

**Entwurf und Charakterisierung des ATLAS Pixel
Front End Chips FED
in strahlenharter DMILL Technologie**

Dissertation
zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von
Mario Ackers
aus
Koblenz

Bonn 2001

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen
Friedrich-Wilhelms-Universität Bonn

1. Referent: Prof. Dr. N. Wermes
2. Referent: Prof. Dr. P. David

Tag der Promotion: 28.06.2001

Inhaltsverzeichnis

1	Einleitung	5
2	LHC und ATLAS	7
2.1	LHC	7
2.2	Der ATLAS Detektor	8
2.2.1	Der innere Detektor	9
2.2.2	Die Kalorimeter	12
2.2.3	Das Myonspektrometer	13
2.2.4	Das Triggersystem	14
3	Die 'Front-End-Familie'	16
4	Halbleiterdetektoren	19
4.1	Signalentstehung in Materie	19
4.1.1	Wechselwirkung geladener Teilchen	19
4.1.2	Wechselwirkungen von Photonen	22
4.2	Funktionsprinzip eines Siliziumdetektors	23
4.2.1	Grundlagen	23
4.2.2	Die Halbleiterdiode als Detektor	24
5	MOSFETs	28
6	Strahlenschäden in Silizium	32
6.1	Schäden am Siliziumsubstrat	32
6.2	Schäden im Siliziumdioxid	34

6.3	Schäden an der Grenzfläche	37
6.4	Auswirkungen auf die Detektoren	37
6.4.1	Erhöhung der Depletionsspannung	37
6.4.2	Leckstromerhöhung	39
6.4.3	Ladungssammlungsineffizienzen	40
6.5	Auswirkungen auf die Elektronik	40
6.5.1	Verschiebung der Schwellenspannung	40
6.5.2	Leckstromerhöhung	42
6.5.3	Reduktion der Ladungsträgerbeweglichkeit	44
6.5.4	Erhöhung des Rauschens	44
6.5.5	SEU und SEL	44
7	Besonderheiten des DMILL Prozesses	46
8	Das Design des FED1	50
8.1	Anforderungen an den FED1	51
8.2	Aufbau und Funktionsweise des FED1	53
8.2.1	Übersicht	53
8.2.2	Die 'Timestamp'-Architektur	56
8.2.3	Das Kommandoregister	57
8.2.4	Das globale Register	58
8.2.5	Der Analogteil	63
8.2.6	Der Pixelkontrollteil	65
8.2.7	Der Ausleseteil	68
8.2.8	Die CEU	69

8.2.9	Die EOC	71
8.2.10	Die RCU	72
8.2.11	Der Resetgenerator	73
9	Messungen am FED1	75
9.1	Der Meßaufbau	75
9.2	Wafermessungen	76
9.2.1	Die Erzeugung von VTH	77
9.2.2	Der digitale Stromverbrauch	77
9.2.3	Das globale Register	78
9.2.4	Das Pixelschieberegister	78
9.2.5	Die Auslekette	80
9.3	Messungen an Einzelchips	82
9.3.1	Messung interner Signale	82
9.3.2	Analoge Messungen ohne Detektor	83
9.3.3	Analoge Messungen mit Detektor	90
9.4	Messungen an modifizierten Einzelchips	93
10	Das Design des FED2	96
10.1	Der FED2D	97
10.2	Der FED2S	97
11	Messungen am FED2	99
11.1	Wafermessungen	99
12	Möglichkeiten und Grenzen einer DSM Technologie bzgl. Strahlenhärte	104

12.1 Das 'Hamming Code' Register	105
12.1.1 Das Funktionsprinzip	105
12.1.2 Die Berechnung des 'Hamming Codes'	106
12.1.3 Bewertung	108
12.2 Das SEU tolerante Register	109
12.2.1 Bewertung	110
13 Zusammenfassung	112
Tabellenverzeichnis	114
Abbildungsverzeichnis	115
Literatur	119

1 Einleitung

Die Frage nach dem Ursprung der Welt, wie wir sie heute kennen, beschäftigt schon seit jeher die Wissenschaftler. Dazu gehören nicht nur die Entdeckung und Vermessung aller Bausteine, aus denen die Materie heute aufgebaut ist, oder deren Wechselwirkungen untereinander, sondern auch, welche Verhältnisse in früheren Zeiten vorlagen. Aus dieser Motivation heraus werden immer leistungsfähigere Teilchenbeschleuniger mit immer höheren Energien gebaut, die inzwischen in Bereiche vordringen, die zeitlich nur noch in der Größenordnung von Femtosekunden nach dem Urknall liegen.

Eine dieser Bestrebungen wird vom europäischen Kernforschungsinstitut CERN (**C**entre **E**uropéen pour la **R**echerche **N**ucléaire) in Genf [1] vorangetrieben in Form des Umbaus des alten LEP (**L**arge **E**lectron **P**ositron **C**ollider) Beschleunigerrings zum neuen LHC (**L**arge **H**adron **C**ollider) [2, 3], der Protonen zur Kollision bringen wird. Mit diesem soll es ermöglicht werden, einige ungeklärte Fragen der modernen Physik zu beantworten. Da wären beispielsweise die Entdeckung und Vermessung des Higgs-Bosons, das für die Herkunft der Masse der Elementarteilchen verantwortlich zeichnet oder die Möglichkeit einer (minimalen) supersymmetrischen Erweiterung des Standardmodells, die zur Existenz von mindestens fünf Higgs-Teilchen führen würde. Bei LEP wurde die untere Massengrenze für das Higgs bis 108 GeV nach oben geschoben, ab dann wird in Zukunft bei LHC bis zu einer Energie von 14 TeV weiter gesucht werden, um evtl. auch bisher unbekannte fundamentale Prozesse zu erforschen.

An den vier Kollisionspunkten bei LHC werden verschiedene Experimente aufgebaut werden, die jeweils für unterschiedliche Anforderungen optimiert werden: ATLAS (**A** **T**oroidal **L**H**C** **A**pparatu**S**) [4, 5] und CMS (**A** **C**ompact **M**uon **S**olenoid) [6] dienen als Mehrzweck pp -Detektoren und sollen das gesamte Forschungspotential des LHC ausschöpfen. LHC-B [7, 8] konzentriert sich ganz auf die Gewinnung neuer Erkenntnisse im Bereich der B -Physik. Schließlich wird die Forschung auf dem Sektor der Physik der Schwerionen, die ebenfalls bei LHC beschleunigt werden können, durch ALICE (**A** **L**H**C** **I**on **C**ollider **E**xperiment) [9, 10] vorangetrieben.

Die bei LHC eingesetzten Detektoren werden Anforderungen leisten bzw. Belastungen ausgesetzt, die zur Zeit an der Grenze des Machbaren liegen. Alle 25 ns gibt es eine pp -Kollision, die wiederum etwa 2000 geladene Teilchen erzeugt, deren Spuren bei ATLAS zu $12 - 15 \mu\text{m}$ genau vermessen werden müssen, wovon allerdings nur ein kleiner Bruchteil von physikalischem Interesse ist. Aufgrund des hohen Teilchenflusses treten somit starke Strahlenbelastungen von etwa $30 - 40 \text{ Mrad}$ in unmittelbarer Nähe des Kollisionpunktes während der geplanten Betriebsdauer von 10 Jahren auf, denen Elektronik und Material ausgesetzt werden. Für die innerste Detektoreinheit bleibt aus diesen Gründen nur die Alternative eines Silizium-Pixel-Detektors, der zum einen wegen seines granularen Aufbaus die erforderliche zweidimensionale Ortsauflösung bei der Trajektorienrekonstruktion liefert und zum anderen die nötige Datenauslesegeschwindigkeit durch die angeschlossene Elektronik ermöglicht. Die Strahlenresistenz des Sensors erreicht man durch Verwendung eines mit Sauerstoff angereicherten n^+ -auf- n -Siliziumdetektors, bei der Ausleselektronik geschieht dies einerseits durch die spezielle Wahl des Herstellungsprozesses, in unserem

Fall dem DMILL-Prozeß der Firma Atmel, ein speziell zur Strahlenhärte ausgelegter Prozeß, andererseits durch besondere Schaltungs- und Layouttechniken.

Gegenstand dieser Arbeit sind das Design des ersten strahlenharten, voll ATLAS-tauglichen Pixelchips FED (**F**ront **E**nd **D**MILL), die Charakterisierung der verwendeten Technologie durch detaillierte elektrische Messungen und die Erwägung der Möglichkeit eines tatsächlichen Einsatzes bei ATLAS.

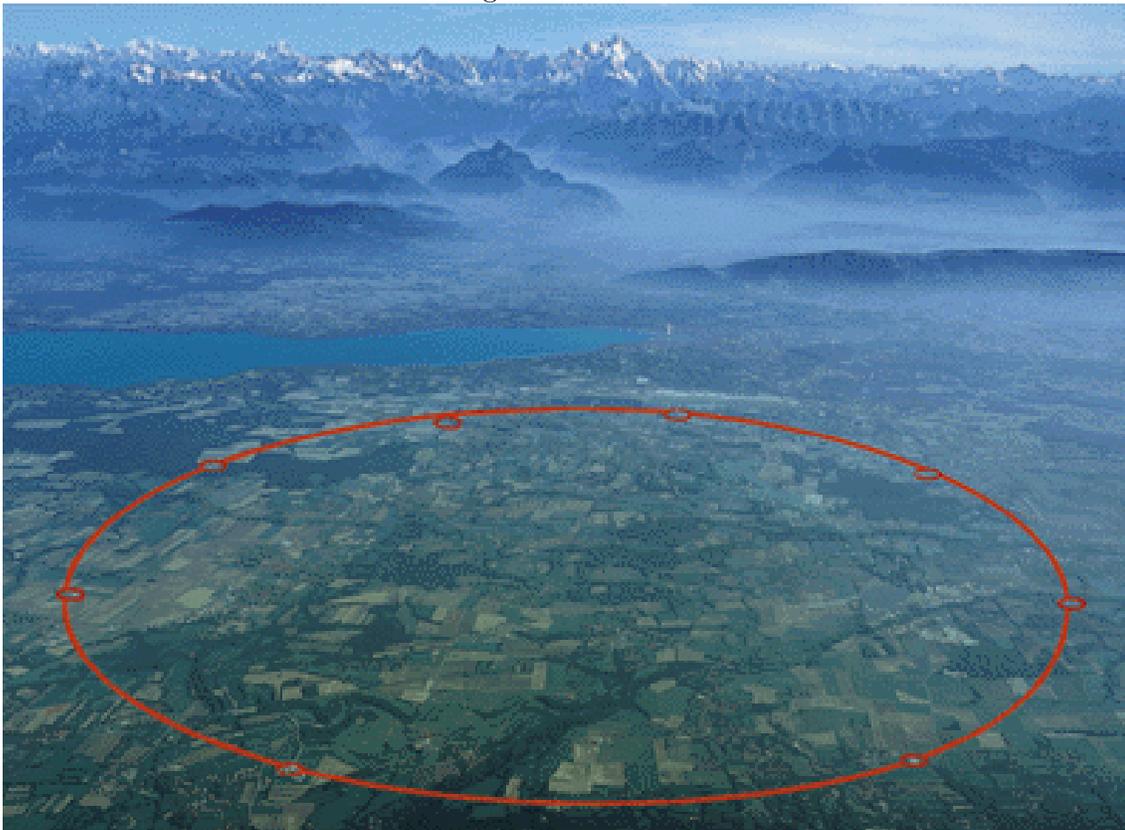
In Kapitel 2 wird ein grober Abriß über den ATLAS-Detektor im allgemeinen und im darauf folgenden Abschnitt eine kurze Übersicht der Historie bereits vorher existierender Auslesechips gegeben. Die wesentlichen Grundlagen bzgl. Halbleiterdetektoren und MOSFETs werden in Kapitel 4 und 5 dargelegt. Kapitel 6 zeigt die Mechanismen und Folgen der Strahlenschädigung für die Siliziumdetektoren und die Ausleseelektronik auf, und Kapitel 7 erläutert die Besonderheiten des DMILL Prozesses. Das eigentliche Design des FED Chips, das sich nach den gestellten Anforderungen richtet, wird ausführlich in Kapitel 8 beschrieben, welches auch auf die Erzielung der Strahlenhärte durch DMILL und den damit verbundenen Herausforderungen bzw. Schwierigkeiten eingeht. Kapitel 9 beinhaltet die zur Charakterisierung des Chips im Hinblick auf eine Massenproduktion in Form von Wafertests durchgeführten Messungen. Ferner werden an dieser Stelle die Tests von einzelnen FED's zur Analyse der aufgetretenen Fehler beschrieben, die hauptsächlich durch die DMILL-Technologie bedingt sind. Kapitel 10 und 11 haben von der Thematik her denselben Inhalt wie Kapitel 8 und 9, beziehen sich jedoch auf eine neue FED-Iteration und decken die Mängel von DMILL eindeutig auf. Zum Abschluß folgt ein Ausblick auf zukunftsweisende DSM (**D**eep **S**ub**M**icron) Technologien, die zu großer Hoffnung für ATLAS Anlaß geben.

2 LHC und ATLAS

2.1 LHC

Der LHC ist ein seit 1994 geplanter pp -Beschleuniger, der in dem bereits existierenden LEP Tunnel installiert und voraussichtlich 2005 fertiggestellt wird. Es handelt sich hierbei um einen Speicherring mit einem Umfang von etwa 27 km und einer Schwerpunktsenergie von 14 TeV , die hoch genug ist, um den für die Higgs-Suche relevanten physikalischen Bereich abzudecken. Bei Inbetriebnahme des Speicherrings wird die Luminosität $10^{33}\text{ cm}^{-2}\text{ s}^{-1}$ betragen und einige Zeit später auf $10^{34}\text{ cm}^{-2}\text{ s}^{-1}$ erhöht werden.

Abbildung 1: Der LHC in Genf

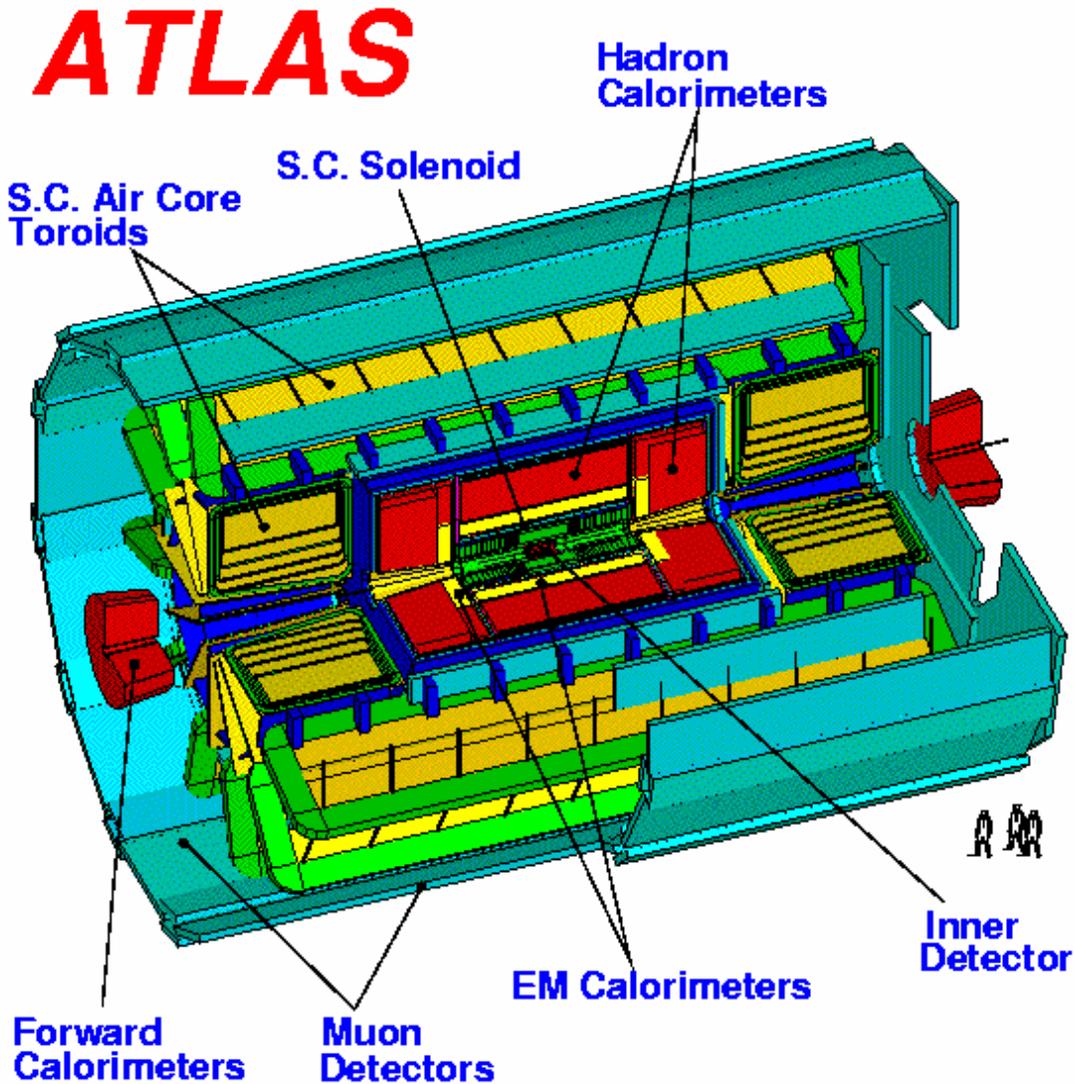


Die Protonen werden zu Paketen von größenordnungsmäßig 10^{11} Teilchen in zwei parallelen Vakuumröhren in entgegengesetzter Richtung beschleunigt und an vier Wechselwirkungspunkten, nämlich dort, wo die Experimente aufgebaut werden, zur Kollision

gebracht. Alle 25 ns wird es zu einer solchen Überkreuzung von Protonenbündeln kommen, die durchschnittlich 25 Wechselwirkungen vollführen und etwa 2000 neue Teilchen erzeugen. Von den vier LHC Detektoren bildet ATLAS mit einer Länge von 42 m , einer Höhe von 22 m und einer Masse von 7000 t das größte Experiment.

2.2 Der ATLAS Detektor

Abbildung 2: Der ATLAS Detektor



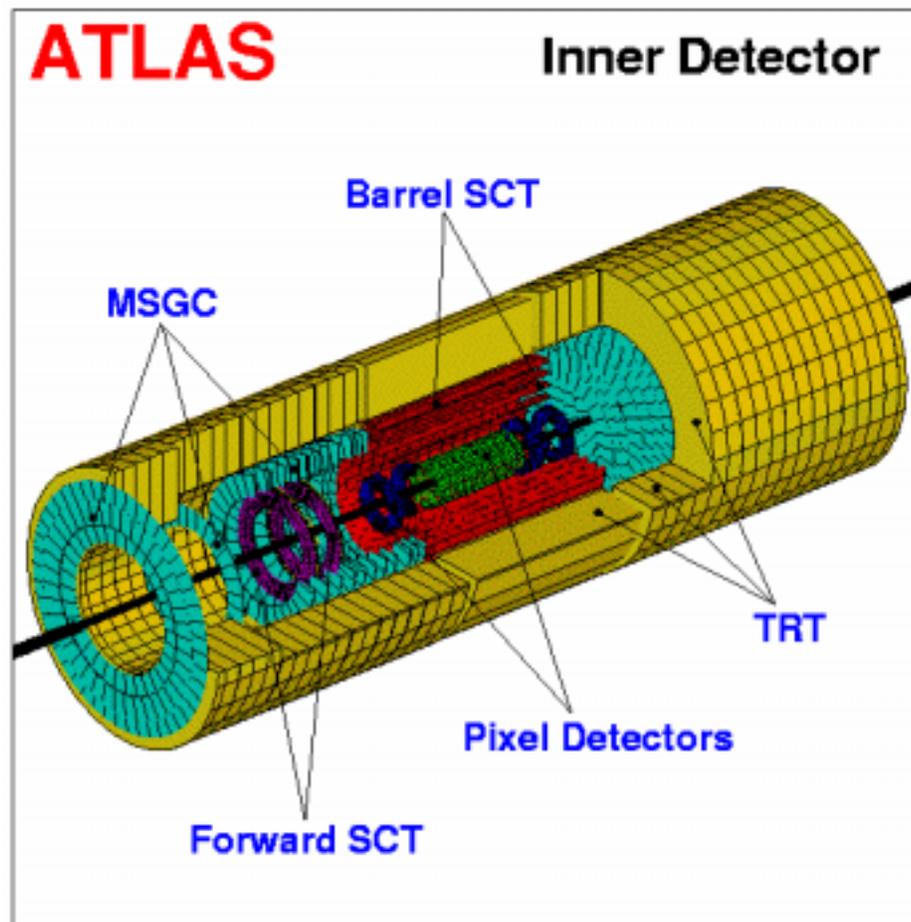
In der Einleitung wurde bereits auf die Motivation zum Bau des ATLAS Detektors eingegangen. Zu nennen sind hier:

- Der experimentelle Nachweis des theoretisch vorhergesagten HIGGS-Bosons und die Ermittlung seiner physikalischen Eigenschaften.
- Die Entdeckung von CP-Verletzungen im B-Kanal während der Phase von LHC mit niedriger Luminosität.
- Die Erforschung der über das Standardmodell hinausgehenden Physik in Form von der Entdeckung supersymmetrischer Teilchen.

ATLAS besteht aus drei, im wesentlichen autonomen, nach dem Zwiebelschalenprinzip angeordneten, Detektorkomponenten, denen ein dreistufiges Triggersystem zur effizienten Datenreduktion angeschlossen ist, welche in den folgenden Abschnitten kurz erläutert werden.

2.2.1 Der innere Detektor

Abbildung 3: Der innere Detektor



Die Spurkammer, die aufgrund ihrer zentralen Lage innerhalb des Gesamtdetektors den Namen 'innerer Detektor' trägt, dient der präzisen Vermessung von Ort bzw. Impuls der erzeugten Teilchen und gibt einen groben Aufschluß über die deponierte Energie. Aufgrund der erwarteten hohen Spurdichten benötigt man zur räumlichen Trennung benachbarter Teilchentrajektorien oder eventueller Sekundärvertizes eine möglichst gute Ortsauflösung mit vielen Spurpunkten in unmittelbarer Nähe des Wechselwirkungspunktes. Als zusätzliche Erschwernis tritt gerade in diesem Bereich eine besonders große Strahlenbelastung einerseits durch, aus der Kollision stammende, hochenergetische Teilchen und andererseits durch, von den Kalorimetern rückgestreute, Neutronen auf.

Aus diesen Gründen wählt man folgenden Aufbau [11, 12]:

- **Der Pixeldetektor**

Drei zylindrische Lagen aus Silizium-Pixeldetektoren umgeben ganz innen den Wechselwirkungspunkt und werden in Vorwärts- bzw. Rückwärtsrichtung von je fünf Scheiben, ebenfalls bestehend aus Pixeldetektoren, abgeschlossen. Die innerste der drei Lagen trägt den Namen B-Lage, weil sie für das sogenannte 'B-Tagging' von großer Bedeutung ist. Sie befindet sich so nah wie möglich am Strahlrohr nur $4,3\text{ cm}$ vom Wechselwirkungspunkt entfernt. Die beiden anderen Lagen heißen Lage 1 bzw. Lage 2 und sind bei Radien von $10,1\text{ cm}$ bzw. $13,2\text{ cm}$ angebracht.

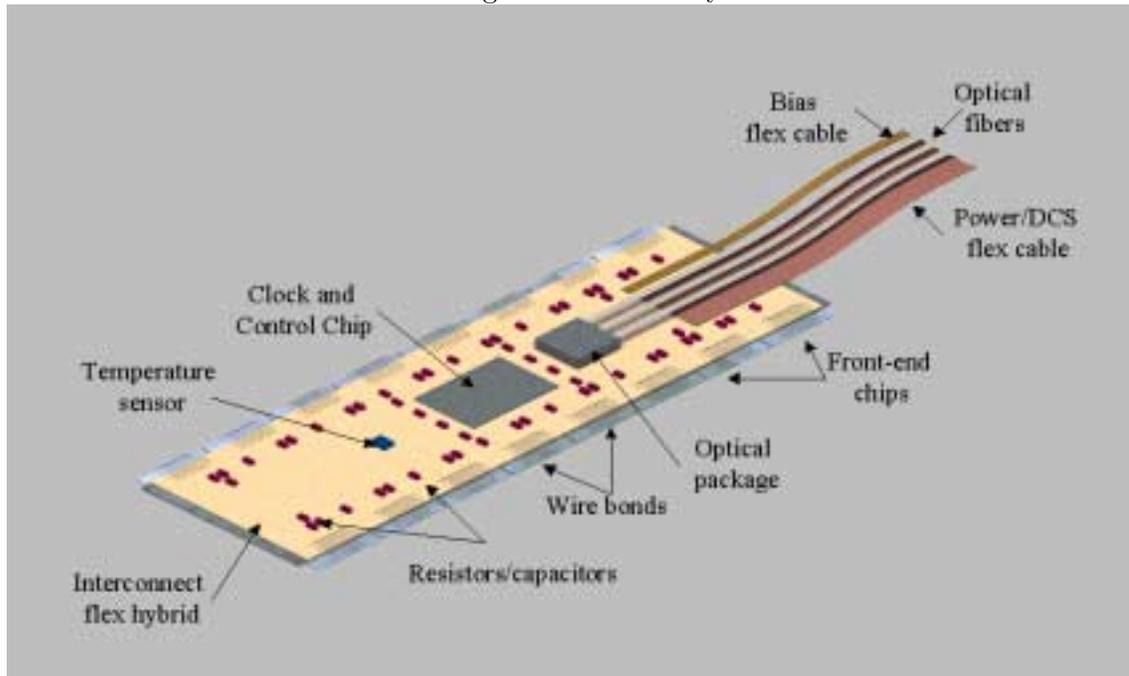
Sowohl die Zylindermäntel als auch die Scheiben bestehen aus insgesamt 2228 identischen Modulen, die jeweils einen Silizium-Pixelsensor und 16 baugleiche **F**ront **E**nd Chips beinhalten, was zu einer Gesamtzahl von $1,4 \times 10^8$ Pixeln führt. Dabei werden Sensorpixel und Elektronikpixel durch Indium- oder Lötzinneügelchen ('Bump Bonds') leitend miteinander verbunden.

Die Koordination und Steuerung der 16 Chips eines Moduls nimmt der MCC (**M**odule **C**ontrol **C**hip) vor, der mittels einer Kaptonfolie ('Flex Hybrid'), die mit Leiterbahnen und passiven elektronischen Komponenten versehen ist, auf den Detektor geklebt wird. Die Verbindung zwischen den Front End Chips und dem 'Flex Hybrid' geschieht mithilfe von winzigen Lötverbindungen ('Wire Bonds'), welche gleichzeitig mangels Stabilität der Drähte auch den Schwachpunkt dieser häufig verwendeten Standardmethode bildet.

Als Alternative soll eventuell in der B-Lage das Konzept des MCM-D (**M**ulti **C**hip **M**odule-**D**eposition) eingesetzt werden, das weitestgehend den Einsatz von 'Wire Bonds' vermeidet. Es handelt sich hierbei um eine Multilagenplatine, die alle notwendigen Kontaktierungen intern schon enthält und mittels 'Bump Bonds' an die Chips angeschlossen wird. Diese Methode befindet sich allerdings noch im Entwicklungsstadium und weist noch eine hohe Defektrate auf.

Die Kommunikation zwischen einzelnen Modulen und der externen Elektronik stellt das sogenannte 'Optopackage' über Glasfaserkabel her. Es enthält eine Laserdiode, den VCSEL (**V**ertical **C**avity **S**urface **E**mitting **D**iode), mit einem dazu passenden Treiberchip, den VDC (**V**CSEL **D**river **C**ircuit), zum Senden der Daten. Der Empfang wird durch eine PIN-Diode und einen Auslesechip, den DORIC (**D**igital **O**ptical **R**eceiver **I**ntegrated **C**ircuit), geleistet.

Abbildung 4: Das 'Flex Hybrid'



Die Stärke der Pixeldetektoren im Vergleich zu Streifendetektoren liegt bei ungefähr vergleichbarer Ortsauflösung darin, daß man einen echt zweidimensionalen Spurpunkt erhält, der durch die angeschlossene Elektronik mit hoher Geschwindigkeit ausgelesen werden kann, was beim Streifendetektor beides nicht möglich ist. Ein weiterer Vorteil wird durch die höhere Strahlenhärte gegeben, die darauf beruht, daß sich die strahlenbedingte Vergrößerung des Detektorleckstroms auf mehr Zellen verteilt.

Mit einer Pixelgröße von $400 \mu m \times 50 \mu m$ erreicht man eine Ortsauflösung von $12 - 15 \mu m$ in $r\phi$ -Richtung und $80 - 100 \mu m$ in z-Richtung.

- **Der Streifendetektor**

Den Pixeldetektor umschließt in analoger Topologie der Silizium-Streifendetektor mit vier Lagen bei Radien von 30, 37, 45 und 52 cm und jeweils zehn Scheiben in Vorwärts- bzw. Rückwärtsrichtung. Seine Ortsauflösung beträgt $20 \mu m$ in $r\phi$ -Richtung und $700 \mu m$ in z-Richtung.

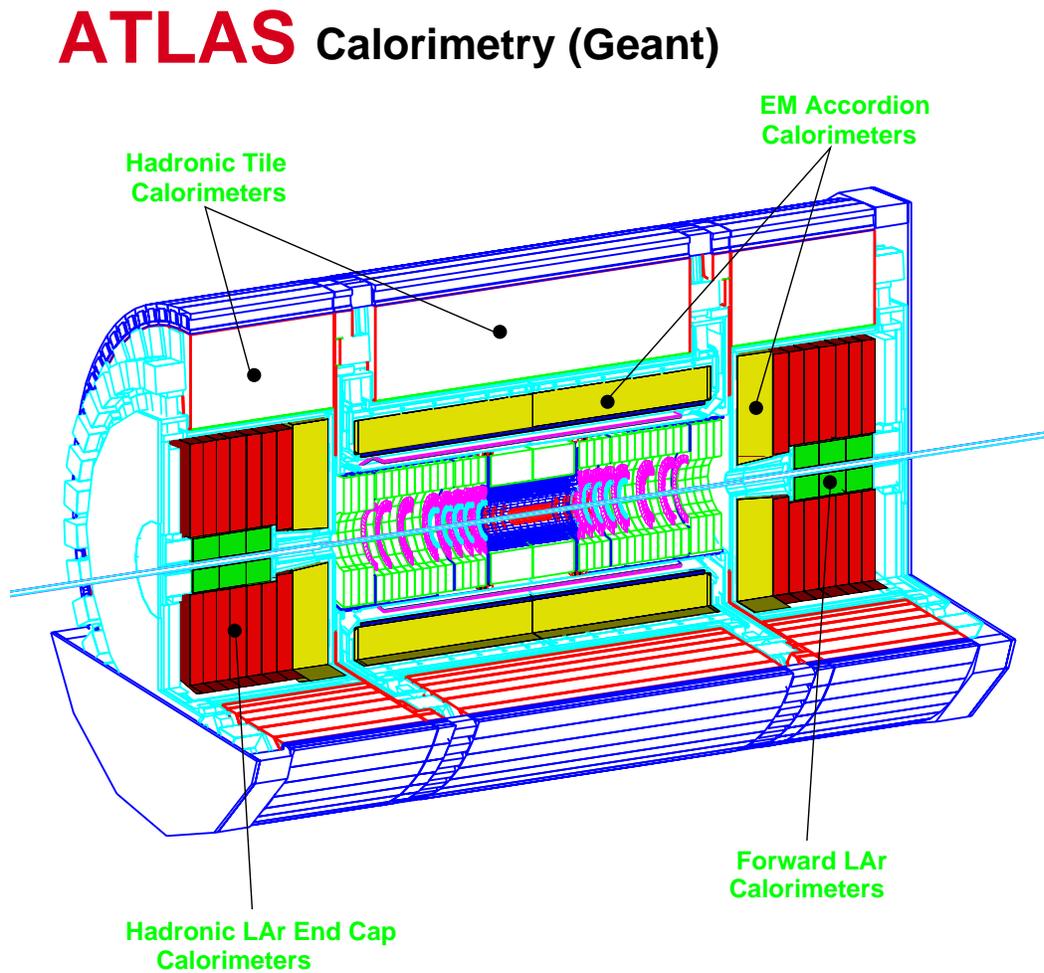
- **Der Übergangsstrahlungsdetektor**

Die äußerste aktive Komponente des inneren Detektors bildet der Übergangsstrahlungsdetektor. Er setzt sich aus 64 konzentrischen Lagen parallel zum Strahlrohr angeordneter Elemente zusammen. Diese bestehen aus Gasdriftkammern mit 4 mm Durchmesser, die in Polypropylenschaum eingebettet sind. Beim Durchgang geladener Teilchen wird dort Übergangsstrahlung erzeugt, die in den Kammern nachgewiesen wird. Man erreicht auf diese Weise etwa 30 Spurpunkte mit einer Ortsauflösung von $130 \mu m$.

Der gesamte Spurdetektor ist zur Impulsbestimmung von einem supraleitenden Solenoiden umgeben, der ein magnetisches Feld von 2 T erzeugt und eine Länge von $6,8\text{ m}$ bzw. einen Radius von $1,15\text{ m}$ erreicht.

2.2.2 Die Kalorimeter

Abbildung 5: Die Kalorimeter



Das Kalorimetersystem [13] bildet die nächste große Detektoreinheit und führt eine präzise Messung der Energie einfallender Teilchen sowie die Bestimmung der fehlenden transversalen Energiekomponente durch. Dies soll auch in Regionen großer Spurdichten ('Jets')

erfolgen, so daß auch hier wiederum sowohl die Strahlenhärte als auch die Auslesegeschwindigkeit von wesentlicher Bedeutung sind.

Das Kalorimeter besteht aus zwei Komponenten:

- **Das elektromagnetische Kalorimeter**

Ziel dieser Einheit ist, neben den oben genannten, die Trennung bzw. Identifikation von Elektronen und Photonen sowie die Separation dieser beiden Teilchen von Hadronen.

Es handelt sich hierbei um ein sogenanntes 'Sampling'-Kalorimeter, das aus zwei alternierenden Schichten aufgebaut ist. Die eine besteht aus einem passiven Absorbermaterial hoher Dichte, Blei im Vorwärts-, Kupfer im zylindrischen Teil, zur Aufschauierung der Teilchen. Die zweite wird durch das eigentliche aktive Nachweismedium gebildet. Aufgrund seiner intrinsischen Strahlenhärte und der schnellen Signalanstiegszeit des Ionisationsstroms von etwa 1 ns fällt hier die Wahl auf flüssiges Argon.

- **Das hadronische Kalorimeter**

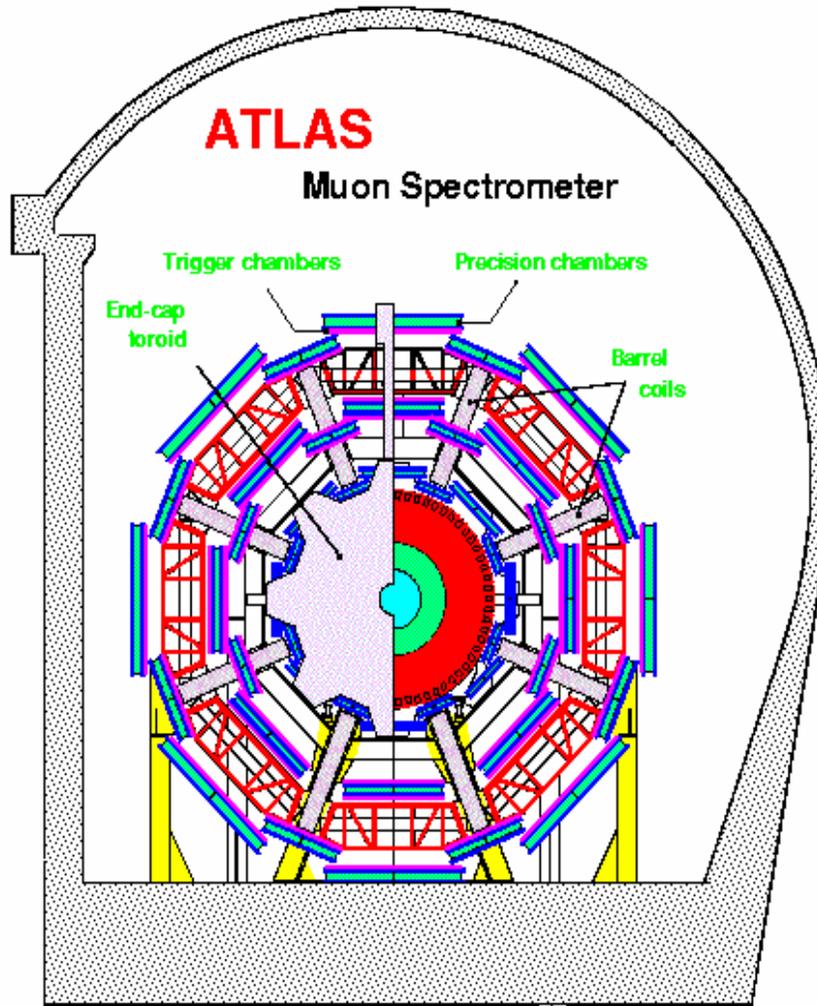
Das hadronische Kalorimeter funktioniert ebenfalls nach der 'Sampling'-Technik und verwendet als Absorber Wolfram, Kupfer und Eisen, wobei letzteres gleichzeitig als Rückführjoch für den Solenoiden dient. Die Entfernung des zylindrischen Teils des Kalorimeters vom Strahlrohr liegt in einem Bereich, der es im Hinblick auf die Strahlenresistenz erlaubt, als aktives Material Plastiksintillatoren zu verwenden.

2.2.3 Das Myonspektrometer

Die Aufgaben des Myonspektrometers [14] liegen in der Identifikation von Myonen, der Bestimmung ihrer Impulse und der Erzeugung eines Triggersignals für die Selektion physikalisch interessanter Ereignisse. Dafür benötigt man zwei verschiedene Komponenten: Erstens Präzisionskammern mit hoher Ortsauflösung, die mithilfe eines toroidalen Magnetfeldes eine genaue Impulsinformation liefern, und zweitens schnelle Triggerkammern, die eine zeitliche Zuordnung der Ereignisse ermöglichen. Beide Teile umschließen den restlichen Detektor in je drei alternierenden Lagen.

Die Präzisionskammern bestehen im zylindrischen Teil aus MDT-Kammern (**M**onitored **D**rift **T**ubes). Dabei handelt es sich um Driftkammern mit 3 cm Durchmesser, die bei geeigneten Betriebsparametern eine Ortsauflösung von $60\text{ }\mu\text{m}$ erreichen. Die Drähte werden in radialer Richtung gespannt, da man zur Ausnutzung des toroidalen Magnetfeldes die z-Komponente besonders genau ermitteln muß. In der Vor- bzw. Rückwärtsrichtung werden neben den MDT-Kammern auch Kathodenstreifenkammern eingesetzt, wobei die jeweils äußerste Lage in 21 m Abstand vom Kollisionspunkt an den Wänden der Experimentierhalle aufgebaut wird, um einen möglichst langen Hebel zu bilden.

Abbildung 6: Das Myonspektrometer



2.2.4 Das Triggersystem

Bereits erwähnt wurde die erste Triggerstufe (Level 1) [15], die in dem Myonsystem und den Kalorimetern erzeugt wird. Bei der Signalerzeugung und der Weiterleitung des Triggers an die anderen Detektorkomponenten vergehen maximal $2,5 \mu s$, währenddessen die Daten in den einzelnen Elementen zwischengespeichert werden müssen. Durch diese Stufe erfolgt eine Reduktion der Daten auf 130 GB/s .

Die zweite Triggerstufe (Level 2) fordert einen Großteil aller Detektordaten an, die sich auf die vom Level 1 selektierten Teilchenspuren beziehen. Nach $1 - 10 \text{ ms}$ wird das zugehörige Triggersignal generiert und an die letzte Stufe (Level 3) weitergeleitet.

Hier werden alle notwendigen Daten aus dem gesamten Detektor zur Ereignisrekonstruktion aufgenommen und mit einer Rate von $10 - 100 \text{ MB/s}$ abgespeichert.

3 Die 'Front-End-Familie'

Abbildung 7: 'Die Front End Familie'

Institute	CPPM + PIB	LBNL
Entstehung		
1996	Bier et Pastis	
1997	MareBo	
1998	FEA/FEC	FEB
1999		FED
2001		FEI

Der in dieser Arbeit vorgestellte FED Chip basiert auf den Erfahrungen, die aus der Produktion einer ganzen Reihe von Prototypen- und Testchips gewonnen wurden, bei deren Entwicklung drei Institute beteiligt waren: Das CPPM (Centre de Physique des Particules de Marseilles) in Frankreich, das LBNL (Lawrence Berkeley National Laboratory) in den USA und das PIB (Physikalisches Institut Bonn). Nachfolgend werden die jüngsten Chips aufgezählt und kurz beschrieben:

- **Der Bier et Pastis**

Im Jahr 1996 wurde in Zusammenarbeit des CPPM und des PIB der 'Bier et Pastis' Pixelchip [16], benannt nach den Nationalgetränken der beiden beteiligten Nationen, entworfen. Dabei handelt es sich um einen Testchip, bestehend aus 12 Spalten mit je 63 Zellen einer Größe von $50 \mu m \times 433,4 \mu m$. Das Design wurde in dem nicht strahlenresistenten $0,8 \mu m$ BiCMOS Prozeß der Firma AMS in Graz hergestellt, in der Hoffnung, daß eine Konversion des Layouts in die strahlenharte DMILL Technologie möglichst einfach vonstatten gehen könnte, was sich jedoch später als Fehleinschätzung herausstellte. Das Funktionsprinzip eines einzelnen Pixels stellt sich wie folgt dar: Die in dem Detektor gesammelte, durch Teilchendurchgang erzeugte, Ladung wird durch einen schnellen, ladungsempfindlichen Vorverstärker integriert und anschließend mithilfe eines Diskriminators digitalisiert. Diese Trefferinformation wird dann als Adresse des getroffenen Pixels in einem 6 Bit breiten Schieberegister, das alle Zellen einer Spalte verbindet, abgelegt und mit der Frequenz der in ATLAS auftretenden Wechselwirkungen von $40 MHz$ an den Rand getaktet. Dort kann dann mittels Adresse und Anzahl der Takte die genaue Trefferlage und -zeit errechnet werden. Eine weitere Datenverarbeitung, die insbesondere bei ATLAS verlangt wird, wurde in diesem Chip nicht implementiert.

- **Der MareBo (Marseille et Bonn)**

1997 folgte wiederum in Zusammenarbeit von CPPM und PIB der strahlenharte Nachfolger des Bier et Pastis, der MareBo [17]. Die verwendete DMILL Technologie ist ebenfalls ein $0,8 \mu m$ BiCMOS Prozeß mit einer vom Hersteller Atmel garantierten Strahlenresistenz bis zu einer Dosis von $10 Mrad$. Bis auf kleine Verbesserungen handelt es sich elektrisch um den selben IC (Integrated Circuit) wie der Bier et Pastis.

- **Der FEA/FEC (Front End AMS bzw. CMOS)**

Der erste voll ATLAS taugliche, allerdings nicht strahlenharte Auslesechip für den Pixeldetektor war 1998 der FEA bzw. der leicht abgewandelte FEC [18], der im Diskriminator anstatt von zwei bipolaren Transistoren eine reine CMOS Variante enthält. Seine Pixelmatrix besteht aus 18 Spalten mit je 160 Zeilen, die eine Zellengröße von $50 \mu m \times 400 \mu m$ besitzt. Das Auslesekonzept des Bier et Pastis wurde übernommen und eine Speicherverwaltung samt Datenserialisierung hinzugefügt.

- **Der FEB (Front End Berkeley)**

Parallel wurde am LBNL der FEB [19] in einem ebenfalls nicht strahlenharten $0,8 \mu m$ Prozeß der Firma Hewlett-Packard submittiert. Dieser beinhaltet auch eine 18×160 Pixelmatrix mit voll ATLAS konformer Funktion, benutzt jedoch im Vergleich zu den oben aufgezählten Chips eine völlig andere Auslesearchitektur, die in Abschnitt 8.2.2 noch eingehend erklärt wird.

