

LowPower Design Methoden  
für  
VLSI CMOS Digitalschaltungen

Dissertation  
zur Erlangung des Grades  
des Doktors der Ingenieurwissenschaften  
der Naturwissenschaftlich-Technischen Fakultät II  
-Physik und Mechatronik-  
der Universität des Saarlandes

von

Dipl.-Ing. Christoph Hitzelberger

Saarbrücken

2007



LowPower Design Methoden  
für  
VLSI CMOS Digitalschaltungen

Dissertation  
zur Erlangung des Grades  
des Doktors der Ingenieurwissenschaften  
der Naturwissenschaftlich-Technischen Fakultät II  
-Physik und Mechatronik-  
der Universität des Saarlandes

von

Dipl.-Ing. Christoph Hitzelberger

Saarbrücken

2007

Tag des Kolloquiums:	23.05.2008	
Dekan:	Univ.-Prof. Dr. rer. nat.	Andreas Schütze
Mitglieder des Prüfungsausschusses:	Univ.-Prof. Dr.-Ing.	Chihao Xu
	Univ.-Prof. Dr.-Ing.	Yiannos Manoli
	Dr. rer. nat.	Ulrich Schmid

*Für meine Familie.*

*Humans think they are smarter than dolphins because we build cars and buildings and start wars etc...and all that dolphins do is swim in the water, eat fish and play around. Dolphins believe that they are smarter for exactly the same reasons. [3]*

# Inhaltsverzeichnis

<b>Danksagung</b>	<b>19</b>
<b>Kurzzusammenfassung</b>	<b>21</b>
<b>Abstract</b>	<b>23</b>
<b>Einleitung und Motivation</b>	<b>24</b>
<b>Technologie</b>	<b>28</b>
<b>1 Ursachen des Leistungsverbrauchs bei digitalen CMOS-Schaltungen</b>	<b>30</b>
1.1 Die Kurzschlussverlustleistung . . . . .	30
1.2 Die statische Verlustleistung . . . . .	34
1.2.1 Der Leakage Strom . . . . .	34
1.2.2 Der Dioden-Leakage Strom . . . . .	34
1.2.3 Der Sub-Threshold Leakage Strom . . . . .	38
1.2.4 Der Gate-Leakage Strom . . . . .	40
1.3 Die dynamische Verlustleistung . . . . .	44
1.4 Weitere Verlustquellen in integrierten Schaltkreisen . . . . .	45
1.5 Qualitativer Vergleich der Verlustleistungen . . . . .	46
<b>2 Reduzierung der statischen Leistungsaufnahme</b>	<b>52</b>
2.1 Multi-Threshold und Dynamic-Threshold CMOS-Technologien . . . . .	52
2.2 Versorgungsspannungsreduzierung bei äquivalenter Performance . . . . .	55
2.3 Vergleich von Gattern in LowLeakage- und HighSpeed-Technologie . . . . .	56
2.4 Simulationsergebnisse . . . . .	58
2.4.1 Vergleich anhand einfacher Gatter . . . . .	58
2.4.2 Frequenzbetrachtung . . . . .	59
2.4.3 Vergleich bei Addiererschaltungen . . . . .	61
2.5 Bewertung der LowLeakage bzw. HighSpeed-Technologie . . . . .	63

2.6	Body-Biasing und Dynamic Threshold CMOS-Technologie . . . . .	64
2.7	Sleep Transistor - Power-Gating . . . . .	67
2.7.1	Vorstellung des Konzepts . . . . .	67
2.7.2	Dimensionierung des Sleeptransistors . . . . .	68
2.7.3	Reduzierung des statischen Leakagestroms durch Power-Gating . . . . .	70
2.8	Reduzierung des Leakagestroms durch Bulkpotentialerhöhung . . . . .	72
2.8.1	Messungen an einer Teststruktur . . . . .	75
2.8.2	Leakagemessungen . . . . .	76
2.8.3	Performancemessungen . . . . .	77
2.9	Bewertung der Technologien zur Beeinflussung der Threshold- spannung . . . . .	79
<b>3</b>	<b>Reduzierung der dynamischen Leistungsaufnahme</b>	<b>80</b>
3.1	Systematische Methoden . . . . .	80
3.1.1	Reduzierung der Schaltaktivität . . . . .	80
3.1.2	Reduzierung der Lastkapazität . . . . .	82
3.2	Dynamische Einstellung der Versorgungsspannung . . . . .	83
3.3	Statistische Verteilung von Technologieparametern . . . . .	83
3.3.1	Flächen- und Abstandsgesetz erster Ordnung . . . . .	84
3.4	Pelgrom Modell . . . . .	85
3.5	Verteilungsfunktion der Parameter . . . . .	88
3.6	Relevante Technologieparameter . . . . .	89
3.6.1	Geometrisches Mismatch . . . . .	91
3.7	Critical-Path-Delay als Performanceindikator . . . . .	93
3.8	Ringinverterkette als Performanceindikator . . . . .	95
3.8.1	Dimensionierung der Inverterkette . . . . .	97
3.9	Evaluierung der RIC-Schwingfrequenz . . . . .	99
3.9.1	Auswertung durch Frequenzangleichung eines VCO . . . . .	101
3.9.2	Sensorinverterkette des Messsystems . . . . .	102
3.9.3	Aufbau und Wirkungsweise des VCO . . . . .	103
3.9.4	Monte-Carlo Analyse des VCO . . . . .	105
3.9.5	Der Spannungsinkrementierer . . . . .	107
3.9.6	Simulationsergebnisse der Auswerteschaltung mit VCO . . . . .	108
3.10	Alternative Auswerteschaltung mit VCO . . . . .	109
3.10.1	Funktionsprinzip der Schaltung . . . . .	109
3.11	Beurteilung der Schaltungskonzepte mit VCO . . . . .	112
3.12	Digitale Erfassung der Ringinverterschwingfrequenz . . . . .	113
3.12.1	Erzeugen eines definierten Zeitintervalls . . . . .	114

3.12.2	Auswerten des Zählerstands . . . . .	115
3.13	Einsparpotential des Systems . . . . .	121
3.14	Erweiterung zum Performance-Driven-Supply-System . . . . .	130
3.14.1	Digital Analog Konverter mit optimiertem Mismatchverhalten . . . . .	130
3.15	Gesamtsystem und Testchip . . . . .	136
3.15.1	Messergebnisse . . . . .	140
3.15.2	Diskrete Umsetzung des Systems . . . . .	143
3.16	Messergebnisse diskrete Lösung . . . . .	146
3.17	Regelung auf konstante RIC-Frequenz . . . . .	151
3.18	Beurteilung des Performance-Driven-Supply Systems . . . . .	154
<b>4</b>	<b>DC-DC Spannungswandler</b>	<b>157</b>
4.1	Der Linearregler . . . . .	159
4.1.1	Topologie des linearen DC-DC Konverters . . . . .	162
4.1.2	Simulationsergebnisse Linearregler . . . . .	164
4.1.3	Einfluss von Strompeaks auf die Regelgüte des linearen DC-DC Konverters . . . . .	169
4.2	PWM-PFM Regler . . . . .	172
4.2.1	Topologie des PWM-PFM-DC-DC-Konverters . . . . .	174
4.2.2	Simulationsergebnisse des PWM-PFM-DC-DC-Konverters . . . . .	176
4.3	Der Buck-Konverter . . . . .	178
4.3.1	Funktionsweise des induktiven Buckkonverters . . . . .	178
4.4	Buck-Konverter mit variablem Taktverhältnis . . . . .	181
4.4.1	Simulationsergebnisse . . . . .	182
4.4.2	Buck-Konverter mit variablem Taktverhältnis und aktiver Diode . . . . .	184
4.5	Buck-Konverter mit Duty-Cycle Generator . . . . .	186
4.5.1	Buck-Konverter mit alternierenden Takten . . . . .	191
4.5.2	Simulationsergebnisse . . . . .	192
4.6	Der Boost-Konverter . . . . .	193
4.6.1	Die Funktionsweise des Boost-Konverters . . . . .	193
4.6.2	PMOS-Ansteuerlogik . . . . .	195
4.6.3	Boost-Konverter mit alternierenden Takten . . . . .	197
4.6.4	Simulationsergebnisse . . . . .	199
4.6.5	Boost-Konverter mit variablem Taktverhältnis . . . . .	201
4.7	Abschließender Vergleich der DC-DC- Spannungswandlerkonzepte . . . . .	203
	<b>Systemintegration der vorgestellten Konzepte</b>	<b>206</b>
	<b>Zusammenfassung und Ausblick</b>	<b>208</b>

---

<b>Anhang</b>	<b>211</b>
<b>A</b>	<b>211</b>
A.1 Dimensionierung VCO . . . . .	211
A.2 Gesamtsystem Auswerteschaltung mit VCO . . . . .	212
A.3 Gesamtsystem alternative Auswerteschaltung mit VCO . . . . .	213
A.4 Differenzverstärker Simulationsergebnisse . . . . .	215
A.5 Komparator Simulationsergebnisse . . . . .	217
A.6 Layouts der Testschaltungen . . . . .	218
A.7 Schaltpläne und Layouts des Testchips . . . . .	221
<b>Literatur</b>	<b>233</b>

# Abbildungsverzeichnis

1	Die-Foto der Pentium4 CPU . . . . .	25
2	Taktpausen bei hoher Temperatur beim Intel Pentium4 . . . . .	25
1.1	Kurzschlussstrom eines CMOS-Inverters ohne Lastkapazität . . . . .	31
1.2	Kurzschlussstrom eines CMOS-Inverters mit Lastkapazität . . . . .	32
1.3	Leakageströme eines gesperrten NMOS-Transistors . . . . .	34
1.4	Querschnitt durch einen CMOS-Inverter mit parasitären Dioden . . . . .	35
1.5	Querschnitt von NMOS-Transistoren in Standard-CMOS- und SOI-Technologie . . . . .	35
1.6	Zwei parallele NMOS-Transistoren in SOI-Technologie [IBM] . . . . .	36
1.7	Qualitativer Performance Vergleich SOI vs. Bulk-CMOS [23] . . . . .	37
1.8	Leistungsverbrauch bei identischer Performance [23] . . . . .	37
1.9	Leistungsverbrauch bei identischer Performance zweier Gatterketten unterschiedlicher Thresholdspannung [46] . . . . .	39
1.10	Temperaturverhalten der Leakage-Ströme . . . . .	41
1.11	Temperaturverhalten der Leakage-Ströme (0.18 $\mu\text{m}$ ) . . . . .	42
1.12	Verhältnis Kurzschlussleistung zu dynamischer Leistungsaufnahme . . . . .	47
1.13	Gesamtleistung nach Gleichung 1.2 und 1.15 . . . . .	48
1.14	Anteil der statischen Leistung an der Gesamtleistung für unterschiedliche Technologien [24] . . . . .	49
1.15	Verhältnis der statischen zur dynamischen Verlustleistung in Abhängigkeit der minimalen Transistorlänge . . . . .	50
1.16	Leistungsdichte der statischen und aktiven Leistung in Abhängigkeit der minimalen Gatelänge . . . . .	51
2.1	Schaltermodell NAND2-Gatter . . . . .	53
2.2	Rise-Time eines CMOS-Inverters nach Gleichung 2.3 . . . . .	55
2.3	<i>falltime</i> eines Inverters abhängig von $V_T$ . . . . .	57
2.4	Vergleich: HS und LL NOR . . . . .	60
2.5	Inverterschaltungen mit Active-Body-Biasing . . . . .	64
2.6	Treiberlayout in Standard CMOS-Technologie . . . . .	65
2.7	Treiberlayout in CMOS-Technologie mit isolierten N-Wannen . . . . .	65

2.8	Treiberlayout in CMOS-Technologie mit isolierten N-Wannen und isolierten P-Bulks . . . . .	65
2.9	CMOS-Schaltung mit Sleep-Transistor . . . . .	67
2.10	Delay einer Gatterkette mit Sleeptransistor in Abhängigkeit der Temperatur . . . . .	68
2.11	Delay und Leakagestrom einer Gatterkette mit Sleeptransistor . . . . .	70
2.12	Delay und Leakagestrom einer Gatterkette mit Sleeptransistor $C_L = 1pF$ . . . . .	71
2.13	Thresholdspannung in Abhängigkeit vom Bulkpotential . . . . .	72
2.14	Leakagestrom einer Gatterkette in Abhängigkeit vom N-Wannenpotential . . . . .	73
2.15	Leakagestrom einer Gatterkette in Abhängigkeit vom Bulkpotential der NMOS- und PMOS-Transistoren . . . . .	74
2.16	Messaufbau zu Messungen mit variablem N-Wannenpotential . . . . .	75
2.17	Leakagestrom einer Teststruktur abhängig vom Bulkpotential . . . . .	76
2.18	Messung des Power-Delay-Produkts einer Teststruktur abhängig vom Bulkpotential . . . . .	77
2.19	Messung der Verzögerungszeit der Teststruktur abhängig vom Bulkpotential . . . . .	78
3.1	Elementare Bauelementegeometrie L, W und D . . . . .	84
3.2	Standardabweichungen der Schwellwertspannung $\sigma(V_T)$ . . . . .	85
3.3	Standardabweichungen der Steilheit $\sigma(\beta)/\beta$ . . . . .	86
3.4	Long Distance Mismatch . . . . .	87
3.5	Delay eines digitalen Treibers abhängig von der Weitenvarianz . . . . .	89
3.6	Delay eines digitalen Treibers abhängig von der Längenvarianz . . . . .	90
3.7	Varianzen der Kanaldimensionen eines MOS-Transistors . . . . .	91
3.8	Thresholdspannungsbeeinflussung durch Kurzkanaleffekte . . . . .	92
3.9	Versorgungsspannungseinstellung durch Critical-Path-Delay . . . . .	93
3.10	Ringinverter Prinzip . . . . .	95
3.11	Histogramm einer RIC bestehend aus drei minimal dimensionierten Invertiern . . . . .	97
3.12	Histogramm einer RIC bestehend aus 127 minimal dimensionierten Invertiern . . . . .	98
3.13	Histogramm einer RIC bestehend aus 13 groß dimensionierten Invertiern . . . . .	98
3.14	Verteilung der physikalischen Parameter einer Chipproduktion über einen längeren Zeitraum . . . . .	99
3.15	Überblick über die Verlustleistungen in CMOS-Schaltungen . . . . .	100
3.16	Schematischer Aufbau der Auswerteschaltung mit VCO . . . . .	101
3.17	Schaltbild des VCO . . . . .	103
3.18	Ausgangsfrequenz des VCO in Abhängigkeit der Eingangsspannung @27°C . . . . .	104

3.19	Histogramm der VCO-Frequenz bei der Eingangsspannung $U_{IN} = 1V$ . . . . .	105
3.20	Spannungsinkrementierer . . . . .	107
3.21	Spannungsverlauf der Ausgangsgrößen . . . . .	108
3.22	Aufbau der Messschaltung mit analoger Auswertung . . . . .	109
3.23	Simulation Spannungsinkrementierer, RIC und VCO . . . . .	110
3.24	Schwingfrequenz des VCO @ $V_{in} = 0.9V$ in Abhängigkeit der Temperatur und der Prozessgrenzen . . . . .	112
3.25	Schaltung zur Erzeugung eines Stoppsignals . . . . .	114
3.26	Einfluss der Temperatur und der physikalischen Parameter auf die RIC-Frequenz . . . . .	115
3.27	Konzept zum Performance Driven Supply . . . . .	116
3.28	Delay einer Gatterkette mit Best-Case Parametern . . . . .	117
3.29	Delay einer Gatterkette mit Worst-Case Parametern . . . . .	118
3.30	Delay einer Gatterkette mit Worst-Case Parametern mit variabler Versorgungsspannung . . . . .	119
3.31	Delay einer Gatterkette mit Worst-Case Parametern mit optimierter Versorgungsspannung . . . . .	119
3.32	Delay der Gatterkette aus Abbildung 3.31 mit Best-Case Parametern mit den für Worst-Case Mismatch bestimmten Versorgungsspannungen . . . . .	122
3.33	Verlustleistung einer Gatterkette mit fester und variabler Versorgungsspannung im typischen Fall (Verhältnis rechts) . . . . .	122
3.34	Verlustleistung einer Gatterkette mit fester und variabler Versorgungsspannung bei Best-Case-Mismatch (Verhältnis rechts) . . . . .	123
3.35	Verlustleistung einer Gatterkette mit fester und variabler Versorgungsspannung bei WorstCase-Parametersatz (Verhältnis rechts) . . . . .	123
3.36	Mismatch eines minimalen NMOS-Transistors . . . . .	124
3.37	falltime CMOS-Inverter . . . . .	124
3.38	falltime CMOS-Inverter $V_T, \beta = normal$ . . . . .	125
3.39	falltime CMOS-Inverter $V_T, \beta = normal, V_{DD} = 1.65V$ . . . . .	126
3.40	Versorgungsspannung für konstantes Delay bei WorstCase Transistorgeometrie . . . . .	127
3.41	Schwingfrequenz einer Ringinverterkette abhängig von der Temperatur und dem Parametersatz . . . . .	128
3.42	Gatterdelay abhängig von der Versorgungsspannung, der Temperatur und dem Parametersatz . . . . .	129
3.43	4bit Digital-Analog Konverter . . . . .	130
3.44	3bit DAC Ersatzschaltbilder gemäß Abbildung 3.43 . . . . .	131
3.45	Ausgangsspannung des DACs nach Abbildung 3.43 bei verschiedenen Spannungen in Abhängigkeit der Temperatur . . . . .	134
3.46	Vergrößerung der Darstellung aus Abbildung 3.45 . . . . .	134

3.47	Statistische Verteilung der Ausgangsspannung des DACs aus Abbildung 3.43 . . . . .	135
3.48	Systemüberblick . . . . .	136
3.49	Layout des Testchips mit Funktionsgruppen . . . . .	137
3.50	Testboard . . . . .	138
3.51	Testboard-Schaltplan . . . . .	139
3.52	Messung der Ringinverterfrequenz in Abhängigkeit von der Temperatur . . . . .	140
3.53	Simulation der Ringinverterfrequenz in Abhängigkeit von der Temperatur analog zur Messung aus Abbildung 3.53 . . . . .	141
3.54	Messung der Frequenz minimal dimensionierter Ringinverter (n=127) in Abhängigkeit von der Temperatur . . . . .	142
3.55	Schaltplan der diskreten Umsetzung der Testschaltung . . . . .	143
3.56	Platine der diskreten Umsetzung . . . . .	144
3.57	Schaltplan der DAC-Platine mit VDD-Erzeugung . . . . .	145
3.58	Kennlinie VDD-Erzeugung mit diskretem Aufbau . . . . .	148
3.59	Differenz zwischen erzeugter und optimaler Versorgungsspannung	149
3.60	relativer Leistungsumsatz bei optimaler und erzeugter Versorgungsspannung . . . . .	150
3.61	Versorgungsspannungseinstellung mit konstanter RIC-Frequenz .	151
3.62	Korrelation RIC-Frequenz-Delay . . . . .	152
3.63	$V_{DD}$ -Erzeugung mit konstanter RIC-Frequenz . . . . .	153
3.64	Versorgungsspannung-Hopping [17] . . . . .	155
4.1	Schematischer Aufbau eines linearen DC-DC-Konverters . . . . .	159
4.2	Schematischer Aufbau Linearer DC-DC-Konverter . . . . .	162
4.3	Schematic des verwendeten Differenzverstärkers . . . . .	163
4.4	Lastelemente . . . . .	166
4.5	Spannungs- und Stromkurven für eine Beschaltung entsprechend Zeile 4 aus Tabelle 4.7 . . . . .	168
4.6	Gate-Potential $U_{GG}$ des Stelltransistors $T_S$ vergrößerte Darstellung aus Abbildung 4.5 . . . . .	169
4.7	Lastelement zur Simulation einer komplexen Digitalschaltung . .	169
4.8	Spannungs- und Stromkurven für eine Beschaltung entsprechend Abbildung 4.7 . . . . .	170
4.9	Vergrößerung aus Abbildung 4.8 . . . . .	171
4.10	Package eines Intel-Pentium 4 Prozessors mit Stützkapazitäten .	171
4.11	Schematischer Aufbau des schaltenden PWM-PFM-Reglers . . .	172
4.12	Transientensimulation PWM-Konverter mit konstanter Last . .	173
4.13	Schaltbild des Komparators . . . . .	174
4.14	Spannungs- und Stromkurven für eine Beschaltung entsprechend Zeile 4 aus Tabelle 4.10 . . . . .	177
4.15	Prinzipschaltbild eines Abwärts-Wandlers . . . . .	178

4.16	Ströme und Spannungen des Buck-Konverters aus Abb. 4.15 . . .	180
4.17	Buck-Konverter mit Komparator zur PWM-Erzeugung . . . . .	181
4.18	Buck-Konverter mit aktiver Diode . . . . .	184
4.19	PWM-Generator . . . . .	186
4.20	Ausgangssignal des PWM-Generators für verschiedene Steuer- spannungen . . . . .	187
4.21	Schema des Buckkonverters mit Duty-Cycle-Generator . . . . .	188
4.22	Simulation des Buckkonverters nach Abbildung 4.21 . . . . .	189
4.23	Vergrößerung der Abbildung 4.22 . . . . .	190
4.24	Buck-Konverter mit alternierenden Takten . . . . .	191
4.25	Erzeugung eines PWM-Signals mit $D=1:7$ bzw. $D=7:1$ . . . . .	192
4.26	Prinzipschaltbild eines Aufwärts-Wandlers . . . . .	193
4.27	Steuersignal und Stromkurve Aufwärts-Wandler . . . . .	194
4.28	Steuerlogik der „aktiven Diode“ (blau) . . . . .	196
4.29	Boost-Konverter mit alternierenden Takten . . . . .	198
4.30	Spannungs- und Stromkurven für eine Beschaltung entsprechend Zeile 4 aus Tabelle 4.17 . . . . .	200
4.31	Schema des geregelten Boost-Konverters mit variablem PWM- Steuersignal . . . . .	201
4.32	Transientensimulation der Schaltung nach Abbildung 4.31 . . . .	202
4.33	Wirkungsgradvergleich der verschiedenen Wandlerkonzepte bei einer Last von $R_L = 1k\Omega$ . . . . .	203
4.34	Wirkungsgradvergleich der verschiedenen Wandlerkonzepte bei einer Last von $R_L = 5k\Omega$ . . . . .	204
4.35	Übersicht der Verlustleistungs-minimierenden Maßnahmen im Sy- stementwurf . . . . .	207
A.1	Dimensionierung der VCO-Schaltung . . . . .	211
A.2	Schematischer Aufbau der digitalen Auswertungsschaltung . . . .	212
A.3	Schematischer Aufbau der analogen Auswertungsschaltung . . . .	213
A.4	Spannungsverläufe für eine Referenzspannung von 900 mV . . . .	214
A.5	Bodeplot des Differenzverstärkers aus Abbildung 4.3 . . . . .	215
A.6	Dimensionierung des Differenzverstärkers aus Abb. 4.3 . . . . .	216
A.7	Bodeplot des Komparators aus Abbildung 4.13 . . . . .	217
A.8	Komparator mit zweifach positivem Feedback . . . . .	217
A.9	Layout des Testboards . . . . .	218
A.10	Layout der diskreten Umsetzung . . . . .	219
A.11	Layout des diskreten 4Bit-DAC mit variabler Kennlinie . . . . .	220
A.12	Layout CMOS-Standardzellen . . . . .	221
A.13	Schaltplan und Layout CMOS-XOR2 . . . . .	222
A.14	Schaltplan und Layout 2Bit Volladdierer . . . . .	223
A.15	Schaltplan und Layout D-FlipFlop . . . . .	224
A.16	Schaltplan und Layout JK-FlipFlop mit asynchronem Reset . . . .	225

---

A.17 Schaltplan und Layout 2-Input-Multiplexer . . . . .	226
A.18 Schaltplan 4-Bit Widerstands-DAC . . . . .	227
A.19 Layout 4-Bit Widerstands-DAC . . . . .	228
A.20 Schaltplan und Layout RIC, groß dimensioniert . . . . .	229
A.21 Layout RIC, minimal dimensioniert . . . . .	229
A.22 Layout Teststruktur für Bulksteuerung . . . . .	230
A.23 Layout Testchip . . . . .	231
A.24 Chipfotografie . . . . .	232

# Tabellenverzeichnis

1	Pentium 4 Kenndaten . . . . .	24
2	Leistungsaufnahme moderner Prozessoren . . . . .	26
3	Technologievergleich ST Microelectronics . . . . .	29
2.1	Leakage-Strom und Verlustleistung einfacher Gatter . . . . .	58
2.2	Verlustleistung in Abhängigkeit der Frequenz . . . . .	59
2.3	8-bit Addierer statischer Fall . . . . .	61
2.4	4-bit Addierer bei unterschiedlichen Signalfrequenzen . . . . .	61
3.1	Vergleich Binärcode - Graycode . . . . .	81
3.2	Prozessparameter der 0.18 $\mu\text{m}$ Technologie S.T. Microelectronics	87
3.3	Kenndaten der Sensorinverterkette . . . . .	102
3.4	Arbeitsbereich des VCO . . . . .	104
3.5	Prozessabhängigkeit der Ausgangsfrequenz des VCO aus Abbil- dung 3.17 . . . . .	105
3.6	Adressdekodierung des 4bit DAC aus 3.43 . . . . .	133
3.7	Zählerstand bei verschiedenen RIC-Frequenzen . . . . .	147
4.1	Simulationsergebnisse des Linearreglers mit $C_S = 1\mu\text{F}$ und $R_L =$ $1\text{k}\Omega$ . . . . .	164
4.2	Wirkungsgrad bei verschiedenen Lasten und Ausgangsspannungen	165
4.3	Variation der Ausgangsspannung bei konstanter Last ( $R_L = 1\text{k}\Omega$ )	165
4.4	Variation der Ausgangsspannung bei konstanter Last ( $R_L = 10\text{k}\Omega$ )	166
4.5	Arten von Versorgungsspannungen . . . . .	167
4.6	Schwankende Eingangsspannungen bei konstanten Lasten . . . . .	167
4.7	Wechselnde Last nach Abbildung 4.4 . . . . .	167
4.8	Variation der Ausgangsspannung bei konstanter Last ( $1\text{k}\Omega$ ) . .	176
4.9	Variation der Ausgangsspannung bei konstanter Last ( $10\text{k}\Omega$ ) . .	176
4.10	Messungen bei wechselnden Bedingungen . . . . .	177
4.11	Variation der Ausgangsspannung bei konstanter Last ( $1\text{k}\Omega$ ) . .	182
4.12	Wirkungsgrad bei verschiedenen Lasten und Ausgangsspannungen	182
4.13	Messungen bei wechselnden Bedingungen . . . . .	183
4.14	Variation der Ausgangsspannung bei konstanter Last ( $1\text{k}\Omega$ ) . .	192
4.15	Variation der Ausgangsspannung bei konstanter Last ( $1\text{k}\Omega$ ) . .	199

---

4.16	Variation der Ausgangsspannung bei konstanter Last ( $5\text{k}\Omega$ ) . . .	199
4.17	Messungen bei wechselnden Bedingungen . . . . .	200

# Danksagung

Mein erster und größter Dank gilt meiner lieben Frau Natascha, die mich in meinem Streben, diese Dissertation zu Ende zu bringen, immer unterstützt hat. Durch das aufmerksame Korrekturlesen hat sie einen praktischen Anteil am Gelingen der vorliegenden Arbeit. Für ihre Geduld und Liebe danke ich der besten Ehefrau von allen. Ein herzlicher Dank, auch wenn sie es noch nicht lesen kann, geht an meine kleine Tochter Nuri, die fortwährend mein Herz erfreut und somit meine Stimmung immer wieder aufhellt. Das gleiche gilt für meine jüngste Tochter Mia, die sich sozusagen zum Ende dieser Dissertation auch noch lebhaft zu Wort gemeldet hat. Meinen Töchtern jedenfalls ist das Konzept von LowPower nicht geläufig.

