf{l},\mathbf{i}} : [0, 1]^D \rightarrow \mathbb{R}, \quad \varphi_{\mathbf{l},\mathbf{i}}(\mathbf{x}) := \prod_{k=1}^D \varphi_{l_k, i_k}(x_k)$$

Analog zum eindimensionalen Fall ist der Vektorraum $V_{\mathbf{l}}$ definiert:

$$V_{\mathbf{l}} := \text{span}\{\varphi_{\mathbf{l},\mathbf{i}} \mid i_k = 0, \dots, 2^{l_k}, k \in \{1, \dots, D\}\}$$

Die Indexmenge $I_{\mathbf{l}}$ wird als kartesisches Produkt über die einzelnen Indexmengen jedes Multiindexeintrags gebildet:

$$I_{\mathbf{l}} := I_{l_1} \times \dots \times I_{l_D}, \quad \mathbf{l}_i \in \mathbf{l}$$

Die mehrdimensionalen hierarchischen Unterräume lassen sich wie folgt definieren:

$$W_{\mathbf{l}} := \text{span}\{\varphi_{\mathbf{l},\mathbf{i}} \mid \mathbf{i} \in I_{\mathbf{l}}\}$$

Diese Unterräume lassen sich wieder zu dem Vektorraum $V_{\mathbf{l}}$ zusammenfügen:

$$V_{\mathbf{l}} = \bigoplus_{l'=0}^{\mathbf{l}} W_{l'} \tag{3.1}$$

Wie auch im eindimensionalen Fall gilt diese Relation nur, wenn die zugrundeliegenden Basisfunktionen bestimmte Glattheitsbedingungen erfüllen. Eine Funktion $f_{\mathbf{l}} \in V_{\mathbf{l}}$ lässt sich nun folgendermaßen darstellen:

$$f_{\mathbf{l}} = \sum_{l'=0}^{\mathbf{l}} \sum_{i \in I_{l'}} \alpha_{l',i} \varphi_{l',i} \tag{3.2}$$

3.2 Sparse Grids

Die grundsätzliche Idee von regulären Sparse Grids basiert auf einer Abwandlung der Darstellung aus Gleichung (3.1), in der nur die Unterräume W_l verwendet werden, die besonders „wichtig“ sind. Die anderen werden weggelassen. Eine Möglichkeit zu beurteilen, welche Unterräume wichtig sind, ist die hierarchischen Überschüsse $\alpha_{l,i}$ aus Gleichung (3.2) zu betrachten. Erfüllen die gewählten Ansatzfunktionen und die zu interpolierende Funktion gewisse Eigenschaften, so ist es möglich abzuschätzen, wie viel ein Unterraum eines bestimmten Levels zur Interpolation beiträgt. Unterräume höheren Levels tragen weniger bei als die geringeren Levels, es ist also möglich sie wegzulassen, da der entstandene Fehler sehr gering ist. Im Folgenden werden Sparse Grids von Level n definiert, indem die Gleichung (3.2) diesem Sachverhalt entsprechend angepasst wird. Es werden also nur Unterräume W_l die zu Levels l gehören, bei denen die Summe der Tupeleinträge, also die 1-Norm $\|l\|_1$, einen bestimmten Wert n nicht überschreitet. Gleichung (3.1) wird dazu angepasst, der Index beschreibt, dass es sich um einen Sparse Grid Raum mit Level n handelt:

$$V_n^s := \bigoplus_{\substack{l' < l \\ \|l'\|_1 \leq n}} W_{l'} \quad (3.3)$$

Das reguläre Sparse Grid ist dann durch folgende Gitterpunkte gegeben:

$$\Omega_n^s := \bigcup_{\substack{l' < l \\ \|l'\|_1 \leq n}} \{x_{l',i} \mid i \in I_{l'}\}$$

Die Funktionen über dem Sparse Grid Raum V_n^s werden $f_n^s(x)$ genannt und können wie in Gleichung (3.2) dargestellt werden:

$$f_n^s = \sum_{\substack{l' < l \\ \|l'\|_1 \leq n}} \sum_{i \in I_{l'}} \alpha_{l',i} \varphi_{l',i}$$

Die Zahl der Gitterpunkte des dünnen Gitters liegt bei $O(h_n^{-1} \cdot \log(h_n^{-1})^{d-1})$, die des vollen Gitters bei $O(h_n^{-d})$. Der Approximationsfehler $\|f - f_n^s\|$ dieses Gitters liegt für geeignete Basisfunktionen im Bereich $O(h_n^2 \log(h_n^{-1})^{d-1})$ [Gar12]. Für volle Gitter liegt dieser Fehler bei $O(h_n^2)$. Das dünne Gitter erreicht also, für geeignete Basisfunktionen, fast die selbe Genauigkeit wie das volle Gitter, während die Zahl der Punkte deutlich geringer bleibt.

3.3 Räumliche Adaptivität

Auch reguläre Sparse Grids unterliegen zu einem gewissen Grad dem Fluch der Dimensionalität. Mit räumlich adaptiven Sparse Grids ist es unter Umständen möglich, eine noch geringere Zahl an Gitterpunkten für eine ähnliche Genauigkeit zu erreichen. Dies geschieht, indem versucht wird die Gitterpunkte entsprechend der zu interpolierenden Funktion zu platzieren. Regionen der zu interpolierenden Funktion, die bereits mit wenigen Gitterpunkten aufgelöst werden können, erhalten auch weniger Punkte, Regionen in denen mehr Punkte zur besseren Auflösung beitragen, erhalten mehr Punkte. Für manche Funktionen ist mit diesem Verfahren die Zahl der Gitterpunkte, die

benötigt wird um den gleichen Fehler wie reguläre Sparse Grids zu erreichen, geringer. Eine beispielhafte Gegenüberstellung eines regulären und eines räumlich adaptiven Sparse Grids ist in Abbildung 4.6 auf Seite 30 dargestellt.

Bei räumlich adaptiven dünnen Gittern werden dem bestehenden Gitter iterative Gitterpunkte hinzugefügt. Dies wird als Verfeinerung des Gitters, beziehungsweise eines Gitterpunktes, bezeichnet. Das hier verwendete Verfahren wählt hierzu die Gitterpunkte, die a-priori den größten Nutzen versprechen. Das Gitter ist dabei folgendermaßen definiert:

$$\Omega := \{x_{l,i} \mid (l,i) \in \mathcal{I}\}$$

Dabei ist \mathcal{I} eine Indexmenge von Level/Index Paaren. Die Indexmenge $\mathcal{I}^{ref} \subseteq \mathcal{I}$ beinhaltet alle Level/Index Paare, bei denen der zugehörige Gitterpunkt noch weiter verfeinert werden kann. Dies ist der Fall, falls noch nicht alle Indizes der Kinder des Gitterpunkts in der Menge \mathcal{I} enthalten sind. Als Kinder eines Gitterpunktes $x_{l,i}$ werden jene Gitterpunkte $x_{l',i'}$ bezeichnet, für die es ein $r \in \{1, \dots, D\}$ gibt, so das folgende Eigenschaften erfüllt sind:

$$\begin{aligned} l_d &= l'_d, & i_d &= i'_d, & \forall d \in \{1, \dots, D\} \setminus \{r\} \\ l'_r &= l_r + 1 \\ i'_r &\in \{2i_r - 1, 2i_r + 1\} \end{aligned}$$

Das Gitter Ω kann nun verfeinert werden, indem die Kinder der Punkte, die am meisten zur Interpolation von f beiträgt, dem Gitter hinzugefügt werden. Eine Möglichkeit diese Punkte zu finden ist das sogenannte „surplus-based criterion“ (Überschuss basierte Kriterium), bei dem der Punkt, dessen hierarchischer Überschuss betragsmäßig am größten ist, verfeinert wird. Diese Verfeinerung wird iteriert bis die gewünschte Zahl an Gitterpunkten erreicht ist. Je nach verwendeten Basisfunktionen gibt es weitere Anforderungen an die Punktmenge, zum Beispiel müssen oft alle Eltern eines Punktes im Gitter enthalten sein [Pfl10].