- **Der FED (Front End DMILL)**

1999 wurde der erste voll ATLAS taugliche Pixelchip FED [20] in strahlenharter DMILL Technologie hergestellt, der ein mögliches Endprodukt für den tatsächlichen Einsatz im Experiment darstellen soll, was in den Kapiteln 8-11 noch eingehend

diskutiert wird. Dieser IC entstand in Kollaboration aller drei oben genannter Institute. Da sowohl FEA/FEC als auch FEB gut funktionieren und beide ihre Stärken und Schwächen aufweisen, sollte der Vorverstärkerteil des FEA/FEC und die Auslesearchitektur des FEB im Konzept übernommen werden. Nachträglich stellte sich heraus, daß dies die einzige Möglichkeit war, die Pixelgröße von $50 \mu m \times 400 \mu m$ einzuhalten, weil der DMILL Prozeß etwa 30 - 40 % mehr Platz benötigt als vergleichbare nicht strahlenharte Technologien. Dies ist neben elektrischen Unterschieden in den Transistoren auch der Hauptgrund dafür, daß die Designs nicht einfach durch Übertragung eines bereits existierenden Layouts entstehen, sondern wieder Schritt für Schritt neu erdacht werden müssen und letztlich ein völlig anderer Chip entsteht.

- **Der FEI (Front End IXX)**

Gegenwärtig arbeiten die drei Institute an der Entwicklung eines DSM Testchips, der bei zwei Firmen submittiert werden soll. Die entscheidenden Vorteile dieser $0,25 \mu m$ Technologien liegen in den winzigen Abmessungen der Strukturen und der dadurch begründeten natürlichen Strahlenhärte, die als Nebenprodukt abfällt. In Kapitel 12 werden diese Sachverhalte noch näher erläutert.

4 Halbleiterdetektoren

Halbleiterdetektoren werden in der Hochenergiephysik hauptsächlich zur genauen Ortsmessung von geladenen Teilchen und Photonen benutzt.

Im folgenden wird der Mechanismus der Detektion aufgezeigt und das Funktionsprinzip eines Siliziumdetektors erläutert.

4.1 Signalentstehung in Materie

Halbleiterdetektoren sind Festkörperionisationskammern, die anstatt eines Zählgases einen Festkörper zur Detektion benutzen. Dabei muß man die physikalischen Abläufe der Signalentstehung für geladene Teilchen und Photonen getrennt betrachten, weil grundlegend unterschiedliche Mechanismen vorliegen. Alle anderen Teilchen können in einem solchen Detektor nicht direkt nachgewiesen werden.

4.1.1 Wechselwirkung geladener Teilchen

Beim Durchgang eines geladenen Teilchens durch Materie finden entlang seiner Bahn im wesentlichen Stoßionisationen mit den umliegenden Atomen statt. Die Wechselwirkungen erfolgen in Form von Coulombabstoßung von den Hüllenelektronen und führen zu deren Anregung oder Entfernung. Zum Nachweis des Primärteilchens verwendet man letztere, aus der Hülle herausgeschlagenen, Elektronen bzw. die Löcher, die zurückbleiben.

Der mittlere Energieverlust des Teilchens pro zurückgelegter Wegstrecke in einem bestimmten Medium wird durch die Bethe-Bloch-Formel beschrieben [21]:

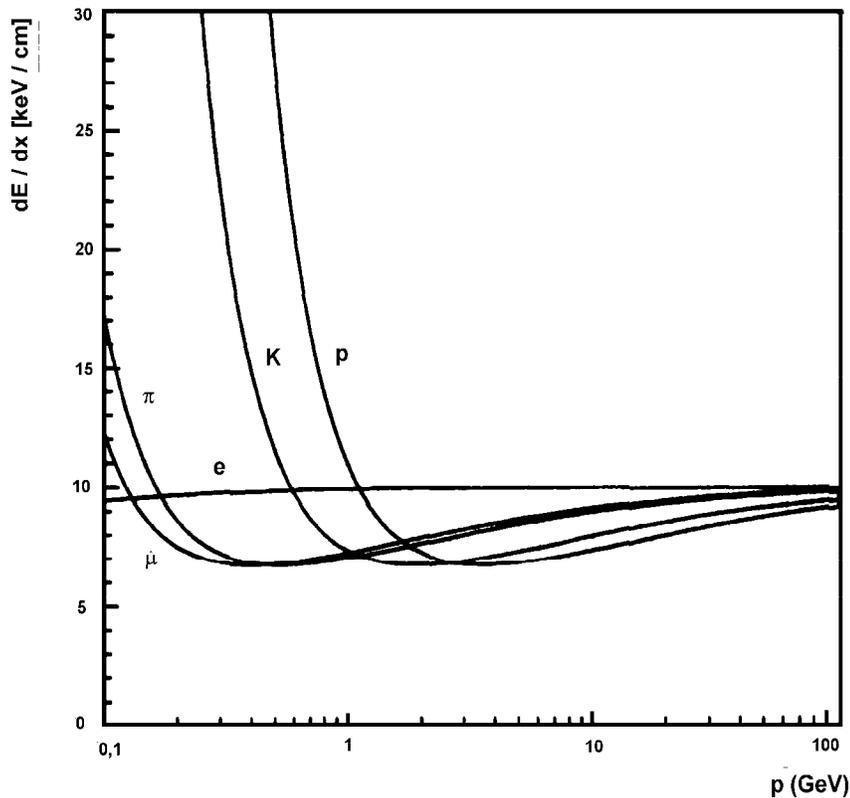
$$-\left\langle \frac{dE}{dx} \right\rangle = K z^2 \frac{Z}{A} \frac{1}{\beta^2} \left[\frac{1}{2} \ln \left(\frac{2m_e c^2 \beta^2 \gamma^2 T_{max}}{I^2} \right) - \beta^2 - \frac{\delta}{2} \right], \quad (4.1)$$

mit

$$\begin{aligned} K &= 4 \pi N_A r_e^2 m_e c^2 = 0,307 \text{ MeV cm}^2 / g \\ N_A &= \text{Avogadrozahl} \\ r_e &= \text{klassischer Elektronenradius} \end{aligned}$$

- m_e = Elektronenmasse
- z = Ladung des einfallenden Teilchens
- Z = Ordnungszahl des Mediums
- A = Massenzahl des Mediums
- β = Geschwindigkeit des einfallenden Teilchens
- $\gamma^2 = (1 - \beta^2)^{-1}$
- T_{max} = maximaler Energieübertrag bei einer Kollision
- I = Mittleres Ionisationspotential (für Si: $I = 172$ eV)
- δ = Dichtekorrektur.

Abbildung 8: Die Bethe-Bloch-Formel als Funktion des Teilchenimpulses für unterschiedliche Teilchen

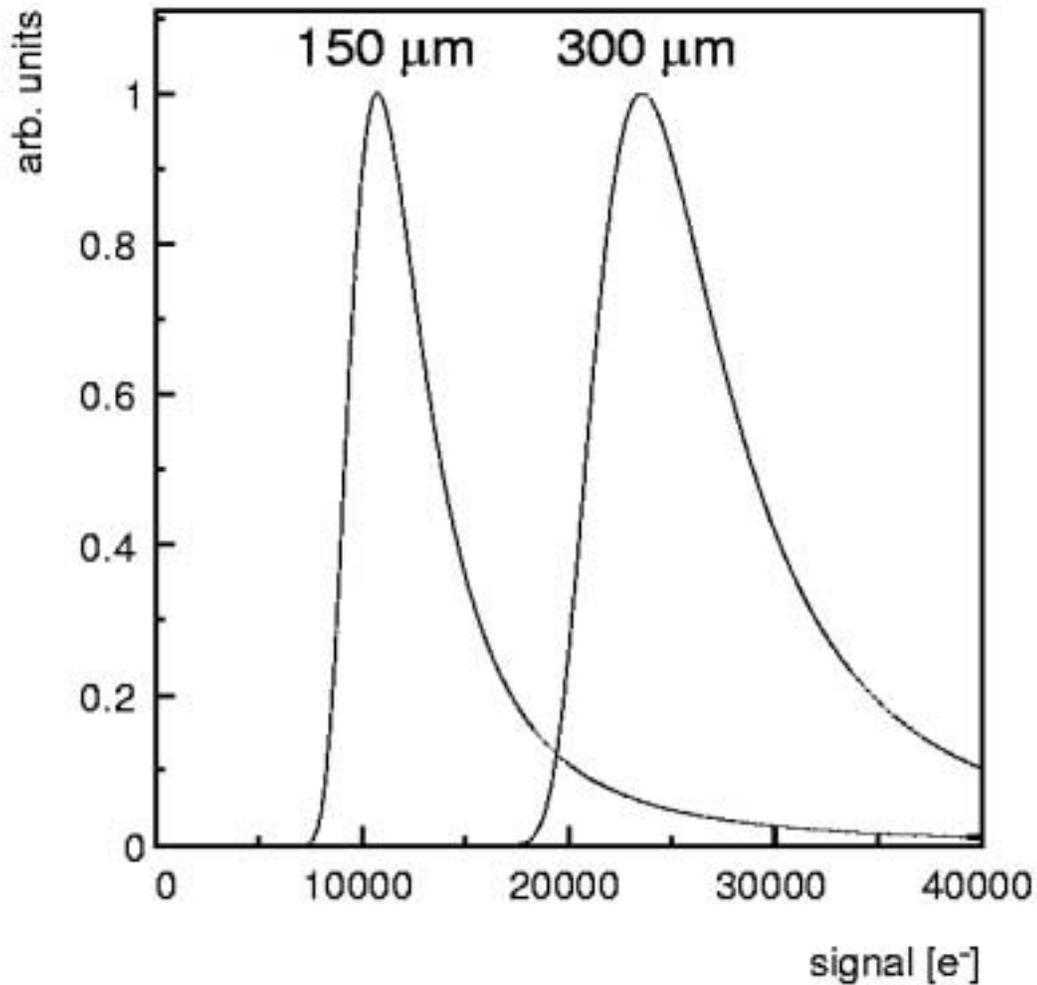


Für kleine Werte von $\beta\gamma$ nimmt der mittlere Energieverlust wie $1/\beta^2$ ab, bis er bei $\beta\gamma \simeq 3 - 4$ ein Minimum annimmt. In dieser Region nennt man die Teilchen MIP's (Minimal Ionizing Particle). Wird $\beta\gamma$ noch größer, erreicht man den Bereich des relativistischen Anstiegs gemäß $\ln(\beta^2\gamma^2)$, der aufgrund von Dichteeffekten in der Materie, die durch den Parameter δ beschrieben werden, schließlich in ein Plateau übergeht. Da dieser

Sättigungswert für Silizium nicht stark von dem Minimalwert abweicht, bezeichnet man alle Teilchen mit $\beta\gamma \geq 3$ als MIP's.

Die Bethe-Bloch-Formel gilt nur für Teilchen, die viel schwerer als Elektronen sind, weil bei ihrer Herleitung genau diese Annahme benutzt wird. Außerdem berücksichtigt sie weder Bremsstrahlung noch die Veränderung der Wirkungsschnitte bei Wechselwirkungen zwischen identischen Teilchen, so daß insbesondere Elektronen nicht exakt dieser Formel folgen.

Abbildung 9: Die Landauverteilung für verschiedene Dicken von Silizium [22]



Die sekundären Elektronen werden durch einen statistischen Prozeß erzeugt, da die Energie des Primärteilchens in vielen einzelnen Potentialstößen an die Hüllenelektronen abge-

geben wird. Durchläuft das nachzuweisende Teilchen eine lange Strecke, also einen 'dicken' Detektor, entspricht die Verteilung des Energieverlustes einer Gaußverteilung, weil sehr viele Stöße erfolgen. Handelt es sich jedoch um einen kurzen Weg bzw. um einen 'dünnen' Detektor mit geringerer Stoßhäufigkeit, verschiebt sich die Verteilung zu kleineren Werten hin, und sie erhält eine asymmetrische Form; man erhält eine Landau-Verteilung. Ursprung für den Ausläufer bei hohen Energien ist die Erzeugung von sogenannten δ -Elektronen, die ihrerseits einen Energieübertrag erhalten, der wiederum zur weiteren Stoßionisation ausreicht.

In einem Siliziumdetektor einer Dicke von $300 \mu m$ werden bei senkrechtem Teilchendurchtritt im Mittel $117 keV$ deponiert. Der wahrscheinlichste Wert liegt etwas niedriger bei $84 keV$ und rührt von der Asymmetrie der Landauverteilung her. Zur Erzeugung eines Elektron-Loch-Paares benötigt man im Mittel $3,62 eV$, so daß ein MIP durchschnittlich 32500 davon erzeugt und im Fall der höchsten Wahrscheinlichkeit 23300.

4.1.2 Wechselwirkungen von Photonen

Bei Photonen handelt es sich um elektrisch neutrale Teilchen, die daher nicht der Coulombwechselwirkung unterliegen, sondern anderen Mechanismen, die nachfolgend aufgeführt sind:

- **Photoeffekt**

Bei einer Photonenergie unterhalb von $50 keV$ tritt hauptsächlich der Photoeffekt auf, bei dem das einfallende Photon seine gesamte Energie auf ein gebundenes Elektron überträgt und selbst verschwindet. Das angeregte Elektron besitzt dann die Energie

$$E = E_{Photon} - E_{Bindung} \quad (4.2)$$

und kann diese wiederum durch weitere Ionisationen im Detektor abgeben. Falls das Photoelektron komplett im Sensor absorbiert wird, kann auf diese Weise eine Energieeichung desselben durchgeführt werden.

- **Paarbildung**

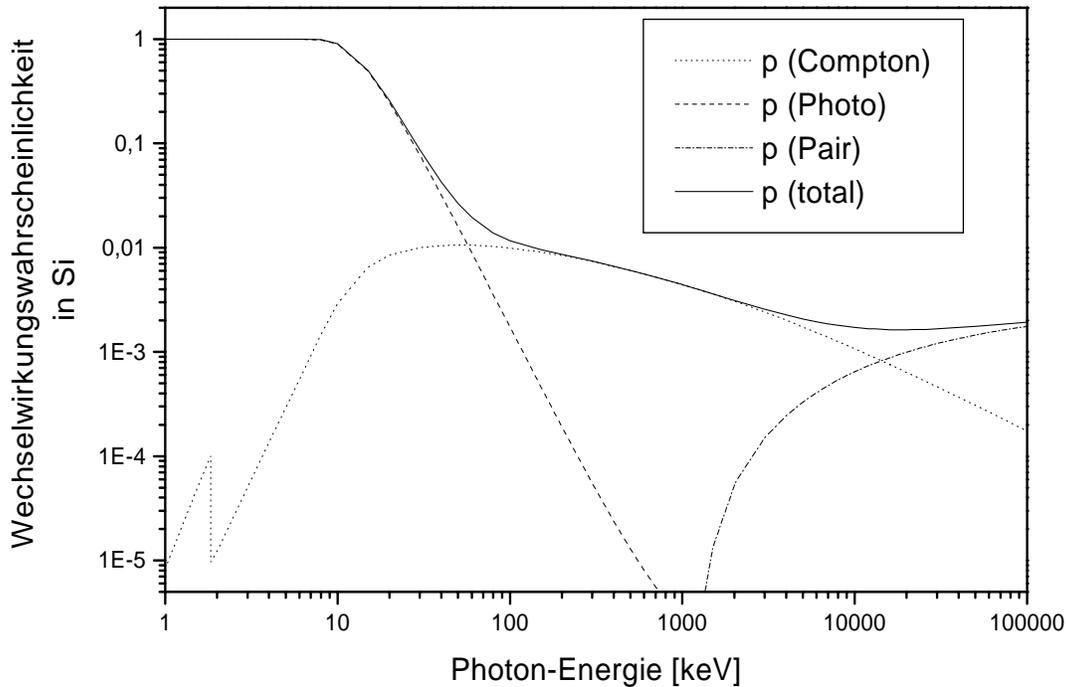
Ab einer Photonenergie oberhalb von $1,022 MeV$ gelangt man in den Bereich, wo man die Ruhemasse eines Elektron-Positron-Paares erzeugen kann. Unter Anwesenheit eines massiven Teilchens, das den Rückstoßimpuls aufnimmt, findet oberhalb von $50 MeV$ hauptsächlich der Mechanismus dieser Paarbildung statt.

- **Comptoneffekt**

Im restlichen Energiebereich tritt im wesentlichen der Comptoneffekt auf, der auf der elastischen Streuung von Photonen an quasifreien Elektronen beruht. Hier erhält das

Comptonelektron beim Stoß nur einen Teil der Photonenergie, welches seinerseits mit einer größeren Wellenlänge weiterfliegt.

Abbildung 10: Die Wechselwirkungswahrscheinlichkeit von Photonen in $300 \mu\text{m}$ Silizium



Die Wechselwirkungswahrscheinlichkeit [23] von hochenergetischen Photonen im GeV -Bereich in einem $300 \mu\text{m}$ dicken Siliziumdetektor liegt bei ATLAS allerdings unter einem Prozent, so daß diese Prozesse an dieser Stelle keiner weiteren Untersuchung bedürfen.

4.2 Funktionsprinzip eines Siliziumdetektors

4.2.1 Grundlagen

Die Einteilung von chemischen Elementen in Leiter, Halbleiter und Isolatoren erfolgt durch Betrachtung des sogenannten Bändermodells und der Fermieenergie E_F . Während

die Elektronen einzelner Atome diskrete Energieniveaus besitzen, führt deren riesige Anzahl innerhalb von makroskopischen Kristallen wegen ihrer gegenseitigen Beeinflussung zur Ausbildung von quasikontinuierlichen Energiebändern, die durch quantenmechanisch verbotene Zonen, in denen sich keine Elektronen befinden, getrennt sind. Zur Charakterisierung der Materialien verwendet man nun das energetisch am höchsten liegende, noch vollbesetzte Band, das man Valenzband nennt, und das darauf folgende Band, das noch freie Plätze besitzt und Leitungsband heißt. Im Falle eines Leiters überlappen sich diese Bänder, und die Fermienergie liegt im selben Energiebereich. Beim Halbleiter und beim Isolator ist das Leitungsband bei einer absoluten Temperatur $T = 0 \text{ K}$ unbesetzt, und die Fermienergie liegt in der Bandlücke, so daß keine elektrische Leitung erfolgen kann. Der Unterschied zwischen beiden liegt ausschließlich in der Breite der Bandlücke, die bei einem Isolator Werte größer als 3 eV annimmt und bei dem Halbleiter Silizium bei $1,12 \text{ eV}$ liegt. Damit ergibt sich nach der Fermi-Dirac-Statistik für die Besetzungswahrscheinlichkeit $F(E)$ des Leitungsbandes [24]

$$F(E) = \frac{1}{1 + e^{\frac{E-E_F}{kT}}}, \quad (4.3)$$

mit

$$k = \text{Boltzmannkonstante} = 1,3806 \times 10^{-23} \text{ J/K}, \quad (4.4)$$

daß im Falle des Isolators praktisch keine Elektronen den Übergang in das Leitungsband durch thermische Anregung schaffen, wogegen dies in Silizium durchaus möglich ist. Darum nennt man Halbleiter auch 'Heißleiter'.

Die Differenz zwischen der $1,12 \text{ eV}$ breiten Bandlücke und der Energie von $3,61 \text{ eV}$, die man zur Generation eines Elektron-Loch-Paares benötigt, wird in Form von Schwingungsenergie an das Kristallgitter (Phononen) übertragen.

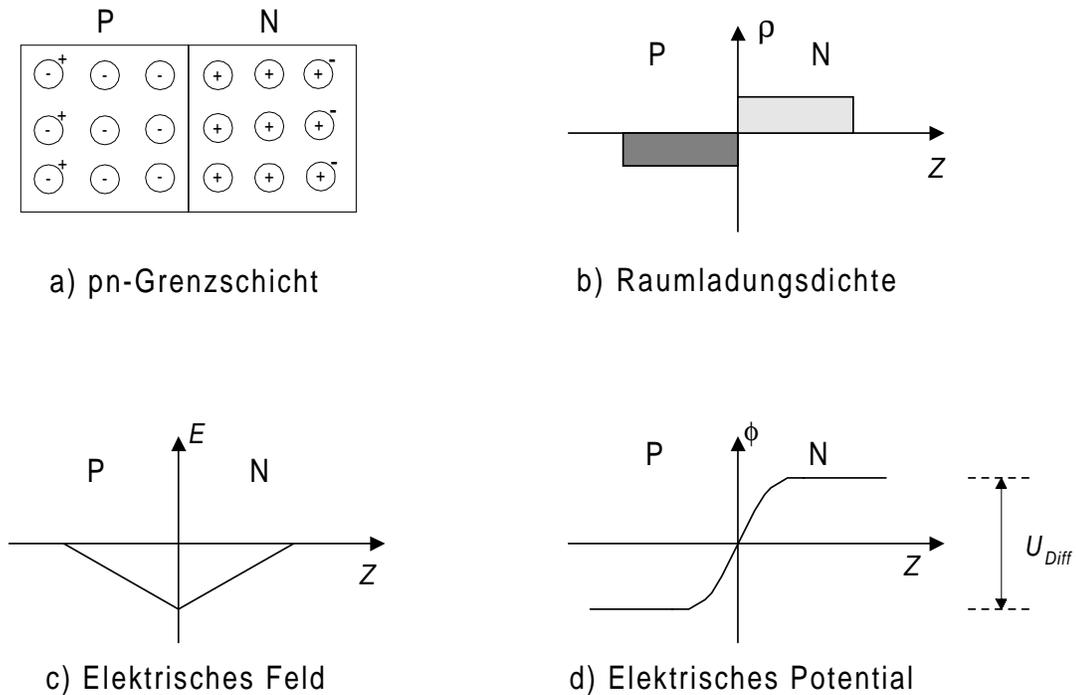
4.2.2 Die Halbleiterdiode als Detektor

In reinem Silizium befinden sich bei Zimmertemperatur größenordnungsmäßig 10^{10} Ladungsträger pro Kubikzentimeter im Leitungsband und entsprechend ebenso viele Löcher im Valenzband. Wenn man diese Zahl mit den 23300 Elektron-Loch-Paaren vergleicht, die ein MIP in $300 \mu\text{m}$ Silizium deponiert, erkennt man, daß damit kein Teilchennachweis erfolgen kann.

Daher verunreinigt man gezielt den Halbleiter mit Fremdatomen aus den benachbarten Hauptgruppen des Periodensystems, meist Bor als (Elektronen-) Akzeptor bzw. Arsen

oder Phosphor als (Elektronen-) Donator. Diesen Vorgang nennt man p -Dotierung bzw. n -Dotierung. Die Folge ist zunächst eine von der Fremdatomanzahl abhängige Leitfähigkeitsverbesserung, weil je nach Typ ein Überschuß an Elektronen (bei n -Typ) bzw. Löchern (bei p -Typ) entsteht, und außerdem beim Kontakt unterschiedlich dotierter Gebiete eine gleichrichtende Wirkung, was im folgenden näher erläutert wird.

Abbildung 11: Die relevanten elektrischen Größen am pn -Übergang [25]



Bei Vorliegen eines pn -Übergangs (Diode) fließt aufgrund der Konzentrationsgefälle der beiden Ladungsträgersorten zunächst ein Diffusionsstrom, der von der p -Zone zur n -Zone gerichtet ist, wenn man das Vorzeichen der Ladungsträger berücksichtigt. Zurück bleiben ortsfeste Raumladungen der Atomrümpfe, und zwar positive auf der n -Seite und negative auf der p -Seite. Im Übergangsbereich entsteht auf diese Weise ein elektrisches Feld, das einen Feldstrom entgegen dem Diffusionsstrom verursacht und so lange wirkt, bis ein dynamisches Gleichgewicht entsteht. Dort existieren keine freie Ladungsträger mehr und man spricht von der Verarmungs-, Raumladungs- oder der Depletionszone, deren Breite d sich wie folgt errechnet:

$$d = \sqrt{\frac{2\epsilon}{q} \frac{N_A + N_D}{N_A N_D} (U_D + U_{ext})}, \quad (4.5)$$

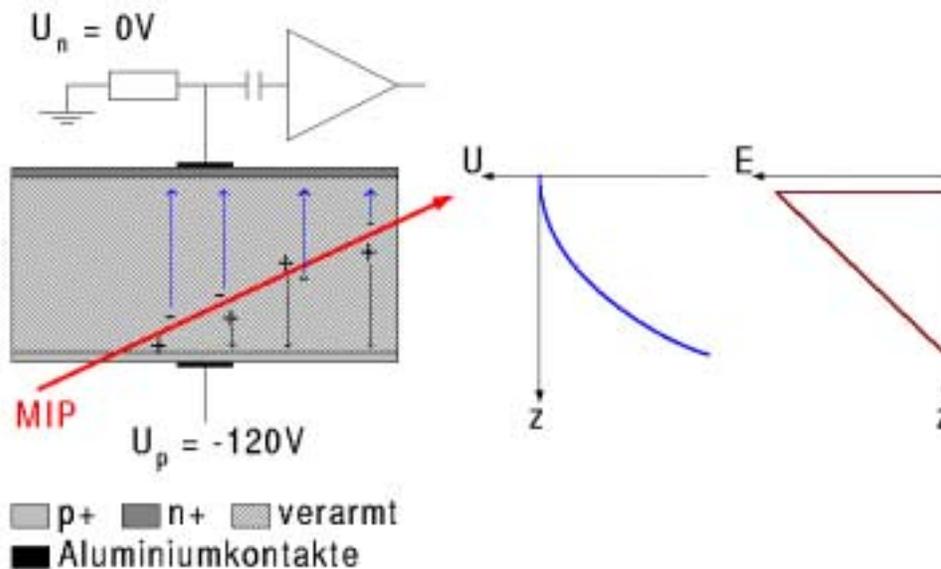
mit

- ϵ = Dielektrizitätskonstante
- q = Ladung
- N_A/N_D = Akzeptoren-/Donatorenkonzentration
- U_D = Diffusionsspannung (für Si: $U_D \sim 0,7 \text{ V}$)
- U_{ext} = extern von p nach n angelegte Spannung.

Man erkennt, daß die Depletionszone für kleinere Werte als $U_{ext} = -0,7 \text{ V}$ (Durchlaßrichtung) verschwindet, so daß in diesem Fall ein Stromfluß von p -Zone nach n -Zone ermöglicht wird. Wählt man größere Werte für U_{ext} (Sperrichtung), wächst die Dicke des Verarmungsbereichs an, und zwar hauptsächlich in das schwächer dotierte Gebiet hinein.

Daher benutzt man in der Regel folgenden Aufbau, wenn man eine Siliziumdiode als Teilchendetektor verwenden will (natürlich kann man jeweils p und n vertauschen):

Abbildung 12: Die pn -Diode als Teilchendetektor



Ein schwach p -dotiertes Substrat ($\sim 10^{12} \text{ cm}^{-3}$) erhält auf einer Seite eine wenige μm dünne hochdotierte n^+ -Zone ($\sim 10^{19} \text{ cm}^{-3}$), so daß man durch Anlegen einer ausreichend großen Sperrspannung praktisch eine Verarmung des gesamten Detektors erreicht und so ein Volumen schafft, welches völlig frei von beweglichen Ladungsträgern ist. Dieses ist letztlich sensitiv für die wenigen Ladungen, die der Durchgang eines MIP's erzeugt. Um zu verhindern, daß der metallische p -Seitenkontakt die Depletionszone berührt (Schottkydiode), was durch Ladungsträgerinjektion einen Diodendurchbruch verursachen würde, dotiert man dort eine p^+ -Schicht, die nicht depletiert werden kann. Für einen typischen $300 \mu\text{m}$ dicken Siliziumdetektor benötigt man eine Sperrspannung von etwa $120 - 140 \text{ V}$.

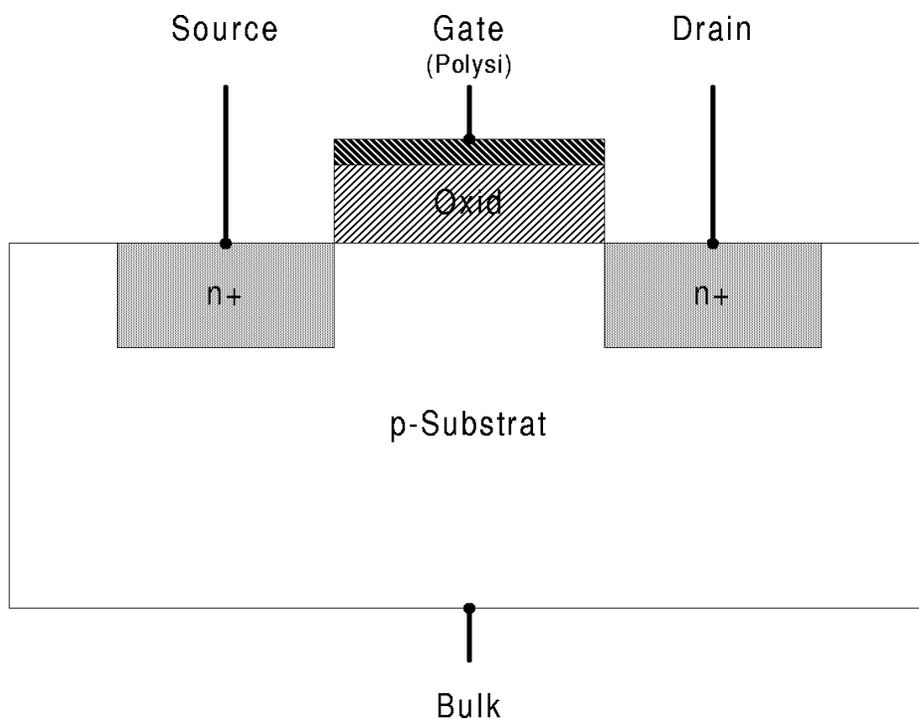
Nun offenbart sich das Funktionsprinzip des Halbleiterdetektors: Die in dem vollständig depletierten Detektor, durch Teilcheneinfall erzeugten, Elektron-Loch-Paare werden durch das elektrische Feld getrennt (können also nicht direkt wieder rekombinieren) und driften zu den jeweiligen Elektroden. Dort entsteht dann während dieser Bewegung in Abhängigkeit der Polarität und Anzahl der Ladungen das weiter zu verarbeitende Signal. Eine zweidimensionale Ortsinformation erhält man einfach, indem man eine Elektrode in Form einer Matrix segmentiert und jede Zelle getrennt ausliest.

5 MOSFETs

Die in dieser Arbeit behandelten Elektronikchips enthalten hauptsächlich MOSFETs (**M**etall **O**xide **S**emiconductor **F**ield **E**ffect **T**ransistor), die in zwei Bauweisen verfügbar sind: Dem n -Kanal MOSFET (NMOS) und dem p -Kanal MOSFET (PMOS). Im Gegensatz zu bipolaren Transistoren erfolgt die Stromleitung bei MOSFETs nur durch eine Ladungsträgersorte, und die Stromsteuerung geschieht nahezu leistungslos durch ein elektrisches Feld.

Ein n -Kanal MOSFET besteht aus einer Polysilizium-Oxid-Silizium Struktur (aus historischen Gründen resultiert die Bezeichnung 'MOS', weil früher statt Polysilizium Aluminium verwendet wurde), die als Steuerelektrode ('Gate') fungiert, und zwei angrenzenden n^+ -dotierten Inseln, die in ein p -dotiertes Substrat ('Bulk') eingebettet sind. Die beiden symmetrischen n^+ -Gebiete werden als 'Source' bzw. 'Drain' bezeichnet, wobei im folgenden die 'Source' immer an das niedrigere Potential angeschlossen wird.

Abbildung 13: Querschnitt durch einen NMOS

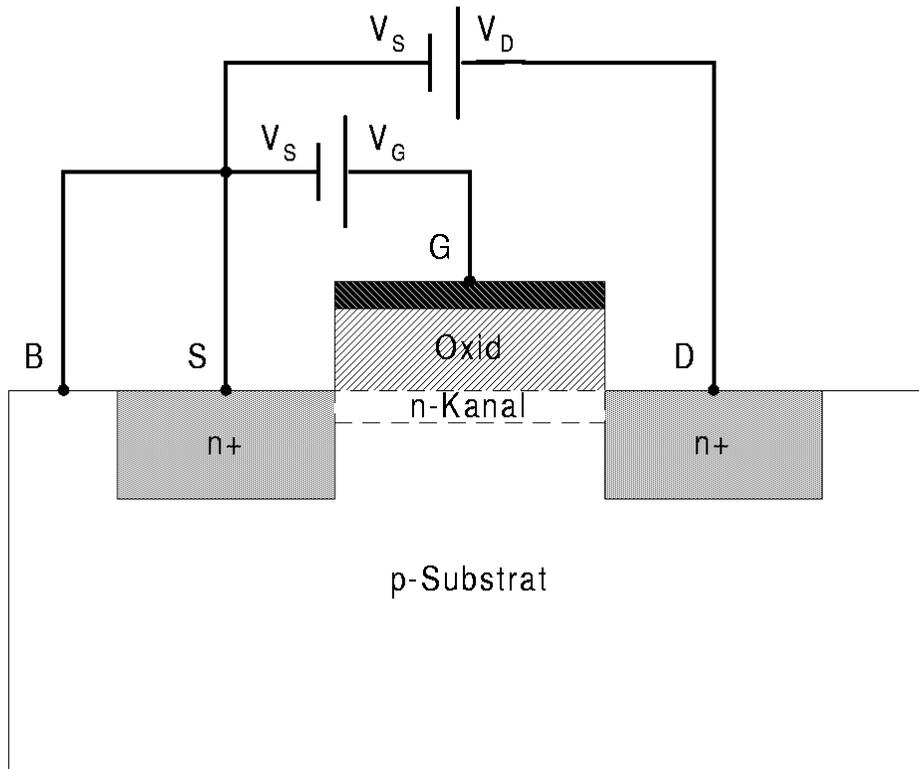


Das 'Gate' bildet einen Plattenkondensator mit Siliziumdioxid als Dielektrikum und beeinflusst durch Anlegen einer Spannung die Ladungsträgeranzahl bzw. Leitfähigkeit zwischen 'Source' und 'Drain' im p -Substrat. Beim Betrieb eines NMOS legt man den 'Bulk' auf

ein festes negatives Potential gegenüber 'Source' und 'Drain', so daß alle Dioden zwischen den verschiedenen Diffusionsgebieten sperren und kein Strom fließt. Sobald die 'Gate'-'Source'-Spannung $V_{GS} = V_G - V_S$ einen bestimmten Schwellenwert V_{THn} , der von den Kristallfehlern an der Oxid-Substrat Grenzfläche und den Eigenschaften der verwendeten Materialien abhängt, überschreitet, bildet sich durch Influenz eine Elektronenschicht (Inversionskanal) direkt unter dem Oxid, die 'Source' und 'Drain' leitend miteinander verbindet. Auch unterhalb der Schwellenspannung V_{THn} bewirkt die 'Gate'-'Source'-Spannung schon einen Stromfluß, der exponentiell mit V_{GS} ansteigt. Dieser Sperrstrom erreicht jedoch in den in dieser Arbeit behandelten Schaltungen vernachlässigbar geringe Werte und wird im folgenden gleich Null gesetzt.

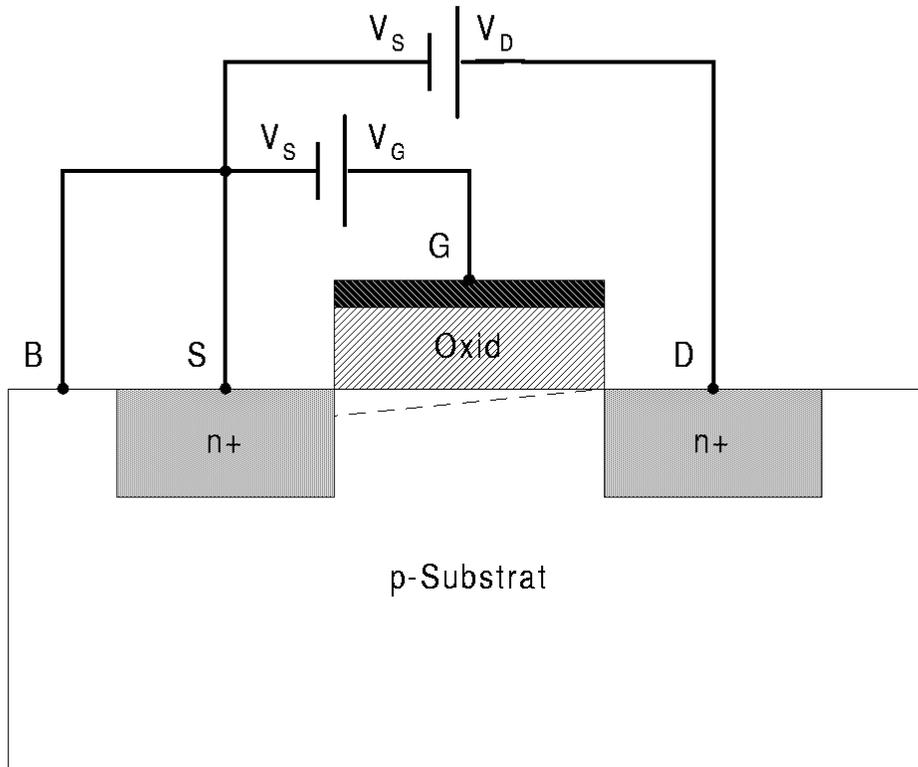
Legt man nun eine positive Spannung $V_{DS} = V_D - V_S$ zwischen 'Drain' und 'Source' an, fließt ein Elektronenstrom durch den n -Kanal. Ebenso wie durch V_{GS} könnte man durch Variation des 'Bulk'-Potentials eine Steuerung des 'Drainstroms' I_D mittels $V_{BS} = V_B - V_S$ bewirken, da genau genommen eine Parallelschaltung eines NMOS und eines parasitären JFETs (**J**unction **F**ET) vorliegt. In den meisten Technologien besteht jedoch nur die Möglichkeit, alle 'Bulks' einer Transistorsorte auf dasselbe Potential zu legen, so daß man gewöhnlich immer den 'Bulk'-Kontakt mit dem 'Source'-Kontakt verbindet und auf diese Weise den sogenannten 'Bulkeffekt' eliminiert.

Abbildung 14: Querschnitt und Betriebsspannungen eines NMOS



Solange V_{DS} bei festem V_{GS} kleine Werte besitzt, verhält sich der NMOS wie ein Widerstand, das heißt I_D wächst proportional zu V_{DS} , und man spricht vom linearen Bereich. Bei weiter ansteigendem V_{DS} wird der Inversionskanal von der 'Drain'-Seite her zunehmend abgeschnürt, bis 'Source' und 'Drain' nicht mehr durch denselben verbunden sind. Durch Drift erreichen jedoch weiterhin die Elektronen noch die 'Drain', allerdings bleibt I_D jetzt annähernd konstant. Man nennt daher den Bereich, wo $V_{DS} \geq V_{GS} - V_{THn}$ gilt, Sättigungsbereich und den Wert $V_{DS_{sat}} = V_{GS} - V_{THn}$ Sättigungsspannung.

Abbildung 15: Querschnitt eines NMOS in Sättigung



Durch das Einschnüren des Kanals verkürzt sich dieser, was einen leichten Anstieg des Drainstroms im Sättigungsbereich bewirkt. Diesen Kanallängenmodulationseffekt nennt man auch Earlyeffekt, der durch den Parameter λ beschrieben wird und sich umgekehrt proportional zur Gatelänge verhält.

Die Ausgangskennlinie $I_D(V_{DS})$ und die Transferkennlinie $I_D(V_{GS})$ eines NMOS lassen sich durch unterschiedlich komplizierte Modelle beschreiben. Für Handberechnungen eignet sich am besten das sogenannte Level-1 Modell [26], dessen Genauigkeit zum Verständnis der Schaltungen im allgemeinen ausreicht. Hiernach ergeben sich folgende quantitativen Zusammenhänge:

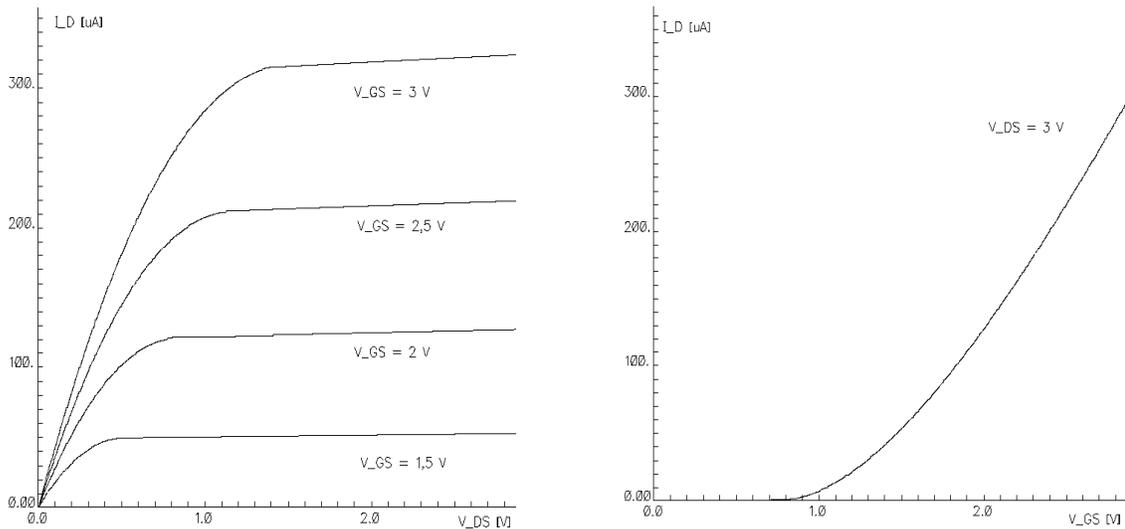
$$I_D = K_n \frac{W}{L} \left(V_{GS} - V_{THn} - \frac{V_{DS}}{2} \right) V_{DS} (1 + \lambda V_{DS}) \quad \text{linearer Bereich} \quad (5.1)$$

$$I_D = \frac{K_n}{2} \frac{W}{L} (V_{GS} - V_{THn})^2 (1 + \lambda V_{DS}) \quad \text{Sättigungsbereich,} \quad (5.2)$$

mit

- $K_n = \mu C_{Ox}$, Transkonduktanzparameter für NMOS (technologieabhängig)
- $\mu =$ Ladungsträgerbeweglichkeit
- $C_{Ox} =$ Gatekapazität pro Fläche
- $W =$ Gatebreite
- $L =$ Gatelänge.

Abbildung 16: Die Kennlinien eines NMOS: Ausgangskennlinienfeld (links) und Eingangskennlinie (rechts)



Alle Aussagen in diesem Kapitel gelten analog im Falle des PMOS, wenn man einfach n -Implantationen und p -Implantationen vertauscht, die Betriebsspannungen umpolt sowie die entsprechenden Technologieparameter wählt. Hierbei fällt besonders der Transkonduktanzparameter ins Gewicht, weil aufgrund der geringeren Löcherbeweglichkeit ein PMOS bei gleichen Einstellungen wie für einen gleich großen NMOS etwa nur die Hälfte bis ein Drittel von I_D liefert.

6 Strahlenschäden in Silizium

Zunächst werden allgemein die bei Bestrahlung auftretenden Schäden sowohl am Siliziumsubstrat selbst als auch an der angrenzenden Passivierungsschicht, bestehend aus Siliziumdioxid, und der dazwischen liegenden Grenzfläche behandelt. Die Mechanismen sind, da es sich um dieselben Materialien handelt, für Sensor- und Elektronikchips identisch, allerdings ergeben sich unterschiedliche Auswirkungen je nach Anwendung und müssen voneinander getrennt betrachtet werden.

Viele der hier aufgeführten Schädigungen können teilweise durch komplizierte Ausheilmechanismen abgeschwächt werden [27, 28, 29]. Dieses Thema soll jedoch nicht im Rahmen dieser Arbeit behandelt werden, sondern lediglich die Ergreifung präventiver Maßnahmen erläutert werden.