Meinem Doktorvater Professor Dr. Chihao Xu danke ich dafür, dass er mir nach seiner Berufung an die Universität des Saarlandes die Möglichkeit gegeben hat, mein Dissertationsthema weiterhin zu verfolgen. Er hat mich stets mit Rat und Tat unterstützt und viel Verständnis auch in schweren Zeiten gezeigt. Für die Erstkorrektur der vorliegenden Dissertation bedanke ich mich herzlich.

Professor Dr. Yiannos Manoli gilt natürlich mein Dank für die Anstellung an seinem Lehrstuhl, welche mir erst die Möglichkeit gab, die vorliegende Dissertation zu erstellen. Nach seinem Weggang zur Universität Freiburg konnte ich mir seiner fachlichen Unterstützung auch im Rahmen von gemeinsamen Industrieprojekten sicher sein. Für die Zweitkorrektur danke ich natürlich besonders.

Ferner bedanke ich mich bei Herrn Professor Dr. Andreas Schütze für die Betreuung in der Zeit, als der Lehrstuhl für Mikroelektronik vakant war und für die Bereitstellung des Klimaschranks.

Dem gesamten Lehrstuhl für Mikroelektronik danke ich für das angenehme Arbeitsklima. Hervorheben möchte ich Herrn Dr. Friedel Gerfers, der mir fachlich zumeist weiterhelfen konnte, genauso wie Herrn Dipl.-Ing. Hans Morgenstern, der dafür sorgte, dass die Software lief und irrtümlich gelöschte Dateien nicht im Orkus der EDV verschwanden. Frau Therese Mues, als der Seele des Lehrstuhls, sei gedankt für die reibungslose Zusammenarbeit und darüber hinaus

auch für das Aufrechterhalten des familiären Betriebsklimas (unzählige leckere Essen miteingegriffen). Dipl.-Ing. Kian Min Soh und Dipl.-Ing. Marc Albrecht danke ich ganz herzlich für die gute freundschaftliche Zusammenarbeit und den Kaffee, sowie für diverse spaßige außeruniversitäre Aktivitäten.

All meinen Studenten, Studien- und Diplomarbeitern, Praktikanten und Hi-Wis, die zum Erfolg dieser Arbeit beigetragen haben, sei hiermit ohne Nennung ihrer Namen auch herzlich gedankt, ebenso meinen Freunden, auf die ich mich stets verlassen kann.

Zu guter Letzt möchte ich meinen Eltern, Klara und Johannes Hitzelberger, der die Vollendung dieser Dissertation leider nicht mehr miterleben darf, danken, die mir den eingeschlagenen Bildungsweg erst ermöglicht haben.

# Kurzzusammenfassung

Die vorliegende Arbeit beschreibt Methoden und Schaltungen zur Reduzierung der Verlustleistung in integrierten CMOS-Schaltungen. Dabei werden die statische und dynamische Verlustleistung betrachtet. Zur Verminderung der statischen Verlustleistung werden Konzepte und Schaltungen vorgestellt, die Einfluss auf die Thresholdspannung einzelner Transistoren nehmen. Dadurch wird eine Reduzierung des Leakagestromes, und damit der statischen Verlustleistung erzielt. Das vorgestellte Konzept hat keinen verschlechternden Einfluss auf die Geschwindigkeit der Schaltung und kann u.a. auch dazu beitragen, den Stand-By Leistungsverbrauch drastisch zu vermindern. Vergleiche mit eingeführten Methoden zur Reduzierung des StandBy-Leakage-Stromes werden angestellt.

Zur Reduzierung der dynamischen Verlustleistung wird als dominanter Faktor die Versorgungsspannung verringert. Hierbei werden Faktoren wie Geschwindigkeitsanspruch an die Schaltung, Herstellungsvariationen der physikalischen Parameter des Chips und Temperatur als Parameter für eine optimale Versorgungsspannung herangezogen. Das vorgestellte System arbeitet selbstständig und kann als vorgefertigtes Makro, ähnlich den Zellen einer Standardzellenbibliothek, dem Designer anhand gegeben werden. Als Sensorwert für das System wird hierbei die Frequenz einer Ringinverterkette herangezogen. Das System bestimmt abhängig von den o.a. Faktoren eine neue Versorgungsspannung, die in den allermeisten Fällen unterhalb der Standardversorgungsspannung liegt.

Als nicht zu beeinflussende Limitierung für die Genauigkeit des Systems wird das geometrische Matching ausgemacht. Unter Berücksichtigung dieses Faktors können Einsparungen im Bereich von  $\approx 20\%$  hinsichtlich des dynamischen Leistungsverbrauchs über eine Chipmarge mit typischer Parameterspreizung erreicht werden.

Im Anschluss werden Schaltungen zur Wandlung der Versorgungsspannung betrachtet. Besonderes Augenmerk wird hierbei auf die Eignung in LowPower-Umgebungen gelegt, weshalb der Eigenleistungsverbrauch der Wandler-schaltungen in den Focus tritt. Konzepte mit geringer Komplexität sind diesbezüglich besonders geeignet. Aufgrund des Einsatzes in LowPower-Applikationen mit möglichem Batterie- oder Akkubetrieb erweist sich unter den betrachteten Low-Power DC-DC-Konvertern der Boost-Konverter als vorteilhaft, da damit die Akkuleistung sehr effizient ausgenutzt werden kann.

Abschließend erfolgt eine Zusammenfassung der Ergebnisse und Überlegungen, welche Untersuchungen zukünftig angestellt werden können, wie z.B. die Integration zusätzlicher Temperatursensoren zur Erhöhung der Genauigkeit, oder die Untersuchung weiterer Prozesse zur Qualifizierung der vorgestellten Systeme zur Verlustleistungsreduktion.

# Abstract

This thesis describes methods and circuits for the reduction of the the energy dissipation in integrated CMOS circuits. The static and dynamic energy dissipation are considered. Concepts and circuits, which influence the threshold-voltage of transistors and lead to a reduction of the leakage-current are presented . These result in a reduction of the static energy dissipation. The presented concept does not impact the circuit-performance and can contribute to drastic decrease in stand-by power consumption. Comparisons with state-of-the-art methods for reducing the leakage-current are employed.

To reduce the dynamic energy dissipation, supply voltage, being the main factor, is reduced. Speed specification, manufacture variations of the physical parameters of the chip and the chip-temperature are the parameters, that determine the optimal supply voltage. The presented system works independently and can be implemented as a macrocell, similarly to the cells of a standard-library. The frequency of a ring inverter chain is considered to be the set value for the system. The system determines a new supply voltage, which is in most cases below the standard supply voltage.

The geometrical mismatch of the integrated transistors limits the accuracy of the system. With consideration of this factor power-savings within the range of approximately 20% of the dynamic power consumption can be achieved, considering a chip margin with typical parameter spreading.

The next chapter deals with circuits for voltage-conversion. Special attention is paid to the suitability in LowPower environments. Therefore the analyses focus in the consumption of the DC-DC converter circuits itself. Concepts with low complexity are particularly suitable. For LowPower applications with (rechargeable) battery power supply the boost converter is particularly interesting, because it can use more battery capacity than buck-converters. Finally a summary of the results is given as are considerations, of what kinds of investigations could be employed in the future.

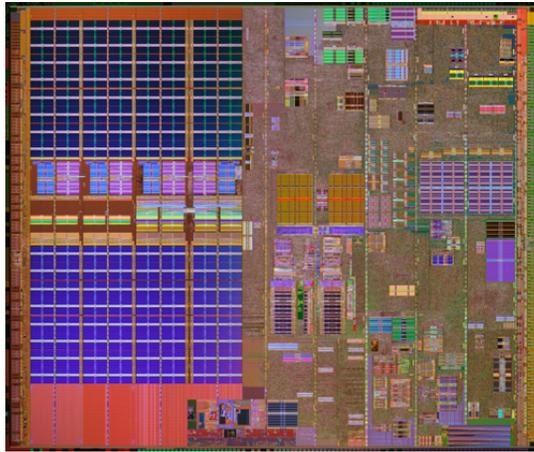
# Einleitung und Motivation

Eine Betrachtung der Applikationen der Mikroelektronik zeigt unterschiedliche Trends. Im Bereich der Prozessoren wird immer mehr Funktionalität bei immer höheren Taktraten umgesetzt, um hohe Rechengeschwindigkeiten zu erreichen. Ein Intel-Pentium4-Prozessor besitzt folgende Kenndaten:

Kenndaten Intel Pentium4-CPU 6xx-Serie	
V-Core	1.25 – 1.4 V
Kernel	Prescott-2M
Technologie	90nm
Anzahl Transistoren	169 Millionen
Die-Size	135mm <sup>2</sup>
CPU-Clock	≤ 3600MHz
Leistungsverbrauch	115W@3.6GHz

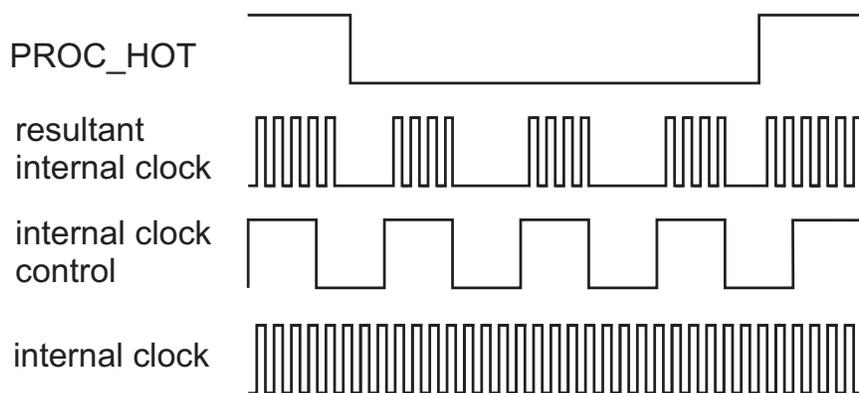
**Tabelle 1:** Pentium 4 Kenndaten

Die enorme Leistungsaufnahme moderner CPUs macht eine aufwändige Kühlung notwendig. Maßnahmen zur Reduzierung des Leistungsverbrauchs werden hierbei nicht im Interesse der Energieeinsparung getätigt, sondern erfolgen, um



**Abbildung 1:** Die-Foto der Pentium4 CPU

eine physikalische Überhitzung und, infolge dessen, einen Ausfall der CPU zu verhindern. Auch wird eine Regelung des Leistungsverbrauchs vorgenommen, um die Lärmbelastigung aufgrund des unvermeidlichen Kühlerventilators zu reduzieren, bzw. diesen ganz durch passive Kühlkörper zu ersetzen. Intel setzt hierbei für seine verschiedenen Prozessorfamilien verschiedene Konzepte ein. Bei dem o.a. Pentium4 Prozessor wird die Temperatur über eine mitintegrierte Diode überwacht. Im Fall einer zu hohen Temperatur werden zunächst Taktpausen eingefügt, um so die Anzahl der Ladungsbewegungen pro Zeiteinheit, und damit einen Hauptgrund für Leistungsverbrauch zu reduzieren (siehe Abbildung 2). Das Signal *PROC – HOT* wird bei hoher Die-Temperatur low. Der eigentliche Coretakt wird dann durch die *AND*-Verknüpfung der Signale *internalclock* und *internalclockcontrol* gebildet. Es entstehen dadurch Taktlücken, der Prozessor setzt weniger Leistung um und kann dadurch abkühlen.



**Abbildung 2:** Taktpausen bei hoher Temperatur beim Intel Pentium4

Es werden zusätzlich Strategien angewandt, die Clock-Frequenz und die Versorgungsspannung in Phasen geringer Aktivität zu reduzieren, um so die Leistungsaufnahme drastisch zu reduzieren. AMD setzt bei seinen modernen Prozessoren beispielsweise die *cool'n'quite*-Technologie ein, bei der sowohl Taktfrequenz, als auch die Betriebsspannung in Phasen niedriger Performance-Ansprüche reduziert werden. Diese Technologie ist softwaregesteuert und reduziert den Leistungsumsatz im Idle-Mode wesentlich (siehe Tabelle 2[45]). Die *cool'n'quite*-Technologie von AMD wird ähnlich auch in mobilen Prozessoren umgesetzt, wo es auf reduzierte Gesamtverlustleistung im Hinblick auf Wärmezeugung und Akku-Betriebsdauer besonders ankommt.

Prozessor	Maximal	Idle-Modus	Idle-Modus und Cool 'n' Quiet
AMD Athlon 64 3200+	89 Watt	27 Watt	10 Watt
Intel Prescott-CPU	über 100 Watt	60 Watt	nicht unterstützt

**Tabelle 2:** Leistungsaufnahme moderner Prozessoren

Die aus Tabelle 2 hervorgehende Leistungseinsparung von 50 Watt im Idle-Mode einem Prozessor mit *coll'n'quite*-Technologie gegenüber eines Prozessors ohne dieses Feature bedeutet eine finanzielle Einsparung von ca. 35€/a bei durchschnittlichem geschäftlichem PC-Gebrauch [45]. Bei Großunternehmen mit mehreren tausend PCs birgt die Reduzierung der Verlustleistung also auch signifikante finanzielle Vorteile.

Es ergeben sich bei High-Performance-Systemen also mehrere Gründe, den Leistungsumsatz zu minimieren:

1. Minimierung der thermischen Belastung
2. Minimierung der akustischen Belästigung
3. Erhöhung der Betriebszeit bei batteriebetriebenen Systemen
4. Minimierung von Betriebskosten

Grundsätzlich gelten diese Beweggründe bei anderen Applikationen auch, nur mit unterschiedlicher Gewichtung. ASICs im mobilen Einsatz, insbesondere im Mobilfunkbereich (Handys) werden auf niedrigsten Leistungsumsatz hin designed, denn lange Betriebszeiten sind hierbei ein klares Kaufargument. Der Unterschied zu modernen CPUs ist hierbei der wesentlich geringere Leistungsverbrauch, natürlich hervorgerufen auch durch geringere Rechenleistung.

Der Kunde legt bei dieser Applikation aber großen Wert auf lange Betriebsdauer ohne Wiederaufladen der Akkumulatoren, weshalb hier Low-Power Fähigkeit ein primäres Designziel ist. Unter anderem durch die Einführung der UMTS-Technologie mit ihren vielfältigen Möglichkeiten wächst aber auch im Mobilfunksektor die Rechenleistung der ASICs immer weiter. Dennoch eine lange unterbrechungsfreie Betriebszeit zu gewährleisten, ist eine große Herausforderung für den Designer.

Andere Applikationen, wie zum Beispiel medizinische Implantate (Herzschrittmacher, Sensoren) dürfen keinen hohen Leistungsumsatz haben, teils zur Vermeidung von Batteriewechseln, die Operationen notwendig machen (Herzschrittmacher), teilweise auch, um gesundheitsschädliche Erwärmung zu vermeiden (Sensoren im Blutkreislauf). Darüber hinaus gibt es medizinische Applikationen, die induktiv mit Energie durch die Haut versorgt werden ([8]). Durch die schwache Energiekopplung hat hierbei ein niedriger Leistungsumsatz absolute Priorität. Ein positiver Nebeneffekt des Powermanagements ist die allgemeine Energieeinsparung. Die enorme Anzahl verbauter Chips multipliziert individuelle Energieeinsparung zu ökologisch relevanten Größenordnungen, wenngleich dies oft kein primäres Designziel ist. Andere klassische Applikationen mit erhöhter LowPower Anforderung sind Satelliten und Raumsonden.

Es gibt also viele Beweggründe, die Energieumsetzung innerhalb eines ASICs zu reduzieren. Die vorliegende Dissertation zeigt dezidiert Methoden zur Verlustleistungsreduktion in LowPower Schaltungen und untersucht die Anwendung von Power-Management Konzepten in Applikationen, die durch ihre geringe Komplexität und Rechenleistung einen geringen Energieumsatz haben. Ziel ist es, auch im kleinen Stil Energie einzusparen. Als Anwendungen kommen alle Applikationen mit schwieriger Energieversorgung in Frage, wie medizinische Implantate, Sensorauswertungsschaltungen ohne feste Energieversorgung (z.B. Reifendrucksensoren), untergeordnete Systeme mit besonders niedrigem Energieverbrauch (z.B. in der Satellitenelektronik oder in autarken Raumfahrzeugen). Hierbei beschränkt sich die vorliegende Arbeit auf die Energieversorgung digitaler Schaltungen. Analoge Schaltungen haben grundsätzlich erhöhte Anforderungen an die Qualität der Versorgungsspannung, weil oftmals die Qualität der Signalverarbeitung von der Qualität der Versorgungsspannung in mehr oder weniger großem Maße abhängt. Die gezeigten Konzepte arbeiten teilweise mit DC-DC-Konvertern, von denen alleine der Linearkonverter geeignet ist, eine hochqualitative (niedriger Ripple) Ausgangsspannung bereitzustellen. Der Linearkonverter ist für LowPower-Applikationen wegen seines prinzipbedingten schlechten Wirkungsgrades aber schlecht geeignet. Die Ansätze für niedrigen Energieumsatz in analogen Schaltungen sind reduzierte, aber konstante Versorgungsspannung und optimiertes Design (Ruhestrome, Komplexität des Systems, Schaltungstechnik). Dies wird in der vorliegenden Arbeit nicht betrachtet. Grundsätzlich sind die vorgestellten Konzepte auch für analoge Schaltungen vorstellbar, zum einen, wenn das Schaltungsdesign eine gute Unterdrückung von Versorgungsschwan-

kungen bietet (PSSR), zum anderen, wenn statt der bevorzugten schaltenden Regler Linearkonverter zum Einsatz kommen, oder der unvermeidliche Ripple durch große Glättkapazitäten auf nicht mehr störende Größen reduziert werden kann. Das bedeutet in den meisten Fällen aber, dass externe Kondensatoren zum Einsatz kommen müssen, was in Anbetracht der Kosten für Beschaffung, Montage und zusätzliche Pads i.d.R. vermieden werden soll. Die Motivation zum Einsparen von Energie sind vielfältig; die in der vorliegenden Arbeit dargelegten Methoden hierzu sind von der jeweiligen Technologie prinzipiell unabhängig. Um jedoch konkrete Aussagen über die Qualität der Methode treffen zu können, werden diese anhand einer State-of-the-Art Technologie qualifiziert.

# Technologie

Die Simulationen und Messungen in der vorliegenden Dissertation beziehen sich fast ausschließlich auf die  $0.18\mu m$  Multi-Threshold Technologie von ST Microelectronics. Die Verwendung dieser Technologie hat mehrere Beweggründe. Zum Zeitpunkt der Erstellung der wesentlichen Teile der Arbeit war diese Technologie state of the art. Im Rahmen zweier BMBF Projekte (MEDEA und MENVOS [29] [7]) wurde mit den damaligen Projektpartnern diese Technologie gewählt, somit standen auch die Designkits für die Cadence Simulations- und Entwurfs-Umgebung zur Verfügung. Die minimale Kanallänge von  $0.18\mu m$  in Verbindung mit der Auswahlmöglichkeit zweier Threshold-Spannungen für NMOS und PMOS Transistoren macht diese Technologie für alle LowPower Applikationen interessant. Trotz der Beschränkung auf die spezielle Technologie sind die Ergebnisse dieser Dissertation nicht nur für diese Technologie gültig, sondern sie lassen sich qualitativ auf andere Technologien übertragen. Die  $0.18\mu m$  Technologie steht mittlerweile für eine State-of-the-Art Technologie, die gleichermaßen für multi purpose digitale und analoge Anwendungen geeignet ist. Tabelle 3 zeigt einen kleinen Technologievergleich mit aktuellen Technologien von ST Microelectronics.

Technology node	$0.18\mu m$ HCMOS8	$0.13\mu m$ HCMOS9	$90nm$ CMOS090
Core Voltage	1.8V	1.2V	1.0V/1.2V
Gate Oxide (Core)	3.2nm	1.7nm/2.0nm	1.6nm/2.2nm
Gate Oxide (I/O, analog)	6.5nm	5.0nm/6.5nm	5.0nm/6.5nm
Physical Gate	$0.17\mu m$	$0.11\mu m$	65nm
Interconnect	Al	Cu	Cu
ILD	k=3.5	k=3.5	k=3
Number of Metal layers	6	6 to 8	6 to 9
Metal pitch	0.64	0.41	0.28
Gate density (k/mm)	85 100 (shrink)	200	430 (hi density) 350 (hi speed)

**Tabelle 3:** Technologievergleich ST Microelectronics

# Kapitel 1

## Ursachen des Leistungsverbrauchs bei digitalen CMOS-Schaltungen

Die Verlustleistung bei digitalen CMOS-Schaltungen teilt sich hauptsächlich in drei Faktoren auf:

- Kurzschlussverlustleistung  $P_{sc}$
- statische Verlustleistung  $P_{stat}$
- dynamische Verlustleistung  $P_{dym}$

Die Gesamtverlustleistung setzt sich aus der Addition der o.a. Verlustleistungen zusammen:

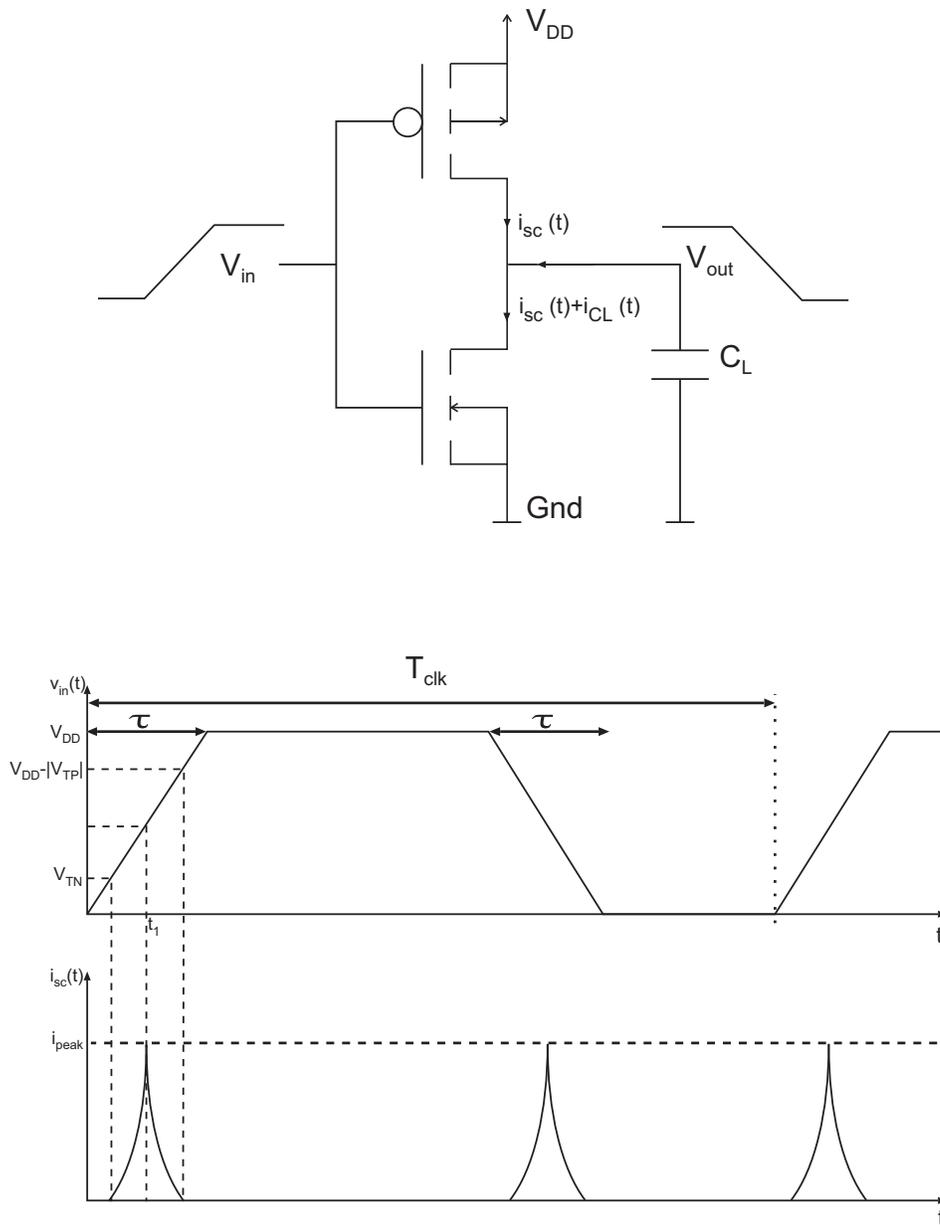
$$P_{loss} = P_{sc} + P_{stat} + P_{dym} \quad (1.1)$$

Anzumerken ist hierbei, dass die Kurzschlussverlustleistung  $P_{sc}$  von ihrem Charakter her auch eine dynamische Verlustleistung darstellt. Mit dynamischer Verlustleistung ist hier die Umladeleistung interner Kapazitäten gemeint.

