

Entwicklung Effizienter Multivariater Lebensdauertests

Efficient Multivariate Lifetime Testing

Der Fakultät Konstruktions-, Produktions- und Fahrzeugtechnik
der Universität Stuttgart
zur Erlangung der Würde eines Doktor-Ingenieurs (Dr.-Ing.) vorgelegte
Abhandlung

von
Marco Arndt, M.Sc.
aus Ravensburg

Hauptberichter: PD Dr.-Ing. habil. Martin Dazer
Mitberichter: Univ.-Prof. Dr.-Ing. Mit Berichter

Tag der mündlichen Prüfung:

Institut für Maschinenelemente der Universität Stuttgart

2026

Vorwort

Thanks for your service.

Kurzfassung/Abstract

-TODO-TODO-TODO-TODO-TODO-TODO-TODO-TODO-TODO-TODO-TODO-

TODO:

-
- Abstract / Kurzfassung / Danksagung
- Bilder prüfen
- "mark changes after your Oral presentation for review"
- "textcolor blue in definitions ändern
- "Print" Funktion aktivieren
- "nocite" Funktion deaktivieren in main Z.269

-TODO-TODO-TODO-TODO-TODO-TODO-TODO-TODO-TODO-TODO-

Inhaltsverzeichnis

Vorwort	iii
Kurzfassung/Abstract	v
Nomenklatur	xi
Abkürzungen	xii
Indizes	xiii
Formelzeichen	xiv
Abbildungsverzeichnis	xix
Tabellenverzeichnis	xxi
1 Einleitung	1
1.1 Forschungsperspektive und Problembeschreibung	2
1.2 Beitrag dieser Arbeit	3
1.3 Aufbau der Arbeit	4
2 Stand der Technik und Forschung	7
2.1 Zuverlässigkeitstechnik und Wahrscheinlichkeitstheorie	7
2.1.1 Begriffe und Definitionen	8
2.1.2 Deskriptive Statistik für Lebensdauerdaten	9
2.1.3 Parametrische Lebensdauermodelle	10
2.1.4 Parameterschätzverfahren	12
2.2 Statistische Versuchsplanung und Modellbildung	17
2.2.1 Grundlagen zur statistischen Versuchsplanung	19
2.2.2 Parameter-Screening	22
2.2.3 Statistische Versuchsplanung zur Lebensdauererprobung .	23
2.2.4 Strategische Vorgehensweisen	29

2.2.5	Optimale Versuchsplanung und Bewertungskriterien	31
2.2.6	Statistische Modellbildung	43
3	Effiziente Testplanung für die multivariate Lebensdauererprobung .	53
3.1	Bewertung des Standes der Forschung und Technik	53
3.1.1	Eignung klassischer Response Surface Designs	53
3.1.2	Limitierungen und Effizienzbegriffe moderner Versuchspla- nung	54
3.1.3	Schlussfolgerung für die Testplanung	56
3.2	Forschungsfragen und Aufbau der Arbeit	57
3.2.1	Aufbau der Arbeit	59
4	Parameter-Screening für multifaktorielle Lebensdauertests .	61
4.1	Identifikation potenzieller Einflussgrößen	62
4.2	Kreativmethoden zum Auswahlprozesse im Parameter-Screening .	63
4.2.1	Strukturierungsmethoden	63
4.2.2	Bewertungsmethoden (Decision Making)	65
4.2.3	Finale Parameter-Diskussion	67
4.3	Randbedingungen in der Parameterauswahl für die Zuverlässigkeit- modellierung	67
4.3.1	Unterscheidung: Robustheit vs. Zuverlässigkeit	68
4.3.2	Anforderungen aus dem Versuchsplan (L-DoE)	68
4.4	Vorgehen zum heuristischen Screening für die Zuverlässigkeitmo- dellierung	69
4.4.1	<i>Phase I:</i> Systemanalyse und Fehlermechanismen	69
4.4.2	<i>Phase II:</i> Heuristische Parameterassessment und Interaktions- analyse	70
4.4.3	<i>Phase III:</i> Abgleich mit dem Versuchsdesign	70
4.5	Zusammenfassung	71
5	Effizienz und Effektivität in der multivariaten Versuchsplanung .	73
5.1	Analyse generischer Design-Abweichungen	74
5.1.1	Methodik zur simulativen Ermittlung von Auswirkungen durch Orthogonalitäts-Abweichungen	76

5.1.2	Sensitivitätsanalyse geometrischer und stochastischer Einflussgrößen	77
5.1.3	Ergebnisse der Sensitivitätsanalyse und Quantifizierung der Trennschärfe	80
5.1.4	Fazit: Statistische Effektivität und ökonomische Relevanz	82
5.2	Ökonomische Bewertung: Das Kostenmodell	83
5.3	Orthogonalitäts-Abweichungen in der Zuverlässigkeitssprognose	83
5.4	Extrapolation Variance Graph (EVG)	83
5.5	Zuverlässigkeitssmodellierung mit Vertrauensbereichen	84
5.6	Simulative Validierung und Methodennachweis	84
5.6.1	Simulationssetup	84
5.6.2	Ergebnisse und Effizienzbewertung	85
6	Fallstudie	87
7	Zusammenfassung und Ausblick	89
7.1	Key Findings	89
7.2	Diskussion	89
7.3	Ausblick	89
Literatur	91	
Anhang	103	
A	Ableitungen zur Maximum-Likelihood-Schätzung	103

Nomenklatur

Abkürzungen

<i>cdf</i>	Cumulative Distribution Function
<i>pdf</i>	Probability Density Function
ALT	Accelerated Lifetime Testing
ANOVA	Varianzanalyse, engl. Analysis of Variance
BEV	Battery Electric Vehicle
CCD	Central Composite Design
CI	Confidence Interval
CP	Center Point
DfR	Design for Reliability
DoE	Design of Experiments
DSD	Definitive Screening Design
DSM	Design-Structure-Matrix
ECU	Electronic Control Unit
EoL	End-of-Life
FBD	Function Block Diagram
FDS	Fraction of Design Space
FEA	Finite-Elemente-Analyse
FTA	Fault-Tree-Analysis
GLL	Generalized Log-Linear
GLM	Generalized Linear Model

ILP	Integer Linear Programming
IMA	Institut für Maschinenelemente
KPI	Leistungsindikator, engl. Key Performance Indicator
L-DoE	Lifetime Design of Experiments
LR-Test	Likelihood-Ratio Test
M7	Seven Management and Planning Tools
MCS	Monte-Carlo-Simulation
MLE	Maximum-Likelihood-Estimation
MMR	Median-Rank-Regression-Methode
OFAT	One Factor At Time
OLS	Ordinary Least Squares
OMARS	Orthogonal Minimally Aliased Response Surface
PH	Proportional Hazard
Q7	Seven Quality Tools
RSD	Response Surface Design
RSM	Response Surface Methodology
SPV	Scaled Prediction Variance
UPV	Unscaled Prediction Variance
VDG	Variance Dispersion Graph
VIF	Varianz-Inflations-Faktor

Indizes

0	Zeitpunkt $t = 0$
	Zentralpunkt
	Regressionskonstante
	Position im Versuchsraum
BC	Box-Cox-Transformation

C	Zentralpunkt
D	Sternpunktabstand CCD
F	Faktorieller Punkt
MC	Monte-Carlo-Simulation
M	Modellbezogen
O	Beobachtet (Observed)
S	Sternpunkt
V	Transformation
a	Simulationsfaktor: Sternpunkt-Abstand
b	Simulationsfaktor: Sternpunkt-Shift
c	Simulationsfaktor: Zentralpunkte
d	Simulationsfaktor: Sternpunkte
error	Fehler 2. Art (Wahrscheinlichkeit für ein falsch-negatives Ergebnis)
e	Simulationsfaktor: System-Rauschen
fail	Ausfallbedingt (Fehlerfolge)
fix	Fix (z. B. Rüstkosten)
f	Simulationsfaktor: Effekt-Verhältnis
g	Simulationsfaktor: Mess-Rauschen
h	Simulationsfaktor: Signifikanzniveau
nc	Nichtzentral (non-central)
sim	Simulationsergebnisse
spec	Spezifisch (pro Prüfling)
test	Versuchs-/Testbezogen
time	Zeitabhängig (pro Zeiteinheit)
total	Gesamt (Summe aller Anteile)
<i>i</i>	Laufvariable
<i>j</i>	Laufvariable

k	Laufvariable (für Faktor-Indizierung)
l	Laufvariable
o	Obere Grenze des Vertrauensbereichs
q	Quantilwert
u	Untere Grenze des Vertrauensbereichs
z	Zensierungsindikator

Formelzeichen

A_{opt}	A-Optimalitätskriterium (Spur-Minimierung)
α	Signifikanzniveau
b	Abstand der Sternpunkte (Axial Points)
β	Weibull-Formparameter (Weibull-Modul)
β	Koeffizientenmatrix
β	Regressionskoeffizient, Modellparameter
C	Fehler 2.Art
C	Kosten (allgemein)
Ω	Kandidatenmenge (Menge aller möglichen Faktorstufenkom- binationen)
$\chi^2(\cdot)$	Chi-Quadrat-Verteilung
$\text{Cov}[\cdot]$	Covarianz-Operator
COVRATIO	Einflussstatistik auf Schätzwerte (Covariance Ratio)
D	Cook's Distanz
d_f	Anzahl der Freiheitsgrade
D_{opt}	D-Optimalitätskriterium (Determinanten-Maximierung)
γ	Binäre Entscheidungsvariable für die Design-Optimierung (1 = Versuchspunkt ausgewählt)
δ	Statusindikator (1=Ausfall, 0=Zensiert)

DFBETAS	Einflussstatistik auf Koeffizienten (Difference in Betas)
DFFITS	Einflussstatistik auf Prädiktionen (Difference in Fits)
$E(\cdot)$	Erwartungswert
E	Effekt
A_{eff}	A-Effizienz
D_{eff}	D-Effizienz (Normierte Determinante)
G_{eff}	G-Effizienz (Verhältnis Modellparameter zu max. Varianz)
I_{eff}	I-Effizienz
V_{eff}	V-Effizienz
λ	Eigenwert (einer Matrix)
	Transformationsparameter (Box-Cox)
ϵ	Fehlerterm-Matrix
ϵ	Fehlerterm
$E(\cdot)$	Erwartungswert
η	Linearer Prädiktor (Logarithmus der charakteristischen Lebensdauer)
$F(\cdot)$	Ausfallwahrscheinlichkeit, Verteilungsfunktion
$f(\cdot)$	Dichtefunktion
\mathbf{F}	Fisher-Informationsmatrix
$G(\cdot)$	Verteilungsfunktion der kleinsten Extremwerte
$g(\cdot)$	Abgeleitete Funktion (für Delta-Methode)
\mathbf{g}'	Gradientenvektor der Funktion g
G_{opt}	G-Optimalitätskriterium (Minimierung max. Prädiktionsvarianz)
$\Gamma(\cdot)$	Gamma-Funktion
\mathbf{H}_f	Hessian-Matrix (der Log-Likelihood-Funktion)
h	Eintrag von \mathbf{H} , Hebelwert
H_0	Nullhypothese

Nomenklatur

H_1	Alternativhypothese
\mathbf{H}	Orthogonale Projektionsmatrix, Prädiktionsmatrix, Hat-Matrix
I_{opt}	I-Optimalitätskriterium (Minimierung durchschn. Prädiktionsvarianz)
k	Anzahl der Modellparameter
$\mathcal{L}(\cdot)$	Likelihood-Funktion
$\Lambda(\cdot)$	Log-Likelihood-Funktion
$\lambda(t)$	Ausfallrate (zeitabhängig)
LR	Likelihood-Ratio
\mathbf{M}	Informationsmatrix $\mathbf{X}'\mathbf{X}$
m	Replikation
μ	Lageparameter
N	Versuchszahl
n	Stichprobenumfang Anzahlvariable
k_f	Anzahl der Faktoren (Dimension)
p	Anzahl der Regressionskoeffizienten
p -Wert	Überschreitungs-, Rest-Irrtumsd wahrscheinlichkeit
p_f	Fraktionsgrad
$power$	Trennschärfe, Power
\Pr	Wahrscheinlichkeit
$\nu(\mathbf{x}_0)$	Prädiktionsvarianz an der Stelle \mathbf{x}_0 (skaliert)
q	Quantilwert
R	Zuverlässigkeit
r	Studentisiertes Residuum
r_{CS}	Cox-Snell-Residuum
r_{Dev}	Standardisiertes Devianz-Residuum

r	Standardisiertes Residuum (bezogen auf die SEV-Verteilung)
\mathbb{R}	Menge der reellen Zahlen
s	Empirische Standardabweichung (von x)
s^2	Empirische Varianz (von x)
$SEV(\cdot, \cdot)$	Verteilung der kleinsten Extremwerte, engl. Smallest Extreme Value
σ	Standardabweichung (theoretisch)
	Skalenparameter
σ^2	Varianz der Lebensdauer (theoretisch)
T	Charakteristische Lebensdauer (Skalenparameter)
t	Zeit, Lebensdauermerimal
	Teststatistik, Signal-to-Noise Ratio
\mathbf{t}	Vektor der Ausfallzeiten
τ	kontinuierliche Zufallsvariable
θ	Modellparameter, Schätzung
Θ	Vektor der Modellparameter
v	Integrationsvariable (Gamma-Funktion)
\mathbf{V}	Varianz-Kovarianz-Matrix
V_{opt}	V-Optimalitätskriterium (Punktuelle Varianz-Minimierung)
$\text{Var}[\cdot]$	Varianz-Operator
$\mathcal{W}(\cdot, \cdot)$	Weibull-Verteilung
w	Hilfsgröße zur Berechnung von Vertrauensbereichen (standardisierte Log-Lebensdauer)
\mathbf{X}	Versuchsplan-Matrix oder Modellmatrix
x	Messwert, Variable, Parameter
\bar{x}	Arithmetischer Mittelwert (empirisch)
\mathbf{x}	Postion im Parameterraum
ξ	Einflussfaktor im Meta-Design der Sensitivitätsanalyse

Nomenklatur

ξ	Vektor der Einflussfaktoren im Meta-Design der Sensitivitätsanalyse
y	Antwortmatrix
y	Systemantwort, -variable
z	Quantil der Standardnormalverteilung
z	Hilfsvariable für die Log-Likelihood-Berechnung
z_e	Standardisierte Variable der Kleinstwertverteilung (SEV); Hilfsgröße für Wald-Konfidenzintervalle (äquivalent zu w)

Abbildungsverzeichnis

2.1	Weibull $f(t)$ für ausgewählte Werte von T und b	13
2.2	Design of Experiments (DoE) Steps gemäß [41]	19
2.3	Parameter-Diagramm (P-Diagramm)	20
2.4	Schematische Darstellung der Effekte: (a) positiver Haupteffekt von Faktor x_1 , (b) Wechselwirkungseffekte zwischen x_1 und x_2	21
2.5	Standard voll-faktorieller Versuchsplan	24
2.6	Einfluss der Schrittweite auf die Approximation des Effekts E	25
2.7	Zentral zusammengesetzter Versuchsplan (CCD) mit $k = 2$	28
2.8	Strategischer Ansatz zu augmentierter Versuchsplanung für die Lebensdauererprobung nach [52, 53]	30
4.1	Funktionsstruktur mit Haupt-, Neben- und Teilfunktionen zur Identifikation von Einflussgrößen (nach Wallace et al. [91])	62
4.2	Erweitertes Ishikawa-Diagramm zur Visualisierung von Interaktionen zwischen Einflussfaktoren (adaptiert nach [73])	64
4.3	Beispielhafte Darstellung einer binären DSM zur Identifikation von Interaktionen zwischen Systemparametern (adaptiert nach [73])	66
4.4	Grid-Analyse zur Klassifizierung von Parametern in aktive, kritische, passive und träge Faktoren	67
4.5	Abgrenzung zwischen initialer Performance-Verteilung (A) und zeitabhängiger Degradation (B) im Screening-Prozess	68
4.6	Methodischer Ablauf des heuristischen Screenings als Vorstufe zum L-DoE	69
4.7	Prozedurale Volumenreduktion des Parameterraums im heuristischen Screening für Lifetime Design of Experiments (L-DoE) (adaptiert nach Arndt et al. [89])	71

Tabellenverzeichnis

2.1	Mathematische Definitionen der Zuverlässigkeitsfunktionen	8
2.2	Deskriptive Kennzahlen im Kontext der Lebensdaueranalyse	10
2.3	Funktionen und Kennzahlen der Weibull-Verteilung	12
2.4	Übersicht, Zielsetzung und mathematische Definition verschiedener Optimalitätskriterien für Versuchspläne	36
2.5	Übersicht, Zielsetzung und mathematische Definition verschiedener Effizienzkriterien für Versuchspläne	38
2.6	Übersicht gängiger Transformationen der Belastungsgrößen zur Abbildung physikalischer Lebensdauermodelle	46
5.1	Nominale Startwerte und Streuparameter des pseudo-physikalischen Simulationsmodells	77
5.2	Definition und Parametrisierung der Untersuchungsfaktoren für die Sensitivitätsanalyse	79
5.3	Zusammenfassende Statistik der Trennschärfe und Identifikation der Haupt-Einflussfaktoren. Trennung in unveränderliche Systemgrößen vs. modifizierbare Designgrößen.	81
5.4	Systematischer Power-Verlust bei Reduktion der Versuchsstützpunkte (Vergleich Augmentiertes Design vs. Minimal-Design)	81

1 Einleitung

Die Absicherung technischer Produkte und Systeme hinsichtlich ihrer Funktionalität bildet einen zentralen Bestandteil der ingenieurwissenschaftlichen Verantwortlichkeiten im Produktentwicklungsprozess. Motiviert durch Produktversprechen gegenüber der potenziellen Käuferschaft sowie bestehender Kunden, zur Wahrung des Selbstverständnisses einer Marke oder rein aufgrund regulatorischer Vorgaben soll hier im Kontext des vorgesehenen Einsatzzweckes Zuverlässigkeitsmanagement betrieben werden. So wird insbesondere aus Marktperspektive erwartet, dass ein (technisches) Produkt - ein Fahrzeug, ein Smartphone, eine Photovoltaikanlage - seine Funktionalität mindestens zum Gewährleistungs- oder Garantiezeitraum uneingeschränkt erfüllt.

Über den gesamten Produktlebenszyklus hinweg unterstützen Methoden der Zuverlässigkeitstechnik dabei, diese Anforderungen systematisch zu erfüllen. Verfahren aus dem Bereich „Safety“, explorative Datenanalysen zur Untersuchung der Produktperformance im Betrieb oder Test, effiziente Versuchsplanung zur Analyse oder zum Nachweis der Lebensdauer am Design, Methoden der beschleunigten Versuchsplanung, engl. [Accelerated Lifetime Testing \(ALT\)](#), der Aufbau probabilistischer Lebensdauermodelle sowie das Risikomanagement im Allgemeinen eignen sich für diese Herausforderung. Das zentrale Ergebnis liegt in der Ermittlung der Ausfallwahrscheinlichkeit als Komplement zur Zuverlässigkeit - also der Wahrscheinlichkeit, dass ein Produkt unter den definierten Randbedingungen eine vorgegebene Zeitdauer ohne funktionskritischen Ausfall übersteht [1]. Üblicherweise soll so nachgewiesen werden, dass das Erzeugnis dem Einfluss einer bestimmten Belastung - beispielsweise einer physikalischen oder elektrischen Kraft, einem Wärmeeintrag oder der Exposition gegenüber einer chemischen Beanspruchung - standhält. Gelingt in der Praxis keine hinreichend genaue Quantifizierung dieser probabilistischen Metrik, so liegen die Ursachen jedoch nicht zwangsläufig allein in ökonomischen Einschränkungen wie dem Zeit- und Kostenbudget für ein erforder-

liches Testing oder einem fehlenden methodischen Know-how - vielmehr könnten *mehrere* Einflussfaktoren auf die Zuverlässigkeit einwirken und sogar Wechselwirkungen ausprägen, ohne dass dies adäquat wahrgenommen oder antizipiert wird.

Moderne Produkte können schlichtweg durch multivariat bedingte Fehlermechanismen ausfallen. Die wirtschaftliche Tragweite dieser Erkenntnis wurde bereits 1996 im renommierten *Forbes Magazine* thematisiert, welches die einschlägigen Erfolge multivariable Testmethoden in der industriellen Praxis hervorhob [2].

1.1 Forschungsperspektive und Problembeschreibung

Da trotz genannter Umstände die Kundenanforderungen und Garantiebedingungen üblicherweise als unveränderlich, teils sogar als zunehmend anspruchsvoller zu verstehen sind, werden Unsicherheiten in der Lebensdauerabsicherung dann meist nur durch präventive Wartungsstrategien, durch Tolerierung von Restrisiken oder durch die Inkaufnahme nachträglicher Schadensbegrenzung behandelt. Der zugrunde liegende Gedanke: ehe ein Produkt, dessen Lebensdauerverhalten nicht quantifizierbar verstanden ist, einen kritischen Verschleißzustand erreicht, wird es im Rahmen eines festgelegten Wartungsintervalls vorsorglich ersetzt. Dabei könnte zugrunde liegen, dass schlichtweg kein physikalisches Modell oder eine ausreichend ausgeprägte empirische Datengrundlage vorhanden ist. Führt auch diese Vorsorge zu erheblichen Regress- oder Kulanzkosten, müssen ausfallschutzorientierte Maßnahmen - bis hin zu Rückrufaktionen - frühzeitig eingeplant und umgesetzt werden, um Image- und Kostenrisiken (wenn auch selten nachhaltig) zu minimieren. Um jenes zu vermeiden, muss also bereits im Vorfeld den verschiedensten Randbedingungen mithilfe der Zuverlässigkeitstechnik begegnet werden. Besonders komplexe Randbedingungen lassen sich beispielsweise durch eine hochgradige Integration von Elektrifizierung und Digitalisierung, verkürzte Entwicklungszyklen, verschärfter Kostendruck, sich per se verändernde Prioritäten aus Marktperspektive, wandelnde Materialauswahl und -komposition, leistungsoptimierte Belastungsszenarien, intensivierte Einsatzbedingungen und nicht zuletzt eine effiziente Ressourcennutzung innerhalb eines Produkts beschreiben - um nur einige zu nennen. Ein einfaches Beispiel verdeutlicht dies: Der komfortable sowie ausfallfreie Betrieb eines Fahr-

zeugs soll einerseits gegenüber verschiedensten Schadensursachen gewährleistet werden; andererseits hängt er maßgeblich vom Funktionserhalt einer inzwischen bis zu dreistelligen Anzahl an **Electronic Control Units (ECUs)** ab - während in der jüngeren Vergangenheit noch eine geringe bis mittlere zweistellige Anzahl mit nur begrenztem Funktionsumfang üblich war [3]. Oder aber der störungsfreie Betrieb digitaler Services im **Battery Electric Vehicle (BEV)** setzt bei Zentralisierung von **ECU**-Funktionen die stetige Funktionsfähigkeit der Traktionsbatterie voraus, durch deren chemische Alterung, verschiedenst beeinflusst, sich jedoch zeitgleich wiederum weitere Unsicherheiten aus verschiedenen Einflüssen eingliedern können. Ein damit betrachtetes System bildet somit ein vielfältig komplexeres Netzwerk aus teils wechselwirkenden oder direkten Eigenschaften, die zu einem Versagen führen könnten.

Damit kann festgehalten werden: Produktdesigns werden angesichts steigender Kundenanforderungen zunehmend raffinierter und im Sinne des Effizienzgedankens optimiert (z.B. durch mehr Sensorik, Rechenleistung und fortschrittliche Assistenzsysteme), gleichzeitig wird jedoch die Art und Weise eines möglichen Funktionsverlustes zunehmend komplexer. Somit rückt die Fähigkeit, Prognosen über die Lebensdauer und Ausfallwahrscheinlichkeit in Abhängigkeit von mehreren Einflussgrößen treffen zu können, zunehmend in den unternehmerischen Fokus. Über die klassische Testplanung im Rahmen der Zuverlässigkeitstechnik hinaus erfordert dies also Methoden der statistischen Versuchsplanung bei gleichzeitiger Berücksichtigung *mehrerer* Einflussfaktoren auf die Lebensdauer. Methoden wie **ALT** und die Lebensdauermodebildung behalten dabei weiterhin ihre Relevanz und bilden einen unverzichtbaren Bestandteil einer fundierten Teststrategie.

1.2 Beitrag dieser Arbeit

Ausgehend von der beschriebenen Problemstellung lässt sich der übergeordnete Beitrag dieser Arbeit wie folgt formulieren: Liegt ein komplexes technisches System vor und soll dieses hinsichtlich seiner Lebensdauer empirisch untersucht werden, um fundierte Prognosen über die Funktionalität im Betrieb treffen zu können, so müssen mehrdimensionale Lebensdaueruntersuchungen nach dem Prinzip des **DoE**

geplant werden. Neben der bloßen Implementierung von mehrdimensionale Tests für die Lebensdauererprobung berücksichtigt dieser Ansatz damit:

- eine effiziente Methodik zur gezielten Vorauswahl relevanter Faktoren aus der Gesamtheit potenzieller Systemparameter - mit dem Ziel, deren signifikanten Einfluss auf die Lebensdauer zu untersuchen;
- die Auswahl geeigneter Strategien und passender Testpläne zur statistisch abgesicherten Quantifizierung von Einflüssen auf die Lebensdauer, in Kombination mit konventionellen Zuverlässigkeitstests wie beispielsweise ALT;
- eine präzise Parameterschätzung zur mathematischen Beschreibung der Effekte auf Basis der als signifikant identifizierten Einflussgrößen;
- die Bilanzierung geeigneter Testpläne im Vergleich zu etablierten, in der Literatur bereits umfangreich diskutierten Versuchsplänen, insbesondere hinsichtlich potenzieller Abweichungen bei nicht-normalverteilten Daten.

1.3 Aufbau der Arbeit

Der allgemeine inhaltliche Aufbau der vorliegenden Arbeit kann Abb. 1 entnommen werden. So folgt auf die in diesem Abschnitt beschriebene Problemstellung sowie Ausführung über den generellen Beitrag der Arbeit weiter in Kapitel 2 der relevante Stand aus aktueller Forschung und Literatur. Kapitel 3 fasst schließlich den Forschungsbedarf zusammen und stellt das Ziel der Arbeit, aus der sich die relevanten Forschungsfragen ergeben, konkret heraus. Kapitel 4 beinhaltet die Vorstellung zu effizienten, qualitativen Screening-Methoden. Hier werden die herausgearbeiteten Vorschläge zu einer Auswahl an heuristischen Methoden für die Selektion der perspektivisch wenig relevanten Faktoren beschrieben, die für die Umsetzung in der experimentellen statistischen Datenerhebung für maßgeblich erachtet werden. Weiter werden in Kapitel 5 darauf die Rahmenbedingungen für die zur statistischen Versuchsplanung neu herausgearbeiteten Versuchsplankonfigurationen für effiziente Lebensdauertests abgeglichen und schließlich bewertet. Als Ergebnis sind neben neuen, effizienten Versuchsplänen auch die relevanten Merkmale beschrieben, die es bedarf, um Versuchspläne im Kontext von Lebensdauertests zu bewerten. Kapitel 6 .

1.3 Aufbau der Arbeit

Abschließend stellt Kapitel 7 eine Zusammenfassung über die methodische Herangehensweise und die erreichten Ergebnisse der Arbeit zusammen und ordnet diese für künftige Forschungsbestrebungen im Bereich der multivariaten Lebensdauer-Versuchsplanung ein.

2 Stand der Technik und Forschung

Dieses Kapitel stellt die für diese Arbeit erforderlichen technischen und methodischen Grundlagen bereit. Zunächst werden in Abschnitt 2.1 zentrale Begriffe und Konzepte der Zuverlässigkeitstechnik sowie das grundlegende statistische Verfahren zur Lebensdauer-Datenanalyse in Kombination mit Versuchsplänen erläutert. Darauf aufbauend folgen in Abschnitt 2.2 die Einführung und die Einordnung von DoE für Lebensdaueruntersuchungen sowie der multivariaten Lebensdauermodellierung aus dem Stand der Technik und der Wissenschaft, die beide für die Entwicklung effizienter Lebensdauerversuchspläne maßgeblich sind. Im Kontext der Lebensdauererprobung umfasst dies insbesondere typische, statistische Versuchspläne sowie Metriken und Indikatoren zur allgemeinen Bewertung der Versuchspläne.

2.1 Zuverlässigkeitstechnik und Wahrscheinlichkeitstheorie

Die Zuverlässigkeitstechnik befasst sich mit der probabilistischen Beschreibung der Lebensdauer technischer Produkte und Systeme sowie der strategischen und statistischen Planung von Lebensdauertests. Ziel ist die statistische Modellierung des Ausfallverhaltens unter Berücksichtigung der Funktionalität des Produkts bei relevanten Randbedingungen. Eine zentrale Aufgabe besteht somit in der statistischen Charakterisierung des Ausfallbegriffs mithilfe deskriptiver Statistik sowie in der Parametrisierung geeigneter Verteilungen zur Abbildung des Lebensdauerverhaltens. Die Modellierung kann - abhängig von den Randbedingungen - auf Basis *einer einzelnen* Belastungsgröße oder *mehrerer* Beanspruchungsparameter erfolgen, die gemeinsam den Produktausfall determinieren. Ein grundlegendes Verständnis des Umgangs mit zufallsverteilten Lebensdauerereignissen ist daher eine elementare Voraussetzung für die statistische Versuchsplanung im Rahmen der Zuverlässigkeitstechnik.

keitstechnik. Weiterführende Konzepte und vertiefte methodische Ansätze zur Zuverlässigkeitstechnik sowie zur statistischen Testplanung sind allen voran in der Standardliteratur von Bertsche und Dazer [1] dargelegt, an deren Vorgehensweise sich die nachfolgende Übersicht orientiert.

2.1.1 Begriffe und Definitionen

Der **Ausfall**, engl. **End-of-Life (EoL)**, eines Produkts bezeichnet das Ende der Lebensdauer durch Verlust der geforderten Funktionalität. Maßgeblich hierfür sind von außen einwirkende **Belastungen** (*einzelne* oder zeitgleich *mehrere* Einflussparameter - Kräfte und Momente im mechanischen Kontext) sowie die daraus resultierenden inneren **Beanspruchungen** (innere Kräfte, Momente und lokale Spannungen). Die Ausfallzeit wird als kontinuierliche Zufallsvariable $\tau > 0$ verstanden. Die fundamentalen Funktionen zur probabilistischen Beschreibung des Lebensdauerverhaltens sind gemeinsam mit deskriptiven Statistiken in Tabelle 2.1 zusammengefasst [1, 4, 5].

Tabelle 2.1: Mathematische Definitionen der Zuverlässigungsfunktionen

Ausfallwahrscheinlichkeit (engl. **Cumulative Distribution Function (cdf)**): Wahrscheinlichkeit eines Ausfalls bis zum Zeitpunkt t .

$$F(t) = \Pr(\tau \leq t) = \int_0^t f(t) dt \quad (2.1)$$

Zuverlässigkeit (engl. Reliability):

Überlebenswahrscheinlichkeit über den Zeitpunkt t hinaus.

$$R(t) = 1 - F(t) = \int_t^\infty f(t) dt \quad (2.2)$$

Tabelle 2.1 (Fortsetzung): Mathematische Definitionen

Wahrscheinlichkeitsdichte (engl. **Probability Density Function (pdf)**):
 Änderungsrate der Ausfallwahrscheinlichkeit (Ausfallintensität).

$$f(t) = \frac{d}{dt} F(t) = \frac{d}{dt} \Pr(\tau \leq t), \quad t \geq 0. \quad (2.3)$$

Ausfallrate (engl. Hazard-Function):

Momentanes Ausfallrisiko, bedingt auf das Überleben bis t .