3.4 Ansatzfunktionen

Bisher wurden die Ansatzfunktionen $\varphi_{l,i}$ nicht näher definiert, es wurde immer nur angenommen, dass sie bestimmte Eigenschaften erfüllen sollen. Nun werden zwei Typen von Ansatzfunktionen vorgestellt, die bei der Erstellung der Ergebnisse benutzt wurden. Eine Ansatzfunktion, die oft verwendet wird, da sie sehr simpel aufgebaut ist und dennoch einige wichtige Eigenschaften erfüllt, ist der lineare B-Spline ersten Grades, auch lineare Hütchenfunktion genannt:

$$\varphi_{l,i}^1(x) := \max\left(1 - \left|\frac{x}{h_l} - i\right|, 0\right)$$

Die Hütchenfunktion ist jedoch nur eingeschränkt nutzbar. Andere Ansatzfunktionen erlauben eine bessere Approximation bestimmter Funktionen, oder ermöglichen, dass Eigenschaften der Zielfunktion, beispielsweise Differenzierbarkeit, auch für das Sparse Grid Surrogat erhalten bleiben.

Beispielsweise können B-Splines höheren Grades genutzt werden. Ein B-Spline $b_{i,\xi}^p$ von Grad p mit Index i lässt sich entsprechend der Cox-de-Boor Rekursionsformel folgendermaßen berechnen:

$$b_{i,\xi}^p(x) := \begin{cases} \frac{x-\xi_i}{\xi_{i+p}-\xi_i} b_{i,\xi}^{p-1}(x) + \frac{\xi_{i+p+1}-x}{\xi_{i+p+1}-\xi_{i-1}} b_{i+1,\xi}^{p-1}(x), & p \geq 1 \\ \chi_{[\xi_i, \xi_{i+1}]}(x), & p = 0 \end{cases}$$

Hier ist $\xi := (\xi_0, \dots, \xi_{m+p})$, $m \in \mathbb{N}$ eine Folge wachsender reeller Zahlen (Knotenfolge) und $\chi_{[\xi_i, \xi_{i+1}]}(x)$ eine Funktion, die genau im Intervall $[\xi_i, \xi_{i+1}]$ den Wert eins annimmt, sonst null [RP19]. Für die folgenden Ergebnisse wurden die B-Splines zusätzlich modifiziert, es wurden Not-a-Knot (NaK) B-Spline Ansatzfunktionen verwendet. Dafür wird der Vektor von Knotenpunkten ξ und die Definition der Splines angepasst [Val19].

Eine Problematik, die bisher vernachlässigt wurde, ist die Randbehandlung. Verwendet man reguläre Sparse Grids ohne speziell angepasste Ansatzfunktionen, so wächst die Zahl der Punkte auf dem Rand mit steigendem Level oder steigender Dimension stark an. Es gibt verschiedene Möglichkeiten dem entgegenzuwirken. Beispielsweise kann die Schrittweite der Gitterpunkte auf dem Rand größer gewählt werden. Eine andere Möglichkeit, die auch für die Erstellung einiger Ergebnisse verwendet wurde, ist es die Randpunkte wegzulassen und modifizierte Ansatzfunktionen zu verwenden. Bei den verwendeten Ansatzfunktion handelt es sich um sogenannte extended NaK B-Splines. Bei diesen werden die Randpunkte weggelassen, mithilfe der ihnen zugeordneten Splines werden jedoch die Splines der inneren Knoten angepasst [RP20].

3.5 Integration mittels Sparse Grids und Stochastic Collocation

3.5.1 Integration

Für die Integration einer Funktion $f(\mathbf{x})$ mithilfe eines regulären oder adaptiven Sparse Grids wird zuerst der jeweilige Interpolant $f_n^s(\mathbf{x})$ erzeugt. Die Integration geschieht nun mithilfe von eindimensionalen Gauß-Legendre Quadraturregeln. Diese werden mithilfe eines Tensorproduktes auf den mehrdimensionalen Fall angepasst. Das Integral aus Gleichung (2.4) lässt sich also folgendermaßen, mit einem Sparse Grid Interpolanten \tilde{R} für die Zielfunktion R , approximieren:

$$\int_{\Omega} \tilde{R}(\xi) \Psi_j(\xi) \varrho(\xi) d\xi \approx \sum_k \tilde{R}(\mathbf{x}_k) \Psi_j(\mathbf{x}_k) \varrho(\mathbf{x}_k) \omega_k,$$

mit \mathbf{x}_k den Stützstellen und ω_k den Gewichten aus der Quadraturregel. Die verwendete Quadraturordnung der Gauß-Legendre Quadratur ist dabei abhängig von den verwendeten Sparse Grid Ansatzfunktionen und der Dichtefunktion ϱ zu wählen [RP20].

3.5.2 Stochastic Collocation

Das Verfahren der Stochastic Collocation erweitert klassische Kollokationsverfahren auf stochastische Systeme. Im folgenden wird nur der Interpolationsansatz der Stochastic Collocation vorgestellt. Es wird erneut ein Surrogat der Zielfunktion gebildet. Im Vergleich zur PCE sind jedoch die Polynombasen nicht festgelegt, sie werden durch Interpolationspolynome ersetzt. Es werden N

Kollokationspunkte verwendet, jedem Kollokationspunkt wird ein mehrdimensionales Polynom L_j zugewiesen. Die Werte der Zielfunktion an den Kollokationspunkten werden als r_j bezeichnet. Die Approximation der Zielfunktion nimmt nun folgende Form an:

$$R(\xi) \cong \tilde{R}(\xi) \sum_{j=0}^N r_j L_j(\xi)$$

Die mehrdimensionale Interpolation kann mittels Sparse Grid gestützten Methoden realisiert werden. Die stochastischen Momente lassen sich nun über die üblichen Formulierungen als Integral über das Surrogat berechnen [RP20]. Der Erwartungswert:

$$E(R) \cong E(\tilde{R}) = \int_{[0,1]^D} \tilde{R}(\xi) \varrho(\xi) d\xi \approx \sum_{j=0}^N \tilde{R}(x_j) \varrho(x_j) \omega_j$$

Und die Varianz:

$$\text{Var}(R) \cong \text{Var}(\tilde{R}) = E(\tilde{R}^2) - E(\tilde{R})^2$$

4 Ergebnisse

Die Methode der Polynomial Chaos Expansion wurde im Rahmen dieser Arbeit in dem Open-Source Software-Framework SG++¹ [Pfl10] implementiert. SG++ ist eine Toolbox, die eine Vielzahl an Algorithmen und Datenstrukturen für die Arbeit mit Sparse Grids bereitstellt. Die Implementierung der Methode ermöglicht die Berechnung der Polynomial Chaos Expansion einer Zielfunktion für beliebige Verteilungen aus Tabelle 2.1, wobei diese nicht zwingend standardisiert sein müssen. Zusätzlich können auch logarithmisch normalverteilte Zufallsvariablen verwendet werden. Die Berechnung der Koeffizienten geschieht mithilfe des vorgestellten Spectral Projection Ansatzes, die Integrale können mit regulären und adaptiven Sparse Grids gelöst werden.