### 1.1 Die Kurzschlussverlustleistung

Die Kurzschlussverlustleistung kommt durch die leitende Verbindung zwischen  $V_{DD}$  und  $V_{SS}$  zustande, die bei CMOS-Schaltungstechnik kurzzeitig entsteht, wenn ein Eingangssignal eines Gatters einen Wechsel des Ausgangsknotens verursacht. Die dadurch umgesetzte Leistung  $P_{sc}$  geht als Verlust in die Leistungsbilanz ein. Exemplarisch für ein digitales Gatter gilt für  $P_{sc}$  eines Inverters [9]:

$$P_{sc} = \frac{\beta}{12} \cdot (V_{DD} - 2V_T)^3 \cdot \frac{\tau}{T_{clk}} \quad (1.2)$$



**Abbildung 1.1:** Kurzschlussstrom eines CMOS-Inverters ohne Lastkapazität

Die Versorgungsspannung geht hierbei mit der dritten Potenz ein, wobei sich eine Erhöhung der Thresholdspannung durch den Faktor 2 in Gleichung 1.2 verhältnismäßig positiv bemerkbar macht. Der Kurzschlussstrom wird demnach durch eine Verringerung der Versorgungsspannung und/oder Erhöhung der Thresholdspannung stark dezimiert.

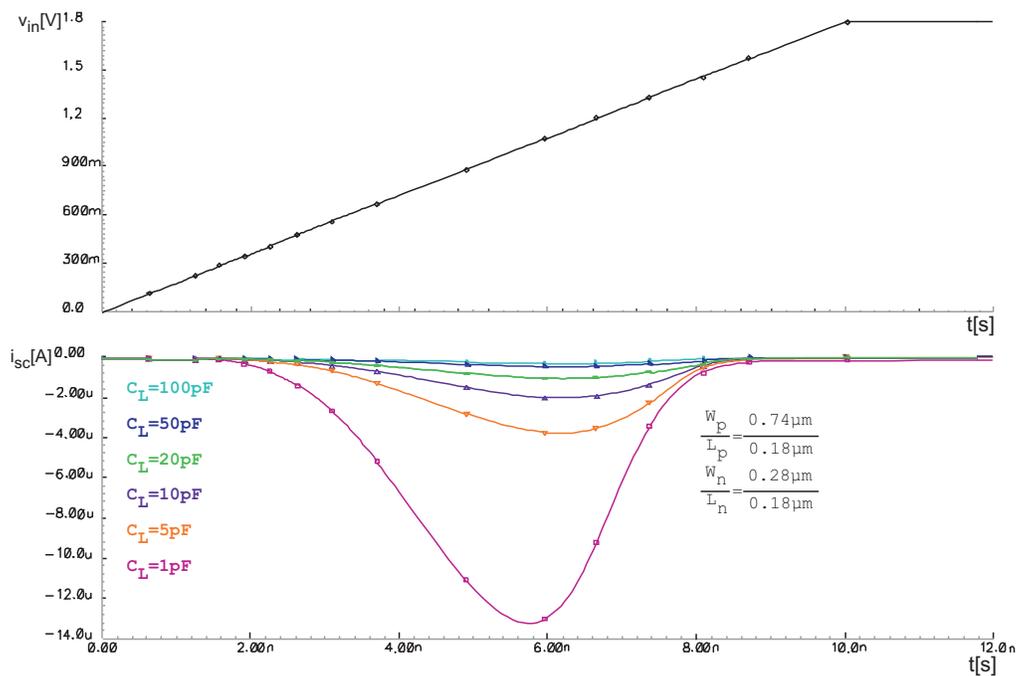
Bei dem häufig in der Realität anzutreffenden Verhältnis

$$V_T = 0.2 \cdot V_{DD} \quad (1.3)$$

wird aus Gleichung 1.2:

$$P_{sc} = \frac{\beta}{12} \cdot 0.216 \cdot (V_{DD})^3 \cdot \frac{\tau}{T_{clk}} \quad (1.4)$$

Der Kurzschlussstrom wird dann kritisch, wenn die rise/fall Zeit des Eingangsknotens groß gegenüber der rise/fall Zeit des Ausgangsknotens wird. Dies führt im Umkehrschluss zu der Überlegung, große Lastkapazitäten zu integrieren, um den Kurzschlussstrom zu minimieren. Unter der Annahme einer großen Lastkapazität und einer kurzen rise/fall Zeit am Eingang ändert sich der Ausgangsknoten beim Wechsel des Logikwertes am Eingang zunächst fast nicht.



**Abbildung 1.2:** Kurzschlussstrom eines CMOS-Inverters mit Lastkapazität

Erst nachdem der Wechsel am Eingang vollständig beendet ist, ändert sich die Spannung signifikant am Ausgangsknoten. Bezugnehmend auf Abbildung 1.1 bedeutet ein Laden des Eingangs ein, durch die grosse Lastkapazität  $C_L$  verzögertes, Entladen des Ausgangs. Durch die spannungserhaltende Wirkung einer großen Lastkapazität ist der Wechsel des Eingangs vollständig beendet, während der Ausgang fast noch die volle Spannung hält. Die Drain-Source Spannung des PMOS-Transistors ist also zum Zeitpunkt  $t_1$  (siehe Abbildung 1.1) sehr klein, weswegen hier nur ein kleiner Strom fließen kann.

Der Kurzschlussstrom wird minimiert. Das geht jedoch mit einer unakzeptablen Erhöhung der Laufzeit, des Platzbedarfs und der dynamischen Verlustleistung einher. Abbildung 1.2 zeigt die Wirkung großer Lastkapazitäten auf den Kurzschlussstrom. Die negative Stromrichtung in Abbildung 1.2 ist eine Folge der Zählrichtung der Simulation und korrespondiert mit Abbildung 1.1. Eine lange Laufzeit durch den Inverter minimiert zwar den Kurzschlussstrom, ist aber unter dem Aspekt, schnelle Schaltungen zu entwerfen, nicht wünschenswert. Ausserdem ist der Ausgang eines Gatters meistens vielfach auch der Eingang eines oder mehrerer anderer Gatter, so dass die Verzögerungszeiten durch einen Pfad in dieser Weise immer länger würden. Ein anzustrebender Kompromiss wäre es, die Gatter mit der Zielsetzung symmetrischer Verzögerungszeiten zu dimensionieren. D.h.: die Ausgangsverzögerung entspricht der Eingangsverzögerung. In diesem Fall ist der typische Anteil der Kurzschlussstromverlustleistung an der Gesamtverlustleistung kleiner als 10% [9]. Wird die Versorgungsspannung  $V_{DD}$  unter die Summe der Beträge der Thresholdspannungen der NMOS- und PMOS-Transistoren gesenkt ( $V_{DD} < V_{Tn} + V_{Tp}$ ), wird der Kurzschlussstrom vollständig eliminiert, da niemals beide Transistoren gleichzeitig leiten können. Bei dem üblichen Verhältnis von Thresholdspannung zur technologiecharakteristischen Versorgungsspannung ( $|V_{Tn,p}| \approx 0.2 \cdot V_{DD}$ ) würde das ein Absenken von  $V_{DD}$  um ca. 60% bedeuten, was die Verzögerungszeiten für die meisten Anwendungen i.d.R. unzulässig erhöht [9].

## 1.2 Die statische Verlustleistung

### 1.2.1 Der Leakage Strom

Ursache statischer Verlustleistung ist der ungewollte Stromfluss durch eigentlich gesperrte Bauelemente. In der CMOS-Technologie gibt es drei Anteile, die die statische Verlustleistung bestimmen:

- Der Dioden-Leakage Strom
- Der Sub-Threshold Leakage Strom
- Der Gate-Leakage Strom

Die Abbildungen 1.3 und 1.4 verdeutlichen diese Anteile.

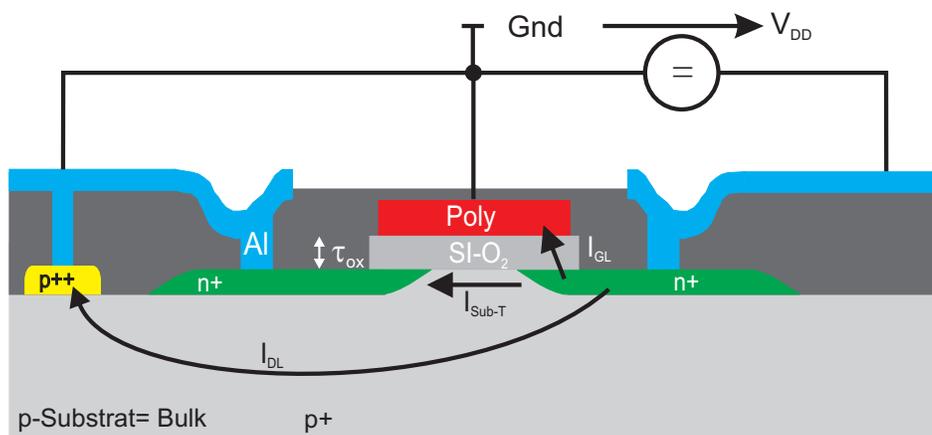


Abbildung 1.3: Leakageströme eines gesperrten NMOS-Transistors

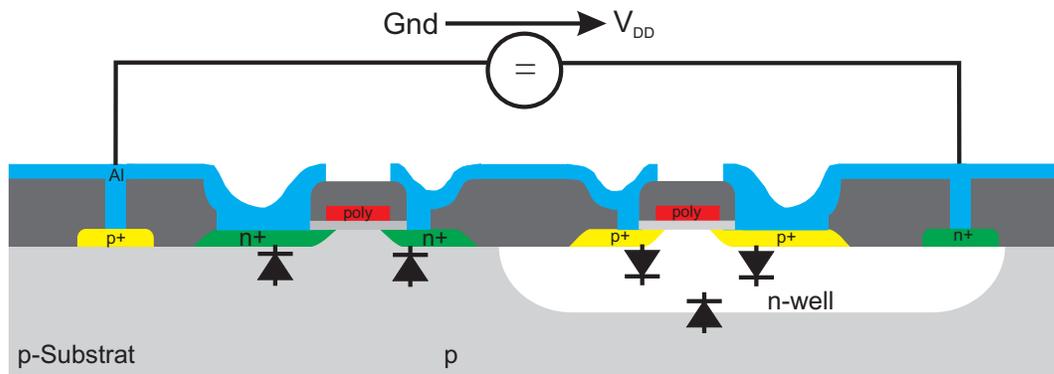
### 1.2.2 Der Dioden-Leakage Strom

In der Standard-CMOS-Technologie entstehen bei der Integration von MOS-Transistoren parasitäre Dioden.

Wie bei jeder Diode kommt es auch bei den parasitären Substrat- und Bulkdioden zu einem Stromfluss in Sperrrichtung. Der Dioden-Leakage Strom berechnet sich wie folgt [9]:

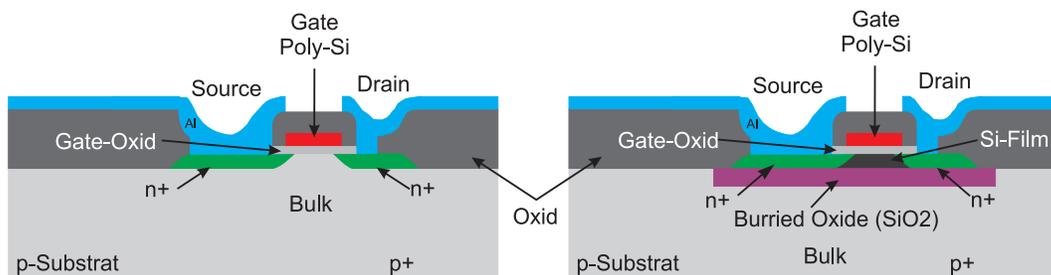
$$I_{DL} = I_s \cdot \left( e^{\frac{V_D}{V_{therm}}} - 1 \right) \quad (1.5)$$

$I_s$  ist der Rückwärtssättigungsstrom,  $V_{therm} = \frac{kT}{q}$  die thermische Spannung ( $V_{therm} = 25mV$  bei  $T = 300k$ ). Der Dioden-Leakage Strom ist relativ unabhängig von der Spannung  $V_D$ , die über der Diode abfällt und ist etwa gleich



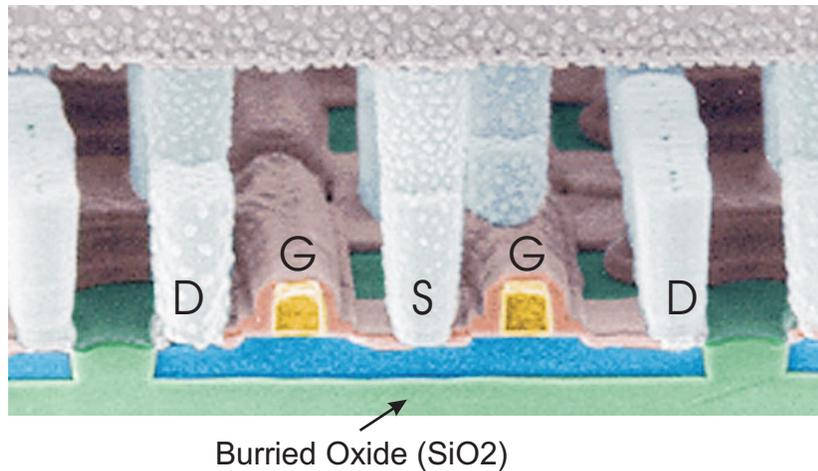
**Abbildung 1.4:** Querschnitt durch einen CMOS-Inverter mit parasitären Dioden

dem Rückwärtssättigungsstrom  $I_S$ . Der Dioden-Leakage Strom ist bei Standard CMOS-Technologien nicht vermeidbar. Eine Möglichkeit, diese Verlustquelle zu minimieren, ist die Verwendung einer sogenannten Silicon-On-Insulator Technologie (SOI). Bei dieser Technologie ist jeder Transistor von einer isolierenden Siliziumoxidschicht umgeben, die einen Stromfluss in Substrat oder Wanne verhindert. Diese Technologie ist durch den erhöhten Herstellungsaufwand jedoch wesentlich teurer als gängige Standard-CMOS-Technologie (P-Well,N-Well oder fortschrittliche Tripple-Well-Technologien).



**Abbildung 1.5:** Querschnitt von NMOS-Transistoren in Standard-CMOS- und SOI-Technologie

Die Abbildungen 1.5 und 1.6 zeigen die isolierende Oxidschicht, die einen Stromfluss nur von Drain oder Source ermöglicht.



**Abbildung 1.6:** Zwei parallele NMOS-Transistoren in SOI-Technologie [IBM]

Durch diese Oxidschicht kommt es bei der SOI-Technologie prinzipbedingt nicht zu LatchUp-Effekten durch parasitäre Thyristoren, welche bei Standard-CMOS-Technologien auftreten können. Ebenfalls reduziert sind die Drain- und Source-Junction Kapazitäten, da die Oxidschicht den Abstand zwischen Drain bzw. Source und dem Substrat erhöht. Dies macht die SOI-Technologie besonders für leakagearme HF-Schaltungen interessant. Das Bulk des NMOS-Transistors aus Abbildung 1.6 ist jetzt nicht mehr das Substrat, sondern ein Silizium-Film zwischen Drain und Source unterhalb des Gateoxids. Dieses neue Bulk kann nicht kontaktiert werden und floatet somit. Die Drain-Bulk bzw. Source-Bulk Spannung hängt also von der Historie der Schaltzustände ab, in Folge dessen variieren auch die Transistoreigenschaften ( $V_T$ ). Abbildung 1.7 zeigt einen Technologievergleich zwischen Standard-CMOS-Technologie (Bulk CMOS) und SOI-Technologie hinsichtlich der relativen Schnelligkeit über die Zeit prognostiziert bezogen auf die Performance von state-of-the-art Standard-CMOS Technologien (double-well) aus dem Jahr 1995. Normiert wurde hierbei auf den Quotienten Geschwindigkeit/Leistungsumsatz.

In Abbildung 1.8 ist der Vorteil der SOI-Technologie im Leistungsverbrauch um den Faktor zwei bis drei bei gleicher Performance gegenüber der Standard-CMOS-Technologie (in Abb. 1.7 und 1.8 mit *Bulk* gekennzeichnet) zu erkennen. Diese Eigenschaften lassen die SOI-Technologie geeignet für Low-Power-High-Performance Schaltungen erscheinen. Vergleiche zwischen Standard Bulk-CMOS Technologie und SOI-Technologie finden sich u.A. in [31] [2]. Ein Nachteil ist der erhöhte Herstellungsaufwand und die damit verbundenen höheren Kosten. Auch gibt es schaltungstechnische Schwierigkeiten durch parasitäre Bipolartransistoren, auf die an dieser Stelle nur hingewiesen sei. In der vorliegenden Arbeit werden nur schaltungstechnische Methoden vorgestellt, die Verlustleistung in den verbreiteten Standard-CMOS-Technologien (Bulk-CMOS: P-Well, double

$V_T$ , in der konkret vorliegenden Technologie auch Tripple-Well) zu reduzieren. Der Bulkdiodesperrstrom ist als solches eine Technologie-inharente Größe, die schaltungstechnisch nicht zu reduzieren ist. Einen positiven Einfluss hat nach Gleichung 1.5 nur eine Verringerung von  $V_D$  und damit verbunden von  $V_{DD}$ , der Versorgungsspannung.

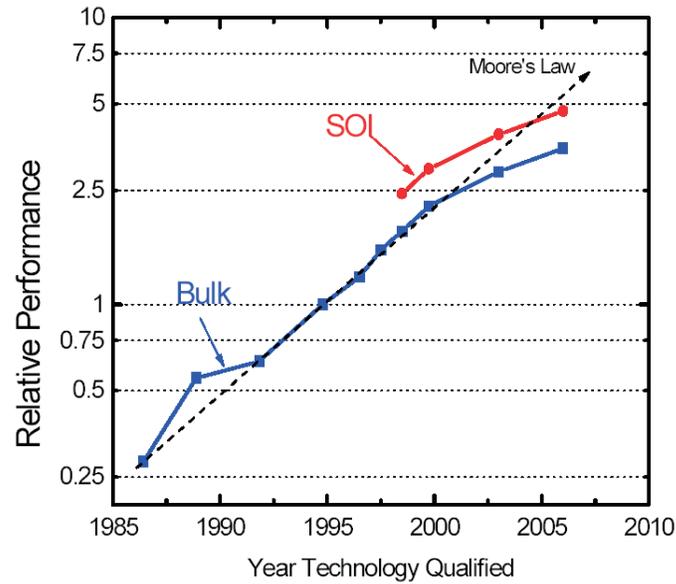


Abbildung 1.7: Qualitativer Performance Vergleich SOI vs. Bulk-CMOS [23]

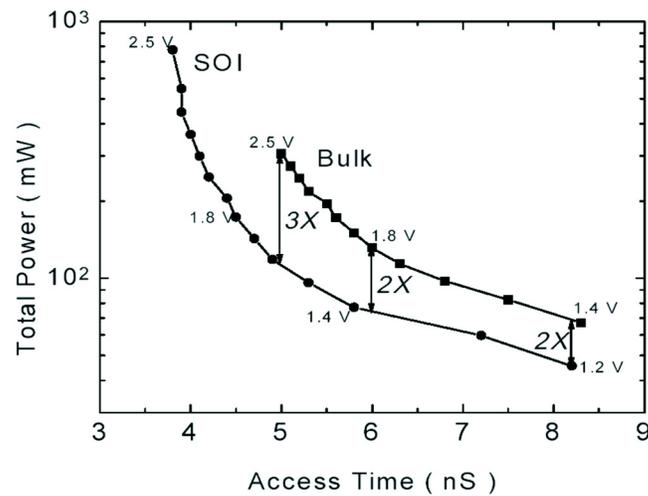


Abbildung 1.8: Leistungsverbrauch bei identischer Performance [23]

### 1.2.3 Der Sub-Threshold Leakage Strom

Bei immer kleiner werdenden Strukturgrößen wird der Sub-Threshold Strom immer bedeutender. Der Sub-Threshold Strom ist der Strom, der über die Drain-Source-Strecke eines MOS-Transistors fließt, wenn der Betrag der Gate-Source Spannung kleiner als der Betrag der Thresholdspannung ist ( $|U_{GS}| < |VT_{n,p}|$ ). Dies ist bei digitalen CMOS-Schaltungen in dem PMOS- oder dem NMOS-Zweig eines CMOS Gatters fast immer der Fall.

Es fließt ein Verluststrom von  $V_{DD}$  über einen gesperrten Transistor nach  $GND$  und erzeugt somit eine statische Verlustleistung. Für den Subthreshold Leakage Strom gilt [20]:

$$I_{DSL} = \mu_0 \frac{\varepsilon_{si}}{\tau_{ox}} \cdot V_{therm} \cdot e^{1.8} \cdot \frac{W}{L} \cdot e^{\frac{U_{GS}-V_T}{n \cdot V_{therm}}} \cdot (1 - e^{\frac{-V_{DS}}{V_{therm}}}) \quad (1.6)$$

$$n = 1 + \Omega \cdot \frac{\tau_{ox}}{D} \quad (1.7)$$

$$\Omega = \frac{\varepsilon_{si}}{\varepsilon_{ox}} \quad (1.8)$$

mit:

$$C_b = \frac{\varepsilon_{si}}{D} \quad (1.9)$$

und:

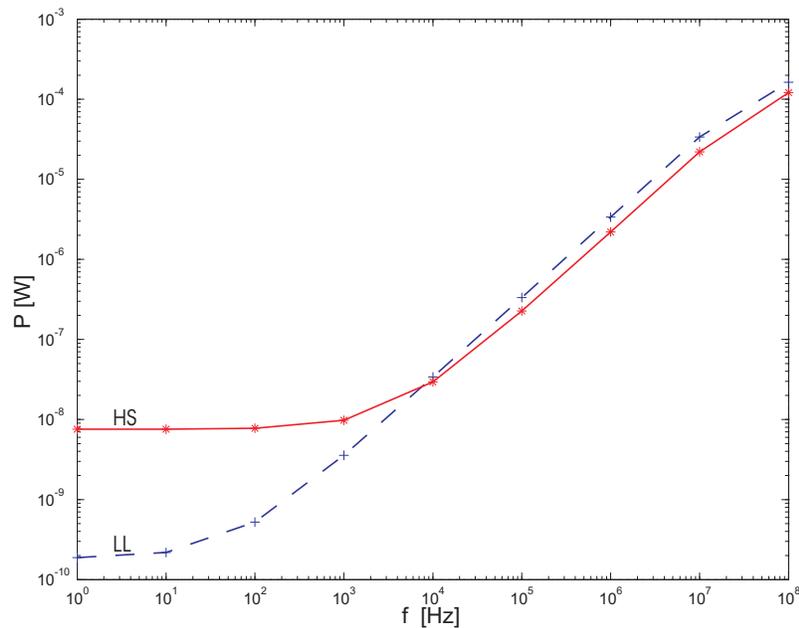
$$C_g = \frac{\varepsilon_{ox}}{\tau_{ox}} \quad (1.10)$$

gilt für:

$$n = 1 + \frac{C_b}{C_g} \quad (1.11)$$

$D$  ist die Breite der Verarmungszone unter dem Kanal (depletion channel width). Für  $V_{DS} \gg V_{therm}$  gilt nach Gleichung 1.6:  $(1 - e^{\frac{-V_{DS}}{V_{therm}}}) \approx 1$ . Dies gilt für einen großen Arbeitsbereich, weswegen näherungsweise der Sub-Threshold Leakage Strom von  $U_{GS}$ , der Thresholdspannung  $V_T$ , der Oxiddicke  $\tau_{ox}$  und dem Technologiefaktor  $n$  abhängt. Für den Schaltungsentwickler sind nur die Größen  $V_{DD}$ ,  $U_{GS}$  und  $V_T$  beeinflussbar. Mit jeder Verkleinerung der Strukturgrößen, speziell bei deep-submicron-Technologien, gilt:  $V_T \downarrow$ ,  $t_{ox} \downarrow$ ,  $U_{GS} \downarrow$ , während  $\varepsilon_{ox}$ ,  $\varepsilon_{si}$ ,  $\mu_0$ ,  $V_{therm}$  und  $\frac{W_{min}}{L_{min}}$  ungefähr konstant bleiben. Über alle Variablen betrachtet vergrößert sich die Sub-Threshold-Leakageleistung von einer zur nächst kleineren (neueren) Technologie näherungsweise um den Faktor 5 [5]. In Anbetracht der durch die kleineren Strukturen stetig verringerten dynamischen Verlustleistung kommt der statischen Verlustleistung eine hohe Bedeutung zu. Verstärkt wird diese Bedeutung noch durch die Forderung nach immer größer werdender Stand-By-Zeit bei mobilen Geräten, wie Mobiltelefonen oder mobilen PCs.