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{\Pr(t < \tau \leq t + \Delta t | \tau > t)}{\Delta t} = \frac{f(t)}{R(t)} \quad (2.4)$$

2.1.2 Deskriptive Statistik für Lebensdauerdaten

Um die theoretischen Parameter der Grundgesamtheit in der Praxis nutzbar zu machen, müssen diese auf Basis empirisch ermittelter Lebensdauerdaten (Stichprobe vom Umfang n mit Werten x_1, \dots, x_n) approximiert werden. Die deskriptive Statistik liefert hierfür Methoden zur Berechnung von **Lageparametern** (Zentrum) und **Streuungsmaßen** (Breite). Tabelle 2.2 fasst die theoretischen Definitionen der Zufallsvariable τ sowie die grundlegenden Zuverlässigungsfunktionen und ihre korrespondierenden empirischen Schätzer kompakt zusammen.

Tabelle 2.2: Deskriptive Kennzahlen im Kontext der Lebensdaueranalyse

Erwartungswert und Mittelwert:

Erwartungswert μ des theoretischen Lageparameters der Grundgesamtheit. Der arithmetische Mittelwert \bar{x} dient als üblicher, **empirischer Schätzer** für eine endliche Stichprobe.

$$\mu = E[\tau] = \int_0^\infty t \cdot f(t) dt \quad \text{geschätzt durch} \quad \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (2.5)$$

Quantile und Median:

Das Quantil t_q definiert den Zeitpunkt, an dem die Verteilung $F(t)$ den Anteil $q \in [0, 1]$ erreicht. Der Median $t_{0.5}$ teilt die Fläche unter der *pdf* in zwei Hälften [6, 7].

$$F(t_q) = q \quad \text{bzw. für Median:} \quad F(t_{0.5}) = 0.5 \quad (2.6)$$

Varianz und Standardabweichung:

Die theoretische Varianz σ^2 (mittlere quadratische Abweichung) wird durch die erwartungstreue empirische Varianz s^2 geschätzt. Die empirische Standardabweichung folgt daraus zu $s = \sqrt{s^2}$.

$$\sigma^2 = E[(\tau - \mu)^2] \approx s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (2.7)$$

2.1.3 Parametrische Lebensdauermodelle

Während die deskriptiven Statistiken \bar{x} und s^2 die zentrale Tendenz und die Streuung der vorliegenden Stichprobe quantifizieren, erlauben sie keine Extrapolation oder die Modellierung der zugrundeliegenden Funktionen $F(t)$ und $f(t)$ der Grundgesamtheit. Um eine prädiktive, mathematische Beschreibung des stochastischen

Ausfallverhaltens zu erhalten, müssen die in Abschnitt 2.1.1 definierten Lebensdaueraufunktionen durch geeignete parametrische Verteilungsmodelle approximiert werden. Andernfalls können nur nichtparametrische Modellierungsansätze zur Schätzung der kumulierten Wahrscheinlichkeit in Überlebensfunktionen wie beispielsweise nach Kaplan und Meier [8] genutzt werden [4, 9]. Die Verteilungsmodelle hingegen bieten eine geschlossene mathematische Form für cdf und pdf und ermöglichen es, das komplexe Ausfallverhalten durch eine geringe Anzahl von Parametern zu charakterisieren.

Weibull-Verteilung

In der Zuverlässigkeitstechnik hat sich die **Weibull-Verteilung** aufgrund ihrer hohen Flexibilität als das am häufigsten verwendete Modell etabliert. Je nach zugrundeliegendem physikalischen Ausfallmechanismus finden jedoch auch andere statistische Verteilungen Anwendung, wie beispielsweise die **Lognormal-Verteilung** (häufig bei Ermüdungs-, Korrosions- oder Diffusionsprozessen), die **Exponential-Verteilung** (zur Modellierung von Zufallsausfällen ohne Alterungseffekte) oder die **Beta-Verteilung** (allgemein zur formenreichen Modellierung von R über dem festen Intervall $[0, 1]$). Für weitere Ausführungen dazu sei an dieser Stelle jedoch auf bereits ausreichend diskutierte Aufbereitungen von Bertsche und Dazer [1], Birolini [5], Yang [6], Rigdon et al. [9] und Hedderich und Sachs [10] verwiesen.

Die (zweiparametrische) Weibull-Verteilung ist das Standardmodell zur Beschreibung der Lebensdauer von technischen Produkten ohne die Berücksichtigung eines möglichen dritten Parameters - der ausfallfreien Zeit t_0 . Sie wird durch den **Formparameter** $b > 0$ (Weibull-Modul) und die **charakteristische Lebensdauer** $T > 0$ (Skalenparameter), welche dem 63,2-ten Perzentil $t_{0,632}$ entspricht, beschrieben. Unabhängig von b gilt hier somit: $F(T) = 1 - e^{-1} \approx 63,2\%$. Folgt die Lebensdauer-Zufallsvariable τ dieser Verteilung, wird dies mathematisch als $\tau \sim W(T, b)$ notiert. Damit ist sie in der Lage, alle drei Phasen der "Badewannenkurve" (Frühhausfälle mit $b < 1$, Zufallsausfälle $b \approx 1$, Verschleißausfälle mit $b > 1$) durch die Wahl ihrer Parametrisierung abzubilden, vgl. Bertsche und Dazer [1]. Die Einheit des Skalenparameters T entspricht der des Messwertes. Die charakteristischen Funktionen und statistischen Momente der Weibull-Verteilung sind in Tabelle 2.3 zusammengefasst.

Tabelle 2.3: Funktionen und Kennzahlen der Weibull-Verteilung

Dichte- und Verteilungsfunktion (*pdf* & *cdf*)

$$f(t) = \frac{b}{T^b} t^{b-1} \exp\left[-\left(\frac{t}{T}\right)^b\right], \quad t > 0 \quad (2.8)$$

$$F(t) = 1 - \exp\left[-\left(\frac{t}{T}\right)^b\right], \quad t > 0 \quad (2.9)$$

Ausfallrate (Hazard-Funktion)

$$\lambda(t) = \frac{b}{T} \left(\frac{t}{T}\right)^{b-1}, \quad t > 0 \quad (2.10)$$

Momente und Gamma-Funktion

Mit der Gamma-Funktion (2.11) $\Gamma(x) = \int_0^\infty v^{x-1} e^{-v} dv$ ergeben sich Erwartungswert μ (vgl. Gleichung (2.5)) und Varianz σ^2 (vgl. Gleichung (2.7)) der Weibull-verteilten Lebensdauer t :

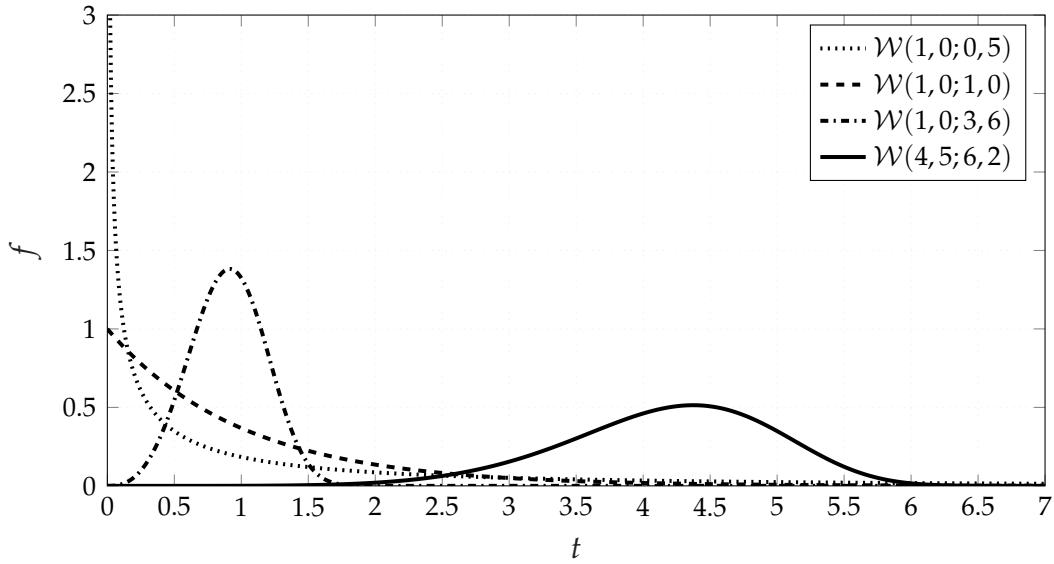
$$\mu = T \cdot \Gamma\left(1 + \frac{1}{b}\right) \quad (2.12)$$

$$\sigma^2 = T^2 \left[\Gamma\left(1 + \frac{2}{b}\right) - \Gamma^2\left(1 + \frac{1}{b}\right) \right] \quad (2.13)$$

Die Flexibilität der Verteilung zeigt sich in den Spezialfällen des Formparameters: Für $b = 1$ geht sie in die Exponentialverteilung über ($\lambda(t) = \text{konst.}$), während sie sich für $b \approx 3,6$ einer Normalverteilung annähert (vgl. Abb. 2.1) [11, 12].

2.1.4 Parameterschätzverfahren

Soll eine geschlossene mathematische Beschreibung des stochastischen Ausfallverhaltens eines Produktes gefunden werden, ist das im vorherigen Abschnitten 2.1.3 definierte parametrische Verteilungsmodell $t \sim \mathcal{W}(T, b)$ zu schätzen. Die Modell-


 Abbildung 2.1: Weibull $f(t)$ für ausgewählte Werte von T und b .

parameter der Grundgesamtheit sind in der praktischen Anwendung jedoch unbekannt. Die zentrale Problemstellung der **Parameterschätzung** besteht somit darin, aus der empirischen Stichprobe bestehend aus n Realisierungen t_1, \dots, t_n der Zufallsvariable τ statistisch fundierte Schätzwerte \hat{T} und \hat{b} zu gewinnen. Diese sind Voraussetzung, um das Lebensdauermodell (z.B. Gleichung (2.9)) zu quantifizieren und prädiktive Aussagen zu Quantilen oder der Zuverlässigkeit $R(t)$ zu ermöglichen. Eine wesentliche Komplikation hierbei sind jedoch das mögliche Auftreten von unvollständigen bzw. **zensierten** Daten sowie *multivariate* Abhängigkeiten der Belastungen zur Messgröße. Während für die Schätzung von Verteilungsparametern einfache Verfahren, wie die **Momentenmethode** oder die **Methode der kleinsten Fehlerquadrate** - engl. **Ordinary Least Squares (OLS)**, die beispielsweise bei der **Median-Rank-Regression-Methode (MMR)** im Wahrscheinlichkeitsnetz Anwendung findet, existieren, sind diese für die umfassende Analyse vielschichtiger Lebensdauerdaten in der Regel unzureichend und hier nur der Vollständigkeit wegen erwähnt - vgl. [1, 13]. Das universell anwendbare und robuste Verfahren, das Herausforderungen wie zensierte Daten und multivariate Modelle inhärent behandelt, ist die **Maximum-Likelihood-Estimation (MLE)** [4, 14].

Maximum-Likelihood-Estimation

Das Grundprinzip der MLE besteht darin, diejenigen Parameterwerte (z.B. \hat{T}, \hat{b}) als Schätzwerte auszuwählen, welche die Wahrscheinlichkeit (engl. Likelihood) maximieren, die empirisch beobachtete Stichprobe (bestehend aus unabhängigen Ausfällen und Zensierungen) zu erhalten. Mathematisch wird die Wahrscheinlichkeit der Realisierung von $\mathbf{t} = (t_1, \dots, t_n)$ einer Stichprobe durch die **Likelihood-Funktion** \mathcal{L} bestimmt. Diese ist eine Funktion des unbekannten Parametervektors $\boldsymbol{\theta}$, der k zu schätzende Parameter enthält (z.B. $\boldsymbol{\theta} = (T, b)$ mit $k = 2$).

Für den vereinfachten Fall, dass die Stichprobe ausschließlich aus n exakten Ausfallereignissen (vollständige Daten) besteht, ist die Likelihood-Funktion \mathcal{L} das Produkt der einzelnen Wahrscheinlichkeitsdichten $f(\cdot)$:

$$\mathcal{L}(\mathbf{t}|\boldsymbol{\theta}) \propto \prod_{i=1}^n f(t_i|\boldsymbol{\theta}). \quad (2.14)$$

Zur Vereinfachung der numerischen Berechnung wird in der Anwendung die **Log-Likelihood-Funktion** Λ verwendet. Durch die Logarithmierung wird das Produkt (Gleichung (2.14)) in eine äquivalente, leichter zu maximierende Summe überführt:

$$\Lambda := \ln(\mathcal{L}(\boldsymbol{\theta})) \propto \sum_{i=1}^n \ln [f(t_i|\boldsymbol{\theta})]. \quad (2.15)$$

Wie zuvor dargelegt, ist dieser vereinfachte Ansatz für Lebensdauerdaten jedoch oft unzureichend, da er das Auftreten von zensierten Daten vernachlässigt. Für die praktische Anwendung existiert jedoch die entsprechende Erweiterung der Likelihood-Funktion um die Differenzierung etwaiger Testausgänge als *Durchläufer*. Dazu wird die Stichprobe als Paarung von t_i, δ_i für \mathbf{t} definiert, wobei t_i der beobachteten Zeit und δ_i einem Statusindikator ($\delta_i = 1$ für einen exakten Ausfall; $\delta_i = 0$ für eine Rechts-Zensierung) entspricht [4, 15]. \mathcal{L} für rechts-zensierte Lebensdauerdaten lautet somit:

$$\mathcal{L}(\mathbf{t}|\boldsymbol{\theta}) \propto \prod_{i=1}^n \left[f(t_i|\boldsymbol{\theta})^{\delta_i} \cdot R(t_i|\boldsymbol{\theta})^{1-\delta_i} \right] \quad (2.16)$$

und definiert die Log-Likelihood Funktion als:

$$\Lambda := \ln(\mathcal{L}(\mathbf{t}|\boldsymbol{\theta})) \propto \sum_{i=1}^n [\delta_i \cdot \ln f(t_i|\boldsymbol{\theta}) + (1 - \delta_i) \cdot \ln R(t_i|\boldsymbol{\theta})]. \quad (2.17)$$

Der Parametervektor $\hat{\boldsymbol{\theta}}$, der den Wert von $\Lambda(\boldsymbol{\theta})$ maximiert, liefert die **MLE**-Werte. Die Schätzwerte repräsentieren die (asymptotisch) effizientesten Schätzwerte für die Parameter der Grundgesamtheit. Dies erfolgt mathematisch durch Nullsetzen k partieller Ableitungen von Λ , sofern mathematisch entsprechende Schätzwerte in geschlossener Form durch $\partial\Lambda/\partial\boldsymbol{\theta} \stackrel{!}{=} 0$ identifiziert werden können [11, 16]. Andernfalls werden numerische Optimierungsalgorithmen, vgl. Newton-Raphson-Verfahren, Patternsearch und vergleichbare, dafür herangezogen - siehe weiterführend [16, 17] sowie detaillierte Untersuchungen von Kremer und Bertsche [18]. An dieser Stelle sei erwähnt, dass systematische Verzerrungen (engl. **Bias**) in $\hat{\boldsymbol{\theta}}$ aufgrund kleiner Stichprobenumfänge auftreten können [19] - jedoch auch korrigierbar sind, vgl. Arbeiten von Hirose [20] und Ross [21].

Die Qualität der Parameterschätzung beeinflusst daraus nicht nur die Prädiktionsgüte zur Schätzung der Lebensdauer oder Zuverlässigkeit - sie bedingt schließlich auch die Effizienz des Schätzverfahrens. Wird im Sinne eines effizienten Verfahrens zur multivariaten Lebensdauermodellbildung eine Methodik gesucht, ist auch die Qualität der Parameterschätzung damit entscheidend. Vertrauensbereiche, oder engl. **Confidence Intervals (CIs)**, können eine Metrik für die Qualität der Modellierung einnehmen, da sie die Unsicherheit oder *Unschärfe* in der Prädiktion bemessen.

Die **MLE** liefert nicht nur die Punktschätzer $\hat{\boldsymbol{\theta}}$, sondern auch die Quantifizierung von deren statistischer Unsicherheit (Präzision). Obwohl verschiedene Ansätze, wie die numerisch anspruchsvolleren Berechnungen nach Likelihood-Ratio-Methode, Bootstrap-Perzentil-Methode oder Monte-Carlo-Approximation existieren, ist das gängigste Verfahren zur Berechnung von **CIs** die Approximation mittels asymptotischer Normalverteilung der **MLE**-Schätzer $\hat{\boldsymbol{\theta}}$ [1, 14]. Dies erfolgt über die **Fisher-Informationsmatrix F**, welche die Information der Stichprobe über die Parameter $\boldsymbol{\theta}$ gemäß Kremer und Bertsche [22] bezüglich des Rechenaufwands und resultierender Modellqualität vergleichsweise effizient quantifiziert. So wird diese Methodik auch in gängiger Applikationen als Standard angewandt, vgl. [6, 14, 16]. In der praktischen Anwendung wird die Fisher-Informationsmatrix auf Basis der resultierenden

Schätzwerte $\hat{\theta} = \theta$ so als Schätzung zur Beobachtung nach \mathbf{F}_O verwendet [14, 23]. Diese ist definiert als die negative **Hesse-Matrix** \mathbf{H}_f der Log-Likelihood-Funktion, ausgewertet an der Stelle der **MLE**-Schätzwerte $\hat{\theta}$:

$$\mathbf{F}_O := -\mathbf{H}_f(\hat{\theta}) = - \left[\frac{\partial^2 \Lambda(\theta)}{\partial \theta_j \partial \theta_l} \right]_{\theta=\hat{\theta}}. \quad (2.18)$$

Die Matrix \mathbf{H}_f entspricht der $(k \times k)$ -Matrix der k zweiten partiellen Ableitungen von Λ (vgl. Gl. (2.17)). Eine Invertierung $\hat{\mathbf{F}}_O^{-1}$ ergibt die geschätzte **Varianz-Kovarianz-Matrix** $\hat{\mathbf{V}}$:

$$\hat{\mathbf{V}} \approx \hat{\mathbf{F}}_O^{-1}. \quad (2.19)$$

Die Diagonalelemente dieser Matrix $\hat{\mathbf{V}}_{jj}$ entsprechen den Varianzen $\text{Var}(\hat{\theta}_j)$ der einzelnen Parameterschätzwerte [14]. Die Nicht-Diagonalelemente $\hat{\mathbf{V}}_{jl}$ (für $j \neq l$) repräsentieren die **Kovarianzen** $\text{Cov}(\hat{\theta}_j, \hat{\theta}_l)$ [6, 14]. Diese Kovarianzen sind von entscheidender Bedeutung, da sie die statistische Abhängigkeit zwischen den Schätzwerten (z.B. zwischen \hat{T} und \hat{b}) quantifizieren, welche für die Berechnung der **CIs** von abgeleiteten Funktionen wie $\hat{R}(t)$ erforderlich sind [4]. Basierend auf der Annahme der asymptotischen Normalität der Schätzer wird ein zweiseitiges $(1 - \alpha)$ -CI für einen einzelnen Parameter $\hat{\theta}_j$ direkt aus dessen Varianz approximiert durch:

$$[\theta_{j,u}, \theta_{j,o}] = \hat{\theta}_j \pm z_{1-\alpha/2} \cdot \sqrt{\hat{\mathbf{V}}_{jj}}, \quad (2.20)$$

wobei $z_{1-\alpha/2}$ dem $(1 - \alpha/2)$ -Quantil der Standardnormalverteilung entspricht. Da Lebensdauerparameter (z.B. T, b) üblicherweise auf $\mathbb{R}_{>0}$ beschränkt sind, werden **CIs** robust über eine Log-Transformation der Parameter berechnet, um physikalisch unmögliche (negative) Intervallgrenzen zu vermeiden [4, 6, 14]:

$$[\theta_{j,u}, \theta_{j,o}] = \hat{\theta}_j \exp \left(\pm z_{1-\alpha/2} \cdot \frac{\sqrt{\hat{\mathbf{V}}_{jj}}}{\hat{\theta}_j} \right). \quad (2.21)$$

Die Berechnung dieser **CIs** für Schätzwerte $g(\hat{\theta})$ zu Größen wie $\hat{R}(t)$ oder \hat{t}_q erfolgt mittels **Delta-Methode** [4, 14]. Dieses auf einer Taylor-Reihenentwicklung basierende Verfahren (Gauß'sche Fehlerfortpflanzung) approximiert die Varianz der Funktion $\hat{g} = g(\hat{\theta})$ unter Einbeziehung der gesamten Varianz-Kovarianz-Matrix.

Dazu wird der **Gradientenvektor** \mathbf{g}' der Funktion g (z.B. $g = R(t)$) bezüglich des k -dimensionalen Parametervektors Θ gebildet:

$$\mathbf{g}' := \left[\frac{\partial g(\Theta)}{\partial \theta_1}, \dots, \frac{\partial g(\Theta)}{\partial \theta_k} \right]_{\Theta=\hat{\Theta}}^T. \quad (2.22)$$

Die approximierte Varianz $\text{Var}(\hat{g})$ der Funktion ergibt sich dann aus:

$$\text{Var}(\hat{g}) \approx \mathbf{g}'^T \mathbf{V} \mathbf{g}'. \quad (2.23)$$

Das Vertrauensintervall für die Funktion \hat{g} wird anschließend unter Verwendung dieser Varianz (bzw. des Standardfehlers $\sqrt{\text{Var}(\hat{g})}$) analog zu Gleichung (2.20) berechnet [6, 14, 24]:

$$[\underline{g}_u, \bar{g}_o] = \hat{g} \pm z_{1-\alpha/2} \sqrt{\text{Var}(\hat{g})}. \quad (2.24)$$

Sollen auch hier nur positive Werte für g berücksichtigt werden, kann eine Logarithmierung in der Berechnung der Vertrauensbereiche analog zu Gl. 2.21 erfolgen [6].

2.2 Statistische Versuchsplanung und Modellbildung

Multivariate Lebensdauertests erfordern definitionsgemäß die Betrachtung mehrerer $k \geq 2$ Einflussfaktoren als Versuchsparameter. Dementsprechend entscheidend ist das Verständnis der wesentlichen Grundlagen im Umgang mit statistischer Versuchsplanung (**DoE**) für Lebensdauerdaten - auch unter dem Begriff **L-DoE** zusammengefasst - sowie der darauffolgenden Lebensdauermodellbildung. Während detaillierte Übersichten zur Historie von **DoE** von anfänglichen einschlägigen Beschreibungen durch Fisher [25], außerdem maßgebliche Weiterentwicklungen durch Box et al. [26] oder evolutionäre Schritte durch Taguchi [27] ausgiebig in Werken von Rigdon et al. [9], Kleppmann [28] und Montgomery [29] beschrieben sind, wird im Folgenden auf die wesentlichen Inhalte für die Forschungsschwerpunkte eingegangen.

Der primäre Anspruch von **DoE** besteht in der effizienten Planung empirischer Datenerhebungen, um Zielgrößen in Abhängigkeit erklärender Variablen zunächst robust zu modellieren und schließlich zu optimieren. Dieses Paradigma lässt sich

auch unter der Begrifflichkeit **Design for Reliability (DfR)** unmittelbar wiedererkennen und so auf die Analyse von Lebensdauer und Zuverlässigkeit übertragen [6, 30]. Da das lokale oder globale Optimum der Lebensdauer- bzw. Zuverlässigkeitsfunktion eines Produktes *a priori* meist unbekannt ist, erfordert dessen Identifikation eine systematische Exploration des Parameterraumes. Eine besondere Herausforderung stellt hierbei zusätzlich die Integration von **ALT** dar: Die Diskrepanz zwischen dem hochbelasteten Testraum (engl. **Design Space**) und dem regulären **Prädiktionsraum** (engl. **Use Space** oder **Field Space**) kann eine **Extrapolation** erforderlich machen, welche die Anforderungen an die Daten- und somit auch an die Designqualität deutlich verschärft. Für umfassendere Ausführungen zu Forschungserkenntnissen in **ALT** sei hierbei insbesondere auf Arbeiten von Meeker et al. [4], Nelson [14], Meeker und Escobar [31] und Elsayed und Zhang [32] verwiesen. Das Unwissen zur tatsächlichen Lage optimaler Antwortwerte und die Möglichkeit, per se einen systematischen Offset zwischen Design- und Field-Space durch **ALT** vorzufinden, stellen Teststrategien nach Best-Guess Ansätzen nachteilig. Hier wird in der industriellen Praxis häufig fälschlicherweise ein **One Factor At Time (OFAT)**-Testing Ansatz gewählt - unabhängig, ob vom Vorhandensein von Lebensdauer- oder anderen Daten, welcher schlichtweg die Wahrscheinlichkeit, Optimalstellen im Parameterraum systematisch zu treffen, senkt und somit gegenüber **L-DoE** nachteilig ist, vgl. [29, 33]. Da nun die geometrische Struktur eines Versuchsplans die erreichbare Modellierungsqualität deterministisch begrenzt, ist eine präzise Bewertung der Plangüte anhand genau dieser Eigenschaft im Vorfeld unerlässlich. Hierfür können objektive **Performance-Indikatoren** sowie mathematische **Optimalitätskriterien** dienen, vgl. [29, 34]. Ergänzend zu rationalen Metriken wie der statistischen **Trennschärfe** (engl. **Power**) und dem Schätzergebnis einer **Koeffizienten-** bzw. Parameterschätzung sind diese Größen damit bestimmend für effiziente multivariate Lebensdauertests.

Vor diesem Hintergrund fokussiert sich dieser Abschnitt auf eine gezielte Auswahl an Grundbegriffen und Metriken für multivariate Testpläne im Kontext von **L-DoE** sowie auf eine Übersicht der für Lebensdauertests geeigneten Versuchspläne, bevor abschließend die statistische Modellbildung beleuchtet wird.

Für eine grundsätzlichere Auseinandersetzung mit konventionellen Methoden und Werkzeugen von **DoE** sei, mit Blick auf den Fokus der vorliegenden Arbeit, hingegen auf die einschlägige Literatur von Kleppmann [28], Siebertz et al. [33], Hinkelmann [35] sowie vornehmlich Montgomery [29] und Myers et al. [36] ver-

wiesen. Diese Werke behandeln intensiv die Inhalte grundsätzlicher statistischer Versuchsplanung, welche um Perspektiven zu L-DoE bereits durch am Institut für Maschinenelemente (IMA) entstandene Dissertationen von Dazer [37], Herzig [38], Grundler [39] und maßgeblich durch Kremer [40] fortschreitend ergänzt wurden. Konsequenterweise werden Hintergründe zum Umgang mit normalverteilten Daten oder Abweichungen davon im Rahmen des DoE, die Diskussion zu einschlägigen Vor- und Nachteilen auch unter Abgrenzung zu Alternativen wie OFAT, die Regressionsmodellierung auf Basis der Varianzanalyse, engl. **Analysis of Variance (ANOVA)**, konventionelle Hypothesentests sowie fundamentale Ausführungen zu ALT in den nachfolgenden Ausführungen nicht explizit betrachtet, sondern als bekannt vorausgesetzt.

2.2.1 Grundlagen zur statistischen Versuchsplanung

Die Anwendung von DoE versteht sich grundsätzlich als Verfahrenskette entlang mehrerer Prozessschritte [29, 41], die beginnend von einer spezifischen Aufgabendefinition in einer statistisch abgesicherten Testentscheidung und Datenmodellierung mündet, vgl. Abbildung 2.2. Das erklärte Ziel ist es, den kausalen Zusammenhang

- 
- ① Definition des Untersuchungsziels
 - ② Identifizierung von Einflussparametern und Systemantwort
 - ③ Auswahl von Faktoren, Stufen und Parameterbereich
 - ④ Auswahl des Versuchsplans
 - ⑤ Durchführung der Versuchsreihe gemäß Versuchsplan
 - ⑥ Statistische Datenanalyse
 - ⑦ Schlussfolgerungen und Empfehlungen/Optimierung

Abbildung 2.2: DoE Steps gemäß [41]

zwischen Einflussfaktoren und Systemantwort funktional abzubilden. Darin abgebildete Einflussfaktoren sollen also per se **statistisch signifikant** und somit relevant für das Systemverhalten sein. So kann beispielsweise die zufallsverteilte Lebensdauer τ in Abhängigkeit von $k \geq 2$ technischen Beanspruchungen zunächst empirisch untersucht und anschließend modelliert sowie optimiert werden (Schritt 1 in Abbildung 2.2).

Im Zentrum der Betrachtung steht damit generell ein technisches **System**, welches abstrakt als Produkt oder Prozess verstanden wird und den Zustand der Ausgangsgröße in Abhängigkeit der definiert. Die zu untersuchende oder zu optimierende Ausgangsgröße wird als **Systemantwort y** (engl. **Response**) bezeichnet. Die gezielt kontrollierbaren und variierten Eingangsgrößen sind **Faktoren (Steuergrößen)**, während nicht kontrollierbare oder unbekannte Einflüsse als **Störgrößen (engl. Noise)** klassifiziert werden (*Schritt 2*, vgl. [28]). Eine visuelle Aufstellung des genannten Zusammenspiels der Parameter kann dem Parameterdiagramm, kurz **P-Diagramm**, in Abbildung 2.3 entnommen werden [29]. Um das Systemverhalten zu charak-

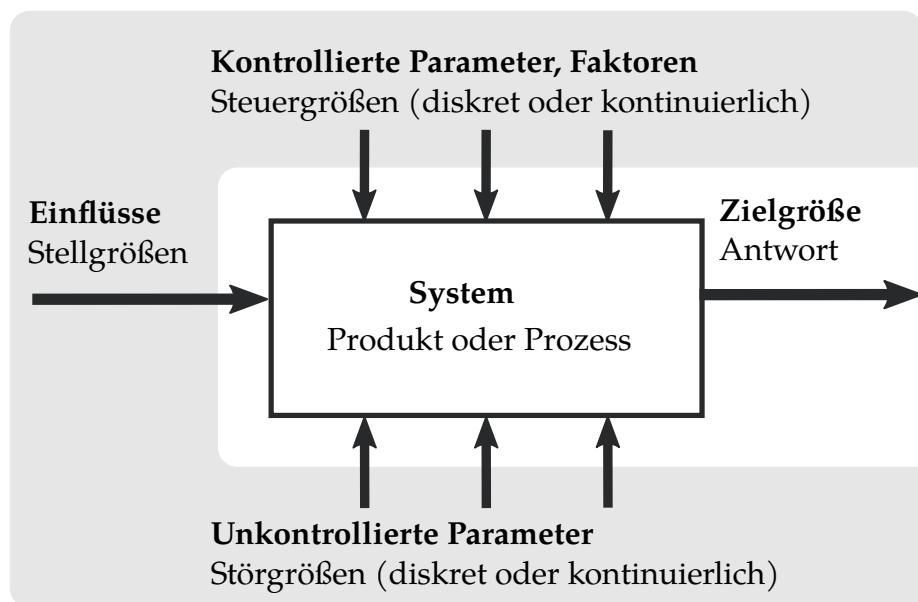


Abbildung 2.3: Parameter-Diagramm (P-Diagramm)

terisieren, werden die Faktoren als kategoriale oder kontinuierliche Parameter im Versuch auf diskreten Werten, den sogenannten **Faktorstufen** (engl. **Level**), variiert (*Schritt 3*). Dies erfolgt in aller Regel in kodierter Darstellung, so entsprechen gemäß der gängigsten Konvention die Stufe -1 der niedrigen und +1 der hohen Einstellstufe. Die planerische Kombination verschiedener Faktorstufen äußert sich in spezifischen **Versuchspunkten** innerhalb des Parameterraums und entspricht der Versuchsplan-Matrix [28, 33]. Der **Versuchsraum** (engl. **Design Space**) wird hierbei durch die Gesamtheit der technisch realisierbaren und im Versuch einstellbaren Parameterkombinationen aufgespannt. Die Auswahl geeigneter statistischer

Versuchspläne (*Schritt 4*) für die Durchführung (*Schritt 5*) wird in Abschnitt 2.2.3 detailliert behandelt.

