

# PDE-restringierte Optimierung in Anwendungen der spanenden Trockenbearbeitung

---

**Dissertation**

zur Erlangung des Grades eines Doktors der Ingenieurwissenschaften  
- Dr.-Ing. -  
des Fachbereichs für Mathematik und Informatik an der Universität Bremen  
von

Heinrich Wernsing

---

1. Gutachter: Prof. Dr. Christof Büskens
2. Gutachter: Dr.-Ing. Jens Sölter

Tag der mündlichen Prüfung: 05. Juli 2018



# Danksagung

Diese Arbeit ist während meiner Tätigkeit am Zentrum für Technomathematik der Universität Bremen entstanden.

Mein besonderer Dank gilt meinem Doktorvater Prof. Dr. Christof Büskens für seine fachliche und persönliche Betreuung. Er verstand es stets, mir neue Denkanstöße und Motivation in hilfreichen Gesprächen und Diskussionen zu geben. Ich danke ihm für sein Vertrauen in meine Arbeit und die lehrreiche Unterstützung.

Mein Dank gilt außerdem den Projektpartnern im Schwerpunktprogramm, Dr.-Ing. Jens Sölter und Maxim Gulpak, für die angenehme und kompetente Zusammenarbeit. Darüber hinaus danke ich Dr.-Ing. Jens Sölter für die Begutachtung dieser Arbeit und Maxim Gulpak für die hilfreichen Korrekturen.

Ich danke der gesamten Arbeitsgruppe Optimierung und Optimale Steuerung für das freundschaftliche Arbeitsklima und stellvertretend Friedrich Kohlmai und Dr. Matthias Knauer für die Anregungen in unzähligen Gesprächen.

Weiterhin bedanke ich mich bei allen, die mir mit ihren Korrekturen zu dieser Arbeit geholfen haben, insbesondere Tobias Linke und Friedrich Kohlmai.

Eine der wichtigsten Grundlagen für diese Dissertation war die Unterstützung durch mein privates Umfeld. Ich bedanke mich aus tiefstem Herzen bei der Liebe meines Lebens Katrin sowie meinen Eltern Michaela und Heinz und meinem Bruder Jan. Danke, dass ihr immer an mich glaubt und hinter mir steht.



# Inhaltsverzeichnis

Abkürzungsverzeichnis . . . . .	v
Symbolverzeichnis . . . . .	v
Abbildungsverzeichnis . . . . .	xii
Tabellenverzeichnis . . . . .	xiii
<b>1 Einleitung</b>	<b>1</b>
1.1 Zielsetzung . . . . .	1
1.2 Aufbau . . . . .	3
1.3 Einordnung des wissenschaftlichen Beitrags . . . . .	4
<b>2 Grundlagen der Optimierung mit PDE-Nebenbedingungen</b>	<b>9</b>
2.1 Lösungskonzepte der PDE-restringierten Optimierung . . . . .	10
2.2 Grundlagen der nichtlinearen Optimierung . . . . .	15
2.2.1 Theorie der nichtlinearen Optimierung . . . . .	15
2.2.2 Algorithmen der nichtlinearen Optimierung . . . . .	21
2.3 Finite-Elemente-Methode und Softwarebibliothek deal.II . . . . .	31
<b>3 Spanende Trockenbearbeitung - Modellierung und Optimierungskonzept</b>	<b>35</b>
3.1 Modellierung der Prozesseinflüsse beim Trockenfräsen . . . . .	36
3.1.1 Eigenspannungsbedingte Effekte - Sub-Modell 1 . . . . .	37
3.1.2 Thermische Effekte - Sub-Modell 2 . . . . .	42
3.2 Vorgehensweise zur Minimierung von Formabweichungen beim Fräsen . .	45
3.3 Modellierung der Prozesseinflüsse beim Bohren . . . . .	47
3.4 Versuchsreihen für die empirischen Modelle . . . . .	48
3.4.1 Experimente für Sub-Modell 1 . . . . .	48
3.4.2 Experimente für Sub-Modell 2 . . . . .	49
3.4.3 Experimente für das thermische Modell des Bohrens . . . . .	49
<b>4 Parameteridentifikation von Fräsprozessen</b>	<b>51</b>
4.1 Parameteridentifikation zur Bestimmung der Wärmestromdichte . . . . .	52
4.2 Allgemeines Wärmeleitungsmodell . . . . .	53
4.3 Problemstellung beim Fräsen . . . . .	58
4.4 Wärmequelle, Diskretisierung und Temperaturdaten . . . . .	61
4.5 Numerische Ergebnisse . . . . .	65
4.5.1 Validierung . . . . .	65
4.5.2 Identifikation von Realdaten . . . . .	69

4.6	Fazit der Parameteridentifikation des Fräsens . . . . .	77
<b>5</b>	<b>Parameteridentifikation von Bohrprozessen</b>	<b>79</b>
5.1	Wärmeleitungsmodell des Bohrens . . . . .	79
5.1.1	SAND-Formulierung . . . . .	81
5.1.2	NAND-Formulierung . . . . .	83
5.2	Randbedingungen, Diskretisierung und Temperaturdaten . . . . .	84
5.3	Numerische Ergebnisse . . . . .	88
5.3.1	Validierung . . . . .	89
5.3.2	Rechenzeitvergleich . . . . .	95
5.3.3	Identifikation von Realdaten . . . . .	100
5.4	Fazit der Parameteridentifikation des Bohrens . . . . .	112
<b>6</b>	<b>Minimierung von Formabweichungen beim Trockenfräsen</b>	<b>113</b>
6.1	Optimierung von quellspannungsbedingten Deformationen . . . . .	114
6.1.1	Quellspannungsmodell . . . . .	115
6.1.2	Optimierungskonzept . . . . .	118
6.1.3	Mathematische Formulierung der Optimierungsprobleme . . . . .	120
6.1.4	Diskretisierung und Modellvalidierung . . . . .	124
6.1.5	Numerische Ergebnisse . . . . .	126
6.2	Optimierung der Gesamtformabweichung . . . . .	133
6.2.1	Kinematisches Modell der Gesamtformabweichungen . . . . .	133
6.2.2	Mathematische Formulierung der Optimierungsprobleme . . . . .	135
6.2.3	Numerische Ergebnisse . . . . .	136
6.2.4	Vergleich zu Messungen realer Experimente . . . . .	140
6.3	Fazit der Minimierung von Formabweichungen . . . . .	141
<b>7</b>	<b>Zusammenfassung und Ausblick</b>	<b>145</b>
7.1	Zusammenfassung . . . . .	145
7.2	Ausblick . . . . .	147
	<b>Literaturverzeichnis</b>	<b>149</b>

## Abkürzungsverzeichnis

BFGS	Broyden-Fletcher-Goldfarb-Shanno
DFG	Deutsche Forschungsgemeinschaft
ESA	Europäische Weltraumorganisation (engl. European Space Agency)
FEM	Finite-Elemente-Methode
IP	Innere-Punkte
Ipopt	Interior Point OPTimizer - Optimierungssoftware
KKT	Karush-Kuhn-Tucker
KM	kinematisches Modell des Materialabtrags
LICQ	Regularitätsbedingungen der linearen Unabhängigkeit (engl. linear independence constraint qualification)
MFCQ	Regularitätsbedingungen von Mangasarian-Fromovitz (engl. Mangasarian-Fromovitz constraint qualification)
NAND	Nested Analysis and Design
NLP	nichtlineares Optimierungsproblem (engl. nonlinear program)
PDE	Partielle Differentialgleichung (engl. Partial Differential Equation)
QP	quadratisches Programm
SAND	Simultaneous Analysis and Design
SPP	Schwerpunktprogramm der DFG
SQP	Sequentielle Quadratische Programmierung
WORHP	We Optimize Really Huge Problems - Optimierungssoftware
ZB	Zustandsbeschränkungen

## Symbolverzeichnis

### Lateinische Symbole

$a_{c_p}, b_{c_p}$	Koeffizienten der temperaturabhängigen spezifischen Wärmekapazität $c_p$
$a_e$	Eingriffsbreite (in mm)
$a_i, b_i, c_i, d_i,$ $e_i, f_i, g_i, h_i$	Koeffizienten der Regressionsmodelle des Fräsens
$a_k, b_k$	Koeffizienten der temperaturabhängigen Wärmeleitfähigkeit $k$
$a_p$	Schnitttiefe (in mm)
$a_{p,eff}$	effektive Schnitttiefe (in mm)
$\mathbf{A}$	gesamte Systemmatrix des linearen Gleichungssystems der FEM
$\hat{\mathbf{A}}, \hat{\mathbf{B}}$	Hilfsmatrizen für die Hessematrix der Lagrange-Funktion
$\mathbf{A}^l, \mathbf{A}^k$	Systemmatrizen für den aktuellen Zustand der Gleichungssysteme beim Fräsen und Bohren

## Symbolverzeichnis

$\tilde{\mathbf{A}}^l, \tilde{\mathbf{B}}^l$	Hilfsmatrizen für die Jacobi-Matrix der Nebenbedingungen
$A_{\text{prim}}$	Fläche der primären Wärmequelle beim Bohren (in $\text{mm}^2$ )
$A_{\text{sek}}$	Fläche der sekundären Wärmequellen beim Bohren (in $\text{mm}^2$ )
$\mathbf{b}, \tilde{\mathbf{b}}$	gesamte rechte Seite des linearen Gleichungssystems der FEM
$B$	Barriere-Funktion
$\hat{B}$	Barriere-Term
$\hat{B}_{\log}$	logarithmischer Barriere-Term
$\mathbf{B}^l, \mathbf{B}^k$	Systemmatrizen für den alten Zustand der Gleichungssysteme beim Fräsen und Bohren
$c_p$	spezifische Wärmekapazität (in $\text{J}/(\text{kg K})$ )
$\mathbf{C}^l$	Systemmatrizen für die Identifikationsparameter des Gleichungssystems zum Zeitpunkt $t_l$ beim Fräsen und Bohren
$C^\infty(\Omega)$	Raum der beliebig oft differenzierbaren Funktionen
$C_0^\infty(\Omega)$	Unterraum von $C^\infty$ mit kompakter Teilmenge ungleich 0
$e$	Einsvektor
$E$	Elastizitätsmodul (in $\text{N}/\text{mm}^2$ )
$\mathbb{E}$	Elastizitätstensor (in $\text{N}/\text{mm}^2$ )
$f$	Funktion des Quellterms der Poisson-Gleichung
$f_G$	allgemeines Kraftfeld
$f_{\text{prim}}$	Höhe der primären Wärmequelle beim Bohren (in mm)
$f_{\text{sek}}$	Höhe der sekundären Wärmequellen beim Bohren (in mm)
$f_V$	Volumenkraft
$f_z$	Zahnvorschub (in mm)
$F$	Zielfunktion
$G$	allgemeine Ungleichungsnebenbedingungen
$\mathbf{G}$	Diagonalmatrix der Ungleichungsnebenbedingungen
$H$	allgemeine Gleichungsnebenbedingungen
$\mathcal{H}$	Hesse-Matrix oder deren Approximation
$H_0^1(\Omega)$	Unterraum von $H^1$ mit Nullrandwert-Funktionen
$H^m(\Omega)$	Raum der $m$ -mal schwach-differenzierbaren Funktionen
$\mathbf{I}$	Einheitsmatrix
$I_A$	Indexmenge der aktiven Ungleichungsnebenbedingungen
$I_A^0$	Teilmenge der Zerlegung von $I_A$ mit Multiplikatoren gleich 0
$I_A^>$	Teilmenge der Zerlegung von $I_A$ mit Multiplikatoren größer 0
$k$	Wärmeleitfähigkeit (in $\text{W}/(\text{m K})$ )
$K$	Plattensteifigkeit (in $\text{N mm}$ )
$K(X)$	Kritischer Kegel von $X$
$l_x$	Plattenlänge in $x$ -Richtung (in mm)
$l_y$	Plattenlänge in $y$ -Richtung (in mm)
$l_z$	Plattenlänge in $z$ -Richtung (in mm)
$L$	Lagrange-Funktion



$L_2(\Omega)$	Raum der quadratisch Lebesgue-integrierbaren Funktionen
$m$	Anzahl aller Nebenbedingungen
$m_g$	Anzahl der Ungleichungsnebenbedingungen
$m_h$	Anzahl der Gleichungsnebenbedingungen
$M_x$	Biegemoment entlang der $x$ -Achse
$M_y$	Biegemoment entlang der $y$ -Achse
$M_{xy}$	Biegemoment in der $xy$ -Ebene
$n$	Spindeldrehzahl
$\vec{n}$	Äußerer Normalenvektor
$n_0^l$	Anzahl der diskreten Zustände im Gebiet $\Gamma_0^l$
$n_b$	Anzahl der Fräsbahnen
$n_d$	Dimension des Raumes
$n_h$	Dimension des Raumes der Finite-Elemente-Ansatzfunktionen
$n_{\text{mess}}$	Anzahl der Messpunkte
$n_p$	Dimension des kontinuierlichen Modells
$n_s$	Anzahl der Zustände eines Modells bzw. Freiheitsgrade der FEM
$n_s^l$	Anzahl der diskreten Zustände zum Zeitpunkt $t_l$
$n_t$	Anzahl der diskreten Zeitpunkte
$n_x$	Anzahl der Optimierungsvariablen
$n_z$	Anzahl der Design-, Steuer- und Identifikationsparameter eines Optimierungsproblems
$n_\Gamma^l$	Anzahl der diskreten Zustände auf dem Rand zum Zeitpunkt $t_l$
$n_\sigma$	Anzahl der diskretisierten Quellspannungswerte
$n_{+,i}^l$	Anzahl der diskreten Zustände im Gebiet $\Gamma_{+,i}^l$
$p$	Multiindex
$P$	Modellgleichungen
$\dot{q}$	Wärmestromdichte (in $\text{W}/\text{mm}^2$ )
$\dot{q}_{\text{ident}}$	zu identifizierende Wärmestromdichte (in $\text{W}/\text{mm}^2$ )
$\dot{q}_{\text{reg},i}$	Wärmestromdichte des $i$ -ten Regressionsmodells beim Fräsen (in $\text{W}/\text{mm}^2$ )
$\dot{q}_{\text{valid}}$	fixierte Wärmestromdichte der Validierung (in $\text{W}/\text{mm}^2$ )
$\dot{Q}_{\text{prim}}$	Wärmeleistung der primären Wärmequelle beim Bohren (in W)
$\dot{Q}_{\text{sek}}$	Summe der Wärmeleistungen der sekundären Wärmequellen beim Bohren (in W)
$r$	Radius des Bohrers (in mm)
$S$	Zustandsvektor eines Modells
$\mathcal{S}$	Abb( $\mathbb{R}^{n_d}, \mathbb{R}^{n_s}$ ), Raum der Abbildungen von $\mathbb{R}^{n_d}$ nach $\mathbb{R}^{n_s}$
$S_{\text{mess}}$	gemessener Zustand
$t$	Zeitvariable (in s)
$t_0$	Startzeitpunkt (in s)
$t_f$	Endzeitpunkt (in s)

## Symbolverzeichnis

$t_l$	$l$ -ter Zeitpunkt (in s)
$T$	Temperaturverteilung (in °C)
$T_0$	Anfangstemperaturverteilung (in °C)
$T_{\text{mess}}$	gemessenes Temperaturfeld (in °C)
$u$	Verschiebungen (in $\mu\text{m}$ ) oder Lösung einer PDE
$u_{\text{ges}}$	Gesamtformabweichung (in $\mu\text{m}$ )
$u_{\text{ges}}^{\text{max}}$	absolutes Maximum der Gesamtformabweichungen (in $\mu\text{m}$ )
$u_{\text{quell}}$	quellspannungsbedingte Verschiebungen (in $\mu\text{m}$ )
$u_{\text{quell}}^{\text{max}}$	absolutes Maximum der quellspannungsbedingten Verschiebungen (in $\mu\text{m}$ )
$u_{\text{therm}}$	thermisch bedingte Verschiebungen (in $\mu\text{m}$ )
$U_\varepsilon$	Epsilon-Umgebung
$v$	Testfunktion
$v_c$	Schnittgeschwindigkeit (in m/min)
$v_f$	Vorschubgeschwindigkeit (in mm/min)
$V_h$	Funktionsraum der Finite-Elemente-Ansatzfunktionen
$w_i$	Gewicht zum $i$ -ten Zustand
$\bar{w}_i$	normiertes Gewicht zum $i$ -ten Zustand
$x, y, z$	Raumkoordinaten (in mm)
$X$	gesamter Optimierungsvektor
$y_{\text{pos}}$	$y$ -Position des heißesten Datenpunktes (in mm)/Optimierungsvariable
$z_0$	Tiefe der Wirkung der Quellspannungen (in mm)
$Z$	Design-, Steuer- oder Identifikationsparameter/Entscheidungsvariable
$\mathcal{Z}$	$\text{Abb}(\mathbb{R}^{n_d}, \mathbb{R}^{n_z})$ , Raum der Abbildungen von $\mathbb{R}^{n_d}$ nach $\mathbb{R}^{n_z}$

## Griechische Symbole

$\alpha$	Bahnwinkel beim Fräsen (in °)
$\alpha_i$	Koeffizient
$\Gamma$	Rand von Gebiet $\Omega$
$\Gamma_0^l$	Randbereich mit adiabaten Bedingungen zum Zeitpunkt $t_l$
$\Gamma_O$	zu bearbeitende Oberfläche des Bauteils
$\Gamma_{O,i}$	von der $i$ -ten Fräsbahn bearbeitete Oberfläche
$\Gamma_U$	Unterseite des Bauteilrandes
$\Gamma_+^l$	Randbereich mit positivem Wärmeeintrag zum Zeitpunkt $t_l$
$\Gamma_{+,i}^l$	Randbereich mit positivem Wärmeeintrag $\dot{q}_{\text{ident},i}$ zum Zeitpunkt $t_l$
$\Delta t$	Zeitschrittweite (in s)
$\vartheta_i$	Zuordnungsfunktion der gemessenen Zustände zum $i$ -ten modellierten Zustand
$\theta$	Faktor des Theta-Verfahrens
$\kappa_x, \kappa_y, \kappa_{xy}$	Krümmungsterme
$\Lambda$	Diagonalmatrix der dualen Variablen

$\lambda$	Lagrange-Multiplikator/duale Variable
$\lambda_G$	Lagrange-Multiplikator/duale Variable für $G$
$\lambda_H$	Lagrange-Multiplikator/duale Variable für $H$
$\lambda_L$	erste Lamé-Konstante (in N/mm <sup>2</sup> )
$\lambda_\mu$	Barriere-Multiplikator
$\mu$	Barriere-Parameter
$\mu_L$	zweite Lamé-Konstante (Schubmodul) (in N/mm <sup>2</sup> )
$\mu_N$	Erwartungswert der Normalverteilung
$\nu$	Poissonzahl
$\rho$	Dichte (in kg/m <sup>3</sup> )
$\Sigma$	zulässige Menge eines Optimierungsproblems
$\boldsymbol{\sigma}$	Quellspannungstensor (in N/mm <sup>2</sup> )
$\sigma$	Quellspannung (Normalspannung) (in N/mm <sup>2</sup> )
$\hat{\sigma}$	transformierte Quellspannung (Normalspannung) (in N/mm <sup>2</sup> )
$\sigma_N$	Standardabweichung der Normalverteilung
$\boldsymbol{\sigma}_{\text{ges}}$	Tensor der Gesamtspannung (in N/mm <sup>2</sup> )
$\tau$	Quellspannung (Torsionsspannung) (in N/mm <sup>2</sup> )
$\hat{\tau}$	transformierte Quellspannung (Torsionsspannung) (in N/mm <sup>2</sup> )
$\varphi_i$	$i$ -te Basisfunktion
$\Phi$	KKT-Gleichungssystem
$\Phi^\mu$	Primal-Dual-Gleichungssystem
$\Omega$	Gebiet, auf dem die PDE definiert ist

### Operatoren

$\text{div}(\cdot)$	Divergenz-Operator
$\circ$	elementweise Multiplikation
$\nabla(\cdot)$	Gradient oder Jacobi-Matrix bzgl. aller Variablen
$\nabla^2(\cdot)$	Hesse-Matrix bzgl. aller Variablen
$\nabla_X(\cdot)$	Gradient oder Jacobi-Matrix bzgl. $X$
$\nabla_X^2(\cdot)$	Hesse-Matrix bzgl. $X$
$\ \cdot\ $	beliebige Norm
$\partial^p$	schwacher Ableitungsoperator mit Multiindex $p$

### Notationen

$(\cdot)_h$	endlich-dimensionale Approximation
$(\cdot)^{(k)}$	Iterierte $k$ einer Größe
$(\cdot)^l$	zeitdiskrete Größe zum Zeitpunkt $t_l$
$(\cdot)_L$	untere Schranke der entsprechenden Variable/Funktion
$(\cdot)^T$	Transponiertes eines Vektors/einer Matrix
$(\cdot)_U$	obere Schranke der entsprechenden Variable/Funktion

## Symbolverzeichnis

$(\cdot)_x$	$x$ -Komponente der entsprechenden Größe
$(\cdot)_y$	$y$ -Komponente der entsprechenden Größe
$(\cdot)_z$	$z$ -Komponente der entsprechenden Größe
$(\cdot)_\phi$	kontinuierliche Funktion der entsprechenden Variable
$(\cdot)^*$	optimierte Variable, Optimierungsergebnis

# Abbildungsverzeichnis

1.1	Einsatz von Kühlschmierstoff in der spanenden Fertigung . . . . .	2
2.1	Algorithmischer Unterschied zwischen SAND und NAND . . . . .	13
2.2	Vorteile von SAND gegenüber NAND . . . . .	13
3.1	Demonstratorbauteil Linearführungsschlitten . . . . .	36
3.2	Bauteil beschrieben durch Gebiet $\Omega$ . . . . .	38
3.3	Funktionsweise des Sub-Modell 1 . . . . .	42
3.4	Bestimmung des Wärmeeintrags . . . . .	43
3.5	Funktionsweise des Sub-Modell 2 . . . . .	45
3.6	Vorgehensweise zur Minimierung der Gesamtformabweichung . . . . .	46
4.1	Geometrie des Eingriffsbogens bzw. der Wärmequelle . . . . .	61
4.2	Temperaturverteilung nach einer Simulation von 50 Schnitten mit sichtbarem Eingriffsbogen . . . . .	62
4.3	Gitter der FEM-Diskretisierung mit Adaption des Eingriffsbogens . . . . .	62
4.4	Thermografieaufnahme eines Fräsexperiments . . . . .	64
4.5	Temperaturfelder des ersten Validierungsszenarios . . . . .	67
4.6	Temperaturfelder des zweiten Validierungsszenarios . . . . .	68
4.7	Temperaturfelder der ersten Identifikation von Realdaten . . . . .	71
4.8	Temperaturfelder der zweiten Identifikation von Realdaten . . . . .	73
4.9	Temperaturfelder der dritten Identifikation von Realdaten . . . . .	75
4.10	Temperaturfelder der vierten Identifikation von Realdaten . . . . .	77
5.1	Gitter der FEM-Diskretisierung . . . . .	85
5.2	Randbereiche der Wärmequellen und der adiabaten Randwerte beim Bohren zum Endzeitpunkt . . . . .	85
5.3	Temperaturverteilung nach einer Simulation des Wärmeeintrags beim Bohren mit fiktiven Wärmestromdichten . . . . .	87
5.4	Thermografieaufnahme eines Bohrexperiments . . . . .	88
5.5	Ergebnis der ersten Validierung mit SAND . . . . .	91
5.6	Vergleich der Ergebnisse von SAND und NAND . . . . .	92
5.7	Temperaturdaten des zweiten Validierungstests . . . . .	93
5.8	Ergebnis der zweiten Validierung mit SAND . . . . .	94
5.9	Über die Berechnungen mit verschiedenen Daten gemittelte Rechenzeiten der Ansätze SAND und NAND für verschiedene Optimierungsaufgaben . . . . .	98
5.10	Rechenzeiten der einzelnen Ansätze im Detail mit linearer Achsenskalierung . . . . .	99

5.11 Probleme bei der Identifikation von Realdaten . . . . .	101
5.12 Ergebnisse der Identifikation des ersten Experiments . . . . .	104
5.13 Identifizierte Wärmestromdichten des ersten Experiments . . . . .	105
5.14 Ergebnisse der Identifikation des zweiten Experiments . . . . .	106
5.15 Identifizierte Wärmestromdichten des zweiten Experiments . . . . .	107
5.16 Ergebnisse der Identifikation des dritten Experiments . . . . .	108
5.17 Identifizierte Wärmestromdichten des dritten Experiments . . . . .	109
5.18 Ergebnisse der Identifikation des vierten Experiments . . . . .	110
5.19 Identifizierte Wärmestromdichten des vierten Experiments . . . . .	111
6.1 Bearbeitungsstrategie mit diskreten Größen (rot) und den kontinuierlichen Variablen (grün) der Prozessstrategien . . . . .	119
6.2 Die drei Grundtypen der berücksichtigten Bearbeitungsstrategien . . . . .	119
6.3 Konvergenz der Quellspannungen bei Gitterverfeinerung . . . . .	125
6.4 Gitter für die Minimierung der quellspannungsbedingten Formabweichungen	126
6.5 Vergleichsstrategie mit resultierender Deformation . . . . .	127
6.6 Ergebnis der Optimierung von Typ <b>A</b> mit 16 Bahnen . . . . .	128
6.7 Vergrößerung der Eingriffsbreiten durch Vergrößerung des Bahnwinkels .	129
6.8 Ergebnis der Optimierung von Typ <b>B</b> mit 40 Bahnen . . . . .	130
6.9 Ergebnis der Optimierung von Typ <b>C</b> mit 24 Bahnen . . . . .	132
6.10 Ermittlung der Gesamtformabweichung in Abhängigkeit von der Schnitttiefe	134
6.11 Ergebnis der Optimierung von Typ <b>A</b> mit 16 Bahnen . . . . .	137
6.12 Ergebnis der Optimierung von Typ <b>B</b> mit 24 Bahnen . . . . .	138
6.13 Ergebnis der Optimierung von Typ <b>C</b> mit 24 Bahnen . . . . .	139
6.14 Vergleich zwischen simulierter und gemessener Formänderung . . . . .	142

# Tabellenverzeichnis

3.1	Materialparameter von 42CrMo4 für die Berechnung von Quellspannungen	39
3.2	Koeffizienten der Regressionsmodelle (3.4) des Sub-Modell 1 . . . . .	40
3.3	Koeffizienten der Regressionsmodelle (3.7) des Sub-Modell 2 . . . . .	44
4.1	Materialparameter von 42CrMo4 für die Wärmeleitungsgleichung . . . . .	63
4.2	Hardware des Rechners . . . . .	65
4.3	Verwendete Maschinenparameter der thermischen Experimente des FräSENS	70
5.1	Temperaturabhängige Materialparameter von 42CrMo4 für die Wärme- leitungsgleichung des Bohrens . . . . .	87
5.2	Hardware des Rechenservers . . . . .	88
5.3	Verwendete Maschinenparameter der thermischen Experimente des Bohrens	102
6.1	Zielfunktionswert und maximale Abweichung nach Optimierung aller neun Strategien . . . . .	133
6.2	Ergebnisübersicht der Minimierung der Gesamtformabweichung . . . . .	140





# Kapitel 1

## Einleitung

### Inhalt

---

1.1	Zielsetzung . . . . .	1
1.2	Aufbau . . . . .	3
1.3	Einordnung des wissenschaftlichen Beitrags . . . . .	4

---

Diese Arbeit ist im Rahmen des Schwerpunktprogramms (SPP) 1480 „Modellierung, Simulation und Kompensation von thermischen Bearbeitungseinflüssen für komplexe Zerspanprozesse“ [13,38] der Deutschen Forschungsgemeinschaft (DFG) entstanden. Innerhalb des SPPs wurde das Projekt „Kompensationsplanung thermischer Prozesseinflüsse beim Trockenfräsen und Trockenbohren“ unter der Leitung von Prof. Dr. Christof Büskens und Dr.-Ing. Jens Sölter bearbeitet. Bearbeitender Projektpartner war Maxim Gulpak. Diese Arbeit konzentriert sich auf die mathematischen Methoden und Ergebnisse des Projektes. Dennoch werden zum besseren Verständnis und zur besseren Einordnung der Ergebnisse in den Gesamtkontext auch ingenieurwissenschaftliche Inhalte vermittelt. Diese Ausführungen basieren größtenteils auf den Arbeiten von Herrn Dr.-Ing. Sölter und Herrn Gulpak. An entsprechender Stelle wird darauf hingewiesen.

### 1.1 Zielsetzung

In der heutigen Zeit ist der Einsatz von spanenden Fertigungsverfahren äußerst vielfältig. Sei es bei der Anfertigung von Zahnersatz, der Produktion von Smartphones, der Herstellung von Küchenarmaturen oder der Fertigung von meterlangen Flugzeugkomponenten - überall kommt die maschinelle Bearbeitung zum Einsatz. Durch die vielseitige Verbreitung der Verfahren steigen auch die Anforderungen an die Präzision. Es ist üblich, in der maschinellen Fertigung Kühlschmierstoffe einzusetzen (siehe Abb. 1.1), um die thermische Belastung des zu fertigenden Werkstücks und des Werkzeugs zu reduzieren. Das liegt darin begründet, dass fast die gesamte Prozess-Energie durch Scherung und Reibung in Wärme umgewandelt wird und sich ohne Kühlschmierstoffe fast

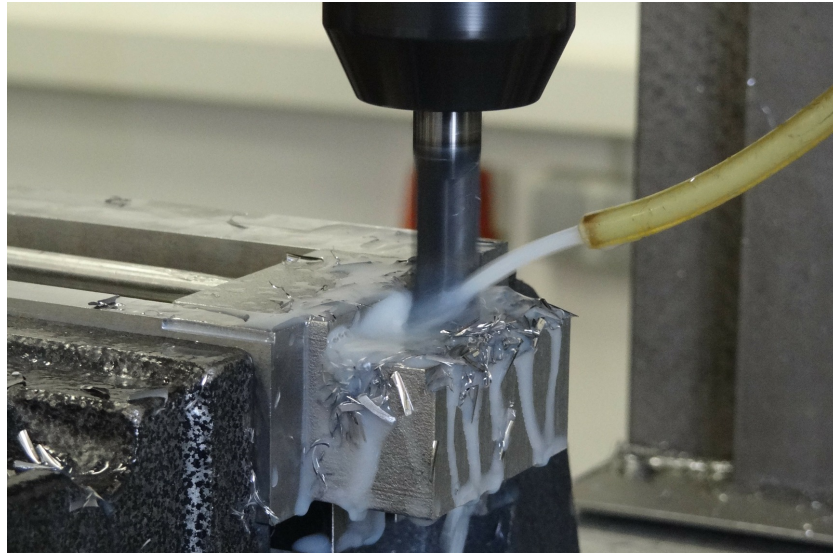


Abbildung 1.1: Einsatz von Kühlschmierstoff in der spanenden Fertigung (Quelle: [14])

ausschließlich auf das Werkstück, das Werkzeug und die Späne verteilt [32]. Durch den Einsatz von Kühlschmierstoffen lassen sich außerdem der Werkzeugverschleiß verringern, höhere Fertigungsgenauigkeiten erzielen und die Qualität der gefertigten Oberfläche verbessern [18]. Dennoch existiert ein gesteigertes Interesse daran, den Kühlmiteleinsatz durch Minimalmengenschmierung zu reduzieren oder gänzlich darauf zu verzichten. Einerseits werden Kosten durch die Beschaffung und Entsorgung verursacht und die Späne verunreinigt, was ein Recycling erschwert. So betragen die zusätzlichen Kosten, die durch das Kühlschmieren verursacht werden, 7 – 17 % der gesamten Produktionskosten [106]. Andererseits fallen neben diesen wirtschaftlichen Interessen vor allem ökologische und gesundheitliche Aspekte, die den Arbeitsplatz betreffen, ins Gewicht.

Um die erforderlichen Qualitätsstandards trotz der erhöhten thermischen Belastung bei der Trockenbearbeitung einzuhalten, ist eine Kompensation von Formabweichungen erforderlich. Durch die zunehmende Digitalisierung von Prozessketten (Industrie 4.0), moderne Sensorik und die Verwendung von Hochleistungsrechnern ist die nichtlineare Optimierung zur Minimierung von Fertigungsfehlern geeigneter denn je. Es benötigt jedoch Vorhersagemodelle, auf deren Grundlage die Fertigung optimiert werden kann. Ein im Projekt „Kompensationsplanung thermischer Prozesseinflüsse beim Trockenfräsen und Trockenbohren“ entstandener neuartiger hybrider Modellansatz ermöglicht neben der Berücksichtigung von Geometrieabweichungen aufgrund thermoelastischer Effekte auch die Vorhersage von Änderungen des Eigenspannungszustands und die damit einhergehenden Biegungen und Torsion. In Abhängigkeit von der Dicke des Bauteils kann sowohl der eine als auch der andere Mechanismus dominieren. Bisher nicht erforschte physikalische Zusammenhänge der Modellierung können durch empirische Regressionsmodelle ergänzt werden, die durch eine breite Datengrundlage aus Experimenten synthetisiert werden. Ein großer Bestandteil dieser Arbeit ist die Bestimmung der nicht direkt messbaren Wärmestromdichten aus Experimenten für die Modellierung der thermischen Effekte. Ein Ziel

dabei ist, die nichtlineare Optimierung zur Lösung der Parameteridentifikationsprobleme anzuwenden. Im Fokus stehen dabei das Fräsen und das Bohren. Darüber hinaus sollen in dieser Arbeit mithilfe des Hybridmodells die Formabweichungen beim Trockenfräsen minimiert bzw. eine ideale Werkstückgeometrie bestmöglich angenähert werden. Dabei werden neben dem Vergleich verschiedener Bearbeitungsstrategien, die durch die Anzahl, Anordnung und Reihenfolge der Fräserbahnen definiert werden, vor allem die Maschinenparameter mit den Methoden der nichtlinearen Optimierung optimiert. Darüber hinaus soll auf eine effiziente und praxisnahe Fertigung geachtet werden.

Zusätzlich zu der ingenieurwissenschaftlichen Anwendung steht aus mathematischer Sicht die Optimierung unter der Berücksichtigung von partiellen Differentialgleichungen (PDE) im Mittelpunkt. Das Interesse an dieser Thematik ist in den letzten Jahren dank steigender Leistung von Rechensystemen stetig größer geworden. Neben den eher theoretischen Betrachtungen von Existenz und Eindeutigkeit der Lösungen PDE-restringierter Optimierungsaufgaben verbleibt die Komplexität und die damit einhergehende Rechenzeit als entscheidende Problemstellung. Es ist weiterhin notwendig, die Effizienz der verwendeten Methodiken zu steigern, um die Modellierung verbessern zu können. Ein vielversprechender Ansatz ist das *Simultaneous Analysis and Design* (SAND)<sup>1</sup>, bei dem die diskretisierte PDE als Nebenbedingung dem Optimierungsproblem hinzugefügt wird. Dafür sollen die geeigneten PDEs und die nötigen Randbedingungen hergeleitet und diskretisiert werden. Zusätzlich kommt dem Ausnutzen der Systemstrukturen der Optimierung bei diesem Ansatz eine besondere Bedeutung zu. Ein häufig verwendeter weiterer Optimierungsansatz ist das *Nested Analysis and Design* (NAND)<sup>2</sup>. Theoretische Überlegungen zeigen, dass das in dieser Arbeit favorisierte SAND bei nichtlinearen PDEs Rechenvorteile aufweisen kann. Neben der erfolgreichen Anwendung des SAND-Ansatzes auf die Aufgabenstellungen der maschinellen Fertigung ist ein Ziel dieser Arbeit einen experimentellen Nachweis für dessen Effizienz zu erbringen. Insgesamt steht bei der PDE-restringierten Optimierung die Methodik sowie die praktikable Umsetzung der Optimierung und nicht das hochpräzise numerische Lösen der PDEs im Vordergrund.

## 1.2 Aufbau

Die vorliegende Arbeit ist wie folgt gegliedert: Im Anschluss an diese Einleitung werden in Kapitel 2 zunächst die Ansätze SAND und NAND zur Lösung PDE-restringierter Optimierungsaufgaben motiviert und deren Vor- und Nachteile diskutiert. Für das numerische Lösen der eingeführten Optimierungsprobleme sind die anschließend in kompakter Form zusammengefassten Optimalitätskriterien von Bedeutung. Diese bilden die Grundlage für die *Sequentielle Quadratische Programmierung* (SQP) und das *Innere-Punkte* (IP)-Verfahren. Mit der abschließend vorgestellten *Finite-Elemente-Methode* (FEM) las-

<sup>1</sup>Das SAND wird auch All-At-Once-Optimierung genannt.

<sup>2</sup>Das NAND wird auch Black-Box-Optimierung genannt.

sen sich PDEs diskretisieren, um sie mit den diskutierten Ansätzen einem Optimierungsproblem hinzuzufügen.

Kapitel 3 beginnt mit der Modellierung von Formänderungen beim Fräsen. In einem hybriden Ansatz werden dabei in zwei Sub-Modellen thermisch bedingte und quellspannungsbedingte Verschiebungen getrennt voneinander betrachtet. Im darauf aufbauenden Optimierungskonzept werden analog zu den Sub-Modellen zwei Optimierungsschritte zur Kompensation der Formänderungen herausgearbeitet. Nach den Betrachtungen über das Fräsen wird anschließend das Bohren thematisiert. Die beiden spannenden Verfahren werden zum Teil empirisch modelliert. Das Kapitel wird beendet mit der Vorstellung der dafür notwendigen Versuchsreihen.

In Kapitel 4 wird zu Beginn die Parameteridentifikation zur Quantifizierung des Wärmeeintrags aus Temperaturdaten vorgestellt. Als thermisches Modell dient dabei die Wärmeleitungsgleichung, deren Randbedingungen und Diskretisierung für das Fräsen definiert werden. Das daraus resultierende PDE-restringierte Optimierungsproblem wird anschließend numerisch gelöst. Die Ergebnisse umfassen eine Validierung mit künstlichen Daten und die Bestimmung von Wärmestromdichten aus den vorgestellten Realdaten.

Kapitel 5 baut auf der Parameteridentifikation aus dem vorherigen Kapitel auf. Dabei wird durch die Berücksichtigung von temperaturabhängigen Materialparametern eine nichtlineare Form der Wärmeleitungsgleichung auf das Bohren angewendet. Durch Formulierung des nichtlinearen Optimierungsproblems in SAND- und NAND-Form lässt sich nach einer Validierung mit künstlichen Daten ein Rechenzeitvergleich beider Ansätze durchführen, dem sich eine Identifikation von Realdaten anschließt.

Kapitel 6 behandelt die Beschreibung und Durchführung der beiden Optimierungsschritte zur Kompensation von Formabweichungen beim Fräsen. Zunächst wird das erste Sub-Modell herangezogen, um Maschinenparameter und Bearbeitungsstrategien hinsichtlich einer spannungsarmen Fertigung zu optimieren. Die numerischen Ergebnisse der Optimierung bilden die Grundlage für die Minimierung der Gesamtformabweichung im zweiten Teil. Hierbei werden die verbliebenen Formabweichungen des ersten Teils mit den thermisch bedingten Verschiebungen des zweiten Sub-Modells superponiert, um anschließend mit der Optimierung von parabelförmigen Fräsbahnen die Gesamtformabweichung zu kompensieren. Beendet wird das Kapitel mit einem Vergleich von gemessenen Formabweichungen realer Experimente und den numerischen Ergebnissen.

Kapitel 7 beendet den Hauptteil der Arbeit mit einer Zusammenfassung und einem Ausblick.

### 1.3 Einordnung des wissenschaftlichen Beitrags

Der Beitrag dieser Arbeit zur Wissenschaft kann in drei Bereiche gegliedert werden. Der erste Bereich ist die Erweiterung der Methodiken zur Bestimmung von Wärmestrom-

dichten in der spanenden Fertigung. Als zweites ist die Erweiterung von Kompensationsstrategien für Formabweichungen beim Trockenfräsen wissenschaftlich interessant. Drittens sind sowohl die Anwendungen in der PDE-restringierten Optimierung als auch der Vergleich der beiden Ansätze SAND und NAND zu nennen.

Der folgende Abschnitt ist in Teilen an [108] angelehnt.

#### **Bestimmung der Wärmestromdichte in der spanenden Fertigung**

Zur Bestimmung und Optimierung von Formabweichungen in der spanenden Fertigung ist die eingebrachte Wärme eine wichtige Größe. Der Wärmefluss über den Rand eines betrachteten Objekts wird durch die Wärmestromdichte quantifiziert. Diese ist experimentell jedoch nicht direkt messbar. Die indirekte Bestimmung der Wärmestromdichte wird in der Literatur häufig als *boundary Inverse Heat Conduction Problem* bezeichnet. Einen guten Überblick über Theorie und Lösungsansätze bieten beispielsweise die Bücher von Beck et al. [9], Alifanov [3], Özisik und Orlande [77] sowie der Artikel von Colaço et al. [28]. Als wichtige Lösungsansätze seien ohne Anspruch auf Vollständigkeit die *Tikhonov-Regularisierung* [99], die *Function Specification Method* von Beck et al. [9] sowie die *iterative Regularisierung* von Alifanov [3] genannt. Diese fundamentalen Methoden werden bis heute verwendet und weiterentwickelt [40, 58, 65, 80, 114, 115].

Viele der gebräuchlichen Methoden verringern oder minimieren die Differenz von gemessenen Daten und simulierten Werten, um den Wärmeeintrag zu bestimmen. In der Praxis werden unterschiedliche Mess- und Simulationsverfahren angewendet. Dyck [34] sowie Pabst [78] und Denkena et al. [31] verwenden die Kalorimetrie bzw. Thermoelemente, um durch die Homogenisierungstemperatur den Wärmeeintrag zu berechnen. Richardson et al. [84] vergleichen die Temperaturen eines analytischen Modells mit den gemessenen Werten von Thermoelementen beim Trockenfräsen, während Biermann et al. [12] die Temperaturen der Thermoelemente mit denen einer 3D-FEM-Simulation beim Tieflochbohren vergleichen. Auch Tai [97] verwendet Thermoelemente beim Tieflochbohren in Kombination mit einer FEM-Simulation. Chen et al. [74] kombinieren die FEM mit Messungen durch Pyrometrie. Sölter und Gulpak verwenden [93] Thermografieaufnahmen und vergleichen einzelne Isothermen dieser Bilder mit denen von simulierten Temperaturfeldern. All diese Ansätze haben gemein, dass der Temperaturabgleich nur über wenige Messpunkte oder sehr kleine Temperaturbereiche erfolgt. Es ist jedoch erstrebenswert, einen großen Bereich an Temperaturdaten zu verwenden, um den Einfluss von Modellvereinfachungen und Messfehlern zu beschränken. Ganze Temperaturfelder werden von Deppermann und Kneer [33] mit einem vereinfachten 1D-Modell für das orthogonale Trockendrehen benutzt. Auch Wernsing et al. [108] vergleichen ganze Temperaturfelder, wobei bei diesem Ansatz die simulierten Temperaturen von Trockenfräsprozessen durch eine zeitabhängige 2D-Finite-Differenzen-Methode berechnet werden und durch eine Parameteridentifikation den Daten angenähert werden. Eine Erweiterung des Vergleichs ganzer Temperaturfelder um eine zeitabhängige 3D-FEM Modellierung erfolgt in [107]. Die in der vorliegenden Arbeit präsentierte Bestimmung der Wärmestromdichte basiert

auf dieser Veröffentlichung. Die Implementierung ist vergleichsweise modular, sodass die Methodik nicht nur auf das Trockenfräsen sondern auch auf das Trockenbohren angewandt werden kann. Methodisch ist dieses Vorgehen der iterativen Regularisierung zuzuordnen. Sie ist dabei dem genannten Ansatz von Alifanov ähnlich. Der Unterschied besteht darin, dass für die Minimierung der Fehlerquadrate die SQP in SAND-Form verwendet wird.

Die Ausführungen dieser Arbeit konzentrieren sich auf die Modellierung sowie die Methodik zur Bestimmung der Wärmestromdichte. Für einen aktuellen Stand der Technik zur Messung von Temperaturen in der spanenden Fertigung sei die Arbeit von Sölter et al. [95] empfohlen.

### **Kompensation von Formabweichungen beim Trockenfräsen**

Wie schon in der Zielsetzung formuliert, ist in den vergangenen Jahren durch höhere Fertigungsanforderungen und den zunehmenden Verzicht auf Kühlschmierstoffe das Interesse an Kompensationsstrategien zur Verringerung von Formabweichungen in der spanenden Fertigung gestiegen. Es existieren einige aktuelle Arbeiten zu dieser Thematik. In der Arbeit von Brinksmeier et al. [20] wird eine Kompensation durch eine geeignete Wahl von Maschinenparametern und der Einspannung beim Drehen von Lagerringen erreicht. Auch Chen et al. [27] und Yu et al. [112] kompensieren durch geeignetes Einspannen. Ein häufiger Ansatz ist neben der Einspannung die Adaption von Maschinenparametern und Werkzeugpfaden. Dabei ist die Verwendung von Metaheuristiken wie genetischen Algorithmen oder Partikelschwarmoptimierung [10, 53, 113] verbreitet.

Die Kompensation von Formabweichungen wird noch wichtiger, wenn bei der Bearbeitung die Menge an Kühlschmierstoffen deutlich reduziert oder sogar ganz darauf verzichtet wird. Aus diesem Grund wurde das Schwerpunktprogramm (SPP) 1480 „Modellierung, Simulation und Kompensation von thermischen Bearbeitungseinflüssen für komplexe Zerspanprozesse“ der DFG [13] ins Leben gerufen, welches insbesondere auch die Trockenbearbeitung thematisiert. Dabei wurden verschiedene Kompensationsansätze verfolgt. Klocke et al. [65] reduzieren den Wärmeeintrag beim Trockendrehen von Getriebewellen, um Formabweichungen zu verringern. In der Dissertation von Iovkov [59] wird der Werkzeugpfad beim Tieflochbohren mit Minimalmengenschmierung aufgrund von FEM-Simulationen angepasst. Auch Zimmermann et al. [117] korrigieren den NC-Code beim Trockendrehen zur Kompensation von geometrischen Abweichungen. In der Arbeit von Niederwestberg [75] wird ebenfalls eine Anpassung des Werkzeugpfades aufgrund von FEM-Berechnungen beim Trockenbohren und -fräsen verfolgt.

Bei allen bekannten Ansätzen werden entweder nur die thermoelastischen Effekte berücksichtigt oder aber nur die eigenspannungsbedingten Verformungen. Beide Effekte können jedoch insbesondere beim Trockenfräsen von dünnwandigen Bauteilen einen relevanten Einfluss haben. Ein Kompensationsansatz für das Trockenfräsen, der beide Aspekte berücksichtigt, ist dem Autor nicht bekannt. Der Ansatz, der in dieser Arbeit gewählt wird,

berücksichtigt sowohl geometrische Abweichungen aufgrund eines ungleichmäßigen Materialabtrags infolge von thermoelastischen Dehnungen als auch Verformungen durch eine Änderung des Eigenspannungszustands. Für die Kompensation wird die nichtlineare Optimierung der Maschinenparameter und ein Vergleich von verschiedenen Bearbeitungsstrategien durchgeführt.

#### **PDE-restringierte Optimierung**

Die Optimierung mit PDEs hat in den vergangenen Jahren in der anwendungsbezogenen Mathematik an Bedeutung zugelegt. Als Indiz dafür seien der „First Sandia Workshop on Large-Scale PDE-Constrained Optimization“ [11] und das Schwerpunktprogramm 1253 „Optimierung mit partiellen Differentialgleichungen“ [70] der DFG genannt. Die Anwendungen sind dabei vielfältig. Angefangen bei dem klassischen Problem der Formoptimierung in der Strömungsmechanik [79] reichen sie heute bis zur Steuerung von Kühlprozessen in der Glasfertigung [56] und der Optimierung von Krebsbehandlungen [4]. Die Lösung PDE-restringierter Optimierungsprobleme erfolgt in der Regel numerisch. Eine elementare Unterscheidung der numerischen Lösungsansätze kann mit Hilfe der direkten und indirekten Verfahren erfolgen. Durch Anwendung eines direkten Verfahrens wird das kontinuierliche Ausgangsproblem zunächst diskretisiert und anschließend optimiert, während bei den indirekten Verfahren zuerst Optimalitätskriterien der kontinuierlichen Optimierung angewendet werden und das numerische Lösen im Anschluss erfolgt. Es existieren Untersuchungen darüber, unter welchen Bedingungen beide Verfahren übereinstimmen (bspw. [72]). Ebenso gibt es Vergleiche der beiden Verfahren (bspw. [2]). Auch wenn keines der beiden Verfahren grundsätzlich besser für die Lösung PDE-restringierter Optimierungsprobleme geeignet ist, überwiegen in der mathematischen Literatur die direkten Verfahren. Als Beispiel für eine Arbeit mit indirektem Ansatz sei [86] genannt. Bei den direkten Verfahren gibt es wiederum verschiedene Abstufungen. So kann bspw. auf der kontinuierlichen Ebene zunächst ein SQP-Ansatz verfolgt werden und anschließend ein direktes Verfahren für die quadratischen Sub-Probleme folgen [56]. Auch existieren Möglichkeiten ein direktes Verfahren nur auf einen Teil des Problems, z. B. die Zustände, anzuwenden und den Rest, z. B. die Steuerungen bei einem Problem der optimalen Steuerung, anfangs in kontinuierlicher Form zu belassen, wie z. B. in dem Artikel von Hinze und Tröltzsch [57]. In dieser Arbeit wird ein klassischer direkter Ansatz verwendet, bei dem alle kontinuierlichen Größen zunächst diskretisiert werden.

Bei der PDE-restringierten Optimierung lassen sich die Verknüpfung von Optimierung und das Lösen der PDE in das SAND und das NAND unterteilen [6, 55]. Beim SAND-Ansatz wird die PDE als Nebenbedingung der Optimierung hinzugefügt, während beim NAND-Ansatz die PDE extern gelöst wird. In den Anwendungen wird der NAND-Ansatz häufig favorisiert, da er sich durch eine gewisse Benutzerfreundlichkeit auszeichnet. Aber auch das SAND besitzt Vorteile, welche innerhalb dieser Arbeit noch ausführlich diskutiert werden. Beim NAND kommt der Berechnung von Ableitungen eine besondere

Bedeutung zu. Diese werden häufig mit finiten Differenzen oder über den *Adjoint Approach* [56] berechnet. In Verbindung mit dem SAND-Ansatz haben sich verschiedene Lösungsmethoden durchgesetzt. Bei der Betrachtung von quadratischen Problemen (quadratische Zielfunktion und lineare PDE bzw. Nebenbedingungen) ohne Ungleichungsnebenbedingungen ergeben die Optimalitätsbedingungen ein lineares Gleichungssystem, welches nur einmal zu lösen ist. Hierbei ist vor allem das effiziente Lösen des Systems im Fokus der Forschung. Einen Schwerpunkt bildet dabei die Vorkonditionierung von iterativen Lösern [52, 83, 88]. Die Vorkonditionierung ist jedoch häufig sehr problemspezifisch, ein universeller Ansatz existiert nicht. Eine Alternative sind die direkten Löser der linearen Algebra. Diese kommen häufig bei den SQP-Verfahren zur Lösung von Optimierungsproblemen mit nichtlinearen Nebenbedingungen zum Einsatz. Die SQP-Methoden benötigen Ableitungsinformationen zweiter Ordnung. Sind diese nicht exakt bekannt, können dabei *Quasi-Newton-Approximationen* zum Einsatz kommen. Ist neben den Werten auch die Struktur der Ableitungen zweiter Ordnung unbekannt, gibt es die Möglichkeit ein reduziertes SQP-Verfahren zu verwenden [25, 76]. Angewandt wird diese Methodik bspw. in [61] und [64].

In dieser Arbeit wird der SAND-Ansatz mit einem herkömmlichen SQP-Verfahren bzw. einem IP-Verfahren sowie einem direkten Löser der linearen Algebra verwendet. Dies verspricht neben einer flexiblen Verwendung auch das effiziente Lösen von Optimierungsproblemen mit PDE-Nebenbedingungen, da der Optimierung analytische Ableitungen zur Verfügung gestellt wurden. Der Schwerpunkt in der vorliegenden Arbeit liegt auf der Anwendung der Methodik im ingenieurwissenschaftlichen Bereich. Es existieren eine Vielzahl von Arbeiten, die die angewandte PDE-restringierte Optimierung zum Thema haben. Erhardt et al. [36] verwenden ein zeitabhängiges, eindimensionales, nichtlineares Modell für die Strömung von Gasen in Pipelines. Brandenburg et al. [17] benutzen eine zeitabhängige, dreidimensionale, nichtlineare PDE für die Formoptimierung. Der Optimierungsansatz ist hier das NAND und es werden keine Ungleichungsnebenbedingungen eingesetzt. Auch Stoll und Wathen [96] verwenden eine zeitabhängige, dreidimensionale PDE, die aber linear ist. Die aus Sicht der Optimierung schwierigste Problemstellung dieser Arbeit ist eine zeitabhängige, dreidimensionale Anwendungsaufgabe mit nichtlinearer PDE und Ungleichungsnebenbedingungen unter Verwendung des SAND-Ansatzes. Eine vergleichbare Veröffentlichung ist dem Autor nicht bekannt. Zudem wird für diese komplexe Problemstellung ein Rechenzeitvergleich der beiden Ansätze SAND und NAND durchgeführt.



# Kapitel 2

## Grundlagen der Optimierung mit PDE-Nebenbedingungen

### Inhalt

---

2.1	Lösungskonzepte der PDE-restringierten Optimierung . . . . .	10
2.2	Grundlagen der nichtlinearen Optimierung . . . . .	15
2.2.1	Theorie der nichtlinearen Optimierung . . . . .	15
2.2.2	Algorithmen der nichtlinearen Optimierung . . . . .	21
2.3	Finite-Elemente-Methode und Softwarebibliothek deal.II . . . . .	31

---

Die Aufgabenstellungen, die in dieser Arbeit behandelt werden, können als Optimierungsprobleme unter Nebenbedingungen aufgefasst werden und sollen mit entsprechenden Methodiken gelöst werden. Deshalb wird zunächst die allgemeine Problemstellung als Optimierungsproblem definiert. Zur besseren Lesbarkeit werden die kontinuierlichen Funktionen und Funktionale mit  $(\cdot)_\phi$  gekennzeichnet, um bei den später häufiger verwendeten diskretisierten Varianten auf einen Index verzichten zu können. Sei dazu  $S_\phi \in \mathcal{S} := \text{Abb}(\mathbb{R}^{n_d}, \mathbb{R}^{n_s})$  eine Abbildung, die  $n_s$  Zustände in einem  $n_d$ -dimensionalen Raum beschreibt. Innerhalb dieser Arbeit wird  $n_d = 4$  sein, da die Zustände wie die Temperatur örtlich und zeitlich betrachtet werden. Des Weiteren seien je nach Anwendung Steuer-, Design- oder Identifikationsfunktionen  $Z_\phi \in \mathcal{Z} := \text{Abb}(\mathbb{R}^{n_d}, \mathbb{R}^{n_z})$  gegeben. Diese Funktionen sind Optimierungsfunktionen, die der Optimierung als Freiheitsgrade dienen. Da aber auch die Zustände Optimierungsfunktionen sein können, werden die Steuer-, Design- oder Identifikationsfunktionen allgemein als Entscheidungsfunktionen bezeichnet. Zusammen mit den Zuständen seien sie durch eine Operatorgleichung  $P_\phi : (\mathcal{S}, \mathcal{Z}) \rightarrow \mathbb{R}^{n_p}$  mit  $P_\phi(S_\phi, Z_\phi) = 0$  gekoppelt, die ein mathematisches Modell darstellt. Dieses kann gewöhnliche Differentialgleichungen, PDEs oder algebraische Gleichungen beinhalten. Das Optimierungskriterium wird durch ein Zielfunktional  $F_\phi : (\mathcal{S}, \mathcal{Z}) \rightarrow \mathbb{R}$  definiert, welches es zu minimieren gilt. Darüber hinaus sollen die Nebenbedingungen  $G_\phi(S_\phi, Z_\phi) \leq 0$  und  $H_\phi(S_\phi, Z_\phi) = 0$  eingehalten werden. Diese werden durch ein Funktional von Ungleichungsnebenbedingungen  $G_\phi : (\mathcal{S}, \mathcal{Z}) \rightarrow \mathbb{R}^{m_g}$  und Gleichungsnebenbedingungen  $H_\phi : (\mathcal{S}, \mathcal{Z}) \rightarrow \mathbb{R}^{m_h}$  definiert. Es können zudem Beschränkungen für die Zustände

und die Entscheidungsfunktionen in der Gestalt  $S_L \leq S_\phi \leq S_U$  bzw.  $Z_L \leq Z_\phi \leq Z_U$  mit  $S_L, S_U \in \mathbb{R}^{n_s}$  und  $Z_L, Z_U \in \mathbb{R}^{n_z}$  auftreten. Das allgemeine Optimierungsproblem hat dann die Form

$$\begin{aligned}
 & \min_{X_\phi \in \mathcal{S} \times \mathcal{Z}} F_\phi(S_\phi, Z_\phi) \\
 & \text{unter } G_\phi(S_\phi, Z_\phi) \leq 0 \\
 & \quad H_\phi(S_\phi, Z_\phi) = 0 \\
 & \quad S_L \leq S_\phi \leq S_U \\
 & \quad Z_L \leq Z_\phi \leq Z_U \\
 & \text{Modell: } P_\phi(S_\phi, Z_\phi) = 0,
 \end{aligned} \tag{2.1}$$

wobei unterschiedliche Wahlmöglichkeiten der Optimierungsfunktion  $X_\phi$  ( $X_\phi = Z_\phi$  oder  $X_\phi = (S_\phi^T, Z_\phi^T)^T$ ) und die Kopplung des Problems mit dem Modell  $P_\phi$  im Laufe dieser Arbeit spezifiziert werden. Ist  $P_\phi$  dabei eine PDE, spricht man auch von einem PDE-restringierten Optimierungsproblem.

In diesem Kapitel werden zunächst die Lösungskonzepte SAND und NAND motiviert. Anschließend werden Optimalitätskriterien vorgestellt, die die Grundlage für die Optimierungsalgorithmen SQP-Verfahren und IP-Verfahren und ihre Implementierungen WORHP und Ipopt sind. Abschließend wird auf die Diskretisierung von PDEs durch die FEM eingegangen.

## 2.1 Lösungskonzepte der PDE-restringierten Optimierung

Die Lösungsansätze der Optimierungsaufgabe (2.1) können in zwei Klassen unterteilt werden: die direkten<sup>1</sup> und die indirekten Verfahren<sup>2</sup>. Die indirekten Verfahren wenden zunächst die notwendigen Kriterien erster Ordnung aus der Theorie der Optimierung kontinuierlicher Funktionen an (z. B. das Minimumprinzip bei einem Problem der optimalen Steuerung), um anschließend das resultierende Gleichungssystem, welches gewöhnliche Differentialgleichungen oder PDEs enthalten kann, zu lösen. Dies erfolgt dann in der Regel durch anschließende numerische Diskretisierung. Die direkten Verfahren diskretisieren das Ausgangsproblem als ersten Schritt, womit das unendlich-dimensionale Problem (2.1) in ein endlich-dimensionales Problem überführt wird. Das Resultat ist ein *nichtlineares Optimierungsproblem* (NLP)<sup>3</sup>, auf welches anschließend notwendige Optimalitätskriterien angewandt werden können, um es mit einem NLP-Löser zu lösen.

Beide Verfahren bieten Vor- und Nachteile (vgl. [23] oder [56]). Auf der einen Seite bieten die indirekten Verfahren zwar eine umfassende Einsicht in die optimale Lösung und

<sup>1</sup>Die direkten Verfahren werden auch *First discretize, then optimize* genannt.

<sup>2</sup>Die indirekten Verfahren werden auch *First optimize, then discretize* genannt.

<sup>3</sup>Ein NLP wird auch *nichtlineares Programm* genannt.

die sogenannten adjungierten Zustände bei hoher Lösungsgenauigkeit. Andererseits sind sie nicht geeignet, um Probleme mit beliebigen Zielfunktionen und Beschränkungen zu lösen. Dies ist insbesondere der Fall, wenn gleichzeitig Zustandsbeschränkungen und Beschränkungen der Entscheidungsvariablen betrachtet werden [56]. Zudem bedeuten die adjungierten Differentialgleichungen einen erhöhten Aufwand für den Anwender. Das in mancher Hinsicht benutzerfreundlichere direkte Verfahren hat in der Vergangenheit an Popularität gewonnen, da die Lösungsungenauigkeit durch die gesteigerte Leistung von Rechnern in Kombination mit feineren Diskretisierungen ausgeglichen werden konnte. Im Zuge dieser Arbeit werden im Folgenden die direkten Verfahren als Lösungsansatz verwendet. Für die Theorie zum Lösen von PDE-restringierten Optimierungsproblemen bzw. Optimalsteuerungsproblemen seien die Bücher von Hinze et al. [56] und Tröltzsch [102] empfohlen. Diese beschreiben Aussagen zur Existenz von Lösungen, Optimalitätsbedingungen und numerische Lösungsansätze. Innerhalb dieser Arbeit werden die Optimierungsprobleme in diskretisierter Form betrachtet.

Das Problem (2.1) reduziert sich mit einer geeigneten Diskretisierung, auf die an entsprechender Stelle noch detaillierter eingegangen wird, zu einem endlich-dimensionalen, nichtlinearen Optimierungsproblem der Form:

$$\begin{aligned}
 & \min_{X \in \mathbb{R}^{n_x}} F(S, Z) \\
 & \text{unter } G(S, Z) \leq 0 \\
 & \quad H(S, Z) = 0 \\
 & \quad S_L \leq S \leq S_U \\
 & \quad Z_L \leq Z \leq Z_U \\
 & \text{Modell: } P(S, Z) = 0
 \end{aligned} \tag{2.2}$$

Dabei sind  $S \in \mathbb{R}^{n_s}$ ,  $Z \in \mathbb{R}^{n_z}$  und  $X \in \mathbb{R}^{n_x}$  ( $X = Z$  oder  $X = (S^T, Z^T)^T$ ) die diskretisierten Variablen des Problems. Diese Vektoren enthalten die diskreten Werte der Funktionen  $S_\phi$ ,  $Z_\phi$  und  $X_\phi$  an den diskreten Punkten im Raum mit einer geeigneten Sortierung. Die Funktionen  $F : \mathbb{R}^{n_s} \times \mathbb{R}^{n_z} \rightarrow \mathbb{R}$ ,  $G : \mathbb{R}^{n_s} \times \mathbb{R}^{n_z} \rightarrow \mathbb{R}^{m_g}$  und  $H : \mathbb{R}^{n_s} \times \mathbb{R}^{n_z} \rightarrow \mathbb{R}^{m_h}$  sind geeignete diskrete Approximationen der Funktionale  $F_\phi$ ,  $G_\phi$  und  $H_\phi$ . Der Modell-Operator  $P_\phi$  wird durch eine geeignete Diskretisierung zu einem Gleichungssystem  $P : \mathbb{R}^{n_s} \times \mathbb{R}^{n_z} \rightarrow \mathbb{R}^{n_s}$ . Die Dimension der Beschränkungen wird an die der Vektoren angepasst:  $S_L, S_U \in \mathbb{R}^{n_s}$  und  $Z_L, Z_U \in \mathbb{R}^{n_z}$ . Beschränkungen in dieser Form werden auch Boxschränken genannt. Sie können in der Regel in Optimierungsalgorithmen effizienter als allgemeine Nebenbedingungen gehandhabt werden.

Die Modelle innerhalb dieser Arbeit liegen als PDE vor. Das Optimierungsproblem (2.2) und die Modell-PDE  $P$  bzw.  $P_\phi$  können auf zwei unterschiedliche Arten miteinander verknüpft werden (vgl. [6]). Der erste Ansatz ist das NAND. Es verwendet einen externen PDE-Löser, um die Modellzustände zu berechnen. Die Zustände gehen dann in diskreter Form in die Zielfunktion und die Nebenbedingungen des Optimierungsproblems ein.

Mathematisch kann diese Vorgehensweise durch

$$\begin{aligned} & \min_{Z \in \mathbb{R}^{n_z}} F(S(Z), Z) \\ & \text{unter } G(S(Z), Z) \leq 0 \\ & \quad H(S(Z), Z) = 0 \\ & \quad Z_L \leq Z \leq Z_U \\ & \quad S_L \leq S(Z) \leq S_U \end{aligned}$$

dargestellt werden, wobei die Zustände  $S(Z)$  durch das externe Lösen der PDE in Abhängigkeit von  $Z$  bestimmt werden. Wichtig ist hierbei, dass die Lösbarkeit der PDE sichergestellt wird. Die Optimierungsvariable  $X \in \mathbb{R}^{n_x}$  des Problems ist durch  $X := Z$  gegeben. Die Beschränkungen für  $Z$  sind Boxschränken. Die Beschränkungen für  $S(Z)$  sind im Allgemeinen nichtlineare Nebenbedingungen.

Der zweite Lösungsansatz ist das SAND. Die PDE wird dabei in diskreter Form als Gleichungssystem den Nebenbedingungen des Optimierungsproblems hinzugefügt:

$$\begin{aligned} & \min_{X \in \mathbb{R}^{n_x}} F(X) \\ & \text{unter } G(X) \leq 0 \\ & \quad H(X) = 0 \\ & \quad P(X) = 0 \\ & \quad X_L \leq X \leq X_U \end{aligned} \tag{2.3}$$

Hierbei sind auch die diskreten Zustände  $S$  in den Optimierungsvariablen enthalten:  $X := (S^T, Z^T)^T$ . In diesem Fall sind auch Zustandsbeschränkungen Boxschränken. Der Unterschied beider Ansätze bei der Verwendung herkömmlicher iterativer NLP-Löser wird in Abb. 2.1 verdeutlicht. Während beim NAND-Ansatz zur Auswertung von Zielfunktion und Nebenbedingungen die PDE in jeder Iteration exakt gelöst werden muss, um die nötigen Zustände zu berechnen, genügt beim SAND-Ansatz pro Iteration eine Auswertung der PDE-Nebenbedingung. NAND bewegt sich bei der Suche nach dem Minimum bezüglich der Zustände ausschließlich im zulässigen Bereich. SAND hingegen iteriert auch über Zustände, die die PDE nicht erfüllen. Das Lösen der PDE erfolgt hier simultan zum Minimieren der Zielfunktion.

Dadurch, dass im SAND keine externen Löser für die PDE verwendet werden, muss die PDE als Nebenbedingung zum Optimierungsalgorithmus hinzugefügt werden. Das bedeutet in der Regel einen erhöhten Implementierungsaufwand. Allerdings besitzt SAND auch eine Reihe von Vorteilen. Im Allgemeinen bieten PDE-Löser dem Benutzer keine Informationen über die Ableitungen der Zustände  $S$  in Abhängigkeit von  $Z$  an, weshalb diese beim NAND numerisch berechnet werden müssen. Durch die selbstständige Implementierung der PDE-Nebenbedingungen können diese im SAND vergleichsweise einfach analytisch berechnet werden. Dies kann eine schnellere Berechnung der Ableitungen sowie eine verbesserte Konvergenz des verwendeten Optimierungsalgorithmus zur Folge haben. Ein weiterer Vorteil ist eine gewisse Robustheit gegenüber lokalen Minima (siehe Abb. 2.2a). Der NAND-Ansatz wird durch die ausschließliche Suche im zulässigen

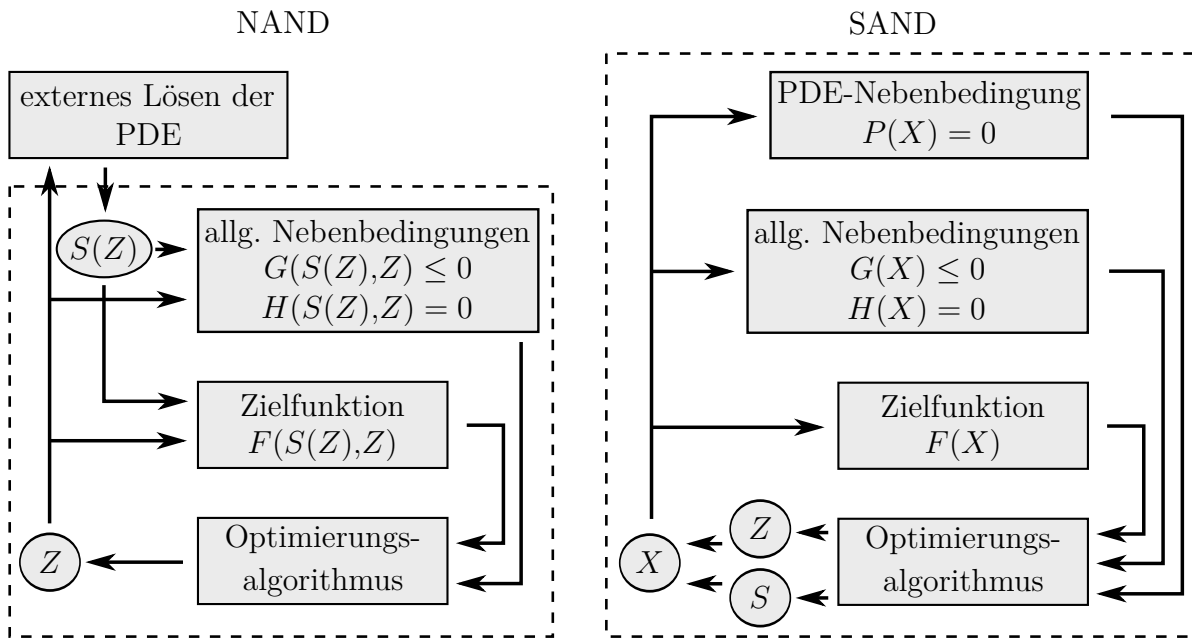
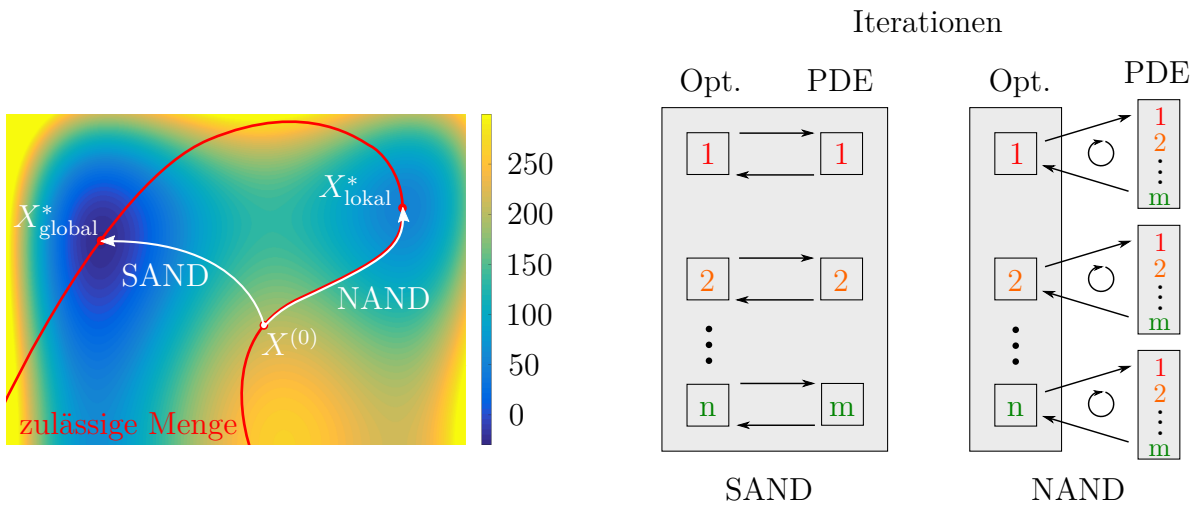


Abbildung 2.1: Algorithmischer Unterschied zwischen SAND und NAND (für eine bessere Übersicht werden hier Boxschränken und Zustandsbeschränkungen nicht dargestellt)



(a) Robustheit gegenüber lokalen Minima

(b) Zusammenspiel von Optimierungsalgorithmus und dem Lösen einer nichtlinearen PDE

Abbildung 2.2: Vorteile von SAND gegenüber NAND

Lösungsraum eingeschränkt und neigt so eher dazu, in lokale Minima zu konvergieren als der SAND-Ansatz, der durch das Verlassen des zulässigen Bereichs eventuell bessere Suchrichtungen nutzen kann. Auch die Behandlung von Zustandsbeschränkungen ist bei

den beiden Ansätzen unterschiedlich. Während SAND die Zustände als Optimierungsvariablen behandelt und Zustandsbeschränkungen als Boxschränken einfach und effizient in das Problem integrieren kann, sind im NAND Zustandsbeschränkungen allgemeine, nichtlineare Nebenbedingungen, für deren Auswertung und Ableitungsberechnung die PDE extern gelöst werden muss. Dies ist in der Regel deutlich rechenintensiver als die Berücksichtigung von Boxschränken.

