

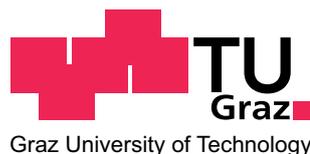
# Optische Vermessung von Halbleiter-Wafern und effiziente Messdatenverarbeitung

Diplomarbeit

durchgeführt von

Thomas Jerman

Institut für Breitbandkommunikation  
der Technischen Universität Graz



Leiter: Univ.-Prof. Dr. Gernot Kubin

Betreuer: Ao. Univ.-Prof. Dr. Erich Leitgeb

Graz, im März 2011

Deutsche Fassung:  
Beschluss der Curricula-Kommission für Bachelor-, Master- und Diplomstudien vom 10.11.2008  
Genehmigung des Senates am 1.12.2008

## EIDESSTÄTTLICHE ERKLÄRUNG

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommene Stellen als solche kenntlich gemacht habe.

Graz, am .....

.....  
(Unterschrift)

Englische Fassung:

## STATUTORY DECLARATION

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

.....  
date

.....  
(signature)

# Zusammenfassung

Die vorliegende Arbeit widmet sich der optischen Vermessung von *Wafern* in der Halbleiterindustrie, sowie der effizienten Verarbeitung von Messdaten zum Zwecke der Positionskorrektur von ebenen als auch von gekrümmten Dünnpwafern. Herkömmliche optische Messsysteme scannen ebene *Wafer* nur in einer Dimension. Im Gegensatz zu Standardwafern haben Dünnpwafer jedoch die Eigenschaft, sich durch innere mechanische Spannungen zu verkrümmen. Mit steigender Waferkrümmung werden die vom optischen System vermessenen Waferkanten mehr und mehr verzerrt abgebildet. Mit Hilfe eines *2D-Wafer-scanners* [1] kann nicht nur die Position der Waferkante, sondern auch die Krümmung des *Wafers* erfasst werden.

In dieser Arbeit steht eine effiziente Messdatenverarbeitung im Vordergrund, welche als Mindestvorgabe Toleranzen von  $\pm 10 \mu\text{m}$  für die Positioniergenauigkeit und  $\pm 0,1^\circ$  für die Winkelausrichtung erreichen soll. Zusätzlich werden die notwendigen Erfordernisse für ein effizientes *Waferprealignment* hinsichtlich seiner mechanischen Umsetzung ergründet. Bei der Entwicklung der verschiedenen Algorithmen zur Messdatenverarbeitung wird neben der Schätzung der Messkurvenparameter auch eine Qualitätsbewertung der Parameterschätzungen durchgeführt. Allen Algorithmen werden zur Beurteilung der *Performance* die maximal erreichbaren Varianzen der geschätzten positionsbeschreibenden Parameter zum objektiven Vergleich gegenübergestellt. Abschließend wird ein Konzept vorgeschlagen, welches sowohl Standard- als auch Dünnpwafer gleichermaßen vermessen und positionieren kann, wobei aktuell übliche Formen und Ausführungen von Wafertypen universell berücksichtigt sind.

# Abstract

This thesis deals with optical measurement of wafers in the semiconductor industry and efficient processing of measured data in order to correct the position of flat and deflected thin wafers. Conventional optical measurement systems scan flat wafers in just one dimension. However, in contrast to standard wafers, thin wafers tend to deflect due to internal mechanical tensions. The more they deflect, the higher the distortion of the optical projection of the wafer edges measured by the optical system. A 2D-Waferscanner [1] is able to measure not only the position of the wafer edge, but also determines the curvature of the wafer.

This thesis examines the necessary conditions for efficient wafer prealignment in mechanical terms, based on the efficient processing of measured data with desired target tolerances of  $\pm 10 \mu\text{m}$  for the positioning and  $\pm 0.1^\circ$  for angle alignment. Different algorithms are developed to process the measured data, curve parameters are estimated, and the quality of these parameter estimates is assessed. For an objective comparison, position describing parameters are furthermore contrasted with values for their maximum possible accuracies. Finally, a method is proposed which is able to handle the measurement and positioning of both standard wafers and thin wafers and which is applicable for all currently prevailing wafer types and forms.

# Danksagung

Mit Fertigstellung dieser Arbeit blicke nicht nur ich auf eine aufregende Studienzeit zurück, sondern auch jene, die mich mit ihrer wertvollen Unterstützung auf diesem Weg begleitet haben. Ich möchte daher meinen aufrichtigen Dank Herrn Dr. Erich Leitgeb aussprechen, der mich bei meinem selbst definierten Diplomarbeitsthema genauso engagiert betreut hat, wie es auch alle anderen Studenten erfahren haben.

Wie es der Zufall wollte, besuchte ich im Jahr 2010 die Lehrveranstaltung *Statistical Signal Processing* bei Herrn Dr. Markus Brandner vom Institut für elektrische Messtechnik und Messsignalverarbeitung. Nicht selten erlaubte ich mir am Ende von Herrn Dr. Brandners Vorlesungen, ihn mit detaillierten Fragen zum Stoff am pünktlichen Verlassen des Hörsaales zu hindern. Ich möchte mich daher herzlich für die vielen Diskussionen und Hilfestellungen begleitend zu meiner Arbeit bedanken.

Nicht unerwähnt soll mein Dank an meinen besten Freund Dipl.-Ing. Robin Priewald bleiben, der mich mit viel Geduld, wertvollen Tipps und praktischen Hinweisen während des Verfassens dieser Arbeit begleitet hat. Er wird stets ein Vorbild für mich sein! Für das Lektorat in deutscher und englischer Sprache bedanke ich mich herzlich bei meiner Freundin Mag<sup>a</sup> Barbara Ranzenhofer, die ich über alles schätze.

Ich bedanke mich auch bei Herrn Heinz Kanduth und Herrn Thomas Nuck von der Firma Semicore GmbH in Villach für die Möglichkeit zur praktischen Umsetzung meiner Arbeit.

Zu guter Letzt schließe ich meine Danksagung mit lieben Grüßen an meine Familie, welche die gesamte Zeit über uneingeschränkt hinter mir gestanden hat. Vielen Dank!

Graz, im März 2011

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
1.1	Hintergrund . . . . .	1
1.2	Motivation . . . . .	2
<b>2</b>	<b>Wafertypen</b>	<b>4</b>
2.1	Waferspezifikationen . . . . .	5
2.2	Modellierung von <i>Waferscans</i> . . . . .	7
2.2.1	<i>Flatwafer</i> . . . . .	8
2.2.2	<i>Wafer mit Notch</i> . . . . .	12
2.2.3	<i>Wafer mit rundem Notch</i> . . . . .	12
<b>3</b>	<b>Circle Fitting</b>	<b>13</b>
3.1	Methode der kleinsten Fehlerquadrate . . . . .	13
3.2	Messmodell für <i>Waferscans</i> . . . . .	16
3.3	Parameterschätzung für $R_0$ . . . . .	17
3.3.1	Kreuzkorrelation . . . . .	18
3.3.2	$R_0$ -Bestimmung mittels <i>CF</i> und <i>FWF</i> . . . . .	20
3.3.3	<i>Cramér-Rao Lower Bound</i> für $\text{var} \{ \hat{R}_0 \}$ . . . . .	23
3.3.4	Konfidenzintervall für $\hat{R}_0$ . . . . .	25
3.4	Parameterschätzung für <i>Waferscans</i> . . . . .	26
3.4.1	Beispiel <i>Flatwafer Fitting (FWF)</i> . . . . .	27
3.4.2	Effizienz von Schätzern . . . . .	29
3.5	Modellvereinfachung . . . . .	30
3.5.1	Reduktion des Modells . . . . .	31
3.5.2	Eigenschaften des reduzierten Modells . . . . .	33

<b>4</b>	<b><i>Sine Wave Fitting</i></b>	<b>34</b>
4.1	Drei Parameter Fitting . . . . .	35
4.1.1	Zwei Parameter Fitting . . . . .	36
4.1.2	Quantisierungsfehler . . . . .	38
4.2	<i>CRLB</i> für $var\{\hat{\delta}\}$ und $var\{\hat{\varphi}_0\}$ . . . . .	40
<b>5</b>	<b><i>Fitting Performance</i></b>	<b>42</b>
5.1	Verzerrte Rauschleistungsschätzung . . . . .	43
5.2	Unverzerrte Rauschleistungsschätzung . . . . .	44
5.3	<i>SWF</i> anhand des Flatwafermodells . . . . .	46
5.4	Vergleich von <i>SWF</i> , <i>CF</i> und <i>FWF</i> . . . . .	49
5.5	Verarbeitung realer Messdaten . . . . .	51
<b>6</b>	<b><i>Flat und Notch Fitting</i></b>	<b>53</b>
6.1	Diskrete Ableitung von <i>Waferscans</i> . . . . .	53
6.2	Parameterschätzung für <i>Flat</i> und <i>Notch</i> . . . . .	55
6.2.1	<i>FF</i> und <i>NF</i> mittels quadratischer Interpolation . . . . .	56
6.2.2	<i>LMS</i> Schätzer für quadratische Flatinterpolation . . . . .	57
6.2.3	<i>CRLB</i> für $var\{\hat{\gamma}\}$ . . . . .	59
<b>7</b>	<b>Dünnpwafer</b>	<b>60</b>
7.1	Korrektur der Waferkrümmung . . . . .	61
7.2	3D-Wafermodelle . . . . .	62
<b>8</b>	<b>Schlussfolgerungen</b>	<b>63</b>
	<b>Literaturverzeichnis</b>	<b>65</b>
	<b>Abbildungsverzeichnis</b>	<b>68</b>
	<b>Tabellenverzeichnis</b>	<b>70</b>

# 1 Einleitung

Im Vorfeld dieser Arbeit wurde ein *2D-Waferscanner* [1] entwickelt, welcher zur optischen Vermessung von Halbleiterwafern dient. Ursprung dieser Entwicklung war die Anforderung, hauchdünne und gekrümmte *Wafer* (Dünnpwafer) in zwei Dimensionen vermessen zu können. Zum Zwecke einer Positionskorrektur (*Prealignment*) von ebenen Standardwafern, erfassen herkömmliche optische Messsysteme den Waferrand nur in einer Dimension. Dabei wird der *Wafer* von einem Rotor (*Chuck*) um  $360^\circ$  gedreht, während das Messgerät den Abstand des Waferrandes zum Rotormittelpunkt misst. Aus der erhaltenen Messkurve kann anschließend die Position des *Wafers* berechnet werden. Im Gegensatz zu Standardwafern haben Dünnpwafer die Eigenschaft, sich durch innere mechanische Spannungen zu verkrümmen. Mit steigender Krümmung verkürzen sich jedoch die vom optischen System abgebildeten Abstände zwischen Waferrand und *Chuck*, was zu verzerrten Messsignalen führt. Mit Hilfe des *2D-Waferscanners* kann nicht nur die Entfernung des Waferrandes zum *Chuck*, sondern auch die Durchbiegung des *Wafers* erfasst werden.

## 1.1 Hintergrund

In einer Vielzahl von Prozessen zur Herstellung integrierter Schaltkreise ist es notwendig, *Wafer* vor ihrer Bearbeitung in Fertigungsanlagen hinreichend genau zu positionieren. Die Anforderungen an die Genauigkeit hängen von den jeweiligen Prozessen ab. Während bei produktionsbegleitenden Messungen elektrischer Parameter von *Wafers* eine Positioniergenauigkeit im Mikrometerbereich ausreicht, so sind für das Sägen der *Wafer*, als einer der letzten Herstellungsprozesse, Genauigkeiten im Nanometerbereich erforderlich [2]. Vor allem aber bei fotolithographischen Prozessen, ist die Positioniergenauigkeit des *Wafers* gegenüber

dem zu übertragenden *Chiplayout* von größter Bedeutung. Das hochgenaue Positionieren von *Wafern* (*Alignment*) wird dabei ausschließlich durch bildgebende Verfahren mit Hilfe von Kameras und Mikroskopen durchgeführt. Bevor ein *Wafer* jedoch dem *Alignment* zugeführt werden kann, wird dieser beim sogenannten *Prealignment* [3][4][5] zum ersten mal messtechnisch erfasst und anschließend für das *Alignment* vorbereitend positioniert. Die einzuhaltenden Toleranzen für die Positionierung liegen durchschnittlich bei  $\pm 10 \mu\text{m}$ . Zusätzlich zur Positionierung des *Wafers* ist auch seine Winkelausrichtung von Bedeutung, für welche üblicherweise Toleranzen von  $\pm 0,1^\circ$  verlangt werden.

## 1.2 Motivation

Das Thema *Waferhandling* ist seit jeher ein heikles, erfordern die empfindlichen *Wafer* doch stets einen behutsamen Transport und dürfen mechanisch keinen großen Belastungen ausgesetzt werden. Um hohe Durchsätze in der Halbleiterfertigung zu erreichen, wird ständig daran gearbeitet, den Grad der Automatisierung zu erhöhen, um letztlich auch den Einfluss menschlicher Fehlerquellen zu reduzieren. Allein die Realität sieht freilich anders aus. Trotz einer Vielzahl von Lösungen für das optische *Alignment* und *Prealignment* von *Wafern*, ist es heutzutage immer noch üblich *Wafer* mit Hilfe mechanischer Vorrichtungen zu positionieren. Selbstverständlich können die *Wafer* dabei beschädigt werden, sodass entweder Ausbrüche am Waferrand entstehen oder sich im schlimmsten Fall Störungen im Kristallgitter des hochreinen Siliziumsubstrates über den ganzen *Wafer* hinweg fortpflanzen und zu vollständigem Bruch führen. In den vergangenen Jahren ist diese Problematik mit den immer dünner werdenden *Wafern* mehr und mehr in den Vordergrund getreten.

Viele Anlagenhersteller berufen sich auf die Standards *SEMI M1-0302* und *SEMI M1.1-M1.15*, welche sowohl mechanische Eigenschaften als auch Materialdicken und Maße von *Wafern* eindeutig spezifizieren. Alle von diesem Standard abweichenden *Wafer* mit veränderten Waferradien, modifizierten Randmarkierungen und vor allem Dünnwafers, müssen von den automatischen Anlagen daher nicht zwingend gehandhabt werden können. Dem Kunden hilft das selbstverständ-

lich wenig, kann er doch technologiebedingt nicht auf das Dünnschleifen seiner *Wafer* verzichten.

So mancher Anlagenhersteller entwickelt zudem seine eigene Lösung für das *Prealignment*, was teilweise zu sehr aufwendigen und komplizierten Realisierungen führt. Obwohl ein erfolgreiches *Prealignment* von standardkonformen *Wafern* in der Praxis schon lange nicht mehr ausreicht, beobachtet man sehr oft, wie sich Anlagenhersteller hauptsächlich auf die Beherrschung ihres Kernprozesses (Dotieren, Messen, Sägen, etc.) konzentrieren und einem stabilen Waferhandling für herkömmliche und dünne *Wafer* weniger Beachtung schenken.

Ziel dieser Arbeit ist einerseits die Feststellung notwendiger Anforderungen für eine effiziente Durchführung des *Waferprealignments* in Hinblick auf die mechanische Umsetzung und andererseits eine effiziente Messdatenverarbeitung, welche sowohl der geforderten Genauigkeit als auch einer praktischen Anwendbarkeit gerecht werden soll. Bei der Entwicklung der verschiedenen Algorithmen zur Messdatenverarbeitung steht neben der Schätzung der Messkurvenparameter vor allem die Qualitätsbewertung der Parameterschätzungen im Vordergrund. Allen Algorithmen wird zur Beurteilung der *Performance* zudem eine theoretisch erreichbare Grenze zum objektiven Vergleich gegenübergestellt. Abschließend wird ein Konzept vorgeschlagen, welches sowohl Standard- als auch Dünnwafers gleichermaßen vermessen und positionieren kann, wobei alle aktuell üblichen Formen und Ausführungen von Wafertypen universell berücksichtigt werden.

## 2 Wafertypen

In der Halbleiterproduktion werden integrierte Schaltkreise wie auch mikroelektromechanische Strukturen (*MEMS*) auf kreisrunden *Wafern* produziert. Als Wafergrundmaterial (Substrat) dienen verschiedene Halbleiter wie Silizium, Germanium, Gallium-Arsenid und viele mehr. Die Herstellung der *Wafer* erfolgt entweder im Tiegelziehverfahren (Czochralski-Verfahren) oder im Zonenschmelzverfahren, mit welchem ein hoch reiner Einkristallstab erzeugt wird. Nach anschließendem Sägen des Einkristallstabes in Scheiben von 500  $\mu\text{m}$  bis 800  $\mu\text{m}$  Dicke werden die so entstandenen *Wafer* geläppt, poliert und hinsichtlich der Kristallorientierung markiert. Diese Markierung wird durch eine seitliche Abflachung (*Flat*) am Waferrand erzeugt. Bei *Wafern* größerer Durchmesser wird die Kristallorientierung durch eine kleine Kerbe (*Notch*) angezeigt, wodurch weniger Verschnitt verursacht wird. Während des gesamten Fertigungsprozesses von integrierten Schaltkreisen dienen *Flat* bzw. *Notch* zur Vorausrichtung (*Prealignment*) der *Wafer* in diversen Fertigungsanlagen. Bei photolithographischen oder messtechnischen Prozessen kommt nach dem *Prealignment* noch eine bildgestützte Feinpositionierung der *Wafer* hinzu. Dabei werden die erzeugten Halbleiterstrukturen auf dem *Wafer* mittels einer Kamera erfasst und durch Bildverarbeitung mit einer Vorgabe verglichen. Diese Positionierung (*Alignment*) ist in der Größenordnung der hergestellten Strukturbreiten durchzuführen und befindet sich im Sub-Mikrometerbereich.

Die Produktion der *Wafer* begann in den 60iger Jahren mit Durchmessern von 1 Zoll bis zu den heute verwendeten 300 mm *Wafern*. Dabei sind kleinere *Wafer* nicht etwa aus den Fertigungsstrassen verschwunden, sondern finden sich vielmehr in Nischenprodukten und bei Spezialsubstraten wieder, welche noch nicht in größeren Durchmessern hergestellt werden. Auf Grund der stetig steigenden Kosteneffizienz geht der Trend weiterhin zu immer größeren Waferdurchmessern.

## 2.1 Waferspezifikationen

Spezifikationen über Form, Durchmesser und Dicke von *Wafers* sind in den Standards *SEMI M1-0302* und *SEMI M1.1-M1.15* beschrieben. *SEMI* ist das Akronym für *Semiconductor Equipment and Materials International* des Fachverbandes der Halbleiterzubehör-, Material- und Flachbildschirmindustrie. In Tabelle 2.1 sind die für das *Prealignment* wichtigsten Kenngrößen aus diesem Standard aufgelistet. Wie aus Tabelle 1 zu erkennen ist, sind diverse Parameter für marktüb-

<b>Durchmesser</b>	100 mm	125 mm	150 mm	200 mm	300 mm
Toleranz	±0,5 mm	±0,5 mm	±0,2 mm	±0,2 mm	±0,2 mm
<b>Dicke</b>	525 µm	625 µm	675 µm	725 µm	775 µm
Toleranz	±20 µm				
<b>Flatlänge I</b>	32,5 mm	42,5 mm	57,5 mm	-	-
Toleranz	±2,5 mm	±2,5 mm	±2,5 mm	-	-
<b>Flatlänge II</b>	18,0 mm	27,5 mm	37,5 mm	-	-
Toleranz	±2,0 mm	±2,5 mm	±2,5 mm	-	-
<b>Notchlänge</b>	-	-	-	4,0 mm	4,0 mm
Toleranz	-	-	-	±0,2 mm	±0,2 mm
<b>Notchtiefe</b>	-	-	-	4,0 mm	4,0 mm
Toleranz	-	-	-	±0,2 mm	±0,2 mm

Tabelle 2.1: Waferspezifikationen

liche *Wafer* standardisiert. Vor allem die angegebenen Toleranzen spielen für *Prealignment* Systeme eine wichtige Rolle, da *Wafer* innerhalb dieser Grenzen richtig erkannt und hinreichend genau positioniert werden müssen. So darf es beispielsweise auf das *Prealignment* eines 150 mm *Wafers* keine Auswirkung haben, ob der *Flat* nun eine Länge von 55 mm oder etwa 60 mm besitzt. Ebenso ist es mit dem Waferdurchmesser. Halbleiterfertigungsanlagen beziehen sich beim *Prealignment* meist auf den Mittelpunkt des *Wafers*, weshalb bei Anlagen, welche *Wafer* verschiedener Größen bearbeiten, Toleranzen im Waferdurchmesser keine Rolle spielen dürfen. Abbildung 2.1 zeigt Wafertypen, wie sie für *Wafer*

mit Durchmessern  $< 125$  mm hergestellt werden. Bei diesen Varianten sind verschiedene Kristallorientierungen mittels primärem und sekundärem *Flat* definiert. Für das *Prealignment* ist stets der primäre *Flat* heranzuziehen. Neben den Standardwafern gibt es jedoch auch solche, welche aus fertigungstechnischen Gründen kleinere Durchmesser als üblich besitzen. Es existieren sogar *Wafer* mit modifizierten Randmarkierungen. Abbildung 2.2 zeigt zwei dieser Spezialformen.

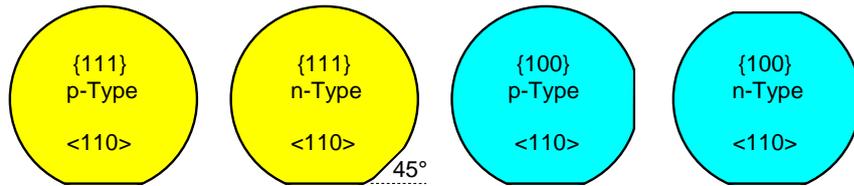


Abbildung 2.1: *Wafer* mit Primär- und Sekundärflat

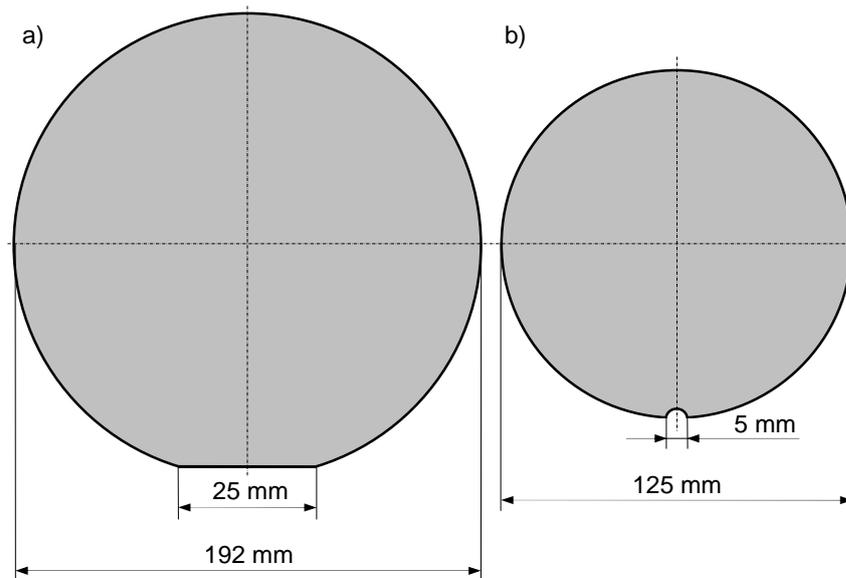


Abbildung 2.2: Nicht standardisierte Spezialwafer

## 2.2 Modellierung von *Waferscans*

Durch Drehung des *Wafers* auf einem *Chuck* und gleichzeitiges optisches Abtasten des Waferrandes entsteht ein *Waferscan*. Üblicherweise wird der *Wafer* während einer vollständigen Rotation von  $360^\circ$  in äquidistanten Winkelschritten abgetastet. Für das Aufstellen eines mathematischen Waferscanmodells ist daher der Einsatz des Polarkoordinatensystems am geeignetsten. Abbildung 2.3 zeigt zunächst einen Kreis als Modell eines fiktiven *Wafers* ohne Randmarkierung. Er wird definiert durch die Länge des Dezentrierungsvektors  $\vec{\delta}$ , den dazugehörigen Nullphasenwinkel  $\varphi_0$  und den Waferradius  $R$ . Die Abtastung des *Wafers* soll entlang der  $0^\circ$ -Geraden des Polarkoordinatensystems durchgeführt werden. Somit entspricht die Radialkoordinate  $r_K(\varphi)$  dem euklidischen Abstand zwischen Koordinatenursprung und Waferrand, welcher von Dezentrierungs- und Radiusvektor aufgespannt wird.

