

# Robuste Optimierung mit Quantilmaßen auf globalen Metamodellen

Inaugural-Dissertation  
zur  
Erlangung des Doktorgrades  
der Mathematisch-Naturwissenschaftlichen Fakultät  
der Universität zu Köln

vorgelegt von

Beate Rhein

aus Geilenkirchen

Berichterstatter: Prof. Dr. Ulrich Trottenberg  
Prof. Dr. Wolfgang Wefelmeyer

Tag der mündlichen Prüfung: 15.1.2014

---

# Danksagung

---

Die vorliegende Dissertation ist zwischen 2008 und 2013 während meiner Zeit als wissenschaftliche Mitarbeiterin am CAD/CAM-Center Cologne der Fachhochschule Köln in Kooperation mit dem Fraunhofer Institut für Wissenschaftliches Rechnen und Algorithmen (SCAI) in Sankt Augustin entstanden.

Ich möchte mich bei meinem Gutachter an der Universität zu Köln, Herrn Prof. Dr. Ulrich Trottenberg, für die wohlwollende Betreuung meines Promotionsverfahrens bedanken. Mein Dank gilt ebenso Herrn Prof. Dr. Wolfgang Wefelmeyer für die bereitwillige Übernahme des Zweitgutachtens.

Ein großes Dankeschön geht an Frau Dr. Tanja Clees, die die wissenschaftliche Betreuung meines Dissertationsprojektes übernommen hat und dessen Fortgang durch viele kritische und inspirierende Fachdiskussionen bereichert hat. Bedanken möchte ich mich auch bei Frau Prof. Dr. Margot Ruschitzka, die die entscheidenden Kontakte herstellte und mich im Hinblick auf meine fachliche, berufliche und persönliche Weiterentwicklung stets gefördert hat.

Dem Team der Abteilung High Performance Analytics des SCAI danke ich für die anregenden Diskussionen und hilfreichen Hinweise. Mein Dank gilt auch meinen Kollegen am CAD/CAM-Center Cologne für die Anregungen zur Anwendbarkeit im Ingenieurbereich und die stets angenehme Arbeitsatmosphäre.

Von Herzen danken möchte ich meinem Mann Bernd und meinen Töchtern Julia und Belinda für die fortwährende Unterstützung während der gesamten Zeit.





---

# Kurzzusammenfassung

---

Die robuste Optimierung bezieht Unsicherheiten der Eingangsvariablen in den Optimierungsprozess ein und bestimmt Lösungen, die unempfindlich gegenüber diesen Schwankungen sind. Sie besitzt einen hohen Anwendungsbezug, da streuende Parameter, wie etwa Fertigungstoleranzen oder Schwankungen in den Materialeigenschaften, in vielen industriellen Anwendungen eine Rolle spielen.

In dieser Dissertation wird eine Methodik für die robuste Optimierung mit Quantilmaßen auf globalen Metamodellen entwickelt. Sie ist zugeschnitten auf die Situation, dass nur wenige Funktionsauswertungen zur Verfügung stehen. In der Praxis ist dies immer dann der Fall, wenn ein komplexes System- oder Prozessverhalten nur über aufwändige Computersimulationen oder Experimente erfasst werden kann.

Der Fokus der Arbeit liegt auf der geeigneten Charakterisierung des Robustheitsverhaltens des Gesamtsystems und seiner effizienten Berechnung. Es wird gezeigt, dass Quantilmaße den Industrieanforderungen an Robustheitsmaße besser genügen als die bisher benutzten Standardmaße. Sie finden geeignetere robuste Bereiche und liefern genauere untere und obere Grenzen für einen vom Anwender nach seinen Bedürfnissen festgelegten Prozentsatz der mittigen Ausgabeverteilung.

Die Methodik approximiert das Systemverhalten mit globalen Metamodellen. Auch das Robustheitsverhalten wird mit Metamodellen für Quantile der Ausgabeverteilungen modelliert. Die effiziente Berechnung der Quantilmaße erfolgt mit einem neuen Algorithmus, der die Haltonfolge mit dem Harrell-Davis-Schätzer kombiniert und eine interne Fehlerabschätzung enthält. Eine numerische Studie zeigt seine Effizienz. Es wird ein Maß für die Toleranz der Quantilmetamodelle entwickelt, das die Grenzen aus dem ursprünglichen Metamodell ableitet. Schließlich unterstützt ein neues iteratives Verfahren den Anwender bei der Auswahl des robusten Optimums. Die Methodik kann mit schnellen Multipolverfahren weiter beschleunigt werden. Dieser Weg wird beschrieben und sein Aufwand abgeschätzt.

Die Gültigkeit und Effizienz der Methodik wird an mehreren realen Anwendungsbeispielen aus der Automobilindustrie demonstriert.



---

# Abstract

---

Robust optimization integrates the input variables' uncertainty in the optimization process and determines solutions, which are insensitive to such variations. It has a high practical orientation, because scattering parameters like production tolerances or variations in the material properties play a role in many industrial applications. In this thesis a methodology for robust optimization with quantile measures on global metamodels is developed. It is tailored to the situation, that only a few function evaluations are available. In practical applications this is the case, whenever a complex system or process behaviour can only be modeled by expensive computer simulations or experiments.

The focus of the thesis is on characterizing the system's robustness in an appropriate way and on calculating it efficiently. It is demonstrated, that quantile measures meet the industry's requirements to robustness measures in a better way than the standard measures used up to now. They find more appropriate robust areas and give more accurate lower and upper limits for a certain percentage of the centred output distribution, which can be fixed according to the needs of the user.

The methodology approximates the system's behaviour with global metamodels. The robustness behaviour is modeled with metamodels for quantiles of the output distributions. The efficient quantile measure computation is done with a new algorithm, which combines Halton sequence and Harrell-Davis estimator and contains an internal error estimation. A numerical study demonstrates its efficiency. A measure for the tolerance of the quantile metamodels is developed, which derives its limits from the original metamodel. Finally a new iterative method supports the user in selecting the robust optimum. The methodology can be accelerated with fast multipole procedures. This way is described and the effort estimated.

The validity and efficiency of the methodology is demonstrated with several real application examples from automotive industry.



---

# Inhaltsverzeichnis

---

<b>Inhaltsverzeichnis</b>	<b>ix</b>
<b>1 Einleitung</b>	<b>1</b>
<b>2 Einführendes Beispiel</b>	<b>7</b>
2.1 Deterministische Optimierung . . . . .	9
2.2 Neuer Ansatz zur robusten Optimierung mit Quantilen . . . . .	11
2.3 Robuste Optimierung mit den Standardmaßen . . . . .	14
2.4 Fazit . . . . .	15
<b>I Methodik</b>	<b>17</b>
<b>3 Grundlagen</b>	<b>19</b>
3.1 Grundlagen aus Wahrscheinlichkeitsrechnung und Statistik . . . . .	19
3.1.1 Grundlagen aus der Wahrscheinlichkeitsrechnung . . . . .	19
3.1.2 Grundlagen aus der Statistik . . . . .	21
3.1.3 Funktionen von Zufallsvariablen . . . . .	21
3.1.4 Eigenschaften der übertragenen Verteilung . . . . .	23
3.2 Versuchspläne . . . . .	23
3.2.1 Monte Carlo-Verfahren . . . . .	25
3.2.2 Latin Hypercube Sampling . . . . .	25
3.2.3 Stratified Sampling . . . . .	26
3.2.4 Centered Stratified Sampling . . . . .	27
3.2.5 Haltonfolge . . . . .	27
3.2.6 Andere Versuchspläne . . . . .	28
3.2.7 Diskussion . . . . .	28
3.3 Nichtparametrische Quantilschätzer . . . . .	32
3.3.1 Empirischer Schätzer . . . . .	32
3.3.2 Kerndichteschätzer . . . . .	34
3.3.3 Umkehrung der Kerndichteverteilungsfunktion . . . . .	35
3.3.4 Glättung des empirischen Schätzers mit Kerndichtefunktionen . . . . .	36
3.3.5 Wahl der Bandweite . . . . .	37

3.3.6	Harrell-Davis-Schätzer . . . . .	38
3.3.7	Sfakianikis-Verginis-Schätzer . . . . .	39
3.3.8	$P^2$ -Algorithmus . . . . .	40
3.3.9	Schätzer für extreme Quantile . . . . .	41
3.3.10	Diskussion . . . . .	41
3.4	Weitere Methoden zur Schätzung von Quantilen . . . . .	43
3.4.1	Taylor-Entwicklung . . . . .	43
3.4.2	Konfidenzintervalle für den Standardschätzer . . . . .	44
3.5	Metamodelle . . . . .	45
3.5.1	Radiale-Basisfunktionen-Modell mit polynomiellm Anteil . . . . .	45
3.5.2	Validierung eines Metamodells . . . . .	46
3.6	Schnelle Multipolverfahren für RBF-Metamodelle . . . . .	47
3.6.1	Erstellung der Baumstruktur . . . . .	48
3.6.2	Multipole Acceptance Criterion . . . . .	49
3.6.3	Approximation der RBF-Summe . . . . .	49
3.6.4	Rekursionsformel für Taylor-Koeffizienten . . . . .	51
3.6.5	Aufwand des Algorithmus . . . . .	51
3.7	Multikriterielle Optimierung . . . . .	51
3.7.1	Lösung einer multikriteriellen Optimierungsaufgabe . . . . .	52
3.7.2	Visualisierung der Paretofront . . . . .	54
3.7.3	Multikriterielle Optimierung auf Metamodellen . . . . .	55
3.8	Clusteranalyse . . . . .	55
3.8.1	Partionierende Verfahren . . . . .	55
3.8.2	Hierarchische Verfahren . . . . .	56
<b>4</b>	<b>Robuste Optimierung</b> . . . . .	<b>57</b>
4.1	Begriff der Robustheit . . . . .	57
4.2	Ablauf einer robusten Optimierung . . . . .	59
4.3	Klassifizierung von Unsicherheiten . . . . .	61
4.4	Formulierung der robusten Optimierungsaufgabe . . . . .	62
4.4.1	Bei einkriteriellen Robustheitsmaßen . . . . .	62
4.4.2	Bei multikriteriellen Robustheitsmaßen . . . . .	62
4.5	Einkriterielle Robustheitsmaße . . . . .	63
4.5.1	Worst-Case-Maß . . . . .	63
4.5.2	Einseitiges Quantil-Maß . . . . .	64
4.5.3	Quantildifferenz-Maß . . . . .	65
4.5.4	Wahrscheinlichkeitsschwellen-Maß . . . . .	66
4.5.5	Erwartungswert-Nutzenfunktion-Maße . . . . .	66
4.5.6	Erwartungswert-Maß . . . . .	67
4.5.7	Standardabweichung-Maß . . . . .	67
4.5.8	Gewichtete Zielfunktion aus Lage- und Streuungsmaß . . . . .	69
4.5.9	Taguchis Signal-Rausch-Verhältnis . . . . .	70
4.5.10	Entropie . . . . .	71
4.6	Multikriterielle Robustheitsmaße . . . . .	72
4.6.1	Erwartungswert und Standardabweichung . . . . .	72
4.6.2	Median und Quantildifferenz . . . . .	73
4.7	Diskussion der Robustheitsmaße . . . . .	75

4.8	Lösung der robusten Optimierungsaufgabe . . . . .	76
4.9	Formulierungen für Nebenbedingungen . . . . .	77
4.10	Abgrenzung zu anderen Forschungsgebieten . . . . .	78
4.10.1	Zuverlässigkeitstheorie . . . . .	78
4.10.2	Design for Six Sigma . . . . .	78
4.10.3	Robustheit von Schätzverfahren . . . . .	79
<b>5</b>	<b>Berechnung von übertragenen Quantilen</b>	<b>81</b>
5.1	Vorüberlegungen . . . . .	81
5.2	Stichprobengröße . . . . .	82
5.3	Versuchsplan und Quantilschätzer . . . . .	85
5.4	Algorithmus für die Schätzung des Quantilmaßes . . . . .	86
<b>6</b>	<b>Robuste Optimierung mit Metamodellen</b>	<b>89</b>
6.1	Ablauf der entwickelten Methodik . . . . .	90
6.1.1	Schritt 1: Metamodell für $f$ . . . . .	91
6.1.2	Schritt 2: Metamodelle für das Quantilmaß . . . . .	91
6.1.3	Schritt 3: Bestimmung der robusten Optima . . . . .	92
6.1.4	Schritt 4: Visualisierung der Paretofront und Entscheidung . . . . .	92
6.2	Einfluss der Metamodelltoleranz auf die Quantilschätzung . . . . .	94
6.3	Diskussion . . . . .	96
<b>7</b>	<b>Aufwand des schnellen Multipolverfahrens</b>	<b>99</b>
7.1	Erstellung des Baumes . . . . .	99
7.1.1	Aufwandsabschätzung für den günstigsten Fall . . . . .	101
7.1.2	Aufwandsabschätzung für den schlechtesten Fall . . . . .	103
7.2	Cluster-Partikel-Algorithmus . . . . .	103
7.2.1	Aufwandsabschätzung für den günstigsten Fall . . . . .	105
7.2.2	Aufwandsabschätzung für den schlechtesten Fall . . . . .	108
7.3	Aufwandsabschätzung für die direkte Summation . . . . .	109
7.4	Vergleich . . . . .	110
7.5	Diskussion . . . . .	111
<b>II Numerische Untersuchungen</b>		<b>113</b>
<b>8</b>	<b>Übertragene Verteilung für einen Punkt</b>	<b>115</b>
8.1	Schätzung der Normalverteilung . . . . .	115
8.2	Testfälle . . . . .	117
8.2.1	Eindimensionale Testfälle . . . . .	118
8.2.2	Zweidimensionale Testfälle . . . . .	119
8.2.3	Anwendungsbeispiel Zweistab-Fachwerk . . . . .	121
8.2.4	Crashverhalten der B-Säule eines Pkw . . . . .	122
8.2.5	Anwendungsbeispiel pars10crits19 . . . . .	123
8.3	Ermittlung der Bandweite des Kerndichteschätzers . . . . .	123
8.3.1	Vorüberlegungen . . . . .	124
8.3.2	Numerische Simulation . . . . .	125
8.4	Vergleich der Quantilschätzer und Versuchspläne . . . . .	128

8.4.1	Vergleich der Versuchspläne . . . . .	129
8.4.2	Vergleich der Quantilschätzer bzgl. der Haltonfolge . . . . .	132
8.4.3	Vergleich von Harrel-Davis-Schätzer und Haltonfolge zu Standardschätzer und Monte-Carlo-Verfahren . . . . .	136
8.5	Test des iterativen Algorithmus . . . . .	138
<b>9</b>	<b>Anwendungen der globalen Methodik</b>	<b>141</b>
9.1	Optimierung der B-Säule eines Pkw . . . . .	141
9.2	Optimierung von pars10crits19 . . . . .	147
<b>10</b>	<b>Toleranzmaß für die Quantilmetamodelle</b>	<b>151</b>
<b>11</b>	<b>Zusammenfassung und Ausblick</b>	<b>157</b>
<b>III</b>	<b>Anhang</b>	<b>159</b>
<b>12</b>	<b>Grafiken zu Kapitel 8</b>	<b>161</b>
12.1	Grafiken zur Bestimmung der Bandweite . . . . .	161
12.1.1	Stichprobengröße $n = 100$ . . . . .	161
12.1.2	Stichprobengröße $n = 1000$ und $n = 10000$ . . . . .	163
12.2	Vergleich der Quantilschätzer . . . . .	164
12.2.1	Stichprobengröße $n = 100$ . . . . .	164
12.2.2	Stichprobengröße $n = 1000$ . . . . .	166
12.2.3	Stichprobengröße $n = 10000$ . . . . .	168
	<b>Abkürzungsverzeichnis</b>	<b>171</b>
	<b>Abbildungsverzeichnis</b>	<b>173</b>
	<b>Tabellenverzeichnis</b>	<b>178</b>
	<b>Literaturverzeichnis</b>	<b>179</b>



# Kapitel 1

---

## Einleitung

---

Bei der Verbesserung von Produkten und Prozessen vertrauen Wirtschaft und Industrie seit langem auf Methoden der mathematischen Optimierung. Sie wird in so unterschiedlichen Bereichen wie der Fahrzeugentwicklung, der Medizin und des Operations Researchs erfolgreich eingesetzt.

Die praktische Umsetzung der berechneten optimalen Kombination der Eingangsvariablen führt allerdings häufig zu kleinen Abweichungen von den nominalen Werten. Dieser Aspekt wird bei der deterministischen Optimierung außer acht gelassen, dabei kommen Unsicherheiten bei Parametern in Anwendungen häufig vor: Bei Bauteilen sind dies z.B. Fertigungstoleranzen oder Streuung in den Materialeigenschaften, bei der Optimierung von Prozessen beispielsweise Schwankungen in der Umgebungstemperatur. Diese Abweichungen in den Eingangsvariablen führen in den realen Produkten und Prozessen somit zu Abweichungen vom Optimum, deren Größen nicht bekannt sind. In den Anwendungen können sie zu einer erheblichen Differenz vom angestrebten Qualitätsziel führen. Die Industrie strebt aber Produkte von gleichmäßiger Qualität an. Sie sollen robust sein, in dem Sinne, sie unempfindlich gegenüber den möglichen Schwankungen auszulegen. Die gewöhnliche deterministische Optimierung führt aber häufig zu einem Design, das empfindlich auf kleine Abweichungen in den Eingangsvariablen reagiert. Bei optimiertem Design werden ja gerade mögliche Reserven aufgelöst. Die resultierenden optimalen Parameterkombinationen liegen dann häufig an den Grenzen ihres Gültigkeitsbereichs, so dass schon eine minimale Abweichung von den optimalen Einstellungen zum Produkt- oder Prozessversagen führen kann.

Aus diesem Grunde gewinnt die *robuste* Optimierung zunehmend an Bedeutung. Sie ist in der Lage, auch Unsicherheiten zu berücksichtigen und ein Optimum so zu bestimmen, dass es möglichst unempfindlich gegenüber diesen Streuungen ist. Sie liefert verschiedene Formulierungen für die mathematische Charakterisierung der Robustheit und Methoden, die quantifizieren, wie sich Schwankungen der Eingangsvariablen auf die Zielvariablen auswirken. Schließlich stellt sie Strategien bereit, die zum robusten Optimum führen. Ihr Ziel ist es, ein Optimum zu finden, das die Optimierungskriterien möglichst gut erfüllt und gleichzeitig unempfindlich gegenüber

den Schwankungen der Eingangsvariablen ist.

Ihre Idee wurde von Genichi Taguchi ab Ende der 1940er Jahre populär gemacht. Sie entwickelte sich unabhängig davon auch aus der stochastischen Optimierung im Operations Research besonders ab 1995. Ihr Hauptanwendungsgebiet liegt heute jedoch im Ingenieurbereich [BS07].

Bei Anwendungen der robusten Optimierung in der Praxis liegt das Systemverhalten in der Regel nicht als analytische Funktionsgleichung vor. Stattdessen kann es durch Simulationen oder Experimente für einzelne Kombinationen von Eingangsvariablen bestimmt werden. So sind z.B. Computersimulationen heute integraler Bestandteil der virtuellen Produktentwicklung. Andererseits sind Simulationen wie auch Experimente zeitintensiv und teuer. Daher wird in der Praxis angestrebt, ihre Anzahl gering zu halten. Diese Arbeit präsentiert einen Ansatz, der speziell auf derartige Situationen zugeschnitten ist.

### Inhalte und Beiträge dieser Dissertation:

Diese Dissertation entwickelt eine neue effiziente Methodik zur robusten Optimierung mit Quantilmaßen auf globalen Metamodellen. Dabei werden Methoden der Numerik, Optimierung und Statistik geschickt zusammengeführt und um neue Vorgehensweisen ergänzt. Die Methodik ermöglicht robuste Optimierung auf der Grundlage weniger echter Funktionsauswertungen. Dabei gibt sie die Streubereiche der Ausgabevariablen genauer an, als die bisher üblichen Methoden.

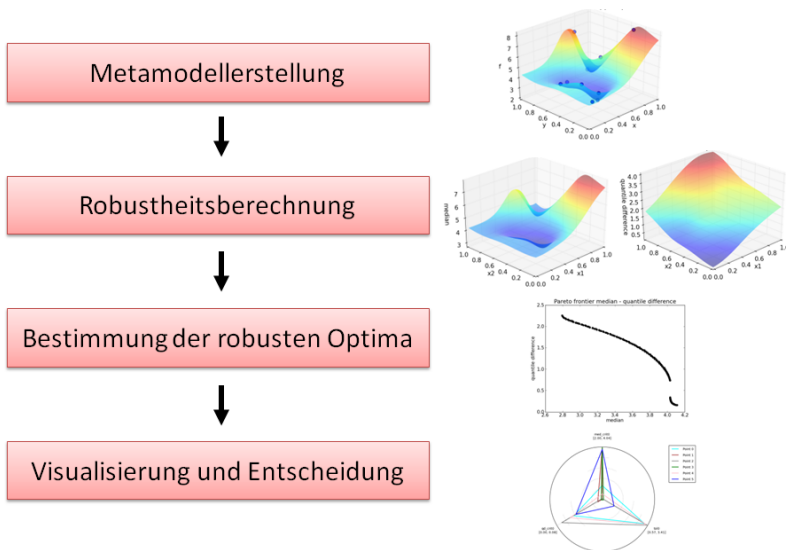


Abb. 1.1: grober Ablauf der Methodik

Der grobe Ablauf der Methodik ist in Abbildung 1.1 dargestellt. Sie umfasst vier Schritte:

- Im ersten Schritt wird das globale Systemverhalten über ein sogenanntes Meta- oder Ersatzmodell angenähert. Es erstellt auf der Grundlage einer Menge von Simulationsergebnissen, die für verschiedene Parameterkombinationen bestimmt wurden, ein Modell des Produkt- oder Prozessverhaltens, also letztlich des Simulationsmodells. Hierher rührt auch die Bezeichnung „Metamodell“, denn es ist das Modell eines Modells. Es kann in Echtzeit ausgewertet werden, so dass eine umfassende Untersuchung des Modellverhaltens möglich ist. Die Metamodellierung ist eine etablierte Technik, wenn nur wenige Simulationen zur Verfügung stehen.
- Im nächsten Schritt wird das Robustheitsverhalten global für das ganze System bestimmt. Dies ist der zentrale Punkt der neu entwickelten Methodik. Die Industrie braucht Robustheitsmaße, die die Streuungsbereiche der Ausgabeparameter möglichst genau angeben, da aus ihnen oft sicherheitsrelevante Grenzen abgeleitet werden. Sie sollen außerdem an den jeweils gewünschten Grad der Robustheit anpassbar sein. Es wird gezeigt, dass die bisher verwendeten Maße Erwartungswert und Standardabweichung Schwächen besitzen. Diese werden durch die hier benutzten Quantilmaße, speziell von Median und symmetrischem oberen und unteren Quantil umgangen. Für die effiziente Berechnung der Quantilmaße wird ein neues Verfahren entwickelt. Innovativ ist auch die globale Charakterisierung des Robustheitsverhaltens über Metamodelle von Quantilen.
- Da sie effizient auswertbar sind, kann im dritten Schritt die Bestimmung der robusten Optima mit Standardalgorithmen der multikriteriellen Optimierung erfolgen.
- Im letzten Schritt werden diese visualisiert und für eine fundierte Entscheidung aufbereitet. Auch dieses Verfahren ist neu.

Im Einzelnen werden Beiträge zu folgenden Punkten gemacht:

1. Neue Vorgehensweise zur effizienten Berechnung des Quantilmaßes  
Die Verwendung des Quantilmaßes zur Charakterisierung von Robustheit ist bisher kaum verbreitet, weil seine Berechnung sehr aufwändig ist. Es wird ein Weg zur Beschleunigung präsentiert, bei dem erstmals fortgeschrittene Versuchspläne mit nichtparametrischen Quantilschätzern kombiniert werden. In einer numerischen Studie wird die Genauigkeit und Effizienz des Verfahrens nachgewiesen und seine Grenzen aufgezeigt (Kap. 8).
2. Neuartiger effizienter iterativer Algorithmus für die Quantilschätzung  
Bisher existiert kein Algorithmus, der die Robustheit eines Systems mit Quantilmaßen bis zu einer vorgegebenen Genauigkeit schätzen kann. Hier wird das beste Verfahren der numerischen Untersuchung für einen effizienten iterativen Algorithmus benutzt, der dies für eine feste Parameterkombination leistet (Kap. 5.4). Sein Aufwand passt sich dem lokalen Verhalten des Systems an. Die Anwendbarkeit wird an Beispielen erläutert. Mit dem Algorithmus kann die Robustheit des Systems für jede beliebige Parameterkombination bestimmt werden.

3. Innovative globale Charakterisierung der Robustheit eines Systems mit Metamodellen von Quantilen

Dieser lokale Ansatz wird erweitert zur globalen Charakterisierung der Robustheit eines Systems. Hierzu werden erstmals Metamodelle für bestimmte Quantile benutzt (Kap. 6). Damit ist es im Gegensatz zu den herkömmlichen Verfahren zur robusten Optimierung möglich, schnell eine Aussage über die Robustheit jeder beliebigen Systemeinstellung zu treffen. Die erstellten Metamodelle sind global und sehr schnell auswertbar. Deshalb können zur Bestimmung der robusten Optima Standardalgorithmen für die multikriterielle Optimierung verwendet werden.

4. Erstmalige Einführung eines Maßes für die Toleranzgrenzen von Quantilmetamodellen

Da bei der globalen Charakterisierung der Robustheit die Quantilmetamodelle auf der Grundlage des Funktionsmetamodells erstellt werden, stellt sich die Frage, wie sich die Ungenauigkeit des ersten Modells auf die darauf basierenden auswirkt. Hierzu existieren bisher keine Ansätze. In dieser Dissertation wird ein neuartiges Kreuzvalidierungsverfahren entwickelt, das diesen Einfluss berechnet und die ungenauen Bereiche der Quantilmetamodelle identifiziert (Kap. 6.2).

5. Innovative Entscheidungsunterstützung durch eine iterative Visualisierung der robusten Optima

Bei mehreren Optimierungskriterien ist der Lösungsraum hochdimensional und die Anzahl der robusten Optima groß. Um dem Entscheider die Auswahl eines Optimums zu erleichtern, wird eine neue Methodik zur Visualisierung der robusten Optima entwickelt (Kap. 6.1.4). Dem Anwender werden zuerst die „grundsätzlichen“ Lösungen in einem Sterndiagramm präsentiert, aus denen er auswählt. Über mehrere Stufen wird er anschließend zur Feinauswahl geleitet. Mathematisch ist hierzu die Clusterung der Pareto-Punkte nötig. Das Verfahren ist speziell auf die robuste Optimierung zugeschnitten und bezieht außerdem die Modelltoleranz als Maß für den Metamodellfehler mit ein.

6. Neue Abschätzung, wann sich der Einsatz von Multipolverfahren zur Beschleunigung der Auswertung von Metamodellen lohnt

Zur Bestimmung des Robustheitsverhaltens des Gesamtsystems sind intensive Auswertungen der Metamodelle erforderlich. Vor kurzem [KW11] wurden Multipolverfahren für eine spezielle Klasse von Metamodellen entwickelt, die den benötigten Aufwand deutlich reduzieren können. Neu erarbeitet wird hier eine genaue Aufwandsabschätzung, aus der abgeleitet wird, wann der Einsatz dieses Multipolverfahrens die Berechnung bei der robusten Optimierung beschleunigt (Kap. 3.6).

Insgesamt wird damit ein innovatives effizientes und praxisnahes Verfahren zur robusten Optimierung präsentiert, das nur wenige echte Funktionsauswertungen benötigt.

**Diese Dissertation gliedert sich wie folgt:**

Zuerst wird anhand eines Anwendungsbeispiels aus der Industrie die Zielsetzung der Dissertation verdeutlicht. Insbesondere werden die Vorteile der Quantilmaße erläutert (Kapitel 2).

Im anschließenden Methodikteil werden die Grundlagen der vielfältigen benutzten Verfahren eingeführt (Kapitel 3). Die robuste Optimierung wird in Kapitel 4 besprochen. Es enthält eine ausführliche Diskussion der bestehenden Robustheitsmaße. Die Bestimmung von Quantilen der Ausgabevariablen für eine feste Kombination von Eingangsvariablen ist Thema von Kapitel 5. Hier wird auch der neue iterative Algorithmus für das Quantilmaß eingeführt. Dieser lokale Ansatz wird im folgenden Kapitel 6 auf den gesamten Designraum erweitert. Es enthält die innovative globale Methodik zur robusten Optimierung mit Metamodellen, das neue Entscheidungsunterstützungsverfahren und das neu entwickelte Kreuzvalidierungsverfahren zur Abschätzung des Einflusses der Modelltoleranz des Funktionsmetamodells auf die Quantilmetamodelle. In Kapitel 7 folgt die genaue Aufwandsabschätzung für das Multipolverfahren.

Die eingeführten Methoden werden anschließend in numerischen Untersuchungen validiert: Der Vergleich verschiedener Versuchspläne und nichtparametrischer Quantilschätzer für die lokale Berechnung des Quantilmaßes ist in Kapitel 8 zu finden. Anwendungen der globalen Methodik werden in Kapitel 9 vorgestellt. Das Toleranzmaß für die Quantilmetamodelle wird in Kapitel 10 untersucht. Danach folgt eine Zusammenfassung der Arbeit (Kapitel 11). Der Anhang (Kapitel 12) enthält einige ergänzende Grafiken.

**Bemerkung:**

Teile dieser Dissertation wurden bereits in [RCR14], [RCR12], [RCR] und [RR13] veröffentlicht.



## Kapitel 2

---

# Einführendes Beispiel

---

Die Automobilindustrie betreibt seit Jahren Gewichtsreduzierung an ihren Produkten. Dieser Trend hat aktuell noch einmal an Bedeutung gewonnen durch die verschärften Richtlinien zum CO<sub>2</sub>-Ausstoß von Kraftfahrzeugen und zur Senkung des Kraftstoffverbrauchs. Auch in der Elektromobilität spielt sie eine große Rolle, da geringeres Gewicht zu einem reduzierten Energiebedarf und damit zu erweiterter Reichweite von Elektroautos beiträgt. Auf der anderen Seite bestimmt die Sicherheit von Fahrzeugen zunehmend den Kaufentscheid von Kunden. Die Gewichtsreduzierung darf nicht zu Lasten des Insassenschutzes gehen.

Zur Analyse des Crashverhaltens von Pkws werden mächtige Simulationsprogramme eingesetzt, die echte Aufprallversuche weitgehend ersetzen konnten. Die Laufzeit solcher Simulationen ist sehr lang. Sie kann je nach Detaillierungsgrad mehrere Tage betragen. Wegen der kurzen Produktinnovationszyklen in der Automobilindustrie kann nur eine begrenzte Anzahl solcher Simulationen durchgeführt werden.

Vor diesem Hintergrund benötigt die Industrie Strategien, die mit wenigen Simulationen eine optimale Auslegung eines Bauteils erlauben. Lange Zeit war es gängige Praxis, pro Simulation einen Parameter zu verändern und aus dem Vergleich zur Vorgängersimulation Schlüsse für die nächste durchzuführende Simulation zu ziehen. Dieses Verfahren verbessert nach und nach die Ergebnisse, schöpft aber nur selten das gesamte Optimierungspotential aus, da das globale Systemverhalten unbekannt bleibt. Als Ausweg hat sich der Einsatz von Metamodellen etabliert. Aus einer Menge von Simulationen für verschiedene Einstellungen der Eingabevariablen wird eine Interpolationsfunktion generiert, die das globale Systemverhalten des Bauteils approximiert. Auf dem Metamodell sind dann umfangreiche Auswertungen möglich. Mathematische Optimierungsmethoden, die die besten Kompromisse aus Reduzierung des Gewichts und Verbesserung des Insassenschutzes finden, können angewendet werden. Dass dieser Schritt nicht ausreicht, um den hohen Anforderungen der Automobilindustrie stand zu halten, soll an einem Anwendungsfall demonstriert werden: Zwischen den beiden Seitentüren eines Pkw befindet sich die sogenannte B-Säule. Sie besteht aus zwei übereinanderliegenden U-Profilen (vgl. Abbildung 2.1). Ihr Verhalten bei einem Seitenaufprall wird in einer Crashesimulation bestimmt,

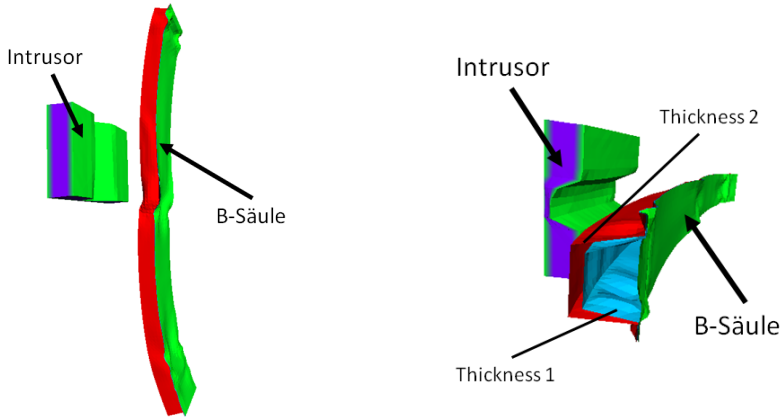


Abb. 2.1: B-Säule mit Intrusor, links Seitenansicht, rechts Draufsicht

in dem ein Intrusor mittig auf die B-Säule trifft. Im Aufprallbereich wird ein Referenzpunkt modelliert, dessen Verschiebung gemessen wird. Diese Größe wird im folgenden *Intrusion* genannt.

Zur Dimensionierung der B-Säule wird die Dicke der beiden U-Profile unabhängig voneinander variiert. Dies sind die beiden Eingabevariablen *Thickness1* und *Thickness2*. Die Säule soll so ausgelegt werden, dass sie einerseits möglichst leicht ist aber andererseits auch einen Seitenaufprall möglichst gut abfängt und ein Eindringen der Fahrgastzelle verhindert. Ziel ist es, die beiden Ausgabevariablen *Mass* und *Intrusion* zu minimieren. Im konkreten Anwendungsfall wurden *Thickness1* und *Thickness2* zwischen 0,5 und 3 mm variiert. Für die Optimierung soll nur der Bereich zwischen 0,8 und 2,7 in Betracht gezogen werden. Es wurden 101 Simulationen mit unterschiedlichen Kombinationen der Eingabevariablen durchgeführt und hieraus für die beiden Ausgabevariablen je ein Metamodell erstellt (vgl. Abbildung 2.2). Das Modell für *Intrusion* zeigt, dass sich die Verschiebung mit steigender Di-

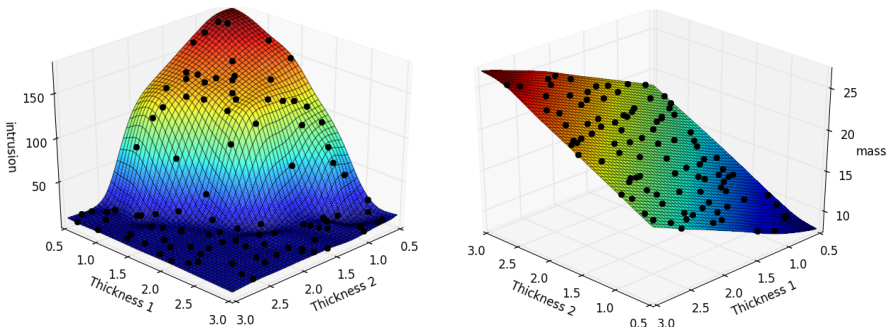


Abb. 2.2: Metamodell für *Intrusion* (links) und *Mass* (rechts) mit Simulationsergebnissen



cke der U-Profile verringert. Der Abfall erfolgt ungefähr in der Diagonalen und ist sehr steil. Übersteigt die Summe der beiden Dicken den Wert von etwa  $3,5 \text{ mm}$  findet keine Verschiebung mehr statt. *Thickness1* und *Thickness2* verhalten sich nicht ganz symmetrisch.

Das Gewicht sinkt linear mit der Abnahme der Dicken. Für *Thickness2* fällt es schneller ab als für *Thickness1*.

*Intrusion* und *Mass* verlaufen genau gegenläufig. Damit das Verhalten beider Variablen trotzdem gut sichtbar ist, ist in den Abbildungen die Ansicht von *Mass* um  $90 \text{ Grad}$  im Gegenuhrzeigersinn gegenüber den Ansicht von *Intrusion* gedreht.

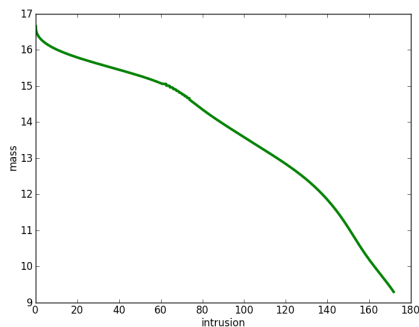


Abb. 2.3: beste Kompromisse für *Intrusion* und *Mass*

## 2.1 Deterministische Optimierung

Für die beiden Ausgabevariablen wird die Menge der besten Kompromisse bestimmt. Sie lässt sich darstellen, in dem man die Werte jeder Variablen auf den Koordinatenachsen gegeneinander aufträgt (vgl. Abb. 2.3). Die zugehörigen Werte der Eingabevariablen sind in den Abbildungen 2.4 und 2.5 als schwarze Punkte eingetragen. Sie verlaufen für *Thickness2* am Rand des zulässigen Bereichs. Bei dieser Betrachtung

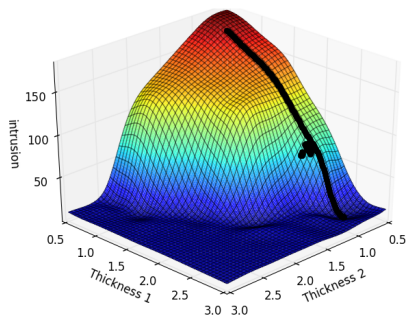


Abb. 2.4: Metamodell für *Intrusion* mit besten Kompromissen

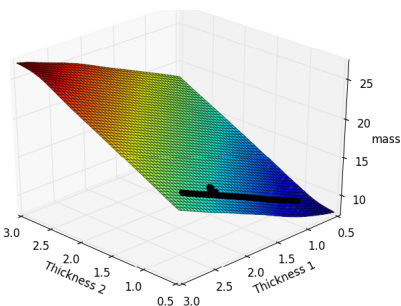


Abb. 2.5: Metamodell für *Mass* mit besten Kompromissen

werden mögliche Abweichungen der Zielvariablen, die durch die Fertigungstoleranzen der Blechdicken hervorgerufen werden, gänzlich außer acht gelassen. Diese können aber zu gravierenden Abweichungen von den nominalen Werten führen. Greift man z.B. aus der Menge der besten Kompromisse den Punkt  $Thickness1 = 2,5$  und  $Thickness2 = 0,8$  heraus, so hat  $Intrusion$  den nominalen Wert 10,47. Nimmt man an, dass die Fertigungstoleranzen der Blechdicken normalverteilt sind mit einer Standardabweichung von 0,05 mm, erhält man für  $Intrusion$  den Erwartungswert  $\mu = 11,84$  bei einer Standardabweichung von  $\sigma = 7,85$ . Diese Streuung ist erheblich zu hoch. Sie darf bei der Auslegung der B-Säule nicht vernachlässigt werden. Auf der

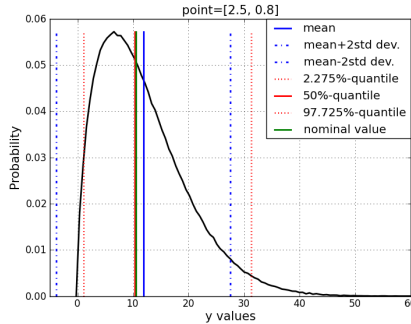


Abb. 2.6: übertragene Verteilung für  $Intrusion$  bei  $Thickness1 = 2,5$  und  $Thickness2 = 0,8$

Grundlage dieser beiden Maße werden in der Industrie häufig die folgenden Grenzen für den Schwankungsbereich abgeleitet:

- $\mu + / - 1\sigma$  überdecken 68% der Verteilung.  
16% liegen unterhalb von  $\mu - \sigma$  und 16% oberhalb von  $\mu + \sigma$
- $\mu + / - 2\sigma$  überdecken 95,45% der Verteilung.  
2,275% liegen unterhalb von  $\mu - 2\sigma$  und 2,275% oberhalb von  $\mu + 2\sigma$
- $\mu + / - 3\sigma$  überdecken 99,73% der Verteilung.  
0,135% liegen unterhalb von  $\mu - 3\sigma$  und 0,135% oberhalb von  $\mu + 3\sigma$

Dahinter steckt implizit die Annahme einer Normalverteilung, für die diese Beziehungen gelten. Würde man diese Grenzen auf die obige Verteilung anwenden, erhielte man solche, die von den tatsächlichen Werten stark abweichen:

$i$	Grenzen aus $\mu + / - i \cdot \sigma$	Intervall aus tatsächlichen Quantilwerten
1	[3, 99; 19, 69]	[4,20; 19,58]
2	[-3, 86; 27, 54]	[1,11; 31,24]
3	[-11, 71; 35, 39]	[0,17; 43,61]

Dies verdeutlicht auch Abbildung 2.6, die die stark verzogene Verteilungsfunktion von  $Intrusion$  für obige Parameterkombination aus  $Thickness1$  und  $Thickness2$  zeigt. Für die Kombination Erwartungswert  $\pm$  Standardabweichung erhält man als untere Grenzen zum Teil negative Werte, die für  $Intrusion$  physikalisch nicht

auftreten können. Ein Maß, das solche Abweichungen zulässt, ist in einem so sensiblen Bereich wie der Crashsimulation nicht brauchbar.

Quantile charakterisieren die Grenzen von Verteilungen wesentlich besser. Sie entsprechen der eigentlichen Intention des Anwenders: Ziel sind zuverlässige obere und untere Grenzen für die Streuung der Ausgabevariablen. Er möchte nicht die Standardabweichung minimieren, sondern die Streuung eines mittigen Prozentsatzes der Ausgabeverteilung. Über die Wahl des Quantils kann er zudem selbst bestimmen, wie hoch dieser Prozentsatz sein soll. Quantile passen sich so dem Sicherheitsbedürfnis des Anwenders an. Dieses Verhalten ist nicht nur an dem ausgewählten Punkt

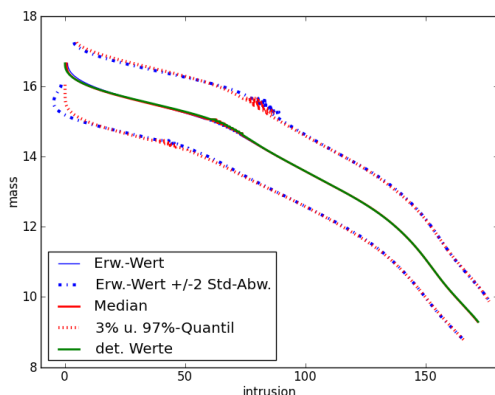


Abb. 2.7: deterministische Paretofront mit Schwankungsbereich

zu beobachten. Abbildung 2.7 zeigt die oberen und unteren Grenzen der deterministischen Paretofront für *Intrusion* errechnet über Erwartungswert  $\pm 2$  Standardabweichungen und über 2,275%-Quantil und 97,725%-Quantil. Man erkennt wieder den großen Schwankungsbereich von *Intrusion* über den gesamten Wertebereich. In der Umgebung der deterministischen Paretofront verhält sich diese Variable ab etwa 10 nahezu linear. Dort treten nur kleine Unterschiede zwischen den beiden Charakterisierungen von Robustheit auf. Im Bereich von *Intrusion* bis 10 ist es nichtlinear. Dort laufen auch die Grenzlinien von Erwartungswert  $\pm 2$  Standardabweichungen und oberem bzw. unterem Quantil auseinander.

## 2.2 Neuer Ansatz zur robusten Optimierung mit Quantilen

Bezieht man die Streuungen der Ausgabevariablen als zusätzliche Ziele mit in die Optimierung ein, führt man eine *robuste* Optimierung durch. Für die Dimensionierung der B-Säule wurde sie beruhend auf den Quantilmaßen für  $q = 0.05$  durchgeführt. Die Abbildungen 2.8 bis 2.11 zeigen die Metamodelle für den Median und die Quantildifferenz von *Masse* und *Intrusion*. Die schwarzen Bereiche sind wieder die Menge der besten Kompromisse, die jetzt sehr viel größer ist.

Die Metamodelle für den Median unterscheiden sich für beide Ausgabevariablen kaum vom ursprünglichen Modell. Die etwas flacheren Bereiche von *Intrusion* gehören nun mit zur Menge der optimalen Kompromisse. Das Metamodell der Quantildifferenz für *Intrusion* gleicht einem Gebirgszug, der diagonal zum Blickwinkel verläuft. Da das Gewicht sich fast linear verhält, ist die Quantildifferenz für *Mass* fast konstant. Nur in den ungenauen Randbereichen, die nicht mit in die Optimierung eingehen, fällt es steil ab.

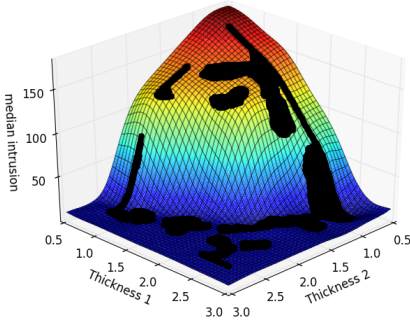


Abb. 2.8: Metamodell für den Median von *Intrusion* mit besten Kompromissen

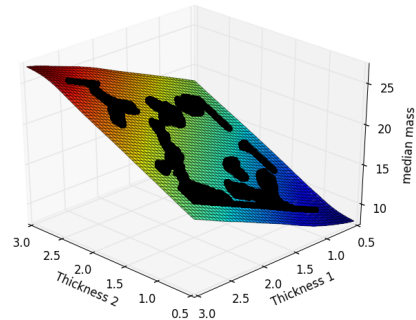


Abb. 2.9: Metamodell für den Median von *Mass* mit besten Kompromissen

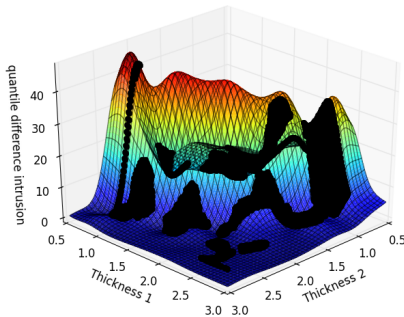


Abb. 2.10: Metamodell für die Quantildifferenz von *Intrusion* mit besten Kompromissen

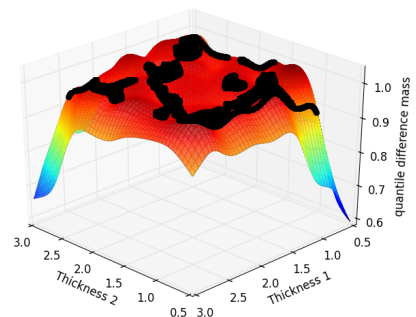


Abb. 2.11: Metamodell für die Quantildifferenz von *Mass* mit besten Kompromissen

Optimum erhält man z.B. den Punkt  $Thickness1 = 0,81$  und  $Thickness2 = 2,44$ , der nun nicht mehr auf der Grenze des zulässigen Bereichs liegt.

	<i>Th.1</i>	<i>Th.2</i>	Med. <i>Int.</i>	QD <i>Int.</i>	Med. <i>Mass</i>	QD <i>Mass</i>
det. Op.	2,5	0,8	10,01	24,7	16,0	1
rob. Op.	0,81	2,44	10,38	15,99	16,81	1

Er besitzt bei gleichem Median für *Intrusion* eine wesentlich kleinere Quantildifferenz. Dies wird erkaufte durch einen geringen Zuwachs an Gewicht. Auch die obere

und untere Grenze des Schwankungsbereich ist nun genau bekannt. Abbildung 2.12 zeigt die zugehörige Verteilung für *Intrusion*. Auch diese ist stark verzerrt. Die

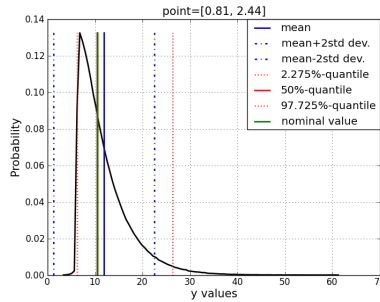


Abb. 2.12: übertragene Verteilung für *Intrusion* bei *Thickness1* = 0,81 und *Thickness2* = 2,44

Menge der besten Kompromisse ist nun 4-dimensional, eine Visualisierung als ganzes ist nicht mehr möglich. Abbildung 2.13 zeigt den Vergleich der Punkte aus der deterministischen Optimierung mit den Medianwerten der robusten Optimierung. Durch die Hinzunahme der Quantildifferenz als Optimierungsziel kommen Punkte hinzu, die zwar für *Intrusion* und *Mass* schlechtere Werte annehmen, aber eine geringere Streuung aufweisen, die aber nicht in der Abbildung visualisiert ist). Die Grafik zeigt, dass dies für *Intrusion* größer als 150 nicht möglich ist. Für *Intrusion* kleiner als 40 sind Punkte mit nur kleinen Abweichungen von der deterministischen Linie vorhanden. Abbildung 2.14 zeigt zusätzlich die Streuungen von *Intrusion*,

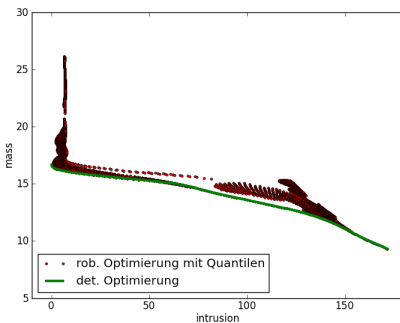


Abb. 2.13: Vergleich der Menge der besten Kompromisse bei deterministischer und robuster Optimierung

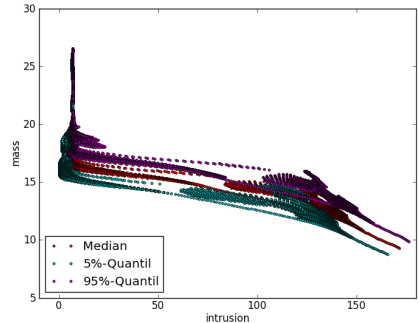


Abb. 2.14: beste Kompromisse bei robuster Optimierung mit Streuungsbereich

gemessen in Quantilen. Für *Intrusion*  $\approx 15$  gibt es eine Reihe von Punkten, die nur eine sehr kleine Schwankung aufweisen. In diesem Fall liegt *Mass* über 20. Im Metamodell ist dies der konstante Bereich von *Intrusion*.

## 2.3 Im Vergleich: Robuste Optimierung mit den Standardmaßen

Zum Vergleich wird nun die robuste Optimierung unter den gleichen Annahmen mit den in der Industrie verbreiteten Standardmaßen Erwartungswert und Standardabweichung durchgeführt. Die Abbildungen 2.15 und 2.16 zeigen die Metamodelle für den Erwartungswert von *Intrusion* und *Mass*. Sie unterscheiden sich kaum von denen für den Median. Auch die Modelle für die Standardabweichung von *Intrusion* (vgl. Abbildung 2.17) und *Mass* (vgl. Abbildung 2.16) weichen in ihrer Form nur wenig von denen für die Quantildifferenz ab. Um die beiden Maße quantitativ zu vergleichen, müssen die Werte der Standardabweichung verdoppelt werden. Sie nimmt dann immer noch kleinere Werte an als die Quantildifferenz. Die schwarzen Berei-

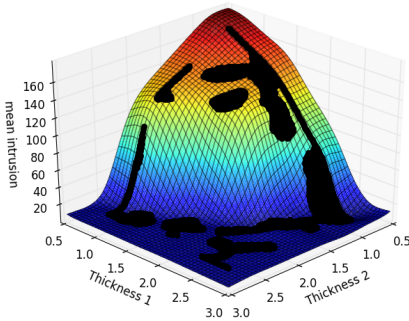


Abb. 2.15: Metamodell für den Erwartungswert von *Intrusion* mit besten Kompromissen

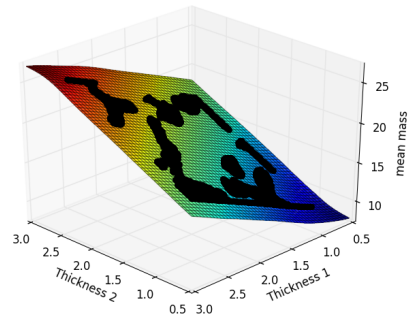


Abb. 2.16: Metamodell für den Erwartungswert von *Mass* mit besten Kompromissen

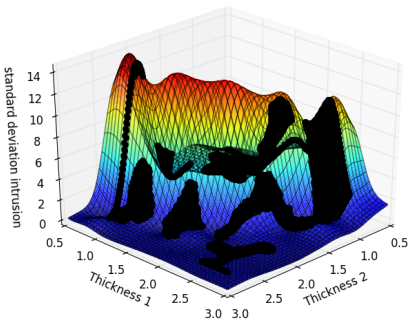


Abb. 2.17: Metamodell für die Quantildifferenz von *Intrusion* mit besten Kompromissen

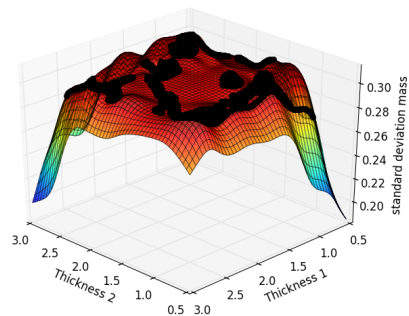


Abb. 2.18: Metamodell für die Quantildifferenz von *Mass* mit besten Kompromissen

che in den Abbildungen markieren wieder die Menge der besten Kompromisse. Sie unterscheidet sich von der mit Median und Quantilmaß berechneten Menge. Dies war zu erwarten, da sie ja auf der Grundlage von unterschiedlichen Modellen berechnet wurde. Wie groß der Unterschied bei den Mengen der besten Kompromisse

für die betrachteten Maße ist, zeigt Abbildung 2.19. Es gibt zwar große Bereiche, die von beiden Maßen empfohlen werden. Manche Bereiche unterscheiden sich aber deutlich, etwa bei  $Thickness1 \approx 1,7$  und  $Thickness2 \approx 2,3$ . Nimmt man an, dass

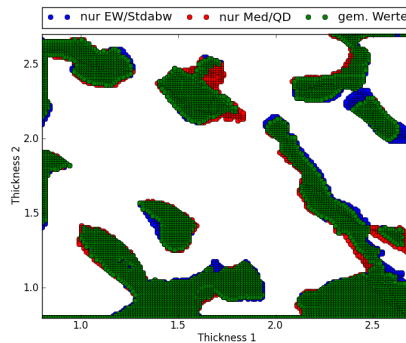


Abb. 2.19: Vergleich der Menge der besten Kompromisse in Abhängigkeit von  $Thickness1$  und  $Thickness2$

das eigentliche Ziel des Anwenders nicht die Minimierung der Standardabweichung ist, sondern die Minimierung eines mittleren Quantilbereichs, ist damit gezeigt, dass die Verwendung von Erwartungswert und Standardabweichung hierfür ungeeignet ist. Die beiden Maße führen nicht nur zu falschen oberen und unteren Grenzen für die Lage der Verteilung, sondern erkennen auch nicht die wirklich robusten Bereiche bezogen auf die Quantildifferenz.

Dieses Anwendungsbeispiel wird in Kapitel 9.1 erweitert. Dort sind die Abweichungen zwischen den Mengen der besten Kompromisse für die verschiedenen Maße noch deutlicher ausgeprägt.

## 2.4 Fazit

Die Einsicht, dass Streuungen von Eingabevariablen mit in die Optimierung einbezogen werden müssen, hat sich inzwischen etabliert. Dabei wird die Verteilung der Ausgabevariablen üblicherweise über die Kennwerte Erwartungswert und Standardabweichung charakterisiert. Das obige Beispiel zeigt, dass selbst in einem einfachen Anwendungsfall diese Maße ungenaue Bereiche angeben. Wird mit ihnen eine robuste Optimierung durchgeführt, werden in den nichtlinearen Bereichen der Übertragungsfunktion ungeeignete Maße zugrunde gelegt. Eine deutliche Verbesserung wird erzielt, wenn stattdessen die Kombination aus Median und Quantildifferenz als Robustheitsmaß benutzt wird. Dieses Maß kann die für die Industrie wichtigen oberen und unteren Schranken von Verteilungen bestimmen. Diese Grenzen sind insbesondere bei der Einhaltung von Sicherheitsschranken, die über Nebenbedingungen angegeben werden, von großer Bedeutung. Durch die Berechnung vieler Quantile kann die Verteilung der Ausgabevariablen angenähert werden. Dies ist nützlich, um Sprünge in der Verteilung zu erkennen. Zum Beispiel ist es beim vorgeschlagenen robusten Optimum für die B-Säule wichtig, die Spitze der Verteilung zu lokalisieren (vgl. Abbildung 2.12).

Das Quantilmaß ist bisher kaum untersucht und wurde in Anwendungen nicht benutzt, da seine Berechnung aufwändig ist. Zentrales Thema dieser Dissertation ist es deshalb, eine Methodik zur robusten Optimierung zu entwickeln, die mit nur wenigen echten Simulationen auskommt und die Streuungsbereiche mit Quantilmaßen bestimmt.



**Teil I**

**Methodik**



## Kapitel 3

---

# Grundlagen

---

Die entwickelte Methodik kombiniert Elemente aus verschiedenen Bereichen der Mathematik und Statistik, deren Grundlagen in diesem Kapitel erläutert werden. Nach einigen grundlegenden Begriffsdefinitionen (Kap. 3.1) werden zunächst mehrere Versuchspläne vorgestellt (Kap. 3.2). Sie werden später für die Schätzung von Quantilen benötigt. Anschließend folgt die Einführung verschiedener nichtparametrischer Quantilschätzer (Kap. 3.3) und weiterer Methoden zur Quantilschätzung (Kap. 3.4). Zentraler Punkt der Methodik ist die Verwendung von Metamodellen. Ihre Grundlagen werden in Kap. 3.5 erläutert. Eine spezielle Klasse von Metamodellen kann effizient mit Hilfe von Multipolverfahren ausgewertet. Es wird in Kap. 3.6 vorgestellt. Kap. 3.7 gibt einen Überblick über die multikriterielle Optimierung, da diese für die Bestimmung von robusten Optima herangezogen wird. In Kap. 3.8 werden Methoden der Clusteranalyse beschrieben, die zur Entscheidungsfindung benutzt werden.

### 3.1 Grundlagen aus Wahrscheinlichkeitsrechnung und Statistik

#### 3.1.1 Grundlagen aus der Wahrscheinlichkeitsrechnung

Für eine eindimensionale Zufallsvariable  $X$  gilt:

Die *Verteilungsfunktion* ( $VF$ ) einer Zufallsvariablen  $X$  wird mit  $F_X(x)$ ,  $x \in \mathbb{R}$ , bezeichnet und ist definiert durch

$$F_X(x) = P(X \leq x)$$

wobei  $\mathbb{R}$  für die Menge der reellen Zahlen steht. Der Index  $X$  wird weggelassen, wenn keine Verwechslung möglich ist.

Ist  $F$  stetig und einmal stückweise stetig differenzierbar, dann existiert die Ableitung

$$p(x) = \frac{dF(x)}{dx}.$$

Sie heißt *Dichtefunktion* ( $DF$ ) von  $X$ .  $p$  ist dann stückweise stetig.

Der *Erwartungswert* von  $X$  ist für eine eindimensionale Zufallsvariable definiert durch

$$E X = \mu = \int_{-\infty}^{\infty} x \cdot p(x) dx$$

die *Varianz* durch

$$\text{Var}(X) = \sigma^2 = \int_{-\infty}^{\infty} (x - E X)^2 \cdot p(x) dx$$

und die *Standardabweichung* durch

$$\sigma = \sqrt{\text{Var}(X)}$$

Für jedes  $q$  mit  $0 < q < 1$  ist das  $q$ -Quantil bestimmt über

$$x_q = F^{-1}(q) = \inf\{x | F(x) \geq q\}$$

Das 0,5-Quantil heißt *Median*.

Die  $q$ -Quantile mit  $q = 0,01, \dots, 0,99$  werden *Perzentile* genannt.

$$Q(q) = F^{-1}(q)$$

wird als *Quantilfunktion* ( $QF$ ) bezeichnet. Die Umkehrfunktion von  $F$  existiert nur, wenn  $F$  streng monoton steigend ist. In allen anderen Fällen, gilt die obige Definition über das Infimum.

Falls die nötigen Ableitungen existieren, kann die erste und zweite Ableitung von  $Q(q)$  ausgedrückt werden durch

$$Q'(q) = \frac{1}{p(Q(q))} = \frac{1}{p(x_q)} \quad \text{und} \quad Q''(q) = -\frac{p'(Q(q))}{p(Q(q))^3} = -\frac{p'(x_q)}{p(x_q)^3}$$

Der Beweis erfolgt mit  $p(x_q) = F'(x_q)$ .

Für eine  $d$ -dimensionale Zufallsvariable  $X = (\Xi_1, \dots, \Xi_d)^T$  heißt  $p(\xi) = p(\xi_1, \dots, \xi_d)$   $d$ -dimensionale *Dichtefunktion*, wenn gilt:

1.  $p(\xi_1, \dots, \xi_d) \geq 0$  für alle  $\xi_1, \dots, \xi_d \in \mathbb{R}$
2.  $\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} p(\xi_1, \dots, \xi_d) d\xi_1 d\xi_2 \dots d\xi_d = 1$ .

Wenn  $\Xi_1, \dots, \Xi_k$  stochastisch unabhängig voneinander sind, gilt:

$$p_X(x) = \prod_{i=1}^d p_{\Xi_i}(\xi_i)$$

In mehreren Dimensionen gilt  $E X = (E \Xi_1, \dots, E \Xi_d)^T$ . Die Definitionen von Varianz und Standardabweichung werden analog übertragen.

### 3.1.2 Grundlagen aus der Statistik

Später werden aus Stichproben verschiedene Parameter von Wahrscheinlichkeitsverteilungen geschätzt. An dieser Stelle werden die wichtigsten Begriffe eingeführt:

Eine *Stichprobe*  $(x_1, \dots, x_n)$  ist eine Beobachtungsreihe, die aus einem Zufallsexperiment gewonnen wird. Für  $i = 1, \dots, n$  wird jeder Stichprobenwert  $x_i$  als Realisierung der Zufallsvariablen  $X_i$  aufgefasst. Damit ist die gesamte Stichprobe eine Realisierung der Zufallsvariablen  $(X_1, X_2, \dots, X_n)$ . Wenn alle Stichprobenelemente  $x_i$   $i = 1, \dots, n$  unabhängig voneinander gezogen werden, sind die zugehörigen  $X_i$   $i = 1, \dots, n$  stochastisch unabhängig voneinander. Sind die Zufallsvariablen zudem identisch verteilt, d.h. besitzen alle die gleiche Verteilungsfunktion  $F(x)$ , spricht man von einer *einfachen unabhängigen Stichprobe*. Im folgenden wird immer von dieser Situation ausgegangen.

Die *empirische Verteilungsfunktion*, die von der Stichprobe  $(x_1, x_2, \dots, x_n)$  erzeugt wird, ist gegeben durch

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{x \geq x_i}, \quad x \in \mathbb{R}$$

wobei  $\mathbf{1}_A$  die Indikatorfunktion zum Ereignis  $A$  ist, d.h.  $\mathbf{1}_A = 1$  falls das Ereignis  $A$  eintritt und  $\mathbf{1}_A = 0$  sonst.

Sei  $x_{(1)}$  die kleinste Zahl der Stichprobe  $x_1, x_2, \dots, x_n$ ,  $x_{(2)}$  die zweitkleinste und  $x_{(n)}$  die größte. Dann ist  $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$  die Anordnung der Stichprobe ihrer Größe nach und der Ausdruck  $(x_{(1)}, x_{(2)}, \dots, x_{(n)})$  heißt *Ordnungsstatistik* der Zufallsstichprobe  $(X_1, X_2, \dots, X_n)$ .

Es ist leicht zu sehen, dass die empirische Verteilungsfunktion  $F_n(x)$  dann auch ausgedrückt werden kann durch

$$F_n(x) = \begin{cases} 0 & \text{für } x < x_{(1)} \\ \frac{k}{n} & \text{für } x_{(k)} \leq x < x_{(k+1)} \quad k = 1 \dots n-1 \\ 1 & \text{für } x > x_{(n)} \end{cases}$$

Nach dem Gesetz der großen Zahlen gilt:

$$F_n(x) \rightarrow F(x) \text{ mit Wahrscheinlichkeit } 1, \text{ wenn } n \rightarrow \infty$$

Aus dem zentralen Grenzwertsatz folgt:

$$\sqrt{n}(F_n(x) - F(x)) \rightarrow \mathcal{N}(0, F(x)(1 - F(x))) \quad \text{für } n \rightarrow \infty \quad (3.1)$$

Die empirische Verteilungsfunktion konvergiert damit punktweise (für festes  $x$ ) gegen die Verteilungsfunktion mit der Konvergenzrate  $1/\sqrt{n}$ .

Der Standardfehler der Schätzung beträgt dabei  $\sqrt{F(x)(1 - F(x))/n}$ .

Die empirische Verteilungsfunktion ist daher ein erwartungstreuer und konsistenter Schätzer der Verteilungsfunktion von  $X$ .

### 3.1.3 Funktionen von Zufallsvariablen

Unsichere Eingangsvariablen lassen sich durch die Verteilung einer Zufallsvariablen  $X$  charakterisieren. Diese Verteilung wird dann von der Funktion  $f$  übertragen. Das

Ergebnis ist nicht mehr ein nominaler Wert, sondern ebenfalls eine Zufallsvariable

$$Y = f(X)$$

Soweit nicht anders vermerkt, sei  $X = (X_1, \dots, X_d)$  eine  $d$ -dimensionale und  $Y = (Y_1, \dots, Y_l)$  eine  $l$ -dimensionale Zufallsvariable.

$X_0$  sei die Verteilung um einen festen Punkt  $x_0$ . Die übertragene Verteilung an der Stelle  $f(x_0)$  wird dann mit  $Y_0 = f(X_0)$  bezeichnet.