Ein weiterer Rechenvorteil ergibt sich bei der Verwendung von Modellen, die durch eine nichtlineare PDE formuliert werden. Diese müssen im Allgemeinen iterativ gelöst werden. Das hat im NAND-Ansatz zur Folge, dass in jeder Hauptiteration des Optimierungsalgorithmus zum einmaligen Lösen der PDE mehrere Iterationen erforderlich sind. Insbesondere bei der Verwendung von Zustandsbeschränkungen und bei der Berechnung von numerischen Ableitungen, bei denen die PDE vielfach in jeder Hauptiteration gelöst werden muss, kann dies rechenintensiv sein (siehe Abb. 2.2b rechts). Da SAND nicht ausschließlich auf zulässige Zustände angewiesen ist, um in der Optimierung voranzuschreiten, kann das Lösen der PDE sukzessive mit den Hauptiterationen erfolgen. Erst in der letzten Iteration wird hier gefordert, dass die Zustände die PDE erfüllen. Das kann im Vergleich zum exakten Lösen der PDE viele Rechenoperationen sparen (siehe Abb. 2.2b links). Dieser theoretische Rechenzeitvorteil vom SAND-Ansatz gegenüber dem NAND-Ansatz wurde bereits von Krause [67] für eine konstruierte Aufgabenstellung ohne Ungleichungsnebenbedingungen und mit einer nichtlinearen, zeitunabhängigen zweidimensionalen PDE untersucht und validiert. In dieser Arbeit wird die Untersuchung für eine anwendungsnahe Problemstellung mit Ungleichungsnebenbedingungen und einer nichtlinearen, zeitabhängigen dreidimensionalen PDE durchgeführt.

Die Tatsache, dass beim SAND-Ansatz wesentlich mehr Optimierungsvariablen berücksichtigt werden, muss kein Nachteil sein. Durch die Verwendung von Lösern, die die Schwachbesetztheit der Problemstrukturen ausnutzen, kann dieser Umstand relativiert werden. Aufgrund dieser Vorteile werden die Optimierungsaufgaben in dieser Arbeit primär mit dem SAND-Ansatz gelöst. Zur Vereinfachung werden im Folgenden die Modellgleichungen  $P$  in den allgemeinen Gleichungsnebenbedingungen  $H$  ergänzt. Dadurch reduziert sich (2.3) zu:

$$\begin{aligned} \min_{X \in \mathbb{R}^{n_x}} \quad & F(X) \\ \text{unter} \quad & G(X) \leq 0 \\ & H(X) = 0 \\ & X_L \leq X \leq X_U \end{aligned} \tag{2.4}$$

Der NAND-Ansatz wird für die Parameteridentifikation des Bohrens (siehe Kapitel 5.3) parallel verwendet um einen Vergleich beider Ansätze durchzuführen. Zudem wird er für einen Teil der Optimierung von Formabweichungen verwendet, da das entsprechende Modell als Black-Box vorliegt.

## 2.2 Grundlagen der nichtlinearen Optimierung

Das folgende Kapitel ist an die Werke von Geiger und Kanzow [45] sowie Nocedal und Wright [76] und an ein unveröffentlichtes Vorlesungsskript von Büskens [22] angelehnt.

Nachdem vorgestellt wurde, wie sich unendlich-dimensionale Optimierungsprobleme durch Diskretisierung in ein endlich-dimensionales NLP überführen lassen, sollen im Folgenden die theoretischen und numerischen Grundlagen für deren Lösung vorgestellt werden. Zunächst wird auf Bedingungen eingegangen, die in einem Optimum gelten und das Auffinden eines solchen erleichtern. Zentrale Sätze sind hierbei das notwendige Optimalitätskriterium erster Ordnung, sowie das hinreichende Kriterium zweiter Ordnung. Mithilfe dieser Kriterien lassen sich numerische Algorithmen herleiten, die das Lösen von Optimierungsproblemen ermöglichen.

### 2.2.1 Theorie der nichtlinearen Optimierung

In diesem Abschnitt werden zentrale Resultate der Theorie der restringierten, nichtlinearen Optimierung vorgestellt. Sie bilden die Grundlage moderner Optimierungsalgorithmen. Die betrachtete Problemklasse, das NLP, wird wie folgt definiert:

#### Definition 2.1 (NLP)

Probleme in der Form

$$\begin{aligned} & \min_{X \in \mathbb{R}^{n_x}} F(X) \\ & \text{unter } G_i(X) \leq 0, \quad i = 1, \dots, m_g \\ & \quad \quad H_j(X) = 0, \quad j = 1, \dots, m_h \end{aligned} \quad (2.5)$$

mit der Zielfunktion  $F : \mathbb{R}^{n_x} \rightarrow \mathbb{R}$  und den Nebenbedingungen  $G : \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{m_g}$  und  $H : \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{m_h}$  werden NLP genannt.

Für die Herleitung von Optimalitätskriterien werden Boxschränken  $X_L \leq X \leq X_U$  zunächst als allgemeine Ungleichungsnebenbedingungen in  $G$  durch  $X_L - X \leq 0$  und  $X - X_U \leq 0$  ergänzt.

Punkte  $X$ , die die Nebenbedingungen aus (2.5) erfüllen, werden zulässige Punkte genannt. Zusammen ergeben sie die zulässige Menge.

#### Definition 2.2 (Zulässige Menge)

Die zulässige Menge des Optimierungsproblems (2.5) ist definiert durch

$$\Sigma := \{X \in \mathbb{R}^{n_x} \mid G(X) \leq 0, H(X) = 0\}.$$

Des Weiteren sei die Epsilon-Umgebung wie folgt definiert:

**Definition 2.3 (Epsilon-Umgebung)**

Die Menge  $U_\varepsilon(X) := \{Y \in \mathbb{R}^{n_x} \mid \|Y - X\| < \varepsilon\}$  mit  $\varepsilon > 0$  wird Epsilon-Umgebung von  $X$  genannt.

Mithilfe der zulässigen Menge lässt sich eine zulässige Lösung in Form eines Minimums von (2.5) charakterisieren.

**Definition 2.4 (globales, lokales, striktes Minimum)**

(i) Ein zulässiger Punkt  $X^*$  wird globales Minimum von (2.5) genannt, wenn

$$F(X^*) \leq F(X) \quad \text{für alle } X \in \Sigma \tag{2.6}$$

erfüllt ist.

(ii) Ein zulässiger Punkt  $X^*$  wird lokales Minimum von (2.5) genannt, wenn es eine Epsilon-Umgebung  $U_\varepsilon(X^*)$  von  $X^*$  gibt, für die

$$F(X^*) \leq F(X) \quad \text{für alle } X \in \Sigma \cap U_\varepsilon(X^*) \tag{2.7}$$

gilt.

(iii) Gilt in (2.6) bzw. (2.7)

$$\begin{aligned} &F(X^*) < F(X) \quad \text{für alle } X \in \Sigma \quad \text{mit } X \neq X^* \\ \text{bzw. } &F(X^*) < F(X) \quad \text{für alle } X \in \Sigma \cap U_\varepsilon(X^*) \quad \text{mit } X \neq X^*, \end{aligned}$$

so spricht man auch von einem strikten globalen bzw. lokalen Minimum.

Um Probleme der Art (2.5) zu lösen, benötigt es Kriterien, um mögliche Minima auszumachen.

**Optimalitätskriterien erster Ordnung**

Im Folgenden seien die Funktionen  $F$ ,  $G_i$  und  $H_j$  einmal stetig-differenzierbar. Zum Auffinden von (lokalen) Lösungen eines NLP ist folgende Funktion eine entscheidende Hilfe.



**Definition 2.5 (Lagrange-Funktion)**

Es sei  $m := m_g + m_h$ . Die Funktion  $L : \mathbb{R}^{n_x} \times \mathbb{R}^m \rightarrow \mathbb{R}$  mit

$$L(X, \lambda) := F(X) + \sum_{i=1}^{m_g} \lambda_i G_i(X) + \sum_{j=1}^{m_h} \lambda_{m_g+j} H_j(X)$$

wird Lagrange-Funktion von (2.5) genannt. Der vektorwertige Multiplikator  $\lambda \in \mathbb{R}^m$  heißt in diesem Zusammenhang Lagrange-Multiplikator.

Mit der Lagrange-Funktion lassen sich zentrale Bedingungen für ein Optimum formulieren. Diese sogenannten *Karush-Kuhn-Tucker* (KKT)-Bedingungen ermöglichen wichtige Aussagen über die Lösungen von nichtlinearen Optimierungsproblemen.

**Definition 2.6 (KKT-Bedingungen, KKT-Punkt)**

(i) Die KKT-Bedingungen sind definiert als

$$\begin{aligned} \nabla_X L(X, \lambda) &= 0 \\ H_j(X) &= 0, \quad j = 1, \dots, m_h \\ \lambda_i \geq 0, \quad G_i(X) \leq 0, \quad \lambda_i G_i(X) &= 0, \quad i = 1, \dots, m_g, \end{aligned}$$

wobei

$$\nabla_X L(X, \lambda) = \nabla F(X) + \sum_{i=1}^{m_g} \lambda_i \nabla G_i(X) + \sum_{j=1}^{m_h} \lambda_{m_g+j} \nabla H_j(X)$$

der Gradient der Lagrange-Funktion bzgl.  $X$  ist.

(ii) Jedes Tupel  $(X^*, \lambda^*)$ , das die Bedingungen aus (i) erfüllt, wird KKT-Punkt genannt.

Neben den KKT-Bedingungen sind sogenannte Regularitätsbedingungen erforderlich, um Optimalitätskriterien zu formulieren. Dafür wird zunächst die Menge der aktiven Indizes eingeführt.

**Definition 2.7 (Indexmenge der aktiven Ungleichungsnebenbedingungen)**

Sei  $X \in \Sigma$  ein zulässiger Punkt von (2.5). Mit

$$I_A(X) := \{i \in \{1, \dots, m_g\} \mid G_i(X) = 0\}$$

wird die Indexmenge der aktiven Ungleichungsnebenbedingungen bezeichnet. Eine Ungleichungsnebenbedingung  $G_i(X)$  heißt entsprechend aktiv, wenn  $G_i(X) = 0$  gilt.

Es existiert eine Reihe von verschiedenen Regularitätsbedingungen, mit denen sich Optimalitätsaussagen zum Problem (2.5) machen lassen. Hier werden folgend zwei vorgestellt, die gebräuchlich im Umgang mit nicht-konvexen Optimierungsproblemen sind.

**Definition 2.8 (Regularitätsbedingungen)**

- (i) Der Punkt  $X \in \Sigma$  genügt den *Regularitätsbedingungen von Mangasarian-Fromovitz* (MFCQ) (engl. Mangasarian-Fromovitz constraint qualification), wenn die Gradienten

$$\nabla H_j(X), \quad j = 1, \dots, m_h$$

linear unabhängig sind und ein Vektor  $d \in \mathbb{R}^{n_x}$  existiert mit

$$\nabla G_i(X)^T d < 0, \quad i \in I_A(X) \quad \text{und} \quad \nabla H_j(X)^T d = 0, \quad j = 1, \dots, m_h.$$

- (ii) Der Punkt  $X \in \Sigma$  genügt den *Regularitätsbedingungen der linearen Unabhängigkeit* (LICQ) (engl.: linear independence constraint qualification), wenn die Gradienten

$$\nabla G_i(X), \quad i \in I_A(X) \quad \text{und} \quad \nabla H_j(X), \quad j = 1, \dots, m_h$$

linear unabhängig sind.

Regularitätsbedingungen und KKT-Bedingungen können in einem wichtigen Satz über die Optimalität eines NLP zusammengebracht werden.

**Satz 2.9 (Notwendiges Optimalitätskriterium erster Ordnung)**

Sei  $X^* \in \mathbb{R}^{n_x}$  ein lokales Minimum des Optimierungsproblems (2.5), das den MFCQ genüge. Dann folgt: Es existiert ein Lagrange-Multiplikator  $\lambda^* \in \mathbb{R}^m$ , sodass  $(X^*, \lambda^*)$  ein KKT-Punkt von (2.5) ist.

**Beweis:** siehe [45] oder [22].

**Bemerkung 2.10**

Es kann gezeigt werden (siehe bspw. [45] oder [22]), dass die MFCQ die LICQ implizieren. Satz 2.9 ist deshalb auch gültig unter den restriktiveren Voraussetzungen der LICQ.

Es sei angemerkt, dass Satz 2.9 zwar die Suche nach einem lokalen Minimum erleichtert, indem er einen Zusammenhang zwischen Minimum und KKT-Punkten herstellt. Jedoch garantiert er nicht, dass ein gefundener KKT-Punkt auch ein lokales Minimum des Problems (2.5) ist. Um diese Aussage treffen zu können, wird ein hinreichendes Kriterium benötigt. Dieses wird in Satz 2.15 folgen. Zunächst soll jedoch eine Aussage über die Eindeutigkeit des Lagrange-Multiplikators getroffen werden.

**Satz 2.11 (Eindeutigkeit des Lagrange-Multiplikators)**

Sei  $X^*$  das lokale Minimum aus Satz 2.9, welches zusätzlich den LICQ genüge. Dann folgt: Der Lagrange-Multiplikator  $\lambda^*$  zum KKT-Punkt  $(X^*, \lambda^*)$  ist eindeutig bestimmt.

**Beweis:** siehe [45], [76] oder [22].

**Optimalitätskriterien zweiter Ordnung**

Es existiert eine weitere notwendige Bedingung für die Optimalität, die jedoch die zweite Ableitung der Lagrange-Funktion verwendet. Im Folgenden seien deshalb die vorher verwendeten Funktionen  $F$ ,  $G_i$  und  $H_j$  zweimal stetig-differenzierbar. Zunächst werden zwei weitere Indexmengen definiert, die eine Zerlegung der Indexmenge  $I_A(X^*)$  darstellen.

**Definition 2.12 (Zerlegung der Indexmenge  $I_A$ )**

Die Indexmenge  $I_A(X^*)$  kann in die beiden Indexmengen

$$\begin{aligned} I_A^>(X^*) &:= \{i \in I_A(X^*) \mid \lambda_i^* > 0\} \\ \text{und } I_A^0(X^*) &:= \{i \in I_A(X^*) \mid \lambda_i^* = 0\} \\ \text{mit } I_A(X^*) &= I_A^>(X^*) \cup I_A^0(X^*) \end{aligned}$$

zerlegt werden.

Generell sind die formulierten Indexmengen aus Definition 2.12 auch von  $\lambda^*$  abhängig. Später werden jedoch die LICQ vorausgesetzt, womit die Lagrange-Multiplikatoren durch  $X^*$  eindeutig bestimmt sind. Für das notwendige Kriterium zweiter Ordnung wird folgende Hilfsmenge definiert.

**Definition 2.13 (Kritischer Kegel)**

Sei  $X^*$  ein lokales Minimum. Mit

$$\begin{aligned} K(X^*) := \{d \in \mathbb{R}^{n_x} \mid &\nabla G_i(X^*)^T d = 0, i \in I_A^>(X^*), \\ &\nabla G_i(X^*)^T d \leq 0, i \in I_A^0(X^*), \\ &\nabla H_j(X^*)^T d = 0, j = 1, \dots, m_h\} \end{aligned}$$

wird der kritische Kegel von  $X^*$  bezeichnet.

Die Menge  $K(X^*)$  lässt sich geometrisch interpretieren. Sie enthält Vektoren jener Richtungen, in die vom KKT-Punkt aus die aktiven Ungleichungsnebenbedingungen mit positiver Komponente des Lagrange-Multiplikators aktiv bleiben, sowie die Gleichungsnebenbedingungen erfüllt bleiben. Für die Richtungen  $d \in K(X^*)$  gilt

$$\lambda_i^* \nabla G_i(X^*)^T d = 0, \quad i = 1, \dots, m_g \quad \text{und} \quad \lambda_{m_g+j}^* \nabla H_j(X^*)^T d = 0, \quad j = 1, \dots, m_h,$$

da entweder  $\nabla G_i(X^*)^T d$  oder  $\lambda_i^*$ , sowie  $\nabla H_j(X^*)^T d$  gleich 0 sind. Damit folgt zusammen mit Satz 2.9, dass

$$\sum_{i=1}^{m_g} \lambda_i^* \nabla G_i(X^*)^T d + \sum_{j=1}^{m_h} \lambda_{m_g+j}^* \nabla H_j(X^*)^T d = \nabla F(X^*)^T d = 0.$$

Das bedeutet, dass  $K(X^*)$  die Richtungen enthält, bezüglich denen aus den Ableitungsinformationen erster Ordnung keine Aussage getroffen werden kann, ob  $F$  steigt oder fällt. Diese Richtungen  $d$  sind deshalb in gewisser Weise „kritisch“ und von besonderer Bedeutung für die Formulierung von Optimalitätskriterien zweiter Ordnung.

**Satz 2.14 (Notwendiges Optimalitätskriterium zweiter Ordnung)**

Sei  $X^* \in \mathbb{R}^{n_x}$  ein lokales Minimum des Optimierungsproblems (2.5), das den LICQ genüge. Dann folgt:

$$d^T \nabla_X^2 L(X^*, \lambda^*) d \geq 0 \quad \text{für alle } d \in K(X^*)$$

Dabei ist der Lagrange-Multiplikator  $\lambda^*$  eindeutig bestimmt.

**Beweis:** siehe [45], [76] oder [22].

Das folgende Resultat liefert eine hinreichende Bedingung an einen KKT-Punkt, damit dieser tatsächlich ein Optimum darstellt.

**Satz 2.15 (Hinreichendes Optimalitätskriterium zweiter Ordnung)**

Sei  $(X^*, \lambda^*) \in \mathbb{R}^{n_x+m}$  ein KKT-Punkt mit

$$d^T \nabla_X^2 L(X^*, \lambda^*) d > 0 \quad \text{für alle } d \in K(X^*), d \neq 0. \tag{2.8}$$

Dann folgt:  $X^*$  ist ein striktes lokales Minimum.

**Beweis:** siehe [45], [76] oder [22].

Es sei darauf hingewiesen, dass in KKT-Punkten, in denen (2.8) nicht erfüllt ist, weiterhin keine Aussage über die Optimalität getroffen werden kann. Des Weiteren reicht die Bedingung (2.8) alleine nicht aus, um ein lokales Minimum zu finden, da Satz 2.15 bereits einen gefundenen KKT-Punkt voraussetzt.

Diese theoretischen Resultate bilden die Grundlage zur Konstruktion vieler Optimierungsalgorithmen. Satz 2.9 hilft mit den KKT-Bedingungen ein potentiell lokales Minimum zu finden. Mit dem hinreichenden Kriterium aus Satz 2.15 lässt sich dann überprüfen, ob ein gefundener KKT-Punkt tatsächlich ein Minimum ist. Im nächsten Abschnitt werden Algorithmen vorgestellt, die auf diesen Resultaten basieren.

## 2.2.2 Algorithmen der nichtlinearen Optimierung

In diesem Abschnitt werden die Grundlagen zweier praxisrelevanter Algorithmen vorgestellt, die für die Lösung von nichtlinearen Optimierungsproblemen weit verbreitet sind: das SQP-Verfahren und das IP-Verfahren.

### SQP-Verfahren

Das SQP-Verfahren ist eines der wichtigsten Verfahren zur Lösung von NLPs. In seiner Grundform wurde es erstmalig von Wilson [109] im Jahr 1963 beschrieben und wird seitdem stetig weiterentwickelt (siehe [46]). Es existieren zahlreiche Variationen des Verfahrens und es ist bis heute Gegenstand aktueller Forschung.

Zur Herleitung des SQP-Verfahrens wird zunächst das Optimierungsproblem (2.5) ohne Ungleichungsnebenbedingungen betrachtet:

$$\begin{aligned} \min_{X \in \mathbb{R}^{n_x}} F(X) \\ \text{unter } H_j(X) = 0, \quad j = 1, \dots, m_h \end{aligned} \quad (2.9)$$

Zum Lösen dieses Problems sind die theoretischen Resultate aus Kapitel 2.2.1 hilfreich. Die Lagrange-Funktion des Problems (2.9) ist gegeben durch:

$$L(x, \lambda) = F(X) + \lambda^T H(X)$$

Dabei ist  $\lambda \in \mathbb{R}^{m_h}$ . Werden die LICQ vorausgesetzt, sind die notwendigen Optimalitätsbedingungen durch die KKT-Bedingungen gegeben (siehe Satz 2.9):

$$\Phi(X, \lambda) := \begin{pmatrix} \nabla_x L(X, \lambda) \\ H(X) \end{pmatrix} = 0 \quad (2.10)$$

Ein sinnvoller Ansatz zur Lösung von 2.9 ist nach einem KKT-Punkt zu suchen, indem das nichtlineare Gleichungssystem  $\Phi(X, \lambda) = 0$  gelöst wird. Eines der bekanntesten Verfahren dafür ist das *Newton-Verfahren*. Das Newton-Verfahren in seiner Grundform ist schnell hergeleitet:

Sei  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  eine stetig differenzierbare Funktion und  $f(x) = 0$  das zu lösende Gleichungssystem. Statt es direkt zu lösen, was in der Regel sehr schwierig ist, wird es iterativ gelöst. Sei dazu  $x^{(k)}$  die  $k$ -te Iterierte der Variablen  $x$ . Um die nächste Iterierte zu erhalten, wird  $f$  um den Punkt  $x^{(k)}$  linearisiert:

$$f(x) \approx f(x^{(k)}) + \nabla f(x^{(k)})(x - x^{(k)}) =: f^{(k)}(x)$$

Anstelle des Ursprungsproblems wird das approximierte Problem  $f^{(k)}(x) = 0$  gelöst. Die Lösung dieses linearisierten Problems wird dann als neue Iterierte  $x^{(k+1)}$  verwendet, was durch Umformen die Iterationsvorschrift

$$x^{(k+1)} = x^{(k)} - \nabla f(x^{(k)})^{-1} f(x^{(k)})$$

ergibt. In der Praxis sieht man in der Regel vom kostspieligen Invertieren des Gradienten von  $f$  ab. Stattdessen wird das lineare Gleichungssystem

$$\nabla f(x^{(k)})d^{(k)} = -f(x^{(k)})$$

gelöst und anschließend die neue Iterierte mit

$$x^{(k+1)} = x^{(k)} + d^{(k)}$$

berechnet. Diese Iterationsvorschrift wird durchgeführt, bis ein Abbruchkriterium erreicht ist, bspw.  $\|f(x^{(k)})\| < \varepsilon$  mit geeignetem  $\varepsilon > 0$ .

Das Newton-Verfahren existiert in vielen Variationen. Auch lassen sich Konvergenzaussagen treffen. So kann gezeigt werden, dass die Regularität von  $\nabla f(x^*)$  in einer Nullstelle  $x^*$  von  $f$  zusammen mit der lokalen Lipschitz-Stetigkeit von  $f$  die quadratische Konvergenz des Newtonverfahrens in einer Umgebung um  $x^*$  impliziert. Für ausführlichere Informationen sei an dieser Stelle das Werk von Geiger und Kanzow [44] empfohlen.

Im Folgenden wird das Newton-Verfahren auf das Gleichungssystem (2.10) angewandt, um einen KKT-Punkt von (2.9) zu finden. Die Iterationsvorschrift dafür lautet

$$X^{(k+1)} := X^{(k)} + \Delta X^{(k)}, \quad \lambda^{(k+1)} := \lambda^{(k)} + \Delta \lambda^{(k)}, \quad (2.11)$$

wobei  $(\Delta X^{(k)}, \Delta \lambda^{(k)})$  eine Lösung des Gleichungssystems

$$\nabla \Phi(X^{(k)}, \lambda^{(k)}) \begin{pmatrix} \Delta X^{(k)} \\ \Delta \lambda^{(k)} \end{pmatrix} = -\Phi(X^{(k)}, \lambda^{(k)}) \quad (2.12)$$

ist. Dabei hat die Jacobi-Matrix von  $\Phi$  die Gestalt

$$\nabla \Phi(X, \lambda) = \begin{pmatrix} \mathcal{H}(X, \lambda) & \nabla H(X)^T \\ \nabla H(X) & 0 \end{pmatrix},$$

wobei

$$\mathcal{H}(X, \lambda) := \nabla_X^2 L(X, \lambda) = \nabla^2 F(X) + \sum_{j=1}^{m_h} \nabla^2 H_j \lambda_j$$

die Hesse-Matrix der Lagrange-Funktion und

$$\nabla H(X) := \begin{pmatrix} \nabla H_1(X)^T \\ \vdots \\ \nabla H_{m_h}(X)^T \end{pmatrix}$$

den Gradienten der Jacobi-Matrix der Nebenbedingungen bezeichnet. Die Iterationsvorschrift in (2.11) und (2.12) wird wegen der Anwendung des Newton-Verfahrens auf die Lagrange-Funktion auch *Lagrange-Newton-Verfahren* genannt. In einigen Fällen ist die analytische Berechnung von  $\nabla_X^2 L$  nicht möglich, oder aber die Hesse-Matrix der Lagrange-Funktion ist gar nicht die beste Wahl für das Verfahren, z. B. wenn sie indefinit ist. Häufig finden dann auch Approximationen von  $\nabla_X^2 L$  Verwendung. Im Folgenden sei deshalb auch  $\mathcal{H}(X, \lambda) \approx \nabla_X^2 L(X, \lambda)$  erlaubt.

Das lineare Gleichungssystem (2.12), das mit  $\mathcal{H}^{(k)} := \mathcal{H}(X^{(k)}, \lambda^{(k)})$ ,  $H^{(k)} := H(X^{(k)})$  und  $F^{(k)} := F(X^{(k)})$  die Form

$$\begin{pmatrix} \mathcal{H}^{(k)} & (\nabla H^{(k)})^T \\ \nabla H^{(k)} & 0 \end{pmatrix} \begin{pmatrix} \Delta X^{(k)} \\ \Delta \lambda^{(k)} \end{pmatrix} = \begin{pmatrix} -\nabla F^{(k)} - (\nabla H^{(k)})^T \lambda^{(k)} \\ -H^{(k)} \end{pmatrix} \quad (2.13)$$

hat, wird wegen der darin repräsentierten KKT-Bedingungen auch *KKT-System* genannt. Das System (2.13) ist mit  $\Delta \lambda^{(k)} = \lambda^{(k+1)} - \lambda^{(k)}$  äquivalent zu

$$\begin{pmatrix} \mathcal{H}^{(k)} & (\nabla H^{(k)})^T \\ \nabla H^{(k)} & 0 \end{pmatrix} \begin{pmatrix} \Delta X^{(k)} \\ \lambda^{(k+1)} \end{pmatrix} = \begin{pmatrix} -\nabla F^{(k)} \\ -H^{(k)} \end{pmatrix}. \quad (2.14)$$

Eine interessante Beobachtung ist, dass das Gleichungssystem (2.14) den KKT-Bedingungen des quadratischen Optimierungsproblems

$$\begin{aligned} \min_{\Delta X^{(k)} \in \mathbb{R}^{n_x}} & (\nabla F^{(k)})^T \Delta X^{(k)} + \frac{1}{2} (\Delta X^{(k)})^T \mathcal{H}^{(k)} \Delta X^{(k)} \\ \text{unter} & H_j^{(k)} + (\nabla H_j^{(k)})^T \Delta X^{(k)} = 0, \quad j = 1, \dots, m_h \end{aligned} \quad (2.15)$$

entspricht. Die Lösung eines Schritts des Lagrange-Newton-Verfahrens lässt sich somit auch als Lösung eines quadratischen Subproblems interpretieren, wenn der Lagrange-Multiplikator der Lösung des Subproblems als Approximation für den Multiplikator im nächsten Schritt des Lagrange-Newton-Verfahrens gewählt wird. Bei diesem Subproblem werden die Gleichungsnebenbedingungen linearisiert und die Zielfunktion in gewisser Weise quadratisch approximiert, wobei die Hesse-Matrix der Zielfunktion durch die Hesse-Matrix der Lagrange-Funktion oder einer Approximation ersetzt wird.

Diese Äquivalenz zwischen quadratischem Problem und Lagrange-Newton-Verfahren motiviert die zusätzliche Behandlung von Ungleichungsnebenbedingungen. Das gleichungsrestringierte quadratische Problem (2.15) wird um Ungleichungsnebenbedingungen erweitert:

$$\begin{aligned} \min_{\Delta X^{(k)} \in \mathbb{R}^{n_x}} & (\nabla F^{(k)})^T \Delta X^{(k)} + \frac{1}{2} (\Delta X^{(k)})^T \mathcal{H}^{(k)} \Delta X^{(k)} \\ \text{unter} & G_i^{(k)} + (\nabla G_i^{(k)})^T \Delta X^{(k)} \leq 0, \quad i = 1, \dots, m_g \\ & H_j^{(k)} + (\nabla H_j^{(k)})^T \Delta X^{(k)} = 0, \quad j = 1, \dots, m_h \end{aligned} \quad (2.16)$$

Dabei wird im Lagrange-Newton-Verfahren das Lösen des Gleichungssystems (2.12) durch das Lösen von (2.16) ersetzt. Diese Vorgehensweise zum Lösen des Optimierungsproblems (2.5) wird SQP-Verfahren genannt. Der Name leitet sich davon ab, dass sequentiell quadratische Optimierungsprobleme gelöst werden. Nachfolgend wird der Pseudocode des Verfahrens beschrieben:

**Algorithmus 2.16 (Lokales SQP-Verfahren)**

- (i) Initialisierung: Wähle  $(X^{(0)}, \lambda^{(0)}) \in \mathbb{R}^{n_x} \times \mathbb{R}^m$ ,  $\mathcal{H}^{(0)} \approx \nabla_X^2 L(X^{(0)}, \lambda^{(0)}) \in \mathbb{R}^{n_x \times n_x}$  symmetrisch und setze  $k = 0$ .
- (ii) Ist  $(X^{(k)}, \lambda^{(k)})$  ein KKT-Punkt von (2.5): STOP.
- (iii) Berechne eine Lösung  $\Delta X^{(k)} \in \mathbb{R}^{n_x}$  von (2.16) mit zugehörigem Lagrange-Multiplikator  $\lambda^{(k+1)} \in \mathbb{R}^m$ .
- (iv) Setze  $X^{(k+1)} \leftarrow X^{(k)} + \Delta X^{(k)}$  und  $k \leftarrow k + 1$ , wähle  $\mathcal{H}^{(k)} \approx \nabla_X^2 L(X^{(k)}, \lambda^{(k)}) \in \mathbb{R}^{n_x \times n_x}$  symmetrisch, gehe zu (ii).

Da das SQP-Verfahren auf dem Newton-Verfahren basiert, stellt sich die Frage, ob ähnliche Konvergenzaussagen auch auf das SQP-Verfahren zutreffen. Diese Frage beantwortet der folgende Satz.

**Satz 2.17**

Sei in Algorithmus 2.16  $\mathcal{H}^{(k)} = \nabla_X^2 L(X^{(k)}, \lambda^{(k)})$  für alle  $k$ . Des Weiteren sei  $(X^*, \lambda^*) \in \mathbb{R}^{n_x} \times \mathbb{R}^m$  ein KKT-Punkt von (2.5), der zusätzlich die folgenden Bedingungen erfüllt:

- (i)  $X^*$  genügt den LICQ.
- (ii) Es gilt die strikte Komplementarität:  $G_i(X^*) + \lambda_i^* \neq 0$ ,  $i = 1, \dots, m_g$ .
- (iii) Es gilt die hinreichende Bedingung zweiter Ordnung (2.8).

Dann existiert ein  $\varepsilon > 0$ , sodass für beliebige Startwerte  $(X^{(0)}, \lambda^{(0)}) \in U_\varepsilon(X^*, \lambda^*)$  alle quadratischen Probleme aus Algorithmus 2.16 eine eindeutige Lösung  $\Delta X^{(k)}$  mit eindeutigem Multiplikator  $\lambda^{(k+1)}$  besitzen. Zudem konvergiert die Folge  $\{(X^{(k)}, \lambda^{(k)})\}$  superlinear gegen  $(X^*, \lambda^*)$ . Sind  $\nabla_X^2 L$ ,  $\nabla_X^2 G_i$ , ( $i = 1, \dots, m_g$ ) und  $\nabla_X^2 H_j$ , ( $j = 1, \dots, m_h$ ) zusätzlich lokal Lipschitz-stetig, so konvergiert  $\{(X^{(k)}, \lambda^{(k)})\}$  sogar quadratisch gegen  $(X^*, \lambda^*)$ .

**Beweis:** siehe [45].

Dieses Konvergenzresultat ist eher von theoretischem Interesse, weil es lediglich eine lokale Konvergenz garantiert. In der Praxis ist man daher eher an einer globalen Variante des SQP-Algorithmus interessiert. Von großem Interesse ist auch, eine geeignete Approximation von  $\nabla_X^2 L(X^{(k)}, \lambda^{(k)})$  zu finden. Es existieren eine Vielzahl von Modifikationen und Erweiterungen für das SQP-Verfahren, die unter anderem diese beiden



Punkte adressieren. An dieser Stelle soll nicht im Detail auf die vielen Variationen eingegangen werden. Es seien abschließend nur ein paar Anmerkungen dazu gemacht:

- Für die Wahl von  $\mathcal{H}^{(k)}$  wird in der Praxis häufig die *modifizierte BFGS<sup>4</sup>-Update-Formel* nach Powell [82] verwendet. Diese garantiert, dass alle  $\mathcal{H}^{(k)}$  positiv definit sind, welches eine wichtige Voraussetzung für die eindeutige Lösbarkeit des quadratischen Subproblems (2.16) ist. Auch lassen sich Konvergenzresultate für diese Wahl der Matrizen  $\mathcal{H}^{(k)}$  herleiten (siehe [81]).
- Die Lösung des quadratischen Subproblems ist in der Regel nicht trivial. In der Praxis kommen hier verschiedene Ansätze zur Anwendung, z. B. die Strategie der aktiven Menge oder IP-Verfahren. Für mehr Details sei hier auf [76] verwiesen.
- Um für den Algorithmus 2.16 nicht nur lokale, sondern auch globale Konvergenz zu gewährleisten, gibt es verschiedene Strategien. Ein weit verbreiteter Ansatz ist die Verwendung einer *Merit-Funktion*<sup>5</sup> innerhalb einer *Armijo-Schrittweitenstrategie*. Für weitere Informationen seien hier [45] und [22] empfohlen.

## WORHP

Eine moderne, praxisnahe Implementierung des SQP-Algorithmus' ist die Softwarebibliothek *WORHP*<sup>6</sup> [24, 104]. Sie ist entwickelt für das Lösen von hochdimensionalen, nichtlinearen und ungleichungsrestringierten Optimierungsproblemen. Dabei wird insbesondere die schwache Struktur des zugrundeliegenden Optimierungsproblems ausgenutzt, um einen hinsichtlich Rechengeschwindigkeit und Speicherbedarf effizienten Algorithmus zu ermöglichen. Hochdimensional bedeutet, dass Probleme mit mehreren Millionen Optimierungsvariablen und Nebenbedingungen mit vertretbarem Speicherbedarf und vertretbarer Rechenzeit gelöst werden können. WORHP verwendet eine Vielzahl von Erweiterungen des SQP-Verfahrens und wird stetig weiterentwickelt (siehe [41–43]). Neben WORHP existieren die Softwareerweiterungen *TransWORHP* [66] und *WORHP Zen* [68]. *TransWORHP* verwendet ein direktes Verfahren für das Lösen von Optimalsteuerungsproblemen mit gewöhnlichen Differentialgleichungen und *WORHP Zen* bietet eine parametrische Sensitivitätsanalyse [21] für ein effizientes Update gestörter Optimierungs- oder Optimalsteuerungsprobleme.

Für die Verwendung von WORHP werden die ersten Ableitungen  $\nabla F$ ,  $\nabla G$  und  $\nabla H$  sowie die zweite Ableitung  $\nabla_{\mathcal{X}}^2 L$  benötigt. Dem Benutzer wird neben der Möglichkeit, diese analytisch anzugeben, die Wahl gelassen, Routinen zur numerischen Berechnung der benötigten Ableitungsinformationen zu verwenden. Beim SQP-Verfahren entstehen in der Regel auf der untersten Lösungsebene lineare Gleichungssysteme. Es besteht die Möglichkeit, diese Ebene zu parallelisieren. WORHP ist für akademische Zwecke frei

<sup>4</sup>Das BFGS ist nach Broyden, Fletcher, Goldfarb und Shanno benannt.

<sup>5</sup>Die Merit-Funktion wird auch *Penalty-Funktion* genannt.

<sup>6</sup>WORHP steht für „We Optimize Really Huge Problems“.

verfügbar und besitzt verschiedene Schnittstellen, wie z. B. eine MATLAB- und C++-Schnittstelle. Ursprünglich für die Europäische Weltraumorganisation ESA entwickelt, wird WORHP heute von hunderten Anwendern weltweit verwendet. Für nähere Informationen wird an dieser Stelle auf die offizielle Homepage [111] verwiesen.

Die Optimierungssoftware WORHP ist ein zentrales Werkzeug dieser Arbeit, mit dem eine Vielzahl der hier auftretenden Optimierungsaufgaben gelöst werden. An entsprechenden Stelle wird auf die Verwendung von WORHP eingegangen.

## IP-Verfahren

Ein weiteres Verfahren zum Lösen von NLPs ist das IP-Verfahren. Der Abschnitt über dieses orientiert sich an der Ausarbeitung von Forsgren et al. [39], die einen guten Überblick über Verfahren im Kontext von nichtlinearer Optimierung gibt. Als Ursprung der IP-Verfahren werden häufig die sogenannten Barriere-Verfahren genannt. Diese wurden lange Zeit hauptsächlich im Kontext von linearer Optimierung betrachtet und standen im Schatten des populäreren Simplex-Verfahrens. Erst der IP-Algorithmus von Karmarkar von 1984 [62] verhalf den IP-Verfahren zu besonderem Forschungsinteresse. Seitdem wurden das Verfahren und seine Variationen fortlaufend weiterentwickelt und sie fanden auch ihre Verwendung in der nichtlinearen Optimierung.

An dieser Stelle soll das IP-Verfahren, genauer gesagt das Primal-Duale IP-Verfahren, mit einem Barriere-Verfahren motiviert werden. Zunächst wird nur der ungleichungsrestringierte Fall von (2.5) betrachtet:

$$\begin{aligned} \min_{X \in \mathbb{R}^{n_x}} F(X) \\ \text{unter } G_i(X) \geq 0, \quad i = 1, \dots, m_g \end{aligned} \quad (2.17)$$

Im Folgenden sei dabei die zweimalige stetige Differenzierbarkeit von  $F$  und  $G$  vorausgesetzt. An dieser Stelle sollen die Nebenbedingungen o. B. d. A. größer-gleich 0 sein, was die folgenden Betrachtungen vereinfacht. Dies kann aber mit  $G_i(X) \geq 0 \Leftrightarrow -G_i(X) \leq 0$  leicht in die Form von (2.5) gebracht werden. Die Basis für die Barriere-Verfahren bildet ein Barriere-Term  $\hat{B}(X)$ , der nur von den Nebenbedingungen abhängt. Zusätzlich konvergiert der Term gegen unendlich für eine beliebige Folge von inneren Punkten der zulässigen Menge  $\Sigma$ , die gegen den Rand der zulässigen Menge  $\partial\Sigma$  konvergiert, also

$$\lim_{k \rightarrow \infty} \hat{B}(X_k) = \infty, \quad X_k \in \text{int}(\Sigma), \quad \lim_{k \rightarrow \infty} X_k \in \partial\Sigma, \quad k \in \mathbb{N}^{\geq 0}.$$

Die wohl gebräuchlichste Wahl für  $\hat{B}$  ist der logarithmische Barriere-Term

$$\hat{B}_{\log}(X) := - \sum_{i=1}^{m_g} \ln G_i(X),$$

welcher in der sogenannten Barriere-Funktion

$$B(X, \mu) := F(X) + \mu \hat{B}_{\log}(X) = F(X) - \mu \sum_{i=1}^{m_g} \ln G_i(X)$$

Verwendung findet, wobei  $\mu > 0$  der Barriere-Parameter ist. Die Barriere-Verfahren lösen anstelle von (2.17) eine Sequenz von Minimierungsaufgaben

$$\min_{X \in \Sigma} B(X, \mu). \quad (2.18)$$

Die zentrale Idee ist, dass für kleine  $\mu$  die Barriere-Funktion der Zielfunktion ähnelt. Dies motiviert die Erwartung, dass für  $\mu \rightarrow 0$  die Lösungen von (2.18) gegen die Lösung von (2.17) konvergieren. Tatsächlich lässt sich dies unter gewissen Voraussetzungen beweisen (siehe [39]). Für positive  $\mu$  sind die Lösungen von (2.18) innere Punkte von  $\Sigma$ , da auf dem Rand der zulässigen Menge

$$B(X, \mu) = \infty, \quad X \in \partial \Sigma \quad (2.19)$$

gilt. Dabei ist der Rand der zulässigen Menge eine Art Barriere, die wegen (2.19) nicht passiert werden kann. Eine notwendige Bedingung für die Lösung von (2.18) ist

$$\begin{aligned} \nabla_X B(X, \mu) &= \nabla F(X) - \sum_{i=1}^{m_g} \frac{\mu}{G_i(X)} \nabla G_i(X) = 0 \\ \Leftrightarrow: \nabla F(X) - \sum_{i=1}^{m_g} \lambda_{\mu, i} \nabla G_i(X) &= 0. \end{aligned} \quad (2.20)$$

Der in (2.20) definierte Multiplikator

$$\lambda_{\mu, i} = \frac{\mu}{G_i(X)} \Leftrightarrow \lambda_{\mu, i} G_i(X) = \mu \quad (2.21)$$

wird Barriere-Multiplikator genannt. Wegen  $\mu > 0$  und  $X \in \Sigma \Leftrightarrow G_i(X) \geq 0$  ist  $\lambda_{\mu, i} > 0$  für  $i = 1, \dots, m_g$ . Dies zusammen mit (2.20) und (2.21) erinnert in gewisser Weise an die KKT-Bedingungen<sup>7</sup> aus Definition 2.6, wobei die Komplementaritätsbedingung (2.21) durch den Wert  $\mu > 0$  gestört wird. Mit kleiner werdendem  $\mu$  nähert sich dies aber der herkömmlichen Komplementarität an.

Die ausschließliche Betrachtung der primalen Variablen  $X$  durch das Lösen von (2.20) hat entscheidende Nachteile (siehe [39]). Ein Ansatz, der sich in der Praxis bewährt hat, ist die unabhängige Betrachtung von primalen und dualen Variablen. Mit der dualen

<sup>7</sup>Die Bedingung (2.20) enthält hier im Gegensatz zu Definition 2.6 ein Minus, da zu Beginn des Kapitels  $G_i(X) \geq 0$  anstatt  $G_i(X) \leq 0$  gefordert wurde. Für diese alternative Formulierung der KKT-Bedingungen siehe [76].

Variablen  $\lambda \in \mathbb{R}^{m_g}$  und der Hinzunahme von (2.21) als Bedingung an die duale Variable lässt sich eine notwendige Bedingung zur Lösung von (2.18) auch mit

$$\Phi^\mu(X, \lambda) := \begin{pmatrix} \nabla F(X) - \nabla G(X)^T \lambda \\ \mathbf{G}(X) \lambda - \mu e \end{pmatrix} = 0 \quad (2.22)$$

angeben. Dabei ist

$$\mathbf{G}(X) := \begin{pmatrix} G_1(X) & & 0 \\ & \ddots & \\ 0 & & G_{m_g}(X) \end{pmatrix} \in \mathbb{R}^{m_g \times m_g} \quad \text{und} \quad e = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^{m_g}.$$

Wendet man zum Lösen des Gleichungssystems (2.22) das Newton-Verfahren an, erhält man das Primal-Duale IP-Verfahren für ungleichungsrestringierte Optimierungsprobleme. Ein Newton-Schritt ist dann gegeben durch

$$\begin{aligned} \nabla \Phi^\mu(X, \lambda) \begin{pmatrix} \Delta X \\ \Delta \lambda \end{pmatrix} &= -\Phi^\mu(X, \lambda) \\ \Leftrightarrow \begin{pmatrix} \mathcal{H}(X, \lambda) & -\nabla G(X)^T \\ \Lambda \nabla G(X)^T & \mathbf{G}(X) \end{pmatrix} \begin{pmatrix} \Delta X \\ \Delta \lambda \end{pmatrix} &= - \begin{pmatrix} \nabla F(X) - \nabla G(X)^T \lambda \\ \mathbf{G}(X) \lambda - \mu e \end{pmatrix} \end{aligned}$$

mit

$$\Lambda := \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_{m_g} \end{pmatrix} \in \mathbb{R}^{m_g \times m_g} \quad \text{und} \quad \mathcal{H}(X, \lambda) := \nabla^2 F(X) - \sum_{i=1}^{m_g} \lambda_i \nabla^2 G_i(X).$$

Um das allgemeine Optimierungsproblem (2.5), welches zusätzlich Gleichungsnebenbedingungen enthält, zu lösen, wird die Minimierungsaufgabe (2.18) um diese erweitert:

$$\begin{aligned} \min_{X \in \Sigma} & B(X, \mu) \\ \text{unter} & H_j(X) = 0, \quad j = 1, \dots, m_h \end{aligned}$$

Die Anwendung der notwendigen KKT-Bedingungen für den gleichungsrestringierten Fall<sup>8</sup> liefert zusammen mit der dualen Variable  $\lambda_G \in \mathbb{R}^{m_g}$  und dem Lagrange-Multiplikator  $\lambda_H \in \mathbb{R}^{m_h}$

$$\Phi^\mu(X, \lambda) := \begin{pmatrix} \nabla F(X) - \nabla G(X)^T \lambda_G - \nabla H(X)^T \lambda_H \\ \mathbf{G}(X) \lambda_G - \mu e \\ H(X) \end{pmatrix} = 0. \quad (2.23)$$

<sup>8</sup>Man verwende Definition 2.6 mit  $m_g = 0$ .

Auch dieses Gleichungssystem kann mit dem Newton-Verfahren gelöst werden. Ein Newton-Schritt angewendet auf (2.23) ergibt dann das lineare Gleichungssystem

$$\begin{pmatrix} \mathcal{H}(X, \lambda) & -\nabla G(X)^T & -\nabla H(X)^T \\ \Lambda_G \nabla G(X)^T & \mathbf{G}(X) & 0 \\ \nabla H(X) & 0 & 0 \end{pmatrix} \begin{pmatrix} \Delta X \\ \Delta \lambda_G \\ \Delta \lambda_H \end{pmatrix} = - \begin{pmatrix} \nabla F(X) - \nabla G(X)^T \lambda_G - \nabla H(X)^T \lambda_H \\ \mathbf{G}(X) \lambda_G - \mu e \\ H(X) \end{pmatrix}.$$

Hierbei ist  $\mathcal{H}(X, \lambda) := \nabla^2 F(X) - \sum_{i=1}^{m_g} \lambda_{G,i} \nabla^2 G_i(X) - \sum_{i=1}^{m_h} \lambda_{H,i} \nabla^2 H_i(X)$ . Das iterative Lösen von (2.23) für sukzessiv kleiner werdendes  $\mu$  ergibt das Primal-Duale IP-Verfahren in seiner Grundform und kann formal wie folgt aufgeschrieben werden:

**Algorithmus 2.18 (Primal-Duales IP-Verfahren)**

- (i) Initialisierung: Wähle  $(X^{(0)}, \lambda_G^{(0)}, \lambda_H^{(0)}) \in \mathbb{R}^{n_x} \times \mathbb{R}^{m_g} \times \mathbb{R}^{m_h}$  mit  $X^{(0)} \in \Sigma$ ,  $\mu^{(0)} > 0$  und setze  $k = 0$ .
- (ii) Ist  $(X^{(k)}, \lambda_G^{(k)}, \lambda_H^{(k)})$  ein KKT-Punkt von (2.5): STOP.
- (iii) Berechne eine Lösung  $(X^{(k+1)}, \lambda_G^{(k+1)}, \lambda_H^{(k+1)})$  von (2.23) mit dem Newton-Verfahren.
- (iv) Wähle  $\mu^{(k+1)}$  mit  $0 < \mu^{(k+1)} < \mu^{(k)}$ , setze  $k \leftarrow k + 1$ , gehe zu (ii).

Auch für dieses Verfahren gibt es vielfältige Modifikationen und Erweiterungen. Wie beim SQP-Verfahren widmen sich einige der Globalisierung der Konvergenz und der Approximation der Hesse-Matrix. Von besonderem Interesse bei praxisnahen Varianten ist eine Liniensuche für das Newton-Verfahren. Diese stellt neben der globalen Konvergenz unter anderem sicher, dass die Iterierten  $(X^{(k)}, \lambda_G^{(k)})$  im zulässigen Bereich bleiben, also, dass  $G(X^{(k)}) > 0$  und  $\lambda_G^{(k)} > 0$  gilt. Für weitere Informationen werden [39], [29] und [105] empfohlen.

**Ipopt**

Eine praxisnahe, frei verfügbare Implementierung des Primal-Dualen IP-Verfahren ist *Ipopt* [105]. Das Software-Paket ist geeignet für die hochdimensionale, nichtlineare Optimierung. Insbesondere eignet es sich für das Lösen von NLPs (2.5). Für Details wird auf die Homepage [60] verwiesen.

Ipopt wird in dieser Arbeit teilweise für die Optimierung eingesetzt. An entsprechender Stelle wird darauf hingewiesen.

### Gleichungsnebenbedingungen und Boxschränken in WORHP und Ipopt

Die NLP-Löser WORHP und Ipopt verwenden anstellen von Ungleichungsnebenbedingungen und Gleichungsnebenbedingungen nur Ungleichungsnebenbedingungen der Form:

$$G_L \leq G(X) \leq G_U \quad (2.24)$$

Dabei sind  $G_L \in \mathbb{R}^{m_g}$  und  $G_U \in \mathbb{R}^{m_g}$  obere bzw. untere Schranke der Nebenbedingungen. Diese Formulierung stellt kein Problem dar, weil die Nebenbedingungen aus (2.5) äquivalent in die Form (2.24) überführt werden können:

$$\begin{aligned} G(X) \leq 0 &\Leftrightarrow -\infty \leq G(X) \leq 0 \\ H(X) = 0 &\Leftrightarrow 0 \leq H(X) \leq 0 \end{aligned}$$

In dieser Arbeit werden weiterhin die Nebenbedingungen

$$\begin{aligned} G(X) &\leq 0 \\ H(X) &= 0 \end{aligned}$$

verwendet. Die äquivalenten Formulierungen wurden bei der Implementierung berücksichtigt.

Bei der Herleitung der Theorie und der Verfahren für NLPs wurden Boxschränken als allgemeine Nebenbedingungen aufgefasst. WORHP und Ipopt können diese in der Form

$$X_L \leq X \leq X_U$$

jedoch besonders effizient berücksichtigen. Deshalb werden sie in der Anwendung von den allgemeinen Nebenbedingungen gesondert betrachtet.

### Abbruchkriterien in WORHP und Ipopt

Die beiden vorgestellten Algorithmen 2.16 und 2.18 zum Lösen von NLPs sind sehr allgemeine Varianten der jeweiligen Verfahren. In der Praxis müssen insbesondere die Abbruchkriterien numerisch formuliert werden. Sie prüfen im Wesentlichen, ob eine mögliche Lösung des Problems ein KKT-Punkt ist. Dabei wird numerisch die Zulässigkeit, das Verschwinden des Gradienten der Lagrange-Funktion sowie die Komplementarität geprüft. In WORHP und Ipopt gibt es für die Prüfung dieser Bedingungen numerische Toleranzen. Diese Werte werden häufig auf  $10^{-6}$  gesetzt. Dies hat sich auch innerhalb dieser Arbeit bewährt.

## 2.3 Finite-Elemente-Methode und Softwarebibliothek deal.II

Die FEM ist ein weitverbreiteter Ansatz zur numerischen Lösung von PDEs. In vielen Anwendungen erhält die FEM den Vorzug gegenüber den Finiten-Differenzen-Verfahren, da insbesondere komplexe Geometrien leichter zu handhaben sind. Die Theorie fußt auf der schwachen Formulierung der zu lösenden PDE und führt zu der Bestimmung sogenannter schwacher Lösungen. Diese unterscheiden sich im Vergleich zu den klassischen Lösungen, deren Approximation man z. B. mit den Finiten-Differenzen-Verfahren erhält, darin, dass diese auf einer Nullmenge nicht differenzierbar sein müssen. Die Approximation der schwachen Problemstellung in einem endlich-dimensionalen Funktionenraum und anschließende numerische Integration führt auf ein Gleichungssystem, welches die gesuchten Koeffizienten der Ansatzfunktion als Lösungsvariablen enthält. Im Folgenden wird die Funktionsweise der FEM kurz skizziert.

### FEM

Der Abschnitt über die FEM orientiert sich an [16], [85] und [90].

Die Methodik der FEM soll am Beispiel der Poisson-Gleichung mit homogenen Dirichlet-Randbedingungen, d. h.

$$\begin{aligned} -\Delta u_\phi &= f_\phi && \text{in } \Omega \\ u_\phi &= 0 && \text{auf } \Gamma, \end{aligned} \tag{2.25}$$

erläutert werden. Dabei ist  $f_\phi : \Omega \rightarrow \mathbb{R}$  die Funktion des Quellterms und  $u_\phi : \Omega \rightarrow \mathbb{R}$  die gesuchte Lösungsfunktion.  $\Omega \subset \mathbb{R}^{n_d}$  ist ein Gebiet, auf dem die PDE definiert ist, und dessen Rand  $\partial\Omega = \Gamma$  ein Lipschitz-Rand sei, d. h. er lässt sich durch eine endliche Anzahl von Lipschitz-stetigen Funktionen darstellen. Für die weiteren Betrachtungen sei zudem  $L_2(\Omega, \mathbb{R}^{n_p})$  der Raum aller Funktionen, die von  $\Omega$  nach  $\mathbb{R}^{n_p}$  abbilden und deren Quadrat über  $\Omega$  Lebesgue-integrierbar ist. Der Raum der beliebig oft differenzierbaren Funktionen sei durch  $C^\infty(\Omega, \mathbb{R}^{n_p})$  gegeben und durch  $C_0^\infty(\Omega, \mathbb{R}^{n_p})$  der Unterraum der Funktionen, die nur in einer kompakten Teilmenge von  $\Omega$  ungleich 0 sind. Zudem gelte  $L_2(\Omega) := L_2(\Omega, \mathbb{R})$  sowie  $C^\infty(\Omega) := C^\infty(\Omega, \mathbb{R})$  und  $C_0^\infty(\Omega) := C_0^\infty(\Omega, \mathbb{R})$ .

Eine Lösung von (2.25) im Sinne der klassischen Differenzierbarkeit lässt sich nur unter sehr starken Voraussetzungen garantieren. Deshalb sind häufig die Konzepte der schwachen Ableitung und Räume schwach differenzierbarer Funktionen hilfreich. Sie lassen sich wie folgt definieren:

#### Definition 2.19 (schwache Ableitung)

Sei  $u_\phi \in L_2(\Omega)$ . Eine Funktion  $w_\phi = \partial^p u_\phi$  wird schwache Ableitung von  $u_\phi$  genannt,

falls  $w_\phi \in L_2(\Omega)$  und

$$\int_{\Omega} v_\phi w_\phi dx = (-1)^{|p|} \int_{\Omega} \partial^p v_\phi u_\phi dx \quad \text{für alle } v_\phi \in C_0^\infty(\Omega)$$

gilt. Dabei ist  $p = (p_1, \dots, p_{n_d}) \in \mathbb{N}_0^{n_d}$ ,  $|p| = \sum_{i=1}^{n_d} p_i$  und  $\partial^p = \frac{\partial^{|p|}}{\partial x_1^{p_1} \dots \partial x_{n_d}^{p_{n_d}}}$ .

**Definition 2.20 (Sobolev-Räume)**

Sei  $m \in \mathbb{N}_0$ . Der Sobolev-Raum  $H^m(\Omega, \mathbb{R}^{n_p})$  ist die Menge aller Funktionen  $u_\phi \in L_2(\Omega, \mathbb{R}^{n_p})$ , deren schwache Ableitungen  $\partial^p u_\phi$  für alle  $|p| \leq m$  existieren. Für  $m = 1$  sei zudem der Sobolev-Raum mit Nullrand definiert durch:

$$H_0^1(\Omega, \mathbb{R}^{n_p}) := \left\{ v_\phi \in H^1(\Omega, \mathbb{R}^{n_p}) : v_\phi|_\Gamma = 0 \right\}$$

Weiter sei  $H^m(\Omega) := H^m(\Omega, \mathbb{R})$  und  $H_0^1(\Omega) := H_0^1(\Omega, \mathbb{R})$ .

Unter der Voraussetzung  $f_\phi \in L_2(\Omega)$  kann gezeigt werden, dass eine eindeutige schwache Lösung  $u_\phi \in H_0^1(\Omega)$  für das Poisson-Problem (2.25) existiert [16]<sup>9</sup>. Um die PDE numerisch mit der FEM zu lösen, wird sie zunächst in die schwache Form überführt. Sie wird dafür mit einer Testfunktion  $v_\phi \in H_0^1(\Omega)$  multipliziert und über  $\Omega$  integriert:

$$-\int_{\Omega} \Delta u_\phi v_\phi dx = \int_{\Omega} f_\phi v_\phi dx$$

Mit  $\Delta u_\phi = \text{div } \nabla u_\phi$  und dem Gaußschen Integralsatz wird daraus:

$$\int_{\Omega} \nabla u_\phi \cdot \nabla v_\phi dx - \int_{\Gamma} \nabla u_\phi \cdot \vec{n} v_\phi ds = \int_{\Omega} f_\phi v_\phi dx$$

Hierbei bezeichnet  $\vec{n}$  die äußere Normale von  $\Omega$ . Da  $v_\phi|_\Gamma = 0$  gilt, ist das Randintegral 0. Die Lösung  $u_\phi$  der PDE muss also

$$\int_{\Omega} \nabla u_\phi \cdot \nabla v_\phi dx = \int_{\Omega} f_\phi v_\phi dx \quad \text{für alle } v_\phi \in H_0^1(\Omega) \tag{2.26}$$

erfüllen. Für die FEM wird dieses unendlich-dimensionale Problem durch die Wahl eines endlich-dimensionalen Lösungsraumes  $V_h$  approximiert. Zunächst wird dafür das Gebiet  $\Omega$  in ein diskretes Gitter  $\Omega_h$  zerlegt. Die Zellen des Gitters sind im dreidimensionalen Raum beispielsweise Tetraeder oder Hexaeder. Mit Hilfe der Eckpunkte und bei Bedarf Punkte der Kanten und im Innern der Zellen lassen sich Basisfunktionen  $\varphi_1, \dots, \varphi_{n_h}$  definieren. Diese werden so gewählt, dass sie jeweils nur auf wenigen dieser Punkte Werte

<sup>9</sup>Wesentlich für den Beweis ist der Satz von Lax-Milgram.



ungleich 0 annehmen, auf  $\Gamma$  den Wert 0 annehmen und stückweise polynomial sind. Häufig werden Lagrange-Polynome verwendet. Es gilt dann  $V_h \subset H_0^1(\Omega)$  und man spricht von sogenannten konformen finiten Elementen. Anstelle von (2.26) wird die diskrete Problemstellung

$$\int_{\Omega} \nabla u_{\phi,h} \nabla v_{\phi,h} dx = \int_{\Omega} f_{\phi} v_{\phi,h} dx \quad \text{für alle } v_{\phi,h} \in V_h \quad (2.27)$$

betrachtet und eine Lösung  $u_{\phi,h} \in V_h$  gesucht. Durch  $V_h \subset H_0^1(\Omega)$  lassen sich Aussagen über die Eindeutigkeit und die Lösbarkeit des kontinuierlichen Problems auch auf das diskrete Problem übertragen. Zudem hängt die Approximationsgüte von  $u_{\phi,h}$  an  $u_{\phi}$  davon ab, wie gut sich  $u_{\phi}$  im diskreten Raum  $V_h$  approximieren lässt. Im Allgemeinen kann man erwarten, dass die Genauigkeit mit der Wahl eines feineren Gitters steigt. Um (2.27) numerisch zu lösen, genügt es, ein  $u_{\phi,h} \in V_h$  mit

$$\int_{\Omega} \nabla u_{\phi,h} \nabla \varphi_i dx = \int_{\Omega} f_{\phi} \varphi_i dx, \quad i = 1, \dots, n_h \quad (2.28)$$

zu finden, da für alle  $v_{\phi,h} \in V_h$  ein  $v \in \mathbb{R}^{n_h}$  mit  $v_{\phi,h} = \sum_{i=1}^{n_h} v_i \varphi_i$  existiert und damit

$$\begin{aligned} (2.28) &\Leftrightarrow \int_{\Omega} \nabla u_{\phi,h} \nabla \varphi_i dx - \int_{\Omega} f_{\phi} \varphi_i dx = 0, \quad i = 1, \dots, n_h \\ &\Rightarrow \sum_{i=1}^{n_h} v_i \left( \int_{\Omega} \nabla u_{\phi,h} \nabla \varphi_i dx - \int_{\Omega} f_{\phi} \varphi_i dx \right) = 0 \\ &\Leftrightarrow \int_{\Omega} \nabla u_{\phi,h} \nabla \left( \sum_{i=1}^{n_h} v_i \varphi_i \right) dx - \int_{\Omega} f_{\phi} \sum_{i=1}^{n_h} v_i \varphi_i dx = 0 \\ &\Leftrightarrow (2.27) \end{aligned}$$

folgt. Auch für die gesuchte Lösung  $u_{\phi,h} \in V_h$  existiert ein  $u \in \mathbb{R}^{n_h}$  mit  $u_{\phi,h} = \sum_{j=1}^{n_h} u_j \varphi_j$ . Damit ist (2.28) äquivalent zu

$$\sum_{j=1}^{n_h} u_j \int_{\Omega} \nabla \varphi_j \nabla \varphi_i dx = \int_{\Omega} f_{\phi} \varphi_i dx, \quad i = 1, \dots, n_h.$$

Dies wiederum wird durch einfache numerische Integration der Basisfunktionen bzw. deren Gradienten über die Zellen des Gitters zu einem linearen Gleichungssystem:

$$\mathbf{A} u = \mathbf{b}(f_{\phi})$$

Dabei ist  $\mathbf{A} \in \mathbb{R}^{n_h \times n_h}$  und  $\mathbf{b}(f_{\phi}) \in \mathbb{R}^{n_h}$ , wobei  $\mathbf{b}$  von  $f_{\phi}$  abhängt. Es gilt  $\mathbf{A}_{ij} = \int_{\Omega} \nabla \varphi_j \nabla \varphi_i dx = 0$ , wenn die beiden Mengen, auf denen  $\varphi_j$  und  $\varphi_i$  ungleich 0 sind,

disjunkt sind. Dies ist durch die Wahl der Basis bei den meisten Paaren von Basisfunktionen der Fall, weshalb  $\mathbf{A}$  schwach besetzt ist. Durch Ausnutzen der Struktur wird das numerische Lösen des diskreten Problems effizient. Das Gleichungssystem kann entweder mit numerischen Verfahren zum Lösen von Gleichungssystemen gelöst werden oder aber als Nebenbedingung zum Optimierungsproblem hinzugefügt werden.

Im Kontext der allgemeinen Formulierungen (2.1) und (2.2) wird durch die FEM aus dem Modell  $P_\phi$  das Gleichungssystem  $P$ . Entsprechend müssen auch für die Zielfunktion  $F_\phi$  sowie die Nebenbedingungen  $G_\phi$  und  $H_\phi$  diskrete Varianten  $F, G$  und  $H$  definiert werden. Dies ist sehr problemspezifisch und geschieht bei der Formulierung der konkreten Optimierungsprobleme in den Kapiteln 4 bis 6. Dort werden auch komplexere PDEs als die Poisson-Gleichung behandelt. In Kapitel 4 wird die zeitabhängige Wärmeleitungsgleichung mit Neumann-Randbedingungen verwendet. Dadurch werden Werte für den Gradienten auf dem Rand vorgeben. In Kapitel 5 wird eine nichtlineare Variante der Wärmeleitungsgleichung betrachtet. Durch die Diskretisierung mit der FEM entsteht daraus ein nichtlineares Gleichungssystem. Zuletzt wird in Kapitel 6.1 mit den Lamé-Navier-Gleichungen noch die PDE einer vektorwertigen Funktion gelöst.

Es existieren zur FEM auch Resultate zur Abschätzung des Diskretisierungsfehlers [16]. Innerhalb dieser Arbeit steht jedoch die Methodik sowie die praktikable Umsetzung der Optimierung im Vordergrund, weshalb auf eine Fehlerabschätzung verzichtet wird.

## deal.II

Die Software *deal.II* [5, 7] ist eine Open-Source-Bibliothek, die die Anwendung der FEM ermöglicht. Mit *deal.II* lassen sich das diskrete Gitter erstellen und die Ansatzfunktionen verwalten. Aus der schwachen Form einer PDE sowie Rand- und Anfangswerten generiert die Software das diskrete Gleichungssystem, für welches auch Routinen zur Lösung bereitgestellt werden. Die Bibliothek zeichnet sich durch eine umfangreiche Dokumentation und Tutorials aus. Innerhalb dieser Arbeit wird *deal.II* im SAND-Ansatz zum Aufstellen des hochdimensionalen, schwach besetzten Gleichungssystems verwendet, um dieses als Nebenbedingung der Optimierung zu lösen. Beim NAND-Ansatz wird das Gleichungssystem nicht der Optimierung hinzugefügt. Im nichtlinearen Fall wird es mit dem Newton-Verfahren extern gelöst. Die bei den Newton-Schritten zu lösenden linearen Gleichungssysteme werden mit dem Paket *UMFPACK* gelöst, das in *deal.II* enthalten ist. Für weitere Informationen zur Software *deal.II* wird auf die Homepage [30] verwiesen.

# Kapitel 3

## Spanende Trockenbearbeitung - Modellierung und Optimierungskonzept

### Inhalt

---

3.1	Modellierung der Prozesseinflüsse beim Trockenfräsen . . . . .	36
3.1.1	Eigenspannungsbedingte Effekte - Sub-Modell 1 . . . . .	37
3.1.2	Thermische Effekte - Sub-Modell 2 . . . . .	42
3.2	Vorgehensweise zur Minimierung von Formabweichungen beim Fräsen	45
3.3	Modellierung der Prozesseinflüsse beim Bohren . . . . .	47
3.4	Versuchsreihen für die empirischen Modelle . . . . .	48
3.4.1	Experimente für Sub-Modell 1 . . . . .	48
3.4.2	Experimente für Sub-Modell 2 . . . . .	49
3.4.3	Experimente für das thermische Modell des Bohrens . . . . .	49

---

Es existieren ökonomische und ökologische Gründe, weshalb in der spanenden Bearbeitung mitunter auf das Kühlschmieren verzichtet wird. Durch die erhöhte Belastung ist jedoch eine geeignete Kompensation notwendig, um die gewünschte Qualität bei der Fertigung sicherzustellen. Dies trifft insbesondere auf die Herstellung hochpräziser Bauteile zu. Das Schwerpunktprogramm (SPP) 1480 „Modellierung, Simulation und Kompensation von thermischen Bearbeitungseinflüssen für komplexe Zerspanprozesse“ [13, 38] wurde von der DFG gefördert, um diese Thematik zu erforschen. Die vorliegende Arbeit ist im Rahmen des Projekts „Kompensationsplanung thermischer Prozesseinflüsse beim Trockenfräsen und Trockenbohren“ entstanden.

Als Demonstratorbauteil diente ein abstraktes Modell eines Linearführungsschlittens (siehe Abb. 3.1), dessen Fräs- und Bohrbearbeitung in abstrakter Form als Beispiele der maschinellen Zerspanung behandelt wurden. Das Ziel des Projektes war es, mit der Entwicklung eines präzisen Hybridmodells und den Methoden der nichtlinearen Optimierung eine Kompensation der erhöhten thermischen Belastungen beim Trockenfräsen und Trockenbohren zu erreichen.