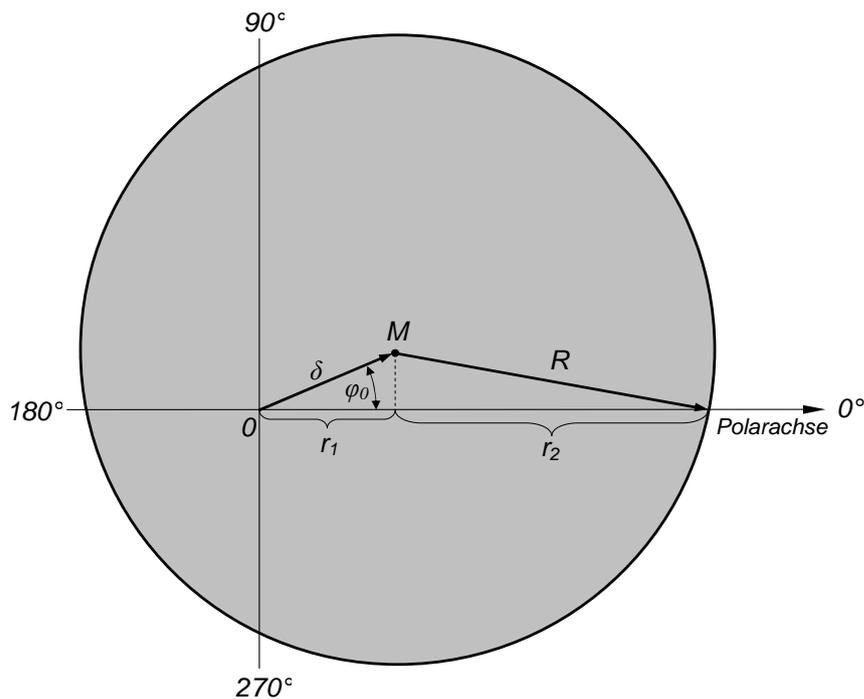


Abbildung 2.3: *Wafer* ohne Randmarkierung

$$r_K(\varphi) = \delta \cos(\varphi + \varphi_0) + r_h(\varphi) \quad (2.1)$$

Dabei wird mit  $r_h(\varphi)$  die Horizontalkomponente des Radiusvektors bezeichnet und mit Hilfe des pythagoräischen Zusammenhanges in (2.2) beschrieben.

$$r_h(\varphi) = \sqrt{R^2 - \delta^2 \sin^2(\varphi + \varphi_0)} \quad (2.2)$$

Wie aus (2.1) hervorgeht, beinhaltet ein *Waferscan* neben der Beschreibung der Horizontalkomponente des Dezentrierungsvektors  $\delta \cos(\varphi + \varphi_0)$  zusätzlich die von  $\varphi$  abhängige Komponente des Radiusvektors. Diese wird besonders durch die von  $\delta$  gewichtete doppeltfrequente  $\sin^2$ -Funktion charakterisiert. Abbildung 2.4 zeigt den Verlauf von  $r_K(\varphi)$  und die Abweichung von einem sinusförmigen Signalverlauf.

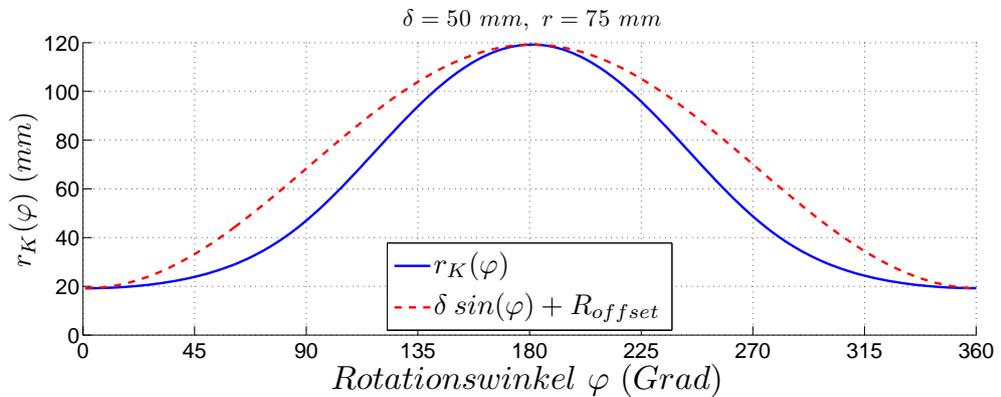


Abbildung 2.4:  $r_K(\varphi)$  verglichen mit sinusförmigem Kurvenverlauf

### 2.2.1 Flatwafer

Zur Vervollständigung des Waferscanmodells soll nun auch die Randmarkierung in Form eines *Flats* betrachtet werden. Abbildung (2.5) zeigt die entsprechende Darstellung mit den notwendigen Parametern. Der *Flat* wird nicht mittels der Sekantenlänge, sondern durch die Flattiefe  $R_D = R - R_F$  beschrieben. Der Flatwinkel  $\gamma$  definiert den Winkel zwischen Dezentrierungsvektor und Flatmitte. Ausgehend von der Geradenvorschrift  $y = kx + d$  wird für die Gerade in Polar-

koordinaten eine Vorschrift für  $r_g(\varphi)$  laut Beziehung (2.3) verwendet.

$$r_g(\varphi) = \frac{d}{\sin(\varphi) - k \cdot \cos(\varphi)} \quad (2.3)$$

Der Ordinatenabschnitt  $d$  setzt sich nach Abbildung (2.5) aus den Teilen I,II und III zusammen. Die Steigung der Flatgerade steht orthogonal auf den Vektor  $\vec{R}_F$ , weshalb für  $k = \tan\left(\frac{\pi}{2} - \phi\right)$  definiert wird. Weiters gilt:  $\phi = \varphi_0 + \gamma$  und  $\vartheta = \varphi + \varphi_0 + \gamma$ .

$$I : \delta \sin(\varphi_0)$$

$$II : R_F \sin(\phi)$$

$$III : \left[ R_F \cos(\vartheta) + \delta \cos(\varphi + \varphi_0) \right] \cdot \tan\left(\frac{\pi}{2} - \phi\right)$$

Die Funktion  $r_{Fa}(\varphi)$  beschreibt nun den allgemeinen Verlauf des *Flats* für  $\phi \neq k\pi$ ,  $k \in \mathbb{Z}$ . Für  $\phi = k\pi$  ist der *Flat* in vertikaler Ausrichtung und Gleichung (2.4) besitzt auf Grund der enthaltenen Tangensfunktionen keine Definitionsmenge. Dieser Spezialfall wird durch die Funktion  $r_{Fv}(\varphi)$  laut Abbildung (2.6) zur Beschreibung vertikaler Flatpositionen berücksichtigt.

$$r_{Fa}(\varphi) = \frac{\delta \sin(\varphi + \varphi_0) + \left[ R_F \cos(\vartheta) + \delta \cos(\varphi + \varphi_0) \right] \cdot \tan\left(\frac{\pi}{2} - \phi\right) + R_F \sin(\vartheta)}{\sin(\varphi) + \tan\left(\frac{\pi}{2} - \phi\right) \cdot \cos(\varphi)} \quad (2.4)$$

$$r_{Fv}(\varphi) = \frac{\sqrt{\left( \delta \cos(\varphi + \varphi_0) + R_F \right)^2 + \left( \delta \sin(\varphi + \varphi_0) \right)^2}}{\cos(\varphi)} \quad (2.5)$$

Das endgültige Wafermodell besteht somit aus der Beschreibung von Kreis (2.1), allgemeinem *Flat* (2.4) und vertikalem *Flat* (2.5). Alle Funktionen besitzen eigene Definitionsbereiche, welche selbst wiederum Funktionen der Parameter  $R$ ,  $\delta$ ,  $\varphi_0$ ,  $R_F$  und  $\gamma$  sind. Durch eine funktionswertabhängige Fallunterscheidung lässt sich mit (2.6) eine einfache Definition für  $r_F(\varphi)$  angeben.

$$r_F(\varphi) := \begin{cases} r_{Fa}(\varphi), & 0 \leq r_{Fa}(\varphi) \leq r_k(\varphi), \phi \neq k\pi, k \in \mathbb{Z} \\ r_{Fv}(\varphi), & 0 \leq r_{Fv}(\varphi) \leq r_k(\varphi), \phi = k\pi \\ r_K(\varphi), & \text{sonst} \end{cases} \quad (2.6)$$

In Abbildung 2.7 ist ein simulierter *Waferscan* auf Basis der Funktion aus (2.6)

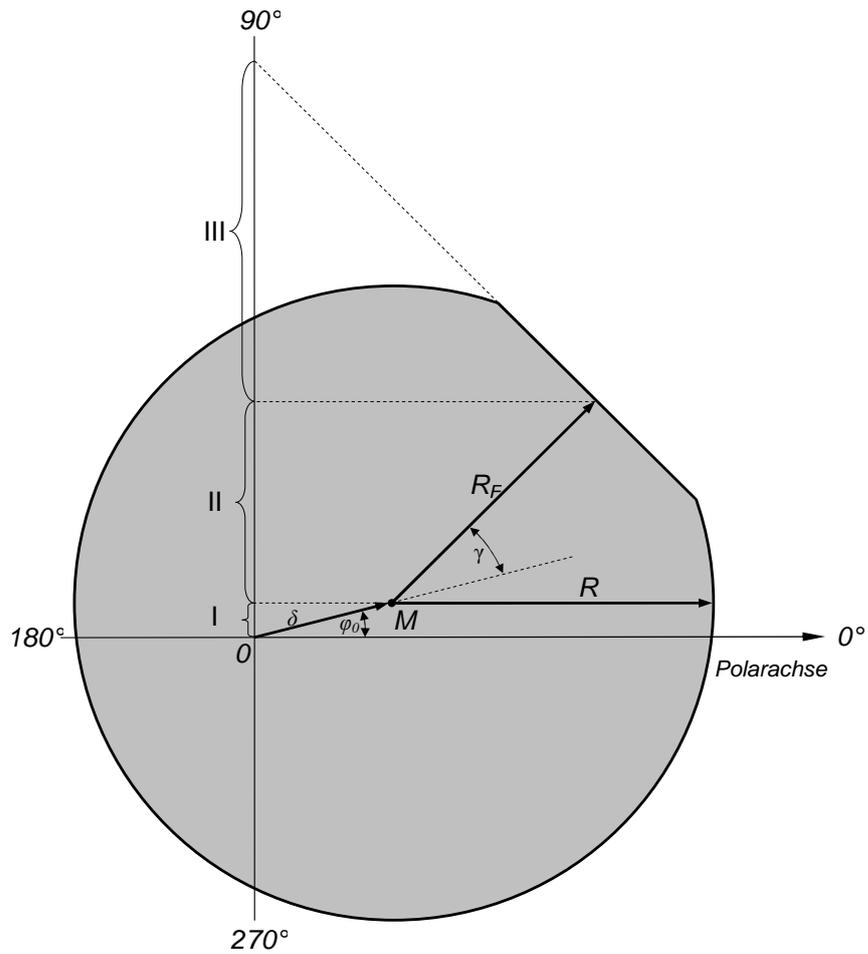


Abbildung 2.5: Wafer mit Flat

gezeigt. Als Ergänzung zum Waferscanmodell mit *Flat* ist in Gleichung (2.7) die Beschreibung für ein Waferscanmodell mit Primär- und Sekundärflat angegeben. Analog zu (2.6) wird für den Sekundärflat  $\bar{\phi} = \varphi_0 + \bar{\gamma}$  sowie  $\bar{\vartheta} = \varphi + \varphi_0 + \bar{\gamma}$  definiert, wobei nun  $\bar{\gamma}$  den Winkel zwischen Dezentrierungsvektor  $\vec{\delta}$  und der Mitte des Sekundärflats beschreibt. Mit  $R_d = R - R_f$  wird die kleinere Tiefe des

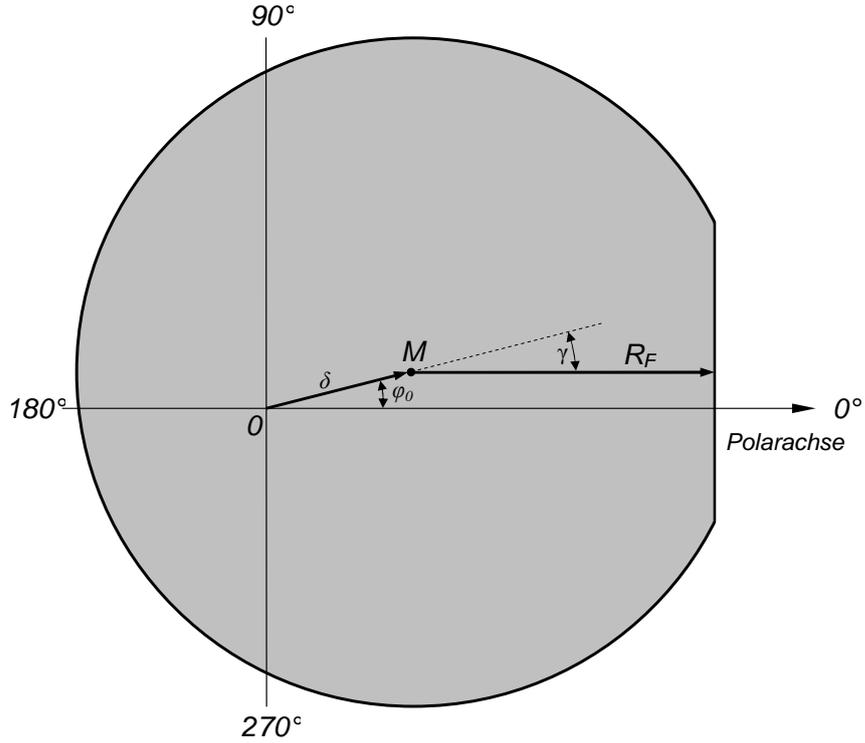


Abbildung 2.6: Wafer mit vertikaler Flatposition

Sekundärflats angeben.

$$r_{2F}(\varphi) := \begin{cases} r_{Fa}(\varphi), & 0 \leq r_{Fa}(\varphi) \leq r_k(\varphi), \phi \neq k\pi, k \in \mathbb{Z} \\ r_{Fv}(\varphi), & 0 \leq r_{Fv}(\varphi) \leq r_k(\varphi), \phi = k\pi \\ r_{fa}(\varphi), & 0 \leq r_{fa}(\varphi) \leq r_k(\varphi), \bar{\phi} \neq k\pi, \\ r_{fv}(\varphi), & 0 \leq r_{fv}(\varphi) \leq r_k(\varphi), \bar{\phi} = k\pi \\ r_K(\varphi), & \text{sonst} \end{cases} \quad (2.7)$$

$$r_{fa}(\varphi) = \frac{\delta \sin(\varphi + \varphi_0) + [R_f \cos(\bar{\vartheta}) + \delta \cos(\varphi + \varphi_0)] \cdot \tan\left(\frac{\pi}{2} - \bar{\phi}\right) + R_f \sin(\bar{\vartheta})}{\sin(\varphi) + \tan\left(\frac{\pi}{2} - \bar{\phi}\right) \cdot \cos(\varphi)} \quad (2.8)$$

$$r_{fv}(\varphi) = \frac{\sqrt{(\delta \cos(\varphi + \varphi_0) + R_f)^2 + (\delta \sin(\varphi + \varphi_0))^2}}{\cos(\varphi)} \quad (2.9)$$

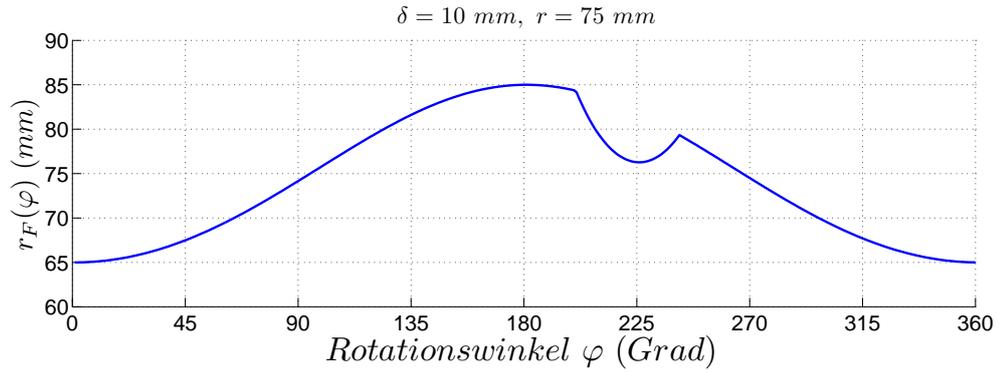


Abbildung 2.7: Darstellung von  $r_F(\varphi)$

### 2.2.2 Wafer mit Notch

Abschließend sind Formbeschreibungen für *Wafer* mit *Notch* (2.10) und halb-kreisförmiger Randmarkierung (2.11) angegeben. Der *Notch* wird beschrieben durch Tiefe  $n_d$  [m], Breite  $n_w$  [m] und  $\gamma$  als Winkel zwischen Dezentrierungsvektor und Notchmitte. Der Waferradius ist wieder mit  $R$  definiert.

$$r_N(\varphi) := \begin{cases} R - \frac{n_d}{2} \left( 1 + \cos\left(\frac{2\pi R}{n_w} \cdot (\varphi - \gamma)\right) \right), & \gamma - \frac{n_w}{2R} \leq \varphi \leq \gamma + \frac{n_w}{2R} \\ R, & \text{sonst} \end{cases} \quad (2.10)$$

### 2.2.3 Wafer mit rundem Notch

Für *Wafer* mit halbkreisförmiger Randmarkierung gilt Definition (2.11), wobei  $r$  den Radius der Randmarkierung beschreibt.

$$r_C(\varphi) := \begin{cases} R - \sqrt{r^2 - R^2 \cdot \sin^2(\varphi - \gamma)}, & \gamma - \frac{r}{R} \leq \varphi \leq \gamma + \frac{r}{R} \\ R, & \text{sonst} \end{cases} \quad (2.11)$$

## 3 Circle Fitting

Eine Vielzahl technischer Anwendungen bedient sich des Einpassens von Kreisen oder Kreisbögen in Messdaten, um die Parameter Radius und Mittelpunkt von Kreisen aus Bildern zu extrahieren [6][7][8]. Bereiche wie Computergrafik, Maschinelles Sehen [9], optische Messtechnik bis hin zur Astronomie, um nur einige wenige zu nennen, wenden Algorithmen des *Circle Fittings* an. In der Bildverarbeitung versucht man möglichst viele Bildpunkte zu identifizieren in welche Kreise mit den Parametern Radius ( $R_n$ ) und Mittelpunkt ( $a_n, b_n$ ) so eingepasst werden, dass die jeweiligen Summen der Fehlerquadrate zwischen ausgewählten Bildpunkten und geschätzten Kreisen minimal werden.

### 3.1 Methode der kleinsten Fehlerquadrate

Abbildung 3.1 zeigt die Bestimmung von Kratern in einer Aufnahme der Mond-Tiefebene *Mare Orientale*. Basis der Minimierungsaufgabe sind  $N$  Datenpunkte  $(x_i, y_i)$  mit  $0 \leq i \leq N - 1$ , welche die Beschreibung der Kreisvorschrift laut Gleichung (3.1) nach Abbildung 3.2 in kartesischen Koordinaten unterstützen sollen.

$$(x_i - a)^2 + (y_i - b)^2 = R^2 \quad (3.1)$$

Mit  $F_c$  nach (3.2) wird jene Funktion definiert, welche die quadrierten geometrischen Distanzen zwischen Bild- und korrespondierenden Kreispunkten summiert.

$$F_c = \sum_{i=0}^{N-1} \left( \sqrt{(x_i - a)^2 + (y_i - b)^2} - R \right)^2 \quad (3.2)$$

Die Minimierung der Funktion  $F_c$  entspricht dem Schätzen der optimalen Kreisparameter  $a, b$  und  $R$ , welche die beste Übereinstimmung zwischen den Datenpunkten und dem eingepassten Kreis erzeugen [6]. Wie aus der Definition der Funktion

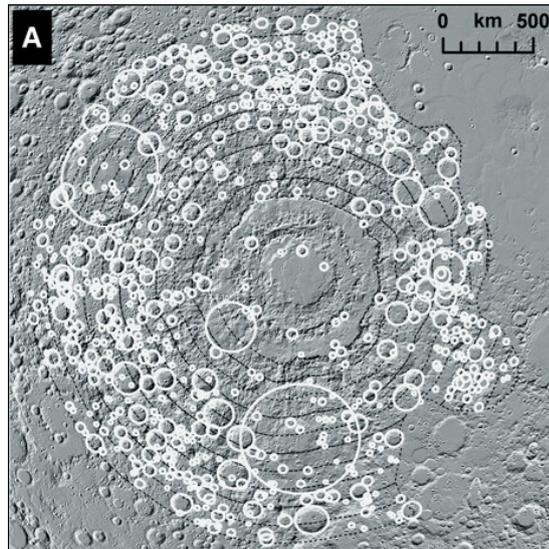
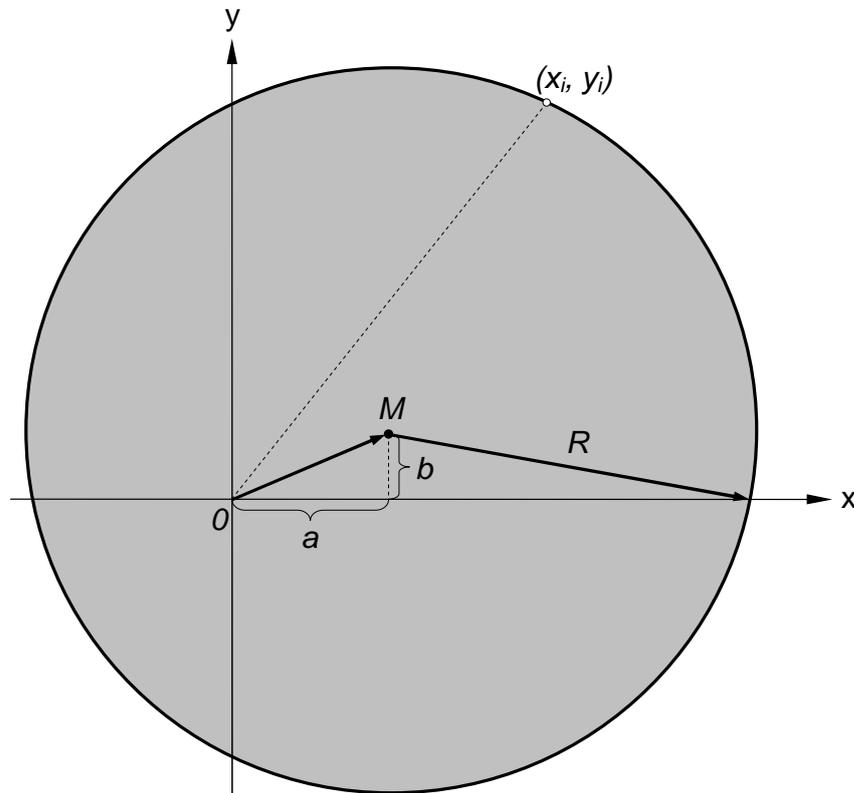


Abbildung 3.1: Krater der Mondtiefebene Orientale aus [10]

hervorgeht, stehen die Parameter jedoch in nichtlinearem Zusammenhang, weshalb keine analytische Lösung dieser Minimierungsaufgabe existiert. Üblicherweise werden iterative Verfahren in Form von numerischen *Least Mean Square* Algorithmen (*LMS*) für nichtlineares *Curve Fitting* zur Lösung herangezogen. Diese stehen in verschiedensten Implementationen wie der *Levenberg-Marquardt*, *Gauß-Newton*, *Line-Search* oder der *Trust-Region* Methode für nichtlineare Minimierung zur Verfügung. Letztere zwei lassen auch die Definition von unteren und oberen Grenzen der einzelnen zu bestimmenden Parameter zu, um Bereiche möglicher Lösungen einschränken zu können. Beinhalten die Messdaten jedoch Ausreißer und grobe Fehler, so gehen diese ungehindert in die Auswertung des vorgegebenen Funktionals ein und erzeugen nicht erwartungstreue Schätzwerte. Es ist daher notwendig, vor der Anwendung von *LMS* Verfahren Messreihen von Ausreißern und groben Fehlern zu bereinigen, wie es beispielsweise der *RANSAC* Algorithmus [11] beschreibt. *RANSAC* steht für *Random Sample Consensus* und wählt aus einem Set von Messwerten durch zufällige Stichproben gerade so viele aus, wie für die Bestimmung des Modells notwendig sind. Die ausgewählten Messwerte werden als ausreißerfrei betrachtet und dienen einer ersten Schätzung der Modellparameter. Im nächsten Schritt wird das Modell anhand der geschätzten Parameter mit allen Messwerten verglichen. Ist die Abweichung zwischen

Abbildung 3.2: Kreisparameter: Radius  $R$  und Mittelpunkt  $M(a,b)$ 

Messwerten und geschätztem Modell kleiner als eine zuvor definierte Schranke, unterstützen die zufällig ausgewählten Messwerte das Modell, andernfalls werden sie zu Ausreißern erklärt. Das Verfahren ist insbesondere bei großen Datenmengen sehr zeitaufwändig, kommt jedoch auf Grund seiner Robustheit gerne zur Anwendung. Aufbauend auf dem *Circle Fitting* Ansatz werden nun notwendige Schritte erarbeitet, um die im Kapitel *Wafer* definierten Modellparameter erfolgreich schätzen zu können. Ziel ist ein robustes und universell einsetzbares System zur Bestimmung von Waferdezentrierung und Orientierung der Randmarkierung von *Wafers* beliebiger Größe und Form (*Flat* oder *Notch*). Stellvertretend für die im vorhergehenden Kapitel beschriebenen Waferformen werden nun anhand des Waferscanmodells  $r_F(\varphi)$  die Eigenschaften der Parameterschätzung untersucht. Wie die gewonnenen Erkenntnisse auf die übrigen Waferformen übertragen werden können, ist Gegenstand späterer Betrachtungen.