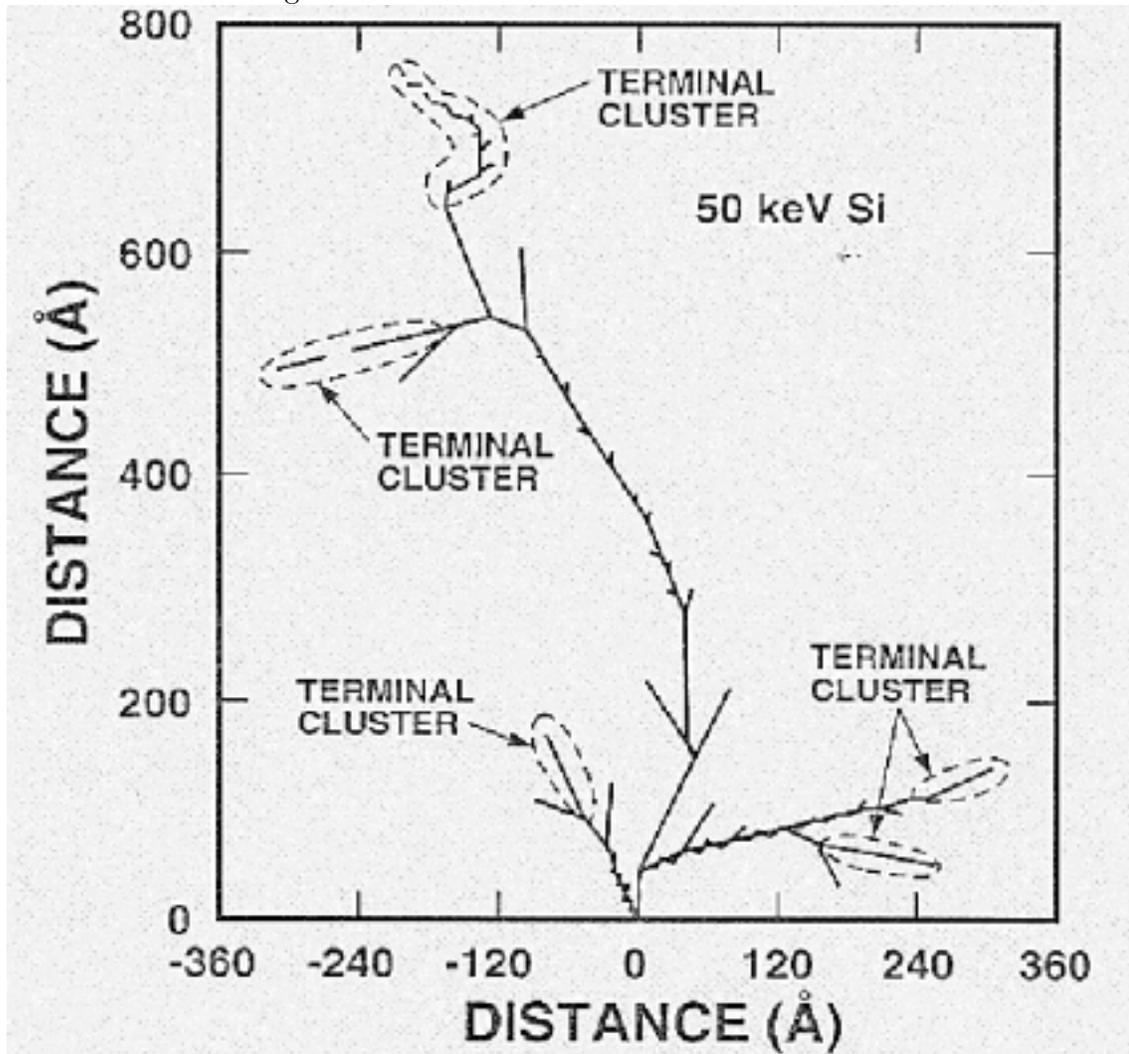
6.1 Schäden am Siliziumsubstrat

Während es sich bei der Ionisation durch den Einfall von Teilchen in den Siliziumkristall um einen reversiblen Prozeß handelt und diese daher keine bleibenden Schäden hervorruft, verursacht der nicht ionisierende Energieverlust NIEL (**N**on **I**onizing **E**nergy **L**oss) [30] der Strahlung bleibende Veränderungen in der Kristallstruktur. Der NIEL hängt ab von der Teilchensorte und -energie und zeigt eine lineare Korrelation zu der Anzahl der Defekte. Daher werden zum besseren Vergleich alle Messungen auf den Wert bezogen, den 1 *MeV* Neutronen erzeugen würden.

Ab einer Deposition von etwa 15 *eV* können einzelne Siliziumatome aus dem Gitter herausgeschlagen werden ('Vacancies') und sich auf Zwischengitterplätzen festsetzen ('Interstitials'). Diese Punktdefekte zeigen sich a priori nicht sehr resistent, bilden jedoch zusammen mit anderen Punktdefekten oder Verunreinigungen im Substrat sehr stabile Komplexe, die jeweils unterschiedliche Energieniveaus und damit auch verschiedene Eigenschaften besitzen. Erhält das zuerst angestoßene Atom PKA (**P**rimary **K**nock-on **A**tom) eine Energie von mehr als etwa 2 *keV*, kann es selbst wiederum Punktdefekte erzeugen und so lokal eine sehr große Anzahl von Kristallversetzungen bewirken, sogenannte 'cluster', die sich anders verhalten als die Komplexe (s. Abbildung 17).

Entscheidend bei all diesen Defekten ist die Lage ihrer Energieniveaus. Nur wenn sie sich in der Bandlücke des Siliziums befinden, ergeben sich für die Funktion als Detektor bzw. IC beobachtbare Folgen, die sich wie folgt äußern: In Silizium werden bei Bestrahlung immer Donatoren vernichtet, aber gleichzeitig nur Akzeptoren generiert. Dieses Phänomen erklärt sich dadurch, daß sich bevorzugt V-P-Komplexe ('Vacancy'-Phosphor) und V-V-Komplexe ('Vacancy'-'Vacancy') bilden [30, 31]. Die Donatoren werden also im V-P-Komplex eingefangen und verlieren ihre Wirkung, während die V-V-komplexe als neue Akzeptoren fungieren.

Abbildung 17: Der NIEL für ein 50 KeV Neutron in Silizium



Aus Messungen erhält man diesen funktionalen Verlauf der effektiven Dotierungskonzentration N_{eff} [21]:

$$N_{eff} = N_{D0} e^{-\delta \Phi} - \beta \Phi, \quad (6.1)$$

mit

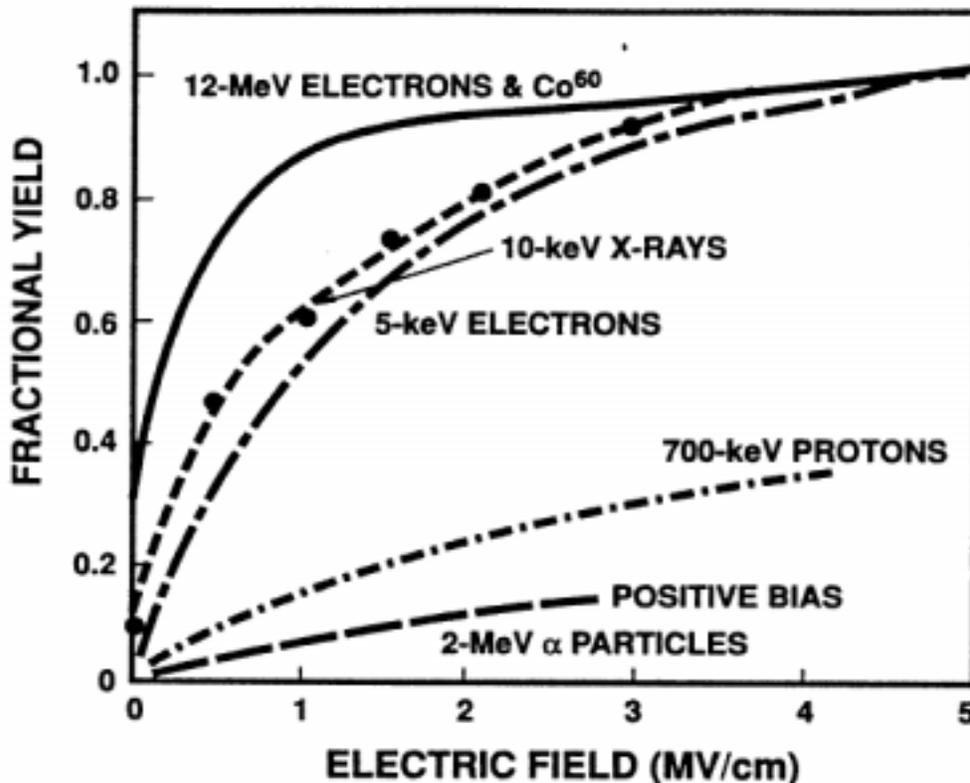
- N_{D0} = Donatorkonzentration vor Bestrahlung
- δ = $6 \times 10^{14} \text{ cm}^2$
- β = $0,03 \text{ cm}^{-1}$
- Φ = Teilchenfluß bezogen auf 1MeV Neutronen.

Der erste Term beschreibt in dieser Formel die Abnahme der Donatorkonzentration, während der zweite die Akzeptorgeneration charakterisiert.

6.2 Schäden im Siliziumdioxid

Der dominante Mechanismus im Siliziumdioxid beruht auf dem Energieverlust der Strahlung in Form von Ionisation, welche eine Energie von etwa 18 eV benötigt. Auch hier treten die oben erwähnten Kristallversetzungen auf, diese sind jedoch im Vergleich zu der Erzeugung von Elektron-Loch-Paaren vernachlässigbar. Die meisten Ladungsträgerpaare rekombinieren sofort wieder, in Anwesenheit eines elektrischen Feldes wird dieser Effekt jedoch mit zunehmender Feldstärke geringer.

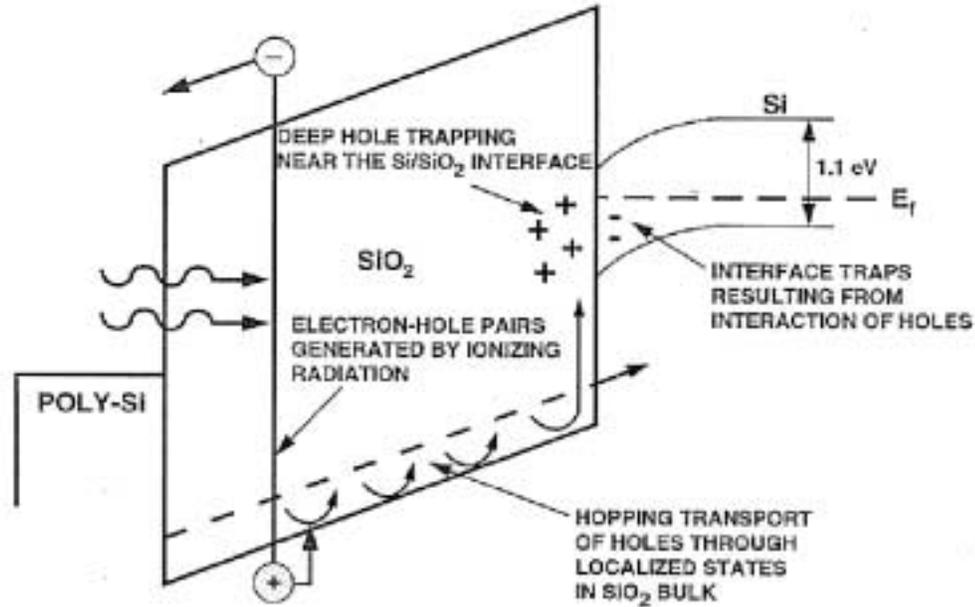
Abbildung 18: Wahrscheinlichkeit der nicht rekombinierten Ladungsträger in SiO_2 für verschiedene Teilchenstrahlung in Abhängigkeit vom elektrischen Feld [32]



Da in gewachsenem Siliziumdioxid nur wenige Elektronen-'Traps' entstehen, diffundieren die restlichen Elektronen aufgrund ihrer hohen Beweglichkeit [30] von $\mu_e = 2 \times 10^{-5} \text{ cm}^2/\text{Vs}$ innerhalb weniger Picosekunden aus dem Oxid hinaus, gemäß der Richtung des angelegten Feldes. Die weniger beweglichen positiven Oxidladungen mit $\mu_{Ox} =$

$20 \text{ cm}^2/\text{Vs}$ verbleiben zunächst an ihren Entstehungsorten und springen dann in Richtung des angelegten Feldes von einem 'Trap' zum nächsten bis zur Grenzfläche, wo sie schließlich in tieferen 'Traps' verbleiben.

Abbildung 19: 'Hopping' Transport der Oxidladungen [33]



Dieser 'Hopping Transport' geschieht durch sogenannte E' -Zentren [31, 34, 35], die von der Struktur des SiO_2 herrühren. Da das Oxid im wesentlichen aus tetraederförmigen SiO_4 -Einheiten besteht, gibt es auch Stellen, an denen Siliziumatome direkt miteinander verbunden sind. Die Oxidladung lagert sich aufgrund ihrer hohen Elektronegativität an eine solche kovalente Bindung an, trennt die Siliziumatome und bildet so eine stabile Konfiguration mit einer positiven Ladung. Die energetisch tiefen 'Traps' kommen daher, daß an der Grenzfläche unterschiedlicher Materialien natürlicherweise immer sehr viele und große Kristalldefekte vorliegen.

Die Oxidladungen bewirken neben dem extern angelegten Feld eine zusätzliche Spannung U_{Ox} , die vom Oxid in Richtung Substrat weist [36]:

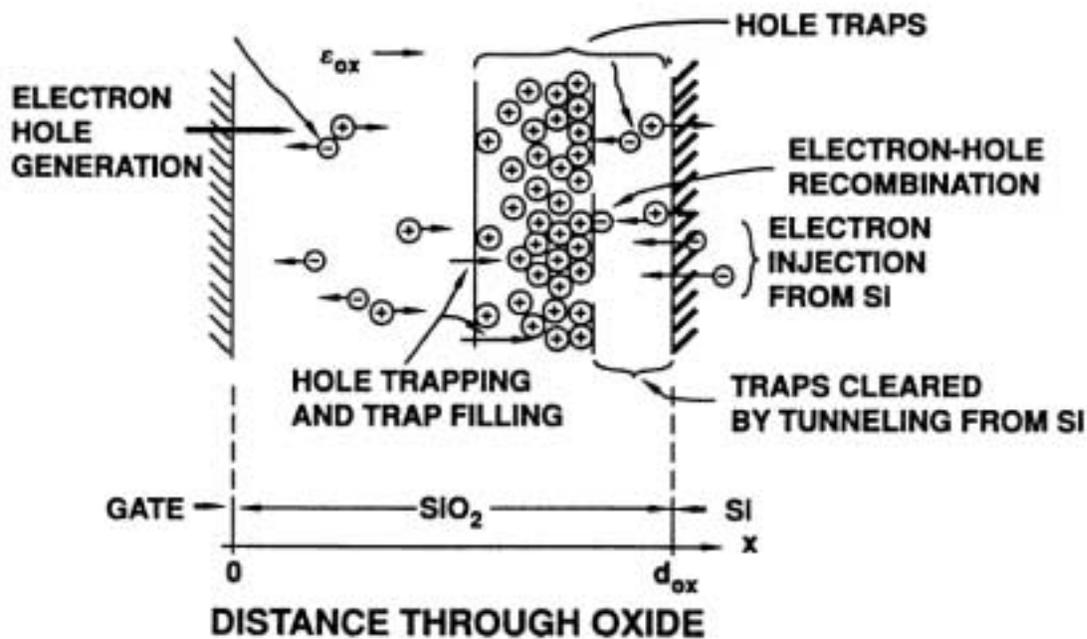
$$U_{Ox} = -\frac{1}{\epsilon_{Ox}} \int_0^{d_{Ox}} x \rho(x) dx = -\frac{q}{\epsilon_{Ox}} d_{Ox} N_{Ox}, \quad (6.2)$$

mit

- ϵ_{Ox} = Dielektrizitätskonstante des Oxids
- d_{Ox} = Dicke des Oxids
- ρ = Oxidladungsdichte
- q = Elementarladung
- N_{Ox} = Anzahl der Oxidladungen/Fläche.

Wenn N_{Ox} durch die Bestrahlung so stark angewachsen ist, daß die Elektronen rekombinieren, bevor sie aus dem Oxid hinausgedriftet sind, stellt sich ein Sättigungswert ein, und der Schaden kann sich nicht mehr vergrößern. Weiterhin läßt sich aus Formel 6.2 ersehen, daß U_{Ox} linear mit der Oxiddicke anwächst.

Abbildung 20: Rekombination der grenzflächennahen Oxidladungen durch Tunneleffekt



Im Hinblick auf Kapitel 12 bleibt noch eine Besonderheit zu erwähnen: Die Oxidladungen verbleiben nicht direkt an der Grenzfläche, sondern etwa 4 – 5 nm davon entfernt [37]. Dieses Phänomen läßt sich durch den Tunneleffekt erklären, wonach Elektronen durch die Potentialbarriere hindurch diffundieren können und anschließend mit den Oxidladungen rekombinieren (s. Abbildung 20).

6.3 Schäden an der Grenzfläche

Auch hier sind wiederum nur die Strahlungseffekte, die durch Ionisation hervorgerufen werden, von Bedeutung. Da bisher diese Grenzflächendefekte noch nicht verstanden sind, dient der Begriff als Definition für alle durch Bestrahlung erzeugten Energieniveaus in der Bandlücke des Siliziums, die direkt an der Grenzfläche zum Oxid lokalisiert sind [38]. Diese Energiezustände haben kontinuierliche Werte und fungieren als Donatoren, wenn sie oberhalb der Mitte der Bandlücke liegen, als Akzeptoren, wenn sie darunter liegen, und als Rekombinations- bzw. Generationszentren, im Falle daß sie exakt in der Mitte liegen.

6.4 Auswirkungen auf die Detektoren

Die strahleninduzierten Schädigungen führen in dreierlei Hinsicht zu einer Verschlechterung des Detektorverhaltens:

6.4.1 Erhöhung der Depletionsspannung

Es werden gemäß der energetischen Lage der Kristalldefekte im Substrat mehr positiv geladene Fehlstellen erzeugt als negative, so daß schwach dotiertes n -Material eine Dotierungsinversion erfährt, während bei einer n^+ -Dotierung der Effekt zu klein dafür ist. Definiert man nun die effektive Dotierungskonzentration im Substrat N_{eff} als

$$N_{eff} = N_D - N_A, \quad (6.3)$$

ersetzt in Formel 4.5 N_A durch N_{eff} und benutzt die Näherungen $U_{ext} \gg U_D$ sowie $N_D \gg N_{eff}$, so erhält man für die Depletionsspannung U_{ext} :

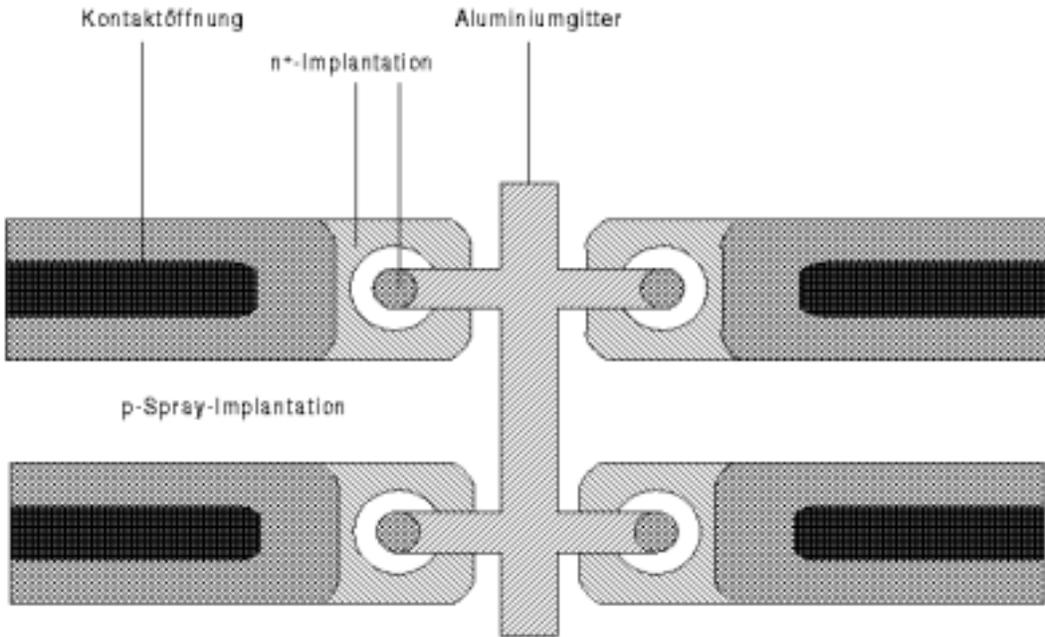
$$U_{ext} = \frac{q d^2}{2\epsilon} N_{eff}. \quad (6.4)$$

Offenbar wächst also die notwendige Depletionsspannung proportional zur effektiven Dotierungskonzentration, so daß man bei ATLAS einen n^+ -auf- n -Detektor benutzt, der während der Betriebszeit langsam zu einem n^+ -auf- p -Detektor konvertiert, um den Maximalwert der Depletionsspannung zu reduzieren.

Damit zu Beginn des Experiments nicht alle Pixel über das n -Substrat durch die Elektronenakkumulationslage miteinander kurzgeschlossen sind, wird vor der n^+ -Implantation

eine gleichmäßige p -Implantation ('p-spray') [39] durchgeführt, die alle Pixel voneinander isoliert. Zur Kontaktierung der Pixel existiert ein Aluminiumgitter, das über eigene n^+ -Kontakte mit dem Detektor verbunden ist. Auf diese Weise kann man mittels 'Punch-Trough-Effekt' die Pixel anschließen und so den Detektor testen (s. Abbildung 21).

Abbildung 21: Layout des Pixeldetektors ('p-spray') [40] mit Kontaktöffnungen, Aluminiumgitter und zugehörigen 'Punch-Through'-Kontakten

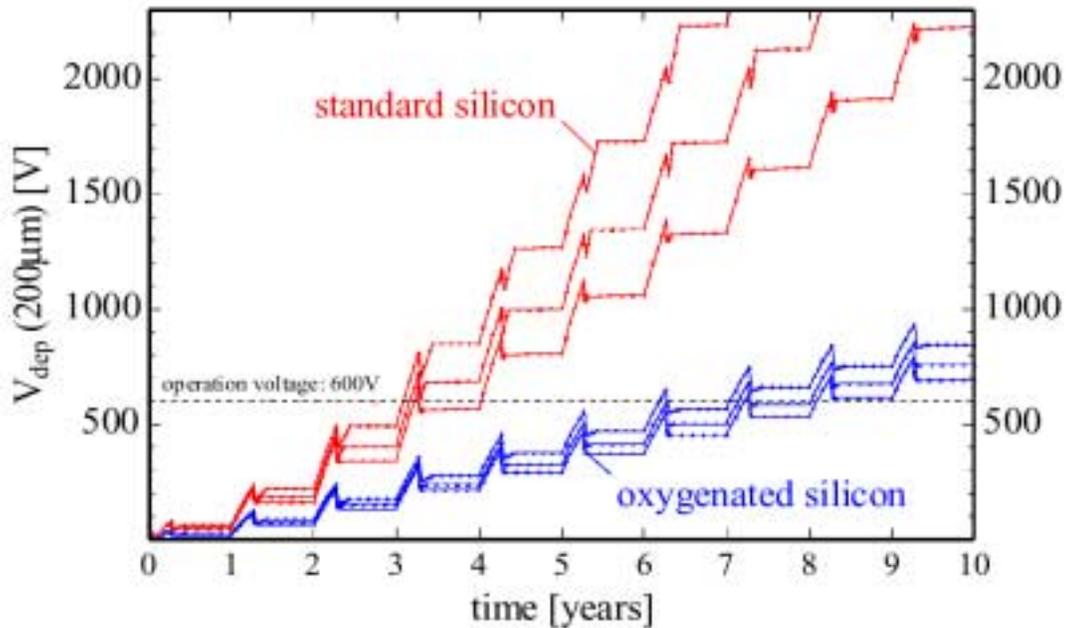


Da eine Erhöhung der Depletionsspannung aber gleichzeitig eine Leckstromerhöhung und damit mehr Kühlaufwand für den Detektor bzw. mehr Leistungsverbrauch bewirkt, stehen bei ATLAS höchstens 600 V zur Verfügung, so daß man letztlich den gesamten Detektor nach etwa 5 Jahren austauschen oder ihn in einem nicht voll depletierten Zustand betreiben müßte. Letzteres hätte natürlich einen dramatischen Einfluß auf die Ladungssammlungseffizienz.

Neuere Untersuchungen der ROSE- und der ATLAS-Kollaboration führen jedoch aus diesem Dilemma hinaus [41, 42, 43]: Bei der Herstellung der Sensoren wird durch Erhöhung des Sauerstoffanteils erreicht, daß die Dotierungsinversion bei Bestrahlung durch geladene Hadronen deutlich langsamer abläuft (a. Abbildung 22), während allerdings bei Neutronenstrahlung keine Verbesserung auftritt.

Da Neutronen hauptsächlich 'Cluster'-Defekte erzeugen, vermutet man, daß der Effekt darauf beruht, daß durch die zusätzlichen Sauerstoffatome donatorartige 'Cluster'-Defekte entstehen, die die unerwünschten Akzeptoren kompensieren, die selbst dann wiederum punktförmig sein müßten [44].

Abbildung 22: Erhöhung der Depletionsspannung beim ATLAS Pixelsensor für Standard Silizium und mit Sauerstoff angereichertem Silizium für drei unterschiedliche Aufwärm-szenarien



Wie schon erwähnt, erzeugen die positiven Oxidladungen an der Grenzschicht zwischen Siliziumsubstrat und -dioxid die der Depletionsspannung entgegengerichtete Spannung U_{Ox} , die bei einer Oxiddicke von $d_{Ox} = 220 \text{ nm}$ in der Größenordnung von 25 V ihren Sättigungswert besitzt [38]. Dieser Effekt führt damit zu einer weiteren Erhöhung der Depletionsspannung, ist jedoch nicht so signifikant wie die Dotierungsinversion.

6.4.2 Leckstromerhöhung

Zwei verschiedene Arten von Leckströmen werden strahlenbedingt vergrößert:

- **Volumenleckstrom**

Die Erzeugung von Generationszentren im Substrat bewirkt eine Erhöhung des Volumenleckstroms, weil sich die Eigenleitung des Siliziums vergrößert. Dieses Problem wird behoben, indem man den Detektor beim Einsatz in ATLAS auf etwa -7° C abkühlt und die Ausleseelektronik so konzipiert, daß sie den verbleibenden

Leckstrom kompensieren kann. Hierbei wird man Werte bis zu etwa 50 nA verkraften können. Der gemessene Volumenleckstrom I_L [21] ergibt sich nach

$$I_L = \alpha \Phi, \quad \alpha = 2 \times 10^{-17} \text{ A cm}^{-1}. \quad (6.5)$$

- **Oberflächenleckstrom**

Die positiven Oxidladungen induzieren eine leitende Elektronenakkumulationsschicht direkt unter der Grenzfläche. Zusammen mit den Generationszentren, die durch die Grenzflächendefekte gebildet werden, entstehen auf diese Weise Oberflächenleckströme, die in der Größenordnung von etwa $2 \mu\text{A cm}^{-2}$ liegen [38]. Über diese Effekte weiß man gegenwärtig noch zu wenig, um prozeßtechnische Maßnahmen bei der Detektorherstellung ergreifen zu können.

6.4.3 Ladungssammlungsineffizienzen

Die Generation von Rekombinationszentren im Substrat nach Bestrahlung führt dazu, daß neben den termisch generierten Ladungen (Leckstrom) auch zu detektierende Ladungen 'getrappt' werden. Die Folge ist eine unerwünschte Verringerung der Signalamplituden um bis zu 20% [44]. Dieses Problem muß wiederum mithilfe der Ausleseelektronik abgefangen werden, indem man die Signale entsprechend verstärkt.

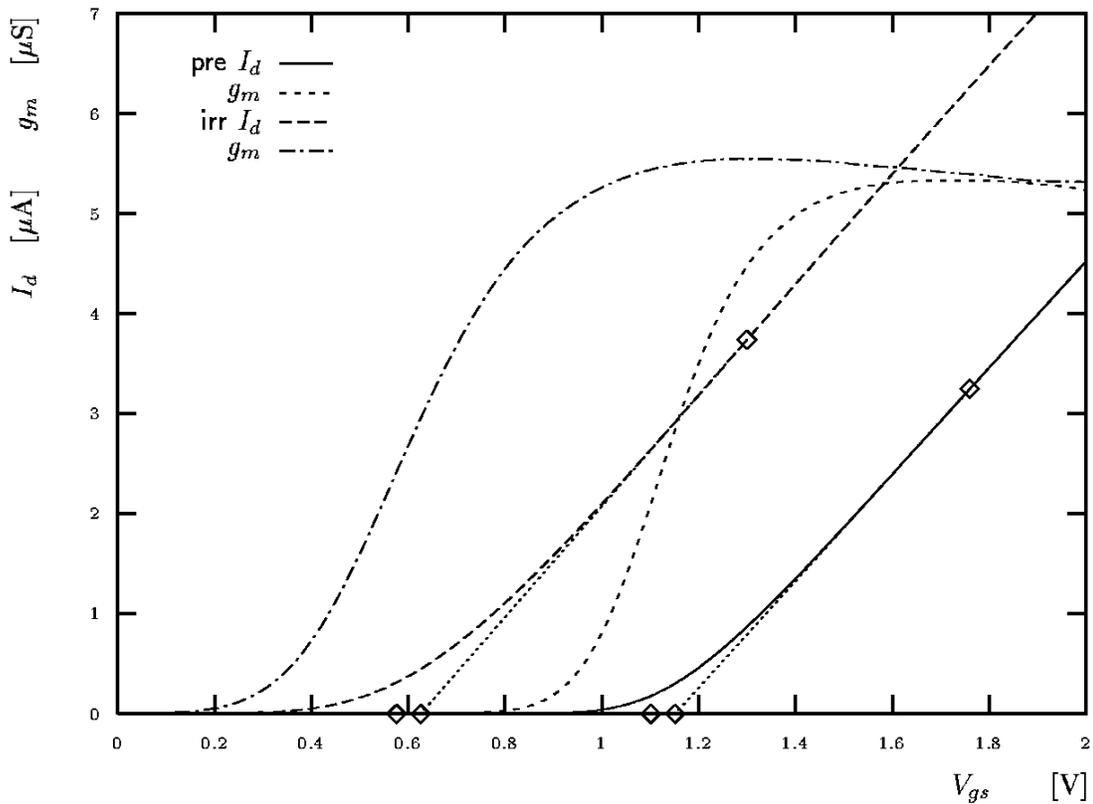
6.5 Auswirkungen auf die Elektronik

Die aufgezählten Strahlenschäden haben für MOSFETs einige wichtige Konsequenzen, die nachfolgend beschrieben werden. Eine Lösungsalternative für die auftretenden Probleme bietet die Verwendung des strahlenharten DMILL Prozesses, dessen Besonderheiten in Kapitel 7 aufgeführt werden.

6.5.1 Verschiebung der Schwellenspannung

Durch die positiven Oxidladungen werden negative Ladungen aus dem darunter liegenden Substrat an die Grenzschicht gesaugt, so daß ein n -leitender Kanal entsteht. Das bedeutet für alle MOSFETs zunächst, daß sich der Wert der Schwellenspannung V_{TH} zu kleineren Werten hin verschiebt, weil man noch die entgegen gerichtete Spannung U_{Ox} gemäß Formel 6.2 hinzuaddieren muß (s. Abbildung 23).

Abbildung 23: Schwellenverschiebung und Transkonduktanz g_m eines NMOS vor und nach Bestrahlung [45]



Bei länger andauernder Bestrahlung tritt neben der durch die Oxidladungen induzierten auch eine durch die Grenzflächenzustände verursachte Verschiebung der Schwellenspannung U_{Gr} auf:

$$U_{Gr} = \frac{q N_{Gr}}{C_{Ox}}, \quad (6.6)$$

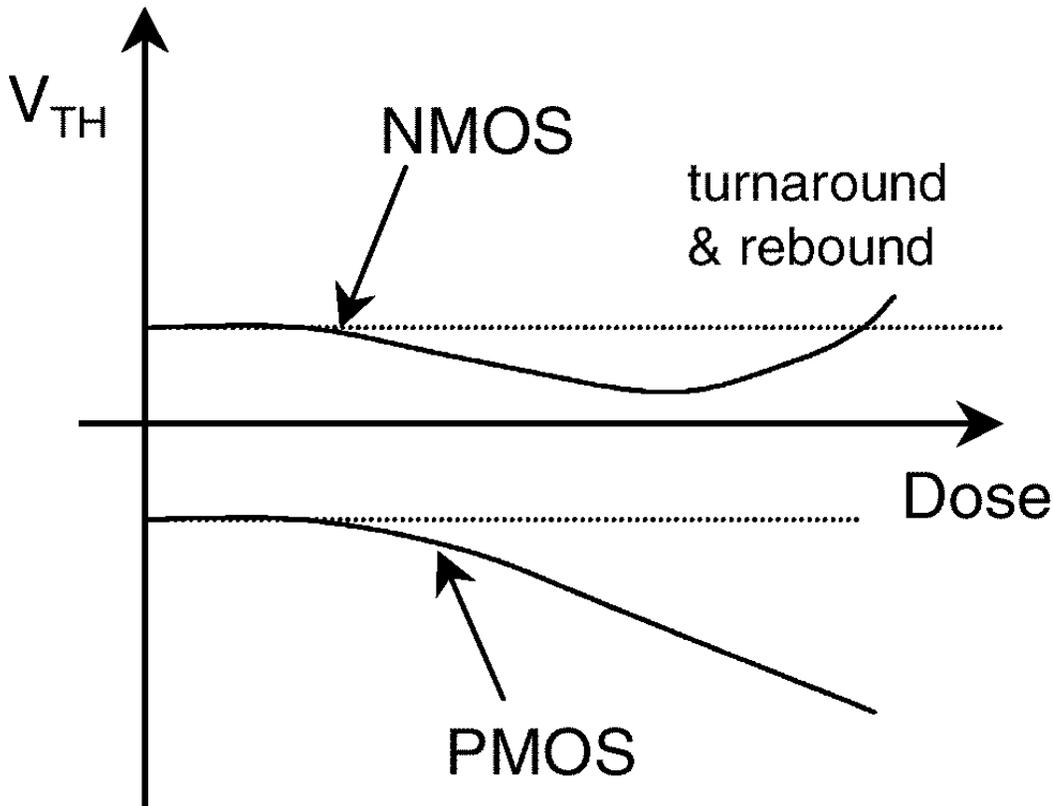
mit

$$N_{Gr} = \text{Anzahl der Grenzflächenladungen/Fläche.} \quad (6.7)$$

Da nach Abschnitt 6.3 die Fehlstellen an der Grenzschicht in allen Energiebereichen auftreten, 'trappen' sie im wesentlichen aufgrund ihrer hohen Anzahl die Majoritätsladungsträger aus dem Siliziumsubstrat, so daß sich ein p -dotiertes Gebiet negativ und ein n -dotiertes Gebiet positiv auflädt. Als Konsequenz zeigt sich beim NMOS ein kompensie-

render Effekt, der bei Bestrahlungsdosen jenseits des Sättigungsbereichs der Oxidladungen eine Erhöhung der Schwellenspannung zeigt ('turnaround'). Es können sogar Werte für V_{TH} erreicht werden, die über denjenigen vor Bestrahlung liegen ('rebound') [36].

Abbildung 24: Schematische Darstellung der Schwellenverschiebung als Funktion der Strahlendosis



Beim PMOS hingegen addieren sich beide Effekte, so daß sich die Schwellenspannung immer zu negativeren Werten hin verschiebt. D. h. ein PMOS wird nach Bestrahlung immer schwächer.

6.5.2 Leckstromerhöhung

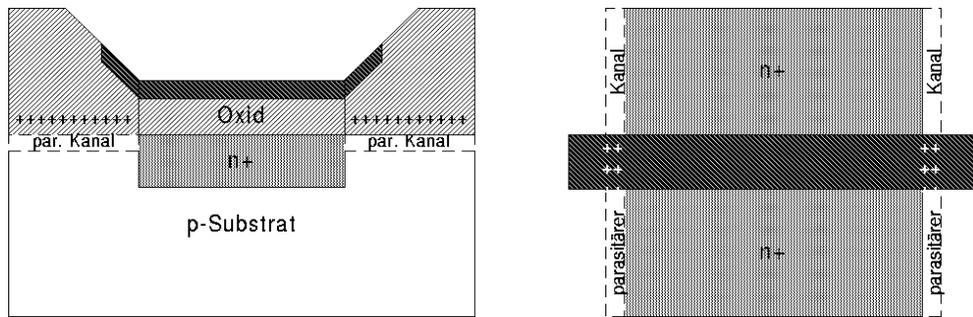
Auch im Elektronikchip werden durch Bestrahlung höhere Leckströme erzeugt, allerdings handelt es sich hierbei um völlig andere Mechanismen, die ausschließlich durch die Generation der Oxidladungen begründet sind. Im Prinzip läßt sich das Phänomen genau wie

die Verschiebung der Schwellenspannung durch U_{Ox} deuten: Die Oxidladungen saugen Elektronen unter die Grenzfläche und erzeugen dort einen leitenden Kanal. Nach Formel 6.2 hängt dieser Vorgang entscheidend von der Dicke des Oxids ab, und da das als Passivierung dienende Feldoxid wesentlich dicker als das Gateoxid ist, fällt der Effekt dort besonders ins Gewicht. Auf diese Weise können sich unerwünschte parasitäre NMOS Transistoren bilden, die als Leckstrompfade fungieren. Dafür existieren zwei Konfigurationen:

- **Einzelner NMOS Transistor**

An den Querrändern eines einzelnen NMOS 'Gates' öffnet sich aufgrund der Verdickung des Oxids ein Kanal, der praktisch außen herum den Strom von 'Source' nach 'Drain' leitet (s. Abbildung 25). Die Folge ist, daß sich der NMOS nicht mehr vollständig abschalten läßt bzw. eine Sperrstromerhöhung.

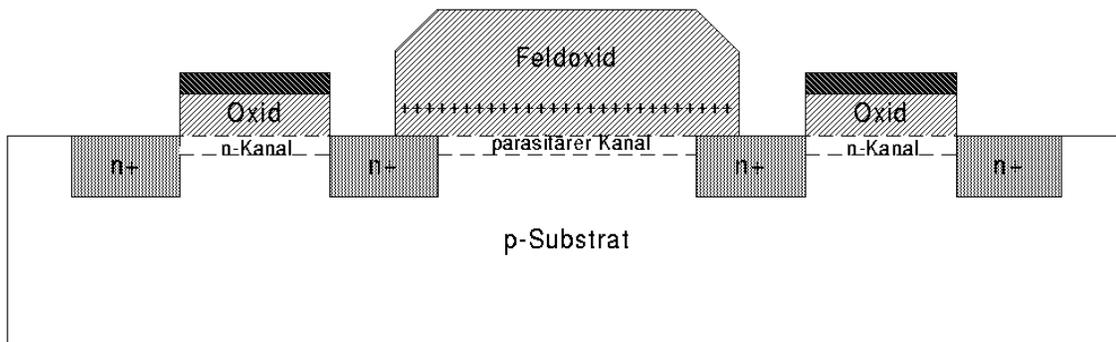
Abbildung 25: Der Leckstrompfad eines einzelnen NMOS um sein Gate herum, links im Querschnitt, rechts als Aufsicht



- **Zwei NMOS Transistoren**

Zwischen jeweils einem 'Source/Drain' Kontakt von zwei verschiedenen Transistoren bildet sich ein Kanal, der ständig Strom führt und somit gespeicherte Ladungen abtransportiert (s. Abbildung 26). Solche Effekte werden durch ringförmige Strukturen vermieden (s. Kapitel 12).

Abbildung 26: Der Leckstrompfad zwischen zwei NMOS Transistoren



6.5.3 Reduktion der Ladungsträgerbeweglichkeit

Die strahlenbedingten Grenzflächen-’Traps’ bewirken eine Reduktion der Beweglichkeit der Ladungsträger im Transistorkanal gemäß [46]

$$\mu = \frac{\mu_0}{1 + \alpha(N_{Gr})}, \quad (6.8)$$

mit

$$\begin{aligned} \mu_0 &= \text{Beweglichkeit vor Bestrahlung} \\ \alpha &= (8 \pm 2) \times 10^{-13} \text{ cm}^2. \end{aligned}$$

Diese hat sowohl eine Verringerung seiner Leitfähigkeit als auch der Transkonduktanz zur Folge. Die Konsequenz ist eine Verringerung des Drainstroms bzw. eine Verlangsamung der Schaltgeschwindigkeit des Transistors, was beim Chipdesign in Form von stärker ausgelegten Bauteilen (größeres W/L) zu beheben ist.

6.5.4 Erhöhung des Rauschens

Die Substratdefekte, die Oxidladungen und die Grenzflächenzustände bilden allesamt durch die Bildung von Energieniveaus in der Siliziumbandlücke Zentren, an denen statistisch Ladungsträger generiert oder absorbiert werden, also eine frequenzunabhängige Rauschquelle (weißes Rauschen). Auch das sogenannte frequenzabhängige $1/f$ -Rauschen, welches meist auf der Wechselwirkung der freien Ladungsträger im Transistorkanal mit Fehlstellen an der Grenzfläche beruht, vergrößert sich ebenfalls aus diesem Grund.

Selbstverständlich treten diese Effekte lokal auch im Sensor auf, mitteln sich jedoch bei einer Integration über die gesamte Detektorfläche zu einem konstanten Leckstrom.

6.5.5 SEU und SEL

Ionisierende Strahlung kann direkt auf Leiterbahnen oder Orte auftreffen, die gespeicherte Ladungen enthalten, und dort die entsprechenden Knoten entladen. An diesen Stellen

kann diese Information verloren gehen und somit den digitalen Wert invertieren. Dieses Phänomen nennt man SEU (**S**ingle **E**vent **U**pset).

Weiterhin besteht die Möglichkeit, daß durch Bestrahlung Strompfade zwischen Versorgungskontakten entstehen und auf diese Weise Schaltkreise durchbrennen [47]. Dieser Vorgang ist als SEL (**S**ingle **E**vent **L**atchup) bekannt.

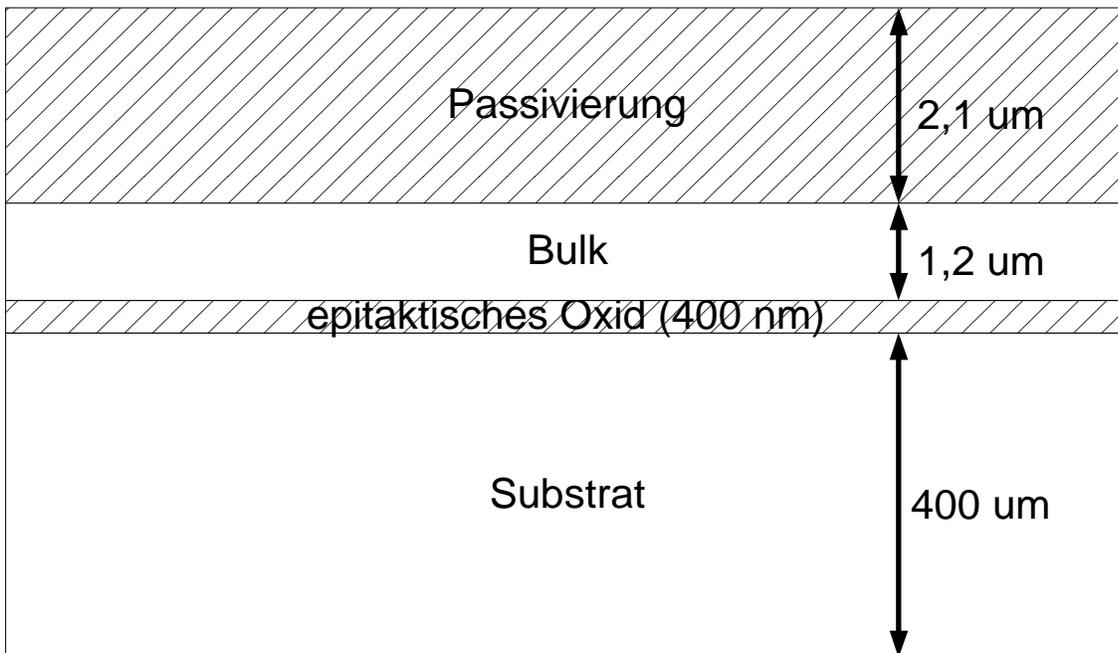
7 Besonderheiten des DMILL Prozesses

Zur Vorbeugung vor den in Kapitel 6 dargelegten Strahlenschädigungen an der Elektronik empfiehlt es sich, bei der Herstellung eine Technologie zu wählen, die intrinsisch die Strahlenhärte gewährleistet. Einer der wenigen Prozesse, die diese Forderung erfüllen, ist DMILL (**D**urci **M**ixte **I**solant **L**ogico **L**inéaire) [48]. Da es sich dabei um eine Entwicklung aus der Militärforschung handelt, die einer gewissen Geheimhaltung unterliegt, sind nur einige wenige Details über DMILL zugänglich.