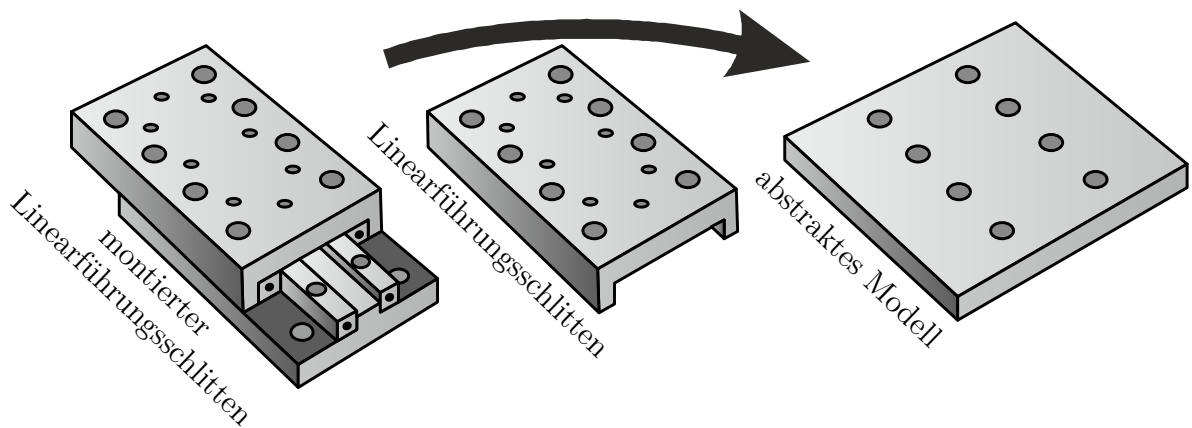


Abbildung 3.1: Demonstratorbauteil Linearführungsschlitten

In diesem Kapitel wird die Modellierung spanender Bearbeitungsprozesse vorgestellt und das Konzept für die Minimierung von Formabweichungen beschrieben. Zunächst wird auf die Modellierung von Fräsprozessen in zwei Sub-Modellen eingegangen. Diese beschreiben zum einen die Formabweichungen aufgrund einer Änderung des Eigenspannungszustands und zum anderen einen ungleichmäßigen Materialabtrag infolge thermischer Dehnungen. Es folgt die Vorstellung eines zweistufigen Optimierungskonzepts für die Kompensation von Fertigungsfehlern beim Fräsen. Abschließend wird das Vorgehen für das Bohren beschrieben und es werden die Versuchsreihen, die für die Modellierung benötigt werden, vorgestellt.

### 3.1 Modellierung der Prozesseinflüsse beim Trockenfräsen

Die Modellierung des Fräsens basiert auf der im Projekt entstandenen Arbeit [51].

Die Modellierung von thermischen Effekten in der maschinellen Fertigung kann in verschiedenen Größenordnungen erfolgen. Während sich die mikroskopische Modellierung auf die Spanbildung konzentriert und dadurch in der Regel nur kurze und kleine Ausschnitte des Gesamtprozesses abbilden kann, ist eine makroskopische Betrachtung besser geeignet, um die Fertigung eines gesamten Bauteils zu modellieren. Es kann bei bestimmten Werkstückgeometrien und Fertigungsaufgaben angenommen werden, dass vor allem zwei Effekte zu einer Beeinträchtigung der Fertigungsgenauigkeit bei der maschinellen Bearbeitung hochpräziser Bauteile führen. Zum einen kommt es durch die induzierte Wärme zu thermoelastischen Dehnungen. Diese können zu einem ungleichmäßigen Materialabtrag führen. Zum anderen treten bei der Zerspanung auch Deformationen auf, die durch eine Veränderung des Eigenspannungszustands hervorgerufen werden. Infolge der maschinellen Bearbeitung kann der Eigenspannungszustand durch plastische Deformationen in den Randschichten und durch das Abtragen von eigenspannungsbehaftete-

tem Material verändert werden [19]. In Abhängigkeit von der Steifigkeit des Werkstücks können die Einflüsse durch Eigenspannungen dominieren. Berücksichtigt werden beide Effekte in einer Art Hybridmodell des Trockenfräsen. Das Hybridmodell setzt sich aus zwei Sub-Modellen zusammen. Sub-Modell 1 beschreibt die eigenspannungsbedingten Effekte, Sub-Modell 2 die thermischen Effekte. Betrachtet wird das Gleichlaufräsen.

#### 3.1.1 Eigenspannungsbedingte Effekte - Sub-Modell 1

Die eigenspannungsbedingten globalen Deformationen sind Biegungen und Torsion des Bauteils aufgrund einer Veränderung des Eigenspannungszustands durch den Fräsprozess. Eine Besonderheit dieser Deformationen ist, dass sie, sofern das Bauteil magnetisch auf einer Spannplatte eingespannt ist, teilweise erst nach dem Ausspannen auftreten. Zudem wirken Biegungen und Torsion global am ganzen Bauteil. Eine Methodik, um die Verformungen zu beschreiben, ist die Verwendung von sogenannten effektiven Quellspannungen [92]. Die Annahme ist dabei, dass die durch die Bearbeitung eingebrachte Veränderung des Eigenspannungszustands äquivalent ist zu einer gleichmäßig verteilten Quellspannung, die in einer oberen Schicht des Bauteils wirkt. Der Rest des Materials wird als spannungsfrei betrachtet. Gulpak et al. [51] und Heinzl et al. [54] haben auch ein mehrschichtiges Quellspannungsmodell entwickelt, das die Berücksichtigung von im Vorhinein mit Eigenspannungen belasteten Bauteilen ermöglicht. In vielen Fällen sind die Eigenspannungen, die vor dem Fräsen im Bauteil vorhanden sind, dominierend und das Kompensationspotential der resultierenden Deformationen kann daher gering sein. Deshalb werden in dieser Arbeit die Bauteile vor der maschinellen Bearbeitung als eigenspannungsfrei angenommen und es wird ein einschichtiges Quellspannungsmodell verwendet. Dies kann in der Praxis zumindest teilweise durch Spannungsarmglühen erreicht werden [8].

Um mit diesem Modellierungsansatz die Minimierung von eigenspannungsbedingten Deformationen zu ermöglichen, ist ein mathematischer Zusammenhang zwischen Maschinenparametern und Quellspannungen erforderlich. Dies kann über ein empirisches, multilineares Regressionsmodell erfolgen.

#### Regressionsmodell

Dieses Vorgehen orientiert sich an der Veröffentlichung [50].

Die Grundidee der Modellierung mit Quellspannungen basiert auf der Plattentheorie von Kirchhoff [100]. Sei dafür ein Gebiet  $\Omega = [0; l_x] \times [0; l_y] \times [0; l_z]$  gegeben, welches das betrachtete Bauteil umfasst, und  $\Gamma = \partial\Omega$  der Rand des Bauteils (vgl. Abb. 3.2). Die Verformung des Bauteils in  $z$ -Richtung sei beispielsweise durch eine Koordinatenmessung der Bauteilunterseite<sup>1</sup>  $\Gamma_U := \Gamma|_{z=0}$  bekannt. Liegen hinreichend viele Messpunkte vor,

<sup>1</sup>Die Unterseite bildet globale Verformungen besser ab, da lokale Verformungen auf der bearbeiteten Seite wirken.

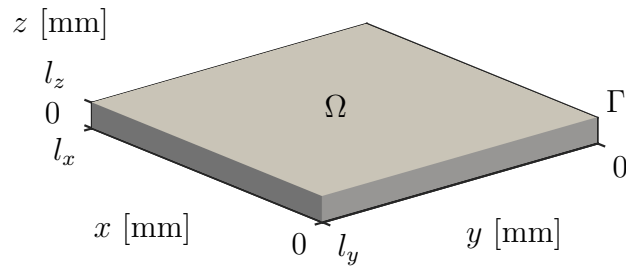


Abbildung 3.2: Bauteil beschrieben durch Gebiet  $\Omega$

lässt sich die tatsächliche Verschiebung durch eine quadratische Funktion der Form

$$u_{\phi,z} : \Gamma_U \rightarrow \mathbb{R}, \quad u_{\phi,z}(x,y) = \alpha_1 + \alpha_2 x + \alpha_3 y + \frac{1}{2} \kappa_x x^2 + \frac{1}{2} \kappa_y y^2 + \frac{1}{2} \kappa_{xy} x y$$

annähern. Dabei sind

$$\kappa_x = \frac{\partial^2 u_{\phi,z}}{\partial x^2}, \quad \kappa_y = \frac{\partial^2 u_{\phi,z}}{\partial y^2} \quad \text{und} \quad \kappa_{xy} = \frac{\partial^2 u_{\phi,z}}{\partial x \partial y}$$

die Krümmungsterme der Verschiebung. Diese können durch die Plattentheorie mit den Biegemomenten  $M_x$ ,  $M_y$  und  $M_{xy}$  einer Platte verknüpft werden:

$$\begin{aligned} M_x &= K (\kappa_x + \nu \kappa_y) \\ M_y &= K (\kappa_y + \nu \kappa_x) \\ M_{xy} &= K(1 - \nu) \kappa_{xy} \end{aligned} \tag{3.1}$$

$$K = \frac{E l_z^3}{12(1 - \nu^2)}$$

Dabei bezeichnet  $K$  die Plattensteifigkeit,  $\nu$  die Poissonzahl und  $E$  den Elastizitätsmodul. Die entsprechenden Werte für den in dieser Arbeit betrachteten Stahl 42CrMo4 sind in Tab. 3.1 zu finden. Es werden die Werte für 20 °C gewählt, da dies in etwa der Durchschnittstemperatur der Bauteile während der Koordinatenmessungen vor und nach der Bearbeitung entspricht. Die Biegemomente  $M_x$  und  $M_y$  wirken entlang der entsprechenden Achsen und das Torsionsmoment  $M_{xy}$  wirkt in der  $x$ - $y$ -Ebene. Die Platte wird als dünn angenommen, was bedeutet, dass  $l_z \ll l_x, l_y$  gilt.

### 3.1 Modellierung der Prozesseinflüsse beim Trockenfräsen

Parameter	Symbol	Wert	Einheit
Elastizitätsmodul	$E$	$212 \cdot 10^3$	N/mm <sup>2</sup>
Poissonzahl	$\nu$	0,3	–

Tabelle 3.1: Materialparameter von 42CrMo4 für die Berechnung von Quellspannungen bei einer Temperatur von 20 °C (Quelle: [103])

Wirkt ein Tensorfeld von Quellspannungen

$$\boldsymbol{\sigma}_\phi : \Omega \rightarrow \mathbb{R}^{3 \times 3}, \quad \boldsymbol{\sigma}_\phi(x, y, z) = \begin{cases} \begin{pmatrix} \sigma_x & \tau_{xy} & \tau_{xz} \\ \tau_{yx} & \sigma_y & \tau_{yz} \\ \tau_{zx} & \tau_{zy} & \sigma_z \end{pmatrix} & \text{für } z \geq l_z - z_0 \\ 0 & \text{sonst} \end{cases}$$

konstant im Bauteil  $\Omega$  von der gefrästen Oberfläche bis zu einer Tiefe  $z_0$ , kann das resultierende Biegemoment durch

$$\begin{aligned} M_{ij} &= \int_{l_z - z_0}^{l_z} \sigma_{\phi, ij} \left( z - \frac{l_z}{2} \right) dz \\ \Leftrightarrow M_{ij} &= \sigma_{\phi, ij} z_0 \left( \frac{l_z - z_0}{2} \right) \end{aligned} \quad (3.2)$$

berechnet werden, wobei  $M_{ij}$  das korrespondierende Biegemoment zum Eintrag  $\sigma_{\phi, ij}$  von  $\boldsymbol{\sigma}_\phi$  ist. Nimmt man nun (3.1) und (3.2) zusammen, lässt sich daraus der gewünschte Zusammenhang herstellen:

$$\begin{aligned} \sigma_x &= \frac{2K}{z_0(z_0 - l_z)} (\kappa_x + \nu \kappa_y) \\ \sigma_y &= \frac{2K}{z_0(z_0 - l_z)} (\kappa_y + \nu \kappa_x) \\ \tau_{xy} &= \frac{2K}{z_0(z_0 - l_z)} \kappa_{xy} (1 - \nu) \end{aligned} \quad (3.3)$$

Hierbei sind  $\sigma_x$  und  $\sigma_y$  die konstanten Quellspannungen, die entlang der entsprechenden Koordinatenachse bis zu einer Tiefe von  $z_0$  wirken, und  $\tau_{xy}$  die Torsionsspannung bis  $z_0$  in der  $x$ - $y$ -Ebene. Es wird angenommen, dass keine Quellspannungen entlang der  $z$ -Richtung auftreten ( $\tau_{xz} = \tau_{yz} = \tau_{zx} = \tau_{zy} = \sigma_z = 0$ ) und dass das Material isotrop ist ( $\tau_{xy} = \tau_{yx}$ ).

Mit (3.3) können aus der Krümmung eines gefertigten Bauteils äquivalente Quellspannungen berechnet werden, die im Bauteil herrschen müssten, um die gleiche Krümmung hervorzurufen. Zusammen mit Experimenten einer Versuchsreihe, bei der die Maschinenstellgrößen variiert werden, kann ein empirisches, multilineares Regressionsmodell aufgestellt werden, das einen Zusammenhang zwischen den Maschinenparametern Vorschubgeschwindigkeit  $v_f$ , Schnittgeschwindigkeit  $v_c$  und Eingriffsbreite  $a_e$  sowie den Quellspannungen schafft:

$$\begin{aligned}\sigma_x &= a_1 + b_1 v_c + c_1 v_f + d_1 a_e + e_1 v_c v_f + f_1 v_c a_e + g_1 v_f a_e + h_1 v_c v_f a_e \\ \sigma_y &= a_2 + b_2 v_c + c_2 v_f + d_2 a_e + e_2 v_c v_f + f_2 v_c a_e + g_2 v_f a_e + h_2 v_c v_f a_e \\ \tau_{xy} &= a_3 + b_3 v_c + c_3 v_f + d_3 a_e + e_3 v_c v_f + f_3 v_c a_e + g_3 v_f a_e + h_3 v_c v_f a_e\end{aligned}\quad (3.4)$$

Zur Bestimmung der freien Koeffizienten  $a_i, b_i, c_i, d_i, e_i, f_i, g_i$  und  $h_i, i = 1, 2, 3$  wurde die Versuchsreihe aus Kapitel 3.4.1 benutzt. Zu jedem Versuch wurden mit (3.3) die drei Quellspannungen berechnet. Dafür wurden mit Koordinatenmessungen die benötigten Krümmungen der gefertigten Bauteile bestimmt (siehe [50]). Mit diesem Datensatz ließen sich dann mittels linearer Ausgleichsrechnung die Koeffizienten des Regressionsmodells ermitteln. Die Abhängigkeit der Quellspannungen von der Schnitttiefe  $a_p$  wird bei diesem Ansatz vernachlässigt, da die Variation der Schnitttiefe in der Optimierung in Kapitel 6.1 nur wenige Mikrometer betragen wird. In diesem Bereich sind die eigenspannungsbedingten Deformationen größtenteils invariant [101]. Die resultierenden Modellkoeffizienten wurden bereits in [51] veröffentlicht und sind in Tab. 3.2 zusammengefasst. Die Eingangsgrößen  $v_c, v_f$  und  $a_e$  wurden auf die gleiche Größenordnung skaliert und haben die Einheiten  $[v_c] = 100\text{m/min}$ ,  $[v_f] = 1\text{m/min}$  und  $[a_e] = 1\text{mm}$ .

Modellkoeffizienten								
$i$	$a_i$	$b_i$	$c_i$	$d_i$	$e_i$	$f_i$	$g_i$	$h_i$
1	-14,71	2,57	5,55	1,88	-1,17	-0,41	-0,34	0,09
2	-27,04	4,06	10,86	3,41	-2,47	-0,66	-1,05	0,27
3	2,86	-1,36	-2,52	-0,76	0,60	0,35	0,39	-0,12

Tabelle 3.2: Koeffizienten der Regressionsmodelle (3.4) des Sub-Modell 1 (Quelle: [51])

Neben einer Bearbeitung der gesamten Werkstückoberfläche mit konstanten Maschinenparametern sollen auch komplexere Bearbeitungsstrategien ermöglicht werden, die aus verschiedenen Fräsoperationen bestehen. Dafür wird die zu bearbeitende Oberfläche  $\Gamma_O := \Gamma|_{z=l_z}$  in  $n_b$  Fräsbahnen eingeteilt. Es ergibt sich eine disjunkte Zerlegung:

$$\Gamma_O = \bigcup_{i=1}^{n_b} \Gamma_{O,i}, \quad \Gamma_{O,i} \cap \Gamma_{O,j} = \emptyset, \quad i, j = 1, \dots, n_b, \quad i \neq j$$



In den  $n_b$  verschiedenen Fräsoperationen wirken jeweils die Maschinenparameter  $a_{e,i}$ ,  $v_{c,i}$  und  $v_{f,i}$ ,  $i = 1, \dots, n_b$ . Diese sind bezüglich einer Fräsbahn, also auf den jeweiligen Teilstücken  $\Gamma_{O,i}$ , konstant. Dadurch ergeben sich mit dem Regressionsmodell (3.4) die Spannungen  $\sigma_x(a_{e,i}, v_{c,i}, v_{f,i})$ ,  $\sigma_y(a_{e,i}, v_{c,i}, v_{f,i})$  und  $\tau_{xy}(a_{e,i}, v_{c,i}, v_{f,i})$  auf dem  $i$ -ten Teilstück. Ist die Fräsbahn um einen Winkel  $-\alpha$  im Vergleich zur Fräsrichtung der Regressionsmodelle gedreht, wird noch eine Achsentransformation notwendig [63]. Im Bereich der  $i$ -ten Bahn herrschen dann gegeben durch Regressionsmodell und Transformation die Spannungen:

$$\begin{aligned}\hat{\sigma}_{x,i} &:= \frac{1}{2} (\sigma_x(a_{e,i}, v_{c,i}, v_{f,i}) + \sigma_y(a_{e,i}, v_{c,i}, v_{f,i})) \\ &\quad + \frac{1}{2} \cos(2\alpha) (\sigma_x(a_{e,i}, v_{c,i}, v_{f,i}) - \sigma_y(a_{e,i}, v_{c,i}, v_{f,i})) + \sin(2\alpha) \tau_{xy}(a_{e,i}, v_{c,i}, v_{f,i}) \\ \hat{\sigma}_{y,i} &:= \frac{1}{2} (\sigma_x(a_{e,i}, v_{c,i}, v_{f,i}) + \sigma_y(a_{e,i}, v_{c,i}, v_{f,i})) \\ &\quad - \frac{1}{2} \cos(2\alpha) (\sigma_x(a_{e,i}, v_{c,i}, v_{f,i}) - \sigma_y(a_{e,i}, v_{c,i}, v_{f,i})) - \sin(2\alpha) \tau_{xy}(a_{e,i}, v_{c,i}, v_{f,i}) \\ \hat{\tau}_{xy,i} &:= -\frac{1}{2} \sin(2\alpha) (\sigma_x(a_{e,i}, v_{c,i}, v_{f,i}) - \sigma_y(a_{e,i}, v_{c,i}, v_{f,i})) + \cos(2\alpha) \tau_{xy}(a_{e,i}, v_{c,i}, v_{f,i})\end{aligned}\tag{3.5}$$

Das Quellspannungstensorfeld  $\sigma_\phi$  wird aus diesen Werten zusammengesetzt:

$$\begin{aligned}\sigma_\phi(x, y, z) &= \begin{pmatrix} \hat{\sigma}_{x,i} & \hat{\tau}_{xy,i} & 0 \\ \hat{\tau}_{xy,i} & \hat{\sigma}_{y,i} & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{für } (x, y) \in \Gamma_{O,i}, z \geq l_z - z_0 \\ \sigma_\phi(x, y, z) &= 0 \quad \text{sonst}\end{aligned}\tag{3.6}$$

Damit das Tensorfeld (3.6) stetig und hinreichend glatt ist, wird zwischen den konstanten Bereichen ein glatter Übergang durch Interpolation<sup>2</sup> geschaffen. Dies erleichtert die angestrebte numerische Optimierung.

### Berechnung der quellspannungsbedingten Formabweichungen

Sind die äquivalenten Quellspannungen des Prozesses bekannt, können in einer FEM-Rechnung die resultierenden quellspannungsbedingten Formabweichungen  $u_{\text{quell}}$  bestimmt werden. Den nötigen mathematischen Zusammenhang schaffen die Lamé-Navier-Gleichungen. Hierbei werden simulativ effektive Quellspannungen innerhalb einer dünnen Schicht im Randbereich der bearbeiteten Oberfläche eingebracht. Eine detaillierte Beschreibung des Vorgehens erfolgt in Kapitel 6.1.1. Ein Vorteil bei diesem Ansatz ist, dass die Simulation nur örtlich und nicht in der Zeit erfolgt. Dies reduziert die Anzahl an diskreten Variablen und damit die Rechenzeit. In Abb. 3.3 ist die Funktionsweise des Sub-Modell 1 zusammengefasst.

<sup>2</sup>Für die Interpolation werden kubische Splines der Software *einspline* [37] verwendet.

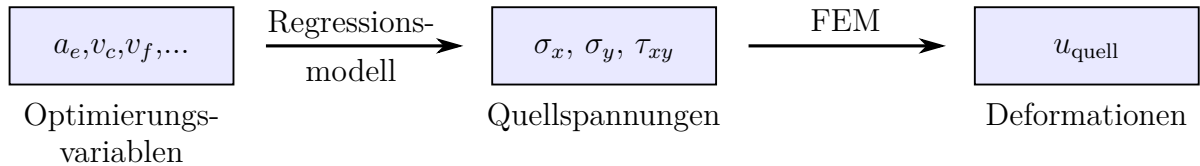


Abbildung 3.3: Funktionsweise des Sub-Modell 1

### 3.1.2 Thermische Effekte - Sub-Modell 2

Den zweiten Teil des Hybridmodells bildet das Sub-Modell 2. Es berücksichtigt die thermoelastischen Effekte, also die thermischen Ausdehnungen des Werkstücks aufgrund der Erwärmung während der Bearbeitung, und den daraus resultierenden ungleichmäßigen Materialabtrag. Die thermoelastischen Effekte beim Fräsen wirken größtenteils lokal. Ein großes Kompensationspotential bietet daher neben einer Reduzierung der eingebrachten Wärmeenergie vor allem eine Anpassung der Schnitttiefe  $a_p$ , um den lokalen Verschiebungen entgegenzuwirken.

Entscheidend für die thermisch bedingten Verschiebungen ist der Wärmeeintrag, der durch die maschinelle Bearbeitung in das Bauteil eingebracht wird. Modelliert wird der Wärmeeintrag durch eine bewegte Wärmequelle in Form des Eingriffsbogens der Schneiden an der Bauteiloberfläche. Physikalisch ist die Wärmequelle eine Wärmestromdichte. Diese ist abhängig von der Wahl der Maschinenparameter. Einen quantitativen Zusammenhang zwischen Wärmestromdichte und Maschinenparameter herzustellen ist maßgeblich für die Erstellung eines exakten Modells. Der Zusammenhang kann über ein empirisches, multilineares Regressionsmodell erfolgen.

#### Regressionsmodell

Dieses Vorgehen orientiert sich an der Veröffentlichung [94].

Die Modellierung der Wärmequelle erfolgt durch eine Einteilung des Eingriffsbogens in Bereiche mit konstanter Wärmestromdichte. Das hat den Vorteil, dass ein (wenn auch grober) Verlauf der Wärmestromdichte entlang des Eingriffsbogens modelliert wird. Bei der größten Eingriffsbreite  $a_e = r$  ( $r$  ist der Werkzeugradius) ergeben sich vier Bereiche (siehe Abb. 3.4). Jeder dieser Bereiche deckt einen Eingriffswinkel von  $22,5^\circ$  ab. Für die vier konstanten Wärmestromdichten wird jeweils ein eigenes Regressionsmodell aufgestellt:

$$\dot{q}_{\text{reg},i} = a_i + b_i v_c + c_i v_f + d_i v_c v_f, \quad i = 1, \dots, 4 \quad (3.7)$$

Die auf einen dieser Bereiche des Werkstücks wirkende Wärmestromdichte  $\dot{q}_{\text{reg},i}$  ist beim gewählten Ansatz nur von der Schnittgeschwindigkeit  $v_c$  und der Vorschubgeschwindigkeit  $v_f$  abhängig. Die Abhängigkeit des Wärmeeintrags von der Schnitttiefe  $a_p$  wird hier vernachlässigt. Sölter et al. [94] haben in Versuchen gezeigt, dass die Schnitttiefe einen nicht unerheblichen Einfluss auf die Wärmestromdichte haben kann, wenn diese

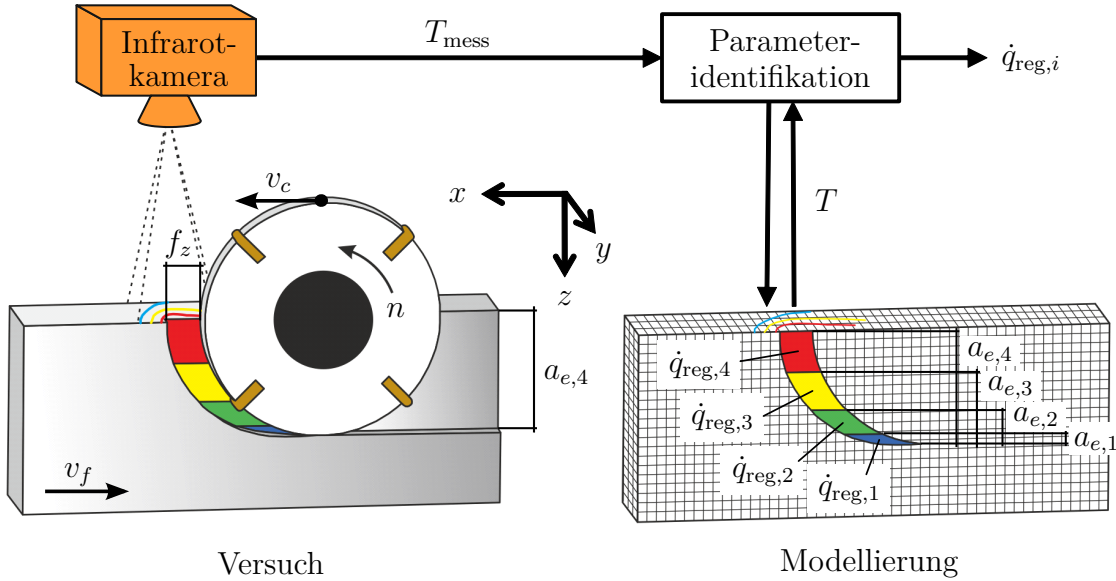


Abbildung 3.4: Bestimmung des Wärmeeintrags

um einen Millimeter variiert wird. Jedoch wird die Variation der Schnitttiefe in der Optimierung in Kapitel 6.2 nur wenige Mikrometer betragen. Eine Interpolation legt dann die Vermutung nahe, dass die Änderung von  $a_p$  in diesem Bereich keinen großen Einfluss auf den Wärmeeintrag hat. Es wird außerdem angenommen, dass der Verlauf der Wärmestromdichte entlang des Eingriffsbogens nicht von der Eingriffsbreite  $a_e$  abhängt, sondern  $a_e$  nur die Breite des Bogens bestimmt.

Für die Berechnung der freien Koeffizienten  $a_i$ ,  $b_i$ ,  $c_i$  und  $d_i$  der Regressionsmodelle wurden die Wärmestromdichten einer Versuchsreihe (siehe Kapitel 3.4.2) ermittelt. Die Maschinenparameter  $v_c$  und  $v_f$  wurden in der Versuchsreihe auf zwei Stufen variiert, die Eingriffsbreite  $a_e$  auf vier Stufen analog zum Eingriffswinkel von  $22,5^\circ$ . Die Schnitttiefe beträgt konstant  $a_p = 1$  mm. Begonnen wurde mit den Versuchen der kleinsten Eingriffsbreite  $a_{e,1}$ . Im Modell gibt es für  $a_{e,1}$  nur einen Bereich mit konstanter Wärmestromdichte (siehe Abb. 3.4, blauer Bereich des Eingriffsbogens). Nachdem die Wärmestromdichten  $\dot{q}_{\text{reg},1}$  dieser Versuche ermittelt wurden, konnten die Koeffizienten des Regressionsmodells  $i = 1$  mittels linearer Ausgleichsrechnung bestimmt werden. Anschließend wurden zu den Versuchen mit der Eingriffsbreite  $a_{e,2}$  die Wärmestromdichten  $\dot{q}_{\text{reg},2}$  des grünen Bereichs (siehe Abb. 3.4) ermittelt. Dabei wurde für den blauen Bereich die Wärmestromdichte  $\dot{q}_{\text{reg},1}$  des ersten Regressionsmodells verwendet. Auf die Berechnung der Koeffizienten des zweiten Modells folgte die nächste Stufe  $a_{e,3}$ . Nach diesem Schema wurde bis  $a_{e,4}$  zur Bestimmung der Koeffizienten von  $\dot{q}_{\text{reg},4}$  fortgefahren.

Die Werte der Koeffizienten der Regressionsmodelle (3.7) wurden bereits in [51] veröffentlicht und sind in Tab. 3.3 zu sehen. Die Eingangsgrößen wurden skaliert und haben die Einheiten  $[v_c] = 100\text{m/min}$  und  $[v_f] = 1\text{m/min}$ .

Was bisher unerwähnt blieb, ist, wie die Wärmestromdichte eines einzelnen Versuchs be-

Modellkoeffizienten ( $\times 100$ )				
$i$	$a_i$	$b_i$	$c_i$	$d_i$
1	-13,45	6,32	5,15	-2,19
2	-9,33	5,22	3,63	-1,77
3	-3,53	2,32	1,34	-0,75
4	-3,55	2,10	1,35	-0,69

Tabelle 3.3: Koeffizienten der Regressionsmodelle (3.7) des Sub-Modell 2 (Quelle: [51])

stimmt werden kann. Sie kann nicht unmittelbar gemessen werden, weshalb das Lösen eines inversen Problems, bei dem Thermografieaufnahmen der Versuche verwendet werden, ein weit verbreiteter Ansatz ist. Eine Möglichkeit, ein inverses Problem zu lösen, ist die Parameteridentifikation, bei der die Differenz zwischen gemessenem und simuliertem Temperaturfeld ( $T_{\text{mess}}$  und  $T$ ) minimiert wird. Die Vorstellung der Parameteridentifikation und ihrer Ergebnisse ist zentraler Bestandteil dieser Arbeit (siehe Kapitel 4).

### Berechnung der thermischen Formabweichungen und der Gesamtformabweichungen

Das Modell der thermischen Formabweichungen wurde im Projekt von den ingenieurwissenschaftlichen Projektpartnern aufgestellt und implementiert. Es wird daher bei der Optimierung als eine Art Black-Box dienen. Gulpak et al. beschreiben es detailliert in [47] und [51]. Die Berechnung von Sub-Modell 2 gliedert sich in drei Schritte: Zunächst werden mit den Regressionsmodellen (3.7) die Wärmestromdichten bestimmt, welche beim Fräsen auf das Bauteil wirken. Sie dienen als Randbedingungen für eine FEM-Simulation einer bewegten Wärmequelle in *ABAQUS*<sup>3</sup>. Die Wärmequelle wird dabei wie in Abb. 3.4 modelliert. Das Ergebnis der FEM-Simulation ist die Temperaturverteilung des Bauteils während eines simulierten Fräsvorgangs. Der zweite Schritt umfasst eine weitere FEM-Berechnung in *ABAQUS*. Durch Simulation der Thermoelastizität lassen sich aus der Temperaturverteilung die thermisch bedingten Knotenverschiebungen des Werkstücks ermitteln. Im dritten Schritt kommt ein kinematisches Modell zum Einsatz. Es modelliert die Kinematik des Werkzeugs zur Bestimmung des Materialabtrags. Zusammen mit den thermisch bedingten Verschiebungen und den quellspannungsbedingten Verschiebungen aus Sub-Modell 1 lässt sich damit die resultierende Bauteilform bzw. die resultierende Geometrieabweichung  $u_{\text{ges}}$  berechnen. Das Sub-Modell 2 wird in Kapitel 6.2 benötigt, um die Gesamtformabweichungen zu minimieren. Abb. 3.5 fasst es zusammen.

<sup>3</sup>ABAQUS ist ein FEM-Programm. Für weitere Informationen wird auf die Homepage [1] verwiesen.

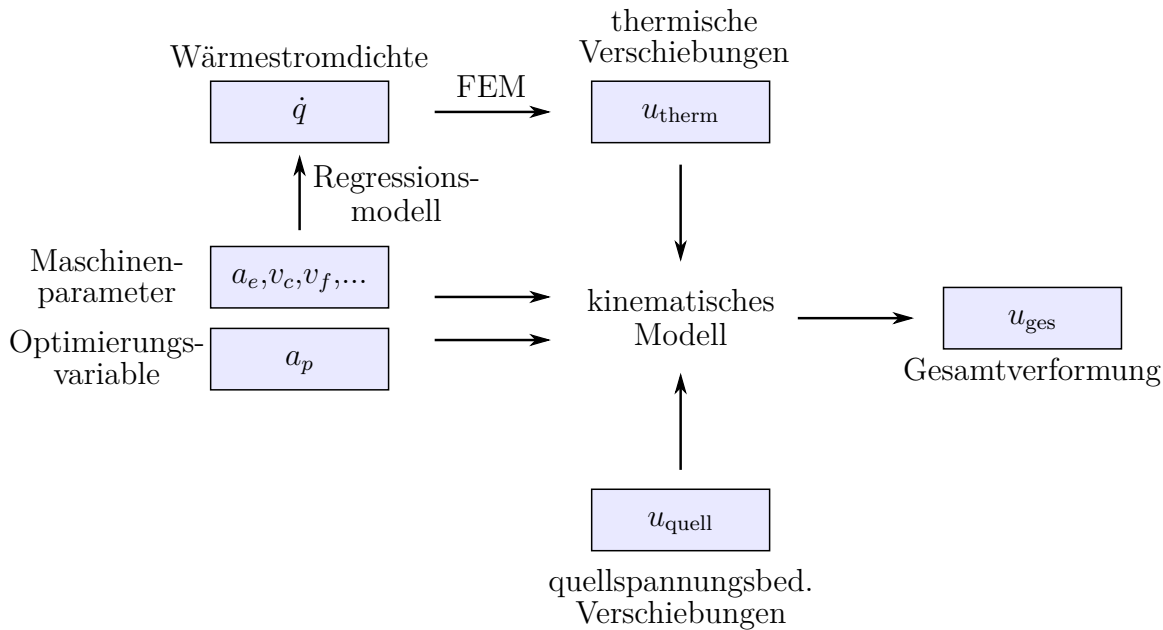


Abbildung 3.5: Funktionsweise des Sub-Modell 2

## 3.2 Vorgehensweise zur Minimierung von Formabweichungen beim Fräsen

Bei der angestrebten Minimierung von Formabweichungen sind die Optimierungsvariablen durch jene Variablen gegeben, die den Fräsprozess beschreiben. Das können neben den Maschinenparametern auch Variablen sein, die die Anzahl, Anordnung oder Reihenfolge der Fräsbahnen bestimmen. Während die Maschinenparameter beim Fräsen kontinuierliche Variablen darstellen, sind beispielsweise die Anzahl an Fräsbahnen oder die Reihenfolge diskrete Größen. Der Fokus dieser Arbeit liegt auf der kontinuierlichen Optimierung. Die diskreten Größen werden durch einen Vergleich verschiedener Frässtrategien untersucht. Diese unterscheiden sich in der Anzahl, Ausrichtung und Reihenfolge der Fräsbahnen. Die kontinuierlichen Variablen werden mit der vorgestellten Methodik der PDE-restringierten Optimierung optimiert.

Die Optimierung aller kontinuierlichen Variablen in einem Schritt und die Minimierung der superponierten Verformungen beider Sub-Modelle ist sehr rechenintensiv. Das liegt an der Komplexität des Sub-Modell 2, welches neben den drei Raumdimensionen auch die Zeitspanne eines vollständigen Fräsprozesses abbildet. Die einmalige Auswertung des Modells dauert bereits mehrere Tage. Diese Problematik wird deshalb durch eine zweistufige Optimierung gelöst, welche die  $a_p$ -Invarianz der Regressionsmodelle ausnutzt. Im ersten Schritt werden die Maschinenparameter  $a_e$ ,  $v_c$  und  $v_f$  und die verschiedenen Frässtrategien hinsichtlich der quellspannungsbedingten Deformationen optimiert. Ziel ist es, eine Bearbeitungsstrategie zu ermitteln, die Biegung und Torsion weitestgehend ver-

meidet. Die Nichtberücksichtigung von  $a_p$  ist unproblematisch, da die Quellsparnungen nahezu unabhängig von diesem Parameter sind und eine Kompensation mit der Schnitttiefe ohnehin nur auf der bearbeiteten Seite des Werkstücks erfolgen kann. Biegung und Torsion wirken sich jedoch auch auf die Rückseite aus. Der zweiten Optimierung geht dann zunächst eine Simulation voraus. Diese berechnet einmalig die thermisch bedingten Verschiebungen des Sub-Modell 2 mit den vorher optimierten Parametern  $a_e^*$ ,  $v_c^*$  und  $v_f^*$  und der optimierten Frässtrategie. Anschließend bilden die thermisch bedingten Verschiebungen und die im ersten Schritt übriggebliebenen quellsparnungsbedingten Deformationen die Grundlage für das kinematische Modell (KM), welches den Materialabtrag simuliert. Mit diesem Modell wird im zweiten Optimierungsschritt die Gesamtformabweichung minimiert, indem der verbleibende Maschinenparameter  $a_p$  optimiert wird. Durch diese zweistufige Optimierung geht zwar die Möglichkeit verloren, die thermisch bedingten Verschiebungen hinsichtlich der Maschinenparameter  $a_e$ ,  $v_c$  und  $v_f$  zu minimieren, da diese bereits im ersten Schritt festgelegt werden. Jedoch wird dadurch der Optimierungsprozess erst praktikabel. Zudem wird davon ausgegangen, dass die Schnitttiefe das größte Kompensationspotential für die thermischen Prozesseinflüsse bietet. Abbildung 3.6 fasst den zweistufigen Optimierungsprozess zusammen. Die Optimierung der quellsparnungsbedingten Deformationen wird in Kapitel 6.1 und die Optimierung der Gesamtverformung in Kapitel 6.2 vorgestellt.

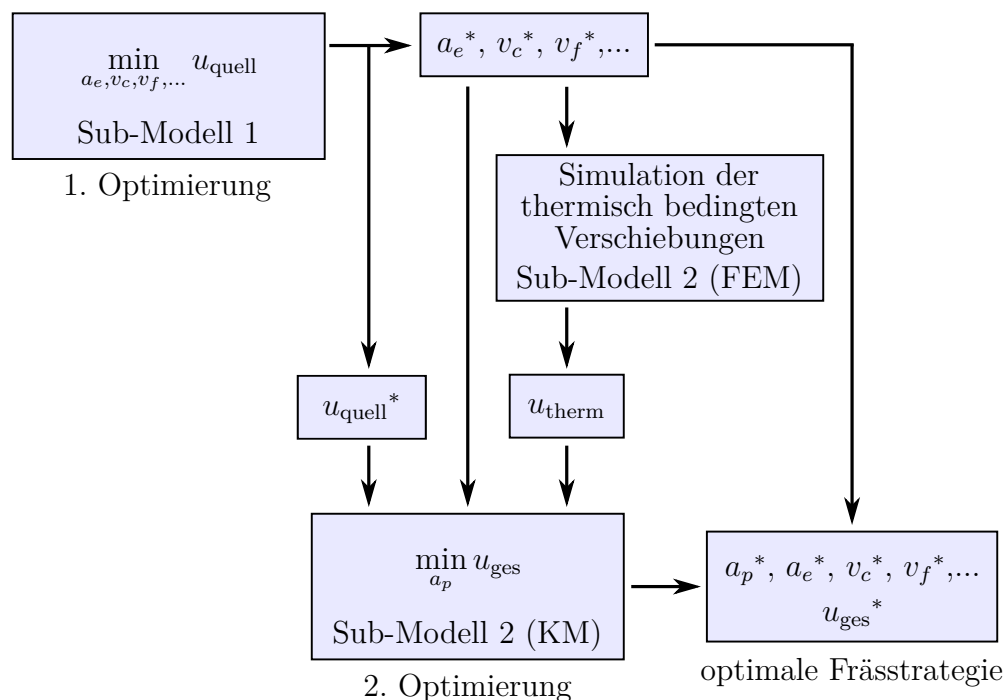


Abbildung 3.6: Vorgehensweise zur Minimierung der Gesamtformabweichung

### 3.3 Modellierung der Prozesseinflüsse beim Bohren

Im Gegensatz zum Trockenfräsen wird das Bohren durch Minimalmengenschmierung unterstützt. Eine Kompensation von thermischen Effekten beim Bohren kann über die Optimierung der Maschinenparameter erfolgen. Diese sind bei den betrachteten Versuchen nur durch Vorschubgeschwindigkeit und Schnittgeschwindigkeit gegeben, da eine Variation des Bohrerradius in der Regel nur in vergleichsweise großen Stufen möglich ist. Eine weitere vielversprechende Kompensationsmöglichkeit bietet beim Bohren mehrerer Löcher im selben Bauteil die Optimierung der Reihenfolge dieser Bohrungen. Letztere Kompensation hat im Rahmen dieser Arbeit allerdings zwei Nachteile. Zum einen zählt die Optimierung einer Reihenfolge zur diskreten Optimierung. Die in dieser Arbeit entwickelten Methodiken der kontinuierlichen Optimierung können hierbei nicht ohne Weiteres zur Anwendung kommen. Zum anderen muss bei der Simulation mehrerer Bohroperationen eine lange Zeitspanne abgebildet werden. Dies erhöht die nötigen Modellberechnungen der Optimierung signifikant, sodass selbst die Rechenkapazität moderner Rechensysteme schnell ausgeschöpft ist. Die Optimierung von Bohrstrategien zur Kompensation von Deformationen befindet sich noch im Entwicklungsstadium.

Innerhalb dieser Arbeit fokussieren sich die Ausführungen zum Bohren auf die Modellierung. Diese ist innerhalb dieser Arbeit der des Fräsens sehr ähnlich. Ein Unterschied besteht darin, dass beim Bohren nur die thermischen Effekte berücksichtigt werden. Aufgrund der Rotationssymmetrie der bearbeiteten Fläche beim Bohren wird davon ausgegangen, dass der Einfluss von eigenspannungsbedingten Deformationen vernachlässigt werden kann. Das Modell orientiert sich deshalb am thermischen Sub-Modell 2 des Fräsens. Auch hier ist ein entscheidender Faktor der Wärmeeintrag.

Die Modellierung des Wärmeeintrags erfolgt durch eine bewegte Wärmequelle innerhalb des Bohrlochs. Die Wärmequelle ist zum einen in eine primäre Wärmequelle, die den Wärmeeintrag durch den Eingriff der Schneide abbildet, eingeteilt und zum anderen in mehrere Bereiche oberhalb der Schneide, die sowohl den Wärmeeintrag der Reibung des Bohrers an der Bohrlochwand simulieren als auch den Wärmeeintrag, der durch den Abtransport der Späne entsteht. Eine ähnliche Methodik wurde von Gulpak und Sölter [49] verwendet. Alle Wärmequellen wirken rotationssymmetrisch und bewegen sich mit der Zeit gemäß der Bohrerbewegung im Loch abwärts. Die Bestimmung der Wärmestromdichten von Experimenten kann wie beim Fräsen durch eine Parameteridentifikation erfolgen. Innerhalb dieser Arbeit liegt der Schwerpunkt des Bohrens auf diesen Berechnungen. Dabei wird ein nichtlineares Wärmeleitungsmodell verwendet und neben der Identifikation von Thermografiebildern ein Vergleich zwischen den Optimierungsansätzen SAND und NAND durchgeführt. Der Abgleich zwischen Experiment und Simulation erfolgt wie beim Fräsen über Temperaturbilder, die von Versuchen aufgezeichnet wurden. Der dafür nötige Versuchsplan ist in Kapitel 3.4.3 beschrieben. Die Details zur Parameteridentifikation werden in Kapitel 5 vorgestellt.

## 3.4 Versuchsreihen für die empirischen Modelle

Für die benötigten Experimentaldaten zur Bestimmung der Quellspannungen von Sub-Modell 1 und zur Bestimmung der Wärmestromdichten der thermischen Modelle des Fräsens und Bohrens wurde auf Versuchsreihen, die im Zuge des SPP-Projekts von den ingenieurwissenschaftlichen Kollegen durchgeführt wurden, zurückgegriffen. Nachfolgend werden die Experimente von Trockenfräsprozessen für die beiden Sub-Modelle und von Bohrprozessen mit Minimalmengenschmierung beschrieben.

### 3.4.1 Experimente für Sub-Modell 1

Die verwendeten Daten basieren auf zwei Versuchsreihen. Versuchsaufbau, Vorgehensweise und Ergebnisse der Versuchsreihen wurden zum einen in [48] und zum anderen in [51, 54] veröffentlicht.

Die Werkstücke der ersten Versuchsreihe [51, 54] aus dem Stahl 42CrMo4 wurden durch Normalglühen vorbehandelt und auf eine Größe von  $l_x = 145,2$  mm,  $l_y = 145,2$  mm und  $l_z = 12$  mm gebracht. Die anschließende Bearbeitung erfolgte durch trockenes Gleichlaufräsen mit dem Vorschub in  $y$ -Richtung. Nach jedem Schnitt wurde das Werkzeug in der Vorschubrichtung zurückgesetzt und um  $a_e$  in  $x$ -Richtung verschoben. Diese Vorgehensweise wurde wiederholt, bis die gesamte Werkstückoberfläche gefräst war. Der Fräser hatte dabei einen Durchmesser von 25 mm und vier mit Titanaluminiumnitrid beschichtete Schneiden aus Wolframcarbid.

Es wurden vier Parametersets gefahren. Dabei wurde die Schnittgeschwindigkeit auf  $v_c = 373$  m/min, die Vorschubgeschwindigkeit auf  $v_f = 2800$  mm/min und die Schnitttiefe auf  $a_p = 1$  mm konstant gehalten. Die Eingriffsbreite wurde mit den Werten  $a_e \in \{0,95$  mm; 3,66 mm; 7,72 mm; 12,5 mm} variiert. Jede Parameterkonstellation wurde fünfmal wiederholt, sodass insgesamt 20 Experimente durchgeführt wurden. Nach der Bearbeitung von fünf Werkstücken wurden die Werkzeugschneiden ausgetauscht.

Um die Daten für die Bestimmung des Regressionsmodells (3.4) zu komplettieren, wurden die genannten Versuche mit einer anderen Versuchsreihe kombiniert, welche in [48] veröffentlicht wurde. In diesen Experimenten wurde neben derselben Variation der Eingriffsbreite  $a_e$  auch die Schnittgeschwindigkeit  $v_c \in \{255$  m/min; 373 m/min} und die Vorschubgeschwindigkeit  $v_f \in \{2000$  mm/min; 2800 mm/min} auf jeweils zwei Stufen variiert. Die Verknüpfung wird dadurch ermöglicht, dass beide Versuchsreihen bei  $v_c = 373$  m/min und  $v_f = 2800$  mm/min die gleiche Parameterkonstellation aufweisen. In dieser Versuchsreihe wurde auch die Schnitttiefe verändert. Innerhalb dieser Arbeit wurden aber nur die Experimente mit  $a_p = 1$  mm benutzt.

Insgesamt konnten durch Koordinatenmessungen aller gefrästen Bauteile die Quellspannungen, die durch den Fräsprozess in die oberste Schicht eingetragen wurden, bestimmt



werden. Die so gewonnenen Daten wurden dazu verwendet, die Koeffizienten des Regressionsmodells (3.4) durch lineare Ausgleichsrechnung zu bestimmen.

#### 3.4.2 Experimente für Sub-Modell 2

Die verwendete Versuchsreihe ist der Arbeit [93] von Sölter und Gulpak entnommen.

Die Werkstücke aus 42CrMo4-Stahl wurden zunächst normalgeglüht und auf eine Größe von  $l_x = 150$  mm,  $l_y = 19$  mm und  $l_z = 20$  mm vorverarbeitet. Anschließend wurde die Bauteilfront, von der die Thermografieaufnahmen entstanden sind, schwarz lackiert, um einen hohen, gleichmäßigen Emissionsgrad zu erreichen. Das Werkzeug mit einem Durchmesser von 25 mm hat vier mit Titanaluminiumnitrid beschichtete Schneiden aus Wolframcarbid. Die eigentliche Bearbeitung wurde dann durch trockenes Gleichlauffräsen realisiert. Dabei wurden die Maschinenparameter variiert - die Schnittgeschwindigkeit mit den Werten  $v_c \in \{255 \text{ m/min}; 373 \text{ m/min}\}$  und die Vorschubgeschwindigkeit mit  $v_f \in \{2000 \text{ mm/min}; 2800 \text{ mm/min}\}$ . Für die Eingriffsbreite wurden die Werte  $a_e \in \{0,95 \text{ mm}; 3,66 \text{ mm}; 7,72 \text{ mm}; 12,5 \text{ mm}\}$  verwendet. In der benannten Versuchsreihe wurde auch die Schnitttiefe verändert, wobei innerhalb dieser Arbeit nur die Experimente mit  $a_p = 1$  mm Verwendung finden. Jede Parameterkonstellation wurde durch fünf Wiederholungen realisiert. Insgesamt standen 80 Experimente zur Verfügung.

Während des Fräsvorgangs wurden thermografische Bilder mit 50 Hz von der gestrichenen Bauteilseite gefertigt. Die Infrarotkamera, die dabei zum Einsatz kam, hat einen Spektralbereich von 3,4–5,0  $\mu\text{m}$ . Die Auflösung beträgt 0,1 °C bei einer Temperatur von 30 °C und die Genauigkeit  $\pm 2$  °C in einem Bereich zwischen 0 °C und 100 °C.

Die so gemessenen Temperaturverteilungen wurden für die Parameteridentifikation in Kapitel 4 benutzt, um die Wärmestromdichten der Experimente zu bestimmen.

#### 3.4.3 Experimente für das thermische Modell des Bohrens

Die Experimente für die Parameteridentifikation von Bohrprozessen entstammen der Arbeit [49] von Gulpak und Sölter.

Nach dem Normalglühen der Werkstücke aus 42CrMo4-Stahl wurden diese auf eine Größe von  $l_x = 82$  mm,  $l_y = 78$  mm und  $l_z = 78$  mm vorverarbeitet. Um einen hohen Emissionsgrad und wenig Reflektion zu erzielen, wurden die Bauteile schwarz lackiert. Zur Bestimmung des Emissionsgrads bei verschiedenen Temperaturen wurden Messvergleiche mit einem Thermoelement durchgeführt.

Die verwendeten Bohrer vom Typ Sandvik CoroDrill 860 sind mit Titanaluminiumnitrid PVD-beschichtet und haben einen Radius von  $r = 4$  mm bzw.  $r = 6$  mm. Für die Anwendung von Minimalmengenschmierung enthalten die Bohrer Löcher zum Einbringen von Kühlflüssigkeit. Der Versuchsplan bestand aus 40 Experimenten, bei denen die

Maschinenparameter Schnittgeschwindigkeit  $v_c \in \{55 \text{ m/min}; 75 \text{ m/min}\}$  und Vorschubgeschwindigkeit  $v_f \in \{400 \text{ mm/min}; 600 \text{ mm/min}\}$  variiert wurden. Neben der Durchführung aller Parameterkombinationen mit den beiden verschiedenen Bohrerradien wurde jeder Versuch fünfmal wiederholt. Nach fünf Versuchen wurde der Bohrer ersetzt, um den Effekt des Verschleißes zu verringern.

Die verwendete Infrarotkamera hat einen Spektralbereich von  $7,5\text{--}14 \mu\text{m}$ . Die Auflösung beträgt  $0,03 \text{ K}$  bei einer Temperatur von  $30 \text{ }^\circ\text{C}$  und die Genauigkeit  $\pm 1 \text{ }^\circ\text{C}$ . Die numerische Auflösung der Daten beträgt jedoch nur  $0,1 \text{ }^\circ\text{C}$ . Die Aufnahmen der Thermografiekamera erfolgten durch ein Schutzrohr. Dies ermöglichte die weitestgehend ungestörte Aufnahme der Bauteilfront. Durch das Schutzrohr war das Bauteil jedoch nur in einem runden Ausschnitt zu sehen, was bei der Verwendung der Daten berücksichtigt werden musste.

Die Thermografiebilder wurden in Kapitel 5.3 verwendet, um den Wärmeeintrag beim Bohren zu bestimmen.

# Kapitel 4

## Parameteridentifikation von Fräsprozessen

### Inhalt

---

4.1	Parameteridentifikation zur Bestimmung der Wärmestromdichte . . .	52
4.2	Allgemeines Wärmeleitungsmodell . . . . .	53
4.3	Problemstellung beim Fräsen . . . . .	58
4.4	Wärmequelle, Diskretisierung und Temperaturdaten . . . . .	61
4.5	Numerische Ergebnisse . . . . .	65
4.5.1	Validierung . . . . .	65
4.5.2	Identifikation von Realdaten . . . . .	69
4.6	Fazit der Parameteridentifikation des Fräsens . . . . .	77

---

Wie in Kapitel 3.1.2 beschrieben, ist der Wärmeeintrag von Zerspanprozessen nicht direkt messbar. Er wird deshalb als Lösung eines Parameteridentifikationsproblems durch einen Abgleich von simulierten und gemessenen Temperaturen bestimmt. In diesem Kapitel wird zunächst die allgemeine Methodik zur Bestimmung des Wärmeeintrags in der spanenden Fertigung beschrieben und die Wärmeleitungsgleichung motiviert. Dieser allgemeine Teil dient auch als Grundlage für die Identifikation von Bohrprozessen, die in Kapitel 5 vorgestellt wird. Anschließend wird die Problemstellung für das Fräsen spezifiziert. Die numerischen Ergebnisse umfassen eine Validierung mit künstlichen Daten und die Bestimmung von Wärmestromdichten realer Experimente. Das Kapitel endet mit einem kurzen Fazit.

Dieses Kapitel orientiert sich an der bereits veröffentlichten Arbeit [107] des Autors.

## 4.1 Parameteridentifikation zur Bestimmung der Wärmestromdichte

Die in diesem Kapitel vorgestellte Methodik ist nicht nur in Anbetracht dieser Arbeit und dem zugehörigen Schwerpunktprogramm von Bedeutung, sondern beschreibt vielmehr eine allgemeine Möglichkeit zur Bestimmung von Wärmestromdichten in der spanenden Fertigung. Das Vorgehen stellt eine Erweiterung der Methodik, die für Sub-Modell 2 (siehe Kapitel 3.1.2) verwendet wurde, dar. Die Verwendung der PDE-restringierten Optimierung mit dem in Kapitel 2.1 vorgestellten SAND-Ansatz verspricht zudem ein effizientes Lösen.

### Mathematische Formulierung der Parameteridentifikation

Die Parameteridentifikation ist eine weit verbreitete Methode, um unbekannte Modellparameter zu bestimmen. Die Anwendungsgebiete sind dabei vielfältig. Echim [35] hat die Parameteridentifikation für die Modellierung einer Luftstrecke eines aufgeladenen Dieselmotors angewendet. Das an die Realität angepasste Modell diente anschließend als Grundlage für die Motorregelung zur Abgasreduzierung. In der Medizin kann die Parameteridentifikation bei der Verbesserung von Verfahren der Krebsbehandlung helfen. So hat Tiesler [98] eine Methodik zur Identifikation von Gewebeparametern aus Temperaturdaten entwickelt, um den Erfolg der Radiofrequenz Ablation zu verbessern. Die Anwendungen reichen bis zu einer Identifikation von rein datenbasierten Modellen, z. B. im Bereich der Energiemaximierung von Kraft-Wärme-Kopplungsanlagen [26].

Bei der Parameteridentifikation wird versucht, einen oder mehrere Modellzustände an gemessene Daten anzugleichen. Dies geschieht über die Minimierung einer Zielfunktion, die eine mathematische Differenz zwischen den gemessenen Werten und den entsprechenden Modellgrößen beschreibt. Es muss davon ausgegangen werden, dass die Messungen in der Praxis mit Fehlern behaftet sind. Wenn man systematische Fehler ausschließen kann, bleiben zufällige Fehler übrig, die durch verschiedene unabhängige Einflussfaktoren entstehen. In der Regel ist diese Art von Fehlern annähernd normalverteilt. Unter dieser Voraussetzung ist auf Grundlage der Statistik die Verwendung einer Least-Squares Zielfunktion, die die quadratischen Abweichungen zwischen den Datenpunkten und den Modellwerten aufsummiert, am sinnvollsten [73, 89]. Ausgehend von der diskretisierten Formulierung (2.4) wird die Zielfunktion der Parameteridentifikation wie folgt definiert:

$$F(S) = \sum_{i=1}^{n_s} \bar{w}_i (S_i - \vartheta_i(S_{\text{mess}}))^2 \quad (4.1)$$

Hierbei bezeichnet  $S_{\text{mess}} \in \mathbb{R}^{n_{\text{mess}}}$  den Vektor der Messgrößen des Zustands  $S$  und  $n_{\text{mess}}$  die Anzahl der Messpunkte.  $\vartheta_i : \mathbb{R}^{n_{\text{mess}}} \rightarrow \mathbb{R}$ ,  $i = 1, \dots, n_s$  sind Funktionen, die zum  $i$ -ten Eintrag des Zustandsvektors  $S$  den korrespondierenden Messpunkt aus  $S_{\text{mess}}$  durch Interpolation berechnen. Zusätzlich werden Gewichte  $w_i \in \{1; 0\}$ ,  $i = 1, \dots, n_s$  eingeführt,

die es ermöglichen gezielt Zustände in der Zielfunktion unberücksichtigt zu lassen. Dies ist beispielsweise zweckmäßig bei Zuständen, zu denen kein Messwert vorhanden ist oder zu denen eine fehlerhafte Messung vorliegt. Zur besseren Vergleichbarkeit werden in der Zielfunktion die normierten Gewichte

$$\bar{w}_i := \frac{w_i}{\sum_{j=1}^{n_s} w_j}, \quad i = 1, \dots, n_s \quad (4.2)$$

verwendet.

Der Zustand  $S$  ist allgemein durch ein PDE-Modell bestimmbar. Dieses hängt zudem vom Entscheidungsvektor  $Z$  ab, welcher bei der Parameteridentifikation aus Identifikationsparametern besteht. Durch Diskretisierung mit der FEM (siehe Kapitel 2.3) wird daraus das allgemeine Gleichungssystem

$$H(S, Z) = 0 \in \mathbb{R}^{n_s}. \quad (4.3)$$

Bei den in diesem Kapitel auftretenden Parameteridentifikationsproblemen wird es neben (4.3) und zusätzlichen Boxschränken keine weiteren Gleichungs- oder Ungleichungsnebenbedingungen geben. Die Boxschränken an den Zustand  $S \in \mathbb{R}^{n_s}$  haben die untere Schranke  $S_L \in \mathbb{R}^{n_s}$  und die obere Schranke  $S_U \in \mathbb{R}^{n_s}$ , die Beschränkungen an  $Z \in \mathbb{R}^{n_z}$  die Schranken  $Z_L \in \mathbb{R}^{n_z}$  und  $Z_U \in \mathbb{R}^{n_z}$ . Das Parameteridentifikationsproblem kann mit der Optimierungsvariablen  $X = (S^T, Z^T)^T \in \mathbb{R}^{n_x}$  in SAND-Form allgemein als folgendes NLP formuliert werden:

$$\begin{aligned} \min_{X \in \mathbb{R}^{n_x}} \quad & \sum_{i=1}^{n_s} \bar{w}_i (S_i - \vartheta_i(S_{\text{mess}}))^2 \\ \text{unter} \quad & H(S, Z) = 0 \\ & S_L \leq S \leq S_U \\ & Z_L \leq Z \leq Z_U \end{aligned} \quad (4.4)$$

In diesem Kapitel werden die Optimierungsprobleme mit dem SAND-Ansatz gelöst, da dieser im Vergleich zum NAND ein effizienteres Lösen verspricht (siehe Kapitel 2.1). Ein Vergleich beider Ansätze folgt für ein komplexeres Problem in Kapitel 5.3.2.

## 4.2 Allgemeines Wärmeleitungsmodell

Zur Herleitung des diskretisierten Modells, welches in der obigen allgemeinen Formulierung durch das Gleichungssystem  $H$  beschrieben wurde, wird zunächst eine kontinuierliche PDE betrachtet, die anschließend durch Diskretisierung in die Form (4.3) gebracht wird.

Zur Beschreibung des Wärmeflusses durch einen Körper wird typischerweise die Wärmeleitungsgleichung verwendet [71]. Die dreidimensionale, instationäre Variante dieser PDE beschreibt über einen Zeitraum  $[t_0; t_f] \subset \mathbb{R}$  die Temperaturverteilung  $T_\phi : \Omega \times [t_0; t_f] \rightarrow \mathbb{R}$  innerhalb eines Gebietes  $\Omega \subset \mathbb{R}^3$ .  $\Omega$  wird im Folgenden den relevanten Ausschnitt des Werkstücks darstellen. Ein äußerer Wärmeeintrag über den Rand  $\Gamma := \partial\Omega$  von  $\Omega$  kann durch eine Neumann-Randwertfunktion  $\dot{q}_\phi : \Gamma \times [t_0; t_f] \rightarrow \mathbb{R}$ , die physikalisch eine Funktion der Wärmestromdichte ist, beschrieben werden. Dabei bedeutet ein positives  $\dot{q}_\phi$  einen positiven Wärmeeintrag in das Bauteil. Zudem sei auch die Anfangstemperaturverteilung  $T_{\phi,0} : \Omega \rightarrow \mathbb{R}$  bekannt. Zusammenfassend lässt sich das wie folgt formulieren:

$$\begin{aligned} \rho c_p(T_\phi) \frac{\partial T_\phi}{\partial t} &= \operatorname{div}(k(T_\phi) \nabla T_\phi) && \text{in } \Omega \\ k(T_\phi) \nabla T_\phi \cdot \vec{n} &= \dot{q}_\phi && \text{auf } \Gamma \\ T_\phi &= T_{\phi,0} && \text{für } t = t_0 \end{aligned} \tag{4.5}$$

Dabei ist  $\vec{n}$  die äußere Normale von  $\Gamma$ . Die Dichte  $\rho$  ist in guter Näherung im relevanten Temperaturbereich konstant, während die spezifische Wärmekapazität  $c_p(T_\phi)$  sowie die Wärmeleitfähigkeit  $k(T_\phi)$  temperaturabhängig sind. Dadurch ist das Problem im Allgemeinen nichtlinear. Mit geeigneten Voraussetzungen kann die Existenz einer Lösung von (4.5) beispielsweise mit dem Galerkin-Verfahren [91] oder der Methode der Semi-Diskretisierung [110] nachgewiesen werden. Beim Nachweis der Eindeutigkeit der Lösung findet das Gronwallsche Lemma [91] Anwendung.

Da die Wärmeleitungsgleichung parabolisch ist, lässt sie sich nicht direkt mit der FEM diskretisieren [85]. Grundsätzlich bieten sich zwei Methoden an, um durch eine spezielle Behandlung der Zeitableitung in (4.5) den verbleibenden elliptischen Teil der PDE mit der FEM zu diskretisieren. Bei der vertikalen Linienmethode<sup>1</sup> wird die Zeitableitung zunächst bei der Diskretisierung nicht berücksichtigt und stattdessen werden nur die drei Raumdimensionen mit der FEM diskretisiert. Dieses Vorgehen führt auf ein großes System von gewöhnlichen Differentialgleichungen. Die Rothe-Methode<sup>2</sup> hingegen diskretisiert als erstes nur die Zeit mit einem geeigneten numerischen Verfahren für gewöhnliche Differentialgleichungen. Dabei entstehen für jeden diskreten Zeitpunkt elliptische PDEs, die anschließend diskretisiert werden können. Der Vorteil der vertikalen Linienmethode ist, dass zur Lösung der gewöhnlichen Differentialgleichungen externe Löser verwendet werden können, während bei der Rothe-Methode die Zeitdiskretisierung vielmehr von Hand erfolgt. Der Vorteil der Rothe-Methode hingegen ist, dass durch die getrennte Diskretisierung des Raums für jeden einzelnen Zeitpunkt das Gitter zeitvariant sein kann. Dies ist bei der vertikalen Linienmethode in der Regel nicht der Fall. Man erhält das gleiche Gitter für alle Zeitpunkte.

Da insbesondere beim Fräsen große Teile der Bauteiloberfläche bearbeitet werden und

<sup>1</sup>Die vertikale Linienmethode wird teilweise auch nur als Linienmethode bezeichnet.

<sup>2</sup>Die Rothe-Methode wird auch horizontale Linienmethode genannt.

daher zu verschiedenen Zeitpunkten die Bereiche von  $\Omega$  unterschiedlich stark beansprucht werden, kann es von Vorteil sein, in Abhängigkeit von der Position des Werkzeugs und des Zeitpunkts lokale Verfeinerungen des Gitters vorzunehmen. Dies ermöglicht an Stellen, die weiter vom Eingriffsbereich entfernt sind und damit kleinere Temperaturgradienten aufweisen, eine Vergrößerung des Gitters. So lassen sich in der Regel eine nicht unerhebliche Menge an Gitterpunkten einsparen. Im Gegenzug wird zwischen den Zeitschritten eine Transformation notwendig, welche die schwach besetzten Systemmatrizen dichter werden lassen kann. Letztlich muss der Anwender anhand von Performanz-Tests entscheiden, ob die Einsparung der Gitterpunkte die dichteren Systemmatrizen aufwiegt. Um für die Parameteridentifikation des Fräsens und Bohrens eine zeitliche Gittervariation zu ermöglichen, wird im Folgenden die Rothe-Methode verwendet. Zunächst wird dafür in (4.5) die Zeit mit einem Einschrittverfahren diskretisiert. Häufig wird das sogenannte Theta-Verfahren verwendet. Dieses stellt eine Verallgemeinerung der Eulerverfahren und des Trapezverfahrens dar und entspricht für  $\theta = 1$  dem expliziten Eulerverfahren, für  $\theta = \frac{1}{2}$  dem Trapezverfahren und für  $\theta = 0$  dem impliziten Eulerverfahren. Der Vorteil des Theta-Verfahrens ist, dass auch nach der Implementierung noch leicht zwischen den verschiedenen Verfahren gewechselt werden kann. Mit einer äquidistanten Diskretisierung der Zeitachse in  $n_t$  Abschnitte ergibt sich:

$$\begin{aligned} \rho c_p(\theta, T_\phi^{l-1}, T_\phi^l) \frac{T_\phi^l - T_\phi^{l-1}}{\Delta t} &= \operatorname{div}(\theta k(T_\phi^{l-1}) \nabla T_\phi^{l-1} + (1 - \theta) k(T_\phi^l) \nabla T_\phi^l) \quad \text{in } \Omega, \quad l = 1, \dots, n_t \\ k(T_\phi^l) \nabla T_\phi^l \cdot \vec{n} &= \dot{q}_\phi^l \quad \text{auf } \Gamma, \quad l = 0, \dots, n_t \\ T_\phi^0 &= T_{\phi,0} \end{aligned} \tag{4.6}$$

Hierbei wurde die Vereinfachung

$$c_p(\theta, T_\phi^{l-1}, T_\phi^l) := \theta c_p(T_\phi^{l-1}) + (1 - \theta) c_p(T_\phi^l)$$

verwendet.  $T_\phi^l : \Omega \rightarrow \mathbb{R}$  ist die Temperaturverteilung zum Zeitpunkt  $t_l := t_0 + l \Delta t$  mit  $\Delta t := \frac{t_f - t_0}{n_t}$  und  $\dot{q}_\phi^l : \Gamma \rightarrow \mathbb{R}$  ist die Funktion der Wärmestromdichte zum Zeitpunkt  $t_l$ .

Das PDE-System (4.6) beschreibt  $n_t$  elliptische PDEs. Zur Lösung des Systems mit der FEM (siehe Kapitel 2.3) wird dessen schwache Formulierung benötigt. Zunächst wird die erste Gleichung von (4.6) mit einer Testfunktion<sup>3</sup>  $v_\phi \in H^1(\Omega)$  multipliziert und über  $\Omega$  integriert:

$$\begin{aligned} &\int_{\Omega} \rho c_p(\theta, T_\phi^{l-1}, T_\phi^l) \frac{T_\phi^l - T_\phi^{l-1}}{\Delta t} v_\phi dx \\ &= \int_{\Omega} \operatorname{div}(\theta k(T_\phi^{l-1}) \nabla T_\phi^{l-1} + (1 - \theta) k(T_\phi^l) \nabla T_\phi^l) v_\phi dx, \quad l = 1, \dots, n_t \end{aligned}$$

<sup>3</sup>Der betrachtete Lösungsraum benötigt hier keine Nullrandwerte, da keine Dirichlet-Bedingungen vorliegen.

Anwendung des Gaußschen Integralsatzes liefert:

$$\begin{aligned}
 & \int_{\Omega} \rho c_p(\theta, T_{\phi}^{l-1}, T_{\phi}^l) \frac{T_{\phi}^l - T_{\phi}^{l-1}}{\Delta t} v_{\phi} dx \\
 &= - \int_{\Omega} (\theta k(T_{\phi}^{l-1}) \nabla T_{\phi}^{l-1} + (1 - \theta) k(T_{\phi}^l) \nabla T_{\phi}^l) \cdot \nabla v_{\phi} dx \\
 & \quad + \int_{\Gamma} (\theta k(T_{\phi}^{l-1}) \nabla T_{\phi}^{l-1} + (1 - \theta) k(T_{\phi}^l) \nabla T_{\phi}^l) \cdot \vec{n} v_{\phi} ds \\
 &= -\theta \int_{\Omega} k(T_{\phi}^{l-1}) \nabla T_{\phi}^{l-1} \cdot \nabla v_{\phi} dx - (1 - \theta) \int_{\Omega} k(T_{\phi}^l) \nabla T_{\phi}^l \cdot \nabla v_{\phi} dx \\
 & \quad + \theta \int_{\Gamma} \underbrace{k(T_{\phi}^{l-1}) \nabla T_{\phi}^{l-1} \cdot \vec{n}}_{=\dot{q}_{\phi}^{l-1}} v_{\phi} ds + (1 - \theta) \int_{\Gamma} \underbrace{k(T_{\phi}^l) \nabla T_{\phi}^l \cdot \vec{n}}_{=\dot{q}_{\phi}^l} v_{\phi} ds, \quad l = 1, \dots, n_t
 \end{aligned}$$

Einsetzen der Neumann-Randwerte führt auf die schwache Form der Wärmeleitungsgleichung:

$$\begin{aligned}
 & \frac{\rho}{\Delta t} \int_{\Omega} c_p(\theta, T_{\phi}^{l-1}, T_{\phi}^l) T_{\phi}^l v_{\phi} dx - \frac{\rho}{\Delta t} \int_{\Omega} c_p(\theta, T_{\phi}^{l-1}, T_{\phi}^l) T_{\phi}^{l-1} v_{\phi} dx \\
 &= -\theta \int_{\Omega} k(T_{\phi}^{l-1}) \nabla T_{\phi}^{l-1} \cdot \nabla v_{\phi} dx - (1 - \theta) \int_{\Omega} k(T_{\phi}^l) \nabla T_{\phi}^l \cdot \nabla v_{\phi} dx \quad (4.7) \\
 & \quad + \theta \int_{\Gamma} \dot{q}_{\phi}^{l-1} v_{\phi} ds + (1 - \theta) \int_{\Gamma} \dot{q}_{\phi}^l v_{\phi} ds, \quad l = 1, \dots, n_t
 \end{aligned}$$

Zusätzlich gilt die Anfangsbedingung

$$T_{\phi}^0 = T_{\phi,0}.$$

Durch Diskretisierung von (4.7) entsteht ein nichtlineares Gleichungssystem, welches anstatt der kontinuierlichen Funktionen die Vektoren  $T^l \in \mathbb{R}^{n_s^l}$ ,  $\dot{q}^l \in \mathbb{R}^{n_{\Gamma}^l}$ ,  $l = 0, \dots, n_t$  und  $T_0 \in \mathbb{R}^{n_s^0}$  enthält. Die Vektoren setzen sich aus den Freiheitsgraden der FEM in einer geeignet sortierten Reihenfolge zusammen. Die Freiheitsgrade entsprechen den diskreten Werten der Funktionen  $T_{\phi}^l$ ,  $\dot{q}_{\phi}^l$  und  $T_{\phi,0}$  an bestimmten Punkten des Gitters. Bei der Wahl linearer finiter Lagrange-Elemente entsprechen diese Punkte genau den Eckpunkten der Gitterzellen. Bei Elementen höherer Ordnung erhält man noch zusätzliche Punkte auf den Zellkanten, -rändern oder auch im Innern der Zellen.  $n_s^l$  ist hierbei die Anzahl an diskreten Werten zum Zeitpunkt  $t_l$  und  $n_{\Gamma}^l$  die Anzahl an diskreten Werten auf dem Rand  $\Gamma$  zum Zeitpunkt  $t_l$ . Durch die Wahl eines zeitvarianten Gitters können diese Vektorlängen für verschiedene  $l$  variieren.