4.1 Methoden

Zunächst werden noch die Metriken vorgestellt, die für die Auswertung der Ergebnisse verwendet wurden. Der Approximationsfehler des Sparse Grid Surrogats f_n^s (sowohl für reguläre als auch für adaptive Gitter) wird durch eine zufällige Stichprobe bestimmt (Monte-Carlo). Dazu wird eine Menge von Punkten $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ uniform gleichverteilt gewählt und der mittlere quadratische Fehler folgendermaßen bestimmt:

$$MSE := \frac{1}{n} \sum_{i=1}^n (f(x_i) - f_n^s(x_i))^2$$

Der Integrationsfehler E_I der verwendeten Integrationsmethode wird jeweils durch den Betrag des numerischen Integrationsfehlers $E(f)$ bestimmt. Er wird folgendermaßen berechnet:

$$E_I = |E(f)| := \left| \int f(\mathbf{x}) d\mathbf{x} - Q(f) \right|$$

wobei $Q(f)$ das Ergebnis der Quadraturmethode ist, $E(f)$ ist die Abweichung von der analytischen Lösung. Der Approximationsfehler des PCE Surrogats wird wieder mithilfe des mittleren quadratischen Fehlers gemessen. Dabei werden die zufällig gewählten Punkte \mathbf{x}_i diesmal entsprechend der transformierten zugrundeliegenden Verteilungen der Polynomial Chaos Expansion gewählt.

$$MSE_{PCE} := \frac{1}{n} \sum_{i=1}^n \left(R(\mathbf{x}_i) - \sum_{j=0}^P \alpha_j \Psi_j(\mathbf{x}_i) \right)^2$$

¹<https://github.com/SGpp/SGpp>

Zusätzlich zu diesen internen Metriken werden verschiedene Ergebnisse dieser PCE Implementierung mit denen der PCE aus dem Open-Source Optimisierungs- und Uncertainty Quantification-Framework DAKOTA [ABD+] verglichen. Außerdem werden diese Ergebnisse noch denen der Methode der Stochastic Collocation gegenübergestellt, die bereits in der SG++ Toolbox implementiert war.

4.2 Gering-dimensionale Beispiele

Polynom von Grad 2

Nun werden zuerst die Ergebnisse der Methode für verschiedene standardisierte Verteilungen anhand Funktionen geringer Dimensionalität dargelegt. Zuerst sei die Zielfunktion $R(\xi)$, $\xi = (\xi_1, \xi_2)$ folgendes Polynom zweiten Grades:

$$R(\xi) = 0.05(\xi_1 - 2)^2 + 2(\xi_2 + 1) \quad (4.1)$$

Da es sich um ein Polynom zweiten Grades handelt, wird die Funktion bereits für die Expansions-Ordnung $p = 2$ genau wiedergegeben, falls die Koeffizienten genau bestimmt werden. Diese Funktion kann außerdem im Sparse Grid Surrogat durch die NaK B-Spline Ansatzfunktionen von Grad drei genau dargestellt werden. Die Verteilungen für die Ergebnisse dieser Zielfunktion sind in Tabelle 4.1 gegeben:

Verteilung	Zufallsvariablen
Normal	$\xi_1, \xi_2 \sim \mathcal{N}(0, 1)$
Uniform	$\xi_1, \xi_2 \sim \mathcal{U}(-1, 1)$
Beta	$\xi_1, \xi_2 \sim Be(1, 2)$
Exponential	$\xi_1, \xi_2 \sim Exp(1)$
Gamma	$\xi_1, \xi_2 \sim \Gamma(2, 1)$

Tabelle 4.1: Verteilung der Zufallsvariablen für Zielfunktion (4.1)

In Abbildung 4.1a wird exemplarisch die Konvergenz des Approximationsfehlers des PCE Surrogats für normalverteilte Parameter gezeigt. Es sind die Ergebnisse für verschiedene Expansions-Ordnungen, Ansatzfunktionen und Gitter-Typen dargestellt. Auch die entsprechenden Ergebnisse der Stochastic Collocation sind abgebildet. Die beiden verwendeten Ansatzfunktionen sind extended NaK B-Splines von Grad eins und drei. Es wurden jeweils sowohl reguläre als auch adaptive Sparse Grids verwendet. Für diese und alle folgenden Ergebnisse wurden bei den räumlich adaptiven Sparse Grids in jeder Iteration zehn Punkte gleichzeitig verfeinert. Die Genauigkeit der Approximation des PCE Surrogats wird hierbei dem Integrationsfehler in Abbildung 4.1b gegenübergestellt. Es ist ersichtlich, dass die Genauigkeit der Integration unter sonst gleichen Bedingungen den Approximationsfehler des PCE Surrogats direkt beeinflusst. Für die Expansions-Ordnung $p = 2$ werden hier die besten Ergebnisse erreicht. Dies liegt daran, dass mit Ordnung $p = 1$ die Zielfunktion noch nicht genau dargestellt werden kann, für die Ordnung $p = 5$ allerdings Oszillationen auftreten können.

Die Konvergenz des relativen Fehlers des berechneten Erwartungswerts und der Varianz sind in Abbildung 4.2a dargestellt. Die Parameter der Funktion sind erneut normalverteilt. Für den Erwartungswert sind die Ergebnisse der Polynomial Chaos Expansion und der Stochastic Collocation für adaptive und reguläre Sparse Grids, jeweils mit extended NaK B-Splines von Grad eins und drei dargestellt. Die Expansions-Ordnung der PCE ist für die Berechnung des Erwartungswerts nicht relevant. Die Ergebnisse mit adaptiven Gittern sind für alle Ansatzfunktionen sowohl für die Polynomial Chaos Expansion, als auch für die Stochastic Collocation besser. Im direkten Vergleich sind die Ergebnisse der Stochastic Collocation für diese Zielfunktion besser als die Ergebnisse der PCE bei sonst gleichen Faktoren. Für diese Ergebnisse lässt sich das durch eine andere Verfeinerung erklären, für die Ergebnisse der PCE wurden immer zehn Gitterpunkte pro Schritt verfeinert, bei der Stochastic Collocation waren es drei. In Abbildung 4.2b ist die Konvergenz des relativen Fehlers der Berechnung der Varianz abgebildet. Dabei wurden für alle Ergebnisse extended NaK B-Splines von Grad drei verwendet. Die Stochastic Collocation erreicht hier deutlich bessere Ergebnisse, sowie eine schnellere Konvergenz, als die Polynomial Chaos Expansion.

In Abbildung 4.3 ist die Konvergenz des Approximationsfehlers des PCE Surrogats und die relativen Fehler der Berechnung des Erwartungswerts und der Varianz für die verschiedenen Verteilungen aus Tabelle 4.1 dargestellt. Die berechneten Erwartungswerte und Varianzen wurden für diese Funktion mit den analytischen Lösungen verglichen. Es wurden extended NaK B-Spline Ansatzfunktionen von Grad drei verwendet. Von den getesteten Ansatzfunktionen lieferten sie die besten Ergebnisse für diese Zielfunktion. Die Expansions-Ordnung ist $p = 2$. Es ist anzumerken, dass die verschiedenen Verteilungen aufgrund unterschiedlicher Definitionsbereiche nicht direkt verglichen werden können. Die Polynomial Chaos Expansionen basierend auf Legendre und Hermite Polynomen erreichen hier, verglichen mit den anderen Verteilungen, eine bessere Konvergenz.