Auf analytischem Weg lässt sie sich nur für spezielle Funktionen  $f$  bestimmen:

Es gilt für eine Dimension ([Haf89]):

Sei  $X$  eine stetige Zufallsvariable und  $y = f(x)$  eine differenzierbare und streng monotone Funktion und  $x = f^{-1}(y)$  ihre Umkehrung.

Ist  $f$  steigend, dann gilt:

$$F_Y(y) = F_X(f^{-1}(y))$$

Ist  $f$  fallend, dann gilt:

$$F_Y(y) = 1 - F_X(f^{-1}(y))$$

In beiden Fällen gilt:

$$p_Y(y) = p_X(f^{-1}(y)) \left| \frac{df^{-1}(y)}{dy} \right|$$

Der Satz gilt auch für differenzierbare und monotone Teile einer Funktion. Für eine beliebige Funktion kann die Verteilung von  $Y$  deshalb durch Aufteilung von  $f$  in streng monotone Teile, Berechnung der übertragenen Dichtefunktion und Zusammensetzen im Bildraum bestimmt werden. Auf diese Weise lässt sich auch zeigen, dass die Übertragung der Standardnormalverteilung durch eine Normalparabel im Nullpunkt eine  $\chi^2$ -Verteilung ergibt :

$$X \sim \mathcal{N}(0, 1) \Rightarrow X^2 \sim \Gamma\left(\frac{1}{2}, 2\right) = \chi_1^2$$

Für höhere Dimensionen lässt sich die übertragene Verteilung analytisch nur für  $d = l$  bestimmen ([Haf89]):

Gegeben sei eine  $d$ -dimensionale stetige Zufallsvariable  $\mathbf{X} = (\Xi_1, \dots, \Xi_d)^T$  und  $y = (y_1, \dots, y_d) = f(\xi_1, \dots, \xi_d)$  eine umkehrbare Funktion, die samt ihrer Umkehrung  $x = f^{-1}(y)$  stetig und einmal stetig differenzierbar ist (Ausnahmen sind höchstens  $d - 1$ -Ränder). Dann gilt für  $Y = f(X)$ :

$$p_Y(y_1, \dots, y_d) = p_X(f_1^{-1}(y), \dots, f_d^{-1}(y)) |J(y_1, \dots, y_d)|$$

$J(y_1, \dots, y_d)$  bezeichnet die Funktional- bzw. Jacobi-Determinante von  $x = f^{-1}(y)$ , die gegeben ist durch

$$J(y_1, \dots, y_d) = \text{Det} \begin{pmatrix} \frac{\partial s_1(y_1, \dots, y_d)}{\partial y_1} & \dots & \frac{\partial s_1(y_1, \dots, y_d)}{\partial y_d} \\ \vdots & \ddots & \vdots \\ \frac{\partial s_d(y_1, \dots, y_d)}{\partial y_1} & \dots & \frac{\partial s_d(y_1, \dots, y_d)}{\partial y_d} \end{pmatrix}$$

Da eine analytische Berechnung der übertragenen Verteilung kaum möglich ist, werden in Kapitel 3.3 Wege für eine numerische Bestimmung aufgezeigt.

### 3.1.4 Eigenschaften der übertragenen Verteilung

Bei der robusten Optimierung soll die Streuung von  $Y = f(X)$  minimiert werden. Hier werden die wichtigsten Eigenschaften der Verteilung von  $Y$  aufgeführt.

Eine Normalverteilung der Zufallsvariablen  $X$  wird nur im Fall einer linearen Funktion  $f$  auch auf die Zufallsvariable  $Y = f(X)$  übertragen. Die Abbildungen 3.1 bis 3.5 zeigen einfache Beispiele, in denen die Verteilung von  $Y$  weder symmetrisch noch normalverteilt ist, obwohl  $X$  es war. Sie zeigen auch noch einige andere wichtige Eigenschaften von übertragenen Dichtefunktionen:

- Falls  $f$  unstetig in  $x_0$  ist, ist die übertragene Verteilung in  $f(x_0)$  ebenfalls unstetig (Abb. 3.1, die übertragenen Dichtefunktionen werden fälschlicherweise als stetig dargestellt).
- Dies gilt auch, wenn  $f$  ein lokales Extremum in  $x_0$  besitzt, das kein globales ist (Abb. 3.2).
- Falls mehrere lokale Extrema im Bereich der Verteilung um  $x_0$  liegen, wird die übertragene Verteilung multimodal (Abb. 3.3).
- Die Verteilung von  $Y$  ist in einem globalen Extremum von  $f$  extrem verzerrt (Abb. 3.4).
- Übertragene Verteilungen können sehr endlastig sein (Abb. 3.5).

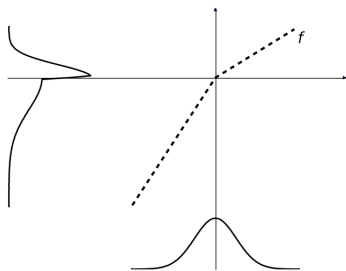


Abb. 3.1:  $f$  unstetig

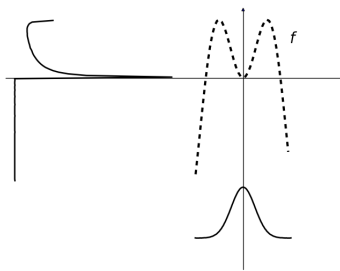
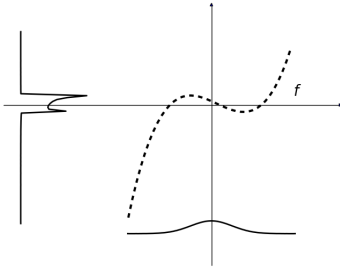
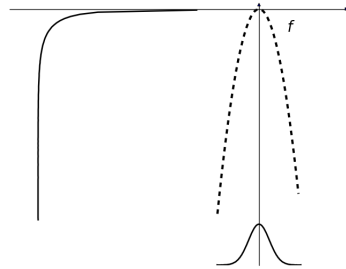
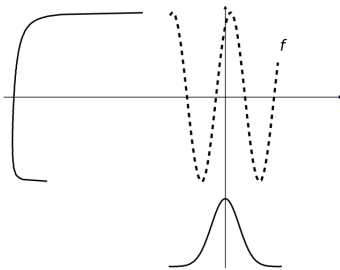


Abb. 3.2:  $f$  mit lokalem Extremum

Diese Beispiele demonstrieren, dass übertragene Verteilungen beliebige Formen annehmen können. Später werden statt der Funktion  $f$  Metamodelle verwendet, die immer stetig sind. Selbst in diesem Fall kann die Dichtefunktion von  $Y$  unstetig, unsymmetrisch und multimodal sein. Im Kapitel 4.5 wird eine geeignete Charakterisierung genauer beleuchtet.

## 3.2 Versuchspläne

Versuchspläne wurden ursprünglich zur Durchführung *physikalischer* Experimente entwickelt, etwa in der Landwirtschaft zur Optimierung von Ernteerträgen oder in

Abb. 3.3:  $f$  mit zwei lokalen ExtremaAbb. 3.4:  $f$  mit globalem ExtremumAbb. 3.5:  $f$  mit globalem Maximum und Minimum

der experimentellen Chemie. Da diese immer mit zufälligen Fehlern behaftet sind, werden die Stichprobenelemente in den klassischen Versuchsplänen oft an den Rändern des Designraums angeordnet, um so den Einfluss dieser Messungenauigkeiten zu verringern.

Später wurden Versuchspläne auch für Computereperimente benutzt. Da diese frei von Fehlern sind - in dem Sinne, dass gleiche Kombinationen von Eingangsvariablen gleiche Ergebnisse liefern - werden hier Pläne benötigt, die den Designraum möglichst gleichmäßig ausfüllen. Aus den Ergebnissen werden dann Metamodelle erstellt. Die in diesem Unterkapitel vorgestellten Versuchspläne gehören zu dieser neueren Generation.

In dieser Arbeit werden Versuchspläne für zwei verschiedene Zwecke benutzt,

- zum einen, um die Stützstellen für die Metamodelle geeignet auszuwählen,
- zum anderen, um eine Stichprobe um einen Punkt  $x_0$  zu verteilen. Für diese Stichprobe werden die Funktionswerte bestimmt und hieraus dann statistische Daten wie Erwartungswert, Standardabweichung oder Quantile berechnet.

Wünschenswerte Eigenschaften von Versuchsplänen sind deshalb



- zur Metamodellerstellung eine gleichmäßige Verteilung der Stichprobenelemente im Designraum und
- zur statistischen Auswertung bei mehreren Dimensionen eine gleichmäßige Verteilung der Stichprobenelemente für die Randverteilungen.

### 3.2.1 Monte Carlo-Verfahren

Das Monte Carlo-Verfahren (MC) [Rub81] ist wohl der einfachste aller Versuchspläne. Wie der Name schon erahnen lässt, entsteht eine Stichprobe durch das zufällige  $n$ -malige Ziehen gemäß der Verteilung von  $X$ . In der Praxis erzeugt ein Zufallszahlengenerator Zahlen im Einheitswürfel  $[0; 1]^d$  (vgl. Abb. 3.6). Diese werden mittels Transformationen in beliebige Verteilungen übertragen. In dieser Dissertation wird die Box-Muller-Transformation zur Erzeugung standardnormalverteilter Zufallsvariablen benutzt. Eine beliebige Normalverteilung kann anschließend über  $y = x \cdot \sigma + \mu$  erreicht werden.

Monte Carlo Stichproben lassen sich leicht erstellen und fortsetzen. Sie füllen den Designraum allerdings nur langsam gleichmäßig aus. Bei kleinen Stichproben können Bereiche außen vor bleiben. In der Literatur finden sich zahlreiche Verbesserungen und Abwandlungen des „reinen“ Monte Carlo-Verfahren, siehe hierzu z.B. [ES00].

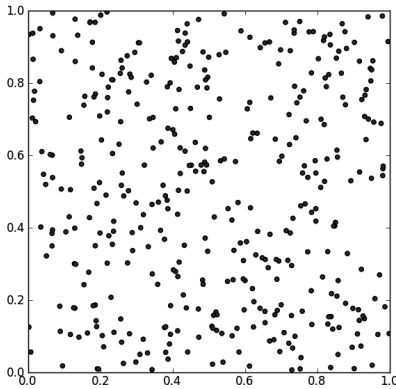


Abb. 3.6: Monte Carlo,  $n = 400$ , Gleichverteilung

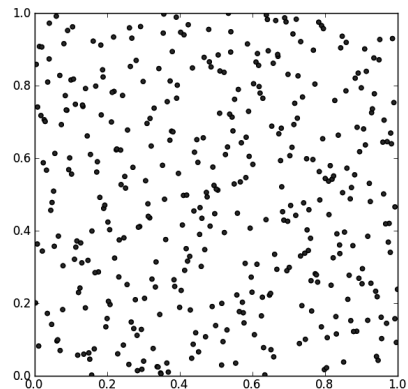


Abb. 3.7: Latin Hypercube Sampling,  $n = 400$ , Gleichverteilung

### 3.2.2 Latin Hypercube Sampling

Das Latin Hypercube Sampling (LHS) ist eine spezielle geschichtete Zufallsstichprobe, die von McKay, Conover und Beckman [MBC79] mit dem Ziel entwickelt wurde, die empirische Verteilungsfunktion sowie Erwartungswert und Varianz einer übertragenen Verteilung effizienter als beim Monte-Carlo-Verfahren zu schätzen. Heute wird es häufig auch für die Erzeugung von Stützstellen für Metamodelle benutzt. Zur seiner Erstellung wird der Designraum für jede Dimension in  $n$  Streifen gleicher Wahrscheinlichkeit unterteilt. Auf diese Weise entstehen  $n^d$  Zellen. Dann wird ein

beliebiger Wert aus einer zufällig selektierten Zelle als erstes Element  $x_1$  ausgewählt. Für jede Dimension werden dann alle Zellen der Streifen, in denen  $x_1$  liegt, von der weiteren Auswahl ausgeschlossen. Aus den verbliebenen Streifen wird wieder zufällig das nächste Element  $x_2$  bestimmt und die zugehörigen Zellen in den Streifen gestrichen. Dieser Vorgang wird bis zum  $n$ -ten Element fortgesetzt (vgl. Abb. 3.7). Bei dieser Vorgehensweise kann es zu Ungleichverteilungen der Stichprobenelemente kommen, z.B. ist ihre diagonale Verteilung bei  $d = 2$  auch ein Latin Hypercube Sampling. In der Praxis wird daher mit LHS gearbeitet, bei denen die Summe der Abstände aller Elemente minimiert wird ([KN05], S.263) oder der minimale Abstand zweier Stichprobenelemente maximiert wird [DK10, LS10, MM95]. Ein Latin Hypercube Sampling lässt sich nachträglich nicht um weitere Elemente erweitern. In diesem Fall muss eine neue Unterteilung des Designraums in Streifen vorgenommen werden.

### 3.2.3 Stratified Sampling

Unter Stratified Sampling (SS) soll hier eine spezielle Form der geschichteten Zufallsstichprobe verstanden werden: Für jede Dimension wird der Designraum aufgeteilt in  $\sqrt[n]{n}$  disjunkte Intervalle gleicher Wahrscheinlichkeit, so dass sich  $n$  Zellen gleicher Wahrscheinlichkeit ergeben. Für die Stichprobe wird aus jeder Zelle ein Wert zufällig ausgewählt (vgl. Abb. 3.8). Daher sind nur Stichprobengrößen von  $a^d$  mit  $a \in \mathbb{N}$  möglich. In einer Dimension besteht kein Unterschied zwischen Latin Hypercube Sampling und Stratified Sampling. Ein Stratified Sampling kann nicht erweitert werden, sondern es muss für eine andere Stichprobengröße neu aufgebaut werden.

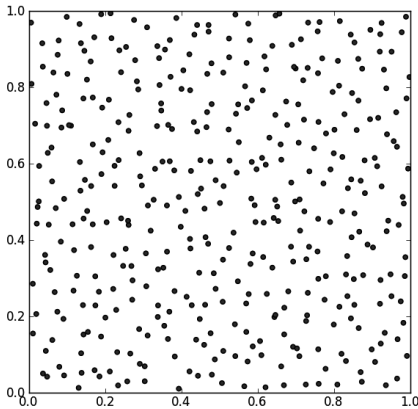


Abb. 3.8: Stratified Sampling,  $n = 400$ , Gleichverteilung

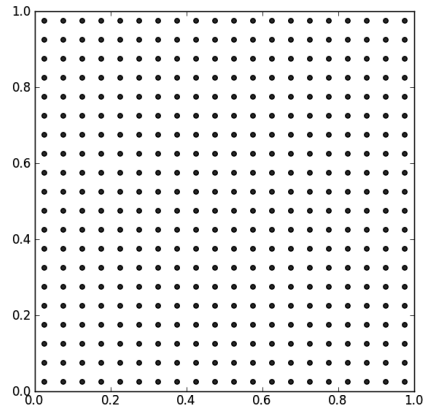


Abb. 3.9: Centered Stratified Sampling,  $n = 400$ , Gleichverteilung

### 3.2.4 Centered Stratified Sampling

Ein Centered Stratified Sampling (CSS) wird genauso aufgebaut wie ein Stratified Sampling mit dem einzigen Unterschied, dass in jeder Zelle der Wert nicht zufällig ausgewählt wird, sondern immer die Mitte der Zelle als Stichprobenelement bestimmt wird. Centered Stratified Sampling ist daher eine deterministische Folge, die keine Zufallskomponente enthält (vgl. Abb. 3.9). Sie ist nicht nachträglich ergänzbar.

### 3.2.5 Haltonfolge

Die Haltonfolge (HAL) [Hal64] gehört zur Klasse der quasi-zufälligen Folgen. Diese wurden ursprünglich zur effizienten numerischen Berechnung von Integralen höherdimensionaler Funktionen entwickelt. Da viele statistische Größen auch über Integrale berechnet werden können, liegt ihre Anwendung für diesen Zweck nahe.

In der Grundform produziert die Haltonfolge eine gleichmäßig verteilte Folge von Zahlen im Intervall  $[0; 1]$ . Sie wird aus der  $p$ -adischen Darstellung der Primzahl  $p$  gewonnen. Jede natürliche Zahl  $g$  kann in der Form

$$g = \sum_{i=0}^r a_i \cdot p^i \quad a_i \in [0; p-1]$$

dargestellt werden. Dann ist die durch

$$\Upsilon_p(g) = \sum_{i=0}^r \frac{a_i}{p^i}$$

definierte Funktion  $\Upsilon_p(g)$  eine Funktion auf  $[0; 1]$  und  $\Upsilon_p(2), \Upsilon_p(3), \dots$  bilden die Haltonfolge (vgl. Abb. 3.10).

Die Erweiterung auf  $d$  Dimensionen erfolgt, in dem  $\Upsilon_p(g)$  für paarweise verschiedene Primzahlen  $p = p_1, p_2, \dots, p_d$  bestimmt und zur *Haltonfolge*

$$(\Upsilon_{p_1}(g), \Upsilon_{p_2}(g), \dots, \Upsilon_{p_d}(g)), \quad g = 2, 3, \dots$$

zusammengesetzt wird.

Ein Maß für die Gleichmäßigkeit der Stichprobenelementeverteilung ist die *Stern-Diskrepanz*:

Gegeben sei eine Folge  $x_1, x_2, \dots, x_n$  über dem Einheitswürfel  $[0; 1]^d$ . Der Quader  $W_t = \{(u_1, \dots, u_d) \mid 0 \leq u_i \leq t_i, i = 1, \dots, d\}$  für  $t = (t_1, \dots, t_d) \in [0; 1]^d$  ist dann eine Teilmenge von  $[0; 1]^d$  mit Volumen  $\mu(W_t) = t_1 t_2 \dots t_d$ . Dann ist die *Stern-Diskrepanz* definiert als

$$D_n^* := D(x_1, \dots, x_n) := \sup_{t \in [0; 1]^d} \left| \mu(W_t) - \frac{1}{n} \sum_{\nu=1}^n 1_{x_\nu \in W_t} \right|.$$

$x_1, x_2, \dots, x_n$  heißt *Folge mit niedriger Diskrepanz*, falls

$$D_n^* = \mathcal{O}\left(\frac{(\log n)^d}{n}\right).$$

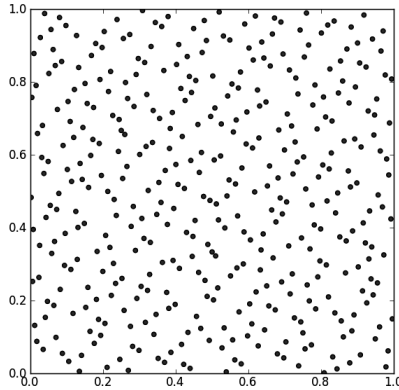


Abb. 3.10: Haltonfolge mit  $p_1 = 2$  und  $p_2 = 3$ ,  $n = 400$ , Gleichverteilung

Die Haltonfolge ist eine Folge mit niedriger Diskrepanz. Ihre Konvergenzrate beträgt  $\mathcal{O}((\log n)^d/n)$  im Gegensatz zu  $\mathcal{O}(1/\sqrt{n})$  bei Monte Carlo ([Caf98]).

Falls in hohen Dimensionen die Primzahlen in ihrer aufsteigenden Reihenfolge zur Konstruktion der Haltonfolge herangezogen werden, treten bei manchen Primzahlpaaren Korrelationen auf. Manteufel ([Man01]) gibt ein Beispiel für  $p_{15} = 47$  und  $p_{16} = 53$ . Dies kann durch den gezielten Ausschluss bestimmter Primzahlen vermieden werden, vgl. hierzu [RA99]. Zudem wird der Raum für kleine Stichproben in hohen Dimensionen nicht gleichmäßig ausgefüllt [HP03].

Einen aktuellen Überblick zu quasi-zufälligen Versuchsplänen im allgemeinen gibt [Lem09]. Zu dieser Gruppe gehört z.B. auch die Sobol-Folge [Sob67]. Eine Abwandlung der Haltonfolge für endliche Stichprobengrößen  $n$  ist die Hammersleyfolge [Ham60]. Zu den Basen  $p_1, p_2, \dots, p_{n-1}$  ist sie definiert durch

$$x_g = \left( \frac{g-1}{n}, \Upsilon_{p_1}(g-1), \Upsilon_{p_2}(g-1), \dots, \Upsilon_{p_{n-1}}(g-1) \right), \quad g = 1, 2, \dots, n$$

Soll der Versuchsplan nachträglich nicht mehr ergänzt werden, ist die Hammersleyfolge der Haltonfolge vorzuziehen, denn die Diskrepanzabschätzung ist um den Faktor  $(p_n - 1) \log n / 2 \log(p_n)$  kleiner als die der entsprechenden Haltonfolge [Nie78].

### 3.2.6 Andere Versuchspläne

Die hier vorgestellten Versuchspläne wurden ausgewählt, da sie am weitesten verbreitet sind. Darüber hinaus gibt es noch eine Reihe weiterer wie orthogonale Felder und Voronoi-Diagramme sowie importance-, directive- und directional-Sampling.

### 3.2.7 Diskussion

In den letzten Jahrzehnten wurden zahlreiche Vergleiche von Versuchsplänen für die Erstellung eines Metamodells veröffentlicht ([SSG06], [ZX10]).

In dieser Arbeit wird hingegen untersucht, welcher Versuchsplan sich für die Schätzung übertragener Quantile am besten eignet. Eine solcher Vergleich findet sich

bisher nicht in der Literatur. Es gibt aber mehrere, die sich mit der Schätzung anderer statistischer Maße der übertragenen Verteilung beschäftigen:

Mc Kay et al. [MBC79] vergleichen MC, LHS und SS. Sie zeigen, dass die Standardschätzer für Erwartungswert und Standardabweichung sowie der empirischen Verteilungsfunktion für diese Versuchspläne unverzerrt sind. Sie geben Formeln für die Varianz der Schätzer an. Sie ist bei SS immer kleiner als bei MC. Die Varianz bei LHS ist für monotone Übertragungsfunktionen kleiner als bei MC. In einem Beispiel schneidet LHS deutlich besser als SS und MC ab. Für LHS beträgt die Konvergenzrate für die empirische Verteilungsfunktion wie bei MC  $1/\sqrt{n}$  [HD03].

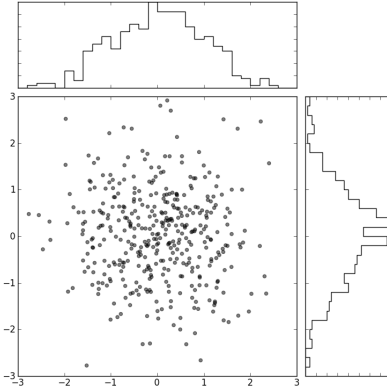


Abb. 3.11: Monte Carlo,  $n = 400$ , Normalverteilung

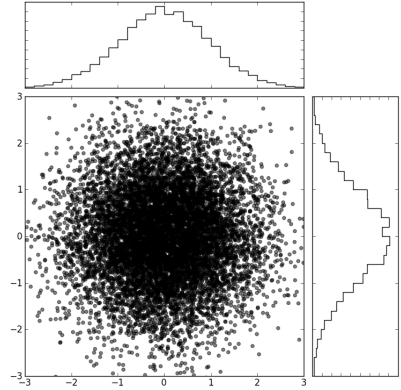


Abb. 3.12: Monte Carlo,  $n = 10000$ , Normalverteilung

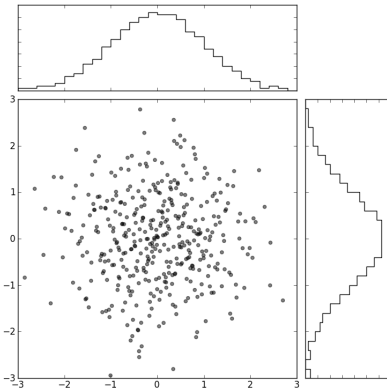


Abb. 3.13: Latin Hypercube Sampling,  $n = 400$ , Normalverteilung

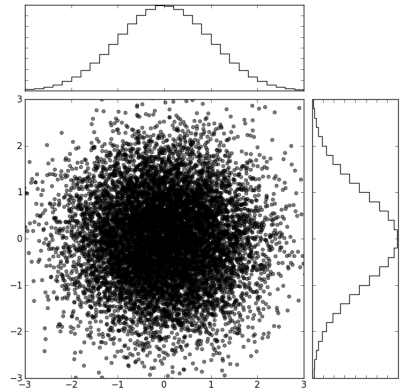


Abb. 3.14: Latin Hypercube Sampling,  $n = 10000$ , Normalverteilung

In [RBGP06] werden MC, LHS und HAL für Testfunktionen miteinander verglichen.

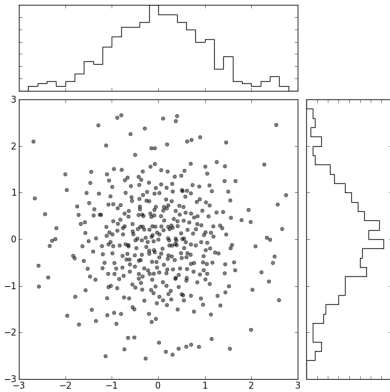


Abb. 3.15: Stratified Sampling,  $n = 400$ , Normalverteilung

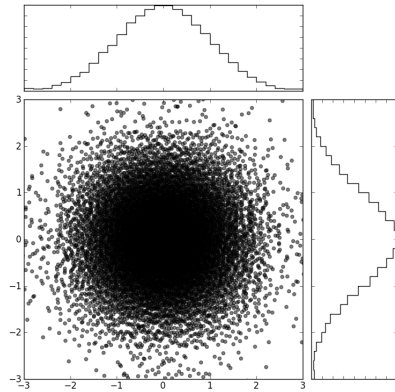


Abb. 3.16: Stratified Sampling,  $n = 10000$ , Normalverteilung

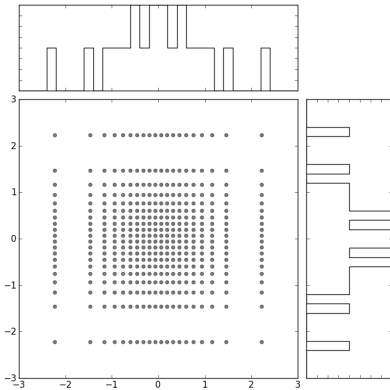


Abb. 3.17: Centered Stratified Sampling,  $n = 400$ , Normalverteilung

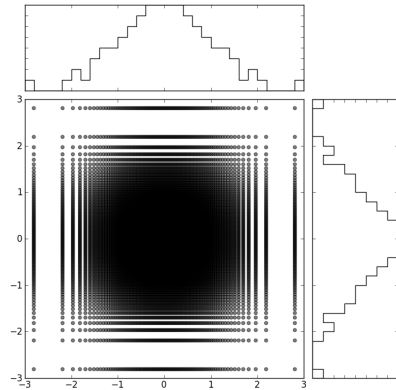


Abb. 3.18: Centered Stratified Sampling,  $n = 10000$ , Normalverteilung

Es wird die Genauigkeit des Erwartungswertes, der Standardabweichung und zweier exceeding-Wahrscheinlichkeiten untersucht. Dabei schneidet LHS besser als MC und dieser besser als HAL ab. Dieses Ergebnis erscheint vor dem Hintergrund der eigenen Untersuchung in Kapitel 8.4 überraschend. Es könnte aber dadurch erklärt werden, dass nur eine Testfunktion und drei Stichproben verwendet wurden.

Huang [HD07] vergleicht MC, LHS und Hammersleyfolge für Erwartungswert und Standardabweichung bei verschiedenen Testfällen und bis zu einer Stichprobengröße von einer Million. Dort sind LHS und Hammersley MC überlegen. Die Hammersleyfolge konvergiert gleichmäßiger.

Die genannten Untersuchungen verwenden meist kleine Stichproben bis höchstens  $n = 100$  und benutzen nur den Standardschätzer. Auf Metamodellen sind aber

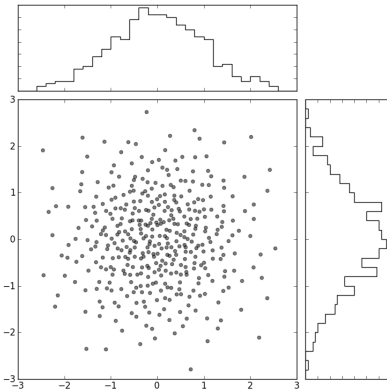


Abb. 3.19: Haltonfolge,  $n = 400$ , Normalverteilung

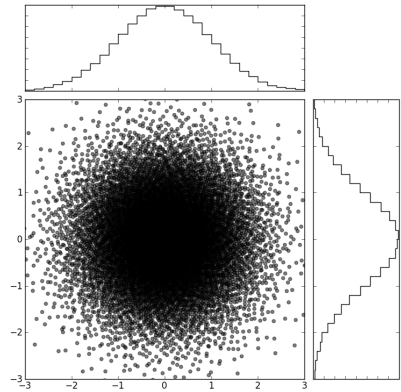


Abb. 3.20: Haltonfolge,  $n = 10000$ , Normalverteilung

wesentlich größere Stichproben möglich und für das Erreichen einer gewünschten Genauigkeit auch nötig. Die Genauigkeit einer Quantilschätzung wurde noch nicht betrachtet, ebenso wenig, wie sich die Kombination mit anderen Schätzern auswirkt. Diese Punkte werden in Kapitel 5 aufgegriffen.

An dieser Stelle werden die Versuchspläne nur visuell hinsichtlich der Gleichmäßigkeit der Verteilung der Punkte im Designraum und der Randverteilungen verglichen. Die Grafiken 3.11 bis 3.20 zeigen die genannten Versuchspläne und ihre Randverteilungen bei zweidimensionaler Standardnormalverteilung jeweils für  $n = 400$  und  $n = 10000$ . LHS, SS und HAL verteilen die Punkte gleichmäßig gemäß der Normalverteilung im Raum, MC ist etwas unregelmäßiger, während CSS ein regelmäßiges Gitter hervorbringt.

Gute Randverteilungen liefert schon bei  $n = 400$  LHS. CSS hat hier deutliche Defizite, die auch bei  $n = 10000$  nicht verschwinden. Für diese Stichprobengröße sind die Randverteilungen wieder bei LHS am besten verteilt, gefolgt von HAL und SS, am Schluss MC und CSS.

Die folgende Tabelle gibt noch einmal einen Überblick über die Eigenschaften der verschiedenen Versuchspläne:

	MC	LHS	SS	CSS	HAL
ergänzzbar	ja	nein	nein	nein	ja
gleichmäßige Randvert.	ja	ja	ja	nein	ja
Konvergenzrate	$\mathcal{O}(\frac{1}{\sqrt{n}})$	$\mathcal{O}(\frac{1}{\sqrt{n}})$	$\mathcal{O}(\frac{1}{\sqrt{n}})$	$\mathcal{O}(\frac{1}{\sqrt{n}})$	$\mathcal{O}(\frac{(\log(n))^d}{n})$

Für kleine Dimensionen besitzt die Haltonfolge die beste Konvergenzrate. Die übrigen Versuchspläne besitzen zwar die gleiche Konvergenzrate, doch für kleine Stichprobengrößen  $n$  sind Latin Hypercube Sampling und Stratified Sampling dem Monte Carlo Verfahren vorzuziehen.

### 3.3 Nichtparametrische Quantilschätzer

Zur Schätzung von Quantilen werden oft parametrische Quantilschätzer benutzt. Dabei wird eine Annahme über die Klasse der unbekanntenen Verteilung  $X$  gemacht und anschließend ihre Parameter aus der Stichprobe  $(x_1, \dots, x_n)$  geschätzt. Mit diesen können dann die Quantile der nun bekannten Verteilung berechnet werden. Wie in Kapitel 3.1.4 gezeigt, können übertragene Verteilungen beliebige Formen annehmen, die zu keiner Verteilungsklasse gehören. Nur in seltenen Fällen lassen sie sich analytisch bestimmen. Daher ist die parametrische Quantilschätzung für diese Anwendung nicht geeignet.

Ohne die Annahme einer Verteilungsklasse kommen die nichtparametrischen Quantilschätzer aus. Sie schätzen ein Quantil direkt aus den Stichprobenwerten. Ihr Anwendungsgebiet liegt vor allem im Finanzbereich zur Berechnung des Standardrisikomaßes *Value at Risk* ([IKK05]).

Es gibt eine Vielzahl von nichtparametrischen Quantilschätzern, von denen die wichtigsten anschließend vorgestellt werden. Weitere finden sich bei [SM90] und [Hut01]. Da die folgenden Schätzer auf die Stichprobenwerte der übertragenen Verteilung, d.h. auf  $y_1, \dots, y_n$  angewendet werden, wird diese Notation verwendet. Die Definitionen gelten analog für jede beliebige Stichprobe  $x_1, \dots, x_n$ .

Für den Rest des Kapitels wird  $l = 1$  angenommen. Wenn es mehr als eine Zielvariable gibt, sind die Ergebnisse für jedes einzelne  $Y_i, i = 1, \dots, l$  gültig.

#### 3.3.1 Empirischer Schätzer

Der *empirische Schätzer* (EMP) oder *Standard-Schätzer für das  $q$ -Quantil* von  $Y$  ist definiert durch:

$$\hat{Y}_{q,n} = \inf\{x, F_{Y_n}(x) > q\} = Y_{([\![q \cdot n]\!] )}$$

wobei  $[x]$  die zu  $x$  nächstgrößere ganze Zahl ist.

Da die empirische Verteilungsfunktion  $F_{Y_n}$  eine Treppenfunktion ist, führt auch ihre verallgemeinerte Umkehrung zu einer linksseitig-stetigen Treppenfunktion.

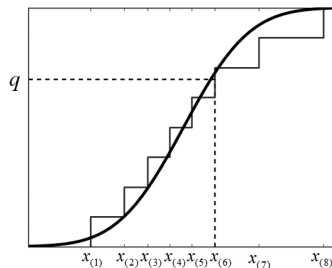


Abb. 3.21: Veranschaulichung des empirischen Schätzers EMP

Abbildung 3.21 veranschaulicht die Bestimmung des empirischen Standardschätzers anhand der Verteilungsfunktion der Standardnormalverteilung. Er ist besonders in



den flachen Bereichen von  $F$  ungenau, d. h. falls  $F'(y)$  klein ist bzw.  $Q'(q)$  groß.  $\hat{Y}_{q,n}$  ist für jede endliche Stichprobengröße  $n$  ein verzerrter Schätzer des  $q$ -Quantils, aber asymptotisch unverzerrt.

Es gilt:

$$\sqrt{n}(\hat{Y}_{q,n} - y_q) \rightarrow \mathcal{N}(0, \text{AVar}_{\text{EMP}}) \quad \text{für } n \rightarrow \infty \quad (3.2)$$

mit der asymptotischen Varianz

$$\text{AVar}_{\text{EMP}} = \frac{q(1-q)}{p^2(y_q)} = q(1-q)Q'^2(q) \quad (3.3)$$

### Näherungsformeln

Über Taylorreihenentwicklung lassen sich für den Erwartungswert und die Varianz von  $\hat{Y}_{q,n}$  folgende asymptotische Ergebnisse zeigen (alle in [DN03], S. 84)

$$\begin{aligned} \mathbb{E}(\hat{Y}_{q,n}) &= y_q - \frac{q(1-q)p'_Y(y_q)}{2(n+2)p_Y^3(y_q)} + \mathcal{O}\left(\frac{1}{n^2}\right) \\ &= Q(q) + \frac{q(1-q)}{2(n+2)}Q''(q) + \mathcal{O}\left(\frac{1}{n^2}\right) \end{aligned} \quad (3.4)$$

$$\begin{aligned} \text{Var}(\hat{Y}_{q,n}) &= \frac{q(1-q)}{(n+2)p_Y^2(y_q)} + \mathcal{O}\left(\frac{1}{n^2}\right) \\ &= \frac{q(1-q)}{(n+2)}Q'(q)^2 + \mathcal{O}\left(\frac{1}{n^2}\right) \end{aligned} \quad (3.5)$$

$$\begin{aligned} \text{Cov}(\hat{Y}_{1-q,n}, \hat{Y}_{q,n}) &= \frac{(1-q)^2}{(n+2)p_Y(y_q)p_Y(y_{1-q})} + \mathcal{O}\left(\frac{1}{n^2}\right) \\ &= \frac{(1-q)^2}{(n+2)}Q'(q)Q'(1-q) + \mathcal{O}\left(\frac{1}{n^2}\right) \end{aligned} \quad (3.6)$$

wobei  $p_Y$  die stetig differenzierbare Dichtefunktion von  $Y$  ist.

Für den *mittleren quadratischen Fehler* MSE erhält man hieraus für den empirischen Standardschätzer die Näherung

$$\begin{aligned} \text{MSE}(\hat{Y}_{q,n}) &= \underbrace{(\mathbb{E}(\hat{Y}_{q,n}) - y_q)^2}_{\mathcal{O}\left(\frac{1}{n^2}\right)} + \text{Var}(\hat{Y}_{q,n}) \\ &= \text{Var}(\hat{Y}_{q,n}) + \mathcal{O}\left(\frac{1}{n^2}\right) \\ &= \frac{1}{n+2} \text{AVar}(\hat{Y}_{q,n}) + \mathcal{O}\left(\frac{1}{n^2}\right) \end{aligned} \quad (3.7)$$

Der Berechnungsaufwand für EMP beträgt  $\mathcal{O}(n \log(n))$ , da die Stichprobe geordnet werden muss. Hierfür ist der Speicherbedarf  $\mathcal{O}(n)$  nötig.

Der Standardschätzer ist bei Verwendung eines Latin Hypercube Versuchsplans ebenfalls unverzerrt und asymptotisch normalverteilt. Seine Varianz ist kleiner gleich der eines Monte Carlo Versuchsplans ([AW98]).

Eine naheliegende Verbesserung des empirischen Schätzers liegt darin, mehrere Ordnungsstatistiken mit in die Schätzung einzubeziehen, z.B. eine Linearkombination aus  $Y_{(\lceil q \cdot (n-1) \rceil)}$  und  $Y_{(\lceil q \cdot n \rceil)}$  zu bilden. Auch in vielen Softwareprogrammen ist der Standard-Schätzer nicht wie oben definiert programmiert, sondern mit kleinen Abweichungen. Einen Überblick und Empfehlungen hierzu gibt [HF96].

Die folgenden Quantilschätzer beziehen *alle* Stichprobenelemente mit in die Schätzung ein.

### 3.3.2 Kerndichteschätzer

Kerndichteschätzung (KDE) wurde ursprünglich von Parzen entwickelt, um die Dichtefunktion aus einer Stichprobe zu bestimmen [Par62]. Die Grundidee besteht darin, um jedes Stichprobenelement einen symmetrischen Kern zu platzieren. Zur Schätzung der Dichtefunktion werden alle Kerne addiert und die Summe normiert (vgl. Abbildung 3.22).

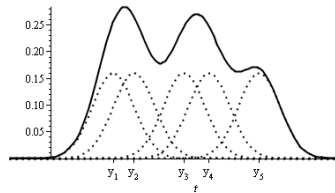


Abb. 3.22: Beispiel einer Kerndichteschätzung für fünf Stichprobenelemente und Gausskern

Sie hat somit die Form

$$\hat{p}(t) = \frac{1}{n} \sum_{i=1}^n k_h(t - y_i) = \frac{1}{nh} \sum_{i=1}^n k\left(\frac{t - y_i}{h}\right)$$

wobei  $k$  eine Kernfunktion ist.  $h$  wird *Glättungsparameter* oder *Bandweite* genannt. Eine *Kernfunktion* ist eine beschränkte und zu 0 symmetrische Dichtefunktion, für die außerdem gilt:

$$|x| \cdot k(x) \rightarrow 0 \quad \text{für} \quad x \rightarrow 0 \quad \text{und} \quad \int x^2 k(x) dx < \infty$$

Gebräuchliche Kernfunktionen sind:

1. Gauss-Kern:

$$k_1(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}t^2\right)$$

2. Epanechnikov-Kern:

$$k_2(t) = \begin{cases} \frac{3}{4}(1 - t^2) & -1 \leq t \leq 1 \\ 0 & \text{sonst} \end{cases}$$

3. Dreieckskern:

$$k_3(t) = \begin{cases} 0 & t < -1 \\ 1 - |t| & -1 \leq t \leq 1 \\ 0 & t > 1 \end{cases}$$

4. gleichverteilter Kern:

$$k_4(t) = \begin{cases} 0 & t < -1 \\ \frac{1}{2} & -1 \leq t \leq 1 \\ 0 & t > 1 \end{cases}$$

5. Cauchy-Kern:

$$k_5(t) = \frac{1}{\pi(1+t^2)}$$

Die Kerndichteschätzung ist in der Lage auch mehrgipflige und verzerrte Dichtefunktionen zu schätzen.

Großen Einfluss auf die Güte der Schätzung hat die Wahl der Bandweite  $h$ . Wird  $h$  groß gewählt, erhält man einen glatten Verlauf der Dichtefunktion. Bei kleinen  $h$  wird er ungleichmäßiger, er ähnelt einem Histogramm mit kleiner Säulenbreite.

### 3.3.3 Umkehrung der Kerndichteverteilungsfunktion

Eine Schätzung für die Verteilungsfunktion erhält man durch Integration von  $\hat{p}$ . Dies führt zu [Azz81]:

$$\hat{F}_n(t) = \int_{-\infty}^t \hat{p}(u) du = \frac{1}{n} \sum_{i=1}^n K_h(t - f(x_i))$$

mit  $K_h(t) = \int_{-\infty}^t k_h(u) du$ .

Zur Schätzung des  $q$ -Quantils ist schließlich die Gleichung

$$\hat{F}_n(t) = q \tag{3.8}$$

zu lösen.

Der zugehörige Schätzer ist dann

$$\hat{Q}_n(q) = \inf\{t : \hat{F}_n(t) \geq q\}, \quad 0 < q < 1$$

Cheng [CS06] zeigt, dass für den mittleren quadratischen Fehler (MSE) asymptotisch gilt

$$\text{MSE}(\hat{Q}_n(q)) = \frac{q(1-q)}{np(y_q)^2} + \frac{h^4 p'(y_q)^2}{4p(y_q)^2} \mu_2(k)^2 - \frac{2h}{n} \frac{1}{p(y_q)} \psi(k)$$

falls

1.  $h \rightarrow 0$  für  $n \rightarrow \infty$
2.  $p$  differenzierbar mit beschränkter Ableitung  $p'$  und

3.  $p'$  stetig in der Nachbarschaft von  $y_q$  mit  $p'(y_q) \neq 0$

mit  $\mu_2(k) = \int_{-\infty}^{\infty} t^2 k(t) dt$  und  $\psi(k) = \int_{-\infty}^{\infty} t \cdot k(t) K(t) dt > 0$ .

Die optimale Bandweite  $\hat{h}_{opt}$  ist dann [Sha98]:

$$\hat{h}_{opt} = \left( \frac{p(y_q) 2\psi(k)}{n(p'(y_q)^2 \mu_2(k)^2)} \right)^{\frac{1}{3}} = \left( \frac{Q'(q)^5 2\psi(k)}{n(Q''(q)^2 \mu_2(k)^2)} \right)^{\frac{1}{3}} \quad (3.9)$$

Dieses Ergebnis zeigt, dass die optimale Bandweite von der Steigung und Krümmung der unbekanntem Quantilfunktion an der Stelle  $q$  abhängt. Sie verändert sich in Bezug auf die Stichprobengröße  $n$  mit  $n^{-1/3}$ . Die Eigenschaften der Kernfunktion gehen über  $\psi$  und  $\mu$  in die Formel ein.

Für den Gauss-Kern erhält man zur Schätzung der Verteilungsfunktion den Ausdruck

$$\hat{F}_{n1}(t) = \frac{1}{2n} \left( \sum_{i=1}^n \operatorname{erf}\left(\frac{1}{2} \frac{\sqrt{2} \cdot (t - f(x_i))}{h}\right) \right) + \frac{1}{2}$$

mit  $\operatorname{erf}(t) = \frac{2}{\sqrt{\pi}} \int_0^t \exp(-u^2) du$  sowie  $\mu_2(k) = 1$  und  $\psi(k) = 1/2\sqrt{\pi}$ .

Die optimale Bandweite beträgt dann

$$\hat{h}_{opt} = \left( \frac{1}{\sqrt{\pi n}} \cdot \frac{p(y_q)}{p'(y_q)^2} \right)^{\frac{1}{3}} = \left( \frac{1}{\sqrt{\pi n}} \cdot \frac{Q'(q)^5}{Q''(q)^2} \right)^{\frac{1}{3}} \quad (3.10)$$

Azzalini [Azz81] erhält ein ähnliches Ergebnis für Kernfunktionen mit kompaktem Träger.

### 3.3.4 Glättung des empirischen Schätzers mit Kerndichtefunktionen

Eine zweite Möglichkeit der Quantilschätzung mit Kernfunktionen besteht darin, Kerne um die Ordnungsstatistiken zu bilden [Par79]:

$$\tilde{Q}_n(q) = \sum_{i=1}^n \left( \int_{\frac{i-1}{n}}^{\frac{i}{n}} \frac{1}{h} \cdot k\left(\frac{t-q}{h}\right) dt \right) \cdot y_{(i)} \quad (3.11)$$

Yang [Yan85] zeigt, dass beide Darstellungen asymptotisch äquivalent sind bzgl. MSE. Sheather und Marron [SM90] erhalten asymptotisch das folgende Ergebnis, unter der Voraussetzung, dass  $p$  nicht symmetrisch oder  $p$  symmetrisch und  $q \neq 0.5$  und  $h \rightarrow 0$  für  $n \rightarrow \infty$ :

$$\operatorname{MSE}(\tilde{Q}_n(q)) = \frac{q(1-q)}{n} Q'(q)^2 + \frac{1}{4} h^4 Q''(q)^2 \mu_2(k)^2 - \frac{2h}{n} (Q'(q))^2 \psi(k) + o\left(\frac{h}{n}\right) + o(h^4)$$

( $\mu_2(k)$  und  $\psi(k)$  wie oben).

Der erste Summand entspricht dabei der Varianz des empirischen Schätzers. Die Gleichung zeigt, dass Kerndichteschätzer eine Verbesserung der Schätzung gegenüber EMP bringen können. Für  $n \rightarrow \infty$  wird der Vorteil verschwinden. Die Verbesserung hängt - neben  $h$  - außerdem von der Steigung und der Krümmung der Quantilfunktion ab.

Falls  $Q'(q) > 0$  gilt, beträgt die optimale Bandweite:

$$\tilde{h}_{opt} = \left( \frac{Q'(q)^2 \cdot 2 \cdot \psi(k)}{n Q''(q)^2 \mu_2(k)^2} \right)^{\frac{1}{3}} \quad (3.12)$$

In die Abschätzung für MSE eingesetzt, ergibt dies:

$$\text{MSE}(\tilde{Q}_n(q)) = \frac{q(1-q)}{n} Q'(q)^2 + \mathcal{O}(n^{-4/3}) \quad (3.13)$$

Für den Gausskern dann:

$$\text{MSE}(\tilde{Q}_n(q)) = \frac{q(1-q)}{n} Q'(q)^2 + \frac{1}{4} h^4 Q''(q)^2 - \frac{h}{n} Q'(q)^2 \frac{1}{\sqrt{\pi}} \quad (3.14)$$

Und:

$$\tilde{h}_{opt} = \left( \frac{Q'(q)^2}{n Q''(q)^2 \sqrt{\pi}} \right)^{\frac{1}{3}} \quad (3.15)$$

Falls  $p$  symmetrisch, existiert keine optimale Bandweite für  $h$  [CS06].

### 3.3.5 Wahl der Bandweite

Die Formeln 3.9 und 3.12 legen zwei unterschiedliche optimale Bandweiten nahe, die asymptotisch äquivalent sind:

$$\hat{h}_{opt} = \left( \frac{Q'(q)^5 \cdot 2 \cdot \psi(k)}{n(Q''(q)^2 \mu_2(k)^2)} \right)^{\frac{1}{3}} = \tilde{h}_{opt} \cdot Q'(q)$$

$$\tilde{h}_{opt} = \left( \frac{Q'(q)^2 \cdot 2 \cdot \psi(k)}{n Q''(q)^2 \mu_2(k)^2} \right)^{\frac{1}{3}}$$

Sie sind optimal für festes  $q$ .

Ein weiteres Maß für die Genauigkeit ist der *mittlere quadratische Integrationsfehler* (mean integrated squared error) MISE.

$$\text{MISE}(\hat{F}_n) = E \left( \int [\hat{F}_n(t) - F(t)]^2 dt \right)$$

Er misst den mittleren quadratischen Fehler über alle  $q$ .

MISE wird minimiert für

$$h_{opt} = \left( \frac{2\psi}{R_1} \right)^{\frac{1}{3}} n^{-\frac{1}{3}} = \left( \frac{2\psi}{\int_0^1 \frac{Q''(q)^2}{Q'(q)^5} dq} \right)^{\frac{1}{3}} n^{-\frac{1}{3}} \quad (3.16)$$

wobei

$$\begin{aligned} R_1 &= \int_{-\infty}^{\infty} (p'(y_q))^2 dy_q \\ &\stackrel{y_q=Q(q), dy_q=Q'(q) dq}{=} \int_0^1 \left( -\frac{Q''(q)}{Q'(q)^3} \right)^2 Q'(q) dq \\ &= \int_0^1 \frac{Q''(q)^2}{Q'(q)^5} dq \end{aligned}$$

die „Rauheit“ der wahren aber unbekanntenen Dichtefunktion  $p$  von  $Y$  ist. Daher muss das optimale  $h$  aus  $\mathcal{O}(n^{-\frac{1}{3}})$  sein. Falls  $p$  die Normalverteilung mit der Standardabweichung  $\sigma$  ist, gilt für den Gausskern [Han04, BHP98]:

$$h_{opt} = 4^{\frac{1}{3}} \cdot \sigma \cdot n^{-\frac{1}{3}}$$

Für den allgemeinen Fall gibt es keine einfache Regel, wie  $h$  am besten zu wählen ist. Verschiedene Vorschläge finden sich z.B. in [SM90, Han04], eine Überblick geben [Tur93, WJ95].

Hjort zeigt, dass bei einer Wahl von  $h$  aus  $\mathcal{O}(n^{-\frac{1}{5}})$  (dies ist die optimale Wahl bei Schätzung der Dichtefunktion, nicht der Verteilungsfunktion), die Schätzung außerhalb jedes Konfidenzintervalls liegen kann [HW01].

### Vergleich der Kernfunktionen

Offensichtlich hat die Wahl der Kernfunktion Einfluss auf die Güte und Glattheit der geschätzten Verteilungsfunktion. Jones ([Jon90]) zeigt, dass bei einer geeigneten Wahl von  $h$  alle Kernfunktionen gute Ergebnisse liefern. Daher wird im folgenden nur der Gausskern untersucht.

Die Gleichung 3.8 muss numerisch gelöst werden. Hierfür wird ein geeigneter Startwert benötigt, damit das Verfahren konvergiert. In den Berechnungen hat es sich als sinnvoll erwiesen, EMP als Startwert zu benutzen. In diesem Fall erhöht sich der theoretisch mögliche Rechenaufwand von  $\mathcal{O}(n)$  auf  $\mathcal{O}(n \log(n))$ . Es wird angenommen, dass  $\text{erf}(x)$  intern gelistet vorliegt.

### 3.3.6 Harrell-Davis-Schätzer

Der Harrell-Davis-Schätzer gehört zur Klasse der  $L$ -Schätzer, die die Quantile aus der gewichteten Summe der Ordnungstatistiken schätzen. Er benutzt den Erwartungswert der  $(n+1)q$ -ten Ordnungsstatistik zur Schätzung des  $q$ -Quantils [HD82]. Damit berechnen sich die Gewichte aus:

$$\omega_{n,i} = \frac{1}{\beta((n+1)q, (n+1)(1-q))} \int_{\frac{i-1}{n}}^{\frac{i}{n}} u^{(n+1)q-1} (1-u)^{(n+1)(1-q)-1} du$$

$i = 1 \dots n$  mit  $\beta(x, y) = \Gamma(x) \cdot \Gamma(y) / \Gamma(x+y)$ .

Der Schätzer für das  $q$ -Quantil ist dann:

$$\text{HD}_q = \sum_{i=1}^n \omega_{n,i} y^{(i)}$$

Seine Berechnung ist aufwändiger als die des Standardschätzers, da zusätzlich die Gewichte bestimmt werden müssen. Falls er immer für eine feste Stichprobengröße  $n$  verwendet wird, empfiehlt es sich, die Gewichte vorher zu berechnen und in einer Datei abzuspeichern. In diesem Fall beträgt der Berechnungsaufwand wie bei EMP  $\mathcal{O}(n \log(n))$  und der Speicherbedarf  $\mathcal{O}(n)$ .

Der Harrell-Davis-Schätzer ist asymptotisch normalverteilt und unverzerrt [DN03, Bro07].

Er ist asymptotisch ein Kerndichteschätzer mit Gausskern gemäß Formel 3.11, dessen Bandweite

$$h_{\text{HD}} = \left( \frac{q(1-q)}{n+1} \right)^{\frac{1}{2}} \quad (3.17)$$

beträgt. Sie ist asymptotisch suboptimal, da  $h \in \mathcal{O}(n^{-1/2})$  und nicht  $\mathcal{O}(n^{-1/3})$ . Dies bewirkt, dass die Gewichte für sehr große Stichproben zu sehr um  $q$  konzentriert sind [SM90].

### 3.3.7 Sfakianikis-Verginis-Schätzer

Sfakianikis und Verginis [SV08] schlagen ebenfalls eine  $L$ -Schätzer vor, benutzen aber für die Gewichtung die Binomialverteilung  $B(n, p)$  mit den Parametern  $n$  und  $q$ . In ihrem Artikel vergleichen sie drei verschiedene Versionen ihres Schätzers. Hier wird nur die Version, die dort am besten abschneidet, nämlich Nr. 3 betrachtet. Sie ist gegeben durch

$$\text{SV3}_q = \sum_{i=1}^n B(i; n, q) y_{(i)} + (2y_{(1)} - y_{(2)}) B(0; n, q) \quad (3.18)$$

Der Sfakianikis-Verginis-Schätzer ist kaum untersucht. Hier erfolgt der Beweis, dass er asymptotisch ebenfalls ein Kerndichteschätzer mit Gausskern und der Bandweite

$$h_{\text{SV3}} = \left( \frac{q(1-q)}{n} \right)^{\frac{1}{2}}, \quad (3.19)$$

ist, denn es gilt:

$$\begin{aligned} B(i; n, q) &= \binom{n}{i} (1-q)^i q^{n-i} \\ &\stackrel{n \rightarrow \infty, nq(1-q) > 9}{\sim} \Phi\left(\frac{i-nq}{\sqrt{nq(1-q)}}\right) - \Phi\left(\frac{i-1-nq}{\sqrt{nq(1-q)}}\right) \\ &= \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{nq(1-q)}} \int_{i-1}^i e^{-\frac{1}{2}\left(\frac{u-nq}{\sqrt{nq(1-q)}}\right)^2} du \\ &\stackrel{v=u/n}{=} \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{nq(1-q)}} \int_{(i-1)/n}^{i/n} e^{-\frac{1}{2}\left(\frac{n(v-q)}{\sqrt{nq(1-q)}}\right)^2} ndv \\ &= \frac{\sqrt{\frac{n}{q(1-q)}}}{\sqrt{2\pi}} \frac{1}{\sqrt{2\pi}} \int_{(i-1)/n}^{i/n} e^{-\frac{1}{2}\left(\frac{(v-q)\sqrt{n}}{\sqrt{q(1-q)}}\right)^2} dv \\ &\stackrel{h=\sqrt{\frac{q(1-q)}{n}}}{=} \frac{1}{h} \frac{1}{\sqrt{2\pi}} \int_{(i-1)/n}^{i/n} e^{-\frac{1}{2}\left(\frac{(v-q)}{h}\right)^2} dv \\ &= \frac{1}{h} \int_{(i-1)/n}^{i/n} k_1\left(\frac{v-q}{h}\right) dv \end{aligned}$$

Dies entspricht der Darstellung von Gleichung 3.11. Die Einschränkung  $nq(1-q) > 9$  ist für festes  $q$  und  $n \rightarrow \infty$  immer erfüllt. Für  $q$  nahe 0 oder 1 wird das asymptotische Verhalten erst später auftreten.

Für den letzten Summanden in Gleichung 3.18 gilt:

$$B(0; n, q) = \binom{n}{0} (1-q)^0 q^{n-0} = q^n \xrightarrow{n \rightarrow \infty} 0$$

Damit ist auch die Bandweite des Sfakianikis-Verginis-Schätzers nicht optimal. Bei sehr großen Stichproben wird sein möglicher Vorteil verschwinden.

Wenn die Gewichte vorher bestimmt werden, ist der Berechnungsaufwand für SV3  $\mathcal{O}(n \log(n))$  und der Speicherbedarf  $\mathcal{O}(n)$ .

### 3.3.8 $P^2$ -Algorithmus

Der  $P^2$ -Algorithmus ist kein Schätzer, sondern eine heuristische Berechnungsmethode zur Approximation von Quantilen. Er wurde von Jain 1985 veröffentlicht ([JC85]). Da er als Bestandteil von Software-Standardbibliotheken weit verbreitet ist, wird er mit in den Vergleich aufgenommen. Weitere ähnliche Schätzer finden sich in [HL84] und [CK08].

Dieser Algorithmus hat den Vorteil, dass auf das Ordnen der Stichprobenwerte verzichtet werden kann. Daher beträgt der Berechnungsaufwand lediglich  $\mathcal{O}(n)$ . Der Speicherbedarf ist auch bei sehr großen Stichproben gering ( $\mathcal{O}(n)$ ), was große Vorteile bei der praktischen Umsetzung bietet.

Zur Schätzung des  $q$ -Quantils bestimmt der Algorithmus fünf Markerpaare, die aus einem Quantilsmarker und einem Positionsmarker bestehen:

Der Quantilsmarker  $q_1$  speichert den kleinsten Stichprobenwert,  $q_2$  die Schätzung des  $q/2$ -Quantil,  $q_3$  die des  $q$ -Quantils,  $q_4$  die des  $(1+q)/2$ -Quantils und  $q_5$  den größten Stichprobenwert.

Die Positionsmarker  $n_i, i = 1, \dots, 5$  beinhalten die Positionen der Quantilsmarker innerhalb der aktuellen Stichprobe. Sie werden mit ihrer idealen Position verglichen. Für den ersten  $q_1$  ist dies offensichtlich  $n_1 = 1$  und für den fünften  $q_5$   $n_5 = n$ . Die übrigen ergeben sich aus der Definition des empirischen Schätzers. Im Artikel wird hierfür  $\hat{Y}_{q,n} = Y_{\lceil q \cdot (n-1) + 1 \rceil}$  benutzt. Die obere Gaussklammer wird aber für die fiktive ideale Position ignoriert.

Für die gesuchten Quantile gilt dann:

$$\begin{aligned} \hat{Y}_{\frac{q}{2},n} &= Y_{(\frac{q}{2} \cdot (n-1) + 1)} \\ \hat{Y}_{q,n} &= Y_{(q \cdot (n-1) + 1)} \\ \hat{Y}_{\frac{1+q}{2},n} &= Y_{(\frac{1+q}{2} \cdot (n-1) + 1)} \end{aligned}$$

Die ideale Position ist der Index dieses empirischen Schätzers:

Für  $q_2$  ist dies  $n_2 = (n-1) \cdot q/2 + 1$ , für  $q_3$   $n_3 = (n-1) \cdot q + 1$  und für  $q_4$   $n_4 = (n-1) \cdot 1 + q/2 + 1$ .

Der Algorithmus belegt die Quantilsmarker mit den ersten fünf Werten der Stichprobe vor. Jedes nachfolgende Stichprobenelement verändert die Positionsmarke  $n_i, i = 1, \dots, 5$  nach folgender Regel:



Für alle Quantilsmarker, die größer als das aktuelle Stichprobenelement sind, wird der Positionsmarker um eins nach oben gesetzt. Dann wird überprüft, ob der Positionsmarker sich noch an der idealen Position befindet. Wenn er sich um mehr als 1 davon entfernt, wird dieser Positionsmarker und die dazugehörige Quantilschätzung korrigiert, in dem durch das ursprüngliche Markerpaar und seine beiden benachbarten Markerpaare eine Parabel gelegt wird. Dann wird der Positionsmarker auf seine ideale Position gesetzt und der dazugehörige Quantilsmarker mit dem Funktionswert der Parabel an dieser Stelle belegt.

Hieraus leitet sich auch der Name  $P^2$ -Algorithmus ab:  $P^2$  steht für „piecewise parabolic formula“ (stückweise parabelförmige Formel).

Für das erste und das fünfte Markerpaar sind keine Umsetzungen nötig, da sie immer an der exakten Position gehalten werden können.

Eine Aussage über die Genauigkeit des Schätzer kann nicht getroffen werden. **Bei Vorabtests zeigte sich, dass der  $P^2$ -Algorithmus bei geordneten Stichproben zu völlig falschen Ergebnissen führt.** Da er im folgenden auf die Funktionswerte nicht-linearer Funktionen zufällig gezogener Stichproben angewendet wird, wird angenommen, dass dieser Fall hier nicht auftritt.

### 3.3.9 Schätzer für extreme Quantile

In der Zuverlässigkeitstheorie werden Ausfallwahrscheinlichkeiten geschätzt. Dies ist die Wahrscheinlichkeit, dass eine Nebenbedingung der Optimierungsaufgabe verletzt wird. Die wichtigsten Methoden sind First-Order-Reliability-Method (FORM) und Second-Order-Reliability-Method (SORM) und Abwandlungen davon [Rac01]. Sie approximieren die Grenzlinie zwischen zulässigem und unzulässigem Bereich durch Polynome erste und zweiter Ordnung. In der Regel werden sehr kleine Quantile geschätzt. Ein Anwendung auf alle Quantile ist prinzipiell möglich. Da sie auf linearen oder quadratischen Approximationen beruhen, sind ihre Schätzungen aber nur für diese Funktionstypen brauchbar [CNNT11].

Die Extremwerttheorie hat ebenfalls spezielle Schätzer für extreme Quantile entwickelt. Diese sind semiparametrisch. Aus den extremen Werten einer Stichprobe werden die Parameter einer Extremwertverteilung geschätzt und hierüber ihre Quantile bestimmt. Sie sind vor allem für  $q < 1/n$  bzw.  $q > 1 - 1/n$  geeignet. Einen Überblick gibt [Emb97].

### 3.3.10 Diskussion

Bei der Vorstellung der einzelnen Schätzer wurden bereits theoretische Ergebnisse für den asymptotischen MSE aufgeführt. Diese werden nun zusammengefasst und verglichen. Für den empirischen Schätzer gilt (vgl. Gleichung 3.7):

$$\text{MSE}_{\text{EMP}} = \frac{1}{n} \text{AVar}_{\text{EMP}} + \mathcal{O}(n^{-2})$$

Für den Kerndichteschätzer ist selbst bei optimaler Wahl der Bandweite asymptotisch nur

$$\text{MSE}_{\text{KD}} = \frac{1}{n} \text{AVar}_{\text{EMP}} + \mathcal{O}(n^{-4/3})$$

möglich (vgl. Gleichung 3.13). Der Vorteil durch die Kerndichteschätzung verschwindet für  $n$  gegen unendlich, wenn auch nur mit der Rate  $\mathcal{O}(n^{-1/3})$  (Differenz zwischen  $\mathcal{O}(n^{-1})$  und  $\mathcal{O}(n^{-4/3})$ ). Asymptotisch wird der Kerndichteschätzer daher schlechter abschneiden. Für endliche Stichprobengrößen kann er die Schätzung allerdings trotzdem verbessern.

Die Differenz beträgt für den Gausskern (vgl. Gleichung 3.14):

$$\text{MSE}_{\text{KDGauss}} - \text{MSE}_{\text{EMP}} = \frac{1}{4}h^4 Q''(q)^2 - \frac{h}{n} Q'(q)^2 \frac{1}{2\sqrt{\pi}}$$

Dieser Ausdruck zeigt, dass die Kerndichteschätzung bei großer Steigung und kleiner zweiter Ableitung von  $Q$  den größten Vorteil gegenüber dem empirischen Schätzer bringen wird.

Für den Harrell-Davis-Schätzer und den Sfakianikis-Verginis-Schätzer gilt asymptotisch (vgl. Gleichungen 3.17 und 3.19):

$$\begin{aligned} \text{MSE}_{\text{HD}} &= \frac{1}{n} \text{AVar}_{\text{EMP}} + \mathcal{O}(n^{-3/2}) \\ \text{MSE}_{\text{SV3}} &= \frac{1}{n} \text{AVar}_{\text{EMP}} + \mathcal{O}(n^{-3/2}) \end{aligned}$$

Der Vorteil, den HD gegenüber EMP haben kann, wird für  $n$  gegen unendlich verschwinden. Da der Vorteil mit der Rate  $\mathcal{O}(n^{-1/2})$  verschwindet, wird er schneller ungeeignet sein als die Kerndichteschätzung bei optimaler Wahl der Bandweite. Später werden die hier aufgeführten Schätzer für Stichproben bis zu einer Größe von 100.000 benutzt. In einer numerischen Untersuchung ist zu klären, wann der theoretische Vorteil von EMP gegenüber den aufwändigeren Schätzern eintritt.

Die Schätzer unterscheiden sich in Berechnungsaufwand und Speicherbedarf:

Bei EMP, HD und SV3 muss die Stichprobe zuerst der Größe nach geordnet werden. Dies erfordert  $\mathcal{O}(n \log(n))$  Berechnungsaufwand und  $\mathcal{O}(n)$  Speicherplatz. Bei EMP wird dann das  $\lceil q \cdot n \rceil$ -te Element bestimmt, für HD und SV3 ist eine gewichtete Summe zu bestimmen ( $\mathcal{O}(n)$ ). Dabei wird vorausgesetzt, dass die Gewichte einmalig vorher bestimmt wurden. Bei der Kerndichteschätzung wird zuerst  $\hat{F}_n(t)$  mit  $\mathcal{O}(n)$  Berechnungsaufwand bestimmt. Anschließend muss die Gleichung  $\hat{F}_n(t) = q$  nach  $q$  aufgelöst werden. Dies geschieht numerisch, z.B. mit dem Newton-Raphson-Verfahren. Der Berechnungsaufwand hängt von der individuellen Problemstellung ab. In den numerischen Tests musste eine geeignete Startwertsuche vorgeschaltet werden, so dass der Aufwand immer höher als bei den anderen Schätzern war. Für P2 müssen nur wenige Marker gespeichert ( $\mathcal{O}(1)$  Speicher) und die Stichprobe einmal durchlaufen werden ( $\mathcal{O}(n)$  Berechnungsaufwand). In Tabelle 3.1 sind die Aufwände noch einmal zusammengefasst.