DMILL ist ein $0,8 \mu\text{m}$ BiCMOS Prozeß mit einer Polysiliziumlage und zwei Aluminiumlagen für Schaltungsverbindungen. Die Herstellerfirma Atmel (ehemals Temic) garantiert eine Resistenz der Elektronik bis zu einer Strahlendosis von 10 Mrad [49]. Die Prozessierung erfordert 24 Maskenschritte [50] im Vergleich zu 15 [51] bei dem vergleichbaren, nicht strahlenharten $0,8 \mu\text{m}$ AMS Prozeß, was schon ein Indiz dafür ist, daß einige Unterschiede bestehen, über die die Firma keine Auskünfte erteilt. Vermutlich handelt es sich dabei um besonders schonende, eventuell sogar ausheilende Maßnahmen, die beispielsweise die Kristallfehlstellen im Siliziumsubstrat und an der Grenzschicht zum Siliziumdioxid minimieren. Dabei handelt es sich jedoch lediglich um (höchstwahrscheinlich richtige) Spekulationen, während die gesicherten Erkenntnisse und Besonderheiten von DMILL sich wie folgt darstellen:

- **SOI Prozeß**

Abbildung 27: Querschnitt durch die SOI-Struktur bei DMILL

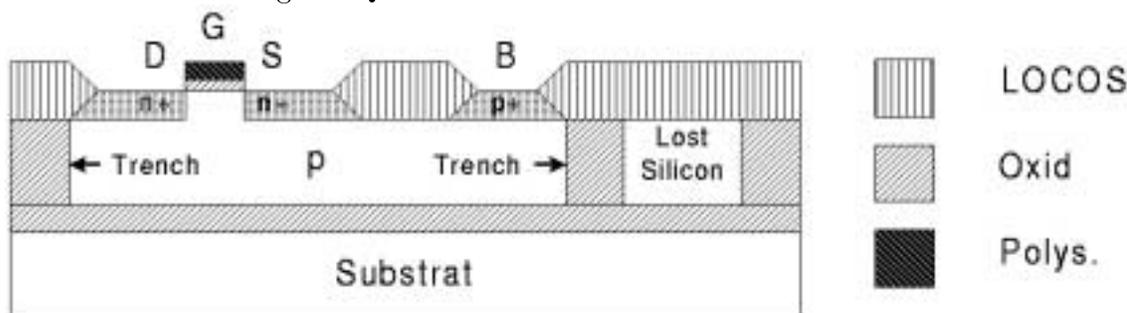


Bei DMILL handelt es sich um einen sogenannten SOI (**Silicon On Insulator**) Prozeß. Das bedeutet, daß das $400\ \mu\text{m}$ dicke Siliziumsubstrat durch eine durchgehende isolierende, epitaktische Siliziumdioxidschicht mit einer Dicke von $400\ \text{nm}$ von den eigentlichen elektronischen Bausteinen, die sich in der obersten $1,2\ \mu\text{m}$ dicken Schicht befinden, abgetrennt wird [52]. Dadurch werden sowohl vor als auch nach Bestrahlung alle Substratschäden eliminiert und damit auch größtenteils das Rauschen und die Volumenleckströme. Nur noch die Anteile, die in dem räumlich wenig ausgedehnten Transistor-’Bulk’ entstehen, beeinflussen dessen Verhalten.

Zusätzlich bleibt noch anzumerken, daß wegen der hohen Dotierungen hier keine Konversion von n nach p stattfindet.

- **’Trenches’**

Abbildung 28: Querschnitt durch einen NMOS mit ’Trenches’



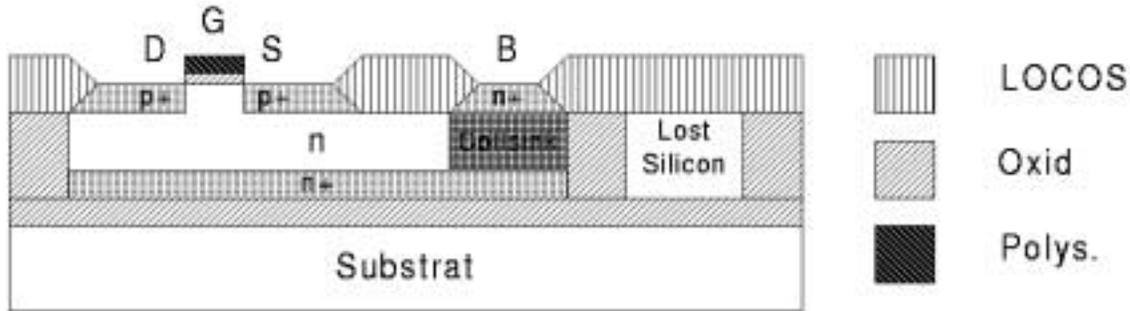
In der Schicht, in der die aktiven Elemente liegen, werden die PMOS Transistoren von den NMOS Transistoren durch tiefe Siliziumdioxidgräben (’Trenches’), die bis zur epitaktischen Oxidschicht reichen, vollständig galvanisch getrennt [53]. Dadurch wird jeglicher ’Latch-Up’ vollständig unterbunden. Bei Bedarf kann man auch besonders kritische Transistoren in analogen Schaltungen einzeln mit ’Trenches’ abtrennen und auf diese Weise zum Beispiel die Leckströme zwischen zwei verschiedenen NMOS verhindern oder möglichst viele Rauschquellen eliminieren. Zwischen zwei ’Trenches’, die einen vorgeschriebenen Abstand von $2\ \mu\text{m}$ aufweisen, befindet sich das sogenannte ’Lost Silicon’, das separat kontaktiert wird und damit eine Barriere für kapazitives Übersprechen in Form einer Äquipotentialfläche bildet.

- **Guardringe**

DMILL stellt zur Erzeugung von Guardringen, die die isolierende Funktion der ’Trenches’ unterstützen, eine besondere Prozeßlage zur Verfügung. Es handelt sich hierbei um eine hoch n^+ -dotierte Schicht (’Collsink’), die im Gegensatz zur oberflächlichen Implantation für ’Source’ bzw. ’Drain’ eines NMOS in der aktiven Schicht den Raum bis hinunter zum epitaktischen Oxid ausfüllt (s. Abbildung 29). Dadurch erhält man wegen der besseren Leitfähigkeit eine ausgeprägtere Potentialwand als mit dem ’Lost Silicon’. Dieses ’Collsink’ kann auch für die Kontaktierung eines PMOS ’Bulks’ benutzt werden, um diesen aus der Tiefe besser auf das Versorgungspotential zu bringen [53].

- LOCOS

Abbildung 29: Querschnitt durch einen PMOS mit 'Collsink' und LOCOS



Der transistornahe Teil des Feldoxids besteht aus einem besonderen Oxid, in das bei der Herstellung Phosphor- oder Arsenbeimischungen erhält (LOCOS) [48]. Diese Implantationen wirken als Donatoren und neutralisieren strahleninduzierte Oxidladungen. Auf diese Weise wird die Bildung beider in Kapitel 6 beschriebenen Konfigurationen von parasitären NMOS Transistoren unterdrückt.

- Gateoxid

Der entscheidende Punkt bei der Schwellenverschiebung der Transistoren ist die Anzahl der Oxidladungen, die proportional zur Dicke des Gateoxids wächst. Bei DMILL beträgt diese nur $17,5 \text{ nm}$ [54]. Als Resultat wird bei DMILL nach einer Strahlendosis von 10 Mrad lediglich eine Schwellenverschiebung von -150 mV beim NMOS bzw. -250 mV beim PMOS erzielt [48].

- Layout

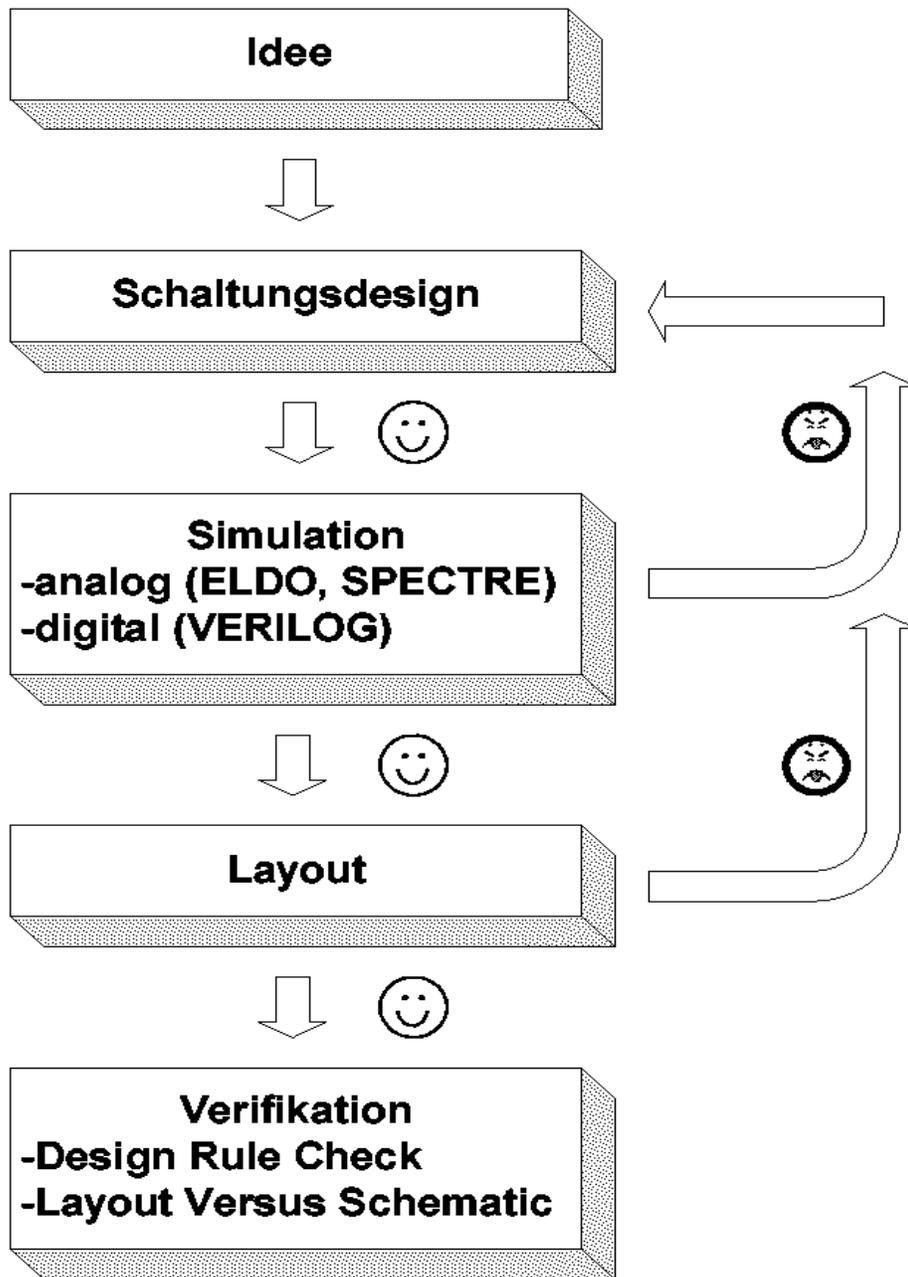
Bei den herkömmlichen Prozessen, die auf einem schwach p -dotierten Substrat basieren, zeigt sich durchweg das Phänomen, daß ein PMOS Transistor minimaler Größe den zwei- bis dreifachen Platz eines entsprechenden NMOS benötigt. Die Ursache dafür liegt in der Notwendigkeit einer zusätzlichen n -dotierten Wanne, die den PMOS 'Bulk' bildet. Bei DMILL entfällt jedoch aufgrund der Abtrennung durch die 'Trenches' letzterer Schritt, weil die Dotierung des 'Bulks' direkt in der richtigen Polarität erfolgt.

Zusätzlich gibt es noch eine layouttechnische Maßnahme, die den Leckstrom eines einzelnen Transistors am Rand seines 'Gates' verringert: Das 'Gate' mit dem darunter liegenden dünnen Oxid wird weit nach außen über das Ende des Inversionskanals hinaus gezogen. Dadurch wird einerseits das Volumen, in dem sich strahleninduzierte Oxidladungen in der Nähe des Kanals anlagern können, drastisch verringert und andererseits das bis zu $2,1 \mu\text{m}$ dicke LOCOS weit vom Kanal entfernt.

Die beiden oben genannten Besonderheiten führen dazu, daß bei DMILL die Verhältnisse umgekehrt liegen: Ein minimaler NMOS braucht im Layout etwa die zweieinhalbfache Fläche eines minimalen PMOS [55]. Aus diesem Grund wird im Design des FED an allen Stellen, an denen es möglich ist, ein PMOS bevorzugt.

8 Das Design des FED1

Abbildung 31: Flußdiagramm zum Ablauf von der Idee bis hin zum Design



Generell stellt das Chipdesign einen iterativen Prozeß dar, der in folgenden Schritten 'Workstation'-basierend abläuft: Zunächst wird die Idee oder die Anforderung in einen elektronischen Schaltplan umgesetzt. Dieser wird dann zur Unterstützung der vorherigen Überlegungen simuliert, entweder mit einem analogen Simulator (z.B. ELDO oder SPECTRE), falls die Schaltung nicht zu groß ist, oder bei komplexen Blöcken mit einem digitalen Simulator (z.B. VERILOG), um die anfallenden Datenmengen und Rechenzeit in einem erträglichen Umfang zu halten. Fördert die Simulation vorher nicht berücksichtigte Probleme zu Tage, muß der Schaltplan neu überdacht und modifiziert werden. Bei

einem zufriedenstellenden Ergebnis nach der Simulation erstellt man anschließend das Layout, das quasi eine Anweisung für den Hersteller bildet, wo welches Bauteil physikalisch platziert werden soll. Auch hier müssen in der Regel bestimmte Anforderungen erfüllt werden, z.B. ein vorgegebenes Platzkontingent. Ist dies nicht durch effizientere Raumaufteilung möglich, muß wiederum der Schaltplan verändert werden, indem man Transistoren oder ganze Gatter einspart bzw. ein völlig neues Konzept erstellt. Genügt das Design nun allen Aspekten, werden im letzten Schritt die Verifikationsprogramme durchgeführt. Dazu gehört einerseits die Überprüfung der vom Hersteller gelieferten Designregeln, die genau alle Mindestgrößen und -abstände der Bauteile vorschreiben. Zum andern gibt es noch ein Programm, das aus dem Layout alle elektronischen Elemente extrahiert und mit dem Schaltplan vergleicht. Auf diese Weise wird sichergestellt, daß sich Layout und Schaltplan exakt entsprechen.

Grundsätzlich ist anzumerken, daß beim Design des FED1 stets ein sorgfältiges Abwägen zwischen Platzangebot (Größe der Pixelzelle $50 \mu m \times 400 \mu m$) und Funktionalität erfolgen muß und man ständig an die Grenzen des Machbaren vorstößt. Besonders die geringe Pixelgröße stellt eine enorme Herausforderung dar und gibt Anlaß zu Kompromissen bezüglich der Verwendung dynamischer Schaltungen im Ausleseteil.

Beim FED1 handelt es sich um ein sehr busintensives Design, in dem mehr als 90 Signale vom Chiprand in die Pixelmatrix geführt werden müssen, was auf den ersten Blick bei einer Technologie mit nur zwei Metallagen ein bedeutender Nachteil zu sein scheint. Erschwerend kommt hinzu, daß gerade im besonders platzkritischen Matrixteil die zweite Metallage zur Abschirmung gegen den Detektor benutzt werden muß, um eine kapazitive Kopplung in denselben zu minimieren. Als Verbindungslagen können also nur die beiden übrigen benutzt werden. Allerdings sind auf dem Chip mehr als 720000 Transistoren bei einer Größe von $7,4 mm \times 11 mm$ realisiert, was einer Belegungsichte von etwa $9000/mm^2$ entspricht. Um die hier vollbrachte Leistung einschätzen zu können, sei zum Vergleich die Belegungsichte des MCC, der ebenfalls in DMILL designed wurde, genannt: Sie beträgt $3500/mm^2$.

Ein Blick in die Schaltpläne und ins Layout zeigt, daß an allen Stellen, wo es möglich und notwendig ist, Transistoren eingespart werden und Leitungen über aktive Flächen hinweg verlegt werden. Durch wohlüberlegten Einsatz der Verbindungslagen und Aufteilung der Topologie kann so der zur Verfügung stehende Platz optimal genutzt werden und sogar die schon erwähnten, angeblich 'Yield' verbessernden Layoutregeln des Herstellers, erfüllt werden.

8.1 Anforderungen an den FED1

Aus den physikalischen Fragestellungen, die Atlas untersuchen soll, und den vom LHC vorgegebenen Betriebsparametern ergeben sich eine Reihe von Anforderungen, denen die Front-End-Elektronik genügen muß:

- **Pixelgröße**

Wegen der geforderten Ortsauflösung σ von $12 \mu m$ in $r\phi$ -richtung und $80 \mu m$ in z -Richtung ergibt sich bei binärer Auslese und Gleichverteilung der Treffer eine Pixelgröße von $50 \mu m \times 300 \mu m$, gemäß der Formel

$$\sigma = \frac{d}{\sqrt{12}}, \quad (8.1)$$

wobei d für die Kantenlänge des Pixels steht. Um die notwendige Logik unterbringen zu können, wird jedoch eine Pixellänge von $400 \mu m$ gewählt, die bei grober analoger Information immer noch die geforderte Ortsauflösung gewährleistet.

- **Nachweiseffizienz**

Um eine möglichst genaue Spurrekonstruktion zu erreichen, sollte der gesamte Pixel-detektor ≥ 3 Spurpunkte pro Teilchenbahn erzeugen, was in einer Nachweiseffizienz von mehr als 95 % pro Pixellage resultiert.

- **Schwelle**

Nach Bestrahlung werden in der nur $200 \mu m$ dünnen B-Lage beim Durchgang eines MIPs nur etwa $6000 e^-$ an Ladungssignal erzeugt. Bei einer eventuellen Ladungsaufteilung auf mehrere Pixel muß der FED1 also bei einer Schwelle von $2000 e^-$ betrieben werden können, d.h. solche Signale noch als Treffer erkennen können.

- **Schwellenvariation**

Das Rauschen im Analogteil der Pixelzellen führt zu einem veränderten Ansprechverhalten der Front-End-Elektronik, was sich letztlich in einer Verschiebung der Schwellenwerte zeigt. Dieses Rauschen soll im Pixel daher unter $300 e^-$ liegen. Durch fabrikationsbedingte Variationen der Bauteile im Pixel ergeben sich ebenfalls unterschiedliche Schwellenwerte von Pixel zu Pixel, so daß hier eine Dispersion von unter $200 e^-$ über den ganzen Chip gefordert wird. Die beiden Anteile addieren sich quadratisch nach

$$\sigma_{gesamt} = \sqrt{\sigma_{Analog}^2 + \sigma_{Dispersion}^2} \quad (8.2)$$

und ergeben damit eine Gesamtschwellenvariation von $360 e^-$.

- **Übersprechen**

Bei niedriger Schwelleneinstellung kann eine hohe Ladungsdeposition im Detektor dazu führen, daß aufgrund kapazitiver Kopplung zwischen benachbarten Pixeln beide Pixel Treffer anzeigen. Für die Ortsauflösung bedeutet dieses Phänomen eine Verbesserung, allerdings sollte wegen der daraus resultierenden höheren Belegungs-dichte die Übersprechwahrscheinlichkeit unter 5-10 % liegen.

- **Leistungsaufnahme**

Die strahlenbedingten Veränderungen im Sensor und in der Elektronik führen dazu, daß man den Pixeldetektor auf $-7^\circ C$ abkühlen muß, wodurch die Leistungsaufnahme der Elektronik auf $40 \mu W$ pro Pixel limitiert wird.

- **Zeitauflösung**

Da bei LHC die Wechselwirkungsrate von 40 MHz vorliegt, muß der FED1 in der Lage sein, mit einer Auflösung von 25 ns Treffer zuzuordnen. Konkret wird gefordert, daß alle Teilchen mit Energien, die zwischen dem 0,2- und 2-fachen eines MIPs liegen, in einem festen Zeitfenster von 25 ns registriert werden können.

- **Totzeit**

Zur Reduktion der Effizienzverluste im Pixel auf etwa 1% muß bei einer Ereignisrate von 40 MHz und einer, durch Monte Carlo Simulationen bestimmten, Trefferwahrscheinlichkeit von 10^{-4} pro Pixel die Totzeit niedriger als $2,5\text{ }\mu\text{s}$ in den beiden äußeren Detektorlagen bzw. $0,5\text{ }\mu\text{s}$ in der B-Lage sein.

- **Leckstromkompensation**

Die strahlenbedingte Leckstromerhöhung im Detektor soll durch den FED1 bis zu einer Größe von 50 nA kompensiert werden können.

- **Zwischenspeicherung und Nullunterdrückung**

Alle Treffer müssen $2 - 3\text{ }\mu\text{s}$ lang auf dem Chip zwischengespeichert werden, bis der Level 1 Trigger das Vorliegen eines physikalisch interessanten Ereignisses meldet. Nur die zu dieser Wechselwirkung gehörenden Treffer sollen letztlich aus dem FED1 ausgelesen werden, während alle anderen gelöscht werden.

- **Strahlenhärte**

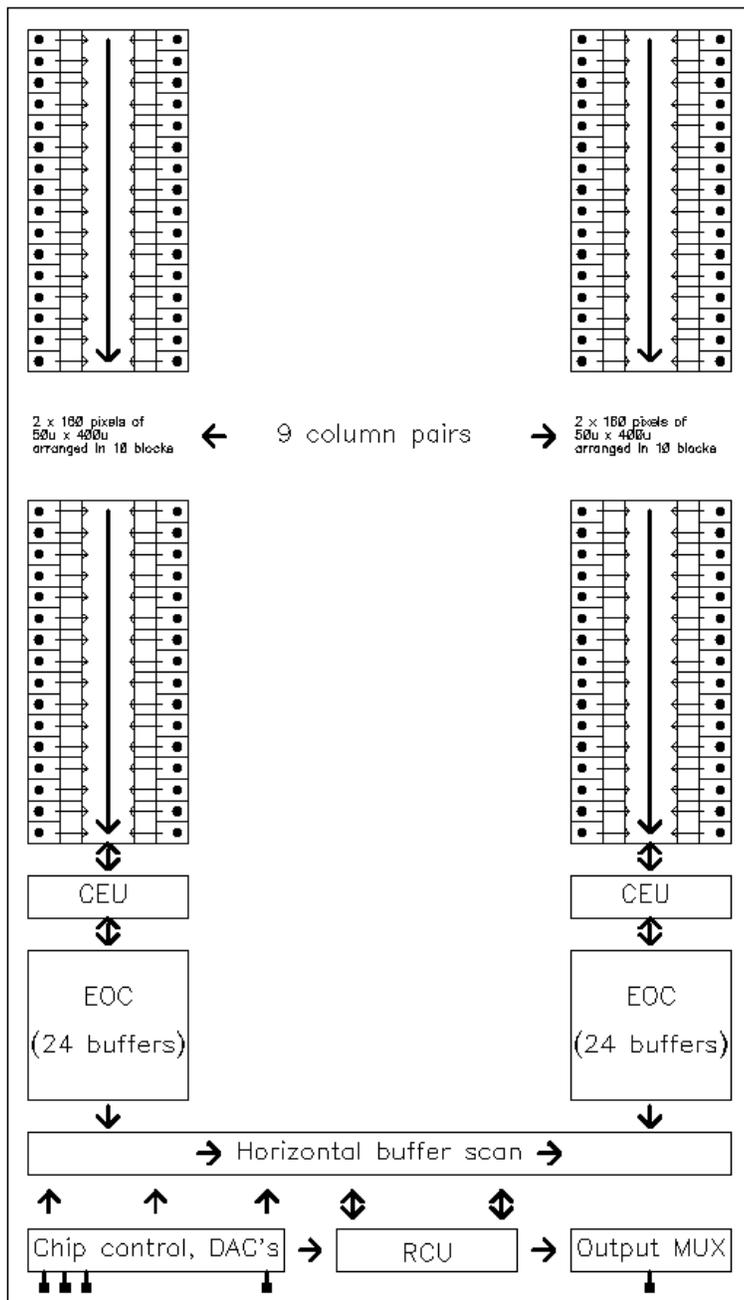
Während einer Betriebszeit von 10 Jahren erwartet man bei ATLAS eine Strahlendosis von $30 - 40\text{ Mrad}$ in den beiden äußeren Detektorlagen bzw. von 50 Mrad in der B-Lage, so daß die Front-End-Elektronik in strahlenharter Technologie hergestellt werden muß.

8.2 Aufbau und Funktionsweise des FED1

8.2.1 Übersicht

Der FED1 Chip besteht aus einer Matrix mit 160 Zeilen und 18 Spalten von Pixelzellen einer Größe von jeweils $50\text{ }\mu\text{m} \times 400\text{ }\mu\text{m}$. Die Pixel sind in neun um die Längsachse spiegel-symmetrischen Doppelspalten angeordnet mit jeweils zehn Blöcken à 2×16 Pixeln. Jeder dieser Pixel enthält folgende Schaltungsblöcke: einen Analogteil, einen Pixelkontrollteil und einen Ausleseteil.

Abbildung 32: Blockschaltbild von FED1



Unterhalb jedes Spaltenpaares folgt die sogenannte CEU (Column **E**nd of Column Arbitration Unit), die die Signale erzeugt, die zur Auslese getroffener Pixel benötigt werden, und die Kommunikation zur darunter liegenden EOC (**E**nd Of Column Logic) herstellt. Die EOC dient der Zwischenspeicherung der Treffer auf dem Chip, bis nach einer im ATLAS Experiment festen Wartezeit ('Latency') von etwa $2 - 3 \mu s$ ein Level 1 Signal das Vorliegen eines interessanten Ereignisses meldet. Alle zu diesem Trigger gehörigen Treffer werden ausgelesen, während Treffer, die nach der 'Latency' keinen Trigger erhalten, sofort gelöscht werden (Nullunterdrückung).

Alle Ausleseaktionen werden durch die einmal pro Chip vorhandene RCU (**R**eadout

Control Unit) gesteuert, die ihrerseits aus mehreren Blöcken besteht, welche für die unterschiedlichen Prozeduren verantwortlich zeichnen.

Zur Versorgung des Chips werden insgesamt zwei analoge Spannungsversorgungen (VDDA und VCCA), eine digitale Spannungsversorgung (VDDD) und zwei Massen (GNDA und GND) benötigt, die alle voneinander getrennt verkabelt sind. Zusätzlich sind für den Betrieb von FED1 7 Ströme und zwei Spannungen notwendig, die allesamt durch DACs (**D**igital **A**nalog **C**onverter) direkt vom FED1 erzeugt werden.

Der FED1 Chip besitzt insgesamt 48 Eingangs- bzw. Ausgangspads, insbesondere einen Dateneingang zum Schreiben und zwei Datenausgänge zum Lesen:

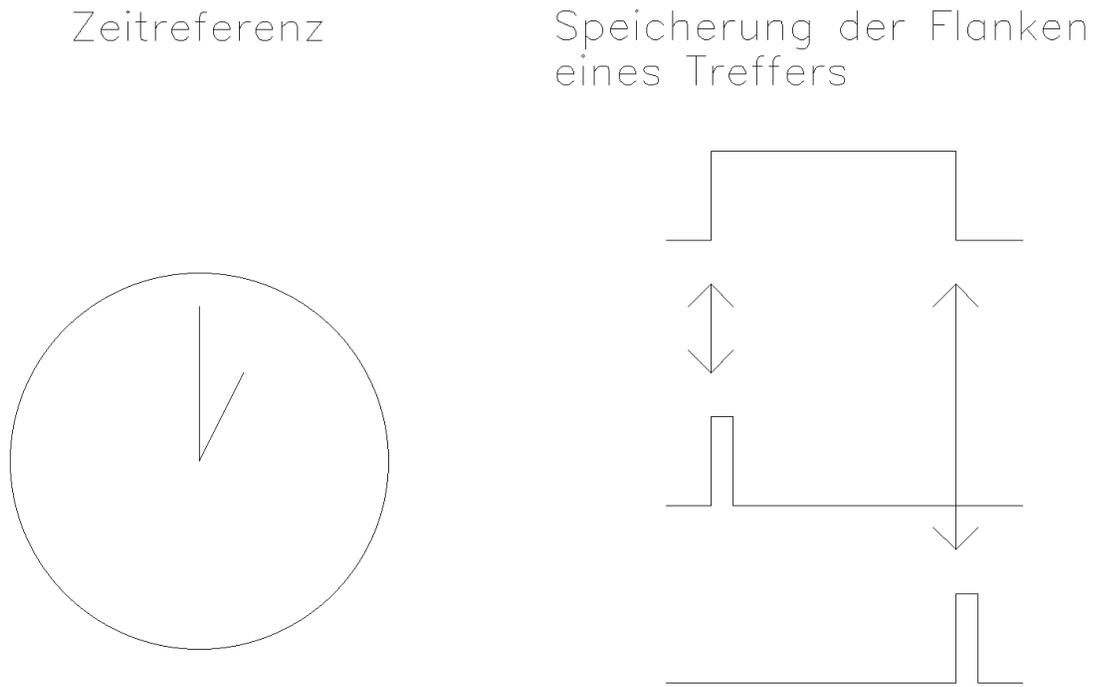
Tabelle 1: Pinbelegung des FED1

Pin	Signal	Bemerkung	Pin	Signal	Bemerkung
1	Guard	Detektorguardring	25	VCCD	Schwelle Diskriminator
2	CCK	5 MHz Kommandoclock	26	-	nicht benutzt
3	CI	Dateneingang	27	ICH	Chopperstrom
4	LD	Laden des Chips	28	ITR	Feinjustierung Diskriminator
5	STR(-)	Injektion, diff.	29	ID	Diskriminatorstrom
6	STR(+)	Injektion, diff.	30	IF	Feedbackstrom
7	GA0	Chipadresse, Bit0	31	IP	Verstärkerbiasstrom
8	GA1	Chipadresse, Bit1	32	IL	Verstärkerlaststrom
9	GA2	Chipadresse, Bit2	33	IPS	Sourcefolgerstrom
10	GA3	Chipadresse, Bit3	34	VDDA	analoge Versorgung (3 V)
11	VDDA	s. Pin 34	35	VCCA	analoge Versorgung (1,5 V)
12	VCCA	s. Pin 35	36	GNDA	analoge Versorgung (0 V)
13	GNDA	s. Pin 36	37	GND	digitale Versorgung (0 V)
14	-	nicht benutzt	38	VDDD	digitale Versorgung (3 V)
15	GND	s. Pin 37	39	RSTb	Chipreset, aktive low
16	VDDD	s. Pin 38	40	DO(-)	Datenausgang, diff.
17	XCK(-)	40 MHz Clock, diff.	41	DO(+)	Datenausgang, diff.
18	XCK(+)	40 MHz Clock, diff.	42	HIT(-)	Datenausgang, diff. (Testpin)
19	LV1(-)	Level 1, diff.	43	HIT(+)	Datenausgang, diff. (Testpin)
20	LV1(+)	Level 1, diff.	44	SEL(-)	Chipselektion, diff. (Testpin)
21	SYNC(-)	Triggerreset, diff.	45	SEL(+)	Chipselektion, diff. (Testpin)
22	SYNC(+)	Triggerreset, diff.	46	AMP	Verstärkerausgang (Testpin)
23	VCAL	analoge Injektion	47	REF	Referenzstrom (Testpin)
24	VTH	Schwelle Diskriminator	48	MESH	Detektorgitter

8.2.2 Die 'Timestamp'-Architektur

Die Auslesearchitektur von FED1 funktioniert nach folgendem Prinzip: Auf dem Chip gibt es einen 7 Bit 'Gray Code'-Zähler, der ständig mit der ATLAS Kollisionsfrequenz von 40 MHz durchläuft und somit eine Art interne Zeitreferenz mit einer Auflösung von 25 ns zur Verfügung stellt. Diese Zeit wird mittels eines Bussystems an alle Pixel verteilt, so daß Beginn und Ende eines eintretenden Treffers sofort in der jeweiligen Zelle bekannt sind und direkt mit den Signalfanken als zwei 7 Bit 'Gray Code' Zahlen gespeichert werden. Diese beiden Zahlen stellen die Zeitmarken dar, woher auch der Name 'Timestamp'-Architektur rührt.

Abbildung 33: Das Prinzip der 'Timestamp'-Architektur



Die getroffenen Pixel werden, sobald die fallende Flanke eintrifft, durch einen zweistufigen 'Dominoscanner' entdeckt und in der CEU gemeldet, die dann den Auslesezyklus generiert. Dieser besteht aus drei Operationen:

- 'Einfrieren der Spalte'

Damit eventuell neu eintretende Treffer den Auslesezyklus nicht stören, sendet die CEU ein Signal, das die gesamte Spalte blockiert. Neue Treffer werden zwar detektiert, können jedoch nicht den 'Scanner' starten.

- **Auslesen der Pixel**

Die CEU sendet das Signal zum Auslesen des obersten Pixels und transferiert die Zeitmarken und Adresse über ein Bussystem in den obersten freien Platz in der EOC, die insgesamt 24 Speicherplätze pro Spaltenpaar besitzt.

- **Löschen des Pixels**

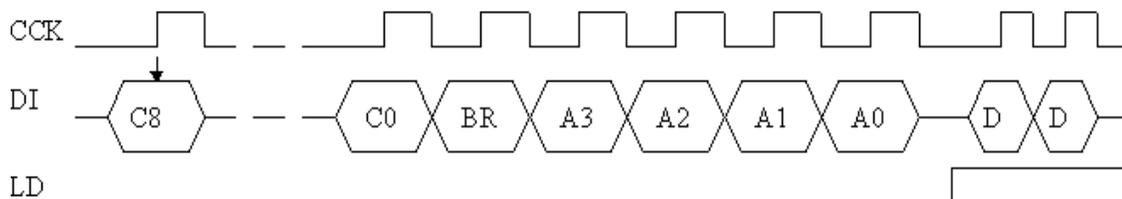
Die Trefferdaten des zuletzt ausgelesenen Pixels werden gelöscht.

Dieser Vorgang wird so lange wiederholt, bis alle zu einem Trigger gehörigen Treffer gefunden wurden. Erst dann wird die 'eingefrorene' Spalte wieder aktiviert. Die Auslesegeschwindigkeit wird über die Taktfrequenz der CEU bestimmt, welche auf 5, 10 oder 20 MHz einstellbar ist.

8.2.3 Das Kommandoregister

Das 15 Bit lange Kommandoregister, das sich im Chipkontrollteil befindet (s. Abbildung 32), ermöglicht die externe Kommunikation mit dem Chip, welche im 'Atlas Pixel Demonstrator' [56] genau spezifiziert ist. Dieses Protokoll benutzt drei Eingangssignale: DI, CCK und LD. Das Übermitteln von Daten läßt sich in zwei Teile aufgliedern: Das Schreiben des eigentlichen Befehls und das Adressieren des Chips. Zunächst ist LD 'aus' und 9 + 1 + 4 Bits werden in der unten dargestellten Reihenfolge an alle 16 Chips auf einem Modul gleichzeitig mit der steigenden Flanke von CCK geschickt. Dabei stehen C8 bis C0 für den Befehl, BR für 'Broadcast' und A3 bis A0 für die Adresse. Nachdem CCK wieder 'aus' ist, wird LD eingeschaltet, und der Chip vergleicht die 4 Adressenbits mit seiner eigenen geographischen Adresse, die fest über GA3 bis GA0 verdrahtet ist. Falls beide Adressen übereinstimmen oder BR eingeschaltet ist, geht SEL an, und der Befehl C8 bis C0 wird in das Kommandoregister geladen. Andernfalls bleibt der Inhalt unverändert bestehen. Mithilfe eines 'Hardresets' (s. Abschnitt 8.2.11) kann das Register gelöscht werden, so daß kein gültiger Befehl vorliegt.

Abbildung 34: Schreiben eines Kommandos



Folgende Kommandos sind im FED1 implementiert:

Tabelle 2: Kommandos des FED1

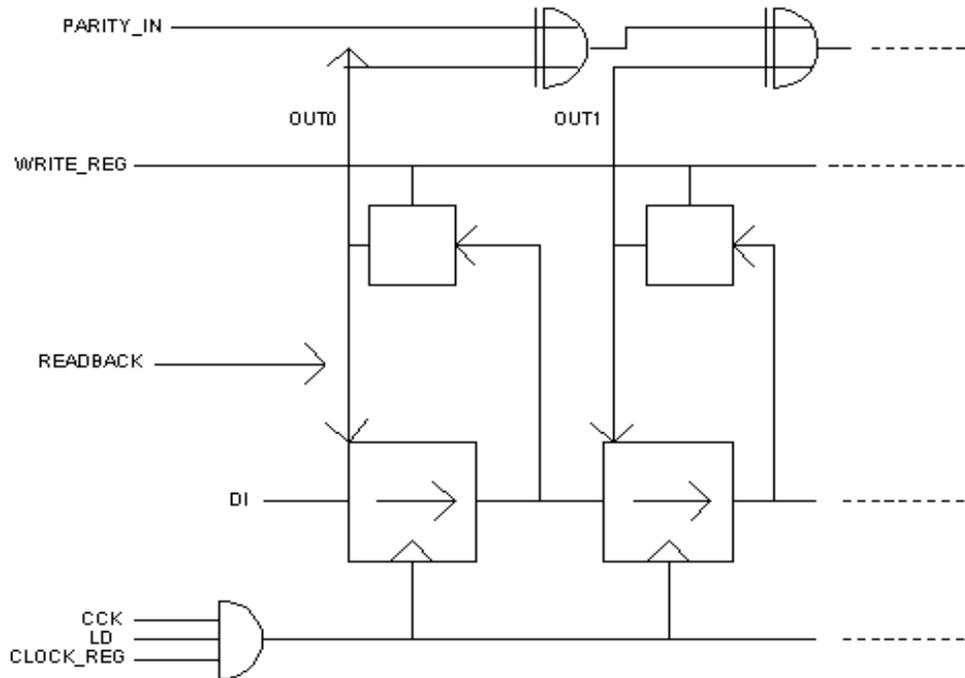
Bit	Name	Funktion
0	CMDRESET	Senden eines 'Softresets' zum Löschen der Ausleseblöcke
1	CLOCKREG	Anschalten der Clock zum Takten des globalen Registers
2	WRITEREG	Laden des globalen Registers
3	READBACK	Zurücklesen des globalen Registers
4	CLOCKPIX	Anschalten der Clock zum Takten des Pixelschieberegisters
5	WRITE2	Laden des 2. Bits zur Feineinstellung der Diskriminatorschwelle
6	WRITEMASK	Laden des Bits zur Maskierung der Pixel
7	WRITE0	Laden des 0. Bits zur Feineinstellung der Diskriminatorschwelle
8	WRITE1	Laden des 1. Bits zur Feineinstellung der Diskriminatorschwelle

8.2.4 Das globale Register

Das globale Register besteht aus einer 109 Bit langen Schieberegisterkette, die alle Blöcke verbindet, welche Kontrollbits zur Einstellung unterschiedlicher Operationsmodi oder Chipparameter beinhalten. Es befindet sich Chipkontrollteil (s. Abbildung 32). Aus layouttechnischen Gründen existiert kein großes zentrales Register, sondern man speichert die Werte direkt dort, wo sie benötigt werden. Auf diese Weise wird der Leitungsaufwand zur Verbindung der einzelnen Bits verringert.

Die Daten werden über DI mit CCK in das Register getaktet, während der Befehl CLOCKREG und LD aktiv sind. Gültig werden diese Werte, wenn anschließend der Befehl WRITEREG an den Chip gesendet wird, indem sie in dazu gehörende 'Latches' geschrieben werden, die eine Art Schattenregister bilden. Das globale Register kann auch ausgelesen werden, indem man das Kommando READBACK aktiviert und so den Inhalt der 'Latches' zurück in das Register transferiert. Taktet man nun analog zum Schreiben die Daten wieder hinaus und wählt den korrekten Ausgang des Ausgangsmultiplexers (s. Kapitelende), so werden diese an DO sichtbar. Auf diese Weise läßt sich überprüfen, ob beim Transfer alles ordnungshalber abgelaufen ist und sich der Registerinhalt nicht verändert hat, beispielsweise durch SEU. Damit diese READBACK Prozedur im Betrieb bei ATLAS nicht ständig wiederholt werden muß, wird die Parität des Registerinhalts durch eine XOR-Kette berechnet. Bei einem Bitflip kann man diesen sofort entweder an DO oder durch eine entsprechende Warnung im EOE Wort (**End Of Event**, s. Abschnitt 8.2.10) erkennen. Durch Senden eines 'Hardresets' wird das Register gelöscht.

Abbildung 35: Blockschaltbild des globalen Registers: Von oben nach unten werden die XOR-Kette, das Schattenregister und das Schieberegister mit den zugehörigen Leitungen dargestellt



In nachfolgender Tabelle wird die Verwendung der Bits im globalen Register aufgelistet:

Tabelle 3: Die Bits im globalen Register des FED1

Bit	Anzahl	Blockname	Funktion
0 - 69	70	Analogregister	Einstellung DACs, Maskierung der Spaltenpaare
70 - 86	17	Hauptregister	Einstellung VCCD, VTH, 'Senseamp', Testmodi
87 - 93	7	Graycodegenerator	Einstellung 'Latency'
94 - 98	5	Selbsttrigger	Aktivierung, Triggerpulsbreite
99 - 102	4	Ausgangsmultiplexer	Auswahl des Signals für DO
103 - 106	4	Hitbusmultiplexer	Auswahl des Signals für HIT
107 - 108	2	Phigenerator	Einstellung der CEU Clock

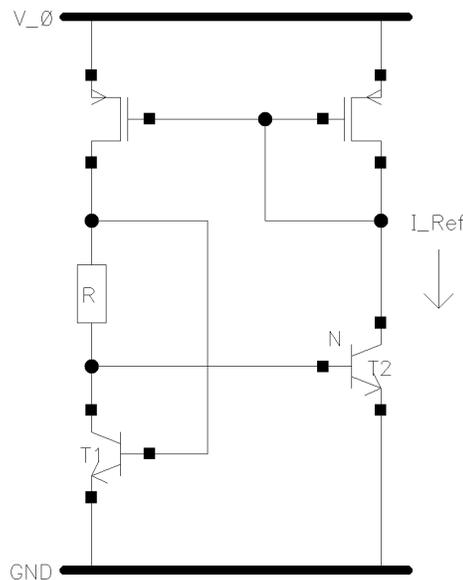
Die oben genannten Blöcke werden nun detaillierter beschrieben:

- **Analogregister**

Dieser Block enthält sieben 8 Bit Strom-DACs, zwei 5 Bit Spannungs-DACs mit Treibern, alle Registerbits, die für deren Einstellung notwendig sind, 9 zusätzliche Bits zum An- bzw. Abschalten jedes einzelnen Spaltenpaares (s. Abschnitt 8.2.6) und fünf unbenutzte Bits.

Zur Erzeugung des Referenzstroms für die Strom-DACs dient eine sogenannte 'Bandgap'-Referenz, deren besondere Eigenschaft darin liegt, daß der von ihr gelieferte Strom I_{Ref} unabhängig von der Versorgungsspannung ist. Ein charakteristisches Merkmal liegt in der Verwendung von bipolaren Transistoren, deren Transferkennlinien exponentiell von der Basis-Emitterspannung abhängen, was sich bei der unten durchgeführten Berechnung von I_{Ref} als äußerst vorteilhaft erweist.

Abbildung 36: Schaltplan der 'Bandgap'-Referenz



In obigem Schaltplan stellt der PMOS Stromspiegel sicher, daß sowohl im linken als auch im rechten Zweig derselbe Strom I_{Ref} fließt. Der Faktor N am bipolaren Transistor T2 bedeutet, daß dieser bei identischen Betriebsparametern den N -fachen Strom im Vergleich zu T1 liefert. Mit diesen Angaben gelten folgende Gleichungen:

$$I_{Ref} = \frac{V_B(T1) - V_B(T2)}{R} \quad (8.3)$$

$$I_{Ref} = I_S \left(e^{\frac{V_B(T1)-GND}{V_T}} - 1 \right) \approx I_S e^{\frac{V_B(T1)-GND}{V_T}} \quad (8.4)$$

$$I_{Ref} = N I_S \left(e^{\frac{V_B(T2)-GND}{V_T}} - 1 \right) \approx N I_S e^{\frac{V_B(T2)-GND}{V_T}}, \quad (8.5)$$

mit

$$\begin{aligned}V_B &= \text{Basispotential} \\I_S &= \text{Transistorsperrstrom} \\V_T &= \frac{kT}{q}.\end{aligned}$$

Daraus ergibt sich für den Referenzstrom:

$$I_{Ref} = \frac{V_T}{R} \ln N. \quad (8.6)$$

Man erkennt lediglich eine lineare Abhängigkeit von V_T , was gleichbedeutend mit einer Temperaturabhängigkeit der Schaltung ist. Da der Pixeldetektor bei einer festen Temperatur von -7°C betrieben werden soll, braucht man diese nicht schaltungstechnisch zu kompensieren.