Bei diesen Applikationen sind lange inaktive Zeiträume die Regel, während deren eine integrierte Schaltung durch die Leakage-Verlustleistung die Nutzungsdauer reduziert. Soft- oder hardwaregesteuerte Powermanagement-Konzepte sind für solche Applikationen ein Weg, die Stand-By-Zeit zu erhöhen. Abbildung 1.9 zeigt zur Veranschaulichung den Leistungsumsatz zweier Ketten aus fünf NOR2-Gattern, einmal mit Transistoren hoher Thresholdspannung ( $LL$ ) und zum Vergleich mit Transistoren niedriger Thresholdspannung ( $HS$ ). Die Versorgungsspannung der HS-Kette wurde hierbei so reduziert, dass sich das gleiche Delay ergibt wie bei der LL-Kette ( $V_{DD-HS} = 1.45V$   $V_{DD-LL} = 1.8V$ ). Die Thresholdspannungen sind ca.  $V_{T-HS} = 0.3V$  und  $V_{T-LL} = 0.5V$ .



**Abbildung 1.9:** Leistungsverbrauch bei identischer Performance zweier Gatterketten unterschiedlicher Thresholdspannung [46]

Bei kleinen Schaltfrequenzen  $<10\text{kHz}$  ergibt sich hier ein Vorteil hinsichtlich des Gesamt-Leistungsverbrauchs bei der Verwendung von Transistoren hoher Thresholdspannung ( $LL$ ), obwohl die Betriebsspannung hier um  $\approx 25\%$  erhöht ist. Die Leakageverlustleistung dominiert hier den Gesamtleistungsverbrauch.

### 1.2.4 Der Gate-Leakage Strom

In modernen sub-micron-CMOS-Technologien sinkt mit der Transistorgröße auch die Gate-Oxid Dicke immer weiter. In der betrachteten Technologie (0.18µm ST Tripple-Well-CMOS) beträgt die Gate-Oxid Dicke  $\tau = 4.5nm$ . Die Gate-Oxid Dicke zukünftiger Technologien wird weiter sinken.

Bei solch kleinen Abständen werden Tunnelereignisse von Ladungsträgern immer wahrscheinlicher. Elektronen tunneln hierbei durch das Gate-Oxid, was zu einem statischen Gate-Strom und damit auch zu einer statischen Gate-Verlustleistung führt. Die Tunnelereignisse finden in den Überlappungszonen zwischen Drain- bzw Source-Gebiet und Gatefläche statt (vergl. Abb. 1.3). Darüber hinaus findet Tunnelleakage zwischen Gate und einem ausgebildeten leitendem Kanal und zwischen Gate und dem Well (Substrat oder n-Wanne in p-Well Technologien) statt. Für die Wahrscheinlichkeit tunnelnder Elektronen durch das Gateoxid gibt das MOS Model Level 1011 [36] folgende Gleichung an:

$$P_{tun}\{V_{ox}, \chi_B, B\} = \begin{cases} \exp\left(-\frac{B}{\chi_B} \cdot \frac{\left[\left(\frac{V_{ox}}{\chi_B}\right)^2 - 3 \cdot \frac{V_{ox}}{\chi_B} + 3\right]}{1 + \left(1 - \frac{V_{ox}}{\chi_B}\right)^{\frac{3}{2}}}\right) & V_{ox} < \chi_B \\ \exp\left(\frac{-B}{V_{ox}}\right) & V_{ox} \geq \chi_B \end{cases} \quad (1.12)$$

Die Tunnelwahrscheinlichkeit ist eine Funktion der Oxidspannung  $V_{ox}$  [V], der Höhe der Tunnelbarriere  $\chi_B$  [V] und einer technologie-inherenten Größe  $B$  [V]. Der Gesamt-Gate-Leakage-Strom wächst exponentiell mit dem Kehrwert der Gateoxiddicke  $\tau_{ox}$  und linear mit der Höhe des elektrischen Feldes zwischen Gate und dem Tunnelziel, also mit der entsprechenden Spannung  $V_{ox} - V_x$ . Während bei jedem Technologieschritt der Sub-Threshold Leakage Strom ca. um den Faktor 5 steigt, vergrößert sich deshalb der Gate-Leakage Strom um den Faktor 500 [5]. Dieses exponentielle Wachstum führt zu einer stetig steigenden Bedeutung der Verlustleistung durch Gate-Leakage mit jedem Technologiesprung.

Bei den in dieser Dissertation betrachteten Technologien ( $L_{min} \geq 0,18\mu m$ ) kann die Gate-Leakage-Verlustleistung nicht simuliert werden, da die Simulationsmodelle diese Möglichkeit nicht bieten. Es kann jedoch davon ausgegangen werden, dass Gate-Leakage hier noch keine besondere Relevanz hat, weswegen hier auch nur der Vollständigkeit halber darauf eingegangen wird.

Die Berechnung des Gate-Leakage Stroms wird im MOS Model Level 903 [35] nicht unterstützt. Erst im MOS Model Level 1010 bzw. 1011 [36] wird das Tunnelverhalten moduliert. Das Designkit der betrachteten 0.18µm Technologie arbeitet jedoch mit dem MOS Model Level 903, so dass Gate-Leakage Ströme bei der Simulation nicht auftreten und immer zu 0 ausgegeben werden. Die Erweiterungen im MOS Model Level 1011 setzen neue Modellkonstanten voraus, die in den gegebenen Modellen der ST-Technologie nicht vorhanden sind.

Es werden die Konstanten  $\chi_{BN}$  bzw.  $\chi_{BP}$  eingeführt (tunnelling barrier height for electrons or holes for  $Si/SiO_2$ -structure,  $\chi_{BN} = 3.1V$ ,  $\chi_{BP} = 4.5V$ ), mit denen die Wahrscheinlichkeit durch das Gateoxid tunnelnder Elektronen berechnet wird und daraus abgeleitet der konkrete Gate-Leakagestrom (für Gate-Source- bzw. Gate-Drain- und Gate-Kanal-Leakage). Das Philips MOS-Transistor Modell wird vom Spectre Simulator verwendet und wird von ST-Microelectronics zum Design-Kit für die verwendete Technologie mitgeliefert. Das Philips MOS-Modell steht in Konkurrenz zu den bekannten BSIM-Modellen der Universität Berkeley ([35],[36],[27]). Gate-Leakage wird ab den BSIM4-Modellen berücksichtigt. Als Besonderheit gilt für den Gate-Leakage Strom, dass er fast temperaturunabhängig ist, da die Tunnelwahrscheinlichkeit einerseits und die elektrische Feldstärke über dem Feldoxid andererseits wenig temperaturabhängig sind. Abbildung 1.10 [44] zeigt eine Abschätzung der Temperaturabhängigkeiten für einen 50nm langen NMOS-Transistor eines experimentellen Deep-Sub-Micron Prozesses.

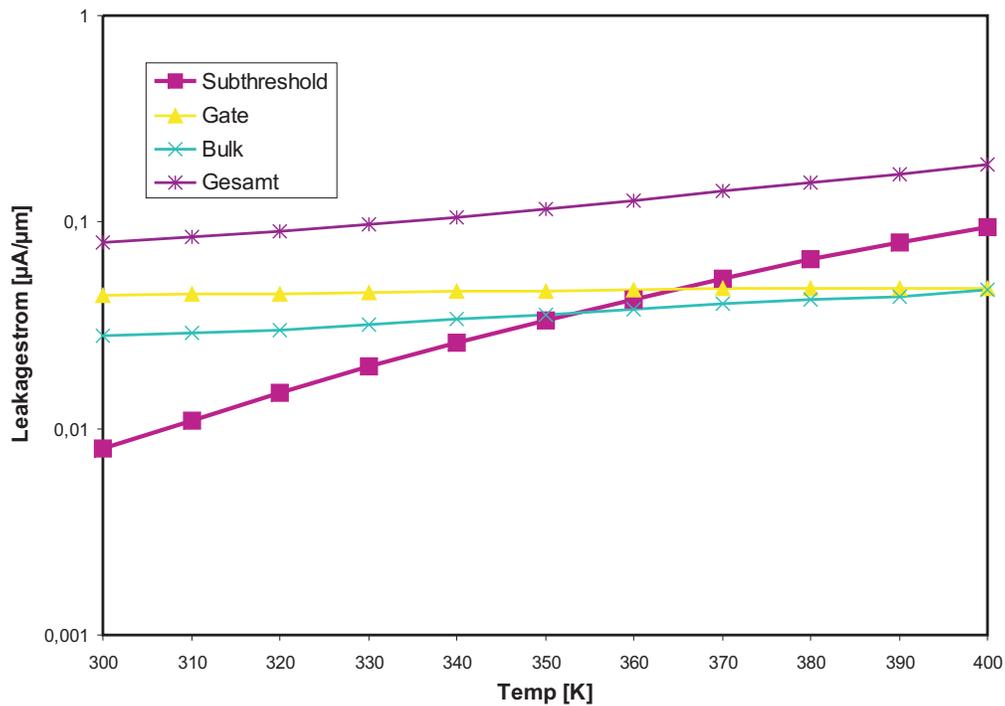


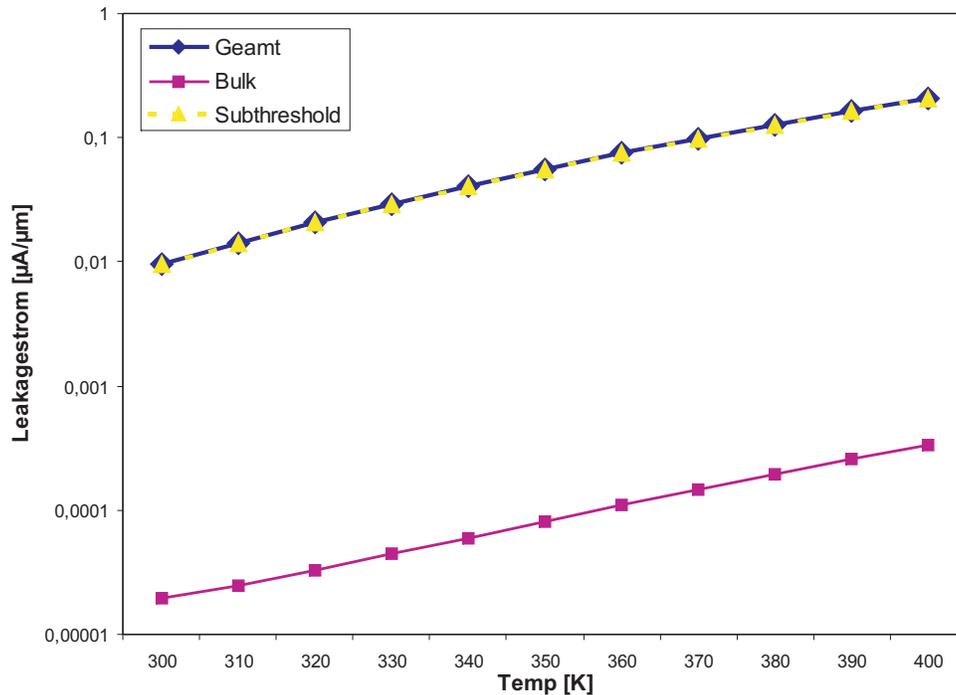
Abbildung 1.10: Temperaturverhalten der Leakage-Ströme

Abbildung 1.11 zeigt zum Vergleich die simulierten Leakageströme eines minimal dimensionierten, gesperrten NMOS-HS Transistors in  $0.18\mu m$ -Technologie. Der Gate-Leakageanteil ist nicht mit angegeben. Im Vergleich zu Abbildung

1.10 ist das Verhältnis von Subthreshold- zu Bulkleakage auffallend, welches zeigt, dass der Bulkleakagestrom in der betrachteten  $0.18\mu\text{m}$  Technologie eine untergeordnete Rolle spielt. Ein direkter Vergleich zwischen Abbildung 1.10 und Abbildung 1.11 ist aufgrund differierender Simulationsmodelle kritisch. Qualitativ lässt sich jedoch eine Zunahme des Gateleakagestroms bei deep-sub-micron Technologien mit exponentiellem Charakter konstatieren. [12] liefert für einen 100nm Prozess für den Gate-Source- und Gate-Drain-Leakagestrom Angaben, [14] erstellt für die Gate-Leakage-Stromdichte folgende Abhängigkeit:

$$J_g = A \cdot \left( \frac{T_{oxref}}{t_{ox}} \right)^{ntox} \cdot \frac{V_g \cdot V_{aux}}{t_{ox}^2} \cdot e^{-B(\alpha - \beta|V_{ox}|)(1 + \gamma|V_{ox}|) \cdot t_{ox}} \quad (1.13)$$

wobei  $V_{aux}$  eine zusätzliche Funktion ist, die je nach Betriebszustand des Transistors und Gateleakagerichtung (Gate-Source, Gate-Drain, Gate-Bulk) variiert. Entscheidend ist der Einfluss von  $t_{ox}$ , der sowohl linear, quadratisch als auch exponentiell eingeht.



**Abbildung 1.11:** Temperaturverhalten der Leakage-Ströme in ST-0.18  $\mu\text{m}$ -Technologie

Der Einfluss des Gateleakage bei der betrachteten  $0.18\mu m$  Technologie lässt sich nur abschätzen. Eine Abschätzhilfe liefern [21] und [22], wo für Technologien mit vergleichbarer Gate-Oxid-Dicke ( $\tau_{ox} \approx 4nm$ ) für die Gatefeldspannung von  $1.8V$  ein Gateleakagestrom von  $I_{G-leak} \approx 10^{-5} \frac{A}{cm^2}$  angegeben wird. Umgerechnet auf die minimale Kanallänge von  $L_{min} = 0.18\mu m$  ergibt sich  $J_G \approx 10^{-8} \frac{\mu A}{\mu m}$ . Dieser Wert ist um Größenordnungen kleiner als die übrigen Leakageströme (vergl. Abbildung 1.11).

## 1.3 Die dynamische Verlustleistung

Für die dynamische Verlustleistung gilt:

$$P_{dym} \sim \alpha \cdot f \cdot C_L \cdot V_{DD}^2 \quad (1.14)$$

In Gleichung 1.14 ist  $f$  der Systemtakt und  $C_L$  die gesamte Lastkapazität. Der Aktivitäts-Faktor  $\alpha$ , der zwischen Null und Eins variieren kann, stellt den durchschnittlichen Anteil der an Umladevorgängen beteiligten Knoten dar. Er belegt die Gesamlastkapazität mit einem statistischen Faktor. Von einem einzelnen schaltenden Gatter ausgehend (also unter Vernachlässigung des Aktivitätsfaktors  $\alpha$ ) gilt näherungsweise für die Umlade-Verlustleistung:

$$P_{dym} = \frac{1}{2} \cdot f \cdot C_L \cdot V_{DD}^2 \quad (1.15)$$

Wie schon aus Gleichung 1.14 ersichtlich, kommt der Versorgungsspannung  $V_{DD}$  wegen des quadratischen Anteils eine dominante Rolle zu. Als Lastkapazität kommt die Summe aller wirksamen Kapazitäten in Frage. Die Summe aller Eingangskapazitäten,  $C_{GS-ges}$ , ist diesbezüglich ein großer Faktor. Darüber hinaus haben die Leitungskapazitäten vor allem komplexer Digitalschaltungen mit ihren langen internen Verbindungen (Signalbus) einen großen Anteil an der wirksamen Lastkapazität. Zusätzlich verzögern diese Leitungskapazitäten die Verarbeitungsgeschwindigkeit, weswegen durch optimiertes Platzieren der Schaltungselemente beim Layout und durch spezielle Routing-Strategien versucht wird, diese Verbindungen so kurz wie möglich zu halten.

Die spezifischen, temperaturabhängigen Widerstände aller an Umladevorgängen beteiligten Strompfade limitieren die Geschwindigkeit der Auf- und Entladevorgänge (es gilt:  $i_{max} = V_{DD}/R_s$ , mit  $R_s =$  effektiver serieller Widerstand). Hierbei sind besonders die Übergangswiderstände bei Kontaktierungen verschiedener Materialien (vergl. Kap. 1.4) kritisch. Die Gesamtverlustleistung ist unabhängig von den zu den Knotenkapazitäten seriellen Widerständen, da der Energiegehalt ( $1/2CU^2$ ) nur von der Kapazität abhängt. Die Lade- bzw. Entladezeit vergrößert sich allerdings mit höheren Widerständen, so dass die Performance der Schaltung beeinflusst wird. Aus diesem Grund erscheint der ohmsche Widerstand auch nicht in Formel 1.15.

## 1.4 Weitere Verlustquellen in integrierten Schaltkreisen

Verluste an parasitären Elementen gehören zu den weniger betrachteten, dennoch vorhandenen Verlustquellen. Hierzu zählen:

- Verluste durch parasitäre Induktivitäten (Leitungs- und Bondinginduktivität)
- Verluste durch parasitäre Kapazitäten
  - Intersignal-Kapazitäten bei Signalleitungen (Busse)
  - Kapazitäten zwischen Signalleitungen und Substrat

Die o.a. Verluste sind nicht Thema der hier beschriebenen Methoden zur Reduzierung der Verlustleistung. Sie sind Gegenstand umfangreicher Untersuchungen in der Literatur und werden durch technologische Maßnahmen, sowie durch schaltungstechnische Maßnahmen (z.B. Reduzierung des Signalswings bei langen Signalleitungen) zu reduzieren versucht. Insbesondere bei großen Digitalschaltungen mit hohen Taktraten (moderne Mikroprozessoren) kommt es zu Schwierigkeiten bei der Versorgungsspannungskonstanz. An den parasitären Induktivitäten treten durch sprunghaften Strombedarf, hervorgerufen durch die massive Parallelverarbeitung von Information, Spannungseinbrüche auf, die ohne den Einsatz zusätzlicher Stützkapazitäten zu Fehlfunktionen führen würden. Durch den immensen Strombedarf solcher Schaltungen haben diese Kapazitäten nicht mehr sinnvoll integrierbare Größenordnungen und müssen in Form externer Kondensatoren appliziert werden, die durch spezielle Montageverfahren (z.B. Flip-Mounting) räumlich nahe zum Verbraucher platziert und elektrisch mit diesem verbunden werden.

## 1.5 Qualitativer Vergleich der Verlustleistungen

Ein Vergleich der vorgestellten Verlustleistungen erweist sich ohne konkrete Beispielapplikation als schwierig. Aufgrund der umfangreichen Abhängigkeiten der einzelnen Verlustquellen von vielen verschiedenen Parametern soll hier nur ein Überblick mit Beispielen aus der Literatur gegeben werden. Genaue Untersuchungen in der Zieltechnologie werden in den entsprechenden folgenden Kapiteln angestellt.

Werden die Gleichungen 1.2 und 1.15 zueinander in Bezug gesetzt, gilt für das Verhältnis von Kurzschlussleistung zu dynamischer (Umlade-) Leistung:

$$\frac{P_{sc}}{P_{dym}} = \frac{\beta}{12} \cdot \frac{(V_{DD} - 2V_T)^3 \cdot \frac{\tau}{T_{clk}}}{\frac{1}{2} \cdot C_L \cdot V_{DD}^2} \quad (1.16)$$

Mit

$$\tau \stackrel{!}{=} \frac{1}{20} \cdot \frac{1}{f} \quad (1.17)$$

und

$$T_{clk} = \frac{1}{f} \quad (1.18)$$

folgt:

$$\frac{P_{sc}}{P_{dym}} = \frac{\beta}{120} \cdot \frac{(V_{DD} - 2V_T)^3}{f \cdot C_L \cdot V_{DD}^2} \quad (1.19)$$

Gleichung 1.17 legt die *rise-* bzw. die *fall-time* eines Gatters auf 1/20 der Periodendauer des Systemtakts fest. Diese Annahme ist willkürlich. Als harte Grenze für  $\tau$  muss  $\tau < \frac{1}{2f}$  gelten. Mit 1.17 folgt für 1.16: Abbildung 1.12 zeigt die grafische Darstellung von 1.19 mit für die Zieltechnologie typischen Werten ( $V_{DD} = 1.8V$ ,  $\beta = 250 \frac{\mu A}{V^2}$ ,  $V_T = 0.36V$ ). Die Graphen für zwei verschiedene Frequenzen  $f$  (Abbildung 1.12 unten) verdeutlichen die dominierende Bedeutung der Kurzschlussleistung für niedrige Frequenzen und kleine Lastkapazitäten.