In der Literatur finden sich mehrere numerische Vergleichsuntersuchungen. Sheather und Marron [SM90] vergleichen verschiedene  $L$ -Schätzer mit Kerndichteschätzern. Dabei schneidet der Harrell-Davis-Schätzer im Durchschnitt am besten ab, dicht gefolgt von den Kerndichteschätzern. Dielman und Pfaffenberger [DLP94] untersuchen HD und EMP. Sie empfehlen HD außer für extreme Quantile mit  $q < 0.02$  bzw.  $q > 0.98$ . In der neueren Untersuchung von Sfakianikis und Verginis [SV08]

Tab. 3.1: Vergleich der Quantilschätzer

Schätzer	Berechnungsaufwand	Speicherbedarf
EMP	$\mathcal{O}(n \log(n))$	$\mathcal{O}(n)$
HD	$\mathcal{O}(n \log(n))$	$\mathcal{O}(n)$
SV3	$\mathcal{O}(n \log(n))$	$\mathcal{O}(n)$
KDE	$\mathcal{O}(n)$ + Aufwand für Newton-Verfahren	$\mathcal{O}(n)$
P2	$\mathcal{O}(n)$	$\mathcal{O}(1)$

wird HD vom dem dort vorgeschlagenen neuen Sfakianikis-Verginis-Schätzer übertroffen. Deshalb wird dieser mit untersucht. Die genannten Artikel benutzen eine Stichprobengröße von höchstens 750.

Inui et al. [IKK05] vergleichen HD und EMP für die im Vergleich zur Normalverteilung endlastigeren  $t$ -Verteilungen. Dabei liefert HD für Stichproben größer als  $n = 1000$  bessere Ergebnisse, allerdings nur um weniger als 10 Prozent. Eine Untersuchung mit Verteilungen, die keiner Klasse angehören, gibt es nicht.

## 3.4 Weitere Methoden zur Schätzung von Quantilen

Neben den nichtparametrischen Quantilschätzern gibt es weitere Methoden zur Schätzung von Quantilen, die in diesem Unterkapitel vorgestellt werden.

### 3.4.1 Taylor-Entwicklung

Zur Berechnung von Erwartungswert und Standardabweichung kann die Taylor-Entwicklung herangezogen werden. Für Quantile ist sie nur eingeschränkt anwendbar:

Für monotonen  $f$  gilt:

$$y_q(f(x_0)) = f(x_q)$$

Zunächst wird die Taylorreihe erster Ordnung

$$f(x) \approx T_1(x) = f(x_0) + f'(x_0)(x - x_0)$$

betrachtet. Für den Median und  $x \sim \mathcal{N}(x_0, \sigma)$  gilt dann

$$y_{0.5}(f(x)) \approx f(x_0)$$

Dies ist die gleiche Näherung wie für den Erwartungswert. Aus der Taylorreihe zweiter Ordnung kann hergeleitet werden:

$$y_{0.5}(f(x)) \approx f(x_0) + \frac{1}{2}f''(x_0)y_{0.5}(X)^2$$

Im allgemeinen kann  $y_{0.5}(X)^2$  nicht analytisch bestimmt werden. Für  $X \sim \mathcal{N}(0, \sigma)$  ist  $X^2$  gamma-verteilt mit den Parametern  $2\sigma^2$  und  $1/2$ . Für den Median dieser Verteilung gibt es keine einfache geschlossene Form. Eine brauchbare Approximationsformel für den Median der übertragenen Verteilung lässt sich mit diesem Ansatz nicht herleiten.

Ein Annäherung für den Median über die Taylorreihen-Entwicklung kann auch auf einem anderen Weg gewonnen werden: Anstatt den Median der Verteilung von  $f(X)$  zu berechnen, wird der Median der durch die Taylorreihe zweiter Ordnung übertragenen Verteilung bestimmt.  $T_2(x)$  ist eine Parabel, für die die übertragene Verteilung analytisch berechnet werden kann.

Angenommen  $X \sim N(x_0, \sigma)$ . Sei  $r$  die Lösung von:

$$\operatorname{erf}\left(\frac{-f'(x_0) + x}{\sqrt{2}f''(x_0)\sigma}\right) + \operatorname{erf}\left(\frac{f'(x_0) + x}{\sqrt{2}f''(x_0)\sigma} - 1\right) = 0$$

Für den Median erhält man

$$y_{0.5}(x_0) \approx \frac{1}{2f''(x_0)}(-f'(x_0) + 2f(x_0)f''(x_0) + r^2)$$

und für die Quantile

$$y_q(x_0) \approx f(x_0) + f'(x_0)(x_q(X) - x_0)$$

Diese Näherung gilt nur für annähernd quadratische Funktionen  $f$ . Im allgemeinen ist es nicht günstig, Quantile über die Taylorreihen-Entwicklung zu berechnen, da die Annäherung im gesamten Schwankungsbereich um eine Stelle  $x_0$  zu schlecht ist.

### 3.4.2 Konfidenzintervalle für den Standardschätzer

Für den Standardschätzer können Konfidenzgrenzen zum Konfidenzniveau  $1 - \alpha$  für das  $q$ -Quantil einer stetigen Zufallsvariablen  $Y$  angegeben werden (z.B. [GC03]). Mit der Ordnungsstatistik der Stichprobe kann das Konfidenzintervall  $[y_{(m_L)}, y_{(m_U)}]$  bestimmt werden über

$$m_L \text{ erfüllt } \sum_{i=0}^{m_L-1} \binom{n}{i} q^i (1-q)^{n-i} = \frac{\alpha}{2}$$

$$\text{und } m_U \text{ erfüllt } \sum_{i=m_U}^n \binom{n}{i} q^i (1-q)^{n-i} = 1 - \frac{\alpha}{2}.$$

Für  $n \cdot q \cdot (1-q) > 9$  kann die Binomialverteilung durch die Normalverteilung approximiert werden. Dies führt zu

$$m_L \approx n \cdot q + 0,5 - z \cdot \sqrt{n \cdot q \cdot (1-q)} \quad \text{und}$$

$$m_U \approx n \cdot q + 0,5 + z \cdot \sqrt{n \cdot q \cdot (1-q)}$$

mit  $\Phi(z) = 1 - \alpha/2$  und  $\Phi$  ist die Verteilungsfunktion der Standardnormalverteilung.

Generell lassen sich Konfidenzintervalle für beliebige Schätzer über Resampling gewinnen (speziell für Quantile z.B. in [BS01]). Da diese Verfahren auf der wiederholten Ziehung von Unterstichproben aus einer Ausgangsstichprobe beruhen, sind sie gerade bei großen Stichproben sehr aufwändig.

### 3.5 Metamodelle

In vielen Anwendungen ist die Funktionsgleichung unbekannt, kann aber für einzelne Punkte exakt bestimmt werden. So ist etwa das Verhalten eines Bauteils nicht als analytische Gleichung bekannt, es kann aber über Simulationen für feste Parameterkombinationen berechnet werden. Diese Situation soll auch hier zugrunde gelegt werden. Man spricht vom *Black-Box-Ansatz*, da keine weitere Annahmen über die spezifische Struktur der Funktionsgleichung vorliegen. Wie bereits in der Einleitung erwähnt wird in dieser Arbeit angenommen, dass es aufwändig ist, exakte Punkte zu bestimmen, wie etwa bei langlaufenden Simulationen oder bei teuren Experimenten. Um dennoch robuste Optimierung betreiben zu können, bietet sich als Ausweg die Metamodellierung an:

*Metamodelle* oder *Ersatzmodelle* bestimmen zu gegebenen Stützstellen  $(x_i, y_i) \quad i = 1, \dots, n$  eine Interpolationsfunktion, die diese Punkte durchläuft. Für beliebige Parameterkombination der Eingabevariablen können dann durch Auswertung des Metamodells Approximationen für die exakten Werte bestimmt werden. Da es als analytische Funktion vorliegt, sind Berechnungen im Vergleich zu Simulationen sehr schnell und billig.

Aus der Fülle von möglichen Metamodellen (siehe [FHT01], [FSK08]) wird in dieser Dissertation nur das mit radialen Basisfunktionen und polynomielltem Anteil benutzt. Alle anderen lassen sich ebenso einsetzen. Weit verbreitet ist z.B. das Kriging-Verfahren [SWMW89] oder neuronale Netzwerke [Hay94]. Wenn die Funktionswerte fehlerbehaftet sind, werden auch Regressionsfunktionen [MPV12] oder Mowing-Least-Squares-Modelle [LS81] benutzt.

#### 3.5.1 Radiale-Basisfunktionen-Modell mit polynomielltem Anteil

Ein *Metamodell mit radialen Basisfunktionen* (RBF) hat die Form

$$\tilde{f}(x) = \sum_{i=1}^n c_i \Phi(|x - x_i|)$$

wobei  $\Phi()$  eine *radiale Basisfunktion* ist, d.h. ihre Werte hängen nur vom Euklidischen Abstand  $|\cdot|$  von  $x$  zu  $x_i$  ab.

Zur Bestimmung der  $c_i$  genügt es, das lineare Gleichungssystem

$$f_j = \sum_{i=1}^n c_i \Phi(|x_j - x_i|) \quad j = 1, \dots, n$$

mit  $f_j = f(x_j)$  zu lösen. Eine direkte Lösung  $c = \Phi^{-1}y$  erhält man durch Inversion der Gram-Matrix  $\Phi_{ij} = \Phi(|x_j - x_i|), i, j = 1, \dots, n$ . Eine geeignete Basisfunktion ist die *multiquadratische* Funktion

$$\Phi(r) = (b^2 + r^2)^{1/2},$$

wobei  $b$  ein Glättungsfaktor für das Verhalten der Funktion nahe  $x$  ist [Har71]. Für sie existiert immer eine Lösung. Allerdings kann das lineare Gleichungssystem schlecht konditioniert sein [Buh03].

Die Konstruktion eines RBF-Metamodells erfordert  $\mathcal{O}(n^2)$  Speicherplatz und  $\mathcal{O}(n^3)$  Operationen. Für multiquadratische Basisfunktionen gibt es Verfahren, die diesen Aufwand deutlich verringern [CBN02]. Zur Verbesserung der Qualität bei linearem oder polynomiellm Verhalten kann ein polynomieller Anteil  $p()$  hinzugefügt werden.

$$\tilde{f}(x) = \sum_{i=1}^n c_i \Phi(|x - x_i|) + p(x)$$

Das Gleichungssystem zur Bestimmung der  $c_i$  muss dann entsprechend erweitert werden (Einzelheiten hierzu in [KN05] oder [Buh03]).

### 3.5.2 Validierung eines Metamodells

Der Fehler eines Metamodells kann nicht berechnet werden, da die exakte Funktion unbekannt ist. Er kann jede Größe annehmen, da die wahre Funktion zwischen zwei Stützstellen beliebig hoch oszillieren kann, während das Metamodell diesen Bereich glättet.

Eine naheliegende Möglichkeit für die Validierung des Metamodells besteht darin, zusätzliche exakte Punkte zu bestimmen und z.B. die Gesamtsumme der quadratischen Abweichungen dieser „Trainingspunkte“ als Maß für die Genauigkeit des Metamodells zu definieren. Diese Vorgehensweise hat einige Nachteile:

1. Die Bestimmung exakter Punkte ist aufwändig. Ziel dieser Dissertation ist es gerade, eine Methodik zu entwickeln, die mit möglichst wenigen exakten Punkten auskommt.
2. Würde man die Trainingspunkte mit in die Erstellung des Metamodells einbeziehen, würde es noch einmal verbessert.
3. Die Fehlerbestimmung erfolgt nur lokal an den Trainingspunkten. In anderen Bereichen des Parameterraums kann der Fehler trotzdem sehr groß sein.

Eine andere Möglichkeit, die Genauigkeit des Metamodells näherungsweise zu bestimmen, bietet das *leave-one-out Kreuzvalidierungsverfahren*:

Wenn man eine Stützstelle eliminiert und aus den restlichen ein neues Metamodell erstellt, gibt die Differenz aus ursprünglichem und neuem Wert an, wie empfindlich das Metamodell dort auf das Weglassen dieser Stützstelle reagiert. Der Wert kann aber auch als Schätzung für den Fehler des Metamodells an dieser Stelle interpretiert werden.

Dieses Verfahren wird nacheinander für alle Stützstellen durchgeführt. Eine Fehlerabschätzung für einen beliebigen Punkt des Parameterraums erhält man durch Interpolation der Stützstellenabschätzungen. Hierzu kann man die Gewichte  $c_i$  des ursprünglichen Metamodells nutzen. Die so errechnete Größe wird auch - in Abgrenzung zum unbekanntem Fehler - *Modelltoleranz* genannt. Sie macht eine gute lokale Fehlerabschätzung im gesamten Designraum möglich. Allerdings gibt es auch Beispiele, die eine kleine Modelltoleranz bei großem tatsächlichen Fehler aufweisen und umgekehrt ([FHT01], S.247).

[NNC12b] geben ein Verfahren an, das für RBF-Modelle mit polynomielltem Anteil eine direkte Berechnung der Modelltoleranz ohne  $n - 1$ -maliges Erstellen eines neuen Metamodells erlaubt.

Das Kreuzvalidierungsverfahren kann verallgemeinert werden, in dem man Teilmengen von Stützstellen eliminiert [FK09].

Falls eine große Modelltoleranz auftritt, sollte das Modell verbessert werden. Dies kann durch das Hinzufügen von Stützstellen geschehen. Um mit möglichst wenigen Punkten auszukommen, wurden verschiedene Ergänzungskriterien entwickelt. Hier sei nur das *expected-improvement*-Kriterium genannt ([KN05], für mehrere Zielkriterien [FK09]). Es berechnet die erwartete Verbesserung für jeden Punkt des Designraums und wählt den Punkt der maximalen Verbesserung aus. Dabei kann zwischen lokaler und globaler Suche skaliert werden. RBF-Modelle können bei einer ungleichmäßigen Verteilung der Stützstellen in Schwingung geraten [Mic86]. Bei einem iterativen Verbessern des Metamodells sollten daher eng beieinanderliegende Stützstellen vermieden werden.

Ein vielversprechender aktueller Ansatz zur Verbesserung der Qualität eines Metamodells ist die *multi-surrogate approximation (MSA) metamodeling method*. Hierbei wird kein einzelnes globales Modell erstellt, sondern viele lokale. Der Designraum wird zunächst geclustert in Bereiche mit ähnlichen Eigenschaften. Für diese wird dann jeweils ein lokales Metamodell erstellt und an den Rändern mit den anderen verbunden [ZX11]. Einen Überblick über die aktuellen Entwicklungen bezüglich Metamodellen gibt [WS07].

### 3.6 Schnelle Multipolverfahren für RBF-Metamodelle

Die Auswertung eines Metamodells ist im Vergleich zu einer Simulation extrem schnell und billig. Zur Bestimmung der übertragenen Verteilung in einem Punkt sind aber mehrere zehntausend Auswertungen nötig, um eine genügend große Genauigkeit zu erzielen. In der im Kapitel 6 beschriebenen Optimierungsprozedur wird eine solche Stichprobe nicht nur für einen Punkt, sondern für alle Stützstellen des Originalmetamodells gezogen. Hier besteht Bedarf zur Beschleunigung der Auswertung von Radialen Basisfunktionen Modellen. Dies ist nur über die Berechnung von Näherungslösungen möglich. Da aber  $f$  schon ein Modell der Realität ist, also schon einen gewissen Fehler besitzt, erscheint es vertretbar, bei der Auswertung noch einmal einen kleinen Fehler  $\epsilon > 0$  zuzulassen, wenn die Auswertung dadurch erheblich beschleunigt werden kann.

In diesem Fall können schnelle Multipolverfahren angewendet werden. Sie wurden ursprünglich zur Lösung von  $n$ -Körper-Systemen entwickelt, speziell zur Berechnung elektrostatischer Potentiale. Drei Ideen liegen ihnen zugrunde:

- Punkte, die nahe beieinander liegen, werden zu einem einzigen Punkt, dem „Zentrum“, zusammengefasst und ihre Eigenschaften in Bezug auf weit entfernt liegende, sogenannte wohlgetrennte, Punkte durch dieses Zentrum repräsentiert.
- Die Zuordnung zu Zentren und die Bestimmung, wann Zentrum und weit entfernte Punkte wohlgetrennt sind, erfolgt über die Erstellung einer Baumstruktur.

- Bei Wohlgetrenntheit werden die Zentren in Reihen entwickelt, man spricht von der Fernfeldentwicklung. Sie wurde zuerst in [GR87] eingeführt.

Diese Prinzipien lassen sich auf radiale Basisfunktionen-Modelle übertragen, da auch bei ihnen der Funktionswert eines Punktes nur von den Abständen zu den Stützstellen abhängig ist. In diesem Kapitel wird ein Ansatz beschrieben, der speziell für RBF-Modelle mit multiquadratischen Basisfunktionen entwickelt wurde [DD12]. In [BK13] wurde er untersucht.

Als Ausgangssituation wird davon ausgegangen, dass ein RBF-Modell für die Menge der Stützstellen  $\{x_i, i = 1, \dots, n\}$  mit  $x_i = (x_{i1}, \dots, x_{id})^T$  erstellt wird. Dann hat das RBF-Modell die Form

$$\tilde{f}(x) = \sum_{i=1}^n c_i \Phi(|x - x_i|)$$

wobei die  $c_i, i = 1, \dots, n$  Gewichte sind und  $\Phi(|x - x_i|)$  eine radiale Basisfunktion und  $|\cdot|$  der Euklidische Abstand ist.

Die Auswertung dieses Metamodells für einen Punkt erfordert offensichtlich einen Aufwand von  $\mathcal{O}(n)$ . Soll die Menge  $\{y_j, j = 1, \dots, m\}$  von  $m$  Auswertungspunkten bestimmt werden, beträgt der Aufwand  $\mathcal{O}(n \cdot m)$ .

Der folgende Treecode-Algorithmus reduziert den Aufwand auf  $\mathcal{O}((n + m) \cdot \log(m))$ . Auswertungspunkte, die nahe beieinander liegen, werden zu Clustern zusammengefasst. Liegen Stützstellen von diesem Cluster genügend weit entfernt, wird - vereinfacht ausgedrückt - der Beitrag dieser Stützstellen für das Cluster zusammengefasst. Dieser Vorgang kann in einer Vorab-Phase durchgeführt werden. Die dabei berechneten Werte werden Momente oder Potentiale genannt. In der Auswertungsphase werden dann die nahe zum Cluster liegenden Stützstellen direkt ausgewertet und die genügend weit entfernten über die berechneten Momente.

Dieser zur Klasse der Schnellen Treecodes gehörende Algorithmus wird verwendet, wenn es viele Auswertungspunkte  $y_j, j = 1, \dots, m$  im Verhältnis zu den Stützstellen  $x_i, i = 1, \dots, n$  gibt. Dann ist es günstiger, nach den Auswertungspunkten zu clustern [Fas07].

### 3.6.1 Erstellung der Baumstruktur

Zunächst wird ein Baum für die  $m$  Auswertungspunkte erstellt. Das Wurzelcluster ist ein rechteckiger Quader, dessen Seiten parallel zu den Koordinatenachsen verlaufen. Er ist der kleinstmögliche Quader, der alle Auswertungspunkte enthält. Nun wird jede Achse genau mittig unterteilt. Auf diese Weise entstehen  $2^d$  Kindercluster der Wurzel. Sie bilden die nächste Stufe des Baumes. Enthält ein Cluster mehr als eine vorher festgelegte Anzahl  $m_0$  von Auswertungspunkten, wird es wieder unterteilt. Dieser Prozess wird solange fortgesetzt, bis alle Cluster höchstens  $m_0$  Punkte enthalten. Das letzte Cluster der Baumstruktur wird Blatt genannt.

Die Baumtiefe  $bt$  ist bei gleichverteilten Auswertungspunkten am kleinsten. Sie beträgt dann

$$bt = \lceil \frac{1}{d} \log_2 \left( \frac{m}{m_0} \right) \rceil$$



und ist proportional zum kleinsten Abstand zweier Auswertungspunkte. Im ungünstigsten Fall kann der Baum beliebig tief werden, es gibt aber Korrekturmaßnahmen, die die Baumtiefe auf  $m$  begrenzen. Der zugehörige Baum besteht dann nur aus einem Ast [APG94].

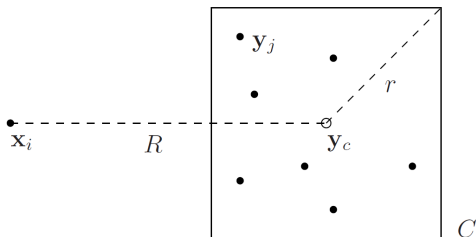


Abb. 3.23: Die Auswertungspunkte  $y_j$  werden unterteilt in Cluster  $C$  mit Zentrum  $y_C$  und Radius  $r$ .  $R$  ist der Partikel-Cluster-Abstand von Partikel  $x_i$  zum Clusterzentrum  $y_C$ . (Grafik entnommen aus [BK13])

Es werden folgende Bezeichnungen eingeführt: Ein Cluster  $C$  besitzt das Zentrum  $y_C$  und den Radius  $r$ . Der Abstand des Clusters um  $y_C$  zur Stützstelle  $x_i$  sei  $R = |y_C - x_i|$  (vgl. Abbildung 3.23).

### 3.6.2 Multiple Acceptance Criterion

Das Multiple Acceptance Criterion (MAC) legt fest, wann eine Stützstelle und ein Cluster weit genug voneinander entfernt sind, damit sie wohlgetrennt sind und die Beiträge der Stützstellen zusammengefasst werden dürfen. Dieser Vorgang wird Nahfeldentwicklung genannt. Die Wahl des MAC hat entscheidenden Einfluss auf die Effizienz des Algorithmus. Deng et al. [DD12] verwenden als MAC für die multiquadratische Basisfunktion  $\Phi(x) = (b^2 + x^2)^{1/2}$  das schon von Krasny [KW11] festgelegte Kriterium:

$$\theta = \frac{r}{\sqrt{R^2 + b^2}} \leq \theta_0$$

Boateng [BK13] verwendet  $r/R \leq \theta_0$ .  $\theta_0$  ist ein benutzerdefinierter Parameter. Gebräuchlich sind  $\theta_0 = 0,5$  bzw.  $0,75$ .

### 3.6.3 Approximation der RBF-Summe

In der RBF-Summe für einen Auswertungspunkt  $y_j$  wird nun unterschieden zwischen Stützstellen, die nahe an  $y_j$  liegen und solchen, die wohlgetrennt sind.

Zur mathematischen Formulierung werden noch folgende Bezeichnungen benötigt: Die Baumtiefe betrage  $bt$  und  $C_1$  bezeichne die Wurzel und  $C_{bt}$  ein Blatt. Dann gehört ein Auswertungspunkt zu einer Reihe von Clustern auf verschiedenen Stufen des Baumes:

$$y_j \in C_{bt} \subseteq \dots \subseteq C_1$$

Das Clusterzentrum zu  $C_i$  werde mit  $y_C^i$  bezeichnet.

Die RBF-Summe lässt sich aufteilen in:

$$\begin{aligned}
 \tilde{f}(y_j) &= \sum_{i=1}^n c_i \Phi(|y_j - x_i|) \\
 &= \sum_{x_i \in D} c_i \Phi(|y_j - x_i|) + \sum_{x_i \notin D} c_i \Phi(|y_j - x_i|) \\
 &= \sum_{x_i \in D} c_i \Phi(|y_j - x_i|) + \sum_{l=1}^{bt} \sum_{x_i \in I_l} c_i \Phi(|y_j - x_i|)
 \end{aligned}$$

Dabei sei  $I_l$  die Menge der Stützstellen  $x_i$ , die von  $C_l$  wohlgetrennt sind, aber nicht von  $C_{l-1}, \dots, C_1$  und  $D$  die Menge der Stützstellen, die nicht wohlgetrennt sind von  $C_{bt}, \dots, C_1$ .

Für die Summanden der zweiten Summe wird eine Taylor-Entwicklung in  $y = y_C^l$  durchgeführt:

$$\begin{aligned}
 \sum_{x_i \in I_l} c_i \Phi(|y_j - x_i|) &\approx \sum_{x_i \in I_l} c_i \sum_{\|k\|=0}^p \frac{1}{k!} D_y^k \Phi(|y_C^l - x_i|) (y_j - y_C^l)^k \\
 &= \sum_{\|k\|=0}^p \sum_{x_i \in I_l} c_i \frac{1}{k!} D_y^k \Phi(|y_C^l - x_i|) (y_j - y_C^l)^k \\
 &= \sum_{\|k\|=0}^p m_k(y_C^l) (y_j - y_C^l)^k \tag{3.20}
 \end{aligned}$$

wobei

$$m_k(y_C^l) = \sum_{x_i \in I_l} c_i a_k(y_C^l, x_i)$$

Momente genannt werden. Die

$$a_k(y_C^l, x_i) = \frac{1}{k!} D_y^k \Phi(|y_C^l - x_i|)$$

sind die Taylor-Koeffizienten. Sie können über eine weiter unten angegebene Rekursionsformel bestimmt werden.

Dabei wird die kartesische Multiindex-Notation benutzt:

$$k = (k_1, \dots, k_d), \quad \|k\| = k_1 + \dots + k_d, \quad k! = k_1! \cdot \dots \cdot k_d!$$

$$y^k = y_1^{k_1} \cdot \dots \cdot y_d^{k_d}, \quad D_y^k = D_{y_1}^{k_1} \cdot \dots \cdot D_{y_d}^{k_d}$$

Gleichung 3.20 heißt Nahfeldentwicklung für die Wechselwirkung zwischen dem Auswertungspunkt  $y_j$  und der Stützstellen  $x_i \in I_l$ .

Die Momente hängen nicht mehr von den Auswertungspunkten  $y_j$  ab, sondern nur noch von den Clusterzentren  $y_C$  und den Stützstellen  $x_i$ .

### 3.6.4 Rekursionsformel für Taylor-Koeffizienten

Zur Berechnung der Taylor-Koeffizienten kann die folgende Rekursionsformel benutzt werden [BK13]. Mit  $a_k(y, x) = a_{k_1, \dots, k_d}(y, x)$  gilt:

$$\begin{aligned} \|k\|(|y-x|^2 + b^2)a_k(y, x) = \\ (2\|k\| - 3) \sum_{i=1}^d (y_i - x_i)a_{k-e_i}(y, x) + (\|k\| - 3) \sum_{i=1}^d a_{k-2e_i}(y, x) \end{aligned} \quad (3.21)$$

mit  $a_0 = \Phi(y_C, x_i)$ ,  $a_k = 0$  falls jedes  $k_i < 0$ ,  $\|k\| = k_1 + \dots + k_d$ .  $e_i$  sind die Einheitsvektoren.

### 3.6.5 Aufwand des Algorithmus

In der Vorabphase kann die Baumstruktur bzgl. der  $m$  Auswertungspunkte bei Gleichverteilung mit  $\mathcal{O}(m \log(m))$  bestimmt werden. Zur Berechnung der Momente  $m_k(y_C)$  muss der Baum im Mittel bis zur durchschnittlichen Baumtiefe  $\log(m)$  durchlaufen werden und zwar  $n$  Mal. Der Aufwand beträgt damit  $\mathcal{O}(n \log(m))$ . In der Berechnungsphase werden die Einflüsse der Momente akkumuliert. Dazu wird im Baum für jeden Auswertungspunkt bis zum zugehörigen Blatt hinabgestiegen. Auf dem Weg dorthin werden die entsprechenden Momente summiert. Da der Baum  $m$  Mal bis zur durchschnittlichen Tiefe durchlaufen wird, beträgt der Aufwand für diese Phase  $\mathcal{O}(m \log(m))$ . Zusammen ergibt dies einen Gesamtaufwand von  $\mathcal{O}((m+n) \log(m))$ .

Bisher existiert in der Literatur nur diese grobe Schätzung des Aufwandes im Falle einer Gleichverteilung der  $m$  Auswertungspunkte. Eine genauere Analyse und ein Vergleich mit der direkten Summation wird im Kapitel 7 durchgeführt. Dort wird ebenfalls die Anwendbarkeit auf die entwickelte Methodik diskutiert.

## 3.7 Multikriterielle Optimierung

Der Begriff „Optimierung“ leitet sich vom lateinischen Wort „optimus“ - „der beste“ ab. Optimierung meint das Bestreben nach Verbesserung und im engeren Sinne die Suche nach der besten Lösung.

Mathematisch wird unter der (deterministischen) Optimierung die Suche nach dem Minimum einer Funktion verstanden. In dieser Dissertation werden folgende Bezeichnungen verwendet:

Die  $d$  Designvariablen  $x = (x_1, \dots, x_d)^T$  werden von der Funktion  $f = (f_1, \dots, f_l)^T$  auf  $l$  Zielvariablen  $y = (y_1, \dots, y_l)^T$  übertragen.

Die deterministische Standardformulierung einer Optimierungsaufgabe lautet dann:

$$\text{Minimiere} \quad y = f(x), \quad (\text{a})$$

$$\text{so dass} \quad g_i(x) \leq 0, \quad i = 1, \dots, I \quad (\text{b})$$

$$\text{und} \quad h_j(x) = 0, \quad j = 1, \dots, J \quad (\text{c})$$

$$\text{und} \quad x_h^L \leq x_h \leq x_h^U, \quad h = 1, \dots, k \quad (\text{d})$$

erfüllt sind.

Die Ungleichheitsnebenbedingungen (b) und die Gleichheitsnebenbedingungen (c) bilden zusammen die *Neben- oder Randbedingungen*. Da die Gleichheitsbedingungen vor der Optimierung eliminiert werden können, werden sie im folgenden nicht mehr berücksichtigt. (d) beschreibt den *Designraum*

$$D = \{x \in \mathbb{R}^n \mid x_h^L \leq x_h \leq x_h^U\}$$

Falls mehrere Zielvariablen  $(y_1, \dots, y_l)^T = (f_1(x), \dots, f_l(x))^T$  vorhanden sind, d.h.  $l > 1$ , spricht man von einer *multikriteriellen* Optimierungsaufgabe. Die Zielvariablen werden dann auch *Kriterien* genannt.

Für eine multikriterielle Optimierungsaufgabe gibt es im allgemeinen keine Lösung, die alle Zielvariablen gleichzeitig minimiert, sondern die Kriterien stehen in Konflikt zueinander. Es gibt z.B. Lösungen, die das eine Kriterium minimieren, aber die anderen Kriterien nicht. Mathematisch entsteht ein Lösungsraum. Um schließlich zu einer einzelnen Lösung zu gelangen, muss der Entscheider auswählen.

Multikriterielle Aufgaben treten in Anwendungen häufig auf. Ein klassischer Zielkonflikt bei der Auslegung von Bauteilen ist beispielsweise die gleichzeitige Minimierung von Masse, Spannung und Kosten.

Später wird gezeigt, dass auch die robuste Optimierung multikriteriell ist.

Weitere Einzelheiten der verschiedenen Ansätze und ihrer Beziehungen untereinander sind in [Ehr99, Mie99] zu finden. Eine gute Einführung in die multikriterielle Optimierung und ihre Algorithmen gibt [Deb05].

### 3.7.1 Lösung einer multikriteriellen Optimierungsaufgabe

Zur Lösung einer multikriteriellen Optimierungsaufgabe gibt es verschiedene Lösungsstrategien, die in den folgenden Unterkapiteln vorgestellt werden. Sie lassen sich klassifizieren nach dem Zeitpunkt, zu dem die Präferenzen des Entscheiders einfließen [HMPY79]:

Bei *a priori*-Strategien werden sie vor der Optimierung festgelegt. Diese werden zuerst vorgestellt. *A posteriori*-Strategien verlangen Entscheidungen erst nach der Optimierung. Der Lösungsraum wird vollständig bestimmt, erst danach wählt der Entscheider aus (folgendes Unterkapitel). In *progressiven* Verfahren geschieht dies während der Optimierung. Der Entscheider legt schrittweise seine Prioritäten fest, hierdurch wird die optimale Menge schrittweise eingeschränkt (letztes Unterkapitel). Der Rechenaufwand kann gegenüber den *a posteriori*-Verfahren reduziert werden.

#### A priori Lösungsstrategien

Eine Überführung in eine einkriterielle Aufgabe ist möglich, wenn man den einzelnen Zielvariablen eine Gewichtung zuordnet und die gewichtete Summe der Zielvariablen minimiert:

$$\tilde{y}(x) = \sum_{i=1}^l w_i f_i(x)$$

Dieser Zugang heißt *präferenz-basierte multikriterielle Optimierung*. Zur Lösung können Standardalgorithmen wie z.B. das Gradientenverfahren benutzt werden. Ein

Nachteil dieser Vorgehensweise liegt darin, dass die Verteilung der Gewichte subjektiv erfolgt und ihre Änderung zu anderen optimalen Lösungen führt.

Eine weitere a priori Lösungsstrategie ist die *Methode der  $\epsilon$ -Begrenzung*. Dabei werden  $l-1$ -Zielfunktionen in Nebenbedingungen umgewandelt, indem man ihnen obere Schranken  $\epsilon_i \quad i = 1, \dots, l-1$  zuweist. Die verbliebene Zielfunktion wird mit deterministischen Optimierungsverfahren minimiert. Die Bestimmung geeigneter Schranken ist schwierig. Bei unpassender Wahl kann es vorkommen, dass keine Lösung gefunden wird.

**A posteriori Lösungsstrategien**

Bei a posteriori Lösungsstrategien wird zunächst der ganze Lösungsraum bestimmt. Dies ist die Menge „bester Kompromisse“ der sich einander widersprechenden Ziele. Dabei ist ein bester Kompromiss ein Punkt, für den eine Verbesserung in einem Kriterium nur durch eine Verschlechterung in einem anderen erkauft werden kann. Zur mathematischen Definition der Menge der besten Kompromisse sind einige Begriffe nötig:

Eine Lösung  $x^{(1)}$  *dominiert* eine andere Lösung  $x^{(2)}$ , wenn die beiden folgenden Bedingungen erfüllt sind:

- Die Lösung  $x^{(1)}$  ist in allen Kriterien nicht schlechter als  $x^{(2)}$ :

$$f_j(x^{(1)}) \leq f_j(x^{(2)}) \text{ für alle } j = 1, \dots, l$$

- Die Lösung  $x^{(1)}$  ist in mindestens einem Kriterium besser als  $x^{(2)}$ :

$$f_j(x^{(1)}) < f_j(x^{(2)}) \text{ für mindestens ein } j \in \{1, \dots, l\}$$

Die so definierte Relation ist transitiv und nicht reflexiv. Damit ist sie eine strikte partielle Ordnung.

Aus einer Menge von Lösungen  $P$  heißt die Menge  $P'$  *nicht-dominierte Menge*, wenn ihre Elemente von keiner Lösung aus  $P$  dominiert werden. Wenn  $P$  den ganzen Designraum umfasst, d.h.  $P = D$ , wird  $P'$  *Pareto-optimale Menge* oder *Paretofront* genannt. Die Paretofront bildet im Zielraum eine Hyperfläche der Dimension  $l - 1$ . Sie ist die gesuchte Menge der besten Kompromisse.

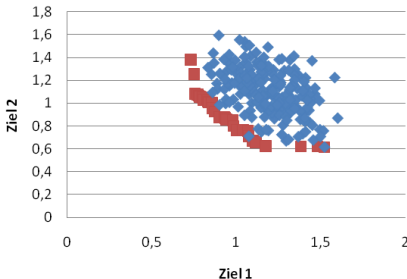


Abb. 3.24: typische Paretofront (rot) für zwei Zielvariablen

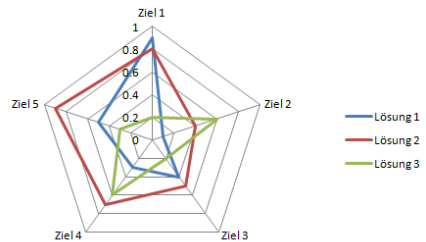


Abb. 3.25: Beispiel für ein Sterndiagramm

Abbildung 3.24 zeigt in rot eine typische Paretofront für zwei Zielvariablen. Diese Lösungen dominieren die blauen Punkte.

Die klassische Methode zu ihrer Bestimmung ist die *Methode der Gewichteten Summen*. Hierbei wird wie bei der präferenz-basierten Optimierung die gewichtete Summe aller Kriterien gebildet. Die optimale Lösung für eine feste Gewichtung ist dann ein Element der Pareto-optimalen Menge. Durch sukzessives Verändern der Gewichtung können daher Punkte der Paretofront bestimmt werden. Der Aufwand ist  $\mathcal{O}(l \cdot n^2)$ . Diese einfache Methode hat zwei große Nachteile [DD97]:

- Das gleichmäßige Verändern der Gewichte muss nicht zu einer gleichmäßigen Verteilung der Pareto-optimalen Punkte auf der Paretofront führen. Gerade bei nichtlinearen Funktionen kann mit dieser Methode nicht sichergestellt werden, dass alle Teile der Paretofront gefunden werden.
- Die Vorgehensweise bestimmt nur den konvexen Teil der Paretofront. Ein nicht-konvexer Teil wird ausgelassen.

Auch die Methode der  $\epsilon$ -Begrenzung lässt sich zur Bestimmung der Paretofront nutzen: Durch sukzessives Verändern der Schranken erhält man Pareto-optimale Punkte auch bei nicht-konvexen Zielräumen.

In dieser Dissertation wird der Algorithmus „ParetoSort“ [HNC10, Ben80] verwendet. Er bestimmt zu einer sehr großen gleichmäßig verteilten Stichprobe der Definitionsmenge alle Funktionswerte. Man erhält so eine Menge  $l$ -dimensionaler Vektoren, die mit einer Divide-and-Conquer-Strategie sortiert werden. Dabei hat sich die rekursive Version als effizient erwiesen, da die Rekursionstiefe gering.

### Progressive Lösungsstrategien

Bei progressiven Verfahren wird der Entscheider an bestimmten Stellen des Optimierungsprozesses aufgefordert, Prioritäten zu setzen und den Lösungsraum einzuschränken. Diese Verfahren beruhen in der Regel auf a posteriori-Verfahren, die nicht vollständig durchgerechnet werden sollen. Auf diese Weise kann der Rechenaufwand erheblich reduziert werden, ohne den Benutzer zu zwingen, sich von Anfang an auf bestimmte Präferenzen festzulegen.

#### 3.7.2 Visualisierung der Paretofront

Für mehr als zwei Kriterien kann die Paretofront nicht mehr auf die in Abb. 3.24 gezeigte Weise dargestellt werden. Eine Möglichkeit, Alternativen für Entscheidungsprobleme mit vielen Kriterien zu visualisieren sind Sterndiagramme [KsZ91]. Ein Beispiel zeigt Abb. 3.25. Jedes Ziel bildet eine Spitze des Sterns. Auf der Achse vom Mittelpunkt zur Spitze wird für jede Alternative der Wert eingetragen und diese über alle Kriterien miteinander verbunden. Auf diese Weise entsteht für jede Alternative ein Stern. Je größer das Volumen des Sterns, desto besser ist der Erfüllungsgrad über alle Kriterien. Wenige Alternativen können auf diese Weise übersichtlich dargestellt werden, bei zu großer Anzahl wird die Unterscheidung schwierig. Aufgrund der Sternform können viele Kriterien einbezogen werden.

Eine andere Möglichkeit der Visualisierung sind selbstorganisierende Karten. Dies sind neuronale Netze mit einem unüberwachten Lernverfahren. Dabei wird die  $m$ -dimensionale Paretofront auf ein zweidimensionales hexagonales Netz projiziert, das eingefärbt wird. Die Erstellung ist aufwändig und die Interpretation der Karten nicht intuitiv. Im Zusammenhang mit robuster Optimierung wurden sie von Shimoyama benutzt [SLJ<sup>+</sup>09].

### 3.7.3 Multikriterielle Optimierung auf Metamodellen

Aufgrund des hohen Berechnungsaufwandes ist eine Paretooptimierung bei lang laufenden Computersimulationen nicht möglich. In dieser Dissertation wird die Paretofront nur auf Metamodellen bestimmt, die sehr schnell auszuwerten sind. Daher werden a posteriori-Verfahren benutzt. Die multikriterielle Optimierung auf Metamodellen ist in der Literatur gut beschrieben, vgl. hierzu z.B. [SPKA01], [SPKA97], [QHS<sup>+</sup>05] und [AC09].

## 3.8 Clusteranalyse

Die Aufgabe der Clusteranalyse ist es, Punkte, die nahe beieinander liegen, in einem Cluster zu gruppieren und alle Punkte in wenigen solchen Clustern zusammenzufassen. Dies erlaubt es, die Teile von Daten zu bestimmen, die sich ähnlich verhalten. Die Clusteranalyse bietet eine kompakte Darstellung der Daten, geeignet für die direkte Analyse durch den Nutzer.

Man unterscheidet partitionierende und hierarchische Clusterverfahren.

### 3.8.1 Partionierende Verfahren

Bei ersteren wird die Anzahl der Cluster vorher festgelegt. Diese Verfahren starten mit einer Partition der Gesamtmenge und tauschen anschließend nach einer Vorschrift Punkte aus. Eine weit verbreitete Cluster-Technik dieser Gruppe ist das  $k$ -means-clustering [HW79].

Der Algorithmus startet mit der Angabe der gewünschten Anzahl  $k$  der Cluster. Als Startwerte werden zufällig  $k$  Punkte als Zentren ausgewählt. Alle Datenpunkte werden dem nächstliegenden Cluster-Zentrum zugeordnet. Als Distanzmaß wird meist der euklidische Abstand benutzt. Dann wird das Zentrum jedes Clusters neu berechnet. Dies geschieht über die Bildung eines Mittelwertes. Hierher rührt auch der Name „ $k$ -means“. Die Punkte werden nun den neuen Zentren zugeordnet. Dieser Vorgang wird iterativ solange wiederholt, bis sich die Cluster-Zuweisungen nicht mehr ändern oder die vorher festgelegte maximale Anzahl von Iterationsschritten erreicht ist.

Der  $k$ -means-Algorithmus bestimmt nicht nur eine Aufteilung der Punkte in Cluster, sondern auch Clusterzentren, die als Repräsentanten eines Clusters genutzt werden können.

$k$ -means benötigt für  $n$  Punkte der Dimension  $l$   $\mathcal{O}(l \cdot n)$  Speicher und hat eine Berechnungskomplexität von  $\mathcal{O}(l \cdot n \cdot k \cdot \text{Anzahl Iterationen})$  ([TSK05], [NNC12b]).

### 3.8.2 Hierarchische Verfahren

Die hierarchischen Clusterverfahren lassen sich in zwei Typen aufteilen:

Bei „Top-Down“-Verfahren gehören am Anfang alle Punkte zu einem Cluster. Über Distanzvergleiche wird es in immer kleinere Cluster aufgeteilt. Es wird beendet, wenn eine festgelegte Anzahl von Schritten erreicht ist oder ein Abbruchkriterium erfüllt ist, etwa wenn der minimale Abstand der Cluster einen vorher festgelegten Wert übersteigt. Anderenfalls würde das Verfahren zu einem Ende kommen, wenn jeder Punkt ein eigenes Cluster bildet.

„Bottom-Up“-Verfahren starten mit dieser Aufteilung. In den Folgeschritten werden benachbarte Cluster zusammengefasst. Auch hier muss ein geeignetes Abbruchkriterium festgelegt werden, da es sonst in einem Cluster enden würde.

Diese Verfahren heißen hierarchisch, da sie eine Hierarchie von Clusteraufteilungen erstellen. Die Baumstruktur kann in Dendogrammen visualisiert werden. Ein Austausch von Punkten zwischen Clustern wie bei den partitionierenden Verfahren findet nicht statt. Die Größe eines Clusters kann bei hierarchischen Verfahren stark unterschiedlich sein. Für die Distanzberechnung der Cluster können verschiedene Distanzmaße benutzt werden, z.B. der durchschnittliche, der maximale oder der minimale Abstand jedes Punktes des einen Clusters zu jedem Punkt des anderen. Hierarchische Verfahren sind speicher- und rechenintensiv: Der Speicherbedarf beträgt  $\mathcal{O}(n^2)$  und die Anzahl der nötigen Operationen bis zu  $\mathcal{O}(n^3)$  [NNC12b].

In dieser Arbeit wird die Clusteranalyse angewendet, um die robusten Optima geeignet zu gliedern und zunächst durch wenige repräsentative Punkte darzustellen. Wegen dieser Eigenschaft, der einfacheren Handhabbarkeit und des geringen Rechenaufwandes wird der  $k$ -means-Algorithmus benutzt.



## Kapitel 4

---

# Robuste Optimierung

---

Die in Kapitel 3.7 beschriebene deterministische Optimierung findet in vielen Bereichen von Wissenschaft und Technik Anwendung. Dabei wird implizit angenommen, dass die nominalen Werte der Eingangsvariablen tatsächlich eingehalten werden. Bei der realen Umsetzung können aber vielfältige Toleranzen auftreten: Im Ingenieurbereich sind dies z.B. Schwankungen der nominalen Geometrie aufgrund von Fertigungstoleranzen, Unsicherheiten bei den Belastungen und Spannungen aufgrund von wechselnden Umwelteinflüssen sowie Materialschwankungen durch schwankende Umgebungsbedingungen bei der Fertigung. Die deterministische Problemformulierung ist daher kein realistisches Modell der Wirklichkeit, da sie die real auftretenden Schwankungen außen vor lässt.

Zudem liegen deterministisch-optimale Designs häufig am Rande des zulässigen Bereichs. Ihre Eigenschaften können sich schon bei einer kleinen Abweichung vom nominalen Wert dramatisch verschlechtern. Solche Designs sind nicht „robust“.

In der Praxis werden deswegen Sicherheitsfaktoren zur Einhaltung der Nebenbedingungen in die Berechnungen aufgenommen. Sie basieren häufig auf Erfahrungswerten (gängige Sicherheitsfaktoren sind 1,2 bis 3), die nicht aus den vorhandenen Streuungen abgeleitet wurden. Alle Unsicherheiten werden so pauschal in einer Zahl konzentriert. Welche Eingabevariablen wie viel zur Streuung beitragen, bleibt ungeklärt, ebenso die Frage, wie stark die Zielvariablen tatsächlich streuen. Die resultierenden optimalen Design sind oft zu konservativ.

### 4.1 Begriff der Robustheit

Eine umfassende Betrachtung bezieht die Streuung der Eingangsvariablen in die Berechnung ein. Entsprechend erhält man für die Zielvariablen nicht mehr nur nominale Werte, sondern ebenfalls streuende Ergebnisse (vgl. Abb. 4.1). Diese müssen zunächst geeignet quantifiziert werden. In einem nächsten Schritt wird dann das Schwankungsverhalten der Zielvariablen bewertet. Ziel ist es, eine sogenannte „robuste“ Lösung zu finden.

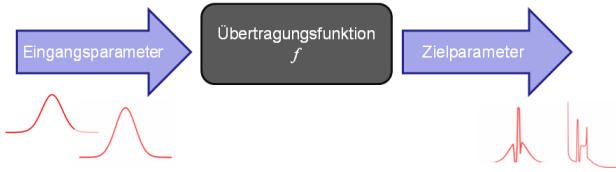


Abb. 4.1: Ein Teilgebiet der robusten Optimierung ist zu bestimmen, wie die Eingangsstreuung von der Funktion auf die Ausgangsstreuung übertragen wird.

Für den Begriff der Robustheit und damit auch des robusten Designs gibt keine allgemein anerkannte Definition. So bezeichnet Taguchi Robustheit als „den Zustand, in der die Technologie-, Produkt- oder Prozessperformance minimal empfindlich ist gegenüber den Faktoren, die diese Variabilität auslösen“[TCT00]. Suh sagt, „robustes Design ist definiert als das Design, das die funktionalen Anforderungen sogar dann erfüllt, wenn die Design-Parameter und die Prozess-Variablen große Toleranzen aufweisen“[Suh01].

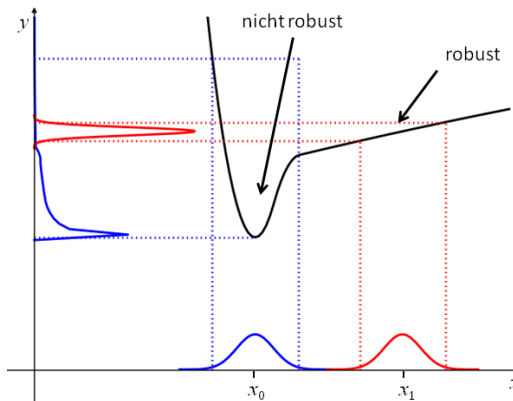


Abb. 4.2: Grundidee des robusten Designs

Allgemein wird ein Design als *robust* bezeichnet, wenn es unempfindlich gegenüber den Schwankungen der Eingangsvariablen ist.

Abbildung 4.2 illustriert diese Idee:

Die Funktion  $f$  hänge nur von der Eingangsvariablen  $x$  ab. Sie unterliege einer Unsicherheit, die durch die Normalverteilung modelliert werden kann. In  $x_0$  ist das deterministische Minimum der Funktion. Dort wird die Streuung der Eingangsvariablen auf einen weiten Bereich der Zielvariablen übertragen (blaue Streuung). Die Streuung der Zielvariablen ist groß.  $x_0$  ist nicht robust. Bei  $x_1$  wird die Eingangsstreuung nur auf einen kleinen Bereich von  $y$  übertragen, da  $f$  in diesem Bereich flach ist (rote Streuung).  $x_1$  ist daher robust. Die robusten Bereiche sind die flachen Regionen einer Funktion.

Abbildung 4.2 veranschaulicht auch das Dilemma der robusten Optimierung:  $x_1$  induziert zwar eine kleinere Streuung als  $x_0$ , die Zielvariable nimmt aber einen größeren Wert an und ist daher weniger optimal. Die kleinere Streuung muss durch einen Verlust an Performance erkaufte werden. Die beiden gegenläufigen Ziele „Minimierung der Funktion“ und „Minimierung der Streuung“ müssen gegeneinander abgewogen werden.

In dieser Arbeit wird unter *robuster Optimierung* folgendes verstanden:

**Robuste Optimierung ist die Suche nach einem Optimum, das die ursprünglichen Zielkriterien im Mittel möglichst gut erfüllt und gleichzeitig möglichst kleine Schwankungen der Zielvariablen aufweist.**

Im Beispiel wurde die übertragene Dichtefunktion berechnet und die Robustheit visuell abgeschätzt. In diesem Kapitel wird dieser Begriff nun mathematisch präzise formuliert. Diese Robustheitsmaße charakterisieren die Zielvariablen und ihre Verteilung. Sie messen, wie groß der negative Einfluss der Streuung der Zielvariablen ist. Hierfür sind verschiedene Definitionen möglich, bei der jede ein anderes Verständnis von Robustheit widerspiegelt.

Mit der bloßen Quantifizierung von Robustheit ist das robuste Optimum jedoch noch nicht gefunden. Die Robustheitsmaße müssen in eine - nun unter Berücksichtigung der Streuungen - formulierte Optimierungsaufgabe integriert werden. Die geeignete Formulierung kann jedoch, wie später gezeigt wird, nicht vom gewählten Robustheitsmaß getrennt werden.

Manchmal wird in der Praxis der folgende weniger aufwändige Weg beschritten:

Es wird eine deterministische Optimierung durchgeführt und für das gefundene optimale Design anschließend die Auswirkungen der Streuungen in den Eingangsvariablen auf die Zielvariablen bestimmt. Wenn die Streuung der Zielvariablen in einem akzeptablen Bereich liegt, wird die Optimierung beendet. Es findet kein Vergleich mit anderen möglichen Lösungen statt. Falls die Unsicherheit in den Zielen zu groß ist, wird ein neuer deterministischer Optimierungslauf gestartet, der ein anderes lokales Optimum liefert und anschließend seine Robustheit überprüft. Dieses Verfahren wird fortgesetzt bis eine akzeptable Lösung gefunden ist. Auf diesem Weg erhält man eine robuste Lösung unter den deterministischen Optima. Bezieht man die Robustheit als Optimierungskriterium von Anfang an mit ein, bestimmt man die tatsächlichen robusten Optima. Sie müssen nicht mit lokalen deterministischen Optima übereinstimmen. Diesen Sachverhalt veranschaulicht Abbildung 4.3: Das globale Minimum der Funktion in  $x_0$  ist nicht robust, das lokale Minimum in  $x_1$  weist eine kleinere Streuung der Zielvariablen auf als  $x_0$ . Im Vergleich dazu ist die in  $x_2$  induzierte Streuung kleiner, ebenso sein Funktionswert.

## 4.2 Ablauf einer robusten Optimierung

Unter robuster Optimierung im weiteren Sinne wird der gesamte Ablauf der Integration von Unsicherheiten in den Optimierungsprozess von der Problemdefinition bis hin zu ihrer Lösung verstanden. Sie läuft in vier Schritten ab ([KN05], S.329):

### 1. Identifikation und Modellierung der Unsicherheit

Im ersten Schritt werden die unsicheren Parameter identifiziert (ggf. durch

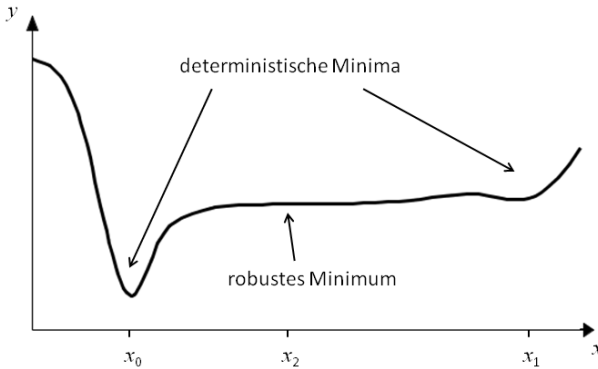


Abb. 4.3: Deterministische Minima und robustes Minimum

Vorschalten einer Sensitivitätsanalyse) und Daten über die Unsicherheit gesammelt. Anschließend wird sie geeignet mathematisch modelliert. Einzelheiten hierzu werden in Kapitel 4.3 beschrieben.

2. Auswahl eines geeigneten Robustheitsmaßes und Formulierung der robusten Optimierungsaufgabe

Robustheit kann auf verschiedene Arten charakterisiert werden. Dabei gibt es zwei Hauptrichtungen:

- a) Viele Maße fassen Robustheit zu *einer* Kennzahl zusammen. Diese Maße charakterisieren entweder die negativen Auswirkungen der Schwankungen oder gewichten Lage und Streuung der Zielvariablen. Anschließend wird dieses Maß minimiert. Die robuste Optimierung hat dann genauso viele Zielvariablen wie die deterministische. Hiermit beschäftigt sich Kapitel 4.5.
- b) Andere Ansätze führen jeweils ein Maß für die Lage der übertragenen Verteilung und eines für ihre Streuung ein und suchen den besten Kompromiss aus diesen beiden Kriterien. Die Optimierungsaufgabe wird in jedem Fall multikriteriell und die Anzahl der Zielvariablen verdoppelt sich. Das Ergebnis ist kein einzelner Wert, sondern die Paretofront, vgl. Kapitel 4.6.

Das benutzte Robustheitsmaß beeinflusst entscheidend das Ergebnis der robusten Optimierung. Die Auswahl sollte sich nach den Präferenzen des Entscheiders richten. Eine Diskussion der vorgestellten Maße erfolgt in Kapitel 4.7.

3. Berechnung der Übertragung der Unsicherheiten auf die Zielvariablen  
Im dritten Schritt wird berechnet, wie sich die im ersten Schritt charakterisierten Streuungen der Eingabevariablen auf die Zielvariablen auswirken. Die entsprechenden Berechnungsmethoden werden jeweils nach der Vorstellung des

Robustheitsmaßes diskutiert. Eine Ausnahme bildet das hier favorisierte Robustheitsmaß basierend auf Quantilen. Wegen seiner zentralen Bedeutung für diese Dissertation wird seine Berechnung in einem eigenen Kapitel 5 vorgestellt.

#### 4. Lösung der robusten Optimierungsaufgabe

Im letzten Schritt wird die robuste Optimierungsaufgabe gelöst. Abhängig von der Anzahl der Zielvariablen und dem gewählten Robustheitsmaß sind ein- oder mehrkriterielle Verfahren zu verwenden. Dies geschieht in Kapitel 4.8.

Häufig wird unter Robustheit auch die zuverlässige Einhaltung der Nebenbedingungen verstanden. Diese Definition entspricht nicht der hier eingeführten von Robustheit als Minimierung der Streuung. Trotzdem werden in Kapitel 4.9 einige Formulierungen für diesen Robustheitsbegriff angegeben.

Eine Abgrenzung der robusten Optimierung zu anderen benachbarten Teilgebieten wird in Kapitel 4.10 vorgenommen.

### 4.3 Klassifizierung von Unsicherheiten

Unsicherheiten lassen sich in zwei Klassen einteilen [Hel97]:

- *Aleatorische Unsicherheit* ist die objektive, nicht-reduzierbare Unsicherheit, die selbst bei ausreichender Information über die Unsicherheit der Eingabevariablen vorhanden ist, z.B. Fertigungstoleranzen, Temperatur- oder Feuchtigkeitsschwankungen. Sie wird auch als stochastische oder zufällige Unsicherheit bezeichnet. Die geeignete Art, diese zu beschreiben, sind Wahrscheinlichkeitsverteilungen.
- *Epistemische* (erkenntnistheoretische) Unsicherheit ist die subjektive und reduzierbare Unsicherheit, die aus einem Mangel an Information über die Unsicherheit der Eingabevariablen stammt. Sie könnte prinzipiell beseitigt werden. Hierzu zählen auch Modellunsicherheiten oder Fehler bei der numerischen Simulation. Epistemische Unsicherheiten lassen sich über Wertebereiche, in denen die Eingangsvariablen schwanken können, oder über Maße aus der fuzzy-logic-Theorie charakterisieren. Ein Überblick über die Ansätze zur epistemischen Unsicherheit gibt [Che00] und [YCDG07]. Andere Autoren [Lan10, KO01, BS07] trennen von den epistemischen Unsicherheiten noch weitere wie Modell-Angemessenheit und Modellunsicherheit oder sie unterscheiden Simulationsfehler, Metamodellfehler und Implementationsfehler [SH08].

Hier wird nur die aleatorische Unsicherheit behandelt. Es wird angenommen, dass sich für jede Eingabevariable die Wahrscheinlichkeitsverteilung angeben lässt.

Die Beschreibung der Unsicherheit der Eingangsvariablen stellt in der Praxis oft ein großes Problem dar. Während die Verteilungen der Fertigungstoleranzen meistens bekannt sind, müssen die der Materialschwankungen erst über aufwändige Versuche ermittelt werden. Zur Absicherung wird dann die robuste Optimierung für verschiedene Eingangsverteilungen durchgeführt. Dies verdeutlicht den Bedarf an effizienten Algorithmen zur robusten Optimierung.

## 4.4 Formulierung der robusten Optimierungsaufgabe

Aus den Designvariablen  $x = (x_1, \dots, x_d)^T$  werden durch die Berücksichtigung der Streuung Zufallsvariablen  $X = (X_1, \dots, X_d)^T$ . Dann sind auch die Zielvariablen Zufallsvariablen  $Y = (Y_1, \dots, Y_l)^T$  mit

$$Y = f(X).$$

Da  $f$  die Streuung überträgt, wird  $f$  auch *Übertragungsfunktion* genannt.

**Es wird angenommen, dass die Eingangsvariablen  $X_i, i = 1, \dots, d$  stochastisch unabhängig sind.**

Im folgenden werden verschiedene Charakterisierungen von Robustheit vorgestellt. Die Definitionen beziehen sich zunächst auf eine feste Parameterkombination  $x_0 = (x_{01}, \dots, x_{0d})^T$ . Die zugehörige Zufallsvariable wird der Einfachheit halber mit  $X$  und die übertragene Verteilung analog mit  $Y = f(X)$  bezeichnet. In der robusten Optimierung werden sie mit anderen Maßen kombiniert, die die Erfüllung der ursprünglichen Optimierungskriterien sicherstellen. Bei der Formulierung der robusten Optimierungsaufgabe muss sich die Tatsache, dass nun Zufallsvariablen optimiert werden sollen, in der Formulierung der Zielvariablen und der Nebenbedingungen wiederfinden.

Entsprechend lautet dann die stochastische Formulierung [PKHLL06]:

$$\text{Minimiere } Y = F(X), \quad (\text{a})$$

$$\text{so dass } G_i(X) \leq 0, \quad i = 1, \dots, I \quad (\text{b})$$

$$\text{und } x_h^L \leq X_h \leq x_h^U, \quad h = 1, \dots, k \quad (\text{c})$$

erfüllt sind.

Dabei sind  $F$  und  $G_i$  Funktionen, die die Eingangsstreuung berücksichtigen.

### 4.4.1 Bei einkriteriellen Robustheitsmaßen

Bei den einkriteriellen Robustheitsmaßen wird jedes ursprüngliche Kriterium  $f_i$  durch jeweils ein Robustheitsmaß  $F_i, i = 1, \dots, l$  ersetzt. Die robuste Formulierung der Zielfunktion lautet dann:

$$\text{Minimiere } Y = F(X) = (F_1(X), \dots, F_l(X))^T \quad (4.1)$$

Die Anzahl der Zielfunktionen bleibt damit gleich. In Kapitel 4.5 werden verschiedene Formulierungen für  $F_i$  vorgestellt.

### 4.4.2 Bei multikriteriellen Robustheitsmaßen

Die multikriteriellen Robustheitsmaße benutzen jeweils ein Maß für die Lage der Verteilung und eines für die Streuung. In diesem Fall verdoppelt sich die Anzahl der Kriterien und es ist in jedem Fall ein multikriterielle Optimierung durchzuführen:

$$\text{Minimiere } Y = F(X) = (L_1(X), \dots, L_l(X), S_1(X), \dots, S_l(X))^T \quad (4.2)$$

Sie werden in Kapitel 4.6 betrachtet.

## 4.5 Einkriterielle Robustheitsmaße

In diesem und im folgenden Unterkapitel werden verschiedene Robustheitsmaße vorgestellt. Ähnliche Übersichten finden sich in [Jur07], [Bau12] und [BS07].

Einkriterielle Robustheitsmaße verwenden die in Gleichung 4.1 angegebene Formulierung der robusten Optimierungsaufgabe.

### 4.5.1 Worst-Case-Maß

Beim Worst-Case-Maß wird berechnet, wie sich die Schwankung von  $X$  schlimmstenfalls auf  $Y$  auswirkt. Es wird die maximale Schwankungsbreite bestimmt [Tro96]:

$$\text{WCM}_N(x) = \sup_{\tilde{x} \in N(x)} f(\tilde{x})$$

Dabei ist  $N(x)$  der Variationsbereich um  $x$ .

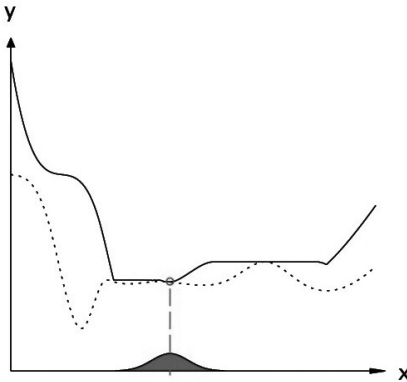


Abb. 4.4: Worst-Case-Maß  $\text{WCM}_N$  für  $N(x) = [\mu - 3\sigma; \mu + 3\sigma]$  (durchgezogene Linie),  $f$  gestrichelte Linie

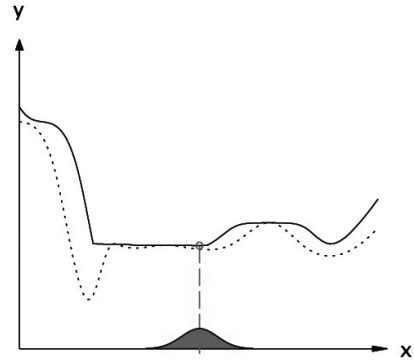


Abb. 4.5: einseitiges Quantilmaß,  $\text{EQM}_{0,9}$  mit  $q = 0,9$  (durchgezogene Linie),  $f$  gestrichelte Linie

Ist der Wertebereich der Eingangsverteilungen beschränkt, wird dieser als Variationsbereich  $N(x)$  benutzt. Bei unbeschränkten Eingangsverteilungen kann das Worst-Case-Maß  $\text{WCM}_N(x)$  beliebig groß werden. Um dies zu verhindern, ist eine Einschränkung des Variationsbereichs von  $x$  auf  $\mu + / - 3\sigma$  bei Normalverteilungen oder  $x_{0,95} - x_{0,05}$  bei beliebigen Verteilungen sinnvoll.

In Abbildung 4.4 wird eine Beispielfunktion  $f$  mit seinem Worst-Case-Maß verglichen. Im Bereich des globalen Minimums von  $f$  nimmt  $\text{WCM}_N(x)$  große Werte an, da  $f$  dort im Schwankungsbereich von  $x$  nach links stark ansteigt. Das Minimum des Worst-Case-Maßes liegt im mittleren flachen Bereich der Funktion. Kleinere Schwankungen im Funktionsverlauf haben keinen Einfluss, z.B. bleiben die Worst-Case-Werte im Bereich des lokalen Maximums der Funktion konstant.  $\text{WCM}_N(x)$  ist an mehreren Stellen nicht differenzierbar.

Offensichtlich ist es ein sehr konservatives Maß, das nur die Grenzen der Eingangsverteilung benutzt. Deswegen findet es vor allem Anwendung für epistemische Unsicherheiten, bei denen oft nur die obere und untere Grenze des Variationsbereichs bekannt ist. Verschiedene Eingangsverteilungen mit dem gleichen Schwankungsbereich ergeben die selben Worst-Case-Werte. Die Form der Ausgangsverteilung bleibt unberücksichtigt.

Es wird auch als *robust counterpart approach* [BS07] oder *robust regularization* [Lew] bezeichnet. Das Worst-Case-Maß in der robusten Optimierung entspricht dem Minimax-Prinzip aus der statistischen Entscheidungstheorie. Es *minimiert* die *maximale* Auswirkung der Streuung im Eingangsparameter  $x$ .