Die aus der Variation resultierende Änderung der Systemantwort quantifiziert den Einfluss des Faktors, der statistisch als **Effekt** E bezeichnet wird und den Mittelwertunterschiede zweier Faktorstufen beschreibt (*Schritte 6-7*). Mittels **Kontrastmethode** wird also die Änderung der Systemantwort über alle durchgeführten Versuche mit jeweiligen Faktorstufen registriert [28, 29]:

$$E = \frac{1}{n|_{+1}} \sum y|_{+1} - \frac{1}{n|_{-1}} \sum y|_{-1} \quad (2.25)$$

So können mittels **DoE** strukturiert, effizient und verbindlich Informationen gewonnen werden, die über die direkten Effekte hinausgehen und differenziert Aufschluss über **Haupteffekte** sowie etwaige **Wechselwirkungen** der Faktoren auf die Antwort des Systems geben - vergleiche Abb. 2.4 sowie Kleppmann [28], Montgomery [29], Siebertz et al. [33] und Kremer [40]. Abbildung 2.4 visualisiert derartige Effekte. So gibt die Darstellung eines Haupteffekts in Abhängigkeit des Vorzeichens und der Steigung (**positiver** oder **negativer** Haupteffekt) sozusagen die Einflussstärke und -richtung wieder, während bei Wechselwirkung der Effekt in Abhängigkeit der Einstellung eines **Co-Faktors** dargestellt wird. Diese Zusammenhänge können

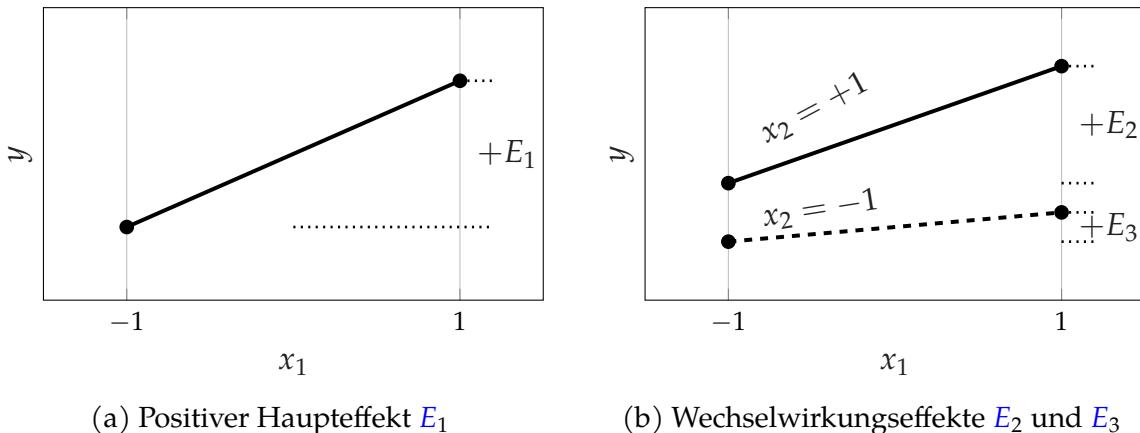


Abbildung 2.4: Schematische Darstellung der Effekte: (a) positiver Haupteffekt von Faktor x_1 , (b) Wechselwirkungseffekte zwischen x_1 und x_2 .

mathematisch positiv oder negativ beschrieben sowie durch Polynomfunktionen höherer Ordnung approximiert werden, um zusätzlich beispielsweise **quadratische**

Effekte oder Mehrfachwechselwirkungen abzubilden. In einem einfachen Fall wird für die lineare Beschreibung des Einflusses von Faktoren auf eine Antwortvariable ein durchschnittlicher Effekt durch eine Regressionskonstante β_0 sowie Haupteffekte durch die Regressionskoeffizienten $\beta_j, j = 0, 1, \dots, k$ zu einem Regressionsmodell erster Ordnung geschätzt. Falls relevant, erfolgt die Ergänzung um die jeweilige Wechselwirkung und einen Fehlerterm ϵ als Zufallsvariable für Abweichungen durch Mess- und Streufehler [9, 29, 36]. Beispielhaft $k = 2$ Faktoren resultiert daraus:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \epsilon. \quad (2.26)$$

Ein Modell zweiter Ordnung enthält zudem quadratische Terme, welche üblicherweise in Optimierungsaufgaben - so auch in der Lebensdauer- und Zuverlässigkeitssanalyse - relevant werden können:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \epsilon. \quad (2.27)$$

Die Gleichung der Modellierung kann so zur einfacheren Handhabung auch in Matrixnotation notiert werden und resultiert in:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad (2.28)$$

wobei $\boldsymbol{\beta}$ als $p \times 1$ Vektor der Regressionskoeffizienten ($p = k + 1$) durch den $n \times 1$ Vektor \mathbf{y} aller Beobachtungen sowie durch die Versuchsplan-Matrix \mathbf{X} ($n \times p$ Einträge) unter Zuhilfenahme eines geeigneten Schätzverfahrens (vgl. Abschnitt 2.1.4) zu ermittelten ist [6, 16]. Sollen zunächst perspektivisch relevante Faktoren für eine versuchstechnische Untersuchung identifiziert werden, kann ein **Parameter-Screening** durchgeführt werden. Dessen Durchführung kann sowohl heuristisch als auch versuchstechnisch erfolgen.

2.2.2 Parameter-Screening

Angesichts der potenziell hohen Komplexität durch Wechselwirkungen und Nichtlinearitäten sind die in Abbildung 2.2 beschriebenen *Schritte* 2-3 als propädeutische Arbeiten für ein effizientes Testdesign zu interpretieren. Methodisch lassen sich diese unter dem Terminus **Screening** subsumieren. Screening-Schritte sind

zwischen der Definition des Untersuchungsziels und der Durchführung der physischen Screening-Experimente angeordnet (vgl. Schritt 3 in Abbildung 2.2 sowie Abschnitt 2.2.3). Daraus folgend dienen Screening-Methoden und -Versuchspläne dem Ziel, Informationsverluste bei einer minimalen Anzahl an Versuchsläufen zu begrenzen und die vitalen (*Steuergrößen*) von den trivialen (*Störgrößen*) Faktoren zu separieren, vgl. Abbildung 2.3.

Im Hinblick auf die Realisierung eines unter Zeit- und Kostenrestriktionen hochgradig effizienten DoE ist die effiziente Ausgestaltung der Screening-Strategie selbst schon von primärem Interesse. In traditionellen DoE-Ansätzen impliziert dies den Einsatz von **Kreativmethoden**, wie sie Standardliteratur von Montgomery [29] aufführen oder exemplarisch durch Kremer [40] und Gundlach [42] zusammengefasst werden. Hierbei ist ein Rückgriff auf Ergebnisse aus Experimenten, die explizit für das Forschungsziel ausgelegt wären, in dieser Phase unter Umständen noch nicht möglich. Es gilt damit zunächst, qualitativ eine rein rational erlesene Sammlung an potenziellen Einflussparametern zu erstellen, um diese dann anhand ihrer geschätzten Einflüsse auf die Systemantwort zu priorisieren. Ansätze aus der Kreativmethodik können dazu genutzt werden und fundieren auf der technischen **Systemanalyse**, die sowohl mit als auch ohne spezifisches Vorwissen über das System erfolgen kann [1]. Hilfsmittel zur Priorisierung einer hier erstellten Parametersammlung können beispielsweise Entscheidungsfindungsprotokolle, wie die **Design-Structure-Matrix (DSM)**, und Methoden aus dem Komplexitätsmanagement, z.B. **Ishikawa-Diagramm**, sein - siehe hierzu auch weiterführende Werke von Mayers [43], Pahl et al. [44], Wu und Hamada [30] und Daenzer und Haberfellner [45] sowie Lindemann et al. [46]. Das Screening liefert somit eine rational festgestellte Auswahl an möglichst wenigen Einflussparametern, die mutmaßlich den entscheidenden Anteil an statistisch begründeter Manipulation der Systemantwort tragen und sich daher für eine Untersuchung in Versuchsplänen qualifizieren. Entsprechend ist daraufhin ein geeigneter Versuchsplan für die physischen Datenerhebungen zu wählen.

2.2.3 Statistische Versuchsplanung zur Lebensdauererprobung

Standardprotokolle aus dem DoE wie der 2^k voll-faktorielle Versuchsplan eignen sich grundsätzlich auch für Lebensdaueruntersuchungen, da sich hier analog zu vergleichbar statistisch verteilter Datenlage Effekte stets als (Mittelwert-) Unterschiede

in der Beobachtung der Systemantwort aus dem Vergleich zweier Einstellstufen eines oder mehrerer Faktoren ergeben [28].

2^k Faktorielle Versuchspläne

Demzufolge kann auch ein Lebensdauer-beeinflussendes Parametersatz - beispielsweise (x_1, x_2) - voll-faktoriell auf zwei Stufen variiert und vollständig kombiniert werden, vgl. Abbildung 2.5. Ein derartiges Setup erlaubt es, die perspektivische

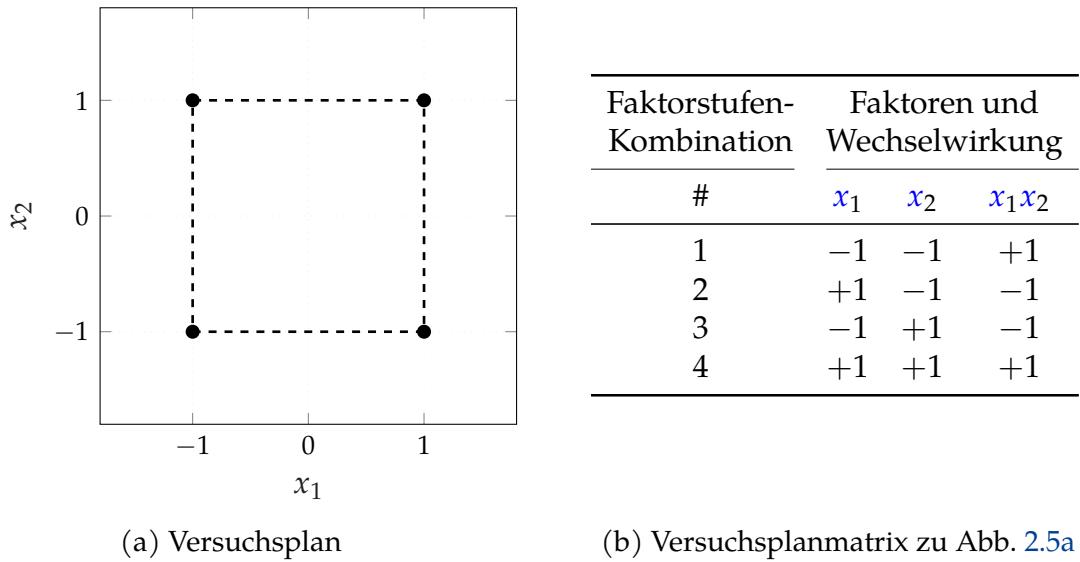


Abbildung 2.5: Standard voll-faktorieller Versuchsplan

Differenz erreichbarer **EoL**-Werte durch niedrige und hohe Beanspruchungswerte der Faktoren zu beobachten [4, 6]. Der voll-faktorielle Versuchsplan bildet somit den Standard-Versuchsplan im **DoE** und fordert bei einmaliger Durchführung (**Replikation $n = 1$**)

$$n = 2^k \quad (2.29)$$

Versuche. Dieser Stichprobenumfang stellt sicher, dass das resultierende Gleichungssystem **gesättigt** ist: Mit n Versuchen lassen sich $n - 1$ Effekte für Hauptfaktoren und Wechselwirkungen eindeutig bestimmen. Von entscheidender Bedeutung für die Aussagekraft des Versuchsplans ist die Wahl der Faktorstufen (vgl. Abbildung 2.6). Die Differenz der gewählten Level muss bereits im Vorfeld definiert werden, sodass signifikante Effekte sicher detektiert werden („**Signal-to-Noise**“), wobei gleichzei-

tig zu geringe Abstände (Rauschen) sowie zu große Intervalle (Gefahr unerkannter Nichtlinearitäten) zu vermeiden sind - vgl. Abbildung 2.6 sowie Kleppmann [28], Wu und Hamada [30] und Siebertz et al. [33]. Auf Basis eines solchen zweistufigen

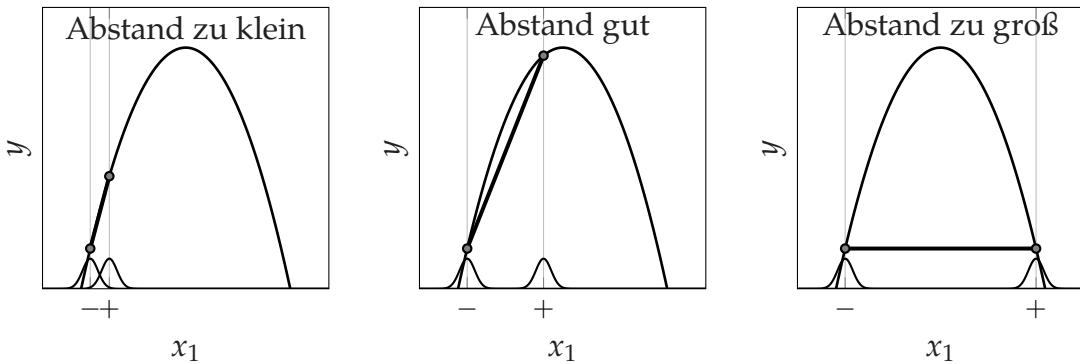


Abbildung 2.6: Einfluss der Schrittweite auf die Approximation des Effekts E

Setups lässt sich der Zusammenhang zwischen Einflussgrößen und Lebensdauer interpolieren und in einem linearen Modell abbilden, welches zudem die Schätzung von Wechselwirkungen erlaubt. Werden davon abweichende Modellterme zur Abbildung der Systemantwort erwartet, berücksichtigen alternative Versuchspläne typischerweise drei bis fünf Faktorstufen. Der voll-faktorielle Versuchsplan nimmt dabei eine entscheidende Schlüsselrolle in der strategischen Modellbildung ein - insbesondere im Hinblick auf Lebensdauerdaten und Zuverlässigkeitstechnik. Darin aufgeführte Versuchspunkte können auf Basis der Struktur ihrer Zuordnung üblicherweise ideal aus Voruntersuchungen übernommen oder durch weiterführende Untersuchungen nachfolgend erweitert werden. Ausgehend vom qualitativen Parameter-Screening (vgl. Abschnitt 2.2.1) ist ohne initiale Experimente oft unklar, welche Faktoren die Antwortvariable, also beispielsweise die Lebensdauer eines Systems, nun tatsächlich signifikant beeinflussen. Folglich ist es essenziell, diese Fragestellung vor der eigentlichen Versuchsumsetzung effizient zu klären.

2^{k-p_f} Fraktionell Faktorielle Versuchspläne

Sind nach Anwendung der Kreativmethoden (qualitatives Screening) weiterhin so viele Einflussfaktoren als relevant eingestuft, dass ein voll-faktorieller Ansatz gemäß Gleichung 2.29 zu einem wirtschaftlich nicht vertretbaren Versuchsumfang führen würde, muss die Strategie hin zu physikalischen Screening-Tests verschärft werden.

Dies empfiehlt sich insbesondere für Systeme mit $k > 5$ Faktoren, um die experimentelle Effizienz zu gewährleisten. Zur Veranschaulichung der Notwendigkeit: Bereits eine einzelne Replikation eines voll-faktoriellen Experiments mit $k = 8$ Faktoren würde $2^8 = 256$ Versuchsdurchläufe erfordern, was in der Lebensdauererprobung meist illusorisch ist.

Für derartige Selektionsaufgaben eignen sich daher **Screening-Versuchspläne**, wie der **teil-faktorielle Versuchsplan** (Fractional Factorial Design) oder alternativ der **Plackett-Burman-Plan** [28, 29, 33]. Bei diesem Ansatz wird lediglich eine selektive Teilmenge (Fraktion) der voll-faktoriellen Versuchsagenda umgesetzt, um mit minimalem Informationsverlust die für den Anwendungsfall signifikanten Effekte zu beschreiben. Mathematisch wird die Anzahl der Versuche dabei auf

$$n = 2^{k-p_f} \quad (2.30)$$

reduziert, wobei p_f den Grad der Fraktionierung (die Anzahl der Generatoren) angibt. Die Validität dieses Vorgehens stützt sich auf zwei fundamentale empirische Postulate [9, 13]:

- Die **Effekthierarchie** besagt, dass Effekte niedrigerer Ordnung – primär Haupteffekte – in der Regel eine größere Amplitude aufweisen und mit höherer Wahrscheinlichkeit signifikant sind als Effekte höherer Ordnung.
- Die **Effektvererbung** impliziert, dass das Auftreten signifikanter Wechselwirkungen oder quadratischer Terme strukturell an die Signifikanz ihrer korrespondierenden Haupteffekte gekoppelt ist.

Eine direkte Konsequenz dieser Reduktion („Der Preis der Einsparung“) ist jedoch, dass sich in teil-faktoriellen Versuchsplänen bestimmte Effekte nicht mehr isoliert betrachten lassen (vgl. Abbildung 2.5a für den Fall einer Fraktionierung). So sind beispielsweise Haupteffekte unter Umständen nicht mehr zweifelsfrei von Wechselwirkungen höherer Ordnung zu unterscheiden. Da sich diese Effekte statistisch überlagern, spricht man von einer **Vermengung** (engl. **Aliasing**). Die Schwere dieser Vermengung wird dabei über die **Auflösung** (engl. Resolution) des Versuchsplans klassifiziert (z. B. Auflösung III, IV oder V). Wird dieser Informationsverlust jedoch bewusst in Kauf genommen und ingenieurwissenschaftlich bewertet, ermöglicht dies eine signifikante Reduktion des Versuchsumfangs, um effizient die dominanten

Faktoren aus der initialen Parametermenge zu isolieren. Diese Vorgehensweise ist von hoher Relevanz, da Lebensdauertests – ob als beschleunigte Prüfung mittels ALT oder unter Feldbedingungen – durch die inhärente Zeitabhängigkeit der Systemantwort erhebliche Kapazitäten binden und Ergebnisse nicht ad hoc verfügbar sind. Im Sinne einer ressourceneffizienten Gesamtstrategie sollten die Screening-Versuche daher idealerweise so konzipiert sein, dass sie nahtlos in einen nachfolgenden, höher aufgelösten Versuchsplan integriert (**augmentiert**) werden können.

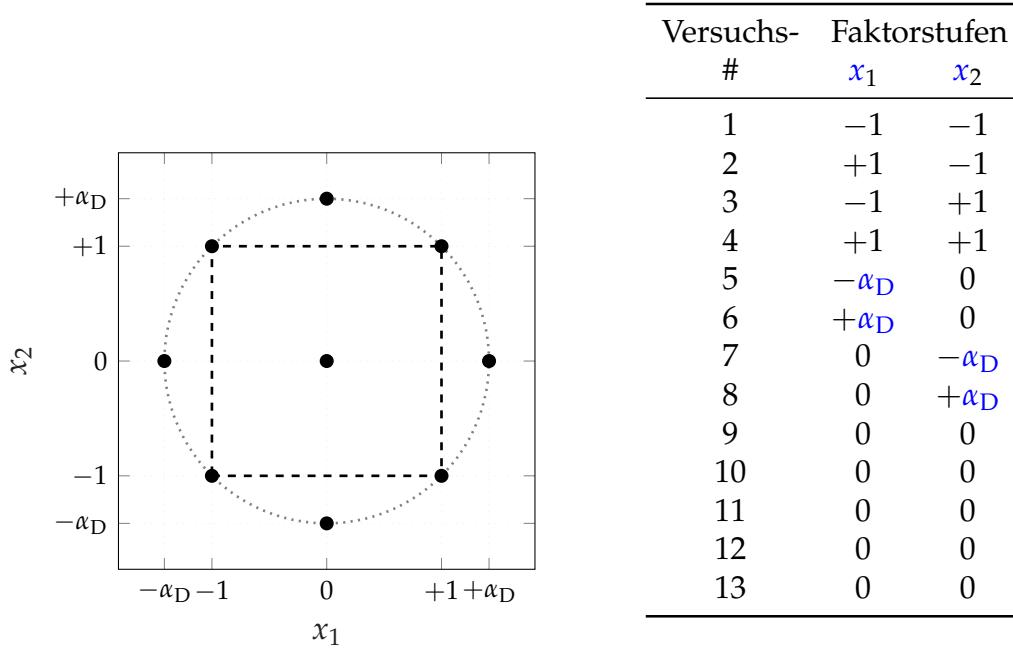
Wirkungsflächenversuchspläne

Die bisher diskutierten Versuchspläne beschränken sich auf die Untersuchung von Faktoren auf jeweils zwei Stufen (± 1). Dies ermöglicht zwar eine effiziente Darstellung linearer Beziehungen und Interaktionen, jedoch ist die Modellierung komplexerer, nicht-linearer Effekte aufgrund fehlender Stützstellen im Versuchsraum damit physikalisch nicht möglich. Perspektivisch ist daher entscheidend, wie die bestehende Datenbasis weitergenutzt und augmentiert werden kann, falls die Systemantwort signifikante **Krümmungen** (engl. **Curvature**) aufweist und die quantitativen Beziehungen zwischen Faktoren und Zielgröße für eine Optimierung detaillierter beschrieben werden müssen.

Die **Response Surface Methodology (RSM)** behandelt als Teildisziplin des DoE derartige Herangehensweisen und bietet hierfür spezielle **Wirkungsflächenversuchspläne** - engl. **Response Surface Designs (RSDs)** - an. Der erste Schritt zur Detektion von Nichtlinearitäten besteht in der Integration von n_C sogenannten **Zentralpunkten**, engl. **Center Points (CPs)**, in den faktoriellen Basisplan. Hierbei werden alle Faktoren auf die kodierte Stufe 0 (die Mitte des Versuchsraums) gesetzt. Weicht der Mittelwert der Systemantwort in den Zentralpunkten signifikant vom Mittelwert über die faktoriellen Eckpunkte ab, deutet dies auf eine Krümmung der Antwortfläche hin [29].

Um diese quadratischen Zusammenhänge explizit zu bestimmen, gilt der **Zentral-Zusammengesetzte-Versuchsplan**, engl. **Central Composite Design (CCD)**, als etablierter Standard. Ein CCD entsteht durch die Augmentierung des ursprünglichen voll- (oder teil-)faktoriellen Plans (den n_F Eckpunkten), vgl. Abbildung 2.7, um:

- eine definierte Anzahl n_C an wiederholten Zentralpunkten (üblicherweise $3 \leq n_C \leq 5$ zur Abschätzung des reinen Fehlers) sowie
- $n_S = 2 \cdot k$ zusätzliche **Sternpunkte** (engl. **Axial Points** oder **Star Points**), die auf den Achsen des Koordinatensystems im Abstand $\pm \alpha_D$ vom Zentrum liegen.



(a) Versuchsplan (CCD)

(b) Versuchsplanmatrix (Stufen) zu Abb. 2.7a

Abbildung 2.7: Zentral zusammengesetzter Versuchsplan (CCD) mit $k = 2$

Während die Zentralpunkte das Vorhandensein quadratischer Effekte validieren, ermöglichen die Sternpunkte deren Wertbestimmung. Die Wahl des Abstands α_D wird primär durch die Geometrie des interessierenden Versuchsraums (**Region of Interest**) diktiert. Betrachtet man diesen Raum als Kugel (**sphärisch**), ist die **Drehbarkeit** (**Rotierbarkeit**, engl. **Rotatability**) ein wesentliches Qualitätsmerkmal. Sie stellt sicher, dass die Varianz der Vorhersage nur vom Abstand zum Zentrum abhängt und invariant gegenüber einer Rotation des Koordinatensystems ist. Um diese Eigenschaft zu gewährleisten, berechnet sich der Abstand α_D in Abhängigkeit der faktoriellen Versuchspunkte n_F zu [29, 47]:

$$\alpha_D = (n_F)^{1/4}. \quad (2.31)$$

Alternativ kann für sphärische Räume auch $\alpha_D = \sqrt{k}$ gewählt werden (**Sphärisches CCD**), wodurch alle Versuchs- und Sternpunkte auf einer Kugeloberfläche liegen [36]. Ist der Versuchsraum hingegen durch harte physikalische Grenzen (z.B. maximale Temperatur) kubisch beschränkt, bietet sich der **flächenzentrierte CCD (Face Centered CCD)** an. Hierbei wird $\alpha_D = 1$ gesetzt, sodass die Sternpunkte direkt auf den Flächenmittnen des Würfels liegen. Dies vereinfacht die Durchführung, da nur drei Faktorstufen ($-1, 0, +1$) benötigt werden, opfert jedoch die Eigenschaft der Rotierbarkeit. Eine effiziente Alternative zum **CCD** stellt das **Box-Behnken-Design** dar [48]. Dieses Design kombiniert 2^k -Faktorielle mit unvollständigen Blockplänen und platziert Versuchspunkte auf den Kantenmittten des Versuchsraums, vermeidet jedoch die extremen faktoriellen Versuchspunkte. Dies kann vorteilhaft für Lebensdauertests sein, bei denen extreme Ecken oft zu verfrühten Ausfällen führen können. Zudem ist es bei korrekter Wahl der Zentralpunkte alias-optimal gegenüber kubischen Modellen [49]. Stößen Standard-Designs (**CCD**, Box-Behnken) aufgrund von Restriktionen im Versuchsraum (**Constraints**), nicht-standardmäßigen Modellierungszielen oder ungewöhnlichen Stichprobenumfängen an ihre Grenzen, empfiehlt sich der Einsatz von **Optimalen Versuchsplänen**, vergleiche übernächsten Abschnitt. Diese computergenerierten Designs minimieren algorithmisch die durchschnittliche Vorhersagevarianz über den gesamten Designraum und können Standard-Designs in ihrer Prädiktionsgüte oft übertreffen [29, 34].

2.2.4 Strategische Vorgehensweisen

In der Praxis entsteht ein **CCD** häufig im Rahmen einer **sequenziellen Versuchstrategie**: Zeigt das initiale lineare Modell Anpassungsmängel (**Lack-of-Fit**), wird der bestehende faktorielle Plan um die Sternpunkte augmentiert, um ein Modell zweiter Ordnung zu fitten [36]. Modelle zweiter Ordnung sind insbesondere in der Lebensdauer- und Zuverlässigkeitstechnik vermehrt von Bedeutung, da sie die Existenz von lokalen Extrema (Minima oder Maxima) der Systemantwort im Versuchsraum detaillierter abbilden können [9, 50]. Prinzipielle Vorgehensweisen, die ein erfolgreiches Umsetzen von **DoE** bzw. **L-DoE** begünstigen sollten, wie **Blockbildung** oder **Randomisierung** (vgl. [28, 33]) seien hier bereits vorausgesetzt. Abbildung 2.8 fasst eine derartige strategische Vorgehensweise zur Realisierung eines

CCD zusammen, wie sie beispielsweise Box und Wilson [51], Box et al. [52] und Bisgaard [53] vorschlagen.

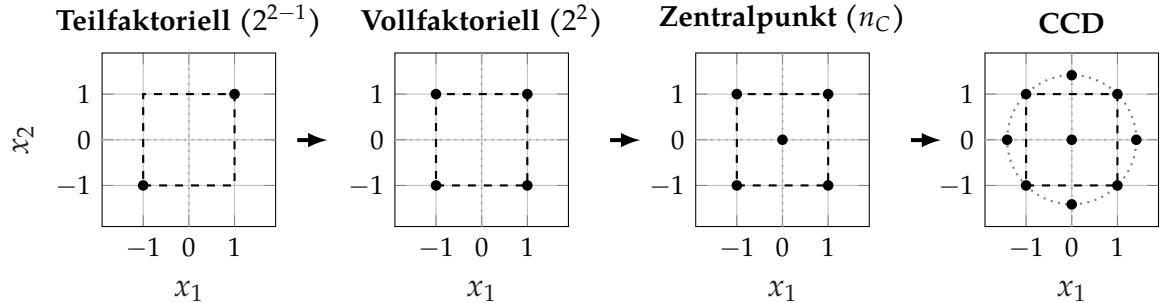


Abbildung 2.8: Strategischer Ansatz zu augmentierter Versuchsplanung für die Lebensdauererprobung nach [52, 53]

Um effizient Versuchspunkte zu allozieren und für sequentiell fortlaufende Versuche weiterverwendet zu werden, wird in Abgleich mit Abbildung 2.2 zunächst ein **teil-faktorieller Versuchsplan** zur Identifikation signifikanter Faktoren durchgeführt. Anschließend wird der Plan zu einem **voll-faktoriellen Versuchsplan** erweitert, um alle Haupeffekte und Wechselwirkungen erster Ordnung zu schätzen. Daraufhin werden **Zentralpunkte** ergänzt, um das Vorhandensein von Krümmungen zu validieren. Abschließend werden bei Bedarf die **Sternpunkte** hinzugefügt, um ein vollständiges **CCD** zu realisieren. Insbesondere in der Lebensdauererprobung ist eine derartige schrittweise Vorgehensweise sinnvoll, um den Versuchsumfang zu minimieren und dennoch eine fundierte Modellierung der Systemantwort zu gewährleisten. Gleichzeitig erlaubt diese Strategie eine flexible Anpassung an unerwartete Ergebnisse in den einzelnen Phasen der Untersuchung, während lokale Optima von Zuverlässigkeiten im Sinne von DfR im Vorfeld ohnehin nicht bekannt sind. In Abhängigkeit der Anpassung zur Wahl eines geeigneten Modells, linear oder quadratisch, kann die Versuchsplanung so stets an die aktuellen Erkenntnisse angepasst werden. Dies folgt im Wesentlichen einer übergeordneten Logik der Modellkomplexität: Gemäß Russell [54] korrespondiert die minimale Anzahl notwendiger Versuchspunkte direkt mit der Anzahl der zu schätzenden Modellparameter p . Intuitiv erfordert bereits die Bestimmung der Parameter einer einfachen Geraden (β_0, β_1) mindestens $N = 2$ unterschiedliche Stützstellen. Soll hingegen ein quadratischer Zusammenhang ($\beta_0, \beta_1, \beta_{11}$) abgebildet werden, sind zwingend mindestens drei Stützstellen erforderlich, um das Gleichungssystem zu lösen. Im Kontext optimi-

maler Versuchspläne definiert nach Pukelsheim [55] der Maßerweiterungssatz von Carathéodory dabei eine theoretische Obergrenze für die notwendigen Stützstellen von $N = p(p + 1)/2 + 1$, beziehungsweise $N = p(p + 1)/2$, sofern das Interesse der Schätzung aller Elemente des Parametervektors β gilt [54]. Es kann beispielsweise auf Basis der Ergebnisse des teil-faktoriellen Plans entschieden werden, ob eine Erweiterung zum voll-faktoriellen Plan überhaupt notwendig ist. Genauso kann die Anzahl der Zentralpunkte an die beobachtete Streuung der Lebensdauerdaten angepasst werden, um eine robuste Schätzung des Fehlers zu gewährleisten. Ist dann eine Extrapolation in Richtung a-priori unbekannter Beanspruchungsniveaus geplant, verfügt der CCD über die notwendigen Stützstellen, um eine adäquate Explorations des Versuchsraums zu ermöglichen. Zuletzt kann selbst bei Extrapolation über die Grenzen des Versuchsraums hinaus auf Basis der modellierten Krümmungen eine fundierte Abschätzung der Systemantwort erfolgen [29, 36] und sogar mit Validierungsversuchen unter realen Einsatzbedingungen hinterlegt werden [50]. So wird eine effiziente und zugleich robuste Lebensdauererprobung ermöglicht, die den Anforderungen der DfR gerecht wird und wirtschaftlich abbildbar bleibt [56].