## 3.2 Messmodell für *Waferscans*

Vorerst bedarf das zugrunde liegende Waferscanmodell einer Anpassung an den realen Messprozess. Um die Praxis bei der Messdatengewinnung im Modell abzubilden, ersetzt die diskrete Abtastung mit  $N$  Werten pro vollständiger Umdrehung den kontinuierlichen Winkel  $\varphi$  aus Gleichung (6). Die Entfernung des Messgeräts (*Scanner*) vom Chuckmittelpunkt wird nach Abbildung (3.3) mit  $R_0$  festgelegt und beschreibt den Offset zwischen Scanner- und Weltkoordinatensystem in Polarachsenrichtung. Die Abtastung des Waferrandes durch den Scanner geschieht also nur in einem beschränkten Bereich, welcher durch die aktive Sensorlänge vorgegeben ist und lediglich die maximale zu erwartende Dezentrierung des *Wafers* abdeckt. Der zu schätzende Parametervektor lautet daher  $\underline{\theta}_F = [\delta, \varphi_0, R, R_F, \gamma, R_0]^T$ . Zudem wird der gesamte Messprozess mit additivem weißen Gauß'schen Rauschen  $w[n]$  der Verteilung  $\mathcal{N}(\mu, \sigma^2)$  modelliert. Die vom Messgerät durchgeführte Diskretisierung mit der Auflösung  $\Delta$  erweitert das Modell um das Quantisierungsrauschen  $q[n]$ , wodurch das diskrete *Sample*  $s_F[n; \underline{\theta}_F]$  durch Gleichung (3.3) beschrieben wird.

$$s_F[n; \underline{\theta}_F] = r_F[n] - R_0 + w[n] + q[n] \quad (3.3)$$

Entgegen der Kreisgleichung (3.1) aus dem *Circle Fitting* Ansatz ist Gleichung (3.3) in Polarkoordinaten definiert und besitzt einen umfangreicheren Parametervektor  $\underline{\theta}_F$ . Die vom *Scanner* erzeugten Messdaten besitzen entsprechend Abbildung (3.3) keinen Bezug zum Koordinatenursprung. Solange  $R_0$  unbekannt ist, kann der Waferradius  $R$  nicht eindeutig bestimmt werden. Daraus folgt, dass auch die restlichen Parameter fehlerhaft geschätzt werden. Es darf jedoch davon ausgegangen werden, dass das Messgerät in Bezug auf das Rotationszentrum (Koordinatenursprung, *Chuckmitte*) eine definierte und fixe Position einhält. Der Parameter  $R_0$  kann daher messtechnisch ermittelt und weiterführend als Konstante des Systems angesehen werden.

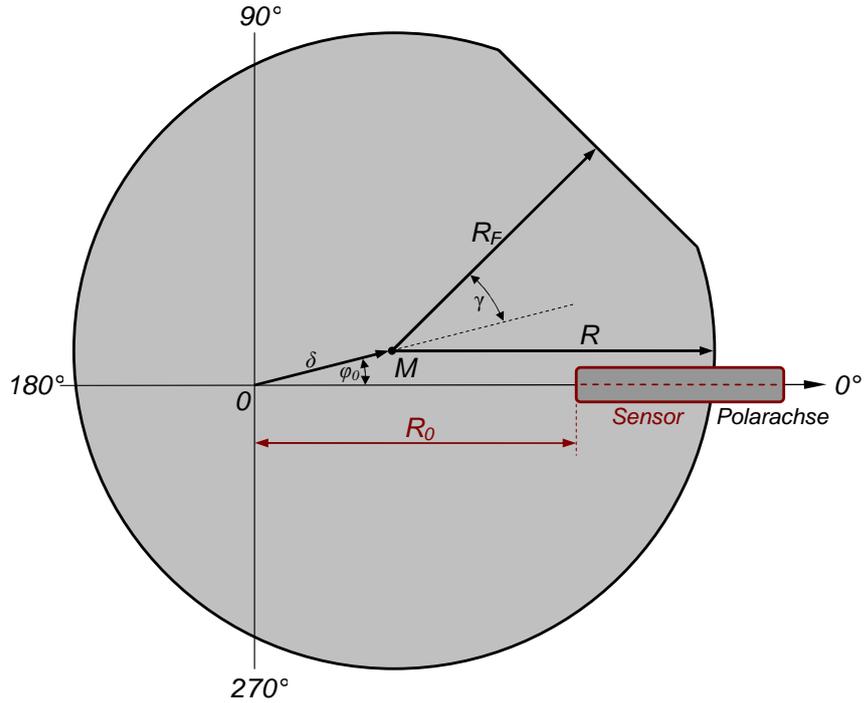


Abbildung 3.3: Position des Messgerätes

### 3.3 Parameterschätzung für $R_0$

Eine Schätzung des *Offsets*  $R_0$  zwischen *Scanner*- und Weltkoordinatensystem wird unter der Verwendung eines Referenzwafers mit *Flat* realisiert, von welchem sowohl der Waferradius  $R$  als auch die Flattiefe  $R_D$  bekannt sind. Der zu schätzende Parametervektor zur Kalibrierung des Systems lautet  $\underline{\theta}_{cal} = [\delta, \varphi_0, \gamma, R_{0,cal}]^T$ . Alle Parameter sind voneinander unabhängig, weshalb eine eindeutige Lösung für  $R_{0,cal}$  berechnet werden kann. Als Startwerte für einen nichtlinearen *Curve Fitting* Algorithmus werden  $\delta$ ,  $\varphi_0$  und  $R_{0,cal}$  unbekannterweise mit Null vorgegeben. Für den Flatwinkel  $\gamma$  muss jedoch ein Startwert gewählt werden, welcher dem tatsächlichen Flatwinkel nahe ist. Hintergrund ist die Funktionsweise des *Curve Fittings*. Durch Variation der Parameter wird die numerische Auswirkung auf die Funktion  $F_c$  aus Gleichung (3.2) berechnet. Wird der Startwert des Flatwinkels in einen Bereich gelegt, welcher sich nicht zumindest teilweise mit dem *Flat* aus den Messdaten überlappt, steht eine Änderung des Flatwinkels in keinem

Zusammenhang mit der Auswertung der Funktion  $F_c$ . Der Gradient bezüglich des Flatwinkels  $\gamma$  ergibt Null, weshalb der *Flat* nicht erfolgreich eingepasst werden kann. Durch mehrfaches Durchführen des *Curve Fittings* mit wanderndem Startwert für den Flatwinkel und anschließender Wahl des Parametervektors mit dem kleinsten Fehlerfunktional kann dieser Umstand umgangen werden. Der bei dieser Variante erhöhte Rechenaufwand kann vermieden werden, wenn alternativ ein Parametervektor mit bereits bekannten Startwerten definiert werden kann. Für welchen Weg auch immer man sich entscheidet, universelle Lösungen zum Auffinden des Flatwinkels  $\gamma$  bzw. der Randmarkierung des *Wafers* können nur durch Auswertung der Messdaten erreicht werden. Ist die Position der Flatmitte in den Messdaten einmal bekannt, stehen die beiden Vorgehensweisen *Circle Fitting* (*CF*) und *Flatwafer Fitting* (*FWF*) zur Verfügung. Beim *Circle Fitting* wird nur das mathematische Modell  $r_K [n; [\delta, \varphi_0, R_{0_{cal}}]^T]$  zur fehlerquadratminimierenden Schätzung der Kreisparameter verwendet, wohingegen beim *Flatwafer Fitting* anhand des mathematischen Modells  $r_F [n; \vartheta_{cal}]$  ein nicht lineares *Fitting* im Sinne einer Fehlerquadratminimierung durchgeführt wird. Die Position der Flatmitte dient beim *CF* zum Ausschluss der Messwerte, welche im Zuge der Parameterschätzung ausgeblendet werden müssen und beim *FWF* als Startwert für  $\gamma$ . Im Folgenden wird das Auffinden des Flatbereiches für das *CF* und *FWF* zur Schätzung von  $R_0$  beschrieben.

### 3.3.1 Kreuzkorrelation

Durch Lokalisierung der Flatmitte in den Messdaten wird ein Startwert für den Flatwinkel  $\gamma$  als Basis für das *FWF* ermittelt. Der Anspruch liegt dabei auf einer Datenverarbeitung, welche den *Flat* besonders unter Berücksichtigung von Rauschsignalen und Messwertausreißern erkennt. Die Berechnung einer Kreuzkorrelation zwischen *Waferscan* und Flatbeschreibung  $r_{F_a} [n]$  erweist sich dabei als einfache Methode, den *Flat* im Messsignal zu lokalisieren. Abbildung 3.4 zeigt einen simulierten *Waferscan*, in welchem sowohl Quantisierungs- als auch Weißes Rauschen berücksichtigt sind. In Abbildung 3.5 ist das Ergebnis einer normierten Kreuzkorrelation aus *Waferscan* und Flatbeschreibung gezeigt. Laut Definition

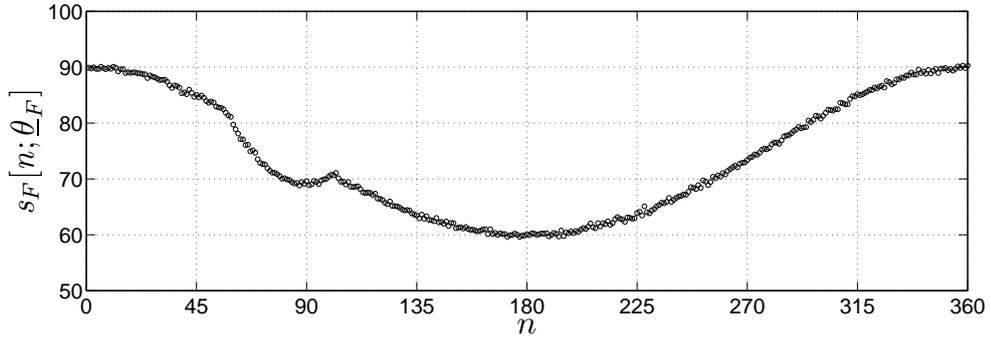


Abbildung 3.4: *Waferscan* mit *Flat*

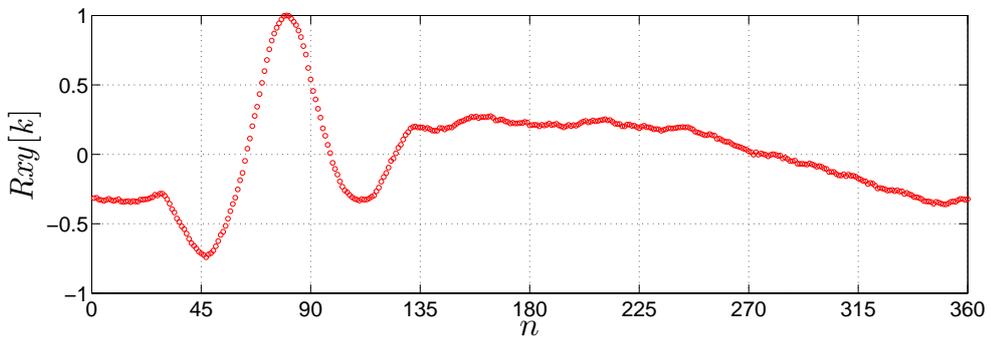


Abbildung 3.5: Kreuzkorrelation aus *Waferscan* und Flatbeschreibung  $r_{F_a}[n]$

des Flatwinkels (siehe Abbildung 3.3) setzt sich der durch Kreuzkorrelation gefundene Flatmittelpunkt jedoch aus  $\varphi_0$  und  $\gamma$  zusammen und kann nicht direkt als Startwert für  $\gamma$  herangezogen werden. Anwendbar ist das Aufsuchen des *Flats* in der Form also nur, wenn für  $\varphi_0 \approx 0$  gilt oder für  $\varphi_0$  auch ein Startwert bekannt ist, welcher ebenfalls mittels einer Kreuzkorrelation gefunden werden kann. Bezüglich Abbildung 3.4 lauten die Startwerte für den Parametervektor des *FWF*  $\underline{\theta}_{cal} = [\delta, \varphi_0, \gamma, R_{0_{cal}}]^T = [0, 0, \pi/2, 0]^T$ . Für das *CF* werden lediglich die Messdaten um den durch Kreuzkorrelation ermittelten Flatbereich  $W_F$  nach Abbildung 3.7 verworfen. Die Startwerte für das *CF* werden mit  $[\delta, \varphi_0, R_{0_{cal}}]^T = [0, 0, 0]^T$  vorgegeben.

### 3.3.2 $R_0$ -Bestimmung mittels $CF$ und $FWF$

Mit Hilfe des nichtlinearen *Curve Fittings* *lsqcurvefit* (MATLAB<sup>®</sup>[12]) wird der gesuchte Parameter  $R_0$  anhand einer Reihe von  $M = 1000$  simulierten *Waferscans* mit additivem Rauschen der Verteilung  $\mathcal{N}(\theta, \sigma_0^2)$  geschätzt. In Abbildung 3.6 wird stellvertretend für die  $M$  Durchläufe das Ergebnis eines  $FWF$  mit zugehörigem *Waferscan* dargestellt, Abbildung 3.7 zeigt das Ergebnis eines  $CF$ , wobei die Messwerte im Flatbereich  $W_F$  nicht berücksichtigt wurden. Abbildung 3.8

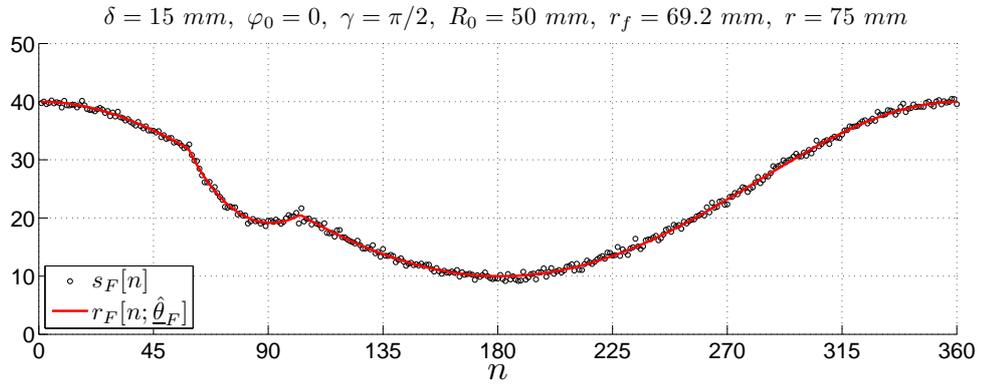


Abbildung 3.6: Flatwafer Fitting ( $FWF$ ) mit  $\sigma_0^2 = 0.05 \text{ mm}^2$

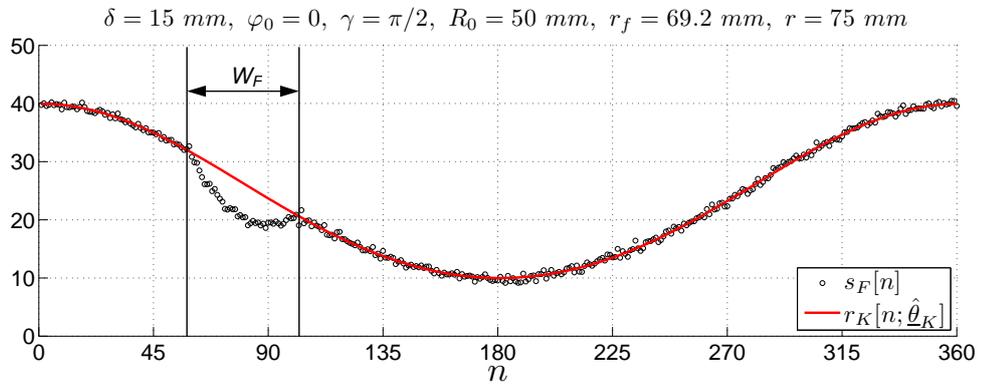


Abbildung 3.7: Circle Fitting ( $CF$ ) mit  $\sigma_0^2 = 0.05 \text{ mm}^2$

zeigt das Histogramm der  $M$  geschätzten Parameter  $\hat{R}_0$  für  $CF$  und  $FWF$ . Der Mittelwert  $\bar{R}_0$  der jeweiligen Verteilungen wird nach (3.4) als arithmetisches Mittel aus den Schätzungen für  $\hat{R}_0$  ermittelt. Sowohl das Histogramm als auch die

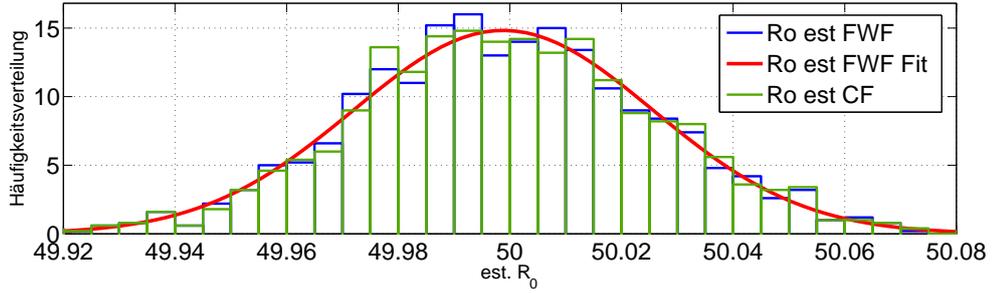


Abbildung 3.8: Histogramm von  $M = 1000$  Parameterschätzungen für  $R_0$

Berechnung zeigt, dass für  $\bar{R}_0 \approx R_0 = 50$  mm gilt.

$$\bar{R}_0 = \frac{1}{M} \sum_{i=0}^{M-1} \hat{R}_{0i} = 49,9988 \text{ mm} \quad (3.4)$$

Auf Basis von  $\bar{R}_0$  wird nach (3.5) die Varianz der Verteilung von  $M$  Parameterschätzungen  $\hat{R}_{0i}$  (Abbildung 3.8) bestimmt [13].

$$\bar{\sigma}^2 = \frac{1}{M-1} \sum_{i=0}^{M-1} (\hat{R}_{0i} - \bar{R}_0)^2 \quad (3.5)$$

Anhand der Simulation zeigt sich der in (3.6) beschriebene Zusammenhang zwischen geschätzter Varianz  $\bar{\sigma}^2$  und der additiven Rauschleistung  $\sigma_0^2$  der simulierten *Waferscans*, welche jeweils  $N$  Messwerte umfassen.

$$\bar{\sigma}^2 \approx \frac{\sigma_0^2}{N} \quad (3.6)$$

Der Standardfehler des arithmetischen Mittels  $\bar{R}_0$  wird nach (3.7) als Quotient aus  $\bar{\sigma}^2$  und der Anzahl der  $M$  Parameterschätzungen gebildet.

$$\text{var} \{ \bar{R}_0 \} = \frac{\bar{\sigma}^2}{M} \quad (3.7)$$

Da in der Praxis nicht tausende Messungen durchgeführt werden wollen, soll der Parameter  $R_0$  anhand einer einzelnen Messkurve geschätzt werden. Das Ergebnis ist ein einziger Schätzwert und es stellt sich die Frage nach der Größe seiner Varianz. Aus Beziehung (3.6) geht hervor, dass die Streuung der geschätzten Parameter  $\hat{R}_{0i}$  einen Zusammenhang mit der Streuung des additiven Rauschens der

*Waferscans* aufweist. Durch Berechnung des quadratischen Fehlers zwischen *Waferscan* und ermitteltem Modell kann auch ein Schätzwert für das im Messsignal enthaltene Rauschen bestimmt werden, wie im Falle des *FWF* in (3.8) angegeben ist.

$$\hat{\sigma}_{FWF}^2 = \frac{1}{N-1} \sum_{n=0}^{N-1} (s_F[n] - r_F[n; \hat{\theta}_{FWF}])^2 \quad (3.8)$$

$$\hat{\sigma}_{CF}^2 = \frac{1}{N - W_F - 1} \sum_{n=0}^{N - W_F - 1} (s_F[n] - r_K[n; \hat{\theta}_{CF}])^2 \quad (3.9)$$

Für die Varianz des geschätzten Parameters  $\hat{R}_0$  wird entsprechend (3.7) der Quotient aus  $\hat{\sigma}^2$  und der Anzahl der Messwerte  $N$  bzw.  $N - W_F$  nach (3.10) und (3.11) gebildet.

$$\text{var}_{FWF} \{ \hat{R}_0 \} \approx \frac{\hat{\sigma}_{FWF}^2}{N} \quad (3.10)$$

$$\text{var}_{CF} \{ \hat{R}_0 \} \approx \frac{\hat{\sigma}_{CF}^2}{N - W_F} \quad (3.11)$$

Die Schätzungen von  $\hat{\sigma}_{FWF}^2$  und  $\hat{\sigma}_{CF}^2$  für verschiedene Messwerte  $N$  werden in Abbildung (3.9) dargestellt. Mit großem Messwertumfang nähern sich die Schätzungen für  $N > 250$  rasch der Rauschleistung  $\sigma_0^2 = 0.05 \text{ mm}^2$ , welche Basis der simulierten *Waferscans* war. Es wird festgehalten, dass durch das nichtlineare

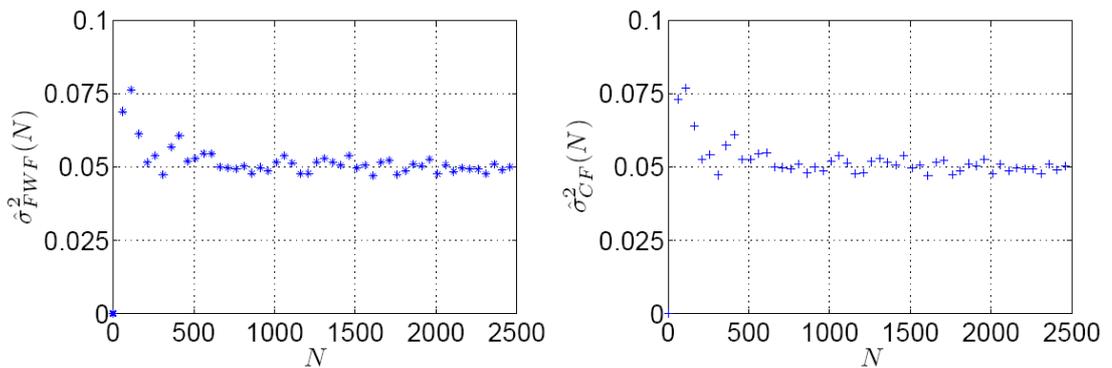


Abbildung 3.9: Varianzschätzung bei *CF* und *FWF*

*Curve Fitting* der Parameter  $R_0$  geschätzt und durch Berechnung des quadratischen Fehlers zwischen Modell und Messkurve auch ein Schätzwert für die Varianz von  $\hat{R}_0$  bestimmt werden kann.