Bei der Parameteridentifikation zur Bestimmung des Wärmeeintrags sollen nicht alle Randwerte  $\dot{q}^l$ ,  $l = 0, \dots, n_t$  identifiziert werden, sondern nur solche, die einen positiven Wärmestrom in das Bauteil aufweisen, z. B. durch einen Kontakt zum Werkzeug. Für die anderen Randwerte werden adiabate Randbedingungen ( $\dot{q} = 0$ ) angenommen. Vorangegangene Simulationen konnten zeigen, dass die Einschränkungen dieser Annahme



vernachlässigbar klein sind. Die Aufteilung des Randes in diese zwei Gruppen ist zeitvariant. Der Rand mit positivem Wärmeeintrag zum Zeitpunkt  $t_l$  wird mit  $\Gamma_+^l$  bezeichnet, der Rand mit adiabaten Bedingungen mit  $\Gamma_0^l = \Gamma \setminus \Gamma_+^l$ . Zusätzlich werden die Randstücke  $\Gamma_+^l$  in  $n_z$  disjunkte Bereiche  $\Gamma_{+,i}^l$  eingeteilt:

$$\begin{aligned}\Gamma_+^l &= \bigcup_{i=1}^{n_z} \Gamma_{+,i}^l \\ \emptyset &= \Gamma_{+,i}^l \cap \Gamma_{+,j}^l, \quad i, j = 1, \dots, n_z, \quad i \neq j, \quad l = 0, \dots, n_t\end{aligned}$$

Innerhalb der Randstücke  $\Gamma_{+,i}^l$  wird die Wärmestromdichte konstant und zeitinvariant modelliert. D. h. in den Bereichen  $\Gamma_{+,i}^l$  und  $\Gamma_{+,i}^k$ ,  $l, k = 1, \dots, n_t$  wirkt der gleiche konstante Randwert. Insgesamt ergeben sich so  $n_z$  zu identifizierende Wärmestromdichten  $\dot{q}_{\text{ident},i}$ , die durch  $i$  den Randstücken  $\Gamma_{+,i}^l$ ,  $l = 1, \dots, n_t$  zugeordnet werden, auf denen sie wirken. Die Vektoren aller Randwerte  $\dot{q}^l$  werden mit den Identifikationsparametern  $\dot{q}_{\text{ident},i}$  dadurch wie folgt definiert:

$$\begin{aligned}\dot{q}^l|_{\Gamma_{+,i}^l} &= (\dot{q}_{\text{ident},i}, \dots, \dot{q}_{\text{ident},i})^T \in \mathbb{R}^{n_{+,i}^l}, \quad i = 1, \dots, n_z \\ \dot{q}^l|_{\Gamma_0^l} &= (0, \dots, 0)^T \in \mathbb{R}^{n_0^l}, \quad l = 0, \dots, n_t\end{aligned} \tag{4.8}$$

Hierbei ist  $n_{+,i}^l$  die Anzahl der diskretisierten Werte im Gebiet  $\Gamma_{+,i}^l$  und  $n_0^l$  die Anzahl im Gebiet  $\Gamma_0^l$ .

Wenn man für  $\theta$ ,  $\rho$ ,  $\Delta t$  und  $T_0$  feste Werte wählt und (4.8) verwendet, können die  $n_t + 1$  Gleichungssysteme, die durch die Diskretisierung entstehen, allgemein<sup>4</sup> durch

$$\begin{aligned}H^l(T^l, T^{l-1}, \dot{q}_{\text{ident}}) &= 0, \quad l = 1, \dots, n_t \\ T^0 &= T_0\end{aligned}$$

formuliert werden. Als ein geschlossenes Gleichungssystem lässt sich dies durch

$$H(T, \dot{q}_{\text{ident}}) = 0$$

darstellen, wobei  $T := ((T^1)^T, \dots, (T^{n_t})^T)^T \in \mathbb{R}^{n_s}$  der Vektor aller diskretisierten Zustände und  $n_s = \sum_{l=1}^{n_t} n_s^l$  dessen Größe ist, sowie  $\dot{q}_{\text{ident}} := (\dot{q}_{\text{ident},1}, \dots, \dot{q}_{\text{ident},n_z})^T \in \mathbb{R}^{n_z}$  den Vektor aller Identifikationsparameter darstellt.

Damit lässt sich das allgemeine Parameteridentifikationsproblem (4.4) mit den Temperaturen als Zustände  $S := T$  und der Wärmestromdichte als Identifikationsparameter

<sup>4</sup>Die Gleichungssysteme werden für das Fräsen und Bohren in den jeweiligen Abschnitten noch detaillierter formuliert.

$Z := \dot{q}_{\text{ident}}$  spezifizieren:

$$\begin{aligned}
 & \min_{X \in \mathbb{R}^{n_x}} \sum_{i=1}^{n_s} \bar{w}_i (T_i - \vartheta_i(T_{\text{mess}}))^2 \\
 & \text{unter } H(T, \dot{q}_{\text{ident}}) = 0 \\
 & T_L \leq T \leq T_U \\
 & \dot{q}_L \leq \dot{q}_{\text{ident}} \leq \dot{q}_U
 \end{aligned} \tag{4.9}$$

Dabei ist  $X = (T^T, \dot{q}_{\text{ident}}^T)^T \in \mathbb{R}^{n_x}$  die Optimierungsvariable und  $\vartheta_i : \mathbb{R}^{n_{\text{mess}}} \rightarrow \mathbb{R}$ ,  $i = 1, \dots, n_s$  sind die Funktionen, die durch Interpolation zum  $i$ -ten Eintrag des Vektors der Temperaturen  $T$  die korrespondierende gemessene Temperatur aus  $T_{\text{mess}} \in \mathbb{R}^{n_{\text{mess}}}$  berechnen. Die Boxschränken von  $T$  und  $\dot{q}_{\text{ident}}$  werden mit  $T_L, T_U \in \mathbb{R}^{n_s}$  bzw.  $\dot{q}_L, \dot{q}_U \in \mathbb{R}^{n_z}$  bezeichnet. Für die Anzahl an Gleichungsnebenbedingungen gilt  $m_h = n_s$ . Sie ist also gleich der Anzahl an Zuständen.

### 4.3 Problemstellung beim Fräsen

Während die vorangegangenen Ausführungen auch als Grundlage für die Parameteridentifikation des Bohrens dienen, wird die Problemstellung nun explizit für das Fräsen spezifiziert. Beim Fräsen wird nur der Fall betrachtet, bei dem alle Materialparameter temperaturunabhängig sind, also  $\rho$  sowie  $c_p$  und  $k$  konstant sind. Der temperaturabhängige Fall wird beim Bohren betrachtet. Durch die konstanten Parameter vereinfacht sich die schwache Form (4.7) zu

$$\begin{aligned}
 & \frac{\rho c_p}{\Delta t} \int_{\Omega} T_{\phi}^l v_{\phi} dx - \frac{\rho c_p}{\Delta t} \int_{\Omega} T_{\phi}^{l-1} v_{\phi} dx \\
 & = -\theta k \int_{\Omega} \nabla T_{\phi}^{l-1} \cdot \nabla v_{\phi} dx - (1 - \theta) k \int_{\Omega} \nabla T_{\phi}^l \cdot \nabla v_{\phi} dx \\
 & + \theta \int_{\Gamma} \dot{q}_{\phi}^{l-1} v_{\phi} ds + (1 - \theta) \int_{\Gamma} \dot{q}_{\phi}^l v_{\phi} ds, \quad l = 1, \dots, n_t.
 \end{aligned}$$

Wendet man auf diese schwache Form die FEM an, entstehen im Gegensatz zur allgemeinen Variante lineare Gleichungssysteme, die mit der Verwendung von (4.8) die Form

$$\begin{aligned}
 \mathbf{A}^l T^l &= \mathbf{B}^l T^{l-1} + \mathbf{C}^l \dot{q}_{\text{ident}}, \quad l = 1, \dots, n_t \\
 T^0 &= T_0
 \end{aligned} \tag{4.10}$$

haben. Dabei sind  $\mathbf{A}^l \in \mathbb{R}^{n_s^l \times n_s^l}$  und  $\mathbf{B}^l \in \mathbb{R}^{n_s^l \times n_s^{l-1}}$  die Systemmatrizen der Zustände sowie  $\mathbf{C}^l \in \mathbb{R}^{n_s^l \times n_z}$  die Systemmatrizen der Randwerte. Die Struktur und die Werte von  $\mathbf{A}^l$ ,  $\mathbf{B}^l$  und  $\mathbf{C}^l$  sind abhängig von der Wahl der finiten Elemente und werden hier aufgrund ihrer Größe und Komplexität nicht näher spezifiziert. Die Systemmatrix  $\mathbf{B}^l$

enthält daneben noch Werte, die die Temperaturwerte  $T^{l-1}$  vom  $l-1$ -ten Gitter auf das  $l$ -te Gitter transformieren. Alternativ können die Gleichungssysteme (4.10) als ein geschlossenes System repräsentiert werden:

$$\underbrace{\begin{pmatrix} \mathbf{A}^1 & & & \\ -\mathbf{B}^2 & \mathbf{A}^2 & & \\ & \ddots & \ddots & \\ & & -\mathbf{B}^{n_t} & \mathbf{A}^{n_t} \end{pmatrix}}_{=: \mathbf{A}} \underbrace{\begin{pmatrix} T^1 \\ T^2 \\ \vdots \\ T^{n_t} \end{pmatrix}}_{= T} = \underbrace{\begin{pmatrix} \mathbf{B}^1 T_0 + \mathbf{C}^1 \dot{q}_{\text{ident}} \\ \mathbf{C}^2 \dot{q}_{\text{ident}} \\ \vdots \\ \mathbf{C}^{n_t} \dot{q}_{\text{ident}} \end{pmatrix}}_{=: \mathbf{b}(\dot{q}_{\text{ident}})} \quad (4.11)$$

$$\Leftrightarrow: \quad H(T, \dot{q}_{\text{ident}}) = 0$$

Dabei ist  $\mathbf{A} \in \mathbb{R}^{n_s \times n_s}$  und  $\mathbf{b} \in \mathbb{R}^{n_s}$ .

Bei der Parameteridentifikation des Fräsen werden keine Boxschränken für die Zustände  $T$  berücksichtigt.<sup>5</sup> Die Identifikationsparameter haben die Schranken  $\dot{q}_L = 0 \in \mathbb{R}^{n_z}$  und  $\dot{q}_U = \infty \in \mathbb{R}^{n_z}$ . Optimierungsproblem (4.9) ist mit dem SAND-Ansatz in diesem Fall in der Form

$$\begin{aligned}
 & \min_{X \in \mathbb{R}^{n_x}} \sum_{i=1}^{n_s} \bar{w}_i (T_i - \vartheta_i(T_{\text{mess}}))^2 \\
 & \text{unter } \mathbf{A} T - \mathbf{b}(\dot{q}_{\text{ident}}) = 0 \\
 & \quad 0 \leq \dot{q}_{\text{ident}} \leq \infty
 \end{aligned} \quad (4.12)$$

zu lösen. Dabei ist  $m_h = n_s$  und  $m_g = 0$ , da die Beschränkungen von  $\dot{q}_{\text{ident}}$  Boxschränken sind und vom Algorithmus separat behandelt werden. Die Optimierungsvariable ist durch  $X = (T^T, \dot{q}_{\text{ident}}^T)^T \in \mathbb{R}^{n_x}$  gegeben.

Das Optimierungsproblem (4.12) ist ein sogenanntes quadratisches Programm (QP), d. h. ein Optimierungsproblem mit quadratischer Zielfunktion und linearen Nebenbedingungen. QPs sind eine Untermenge der NLPs, weshalb auch QPs mit dem NLP-Löser WORHP gelöst werden können. Dabei ist es in der Regel sowohl für die Performanz als auch für die Konvergenz von Vorteil, dem zugrundeliegenden SQP-Verfahren analytische

<sup>5</sup>Eine Untersuchung der Parameteridentifikation mit Zustandsbeschränkungen erfolgt für das Bohren in Kapitel 5.



## 4.4 Wärmequelle, Diskretisierung und Temperaturdaten

Der Wärmeeintrag des Fräsprozesses wird makroskopisch durch eine bewegte Wärmequelle modelliert. Dabei wird kein Materialabtrag modelliert und das Bauteil in seiner gefertigten Geometrie betrachtet. In jedem diskreten Zeitschritt wird der Schnitt bzw. der daraus resultierende Wärmeeintrag einer der vier Werkzeugschneiden auf einmal durchgeführt. Die Kontaktfläche zwischen Schneide und Bauteil bei einem Schnitt hat die Form einer Sichel bzw. eines Bogens (siehe Abb. 4.1). Diese Form korrespondiert zur Trajektorie der Schneide und wird durch die Schnittgeschwindigkeit, die Vorschubgeschwindigkeit und die Eingriffsbreite bestimmt. Der Eingriffsbogen beschreibt den Bereich des Wärmeeintrags innerhalb eines Zeitschritts und wird durch  $\Gamma_+^l$  gekennzeichnet (siehe Kapitel 4.2). Nach jedem Zeitschritt wird der Eingriffsbogen entlang des Vorschubs um die Länge des Zahnvorschubs  $f_z$  verschoben.

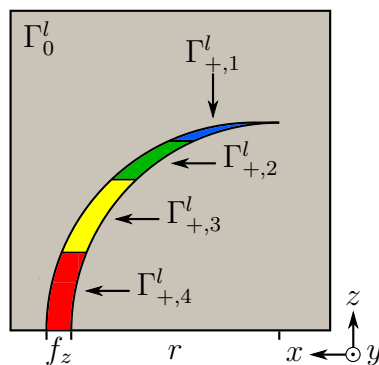


Abbildung 4.1: Geometrie des Eingriffsbogens bzw. der Wärmequelle für  $a_e = r$ , aufgeteilt in  $n_z = 4$  Bereiche

Der Eingriffsbogen  $\Gamma_+^l$  wird gemäß der Anzahl an Identifikationsparametern in  $n_z$  Bereiche  $\Gamma_{+,i}^l$ ,  $i = 1, \dots, n_z$  eingeteilt, die jeweils den gleichen Eingriffswinkel abdecken (siehe Abb. 4.1). Es wird angenommen, dass jeder Schnitt den gleichen Wärmeeintrag bewirkt, d. h. dass der Wärmestrom in den Bereichen  $\Gamma_{+,i}^l$ ,  $l = 1, \dots, n_t$  über die Zeit konstant ist. Außerhalb des Bogens, auf der Randfläche  $\Gamma_0^l$ , werden adiabate Randbedingungen angenommen. Es gelten die Neumann-Randbedingungen (4.8).

Eine komplette Simulation setzt sich aus einer Sequenz einzelner Schnitte zusammen. Abbildung 4.2 zeigt die Temperaturverteilung einer Simulation von 50 Schnitten mit den beschriebenen Randbedingungen, wobei fiktive Werte für  $\dot{q}_{\text{ident}}$  angenommen wurden. Gut sichtbar ist die erhöhte Temperatur innerhalb von  $\Gamma_+^{n_t}$  und den dadurch angedeuteten Eingriffsbogen.

Performanz-Tests haben gezeigt, dass es sinnvoll ist, die Diskretisierung und das daraus entstehende Gitter zeitlich an die Randbedingungen anzupassen. In Abb. 4.3 ist beispielhaft ein diskretes Gitter zu sehen. Die Feinheit wird mit der Nähe zum Eingriffsbogen angepasst. Das zeitvariante Gitter adaptiert die Bewegung der Wärmequelle, sodass

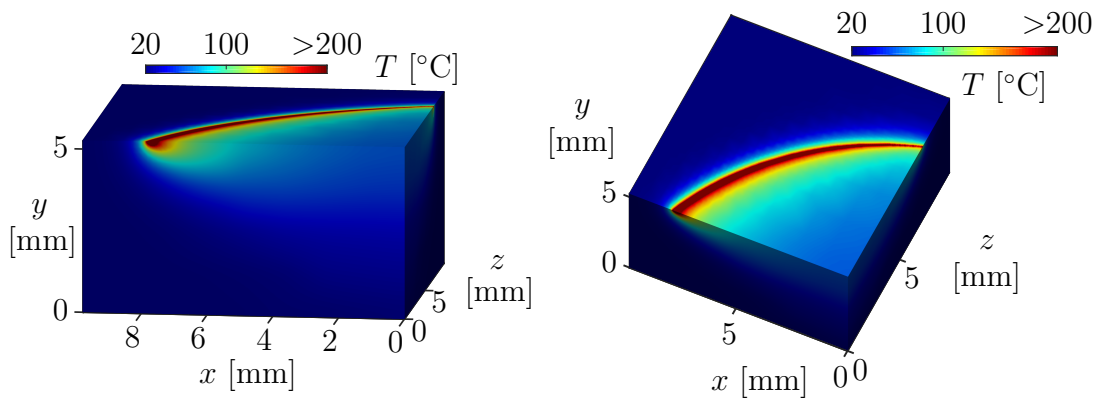


Abbildung 4.2: Temperaturverteilung nach einer Simulation von 50 Schnitten mit sichtbarem Eingriffsbogen

zum einen in Bereichen mit großen Temperaturgradienten eine präzise Berechnung erfolgt und zum anderen die Anzahl an Gitterpunkten in Bereichen niedriger Gradienten reduziert werden kann.<sup>6</sup> Auf der linken Seite von Abb. 4.3 ist zudem eine Verfeinerung der Bauteilfront zu erkennen. Dies ist die Seite, auf der der Temperaturabgleich mit den Thermografieaufnahmen erfolgt. Die erhöhte Dichte an Gitterpunkten ermöglicht einen hochauflösenden Vergleich der Temperaturfelder.

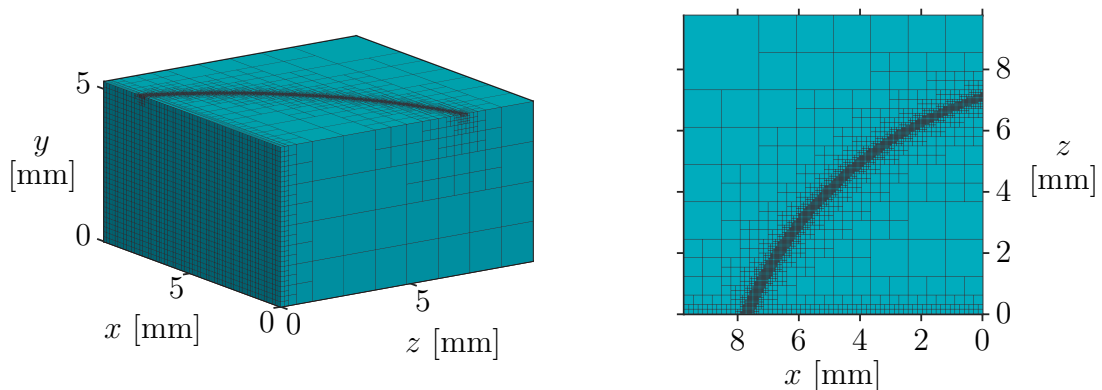


Abbildung 4.3: Gitter der FEM-Diskretisierung mit Adaption des Eingriffsbogens

Die Diskretisierung der FEM erfolgt mit linearen Lagrange-Ansatzfunktionen auf Hexaedern (vgl. Kapitel 2.3). Die Freiheitsgrade des Gleichungssystems entsprechen so den diskreten Werten der Temperaturverteilung an den Gitterpunkten. Die Zellen haben in der feinsten Größe in Abhängigkeit von den Maschinenparametern einen Durchmesser von etwa der Hälfte bis zu dem Doppelten des Zahnvorschubs  $f_z$ . Der betrachtete Ausschnitt des Werkstücks  $\Omega$  entspricht in  $x$ - und  $y$ -Koordinatenrichtung dem der Temperaturbilder, die eine Größe von  $9,76 \times 5,24 \text{ mm}^2$  haben. Die Ausdehnung von  $\Omega$  in  $z$ -Richtung wird über die Eingriffsbreite hinaus etwas vergrößert, um einen Wärmefluss

<sup>6</sup>Auf eine Gitteradaption auf Grundlage von Fehlerschätzern wird in dieser Arbeit verzichtet.

in die Tiefe zu ermöglichen.  $\Omega$  hat insgesamt eine Größe zwischen  $9,76 \times 5,24 \times 2,62 \text{ mm}^3$  und  $9,76 \times 5,24 \times 15 \text{ mm}^3$ . Es werden simulativ in Abhängigkeit vom Zahnvorschub und der Vorschubgeschwindigkeit  $n_t = 36$  bis  $n_t = 73$  Schnitte erfasst, was zu der gleichen Anzahl an Zeitschritten führt und eine Zeitspanne zwischen  $t_f - t_0 \approx 0,16 \text{ s}$  und  $t_f - t_0 \approx 0,23 \text{ s}$  abbildet. Insgesamt ergeben sich für die Parameteridentifikation in Abhängigkeit von der Eingriffsbreite  $a_e$  und der Anzahl an Zeitschritten  $n_t$  Gitter mit insgesamt  $n_s = 92\,495$  bis  $n_s = 254\,403$  diskreten Zuständen bzw. Gitterpunkten. Die für die Wärmeleitungsgleichung benötigten Materialparameter spezifische Wärmekapazität  $c_p$ , Dichte  $\rho$  und Wärmeleitfähigkeit  $k$  können Tab. 4.1 entnommen werden. Bei der Parameteridentifikation des Fräsens werden diese als temperaturunabhängig angenommen. Es werden die Werte für  $30^\circ\text{C}$  gewählt, da dies in etwa der Durchschnittstemperatur der gemessenen Temperaturfelder entspricht. Der Temperaturvergleich wird zum letzten Zeitpunkt der Simulation durchgeführt, wenn die Position der Wärmequelle der Position des Eingriffsbogens in den Thermografieaufnahmen entspricht. Räumlich liegen die gemessenen Temperaturen in der  $x$ - $y$ -Ebene mit  $z = 0$ . Das bedeutet, dass die modellierten Temperaturen  $T^{n_t}|_{z=0}$  mit den gemessenen Temperaturen  $T_{\text{mess}}$  verglichen werden.

Parameter	Symbol	Wert	Einheit
spezifische Wärmekapazität	$c_p$	465,4	J/(kg K)
Dichte	$\rho$	7830,0	kg/m <sup>3</sup>
Wärmeleitfähigkeit	$k$	45,1	W/(m K)

Tabelle 4.1: Materialparameter von 42CrMo4 für die Wärmeleitungsgleichung bei einer Temperatur von  $30^\circ\text{C}$  (Quelle: [103])

Die verwendeten Temperaturdaten wurden mit einer Kamera aufgenommen, die für niedrige Temperaturen kalibriert wurde. Die Aufnahmen zeigen generell keine Temperaturen über  $70^\circ\text{C}$ , wenngleich bei einem Fräsexperiment mit höheren Temperaturen zu rechnen ist. Es wird angenommen, dass die Genauigkeit der gemessenen Temperaturfelder im unteren zweistelligen  $^\circ\text{C}$ -Bereich am höchsten ist. Bei den Experimenten wurde eine Blende eingesetzt, um Wärmestrahlungen des Werkzeugs abzuschirmen und die Kamera vor Spänen zu schützen. Aufgrund der Nähe von Werkzeug und Blende zur Linse der Kamera ist davon auszugehen, dass Störungen die Aufnahmen beeinflussen. Die in Abb. 4.4 gezeigte Aufnahme legt die Vermutung nahe, dass die Temperaturen der eigentlichen Bauteilfront durch Wärmestrahlung des Werkzeugs oder eine Abschirmung der Strahlung durch die Blende teilweise verfälscht werden. Diese Datenfehler betreffen den oberen Bereich der Temperaturfelder. Um diesem Zusatzwissen in der Parameteridentifikation gerecht zu werden, wird nur ein Teil der Thermografieaufnahmen für die Identifikation verwendet. Zum einen finden nur die unteren 90% der Bilder Verwendung. Zum anderen haben sich die Temperaturdaten unterhalb von  $20^\circ\text{C}$  und oberhalb von  $40^\circ\text{C}$  als besonders fehleranfällig erwiesen, sodass nur Temperaturen innerhalb dieses Bereichs verwendet werden. Die normierten Gewichte  $\bar{w}_i$ ,  $i = 1, \dots, n_s$  der Zielfunktion (4.12) werden entsprechend angepasst, indem die nicht normierten Gewichte  $w_i$ ,  $i = 1, \dots, n_s$  (siehe

(4.2)) berücksichtigter Zustände bzw. Datenpunkte auf 1 gesetzt werden und die der unberücksichtigten auf 0.

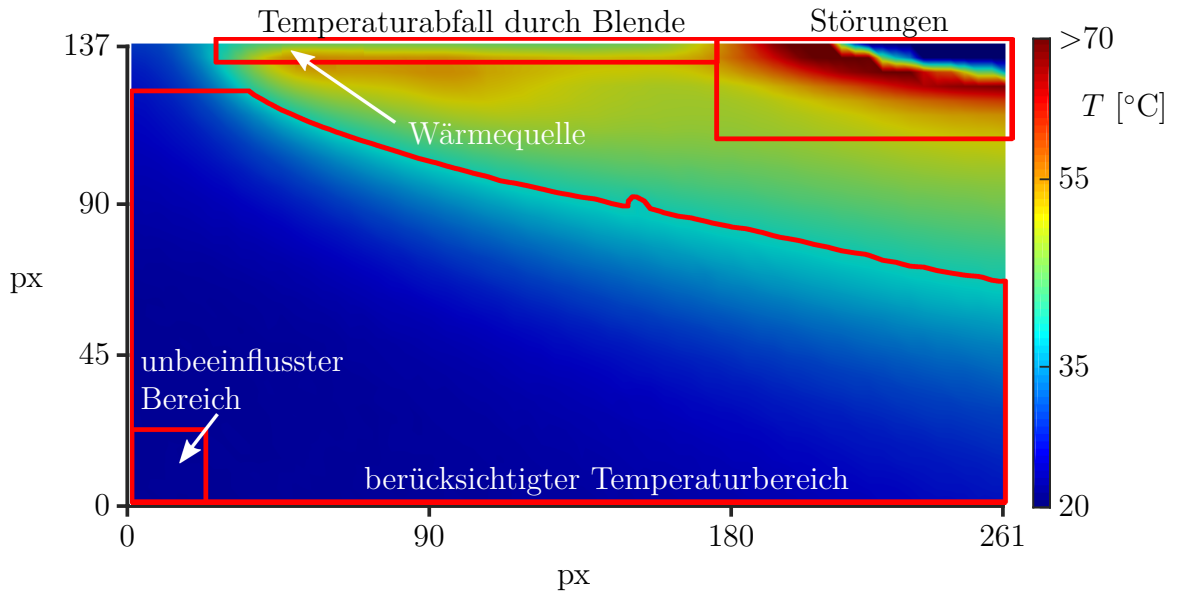


Abbildung 4.4: Thermografieaufnahme eines Fräsexperiments

Die für das SQP-Verfahren (siehe Algorithmus 2.16) benötigte Startschätzung der primalen und dualen Optimierungsvariablen  $(X^{(0)}, \lambda^{(0)})$  wird für alle Parameteridentifikationen des Fräsens wie folgt gewählt:

$$\begin{aligned} T_i^{(0)} &= 100 \text{ °C}, & i &= 1, \dots, n_s \\ \dot{q}_{\text{ident},1}^{(0)} &= 200 \text{ W/mm}^2 \\ \dot{q}_{\text{ident},i}^{(0)} &= 0 \text{ W/mm}^2, & i &= 2, \dots, n_z \\ \lambda_i^{(0)} &= 0, & i &= 1, \dots, m \end{aligned}$$

Vorherige Tests konnten empirisch zeigen, dass die Wahl der Startschätzung keinen Einfluss auf die Lösungsgüte hat, weshalb sich diese triviale Startschätzung anbietet.

Im Vorfeld der numerischen Auswertungen wurde die Parameteridentifikation und insbesondere das Lösen des Gleichungssystems (4.11) für verschiedene  $\theta$  ( $\theta \in \{0; 0,5; 1\}$ ) des Theta-Verfahrens und verschiedene Randbedingungen getestet. Dabei stellte sich heraus, dass die explizite Variante ( $\theta = 1$ ) die geringste Lösungsgenauigkeit aufweist. Die Temperaturverteilung wies dabei teilweise eine zeitliche Oszillation auf, möglicherweise weil die Differentialgleichung bezüglich der Zeit steif ist. Der Unterschied der Genauigkeit für das Trapezverfahren ( $\theta = 0,5$ ) und das implizite Eulerverfahren ( $\theta = 0$ ) ist sehr gering. Beide Verfahren liefern gute Resultate ohne Oszillation. Mit dem Trapezverfahren ergibt sich jedoch eine höhere Berechnungsdauer, da die Systemmatrizen dichter besetzt sind. Für die numerischen Berechnungen wird daher  $\theta = 0$  gewählt.



## 4.5 Numerische Ergebnisse

Die Parameteridentifikation des Fräsens wird in zwei Problemstellungen aufgeteilt. Der erste Teil umfasst eine Validierung, in der mit künstlichen Daten die grundsätzliche Funktionsfähigkeit der Methodik überprüft wird. Im zweiten Teil wird die vorgestellte Parameteridentifikation auf Realdaten angewendet, um den Wärmeeintrag von Versuchen zu quantifizieren. Das System, das für die numerischen Berechnungen verwendet wurde, ist ein leistungsstarker Desktop-Rechner mit Debian 8.1. Die Hardware ist in Tab. 4.2 zu sehen.

Hardware	Spezifikation
CPU	Intel Core i7-3770 @3.40 GHz (insg. 8 logische Kerne)
RAM	32GB
HDD	900GB
GPU	NVidia GeForce GT 610

Tabelle 4.2: Hardware des Rechners

Im Folgenden werden Variablen, die das Ergebnis einer Optimierung darstellen, mit einem  $(\cdot)^*$  gekennzeichnet (vgl. Definition 2.6).

### 4.5.1 Validierung

Die Validierung dient dazu, die Funktionalität des gewählten Ansatzes zu testen. Der gewählte Ansatz umfasst neben der Modellierung durch die bewegte Wärmequelle und der Aufstellung eines Parameteridentifikationsproblems als NLP auch das numerische Lösen mit der Optimierungssoftware WORHP. Für die Validierung werden in zwei Szenarien künstliche Daten mit einer fixierten Wärmestromdichte  $\dot{q}_{\text{valid}}$  durch einmaliges Lösen des Gleichungssystems (4.11) erzeugt. Dabei sind die Randwerte durch  $\dot{q}^l|_{\Gamma_+^l} = \dot{q}_{\text{valid}}$  gegeben ( $n_z = 1$ ). Um zu testen, wie die Methodik auf Störungen in den Daten reagiert, wird das simulierte Temperaturfeld  $T_{\text{sim}}$  anschließend verrauscht. Dafür werden die Daten mit einem Zufallswert der Normalverteilung addiert. Diese hat als Erwartungswert  $\mu_N = 0$ . Die Standardabweichung  $\sigma_N$  der Normalverteilung beträgt bei den zwei betrachteten Szenarien 10 % bzw. 20 % vom zu störenden Temperaturwert. Mit der Parameteridentifikation wird anschließend versucht, aus den gestörten Daten auf den fixierten Wert  $\dot{q}_{\text{valid}}$  zu schließen.

Um den Einfluss der Zufallszahlen zu reduzieren, bietet sich eine Monte-Carlo-Simulation an. In dieser wird das Experiment mit verschiedenen Zufallszahlen häufig wiederholt und der Durchschnitt aller Experimente betrachtet. Da die Identifikation eines verrauschten Temperaturfeldes bereits mehrere Stunden dauert, wird die Monte-Carlo-Simulation mit

zehn Wiederholungen für jedes Szenario durchgeführt. Dies stellt einen Kompromiss zwischen Berechnungsdauer und Reduzierung des Zufalls dar. Variablen des  $j$ -ten Versuchs der Monte-Carlo-Simulation werden mit  $(\cdot)_{[j]}$  gekennzeichnet.

### Erstes Szenario

Die erste Identifikation wurde mit einer Eingriffsbreite von  $a_e = 0,95$  mm durchgeführt. Um über die Eingriffsbreite hinaus einen Wärmefluss in die Tiefe zu ermöglichen, ist die Größe des Bauteilausschnitts  $\Omega = [0; 9,76] \times [0; 5,24] \times [0; 2,62]$  mm<sup>3</sup>. Die Vorschubgeschwindigkeit beträgt  $v_f = 2800$  mm/min und die Schnittgeschwindigkeit beträgt  $v_c = 373$  m/min. Dies führt zu einem Zahnvorschub von  $f_z \approx 0,15$  mm und 52 Schnitten, die in der Simulation berücksichtigt werden. Es wird eine Zeitspanne von  $t_f - t_0 \approx 0,16$  s abgedeckt. Die diskreten Temperaturen  $T_0$  des Werkstücks zum Zeitpunkt  $t_0$  betragen konstant 20 °C und die zu identifizierende Wärmestromdichte beträgt  $\dot{q}_{\text{valid}} = 140$  W/mm<sup>2</sup>. Die mit diesem Wert simulierten Daten  $T_{\text{sim}}$  (siehe Abb. 4.5a) wurden für die zehn Wiederholungen der Monte-Carlo-Simulation jeweils mit einem normalverteiltem Rauschen von 10 % gestört. In Abb. 4.5b sind, beispielhaft für den ersten Versuch, die verrauschten Daten  $T_{\text{mess}[1]}$  der Ebene  $z = 0$  mm zum Zeitpunkt  $t = t_f$  und der für die Identifikation verwendete Temperaturbereich zu sehen. Der rote Rand markiert die Temperaturen zwischen 20 °C und 40 °C und die unteren 90 % des Bildes. Die Diskretisierung der Temperaturfelder führt zusammen mit dem Identifikationsparameter  $\dot{q}_{\text{ident}}$  auf  $n_x = 126\,051$  Optimierungsvariablen. Zum Lösen des Problems wurde das Optimierungsproblem (4.12) mit den beschriebenen Werten in der Optimierungssoftware WORHP implementiert.

Die durchschnittliche Dauer des Optimierungsalgorithmus im ersten Szenario beträgt für durchschnittlich 18 Hauptiterationen 1,6 h. Ein Ausschnitt des durch die Identifikation rekonstruierten Temperaturfeldes  $T_{[1]}^*$  ist in Abb. 4.5c für den ersten Versuch der Monte-Carlo-Simulation zu sehen. Ein Unterschied zwischen diesem Temperaturfeld und den Originaldaten in Abb. 4.5a ist nur schwer auszumachen. Insbesondere auch der obere rechte Bereich, welcher nicht in der Zielfunktion berücksichtigt wird, wird präzise rekonstruiert. Dies wird durch die geringe mittlere absolute Abweichung von durchschnittlich

$$\frac{1}{10} \sum_{j=1}^{10} \left( \frac{1}{n_s} \sum_{i=1}^{n_s} |T_{i[j]}^* - T_{\text{sim},i}| \right) \approx 0,70 \text{ °C}$$

bestätigt und entspricht einer mittleren relativen Abweichung von 0,8 %. Es sei an dieser Stelle erwähnt, dass die Identifikation nur mit einem Bruchteil (0,9 %) der ursprünglichen Daten ( $x < 0,9 \cdot 5,24$  mm,  $z = 0$  mm,  $t = t_f$ ,  $20 \text{ °C} < T_{\text{mess}} < 40 \text{ °C}$ ) durchgeführt wurde, die erwähnten Abweichungen aber über alle Datenpunkte berechnet wurden. Von der vorher festgelegten Wärmestromdichte  $\dot{q}_{\text{valid}} = 140$  W/mm<sup>2</sup> weicht die durchschnittlich

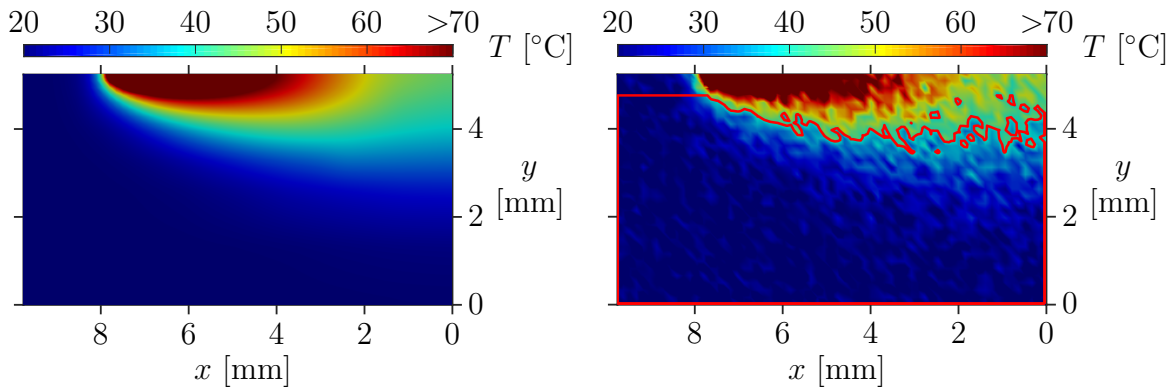
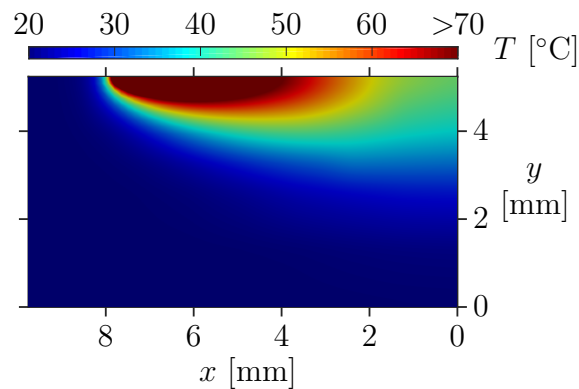
(a) ungestörte Temperaturdaten  $T_{\text{sim}}^{n_t}|_{z=0}$ (b) verrauschte Temperaturdaten  $T_{\text{mess}[1]}$ (c) identifiziertes Temperaturfeld  $T_{[1]}^{n_t^*}|_{z=0}$ 

Abbildung 4.5: Temperaturfelder des ersten Validierungsszenarios

identifizierte Wärmestromdichte

$$\frac{1}{10} \sum_{j=1}^{10} \dot{q}_{\text{ident}[j]}^* \approx 139,26 \text{ W/mm}^2$$

um  $0,74 \text{ W/mm}^2$  ab. Auch dies entspricht nur einer relativen Abweichung von  $0,5\%$ . Die größte absolute Abweichung zwischen simulierten Temperaturdaten und identifiziertem Temperaturfeld beträgt durchschnittlich  $3,4 \text{ }^\circ\text{C}$  und tritt in unmittelbarer Nähe zur Wärmequelle auf. Dies entspricht den Erwartungen, da dort aufgrund der hohen Wärmestromdichte die größten Temperaturgradienten auftreten und die Lösung der PDE am sensitivsten ist.

## Zweites Szenario

Das zweite Szenario simuliert einen Fräsprozess mit einer Eingriffsbreite von  $a_e = 12,5 \text{ mm}$ , einer Vorschubgeschwindigkeit von  $v_f = 2000 \text{ mm/min}$  und einer Schnitt-

geschwindigkeit von  $v_c = 255 \text{ m/min}$ . Dafür ist der Ausschnitt des Werkstücks auf  $\Omega = [0; 9,76] \times [0; 5,24] \times [0; 15] \text{ mm}^3$  vergrößert. Der resultierende Zahnvorschub beträgt  $f_z \approx 0,15 \text{ mm}$ . Die Zeitspanne von  $t_f - t_0 \approx 0,23 \text{ s}$  wird in 50 Zeitschritte zerlegt, wobei die gleiche Anzahl an Schnitten simuliert wird. Die diskreten Temperaturen  $T_0$  des Werkstücks zum Zeitpunkt  $t_0$  betragen konstant  $20^\circ\text{C}$ . Die Temperaturdaten wurden mit einer Wärmestromdichte von  $\dot{q}_{\text{valid}} = 70 \text{ W/mm}^2$  erzeugt und anschließend für die Monte-Carlo-Simulation zehnmal mit einem normalverteiltem Rauschen gestört. Dieses weist eine erhöhte Standardabweichung von 20 % des Datenwertes auf. Ein Ausschnitt ( $z = 0$ ,  $t = t_f$ ) der Originaldaten sowie die verrauschten Daten für den ersten Versuch sind in Abb. 4.6a bzw. 4.6b zu sehen. Auch hier sind die verwendeten Daten ( $x < 0,9 \cdot 5,24 \text{ mm}$ ,  $z = 0 \text{ mm}$ ,  $t = t_f$ ,  $20^\circ\text{C} < T_{\text{mess}} < 40^\circ\text{C}$ ) mit einem roten Rand gekennzeichnet. Für die Optimierung mit WORHP werden  $n_x = 186\,154$  Optimierungsvariablen verwendet.

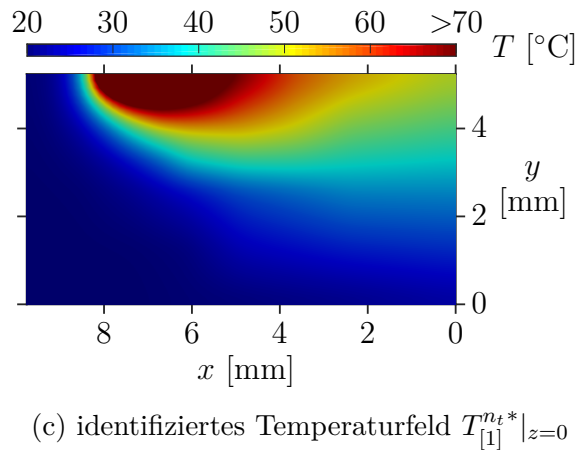
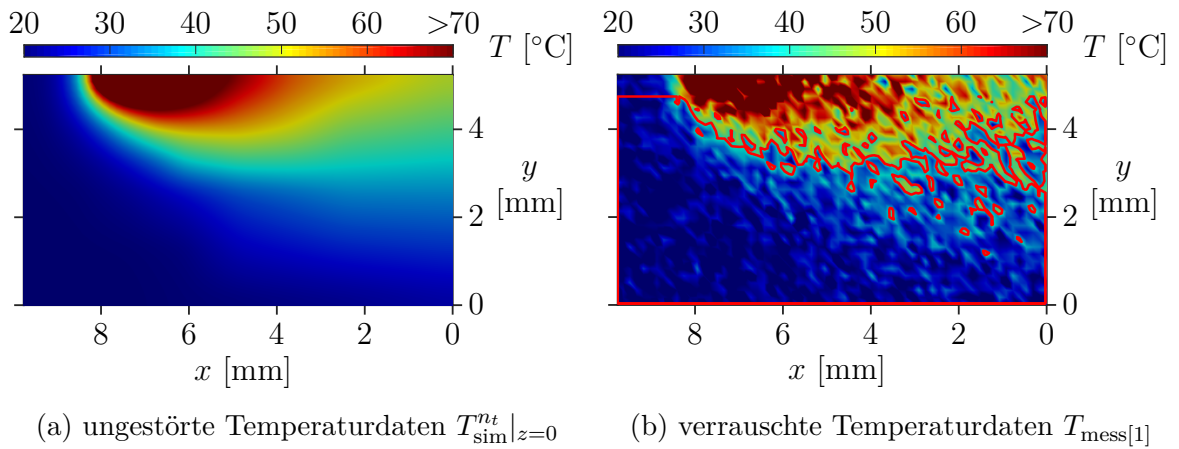


Abbildung 4.6: Temperaturfelder des zweiten Validierungsszenarios

Der Optimierungsalgorithmus benötigt durchschnittlich 20 Hauptiterationen und 4,5 h zum Auffinden der Lösung. Das rekonstruierte Temperaturfeld des ersten Versuchs ist für  $z = 0 \text{ mm}$  und  $t = t_f$  in Abb. 4.6c zu sehen. Erneut ist die Datenrekonstruktion trotz des erhöhten Rauschens erfolgreich. Originaldaten und identifizierte Temperatu-

ren stimmen optisch überein. Dies trifft auch auf den oberen rechten Bereich zu, der im Optimierungsprozess unberücksichtigt blieb. Die mittlere Abweichung beträgt im Durchschnitt

$$\frac{1}{10} \sum_{j=1}^{10} \left( \frac{1}{n_s} \sum_{i=1}^{n_s} |T_{i[j]}^* - T_{\text{sim},i}| \right) \approx 0,61 \text{ } ^\circ\text{C}.$$

Die mittlere relative Abweichung von 0,7% ist im Vergleich zum ersten Szenario sogar geringer. Auch hier wurde verglichen mit der Anzahl an allen simulierten Datenpunkten mit 0,6% nur ein kleiner Teil für die Identifikation verwendet. Die identifizierte Wärmestromdichte ist mit durchschnittlich

$$\frac{1}{10} \sum_{j=1}^{10} \dot{q}_{\text{ident}[j]}^* \approx 69,48 \text{ W/mm}^2$$

um  $0,52 \text{ W/mm}^2$  von der ursprünglichen Wärmestromdichte  $\dot{q}_{\text{valid}} = 70 \text{ W/mm}^2$  entfernt. Dies entspricht einer relativen Abweichung von 0,7%. Die größte Abweichung zwischen Daten und identifizierten Temperaturen tritt erneut in der näheren Umgebung der Wärmequelle auf und beträgt  $4,5 \text{ } ^\circ\text{C}$  im Durchschnitt über die zehn Versuche der Monte-Carlo-Simulation.

Die Optimierung dieses zweiten Szenarios liefert trotz des höheren Rauschens von 20% im Vergleich zum ersten Szenario ähnliche Resultate. Die kleinere Wärmestromdichte im zweiten Szenario führt zu kleineren Temperaturgradienten, was möglicherweise das numerische Lösen der PDE erleichtert und das erhöhte Rauschen kompensiert. Insgesamt sind die Resultate der Validierung zufriedenstellend. Insbesondere konnte gezeigt werden, dass für eine erfolgreiche Identifikation schon kleine Datenausschnitte ausreichen. Dies motiviert die Erprobung des Identifikationsansatzes an Realdaten.

## 4.5.2 Identifikation von Realdaten

In diesem Abschnitt wird die vorgestellte Methodik auf Thermografieaufnahmen realer Experimente angewendet. Die Daten stammen aus den Experimenten, die in Kapitel 3.4.2 beschrieben werden. Die Versuche unterscheiden sich in der Kombination der verwendeten Maschinenparameter. Diese sind in Tab. 4.3 abgebildet. Die Schnittgeschwindigkeit  $v_c$  und die Vorschubgeschwindigkeit  $v_f$  wurden auf zwei Stufen variiert, die Eingriffsbreite auf vier Stufen. Jede Kombination der drei Maschinenparameter wurde fünfmal wiederholt. Insgesamt stehen die Bilder von 80 Experimenten zur Verfügung. Für diese Arbeit werden davon einzelne Versuche ausgewählt und vorgestellt. Die Anfangstemperatur  $T_0$  des Bauteils zu Beginn des Fräsprozesses ist aus den Experimenten nicht genau bekannt. Deshalb wird die Temperatur aus dem linken unteren Bereich der Bilder (siehe Abb. 4.4) ausgelesen, da dieses Gebiet am wenigsten vom Wärmeeintrag beeinflusst ist und davon auszugehen ist, dass die Temperaturen in diesem Bereich nahezu der Anfangstemperatur entsprechen. Die Anzahl an Identifikationsparametern

Stufe	$v_c$ [m/min]	$v_f$ [mm/min]	$a_e$ [mm]
1	255	2000	0,95
2	373	2800	3,66
3	-	-	7,72
4	-	-	12,5

Tabelle 4.3: Verwendete Maschinenparameter der thermischen Experimente des Fräsens

wird der Modellierung von Sub-Modell 2 (siehe Kapitel 3.1.2) angepasst. Das bedeutet, dass bei der Eingriffsbreite von Stufe 1 ( $a_e = 0,95$  mm) ein Identifikationsparameter verwendet wird und ein konstanter Bereich für den Wärmeeintrag entsteht. Bei Stufe 4 ( $a_e = 12,5$  mm) gilt  $n_z = 4$  und es entstehen vier Bereiche verschiedenen Wärmeeintrags. Der Unterschied zur Modellierung in Sub-Modell 2 ist hier, dass die Wärmestromdichten der kleineren Stufen nicht auf die nächste Stufe übertragen werden, sondern für jede Maschinenparameterkombination die Wärmestromdichten aller Bereiche separat ermittelt werden. Dies hat den Vorteil, dass für unterschiedliche Eingriffsbreiten auch unterschiedliche Tiefenprofile von Wärmestromdichten möglich sind. D. h. in den Bereichen  $\Gamma_{+,1}^l$ ,  $\Gamma_{+,2}^l$  und  $\Gamma_{+,3}^l$  sind jeweils verschiedene Wärmestromdichten für verschiedene Stufen der Eingriffsbreite erlaubt.

### Erstes Experiment

Das erste Experiment, von dem die Wärmestromdichte mittels Parameteridentifikation ermittelt wurde, wurde mit den Maschinenparametern  $a_e = 0,95$  mm,  $v_c = 373$  m/min und  $v_f = 2800$  mm/min durchgeführt. Der modellierte Bauteilausschnitt ist dadurch  $\Omega = [0; 9,76] \times [0; 5,24] \times [0; 2,62]$  mm<sup>3</sup> groß. Die Schnitt- und Vorschubgeschwindigkeit führen auf einen Zahnvorschub von  $f_z \approx 0,15$  mm und zu  $n_t = 52$  Zeitschritten in der Simulation. Die diskreten Temperaturen aller Zeitschritte und der Identifikationsparameter  $\dot{q}_{\text{ident}}$  ergeben zusammen  $n_x = 126051$  Optimierungsvariablen. Die betrachtete Zeitspanne beträgt  $t_f - t_0 \approx 0,16$  s und die aus den Daten abgelesene Anfangstemperatur 20,0 °C. Das während des Experiments aufgenommene Temperaturfeld ist in Abb. 4.7a zu sehen. Der rote Bereich markiert erneut die Messpunkte, die für die Identifikation benutzt wurden ( $x < 0,9 \cdot 5,24$  mm,  $z = 0$  mm,  $t = t_f$ ,  $20$  °C  $< T_{\text{mess}} < 40$  °C). Dieser ist vergleichsweise groß, da das Bild größtenteils niedrige Temperaturwerte aufweist, die unter 40 °C liegen. Die Durchschnittstemperatur der Daten liegt bei 27,5 °C. Im Temperaturfeld treten im oberen Bereich ( $y > 4,5$  mm) Störungen auf, die schon in Abb. 4.4 beschrieben wurden. Die Ursache wird hier im Bereich  $4$  mm  $< x < 8$  mm in der Abschirmung der Wärmestrahlung durch die Blende vermutet, die in den Experimenten zum Abdecken des Werkzeugs verwendet wurde. Im Bereich  $0$  mm  $< x < 4$  mm überlagern möglicherweise Wärmestrahlungen des Werkzeugs das Temperaturfeld des Bauteils.

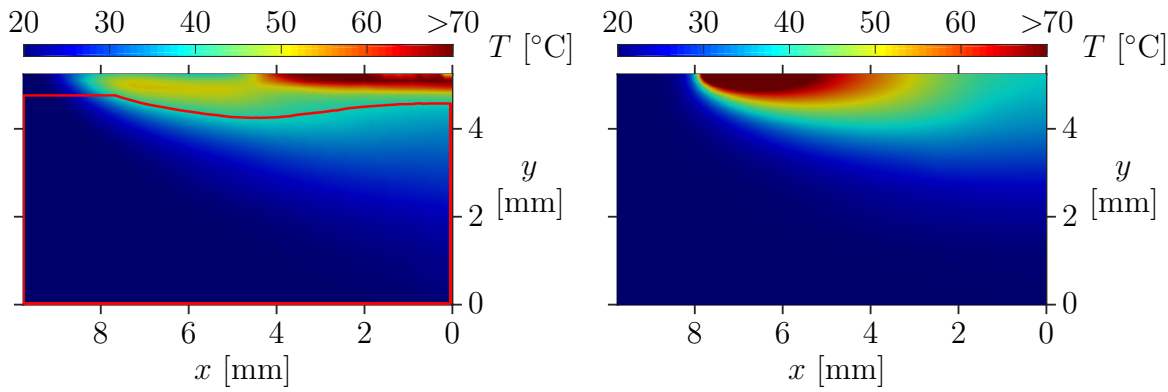
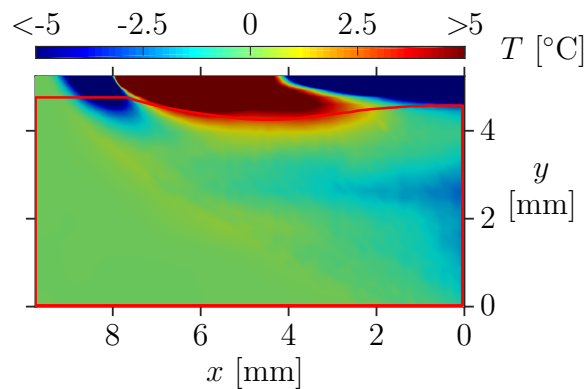
(a) Temperaturdaten  $T_{\text{mess}}$ (b) identifiziertes Temperaturfeld  $T^{n_t^*}|_{z=0}$ (c) Temperaturdifferenz  $T^{n_t^*}|_{z=0} - T_{\text{mess}}$ 

Abbildung 4.7: Temperaturfelder der ersten Identifikation von Realdaten

Der Identifikationsprozess mit der Software WORHP benötigt 13 Hauptiterationen und dauert 1 h. Die identifizierte Wärmestromdichte beträgt  $\dot{q}_{\text{ident}}^* \approx 106,8 \text{ W/mm}^2$ . Das identifizierte Temperaturfeld für  $z = 0 \text{ mm}$  und zum Zeitpunkt  $t = t_f$  ist in Abb. 4.7b zu finden. Es reproduziert die gemessenen Temperaturen in großen Bereichen, insbesondere den Bereich der verwendeten Temperaturen (roter Rand). Dies wird durch Abb. 4.7c gestützt. Sie zeigt die Differenz zwischen identifizierten und gemessenen Temperaturen. Der Großteil der Temperaturdifferenzen befindet sich in einem Bereich zwischen  $-1^\circ\text{C}$  und  $1^\circ\text{C}$ . Die größten Abweichungen treten im oberen Bereich der Thermografieaufnahme auf. Die zuvor diskutierten Effekte am oberen Bildrand werden durch das verwendete Modell nicht berücksichtigt, weshalb eine perfekte Identifikation nicht erwartet werden kann. Vielmehr kann diese Methodik auch als Möglichkeit angesehen werden, Datenbereiche zu rekonstruieren, die von der Bildgebung nicht erfasst werden. Die große Abweichung im Bereich der Wärmequelle lässt vermuten, dass die Temperaturen von der Wärmebildkamera nicht vollständig korrekt erfasst wurden, da in der unmittelbaren Umgebung des Schneideneingriffs mit Temperaturen jenseits von  $100^\circ\text{C}$  zu rechnen ist, die Aufnahmen aber im Maximum nur Temperaturen bis  $70^\circ\text{C}$  aufweisen. Wie zu

erwarten gelingt die größte Übereinstimmung zwischen Daten und simulierten Werten im Inneren des rot umrandeten Bereichs. Die Differenz liegt hier häufig sehr nah bei 0. Die mittlere absolute Abweichung beträgt hier nur

$$\sum_{i=1}^{n_s} \bar{w}_i |T_i^* - \vartheta_i(T_{\text{mess}})| \approx 0,8 \text{ } ^\circ\text{C}.$$

Die mittlere relative Abweichung beträgt

$$\sum_{i=1}^{n_s} \bar{w}_i \left| \frac{T_i^* - \vartheta_i(T_{\text{mess}})}{\vartheta_i(T_{\text{mess}})} \right| \cdot 100 \% \approx 3,0 \%.$$

Dies wird durch den niedrigen Zielfunktionswert von  $F(X^*) \approx 1,2 \text{ } ^\circ\text{C}^2$  bekräftigt. Er beschreibt die mittlere quadratische Abweichung der identifizierten Temperaturen von den gemessenen.

Die vergleichsweise niedrigen Temperaturen des Temperaturfeldes in Abb. 4.7b gehen aus der kleinen Eingriffsbreite hervor. Da sich nur  $a_e = 0,95 \text{ mm}$  von  $r = 12,5 \text{ mm}$  im Eingriff befinden, ist die Fläche, durch die Wärme ins Bauteil eingetragen wird, verhältnismäßig klein. Dies kann nur in Teilen von der hohen Wärmestromdichte kompensiert werden. Der im Vergleich zu den anderen Identifikationen (siehe nachfolgende Ergebnisse) in  $x$ -Richtung langgezogene Bereich, in denen Temperaturen über  $70 \text{ } ^\circ\text{C}$  erreicht werden, ist begründet durch die Form der Wärmequelle. Bei kleinem  $a_e$  befindet sich der Bereich des Bogens, der die Wärmequelle in  $x$ -Richtung breit werden lässt, näher an der Bauteilfront.

## Zweites Experiment

Im zweiten Experiment ist die Eingriffsbreite auf  $a_e = 7,72 \text{ mm}$  erhöht, während die Schnittgeschwindigkeit auf  $v_c = 255 \text{ m/min}$  reduziert ist. Die Vorschubgeschwindigkeit ist mit  $v_f = 2800 \text{ mm/min}$  unverändert. Für die erhöhte Eingriffsbreite ist ein in  $z$ -Richtung vergrößertes Bauteil  $\Omega = [0; 9,76] \times [0; 5,24] \times [0; 9,76] \text{ mm}^3$  nötig. Der Zahnvorschub  $f_z \approx 0,22 \text{ mm}$  ist aufgrund der geringeren Schnittgeschwindigkeit größer als im ersten Experiment. Die Anzahl der Schnitte wird dadurch ebenfalls reduziert, was in  $n_t = 36$  Zeitschritten resultiert und eine Zeitspanne von  $t_f - t_0 \approx 0,16 \text{ s}$  abdeckt. Die Anfangstemperatur des Bauteils zu Beginn des simulierten Fräsprozesses beträgt  $23,0 \text{ } ^\circ\text{C}$ . Das gemessene Temperaturfeld, welches mit dem Endtemperaturfeld der Simulation verglichen wird, ist in Abb. 4.8a zu sehen. Der rot umrandete Datenbereich mit den für die Identifikation verwendeten Messpunkten ist hier etwas kleiner, wenn auch immer noch vergleichsweise groß. Die Durchschnittstemperatur ist mit  $29,4 \text{ } ^\circ\text{C}$  etwas größer. Der obere Bildrand im Bereich  $4 \text{ mm} < x < 8 \text{ mm}$  scheint nicht durch Abschirmungen der Blende gestört zu sein, da kein Temperaturabfall zu erkennen ist. Die Erwärmung im Bereich  $y > 5 \text{ mm}$  und  $x < 3 \text{ mm}$  könnte erneut durch Wärmestrahlungen des Werkzeugs hervorgerufen werden.



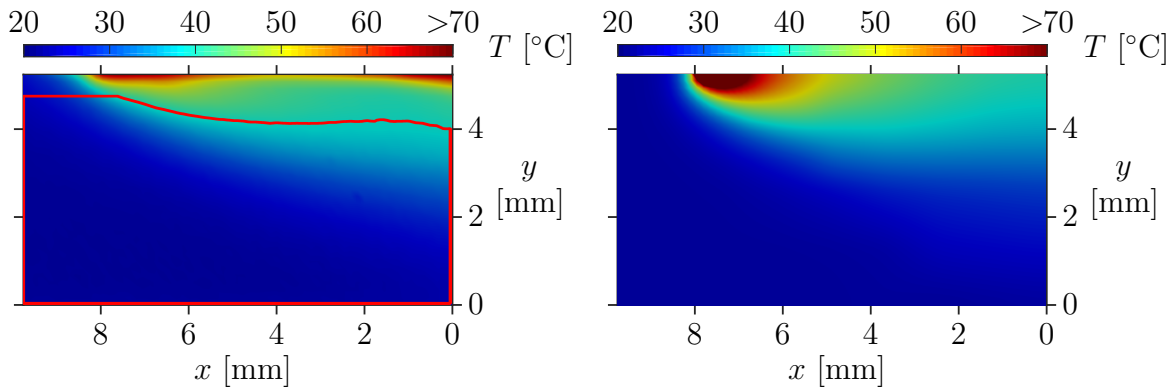
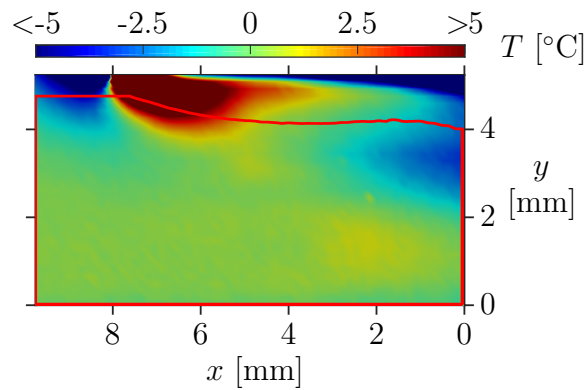
(a) Temperaturdaten  $T_{\text{mess}}$ (b) identifiziertes Temperaturfeld  $T^{n_t^*}|_{z=0}$ (c) Temperaturdifferenz  $T^{n_t^*}|_{z=0} - T_{\text{mess}}$ 

Abbildung 4.8: Temperaturfelder der zweiten Identifikation von Realdaten

Der Parameteridentifikation dieses Experiments stehen  $n_x = 98356$  Optimierungsvariablen zur Verfügung, wovon  $n_z = 3$  die zu identifizierenden Wärmestromdichten sind. Die Anzahl ist aufgrund der geringen Anzahl an Zeitschritten trotz des größeren Bauteilausschnitts im Vergleich zum ersten Experiment reduziert. Die Optimierung findet nach 0,9 h mit

$$\dot{q}_{\text{ident}}^* \approx \begin{pmatrix} 54,8 \\ 66,8 \\ 34,3 \end{pmatrix} \text{ W/mm}^2$$

ein Optimum und benötigt dafür 13 Hauptiterationen. Diese Werte sind verglichen zur Wärmestromdichte des ersten Experiments deutlich kleiner. Das liegt daran, dass durch die große Eingriffsbreite und den großen Zahnvorschub eine große Fläche des Eingriffsbogens entsteht. Das aus der Identifikation resultierende Temperaturfeld ist in Abb. 4.8b dargestellt. Es approximiert die Daten zu großen Teilen. Abbildung 4.8c zeigt die Differenz zwischen identifizierten und gemessenen Temperaturen. Während der für die

Identifikation verwendete Temperaturbereich erneut fast ausschließlich Abweichungen kleiner  $1^\circ\text{C}$  aufweist, treten im oberen Bildbereich auch Abweichungen auf, die größer sind als  $5^\circ\text{C}$ . Der Bereich der Wärmequelle in der Simulation ist wie zuvor deutlich wärmer als der entsprechende Datenbereich, der nur Temperaturen bis  $70^\circ\text{C}$  beinhaltet. Die Bereiche links und rechts der Wärmequelle hingegen sind in der Thermografieaufnahme wärmer. Dennoch liegen auch außerhalb des rot umrandeten Bereichs Messung und Daten häufig nah beieinander. Im Inneren des Bereichs wird eine Approximation erzielt, die größtenteils eine Abweichung in der Nähe von  $0^\circ\text{C}$  aufweist. Die mittlere absolute Abweichung liegt hier bei

$$\sum_{i=1}^{n_s} \bar{w}_i |T_i^* - \vartheta_i(T_{\text{mess}})| \approx 0,7^\circ\text{C}$$

und die mittlere relative Abweichung bei

$$\sum_{i=1}^{n_s} \bar{w}_i \left| \frac{T_i^* - \vartheta_i(T_{\text{mess}})}{\vartheta_i(T_{\text{mess}})} \right| \cdot 100\% \approx 2,3\%.$$

Der optimale Zielfunktionswert  $F(X^*) \approx 1,3^\circ\text{C}^2$  bestätigt dieses Resultat.

Der in Abb. 4.8b kleine Bereich mit Temperaturen über  $70^\circ\text{C}$  ist begründet durch die verhältnismäßig kleine identifizierte Wärmestromdichte. Durch die große Eingriffsbreite ist der Temperaturbereich zwischen  $40^\circ\text{C}$  und  $50^\circ\text{C}$ , der hinter der Wärmequelle hergezogen wird, vergrößert. Dies sorgt zusammen mit der hohen Starttemperatur für ein Temperaturfeld mit vergleichsweise hohen Temperaturen.

### Drittes Experiment

Das dritte Experiment, dessen Wärmestromdichte und Temperaturfeld mit der Parameteridentifikation identifiziert werden, besitzt die Maschinenparameter  $a_e = 3,66\text{ mm}$ ,  $v_c = 373\text{ m/min}$  und  $v_f = 2000\text{ mm/min}$ . Der Ausschnitt des Bauteils, der vom Modell für die Simulation erfasst wird, beträgt  $\Omega = [0; 9,76] \times [0; 5,24] \times [0; 5,24]\text{ mm}^3$ . Die hohe Schnitt- und kleine Vorschubgeschwindigkeit führen mit  $f_z \approx 0,11\text{ mm}$  auf den kleinsten Zahnvorschub der gesamten Versuchsreihe. Dadurch ist die Anzahl an Zeitschritten  $n_t = 73$  am höchsten. Die simulativ erfasste Zeitspanne beträgt  $t_f - t_0 \approx 0,23\text{ s}$ . Die Anfangstemperatur, abgelesen aus der linken unteren Ecke der Temperaturdaten, beträgt  $25,1^\circ\text{C}$ . Das zugehörige Bild der Infrarotkamera ist in Abb. 4.9a zu sehen. Der für die Identifikation relevante Datenbereich mit rotem Rand ist bei diesem Experiment deutlich kleiner als in den beiden Experimenten zuvor. Dies spiegelt sich auch in der hohen Durchschnittstemperatur von  $37,2^\circ\text{C}$  wider. Im Bereich von  $y > 4,5\text{ mm}$  und  $x < 8\text{ mm}$  könnten erneut die erwähnten Effekte für Störungen im Temperaturfeld verantwortlich sein. Der Temperaturabfall am oberen Rand im Bereich  $x < 8\text{ mm}$  ist hier jedoch nicht so stark ausgeprägt. Die Störung in der oberen rechten Ecke des Bildes ist dagegen sehr

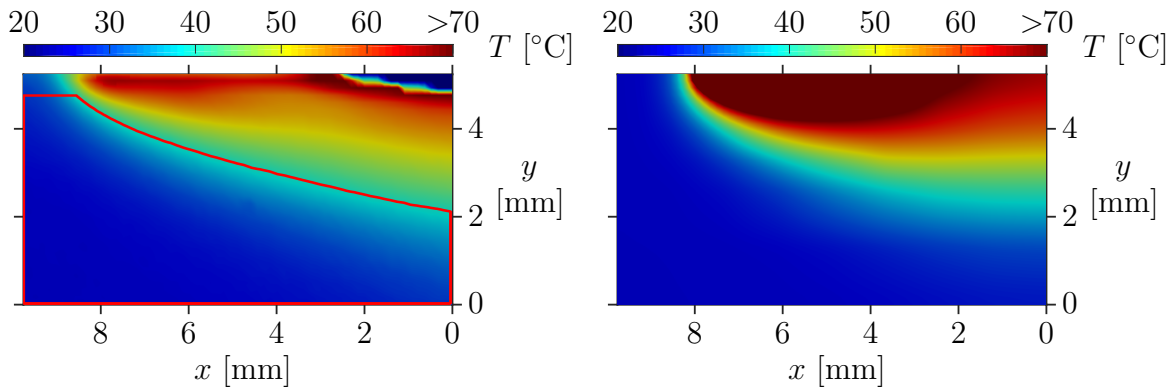
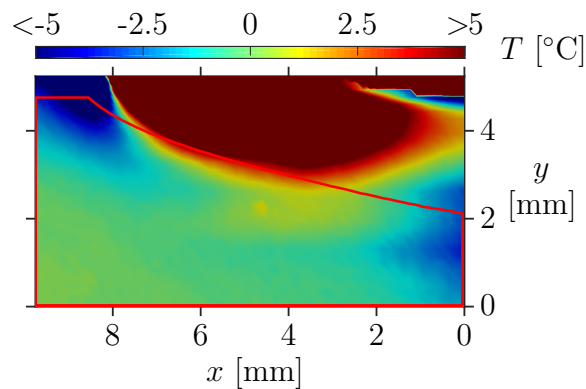
(a) Temperaturdaten  $T_{\text{mess}}$ (b) identifiziertes Temperaturfeld  $T^{nt*}|_{z=0}$ (c) Temperaturdifferenz  $T^{nt*}|_{z=0} - T_{\text{mess}}$ 

Abbildung 4.9: Temperaturfelder der dritten Identifikation von Realdaten

groß. Diese zum Inneren hin sehr hohe und nach außen stark abfallende Temperatur lässt sich nicht ausschließlich durch Wärmeleitung im Bauteil oder Wärmestrahlung des Werkzeugs erklären. Die genaue Ursache bleibt unklar.

Der Optimierung stehen bei diesem Versuch zwei Identifikationsvariablen und insgesamt  $n_x = 251993$  Optimierungsvariablen zur Verfügung. Es dauert 4,4 h bis zur Konvergenz nach 17 Hauptiterationen. Durch die vielen Zeitschritte ist sowohl die Anzahl der Variablen als auch die Rechenzeit erhöht. Das Ergebnis beinhaltet optimierte Wärmestromdichten von

$$\dot{q}_{\text{ident}}^* \approx \begin{pmatrix} 108,4 \\ 155,6 \end{pmatrix} \text{ W/mm}^2$$

und den in Abb. 4.9b gezeigten Ausschnitt des identifizierten Temperaturfeldes. Es approximiert die Daten im rot umrandeten Bereich fast vollständig in einem Toleranzbereich von  $1^\circ\text{C}$ . Zum Großteil sind die Abweichungen merklich kleiner als  $1^\circ\text{C}$  und nah bei 0. Gestützt wird dies durch die in Abb. 4.9c gezeigten Differenzen. Anders sieht

dies außerhalb der bei der Identifikation berücksichtigten Temperaturen aus. Durch die hohen gemessenen Temperaturen und der damit einhergehenden hohen identifizierten Wärmestromdichte entsteht im simulierten Temperaturfeld ein großer Bereich mit Temperaturen über 70 °C. Da die Wärmebilder aber nur Temperaturen kleiner als 70 °C wiedergeben, sind die Abweichungen hier vergleichsweise groß. Auch der erkennbare Datenfehler in der oberen rechten Ecke des Bildes mündet in vergleichsweise hohen Differenzen.