2.2.5 Optimale Versuchsplanung und Bewertungskriterien

Wird maximale Flexibilität jenseits starrer Standard-Designs wie dem voll-faktoriellen Setup oder dem CCD verlangt, können **optimale Versuchspläne** einen leistungsfähigen, effizienten Lösungsansatz für *multivariate* Versuchsplanung bieten [29, 36, 57]. Diese verfolgen das Ziel, die Versuchspunkte algorithmisch so im Versuchsraum zu positionieren, dass spezifische Qualitätskriterien - allen voran die Verteilung der (um die Versuchszahl N skalierten) **Prädiktionsvarianz**, engl. **Scaled Prediction Variance (SPV)**,

$$\nu(\mathbf{x}_0) = \frac{N\text{Var}[\hat{y}(\mathbf{x}_0)]}{\sigma^2} = \mathbf{x}_0' \left(\frac{\mathbf{M}}{N} \right)^{-1} \mathbf{x}_0 = N \mathbf{x}_0' \mathbf{M}^{-1} \mathbf{x}_0 \quad (2.32)$$

optimiert werden [29, 36]. Die Prädiktionsvarianz ist ein dimensionsloses Maß, das ausschließlich über die geometrische Anordnung der N Versuchspunkte (der Matrix \mathbf{X}) in der **Informationsmatrix** (auch **Momentenmatrix**)

$$\mathbf{M} \stackrel{m=1}{=} \mathbf{X}' \mathbf{X} \quad (2.33)$$

definiert und damit positionsabhängig von einer Referenz $\mathbf{x}_0' = [1, \mathbf{x}_{01}, \dots, \mathbf{x}_{0k}]$ im Parameterraum ist [36]. Sie ist damit unabhängig von der tatsächlichen Streuung der Messdaten (σ^2). Weiter wird sie über die Anzahl der Versuchspunkte N skaliert, um Vergleiche zwischen Versuchsplänen mit unterschiedlicher Stichprobengröße zu ermöglichen [34, 36] - kann aber auch unskaliert betrachtet werden: als **Unscaled Prediction Variance (UPV)** = $\text{Var}[\hat{y}(\mathbf{x}_0)]/\sigma^2 = \mathbf{x}_0' \mathbf{M}^{-1} \mathbf{x}_0$, vgl. Montgomery [29] und Myers et al. [58]. Zusammen mit der Varianz des Versuchsfehlers σ^2 (vgl. Gleichung 2.35) erlaubt die Prädiktionsvarianz somit eine direkte Quantifizierung der **Unsicherheit in der Modellvorhersage** $\text{Var}[\hat{y}(\mathbf{x}_0)]$ an einer beliebigen Stelle \mathbf{x}_0 im Versuchsräum. Um sie über verschiedene Stellen im Parameterraum zu bewerten, konnten allen voran Zahran et al. [59] und Giovannitti-Jensen und Myers [60] mit dem **Variance Dispersion Graph (VDG)** oder **Fraction of Design Space (FDS)-Plot** Visualisierungsstrategien entwickeln. Eine effiziente Versuchsplanung für multivariate Lebensdaueruntersuchungen muss sich neben pragmatischen Beweggründen maßgeblich an diesen Größen orientieren, weshalb im Folgenden eine Übersicht der gängigen Metriken und Qualitätsmerkmale gegeben wird.

Der Einsatz optimaler Versuchspläne ist zunächst prädestiniert für Szenarien, in denen klassische Pläne an ihre Grenzen stoßen - sei es durch physikalische Randbedingungen welche bestimmte Faktorstufenkombinationen ausschließen, oder durch strikte Limitierungen der verfügbaren Versuchskapazität [28, 34]. Derart klassische Versuchspläne wie der voll-faktorielle Versuchsplan sind aufgrund ihrer Versuchspunktanordnung stets **orthogonal** [9, 29]. Ein Versuchsplan wird als orthogonal bezeichnet, wenn keine Korrelation zwischen jeweils zwei Spalten der Versuchsplan-Matrix vorliegt - deren Skalarprodukte also jeweils null ergeben:

$$\langle \mathbf{x}_i, \mathbf{x}_j \rangle = 0 \quad \text{für alle } i \neq j. \quad (2.34)$$

Oder in anderen Worten: entspricht die Informationsmatrix \mathbf{M} einer Diagonalmatrix, ist der Versuchsplan orthogonal [9, 13]. So können Effekte eindeutig identifiziert werden - die Vektoren der Faktorstufenkombinationen sind linear unabhängig und Effekte lassen sich unverzerrt schätzen. Dies stellt einen Versuchsplan also zunächst einmal qualitativ günstig dar.

Zudem liegt **Ausgewogenheit** vor, sofern für einen jeweiligen Faktor alle anderen Faktoreinstellungen gleichmäßig aufgeteilt sind [33]. Damit wird Varianzhomoge-

nität und Gleichbehandlung der Faktoren gewährleistet, sodass die Schätzung der Effekte unverzerrt erfolgt.

Mathematisch fundiert die Bewertung der Schätzgenauigkeit auf der Inversen der Informationsmatrix, welche als **Varianz-Kovarianz-Matrix** (oder **Dispersionsmatrix**) $(\mathbf{X}'\mathbf{X})^{-1}$ der Regressionskoeffizienten definiert ist:

$$\text{Var}(\hat{\boldsymbol{\beta}}) = \sigma^2 (\mathbf{X}'\mathbf{X})^{-1}, \quad (2.35)$$

wobei wie gehabt σ^2 die Varianz des Versuchsfehlers darstellt. Die Diagonalelemente der Matrix $(\mathbf{X}'\mathbf{X})^{-1}$ stehen hierbei in direktem Zusammenhang mit dem **Varianz-Inflations-Faktor (VIF)** [28, 33]. Der **VIF** dient als Maßzahl für Multikollinearität und quantifiziert den Faktor, um den sich die Varianz eines geschätzten Koeffizienten im Vergleich zu einem vollständig orthogonalen Design aufgrund von Korrelationen zwischen den Faktoren erhöht (vergleiche auch **Konditionszahl**) [10]. Während bei orthogonalen Plänen (Idealfall) ein **VIF** = 1 vorliegt, deuten hohe Werte (typischerweise > 5 oder > 10) auf eine instabile Modellschätzung hin [13, 28].

Ergänzend zur globalen Bewertung der Multikollinearität durch den **VIF** erlaubt die Betrachtung der sogenannten **Prädiktionsmatrix** (engl. auch **Hat-Matrix**)

$$\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \quad (2.36)$$

eine lokale Diagnose der Versuchspunkte. Die Diagonalelemente $h_{ii} \in \left[\frac{1}{n}, 1\right]$ dieser Projektionsmatrix, bezeichnet als **Hebelwerte** (engl. **Leverage**), quantifizieren den Einfluss eines einzelnen Versuchslaufs auf die Modellvorhersage. Punkte mit hohen Hebelwerten (typischerweise $h_{ii} > 2p/n$) befinden sich geometrisch weit vom Zentrum des Versuchsraums entfernt und dominieren die Regression, was den Plan anfällig für Ausreißer in diesen spezifischen Einstellungen macht [13, 33].

Während Hebelwerte jedoch lediglich das *Potenzial* einer Beobachtung zur Modellbeeinflussung aufgrund ihrer geometrischen Exponiertheit indizieren, quantifizieren Einflussstatistiken die *tatsächliche* Auswirkung auf die Regressionsparameter und die Vorhersagegüte. Die Identifikation solcher Beobachtungen erfolgt methodisch durch den iterativen Ausschluss des i -ten Datensatzes (*Leave-One-Out*-Methodik) und den Vergleich der resultierenden Modellstatistiken mit dem ursprünglichen Modell basierend auf n Beobachtungen [13, 61]. Als globales Maß

für den Einfluss der i -ten Beobachtung auf den Vektor aller geschätzten Regressionskoeffizienten $\hat{\beta}$ dient die **Cook-Distanz** (D_i). Sie verknüpft die Information des intern studentisierten Residuums r_i mit dem Hebelwert h_{ii} :

$$D_i = \frac{r_i^2}{p} \cdot \frac{h_{ii}}{1 - h_{ii}}. \quad (2.37)$$

Ein hoher Wert (typischerweise $D_i > 1$ oder $D_i > 4/n$) signalisiert, dass das Entfernen der Beobachtung zu einer signifikanten Verschiebung der Modellparameter führen würde, da der Punkt sowohl weit vom Zentrum liegt als auch einen großen Fehler aufweist [61, 62].

Für eine differenzierte Analyse auf Ebene der einzelnen Modellterme wird die Metrik **DFBETAS** (Difference in Betas, Standardized) herangezogen. Sie misst die standardisierte Änderung eines spezifischen Regressionskoeffizienten β_j bei Ausschluss der i -ten Beobachtung:

$$\text{DFBETAS}_{j,i} = \frac{\hat{\beta}_j - \hat{\beta}_{j(-i)}}{\sqrt{\hat{\sigma}_{(-i)}^2 (\mathbf{X}'\mathbf{X})_{jj}^{-1}}}, \quad (2.38)$$

wobei $\hat{\beta}_{j(-i)}$ den Koeffizienten ohne die i -te Beobachtung und $\hat{\sigma}_{(-i)}^2$ die entsprechende Fehlerquadratsschätzung darstellt. Hiermit lässt sich prüfen, ob einzelne Versuche die Schätzung eines Effekt-Terms verzerren (kritischer Schwellenwert oft $> 2/\sqrt{n}$) [13, 61].

Ergänzend beschreibt **DFFITS** (Difference in Fits) die Änderung des Vorhersagewertes an der Stelle i selbst, normiert auf die Standardabweichung der Anpassung:

$$\text{DFFITS}_i = \frac{\hat{y}_i - \hat{y}_{i(-i)}}{\sqrt{\hat{\sigma}_{(-i)}^2 h_{ii}}}. \quad (2.39)$$

Abschließend bewertet die **COVRATIO** den Einfluss auf die Präzision der Schätzung. Sie setzt die Determinante der Varianz-Kovarianz-Matrix ohne die i -te Beobachtung ins Verhältnis zur ursprünglichen Matrix:

$$\text{COVRATIO}_i = \frac{\det((\mathbf{X}'_{(-i)}\mathbf{X}_{(-i)})^{-1})}{\det((\mathbf{X}'\mathbf{X})^{-1})}. \quad (2.40)$$

Werte, die signifikant von 1 abweichen (Grenzen $1 \pm 3p/n$), zeigen an, dass die Beobachtung die Konfidenzbereiche der Koeffizienten unverhältnismäßig beeinflusst [61].

Da in restriktiven Versuchsräumen Orthogonalität oft nicht erreichbar ist, zielen optimale Versuchspläne darauf ab, durch Minimierung spezifischer Eigenschaften der Dispersionsmatrix $(\mathbf{X}'\mathbf{X})^{-1}$ den Informationsgehalt trotz Korrelationen zu maximieren. Die Generierung solcher maßgeschneiderten Pläne erfolgt algorithmisch unter Maximierung spezifischer statistischer Gütekriterien. Eine Übersicht der einschlägigen **Optimalitätskriterien** - oder einfach **Optimalitäten** - welche den quantitativen Vergleich zur Standard-Methodik ermöglichen, ist in Tabelle 2.4 zusammengefasst [29, 34]. Sie bilden mitunter die Grundlage für diverse weitere Optimalitäten und hybride Ansätze und lassen sich prinzipiell in Modellierungskriterien (**A-, D-Optimalität**) und Prädiktionskriterien (**G-, I-, U-, V- Optimalität**) unterteilen [9, 36].

Jenseits praktischer Kriterien sind hier der Vollständigkeit halber auch eine Teilmenge eher theoretischer Optimalitäten zu nennen. Die **E-Optimalität**, führt zur Minimierung des maximalen Eigenwerts der Dispersionsmatrix, was eine Worst-Case-Absicherung zur Varianzschätzung begünstigt - vergleiche Russell [54] und Boyd und Vandenberghe [63]. Für explorative Zielsetzungen ohne starre Modellannahmen kann nach Atkinson [64] die **S-Optimalität** relevant sein, da hier durch die Maximierung der euklidischen Abstände benachbarter Punkte eine gleichmäßige Raumfüllung (*Space-Filling*) sichergestellt werden kann. So können für sich ändernde Modelle entlang einzelner Parameter wie beispielsweise $\hat{\mathbf{T}} \in \hat{\Theta}$ Sensitivitätsanalysen effizient durchgeführt werden. Eine Verallgemeinerung der prädiktionsorientierten Kriterien stellt die **Q-Optimalität** dar, welche analog zur **I-Optimalität** die integrierte Prädiktionsvarianz minimiert, jedoch mittels Gewichtsfunktionen eine differenzierte Priorisierung spezifischer Regionen im Versuchsraum erlaubt, vgl. Goos und Jones [34]. Liegt der Fokus hingegen isoliert auf einer Teilmenge der Modellparameter - etwa zur Trennung von Haupteffekten und Blockeffekten oder von Termen erster und zweiter Ordnung - ermöglicht die **D_s-Optimalität** als Derivat der **D-Optimalität** eine gezielte Maximierung der Schätzgüte für genau dieses jeweilige Subset der Parametergruppe - siehe auch Werke von Goos und Jones [34], Myers et al. [36] und Atkinson et al. [65].

Tabelle 2.4: Übersicht, Zielsetzung und mathematische Definition verschiedener Optimalitätskriterien für Versuchspläne

Kriterium	Zielgröße, Beschreibung und Definition
A-Optimalität	<p>Minimierung der Spur der inversen Informationsmatrix [9, 28, 29].</p> <p><i>Ziel: Maximale Präzision der einzelnen Parameter im Durchschnitt.</i></p> $A_{\text{opt}} = \min \left(\text{spur} \left(\mathbf{M}^{-1} \right) \right) = \min \left(\sum_{j=1}^k \text{Var} (\hat{\beta}_j) \right) \quad (2.41)$
D-Optimalität	<p>Maximierung der Determinante der Informationsmatrix \mathbf{M}. Da das Volumen des Vertrauensellipsoids der Koeffizientenschätzwerte umgekehrt proportional zur Quadratwurzel der Determinante ($\sqrt{\det(\mathbf{M})}$) ist, minimiert dieses Kriterium das Unsicherheitsvolumen im Parameterraum [28, 29, 36, 65].</p> <p><i>Ziel: Minimierung des Volumens des gemeinsamen Vertrauensbereichs aller Modellparameter.</i></p> $D_{\text{opt}} = \max (\det(\mathbf{M})) = \min \left(\det \left(\mathbf{M}^{-1} \right) \right) \quad (2.42)$
G-Optimalität	<p>Minimierung der maximalen skalierten Prädiktionsvarianz $\nu(\mathbf{x}_0)$ im relevanten Parameterraum [29, 36, 54].</p> <p><i>Ziel: Qualität der Vorhersage an der ungünstigsten Stelle sichern.</i></p> $G_{\text{opt}} = \min \left(\max_{\mathbf{x}_0 \in \mathbb{R}^k} \nu(\mathbf{x}_0) \right) \quad (2.43)$

Fortsetzung auf der nächsten Seite...

Tabelle 2.4 (Fortsetzung): Übersicht der Optimalitätskriterien

Kriterium	Zielgröße, Beschreibung und Definition
I-Optimalität	<p>Minimierung der durchschnittlichen (<i>integrierten</i>) Prädiktionsvarianz über den Parameterraum - auch als U-Optimalität im ingenieurwissenschaftlich-statistischen Sinn bekannt: Optimierung bzgl. einer <i>uniformen</i> Prädiktionsvarianz [9, 65, 66].</p> <p><i>Ziel: Optimale Vorhersagegüte (SPV) im Mittel über den gesamten Raum - z.B. bei initialer Parametrisierung eines multivariaten LebensdauermodeLLs.</i></p> $I_{\text{opt}} = \min \left(\frac{N}{\int_{\mathbb{R}^k} d\mathbf{x}} \int_{\mathbb{R}^k} v(\mathbf{x}_0) d\mathbf{x} \right) \quad (2.44)$
V-Optimalität	<p>Minimierung der durchschnittlichen Prädiktionsvarianz über ein diskretes Set von N Punkten [29, 34, 58].</p> <p><i>Ziel: Optimale Vorhersagegüte (SPV) an spezifischen i Stellen - z.B. bei L-DoE-Versuchspunkten oder spezifischen Nennlasten.</i></p> $V_{\text{opt}} = \min \left(\frac{1}{N} \sum_{i=1}^N v(\mathbf{x}_i) \right) \quad (2.45)$

Zur vergleichbaren Bewertung unterschiedlicher Designs unabhängig von der Skalierung werden normierte **Effizienzen** herangezogen [29]. Aus Tabelle 2.4 lassen sich die **A-Effizienz** A_{eff} , **D-Effizienz** D_{eff} , **G-Effizienz** G_{eff} , **I-Effizienz** I_{eff} und **V-Effizienz** V_{eff} ableiten, welche jeweils die Güte $[0, 1]$ eines Versuchsplans quantifizieren - vergleiche Tabelle 2.5.

Tabelle 2.5: Übersicht, Zielsetzung und mathematische Definition verschiedener Effizienzkriterien für Versuchspläne

Kriterium	Zielgröße, Beschreibung und Definition
A-Effizienz	Maß für die durchschnittliche Präzision der Regressionskoeffizienten. [34, 36].
	$A_{\text{eff}} = 100 \cdot \frac{p}{\text{spur}(\mathbf{N} \cdot \mathbf{M}^{-1})} \quad (2.46)$
D-Effizienz	Maß für den Informationsgehalt der Informationsmatrix \mathbf{M} , definiert über die Determinante, umgekehrt proportional zum Volumen des Vertrauensellipsoids der Parameterschätzwerte und durch die Potenzierung mit $1/p$ pro Schätzparameter normiert. [34, 36].
	$D_{\text{eff}} = 100 \cdot \frac{\det(\mathbf{M})^{1/p}}{N} \quad (2.47)$
G-Effizienz	Maß für die Vorhersagegüte im ungünstigsten Fall innerhalb des Versuchsräums. Da für die maximale SPV die theoretische Untergrenze $v(\mathbf{x}_0)_{\max} \geq p$ gilt, beschreibt dieses Kriterium das Verhältnis zwischen idealer und realisierter maximaler Varianz [36].
	$G_{\text{eff}} = 100 \cdot \frac{p}{v(\mathbf{x}_0)_{\max}} \quad (2.48)$
I-Effizienz	Maß für die durchschnittliche Vorhersagegüte über den gesamten Versuchsräum [36].
	$I_{\text{eff}} = 100 \cdot \frac{p}{\frac{1}{\int_R d\mathbf{x}} \int_R v(\mathbf{x}_0) d\mathbf{x}} \quad (2.49)$

Fortsetzung auf der nächsten Seite...

Tabelle 2.5 (Fortsetzung): Übersicht der Effizienzkriterien

Kriterium	Zielgröße, Beschreibung und Definition
V-Effizienz	Maß für die durchschnittliche Vorhersagegüte über ein diskretes Set von N Punkten analog zu Gleichung 2.49 [36].

$$V_{\text{eff}} = 100 \cdot \frac{1}{N} \sum_{i=1}^N \frac{p}{v(\mathbf{x}_0)(\mathbf{x}_i)} \quad (2.50)$$

Forschungsschwerpunkte in der RSM

Während die klassische sequentielle Strategie (Screening gefolgt von Augmentierung zum CCD oder Anpassung von Parameterräumen mit optimalen Versuchsplänen) die methodischen Risiken minimiert, erfordert sie oft einen hohen administrativen und zeitlichen Aufwand durch mehrere Versuchphasen - insbesondere bei technischen Versuchsanlagen mit hohem Rüst-Aufwand. Als Antwort darauf entwickelten Jones und Nachtsheim [49] zunächst die Klasse der **Definitive Screening Designs (DSDs)**, um auch bei größeren Faktorenzahlen (typisch $k \geq 4$) Haupteffekte von nicht-linearen Effekten in einem Schritt zu trennen. Da DSDs jedoch eine starre Struktur aufweisen und oft nur eine geringe statistische Power zur Schätzung quadratischer Effekte bieten, etablierte sich in jüngster Forschung die umfassendere Klasse der **Orthogonal Minimally Aliased Response Surface (OMARS)-Designs** (*Orthogonal Minimally Aliased Response Surface Designs*), zu der Goos [67] zuletzt eine detaillierte Übersichtsarbeit vorlegten. Núñez Ares und Goos [68] erarbeiteten hierfür einen Katalog verschiedener OMARS-Designs, die im Gegensatz zu klassischen Plänen eine flexible Wahl der Versuchszahl n ermöglichen. Diese Designs zielen spezifisch auf den fließenden Übergang von Screening zu RSD in einem einzigen experimentellen Schritt ab (*One-Step Approach*). Dazu basieren sie auf der Variation aller quantitativen Faktoren auf drei Stufen ($-1, 0, +1$), um quadratische Zusammenhänge prinzipiell abzubilden. Die statistische Konstruktion garantiert dabei zwei fundamentale Orthogonalitätseigenschaften, die sie von klassischen Optimal-Designs abgrenzen: Erstens sind alle Haupteffekte orthogonal zueinander, was eine

unabhängige Schätzung der linearen Einflüsse sichert. Zweitens sind die Haupteffekte vollständig unkorreliert mit sämtlichen Effekten zweiter Ordnung (sowohl Zweifach-Wechselwirkungen als auch quadratische Terme). Das verbleibende Aliasing (Vermengung) beschränkt sich somit ausschließlich auf die Gruppe der Effekte zweiter Ordnung untereinander und wird algorithmisch minimiert (engl. *Minimally Aliased*). Dies erlaubt eine signifikante Flexibilisierung der Stichprobengröße im Vergleich zu starr definierten Plänen wie dem **CCD**, bei gleichzeitigem Erhalt der robusten Schätzbarkeit der primären Haupteffekte. Die Generierung eines solchen **OMARS**-DesignsST erfolgt dabei nicht über geometrische Konstruktionsvorschriften (wie beim **CCD**), sondern algorithmisch durch die Lösung eines ganzzahligen linearen Optimierungsproblems (engl. **Integer Linear Programming (ILP)**). Basierend auf der Kandidatenmenge Ω (üblicherweise entsprechend einem vollfaktoriellen 3^{k_f} Plan) wird ein Vektor binärer Entscheidungsvariablen $\gamma \in \{0,1\}^{|\Omega|}$ gesucht, der folgende mathematische Restriktionen simultan erfüllt:

- **Einhaltung des Stichprobenumfangs:** Die Summe der ausgewählten Versuchspunkte muss exakt der gewünschten Anzahl n entsprechen:

$$\sum_{i=1}^{|\Omega|} \gamma_i = n. \quad (2.51)$$

- **Orthogonalität der Haupteffekte:** Jedes Paar von Haupteffekt-Spalten j und l muss unkorreliert sein. Für die kodierten Faktorstufen $x_{ij} \in \{-1, 0, +1\}$ des i -ten Kandidaten gilt:

$$\sum_{i=1}^{|\Omega|} \gamma_i \cdot x_{ij} \cdot x_{il} = 0 \quad \forall 1 \leq j < l \leq k_f. \quad (2.52)$$

- **OMARS-Eigenschaft (Orthogonalität Haupteffekt zu Effekten zweiter Ordnung):** Dies ist das definierende Merkmal. Jeder Haupteffekt j muss orthogonal zu jedem Effekt zweiter Ordnung (Interaktion $x_j x_k$ oder quadratischer Term x_l^2) sein:

$$\sum_{i=1}^{|\Omega|} \gamma_i \cdot x_{ij} \cdot (x_{il} \cdot x_{ik}) = 0 \quad \forall j, l, k \in \{1, \dots, k_f\}. \quad (2.53)$$

Unter Einhaltung dieser harten Nebenbedingungen wird anschließend jene Design-Variante ausgewählt, welche die Aliasing-Struktur der Effekte zweiter Ordnung untereinander minimiert (z. B. durch Maximierung von D_{opt} , vgl. Gleichung 2.42).

Signifikanz und Trennschärfe

Zuletzt muss sichergestellt werden, dass der Versuchsplan über eine ausreichende **Trennschärfe** (engl. *power*) verfügt. Die Trennschärfe beschreibt als Metrik diejenige Wahrscheinlichkeit, mit der ein existierender Einfluss durch Faktoren vor dem Hintergrund des experimentellen Rauschens (vergleiche ***t*-Statistik, Signal-to-Noise Ratio, Signifikanztests** nach Kleppmann [28], Montgomery [29] und Goos und Jones [34])) korrekt als **signifikanter** Effekt E identifiziert wird [28, 33]. Insofern ist ein Effekt in praktischem Umfeld natürlich problemabhängig (Streuungscharakteristik der Lebensdauer, realisierbare Stichprobengröße, etc.) und meist von Bedeutung für $E \geq 2\sigma$, wobei Trennschärfewerte $\geq 80\%$ als befriedigend gelten [9, 29]. Insbesondere bei reduzierten Versuchsplänen und unter dem Einfluss stochastischer Lebensdauerstreuung ist die Power-Analyse jedoch essenziell, um das Risiko von falsch-negativen Schlussfolgerungen zu minimieren. Letztendlich kann die Trennschärfe somit als Garantiemetrik verstanden werden, die sicherstellt, dass ein geplanter Versuchsaufbau in der Lage ist, Effekte von praktischer Relevanz zu detektieren [36, 69] - was vornehmlich in Bezug auf Lebensdauertests aufgrund genannter wirtschaftlicher Randbedingungen von hoher Bedeutung ist.

Formalisiert wird dies im Rahmen **statistischer Hypothesentests** [28, 29]. Die Zielsetzung besteht darin, den Nachweis zu erbringen, ob ein bestimmter Faktor x_j einen signifikanten Einfluss auf die Systemantwort y ausübt. Dazu werden zwei konkurrierende Hypothesen formuliert: Die **Nullhypothese** H_0 unterstellt, dass kein Zusammenhang besteht, der korrespondierende Regressionskoeffizient β_j (vgl. Abschnitt 2.2.6) also den Wert Null annimmt:

$$H_0 : \beta_j = 0 \quad \text{gegen} \quad H_1 : \beta_j \neq 0. \quad (2.54)$$

Die **Alternativhypothese** H_1 postuliert hingegen einen signifikanten Effekt ($\beta_j \neq 0$). Die Entscheidung über die Annahme oder Ablehnung von H_0 basiert auf dem ***p*-Wert**. Dieser quantifiziert die Wahrscheinlichkeit, unter der Annahme der Gültigkeit von H_0 die beobachteten Daten (oder extremere Ergebnisse) zu erhalten.

Unterschreitet der p -Wert das a priori definierte **Signifikanzniveau** α (üblicherweise $\alpha = 0,05$), wird H_0 zugunsten der Alternativhypothese H_1 verworfen. Der Fehler, H_0 fälschlicherweise abzulehnen, obwohl kein Effekt vorliegt, wird als Fehler 1. Art bezeichnet; seine Wahrscheinlichkeit ist durch α begrenzt. Die **Trennschärfe** fokussiert hingegen auf den Fehler 2. Art (oft mit β_{error} notiert), welcher das Risiko beschreibt, einen tatsächlich vorhandenen Effekt nicht zu erkennen (falsch-negative Entscheidung). Die Power folglich ist definiert als das Komplement dieses Fehlers ($1 - \beta_{\text{error}}$) und entspricht somit der Wahrscheinlichkeit, H_0 korrekterweise abzulehnen, wenn H_1 wahr ist:

$$\text{power} = \Pr(H_0 \text{ ablehnen} \mid H_1 \text{ ist wahr}) = 1 - \beta_{\text{error}}. \quad (2.55)$$

In idealisierten, orthogonalen Standard-Versuchsplänen unter der Annahme normalverteilter Residuen $\epsilon \sim \mathcal{N}(0, \sigma^2)$ lässt sich die Trennschärfe für klassische **OLS**-Regressionsmodelle exakt analytisch bestimmen. Die power für den Signifikanztest eines einzelnen Regressionskoeffizienten β_j berechnet sich aus der Wahrscheinlichkeit, dass die Teststatistik den kritischen Wert der zentralen Prüfverteilung überschreitet. Unter der Gültigkeit der Alternativhypothese H_1 folgt die Teststatistik jedoch einer verschobenen, *nicht-zentralen* Verteilung (z.B. der nicht-zentralen t -Verteilung). Der Grad dieser Verschiebung wird durch den sogenannten **Nichtzentralitätsparameter** δ_{nc} quantifiziert [29, 70]. Für die Überprüfung eines individuellen Modellparameters ist dieser definiert als:

$$\delta_{\text{nc}} = \frac{|\beta_j|}{\sigma \sqrt{(\mathbf{X}'\mathbf{X})_{jj}^{-1}}} \quad (2.56)$$

wobei $|\beta_j|$ die zu detektierende wahre Effektstärke, σ die Standardabweichung des experimentellen Rauschens und $(\mathbf{X}'\mathbf{X})_{jj}^{-1}$ das j -te Diagonalelement der inversen Informationsmatrix \mathbf{M}^{-1} (vgl. Abschnitt 2.2.5) darstellt [36, 69]. Anhand von Gleichung 2.56 wird der mathematische Zusammenhang der Effektivität deutlich: Die Trennschärfe, welche monoton mit δ_{nc} wächst, steigt bei größeren Effekten und sinkt bei stärkerem Rauschen.

Die entscheidende Rolle des Versuchsplans manifestiert sich im Term $(\mathbf{X}'\mathbf{X})_{jj}^{-1}$. In einem vollständig orthogonalen Versuchsplan ist die Informationsmatrix $\mathbf{M} = \mathbf{X}'\mathbf{X}$

eine reine Diagonalmatrix. Die Schätzer der Koeffizienten sind somit linear unabhängig und das Diagonalelement der Inversen minimiert sich. Sobald jedoch geometrische Verzerrungen auftreten und die Orthogonalität verloren geht, entstehen signifikante Kovarianzen (Werte abseits der Hauptdiagonalen in \mathbf{M}). Bei der Invertierung der Matrix führt dies zu einer Inflation der Varianzen, wodurch der Wert $(\mathbf{X}'\mathbf{X})_{jj}^{-1}$ anwächst. Dies vergrößert den Nenner in Gleichung 2.56, reduziert den Nichtzentralitätsparameter δ_{nc} und mindert somit unmittelbar die analytische Trennschärfe.

Solche analytischen Berechnungen, die auf der exakten Bestimmung von δ_{nc} beruhen, verkomplizieren sich erheblich, sobald asymmetrische Design-Modifikationen und Mehrfach-Korrelationen greifen. Im Kontext der Lebensdauerdatenanalyse, bei der zudem nicht-normalverteilte und censierte Daten vorliegen, lassen sich diese Wahrscheinlichkeiten folglich ohnehin nicht mehr über klassische **t-Tests** ableiten. Aus der Arbeit von Nelder und Wedderburn [71] kommen hier **Verallgemeinerte Lineare Modelle**, engl. **Generalized Linear Models (GLMs)** zum Einsatz, bei denen die Signifikanz der Koeffizienten β_j vorzugsweise über den **Likelihood-Ratio Test (LR-Test)** (oder approximativ über **Wald-Tests** - vergleiche hierzu Arbeiten von Meeker et al. [4], Rigdon et al. [9], Montgomery [29] und McCulloch und Searle [72]) ermittelt wird. Da für diese komplexen Verteilungsmodelle - insbesondere unter dem Einfluss von Orthogonalitätsabweichungen und bei kleinen Stichprobenumfängen n - keine geschlossenen analytischen Lösungen zur Berechnung der Trennschärfe existieren, erfolgt die Untersuchung in der modernen Praxis numerisch mittels **Monte-Carlo-Simulation (MCS)** [40, 73, 74]. Hierbei wird der geplante Versuchsplan virtuell vielfach durchlaufen (z.B. $n_{MC} = 10^4$ Simulationsläufe). Für jeden Lauf werden basierend auf einem angenommenen Modell Ausfallzeiten generiert, verrauscht bzw. mit Streuung überlagert und/oder censiert. Anschließend erfolgt die Modellbildung und Hypothesenprüfung. Der Anteil der Simulationsläufe, in denen der definierte Effekt korrekt als signifikant (p -Wert $< \alpha$) erkannt wird, entspricht der geschätzten **power** des Designs.