Alle Strom-DACs generieren auf dem Chip Ströme im nominellen Bereich von $0 - 255 \mu\text{A}$ in $1 \mu\text{A}$ großen Schritten; sie benutzen alle dieselbe interne 'Bandgap'-Referenz bei der Stromerzeugung. Die Ausgangsströme der DACs werden in die 'Bias'-Schaltkreise, die aus Stromspiegeln bestehen, geführt und dort geeignet heruntergeteilt, um schließlich die Versorgung der Analogteile der einzelnen Pixel zu bilden bzw. einen Injektionspuls für die analoge Testinjektion zu erzeugen.

Die Strom-DACs selbst bestehen jeweils aus einer 8×8 -Matrix von layouttechnisch identischen $4 \mu\text{A}$ Stromquellen, einer $2 \mu\text{A}$ und einer $1 \mu\text{A}$ Quelle. Beim Durchfahren eines DACs werden die identischen Stromquellen nach folgendem Schema eingeschaltet: Innerhalb einer mittleren Spalte wird zunächst eine in der Mitte liegende Quelle aktiviert, dann folgen jeweils nach oben bzw. unten alternierend die angrenzenden Quellen. Wenn eine ganze Spalte 'an' ist, wird auf die gleiche Weise nach rechts und links mit den Spalten verfahren.

Sinn dieser komplizierten Anordnung ist die Kompensation von herstellungsbedingten Inhomogenitäten, die zu einem nicht linearen Verhalten des DACs führen. Auf diese Weise kann eine über die DAC-Fläche lineare oder sogar parabolische Verteilung der Transistorstärken ausgeglichen [57] und eine höchst präzise Funktion gewährleistet werden. Aus Platzgründen wurden PMOS-Stromquellen und nur 63 statt 255 Einheitszellen benutzt, denn man ist im Layout durch die Breite eines Spaltenpaares limitiert.

Eine weitere Besonderheit besteht darin, daß die einzelnen Ströme der DACs eigentlich nicht ausgeschaltet, sondern nach Masse geleitet werden. Das bedeutet, dass die Transistoren immer leitend sind und damit geringere Strahlenschädigungen erfahren werden [58].

Die folgende Tabelle zeigt die Divisoren der Stromteilung und die Grundeinstellungen der Ströme bei einem üblichen DAC Wert von 64:

Tabelle 4: Die Ströme des FED1 bei einem DAC Wert von 64

Strom	Divisor	Wert [μA]
ITR	54	1,2
ID	9	7,1
IF	14400	0,0044
IP	5,4	11,6
IL	27	2,4
IPS	27	2,4
ICH	1	64

Die Spannungs-DACs bestehen im Prinzip aus einstellbaren Widerstandsketten, die einen Spannungsteiler bilden. In 32 Schritten mit einer Größe von etwa 30 mV können diese im Bereich zwischen 1 V –2 V verändert werden und so die Spannungen zur globalen Einstellung der Diskriminatorschwellen liefern.

Die Knoten zwischen den Strom-DACs und den 'Bias'-Schaltkreisen sind an den vorgesehenen Ausgangs-'Pads' zugänglich und können dort mithilfe eines Amperemeters nach Masse gemessen werden. Das gleiche gilt für die Ausgangsspannungen der Spannungs-DACs. Ebenso existiert die Möglichkeit, die DACs abzuschalten und über diese 'Pads' extern Ströme bzw. Spannungen aufzuprägen, falls ein Defekt vorliegt.

- **Hauptregister**

Das Hauptregister trägt seinen Namen nur aus historischen Gründen, weil es bei früheren Prototypen nur dieses eine globale Register gab und keine Aufspaltung in Blöcke vorhanden war. Seine Länge beträgt 17 Bits, und es speichert Werte für diverse Chipeinstellungen (s. Tabelle 3).

- **Graycodegenerator**

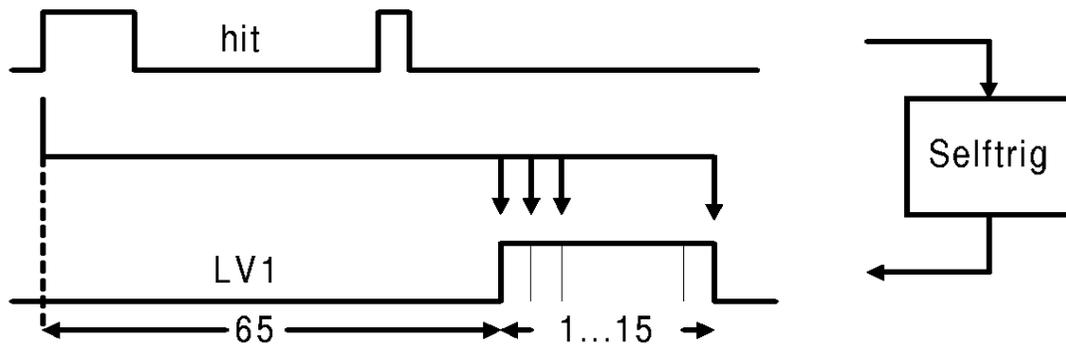
Alle Zeitmarken, die in der Matrix verteilt werden, werden durch den Graycodegenerator erzeugt. Da sich bei der 'Gray Code' Zählweise von einem Wert zum nächsten immer nur ein Bit verändert, empfiehlt sich die Verwendung desselben, um bei der Trefferaufnahme im Pixel immer nur höchstens einen Zähltakt falsch liegen zu können, falls man gerade beim Weiterspringen des Zählers speichert.

Dieser Block enthält einen mit XCK getakteten 7 Bit Binärzähler zur Erzeugung der Zeitmarken, einen 7 Bit Addierer zur Generation der 'Latency' verzögerten Zeitmarken, einen Binärcode zu 'Gray Code' Konverter und sieben Registerbits zur Speicherung des 'Latency'-Wertes.

- **Selbsttrigger**

Im normalen ATLAS Betriebsmodus schickt der MCC die Level 1 Trigger zum Front End Chip. Für Labormessungen mit radioaktiven Quellen muß man diese Trigger auf andere Weise erzeugen, und zwar exakt zu einem Zeitpunkt der um die 'Latency' verzögert ist, um überhaupt Trefferdaten aus dem Chip auslesen zu können.

Abbildung 37: Die Funktionsweise des Selbsttriggers



Eine elegante Lösung bietet der Selbsttrigger, der diese Aufgabe direkt auf dem Chip erfüllt. Er verwendet das Hitbussignal, das Treffer in der Matrix sofort meldet (s. Abschnitt 8.2.6), verzögert es um 65 XCK Takte und sendet es als Level 1 Signal mit einer programmierbaren Länge zurück an die Matrix. Wenn nun die 'Latency' auf den richtigen Wert festgelegt wurde, erhält der Chip einen korrekten Trigger.

In diesem Block werden fünf Registerbits benötigt, um den Selbsttriggermodus zu aktivieren bzw. deaktivieren und die Länge des Level 1 Signals zwischen 1 und 15 XCK Takten zu programmieren.

- **Ausgangsmultiplexer/Hitbusmultiplexer**

Es gibt zwei 16 zu 1 Multiplexer mit identischen Auswahlcodes, um sich an den Datenausgängen des Chips verschiedene wichtige interne Signale anschauen zu können: Der Ausgangsmultiplexer ist mit DO, der Hitbusmultiplexer mit HIT verbunden. Aus Platzgründen im Layout können statt 16 jeweils nur 8 Signale an die Ausgangspads weitergeleitet werden.

- **Phigenerator**

Der Phigenerator erzeugt aus der 40 MHz XCK eine 5 MHz Clock für die Überlaufkontrolle der EOC (s. Abschnitt 8.2.9), eine 10 MHz Clock für die RCU und eine zweiphasige Clock für die CEU. Letztere liegt mit einer Frequenz von 20 MHz, 10 MHz oder 5 MHz vor, welche mittels zwei Registerbits ausgewählt werden kann.

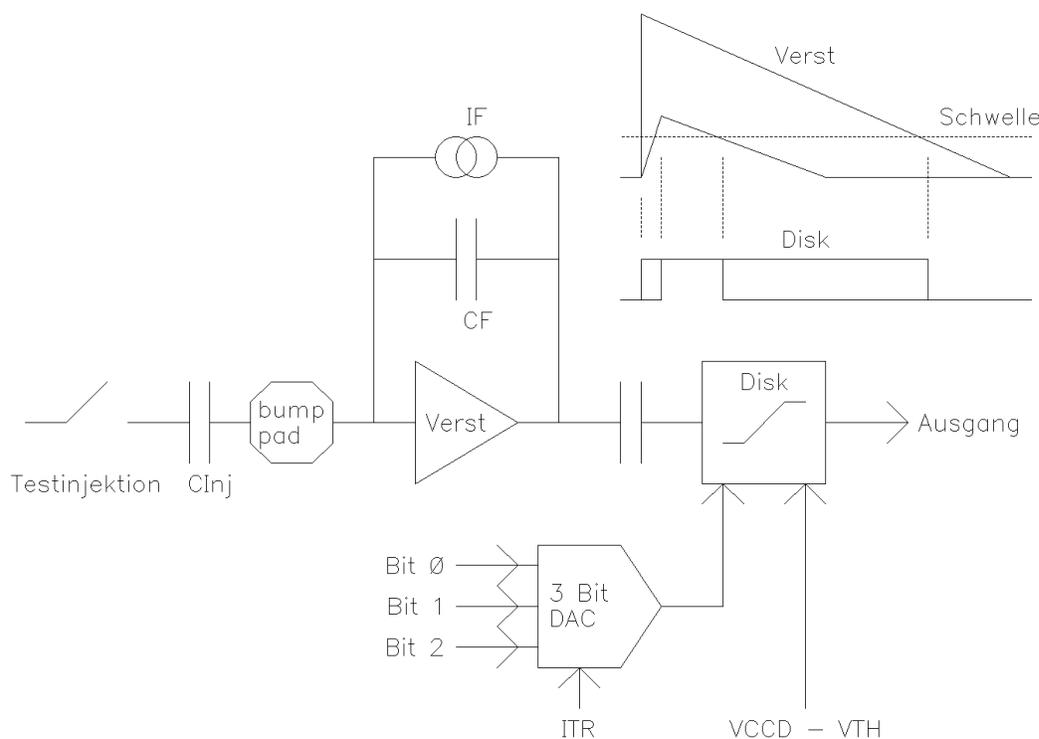
8.2.5 Der Analogteil

Der Analogteil enthält einen ladungsempfindlichen Vorverstärker, der über die Kapazität $C_F = 2,4 fF$ im Detektor erzeugte Signale aufintegriert. Durch den konstanten Strom IF wird der Kondensator C_F wieder entladen bzw. ein Detektorleckstrom bis zu 50 nA kompensiert. Auf diese Weise entsteht ein Verstärkersignal, dessen Breite sich linear zu seiner

Höhe verhält. Dieser Puls wird nun wechselstrommäßig an den Diskriminator weitergegeben und digitalisiert. Die Einstellung der Schwelle erfolgt einerseits grob global auf dem Chip mithilfe der Differenz der Spannungen V_{CCD} und V_{TH} . Andererseits ermöglicht ein 3-Bit-DAC mit einer durch ITR stromgesteuerten, variablen Schrittweite eine lokale Feinjustierung in jedem einzelnen Pixel, um ein möglichst homogenes Ansprechverhalten der gesamten Matrix zu erzielen.

In der unteren Abbildung sind neben einem Blockschaltbild des Analogteils auch die Ausgangssignale des Verstärkers und des Diskriminators dargestellt. Durch die bereits erwähnte Proportionalität zwischen der Höhe des Verstärkersignals und der Zeitspanne TOT (**T**ime **O**ver **T**hreshold), während der der Diskriminator feuert, erhält man eine grobe Information über die im Detektor deponierte Ladungsmenge.

Abbildung 38: Blockschaltbild des Analogteils

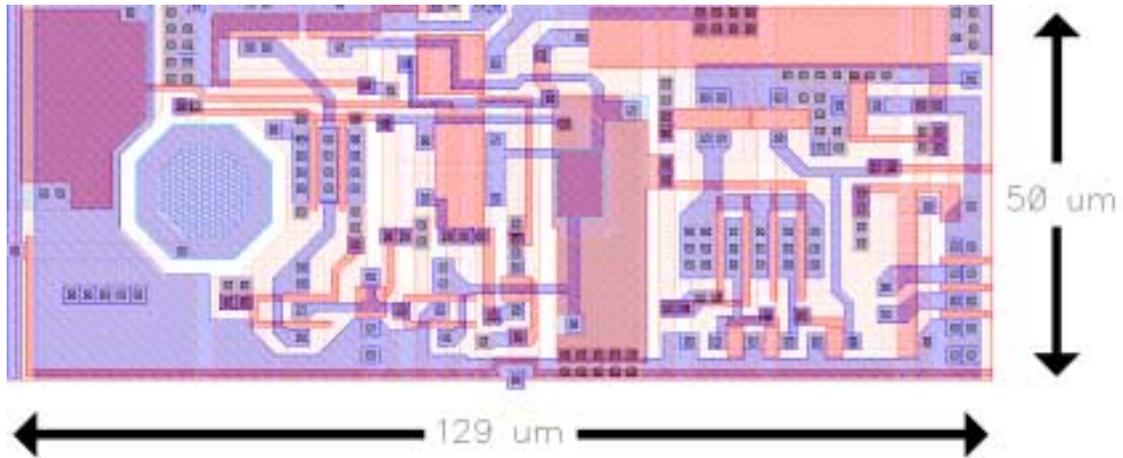


Weiterhin befinden sich im Analogteil eine Testkapazität von 12 fF , mit der durch Anlegen eines Spannungssprungs Ladung in den Vorverstärker injiziert werden kann, und das 'Bumpbondpad' zur Kontaktierung des Sensors. Dieser Spannungssprung erfolgt entweder intern durch einen auf dem Chip implementierten Pulsgenerator ('Chopper') oder extern über das VCAL-Pad. Der Vorverstärkerausgang des rechten unteren Pixels besitzt zu Testzwecken eine Verbindung zum AMP-Pad, an dem man die Funktionsfähigkeit kontrollieren kann.

Der Forderung, für diese Zelle ein besonders kompaktes Layout zu erstellen, damit die vorgegebene Pixelgröße ausreicht, ist nur äußerst schwierig nachzukommen. Neben den

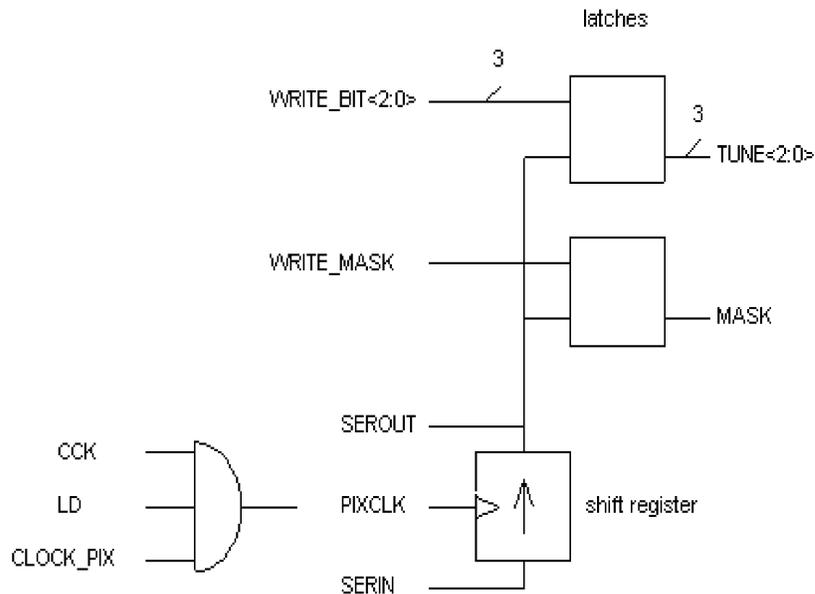
Transistoren müssen nicht nur einige Kondensatoren realisiert werden, die in DMILL eine große Fläche belegen, sondern auch ein bedeutender Teil für das 'Bump Bond Pad' reserviert werden. Zusammen mit den Versorgungsleitungen, die über die gesamte Pixelmatrix verteilt und möglichst über wenig aktive Bereiche geführt werden müssen, ist das Ergebnis bei einer Gesamtgröße von $50 \mu\text{m} \times 129 \mu\text{m}$ um etwa 10 % kleiner als im FEC, obwohl der DMILL Prozeß an sich eine geringere Integrationsdichte besitzt.

Abbildung 39: Layout des Analogteils



8.2.6 Der Pixelkontrollteil

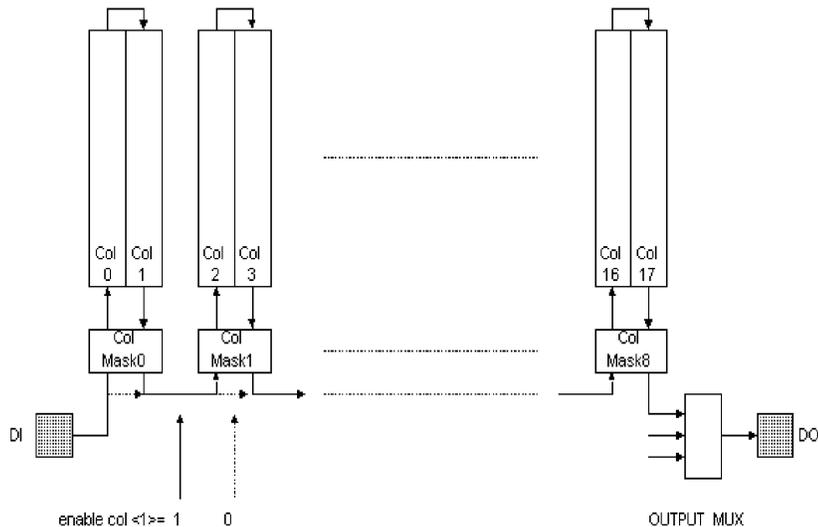
Abbildung 40: Blockschaltbild des Pixelkontrollteils



Der Pixelkontrollteil besteht aus vier statischen Speicherzellen ('Latch'), in denen zum einen die drei Bits zur Einstellung des oben erwähnten DACs abgelegt werden. Zum anderen kann ein eventuell defekter Pixel maskiert werden und wird somit von allen nachfolgenden Schaltungsteilen ignoriert. Die Werte werden mithilfe eines semistatischen Schieberegisters mit einer nicht überlappenden, zweiphasigen 5 MHz Clock in die Matrix transportiert, wobei jeder Pixelkontrollteil ein Bit dieses Registers beinhaltet (s. Abbildung 40).

Die einzelnen 'Flipflops' des Pixelschieberegisters werden folgendermaßen verknüpft: Die seriellen Daten treffen zu Beginn des Schreibvorgangs in der linken unteren Ecke der Matrix ein (Spalte 0, Zeile 0). Dann werden diese innerhalb derselben Spalte nach oben getaktet (bis Spalte 0, Zeile 159) und weiter in die benachbarte Spalte (Spalte 1, Zeile 159) geschoben, von wo sie wieder nach unten (Spalte 1, Zeile 0) gebracht werden, usw. Die letzte Zelle bildet der rechte untere Pixel (Spalte 17, Zeile 0). Von dort wird der serielle Registerausgang über den Ausgangsmultiplexer zu DO gebracht.

Abbildung 41: Blockschaltbild des Pixelregisters mit Spaltenpaarmaskierung



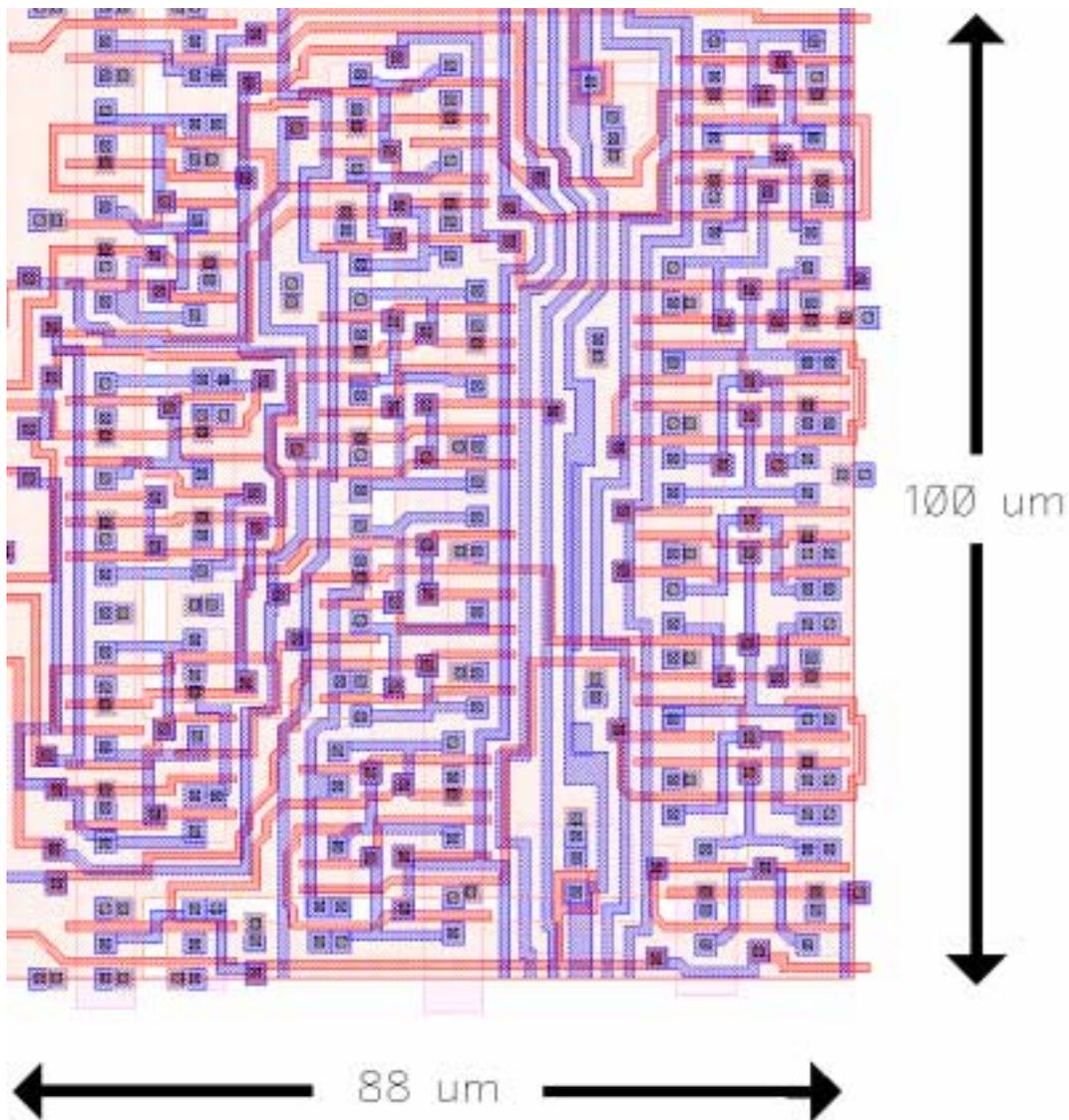
Um beispielsweise die Maske zum Abschalten gewisser Pixel zu laden, wird folgende Prozedur durchgeführt: Der Befehl CLOCKPIX wird an das Kommandoregister geschickt und 2880 Datenbits in das Register transferiert. Dann aktiviert man das Kommando WRITEMASK und lädt damit die 'Latches', so daß die Maske gültig wird. Zum Schluß wird durch Senden eines beliebigen anderen Kommandos WRITEMASK wieder deaktiviert.

Zu Test- und Diagnosezwecken stehen im Pixelkontrollteil eine digitale Injektionsvorrichtung und ein Hitbussystem zur Verfügung, das aus einer schnellen ODER-Verknüpfung ('wired OR') der Diskriminatorausgänge aller Pixel besteht und sofort den Eintritt eines Treffers in der Matrix meldet.

Das Design dieses Blocks stellt topologisch eines der schwierigsten des gesamten Chips

dar: Einerseits müssen in der Längsrichtung 8 Signale über die gesamte Spalte geführt werden. Andererseits werden 5 Signale zum Analogteil geführt, während nur eines auf der anderen Seite in der Auslese benötigt wird, so daß hier eine natürliche Asymmetrie besteht. Um eine besonders Platz sparende Anordnung zu realisieren, wird das Layout von zwei übereinander liegenden Pixelkontrollteilen miteinander verwoben. Mit der nun zur Verfügung stehenden Zellenhöhe von $100\ \mu\text{m}$ ist es möglich, die aktive Fläche in drei Doppelreihen von Transistoren aufzuteilen und den großen, geschlossenen Hitbus Transistor unter den Längsbussen zu verstecken. Allerdings kann kein statisches 'Flipflop' für das Register untergebracht werden, sondern lediglich die oben bereits erwähnte, semistatische, jedoch sehr konservative Version. Die benötigte Fläche beträgt $100\ \mu\text{m} \times 88\ \mu\text{m}$, was im Vergleich zum FEC, der dieselbe Schaltung benutzt, eine Verschlechterung um 20 % bedeutet.

Abbildung 42: Layout des Pixelkontrollteils



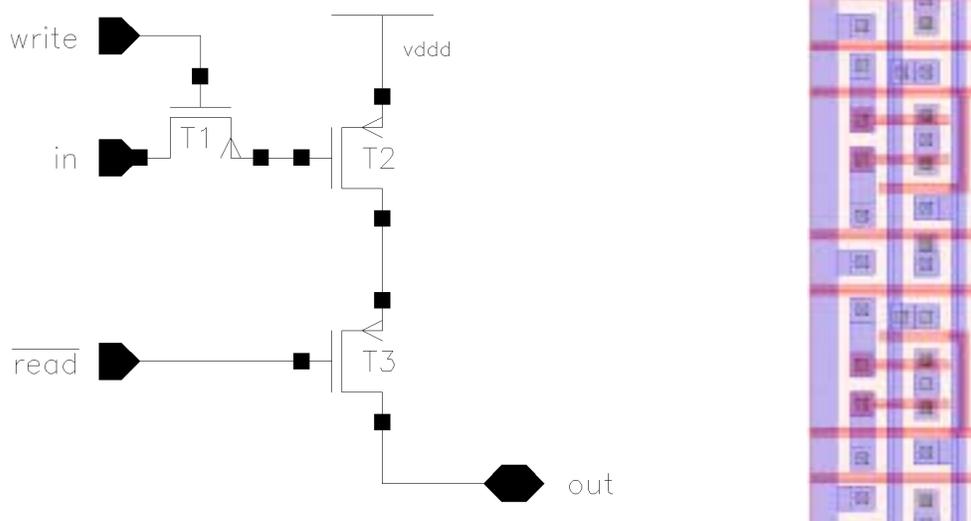
Für den Fall, daß einige Spaltenpaare Probleme zeigen, beispielsweise durch ein defektes 'Flipflop' im Pixelregister oder im Ausleseteil, wird ein Block implementiert, der einzelne dieser Spaltenpaare maskiert und damit die Möglichkeit bietet, wenigstens den restlichen Teil des Chips ohne das defekte Spaltenpaar zu betreiben. Wenn ein Spaltenpaar 'an' ist, funktioniert alles normal, wenn es 'aus' ist, wird der Registerausgang des links davon liegenden Spaltenpaares mit dem Registeringang des rechts davon liegenden Spaltenpaares verbunden. Somit wird das Register um 320 Bits verkürzt, was beim Beschreiben berücksichtigt werden muß. Zusätzlich wird der Hitbus des betreffenden Spaltenpaares abgeschaltet und die zugehörige EOC separiert.

8.2.7 Der Ausleseteil

Im Ausleseteil befinden sich eine Logik zur Treffererkennung, ein zweistufiger 'Domino-scanner' zur Treffersuche innerhalb einer Spalte und ein Block zur Trefferspeicherung im jeweiligen Pixel. Mithilfe eines 8 Bit ROMs (**R**ead **O**nly **M**emory), das sich zwei nebeneinander liegende Pixel in einem Spaltenpaar teilen, wird die jeweilige Pixeladresse fest verdrahtet, so daß die Zeile des getroffenen Pixels innerhalb dieses Spaltenpaares bekannt ist. Die zugehörige Zeitinformation wird in einem 14 Bit RAM (**R**andom **A**ccess **M**emory) abgelegt.

Diese Zelle unterliegt schaltungstechnisch wohl den meisten Kompromissen. Alle Speicherzellen bestehen wegen der geringen Größe ausschließlich aus PMOS Transistoren und sind in dynamischer Logik implementiert. Pro RAM Bit werden lediglich 3, pro ROM Bit sogar nur ein Transistor verwendet.

Abbildung 43: Schaltplan eines und Layout von 4 RAMs



Man erkennt, daß das Layout äußerst platzsparend umgesetzt wird, indem ein Großteil der Fläche mit Metall (blau) bedeckt und die Transistoren darunter verlegt werden. Dies ist zwingend notwendig, um letztlich die $400 \mu m$ Pixellänge einzuhalten.

Aus dem Schaltplan läßt sich entnehmen, daß vom Eingang her nur eine gute logische '1' geschrieben werden kann, die beim Lesen bewirkt, daß kein Strom zum Ausgang fließt. Versucht man eine logische '0' zu schreiben, wird aus $0 V$ hinter dem Eingangstransistor T1 ein Signal von etwa $0,7 - 1 V$ erzeugt, welches der PMOS Schwellenspannung entspricht. Dieser Wert wird nun auf der Gatekapazität von T2 dynamisch gespeichert und bewirkt beim Lesen einen Ausgangsstrom von etwa $20 \mu A$. Berücksichtigt man nun, daß am Ausgang dieser Zelle alle 160 Pixel einer Spalte durch eine etwa $8 mm$ lange Leitung miteinander verbunden sind, läßt sich leicht einsehen, daß die parasitäre kapazitive Last von etwa $10 pF$ zu groß ist, um eine nennenswerte Spannungsänderung auf diesem Datenbus zu erzeugen. Man erhält nach der Formel

$$U = \frac{I t}{C} \quad (8.7)$$

durch Einsetzen obiger Werte einen Spannungshub von $50 mV$ bei der maximalen Auslesefrequenz von $20 MHz$. Da der Sachverhalt nach Bestrahlung aufgrund der verschlechterten Transistoreigenschaften noch problematischer wird, sind zur Lösung dieses Problems noch einige Bemühungen notwendig, die im nächsten Abschnitt beschrieben werden.

Die interne Auslese erfolgt nicht mit der ATLAS Frequenz von $40 MHz$, weil das 'Durchscannen' der Matrix mehr Zeit benötigt, was jedoch wegen der geringen Trefferwahrscheinlichkeit von 10^{-4} pro Pixel kein Problem darstellt. Um ein Auslesen mit $20 MHz$ zu erreichen, sucht der 'Scanner' nicht jeden Pixel einer Spalte einzeln durch, sondern sieht in der ersten Stufe in den bereits erwähnten 16er Blöcken nach, ob dort Treffer vorliegen. Bei einem positiven Ergebnis in einem Block, geht er eine Stufe tiefer und untersucht die Einzelpixel. Statt 160 logische Gatter werden also maximal $10 + 16$ durchlaufen, was bei einer Durchlaufzeit von etwa $1 - 2 ns$ pro Gatter gerade ausreicht. Darum ist die Auslesefrequenz vorsichtshalber auf die kleineren Werte von $5 MHz$ bzw. $10 MHz$ einstellbar, was sich nach Bestrahlung noch als notwendig herausstellen könnte.

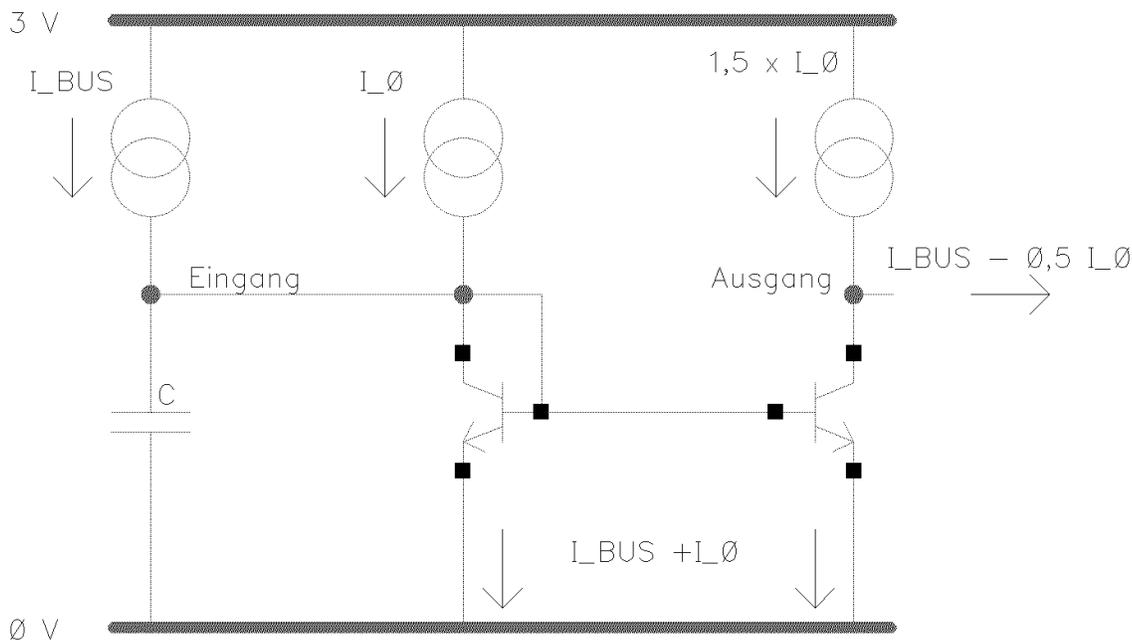
8.2.8 Die CEU

Wie bereits erwähnt, liefert die CEU die Auslesesignale für das zugehörige Spaltenpaar und füllt die Trefferdaten von oben nach unten in die EOC Speicherplätze. In diesem Zusammenhang liest sie alternierend die linke bzw. rechte Spalte aus und fügt die entsprechende Information in Form eines neunten Adressbits hinzu.

Die andere Aufgabe dieses Blocks besteht darin, die im vorigen Abschnitt erläuterten geringen Spannungs- bzw. Stromänderungen auf den Auslesebussen zu verarbeiten und in

ein 'full swing' CMOS Signal umzuwandeln. Dieses Problem wird durch den sogenannten 'Senseamp' gelöst.

Abbildung 44: Schaltplan des 'Senseamps'



Der 'Senseamp' funktioniert im Prinzip wie folgt: Um beim Beschreiben des Datenbusses nicht den ohnehin schon niedrigen Strom I_{BUS} zu verbrauchen, um dessen große parasitäre Kapazität C aufzuladen, übernimmt dies der kleine Strom I_0 , der mit einer programmierbaren Stromquelle erzeugt wird. Auf diese Weise bleibt also am Eingang das Potential annähernd konstant. Der gesamte Strom $I_{BUS} + I_0$ wird nun über einen bipolaren Stromspiegel in den Zweig mit einem stehenden Strom von $1,5 \times I_0$ umgeleitet, so daß am Ausgang der Strom $I_{BUS} - I_0$ für das Schalten des nächsten logischen Gatters mit kleiner Kapazität zur Verfügung steht. Das führt dazu, daß beim Schreiben einer logischen '1' ($I_{BUS} = 0 \mu A$) der Ausgang des 'Senseamps' auf logisch '0' geht, während das Schreiben einer logischen '0' ($I_{BUS} = 20 \mu A$) eine Erzeugung einer logischen '1' am Ausgang bewirkt. Das invertierende Verhalten des RAMs wird somit kompensiert.

Die Vorteile dieser Schaltung liegen in dem niedrigen Spannungshub auf dem Datenbus, der geringen Sensitivität auf die parasitäre Buskapazität C , der theoretisch ebenfalls sehr geringen Abhängigkeit vom Strom I_0 und nicht zuletzt dem niedrigen Leistungsverbrauch. Wegen der geringen Eingangsimpedanz der bipolaren Transistoren gegenüber MOSFETs werden diese hier an der einzigen Stelle im gesamten digitalen Teil des Chips implementiert.

Das Design der 'Senseamps' stellt einen der schwierigsten Punkte überhaupt dar, weil die auftretenden Ströme und Kapazitäten nur schwer zu evaluieren sind und letztlich aus

Schätzungen bestehen. Dynamische Knoten, parasitäre Effekte und Veränderungen nach Bestrahlung werden durch die Simulationsmodelle nur teilweise oder gar nicht berücksichtigt, so daß eine möglichst robuste und von diesen Parametern unabhängige Schaltung erfunden werden mußte. Für die Funktionstüchtigkeit des gesamten Chips ist dies ein entscheidender Punkt. Weiterhin benötigt ein bipolarer Transistor im Layout etwa den achtfachen Platz eines NMOS, was bei der Verwendung von 22 'Senseamps', also 44 *npn*-Transistoren wiederum einige Schwierigkeiten aufwirft. Nicht umsonst ist also die CEU der letzte 'Layout'-Block des Chips, der fertiggestellt wurde.

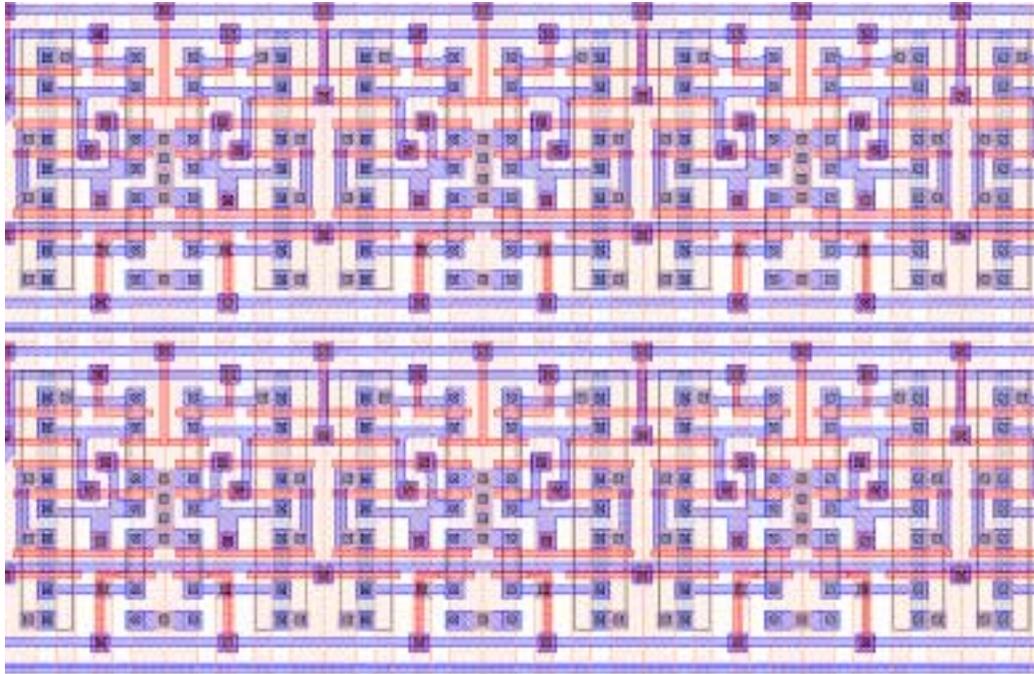
8.2.9 Die EOC

Die EOC besteht aus einer intelligenten Speicherlogik, die 24 Zellen pro Spaltenpaar enthält. Dort werden die aus der Pixelmatrix ausgelesenen Trefferdaten (Adresse und Zeitmarken) von oben nach unten zwischengespeichert, bis ein Level 1 Trigger eintrifft. Die Zeitdifferenz zwischen der Entstehung eines Treffers und der Ankunft des zugehörigen Level 1 muß exakt der von ATLAS vorgegebenen 'Latency' entsprechen, damit ein physikalisch interessantes, gültiges Ereignis vorliegt. Dieser Zeitvergleich wird in der EOC direkt vorgenommen, und die Trefferdaten werden sofort gelöscht, wenn es keinen passenden Trigger gibt, damit wieder Platz für neu ankommende Daten entsteht. Falls dieses Triggersignal jedoch exakt nach der vorgegebenen 'Latency' ankommt, wird ein schneller horizontaler 'Scanner' gestartet, der die EOC aller Spalten durchsucht. Die jeweils gültigen Daten werden der Reihe nach (zuerst am weitesten links liegende Spalte, dann oberster EOC Platz) an den Serialisierer (s. Abschnitt 8.2.10) geleitet und 4 Adressenbits für die Spaltenpaaradresse hinzugefügt.

Für den Fall, daß während des Betriebs kurzzeitig die EOC eines Spaltenpaares gefüllt wird, wird eine Überlaufkontrolle implementiert, die im EOE (s. Abschnitt 8.2.10) Wort eine Warnung hervorruft. Das bedeutet, daß die gespeicherten Daten zwar korrekt sind, aber eventuell nicht alle Treffer erfaßt wurden.

Auch dieser Block stellt den Designer vor einige Probleme. Aufgrund von Simulationen sollten mindestens 30 EOC Speicherplätze pro Spaltenpaar zur Verfügung stehen, um die Nachweiseffizienz von 95 % in der B-Lage zu gewährleisten [59]. Wiederum muß also auf einer vorgegebenen Fläche von $1200 \mu m \times 800 \mu m$ ein möglichst kompaktes Layout erstellt werden. Dabei besteht die größte Schwierigkeit in der Einhaltung der Breite, weil einerseits eine Vielzahl von Daten- und Versorgungsleitungen kaum über diesen Bereich passen und andererseits die aktive Fläche pro Speichereinheit sinnvollerweise rechteckig sein soll, damit beim Aneinanderfügen mehrerer Zellen kein Platz verschwendet wird. Letztlich muß sogar die Speicherung der Zeitmarken für die steigenden Trefferflanken in das FIFO (s. Abschnitt 8.2.10) ausgelagert werden, weil kein Platz mehr für die Datenbusse vorhanden ist.

Abbildung 45: Ausschnitt aus dem Layout der EOC



8.2.10 Die RCU

Die RCU kontrolliert den Datentransfer von der EOC zum DO 'Pad' und besteht im wesentlichen aus einem FIFO (**F**irst **I**n **F**irst **O**ut), dem TOT Generator, dem Serialisierer und dem EOE Generator.

- **Das FIFO**

Das FIFO ist eine Speichereinheit mit 16 Plätzen, die 16 Level 1 Triggernummern, also ebenso viele physikalisch interessante Ereignisse, speichern kann. Da in der EOC zu wenig Platz vorhanden ist, werden die zu den Triggern gehörenden Zeitmarken und eventuell auftretende EOC Warnungsmeldungen auch im FIFO abgelegt, weil in diesem Bereich die Integrationsdichte niedriger ist. Zweck dieses Blocks ist, daß man verhindert, Level 1 Trigger bzw. Trefferdaten zu verlieren, wenn im Falle vieler Treffer die Auslese vorhergehender Trigger noch laufen sollte. Weiterhin initiiert das FIFO die Auslese des Chips, so lange es abzuarbeitende Trigger enthält. In diesem Block sind auch die beiden 4 Bit Zähler zur Erzeugung der ankommenden und der zu verarbeitenden Triggernummern implementiert.

- **Der TOT Generator**

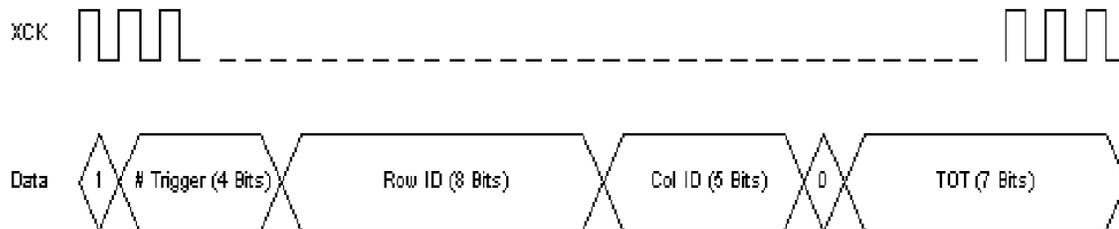
Da die Zeitmarken der steigenden und fallenden Flanke eines Treffers bekannt sind, kann man deren Differenz dazu benutzen, eine grobe analoge Information zu extrahieren, das sogenannte TOT. Dieses gibt einen Hinweis auf die Ladungsmenge, die

im Vorverstärker deponiert wurde. Der TOT Generator berechnet die Differenz der beiden 'Gray Code' Zahlen und wandelt diese in einen binären TOT Wert um.