### 3.3.3 Cramér-Rao Lower Bound für $\text{var} \{ \hat{R}_0 \}$

Anhand der vorangegangenen Simulation wurde näherungsweise ein Bezug zwischen dem geschätzten Parameter  $\hat{R}_0$  und seiner Varianz hergestellt. Interessant ist nun die Frage nach einer Grenze dieser Varianz. Wie genau kann dieser Parameter bestenfalls geschätzt werden? Eine analytische Methode zur Lösung dieser Aufgabe erhält man durch Aufstellen der *Cramér-Rao* Ungleichung. Im englischen Sprachraum als *Cramér-Rao Lower Bound* (*CRLB*) bezeichnet, beschreibt sie eine untere Schranke für die Varianz eines geschätzten Parameters [14]. Basierend auf der *Fisher Information*, welche Aussagen über die bestmögliche Qualität von Parameterschätzungen macht, ist in (3.12) die *Cramér-Rao Lower Bound* definiert als die Inverse der *Fisher Information*.

$$\text{CRLB: } \text{var} \{ \hat{\theta} \} \geq \frac{1}{I(\theta)} \quad (3.12)$$

Ausgehend vom *Circle Fitting* Ansatz wird nun die untere Schranke der Varianz für  $\hat{R}_0$  bestimmt. Für einen *Waferscan* auf Basis eines Kreises dient für den diskreten Fall Gleichung (2.1) als Vorgabe. Der *Flat* wird nicht mit berücksichtigt. Die diskrete Beschreibung des *Waferscans* ohne *Flat* nach (3.13) beinhaltet die gesuchten Parameter  $\delta$ ,  $\varphi_0$  und  $R_0$ . Der Waferradius  $R$  gilt nach wie vor als bekannt. Für die Standardabweichung von  $w[n]$  gilt  $\sigma_0 > \Delta$ , weshalb das Quantisierungsrauschen vernachlässigt werden darf.

$$s[n; \underline{\theta}] = \delta \cos\left(\frac{2\pi}{N}n + \varphi_0\right) + \sqrt{R^2 - \delta^2 \sin^2\left(\frac{2\pi}{N}n + \varphi_0\right)} - R_0 + w[n] \quad (3.13)$$

Im Falle des gesuchten Parametervektors  $\theta = [\delta, \varphi_0, R_0]$  wird die *Fisher Information* nach (3.14) in Matrixform ausgewertet. Die im Signal enthaltene Rauschleistung wird mit  $\sigma^2$  berücksichtigt.

$$I(\theta)_{ij} = \frac{1}{\sigma_0^2} \left( \frac{\partial s_{\underline{\theta}}}{\partial \theta_i} \right) \left( \frac{\partial s_{\underline{\theta}}}{\partial \theta_j} \right)^T = \frac{1}{\sigma_0^2} \sum_{n=0}^{N-1} \frac{\partial s[n; \underline{\theta}]}{\partial \theta_i} \frac{\partial s[n; \underline{\theta}]}{\partial \theta_j} \quad (3.14)$$

Um die partiellen Ableitungen nicht unnötig kompliziert werden zu lassen, wird der Parameter der Dezentrierung mit  $0 < \delta < 1$  mm eingeschränkt. Für den Radius gilt  $R \gg 1$  mm, sodass sich ausgehend von Gleichung (3.13) der Wurzelterm

in (3.15) zu  $R$  vereinfachen lässt.

$$s[n; \underline{\theta}] = \delta \cos\left(\frac{2\pi}{N}n + \varphi_0\right) + R - R_0 + w[n] \quad (3.15)$$

Die jeweiligen Ableitungen lauten:

$$\begin{aligned} I(\underline{\theta})_{11} &= \frac{1}{\sigma_0^2} \sum_{n=0}^{N-1} \cos^2\left(\frac{2\pi}{N}n + \varphi_0\right) \approx \frac{N}{2\sigma_0^2} \\ I(\underline{\theta})_{12} = I(\underline{\theta})_{21} &= \frac{-1}{\sigma_0^2} \sum_{n=0}^{N-1} \delta \sin\left(\frac{2\pi}{N}n + \varphi_0\right) \cos\left(\frac{2\pi}{N}n + \varphi_0\right) \approx 0 \\ I(\underline{\theta})_{13} = I(\underline{\theta})_{31} &= \frac{1}{\sigma_0^2} \sum_{n=0}^{N-1} 1 \cdot \cos\left(\frac{2\pi}{N}n + \varphi_0\right) \approx 0 \\ I(\underline{\theta})_{22} &= \frac{1}{\sigma_0^2} \sum_{n=0}^{N-1} \delta^2 \sin^2\left(\frac{2\pi}{N}n + \varphi_0\right) \approx \frac{N\delta^2}{2\sigma_0^2} \\ I(\underline{\theta})_{23} = I(\underline{\theta})_{32} &= \frac{-1}{\sigma_0^2} \sum_{n=0}^{N-1} 1 \cdot \delta \sin\left(\frac{2\pi}{N}n + \varphi_0\right) \approx 0 \\ I(\underline{\theta})_{33} &= \frac{1}{\sigma_0^2} \sum_{n=0}^{N-1} 1 \approx \frac{N}{\sigma_0^2} \end{aligned}$$

Die Matrix der Fisher Information ergibt sich somit zu:

$$I(\underline{\theta}) = \frac{1}{\sigma_0^2} \begin{pmatrix} \frac{N}{2} & 0 & 0 \\ 0 & \frac{N\delta^2}{2} & 0 \\ 0 & 0 & N \end{pmatrix} \quad (3.16)$$

Durch Invertieren der *Fisher Matrix* (3.16) gewinnt man mit den Einträgen aus der Hauptdiagonale die unteren Schranken für die Varianz der geschätzten Parameter von  $\theta$  laut (3.17).

$$CRLB: \frac{1}{I(\underline{\theta})} \geq \sigma_0^2 \begin{pmatrix} \frac{2}{N} & 0 & 0 \\ 0 & \frac{2}{N\delta^2} & 0 \\ 0 & 0 & \frac{1}{N} \end{pmatrix} \quad (3.17)$$

Die kleinste erreichbare Varianz für den gesuchten Parameter  $\hat{R}_0$  ist mit (3.18) eine Bestätigung für die vorhergehenden Simulationen und deckt sich mit der Abschätzung der Varianz nach (3.10).

$$\text{var} \{ \hat{R}_0 \} \geq \frac{\sigma_0^2}{N} \quad (3.18)$$

Bei genauer Betrachtung der *CRLB* Matrix in (3.17) lassen sich die Zusammenhänge zwischen den einzelnen Parametern interpretieren. So ist beispielsweise die Varianz von  $\hat{\delta}$  doppelt so groß gegenüber der von  $\hat{R}_0$ , wenngleich beide voneinander unabhängig sind. Für die Bestimmung der Varianz von  $\hat{R}_0$  könnte man daher auch  $\delta = 0$  setzen. In diesem Fall beinhaltet das Modell nur mehr die Konstanten  $R$  und  $R_0$ . Spätestens jetzt wird deutlich, dass die Schätzung von  $R_0$  dem Berechnen eines arithmetischen Mittels gleicht und die Varianz von  $\hat{R}_0$  der dazugehörige Standardfehler ist. Allen Parametern gemein ist selbstverständlich eine sinkende Varianz mit steigender Anzahl der Messwerte  $N$ . Die Varianz für den Nullphasenwinkel  $\hat{\varphi}_0$  ist zusätzlich abhängig von der Dezentrierung und verringert sich mit größer werdendem  $\delta$ . Durch die Herleitung der *Cramér-Rao Lower Bound* kann nun eine Aussage getroffen werden, wie genau unser gesuchter *Offset*  $R_0$  bestmöglich geschätzt werden kann.

### 3.3.4 Konfidenzintervall für $\hat{R}_0$

Kehren wir gedanklich zur Simulation der  $M = 1000$  Parameterschätzungen für  $R_0$  zurück. Die geschätzten Parameter bilden das in Abbildung 3.8 dargestellte Histogramm. Obwohl der Mittelwert  $\bar{R}_0$  mit guter Näherung bestimmt wurde, gibt es Schätzwerte für  $R_0$  welche außerhalb von beispielsweise  $\bar{R}_0 \pm \sigma$  liegen. Schätzwerte, welche dem tatsächlichen Parameter nahe sind, treten mit hoher Wahrscheinlichkeit auf, jene die dem tatsächlichen Wert fern sind, treten mit geringerer Wahrscheinlichkeit auf. Mit Hilfe des Konfidenzintervalls lässt sich nun eine Aussage über die Präzision der Lageschätzung eines Parameters treffen, indem es einen Bereich um den geschätzten Parameterwert angibt, welcher mit einer zuvor festgelegten Wahrscheinlichkeit die wahre Lage des Parameters trifft. Diese Wahrscheinlichkeit wird als Konfidenzniveau bezeichnet.

$$\left[ \hat{R}_0 - t_{(1-\frac{\alpha}{2}, N-1)} \frac{\sigma_0}{\sqrt{N}}; \hat{R}_0 + t_{(1-\frac{\alpha}{2}, N-1)} \frac{\sigma_0}{\sqrt{N}} \right] \quad (3.19)$$

Für ein Konfidenzniveau von 95%, was einer Signifikanz von 5% entspricht, und einem Stichprobenumfang von  $N=100$ , erhält man aus der *Student t* Verteilung

aus Tabelle 3.1 einen Wert von 1.984.

$$\left[ \hat{R}_0 - 1,984 \cdot \frac{\sigma_0}{\sqrt{N}}; \hat{R}_0 + 1,984 \cdot \frac{\sigma_0}{\sqrt{N}} \right] \quad (3.20)$$

Das in (3.20) angegebene Konfidenzintervall bedeutet, dass das berechnete In-

n	$\alpha = 0,1$	$\alpha = 0,05$	$\alpha = 0,025$	$\alpha = 0,01$	$\alpha = 0,005$
30	1,310	1,697	2,042	2,457	2,750
40	1,303	1,684	2,021	2,423	2,704
60	1,296	1,671	2,000	2,390	2,660
100	1,290	1,660	1,984	2,364	2,626
200	1,286	1,653	1,972	2,345	2,601
300	1,284	1,650	1,968	2,339	2,592
400	1,284	1,649	1,966	2,336	2,588
500	1,283	1,648	1,965	2,334	2,586
$\infty$	1,282	1,645	1,960	2,327	2,576

Tabelle 3.1: Quantile der *Student t* Verteilung

tervall mit einer Wahrscheinlichkeit von 95% den wahren Wert beinhaltet. Für andere Konfidenzniveaus sind entsprechende Werte aus der Tabelle zu entnehmen. Durch die Angabe des Konfidenzniveaus wird deutlich, dass die Schätzung eines Parameters immer mit einer Wahrscheinlichkeit verbunden ist, die gewünschte Genauigkeit nicht zu erreichen.

### 3.4 Parameterschätzung für *Waferscans*

Jede erfolgreiche Parameterschätzung basiert auf einem mathematischen Modell, welches den realen Messprozess möglichst genau beschreibt. Das für *Flatwafer* erstellte Modell aus Gleichung (3.3) beschreibt die Messdaten des *Scanners* anhand des Flatwafermodells  $r_F[n]$ . Der *Offset* zwischen Welt- und Scannerkoordinatensystem  $R_0$  wird ebenso berücksichtigt, wie die Störsignale weißes Gauß'sches Rauschen  $w[n]$  und Quantisierungsrauschen  $q[n]$ . Bevor die Parameterschätzung des Flatwafermodells durchgeführt werden kann, müssen die vom *Scanner* gelieferten

Daten erst mittels des *Offsets*  $R_0$  korrigiert werden. Wie dieser *Offset* im Zuge einer Kalibrierung für die Messanordnung zu bestimmen ist, wird ausführlich in Kapitel 3.3 behandelt. Die Parameterschätzung von *Waferscans* beschäftigt sich nun mit der Schätzung der Parameter des Flatwafermodells  $r_F[n]$ , welche auf Basis einer numerischen Fehlerquadratminimierung durchgeführt wird. Die praktische Umsetzung erfolgt mit einem beliebigen numerischen *Solver* für nicht-lineares *Curve Fitting* wie dem *lsqcurvefit*, welcher unter *MATLAB* verfügbar ist.

### 3.4.1 Beispiel *Flatwafer Fitting (FWF)*

An dieser Stelle zeigt ein kurzes *MATLAB Script* die praktische Umsetzung des *FWF*. Der verwendete *MATLAB Solver* *lsqcurvefit* berechnet auf Basis des diskreten Messsignals die numerische Minimierung der übergebenen Funktion  $r_F[n]$ . Für das *Curve Fitting* muss lediglich der Vektor mit den Messdaten als auch die Funktion  $r_F[n]$  in Form eines M-Files dem Solver übergeben werden. Abbildung (3.10) zeigt ein *MATLAB Script* mit allen zur Durchführung der Parameterschätzung notwendigen Schritten. In Abbildung (3.11) ist die zugehörige Implementation dieses Modells angegeben. In einer Simulation von  $M = 1000$  Parameterschätzungen wurden die in Abbildung (3.12) dargestellten Histogramme aller geschätzten Parameter ermittelt. Die Leistung des additiven Rauschsignals beträgt  $0.05 \text{ mm}^2$ . Quantisierungsrauschen wurde in dieser Simulation nicht berücksichtigt. Tabelle 3.2 beinhaltet die Ergebnisse und stellt den Parametern  $\delta$  und  $\varphi_0$

	$\mu$	$\hat{\mu}$	$\hat{\sigma}$	$\sqrt{CRLB(\cdot)}$
$\delta$	15,0 mm	14,99 mm	0,0336 mm	0,0167 mm
$\phi_0$	1,5708 rad	1,5709 rad	0,0023 rad	0,0011 rad
$R_F$	69,2 mm	69,197 mm	0,0774 mm	-
$\gamma$	1,0472 rad	1,0472 rad	0,0048 rad	-
$R$	75,0 mm	75,001 mm	0,025 mm	-

Tabelle 3.2: Auswertung der *Flatwafer* Parameterschätzung

```

% Nonlinear LMS curve fitting
% for waferscans with flat

% N = number of samples
N = 360;
% Noise variance [mm²]
Un = 0.05;
wn = sqrt(Un)*randn(1,N);
% n vector from 0 to 2*pi (N-steps)
n = 0:(2*pi/N):(2*pi)-(2*pi/N);
% Parameter vector x = [delta, phi0, RF, gamma, R];
x = [15, pi/2, 69.2, pi/3, 75];
% Initial guess
x0 = [0, pi/3, 0, pi/2, 0];
% Lower bound for parameter estimation
lb = [0, 0, 60, 0, 50];
% Upper bound for parameter estimation
ub = [50, 2*pi, 70, 2*pi, 100];
% Waferscan simulation + AWGN
r = flatwaferscan(x,t) + wn;
% Start nonlinear curve fitting
[x_est,resnorm] = lsqcurvefit(@flatwaferscan,x0,n,r,lb,ub);

```

Abbildung 3.10: *MATLAB Script* für das *Flatwafer Fitting (FWF)*

die in Kapitel 3.3.3 hergeleiteten unteren Grenzen der jeweiligen Standardabweichungen gegenüber. Der *Solver* iteriert solange, bis eine untere Schranke bezüglich der Abweichung zwischen vorgegebenem und geschätztem Modell unterschritten wird. Allein diese Feststellung genügt, um aussagen zu können, dass eine optimale Lösung mit einem iterativen *Solver* nur angenähert werden kann. Eine optimale Lösung kann mit jenem erwartungstreuen Schätzer erreicht werden, welcher die kleinste Varianz für alle möglichen Werte der geschätzten Parameter liefert. So ein Schätzer wird als *Minimum Variance Unbiased Estimator (MVUE)* bezeichnet. Erreicht der Schätzer zudem noch die bestmögliche Varianz entsprechend der *Cramer-Rao* Ungleichung, wird der Schätzer als effizient bezeichnet.

```

function r = flatwaferScan(x,n)
% Calculation of radial waferScan including flat
% Parameter list: delta, phi0, rf, gamma, r

% R0 = Offset between coordinate system and wafer scanner [mm]
R0 = 50;

for k = 1:length(n)

% Calculation of circle
rK = x(1)*cos(n(k)-x(2))+sqrt(x(5)^2-x(1)^2*sin(n(k)-x(2))^2);

% Calculation of flat
rF = (x(1)*sin(x(2))+(x(3)*cos(x(2)+x(4))+x(1)*cos(x(2)))*tan(pi/2-
x(2)-x(4))+x(3)*sin(x(2)+x(4)))/(sin(n(k))+tan(pi/2-x(2)-
x(4))*cos(n(k)));

    if rF > 0
        if rK < rF
            r(k) = rK;
        else
            r(k) = rF;
        end
    else
        r(k) = rK;
    end
end

% Subtraction of R0
r = r - R0;

```

Abbildung 3.11: *MATLAB* Implementation der Funktion  $r_F$

### 3.4.2 Effizienz von Schätzern

Für einen unbekanntem Parameter  $\theta$  und der zugehörigen *Fisher Information*  $I(\theta)$  ist die Effizienz eines erwartungstreuen Schätzers  $\hat{\theta}$  nach 3.21 definiert.

$$e(\hat{\theta}) = \frac{1}{I(\theta) \cdot \text{var}\{\hat{\theta}\}} \quad (3.21)$$

Der Schätzer gilt als effizient wenn  $e(\hat{\theta}) = 1$  ist, was erreicht wird wenn seine Varianz laut (3.22) gleich der Inversen der *Fisher Information* ist.

$$\text{var}\{\hat{\theta}\} = I(\theta)^{-1} \quad (3.22)$$

Üblicherweise ist es jedoch ausreichend wenn der Schätzer asymptotisch effizient ist, sodass seine Effizienz mit wachsendem Stichprobenumfang  $N$  gegen Eins

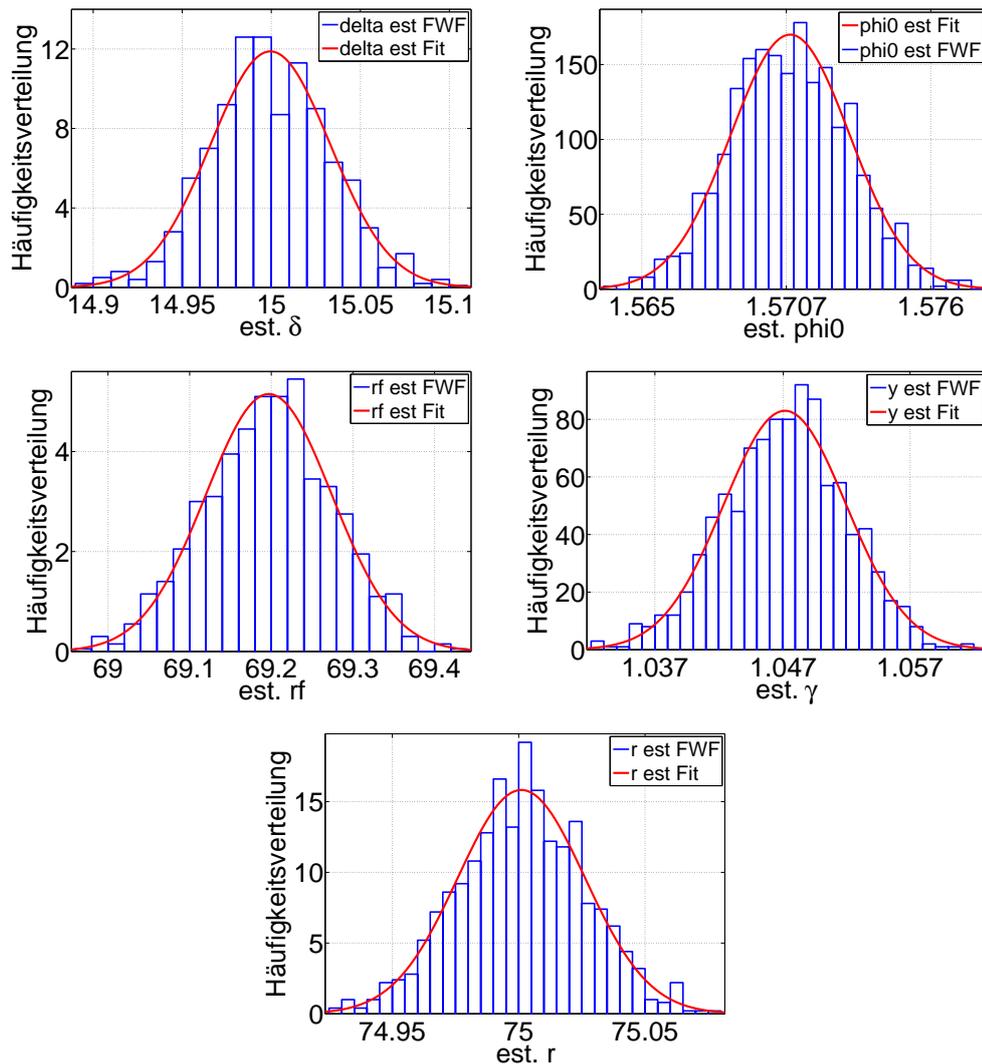


Abbildung 3.12: Häufigkeitsverteilungen aller *FWF* Parameter

konvergiert. Asymptotisch effiziente Schätzer können in vielen Fällen durch die *Maximum Likelihood* Methode realisiert werden.

### 3.5 Modellvereinfachung

Nach erfolgreichem Bestimmen des Parametervektors  $\underline{\theta}_F$ , kann das eigentliche Ziel der Waferzentrierung durchgeführt werden. Der Wafermittelpunkt ist durch die Parameter  $\delta$  und  $\varphi_0$  eindeutig bestimmt und kann korrigiert werden. In ei-

nem zweiten Schritt wird der *Wafer* um den Flatwinkel  $\gamma$  in die gewünschte Lage gedreht. Dabei wird klar, dass das vollständige Bestimmen des Waferscanmodells mehr Parameter erfordert, als für die Dezentrierung benötigt werden. Darüber hinaus beschreibt das Waferscanmodell  $r_F(\varphi)$  lediglich *Wafer* mit primärem *Flat*. Für alle anderen Wafertypen müssen die jeweiligen Modelle herangezogen werden, wobei jedes seinen eigenen Parameterschätzer (Schätzer) benötigt. Auf Grund der Forderung nach einem für alle Wafertypen universell einsetzbaren System käme zur Aufgabe der Parameterschätzung auch noch eine Klassifizierung des zu verwendenden Waferscanmodells hinzu. Dieser Aufwand soll mit einem vereinfachten Modell umgangen werden, welches es ermöglicht, *Wafer* unabhängig vom Typ zu behandeln. Es beruht auf der Gegebenheit, dass zuerst die Waferzentrierung und erst danach die Ausrichtung der Randmarkierung ausgeführt wird. Für die Waferzentrierung sind nur Messdaten des kreisförmigen *Wafers* von Interesse, weshalb der *Wafer* als homogene kreisförmige Scheibe beschrieben und jede Art von Randmarkierung im Modell ignoriert werden soll. Merkmale wie *Flat*, *Notch* u.a. werden dabei als grobe Ausreißer aufgefasst. Es besteht nun die Aufgabe, diese Ausreißer zu identifizieren und auszuschließen. Die Forderung nach einer Ausreißeridentifizierung wird in ihrer Notwendigkeit dadurch unterstützt, dass im realen Fertigungsprozess auch *Wafer* mit fehlerhaften Rändern, sogenannten Randausbrüchen, erfolgreich gehandhabt werden müssen.

### 3.5.1 Reduktion des Modells

Das bereits in Abbildung 2.3 vorgestellte Modell in Form eines fiktiven *Wafers* soll als Grundlage des vereinfachten Modells dienen. Die daraus abgeleitete mathematische Beschreibung für den Messprozess in Gleichung (3.23) enthält den verkürzten Parametervektor  $\underline{\theta}_K = [\delta, \varphi_0, R, R_0]^T$ .

$$s_K[n; \underline{\theta}_K] = \delta \cos\left(\frac{2\pi}{N}n + \varphi_0\right) + \sqrt{R^2 - \delta^2 \sin^2\left(\frac{2\pi}{N}n + \varphi_0\right)} - R_0 + w[n] + q[n] \quad (3.23)$$

Die Schätzung dieser Parameter basiert auf dem Einpassen eines Kreises in die Messdaten  $s_K[n; \underline{\theta}_0]$ , sodass der quadratische Fehler zwischen Modell und Mess-

daten minimal wird. Diese Problemklasse bezeichnet man als *Circle Fitting*, in welcher die Methode der kleinsten Fehlerquadrate (*Least Mean Square, LMS*) Verwendung findet. Mathematisch entspricht dies nun der Minimierung des Funktionals

$$\mathcal{F}_{\underline{\theta}_K} = \sum_{n=0}^{N-1} \left( s_K[n; \underline{\theta}_K] - r_K[n; \underline{\theta}_0] \right)^2 \quad (3.24)$$

mit  $\underline{\theta}_0 = [\delta, \varphi_0, R]$ . Unter der Annahme weißen Gauß'schen Rauschens mit Standardabweichung  $\sigma_0 > \Delta$  [15][16] und Mittelwert  $\mu = 0$  ist der LMS Algorithmus einem *Maximum Likelihood* Schätzer äquivalent. Für das in Gleichung (3.24) gegebene Funktional existiert jedoch keine geschlossene Lösung, da die Minimierungsaufgabe ein nichtlineares Problem darstellt. Bekannte Algorithmen sind entweder iterativer Natur und aufwendig oder führen grundsätzlich nur zu Näherungslösungen. In Hinblick auf einen einfachen, schnellen und robusten Algorithmus zur Lösung der Minimierungsaufgabe soll das Modell des Messprozesses auf Basis der Signalkomponenten untersucht werden. Sowohl der Kosinusterm als auch der Wurzelterm in (3.23) beinhalten die Parameter  $\delta$  und  $\varphi_0$ , wodurch das Modell überbestimmt ist. Wie aus der Gleichung hervorgeht, wird das Modell im Wesentlichen durch die Grundschiwingung  $\delta \cos\left(\frac{2\pi}{N}n + \varphi_0\right)$  und die  $\delta \sin^2\left(\frac{2\pi}{N}n + \varphi_0\right)$ -Funktion geprägt, welche zueinander die Eigenschaft der Orthogonalität besitzen. Es liegt nahe, die Modellbeschreibung auf die wesentliche Information zu reduzieren, indem  $R_0$  und der Wurzelterm in Gleichung (3.23) eliminiert werden. Bildet man aus  $s_K[n; \underline{\theta}_0]$  eine um  $\pi$  bzw.  $N/2$  verschobene Kopie [17] und vereinfacht die Terme, so erhält man:

$$\delta \cos\left(\frac{2\pi}{N}n + \varphi_0 + \frac{N}{2}\right) = -\delta \cos\left(\frac{2\pi}{N}n + \varphi_0\right) \quad (3.25)$$

und

$$\sqrt{R^2 - \delta^2 \sin^2\left(\frac{2\pi}{N}n + \varphi_0 + \frac{N}{2}\right)} = \sqrt{R^2 - \delta^2 \sin^2\left(\frac{2\pi}{N}n + \varphi_0\right)}. \quad (3.26)$$

Während der *cos*-Term (3.25) nach dem Verschieben dem Originalterm in Gleichung (3.23) mit negativem Vorzeichen entspricht, zeigt sich für den  $\pi$ -periodischen Wurzelterm (3.26) keine Veränderung zu Gleichung (3.23). Nach Subtraktion vom ursprünglichen, nicht verschobenen Modell mit der verschobenen

Kopie laut Gleichung (3.27)

$$s[n; \underline{\theta}] = s_K[n; \theta_0] - s_K[n + N/2; \theta_0] \quad (3.27)$$

verbleiben nur mehr der *cos*-Term und die resultierenden Rauschsignale  $\tilde{w}[n]$  und  $\tilde{q}[n]$ , welche das neue reduzierte Modell (3.28) mit dem Parametervektor  $\underline{\theta} = [\delta, \varphi_0]^T$  bilden.

$$s[n; \underline{\theta}] = 2 \delta \cos\left(\frac{2\pi}{N}n + \varphi_0\right) + \tilde{w}[n] + \tilde{q}[n] \quad (3.28)$$

Durch die Subtraktion der originalen mit der verschobenen Messkurve in Gleichung (3.27) folgt, dass unter Voraussetzung einer äquidistanten Abtastung der Signalmittelwert nach Gleichung (3.29) stets Null muss, da sich jede Form von Offset in den Daten aufhebt. Somit wird auch jeder von Null verschiedene Mittelwert  $\mu$  des Messrauschens eliminiert. Für die Verteilung von  $\tilde{w}[n]$  muss  $\tilde{\mu} = 0$  gelten!

$$\sum_{n=0}^{N-1} s[n; \underline{\theta}] \stackrel{!}{=} 0 \quad (3.29)$$

Mit  $\tilde{q}[n]$  wird das resultierende deterministische Quantisierungsrauschen beschrieben. In Hinsicht auf ein lineares Modell wird die *cos*-Funktion aus (3.28) in zwei Terme aufgespalten. Der resultierende Parametervektor lautet  $\underline{\theta}^* = [\delta_x, \delta_y]^T$ .

$$s[n; \underline{\theta}^*] = 2 \left( \delta_x \cos\left(\frac{2\pi}{N}n\right) + \delta_y \sin\left(\frac{2\pi}{N}n\right) + \frac{\tilde{w}[n]}{2} + \frac{\tilde{q}[n]}{2} \right) \quad (3.30)$$

### 3.5.2 Eigenschaften des reduzierten Modells

Das bereits vereinfachte Modell wurde um zwei weitere Parameter reduziert und gehört somit nicht mehr der Problemklasse *Circle Fitting* an. Es ist auf die Beschreibung des Wafermittelpunktes beschränkt und erfüllt den Wunsch sowohl vom Waferradius  $R$  als auch vom *Scanneroffset*  $R_0$  unabhängig zu sein. Da letzterer nicht mehr Teil des Modells ist, hat der Abstand des Messgerätes zum Rotationszentrum (*Chuck*) keinen Einfluss mehr. Das reduzierte Modell ist nun Teil der Problemklasse *Sine Wave Fitting (SWF)*, linear bezüglich der Parameter und benötigt keine aufwendigen iterativen und approximativen Lösungsverfahren.