2.2.6 Statistische Modellbildung

Die statistische Modellbildung transformiert die durch den Versuchsplan generierte Datenbasis in einen funktionalen, empirisch basierten Zusammenhang zwischen den Einflussfaktoren x_i und der Lebensdauerantwort. Aufgrund der in Abschnitt 2.1

dargelegten Eigenschaften von Lebensdauerdaten (Nicht-Normalverteilung, Zensierung) sind klassische Regressionsverfahren (**OLS**) hier nicht zulässig. Stattdessen kommt i.d.R. das Framework der **GLM** zur Anwendung, welches die lineare Prädiktion mit der zugrundeliegenden Verteilung verknüpft [4, 6, 9]. Hinsichtlich der Modellauswahl können in klassischen Ansätzen der Versuchsplanung dabei oft auf Potenztransformationen der Antwortvariablen in Form $y^{(\lambda_{BC})}$ oder $\ln(y)$ (für $\lambda_{BC} = 0$ mit geometrischem Mittelwert der Beobachtung $\dot{y} = \exp\left[\frac{1}{n} \sum_{i=1}^n \ln(y_i)\right]$) zurückgegriffen werden - wie die **Box-Cox-Transformation** nach Montgomery et al. [13] und Box und Cox [75]. So wird Varianzhomogenität und Normalverteilung approximiert zu:

$$y_i^{(\lambda_{BC})} = \begin{cases} \frac{y_i^{\lambda_{BC}} - 1}{\lambda_{BC} y_i^{\lambda_{BC} - 1}}, & \text{für } \lambda_{BC} \neq 0, \\ \dot{y} \ln(y_i), & \text{für } \lambda_{BC} = 0. \end{cases} \quad (2.57)$$

Im Kontext der expliziten Lebensdaueranalyse ist dieser generische Ansatz jedoch weniger gebräuchlich, da die physikalisch motivierte Verteilungsannahme (z.B. Weibull) bereits eine inhärente Transformation der Lebensdauerzeit impliziert - vgl. Abschnitt 2.1.3, Gleichung 2.8 sowie Yang [6] und Wu und Hamada [30]. Wird also für die Lebensdauerverteilung die Weibull-Verteilung (Abschnitt 2.1.3) angenommen, so ergibt sich für $j = 0, \dots, k$ betrachtete Parameter aus beispielsweise Gleichung 2.27 und n Beobachtungen analog zu Gleichung 2.8 der Erwartungswert der charakteristischen Lebensdauer als **Link-Funktion** der Einflussfaktoren x_i und des unbekannten Koeffizientenvektors $\hat{\beta}$ [6, 30, 76]:

$$E_{\theta_x}(\ln(T)) = y = \beta_0 + \sum_{j=1}^k \beta_j x_j + \dots + \epsilon. \quad (2.58)$$

In der Lebensdauer-Datenanalyse werden insbesondere zwei weitere Modellierungsansätze verwendet: neben den **GLM**-Ansätzen insbesondere der **Proportional Hazard (PH)-Modell**-Ansatz. Jedoch werden **PH**-Modelle vornehmlich in der Biostatistik verwendet und spielen hier ihre Stärke bei der Abbildung zeitabhängiger Kovariate aus [77, 78]:

$$f(t, \mathbf{X}) = \lambda(t)(t, \mathbf{X}) \cdot R(t, \mathbf{X}) = b \cdot t^{b-1} \cdot e^{[y - \frac{b}{b} e^y]}. \quad (2.59)$$

Weibull-GLM

Im ingenieurwissenschaftlichen Kontext - maßgeblich gestützt durch Ausführungen in Meeker et al. [4], Yang [6], Wu und Hamada [30], Myers et al. [58], Kremer et al. [78], Dobson und Barnett [79] und Stufken und Yang [80] - finden hingegen **Log-Location-Scale-Modelle**, also **GLMs**, überwiegend Anwendung. Obwohl für die Weibull-Verteilung eine mathematische Äquivalenz zwischen beiden Ansätzen besteht, sprechen drei fundamentale Gründe für deren Verwendung. **Physikalischer Bezug:** Gängige physikalische Beschleunigungsmodelle (Arrhenius, Inverses Potenzgesetz) beschreiben die Änderung der *Zeit* bis zum Ausfall ($\ln(\hat{T})$) und nicht der Ausfallrate. Die Log-Linearisierung dieser Gesetze führt direkt auf die Regressionsgleichung der Location-Scale-Modelle. **Universalität der Auswertung:** Das Modell bietet einen einheitlichen mathematischen Rahmen, der nicht nur für die Weibull-Verteilung, sondern identisch auch für die Lognormal-Verteilung und Normalverteilung gültig ist. Das **PH**-Modell hingegen ist für Lognormal-Verteilungen mathematisch nicht geschlossen anwendbar. **Fokus auf Lebensdauer-Quantile:** Im Engineering und **DfR** liegt das Interesse primär auf der Prädiktion von Ausfallzeiten für geringe Ausfallanteile (z.,B. t_{10}). Diese Größen sind direkte Ergebnisse der Location-Scale-Gleichung, während sie im **PH**-Modell nur indirekt über die Invertierung der Hazard-Funktion zugänglich sind. Folglich wird eine zufallsverteilte Lebensdauer τ für ein Weibull-**GLM** mit der Dichtefunktion

$$f(t, \mathbf{X}) = b \cdot t^{b-1} e^{-b(\beta_0 + \sum_{j=1}^k \beta_j x_j)} \cdot e^{-t^b e^{-b(\beta_0 + \sum_{j=1}^k \beta_j x_j)}}, \quad t > 0. \quad (2.60)$$

beschrieben und analog zu Gleichungen 2.14–2.17 in die **Generalized Log-Linear (GLL)-Funktion** überführt [6, 54, 81]:

$$\begin{aligned} \Lambda &= \ln(\mathcal{L}(t_i; b, \beta_0, \dots, \beta_k)) \\ &\propto \sum_{i=1}^n \left[\ln(b) - b \left(\beta_0 + \sum_{j=1}^k \beta_j x_{ij} \right) + (b-1) \ln(t_i) - t_i^b e^{-b(\beta_0 + \sum_{j=1}^k \beta_j x_{ij})} \right]. \end{aligned} \quad (2.61)$$

Werden außerdem (rechts-) zensierte Daten berücksichtigt, so ergibt sich analog zu Gleichung 2.17 über Gleichung 2.61 die erweiterte GLL-Funktion [4, 6]:

$$\Lambda_z \propto \Lambda - (1 - \delta_i) \cdot t_i^b e^{-b(\beta_0 + \sum_{j=1}^k \beta_j x_{ij})}. \quad (2.62)$$

Sollen über die rein empirische Datenauswertung hinaus bekannte physikalische Gesetzmäßigkeiten im Modell berücksichtigt werden, ist die lineare Struktur des Prädiktors entsprechend anzupassen. Dies erfolgt durch eine Transformation der physikalischen Beanspruchungsgrößen x in den Modellraum. Anstelle der direkten Verwendung des Einflussfaktors in der GLL-Beziehung (Gleichung 2.60) wird eine transformierte Größe x_V eingesetzt, welche die physikalische Natur des Schädigungsmechanismus abbildet. Gängige Transformationen umfassen hierbei die reziproke Anpassung für thermische Lasten oder logarithmische Ansätze für elektrische Potenzmodelle, vgl. Tabelle 2.6. Eine Übersicht verschiedener Modellierungsansätze

Tabelle 2.6: Übersicht gängiger Transformationen der Belastungsgrößen zur Abbildung physikalischer Lebensdauermodelle

Art	Transformation	Lebensdauermodell	Anwendungsfall
Keine	$x_V = x$	Exponential / Wöhler	mechanisch
Reziprok	$x_V = 1/x$	Arrhenius	thermisch
Logarithmisch	$x_V = e^x$	Potenzmodell	elektrisch

(z.B. **Arrhenius-Modell**, **Inverses Potenzgesetz**, **Eyring**, **Coffin-Manson-Modell**, etc.) und deren Einbindung in die GLM-Struktur findet sich in den Arbeiten von Meeker et al. [4], Rigdon et al. [9] und Modarres et al. [77]. Die Schätzung der unbekannten Parameter $\hat{\beta}$ und \hat{b} erfolgt anschließend mittels MLE durch Maximierung der Log-Likelihood-Funktion (Gleichung 2.62) mittels beispielsweise **Newton-Raphson-Verfahren** bzw. **Pattern-Search-Algorithmen** oder **Gradientenverfahren**, [6, 17]. Geeignete numerische Optimierungsverfahren wurden in Abschnitt 2.1.4 erwähnt und unter anderem durch Vorarbeiten von Kremer und Bertsche [18] evaluiert.

Signifikanzanalyse

Liegen die Parameterschätzungen $\hat{\theta}_i$ vor, erfolgt die Bewertung ihrer statistischen Relevanz mittels **LR-Test**. Dieses Verfahren stellt das methodische Äquivalent zur **ANOVA** bei normalverteilten Daten dar und prüft, ob die Aufnahme eines Parameters die Modellanpassung signifikant verbessert. Der Test basiert auf dem Vergleich der maximalen Likelihoods des vollständigen Modells $\mathcal{L}(\hat{\theta}_i)$ gegenüber einem reduzierten Modell $\mathcal{L}(\hat{\theta}_{-i})$, aus dem der zu untersuchende Parameter eliminiert wurde [22, 40]. Die Teststatistik LR quantifiziert den Informationsverlust durch die Modellreduktion und ist definiert als **Devianz-Funktion** [4, 82]:

$$LR = -2 \ln \left(\frac{\mathcal{L}(\hat{\theta}_{-i})}{\mathcal{L}(\hat{\theta}_i)} \right) = 2 \cdot [\ln(\mathcal{L}(\hat{\theta}_i)) - \ln(\mathcal{L}(\hat{\theta}_{-i}))]. \quad (2.63)$$

Unter der Nullhypothese H_0 , dass der ausgeschlossene Parameter keinen Einfluss hat, folgt LR asymptotisch einer χ^2 -Verteilung mit d_f Freiheitsgraden (Anzahl der ausgeschlossenen Parameter, hier meist $d_f = 1$). Die Signifikanzentscheidung erfolgt über den **p-Wert**, der die Wahrscheinlichkeit angibt, einen Testwert $\geq LR$ zu beobachten:

$$p\text{-Wert} = \Pr(\chi^2(d_f) \geq LR). \quad (2.64)$$

Ist dieser **p-Wert** kleiner als das definierte Signifikanzniveau $p\text{-Wert} < \alpha$, wird der Parameter als signifikant in das Modell aufgenommen.

Erwartungswerte und Konfidenzintervalle

Nach der Ermittlung der optimalen Modellparameter $\hat{\theta}$ ist die Quantifizierung der mit diesen Schätzungen verbundenen Unsicherheit essenziell. Analytisch erfolgt dies über die Betrachtung der Krümmung der Log-Likelihood-Funktion Λ im Maximum. Je stärker die Krümmung, desto präziser ist die Schätzung (geringere Varianz). Mathematisch erfolgt dies durch die drei Schritte: Wahl der statistischen Zielgröße (R, F), Berechnung des Standard-Fehlers (via F analog zu Gleichung 2.18) und Vertrauensbereich-Schätzung / Wald-Statistik (Gleichung 2.20 bzw. 2.21). So wird dies analog zu Gleichungen 2.18-2.19 durch die Hesse-Matrix H_f beschrieben, welche die zweiten partiellen Ableitungen der Log-Likelihood-Funktion nach den

Parametern θ_i (also β_j und b) enthält. Der negative Erwartungswert dieser Matrix führt damit zur Erwartung der Fisher-Informationsmatrix \mathbf{F} [25, 83]:

$$\mathbf{F}_O := -\mathbf{H}_f(\hat{\theta}) = \begin{bmatrix} -\frac{\partial^2 \mathcal{L}}{\partial \theta_1^2} & -\frac{\partial^2 \mathcal{L}}{\partial \theta_1 \partial \theta_2} & \cdots & -\frac{\partial^2 \mathcal{L}}{\partial \theta_1 \partial \theta_k} \\ -\frac{\partial^2 \mathcal{L}}{\partial \theta_2 \partial \theta_1} & -\frac{\partial^2 \mathcal{L}}{\partial \theta_2^2} & \cdots & -\frac{\partial^2 \mathcal{L}}{\partial \theta_2 \partial \theta_k} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{\partial^2 \mathcal{L}}{\partial \theta_k \partial \theta_1} & -\frac{\partial^2 \mathcal{L}}{\partial \theta_k \partial \theta_2} & \cdots & -\frac{\partial^2 \mathcal{L}}{\partial \theta_k^2} \end{bmatrix}. \quad (2.65)$$

Die asymptotische Varianz-Kovarianz-Matrix $\hat{\mathbf{V}}$ der Parameterschätzer wird anschließend durch die Inversion der Fisher-Informationsmatrix für $\hat{\theta} = \theta$ approximiert [6, 83]:

$$\hat{\mathbf{V}} \approx \mathbf{F}_O^{-1} = \begin{bmatrix} \hat{\text{Var}}(\hat{\theta}_1) & \hat{\text{Cov}}(\hat{\theta}_1, \hat{\theta}_2) & \cdots & \hat{\text{Cov}}(\hat{\theta}_1, \hat{\theta}_k) \\ \hat{\text{Cov}}(\hat{\theta}_2, \hat{\theta}_1) & \hat{\text{Var}}(\hat{\theta}_2) & \cdots & \hat{\text{Cov}}(\hat{\theta}_2, \hat{\theta}_k) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\text{Cov}}(\hat{\theta}_k, \hat{\theta}_1) & \hat{\text{Cov}}(\hat{\theta}_k, \hat{\theta}_2) & \cdots & \hat{\text{Var}}(\hat{\theta}_k) \end{bmatrix}. \quad (2.66)$$

Die Hauptdiagonalelemente dieser Matrix liefern die geschätzten Varianzen $\hat{\text{Var}}(\hat{\theta}_j)$ der einzelnen Modellparameter. Unter der Annahme der asymptotischen Normalverteilung der MLE beschreibt Yang [6] neben weiteren Alternativen aus Meeker et al. [4] und Kremer und Bertsche [22] die symmetrischen $(1 - \alpha)$ -Konfidenzintervalle (Wald-Intervalle) wie nachfolgend - vgl. Gleichung 2.21:

$$[\theta_{j,u}, \theta_{j,o}] = \hat{\theta}_j \pm z_{1-\alpha/2} \cdot \sqrt{\hat{\text{Var}}(\hat{\theta}_j)} \quad \text{bzw. } \dots = \hat{\theta}_j \exp \left(\pm z_{1-\alpha/2} \cdot \frac{\sqrt{\hat{\mathbf{V}}_{jj}}}{\hat{\theta}_j} \right). \quad (2.67)$$

Vertrauensbereiche für Funktionswerte und Weibull-GLM

Oft ist nicht nur die Unsicherheit der Modellparameter selbst, sondern die der daraus abgeleiteten Funktionen - wie der Zuverlässigkeit $R(t)$ oder der Ausfallwahrscheinlichkeit $F(t)$ - von Interesse. Zur analytischen Ermittlung dieser sogenannten Fisher-Vertrauensbereiche führt Yang [6] zunächst eine Hilfsgröße w ein - vergleiche auch z_e nach Meeker et al. [4]. Diese normiert die logarithmische Abweichung der betrachteten Zeit t von der geschätzten charakteristischen Lebensdauer \hat{T} mit dem

Formparameter \hat{b} (vergleiche auch Bain und Engelhardt [24] für weitere Abschätzungen):

$$\hat{w} = \hat{b} \cdot \ln\left(\frac{t}{\hat{T}}\right). \quad (2.68)$$

Die Varianz dieser Größe lässt sich mittels Taylor-Reihenentwicklung (Delta-Methode, vgl. Abschnitt 2.1.4) approximativ aus den Varianzen und Kovarianzen der Parameterschätzer und analog zu Gleichungen 2.22 und 2.23 bestimmen:

$$\hat{\text{Var}}(\hat{w}) = \left(\frac{\hat{b}}{\hat{T}}\right)^2 \hat{\text{Var}}(\hat{T}) + \left(\frac{\hat{w}}{\hat{b}}\right)^2 \hat{\text{Var}}(\hat{b}) - \frac{2\hat{w}}{\hat{T}} \hat{\text{Cov}}(\hat{T}, \hat{b}). \quad (2.69)$$

Basierend auf der asymptotischen Normalverteilung ergeben sich die Vertrauensgrenzen $[\hat{w}_u, \hat{w}_o]$ gemäß:

$$[\hat{w}_u, \hat{w}_o] = \hat{w} \pm z_{1-\alpha/2} \cdot \sqrt{\hat{\text{Var}}(\hat{w})}. \quad (2.70)$$

Die Rücktransformation dieser Grenzen in den Wertebereich der Ausfallwahrscheinlichkeit erfolgt anschließend über die Verteilungsfunktion der kleinsten Extremwerte $G = 1 - \exp[-\exp(\hat{w})]$, woraus sich direkt die Konfidenzintervalle der Zuverlässigkeit ableiten lassen. Im Fall von GLM ist zu berücksichtigen, dass die geschätzte charakteristische Lebensdauer \hat{T} keine unabhängige Konstante ist, sondern funktional entsprechend Gleichung 2.58 von den Kovariaten x und dem Parametervektor $\hat{\beta}$ abhängt. Die für Gleichung 2.69 benötigte Varianz $\hat{\text{Var}}(\hat{T})$ muss folglich unter Berücksichtigung der gesamten Kovarianzmatrix der Regressionskoeffizienten aus Gleichung 2.66 berechnet werden. Dies erfolgt durch die allgemeine Fehlerfortpflanzung nach Gauß:

$$\hat{\text{Var}}(\hat{T}) \approx \sum_{i=1}^k \left(\frac{\partial \hat{T}}{\partial \hat{\beta}_i} \right)^2 \hat{\text{Var}}(\hat{\beta}_i) + \sum_{i=1}^k \sum_{\substack{j=1 \\ j \neq i}}^k \left(\frac{\partial \hat{T}}{\partial \hat{\beta}_i} \right) \left(\frac{\partial \hat{T}}{\partial \hat{\beta}_j} \right) \hat{\text{Cov}}(\hat{\beta}_i, \hat{\beta}_j). \quad (2.71)$$

Modellgüte und Residuenanalyse

Die Validität eines parametrischen Lebensdauermodells (hier Weibull-GLM) beruht auf drei fundamentalen Annahmen, die es nach der Parameterschätzung zu verifizieren gilt: Erstens folgen die Lebensdauern der angenommenen Verteilung. Zweitens

sind die Lebensdauern statistisch unabhängig. Drittens besteht ein log-linearer Zusammenhang zwischen dem Skalenparameter und den Einflussgrößen, wobei der Formparameter in einfachen Modellen als konstant über den gesamten Versuchsräum angenommen wird. Die Beurteilung dieser Modellgüte erfordert bei Lebensdauerdaten, bedingt durch Nicht-Normalverteilung und Zensierung, eine Anpassung klassischer Konzepte zur Residuenanalyse ($y - \hat{y}$) [28, 33]. Stattdessen werden transformierte Residuen betrachtet, die den Bezug zur Verteilung der kleinsten Extremwerte **SEV** herstellen, welcher der logarithmierten Weibull-Verteilung zugrunde liegt. Analog zur ***z***-Transformation bei der Normalverteilung lassen sich für jede Beobachtung t_i standardisierte Residuen r_i berechnen. Dabei ersetzt der logarithmierte Schätzwert der charakteristischen Lebensdauer $\ln(\hat{T}_i)$ (vgl. Gleichung 2.58) den Lageparameter μ und der Kehrwert des Formparameters $1/\hat{b}$ die Skalierung σ . Dies entspricht exakt der in diesem Abschnitt eingeführten Hilfsgröße w :

$$r_i = \frac{\ln(t_i) - \ln(\hat{T}_i)}{1/\hat{b}} = \hat{b} \cdot \ln\left(\frac{t_i}{\hat{T}_i}\right). \quad (2.72)$$

Ist das Modell korrekt spezifiziert, müssen diese Residuen (unter Berücksichtigung der Zensierung) einer Verteilung der kleinsten Extremwerte **SEV** ($\mu = 0, \sigma = 1$) folgen. Eine allgemeinere Form der Validierung bieten die **Cox-Snell-Residuen** r_{CS} [4, 9]. Sie stellen eine Transformation der standardisierten Residuen dar und entsprechen dem Wert der kumulativen Hazard-Funktion an der Stelle der Beobachtung. Für das Weibull-Modell gilt der Zusammenhang:

$$r_{CSi} = -\ln(\hat{R}(t_i)) = \exp(r_i) = \left(\frac{t_i}{\hat{T}_i}\right)^{\hat{b}}. \quad (2.73)$$

Folgt das angepasste Modell den Daten, so verhalten sich die Residuen r_{CS} wie eine Stichprobe aus einer Standard-Exponentialverteilung mit Mittelwert 1. Diese Eigenschaft erlaubt eine grafische Validierung mittels Probability-Plots: Liegen die geplotteten Residuen auf einer Geraden im Exponential-Wahrscheinlichkeitsnetz, kann von einer adäquaten Modellpassung ausgegangen werden. Zur Identifikation von Ausreißern oder einflussreichen Beobachtungen eignen sich hingegen standardisierte Devianz-Residuen r_{Dev} . Im Gegensatz zu den stets positiven Cox-Snell-Residuen

sind diese symmetrischer um Null verteilt und folgen asymptotisch einer Standardnormalverteilung:

$$r_{\text{Dev}i} = \text{sgn}(r_{\text{CS}i} - 1) \cdot \sqrt{2 [r_{\text{CS}i} - \delta_i - \delta_i \ln(r_{\text{CS}i})]}. \quad (2.74)$$

Neben der Verteilungsprüfung dienen Streudiagramme dieser Residuen zur Aufdeckung struktureller Modelldefizite. Dabei werden r_{CS} oder r_{Dev} gegen die prädizierten Werte oder gegen einzelne Einflussfaktoren aufgetragen und gegenüber Trendfreiheit und Varianzhomogenität bewertet. Besteht initiale Unsicherheit über die zugrundeliegende Physik, ermöglicht zudem der Vergleich den Wahrscheinlichkeitsnetzen der Residuen eine empirische Selektion: Dasjenige Modell, dessen Punktegraphen eine geringere Krümmung aufweisen und besser der theoretischen Geraden folgen, ist statistisch zu bevorzugen.

3 Effiziente Testplanung für die multivariate Lebensdauererprobung

3.1 Bewertung des Standes der Forschung und Technik

Der in Kapitel 2 dargelegte Stand der Forschung und Technik zur Versuchsplanungs-methodik verdeutlicht, dass sowohl die Zuverlässigkeitstechnik als auch DoE für sich genommen etablierte Disziplinen mit jeweils umfangreichen Methodenreper-toire darstellen. Die Schnittmenge beider Felder, die effiziente Planung multivaria-ter Lebensdauerversuche unter Berücksichtigung nicht-normalverteilter Daten und Zensierung (Weibull-GLM), weist jedoch spezifische Lücken auf, die den Bedarf an weiterführender Forschung begründen.

3.1.1 Eignung klassischer Response Surface Designs

Klassische Versuchspläne der RSM, insbesondere das CCD, haben sich als robuster Standard für die Modellierung nicht-linearer Zusammenhänge bewährt [9, 28]. Ihre Stärken liegen in der sequenziellen *Augmentierbarkeit*, der *Rotierbarkeit* und der Fähigkeit zur unabhängigen Schätzung des reinen Fehlers (vgl. Abschnitt 2.2.3). Zudem lässt sich aus dem am IMA bislang generierten Wissensstand nach Kremer [40] ableiten, welche methodischen Vorgehen im Rahmen des L-DoE sich bestenfalls dazu eigenen, adäquate Schätzungen für Modellparameter [18] und Vertrauensbe-reiche [22] zu erhalten. Für die spezifischen Anforderungen der Lebensdauerer-probung, die häufig eine Extrapolation der Ergebnisse aus einem hochbelasteten ALT-Testraum in einen niedriger belasteten Anwendungsraum erfordert, ist die starre Symmetrie des klassischen CCD jedoch nicht zwangsläufig optimal. Ein geeig-neter, *effizienter* Umgang mit dieser Herausforderung ist bislang nicht etabliert. Es

existieren zwar bereits vereinzelte Ansätze zur Modifikation von **CCDs**, wie die Arbeiten von Box und Wilson [51], Ahn [84] und Donev [85] und Ardakani et al. [86] zeigen. Diese fokussieren sich jedoch zumeist auf generische Optimalitätskriterien oder spezifische geometrische Restriktionen und adressieren nicht explizit die Herausforderungen der probabilistischen Lebensdauerprognose unter Extrapolation im Kontext der statistischen **Trennschärfe** (*power*). Obgleich für die Anwendung im Maschinen- und Fahrzeugbau von entscheidender Bedeutung, existiert bislang kein etabliertes Vorgehen, um die Versuchspunkt-Konstellation gezielt in Prognoserichtung zu optimieren und so die Extrapolationsfähigkeit zu maximieren, ohne die Balance zwischen Modellgüte und Versuchsaufwand zu gefährden.

3.1.2 Limitierungen und Effizienzbegriffe moderner Versuchsplanung

Der Einsatz computergenerierter optimaler Versuchspläne erscheint zunächst als logischer Schritt, um maßgeschneiderte Designs für Lebensdauererprobungen zu entwerfen. Im Kontext multivariater Lebensdauertests (**ALT**) erfährt der Begriff der **Effizienz** jedoch eine signifikante Komplexitätssteigerung gegenüber der klassischen linearen Versuchsplanung, die sich in den drei Dimensionen *Modellabhängigkeit*, *zeitliche Dynamik* und *strukturelle Flexibilität* aufgliedert. Formal definiert sich die Effizienz eines Versuchsplans einerseits analytisch über die Informationsmatrix $\mathbf{M}(\mathbf{X}, \boldsymbol{\theta})$ [74, 80], andererseits generisch über das Verhältnis von Leistungsfähigkeit zur Informationsbeschaffung - ein Anspruch, der nach Thommen et al. [87] in einem fundamentalen Spannungsfeld zum Grundsatz der Vollständigkeit steht. Damit determiniert die Effizienz maßgeblich die **Effektivität** und Wirtschaftlichkeit der Lebensdauererprobung und in letzter Instanz die Verlässlichkeit der Modellierung (vgl. Abschnitt 2.2.5: *power*) sowie die Prognosegüte der Zuverlässigkeit $R(t)$.

Modellabhängigkeit und Lokale Optimalität

Erstens stößt die zugrundeliegende Modellierung mittels Weibull-**GLMs** auf ein fundamentales Problem, das von Khuri [74] und Stufken und Yang [80] hinlänglich diskutiert wird. Im Gegensatz zu klassischen linearen Modellen (unter Normalverteilungsannahme) hängt die Fisher-Informationsmatrix \mathbf{F} bei **GLMs** nicht allein von

den Einstellungen der Faktoren \mathbf{X} ab, sondern vollständig von den wahren, aber unbekannten Modellparametern $\boldsymbol{\theta}$ (insbesondere dem Formparameter \hat{b} und via der Link-Funktion dem Lageparameter \hat{T}):

$$\mathbf{M}(\mathbf{X}, \boldsymbol{\theta}) = \mathbb{E} \left[-\frac{\partial^2 \mathcal{L}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2} \right]. \quad (3.1)$$

Daraus resultiert das Paradoxon, dass zur Konstruktion eines optimalen Plans bereits Kenntnisse über die zu ermittelnden Parameter vorliegen müssten. Zwar existieren auf Basis von Exponential- oder Poissonverteilung diverse Algorithmen zur numerischen Bestimmung optimaler Versuchspläne für **GLMs** [74], und auch für Experimente mit binärer Antwortstruktur (logistische oder Probit-Modelle) wurden Ansätze entwickelt [88]. Diese Verfahren setzen jedoch stets voraus, dass die wahren Parameterwerte $\boldsymbol{\theta}$ bekannt sind oder verlässlich geschätzt werden können – beispielsweise mittels initialem *best guess*. Folglich sind solche Designs, wie Stufken und Yang [80] ausführen, allenfalls nur **lokal optimal**. Da die Parameter im Vorfeld einer Lebensdaueruntersuchung unbekannt sind, existieren keine universell optimalen **GLM**-Designs für die Weibull-Lebensdauermodellierung. Ein Design, das für einen angenommenen Parametersatz optimal ist, kann sich bei Abweichung der realen Werte als ineffizient erweisen („The designs may be poor if the choice of values is far from the true parameter values“, Stufken und Yang [80]).

Zeitliche und Strukturelle Effizienz

Zweitens muss die **zeitliche Effizienz** berücksichtigt werden. Da die Information in Lebensdauertests stochastisch über die Zeit generiert wird und womöglich durch Zensierung limitiert ist, wird die Varianz der Schätzung maßgeblich durch die Anzahl der tatsächlich ausgefallenen Einheiten getrieben [16]. Ein Versuchsplan gilt folglich nur dann als effizient, wenn er die erwartete Testdauer bei gegebener Präzision minimiert. Drittens bietet die Klasse der **OMARS**-Designs eine Antwort auf die Forderung nach **struktureller Effizienz** und Flexibilität. Im direkten Vergleich zum etablierten **CCD**, das durch seine geometrische Konstruktion eine starre Anzahl an Versuchen fordert, ermöglichen **OMARS**-Designs zwar eine Entkopplung von Faktoren- und Versuchsanzahl (z.B. $N = 40$ statt $N \geq 45$ für $k = 6$). Sie resultieren jedoch in einer finalen Versuchspunktverteilung, die sich im Kontext von

Aufwänden in EoL-Testing rational nur schwer sequenziell erweitern oder anpassen lässt. Strategisch verfolgen sie einen *One-Step*-Ansatz [67], der Screening und Optimierung integriert, um Rüstzeiten und zeitliche Drifts zu vermeiden. Der Preis für diese Effizienz ist ein kontrolliertes Maß an *Aliasing* zwischen Effekten zweiter Ordnung. Damit qualifizieren sich diese Designs insbesondere für ressourcenbeschränkte Untersuchungen, wenngleich das CCD seine Stärke in der Robustheit und sequenziellen Kontrollierbarkeit behält.

3.1.3 Schlussfolgerung für die Testplanung

Aus der Diskrepanz zwischen theoretischem Anspruch (optimale GLM-Designs) und praktischer Anwendbarkeit (unbekannte Parameter) leitet sich die Notwendigkeit einer pragmatischen und zugleich statistisch fundierten Vorgehensweise ab. Anstatt sich auf die Suche nach einem theoretisch *optimalen GLM-Design* zu versteifen, das aufgrund der Parameterabhängigkeit in der Praxis kaum robust umsetzbar ist, erscheint die Adaption bewährter Standard-Designs (CCD) vielversprechender. Dabei dürfen die etablierten Optimalitätskriterien jedoch nicht als alleiniges Entwurfsziel fungieren, sondern müssen durch praxisrelevante Benchmark-Metriken ergänzt werden:

- **Modellierungsgüte (A-, D-Kriterien):** Diese Metriken sichern ab, dass die Parameterschätzung (Bestimmung von Θ) auch im modifizierten Design statistisch solide und präzise bleibt.
- **Prognosefähigkeit (G-, I-, V-Kriterien):** Diese Metriken gewinnen im Kontext der Lebensdauererprobung an primärer Bedeutung, da das Ziel die präzise Vorhersage der Zuverlässigkeit $R(t)$ unter Einsatzbedingungen (Extrapolation) ist.
- **Testdauer und Ressourceneffizienz:** Diese Metriken berücksichtigen die praktischen Rahmenbedingungen von Lebensdauertests, indem sie die erwartete Testdauer und den Ressourcenverbrauch minimieren, ohne die statistische Aussagekraft unzulässig zu beschneiden.
- **Trennschärfe (Power):** Sie fungiert als die entscheidende „Erfolgsgarantie“ des Tests. Insbesondere bei kostenintensiven EoL-Versuchen ist es ökonomisch

riskant, ein Design zu wählen, das zwar theoretisch optimal schätzt, aber praktisch relevante Effekte mit hoher Wahrscheinlichkeit übersieht. Die Power muss daher als primäres Entscheidungskriterium gegen Einsparpotenziale abgewogen werden.