- **Der Serialisierer**

Der Serialisierer besteht im wesentlichen aus einem, mit XCK getakteten, Schieberegister, das parallel ankommene Trefferdaten in ein 26 Bit langes Wort umwandelt und diese nach folgendem Schema an DO hinaustaktet: (Das erste Bit erscheint zuerst am Ausgang)

Abbildung 46: Datenprotokoll des Serialisierers



- **Der EOE Generator**

Um die Daten eines Triggers besser von Daten des nachfolgenden Triggers trennen zu können, wird zwischen beiden vom Chip ein sogenanntes EOE Wort gesendet. In diesem Wort wird einfach die Zeilennummer des 26 Bit Stroms durch eine nicht existierende Zeilennummer ersetzt.

Tabelle 5: Die EOE Wörter (hexadezimale Schreibweise)

Adresse	Erklärung
F0	normales EOE Wort
E1	EOE plus EOC Überlauf
E2	EOE plus Fehler in der Parität der Daten im globalen Register
E2	EOE plus beide oben genannten Fehler

8.2.11 Der Resetgenerator

Es gibt zwei Arten, den FED1 in einen definierten Zustand zurückzusetzen:

- 'Softreset'.
- 'Hardreset'.

Der 'Softreset' löscht nur die Blöcke im FED1, die zu der Ausleseketten gehören: Pixel, EOC und RCU. Der 'Hardreset' setzt zusätzlich noch das Kommandoregister und das globale Register zurück.

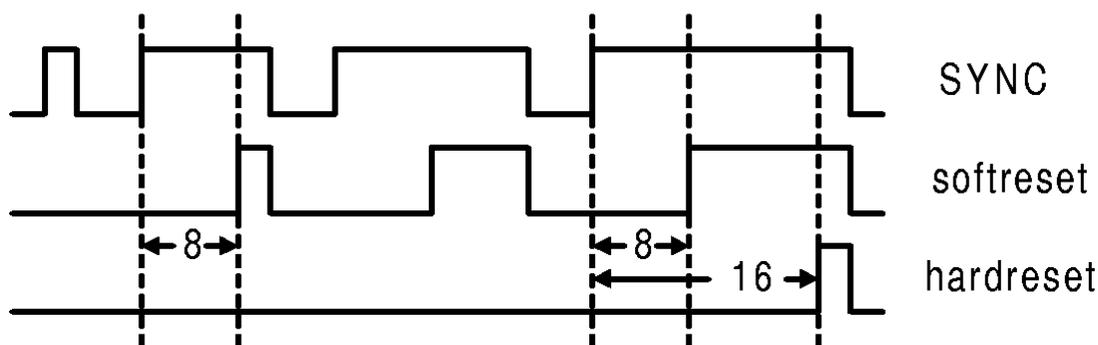
Es bestehen folgende Möglichkeiten, einen 'Softreset' zu erzeugen:

- Der Befehl CMDRESET wird an den FED1 geschickt.
- Ein kurzer SYNC (mindestens 9 XCK Perioden) wird an den FED1 geschickt, wobei die Länge des 'Softreset' identisch mit der des SYNC Signals, vermindert um 8 XCK Takte ist (s. Abbildung 47).
- RSTb wird aktiviert.

Um einen 'Hardreset' zu initiieren, existieren diese Wege:

- Ein langer SYNC (mindestens 17 XCK Perioden) wird an den FED1 geschickt, wobei die Länge des 'Hardreset' identisch mit der des SYNC Signals, vermindert um 16 XCK Takte ist. Ein 'Softreset' wird automatisch 8 XCK Perioden vorher gesendet (s. Abbildung 47).
- RSTb wird aktiviert.

Abbildung 47: Das Zurücksetzen des FED1 mithilfe des SYNC



9 Messungen am FED1

In Kapitel 8 wurde der Aufbau von FED1 detailliert vorgestellt und die kritischen Blöcke diskutiert. Da jedoch kein Simulationsmodell perfekt und die Komplexität der elektronischen Schaltungen sehr hoch ist, treten im allgemeinen Diskrepanzen zwischen den erwarteten und den realen Ergebnissen auf. Ein weiterer Faktor besteht darin, daß in der Simulation parasitäre Effekte und durch Variationen im Herstellungsprozeß oder Strahlenschädigung induzierte Schwankungen der Transistorparameter gar nicht oder nur unzureichend berücksichtigt werden.

Aus diesem Grund sind umfassende Messungen, die die Eigenschaften der Chips charakterisieren bzw. die gestellten Anforderungen überprüfen, zwingend notwendig. Zum anderen wird aufgrund von Prozeßschwankungen in der Herstellung nicht jeder Chip einer Produktion gleich gut funktionieren, so daß man Aufschluß über die Größe der Ausbeute gewinnen muß. Im folgenden werden die Messungen der FED1 Submission vom August 1999 vorgestellt.

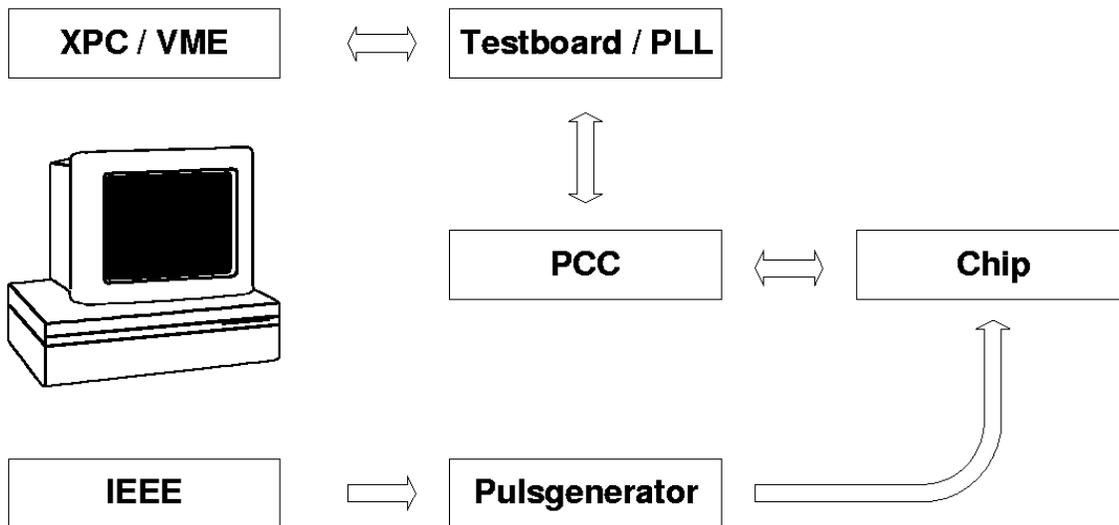
9.1 Der Meßaufbau

Zur Messung der Front-End-Chips stehen zwei verschiedene Testsysteme zur Verfügung, die, je nach Art der Tests, ihre Stärken und Schwächen aufweisen und daher parallel benutzt werden.

Das eine System besteht aus einem PC, der die Steuersequenzen für den FED1 generiert und die zurückkommenden Daten empfängt, bzw. speichert und verarbeitet. Die dazu benötigten Routinen wurden in der Programmiersprache C/C++ geschrieben. Über eine XPC-Karte werden die Signale an ein 'Testboard' [60] geleitet, das als Kommunikationsschnittstelle zwischen dem PC und dem zu testenden Chip dient. Dieses erzeugt mithilfe eines FPGA (**F**ield **P**rogrammable **G**ate **A**rray) die 40 MHz Clock und die Level 1 Triggersignale, bzw. zählt die Anzahl der auf dem Hitbus registrierten Treffer, um als Kontrollmöglichkeit unabhängig von der Auslesearchitektur des Chips messen zu können. Die Weiterleitung der Signale vom 'Testboard' an den Chip übernimmt die PCC (**P**ixel **C**ontrol **C**ard) [61], die außerdem auch die Versorgungsspannungen VDDD, VDDA, VCCA für den FED1 erzeugt. Diese können über Potentiometer einzeln eingestellt und die zugehörigen Stromaufnahmen gemessen werden. Die Erzeugung von externen Spannungsstufen zur Ladungsinjektion über die Injektionskapazitäten des Chips geschieht mittels eines Pulsgenerators, der über eine IEEE-Schnittstelle vom PC angesteuert werden kann.

Das zweite Meßsystem enthält statt der XPC-Karte und dem 'Testboard' eine VME-Schnittstelle und eine PLL (**P**ixel **L**ow **L**evel Card), wobei die zugehörige Software ebenfalls eine andere als beim ersten System ist.

Abbildung 48: Skizze des Meßaufbaus



Der Vorteil des VME-basierten Systems besteht in der höheren Geschwindigkeit bei der Durchführung der Messungen, bietet allerdings im Gegensatz zum anderen System kaum Möglichkeiten zur Fehlerdiagnose auf elementarem Niveau.

Schließlich muß noch die physikalische Verbindung zum Chip selbst hergestellt werden. Der Hersteller liefert diese in Form von dünnen Siliziumscheiben ('Wafer') mit einem Durchmesser von 6 Zoll. In diesem Fall wird der 'Wafer' zum Messen durch ein Vakuum auf einer 'Probestation' gehalten und mit einer Nadelkarte, die genau 48 auf die Anschlußfelder des Chips passende Nadeln enthält, kontaktiert. Zusätzlich befinden sich darauf noch einige Filter und Treiber für Versorgungsspannungen und Signale.

Weiterhin besteht über eine SCS (Single Chip Support) Karte die Möglichkeit einzelne, aus dem 'Wafer' herausgeschnittene Chips durchzumessen. Dazu werden die 48 Anschlüsse des FED1 mithilfe von 'Wire Bonds' mit den Leiterbahnen der Karte verbunden und der Chip selbst auf die Karte geklebt.

Diese Methode hat den Vorzug, daß der Aufbau insgesamt rauschärmer ist als bei der Verwendung der 'Probestation'. Andererseits sind auf diese Weise kontaktierte Chips nicht mehr weiter verwendbar.

9.2 Wafermessungen

Die Wafermessungen an den FED1 Chips liefern Ergebnisse, die Erklärungen bedürfen, welche Schritt für Schritt mühsam herausgefiltert werden müssen. Zunächst wird an ein-

zelen Chips auf einem 'Wafer' grob untersucht, welche Blöcke funktionieren und wo Probleme liegen, um dann systematisch die gesamte Produktion statistisch zu vermessen und geeignete Kandidaten für die weiteren Einzelmessungen zu identifizieren. In den folgenden Abschnitten werden nun die einzelnen Fehlerquellen aufgezählt.

9.2.1 Die Erzeugung von VTH

Die Erzeugung der Diskriminatorspannung VTH ist weder auf dem Chip noch extern über das 'Eingangspad' möglich, so daß der Diskriminator nicht betrieben werden kann und damit schon von vorneherein kein einziger Chip bei ATLAS einsetzbar ist. Die Ursache liegt in einem Fehler im Verifikationsprogramm, das vom Hersteller zur Verfügung gestellt wird und resultiert darin, daß im Layout zwei Kondensatoren systematisch kurzgeschlossen sind, während der entsprechende Schaltplan korrekt ist. Diese Diskrepanz hätte jedoch durch die LVS Überprüfung entdeckt werden müssen. Da zur Umgehung dieses Problems Modifikationen mit fokussiertem Ionenstrahl FIB (**F**ocussed **I**on **B**eam) notwendig sind, die man nur an Einzelchips vornehmen kann, werden bei den Wafertests nur digitale Messungen durchgeführt und die analoge Versorgung ausgeschaltet.

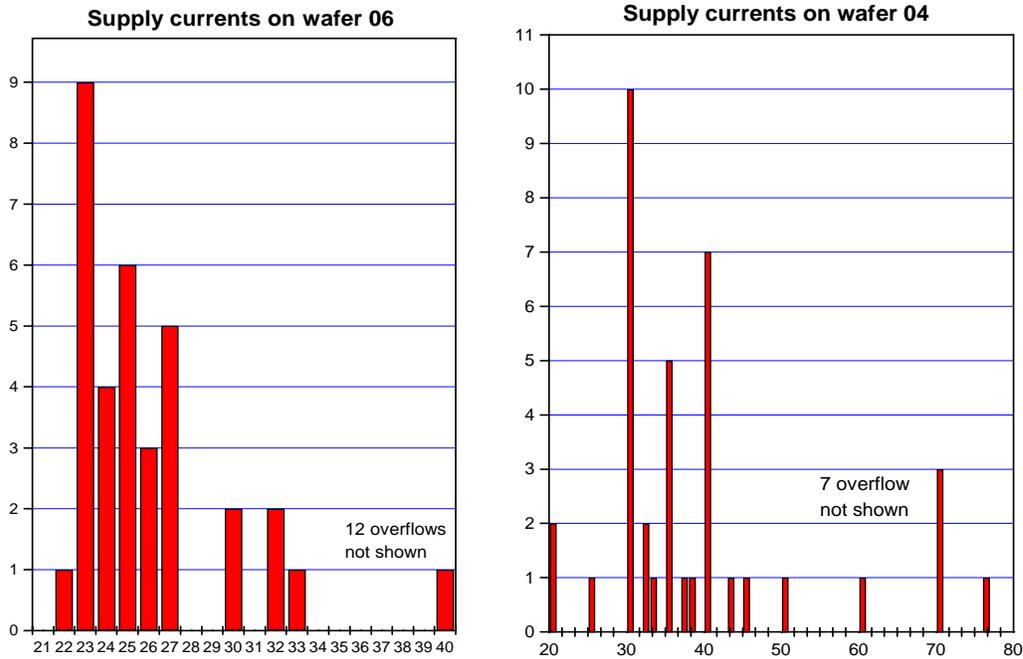
9.2.2 Der digitale Stromverbrauch

Die Größe des Stromverbrauchs im digitalen Teil der Chips gibt die ersten Hinweise auf die Homogenität einzelner 'Wafer' bzw. Variationen unterschiedlicher 'Wafer' und den 'Yield'. Exemplarisch sind in Abbildung 49 die Werte für zwei 'Wafer' histogrammiert.

Man sieht, daß bei 'Wafer' 06 die meisten Chips im Bereich $20\text{ mA} - 30\text{ mA}$ liegen, was auch den aus der Simulation erwarteten Werten entspricht. 'Wafer' 04 dagegen zeigt eine deutlich breitere Verteilung mit insgesamt höheren Strömen.

Dieses Ergebnis zeigt, daß bei der Produktion einige Parameter nicht unter guter Kontrolle sind und der 'Yield' stark vom jeweiligen 'Wafer' abhängt.

Abbildung 49: Der digitale Stromverbrauch der Chips auf zwei verschiedenen 'Wafern'



9.2.3 Das globale Register

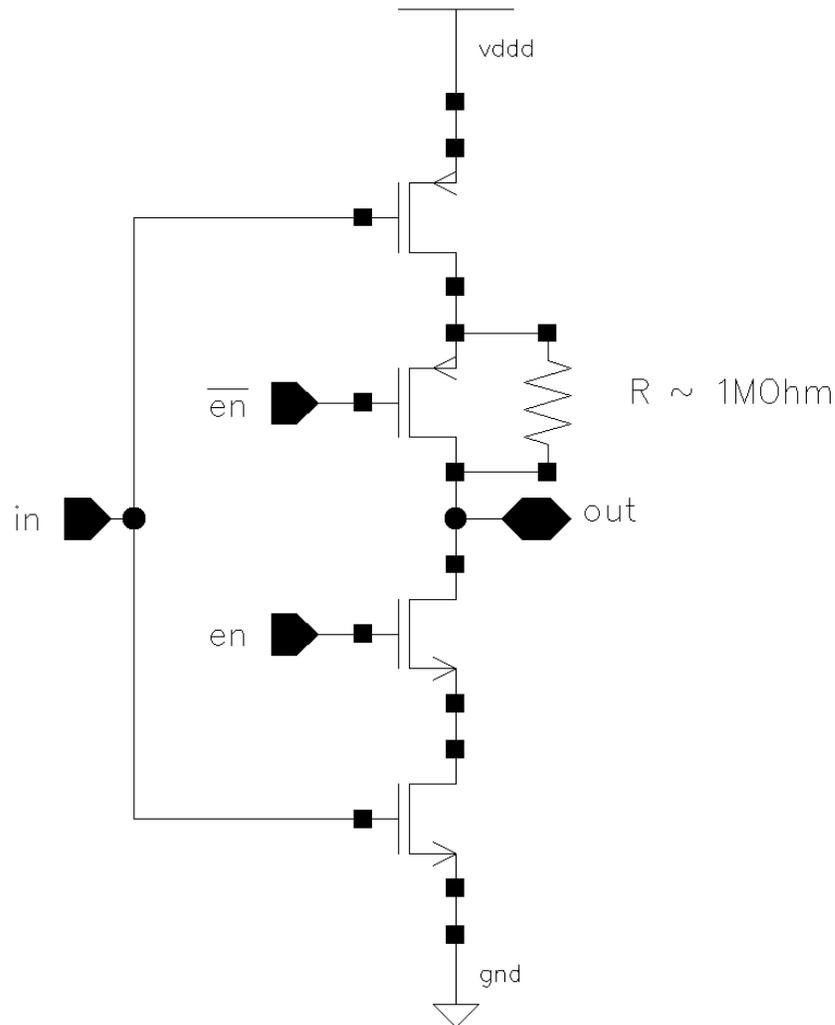
Der Test des globalen Registers erfolgt nach folgendem Prinzip: Man taktet ein bekanntes Datenmuster hinein, führt die READBACK Prozedur aus und taktet den Registerinhalt wieder nach außen. Beim Vergleich der gesendeten und empfangenen Daten sollte sich dann keine Diskrepanz zeigen, wenn das Register einwandfrei funktioniert. Damit hat man gleichzeitig auch sichergestellt, daß das Kommandoregister funktioniert, weil sonst kein Ansprechen des Chips möglich ist. Das Ergebnis zeigt bei 'Wafer' 06, daß von 45 getesteten Chips 39 ein intaktes globales Register besitzen, während dies bei 'Wafer' 04 nur deren 28 sind. Die Vermutung von Abschnitt 9.2.2 bestätigt sich also.

9.2.4 Das Pixelschieberegister

Analog zum vorigen Test verfährt man mit dem Pixelschieberegister, außer, daß hier keine READBACK Prozedur erfolgt, weil sie in diesem Register nicht vorhanden ist. Von den Chips mit intaktem globalen Register bleiben nach dieser Messung noch 13 bzw. 8 Chips

auf 'Wafer' 06 bzw. 'Wafer' 04 übrig, die in allen Doppelspalten ein funktionierendes Pixelschieberegister enthalten. Dabei tritt in etwa 5 - 10 % aller Fälle ein besonderer Defekt auf: Beim Vergleich der Eingangs- und Ausgangsdaten stellt man fest, daß in jedem Block aus logischen '1'en jeweils die letzte '1' verloren geht. Dieses Phänomen kann fast immer auskuriert werden, wenn man die Taktfrequenz erhöht, was den Verdacht aufkommen läßt, daß es teilweise Schwierigkeiten mit dem dynamischen Knoten im semistatischen Register gibt. Dort wird lediglich während der Schiebephase der Registerinhalt auf der Eingangskapazität eines CMOS-Inverters gespeichert. Allerdings ist bei korrekter Funktionsweise aller Transistoren und der normalen Taktfrequenz von 5 MHz die Zeit zu kurz, um den dynamisch gespeicherten Registerwert durch Leckströme zu entladen. Zudem sollte dann auch der Verlust von logischen '0'en erfolgen. Der einzig mögliche Ansatz zu einer Erklärung liegt darin, daß der in der unteren Abbildung gezeigte Inverter einen hochohmigen Kurzschluß im Bereich einiger $M\Omega$ aufweist, und zwar genau an der unten gezeigten Stelle. Interpretieren läßt sich dieser Kurzschluß einfach als ein PMOS, der im abgeschalteten Zustand bis zu $1 \mu A$ Strom fließen läßt, also ein defekter Transistor.

Abbildung 50: Der Kurzschluß im Pixelschieberegister



Simulationen können diesen Effekt exakt reproduzieren. Zusätzlich zeigen diese, daß allgemein in dieser Hinsicht defekte Transistoren nur Auswirkungen auf dynamische Schaltungen haben können. Fehler in statischer Elektronik können nur sichtbar werden, wenn diese Kurzschlüsse etwa 2 - 3 Größenordnungen niedriger, also im 10 $k\Omega$ Bereich liegen. Die logische Folge ist also, daß die Möglichkeit besteht, daß sehr viel mehr Transistoren kaputt sind, aber nicht explizit entdeckt werden können. Zu dem Problem des niedrigen 'Yields' gesellt sich nun also vermutlich noch ein prinzipielles Problem der DMILL Technologie.

9.2.5 Die Ausleseketten

Zur Überprüfung der Auslesearchitektur wird über den auf den Chips implementierten Testmechanismus in jeden einzelnen Pixel nacheinander digital eine feste Anzahl von Treffern einer bestimmten Länge injiziert und am Datenausgang ausgelesen. Im Normalfall sollten dann, wenn das Level 1 Signal korrekt gesendet wird, genau die Trefferdaten wieder am DO sichtbar werden.

Diese Messung zeigt zunächst einmal, daß man erst gültige Daten am Ausgang beobachten kann, wenn man die digitale Versorgungsspannung VDDD auf mindestens 4,8 V erhöht, was den Simulationsergebnissen, wonach der Chip bei 3 V einwandfrei funktionieren sollte, widerspricht. Anhand von Messungen an einzelnen Chips (s. Abschnitt 9.3) kann dieses Phänomen jedoch geklärt werden.

Zum anderen tritt in diesen Messungen ein breites Spektrum verschiedener Fehler auf:

- **Defekte Adressenbits**

Innerhalb eines Spaltenpaares feuern beispielsweise nur Pixel mit gerader Zeilennummer, was entweder darauf beruht, daß die zugehörige Datenleitung auf dem Chip unterbrochen ist oder defekte Transistoren im ROM vorliegen.

- **Tote Pixel**

Nach der digitalen Injektion und Sendung eines korrekten Triggers antwortet der betreffende Pixel nicht. Merkwürdigerweise ändert sich auf ein und demselben Chip das Muster der toten Pixel innerhalb der Matrix, wenn man dieselbe Messung mehrmals durchführt. Dieser Effekt steht, wie sich viel später erst herausstellt (s. Abschnitt 9.4) im direkten Zusammenhang mit dem folgenden Punkt.

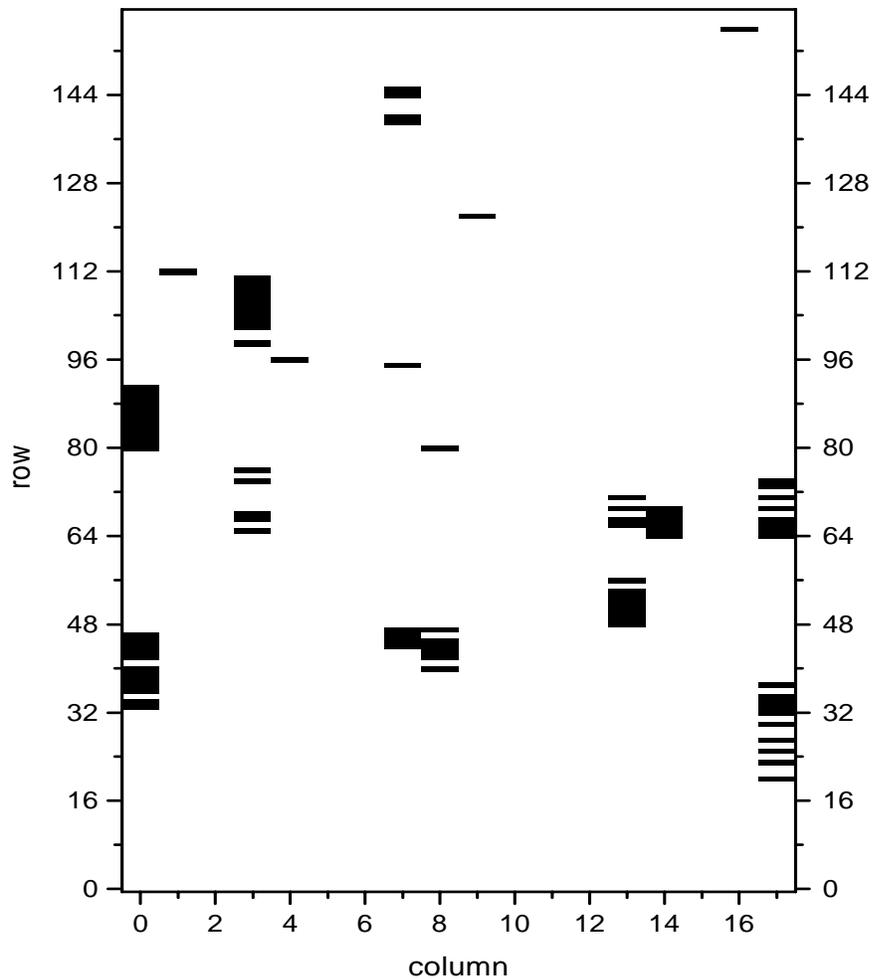
- **Oszillierende Pixel**

Oftmals tritt das Phänomen auf, daß bestimmte Pixel bei einer Injektion nicht antworten, aber bewirken, daß in derselben Spalte plötzlich der unterste Pixel eine willkürliche Anzahl von Treffern liefert und zwar mit Zeilennummer und TOT gleich Null. Meistens scheint dann die gesamte Spalte außer dem untersten Pixel tot zu

sein. Ändert man die Auslesefrequenz, so ändern sich auch die Pixel, die diese Eigenschaft besitzen. Dieses Problem kann ebenfalls nach langwierigen Simulationen und Messungen an FIB modifizierten Einzelchips aufgeklärt werden (s. Abschnitt 9.4).

Unter Berücksichtigung dieser Defekte bleibt zum Schluß kein einziger Chip auf sechs 6 Zoll 'Wafeln' übrig, der akzeptabel ist. Allerdings beweist folgende Trefferkarte eines der besten gefundenen Chips, in der die Pixelmatrix dargestellt ist und alle defekten Pixel in schwarz gehalten sind, daß die Auslesearchitektur prinzipiell funktioniert. Ansonsten würden auf diesem Chip nicht etwa 2800 der 2880 Pixel einwandfrei ihre Daten ausgeben. Das heißt, vom Gesichtspunkt des Designs gesehen, muß die implementierte Logik in Ordnung sein.

Abbildung 51: Die Trefferkarte eines der besten FED1 Chips



9.3 Messungen an Einzelchips

Zur weiteren Untersuchung der bisher unverstandenen Phänomene werden die besten gefundenen Chips aus den 'Wafern' herausgeschnitten und einzeln vermessen. Um auch Informationen über die analogen Schaltungsblöcke zu gewinnen, wird auf diesen Chips mittels FIB der Generator der Spannung V_{TH} isoliert, indem dessen Zuleitung an einer geeigneten Stelle durchgeschnitten wird. Nun kann diese Spannung durch ein externes Netzgerät über das zugehörige 'Eingangspad' eingestellt werden.

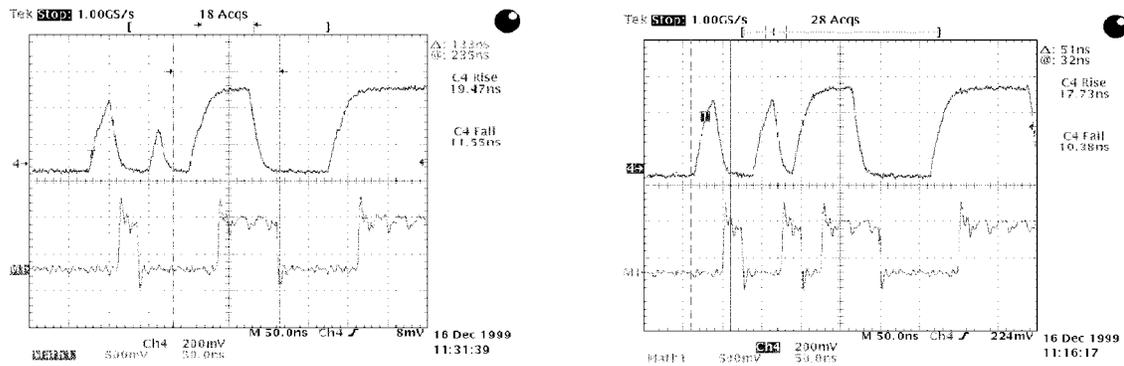
9.3.1 Messung interner Signale

Neben den 48 Anschlüssen für den Chip wurden in weiser Voraussicht zu diagnostischen Zwecken rund um jeden Chip kleine 'Metallpads' im Layout verteilt, die mit internen Signalleitungen verbunden wurden. Mithilfe von aktiven 'Probenadeln', die einen Verstärker in der Spitze enthalten, können diese 'Pads' kontaktiert und die zeitlichen Verläufe der Signale auf einem Oszilloskop visualisiert werden. Eine besondere Eigenschaft dieser aktiven Nadeln besteht in der geringen Eingangskapazität von $0,1 \text{ pF}$ [62], so daß sogar dynamische Knoten nicht durch diese Kapazität entladen werden können und somit ebenfalls meßbar sind.

Das Ergebnis dieser Messungen zeigt, daß alle Signale generell mehr oder weniger deutlich höhere Anstiegszeiten benötigen, als dies in den Simulationen, die aus dem Layout extrahierte parasitäre Effekte mitberücksichtigen, der Fall ist. Bei fast allen Signalen stellt dieser Effekt jedoch kein Problem für die Funktion des Chips dar, weil deren Geschwindigkeit trotzdem noch ausreichend ist, mit einer Ausnahme: Der Ausgang des Serialisierers. Dieser zeigt eine Anstiegszeit von 18 ns , was bei einer Taktfrequenz von 40 MHz offensichtlich zu langsam ist. Die Folge ist, daß beim Heraustakten der Daten einzelne Bits nicht korrekt übertragen werden und damit bei der Trefferauslese unsinnige Informationen an den Chipausgang gelangen. Dieses Problem kann behoben werden, indem man die digitale Versorgungsspannung auf mindestens $4,8 \text{ V}$ erhöht, um diesem speziellen Signal etwas mehr Geschwindigkeit zu geben.

In Abbildung 52 ist oben der Ausgang des Serialisierers, unten der Datenausgang des Chips, gemessen mit einem Oszilloskop, für zwei unterschiedliche digitale Versorgungsspannungen dargestellt. Im linken Bild ist erkennbar, daß V_{DDD} mit $4,5 \text{ V}$ offenbar noch zu niedrig ist, um die Daten richtig zu übertragen, während rechts bei $V_{DDD} = 4,8 \text{ V}$ alles korrekt verläuft.

Abbildung 52: Das Ausgangssignal des Serialisierers (oben) verglichen mit dem Chipausgang (unten) bei 4,5 V (links) bzw. bei 4,8 V (rechts): Links wird eine '1' nicht korrekt übertragen



Für die systematisch zu geringen Signalgeschwindigkeiten bleiben im Prinzip nur zwei Erklärungen übrig: Entweder liefern die Transistoren weniger Strom, als es die Simulationen vorhersagen, oder die parasitären Kapazitäten sind größer, als es aus den Angaben der Herstellungsparameter ersichtlich ist.

Erstere Ursache wird anhand von Messungen an Teststrukturen überprüft; diese geben jedoch keinen Hinweis darauf.

Also bleibt nur letztere Möglichkeit, die an einer Untersuchung am LBNL bestätigt wird. Der Ausgang des Serialisierers ist über eine etwa 1 mm lange Metalleitung mit dem Ausgangsmultiplexer verbunden, die laut Herstellerdaten eine parasitäre Kapazität von etwa 1 pF darstellen soll. Um die beobachtete Langsamkeit dieses Signals erklären zu können, benötigt man jedoch etwa 3 pF, was eine nicht akzeptable Diskrepanz zum Herstellerwert bedeutet. Am LBNL wird diese Leitung mittels FIB durchtrennt und ein 'Metallpad' angebracht, so daß der Ausgang des Serialisierers mit einer Nadel gemessen werden kann. Nach dieser Modifikation zeigt das Signal auch bei der nominellen Versorgungsspannung von 3 V die erwartete Geschwindigkeit und löst damit auch ein Problem aus Abschnitt 9.2.5.

9.3.2 Analoge Messungen ohne Detektor

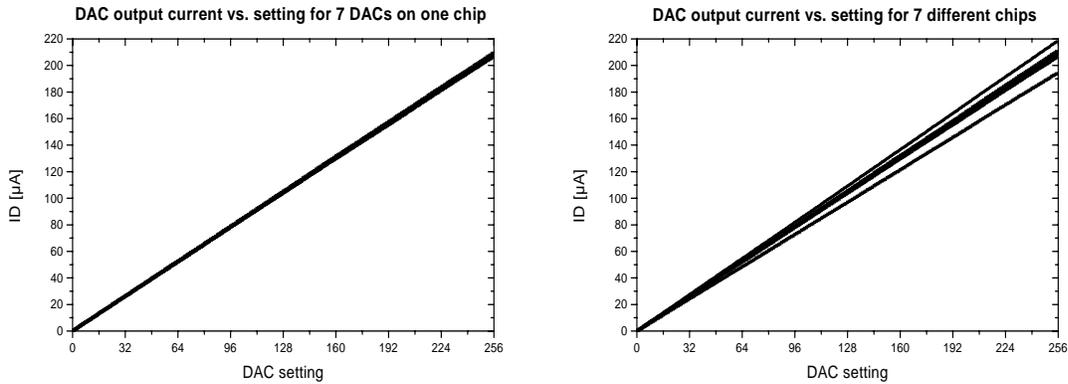
Zur weiteren Charakterisierung der Chips werden nun die analogen Schaltungsblöcke vermessen.

- **Die Strom-DACs**

Die 8 Bit Strom-DACs der Chips werden durchgefahen, um die Monotonie, Linearität und Genauigkeit zu überprüfen. Dabei zeigen alle sieben DACs auf einem Chip

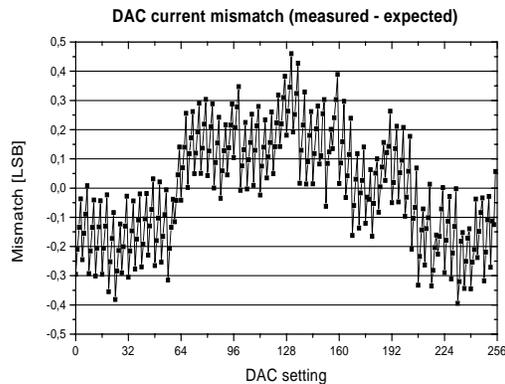
praktisch keine Variation der Ströme untereinander; auch von Chip zu Chip sind die Unterschiede nur gering.

Abbildung 53: Die gemessenen DAC Ströme eines (links) bzw. mehrerer FED1 Chips (rechts)



Die Präzision jedes DACs liegt sogar bei über 9 Bit, wenn man die erwarteten mit den gemessenen Werten vergleicht (s. Abbildung 54).

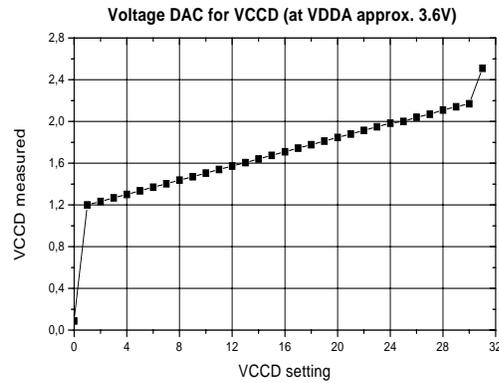
Abbildung 54: Der Unterschied zwischen erwarteten und gemessenen DAC Strömen



- **Der Spannungs-DAC für VCCD**

Analog zum vorigen Abschnitt wird mit dem 5 Bit DAC zur Einstellung der Diskriminatorspannung VCCD verfahren. Die Messung zeigt exakt das erwartete Verhalten (s. Abbildung 55).

Abbildung 55: Die gemessenen DAC Spannungen eines FED1 Chips



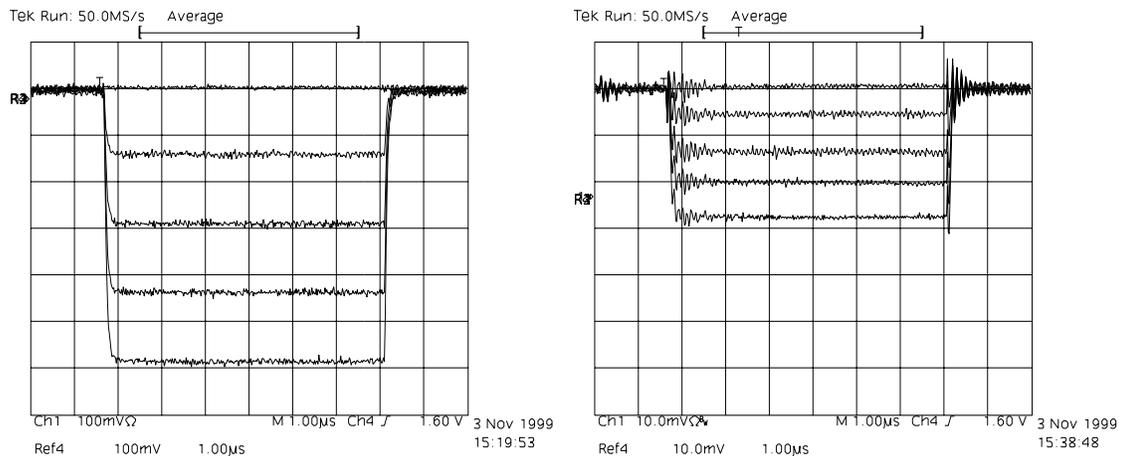
- **Der interne Pulsgenerator**

Der auf dem Chip realisierte interne Pulsgenerator ('Chopper') erzeugt zum Testen des Vorverstärkers steile Spannungssprünge. Die Höhe dieser Stufen sind über einen der Strom-DACs einstellbar und sollen ein lineares Verhalten zeigen. Der Pulsgenerator deckt zwei Bereiche bezüglich der äquivalenten injizierten Ladung ab: Bis etwa $70000 e^-$ bzw. bis $7000 e^-$ in 256 äquidistanten Schritten.

Abbildung 56: Die gemessenen Spannungssprünge des 'Choppers'

Hoher Bereich: Strom-DAC Werte 4, 8, 12, 16 (links)

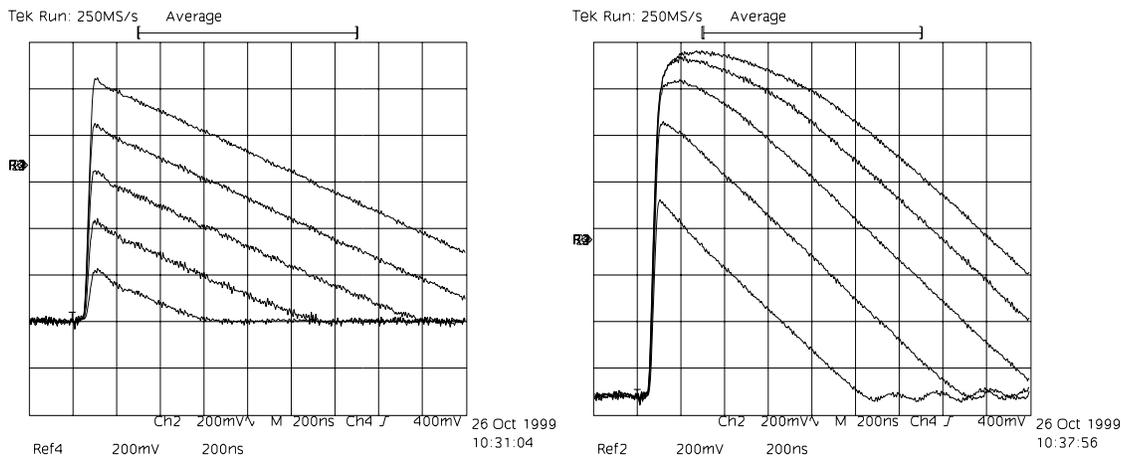
Niedriger Bereich: Strom-DAC Werte 2, 4, 6, 8 (rechts)



- **Der Vorverstärker**

Am Testausgang des rechten unteren Pixels in der Matrix werden die Ausgangssignale des Vorverstärkers für diverse Einstellungen gemessen, um dessen Eigenschaften zu bestimmen.

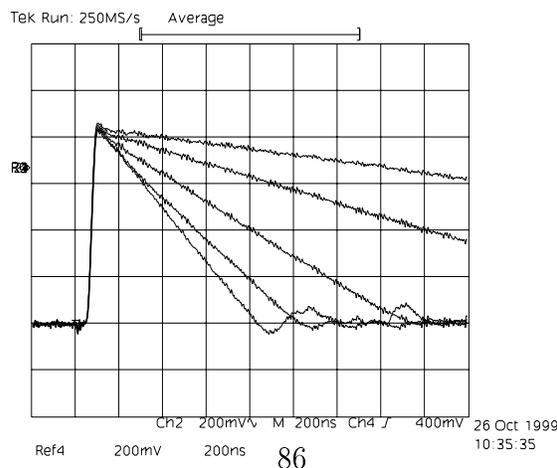
Abbildung 57: Die gemessenen Vorverstärkersignale für verschiedene Ladungsinjektionen: Links wird der 'Chopper'-DAC von unten nach oben auf 16, 32, 48, 64 und 80 eingestellt, wobei der Wert 80 der Ladung entspricht, die ein MIP im Detektor erzeugt. Rechts entsprechen die Kurven den DAC Werten 60, 90, 120, 150 und 180.



Man findet die beabsichtigte Dreiecksform des Ausgangssignals und den linearen Zusammenhang zwischen injizierter Ladung und Signalbreite, der die Voraussetzung für das TOT bildet. Für sehr hohe Ladungen ab etwa $50000 e^-$ geht der Verstärker in Sättigung.

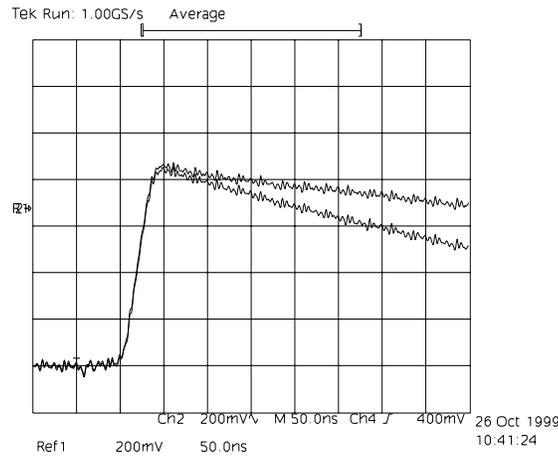
In der nächsten Messung erkennt man das Verhalten des Vorverstärkers für verschiedene Einstellungen des Rückkopplungsstroms IF, der die Zeit bis zur Rückkehr zur Ausgangsspannung bestimmt. Auf diese Weise kann die Totzeit des Pixels verkürzt werden (schnelles 'Shaping') oder die TOT Auflösung verbessert werden (langsames 'Shaping'). Erfreulicherweise sieht man kein ballistisches Defizit ('Shaping Loss') durch Erhöhung von IF, so daß die Verstärkung unabhängig von diesem Parameter ist. Unter dem 'Shaping Loss' versteht man das Phänomen, daß während der Phase, in der der Vorverstärker Ladungssignale aufintegriert, schon wieder eine Entladung durch den Rückkopplungsstrom IF erfolgt, so daß der Vorverstärkerausgang für hohe Werte von IF nicht mehr den Maximalausschlag erreicht.

Abbildung 58: Die gemessenen Vorverstärkersignale für verschiedene Ströme IF: Die DAC Einstellungen betragen von unten nach oben 128, 96, 64, 32 und 16.



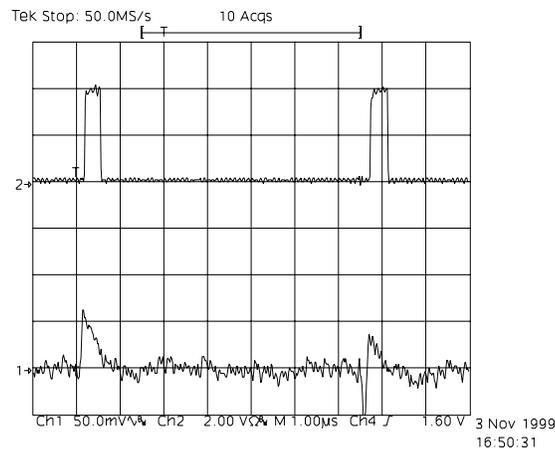
Auch die Anstiegszeit zeigt keine Abhängigkeit vom 'Shaping' und liegt bei etwa 30 ns , was mit der Simulation übereinstimmt.