## 4 *Sine Wave Fitting*

Unter *Sine Wave Fitting* (*SWF*) versteht man in diesem Zusammenhang das Einpassen einer Sinus-Funktion in ein winkel- und wertediskretes Messsignal, für welches das Modell laut Gleichung (4.1) gilt.

$$s[n] = A_0 \cos\left(\frac{2\pi}{N}n\right) + B_0 \sin\left(\frac{2\pi}{N}n\right) + C_0 + w[n] + q[n] \quad (4.1)$$

Prägend für den Begriff des *Sine Wave Fittings* [18] [19] war der IEEE Standard *1057-1994* [20], welcher sich ausführlich mit der Definition einheitlicher Terminologien und adäquater Testmethoden zur Evaluierung der Eigenschaften von Analog Digital Umsetzern (*ADU*) befasst. Anhand sinusförmiger Testsignale, werden Methoden vorgestellt, um Fehler von *ADU*'s messtechnisch erfassen und gegebenenfalls korrigieren zu können, wie es beispielsweise im Falle von Verstärkungs- und Offsetfehlern notwendig ist. Eine aktuelle Ausgabe des ursprünglichen Standards steht unter dem Titel *IEEE Standard for Terminology and Test Methods for Analog-to-Digital Converters, IEEE 1241-2000* [21] zur Verfügung. Die in diesen Standards beschriebenen Methoden sehen das Versorgen der zu untersuchenden Analog Digital Umsetzer mit sinusförmigen Testsignalen vor. Die am Ausgang der Umsetzer zur Verfügung stehenden diskreten Signale werden durch ein Drei oder Vier Parameter Fitting untersucht. Durch Vergleichen der geschätzten mit den tatsächlichen Modellparametern können Genauigkeit und *Performance* von *ADU*'s festgestellt werden. Beide Standards sind für die Schätzung des reduzierten Modells aus Kapitel 3.5.2 deshalb relevant, weil sie auf demselben mathematischen Modell beruhen. Die Berücksichtigung der Signalquantisierung stellt eines der Hauptmerkmale der Schätzaufgabe dar und wird im Standard ebenfalls behandelt. Unter der Voraussetzung einer bekannten Signalfrequenz beschreiben die Standards eine analytische Lösung für das Drei Parameter Fitting. Für unbekannte Signalfrequenzen ist das iterative Vier Parameter Fitting anzuwenden. Da

beim Scannen von *Wafeln* die Signalfrequenz bzw. die Anzahl der Messwerte pro Umdrehung exakt bekannt ist, kann in diesem Fall das Drei Parameter Fitting angewendet werden.

## 4.1 Drei Parameter Fitting

Unter der Voraussetzung, dass die Signalfrequenz des Messsignals laut Definition (4.1) bekannt ist, kann ein Drei Parameter Fitting durchgeführt werden. Ausgehend vom Datenvektor  $\underline{s}$ , welcher die gemessenen Werte  $[s_0, s_1, \dots, s_{N-1}]^T$  beinhaltet, und den zugehörigen diskreten Winkelschritten  $n_0, n_1, \dots, n_{N-1}$ , wird in den IEEE Standards 1057-1994 [20] und 1241-2000 [21] ein Algorithmus zur Schätzung des Parametervektors  $\underline{\theta} = [A_0, B_0, C_0]^T$  beruhend auf Matrix Operationen vorgestellt. Dieser minimiert die Summe der Fehlerquadrate aus Gleichung (4.2),

$$\sum_{n=0}^{N-1} \left[ s_n - A_0 \cos\left(\frac{2\pi}{N}n\right) - B_0 \sin\left(\frac{2\pi}{N}n\right) - C_0 \right]^2 = e^T e \rightarrow \min \quad (4.2)$$

wobei die  $N \times 3$  Matrix  $\underline{\mathbf{D}}$  laut (4.3) definiert wird.

$$\underline{\mathbf{D}} = \begin{pmatrix} \cos \omega n_0 & \sin \omega n_0 & 1 \\ \cos \omega n_1 & \sin \omega n_1 & 1 \\ \vdots & \vdots & \vdots \\ \cos \omega n_{N-1} & \sin \omega n_{N-1} & 1 \end{pmatrix} \quad (4.3)$$

Die zu minimierende Fehlerquadratsumme  $e^T e$  aus Gleichung (4.2) wird in Matrixnotation laut Gleichung (4.4) angegeben.

$$(\underline{s} - \underline{\mathbf{D}} \hat{\underline{\theta}})^T (\underline{s} - \underline{\mathbf{D}} \hat{\underline{\theta}}) = \underline{e}^T \underline{e} \rightarrow \min \quad (4.4)$$

Gleichungen (4.5) bis (4.7) beschreiben die Ableitungen der Fehlerquadratsumme in Vektornotation.

$$\nabla_{\hat{\underline{\theta}}}(\underline{e}^T \underline{e}) = \nabla_{\hat{\underline{\theta}}}(\underline{e}^T) \underline{e} + \nabla_{\hat{\underline{\theta}}}(\underline{e}^T) \underline{e} = 2 \nabla_{\hat{\underline{\theta}}}(\underline{e}^T) \underline{e} \quad (4.5)$$

$$\nabla_{\hat{\underline{\theta}}}(\underline{e}^T) = \nabla_{\hat{\underline{\theta}}}(\underline{s} - \underline{\mathbf{D}} \hat{\underline{\theta}})^T = -\nabla_{\hat{\underline{\theta}}}(\hat{\underline{\theta}}^T \underline{\mathbf{D}}^T) \quad (4.6)$$

$$-\nabla_{\hat{\theta}}(\hat{\theta}^T \underline{\mathbf{D}}^T) = -\underbrace{\nabla_{\hat{\theta}}(\hat{\theta}^T)}_I \underline{\mathbf{D}}^T - \underbrace{\nabla_{\hat{\theta}}(\underline{\mathbf{D}})}_0 \hat{\theta} = -\underline{\mathbf{D}}^T \quad (4.7)$$

Die eigentliche Minimierung erfolgt nun durch Null setzen der Ableitung von  $e^T e$ .

$$\nabla_{\hat{\theta}}(\underline{\mathbf{e}}^T \underline{\mathbf{e}}) = -2\underline{\mathbf{D}}^T(\underline{\mathbf{s}} - \underline{\mathbf{D}} \hat{\theta}) \stackrel{!}{=} 0 \quad (4.8)$$

$$\underline{\mathbf{D}}^T \underline{\mathbf{D}} \hat{\theta} = \underline{\mathbf{D}}^T \underline{\mathbf{s}} \quad (4.9)$$

Die Parameterschätzung auf Basis der Fehlerquadratminimierung aus Gleichung (4.9) kann nun mittels Vorschrift (4.10) angegeben werden.

$$\hat{\theta} = (\underline{\mathbf{D}}^T \underline{\mathbf{D}})^{-1} (\underline{\mathbf{D}}^T \underline{\mathbf{s}}) \quad (4.10)$$

Eine Gegenüberstellung der Parametervektoren des *SWF* und des reduzierten Modells aus (3.30) ist in (4.11) dargestellt. Mit  $C_0 = 0$  darf beim reduzierten Modell auf Grund der Definition aus (3.29) auf die Schätzung dieser Konstante verzichtet werden.

$$\begin{pmatrix} A_0 \\ B_0 \\ C_0 \end{pmatrix} = \begin{pmatrix} \delta_x \\ \delta_y \\ 0 \end{pmatrix} \quad (4.11)$$

Die Auswertung von *Waferscans* vereinfacht sich somit auf die Schätzung der für das *Prealignment* relevanten Dezentrierungsparameter  $\delta_x$  und  $\delta_y$ .

### 4.1.1 Zwei Parameter Fitting

Die in Kapitel 3.5 beschriebene Modellvereinfachung sieht in einem ersten *Prealignment*vorgang die Schätzung und mechanische Korrektur der Dezentrierung unabhängig vom Wafertyp vor. *Flat* oder *Notch* werden dabei nicht berücksichtigt. In einem optionalen zweiten Schritt kann dann die Randmarkierung vom bereits zentrierten *Wafer* gesucht werden. Je nach Vorgabe wird anschließend die Randmarkierung auf die gewünschte Winkelposition gedreht. Diese Vorgangsweise erweist sich auch deshalb als vernünftig, weil nicht in allen Fällen ein vollständiges *Prealignment* mit Winkelausrichtung des *Wafers* verlangt wird. Durch den Verzicht auf die Schätzung aller Waferparameter wird der Rechenaufwand

drastisch verringert und auf die einzig interessante Information der Dezentrierung reduziert. Zusätzlich ist das reduzierte Modell universell auf alle Wafertypen anwendbar, da jede Art von Randmarkierung unberücksichtigt bleibt. Bei einer vollständigen Schätzung aller Modellparameter muss nach erfolgter Positionskorrektur der *Wafer* ohnehin noch zur gewünschten Winkelposition gedreht werden. Das vollständige Prealigmentverfahren unterscheidet sich demnach vom reduzierten nur dadurch, dass bei ersterem die Winkelposition schon vor dem Drehen des *Wafers* auf seine Endposition bekannt ist und beim reduzierten die Randmarkierung während der Drehung zur Endposition erst gesucht werden muss. Da die Drehung des *Wafers* in jedem Fall notwendig ist, liegt es nahe die Mittelpunktbestimmung des *Wafers* und die Winkelschätzung separat durchzuführen.

Abgeleitet vom Drei Parameter Fitting wird die Schätzung der Parameter  $\delta_x$  und  $\delta_y$  mittels der  $N \times 2$  Matrix  $\mathbf{D}^*$  aus (4.12) und des reduzierten Parametervektors  $\hat{\underline{\theta}}^*$  nach (4.13) durchgeführt.

$$\mathbf{D}^* = \begin{pmatrix} \cos \omega n_0 & \sin \omega n_0 \\ \cos \omega n_1 & \sin \omega n_1 \\ \vdots & \vdots \\ \cos \omega n_{N-1} & \sin \omega n_{N-1} \end{pmatrix} \quad (4.12)$$

$$\hat{\underline{\theta}}^* = \begin{pmatrix} \hat{\delta}_x \\ \hat{\delta}_y \end{pmatrix} \quad (4.13)$$

Für den unbekannt Parametervektor  $\hat{\underline{\theta}}^*$  gilt nach (4.10) der in Gleichung (4.14) angegebene fehlerquadratminimierende Schätzer.

$$\hat{\underline{\theta}}^* = (\mathbf{D}^{*T} \mathbf{D}^*)^{-1} (\mathbf{D}^{*T} \underline{s}) \quad (4.14)$$

Die Darstellung des Wafermittelpunktes in Polarform ist für den Dezentrierungsvektor  $\vec{\delta}$  in (4.15) angegeben.

$$|\hat{\delta}| = \sqrt{\hat{\delta}_x^2 + \hat{\delta}_y^2} \quad (4.15)$$

Für die Berechnung des Nullphasenwinkels  $\varphi_0$  müssen Fallunterscheidungen laut (4.16) berücksichtigt werden, da durch das Verhältnis von  $\delta_y/\delta_x$  die Information

des zugehörigen Quadranten verloren geht.

$$\hat{\varphi}_0 = \begin{cases} \tan^{-1}\left(\frac{\delta_y}{\delta_x}\right) & \delta_x > 0 \\ \pi - \tan^{-1}\left(\frac{-\delta_y}{\delta_x}\right) & \delta_x < 0 \\ +\pi/2 & \delta_x = 0, \delta_y > 0 \\ -\pi/2 & \delta_x = 0, \delta_y < 0 \end{cases} \quad (4.16)$$

### 4.1.2 Quantisierungsfehler

Die Methode der kleinsten Fehlerquadrate ist bekannt für ihre hervorragenden Ergebnisse und gilt als eine der besten. Die Eigenschaft dieser Methode einem *Maximum Likelihood* Schätzer äquivalent zu sein, unterliegt jedoch gewissen Voraussetzungen [6]. Die Abweichungen zwischen dem wahren Sinus-Verlauf und den Messwerten müssen zufällig und voneinander unabhängig sein [22] [23]. Die Verteilung der Fehler muss einen Mittelwert  $\mu = 0$  haben und gaußverteilt sein. Im Falle der *SWF* sind diese notwendigen Voraussetzungen grundsätzlich jedoch nicht erfüllt. Der Quantisierungsfehler ist teilweise deterministisch und die Messwerte sind stark voneinander abhängig. Vor allem an den Scheitelpunkten der Sinus-Funktion treten größere Quantisierungsfehler auf. Zur Verdeutlichung zeigt Abbildung 4.1 ein Sinussignal vor und nach der Quantisierung. Der Quantisierungsfehler ist weitgehend sägezahnähnlich und gleichförmig verteilt. Für Bereiche maximaler Amplituden trifft das jedoch nicht mehr zu. Abbildung 4.2 stellt Histogramme von Quantisierungsfehlern beruhend auf den Quantisierungsstufen  $\Delta_1$  und  $\Delta_2$  dar. Für große Quantisierungsstufen  $\Delta_1$  zeigt das Histogramm  $H_1$ , dass keine gleichförmige Häufigkeitsverteilung des Quantisierungsfehlers mehr vorliegt. Mit der kleineren Quantisierungsstufenhöhe  $\Delta_2$  kann  $H_2$  hingegen näherungsweise als gleichförmig verteilt angesehen werden. Um den Einfluss des Quantisierungsrauschens gering zu halten, ist also darauf zu achten für die Analog Digital Umsetzung eine ausreichend hohe Auflösung gewährleisten zu können, welche um ein Vielfaches unter der maximalen Signalamplitude liegen muss. Sollte aus welchen Gründen auch immer die Auflösung des Umsetzers nicht groß genug sein,

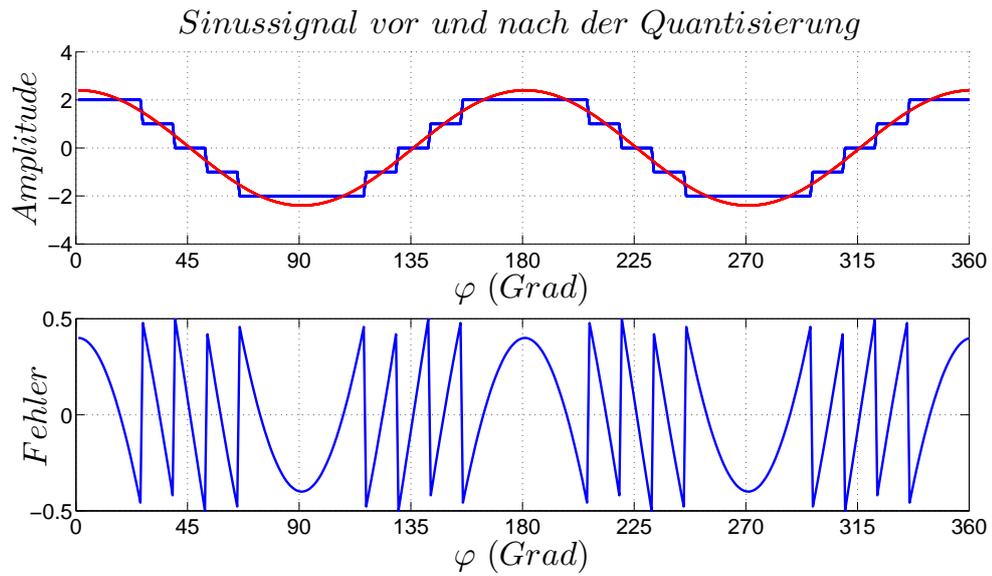


Abbildung 4.1: Ideal quantisiertes Sinussignal (Amplitude:  $2.5 \Delta$ ,  $\Delta=1$ )

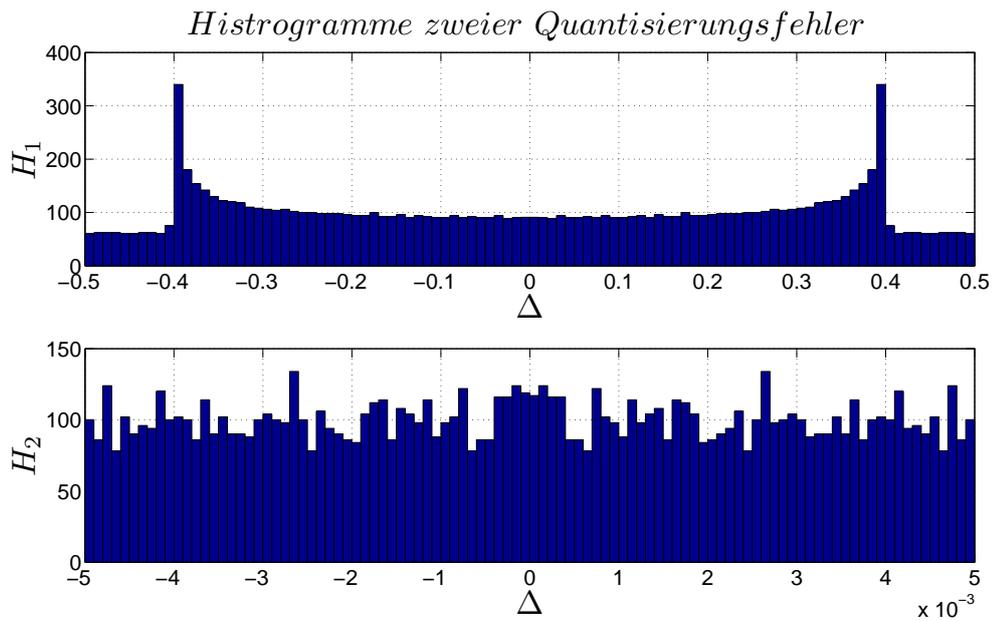


Abbildung 4.2: Histogramme für  $\Delta_1 = 1$  und  $\Delta_2 = 0.01$

um eine annähernd gleichförmige Verteilung erreichen zu können, wird in [24] das *SWF* für geringe Auflösungen von Analog Digital Umsetzern beschrieben. Auf Basis der dort untersuchten Histogramme des Quantisierungsfehlers wird vorge-

schlagen, die digitalisierten Werte aus den Scheitelbereichen des sinusförmigen Signals vor der Parameterschätzung laut Abbildung 4.3 zu verwerfen.

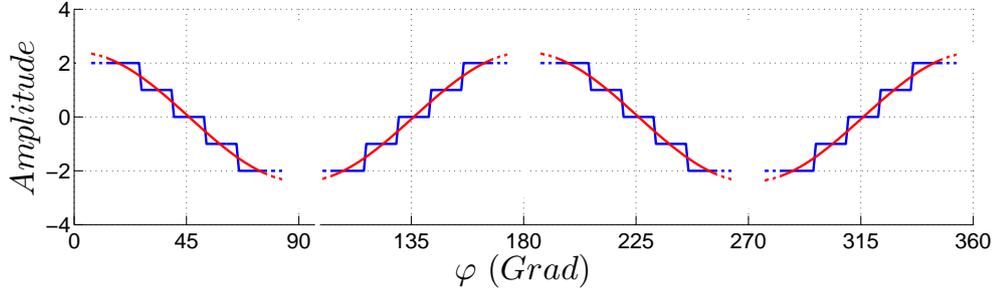


Abbildung 4.3: Verwerfen von *Samples* bei großen Quantisierungsstufen

## 4.2 *CRLB* für $\text{var} \{ \hat{\delta} \}$ und $\text{var} \{ \hat{\varphi}_0 \}$

Für das in Gleichung (3.28) beschriebene reduzierte Modell

$$s[n; \theta] = 2 \delta \cos\left(\frac{2\pi}{N}n + \varphi_0\right) + \tilde{w}[n] + \tilde{q}[n]$$

wird nun die untere Grenze der Varianz für  $\hat{\delta}$  und  $\hat{\varphi}_0$  untersucht. Sowohl das im Modell beschriebene additive weiße Gauß'sche Rauschen  $\tilde{w}[n]$  als auch das Quantisierungsrauschen  $\tilde{q}[n]$ , ist bei der Bestimmung der *CRLB* zu berücksichtigen [25][26]. Für das nach Definition (3.29) zwingend mittelwertfreie Gauß'sche Rauschsignal  $\tilde{w}[n]$  gilt die Verteilung  $\mathcal{N}(\theta, \sigma_\theta^2)$ . Der Quantisierungsfehler  $q[n]$  wird unter der Voraussetzung ausreichender Auflösung und unter der Bedingung  $\sigma_0 > \Delta$  mit Verweis auf [22] als Zufallsvariable mit Mittelwert Null und Varianz  $\Delta^2/12$  modelliert. Die Verteilung des Quantisierungsrauschens  $q[n]$  im Intervall  $[-\Delta/2, \Delta/2]$  kann daher mit hinreichender Genauigkeit als gleichförmig angenommen werden. Die Vorschrift zur Bestimmung der *CRLBs* von  $\hat{\delta}$  und  $\hat{\varphi}_0$  wird daher nach Definition (4.17) angegeben.

$$CRLB: \text{var} \{ \hat{\theta} \} \geq \frac{\sigma_0^2 + \frac{\Delta^2}{12}}{\sum_{n=0}^{N-1} \left( \frac{\partial s[n; \theta]}{\partial \theta} \right)^2} \quad (4.17)$$

Nach (4.17) sind in (4.18) und (4.19) die partiellen Ableitungen für die *CRLBs* von  $\hat{\delta}$  und  $\hat{\varphi}_0$  angegeben [26].

$$CRLB: \text{var} \{ \hat{\delta} \} \geq \frac{\sigma_0^2 + \frac{\Delta^2}{12}}{\sum_{n=0}^{N-1} \cos^2\left(\frac{2\pi}{N}n + \varphi_0\right)} \approx \frac{2(\sigma_0^2 + \frac{\Delta^2}{12})}{N} \quad (4.18)$$

$$CRLB: \text{var} \{ \hat{\varphi}_0 \} \geq \frac{\sigma_0^2 + \frac{\Delta^2}{12}}{\sum_{n=0}^{N-1} \delta^2 \sin^2\left(\frac{2\pi}{N}n + \varphi_0\right)} \approx \frac{2(\sigma_0^2 + \frac{\Delta^2}{12})}{N\delta^2} \quad (4.19)$$

## 5 *Fitting Performance*

Dieses Kapitel widmet sich der Untersuchung der bisher vorgestellten *Curve Fitting* Varianten. Der Anspruch an die Parameterschätzung liegt bei der Schätzung der Modellparameter  $\delta$  und  $\varphi_0$  sowie deren Varianzen. Die *Performance* für das *Sine Wave Fitting* (*SWF*), das *Circle Fitting* (*CF*) und das *Flatwafer Fitting* (*FWF*) sollen vorerst anhand mathematischer Modelle untersucht werden, um die geschätzten Werte denen der ursprünglichen Modellparameter gegenüberstellen zu können. Anschließend werden die erarbeiteten Erkenntnisse zur Schätzung eines tatsächlichen *Waferscans* herangezogen. Beginnend mit dem *SWF* soll zunächst die Frage beantwortet werden, wie genau diese Parameterschätzung überhaupt durchgeführt werden kann. Als Mittel dieser Qualitätsbeurteilung dient eine Varianzschätzung der Parameter  $\hat{\delta}$  und  $\hat{\varphi}_0$ . Eine Aussage über die kleinste erreichbare Varianz der Parameter wird mit Hilfe der in Kapitel 4.2 hergeleiteten *CRLBs* der gesuchten Parameter gefunden. Diese untere Grenze beruht nach (4.18) und (4.19) auf der, dem Messsignal überlagerten, Rauschleistung  $\sigma_0^2$  und dem Quantisierungsrauschen  $\Delta^2/12$  [26]. Im Falle des Quantisierungsrauschens ist davon auszugehen, dass die Auflösung des verwendeten Messgerätes bekannt ist. Für das additive weiße Gauß'sche Rauschen liegt im Allgemeinen keine Information vor. Zur Bestimmung der Varianzen der geschätzten Parameter ist somit zusätzlich eine Schätzung der im Messsignal enthaltenen Rauschleistung erforderlich. Um die grundlegende *Performance* des *SWF* erarbeiten zu können, wird in einem ersten Schritt auf das ursprüngliche reduzierte mathematische Modell  $s[n; \underline{\theta}]$  aus (3.28) zurückgegriffen.