Die mittlere absolute Abweichung innerhalb des roten Randes ist mit

$$\sum_{i=1}^{n_s} \bar{w}_i |T_i^* - \vartheta_i(T_{\text{mess}})| \approx 0,9 \text{ °C}$$

etwas höher als zuvor. Die korrespondierende relative Abweichung ist

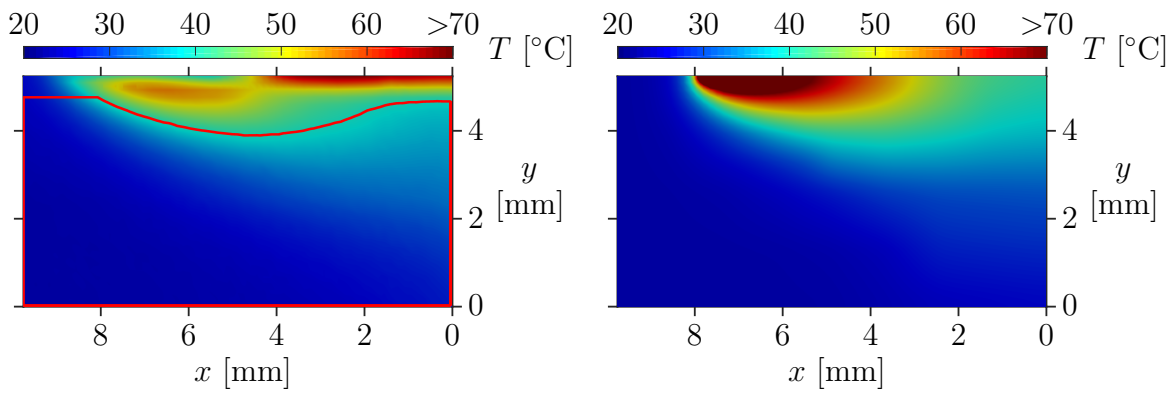
$$\sum_{i=1}^{n_s} \bar{w}_i \left| \frac{T_i^* - \vartheta_i(T_{\text{mess}})}{\vartheta_i(T_{\text{mess}})} \right| \cdot 100 \% \approx 3,0 \%$$

Der optimale Zielfunktionswert von  $F(X^*) \approx 1,9 \text{ °C}^2$  verdeutlicht, dass dieser dritte Datensatz schwieriger zu approximieren ist als die bisherigen.

#### Viertes Experiment

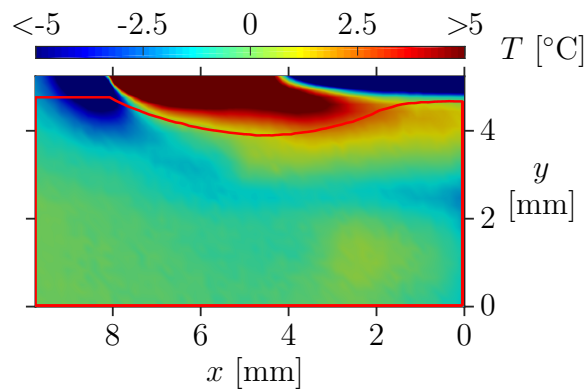
Das letzte Experiment, welches in diesem Abschnitt mit der Parameteridentifikation untersucht wird, wurde mit den Maschinenparametern  $a_e = 0,95 \text{ mm}$ ,  $v_c = 373 \text{ m/min}$  und  $v_f = 2000 \text{ mm/min}$  durchgeführt. Es unterscheidet sich damit nur in der Vorschubgeschwindigkeit vom ersten Experiment. Das Modell besitzt den gleichen Bauteilausschnitt  $\Omega = [0; 9,76] \times [0; 5,24] \times [0; 2,62] \text{ mm}^3$ . Der Zahnvorschub beträgt  $f_z \approx 0,11 \text{ mm}$ . Durch die kleinere Stufe der Vorschubgeschwindigkeit gibt es 73 Schnitte, die in  $n_t = 73$  Zeitschritten simuliert werden. Die simulierte Dauer des Fräsprozesses ist  $t_f - t_0 \approx 0,23 \text{ s}$ . Die Anfangstemperatur zu Beginn der Simulation ist 23,6 °C. Abbildung 4.10a zeigt das Temperaturfeld, das während des Versuchs mit einer Infrarotkamera aufgenommen wurde. Es ähnelt der Aufnahme des ersten Experiments in Abb. 4.7a stark. Insgesamt weist es aber eine höhere Durchschnittstemperatur von 31,8 °C auf. Der rot umrandete Bereich der Temperaturen unter 40 °C ist dadurch etwas kleiner. Die Effekte im oberen Bildbereich sind ähnlich zum ersten Versuch. Die Störung oben rechts im Bild erscheint etwas kleiner, der Bereich der Wärmequelle etwas wärmer.

Die Software WORHP benötigt 1,4 h und 18 Hauptiterationen, um in einer Optimierung mit  $n_x = 164795$  Optimierungsvariablen das Minimum zu finden. Die identifizierte Wärmestromdichte beträgt  $\dot{q}_{\text{ident}}^* \approx 122,7 \text{ W/mm}^2$  und ist damit etwas höher als im ersten Experiment. So ist auch das identifizierte Temperaturfeld in Abb. 4.10b etwas wärmer. Erneut ist die Approximation größtenteils nicht weiter als 1 °C von den Daten entfernt. Insbesondere der rot umrandete Bereich weicht nur in wenigen Punkten davon ab, wie



(a) Temperaturdaten  $T_{\text{mess}}$

(b) identifiziertes Temperaturfeld  $T^{n_t^*}|_{z=0}$



(c) Temperaturdifferenz  $T^{n_t^*}|_{z=0} - T_{\text{mess}}$

Abbildung 4.10: Temperaturfelder der vierten Identifikation von Realdaten

in Abb. 4.10c zu sehen ist. In diesem Bereich beträgt die mittlere absolute Abweichung des identifizierten Temperaturfeldes von den Daten

$$\sum_{i=1}^{n_s} \bar{w}_i |T_i^* - \vartheta_i(T_{\text{mess}})| \approx 0,8 \text{ } ^\circ\text{C}$$

bzw.

$$\sum_{i=1}^{n_s} \bar{w}_i \left| \frac{T_i^* - \vartheta_i(T_{\text{mess}})}{\vartheta_i(T_{\text{mess}})} \right| \cdot 100 \% \approx 2,7 \%$$

Der optimale Zielfunktionswert  $F(X^*) \approx 1,2 \text{ } ^\circ\text{C}^2$  bekräftigt die geringe Abweichung.

## 4.6 Fazit der Parameteridentifikation des Fräsens

Die Parameteridentifikation ist eine Methodik, die dazu eingesetzt werden kann, großflächig gemessene Temperaturfelder zu approximieren und nicht messbare physikalische

Größen wie die Wärmestromdichte zu identifizieren. Durch die Minimierung einer Zielfunktion werden dabei Modellzustände möglichst gut an die Messgrößen angepasst. Der hier gewählte SAND-Ansatz für die PDE-restringierte Optimierung scheint, wenn auch ohne Vergleichsmöglichkeit<sup>7</sup>, effizient, da die Optimierungsaufgaben mit mehr als 100 000 Variablen in wenigen Stunden<sup>8</sup> gelöst werden.

Die zunächst durchgeführte Validierung ist ein Indiz für die Funktionstüchtigkeit der Methodik. Modellbasierte Daten, die mit einem normalverteilten Rauschen von bis zu 20 % gestört werden, lassen sich nahezu perfekt rekonstruieren. Dies gelingt trotz der Tatsache, dass ein Großteil der Daten bei der Identifikation unberücksichtigt bleibt. Ein Bruchteil der Messpunkte genügt, um auch die Daten, die nicht in die Zielfunktion einfließen, zu identifizieren. Dies ist insbesondere für die Identifikation von realen Prozessen interessant, da häufig im Kontext echter Versuche die Datenlage nicht vollständig ist.

Dies ist der Fall bei den anschließend untersuchten Fräsexperimenten. Die Temperaturaufnahmen zeigen lediglich das Temperaturfeld einer Oberflächenseite des Werkstücks zu einem Zeitpunkt. Die vorhandenen Bilder sind beeinträchtigt durch Blenden und zusätzliche Wärmequellen außerhalb des modellierten Bauteils. Durch die unmittelbare Nähe des Werkzeugs zur gemessenen Oberfläche ist es unmöglich, ungestörte Temperaturfelder ausschließlich vom Bauteil aufzunehmen. Es erschien daher sinnvoll, nur einen Teil dieser Daten aus einem weniger störungsanfälligen Bereich zu verwenden. Die Rekonstruktion der Temperaturverteilung dieser Bereiche gelingt zu einem hohen Maße. Die Abweichungen sind gering und liegen im Mittel bei unter einem Grad Celsius. Darüber hinaus ist die verwendete Methodik eine Möglichkeit, fehlerhafte oder durch andere thermische Effekte überlagerte Daten entsprechend der Modellgenauigkeit zu korrigieren. Das identifizierte Temperaturfeld kann darüber Aufschluss geben, wie der tatsächliche Zustand des Bauteils zum Aufnahmezeitpunkt war. Zusammenfassend ist die SAND-basierte Parameteridentifikation eine geeignete Erweiterung des Spektrums der Methoden zur indirekten Bestimmung von Wärmestromdichten beim Fräsen.

---

<sup>7</sup>Ein direkter Vergleich zwischen SAND- und NAND-Ansatz erfolgt für das Bohren in Kapitel 5.3.

<sup>8</sup>Davon werden viele Optimierungsaufgaben in unter einer Stunde gelöst.

# Kapitel 5

## Parameteridentifikation von Bohrprozessen

### Inhalt

---

5.1	Wärmeleitungsmodell des Bohrens . . . . .	79
5.1.1	SAND-Formulierung . . . . .	81
5.1.2	NAND-Formulierung . . . . .	83
5.2	Randbedingungen, Diskretisierung und Temperaturdaten . . . . .	84
5.3	Numerische Ergebnisse . . . . .	88
5.3.1	Validierung . . . . .	89
5.3.2	Rechenzeitvergleich . . . . .	95
5.3.3	Identifikation von Realdaten . . . . .	100
5.4	Fazit der Parameteridentifikation des Bohrens . . . . .	112

---

Zur Bestimmung des Wärmeeintrags beim Fräsen wurde bereits die Methodik der Parameteridentifikation eingesetzt. In diesem Kapitel wird die Parameteridentifikation auf das Bohren angewendet. Die numerischen Auswertungen umfassen drei Analysen. Zunächst wird eine Validierung der Ansätze SAND und NAND durchgeführt, indem künstliche Daten identifiziert werden. Anschließend erfolgt ein empirischer Vergleich der Berechnungszeiten zwischen den beiden Ansätzen. Zuletzt wird die Bestimmung der Wärmestromdichten aus experimentellen Daten von Bohrprozessen durchgeführt.

### 5.1 Wärmeleitungsmodell des Bohrens

Das mathematische Modell zur Beschreibung des Wärmeeintrags und des Wärmeflusses im Bauteil ist durch die Wärmeleitungsgleichung gegeben. In Kapitel 4.2 wurde für die numerische Diskretisierung bereits die schwache Form (4.7) hergeleitet. Ausgehend

von dieser Gleichung wird das Wärmeleitungsmodell für das Bohren beschrieben. Um die nachfolgenden Betrachtungen übersichtlicher zu halten und die Implementierung zu erleichtern, wird für die Zeitdiskretisierung das allgemeine Theta-Verfahren mit  $\theta = 0$  spezifiziert:

$$\begin{aligned} & \frac{\rho}{\Delta t} \int_{\Omega} c_p(T_{\phi}^l) T_{\phi}^l v_{\phi} dx - \frac{\rho}{\Delta t} \int_{\Omega} c_p(T_{\phi}^l) T_{\phi}^{l-1} v_{\phi} dx \\ &= - \int_{\Omega} k(T_{\phi}^l) \nabla T_{\phi}^l \cdot \nabla v_{\phi} dx + \int_{\Gamma} \dot{q}_{\phi}^l v_{\phi} ds, \quad l = 1, \dots, n_t \end{aligned} \quad (5.1)$$

Dies entspricht einer impliziten Euler-Diskretisierung. Die implizite Variante hat sich im Vergleich zum expliziten Verfahren bei der Lösung der Parameteridentifikation des Fräsens als stabiler und genauer erwiesen. Zudem war es effizienter als das Trapez-Verfahren.

Die Modellierung erfolgt im Gegensatz zum Fräsen mit einem nichtlinearen Wärmeleitungsmodell. Dafür werden die Materialparameter  $c_p$  und  $k$  linear abhängig von der lokalen Temperatur  $T_{\phi}^l$  gewählt:

$$\begin{aligned} c_p(T_{\phi}^l) &= a_{c_p} T_{\phi}^l + b_{c_p} \\ k(T_{\phi}^l) &= a_k T_{\phi}^l + b_k, \quad l = 1, \dots, n_t \end{aligned} \quad (5.2)$$

Dies lässt einerseits eine genauere Berechnung zu, andererseits ist dadurch der empirische Vergleich der Berechnungszeiten zwischen dem SAND- und dem NAND-Ansatz besonders interessant, da wie in Kapitel 2.1 angeführt mögliche Rechenvorteile des SAND-Ansatzes entstehen können. Mit der Abhängigkeit (5.2) ergibt sich (5.1) zu

$$\begin{aligned} & \frac{\rho}{\Delta t} \int_{\Omega} (a_{c_p} T_{\phi}^l + b_{c_p}) T_{\phi}^l v_{\phi} dx - \frac{\rho}{\Delta t} \int_{\Omega} (a_{c_p} T_{\phi}^l + b_{c_p}) T_{\phi}^{l-1} v_{\phi} dx \\ &= - \int_{\Omega} (a_k T_{\phi}^l + b_k) \nabla T_{\phi}^l \cdot \nabla v_{\phi} dx + \int_{\Gamma} \dot{q}_{\phi}^l v_{\phi} ds, \quad l = 1, \dots, n_t. \end{aligned} \quad (5.3)$$

Im Gegensatz zum Fräsen ist beim Bohren die Änderung der Randflächen  $\Gamma_{+}^l$ , die einen positiven Wärmeeintrag erfahren, von einem Zeitschritt auf den Nächsten vergleichsweise klein. Eine Gitteranpassung für jeden Zeitschritt verspricht hier keine große Einsparung an Gitterpunkten bzw. diskreten Zuständen. Im Gegenzug würde die nötige Transformation der Zustände vom Gitter des vorherigen Zeitpunkts auf den folgenden die Systemmatrizen dichter besetzen. Aus diesem Grund wird beim Bohren auf eine Gitteranpassung verzichtet und für jeden Zeitschritt die gleiche Diskretisierung gewählt.

Durch die Diskretisierung (siehe Kapitel 2.3) entstehen für  $l = 1, \dots, n_t$  Gleichungssysteme, welche quadratisch bezüglich der Temperatur und linear bezüglich der Wärme-



stromdichte sind. Mit (4.8) können diese durch

$$\begin{aligned}
 & \left[ \mathbf{A}^1 T^l | \mathbf{A}^2 T^l | \dots \right] T^l + \mathbf{A}^0 T^l - \left[ \mathbf{B}^1 T^l | \mathbf{B}^2 T^l | \dots \right] T^{l-1} - \mathbf{B}^0 T^{l-1} - \mathbf{C}^l \dot{q}_{\text{ident}} = 0 \\
 & T^0 = T_0 \\
 \Leftrightarrow & \sum_{k=1}^{n_s^1} \mathbf{A}^k T^l T_k^l + \mathbf{A}^0 T^l - \sum_{k=1}^{n_s^1} \mathbf{B}^k T^l T_k^{l-1} - \mathbf{B}^0 T^{l-1} - \mathbf{C}^l \dot{q}_{\text{ident}} = 0 \\
 & T^0 = T_0 \\
 \Leftrightarrow & H^l(T^l, T^{l-1}, \dot{q}_{\text{ident}}) = 0 \\
 & T^0 = T_0 \\
 \Leftrightarrow & H(T, \dot{q}_{\text{ident}}) = 0
 \end{aligned} \tag{5.4}$$

dargestellt werden. Hierbei sind  $\mathbf{A}^k \in \mathbb{R}^{n_s^1 \times n_s^1}$ ,  $\mathbf{B}^k \in \mathbb{R}^{n_s^1 \times n_s^1}$ ,  $k = 1, \dots, n_s^1$  und  $\mathbf{C}^l \in \mathbb{R}^{n_s^1 \times n_z}$ ,  $l = 1, \dots, n_z$  Systemmatrizen die von der Wahl der finiten Elemente abhängen. Sie werden aufgrund ihrer Komplexität und Größe nicht weiter spezifiziert. Da Gitter und Diskretisierung zeitinvariant sind, sind auch die Systemmatrizen nicht vom Zeitpunkt abhängig. Es gilt  $n_s^1 = n_s^2 = \dots = n_s^{n_t}$ .

### 5.1.1 SAND-Formulierung

Das Gleichungssystem (5.4) dient als Modell für die Parameteridentifikation. Beim SAND-Ansatz wird es in Form von Gleichungsnebenbedingungen der Optimierung hinzugefügt. Zusätzlich werden Boxschränken mit  $T_L = 20 \text{ °C}$ ,  $T_U = \infty$  sowie  $\dot{q}_L = 0$  und  $\dot{q}_U = \infty$  berücksichtigt. Ausgehend von (4.9) lautet die SAND-Formulierung des Parameteridentifikationsproblems mit der Definition von  $H$  aus (5.4):

$$\begin{aligned}
 & \min_{X \in \mathbb{R}^{n_x}} \sum_{i=1}^{n_s} \bar{w}_i (T_i - \vartheta_i(T_{\text{mess}}))^2 \\
 & \text{unter } \sum_{k=1}^{n_s^1} \mathbf{A}^k T^l T_k^l + \mathbf{A}^0 T^l - \sum_{k=1}^{n_s^1} \mathbf{B}^k T^l T_k^{l-1} - \mathbf{B}^0 T^{l-1} - \mathbf{C}^l \dot{q}_{\text{ident}} = 0, \quad l = 1, \dots, n_t \\
 & T^0 = T_0 \\
 & T_L \leq T \leq T_U \\
 & \dot{q}_L \leq \dot{q}_{\text{ident}} \leq \dot{q}_U
 \end{aligned} \tag{5.5}$$

Hierbei ist  $m_h = n_s$  und  $m_g = 0$ , da die Beschränkungen von  $T$  und  $\dot{q}_{\text{ident}}$  Boxschränken sind und vom Algorithmus separat behandelt werden. Für die Optimierungsvariable ergibt sich  $X = (T^T, \dot{q}_{\text{ident}}^T)^T \in \mathbb{R}^{n_x}$ .

Auch beim Bohren sollen die algorithmischen Vorteile durch das Angeben analytischer Ableitungen ausgenutzt werden. Der Gradient der Zielfunktion von (5.5) hat die Form

$$\nabla F(X) = \begin{pmatrix} 2\bar{w}_1(T_1 - \vartheta_1(T_{\text{mess}})) \\ 2\bar{w}_2(T_2 - \vartheta_2(T_{\text{mess}})) \\ \vdots \\ 2\bar{w}_{n_s}(T_{n_s} - \vartheta_{n_s}(T_{\text{mess}})) \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^{n_x}.$$

Die Jacobi-Matrix von Gleichungssystem (5.4) ist gegeben durch

$$\nabla H(X) = \begin{pmatrix} \nabla H^1(X) \\ \nabla H^2(X) \\ \vdots \\ \nabla H^{n_t}(X) \end{pmatrix} = \begin{pmatrix} \tilde{\mathbf{A}}^1 & & & -\mathbf{C}^1 \\ \tilde{\mathbf{B}}^2 & \tilde{\mathbf{A}}^2 & & -\mathbf{C}^2 \\ & \ddots & \ddots & \vdots \\ & & \tilde{\mathbf{B}}^{n_t} & \tilde{\mathbf{A}}^{n_t} & -\mathbf{C}^{n_t} \end{pmatrix} \in \mathbb{R}^{n_s \times n_x}, \quad (5.6)$$

mit  $\tilde{\mathbf{A}}^l = \sum_{k=1}^{n_s} \mathbf{A}^k T_k^l + [\mathbf{A}^1 T^l | \mathbf{A}^2 T^l | \dots] + \mathbf{A}^0 - \sum_{k=1}^{n_s} \mathbf{B}^k T_k^{l-1}$   
 und  $\tilde{\mathbf{B}}^l = -[\mathbf{B}^1 T^l | \mathbf{B}^2 T^l | \dots] - \mathbf{B}^0.$

Die Hesse-Matrix der Lagrange-Funktion

$$\nabla_X^2 L(X, \lambda) = \nabla^2 F(X) + \sum_{k=1}^{n_s} \lambda_k \nabla^2 H_k(X)$$

hat die Form

$$\nabla_X^2 L(X, \lambda) = \begin{pmatrix} \nabla_T^2 F(X) + \sum_{k=1}^{n_s} \lambda_k \nabla_T^2 H_k(X) & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{n_x \times n_x}$$

und ist damit nur von  $T$  abhängig, da  $F$  unabhängig und  $H$  linear abhängig von  $\dot{q}_{\text{ident}}$  ist. Die Hesse-Matrix der Zielfunktion bzgl.  $T$  ist

$$\nabla_T^2 F(X) = \begin{pmatrix} 2\bar{w}_1 & & & \\ & 2\bar{w}_2 & & \\ & & \ddots & \\ & & & 2\bar{w}_{n_s} \end{pmatrix} \in \mathbb{R}^{n_s \times n_s}$$

und die Summe der Hessematrizen des Gleichungssystems bzgl.  $T$  ist durch

$$\sum_{k=1}^{n_s} \lambda_k \nabla_T^2 H_k(X) = \begin{pmatrix} \hat{\mathbf{A}} & \hat{\mathbf{B}}^T & & & \\ \hat{\mathbf{B}} & \hat{\mathbf{A}} & \cdots & & \\ & & \cdots & \cdots & \hat{\mathbf{B}}^T \\ & & & \hat{\mathbf{B}} & \hat{\mathbf{A}} \end{pmatrix} \in \mathbb{R}^{n_s \times n_s},$$

$$\text{mit} \quad \hat{\mathbf{A}}_{ij} = \sum_{k=1}^{n_s^1} \lambda_k \mathbf{A}_{kj}^i + \sum_{k=1}^{n_s^1} \lambda_k \mathbf{A}_{ki}^j, \quad i, j = 1, \dots, n_s^1$$

$$\text{und} \quad \hat{\mathbf{B}}_{ij} = - \sum_{k=1}^{n_s^1} \lambda_k \mathbf{B}_{ki}^j, \quad i, j = 1, \dots, n_s^1$$

gegeben.

### 5.1.2 NAND-Formulierung

Die in diesem Kapitel verwendete NAND-Formulierung des Parameteridentifikationsproblems hat die Form

$$\begin{aligned} \min_{\dot{q}_{\text{ident}} \in \mathbb{R}^{n_s}} & \sum_{i=1}^{n_s} \bar{w}_i (T_i(\dot{q}_{\text{ident}}) - \vartheta_i(T_{\text{mess}}))^2 \\ \text{unter} & \quad T_L \leq T(\dot{q}_{\text{ident}}) \leq T_U \\ & \quad \dot{q}_L \leq \dot{q}_{\text{ident}} \leq \dot{q}_U, \end{aligned} \quad (5.7)$$

wobei die Temperaturen  $T$  in Abhängigkeit von  $\dot{q}_{\text{ident}}$  durch das Lösen von (5.4) außerhalb der Optimierung bestimmt werden. Die Beschränkungen sind durch  $T_L = 20^\circ\text{C}$ ,  $T_U = \infty$  sowie  $\dot{q}_L = 0$  und  $\dot{q}_U = \infty$  gegeben. Die Anzahl der Nebenbedingungen<sup>1</sup> ist  $m_h = 0$  und  $m_g = 2 n_s$ . Die Beschränkungen von  $T$  sind durch das Lösen von (5.4) nichtlinear. Die Beschränkungen von  $\dot{q}_{\text{ident}}$  sind Boxschränken und werden separat gehandhabt.

An dieser Stelle wird ein entscheidender Unterschied zur SAND-Formulierung deutlich. Während die Temperaturbeschränkungen im SAND-Ansatz einfach zu handhabende Boxschränken sind, sind sie hier hochgradig nichtlineare Nebenbedingungen. Dies liegt an der Abhängigkeit der Temperaturen von den Wärmestromdichten über den Lösungsoperator von (5.4). Gleiches gilt für die Abhängigkeit der Zielfunktion von  $\dot{q}_{\text{ident}}$ , was die Berechnung von Ableitungen erschwert. Bei der Optimierung wird auf numerische Ableitungen zurückgegriffen, die vom Algorithmus bereitgestellt werden. Auch die Struktur dieser Formulierung des Optimierungsproblems ist unterschiedlich. Es hat im Gegensatz zum SAND-Ansatz wenig Optimierungsvariablen und ist dicht besetzt.

<sup>1</sup>Durch Umformen einer Ungleichungsnebenbedingung mit beidseitigen Schranken in die Form (2.5) verdoppelt sich die Anzahl.

## 5.2 Randbedingungen, Diskretisierung und Temperaturdaten

Die Geometrie des modellierten Bauteils beim Bohren ist der von den Experimenten aus Kapitel 3.4.3 nachempfunden. Der Bohrkanal hat zum Zeitpunkt der Thermografieaufnahme eine Tiefe von 35 mm. Der Durchmesser entspricht dem des genutzten Bohrers und beträgt 8 mm oder 12 mm. Der Abstand des Lochs zur vorderen Außenwand des Bauteils, auf dem die Thermografiebilder aufgenommen wurden, ist bei den einzelnen Versuchen unterschiedlich und wurde jeweils gemessen. Er variiert zwischen 0,35 mm und 0,88 mm. Das modellierte Bauteil umfasst einen Ausschnitt der echten Bauteile aus den Versuchen und hat insgesamt die Maße von  $25 \times 47 \times 42 \text{ mm}^3$ , was einen ausreichenden Wärmefluss in alle Koordinatenrichtungen zulässt. Zudem deckt die vordere Ebene ( $z = 0 \text{ mm}$ ), in der der Temperaturabgleich mit den Daten erfolgt, einen Großteil des Thermografiebildes ab.

Wie bereits in Kapitel 5.1 erläutert, findet beim Bohren keine Adaption des Gitters zwischen den Zeitschritten statt. Vielmehr ergibt sich ein Gitter für alle Zeitschritte. Die Feinheit des Gitters wird dabei lokal variiert. Berücksichtigt wird insbesondere das Innere des Bohrkanals, in dem aufgrund des Wärmeeintrags hohe Temperaturgradienten zu erwarten sind, und die Bauteilfront, auf der der Temperaturabgleich erfolgt. Die Diskretisierung der FEM erfolgt mit linearen Lagrange-Elementen auf Hexaedern. Die diskreten Temperaturwerte in den Eckpunkten der Zellen entsprechen so den Freiheitsgraden des Gleichungssystems. Die kleinsten Zellen haben einen Durchmesser von etwa 1 mm. Bei der Diskretisierung des Bauteils lässt sich eine Spiegelsymmetrie des Prozesses in der  $y$ - $z$ -Ebene ausnutzen, was in etwa die Hälfte an Gitterpunkten bzw. diskreten Zuständen einspart. Dennoch ist das resultierende Gitter beim Bohren im Vergleich zum Fräsen gröber. Das ist zum einen der komplexeren Geometrie geschuldet, die durch das Bohrloch entsteht, zum anderen ist die Prozessdauer beim Bohren länger. Deshalb ist es hinsichtlich der angestrebten Optimierung praktikabel, mehrere Schnitte der Bohrer-schneide in einem diskreten Zeitschritt zusammenzufassen. Um in der feinsten Diskretisierung etwa 100 000 diskrete Zustände zu erhalten, wird eine Anzahl von bis zu  $n_t = 24$  Zeitschritten gewählt. Dies hat zur Folge, dass ein Zeitschritt etwa 7 bis 22 Schnitte zusammenfasst. Die genaue Zahl ist abhängig vom Zahnvorschub  $f_z$ , welcher wiederum durch die übrigen Maschinenparameter bestimmt wird. Die zeitliche Diskretisierung erscheint zunächst vergleichsweise grob. Jedoch konnten Langenhorst et al. [69] für das Fräsen zeigen, dass eine solche Modellreduktion die Rechenzeit bei ähnlicher Modellgüte verringern kann. Auch innerhalb dieser Arbeit wird sich später zeigen, dass ein gutes Gleichgewicht zwischen der Güte der Identifikationsresultate und der Rechenzeit der Optimierung entsteht. Insgesamt wird damit abhängig von der Vorschubgeschwindigkeit eine Zeitspanne von  $t_f - t_0 = 3,5 \text{ s}$  bzw.  $5,25 \text{ s}$  abgedeckt. Die Anzahl der diskreten Temperaturen aller Zeitschritte beträgt bis zu  $n_s = 112\,896$ . In Abb. 5.1 ist beispielhaft ein diskretisiertes Bauteil zu sehen.

Die Modellierung des Wärmeeintrags beim Bohren ist der des Fräsens ähnlich und erfolgt

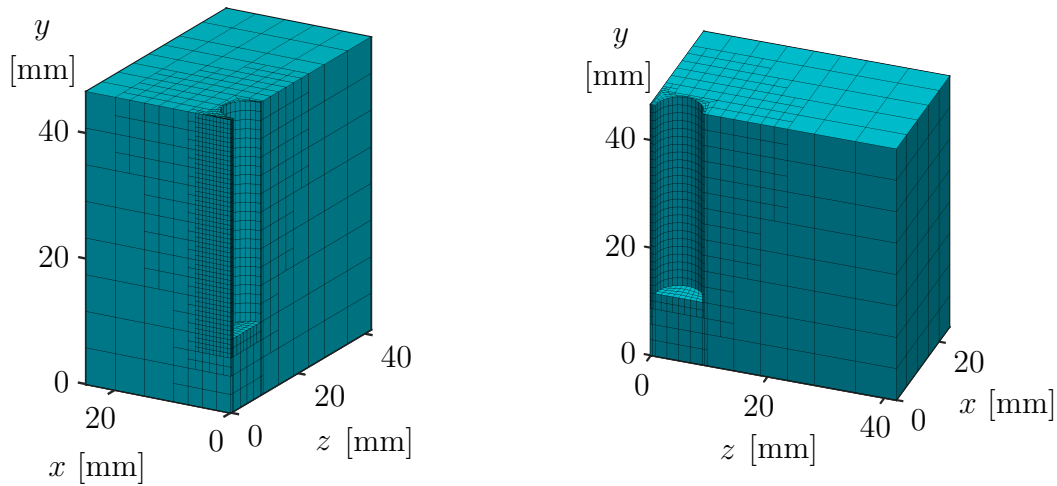


Abbildung 5.1: Gitter der FEM-Diskretisierung

makroskopisch durch eine bewegte Wärmequelle. Das Bauteil wird mit vorhandenem Bohrloch diskretisiert, d. h. es wird kein Materialabtrag simuliert, sondern die Wärmestromdichte direkt auf den Bohrlochrand aufgetragen. Wie in Abb. 5.2 gezeigt, gliedert sich die bewegte Wärmequelle in  $n_z$  Bereiche. Der erste Bereich umfasst eine primäre Wärmequelle  $\dot{q}_{ident,1}$ , welche den direkten Wärmeeintrag durch die Bohrerschneide modelliert. Diese wandert während der Simulation mit der Vorschubgeschwindigkeit analog zur Bohrerbewegung das Bohrloch nach unten. Ihr schließen sich sekundäre Wärmequellen  $\dot{q}_{ident,i}$ ,  $i = 2, \dots, n_z$  an, die sowohl die Reibung des Bohrers als auch einen möglichen Wärmeaustausch der durch die Wendeln abtransportierten Späne mit der Innenwand

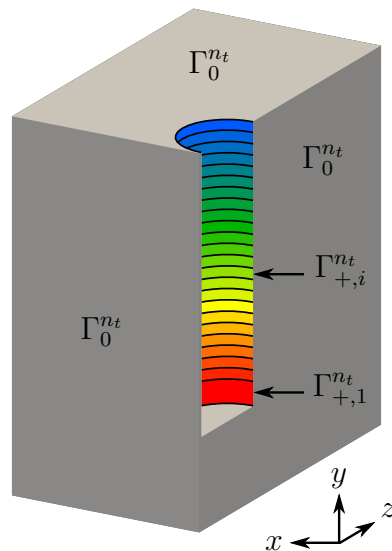


Abbildung 5.2: Randbereiche der Wärmequellen und der adiabaten Randwerte beim Bohren zum Endzeitpunkt

modellieren. Die primäre Wärmequelle wird durch einen Ring im Innern des Bohrlochs dargestellt. Die Höhe des Rings  $f_{\text{prim}}$  ergibt sich aus der Bohrlochtiefe von 35 mm und der Anzahl an Zeitschritten:

$$f_{\text{prim}} = \frac{35 \text{ mm}}{n_t} \quad (5.8)$$

Dies entspricht der Anzahl der Schnitte eines Zeitschritts multipliziert mit dem Zahnvorschub  $f_z$ . Die sekundären Wärmequellen bestehen aus Ringen, die bis an den oberen Rand des Werkstücks reichen (siehe Abb. 5.2). Dies ermöglicht die Identifikation einer nach oben hin abnehmenden Wärmeintensität zur Berücksichtigung des Abkühlens der Späne. Die Höhe der Ringe der sekundären Wärmequellen ist jeweils

$$f_{\text{sek}} = \frac{35 \text{ mm} - f_{\text{prim}}}{n_z - 1}. \quad (5.9)$$

Die Bereiche unterhalb der primären Wärmequelle sowie auch alle äußeren Flächen des Werkstücks erhalten in der Simulation adiabate Randbedingungen. Es wird angenommen, dass in jedem Zeitschritt der gleiche Wärmeeintrag erfolgt, also die Wärmestromdichten der einzelnen Bereiche über die Zeit konstant sind. Mathematisch werden die Randwerte wie folgt zusammengefasst:

$$\begin{aligned} \dot{q}^l|_{\Gamma_{+,1}^l} &= \dot{q}_{\text{ident},1} \\ \dot{q}^l|_{\Gamma_{+,i}^l} &= \dot{q}_{\text{ident},i}, \quad i = 2, \dots, n_z \\ \dot{q}^l|_{\Gamma_0^l} &= 0, \quad l = 0, \dots, n_t \end{aligned} \quad (5.10)$$

In Abb. 5.3 ist die Temperaturverteilung zum Endzeitpunkt einer Simulation zu sehen. Es wurden dabei fiktive Werte für die Wärmequellen angenommen. Zu erkennen ist der erhöhte Wärmeeintrag der primären Wärmequelle und die nach oben abfallende Intensität. Dadurch bildet sich in der vorderen Ebene ( $z = 0$ ) eine Art Flammenprofil der Isothermen aus. Die beiden Materialparameter spezifische Wärmekapazität und Wärmeleitfähigkeit sind gemäß (5.2) temperaturabhängig. Die Dichte ist in guter Näherung im relevanten Temperaturbereich konstant. Um die Parameter  $a_{c_p}$ ,  $b_{c_p}$ ,  $a_k$  und  $b_k$  zu bestimmen werden die Werte aus Tab. 5.1 verwendet und eine Ausgleichsgerade bestimmt. Die resultierenden Werte sind

$$\begin{aligned} a_{c_p} &\approx 0,3988, \quad b_{c_p} \approx 454,1311 \\ a_k &\approx -0,0057, \quad b_k \approx 45,3787. \end{aligned}$$

Die für die Identifikation nötigen Thermografieaufnahmen stammen aus der Versuchreihe, welche in Kapitel 3.4.3 beschrieben wurde. Abbildung 5.4 zeigt exemplarisch die Aufnahme eines Experiments. Obwohl die Daten auch Temperaturen bis zu 200 °C beinhalten, werden hier nur Temperaturen bis 70 °C dargestellt. Dies dient der Hervorhebung

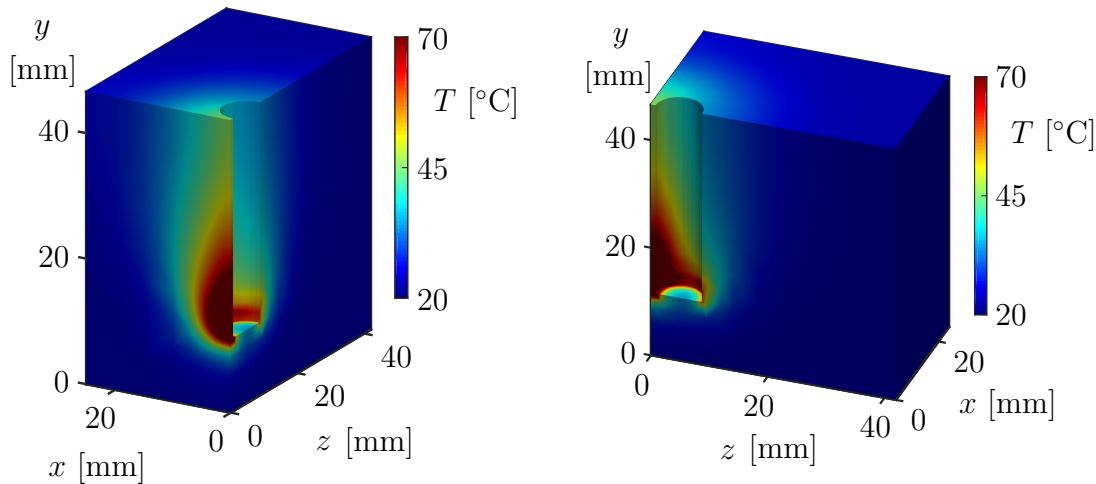


Abbildung 5.3: Temperaturverteilung nach einer Simulation des Wärmeeintrags beim Bohren mit fiktiven Wärmestromdichten

Parameter	Symbol	Werte			Einheit
		20	100	200	
spezifische Wärmekapazität	$c_p$	461	496	533	J/(kg K)
Wärmeleitfähigkeit	$k$	45,1	45,1	44,1	W/(m K)
Dichte	$\rho$	7830			kg/m <sup>3</sup>

Tabelle 5.1: Temperaturabhängige Materialparameter von 42CrMo4 für die Wärmeleitungsgleichung des Bohrens (Quelle: [103])

der niedrigen Temperaturbereiche. So wird das im Experiment verwendete Schutzrohr erkennbar, mit dem die Kamera vor Spänen geschützt wurde. Auch werden leichte Störungen sichtbar. Mit einem weißen Rand ist der Bereich der Daten gekennzeichnet, der für die Identifikation verwendet wird. Dabei werden nur Temperaturen berücksichtigt, die das Bauteil und nicht das Schutzrohr zeigen und die das modellierte Bauteil überdecken. Der Bereich unterhalb der gestrichelten weißen Linie ist jener Bereich, der für die Validierung verwendet wurde (siehe Kapitel 5.3.1). Die markierten Daten werden mit den modellierten Temperaturen zum Endzeitpunkt der Simulation in der  $x$ - $y$ -Ebene mit  $z = 0$  verglichen. In der Zielfunktion von (5.5) bzw. (5.7) werden jene Gewichte  $w_i$ ,  $i = 1, \dots, n_s$  (siehe (4.2)) gleich 1 gesetzt, deren Zustände diesen Einschränkungen entsprechen. Die übrigen Temperaturen werden mit 0 gewichtet.

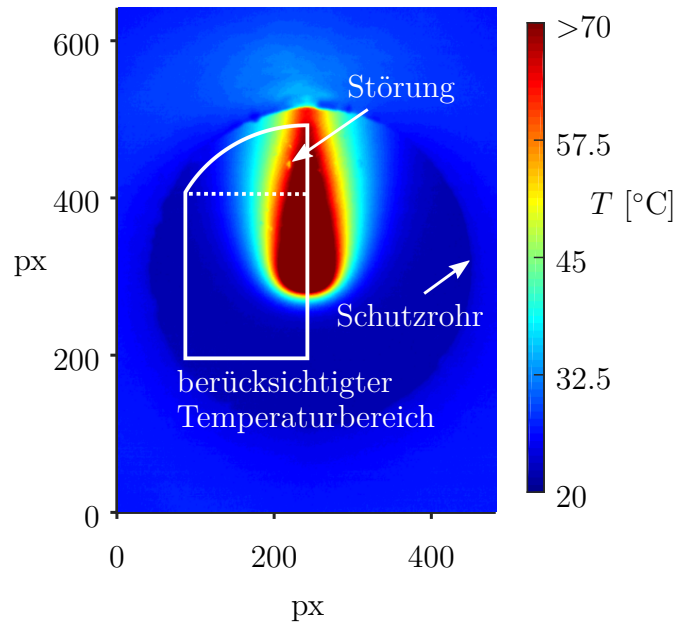


Abbildung 5.4: Thermografieaufnahme eines Bohrexperiments

### 5.3 Numerische Ergebnisse

Die Berechnungen der Parameteridentifikation des Bohrens wurden auf einem leistungsstarken Rechner mit Ubuntu 16.04 durchgeführt. Die Hardware ist in Tab. 5.2 zu sehen. Die numerischen Auswertungen teilen sich in drei Bereiche: eine Validierung der Ansätze SAND und NAND, ein Rechenzeitvergleich beider Ansätze sowie die Identifikation von Realdaten. Die Validierung der Parameteridentifikation des Bohrens wird an künstlich erzeugten sowie zusätzlich an verrauschten Daten validiert. Dabei wird die Validierung mit verrauschten Daten wie beim Fräsen durch eine Monte-Carlo-Simulation durchgeführt. Für den Rechenzeitvergleich der Ansätze SAND und NAND werden die Ergebnisse aus der Validierung um eine Vielzahl weiterer Rechnungen ergänzt. Dabei wird nicht nur das Optimierungsproblem variiert, sondern auch die Datengrundlage, indem sowohl künstliche und verrauschte Daten als auch Realdaten Verwendung finden.

Hardware	Spezifikation
CPU	2x Intel Xeon E5-2637 v3 @3.50 GHz (insg. 16 logische Kerne)
RAM	24x 32GB Samsung DDR4-2133 CL15 (insg. 768GB RAM)
HDD	2x 128GB SSDs + 4x 1TB HDDs
GPU	NVidia GeForce GT 610

Tabelle 5.2: Hardware des Rechnerservers



Die Rechenzeiten der verschiedenen Identifikationsaufgaben ergeben eine breite Grundlage zur Beurteilung der Performanz der Ansätze. Der dritte Teil umfasst die Identifikation von Realdaten. Mit den Erkenntnissen aus dem Rechenzeitvergleich wird dafür die performanteste Methodik ausgewählt und auf die Thermografiebilder von Experimenten angewandt.

### 5.3.1 Validierung

Beim Bohren wird, ähnlich dem Vorgehen beim Fräsen, durch die Identifikation von künstlichen und verrauschten Daten die Funktionalität der Methodik und der Implementierung beider Optimierungsansätze erprobt. Die künstlichen Daten werden durch Lösen des Gleichungssystems (5.4) erzeugt. Die Randwerte werden dabei mit

$$\begin{aligned} \dot{q}^l|_{\Gamma_{+,1}^l} &= \dot{q}_{\text{valid},1} \\ \dot{q}^l|_{\Gamma_{+,i}^l} &= \dot{q}_{\text{valid},i}, \quad i = 2, \dots, n_z, \quad l = 1, \dots, n_t \end{aligned}$$

fixiert. Die Wärmestromdichten der Validierung  $\dot{q}_{\text{valid},i}$ ,  $i = 1, \dots, n_z$  gilt es anschließend durch die Parameteridentifikation zu rekonstruieren. Die Simulation zur Erzeugung der Daten bildet einen Bohrprozess mit  $v_c = 55$  m/min und  $v_f = 400$  mm/min ab, woraus ein Zahnvorschub von  $f_z \approx 0,09$  mm resultiert. Der Radius des Bohrers beträgt  $r = 4$  mm und die Wandstärke zwischen Loch und vorderer Messebene 0,86 mm. Die Tiefe des Lochs zum Zeitpunkt des Temperaturabgleichs ist 35 mm. Damit wird eine Zeitspanne von  $t_f - t_0 = 5,25$  s benötigt um 383 Schnitte abzudecken. Die Zeit wird in  $n_t = 12$  diskrete Schritte unterteilt. Der Ausschnitt des Bauteils hat eine Größe von  $\Omega = [0; 25] \times [0; 47] \times [0; 42]$  mm<sup>3</sup>. Die simulierten Anfangstemperaturen  $T_0$  des Bauteils betragen bei der Validierung konstant 22 °C.

Der Bereich des Wärmeeintrags wird in  $n_z = 6$  Bereiche eingeteilt, woraus sich sechs Parameter ergeben, die es zu identifizieren gilt. Die räumliche Diskretisierung in Kombination mit der zeitlichen Diskretisierung führt auf  $n_s = 56712$  Zustände. Der Teil der künstlichen Daten, der für die Validierung in der Zielfunktion verwendet wird, wird an den Realdatenfall angepasst. Verglichen werden die Temperaturzustände des Modells mit den künstlichen Temperaturdaten zum Endzeitpunkt  $t_f$  in der  $x$ - $y$ -Ebene mit  $z = 0$  mm. Dabei wird nur der Ausschnitt an Daten berücksichtigt, welcher in Abb. 5.4 innerhalb des weißen Rahmens zu sehen ist und sich unter der weiß gestrichelten Linie befindet. Dies ist dem Vorgehen beim Fräsen nachempfunden und soll verhindern, dass Störungen durch das Schutzrohr in die Daten aufgenommen werden. Ein Ausschluss von hohen Temperaturdaten wie beim Fräsen ist beim Bohren nicht notwendig, da durch den günstigen Versuchsaufbau keine Schneiden oder andere Störfaktoren in das Bild hineinwirken. Auch der abgebildete Temperaturbereich ist deutlich größer als der beim Fräsen.

Betrachtet werden für die Validierung zwei verschiedene Szenarien. Im ersten Szenario werden ungestörte Simulationsdaten verwendet, um die fixierten Wärmestromdichten zu rekonstruieren. In einem zweiten Szenario werden die Daten vor der Identifikation durch ein Rauschen gestört. Dies erfolgt für jeden Datenpunkt durch Addition eines Zufalls-werts, der Normalverteilung. Die Verteilung hat den Erwartungswert  $\mu_N = 0$ . Die Standardabweichung  $\sigma_N$  beträgt 10 % des zu störenden Temperaturwerts. Um den Einfluss der Zufallszahlen zu reduzieren, wird dieser Versuch in einer Monte-Carlo-Simulation zehnmal wiederholt. Dies ist ein Kompromiss zwischen Berechnungsdauer, die bei einem Einzelversuch bereits mehrere Stunden beträgt, und der Reduzierung des Zufalls. Variablen des  $j$ -ten Versuchs der Monte-Carlo-Simulation werden mit  $(\cdot)_{[j]}$  gekennzeichnet.

Für das numerische Lösen wurden die Parameteridentifikationsprobleme (5.5) für den SAND-Ansatz und (5.7) für den NAND-Ansatz mit den beschriebenen Konfigurationen in WORHP implementiert. In vorangegangenen Tests konnte empirisch gezeigt werden, dass die Lösungsgüte nicht von der Wahl der Startschätzung  $(X^{(0)}, \lambda^{(0)})$  (siehe Algorithmus 2.16) für das SQP-Verfahren in WORHP abhängt. Deshalb ist eine triviale Startschätzung für die Validierung naheliegend:

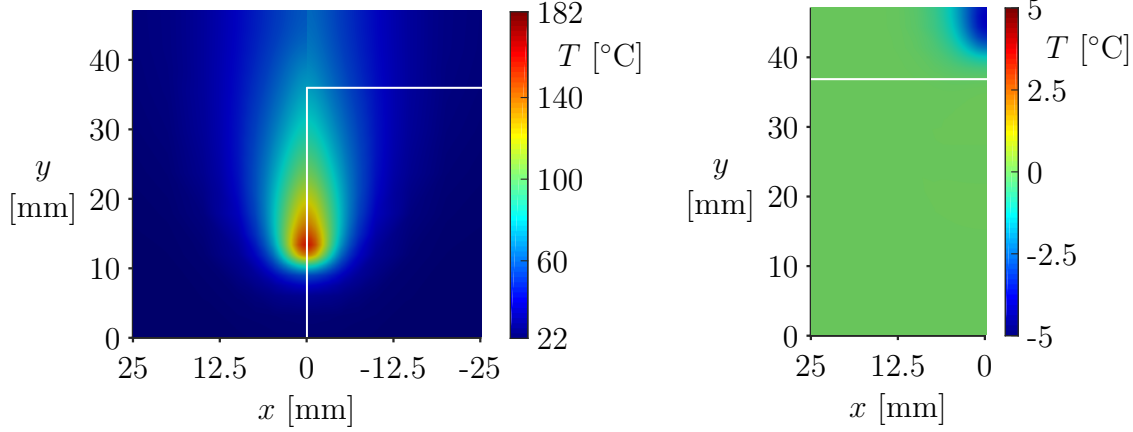
$$\begin{aligned} T_i^{(0)} &= 22 \text{ °C}, & i &= 1, \dots, n_s \\ \dot{q}_{\text{ident},i}^{(0)} &= 0 \text{ W/mm}^2, & i &= 1, \dots, n_z \\ \lambda_i^{(0)} &= 0, & i &= 1, \dots, m. \end{aligned} \tag{5.11}$$

### Erstes Szenario

Das erste Szenario umfasst die Validierung durch die Parameteridentifikation von ungestörten künstlichen Daten für den SAND- und NAND-Ansatz. Dafür wurde der Vektor der zu identifizierenden Wärmestromdichten auf

$$\dot{q}_{\text{valid}} = (2,5 \quad 0,25 \quad 0,125 \quad 0,08\bar{3} \quad 0,0625 \quad 0,05)^T \text{ W/mm}^2$$

festgelegt. Die durch Lösen der Wärmeleitungsgleichung erzeugten künstlichen Daten ( $T_{\text{mess}} = T_{\text{sim}}^{n_t}|_{z=0}$ ) wurden anschließend zunächst mit dem SAND-Ansatz rekonstruiert. Die Laufzeit dieser Parameteridentifikation beträgt für 19 Hauptiterationen 1,2h. Abbildung 5.5a zeigt einen Ausschnitt ( $t = t_f$ ,  $z = 0$  mm) der identifizierten Temperaturen sowie die Temperaturdaten. Zum besseren Vergleich werden beide Temperaturfelder in einem Koordinatensystem abgebildet, wobei die Daten an der Achse mit  $x = 0$  mm gespiegelt wurden. Die rekonstruierten Temperaturen sind im Bereich  $x > 0$  mm und die Daten im Bereich  $x < 0$  mm zu sehen. Mit einem weißen Rahmen sind jene Daten gekennzeichnet, die für die Identifikation verwendet wurden. Das Ausschließen einiger Daten ist durch die noch folgende Identifikation von Realdaten motiviert, bei der aufgrund von Störungen nicht alle gemessenen Temperaturen verwendet werden können. Ein optischer Vergleich der Temperaturfelder legt eine gelungene Rekonstruktion der



(a) links ( $x > 0$  mm):  $T^{n_t^*}|_{z=0}$ , rechts ( $x < 0$  mm):  
 $T_{\text{mess}}$  (gespiegelt)

(b)  $T^{n_t^*}|_{z=0} - T_{\text{mess}}$

Abbildung 5.5: Ergebnis der ersten Validierung mit SAND

Daten nahe. Bis auf kleine Unterschiede im oberen Bereich in der Nähe von  $x = 0$  mm sind keine Abweichungen zu erkennen. Diese Beobachtung wird durch die Abb. 5.5b gestützt. Sie zeigt die Differenz beider Temperaturfelder. Bis auf den genannten Bereich ist die Abweichung der identifizierten Temperaturen von den gemessenen auch quantitativ 0. Die größte absolute Abweichung tritt in der Ecke  $x = 0$  mm und  $y = 47$  mm auf und beträgt  $4,2^\circ\text{C}$ . Der Vektor der identifizierten Wärmestromdichten ist

$$\dot{q}_{\text{ident}}^* \approx (2,5 \quad 0,25 \quad 0,125 \quad 0,084 \quad 0,065 \quad 0,004)^T \text{ W/mm}^2.$$

Die relativen Abweichungen dieser Werte von den fixierten Werten sind

$$|\dot{q}_{\text{valid}} - \dot{q}_{\text{ident}}^*| \circ \frac{1}{\dot{q}_{\text{valid}}} \cdot 100\% \approx (0,01 \quad 0,07 \quad 0,18 \quad 0,34 \quad 3,9 \quad 91,4)^T \%.$$

Hierbei bezeichne  $\circ$  den Operator der elementweisen Multiplikation. Auffällig ist der Anstieg der Fehler mit steigender Höhe der Position der Wärmequelle. Während die Abweichung der primären Wärmequelle vernachlässigbar klein ist, steigt der Fehler bis zum vorletzten Wert auf ein geringes Niveau an. Durch die Bewegung der Wärmequelle verkürzt sich mit der Höhe der Position der Wärmequelle die Wirkungsdauer dieser im Prozess und verringert damit den Einfluss auf die Zustände und damit auf die Zielfunktion. Der letzte Wert weist dadurch eine erhebliche Abweichung auf. Dies korreliert mit der erhöhten Abweichung im oberen Bereich der Temperaturfelder. Die letzte sekundäre Wärmestromdichte  $\dot{q}_{\text{ident},6}$  wirkt erst am Ende der betrachteten Zeitspanne und nur in jenem oberen Bereich, welcher vom Temperaturvergleich ausgeschlossen wurde. Eine erfolgreiche Identifikation ist dort deshalb nicht möglich. Der Einfluss auf den Gesamtfehler ist jedoch gering. Die mittlere absolute Abweichung über alle Datenwerte aller

Zeitschritte liegt lediglich bei

$$\frac{1}{n_s} \sum_{i=1}^{n_s} |T_i^* - T_{\text{sim},i}| \approx 0,04 \text{ } ^\circ\text{C}.$$

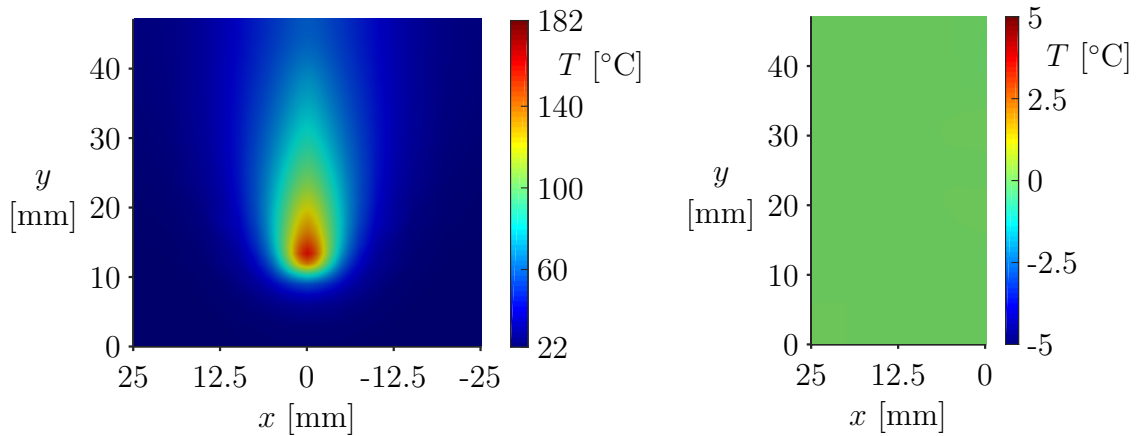
Die mittlere relative Abweichung liegt bei

$$\frac{1}{n_s} \sum_{i=1}^{n_s} \left| \frac{T_i^* - T_{\text{sim},i}}{T_{\text{sim},i}} \right| \cdot 100 \% \approx 0,07 \%.$$

Die Identifikation wurde wie bei der Validierung des Fräsens nur mit einem Bruchteil (0,6 %) der ursprünglichen Daten durchgeführt, während die erwähnten Abweichungen aber über alle Datenpunkte berechnet wurden.

Als nächstes wurde die gleiche Identifikationsaufgabe mit dem NAND-Ansatz gelöst. Der Algorithmus benötigt 2,4 h und liefert ein nahezu identisches Ergebnis zum SAND-Ansatz. Zum Vergleich sind in Abb. 5.6a Ausschnitte ( $t = t_f$ ,  $z = 0 \text{ mm}$ ) der Lösungen beider Ansätze in einem Bild zu sehen. Links ( $x > 0 \text{ mm}$ ) wurde das Temperaturfeld vom SAND-Ansatz übernommen. Rechts ( $x < 0 \text{ mm}$ ) sind die resultierenden Temperaturen des NAND-Ansatzes gespiegelt zu sehen. Optisch ist kein Unterschied auszumachen. Auch die Differenz der Temperaturfelder, abgebildet in Abb. 5.6b, zeigt quantitativ keine Unterschiede. Die maximale absolute Differenz beträgt  $0,1 \text{ } ^\circ\text{C}$ . Die identifizierten Wärmestromdichten des NAND-Ansatzes sind mit

$$\dot{q}_{\text{ident}}^* \approx (2,5 \quad 0,25 \quad 0,125 \quad 0,083 \quad 0,066 \quad 0,005)^T \text{ W/mm}^2$$



(a) Temperaturfelder  $T^{nt*}|_{z=0}$ , links ( $x > 0$ ): SAND, (b) Differenz der Temperaturfelder von rechts ( $x < 0$ ): NAND (gespiegelt) SAND und NAND (ungespiegelt)

Abbildung 5.6: Vergleich der Ergebnisse von SAND und NAND

sehr ähnlich zu den entsprechenden Werten des SAND-Ansatzes. Es treten die gleichen Probleme mit der Identifikation des oberen Temperaturbereichs und der letzten sekundären Wärmestromdichte auf. Auch hier hat dies jedoch nur einen untergeordneten Einfluss auf das Gesamtergebnis.

## Zweites Szenario

Beim zweiten Szenario der Validierung werden die Daten für die zehn Wiederholungen der Monte-Carlo-Simulation vor der Identifikation jeweils mit einem normalverteiltem Rauschen von 10% verfälscht. Dieses Vorgehen soll über die eigentliche Funktionalität hinaus die Robustheit der Methodik gegenüber Messfehlern testen. Abbildung 5.7 zeigt sowohl einen Ausschnitt ( $z = 0$  mm,  $t = t_f$ ) der ursprünglichen Daten der Simulation ( $x < 0$  mm) als auch, beispielhaft für den ersten Versuch der Monte-Carlo-Simulation, die gestörten Daten ( $x > 0$  mm), welche für die Identifikation verwendet wurden (weißer Rahmen). Der Vektor der zu identifizierenden Wärmestromdichten ist durch

$$\dot{q}_{\text{valid}} = (2 \quad 0,2 \quad 0,1 \quad 0,0\bar{6} \quad 0,05 \quad 0,04)^T \text{ W/mm}^2$$

gegeben.

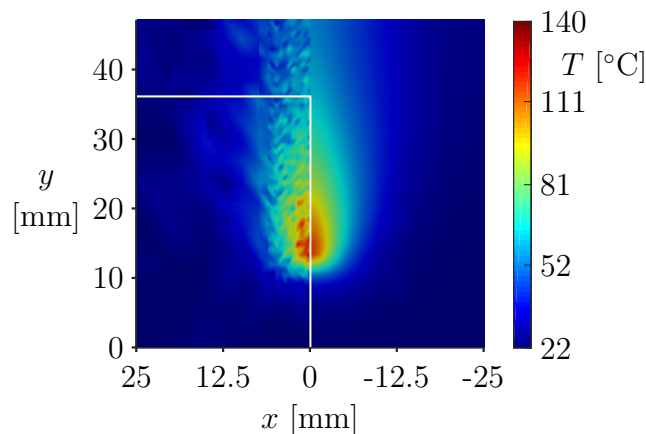
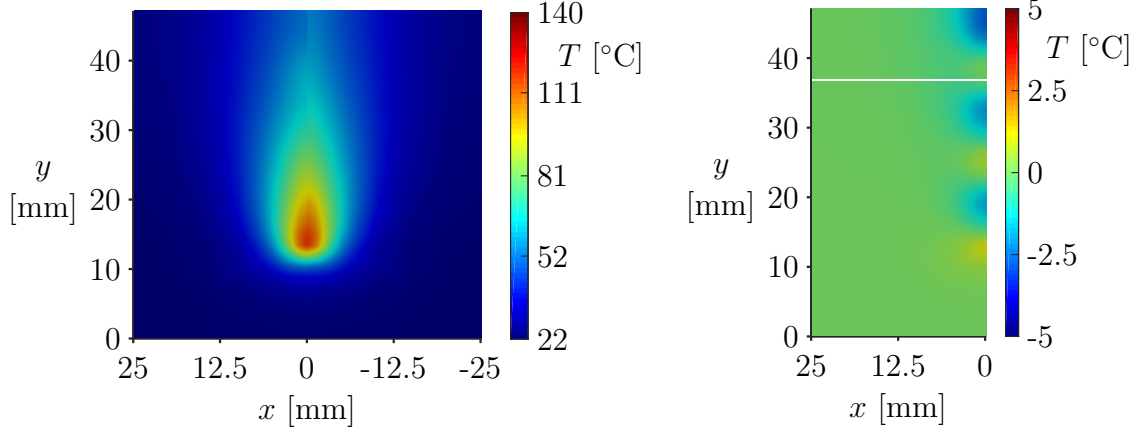


Abbildung 5.7: Temperaturdaten des zweiten Validierungstests, links ( $x > 0$  mm):  $T_{\text{mess}[1]}$ , rechts ( $x < 0$  mm):  $T_{\text{sim}}^{nt}|_{z=0}$  (gespiegelt)

Als erstes wurde der SAND-Ansatz getestet. Durchschnittlich über die zehn Versuche beträgt die Rechenzeit 1,5 h für 21 Hauptiterationen. Ein Ausschnitt des identifizierten Temperaturfeldes für den ersten Versuch ist in Abb. 5.8a zu sehen ( $x > 0$  mm). Zum Vergleich sind die ungestörten Daten im gleichen Koordinatensystem in gespiegelter Form dem Resultat gegenübergestellt ( $x < 0$  mm). Wie beim ersten Validierungstest, bei dem die Daten unverfälscht identifiziert wurden, treten auch hier im oberen Abschnitt des Bildes die größten Unterschiede auf. Ansonsten sind weitere Abweichungen optisch kaum

erkennbar. In Abb. 5.8b, welche die Differenz der gezeigten Temperaturfelder darstellt, werden kleinere Abweichungen im Bereich der primären Wärmequelle und unmittelbar darüber sichtbar. Die relative Abweichung ist hier jedoch deutlich geringer als am oberen Rand des Bildes, da die Temperaturen in der Nähe der Wärmequelle sehr hoch sind.



(a) links ( $x > 0$ ):  $T_{[1]}^{n_t^*}|_{z=0}$ , rechts ( $x < 0$ ):  $T_{sim}^{n_t}|_{z=0}$  (gespiegelt) (b)  $T_{[1]}^{n_t^*}|_{z=0} - T_{sim}^{n_t}|_{z=0}$

Abbildung 5.8: Ergebnis der zweiten Validierung mit SAND

Die maximale absolute Abweichung tritt bei allen zehn Versuchen in der oberen rechten Ecke auf und beträgt durchschnittlich  $4,8^\circ\text{C}$ . Während die Abweichungen im oberen Bildbereich erneut durch die fehlenden Daten begründet werden können, werden die Fehler nahe der Wärmequelle möglicherweise durch das Rauschen verursacht. Dennoch sind die mittlere absolute sowie relative Abweichung über alle Datenwerte aller Zeitschritte mit durchschnittlich

$$\frac{1}{10} \sum_{j=1}^{10} \left( \frac{1}{n_s} \sum_{i=1}^{n_s} |T_{i[j]}^* - T_{sim,i}| \right) \approx 0,3^\circ\text{C}$$

und

$$\frac{1}{10} \sum_{j=1}^{10} \left( \frac{1}{n_s} \sum_{i=1}^{n_s} \left| \frac{T_{i[j]}^* - T_{sim,i}}{T_{sim,i}} \right| \cdot 100\% \right) \approx 0,5\%$$

gering. Auch hier wurde die Identifikation nur mit einem Bruchteil (0,6%) der ursprünglichen Daten durchgeführt. Die erwähnten Abweichungen wurden jedoch über alle Datenpunkte berechnet. Der durchschnittlich identifizierte Vektor der Wärmestromdichten ist

$$\frac{1}{10} \sum_{j=1}^{10} \dot{q}_{ident[j]}^* \approx (1,997 \quad 0,196 \quad 0,1 \quad 0,077 \quad 0,031 \quad 0,001)^T \text{ W/mm}^2.$$

Die Abweichungen vom fixierten Vektor sind gering aber höher als im ersten Szenario. Dies war aufgrund des zusätzlichen Rauschens zu erwarten. Die Abweichung der letzten

sekundären Wärmestromdichte ist erneut aufgrund von fehlenden Daten am höchsten. Insgesamt sind die Rekonstruktionen trotz starkem Rauschen erfolgreich.

Der NAND-Ansatz benötigt durchschnittlich 3,4 h bis zur Konvergenz und kommt im Vergleich zum SAND-Ansatz ein weiteres Mal bis auf numerische Ungenauigkeiten zum gleichen Ergebnis. Optisch sind die Temperaturfelder der zehn Versuche deckungsgleich, weshalb auf eine Visualisierung an dieser Stelle verzichtet wird. Die identifizierten Wärmestromdichten betragen durchschnittlich

$$\frac{1}{10} \sum_{j=1}^{10} \dot{q}_{\text{ident}[j]}^* \approx (1,998 \quad 0,196 \quad 0,1 \quad 0,077 \quad 0,031 \quad 0,001)^T \text{ W/mm}^2$$

und sind nahezu identisch zu den Werten des SAND-Ansatzes. Auch die numerischen Werte, welche die Abweichungen quantifizieren, sind annähernd gleich. Es wird deshalb davon ausgegangen, dass beide Ansätze hinsichtlich der Resultate dieser Art von Optimierungsproblemen qualitativ nicht zu unterscheiden sind. Lediglich die verschiedenen Rechenzeiten lassen einen Unterschied beider Ansätze vermuten. Dies wird im folgenden Kapitel untersucht.

### 5.3.2 Rechenzeitvergleich

Die Optimierung mit PDEs hat in den letzten Jahren durch die stetige Weiterentwicklung der Rechenleistung von Computersystemen an Bedeutung gewonnen. Dennoch ist die Rechenzeit im Vergleich zur Optimierung mit gewöhnlichen Differentialgleichungen weiterhin ein entscheidender Faktor. Dies hat seinen Grund in den zusätzlichen Dimensionen, die mit einem numerischen Verfahren zu diskretisieren sind. Betrachtet man ein vollständiges, zeitvariantes Modell - wie bei der betrachteten Parameteridentifikation des Bohrens - ergeben sich drei Raumdimensionen und eine Zeitdimension. Im Vergleich zur Optimierung mit gewöhnlichen Differentialgleichungen, bei der in der Regel das Modell nur zeitlich diskretisiert wird, resultiert dies in einer deutlich erhöhten Anzahl an diskreten Zuständen. Das numerische Lösen solcher Probleme ist daher häufig eine Abwägung zwischen der Genauigkeit des diskreten Modells und der Laufzeit des Lösungsalgorithmus. Für eine möglichst feine Diskretisierung des Modells ist es daher sinnvoll, ein schnelles Verfahren zu verwenden.

Die beiden Ansätze SAND und NAND sind numerische, direkte Verfahren zur Lösung von PDE-restringierten Optimierungsaufgaben. In diesem Kapitel erfolgt ein Vergleich der Performanz beider Ansätze am Beispiel der Parameteridentifikation des Bohrens. Die beiden zugehörigen Optimierungsaufgaben sind in den Gleichungen (5.5) und (5.7) formuliert. Das SAND-Optimierungsproblem ist ein hochdimensionales, nichtlineares und schwach besetztes Optimierungsproblem. Das Lösen erfolgt mit der Software WORHP. Hierbei werden die in Kapitel 5.1.1 beschriebenen analytischen Ableitungen verwendet. Dies verspricht eine deutliche Effizienzsteigerung im Gegensatz zur Verwendung von

numerisch berechneten Ableitungen. Der Optimierungsalgorithmus verwendet im Hintergrund Routinen zum Lösen linearer Gleichungssysteme. Auf dieser Ebene besteht die Möglichkeit der Parallelisierung des Verfahrens. Während bei kleinen Problemgrößen durch den Overhead der Parallelisierung ein Rechenzeitverlust entstehen kann, ist das SAND-Problem durch die hochdimensionale Problemgröße für diese Art der Parallelisierung gut geeignet.

Das Lösen des NAND-Optimierungsproblems erfolgt auf zwei Ebenen. Auf der ersten Ebene wird mit WORHP die eigentliche Optimierung durchgeführt. Da die Optimierungsvariable nur die Identifikationsparameter  $\dot{q}_{\text{ident}}$ , nicht jedoch die Zustände  $T$  enthält, ist das Problem niedrig-dimensional und dicht besetzt. Die Parallelisierung der linearen Algebra innerhalb des SQP-Verfahrens ist deshalb nicht effizienzsteigernd. Das Optimierungsproblem ist zudem hochgradig nichtlinear, da die Zustände in der Zielfunktion und den Zustandsbeschränkungen (ZB) durch den Lösungsoperator der PDE mit den Optimierungsvariablen verknüpft sind. Die Bereitstellung von analytischen Ableitungen innerhalb von WORHP ist dadurch sehr schwierig, weshalb statt dem direkten analytischen Ableiten zwei Alternativen interessant sind. Zum einen der Adjoint Approach, bei dem eine weitere sogenannte adjungierte PDE zu lösen ist, und zum anderen die numerische Differentiation. Da innerhalb von WORHP bereits Routinen für das numerische Ableiten zur Verfügung stehen, werden diese für den NAND-Ansatz verwendet.