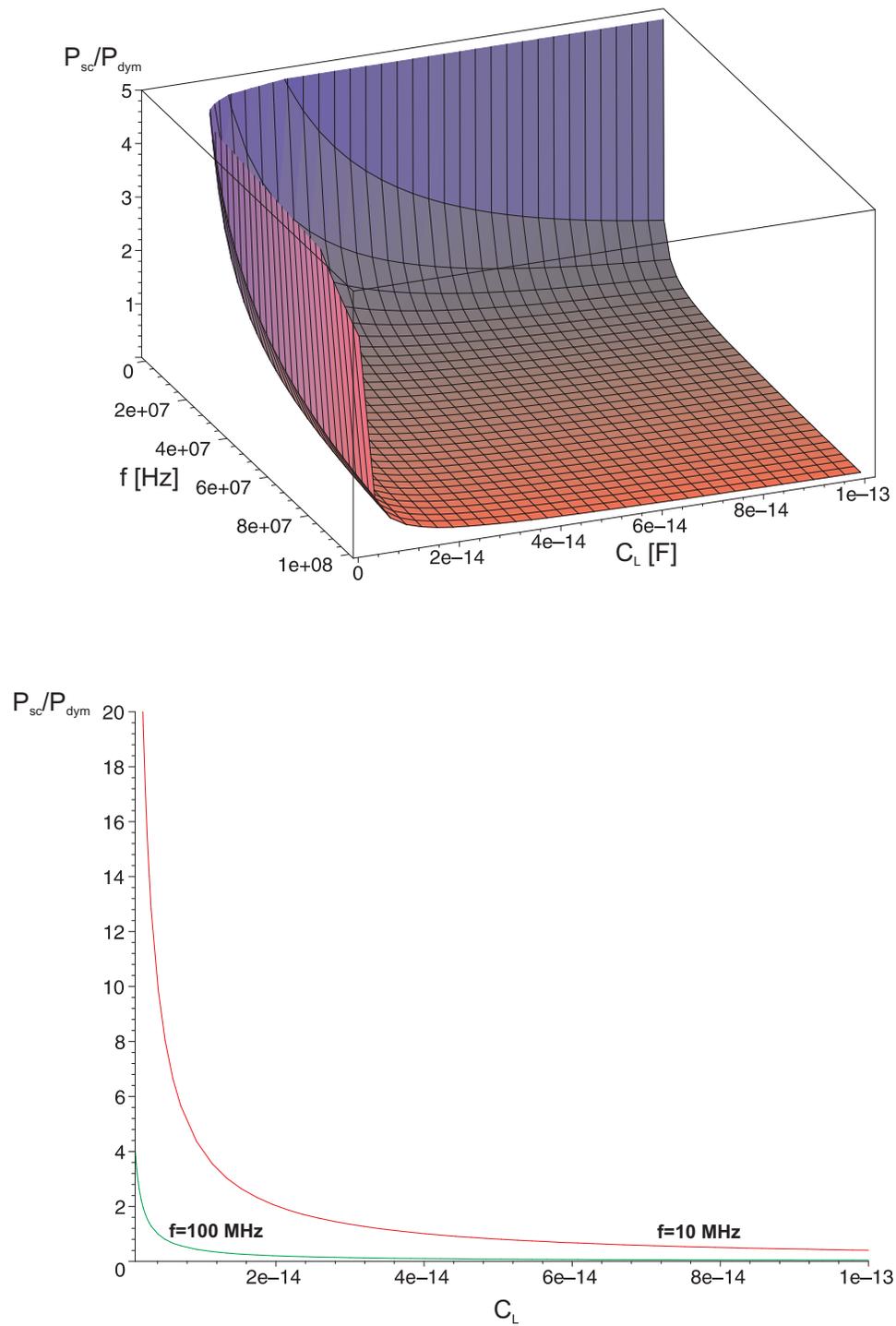


Abbildung 1.12: Verhältnis Kurzschlussleistung zu dynamischer Leistungsaufnahme

In der Praxis sind jedoch nicht alle Verhältnisse von  $C_L$  und  $f$  erreichbar, da sie sich gegenseitig bedingen. Mit schnellerer *rise-* bzw. *fall-time* wird das Verhältnis von Kurzschlussleistung zu dynamischer Leistungsaufnahme kleiner. Im Hinblick auf die Gesamtleistung eines Gatters interessant erscheint die Addition der Kurzschluss- und der dynamischen Leistung (Abb. 1.13):

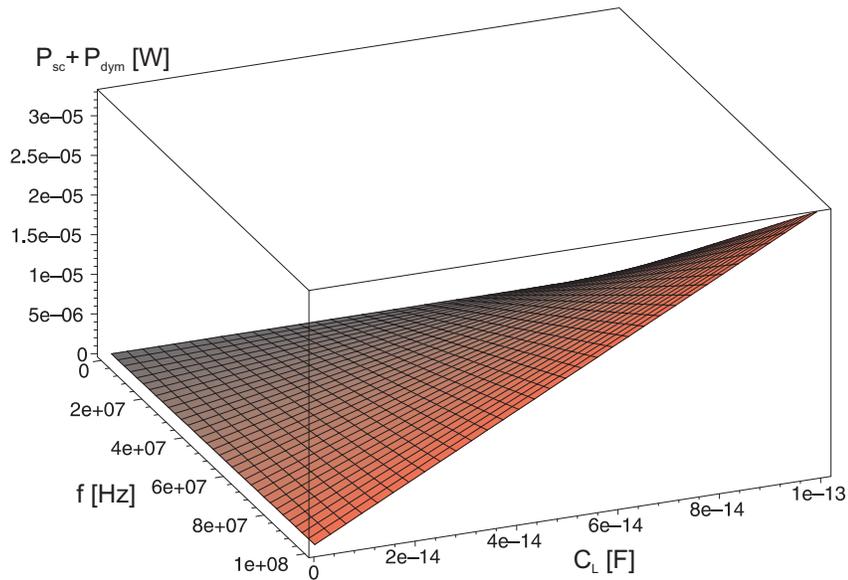
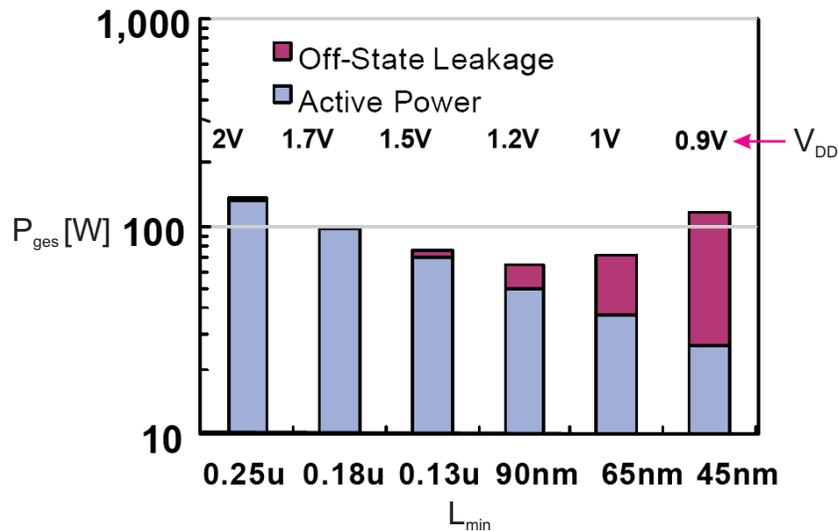


Abbildung 1.13: Gesamtleistung nach Gleichung 1.2 und 1.15

Während die Kurzschlussleistung nach Gleichung 1.2 für ein bestimmtes Verhältnis von  $\tau_{r,f}$  zu  $T$  einen konstanten Wert annimmt, skaliert die dynamische Leistung sowohl mit der Lastkapazität  $C_L$ , als auch mit der Frequenz  $f$ . Dies führt bei der Gesamtleistungsberechnung unter Vernachlässigung der Leakageleistung zu einem in Abbildung 1.13 zu beobachtbaren Maximum bei großen Lastkapazitäten und hohen Taktfrequenzen. Anzustreben ist also bei gegebenem Takt, die Lastkapazität so gering wie möglich zu halten bzw. die Taktfrequenz nicht höher als notwendig zu gestalten. Abbildung 1.12 zeigt den dominierenden Einfluss der dynamischen Leistung bei der Gesamtleistungsaufnahme. Wird diese Aussage mit Gleichung 1.15 in Verbindung gebracht, unterstreicht dies die Bemühungen, die Versorgungsspannung so niedrig wie möglich zu halten.

Der Einfluss der statischen Leistungsaufnahme ist komplexer. Kapitel 1.2 unterstreicht die zunehmende Bedeutung der statischen Leistungsaufnahme mit jedem Technologiesprung. Eine Abschätzung des Anteils der statischen Leistungsaufnahme für verschiedene Technologien liefert [24] (Abbildung 1.14):

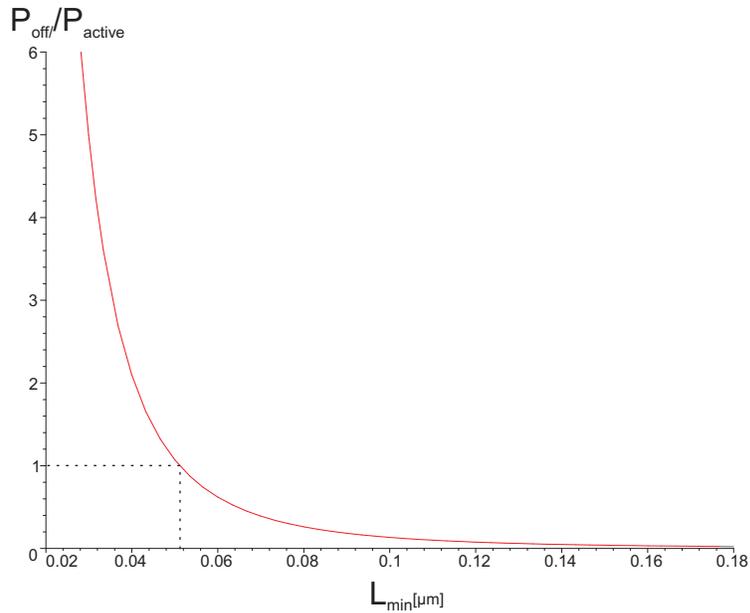


**Abbildung 1.14:** Anteil der statischen Leistung an der Gesamtleistung für unterschiedliche Technologien [24]

Das Schaubild nach Abbildung 1.14 bezieht sich auf eine große Digitalschaltung mit einer normierten Gesamttransistorweite (bei minimaler Länge) von  $\sum W_i = 30m$  (ca. 100 Millionen Transistoren in  $0.18 \mu m$  Technologie). Deutlich ist hier der immer dominierender werdende Anteil der statischen Verlustleistung für jede Verkleinerung der minimalen Transistorlänge zu erkennen. Abbildung 1.15 zeigt das Verhältnis von statischer Verlustleistung ( $P_{off}$ ) zu dynamischer Verlustleistung ( $P_{active}$ ) abgeleitet aus Abbildung 1.14 in Abhängigkeit der Technologiegröße  $L_{min}$ . Die den Daten aus Abbildung 1.14 angeglichene Funktion  $\frac{P_{off}}{P_{active}}$  ergibt sich zu:

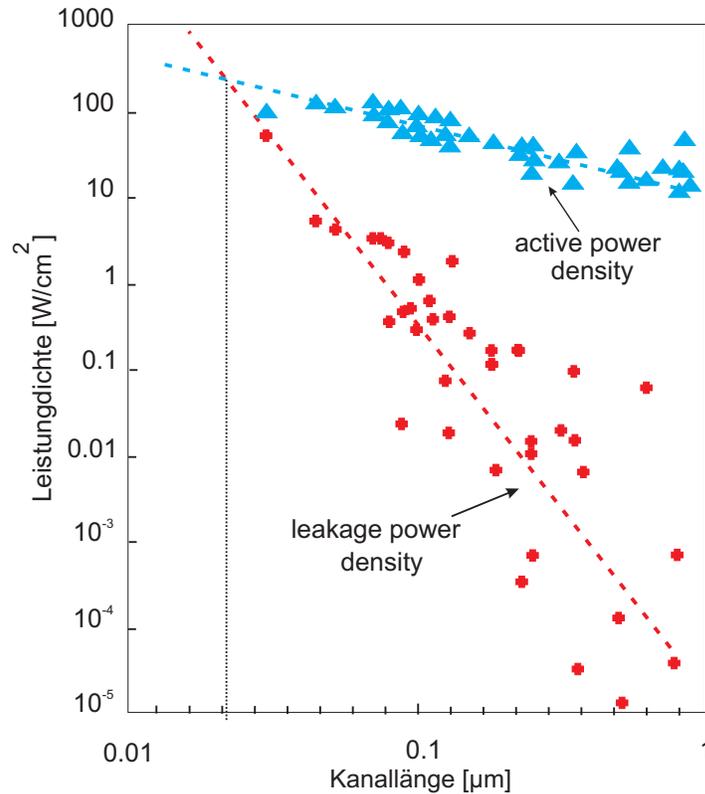
$$\frac{P_{off}}{P_{active}} = \frac{11.6 (\mu m)^3 \cdot 10^{-5}}{0.87 \cdot L_{min}^3} \quad (1.20)$$

wobei  $L_{min}$  in  $\mu m$  angegeben ist.



**Abbildung 1.15:** Verhältnis der statischen zur dynamischen Verlustleistung in Abhängigkeit der minimalen Transistorlänge

Einen noch besseren Eindruck liefert Abbildung 1.16 [33]. Durch Untersuchungen an vorhandenen Prozessen und Extrapolation findet sich eine Grenze von  $L_{min} = 20nm$ , für welche der Gleichstand der aktiven Leistung mit der Leakageleistung prognostiziert wird ( $T_{junction} = 20^{\circ}C$ ). Zieht man in Abbildung 1.16 die Chipwärmung auch in Betracht, erhöht sich die Leakage-Leistung, vor allem die Subthreshold-Leistung (vergl. Abbildung 1.10), so dass der Schnittpunkt der beiden Leistungsgeraden in Richtung größerer Kanallänge verschoben und mit  $L_{min} = 50nm$  prognostiziert wird [33], was durch die Abbildung 1.15 bestätigt wird.



**Abbildung 1.16:** Leistungsdichte der statischen und aktiven Leistung in Abhängigkeit der minimalen Gatelänge

# Kapitel 2

## Reduzierung der statischen Leistungsaufnahme

### 2.1 Multi-Threshold und Dynamic-Threshold CMOS-Technologien

In diesem Kapitel sollen Schaltungstechniken untersucht werden, die die Thresholdspannung der Transistoren gezielt beeinflussen. Hierzu zählen:

- Verwendung von HighSpeed HS bzw. LowLeakage LL Gattern (s. Kap. 2.2)
- Designen von Gattern mit Transistoren unterschiedlicher Thresholdspannung
- Dynamische Beeinflussung der Thresholdspannung durch Bulkpotentialsteuerung (s. Kap. 2.6 und 2.2)

Das Ausstatten einzelner Gatter mit Transistoren unterschiedlicher Thresholdspannung basiert auf der Untersuchung der statistischen Schaltungshäufigkeit. Bei einem NAND2 Gatter z.B. gibt es doppelt so viele Pfade vom Ausgangsknoten nach VDD als nach GND. Unter der Annahme statistisch gleich wahrscheinlicher Eingangssignalkombinationen ist es demzufolge doppelt so wahrscheinlich, dass auf eine berechnete 1 am Ausgang erneut eine 1 erzeugt wird (bei Gattern mit mehr Eingängen, wie NAND3, NOR3 etc., erhöhen sich die Wahrscheinlichkeiten entsprechend). Also erscheint es sinnvoll, die P-MOS Zweige des NAND Gatters mit Transistoren niedriger Thresholdspannung (sog. HighSpeed- bzw. HS-Typen) auszustatten (entsprechend die NMOS Zweige bei NOR-Gattern). Die sperrenden NMOS-Transistoren werden entsprechend mit hoher Thresholdspannung (sog. LowLeakage- bzw LL-Typen) ausgestattet. Diese tragen zu einem entsprechend kleinen Leakagestrom in drei von vier Fällen bei. Nachteilig erweist sich hierbei die längere Berechnungsdauer für eine 0 am Ausgang

und den in diesem Fall nicht reduzierten statischen Strom. Abbildung 2.1 verdeutlicht das Schaltfälle für ein Schaltermodell eines NAND2-Gatters. In drei von vier Fällen (Abbildung 2.1 links) profitieren sowohl Schaltgeschwindigkeit, als auch Subthreshold-Leakage-Strom. Das komplexe Schaltverhalten (Timing) muss beim Design berücksichtigt, und durch Simulation die Einhaltung der Performance gesichert werden.

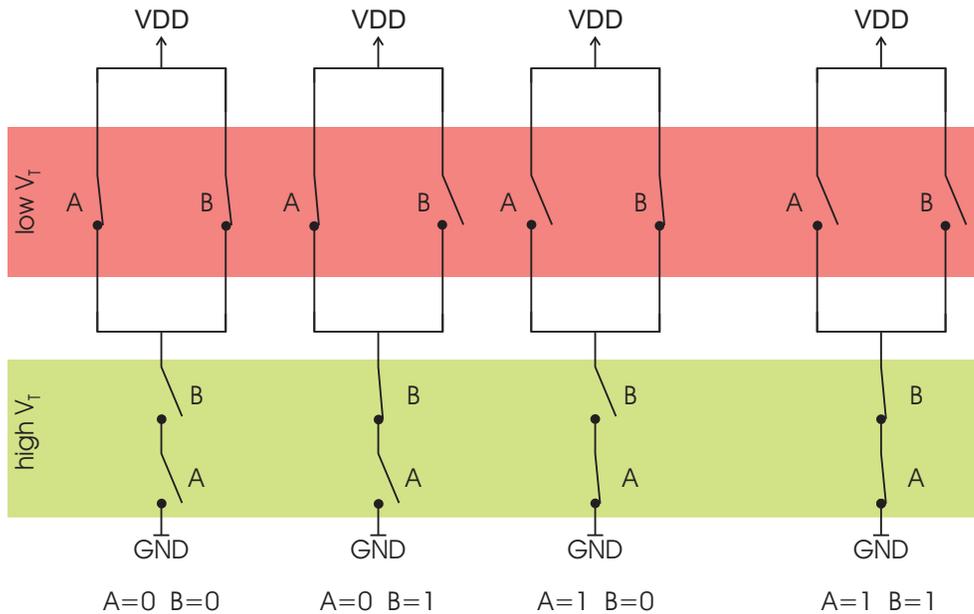


Abbildung 2.1: Schaltermodell NAND2-Gatter

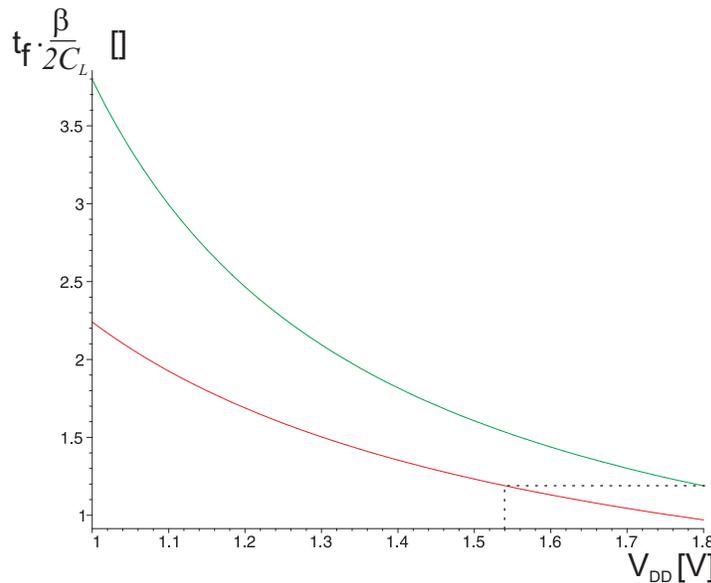
Die Anwendung dieses Konzepts erweist sich als schwierig aufgrund der von den Eingangskombinationen abhängigen Laufzeiten. Es muss eine volle statistische Analyse der Schalthäufigkeiten vorgenommen werden. Der kritische Pfad als Designindiz für die Performance einer Schaltung kann durch die abhängigen Laufzeiten im Betrieb variieren. Soll eine automatische Synthetisierung eines VHDL-Codes erfolgen, müssen alle Gatter in vier verschiedenen Versionen vorliegen (pro Zweig HS oder LL, bzw voll-HS oder voll-LL). Wegen des hohen Aufwands und des zu erwartenden kleinen Gewinns wird dieses Konzept nicht weiter verfolgt.

Eine andere Vorgehensweise ist der gemischte Einsatz von Gattern, die komplett mit LL-Transistoren aufgebaut sind mit Gattern, die vollständige mit HS-Transistoren aufgebaut sind. Bei der Synthese eines HDL Codes werden alle Standardzellen in der Ausführung LL gewählt. Anschließend wird der kritische Pfad durch Simulation evaluiert. Ergibt die Simulation der aus dem Layout extrahierten Schaltung, dass die Spezifikationen hinsichtlich der Verzögerungszeiten nicht erfüllt werden, werden die Gatter im kritischen Pfad sukzessive

durch HS-Typen ersetzt. Danach wird erneut extrahiert und die Spezifikationen überprüft. Eventuell ergeben sich durch die Ersetzung einzelner Gatter andere kritische Pfade, was bei der Resynthese berücksichtigt werden muss. Dieses Verfahren wird solange angewandt, bis die Spezifikation erfüllt ist. Hierbei ist darauf zu achten, dass das Ersetzen einzelner Gatter nicht zu Timing-Problemen, wie Races, führt. Dieses Vorgehen passt die Performance der Schaltung an die Spezifikation an.

## 2.2 Versorgungsspannungsreduzierung bei äquivalenter Performance

Das Verwenden von zwei unterschiedlichen N- und PMOS-Transistortypen stellt einen nicht unerheblichen technologischen Mehraufwand dar. Interessant ist der Ansatz, nur N- und PMOS-Transistoren einer Schwellenspannung zu verwenden. Da die Performance einer Schaltung nicht nur durch die Schwellenspannung bestimmt wird, sondern in starkem Maße auch von der Versorgungsspannung, wird untersucht, ob die HS-Technologie mit reduzierter Versorgungsspannung ähnliches Einsparungspotential hinsichtlich statischer und dynamischer Verlustleistung bietet. Betrachtet werden Standard-CMOS Schaltungen jeweils in LL- und HS-Technologie. Die Gatter werden im Hinblick auf symmetrisches Umschaltverhalten dimensioniert. Die absolute Größe der Transistoren ist bei LL-Gattern und HS-Gattern jeweils identisch, um aussagefähige Vergleiche durchführen zu können. Die Umschaltgeschwindigkeit der Gatter wird nun durch Absenken von  $V_{DD}$  der HS-Gatter an die der LL-Gatter angeglichen. Bei der betrachteten Technologie stellt sich eine Performanceäquivalenz der HS-Gatter bei einer Betriebsspannung von  $1.45V$  ein. Theoretisch lässt sich diese Spannung aus Gleichung 2.3 für die Verzögerungszeit eines Inverters herleiten.



**Abbildung 2.2:** Rise-Time eines CMOS-Inverters nach Gleichung 2.3

Abbildung 2.2 zeigt die normierte Verzögerungszeit eines Inverters nach Gleichung 2.3 (hier  $t_{fall}$ ,  $t_{rise}$  analog) für die Thresholdspannungen  $V_{TLL} = 0.5V$  (grün) und  $V_{THS} = 0.3V$  (orange). Diese Thresholdspannungen spiegeln in etwa die Verhältnisse der Zieltechnologie wieder. Markiert ist die Spannung

$V_{DD_{HS}} \simeq 1.54V$ , bei der die Verzögerungszeit des HS-Inverters die gleiche ist, wie die des LL-Inverters bei  $V_{DD_{LL}} = 1.8V$ , was in Anbetracht der einfachen Modellierung eine gute Übereinstimmung mit dem simulierten Wert darstellt. Diese Verringerung der Versorgungsspannung geht in alle Verlustleistungen verbessernd ein. Dennoch werden mit LL-Gattern geringere Leakage Verlustleistungen erzielt. Die Gesamtverlustleistung ist jedoch stark frequenzabhängig, da die dynamische Verlustleistung frequenzabhängig ist (vergl.: Gleichung 1.14).

## 2.3 Vergleich von Gattern in LowLeakage- und HighSpeed-Technologie

MOS-Transistoren in LowLeakage(LL)-Technologie besitzen eine höhere Threshold - Spannung  $U_T$  als HighSpeed-Typen.

Deshalb tritt bei ihnen ein geringerer Leakage-Strom im Sperrbereich auf (vergleiche Gleichung 1.6); die Verzögerungszeit steigt aber durch ein geringeres  $U_{GS_{eff}} = U_{GS} - V_T$ , was durch die Herleitung der Gleichung für die Verzögerungszeit eines CMOS-Inverters als Referenz für jedes beliebige CMOS-Gatter verdeutlicht werden kann. Das Entladen einer Lastkapazität  $C_L$  von  $0.9V_{DD}$  bis  $V_{DD} - V_T$  wird hierbei als *falltime*  $t_f$  bezeichnet und berechnet sich wie folgt:

Für den Inverter gilt:  $-\frac{C_L}{I_D} \cdot dU_{out} = dt$   
mit:

$$\begin{aligned} Q_C &= C_L U_{out} \\ \Rightarrow \frac{d}{dt} Q_C &= I_D = C_L \frac{d}{dt} U_{out} \\ \Rightarrow \frac{I_D}{C_L} &= \frac{d}{dt} U_{out} \end{aligned} \quad (2.1)$$

lässt sich  $t_f$  dann folgendermaßen bestimmen:

$$t_f = \int_{V_{DD}-U_T}^{0.9 \cdot V_{DD}} \frac{2C_L}{\beta_n} \cdot \frac{dU_{out}}{U_{GS_{eff}}^2} + \int_{0.1 \cdot V_{DD}}^{V_{DD}-U_T} \frac{2C_L}{\beta_n} \cdot \frac{dU_{out}}{(2U_{GS_{eff}} - U_{out})U_{out}} \quad (2.2)$$

Nach der Integration ergibt sich für die *falltime* (*risetime* analog) eines Inverters [52]:

$$t_f = \frac{2C_L}{\beta} \cdot \left[ \frac{U_T - 0.1V_{DD}}{(U_{GS_{eff}})^2} + \frac{1}{2U_{GS_{eff}}} \cdot \ln \left( \frac{V_{DD} - U_T}{2U_{GS_{eff}} - U_{GS_{eff}}} \frac{2U_{GS_{eff}} - 0.1V_{DD}}{0.1V_{DD}} \right) \right] \quad (2.3)$$

Diese Gleichung ist in Abb. 2.3 grafisch dargestellt. Dabei liegt  $V_{T0}$  bei realen Transistoren etwa zwischen  $0.2$  und  $0.5 \cdot V_{DD}$  (ST.  $0.18\mu m$ ,  $V_{DD}=1.8$  V,  $V_{T0}$  ca.  $500$  mV (LL)). Die zur grafischen Darstellung in Gleichung 2.3 eingesetzten Werte ( $\beta_0=264 \frac{\mu A}{V^2}$ ) sind dem MOS-Modell Level 903 entnommen. Zusätzlich wird  $C_L$  zu  $1pF$  und  $W/L$  zu  $\frac{W}{L} = \frac{0.28}{0.18}$ , also dem in dieser Technologie kleinst möglichen Wert, angenommen.

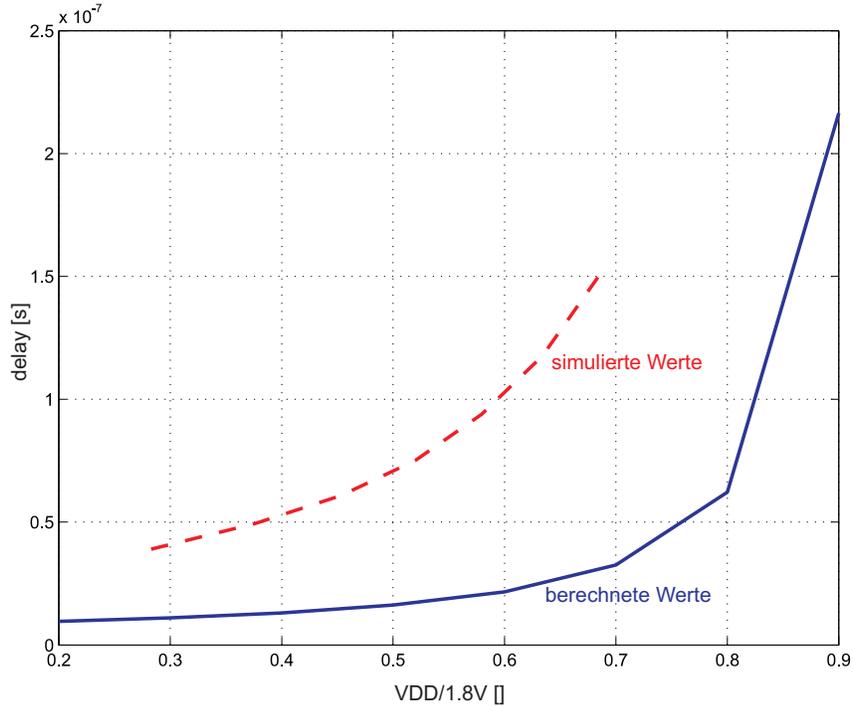


Abbildung 2.3: *falltime* eines Inverters abhängig von  $V_T$

Ein Vergleich der theoretischen zu den in einer Simulation ermittelten Werten zeigt deutliche Abweichungen im Kurvenverlauf, welche in der genaueren Modellierung bei der Simulation begründet liegt.