Bei diesem sehr konservativen Ansatz wird von harten Nebenbedingungen ausgegangen. In diesem Fall geben [BTN02] und [Zha07] Lösungen für einfache Optimierungsprobleme an. Im wesentlichen sind dies lineare Optimierungsprobleme mit quadratischen oder konischen Nebenbedingungen. Sie werden als *robust linear quadratic* oder *robust conic programming* bezeichnet. Einen aktuellen Überblick über die umfangreiche Literatur geben [BTGN09] und [BBC11]. Stinstra [SH08] wendet das Worst-Case-Maß auf Kriging-Metamodelle an.

### 4.5.2 Einseitiges Quantil-Maß

Die Verteilung von  $Y$  kann auch über ein oberes Quantil charakterisiert werden [Jur07, BS07]:

$$\text{EQM}_q(x) = y_q(x) = \inf\{y \mid P(f(X) \leq y) \geq q\} \quad 0 \leq q \leq 1$$

Dies ist eine Verallgemeinerung des Worst-Case-Maßes, bei dem das Quantil für  $q = 1$  bestimmt wird. Im Vergleich zu ihm ist es weniger konservativ, da  $1 - q$  Prozent der Ausgangsverteilung unberücksichtigt bleiben.

Für unterschiedliche Werte von  $q$  nimmt es andere Werte an. Über die Wahl von  $q$  kann daher gesteuert werden, welches „Sicherheitsniveau“ garantiert werden soll, denn  $\text{EQM}_q(x)$  sagt per definitionem aus, dass  $q$  Prozent der Ausgangsverteilung unterhalb dieses Wertes liegen.

Beim einseitigen Quantilmaß wird nicht die Größe der Streuung von  $Y$  betrachtet, sondern nur die negativen Auswirkungen nach oben hin. Deswegen sind nur große Werte für  $q$  sinnvoll.

Abbildung 4.5 zeigt das einseitige 90%-Quantilmaß und die Funktion  $f$ . Das Minimum von  $\text{EQM}_q$  liegt im mittleren flachen Bereich der Funktion. Im Vergleich zum Worst-Case-Maß reagiert es empfindlicher auf kleine Schwankungen im Funktionsverlauf. Es besitzt konstante Bereiche und nicht-differenzierbare Stellen.

Ein Sonderfall ist die Wahl von  $q = 0,5$ . In diesem Fall wird für jeden  $x$ -Wert der Median der übertragenen Verteilung bestimmt. Es wird die Lage der Verteilung minimiert, nicht deren Streuung.

Dieses Maß zeigt Abbildung 4.6. Der Verlauf von  $\text{EQM}_{0,5}$  folgt in flachen Bereichen dem Funktionsverlauf, die Extrema sind nicht so stark ausgeprägt. Als Funktion von  $x$  ist es differenzierbar. Das Minimum stimmt mit dem globalen Minimum von  $f$  überein.



Praxisanwendungen dieser Formulierung im Bereich der robusten Optimierung sind bisher nicht bekannt.

In [BC10] wird das einseitige Quantilmaß für ein Einstufungs- und Auswahlverfahren zum Vergleich einer endlichen Menge stochastischer Systeme benutzt. Angün [Ang11] führt den in der Finanzwelt als Risikomaß verwendeten *Average Value-at-Risk* als spezielles einseitiges Quantilmaß ein. Es gilt:

$$\begin{aligned} \text{AVAR}_q(x) &= E(Y|Y > y_q(x)) \\ &= y_q(x) + \underbrace{E(Y - y_q(x)|Y > y_q(x))}_{\text{mittlere Überschreitung im Überschreitungsfall}} \end{aligned}$$

Der zweite Summand misst die negativen Auswirkungen der  $1 - q\%$  der Verteilung, die das  $q$ -Quantil übersteigen.

**Berechnungsmethoden**

Methoden zur Berechnung von Quantilen wurden schon in Kapitel 3.3 beschrieben. Da sie eine zentrale Rolle in der hier beschriebenen Methodik zur robusten Optimierung einnehmen, wird in Kapitel 5 untersucht, wie sich ihre Berechnung effizienter gestalten lässt.

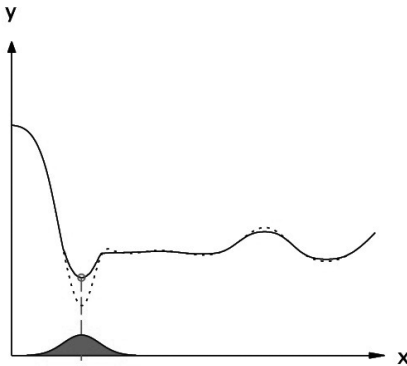


Abb. 4.6: Median  $y_{0,5}$  (durchgezogene Linie),  $f$  gestrichelte Linie

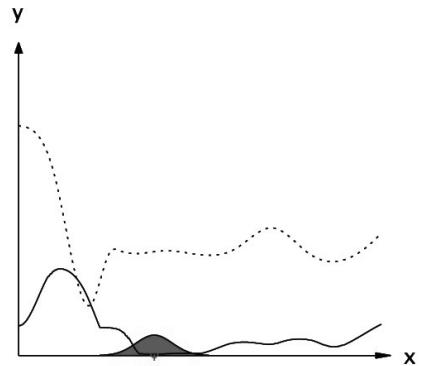


Abb. 4.7: Quantildifferenz  $QD_Y$  mit  $q = 0,05$  (durchgezogene Linie),  $f$  gestrichelte Linie

**4.5.3 Quantildifferenz-Maß**

Die *Quantildifferenz* zweier symmetrische Quantile, auch *Quantilabstand* genannt

$$QD_q(x) = y_q(x) - y_{1-q}(x)$$

ist ein Maß für die Streuung einer Verteilung und kann damit auch als Robustheitsmaß verwendet werden [DSC04].

Auch bei diesem Maß kann der Entscheider den gewünschten Grad an Robustheit über die Wahl von  $q$  steuern. Mit steigendem  $q$  werden immer größere Teile der übertragenen Verteilung mit in die Berechnung einbezogen. Für  $q = 1$  erhält man wieder das Worst-Case-Maß.

Dieses Maß misst allein die Streuung von  $f(X)$ . Der Funktionswert von  $x$  oder ein anderes Maß für die Lage dieser Verteilung bleiben außen vor.

In Abbildung 4.7 wurde die Quantildifferenz für  $q = 0,05$  der Beispielfunktion bestimmt. Sie wird groß in den Bereichen, in denen die Funktion stark ansteigt oder fällt. Sein Minimum liegt im flachen Bereich von  $f$ . Für diese Beispielfunktion ist der zugehörige Funktionswert relativ klein. Es lässt sich leicht eine andere Funktion konstruieren, deren flachster Bereich sehr große Funktionswerte aufweist. In diesem Fall müsste eine robuste Optimierung mit der Quantildifferenz als Robustheitsmaß mit einem großen Verlust an Optimalität in den ursprünglichen Zielkriterien erkauft werden.

Es wird in [EPM11] verwendet.

#### 4.5.4 Wahrscheinlichkeitsschwellen-Maß

Mit dem einseitigen Quantilmaß eng verwandt ist das Wahrscheinlichkeitsschwellen-Maß [BS07]. Zu einem vorgegebenen Schwellenwert  $s$  wird  $P(Y \leq s)$  bestimmt, d.h. die Wahrscheinlichkeit, das die übertragene Verteilung unterhalb von  $s$  liegt.

Ziel der Optimierung ist dann,  $x$  so zu bestimmen, dass diese Wahrscheinlichkeit möglichst groß wird. Da alle Robustheitsmaße einheitlich minimiert werden sollen, wird hier der negative Wert als Robustheitsmaß definiert:

$$\text{WSM}_s(x) = -P(Y \leq s)$$

Dieses Maß kann nur Werte zwischen 0 und -1 annehmen. Es ist abhängig vom gewählten Schwellenwert  $s$ . Der Entscheider kann hierüber den gewünschten Grad an Robustheit bestimmen. Wird  $s$  allerdings zu groß gewählt, liegt die gesamte übertragene Verteilung unterhalb dieser Grenze und man erhält den Wert -1 für  $\text{WSM}_s(x)$ . Wird  $s$  kleiner als das absolute Minimum von  $f$  gewählt, ist es konstant 0. Die Streuung der übertragenen Verteilung geht nicht direkt in dieses Maß ein.

In Abbildung 4.8 ist das Schwellenmaß für die Beispielfunktion gezeigt. Es hat einen welligeren Verlauf als die einseitigen Quantilmaße. Das Optimum liegt im Bereich des lokalen Minimums.

Anwendungen dieses Maßes fehlen bisher.

#### 4.5.5 Erwartungswert-Nutzenfunktion-Maße

Die Idee dieses Maßes beruht auf der Einführung einer Nutzenfunktion  $U$  für  $f(X)$  [BS07, DK04]. Der daraus bestimmte Erwartungswert ist dann ein Robustheitsmaß:

$$\text{ENM}(x) = E(U(f(X)))$$

Definiert man  $U(Y) = U(f(X)) = \text{sign}(f)|f|^k$  erhält man *Momenten-Maße*.

Bei  $k = 1$  ist dies der Erwartungswert von  $f$ .

Für  $k < 1$  werden die mittleren Verteilungswerte stärker gewichtet, während für  $k > 1$  der Einfluss der extremen Werte verstärkt wird.

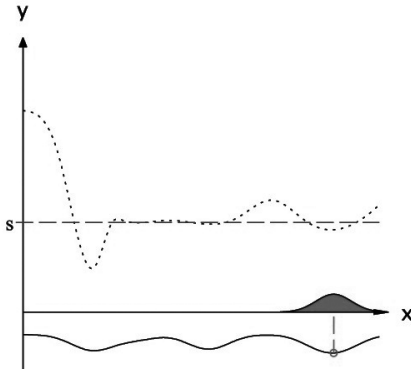


Abb. 4.8: Schwellenmaß  $WSM_s$ ,  $s = 4$  (durchgezogene Linie),  $f$  gestrichelte Linie

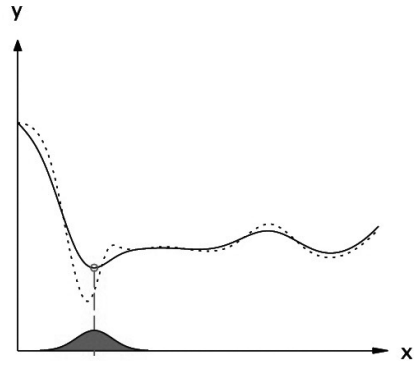


Abb. 4.9: Erwartungswertmaß,  $\mu_Y$  (durchgezogene Linie),  $f$  gestrichelte Linie

Sinnvolle Nutzenfunktionen sind auch solche, die zu Streuungsmaßen führen. Benutzt werden z.B. Varianz, Standardabweichung oder relative Varianzmaße. Darüber hinaus gibt es Nutzenfunktionen, die speziell auf die Eigenheiten der Anwendung zugeschnitten sind.

Auf die beiden gebräuchlichsten Erwartungswert-Nutzenfunktion-Maße wird kurz genauer eingegangen:

### 4.5.6 Erwartungswert-Maß

Das Erwartungswertmaß ist definiert durch

$$EWM(x) = E(f(X))$$

Es ist ein Maß für die Lage der übertragenen Verteilung, d.h. für ihren durchschnittlichen Wert. Die Streuung bleibt unberücksichtigt.

Abbildung 4.9 zeigt  $EWM(x)$ . Es folgt in Teilen dem Verlauf von  $f$ , glättet aber stärkere Ausschläge. Das Minimum liegt nicht im absoluten Minimum von  $f$ , sondern kurz daneben, da in der Nähe eines Extremums schiefe Verteilungen entstehen, die dieses Maß nicht gut repräsentieren kann. Das Erwartungswert-Maß ist überall differenzierbar.

### 4.5.7 Standardabweichung-Maß

Die Standardabweichung der übertragenen Verteilung

$$\sigma_Y(x) = (E(f(X) - E(f(X)))^2)^{\frac{1}{2}}$$

ist ein auf einem Erwartungswert beruhendes Maß für die Streuung von  $f(X)$ . Wie bei der Quantildifferenz bleibt die Lage der Verteilung unberücksichtigt. In Abbildung 4.10 wurde  $\sigma_Y(x)$  für die Beispielfunktion berechnet. Da alleine die Streuung minimiert wird, liegt das Minimum im flachen Bereich von  $f$ .

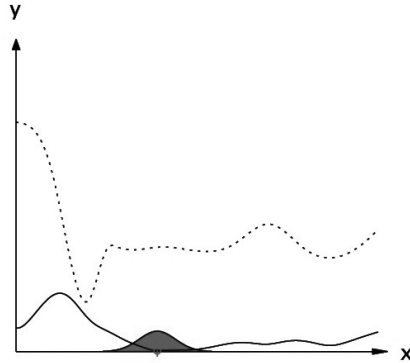


Abb. 4.10: Standardabweichung  $\sigma_Y$  (durchgezogene Linie),  $f$  gestrichelte Linie

### Berechnungsmethoden

Unter den verschiedenen Methoden zu ihrer Berechnung ist die Annäherung über die Taylor-Reihe am weitesten verbreitet.

Falls  $f$  stetig und zweimal differenzierbar ist, gilt für die Verteilung um ein festes  $x_0$  [CATM96, GLK02]:

$$\begin{aligned}\mu_{Y_0}(x_0) &\approx f(x_0) && \text{(erste Ordnung) oder} \\ \mu_{Y_0}(x_0) &\approx f(x_0) + \frac{1}{2} \sum_{i=1}^d f_i''(x_0) \sigma_i^2 && \text{(zweite Ordnung)}\end{aligned}$$

Die Taylor-Entwicklung für die Varianz lautet:

$$\begin{aligned}\sigma_{Y_0}^2(x_0) &\approx \sum_{i=1}^d (f_i'(x_0))^2 \sigma_i^2 && \text{(erste Ordnung) oder} \\ \sigma_{Y_0}^2(x_0) &\approx \sum_{i=1}^d (f_i'(x_0))^2 \sigma_i^2 + \frac{1}{2} \sum_{j=1}^d \sum_{i=1}^d (f_{ij}''(x_0))^2 \sigma_i^2 \sigma_j^2 && \text{(zweite Ordnung)}\end{aligned}$$

Die Taylor-Approximation ist nur dann befriedigend, wenn die Varianz von  $X$  klein ist und die Funktion nur mäßig nichtlinear. Eine Entwicklung der Taylor-Reihe bis zu einer höheren Ordnung muss nicht zu einer größeren Genauigkeit führen [PCG07]. Ein Nachteil ist die teure Berechnung der Ableitungen. Andere Methoden wie die sigma-point-Methode [PCG07], Gauss-Hermite Integration [HD06] oder Gauss'sche Quadratur [SK02] vermeiden dies.

Für höhere Dimensionen ist es besser, mit Stichproben um  $x_0$  gemäß der Verteilung der Eingabevariablen zu arbeiten und aus ihnen die ersten beiden Momente zu schätzen.

[Jin04] gibt analytische Formeln für Erwartungswert und Varianz der übertragenen Verteilung bei RBF-Modellen mit Gauss'scher Basisfunktion und gleich- oder normalverteilter Eingangsverteilung an. Die Berechnung erfolgt über die Tensor-darstellung des Metamodells und mehrere Zwischenschritte. Sie ist aufwändig und wurde in der Wissenschaft nicht weiter verfolgt.

**4.5.8 Gewichtete Zielfunktion aus Lage- und Streuungsmaß**

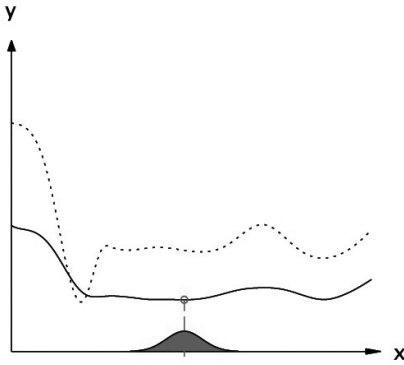


Abb. 4.11: gewichtete Zielfunktion,  $0,5\mu_Y + 0,5\sigma_Y$  (durchgezogene Linie),  $f$  gestrichelte Linie

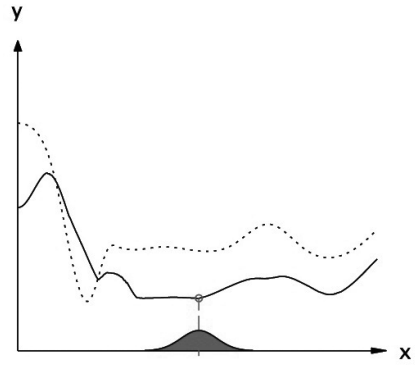


Abb. 4.12: gewichtete Zielfunktion  $0,5y_{0,5} + 0,5(y_{0,95} - y_{0,05})$  (durchgezogene Linie),  $f$  gestrichelte Linie

Die bisher vorgestellten Robustheitsmaße messen auf unterschiedliche Arten entweder die Streuung der übertragenen Verteilung oder ihren mittleren Wert. Dabei sind diese beiden Ziele oft gegenläufig. So befindet sich für die Beispielfunktion das Minimum der klassischen Lagemaße Median und Erwartungswert nahe beim globalen Minimum von  $f$ , während der kleinste Wert für die Streuungsmaße Quantildifferenz und Standardabweichung im flachen Bereich der Funktion liegen. Die Lagemaße entsprechen der „mittleren“ Performance der Funktion. Dagegen messen die Streuungsmaße einzig die „Robustheit“, sie lassen völlig die Lage unberücksichtigt. Ein Entscheider sucht häufig einen Ausgleich zwischen diesen beiden Zielen. Eine Möglichkeit, diesen herzustellen, ist die Bildung einer gewichteten Zielfunktion (vgl. Kapitel 3.7.1).

Dem Lagemaß wird ein Gewicht  $\alpha \in [0; 1]$  zugeordnet. Das Streuungsmaß wird entsprechend mit  $1 - \alpha$  gewichtet. Das Robustheitsmaß hat dann die Form

$$GZ_\alpha(x) = \alpha \cdot \text{Lagemaß} + (1 - \alpha) \cdot \text{Streuungsmaß}$$

Über die Wahl von  $\alpha$  kann der Entscheider seine eigenen Präferenzen einfließen lassen. Für  $\alpha = 0$  wird nur das Streuungsmaß berechnet, für  $\alpha = 1$  nur das Lagemaß. Statt der absoluten Maße können auch relative verwendet werden, um gleiche Größenordnungen zu erhalten. So wird in [LP01] das Maß

$$GZE_{rel}(x) = \alpha \frac{E(f(X))}{E(f(X))^*} + (1 - \alpha) \frac{\sigma_Y(x)}{\sigma_Y(x)^*}$$

verwendet, wobei  $E(f(X))^*$  der minimale Erwartungswert und  $\sigma_Y(x)^*$  die minimale Standardabweichung ist, wenn dieses Kriterium alleine betrachtet wird.

Als Lage- und Streuungsmaße kommen die Paare Erwartungswert und Standardabweichung oder Median und Quantildifferenz in Betracht:

### Erwartungswert und Standardabweichung

$$\text{GZE}(x) = \alpha E(f(x)) + (1 - \alpha)\sigma_Y(x)$$

Abbildung 4.11 zeigt die gewichtete Zielfunktion für Erwartungswert und Standardabweichung mit  $\alpha = 0,5$ . Das Minimum befindet sich im flachen Teil von  $f$  ist aber gegenüber der alleinigen Betrachtung der Standardabweichung weiter nach rechts in Richtung des dortigen lokalen Minimums gewandert.

### Median und Quantildifferenz

$$\text{GZM}(x) = \alpha \text{EQM}_{0,5}(x) + (1 - \alpha) \text{QD}_q(x)$$

In Abbildung 4.12 wurden Median und Quantildifferenz für  $q = 0,95$  jeweils gleich gewichtet. Auch dieses Robustheitsmaß wird im flachen Bereich von  $f$  minimal, im Vergleich zu  $\text{QD}_{0,95}$  liegt es näher am mittleren lokalen Minimum.

## 4.5.9 Taguchis Signal-Rausch-Verhältnis

Taguchi kommt das Verdienst zu, im Ingenieurbereich das Bewusstsein für den Einfluss von Unsicherheiten auf die Qualität eines Produktes geweckt zu haben. Seine Methoden werden noch heute in vielen Bereichen der Qualitätssicherung eingesetzt. Ausführliche Beschreibungen finden sich z.B. in [Tag86] und [Pha89].

Taguchi unterteilt den Entwicklungsprozess in drei Stufen:

- Im *Systemdesign* werden die Variablen  $x_i, i = 1, \dots, d$ , die das Zieldesign beeinflussen (die sog. Kontrollvariablen), und diejenigen, die unsicher sind  $\xi_j, j = 1, \dots, m$  (sog. Rausch-Variablen), identifiziert und quantifiziert sowie das grobe Design des Bauteils erstellt.
- Das *Parameterdesign* bestimmt die optimalen Werte für die Kontrollvariablen.
- Der letzte Schritt ist das *Toleranzdesign*. Hier geschieht eine Feinabstimmung der im zweiten Schritt gefundenen Parameterwerte im Hinblick auf die Minimierung der Toleranzen der Rausch-Variablen.

Bei seiner Entwicklung geht Taguchi davon aus, dass zur Bestimmung des optimalen Designs nur physikalische Versuche möglich sind. Deswegen benutzt er in den Phasen zwei und drei spezielle reduzierte Versuchspläne, die auf orthogonalen Feldern beruhen. Die Anzahl der benötigten Versuche steigt dabei exponentiell mit der Anzahl der Eingangsvariablen. Mit dem Aufkommen von Computersimulationen wurden auch Monte-Carlo-Verfahren benutzt.

Zur Charakterisierung der Streuung werden drei Varianten der mittleren quadratischen Abweichung (MSD, mean-square deviation) von einem Zielwert  $y_{Ziel}$  eingeführt.

$$\text{MSD}(x_0) = \frac{1}{m} \sum_{j=1}^m (f(x_0, \xi_j) - y_{Ziel})^2$$

Wenn das Zieldesign möglichst klein sein soll, vereinfacht sich der Ausdruck zu

$$\text{MSD}(x_0) = \frac{1}{m} \sum_{j=1}^m f^2(x_0, \xi_j)$$

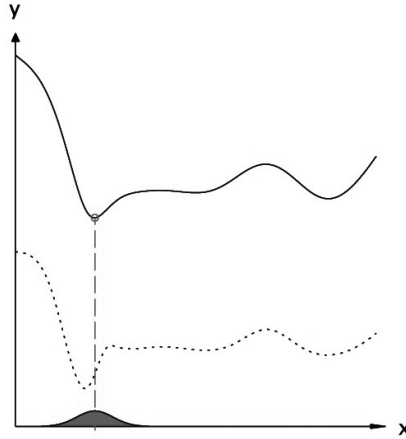


Abb. 4.13: Taguchis S/N (durchgezogene Linie),  $f$  gestrichelte Linie

Wenn es möglichst groß sein soll, schlägt Taguchi

$$\text{MSD}(x_0) = \frac{1}{m} \sum_{j=1}^m \frac{1}{f^2(x_0, \xi_j)}$$

vor. Als Robustheitsmaß wird schließlich das *Signal-Rausch-Verhältnis* bzw. *Taguchis Verlustfunktion* definiert als:

$$S/N = 10 \cdot \log_{10}(\text{MSD}(x_0))$$

Diese Methode ist in den letzten Jahrzehnten kontrovers diskutiert worden, siehe z.B. [BT87, NAM<sup>+</sup>92, Tro96]. So ist die Logarithmierung von  $\text{MSD}(x_0)$  nur eine monotone Transformation, die die Lage der robusten Optima nicht verändert. Falls Kontroll- und Rauschvariablen übereinstimmen, gilt für die zweite Variante:

$$\text{MSD}(x_0) = E(f^2(x_0))$$

Taguchis Robustheitsmaß ist dann das zweite Momenten-Maß. Abbildung 4.13 zeigt S/N für die Beispielfunktion.

#### 4.5.10 Entropie

Hurtado [Hur08] schlägt die Entropie von  $Y$  als Robustheitsmaß vor.

$$H_Y(x_0) = - \int p_Y(t) \ln p_Y(t) dt.$$

Eine genauere Untersuchung gibt es hierzu bisher nicht. Seine Berechnung ist aufwändig, da es die Dichtefunktion von  $Y$  benötigt.

## 4.6 Multikriterielle Robustheitsmaße

Multikriterielle Robustheitsmaße benutzen ein Maß  $L(X)$  für die Lage der Verteilung von  $Y$  und eines  $S(X)$  für seine Streuung. Hierfür kann jeweils auf einkriterielle Robustheitsmaße zurückgegriffen werden:

- Für die Lage kommen der nominale Funktionswert  $y$ , sein Erwartungswert  $\mu_Y$  oder der Median  $y_{0,5}$  in Betracht.
- Für die Streuung werden die Varianz  $\text{Var}(Y)$ , die Standardabweichung  $\sigma_Y$  und die Quantildifferenz  $\text{QD}_Y$  benutzt. Möglich sind auch  $F_N$  oder  $H_Y$  aus 4.5.

Wenn die Größenordnung der Zielvariablen verschieden ist, empfiehlt es sich, diese Maße durch Division durch eine Referenzgröße, beispielsweise des absoluten Minimums, in relative Maße umzuwandeln.

Bei Verwendung von  $L(X)$  und  $S(X)$  wird die robuste Optimierungsaufgabe multi-kriteriell wie in Gleichung 4.2 formuliert.

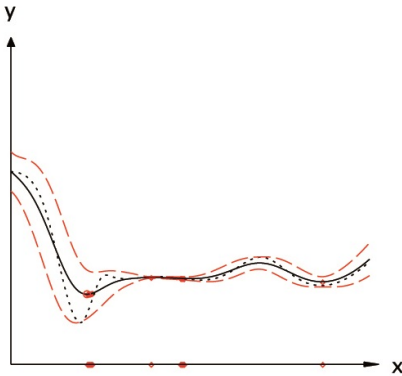


Abb. 4.14: Erwartungswert (durchgezogene Linie),  $f$  gestrichelte Linie,  $\mu \pm \sigma$  (lang gestrichelt) und Pareto-optimale Punkte

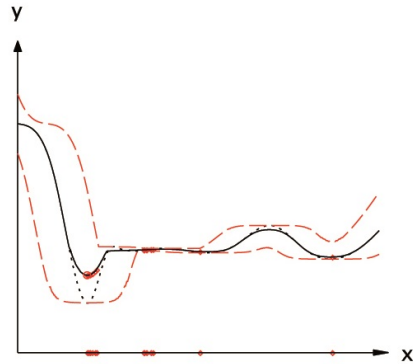


Abb. 4.15: Median (durchgezogene Linie),  $f$  gestrichelte Linie,  $y_{0,05}$  und  $y_{0,95}$  (lang gestrichelt) sowie Pareto-optimale Punkte

### 4.6.1 Erwartungswert und Standardabweichung

Es wird die Paretofront zu

$$\mu_Y(x) = E(Y)$$

als Lageparameter und die Varianz oder die Standardabweichung

$$\sigma_Y^2(x) = \text{Var}(Y) \quad \text{oder} \quad \sigma_Y(x) = \sqrt{\text{Var}(Y)}$$

als Streuungsmaß bestimmt.

Anwendungen dieses Maßes sind zu finden in [Das00], [Jen06], [DK04], [ZFM05] und [CATM96].



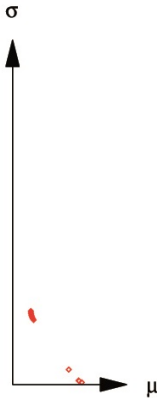


Abb. 4.16: Paretofront für Erwartungswert und Standardabweichung

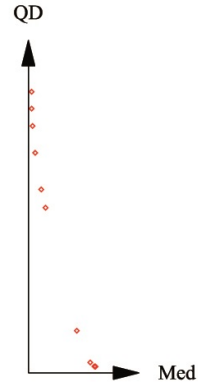


Abb. 4.17: Paretofront für Median und Quantildifferenz

Abbildung 4.14 zeigt für die Beispielfunktion  $\mu_Y$  und  $\mu_Y + / - \sigma_Y$  sowie die Pareto-optimalen Punkte, markiert auf der Erwartungswertfunktion und auf der  $x$ -Achse. Sie sind die besten Kompromisse zwischen Minimierung der Erwartungswertfunktion und Minimierung der Standardabweichungsfunktion. Das absolute Minimum von  $f$  gehört nicht dazu, da es nicht mit dem Minimum der Erwartungswertfunktion übereinstimmt. Im flachen Bereich werden einige Punkte gefunden, ebenso im rechten lokalen Minimum. Die Grenzen  $\mu_Y + / - \sigma_Y$  dürfen nicht so interpretiert werden, dass die übertragene Verteilung an allen Stellen symmetrisch zur Erwartungswertfunktion verläuft. In Abbildung 4.16 ist die zugehörige Paretofront zu sehen.

#### 4.6.2 Median und Quantildifferenz

Wie schon gesehen ist der Median  $y_{0,5}$  ein Maß für die Lage der Verteilung und die Quantildifferenz zweier symmetrische Quantile

$$QD_Y(x_0) = y_q(x_0) - y_{1-q}(x_0)$$

das zugehörige Streuungsmaß.

Dieses Maß ist in der Literatur kaum gebräuchlich. Du [DSC04], Mourelatos [ML06] und Huang [HD07] schlagen die Quantildifferenz für den Bereich der zuverlässigkeitsbasierten Optimierung vor, benutzen aber den Erwartungswert als Lageparameter. In Abbildung 4.15 wird der Median sowie  $y_{0,05}$  und  $y_{0,95}$  gezeigt. Für jeden  $x$ -Wert ist damit eine untere und obere Grenze für die übertragene Verteilung gegeben. Es ist deutlich zu sehen, dass die Intervallbreite stark schwankt und die Verteilung nicht symmetrisch zum Median verläuft. Die zugehörigen Pareto-optimalen Punkte sind auf der Medianfunktion und der  $x$ -Achse markiert. Sie unterscheiden sich von denen für Erwartungswert und Standardabweichung. Das absolute Minimum von  $f$  gehört zu den Pareto-optimalen Punkten. Die Größenangabe für die Streuung ist für beide Maße unterschiedlich. Dies wird in den Abbildung 4.16 und 4.17 deutlich: Sie zeigen

im gleichen Maßstab die Paretofronten für Erwartungswert und Standardabweichung sowie für Median und Quantildifferenz. Aufgrund seiner Konstruktion geben nur die Quantildifferenzen die Größe der wahren Streuung wider.

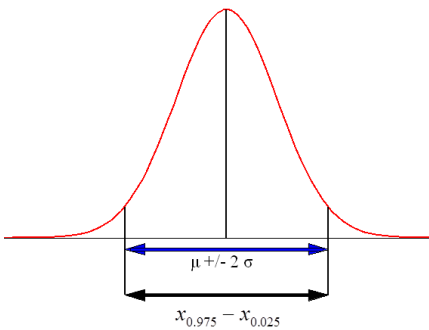


Abb. 4.18: Vergleich von  $\mu + / - 2\sigma$  mit Median und Quantildifferenz für eine Normalverteilung

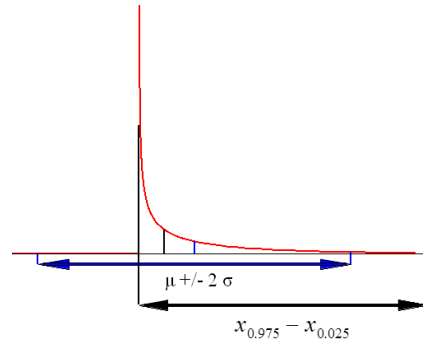


Abb. 4.19: Vergleich von  $\mu + / - 2\sigma$  mit Median und Quantildifferenz für eine schiefe Verteilung

Das Quantilmaß bietet gegenüber Erwartungswert und Standardabweichung einige Vorteile:

- Der Median liegt auf der 50-Prozent-Grenze der Dichtefunktion. Er ist im Gegensatz zum Erwartungswert unempfindlich gegenüber Ausreißern und repräsentiert bei schiefen Verteilungen deren Mitte besser. Diese treten gerade in der Umgebung von lokalen Extrema der Übertragungsfunktion auf (siehe Abb. 3.5).
- Die Quantildifferenz kann direkt interpretiert werden:  $2q$  Prozent der Verteilung liegen außerhalb ihres Bereichs. Das untere und das obere Quantil geben die Grenzen für das Intervall an, in dem sich die mittleren  $1 - 2q$  Prozent der Verteilung befinden. Sie liegen definitionsgemäß immer innerhalb der Verteilung.
- Aus den Maßen Erwartungswert und Standardabweichung können keine Grenzen abgeleitet werden. In der Praxis wird dies dennoch oft getan, in dem man eine Normalverteilung zugrunde legt. Dann entspräche  $\mu - 2\sigma$  ungefähr dem 3%- und  $\mu + 2\sigma$  dem 97%-Quantil. Diese Grenzen sind nicht genau. Besonders bei schiefen Verteilungen geben sie falsche Bereiche an. Die Abbildung 4.18 zeigt, dass die Grenzen abgeleitet aus Quantilen und aus Erwartungswert  $+/- 2 \cdot$  Standardabweichung für Normalverteilungen gleich sind. Als Beispiel für eine schiefe Verteilung ist in Abb. 4.19 die übertragene Verteilung von  $f(x) = x^2$  für  $x = 0$  abgebildet. Die Grenzen für  $\mu + / - 2\sigma$  liegen weit von den tatsächlichen Werten entfernt. Sie umschließen Bereiche, die außerhalb der Verteilung liegen.

Die Unterschiede zwischen den beiden Maßen werden auch in Abbildung 4.20 deutlich: Dort ist als Annäherung an die übertragene Verteilung jeweils die

Häufigkeitsverteilung einer Stichprobe für verschiedene  $x$ -Werte der Beispielfunktion  $f$  abgedruckt. Alle drei Verteilungen sind schief, so dass die Kombination  $\mu + / - 2\sigma$  einen ungeeigneten Bereich angibt, der die tatsächliche Streuung überschätzt.

- Über die Wahl von  $q$  kann der Entscheider bestimmen, wie viel Prozent der Verteilung er bei der Optimierung außen vor lassen will. Er kann die Robustheitsberechnung seinen individuellen Anforderungen an den Grad der Robustheit anpassen.

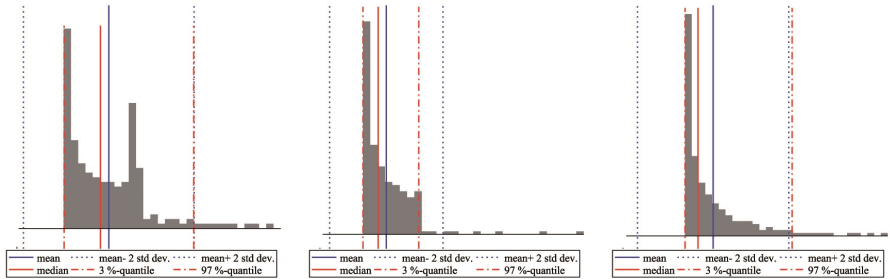


Abb. 4.20: übertragene Verteilung im absoluten Minimum (links), im mittleren flachen Bereich (Mitte) und im größten Pareto-Punkt von  $f$  (rechts)

## 4.7 Diskussion der Robustheitsmaße

Robustheitsmaße charakterisieren die übertragene Verteilung durch ein oder zwei Kennzahlen. Dabei können alle Maße, die für Wahrscheinlichkeitsverteilungen eingeführt wurden, auch hier verwendet werden. Die Aufzählung in den vorangegangenen Unterkapiteln ist daher keineswegs vollständig. Jedes Maß steht dabei für eine andere spezifische Eigenschaft der übertragenen Verteilung. Dies soll am Beispiel der Anforderungen an die Qualität eines Produktes demonstriert werden:

1. Falls die Qualität eines Produktes im Durchschnitt möglichst gut sein soll, sollte das Erwartungswertmaß oder das Medianmaß gewählt werden. Dabei ist aber zu beachten, dass die Streuung nicht betrachtet wird und daher die Qualität stark schwanken kann.
2. Wird dagegen gefordert, dass die Qualität eines Produktes möglichst gleichmäßig sein soll, kann das Standardabweichungsmaß oder die Quantildifferenz benutzt werden. Die mittlere Qualität des Bauteils bleibt unberücksichtigt.
3. Das Worst-Case-Maß wird benutzt, wenn das schlechtest-mögliche Produkt möglichst gut sein soll. Es ist sehr konservativ, da lediglich der größte Schaden betrachtet wird. Die Wahl dieses Maß ist angezeigt, wenn der Worst Case in jedem Fall verhindert werden soll.

4. Falls für einen bestimmten Prozentsatz des Produktes die Qualität möglichst gut sein soll, wird das einseitige Quantilmaß verwendet. Hierdurch werden die negativen Auswirkungen begrenzt.
5. Falls der Verlust an Qualität durch eine Funktion  $U$  ausgedrückt werden kann, ist es günstig, diese zu verwenden. In der Praxis ist es oft schwierig, eine solche Funktion zu bestimmen.

Die Wahl eines geeigneten Robustheitsmaßes hängt also von den spezifischen Anforderungen der Anwendung ab.

In der Einführung dieses Kapitels wurde Robustheit aber schon enger gefasst. Nach dieser Definition soll ein Robustheitsmaß dann als geeignet gelten, wenn es sowohl die Lage als auch die Streuung der übertragenen Verteilung berücksichtigt. Diese enger gefaßte Definition erfüllen eine Reihe der oben aufgeführten Maße nicht:

- Das Erwartungswertmaß und das Medianmaß beziehen die Streuung nicht mit ein.
- Das Worst-Case-Maß, das Signal-Rausch-Verhältnis, das Standardabweichungsmaß und die Quantildifferenz vernachlässigen die Lage der übertragenen Verteilung.
- Das einseitige Quantilmaß und das Wahrscheinlichkeitsschwellenmaß begrenzen nur die negativen Auswirkungen der Streuung. Eine Betrachtung der Lage oder der gesamten Streuung findet nicht statt.

Es bleiben lediglich die gewichtete Zielfunktion aus Lage- und Streuungsmaß sowie die multikriteriellen Robustheitsmaße. Der erste Fall ist wesentlich weniger aufwändig zu lösen, aber der Entscheider muss vor Beginn der Optimierung eine Gewichtung zwischen Lage- und Streuungsmaß vornehmen. Dies ist schwierig. Das Ergebnis ist ein einzelner Punkt. Es bleibt für den Entscheider unklar, ob mit etwas mehr Verzicht an Optimalität viel mehr an Robustheit erreicht werden kann oder umgekehrt. Die Minimierung der multikriteriellen Robustheitsmaße ist sehr aufwändig, da die Paretofront zu bestimmen ist. Dafür erhält der Entscheider wesentlich mehr Informationen: Er kann erkennen, wie „teuer“ er Robustheit erkaufen muss, d.h. wie sehr er vom Minimum der ursprünglichen Optimierungsziele abrücken muss, um Unempfindlichkeit gegenüber Schwankungen der Eingangsvariablen zu erreichen. So kann er verschiedene Kompromisse gegeneinander abwägen. Daher ist dieser Ansatz vorzuziehen.

In Kapitel 6 wird ein Weg beschrieben, wie Paretofronten für das multikriterielle Quantilmaß mit erträglichem Aufwand berechnet werden können.

Tabelle 4.1 fasst die verschiedenen Eigenschaften der Robustheitsmaße noch einmal zusammen.

Das Worst-Case-Maß ist für epistemische Unsicherheiten geeignet, während alle anderen auf aleatorische Unsicherheiten zugeschnitten sind.

## 4.8 Lösung der robusten Optimierungsaufgabe

Im vorigen Abschnitt wurden verschiedene Robustheitsmaße vorgestellt. Falls ein einkriterielles Maß ausgewählt wird und nur ein Zielkriterium vorliegt, bleibt die zu

Tab. 4.1: Übersicht der Robustheitsmaße und ihres Minimierungszieles

Robustheitsmaß	Minimierung
Worst-Case-Maß	des schlimmsten Falles
einseitiges Quantil-Maß	der Streuung „nach oben“
Quantildifferenz-Maß	der Streuung
Wahrscheinlichkeitsschwellen-Maß	der Streuung „nach oben“
Erwartungswert-Maß	des Mittelwertes
Standardabweichung-Maß	der Streuung
gewichtete Zielfunktion	des Mittelwertes und der Streuung
multikriterielle Robustheitsmaße	des Mittelwertes und der Streuung

lösende Optimierungsaufgabe eindimensional. In allen anderen Fällen ist ein multikriterielles Problem zu lösen. Dies kann mit den in Kapitel 3.7 beschriebenen Methoden geschehen. Diese benötigen in der Regel sehr viele Funktionsauswertungen. In dieser Dissertation wird, wie bereits mehrfach erwähnt, angenommen, dass eine Auswertung sehr teuer ist. Ein effizientes Verfahren zu Lösung der Robusten Optimierungsaufgabe für diese Aufgabenstellung wird in Kapitel 6 vorgestellt.

## 4.9 Formulierungen für Nebenbedingungen

Bei den Formulierungen für die Robustheitsmaße wurden die Nebenbedingungen stets außer acht gelassen. Bei einem robusten Design muss aber berücksichtigt werden, wie sich die Variation der Eingangsvariablen auf die Einhaltung der Nebenbedingungen auswirkt. Dies wird in diesem Kapitel betrachtet.

Gleichheitsbedingungen müssen vor der Erweiterung einer deterministischen Optimierungsaufgabe auf eine stochastische in Ungleichheitsbedingungen umformuliert werden, da sie von Zufallsvariablen nicht erfüllt werden können. Man unterscheidet zwischen harten und weichen Ungleichheitsbedingungen [Das00].

Harte Ungleichheitsbedingungen müssen z.B. aus physikalischen Gründen *immer* erfüllt sein. In diesem Fall lautet die stochastische Formulierung:

$$g_i(X) \leq 0, \quad i = 1, \dots, I$$

Eine Überblick über den Stand der Wissenschaft zur Einhaltung von harten Nebenbedingungen gibt [BBC11].

Bei weichen Ungleichheitsbedingungen ist diese Worst-Case-Formulierung zu konservativ. Für diese Fälle kann die Festlegung einer oberen Grenze  $P_{max}$  für die *Ausfallrate (probability of failure)*  $P_f$  benutzt werden [ML06]:

$$P_f = P(g_i(X) \geq 0) < P_{max} \quad i = 1, \dots, I \quad (4.3)$$

Diese Formulierung wird bei der *zuverlässigkeitsbasierten Optimierung* benutzt, die im nächsten Unterkapitel besprochen wird.

Eine andere Möglichkeit der stochastischen Formulierung der Nebenbedingungen besteht darin, eine Kostenfunktion  $C$  zu definieren, die die Kosten für eine Verletzung der Ungleichheitsbedingungen angibt. Die erwarteten Kosten dürfen dann eine

obere Grenze  $C_{max,i}$  nicht überschreiten [Jur07]:

$$E(C(g_i(X))) \leq C_{max,i} \quad i = 1, \dots, I$$

Die Angabe einer geeigneten Kostenfunktion ist schwierig. Falls die Funktion

$$H(x) = \begin{cases} 0 & x \leq 0 \\ 1 & \text{sonst} \end{cases}$$

gewählt wird, erhält man wieder die Gleichung 4.3.

Weit verbreitet ist die Formulierung

$$E(g_i(X)) + c\sigma_{g_i(X)} \leq 0$$

mit  $c$  konstant, z.B.  $c = 3$  [CATM96, DK04]. Hier geht implizit die Annahme ein, dass die  $g_i(X)$  normalverteilt sind und über die Wahl von  $c$  ein Konfidenzbereich gewählt wird.

In den folgenden Kapiteln wird angenommen, dass eine Optimierungsaufgabe ohne Nebenbedingungen vorliegt.

## 4.10 Abgrenzung zu anderen Forschungsgebieten

### 4.10.1 Zuverlässigkeitstheorie

Ein benachbartes Gebiet ist die *Zuverlässigkeitstheorie*. Sie sucht nach einem Design, dessen Ausfallwahrscheinlichkeit möglichst klein ist. Auch hier werden Quantile von übertragenen Verteilungen berechnet. Im Unterschied zum robusten Design, bei dem Quantile bis zu  $q = 0,95$ , bei moderat nichtlinearen Funktionen bis höchstens  $q = 0,99$  berechnet werden, sind die angestrebten Ausfallraten wesentlich kleiner und die Quantile daher wesentlich größer, z.B.  $q = 0,999$ . Eine Übertragung der Methoden führt zu unbefriedigenden Ergebnissen. Ziel ist neben der Minimierung der ursprünglichen Zielkriterien nicht deren Streuung zu verkleinern, sondern die Einhaltung der Nebenbedingungen unter eine Grenze zu drücken. Aus den Nebenbedingungen  $g_i(x) \leq 0, i = 1, \dots, I$  wird

$$P_f = P(g_i(X) > 0) \leq P^U \quad i = 1, \dots, I$$

wobei  $P_f$  die Ausfallwahrscheinlichkeit ist und  $P^U$  die obere zulässige Grenze hierfür. Die  $g_i$  heißen dann Grenzzustandsfunktionen. Einen Überblick geben [Rac01] und [DC00]. Dieses Gebiet zählt nicht zur klassischen robusten Optimierung. Einzelheiten zur Abgrenzung finden sich in [SJ08] und [DK04].

### 4.10.2 Design for Six Sigma

Das *Design for Six Sigma* fordert, die Ausfallwahrscheinlichkeit von Produkten unter die Grenze  $6\sigma$  der Normalverteilung, d.h. unter 0,000002 Prozent zu drücken. Die  $6\sigma$  teilen sich auf in  $4,5\sigma$  Ausfallwahrscheinlichkeit des Produktes, dies entspricht 3,4 Teilen pro einer Million und  $1,5\sigma$  für die Mittelwertverschiebung, die bei länger laufenden Fertigungsprozessen auftritt. Design for Six Sigma ist eine von Motorola eingeführte Methode des Qualitätsmanagements, die alle Teile eines Unternehmens mit verschiedenen Maßnahmen auf das Qualitätsziel ausrichtet. Mathematische Methoden spielen eine untergeordnete Rolle [KYG04].

### 4.10.3 Robustheit von Schätzverfahren

In der Statistik wird der Begriff der Robustheit im Zusammenhang mit Schätzverfahren verwendet. Ein Verfahren heißt robust, wenn es unempfindlich gegenüber kleinen Änderungen der Stichprobenwerte ist [HRRS11]. Insofern liegt hier die gleiche Idee von Robustheit zugrunde.

So ist z.B. der Erwartungswert nicht robust, da er sich durch einzelne Ausreißer stark ändern kann. Der Median dagegen ist ein robustes Maß für die Lage einer Verteilung. Die Quantildifferenz ist in diesem Sinne ebenfalls robuster als die Standardabweichung.





## Kapitel 5

---

# Berechnung von übertragenen Quantilen

---

### 5.1 Vorüberlegungen

Im vorangegangenen Kapitel wurden Robustheitsmaße ausführlich verglichen und eine Empfehlung für das bisher kaum benutzte Quantilmaß ausgesprochen. Ein Grund für seine geringe Verbreitung liegt sicherlich in der aufwändigen Bestimmung. Deswegen werden in diesem Kapitel verschiedene Methoden für seine *punktweise* Berechnung besprochen, d.h. es wird angenommen, dass die übertragenen Quantile für eine feste Kombination der Eingangsvariablen bestimmt werden. Im nächsten Kapitel erfolgt dann die Erweiterung auf den ganzen Designraum.

Dabei wird von folgenden Prämissen ausgegangen:

- Die Methode soll universell einsetzbar sein, sowohl für kleine als auch für hohe Dimensionen.
- Sie soll auch für hochgradig nichtlineare Funktionen und solche, die stark variieren, genaue Ergebnisse liefern.
- Die Methode soll auf Metamodelle angewendet werden.

Hierfür bietet es sich an, eine Sampling-Methode zu verwenden:

- Bei Metamodellen ist eine einzelne Berechnung sehr billig. Daher können die zu erwartenden hohen Stichprobengrößen (vgl. Abschätzung in Abbildung 5.2) verarbeitet werden. Andererseits wird der Ansatz im nächsten Kapitel auf viele Stützstellen erweitert. Ein einfacher Monte-Carlo-Ansatz wird diesen Anforderungen nicht mehr gerecht.
- Auf die Entwicklung von Näherungsverfahren über partielle Ableitungen oder andere dimensionsabhängige Verfahren wird verzichtet, da sie dem „curse of dimensionality“, dem „Fluch der Dimension“ unterliegen.

- Importance sampling-Verfahren können auch bei hohen Dimensionen eingesetzt werden. Hierfür existieren schon theoretische Ansätze zur Bestimmung von Quantilen mit Metamodellen [CGI08]. Die dort beschriebene Vorgehensweise bietet Vorteile bei sehr stark schwingenden Übertragungsfunktionen, deren Schwingung durch ein Metamodell geglättet wird. In solchen Fällen sollte auf dieses Verfahren zurückgegriffen werden.
- Metamodelle liefern stetige Funktionen. In Kapitel 3.1.4 wurde für einfache Funktionen gezeigt, dass die resultierende übertragene Verteilung trotzdem unstetig sein kann. Auch in Anwendungen kann man häufig unstetige oder sehr ungleichmäßige übertragene Verteilungen beobachten. Beispiele hierfür werden in Kapitel 8.2 angeführt. Die Methode sollte auch für solche Fälle gute Ergebnisse liefern.
- Robustheitsmaße charakterisieren eine ganze Verteilung mit wenigen Zahlen. Dies beinhaltet immer einen Informationsverlust. Daher kann es günstig sein, schnell die ganze Verteilung zu berechnen. Dies sollte mit den angestrebten Methode möglich sein.

Wenn Quantile über Stichproben geschätzt werden, lassen sich drei Parameter variieren:

1. die Stichprobengröße
2. die Verteilung der Stichprobenelemente, d.h. der Versuchsplan
3. der Quantilschätzer

Ziel ist es, zu erforschen, wie diese miteinander kombiniert werden sollten, damit mit wenig Aufwand Quantile bis zu einer vorgegebenen Genauigkeit effektiv geschätzt werden können.

Dazu wird in Kapitel 5.2 eine Abschätzung der nötigen Stichprobengröße zur Erreichung einer vorgegebenen Genauigkeit hergeleitet. Kapitel 5.3 beschäftigt sich mit dem Zusammenspiel von Versuchsplan und Quantilschätzer. Beides wird in Kapitel 5.4 zu einem Algorithmus zusammengeführt.

## 5.2 Stichprobengröße

Zur Herleitung einer Formel für die nötige Stichprobengröße des Standardschätzers sind die folgenden Ergebnisse nötig:

Für die asymptotische gemeinsame Verteilung von Quantilen gilt ([DN03], S.288): Sei  $0 < q_1 < q_2 < \dots < q_d < 1$  und  $0 < p(y_{q_j}) < \infty$  für  $j = 1, \dots, d$ . Dann gilt:

$$\sqrt{n}(\hat{Y}_{q_1, n} - y_{q_1}, \dots, \hat{Y}_{q_d, n} - y_{q_d})^T \rightarrow \mathcal{N}(0, \sigma^2) \quad \text{für } n \rightarrow \infty$$

mit Kovarianzmatrix

$$\sigma^2 = \left( \frac{q_j(1 - q_{j'})}{p(y_j)p(y_{j'})} \right)_{1 \leq j \leq j' \leq d} \quad (5.1)$$

Für  $d = 1$  erhält man die asymptotische Varianz des empirischen Schätzers wie in Gleichung 3.3:

$$\sqrt{n}(\hat{Y}_{q,n} - y_q) \rightarrow \mathcal{N}(0, \sigma^2) \quad \text{für } n \rightarrow \infty \quad (5.2)$$

mit

$$\text{AVar}_{\text{EMP}} = \frac{q(1-q)}{p_Y^2(y_q)} = q(1-q) Q'_Y(q)^2$$

Die Konvergenzrate des empirischen Schätzers ist daher  $1/\sqrt{n}$ . Seine Varianz  $\sigma^2$  hängt nicht von der Dimension von  $Y$  ab, sondern vom zu bestimmenden Quantil und der Steigung der Quantilfunktion an dieser Stelle.

$q(1-q)$  besitzt ein Maximum von 0,25 bei  $q = 0,5$  und ist klein für  $q$  nahe 0 oder 1. Der bestimmende Term von  $\text{AVar}_{\text{EMP}}$  ist jedoch  $Q'(q) = 1/p_Y(y_q)$ , das quadratisch in die Formel eingeht. Je größer  $Q'(q)$  ist, desto größer ist der Fehler. In den Abbildungen 3.1 bis 3.5 wurde bereits gezeigt:

- Wenn  $x_0$  in einem monoton steigenden oder fallendem Bereich von  $f$  liegt, ist  $p_Y(y_q)$  klein in den Enden der Verteilung und groß in der Mitte. Für  $q$  nahe 0 oder 1 sind in diesem Fall große Stichproben nötig.
- Wenn  $x_0$  im Bereich eines globalen Extremums  $x_{ex}$  liegt, konzentriert sich die Verteilung von  $Y = f(X_0)$  in der Nähe des Funktionswertes des Extremums  $f(x_{ex})$ . Für eine gute Schätzung der Quantile nahe  $f(x_{ex})$  sind nur kleine Stichproben nötig, sogar wenn  $q$  nahe 0 (bei Maximum) oder 1 (bei Minimum) liegt.
- Bei schwingenden Funktionen kann es sogar vorkommen, dass  $p_Y$  an den Enden größer ist als in der Mitte. In diesem Fall sind in der Mitte größere Stichproben als an den Enden zu benutzen, um die gleiche Genauigkeit zu erzielen.

Die obigen Beispiele gelten für eindimensionale Funktionen. In höheren Dimensionen kommt der zentrale Grenzwertsatz zur Geltung: Die übertragenen Verteilungen werden sich immer mehr einer Normalverteilung annähern je größer die Dimension wird.

Die meisten Verteilungen laufen an ihren Enden flach aus, so dass  $Q'$  hier groß wird. Abbildung 5.1 zeigt die Varianz des Standardschätzers für die Standardnormalverteilung.

Aus Gleichung 5.1 folgt für die asymptotische Verteilung der Quantildifferenz  $\text{QD}_y = y_q - y_{1-q}$  mit  $q \neq 1-q$

$$\sqrt{n}(\hat{Y}_{q,n} - \hat{Y}_{1-q,n} - (y_q - y_{1-q})) \rightarrow \mathcal{N}(0, \sigma_{\hat{Y}_{q,n} - \hat{Y}_{1-q,n}}^2) \quad \text{für } n \rightarrow \infty$$

mit

$$\sigma_{\hat{Y}_{q,n} - \hat{Y}_{1-q,n}}^2 = (1-q) \left( \frac{q}{p^2(y_q)} - \frac{2(1-q)}{p(y_q)p(y_{1-q})} + \frac{q}{p^2(y_{1-q})} \right)$$

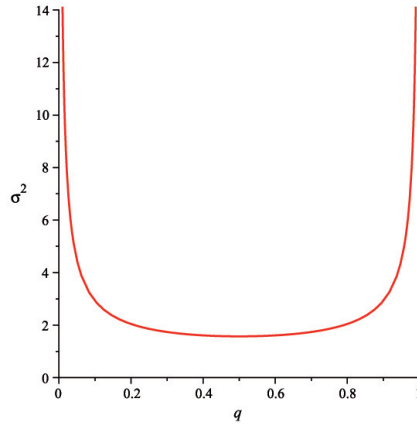
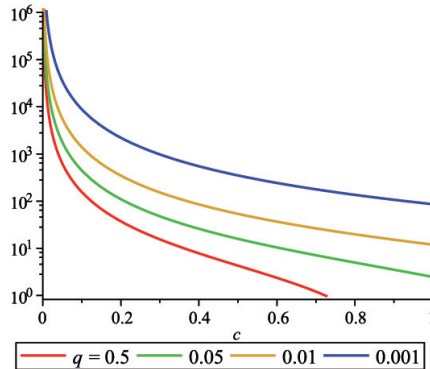


Abb. 5.1: Varianz des Standardschätzers für die Standardnormalverteilung

und für die empirische Quantildifferenz

$$\begin{aligned}
 \text{MSE}(\hat{Y}_{q,n} - \hat{Y}_{1-q,n}) &= \underbrace{(\mathbb{E}(\hat{Y}_{q,n} - \hat{Y}_{1-q,n}) - (y_q - y_{1-q}))^2}_{\mathcal{O}(\frac{1}{n^2}) \text{ vgl. Gleichung 3.4}} \\
 &\quad + \text{Var}(\hat{Y}_{q,n}) - 2 \text{Cov}(\hat{Y}_{1-q,n}, \hat{Y}_{q,n}) + \text{Var}(\hat{Y}_{1-q,n}) \\
 &= \frac{1-q}{n+2} \left( \frac{q}{p_Y^2(y_q)} - \frac{2(1-q)}{p(y_q)p(y_{1-q})} + \frac{q}{p^2(y_{1-q})} \right) + \mathcal{O}\left(\frac{1}{n^2}\right)
 \end{aligned}$$

Abb. 5.2: nötige Stichprobengröße bei der Standardnormalverteilung für verschiedene  $q$ 

Aus Gleichung 3.7 folgt für die *Wurzel aus dem mittleren quadratischen Fehler* (root-

mean-square error (RMSE) :

$$\text{RMSE}(\hat{Y}_{q,n}) \approx \frac{\sqrt{q(1-q)}}{\sqrt{n+2} \cdot p_Y(y_q)} = \frac{\sqrt{q(1-q)}}{\sqrt{n+2}} Q'(q) \quad (5.3)$$

Hieraus erhält man eine Abschätzung für die nötige Stichprobengröße  $n$ , wenn RMSE kleiner als eine vorgegebene Schranke  $c$  ist:

$$\text{RMSE}(\hat{Y}_{q,n}) < c \Leftrightarrow n > \frac{q(1-q)}{p_Y^2(y_q) \cdot c^2} - 2 = \frac{q(1-q)}{c^2} Q'(q)^2 - 2 \quad (5.4)$$

Die Ungleichung 5.4 gilt für zufällige Stichproben. Sie darf nur auf den Monte Carlo-Versuchsplan angewendet werden. Für die anderen Versuchspläne sind keine Abschätzungen bekannt. Aus den Voraussetzungen für die Ausgangsgleichung 3.7 folgt, dass die Quantilfunktionen  $Q$  stetig sein muss.

Beispielhaft werden für die Standardnormalverteilung und  $c = 10\%$  einige Werte berechnet: Die nötige Stichprobengröße  $n$  ist 155 für  $q = 0,5$ , 444 für  $q = 0,05$ , 1391 für  $q = 0,01$  und 8809 für  $q = 0,001$ . Die Stichprobengröße steigt quadratisch in  $c$ . Für  $c = 1\%$  und  $q = 0,5$  ist bereits  $n = 15705$  erforderlich, für  $q = 0,05$   $n = 44654$ , für  $q = 0,01$   $n = 139368$  und schließlich für  $q = 0,001$   $n = 881161$ .

Abbildung 5.2 zeigt die nötige Stichprobengröße für die Schätzung der Standardnormalverteilung in Abhängigkeit von  $c$  für verschiedene  $q$ .

Für die Quantildifferenz gilt:

$$\begin{aligned} \text{MSE}(\hat{Y}_{q,n} - \hat{Y}_{1-q,n}) &\leq \text{MSE}(\hat{Y}_{q,n}) + \text{MSE}(\hat{Y}_{1-q,n}) < 2c^2 \\ \Rightarrow \text{RMSE}(\hat{Y}_{q,n} - \hat{Y}_{1-q,n}) &< \sqrt{2}c \end{aligned}$$

Eine alternative Methode zur Bestimmung eines Quantils bis zu einer vorgegebenen Genauigkeit ist bei stetiger Dichtefunktion auch über die in Kapitel 3.4.2 eingeführten Konfidenzintervalle möglich: Für  $q$  und  $n$  werden zunächst  $m_L$  und  $m_U$  bestimmt. Die Differenz der zugehörigen Ordnungsstatistiken  $y_{(m_L)} - y_{(m_U)}$  ergibt die absolute Breite des Konfidenzintervalls.

$$\frac{y_{(m_L)} - y_{(m_U)}}{2\hat{Y}_{q,n}}$$

ist dann ein Maß für die relative Abweichung vom genauen Wert. Da sich die nötige Stichprobengröße nicht aus dem Konfidenzintervall abschätzen lässt, wird dieser Weg nicht weiter verfolgt.

### 5.3 Versuchsplan und Quantilschätzer

Die andere mögliche Stellschraube für die Berechnung der übertragenen Verteilung ist die Auswahl des Schätzers für die Quantile. In Kapitel 3.3 wurden bereits mehrere nichtparametrische Quantilschätzer vorgestellt, von denen der empirische Standardschätzer asymptotisch die beste Konvergenzrate besitzt. Die dort aufgeführten

theoretischen Ergebnisse sowie die dort zitierten numerischen Untersuchungen zeigen aber, dass Kerndichteschätzer, Harrell-Davis-Schätzer und Sfakianikis-Verginis-Schätzer für endliche Stichproben Vorteile bringen können.

Mit Ausnahme des empirischen Standardschätzers wurden die nichtparametrischen Quantilschätzer bisher nicht für die Bestimmung von übertragenen Quantilen eingesetzt, sondern nur auf Stichproben angewendet, deren Verteilung nicht beeinflusst werden kann.

In diesem Kontext kann aber die Stichprobe  $x_1, \dots, x_n$  gemäß einem Versuchsplan gezogen werden. Aus der Stichprobe  $f(x_1), \dots, f(x_n)$  wird dann das Quantil berechnet.

Daher ist es sinnvoll zu untersuchen, wie sich verschiedene Versuchspläne für  $x$  auf die Genauigkeit der Schätzung von  $f(x)$  auswirken.

Dazu wird in Kapitel 8 eine numerische Untersuchung durchgeführt. Dort wird eine Empfehlung für die Kombination aus Haltonfolge und Harrell-Davis-Schätzer gegeben. Dieses Ergebnis wird benutzt, um einen effizienten Algorithmus für die Schätzung eines übertragenen Quantils bis zu einer vorgegebenen Genauigkeit zu entwickeln:

## 5.4 Algorithmus für die Schätzung des Quantilmaßes

Die Idee des Algorithmus besteht darin, die Abschätzung 5.3 zu nutzen, um die Genauigkeit der Schätzung zu bestimmen und ggf. mit Ungleichung 5.4 die nötige Stichprobengröße zu bestimmen. Das Verfahren ist iterativ. In jedem Iterationsschritt werden die Quantile für  $q$ ,  $0,5$  und  $1 - q$  bestimmt.

Es werden die folgenden Schritte ausgeführt, der Pseudocode ist in Algorithmus 1 angegeben.

Die Berechnung beginnt mit einer kleinen Stichprobe. In den Tests erwies sich  $n = 100$  für  $0,05 < q < 0,95$  als ausreichend. Bei extremeren Quantilen sollte man mindestens mit  $n = 1000$  beginnen.

Zur Quantilschätzung wird eine Stichprobe  $x_1, \dots, x_n$  nach der Haltonfolge verteilt und anschließend die Funktionswerte  $y_1 = f(x_1), \dots, y_n = f(x_n)$  bestimmt. Nun werden die Ordnungsstatistiken  $y_{(1)}, \dots, y_{(n)}$  gebildet, um daraus mit dem Harrell-Davis-Schätzer die Quantile zu schätzen.

Die zur Fehlerschätzung nötige Steigung  $Q'$  wird über den zentralen Differenzenquotienten bestimmt. Hierfür werden zusätzlich die Quantile für  $q - 0,04$ ,  $q + 0,04$ ,  $0,46$ ,  $0,54$ ,  $1 - q - 0,04$  und  $1 - q + 0,04$  berechnet. Dies kostet nur  $12n$  zusätzliche Operationen.  $+/- 0,04$  erwies sich bei Tests als günstig. Bei  $q < 0,05$  wird vom zentralen Quotienten abgewichen.

Nun kann die Steigung von  $Q$  an den drei Quantilwerten abgeschätzt werden. Mit der Formel 5.3 lässt sich der Fehler approximieren. Im Algorithmus wird eine abgewandelte Version benutzt. Wegen  $\sqrt{q(1-q)} < 1$  für  $0 < q < 1$  und  $n + 2 \approx n$  für große  $n$  wird  $1/\sqrt{n} \cdot Q'$  als Schätzung für den Fehler berechnet. Da die Approximation der Steigung für kleine Stichproben schlecht ist, werden Sicherheitsfaktoren benutzt, die mit steigenden Iterationsschritten sinken.

Liegt die Fehlerschätzung über der vom Benutzer vorgegebenen Grenze  $c$ , wird die nötige Stichprobengröße  $n_{nec}$  berechnet. Sie entspricht dem Vierfachen der nach

Formel 5.4 erforderlichen Größe. Um zu vermeiden, die Gewichte für den Harrell-Davis-Schätzer jeweils neu zu berechnen, werden diese für bestimmte Stichprobengrößen vorher berechnet und einer Datei gespeichert.  $n_{nec}$  wird auf die nächstgrößere abgespeicherte Zahl umgesetzt. Die Haltonfolge wird um die zusätzlichen Elemente ergänzt und die Quantilschätzung startet von vorne.

In Kapitel 8.5 wird verifiziert, dass der Algorithmus die Quantile bis zur gegebenen Genauigkeit schätzt und die Stichprobengröße gemäß dem lokalen Verhalten der Quantilfunktion anpasst.