## 5.1 Verzernte Rauschleistungsschätzung

Anhand einer Simulation wird dem reduzierten Modell ein künstliches Rauschsignal  $w[n]$  der Leistung  $\sigma_0^2 = 0.001 \text{ mm}^2$  überlagert. *Flat* oder *Notch* werden zunächst nicht modelliert. Die Auflösung des Quantisierers wird mit  $\Delta = 50 \text{ }\mu\text{m}$  festgelegt. Die in Abbildung 5.1 und 5.2 dargestellten Histogramme, zeigen die Verteilungen von  $K = 250$  Differenzen zwischen vorgegebener Rauschleistung  $\bar{\sigma}_0^2$  und geschätzter Rauschleistung  $\bar{\sigma}^{*2}$  verschiedener Parameterschätzer. Jeder Schätzwert für die Rauschleistung wird in der Simulation als Mittelwert aus  $M = 100$  Parameterschätzungen bestimmt. Der Simulation zufolge, sind die

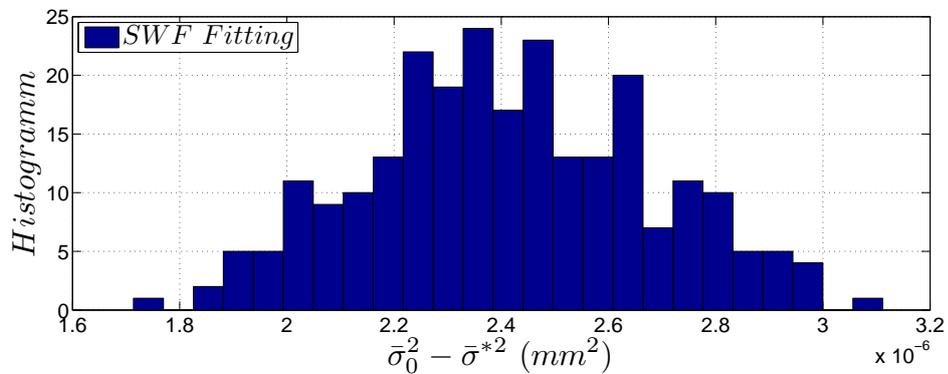


Abbildung 5.1: Verteilung von  $\bar{\sigma}_0^2 - \bar{\sigma}^{*2}$  des *SWF*

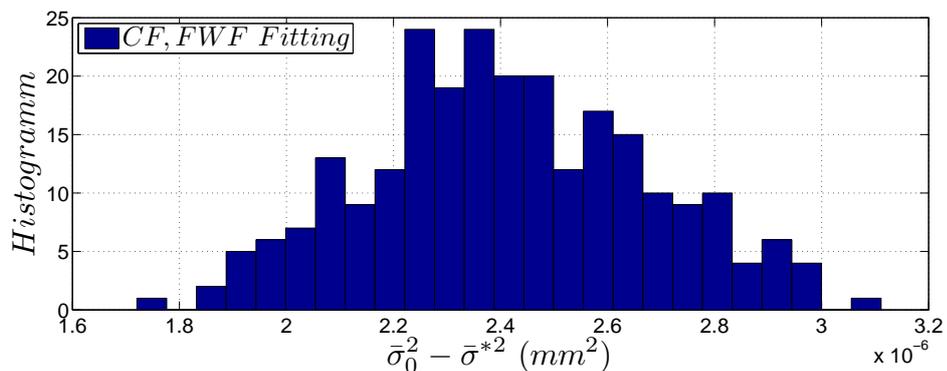


Abbildung 5.2: Verteilung von  $\bar{\sigma}_0^2 - \bar{\sigma}^{*2}$  von *CF* und *FWF*

Differenzen  $\bar{\sigma}_0^2 - \bar{\sigma}^{*2}$  aller Schätzer strikt positiv. Im Mittel liegt die geschätzte Rauschleistung also unter der tatsächlichen. Der Grund für dieses Ergebnis ist

nicht etwa ein Fehler, sondern die Eigenschaft des Schätzers eine Lösung zu finden, welche die Fehlerquadratsumme minimiert. Genau diese Fehlerquadratsumme ist der Schätzwert für die Rauschleistung  $\sigma^{*2}$ . Sie berechnet sich aus den quadrierten Differenzen zwischen geschätztem Modell und den ursprünglichen Daten laut Gleichung (5.1).

$$\sigma^{*2} = \frac{1}{N-1} \sum_{n=0}^{N-1} (s[n; \hat{\theta}] - r[n])^2 \quad (5.1)$$

Die Diskrepanz liegt in der Auswertung der Simulation bzw. der Interpretation der Ergebnisse. Dem simulierten Messdatenvektor wurde künstlich ein zufälliges Rauschsignal hinzu addiert. Die Schätzer bestimmen nun aber nicht die ursprünglichen Modellparameter, sondern jene, welche Modell samt Rauschen am besten annähern. Die Abweichung zwischen tatsächlichem und geschätztem Wert wird Verzerrung oder Bias genannt. Der in Gleichung (5.2) definierte Bias ist nun die Differenz zwischen dem Erwartungswert für die Rauschleistung von  $w[n]$  und dem Erwartungswert für die geschätzte Rauschleistung  $\sigma^{*2}$ . Nur wenn die Anzahl der Messwerte  $N$  gegen unendlich strebt, wird die geschätzte Rauschleistung auch der tatsächlichen entsprechen, was als asymptotische Erwartungstreue bezeichnet wird. Sowohl das *SWF* als auch das *CF* und *FWF* sind für eine endliche Anzahl von Messwerten verzerrte Schätzer für die Rauschleistung, dessen Schätzung im Mittel unter der tatsächlichen Rauschleistung liegt.

$$\text{Bias}(\sigma^{*2}) = E\{w^2[n]\} - E\{(s[n; \hat{\theta}] - r[n])^2\} > 0 \quad (5.2)$$

## 5.2 Unverzerrte Rauschleistungsschätzung

In der Praxis können immer nur endlich viele Messwerte aufgenommen werden, weshalb nun die Abhängigkeit der geschätzten Rauschleistung in Bezug auf die Anzahl der Messdaten untersucht werden soll. Abbildung 5.3 zeigt die geschätzte Rauschleistung  $\sigma^{*2}$  für die *Curve Fitting* Varianten *SWF*, *CF* und *FWF*, welche mit der Differenz  $\sigma_0^2 - \sigma^{*2}$  normiert ist, wobei  $\sigma_0^2$  die in der Simulation vorgegebene Rauschleistung darstellt. Mit guter Näherung dürfen zwischen den normierten Rauschleistungen und der Anzahl der Messwerte  $N$ , lineare Zusammenhänge abgelesen werden. Die Steigung beim *CF* und *FWF* liegt ungefähr bei 1/2, die des

$SWF$  beträgt ca.  $1/4$ . Dieses Verhalten ist in den Gleichungen (5.4) für das  $CF$

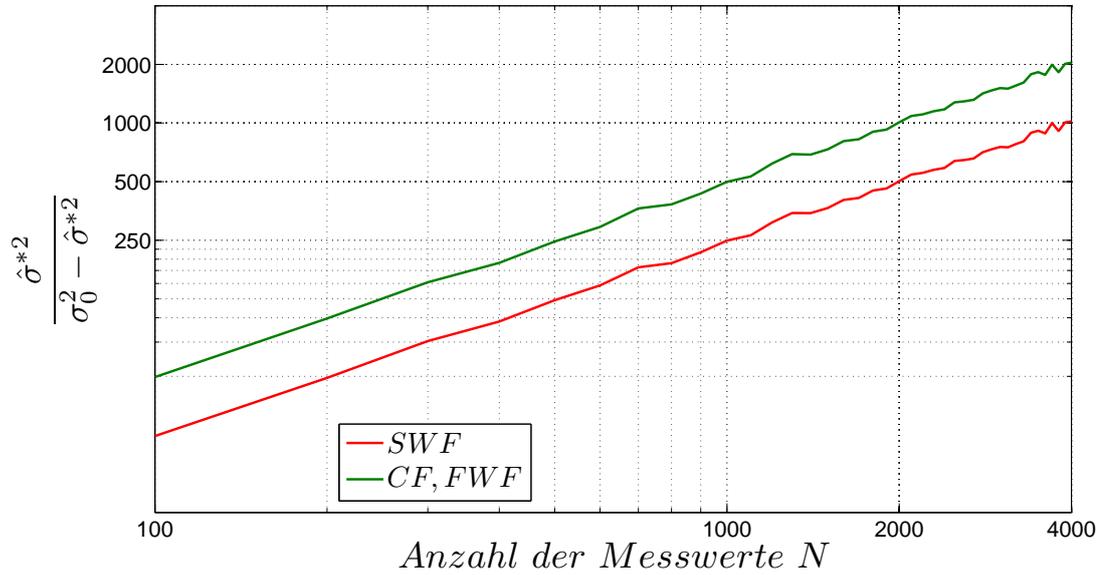


Abbildung 5.3: Normiertes  $\hat{\sigma}^{*2}$  von  $SWF$ ,  $CF$  und  $FWF$

und  $FWF$  sowie in Gleichung (5.3) für das  $SWF$  festgehalten.

$$SWF : \frac{\sigma^{*2}}{\sigma_0^2 - \sigma^{*2}} \approx \frac{1}{4}N \quad (5.3)$$

$$CF, FWF : \frac{\sigma^{*2}}{\sigma_0^2 - \sigma^{*2}} \approx \frac{1}{2}N \quad (5.4)$$

Durch Umformung auf die Rauschleistung  $\sigma_0^2$  können mit den Gleichungen (5.5) und (5.6) Korrekturfaktoren für die Rauschleistungsschätzung, ähnlich der korrigierten Stichprobenvarianz, angegeben werden.

$$SWF : \sigma_0^2 \approx \sigma^{*2} \left(1 + \frac{4}{N}\right) \quad (5.5)$$

$$CF, FWF : \sigma_0^2 \approx \sigma^{*2} \left(1 + \frac{2}{N}\right) \quad (5.6)$$

Für das  $SWF$ , zeigt Abbildung 5.4 die Verteilung der Differenzen zwischen tatsächlicher Rauschleistung  $\sigma_0^2$  und den korrigierten Rauschleistungsschätzungen  $\sigma^{*2}$ . Die Differenzen sind annähernd mittelwertfrei verteilt, sodass die korrigierte Schätzung für die Rauschleistung nun als erwartungstreu bezeichnet werden darf.

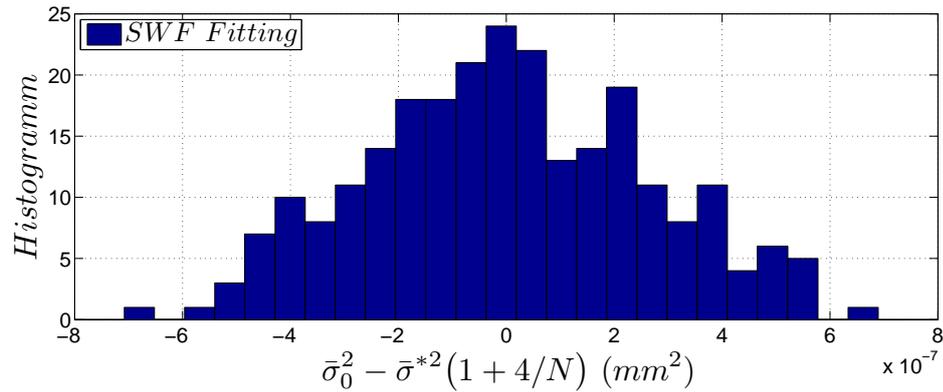


Abbildung 5.4: Korrigierte Rausleistungsschätzung für das *SWF*

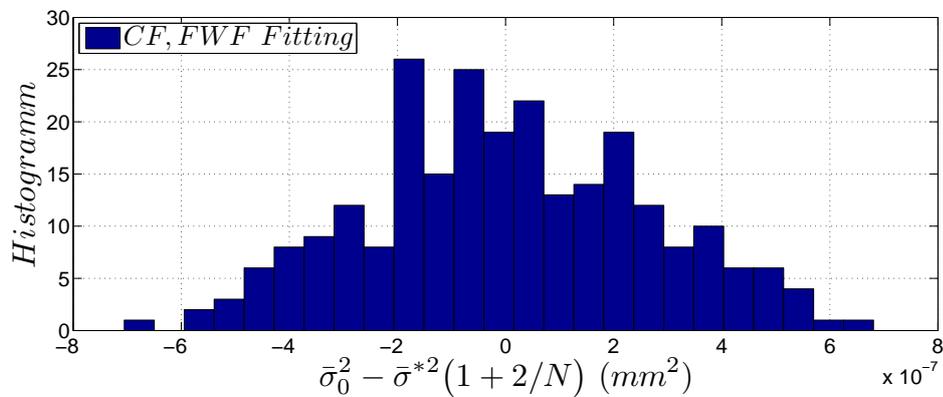


Abbildung 5.5: Korrigierte Rausleistungsschätzung für *F* und *FWF*

In Abbildung 5.5 wird das Ergebnis für die korrigierte Schätzung der Rausleistung für das *CF* und *FWF* gezeigt. Mit Hilfe dieser korrigierten Schätzungen der Rausleistungen können sogleich auch die Varianzen der geschätzten Parameter realistischer angegeben werden.

### 5.3 *SWF* anhand des Flatwafermodells

Die vorhergehenden Simulationen zeigen die korrigierte Schätzung der Rausleistung für das reduzierte Modell. Nun soll die Qualität der Parameterschätzung anhand einer Messkurve mit Randmarkierung untersucht werden. Abbildung 5.6 zeigt das Ergebnis eines *SWF* anhand eines Flatwafermodells. Wie es das *SWF*

verlangt, wird zwar die Differenz zwischen dem Messdatenvektor  $s_F[n]$  und seiner verschobenen Kopie berechnet, die Flatbereiche, welche nicht im reduzierten Modell berücksichtigt werden, führen jedoch zum Versagen der Schätzung. Trotz

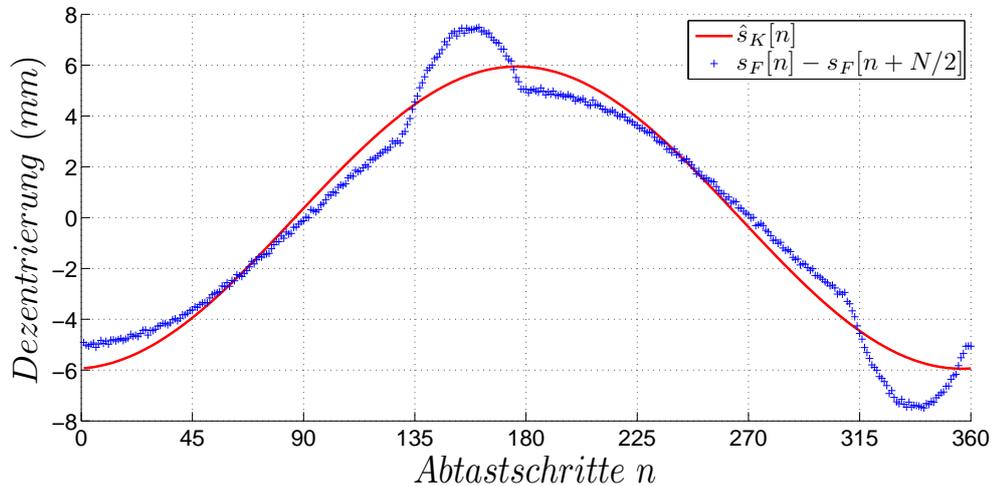


Abbildung 5.6: Mislungenes  $SWF$  eines *Flatwafer Scans*

der großen Abweichung zwischen den beiden Signalverläufen, stellt die Schätzung jene optimale Lösung dar, welche in Summe die kleinste quadrierte Abweichung zum Messsignal besitzt. Da im  $SWF$  keine Randmarkierung berücksichtigt ist und der analytische Schätzer stets alle im Messdatenvektor vorhandenen Einträge zur Bildung des Ergebnisses verwendet, führen Ausreißer oder Abweichungen des mathematischen Modells zu schwerwiegenden Verfälschungen des Ergebnisses. Für das  $SWF$ , sowie für jedes andere  $LMS$  Verfahren auch, ist es daher zwingend notwendig die Messdaten vor zu verarbeiten, sodass keine groben Ausreißer die Schätzung beeinflussen können. Abbildung 5.7 zeigt das um die Flatbereiche verkürzte Messsignal, welches nun dem mathematischen Modell des  $SWF$  genügt. Die Auswahl nicht gültiger *Samples* kann beispielsweise durch das in Kapitel 3.3.1 beschriebene Kreuzkorrelieren zwischen Messkurve und Flatverlauf durchgeführt werden. In jedem Fall bedeutet es jedoch einen Informationsverlust, welcher im Falle eines 6 Zoll *Wafers* bei ca. 30 Prozent der Daten liegt. Dementsprechend steigt auch die Varianz der geschätzten Parameter wie in Abbildung 5.8 für  $\hat{\delta}$  und  $\hat{\varphi}_0$  unter Berücksichtigung eines additiven weißen Gauß'schen Rauschens von  $\sigma_0^2 = 0.001 \text{ mm}^2$  und einer Auflösung des Quantisierers von  $\Delta = 0.05 \text{ mm}$  darge-

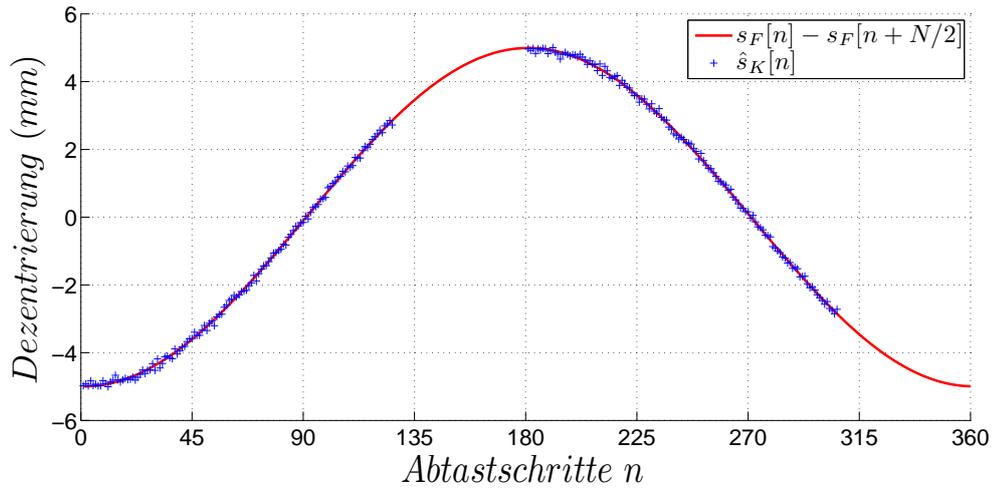


Abbildung 5.7: Erfolgreiches *SWF* anhand eines reduzierten Messdatenvektors

stellt ist.

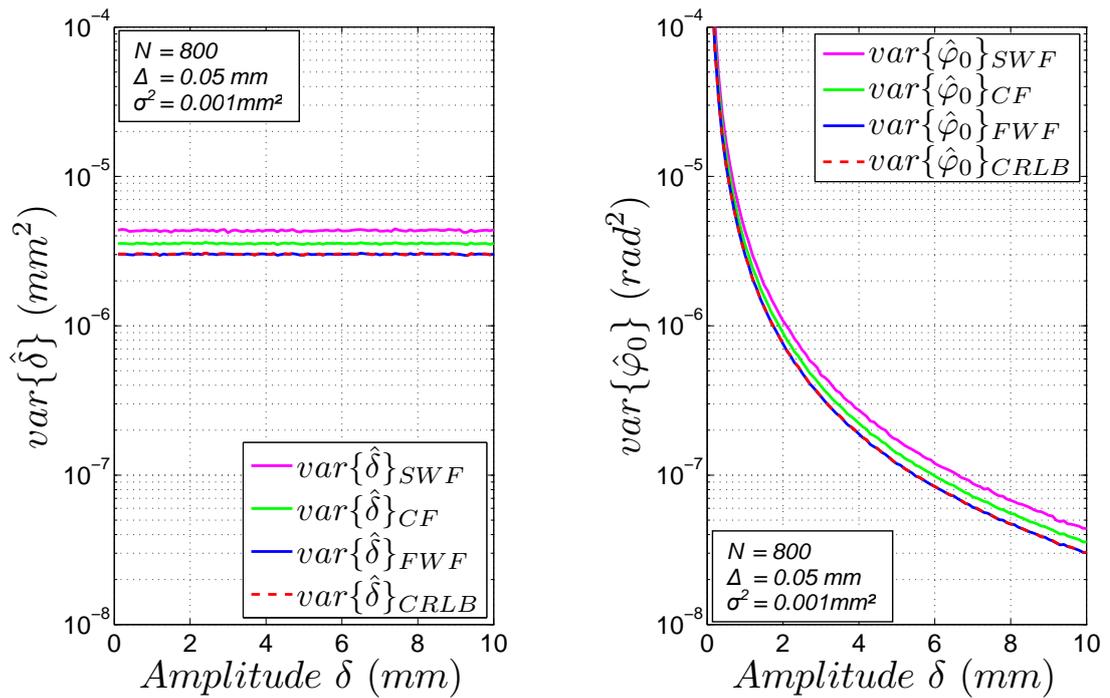


Abbildung 5.8:  $var\{\hat{\delta}\}$  und  $var\{\hat{\varphi}_0\}$  für *SWF*, *CF* und *FWF* für *Flatwafer*

## 5.4 Vergleich von *SWF*, *CF* und *FWF*

Die Diagramme in Abbildung 5.8 zeigen neben den geschätzten Varianzen für  $\hat{\delta}$  und  $\hat{\varphi}_0$  verschiedener Schätzer, die Grenzen der kleinsten theoretisch erreichbaren Varianzen (*CRLB*). Interessant ist nun die Frage, nach den Gründen für die Unterschiede zwischen den jeweiligen Schätzern. Die Bestimmung der theoretischen Grenzen für  $\text{var} \{ \hat{\delta} \}_{CRLB}$  und  $\text{var} \{ \hat{\varphi}_0 \}_{CRLB}$  beruht auf dem reduzierten Modell mit dem vollständigen Messdatenvektor der Länge  $N$ . Da beim *SWF* jedoch 30 Prozent der Messdaten verworfen werden müssen, erreicht die geschätzte Varianz der Parameter die theoretische Grenze nicht. Für das *CF* sind nur rund 15 Prozent der Messdaten zu verwerfen. Das *FWF* verwendet als einziges Verfahren den gesamten Messdatenvektor. Der Abstand der erreichbaren Varianzen des *CF* und *FWF* zu den Varianzen des *SWF* beruht nun auf den Eigenschaften des reduzierten Modells. Bei Vernachlässigung von *Flat* oder *Notch*, erhält man durch die Subtraktion aus ursprünglichen und verschobenen Messdaten ein Signal mit doppelter Amplitude laut Gleichung (3.28). Die Leistung des Kosinus-Signales vervierfacht sich dabei, während sich die unkorrelierte Rauschleistung lediglich verdoppelt. Damit einhergehend wird eine Verbesserung des Signal Rausch Abstandes (*Signal to Noise Ratio, SNR*) um den Faktor zwei erreicht. Zugleich werden aber weniger als die Hälfte der Messdaten  $N$  verwendet, weshalb sich letztlich eine geringere erreichbare Varianz für das *SWF* ergibt. Bei äquidistanter Abtastung entspricht die Differenz der ursprünglichen Messkurve mit ihrer verschobenen Kopie einer Subtraktion jedes einzelnen Messwertes mit seinem am Waferrand um  $180^\circ$  gegenüberliegend aufgenommenen Messwert. Dadurch wird nicht nur der Signal Rausch Abstand verbessert, sondern auch ein Dither Effekt erzeugt, welcher durch zufälliges Messrauschen die inneren statistischen Abhängigkeiten zwischen benachbarten Messwerten reduziert. Abbildung 5.9 zeigt die Varianzschätzungen der verschiedenen Verfahren anhand eines *Notchwafers*. In diesem Fall werden für das *SWF* nur rund 2 Prozent der Messdaten verworfen, für die beiden anderen Verfahren nur halb so viele. Im Falle des *SWF* erreicht die geschätzte Varianz beinahe die theoretisch vorhergesagte Grenze. Die kleinen Unterschiede zwischen *SWF*, *CF* und *FWF* sind praktisch zu vernachlässigen.