Zweidimensionale trigonometrische Funktion

Als nächstes werden die Ergebnisse mit einer weiteren zweidimensionalen Funktion, jetzt mit trigonometrischen Elementen vorgestellt:

$$R(\xi) = \sin(\xi_1) + 3 \xi_1 \sin^3(\xi_2) \quad (4.2)$$

Sie ist in Abbildung 4.4 dargestellt. Da diese Funktion eine Sinusfunktion enthält kann sie nicht mehr exakt durch Polynome dargestellt werden, sowohl vom PCE Surrogat, als auch durch die verwendeten Ansatzfunktionen im Sparse Grid Interpolant. Die Verteilungsfunktionen werden hier so gewählt, dass die Parameter nur Werte in $[-1, 1]$ annehmen können. Die Verteilungen für die Analyse dieser Zielfunktion sind in Gleichung (4.3) aufgeführt.

$$\xi_1, \xi_2 \sim \mathcal{N}(0, 0.1) \quad (4.3)$$

Hier wird der Approximationsfehler des PCE Surrogats dem Approximationsfehler des Sparse Grid Surrogats gegenübergestellt. Bei der vorigen Funktion wurde statt dem Approximationsfehler des Sparse Grid Surrogats der Integrationsfehler betrachtet, da bei jener der Approximationsfehler mit den verwendeten Ansatzfunktionen null beträgt. Für die hier behandelte Funktion ist der Approximationsfehler des Sparse Grid Surrogats jedoch nicht null, die Konvergenz ist in Abbildung 4.5b abgebildet. Dieser wird hier dem Approximationsfehler des PCE Surrogats für normalverteilte

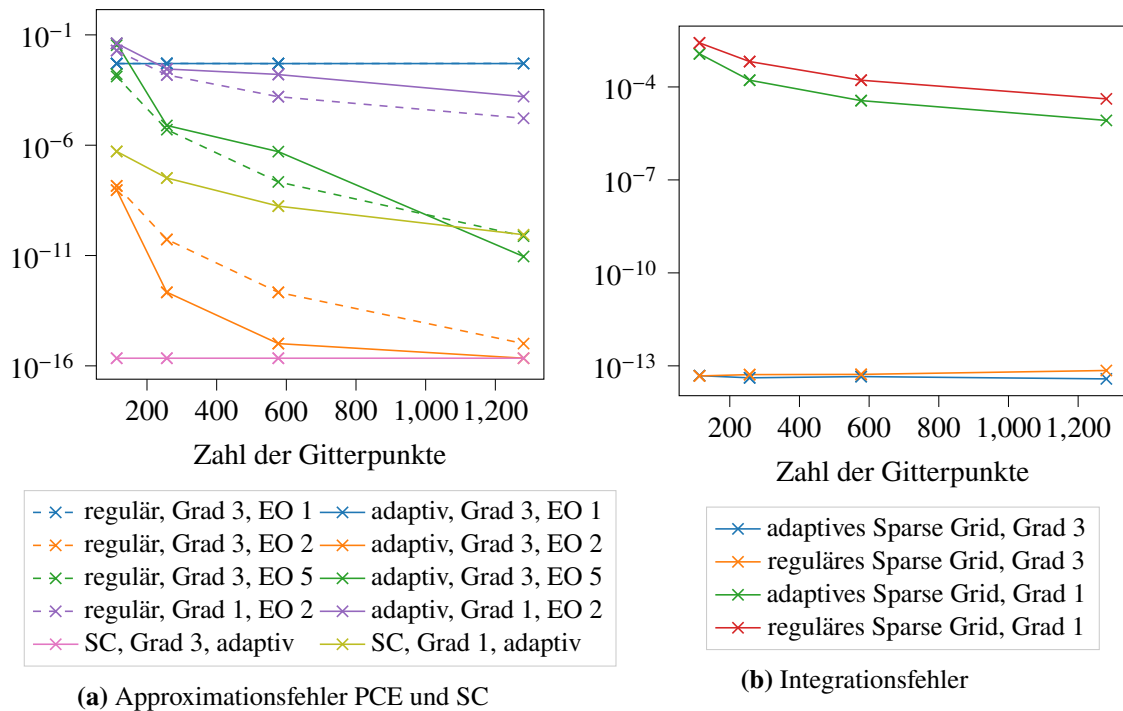


Abbildung 4.1: Approximationsfehler der UQ Methoden und Integrationsfehler der verwendeten Sparse Grids mit extended NaK Ansatzfunktion verschiedenen Grades. Ergebnisse der PCE mit Expansions-Ordnung (EO) eins, zwei und fünf. Ergebnisse für Gleichung (4.1).

Parameter entsprechend Gleichung (4.3) gegenübergestellt. Dabei gilt zu beachten, dass der PCE Approximationsfehler entsprechend der zugrundeliegenden Verteilungsfunktion ermittelt wurde, während der Approximationsfehler des Sparse Grid Surrogats uniform verteilt bestimmt wurde. Die hier schlechtere Konvergenz des PCE Approximationsfehlers mit adaptiven Sparse Grids lässt sich dadurch erklären, dass die iterativ hinzugefügten Gitterpunkte keine Informationen über die Verteilungen berücksichtigen und nur angepasst an die Zielfunktion und die Polynombasis auf dem gesamten Definitionsbereich verteilt werden. Es kann also vorkommen, dass die mehrdimensionale Dichtefunktion der Verteilungen an einer Stelle, an der besonders viele Gitterpunkte adaptiv platziert wurden, nahezu null ist und die genaue Auflösung an dieser Stelle nur eine geringe Auswirkung auf die stochastischen Momente hat.

Modifizierte zweidimensionale trigonometrische Funktion

Die letzte gering-dimensionale Funktion, die mit den Ergebnissen der vorigen beiden verglichen wird, ist die folgende:

$$R(\xi) = \sin(\xi_1) + 3 \xi_1 \sin^3(\xi_2) + 5 \cdot e^{-100 \cdot ((\xi_1 - 0.3)^2 + (\xi_2 - 0.3)^2)} \quad (4.4)$$

Der betrachtete Definitionsbereich ist erneut $[-1, 1]^2$. Sie ist aus der vorigen Zielfunktion gebildet, auf diese wurde noch eine Funktion addiert, die an dem Punkt $(0.3, 0.3)$ eine „Spitze“ hat und die fast im ganzen restlichen Definitionsbereich den Wert null annimmt. Sie ist in Abbildung 4.7