Eine effiziente Testplanung für die multivariate Lebensdauererprobung erfordert folglich weit mehr als die Maximierung theoretischer Informationsmaße. Sie verlangt einen strategischen Kompromiss: Das Design muss flexibel genug sein, um Extrapolationen durch **ALT** gezielt zu unterstützen (Prognosegüte), robust genug, um die Unsicherheiten der nicht-linearen **GLM**-Parameterschätzung abzufangen (Modellierungsgüte), und dabei eine statistische Trennschärfe gewährleisten, die das Investitionsrisiko des Versuchs rechtfertigt. Nur wenn die Wahrscheinlichkeit der Effektdetektion (*power*) in einem ökonomisch vertretbaren Verhältnis zu den Testkosten steht, ist die Bedingung der Effizienz erfüllt und Effektivität gewährleistet.

3.2 Forschungsfragen und Aufbau der Arbeit

Das übergeordnete Ziel dieser Arbeit, die Entwicklung effizienter multivariater Lebensdauerversuche, fußt damit auf der Prämisse, dass etablierte Standard-Designs nicht zwingend durch vollkommen neuartige Algorithmen ersetzt, sondern vielmehr intelligent adaptiert werden müssen. Als methodischer Anker und Referenzpunkt dient hierbei das **CCD**. Dieser Konsens begründet sich auf einer Synthese aus verfahrenstechnischen, physikalischen und statistischen Vorzügen, die das **CCD** für die Zuverlässigkeitssmodellierung im Maschinen- und Fahrzeugbau prädestinieren. Dies lässt sich aus praktischen Erfahrungen wie beispielsweise nach Rigdon et al. [9], Kleppmann [28] und Myers et al. [36] sowie einer Vielzahl wissenschaftlicher Publikationen ableiten (vgl. Kapitel 2) und umfasst insbesondere folgende Aspekte:

1. **Sequenzielle Augmentierbarkeit:** Das Design unterstützt ideal den ökonomischen Zwang zur Ressourceneffizienz. Es ermöglicht eine schrittweise Erweiterung vom Screening über lineare Modelle bis hin zur quadratischen Modellierung. Dies erlaubt den Abbruch oder die Ausweitung von Versuchen basierend auf Zwischenergebnissen und minimiert das Risiko von Fehlinvestitionen bei zeitintensiven Lebensdauertests [1].

2. **Universelle Prädiktionseigenschaften:** Da der exakte Betriebsbereich oder das Optimum der Zuverlässigkeit a priori oft unbekannt sind und Tests beschleunigt (ALT, Offset zwischen Test- und Feldraum) stattfinden, bietet die Rotierbarkeit des Designs günstige Voraussetzungen für eine richtungsunabhängige Prädiktionsgüte [4, 29].
3. **Wissenschaftliche Fundierung:** Das CCD ist als breit diskutierter und gut erforschter Standard etabliert, was die Akzeptanz und Vergleichbarkeit der Ergebnisse im wissenschaftlichen und industriellen Umfeld sicherstellt, vgl. Rigdon et al. [9], Kleppmann [28] und Montgomery [29].
4. **Validierung physikalischer Modelle:** Durch die Variation auf fünf Faktorstufen ($-\alpha_D, -1, 0, +1, +\alpha_D$) ermöglicht das Design - im Gegensatz zu einfacheren Plänen - die Überprüfung der Gültigkeit physikalischer Beschleunigungsgesetze (z.B. Arrhenius) und die Detektion von Linearitätsabweichungen bei hohen Lasten [4, 6, 77].
5. **Schätzung des reinen Fehlers:** Die Integration mehrfacher Zentralpunkte erlaubt eine robuste Schätzung der natürlichen Streuung (*Pure Error*) unabhängig vom Anpassungsfehler des Modells (*Lack of Fit*). Dies ist für die Berechnung verlässlicher Vertrauensbereiche der Zuverlässigkeit essenziell [4].
6. **Orthogonale Blockbildung:** Das Design lässt sich exzellent in orthogonale Blöcke unterteilen [9, 29]. Dies ist für Langzeitversuche von entscheidender Bedeutung, um zeitliche Trends durch beispielsweise Chargenunterschiede statistisch zu bereinigen, ohne die Parameterschätzung zu verzerrn.
7. **Robustheit gegen Datenverlust:** Aufgrund seiner symmetrischen Struktur weist das CCD eine hohe Toleranz gegenüber fehlenden Werten (*Missing Data*) auf, was die Auswertbarkeit des Versuchs auch bei unvorhergesehenen Ausfällen oder technischer Zensierung auf dem Prüfstand weitgehend sichert [29, 36].

Ausgehend von diesem Status quo zielt die vorliegende Arbeit darauf ab, dieses bewährte Grundkonstrukt nicht zu verwerfen, sondern es gezielt für die spezifischen Anforderungen der Lebensdauerprognose zu modifizieren. Dies führt zur zentralen

Forschungsfrage dieser Arbeit:

Wie können effiziente, multivariate Testdesigns (RSDs) entwickelt und geplant werden, um eine beschleunigte Lebensdauererprobung zum Zweck einer multivariaten Zuverlässigkeitssmodellierung umzusetzen?

Zur Beantwortung dieser Hauptfragestellung werden folgende Teilfragestellungen abgeleitet:

1. **Adaptionsbedarf und -strategie:** Wie können konventionelle Planungsstrategien und Testdesigns - allen voran CCDs - adäquat zum Zweck der beschleunigten Lebensdauererprobung (ALT) angepasst werden?
2. **Bewertungsmethodik:** Wie können die Auswirkungen derartiger Anpassungen (Manipulationen der Orthogonalität) an CCDs auf die Prognosegüte quantitativ charakterisiert und vergleichbar bewertet werden?

3.2.1 Aufbau der Arbeit

Der weitere Aufbau der Arbeit orientiert sich an der logischen Abfolge von der Parameteridentifikation bis zur Design-Optimierung. In **Kapitel 4** wird zunächst das Themenfeld des Screenings adressiert. Hier wird ein Vorschlag zu einem heuristischen Screening-Ansatz vorgestellt, der darauf abzielt, die Effizienz im Prozess der Parameterselektion gezielt für Lebensdauer-beeinflussende Faktoren zu steigern und somit die Basis für nachfolgende RSDs unter dem Prädikat *effizient* zu legen. Grundsätzlich neu ist hierbei die Behandlung mutmaßlicher Wechselwirkungen der Lebensdauer-beeinflussenden Faktoren sowie die Berücksichtigung von Änderungen in der signifikanten Aktivmenge der Einflussparameter durch den Screening-Prozess. Darauf aufbauend widmet sich **Kapitel 5** der generischen Untersuchung und Bewertung von Versuchsplanabweichungen, ausgehend vom CCD als Referenz. Zwar erfolgt die Demonstration der Methodik exemplarisch am CCD, der Anspruch auf Gültigkeit erstreckt sich jedoch auch auf andere RSDs. Dies begründet sich durch die in Abschnitt **2.2.5** dargelegten Eigenschaften, welche universell auf der Fisher-Informationsmatrix sowie der empirischen Datenlage basieren. In diesem Kontext werden ein Kostenmodell sowie ein neu entwickeltes Tool zur Bewertung der Prädik-

3 Effiziente Testplanung für die multivariate Lebensdauererprobung

tionsgüte in Abhängigkeit der **SPV** eingeführt. Daraus abgeleitet werden konkrete Strategieempfehlungen für die Anpassung von **CCDs**, etwa zur Optimierung der Extrapolationsfähigkeit, vorgestellt und diskutiert. Abschließend ist in **Kapitel 6** die Vorstellung einer Fallstudie dokumentiert, um die erarbeitete Methodik an einem praxisnahen Anwendungsbeispiel zu veranschaulichen und die theoretischen Erkenntnisse in den operativen Kontext zu transferieren.

4 Parameter-Screening für multifaktorielle Lebensdauertests

Die in Kapitel 3 hergeleitete Notwendigkeit effizienter Testdesigns setzt voraus, dass die Anzahl der zu untersuchenden Faktoren k auf ein handhabbares Maß begrenzt ist. Da die Komplexität und der Versuchsumfang exponentiell mit der Anzahl der Faktoren steigen (vgl. Gleichung 2.29), ist eine präzise Vorselektion der Einflussgrößen entscheidend. Klassische Ansätze der Versuchsplanung setzen hierfür oft auf experimentelle Screening-Pläne. Im Kontext der Lebensdauererprobung führen diese jedoch zu einem Paradoxon: Um experimentell zu prüfen, ob ein Parameter die Lebensdauer beeinflusst, müssten gleichermaßen bei konventionellen Screening-designs (z.B. 2^{k-p_f} Designs, Plackett-Burman-Designs) sowie modernen RSDs (z.B. OMARS-Designs) bereits zeitintensive EoL-Tests durchgeführt werden, was den Effizienzvorteil des Screenings zunichtemacht.

Dieses Kapitel stellt daher einen effizienten methodischen Ansatz zum **heuristischen Screening** vor. Ziel ist es, basierend auf Expertenwissen und systematischer Analyse eine qualitative Reduktion des Parameterraums vorzunehmen, *bevor* physische Versuche gestartet werden [89]. Dabei liegt der Fokus explizit auf der Unterscheidung zwischen bloßer Robustheit (zum Zeitpunkt $t = 0$) und echter Zuverlässigkeit (über die Zeit $t > 0$). Aus der Perspektive der Zeitdomäne ist die differenzierte Klassifizierung der Einflussfaktoren für die Konzeption zuverlässigkeitstechnischer Untersuchungen von entscheidender Bedeutung. Schließlich sind stochastische Degradationspfade entsprechend Abschnitt 2.1.1 und die daraus resultierende Überlebenswahrscheinlichkeit $R(t)$ nicht statisch zu betrachten, sondern zeitvarianten Interaktionsstrukturen und einer sich verschiebenden Effekthierarchie unterlegen. Folglich kann das Parameterserset, welches das Systemverhalten in fortgeschrittenen Phasen des Lebensdauerzyklus determiniert, signifikant von jenen Faktoren divergieren, die lediglich die initiale Performanceverteilung des Systems bei $t = 0$ dominieren. Die Auswahl der relevanten Kovariaten ist somit inhärent dem

Einfluss der Zeit ausgesetzt. Um dieser Dynamik gerecht zu werden, synergisiert der vorgestellte Ansatz teils bestens etablierte Kreativtechniken (z.B. Brainstorming, Fehlerbaum-Analysen - engl. **Fault-Tree-Analysis (FTA)**, Ishikawa-Diagramme) mit strukturierten Bewertungswerkzeugen (z.B. **DSM**, Grid-Analyse), um eine belastbare Identifikation und Priorisierung der kritischen Einflussgrößen zu gewährleisten.

4.1 Identifikation potenzieller Einflussgrößen

Die Basis eines jeden Screenings bildet die vollständige Erfassung aller potenziellen Einflussgrößen. Um eine lückenlose Identifikation zu gewährleisten, ist eine strukturierte Systemanalyse unerlässlich, wie sie in der Zuverlässigkeitstechnik nach Bertsche und Dazer [1] etabliert ist. Ein zentrales Werkzeug hierfür ist das **Funktionsblockdiagramm**, engl. **Function Block Diagram (FBD)** [42, 46, 90]. Es abstrahiert das technische System auf seine Ein- und Ausgangsgrößen, klassifiziert nach den Flussgrößen *Energie*, *Stoff* und *Signal* (vgl. Abbildung 4.1) [44]. Innerhalb der Systemgrenzen werden Haupt- und Nebenfunktionen definiert, deren Nichterfüllung direkt zu potenziellen Ausfallmechanismen führt. Zur initialen Sammlung

Abbildung 4.1: Funktionsstruktur mit Haupt-, Neben- und Teilfunktionen zur Identifikation von Einflussgrößen (nach Wallace et al. [91])

der Parameter (Informationsbeschaffung) eignen sich, aufbauend auf dem **FBD**, entsprechend Montgomery [92] und Brückner [93] klassische Methoden des Qualitätsmanagements (**Seven Quality Tools (Q7)** and **Seven Management and Planning Tools (M7)**) sowie etliche Kreativtechniken:

- **Literaturrecherche:** Analyse bestehender Ausfallmodi und physikalischer Wirkmechanismen (z.B. Arrhenius, Wöhler) aus vergleichbaren Anwendungen aber auch simulativen Untersuchungen (**Finite-Elemente-Analyse (FEA)**) [94].
- **Brainstorming** bzw. **ABC-Brainstorming** ermöglichen es interdisziplinären Expertenteams, intuitives Erfahrungswissen zu explizieren [43, 45].
- **Delphi-Methode:** Durch mehrstufige, anonymisierte Befragungen von Experten können subjektive Einschätzungen objektiviert und Konsens über potenzielle Einflussgrößen erzielt werden [45].

- **Morphologische Analyse:** Durch die Zerlegung des Systems in Teilfunktionen und die Variation von Lösungsprinzipien werden systematisch konstruktive Einflussparameter aufgedeckt [95, 96].

Das Ergebnis dieses Schrittes ist eine unsortierte, aber möglichst vollständige Liste (Parameter-Pool) aller Größen, die potenziell auf das System einwirken.

4.2 Kreativmethoden zum Auswahlprozesse im Parameter-Screening

Kreativmethoden zum Auswahlprozess im Parameter-Screening Nach der Identifikation muss der Parameter-Pool zunächst strukturiert und anschließend bewertet werden, um die vitalen Einflussgrößen (*vital few*) von den trivialen (*trivial many*) zu trennen, vergleiche Arndt et al. [89]. Insbesondere basierend auf Arbeiten von Gundlach [42], Mayers [43] und Kremer [97] werden im Folgenden geeignete Methoden vorgestellt, die sich für das heuristische Screening im Kontext von zeitvarianten Parametersets in der Lebensdaueranalyse eignen und darüber hinaus auf Randbedingungen der Versuchsplanung Rücksicht nehmen.

4.2.1 Strukturierungsmethoden

Nachdem durch die vorangegangenen Schritte der Informationsbeschaffung eine möglichst vollständige Sammlung aller potenziellen Systemparameter, Inputs und Outputs generiert wurde, bedarf es einer systematischen Strukturierung auf Basis eines vertieften Systemverständnisses. Dieser Schritt darf sich nicht auf eine reine Clusterung in Parametergruppen beschränken. Vielmehr gilt es, die *gegenseitigen Beeinflussungen* der Größen zu identifizieren und zu offenbaren. Insbesondere für die Lebensdaueranalyse stellt dies einen entscheidenden Mehrwert dar, der im Einklang mit den Prinzipien des L-DoE steht [29, 40]: Phänomene wie Alterungseffekte oder die zeitvariante Änderung von Materialeigenschaften sind physikalisch oft nicht trivial durch Einzelgrößen beschreibbar, sondern resultieren maßgeblich aus Interaktionen (vgl. Abbildung 4.5). Um diese komplexen, teils zeitabhängigen Zusammenhänge für die weitere Selektion greifbar zu machen, eignen sich primär grafische Methoden, die über die reine Auflistung hinausgehen:

- **Affinitätsdiagramm:** Als Werkzeug der M7 dient es der thematischen Clusterung der oft unstrukturierten Ergebnisse aus Brainstorming-Sessions [93]. Es ordnet Parameter übergeordneten Kategorien zu und deckt erste logische Gruppierungen auf, womit es als ideale Vorstufe für detailliertere Analysen (z.B. Ishikawa) fungiert.
- **Mind-Mapping:** Diese Methode ermöglicht eine hierarchische Gliederung der Einflussgrößen ausgehend vom Untersuchungsziel (Wurzel) über Haupt- zu Nebenparametern (Äste). Der entscheidende Vorteil für das Screening liegt in der Möglichkeit, *Querbeziehungen* durch Verbindungslien zwischen den Ästen zu visualisieren, wodurch Interdependenzen abseits der direkten Hierarchie offensichtlich werden [98].
- **Erweitertes Ishikawa-Diagramm:** Das klassische Ursache-Wirkungs - Diagramm (Fischgräten-Diagramm) strukturiert Parameter traditionell nach 6M (Mensch, Maschine, Material, Methode, Mitwelt, Messung). Für das heuristische Screening wird es dahingehend modifiziert, dass nicht nur statische Haupteffekte, sondern explizit **Interaktionen** durch *Querverbindungen* zwischen den Ästen visualisiert werden (vgl. Abbildung 4.2) [45, 73].
- **Interdependenz-Netzwerke (Vernetztes Denken):** Für hochkomplexe Systeme bietet sich die Modellierung als gerichteter Graph an. Hierbei werden Parameter als Knoten und ihre Wirkbeziehungen als Kanten dargestellt [99]. Dies erlaubt nicht nur die Abbildung der Wirkrichtung (positiv/negativ), sondern auch die Integration zeitlicher Dynamiken und Intensitäten, was für die Modellierung von Degradationsprozessen vorteilhaft ist.
- **ABC-Analyse:** Zur Reduktion der Komplexität klassifiziert die ABC-Analyse die Parameter nach dem Pareto-Prinzip in Klassen hoher (A), mittlerer (B) und geringer (C) Relevanz. Dies dient als erster Filter, um den Fokus auf die vermuteten Haupttreiber der Lebensdauer zu lenken [29].

Abbildung 4.2: Erweitertes Ishikawa-Diagramm zur Visualisierung von Interaktionen zwischen Einflussfaktoren (adaptiert nach [73])

Die hierdurch erreichte Transparenz über die Vernetzung der Parameter bildet die notwendige Basis für den nächsten Schritt: Die Überführung der qualitativen Struktur in eine quantitative Priorisierung, um die finalen Faktoren für den Versuchsplan zu selektieren.

4.2.2 Bewertungsmethoden (Decision Making)

Aufbauend auf der qualitativen Strukturierung der Einflussgrößen (vgl. Abschnitt 4.2.1) ist nun eine analytische Bewertung erforderlich, um die *vital few* von den *trivial many* zu separieren. Ziel ist es, die gesammelten Parameter in eine Rangfolge zu bringen, die sowohl deren vermutete Relevanz für die Lebensdauer als auch deren Interaktionspotenzial widerspiegelt. Hierfür eignen sich matrixbasierte Ansätze, die eine systematische Paarvergleichung erzwingen und subjektive Einschätzungen objektivierbar machen.

Design-Structure-Matrix (DSM)

Die **DSM** ist eine quadratische $n \times n$ -Matrix, in der alle n identifizierten Systemparameter sowohl in den Zeilen als auch in den Spalten aufgetragen sind. Ein Eintrag y_{ij} in der Matrix symbolisiert dabei eine gerichtete Abhängigkeit: Der Parameter in Zeile i beeinflusst den Parameter in Spalte j [46, 97]. Für das Screening im Kontext der Zuverlässigkeit wird die Matrix häufig um eine Zielgrößen-Spalte erweitert, um den direkten Einfluss auf die Lebensdauer t zu erfassen. Die Bewertung kann auf zwei Detailebenen erfolgen:

- **Binäre DSM:** Erfasst lediglich die Existenz einer Interaktion ($y_{ij} \in \{0,1\}$). Dies eignet sich für frühe Phasen mit geringem Detailwissen, um prinzipielle Vernetzungen aufzudecken.
- **Numerische DSM:** Gewichtet die Stärke des Einflusses, z.B. auf einer Skala von 0 (kein Einfluss) bis 3 (starker Einfluss). Dies erlaubt eine differenzierte Priorisierung.

Ein wesentlicher Vorteil der **DSM** ist die Aufdeckung von Asymmetrien: Ein Parameter A kann Parameter B stark beeinflussen, ohne dass B signifikant auf A zurückwirkt ($y_{AB} \neq y_{BA}$). Solche *aktiven* Parameter sind potenzielle Steuergrößen für den

Versuchsplan [40, 89]. Zudem ermöglicht die Matrix durch algorithmische Partitionierung (Clustering) das Identifizieren von Parametergruppen, die eng miteinander interagieren, was Hinweise auf physikalische Wirkmechanismen liefert [46].

Abbildung 4.3: Beispielhafte Darstellung einer binären DSM zur Identifikation von Interaktionen zwischen Systemparametern (adaptiert nach [73])

Grid-Analyse (Portfolio-Analyse)

Die Grid-Analyse visualisiert die Ergebnisse der numerischen **DSM** in einem zweidimensionalen Portfolio-Diagramm (vgl. Abbildung 4.4). Hierfür werden für jeden Parameter i zwei Kennzahlen berechnet [43]:

1. **Aktivsumme (Active Sum):** Die Summe der Zeileneinträge ($\sum_j y_{ij}$). Sie ist ein Maß dafür, wie stark der Parameter das Gesamtsystem treibt.
2. **Passivsumme (Passive Sum):** Die Summe der Spalteneinträge ($\sum_i y_{ij}$). Sie beschreibt, wie stark der Parameter selbst von anderen Größen beeinflusst wird (Reaktivität).

Die Positionierung im Diagramm erlaubt eine Klassifizierung in vier Quadranten, aus denen sich direkte Handlungsempfehlungen für das **DoE** ableiten:

- **Aktive Parameter (hohe Aktiv-, niedrige Passivsumme):** Diese Größen qualifizieren sich als ideale Steuerfaktoren (x) für den Versuchsplan, da sie das System dominieren, ohne selbst instabil zu sein.
- **Kritische Parameter (hohe Aktiv- und Passivsumme):** Diese Faktoren weisen eine starke Vernetzung auf. Sie sind relevant für die Lebensdauer, bergen aber aufgrund ihrer Abhängigkeiten ein hohes Risiko für unerwünschte Wechselwirkungen. Ihre Überwachung im Versuch ist von besonderer Bedeutung.
- **Passive Parameter (niedrige Aktiv-, hohe Passivsumme):** Diese Größen eignen sich primär als Indikatoren oder Antwortgrößen (y), da sie empfindlich auf Änderungen im System reagieren, während ihre Eignung als Steuerfaktoren als gering einzustufen ist.

4.3 Randbedingungen in der Parameterauswahl für die Zuverlässigkeitssmodellierung

- **Träge Parameter (niedrige Summen):** Diese Faktoren spielen eine untergeordnete Rolle und lassen sich im Screening oft vernachlässigen oder als Konstanten fixieren.

Zur Abbildung der Unsicherheit heuristischer Schätzungen ist eine Erweiterung der Grid-Analyse um Konfidenz-Vektoren möglich [89]. Hierbei geben Experten für jede Bewertung an, ob diese auf gesichertem Wissen (Daten, Literatur) oder Intuition beruht. Große Differenzen zwischen *gesicherter* und *intuitiver* Position im Grid indizieren Wissenslücken, deren Schließung durch experimentelle Voruntersuchungen vor der Aufnahme des Parameters in ein aufwendiges L-DoE erforderlich ist.

Abbildung 4.4: Grid-Analyse zur Klassifizierung von Parametern in aktive, kritische, passive und träge Faktoren

4.2.3 Finale Parameter-Diskussion

Die Ergebnisse aus DSM und Grid-Analyse bilden die Entscheidungsgrundlage für die finale Selektion der Versuchs-Faktoren [89]. In diesem Diskurs erfolgt eine Abwägung der heuristischen Erkenntnisse gegen die harten Restriktionen des Versuchsplans (Budget, Planungsumfang, Messbarkeit). Ziel ist ein Kompromiss zwischen der vollständigen Abbildung aller wirksamen Mechanismen (Modellgüte) und der Reduktion auf eine handhabbare Anzahl an Faktoren (Effizienz). Als relevant identifizierte, aber aus Budgetgründen nicht variierbare Faktoren sind explizit als Konstanten zu dokumentieren oder als Noise-Faktoren in einer Robustheitsbetrachtung zu berücksichtigen.

4.3 Randbedingungen in der Parameterauswahl für die Zuverlässigkeitssmodellierung

Die Anwendung der oben genannten Methoden muss im Kontext der Lebensdaueranalyse spezifische Randbedingungen berücksichtigen, die sich fundamental von der klassischen Robustheitsoptimierung unterscheiden [42, 43, 100].

4.3.1 Unterscheidung: Robustheit vs. Zuverlässigkeit

Ein häufiges Missverständnis im Screening ist die Gleichsetzung von initialer Streuung der Produktperformance mit Lebensdauerstreuung. Wie in Abbildung 4.5 dargestellt, betrachtet die klassische Robustheitsanalyse nach Klein [101] im Verständnis von Taguchi/Shainin oft nur die Verteilung der Systemantwort zum Zeitpunkt $t = 0$. Die Zuverlässigkeitstechnik fokussiert jedoch auf die **Degradation** über die Zeit. Faktoren, die initial kaum Einfluss auf die Performance haben (z.B. Materialalterung, Verschleiß, Wechselwirkungen), können im Zeitverlauf dominante Treiber für Ausfallmechanismen werden. Heuristisches Screening für L-DoE muss daher explizit Parameter priorisieren, die *zeitabhängige* Schädigungsmechanismen treiben, selbst wenn sie *initial inaktiv* erscheinen.

Abbildung 4.5: Abgrenzung zwischen initialer Performance-Verteilung (A) und zeitabhängiger Degradation (B) im Screening-Prozess

4.3.2 Anforderungen aus dem Versuchsplan (L-DoE)

Die im heuristischen Prozess ausgewählten Faktoren werden darauffolgend im statistischen Versuchsplan (DoE) schließlich verarbeitet. Daraus ergeben sich zunächst harte Restriktionen für die Auswahl:

- **Einstell- und Regelbarkeit:** Nur Parameter, die im Versuch aktiv und präzise auf verschiedene Stufen eingestellt werden können, qualifizieren sich als Faktoren. Nicht steuerbare Größen lassen sich sonst in *Noise* oder *Co-Faktoren* einordnen.
- **Beschleunigbarkeit:** Für ALT werden schließlich nur Faktoren gewählt, die eine physikalische Beschleunigung der Schädigung ermöglichen (z.B. Temperatur, mechanische Spannung), ohne den Fehlermechanismus zu verändern.
- **Kosten und Komplexität:** Da der Stichprobenumfang bei voll-faktoriellen Plänen mit $n \propto m \cdot N^k$ wächst, vergleiche Abschnitte 2.2.3, 2.2.4, zwingt die Ökonomie zur drastischen Reduktion von k . Gleicher Verhalten zeigt sich wie schon erwähnt auch bei konventionellen und modernisierten Algorithmen von RSDs.

4.4 Vorgehen zum heuristischen Screening für die Zuverlässigkeitssmodellierung

Das heuristische Screening definiert bestenfalls also ein wirtschaftlich abbildbares Maximum an umsetzbaren Faktoren $k_{\max} \leq 2 \dots 4$, vgl. [29, 33].

- **Design Resolution (Auflösung):** Für den Fall, dass ultimativ dennoch viele vermeintlich relevante Faktoren aus dem Screening-Prozess resultieren, verbleiben nur Screening-Designs für die weitere experimentell empirische Analyse. Das Screening muss vorab klären, welche Interaktionen physikalisch plausibel sind, um eine geeignete Auflösung (Resolution III, IV oder V) oder Konfiguration von Screening-RSDs zu wählen und Aliasing von Haupteffekten mit wichtigen Wechselwirkungen zu minimieren bzw. zu vermeiden.

Mit den genannten Anforderungen zur methodischen Anwendung von L-DoE ermöglichen schließlich die in Abschnitt 4.2 identifizierten heuristischen Methoden für das Parameter-Screening ein strukturiertes Vorgehen zur effizienten Auswahl relevanter Einflussgrößen [89].

4.4 Vorgehen zum heuristischen Screening für die Zuverlässigkeitssmodellierung

Basierend auf der Analyse der Methoden und Randbedingungen wird in Anlehnung an Arndt et al. [89] folgendes prozedurales Vorgehen für das Parameter-Screening in der Zuverlässigkeitssmodellierung vorgeschlagen - vergleiche auch Abbildung 4.6:. Dieser Ansatz fungiert als methodische Barriere vor dem eigentlichen ressourcenintensiven Lebensdauerversuch, um die Effizienz der Datenerhebung zu maximieren.

Abbildung 4.6: Methodischer Ablauf des heuristischen Screenings als Vorstufe zum L-DoE

4.4.1 Phase I: Systemanalyse und Fehlermechanismen

Der Prozess initiiert nicht mit der isolierten Betrachtung von Parametern, sondern mit der Identifikation des relevanten Ausfallmechanismus. Basierend auf dem FBD (Abschnitt 4.1) erfolgt die Definition des spezifischen **Failure Mode** (z.B. Zahnbruch vs. Abrieb bei einem Riemen). Da unterschiedliche Fehlermodi in der Regel von

divergenten Parametersets getrieben werden, führt eine Vermischung unweigerlich zu Rauschen in den Daten und ist zu vermeiden.

4.4.2 Phase II: Heuristische Parameterassessment und Interaktionsanalyse

Im zweiten Schritt unterliegt der Parameter-Pool einer Filterung durch die in Abschnitt 4.2 vorgestellten Methoden (z.B. Ishikawa-Diagramm in Kombination mit Grid-Analyse). Arndt et al. [89] stellen hierzu durch Bewertung der Eignung der Methoden im Kontext der Lebensdaueranalyse fest, dass vornehmlich die Kombination aus erweitertem Ishikawa-Diagramm (zur Visualisierung von Interaktionen) und DSM (zur quantitativen Bewertung der Vernetzung) eine belastbare Grundlage für die Priorisierung der Faktoren darstellt. Kritisch ist hierbei die Bewertung der Interaktionsdichte in Bezug auf die Lebensdauer als Zielgröße. Faktoren, die im Ishikawa-Diagramm oder der DSM eine hohe Vernetzung aufweisen (korrespondierend mit einer hohen Aktivsumme), erfahren eine Priorisierung, da sie mit hoher Wahrscheinlichkeit - insbesondere wechselwirkend über jedenfalls Abschnitte der Lebensdauer - als Katalysatoren für Degradationsprozesse fungieren. Das Ergebnis dieses Prozessschrittes ist eine hierarchische Ordnung der Faktoren nach ihrer vermuteten Relevanz für die Lebensdauer.

4.4.3 Phase III: Abgleich mit dem Versuchsdesign

Im finalen Schritt erfolgt der Abgleich der priorisierten Faktorenliste mit den Restriktionen des gewählten L-DoE (vgl. Abschnitt 3.2 sowie Abschnitt 4.3.2). Hierbei sind folgende Prüffragen essenziell:

- Korrespondiert die Anzahl der Faktoren k mit dem verfügbaren Budget für n Versuche?
- Ist die physikalische Unabhängigkeit der Faktoreinstellung (Orthogonalität) gewährleistet?
- Ermöglichen die gewählten Stufenabstände eine signifikante, detektierbare Änderung der Lebensdauer (Power > 80%), vgl. Abschnitt 2.2.1?

Lediglich Faktoren, die diesen Filter passieren, werden in das finale L-DoE-Modell integriert. Parameter, die als relevant klassifiziert wurden, jedoch aus Budgetgründen entfallen, sind als konstante Randbedingungen zu dokumentieren oder als Noise-Faktoren in Robustheitsanalysen auszulagern.

Das Resultat dieses dreistufigen Vorgehens ist eine kondensierte Parametermenge, die das Optimum zwischen notwendigem Informationsgehalt und vertretbarem experimentellem Aufwand repräsentiert. Wie in Abbildung 4.7 visualisiert, wird das Volumen der zu betrachtenden Einflussgrößen von der initialen Systemanalyse (Gesamtmenge) über die heuristische Bewertung bis hin zum finalen Design-Matching kontinuierlich auf das für das L-DoE essenzielle Maß verdichtet.

Abbildung 4.7: Prozedurale Volumenreduktion des Parameterraums im heuristischen Screening für L-DoE (adaptiert nach Arndt et al. [89])

4.5 Zusammenfassung

Das in diesem Kapitel vorgestellte heuristische Screening schließt die methodische Lücke zwischen der qualitativen Systemanalyse und der quantitativen Versuchsplanning. Durch die strukturierte Nutzung von Expertenwissen (z.B. mittels DSM und erweiterten Ishikawa-Diagrammen) lässt sich die Anzahl der Versuchs-Faktoren k rational reduzieren, ohne relevante Interaktionen zu vernachlässigen. Entscheidend ist dabei der vollzogene Perspektivwechsel von der Betrachtung der initialen Performance-Streuung hin zur Analyse der zeitabhängigen Degradation eines Performancemarkals. Dieses Vorgehen schafft die notwendige Voraussetzung für die Anwendbarkeit der in Kapitel 3 diskutierten effizienten Versuchspläne für Lebensdauererprobungen, da diese ihre Stärken erst bei einer begrenzten Anzahl an Faktoren ($k \approx 3 \dots 5$) wirtschaftlich ausspielen. Die methodische Vorarbeit des Screenings fungiert somit als zentraler Hebel für die Effizienz des gesamten Lebensdauerversuchs.