Die zweite Lösungsebene umfasst die Berechnung der von der Optimierung benötigten Zustände für ein festes  $\dot{q}_{\text{ident}}$  durch externes Lösen der PDE. Dafür werden die  $n_t$  nichtlinearen Gleichungssysteme (5.4) sequentiell durch ein Newton-Verfahren gelöst. Das Newton-Verfahren benötigt dafür die Ableitung des Gleichungssystems (5.4) bzgl. der Zustände  $\nabla_{T^l} H^l$ . Häufig wird auch diese numerisch berechnet. Durch die Implementierung des SAND-Ansatzes ist die Ableitung jedoch bereits als Teil der Jacobi-Matrix (5.6) vorhanden. Die benötigte Ableitung ist durch  $\nabla_{T^l} H^l(T^l) = \tilde{\mathbf{A}}^l$  gegeben. Die entsprechende Implementierung kommt beim NAND-Ansatz zum Einsatz. Ein Newton-Schritt wird durch Lösen eines linearen Gleichungssystems vollzogen. Da das resultierende lineare Gleichungssystem unsymmetrisch, nicht positiv definit und schwach besetzt ist, kommt für die Lösung eine Software zum Einsatz, welche für diese Eigenschaften gut geeignet ist. Sie heißt UMFPACK und ist Teil der FEM-Software deal.II, welche für die Diskretisierung der PDE verwendet wurde. Eine Parallelisierung der linearen Algebra erzeugt auch hier so viel Overhead, dass dadurch keine Geschwindigkeitssteigerung erreicht werden kann.

Für beide Ansätze sei erwähnt, dass bei der Implementierung zwar ein besonderes Augenmerk auf die Rechengeschwindigkeit gelegt wurde, jedoch ein Vergleich mit Ansätzen mit kommerzieller Software wie z. B. dem FEM-Solver ABAQUS schwer fällt, da in solcher Software oft jahrelange Entwicklung ganzer Unternehmen steckt und diese Software hinsichtlich der Performanz optimiert wurde.

Der Rechenzeitvergleich vom SAND- und NAND-Ansatz erfolgt durch einen Laufzeitvergleich vieler verschiedener Optimierungsaufgaben. Die Aufgaben unterscheiden sich



in der Anzahl an Zuständen aufgrund einer Variation der Gitterdichte, in der Anzahl der Identifikationsparametern, darin, ob Zustandsbeschränkungen verwendet werden, sowie in den zu identifizierenden Daten. Die Anzahl an diskreten Zuständen  $n_s$  wird dabei zwischen 5736 und 70 200 auf acht Stufen variiert, wobei die genaue Anzahl von der Diskretisierung abhängt. Die Anzahl der Identifikationsparameter beträgt  $n_z = 6, 12$  oder 24. Die verwendeten Daten sind entweder künstlich erzeugt, künstlich erzeugt und mit einem Rauschen versehen oder das Thermografiebild eines Experiments. Die künstlichen Daten sind jene aus dem vorangegangenen Kapitel 5.3.1 der Validierung. Die Thermografieaufnahme stammt aus der Versuchsreihe, welche in Kapitel 3.4.3 beschrieben wurde. Das zugehörige Experiment besitzt die Maschinenparameter  $v_c = 55$  m/min,  $v_f = 400$  mm/min sowie  $r = 4$  mm und wird im nächsten Kapitel der Realdatenidentifikation genauer untersucht. Ein Problem, welches bei der Verwendung von Realdaten auftritt, ist, dass die Position dieser auf dem Bauteil nicht genau bekannt ist, da die relative Position der Thermografiekamera zum Bauteil von Experiment zu Experiment leichte Translationen erfahren hat<sup>2</sup>. Die Positionierung der Daten in  $x$ -Richtung ist dabei vergleichsweise einfach, da die Bilder bis auf Messungenauigkeiten symmetrisch sind und die Symmetrieachse leicht auf die zugehörige Bauteilposition  $x = 0$  gelegt werden kann. Die richtige  $y$ -Position zu bestimmen ist schwieriger, da sie nicht einfach aus dem Bild abgelesen werden kann. Es wird davon ausgegangen, dass die heißeste Position der Bauteilfront auf Höhe des unteren Endes des Bohrlochs auftritt. Daher wird das Temperaturbild so positioniert, dass die  $y$ -Position des heißesten Punktes der Temperaturdaten mit dieser Position übereinstimmt. Die Startschätzung  $(X^{(0)}, \lambda^{(0)})$  für die Optimierung wird für alle Probleme wie bei der Validierung (siehe (5.11)) gewählt.

Auch für diese Berechnungen wird das System aus Tab. 5.2 verwendet. Für die SAND-Berechnungen kommen fünf der 16 Kerne zum Einsatz, sodass drei Berechnungen parallel ausgeführt werden können und ein Kern für Systemprogramme verbleibt. Die Rechnungen mit dem NAND-Ansatz erfolgen auf einem Kern, da die Parallelisierung wie bereits erwähnt keine Leistungssteigerung bzw. sogar einen Overhead bewirkt.

Für die Bewertung der Rechenzeiten der jeweiligen Optimierungsprobleme wurden die Laufzeiten der drei verschiedenen Datenquellen jeweils gemittelt. So fallen einzelne Ausreißer, die aufgrund von vereinzelt numerischen Schwierigkeiten beim Lösen entstehen können, nicht so stark ins Gewicht. Abbildung 5.9 zeigt diese gemittelten Rechenzeiten für beide Optimierungsansätze und für alle Aufgabenstellungen in einem Balkendiagramm. In blau sind die Rechenzeiten des SAND-Ansatzes und in grün die entsprechenden Werte des NAND-Ansatzes zu sehen. Die Werte gleicher Szenarien sind übereinander aufgetragen. Die Auswertung gliedert sich durch die Variation der Anzahl an diskreten Zuständen und an Identifikationsparametern in 24 Paare. Die Paare zeigen jeweils die Laufzeit ohne Zustandsbeschränkungen in heller Farbe und mit Zustandsbeschränkungen in dunkler Farbe nebeneinander. Die gemittelten Rechenzeiten reichen beim SAND-Ansatz von 1,7 min bis zu 3,2 h und beim NAND-Ansatz von 9,7 min bis zu 61,5 h. Die Berechnungen mit dem SAND-Ansatz sind durchweg schneller, wobei es

<sup>2</sup>Eine Rotation der Daten wurde nicht festgestellt.

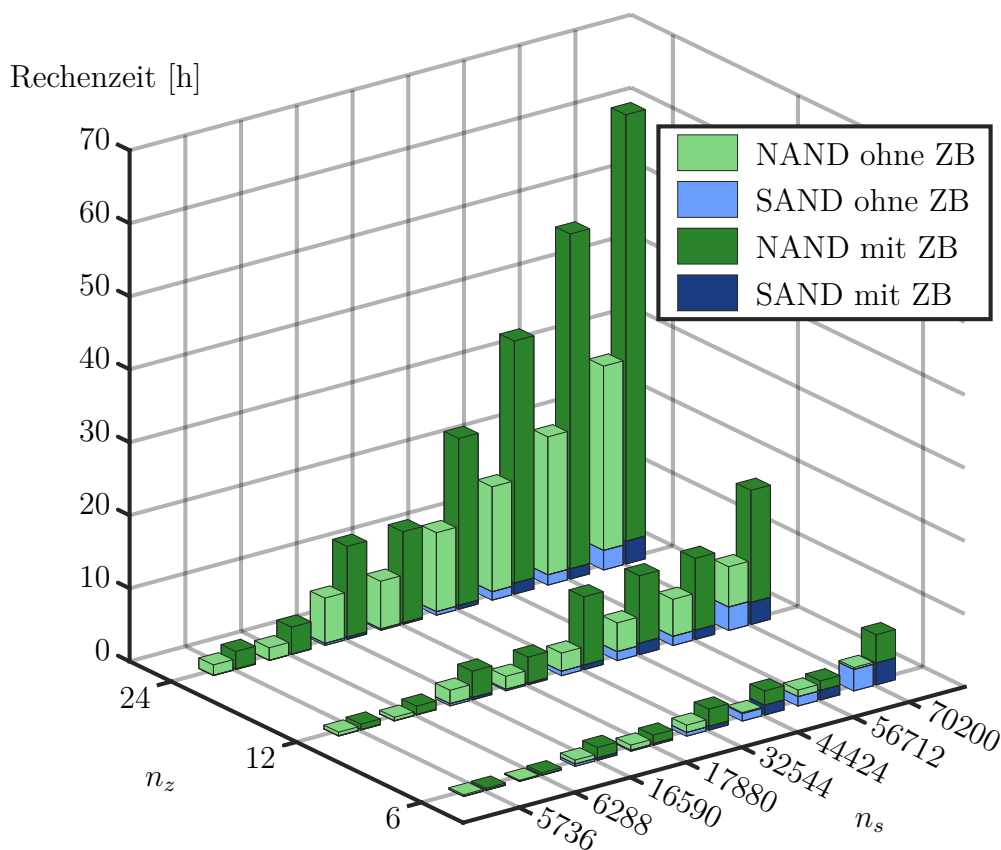


Abbildung 5.9: Über die Berechnungen mit verschiedenen Daten gemittelte Rechenzeiten der Ansätze SAND und NAND für verschiedene Optimierungsaufgaben

bei den ungemittelten Laufzeiten der einzelnen Daten-Szenarien insgesamt vier Sonderfälle gibt, bei denen der NAND-Ansatz minimal schneller ist. Diese Fälle treten jedoch nur bei  $n_z = 6$  und ohne Zustandsbeschränkungen auf. Insgesamt werden die absoluten Unterschiede der Laufzeiten mit steigendem  $n_s$  und  $n_z$  deutlich größer. Während die Differenz für  $n_s = 6288$ ,  $n_z = 6$  und ohne Zustandsbeschränkungen nur 7,5 min beträgt, ist sie für  $n_s = 70200$ ,  $n_z = 24$  und mit Zustandsbeschränkungen bereits 58,4 h. Die relativen Laufzeitunterschiede reichen von 9% bis 8971%. Die Anzahl an Hauptiterationen, die bis zur Konvergenz benötigt werden, ist bei beiden Ansätzen sehr ähnlich. Der SAND-Ansatz benötigt mit durchschnittlich 25,7 nur unwesentlich weniger Hauptiterationen als der NAND-Ansatz mit durchschnittlich 26,1. Der Geschwindigkeitsvorteil von SAND muss also durch die Berechnungsdauer der einzelnen Iterationen begründet sein und nicht durch eine schnellere Konvergenz.

Abbildung 5.10 zeigt die Laufzeiten der Ansätze in getrennten Koordinatensystemen und mit linear skalierten Achsen. In Abb. 5.10a wird deutlich, dass die Rechenzeit des NAND-Ansatzes sowohl mit steigendem  $n_s$  und  $n_z$  als auch mit der Hinzunahme von Zustandsbeschränkungen größer wird. Dies lässt sich nachvollziehen: Da das zugrundeliegende Optimierungsproblem sehr klein ist ( $n_z$  ist gleich der Anzahl der Optimierungs-

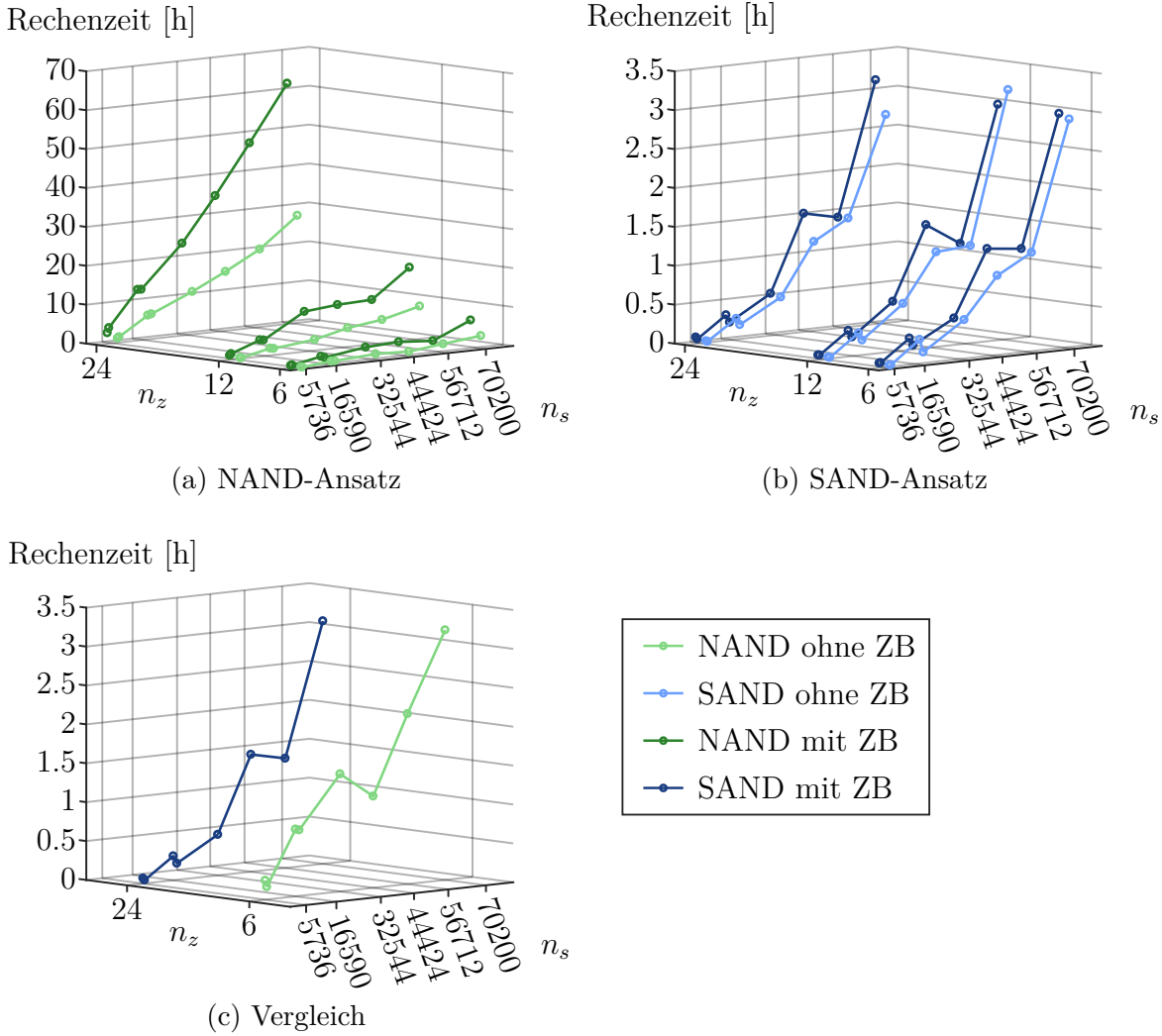


Abbildung 5.10: Rechenzeiten der einzelnen Ansätze im Detail mit linearer Achsenskalierung

variablen  $n_x$ ), wird der Großteil der Berechnungen für das Lösen der PDE durch das Newton-Verfahren benötigt. Wenn  $n_z$  größer wird, wachsen auch die Systemmatrizen innerhalb von WORHP und die Häufigkeit der Berechnungen der finiten Differenzen mit dem Newton-Verfahren steigt. Auch die Hinzunahme von Zustandsbeschränkungen erhöht die Anzahl der erforderlichen Lösungen mit dem Newton-Verfahren. Die Beschränkungen sind beim NAND-Ansatz nichtlineare Nebenbedingungen, welche das Lösen der PDE zur Ermittlung der Zustände voraussetzen. Im Unterschied dazu legt die Anzahl an Zuständen  $n_s$  die Größe des PDE-Gleichungssystems fest und somit auch die Berechnungsdauer der einzelnen Newton-Verfahren.

Beim SAND-Ansatz ergeben sich andere Implikationen (siehe Abb. 5.10b). Während die Anzahl an Zuständen  $n_s$  einen großen Einfluss auf die Rechenzeit hat, ist der Einfluss der Identifikationsparameter  $n_z$  und der Einfluss durch die zusätzliche Berücksichtigung von

Zustandsbeschränkungen vernachlässigbar klein. Die Größe des Optimierungsproblems ( $n_x = n_s + n_z$ ) wird maßgeblich durch  $n_s$  bestimmt, da  $n_z$  im Vergleich sehr klein ist ( $n_z \ll n_s$ ). So hat  $n_s$  einen sehr viel größeren Einfluss auf die Rechenzeit als  $n_z$ . Die Zustandsbeschränkungen können, da die Zustände Optimierungsvariablen sind, durch Boxschränken berücksichtigt werden. Diese sind sehr viel effizienter in der Optimierung zu berücksichtigen als die nichtlinearen Zustandsbeschränkungen beim NAND-Ansatz und haben deshalb keinen erkennbaren Anstieg der Rechenzeit zur Folge.

Dadurch, dass der SAND-Ansatz keine feststellbare Beeinträchtigung der Rechenzeit bezüglich der Anzahl der Identifikationsparameter sowie der Verwendung von Zustandsbeschränkungen aufweist, hat er entscheidende Vorteile gegenüber dem NAND-Ansatz. Die Rechenzeiten sind signifikant kürzer, insbesondere bei großem  $n_z$  und großem  $n_s$ . Abbildung 5.10c zeigt, dass sogar die größten Optimierungsprobleme vom SAND-Ansatz in etwa der gleichen Zeit gelöst werden wie die kleinsten Optimierungsprobleme vom NAND-Ansatz. Insgesamt stützen die Resultate die theoretischen Überlegungen aus Kapitel 2.1. Die Verwendung von SAND ermöglicht daher eine feinere Diskretisierung des Bauteils sowie eine feinere Auflösung der Wärmequelle. Der SAND-Ansatz ist daher die favorisierte Methodik für die Identifikation der Realdaten.

### 5.3.3 Identifikation von Realdaten

Bei der Validierung und den Berechnungen zur Laufzeitmessung konnten Erkenntnisse gewonnen werden, die kleinere Anpassungen für die Identifikation mit Realdaten sinnvoll machen. So ließ sich beobachten, dass die Temperaturen im oberen Bereich der Bilder nicht zufriedenstellend rekonstruiert werden konnten (siehe Kapitel 5.3.1). Dies liegt daran, dass die Daten, die für die Identifikation verwendet werden, aufgrund der Überlagerungen mit dem Schutzrohr großzügig nach oben beschränkt wurden. Die Beschränkung erfolgte wie beim Fräsen durch Beschneidung des  $y$ -Bereichs. Um für die Identifikation jedoch einen möglichst großen Bereich der Daten zur Verfügung stellen zu können, soll für die Realdatenidentifikation der berücksichtigte Bereich erweitert werden. Die obere Grenze dieses Bereichs wird, wie in Abb. 5.4 durch den weißen Rand dargestellt, der runden Geometrie des Schutzrohrs angepasst.

Bei den Berechnungen zum Rechenzeitvergleich wurde die Problematik der Positionierung der Temperaturdaten auf dem Bauteil in  $y$ -Richtung thematisiert. Beim dort gewählten Ansatz wurde der heißeste Punkt des Thermografiebildes mit der Höhe des unteren Endes des Bohrlochs zusammengelegt. Abbildung 5.11a zeigt auf der linken Seite ( $x > 0$  mm) ein durch die Parameteridentifikation rekonstruiertes Temperaturfeld aus dem Rechenzeitvergleich und auf der rechten Seite ( $x < 0$  mm) die zugehörigen Daten. Die Wahl der  $y$ -Position der Daten scheint nicht genau genug zu sein, da eine deutliche Verschiebung der Temperaturfelder in  $y$ -Richtung zu erkennen ist. Auch ein weiterer Versuch, bei dem die Positionen der heißesten Temperaturen von Daten und simulierten Temperaturfeldern zusammenfallen, lieferte nicht das gewünschte Resultat. Für die

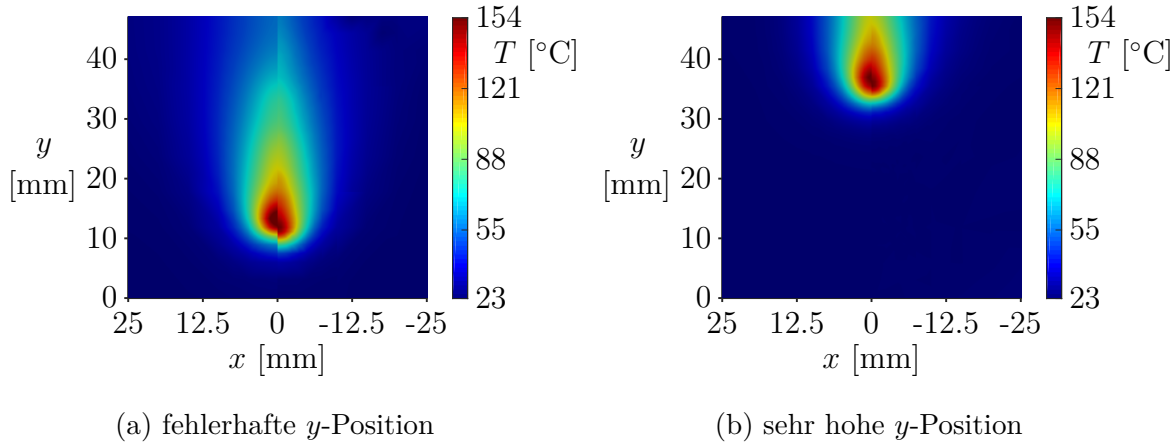


Abbildung 5.11: Probleme bei der Identifikation von Realdaten

Identifikation von Realdaten wird deshalb eine dritte Variante verfolgt, bei der eine neue Optimierungsvariable  $y_{\text{pos}} \in \mathbb{R}$  in das Problem aufgenommen wird:

$$X = (T^T, \dot{q}_{\text{ident}}^T, y_{\text{pos}}^T)^T \in \mathbb{R}^{n_x}$$

Die Variable bestimmt, auf welche  $y$ -Position des Bauteils der heißeste Datenpunkt gelegt wird. Sie wird zusammen mit den anderen Parametern optimiert. Dies ermöglicht dem Optimierungsalgorithmus das Verschieben der Daten an eine Position, an der sich die bestmögliche Übereinstimmung mit den simulierten Temperaturen ergibt. Das Hinzufügen der Variable  $y_{\text{pos}} \in \mathbb{R}$  erhöht die Anzahl an Optimierungsvariablen in WORHP auf  $n_x = n_s + n_z + 1$  und erfordert die Anpassung der Systemmatrizen. Da die Temperaturdaten in der Zielfunktion, nicht aber in den Nebenbedingungen, Verwendung finden, betrifft dies den Vektor  $\nabla F(X)$  und die Matrix  $\nabla_X^2 L(X, \lambda)$  bzw. den Summanden  $\nabla^2 F(X)$  der Hesse-Matrix. Die neue Optimierungsvariable  $y_{\text{pos}}$  beeinflusst die Interpolation der Daten auf dem Gitter, die durch die Funktion  $\vartheta(T_{\text{mess}})$  erfolgt. Dadurch, dass diese Interpolation durch Fremdsoftware<sup>3</sup> erfolgt, ist ein analytisches Ableiten schwierig. Deshalb werden die analytischen Systemmatrizen aus Kapitel 5.1.1 um numerische Ableitungen durch finite Differenzen ergänzt. Der Vektor  $\nabla F(X)$  erhält einen weiteren Eintrag, die Matrix  $\nabla_X^2 L(X, \lambda)$  eine neue Zeile sowie eine neue Spalte und die Matrix  $\nabla H(X)$  eine neue Spalte, die jedoch gleich 0 ist.

Die Verwendung der  $y$ -Position der Daten als Optimierungsvariable erfordert eine notwendige Erweiterung des Optimierungsproblems. Da der Algorithmus die Freiheit besitzt, die Daten beliebig in  $y$ -Richtung zu verschieben, neigt er dazu, die Daten nach oben zu schieben (siehe Abb. 5.11b). Dies ist für die Minimierung der Zielfunktion sinnvoll, weil so der Bereich kleiner Temperaturen vergrößert wird. Dieser Bereich ist leichter zu identifizieren als ein Bereich hoher Temperaturen. Die unter dem heißen Bereich liegende primäre sowie einige sekundäre Wärmequellen werden dabei auf 0 gesetzt, um

<sup>3</sup>Für die Interpolation werden kubische Splines der Software `inspline` [37] verwendet.

den niedrigen Temperaturen zu entsprechen. Die höher liegenden Wärmequellen werden entsprechend erhöht. Allein mit Boxschränken an die Variable  $y_{\text{pos}}$  kann diesem Verhalten nicht entgegengewirkt werden, da eine sinnvolle Wahl der Schranken schwer zu ermitteln ist. Stattdessen genügt es, kleine primäre Wärmeintensitäten zu verhindern. Eine große primäre Wärmequelle ist nicht mit einer hohen Datenposition in Einklang zu bringen. Deshalb wird nachfolgend gefordert, dass die primäre Wärmequelle die größte aller Wärmequellen ist:

$$\dot{q}_{\text{ident},1} \geq \dot{q}_{\text{ident},i} \Leftrightarrow \dot{q}_{\text{ident},1} - \dot{q}_{\text{ident},i} \geq 0, \quad i = 2, \dots, n_z$$

Diese Bedingungen schränken den Lösungsraum nicht zusätzlich ein, da davon auszugehen ist, dass die primäre Wärmequelle ohnehin größer als die sekundären Wärmequellen ist. Die neuen Bedingungen erweitern das Optimierungsproblem um  $m_g = n_z - 1$  Ungleichungsnebenbedingungen  $G(X)$ . Die Ableitungen für den Optimierungsalgorithmus sind

$$\nabla G(X) = \begin{pmatrix} 0 & \cdots & 0 & 1 & -1 & & 0 \\ \vdots & & \vdots & \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & & & -1 & 0 \end{pmatrix} \in \mathbb{R}^{m_g \times n_x}$$

und  $\nabla^2 G_i(X) = 0 \in \mathbb{R}^{n_x \times n_x}, \quad i = 1, \dots, m_g.$

Die für die Identifikation verwendeten Thermografieaufnahmen stammen aus den Experimenten, die in Kapitel 3.4.3 beschrieben wurden. In der Versuchsreihe wurden die Maschinenparameter Schnittgeschwindigkeit  $v_c$ , Vorschubgeschwindigkeit  $v_f$  sowie der Bohrradius  $r$  variiert. Die verwendeten Werte sind in Tab. 5.3 zu sehen. Alle möglichen Parameterkonstellationen wurden in fünf Experimenten wiederholt, sodass insgesamt die Thermografiebilder von 40 Bohrversuchen vorliegen. Davon werden vier Ergebnisse vorgestellt.

Stufe	$v_c$ [m/min]	$v_f$ [mm/min]	$r$ [mm]
1	55	400	4
2	75	600	6

Tabelle 5.3: Verwendete Maschinenparameter der thermischen Experimente des Bohrens

Die für das Modell benötigte Anfangstemperatur  $T_0$  liegt zu den Experimenten nicht exakt vor, sodass die Temperatur wie beim Fräsen aus den größtenteils unbeeinflussten Bereichen der Bilder entnommen wird. Die Anzahl der Identifikationsparameter beträgt bei allen Berechnungen  $n_z = 24$ . Die Anzahl an simulierten Zeitschritten beträgt ebenfalls  $n_t = 24$ . Die Anzahl der diskreten Zustände aller Zeitschritte variiert mit dem

Bohrerradius und beträgt  $n_s = 104\,424$  bzw.  $n_s = 112\,896$ . Bei der Identifikation von Realdaten haben Tests gezeigt, dass es sinnvoll sein kann, eine zulässige Startschätzung  $(X^{(0)}, \lambda^{(0)})$  für das SQP-Verfahren in WORHP zu wählen. Deshalb wird einmalig eine fiktive Temperaturverteilung  $T_{\text{sim}}$  durch Lösen des Modells simuliert. Diese wird dann mit den zugehörigen Wärmestromdichten als Startschätzung gewählt:

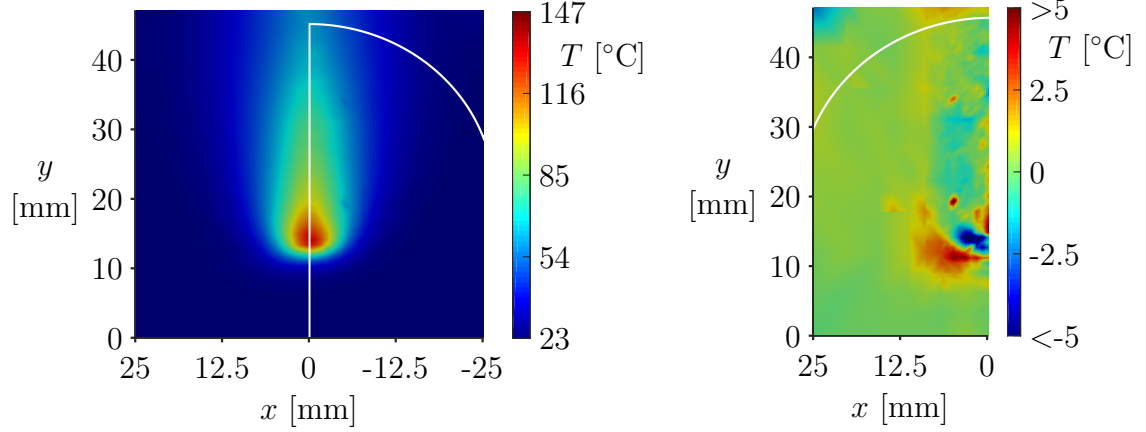
$$\begin{aligned} T^{(0)} &= T_{\text{sim}} \\ \dot{q}_{\text{ident},1}^{(0)} &= 3,5 \text{ W/mm}^2 \\ \dot{q}_{\text{ident},2}^{(0)} &= 0,35 \text{ W/mm}^2 \\ \dot{q}_{\text{ident},i}^{(0)} &= 0,035 \text{ W/mm}^2, \quad i = 3, \dots, n_z \\ \lambda_i^{(0)} &= 0, \quad i = 1, \dots, m \end{aligned}$$

Für die Startschätzung wird  $y_{\text{pos}}^{(0)}$  so gewählt, dass die größten Temperaturwerte von  $T_{\text{sim}}$  und  $T_{\text{mess}}$  auf der gleichen Position liegen.

### Erstes Experiment

Die erste Optimierung erfolgte mit den Daten eines Experiments, bei dem zwischen Bohrloch und Bauteilfront eine 0,65 mm dünne Wand entstanden ist und welches mit einer Schnittgeschwindigkeit von  $v_c = 55 \text{ m/min}$ , einer Vorschubgeschwindigkeit von  $v_f = 600 \text{ mm/min}$  und einem Bohrerradius von  $r = 6 \text{ mm}$  durchgeführt wurde. Dies führt auf einen Zahnvorschub von  $f_z \approx 0,21 \text{ mm}$  und mit der Bohrlochtiefe von 35 mm auf eine Prozesszeit von  $t_f - t_0 = 3,5 \text{ s}$ . Mit den  $n_t = 24$  Zeitschritten werden im Modell je sieben Schnitte in einem Zeitschritt zusammengefasst. Das gespiegelte Thermografiebild ist in Abb. 5.12a auf der rechten Seite ( $x < 0 \text{ mm}$ ) zu sehen. Der weiße Rand markiert jenen Bereich, welcher für die Identifikation verwendet wurde. Aus diesen Daten kann eine Anfangstemperatur von  $T_0 \approx 22,84 \text{ }^\circ\text{C}$  und eine Durchschnittstemperatur von  $57,4 \text{ }^\circ\text{C}$  ermittelt werden.

Der Optimierung mit WORHP stehen insgesamt  $n_x = 104\,449$  Optimierungsvariablen zur Verfügung. Sie benötigt 2,2 h und 15 Hauptiterationen zur Konvergenz. Das rekonstruierte Temperaturfeld wird in Abb. 5.12a, linke Seite ( $x > 0 \text{ mm}$ ) gezeigt. Es weist optisch eine hohe Übereinstimmung mit den Daten auf. Es treten lediglich kleine Unterschiede in der Form des Flammenprofils auf, welches im rekonstruierten Temperaturfeld gleichmäßiger ist. Es ist zudem zu erkennen, dass kleinere Bereiche ( $x \approx -5 \text{ mm}$ ,  $y \approx 19 \text{ mm}$  und  $x \approx -5 \text{ mm}$ ,  $y \approx 34 \text{ mm}$ ) mit Störungen in den Daten durch die Identifikation eliminiert werden können. Auffällig ist im Vergleich zur Validierung, dass durch die Erweiterung des verwendeten Datenbereichs im oberen Teil des Bildes eine deutlich bessere Rekonstruktion ermöglicht wird. Auch die anderen Maßnahmen scheinen die Identifikation zu begünstigen. Die optimierte  $y$ -Position des heißesten Datenpunktes  $y_{\text{pos}}^* \approx 0,0134$  wurde gut an das identifizierte Temperaturfeld angepasst. Abbildung



(a) links ( $x > 0$  mm):  $T^{n_t^*}|_{z=0}$ , rechts ( $x < 0$  mm):  $T_{\text{mess}}$  (gespiegelt) (b)  $T^{n_t^*}|_{z=0} - T_{\text{mess}}$

Abbildung 5.12: Ergebnisse der Identifikation des ersten Experiments

5.12b stützt diesen Eindruck. Es zeigt die Differenz von simulierten und gemessenen Temperaturen sowie den Bereich, in dem sich die berücksichtigten Daten befinden. Es überwiegt ein großer Bereich, der eine Differenz nahe 0 aufweist. In der Nähe der primären Wärmequelle sind die Abweichungen am größten. Dies ist nachvollziehbar, da dort auch absolut die höchsten Temperaturen auftreten. Die Differenz beträgt im Minimum  $-7,2^\circ\text{C}$  und im Maximum  $6,4^\circ\text{C}$ . Auch die erwähnten Störungen in den Daten sind in der Differenzen-Abbildung gut zu erkennen. Der Datenbereich, welcher der Identifikation nicht zur Verfügung stand, wird mit Ausnahme des Randes gut rekonstruiert. Die Abweichung in der Ecke ( $x = 25$  mm,  $y = 47$  mm) deutet auf das Schutzrohr hin. Die mittlere absolute Abweichung der berücksichtigten Temperaturdaten von den entsprechenden simulierten Temperaturen beträgt

$$\sum_{i=1}^{n_s} \bar{w}_i |T_i^* - \vartheta_i(T_{\text{mess}})| \approx 0,8^\circ\text{C}.$$

Dies entspricht einer mittleren relativen Abweichung von

$$\sum_{i=1}^{n_s} \bar{w}_i \left| \frac{T_i^* - \vartheta_i(T_{\text{mess}})}{\vartheta_i(T_{\text{mess}})} \right| \cdot 100\% \approx 1,7\%.$$

Die niedrigen Abweichungen werden vom finalen Zielfunktionswert  $F(X^*) \approx 1,8^\circ\text{C}^2$  bestätigt. Der Zielfunktionswert beschreibt die mittlere quadratische Abweichung.

Die identifizierten Wärmestromdichten sind in Abb. 5.13 mit der zugehörigen  $y$ -Position der Wärmequelle im letzten Zeitschritt zu sehen. Die primäre Wärmestromdichte beträgt  $\dot{q}_{\text{ident},1}^* \approx 1,94 \text{ W/mm}^2$ . Auffällig ist, dass die erste sekundäre Wärmequelle  $\dot{q}_{\text{ident},2}^*$  den



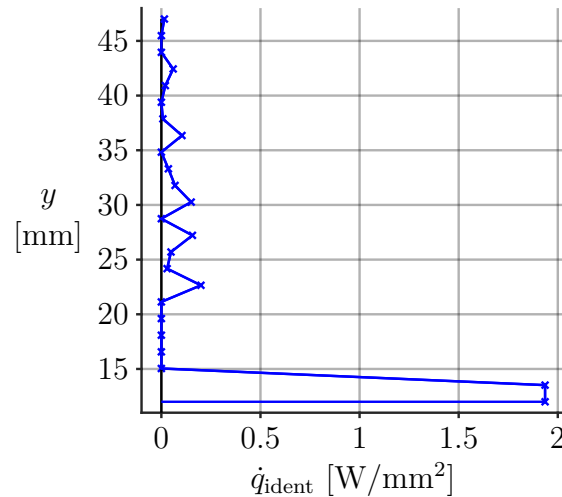


Abbildung 5.13: Identifizierte Wärmestromdichten des ersten Experiments

gleichen Wert wie die primäre Wärmequelle besitzt. Dies lässt sich vermutlich dadurch erklären, dass durch das Zusammenfassen mehrerer Schnitte in einen Zeitschritt eine klare Abgrenzung zwischen Schnittbereich und dem Bereich des sekundären Wärmeeintrags im Modell nicht möglich ist. Hier könnte eine feinere zeitliche Auflösung Abhilfe schaffen. Dies würde allerdings mit einer Erhöhung der Rechenzeit einhergehen. Es fällt zudem auf, dass die sekundäre Wärmequelle nicht monoton abfällt, sondern es auch Abschnitte gibt, in denen die Intensität nahe 0 liegt. Auch hier ist mit numerischen Ungenauigkeiten zu rechnen. Es könnte jedoch darauf hinweisen, dass es Bereiche im Bohrkanaal gibt, die stärker beansprucht werden als andere. Die Reibung könnte bei einem nicht perfekt symmetrischen Bohrer ungleichmäßig verteilt sein oder aber die Dynamik der abtransportierten Späne könnte für einen ungleichmäßigen Wärmeeintrag sorgen. Ansonsten entspricht das Ergebnis den Erwartungen. Die primäre Wärmequelle ist deutlich größer als die durchschnittliche sekundäre Wärmequelle. Auch die Intensität der sekundären Wärmequelle nimmt im Mittel nach oben hin ab. Dass die Wärmestromdichten im Vergleich zu den Resultaten beim Fräsen (siehe Kapitel 4.5) deutlich kleiner sind, liegt an der Modellierung. Beim Bohren werden mehrere Zeitschritte zusammengefasst, weshalb die Fläche, über die die Wärme ins Bauteil eingebracht wird, größer ist. Dies wird durch eine kleinere identifizierte Wärmestromdichte kompensiert.

Eine weitere Größe, um spanende Fertigungsprozesse zu quantifizieren, ist die Wärmeleistung. Sie berechnet sich aus dem Produkt der Wärmestromdichte und der Fläche, auf der diese wirkt. Sie ist im Allgemeinen zeitvariant. Die Leistung der primären Wärmequelle  $\dot{Q}_{\text{prim}}$  ist hier jedoch konstant, da die Wärmestromdichte und die Größe der Fläche  $A_{\text{prim}}$ , auf der sie wirkt, zeitunabhängig sind. Sie beträgt

$$\dot{Q}_{\text{prim}} = \dot{q}_{\text{ident},1} * A_{\text{prim}} = \dot{q}_{\text{ident},1} * 2 \pi r f_{\text{prim}} \approx 106,4 \text{ W}.$$

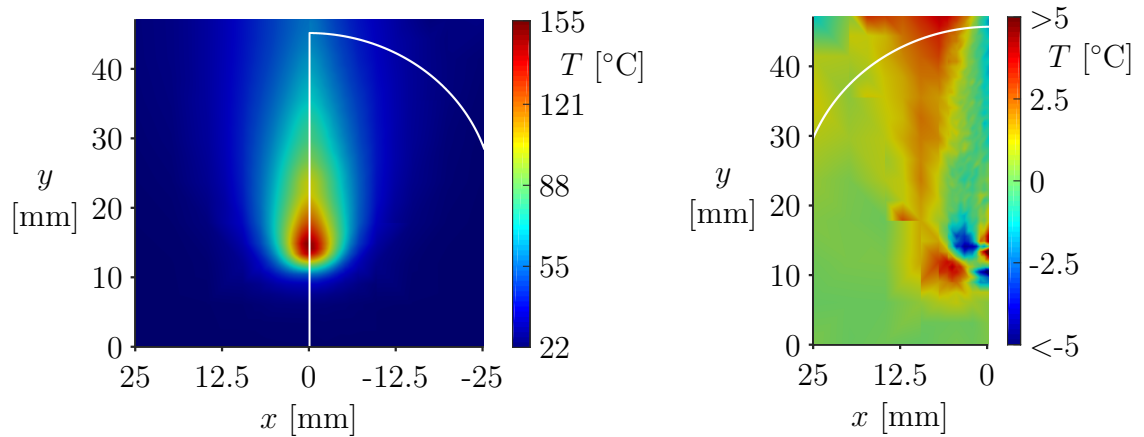
Die Summe der Leistungen der sekundären Wärmequellen  $\dot{Q}_{\text{sek}}$  nimmt mit der Zeit zu, da die Anzahl der Wärmequellen, die auf das Bauteil wirken, im Laufe des Prozesses

zunimmt. Zum Endzeitpunkt beträgt die Leistung

$$\dot{Q}_{\text{sek}} = \sum_{i=2}^{n_z} \dot{q}_{\text{ident},i}^* A_{\text{sek}} = \sum_{i=2}^{n_z} \dot{q}_{\text{ident},i}^* 2 \pi r f_{\text{sek}} \approx 155,7 \text{ W}.$$

## Zweites Experiment

Das zweite Experiment, auf das die Parameteridentifikation angewandt wurde, besitzt die Maschinenparameter  $v_c = 55 \text{ m/min}$ ,  $v_f = 400 \text{ mm/min}$  und  $r = 4 \text{ mm}$ . Dies hat einen Zahnvorschub von  $f_z \approx 0,09 \text{ mm}$  zur Folge. Die Prozesszeit beträgt  $t_f - t_0 = 5,25 \text{ s}$  und die Wandstärke zwischen Vorderseite und Bohrloch  $0,53 \text{ mm}$ . In einem Zeitschritt werden 16 Schnitte zusammengefasst. Die ermittelte Anfangstemperatur beträgt  $T_0 \approx 21,8^\circ\text{C}$ . Die Temperaturdaten sind gespiegelt in Abb. 5.14a auf der rechten Seite ( $x < 0 \text{ mm}$ ) zu sehen. Die Durchschnittstemperatur der Daten beträgt  $64,9^\circ\text{C}$ .



(a) links ( $x > 0 \text{ mm}$ ):  $T^{n_t^*}|_{z=0}$ , rechts ( $x < 0 \text{ mm}$ ):  $T_{\text{mess}}$  (gespiegelt) (b)  $T^{n_t^*}|_{z=0} - T_{\text{mess}}$

Abbildung 5.14: Ergebnisse der Identifikation des zweiten Experiments

Die linke Seite ( $x > 0 \text{ mm}$ ) der Abb. 5.14a zeigt das resultierende Temperaturfeld nach der Optimierung mit WORHP. Insgesamt gibt es  $n_x = 112921$  Optimierungsvariablen. Die Anzahl an Hauptiterationen ist mit 29 vergleichsweise hoch, weshalb auch die Berechnungsdauer mit 4,6 h im Vergleich zu den anderen Experimenten erhöht ist. Die optische Übereinstimmung zwischen Daten und simulierten Werten ist im Vergleich zum ersten Versuch sogar noch etwas besser. Das könnte daran liegen, dass die Daten optisch weniger Fehler aufweisen. Es treten keine optischen Störungen auf und auch die geometrische Form des gemessenen Temperaturfeldes ist gleichmäßiger. Sowohl die Rekonstruktion des oberen Datenbereichs als auch die Optimierung der  $y$ -Position der Daten ( $y_{\text{pos}}^* \approx 0,0139 \text{ mm}$ ) sind erfolgreich. Die Differenz der beiden Temperaturfelder ist in Abb. 5.14b zu sehen. Dabei fällt zunächst auf, dass die Bereiche, in denen die

Differenz höher als  $1\text{ }^\circ\text{C}$  ist, größer als im ersten Versuch sind. Die minimale Differenz ist mit  $-6,5\text{ }^\circ\text{C}$  jedoch kleiner, die maximale Differenz mit  $6,6\text{ }^\circ\text{C}$  in etwa gleich groß. Diese Extremwerte treten erneut in unmittelbarer Nähe der Wärmequelle auf. Der in  $y$ -Richtung langgezogene Bereich um  $x \approx 10\text{ mm}$  mit einer positiven Differenz größer  $1\text{ }^\circ\text{C}$  verdeutlicht, dass das rekonstruierte Flammenprofil insgesamt etwas breiter ist als im Thermografiebild. Die simulierten Temperaturen weichen im absoluten Mittel von den berücksichtigten Temperaturdaten um

$$\sum_{i=1}^{n_s} \bar{w}_i |T_i^* - \vartheta_i(T_{\text{mess}})| \approx 1,0\text{ }^\circ\text{C}$$

ab. Die mittlere relative Abweichung beträgt

$$\sum_{i=1}^{n_s} \bar{w}_i \left| \frac{T_i^* - \vartheta_i(T_{\text{mess}})}{\vartheta_i(T_{\text{mess}})} \right| \cdot 100\text{ mm} \approx 1,9\%$$

und der Zielfunktionswert  $F(X^*) \approx 1,9\text{ }^\circ\text{C}^2$ . Diese Werte bestätigen den Eindruck, dass dieses zweite Resultat etwas schlechter ist als das Erste, insgesamt ist es jedoch zufriedenstellend.

Die identifizierte primäre Wärmequelle hat den Wert  $\dot{q}_{\text{ident},1}^* \approx 3,82\text{ W/mm}^2$ . In Abb. 5.15 sind sowohl dieser als auch die Werte der sekundären Wärmequellen aufgetragen. Die Werte erfüllen die Erwartungen. Die primäre Wärmequelle ist deutlich größer als alle anderen. Die Intensität der sekundären Wärmequelle nimmt im Mittel nach oben hin ab. Im ersten Versuch ist die primäre Wärmequelle in etwa halb so groß, obwohl die Temperaturfelder beider Experimente eine ähnliche Geometrie und Maximaltemperatur aufweisen. Der große Unterschied zwischen den primären Wärmequellen kommt wohl

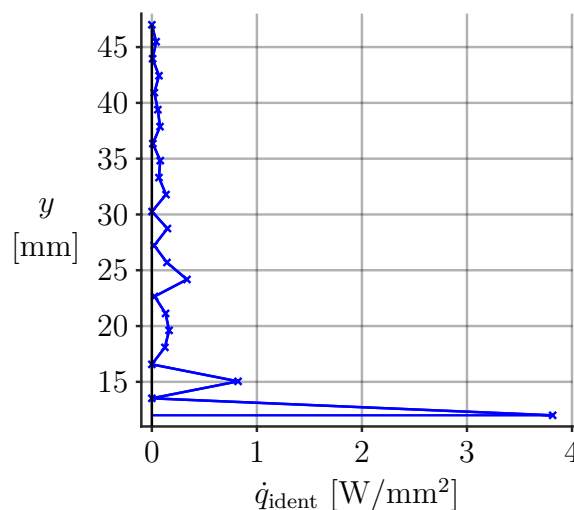


Abbildung 5.15: Identifizierte Wärmestromdichten des zweiten Experiments

dadurch zustande, dass im ersten Versuch die erste sekundäre Wärmequelle sehr groß ist und damit die vergleichsweise kleine primäre Wärmequelle kompensiert wird.

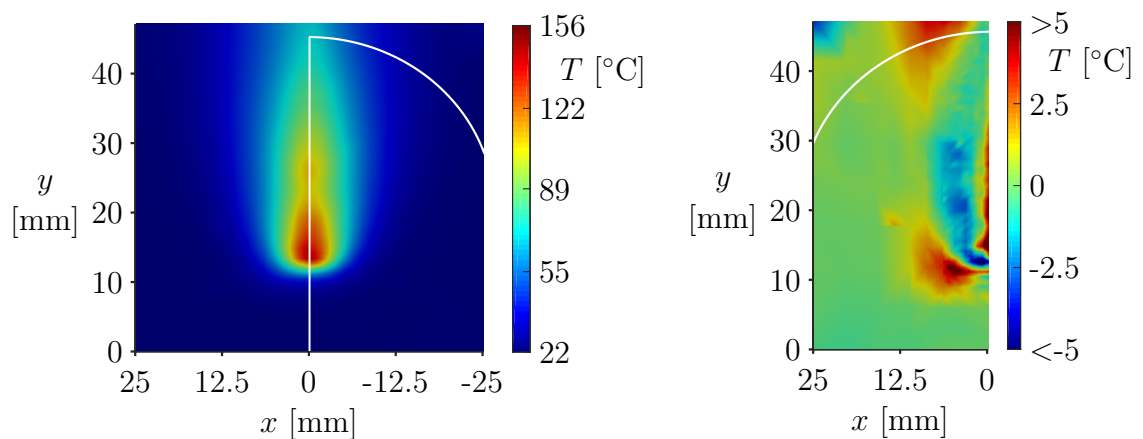
Die Wärmeleistungen betragen im zweiten Versuch

$$\dot{Q}_{\text{prim}} \approx 139,9 \text{ W} \quad \text{und} \quad \dot{Q}_{\text{sek}} \approx 91,0 \text{ W}.$$

Die Gesamtleistung ist hier trotz der großen Wärmestromdichten geringer als im ersten Experiment. Dies liegt an den Flächen  $A_{\text{prim}}$  und  $A_{\text{sek}}$ , die durch den Radius  $r = 4 \text{ mm}$  kleiner sind.

### Drittes Experiment

Die Schnittgeschwindigkeit des dritten Experiments beträgt  $v_c = 75 \text{ m/min}$ , die Vorschubgeschwindigkeit  $v_f = 600 \text{ mm/min}$  und der Bohrerradius  $r = 6 \text{ mm}$ . Der Zahnvorschub ist dadurch  $f_z \approx 0,15 \text{ mm}$ . Die Dicke der Wand zwischen Vorderseite und Bohrloch ist  $0,5 \text{ mm}$  und die Anfangstemperatur  $T_0 \approx 22,5 \text{ }^\circ\text{C}$ . Die  $n_t = 24$  Zeitschritte umfassen eine Prozessdauer von  $t_f - t_0 = 3,5 \text{ s}$ . Jeder Zeitschritt bildet dabei 10 Schnitte ab. Das Thermografiebild des Experiments bei einer Bohrlochtiefe von  $35 \text{ mm}$  ist in gespiegelter Form in Abb. 5.16a auf der rechten Seite ( $x < 0 \text{ mm}$ ) zu sehen. Die Daten haben eine Durchschnittstemperatur von  $66,8 \text{ }^\circ\text{C}$ . Auffällig ist die langgezogene Geometrie der warmen Bereiche. Diese wird sicherlich dadurch begünstigt, dass die Wandstärke bei diesem Experiment vergleichsweise gering ist. Zudem nimmt die Temperatur ausgehend von der wärmsten Stelle ( $x \approx 0 \text{ mm}$ ,  $y \approx 14 \text{ mm}$ ) zunächst nach oben hin ab und steigt dann noch einmal an, was ein zweites, jedoch kleineres Temperaturextremum im Bereich  $x \approx 0 \text{ mm}$  und  $y \approx 27 \text{ mm}$  zur Folge hat.



(a) links ( $x > 0$ ):  $T^{n_t*}|_{z=0}$ , rechts ( $x < 0$ ):  $T_{\text{mess}}$  (gespiegelt)

(b)  $T^{n_t*}|_{z=0} - T_{\text{mess}}$

Abbildung 5.16: Ergebnisse der Identifikation des dritten Experiments

Die Optimierung von  $n_x = 104\,449$  Optimierungsvariablen dauert 2,5 h und benötigt 17 Hauptiterationen. Abbildung 5.16a zeigt auf der linken Seite ( $x > 0$  mm) das rekonstruierte Temperaturfeld der Parameteridentifikation. Die optimierte  $y$ -Position  $y_{\text{pos}}^* \approx 0,0133$  mm ermöglicht optisch eine hohe Übereinstimmung mit den Daten. Die Differenz beider Temperaturfelder ist in Abb. 5.16b dargestellt. Die Extremwerte der Differenzen sind  $-7,9^\circ\text{C}$  und  $9,1^\circ\text{C}$  und befinden sich in der Nähe der primären Wärmequelle. Die Werte sind höher als in den Experimenten zuvor. Dafür sind weite Teile des Differenzen-Temperaturfeldes annähernd 0. Die mittleren Abweichungen betragen

$$\sum_{i=1}^{n_s} \bar{w}_i |T_i^* - \vartheta_i(T_{\text{mess}})| \approx 1,3^\circ\text{C}$$

bzw.

$$\sum_{i=1}^{n_s} \bar{w}_i \left| \frac{T_i^* - \vartheta_i(T_{\text{mess}})}{\vartheta_i(T_{\text{mess}})} \right| \cdot 100 \approx 2,1\%.$$

Durch die quadratischen Differenzen in der Zielfunktion haben die hohen Minimal- und Maximalwerte der Differenzen auch einen hohen finalen Zielfunktionswert zur Folge. Dieser ist  $F(X^*) \approx 2,9^\circ\text{C}^2$ .

Die identifizierten Wärmestromdichten sind in Abb. 5.17 zu sehen. Die primäre Wärmestromdichte beträgt  $\dot{q}_{\text{ident},1}^* \approx 3,53 \text{ W/mm}^2$ . Dem in den Daten erkennbaren lokalen Temperaturmaximum ( $x \approx 0$  mm und  $y \approx 27$  mm) wird dadurch Rechnung getragen, dass die identifizierte sekundäre Wärmequelle bei  $y \approx 27$  mm ebenfalls einen vergleichsweise hohen Wert aufweist. An dieser Stelle entsteht möglicherweise im Bohrkanal besonders viel Reibung oder es findet ein erhöhter Wärmeaustausch mit den Spänen statt. Ab diesem zweiten Extremum nimmt die Intensität der sekundären Wärmequelle im Durchschnitt ab.

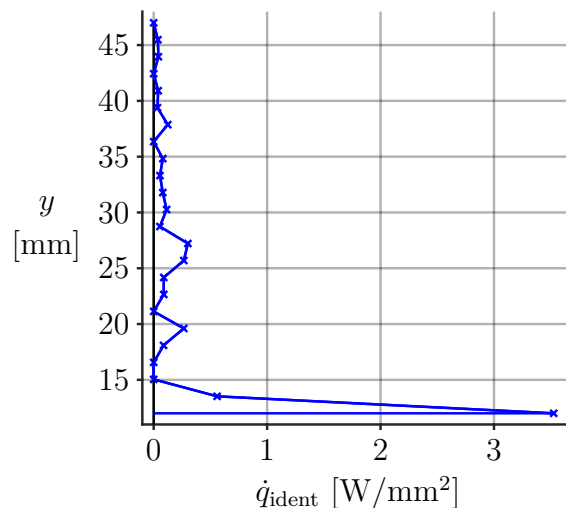


Abbildung 5.17: Identifizierte Wärmestromdichten des dritten Experiments

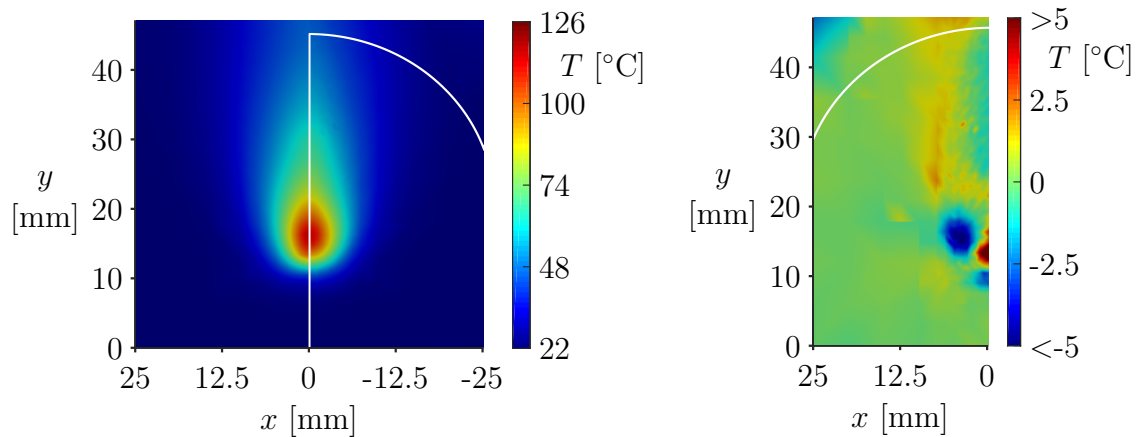
Im dritten Versuch sind die identifizierten Wärmeleistungen

$$\dot{Q}_{\text{prim}} \approx 193,9 \text{ W} \quad \text{und} \quad \dot{Q}_{\text{sek}} \approx 127,3 \text{ W}.$$

Die Gesamtleistung ist aufgrund der großen Wärmestromdichten höher als im ersten Versuch und aufgrund der größeren Fläche ( $r = 6 \text{ mm}$ ) höher als im zweiten Versuch. Die hohe Leistung wird durch das vergleichsweise warme Temperaturbild (Durchschnittstemperatur:  $66,8^\circ\text{C}$ ) bestätigt.

### Viertes Experiment

Die Temperaturdaten des vierten Experiments sind in Abb. 5.18a rechts ( $x < 0 \text{ mm}$ ) zu sehen. Die Durchschnittstemperatur der Daten beträgt  $52,3^\circ\text{C}$ , die aus den Daten ermittelte Anfangstemperatur  $T_0 \approx 21,7^\circ\text{C}$ . Der Bohrprozess wurde mit den Maschinenparametern  $v_c = 75 \text{ m/min}$ ,  $v_f = 400 \text{ mm/min}$  sowie  $r = 4 \text{ mm}$  durchgeführt. Dabei ergibt sich eine Wandstärke von  $0,85 \text{ mm}$  und ein Zahnvorschub von  $f_z \approx 0,07 \text{ mm}$ . Der kleine Zahnvorschub  $f_z$  führt dazu, dass 22 Schnitte in einem Zeitschritt zusammengefasst werden. Die Prozessdauer beträgt  $t_f - t_0 = 5,25 \text{ s}$ .



(a) links ( $x > 0$ ):  $T^{n_t^*}|_{z=0}$ , rechts ( $x < 0$ ):  $T_{\text{mess}}$  (gespiegelt)

(b)  $T^{n_t^*}|_{z=0} - T_{\text{mess}}$

Abbildung 5.18: Ergebnisse der Identifikation des vierten Experiments

Die Parameteridentifikation benötigt 2,7 h, um in 15 Hauptiterationen die  $n_x = 112921$  Variablen zu optimieren. Die Datenrekonstruktion ist in Abb. 5.18a auf der linken Seite ( $x > 0 \text{ mm}$ ) zu sehen. Die optische Übereinstimmung ist nicht ganz so hoch wie in den Versuchen zuvor. Durch den optimierten Wert  $y_{\text{pos}}^* = 0,0157 \text{ mm}$  gelingt es, die Geometrie der Daten im Bereich  $y > 15 \text{ mm}$  großflächig zu rekonstruieren. Im Bereich darunter entsteht jedoch eine leichte Verschiebung der Geometrien. In einem Teil der Daten in der Nähe vom Punkt  $x \approx 5 \text{ mm}$  und  $y \approx 31 \text{ mm}$  treten leichte Störungen auf, die durch

die Rekonstruktion kompensiert werden können. Die Differenzen-Abbildung 5.18b offenbart eine bessere Rekonstruktion als es der optische Vergleich vermuten lässt. Bis auf die Bereiche in der Nähe der primären Wärmequelle ist eine großflächige Übereinstimmung der Temperaturen erkennbar. Der Abfall der Differenz in der Ecke  $x = 25 \text{ mm}$  und  $y = 47 \text{ mm}$  deutet das Schutzrohr an. Die minimale Differenz beträgt  $-4,9 \text{ }^\circ\text{C}$ , die maximale  $8,5 \text{ }^\circ\text{C}$ . Die mittlere absolute Abweichung der simulierten Temperaturen von den gemessenen ist

$$\sum_{i=1}^{n_s} \bar{w}_i |{}^*T_i - \vartheta_i(T_{\text{mess}})| \approx 1,0 \text{ }^\circ\text{C}.$$

Die mittlere relative Abweichung beträgt

$$\sum_{i=1}^{n_s} \bar{w}_i \left| \frac{{}^*T_i - \vartheta_i(T_{\text{mess}})}{\vartheta_i(T_{\text{mess}})} \right| \cdot 100 \approx 1,8 \text{ } \%$$

Der optimale Zielfunktionswert  $F(X^*) \approx 2,0 \text{ }^\circ\text{C}^2$  bestätigt die niedrigen Werte.

Die Wärmestromdichten sind in Abb. 5.19 visualisiert. Neben dem hohen Wert für  $\dot{q}_{\text{ident},1}^* \approx 3,9 \text{ W/mm}^2$  ist auch der Wert  $\dot{q}_{\text{ident},4}^* \approx 1,0 \text{ W/mm}^2$  der sekundären Wärmequelle vergleichsweise hoch. Zusammen sorgt dies für eine relativ große und runde Geometrie des Bereichs mit Temperaturen über  $90 \text{ }^\circ\text{C}$ . Dadurch, dass die Temperaturen nach oben schnell abfallen, sind die Werte der sekundären Wärmequelle ansonsten sehr klein.

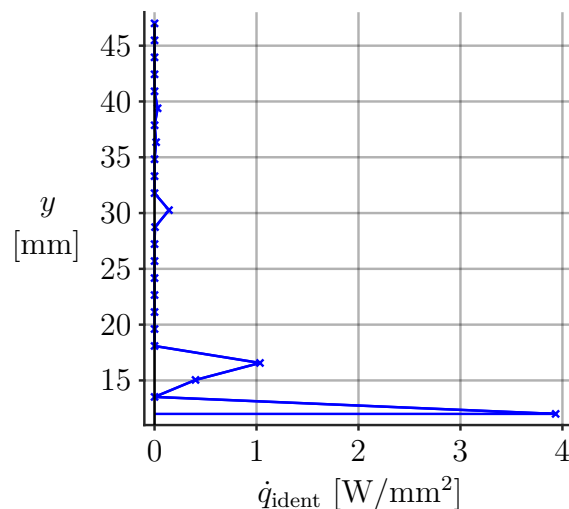


Abbildung 5.19: Identifizierte Wärmestromdichten des vierten Experiments

Auch die Gesamtleistung ist wegen den vergleichsweise kleinen Temperaturen (Maximaltemperatur  $126 \text{ }^\circ\text{C}$ , Durchschnittstemperatur  $52,3 \text{ }^\circ\text{C}$ ) und der kleineren Fläche ( $r = 4 \text{ mm}$ ) gering. Die primäre und die sekundäre Wärmeleistung betragen

$$\dot{Q}_{\text{prim}} \approx 144,2 \text{ W} \quad \text{und} \quad \dot{Q}_{\text{sek}} \approx 59,7 \text{ W}.$$

## 5.4 Fazit der Parameteridentifikation des Bohrens

Die Parameteridentifikation zur Ermittlung von Wärmestromdichten konnte nach dem Fräsen auch erfolgreich auf das Bohren angewandt werden. Dabei kamen die beiden Ansätze SAND und NAND zum Einsatz. Zunächst wurde eine Validierung durchgeführt, bei der künstliche Daten und künstliche Daten mit Rauschen verwendet wurden. Bis auf den Datenbereich, welcher für die Identifikation ausgespart wurde, ist die Rekonstruktion der Daten nahezu perfekt. Beide Ansätze liefern numerisch sehr ähnliche Resultate, sodass bezüglich der Qualität der Ergebnisse kein Ansatz gegenüber dem anderen zu bevorzugen ist. Als nächstes wurden die Rechenzeiten beider Ansätze miteinander verglichen. Dabei konnten die theoretischen Überlegungen bestätigt werden. Der SAND-Ansatz hat in den getesteten Optimierungsaufgaben einen deutlichen Geschwindigkeitsvorteil gegenüber dem NAND-Ansatz. Dies ist insbesondere der Fall, wenn die Anzahl an Identifikationsparametern hoch ist und wenn Zustandsbeschränkungen verwendet werden. Im Mittel ist der SAND-Ansatz in etwa um den Faktor 17 schneller als der NAND-Ansatz.

Diese deutlichen Ergebnisse des Rechenzeitvergleichs gaben Anlass, bei der Realdaten-Identifikation den SAND-Ansatz zu verwenden. Dafür wurde die Methodik zunächst erweitert. Zum einen wurde der Datenbereich, der für die Identifikation verwendet wird, vergrößert und an die Geometrie des runden Schutzrohrs angepasst, um die Datenrekonstruktion im oberen Teil des Bildes zu verbessern. Zum anderen wurde die Identifikation um die  $y$ -Position der Daten erweitert, da diese nicht präzise genug aus den Experimenten bekannt ist. Für die Identifikation standen die Thermografiebilder von 40 Experimenten zur Verfügung. Die Ergebnisse der Identifikation von vier Experimenten wurden vorgestellt. Dabei konnte festgestellt werden, dass die Erweiterungen die Qualität der Ergebnisse begünstigen. Die Rekonstruktion des oberen Datenbereichs ist genauso zufriedenstellend wie die Adaption der  $y$ -Position der Daten an das simulierte Temperaturfeld. Bei allen vier Experimenten gibt es sowohl optisch als auch numerisch eine hohe Übereinstimmung zwischen Rekonstruktion und Daten. Die identifizierten Verläufe der Wärmestromdichten im Bohrkanal entsprechen den Erwartungen. Die primäre Wärmequelle hat die größte Intensität, die Intensität der sekundären Wärmequelle nimmt im Mittel mit größerer Entfernung zur Schneide ab. Die Ergebnisse deuten jedoch an, dass der Verlauf der sekundären Wärmequelle in der Realität möglicherweise nicht monoton ist.

Wie erwartet führt die Nichtlinearität der PDE zu einer langsameren Konvergenz der Parameteridentifikation. Während beim Fräsen mit einer linearen PDE im Durchschnitt 14,7 Hauptiterationen bis zur Konvergenz benötigt werden, liegt dieser Wert beim Bohren im Mittel bei 23,1.

Insgesamt ist die Parameteridentifikation eine geeignete und effiziente Methodik, um auch mit einem nichtlinearen Modell Wärmestromdichten von Zerspanprozessen zu bestimmen.



# Kapitel 6

## Minimierung von Formabweichungen beim Trockenfräsen

### Inhalt

---

6.1	Optimierung von quellspannungsbedingten Deformationen . . . . .	114
6.1.1	Quellspannungsmodell . . . . .	115
6.1.2	Optimierungskonzept . . . . .	118
6.1.3	Mathematische Formulierung der Optimierungsprobleme . .	120
6.1.4	Diskretisierung und Modellvalidierung . . . . .	124
6.1.5	Numerische Ergebnisse . . . . .	126
6.2	Optimierung der Gesamtformabweichung . . . . .	133
6.2.1	Kinematisches Modell der Gesamtformabweichungen . . . .	133
6.2.2	Mathematische Formulierung der Optimierungsprobleme . .	135
6.2.3	Numerische Ergebnisse . . . . .	136
6.2.4	Vergleich zu Messungen realer Experimente . . . . .	140
6.3	Fazit der Minimierung von Formabweichungen . . . . .	141

---

Die zentrale Aufgabenstellung dieser Arbeit ist die Minimierung von Formabweichungen beim Trockenfräsen. Das Hybridmodell (siehe Kapitel 3.1) wurde im SPP 1480 [13, 38] entwickelt, um Vorhersagen über Deformationen treffen zu können. Im Hybridmodell werden in zwei Sub-Modellen zwei verschiedene Effekte des Fräsens abgebildet. In Sub-Modell 1 werden die Formabweichungen, die aufgrund einer Änderung des Eigenspannungszustands auftreten, durch das Einbringen von Quellspannungen in eine randnahe Schicht des diskretisierten Bauteils modelliert. Verformungen dieser Art wirken größtenteils global auf das Bauteil. Es ist zweckmäßig, eine spannungsarme Bearbeitung anzustreben, da eine Kompensation durch Anpassung des Fräspfades nur lokal erfolgt und wenig Einfluss auf die Torsion eines Bauteils oder Formabweichungen auf der Rückseite

hat. Anders verhält es sich bei Sub-Modell 2, in welchem die lokalen thermischen Dehnungen und der daraus resultierende ungleichmäßige Materialabtrag modelliert wird. Hier spielt die Reduzierung der eingebrachten thermischen Energie durch Anpassung von Vorschubgeschwindigkeit, Schnittgeschwindigkeit und Eingriffsbreite eine zweitrangige Rolle. Da die lokalen thermischen Dehnungen größtenteils orthogonal zur Oberfläche erfolgen, birgt eine Adaption der Schnitttiefe das größte Kompensationspotential. Eine Aufteilung der Minimierung der Formabweichung in zwei Optimierungsphasen (wie in Abb. 3.6 in Kapitel 3.2 beschrieben) scheint deshalb sinnvoll. Dies gilt insbesondere, wenn zudem die Komplexität und die damit einhergehende Rechenzeit in Betracht gezogen wird. Die Aufteilung in zwei Optimierungsphasen erfordert lediglich eine einmalige Simulation mit dem Modell der thermisch bedingten Verschiebungen. Diese Berechnung ist sehr rechenintensiv und benötigt auf einem modernen Rechner bereits mehrere Tage. In einer Optimierung, in der das thermische Modell Nebenbedingung ist, sind eine Vielzahl von Modellberechnungen nötig. Diese äußerst rechenintensive Optimierung wird durch die Aufteilung in zwei Optimierungsphasen vermieden.

In diesem Kapitel werden zunächst die quellspannungsbedingten Deformationen minimiert, indem hinsichtlich einer spannungsarmen Bearbeitung bzw. einer Bearbeitung, in der sich Spannungen ausgleichen können, optimiert wird. Dabei werden verschiedene Maschinenparameter sowie verschiedene Frässtrategien berücksichtigt. Anschließend werden die verbliebenen Formabweichungen aus der ersten Optimierung mit den thermischen Dehnungen superponiert. Auf Grundlage dieser Verschiebungen wird dann die Kinematik des Werkzeugs angepasst, indem die Schnitttiefe optimiert wird. Das Kapitel schließt mit einem Vergleich zwischen den numerischen Ergebnissen und Messungen realer Experimente.

Dieses Kapitel orientiert sich an der Projektveröffentlichung [51].

## **6.1 Optimierung von quellspannungsbedingten Deformationen**

Die Deformationen, die aufgrund einer Änderung des Eigenspannungszustands hervorgerufen werden, treten überwiegend erst nach dem Ausspannen des Werkstücks auf. Das liegt daran, dass das Bauteil zunächst in seiner ursprünglichen Form gehalten wird. Nach dem Ausspannen treten dann Biegungen und Torsion auf, um ein Spannungsgleichgewicht herzustellen. In der Simulation wird dieser Prozess durch das Einbringen von Quellspannungen in eine randnahe Zone modelliert. Die Berechnung der daraus resultierenden Deformationen erfolgt mit den Lamé-Navier-Gleichungen. Ein Zusammenhang zwischen Prozessgrößen und Quellspannungen besteht durch das Regressionsmodell (3.4) aus Kapitel 3.1.1. Dies ermöglicht die Optimierung der Maschinenparameter Vorschubgeschwindigkeit, Schnittgeschwindigkeit und Eingriffsbreite. Zusätzlich werden verschie-

dene Frässtrategien betrachtet. Dabei wird die Anzahl, Anordnung und Reihenfolge der Fräsbahnen variiert.

### 6.1.1 Quellspannungsmodell

Das Regressionsmodell der Quellspannungen, welches als Teil des Sub-Modell 1 (siehe Abb. 3.3 in Kapitel 3.1.1) entwickelt wurde, ermöglicht es, einem Fräsprozess äquivalente Quellspannungen zuzuordnen. Aus diesen lassen sich durch Lösen einer PDE quellspannungsbedingte Deformationen berechnen. Im Folgenden wird der mathematische Zusammenhang beschrieben.