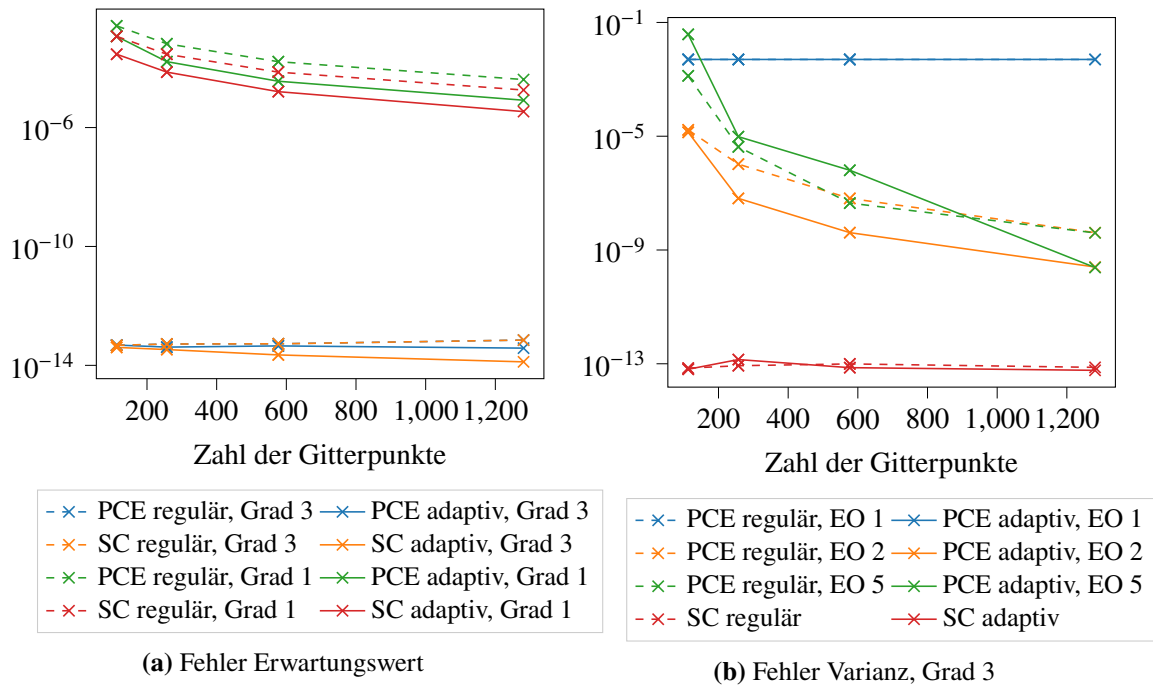


Abbildung 4.2: Fehler Erwartungswert und Varianz Normalverteilung für PCE und SC. Es wurden extended NaK B-Splines verschiedenen Grades verwendet. Ergebnisse für Gleichung (4.1).

dargestellt. Wieder enthält die Funktion trigonometrische Elemente und ist somit weder im PCE Surrogat, noch im Sparse Grid Surrogat mit den hier verwendeten Ansatzfunktionen exakt darstellbar. Die hier verwendeten Verteilungsfunktionen entsprechen wieder denen aus Gleichung (4.3). In Abbildung 4.6 sind das reguläre und das adaptive Sparse Grid mit jeweils 300 Gitterpunkten dargestellt. Hierbei ist die Skalierung zu beachten: Der Definitionsbereich der Funktion ist $[-1, 1]^2$, er wurde auf $[0, 1]^2$ skaliert. Wie angenommen werden bei dem adaptiven Sparse Grid mehr Punkte im Bereich der „Spitze“ hinzugefügt.

4.3 Short Column

Als nächstes wird die Polynomial Chaos Expansion auf ein Beispielproblem aus der Uncertainty Quantification angewandt. Die „Short Column“ Funktion behandelt ein stabförmiges Bauteil mit einem konstanten rechteckigen Querschnitt, mit Breite $b = 5$ und Länge $h = 15$. Die Festigkeit Y dieses Bauteils ist unbekannt und wird durch eine Zufallsvariable modelliert. Auf dieses Bauteil wirken zwei Kräfte, das Biegemoment M und eine Druckkraft P , die ebenfalls durch Zufallsvariablen gegeben sind. Die Funktion modelliert das Versagen des Bauteils unter verschiedenen Kräften [EWC08]. Sie ist folgendermaßen definiert:

$$f(x) = 1 - \frac{4M}{bh^2Y} - \frac{P^2}{b^2h^2Y^2} \quad (4.5)$$

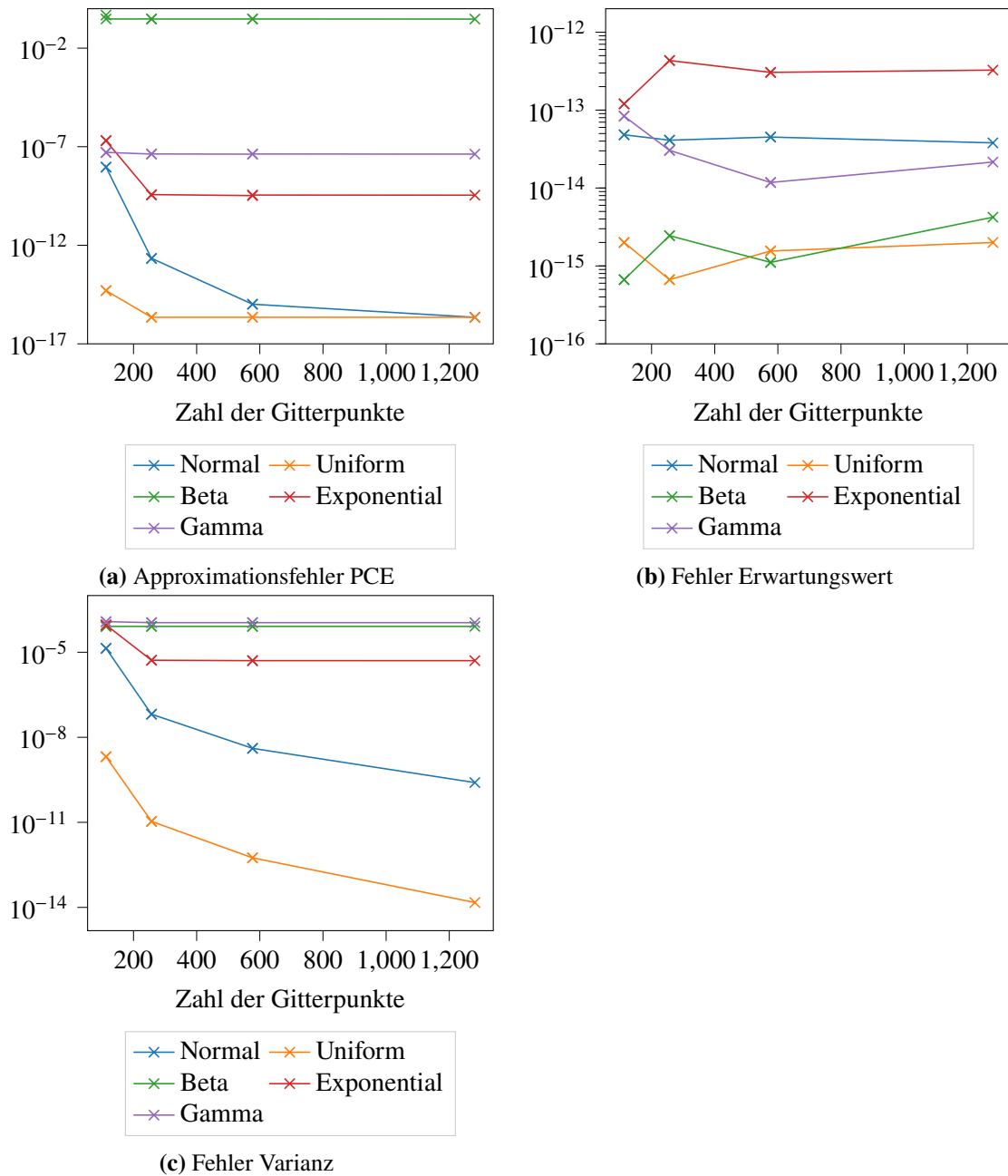


Abbildung 4.3: Approximationsfehler und relativer Fehler des Erwartungswerts und der Varianz für alle Verteilung aus Tabelle 4.1. Alle Ergebnisse für PCE mit adaptiven Sparse Grids und extended NaK B-Spline Ansatzfunktionen von Grad drei. Ergebnisse für Gleichung (4.1).