HighSpeed(HS)-Transistoren besitzen eine kleinere Threshold - Spannung. Die HS-Technologie bietet dadurch, bei identischer Dimensionierung, zwar kürzere Schaltzeiten als die LL-Technologie, durch den höheren Leakage-Strom steigt aber auch die statische Verlustleistung, sowie durch die höheren Ladeströme auch die dynamische Verlustleistung. Durch Absenkung der Versorgungsspannung von 1.8V auf 1.45V bei der HS-Technologie werden die gleichen Umschaltzeiten erreicht wie bei der Verwendung von LowLeakage-Technologie mit einer Versorgungsspannung von  $V_{DD} = 1.8$  V. Dies zeigen folgende Simulationsergebnisse.

## 2.4 Simulationsergebnisse

Um die beiden Technologien hinsichtlich Tauglichkeit in LowPower - Anwendungen vergleichen zu können, wird die Performance beider Schaltungstechnologien angeglichen, indem die Versorgungsspannung bei den HS-Gattern solange abgesenkt wird, bis diese die gleiche Verzögerungszeit wie die LL-Gatter bei der Standardversorgungsspannung  $V_{DD} = 1.8$  V aufweisen. Es stellt sich somit für die HS-Gatter ein  $V_{DD}$  von 1.45 V ein. Die Verzögerungszeit beträgt dann jeweils 11.32ns. Ohne diese Angleichung beträgt sie bei den HS-Gattern nur ca. 9.6ns. Die genannten Werte gelten näherungsweise für alle betrachteten digitalen Gatter (NAND, NOR, Inverter, XOR), weil die Transistoren der Gatter jeweils so dimensioniert werden, dass sie die gleiche Performance besitzen wie der Inverter. Als Eingangssignale dienen Rechtecksignale der angegebenen Frequenz; der Ausgang wird mit einer Kapazität von 1pF belastet. Für nachfolgende Simulationen gilt:  $V_{DD_{HS}} = 1.45$  V,  $V_{DD_{LL}} = 1.8$  V.

### 2.4.1 Vergleich anhand einfacher Gatter

Der Vergleich der beiden Technologien wird anhand von einfachen digitalen Gattern wie einem Inverter, einem NAND2- und einem NOR2-Gatter durchgeführt. Die Ergebnisse einer Simulation mit Ausgangslast  $C_L = 1$  pF und einer Simulationsfrequenz von 1 MHz sind in Tabelle 2.1 zusammengestellt.

	$I_{leak}$ in A		$P_{stat}$ in W		$P_{ges}$ in W $\approx P_{dyn}$	
	HS @1.45V	LL @1.8V	HS @1.45V	LL @1.8V	HS @1.45V	LL @1.8V
Inverter	9.45e-10	1.52e-11	1.37e-9	2.73e-11	1.05e-6	1.62e-6
NAND	6.19e-10	3.05e-12	8.98e-10	5.49e-12	5.26e-7	8.13e-7
NOR	1.89e-9	3.64e-11	2.74e-9	6.54e-11	1.05e-6	1.63e-6

**Tabelle 2.1:** Leakage-Strom und Verlustleistung einfacher Gatter

Der Leakage-Strom und damit auch die statische Verlustleistung ist, wie zu erwarten, bei der LL-Technologie bei jedem Gatter am geringsten. Bei Betrachtung der gesamten Leistung dominiert der dynamische Anteil an der Verlustleistung. Durch die Reduzierung der Versorgungsspannung erreicht hier die HighSpeed-Technologie kleinere Verlustleistungen. Obige Tabelle zeigt die geringe Bedeutung des statischen Leistungsanteils schon bei Frequenzen, die weit unterhalb üblicher Werte liegen: weniger als 0.1 Promille.

### 2.4.2 Frequenzbetrachtung

Interessant ist die Frage, ab welcher Schaltfrequenz (am Gattereingang) die HS-Technologie mit reduzierter Versorgungsspannung geringere GesamtLeistungsverluste produziert als die LL-Technologie mit identischer Schaltgeschwindigkeit. Hierzu wurde eine Kette aus 5 hintereinander geschalteten Standard-Gattern jeweils in HS- und LL-Technologie aufgebaut und die Gesamtleistung bei unterschiedlichen Eingangssignalfrequenzen ermittelt. Die Simulationsergebnisse sind in Tabelle 2.2 zusammengestellt und in der Abbildung 2.4 grafisch dargestellt. Dabei ist  $V_{DD_{HS}} = 1.45V$  und  $V_{DD_{LL}} = 1.8V$ .

Frequenz in Hz	Verlustleistung $P_V$ in W	
	LL @1.8 V	HS @ 1.45 V
1	1.88e-10	7.55e-9
10	2.18e-10	7.57e-9
100	5.23e-10	7.77e-9
1K	3.57e-9	9.57e-9
10K	3.40e-8	2.95e-8
100K	3.34e-7	2.27e-7
1M	3.38e-6	2.2e-6
10M	3.38e-5	2.20e-5
100M	1.623e-4	1.21e-4

**Tabelle 2.2:** Verlustleistung in Abhängigkeit der Frequenz

Abbildungen 2.4 zeigt, dass bei Frequenzen unterhalb von ca. 10 kHz die LowLeakage-Technologie weniger Gesamtverlustleistung umsetzt als die High-Speed-Technologie. Bei Frequenzen ab ca. 10kHz ist die HighSpeed-Technologie im Vorteil. Schaltfrequenzen von wenigen Kilohertz kommen in modernen digitalen Systemen nicht vor, weswegen für die Verwendung der LowLeakage-Technologie bei oberflächlicher Betrachtungsweise wenig spricht.

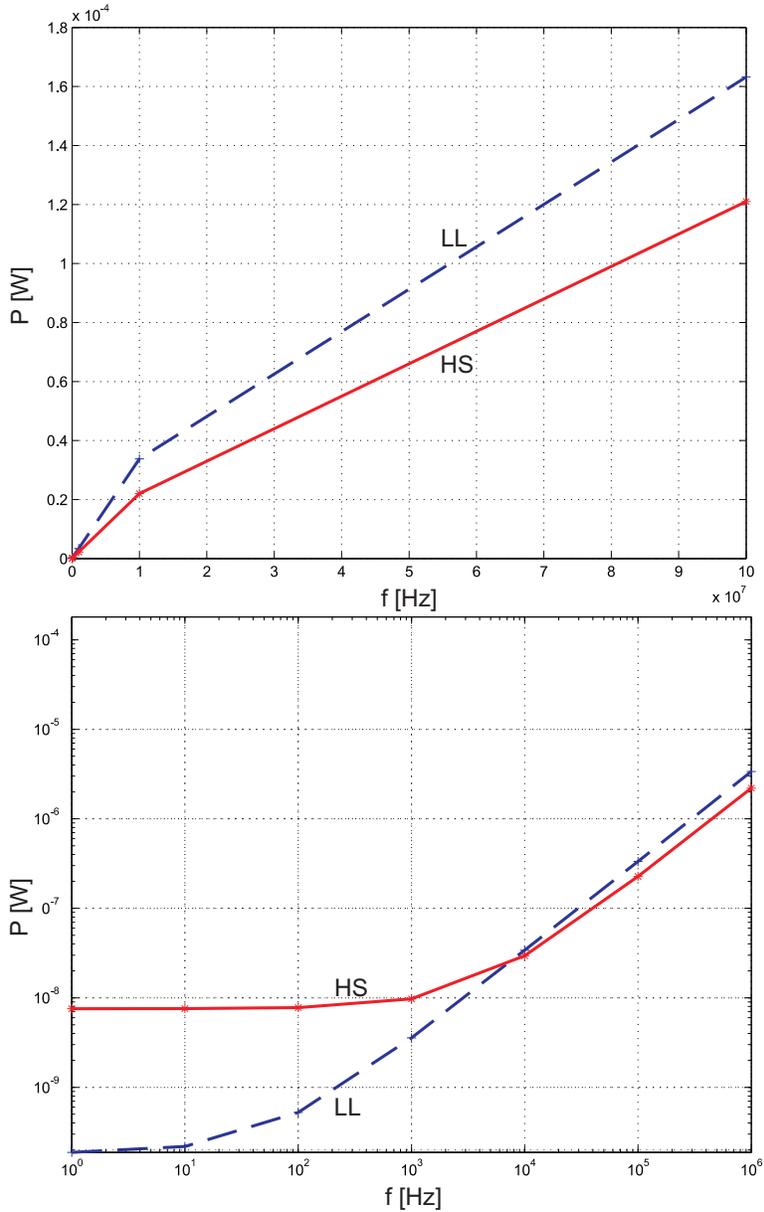


Abbildung 2.4: Vergleich: HS und LL NOR

### 2.4.3 Vergleich bei Addiererschaltungen

Als Beispiel für eine komplexere Schaltung wird nun eine Addiererschaltung betrachtet. Als Addiererstruktur wird bei den Simulationen der 8bit Ripple-Carry-Addierer gewählt. Ein Vergleich der Addierer für den Fall konstanter Eingänge (alle Eingänge liegen auf Massepotential) zeigt die verhältnismäßige kleine statische Verlustleistung der LowLeakage Technologie.

Wiederum findet der Vergleich bei angenäherten Verzögerungszeiten statt ( $V_{DD-LL} = 1.8V$  und  $V_{DD-HS} = 1.45V$ ,  $\tau = 11.9ns$ ).

	LL @ 1.8 V	HS @ 1.45 V
$I_{leak}$ in A	1.113e-9	63.57e-9
$P_V$ in W	2.0e-9	92.2e-9

**Tabelle 2.3:** 8-bit Addierer statischer Fall

Um die Simulationszeit zu reduzieren, wird bei den nächsten Vergleichen ein 4bit-Addierer betrachtet, dessen Eingänge sich in der Art periodisch ändern, dass eine maximale Schaltaktivität erreicht wird. In Tabelle 2.4 ist die Verlustleistung in Abhängigkeit von der Eingangssignalfrequenz angegeben, weil diese allein auch aussagekräftig ist.

	LL @ 1.8 V	HS @ 1.45 V
$P_V$ in W bei 20kHz	8.94e-8	1.0237e-7
$P_V$ in W bei 1MHz	4.41e-6	2.922e-6
$P_V$ in W bei 2MHz	8.824e-6	5.80e-6

**Tabelle 2.4:** 4-bit Addierer bei unterschiedlichen Signalfrequenzen

Analog zum Ergebnis der Einzelgatterbetrachtung zeigt sich das Einsparpotential hinsichtlich der Gesamtverlustleistung der HighSpeed-Technologie mit reduzierter Versorgungsspannung. Die LowLeakage-Technologie mit identischen Schaltzeiten besitzt diesbezüglich nur bei irrelevant kleinen Signalfrequenzen einen Vorteil. Ein Hinweis auf die dominante dynamische Verlustleistung bei höheren Schaltfrequenzen ist die annähernde Verdopplung der Gesamtverlustleistung bei 2-facher Schaltfrequenz.

Wird die Gesamtverlustleistung bei  $1MHz$  als Referenz benutzt und auf eine Schaltfrequenz von  $20kHz$  skaliert ( $P_V \cdot 0.02$ ), wird der Anteil nicht dynamischer Leistung ersichtlich (LL:  $82nW$  zu  $89.4nW$ ; HS:  $58.44nW$  zu  $102.37nW$  vergl. Tabelle 2.4). Somit liegt der Anteil statischer Verlustleistung bei  $20kHz$  für die LowLeakage Schaltung nur bei  $8.2\%$ , bei der HighSpeed Schaltung dagegen bei  $42.9\%$ .

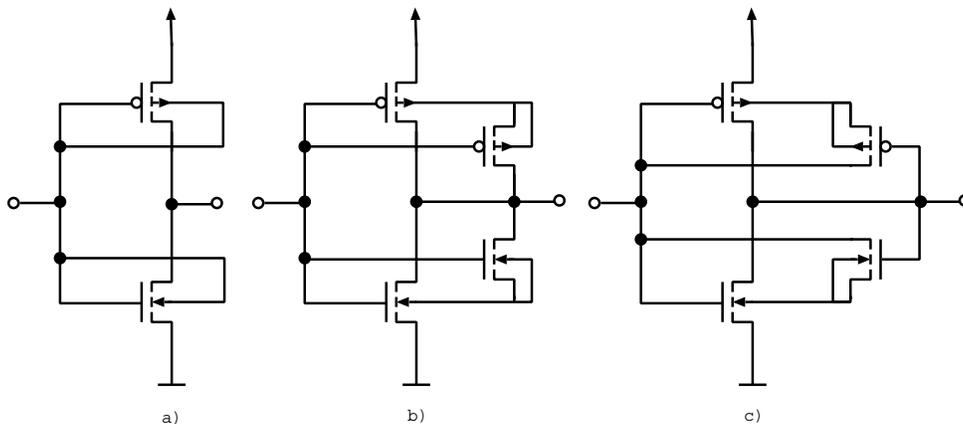
## 2.5 Bewertung der LowLeakage bzw. HighSpeed-Technologie

Durch Reduktion der Versorgungsspannung und Verwendung von Transistoren ausschließlich niedriger Thresholdspannung (HS-Technologie) wird eine signifikante Verringerung des Gesamtleistungsverbrauchs erreicht, bei ansonsten gleichen Schaltzeiten wie bei der Verwendung von LowLeakage Gattern mit Standardversorgungsspannung. Der damit einhergehende niedrigere Signalrauschabstand ist für digitale Anwendungen unerheblich. Bei reduzierten Anforderungen an die Schaltfrequenz kann die Versorgungsspannung sowohl bei Verwendung von HighSpeed- als auch von LowLeakage-Technologie weiter verringert werden (vergl. Kap 3.2), so dass bei üblichen Signalfrequenzen ( $>100\text{kHz}$ ) die HighSpeed-Technologie immer zu bevorzugen ist. Nur bei sehr kleinen Signalfrequenzen bzw. im StandBy Betrieb erzielt die LowLeakage Technologie geringere Gesamtverlustleistungen. Es muss also eine Abwägung zwischen zu erwartender StandBy Zeit und Normalbetrieb vorgenommen werden.

Bei Priorität kleiner Leakage-Verlustleistungen ist die LowLeakage-Technologie immer vorzuziehen. Für künftige Technologien wird die Leakageleistung jedoch exponentiell ansteigen, so dass für diese die Einsatzfrequenz, unter welcher ein Einsatz von Transistoren höherer Thresholdspannung lohnend ist, steigen wird. Die Verwendung von LowLeakage Transistoren ist dann angebracht, wenn minimaler Leistungsverbrauch Priorität über die Schaltfrequenz hat, oder, bei gemischtem Einsatz von LowLeakage und HighSpeed-Transistoren, um die Performance sukzessive an die Spezifikation anzupassen (vergl Kapitel 2.2). Bei gleicher Versorgungsspannung erreicht die LowLeakage-Technologie nämlich kleinere Gesamtleistungsverluste als die HighSpeed-Technologie bei Inkaufnahme einer reduzierten Taktfrequenz und somit einer kleineren Rechenleistung. Somit kommt die Anwendung der LowLeakage-Technologie immer dann in Frage, wenn die Taktanforderung so gering ist, dass die dazu erforderliche Versorgungsspannung bei Verwendung der HighSpeed-Technologie unrealistisch klein wäre, was einen stark reduzierten Störabstand zur Folge hätte. Darüber hinaus bestünde die Gefahr, dass die Levelshifter außerhalb ihrer Spezifikation arbeiten, wenn die Signalhübe sehr klein werden. Die Kombination von geringem Leakagestrom im StandBy Betrieb und hoher Signalverarbeitungsgeschwindigkeit wird im Folgenden untersucht.

## 2.6 Body-Biasing und Dynamic Threshold CMOS-Technologie

Unter dem Begriff *(Active)-Body-Biasing* bzw. *Dynamic-Threshold-CMOS* (DT-CMOS) versteht man eine CMOS-Schaltungstechnik, bei der abhängig vom Eingangs- und Ausgangspotential eines Gatters das Bulkpotential beschaltet wird. Bei vielen Standard-CMOS Technologien, insbesondere bei Technologien, die bevorzugt für Digitalschaltungen verwendet werden, sind die NMOS Transistoren nicht vom Substrat isoliert, so dass deren Bulkpotential nur global veränderlich ist. Die in der Literatur besprochenen Konzepte zum Thema Active Body-Bias (z.B. [34]) beziehen sich aus diesem Grunde vorwiegend auf SOI-Technologien. Bei der in der vorliegenden Dissertation verwendeten Technologie ( $0.18\mu\text{m}$  tripple well) sind isolierte Wannen für NMOS-Transistoren möglich, weswegen die Konzepte zum Active-Body-Biasing auf diese in Grenzen übertragbar sind. Abbildung 2.5 zeigt exemplarisch in der Literatur vorgeschlagene Inverterschaltungen mit Active-Body-Biasing.



**Abbildung 2.5:** Inverterschaltungen mit Active-Body-Biasing

Ein kurzer Überblick über weitere Body-Biasing-Konzepte findet sich unter anderem in [34]. Allen Konzepten, die aktiv Einfluss auf die Thresholdspannung nehmen, indem sie das Bulk des Transistors ansteuern, ist gemeinsam, dass sie einen stark erhöhten Flächenbedarf besitzen, wenn sie in Standard-CMOS-Prozessen realisiert werden sollen. Bei diesen ist das Bulk der NMOS-Transistoren normalerweise identisch mit dem Substrat. Soll nun aktiv auf das individuelle Bulkpotential jedes Transistors Einfluss genommen werden, müssen die NMOS-Transistoren einzeln vom Substrat isoliert werden, was prinzipiell durch Isolationsschichten und tripple-well Technologie möglich ist. Der Platzbedarf durch die Isolationsschicht steigt aber im Vergleich zur Standardlösung drastisch.

Gleiches gilt für die PMOS-Transistoren. Soll auf das Bulk eines einzelnen Transistors individuell zugegriffen werden, muss jeder PMOS-Transistor eine eigene N-Wanne erhalten, mit dem sich durch die Design-Rules ergebenen Flächenzuwachs durch Abstands- und Kontaktregeln. Die Abbildungen 2.6, 2.7 und 2.8 sind zur besseren Vergleichbarkeit im gleichen Maßstab dargestellt und verdeutlichen diesen Umstand anhand eines minimalen CMOS-Inverters. Der Platzbedarf steigt, ausgehend vom Standardlayout, um 80% (Abbildung 2.7) bzw. um 1500% (Abbildung 2.8).

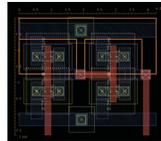


Abbildung 2.6: Treiberlayout in Standard CMOS-Technologie

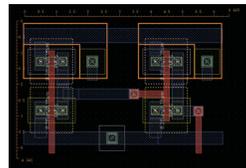


Abbildung 2.7: Treiberlayout in CMOS-Technologie mit isolierten N-Wannen

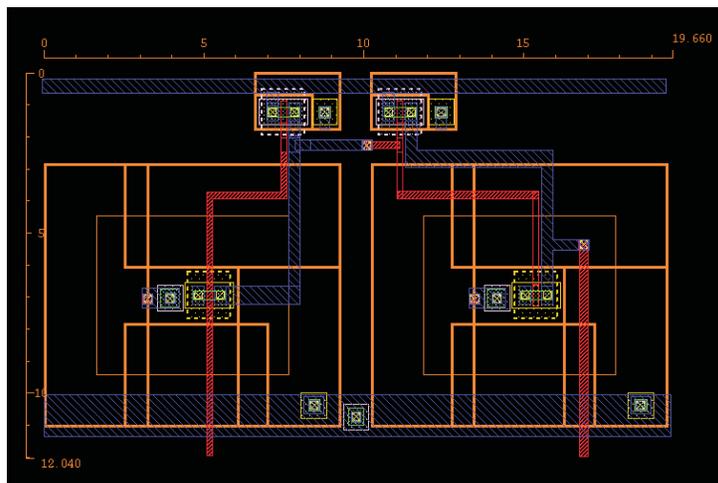


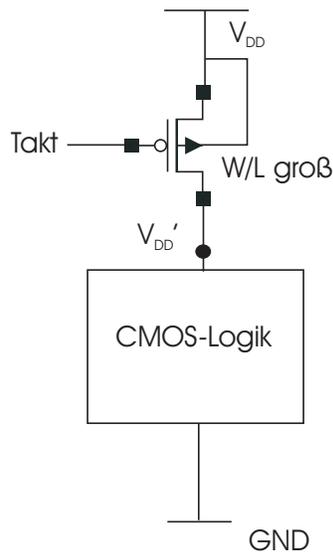
Abbildung 2.8: Treiberlayout in CMOS-Technologie mit isolierten N-Wannen und isolierten P-Bulks

Diese Verhältnisse verschlechtern sich noch weiter, wenn Schaltlogik zur Ansteuerung der Bulks mitintegriert werden muss, was in den o.a. Beispiellayouts noch nicht der Fall ist. Ausgehend von hoch integrierten digitalen Schaltungen sind diese Flächenzuwächse nicht akzeptabel. Ursache für den enormen Platzbedarf der Schaltung nach Abbildung 2.8 sind die Design-Regeln, die bei der Implementierung einer isolierenden Schicht im P-Substrat Anwendung finden. Anders als bei der SOI-Technologie stellen vom Substrat isolierte Bulks in CMOS-Technologie eine große Kapazität gegenüber dem Substrat dar. Eine aktive Umladung des Bulk im Takt der Gate-Signale zieht damit eine wesentlich erhöhte dynamische Verlustleistung nach sich. Dies macht die aktive Ansteuerung der Bulks in der Zieltechnologie unrentabel.

## 2.7 Sleep Transistor - Power-Gating

### 2.7.1 Vorstellung des Konzepts

Ein heute schon in der Praxis eingesetztes Konzept zur Reduzierung der Verlustleistung ist die Einführung eines sogenannten *Sleep-Transistors*. Bei diesem handelt es sich um einen zwischen Spannungsversorgung und eigentlichem Logikteil eingebauten LowLeakage PMOS-Transistor, der die Schaltung im passiven Fall von der Versorgungsspannung abtrennt und im aktiven Fall nach Möglichkeit keinen weiteren Einfluss auf die Funktion hat. Ein prinzipieller Aufbau solcher Schaltungen ist in Abbildung 2.9 dargestellt.

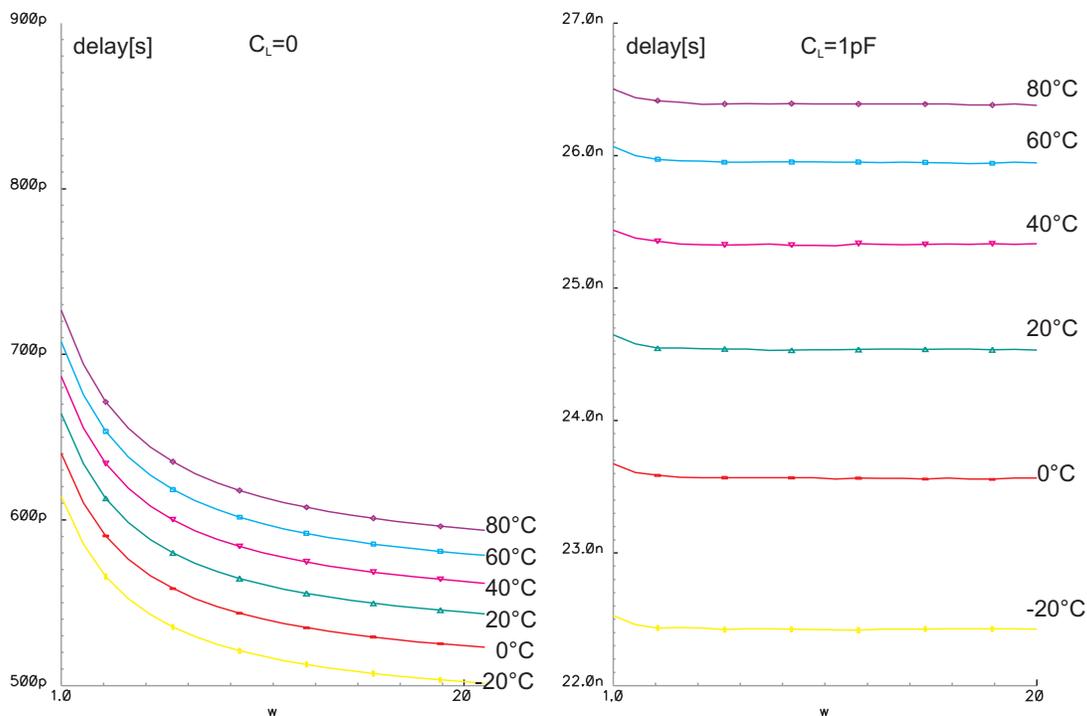


**Abbildung 2.9:** CMOS-Schaltung mit Sleep-Transistor

Die Verwendung eines NMOS-Transistors in LowLeakage Technologie, der den Logikteil von der Groundrail abtrennt, ist prinzipiell auch möglich. Bei Standard-Digital-Schaltungen werden aber aus Platzgründen meist NMOS-Transistoren verwendet, deren Bulk mit dem Substrat identisch ist. Durch den unvermeidlichen Spannungsabfall über dem Sleeptransistor kommt es somit auch zu einer Erhöhung der Thresholdspannung. Aus diesem Grund werden hier nur Konzepte mit PMOS-Sleep-Transistor betrachtet.