Sei zunächst das Tensorfeld  $\boldsymbol{\sigma}_{\phi, \text{ges}} : \Omega \rightarrow \mathbb{R}^{3 \times 3}$  gegeben, das eine Gesamtspannung in einem Gebiet  $\Omega$  beschreibt. Die Gleichgewichtsbedingung der linearen Elastizitätstheorie [87] setzt die Gesamtspannung mit einer Volumenkraft  $f_V : \Omega \rightarrow \mathbb{R}^3$  in Verbindung:

$$-\operatorname{div} \boldsymbol{\sigma}_{\phi, \text{ges}} = f_V$$

Für die spätere Diskretisierung wird diese Gleichung in die schwache Formulierung überführt. Dafür werden beide Seiten der Gleichung mit einer Testfunktion  $v_\phi$ , dessen Funktionenraum später festgelegt wird, multipliziert und über  $\Omega$  integriert:

$$-\int_{\Omega} \operatorname{div} \boldsymbol{\sigma}_{\phi, \text{ges}} \cdot v_\phi \, dx = \int_{\Omega} f_V \cdot v_\phi \, dx$$

Der Gaußsche Integralsatz liefert anschließend:

$$\int_{\Omega} \boldsymbol{\sigma}_{\phi, \text{ges}} : \nabla v_\phi \, dx - \int_{\Gamma} \boldsymbol{\sigma}_{\phi, \text{ges}} \vec{n} \cdot v_\phi \, ds = \int_{\Omega} f_V \cdot v_\phi \, dx$$

Dabei ist das Frobenius-Skalarprodukt zweier Tensorfelder  $\mathbf{A}, \mathbf{B} : \Omega \rightarrow \mathbb{R}^{n \times m}$  durch

$$\mathbf{A} : \mathbf{B} := \sum_{i=1}^n \sum_{j=1}^m A_{ij} B_{ij}$$

definiert. Die Annahme, dass auf der Werkstückoberfläche  $\Gamma$  keine externen Lasten wirken, kann mit dem äußeren normalen Vektor  $\vec{n}$  durch die Neumann-Bedingung

$$\boldsymbol{\sigma}_{\phi, \text{ges}} \vec{n} = 0 \quad \text{auf } \Gamma \tag{6.1}$$

beschrieben werden, was die Gleichung zu

$$\int_{\Omega} \boldsymbol{\sigma}_{\phi, \text{ges}} : \nabla v_\phi \, dx = \int_{\Omega} f_V \cdot v_\phi \, dx$$

vereinfacht. Durch das Hookesche Gesetz lassen sich die Gesamtspannungen durch Verschiebungen  $u_\phi : \Omega \rightarrow \mathbb{R}^3$  ersetzen:

$$\int_{\Omega} \mathbb{E} \frac{1}{2} (\nabla u_\phi + (\nabla u_\phi)^T) : \nabla v_\phi \, dx = \int_{\Omega} f_V \cdot v_\phi \, dx$$

Dabei ist  $\mathbb{E} \in \mathbb{R}^{3 \times 3 \times 3 \times 3}$  der materialabhängige Elastizitätstensor vierter Stufe. Nimmt man einen homogenen und isotropen Körper an, vereinfacht sich die Gleichung aufgrund der besonderen Struktur von  $\mathbb{E}$  zu

$$\begin{aligned} & \int_{\Omega} \left( \lambda_L \operatorname{div} u_\phi \mathbf{I} + \mu_L (\nabla u_\phi + (\nabla u_\phi)^T) \right) : \nabla v_\phi \, dx = \int_{\Omega} f_V \cdot v_\phi \, dx \\ \Leftrightarrow & \lambda_L \int_{\Omega} \operatorname{div} u_\phi \operatorname{div} v_\phi \, dx + \mu_L \int_{\Omega} \nabla u_\phi : \nabla v_\phi \, dx + \mu_L \int_{\Omega} (\nabla u_\phi)^T : \nabla v_\phi \, dx = \int_{\Omega} f_V \cdot v_\phi \, dx. \end{aligned} \quad (6.2)$$

Dabei ist  $\mathbf{I} \in \mathbb{R}^{3 \times 3}$  die Einheitsmatrix. Die Variablen  $\lambda_L$  und  $\mu_L$ <sup>1</sup> sind die skalaren Lamé-Konstanten, welche im isotropen Fall durch Poissonzahl  $\nu$  und Elastizitätsmodul  $E$  (siehe Tab. 3.1) ausgedrückt werden können:

$$\lambda_L = \frac{\nu E}{(1 + \nu)(1 - 2\nu)}, \quad \mu_L = \frac{E}{2(1 + \nu)}$$

Gleichung (6.2) ist die schwache Form der stationären, dreidimensionalen Lamé-Navier-Gleichungen<sup>2</sup> [63]. Die Volumenkraft  $f_V$  kann sich aus mehreren Anteilen zusammensetzen. Im Folgenden seien diese durch das Quellspannungstensorfeld<sup>3</sup>  $\sigma_\phi : \Omega \rightarrow \mathbb{R}^{3 \times 3}$  und einen allgemeinen Teil  $f_G$  gegeben:

$$f_V = f_G - \operatorname{div} \sigma_\phi$$

Der allgemeine Teil kann beispielsweise das Schwerfeld der Gravitation beinhalten. Insgesamt ergibt sich so ein Zusammenhang zwischen Quellspannungen und Verschiebungen:

$$\begin{aligned} & \lambda_L \int_{\Omega} \operatorname{div} u_\phi \operatorname{div} v_\phi \, dx + \mu_L \int_{\Omega} \nabla u_\phi : \nabla v_\phi \, dx + \mu_L \int_{\Omega} (\nabla u_\phi)^T : \nabla v_\phi \, dx \\ & = \int_{\Omega} f_G \cdot v_\phi \, dx - \int_{\Omega} \operatorname{div} \sigma_\phi \cdot v_\phi \, dx \end{aligned} \quad (6.3)$$

In der Theorie kann unter bestimmten Voraussetzungen die Existenz einer eindeutigen Lösung der PDE (6.3) gezeigt werden. Insbesondere müssen dafür geeignete Randbedingungen formuliert werden. So ist das Unterbinden einer Verschiebung auf einer

<sup>1</sup>Die Variable  $\mu_L$  wird auch Schubmodul genannt.

<sup>2</sup>Die Lamé-Navier-Gleichungen sind auch als Navier-Cauchy-Gleichungen oder lineare Elastizitätsgleichungen bekannt.

<sup>3</sup>Hierbei wird auf die Indexschreibweise  $\sigma_{\phi, \text{quell}}$  verzichtet, da im Folgenden ausschließlich Spannungen in Form von Quellspannungen betrachtet werden.

Randfläche mit positivem Oberflächenmaß notwendig, damit eine eindeutige Lösung existiert [16]. Beim numerischen Lösen sollen dem Körper  $\Omega$  die größtmöglichen Freiheiten zur Verformung gegeben werden und gleichzeitig Translationen und Rotationen unterbunden werden. Dazu muss das Werkstück in sechs Freiheitsgraden in mindestens drei Punkten, die nicht auf einer Linie liegen, eingeschränkt werden. Das Gebiet  $\Omega$  sei nachfolgend durch ein Werkstück mit den Maßen  $\Omega = [0; 145,2] \times [0; 145,2] \times [0; 11]$  mm<sup>3</sup> gegeben. Alle Bedingungen können damit durch Dirichlet-Bedingungen mit einem hinreichend kleinem  $\varepsilon > 0$  formuliert werden:

$$\begin{aligned} u_\phi &= 0 \quad \text{auf } U_\varepsilon((0 \ 145,2 \ 0)^T) \cap \Gamma \\ u_{\phi,x} = u_{\phi,z} &= 0 \quad \text{in } (0 \ 0 \ 0)^T \\ u_{\phi,z} &= 0 \quad \text{in } (145,2 \ 0 \ 0)^T \end{aligned} \quad (6.4)$$

Mit diesen Bedingungen lässt sich auch der Funktionenraum für  $u_\phi$  und  $v_\phi$  festlegen:

$$u_\phi, v_\phi \in \left\{ \begin{aligned} w_\phi \in H^1(\Omega; \mathbb{R}^3) : w_\phi|_{U_\varepsilon((0 \ 145,2 \ 0)^T) \cap \Gamma} &= 0, \\ w_{\phi,z}|_{(0 \ 0 \ 0)^T \cup (145,2 \ 0 \ 0)^T} &= 0, \\ w_{\phi,x}|_{(0 \ 0 \ 0)^T} &= 0 \end{aligned} \right\}$$

Da das Modell aus den Gleichungen (6.3) und (6.4) im Folgenden nur die quellspannungsbedingten Verformungen abbilden soll, wird der allgemeine Teil der Volumenkraft  $f_G$  nicht weiter berücksichtigt. Das allgemeine Verschiebungsfeld  $u_\phi$  wird so zum quellspannungsbedingten Verschiebungsfeld  $u_{\phi,\text{quell}} : \Omega \rightarrow \mathbb{R}^3$ . Zusammen mit dem Regressionsmodell (3.4), der Transformation (3.5) und dem Quellspannungstensorfeld (3.6) ergibt das den gesuchten Zusammenhang zwischen Maschinenparameter und den Verschiebungen aufgrund einer Änderung des Eigenspannungszustands:

$$\begin{aligned} &\lambda_L \int_\Omega \operatorname{div} u_{\phi,\text{quell}} \operatorname{div} v_\phi \, dx + \mu_L \int_\Omega \nabla u_{\phi,\text{quell}} : \nabla v_\phi \, dx + \mu_L \int_\Omega (\nabla u_{\phi,\text{quell}})^T : \nabla v_\phi \, dx \\ &= - \int_\Omega \operatorname{div} \boldsymbol{\sigma}_\phi(a_e, v_c, v_f) \cdot v_\phi \, dx \end{aligned} \quad (6.5)$$

Dabei seien im Folgenden  $a_e, v_c, v_f \in \mathbb{R}^{n_b}$  die Vektoren der Maschinenparameter, welche die skalaren Maschinenparameter der  $n_b$  Fräsbahnen enthalten.

Aus der Diskretisierung von (6.5) mit der FEM (siehe Kapitel 2.3) resultiert ein Gleichungssystem:

$$\mathbf{A} u_{\text{quell}} = \tilde{\mathbf{b}}(\boldsymbol{\sigma}) = \mathbf{b}(a_e, v_c, v_f) \quad (6.6)$$

Hierbei bezeichnet  $u_{\text{quell}} \in \mathbb{R}^{n_s}$  den Vektor der diskretisierten Werte der quellspannungsbedingten Verschiebungen in geeignet sortierter Form und  $\mathbf{A} \in \mathbb{R}^{n_s \times n_s}$  die Systemmatrix, die von der Wahl der finiten Elemente abhängt. Die rechte Seite  $\tilde{\mathbf{b}} : \mathbb{R}^{n_\sigma} \rightarrow \mathbb{R}^{n_s}$  ist

eine nichtlineare Funktion, die ebenfalls aus der Diskretisierung resultiert. Der Vektor  $\sigma \in \mathbb{R}^{n_\sigma}$  enthält die diskretisierten Werte von  $\sigma_\phi$ . Eine weitere Formulierung, welche für die Optimierung verwendet wird, ergibt sich mit der rechten Seite  $\mathbf{b} : \mathbb{R}^{n_z} \rightarrow \mathbb{R}^{n_s}$ , die sowohl durch die FEM als auch durch das Regressionsmodell und die Interpolation (siehe Kapitel 3.1.1) bestimmt wird.  $\mathbf{A}$  und  $\mathbf{b}$  bzw.  $\tilde{\mathbf{b}}$  werden aufgrund ihrer Komplexität und Größe nicht weiter spezifiziert. Das diskretisierte Gleichungssystem (6.6) wird für die Optimierung als Modell verwendet, indem es im SAND-Ansatz als Nebenbedingung dem Optimierungsproblem hinzugefügt wird.

## 6.1.2 Optimierungskonzept

Die Optimierung mit Sub-Modell 2 hat zum Ziel, Frässtrategien und Maschinenparameter zu optimieren, um die quellspannungsbedingten Formabweichungen zu minimieren. Zum besseren Verständnis werden folgende zwei Begriffe eingeführt:

Die **Bearbeitungsstrategie** fasst verschiedene Variablen einer Gesamtbearbeitung zusammen. Dies sind die Anzahl, Anordnung und Reihenfolge der einzelnen Fräsprozesse. Es gibt unzählige Bearbeitungsstrategien, um die gleiche Fertigung eines Bauteils zu erzielen. Zu beachten bei der Wahl einer Bearbeitungsstrategie ist, dass diese mit ihren Einzelprozessen auch die gesamte Bauteiloberfläche abdeckt und dass das Fräsen im Gleichlauf erfolgt.

Die **Prozessstrategie** fasst im Gegenzug die Maschinenparameter zusammen, die die einzelnen Fräsprozesse beschreiben. Zu einer Bearbeitungsstrategie gehören somit  $n_b$  Prozessstrategien. Da innerhalb eines Einzelprozesses die Maschinenparameter konstant bleiben, besteht die  $i$ -te Prozessstrategie aus den Maschinenparametern  $a_{e,i}$ ,  $v_{c,i}$  und  $v_{f,i}$ ,  $i = 1, \dots, n_b$  (siehe Kapitel 3.2).

Die Optimierung von Bearbeitungs- und Prozessstrategien beinhaltet wie in Abb. 6.1 dargestellt sowohl kontinuierliche als auch diskrete Optimierungsvariablen. Diskrete Variablen sind beispielsweise die Anzahl und Reihenfolge der Fräsprozesse innerhalb der Gesamtbearbeitung. Kontinuierliche Variablen sind die Maschinenparameter sowie die Richtung der Fräsbahnen, die durch einen Winkel beschrieben werden kann. Eine Bearbeitungsstrategie besteht größtenteils aus diskreten Größen, während sich eine Prozessstrategie ausschließlich aus kontinuierlichen Variablen zusammensetzt. Der Fokus dieser Arbeit liegt auf der Problemklasse der NLPs und dem Lösen dieser Probleme mit dem SAND-Ansatz. Die Optimierung der Bearbeitungsstrategien kann nicht ohne Weiteres mit dieser Methodik der kontinuierlichen Optimierung erfolgen. Die diskreten Variablen werden innerhalb dieser Arbeit untersucht, indem Bearbeitungsstrategien, die sich in der Anzahl, Anordnung und Reihenfolge voneinander unterscheiden, miteinander verglichen werden. Mögliche kontinuierliche Variablen einer Bearbeitungsstrategie sowie die Maschinenparameter der Prozessstrategien werden mit der nichtlinearen Optimierung im SAND-Ansatz optimiert. Die diskreten Größen werden während der kontinuierlichen Optimierung fixiert.

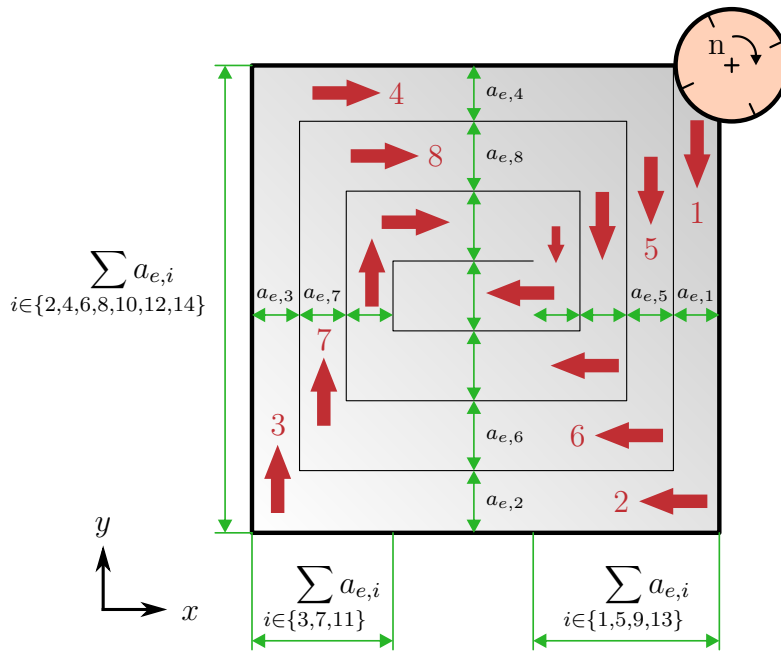


Abbildung 6.1: Bearbeitungsstrategie mit diskreten Größen (rot) und den kontinuierlichen Variablen (grün) der Prozessstrategien

### Die verschiedenen Bearbeitungsstrategien

Für die Analyse der diskreten Variablen werden verschiedene Bearbeitungsstrategien ausgewählt und miteinander verglichen. Um die Anzahl dieser in einem praktikablen Umfang zu halten, wurde bei der Auswahl auf das ingenieurwissenschaftliche Know-How der Projektkollegen zurückgegriffen [51]. Abbildung 6.2 zeigt die drei Grundtypen der Bearbeitungsstrategien, die die Reihenfolge und Anordnung der einzelnen Fräsbahnen festlegen.

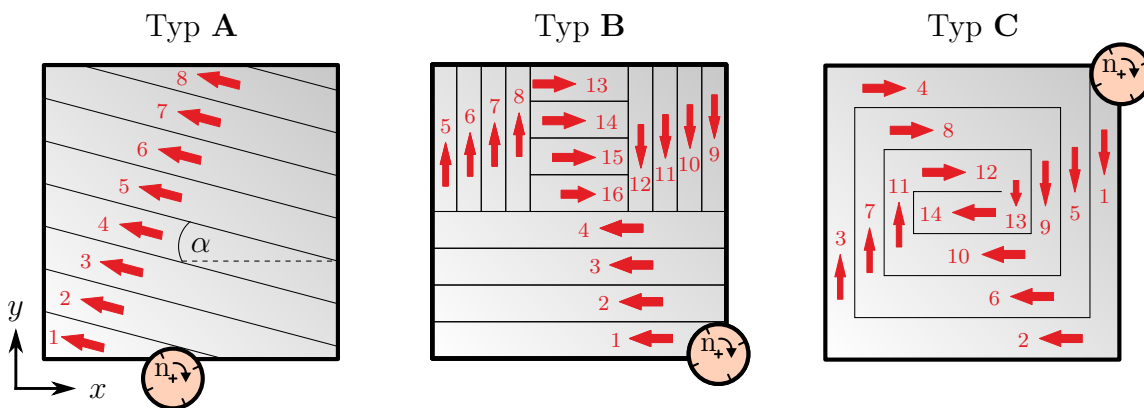


Abbildung 6.2: Die drei Grundtypen der berücksichtigten Bearbeitungsstrategien

- Typ **A** ist eine triviale Strategie paralleler Fräsbahnen, die um einen Winkel  $\alpha$  relativ zum Werkstück rotiert werden können. So wird ermöglicht, dass die Hauptspannungsrichtung mit den Koordinatenachsen zusammenfällt. Dies kann die Minimierung der Formabweichungen begünstigen.
- Typ **B** setzt sich aus vier Bereichen unterschiedlicher Fräsrichtungen zusammen. Die Richtungen sind dabei parallel zu den Bauteilkanten. Die Einteilung der Oberfläche in die vier Bereiche kann dabei flexibel gestaltet werden. Dieser Grundtyp soll der Optimierung ein hohes Maß an Freiheit einräumen.
- Typ **C** orientiert sich an einer ökonomischen Bearbeitung, indem die Gesamtlänge des Fräspfades möglichst kurz gewählt wird und eine Bearbeitung ohne Unterbrechung ermöglicht wird. Diese Art der Bearbeitung ist in der Industrie weit verbreitet. Es ist davon auszugehen, dass durch Anpassen der Maschinenparameter eine Reduzierung der Formabweichung erreicht werden kann.

Diese drei Grundtypen werden in der Anzahl an Fräsbahnen variiert, wodurch auch die mittlere Eingriffsbreite variiert. Grundtyp **A** erhält  $n_b = 12, 16$  oder  $24$  Pfade, die Typen **B** und **C**  $n_b = 24, 32$  oder  $40$  Pfade. Insgesamt ergeben sich so neun verschiedene Strategien, die sowohl hinsichtlich der Verformung als auch der Bearbeitungseffizienz miteinander verglichen werden.

### 6.1.3 Mathematische Formulierung der Optimierungsprobleme

Für die Optimierung der neun Bearbeitungsstrategien wird die Problemstellung als NLP formuliert, bei dem das diskretisierte Modell als Nebenbedingung berücksichtigt wird. Die Optimierungsvariablen sind neben den Zuständen die Prozessgrößen und bei Typ **A** zusätzlich der Winkel  $\alpha \in \mathbb{R}$ , der den Rotationswinkel der Fräsbahnen beschreibt. Die Prozessgrößen bestehen aus den Vektoren der Maschinenparameter:  $a_e, v_c, v_f \in \mathbb{R}^{n_b}$ . Um den Vektor der quellspannungsbedingten Formabweichungen  $u_{\text{quell}} \in \mathbb{R}^{n_s}$  zu minimieren, wird eine quadratische Zielfunktion gewählt. Zusammen mit dem Modell (6.6) ergibt sich:

$$\begin{aligned} \min_{X \in \mathbb{R}^{n_x}} F(u_{\text{quell}}) &:= \frac{1}{n_s} \sum_{i=1}^{n_s} u_{\text{quell},i}^2 \\ \text{unter } \mathbf{A} u_{\text{quell}} - \mathbf{b}(a_e, v_c, v_f, (\alpha)) &= 0 \end{aligned} \quad (6.7)$$

Die Optimierungsvariable ist dabei durch  $X := \left( u_{\text{quell}}^T, a_e^T, v_c^T, v_f^T, (\alpha) \right)^T \in \mathbb{R}^{n_x}$  mit  $n_x = n_s + 3n_b(+1)$  gegeben, wobei die Variable  $\alpha$  nur bei Typ **A** Verwendung findet. Zu beachten ist, dass  $\mathbf{b}(a_e, v_c, v_f, (\alpha))$  und damit die Nebenbedingungen durch das Regressionsmodell und die Spline-Interpolation des Spannungsfeldes bzgl. der Maschinenparameter hochgradig nichtlinear sind.

In diesem ersten Optimierungsschritt treten globale Verformungen auf, die sich im zweiten Schritt mit der Optimierung der Schnitttiefe nur teilweise kompensieren lassen. Dies



## 6.1 Optimierung von quellspannungsbedingten Deformationen

ist insbesondere bei einer Torsion des Bauteils der Fall. Vorangegangene Optimierungstests konnten zeigen, dass sich eine quellspannungsbedingte Torsion durch das Anpassen der Maschinenparameter vermeiden lässt. Deshalb soll eine zusätzliche Nebenbedingung die Torsion des Bauteils verhindern. Die  $z$ -Verschiebung des Bauteils ist in drei Ecken (siehe (6.4)) gelagert, sodass die Torsion vollständig durch die  $z$ -Verschiebung in der vierten, ungelagerten Ecke quantifiziert wird. Diese Verschiebung sei o. B. d. A. der erste Eintrag  $u_{\text{quell},1}$  des Verschiebungsvektors  $u_{\text{quell}}$ . Damit lässt sich die Boxschränke

$$-0,5 \leq u_{\text{quell},1} \leq 0,5 \text{ [\mu m]} \quad (6.8)$$

formulieren, welche eine Toleranz von  $\pm 0,5 \mu\text{m}$  für die Torsion erlaubt.

Um eine Extrapolation der Regressionsmodelle (3.4) und (3.7) zu vermeiden, müssen die Maschinenparameter beschränkt werden. Die Gültigkeitsbereiche der Eingriffsbreite  $a_e \in \mathbb{R}^{n_b}$ , der Schnittgeschwindigkeit  $v_c \in \mathbb{R}^{n_b}$  und der Vorschubgeschwindigkeit  $v_f \in \mathbb{R}^{n_b}$  sind durch die Boxschränken

$$\begin{aligned} 0,95 &\leq a_{e,i} \leq 12,5 \text{ [mm]} \\ 255 &\leq v_{c,i} \leq 373 \text{ [m/min]} \\ 2000 &\leq v_{f,i} \leq 2800 \text{ [mm/min]}, \quad i = 1, \dots, n_b \end{aligned} \quad (6.9)$$

gegeben.

Zusätzlich muss gewährleistet werden, dass die Bearbeitungsstrategien von Typ **A**, **B** und **C** mit ihren einzelnen Prozessstrategien die gesamte Bauteiloberfläche abdecken, wie es in Abb. 6.2 dargestellt wird. Dies kann durch Nebenbedingungen an die Eingriffsbreiten realisiert werden. Für Typ **A** muss die Summe aller Pfade unter Berücksichtigung des Winkels  $\alpha$  die Bauteillänge von 145,2 mm erreichen. Dabei genügt es aufgrund der quadratischen Form der Bauteiloberfläche Winkel zwischen  $0^\circ$  und  $45^\circ$  zu betrachten. Bei Typ **B** muss die Summe der horizontalen Fräsbahnen die Bauteillänge abdecken. Die Summe der vertikalen Bahnen kann dabei zwischen 0 mm und der Bauteillänge variieren. Gleiches gilt für die horizontalen und vertikalen Fräsbahnen von Typ **C** (siehe

Abb. 6.1). Mathematisch lassen sich diese Bedingungen wie folgt zusammenfassen:

Typ	Nebenbedingungen	Einheit
<b>A</b>	$\sum_{i=1}^{n_b} a_{e,i} = 145,2 (\sin \alpha + \cos \alpha)$	mm
	$0 \leq \alpha \leq 45$	°
<b>B</b>	$\sum_{i=1}^{\frac{1}{4}n_b} a_{e,i} + \sum_{i=\frac{3}{4}n_b+1}^{n_b} a_{e,i} = 145,2$	mm
	$0 \leq \sum_{i=\frac{1}{4}n_b+1}^{\frac{3}{4}n_b+1} a_{e,i} \leq 145,2$	mm
<b>C</b>	$\sum_{i=1}^{\frac{n_b}{2}} a_{e,2i} = 145,2$	mm
	$0 \leq \sum_{i=1}^{\frac{n_b}{2}} a_{e,2i-1} \leq 145,2$	mm

(6.10)

Das gesamte Optimierungsproblem setzt sich in der SAND-Formulierung aus (6.7) - (6.10) zusammen, wobei  $H$  und  $G$  entsprechend der angeführten Gleichungen bzw. Ungleichungen definiert werden. Mathematisch ist dieses Problem den Problemen der Parameteridentifikation aus den Kapiteln 4 und 5 sehr ähnlich, da es eine quadratische Zielfunktion sowie nichtlineare Nebenbedingungen aus der Diskretisierung einer PDE enthält. Die Boxschränken in (6.8) und (6.9) sowie die zweite Ungleichung von Typ **A** in (6.10) werden separat behandelt, sodass sich für Typ **A**  $m_h = n_s + 1$  und  $m_g = 0$  sowie für Typ **B** und **C**  $m_h = n_s + 1$  und  $m_g = 2$  ergeben.<sup>4</sup>

Auch bei dieser Problemstellung sollen dem Lösungsalgorithmus Ableitungen der Zielfunktion und der Nebenbedingungen zur Verfügung gestellt werden. Der Gradient der Zielfunktion in (6.7) hat die Form:

$$\nabla F(X) = \begin{pmatrix} \frac{2}{n_s} u_{\text{quell},1} \\ \frac{2}{n_s} u_{\text{quell},2} \\ \vdots \\ \frac{2}{n_s} u_{\text{quell},n_s} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^{n_x}$$

<sup>4</sup>Durch Umformen einer Ungleichungsnebenbedingung mit beidseitigen Schranken in die Form (2.5) verdoppelt sich die Anzahl.

Die ersten  $n_s$  Zeilen der Jacobi-Matrix von  $H$  sind durch

$$\begin{pmatrix} \nabla H_1(X)^T \\ \nabla H_2(X)^T \\ \vdots \\ \nabla H_{n_s}(X)^T \end{pmatrix} = \left( \underbrace{\mathbf{A}}_{\in \mathbb{R}^{n_s \times n_s}} \quad \underbrace{-\nabla_Z \mathbf{b}}_{\in \mathbb{R}^{n_s \times n_z}} \right) \in \mathbb{R}^{n_s \times n_x}$$

gegeben. Dabei wird die Jacobi-Matrix  $\nabla_Z \mathbf{b}$  bzgl. der Entscheidungsvariablen  $Z = (a_e^T, v_c^T, v_f^T, (\alpha))^T \in \mathbb{R}^{n_z}$  numerisch durch finite Differenzen bestimmt, da ein analytisches Differenzieren durch die Verwendung von Fremdsoftware bei der Interpolation schwierig ist. Hinzu kommt noch eine weitere Zeile  $\nabla H_{n_s+1}(X)^T \in \mathbb{R}^{n_x}$ , die abhängig von der gewählten Bearbeitungsstrategie ist:

$$\begin{aligned} \mathbf{A}: \nabla H_{n_s+1}(X)^T &= (0 \cdots 0 \ 1 \cdots & \cdots 1 \ 0 \cdots 0 \ 0) \\ \mathbf{B}: \nabla H_{n_s+1}(X)^T &= (0 \cdots 0 \ 1 \cdots 1 \ 0 \cdots \cdots 0 \ 1 \cdots 1 \ 0 \cdots 0) \\ \mathbf{C}: \nabla H_{n_s+1}(X)^T &= \underbrace{(0 \cdots 0 \ 0 \ 1 \ 0 \ 1 \cdots)}_{\in \mathbb{R}^{n_s}} \underbrace{\cdots \cdots \cdots 0 \ 1 \ 0 \cdots 0}_{\in \mathbb{R}^{n_b}} \underbrace{\cdots \cdots 0 \ 1 \ 0 \cdots 0}_{\in \mathbb{R}^{2n_b}} \end{aligned}$$

Die Jacobi-Matrix der Ungleichungsnebenbedingungen  $\nabla G(X) \in \mathbb{R}^{2 \times n_x}$  ist für Typ **B** und **C** durch

$$\begin{aligned} \mathbf{B}: \nabla G(X) &= \begin{pmatrix} 0 \cdots 0 \ 0 \cdots 0 \ 1 \cdots & \cdots 1 \ 0 \cdots 0 \ 0 \cdots 0 \\ 0 \cdots 0 \ 0 \cdots 0 \ -1 \cdots & \cdots -1 \ 0 \cdots 0 \ 0 \cdots 0 \end{pmatrix} \\ \mathbf{C}: \nabla G(X) &= \begin{pmatrix} 0 \cdots 0 \ 1 \ 0 \ 1 \ 0 \cdots & \cdots 1 \ 0 \ 0 \cdots 0 \\ 0 \cdots 0 \ -1 \ 0 \ -1 \ 0 \cdots & \cdots -1 \ 0 \ 0 \cdots 0 \end{pmatrix} \\ &\underbrace{\hspace{10em}}_{\in \mathbb{R}^{n_s}} \underbrace{\hspace{10em}}_{\in \mathbb{R}^{n_b}} \underbrace{\hspace{10em}}_{\in \mathbb{R}^{2n_b}} \end{aligned}$$

gegeben. Typ **A** kommt ohne Ungleichungsnebenbedingungen aus. Die Hesse-Matrix der Lagrange-Funktion hat aufgrund der Linearität der Nebenbedingungen (6.10) die Form

$$\nabla_X^2 L(X, \lambda) = \nabla^2 F(X) + \sum_{k=1}^{n_s} \lambda_k \nabla^2 H_k(X).$$

Die Zielfunktion  $F$  ist nicht von  $Z$  abhängig, weshalb ihre Hesse-Matrix eine Diagonal-

matrix ist:

$$\nabla^2 F(X) = \begin{pmatrix} \frac{2}{n_s} & & & 0 & \cdots & 0 \\ & \frac{2}{n_s} & & \vdots & & \vdots \\ & & \ddots & & & \\ & & & \frac{2}{n_s} & \vdots & \\ 0 & \cdots & & \cdots & 0 & \\ \vdots & & & & & \vdots \\ 0 & \cdots & & & \cdots & 0 \end{pmatrix} \in \mathbb{R}^{n_x \times n_x}$$

Die Hesse-Matrizen der Gleichungsnebenbedingungen sind unabhängig von  $u_{\text{quell}}$  und haben die Gestalt

$$\nabla^2 H_i(X) = \begin{pmatrix} 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ & & & 0 & \cdots & 0 \\ & & & \nabla_{Z_1 Z_1}^2 \mathbf{b}_i(Z) & \cdots & \nabla_{Z_1 Z_{n_z}}^2 \mathbf{b}_i(Z) \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & \nabla_{Z_{n_z} Z_1}^2 \mathbf{b}_i(Z) & \cdots & \nabla_{Z_{n_z} Z_{n_z}}^2 \mathbf{b}_i(Z) \end{pmatrix} \in \mathbb{R}^{n_x \times n_x}, \quad i = 1, \dots, n_s.$$

Auch hier werden die zweiten Ableitungen  $\nabla_{Z_k Z_l}^2 \mathbf{b}_i(Z)$ ,  $k, l = 1, \dots, n_z$ ,  $i = 1, \dots, n_s$  durch finite Differenzen bestimmt, da ein analytisches Ableiten durch die Interpolation mit Fremdsoftware erschwert wird.

### 6.1.4 Diskretisierung und Modellvalidierung

Innerhalb dieses Abschnitts soll eine Validierung des Modells erfolgen, die dabei hilft, die Feinheit der Diskretisierung zu wählen. Das zu diskretisierende Gebiet  $\Omega = [0; 145,2] \times [0; 145,2] \times [0; 11] \text{ mm}^3$  beschreibt das Bauteil, welches ausgehend von 12 mm Höhe mit einer Schnitttiefe von  $a_p = 1 \text{ mm}$  bearbeitet wurde. Die Diskretisierung erfolgt durch die FEM mit quadratischen Lagrange-Elementen auf Hexaedern. Um das Modell zu validieren, wird zunächst ein konstantes Quellspannungstensorfeld mit fiktiven Werten gewählt:

$$\begin{aligned} \boldsymbol{\sigma}_\phi(x, y, z) &= \begin{pmatrix} 3 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ N/mm}^2 \quad \text{für } z \geq 10 \text{ mm} \\ \boldsymbol{\sigma}_\phi(x, y, z) &= 0 \quad \text{N/mm}^2 \quad \text{sonst} \end{aligned} \quad (6.11)$$

## 6.1 Optimierung von quellspannungsbedingten Deformationen

Dieses wird anschließend diskretisiert und in  $\sigma \in \mathbb{R}^{n_\sigma}$  geeignet sortiert zusammengefasst. Durch Lösen des Gleichungssystems (6.6) können die resultierenden quellspannungsbedingten Verschiebungen  $u_{\text{quell}}$  berechnet werden. Mit einer Approximation der verformten Oberfläche durch eine quadratische Funktion können wie in Kapitel 3.1.1 über die Plattentheorie die Quellspannungen berechnet werden. Diese sollten bei validem Modell mit feiner werdender Diskretisierung gegen die konstanten Quellspannungen des Tensorfeldes (6.11) konvergieren.

Zunächst konnte festgestellt werden, dass die besten Approximationen des Tensorfeldes erzielt werden, wenn die Anzahl der diskreten Segmente der  $z$ -Koordinate ein Vielfaches von 11 entspricht. Dies ist nachvollziehbar, da die Höhe der Schicht, in der die Quellspannungen eingebracht werden, ein Elftel der Plattendicke entspricht ( $l_z = 11$  mm,  $z_0 = 1$  mm). Für die Validierung wurde diese Größe deshalb auf den Wert 11 fixiert und die Anzahl der diskreten Segmente der  $x$ - und  $y$ -Koordinate zwischen 3 und 26 variiert. Dies führt zu einer Anzahl von 3381 bis 193 821 diskretisierten Freiheitsgraden. Abbildung 6.3 zeigt die mit der Plattentheorie bestimmten Quellspannungen bei Verfeinerung des Gitters. Es lässt sich eine schnelle Konvergenz beobachten. Ab einer Anzahl von elf Segmenten für die Diskretisierung jeder Koordinatenrichtung ist die Abweichung kleiner als  $0,01$  N/mm<sup>2</sup> bzw. 1 %.

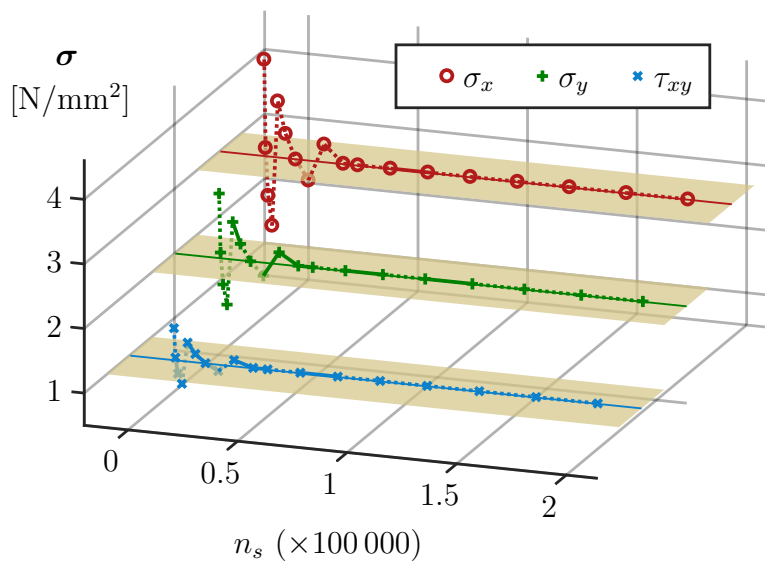


Abbildung 6.3: Konvergenz der Quellspannungen bei Gitterverfeinerung

Zusätzlich ist zu beachten, dass die Oberfläche  $\Gamma_O$  in Fräsbahnen unterteilt wird. Um dies geeignet abbilden zu können, werden die  $x$ - und  $y$ -Koordinate für die Optimierung mit 16 Segmenten diskretisiert. Zusammen mit der Diskretisierung der  $z$ -Koordinate durch elf Segmente ergeben sich 75 141 Freiheitsgrade. Abbildung 6.4 zeigt das Gitter, das durch die Diskretisierung mit den genannten Werten entsteht.

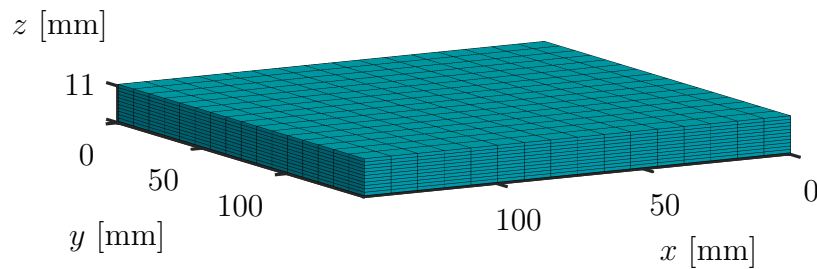


Abbildung 6.4: Gitter für die Minimierung der quellspannungsbedingten Formabweichungen

### 6.1.5 Numerische Ergebnisse

Zur vereinfachten Darstellung wird die folgende Variable eingeführt:

$$u_{\text{quell}}^{\max} := \max_{i=1, \dots, n_s} |u_{\text{quell}, i}|$$

In Kapitel 6.1.2 wurde bereits das Optimierungskonzept zur Minimierung von quellspannungsbedingten Formabweichungen vorgestellt. Dieses beinhaltet die Optimierung von insgesamt neun Bearbeitungsstrategien als NLP in der SAND-Formulierung. Die mathematische Problemstellung wurde dafür in Kapitel 6.1.3 formuliert. In diesem Abschnitt werden die Ergebnisse der numerischen Optimierung von drei Bearbeitungen im Detail und eine Gesamtübersicht präsentiert. Die Berechnungen wurden auf dem Rechner aus Tab. 5.2 durchgeführt. Um die Ergebnisse besser bewerten zu können wird zunächst eine Bearbeitungsstrategie als Vergleichsgröße vorgestellt.

#### Triviale Bearbeitungsstrategie als Vergleichsgröße

Als Vergleich für die Optimierungsergebnisse soll eine triviale Bearbeitungsstrategie dienen. Diese ist vom Typ **A** mit einem Bahnwinkel von  $\alpha = 0^\circ$  und  $n_b = 12$  Einzelprozessen sowie konstanten Maschinenparametern:

$$a_{e,i} = \frac{145.2}{n_b} = 12,1 \text{ mm}, \quad v_{c,i} = 300 \text{ m/min}, \quad v_{f,i} = 2400 \text{ mm/min}, \quad i = 1, \dots, n_b$$

Abbildung 6.5 zeigt die Fräsbahnen dieser Strategie sowie die daraus resultierenden quellspannungsbedingten Verschiebungen durch einmaliges Lösen des Modells (6.6). Die Verschiebungen wurden zur besseren Visualisierung mit dem Faktor 10 000 verstärkt zur

## 6.1 Optimierung von quellspannungsbedingten Deformationen

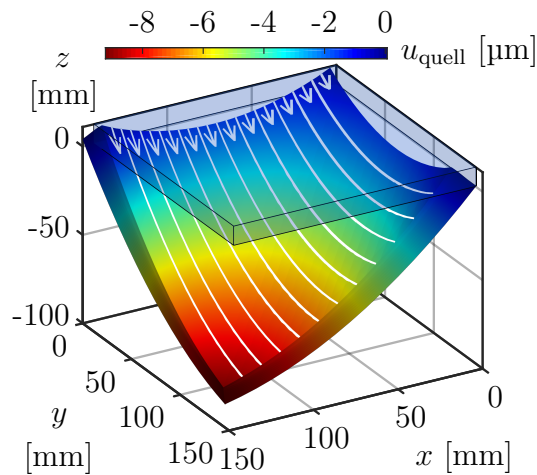


Abbildung 6.5: Vergleichsstrategie mit resultierender Deformation

Position des Bauteils addiert. Als Orientierung ist auch das Bauteil ohne Deformation dargestellt.

Die Abweichung dieser Bearbeitung ist vergleichsweise groß. Der Zielfunktionswert, der zur mittleren quadratischen Abweichung korrespondiert, beträgt  $F(X) \approx 7,8 \mu\text{m}^2$  und die maximale Abweichung von der unverformten Geometrie beträgt  $u_{\text{quell}}^{\text{max}} \approx 9,2 \mu\text{m}$ . Die maximale Abweichung tritt in der ungelagerten Ecke auf, was eine große Torsion des Bauteils zur Folge hat.

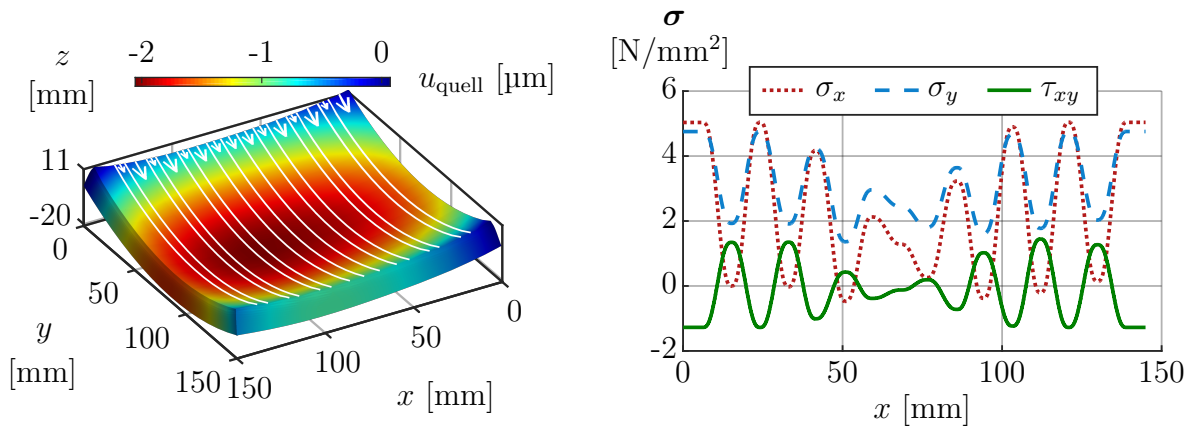
Die Optimierung dieser Strategie (Typ **A** mit  $n_b = 12$ ) scheitert, da eine Torsion des Bauteils nicht verhindert werden kann (vgl. Nebenbedingung (6.8)). Die wenigen Pfade ( $n_b = 12$ ) haben eine große durchschnittliche Eingriffsbreite zur Folge, was wiederum zu einer negativen Torsionsspannung in großen Teilen des Werkstücks führt. Die negativen Spannungen können nicht durch schmale Fräsbahnen und die damit einhergehenden positiven Torsionsspannungen kompensiert werden. Deshalb ist eine Erhöhung der Anzahl an Fräsbahnen sinnvoll.

### Typ A, 16 Bahnen

Für die Optimierung von Typ **A** mit  $n_b = 16$  Fräsbahnen werden die Startwerte

$$\begin{aligned}
 a_{e,i}^{(0)} &= 9,075 \text{ mm} \\
 v_{c,i}^{(0)} &= 255 \text{ m/min} \\
 v_{f,i}^{(0)} &= 2800 \text{ mm/min}, \quad i = 1, \dots, n_b \\
 \alpha^{(0)} &= 0^\circ
 \end{aligned} \tag{6.12}$$

gewählt. Die Startschätzung für den Verschiebungsvektor  $u_{\text{quell}}^{(0)}$  wird durch das Lösen des Gleichungssystems (6.6) mit den Werten aus (6.12) bestimmt. Dies hat den Vorteil, dass die Optimierung mit Ausnahme von Nebenbedingung (6.8) im zulässigen Bereich starten kann. Die Optimierung mit Ipopt benötigt 18,9h und 152 Iterationen um zu konvergieren. In Abb. 6.6a sind die optimierten Fräsbahnen sowie die resultierenden Verschiebungen zu sehen. Die Torsion kann eliminiert werden, indem Prozesse mit brei-



(a) Optimierte Fräsbahnen und die resultierenden quellspannungsbedingten Verschiebungen (b) Spannungsverlauf im Bauteil entlang der  $x$ -Achse

Abbildung 6.6: Ergebnis der Optimierung von Typ **A** mit 16 Bahnen

ten und schmalen Eingriffsbreiten alternieren. So können sich negative und positive Torsionsspannungen  $\tau_{xy}$  neutralisieren (siehe Abb. 6.6b).

Die Normalspannungen lassen sich nicht gänzlich neutralisieren, da die optimierten Maschinenparameter fast ausschließlich positive Spannungen hervorrufen. Dadurch entsteht eine leichte Biegung des Bauteils. Dies ist entlang der  $y$ -Achse stärker ausgeprägt, da auch die Werte für  $\sigma_y$  im Mittel größer sind als die Werte für  $\sigma_x$  (siehe Abb. 6.6b). Die optimierten Eingriffsbreiten liegen im Intervall  $a_{e,i}^* \in [5.1; 12.5]$  mm,  $i = 1, \dots, n_b$ . Die optimierten Werte der anderen Maschinenparameter liegen numerisch auf den Schranken des zulässigen Bereichs (siehe (6.9)). Die Schnittgeschwindigkeiten nehmen alle den Wert der unteren Schranke an:  $v_{c,i}^* \approx 255$  m/min,  $i = 1, \dots, n_b$ . Die Vorschubgeschwindigkeiten wechseln in Abhängigkeit der Eingriffsbreite zwischen unterer und oberer Schranke:  $v_{f,i}^* \approx 2000$  mm/min oder  $v_{f,i}^* \approx 2800$  mm/min,  $i = 1, \dots, n_b$ . Die Ursache dieses Verhaltens ist die Linearität des Regressionsmodells und die Gegebenheit, dass die Schnittgeschwindigkeiten und die Vorschubgeschwindigkeiten im Gegensatz zu den Eingriffsbreiten keine weiteren Nebenbedingungen einhalten müssen. Der optimierte Bahnwinkel ist  $\alpha^* \approx 0^\circ$ . Eine Vergrößerung des Winkels würde dazu führen, dass die durchschnittliche Eingriffsbreite erhöht werden müsste (siehe Abb. 6.7). Dies hätte durch die Eigenschaften des Regressionsmodells eine erhöhte Biegung zur Folge, was durch die Optimierung vermieden wird.



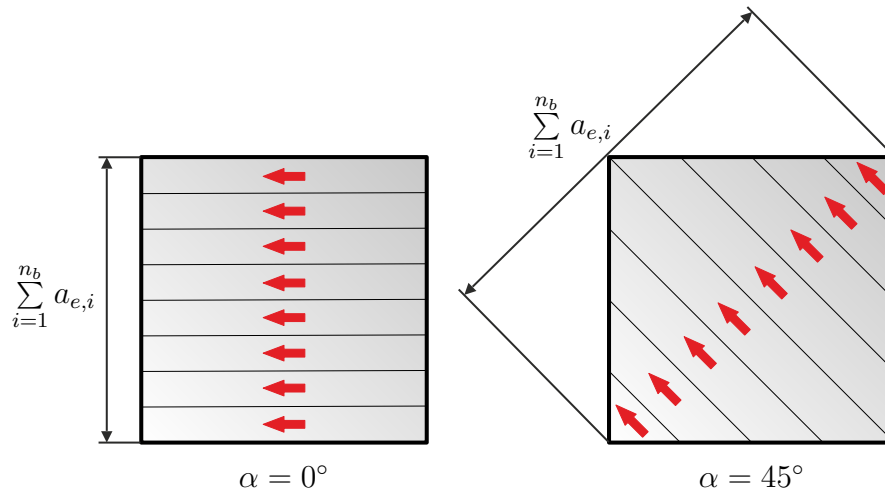


Abbildung 6.7: Vergrößerung der Eingriffsbreiten durch Vergrößerung des Bahnwinkels

Im Vergleich zur trivialen Bearbeitungsstrategie weist die optimierte Strategie vom Typ **A** mit 16 Bahnen eine deutlich reduzierte Verformung auf. Eine Torsion des Bauteils kann verhindert werden. Die maximale Abweichung kann um 77 % auf  $u_{\text{quell}}^{\text{max}} \approx 2,1 \mu\text{m}$  verringert werden, der Zielfunktionswert sogar um 91 % auf  $F(X^*) \approx 0,7 \mu\text{m}^2$ . Aus ökonomischer Sicht ist diese Strategie jedoch nicht von Vorteil. Die Anzahl an Fräsprozessen ist leicht erhöht und die Werkzeugposition muss nach jedem Einzelprozess verschoben werden.

### Typ B, 40 Bahnen

Die Startschätzung für die Optimierung von Strategie **B** mit  $n_b = 40$  Einzelprozessen ist durch

$$\begin{aligned}
 a_{e,i}^{(0)} &= 7,26 \text{ mm}, & i &= 1, \dots, \frac{n_b}{4}, \frac{3n_b}{4} + 1, \dots, n_b \\
 a_{e,i}^{(0)} &= 4,84 \text{ mm}, & i &= \frac{n_b}{4} + 1, \dots, \frac{3n_b}{4} \\
 v_{c,i}^{(0)} &= 255 \text{ m/min}, & i &= 1, \dots, n_b \\
 v_{f,i}^{(0)} &= 2800 \text{ mm/min}, & i &= 1, \dots, n_b
 \end{aligned} \tag{6.13}$$

gegeben. Dabei wird  $u_{\text{quell}}^{(0)}$  erneut durch das Lösen von (6.6) mit den Werten aus (6.13) bestimmt. Der Optimierungsalgorithmus benötigt 108 Iterationen und 17,2h bis zur Konvergenz. Das Ergebnis der Optimierung ist in Abb. 6.8a zu sehen. Die Abb. zeigt das Bauteil mit den resultierenden Verschiebungen und die optimierten Fräsbahnen. Dabei fällt auf, dass der Bereich von Fräsbahnen entlang der  $x$ -Richtung in der Ecke  $x = 145,2 \text{ mm}$  und  $y = 0 \text{ mm}$  des Bauteils nahezu verschwindet. Darüber hinaus sind auch hier in großen Teilen der Oberfläche alternierend schmale und breite Eingriffsbreiten erkennbar. Dadurch, dass der Optimierung vergleichsweise viele Einzelprozesse zur Verfügung stehen, kann der zulässige

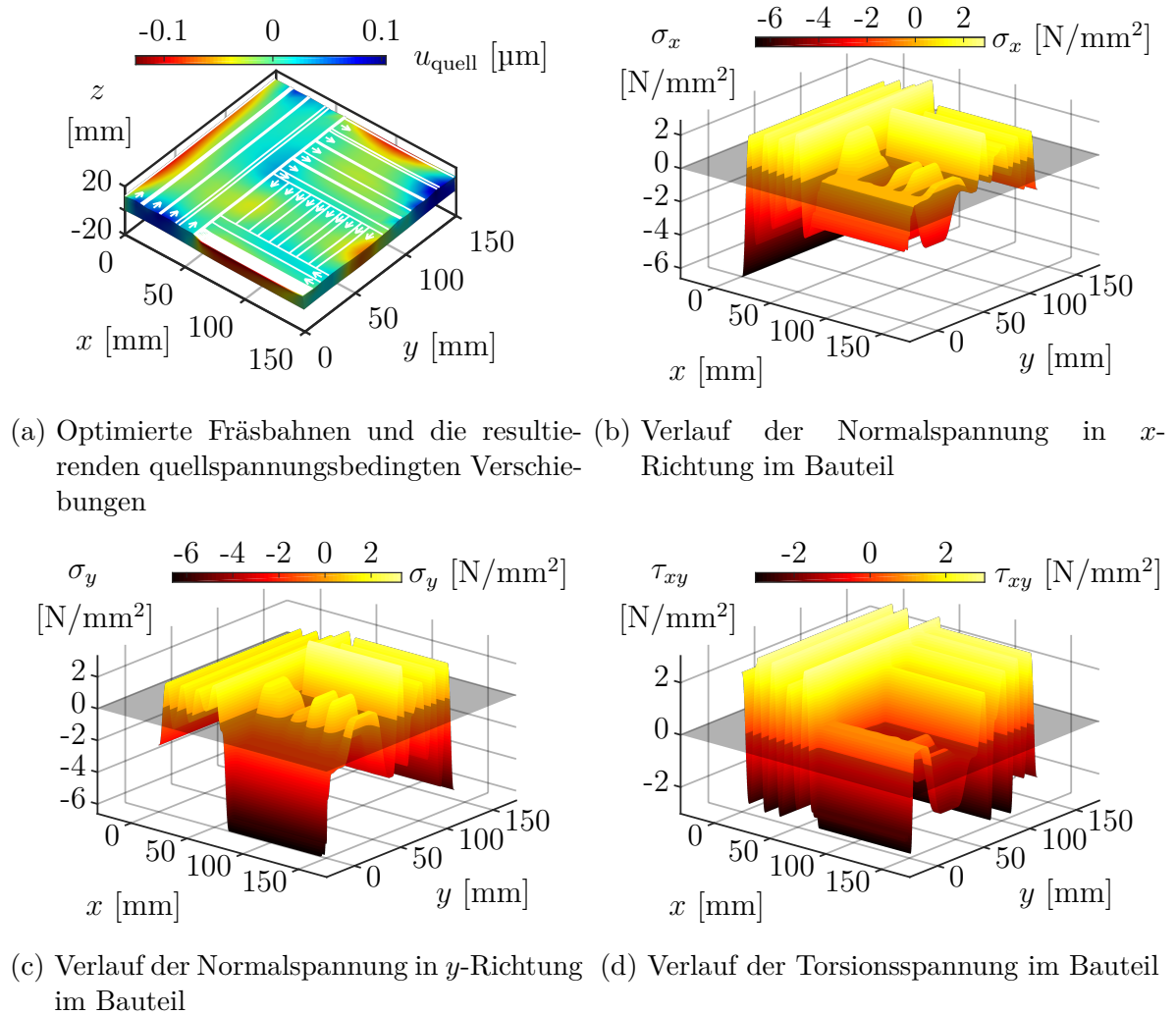


Abbildung 6.8: Ergebnis der Optimierung von Typ **B** mit 40 Bahnen

Bereich der Eingriffsbreiten ( $a_{e,i}^* \in [0.95; 12.5]$  mm,  $i = 1, \dots, n_b$ ) in Gänze ausgenutzt werden und die Nebenbedingungen (6.10) sind weniger restriktiv. Die Maschinenparameter  $v_c$  und  $v_f$  sind erneut fast ausschließlich gegen ihre Schranken konvergiert.

Durch die Vielzahl an Fräsprozessen ergibt sich ein großes Optimierungspotential. Das Maximum der quellspannungsbedingten Verformung kann gegenüber der trivialen Strategie um 99 % auf  $u_{\text{quell}}^{\text{max}} \approx 0,1$   $\mu\text{m}$  reduziert werden. Der Zielfunktionswert ist annähernd 0 ( $F(X^*) \approx 0,0005$   $\mu\text{m}^2$ ). Infolge der geringen mittleren Eingriffsbreite treten im Gegensatz zur Optimierung von Typ **A** mit 16 Bahnen auch negative Normalspannungen auf (siehe Abb. 6.8b und 6.8c). Dies ermöglicht einen Ausgleich von negativen und positiven Werten der Normalspannungen sowohl in  $x$ - als auch in  $y$ -Richtung. Eine Biegung des Bauteils wird fast vollständig vermieden. Die Torsion wird durch einen Ausgleich der Werte der Torsionsspannungen (siehe Abb. 6.8d) erneut eliminiert.

Da gewöhnliche Bearbeitungszentren nur über eine Bearbeitungsgenauigkeit von  $1\ \mu\text{m}$  verfügen, weist diese optimierte Strategie deutlich kleinere Verformungen auf als es erforderlich ist. Mit 40 Fräsbahnen und dem häufigen Verschieben der Werkzeugposition ist diese Strategie jedoch die ineffizienteste aller hier betrachteten Strategien.

### Typ C, 24 Bahnen

Die Startwerte für die Optimierung werden mit Ausnahme der Eingriffsbreiten wie zuvor (siehe (6.13)) gewählt. Für die Eingriffsbreiten gilt:

$$\begin{aligned} a_{e,2i}^{(0)} &= 12,1\ \text{mm} \\ a_{e,2i-1}^{(0)} &= 8,0\overline{6}\ \text{mm}, \quad i = 1, \dots, \frac{n_b}{2} \end{aligned} \tag{6.14}$$

Nach 24 Iterationen und 5,4 h konvergiert der Optimierungsalgorithmus. Abbildung 6.9a zeigt die resultierenden Formabweichungen und die optimierten Fräsbahnen. Durch die vergleichsweise geringe Anzahl an Einzelprozessen bleibt der Optimierung nur wenig Spielraum zur Variation der Eingriffsbreite. Die optimierten Eingriffsbreiten nehmen größtenteils den Wert der oberen Schranke ( $12,5\ \text{mm}$ ) an. Nur wenige Werte sind kleiner. Eine Eingriffsbreite nimmt den Wert der unteren Schranke an ( $0,95\ \text{mm}$ ). Die Schnittgeschwindigkeiten konvergieren ausschließlich gegen die obere Schranke:  $v_{c,i}^* = 373\ \text{m/min}$ ,  $i = 1, \dots, n_b$ ; die Vorschubgeschwindigkeiten konvergieren gegen die untere Schranke:  $v_{f,i}^* = 2000\ \text{mm/min}$ ,  $i = 1, \dots, n_b$ .

Im Allgemeinen sind die Verformungen bei dieser optimierten Strategie gering, wenn auch etwas höher als bei der Strategie zuvor. Es treten leichte Biegungen auf, die zu einer maximalen Formabweichung von  $u_{\text{quell}}^{\text{max}} \approx 1,1\ \mu\text{m}$  führen. Dies entspricht einer Reduzierung um 88%. Der Zielfunktionswert kann sogar um 97% auf  $F(X^*) \approx 1,1\ \mu\text{m}^2$  reduziert werden. Die Biegungen lassen sich aufgrund der großen mittleren Eingriffsbreite nicht eliminieren, können aber in großen Teilen des Bauteils auf einem niedrigen Niveau gehalten werden (siehe Abb. 6.9b und 6.9c). Dabei sind die Biegungen in  $x$ - bzw.  $y$ -Richtung in etwa gleich stark ausgeprägt. Strategien vom Typ C haben den Vorteil, dass sich in etwa gleich große Bereiche vertikaler und horizontaler Fräsbahnen ergeben. Bei einer Drehung der Fräsbahn um  $90^\circ$  wechselt das Vorzeichen der Torsionsspannung durch die erforderliche Achsentransformation (siehe (3.5)). Die gegensätzlichen Torsionsspannungen können sich so im Mittel eliminieren (siehe Abb. 6.9d). Deshalb sind Strategien vom Typ C besonders geeignet, um eine Torsion des Bauteils zu verhindern.

Die erforderliche Fertigungsgenauigkeit ( $1\ \mu\text{m}$  für gewöhnliche Bearbeitungszentren) erreicht auch diese Strategie. Sie ist zudem die wirtschaftlichste aller in dieser Arbeit betrachteten Strategien, da ein zusammenhängender Werkzeugpfad mit einer mittleren Eingriffsbreite nahe dem Werkzeugradius entsteht.

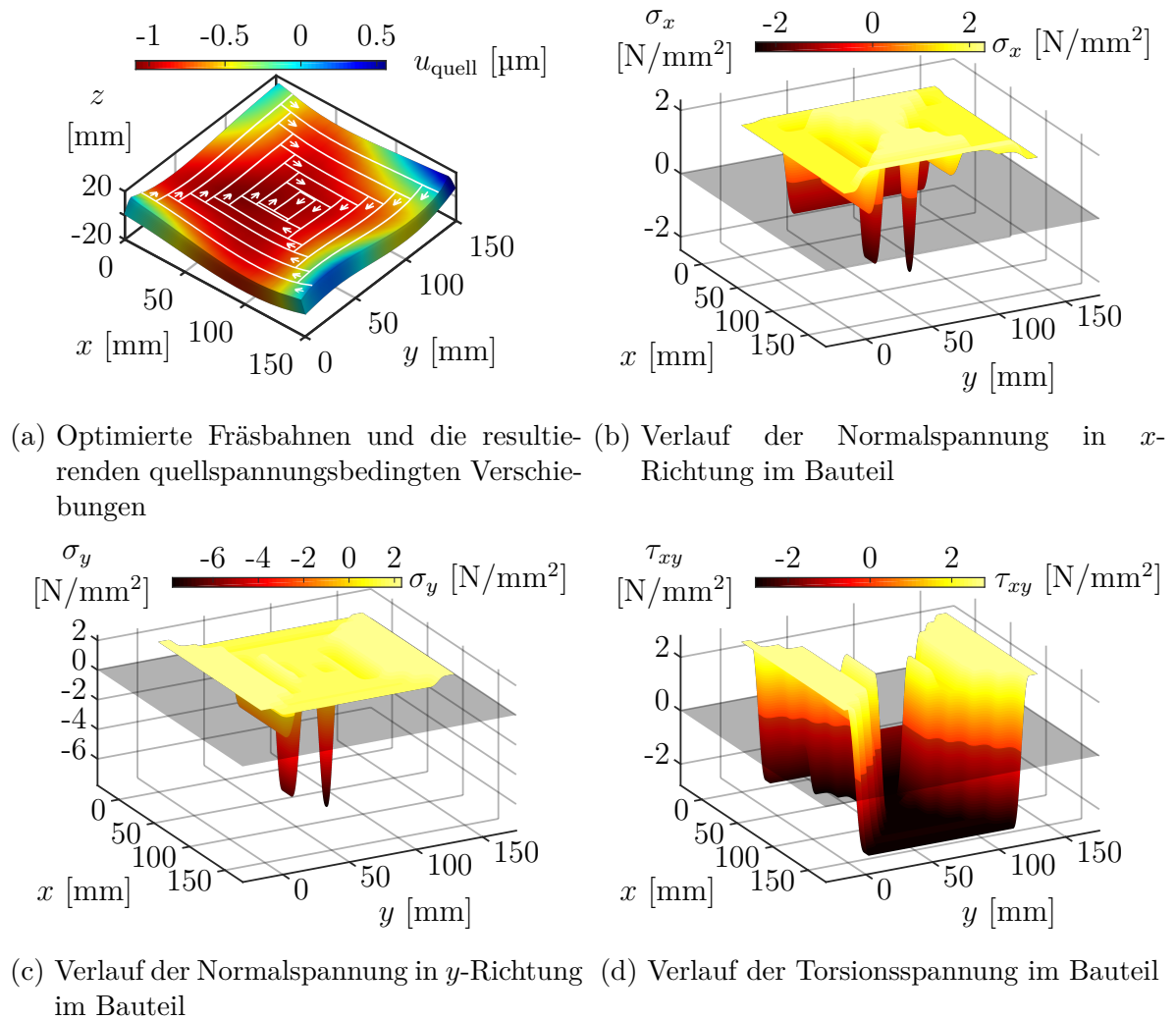


Abbildung 6.9: Ergebnis der Optimierung von Typ **C** mit 24 Bahnen

## Übersicht

Eine Übersicht zu den Ergebnissen der Optimierung aller betrachteten neun Strategien ist in Tab. 6.1 zu finden. Die gerundeten Prozentzahlen geben die Reduzierung im Vergleich zur trivialen Strategie an. Bis auf Typ **A** mit 12 Bahnen sind die Zielfunktionswerte und die maximalen Formabweichungen der übrigen Strategien sehr zufriedenstellend. Viele Ergebnisse übertreffen die gewünschte Fertigungsgenauigkeit hinsichtlich der quellspannungsbedingten Formabweichungen sogar deutlich. Je mehr Einzelprozesse eine Strategie besitzt, desto besser sind die Resultate, da der Optimierung mehr Freiheitsgrade zur Verfügung gestellt werden. Typ **B** und **C** mit 40 Bahnen weisen die geringste Maximalabweichung auf. Aufgrund der geringen Deformationen kann auch der Fertigungseffizienz eine besondere Bedeutung beigemessen werden. Hinsichtlich der Wirtschaftlichkeit sind Strategien mit wenigen Bahnen zu bevorzugen sowie Strategien

vom Typ **C**, die ohne Unterbrechung des Fräspfades auskommen. So stellt Typ **C** mit 24 Bahnen die effizienteste Art der Fertigung dar.

Typ A			Typ B			Typ C		
$n_b$	$F(X^*)$	$u_{\text{quell}}^{\text{max}}$	$n_b$	$F(X^*)$	$u_{\text{quell}}^{\text{max}}$	$n_b$	$F(X^*)$	$u_{\text{quell}}^{\text{max}}$
12	keine Lösung		24	0,3 $\mu\text{m}^2$ (-96 %)	1,9 $\mu\text{m}$ (-79 %)	24	0,2 $\mu\text{m}^2$ (-97 %)	1,1 $\mu\text{m}$ (-88 %)
16	0,7 $\mu\text{m}^2$ (-91 %)	2,1 $\mu\text{m}$ (-77 %)	32	0,001 $\mu\text{m}^2$ (-100 %)	0,2 $\mu\text{m}$ (-98 %)	32	0,002 $\mu\text{m}^2$ (-100 %)	0,2 $\mu\text{m}$ (-98 %)
20	0,015 $\mu\text{m}^2$ (-100 %)	0,5 $\mu\text{m}$ (-95 %)	40	0,0005 $\mu\text{m}^2$ (-100 %)	0,1 $\mu\text{m}$ (-99 %)	40	0,0004 $\mu\text{m}^2$ (-100 %)	0,1 $\mu\text{m}$ (-99 %)

Tabelle 6.1: Zielfunktionswert und maximale Abweichung nach Optimierung aller neun Strategien

## 6.2 Optimierung der Gesamtformabweichung

Die Minimierung der Formabweichungen beim Fräsen gliedert sich in zwei Teile. Im vorangegangenen Abschnitt wurden die Maschinenparameter Eingriffsbreite, Schnittgeschwindigkeit und Vorschubgeschwindigkeit optimiert sowie verschiedene Bearbeitungstypen miteinander verglichen. Dabei wurden zunächst nur die quellspannungsbedingten Formabweichungen betrachtet. In diesem Kapitel werden die Gesamtformabweichungen betrachtet. Diese umfassen neben den quellspannungsbedingten Deformationen, welche aus dem ersten Optimierungsschritt übrig geblieben sind, die thermischen Formabweichungen, die aufgrund eines ungleichmäßigen Materialabtrags entstehen. Hierfür wird die Schnitttiefe unter Berücksichtigung eines kinematischen Werkzeugmodells optimiert.

### 6.2.1 Kinematisches Modell der Gesamtformabweichungen

Die Beschreibung des Modells orientiert sich an den Arbeiten von Gulpak et al. [47,51].

Das Modell zur Berechnung der Gesamtformabweichung gliedert sich in mehrere Schritte (siehe Abb. 3.5). Zunächst wird aus den optimierten Maschinenparametern  $a_e^*$ ,  $v_c^*$  und  $v_f^*$  sowie  $\alpha^*$  (Typ **A**) aus dem ersten Optimierungsschritt mit dem thermischen Regressionsmodell der Wärmeeintrag der gesamten Bearbeitung berechnet. Dieser dient als Grundlage für eine erste FEM-Simulation, welche die Wärmeverteilung im Bauteil berechnet. Basierend darauf können in einer zweiten FEM-Rechnung die thermisch

bedingten Verschiebungen  $u_{\text{therm},z}$  ermittelt werden. Durch Subtraktion der quellspannungsbedingten Verschiebungen in  $z$ -Richtung  $u_{\text{quell},z}$ , die in der ersten Optimierung übrig geblieben sind, ergibt sich die resultierende Verschiebung in  $z$ -Richtung. Die quellspannungsbedingten Verschiebungen werden subtrahiert, da sie eine gegensätzliche Wirkung zu den thermisch bedingten Verschiebungen haben. Die thermisch bedingten Verschiebungen wirken unmittelbar während des Prozesses. Eine positive Verschiebung in  $z$ -Richtung erfordert zur Kompensation eine positive Höhenänderung des Werkzeugs. Die quellspannungsbedingten Verschiebungen wirken erst nach dem Ausspannen des Werkstücks. Eine positive Verschiebung in  $z$ -Richtung nach dem Fräsen kann durch eine negative Höhenänderung des Werkzeugs während des Prozesses kompensiert werden.