5 Effizienz und Effektivität in der multivariaten Versuchsplanung

Nach der in Kapitel 4 beschriebenen qualitativen Vorselektion der Einflussgrößen ist nun die quantitative Planung durch die Entwicklung adäquater, effizienter und multivariater Versuchspläne erforderlich. Ziel ist es, etablierte Standard-RSDs gezielt an die spezifischen Anforderungen der Lebensdauerprognose anzupassen, ohne deren statistische Integrität unzulässig zu kompromittieren. In Anlehnung an die Ausführungen aus Abschnitt 3.2 gilt dabei fortwährend die Prämisse, dass insbesondere das CCD aufgrund seiner sequenziellen Augmentierbarkeit (Screening → lineares Modell → quadratische Modellterme, vgl. Abbildung 2.8) im Kontext kostenintensiver EoL-Tests einen entscheidenden operativen Vorzug bietet und daher im Fokus der vorliegenden Untersuchung steht.

Dieses Kapitel widmet sich in zwei Schritten zunächst der systematischen Analyse geometrischer Design-Abweichungen und deren Auswirkung auf die statistische Güte (siehe Abschnitt 5.1). Unabhängig von einer möglichen Beschleunigung durch ALT wird zuerst untersucht, inwieweit typische Manipulationen der Design-Geometrie wie variierende Koordinaten oder asymmetrische Replikationen der Versuchspunkte sowie mangelnde Robustheit in der Versuchssteuerung die Performance beeinflussen. Um diese komplexen Abhängigkeiten generisch und losgelöst von spezifischen Produktdaten zu untersuchen, ist der Einsatz numerischer Simulationen immer dann erforderlich, wenn die Parameterschätzung betreffende Metriken erfasst werden sollen. Mithilfe von MCSs und fiktiven physikalischen Modellen wird die Entwicklung dieser Verhältnismäßigkeiten somit systematisch aufgezeigt, siehe Abschnitt 5.1.1. Erst so werden **Auswirkungen von Orthogonalitäts-Abweichungen** quantifizierbar und damit die Grundlage für die Auswahl von Design-Alternativen geschaffen.

In einem zweiten Schritt werden als signifikant und dominant erfasste Einflüsse durch Design-Modifikationen hinsichtlich ihrer generischen Auswirkung auf die

Modellierung bewertet (Abschnitt 5.1.2). Sie sollen im weiteren Verlauf der Untersuchung als Grundlage für die Auswahl vielversprechender Design-Alternativen dienen und im Anschluss ökonomisch bewertet. Um hierbei valide Aussagen über die Effektivität einer Konfiguration zu treffen, wird neben der bloßen Versuchsplan-Performance oder Schätzgüte maßgeblich die **statistische Trennschärfe** (*power*) als übergeordnete Metrik herangezogen. Sie fungiert als Indikator für die Erfolgsgarantie der Testagenda, indem sie die Wahrscheinlichkeit beziffert, mit der real existierende Effekte tatsächlich mithilfe des jeweiligen Versuchsplans als signifikant erkannt werden, vgl. Abschnitt 2.2.5. Was in der allgemeinen DoE-Methodik oft als nachgeordnete Größe behandelt wird [9, 29], avanciert bei kostenintensiven EoL-Tests zum **Leistungsindikator, engl. Key Performance Indicator (KPI)**: Jede Manipulation des Designs muss zwingend gegen den potenziellen Verlust an der Detektionswahrscheinlichkeit hinsichtlich der Effekte auf die Lebensdauer innerhalb des GLM abgewogen werden.

Darauf aufbauend führt Abschnitt 5.2 ein Kostenmodell ein, um die monetären Konsequenzen derartiger Modifikationen zu quantifizieren. Die Bewertung der Versuchspläne erfolgt dabei streng nach der Kerndefinition von Effizienz als Maß für die Leistungsfähigkeit: Welche Information wird perspektivisch generiert und welcher monetäre Aufwand steht dem gegenüber? Daraus leitet sich die zentrale ökonomische Leitfrage ab: *Wie viel Trennschärfe ist finanziert?* Erst so wird quantifiziert, inwiefern zufällige oder möglicherweise bewusste Design-Modifikationen die **Effizienz** des RSD verändern und zu welchem Maß die **Effektivität** dabei beeinträchtigt wird.

Zur Bewertung der Prognosegüte wird abschließend eine Erweiterung bestehender Visualisierungsmethoden vorgestellt (Abschnitt 5.4), bevor die Gesamtmethodik in der numerischen Simulationsstudie anhand der Zuverlässigkeitssmodellierung mit Aussagesicherheit final bewertet wird (Abschnitt 5.5).

5.1 Analyse generischer Design-Abweichungen

Generell können sich Orthogonalitäts-Abweichungen durch physikalische Restriktionen der technischen Versuchsausführung, die spezifische Performance und Funktionalität des Systems oder Produkts prinzipiell oder bei ALT durch möglicherweise

bewusste Manipulation des Versuchsraums zur Steigerung der Extrapolationsgüte sein. Insbesondere im Kontext der Lebensdauerprognose sind solche Eingriffe praxisrelevant und begründet: Die Verschiebung der Versuchspunkte in Richtung höherer Lastniveaus verkürzt die Prüfzeit und senkt die Kosten, während eine Verortung nahe dem Operationsbereich die Modellanpassung verbessert und die Extrapolationsunsicherheit reduziert.

Allerdings bergen solche Modifikationen das Risiko, die statistische Integrität des Designs zu beeinträchtigen. Im Idealfall sind die Design-Matrix-Spalten eines **CCD** orthogonal zueinander, was eine unabhängige Schätzung der Modellparameter ermöglicht - vgl. Abschnitt 2.2.3. Geometrische Verzerrungen führen zwangsläufig zu einem Verlust dieser Orthogonalität (vgl. Gleichung 2.34). Besonders kritisch ist, dass die Auswirkungen peripherer Parameterstufen (Sternpunkte) im augmentierten Ansatz erst spät im sequenziellen Prozess sichtbar werden - nämlich dann, wenn das initiale faktorielle Design erweitert wird. Daher ist es essenziell, die Robustheit des **CCD** gegenüber typischen Design-Abweichungen systematisch zu quantifizieren. Entgegen bereits erfolgter Untersuchungen durch Box [102], Ahn [84], Donev [85], Zhang und Yue [103] sowie Jang [104] oder vergleichbarer liegt der Fokus hierbei allerdings nicht auf der reinen geometrischen Verzerrung und der Analyse sowie Bewertung durch geeignete Optimalitätskriterien (entsprechend Abschnitt 2.2.5). Die hier erwähnten Beiträge konzentrieren sich primär auf die geometrische Charakterisierung der Design-Abweichungen und deren Einfluss auf die Struktur der Design-Matrix bzw. geeigneter Optimalitätskriterien. Vielmehr erfolgt hier eine Quantifizierung der Auswirkungen auf die Schätzpräzision und insbesondere die Trennschärfe im Kontext von Lebensdauerprognosen. Die Bewertung mittels Optimalitätskriterien wäre aus genannten Gründen ohnehin nicht zielführend, da im Umkehrschluss die Erfüllung der Kriterien im Planungsprozess erneut Unkenntnis über die tatsächlichen Auswirkungen auf die Schätzpräzision und Trennschärfe erzeugen würde. Dies hat zur Folge, dass zunächst eine numerische Simulation erforderlich ist, um die Auswirkungen von Design-Abweichungen auf die statistische Güte quantifizieren zu können, bevor überhaupt eine Bewertung der Design-Alternativen erfolgen kann.

5.1.1 Methodik zur simulativen Ermittlung von Auswirkungen durch Orthogonalitäts-Abweichungen

Um die Auswirkungen von Design-Abweichungen auf die statistische Güte quantifizieren zu können, bedarf es einer kontrollierten Untersuchungsumgebung, in der - im Gegensatz zu physischen Experimenten - die *wahre* Natur des Systems bekannt ist. Durch den Vergleich von beiderseits Modellierungs- als auch Analyse-Ergebnissen mit einer vorab definierten Wahrheit (engl. *Ground Truth*) lassen sich **KPIs** wie die Schätzpräzision und allen voran die *power* bestimmen. Zu diesem Zweck wird für die Untersuchungen eine allgemeingültige Simulationsumgebung geschaffen, die auf einem scheinbar-physikalischen Modell basiert [105, 106].

Pseudo-physikalisches Simulationsmodell

Das zugrundeliegende Systemverhalten im Sinne einer *Ground Truth* wird **global** durch ein polynomiales Regressionsmodell zweiter Ordnung approximiert. Dies entspricht der Struktur, die auch als Ergebnis der **RSM**-Untersuchung erwartet wird (vgl. Gleichung 2.27) und möglicherweise im Zuge einer Zuverlässigkeitsuntersuchung eines technischen Produktes resultieren kann. Um die Anforderungen an eine multivariate Untersuchung ($k \geq 2$) zu erfüllen und die Komplexität handhabbar zu halten, wird im Basisszenario von zwei Faktoren (x_1, x_2) ausgegangen. Das Modell berücksichtigt neben den Haupteffekten explizit die für die Lebensdaueranalyse kritischen Wechselwirkungen ersten Grades sowie quadratische Terme zur Abbildung von Nichtlinearitäten:

$$y_{\text{sim}} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \epsilon. \quad (5.1)$$

Um ein realistisches, stochastisches Verhalten abzubilden, sind die Modellparameter nicht deterministisch fixiert. Den Regressionskoeffizienten β_j werden zwar nominale Startwerte zugewiesen (vgl. Tabelle 5.1), diese werden jedoch in jedem Simulationslauf mit einer definierten Streuung überlagert. Zusätzlich wird die resultierende Systemantwort y_{sim} mit einer Reststreuung ϵ beaufschlagt, um die inhärente **aleatorische Unsicherheit** des Systems (Messfehler, unbekannte Einflüsse) abzubilden. Die Default-Werte für die Koeffizienten sowie die vorgegebene Standardabweichung

chung σ_{sim} , welche als Streuparameter für das stochastische Basis-Verhalten dient, sind in Tabelle 5.1 zusammengefasst.

Tabelle 5.1: Nominale Startwerte und Streuparameter des pseudo-physikalischen Simulationsmodells

Modellparameter	Default-Wert
Regressionskonstante β_0	30
Regressionskoeffizienten β_j (für $j \in \{1, 2, 12, 11, 22\}$)	10
Standardabweichung der Systemantwort σ_{sim}	1

Simulationsalgorithmus und Monte-Carlo-Ansatz

Das Modell der *Ground Truth* wird zunächst mit den zu untersuchenden Versuchsplänen gekoppelt. Für jeden Stützpunkt des Versuchsplans (Versuchspunkte) werden die korrespondierenden Systemantworten simuliert. Anschließend erfolgt die Rückrechnung des Regressionsmodells mittels **MLE**, entsprechend der ausgeführten Methodik aus Abschnitt 2.1.4. Um statistisch belastbare Aussagen zu treffen und **epistemische Unsicherheiten** auf einen minimalen numerischen Approximationsfehler zu reduzieren, wird das Verfahren in eine **MCS** mit $n_{\text{MC}} = 10.000$ Iterationen eingebettet - vergleiche Abschnitt 2.2.5 für entsprechende Hintergründe. Dieser Ansatz erlaubt die Bewertung zweier zentraler Kriterien:

1. **Präzision der Parameterschätzung:** Durch den (Mittelwert-) Vergleich der geschätzten Parameter $\hat{\beta}$ mit den *wahren* Koeffizienten β_{true} lassen sich Bias und Varianz der Schätzung ermitteln.
2. **Bestimmung der Trennschärfe power:** Für jeden Simulationslauf wird geprüft, ob die signifikanten Effekte korrekt erkannt wurden ($p\text{-Wert} < \alpha$). Das Verhältnis der Treffer zur Gesamtzahl der Simulationen ergibt die empirische **power** (vgl. Gleichung 2.55).

5.1.2 Sensitivitätsanalyse geometrischer und stochastischer Einflussgrößen

Um zunächst die Robustheit des methodischen Ansatzes gegenüber den eingangs erwähnten Einflussgrößen zu bewerten, wird eine Sensitivitätsanalyse durchgeführt,

welche methodische einem **DoE on Experimental Design**-Ansatz folgt: Die Eigenschaften des Versuchsplans sowie die der *Ground Truth* werden als Faktoren eines übergeordneten Meta-Designs derart variiert, sodass deren Einfluss auf die Effektivität der Schätzung und damit insbesondere die Trennschärfe der Effekte im **GLM** quantifiziert werden kann [105]. Infrage kommt dabei eine aus praktischer Überlegung erfasste Auswahl an sowohl geometrische Manipulationen als auch stochastische Einflüsse sowie methodische Parameter und deren Interaktionen:

1. Manipulation von Sternpunkten über deren axialen Abstand,
2. Manipulation von Sternpunkten durch co-faktorielle Korrelation,
3. Variation der Anzahl der Zentralpunkte,
4. Variation der Anzahl der Sternpunkte,
5. Variation der Standardabweichung der Effektstärken (System-Rauschen),
6. Änderung des relativen Verhältnisses von Effektstärken untereinander,
7. Variation der Standardabweichung der Systemantwort (Mess-Rauschen),
8. Variation des Signifikanzniveaus.

Insbesondere die Berücksichtigung der Verhältnismäßigkeit von Effektstärken sowie der Rauschparameter ist von zentraler Bedeutung, da sie die Signal-zu-Rausch-Verhältnisse im Modell direkt beeinflussen und damit die Detektionswahrscheinlichkeit der Effekte maßgeblich steuern, gleichzeitig aber eine gegen unendlich tendierende Variationsmöglichkeit bieten, die in einem reinen Screening-Ansatz nicht abgedeckt werden könnte. Zur formalen Beschreibung des Untersuchungsraums werden diese acht Einflussgrößen, welche sowohl deterministische geometrische Manipulationen als auch stochastische kontinuierliche Einflussgrößen abbilden, im Vektor $\xi = [\xi_a, \dots, \xi_h]'$ zusammengefasst und in Tabelle 5.2 definiert. Etwaige Ergänzungen wie die Variation der Anzahl der Faktoren, die Manipulation der faktoriellen Versuchspunkte oder die Einführung von Lebensdauermodellen werden dabei bewusst ausgeschlossen, um die Komplexität des Meta-Designs handhabbar zu halten und die Interpretation der Ergebnisse durch eine pragmatische Auswahl und im Fokus der in Kapitel 3 postulierten Zielsetzung zu erleichtern.

So werden schließlich Haupteffekte, Interaktionen sowie quadratische Effekte auf die Trennschärfe als wesentlicher Performance-KPI und damit Metrik zur Effektivität ermittelt.

5.1 Analyse generischer Design-Abweichungen

Tabelle 5.2: Definition und Parametrisierung der Untersuchungsfaktoren für die Sensitivitätsanalyse

	Beschreibung	Wertebereich	Charakteristik
ξ_a	Sternpunkt-Abstand (α_D) Variation des Abstands der Sternpunkte zum Zentrum (z. B. flächenzentriert, rotierbar und extendiert).	[1,00; $2 \cdot \alpha_D$]	Kont. (CCD)
ξ_b	Sternpunkt-Shift Asymmetrische / Co-Faktorielle Verschiebung eines Sternpunkts zur Simulation geometrischer Verzerrungen und Faktoren-Korrelation.	[0,10; α_D]	Kont. (CCD)
ξ_c	Zentralpunkte Variation der Anzahl der Zentralpunkte $n_C > 0$.	[1,00; 8,00]	Diskret (Faktoriell)
ξ_d	Sternpunkte Variation durch Entfall von maximal 50 % der Sternpunkte (z. B. durch Zensierung oder Funktionalitätsverlust).	[2,00; 4,00]	Diskret (Faktoriell)
ξ_e	System-Rauschen Verhältnis zwischen pseudophysikalischen Modellkoeffizienten und Systemabweichung bzw. Effektsreuung.	[0,80; 1,20]	Kont. (CCD)
ξ_f	Effekt-Verhältnis Verhältnis zweier Haupteffekte zueinander: zwischen gleichwertiger Effektstärke oder einem Anteil von 10 %.	[0,10; 1,00]	Kont. (CCD)
ξ_g	Mess-Rauschen Verhältnis zwischen den Modellkoeffizienten und dem Wert des Streuparameters der Systemantwort.	[0,80; 1,20]	Kont. (CCD)
ξ_h	Signifikanzniveau Das gewählte Signifikanzniveau.	[0,02; 0,08]	Kont. (CCD)

5.1.3 Ergebnisse der Sensitivitätsanalyse und Quantifizierung der Trennschärfe

Die simulative Untersuchung der Trennschärfe (*power*) über den durch ξ definierten Untersuchungsraum offenbart eine signifikante Dichotomie der Einflussgrößen. Es muss zwischen **dispositiven Designfaktoren** (Geometrie ξ_a , Stützstellen ξ_c, ξ_d), die vom Versuchsplaner aktiv modifiziert werden können, und **externen Restriktionen** (Rauschen ξ_e, ξ_g , Signifikanzniveau ξ_h), die durch das physikalische Testsystem oder normative Vorgaben als unveränderlich zu betrachten sind, unterschieden werden.

Betrachtet man zunächst die globale **Modell-Power** ($power_M$), so zeigt sich eine extrem hohe Robustheit gegenüber allen Variationen mit einem Mittelwert von $power \approx 99,7\%$. Diese pauschale Metrik verdeckt jedoch, dass das Verhalten spezifischer Einzel-Effekte – insbesondere die Identifikation der für die Lebensdauerprognose kritischen quadratischen Terme – deutlich sensitiver ausfällt. Wie Tabelle 5.3 zeigt, liegt die mittlere Trennschärfe für quadratische Terme ($power_{\beta_{11}}, power_{\beta_{22}}$) nur noch im Bereich von 70 % bis 79 %.

Die entscheidende Erkenntnis liefert jedoch der Blick auf die **Sensitivitätsindizes** (Spalte „Dominante Sensitivität“). Für die Mehrheit der untersuchten Koeffizienten ($power_M, power_{\beta_1}, power_{\beta_{12}}, power_{\beta_{11}}$) ist nicht etwa eine Reduktion der Versuchspunkte der limitierende Faktor, sondern das **Messrauschen** (ξ_g). Mit einem Sensitivitätsindex von bis zu $E = 1,99$ (bei $power_{\beta_{11}}$) dominiert die Güte der Messtechnik die Erfolgswahrscheinlichkeit des Tests. Da das Rauschen (ξ_g, ξ_e) sowie das Signifikanzniveau (ξ_h) in der praktischen Anwendung bei EoL-Tests an bestehenden Prüfständen meist feste Randbedingungen darstellen, definieren sie eine natürliche Obergrenze der erreichbaren Trennschärfe, die auch durch maximale Versuchspläne nicht durchbrochen werden kann.

Dies birgt ein erhebliches ökonomisches Potenzial: Wenn die Trennschärfe ohnehin primär durch das System- und Messrauschen limitiert wird, ist zu prüfen, ob Design-Vereinfachungen (Reduktion von ξ_c und ξ_d) überhaupt zu einer *signifikanten* weiteren Verschlechterung führen. Tabelle 5.4 quantifiziert diesen „Design-Penalty“. Es zeigt sich, dass der Verzicht auf Zentralpunkte (ξ_c) und Sternpunkte (ξ_d) zwar die Power reduziert (z. B. $\Delta power \approx -13\%$ für $power_{\beta_{11}}$), dieser Einfluss jedoch in der gleichen Größenordnung liegt wie die unvermeidbaren Rauscheffekte.

5.1 Analyse generischer Design-Abweichungen

Tabelle 5.3: Zusammenfassende Statistik der Trennschärfe und Identifikation der Haupt-Einflussfaktoren. Trennung in unveränderliche Systemgrößen vs. modifizierbare Designgrößen.

Koeffizient	Mittelwert	Wertebereich	Dominante Sensitivität	
			Faktor (Kategorie)	E
$power_M$	99,70 %	[94,51; 100,00]	ξ_g (System)	0,104
$power_{\beta_0}$	97,86 %	[76,85; 100,00]	ξ_e (System)	0,515
$power_{\beta_1}$	78,55 %	[59,99; 90,77]	ξ_g (System)	1,496
$power_{\beta_2}$	65,98 %	[37,34; 84,06]	ξ_f (System)	2,530
$power_{\beta_{12}}$	72,58 %	[52,47; 85,90]	ξ_g (System)	1,930
$power_{\beta_{11}}$	70,94 %	[32,59; 86,81]	ξ_g (System)	1,988
$power_{\beta_{22}}$	79,31 %	[33,37; 100,00]	ξ_a (Design)	3,724

Tabelle 5.4: Systematischer Power-Verlust bei Reduktion der Versuchsstützpunkte (Vergleich Augmentiertes Design vs. Minimal-Design)

Koeffizient	Effekt ξ_c (8 → 1 CP)	Effekt ξ_d (4 → 2 SP)
$power_{\beta_0}$	-4,45 %	-0,30 %
$power_{\beta_1}$	-2,43 %	-5,99 %
$power_{\beta_2}$	+1,07 %*	-7,95 %
$power_{\beta_{11}}$	-13,87 %	-12,80 %
$power_{\beta_{22}}$	n. a.	-6,07 %

* Positiver Wert resultiert aus Varianzverschiebungen bei instabilen Designs.

Einzig der quadratische Koeffizient $power_{\beta_{22}}$ bildet eine Ausnahme: Hier ist der geometrische Abstand der Sternpunkte (ξ_a) mit einem extremen Index von 3,72 der dominierende Faktor (vgl. Tabelle 5.3). Für alle anderen Terme gilt: Da die Power durch die physikalischen Randbedingungen (ξ_e, ξ_g) maßgeblich vorbestimmt ist, besteht Spielraum zur Anpassung der Faktoren ξ_c und ξ_d in Bereiche, die aus Aspekten der Versuchszeit und des Budgets attraktiv sind, ohne die Effektivität des Tests unverhältnismäßig stärker zu kompromittieren, als es das Messsystem ohnehin tut.

5.1.4 Fazit: Statistische Effektivität und ökonomische Relevanz

Aus den numerischen Analysen lässt sich ein zentrales Fazit für die Lebensdauererprobung ableiten: Das **CCD** erweist sich als äußerst robust gegenüber moderaten geometrischen Modifikationen. Die statistische Effektivität, repräsentiert durch die globale und modellspezifische Trennschärfe, wird primär durch nicht beeinflussbare, exogene System- und Messrauschen-Parameter limitiert. Eine gezielte Reduktion der Versuchsumfänge - wie der partielle Wegfall von Zentral- oder Sternpunkten - bewirkt zwar einen messbaren isolierten Verlust an Trennschärfe, diesen *Design-Penalty* gilt es jedoch stets in Relation zur ohnehin limitierenden Wirkung des experimentellen Rauschens zu setzen.

Für die praktische **EoL**-Versuchsplanung bedeutet dies einen Paradigmenwechsel: Anstatt **RSDs** um jeden Preis auf strenge geometrische Orthogonalität und eine maximale Stützstellenanzahl zu trimmen, können und sollten Versuchspläne flexibel an betriebliche Limitierungen adaptiert werden. Der bewusste Tausch einer geringfügig reduzierten statistischen Trennschärfe gegen eine potenziell massive Einsparung an Versuchszeit und finanziellen Ressourcen stellt folglich einen validen Optimierungsansatz dar. Damit ist die methodische Grundlage geschaffen, um Modifikationen am Versuchsplan im folgenden Abschnitt in einem dedizierten Kostenmodell monetär zu bewerten und der statistischen Leistungsfähigkeit gegenüberzustellen. Das Ziel ist eine holistische Bewertungsmetrik, die den Informationsgewinn direkt mit dem wirtschaftlichen Aufwand verknüpft.

5.2 Ökonomische Bewertung: Das Kostenmodell

Eine rein statistische Bewertung greift in der industriellen Praxis zu kurz. Eine Design-Modifikation, die die Prädiktionsvarianz um 5% senkt, aber die Versuchskosten verdoppelt, ist ineffizient. Daher wird ein Kostenmodell eingeführt, das die monetären Aufwände eines Lebensdauerversuchs in Abhängigkeit der Design-Parameter abbildet. Die Gesamtkosten C_{total} setzen sich additiv aus fixen und variablen Anteilen zusammen:

$$C_{\text{total}} = C_{\text{fix}} + (n \cdot C_{\text{spec}}) + \left(\sum_{i=1}^n t_{\text{test}i}(x_i) \cdot C_{\text{time}} \right) + C_{\text{fail}}. \quad (5.2)$$

Hierbei beschreibt C_{fix} die Rüstkosten, C_{spec} die Herstellkosten pro Prüfling und C_{time} die laufenden Betriebskosten des Prüfstands pro Zeiteinheit. Der Term $t_{\text{test}i}(x_i)$ ist die entscheidende, vom Versuchsplan abhängige Größe: Durch **ALT** (hohe Lastniveaus x) sinkt die Prüfzeit, was C_{total} reduziert. Gleichzeitig steigt jedoch durch aggressive Lasten potenziell der Energiebedarf oder das Risiko nicht-repräsentativer Ausfallmechanismen (C_{fail}). Dieses Modell ermöglicht es, Design-Alternativen nicht nur nach ihrer statistischen Effizienz (D-, G-Effizienz), sondern nach ihrer *ökonomischen Effizienz* (Information pro Euro) zu bewerten. Dieses Modell ermöglicht es, Design-Alternativen nicht nur nach ihrer statistischen Effizienz (D-, G-Effizienz), sondern nach ihrer *ökonomischen Effizienz* (Information pro Euro) zu bewerten.

5.3 Orthogonalitäts-Abweichungen in der Zuverlässigskeitsprognose

5.4 Extrapolation Variance Graph (EVG)

Zur Beurteilung der Vorhersagegüte im gesamten Versuchsraum haben sich der **VDG** und der **FDS-Plot** etabliert. Diese Metriken bewerten die Verteilung der skalierten Prädiktionsvarianz (**SPV**) jedoch oft global oder über sphärische Distanzen gemittelt. Für die Lebensdauerprognose, bei der oft eine Extrapolation in einen spe-

zifischen, eng begrenzten Anwendungsbereich ("Corner Case") von Interesse ist, sind diese globalen Maße oft unzureichend.

Daher wird in dieser Arbeit eine neue, erweiterte Auswertungsmethode vorgestellt: Das **Extrapolations-spezifische SPV-Mapping**. Im Gegensatz zum klassischen VDG, der die Varianz über Radien mittelt, visualisiert diese Methode die Isolinien der Prädiktionsvarianz explizit in Richtung der relevanten Belastungsvektoren. Dies ermöglicht die gezielte Bewertung der Design-Güte exakt in jenen Bereichen des Parameterraums, in denen die spätere Feldprognose erfolgen soll, und deckt Schwachstellen auf, die in globalen FDS-Plots verborgen blieben.

5.5 Zuverlässigkeitsmodellierung mit Vertrauensbereichen

5.6 Simulative Validierung und Methodennachweis

Aufgrund der unendlichen Kombinationsmöglichkeiten von Einflussfaktoren, Lastniveaus und physikalischen Modellen ist eine generische Sensitivitätsanalyse für alle denkbaren Anwendungsfälle nicht zielführend. Stattdessen wird die Anwendbarkeit und Leistungsfähigkeit der vorgeschlagenen Methodik am Beispiel eines generischen Referenzsystems mittels Monte-Carlo-Simulation (MC) nachgewiesen.

5.6.1 Simulationssetup

Als Referenzsystem dient ein Weibull-verteiltes System ($b = 2$, $T = 1000$ h), dessen Lebensdauerverhalten durch ein log-lineares Modell mit zwei Hauptfaktoren und einer Wechselwirkung beschrieben wird. Im Rahmen der MC-Simulation werden $n_{MC} = 10.000$ virtuelle Versuchsreihen erzeugt, wobei für jeden Lauf: 1. Lebensdauerdaten basierend auf dem definierten Modell und den Versuchsplan-Stützstellen (modifiziertes CCD vs. Benchmark-CCD) generiert werden. 2. Eine stochastische Streuung sowie Zensierung (Typ I/II) aufgeprägt wird. 3. Die Modellparameter mittels **MLE** zurückgeschätzt werden.

5.6.2 Ergebnisse und Effizienzbewertung

Die Auswertung erfolgt multidimensional anhand der in den vorangegangenen Abschnitten definierten Metriken:

- **Power:** Wie hoch ist die Wahrscheinlichkeit, die definierten Effekte trotz Design-Manipulation korrekt zu detektieren?
- **Parameterschätzgüte:** Wie verhalten sich Bias und Varianz der Schätzer (\hat{b} , \hat{T}) bei Verlust der Orthogonalität?
- **Ökonomische Bilanz:** Welches Design liefert die geforderte Präzision bei minimalen Gesamtkosten gemäß Gleichung 5.2?

Durch den Vergleich der gesampleten Versuchszeiten (Time-to-Failure) kann valide auf die Effizienz im ALT-Kontext geschlossen werden. Es wird gezeigt, dass ein gezielt manipuliertes CCD, bei dem Sternpunkte in Richtung der Hauptbelastung verschoben werden, die Prognoseunsicherheit im Anwendungsbereich signifikant senken kann, ohne die maximale Versuchsanzahl oder das Budget des konventionellen Benchmarks zu überschreiten. Damit wird der Nachweis erbracht, dass die entwickelte Methodik zur Bewertung und Anpassung von Versuchsplänen (Kostenmodell + SPV-Visualisierung + Orthogonalitäts-Check) ein robustes Werkzeug zur Steigerung der Testeffizienz darstellt.

6 Fallstudie

7 Zusammenfassung und Ausblick

7.1 Key Findings

7.2 Diskussion

7.3 Ausblick

Literatur

- [1] B. Bertsche und M. Dazer. *Zuverlässigkeit im Fahrzeug- und Maschinenbau: Ermittlung von Bauteil- und System-Zuverlässigkeiten*. 4. Auflage. Lehrbuch. Berlin und Heidelberg: Springer Vieweg, 2022. ISBN: 978-3-662-65023-3. URL: <http://www.springer.com/>.
- [2] R. Koselka. „The New Mantra: MVT“. In: *Forbes* (11.03.1996), S. 114–118.
- [3] *DAT Report 2025: Kurzbericht*. Ostfildern, Deutschland, Februar 2025. URL: https://www.dat.de/fileadmin/protected/DAT_Report/2025/DAT-Report-2025-Kurzbericht.pdf.
- [4] W. Q. Meeker, L. A. Escobar und F. G. Pascual. *Statistical methods for reliability data*. Second edition. Wiley series in probability and statistics. Hoboken, NJ: Wiley, 2022. ISBN: 978-1-118-11545-9.
- [5] A. Birolini. *Reliability Engineering*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2017. ISBN: 978-3-662-54208-8. doi: [10.1007/978-3-662-54209-5](https://doi.org/10.1007/978-3-662-54209-5).
- [6] G. Yang. *Life cycle reliability engineering*. Hoboken, N.J.: Wiley, 2007. ISBN: 9780471715290. URL: <http://www.loc.gov/catdir/enhancements/fy0740/2006019150-b.html>.
- [7] L. Fahrmeir, C. Heumann, R. Künstler, I. Pigeot und G. Tutz. *Statistik*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2016. ISBN: 978-3-662-50371-3. doi: [10.1007/978-3-662-50372-0](https://doi.org/10.1007/978-3-662-50372-0).
- [8] E. L. Kaplan und P. Meier. „Nonparametric Estimation from Incomplete Observations“. In: *Journal of the American Statistical Association* 53.282 (1958), S. 457. ISSN: 01621459. doi: [10.2307/2281868](https://doi.org/10.2307/2281868).
- [9] S. E. Rigdon, R. Pan, D. C. Montgomery und L. J. Freeman. *Design of experiments for reliability achievement*. Hoboken, NJ: Wiley, 2022. ISBN: 9781119237693. doi: [10.1002/9781119237754](https://doi.org/10.1002/9781119237754).