### 2.7.2 Dimensionierung des Sleeptransistors

Damit der Sleep-Transistor keinen weiteren Einfluss auf die Performance des Logikteils hat, muss sein W/L-Verhältnis sehr groß gewählt werden. Dann ist die Spannung  $U_{DS}$  über dem Sleep-Transistor klein und der Strom durch die CMOS-Logik im aktivem Betrieb wird nur unwesentlich begrenzt. Betrachtet wird eine Schaltung aus 12 seriellen Standardgattern. Ohne Sleeptransistor hat diese Schaltung ein delay von  $\tau_d = 520ps$ . Durch Einführen eines Sleeptransistors vergrößert sich das Delay, da, egal wie weit der Sleeptransistor dimensioniert ist, immer ein (wenn auch geringes)  $U_{DS}$  an ihm abfällt. Abbildung 2.10 zeigt das Delay der o.a. Schaltung mit Sleeptransistor abhängig von der Weite des Sleeptransistors bei minimaler Länge ohne zusätzliche Lastkapazität (links) und mit einer zusätzlichen Lastkapazität von  $C_L = 1pF$  (rechts).



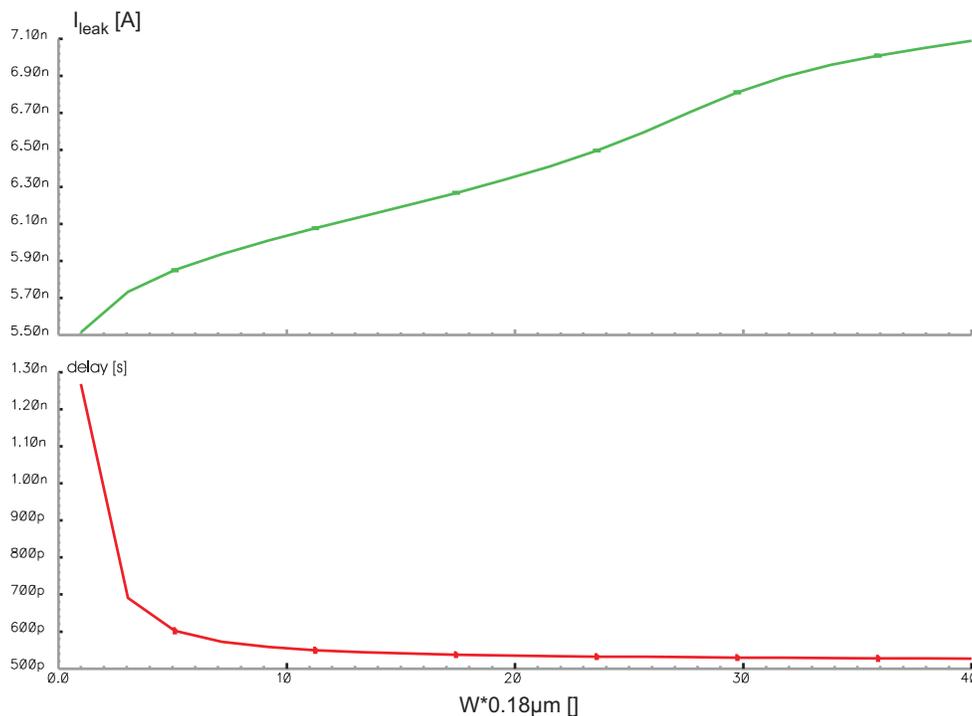
**Abbildung 2.10:** Delay einer Gatterkette mit Sleeptransistor in Abhängigkeit der Temperatur

Anhand von Abbildung 2.10 ist ersichtlich, dass die Dimensionierung des Sleeptransistor sowohl von Temperatur, als auch von der Last abhängig ist. Da der Spannungsabfall  $U_{DS-sleep}$  über dem Sleeptransistor lastabhängig ist, gelten für andere Lastströme auch andere Dimensionierungen. Das erklärt den Unterschied in der optimalen Dimensionierung des Sleeptransistors in Abbildung 2.10. Der Einfluss der Temperatur entspricht der Abhängigkeit des ON-Widerstands des Sleeptransistors. Je höher die Transistortemperatur ist, umso größer wird der

ON-Widerstand. Dies gilt aber für alle Transistoren, so dass sich die Dimensionierungsempfehlung für den Sleeptransistor nicht mit der Temperatur ändert, da sich die Schaltströme im Logikteil entsprechend mit ändern. Die Gesamtverzögerung skaliert absolut, die qualitative Veränderung durch den Sleeptransistor bleibt in etwa gleich.

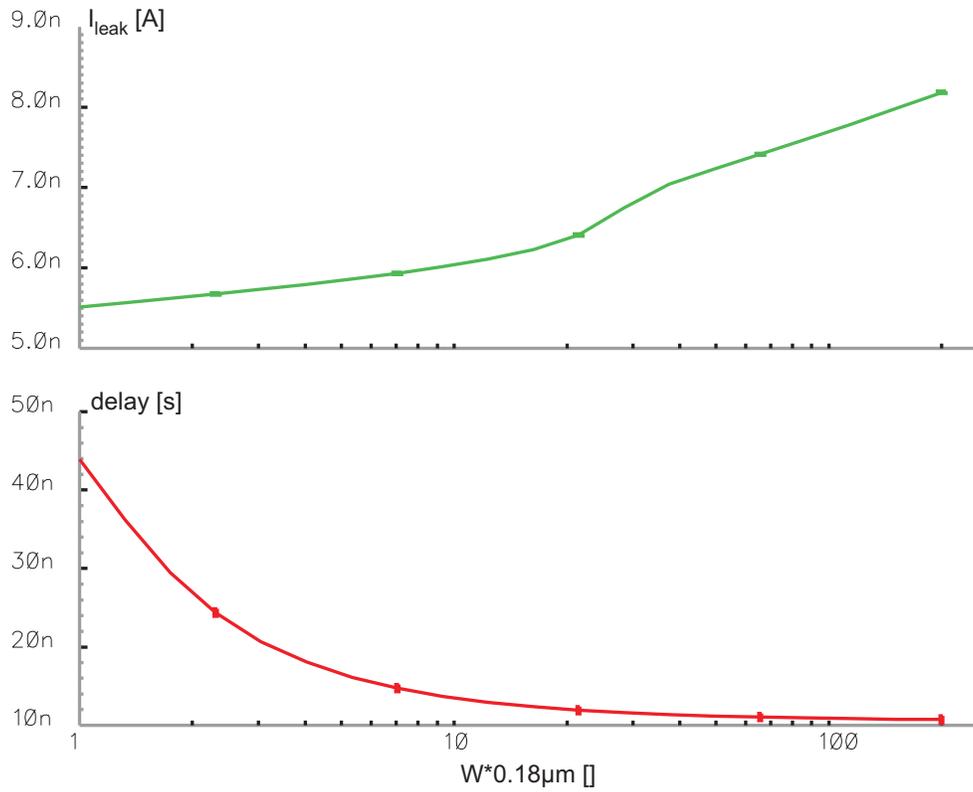
### 2.7.3 Reduzierung des statischen Leakagestroms durch Power-Gating

Soll der Sleeptransistor im Betriebsmodus möglichst keinen Einfluss auf die Schaltung haben, so soll er im Stand-By Modus den Leakage Strom von  $V_{DD}$  nach  $GND$  reduzieren. Dies wird zum einen durch die Verwendung eines Sleeptransistors in Low-Leakage-Technologie (hohe Thresholdspannung) erreicht, zum andern durch den Stack-Effekt bei der Serienschaltung von Transistoren [19]. Die Abbildungen 2.11 und 2.12 verdeutlichen den Zusammenhang zwischen Dimensionierung des Sleeptransistors, Einfluss auf das Delay der Schaltung und Reduzierung des Leakagestroms.



**Abbildung 2.11:** Delay und Leakagestrom einer Gatterkette mit Sleeptransistor

Die Reduzierung des Leakagestroms ist umgekehrt proportional zur Weite des Sleeptransistors. Da die Weite des Sleeptransistors auch umgekehrt proportional zur Erhöhung des Delays des Logikteils ist, entsteht hier ein trade-off, der darüberhinaus auch noch vom Laststrom abhängig ist. Die Dimensionierung des Sleeptransistors ist somit keine triviale Aufgabe; das sogenannte *optimum sizing* des Sleep-Transistors ist Thema vieler Veröffentlichungen [25].



**Abbildung 2.12:** Delay und Leakagestrom einer Gatterkette mit Sleeptransistor  $C_L = 1pF$

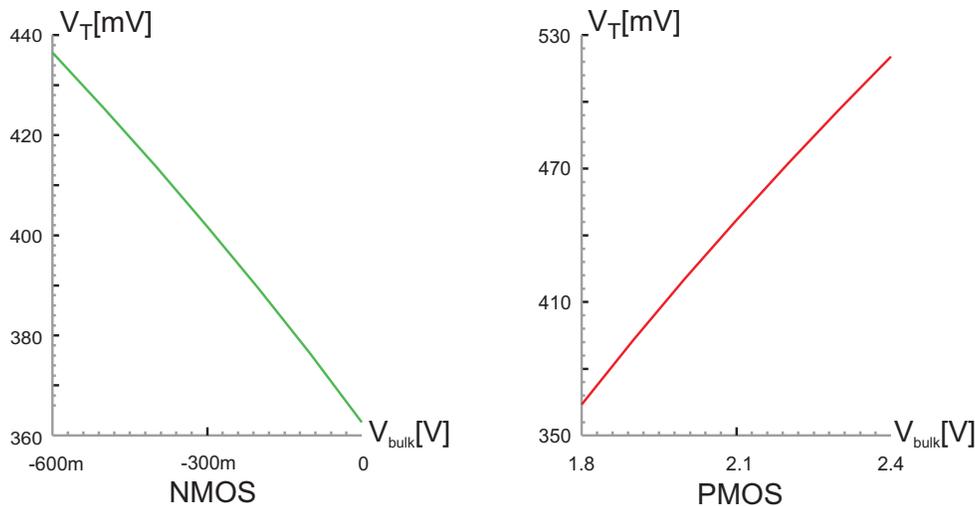
## 2.8 Reduzierung des Leakagestroms durch Bulkpotentialerhöhung

Eine weitere Methode, den Drain-Source-Leakage Strom zu reduzieren ist die dynamische Erhöhung der Thresholdspannung der Transistoren während des StandBy Betriebes. Da die Thresholdspannung von der Source-Bulk-Spannung abhängt (vergl. Gleichung 2.4), wird durch eine Erhöhung des Bulkpotentials auch eine Erhöhung der Thresholdspannung bewirkt. Nach Gleichung 1.6 sinkt somit auch der Subthreshold-Leakage Strom. Für die Thresholdspannung eines MOS-Transistors in Abhängigkeit seiner Source-Bulk-Spannung gilt näherungsweise:

$$V_T = V_{T0} + \gamma \cdot \sqrt{U_{SB} + 2 \cdot U_D} - \sqrt{U_{diff}} \quad (2.4)$$

$\gamma$  ist hierbei der technologieinherente Substratfaktor [ $1/\sqrt{V}$ ] und  $U_{diff} = \frac{kT}{q} \cdot \ln \frac{N_A N_D}{n_i^2}$  die Diffusionsspannung.

Abbildung 2.13 zeigt den Zusammenhang zwischen Bulkpotential und Thresholdspannung für minimal dimensionierte NMOS und PMOS-Transistoren der Zieltechnologie bei einer Versorgungsspannung von  $V_{DD} = 1.8V$ .

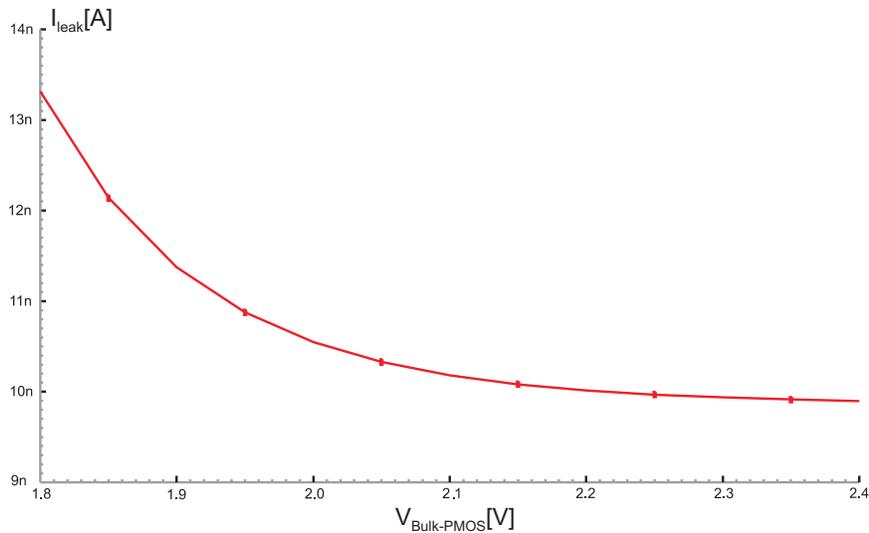


**Abbildung 2.13:** Thresholdspannung in Abhängigkeit vom Bulkpotential

Erneut wird aufgrund der einfacheren Kontaktierbarkeit der N-Wannen von PMOS-Transistoren nur der Einfluss des Bulkpotentials der PMOS-Transistoren auf den Leakagestrom untersucht. Zur Verifizierung des Einsparpotentials wird

## 2.8 Reduzierung des Leakagestroms durch Bulkpotentialerhöhung 73

die Schaltung aus 2.7.2 erneut untersucht, nun aber mit erhöhtem Bulkpotential für die PMOS-Transistoren. Abbildung 2.14 verdeutlicht den Einfluss des Bulkpotentials auf den Leakagestrom.

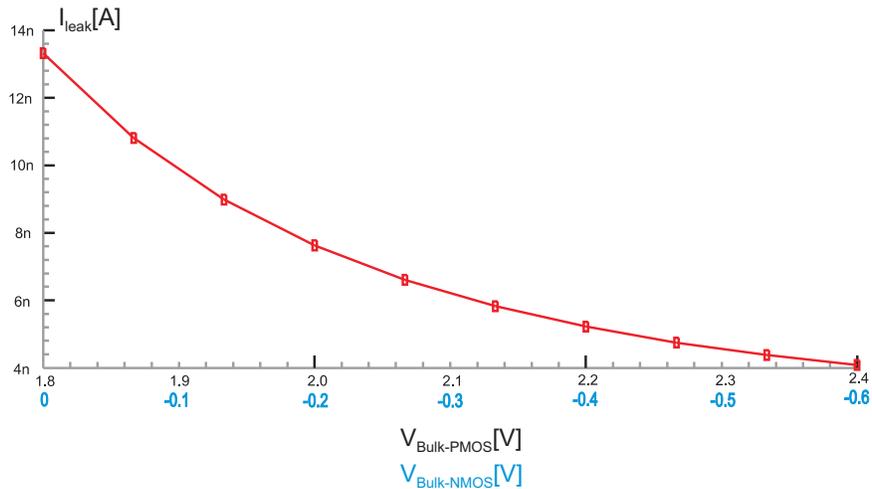


**Abbildung 2.14:** Leakagestrom einer Gatterkette in Abhängigkeit vom N-Wannenpotential

Das kleinere Einsparpotential gegenüber der Sleep-Transistor-Lösung ist zu bemerken (vergl. Abbildung 2.10 bzw. 2.11). Durch die Beschaltung der N-Wannen der Gatterkette sind im stationären Zustand ungefähr die Hälfte aller PMOS-Transistoren im leitenden Zustand. Auf sie hat die Bulkpotentialanhebung keinen relevanten Leakage-reduzierenden Einfluss. Hinzu kommt der Stackeffekt bei der Sleep-Transistor-Lösung, der den Leakagestrom weiter verkleinert.

## 2.8 Reduzierung des Leakagestroms durch Bulkpotentialerhöhung 74

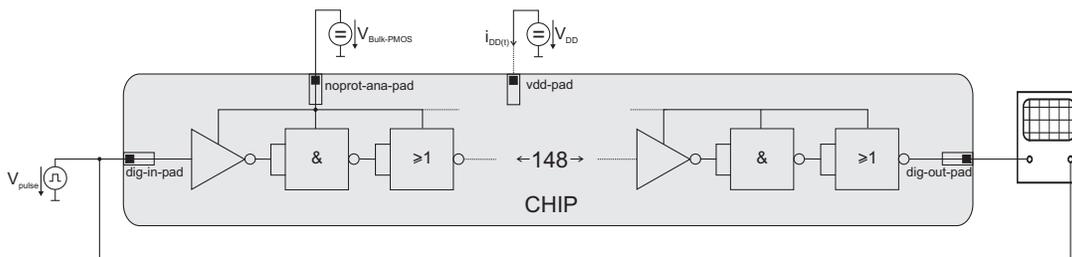
Wird das Bulkpotential der NMOS-Transistoren in einem zweiten Schritt im SleepMode unter das GND-Potential abgesenkt (Abb. 2.15), dann ergeben sich für die Leakageerduzierung bessere Werte. Technisch ist diese Vorgehensweise aber problematisch, da die Bulks der NMOS-Transistoren mit dem Substrat übereinstimmen, welches normalerweise auf GND-Potential liegt. Das GND-Potential wird in der Regel aber über Platinen hinweg verteilt, so dass eine Auftrennung des Substrats hier schwierig zu realisieren ist. Isolierte NMOS-Transistoren können in Tripple-Well-Prozessen realisiert werden. Wenn, wie in diesem Fall, eine Vielzahl von Transistoren auf demselben Bulkpotential liegen, ist der Flächenzuwachs, anders als in Abbildung 2.7, nicht besonders hoch.



**Abbildung 2.15:** Leakagestrom einer Gatterkette in Abhängigkeit vom Bulkpotential der NMOS- und PMOS-Transistoren

### 2.8.1 Messungen an einer Teststruktur

Zur Verifizierung der Wirksamkeit der Bulkpotentialanhebung der PMOS-Transistoren auf die Reduktion des Leakagesstroms wird eine Messung einer Teststruktur vorgenommen. Diese besteht aus 150 hintereinandergeschalteten Standard-CMOS-Gattern in  $0.18\mu\text{m}$  Tripple-Well Technologie im Standardzellenentwurf. Alle PMOS-Transistoren liegen in der gleichen N-Wanne, deren Potential über einen Pad frei wählbar ist. Die unrealistisch große Anzahl von hintereinandergeschalteten Gattern wurde aus Gründen der besseren Messbarkeit gewählt. Die Ergebnisse sind ohne Einschränkung auch auf kleinere Pfadtiefen übertragbar. Abbildung 2.16 verdeutlicht den Messaufbau.

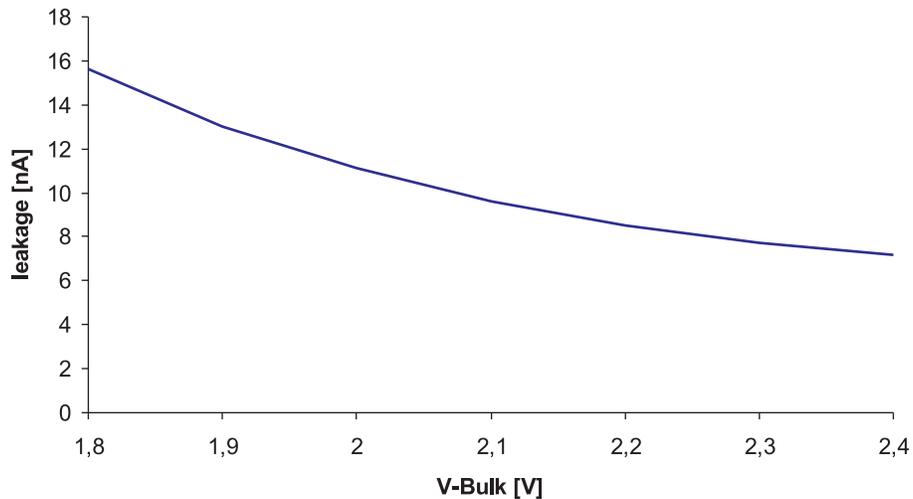


**Abbildung 2.16:** Messaufbau zu Messungen mit variablem N-Wannenpotential

Die Messung des Delays ist nicht trivial, da Pulsgenerator und Chippads jeweils verschiedene Treiberfähigkeiten besitzen. Die Messergebnisse sind hierbei also nur Anhaltspunkte, die durch Vergleich mit der Simulation besser bewertet werden können.

### 2.8.2 Leakagemessungen

Zur Leakagestrommessung wird das Bulkpotential schrittweise erhöht und die Stromaufnahme im stationären Zustand (Eingang ändert sich nicht) gemessen.

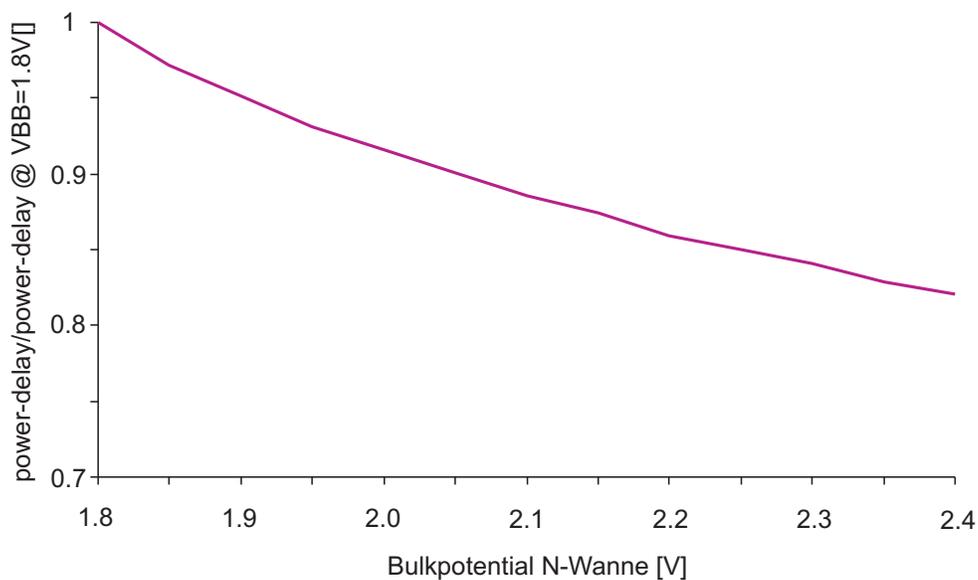


**Abbildung 2.17:** Leakagestrom einer Teststruktur abhängig vom Bulkpotential

Der positive Einfluss der Bulkpotentialerhöhung ist hierbei klar zu erkennen. Die Stromaufnahme der Testschaltung sinkt im idle-mode um ca. 54 %. Anzumerken ist, dass das Aufladen der N-Wanne auf ein höheres Potential selbst eine Ladeleistung darstellt. Um ein günstiges Verhältnis von Einsparpotential durch die Bulkpotentialerhöhung und der dazu notwendigen Ladeleistung zu erhalten, sollte eine Bulkpotentialerhöhung nur für längere idle-Phasen durchgeführt werden.

### 2.8.3 Performancemessungen

Nun wird das Delay der Gatterkette in Abhängigkeit des Bulkpotentials gemessen. Aufgrund der Thesholdvoltageverschiebung ausschließlich der PMOS-Transistoren werden die Ladevorgänge der Ausgangsknoten mit steigendem Bulkpotential verzögert. Dies führt zu einer Unsymmetrie der Schaltvorgänge, die sich über die große Gatteranzahl aber ausmittelt. Interessant ist hierbei das Power-Delay-Produkt, welches ein Maß für die Effizienz der Schaltung darstellt. Die Messung, die in Abbildung 2.18 dargestellt ist, verdeutlicht die Reduktion des Power-Delay-Produkts mit steigendem N-Wannen Bulkpotential. Als Stimuli ist hierbei ein Rechtecksignal der Amplitude von  $1.8V$  und der Frequenz  $f_s = 1MHz$  gegeben. Die Werte sind für ein Standardbulkpotential von  $V_{BB} = 1.8V$  bei der Standardversorgungsspannung von  $V_{DD} = 1.8V$  normiert.