Diese rechenintensive Simulationskette wird für jede Bearbeitungsstrategie einmalig durchgeführt. Auf Basis der Gesamtverschiebungen erfolgt anschließend die zweite Optimierung, die die Schnitttiefe  $a_p$  hinsichtlich einer möglichst ebenen Oberfläche optimiert. Dafür wird ein kinematisches Modell (KM), das die Werkzeuggeometrie und -bewegung abbildet, verwendet, um den Materialabtrag in Abhängigkeit von der Schnitttiefe  $a_p$  zu berechnen. Im Modell werden die Oberflächenknoten der FEM-Simulation sowie ihre Position mit dem zugehörigen Verschiebungsvektor  $u_{\text{therm},z} - u_{\text{quell},z}$  addiert und mit der Kinematik des Werkzeugs wird die effektive Schnitttiefe  $a_{p,\text{eff}}$  bestimmt. Für jeden Knoten wird dabei über alle Zeitschritte die größte effektive Schnitttiefe berücksichtigt, um abschließend die resultierende Oberflächengeometrie bzw. die resultierende Gesamtverformung  $u_{\text{ges},z}$  zu erhalten (siehe Abb. 6.10). Dieses kinematische Modell wurde von

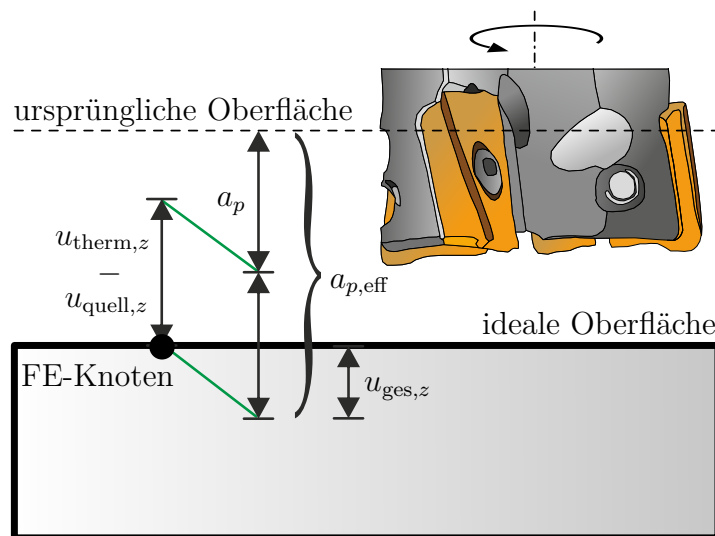


Abbildung 6.10: Ermittlung der Gesamtformabweichung in Abhängigkeit von der Schnitttiefe

den Projektkollegen aus den Ingenieurwissenschaften entwickelt und dient als eine Art

Black-Box-Modell

$$u_{\text{ges},z} = \text{KM}(a_p), \quad (6.15)$$

das für die Optimierung verwendet wird.

## 6.2.2 Mathematische Formulierung der Optimierungsprobleme

Bei der zweiten Optimierung ist das Ziel, eine möglichst ebene Oberfläche zu erhalten, die auf einer Höhe von 11 mm liegt. Verwendet werden drei Bearbeitungsstrategien, die bereits im ersten Optimierungsschritt hinsichtlich  $a_e$ ,  $v_c$  und  $v_f$  sowie für Typ **A** auch hinsichtlich  $\alpha$  optimiert wurden. Die Optimierungsvariablen sind die Schnitttiefen der einzelnen Fräsprozesse. Dabei wird die Schnitttiefe pro Bahn als Parabel modelliert. Dies hat den Vorteil, dass als Ergebnis der Optimierung ein glatter Fräspfad entsteht, der auf einem CNC-Bearbeitungszentrum realisiert werden kann. Die einzelnen Parabeln können durch drei Variablen beschrieben werden. Sie stellen die Höhe in  $z$ -Richtung bzw. die Schnitttiefe am Anfang, in der Mitte und am Ende der Fräsbahn dar. Bei  $n_b$  Einzelprozessen ergeben sich  $n_x = 3n_b$  Optimierungsvariablen, die in einem Vektor zusammengefasst werden:  $X = a_p \in \mathbb{R}^{3n_b}$ . Aus diesen lässt sich mit dem kinematischen Black-Box-Modell die Gesamtformabweichung bestimmen. Die Werte an den diskreten Punkten der Oberfläche lassen sich in einem Vektor  $u_{\text{ges},z} \in \mathbb{R}^{n_s}$  zusammenfassen. Um die Deformation zu minimieren, wird in der Zielfunktion die Summe der quadratischen Formabweichungen berücksichtigt. Daneben kann es aus Stabilitätsgründen von Vorteil sein, einen Regularisierungsterm zu addieren. Bei schmalen Fräsbahnen besteht die Möglichkeit, dass diese von nachfolgenden Fräsprozessen überlagert werden, d. h. dass die Oberfläche ein zweites Mal bearbeitet wird. Bezüglich des kinematischen Modells wird dadurch die schmale Bahn redundant. Durch den zusätzlichen Term wird sichergestellt, dass sich die Bahnen nahe der Oberfläche befinden und damit alle Einzelprozesse den durch die Bearbeitungsstrategie vorgesehenen Materialabtrag ausführen. Insgesamt ist das Optimierungsproblem durch

$$\min_{a_p \in \mathbb{R}^{3n_b}} \sum_{i=1}^{n_s} u_{\text{ges},z,i}(a_p)^2 + c \sum_{i=1}^{3n_b} (a_{p,i} - 1 \text{ mm})^2 \quad (6.16)$$

gegeben. Hierbei entspricht  $a_p = 1 \text{ mm}$  der Höhe der idealen Oberfläche. Der Gewichtungsfaktor  $c \in \mathbb{R}$  wird hinreichend klein gewählt, damit die Minimierung der Formabweichung dominiert. Da die Modellberechnungen  $u_{\text{ges},z,i}(a_p)$  außerhalb der Optimierung in einer Black-Box erfolgen und es keine weiteren Nebenbedingungen gibt, ist das Optimierungsproblem (6.16) unrestringiert ( $m_h = 0$  und  $m_g = 0$ ). Die Optimierungssoftware WORHP, welche zum Lösen des Optimierungsproblems verwendet wird, ist auf NLPs spezialisiert. Es wird sich jedoch zeigen, dass sie auch zum Lösen dieser unrestringierten Probleme gut geeignet ist. Es wird dabei auf die Routinen der numerischen Differentiation zurückgegriffen, da die analytischen Ableitungen des kinematischen Modells nur schwer zu bestimmen sind.

### 6.2.3 Numerische Ergebnisse

Zur vereinfachten Darstellung wird die folgende Variable eingeführt:

$$u_{\text{ges}}^{\max} := \max_{i=1, \dots, n_s} |u_{\text{ges},i}|$$

Für die numerische Minimierung der Gesamtformabweichung wurde das System aus Tab. 4.2 verwendet. Es werden jeweils die effizientesten Bearbeitungsstrategien der drei Typen aus Kapitel 6.1.5 ausgewählt, da alle Optimierungen (außer Strategie **A** mit 12 Bahnen) die quellspannungsbedingten Formabweichungen zufriedenstellend reduzieren konnten. Die effizientesten Bearbeitungsstrategien sind Typ **A** mit 16 Bahnen<sup>5</sup> sowie Typ **B** und **C** mit 24 Bahnen. Die Maschinenparameter mit Ausnahme der Schnitttiefe erhalten die optimierten Werte  $a_e^*$ ,  $v_c^*$  und  $v_f^*$  sowie bei Typ **A**  $\alpha^*$  aus Kapitel 6.1.5. Die Startschätzung der Optimierungsvariable  $a_p$  wird konstant  $1 \mu\text{m}$  über der idealen Oberfläche gewählt:

$$a_p^{(0)} = 0,999 \text{ mm} \in \mathbb{R}^{3n_b} \quad (6.17)$$

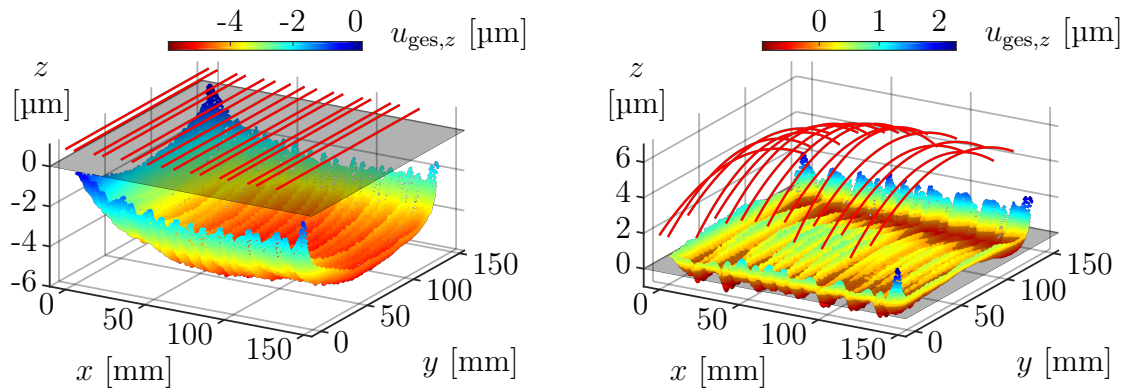
Dies wirkt der thermischen Ausdehnung bereits etwas entgegen.

#### Typ A, 16 Bahnen

Für die Optimierung der Bearbeitungsstrategie **A** mit 16 Bahnen wurde zunächst von den Projektpartnern eine mehrtägige Simulation der thermisch bedingten Verschiebungen durchgeführt. Anschließend wurden diese mit den quellspannungsbedingten Verschiebungen verrechnet. Abbildung 6.11a zeigt die resultierende Deformation der Bauteiloberfläche durch eine Bearbeitung mit konstanter Schnitttiefe  $a_p = 0,999 \text{ mm}$  sowie die ideale Oberfläche. In Rot sind die dafür verwendeten Fräspfade visualisiert. Neben der genannten Schnitttiefe wurden die Maschinenparameter  $a_e^*$ ,  $v_c^*$ ,  $v_f^*$  und  $\alpha^*$  der ersten Optimierung verwendet. Dabei überlagern sich zwei Effekte: Zum einen führen die Quellspannungen nach dem Ausspannen zu einem Absenken des Zentrums des Werkstücks (siehe Abb. 6.6a). Zum anderen bewirken die thermisch bedingten Verschiebungen in positiver  $z$ -Richtung einen erhöhten Materialabtrag, der nach dem Abkühlen des Materials ein weiteres Absenken der Oberfläche zur Folge hat. Der thermische Effekt nimmt durch die Wärmeakkumulation mit voranschreitendem Prozess in  $x$ -Richtung sowie zur Mitte des Werkstücks zu. Die maximale Abweichung beträgt  $u_{\text{ges}}^{\max} \approx 6,2 \mu\text{m}$ , wobei etwa ein Drittel durch Quellspannungen und zwei Drittel durch einen ungleichmäßigen Materialabtrag hervorgerufen werden. Der Zielfunktionswert dieser hinsichtlich der Schnitttiefe trivialen Bearbeitung ist  $F(X) \approx 1,7 \mu\text{m}^2$ .

<sup>5</sup>Für die effizientere Strategie **A** mit 12 Bahnen konvergiert die Optimierung aufgrund der Torsionsnebenbedingung (6.8) nicht.





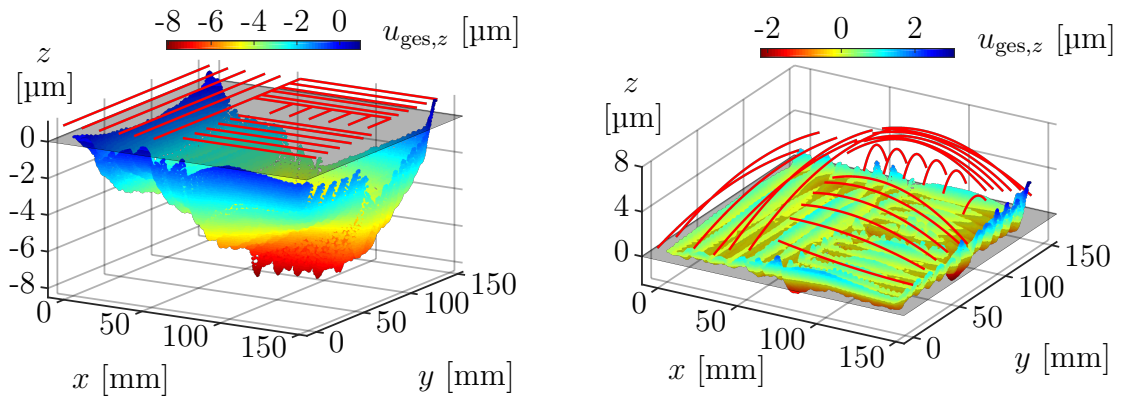
(a) Fräsbahnen und resultierende Oberfläche (b) Optimierte Fräsbahnen und resultierende Oberfläche bei konstanter Schnitttiefe  $a_p^{(0)}$

Abbildung 6.11: Ergebnis der Optimierung von Typ **A** mit 16 Bahnen

Für die Optimierung dieser Bearbeitungsstrategie wurde die Software WORHP mit  $n_x = 48$  Optimierungsvariablen und den Startwerten aus (6.17) verwendet. Sie benötigt lediglich 2 min, um nach 37 Hauptiterationen zu konvergieren. Die geringe Berechnungsdauer wird durch die kleine Problemgröße und das performante kinematische Modell erreicht. Die optimierten Fräspfade sowie die resultierende Oberflächenverformung sind in Abb. 6.11b zu sehen. Die gewölbten Fräsbahnen kompensieren die Verschiebungen in  $z$ -Richtung in weiten Teilen. Die größten Abweichungen treten in den Ecken des Werkstücks und am Rand auf. Hier sind die Verschiebungen in  $z$ -Richtung vergleichsweise klein, da die thermisch bedingten Verschiebungen zu einem Großteil auch in  $x$ - und  $y$ -Richtung erfolgen können und auch die quellspannungsbedingten Verschiebungen in den Ecken und am Rand gering sind. Die Gradienten der Verschiebungen sind dadurch besonders groß. Eine Kompensation durch parabelförmige Bahnen ist dort nur mit Einschränkungen möglich. Dennoch kann die maximale Abweichung im Vergleich zur konstanten Bearbeitung um 62% auf  $u_{\text{ges}}^{\text{max}} \approx 2,3 \mu\text{m}$  reduziert werden. Dies kommt der Fertigungsgenauigkeit von konventionellen Bearbeitungszentren sehr nahe. Der Zielfunktionswert wird auf  $F(X^*) \approx 0,04 \mu\text{m}^2$  verringert.

### Typ B, 24 Bahnen

Als zweites wird die Strategie vom Typ **B** mit 24 Bahnen betrachtet. Nach der Simulation der thermisch bedingten Verschiebungen konnte zusammen mit den quellspannungsbedingten Verschiebungen eine Simulation des kinematischen Modells durchgeführt werden. Die Schnitttiefe betrug dabei  $a_p = 0,999 \text{ mm}$ . Die resultierende Oberfläche des Bauteils sowie die Fräspfade sind in Abb. 6.12a zu sehen. Auffällig sind die steigenden Abweichungen mit Voranschreiten der Gesamtbearbeitungen. Die größten Abweichungen treten im Bereich  $x > 70 \text{ mm}$  auf. Die Bahnen sind dort zudem vergleichsweise kurz, sodass viele Prozesse auf kleinem Raum ausgeführt werden. Es entsteht eine erhöhte



(a) Fräsbahnen und resultierende Oberfläche bei konstanter Schnitttiefe  $a_p^{(0)}$  (b) Optimierte Fräsbahnen und resultierende Oberfläche

Abbildung 6.12: Ergebnis der Optimierung von Typ **B** mit 24 Bahnen

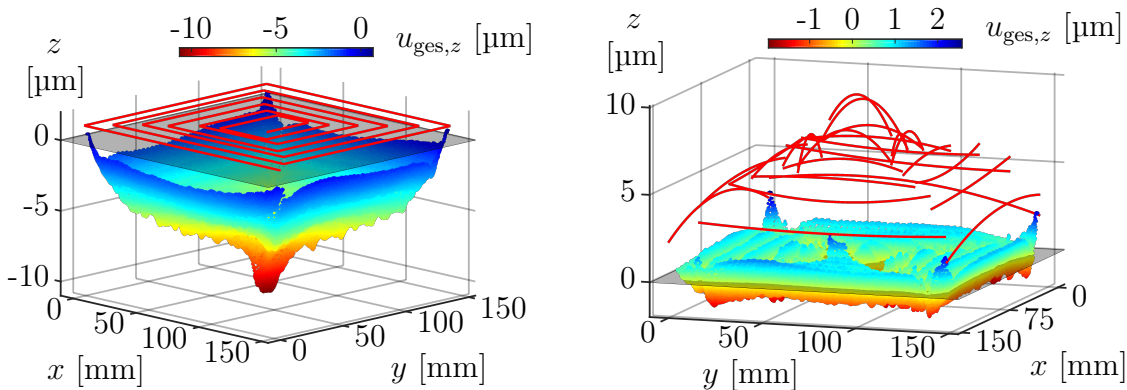
Wärmeakkumulation, die zu einem erhöhten Materialabtrag führt. Der Anteil der thermisch bedingten Formabweichungen ist deshalb mit 75 % größer als im Versuch zuvor. Auch die Maximalabweichung von  $u_{\text{ges}}^{\text{max}} \approx 8,1 \mu\text{m}$  übertrifft die vorherige. Der Zielfunktionswert ist aufgrund des Bereichs  $x < 70 \text{ mm}$  mit vergleichsweise kleinen Abweichungen auf  $F(X) \approx 1,5 \mu\text{m}^2$  reduziert.

Der Optimierung mit WORHP stehen  $n_x = 72$  Optimierungsvariablen mit den Startwerten aus (6.17) zur Verfügung. Nach 7 min und 33 Hauptiterationen konvergiert der Algorithmus. Durch die zusätzlichen Fräsprozesse ist die Optimierungsdauer länger als zuvor, aber weiterhin sehr gering. Die optimierten Fräspfade und die resultierende Oberfläche sind in Abb. 6.12b zu sehen. Erneut ist gut zu erkennen, dass die Fräsbahnen in etwa um jenen Wert angehoben werden, welcher den Formabweichungen der konstanten Bearbeitung (siehe Abb. 6.12a) entspricht. Unter Berücksichtigung von Abb. 6.10, die die Funktionsweise des kinematischen Modells illustriert, ist diese Kompensation nachvollziehbar. Zu den größten Abweichungen von der Idealgeometrie kommt es wie zuvor am Rand. Das veränderte Materialverhalten bezüglich der Ausdehnung sorgt auch hier für große Verschiebungsgradienten in Richtung Bauteilmitte. Diese lassen sich durch parabolische Schnitttiefen nicht gänzlich kompensieren. So verbleibt eine maximale Restabweichung von  $u_{\text{ges}}^{\text{max}} \approx 3,1 \mu\text{m}$ . Dies ist eine Reduzierung von 62 % verglichen zur Bearbeitung mit konstanter Schnitttiefe. Zu bemerken ist, dass die Abweichung im Mittel sehr viel kleiner ausfällt. Der optimale Zielfunktionswert beträgt  $F(X^*) \approx 0,07 \mu\text{m}^2$ .

Im Vergleich zur ersten Strategie vom Typ **A** ist diese Strategie vom Typ **B** schlechter zur hochpräzisen Fertigung geeignet. Zum einen ist sie durch die häufigen Unterbrechungen der Einzelprozesse weniger effektiv, zum anderen sind sowohl die Zielfunktion als auch die maximale Abweichung größer.

### Typ C, 24 Bahnen

Die letzte Strategie, für welche die Schnitttiefen optimiert wurden, ist vom Typ **C** mit 24 Bahnen. Das Simulationsergebnis mit konstanten Schnitttiefen  $a_p = 0,999$  mm wird in Abb. 6.13a gezeigt. Dadurch, dass die Einzelprozesse in der Mitte des Bauteils zusammenlaufen, entsteht im Zentrum eine sehr große Wärmeakkumulation, die zu einem erhöhten Materialabtrag und zu einem starken Absenken der Bauteiloberfläche führt. Die quellspannungsbedingten Effekte haben an der Maximalabweichung von  $u_{\text{ges}}^{\text{max}} \approx 10,7$   $\mu\text{m}$  mit etwa 10 % nur einen untergeordneten Anteil. Auch die Zielfunktion der Bearbeitung mit konstanten Schnitttiefen hat dadurch einen erhöhten Wert:  $F(X) \approx 2,0$   $\mu\text{m}^2$ .



(a) Fräsbahnen und resultierende Oberfläche bei konstanter Schnitttiefe  $a_p^{(0)}$  (b) Optimierte Fräsbahnen und resultierende Oberfläche

Abbildung 6.13: Ergebnis der Optimierung von Typ **C** mit 24 Bahnen

Die Optimierung benötigt 37 Hauptiterationen und 9 min, um die  $n_x = 72$  Schnitttiefen zu optimieren. Wie in Abb. 6.13b zu sehen ist, gelingt es erneut, die Abweichungen durch erhöhte Fräsbahnen zu kompensieren. Trotz der großen Wärmeakkumulation im Zentrum des Bauteils treten auch hier die größten Abweichungen am Rand auf. Die maximale Abweichung beträgt  $u_{\text{ges}}^{\text{max}} \approx 2,6$   $\mu\text{m}$ , was einer Reduzierung von 76 % im Vergleich zur Bearbeitung mit konstanten Schnitttiefen entspricht. Der Zielfunktionswert beträgt  $F(X^*) \approx 0,1$   $\mu\text{m}^2$ . Zu bemerken ist, dass diese in der Industrie häufig eingesetzte Bearbeitungsstrategie besonders von einer Optimierung der Schnitttiefe profitiert, da die Wärmeentwicklung zum Ende des Prozesses in der Mitte des Bauteils äußerst hoch ist.

Insgesamt liefert diese Strategie ähnlich geringe Abweichungen wie Strategie **A** mit 16 Bahnen. Jedoch ist diese Strategie vom Typ **C** die effizienteste aller betrachteten Strategien, weshalb sie für die Anwendung zu bevorzugen ist.

### Übersicht

Insgesamt konnte mit der Optimierung der Maschinenparameter eine deutliche Verringerung der Formabweichungen erreicht werden. Tabelle 6.2 gibt eine Übersicht über die

erzielten Ergebnisse. Zunächst wurden mit Sub-Modell 1 nur die quellspannungsbedingten Formabweichungen betrachtet. Durch Verwendung einer trivialen Bearbeitungsstrategie mit parallelen, breiten Fräsbahnen (Typ A) wurde eine Maximalabweichung von  $9,2\ \mu\text{m}$  verursacht. Durch einen Wechsel der Strategie zu Typ C, welche zudem eine effizientere Bearbeitung ermöglicht, konnte die Abweichung um 47% verringert werden. Durch die Optimierung der Maschinenparameter Eingriffsbreite, Schnittgeschwindigkeit und Vorschubgeschwindigkeit konnte insgesamt eine Reduzierung der maximalen quellspannungsbedingten Formabweichung um 88% auf  $1,1\ \mu\text{m}$  erzielt werden.

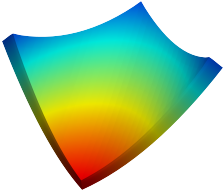
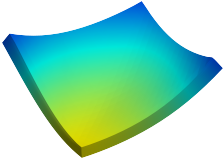

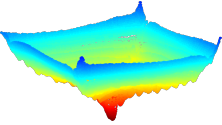
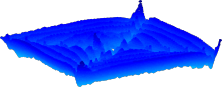
	Typ A	Typ C		
Z	-	-	$a_e, v_c, v_f$	$a_e, v_c, v_f, a_p$
Sub-Modell 1				
	$u_{\text{quell}}^{\text{max}}$ 9,2 $\mu\text{m}$	5,8 $\mu\text{m}$ (-47%)	1,1 $\mu\text{m}$ (-88%)	
Sub-Modell 2				
	$u_{\text{ges}}^{\text{max}}$		10,7 $\mu\text{m}$	2,6 $\mu\text{m}$ (-76%)

Tabelle 6.2: Ergebnisübersicht der Minimierung der Gesamtformabweichung

Mit Sub-Modell 2 werden zusätzlich die thermischen Formabweichungen betrachtet. Ohne Anpassung der Schnitttiefe, aber unter Berücksichtigung der zuvor optimierten Maschinenparameter, resultiert eine maximale Gesamtabweichung von  $10,7\ \mu\text{m}$ . Mit der Optimierung parabolischer Schnitttiefen kann der ungleichmäßige Materialabtrag gut kompensiert werden. Die Gesamtabweichung lässt sich so um 76% verringern. Übrig bleibt eine Restabweichung von  $2,6\ \mu\text{m}$  von der Idealgeometrie.

### 6.2.4 Vergleich zu Messungen realer Experimente

Von den ingenieurwissenschaftlichen Projektpartnern wurden die optimierten Frässtrategien in Experimenten erprobt. Die Ergebnisse dazu wurden in [51] veröffentlicht und werden hier kurz wiedergegeben. Für die Durchführung der Experimente wurden die von der Optimierung berechneten Maschinenparameter sowie die zugehörige Bearbeitungsstrategie in CNC-Code übersetzt. Die Fräspfade wurden dabei durch lineare Segmente

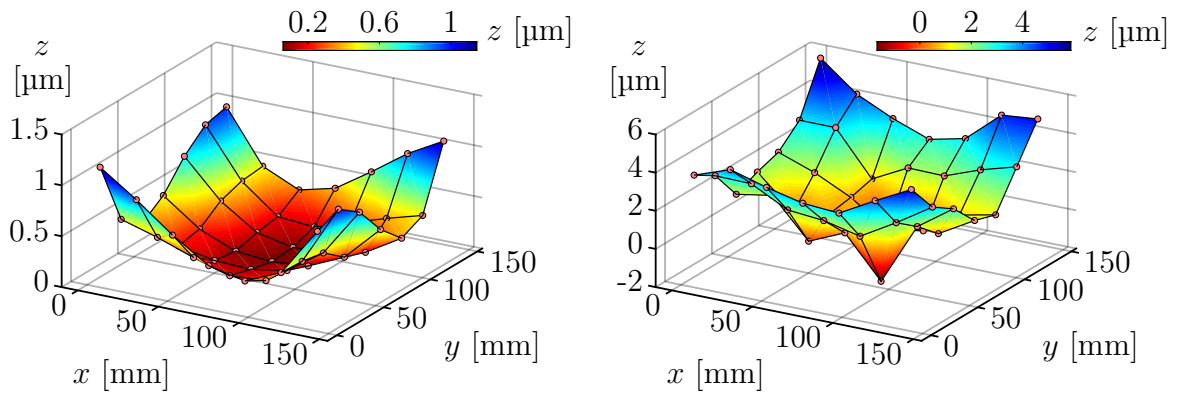
von mindestens 1  $\mu\text{m}$  Länge angenähert. Neben Bauteilen, die bereits vor der Bearbeitung mit Quellspannungen belastet waren, wurden auch Bauteile betrachtet, die vor der Bearbeitung spannungsarmgeglüht wurden. Letztere Untersuchung ist im Kontext dieser Arbeit relevant, da das verwendete Modell Bauteile annimmt, die vor dem Fräsen spannungsfrei sind. Die Geometrie jedes Bauteils wurde dreimal gemessen. Die obere, bearbeitete Seite wurde nach dem Fräsen gemessen, um einen Vergleich zu den Resultaten der Optimierung und den Vorhersagen des Modells zu ermöglichen. Die untere Seite wurde vor und nach dem Fräsen gemessen. Ein Vergleich beider Messungen lässt einen Rückschluss auf die Formänderungen durch den Bearbeitungsprozess zu.

Ein Vergleich der Gesamtformabweichung der bearbeiteten Oberfläche zwischen Modellvorhersage der Optimierung und der Messung nach den Experimenten ergibt Abweichungen von bis zu 30  $\mu\text{m}$ . Neben thermischen Abweichungen, die auch eine im Modell nicht berücksichtigte Ausdehnung des Werkzeugs beinhalten könnten, wird die Ursache vor allem in der Einspannung des Bauteils durch eine Magnetspannplatte vermutet. Dabei sind insbesondere zwei Effekte von Bedeutung, die bei einem Bauteil auftreten können, das vor der Bearbeitung eine Biegung der Unterseite aufweist. Zum einen kann die Biegung des Bauteils durch die magnetische Kraft der Spannplatte temporär reduziert werden. Nach dem Ausspannen federt das Bauteil in seine ursprüngliche Form zurück und weist dadurch Formabweichungen auf. Dieser Effekt konnte durch Messungen der Geometrie vor und nach dem Einspannen bestätigt werden. Zum anderen kann das Bauteil durch eine konvexe Biegung, die durch die Magnetkraft nicht gänzlich entfernt wurde, auf der Spannplatte rollen. Durch zur Oberfläche orthogonale Prozesskräfte können so die Formabweichungen der Biegung der Unterseite zum Teil auf die gefertigte Oberfläche übertragen werden.

Durch die Betrachtung der Werkstückunterseite konnte die Vorhersagegenauigkeit des Quellspannungsmodells separat überprüft werden. Abbildung 6.14 zeigt einen Vergleich zwischen simulierter quellspannungsbedingter Formänderung und gemessener Formänderung der Unterseite nach der Bearbeitung eines spannungsarmgeglühten Bauteils mit der optimierten Strategie vom Typ C. Es ist eine hohe Übereinstimmung festzustellen. Die Abweichungen betragen im Maximum nur 4  $\mu\text{m}$  und weisen qualitativ eine ähnliche Biegung auf. Insbesondere konnte durch die optimierte Bearbeitung sowohl in der Simulation als auch im Experiment eine Torsion des Bauteils vermieden werden. Insgesamt verdeutlichen diese Auswertungen das große Potential des Hybridmodells mit zum Teil guten Modellvorhersagen. Insbesondere zeigen die Untersuchungen Erweiterungsmöglichkeiten auf, um die Modellgenauigkeit weiter zu steigern.

## 6.3 Fazit der Minimierung von Formabweichungen

Für die Minimierung von Formabweichungen beim Trockenfräsen wurde ein neuartiges Hybridmodell verwendet. Dieses betrachtet neben den thermischen Dehnungen, die zu einem ungleichmäßigen Materialabtrag führen können, auch Biegungen und Torsion von



(a) simulierte quellspannungsbedingte Formänderung (b) gemessene Formänderung der Unterseite nach der Bearbeitung

Abbildung 6.14: Vergleich zwischen simulierter und gemessener Formänderung

dünnen Bauteilen, welche durch eine Änderung des Eigenspannungszustands hervorgerufen werden. Bei letzterem Sub-Modell wurden die Bauteile als vor der Bearbeitung eigenspannungsfrei angenommen, da so das größtmögliche Optimierungspotential ausgeschöpft werden kann. In der Praxis wurde ein spannungsfreier Zustand durch das Spannungsarmglühen angenähert. Die Minimierung der Formabweichung beider Sub-Modelle wurde auf zwei Optimierungsschritte aufgeteilt, um eine praktikable Rechenzeit zu gewährleisten. Zunächst wurden die Eingriffsbreite, Schnittgeschwindigkeit und Vorschubgeschwindigkeit optimiert und verschiedene Bearbeitungsstrategien miteinander verglichen, um die quellspannungsbedingten Formabweichungen zu minimieren. Es konnte gezeigt werden, dass spannungsfreie Bauteile so bearbeitet werden können, dass nahezu keine Biegungen und Torsion durch das Fräsen resultieren. Die Geometrieabweichungen sind so gering, dass ein besonderes Augenmerk auf die Effizienz der Strategie gelegt werden konnte. Spiralbahnen mit wenigen Einzelprozessen und optimierten Maschinenparametern sind nicht nur besonders ökonomisch, sondern auch gut geeignet um Formabweichungen größtenteils zu verhindern. Es konnte gezeigt werden, dass der schon bei der Parameteridentifikation verwendete SAND-Ansatz für die PDE-restringierte Optimierung auch für die Minimierung von Formabweichungen effizient und zuverlässig ist.

Anschließend dienten die verbliebenen quellspannungsbedingten Formabweichungen zusammen mit den simulierten thermischen Dehnungen als Grundlage für eine zweite Optimierung. Dadurch, dass dafür nur eine einmalige Simulation der thermoelastischen Effekte nötig ist und die Optimierung als Simulationsgrundlage ein performantes kinematisches Werkzeugmodell verwendet, konnte die Berechnungsdauer entscheidend gesenkt werden. Die Minimierung der Gesamtformabweichung erfolgte durch die Optimierung der Schnitttiefen der einzelnen Fräsbahnen, welche als Parabeln modelliert wurden. Durch die Verwendung von konkaven Parabeln mit hohen Scheitelpunkten konnten die thermischen Dehnungen und die restlichen quellspannungsbedingten Verschiebungen hervorragend kompensiert werden. Auch hier war weniger die Wahl der Bearbeitungsstrategie

als die Anpassung der Fräsbahnen entscheidend. Daher sind effiziente Bearbeitungen bei der Wahl der Strategie zu bevorzugen. Letztendlich konnte eine Minimierung der Maximalabweichung von der Idealgeometrie auf etwa  $2\ \mu\text{m}$  realisiert werden. Dies ist nahe der Bearbeitungsgenauigkeit konventioneller Bearbeitungszentren. Die Aufteilung der Optimierung in zwei Schritte scheint in diesem Fall gerechtfertigt, da diese das Optimierungspotential nur geringfügig einschränkt, im Gegenzug aber eine deutliche Senkung der Rechenzeit ermöglicht wurde.

Der Vergleich der Optimierungsergebnisse mit den Resultaten realer Experimente macht deutlich, dass manche Effekte nicht vernachlässigbar sind. So ist das Verhalten des Bauteils durch die magnetische Einspannung in Kombination mit den Prozesskräften möglicherweise relevant für die resultierende Formabweichung. Auch die thermischen Ausdehnungen des Werkzeugs könnten einen signifikanten Anteil haben. Diese und vorherige experimentelle Untersuchungen haben zudem gezeigt, dass die Vorhersagen der einzelnen Sub-Modelle qualitativ und quantitativ zutreffend sind. Formänderungen aufgrund einer Änderung des Eigenspannungszustands und der erhöhte Materialabtrag aufgrund thermischer Dehnungen in Kombination mit einem präzisen kinematischen Werkzeugmodell können gut präzisiert werden. Zudem ermöglicht das Quellspannungsmodell in Kombination mit der Optimierung die Fertigung eines torsionsfreien Bauteils.





# Kapitel 7

## Zusammenfassung und Ausblick

### Inhalt

---

7.1 Zusammenfassung . . . . .	145
7.2 Ausblick . . . . .	147

---

Dieses Kapitel fasst die Arbeit abschließend zusammen und gibt einen Ausblick zu möglichen Folgearbeiten. Ein detaillierteres Fazit zu den numerischen Ergebnissen ist am Ende der jeweiligen Kapitel zu finden.

### 7.1 Zusammenfassung

In der spanenden Bearbeitung ist durch hochpräzise Fertigungsanforderungen bei gleichzeitiger Reduktion oder gänzlichem Auslassen von Kühlschmierung eine Kompensation von Prozesseinflüssen, die zu Ungenauigkeiten führen, unerlässlich. Bei der Betrachtung des thermischen Bauteilverhaltens ist der Wärmeeintrag eine entscheidende Größe, welche aber nicht direkt gemessen werden kann. Ein Verfahren für die inverse Bestimmung von Wärmestromdichten ist für die empirische Modellierung erforderlich. Für die geometrischen Formabweichungen von dünnwandigen Bauteilen ist neben dem thermoelastischen Verhalten auch die Änderung des Eigenspannungszustands von Bedeutung. In der Regel sind beide Modelle sehr komplex und rechenintensiv. Kommen für die Kompensation Optimierungsansätze zum Einsatz, ist deshalb die Effizienz des Verfahrens von besonderer Wichtigkeit.

Die vorliegende Arbeit konnte zu dieser Thematik in drei Bereichen wesentliche Resultate erzielen. Erstens konnten die Wärmestromdichten von Fräs- und Bohrprozessen durch eine PDE-restringierte Parameteridentifikation erfolgreich bestimmt werden. Zweitens konnte ein praxisnaher Vergleich der Optimierungsansätze SAND und NAND die Theorie zur Laufzeit der Ansätze bestätigen. Drittens umfassen die Resultate die Entwicklung und Anwendung eines zweistufigen Optimierungsverfahrens zur prädiktiven Minimierung von Formabweichungen beim Trockenfräsen.

Für die indirekte Bestimmung von Wärmestromdichten beim Fräsen wurde die Parameteridentifikation als PDE-restringiertes Optimierungsproblem durch künstliche verauschte Daten erfolgreich validiert und danach für die Identifikation von Thermografieaufnahmen aus den vorgestellten Experimenten verwendet. Die Optimierung mit dem SAND-Ansatz stellte sich zur Bestimmung des Wärmeeintrags von Realversuchen als effizient und robust heraus. Eine geeignete Gewichtung der Daten ermöglichte dabei zudem die Rekonstruktion von fehlerhaften sowie nicht gemessenen Bereichen. Für die Identifikation von Wärmestromdichten beim Bohren wurde die nichtlineare Variante der Wärmeleitungsgleichung mit temperaturabhängigen Materialparametern entwickelt. Nach der erfolgreichen Validierung der Parameteridentifikation mit nichtlinearem Modell erzielte die Identifikation anhand gemessener Temperaturfelder auch beim Bohren überzeugende Ergebnisse. Da der Versuchsaufbau der Bohrexperimente eine Messung mit weniger Fehlerquellen ermöglichte, ist die optische und numerische Übereinstimmung von rekonstruierten und gemessenen Temperaturen im Vergleich zum Fräsen noch höher. Es konnten mittlere absolute Abweichungen von unter  $1\text{ }^{\circ}\text{C}$  und mittlere relative Abweichungen von unter  $2\%$  erreicht werden. Auch die zusätzliche Identifikation der  $y$ -Position der Aufnahmen wurde äußerst zufriedenstellend gelöst.

Im Kontext der Parameteridentifikation des Bohrens wurde zusätzlich ein Vergleich der Ansätze SAND und NAND durchgeführt. Dabei konnte hinsichtlich der Qualität der Lösungen kein entscheidender Unterschied festgestellt werden. Beim Vergleich der Rechenzeiten ergab sich ein anderes Bild. Hier hatte das SAND gegenüber dem NAND deutliche Vorteile. Der SAND-Ansatz erreichte fast durchweg schnellere Berechnungen als der NAND-Ansatz. Im Extremfall bestand ein Geschwindigkeitsunterschied vom Faktor 90. Aufgrund einer ähnlichen Konvergenzrate der Ansätze konnte der Unterschied auf die Berechnungsdauer der einzelnen Iterationen zurückgeführt werden. Insgesamt entsprechen die Beobachtungen den theoretischen Überlegungen, wonach der NAND-Ansatz mit steigender Anzahl an Identifikationsparametern und bei der Verwendung von Zustandsbeschränkungen deutlich stärker an Effizienz einbüßt als der SAND-Ansatz.

Der letzte Teil der Arbeit behandelt die Kompensation von Formabweichungen beim Trockenfräsen. Für eine effiziente Berechnung wurde eine Optimierung in zwei Schritten durchgeführt. Als erstes erfolgte die Minimierung der quellspannungsbedingten Deformationen durch Optimierung der Eingriffsbreite, Schnittgeschwindigkeit und Vorschubgeschwindigkeit sowie durch einen Vergleich verschiedener Bearbeitungsstrategien. Es wurde beobachtet, dass ein Tordieren des Bauteils bei acht von neun Strategien verhindert werden konnte. Mit steigender Anzahl an Fräsbahnen konnte auch die Biegung fast vollständig eliminiert werden. Die quellspannungsbedingte Maximalabweichung einer trivialen Bearbeitung von  $9,2\text{ }\mu\text{m}$  konnte auf bis zu  $0,1\text{ }\mu\text{m}$  reduziert werden. Die Wahl der Strategie spielte dabei eine untergeordnete Rolle, weshalb effiziente Strategien aus Spiralen mit wenigen Windungen empfehlenswert sind. Für die Optimierung der Gesamtformabweichung wurde nach einer Simulation der thermischen Dehnungen ein kinematisches Werkzeugmodell verwendet, um die Schnitttiefe zu optimieren. Die Ergebnisse zeigen, dass parabelförmige Bahnen sehr gut geeignet sind, um den thermischen Dehnungen entgegenzuwirken. Insgesamt konnte die maximale Gesamtabwei-

chung von  $10,7\ \mu\text{m}$  bei einer Strategie mit konstanten Schnitttiefen auf  $2,6\ \mu\text{m}$  durch Optimierung der Schnitttiefen reduziert werden. Der Vergleich der Ergebnisse mit realen Experimenten ergab, dass mit den einzelnen Submodellen eine gute Vorhersage von Formabweichungen quellspannungsfreier Bauteile möglich ist und in Kombination mit der Optimierung torsionsfreie Bauteile gefertigt werden können. Für die Superposition beider Effekte sollten noch weitere Einflüsse wie bspw. die Einspannung des Bauteils oder die Werkzeugdehnung berücksichtigt werden.

## 7.2 Ausblick

Neben den diskutierten wissenschaftlichen Resultaten bietet diese Arbeit auch Grundlage für weitere Forschung.

Bei der Parameteridentifikation von Fräsprozessen wurde im Vergleich zu früheren Verfahren (siehe Kapitel 1.3) ein Großteil der Daten berücksichtigt. Ein kleiner Teil davon blieb aber weiterhin unberücksichtigt, da Störungen die Qualität der Messungen beeinträchtigten. Interessant wäre, ob Thermografieaufnahmen mit geringeren Messfehlern eine höhere Übereinstimmung über die gesamte Aufnahme ermöglichen. Dafür wäre zu untersuchen, inwieweit ein weniger störanfälliger Versuchsaufbau zu realisieren ist.

Die Implementierung des SAND- und NAND-Ansatzes für die Identifikation von Bohrprozessen erfolgte unter Berücksichtigung einer effizienten Umsetzung, insbesondere durch Berücksichtigung analytischer Ableitungen auf SQP-Ebene beim SAND-Ansatz und auf Ebene des Newton-Verfahrens beim NAND-Ansatz. Es existieren noch weitere Varianten bei der Umsetzung und Implementierung der Ansätze (siehe Kapitel 1.3). Ein Rechenzeitvergleich des NAND mit dem Adjoint Approach oder des SAND-Ansatzes mit einem geeignet vorkonditionierten iterativen Lösungsverfahren der linearen Algebra wären vielversprechende Folgeuntersuchungen.

Die Minimierung von Formabweichungen beim Fräsen wurde in zwei Schritten realisiert. Dadurch wurde eine besonders effiziente Optimierung ermöglicht und Formabweichungen von über  $10\ \mu\text{m}$  konnten nahezu eliminiert werden. Es sind jedoch spanende Verfahren und Werkzeug-Werkstück-Kombinationen denkbar, bei denen mit einer größeren Formabweichung zu rechnen ist. In diesen Fällen wäre die Implementierung einer Optimierung, die alle Effekte und Maschinenparameter gleichzeitig berücksichtigt, interessant. Dafür wäre es insbesondere wichtig, auf eine performante Umsetzung zu achten und besonders leistungsstarke Rechensysteme zu verwenden.

Durch die Umsetzung der optimierten Frässtrategien in realen Experimenten konnten die Vorhersagemodelle teilweise validiert werden. Damit die simulierten Formabweichungen noch besser auf reale Experimente übertragen werden können, ist eine Erweiterung der verwendeten Modelle sinnvoll. So könnte die Art der Einspannung einen Einfluss auf die Bearbeitung haben. Auch die Berücksichtigung von Deformationen des Werkzeugs könnten die Modellierungsgüte steigern. Des Weiteren könnte die Verwendung eines Modells,

das vor der Bearbeitung inhärente Quellspannungen berücksichtigt (siehe bspw. [51]), die Vorhersagegenauigkeit weiter verbessern.

Durch die Bestimmung des Wärmeeintrags von Bohrprozessen wurde in dieser Arbeit ein grundlegender Beitrag für die Modellierung von thermischen Formabweichungen beim Bohren geschaffen. Darauf aufbauend ist eine Minimierung der Formabweichungen denkbar. Wird dabei ein langer Zeithorizont mit mehreren Bohrungen in Reihe betrachtet, ist besonders auf eine effiziente Realisierung der Optimierung zu achten. Bei der Verwendung von diskreten Variablen, z. B. bei der Optimierung von Bohrreihenfolgen, könnten die Methoden der gemischt-ganzzahligen Optimierung verwendet werden.

Die Diskretisierung der PDEs erfolgte in dieser Arbeit mit lokal verfeinerten Gittern. Die Auflösung wurde dabei an einer performanten Optimierung ausgerichtet. Eine mögliche Erweiterung dieses Ansatzes ist eine adaptive Verfeinerung der Gitter auf der Grundlage eines FEM-Fehlerschätzers. In Kombination mit einer Multilevel-Optimierung [15, 116] wäre dadurch eine interessante Folgearbeit gegeben.

All diese Perspektiven zeigen auf, dass aufbauend auf dieser Arbeit noch weiteres Forschungspotential in der PDE-restringierten Optimierung und der Optimierung von spanenden Fertigungsverfahren vorhanden ist.

# Literaturverzeichnis

- [1] *ABAQUS UNIFIED FEA*. <https://www.3ds.com/de/produkte-und-services/simulia/produkte/abaqus/>, 2017. – [Online; abgerufen am 29. März 2017]
- [2] ABRAHAM, F. ; BEHR, M. ; HEINKENSCHLOSS, M.: The effect of stabilization in finite element methods for the optimal boundary control of the Oseen equations. In: *Finite Elements in Analysis and Design* 41 (2004), Nr. 3, S. 229 – 251
- [3] ALIFANOV, O. M.: *Inverse Heat Transfer Problems*. Berlin, Heidelberg : Springer Berlin Heidelberg, 1994
- [4] ALTROGGE, I. ; BÜSKENS, C. ; KRÖGER, T. ; PEITGEN, H.-O. ; PREUSSER, T. ; TIESLER, H.: Modeling, Simulation and Optimization of Radio Frequency Ablation. In: *Constrained Optimization and Optimal Control for Partial Differential Equations*. Basel : Springer, 2012, S. 487–499
- [5] ARNDT, D. ; BANGERTH, W. ; DAVYDOV, D. ; HEISTER, T. ; HELTAI, L. ; KRONBICHLER, M. ; MAIER, M. ; PELTERET, J.-P. ; TURCK SIN, B. ; WELLS, D.: The deal.II Library, Version 8.5. In: *Journal of Numerical Mathematics* (2017)
- [6] ARORA, J. S. ; WANG, Q.: Review of formulations for structural and mechanical system optimization. In: *Structural and Multidisciplinary Optimization* 30 (2005), Nr. 4, S. 251–272
- [7] BANGERTH, W. ; HARTMANN, R. ; KANSCHAT, G.: deal.II – a General Purpose Object Oriented Finite Element Library. In: *ACM Trans. Math. Softw.* 33 (2007), Nr. 4, S. 1–27
- [8] BARGEL, H.-J.: *Werkstoffkunde*. 11. Aufl. Berlin : Springer, 2012 (Springer-Lehrbuch)
- [9] BECK, J. V. ; BLACKWELL, B. ; CLAIR, C. R. S.: *Inverse Heat Conduction: Ill-Posed Problems*. New York : Wiley, 1985
- [10] BHARATHI RAJA, S. ; BASKAR, N.: Particle swarm optimization technique for determining optimal machining parameters of different work piece materials in turning operation. In: *The International Journal of Advanced Manufacturing Technology* 54 (2011), Nr. 5, S. 445–463
- [11] BIEGLER, L. T. ; HEINKENSCHLOSS, M. ; GHATTAS, O. ; BLOEMEN WAANDERS, B. van (Hrsg.): *Large-Scale PDE-Constrained Optimization*. Berlin, Heidelberg : Springer, 2003 (Lecture Notes in Computational Science and Engineering)

- [12] BIERMANN, D. ; IOVKOV, I. ; BLUM, H. ; RADEMACHER, A. ; TAEBI, K. ; SUTTMEIER, F. T. ; KLEIN, N.: Thermal Aspects in Deep Hole Drilling of Aluminium Cast Alloy Using Twist Drills and MQL. In: *Procedia CIRP* 3 (2012), S. 245–250
- [13] BIERMANN, D. ; HOLLMANN, F. (Hrsg.): *Thermal Effects in Complex Machining Processes: Final Report of the DFG Priority Programme 1480*. Cham : Springer International Publishing, 2018 (Lecture Notes in Production Engineering)
- [14] *Bild zur freien Nutzung (Creative Commons Lizenz), Fotograf: Beeki*. [https://cdn.pixabay.com/photo/2013/02/22/08/20/drilling-84848\\_960\\_720.jpg](https://cdn.pixabay.com/photo/2013/02/22/08/20/drilling-84848_960_720.jpg), 2013. – [Online; abgerufen am 12. März 2018]
- [15] BORZI, A. ; SCHULZ, V.: Multigrid Methods for PDE Optimization. In: *SIAM Review* 51 (2009), Nr. 2, S. 361–395
- [16] BRAESS, D.: *Finite Elemente - Theorie, schnelle Löser und Anwendungen in der Elastizitätstheorie*. Berlin, Heidelberg : Springer, 2007
- [17] BRANDENBURG, C. ; LINDEMANN, F. ; ULBRICH, M. ; ULBRICH, S.: Advanced Numerical Methods for PDE Constrained Optimization with Application to Optimal Design in Navier Stokes Flow. In: *Constrained Optimization and Optimal Control for Partial Differential Equations*. Basel : Springer, 2012, S. 257–275
- [18] BRINKSMEIER, E. ; HEINZEL, C.: Aufgaben und Auswahl der Kühlschmierstoffe. In: *DIF-Tagung*, Deutsches Industrie Forum für Technologie, 1995, S. 1–32
- [19] BRINKSMEIER, E. ; SÖLTER, J.: Prediction of shape deviations in machining. In: *CIRP Annals - Manufacturing Technology* 58 (2009), Nr. 1, S. 507–510
- [20] BRINKSMEIER, E. ; SÖLTER, J. ; GROTE, C.: Distortion Engineering – Identification of Causes for Dimensional and Form Deviations of Bearing Rings. In: *CIRP Annals* 56 (2007), Nr. 1, S. 109 – 112
- [21] BÜSKENS, C.: *Optimierungsmethoden und Sensitivitätsanalyse für optimale Steuerprozesse mit Steuer- und Zustands-Beschränkungen*, Westfälische Wilhelms-Universität Münster, Diss., 1998
- [22] BÜSKENS, C.: *Optimierung*. 2008. – unveröffentlichtes Vorlesungsskript, Universität Bremen
- [23] BÜSKENS, C.: *Optimale Steuerung und Regelung dynamischer Systeme*. 2010. – unveröffentlichtes Vorlesungsskript, Universität Bremen
- [24] BÜSKENS, C. ; WASSEL, D.: The ESA NLP solver WORHP. In: *Modeling and Optimization in Space Engineering* Bd. 73. Berlin : Springer, 2013 (Optimization and Its Applications), S. 85–110
- [25] BYRD, R. H. ; NOCEDAL, J.: An analysis of reduced Hessian methods for constrained optimization. In: *Mathematical Programming* 49 (1990), Nr. 1, S. 285–323

- [26] CHEN, S. Q. Q.: *Datenbasierte Modellierung und Optimierung von Kraft-Wärme-Kopplungsanlagen*, Universität Bremen, Diss., 2017
- [27] CHEN, W. ; NI, L. ; XUE, J.: Deformation control through fixture layout design and clamping force optimization. In: *The International Journal of Advanced Manufacturing Technology* 38 (2007), Nr. 9, S. 860–867
- [28] COLAÇO, M. J. ; ORLANDE, H. R. B. ; DULIKRAVICH, G. S.: Inverse and optimization problems in heat transfer. In: *Journal of the Brazilian Society of Mechanical Sciences and Engineering* 28 (2006), S. 1 – 24
- [29] CURTIS, F. E.: A penalty-interior-point algorithm for nonlinear constrained optimization. In: *Mathematical Programming Computation* 4 (2012), Nr. 2, S. 181–209
- [30] *deal.II — an open source finite element library*. <http://www.dealii.org/>, 2017. – [Online; abgerufen am 08. Mai 2017]
- [31] DENKENA, B. ; SCHMIDT, A. ; MAASS, P. ; NIEDERWESTBERG, D. ; NIEBUHR, C. ; VEHMEYER, J.: Prediction of Temperature Induced Shape Deviations in dry Milling. In: *Procedia CIRP* 31 (2015), S. 340 – 345. – 15th CIRP Conference on Modelling of Machining Operations (15th CMMO)
- [32] DENKENA, B. ; TÖNSHOFF, H. K.: *Spanen: Grundlagen*. 3., bearb. und erw. Aufl. Berlin, Heidelberg : Springer, 2011 (VDI-Buch)
- [33] DEPPERMAN, M. ; KNEER, R.: Determination of the heat flux to the workpiece during dry turning by inverse methods. In: *Production Engineering* 9 (2015), Nr. 4, S. 465–471
- [34] DYCK, M.: *Beitrag zur Analyse thermisch bedingter Werkstückdeformationen in Trockenbearbeitungsprozessen*, Universität Karlsruhe, Diss., 2007
- [35] ECHIM, M.: *Modellbasierte optimale Mehrgrößenregelung eines aufgeladenen Dieselmotors mittels Methoden der nichtlinearen Optimierung*, Universität Bremen, Diss., 2011
- [36] EHRHARDT, K. ; STEINBACH, M. C.: Nonlinear Optimization in Gas Networks. In: *Modeling, Simulation and Optimization of Complex Processes*. Berlin, Heidelberg : Springer, 2005, S. 139–148
- [37] ESLER, K. P.: *einspline*. <http://www.einspline.sourceforge.net/index.shtml>, 2017. – [Online; abgerufen am 06. November 2017]
- [38] FORSCHUNGSGEMEINSCHAFT, Deutsche: *SPP 1480: Modellierung, Simulation und Kompensation von thermischen Bearbeitungseinflüssen für komplexe Zerspanprozesse*. <http://www.cutsim.de/de/index.html>, <http://gepris.dfg.de/gepris/projekt/130379584>, 2017. – [Online; abgerufen am 20. April 2018]
- [39] FORSGREN, A. ; GILL, P. E. ; WRIGHT, M. H.: Interior methods for nonlinear optimization. In: *SIAM Review* 44 (2002), S. 525–597

- [40] GARDAREIN, J.-L. ; CORRE, Y. ; RIGOLLET, F. ; NILIOT, C. L. ; REICHLER, R. ; ANDREW, P.: Thermal quadrupoles approach for two-dimensional heat flux estimation using infrared and thermocouple measurements on the JET tokamak. In: *International Journal of Thermal Sciences* 48 (2009), Nr. 1, S. 1 – 13
- [41] GEFFKEN, S.: *Effizienzsteigerung numerischer Verfahren der nichtlinearen Optimierung*, Universität Bremen, Diss., 2017
- [42] GEFFKEN, S. ; BÜSKENS, C.: Feasibility refinement in sequential quadratic programming using parametric sensitivity analysis. In: *Optimization Methods and Software* (2016)
- [43] GEFFKEN, S. ; BÜSKENS, C.: Worhp Multi-Core Interface, Parallelisation Approaches for an NLP Solver. In: *6th International Conference on Astrodynamics Tools and Techniques (ICATT)*, Darmstadt, 2016, S. 1–7
- [44] GEIGER, C. ; KANZOW, C.: *Numerische Verfahren zur Lösung unrestringierter Optimierungsaufgaben*. Springer, 1999 (Springer-Lehrbuch)
- [45] GEIGER, C. ; KANZOW, C.: *Theorie und Numerik restringierter Optimierungsaufgaben*. Berlin : Springer, 2002
- [46] GILL, P. E. ; WONG, E.: Sequential Quadratic Programming Methods. In: *Mixed Integer Nonlinear Programming* Bd. 154. Springer New York, 2012, S. 147–224
- [47] GULPAK, M. ; SÖLTER, J.: Development and validation of a hybrid model for the prediction of shape deviations in dry machining processes. In: *Procedia CIRP* 31 (2015), S. 346–351
- [48] GULPAK, M. ; SÖLTER, J.: Development of a hybrid model for the prediction of shape deviations in milling. In: *Materialwissenschaft und Werkstofftechnik* 47 (2016), Nr. 8, S. 718–725
- [49] GULPAK, M. ; SÖLTER, J.: Thermal Modelling of Drilling Steel. In: *WGP Congress 2016* Bd. 1140, Trans Tech Publications, 2016 (Advanced Materials Research), S. 205–212
- [50] GULPAK, M. ; SÖLTER, J. ; BRINKSMEIER, E.: Prediction of Shape Deviations in Face Milling of Steel. In: *Procedia CIRP* 8 (2013), S. 15–20
- [51] GULPAK, M. ; WERNESING, H. ; SÖLTER, J. ; BÜSKENS, C.: Compensation Strategies for Thermal Effects in Dry Milling. In: *Thermal Effects in Complex Machining Processes: Final Report of the DFG Priority Programme 1480*. Cham : Springer International Publishing, 2018 (Lecture Notes in Production Engineering), S. 251–288
- [52] HABER, E. ; ASCHER, U. M.: Preconditioned all-at-once methods for large, sparse parameter estimation problems. In: *Inverse Problems* 17 (2001), Nr. 6



- [53] HAO, W. ; HONGTAO, Z. ; QIANJIAN, G. ; XIUSHAN, W. ; JIANGUO, Y.: Thermal error optimization modeling and real-time compensation on a CNC turning center. In: *Journal of Materials Processing Technology* 207 (2008), Nr. 1, S. 172 – 179
- [54] HEINZEL, C. ; SÖLTER, J. ; GULPAK, M. ; RIEMER, O.: An analytical multilayer source stress approach for the modelling of material modifications in machining. In: *CIRP Annals - Manufacturing Technology* 66 (2017), S. 531–534
- [55] HERZOG, R. ; KUNISCH, K.: Algorithms for PDE-constrained optimization. In: *GAMM-Mitteilungen* 33 (2010), Nr. 2, S. 163–176
- [56] HINZE, M. ; PINNAU, R. ; ULBRICH, M. ; ULBRICH, S.: *Optimization with PDE Constraints*. Springer, 2009 (Mathematical Modelling)
- [57] HINZE, M. ; TRÖLTZSCH, F.: Discrete concepts versus error analysis in PDE-constrained optimization. In: *GAMM-Mitteilungen* 33 (2010), Nr. 2, S. 148–162
- [58] HUANG, C.-H. ; JAN, L.-C. ; L., R. ; SHIH, A. J.: A three-dimensional inverse problem in estimating the applied heat flux of a titanium drilling – Theoretical and experimental studies. In: *International Journal of Heat and Mass Transfer* 50 (2007), Nr. 17, S. 3265 – 3277
- [59] IOVKOV, I.: *Technologische Analyse des Tiefbohrens mit Minimalmengenschmierung und simulationsbasierte Kompensation des Mittenverlaufs*, Technische Universität Dortmund, Diss., 2016
- [60] *Ipopt - Interior Point OPTimizer*. <https://projects.coin-or.org/Ipopt>, 2017. – [Online; abgerufen am 16. März 2017]
- [61] ITLE, G. C. ; SALINGER, A. G. ; PAWLOWSKI, R. P. ; SHADID, J. N. ; BIEGLER, L. T.: A tailored optimization strategy for PDE-based design: application to a CVD reactor. In: *Computers & Chemical Engineering* 28 (2004), Nr. 3, S. 291–302
- [62] KARMARKAR, N.: A New Polynomial-time Algorithm for Linear Programming. In: *Proceedings of the Sixteenth Annual ACM Symposium on Theory of Computing*. New York : ACM, 1984, S. 302–311
- [63] KIENZLER, R. ; SCHRÖDER, R.: *Einführung in die Höhere Festigkeitslehre*. Heidelberg : Springer, 2009
- [64] KIM, H. K. ; LEE, J. H. ; HIELSCHER, A. H.: PDE-Constrained Fluorescence Tomography With the Frequency-Domain Equation of Radiative Transfer. In: *IEEE Journal of Selected Topics in Quantum Electronics* 16 (2010), Nr. 4, S. 793–803
- [65] KLOCKE, F. ; KNEER, R. ; BURGHOLD, M. ; DEPPERMAN, M. ; PENG, B. ; PULS, H.: Modelling and Compensation of Thermoelastic Workpiece Deformation in Dry Cutting. In: *Thermal Effects in Complex Machining Processes: Final Report of the DFG Priority Programme 1480*. Cham : Springer International Publishing, 2018 (Lecture Notes in Production Engineering), S. 63–94

- [66] KNAUER, M. ; BÜSKENS, C.: From WORHP to TransWORHP. In: *Proceedings of the 5th International Conference on Astrodynamics Tools and Techniques*. Noordwijk, 2012 (5th International Conference on Astrodynamics Tools and Techniques)
- [67] KRAUSE, D.: *Parameteridentifikation bei partiellen Differentialgleichungen*. 2015. – unveröffentlichte Masterarbeit, Universität Bremen
- [68] KUHLMANN, R. ; GEFFKEN, S. ; BÜSKENS, C.: WORHP Zen: Parametric Sensitivity Analysis for the Nonlinear Programming Solver WORHP. In: *Operations Research Proceedings*. Berlin, 2017 (International Conference on Operations Research 2017). – [zur Veröffentlichung angenommen]
- [69] LANGENHORST, L. ; GULPAK, M. ; SÖLTER, J. ; RIEMER, O.: Effects of model reduction on simulated temperature fields in milling. In: *Procedia CIRP* 58 (2017), S. 511–516
- [70] LEUGERING, G. ; ENGELL, S. ; GRIEWANK, A. ; HINZE, M. ; RANNACHER, R. ; SCHULZ, V. ; ULBRICH, M. ; ULBRICH, S. (Hrsg.): *Constrained Optimization and Optimal Control for Partial Differential Equations*. Basel : Springer, 2012 (International Series of Numerical Mathematics)
- [71] LEWIS, R. W. ; MORGAN, K. ; THOMAS, H. R. ; SEETHARAMU, K.: *The Finite Element Method in Heat Transfer Analysis*. John Wiley & Sons, 1996
- [72] LEYKEKHMAN, D.: Investigation of Commutative Properties of Discontinuous Galerkin Methods in PDE Constrained Optimal Control Problems. In: *Journal of Scientific Computing* 53 (2012), Nr. 3, S. 483–511
- [73] LUDWIG, R.: *Methoden der Fehler- und Ausgleichsrechnung: Lehrbuch für Studenten aller naturwissenschaftlichen und technischen Fachrichtungen ab 3. Semester*. Wiesbaden : Vieweg+Teubner, 1969
- [74] MING, C. ; FANGHONG, S. ; HAILI, W. ; RENWEI, Y. ; ZHENGHONG, Q. ; SHUQIAO, Z.: Experimental research on the dynamic characteristics of the cutting temperature in the process of high-speed milling. In: *Journal of Materials Processing Technology* 138 (2003), Nr. 1–3, S. 468–471
- [75] NIEDERWESTBERG, D.: *Prognose und Kompensation der temperaturbedingten Werkstückverlagerungen beim Trockenfräsen*, Leibniz Universität Hannover, Diss., 2016
- [76] NOCEDAL, J. ; WRIGHT, S. J.: *Numerical optimization*. 2. ed. New York : Springer, 2006 (Springer series in operation research and financial engineering)
- [77] ÖZISIK, M. N. ; ORLANDE, H. R. B.: *Inverse Heat Transfer: Fundamentals and Applications*. New York : Taylor & Francis, 2000
- [78] PABST, R.: *Mathematische Modellierung der Wärmestromdichte zur Simulation des thermischen Bauteilverhaltens bei der Trockenbearbeitung*, Universität Karlsruhe, Diss., 2009

- [79] PIRONNEAU, O.: On optimum design in fluid mechanics. In: *Journal of Fluid Mechanics* 64 (1974), Nr. 1, S. 97–110
- [80] POURGHOLI, R. ; ROSTAMIAN, M.: A numerical technique for solving IHCPs using Tikhonov regularization method. In: *Applied Mathematical Modelling* 34 (2010), Nr. 8, S. 2102 – 2110
- [81] POWELL, M. J. D.: The Convergence of Variable Metric Methods for Nonlinearly Constrained Optimization Calculations. In: *Nonlinear Programming 3*. New York : Academic Press, 1978, S. 27–63
- [82] POWELL, M. J. D.: A fast algorithm for nonlinearly constrained optimization calculations. In: *Numerical Analysis*. Berlin, Heidelberg : Springer, 1978, S. 144–157
- [83] REES, T. ; STOLL, M. ; WATHEN, A.: All-at-once preconditioning in PDE-constrained optimization. In: *Kybernetika* 46 (2010), Nr. 2, S. 341–360
- [84] RICHARDSON, D. J. ; KEAVEY, M. A. ; DAILAMI, F.: Modelling of cutting induced workpiece temperatures for dry milling. In: *International Journal of Machine Tools and Manufacture* 46 (2006), Nr. 10, S. 1139–1145
- [85] RICHTER, T.: *Die Finite Elemente Methode für partielle Differentialgleichungen*. <http://www.numerik.uni-hd.de/~richter/pdf-files/numerik2.pdf>. Version: 2012. – Vorlesungsskript, Universität Heidelberg
- [86] ROZZA, G. ; MANZONI, A. ; NEGRI, F.: Reduction strategies for PDE-constrained optimization problems in haemodynamics. In: *Proceedings of the 6th European Congress on Computational Methods in Applied Sciences and Engineering*. Wien : Technische Universität Wien, 2012, S. 1748–1769
- [87] SADD, M. H.: *Elasticity: Theory, Applications, and Numerics*. Burlington : Elsevier Science, 2009
- [88] SCHÖBERL, J. ; ZULEHNER, W.: Symmetric Indefinite Preconditioners for Saddle Point Problems with Applications to PDE-Constrained Optimization Problems. In: *SIAM J. Matrix Anal. Appl.* 29 (2007), Nr. 3, S. 752–773
- [89] SCHITTKOWSKI, K.: *Numerical Data Fitting in Dynamical Systems: A Practical Introduction with Applications and Software*. Norwell, MA, USA : Kluwer Academic Publishers, 2002
- [90] SCHMIDT, A.: *Numerik partieller Differentialgleichungen*. 2010. – unveröffentlichte Vorlesungsmitschrift, Universität Bremen
- [91] SCHWEIZER, B.: *Partielle Differentialgleichungen: Eine anwendungsorientierte Einführung*. Berlin, Heidelberg : Springer, 2013
- [92] SÖLTER, J.: *Ursachen und Wirkmechanismen der Entstehung von Verzug infolge spanender Bearbeitung*, Universität Bremen, Diss., 2010. – Shaker Verlag Aachen

- [93] SÖLTER, J. ; GULPAK, M.: Heat partitioning in dry milling of steel. In: *CIRP Annals - Manufacturing Technology* 61 (2012), Nr. 1, S. 87–90
- [94] SÖLTER, J. ; GULPAK, M. ; BRINKSMIEIER, E.: Modellentwicklung zur Minimierung von Geometrieabweichungen in der Trockenbearbeitung. In: *Zeitschrift für wirtschaftlichen Fabrikbetrieb* 107 (2012), Nr. 4, S. 224–228
- [95] SÖLTER, J. ; FROHMÜLLER, R. ; WIRBSER, H.: Temperature Measurements and Heat Partitioning in Machining Processes. In: *Thermal Effects in Complex Machining Processes: Final Report of the DFG Priority Programme 1480*. Cham : Springer International Publishing, 2018 (Lecture Notes in Production Engineering), S. 5–21
- [96] STOLL, M. ; WATHEN, A.J.: All-at-once solution of time-dependent Stokes control. In: *Journal of Computational Physics* 232 (2013), Nr. 1, S. 498–515
- [97] TAI, L.-J.: *Thermal Modeling of Workpiece Temperature and Distortion in MQL Deep-Hole Drilling*, University of Michigan, Diss., 2011
- [98] TIESLER, H.: *Identification of Material Parameters from Temperature Measurements in Radio Frequency Ablation*, Universität Bremen, Diss., 2011
- [99] TIKHONOV, A. N. ; ARSEININ, V. I. A.: *Solutions of ill-posed problems*. Washington D.C. : Winston, 1977
- [100] TIMOSHENKO, S. ; WOINOWSKY-KRIEGER, S.: *Theory of plates and shells*. 2nd ed. New York : McGraw-Hill, 1959
- [101] TÖNSHOFF, H. K.: *Eigenspannungen und plastische Verformungen im Werkstück durch spanende Bearbeitung*, TH Hannover, Diss., 1966
- [102] TRÖLTZSCH, F.: *Optimale Steuerung partieller Differentialgleichungen: Theorie, Verfahren und Anwendungen*. 2. Aufl. Wiesbaden : Vieweg + Teubner, 2009
- [103] VEREIN DEUTSCHER EISENHÜTTENLEUTE (Hrsg.): *Taschenbuch der Stahl-Eisen-Werkstoffblätter*. Düsseldorf : Verlag Stahleisen GmbH, 1997
- [104] WASSEL, D.: *Exploring novel designs of NLP solvers: Architecture and Implementation of WORHP*, Universität Bremen, Diss., 2013
- [105] WÄCHTER, A. ; BIEGLER, L. T.: On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. In: *Mathematical Programming* 106 (2006), Nr. 1, S. 25–57
- [106] WEINERT, K. ; INASAKI, I. ; SUTHERLAND, J. W. ; WAKABAYASHI, T.: Dry Machining and Minimum Quantity Lubrication. In: *CIRP Annals* 53 (2004), Nr. 2, S. 511–537
- [107] WERNING, H. ; BÜSKENS, C.: Parameter Identification for Finite Element Based Models in Dry Machining Applications. In: *Procedia CIRP* 31 (2015), S. 328–333

- [108] WERNING, H. ; GULPAK, M. ; BÜSKENS, C. ; SÖLTER, J. ; BRINKSMEIER, E.: Enhanced method for the evaluation of the thermal impact of dry machining processes. In: *Production Engineering* 8 (2014), Nr. 3, S. 291–300
- [109] WILSON, R. B.: *A Simplicial Method for Convex Programming*, Harvard University, Diss., 1963
- [110] WOLFF, M.: *Partielle Differentialgleichungen und Sobolev-Räume*. 2016. – unveröffentlichtes Vorlesungsskript, Universität Bremen
- [111] *WORHP - Large-scale Sparse Nonlinear Optimization*. <http://www.worhp.de>, 2017. – [Online; abgerufen am 10. März 2017]
- [112] YU, J.-H. ; CHEN, Z.-T. ; JIANG, Z.-P.: A control process for machining distortion by using an adaptive dual-sphere fixture. In: *The International Journal of Advanced Manufacturing Technology* 86 (2016), Nr. 9, S. 3463–3470
- [113] ZAIN, A.M. ; HARON, H. ; SHARIF, S.: Application of GA to optimize cutting conditions for minimizing surface roughness in end milling machining process. In: *Expert Systems with Applications* 37 (2010), Nr. 6, S. 4650 – 4659
- [114] ZHANG, H. ; WANG, W. ; ZHOU, D. ; MA, F. ; LU, B. ; ZHOU, L.: A Study for Initial Solidification of Sn-Pb Alloy during Continuous Casting: Part I. The Development of the Technique. In: *Metallurgical and Materials Transactions B* 45 (2014), Nr. 3, S. 1038–1047
- [115] ZHANG, H. ; WANG, W. ; ZHOU, L.: Calculation of Heat Flux Across the Hot Surface of Continuous Casting Mold Through Two-Dimensional Inverse Heat Conduction Problem. In: *Metallurgical and Materials Transactions B* 46 (2015), Nr. 5, S. 2137–2152
- [116] ZIEMS, J. C. ; ULBRICH, S.: Adaptive Multilevel Inexact SQP Methods for PDE-Constrained Optimization. In: *SIAM Journal on Optimization* 21 (2011), Nr. 1, S. 1–40
- [117] ZIMMERMANN, M. ; SCHINDLER, S. ; STEINMANN, P. ; AURICH, J. C.: Compensation of Thermo-mechanically Induced Workpiece and Tool Deformations During Dry Turning. In: *Procedia CIRP* 46 (2016), S. 31 – 34