Abbildung 59: Die gemessene Anstiegszeit für verschiedene Ströme IF: Die DAC Werte von unten nach oben betragen 80 und 40.



Zur Überprüfung, ob bei einem Treffer der Diskriminator feuert, betrachtet man das Hitbussignal, das eine ODER-Verknüpfung aller Diskriminatoren auf dem Chip darstellt. Die untere Abbildung zeigt die gewünschte Korrelation zwischen Vorverstärker und Diskriminator.

Abbildung 60: Die Korrelation zwischen Vorverstärker (unten) und Diskriminator (oben)



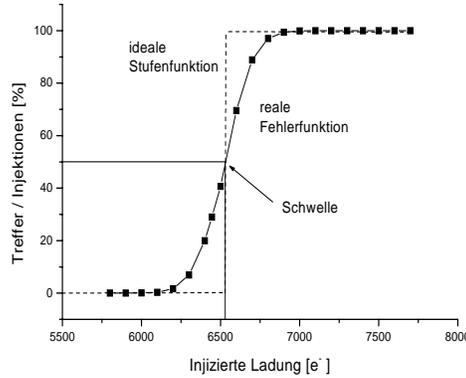
- **Schwellenmessungen**

Die Schaltschwelle des Diskriminators wird durch die Differenz der beiden Spannungen VCCD und VTH eingestellt, wobei eine große Differenz eine hohe Schwelle verursacht. Im idealen Fall sollten durch Ladungsinjektion generierte Vorverstärker-signale, die unterhalb der Diskriminatorschwelle liegen, keinen Einfluß auf denselben haben, während alle darüber liegenden Signale ein Feuern des Diskriminators verursachen sollten. Die graphische Auftragung der Feuerwahrscheinlichkeit des Diskri-

minators in Abhängigkeit von der injizierten Ladung sollte also am Schwellenwert von 0 % auf 100 % springen.

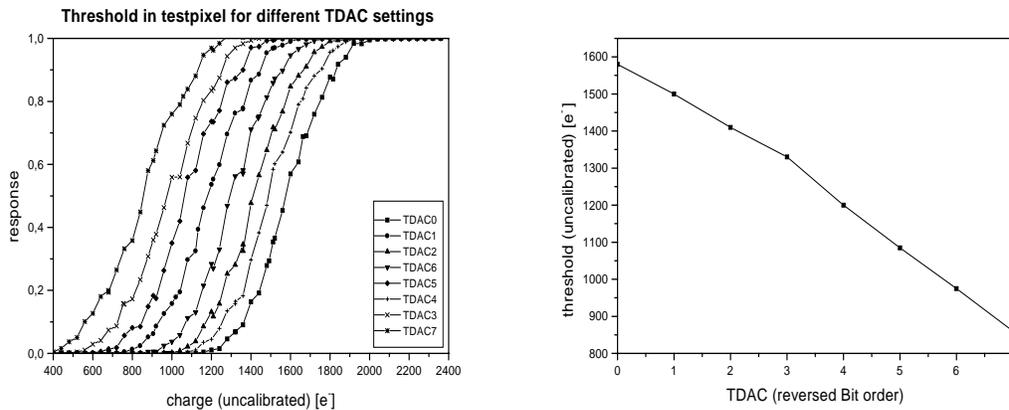
Durch das elektronische Rauschen des Vorverstärkers können jedoch unterhalb der Schwelle liegende Signale bereits detektiert bzw. darüber liegende nicht als Treffer erkannt werden. Dieses Verhalten wird durch eine Funktion beschrieben, in der das gaußförmige Rauschen des Verstärkers mit obiger idealen Stufenfunktion gefaltet wird. Das Ergebnis ist eine sogenannte 'S-Kurve', aus der man am 50 % Wert den Schwellenwert und aus der Steigung in diesem Punkt den Rauschwert ablesen kann.

Abbildung 61: Vergleich zwischen idealer und gemessener Schwellenkurve [40]



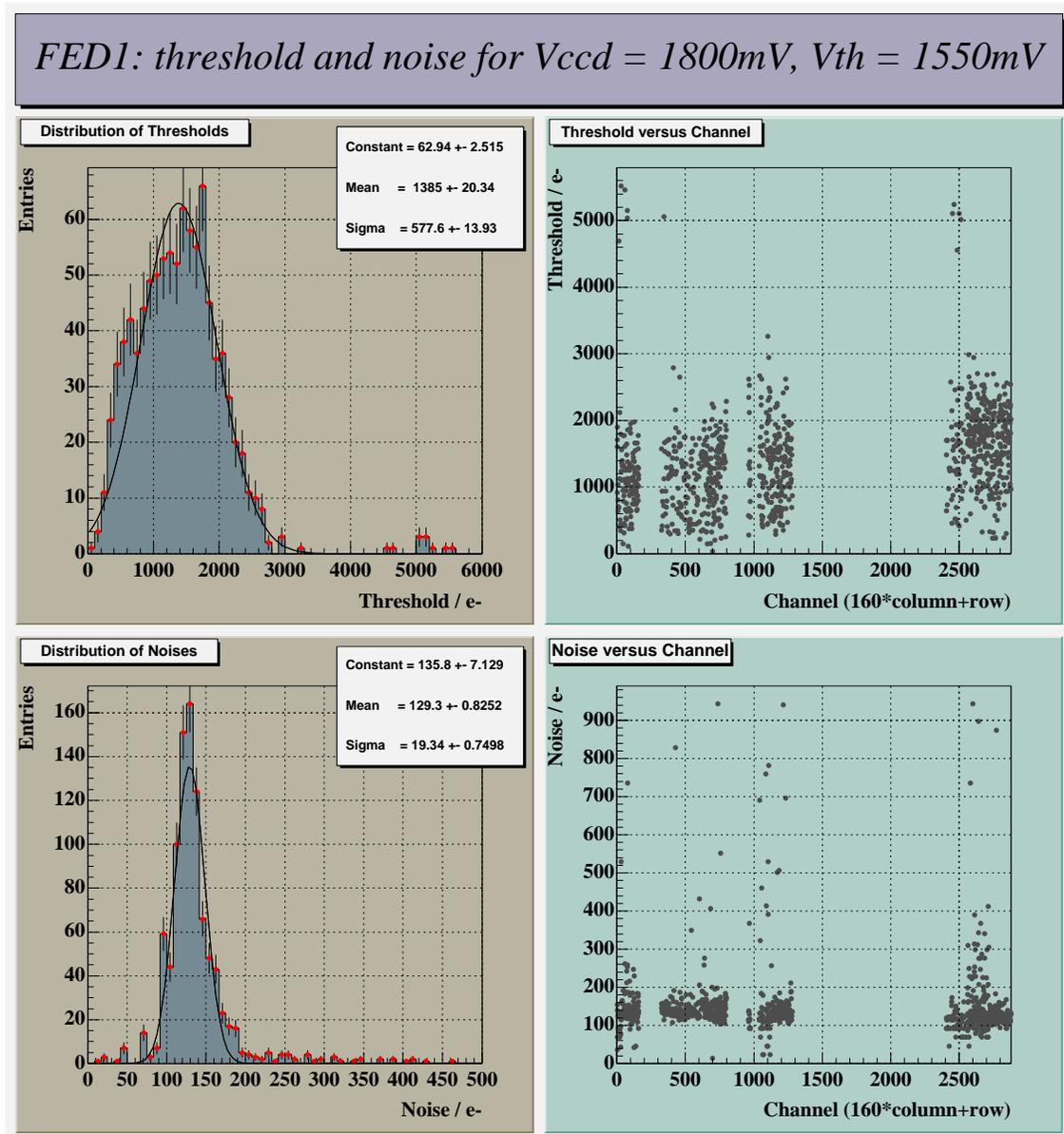
In jedem einzelnen Pixel besteht weiterhin die Möglichkeit, durch die Herstellung oder Verteilung der Spannungsversorgung auf dem Chip bedingte Variationen der Schwellen über einen 3 Bit DAC, dessen Schrittweite durch den Strom ITR eingestellt werden kann, auszugleichen. Dies ist in der unteren Abbildung graphisch dargestellt.

Abbildung 62: Die gemessene Schwelle eines Pixels für verschiedene Einstellungen des 3 Bit DACs mit den zugehörigen 'S-Kurven' (links) und den Auswirkungen auf den Schwellenwert (rechts)



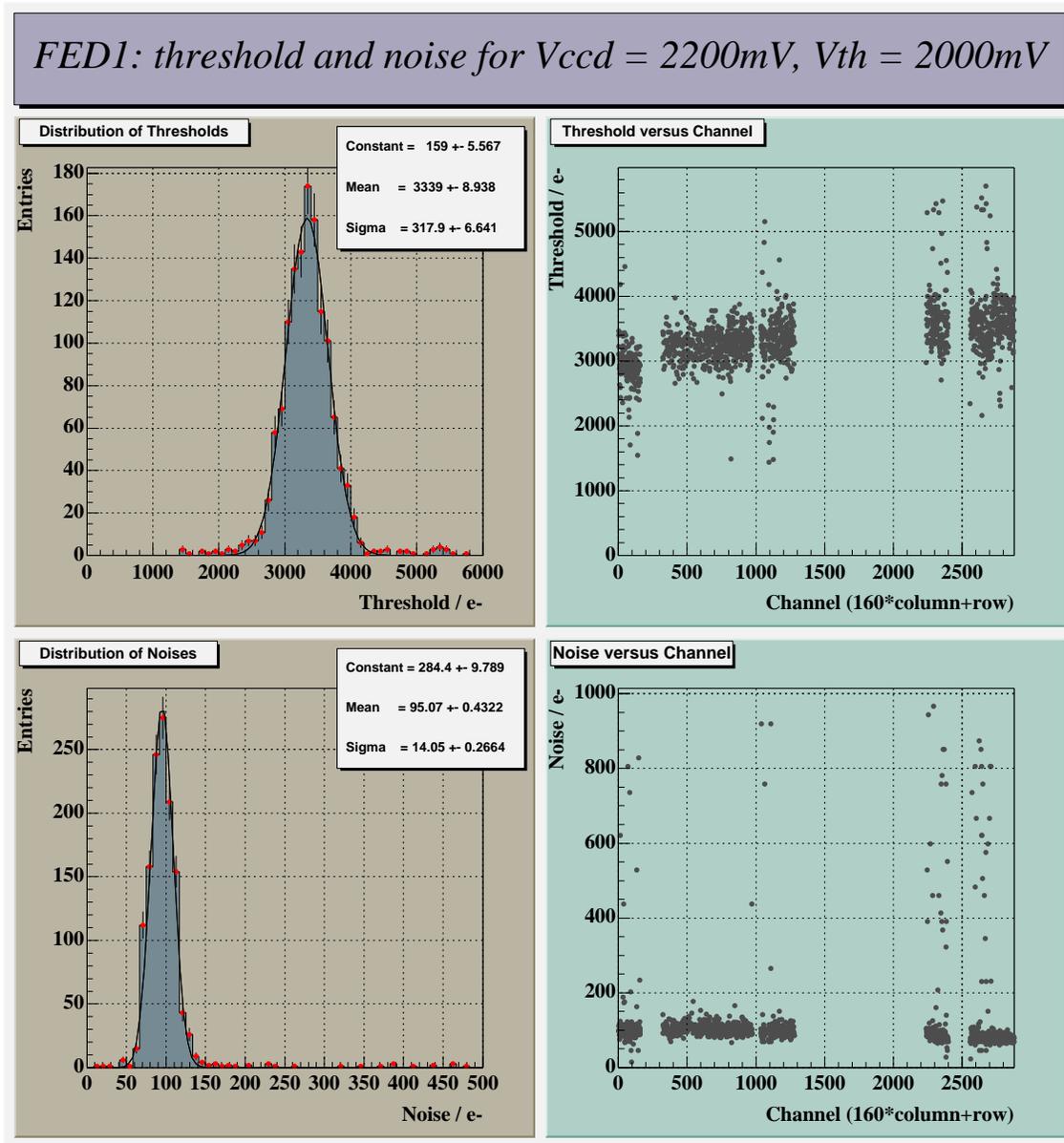
Abschließend wird bei einem kompletten, funktionierenden Chip mit einigen toten Spalten eine Schwellen- und Rauschmessung durchgeführt. Man erkennt, daß die Einstellung einer Schwelle von unter $2000 e^-$, wie sie bei ATLAS gefordert wird (s. Abschnitt 8.1), prinzipiell möglich ist, während die verlangte Schwellendispersion mit fast $600 e^-$ deutlich über dem Wert von $200 e^-$ liegt. Eine Feinjustierung der Schwellen wird bei keinem Chip systematisch vorgenommen, weil die FED1 Chips sich so instabil über längere Meßzeiträume verhalten, daß sich dies als unmöglich erweist. Das elektronische Rauschen erfüllt mit etwa $130 e^-$ die Anforderungen von ATLAS.

Abbildung 63: Schwellen- und Rauschmessung eines FED1 Chips bei niedriger Schwelle



Auch bei einer höheren Schwelleneinstellung bleibt die Dispersion noch bei über $300 e^-$, während sich das Rauschen sogar noch unter $100 e^-$ verbessert.

Abbildung 64: Schwellen- und Rauschmessung eines FED1 Chips bei höherer Schwelle

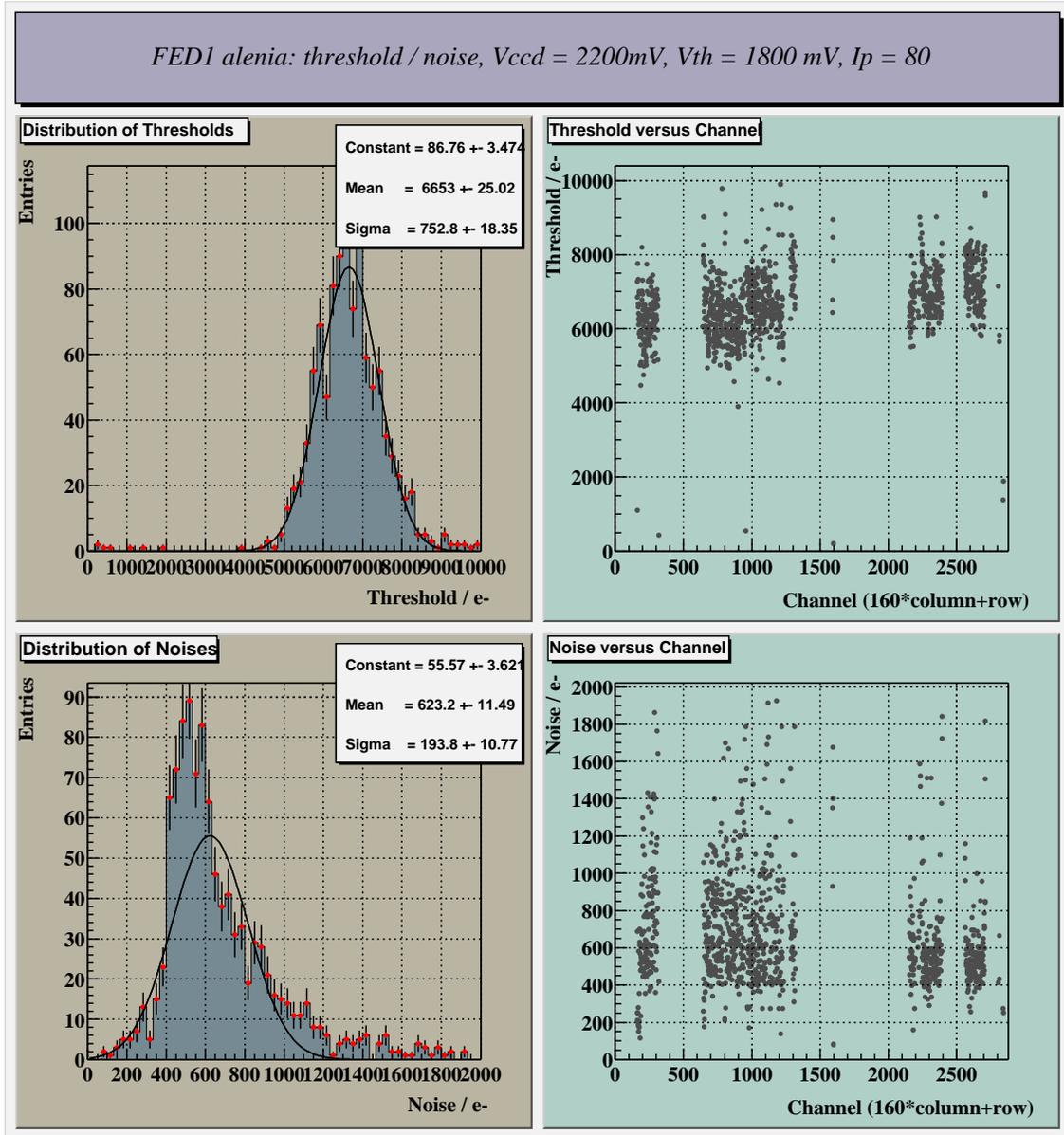


9.3.3 Analoge Messungen mit Detektor

Im nächsten Schritt wird ein FED1 Chip, der an einen Detektor gebondet wurde, vermessen, um die Eigenschaften eines kompletten Hybrids mit denen eines nackten ICs vergleichen zu können.

Bei der unten abgebildeten Schwellenmessung wird eine etwas höhere Schwelle von etwa $6600 e^-$ eingestellt, um ein stabiles Verhalten des Chips zu erzielen. Sowohl das elektronische Rauschen mit über $600 e^-$ als auch die Schwellendispersion mit etwa $750 e^-$ liegen in für ATLAS inakzeptablen Größenordnungen und weit höher als bei Messungen an Chips ohne Detektor.

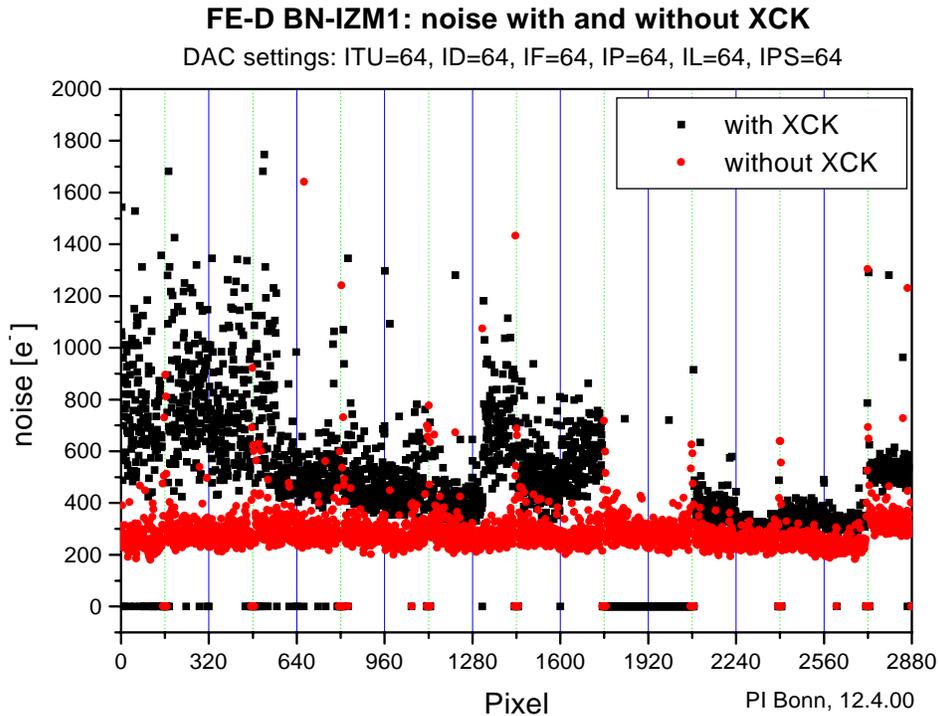
Abbildung 65: Schwellen- und Rauschmessung eines FED1 Chips mit Detektor



Dieses Problem kann nur von der kapazitiven Kopplung zwischen Chip und Detektor herrühren. Vermutlich sprechen schnelle digitale Signale oder Spannungsspitzen auf der digitalen Versorgungsspannung auf den Sensor über, obwohl eine komplette Metallage im Layout zur Abschirmung verwendet wird. Zur Verifikation dieser These wird eine Rauschmessung im normalen Betrieb des Chips vorgenommen und in einer weiteren Messung die $40 MHz$ XCK Clock abgeschaltet, um bei gleichen Einstellungen einen Vergleich ziehen

zu können. Das Ausschalten der Clock bewirkt, daß der Graycodegenerator inaktiv wird und somit die Busse, die dessen Ausgänge ('Timestamps') in die Pixelmatrix führen, ruhig bleiben. Weiterhin wird der Serialisierer abgestellt, so daß man das Hitbussignal zur Messung verwenden muß.

Abbildung 66: Rauschmessung eines FED1 Chips mit und ohne XCK



In obiger Abbildung erkennt man deutlich den Unterschied zwischen ein- und ausgeschalteter Clock. Durch die Analyse der zugehörigen Daten ergibt sich mit laufender XCK ein Rauschen von $532 \pm 214 e^-$, während das Rauschen sich im anderen Fall mit $287 \pm 102 e^-$ drastisch verringert.

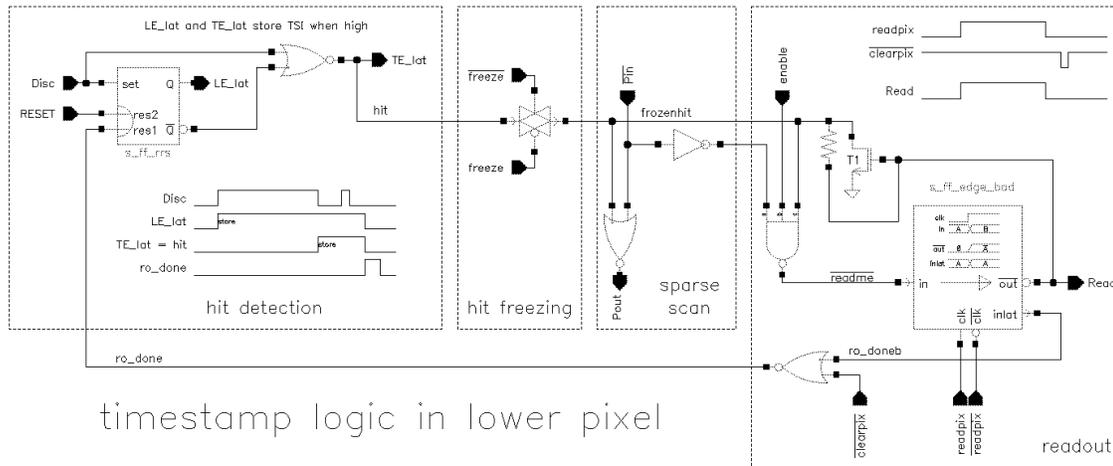
In diesem Fall reicht es daher offenbar nicht aus, eine Metallage im 'Layout' des Chips als Abschirmung gegen den Sensor zu verwenden. Mehr Möglichkeiten bietet der DMILL Prozeß jedoch nicht, so daß dieser Nachteil im Hinblick auf die ATLAS Anforderungen bzgl. des Rauschens und der Schwellenvariation ein weiteres ernstes Problem darstellt.

Als positives Ergebnis läßt sich die Tatsache anführen, daß alle analogen Schaltungsteile erwartungsgemäß funktionieren, wenn auch nicht in jedem Pixel, was wiederum als 'Yield' Problem zu deuten ist.

9.4 Messungen an modifizierten Einzelchips

Dieser Abschnitt liefert die Erklärung der Problematik, die in der Ausleseketten bzgl. oszillierender Pixel auftritt (s. Abschnitt 9.2.5). Dazu sind einige Vorüberlegungen notwendig. In Abschnitt 8.2.2 wurde bereits erklärt, wie die Auslese in FED1 funktioniert. Der zugehörige Schaltplan der Logik im Pixel ist in der unteren Abbildung dargestellt.

Abbildung 67: Der Schaltplan der Ausleselogik im Pixel



Im Normalfall liegen 'frozenhit', 'readpix' und 'clearpix' auf logisch '0', bis der Diskriminator feuert. Mit der fallenden Flanke des Treffers gehen 'hit' bzw. 'frozenhit' auf '1' und starten den 'Scanner'. Sobald dieser am Spaltenende ankommt, setzt die CEU 'freeze' auf '1' und öffnet dadurch den angeschlossenen komplementären Schalter. Ab diesem Zeitpunkt wird auf 'frozenhit' dynamisch eine '1' gespeichert, die bei korrekter Funktionsweise gehalten wird, bis die CEU 'readpix' auf '1' setzt und damit die Trefferdaten aus dem RAM bzw. ROM vom Pixel in die EOC transferiert. Gleichzeitig wird 'frozenhit' gelöscht. Durch Senden des 'clearpix' Signals wird dann nur genau der Pixel gelöscht, der gerade ausgelesen wurde. Anschließend werden alle drei Steuersignale wieder auf '0' gesetzt.

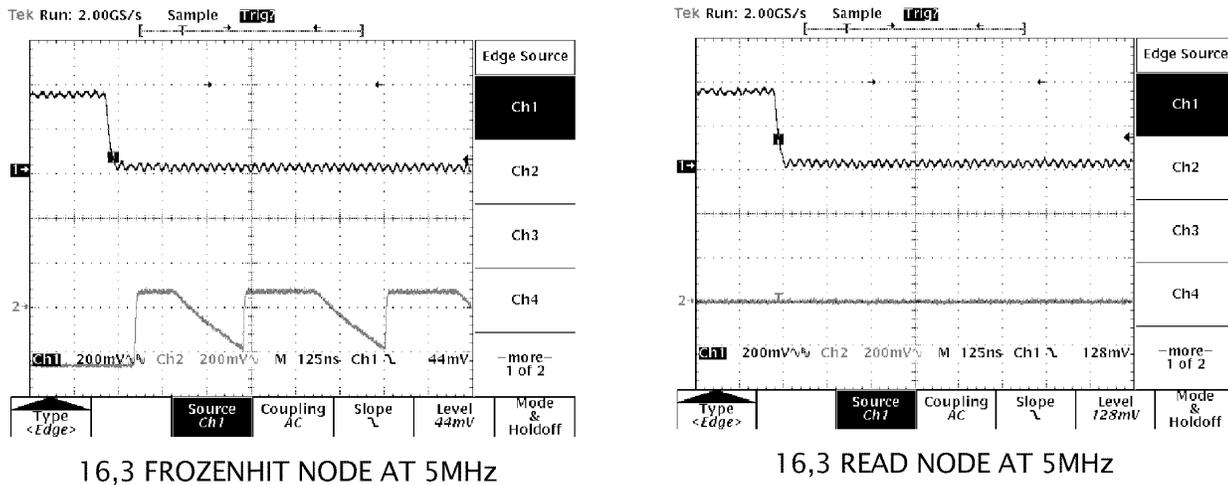
Nimmt man nun an, daß ähnlich wie im Pixelschieberegister die dynamische Speicherung eines Signals verantwortlich für die Oszillationen ist, bleibt als einzige Fehlerquelle nur der NMOS T1, der das Löschen von 'frozenhit' vornimmt, übrig. Wenn im Falle eines Treffers nach der Sendung von 'freeze' die auf 'frozenhit' dynamisch gespeicherte '1' über einen hochohmigen Kurzschluß, der wiederum im Bereich $M\Omega$ oder niedriger liegt, auf '0' gezogen wird, würde die CEU trotzdem 'readpix' auf '1' setzen und versuchen, einen durch den 'Scanner' markierten Pixel auszulesen. Da aber nun 'frozenhit' nicht mehr aktiv ist, zeigt der 'Scanner' auf keinen Pixel, und es werden Nullen als Adresse und TOT ausgelesen. Auch 'clearpix' wirkt lediglich auf einen so ausgewählten Pixel, so daß der ursprünglich getroffene Pixel nicht gelöscht werden kann und die Auslesesequenz erneut startet. Dieser Kreislauf kann nur durch Zurücksetzen des Chips angehalten werden.

Dieses Phänomen hängt stark von der selektierten Auslesefrequenz bzw. der Stärke des Kurzschlusses ab und kann bei einer Erhöhung dieser Frequenz oder des Kurzschlußwiderstands ausgeheilt werden. Dagegen führt eine Variation in entgegengesetzter Richtung dazu, daß der Pixel komplett tot erscheint, weil sich 'frozenhit' zu schnell entlädt, um noch die Sendung von 'readpix' zu veranlassen. Alle diese Fälle werden für jeweils unterschiedliche Auslesefrequenzen beobachtet, so daß man als Konsequenz schließen muß, daß ein ganzes Spektrum solcher Kurzschlüsse bzw. Transistoren unterschiedlicher Qualität vorliegen.

Diese Theorie kann zunächst durch detaillierte Simulationen und später auch durch direkte Messungen von 'freezepix', 'readpix' und 'clearpix' mithilfe der 'Probepads' exakt in allen Aspekten verifiziert werden.

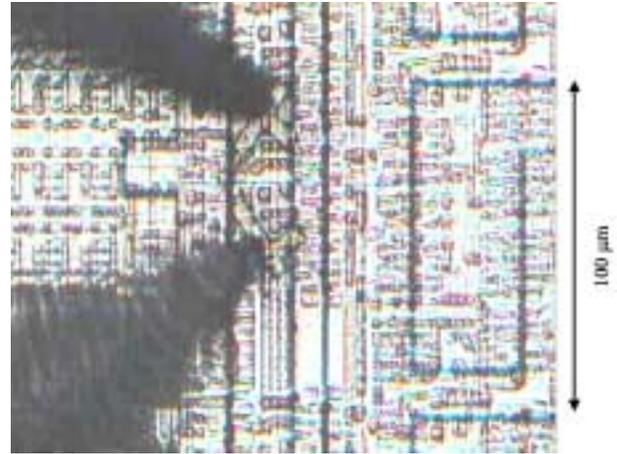
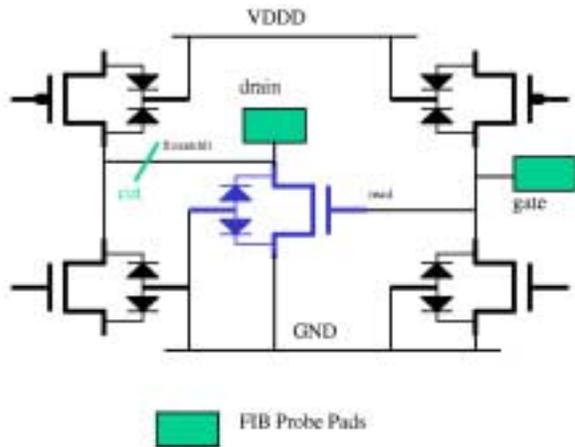
In der unteren Abbildung sind exemplarisch zwei Oszillogramme dargestellt, die das Wegdriften und Oszillieren des 'frozenhit' und das daraus resultierende Ausbleiben des 'readpix' zeigen.

Abbildung 68: Oszillogramme der Auslesesignale eines scheinbar toten Pixels: 'frozenhit' (links) und 'readpix' (rechts)



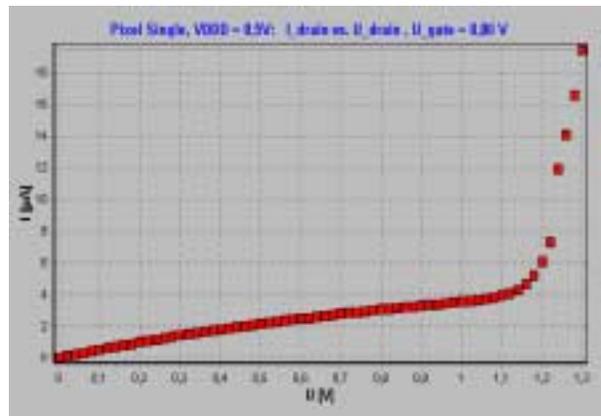
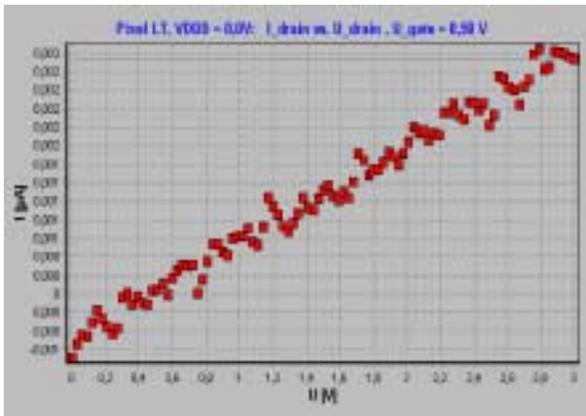
Um noch einen weiteren Beweis für die angestellten Überlegungen liefern zu können, werden an einigen FED1 Chips folgende FIB Modifikationen vorgenommen: Bei defekten bzw. intakten Pixeln wird die 'frozenhit' Leitung durchgeschnitten und 'Metallpads' an 'Drain' und 'Gate' des fraglichen NMOS angebracht.

Abbildung 69: Schaltplan (links) und Foto (rechts) der FIB modifizierten Pixellogik mit 'Probenadeln' und 'Pads' (kreuzförmig)



Unter Verwendung von 'Probenadeln' kann nun durch Messung des 'Drainstroms' und der 'Drain'-'Source'-Spannung der Kanalwiderstand des NMOS ermittelt werden:

Abbildung 70: Die Kanalwiderstände des NMOS in der Ausleselogik in einem 'guten' Pixel (links) und in einem 'schlechten' Pixel (rechts)



Die Messung liefert für einen 'guten' Pixel einen Widerstand von $1\text{ G}\Omega$ bzw. $250\text{ k}\Omega$ für einen 'schlechten' Pixel. Der Anstieg im rechten Bild bei einer Spannung von $U_D - V_{DD} = 0,7\text{ V}$ rührt daher, daß sich eine der 'Bulkdioden' eines angeschlossenen PMOS öffnet.

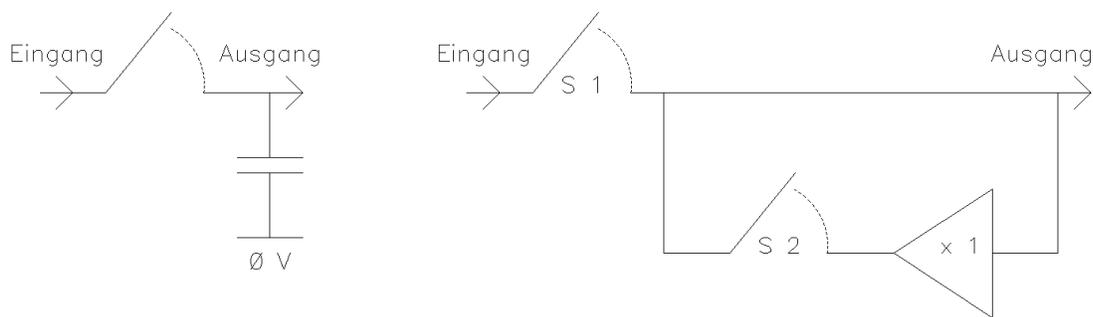
Zusammen mit den vorigen Ergebnissen ist nun zweifelsfrei erwiesen, daß der DMILL Prozeß zumindest in dieser Produktion eklatante Technologiefehler aufweist, die einen Einsatz bei ATLAS nicht zulassen.

10 Das Design des FED2

Aufgrund der ernüchternden Ergebnisse aus Kapitel 9 könnte man DMILL schon für den Einsatz bei ATLAS disqualifizieren. Es besteht jedoch die Hoffnung, daß es sich um eine einmalige Fehlproduktion handelt, so daß man sich entschließt, einen zweiten Versuch zu wagen. Um etwas mehr Gewißheit speziell über die Verwendung von dynamischer Logik zu erlangen, werden bei diesem Design zwei Versionen eines Front-End Chips realisiert: Ein leicht modifizierter FED1, der sogenannte FED2D, und eine 'statische' Alternative, der FED2S.

Der Unterschied zwischen dynamischer und statischer Logik wird nun am Beispiel einer Speicherzelle erläutert:

Abbildung 71: Der Unterschied zwischen dynamischer und statischer Logik am Beispiel einer dynamischen (links) und einer statischen Speicherzelle (rechts)



In der linken Abbildung sei zunächst der Schalter geschlossen. In diesem Fall bewirkt ein Eingangssignal, daß eine entsprechende Ladung auf den Kondensator fließen kann und am Ausgang als Spannung abgegriffen werden kann. Öffnet man nun den Schalter, so bleibt diese Ladung auf dem Kondensator gespeichert und damit auch die zugehörige Spannung. Durch Leckströme wird der Kondensator wieder entladen, weil kein Mechanismus für eine Auffrischung der gespeicherten Ladung vorhanden ist, und man spricht von dynamischer Speicherung.

Zum Vergleich ist rechts eine statische Alternative dargestellt. Zunächst sei der Schalter S1 geschlossen, S2 geöffnet. Das Eingangssignal liegt in diesem Fall direkt am Ausgang vor. S1 wird nun geöffnet, während gleichzeitig S2 geschlossen wird. Das Ausgangssignal frisst sich nun ständig über den 1:1 Verstärker auf und bleibt daher fest gespeichert.

10.1 Der FED2D

Beim FED2D wird im Vergleich zum FED1 logisch nichts verändert. Lediglich an einzelnen Blöcken werden kleine Modifikationen vorgenommen, die nachfolgend aufgelistet werden:

- **Der Kurzschluß im VTH-Generator**

Das in Abschnitt 9.2.1 geschilderte Problem des Kurzschlusses zweier Kondensatoren kann leicht behoben werden, indem diese im Layout in zwei getrennte Wannan verlegt werden.

- **Der dynamische Bereich der Spannungs-DACs**

Die dynamischen Bereiche für die beiden Diskriminatorspannungen werden gemäß der Meßergebnisse angepaßt. Diese sind nun zwischen $1,7 V$ und $2 V$ in $10 mV$ Schritten einstellbar.

- **Treiber für diverse interne Signale**

Aufgrund der in Abschnitt 9.3.1 gemessenen mangelnden Geschwindigkeit der Signale werden im Layout alle Zuleitungen überprüft. Bei denen, die eine Länge von etwa $800 \mu m$ überschreiten, werden, falls ausreichend Platz vorhanden ist, stärkere Treiber implementiert. Dies gilt insbesondere für den Ausgang des Serialisierers, der dafür verantwortlich ist, daß der FED1 nicht bei einer digitalen Versorgungsspannung von $3 V$ betrieben werden kann.

10.2 Der FED2S

Der FED2S unterliegt zusätzlich zu den Verbesserungen, die in FED2D vorgenommen werden, noch einigen weiteren Modifikationen, um zu überprüfen, ob DMILL ein prinzipielles Problem mit dynamischen Schaltungen aufweist, die im FED2S dann nicht mehr sichtbar sein sollten.

- **Das Pixelschieberegister**

Ein limitierender Faktor des 'Yields' liegt in der Funktionstüchtigkeit des Pixelschieberegisters (s. Abschnitt 9.2.4). Aus diesem Grund wird das semistatische 'Flipflop' durch eine statische Version ersetzt, wie sie auch im gut arbeitenden globalen Register verwendet wird.

- **Die Ausleselogik im Pixel**

Die Ergebnisse in Abschnitt 9.4 geben zum Anlaß, auch die Ausleselogik im Pixel als volle statische Version zu implementieren. Dies betrifft aus Platzgründen jedoch nicht den Bereich zur Speicherung der Adressen und 'Timestamps'.

Die Änderungen im Bereich der Pixelmatrix erfolgen allesamt durch die Implementation zusätzlicher Transistoren, so daß der im FED1 schon komplett ausgenutzte Platz nicht dafür ausreicht. Die Konsequenz ist, daß andere Schaltungsteile herausgenommen werden müssen, die zunächst erübrigt werden können, um diese dann bei einer eventuell nachfolgenden Submission gegen die Testfunktionen im Pixel (digitale Injektionsmöglichkeit, Hitbus) austauschen zu können. Die Wahl fällt auf die drei 'Latches' zur Speicherung der Werte für den 3 Bit DAC, der die Feinjustierung der Diskriminatorschwellen vornimmt. Damit ist von vorneherin klar, daß der FED2S nicht bei ATLAS einsetzbar sein wird, sondern ausschließlich Testzwecken dient.

11 Messungen am FED2

Die zweite Submission des Front End Chips erfolgte im Juli 2000. Die Bestellung umfaßte 8 'Wafer', die mit Standard Prozeßparametern hergestellt wurden, und 16 weitere 'Wafer', die einige Variationen aufwiesen. Die Firma Atmel wollte mit diesen 'Testwafern' ihre Produktion überprüfen und Hinweise auf die Ursachen für die defekten Transistoren finden. Diesbezüglich existierte die gemäß den Messungen aus Abschnitt 9.4 sehr unwahrscheinliche Hypothese, daß die 'Gates' der Transistoren nicht korrekt funktionierten. Daher wurden drei Herstellungsparameter verändert, die unterschiedlich miteinander kombiniert wurden:

- Die Gatelänge wurde um $\pm 50 \text{ nm}$ variiert.
- Das Ätzen nach dem Aufbringen des Polysiliziums wurde in zwei Stufen verstärkt.
- Das Ätzen nach der Implementation von Kontakten zwischen Polysilizium und erster Metallage wurde in zwei Stufen verstärkt.

11.1 Wafermessungen

Die 16 prozeßvariierten 'Wafer' werden zusammen mit den 8 Standard 'Wafern' wie in Kapitel 9.2 digitalen Messungen unterzogen, um die Ausbeuten vergleichen zu können. Die Ergebnisse sind in den Tabellen auf den folgenden Seiten zusammengefaßt. Die Kategorie 'Perfekte Chips' beinhaltet Chips, deren Register alle komplett funktionieren und deren Pixel ausnahmslos perfekte Effizienz bei digitaler Injektion aufweisen, während in der Spalte 'Gute Chips' pro Spalte 10 Pixel erlaubt werden, die nicht zu 100% digitale Treffer erkennen.

Man sieht, daß die Ausbeute bei den dynamischen Chips ähnlich enttäuschend ist wie bei der ersten Submission: Es gibt keinen einzigen digital perfekten Chip und nur eine Handvoll akzeptabel funktionierende. Diese Tatsache erübrigt weitergehende analoge Messungen am FED2D.

Weiterhin läßt sich hier auch kein Effekt durch die Prozeßvariationen erkennen, was bedeutet, daß die 'Gates' offenbar nicht für die Existenz defekter Transistoren verantwortlich sind.

Tabelle 6: Digitale Meßergebnisse FED2D, variierte 'Wafer'

FED2D, 43 Chips pro 'Wafer'				
'Wafer'	Gute Chips		Perfekte Chips	
Nummer	Anzahl	Ausbeute [%]	Anzahl	Ausbeute [%]
02	0	0	0	0
03	1	2	0	0
04	2	5	0	0
05	4	9	0	0
06	0	0	0	0
07	1	2	0	0
08	1	2	0	0
10	0	0	0	0
11	0	0	0	0
13	0	0	0	0
15	0	0	0	0
16	2	5	0	0
18	0	0	0	0
19	0	0	0	0
21	1	0	0	0
22	0	0	0	0

Tabelle 7: Digitale Meßergebnisse FED2D, Standard 'Wafer'

FED2D, 43 Chips pro 'Wafer'				
'Wafer'	Gute Chips		Perfekte Chips	
Nummer	Anzahl	Ausbeute [%]	Anzahl	Ausbeute [%]
03	1	2	0	0
04	3	7	0	0
05	0	0	0	0
06	2	5	0	0
07	2	5	0	0
08	1	2	0	0
09	3	7	0	0
10	0	0	0	0

Tabelle 8: Digitale Meßergebnisse FED2S, variierte 'Wafer'

FED2S, 45 Chips pro 'Wafer'				
'Wafer'	Gute Chips		Perfekte Chips	
Nummer	Anzahl	Ausbeute [%]	Anzahl	Ausbeute [%]
02	20	44	4	9
03	18	40	7	16
04	22	49	5	11
05	19	42	2	4
06	18	40	10	22
07	16	36	1	2
08	20	44	0	0
10	2	4	1	2
11	19	42	10	22
13	18	40	0	0
15	15	33	6	13
16	2	4	2	4
18	19	42	8	18
19	15	33	0	0
21	7	16	5	11
22	0	0	0	0

Tabelle 9: Digitale Meßergebnisse FED2S, Standard 'Wafer'

FED2S, 45 Chips pro 'Wafer'				
'Wafer'	Gute Chips		Perfekte Chips	
Nummer	Anzahl	Ausbeute [%]	Anzahl	Ausbeute [%]
03	23	51	4	9
04	28	62	15	33
05	29	64	8	18
06	19	42	1	2
07	27	60	4	9
08	23	51	9	20
09	17	38	3	7
10	22	49	1	2

Ein signifikanter Unterschied offenbart sich jedoch, wie erwartet, im Vergleich zwischen FED2D und FED2S. Die 'statische' Version liefert immerhin eine Ausbeute guter Chips um 40% bei den variierten 'Wafern' bzw. sogar zwischen 40% und 60% bei den Standard 'Wafern'.