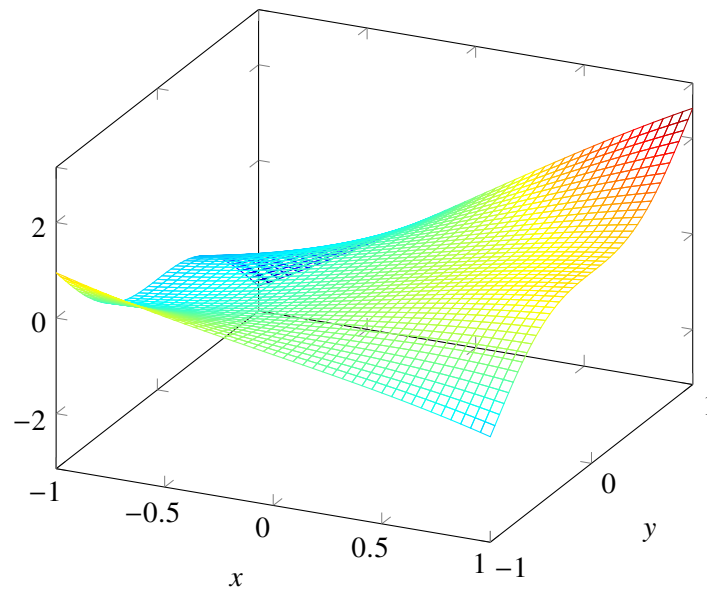


Abbildung 4.4: Gleichung (4.2)

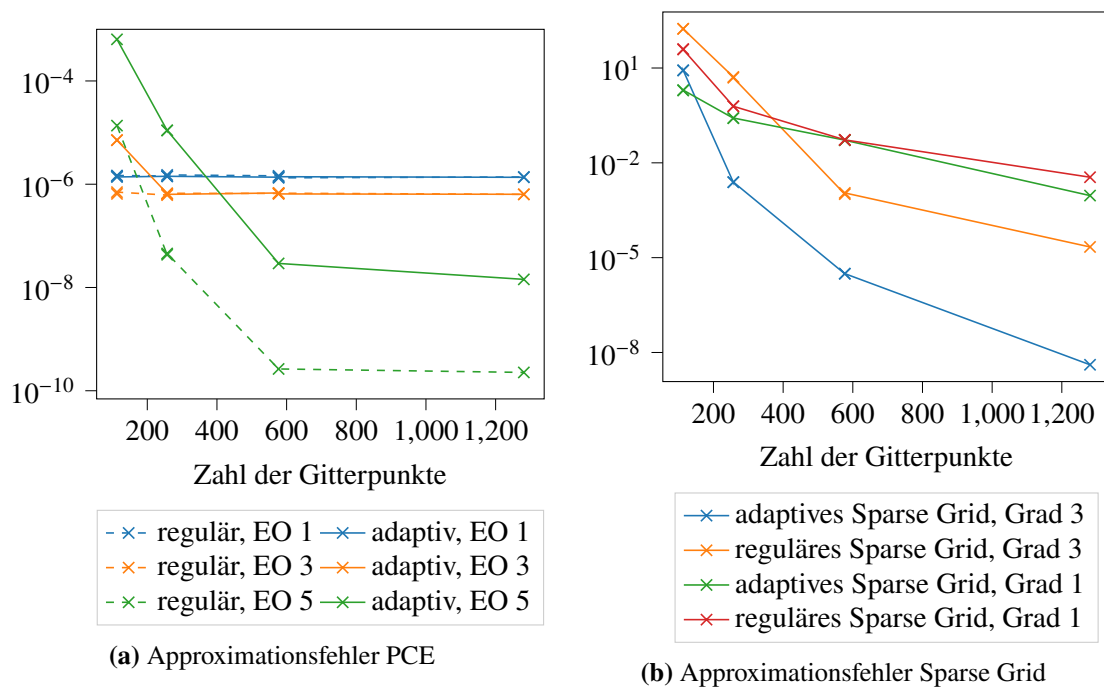


Abbildung 4.5: Approximationsfehler PCE für verschiedene Expansions-Ordnungen (EO) und Approximationsfehler Sparse Grid. Zu beachten: Die Fehler wurden unterschiedlich bestimmt. Ergebnisse für Gleichung (4.2).

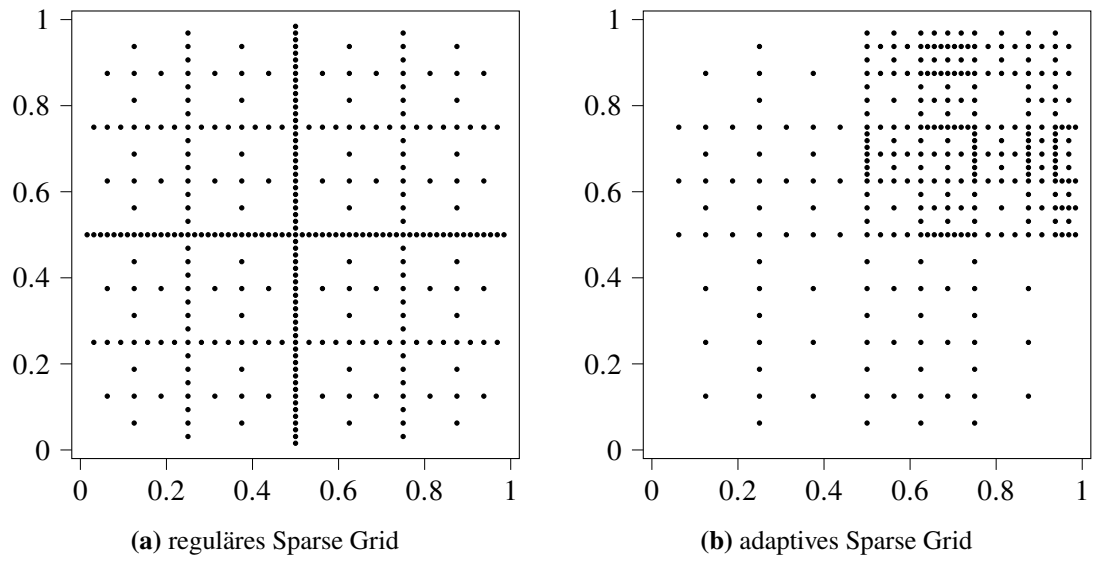


Abbildung 4.6: Vergleich eines regulären und adaptiven Sparse Grids für Gleichung (4.4)

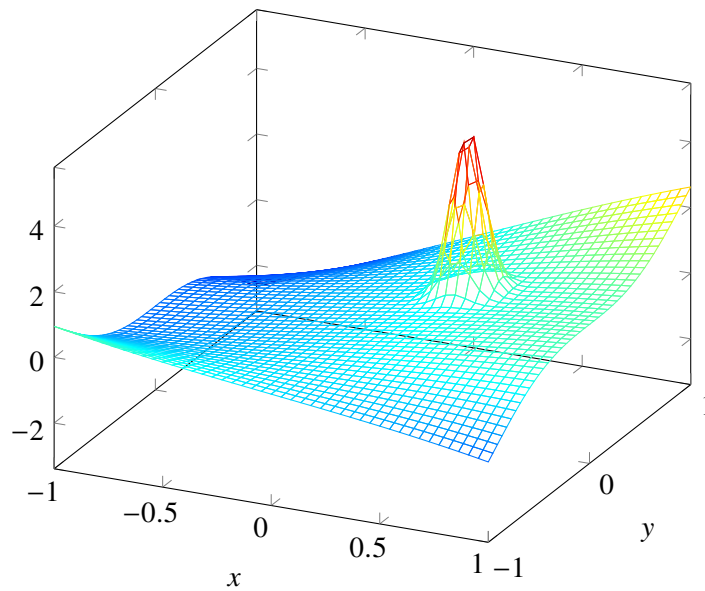


Abbildung 4.7: Gleichung (4.4)

Die Verteilungen der drei Parameter sind in Tabelle 4.2 gegeben. Für das hier gezeigte Beispiel sind die Parameter unabhängig voneinander.