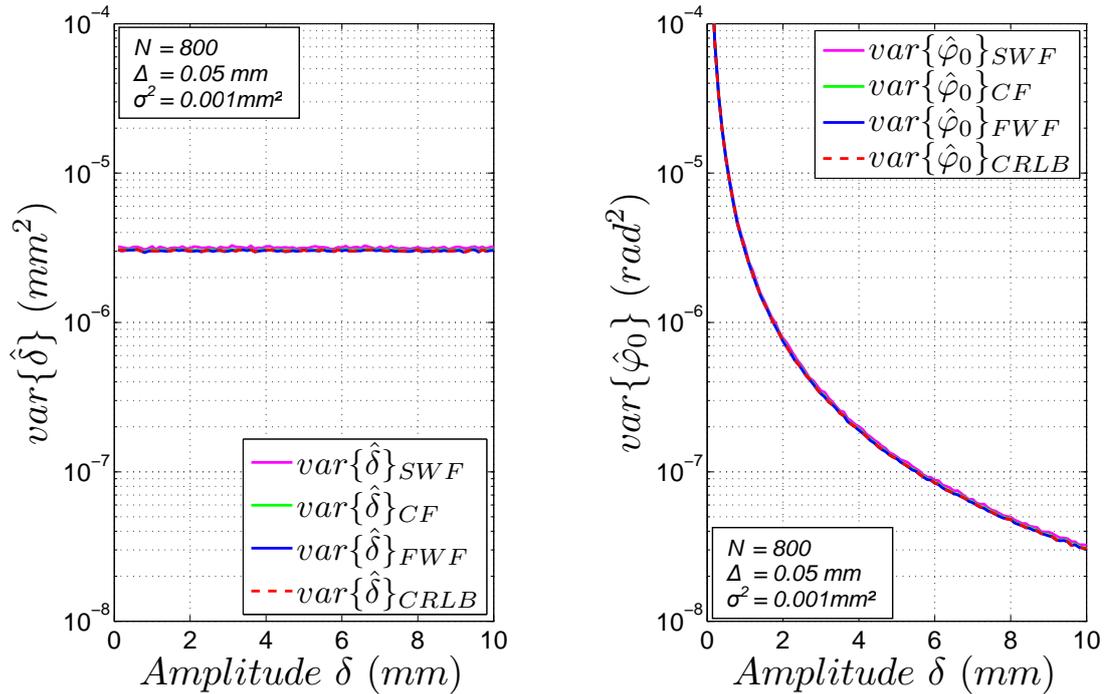


Abbildung 5.9:  $\text{var}\{\hat{\delta}\}$  und  $\text{var}\{\hat{\varphi}_0\}$  für SWF, CF und FWF für Notchwafer

Es soll an dieser Stelle festgehalten werden, dass auf Grund der fehlerquadratminimierenden Eigenschaften aller Verfahren, jedes einzelne eine optimale Lösung erreicht, sodass alle Verfahren die für sie geltenden unteren Grenzen auch annähernd erreichen. Berücksichtigt man bei der Berechnung der *Cramér-Rao Lower Bound* für die Varianzen der Parameter des CF und FWF den ursprünglichen Signal Rausch Abstand, so stellt man fest, dass sich beide Verfahren, ähnlich dem SWF, ihrer theoretischen Grenze annähern. Nun weiß man aus Erfahrung, dass das Erreichen theoretischer Grenzen schwierig ist und spätestens in der Praxis scheitert. Dennoch werden hier nahezu perfekte Ergebnisse präsentiert. Um mit den LMS Verfahren die Qualität von *Maximum Likelihood* Schätzern erreichen zu können, wurden zu Beginn notwendige Forderungen an die Rauschsignale gestellt. Da in den Simulationen genau diese Bedingungen erfüllt wurden, entsprechen die Ergebnisse nun auch den Erwartungen. Abbildung 5.10 zeigt Signal Rausch Abstände für verschiedene Schätzer bei unterschiedlichen Auflösungen des Quantisierers. Für Dezentrierungen unter  $0,1$  mm sind die Signal Rausch Abstände

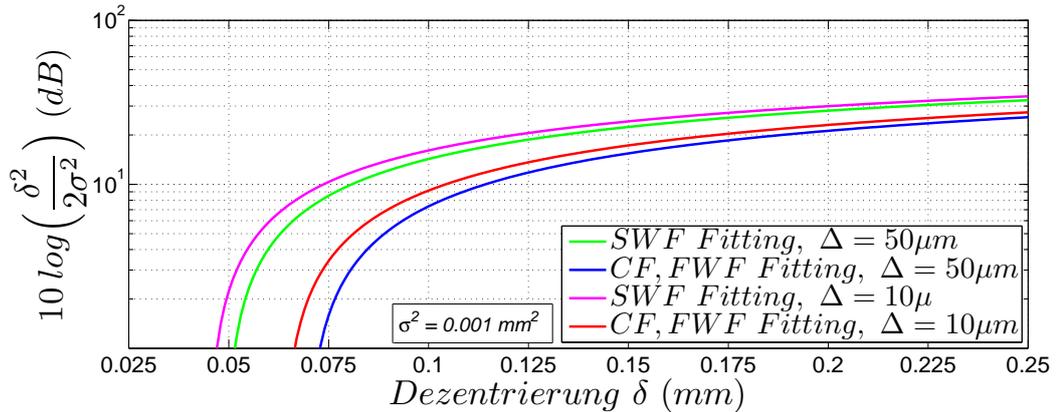


Abbildung 5.10: SNR der Messdaten für verschiedene Schätzer

allgemein schlecht, was verdeutlicht, dass brauchbare Parameterschätzungen von *Waferscans* mit größeren Dezentrierungen besser durchgeführt werden können. Die Schätzung der Varianz von  $\hat{\varphi}_0$  unterliegt mit  $1/\delta^2$  zudem einer starken Abhängigkeit zur Dezentrierung und zeigt, dass der Nullphasenwinkel umso genauer geschätzt werden kann, je höher die Dezentrierung ist.

## 5.5 Verarbeitung realer Messdaten

Mit Hilfe der zuvor erarbeiteten Werkzeuge werden nun die gesuchten Parameter als auch deren Varianzen anhand wahrer Messdaten bestimmt. Abbildung 5.11 zeigt einen *Waferscan* eines 6 Zoll *Wafers* mit *Flat*, welcher mit einem im Vorfeld dieser Arbeit entwickelten *2D-Waferscanner* [1] aufgenommen wurde. Die Annahmen für die Rauschsignale bei allen zuvor durchgeführten Simulationen erweisen sich für praktische Messungen als durchaus gerechtfertigt, was auf Grund der Qualität der Messkurve zu erkennen ist. Bis auf die Quantisierungsfehler sind kaum Messwertausreißer festzustellen. Der *Waferscan* entstand durch äquidistante Abtastung mit  $N = 800$  Samples und einer Auflösung von  $\Delta = 50 \mu\text{m}$ . Tabelle 5.1 listet die Ergebnisse des *SWF*, *CF* und *FWF* auf. Tabelle 5.2 listet die Ergebnisse des *SWF*, *CF* und *FWF* anhand des 6 Zoll *Waferscans* auf, für welche die Toleranzangabe der Schätzwerte mit  $\pm 3\sigma$  angegeben ist. Laut [27] ist jedoch die Schätzung der Phase einer Sinus- bzw. Kosinuskurve immer verzerrt, unabhän-

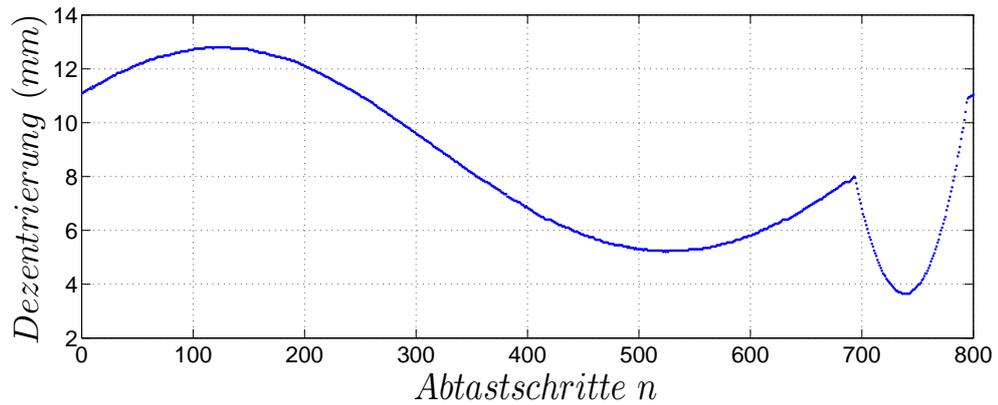


Abbildung 5.11: *Waferscan* eines 6 Zoll *Wafers* mit *Flat*

	$\hat{\sigma}^2$ mm <sup>2</sup>	$\hat{\delta}$ mm	$\hat{\sigma}(\hat{\delta})$ mm <sup>2</sup>	$\hat{\varphi}_0$ [rad]	$\hat{\sigma}(\hat{\varphi}_0)$ [rad]	N
<i>SWF</i>	$2,691 \cdot 10^{-4}$	3,7828	$1,386 \cdot 10^{-3}$	-0,9695	$9,689 \cdot 10^{-5}$	280
<i>CF</i>	$7,022 \cdot 10^{-4}$	3,7821	$1,427 \cdot 10^{-3}$	-0,9684	$9,974 \cdot 10^{-5}$	680
<i>FWF</i>	$1,022 \cdot 10^{-3}$	3,7820	$1,598 \cdot 10^{-3}$	-0,9686	$1,117 \cdot 10^{-4}$	800

Tabelle 5.1: Geschätzte Varianzen für  $\hat{\delta}$  und  $\hat{\varphi}_0$  des 6 Zoll *Waferscans*

	$\hat{\delta}$ [mm]	$\hat{\varphi}_0$ [rad]
<i>SWF</i>	$3,7828 \pm 4,2 \cdot 10^{-3}$	$-0,9695 \pm 290 \cdot 10^{-6}$
<i>CF</i>	$3,7821 \pm 4,3 \cdot 10^{-3}$	$-0,9684 \pm 299 \cdot 10^{-6}$
<i>FWF</i>	$3,7820 \pm 4,8 \cdot 10^{-3}$	$-0,9686 \pm 335 \cdot 10^{-6}$

Tabelle 5.2: Schätzwerte für  $\hat{\delta}$  und  $\hat{\varphi}_0$  des 6 Zoll *Waferscans*

gig von der Qualität der geschätzten Rauschleistung, für welche ein unverzerrter Schätzer angegeben werden konnte.

## 6 *Flat und Notch Fitting*

Dieses Kapitel widmet sich dem *Flat Fitting* (*FF*) und *Notch Fitting* (*NF*), wobei vorausgesetzt wird, dass sich der Wafermittelpunkt bereits im Zentrum des *Chucks* befindet. Wie in 3.3.1 beschrieben, kann mittels Kreuzkorrelation zwischen Messdaten und Verlauf der Randmarkierung der Winkel  $\gamma$  ermittelt werden. Dieser Vorgang ist nicht nur für das Auffinden von *Flat* oder *Notch* von Bedeutung, sondern dient auch der Selektion jener Messwerte, welche für das *SWF* oder *CF* herangezogen werden dürfen. Nachteilig an der Kreuzkorrelation ist die Genauigkeit der Bestimmung des Winkels der Randmarkierung  $\gamma$ , welche bestenfalls die Größenordnung der halben Winkelschrittweite erreicht. Will man den Winkel für die Position von *Flat* oder *Notch* beispielsweise auf eine Toleranz von  $\pm 0,1^\circ$  bestimmen, so bräuchte man mindestens  $3600/2$  Abtastwerte pro Umdrehung. Der Rechenaufwand wäre enorm, weil die gesamten Messdaten verarbeitet werden müssten, um anschließend die Position des Maximalwertes im Ergebnisvektor als Mittelpunkt für den *Flat* oder den *Notch* identifizieren zu können (siehe Abbildung 3.5). Sowohl für die Auswahl der Messwerte für das *SWF* oder *CF* als auch für das Bestimmen des Winkels für *Flat* oder *Notch* wird vorgeschlagen, eine diskrete Ableitung der Messdaten durchzuführen.

### 6.1 Diskrete Ableitung von *Waferscans*

Der Vorteil dieser Methode liegt in der einfachen Durchführung laut Gleichung (6.1) und ihrer universellen Einsetzbarkeit, bei welcher die Differenzen zwischen den einzelnen Abtastwerten begleitend zur Messung berechnet werden können.

$$\frac{ds_F[n]}{dn} \approx s_F[n] - \frac{s_F[n - \Delta n]}{\Delta n} \quad (6.1)$$

Gleichung (6.2) reduziert die diskrete Ableitung auf die Bildung einfacher Differenzen aufeinander folgender Messwerte.

$$\Delta_n s_F[n] = s_F[n] - s_F[n - 1], \quad \Delta n = 1 \quad (6.2)$$

Jede Differenz, welche über einem maximal zu erwartenden Grenzwert liegt, weist darauf hin, dass die dahinter stehenden Messwerte nicht für das *SWF* oder *CF* verwendet werden dürfen. Dabei ist es irrelevant wodurch erhöhte Differenzen verursacht werden, sodass auch eventuelle Kantenausbrüche bei *Wafers* oder Messwertausreißer von der weiteren Datenverarbeitung ausgeschlossen werden können. Negative Differenzen bestimmter Höhe markieren den Beginn einer Unregelmä-

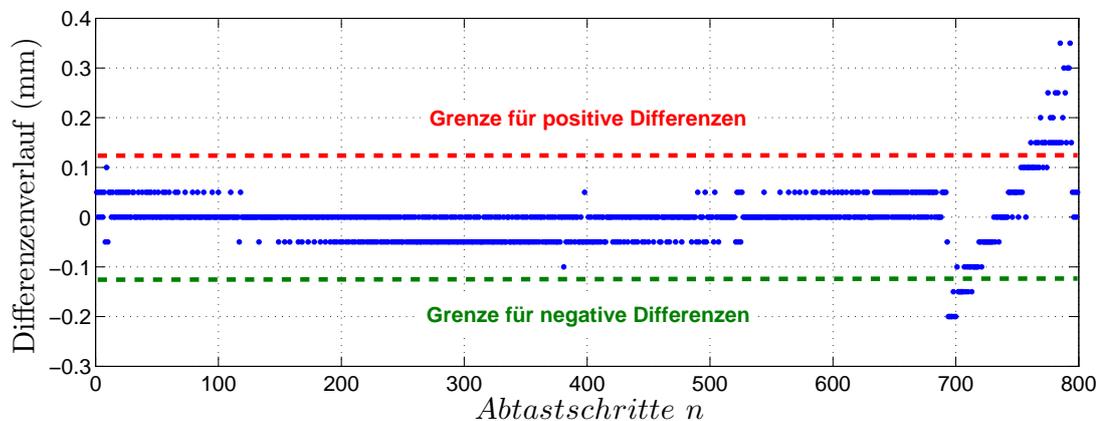


Abbildung 6.1: Differenzenverlauf des *Waferscans* aus Abbildung 5.11

Bigkeit im *Waferscan*, positive Differenzen definieren ihr Ende, wodurch auch die Länge einer Abweichung von der Kreisform des *Wafers* festgestellt werden kann. Für *Wafer* mit *Flat* ist der Primärflat durch die längste Flachstelle des *Wafers* definiert. Der *Notch* hingegen besitzt unabhängig vom *Wafer*durchmesser immer dieselbe Form, weshalb seine Länge bezüglich der Winkelschrittweite bekannt und ebenfalls einfach zu ermitteln ist. Abbildung 6.1 zeigt den Differenzenverlauf des realen *Waferscans* aus dem vorhergehenden Kapitel. Durch Aufsuchen mehrerer aufeinander folgender Differenzen, welche unter der Grenze für die minimale negative Differenz liegen, kann der Start der Randmarkierung erkannt werden. Mehrere aufeinander folgende positive Differenzen kennzeichnen ihr Ende. Durch Mittelung von Start- und Endwert kann zudem auch einfach ein Startwert für

den Flatwinkel  $\gamma$  im Falle des *FWF* ermittelt werden, sodass auf die Kreuzkorrelation verzichtet werden kann. Für *Waferscans* von bereits zentrierten *Wafern* können mit dieser Methode zur Bestimmung des Winkels von *Flat* oder *Notch* zudem dieselben Genauigkeiten erreicht werden, wie mit Hilfe der Kreuzkorrelation, weshalb diese Methode in der Praxis vor allem dort Verwendung findet, wo die mechanische Winkelauflösung der Rotationseinheit gering ist und eine genauere Positionierung des Flatwinkels ohnehin nicht umgesetzt werden kann.

## 6.2 Parameterschätzung für *Flat* und *Notch*

Genauere Schätzungen des Flatwinkels, wie sie mit dem *FWF* erreicht werden können, sollen nun auch für bereits zentrierte *Wafer* durchgeführt werden. In Gleichung (6.3) ist die vereinfachte Beschreibung eines *Flats* angegeben, welche bei zentrierten *Wafern* nur mehr vom Waferdurchmesser  $R$  und dem Flatwinkel  $\gamma$  abhängt.

$$\bar{r}_{F^*}\left(\frac{2\pi}{N}n\right) = \frac{R}{\cos\left(\frac{2\pi}{N}n + \gamma\right)} \quad (6.3)$$

$$\bar{r}_F\left(\frac{2\pi}{N}n\right) := \begin{cases} \bar{r}_{F^*}\left(\frac{2\pi}{N}n\right), & \bar{r}_{F^*}\left(\frac{2\pi}{N}n\right) \leq R \\ R, & \text{sonst} \end{cases} \quad (6.4)$$

Die vollständige Definition des *Waferscans* ohne Dezentrierung ist in Gleichung (6.4) angegeben. Das zugehörige Messmodell in Gleichung (6.5) berücksichtigt in Anlehnung an (3.3) wieder den *Scanneroffset*  $R_0$ , das AWGN-Rauschen  $w[n]$  und das Quantisierungsrauschen  $q[n]$ .

$$\bar{s}_F[n] = \bar{r}_F[n] - R_0 + w[n] + q[n] \quad (6.5)$$

Abbildung (6.2) zeigt den Verlauf eines zentrierten *Wafers* laut Messmodell aus Gleichung (6.5) mit  $R = 75$  mm,  $R_0 = 66$  mm und  $\gamma = \pi/2$ . Auf Basis dieses Messmodells sind zur Schätzung des Flatwinkels bzw. des Notchwinkels  $\gamma$  also auch der Waferradius  $R$  und der *Scanneroffset*  $R_0$  zu schätzen. Da mit den Methoden *SWF* und *CF* Waferradius und *Scanneroffset* nicht ermittelt werden müssen, ist auch für das Schätzen des Flatwinkels bzw. Notchwinkels eine Methode gefragt, welche von diesen Parametern unabhängig ist. Eine mögliche Lösung kann

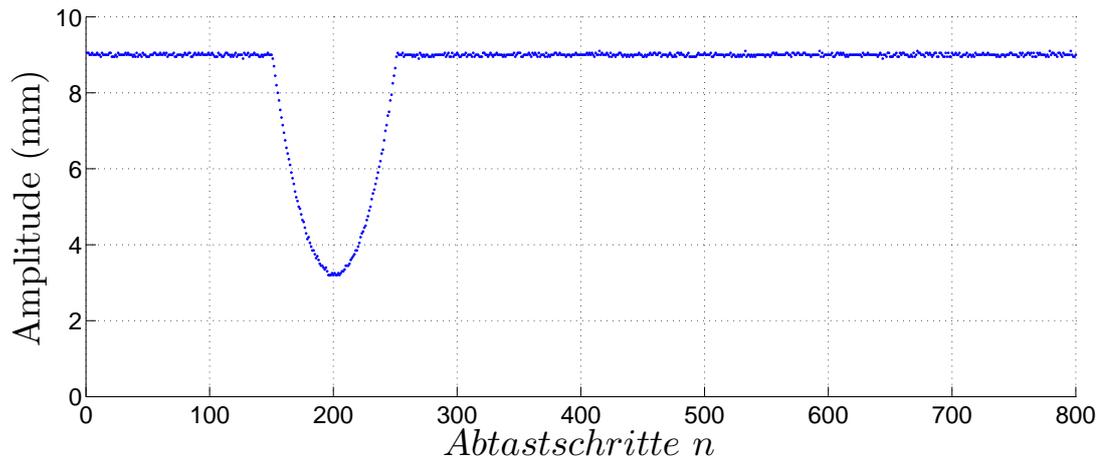


Abbildung 6.2: Flatkurve eines zentrierten *Wafers*

durch das Einpassen einer quadratischen Funktion in *Flat* oder *Notch* erreicht werden. Anschließend ergibt die Auswertung des Scheitelpunktes den Winkel  $\gamma$ .

### 6.2.1 *FF* und *NF* mittels quadratischer Interpolation

Mit Hilfe einer herkömmlichen quadratischen Funktion, welche wiederum mit einem *Least Mean Square* Algorithmus (*LMS*) in die Flatkurve oder den Notchverlauf eingepasst werden soll, kann der Winkel  $\gamma$  vor allem unabhängig vom *Scanneroffset*  $R_0$  ermittelt werden. Gleichung (6.6) definiert das zur Interpolation der Messkurven verwendete Modell, wobei  $n_A$  den Anfang der Randmarkierung und  $n_E$  das Ende derselben markiert.

$$s[n] = an^2 + bn + c + w[n] + q[n], \quad n_A < n < n_E \quad (6.6)$$

Der Scheitelpunkt der quadratischen Kurve ist nun äquivalent zum Mittelpunkt der Randmarkierung, welcher den Winkel  $\gamma$  definiert. Durch Null setzen der Ableitung der quadratischen Funktion wird auf einfache Weise der Scheitelpunkt

ermittelt.

$$\frac{ds[n]}{dn} = 2an + b \stackrel{!}{=} 0 \quad (6.7)$$

$$n_\gamma = -\frac{b}{2a} \quad (6.8)$$

Abbildung (6.3) zeigt wie diese Interpolation mittels kleinster Fehlerquadrat-schätzung durchgeführt werden kann.

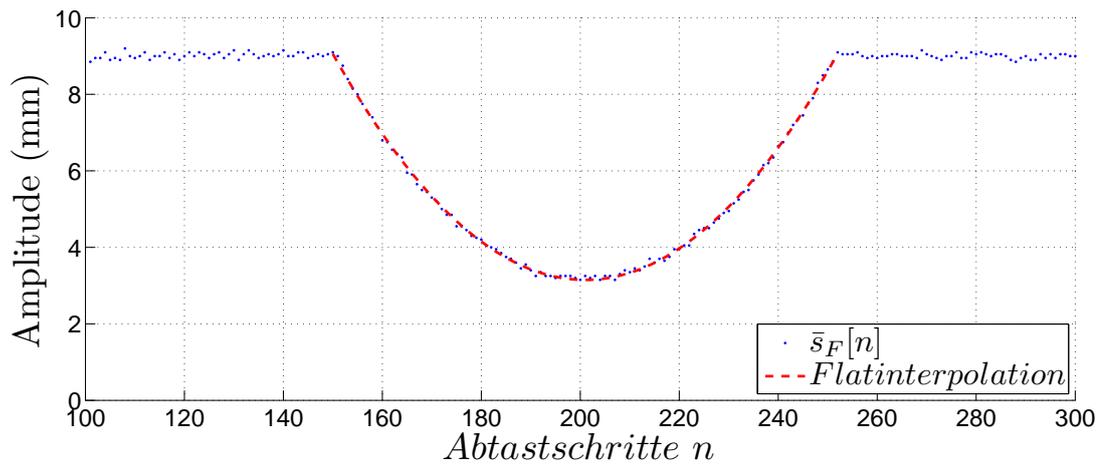


Abbildung 6.3: Interpolation des Flatbereiches aus Abbildung 6.2

### 6.2.2 LMS Schätzer für quadratische Flatinterpolation

Für die Interpolation der Randmarkierung durch eine quadratische Funktion wird die in Gleichung (6.6) definierte Funktion herangezogen. Ähnlich dem *SWF* kann auch für die Interpolation der Randmarkierung ein lineares Gleichungssystem angegeben werden. Mit  $\bar{s}_F[n] = [s_A, s_{A+1}, \dots, s_E]^T$  wir der Messdatenvektor und mit  $[n_A, n_{A+1}, \dots, n_E]^T$  der Vektor mit den zugehörigen Winkelschritten angegeben. Der Parametervektor lautet  $\underline{\theta}_q = [a, b, c]^T$ . Das lineare Gleichungssystem hat dabei dieselbe Struktur wie beim *SWF*, sodass die  $N_F \times 3$  Matrix  $\mathbf{K}$  laut Definition (6.9) aufgestellt wird, wobei  $N_F$  für die Länge der Randmarkierung im

Messdatenvektor steht.

$$\underline{\mathbf{K}} = \begin{pmatrix} n_A^2 & n_A & 1 \\ n_{A+1}^2 & n_{A+1} & 1 \\ \vdots & \vdots & \vdots \\ n_E^2 & n_E & 1 \end{pmatrix} \quad (6.9)$$

Der in Gleichung (6.10) angegebene analytische *LMS* Schätzer weist demnach auch dieselbe Struktur wie der des *SWF* auf.

$$\hat{\underline{\theta}}_q = (\underline{\mathbf{K}}^T \underline{\mathbf{K}})^{-1} (\underline{\mathbf{K}}^T \underline{\bar{s}}_F) \quad (6.10)$$

Mit Hilfe des geschätzten Parametervektors  $\hat{\underline{\theta}}_q = [a, b, c]^T$  ist es nun möglich den *Flat*, wie in Abbildung 6.3 gezeigt, quadratisch zu interpolieren. Die Summe der Fehlerquadrate zwischen gemessenem Flatverlauf und der interpolierten Kurve kann wiederum dafür herangezogen werden, um die Rauschleistung des Messsignals zu schätzen. In Hinblick auf die Berechnung des Scheitelpunkts des *FF*, laut Gleichung (6.8), ist die Schätzung des Parameters  $c$  jedoch nicht erforderlich. Dementsprechend reduziert sich die Matrix  $\underline{\mathbf{K}}$  zur  $N_F \times 2$  Matrix  $\underline{\mathbf{K}}^*$  laut (6.11), der Parametervektor lautet  $\underline{\theta}_q^* = [a, b]^T$ . Eine Vollständige Rekonstruktion des *Flats* auf Basis der quadratischen Funktion ist nun jedoch nicht mehr möglich. Ebenso kann die Rauschleistung des Messsignals auch nicht mehr durch die Differenz zwischen Messdaten und deren Interpolation geschätzt werden. Da beim *SWF* und *CF* jedoch bereits eine Schätzung der Rauschleistung durchgeführt wurde, darf diese auch für das *FF* herangezogen werden.

$$\underline{\mathbf{K}}^* = \begin{pmatrix} n_A^2 & n_A \\ n_{A+1}^2 & n_{A+1} \\ \vdots & \vdots \\ n_E^2 & n_E \end{pmatrix} \quad (6.11)$$

Der Vollständigkeit halber wird der reduzierte analytische Schätzer für die quadratische Interpolation in Gleichung (6.12) angegeben.

$$\hat{\underline{\theta}}_q^* = (\underline{\mathbf{K}}^{*T} \underline{\mathbf{K}}^*)^{-1} (\underline{\mathbf{K}}^{*T} \underline{\bar{s}}_F) \quad (6.12)$$