Die Abschätzung 3.7 gilt nur, falls die Dichtefunktion an dieser Stelle stetig differenzierbar ist. Diese Einschränkung kann nicht umgangen werden. Der vorgegebene Schätzfehler kann an Stellen, an denen die Dichtefunktion nicht stetig differenzierbar ist, möglicherweise nicht eingehalten werden. Die Abschätzung wurde nur für den Monte Carlo Versuchsplan bewiesen. Die numerische Untersuchung rechtfertigt aber den Einsatz des Halton-Versuchsplans, der dort bessere Ergebnisse als MC erzielt.

```

1 Algorithm Berechne  $Q(q)$  ()
   input :  $qmenge = \{0.05, 0.5, 0.95\}$ 
   input :  $c$ 
   output:  $Q(q)$ 

2    $n_{Anfang} = 0;$ 
3    $n_{Ende} = 100;$ 
4    $AnzIt = 0;$ 
5    $qmengeneu = \{qmenge[i] - 0.04, qmenge[i], qmenge[i] + 0.04 \text{ für } i = 1, 2, 3\}$ 
   (für Schätzung von  $Q'$ );
6   while  $AnzIt < 4$  do
7     Bestimme die Stichprobenelemente von  $n_{Anfang}$  bis  $n_{Ende}$  gemäß
     Haltonfolge;
8     Bestimme die Funktionswerte hierzu;
9     Ordne die gesamte Stichprobe;
10    Schätze alle Quantile aus  $qmengeneu$  mit dem Harrell-Davis-Schätzer;
11    Schätze  $Q'$  für alle Elemente aus  $qmenge$  über den Differenzenquotient

           
$$Q'(qmenge[i]) = \frac{Q(qmenge[i] + 0.04) - Q(qmenge[i] - 0.04)}{0.08}$$


     für  $i = 1, 2, 3;$ 
12    if  $AnzIt = 0$  then
13      |  $error[i] = 6 \frac{1}{\sqrt{n_{Ende}}} Q'(qmenge[i])$  für  $i = 1, 2, 3$  (Sicherheitsfaktor 6)
14    else if  $AnzIt = 1$  then
15      |  $error[i] = 2 \frac{1}{\sqrt{n_{Ende}}} Q'(qmenge[i])$  für  $i = 1, 2, 3$  (Sicherheitsfaktor 2)
16    else
17      |  $error[i] = \frac{1}{\sqrt{n_{Ende}}} Q'(qmenge[i])$  für  $i = 1, 2, 3$  (ab der dritten
      | Iteration kein Sicherheitsfaktor mehr)
18    end
19    if  $error[i] > c$  then
20      |
           
$$n_{nec} = \frac{4q(1-q)}{c^2} Q'(qmenge[i])^2 \text{ für } i = 1, 2, 3$$

21    else
22      | fertig
23    end
24     $n_{Anfang} = n_{Ende};$ 
25    setze  $n_{nec}$  auf die nächstgrößere Stichprobengröße um, für die die
     Gewichte des Harrell-Davis-Schätzers abgespeichert sind;
26     $n_{Ende} = n_{nec};$ 
27     $AnzIt = AnzIt + 1$ 
28  end

```

**Algorithm 1:** Algorithmus zur Bestimmung der Quantile aus  $qmenge$



## Kapitel 6

---

# Robuste Optimierung mit Metamodellen

---

Robuste Optimierung erfordert viele Auswertungen der Funktion  $f$ . In Ungleichung 5.4 wurde eine Abschätzung der nötigen Stichprobengröße für die Quantilberechnung an nur einer Stelle der Funktion angegeben. Ein direkter Optimierungsalgorithmus würde solche Berechnungen an vielen Stellen erfordern. Da in dieser Dissertation angenommen wird, dass eine Funktionsauswertung teuer ist, verbietet sich eine direkte Robustheitsberechnung auf  $f$ . Metamodelle bieten einen naheliegenden Ausweg. Ihre Integration in eine robuste Optimierungsprozedur ist auf mehrere Arten möglich:

- Sie werden lediglich als Ersatz für die Funktion  $f$  benutzt. Anschließend werden die robusten Optima mit Suchalgorithmen bestimmt.  
In der Literatur finden sich mehrere Beispiele für diesen Ansatz: Jin [JDC03] vergleicht verschiedene Metamodelle und verschiedene Formulierungen für eine gewichtete Zielfunktion aus Erwartungswert und Standardabweichung. In [Jin04] gibt er analytische Formeln für Erwartungswert und Varianz der übertragenen Verteilung bei verschiedenen Metamodellarten an. Da der Berechnungsaufwand hierfür sehr hoch ist, eignen sie sich nicht für eine effiziente Anwendung. Shimoyama [SLJ<sup>+</sup>09] benutzt ein Kriging-Metamodell für das Systemverhalten. Als multikriterielles Robustheitsmaß wird Erwartungswert und Standardabweichung verwendet. Die Paretofront wird mit evolutionären Mehrziel-Algorithmen bestimmt.
- In anderen Ansätzen geht der Einsatz von Metamodellen weiter: Sie modellieren neben  $f$  auch das Robustheitsverhalten.  
Vining [VM90] und Chen [CATM96] erstellen zuerst ein Metamodell für  $f$  und anschließend je eines für Erwartungswert und Standardabweichung der übertragenen Verteilung. Die Bestimmung des robusten Optimums erfolgt über die Paretofront. Lee [LP06] erstellt ein Metamodell für das Systemverhalten. Hieraus wird mittels Monte Carlo Simulationen die Standardabweichung der Ausgabevariablen bestimmt und ein weiteres Metamodell gebildet. Die multikriterielle Optimierungsaufgabe wird in eine einkriterielle verwandelt, in dem die Standardabweichung minimiert wird und für den Funktionswert eine obere

Grenze einzuhalten ist. Dellino [DKM12] verwendet für ein eindimensionales Beispiel ein Kriging-Metamodell und benutzt dieses, um punktweise Erwartungswert und Standardabweichung der übertragenen Verteilung zu schätzen und hieraus wiederum Metamodelle zu erstellen. Sie bestimmt anschließend die Paretofront hierzu.

Dieser zweite Ansatz wird in der folgenden Methodik aufgegriffen und auf das Quantilmaß übertragen. Die Berechnung von Quantilen auf Metamodellen ist kaum untersucht. In [Sud12] ist eine Methode zur Bestimmung der Ausfallwahrscheinlichkeit für Kriging-Metamodelle beschrieben. Ihre Konvergenzgeschwindigkeit entspricht der der Monte-Carlo-Simulation.

Mit Metamodellen lässt sich das Robustheitsverhalten für den ganzen Designraum beschreiben. Da sie sehr schnell ausgewertet werden können, ist es möglich, auch aufwändige Berechnungen durchzuführen, wie sie für die Bestimmung der übertragenen Quantile und der Paretofront nötig sind. Es müssen keine speziellen Algorithmen entwickelt werden, sondern die in Kapitel 3.7.1 beschriebenen Verfahren können benutzt werden.

## 6.1 Ablauf der entwickelten Methodik

**Im folgenden wird von einer Optimierung ohne Nebenbedingungen ausgegangen.**

Wie bereits in der Einleitung skizziert, verläuft die Methodik zur robusten Optimierung mit Metamodellen mit Hilfe von Quantilmaßen in vier Schritten:

1. Im ersten Schritt wird das Systemverhalten  $f$  durch ein Metamodell approximiert.
2. Im zweiten Schritt werden hiermit Metamodelle für den übertragenen Median und die Quantildifferenz erstellt.
3. Diese werden im dritten Schritt benutzt, um die Paretofront der robusten Optima zu bestimmen.
4. Im letzten Schritt werden die robusten Optima über ein mehrstufiges Verfahren visualisiert.

Diese Methodik macht es möglich, aus wenigen Stützstellen das Systemverhalten und das Robustheitsverhalten des Systems für den ganzen Designraum zu modellieren.

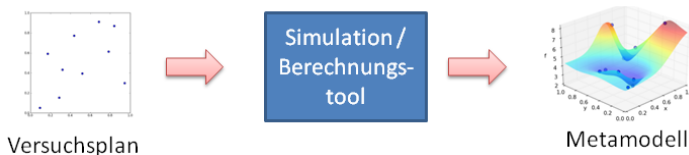


Abb. 6.1: Schritt 1: Metamodell für  $f$

### 6.1.1 Schritt 1: Metamodell für $f$ (Abb. 6.1)

Für eine Anfangsmenge von Stützstellen  $(x_1, \dots, x_n)$  werden über Computerexperimente die Funktionswerte  $(f(x_1), \dots, f(x_n)) = (y_1, \dots, y_n)$  ermittelt. Die Verteilung der Stützstellen sollte dabei nach einem geeigneten Versuchsplan wie Latin Hypercube Sampling oder Haltonfolge erfolgen.

Anschließend wird aus den Wertepaaren  $((x_1, y_1), \dots, (x_n, y_n))$  ein Metamodell  $\tilde{f}$  für die Originalfunktion  $f$  erstellt. Hierzu ist das Radiale Basisfunktionen-Modell mit polynomieller Anteil passend. Das Metamodell wird validiert und nötigenfalls um zusätzliche Stützstellen ergänzt. Im folgenden wird angenommen, dass schon der erste Stützstellensatz den Genauigkeitsanforderungen des Benutzers genügt.

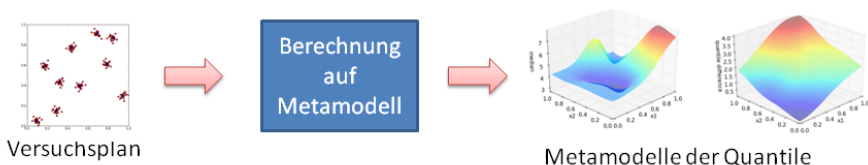


Abb. 6.2: Schritt 2: Robustheitsberechnung

### 6.1.2 Schritt 2: Metamodelle für den Median und die Quantildifferenz (Abb. 6.2)

Zur Quantifizierung der Robustheit wird das multikriterielle Quantilmaß benutzt. An den Funktionswerten der Stützstellen werden ein unteres Quantil  $y_l(x_i)$ , der Median  $y_m(x_i)$  und ein oberes Quantil  $y_u(x_i)$   $i = 1, \dots, n$  mit dem Algorithmus 1 geschätzt, d.h. es wird um jede Stützstelle eine Stichprobe gezogen und diese ggf. erhöht bis für jede Stützstelle die vorher festgelegte Genauigkeit in der Schätzung erreicht ist. Es werden die Stützstellen benutzt, da man annehmen kann, dass das Metamodell in der Nähe dieser Punkte am besten die wahre Funktion approximiert. Die berechneten Werte dienen als Stützstellen für zwei neue Metamodelle: Aus den  $(x_1, \dots, x_n)$  und  $(y_u(x_1) - y_l(x_1), \dots, y_u(x_n) - y_l(x_n))$  wird ein Metamodell  $\tilde{f}_{qd}$  für die Quantildifferenz der übertragenen Verteilung für den gesamten Designraum erstellt. Analog sind die Medianwerte die Stützstellen für ein Metamodell  $\tilde{f}_m$  für den Median. Mit diesen ist es jetzt möglich, sehr schnell Median und Quantildifferenz im gesamten Designraum zu berechnen.

Auch diese Metamodelle müssen validiert werden. Falls eines nicht genau genug ist, können auf zwei Arten Stützstellen ergänzt werden:

- Es wird zu Schritt 1 zurückgekehrt. Im ursprünglichen Metamodell werden neue Punkte durch zusätzliche Computerexperimente ergänzt. Dieser Weg ist aufwändig, würde aber die Qualität aller Metamodelle erhöhen.
- In diesem Schritt werden Punkte hinzu genommen, deren Funktionswerte nicht über  $f$  sondern über das Metamodell  $\tilde{f}$  bestimmt werden. Diese Vorgehensweise ist sehr effizient. Ob ein Nachteil in der Genauigkeit besteht, kann pauschal nicht vorhergesagt werden, da die Genauigkeit der Quantilschätzung von der

Güte der Approximation über den gesamten Bereich der Eingangsverteilung abhängt.

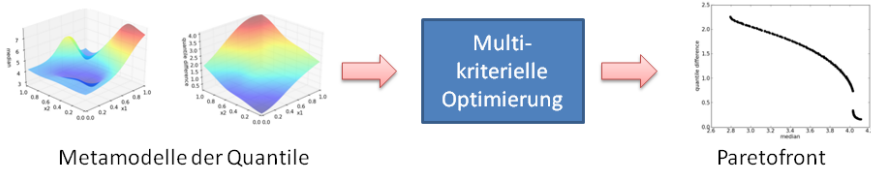


Abb. 6.3: Schritt 3: Bestimmung der robusten Optima

### 6.1.3 Schritt 3: Bestimmung der robusten Optima (Abb. 6.3)

Die robuste Optimierungsaufgabe lautet nun

$$\text{Minimiere } \tilde{f}_m(y) \quad \text{und} \quad \tilde{f}_u(y) - \tilde{f}_l(y) \quad (6.1)$$

wobei  $\tilde{f}_m(y)$  den Mittelwert der übertragenen Verteilung modelliert und  $\tilde{f}_u(y) - \tilde{f}_l(y)$  die Streuung.

Da beide Zielfunktionen Metamodelle sind, die den gesamten Designraum abdecken und deren Auswertung sehr billig ist, können zur Lösung des Optimierungsproblems Standard-Algorithmen zur Paretooptimierung eingesetzt werden.

Ist der Zielraum mehrdimensional wird je Dimension ein Ausgangsmetamodell generiert, aus dem dann mit den beschriebenen Schritten je Dimension zwei Quantil-Metamodelle erstellt werden. Bei der robusten Optimierungsaufgabe erhöht sich entsprechend die Anzahl der Zielfunktionen. Es bleibt aber bei der Anwendbarkeit der Pareto-Standard-Algorithmen.

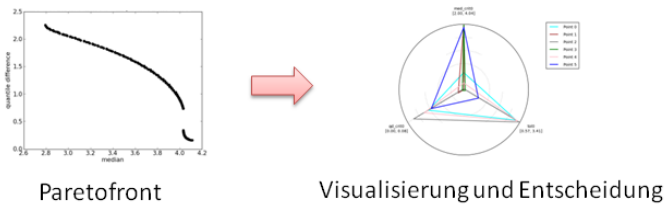


Abb. 6.4: Schritt 4: Visualisierung der Paretofront und Entscheidung

### 6.1.4 Schritt 4: Visualisierung der Paretofront und Entscheidung (Abb. 6.4)

In Kapitel 3.7.2 wurde bereits die Visualisierung von Paretofronten mit Sterndiagrammen vorgestellt. Sie ist mit einigen Erweiterungen auf die hier berechneten robusten Optima übertragbar.

Bei der Verwendung von Metamodellen kommt der Aspekt ihrer Genauigkeit hinzu. Der Entscheider sollte abschätzen können, wie groß der Fehler seiner Modelle in den robusten Optima ist. Bei zwei annähernd gleichen Lösungen sollte er sich für die mit der größeren Modellgenauigkeit entscheiden. Diese zusätzliche Information lässt sich leicht in Sterndiagramme als weitere Spitzen integrieren. Sie wird nur informativ angegeben und ist nicht Teil der Optimierungsprozedur. Angegeben wird die Toleranz des Ursprungsmetamodells  $\hat{f}$ . Alternativ könnte auch die der Quantilmetamodelle einbezogen werden. Im folgenden Unterkapitel wird zudem eine Vorgehensweise zur Bestimmung der Toleranz der Quantilmetamodelle basierend auf Kreuzvalidierung beschrieben. Verwendet man eine dieser Möglichkeiten würde sich die Anzahl der Spitzen weiter erhöhen und die Analyse der Diagramme erschweren. Testanalysen haben gezeigt, dass sich die Toleranz des ursprünglichen Metamodells und die der Quantilmetamodelle ähnlich verhalten. Darüber hinaus würde sich der Berechnungsaufwand erhöhen.

Die Skalierung der Achsen des Sterndiagramms erfolgt mit absoluten Werten, da diese durch den Entscheider direkt interpretiert werden können. Um der unterschiedlichen Größenordnung der einzelnen Kriterien und ihrer Schwankungsbreite Rechnung zu tragen, wird für jede Achse der kleinste Wert aller robusten Optima dieses Kriteriums in den Mittelpunkt gelegt, der größte Wert entspricht der Spitze. Diese Extrema werden neben jeder Achse angegeben.

Die Paretofront besteht bei robusten Optimierungsproblemen meist aus sehr vielen Punkten. Würde man sie in einem einzigen Sterndiagramm darstellen, wäre es zu unübersichtlich, denn diese Art der Visualisierung eignet sich nur für wenige Alternativen. Deswegen wird ein sequentielles Verfahren zur Auswahl eines robusten Optimums benutzt:

Die Paretofront wird mit dem im Kapitel 3.8 vorgestellten  $k$ -means-Algorithmus geclustert und nur die Clusterzentren im Sterndiagramm angezeigt. Der Entscheider kann nun zwischen wenigen „grundsätzlichen Lösungen“ auswählen. Wenn er sich auf eine Lösung festgelegt hat, werden ihm im nächsten Schritt alle Elemente, die zu diesem Clusterzentrum gehören, angezeigt. Sind auch dies zu viele, werden diese zuerst geclustert. So kann er seine Auswahl schrittweise eingrenzen und schließlich eine Feinauswahl treffen. Die Einzelheiten können Algorithmus 2 entnommen werden.

Danach werden die Parameterwerte für die Eingabevariablen zur gewählten Kombination ausgegeben. Es ist ratsam, mit ihnen eine abschließende Funktionsauswertung durchzuführen.

Diese Vorgehensweise ermöglicht es dem Benutzer, aus dem gesamten Lösungsraum auszuwählen, ohne vorher Prioritäten festlegen zu müssen. Bei der Visualisierung repräsentieren im ersten Bild die Clusterzentren einen weiten Bereich der möglichen robusten Optima. Insbesondere einzelne isolierte Lösungen sind dort nicht sichtbar. Auf diese stößt der Benutzer erst auf den letzten Auswahlbildern. Dies stellt keinen Nachteil dar, da es ratsam ist, das robuste Optimum in einem größeren robusten Bereich zu wählen. Diese werden durch die Clusterung adäquat dargestellt.

Die in den Sterndiagrammen dargestellten Clusterzentren gehören nicht zu den Paretopunkten. Deswegen werden die später ausgewählten robusten Optima bessere Optimalitätseigenschaften haben als in den ersten Bildern angezeigt. Die erreichbaren Wertebereiche lassen sich aber schon an den angezeigten Grenzen ablesen. Dieser Umstand sollte dem Benutzer bewusst sein.

```

1 Algorithm Wähle robustes Optimum aus()
   input : punkte = Paretofront
   input : n = Anzahl der robusten Optima
2   Min. und Max. für jedes Kriterium der Paretofront bestimmen (für
   Achsenskalierung und -beschriftung);
3   maxAnzCl = 6 (maximale Anzahl der Alternativen, die im
   Sterndiagramm angezeigt werden sollen);
4   while n > maxAnzCl do
5     Bestimme die Clusterzentren cz aus punkte;
6     Bestimme die Modelltoleranzen zu cz;
7     Gebe Sterndiagramm für cz aus (mit Modelltoleranzen und skaliert
   mit Min. und Max.);
8     Entscheider wählt Alternative i aus;
9     punkte = Elemente in Cluster i;
10    n = Anzahl der Elemente in Cluster i;
11  end

```

**Algorithm 2:** Auswahl des robusten Optimums

## 6.2 Einfluss der Metamodelltoleranz auf die Quantilschätzung

In Schritt 2 der Methodik werden Metamodelle für Median und Quantildifferenz der übertragenen Verteilung gebildet. Die Werte der Stützstellen beruhen auf Schätzungen der übertragenen Quantile auf Grundlage des Metamodells für  $f$ . Es stellt sich nun die Frage, ob aus der Modelltoleranz von  $\tilde{f}$  Schlüsse gezogen werden können bezüglich der Genauigkeit oder der Modelltoleranz der Modelle für  $\tilde{f}_m$  und  $\tilde{f}_u - \tilde{f}_l$ . Analog zum im Kapitel 3.5.2 beschriebenen Kreuzvalidierungsverfahren für die Metamodelltoleranz kann man eines zur Schätzung der Quantiltoleranz gewinnen:

Im Ausgangsmetamodell  $\tilde{f}$  wird eine Stützstelle  $(x_i, y_i)$  eliminiert und hieraus ein Metamodell  $\tilde{f}_{\text{ohne } x_i}$  erstellt. Anschließend werden die übertragenen Quantile auf der Grundlage von  $\tilde{f}_{\text{ohne } x_i}$  bestimmt. Dieser Vorgang wird für alle Stützstellen  $i = 1, \dots, n$  durchgeführt, so dass man  $n$  Quantilschätzungen für jede Stützstelle erhält. Hiervon wird die größte als obere Toleranzgrenze verwendet, die kleinste als untere. Um auch Toleranzgrenzen für alle anderen Punkte des Metamodells zu erhalten, wird jeweils aus den unteren und oberen Grenzen ein Metamodell gebildet. Diese Vorgehensweise wird in Algorithmus 3 wiedergegeben.

Die Toleranzgrenzen müssen den wahren Wert nicht einschließen. Sie sind vielmehr analog zur Modelltoleranz zu interpretieren: Große Toleranzbereiche zeigen an, dass der geschätzte Wert möglicherweise einen großen Fehler aufweist. Enge Grenzen weisen auf genaue Werte hin.

In Kapitel 10 wird dieser Ansatz zur Angabe von Toleranzgrenzen der Quantilmotamodelle numerisch untersucht. In der Literatur gibt es bisher keinen weiteren.

Eine ähnliche Vorgehensweise beschreibt Dellino [DKM12] für Konfidenzgrenzen von Metamodellen für Erwartungswert und Standardabweichung:

Eine Menge von Stützstellen dient als Grundlage für ein Kriging-Metamodell. Anschließend wird eine größere Stichprobe um jede Stützstelle gemäß der Verteilung der unsicheren Eingabeparameter gezogen und hieraus Erwartungswert und Standardabweichung geschätzt. Aus den approximierten Werten wird dann die Paretofront

```

1 Algorithm Bestimme Toleranzgrenzen()
   input : Stützstellen  $(x_i, y_i), i = 1, \dots, n$ 
   output: Metamodelle der Toleranzgrenzen für das Median- und das
           Quantildifferenz-Metamodell
2   Bestimme Metamodell  $\tilde{f}$ ;
3   for  $k = 1, \dots, n$  do
4     Streiche  $(x_k, y_k)$  aus der Menge der Stützstellen;
5     Bestimme aus den verbliebenen Stützstellen ein Metamodell
        $\tilde{f}_{\text{ohne } x_k}$ ;
6     Bestimme die übertragenen Quantile
        $y_{0,5j}^{\text{ohne } x_k}, y_{qj}^{\text{ohne } x_k}, y_{1-qj}^{\text{ohne } x_k} j = 1, \dots, n$ ;
7   end
8   for  $j = 1, \dots, n$  do
9     Bestimme  $y_{0,5j}^{\max} = \max(y_{0,5j}^{\text{ohne } x_k} \text{ für } k = 1, \dots, n)$ ;
10    Bestimme  $y_{0,5j}^{\min} = \min(y_{0,5j}^{\text{ohne } x_k} \text{ für } k = 1, \dots, n)$ ;
11    Bestimme  $y_{qj}^{\max} = \max(y_{qj}^{\text{ohne } x_k} \text{ für } k = 1, \dots, n)$ ;
12    Bestimme  $y_{qj}^{\min} = \min(y_{qj}^{\text{ohne } x_k} \text{ für } k = 1, \dots, n)$ ;
13    Bestimme  $y_{1-qj}^{\max} = \max(y_{1-qj}^{\text{ohne } x_k} \text{ für } k = 1, \dots, n)$ ;
14    Bestimme  $y_{1-qj}^{\min} = \min(y_{1-qj}^{\text{ohne } x_k} \text{ für } k = 1, \dots, n)$ ;
15    Bestimme  $QD_{qj}^{\max} = y_{1-qj}^{\max} - y_{qj}^{\min}$ ;
16    Bestimme  $QD_{qj}^{\min} = y_{1-qj}^{\min} - y_{qj}^{\max}$ ;
17  end
18  Bestimme das Metamodell  $\tilde{f}_m^{\max}$  aus  $y_{0,5j}^{\max} j = 1, \dots, n$ ;
19  Bestimme das Metamodell  $\tilde{f}_m^{\min}$  aus  $y_{0,5j}^{\min} j = 1, \dots, n$ ;
20  Bestimme das Metamodell  $\tilde{f}_{qd}^{\max}$  aus  $QD_{qj}^{\max} j = 1, \dots, n$ ;
21  Bestimme das Metamodell  $\tilde{f}_{qd}^{\min}$  aus  $QD_{qj}^{\min} j = 1, \dots, n$ ;

```

**Algorithm 3:** Toleranzgrenzen für die übertragenen Quantile

bestimmt. Dieser Ansatz erwies sich in ihrem Artikel zwar als weniger genau als der Ansatz, Erwartungswert und Standardabweichung direkt aus einer Stichprobe zu schätzen. Dies lag an einer zu kleinen Stichprobe für den unsicheren Eingabeparameter (dort  $n = 100$ ). In Kapitel 10 wird die gleiche Beispielfunktion benutzt. Ein RBF-Metamodell mit den gleichen Stützstellen liefert bei  $n = 10000$  ein nahezu exaktes Ergebnis.

Zur Überprüfung der Güte des Metamodells verwendet Dellino Kreuzvalidierung. Allerdings werden keine Maße angegeben, die Entscheidung über die Güte erfolgt visuell. Ein Vertrauensbereich für einzelne Punkte der Paretofront wird über nicht-parametrisches verteilungsfreies bootstrapping gewonnen:

Aus der Menge der Stützstellen wird durch Ziehen mit Zurücklegen eine neue Stützstellenmengen gezogen. Mit ihr wird die oben beschriebene Vorgehensweise erneut durchgeführt, so dass eine zweite Paretofront entsteht. Dieser Vorgang wird wiederholt durchgeführt. Am Ende entsteht eine Menge von Paretofronten. Eine Konfidenzregion für einen Punkt der Paretofront wird gewonnen, in dem für die Menge der

zugehörigen bootstrapped Paretopunkte symmetrische Quantile geschätzt werden. Da beim bootstrapping sehr viele Wiederholungen nötig sind, ist dieses Verfahren wesentlich aufwändiger als die oben beschriebene Vorgehensweise. Es kann nicht garantieren, dass der wahre Paretopunkt innerhalb der Konfidenzregion liegt.

### 6.3 Diskussion

In diesem Kapitel wurde eine durchgängige Methodik für die robuste Optimierung bei wenigen Funktionsauswertungen vorgestellt. Das Verfahren benötigt als Eingabe nur die folgenden Minimalangaben: die Eingabevariablen, ihre Wertebereiche und ihr Schwankungsverhalten sowie die Ausgabevariablen. Der Anwender gibt zudem an, wie viele echte Funktionsauswertungen er durchführen möchte. Über die Wahl von  $q$  kann er den Grad der Robustheit bestimmen. Diesen kann er auch nachträglich variieren. Dann sind keine zusätzlichen Funktionsauswertungen sondern nur erneute Berechnungen auf dem Metamodell nötig.

Die Methodik beschränkt sich nicht nur auf normalverteilte Eingangsverteilungen, sondern sie kann analog auch auf andere Eingangsverteilungen angewendet werden. Die Stichproben um die Stützstellen müssen dann nur gemäß der Eingangsverteilung gezogen werden.

Nach diesen Anfangseingaben läuft der Rest der Methodik automatisch bis zur Entscheidungsfindung ab. Es wird eine umfassende robuste Optimierung durchgeführt, die die *globale* Paretofront bestimmt im Gegensatz zu anderen Suchverfahren, die in lokalen Fronten stecken bleiben können.

Das Verfahren eröffnet dem Anwender die Möglichkeit, umfassend die besten Kompromisse gegeneinander abzuwägen. Dies ist wichtig bei Optimierungen, bei denen vorher noch keine Vorstellung davon vorhanden ist, welche Kompromisse überhaupt möglich sind, etwa bei der Dimensionierung von Bauteilen. Dort müssen oft noch zusätzliche Restriktionen berücksichtigt werden, die sich nicht oder nur sehr aufwändig mathematisch ausdrücken lassen, etwa die Einschränkungen durch bestimmte Fertigungsmöglichkeiten.

Durch die Benutzung von Quantilen können die Schwankungsbereiche für jede Zielgröße genau angegeben werden. Auch dies ist in der Praxis sehr erwünscht.

Die Genauigkeit der Berechnungen hängt entscheidend von der Güte des Ursprungsmetamodells ab, da dieser Fehler auf die Quantilschätzung übertragen wird. Daher sollte  $\tilde{f}$  sorgfältig validiert werden und ggf. durch zusätzliche Stützstellen verbessert werden. Über die Berechnung der Modelltoleranz für die Quantilmetamodelle bzw. über die in Kapitel 6.2 beschriebene Kreuzvalidierungsprozedur können ungenaue Bereiche der Quantilschätzung identifiziert werden. Mögliche Wege zur Verbesserung ihre Genauigkeit wurden bereits angesprochen: Falls  $\tilde{f}$  genau genug ist, genügt es, für zusätzliche Punkte die übertragenen Quantile zu bestimmen und diese den Stützstellen für die Quantilmetamodelle hinzu zu fügen. Ein echtes Fehlermaß existiert für Metamodelle jedoch nicht. Dies spricht nicht gegen die Anwendung von Metamodellen: Falls einzelne Funktionsauswertungen aufwändig sind, sind sie das Mittel der Wahl. Bei mäßig nichtlineares Systemverhalten führen sie zu befriedigenden Ergebnissen. Die Verwendung von Quantilmetamodellen ist nur die konsequente Fortsetzung dieses Weges.



Die vorgestellte Methodik ist sehr effizient, da sie nur eine minimale Anzahl von Funktionsauswertungen benötigt. Die anschließenden Berechnungen auf den Meta-modellen sind zwar umfangreich, aber im Vergleich zu einer Funktionsauswertung im allgemeinen wesentlich schneller. Die Anzahl der Berechnungen auf den Meta-modellen wurde im Vergleich zum Standardansatz an mehreren Stellen wesentlich reduziert:

- An jeder Stützstelle wird für die Bestimmung der Quantile nur die lokal nötige Stichprobengröße benutzt.
- Die numerische Untersuchung in Kapitel 8 wird zeigen, dass der Harrell-Davis-Schätzer dem empirischen Standardschätzer im Bereich der verwendeten Stichprobengrößen überlegen ist.
- Die Haltonfolge reduziert im Vergleich zum Monte Carlo-Verfahren die nötige Stichprobengröße bei kleinen Dimensionen deutlich. Sie ist aber ebenso nachträglich ergänzbar, so dass bei Erhöhung der Stichprobengröße nicht eine neue Stichprobe gezogen werden muss.

Im folgenden Kapitel wird zudem eine Möglichkeit der weiteren Beschleunigung vorgestellt.

Ein zusätzlicher Zeitgewinn lässt sich durch die Parallelisierung der Methodik erzielen:

- Bei der Bestimmung der Stützstellen können die Funktionsauswertungen unabhängig voneinander durchgeführt werden.
- Die Quantile für jede Stützstelle können parallel berechnet werden.



## Kapitel 7

---

# Aufwandsabschätzung für das schnelle Multipolverfahren

---

Die im vorigen Kapitel eingeführte Methodik führt sehr viele Auswertungen des Metamodells zur Berechnung der übertragenen Quantile durch. Bei dem hier benutzten multiquadratischen Radiale Basisfunktionen-Modell beträgt der Aufwand für die  $m$ -fache Auswertung mit  $n$  Stützstellen  $\mathcal{O}(m \cdot n)$ . Mit dem schnellen Multipolverfahren aus Kapitel 3.6 lässt er sich auf  $\mathcal{O}((m+n) \log(m))$  reduzieren. Eine genauere Aufwandsabschätzung fehlt in der Literatur, insbesondere eine Worst-Case-Abschätzung. Deswegen werden an dieser Stelle die Aufwände genauer bestimmt und mit der direkten Auswertung verglichen. Es wird untersucht, unter welchen Bedingungen eine Aufwandseinsparung beim Einsatz dieses Verfahrens in der hier entwickelten Methodik zur robusten Optimierung möglich ist.

Bei der Abschätzung wird der polynomielle Anteil des RBF-Modells außer acht gelassen, da er bei beiden Vorgehensweisen gleich hoch ist. Eine Integration in die Multipol-Entwicklung ist nicht möglich.

Zuerst werden die Aufwände der Erstellung eines Baumes für die  $m$  Auswertungspunkte im günstigsten und im schlechtesten Fall bestimmt (Kapitel 7.1, Algorithmus 4). Danach folgen Aufwandsabschätzungen für die Extremsituationen beim eigentlichen Cluster-Partikel-Algorithmus (Kapitel 7.2, Algorithmus 5). Die Komplexität der direkten Summation wird in Kap. 7.3 betrachtet. Vergleichende Berechnungen bietet Kap. 7.4, aus denen schließlich in Kap. 7.5 Empfehlungen für den Einsatz zur Bestimmung von Quantilmetamodellen gegeben werden.

### 7.1 Erstellung des Baumes

Zunächst wird eine Baumstruktur erstellt, die alle Auswertungspunkte enthält. In Algorithmus 4 ist der Pseudocode angegeben, der allen  $y_j, j = 1, \dots, m$  Clusterzentren  $y_C$  zuordnet und diese in einem Baum abspeichert. Die Höchstzahl der Partikel pro Blatt ist auf  $m_0$  begrenzt. Ein Cluster hat die Form eines Quaders. Für jede

Verfeinerungsstufe des Baumes wird abgespeichert, welchen Radius  $r$  und welches Zentrum  $y_C$  das zugehörige Cluster  $C$  besitzt.

```

1 Algorithm Baum erstellen()
   | input :  $y_1, \dots, y_m$  Partikel = Auswertungspunkte
   | output: Baum
2   for  $j = 1$  to  $m$  do
3     | Baum einfügen( $j, root$ )
4   end
5   Durchlaufe den Baum (z.B. mit Breitensuche);
6   Lösche alle leeren Blätter (der Baum könnte Blätter besitzen, deren
   | Geschwister nicht leer sind);
7 Procedure: Baum einfügen( $j, k$ )
8   | if der Unterbaum mit Wurzel  $k$  enthält mehr als  $m_0$  Partikel then
9     | Bestimme Kind  $c$  von Wurzel  $k$ , in dem  $j$  liegt;
10    | Baum einfügen( $j, c$ );
11   | else
12     | if der Unterbaum mit Wurzel  $k$  enthält genau  $m_0$  Partikel then
13       | Füge  $k$   $2^d$  Kinder hinzu;
14       | Bestimme  $y_C$  und  $r$  für alle Kinder;
15       | Füge die schon vorhandenen Partikel in die richtigen Kinder
16       | ein;
17       | Sei  $c$  das Kind, in dem Partikel  $j$  liegt;
18       | Baum einfügen( $j, c$ );
19     | else
20       | (der Unterbaum mit Wurzel  $k$  enthält weniger als  $m_0$ 
21       | Partikel,  $k$  ist Blatt);
22       | Speichere Partikel  $j$  in  $k$ 
23     | end
24   | end

```

**Algorithm 4:** Algorithmus für die Erstellung der Baumstruktur

Für die Erstellung des Baumes werden zwei extreme Situationen miteinander verglichen:

- günstigster Fall:  
Die Partikel sind im Designraum gleichmäßig verteilt. Das Ergebnis wird ein ausgewogener Baum mit gleichmäßig vielen Ästen sein. In diesem Fall besitzt jeder Knoten  $2^d$  Äste. Dies ergibt für eine Baumtiefe  $bt$   $2^{bt \cdot d}$  Blätter. Im günstigsten Fall enthalten sie alle genau  $m_0$  Elemente. Wegen  $2^{bt \cdot d} \cdot m_0 = m$  folgt für die Baumtiefe

$$bt = \frac{1}{d} \log_2\left(\frac{m}{m_0}\right)$$

- schlechtester Fall:  
Die Partikel sind an einer Stelle des Designraums stark konzentriert. Wie schon in Kapitel 3.6 erwähnt, entsteht im schlimmsten Fall ein Baum mit nur einem Ast, für den gilt:

$$bt = m$$

Abbildung 7.1 zeigt ein Beispiel für die Verteilung der Punkte im schlechtesten Fall.

### 7.1.1 Aufwandsabschätzung für den günstigsten Fall

Die Aufwandsabschätzung wird in drei Teile unterteilt. Aufwand für

- a) Partikel in den Baum einfügen
- b) Unterbäume einfügen
- c)  $y_C$  und  $r$  bestimmen.

zu a)

Für  $m \leq m_0$  werden die Partikel der Wurzel hinzugefügt. Pro Partikel sind zwei Vergleiche nötig. Für die Partikel  $y_{m_0+1}, \dots, y_{m_0 2^d}$  sind drei Vergleiche nötig, für die nächsten  $y_{m_0 2^{d+1}}, \dots, y_{m_0 2^{2d}}$  vier usw.

$m$	$bt$	Anzahl Operationen
$m_0$	0	$2m_0 = a_0$
$m_0 2^d$	1	$a_0 + 3(m_0 2^d - m_0) = a_1$
$m_0 2^{2d}$	2	$a_1 + 4(m_0 2^{2d} - m_0 2^d) = a_2$
$\vdots$	$\vdots$	$\vdots$
$m_0 2^{id}$	$i$	$a_i + (i+2)(m_0 2^{id} - m_0 2^{(i-1)d})$

Aus der Tabelle ist ersichtlich, dass sich die Anzahl der Operationen über die folgende Summe bestimmen lässt: \*

$AnzOp$

$$\begin{aligned}
 &= 2m_0 + \sum_{j=1}^{bt} (j+2)(m_0 2^{jd} - m_0 2^{(j-1)d}) \\
 &= \frac{m_0 (2^{bt d} 4^d bt - 2^{1+bt d+d} bt - 5 2^{(bt+1)d} + 2^{bt d} bt + 3 2^{bt d} + 2^d - 1 + 2^{1+bt d} 4^d)}{(2^d - 1)^2}
 \end{aligned}$$

Mit  $bt = \frac{1}{d} \log_2(\frac{m}{m_0})$  bzw.  $2^{bt d} = \frac{m}{m_0}$  folgt:

$$\begin{aligned}
 &AnzOp \tag{7.1} \\
 &= \frac{m_0}{(2^d - 1)^2} \left( \frac{m}{m_0} 4^d \frac{1}{d} \log_2\left(\frac{m}{m_0}\right) - 2^{1+d} \frac{m}{m_0} \frac{1}{d} \log_2\left(\frac{m}{m_0}\right) - 5 2^d \frac{m}{m_0} \right. \\
 &\quad \left. + \frac{m}{m_0} \frac{1}{d} \log_2\left(\frac{m}{m_0}\right) + 3 \frac{m}{m_0} + 2^d - 1 + 2 \frac{m}{m_0} 4^d \right) \\
 &= \frac{1}{(2^d - 1)^2} \left( m \frac{1}{d} \log_2\left(\frac{m}{m_0}\right) (2^{2d} - 2^{d+1} + 1) + m(-5 2^d + 3 + 2^{2d+1}) \right. \\
 &\quad \left. + (2^d - 1)m_0 \right)
 \end{aligned}$$

---

\*Diese Termumformung und weitere wurden mit dem Computer-Algebra-System MAPLE bestimmt.

Der Aufwand für die Erstellung des Baumes liegt damit auch im günstigsten Fall in  $\mathcal{O}(m \log(m))$ . Mit Hilfe von Gleichung 7.1 lässt er sich genau berechnen.

Für  $d = 2$  beträgt er:

$$\text{AnzOp} = \frac{1}{2}m \log_2\left(\frac{m}{m_0}\right) + \frac{5}{3}m + \frac{1}{3}m_0$$

und für  $d = 3$ :

$$\text{AnzOp} = \frac{1}{3}m \log_2\left(\frac{m}{m_0}\right) + \frac{13}{7}m + \frac{1}{7}m_0$$

zu b)

Wenn ein Unterbaum eingefügt wird, sind folgende Schritte notwendig:

- $2^d$  Kinder einfügen ( $2^d$  Operationen)
- die  $m_0$  vorhandenen Partikel in die Kinder einfügen (1 Operation)
- den neuen Partikel einfügen (ebenfalls 1 Operation)

Dies sind  $m_0 + 1 + 2^d$  Operationen für das Einfügen eines Unterbaumes. Im günstigsten Fall werden

$$\sum_{j=0}^{bt-1} 2^{j \cdot d} = \frac{2^{d \cdot bt} - 1}{2^d - 1} = \left(\frac{m}{m_0} - 1\right) \frac{1}{2^d - 1}$$

Unterbäume eingefügt. Damit beträgt der Aufwand für b)

$$(2^d + m_0 + 1) \cdot \left(\frac{m}{m_0} - 1\right) \frac{1}{2^d - 1}$$

zu c)

$y_C$  und  $r$  müssen für jeden Unterbaum bestimmt werden:

$$2 \cdot \left(\frac{m}{m_0} - 1\right) \frac{1}{2^d - 1}$$

Insgesamt beträgt damit der Aufwand für die Erstellung des Baumes im günstigsten Fall:

$$\begin{aligned} \text{AnzOp} &= (2^d + m_0 + 3) \left(\frac{m}{m_0} - 1\right) \frac{1}{2^d - 1} + \frac{1}{(2^d - 1)^2} \\ &\quad \left(m \log_2\left(\frac{m}{m_0}\right) \frac{1}{d} (2^{2d} - 2^{d+1} + 1) + m(-5 \cdot 2^d + 3 + 2^{2d+1}) + (2^d - 1)m_0\right) \end{aligned}$$

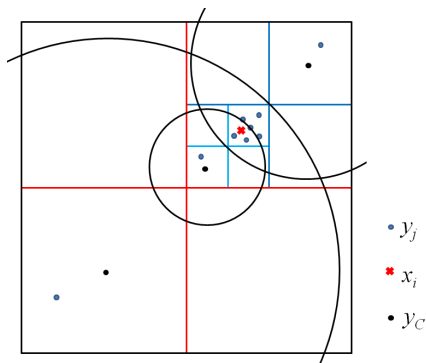


Abb. 7.1: ungünstigste Verteilung der  $x_i$ , die Kreise sind die Grenzen, bis zu denen MAC bzgl.  $y_C$  nicht erfüllt ist

### 7.1.2 Aufwandsabschätzung für den schlechtesten Fall

Im schlechtesten Fall konzentrieren sich die Auswertungspunkte etwa in der Mitte des Designraums. Sie sind so verteilt, dass durch Teilung einer Box ein Punkt separiert wird und alle übrigen Punkte in der neuen kleineren Box liegen (vgl. Abbildung 7.1).

Bei der Erstellung des Baumes werden zunächst  $m_0$  Punkte an der Wurzel eingefügt (jeweils 2 Operationen). Für den nächsten Punkte sind dann 3 Operationen nötig und ein Unterbaum muss eingefügt werden. Ein Unterbaum enthält einen Partikel, ein Unterbaum  $m_0$  Partikel und alle übrigen sind leer. Um den nächsten Punkt einzufügen, sind dann vier Operationen nötig. Er wird im Unterbaum einsortiert, der schon  $m_0$  Punkte enthält, so dass erneut ein Unterbaum eingefügt werden muss. Diese Abfolge wiederholt sich bis alle  $m$  Punkte eingefügt sind. Die Anzahl der nötigen Operationen beträgt dann:

$$\begin{aligned}
 \text{AnzOp} &= 2m_0 + \sum_{j=1}^{m-m_0} (j+2) \left( \underbrace{2^d + m_0 + 1}_{\text{Unterbaum einfügen}} + \underbrace{2}_{y_{C,r} \text{ bestimmen}} \right) \\
 &= 2m_0 + (2^d + m_0 + 3) \left( \frac{1}{2} (m - m_0 + 1)^2 + \frac{3}{2} m - \frac{3}{2} m_0 - \frac{1}{2} \right)
 \end{aligned}$$

Damit liegt der Aufwand im schlechtesten Fall bei  $\mathcal{O}(m^2)$ .

## 7.2 Cluster-Partikel-Algorithmus

Im Cluster-Partikel-Algorithmus wird aufbauend auf der in Kapitel 7.1 erstellten Baumstruktur die Auswertung des Metamodells für die Partikel durchgeführt. In Algorithmus 5 ist der Pseudocode für den Cluster-Partikel-Algorithmus angegeben. Zuerst wird die Prozedur „Berechne Momente( $x_i, C$ )“ für jede Stützstelle  $x_i$  mit der Wurzel des Baumes aufgerufen. Für jedes  $x_i$  wird der Baum so weit durchlaufen, bis der zugehörige Quader und  $x_i$  wohlgetrennt sind und die zugehörigen Taylorkoeffizienten  $a_k$  bestimmt. Die Momente für jedes Quaderzentrum  $y_C$  werden nach und

```

1 Algorithm Cluster-Partikel-Algorithmus()
   input :  $y_1, \dots, y_m$  Auswertungspunkte
   input :  $x_1, \dots, x_n$  Stützstellen
   input :  $c_1, \dots, c_n$  Gewichte des RBF-Modells
   input :  $\theta_0$  MAC-Parameter
   input :  $p$  Ordnung des Taylor-Polynoms
   output:  $f(y_j)$   $j = 1, \dots, m$ 
2 Baum erstellen (bilde Baumstruktur);
3 for  $i = 1$  to  $n$  do
4   | Berechne Momente( $x_i, root$ )
5   end
6 Akkumuliere( $root, []$ ) ;
7 Procedure: Berechne Momente( $x_i, C$ )
8   if  $C$  ist Blatt then
9   | direkte Summation  $f(y_j) = f(y_j) + c_i \Phi(|y_j - x_i|)$ 
10  else
11  | if  $x_i, y_C$  erfüllen  $MAC(\theta_0)$  then
12  | | berechne  $a_k(y_C, x_i)$  ;
13  | | update Moment  $m_k(y_C) = m_k(y_C) + c_i a_k(y_C, x_i)$ 
14  | else
15  | | for  $k = 1$  to Anzahl der Kinder von C do
16  | | | Berechne Momente( $x_i, \text{Kind } k \text{ von } C$ )
17  | | end
18  | end
19  end
20 (erste Summe von Gleichung 2 bestimmt,  $m_k(y_C)$  bestimmt)
21 Procedure: Akkumuliere( $C, H$ )
22 if  $C$  ist Blatt then
23 | for jede Zelle  $\tilde{C}$  in der Hierarchie von H do
24 | |  $f(y_j) = f(y_j) + \sum_{\|k=0\|}^p m_k(\tilde{C})(y_j - y_{\tilde{C}})^k$ 
25 | end
26 else
27 | for jedes Kind  $\tilde{C}$  von C do
28 | | Akkumuliere( $\tilde{C}, [H.C]$ )
29 | end
30 end

```

**Algorithm 5:** Cluster-Partikel-Algorithmus zur schnellen Auswertung eines RBF-Metamodells (entnommen aus [DD12])

nach aufsummiert. Diese Schritte könnten vorab unabhängig von den  $y_i$  geschehen. Um zusätzliche Baumdurchläufe zu sparen, werden in der Prozedur aber auch schon die direkten Summationen bzgl. der Partikel  $y_j$  im Quader durchgeführt, wenn der Quader und  $x_i$  nicht wohlgetrennt sind.

Nach Durchlauf der *for*-Schleife sind

1. die direkten Summationen im Nahfeld durchgeführt, d.h. für alle  $y_j$  wurden die Basisfunktionen aufaddiert und
2. alle Momente  $m_k$  bestimmt.



In der Prozedur „Akkumuliere( $C, H$ )“ werden nun noch die Momente des Fernfeldes zu den  $y_j$  addiert.

Für die Erstellung des günstigsten und schlechtesten Falles muss die Verteilung der Stützstellen  $x_i, i = 1, \dots, n$  zusätzlich zur Verteilung der Auswertungspunkte  $y_j, j = 1, \dots, m$  berücksichtigt werden.

- Der günstigste Fall ist derjenige, bei dem in der Prozedur „Berechne Momente“ der Durchlauf des Baumes auf einer möglichst hohen Ebene beendet wird, d.h. wenn in Zeile 11 das MAC-Kriterium erfüllt ist. Dies ist erfüllt, wenn die Stützstellen in den Ecken des Designraums liegen (vgl. Abbildung 7.2). Zum Nahfeld gehören dann nur die wenigen unmittelbaren Nachbarzellen. Als MAC wird  $r \leq 0,5R$  benutzt.
- Im schlechtesten Fall muss bis in die Blätter der untersten Stufe des Baumes verzweigt werden. Dies bedeutet für den schlechtesten Baum aus Kapitel 7.1, dass auch die Stützstellen in diesen Blättern liegen.

### 7.2.1 Aufwandsabschätzung für den günstigsten Fall

#### Aufwand für „Berechne Momente“

Zunächst wird die Anzahl der Vergleiche bestimmt, die im günstigsten Fall nötig sind. Die Prozedur „Berechne Momente“ wird zuerst für ein  $x_i$  und die Wurzel aufgerufen. Nach zwei Vergleichen wird in alle  $2^d$  Kinder der Wurzel verzweigt. Für  $2^d - 1$  von ihnen ist MAC erfüllt (2 Vergleiche je Kind), für das verbliebene muss in die  $2^d$  Kinder abgestiegen werden usw. bis schließlich das Blatt, in dem  $x_i$  liegt, erreicht ist. Dies sind

$$\begin{aligned} 2 + \underbrace{2(2^d - 1) + 2(2^d - 1) + \dots + 2(2^d - 1)}_{bt \text{ Mal}} + 2^d bt &= 1 + 2bt(2^d - 1) + 2^d bt \\ &= bt(2^{d+1} + 2^d - 1) + 2 \end{aligned}$$

Vergleiche für ein  $x_i$ . Für alle  $n$  Stützstellen ergibt sich mit  $bt = \frac{1}{d} \log_2(\frac{m}{m_0})$ :

$$\text{AnzOp} = n \left( \frac{1}{d} \log_2 \left( \frac{m}{m_0} \right) (2^{d+1} + 2^d - 1) + 2 \right) \quad (7.2)$$

Bevor die Momente berechnet werden können, müssen die  $a_k$  für  $\|k\| = 0, \dots, p$  über die Rekursionsformel 3.21 bestimmt werden. Sie lautet nach  $a_k$  aufgelöst:

$$\begin{aligned} a_k &= \frac{2\|k\| - 3}{\|k\|(|y - x|^2 + b^2)} \sum_{i=1}^d (y_i - x_i) a_{k - e_i} \\ &\quad + \frac{\|k\| - 3}{\|k\|(|y - x|^2 + b^2)} \sum_{i=1}^d (y_i - x_i) a_{k - 2e_i} \end{aligned}$$

wobei  $|\cdot|$  die Euklidische Norm ist. Für die Bestimmung eines  $a_k$  sind dann  $4d + 2$  Additionen oder Subtraktionen und  $2d + 4$  Multiplikationen nötig.

Es gibt  $\binom{p+d-1}{d-1}$  Multi-Indizes mit  $\|k\| = p$ . Die Anzahl der Summanden der Summe von 0 bis  $p$  beträgt dann:

$$\sum_{j=0}^p \binom{j+d-1}{d-1} = \binom{p+d}{d} \tag{7.3}$$

Damit beträgt der Aufwand der Berechnung von  $a_k(y_C, x_i)$   $\binom{p+d}{d}(4d+2)$  Additionen und  $\binom{p+d}{d}(2d+4)$  Multiplikationen.

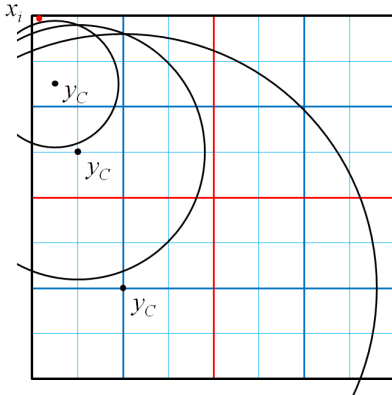


Abb. 7.2: günstigste Verteilung der  $x_i$ , in der äußersten Ecke des Designraums, die Kreise sind die Grenzen, bis zu denen MAC bzgl.  $y_C$  nicht erfüllt ist

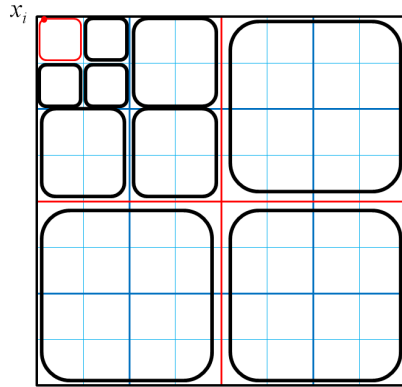


Abb. 7.3: direkte Summation (rot) und Boxen, für die MAC erfüllt ist (schwarz)

Wenn alle Stützstellen in den Eckboxen liegen, werden höchstens  $2^d$  Boxen direkt summiert, für  $2^d - 1$  Boxen der zweituntersten Ebene ist MAC erfüllt, dann für die  $2^d - 1$  Boxen der drittuntersten Ebene usw. (vgl. Abbildung 7.3).

Die  $a_k$  sind  $(bt - 2)(2^d - 1)$  Mal zu berechnen, für die  $m_k$  werden ebensoviele Multiplikationen und Additionen ausgeführt.

Für die direkte Summation sind dann in der Prozedur „Berechne Momente“  $2^d$  Multiplikationen,  $2^d$  Additionen und  $2^d$  Auswertungen der Basisfunktion nötig. Eine Basisfunktion kann mit  $d + 1$  Multiplikationen,  $d$  Additionen und einer Wurzelberechnung ausgewertet werden.

Diese Berechnungen müssen für alle  $n$  Stützstellen durchlaufen werden.

Damit ergibt sich zusätzlich zu den Vergleichen:

$$\begin{aligned}
 \text{AnzMult} &= \underbrace{n}_{\text{Anz. Stützstellen}} \cdot \underbrace{\binom{p+d}{d}}_{\text{Anz. } a_k} \underbrace{(2d+4)}_{a_k} \underbrace{(bt-2)}_{\text{Abstiegstiefe auf jeder Stufe}} \underbrace{(2^d-1)}_{(2^d-1)} \\
 &\quad + \underbrace{(bt-2)(2^d-1)}_{m_k} + \underbrace{2^d+d+1}_{\text{dir. Summ.}} \\
 &= n \left( \frac{1}{d} \log_2 \left( \frac{m}{m_0} \right) - 2 \right) (2^d - 1) \left( \binom{p+d}{d} (2d+4) + 1 \right) + n(2^d + d + 1) \\
 \text{AnzAdd} &= n \cdot \left( \binom{p+d}{d} (4d+2)(bt-2)(2^d-1) \right. \\
 &\quad \left. + (bt-2)(2^d-1) + 2^d + d \right) \\
 &= n \left( \frac{1}{d} \log_2 \left( \frac{m}{m_0} \right) - 2 \right) (2^d - 1) \left( \binom{p+d}{d} (4d+2) + 1 \right) + n(2^d + d) \\
 \text{AnzWur} &= 2^d
 \end{aligned}$$

Der Aufwand liegt damit in  $\mathcal{O}(n \log(m))$ . Der Aufwand bzgl.  $p$  beträgt  $\mathcal{O}(p^{d+1})$ .

### Aufwand für „Akkumuliere“

Die Prozedur „Akkumuliere“ steigt bis in jedes Blatt hinab, dann wird in der Hierarchie nach oben akkumuliert. Die  $2^{bt-d}$  Blätter besitzen alle die Tiefe  $bt$ , so dass insgesamt

$$2 \cdot 2^{bt-d} bt = 2 \frac{m}{m_0} \frac{1}{d} \log_2 \left( \frac{m}{m_0} \right)$$

Vergleiche durchgeführt werden müssen.

Die Auswertung des Ausdrucks in Zeile 24 benötigt

$$\sum_{j=0}^p \binom{j+d-1}{d-1} (j+1) = \frac{(p+1)(1+(p+1)d)}{d(d+1)} \binom{p+d}{d-1}$$

Multiplikationen und  $\binom{d+p}{d}$  Additionen. Der Baum besitzt  $2^{bt-d}$  Blätter. Die Hierarchietiefe ist immer gleich  $bt$ . Insgesamt:

$$\begin{aligned}
 \text{AnzMult} &= \frac{(p+1)(1+(p+1)d)}{d(d+1)} \binom{p+d}{d-1} 2^{bt-d} bt \\
 &= \frac{m}{m_0} \frac{1}{d} \log_2 \left( \frac{m}{m_0} \right) \frac{(p+1)(1+(p+1)d)}{d(d+1)} \binom{p+d}{d-1} \\
 \text{AnzAdd} &= \binom{p+d}{d} 2^{bt-d} bt \\
 &= \frac{m}{m_0} \frac{1}{d} \log_2 \left( \frac{m}{m_0} \right) \binom{p+d}{d}
 \end{aligned}$$

Der Aufwand liegt damit in  $\mathcal{O}(m \log(m))$ . Der Aufwand bzgl.  $p$  beträgt auch hier  $\mathcal{O}(p^{d+1})$ .

### Aufwand insgesamt

Fasst man die nötigen Berechnungsschritte aus „Erstelle Baum“ , „Berechne Momente“ und „Akkumuliere“ zusammen, erhält man insgesamt:

*AnzOp*

$$\begin{aligned} &= (2^d + m_0 + 3) \left( \frac{m}{m_0} - 1 \right) \frac{1}{2^d - 1} + \frac{1}{(2^d - 1)^2} \\ &\quad \left( m \log_2 \left( \frac{m}{m_0} \right) \frac{1}{d} (2^{2d} - 2^{d+1} + 1) + m(-5 \cdot 2^d + 3 + 2^{2d+1}) + (2^d - 1)m_0 \right) \\ &\quad + n \left( \frac{1}{d} \log_2 \left( \frac{m}{m_0} \right) (2^{d+1} + 2^d - 1) + 2 \right) \\ &\quad + 2 \frac{m}{m_0} \frac{1}{d} \log_2 \left( \frac{m}{m_0} \right) \end{aligned}$$

$$\begin{aligned} \textit{AnzMult} &= n \left( \frac{1}{d} \log_2 \left( \frac{m}{m_0} \right) - 2 \right) (2^d - 1) \binom{p+d}{d} (2d+4) + n(2^d + d + 1) \\ &\quad + \frac{m}{m_0} \frac{1}{d} \log_2 \left( \frac{m}{m_0} \right) \frac{(p+1)(1+(p+1)d)}{d(d+1)} \binom{p+d}{d-1} \end{aligned}$$

$$\begin{aligned} \textit{AnzAdd} &= n \left( \frac{1}{d} \log_2 \left( \frac{m}{m_0} \right) - 2 \right) (2^d - 1) \binom{p+d}{d} (4d+2) + n(2^d + d) \\ &\quad + \frac{m}{m_0} \frac{1}{d} \log_2 \left( \frac{m}{m_0} \right) \binom{p+d}{d} \end{aligned}$$

$$\textit{AnzWur} = 2^d$$

Der Gesamtaufwand liegt auch im günstigsten Fall in  $\mathcal{O}((m+n) \log(m))$ . Über die obigen Formeln kann er nun genau bestimmt werden.

## 7.2.2 Aufwandsabschätzung für den schlechtesten Fall

### Aufwand für „Berechne Momente“

Im schlechtesten Fall liegen die Stützstellen in der kleinsten Box, d.h. im untersten Blatt. Dann ist für die  $m_0$  Elemente dieser Box direkte Summation erforderlich. Abbildung 7.1 zeigt, dass es möglich ist, die Stützstellen  $x_i$  so zu platzieren, dass MAC auch für die größeren Boxen des Baumes nicht erfüllt ist. Auch für die übrigen Auswertungspunkte ist direkte Summation notwendig und es werden keine Momente berechnet.

Der Aufwand ist daher mindestens so hoch wie für die direkte Summation, die in Kapitel 7.3 behandelt wird.

Da der Baum für jede Stützstelle einmal bis zum Blatt hinabgestiegen wird und diese im schlechtesten Fall  $m$  beträgt, sind zusätzlich noch

$$\text{AnzOp} = n \cdot m$$

Vergleiche nötig.

### Aufwand für „Akkumuliere“

In der Prozedur „Akkumuliere“ wird der Baum erneut bis zum Blatt durchlaufen und anschließend wieder zur Wurzel zurückgegangen. Dies entspricht

$$\text{AnzOp} = 2m$$

### Aufwand insgesamt

Der Aufwand beträgt

$$\begin{aligned} \text{AnzOp} &= 2m_0 + (2^d + m_0 + 3)\left(\frac{1}{2}(m - m_0 + 1)^2 + \frac{3}{2}m - \frac{3}{2}m_0 - \frac{1}{2}\right) \\ &\quad + n \cdot m + 2m \end{aligned}$$

und zusätzlich wie bei direkter Summation (vgl. Kapitel 7.3):

$$\begin{aligned} \text{AnzMult} &= m \cdot n \cdot (d + 1) \\ \text{AnzAdd} &= m \cdot (n - 1 + n \cdot d) = m \cdot (n(d + 1) - 1) \\ \text{AnzWur} &= m \cdot n \end{aligned}$$

## 7.3 Aufwandsabschätzung für die direkte Summation

Bei der direkten Summation ist der Ausdruck

$$y = \sum_{i=1}^n c_i \Phi(|y - x_i|) \quad (7.4)$$

$m$  Mal auszuwerten. Die Auswertung des Ausdrucks

$$\Phi(|y_j - x_i|) = \sqrt{b^2 + \sum_{k=1}^d (y_{jk} - x_{jk})^2}$$

erfordert  $d + 1$  Multiplikationen,  $d$  Additionen und eine Wurzelberechnung. Für den Ausdruck in Gleichung 7.4 benötigt man dann  $n \cdot (d + 1)$  Multiplikationen,  $(n - 1) + n \cdot d$  Additionen und  $n$  Wurzelberechnungen. Für  $m$  Auswertungspunkte ergibt dies insgesamt:

$$\begin{aligned} \text{AnzMult} &= m \cdot n \cdot (d + 1) \\ \text{AnzAdd} &= m \cdot (n - 1 + n \cdot d) = m \cdot (n(d + 1) - 1) \\ \text{AnzWur} &= m \cdot n \end{aligned}$$

## 7.4 Vergleich

Die obige Untersuchung zeigt, dass der Aufwand für das Multipolverfahren selbst im günstigsten Fall in  $\mathcal{O}((m+n) \log(m))$  liegt. Nimmt man  $p$  und  $d$  zur Betrachtung hinzu, ergibt sich  $(n+m) \log(m) p^d$  als größter Faktor bei den Multiplikationen und Additionen. Im Vergleich dazu kommt die direkte Summation auf  $n \cdot m \cdot d$ . Die Anzahl der Eingabevariablen  $d$  geht einmal exponentiell und einmal linear in den Aufwand ein. Ob sich die Anwendung des Multipolverfahrens lohnt, hängt damit von der Dimension des Problems ab. Je mehr Eingabevariablen vorhanden sind, desto aufwändiger wird die Berechnung.

Für den schlechtesten Fall, d.h. bei ungünstiger Verteilung der Auswertungspunkte und Stützstellen zueinander, wurde gezeigt, dass der Aufwand für das schnelle Multiverfahren *immer* größer als für die direkte Summation ist.

Die aufgeführten Aufwände werden nun anhand von konkreten Zahlen verglichen. In den Anwendungen ist  $p = 20$  und  $m_0 = 500$  üblich. Außerdem wird angenommen, dass die Stichprobengröße um jede Stützstelle 50.000 bzw. 100.000 beträgt. Zur Vereinfachung wird die Anzahl aller Operationen addiert. Für diese Annahmen lässt sich nun in Abhängigkeit von der Dimension  $d$  bestimmen, ab welcher Stützstellenanzahl  $n$  das Multipolverfahren günstiger ist. Das Ergebnis ist in der folgenden Tabelle aufgeführt:

$$m = 50000 \cdot n$$

$d$	1	2	3	4	5	6	7	8	9
$n$	4	7	29	148	687	3009	8634	36933	>100.000

$$m = 100000 \cdot n$$

$d$	1	2	3	4	5	6	7	8	9
$n$	4	8	35	144	649	2680	7373	28173	>100.000

Je größer die Stichproben um die Stützstellen sind, desto eher lohnt sich das Multipolverfahren.

Bei Metamodellen für die robuste Optimierung sind mehr als 100 Stützstellen selten. Ein Anwendung des schnellen Multipolverfahrens in diesem Bereich lohnt sich daher nur bis zur Dimension 3.

Nimmt man an, dass für ein Metamodell 100 Stützstellen bestimmt werden und  $m = 50000n$  gilt, kann man nicht nur die Gesamtanzahl, sondern auch die Aufwände für die einzelnen Operationen vergleichen. Abbildung 7.4 zeigt dies für  $d = 3$  und  $d = 4$ . Die Anzahl der Multiplikationen nimmt beim günstigsten Fall bei steigender Dimension sehr stark zu. Für  $d = 4$  ist sie deutlich größer als für die direkte Summation. Zwar müssen hier wesentlich mehr aufwändige Wurzelbestimmungen und einfache Additionen durchgeführt werden, doch insgesamt ist das Multipolverfahren hier teurer in der Berechnung.

Mit den gewonnenen Formeln kann dieser Vergleich auch für andere Zahlenkombinationen durchgeführt werden.

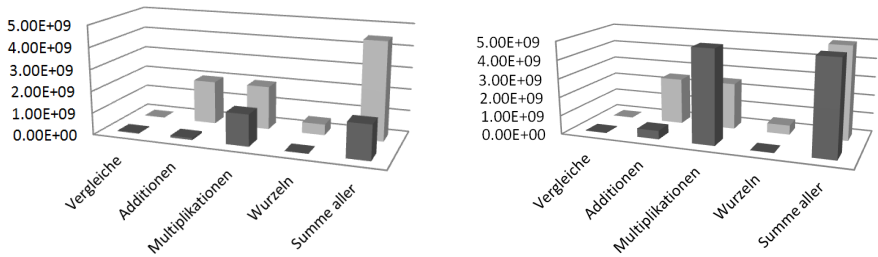


Abb. 7.4: Aufwandsvergleich für  $d = 3$  (links) und  $d = 4$  (rechts), dunkelgrau = günstigster Fall, hellgrau = direkte Summation

## 7.5 Diskussion

Boateng und Krasny [BK13] wenden das gleiche schnelle Multipolverfahren auf die Berechnung von elektrostatischen Potentialen im  $\mathbb{R}^3$  an. Bei der Baumstruktur werden je Blatt bis 500 Auswertungspunkte zugelassen. In den dort aufgeführten numerischen Tests ist der relative Fehler unabhängig von  $n$  und  $m$ . In vielen Fällen ist der Treecode-Algorithmus schneller als direkte Summation. Die Laufzeit hängt fast linear von  $n$  und unterlinear von  $m$  ab. Der Gesamtlaufzeit wird daher vom Term  $\mathcal{O}(n \log(m))$  dominiert, also von der Phase „Berechnung der Momente“. Im ersten Testfall ist der Algorithmus deutlich schneller als die direkte Summation. Allerdings wurde eine niedrige Ordnung von  $p$  gewählt, so dass der relative Fehler ca. 0,5% betrug. In Testfall 2 ist der Treecode auch für hohe  $p$  deutlich schneller. Der Speicherbedarf ist bis zu vier Mal so groß wie bei der direkten Summation.