Dabei muß man bemerken, daß die 'Testwafer' 02 bis 06 nominell ebenfalls nach Standardparametern hergestellt wurden, so daß diese den Standard 'Wafern' entsprechen sollten. Dies trifft im Rahmen der vorliegenden Statistik auch zu, wenn man die stark schwankenden Zahlen der perfekten Chips außer Acht läßt.

Im Hinblick auf den Einfluß der Prozeßvariationen erkennt man, daß die 'Testwafer' 10, 16, 21 und 22 eine deutlich geringere Ausbeute guter Chips besitzen. Diese 'Wafer' weisen als gemeinsame Eigenschaft eine vergrößerte Gatelänge auf. Zusätzlich findet man bei den 'Wafern' 07, 08, 13 und 19 einen signifikant niedrigen 'Yield' an perfekten Chips. Diese 'Wafer' gehören zu der Gruppe, die ein verkürztes 'Gate' gemeinsam hat. Es scheint also, daß bei den FED2S Chips die Variation der Gatelängen einen negativen Einfluß auf die Ausbeute hat, unabhängig davon, in welche Richtung sie verläuft. Vermutlich liegt der Grund dafür darin, daß bestimmte Schaltungsteile des Chips schlechter funktionieren, wenn der Strom in den Transistoren niedriger ausfällt, während andere Schaltungsteile bei erhöhten Strömen problematisch reagieren. Alle anderen Variationen bewirken keine beobachtbaren Veränderungen der Ausbeute.

Aufgrund des ermutigenden digitalen 'Yields' beim FED2S werden nun bei den Standard 'Wafern' auch analoge Injektionen mithilfe des internen Pulsgenerators vorgenommen, um eine Gesamtausbeute zu ermitteln und für weiterführende Messungen geeignete Chips zu identifizieren.

Man kann vorwegnehmen, daß kein einziger Chip mit perfektem Verhalten der Analogteile aller Pixel existiert, wohl aber einige als 'gut' zu klassifizierende, die von der Anzahl her ausreichen, 4 - 6 Module aufzubauen.

Tabelle 10: Analoge Meßergebnisse FED2S, Standard 'Wafer'

FED2S, 45 Chips pro 'Wafer'		
'Wafer'	Gute Chips	
Nummer	Anzahl	Ausbeute [%]
03	9	20
04	11	24
05	9	20
06	12	27
07	18	40
08	17	38
09	2	4
10	1	2

Das Kriterium für einen 'guten' FED2S, das in der voranstehenden Tabelle verwendet wird, fordert weniger als 10 Pixel pro Spalte, die eine Effizienz außerhalb von 80% - 120% der injizierten Treffer aufweisen.

Als Fazit ergibt sich, daß offenbar nicht die modifizierten Prozeßparameter für die Existenz der defekten Transistoren verantwortlich zeichnen. Über die wirklichen Ursachen herrscht gegenwärtig noch Unklarheit.

Weiterhin zeigt die zweite Submission, daß es sich bei dem Problem der defekten Transistoren nicht um eine einmalige Fehlproduktion handelt, sondern um eine prinzipielle Schwäche des DMILL Prozesses. Auch die Implementierung eines 'statischen' Designs bewirkt keine zufriedenstellende Ausbeute.

Als positiver Aspekt läßt sich zweifellos anführen, daß aufgrund der Existenz digital perfekter bzw. analog guter Chips das Design an sich funktioniert.

12 Möglichkeiten und Grenzen einer DSM Technologie bzgl. Strahlenhärte

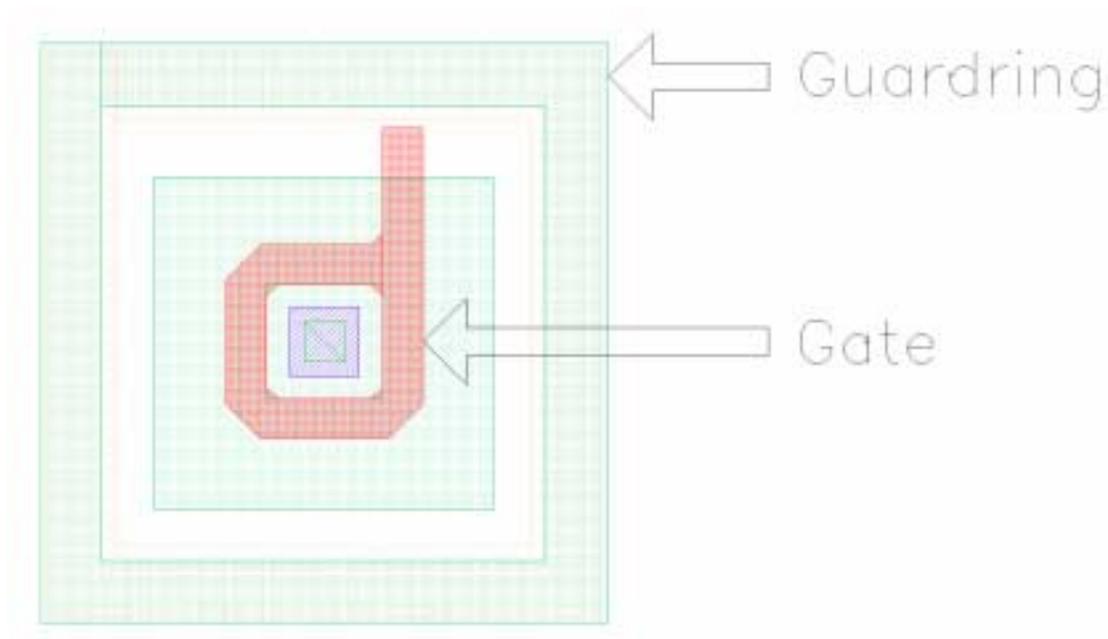
Nachdem der Einsatz von DMILL für ATLAS im Hinblick auf die Ergebnisse aus Kapitel 9 und 11 ausgeschlossen werden muß, bleibt als einzige Alternative für die Entwicklung eines strahlenharten Front End Chips die Verwendung von DSM (**D**eep **S**ub**M**icron) Technologien übrig. Diese haben neben der minimalen Transistorkanallänge von $0,24 \mu m$ [63], [64] bzw. der insgesamt winzigen Strukturgrößen aller Bauteile den entscheidenden Vorteil intrinsischer Strahlenresistenz. Aufgrund der geringen 'Gatedicke' von $6,2 nm$ können die strahleninduzierten Oxidladungen durch tunnelnde Elektronen komplett neutralisiert werden (s. Abschnitt 6.2), so daß keine Verschiebung der Transistorschwellenspannungen auftritt. Den Effekt der beiden parasitären NMOS Konfigurationen (s. Abschnitt 6.4.2) eliminiert man durch spezielle Layouttechniken, indem man die NMOS Transistoren mit kreisförmigen, geschlossenen 'Gates' versieht, so daß kein 'Drainstrom' um das 'Gate' herumfließen kann, bzw. einen zusätzlichen p^+ -'Guardring' hinzufügt, der die Bildung eines NMOS über das Feldoxid verhindert (s. Abbildung 72). Dies wird erst durch die um etwa einen Faktor 9 höhere Belegungsichte und dem daraus resultierenden enorm vergrößerten Platzangebot dieser Prozesse möglich.

Weitere Vorteile liegen darin, daß die mithilfe solcher Technologien hergestellten Schaltungen wesentlich schneller arbeiten und mit der niedrigen Versorgungsspannung von $2 V$ betrieben werden können. Für ATLAS bedeutet dies eine Verringerung des Kühlaufwandes und damit eine Ersparnis elektrischer Leistung. Diese könnte man beispielsweise zusätzlich für die Front End Chips verwenden, um mehr Freiheiten und Möglichkeiten beim Design zu eröffnen. Weiterhin bieten diese industriellen Prozesse eine schnelle Produktion großer Stückzahlen mit sehr hohen Ausbeuten, so daß damit auch immense Kosten eingespart werden.

Neben all diesen hervorragenden Eigenschaften existiert leider eventuell auch eine Schwachstelle, die es in zukünftigen Untersuchungen noch zu evaluieren gilt: Die dünnen 'Gates' bestehen nur aus einigen 10 Atomlagen. Nun stellt sich die Frage, ob hochenergetische Hadronen in der Lage sind, diese 'Gates' zu zerstören, indem sie Löcher hineinschießen. Dieses Problem wurde bisher noch nicht erforscht und muß schnellstmöglich geklärt werden.

Als konkretes Beispiel, das gleichzeitig einen Ausblick über die weiteren Schritte in der Front End Elektronik Entwicklung für ATLAS geben soll, wird in den nächsten Abschnitten das Design einer strahlenharten Speicherzelle erläutert, wie es in einer DSM Technologie umgesetzt werden kann. Die Konzentration liegt hierbei darauf, wie ein Speicher vor SEU geschützt werden kann.

Abbildung 72: Layout eines strahlenharten NMOS in DSM Technologie



12.1 Das 'Hamming Code' Register

In Abschnitt 8.2.4 wurde bereits erwähnt, daß man durch die Berechnung der Parität eines Datensatzes Bitumkehrungen detektieren kann. Jetzt geht man einen Schritt weiter und möchte eine Schaltung entwerfen, die solche Fehler automatisch korrigiert. Dabei bedient man sich des sogenannten 'Hamming Codes' [65], benannt nach seinem Erfinder Richard W. Hamming, der auch in industriellen Speicherbausteinen eingesetzt wird.

12.1.1 Das Funktionsprinzip

Das Funktionsprinzip stellt sich wie folgt dar: Zunächst berechnet man nach einer bestimmten Vorschrift aus den zu speichernden Daten einen Binärkode, der dieses Datenwort beschreibt und als korrekt identifiziert. Dieser 'Hamming Code' wird an das Datenwort angehängt und beide zusammen in einem Register gespeichert. Während nun die Daten im Speicher liegen, wird fortlaufend eine regelmäßige erneute Berechnung des Codes anhand dieser abgelegten Daten vorgenommen, der durch den Vergleich mit dem gespeicherten Code sofort eventuell auftretende Diskrepanzen entdeckt. Stimmen die beiden 'Hamming Codes' überein, befindet sich alles in korrektem Zustand, andernfalls ist ein Datenbit

falsch. Um entscheiden zu können, um welches Bit es sich handelt, muß der Code eine gewisse 'Intelligenz' aufweisen, so daß man an dessen Wert die fehlerhafte Stelle erkennen kann, die dann gezielt korrigiert wird.

Das Problem dieser Methode liegt darin, daß natürlich auch mehr als ein einzelnes Datenbit durch SEU betroffen sein kann, so daß die Korrektur nicht mehr ohne weiteres funktioniert. Zur Verbesserung verfährt man ebenso wie mit dem 'Hamming Code' zusätzlich mit der Parität der Daten, die vor der Speicherung berechnet wird, den Daten hinzugefügt wird usw.

Auf diese Weise befindet man sich in der Lage, einzelne Bitfehler korrigieren zu können, während man zweifache Bitfehler immerhin noch detektieren kann und das Register neu laden muß. Diese Information gewinnt man anhand der Kombination von Code und Parität: Wenn beide korrekt sind, gibt es keine Fehler. Falls die Parität stimmt und der Code falsch ist, handelt es sich um eine gerade Anzahl von Bitfehlern innerhalb der Daten; bei umgekehrten Vorzeichen entsprechend um eine ungerade Zahl von Bitfehlern. Nimmt man nun an, daß die SEU Rate nicht zu hoch liegt, was durch geringe Registerlängen unterstützt wird, kann man annehmen, daß keine Fehler höherer Ordnung mehr auftreten, so daß keine weiteren Maßnahmen notwendig sind.

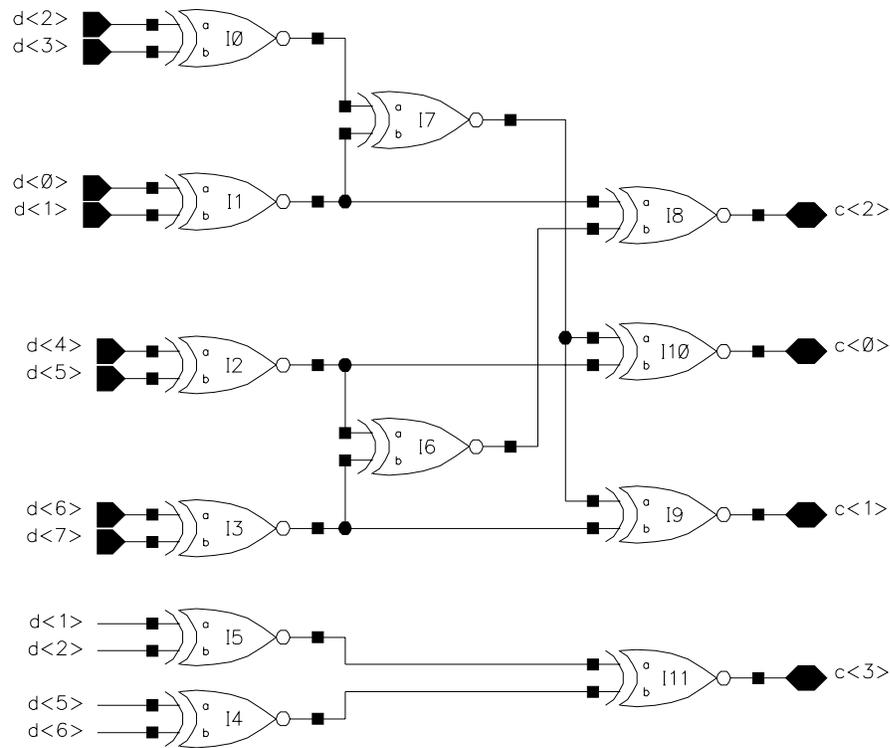
12.1.2 Die Berechnung des 'Hamming Codes'

Zunächst einmal gilt es zu ermitteln, wie viele Code Bits man benötigt, um ein Datenwort einer bestimmten Länge zu schützen. Da nicht nur die Daten mit der Bitanzahl d , sondern auch der Code mit der Bitanzahl c gespeichert wird, muß man insgesamt ein Wort der Länge $d + c$ vor SEU schützen. Außerdem möchte man in der Lage sein, genau das falsche Bit zu identifizieren, so daß man jeder Fehlerstelle genau eine Zahl des Codes zuordnen muß.

Mit einer c -stelligen Binärzahl kann man 2^c verschiedene Zustände erzeugen. Reserviert man nun den Wert '0' für das korrekte Datenwort, ergibt sich, daß mithilfe eines c Bit 'Hamming Codes' höchstens $2^c - 1 - c$ Datenbits geschützt werden können (oder weniger, wenn man nicht alle Codezahlen ausnutzt).

Als Beispiel betrachten wir nun ein 8 Bit Datenwort $d_0...d_7$, das durch einen 4 Bit 'Hamming Code' $c_0...c_3$ geschützt werden soll. Bei dessen Berechnung, die im wesentlichen aus der Kombination der Paritäten unterschiedlicher Datenbits erfolgt, ist darauf zu achten, daß man diese möglichst effizient, also mit wenig Schaltungsaufwand vornimmt. In den unteren Abbildungen sind der entsprechende Schaltplan, der ausschließlich XORs beinhaltet, mit der zugehörigen Funktionstabelle dargestellt.

Abbildung 73: Schaltplan und Funktionstabelle zur Berechnung des 'Hamming Codes'



d \emptyset	d1	d2	d3	d4	d5	d6	d7	c \emptyset	c1	c2	c3	
1	1	1	1	1	1	\emptyset	\emptyset	1	\emptyset	\emptyset	\emptyset	c \emptyset equation
1	1	1	1	\emptyset	\emptyset	1	1	\emptyset	1	\emptyset	\emptyset	c1 equation
1	1	\emptyset	\emptyset	1	1	1	1	\emptyset	\emptyset	1	\emptyset	c2 equation
\emptyset	1	1	\emptyset	\emptyset	1	1	\emptyset	\emptyset	\emptyset	\emptyset	1	c3 equation

Die Tabelle entsteht dadurch, daß man alle 12 zu schützenden Bits (8 Datenbits + 4 Codebits) in einer Zeile aufträgt und jeder Spalte einen anderen, von 0 verschiedenen, 4 Bit Code zuordnet. Dabei wird hier in den rechten Spalten die 4×4 -Einheitsmatrix verwendet und in den anderen Spalten geschickt kombiniert, um Schaltungslogik einzusparen. Es ist zu beachten, daß es sich hier nicht um eine Wertetabelle, sondern lediglich um eine Vorschrift handelt, wie der 'Hamming Code' $c_0 \dots c_3$ zu ermitteln ist.

Der 'Hamming Code' c_i berechnet sich, indem man in der i -ten Zeile die Gesamtparität aller Datenspalten (nur d -Spalten, keine c -Spalten) mit einem Eintrag '1' ermittelt. Bei einem auftretenden 'Bitflip' vergleicht man dann einfach den Code des korrekten, ur-

sprünglich gespeicherten, Datenworts mit dem Code des falschen Datenworts, indem man eine bitweise XOR Operation durchführt. Auf diese Weise entsteht ein Muster, welches durch den Vergleich mit den Spalten in der Tabelle direkt das korrupte Bit identifiziert.

Nehmen wir zum Beispiel an, daß das zu schützende Datenwort $d_0 = 1, d_1 \dots d_7 = 0$ lautet. Nach der Tabelle in Abbildung 73 errechnet sich der zugehörige Code zu $c_0 \dots c_3 = 1110$. Findet man nun eine Veränderung der Codebits gemäß $c_0 \dots c_3 = 1000$ vor, so ergibt der bitweise Vergleich (XOR Berechnung) mit dem ursprünglichen Code das Muster '0110'. Anhand der Tabelle erkennt man also, daß das Datenbit d_7 fehlerhaft ist, weil dieses als einziges nur zur Berechnung von c_1 und c_2 benutzt wurde. Der Wert von d_7 hat sich demnach von '0' nach '1' verändert.

12.1.3 Bewertung

Die bisherigen Ausführungen stellen einen recht komplizierten Sachverhalt dar, der sich auch in der Komplexität des zugehörigen Schaltplans manifestiert. Demzufolge birgt eine solche Logik leicht die Gefahr, Fehler beim Design einzubringen. Im Vergleich zu einem normalen 8 Bit Schieberegister, das aus 8 simplen 'Flipflops' besteht, setzt sich ein 'Hamming Code' Register mit automatischer Fehlerkorrektur, das 8 Datenbits schützt, aus folgenden Blöcken zusammen:

- 13 Speicher-'Flipflops' (8 Daten + 4 Code + 1 Parität)
- Logik zur Berechnung des 'Hamming Codes'
- Logik zum Vergleich beider Codes
- Decoder zur Identifikation des falschen Datenbits
- 5 'Flipflops' zur Speicherung des veränderten 'Hamming Codes' und der Parität
- Logik zum Überschreiben der 13 Speicher-'Flipflops' im Falle eines Fehlers

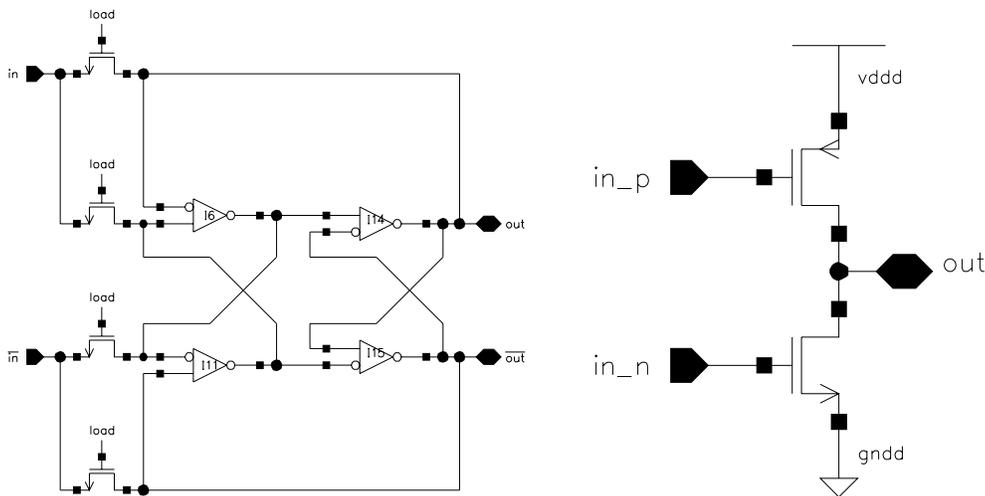
Diese Liste läßt erahnen, daß sich die Größe des zugehörigen Layouts mindestens verdoppeln wird, was dazu führt, daß ein solches Register selbst unter Verwendung eines DSM Prozesses nur für das globale Register verwendbar wäre. Daher wird im nächsten Abschnitt eine bessere Alternative vorgestellt.

12.2 Das SEU tolerante Register

Das SEU tolerante Register folgt einem völlig anderen Ansatz: Das Ziel liegt darin, Fehler in den Datenbits a priori als Eigenschaft der verwendeten elektronischen Schaltung zu verhindern. Dabei benutzt man ausschließlich logische Schaltungen, die von zwei Eingangsvariablen abhängen und verwendet jedes Ausgangssignal zweimal als Eingangssignal (doppelte Rückkopplung). Auf diese Weise kann eine Änderung des logischen Werts auf einer Leitung niemals dazu führen, daß der gesamte Schaltkreis den gespeicherten Wert verliert.

In der unteren Abbildung ist der von T. Calin [66] patentierte Schaltplan eines SEU toleranten 'Latches' dargestellt, wobei der Inhalt der Invertersymbole mit zwei Eingängen einfach aus einem Inverter mit getrennt herausgeführten 'Gates' besteht.

Abbildung 74: Schaltplan eines SEU toleranten 'Latches' (links) und des verwendeten Inverters (rechts)



Aus der intensiven Betrachtung des Schaltplans und detaillierten Simulationen gewinnt man die Erkenntnis, daß man auf drei Aspekte achten muß, wenn man alle möglichen Fälle durchspielt:

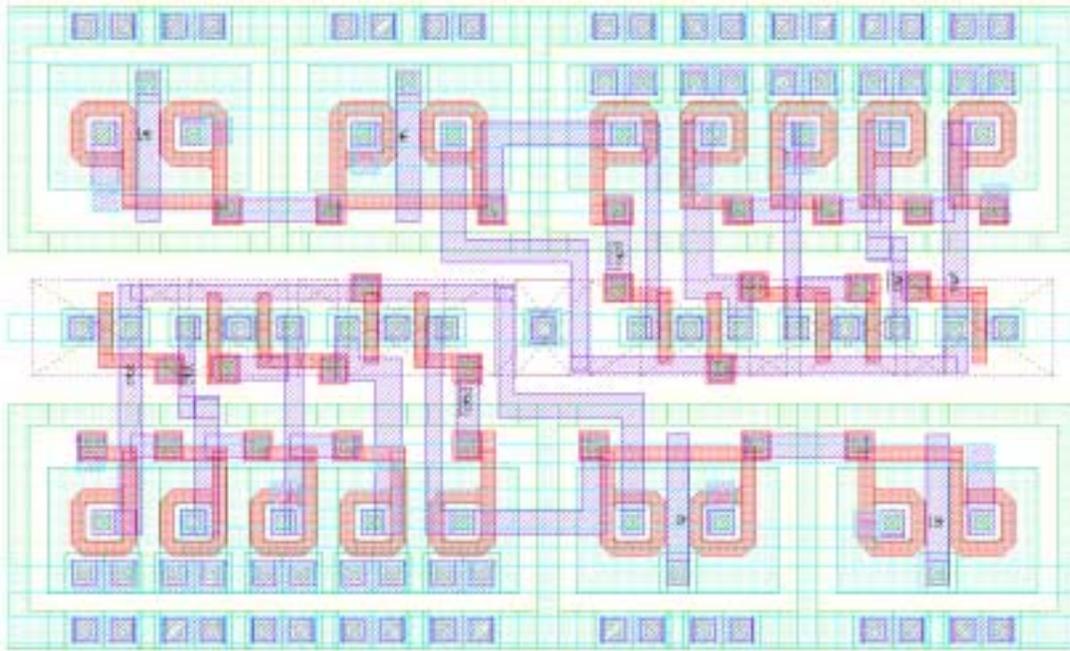
- Die Stärke des Eingangssignals muß groß genug sein, um einen beliebigen logischen Wert in das 'Latch' zu laden.
- Eine Transistorsorte muß immer stärker sein als die andere. In unserem Fall fällt die Wahl auf den NMOS, weil er aufgrund des kreisförmigen 'Gates' und des Transkonduktanzparameters dafür prädestiniert ist.

- Nur die abgebildete Konfiguration funktioniert. Bei Vertauschung der Eingänge eines oder mehrerer Inverter kann ein SEU den Verlust der gespeicherten Information verursachen.

12.2.1 Bewertung

Die Vorteile dieser Schaltung gegenüber dem 'Hamming Code' Register liegen auf der Hand: Der Schaltplan ist einfach und funktioniert sicher, und man erhält ein sehr kompaktes Layout mit der Größe von $22 \mu\text{m} \times 13 \mu\text{m}$, das vergleichbar mit einem normalen statischen 'Latch' ist. Aus diesem Grund kann man ein strahlenhartes Register sowohl für das globale Register als auch für das Pixelregister auf einem DSM Chip implementieren.

Abbildung 75: Layout zweier SEU toleranter 'Latches'



Insgesamt ergeben sich also durch die Verwendung von DSM Technologien eine Vielzahl neuer Möglichkeiten im Hinblick auf die Entwicklung eines strahlenharten Auslesechips für den ATLAS Pixeldetektor. Nach dem Ausscheiden des DMILL Prozesses bieten diese allerdings auch die einzige Alternative. Gegenwärtig werden bereits die ersten Testchips bei

verschiedenen Herstellern submittiert, um neben der allgemeinen Funktion der Elektronik möglichst bald auch Aufschluß über das Problem der strahleninduzierten Zerstörung von 'Transistorgates' zu gewinnen. Vermutlich wird im Sommer 2001 der erste Front-End Chip FEI für ATLAS hergestellt werden.

13 Zusammenfassung

Im Rahmen dieser Arbeit wurde der erste voll ATLAS taugliche, strahlenharte Auslesechip FED für den Pixeldetektor in DMILL Technologie entwickelt und anhand von detaillierten Messungen charakterisiert. Der Schwerpunkt lag auf dem Design einer mit den Anforderungen von ATLAS konformen Auslesearchitektur und der Implementierung einer groben analogen Trefferinformation (TOT). Die größte Herausforderung bildete die vorgegebene Pixelgröße von $50 \mu\text{m} \times 400 \mu\text{m}$, die nur äußerst schwierig zu realisieren war. Zusammen mit den Maßnahmen zur Erzielung der Strahlenresistenz, die schaltungs- und layouttechnisch vorgenommen wurden, wurden die Designer immer wieder an die Grenze des Machbaren geführt (s. Kapitel 8).

Die Funktionsfähigkeit der Chips zweier Submissionen im August 1999 und im Juli 2000 wurde anhand von Messungen an 'Wafern' und Einzelchips überprüft (s. Kapitel 9 und 11). Dabei trat eine Vielzahl von technologisch bedingten Defekten und Problemen auf, die in mühsamer, langwieriger Kleinarbeit durch Simulationen und Messungen an FIB modifizierten Chips geklärt wurden:

- Die parasitären Kapazitäten sind in der Realität größer als die in den Technologieparametern angegebenen Werte. Daraus folgt, daß alle Signale generell langsamer als erwartet sind.
- Die Herstellung liefert eine hohe Anzahl defekter Transistoren, die im ausgeschalteten Zustand Ströme im Mikroamperebereich führen. Diese Tatsache schließt die Verwendung von dynamischen Schaltungen aus, welche im Hinblick auf die Einhaltung der vorgegebenen Pixelgröße von $50 \mu\text{m} \times 400 \mu\text{m}$ jedoch absolut notwendig ist.
- Im Bereich der Pixelmatrix wurde eine Metallage als Abschirmung zwischen Elektronikchip und Sensor verwendet, die jedoch das kapazitive Übersprechen von Ladung in den Detektor nicht verhindern kann. Da DMILL insgesamt nur drei Verbindungslagen besitzt, existieren in dieser Technologie keine weiteren Maßnahmen, die der Designer diesbezüglich treffen kann.
- Die Messungen ergaben bezüglich des elektrischen Verhaltens der Chips starke Schwankungen von 'Wafer' zu 'Wafer'. Dies deutet auf Herstellungsparameter hin, die bei DMILL nicht unter Kontrolle sind.

Diese Fehler der DMILL Technologie führten dazu, daß die beiden Submissionen insgesamt eine Ausbeute von 0% lieferten. DMILL kann aus diesem Grund nicht für den Bau des Pixelauslesechips bei ATLAS verwendet werden.

Die einzige mögliche Alternative bietet der Einsatz von DSM Technologien, deren Vorzüge in Kapitel 12 dargelegt wurden. Bei zwei Firmen wurden bereits zwei Testchips submitted, die wiederum in Zusammenarbeit des CPPM, LBNL und PIB entwickelt wurden.

Voraussichtlich im Sommer diesen Jahres werden diese anhand von Messungen charakterisiert worden sein, so daß man beurteilen kann, ob sich die großen Hoffnungen, die in DSM Technologien gesetzt werden, bestätigen.

Tabellenverzeichnis

1	Pinbelegung des FED1	55
2	Kommandos des FED1	58
3	Die Bits im globalen Register des FED1	59
4	Die Ströme des FED1 bei einem DAC Wert von 64	62
5	Die EOE Wörter (hexadezimale Schreibweise)	73
6	Digitale Meßergebnisse FED2D, variierte 'Wafer'	100
7	Digitale Meßergebnisse FED2D, Standard 'Wafer'	100
8	Digitale Meßergebnisse FED2S, variierte 'Wafer'	101
9	Digitale Meßergebnisse FED2S, Standard 'Wafer'	101
10	Analoge Meßergebnisse FED2S, Standard 'Wafer'	102

Abbildungsverzeichnis

1	Der LHC in Genf	7
2	Der ATLAS Detektor	8
3	Der innere Detektor	9
4	Das 'Flex Hybrid'	11
5	Die Kalorimeter	12
6	Das Myonspektrometer	14
7	Die 'Front End Familie'	16
8	Die Bethe-Bloch-Formel als Funktion des Teilchenimpulses	20
9	Die Landauverteilung für verschiedene Dicken von Silizium	21
10	Die Wechselwirkungswahrscheinlichkeit von Photonen in $300 \mu m$ Silizium .	23
11	Die relevanten elektrischen Größen am pn -Übergang	25
12	Die pn -Diode als Teilchendetektor	26
13	Querschnitt durch einen NMOS	28
14	Querschnitt und Betriebsspannungen eines NMOS	29
15	Querschnitt eines NMOS in Sättigung	30
16	Die Kennlinien eines NMOS	31
17	Der NIEL für ein 50 KeV Neutron in Silizium	33
18	Wahrscheinlichkeit der nicht rekombinierten Ladungsträger in SiO_2	34
19	'Hopping' Transport der Oxidladungen	35
20	Rekombination der grenzflächennahen Oxidladungen durch Tunneleffekt . .	36
21	Layout des Pixeldetektors ('p-spray')	38
22	Erhöhung der Depletionsspannung beim ATLAS Pixelsensor	39

23	Schwellenverschiebung und g_m eines NMOS vor und nach Bestrahlung . . .	41
24	Die Schwellenverschiebung als Funktion der Strahlendosis	42
25	Der Leckstrompfad eines einzelnen NMOS um sein Gate herum	43
26	Der Leckstrompfad zwischen zwei NMOS Transistoren	43
27	Querschnitt durch die SOI-Struktur bei DMILL	46
28	Querschnitt durch einen NMOS mit 'Trenches'	47
29	Querschnitt durch einen NMOS mit 'Collsink' und LOCOS	48
30	Layout eines minimalen NMOS bzw. PMOS im Vergleich	49
31	Flußdiagramm zum Ablauf von der Idee bis hin zum Design	50
32	Blockschaltbild von FED1	54
33	Das Prinzip der 'Timestamp'-Architektur	56
34	Schreiben eines Kommandos	57
35	Blockschaltbild des globalen Registers	59
36	Schaltplan der 'Bandgap'-Referenz	60
37	Die Funktionsweise des Selbsttriggers	63
38	Blockschaltbild des Analogteils	64
39	Layout des Analogteils	65
40	Blockschaltbild des Pixelkontrollteils	65
41	Blockschaltbild des Pixelregisters mit Spaltenpaarmaskierung	66
42	Layout des Pixelkontrollteils	67
43	Schaltplan eines und Layout von 4 RAMs	68
44	Schaltplan des 'Senseamps'	70
45	Ausschnitt aus dem Layout der EOC	72
46	Datenprotokoll des Serialisierers	73

47	Das Zurücksetzen des FED1 mithilfe des SYNC	74
48	Skizze des Meßaufbaus	76
49	Der digitale Stromverbrauch der Chips auf zwei verschiedenen 'Wafern' . . .	78
50	Der Kurzschluß im Pixelschieberegister	79
51	Die Trefferkarte eines der besten FED1 Chips	81
52	Das Ausgangssignal des Serialisierers	83
53	Die gemessenen DAC Ströme eines bzw. mehrerer FED1 Chips	84
54	Der Unterschied zwischen erwarteten und gemessenen DAC Strömen	84
55	Die gemessenen DAC Spannungen eines FED1 Chips	85
56	Die gemessenen Spannungssprünge des 'Choppers'	85
57	Die gemessenen Vorverstärkersignale für verschiedene Ladungen	86
58	Die gemessenen Vorverstärkersignale für verschiedene Ströme IF	86
59	Die gemessene Anstiegszeit für verschiedene Ströme IF	87
60	Die Korrelation zwischen Vorverstärker und Diskriminator	87
61	Vergleich zwischen idealer und gemessener Schwellenkurve	88
62	Die gemessene Schwelle eines Pixels	88
63	Schwellen- und Rauschmessung eines FED1 Chips bei niedriger Schwelle . .	89
64	Schwellen- und Rauschmessung eines FED1 Chips bei höherer Schwelle . . .	90
65	Schwellen- und Rauschmessung eines FED1 Chips mit Detektor	91
66	Rauschmessung eines FED1 Chips mit und ohne XCK	92
67	Der Schaltplan der Ausleselogik im Pixel	93
68	Oszillogramme der Auslesesignale eines scheinbar toten Pixels	94
69	Schaltplan und Foto der FIB modifizierten Pixellogik	95
70	Die Kanalwiderstände des NMOS in der Ausleselogik im Pixel	95

71	Der Unterschied zwischen dynamischer und statischer Logik	96
72	Layout eines strahlenharten NMOS in DSM Technologie	105
73	Schaltplan und Funktionstabelle zur Berechnung des 'Hamming Codes' . .	107
74	Schaltplan eines SEU toleranten 'Latches'	109
75	Layout zweier SEU toleranter 'Latches'	110

Literatur

- [1] <http://cern.web.cern.ch/CERN>
- [2] <http://lhc.web.cern.ch/lhc>
- [3] *The LHC Conceptual Design Report - The Yellow Book*, CERN/AC/95-05 (1995)
- [4] <http://atlas.web.cern.ch/Atlas>
- [5] *ATLAS Technical Proposal*, CERN/LHCC/94-43 (1994)
- [6] <http://cmsinfo.cern.ch/Welcome.html>
- [7] <http://lhcb.web.cern.ch/lhcb>
- [8] *LHCB Technical Proposal*, CERN/LHCC/98-4 (1998)
- [9] <http://www1.cern.ch/ALICE/Welcome.html>
- [10] *ALICE Technical Proposal*, CERN/LHCC/95-71 (1995)
- [11] *Inner Detector Technical Design Report*, Vol I + II, CERN/LHCC/97-16 + 97-17 (1997)
- [12] *Pixel Detector Technical Design Report*, CERN/LHCC/98-13 (1998)
- [13] *Liquid Argon Technical Design Report*, CERN/LHCC/96-42 (1996)
- [14] *Muon Spectrometer Technical Design Report*, CERN/LHCC/97-22 (1997)
- [15] *Level-1-Trigger Technical Design Report*, ATLAS TDR-12 (1998)
- [16] C. Berg et al., *Bier et Pastis, a Pixel Front End Chip for ATLAS at LHC*, Nucl. Instr. and Meth. A 379 (1998)
- [17] L. Blanquard et al., *Marebo, a Full Radhard Pixel Detector Prototype for ATLAS*
- [18] <http://atlas.physik.uni-bonn.de/electronics/pirate>
- [19] K. Einsweiler et al., *FE-B Front-End Guide*, V1.0 (1999)
- [20] <http://atlas.physik.uni-bonn.de/electronics/fed>
- [21] *Physical Review D, Particles and Fields*, Volume 50 (1994)
- [22] P. Middelkamp, *Tracking with active pixel detectors*, Dissertation an der Universität Wuppertal, WUB-DIS 96-23 (1996)
- [23] W. Neeser, *Entwicklung eines Bioscope-Systems aus DEPFET-Matrizen*, Dissertation an der Universität Bonn, BONN-IR 00-06 (2000)
- [24] S. M. Sze, *Physics of semiconductor devices*, John Wiley & Sons (1981)

- [25] Boris A. Raith, *Entwicklung hybrider Pixel-Detektoren für Proton-Proton Kollisionen im ATLAS Experiment am Large Hadron Collider am CERN*, Dissertation an der Universität Bonn, BONN-IR-98-17 (1988)
- [26] C. T. Sah, *Characteristics of the Metall-Oxide-Semiconductor Transistor*, IEEE Transactions on Electron Devices 11 (1964)
- [27] R. Wunstorf et al., *Investigation of donor and acceptor removal and long term annealing in silicon with different boron/phosphorus ratios*, Nucl. Inst. and Meth. A377 (1996)
- [28] E. Fretwurst et al., *Reverse annealing of the effective impurity concentration and long term operational scenario for silicon detectors in future collider experiments*, Nucl. Inst. and Meth. A342 (1994)
- [29] H.-J. Ziock et al., *Temperature dependance of the radiation induced change of depletion voltage in silicon PIN detectors*, Nucl. Inst. and Meth. A342 (1994)
- [30] R. Wunstorf, *Systematische Untersuchungen zur Strahlenresistenz von Silizium-Detektoren für die Verwendung in Hochenergiephysik-Experimenten*, Nucl. Inst. and Meth. A377 (1996) Dissertation an der Universität Hamburg, DESY FH1K-92-01 (1992)
- [31] I. Yoshii, *Radiation effects on MOS devices and radiation-hard CMOS technologies*, Nucl. Inst. and Meth. A342 (1994)
- [32] L. Blanquart, *Effets des radiations sur les composants*
- [33] McLee et al., *Electron-hole generation, transport, and trapping in SiO₂*, in [67]
- [34] F. J. Feigl et al., *Oxygen Vacancy Model for the E' Center in SiO₂*, Solid State Commun. 14 (1974)
- [35] P. M. Lenahan, P. V. Dressendorfer, *Hole traps and trivalent silicon centers in metal/oxide/silicon devices*, J. Appl. Phys. 55 (1984)
- [36] A. Giraldo, *Evaluation of Deep Submicron Technologies with Radiation Tolerant Layout for Electronics in LHC Environments*, Dissertation an der Universität Padua (1998)
- [37] Boetsch et al., *Saturation of threshold voltage shift in MOSFET's at high total dose*, IEEE Trans. Nucl. Sci., Vol. NS-31 (1986)
- [38] R. Wunstorf, *Radiation Hardness of Silicon Detectors: Current Status*, Invited Paper presented at the IEEE Nucl. Sci. Symp. (1996)
- [39] T. Rohe et al., *Design and Test of Pixel Sensors for the ATLAS Pixel Detector*, submitted to Nucl. Instr. and Meth.
- [40] T. Stockmanns, *Messungen an der PIRATE Front-End-Elektronik für den ATLAS-Pixeldetektor*, Diplomarbeit an der Universität Bonn (2000)

- [41] ROSE Collaboration., *RD48 Status Report*, CERN/LHCC 97-39 (1997)
- [42] ROSE Collaboration., *RD48nd Status Report*, CERN/LHCC 98-39 (1998)
- [43] ROSE Collaboration., *RD48rd Status Report*, CERN/LHCC 2000-009 (1999)
- [44] R. Wunstorf, *Radiation tolerant sensors for the ATLAS pixel detector*
- [45] S. Boettcher, *Study of the Radiation Damage in Analog CMOS Pipelines, MOS Transistors and MOS Capacitors*, Dissertation an der Universität Hamburg (1996)
- [46] F. W. Sexton, J. R. Schwank, *Correlation of Radiation Effects in Transistor and Integrated Circuits*, IEEE Trans. on Nucl. Sci. (1998)
- [47] P. R. Gray, R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, John Wiley & Sons (1993)
- [48] T. Corbiere, *DMILL, Mixed Analog/Digital Rad-Hard Technology*, Vortrag DMILL Seminar in Genf (1998)
- [49] *DMILL mixed analog/digital Radiation Hard BiCMOS*, Prospekt der Firma Temic (1998)
- [50] C. Le Moullic, F. Lainat, *DMILL Overview*, Vortrag DMILL Seminar in Genf (1998)
- [51] *0,8 μm BiCMOS Process Parameters* (1997)
- [52] E. Delagnes, *DMILL Technology Overview*, Vortrag DMILL Seminar in Genf (1998)
- [53] G. Festes, *DMILL Design Kit*, Vortrag DMILL Seminar in Genf (1998)
- [54] *DMILL Electrical Design Rules* (1996-1999)
- [55] *DMILL Topological Design Rules* (1996-1999)
- [56] G. Darbo, K. Einsweiler, P. Fischer *Atlas Pixel Demonstrator* (1999)
- [57] J. Bastos et al., *A 12-Bit Intrinsic Accuracy High-Speed CMOS DAC*, IEEE Journal of Solid-State Circuits, Vol. 33 (1998)
- [58] *Mündliche Mitteilung von Peter Fischer* (1999)
- [59] D. Calvet, *Simulation Results on Readout Architectures*, Vortrag ATLAS-Woche (1999)
- [60] Silicon Solutions, *The Blue Board Testsystem, Userguide*
- [61] M. Ziolkowsky, *Pixel Control Card Specification*, Universität Siegen Ver. 2.3 (1998)
- [62] <http://www.picoprobe.com>
- [63] *CMOS 6SF & CMS 6SFS Design Manual* (2000)
- [64] *TSMC 0.25 μm Mixed Signal 2P5M/1P5M+ Salicide 2.5 V/3.3 V Design Rule* (1999)

- [65] Robert J. McEliece *The Reliability of Computer Memories*
- [66] T. Calin et al., *Upset Hardened Memory Design for Submicron CMOS Technology*, IEEE Trans. Nucl. Sci., Vol. NS-43 (1996)
- [67] T. P. Ma, P. V. Dressendorfer, *Ionizing radiation effects in MOS devices and circuits*, John Wiley & Sons (1989)

Danksagung

Zunächst danke ich Herrn Prof. Dr. Norbert Wermes für die interessante Themenstellung dieser Doktorarbeit und die gute Betreuung.

Selbstverständlich gilt mein Dank auch Dr. Peter Fischer, der mir stets unermüdlich mit Rat und Tat zur Seite stand bei allen erdenklichen Problemen, die während meiner Zeit in der Arbeitsgruppe auftraten.

Ebenso stand mir mein Kollege Giacomo Comes immer mit wertvollen Ratschlägen und nützlichen Anregungen zur Verfügung, wenn es um die Besonderheiten des Cadence Programms ging.

Weiterhin verdient die stete Hilfsbereitschaft meiner Kommilitonen aus der ATLAS Gruppe im FAUST Labor bei immer wieder aufgetretenen Problemen während der Durchführung der Chipmessungen meinen Dank.

Dank gebührt auch allen Kollegen aus der Hardwaregruppe für die angenehme Arbeitsatmosphäre während meiner Promotion.

Nicht zuletzt danke ich meiner Familie, die mir das Studium ermöglichte und mich in allen Belangen unterstützte.