Parameter	Zufallsvariablen
Festigkeit Y	$Y \sim \text{Log}(5, .5)$
Biegemoment M	$M \sim \mathcal{N}(2000, 400)$
Druckkraft P	$P \sim \mathcal{N}(500, 100)$

Tabelle 4.2: Verteilungen der Parameter der Short Column Function

Für diese Funktion wird zuerst der Approximationsfehler des PCE Surrogats mit dem Approximationsfehler der Stochastic Collocation verglichen. Dabei werden extended NaK B-Splines von Grad drei und fünf verglichen. Die Approximationsfehler des PCE Surrogats und der Stochastic Collocation für reguläre und adaptive Gitter ist in Abbildung 4.8a für B-Splines von Grad drei dargestellt. In Abbildung 4.8b werden die Approximationsfehler für B-Splines von Grad fünf gegenübergestellt. Für diese Zielfunktion ist der Unterschied zwischen extended NaK B-Splines von Grad drei und fünf für die Expansion-Ordnung $p = 5$ klar erkennbar. Dies liegt daran, dass die Splines von Grad drei das Produkt der Zielfunktion und der höhergradigen Basispolynome nicht mehr genau approximieren können. Für die anderen Expansions-Ordnungen lässt sich dies nicht beobachten, sie Konvergieren jedoch etwas schneller. Der Approximationsfehler der Stochastic Collocation ist wieder deutlich besser.

In Abbildung 4.9 sind der relative Fehler des Erwartungswerts und der Varianz dargestellt. Erneut werden die implementierte Polynomial Chaos Expansion und die Stochastic Collocation aus SG++ verglichen. Zusätzlich werden die Ergebnisse noch mit denen der Polynomial Chaos Expansion mit Sparse Grids aus DAKOTA verglichen. Die berechneten relativen Fehler für den Erwartungswert und die Varianz werden hier mithilfe einer Referenzlösung erstellt. Die Referenzlösung für den Erwartungswert wurde in DAKOTA ermittelt, dazu wurde der Erwartungswert mithilfe einer PCE auf einem Sparse Grid mit Level 8 berechnet. Diese Lösung wurde mit einer weiteren Referenzlösung, die mit einer PCE mit räumlich adaptiven Sparse Grids mit 3000 Gitterpunkten berechnet wurde, verglichen. Der absolute Fehler zwischen diesen beiden Lösungen ist kleiner als 10^{-12} . Die Referenzlösung für die Varianz wurde analog bestimmt, sie wurde jedoch mit der Stochastic Collocation verglichen. Der absolute Fehler zwischen den beiden Lösungen ist kleiner als 10^{-8} .

Der relative Fehler des Erwartungswerts der Polynomial Chaos Expansion mit adaptiven Sparse Grids konvergiert für dieses Problem besser als der des Erwartungswerts der Stochastic Collocation. Er konvergiert ähnlich wie die PCE mit regulären Sparse Grids. Der relative Fehler des Erwartungswerts der mit DAKOTA bestimmt wurde konvergiert am schnellsten. Auch bei der Berechnung der Varianz konvergiert DAKOTA für diese Funktion am schnellsten. Die Polynomial Chaos Expansion mit regulären Sparse Grids erreicht eine ähnlich gute Konvergenz wie die Stochastic Collocation mit räumlich adaptiven Gittern.

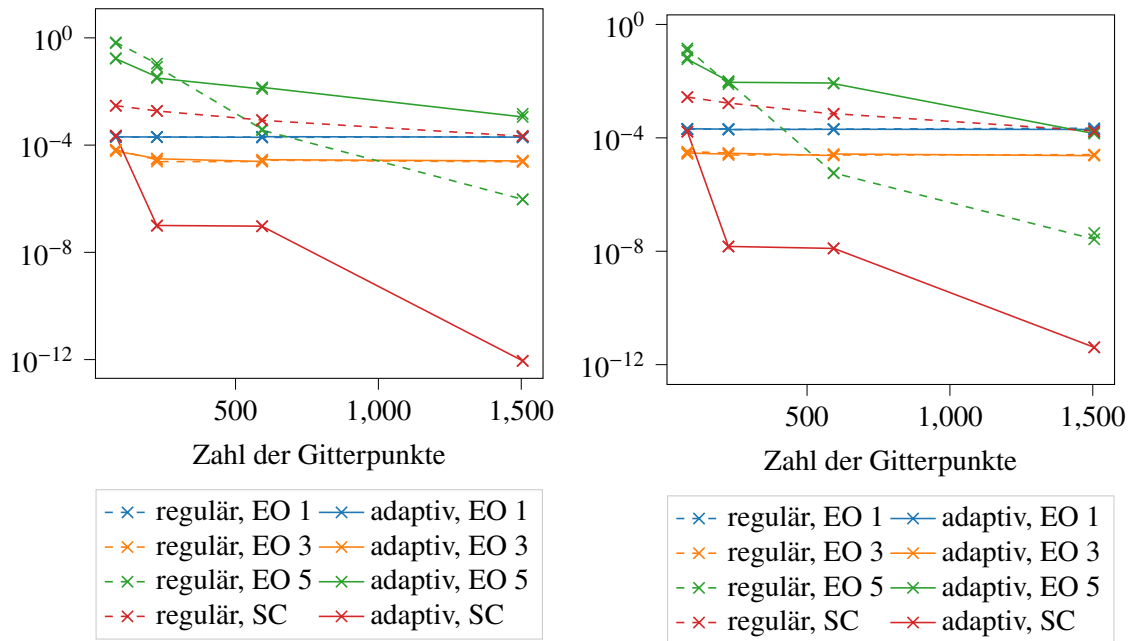


Abbildung 4.8: Approximationsfehler PCE und SC für extended NaK B-Splines von Grad drei und fünf. Ergebnisse für Gleichung (4.5).

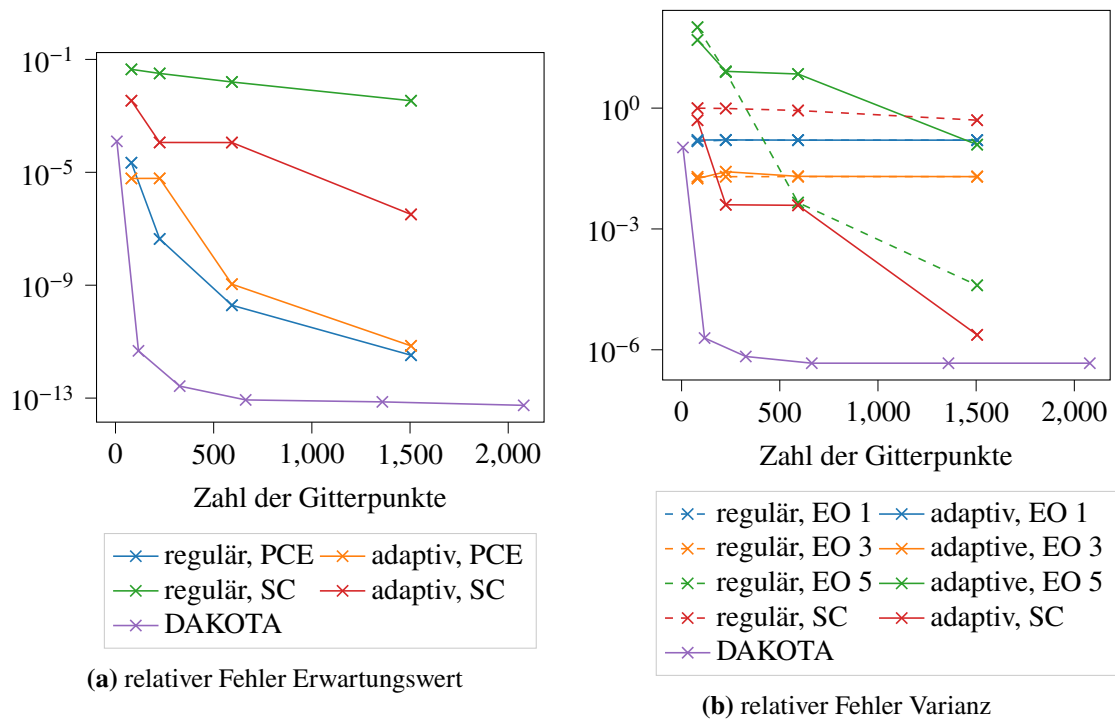


Abbildung 4.9: Relativer Fehler der Berechnung von Erwartungswert und Varianz, ermittelt mithilfe einer Referenzlösung. Ergebnisse für Gleichung (4.5).