- [10] J. Hedderich und L. Sachs. *Angewandte Statistik*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2020. ISBN: 978-3-662-62293-3. doi: [10.1007/978-3-662-62294-0](https://doi.org/10.1007/978-3-662-62294-0).
- [11] H. Rinne. *The Weibull Distribution: A Handbook*. 1st ed. London: CRC Press LLC, 2008. ISBN: 978-1-4200-8743-7. URL: <https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=1633687>.
- [12] D. Kececioglu. *Reliability & life testing handbook*. Lancaster, PA: DEStech Publications, 2002. ISBN: 1-932078-03-7.
- [13] D. C. Montgomery, E. A. Peck und G. G. Vining. *Introduction to linear regression analysis*. Sixth edition. Wiley series in probability and statistics. Hoboken, New Jersey: Wiley, 2021. ISBN: 978-1-119-57872-7.
- [14] W. Nelson. *Accelerated Testing: Statistical Models, Test Plans, and Data Analysis*. 1st ed. Bd. v.344. Wiley Series in Probability and Statistics Series. Newark: John Wiley & Sons Incorporated, 1990. ISBN: 978-0471697367. URL: <https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=469757>.
- [15] J. D. Kalbfleisch. *The statistical analysis of failure time data*. 2nd ed. Wiley series in probability and statistics. Hoboken, N.J: J. Wiley, 2002. ISBN: 047136357X. doi: [10.1002/9781118032985](https://doi.org/10.1002/9781118032985).
- [16] W. Nelson. *Applied Life Data Analysis*. Bd. v.577. Wiley series in probability and statistics. Hoboken: John Wiley & Sons Inc, 2005. ISBN: 9780471644620. URL: <https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=226561>.
- [17] H. Qiao und C. P. Tsokos. „Parameter estimation of the Weibull probability distribution“. In: *Mathematics and Computers in Simulation* 37.1 (1994), S. 47–55. ISSN: 0378-4754. doi: [10.1016/0378-4754\(94\)90058-2](https://doi.org/10.1016/0378-4754(94)90058-2).
- [18] A. Kremer und B. Bertsche. „A New Approach for Parametrizing Multidimensional Lifetime Models“. In: *Proceedings of the 29th European Safety and Reliability Conference (ESREL)*. Hrsg. von M. Beer und E. Zio. Europe und Singapore: European Safety and Reliability Association und Research Publishing Services, 2019, S. 715–722. ISBN: 978-981-11-2724-3. doi: [10.3850/978-981-11-2724-3{\textunderscore}0443-cd](https://doi.org/10.3850/978-981-11-2724-3{\textunderscore}0443-cd).

- [19] R. B. Abernethy. *The new Weibull handbook: Reliability & statistical analysis for predicting life, safety, risk, support costs, failures, and forecasting warranty claims, substantiation and accelerated testing, using Weibull, log normal, Crow-AMSAA, Probit, and Kaplan- meier Models*. 5. ed. North Palm Beach, Fla.: R.B. Abernethy, 2006. ISBN: 0965306232.
- [20] H. Hirose. „Bias correction for the maximum likelihood estimates in the two-parameter Weibull distribution“. In: *IEEE Transactions on Dielectrics and Electrical Insulation* 6.1 (1999), S. 66–68. ISSN: 10709878. DOI: [10 . 1109 / 94 . 752011](https://doi.org/10.1109/94.752011).
- [21] R. Ross. „Bias and standard deviation due to Weibull parameter estimation for small data sets“. In: *IEEE Transactions on Dielectrics and Electrical Insulation* 3.1 (1996), S. 28–42. ISSN: 10709878. DOI: [10 . 1109 / 94 . 485512](https://doi.org/10.1109/94.485512).
- [22] A. Kremer und B. Bertsche. „A Methodology for Consideration of Uncertainty in Lifetime Design of Experiments“. In: *2019 Annual Reliability and Maintainability Symposium (RAMS)*. IEEE, 2019, S. 1–6. ISBN: 978-1-5386-6554-1. DOI: [10 . 1109 / RAMS . 2019 . 8768906](https://doi.org/10.1109/RAMS.2019.8768906).
- [23] J. F. Lawless. *Statistical models and methods for lifetime data*. 2. ed. Wiley series in probability and statistics. Hoboken, NJ: Wiley-Interscience, 2003. ISBN: 0471372153. URL: [http : / / www . loc . gov / catdir / bios / wiley044 / 2002151805 . html](http://www.loc.gov/catdir/bios/wiley044/2002151805.html).
- [24] L. J. Bain und M. Engelhardt. *Statistical Analysis of Reliability and Life-Testing Models*. Routledge, 2017. ISBN: 9780203738733. DOI: [10 . 1201 / 9780203738733](https://doi.org/10.1201/9780203738733).
- [25] R. A. Fisher. *The Design of Experiments*. Edinburgh: Oliver & Boyd, 1935.
- [26] G. E. P. Box, W. G. Hunter und J. S. Hunter. *Statistics for experimenters: An introduction to design, data analysis, and model building*. Wiley series in probability and mathematical statistics. New York: Wiley, 1978. ISBN: 0471093157. URL: [http : / / www . loc . gov / catdir / description / wiley033 / 77015087 . html](http://www.loc.gov/catdir/description/wiley033/77015087.html).
- [27] G. Taguchi. *Taguchi's quality engineering handbook*. Hoboken, NJ: Wiley, 2005. ISBN: 9780471413349. DOI: [10 . 1002 / 9780470258354](https://doi.org/10.1002/9780470258354).
- [28] W. Kleppmann. *Versuchsplanung: Produkte und Prozesse optimieren*. 10., überarbeitete Auflage. Praxisreihe Qualität. München: Hanser, 2020. ISBN: 978-3-446-46397-4.

- [29] D. C. Montgomery. *Design and Analysis of Experiments*. 10th. Hoboken, NJ, USA: John Wiley & Sons, Inc, 2020. ISBN: 9781119722106.
- [30] C.-F. Wu und M. Hamada. *Experiments: Planning, analysis and optimization*. Third edition. Wiley series in probability and statistics. Hoboken, NJ, USA: Wiley, 2021. ISBN: 9781119470106. URL: <https://zbmath.org/?q=an%3A1460.62003>.
- [31] W. Q. Meeker und L. A. Escobar. „A Review of Recent Research and Current Issues in Accelerated Testing“. In: *International Statistical Review / Revue Internationale de Statistique* 61.1 (1993), S. 147. ISSN: 03067734. DOI: [10.2307/1403600](https://doi.org/10.2307/1403600).
- [32] E. A. Elsayed und H. Zhang. „Design of PH-based accelerated life testing plans under multiple-stress-type“. In: *Reliability Engineering & System Safety* 92.3 (2007), S. 286–292. ISSN: 09518320. DOI: [10.1016/j.ress.2006.04.016](https://doi.org/10.1016/j.ress.2006.04.016).
- [33] K. Siebertz, D. van Bebber und T. Hochkirchen. *Statistische Versuchsplanung: Design of Experiments (DoE)*. 2. Auflage. VDI-Buch. Berlin und Heidelberg: Springer Vieweg, 2017. ISBN: 978-3-662-55742-6.
- [34] P. Goos und B. A. Jones. *Optimal Design of Experiments: A Case Study Approach*. 1st. Hoboken, NJ: John Wiley & Sons, Inc, 2011. ISBN: 9780470744611.
- [35] K. Hinkelmann. *Design and Analysis of Experiments - Volume 3: Special Designs and Applications*. 1st. John Wiley & Sons, Inc, 2012. ISBN: 978-0470-53068-9.
- [36] R. H. Myers, D. C. Montgomery und C. M. Anderson-Cook. *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*. 4. Aufl. Hoboken, NJ, USA: John Wiley & Sons, Inc, 2016. ISBN: 978-1-118-91601-8.
- [37] M. Dazer. *Zuverlässigkeitstestplanung mit Berücksichtigung von Vorwissen aus stochastischen Lebensdauerberechnungen*. 2019. DOI: [10.18419/OPUS-10501](https://doi.org/10.18419/OPUS-10501).
- [38] T. Herzig. *Anforderungsgerechte Produktauslegung durch Planung effizienter beschleunigter Zuverlässigkeitstests*. 2021. DOI: [10.18419/OPUS-12001](https://doi.org/10.18419/OPUS-12001).
- [39] A. Grundler. *Zuverlässigkeitssnachweis von Systemen durch Komponenten- und Systemtests unter Berücksichtigung von Vorkenntnissen*. 2024. DOI: [10.18419/OPUS-14393](https://doi.org/10.18419/OPUS-14393).

- [40] A. Kremer. „Statistische Versuchsplanung in der Lebensdauererprobung mit Vertrauensintervallen“. Diss. 2021. doi: [10.18419/OPUS-11728](https://doi.org/10.18419/OPUS-11728).
- [41] D. E. Coleman und D. C. Montgomery. „A Systematic Approach to Planning for a Designed Industrial Experiment“. In: *Technometrics* 35.1 (1993), S. 1. issn: 0040-1706. doi: [10.2307/1269280](https://doi.org/10.2307/1269280).
- [42] C. Gundlach. „Entwicklung eines ganzheitlichen Vorgehensmodells zur problemorientierten Anwendung des statistischen Versuchsplanung“. Zugl.: Kassel, Univ., Diss, 2004. Kassel. url: <http://d-nb.info/971843546/34>.
- [43] B. Mayers. *Prozeß- und Produktoptimierung mit Hilfe der statistischen Versuchsmethodik*. Zugl.: Aachen, Techn. Hochsch., Diss., 1997. Als Ms. gedr. Bd. 97,9. Berichte aus der Produktionstechnik. Aachen: Shaker, 1997. isbn: 3826524977.
- [44] G. Pahl, W. Beitz, J. Feldhusen und K.-H. Grote. *Engineering design: A systematic approach*. 3. ed. London: Springer, 2007. isbn: 978-1-84628-318-5. doi: [10.1007/978-1-84628-319-2](https://doi.org/10.1007/978-1-84628-319-2).
- [45] W. F. Daenzer und R. Haberfellner, Hrsg. *Systems Engineering: Methodik und Praxis*. 11., durchges. Aufl. Zürich: Verl. Industrielle Organisation, 2002. isbn: 385743998X.
- [46] U. Lindemann, M. Maurer und T. Braun. *Structural Complexity Management: An Approach for the Field of Product Design*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008. isbn: 978-3-540-87888-9. doi: [10.1007/978-3-540-87889-6](https://doi.org/10.1007/978-3-540-87889-6).
- [47] G. E. P. Box und J. S. Hunter. „Multi-Factor Experimental Designs for Exploring Response Surfaces“. In: *The Annals of Mathematical Statistics* 28.1 (1957), S. 195–241.
- [48] G. E. P. Box und D. W. Behnken. „Some New Three Level Designs for the Study of Quantitative Variables“. In: *Technometrics* 2 (1960), S. 455–475. issn: 0040-1706.
- [49] B. Jones und C. J. Nachtsheim. „Efficient Designs With Minimal Aliasing“. In: *Technometrics* 53.1 (2011), S. 62–71. issn: 0040-1706. doi: [10.1198/TECH.2010.09113](https://doi.org/10.1198/TECH.2010.09113).

- [50] A. M. Dean, D. Voss und D. Draguljić. *Design and analysis of experiments*. Second edition. Springer texts in statistics. Cham: Springer, 2017. ISBN: 978-3-319-52248-7. doi: [10.1007/978-3-319-52250-0](https://doi.org/10.1007/978-3-319-52250-0).
- [51] G. E. P. Box und K. B. Wilson. „On the Experimental Attainment of Optimum Conditions“. In: *Journal of the Royal Statistical Society. Series B (Methodological)* 13.1 (1951), S. 1–45. ISSN: 00359246. URL: <http://www.jstor.org/stable/2983966> (besucht am 23.03.2023).
- [52] G. E. P. Box, S. Bisgaard und C. Fung. „An explanation and critique of taguchi's contributions to quality engineering“. In: *Quality and Reliability Engineering International* 4.2 (1988), S. 123–131. ISSN: 0748-8017. doi: [10.1002/qre.4680040207](https://doi.org/10.1002/qre.4680040207).
- [53] S. Bisgaard. „A Conceptual Framework for the use of Quality Concepts and Statistical Methods in Product Design“. In: *Journal of Engineering Design* 3.1 (1992), S. 31–47. ISSN: 0954-4828. doi: [10.1080/09544829208914746](https://doi.org/10.1080/09544829208914746).
- [54] K. G. Russell. *Design of experiments for generalized linear models*. Chapman and Hall/CRC Interdisciplinary Statistics Ser. Boca Raton, London und New York: CRC Press Taylor & Francis Group, 2019. ISBN: 9780429057489. doi: [10.1201/9780429057489](https://doi.org/10.1201/9780429057489).
- [55] F. Pukelsheim. *Optimal design of experiments: Originally published: New York : J. Wiley, 1993*. Classic ed. Bd. 50. Classics in applied mathematics. Philadelphia, Pa.: Society for Industrial and Applied Mathematics (SIAM 3600 Market Street Floor 6 Philadelphia PA 19104), 2006. ISBN: 9780898716047. doi: [10.1137/1.9780898719109](https://doi.org/10.1137/1.9780898719109).
- [56] M. Risberg Ellekjær und S. Bisgaard. „The use of experimental design in the development of new products“. In: *International Journal of Quality Science* 3.3 (1998), S. 254–274. ISSN: 1359-8538. doi: [10.1108/13598539810229230](https://doi.org/10.1108/13598539810229230).
- [57] W. Kleppmann. *Versuchsplanung: Produkte und Prozesse optimieren*. 9., überarbeitete Auflage. Hanser eLibrary. München: Hanser, 2016. ISBN: 978-3-446-44716-5. doi: [10.3139/9783446447172?locatt=mode:legacy](https://doi.org/10.3139/9783446447172?locatt=mode:legacy).

- [58] R. H. Myers, D. C. Montgomery, G. G. Vining und T. J. Robinson. *Generalized linear models: With applications in engineering and the sciences*. 2. ed. Wiley series in probability and statistics. Hoboken, NJ: Wiley, 2010. ISBN: 9780470454633. doi: [10.1002/9780470556986](https://doi.org/10.1002/9780470556986).
- [59] A. Zahran, C. M. Anderson-Cook und R. H. Myers. „Fraction of Design Space to Assess Prediction Capability of Response Surface Designs“. In: *Journal of Quality Technology* 35.4 (2003), S. 377–386. ISSN: 0022-4065.
- [60] A. Giovannitti-Jensen und R. H. Myers. „Graphical Assessment of the Prediction Capability of Response Surface Designs“. In: *Technometrics* 31.2 (1989), S. 159–171. ISSN: 0040-1706.
- [61] D. A. Belsley. *Regression diagnostics: Identifying influential data and sources of collinearity*. Wiley series in probability and statistics. New York: Wiley, 2004. ISBN: 9780471725145. URL: <https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=226535>.
- [62] L. Fahrmeir, T. Kneib und S. Lang. *Regression: Modelle, Methoden und Anwendungen*. 2. Auflage. Statistik und ihre Anwendungen. Heidelberg und Berlin: Springer, 2009. ISBN: 978-3-642-01836-7. doi: [10.1007/978-3-642-01837-4](https://doi.org/10.1007/978-3-642-01837-4).
- [63] S. P. Boyd und L. Vandenberghe. *Convex optimization*. Cambridge, New York, NY und Port Melbourne: Cambridge University Press, 2004. ISBN: 978-0-521-83378-3. URL: <http://www.loc.gov/catdir/description/cam041/2003063284.html>.
- [64] A. C. Atkinson. „The Design of Experiments to Estimate the Slope of a Response Surface“. In: *Biometrika* 57.2 (1970), S. 319. ISSN: 00063444. doi: [10.2307/2334839](https://doi.org/10.2307/2334839).
- [65] A. C. Atkinson, A. N. Donev und R. D. Tobias. *Optimum Experimental Designs, with SAS*. 1st. New York: Oxford University Press Inc, 2007. ISBN: 9780199296590.
- [66] E. M. Monroe, R. Pan, C. M. Anderson-Cook, D. C. Montgomery und C. M. Borror. „A generalized linear model approach to designing accelerated life test experiments“. In: *Quality and Reliability Engineering International* 27.4 (2011), S. 595–607. ISSN: 0748-8017. doi: [10.1002/qre.1143](https://doi.org/10.1002/qre.1143).

- [67] P. Goos. „OMARS Designs for Factor Screening and Response Surface Experimentation in One Step: A Review“. In: *WIREs Computational Statistics* 17.2 (2025). issn: 1939-5108. doi: [10.1002/wics.70018](https://doi.org/10.1002/wics.70018).
- [68] J. Núñez Ares und P. Goos. „Enumeration and Multicriteria Selection of Orthogonal Minimally Aliased Response Surface Designs“. In: *Technometrics* 62.1 (2020), S. 21–36. issn: 0040-1706. doi: [10.1080/00401706.2018.1549103](https://doi.org/10.1080/00401706.2018.1549103).
- [69] J. Cohen. *Statistical power analysis for the behavioral sciences*. Second edition. Hillsdale, NJ: Lawrence Erlbaum Associates Publishers, 1988. isbn: 9781134742707. doi: [10.4324/9780203771587](https://doi.org/10.4324/9780203771587).
- [70] M. H. Kutner, C. J. Nachtsheim, J. Neter und W. Li. *Applied Linear Statistical Models*. 5. ed. McGraw-Hill/Irwin series Operations and decision sciences. Boston, Mass.: McGraw-Hill Irwin, 2005. isbn: 0072386886. url: <http://www.loc.gov/catdir/bios/mh042/2004052447.html>.
- [71] J. A. Nelder und R. W. M. Wedderburn. „Generalized Linear Models“. In: *Journal of the Royal Statistical Society. Series A (General)* 135.3 (1972), S. 370. issn: 00359238. doi: [10.2307/2344614](https://doi.org/10.2307/2344614).
- [72] C. E. McCulloch und S. R. Searle. *Generalized, linear, and mixed models*. A Wiley-Interscience publication. New York und Weinheim: Wiley, 2001. isbn: 9780471193647. doi: [59430](https://doi.org/10.1002/0471193647).
- [73] M. Arndt, P. Mell und M. Dazer. „Generic effects of deviations from test design orthogonality on test power and regression modelling of Central-Composite Designs“. In: *Proceedings of the 16th International Probabilistic Safety Assessment and Management Conference (PSAM16)*. Bd. PSAM 16. Honolulu, HI, USA, 2022.
- [74] A. Khuri. *Response Surface Methodology and Related Topics*. 1st. World Scientific Publishing Co Pte Ltd, 2006. isbn: 9789812564580.
- [75] G. E. P. Box und D. R. Cox. „An Analysis of Transformations“. In: *Journal of the Royal Statistical Society Series B: Statistical Methodology* 26.2 (1964), S. 211–243. issn: 1369-7412. doi: [10.1111/j.2517-6161.1964.tb00553.x](https://doi.org/10.1111/j.2517-6161.1964.tb00553.x).

- [76] M. V. Wüthrich und M. Merz. *Statistical Foundations of Actuarial Learning and its Applications*. Cham: Springer International Publishing, 2023. ISBN: 978-3-031-12408-2. doi: [10.1007/978-3-031-12409-9](https://doi.org/10.1007/978-3-031-12409-9).
- [77] M. Modarres, M. Nuri-Amiri und C. Jackson. *Probabilistic physics of failure approach to reliability: Modeling, accelerated testing, prognosis and reliability assessment*. Performability engineering series. Beverly, MA und Hoboken, NJ: Scrivener Publishing und John Wiley & Sons Inc, 2017. ISBN: 9781119388630.
- [78] A. Kremer, L. Dürösö und B. Bertsche. „Reliability Prediction using Design of Experiments“. In: *Proceedings of the 30th European Safety and Reliability Conference and 15th Probabilistic Safety Assessment and Management Conference*. Hrsg. von P. Baraldi, F. Di Maio und E. Zio. Singapore: Research Publishing Services, 2020, S. 1501–1507. ISBN: 978-981-14-8593-0. doi: [10.3850/978-981-14-8593-0{\textunderscore}3538-cd](https://doi.org/10.3850/978-981-14-8593-0{\textunderscore}3538-cd).
- [79] A. J. Dobson und A. G. Barnett. *An Introduction to Generalized Linear Models*. Fourth edition. Texts in statistical science. Boca Raton, FL: CRC Press, 2018. ISBN: 9781315182780. doi: [10.1201/9781315182780](https://doi.org/10.1201/9781315182780).
- [80] J. Stufken und M. Yang. „Optimal Designs for Generalized Linear Models“. In: *Design and analysis of experiments*. Hrsg. von K. Hinkelmann. Bd. 3. Wiley series in probability and statistics. Hoboken, NJ: Wiley-Interscience, 2012, S. 137–164. ISBN: 9780470530689. doi: [10.1002/9781118147634.ch4](https://doi.org/10.1002/9781118147634.ch4).
- [81] M. Aitkin und D. Clayton. „The Fitting of Exponential, Weibull and Extreme Value Distributions to Complex Censored Survival Data Using GLIM“. In: *Journal of the Royal Statistical Society Series C: Applied Statistics* 29.2 (1980), S. 156. ISSN: 0035-9254. doi: [10.2307/2986301](https://doi.org/10.2307/2986301).
- [82] A. C. Rencher und G. B. Schaalje. *Linear models in statistics*. 2. ed. Hoboken, NJ: Wiley-Interscience, 2008. ISBN: 9780471754985. URL: <http://www.loc.gov/catdir/enhancements/fy0826/2007024268-b.html>.
- [83] D. Rasch und D. Schott. *Mathematical statistics*. Hoboken, NJ: John Wiley & Sons Inc, 2018. ISBN: 9781119385288.
- [84] H. Ahn. „Central Composite Design for the Experiments with Replicate Runs at Factorial and Axial Points“. In: *Industrial Engineering* 349 (2015), S. 969–979.

- [85] A. N. Donev. „Design of experiments in the presence of errors in factor levels“. In: *Statistical Planning and Inference* 126.2 (2004), S. 569–585.
- [86] M. K. Ardakani, D. Das, S. S. Wulf und T. J. Robinson. „Estimation in Second-Order Models with Errors in the Factor Levels“. In: *Commun. in Statistics* 40.9 (2011), S. 1573–1590. doi: [10.1080/03610921003637421](https://doi.org/10.1080/03610921003637421).
- [87] J.-P. Thommen, A.-K. Achleitner, D. U. Gilbert, D. Hachmeister, S. Jarchow und G. Kaiser. *Allgemeine Betriebswirtschaftslehre Arbeitsbuch: Repetitionsfragen - Aufgaben - Lösungen*. 9., überarbeitete und aktualisierte Auflage. Wiesbaden und Heidelberg: Springer Gabler, 2022. isbn: 978-3-658-31798-0. doi: [10.1007/978-3-658-31799-7](https://doi.org/10.1007/978-3-658-31799-7).
- [88] M. Yang, B. Zhang und S. Huang. „Optimal designs for generalized linear models with multiple design variables“. In: *Statistica Sinica* 21.3 (2011), S. 1415–1430. issn: 10170405. doi: [10.5705/ss.2009.115](https://doi.org/10.5705/ss.2009.115).
- [89] M. Arndt, M. Dazer, W. Raither und B. Bertsche. „Parameter assessment for reliability modeling of machine components using heuristic screening“. In: *Forschung im Ingenieurwesen* 87.4 (2023), S. 1347–1370. issn: 0015-7899. doi: [10.1007/s10010-023-00711-5](https://doi.org/10.1007/s10010-023-00711-5).
- [90] H. Krallmann, A. Bobrik und O. Levina. *Systemanalyse im Unternehmen: Prozessorientierte Methoden der Wirtschaftsinformatik*. 6., überarb. und erw. Aufl. München: Oldenbourg-Verl., 2013. isbn: 978-3-486-72982-5. doi: [10.1524/9783486729825](https://doi.org/10.1524/9783486729825).
- [91] K. Wallace, G. Pahl, W. Beitz, J. Feldhusen, K.-H. Grote und L. T. M. Blessing, Hrsg. *Engineering design: A systematic approach*. 3. ed. London: Springer, 2007. isbn: 1846283183. url: <http://site.ebrary.com/lib/librarytitles/docDetail.action?docID=10230457>.
- [92] D. C. Montgomery. *Introduction to statistical quality control*. Eighth edition, EMEA edition. Hoboken, NJ: Wiley, 2020. isbn: 9781119657118.
- [93] C. Brückner. *Qualitätsmanagement: Das Praxishandbuch für die Automobilindustrie*. 2., vollständig überarbeitete und erweiterte Auflage. München: Hanser, 2019. isbn: 978-3-446-45575-7. doi: [10.3139/9783446459724](https://doi.org/10.3139/9783446459724).

- [94] A. Breiing und R. Knosala. *Bewerten technischer Systeme: Theoretische und methodische Grundlagen bewertungstechnischer Entscheidungshilfen*. Berlin und Heidelberg: Springer, 1997. ISBN: 978-3-642-63908-1. doi: [10.1007/978-3-642-59229-4](https://doi.org/10.1007/978-3-642-59229-4).
- [95] S. Albers. *Handbuch Technologie- und Innovationsmanagement: Strategie — Umsetzung — Controlling*. SpringerLink Bücher. Wiesbaden und s.l.: Gabler Verlag, 2005. ISBN: 978-3-322-90787-5. doi: [10.1007/978-3-322-90786-8](https://doi.org/10.1007/978-3-322-90786-8).
- [96] G. Thompson. *Improving maintainability and reliability through design*. London: Professional Engineering Publ, 1999. ISBN: 1860581358.
- [97] A. Kremer. „Ermittlung von lebensdauerbeeinflussenden Faktoren bei Zahnrämengetrieben“. In: *antriebstechnik* 2018.11 (11), S. 148–153. URL: <https://digital.antriebstechnik.de/antriebstechnik-11-2018/62195350>.
- [98] T. Buzan und B. Buzan. *The mind map book: Unlock your creativity, boost your memory, change your life*. 1. publ. Harlow u.a.: Pearson, BBC Active, 2010. ISBN: 9781406647167. URL: <https://permalink.obvsg.at/AC11808619>.
- [99] P. Tittmann. *Graphentheorie: Eine anwendungsorientierte Einführung : mit 115 Bildern, zahlreichen Beispielen und 92 Aufgaben*. 3., aktualisierte Auflage. Mathematik-Studienhilfen. München: Hanser, 2019. ISBN: 978-3-446-46052-2. doi: [10.3139/9783446465039](https://doi.org/10.3139/9783446465039).
- [100] B. Klein. *Versuchsplanung - DoE: Einführung in die Taguchi/Shainin-Methodik*. 3rd ed. München: De Gruyter, 2011. ISBN: 978-3-486-71136-3. doi: [10.1524/9783486711363](https://doi.org/10.1524/9783486711363).
- [101] B. Klein. *Versuchsplanung - DoE: Einführung in die Taguchi/Shainin-Methodik*. 4. Auflage. Studium. München: De Gruyter Oldenbourg, 2014. ISBN: 978-3-486-77842-7. doi: [10.1524/9783110343847](https://doi.org/10.1524/9783110343847).
- [102] G. E. P. Box. „The Effects of Errors in the Factor Levels and Experimental Design“. In: *Technometrics* 5.2 (1963), S. 247–262. ISSN: 0040-1706. URL: <https://doi.org/10.2307/1266066>.
- [103] M.-J. Zhang und R.-X. Yue. „Optimal designs for homoscedastic functional polynomial measurement error models“. In: *AStA Advances in Statistical Analysis* 105.3 (2021), S. 485–501. ISSN: 1863-8171. doi: [10.1007/s10182-021-00399-4](https://doi.org/10.1007/s10182-021-00399-4).

- [104] D.-H. Jang. „Measures for Evaluating Non-Orthogonality of Experimental Designs“. In: *Communications in Statistics - Theory and Methods* 31.2 (2002), S. 249–260. ISSN: 0361-0926. doi: [10.1081/STA-120002649](https://doi.org/10.1081/STA-120002649).
- [105] M. Arndt und M. Dazer. *Power and Accuracy Characterization of Regression Models in The Presence of Orthogonality Deviations in CCDs*. In: *6th International Conference on Design of Experiments (ICDOE 2022)*. 8.-11.05.2023. URL: <https://www.memphis.edu/msci/icodoe22/index.php>.
- [106] M. Arndt, P. Mell und M. Dazer. „Generic effects of deviations from test design orthogonality on test power and regression modelling of Central-Composite Designs“. In: *Probabilistic Safety Assessment and Management PSAM 16, June 26-July 1, 2022, Honolulu, Hawaii*. URL: <https://www.iapsam.org/PSAM16/papers/MA19-PSAM16.pdf>.

Anhang A Ableitungen zur Maximum-Likelihood-Schätzung

A.1 Modell Daefinition (Modell zweiter Ordnung)

Für ein System mit $k = 2$ Einflussfaktoren (x_1, x_2) basierend auf einem **CCD** ist der lineare Prädiktor η_i definiert. Er beschreibt den Logarithmus der charakteristischen Lebensdauer wie folgt:

$$\eta_i = \ln(T_i) = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \beta_{12} x_{i,1} x_{i,2} + \beta_{11} x_{i,1}^2 + \beta_{22} x_{i,2}^2 \quad (1)$$

Der zugehörige Parametervektor β und der Regressorvektor x_i lauten:

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_{12} \\ \beta_{11} \\ \beta_{22} \end{bmatrix}, \quad x_i = \begin{bmatrix} 1 \\ x_{i,1} \\ x_{i,2} \\ x_{i,1} x_{i,2} \\ x_{i,1}^2 \\ x_{i,2}^2 \end{bmatrix} \quad (2)$$

A.2 Log-Likelihood-Funktion

Wir führen zunächst den Indikator δ_i ein ($\delta_i = 1$ für Ausfall, 0 für Zensierung). Zusätzlich definieren wir die Hilfsvariable z_i :

$$z_i = t_i^b \exp(-b\eta_i) \quad (3)$$

Damit lautet die Log-Likelihood-Funktion Λ :

$$\Lambda = \sum_{i=1}^n [\delta_i (\ln(b) - b\eta_i + (b-1) \ln(t_i)) - z_i] \quad (4)$$

A.3 Score-Vektor (Erste Ableitungen)

Die partiellen Ableitungen (Gradient) dienen zur Maximierung der Funktion. Die Ableitung nach den Regressionskoeffizienten für $j \in \{0, 1, 2, 12, 11, 22\}$ lautet:

$$\frac{\partial \Lambda}{\partial \beta_j} = b \sum_{i=1}^n (z_i - \delta_i) x_{i,j} \quad (5)$$

Die Ableitung nach dem Formparameter b ergibt sich zu:

$$\frac{\partial \Lambda}{\partial b} = \sum_{i=1}^n \left[\frac{\delta_i}{b} + (\delta_i - z_i)(\ln(t_i) - \eta_i) \right] \quad (6)$$

A.4 Hesse-Matrix (Zweite Ableitungen)

Die Elemente der Hesse-Matrix werden für die Fisher-Information und den Newton-Raphson-Algorithmus benötigt. Die zweite Ableitung nach den Koeffizienten β (für Indizes j und l) ist:

$$\frac{\partial^2 \Lambda}{\partial \beta_j \partial \beta_l} = -b^2 \sum_{i=1}^n z_i x_{i,j} x_{i,l} \quad (7)$$

Die zweite Ableitung nach dem Formparameter b lautet:

$$\frac{\partial^2 \Lambda}{\partial b^2} = - \sum_{i=1}^n \left[\frac{\delta_i}{b^2} + z_i (\ln(t_i) - \eta_i)^2 \right] \quad (8)$$

Die gemischte Ableitung nach β und b berechnet sich wie folgt:

$$\frac{\partial^2 \Lambda}{\partial \beta_j \partial b} = \sum_{i=1}^n x_{i,j} [(z_i - \delta_i) + bz_i (\ln(t_i) - \eta_i)] \quad (9)$$