Deng [DD12] wendet das Verfahren auf die Auswertung von RBF-Modellen an. Er nimmt in der numerischen Untersuchung an, dass die Stützstellen gleich den Auswertungspunkten sind. Dann ist  $m = n$  und die Zeitkomplexität beträgt  $\mathcal{O}(n \log(n))$ . Er untersucht nur  $d = 2$  und  $d = 3$ . Der angegebene Aufwand gilt für den Fall, dass die Auswertungspunkte gleichmäßig im Designraum verteilt sind. Dann wird eine Baumtiefe von  $\log_2(m)$  erreicht.

Diese Ergebnisse lassen sich nur mit Einschränkungen auf die Anwendung der schnellen Multipol-Methode für die Berechnung der übertragenen Quantile an den Stützstellen des Metamodells übertragen: Hier werden die Auswertungspunkte gemäß der Eingangsverteilung um die Stützstellen verteilt sein. Dies ist keine Gleichverteilung! Bei eng beieinanderliegenden Stützstellen werden die Auswertungspunkte den Designraum zumindest ausfüllen. Liegen die Stützstellen weit auseinander und ist die Streuung der Eingangsverteilungen klein, wird es auch Teile des Designraums geben, die keine Auswertungspunkte enthalten.

Da die Baumtiefe proportional zum kleinsten Abstand zweier Auswertungspunkte ist, sind sehr große Baumtiefen zu erwarten.

Bei den Aufwandsbetrachtungen für das schnelle Multipolverfahren darf die Dimension  $d$  der Aufgabe nicht außer acht gelassen werden. Sie beeinflusst den Aufwand wesentlich. Für sehr kleine Dimensionen wird sich der Einsatz auch in der vorgestellten Methodik lohnen. Bei höheren Dimensionen hängt die Entscheidung von der

Verteilung und Anzahl der Stützstellen und Auswertungspunkte ab. Nimmt man an, dass nicht mehr als 100 Stützstellen bestimmt werden und die Stichproben bis zu 100.000 Elemente enthalten, lohnt sich der Einsatz im günstigsten Fall bis zu drei Dimensionen.



## Teil II

# Numerische Untersuchungen



## Kapitel 8

---

# Übertragene Verteilung für einen Punkt

---

Im vorangegangenen theoretischen Teil der Arbeit wurde eine Methodik zur Bestimmung von robusten Optima mit Hilfe von Metamodellen vorgeschlagen. Ihre Anwendbarkeit wird in den nun folgenden numerischen Untersuchungen verifiziert. Zunächst wird nicht der gesamte Designraum betrachtet, sondern jeweils eine feste Parameterkombination der Eingangsvariablen.

Die Gültigkeit der gesamten Methodik wird an zwei Anwendungsfällen in Kapitel 9 demonstriert.

In diesem Kapitel wird angenommen, dass die Übertragungsfunktion  $f$  bekannt ist. Es werden verschiedene Methoden zur Berechnung der Verteilung von  $Y_0 = f(X_0)$  für festes  $x_0$  miteinander verglichen.

In einer numerischen Untersuchung wird geklärt, welche Kombination aus Versuchsplan und Quantilschätzer die beste Schätzung liefert. Die untersuchten Testfälle hierzu werden in Kapitel 8.2 vorgestellt. Der Kerndichteschätzer benötigt die Festlegung einer Bandweite  $h$ . Für sie wird eine einfache Formel ermittelt (Kapitel 8.3), die beim anschließenden Vergleich der verschiedenen Quantilschätzer Anwendung findet (Kapitel 8.4).

### 8.1 Schätzung der Normalverteilung

Zunächst wird der einfachste Fall betrachtet: In einer Dimension wird die Standardnormalverteilung durch die Funktion  $f(x) = x$  übertragen. Die resultierende Verteilung ist offensichtlich wieder die Standardnormalverteilung.

Für die deterministischen Versuchspläne CSS und HA wird eine Stichprobe vom Umfang 1000 gezogen und alle Perzentile geschätzt. Für die nichtdeterministischen SS (in einer Dimension gleich LHS) und MC werden jeweils 100 Stichproben vom Umfang 1000 gezogen und für alle Perzentile jeweils der Erwartungswert gebildet.

In diesem Kapitel wird der Fehler lokal betrachtet. Für jedes einzelne Quantil werden die unterschiedlichen Versuchspläne und Schätzer miteinander verglichen.

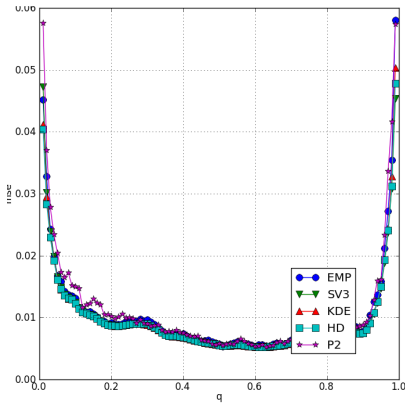


Abb. 8.1: MSE für Monte Carlo

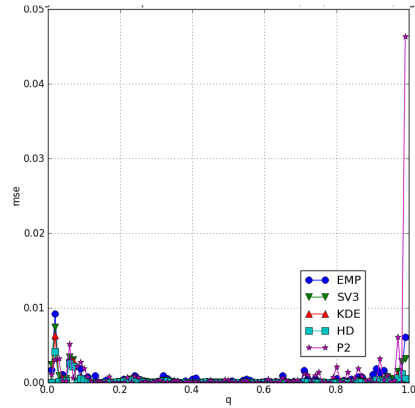


Abb. 8.2: MSE für Haltonfolge

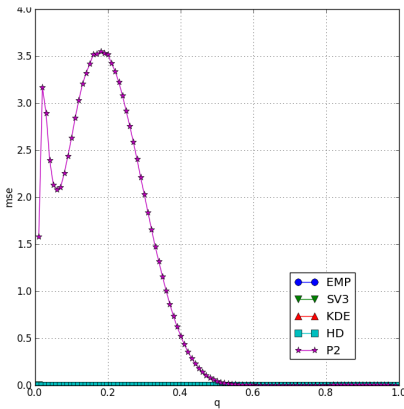


Abb. 8.3: MSE für Stratified Sampling

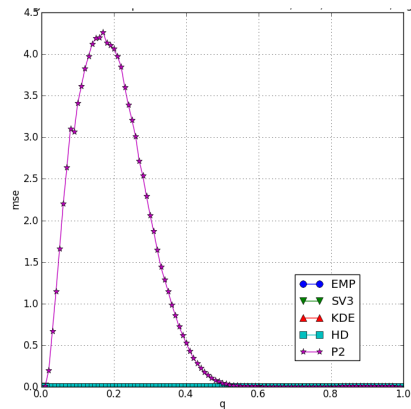


Abb. 8.4: MSE für Centered Stratified Sampling

In den Abbildungen 8.1 bis 8.4 ist jeweils der mittlere quadratische Fehler der verschiedenen Versuchspläne aufgeführt. Für CSS und SS zeigen sie einen großen Fehler für P2 bei den Quantilen kleiner 0, 5. Dies liegt an ihrer Konstruktion: In einer Dimension sind die Stichprobenwerte der Größe nach geordnet. P2 ist für diese Versuchspläne in diesem Quantilbereich nicht geeignet. Um die anderen Schätzer besser vergleichen zu können, zeigen die Abbildungen 8.5 bis 8.8 die Versuchspläne noch einmal ohne P2 und in logarithmischer Skala. Bei allen übrigen Schätzern nimmt der mittlere quadratische Fehler deutlich für extreme Quantile zu. Er ist ungefähr symmetrisch zum Median.

Bei MC sind alle Schätzer etwa gleich gut. Bei HA oszilliert die Schätzung stark in Abhängigkeit von  $q$ . SS und CSS liefern ein ähnliches Bild: Für den Median bringt KDE die beste Schätzung, bei den übrigen kleinen bis mittleren Quantilen

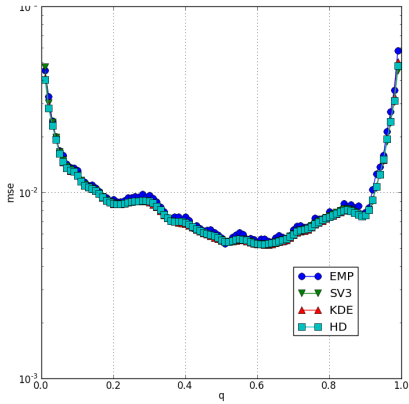


Abb. 8.5: MSE für Monte Carlo Versuchsplan ohne P2

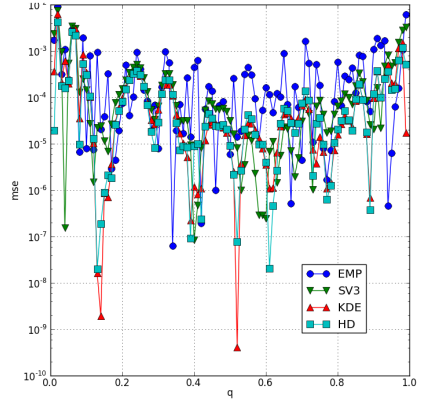


Abb. 8.6: MSE für Haltonfolge, ohne P2

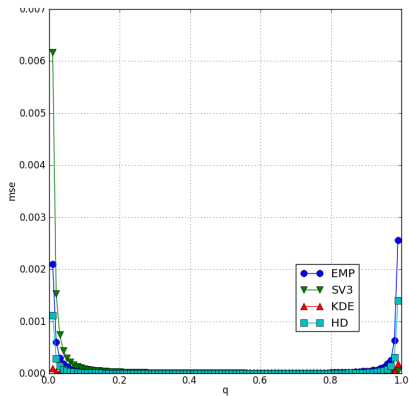


Abb. 8.7: MSE für Stratified Sampling, ohne P2

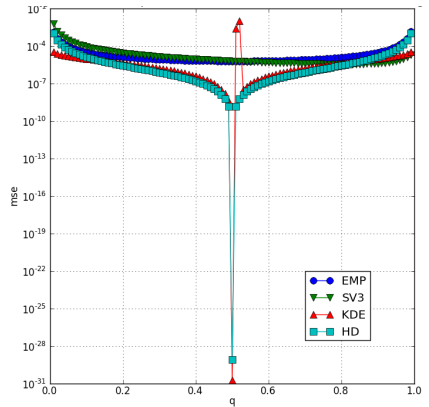


Abb. 8.8: MSE für Centered Stratified Sampling, ohne P2

HD und dann EMP. SV3 verläuft unsymmetrisch und ist für sehr große Quantile am besten und für sehr kleine am schlechtesten. Ansonsten gilt für extreme Quantile die Reihenfolge KDE ist besser als HD ist besser als EMP.

Schon dieses einfache Beispiel zeigt, dass die Genauigkeit der Schätzung sowohl vom Versuchsplan als auch vom Quantilschätzer beeinflusst wird. Darüber hinaus hängt sie von Verhalten von  $Q(q)$  ab.

## 8.2 Testfälle

In der numerischen Untersuchung werden zwei unterschiedliche Aspekte untersucht: Einerseits bestimmen die Quantilschätzer die Verteilung alleine aus den Werten der Stichprobe. Die Eigenschaften der Übertragungsfunktion gehen nur über die  $y_i$  in die

Schätzung ein. Liefern zwei verschiedene Übertragungsfunktionen die gleiche Stichprobe, wird auch die Quantilschätzung gleich ausfallen. Daher sind *unterschiedliche induzierte Verteilungen* zu berücksichtigen. Andererseits wird die Stichprobe nach verschiedenen Versuchsplänen gezogen. Ihre Unterschiede werden bei *unterschiedlichen Übertragungsfunktionstypen* und *Dimensionen* deutlich.

### 8.2.1 Eindimensionale Testfälle

Für eine Dimension werden die folgenden Testfälle betrachtet:

1. Die Übertragungsfunktion ist linear,

$$f_1(x) = 2x \quad \text{mit} \quad x_0 = 0$$

2. Die Übertragung findet im Minimum einer Funktion statt, hier wird gewählt

$$f_2(x) = x^2 \quad \text{mit} \quad x_0 = 0$$

3. Die Übertragung findet ebenfalls im Minimum einer Funktion statt, aber sie ist unsymmetrisch im Nullpunkt und dort nicht differenzierbar. Sie wird auch in [BS07] untersucht.

$$f_3(x) = \begin{cases} -x & x \leq 0 \\ \sqrt{x} & 0 < x \end{cases} \quad \text{mit} \quad x_0 = 0$$

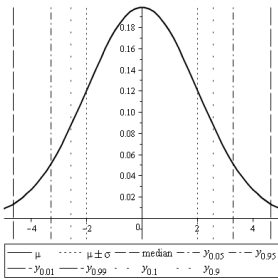


Abb. 8.9: Dichtefunktion von Testfall 1

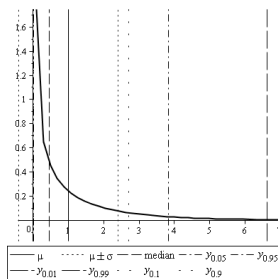


Abb. 8.10: Dichtefunktion von Testfall 2

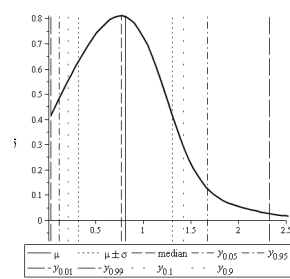


Abb. 8.11: Dichtefunktion von Testfall 3

$X$  sei standardnormalverteilt. Die „wahre“ Verteilung wird über eine Stichprobe vom Umfang 10.000.000, verteilt nach der Haltonfolge, ermittelt. Die Abb. 8.9 bis 8.11 zeigen die resultierenden Dichtefunktionen.

Die beiden ersten Testfälle erzeugen zwei gegensätzliche Verteilungen:

- Testfall 1 lässt die Eingangsverteilung unverändert: Sie ist symmetrisch und flach auslaufend an den Rändern.

- Testfall 2 „klappt“ einen Arm der Normalverteilung auf den anderen und zieht beide in die Länge. Dadurch entsteht die extrem verzerrte Chi-Quadrat-Verteilung, die auf der linken Seite unendlich stark ansteigt.
- Testfall 3 ist interessant, da die übertragene Verteilung nur moderat unsymmetrisch ist und an der linken Seite abrupt endet.

### 8.2.2 Zweidimensionale Testfälle

In zwei Dimensionen werden fünf Testfunktionen für den Referenzpunkt  $x_0 = (1, 1)$  untersucht. Es handelt sich um nichtlineare Funktionen, bei denen sich der Referenzpunkt in der Nähe eines Extremums befindet (vgl. Abb. 8.12 bis 8.15). Die Abbildungen 8.16 bis 8.20 zeigen die übertragenen Dichtefunktionen. Die ersten vier sind [PCG07] entnommen, der zehnte Testfall ist eine lineare Funktion.

4.  $f_4(x) = \sin(x_1 - 0, 21) \cdot \sin(x_2 - 0, 21)$
5.  $f_5(x) = 0, 5x_1^2 - 1, 5x_1 + 0, 7x_2^2 - 1, 2x_2 + 1, 05$
6.  $f_6(x) = 1, 7x_1^3 + 1, 3x_1^2 - 2, 4x_1 - 0, 5x_2^3 + 3, 2x_2^2 - 1, 6x_2 + 0, 1x_1^2x_2 - 0, 2x_1x_2^2 + 1, 6x_1x_2 - 12, 8$
7.  $f_7(x) = 0, 4x_1^2 + 0, 7x_1 + 0, 5x_2^3 - 1, 1x_2^2 - 0, 9x_2 - 1, 3x_2^2x_1$
8.  $f_8(x) = 2x_1 + 3x_2$

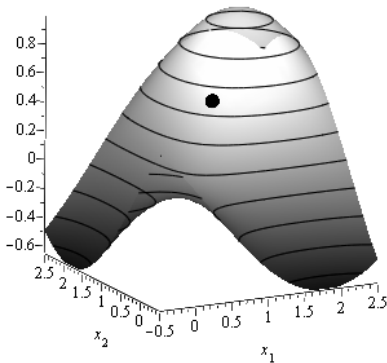


Abb. 8.12:  $f_4$

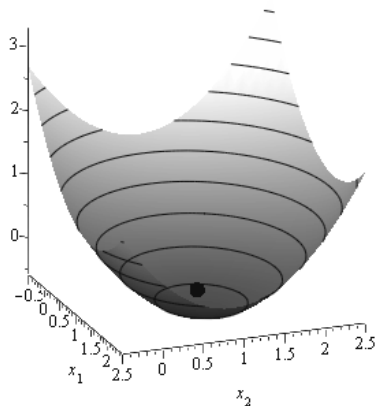


Abb. 8.13:  $f_5$

Hier gelte  $X \sim \mathcal{N}(0; 0, 5)$ . Auch hier wird die Referenzverteilung über eine Haltonfolge vom Umfang 10.000.000 bestimmt.

Die Dichtefunktionen der Testfälle unterscheiden sich erheblich:

- Testfall 4 ist ein Produkt von Sinusfunktionen und daher beschränkt. Die Dichtefunktion enthält einen „Haken“ bei Null und ist bis Null konvex.

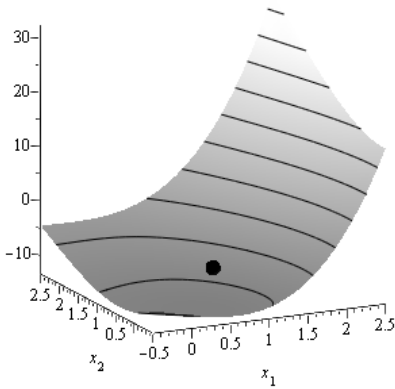
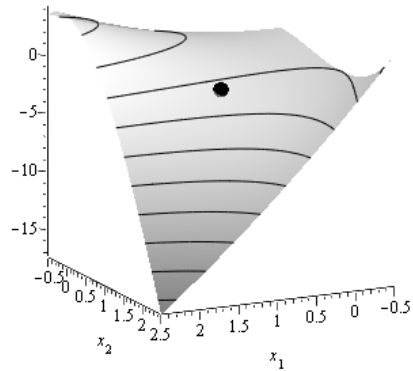
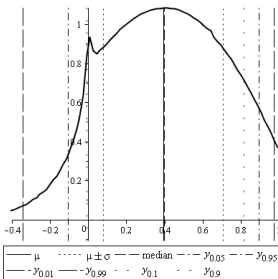
Abb. 8.14:  $f_6$ Abb. 8.15:  $f_7$ 

Abb. 8.16: Dichtfunktion von Testfall 4

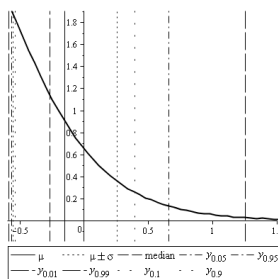


Abb. 8.17: Dichtfunktion von Testfall 5

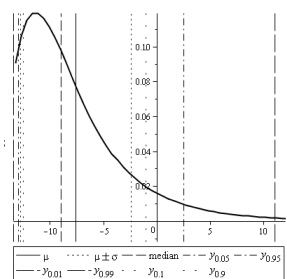


Abb. 8.18: Dichtfunktion von Testfall 6

- Testfall 5 ist ein Polynom zweiten Grades, bei dem  $x_1$  und  $x_2$  unabhängig voneinander sind. Da sich der Referenzpunkt in der Nähe des Minimums befindet, ergibt sich eine rechtsschiefe Verteilung, allerdings weniger ausgeprägt als bei Testfall 2 .
- Testfall 6 ist eine Polynom 3. Grades mit starken Wechselwirkungen zwischen den Parametern. Auch hier befindet sich der Referenzpunkt nahe am Minimum. Die resultierende Verteilung ist eingipflig und läuft rechts sehr flach aus. Sie umfasst den größten Wertebereich der zweidimensionalen Testfälle.
- Bei Testfall 7 ist  $f$  wieder ein Polynom 3. Grades. Diesmal befindet sich der Referenzpunkt in der Nähe eines flachen Maximums. Daher ist die induzierte Dichtfunktion leicht linksschief.
- Da die achte Testfunktion linear ist, ist die übertragene Verteilung normalverteilt.



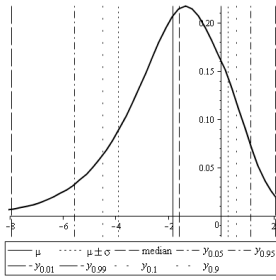


Abb. 8.19: Dichtefunktion von Testfall 7

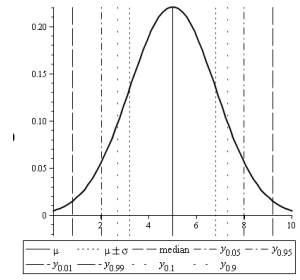


Abb. 8.20: Dichtefunktion von Testfall 8

Polynome vierten Grades und höher kommen in Anwendungen selten vor. Daher werden sie hier nicht untersucht.

### 8.2.3 Anwendungsbeispiel Zweistab-Fachwerk

Dieses Beispiel ist [JDC03] entnommen und wurde in [HD07] zur Berechnung des 0,05- und 0,95-Quantils benutzt. Es beschreibt eine Zweistab-Fachwerk (zsf) wie in Abbildung 8.21.

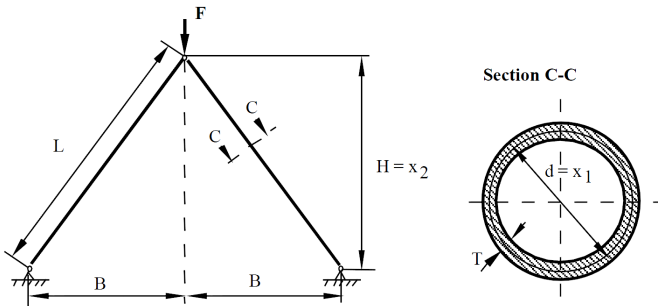


Abb. 8.21: Zweistab-Fachwerk, Grafik entnommen aus [JDC03]

Die Funktion

$$f(x_1, x_2, x_3, x_4, x_5, x_6) = \frac{x_1 \sqrt{x_4^2 + x_5^2}}{2\pi x_3 x_6 x_5} - \frac{\pi^2 x_2 (x_3^2 + x_6^2)}{8(x_4^2 + x_5^2)}$$

beschreibt die Knickbeanspruchung in den beiden Balken, wobei  $x_1$  die äußere Kraft  $F$  ist,  $x_2$  das Elastizitätsmodul  $E$ ,  $x_3$  die Dicke der Querschnitte  $T$ ,  $x_4 = B$  die Breite der Struktur,  $x_5$  die Höhe des Zweistab-Fachwerks  $H$  und  $x_6$  der innere nominale Durchmesser der Querschnitte.

Die Variablen sind alle normalverteilt mit den folgenden Parametern:

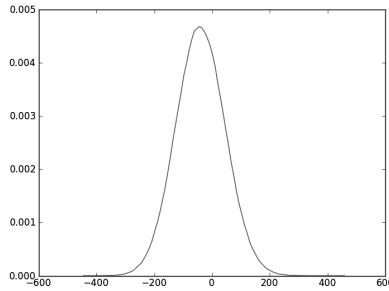


Abb. 8.22: Dichtefunktion für Zweistab-Fachwerk

Variable	Erwartungswert	Standardabweichung
$x_1$	150000 N	30000 N
$x_2$	210000 N/mm <sup>2</sup>	21000 N/mm <sup>2</sup>
$x_3$	2,5 mm	0,1 mm
$x_4$	750 mm	5 mm
$x_5$	720 mm	5 mm
$x_6$	40 mm	1 mm

Hierdurch wird die in Abbildung 8.22 dargestellte übertragene Dichtefunktion erzeugt. Sie ist über einen großen Wertebereich gestreckt und annähernd normalverteilt.

Mit diesem 6-dimensionalen Testfall lassen sich die Versuchspläne für höhere Dimensionen miteinander vergleichen.

### 8.2.4 Crashverhalten der B-Säule eines Pkw

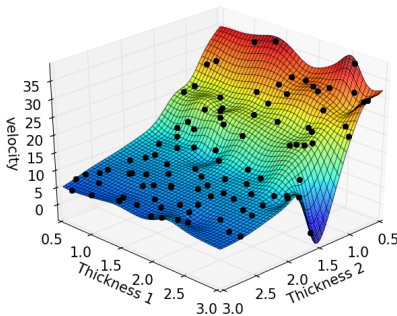


Abb. 8.23: Metamodell für Geschwindigkeit mit Stützstellen

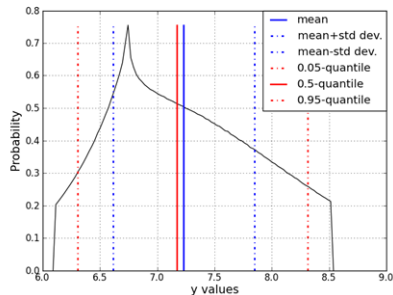


Abb. 8.24: Dichtefunktion von *Geschwindigkeit* für (2,4; 2,9)

Das Anwendungsbeispiel aus dem einführenden Kapitel 2, das Crashverhalten der B-Säule eines Pkw, wird an dieser Stelle wieder aufgegriffen (weitere Einzelheiten

in [NNC12a] und [NNC12b]). Hier wird nun eine andere Ausgabevariable beobachtet: Für einen Referenzpunkt, der sich in der Mitte der B-Säule befindet, wird die maximale Geschwindigkeit während des Crashes im Kontaktbereich bestimmt.

Es werden 101 Simulationen durchgeführt, die gemäß dem Latin Hypercube Sampling verteilt sind. Hieraus wird das Metamodell für die Ausgabevariable erstellt. Abbildung 8.23 zeigt es zusammen mit seinen Stützstellen. Wie man leicht erkennen kann, weist es ein stark nichtlineares Verhalten auf. Dieses Metamodell wird hier als Testfunktion benutzt.

Es sei  $Thickness\ 1 = 2,4$  und  $Thickness\ 2 = 2,9$ . Es wird angenommen, dass beide Eingangsvariablen normalverteilt sind mit der Standardabweichung 0,1. Abbildung 8.24 zeigt die Dichtefunktion der übertragenen Verteilung. Sie verläuft sehr spitz und endet zu beiden Seiten abrupt.

In Kapitel 9.1 wird dieses Beispiel noch einmal aufgegriffen und erweitert.

### 8.2.5 Anwendungsbeispiel pars10crits19

pars10crits19 ist ein Anwendungsbeispiel aus der Automobilindustrie mit 10 Eingabevariablen und 19 Ausgabevariablen. Hierfür wurden 101 Stützstellen bestimmt und ein Metamodell etabliert. Dieses wird hier als Übertragungsfunktion angenommen. Für die numerische Untersuchung werden die Eingabevariablen auf den Wert (1,3; 0,9; 2,3; 1,3; 1,8; 1,36; 1,36; 0,9; 1,2; 1,1) gesetzt. Sie seien normalverteilt mit der Standardabweichung 0,01. Es wird die übertragene Verteilung für Kriterium 3 untersucht. Nach dem zentralen Grenzwertsatz nähern sich übertragene Verteilungen für steigendes  $d$  immer mehr der Normalverteilung an, so auch die in Abbildung 8.25 dargestellte von Kriterium 3. Es ist wichtig, auch in diesem Fall das Quantilmaß zu benutzen, da die übertragenen Verteilungen verzogen oder entlastigt sein können. Nur die Quantile bestimmen die genauen Grenzen.

Das Beispiel wird in Kapitel 9.2 fortgeführt.

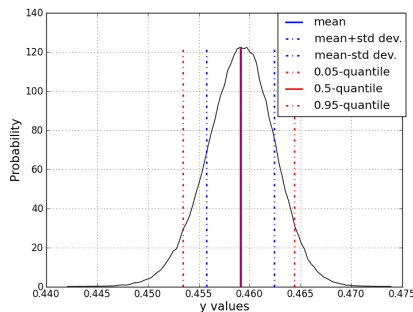


Abb. 8.25: Dichtefunktion von pars10crits19 für die gewählte Parameterkombination

## 8.3 Ermittlung der Bandweite des Kerndichteschätzers

Im Kapitel 3.3.2 wurden die Kerndichteschätzer eingeführt und zwei asymptotisch äquivalente Formen vorgestellt. Für die numerische Untersuchung wird nun die in

Kapitel 3.3.3 beschriebene Form mit dem Gausskern verwendet.

Entscheidend für die Genauigkeit des Schätzers ist die richtige Wahl der Bandweite  $h$ . In Kapitel 8.3.1 werden dazu einige Vorüberlegungen zusammengetragen, bevor in Kapitel 8.3.2 numerische Tests zu seiner Bestimmung durchgeführt werden.

### 8.3.1 Vorüberlegungen

In Kapitel 3.3.2 wurden schon theoretische Ergebnisse für die optimale Bandweite des hier verwendeten Schätzers  $\hat{Q}_n(q)$  mit Gausskern angegeben.

Soll der mittlere quadratische Fehler des Schätzers  $\hat{Q}_n(q)$  mit Gausskern für die Schätzung von  $y_q$  minimiert werden, beträgt die optimale Bandweite (vgl. Gleichung 3.10):

$$\hat{h}_{opt} = \left( \frac{1}{\sqrt{\pi n}} \cdot \frac{p(y_q)}{p'(y_q)^2} \right)^{\frac{1}{3}} = \left( \frac{1}{\sqrt{\pi n}} \cdot \frac{Q'(q)^5}{Q''(q)^2} \right)^{\frac{1}{3}} = \left( \frac{1}{\sqrt{\pi n}} \cdot \frac{1}{\frac{Q''(q)^2}{Q'(q)^5}} \right)^{\frac{1}{3}}$$

Sie hängt vom Quotienten der ersten beiden Ableitungen von  $Q$  ab. Diese sind unbekannt.

$\hat{h}_{opt}$  lässt sich für die beiden ersten Testfälle analytisch bestimmen. In Abbildung 8.26 ist es in Abhängigkeit von  $q$  gezeigt.  $\hat{h}_{opt}$  nimmt für Testfall 1 nur Werte größer als 1 an. Je näher  $q$  an 0,5 rückt, desto größer ist  $\hat{h}_{opt}$ . Für  $q = 0,5$  existiert kein optimales  $h$ . Für Testfall 2 kann  $\hat{h}_{opt}$  alle Werte zwischen 0 und unendlich annehmen.

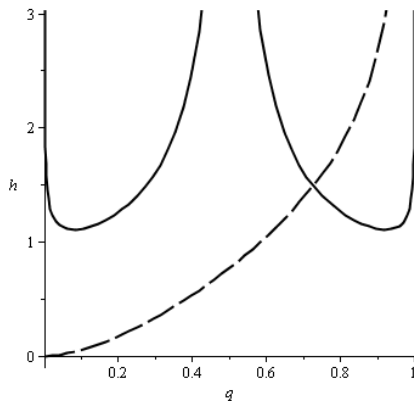


Abb. 8.26:  $\hat{h}$  für  $n = 1$  und Standardnormalverteilung (durchgezogene Linie) und Chi-Quadratverteilung mit  $\mu = 1$  (gestrichelte Linie)

Die Bestimmung von  $\hat{h}_{opt}$  ist aufwändig, da neben dem eigentlichen Ziel,  $Q(q)$  zu bestimmen, auch  $Q'$  und  $Q''$  geschätzt werden müssen, was den Rechenaufwand zusätzlich erhöht.

Daher wird hier auf eine lokale Bestimmung von  $\hat{h}_{opt}$  in Abhängigkeit von  $q$  zur Minimierung des Fehlermaßes MSE verzichtet, um die Berechnung hierdurch nicht zu verlangsamen.

Stattdessen soll der Fehler der Schätzung über die gesamte Verteilung betrachtet werden. Dies erlaubt es,  $q$  zu variieren, ohne eine neue Bandweite bestimmen zu müssen. Als Fehlermaß wird daher der mittlere quadratische Integrationsfehlers über alle  $q$  MISE benutzt.

Bzgl. MISE gilt (vgl. 3.16):

$$h_{opt} = \left(\frac{1}{R_1}\right)^{\frac{1}{3}} n^{-\frac{1}{3}} = \left(\frac{1}{\sqrt{\pi} \int_0^1 \frac{Q''(q)^2}{Q'(q)^5} dq}\right)^{\frac{1}{3}} n^{-\frac{1}{3}}$$

Auch hier lässt sich die optimale Bandweite für die beiden ersten Testfälle und  $X$  standardnormalverteilt analytisch bestimmen:

Für die lineare Funktion gilt  $h_{opt} = 4^{\frac{1}{3}} \sigma n^{-1/3} \approx 1,59 \sigma n^{-1/3}$  (siehe Gleichung 3.10).

Für die quadratische Funktion ist  $Y$  chi-quadrat-verteilt mit  $\mu = 1$ . Die Dichtefunktion ist dann

$$p_Y(y) = \begin{cases} 0 & t < 0 \\ \frac{1}{2} \frac{\sqrt{2} e^{-\frac{1}{2}t}}{\sqrt{\pi} \sqrt{t}} & \text{sonst} \end{cases} \quad (8.1)$$

und

$$p'_Y(y) = \begin{cases} 0 & t < 0 \\ \text{undefiniert} & t = 0 \\ -\frac{1}{4} \frac{\sqrt{2} e^{-1/2 t}}{\sqrt{\pi} t^{3/2}} - \frac{1}{4} \frac{\sqrt{2} e^{-1/2 t}}{\sqrt{\pi} \sqrt{t}} & 0 < t \end{cases} \quad (8.2)$$

Dies ergibt:

$$\int p'_Y(y)^2 dy = \frac{1}{4\pi} \left(-\frac{1}{4} \int \frac{e^{-t}}{t} dt - \frac{3e^{-t}}{4t} - \frac{e^{-t}}{4t^2}\right) \quad (8.3)$$

Für  $R_1$  erhält man schließlich:

$$R_1 = \infty$$

Da  $R_1$  im Nenner von  $h_{opt}$  steht, sollte die Bandweite möglichst klein gewählt werden. Diese beiden Beispiele zeigen, dass ein guter Kompromiss zwischen 0 und 1,59 liegen wird.

### 8.3.2 Numerische Simulation

$h_{opt}$  für den linearen Fall legt den Ansatz

$$h = a \cdot \hat{\sigma} \cdot n^{-1/3}$$

für die Bandweite nahe, wobei  $\hat{\sigma}$  die Standardabweichung der Stichprobe ist. Die Vorüberlegungen haben gezeigt, dass  $a = 1,59$  für Testfall 1 optimal ist, während für Testfall 2  $h$  möglichst klein sein sollte. Zur genaueren Untersuchung werden für  $a$  nacheinander alle Elemente der Menge  $A = \{0, 2; 0, 5; 1; 1, 5; 1, 59; 2\}$  eingesetzt. Als Fehlermaß wird als Annäherung an MISE benutzt:

$$\widehat{\text{MISE}} = \frac{1}{99} \sum_{q=0,01}^{0,99} (\hat{Q}(q) - y_q)^2$$

Tab. 8.1: Übersicht der durchgeführten Berechnungen zur Wahl der Bandweite

Testfall	$n$	Abbildung	$n$	Abbildung	$n$	Abbildung
1	100	12.1	1000	12.11	10000	12.12
2	100	12.2	1000	12.14	10000	12.15
3	100	12.3	1000	12.16	10000	12.17
4	100	12.4				
5	100	12.5				
6	100	12.6				
7	100	12.7				
8	100	12.8				
zsf	64	12.9				
B-Säule	100	12.10				
pars10crits19	1024	12.13				

Für jede Kombination aus Testfall und Versuchsplan wird  $\widehat{MISE}$  berechnet. Für  $n = 100$  werden für die zufallsabhängigen Versuchspläne MC, LHS und SS jeweils 100 verschiedene Stichproben gezogen und der Erwartungswert von  $\widehat{MISE}$  gebildet. Um zu kontrollieren, ob sich dasselbe Bild auch für größere Stichproben ergibt, werden die gleichen Berechnungen noch einmal für  $n = 1000$  und  $10000$  durchgeführt. Da es ein Kontrollrechnung ist, werden jeweils nur zwei Stichproben gezogen. Die Ergebnisse für MC und SS dürfen daher nicht verallgemeinert werden.

Die daraus resultierenden mittleren quadratischen Integrationsfehler sind im Kapitel 12 abgebildet. Eine Übersicht der durchgeführten Berechnungen und die detaillierten Referenzen zu den Abbildungen gibt Tabelle 8.1.

## Auswertung

Bei der Auswertung wird zwischen einer und mehreren Dimensionen unterschieden. In Tabelle 8.2 sind jeweils die optimalen Bandweiten für die verschiedenen Testfälle einer Dimension aufgeführt.

Tab. 8.2: Übersicht der optimalen Bandweiten für eine Dimension

Testfall	MC	HA	CSS	SS/ LHS
1	1,5	0,5	0,2	0,5
2	0,5	0,5	0,5	0,5
3	1	1	0,5	0,5

Die numerisch bestimmte optimale Bandweite hängt für Testfall 1 und  $n = 100$  vom Versuchsplan ab: Für CSS ist 0,2 optimal, für SS und HA 0,5 und für MC 1,5. Dies lässt sich so interpretieren: Je „ungleichmäßiger“ ein Versuchsplan die Stichprobe verteilt, desto größer sollte die Bandweite sein.  $\widehat{MISE}$  ändert sich für MC im Bereich zwischen 0,5 und 1,59 nur wenig.

Für Testfall 2 und  $n = 100$  ist 0,5 die beste Wahl für alle Versuchspläne.

Im dritten Testfall ist für die gleichmäßigen Versuchspläne CSS und SS  $h = 0,5$  optimal, für die ungleichmäßigeren MC und HA  $h = 1$ .

Tab. 8.3: Übersicht der optimalen Bandweiten bei unterschiedlichen Stichproben-  
größen

Plan <i>n</i>	Testfall 1			Testfall 2			Testfall 3		
	100	1000	10000	100	1000	10000	100	1000	10000
MC	1,5	2	1	0,5	0,2	0,2	1	2	1,5
SS	0,5	0,2	0,2	0,5	0,2	0,2	0,5	0,2	0,2
HA	0,5	1	0,2	0,5	0,2	0,2	1	1	1
CSS	0,2	0,2	0,5	0,5	0,2	0,2	0,5	0,2	0,2

Tabelle 8.3 zeigt das Ergebnis der Kontrollrechnungen für  $n = 1000$  und  $10000$ . Beim Testfall 1 und  $n = 1000$  ist  $a = 0,2$  für SS und CSS am günstigsten, für HA  $a = 1$  und für MC  $a = 2$ . Bei  $10000$  Punkten bleibt  $a = 0,2$  für SS die beste Wahl, ebenso für HA. Für CSS sollte  $a = 0,5$  und für MC  $a = 1$  gewählt werden. Der für sehr große Stichproben theoretisch beste Wert von  $1,59$  wird nicht erreicht.

Im zweiten Testfall und  $n = 1000$  oder  $10000$  ist für alle Versuchspläne  $a = 0,2$  die beste Wahl. Bei sehr großen Stichproben wird die optimale Schrittweite kleiner, da sie sich der theoretisch besten Schrittweite, hier „möglichst klein“ annähert.

Insgesamt lässt sich keine in allen Fällen beste Wahl für  $a$  identifizieren. Als Kompromiss wird **bei den folgenden Rechnungen  $a = 0,5$  für eine Dimension zugrunde gelegt.**

Tab. 8.4: Vergleich der Versuchspläne für eine Dimension

Testfall	<i>n</i>	$\widehat{MISE}$ der Versuchspläne
1	100	CSS < SS < HA « MC
1	1000	SS < CSS < HA « MC
1	10000	SS < HA < CSS « MC
2	100	HA < CSS < SS « MC
2	1000	CSS < SS < HA « MC
2	10000	CSS ~ SS < HA « MC
3	100	CSS < SS < HA < MC
3	1000	CSS < SS < HA < MC
3	10000	CSS ~ SS < HA < MC

Die Grafiken im Anhang zeigen, dass die Reihenfolge der Versuchspläne bzgl.  $\widehat{MISE}$  für jede Wahl von  $h$  gleich ist. Sie ist in Tabelle 8.4 für die eindimensionalen Testfälle in Abhängigkeit von der Stichprobengröße aufgeführt. MC hat immer den größten Fehler. CSS und SS sind meistens besser als HA. Die Reihenfolge bleibt bei steigender Stichprobengröße erhalten, nur bei Testfall 1 und  $n = 1000$  ist bei geringer Bandweite CSS besser und bei größerer HA.

In höheren Dimensionen schwankt die optimale Wahl für  $h$  zwischen  $0,2$  und  $1,5$ . Während für die schiefen Verteilungen der Testfälle 5 und 6 kleinere Bandweiten optimal sind, benötigen die eher symmetrischen Verteilungen der Testfälle 4, 7 und Zweistab-Fachwerk eher größere Bandweiten. Die Verteilungen der Testfälle 1 und 8 sind normalverteilt. Sie ergeben bis auf MC auch die gleichen optimalen Bandweiten für  $h$ . Ein Blick auf die Abb. 12.4 bis 12.10 zeigt, dass die Unterschiede zwischen  $a = 0,5$  und  $a = 1$  gering sind.

Tab. 8.5: Übersicht der optimalen Bandweiten für mehrere Dimensionen

Testfall	MC	HA	CSS	SS	LHS
4	1	0,5	1	0,5	1
5	1	0,5	0,5	0,5	0,5
6	1	0,5	0,2	0,2	1
7	1	1	0,5	1	1
8	1	0,5	0,2	0,5	1
Zweistab-Fachwerk	1,5	1	1,5	0,5	1
B-Säule	1	0,5	0,5	1	1
pars10crits19	1,5	0,2	2	0,2	1

### Daher wird für mehrere Dimensionen $a = 1$ gewählt.

Ein weiterer Grund für diese Wahl von  $a$  liegt in der numerischen Stabilität der Berechnung. Die Berechnung wird instabil, wenn  $\hat{F}_n(t)$  sehr stark ansteigt und die Bandweite klein ist. Die Bestimmung eines geeigneten Startwertes für die Lösung von  $\hat{F}_n(t) - q = 0$  ist in diesem Fall schwierig. Oft führen erst mehrere Versuche zu einer Lösung. Dieses Problem trat in der numerischen Simulationen häufig bei  $a = 0,2$  und  $0,5$  auf. Da durch die mehrmaligen Versuche der Rechenaufwand erheblich vergrößert wird, eignet sich diese Schrittweite nicht für einen effizienten Algorithmus. Ab  $a = 1$  ist die Berechnung stabil.

Tab. 8.6: Vergleich der Versuchspläne für  $a = 1$ 

Testfall	mse der Versuchspläne
4	HA < LHS < SS < CSS < MC
5	HA < LHS < CSS < MC < SS
6	HA < LHS < CSS < SS < MC
7	HA < CSS < LHS < SS < MC
8	HA < CSS < LHS < SS < MC
Zweistab-Fachwerk	HA < CSS < LHS < MC < SS
B-Säule	SS < CSS < HA < LHS < MC
pars10crits19	LHS < MC < HA < CSS < SS

Unabhängig von der Bandweite ergibt sich für die meisten Testfälle die gleiche Reihenfolge der Versuchspläne. (Ausnahmen bilden nur LHS für Testfall 7 und 8 sowie CSS für zsf.) In mehreren Dimensionen ist HA immer der beste Versuchsplan. Die folgenden Ränge hängen vom Testfall ab. Insgesamt sind LHS und CSS besser als SS und MC. Ein genauerer Vergleich, der auch die anderen Schätzer einschließt, folgt im nächsten Kapitel.

## 8.4 Vergleich der Quantilschätzer und Versuchspläne

Ziel der numerischen Untersuchung dieses Kapitels ist es, die beste Kombination aus Quantilschätzer und Versuchsplan für die Berechnung der übertragenen Verteilung für einen Punkt zu bestimmen. Betrachtet werden die Versuchspläne MC, SS, CSS, LHS und HA und die Quantilschätzer EMP, SV3, KDE mit der im Kapitel 8.3



Tab. 8.7: Übersicht der durchgeführten Berechnungen zum Vergleich der Versuchspläne und Schätzer

Testfall	$n$	Abbildung	$n$	Abbildung	$n$	Abbildung
1	100	12.18	1000	12.28	10000	12.39
2	100	12.19	1000	12.29	10000	12.40
3	100	12.20	1000	12.30	10000	12.41
4	100	12.21	1024	12.31	10000	12.42
5	100	12.22	1024	12.32	10000	12.43
6	100	12.23	1024	12.33	10000	12.44
7	100	12.24	1024	12.34	10000	12.45
8	100	12.25	1024	12.35	10000	12.46
zsf	64	12.26	729	12.36	4096	12.47
B-Säule			1024	12.37	10000	12.48
pars10crits19			1024	12.38	59049	12.49

bestimmten Bandweite sowie HD und P2. Es werden wieder die Testfälle aus Kapitel 8.2 benutzt.

Für eine Dimension werden  $n = 100$ ,  $n = 1000$  und  $n = 10000$  miteinander verglichen. Da für SS und LHS nur Stichprobengrößen von  $a^k$  mit  $a \in \mathbb{N}$  möglich sind, werden für zwei Dimensionen  $n = 100$ ,  $n = 1024$  und  $n = 10000$  berechnet. Beim 6-dimensionalen Testfall zsf beträgt die Stichprobengröße  $n = 2^6 = 64$ ,  $n = 3^6 = 729$  und  $n = 4^6 = 4096$ . Beim 10-dimensionalen Testfall pars10crits19 wird  $n = 2^{10} = 1024$  betrachtet. Der Fehler für  $n = 3^{10} = 59049$  lag unterhalb der Messgenauigkeit für den exakten Wert. Daher wird diese Berechnung in den Auswertungen nicht berücksichtigt.

Bei  $n = 100$  werden für MC, SS und LHS jeweils 100 Stichproben berechnet und ihr Erwartungswert gebildet, darüber 2 Stichproben. Tabelle 8.7 listet alle durchgeführten Simulationen auf und enthält die Verweise auf die Einzelgrafiken im Kapitel 12.

Als Fehlermaß wird wieder  $\widehat{\text{MISE}}$  benutzt.

### 8.4.1 Vergleich der Versuchspläne

Schon ein kurzer Blick auf die Abbildungen 12.18 bis 12.46 im Anhang zeigt, dass der Fehler für den Schätzer P2 oft wesentlich größer ist als für die übrigen Quantilschätzer. Es ist offensichtlich, dass P2 nicht der beste Schätzer sein wird. Deswegen wird er von den nachfolgenden Betrachtungen ausgenommen.

Betrachtet man nun nur noch die übrigen Schätzer fällt auf, dass sich die Linien für die einzelnen Versuchspläne - bis auf wenige Ausnahmen - nicht schneiden, sondern oft nahezu parallel verlaufen. Die Güte eines Versuchsplans ist somit weitgehend unabhängig vom Quantilschätzer. Daher lässt sich für jede Kombination aus Testfall und Stichprobengröße ein Rangfolge angeben, die unabhängig vom Quantilschätzer ist. Die Tabelle 8.8 gibt das Ranking für eine Dimension und Tabelle 8.9 das für höhere Dimensionen wieder.

Diese Übersicht zeigt, dass die Güte des Versuchsplans selbst bei Beschränkung auf einen Testfall abhängig von der Stichprobengröße ist. So ist z.B. bei der B-Säule

CSS für  $n = 100$  der beste Versuchsplan, für  $n = 1024$  aber HA.

Es gibt keinen Versuchsplan, der in allen untersuchten Kombinationen am besten abschneidet. Dies war auch nicht zu erwarten, da die deterministischen Pläne CSS und HA zufällig ungünstig verteilt sein könnten.

Tab. 8.8: Vergleich der Versuchspläne (außer P2) für eine Dimension

Testfall	n	$\widehat{\text{MISE}}$ der Versuchspläne
1	100	CSS $\approx$ SS < HA « MC
1	1000	SS < CSS < HA « MC
1	10000	CSS $\approx$ SS < HA « MC
2	100	HA < CSS < SS < MC
2	1000	CSS $\approx$ SS < HA < MC
2	10000	CSS $\approx$ SS < HA < MC
3	100	CSS < SS < HA < MC
3	1000	CSS < SS < HA < MC
3	10000	CSS < SS < HA < MC

In einer Dimension ist CSS - bis auf Testfall 1 bei  $n = 1000$  und Testfall 2 bei  $n = 100$  - der beste Versuchsplan. Fast genauso gut ist SS, mit Abstand folgt HA. In allen Fällen ist MC der Versuchsplan mit dem größten Fehler. Für eine Dimension bringt die gleichmäßige Verteilung der Stichprobenwerte, die CSS und SS vornehmen, Vorteile, unabhängig von der Form der übertragenen Dichtefunktion.

Tab. 8.9: Vergleich der Versuchspläne (außer P2) für mehrere Dimensionen

Testfall	n	$\widehat{\text{MISE}}$ der Versuchspläne
4	100	HA < LHS < SS < CSS < MC
4	1024	HA < LHS < SS < CSS $\approx$ MC
4	10000	HA < SS < LHS < MC < CSS
5	100	HA < LHS < CSS < MC $\approx$ SS
5	1024	HA < LHS < MC < SS < CSS
5	10000	HA < LHS $\approx$ MC < SS $\approx$ CSS
6	100	HA < LHS < CSS (außer EMP) < SS < MC
6	1024	HA < LHS < MC < CSS < SS
6	10000	HA < LHS < SS < MC < CSS
7	100	HA < CSS < LHS < SS < MC
7	1024	HA < CSS $\approx$ SS < LHS < MC
7	10000	HA < CSS (außer EMP) < SS < LHS < MC
8	100	HA < CSS < SS $\approx$ LHS < MC
8	1024	HA < CSS < LHS < SS < MC
8	10000	CSS < SS < HA < LHS < MC
zsf	64	HA < CSS < LHS (außer SV3) < MC < SS
zsf	729	HA < LHS < MC < CSS < SS
zsf	4096	HA < LHS < MC < CSS < SS
B-Säule	100	CSS < SS < HA < LHS < MC
B-Säule	1024	HA < CSS < SS < LHS < MC
B-Säule	10000	HA < SS < CSS < LHS < MC
pars10crits19	1024	LHS < MC < HA < CSS < SS

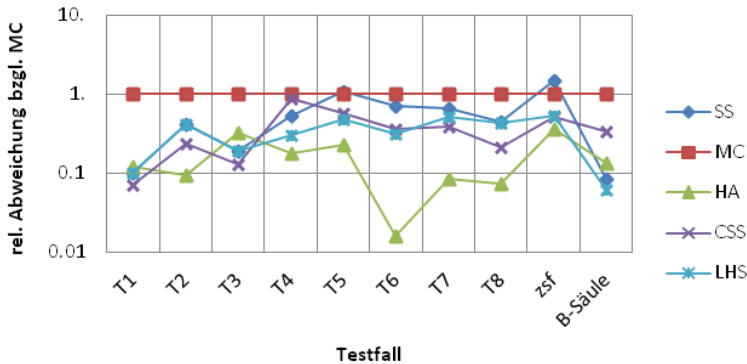


Abb. 8.27: Übersicht aller Testfälle für HD,  $n = 100$

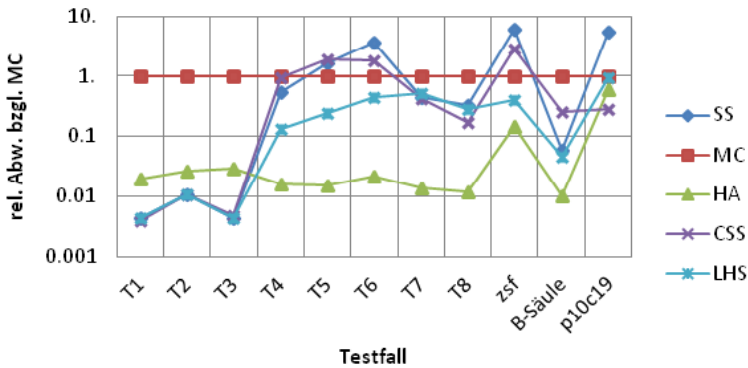


Abb. 8.28: Übersicht aller Testfälle für HD,  $n = 10000$

In höheren Dimensionen ist HA der beste Versuchsplan. Er weist bis auf die Kombinationen Testfall 8 mit  $n = 10000$ , B-Säule mit  $n = 100$  und pars10crits19 mit  $n = 1024$  den kleinsten Fehler auf. Der schlechteste Versuchsplan ist meistens MC. Er liegt in 12 von 22 Fällen am Ende der Rangfolge und ist nur ein Mal zweitbest. Eine Reihenfolge der verbliebenen Versuchspläne lässt sich schwer erstellen: LHS ist 11 Mal mindestens zweitbest. Plan, CSS 9 Mal und SS 4 Mal. Besonders gut schneidet CSS bei den eher symmetrischen Testfällen 7, 8 und B-Säule ab. LHS hat gute Werte bei den schiefen Testfällen 4, 5 und 6. Obwohl auch zsf und pars10crits19 eine annähernd normalverteilte übertragene Dichtefunktion haben, ist dort LHS besser als CSS. Dies liegt an der höheren Dimension der Testfälle. Auch bei der Verwendung eines Versuchsplans zur Erstellung eines Metamodells ist LHS in höheren Dimensionen CSS überlegen.

Einen zusätzlichen quantitativen Vergleich bietet die Abbildung 8.27. Dort ist die prozentuale Abweichung gegenüber MC für jeden Testfall jeweils für den Harrell-Davis-Schätzer und  $n = 100$  dargestellt. Für diesen Schätzer verbessert die Haltonfolge die Schätzung um mindestens 60 Prozent. Für T6 beträgt sie sogar 98 Prozent.

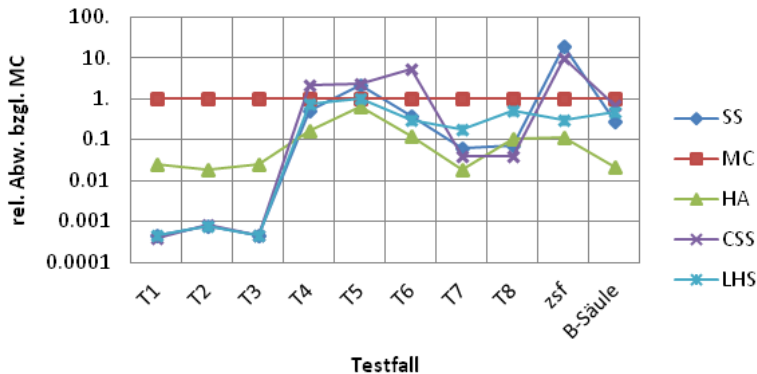


Abb. 8.29: Übersicht aller Testfälle für HD,  $n = 10000$

Abbildung 8.28 zeigt den gleichen Sachverhalt für  $n = 1000$ . Dort beträgt die relative Verbesserung durch HA in den Testfällen bis zu zwei Dimensionen mindestens 98 Prozent. Bei den höherdimensionalen Testfällen fällt der Gewinn kleiner aus: 85% für das Zweistabfachwerk und 41% für pars10crits19.

Für  $n = 10000$  ist HA um mindestens 38 Prozent besser als MC (Abbildung 8.29). Kein anderer Versuchsplan bringt über alle untersuchten Stichprobengrößen und alle Testfälle eine solch konstante Verbesserung.

Über alle Testfälle und alle Stichprobengrößen betrachtet ist damit die Haltonfolge der geeignetste Versuchsplan.

**Daher wird in der Methodik zur robusten Optimierung die Haltonfolge als Versuchsplan für die Quantilschätzung benutzt.**

Sie hat gegenüber LHS, CSS und SS den zusätzlichen Vorteil, dass sie beliebig erweiterbar ist.

#### 8.4.2 Vergleich der Quantilschätzer bzgl. der Haltonfolge

Im folgenden werden die verschiedenen Quantilschätzer untereinander verglichen. Da schon die Haltonfolge als bester Versuchsplan identifiziert wurde, beziehen sich die folgenden Auswertungen nur noch auf sie.

Die Tabellen 8.10 und 8.11 zeigen die Rangfolge der Quantilschätzer für diesen Versuchsplan.

Der Schätzer P2 schneidet in 18 von 31 Fällen am schlechtesten ab, er erreicht nur einmal den zweiten Platz. EMP ist 19 Mal schlechter als die drei übrigen. Für HD, KDE und SV3 lässt sich aus den Rangfolgen keine Gesetzmäßigkeit erkennen. Eine bessere Übersicht bietet sich, wenn man zu allen Schätzern die relative Abweichung gegenüber EMP für alle Testfälle vergleicht. Abbildung 8.30 zeigt dies für  $n = 100$ , Abb. 8.31 für  $n = 1000$  und Abb. 8.32 für  $n = 10000$ .

Der Vergleich für  $n = 100$  zeigt:

- P2 ist bis auf zwei Testfälle schlechter als EMP und zwar um bis zu 450 Prozent.

Tab. 8.10: Vergleich der Schätzer bzgl. HAL für eine Dimension

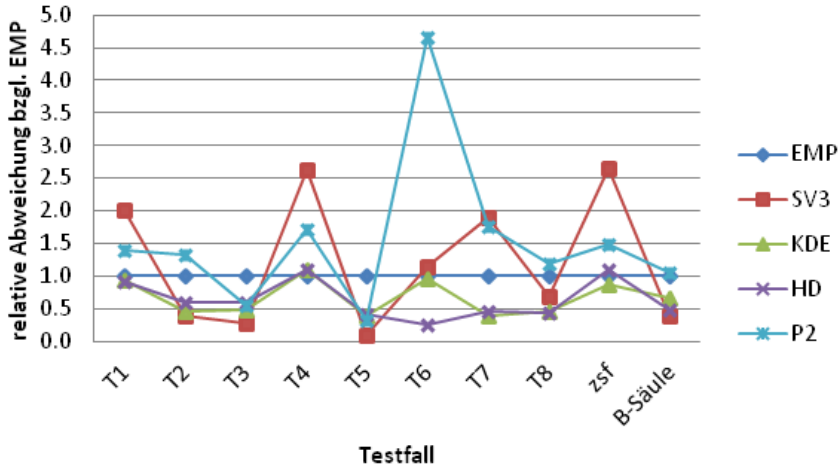
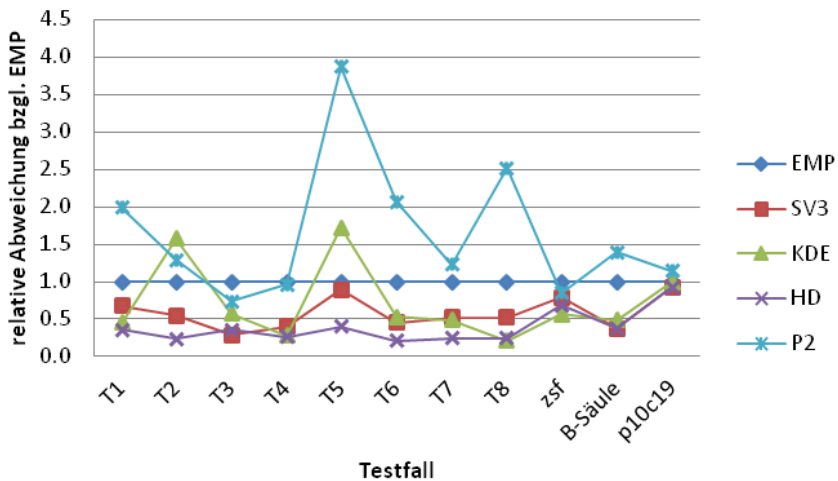
Testfall	$n$	$\widehat{\text{MISE}}$ der Schätzer
1	100	KDE = HD < EMP < P2 < SV3
1	1000	HD < KDE < SV3 < EMP < P2
1	10000	KDE < SV3 < HD < EMP < P2
2	100	SV3 < KDE < HD < EMP < P2
2	1000	HD < SV3 < EMP < P2 < KDE
2	10000	SV3 < HD < P2 < EMP < KDE
3	100	SV3 < KDE < P2 < HA < EMP
3	1000	SV3 < HD < KDE < P2 < EMP
3	10000	HD < KDE < SV3 < EMP < P2

Tab. 8.11: Vergleich der Schätzer bzgl. HAL für mehrere Dimensionen

Testfall	$n$	$\widehat{\text{MISE}}$ der Schätzer
4	100	EMP < KDE = HD < P2 < SV3
4	1024	HD < KDE < SV3 < P2 < EMP
4	10000	SV3 < KDE < HD < P2 < EMP
5	100	SV3 < P2 < KDE = HD < EMP
5	1024	HD < SV3 < EMP < KDE < P2
5	10000	EMP < HD < SV3 < KDE < P2
6	100	HD < KDE = EMP < SV3 < P2
6	1000	HD < SV3 < KDE < EMP < P2
6	10000	EMP < HD < KDE < SV3 < P2
7	100	KDE < HD < EMP < P2 < SV3
7	1000	HD < KDE < SV3 < EMP < P2
7	10000	SV3 < KDE < HD < EMP < P2
8	100	HD < KDE < SV3 < EMP < P2
8	1000	KDE < HD < SV3 < EMP < P2
8	10000	KDE < SV3 < HD < EMP < P2
zsf	100	KDE < EMP < HD < P2 < SV3
zsf	729	KDE < HD < SV3 < P2 < EMP
zsf	4096	KDE < HD < SV3 < EMP < P2
B-Säule	100	KDE < EMP < HD < P2 < SV3
B-Säule	1024	SV3 < HD < KDE < EMP < P2
B-Säule	10000	SV3 < KD < HD < EMP < P2
pars10crits19	1024	SV3 < HD < KDE < EMP < P2

- SV3 ist fünf Mal bester Schätzer, aber in fünf Fällen schlechter als EMP. Auch hier ist die Verschlechterung deutlich. Wünschenswert wäre allerdings ein Schätzer, der gleichmäßig in allen Situationen eine Verbesserung erzielt.
- KDE ist EMP immer mindestens gleichwertig. Es verbessert die Schätzung oft um 50 Prozent.
- HD ist ähnlich gut wie KDE, bei T6 sogar wesentlich besser. Er ist zwei Mal um 10 Prozent schlechter als EMP.

Für  $n = 1000$  ergibt sich:

Abb. 8.30: Vergleich aller Testfälle für die Haltonfolge,  $n = 100$ Abb. 8.31: Vergleich aller Testfälle für die Haltonfolge,  $n = 1000$ 

- P2 ist wieder bis auf zwei Testfälle schlechter als EMP und zwar um bis zu 400 Prozent.
- KDE ist zwei Mal schlechter als EMP.
- SV3 ist immer besser als EMP.
- HD liefert ebenfalls immer bessere Ergebnisse als EMP. Er ist SV3 und KDE mindestens gleichwertig. In sieben der elf Fälle ist seine Schätzung die genaueste.

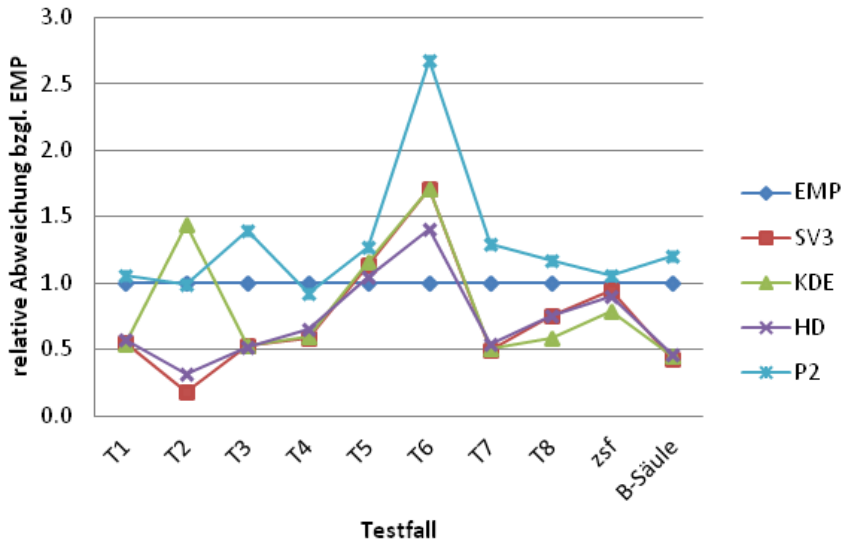


Abb. 8.32: Vergleich aller Testfälle für die Haltonfolge,  $n = 10000$

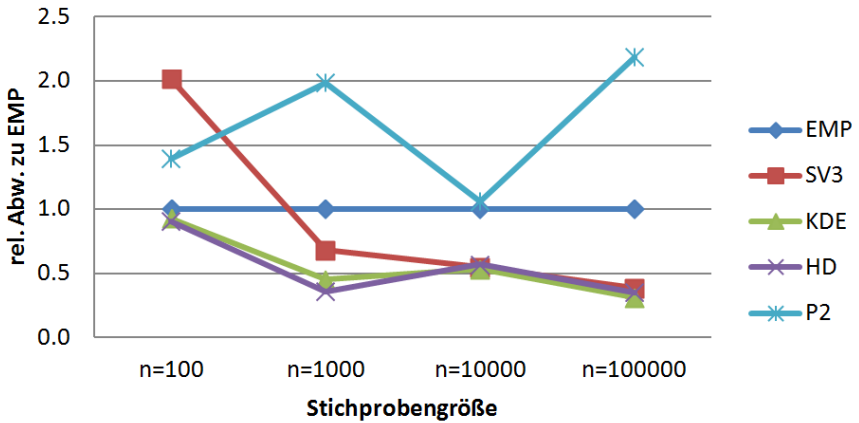


Abb. 8.33: Vergleich der Stichprobengröße für Testfall 1 und Haltonfolge

Für  $n = 10000$  gilt:

- P2 ist in acht von zehn Fällen der schlechteste Schätzer. Die Abweichung gegenüber EMP beträgt bis zu 270 Prozent.
- Die Schätzung mit KDE ist in drei Fällen der von EMP unterlegen. In den übrigen Situationen ist sie deutlich besser.
- Die Ergebnisse von SV3 sind denen von KDE sehr ähnlich. Eine Ausnahme

bildet T2, wo SV3 wesentlich besser ist.

- HD liefert ähnlich gute Schätzungen wie SV3.