### 6.2.3 CRLB für $\text{var} \{\hat{\gamma}\}$

Wie schon bei allen vorherigen Schätzungen will auch für die Bestimmung des Flatwinkels die Varianz geschätzt werden. Dafür werden die CRLBs der Parameter  $a$  und  $b$  bestimmt, welche zur Berechnung des Scheitelwertes dienen [14]. In einem zweiten Schritt ist auf Basis dieser Varianzen die Varianz des Scheitelwertes zu berechnen.

$$\begin{aligned} I(\underline{\theta})_{11} &= \frac{1}{\sigma_0^2} \sum_{n=0}^{N-1} n^4 = \frac{1}{\sigma_0^2} \frac{N(N-1)(2N-1)(3N^2-3N-1)}{30} \\ I(\underline{\theta})_{12} &= I(\underline{\theta})_{21} = \frac{1}{\sigma_0^2} \sum_{n=0}^{N-1} n^3 = \frac{1}{\sigma_0^2} \left( \frac{N(N-1)}{2} \right)^2 \\ I(\underline{\theta})_{22} &= \frac{1}{\sigma_0^2} \sum_{n=0}^{N-1} n^2 = \frac{1}{\sigma_0^2} \frac{N(N-1)(2N-1)}{6} \end{aligned}$$

Die Matrixinversion der Fisher Information  $I(\underline{\theta})$  ergibt sich, nach deutlichen Vereinfachungen, zu:

$$CRLB: \frac{1}{I(\underline{\theta})} \geq \sigma_0^2 \begin{pmatrix} \frac{1}{N^5} & -\frac{1}{N^4} \\ -\frac{1}{N^4} & \frac{1}{N^3} \end{pmatrix} \quad (6.13)$$

Die Varianzen der geschätzten Parameter ergeben sich somit näherungsweise zu  $\text{var} \{\hat{a}\} = \hat{\sigma}_a^2 \approx \sigma_0^2/N^5$  und zu  $\text{var} \{\hat{b}\} = \hat{\sigma}_b^2 \approx \sigma_0^2/N^3$ , wobei  $\sigma_0^2$  für die im Messsignal enthaltene Rauschleistung steht. Der Scheitelpunkt im Messdatenvektor wird auf Basis der Beziehung (6.8) mit  $n_\gamma = -\hat{b}/2\hat{a}$  berechnet. Die Varianz  $\text{var} \{\hat{n}_\gamma\}$  ergibt sich nun aus der Summe der beiden Varianzen  $\text{var} \{\hat{a}\}$  sowie  $\text{var} \{\hat{b}\}$  und wird laut Beziehung (6.14) berechnet.

$$\text{var} \{\hat{n}_\gamma\} \approx \hat{\sigma}_a^2 \left( \frac{\partial n_\gamma}{\partial \hat{a}} \right)^2 + \hat{\sigma}_b^2 \left( \frac{\partial n_\gamma}{\partial \hat{b}} \right)^2 \approx \hat{\sigma}_a^2 \left( \frac{\hat{b}}{2\hat{a}^2} \right)^2 + \hat{\sigma}_b^2 \left( \frac{1}{2\hat{a}} \right)^2 \quad (6.14)$$

Für die Berechnung der Varianz des geschätzten Winkels  $\hat{\gamma}$  im Bogenmaß, ist Gleichung (6.14) mit dem Faktor  $(2\pi/N)^2$  zu multiplizieren.

$$\text{var} \{\hat{\gamma}\} \approx \text{var} \{\hat{n}_\gamma\} \cdot \left( \frac{2\pi}{N} \right)^2 \quad (6.15)$$

## 7 Dünnpwafer

In einer Vielzahl technischer Anwendungen müssen *Wafer* produktionsbedingt dünn geschliffen werden. Die Folge sind flexible *Wafer*, welche sich durch mechanische Materialspannungen derart verformen, dass sie Krümmungen von mehreren Millimetern erreichen können. Abbildung 7.1 zeigt einen solchen Dünnpwafer mit

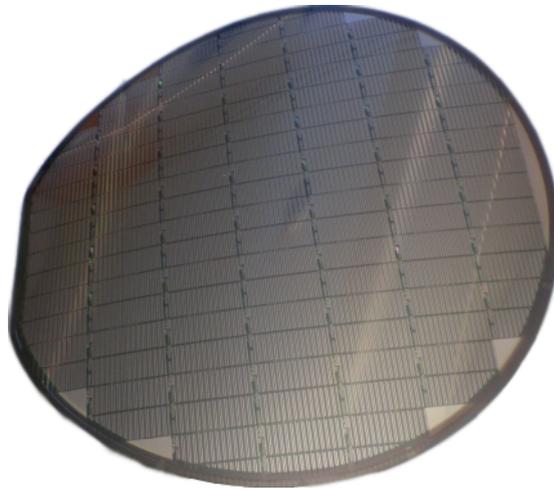
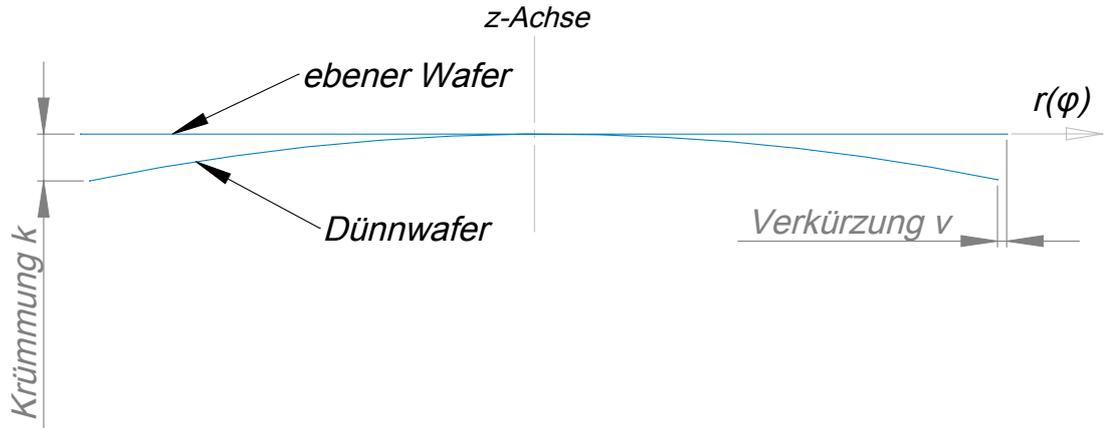


Abbildung 7.1: Dünnpwafer (6 Zoll) mit *Flat*, Krümmung: 6 mm

einer Stärke von nur  $90\ \mu\text{m}$ . Durch die Krümmung verzerrt sich die Form des *Wafers*, sodass die Projektion seines Umfanges in der Ebene  $r(\varphi)$  elliptisch wird. Die Abbildung der Waferkante verkürzt sich und weicht somit von der ursprünglichen Kreisform des *Wafers* ab. Herkömmliche optische Messverfahren können die Krümmung nicht erfassen, wodurch es zu Verzerrungen im *Waferscan* und zu Fehlern in der Parameterschätzung kommt. Abbildung 7.2 definiert die Begriffe Krümmung  $k$  und Verkürzung  $v$ . Mit Hilfe eines *2D-Waferscanners* kann die Kante des *Wafers* zusätzlich in Richtung der z-Achse vermessen werden. Mit einem quadratisch modellierten Verlauf für die Krümmung kann ein komplettes dreidimensionales Modell für gekrümmte Dünnpwafer erstellt werden. Auf Basis der

Abbildung 7.2: Krümmung  $k$  und Verkürzung  $v$  von Dünnwafern

Krümmungsinformation kann auch die auftretende Verkürzung der Waferkante bestimmt werden, wodurch die bereits vorgestellten Algorithmen *SWF*, *CF*, *FWF* sowie des *FF* unverändert angewendet werden können.

## 7.1 Korrektur der Waferkrümmung

Durch die Berechnung der Waferkrümmung und die Entfernungsmessung der Waferkante  $r(\varphi)$  zur  $z$ -Achse wird der Verlauf der Waferkante in Zylinderkoordinaten darstellbar. Der maximale Krümmung des Waferrandes wird nach Abbildung 7.2 mit  $z(\varphi) = \alpha(\varphi) r^2(\varphi)$  quadratisch modelliert. Der Krümmungsparameter  $\alpha(\varphi)$  wird demnach laut Gleichung (7.1) berechnet.

$$\alpha(\varphi) = \frac{z(\varphi)}{r^2(\varphi)} \quad (7.1)$$

Berechnet man für jede Schnittebene die Länge der gekrümmten Waferlinie bezüglich des projizierten verkürzten Waferrandes  $r_p$  nach Gleichung (7.2), erhält man auf einfache Weise jenen Abstand, welchen die Waferkante zur  $z$ -Achse hätte, wäre der *Wafer* nicht gekrümmt.

$$\bar{r}(\varphi) = \int_0^{r_p} \sqrt{1 + \left(\frac{dz(r)}{dr}\right)^2} dr \quad (7.2)$$

Für die Beschreibung der Waferkrümmung  $z(r) = \alpha \cdot r^2$  kann nach Einsetzen in (7.2) eine Korrektur des gemessenen Abstandes zwischen  $z$ -Achse und Waferrand

laut Vorschrift (7.3) berechnet werden.

$$\bar{r} = \frac{r_p}{2} \sqrt{1 + 4 \alpha^2 r_p^2} + \frac{1}{4\alpha} \operatorname{arsinh}(2 \alpha r_p) \quad (7.3)$$

Für diese Berechnung ist das Bekanntsein des *Scanneroffsets*  $R_0$  jedoch zwingend erforderlich, um die Entfernung des gemessenen Waferrandes  $\bar{r}(\varphi)$  in Bezug zur z-Achse bzw. zum Chuckmittelpunkt angeben zu können.

## 7.2 3D-Wafermodelle

Auf Basis der Messdaten des *2D-Waferscanners* lassen sich nun auf einfache Weise 3D-Modelle von gekrümmten *Wafers* darstellen, wie das im Zuge dieser Arbeit entstandene Programm visualisiert (siehe Abbildung 7.3). Diese Visualisierung ermöglicht somit eine einfache Darstellung der Krümmungen von Dünnwafern und stellt diverse Randausbrüche von *Wafers* anschaulich dar, sodass im laufenden Fertigungsprozess alle Waferparameter für die Qualitätskontrolle erfasst, angezeigt und für statistische Zwecke aufgezeichnet werden können.

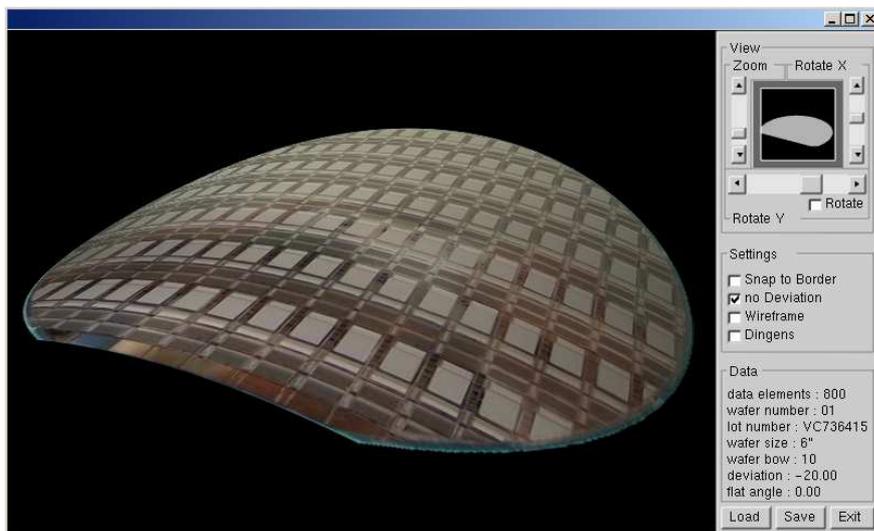


Abbildung 7.3: 3D-Modell eines gekrümmten Dünnwafers [28]

## 8 Schlussfolgerungen

Nach Definition verschiedener Wafertypen wurde eine vollständige mathematische Beschreibung eines dezentrierten *Flatwafers* angegeben. Stellvertretend für alle Wafertypen wurde anhand des *Flatwafers* die Schätzung all seiner Parameter gezeigt. Dabei wurde ein herkömmliches iteratives nichtlineares *Curve Fitting* auf Basis kleinster Fehlerquadrate verwendet. Nachteilig an dieser Vorgehensweise ist jedoch die zwingende Notwendigkeit, das zu verwendende Wafermodell a priori zu kennen. Bevor die Waferscans nicht analysiert sind, ist auch nicht bekannt, welches Modell zur Parameterschätzung herangezogen werden muss.

In einem zweiten Schritt wurde das mathematische Modell für *Wafer* vereinfacht, indem zunächst auf die Modellierung jeglicher Randmarkierungen verzichtet wird. Dieses Modell kann nun für jeden *Wafer* angewendet werden und mit dem *Sine Wave Fitting (SWF)* zudem analytisch geschätzt werden. Die Selektion der Messdaten, welche zur Parameterschätzung herangezogen werden dürfen, erfolgt durch Auswertung der Differenzen zwischen den einzelnen diskreten Messwerten. Diese Methode ist ebenfalls unabhängig vom Wafermodell und der Art der Randmarkierung (*Flat* oder *Notch*) und dient darüber hinaus auch universell zur Detektion von Randausbrüchen oder eventuellen Messwertausreißern. Eine Konsequenz dieser Vorgehensweise ist nun, dass die Schätzung des Wafermittelpunktes ( $\hat{\delta}$ ,  $\hat{\varphi}$ ) und die Schätzung des Winkels  $\hat{\gamma}$  von *Flat* bzw. *Notch* in zwei getrennten Schritten ausgeführt werden muss. Erst nach der Korrektur der Waferdezentrierung kann die Randmarkierung in einem zweiten Scandurchgang vermessen und positioniert werden. Im Sinne einer effizienten Umsetzung des *Prealignments* ist jedoch genau diese Methode in der Praxis von Vorteil, da die Anforderungen an die mechanische Umsetzung so am geringsten ausfallen. Um beispielsweise das *Prealignment* mit den Daten des *Flatwafer Fittings FWF* direkt umsetzen zu können, braucht man eine X-Y-Einheit, welche den *Wafer* positioniert, und einen Rotor, welcher den *Wafer* auf die gewünschte Winkelposition

dreht. Im Falle des universellen zweistufigen Vorgehens wird lediglich der Rotor benötigt, um das *Scannen* des *Wafers* zu realisieren. Die Korrektur der Waferdezentrierung kann von dem Anlagenteil durchgeführt werden, der den *Wafer* dem *Prealignment* zuführt (Roboter, Transfersystem). Ist der *Wafer* zentriert, wird in einem zweiten Schritt nur mehr die Winkelausrichtung korrigiert. Die vorgeschlagenen analytischen Algorithmen *Sine Wave Fitting* (*SWF*) und *Flat Fitting* (*FF*) sind für alle Wafertypen einsetzbar und einfach zu implementieren. Die Untersuchung der Parametervarianzen für  $\hat{\delta}$ ,  $\hat{\varphi}$  und  $\hat{\gamma}$  vervollständigt die Parameterschätzung und stellt die Basis für eine erfolgreiche praktische Umsetzung dieser universellen *Prealignment*-methode dar.

Mit Hilfe der in dieser Arbeit entworfenen Algorithmen, soll in Zukunft die Messdatenverarbeitung im *2D-Waferscanner* integriert werden, sodass ein kompaktes Messgerät zu Verfügung steht, welches nicht nur Messwerte liefert, sondern zusätzlich auch die gewünschten Parameter schätzt. Die praktischen Messungen mit dem *2D-Waferscanner* konzentrierten sich bisher auf die Vermessung von ebenen *Wafers*, weshalb künftig Messungen mit Dünnwafers durchzuführen sind. Dabei wird zu untersuchen sein, mit welcher Qualität die mathematische Korrektur von Dünnwaferscans durchgeführt werden kann.

## Literaturverzeichnis

- [1] T. Jerman, *2D Waferscanner*, Informationstechnik Projekt am IBK der TU Graz, 2008.
- [2] H. T. Kim, H. Yang, and S. Kim, *The Wafer Alignment algorithm regardless of rotational center*, International Conference on Mechatronics, IEEE, 2006.
- [3] C. Huang, Q. Cao, and Z. Fu, *A New Wafer Prealigner Based on Multi-sensor Fusion*, Proceedings of the 7<sup>th</sup> World Congress on Intelligent Control and Automation, 2008.
- [4] H. S. Lee, J. W. Jeon, J.-W. Kim, S.-J. Chung, and J.-E. Byun, *The Development of a 12-inch Wafer Prealigner with Optical Character Recognition*, The 27<sup>th</sup> Annual Conference of the IEEE Industrial Electronics Society, IECON, 2001.
- [5] Q. Dongsheng, Q. Suilong, R. Weibin, S. Yixu, and Z. Yannan, *Design and Experiment of the Wafer Pre-alignment System*, Proceedings of the International Conference on Mechatronics and Automation, IEEE, 2007.
- [6] N. I. Chernov and C. Lesort, *Least Squares Fitting of Circles*, Journal of Mathematical Imaging and Vision, Vol. 23, Issue 3, 2005.
- [7] E. E. Zelinker and I. V. L. Clarkson, *A Statistical Analysis of Least-Squares Circle-Centre Estimation*, Proceedings of the 3<sup>rd</sup> International Symposium on Signal Processing and Information Technology, IEEE, 2003.
- [8] D. Umbach and K. N. Jones, *A Few Methods for Fitting Circles to Data*, Transactions on Instrumentation and Measurement, IEEE, 2005.

- [9] N. C. Stache and H. Zimmer, *Robust Circle Fitting in Industrial Vision for Process Control of Laser Welding*, Proceedings of the 11<sup>th</sup> International Student Conference on Electrical Engineering POSTER, 2007.
- [10] J. W. Head, *Global Distribution of Large Lunar Craters: Implications for Resurfacing and Impactor Populations*, 2010. [Online]. Available: <http://www.sciencemag.org/content/329/5998/1504.abstract>
- [11] M. A. Fischler and R. C. Bolles, *Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography*, Comm. of the ACM, Vol 24, pp 381-395, 1981.
- [12] MATLAB<sup>®</sup>, *version 7.10.0 (R2010a)*. Natick, Massachusetts: The MathWorks Inc., 2010.
- [13] S. B. Vardeman and C.-S. Lee, *Likelihood-based statistical estimation from quantized data*, Transactions on Instrumentation and Measurement, IEEE, 2005.
- [14] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice-Hall, 1993.
- [15] K. Hejn and A. Pacut, *Generalized Model of the Quantization Error - A Unified Approach*, Transactions on Instrumentation and Measurement , IEEE, 1996.
- [16] A. B. Sripad and D. L. Snyder, *A Necessary and Sufficient Condition for Quantization Errors to be Uniform and White*, IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. ASSP-25, 1977.
- [17] V. W. Volovich, *Semiconductor Object Pre-aligning Apparatus*, U.S. Patent 5,238,354, 1993.
- [18] A. Sárhegyi and I. Kollár, *Robust Sine Wave Fitting in ADC Testing*, Instrumentation and Measurement Technology Conference, IMTC '06, IEEE, 2006.

- [19] P. Händel, *Evaluation of a Standardized sine wave fit Algorithm*, Department of Signals, Sensors and Systems, Royal Institute of Technology, SE-100 44 Stockholm, Sweden, 2000.
- [20] *IEEE Standard for digitizing waveform recorders*, IEEE Standard 1057, 1994.
- [21] *IEEE Standard for Terminology and Test Methods for Analog-to-Digital Converters*, IEEE Standard 1241, 2000.
- [22] A. Moschitta and P. Carbone, *Cramér-Rao Lower Bound for Parametric Estimation of Quantized Sinewaves*, Transactions on Instrumentation and Measurement, IEEE, 2007.
- [23] M. F. Wagdy and W.-M. NG, *Validity of Uniform Quantization Error Model for Sinusoidal Signals Without and with Dither*, IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. ASSP-25, 1977.
- [24] I. Kollár and J. J. Blair, *Improved Determination of the Best Fitting Sine Wave in ADC Testing*, Proceedings of the Instrumentation and Measurement Technology Conference, IMTC '04, IEEE, 2004.
- [25] A. Moschitta and P. Carbone, *Statistical Efficiency of Sinewave Fitting when using non-linear Quantizers*, Proceedings of the Instrumentation and Measurement Technology Conference, IMTC '05, IEEE, 2005.
- [26] F. C. Alegria, *Uncertainty of the Estimates of Sine Wave Fitting of Digital Data in the Presence of Additive Noise*, Proceedings of the Instrumentation and Measurement Technology Conference, IMTC '06, IEEE, 2006.
- [27] K. Peters and S. Kay, *Unbiased estimation of the phase of a sinusoid*, Proceedings of the Acoustics, Speech, and Signal Processing Conference, ICASSP '04, IEEE, 2004.
- [28] P. Kogelnig, *Visual-microscan - Programm zur 3D-Darstellung von Wafermodellen*, TU Wien, 2007.

# Abbildungsverzeichnis

2.1	Wafer mit Primär- und Sekundärflat . . . . .	6
2.2	Nicht standardisierte Spezialwafer . . . . .	6
2.3	Wafer ohne Randmarkierung . . . . .	7
2.4	$r_K(\varphi)$ verglichen mit sinusförmigem Kurvenverlauf . . . . .	8
2.5	Wafer mit Flat . . . . .	10
2.6	Wafer mit vertikaler Flatposition . . . . .	11
2.7	Darstellung von $r_F(\varphi)$ . . . . .	12
3.1	Krater der Mondtiefebene Orientale aus [10] . . . . .	14
3.2	Kreisparameter: Radius R und Mittelpunkt M(a,b) . . . . .	15
3.3	Position des Messgerätes . . . . .	17
3.4	Waferscan mit Flat . . . . .	19
3.5	Kreuzkorrelation aus Waferscan und Flatbeschreibung $r_{F_a}[n]$ . . . . .	19
3.6	Flatwafer Fitting (FWF) mit $\sigma_0^2 = 0.05 \text{ mm}^2$ . . . . .	20
3.7	Circle Fitting (CF) mit $\sigma_0^2 = 0.05 \text{ mm}^2$ . . . . .	20
3.8	Histogramm von $M = 1000$ Parameterschätzungen für $R_0$ . . . . .	21
3.9	Varianzschätzung bei CF und FWF . . . . .	22
3.10	MATLAB Script für das Flatwafer Fitting (FWF) . . . . .	28
3.11	MATLAB Implementation der Funktion $r_F$ . . . . .	29
3.12	Häufigkeitsverteilungen aller FWF Parameter . . . . .	30
4.1	Ideal quantisiertes Sinussignal (Amplitude: $2.5 \Delta$ , $\Delta=1$ ) . . . . .	39
4.2	Histogramme für $\Delta_1 = 1$ und $\Delta_2 = 0.01$ . . . . .	39
4.3	Verwerfen von Samples bei großen Quantisierungsstufen . . . . .	40
5.1	Verteilung von $\bar{\sigma}_0^2 - \bar{\sigma}^{*2}$ des SWF . . . . .	43
5.2	Verteilung von $\bar{\sigma}_0^2 - \bar{\sigma}^{*2}$ von CF und FWF . . . . .	43

5.3	Normiertes $\hat{\sigma}^{*2}$ von <i>SWF</i> , <i>CF</i> und <i>FWF</i> . . . . .	45
5.4	Korrigierte Rauschleistungsschätzung für das <i>SWF</i> . . . . .	46
5.5	Korrigierte Rauschleistungsschätzung für <i>F</i> und <i>FWF</i> . . . . .	46
5.6	Misslungenes <i>SWF</i> eines <i>Flatwafer Scans</i> . . . . .	47
5.7	Erfolgreiches <i>SWF</i> anhand eines reduzierten Messdatenvektors . .	48
5.8	$var \{ \hat{\delta} \}$ und $var \{ \hat{\varphi}_0 \}$ für <i>SWF</i> , <i>CF</i> und <i>FWF</i> für <i>Flatwafer</i> . . .	48
5.9	$var \{ \hat{\delta} \}$ und $var \{ \hat{\varphi}_0 \}$ für <i>SWF</i> , <i>CF</i> und <i>FWF</i> für <i>Notchwafer</i> . .	50
5.10	<i>SNR</i> der Messdaten für verschiedene Schätzer . . . . .	51
5.11	<i>Waferscan</i> eines 6 Zoll <i>Wafers</i> mit <i>Flat</i> . . . . .	52
6.1	Differenzenverlauf des <i>Waferscans</i> aus Abbildung 5.11 . . . . .	54
6.2	Flatkurve eines zentrierten <i>Wafers</i> . . . . .	56
6.3	Interpolation des Flatbereiches aus Abbildung 6.2 . . . . .	57
7.1	Dünnyafer (6 Zoll) mit <i>Flat</i> , Krümmung: 6 mm . . . . .	60
7.2	Krümmung $k$ und Verkürzung $v$ von Dünnyafern . . . . .	61
7.3	3D-Modell eines gekrümmten Dünnyafers [28] . . . . .	62

# Tabellenverzeichnis

2.1	Waferspezifikationen . . . . .	5
3.1	Quantile der <i>Student t</i> Verteilung . . . . .	26
3.2	Auswertung der <i>Flatwafer</i> Parameterschätzung . . . . .	27
5.1	Geschätzte Varianzen für $\hat{\delta}$ und $\hat{\varphi}_0$ des 6 Zoll <i>Waferscans</i> . . . . .	52
5.2	Schätzwerte für $\hat{\delta}$ und $\hat{\varphi}_0$ des 6 Zoll <i>Waferscans</i> . . . . .	52