5 Zusammenfassung und Ausblick

In dieser Arbeit wurde die Polynomial Chaos Expansion mit räumlich adaptiven und regulären Sparse Grids in dem Open-Source Software-Framework SG++ implementiert. Die Polynomial Chaos Expansion ist eine Methode aus der Uncertainty Quantification, bei der mithilfe von orthogonalen Polynomen ein Surrogat für eine Zielfunktion erstellt wird. Aus den Koeffizienten dieser Expansion lassen sich dann einfach und effizient die stochastischen Momente der Zielfunktion bestimmen. Um die Koeffizienten der Expansion zu berechnen, wurde ein Spectral Projection Ansatz verwendet. Bei diesem müssen hochdimensionale Integrale gelöst werden. Da traditionelle Integrationsmethoden unter dem Fluch der Dimensionalität leiden, wurde die Integration mithilfe von räumlich adaptiven und regulären Sparse Grids gelöst. Zuletzt wurden einige Ergebnisse der Implementierung vorgestellt und mit den Ergebnissen einer anderen Methode aus der Uncertainty Quantification, der Stochastic Collocation, verglichen. Diese Ergebnisse wurden zusätzlich noch mit denen der Polynomial Chaos Expansion mit Sparse Grids des Open-Source Optimisierungs- und Uncertainty Quantification Frameworks DAKOTA verglichen.

Die Ergebnisse zeigten, dass die Polynomial Chaos Expansion mit räumlich adaptiven Sparse Grids für bestimmte Funktionen eine verbesserte Konvergenz des Fehlers bieten kann als die Polynomial Chaos Expansion mit regulären Sparse Grids. Diese verbesserte Konvergenz gilt sowohl für den Approximationsfehler des PCE Surrogats, als auch für den Fehler bei der Berechnung der stochastischen Momente. Eine Ausnahme, bei der die adaptiven Sparse Grids im Kontext der Uncertainty Quantification schlechter abschneiden als die regulären, liegt vor, wenn die mehrdimensionale Dichtefunktion an der Stelle die adaptiv verfeinert wird einen vergleichsweise geringen Wert oder sogar null annimmt. Dies liegt daran, dass das verwendete Kriterium, nach welchem die Gitterpunkte verfeinert werden, keine Informationen über die zugrundeliegende Dichtefunktion berücksichtigt. Dass das hier verwendete Kriterium nur die Zielfunktion berücksichtigt hat auch Vorteile: Der Approximationsfehler des Sparse Grids ist dadurch besser, für diesen ist die Dichtefunktion nicht relevant.

Weiterführend kann untersucht werden, ob durch die Wahl eines anderen Kriteriums, welches Informationen über die Dichtefunktion verwendet, die Polynomial Chaos Expansion mit räumlich adaptiven Sparse Grids bessere Ergebnisse erzielen kann. Auch hier könnte man die Ergebnisse wieder mit denen einer Stochastic Collocation mit adaptiven Sparse Grids vergleichen, die auch entsprechend diesem Kriterium verfeinert.

Literaturverzeichnis

- [ABD+] B. Adams, W. Bohnhoff, K. Dalbey, M. Ebeida, J. Eddy, M. Eldred, G. Geraci, R. Hooper, P. Hough, K. Hu, J. Jakeman, M. Khalil, K. Maupin, J. Monschke, E. Ridgway, A. Rushdi, J. Stephens, L. Swiler, D. Vigil, T. Wildey, J. Winokur. *Dakota, A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis: Version 6.11 User's Manual* (zitiert auf S. 24).
- [EB09] M. Eldred, J. Burkardt. „Comparison of non-intrusive polynomial chaos and stochastic collocation methods for uncertainty quantification“. In: *47th AIAA aerospace sciences meeting including the new horizons forum and aerospace exposition*. 2009, S. 976 (zitiert auf S. 10, 13).
- [EWC08] M. Eldred, C. Webster, P. Constantine. „Evaluation of non-intrusive approaches for Wiener-Askey generalized polynomial chaos“. In: *49th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, 16th AIAA/ASME/AHS Adaptive Structures Conference, 10th AIAA Non-Deterministic Approaches Conference, 9th AIAA Gossamer Spacecraft Forum, 4th AIAA Multidisciplinary Design Optimization Specialists Conference*. 2008, S. 1892 (zitiert auf S. 27).
- [Gar12] J. Garcke. „Sparse grids in a nutshell“. In: *Sparse grids and applications*. Springer, 2012, S. 57–80 (zitiert auf S. 17, 19).
- [LK10] O. Le Maître, O. M. Knio. *Spectral methods for uncertainty quantification: with applications to computational fluid dynamics*. Springer Science & Business Media, 2010 (zitiert auf S. 9).
- [Pfl10] D. Pflüger. „Spatially Adaptive Sparse Grids for High-Dimensional Problems“. Diss. Dissertation, 2010 (zitiert auf S. 15, 17, 20, 23).
- [RP19] M. F. Rehme, D. Pflüger. „Active Subspaces with B-Spline Surrogates on Sparse Grids“. In: *3rd ECCOMAS Thematic Conference on Uncertainty Quantification in Computational Sciences and Engineering, Crete, Greece*. 2019 (zitiert auf S. 21).
- [RP20] M. F. Rehme, D. Pflüger. „Stochastic Collocation with hierarchical extended B-Splines on Sparse Grids“. In: *Proceedings of Approximation Theory XVI*. (to be published). 2020 (zitiert auf S. 21, 22).
- [Sud15] B. Sudret. „Polynomial chaos expansions and stochastic finite element methods“. In: *Risk and reliability in geotechnical engineering* (2015), S. 265–300 (zitiert auf S. 12, 13, 16).
- [Val19] J. Valentin. „B-splines for sparse grids: Algorithms and application to higher-dimensional optimization“. Diss. 2019 (zitiert auf S. 17, 18, 21).

- [XK02] D. Xiu, G. E. Karniadakis. „The Wiener–Askey polynomial chaos for stochastic differential equations“. In: *SIAM journal on scientific computing* 24.2 (2002), S. 619–644 (zitiert auf S. 10).

Alle URLs wurden zuletzt am 2. November 2020 geprüft.

Erklärung

Ich versichere, diese Arbeit selbstständig verfasst zu haben. Ich habe keine anderen als die angegebenen Quellen benutzt und alle wörtlich oder sinngemäß aus anderen Werken übernommene Aussagen als solche gekennzeichnet. Weder diese Arbeit noch wesentliche Teile daraus waren bisher Gegenstand eines anderen Prüfungsverfahrens. Ich habe diese Arbeit bisher weder teilweise noch vollständig veröffentlicht. Das elektronische Exemplar stimmt mit allen eingereichten Exemplaren überein.

Ort, Datum, Unterschrift