Über alle Testfälle und Stichprobengrößen betrachtet ist HD der konstanteste Schätzer. Er verbessert die Schätzung gegenüber EMP um durchschnittlich etwa 50 Prozent. Zudem ist seine Berechnung nur unwesentlich aufwändiger, wenn die Gewichte vorher bestimmt wurden. KDE hat Schwierigkeiten bei den extrem schiefen Verteilungen von T2, T5 und T6. Hier ist die Wahl der Bandweite zu groß. Eine weiterer Nachteil ist seine aufwändige Berechnung. Auch SV3 liefert im Durchschnitt bessere Schätzungen als EMP. Er hat Schwächen bei kleinen Stichprobengrößen. P2 ist EMP eindeutig unterlegen.

**Daher wird im folgenden der Harrell-Davis-Schätzer benutzt.**

In Kapitel 3.3 wurde gezeigt, dass der theoretische Vorteil, den die Schätzer KDE und HD gegenüber EMP besitzen, asymptotisch verschwinden wird. Die bisherigen Untersuchungen zeigen, dass dies bei  $n = 10000$  noch nicht der Fall ist. Für den Testfall 1 wurde auch die Stichprobengröße  $n = 100.000$  berechnet. Die drei Schätzer HD, KDE und SV3 verbessern auch für diese Stichprobengröße die Schätzung deutlich (vgl. Abb. 8.33).

Eine Verwendung des Harrell-Davis-Schätzers bis zur Stichprobengröße von 100.000 ist daher sinnvoll.

### 8.4.3 Vergleich von Harrel-Davis-Schätzer und Haltonfolge zu Standard-schätzer und Monte-Carlo-Verfahren

In den beiden vorangegangenen Unterkapiteln wurden Versuchsplan und Quantilschätzer getrennt voneinander betrachtet. Hier werden nun beide zusammen untersucht. Es soll eine Aussage darüber getroffen werden, wie groß der Vorteil ist, der gegenüber dem Standardverfahren -MC und EMP- erzielt wird und ob der Versuchsplan oder der Schätzer den größten Beitrag dazu leistet.

Dazu wird die relative Abweichung bezüglich MC und EMP für die Kombinationen MC und HD, HAL und EMP und schließlich HAL und HD über alle Testfälle betrachtet. Abbildung 8.34 zeigt dies für die Stichprobengröße  $n = 100$ , Abb. 8.35 für  $n = 1000$  und Abb. 8.36 für  $n = 10000$ .

Die alleinige Verwendung des Harrell-Davis-Schätzers beim Standardversuchsplan Monte Carlo bringt bei kleiner Stichprobengröße ( $n = 100$ ) eine Verbesserung von ca. 20%. Bei größeren Stichproben beträgt sie unter 10%. Benutzt man die Haltonfolge und verwendet den Standardschätzer EMP wird bei  $n = 100$  eine Verbesserung von 80% erreicht. Bei größeren Stichproben sind sogar 90% möglich, wobei es Ausnahmen gibt. Die Kombination des Harrell-Davis-Schätzer mit der Haltonfolge bringt nur noch eine Verbesserung von wenigen Prozenten, vor allem bei großen Stichproben. Da die Konvergenzgeschwindigkeit bezüglich der Stichprobengröße sehr gering ist ( $1/\sqrt{n}$ ), ist eine kleine Verbesserung der Schätzung auf anderem Wege nur durch eine erhebliche Vergrößerung der Stichprobe zu erreichen. Die Verwendung des Harrell-Davis-Schätzer ist daher sinnvoll. Die theoretischen Ergebnisse zur Konvergenzrate der Schätzer zeigen, dass für genügend große Stichproben EMP der beste



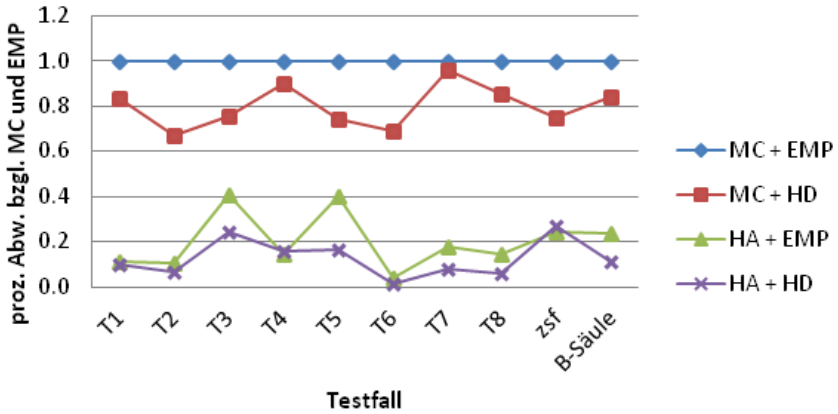


Abb. 8.34: prozentualer Vergleich bezüglich Monte Carlo Versuchsplan und empirischem Schätzer,  $n = 100$

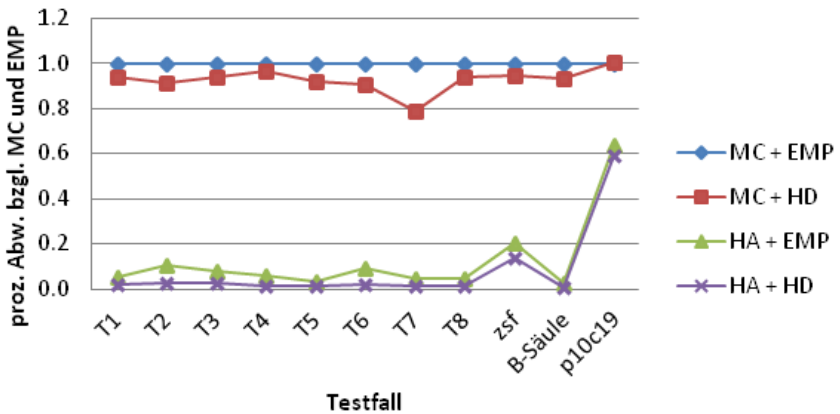


Abb. 8.35: prozentualer Vergleich bezüglich Monte Carlo Versuchsplan und empirischem Schätzer,  $n = 1000$

Schätzer sein wird. Die numerische Untersuchung ergibt, dass dies für Stichproben-  
größen bis zu 100.000 noch nicht der Fall ist. Bei Verwendung wesentlich größerer  
Stichproben sollte EMP benutzt werden.

Den größten Beitrag zur Verbesserung einer Schätzung leistet die Verwendung der  
Haltonfolge. Da die Konvergenzrate der Haltonfolge dimensionsabhängig ist, wird  
der Vorteil mit steigendem  $d$  verschwinden. Bis zu 6 Dimensionen ist dies noch nicht  
der Fall.

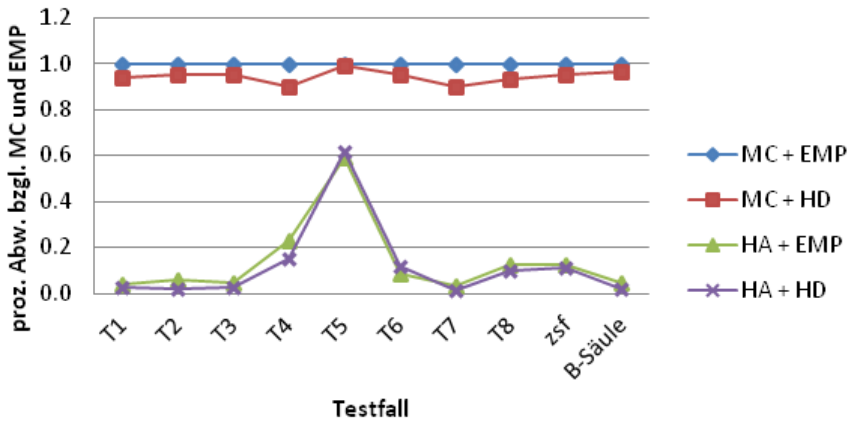


Abb. 8.36: prozentualer Vergleich bezüglich Monte Carlo Versuchsplan und empirischem Schätzer,  $n = 10000$

## 8.5 Test des iterativen Algorithmus

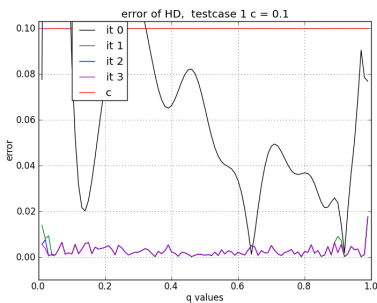


Abb. 8.37: Testfall 1

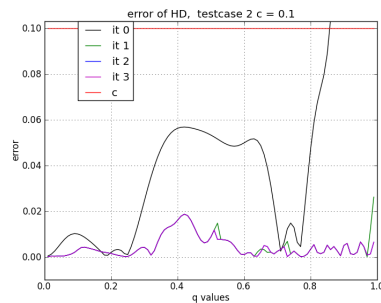


Abb. 8.38: Testfall 2

Die Grafiken 8.37 bis 8.44 zeigen für jeden Testfall die Größe des Fehlers bei jedem Iterationsschritt. Die vorgegebene Fehlergrenze wird für alle Testfälle eingehalten. Der Algorithmus ist richtig ausbalanciert: Er erkennt, wenn der Fehler zu groß ist, z. B. bei Testfall 2 und  $q$  nahe 1. Die Erhöhung der Stichprobengröße erfolgt im richtigen Verhältnis: Sie bringt eine deutliche Verbesserung der Genauigkeit ohne bei Extremfällen weitere Erhöhungen überflüssig zu machen.

In allen Testfällen wird die Fehlergrenze schon nach dem zweiten Iterationsschritt eingehalten. Tests mit weniger scharfen Fehlerschätzungen im Algorithmus führten aber in Einzelfällen zu schlechten Schätzungen. In Einzelfällen durchläuft er alle vier Iterationsschritte, z.B. bei Testfall 1 für kleine  $q$ . Der unregelmäßige Verlauf des tatsächlichen Fehlers im ersten Iterationsschritt zeigt wie schlecht die Anfangsschätzung ist.

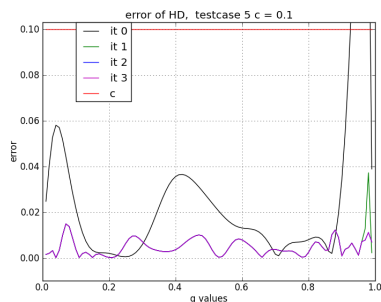


Abb. 8.39: Testfall 3

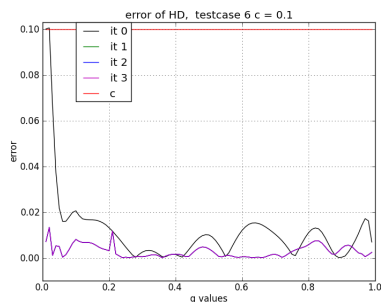


Abb. 8.40: Testfall 4

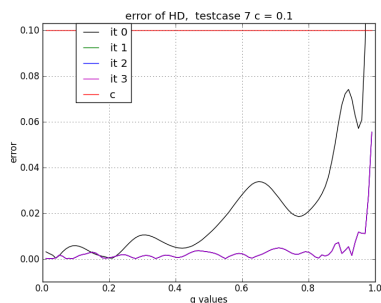


Abb. 8.41: Testfall 5

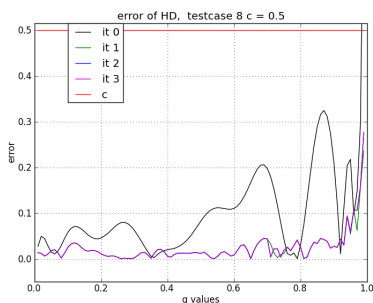


Abb. 8.42: Testfall 6

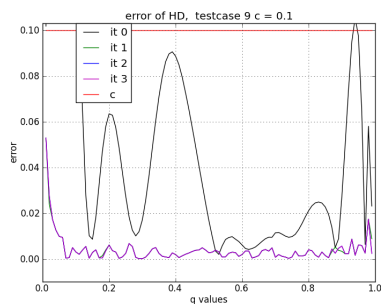


Abb. 8.43: Testfall 7

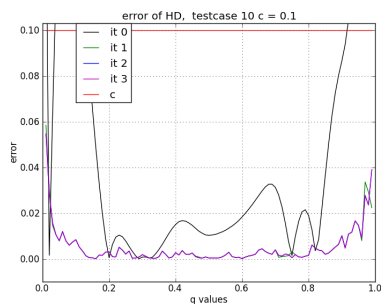


Abb. 8.44: Testfall 8

Der Algorithmus ist konservativ ausgelegt, denn der tatsächliche Schätzfehler liegt immer deutlich unter der vorgegebenen Grenze.



## Kapitel 9

---

# Anwendungen der neu entwickelten globalen Methodik

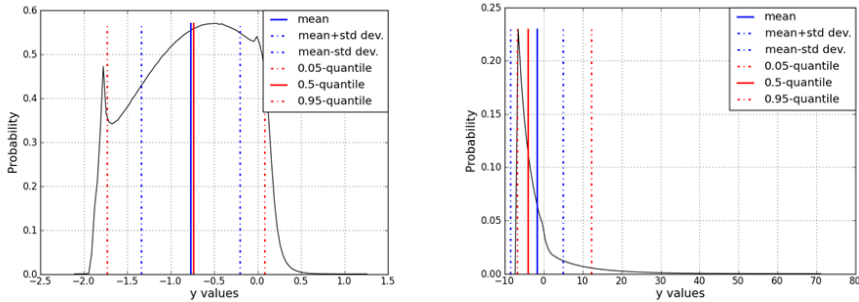
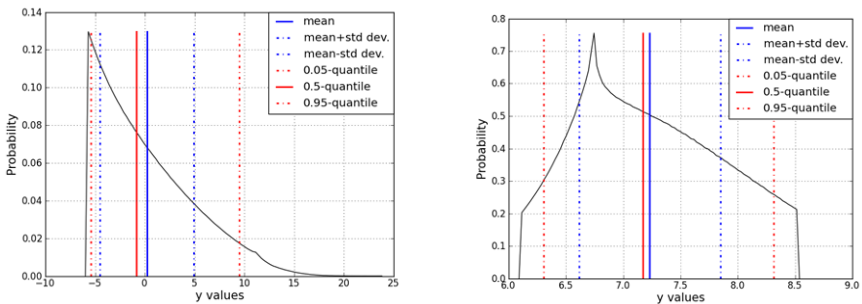
---

In Kapitel 6 wurde die Methodik vorgestellt, die es ermöglicht, eine robuste Optimierung mit wenigen echten Funktionsauswertungen durchzuführen. Diese Situation liegt in der Praxis häufig vor: Viele naturwissenschaftliche Sachverhalte können über Computersimulationen analysiert werden, etwa Wetter- und Klimavorhersagen oder Verkehrs- und Logistik-Simulationen. Im Ingenieurbereich werden z.B. Finite-Elemente-Methoden (FEM), Strömungs- und Fließsimulationen durchgeführt. Solche Computerexperimente dauern oft sehr lange, so dass nur eine begrenzte Anzahl von ihnen im Entwicklungsprozess durchgeführt werden kann. Die entwickelte Methodik ist auf den Einsatz in diesen Gebieten zugeschnitten. Die beiden hier vorgestellten Anwendungsbeispiele kommen auch aus diesem Bereich: Die Simulation des Crashverhaltens von Kraftfahrzeugen durch FEM-Simulationen ist besonders anspruchsvoll. In Kapitel 9.1 wird eine robuste Optimierung des Crashverhaltens einer Pkw-B-Säule durchgeführt. Auch das zweite Beispiel `pars10crits19` stammt aus der Automobilindustrie (Kapitel 9.2).

### 9.1 Optimierung der B-Säule eines Pkw

Das Beispiel aus den Kapiteln 2 und 8.2.4 wird an dieser Stelle wieder aufgegriffen und erweitert. Betrachtet werden die beiden Ausgabevariablen *Intrusion* und *Geschwindigkeit*. Das Gewicht wird nicht hinzugenommen, damit die Ergebnisse der Optimierung nachvollziehbar bleiben.

Gesucht sind die robusten Optima für die beiden Ausgabevariablen *Geschwindigkeit* und *Intrusion*. Für die Eingabevariablen sind alle Parameterkombinationen zu bestimmen, die die besten Kompromisse aus Minimierung der Ausgabenvariablen und ihrer Schwankung darstellen. Dabei wird die in Kapitel 6 vorgestellte Vorgehensweise angewendet.

Abb. 9.1: Beispiele für Dichtefunktionen von *Intrusion*Abb. 9.2: Beispiele für Dichtefunktionen von *Geschwindigkeit*

Es werden 101 Simulationen durchgeführt, die gemäß dem Latin Hypercube Sampling verteilt sind. Hieraus werden die Metamodelle für die Ausgabevariablen erstellt. Sie wurden schon in Abb. 2.2 für *Intrusion* und in Abb. 8.23 für *Geschwindigkeit* abgebildet.

Für die Robustheitsuntersuchung wird angenommen, dass beide Eingangsvariablen normalverteilt sind mit der Standardabweichung  $0,1 \text{ mm}$ . In den Abbildungen 9.1 und 9.2 sind einige Beispiele von Dichtefunktionen der Zielvariablen angegeben. Sie illustrieren, dass die übertragene Verteilung unsymmetrisch und stark verzerrt sein kann. Für die Quantildifferenz wird  $q = 0,05$  gewählt, d.h. es wird das 0,95- und das 0,05-Quantil berechnet, ebenso der Median.

Gemäß der entwickelten Methodik wird der Algorithmus 1 mit  $c = 0,01$  für die Stützstellen des ersten Metamodells durchgeführt. Danach sind die übertragenen Quantile für 0,05, 0,5 und 0,95 an diesen Punkten bestimmt. Hiermit werden zwei Metamodelle erstellt:

Aus den Medianwerten  $\tilde{f}_m$  und aus den Differenzen der oberen und unteren Quantile  $\tilde{f}_{qd}$ . Sie werden benutzt, um die Paretofront für die vier Zielvariablen *Median Geschwindigkeit*, *Median Intrusion*, *Quantildifferenz Geschwindigkeit* und *Quantildifferenz Intrusion* zu berechnen. Dazu wird eine Stichprobe der Größe 20000 im gesamten Designraum verteilt und die zugehörigen Metamodellwerte bestimmt. Der Standardalgorithmus „Pareto-sort“ findet dann 1484 Pareto-optimale Punkte. Die

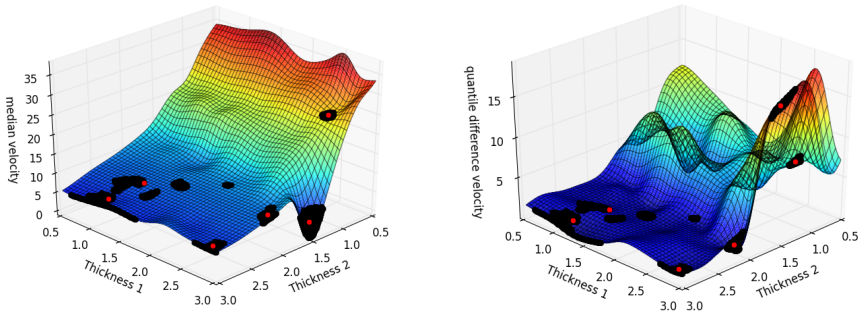


Abb. 9.3: Metamodelle für *Geschwindigkeit*: Median und Quantildifferenz, in schwarz robuste Optima, in rot die Cluster-Zentren

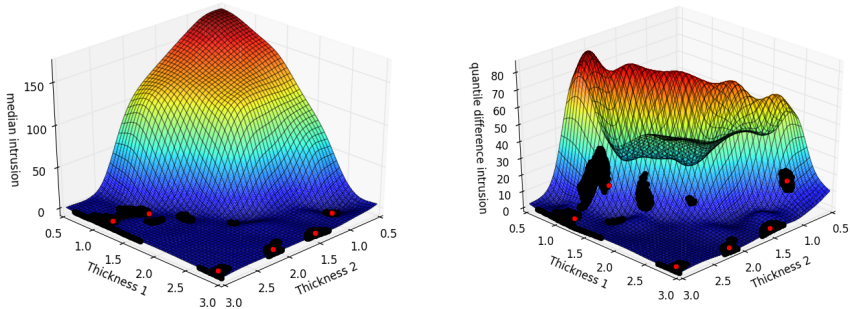


Abb. 9.4: Metamodelle für *Intrusion*: Median und Quantildifferenz, in schwarz robuste Optima, in rot die Cluster-Zentren

Abbildungen 9.3 und 9.4 zeigen die vier Metamodelle zusammen mit den Pareto-optimalen Punkten.

Die Median-Metamodelle weichen kaum von den ursprünglichen Modellen ab, sie „glätten“ den Kurvenverlauf ein wenig. Die Metamodelle für die Quantildifferenz nehmen kleine Werte für die flachen Bereiche des Originalmodells an und große bei starken Änderungen. So ist z.B. die Quantildifferenz für *Geschwindigkeit* im Bereich des spitzen Minimums bei (3; 1, 5) am größten. Bei einer Standardabweichung von 0,1 mm variiert die Verteilung im wesentlichen um +/- 0,3 mm. Dies reicht aus, um starke Schwankungen der Zielvariablen hervorzurufen. Bei beiden Kriterien liegt die maximale Quantildifferenz ungefähr bei der Hälfte des Wertebereiches. *Geschwindigkeit* und *Intrusion* reagieren sehr empfindlich auf Schwankungen der Eingangsparameter.

Für die Verteilung der Paretopunkte ist es typisch, dass sich robuste Bereiche bilden und nicht nur einzelne Punkte. Es gehören immer die deterministischen Minima der einzelnen Zielvariablen dazu. So hat z.B. der Median von *Intrusion* sein Minimum nahe 2,5 mm für *Thickness 1* und 1 mm für *Thickness 2*. Alle Metamodelle besitzen kleine Werte für *Thickness 1* zwischen 2 und 3 mm und für *Thickness 2* ebenfalls zwischen 2 und 3 mm. Robuste Bereiche befinden sich auch im Gebiet *Thickness*

$1 \in [0, 5; 1, 5]$  und *Thickness*  $1 \in [2; 3]$ , da dort Median und Quantildifferenz von *Geschwindigkeit* und der Median von *Intrusion* klein sind. Die Quantildifferenz von *Intrusion* nimmt zwar auch für *Thickness* 1 und *Thickness*  $2 \in [0; 0, 5]$  kleine Werte an. Dort befinden sich aber keine robusten Optima, da die drei anderen Modelle dort groß werden.

Eine Auswahl des robusten Optimums ist für diesen Anwendungsfall auf der Grundlage der gerade besprochenen Abbildungen möglich. Bei mehr als zwei Eingabevariablen können die Metamodelle nicht mehr auf diese Art veranschaulicht werden. Deswegen wurde als letzter Schritt der vorgestellten Methodik die Visualisierung der robusten Optima über Sterndiagramme entwickelt. Diese wird nun für die B-Säule durchgeführt.

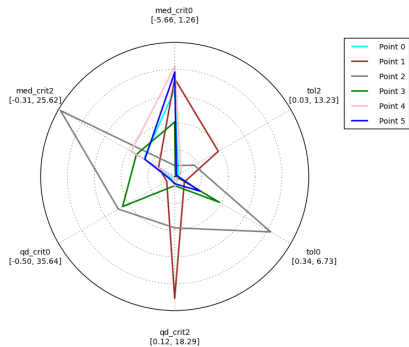


Abb. 9.5: Sterndiagramm für die sechs Clusterzentren

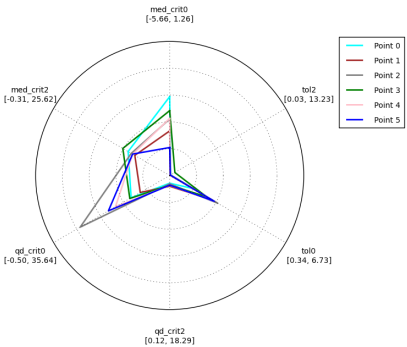


Abb. 9.6: Sterndiagramm mit Zentren für Cluster 3

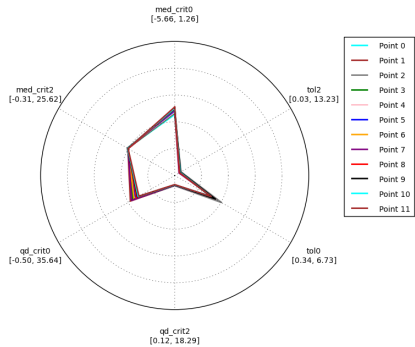


Abb. 9.7: Sterndiagramm mit Elementen von Clusterzentrum 3 aus Abb. 9.6

Als erstes werden die robusten Optima zu sechs Clustern zusammengefasst und ihre Zentren bestimmt. Dies sind die roten Punkte in den Abbildungen 9.3 und 9.4. Jedes Clusterzentrum repräsentiert einen oder mehrere robuste Bereiche.

Nur die Clusterzentren werden in einem Sterndiagramm dargestellt (Abbildung 9.5). Beispielsweise entspricht der Punkt 1 (roter Stern) dem kleinsten Wert von *Geschwindigkeit* (= crit2). Er steht für den robusten Bereich um das absolute Minimum bei (3; 1, 5). Die Quantildifferenz für diese Variable ist dort am größten.



Das Intervall an jeder Spitze des Sterns gibt den Wertebereich an, den die Paretopunkte für dieses Kriterium annehmen. Nur in diesem Bereich liegen robuste Optima. Sie geben dem Entscheider an, in welchem Rahmen sich die Optima bewegen bzw. welche Werte für die Ausgabevariablen erreichbar sind und welche nicht. Außerdem werden als zusätzliche Spitzen die Toleranzen der ursprünglichen Metamodelle angegeben. Ein Blick auf Abbildung 9.5 zeigt, dass sie sich lokal stark unterscheiden können.

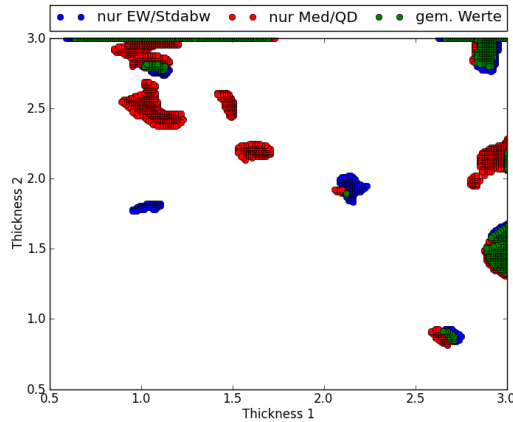


Abb. 9.8: Vergleich der Paretofronten für Erwartungswert und Standardabweichung sowie Median und Quantildifferenz von *Intrusion* und *Geschwindigkeit* in Abhängigkeit von *Thickness1* und *Thickness2*

Der Entscheider kann nun die einzelnen Lösungen analysieren und gegeneinander abwägen. Erst an dieser Stelle fließen seine Präferenzen in den Entscheidungsprozess ein. Er könnte sich z.B. für Punkt 3 (grüner Stern) entscheiden, wenn er beide Mediane gleichermaßen minimieren möchte und die dazugehörigen Quantildifferenzen nicht zu groß werden sollen. (Punkt 3 entspricht hier den Eingabevariablen  $Thickness1 \approx 1,3$  und  $Thickness2 \approx 2,0$ . Zu diesem Cluster gehören auch die robusten Bereiche weiter rechts in den Abbildungen.)

Nachdem der Entscheider eine Auswahl getroffen hat, wird die weitere Suche auf die robusten Optima dieses Clusters beschränkt. Da es mehr als sechs Punkte enthält, wird erneut geclustert. Die neu berechneten Clusterzentren werden dem Entscheider wieder in einem Sterndiagramm präsentiert (Abbildung 9.6). Hier kann er nun seine Abwägungen verfeinern. So weist Punkt 2 etwa eine große Quantildifferenz für *Intrusion* auf und könnte deswegen ausscheiden. Angenommen er wählt Punkt 3, gelangt er im nächsten Schritt zu den Elementen dieses Clusters (Abbildung 9.7). Aus diesen kann nun die endgültige Auswahl getroffen werden. Die einzelnen Alternativen unterscheiden sich jetzt nur noch marginal, denn sie sind Teil eines robusten Bereichs. Zuletzt werden die zugehörigen Werte für die Eingabevariablen ausgegeben. In diesem Fall erhält man den robusten Bereich bei  $Thickness1 = 2,08$  und  $Thickness2 = 1,89$ .

Auch für dieses Optimierungsproblem sollen die Ergebnisse mit denen der Optimierung auf Basis von Erwartungswert und Standardabweichung verglichen werden.

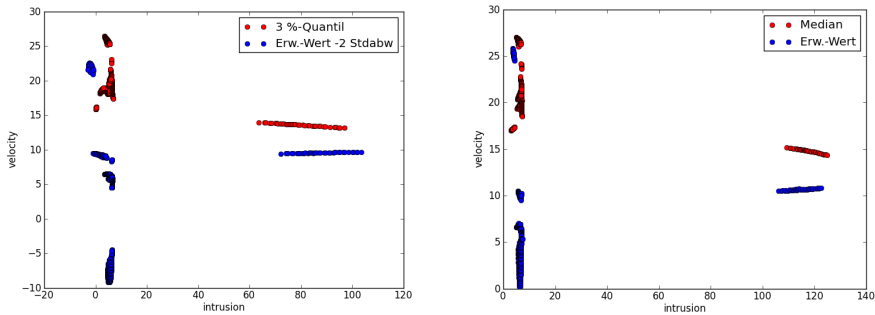


Abb. 9.9: Paretofront für Erwartungswert und Standardabweichung, Vergleich der unteren Grenze (links) und des Mittelwertes (rechts) mit den Quantilen

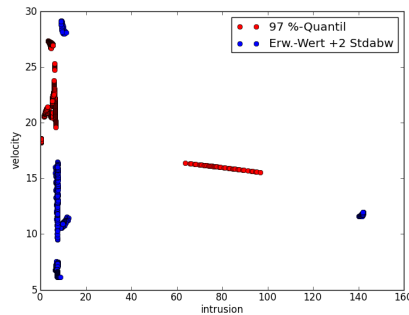


Abb. 9.10: Paretofront für Erwartungswert und Standardabweichung, Vergleich der oberen Grenze mit den Quantilen

Die Abbildung 9.8 zeigt die robusten Bereiche für die beiden Maße in Abhängigkeit von *Thickness1* und *Thickness2*. Sie besitzen kaum Überschneidungen. Die unterschiedlichen Maße finden unterschiedliche robuste Bereiche. Die für  $\mu$  und  $\sigma$  berechneten sind für diese Maße zwar optimal, entsprechen aber nicht der für die Industrie wichtigen Interpretation von Robustheit.

Die Paretofront für Erwartungswert und Standardabweichung soll nun genauer analysiert werden. Dazu wird sie an den Pareto-optimalen Punkten mit den Medianen und den Quantilen verglichen, die sich für diese Parameterkombinationen ergeben (siehe Abbildungen 9.9 und 9.10). Die untere 3%-Quantil-Grenze wird von  $\mu - 2\sigma$  tiefer eingeschätzt, als sie tatsächlich ist. Für *Geschwindigkeit* ergeben sich sogar negative Werte, die physikalisch nicht möglich sind. Die Werte sind sowohl für *Geschwindigkeit* als auch für *Intrusion* verschoben. Die Differenz ist nicht einheitlich, sondern hängt vom konkreten Punkt ab. Bei Erwartungswert und Median sowie  $\mu + 2\sigma$  und 97%-Quantil-Grenze sind die gleichen Phänomene zu beobachten.

Wenn man *Mass* als zusätzliches Optimierungskriterium hinzu nimmt, wird die Paretofront deutlich größer. In Abbildung 9.11 werden beide Maße für diese Situation verglichen. Auch hier unterscheiden sich die robusten Bereiche signifikant.

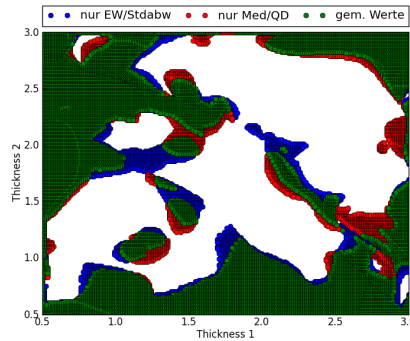


Abb. 9.11: Vergleich der Paretofronten von *Intrusion*, *Geschwindigkeit* und *Mass* für Erwartungswert und Standardabweichung sowie Median und Quantildifferenz in Abhängigkeit von *Thickness1* und *Thickness2*

## 9.2 Optimierung von pars10crits19

Das in Kapitel 8.2.5 eingeführte Beispiel pars10crits19 wird hier noch einmal aufgegriffen. Für die robuste Optimierung werden nun die Kriterien 3, 4 und 12 ausgewählt. Die Pearsonsche Korrelationsmatrix in Abbildung 9.12 zeigt, dass nahezu alle Eingabevariablen Einfluss auf diese Parameter haben. Eine genauere Analyse des Metamodells ergibt zudem, dass die Beziehung zu einigen Variablen nicht monoton ist.

Es wurden 101 Stützstellen bestimmt und ein Metamodell etabliert. Die Eingabevariablen seien normalverteilt mit Standardabweichung 0,01. An den Stützstellen werden die übertragenen Quantile berechnet. Die oberen Grenzen  $c$  für die Genauigkeit der Schätzung werden abhängig von der Größenordnung der Variablen gewählt. Sie liegen zwischen 0,001 und 1. Mit den gewonnenen Quantilwerten werden Metamodelle für den Median und die Quantildifferenz erstellt. Die Modelle werden validiert. Anschließend wird die Paretofront bestimmt. Durch nicht-dominiertes Sortieren von 20000 zufällig ausgewählten Punkten erhält man 1885 Pareto-optimale Punkte. Diese werden zuerst in 6 Clusterzentren zusammengefasst (Abbildung 9.13). Es wird angenommen, dass der Entscheider Cluster 2 auswählt. Für die Punkte dieses Cluster werden erneut 6 Clusterzentren bestimmt und visualisiert (Abbildung 9.14).

Angenommen, nun wird Zentrum 0 ausgewählt. Dieses Cluster enthält nun noch 10 Paretopunkte, die in Abbildung 9.16 gezeigt werden.

Dieses Beispiel zeigt, dass die Methodik auch für Optimierungsprobleme höherer Dimension anwendbar ist.

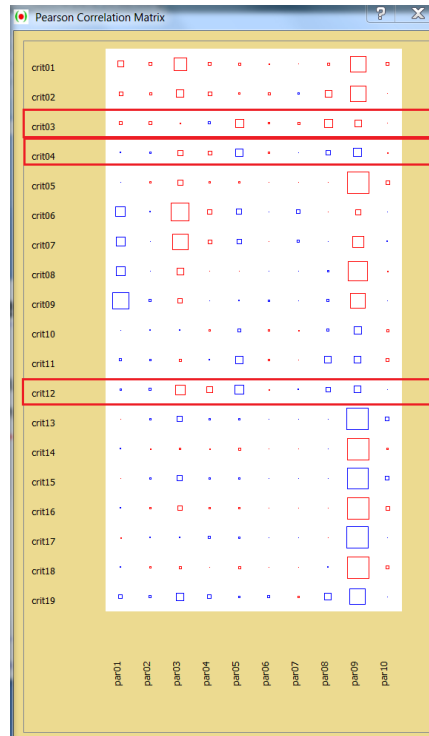


Abb. 9.12: Korrelationsmatrix zu pars10crits19, rot  $\hat{=} r > 0$ , blau  $\hat{=} r < 0$

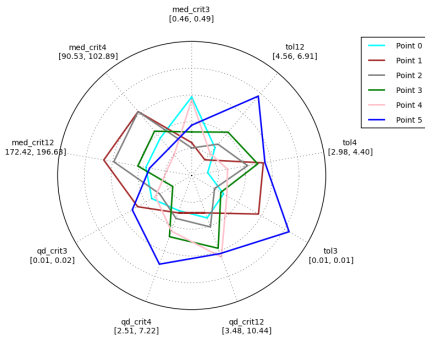


Abb. 9.13: Sterndiagramm für die sechs Clusterzentren

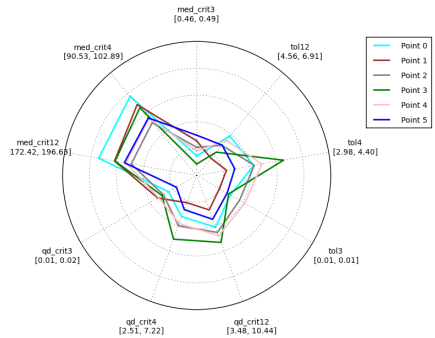


Abb. 9.14: Sterndiagramm mit Zentren für Cluster 2

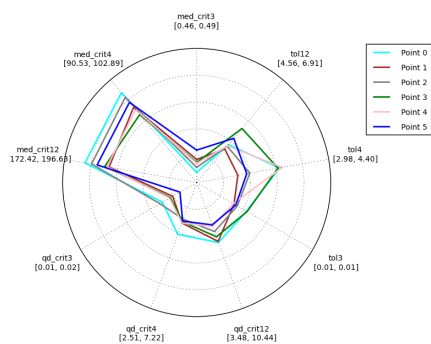


Abb. 9.15: Sterndiagramm mit Zentren zu Cluster 0 aus Abb. 9.14

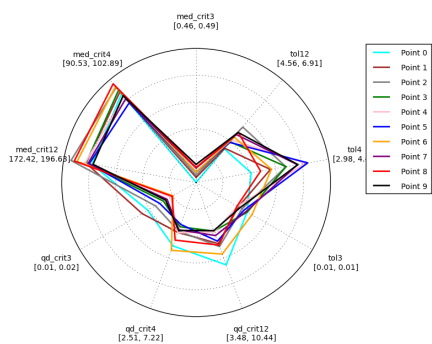


Abb. 9.16: Sterndiagramm mit Elementen von Clusterzentrum 0 aus Abb. 9.15



## Kapitel 10

---

# Toleranzmaß für die Quantilmetamodelle

---

In Kapitel 6.2 wurde ein Maß beschrieben, das den Einfluss der Toleranz des ursprünglichen Metamodells auf die Genauigkeit der Quantilmetamodelle misst. Analog zur Ermittlung der Metamodelltoleranz wurde ein Kreuzvalidierungsverfahren vorgestellt, das Toleranzgrenzen für das Median- und das Quantildifferenzmetamodell ausgehend vom ursprünglichen Metamodell ermittelt. Dieser Ansatz soll nun numerisch untersucht werden. Dazu wird ein Anwendungsbeispiel aus der Literatur benutzt, bei dem sich die exakten Werte analytisch bestimmen lassen.

Im Operations Research wird das folgende Lagerhaltungsmodell benutzt, um die optimale Bestellmenge zu bestimmen:

$$C = \underbrace{\frac{a \cdot K}{Q} + a \cdot c}_{\text{Beschaffungskosten}} + \underbrace{\frac{h \cdot Q}{2}}_{\text{Lagerhaltungskosten}}$$

mit  $C$  Gesamtkosten,  $a$  Jahresbedarf,  $K$  Fixkosten pro Bestellung,  $c$  Kosten pro produzierter Einheit,  $h$  kalkulatorischer Zins- und Lagerhaltungskostensatz und  $Q$  Bestellmenge. Dann gilt für die optimale Bestellmenge  $Q_{opt}$ , die  $C$  minimiert:

$$Q_{opt} = \sqrt{\frac{2aK}{h}}$$

Dieses Beispiel wurde schon in von Dellino mit den folgenden Werten verwendet [DKM12]:  $a = 8000$  Einheiten,  $K = 12000$  Euro,  $c = 10$  Euro pro Stück,  $h = 30\%$ . Dann folgt:

$$Q_{opt} = 25298$$

Nun wird angenommen, dass der Bedarf unsicher ist, aber dass gilt

$$a \sim \mathcal{N}(\mu_a; \sigma_a)$$

mit  $\mu_a = 8000$  und  $\sigma_a = 800$ .

Da  $C$  in  $a$  linear ist, folgt:

$$\begin{aligned} y_{0.5}(C) = E(C) &= \frac{\mu_a \cdot K}{Q} + \mu_a \cdot c + \frac{h \cdot Q}{2} \\ \sigma(C) &= \frac{\sigma_a \cdot K}{Q} + \sigma_a \cdot c \\ QD(C) &= \frac{QD(C) \cdot K}{Q} + QD(C) \cdot c \end{aligned}$$

Aus dem deterministischen Optimierungsproblem wird nun ein stochastisches und multikriterielles: Dellino bestimmt die Paretofront für die Minimierung von Erwartungswert und Standardabweichung. An dieser Stelle wird zusätzlich auch die Minimierung von Median und Quantildifferenz mit einbezogen. Wegen der Linearität von  $C$  in  $a$  kann auch sie exakt bestimmt werden (vgl. Abbildung 10.1).

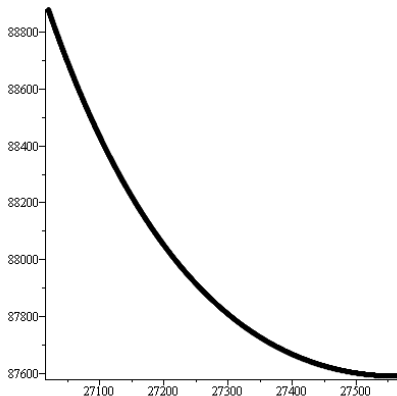


Abb. 10.1: Paretofront der Kostenfunktion für Median und Quantildifferenz

In der Praxis ist die Funktionsgleichung für  $C$  unbekannt. Es wird angenommen, dass sie für bestimmte Kombinationen von  $a$  und  $Q$  bestimmt werden kann. Für die numerische Untersuchung soll gelten  $a \in [5600; 10400] = [\mu_a - 3\sigma_a; \mu_a + 3\sigma_a]$  und  $Q \in [15000; 45000]$ . Abweichend von den bisherigen Beispielen werden nicht die übertragenen Verteilungen für  $C$  bestimmt, sondern die bedingten übertragenen Verteilungen von  $C$  unter  $Q$ . Da  $C$  in  $a$  linear ist, lassen sich die bedingten Quantile exakt berechnen. Es werden 100 Parameterkombination für  $(a, Q)$ , die nach der Haltonfolge verteilt sind, bestimmt und hieraus das Ursprungsmetamodell  $\tilde{f}$  erstellt. Zur Berechnung der bedingten Quantile von  $C$  unter  $Q$  wird an 25 gleichverteilten Punkten von  $Q$  jeweils eine gemäß  $a$  verteilte Stichprobe von 10000 in Richtung von  $a$  gezogen. Aus diesen Stichproben werden dann die bedingten Quantile mit dem Harrell-Davis-Schätzer bestimmt. Man erhält an 25 Stellen von  $Q$  die bedingten Quantile von  $C$  in Bezug auf  $a$  unter  $Q$ .

Im Unterschied zur in Algorithmus 3 beschriebenen Kreuzvalidierungsprozedur werden in Schritt 9 bis Schritt 16 nicht die übertragenen Quantile an den Stützstellen,



sondern die bedingten übertragenen Quantile von  $a$  unter  $Q$  an 25 gleichverteilten Stellen von  $Q$  berechnet.

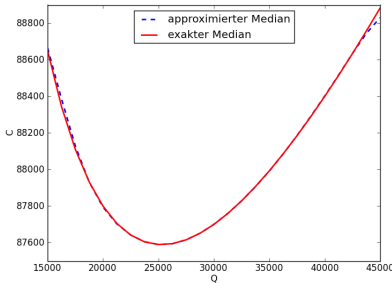


Abb. 10.2: Vergleich exakter und approximierter Median

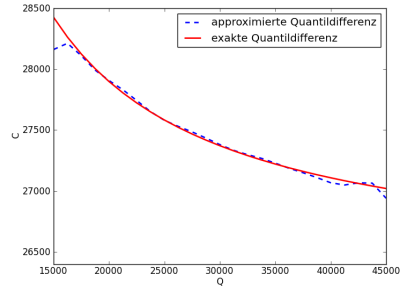


Abb. 10.3: Vergleich exakte und approximierter Quantildifferenz

Abbildung 10.2 zeigt den exakten und den geschätzten Median in Abhängigkeit von  $Q$ . Die Approximation ist sehr gut. Nur am rechten Rand des Definitionsbereichs kommt es zu einer größeren Abweichung.

Die Schätzung der Quantildifferenz (siehe Abbildung 10.3) gelingt weniger gut: An den beiden Rändern weicht sie stark vom exakten Wert ab.

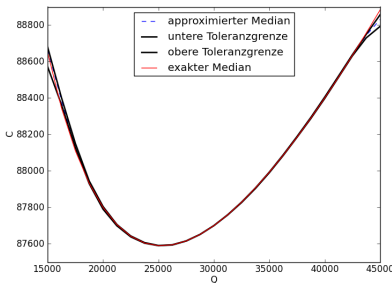


Abb. 10.4: exakter und approximierter Median mit Toleranzgrenzen

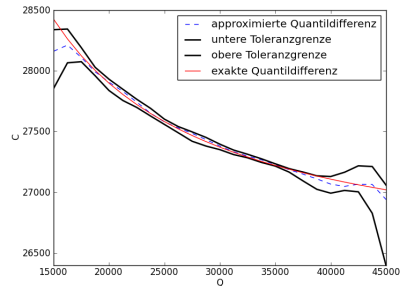


Abb. 10.5: exakte und approximierter Quantildifferenz mit Toleranzgrenzen

Die Toleranzgrenzen für den Median sind in Abbildung 10.4 dargestellt. Sie verlaufen eng um den geschätzten Median, nur an den Rändern werden sie größer. Der exakte Wert wird nicht in allen Bereichen eingeschlossen: Zwischen 16000 und 18000 sowie ab 43000 liegt er außerhalb der Toleranzgrenzen.

Für die Quantildifferenz schließen die Toleranzgrenzen den exakten Wert ein bis auf einen kleinen Bereich zwischen 15000 und 16000 (vgl. Abbildung 10.5).

Abbildung 10.6 zeigt, dass die exakte und die approximierter Paretofront innerhalb der geschätzten Toleranzgrenzen liegen. Der Bereich für rechts vom Minimum des Medians in  $Q = 27566$  bildet die Paretofront. Dort schließen die Toleranzgrenzen

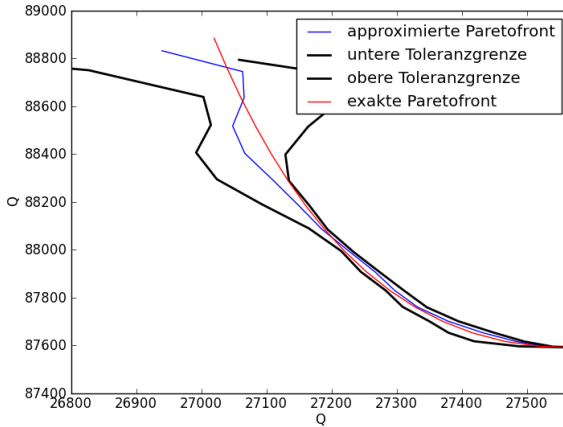


Abb. 10.6: exakte und approximierte Paretofront mit Toleranzgrenzen

für die Quantildifferenz den wahren Wert ein und bilden einen solch weiten Bereich, dass auch der falsche Toleranzbereich für den Median nicht ins Gewicht fällt.

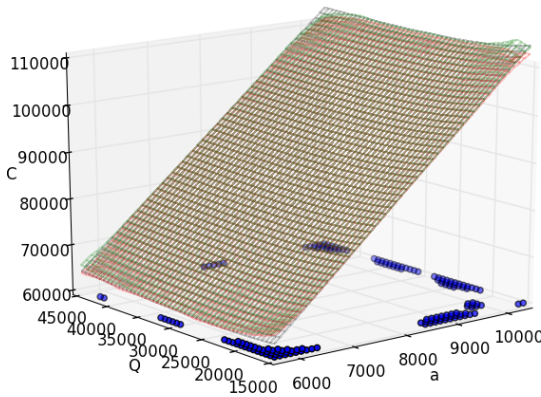


Abb. 10.7: exakte Kostenfunktion (schwarz), Metamodell + Toleranz (grün), Metamodell - Toleranz (rot), Bereiche, in denen die exakte Funktion außerhalb des Bereichs Metamodell +/- Toleranz liegt (blaue Punkte)

Insgesamt veranschaulicht dieses Beispiel, dass mit der Kreuzvalidierungsprozedur Toleranzbereiche für die Quantilmetamodelle bestimmt werden können. Es zeigt aber auch, dass nicht garantiert werden kann, dass die wahren Werte tatsächlich von den Bereichen eingeschlossen werden. Die Toleranzwerte sollten vielmehr qualitativ interpretiert werden: Große Toleranzen deuten darauf hin, dass die Schätzung an dieser Stelle ungenau sein könnte, kleine zeigen eine genaue Schätzung an.

Diese Interpretation wird vom hier präsentierten Beispiel bestätigt: Die weiten Be-

reiche jeweils an den Rändern von Median und Quantildifferenz zeigen die Ungenauigkeit richtig an. Bei der Quantildifferenz gibt es auch um  $Q = 40.000$  einen größeren Schätzfehler, der über weitere Toleranzgrenzen angezeigt wird.

Diese Interpretation folgt der der Kreuzvalidierungsprozedur für Metamodelle. Auch sie kann nicht gewährleisten, dass die wahren Werte vom Toleranzbereich eingeschlossen werden:

Abbildung 10.7 zeigt die Funktionsgleichung, das Metamodell + Modelltoleranz und das Metamodell – Modelltoleranz für  $C(a, Q)$ . Da sie sich kaum unterscheiden, sind in der  $a$ - $Q$ -Ebene die Bereiche, in denen die wahren Werte nicht im angegebenen Toleranzbereich liegen, blau markiert. Für dieses Beispiel liefert Kreuzvalidierung jeweils an den Rändern falsche Bereiche.



## Kapitel 11

---

# Zusammenfassung und Ausblick

---

In dieser Dissertation wurde eine innovative Methodik zur robusten Optimierung mit Metamodellen unter Verwendung von Quantilmaßen entwickelt. Sie umfasst Elemente der Numerik, der Optimierung und der Statistik, die miteinander zu einer neuen Vorgehensweise kombiniert und um neue Inhalte ergänzt wurden. Verschiedene theoretische Aspekte wurden untersucht und die Relevanz der Methodik für die Praxis an mehreren Beispielen erläutert.

Die wichtigsten Ergebnisse werden hier noch einmal zusammengefasst und ein Ausblick auf die zukünftige Forschung gegeben:

Die globale Methodik benötigt nur so viele Funktionsauswertungen, wie nötig sind, um ein hinreichend validiertes Metamodell zu erstellen. Die genaue Anzahl hängt von der Nichtlinearität der Übertragungsfunktion ab. Damit entspricht sie den Anforderungen der Industrie, möglichst wenige Simulationen oder Experimente durchführen zu müssen. Ist das Ursprungsmetamodell genau genug, läuft die Methodik bis zur Bestimmung der robusten Optima automatisch ab. Die Ausführung kann an mehreren Stellen parallelisiert werden.

Es wurde gezeigt, dass das bisherige Standardmaß für Robustheit Schwächen aufweist, die besonders bei Industrieanwendungen zum Tragen kommen. Stattdessen werden in der Methodik Quantilmaße verwendet: Sie geben die untere und obere Grenze des Schwankungsbereichs der Ausgabevariablen genauer als das bisherige Standardmaß an. Dadurch verbessert sich die Aussagekraft der in der Optimierungsphase bestimmten robusten Bereiche. Die Quantilmaße passen sich dem gewünschten Grad an Robustheit an.

Zur effizienten Berechnung übertragener Quantile wurde ein neues Verfahren entwickelt. Es besteht aus der Kombination der Haltonfolge als Versuchsplan für die Verteilung der Stichprobe um die Eingabevariablen mit dem Harrell-Davis-Schätzer für die Schätzung des zugehörigen Quantils der Ausgabevariable. In einer numerischen Studie wurde bewiesen, dass diese Kombination der Standardprozedur und weiteren untersuchten Alternativen deutlich überlegen ist. Sie kommt mit kleineren Stichprobengrößen aus und beschleunigt die Berechnung wesentlich. Ein zusätzlicher Vorteil ist, dass mit diesem Verfahren auch die gesamte Verteilungsfunktion

schnell berechnet werden kann. Es wurde in einen neuen iterativen Algorithmus integriert, der punktweise Quantile der Ausgabevariablen bis zu einer vorgegebenen Genauigkeit schätzt.

Zur Abschätzung der globalen Genauigkeit der Methodik wurde ein Maß für die Toleranz der Quantilmetamodelle entwickelt. Es handelt sich um ein erweitertes Kreuzvalidierungsverfahren, das auf den einzigen „sicheren“ Werten, nämlich den Stützstellen des ursprünglichen Metamodells, basiert. Es identifiziert die unsicheren Bereiche der Quantilmetamodelle.

Die entwickelte Methodik kann beliebige Metamodelle verwenden, die intensiv ausgewertet werden. Eine Beschleunigung der Auswertung ist für die Klasse der radialen Basisfunktionenmodelle durch die Verwendung von Multipolverfahren möglich. Dieser Weg wurde beschrieben und über eine genaue Aufwandsabschätzung gezeigt, wann er Vorteile bringt.

Neu eingeführt wurde eine Vorgehensweise zur Unterstützung des Anwenders bei der konkreten Auswahl des robusten Optimums. Dabei werden die robusten Optima geclustert und in speziell angepassten Stern diagrammen visualisiert.

Ein Thema zukünftiger Forschung wird die Kombination der entwickelten Methodik mit dem Verfeinern von Metamodellen sein, um die Anzahl der nötigen Simulationen weiter zu reduzieren. Die Grundidee sieht wie folgt aus: Für ein grobes Metamodell werden die nicht robusten Bereiche bestimmt und diese von der weiteren Bearbeitung ausgeschlossen. Verfeinert durch zusätzliche Stützstellen werden nur die übriggebliebenen Bereiche. Dieses Verfahren wird sequentiell durchgeführt. Voraussetzung ist eine zuverlässige Erfassung der robusten Bereiche für grobe Metamodelle.

Im Hinblick auf Robustheitsmaße sollte folgender Aspekt weiter beleuchtet werden: In die robuste Optimierung mit Median und Quantildifferenz, aber auch mit den Standardmaßen, geht nicht ein, ob die Streuung der Ausgabevariablen sich stärker nach oben oder nach unten auswirkt, sondern lediglich die Breite eines mittigen Bereichs der Ausgabeverteilung. Über die Analyse der Grenzen kann der Anwender dies zwar nachvollziehen. Eine Integration dieser Eigenschaft in das Robustheitsmaß wäre aber wünschenswert, denn wenn eine Ausgabeverteilung zum Minimierungsziel hin verzogen ist, wird dies in Anwendungen positiv gewertet. Denkbar ist beispielsweise ein Gewichtungsfaktor für die Quantildifferenz, der eine Verzerrung in die eine Richtung belohnt und in die andere bestraft.

Um die entwickelte Methodik in der Automobilindustrie zu etablieren, ist es nötig, diese möglichst weit zu automatisieren. Die Durchführung der Simulationen gemäß einem vorgegebenen Versuchsplan und die Zusammenführung der Daten für die Stützstellen lässt sich z.B. über Skripte steuern. Dieser Weg wird aktuell in einem entsprechenden Projekt für die Bauteildimensionierung mit Finite-Elemente-Simulationen beschritten.

**Teil III**

**Anhang**





## Kapitel 12

---

# Grafiken zu Kapitel 8

---

Alle Grafiken dieses Kapitels zeigen auf der  $y$ -Achse den Fehler  $\widehat{MISE}$ . Die Ordinate ist logarithmisch skaliert, bis auf Kapitel 12.2.1. Dort ist sie kartesisch.

Die absolute Größe von  $\widehat{MISE}$  hängt stark vom Wertebereich ab, den die Übertragungsfunktion im Bereich um den untersuchten Punkt abdeckt. Insbesondere die sehr hohen Werte für das Zweistab-Fachwerk sind damit zu erklären.

## 12.1 Grafiken zur Bestimmung der Bandweite

### 12.1.1 Stichprobengröße $n = 100$

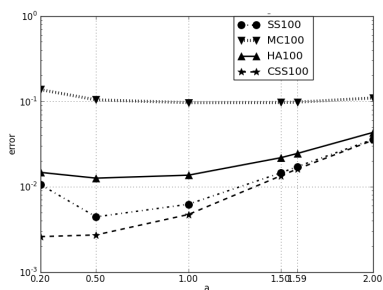


Abb. 12.1: Testfall 1,  $n = 100$

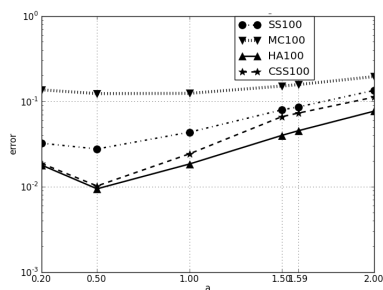


Abb. 12.2: Testfall 2,  $n = 100$

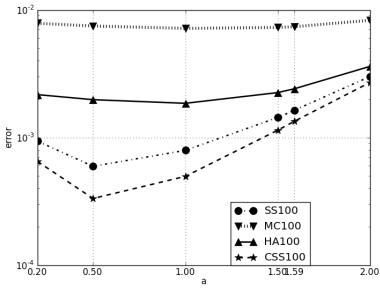


Abb. 12.3: Testfall 3,  $n = 100$

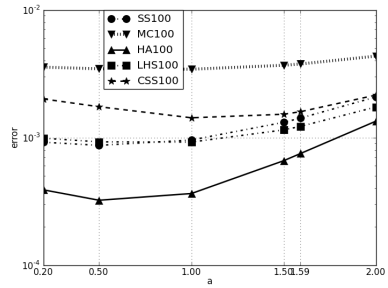


Abb. 12.4: Testfall 4,  $n = 100$

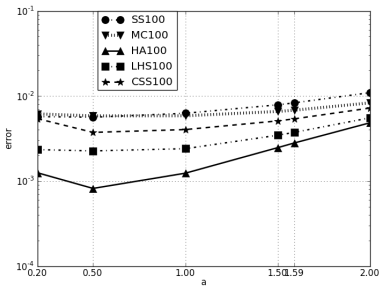


Abb. 12.5: Testfall 5,  $n = 100$

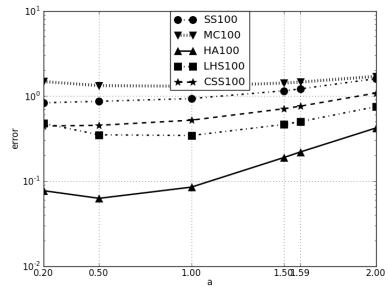


Abb. 12.6: Testfall 6,  $n = 100$

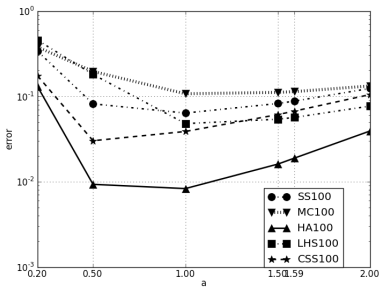


Abb. 12.7: Testfall 7,  $n = 100$

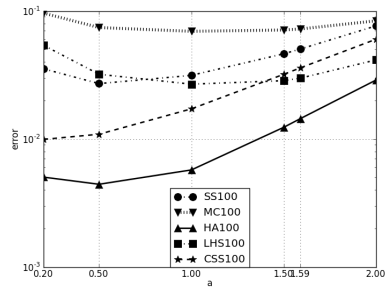


Abb. 12.8: Testfall 8,  $n = 100$

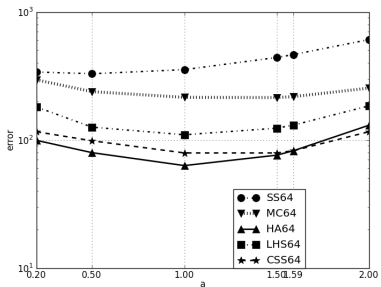


Abb. 12.9: Testfall Zweistab-Fachwerk,  $n = 64$

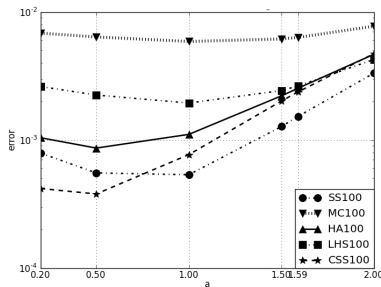


Abb. 12.10: Testfall B-Säule,  $n = 100$

12.1.2 Stichprobengröße  $n = 1000$  und  $n = 10000$

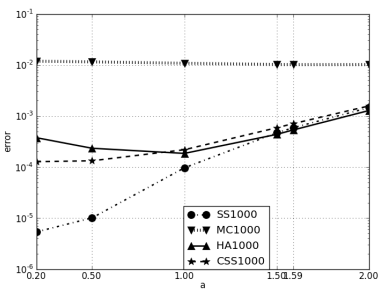


Abb. 12.11: Testfall 1,  $n = 1000$

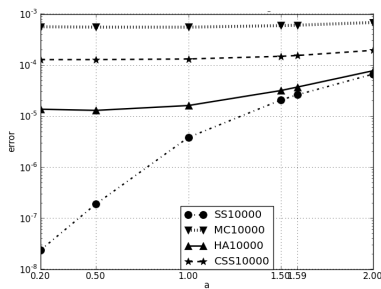


Abb. 12.12: Testfall 1,  $n = 10000$

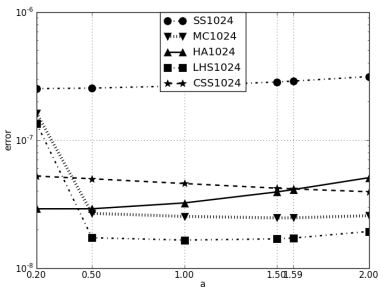


Abb. 12.13: Testfall pars10crits19,  $n = 1024$

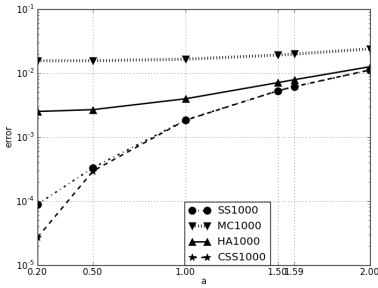


Abb. 12.14: Testfall 2,  $n = 1000$

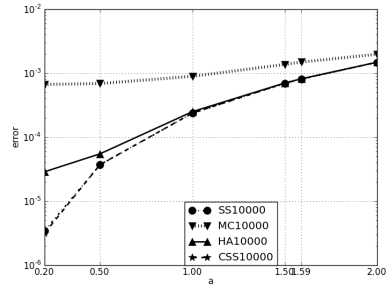


Abb. 12.15: Testfall 2,  $n = 10000$

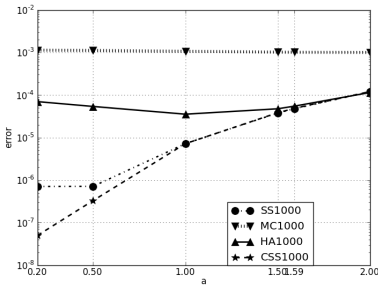


Abb. 12.16: Testfall 3,  $n = 1000$

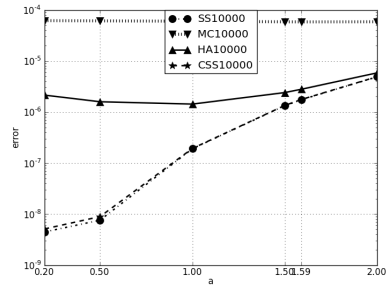


Abb. 12.17: Testfall 3,  $n = 10000$

## 12.2 Vergleich der Quantilschätzer

Die Skalierung der  $y$ -Achse ist für  $n = 100$  kartesisch, für  $n = 1000$  und  $n = 10000$  logarithmisch.

### 12.2.1 Stichprobengröße $n = 100$

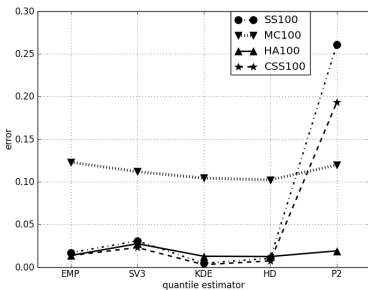


Abb. 12.18: Testfall 1,  $n = 100$

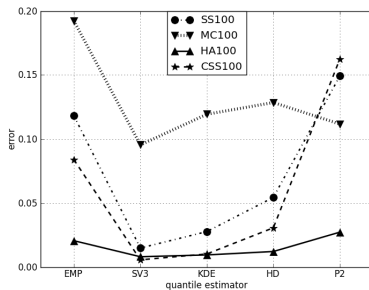


Abb. 12.19: Testfall 2,  $n = 100$

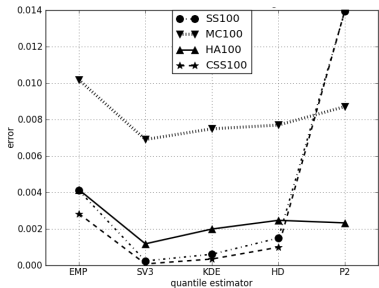


Abb. 12.20: Testfall 3,  $n = 100$

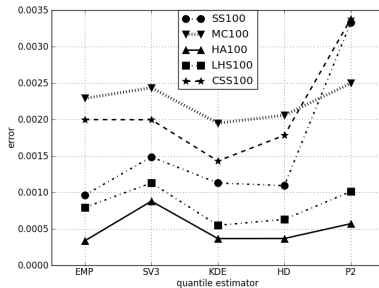


Abb. 12.21: Testfall 4,  $n = 100$

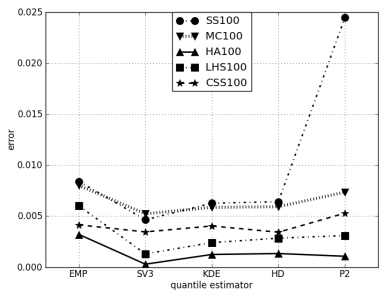


Abb. 12.22: Testfall 5,  $n = 100$

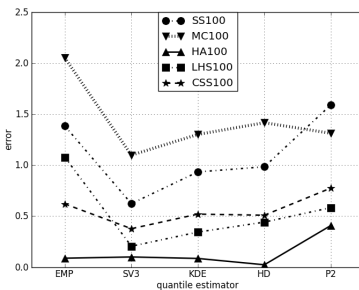


Abb. 12.23: Testfall 6,  $n = 100$

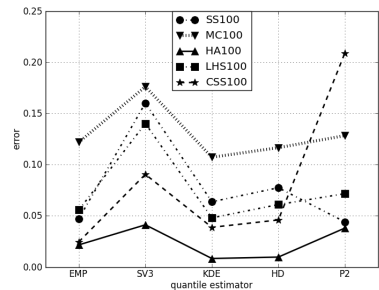


Abb. 12.24: Testfall 7,  $n = 100$

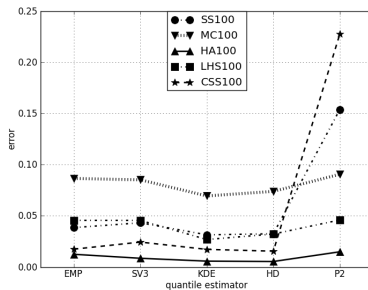


Abb. 12.25: Testfall 8,  $n = 100$

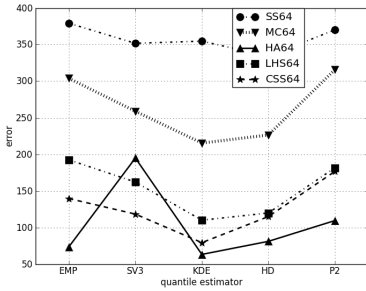


Abb. 12.26: Testfall zsf,  $n = 64$

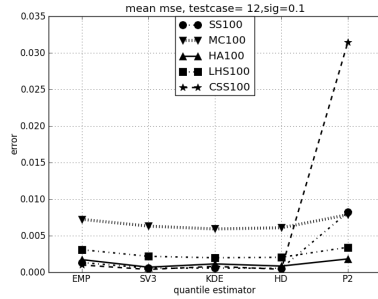


Abb. 12.27: Testfall B-Säule,  $n = 100$

### 12.2.2 Stichprobengröße $n = 1000$

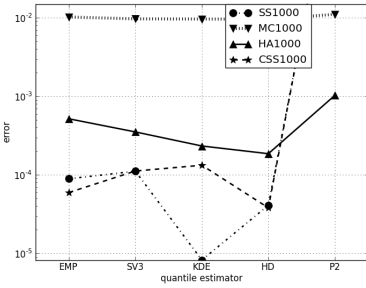


Abb. 12.28: Testfall 1,  $n = 1000$

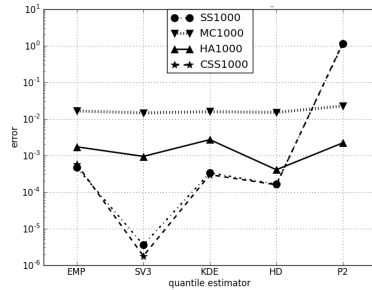


Abb. 12.29: Testfall 2,  $n = 1000$

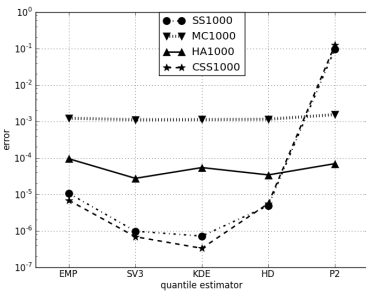


Abb. 12.30: Testfall 3,  $n = 1000$

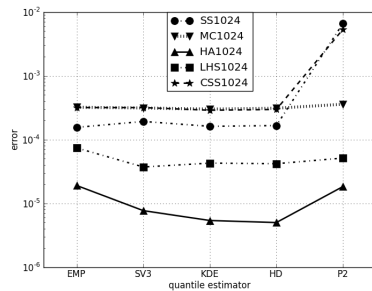


Abb. 12.31: Testfall 4,  $n = 1024$

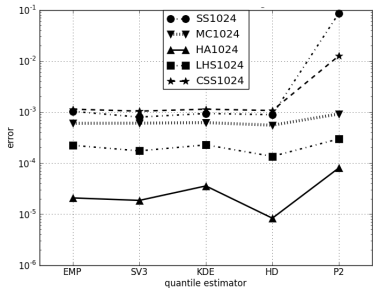


Abb. 12.32: Testfall 5,  $n = 1024$

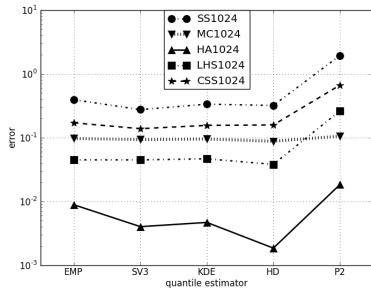


Abb. 12.33: Testfall 6,  $n = 1024$

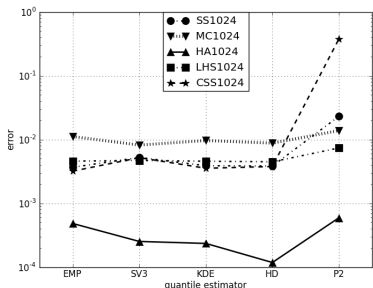


Abb. 12.34: Testfall 7,  $n = 1024$

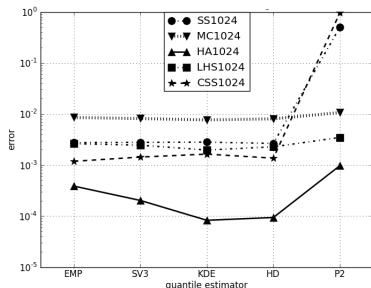


Abb. 12.35: Testfall 8,  $n = 1024$

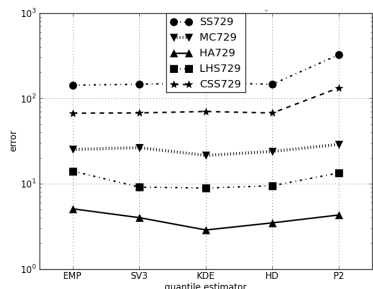


Abb. 12.36: Testfall zsf,  $n = 729$

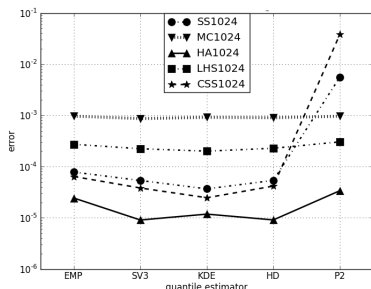


Abb. 12.37: Testfall B-Säule,  $n = 1024$

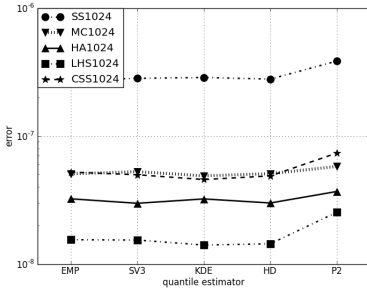


Abb. 12.38: Testfall pars10crits19,  $n = 1024$

### 12.2.3 Stichprobengröße $n = 10000$

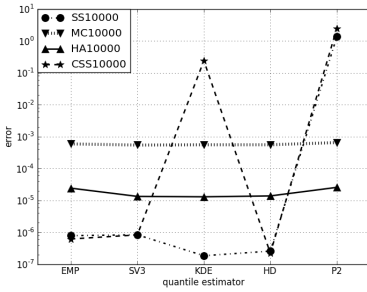


Abb. 12.39: Testfall 1,  $n = 10000$

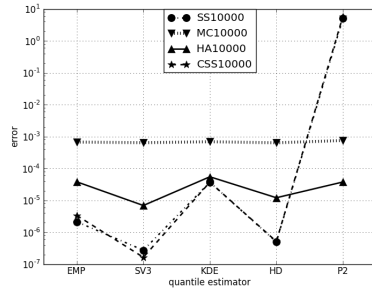


Abb. 12.40: Testfall 2,  $n = 10000$

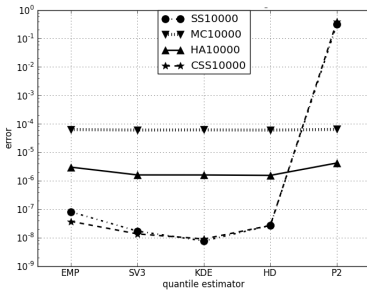


Abb. 12.41: Testfall 3,  $n = 10000$

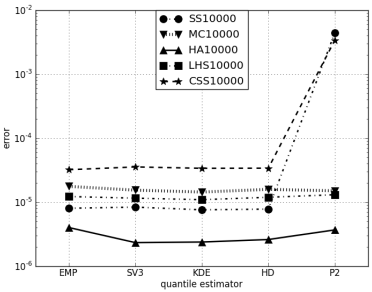


Abb. 12.42: Testfall 4,  $n = 10000$



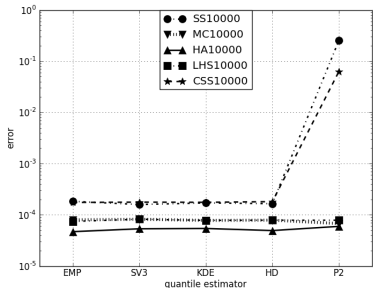


Abb. 12.43: Testfall 5,  $n = 10000$

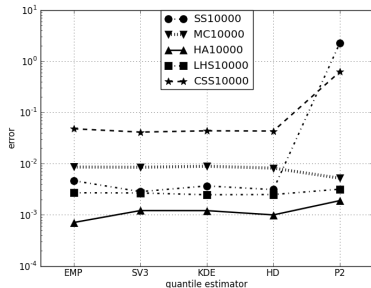


Abb. 12.44: Testfall 6,  $n = 10000$

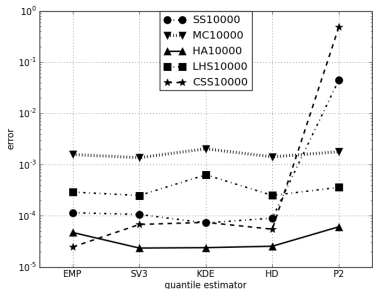


Abb. 12.45: Testfall 7,  $n = 10000$

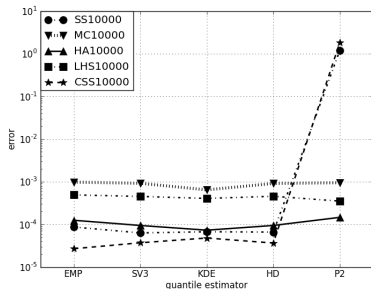


Abb. 12.46: Testfall 8,  $n = 10000$

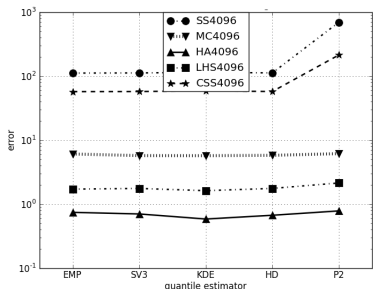


Abb. 12.47: Testfall zsf,  $n = 4096$

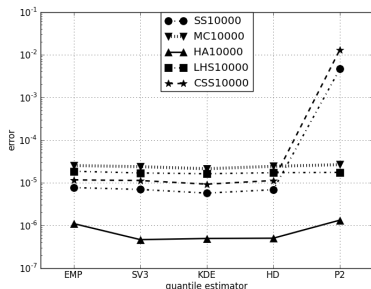


Abb. 12.48: Testfall B-Säule,  $n = 10000$

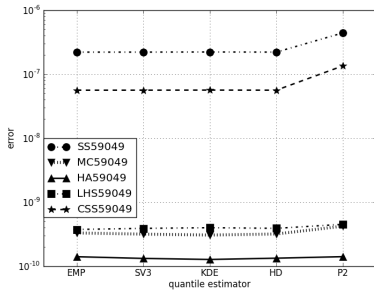


Abb. 12.49: Testfall pars10crits19,  $n = 59049$

---

# Abkürzungsverzeichnis

---

## Metamodelle

$d$	Anzahl der Eingabevariablen
$l$	Anzahl der Zielvariablen
$f$	Übertragungsfunktion
$\tilde{f}$	Metamodell von $f$
$\Phi$	radiale Basisfunktion
$x_i, i = 1, \dots, n$	Stützstellen des Metamodells

## Wahrscheinlichkeitsrechnung und Statistik

$n$	Stichprobengröße
$X, Y$	Zufallsvariable
$F$	Verteilungsfunktion
$p$	Dichtefunktion
$Q$	Quantilfunktion
$x_q$	$q$ -Quantil
$\hat{Y}_{q,n}$	Standard-Schätzer für das $q$ -Quantil
$X_0$	Zufallsvariable für $x_0 = (x_{01}, \dots, x_{0d})^T$
$Y_0$	Zufallsvariable für $f(X_0)$
$E(X)$ bzw. $\mu_X$	Erwartungswert von $X$
$\text{Var}(X)$ bzw. $\sigma_X^2$	Varianz von $X$
$\sigma_X$	Standardabweichung von $X$

## Fehlermaße

MSE	mittlerer quadratischer Fehler
RMSE	Wurzel des mittleren quadratischen Fehlers
MISE	mittlerer quadratischer Integrationsfehler
$\widehat{\text{MISE}}$	Annäherung an MISE vgl. 8.3.2

## Versuchspläne

MC	Monte Carlo
LHS	Latin Hypercube

SS	Stratified Sampling
CSS	Centered Stratified Sampling
HAL	Haltonfolge
Quantilschätzer	
EMP	empirischer Standardschätzer
KDE	Kern dichteschätzer
HD	Harrell-Davis-Schätzer
SV3	Sfakianikis-Verginis-Schätzer
$P^2$	Schätzung mit $P^2$ -Algorithmus
Quantilmaße	
WCM	Worst-Case-Maß
EQM	einseitiges Quantil-Maß
QD <sub>Y</sub>	Quantildifferenz aus $y_q - y_{1-q}$
PSM	Wahrscheinlichkeitsschwellenmaß
ENM	Erwartungswert-Nutzenfunktion-Maß
EWM	Erwartungswert-Maß
GZ	Gewichtete-Zielfunktion-Maß
GZE	GZ bzgl. Erwartungswert und Standardabweichung
GZM	GZ bzgl. Median und Quantildifferenz
S / N	Taguchis Signal-Rausch-Verhältnis
H	Entropie-Maß
Multipolverfahren	
$n$	Anzahl der Stützstellen
$m$	Anzahl der Auswertungspunkte
$x_i, i = 1, \dots, n$	Partikel, entsprechen den Stützstellen des Metamodells
$y_j, j = 1, \dots, m$	Auswertungspunkte, entsprechen der Gesamtheit aller Stichproben um die Stützstellen
$bt$	Baumtiefe
$y_C$	Clusterzentrum
$r$	Radius eines Clusters
$R$	Abstand Partikel zu Cluster

---

# Abbildungsverzeichnis

---

1.1	grober Ablauf der Methodik . . . . .	2
2.1	B-Säule mit Intrusor . . . . .	8
2.2	Metamodell für <i>Intrusion</i> und <i>Mass</i> mit Simulationsergebnissen . . . . .	8
2.3	beste Kompromisse für <i>Intrusion</i> und <i>Mass</i> . . . . .	9
2.4	Metamodell für <i>Intrusion</i> mit besten Kompromissen . . . . .	9
2.5	Metamodell für <i>Mass</i> mit besten Kompromissen . . . . .	9
2.6	übertragene Verteilung für <i>Intrusion</i> im optimalen Punkt . . . . .	10
2.7	deterministische Paretofront mit Schwankungsbereich . . . . .	11
2.8	Metamodell für den Median von <i>Intrusion</i> mit besten Kompromissen . . . . .	12
2.9	Metamodell für den Median von <i>Mass</i> mit besten Kompromissen . . . . .	12
2.10	Metamodell für Quantildifferenz von <i>Int.</i> mit besten Kompromissen . . . . .	12
2.11	Metamodell für Quantildifferenz von <i>Mass</i> mit besten Kompromissen . . . . .	12
2.12	übertragene Verteilung für <i>Intrusion</i> im robusten Optimum . . . . .	13
2.13	Menge der besten Kompromisse bei det. und robuster Optimierung . . . . .	13
2.14	beste Kompromisse bei robuster Optimierung mit Streuungsbereich . . . . .	13
2.15	Metamodell für Erwartungswert von <i>Int.</i> mit besten Kompromissen . . . . .	14
2.16	Metamodell für Erwartungswert von <i>Mass</i> mit besten Kompromissen . . . . .	14
2.17	Metamodell für Standardabw. von <i>Int.</i> mit besten Kompromissen . . . . .	14
2.18	Metamodell für Standardabw. von <i>Mass</i> mit besten Kompromissen . . . . .	14
2.19	Vergleich der Menge der besten Kompromisse . . . . .	15
3.1	übertragene Verteilung für $f$ unstetig . . . . .	23
3.2	übertragene Verteilung für $f$ mit lokalem Extremum . . . . .	23
3.3	übertragene Verteilung für $f$ mit zwei lokalen Extrema . . . . .	24
3.4	übertragene Verteilung für $f$ mit globalem Extremum . . . . .	24
3.5	übertragene Verteilung für $f$ globalem Maximum und Minimum . . . . .	24
3.6	Monte Carlo, $n = 400$ , Gleichverteilung . . . . .	25
3.7	Latin Hypercube Sampling, $n = 400$ , Gleichverteilung . . . . .	25
3.8	Stratified Sampling, $n = 400$ , Gleichverteilung . . . . .	26
3.9	Centered Stratified Sampling, $n = 400$ , Gleichverteilung . . . . .	26

3.10	Haltonfolge mit $p_1 = 2$ und $p_2 = 3$ , $n = 400$ , Gleichverteilung . . . . .	28
3.11	Monte Carlo, $n = 400$ , Normalverteilung . . . . .	29
3.12	Monte Carlo, $n = 10000$ , Normalverteilung . . . . .	29
3.13	Latin Hypercube Sampling, $n = 400$ , Normalverteilung . . . . .	29
3.14	Latin Hypercube Sampling, $n = 10000$ , Normalverteilung . . . . .	29
3.15	Stratified Sampling, $n = 400$ , Normalverteilung . . . . .	30
3.16	Stratified Sampling, $n = 10000$ , Normalverteilung . . . . .	30
3.17	Centered Stratified Sampling, $n = 400$ , Normalverteilung . . . . .	30
3.18	Centered Stratified Sampling, $n = 10000$ , Normalverteilung . . . . .	30
3.19	Haltonfolge, $n = 400$ , Normalverteilung . . . . .	31
3.20	Haltonfolge, $n = 10000$ , Normalverteilung . . . . .	31
3.21	Veranschaulichung des empirischen Schätzers EMP . . . . .	32
3.22	Beispiel einer Kerndichteschätzung . . . . .	34
3.23	Veranschaulichung von Radius und Partikel-Cluster-Abstand . . . . .	49
3.24	Beispiel einer Paretofront für zwei Zielvariablen . . . . .	53
3.25	Beispiel für ein Sterndiagramm . . . . .	53
4.1	Ausgangssituation für robuste Optimierung . . . . .	58
4.2	Grundidee des robusten Designs . . . . .	58
4.3	Deterministische Minima und robustes Minimum . . . . .	60
4.4	Worst-Case-Maß . . . . .	63
4.5	einseitiges Quantilmaß . . . . .	63
4.6	Median . . . . .	65
4.7	Quantildifferenz . . . . .	65
4.8	Schwellenmaß . . . . .	67
4.9	Erwartungswertmaß . . . . .	67
4.10	Standardabweichung . . . . .	68
4.11	gewichtete Zielfunktion, $0,5\mu_Y + 0,5\sigma_Y$ . . . . .	69
4.12	gewichtete Zielfunktion $0,5y_{0,5} + 0,5(y_{0,95} - y_{0,05})$ . . . . .	69
4.13	Taguchis S/N . . . . .	71
4.14	Erwartungswert und Pareto-optimale Punkte . . . . .	72
4.15	Median und Pareto-optimale Punkte . . . . .	72
4.16	Paretofront für Erwartungswert und Standardabweichung . . . . .	73
4.17	Paretofront für Median und Quantildifferenz . . . . .	73
4.18	Vergleich von $\mu + / - 2\sigma$ mit Median und Quantildifferenz . . . . .	74
4.19	Vergleich von $\mu + / - 2\sigma$ mit Median und Quantildifferenz . . . . .	74
4.20	Beispiele für übertragene Verteilungen von $f$ . . . . .	75
5.1	Varianz des Standardschätzers für die Standardnormalverteilung . . . . .	84
5.2	nötige Stichprobengröße bei der Standardnormalverteilung . . . . .	84
6.1	Schritt 1: Metamodell für $f$ . . . . .	90
6.2	Schritt 2: Robustheitsberechnung . . . . .	91
6.3	Schritt 3: Bestimmung der robusten Optima . . . . .	92
6.4	Schritt 4: Visualisierung der Paretofront und Entscheidung . . . . .	92
7.1	ungünstigste Verteilung der $x_i$ . . . . .	103

7.2	günstigste Verteilung der $x_i$ . . . . .	106
7.3	direkte Summation und Boxen, für die MAC erfüllt ist . . . . .	106
7.4	Aufwandsvergleich für $d = 3$ und $d = 4$ . . . . .	111
8.1	MSE bei Schätzung der Normalverteilung, MC . . . . .	116
8.2	MSE bei Schätzung der Normalverteilung, HAL . . . . .	116
8.3	MSE bei Schätzung der Normalverteilung, SS . . . . .	116
8.4	MSE bei Schätzung der Normalverteilung, CSS . . . . .	116
8.5	MSE bei Schätzung der Normalverteilung, MC, ohne P2 . . . . .	117
8.6	MSE bei Schätzung der Normalverteilung, HAL, ohne P2 . . . . .	117
8.7	MSE bei Schätzung der Normalverteilung, SS, ohne P2 . . . . .	117
8.8	MSE bei Schätzung der Normalverteilung, CSS, ohne P2 . . . . .	117
8.9	Dichtefunktion von Testfall 1 . . . . .	118
8.10	Dichtefunktion von Testfall 2 . . . . .	118
8.11	Dichtefunktion von Testfall 3 . . . . .	118
8.12	Funktionsgraph von $f_4$ . . . . .	119
8.13	Funktionsgraph von $f_5$ . . . . .	119
8.14	Funktionsgraph von $f_6$ . . . . .	120
8.15	Funktionsgraph von $f_7$ . . . . .	120
8.16	Dichtefunktion von Testfall 4 . . . . .	120
8.17	Dichtefunktion von Testfall 5 . . . . .	120
8.18	Dichtefunktion von Testfall 6 . . . . .	120
8.19	Dichtefunktion von Testfall 7 . . . . .	121
8.20	Dichtefunktion von Testfall 8 . . . . .	121
8.21	Zweistab-Fachwerk . . . . .	121
8.22	Dichtefunktion für Zweistab-Fachwerk . . . . .	122
8.23	Metamodell für Geschwindigkeit mit Stützstellen . . . . .	122
8.24	Dichtefunktion von <i>Geschwindigkeit</i> für (2,4; 2,9) . . . . .	122
8.25	Dichtefunktion von pars10crits19 für die gewählte Parameterkombination . . . . .	123
8.26	Vergleich der optimalen Bandweite für Standardnormalverteilung und Chi-Quadratverteilung . . . . .	124
8.27	Übersicht aller Testfälle für HD, $n = 100$ . . . . .	131
8.28	Übersicht aller Testfälle für HD, $n = 1000$ . . . . .	131
8.29	Übersicht aller Testfälle für HD, $n = 10000$ . . . . .	132
8.30	Vergleich aller Testfälle für die Haltonfolge, $n = 100$ . . . . .	134
8.31	Vergleich aller Testfälle für die Haltonfolge, $n = 1000$ . . . . .	134
8.32	Vergleich aller Testfälle für die Haltonfolge, $n = 10000$ . . . . .	135
8.33	Vergleich der Stichprobengröße für Testfall 1 und Haltonfolge . . . . .	135
8.34	prozentualer Vergleich bezüglich MC und EMP, $n = 100$ . . . . .	137
8.35	prozentualer Vergleich bezüglich MC und EMP, $n = 1000$ . . . . .	137
8.36	prozentualer Vergleich bezüglich MC und EMP, $n = 10000$ . . . . .	138
8.37	iterativer Algorithmus, Testfall 1 . . . . .	138
8.38	iterativer Algorithmus, Testfall 2 . . . . .	138
8.39	iterativer Algorithmus, Testfall 3 . . . . .	139
8.40	iterativer Algorithmus, Testfall 4 . . . . .	139
8.41	iterativer Algorithmus, Testfall 5 . . . . .	139

8.42	iterativer Algorithmus, Testfall 6 . . . . .	139
8.43	iterativer Algorithmus, Testfall 7 . . . . .	139
8.44	iterativer Algorithmus, Testfall 8 . . . . .	139
9.1	Beispiele für Dichtefunktionen von <i>Intrusion</i> . . . . .	142
9.2	Beispiele für Dichtefunktionen von <i>Geschwindigkeit</i> . . . . .	142
9.3	Metamodelle für <i>Geschwindigkeit</i> : Median und Quantildifferenz . . . . .	143
9.4	Metamodelle für <i>Intrusion</i> : Median und Quantildifferenz . . . . .	143
9.5	Sterndiagramm für die sechs Clusterzentren . . . . .	144
9.6	Sterndiagramm mit Zentren für Cluster 3 . . . . .	144
9.7	Sterndiagramm mit Elementen von Clusterzentrum 3 aus Abb. 9.6 . . . . .	144
9.8	Vergleich der Paretofronten für <i>Intrusion</i> und <i>Geschwindigkeit</i> . . . . .	145
9.9	Paretofront für $\mu$ und $\sigma$ , Vergleich der Grenzen mit den Quantilen . . . . .	146
9.10	Paretofront für $\mu$ und $\sigma$ , Vergleich der Grenzen mit den Quantilen . . . . .	146
9.11	Vergleich der Paretofronten für <i>Int.</i> , <i>Geschw.</i> und <i>Mass</i> . . . . .	147
9.12	Korrelationsmatrix zu pars10crits19 . . . . .	148
9.13	Sterndiagramm für die sechs Clusterzentren . . . . .	148
9.14	Sterndiagramm mit Zentren für Cluster 2 . . . . .	148
9.15	Sterndiagramm mit Zentren zu Cluster 0 aus Abb. 9.14 . . . . .	149
9.16	Sterndiagramm mit Elementen von Clusterzentrum 0 aus Abb. 9.15 . . . . .	149
10.1	Paretofront der Kostenfunktion für Median und Quantildifferenz . . . . .	152
10.2	Vergleich exakter und approximierter Median . . . . .	153
10.3	Vergleich exakte und approximierter Quantildifferenz . . . . .	153
10.4	exakter und approximierter Median mit Toleranzgrenzen . . . . .	153
10.5	exakte und approximierter Quantildifferenz mit Toleranzgrenzen . . . . .	153
10.6	exakte und approximierter Paretofront mit Toleranzgrenzen . . . . .	154
10.7	Vergleich exakte Funktion und Toleranzbereiche des Metamodells . . . . .	154
12.1	Vergleich der Bandweiten für Testfall 1, $n = 100$ . . . . .	161
12.2	Vergleich der Bandweiten für Testfall 2, $n = 100$ . . . . .	161
12.3	Vergleich der Bandweiten für Testfall 3, $n = 100$ . . . . .	162
12.4	Vergleich der Bandweiten für Testfall 4, $n = 100$ . . . . .	162
12.5	Vergleich der Bandweiten für Testfall 5, $n = 100$ . . . . .	162
12.6	Vergleich der Bandweiten für Testfall 6, $n = 100$ . . . . .	162
12.7	Vergleich der Bandweiten für Testfall 7, $n = 100$ . . . . .	162
12.8	Vergleich der Bandweiten für Testfall 8, $n = 100$ . . . . .	162
12.9	Vergleich der Bandweiten für Testfall Zweistab-Fachwerk, $n = 64$ . . . . .	163
12.10	Vergleich der Bandweiten für Testfall B-Säule, $n = 100$ . . . . .	163
12.11	Vergleich der Bandweiten für Testfall 1, $n = 1000$ . . . . .	163
12.12	Vergleich der Bandweiten für Testfall 1, $n = 10000$ . . . . .	163
12.13	Vergleich der Bandweiten für pars10crits19, $n = 1024$ . . . . .	163
12.14	Vergleich der Bandweiten für Testfall 2, $n = 1000$ . . . . .	164
12.15	Vergleich der Bandweiten für Testfall 2, $n = 10000$ . . . . .	164
12.16	Vergleich der Bandweiten für Testfall 3, $n = 1000$ . . . . .	164
12.17	Vergleich der Bandweiten für Testfall 3, $n = 10000$ . . . . .	164
12.18	Vergleich der Quantilschätzer für Testfall 1, $n = 100$ . . . . .	164



12.19	Vergleich der Quantilschätzer für Testfall 2, $n = 100$ . . . . .	164
12.20	Vergleich der Quantilschätzer für Testfall 3, $n = 100$ . . . . .	165
12.21	Vergleich der Quantilschätzer für Testfall 4, $n = 100$ . . . . .	165
12.22	Vergleich der Quantilschätzer für Testfall 5, $n = 100$ . . . . .	165
12.23	Vergleich der Quantilschätzer für Testfall 6, $n = 100$ . . . . .	165
12.24	Vergleich der Quantilschätzer für Testfall 7, $n = 100$ . . . . .	165
12.25	Vergleich der Quantilschätzer für Testfall 8, $n = 100$ . . . . .	165
12.26	Vergleich der Quantilschätzer für Testfall zsf, $n = 64$ . . . . .	166
12.27	Vergleich der Quantilschätzer für Testfall B-Säule, $n = 100$ . . . . .	166
12.28	Vergleich der Quantilschätzer für Testfall 1, $n = 1000$ . . . . .	166
12.29	Vergleich der Quantilschätzer für Testfall 2, $n = 1000$ . . . . .	166
12.30	Vergleich der Quantilschätzer für Testfall 3, $n = 1000$ . . . . .	166
12.31	Vergleich der Quantilschätzer für Testfall 4, $n = 1024$ . . . . .	166
12.32	Vergleich der Quantilschätzer für Testfall 5, $n = 1024$ . . . . .	167
12.33	Vergleich der Quantilschätzer für Testfall 6, $n = 1024$ . . . . .	167
12.34	Vergleich der Quantilschätzer für Testfall 7, $n = 1024$ . . . . .	167
12.35	Vergleich der Quantilschätzer für Testfall 8, $n = 1024$ . . . . .	167
12.36	Vergleich der Quantilschätzer für Testfall zsf, $n = 729$ . . . . .	167
12.37	Vergleich der Quantilschätzer für Testfall B-Säule, $n = 1024$ . . . . .	167
12.38	Vergleich der Quantilschätzer für Testfall pars10crits19, $n = 1024$ . . . . .	168
12.39	Vergleich der Quantilschätzer für Testfall 1, $n = 10000$ . . . . .	168
12.40	Vergleich der Quantilschätzer für Testfall 2, $n = 10000$ . . . . .	168
12.41	Vergleich der Quantilschätzer für Testfall 3, $n = 10000$ . . . . .	168
12.42	Vergleich der Quantilschätzer für Testfall 4, $n = 10000$ . . . . .	168
12.43	Vergleich der Quantilschätzer für Testfall 5, $n = 10000$ . . . . .	169
12.44	Vergleich der Quantilschätzer für Testfall 6, $n = 10000$ . . . . .	169
12.45	Vergleich der Quantilschätzer für Testfall 7, $n = 10000$ . . . . .	169
12.46	Vergleich der Quantilschätzer für Testfall 8, $n = 10000$ . . . . .	169
12.47	Vergleich der Quantilschätzer für Testfall zsf, $n = 4096$ . . . . .	169
12.48	Vergleich der Quantilschätzer für Testfall B-Säule, $n = 10000$ . . . . .	169
12.49	Vergleich der Quantilschätzer für Testfall pars10crits19, $n = 59049$ . . . . .	170

---

# Tabellenverzeichnis

---

3.1	Vergleich der Quantilschätzer . . . . .	43
4.1	Übersicht der Robustheitsmaße und ihres Minimierungszieles . . . . .	77
8.1	Übersicht der durchgeführten Berechnungen zur Wahl der Bandweite . .	126
8.2	Übersicht der optimalen Bandweiten für eine Dimension . . . . .	126
8.3	Übersicht der optimalen Bandweiten bei unterschiedlichen Stichproben- größen . . . . .	127
8.4	Vergleich der Versuchspläne für eine Dimension . . . . .	127
8.5	Übersicht der optimalen Bandweiten für mehrere Dimensionen . . . . .	128
8.6	Vergleich der Versuchspläne für $a = 1$ . . . . .	128
8.7	Übersicht der durchgeführten Berechnungen zum Vergleich der Versuchs- pläne und Schätzer . . . . .	129
8.8	Vergleich der Versuchspläne (außer P2) für eine Dimension . . . . .	130
8.9	Vergleich der Versuchspläne (außer P2) für mehrere Dimensionen . . . . .	130
8.10	Vergleich der Schätzer bzgl. HAL für eine Dimension . . . . .	133
8.11	Vergleich der Schätzer bzgl. HAL für mehrere Dimensionen . . . . .	133

---

# Literaturverzeichnis

---

- [AC09] ALLEN, M. S. ; CAMBEROS, J. A.: Comparison of uncertainty propagation/response surface techniques for two aeroelastic systems. In: *50th AIAA Structures, Structural Dynamics, and Materials Conference, Palm Springs, California, 2009* [cited at p. 55]
- [Ang11] ANGÜN, E.: A risk-averse approach to simulation optimization with multiple responses. In: *Simulation Modelling Practice and Theory* 19 (2011), Nr. 3, S. 911–923 [cited at p. 65]
- [APG94] ALURU, S. ; PRABHU, G. M. ; GUSTAFSON, J.: Truly distribution-independent algorithms for the N-body problem. In: *Proceedings of the 1994 ACM/IEEE conference on Supercomputing* IEEE Computer Society Press, 1994, S. 420–428 [cited at p. 49]
- [AW98] AVRAMIDIS, A. N. ; WILSON, J. R.: Correlation-induction techniques for estimating quantiles in simulation experiments. In: *Operations Research* 46 (1998), Nr. 4, S. 574–591 [cited at p. 33]
- [Azz81] AZZALINI, A.: A note on the estimation of a distribution function and quantiles by a kernel method. In: *Biometrika* 68 (1981), Nr. 1, S. 326–328 [cited at p. 35, 36]
- [Bau12] BAUDOUI, V.: *Optimisation robuste multiobjectifs par modèles de substitution*, ISAE-Institut Supérieur de l’Aéronautique et de l’Espace, Diss., 2012 [cited at p. 63]
- [BBC11] BERTSIMAS, D. ; BROWN, D. B. ; CARAMANIS, C.: Theory and applications of robust optimization. In: *SIAM review* 53 (2011), Nr. 3, S. 464–501 [cited at p. 64, 77]
- [BC10] BATUR, D. ; CHOUBINEH, F.: A quantile-based approach to system selection. In: *European Journal of Operational Research* 202 (2010), Nr. 3, S. 764–772 [cited at p. 65]
- [Ben80] BENTLEY, J. L.: Multidimensional divide-and-conquer. In: *Communications of the ACM* 23 (1980), Nr. 4, S. 214–229 [cited at p. 54]

- [BHP98] BOWMAN, A. ; HALL, P. ; PRVAN, T.: Bandwidth selection for the smoothing of distribution functions. In: *Biometrika* 85 (1998), Nr. 4, S. 799–808 [cited at p. 38]
- [BK13] BOATENG, H. A. ; KRASNY, R.: Comparison of Treecodes for Computing Electrostatic Potentials in Charged Particle Systems with Disjoint Targets and Sources. In: *Journal of Computational Chemistry* 34 (2013), Nr. 25, S. 2159–2167 [cited at p. 48, 49, 51, 111]
- [Bro07] BRODIN, E.: *Extreme value statistics and quantile estimation with applications in finance and insurance*. Chalmers University of Technology, 2007 [cited at p. 38]
- [BS01] BARTON, R. R. ; SCHRUBEN, L. W.: Resampling methods for input modeling. In: *Simulation Conference, 2001. Proceedings of the Winter Bd. 1* IEEE, 2001, S. 372–378 [cited at p. 44]
- [BS07] BEYER, H.-G. ; SENDHOFF, B.: Robust optimization - A comprehensive survey. In: *Computer Methods in Applied Mechanics and Engineering* 196 (2007), Nr. 33-34, S. 3190–3218 [cited at p. 2, 61, 63, 64, 66, 118]
- [BT87] BYRNE, D. M. ; TAGUCHI, S.: The taguchi approach to parameter design. In: *40th Annual Quality Congress Transactions* (1987), S. 19–26 [cited at p. 71]
- [BTGN09] BEN-TAL, A. ; GHAOUI, L. E. ; NEMIROVSKI, A.: *Robust optimization*. Princeton, NJ : Princeton Univ. Press, 2009 (Princeton series in applied mathematics) [cited at p. 64]
- [BTN02] BEN-TAL, A. ; NEMIROVSKI, A.: Robust optimization—methodology and applications. In: *Mathematical Programming* 92 (2002), Nr. 3, S. 453–480 [cited at p. 64]
- [Buh03] BUHMANN, M. D.: *Radial Basis Functions: Theory and Implementations*. Cambridge University Press, 2003 [cited at p. 45, 46]
- [Caf98] CAFLISCH, R. E.: Monte Carlo and quasi-Monte Carlo methods. In: *Acta Numerica* 7 (1998), S. 1–49 [cited at p. 28]
- [CATM96] CHEN, W. ; ALLEN, J. K. ; TSUI, K.-L. ; MISTREE, F.: A Procedure For Robust Design: Minimizing Variations Caused By Noise Factors And Control Factors. In: *ASME Journal of Mechanical Design* 118 (1996), Nr. 4, S. 478–485 [cited at p. 68, 72, 78, 89]
- [CBN02] CHERRIE, J. ; BEATSON, R. ; NEWSAM, G.: Fast Evaluation of Radial Basis Functions: Methods for Generalized Multiquadrics in  $\mathbb{R}^n$ . In: *SIAM Journal on Scientific Computing* 23 (2002), Nr. 5, S. 1549–1571 [cited at p. 46]
- [CGI08] CANNAMELA, C. ; GARNIER, J. ; IOOSS, B.: Controlled stratification for quantile estimation. In: *The Annals of Applied Statistics* 2 (2008) [cited at p. 82]

- [Che00] CHEN, S. Q.: *Comparing probabilistic and fuzzy set approaches for design in the presence of uncertainty*, Virginia Polytechnic Institute and State University, Diss., 2000 [cited at p. 61]
- [CK08] CHEN, E. J. ; KELTON, W. D.: Estimating steady-state distributions via simulation-generated histograms. In: *Computers and Operations Research* 35 (2008), Nr. 4, S. 1003 – 1016 [cited at p. 40]
- [CNNT11] CLEES, T. ; NIKITIN, I. ; NIKITINA, L. ; THOLE, C.-A.: Nonlinear Metamodeling and Robust Optimization in Automotive Design. In: *SIMULTECH*, SciTePress, 2011, S. 483–491 [cited at p. 74]
- [CS06] CHENG, M.-Y. ; SUN, S.: Bandwidth selection for kernel quantile estimation. In: *Journal of the Chinese Statistical Association* 44 (2006), Nr. 44, S. 271–295 [cited at p. 35, 37]
- [Das00] DAS, I.: Robust optimization for constrained, nonlinear programming problems. In: *Engineering Optimization* 32 (2000), Nr. 5, S. 585–618 [cited at p. 72, 77]
- [DC00] DU, X. ; CHEN, W.: Towards a better understanding of modeling feasibility robustness in engineering design. In: *Journal of Mechanical Design* 122 (2000), Nr. 4, S. 385–394 [cited at p. 78]
- [DD97] DAS, I. ; DENNIS, J. E.: A closer look at drawbacks of minimizing weighted sums of objectives for Pareto set generation in multicriteria optimization problems. In: *Structural and Multidisciplinary Optimization* 14 (1997), S. 63–69 [cited at p. 54]
- [DD12] DENG, Q. ; DRISCOLL, T. A.: A Fast Treecode for Multiquadric Interpolation with Varying Shape Parameters. In: *SIAM Journal on Scientific Computing* 34 (2012), Nr. 2, S. A1126–A1140 [cited at p. 48, 49, 104, 111]
- [Deb05] DEB, K.: Multi-Objective Optimization. In: BURKE, E. K. (Hrsg.) ; KENDALL, G. (Hrsg.): *Search Methodologies*. Springer US, 2005 [cited at p. 52]
- [DK04] DOLTSINIS, I. ; KANG, Z.: Robust design of structures using optimization methods. In: *Computer Methods in Applied Mechanics and Engineering* 193 (2004), Nr. 23, S. 2221–2237 [cited at p. 66, 72, 78]
- [DK10] DALBEY, K.R. ; KARAYSTINOS, G.N.: Fast generation of spacefilling Latin hypercube sample designs. In: *Proc. of the 13th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, 2010 [cited at p. 26]
- [DKM12] DELLINO, G. ; KLEIJNEN, J. P. C. ; MELONI, C.: Robust optimization in simulation: Taguchi and Krige combined. In: *INFORMS Journal on Computing* 24 (2012), Nr. 3, S. 471–484 [cited at p. 90, 94, 151]
- [DLP94] DIELMAN, T. ; LOWRY, C. ; PFAFFENBERG, R.: A comparison of quantile estimators. In: *Communications in Statistics - Simulation and Computation* 23 (1994), S. 355–371 [cited at p. 42]

- [DN03] DAVID, H. A. ; NAGARAJA, H. N.: *Order Statistics*. Wiley, New York, 2003 [cited at p. 33, 38, 82]
- [DSC04] DU, X. ; SUDJANTO, A. ; CHEN, W.: An Integrated Framework for Optimization Under Uncertainty Using Inverse Reliability Strategy. In: *Journal of Mechanical Design* 126 (2004), Nr. 4, S. 562–570 [cited at p. 65, 73]
- [Ehr99] EHRGOTT, M.: *Multicriteria Optimization*. Springer, Berlin, 1999 [cited at p. 52]
- [Emb97] EMBRECHTS, P.: *Modelling extremal events: for insurance and finance*. Bd. 33. Springer, 1997 [cited at p. 41]
- [EPM11] EICKHOFF, M. ; PAWLIKOWSKI, K. ; MCNICKLE, D.: Parallel Sequential Estimation of Quantiles During Steady State Simulation. (2011) [cited at p. 66]
- [ES00] EVANS, M. ; SWARTZ, T.: *Approximating integrals via Monte Carlo and deterministic methods*. Oxford University Press, 2000 [cited at p. 25]
- [Fas07] FASSHAUER, G. E.: *Meshfree approximation methods with MATLAB*. Bd. 6. World Scientific Publishing Company, 2007 [cited at p. 48]
- [FHT01] FRIEDMAN, J. ; HASTIE, T. ; TIBSHIRANI, R.: *The elements of statistical learning*. Bd. 1. Springer Series in Statistics, 2001 [cited at p. 45, 46]
- [FK09] FORRESTER, A. ; KEANE, A.: Recent advances in surrogate-based optimization. In: *Progress in Aerospace Sciences* 45 (2009), Nr. 1, S. 50–79 [cited at p. 47]
- [FSK08] FORRESTER, A. ; SOBESTER, A. ; KEANE, A.: *Engineering design via surrogate modelling: a practical guide*. Wiley, 2008 [cited at p. 45]
- [GC03] GIBBONS, J. D. ; CHAKRABORTI, S.: *Nonparametric statistical inference*. Marcel Dekker Inc., New York, 2003 [cited at p. 44]
- [GLK02] GREEN, L. L. ; LIN, H.-Z. ; KHALESSI, M.: Probabilistic Methods For Uncertainty Propagation Applied To Aircraft Design. In: *20th AIAA Applied Aerodynamics Conference, Fluid Dynamics and Collocated Conferences, 2002* [cited at p. 68]
- [GR87] GREENGARD, L. ; ROKHLIN, V.: A fast algorithm for particle simulations. In: *Journal of computational physics* 73 (1987), Nr. 2, S. 325–348 [cited at p. 48]
- [Haf89] HAFNER, R.: *Wahrscheinlichkeitsrechnung und Statistik*. Springer-Verlag, Wien, New York, 1989 [cited at p. 22]
- [Hal64] HALTON, J. H.: Algorithm 247: Radical-inverse quasi-random point sequence. In: *Commun. ACM* 7 (1964), Nr. 12, S. 701–702. – ISSN 0001–0782 [cited at p. 27]

- [Ham60] HAMMERSLEY, J. M.: Monte Carlo methods for solving multivariable problems. In: *Ann. New York Acad. Sci.* 86 (1960), S. 844–874 [cited at p. 28]
- [Han04] HANSEN, B. E.: *Bandwidth Selection for Nonparametric Distribution Estimation*. 2004 [cited at p. 38]
- [Har71] HARDY, R. L.: Multiquadric equations of topography and other irregular surfaces. In: *Journal of geophysical research* 76 (1971), Nr. 8, S. 1905–1915 [cited at p. 45]
- [Hay94] HAYKIN, S.: *Neural networks: a comprehensive foundation*. Prentice Hall PTR, 1994 [cited at p. 45]
- [HD82] HARRELL, F. E. ; DAVIS, C. E.: A new distribution-free quantile estimator. In: *Biometrika* 69 (1982), S. 635–640 [cited at p. 38]
- [HD03] HELTON, J. C. ; DAVIS, F. J.: Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. In: *Reliability Engineering and System Safety* 81 (2003), Nr. 1, S. 23–69 [cited at p. 29]
- [HD06] HUANG, B. ; DU, X.: A robust design method using variable transformation and Gauss—Hermite integration. In: *International Journal for Numerical Methods in Engineering* 66 (2006), S. 1841–1858 [cited at p. 68]
- [HD07] HUANG, B. ; DU, X.: Analytical robustness assessment for robust design. In: *Structural and Multidisciplinary Optimization* 34 (2007), S. 123–137 [cited at p. 30, 73, 121]
- [Hel97] HELTON, J.C.: Uncertainty and sensitivity analysis in the presence of stochastic and subjective uncertainty. In: *Journal of Statistical Computation and Simulation* 57 (1997), Nr. 1-4, S. 3–76 [cited at p. 61]
- [HF96] HYNDMAN, R.J. ; FAN, Y.: Sample quantiles in statistical packages. In: *American Statistician* 50 (1996), Nr. 4, S. 361–365 [cited at p. 34]
- [HL84] HEIDELBERGER, P. ; LEWIS, P.: Quantile estimation in dependent sequences. In: *Operations Research* 32 (1984), Nr. 1, S. 185–209 [cited at p. 40]
- [HMPY79] HWANG, C.L. ; MASUD, A.S.M. ; PAIDY, S.R. ; YOON, K.: *Multiple Objective Decision Making - Methods and Applications: A State-of-the-Art Survey*. Springer-Verlag, Berlin/Heidelberg/New York, 1979 [cited at p. 52]
- [HNC10] HORNUNG, N. ; NIKITINA, L. ; CLEES, T.: Multi-objective optimization using surrogate functions. In: *Proc. 2nd International Conference on Engineering Optimization*, 2010 [cited at p. 54]
- [HP03] HESS, S. ; POLAK, J.: An alternative method to the scrambled Halton sequence for removing correlation between standard Halton sequences in high dimensions. In: *ERSA conference papers* European Regional Science Association, 2003 [cited at p. 28]

- [HRRS11] HAMPEL, F. R. ; RONCHETTI, E. M. ; ROUSSEEUW, P. J. ; STAHEL, W. A.: *Robust statistics: the approach based on influence functions*. Bd. 114. Wiley. com, 2011 [cited at p. 79]
- [Hur08] HURTADO, J.: structural robustness and its relationship to reliability. In: TSOMPANAKIS, Yiannis (Hrsg.): *Structural design optimization considering uncertainties*. Taylor and Francis, 2008 [cited at p. 71]
- [Hut01] HUTSON, A. D.: Rational Spline Estimators of the Quantile Function. In: *Communications in Statistics - Simulation and Computation* 30 (2001), S. 377–390 [cited at p. 32]
- [HW79] HARTIGAN, J. A. ; WONG, M. A.: Algorithm AS 136: A k-means clustering algorithm. In: *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 28 (1979), Nr. 1, S. 100–108 [cited at p. 55]
- [HW01] HJORT, N. L. ; WALKER, S. G.: A note on kernel density estimators with optimal bandwidths. In: *Statistics and Probability Letters* 54 (2001), Nr. 2, S. 153 – 159 [cited at p. 38]
- [IKK05] INUI, K. ; KIJIMA, M. ; KITANO, A.: VaR is subject to a significant positive bias. In: *Statistics and Probability Letters* 72 (2005), Nr. 4, S. 299 – 311 [cited at p. 32, 43]
- [JC85] JAIN, R. ; CHLAMTAC, I.: The P2 algorithm for dynamic calculation of quantiles and histograms without storing observations. In: *Commun. ACM* 28 (1985), Nr. 10, S. 1076–1085 [cited at p. 40]
- [JDC03] JIN, R. ; DU, X. ; CHEN, W.: The use of metamodeling techniques for optimization under uncertainty. In: *Structural and Multidisciplinary Optimization* 25 (2003), Nr. 2, S. 99–116 [cited at p. 89, 121]
- [Jen06] JENSEN, H. A.: Structural optimization of non-linear systems under stochastic excitation. In: *Probabilistic engineering mechanics* 21 (2006), Nr. 4, S. 397–409 [cited at p. 72]
- [Jin04] JIN, R.: *Enhancements of metamodeling techniques in engineering design*, University of Illinois at Chicago, Diss., 2004 [cited at p. 68, 89]
- [Jon90] JONES, M. C.: The performance of kernel density functions in kernel distribution function estimation. In: *Statistics and Probability Letters* 9 (1990), Nr. 2, S. 129–132 [cited at p. 38]
- [Jur07] JURECKA, F.: *Robust Design Optimization Based on Metamodeling Techniques*. Shaker Verlag, 2007 [cited at p. 63, 64, 78]
- [KN05] KEANE, A. J. ; NAIR, P. B.: *Computational Approaches for Aerospace Design*. John Wiley and Sons, Ltd, 2005 [cited at p. 26, 46, 47, 59]
- [KO01] KENNEDY, M. ; O'HAGAN, A.: Bayesian calibration of computer models. In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 63 (2001), Nr. 3, S. 425–464 [cited at p. 61]



- [KsZ91] KASANEN, E. ; ÖSTERMARK, R. ; ZELENY, M.: Gestalt system of holistic graphics: New management support view of MCDM. In: *Computers and Operations Research* 18 (1991), Nr. 2, S. 233 – 239 [cited at p. 54]
- [KW11] KRASNY, R. ; WANG, L.: Fast evaluation of multiquadric RBF sums by a Cartesian treecode. In: *SIAM Journal on Scientific Computing* 33 (2011), Nr. 5, S. 2341–2355 [cited at p. 4, 49]
- [KYG04] KOCH, P.N. ; YANG, R.-J. ; GU, L.: Design for six sigma through robust optimization. In: *Structural and Multidisciplinary Optimization* 26 (2004), S. 235–248 [cited at p. 78]
- [Lan10] LANGEWISCH, D. R.: *Uncertainty and sensitivity analysis for long-running computer codes: a critical review*, Massachusetts Institute of Technology, Diss., 2010 [cited at p. 61]
- [Lem09] LEMIEUX, C.: *Monte Carlo and Quasi-Monte Carlo Sampling*. Springer, 2009 [cited at p. 28]
- [Lew] LEWIS, A.: Robust regularization. – Forschungsbericht [cited at p. 64]
- [LP01] LEE, K.-H. ; PARK, G.-J.: Robust optimization considering tolerances of design variables. In: *Computers and Structures* 79 (2001), Nr. 1, S. 77–86 [cited at p. 69]
- [LP06] LEE, K.-H. ; PARK, G.-J.: A global robust optimization using Kriging based approximation model. In: *JSME International Journal Series C* 49 (2006), Nr. 3, S. 779–788 [cited at p. 89]
- [LS81] LANCASTER, P. ; SALKAUSKAS, K.: Surfaces generated by moving least squares methods. In: *Mathematics of computation* 37 (1981), Nr. 155, S. 141–158 [cited at p. 45]
- [LS10] LEVY, S. ; STEINBERG, D. M.: Computer experiments: a review. In: *ASA Advances in Statistical Analysis* 94 (2010), S. 311–324 [cited at p. 26]
- [Man01] MANTEUFEL, R. D.: Distributed hypercube sampling algorithm. In: *Third AIAA non-deterministic approaches forum paper AIAA-2001-1673, 42nd structures, structural dynamics, and materials conference, Seattle, WA. [6] Beachkofski BK, Grandhi RV, 2001* [cited at p. 28]
- [MBC79] MCKAY, M. D. ; BECKMAN, R. J. ; CONOVER, W. J.: A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. In: *Technometrics* 21 (1979), S. 239–245 [cited at p. 25, 29]
- [Mic86] MICHELLI, A.: Interpolation of scattered data: distance matrices and conditionally positive definite functions. In: *Constructive Approximation* 2 (1986), S. 11–22 [cited at p. 47]
- [Mie99] MIETTINEN, K.: *Nonlinear Multiobjective Optimization*. Kluwer, Boston, MA., 1999 [cited at p. 52]

- [ML06] MOURELATOS, Z. P. ; LIANG, J.: A Methodology for Trading-Off Performance and Robustness Under Uncertainty. In: *Journal of Mechanical Design* 128 (2006), Nr. 4, S. 856–863 [cited at p. 73, 77]
- [MM95] MORRIS, M. D. ; MITCHELL, T. J.: Exploratory designs for computational experiments. In: *Journal of statistical planning and inference* 43 (1995), Nr. 3, S. 381–402 [cited at p. 26]
- [MPV12] MONTGOMERY, D. C. ; PECK, E. A. ; VINING, G. G.: *Introduction to linear regression analysis*. Bd. 821. Wiley, 2012 [cited at p. 45]
- [NAM<sup>+</sup>92] NAIR, V. N. ; ABRAHAM, B. ; MACKAY, J. ; BOX, G. ; KACKER, R. N. ; LORENZEN, T. J. ; LUCAS, J. M. ; MYERS, R. H. ; VINING, G. G. ; NELDER, J. A. ; PHADKE, M. S. ; SACKS, J. ; WELCH, W. J. ; SHOEMAKER, A. C. ; TSUI, K. L. ; S. ; WU, C. F. J.: Taguchi's Parameter Design: A Panel Discussion. In: *Technometrics* 34 (1992), Nr. 2, S. 127–161 [cited at p. 71]
- [Nie78] NIEDERREITER, H.: Quasi-Monte Carlo methods and pseudo-random numbers. In: *Bull. Amer. Math. Soc* 84 (1978), Nr. 6, S. 957–1041 [cited at p. 28]
- [NNC12a] NIKITIN, I. ; NIKITINA, L. ; CLEES, T.: Nonlinear Metamodeling of Bulky Data and Applications in Automotive Design. In: *Progress in Industrial Mathematics at ECMI 2010*. Springer Berlin Heidelberg, 2012 [cited at p. 123]
- [NNC12b] NIKITIN, I. ; NIKITINA, L. ; CLEES, T.: Stochastic Analysis and Nonlinear Metamodeling of Crash Test Simulations and their Application in Automotive Design. In: *Computational Engineering: Design, Development and Applications*. Nova Science, 2012 [cited at p. 47, 55, 56, 123]
- [Par62] PARZEN, E.: On estimation of a probability density function and mode. In: *Annals of Mathematical Statistics* 33 (1962), S. 1065–1076 [cited at p. 34]
- [Par79] PARZEN, E.: Nonparametric Statistical Data Modeling. In: *Journal of the American Statistical Association* 74 (1979), Nr. 365, S. 105–121 [cited at p. 36]
- [PCG07] PADULO, M. ; CAMPOBASSO, M. S. ; GUENOV, M. D.: Comparative analysis of uncertainty propagation methods for robust engineering design. In: *Int. Conf. Eng. Design ICED07. Paris, 2007* [cited at p. 68, 119]
- [Pha89] PHADKE, M. S.: *Quality Engineering Using Robust Design*. Prentice Hall PTR, 1989 [cited at p. 70]
- [PKHLL06] PARK, G.-J. ; K.-H.HWANG ; LEE, T.-H. ; LEE, K.-H.: Robust Design: An Overview. In: *AIAA Journal* 44 (2006), Januar, S. 181–191 [cited at p. 62]

- [QHS<sup>+</sup>05] QUEIPO, N. V. ; HAFTKA, R. T. ; SHYY, W. ; GOEL, T. ; VAIDYANATHAN, R. ; TUCKER, P. K.: Surrogate-based analysis and optimization. In: *Progress in Aerospace Sciences* 41 (2005), Nr. 1, S. 1–28 [cited at p. 55]
- [RA99] ROBINSON, D.G. ; ATCITY, C.: Comparison of Quasi- and Pseudo-Monte Carlo Sampling for Reliability and Uncertainty Analysis. In: *Proceedings of the AIAA Probabilistic Methods Conference, St. Louis MO, AIAA99-1589*, 1999 [cited at p. 28]
- [Rac01] RACKWITZ, R.: Reliability analysis—a review and some perspectives. In: *Structural Safety* 23 (2001), Nr. 4, S. 365 – 395 [cited at p. 41, 78]
- [RBGP06] ROMERO, V. J. ; BURKARDT, J. V. ; GUNZBURGER, M. D. ; PETERSON, J. S.: Comparison of pure and „Latinized“ centroidal Voronoi tessellation against various other statistical sampling methods. In: *Reliability Engineering and System Safety* 91 (2006), Nr. 10–11, S. 1266 – 1280 [cited at p. 29]
- [RCR] RHEIN, B. ; CLEES, T. ; RUSCHITZKA, M.: A Simulation framework for robust optimization based on metamodels. In: *Nafems World Congress 2013, Salzburg, Austria* [cited at p. 5]
- [RCR12] RHEIN, B. ; CLEES, T. ; RUSCHITZKA, M.: Uncertainty quantification using nonparametric quantile estimation and metamodeling. In: *CD-ROM Proceedings of the 6th European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS 2012)*, 2012 [cited at p. 5]
- [RCR14] RHEIN, B. ; CLEES, T. ; RUSCHITZKA, M.: Robustness measures and numerical approximation of the cumulative density function of response surfaces. In: *Communications in Statistics - Simulation and Computation* 43 (2014), Nr. 1, S. 1–17 [cited at p. 5]
- [RR13] RHEIN, B. ; RUSCHITZKA, M.: Virtual Robust Prototype – robust optimization for long-running simulations. In: *XXV. Deutsch-Polnisches Wissenschaftliches Seminar, Köln*, 2013 [cited at p. 5]
- [Rub81] RUBINSTEIN, R. Y.: Simulation and the Monte Carlo method. In: *John Wiley and Sons, New York* (1981) [cited at p. 25]
- [SH08] STINSTRAS, E. ; HERTOGE, D. D.: Robust optimization using computer experiments. In: *European Journal of Operational Research* 191 (2008), Nr. 3, S. 816–837 [cited at p. 61, 64]
- [Sha98] SHANKAR, B.: *An optimal choice of bandwidth for perturbed sample quantiles*, Texas Tech University, Diss., 1998 [cited at p. 36]
- [SJ08] SCHUËLLER, G. I. ; JENSEN, H. A.: Computational methods in optimization considering uncertainties—an overview. In: *Computer Methods in Applied Mechanics and Engineering* 198 (2008), Nr. 1, S. 2–13 [cited at p. 78]

- [SK02] SEO, H.S. ; KWAK, B.M.: Efficient statistical tolerance analysis for general distributions using three-point information. In: *International Journal of Production Research* 40 (2002), Nr. 4, S. 931–944 [cited at p. 68]
- [SLJ<sup>+</sup>09] SHIMOYAMA, K. ; LIM, J. N. ; JEONG, S. ; OBAYASHI, S. ; KOISHI, M.: Practical implementation of robust design assisted by response surface approximation and visual data-mining. In: *Journal of Mechanical Design* 131 (2009), Nr. 6, S. 061007 [cited at p. 55, 89]
- [SM90] SHEATHER, S. J. ; MARRON, J. S.: Kernel Quantile Estimators. In: *Journal of the American Statistical Association* 85 (1990), S. 410–416 [cited at p. 32, 36, 38, 39, 42]
- [Sob67] SOBOL, I. M.: On the distribution of points in a cube and the approximate evaluation of integrals. In: *U.S.S.R. Comput. Math. Math. Phys.* 7 (1967), S. 86–112 [cited at p. 28]
- [SPKA97] SIMPSON, T. ; PEPLINSKI, J. ; KOCH, P. N. ; ALLEN, J. K.: On the use of statistics in design and the implications for deterministic computer experiments. In: *Design Theory and Methodology-DTM'97* (1997), S. 14–17 [cited at p. 55]
- [SPKA01] SIMPSON, T. ; PEPLINSKI, J. ; KOCH, P. N. ; ALLEN, J. K.: Metamodels for computer-based engineering design: survey and recommendations. In: *Engineering with computers* 17 (2001), Nr. 2, S. 129–150 [cited at p. 55]
- [SSG06] SWILER, L. P. ; SLEPOY, R. ; GIUNTA, A. A.: Evaluation of sampling methods in constructing response surface approximations. In: *AIAA Paper 2006–1827*, 2006 [cited at p. 28]
- [Sud12] SUDRET, B.: Meta-models for structural reliability and uncertainty quantification. In: *Proc. 11th ASCE Specialty Conference on Probabilistic Mechanics and Structural Reliability, Notre Dame, USA, 2012* [cited at p. 90]
- [Suh01] SUH, N.P.: *Axiomatic Design: Advances and Applications*. Oxford University Press, 2001 [cited at p. 58]
- [SV08] SFAKIANAKIS, M. E. ; VERGINIS, D. G.: A New Family of Nonparametric Quantile Estimators. In: *Communications in Statistics - Simulation and Computation* 37 (2008), S. 337–345 [cited at p. 39, 42]
- [SWMW89] SACKS, J. ; WELCH, W. J. ; MITCHELL, T. J. ; WYNN, H. P.: Design and analysis of computer experiments. In: *Statistical science* 4 (1989), Nr. 4, S. 409–423 [cited at p. 45]
- [Tag86] TAGUCHI, G.: *Introduction to Quality Engineering : Designing Quality into Products and Processes*. Tokyo, Japan : Asian Productivity Organization, 1986 [cited at p. 70]
- [TCT00] TAGUCHI, G. ; CHOWDHURY, S. ; TAGUCHI, S.: *Robust Engineering-Learn How to Boost Quality While Reducing Costs and Time to Market*. New York : McGraw-Hill, 2000 [cited at p. 58]

- [Tro96] TROSSET, M.: Taguchi and Robust Optimization / Department of Computational and Applied Mathematics, Rice University. 1996 (31). – Forschungsbericht [cited at p. 63, 71]
- [TSK05] TAN, P.-N. ; STEINBACH, M. ; KURMAN, V.: *Introduction to Data Mining*. Addison-Wesley, 2005 [cited at p. 55]
- [Tur93] TURLACH, B. A.: Bandwidth Selection in Kernel Density Estimation: A Review. In: *Discussion paper 9317, Institut de Statistique, UCL Louvainla Neuve, Belgium*, 1993 [cited at p. 68]
- [VM90] VINING, G. G. ; MYERS, R. H.: Combining Taguchi and response surface philosophies: a dual response approach. In: *Journal of quality technology* 22 (1990), Nr. 1 [cited at p. 89]
- [WJ95] WAND, M. M. P. ; JONES, M. M. C.: *Kernel smoothing*. Bd. 60. Crc Press, 1995 [cited at p. 38]
- [WS07] WANG, G. G. ; SHAN, S.: Review of metamodeling techniques in support of engineering design optimization. In: *Journal of mechanical design* 129 (2007), S. 370–380 [cited at p. 47]
- [Yan85] YANG, S.-S.: A Smooth Nonparametric Estimator of a Quantile Function. In: *Journal of the American Statistical Association* 80 (1985), Nr. 392, S. 1004–1011 [cited at p. 36]
- [YCDG07] YOUN, B. D. ; CHOI, K. K. ; DU, L. ; GORSICH, D.: Integration of Possibility-Based Optimization and Robust Design for Epistemic Uncertainty. In: *Journal of Mechanical Design* 129 (2007), Nr. 8, S. 876–882 [cited at p. 61]
- [ZFM05] ZANG, C. ; FRISWELL, M. I. ; MOTTERSHEAD, J. E.: A review of robust optimal design and its application in dynamics. In: *Computers and Structures* 83 (2005), Nr. 4, S. 315–326 [cited at p. 72]
- [Zha07] ZHANG, Y.: General robust-optimization formulation for nonlinear programming. In: *Journal of optimization theory and applications* 132 (2007), Nr. 1, S. 111–124 [cited at p. 64]
- [ZX10] ZHAO, D. ; XUE, D.: A comparative study of metamodeling methods considering sample quality merits. In: *Structural and Multidisciplinary Optimization* 42 (2010), Nr. 6, S. 923–938 [cited at p. 28]
- [ZX11] ZHAO, D. ; XUE, D.: A multi-surrogate approximation method for metamodeling. In: *Engineering with Computers* 27 (2011), S. 139–153 [cited at p. 47]

## Erklärung

Ich versichere, daß die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit - einschließlich Tabellen, Karten und Abbildungen -, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; daß diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; daß sie - abgesehen von unten angegebenen Teilpublikationen - noch nicht veröffentlicht worden ist sowie, daß ich eine solche Veröffentlichung vor Abschluß des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen der Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. Ulrich Trottenberg betreut worden.

Nachfolgend genannte Teilpublikationen liegen vor:

- RHEIN, B.; CLEES, T.; RUSCHITZKA M.: A Simulation framework for robust optimization based on metamodels. In: *Nafems World Congress 2013, Salzburg, Austria*
- RHEIN, B.; CLEES, T.; RUSCHITZKA M.: Uncertainty quantification using non-parametric quantile estimation and metamodeling, In: *CD-ROM Proceedings of the 6th European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS 2012)*, 2012
- RHEIN, B.; RUSCHITZKA M.: Virtual Robust Prototype – robust optimization for long-running simulations, In: *XXV. Deutsch-Polnisches Wissenschaftliches Seminar, Köln*, 2013
- RHEIN, B.; CLEES, T.; RUSCHITZKA M.: Robustness measures and numerical approximation of the cumulative density function of response surfaces, In: *Communications in Statistics - Simulation and Computation* 43 (2014), Nr. 1, S. 1-17

---

Datum

---

Unterschrift