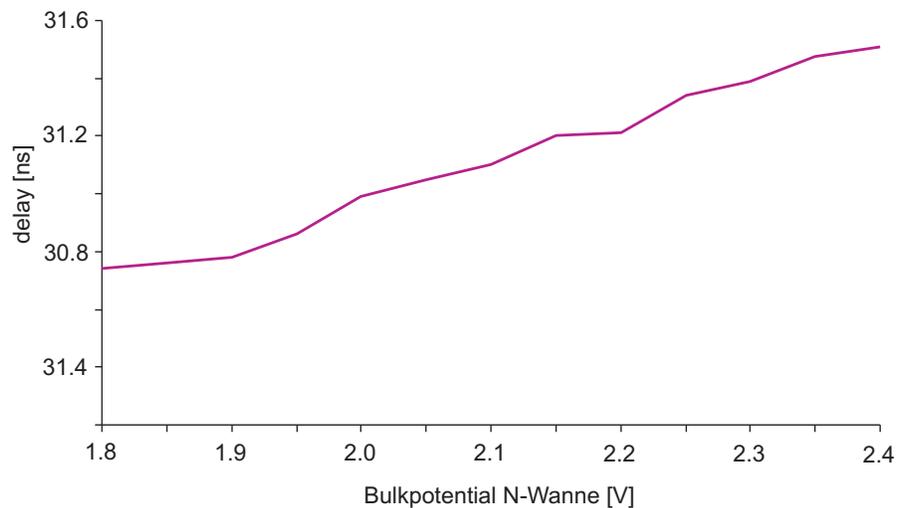
**Abbildung 2.18:** Messung des Power-Delay-Produkts einer Teststruktur abhängig vom Bulkpotential

Aus Abbildung 2.18 ist abzulesen, dass eine Steuerung der Bulkpotentialspannung der PMOS-Transistoren dazu genutzt werden kann, die Performance der Schaltung an die Spezifikationen anzugleichen und somit Leistung einzusparen. Andererseits wird durch die Erhöhung des Bulkpotentials in Idle-Zeiten der statische Leistungsverbrauch minimiert (Abbildung 2.17).

Der Messaufbau ermöglicht eine direkte Messung des statischen Stroms in die Gatterkette nicht, da über die Versorgungspads und die Padkranzversorgung ein nicht beeinflussbarer zusätzlicher statischer Strom fließt. Die Änderung des gemessenen Stroms rührt allerdings alleine von der Erhöhung des N-Wannen-

## 2.8 Reduzierung des Leakagestroms durch Bulkpotentialerhöhung 78

Potentials. Die Performance (also die Verzögerungszeit der Gatterkette) wird durch das Bulkpotential, wie in Abbildung 2.19 gezeigt, beeinflusst.



**Abbildung 2.19:** Messung der Verzögerungszeit der Teststruktur abhängig vom Bulkpotential

Die Messung der Thresholdspannung ist für die Teststruktur nicht vorgesehen. Anhaltspunkte für die Thresholdspannung bei 27°C liefert Abbildung 2.13.

## 2.9 Bewertung der Technologien zur Beeinflussung der Thresholdspannung

Wegen ihrer guten Eignung zur Reduktion der Gesamtverlustleistung ist die Verwendung von Transistoren ausschließlich niedriger Thresholdspannung (HighSpeed-Technologie) bei reduzierter Versorgungsspannung anzuraten (vergl. Kapitel 2.2). Durch die dynamische Erhöhung der Thresholdspannung lassen sich die Vorteile der LowLeakage-Technologie (niedriger Leakagestrom im StandBy-Modus) und der HighSpeed-Technologie (niedrige dynamische Verlustleistung bei reduzierter Versorgungsspannung) miteinander kombinieren. Die sogenannte *(Active)-Body-Biasing* bzw. *Dynamic-Threshold-CMOS-Technologie* sind für Standard-CMOS-Technologien wenig geeignet, da sie eine stark vergrößerte Layoutfläche beanspruchen. Durch eventuell zusätzliche Transistoren zur aktiven Ansteuerung der Bulks bzw. durch das Laden und Entladen der Bulks, die in Standard-CMOS-Technologie selbst einen hohen Kapazitätsbelag aufweisen, wird die Gesamtlastkapazität erhöht, was wiederum zu ansteigender dynamischer Verlustleistung führt.

Die Sleep-Transistor-Technologie zeigt für die Reduzierung des Leakagestroms die besten Ergebnisse. Nachteilig hierbei ist, dass im SleepMode keine Verbindung zu einer oder beiden Versorgungsrails besteht, so dass der momentane Zustand der Schaltung nicht getrieben wird. Je nach Laststrom, und damit nach Komplexität und Spezifikationen, entsteht durch den Sleep-Transistor ein nicht unerheblicher zusätzlicher Platzbedarf.

Die Erhöhung des Bulkpotentials der PMOS-Transistoren reduziert den Leakagestrom systembedingt nur ca. halb so viel, wie bei der Sleep-Transistor-Technologie, die Verbindung zur Versorgungsspannung bleibt aber während des SleepModes bestehen. Werden auch die NMOS-Transistoren in dieses Konzept miteinbezogen, erreicht die Reduzierung des Leakagestroms bessere Werte als die Sleep-Transistor-Technologie. Nachteilig erweist sich dabei das für NMOS-Transistoren in N-Well Prozessen das Bulkpotential nur global, also für den ganzen Chip beeinflusst werden kann, ohne zusätzliche isolierte P-Wannen zu integrieren, deren Platzbedarf hoch ist. Die Verwendung von einer, oder zwei zusätzlichen Versorgungsspannungen macht es notwendig, einen oder zwei zusätzliche Pads bereitzustellen. Die Generierung der zusätzlichen Spannungen ist grundsätzlich auch OnChip möglich, jedoch ist der Aufwand hierzu beträchtlich. Im PCB-Verbund sind mehrere Spannungsebenen ohnehin meistens üblich, so dass diese Spannungen auch extern angelegt werden können. Der Vorteil der Bulkpotentialerhöhung liegt in der permanenten getriebenen Verbindung aller Knoten zu den Versorgungsspannungen, so dass im idle Mode der Schaltungszustand ohne Refreshment erhalten bleibt.

# Kapitel 3

## Reduzierung der dynamischen Leistungsaufnahme

### 3.1 Systematische Methoden

Neben schaltungstechnischen und technologischen Methoden zur Reduzierung der Verlustleistung bergen systematische Methoden ein wichtiges Einsparpotential. Es gilt für die dynamische Verlustleistung:

$$P_{dym} \sim \alpha \cdot f \cdot C_L \cdot V_{DD}^2 \quad (3.1)$$

Anhand der quadratischen Abhängigkeit von  $V_{DD}$  wird die große Bedeutung der Reduzierung der Versorgungsspannung im Hinblick auf Minimierung der Verlustleistung deutlich. Die Faktoren  $\alpha$  und  $f$  stellen Größen dar, die der Systementwerfer direkt beeinflussen kann. Die effektive Größe  $C_L$  ist u.a. eine Folge der Systemkomplexität und somit vom Systementwurf beeinflussbar. Schlussendlich bestimmt  $V_{DD}$  die Schaltgeschwindigkeit der Gatter mit, also auch die minimale Taktfrequenz, die u.a. von einem optimierten Systementwurf reduziert werden kann. Der Systementwurf kann dementsprechend die Verlustleistung reduzieren. Systematische Methoden hierzu finden sich reichlich in der Literatur. Im Folgenden sollen nur einige Beispiele dazu genannt werden.

#### 3.1.1 Reduzierung der Schaltaktivität

Die Reduzierung der Schaltaktivität geht oft auch mit einer Reduzierung der Komplexität der Schaltung, also auch mit niedrigerer Lastkapazität  $C_L$  und kleinerem Platzbedarf einher. Bei booleschen Operationen wird dies z.B. durch Logikminimierung erreicht. Systematisch wird die Schaltaktivität durch optimierte Algorithmen erreicht. Ein Beispiel für ein optimiertes System stellt die Anwendung eines optimierten Gray-Codes auf Adressierungsbusse von Mikroprozessoren dar [41]. Busse mit ihren langen verzweigten Wegen sind hochkapazitive Strukturen,

bei denen eine Reduzierung der Schalthäufigkeit den Leistungsumsatz besonders vermindert. Der normale Programmablauf produziert häufig aufeinanderfolgende Adressen. Bei normalen binären Programmzählern werden hierbei mehrere Busleitungen umgeschaltet, was eine hohe Verlustleistung durch Ladungsänderung der Kapazitäten der Busleitungen zur Folge hat. Wird z.B. dezimal von Programmzeile 3 nach Programmzeile 4 umgeschaltet, bedeutet das binär ein Umschalten von 3 Busleitungen (...011  $\rightarrow$  ...100). Ein Ansatz, diese Umschaltvorgänge zu Minimieren besteht darin, die Adressen aufeinanderfolgender Programmzeilen mit einem Gray-Code zu belegen. Ein Gray-Code ist ein binärer Zählablauf, bei dem sich von Zeile zu Zeile immer nur eine Ziffer ändert. Die Tabelle 3.1.1 verdeutlicht dies für einen 4-Bit Binärzähler.

Binär-Code	Gray-Code
0000	0000
0001	0001
0010	0011
0011	0010
0100	0110
0101	0111
0110	0101
0111	0100
1000	1100
1001	1101
1010	1111
1011	1110
1100	1010
1101	1011
1110	1001
1111	1000
30 Umladungen	16 Umladungen
68 Gatter	72 Gatter
Fläche 100%	Fläche 114.5%

**Tabelle 3.1:** Vergleich der Umschaltvorgänge Binär-Code - Graycode

Der Reduzierung der Schaltvorgänge um annähernd 50 % steht eine Vergrößerung der benötigten Fläche von 14,5 % gegenüber. Unter der Annahme eines gleichen Kapazitätsbelages für beide Zähler ergibt sich ein Verminderung von  $\approx 57\%$  der dynamischen Verlustleistung im Programmzähler bei Verwendung eines Gray-Code Programmzählers. Die Komplexität eines Gray-Code Programmzählers steigt exponentiell mit der Wortbreite der Adresse an. Wie in [41] gezeigt

wird, ist es möglich durch Verwendung sog. segmentierter Gray-Code Zähler diesen Nachteil gegenüber Binärzählern, deren Komplexität linear mit der Wortbreite skaliert, zu vermeiden.

Einsparungen in der Schalthäufigkeit durch intelligente Algorithmen lassen sich auch in anderen Schaltungsteilen erreichen. Erwähnt werden sollen hier nur verschiedene Addiererstrukturen.

### 3.1.2 Reduzierung der Lastkapazität

Die Reduzierung der Lastkapazität durch systematische Maßnahmen geschieht in erster Linie durch Verringerung der Komplexität der Schaltungen (vergl. hierzu auch Kapitel 3.1.1). Neben der Anwendung optimierter Algorithmen kann alternative Schaltungstechnik die Anzahl der Transistoren für eine bestimmte Funktionalität verringern. Somit wird die wirksame Lastkapazität verringert. Durch Verkleinerung der Strukturgrößen bei einem Technologiesprung wird i.d.R. auch eine Reduzierung der Lastkapazitäten erreicht.

## 3.2 Dynamische Einstellung der Versorgungsspannung

Unter dynamischer Einstellung der Versorgungsspannung sind alle Konzepte zu verstehen, die  $V_{DD}$  parameterabhängig verändern. Als Parameter kommen hierfür in Frage:

- Ansprüche an die Verzögerungszeit  $\rightarrow$  Performance
- Umgebungsvariablen wie z.B. die Temperatur
- Technologieparameter

Im Rahmen dieser Dissertation wird ein System vorgestellt, das eine parameterabhängige Referenzspannung zur Verfügung stellt.

## 3.3 Statistische Verteilung von Technologieparametern

Durch statistische Schwankungen der Prozessparameter ( $\beta_0$  und  $V_{T0}$ ), hervorgerufen durch Variationen der Schichtdicken (z.B. der Oxidschichtdicke) und Unterschiede in der Stärke des Dotierungsgrades, gibt es Abweichungen der Technologieparameter zwischen Chips einer Charge eines Herstellungsprozesses. Um ein einwandfreies Funktionieren der Chips zu gewährleisten, ohne übermäßigen Produktionsausschuss in Kauf nehmen zu müssen, werden die digitalen Schaltungen entsprechend so designed, dass die Funktion auch bei ungünstig variierenden Technologieparametern innerhalb der Spezifikation gesichert ist. Dies birgt den Nachteil in sich, dass die Schaltungen auf Chips mit günstig variierenden Prozessparametern die Spezifikationen übertreffen. Diese Chips könnten mit einer niedrigeren Versorgungsspannung betrieben werden und somit Leistung einsparen. Um dieses Einsparpotential auszunutzen, müssen die Technologieparameter für jeden Chip erfasst werden. Dies ist für einzelne Größen, wie der Thresholdspannung  $V_{T0,n,p}$  oder die Steilheit der Transistoren  $\beta_{0,n,p}$  nur schwierig möglich (vergl.: [26], [30] und [48]).

Die meisten statistischen Modelle der monolithischen Integration enthalten eine Abhängigkeit von den geometrischen Abmessungen der betrachteten Bauteile. Im einfachsten Fall sind dies die Kantenabmessungen  $W$  und  $L$ , sowie die Distanz  $D$  zwischen zwei zu vergleichenden Bauelementen.

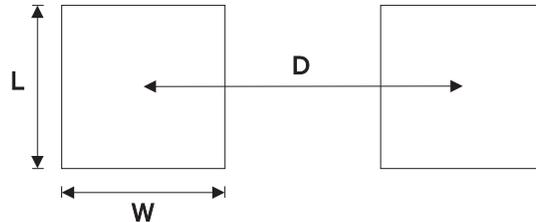


Abbildung 3.1: Elementare Bauelementengeometrie  $L$ ,  $W$  und  $D$

### 3.3.1 Flächen- und Abstandsgesetz erster Ordnung

Das von Lakshmikummar 1986 erstmalig veröffentlichte Statistikmodell für monolithisch integrierte MOS-Transistoren wird als Flächengesetz erster Ordnung bezeichnet [26]. Es beschreibt die relative statistische Paarigkeit von Parametern zweier direkt benachbarter MOS-Transistoren. Das Flächengesetz besagt, dass sich die relative Standardabweichung eines Transistorparameters  $P$  (z.B. Thresholdspannung  $V_T$ ) umgekehrt proportional zur effektiven Fläche des Transistors verhält.

$$\frac{\sigma_{(P)}^2}{P^2} = \frac{A_P^2}{W_{eff} \cdot L_{eff}} \quad (3.2)$$

Der Proportionalitätsfaktor  $A_P$  wird als Flächen-Matchingparameter bezeichnet und ist eine technologieabhängige Größe. Pelgrom ergänzte 1989 dieses Statistikmodell [30], indem er die räumliche Distanz  $D$  zwischen zwei betrachteten Bauteilen berücksichtigte. Die relative Standardabweichung der einzelnen Transistorparameter setzt sich dabei aus zwei Anteilen zusammen. Der erste Anteil besteht aus dem oben genannten Flächengesetz, während der zweite Anteil proportional mit der Distanz  $D$  anwächst (Distanzgesetz). Aufgrund der statistischen Unabhängigkeit beider Fehleranteile addieren sich die Anteile geometrisch.

$$\frac{\sigma_{(P)}^2}{P^2} = \frac{A_P^2}{W_{eff} \cdot L_{eff}} + S_{(P)}^2 \cdot D^2 \quad (3.3)$$

Der Proportionalitätsfaktor  $S_{(P)}$  wird als Distanz-Matchingparameter bezeichnet. Er ist, wie der Flächen-Matchingparameter  $A_{(P)}$ , eine technologieabhängige Größe.

### 3.4 Pelgrom Modell

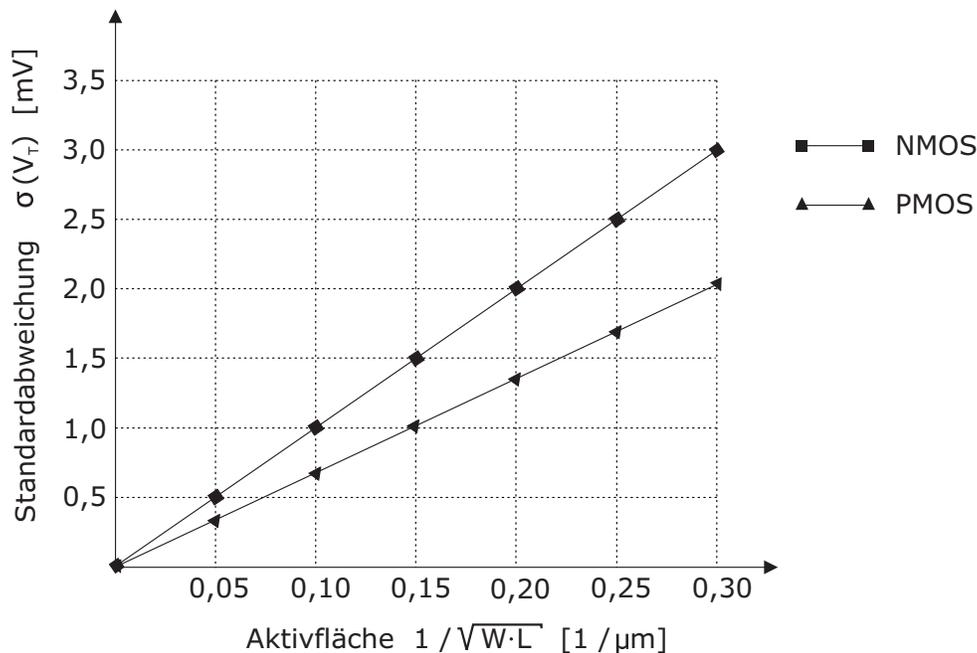
Pelgrom beschreibt in [30] statistische Betrachtungen und Messungen, um eine Aussage über das Matching in verschiedenen CMOS-Technologien zu erhalten. Wird eine Anzahl von nominal identischen Transistoren betrachtet, zeigen die Transistorparameter ein Matchingverhalten nach dem oben schon erwähnten Flächen-Abstands-Gesetz. Für die Transistor-beschreibenden Parameter Threshholdspannung  $V_T$  und Steilheit  $\beta$  erhält man folgende Abhängigkeiten

$$\sigma^2(V_T) = \frac{A_{V_T}^2}{W \cdot L} + S_{V_T}^2 \cdot D^2 \quad (3.4)$$

$$\frac{\sigma^2(\beta)}{\beta^2} = \frac{A_\beta^2}{W \cdot L} + S_\beta^2 \cdot D^2 \quad (3.5)$$

, wobei  $A_\beta$ ,  $A_{V_T}$ ,  $S_\beta$  und  $S_{V_T}$  technologieabhängige Konstanten (Prozessparameter) sind.

Der zweite Term dieser Gleichungen wird üblicherweise vernachlässigt, ausgenommen für sehr große Distanzen zwischen den einzelnen Schaltungselementen.



**Abbildung 3.2:** Standardabweichungen der Schwellwertspannung  $\sigma(V_T)$

Die Abbildungen (3.2) und (3.3) zeigen die Standardabweichungen der Threshholdspannung  $\sigma(V_T)$  und der Steilheit  $\sigma(\beta)/\beta$  in Abhängigkeit von der Quadratwurzel der Aktivfläche  $1/\sqrt{W \cdot L}$  für NMOS- und PMOS Transistoren der  $0.18 \mu\text{m}$  CMOS-Technologie von S.T. Microelectronics.

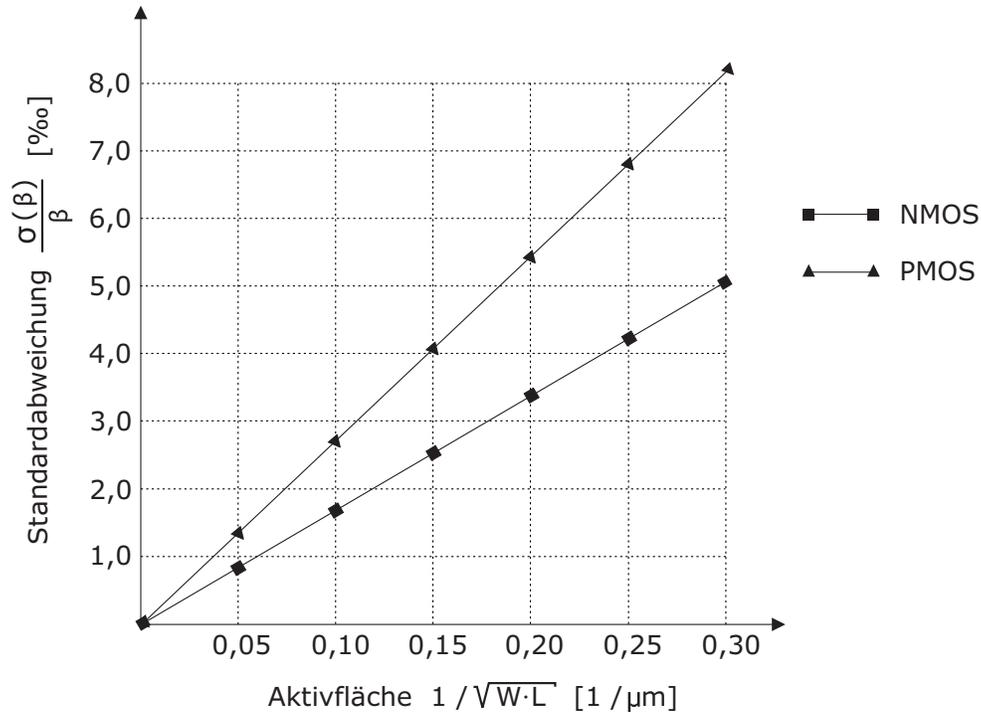
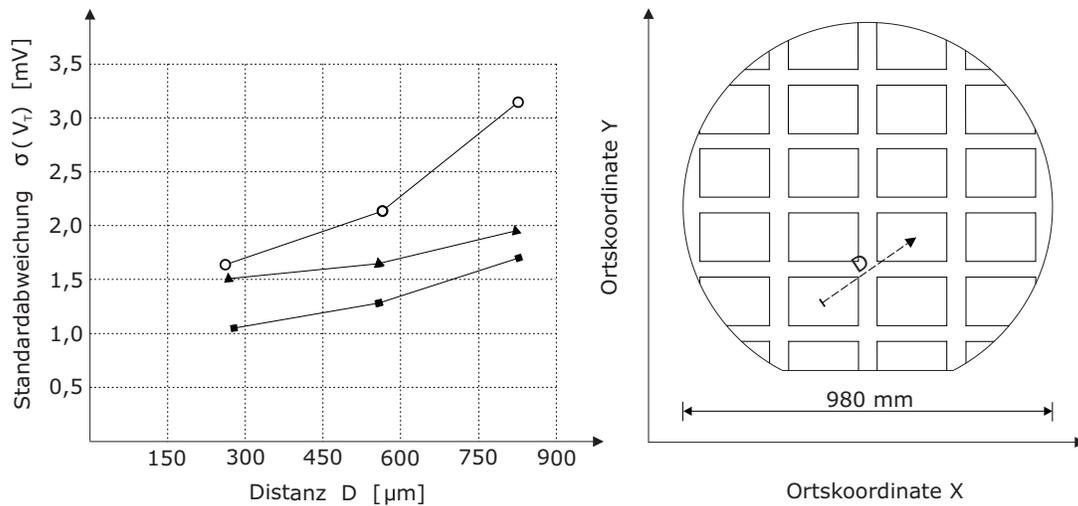


Abbildung 3.3: Standardabweichungen der Steilheit  $\sigma(\beta)/\beta$

Eine Übersicht der technologieabhängigen Prozessparameter der  $0.18 \mu\text{m}$  Standard CMOS-Technologie von S.T. Microelectronics liefert Tabelle (3.2). Die Distanz-Matching Parameter  $S_{V_T}$  und  $S_\beta$  werden für diese Technologie zu Null angegeben. Dies entspricht natürlich nicht der Realität. Untersuchungen früherer Technologien von S.T. Microelectronics haben aber gezeigt, dass weder für die Thresholdspannung  $V_T$ , noch für die Steilheit  $\beta$  signifikante Matchingeffekte über einen Distanzbereich von  $1 \mu\text{m}$  bis  $1400 \mu\text{m}$  festzustellen waren. Da für die  $0.18 \mu\text{m}$  Technologie von S.T. Microelectronics von einem noch besseren Matchingverhalten auszugehen ist, wurde hier auf die Untersuchung der Distanz-Matching Parameter verzichtet. Für Technologien anderer Firmen, z.B. X-FAB Wafermap, existieren Untersuchungen über Long Distance Matching zwischen größeren Wafergebieten [28] (Abbildung 3.4). Die gemessenen Standardabweichungen weisen keinen linearen Zusammenhang mit der Distanz auf. Die Parameterschwankungen über große Entfernungen scheinen komplexer und nicht mehr durch das einfache Flächen-Distanz-Gesetz erster Ordnung approximierbar.

	Gültigkeitsbereich	N-MOS	P-MOS	Einheit
$A_{VT}$	$0.325 \mu m^2 < W \cdot L$	10	6.8	$mV \cdot \mu m$
$A_\beta$	$1.3 \mu m < W < 50.9 \mu m$ $0.25 \mu m < L < 12.5 \mu m$	1.69	2.72	$\% \cdot \mu m$
$S_{VT}$	$0.325 \mu m^2 < W \cdot L$	0	0	$\mu V \cdot \mu m$
$S_\beta$	$1.3 \mu m < W < 50.9 \mu m$ $0.25 \mu m < L < 12.5 \mu m$	0	0	$10^{-6} \cdot \mu m$

**Tabelle 3.2:** Prozessparameter der 0.18  $\mu m$  Technologie S.T. Microelectronics



**Abbildung 3.4:** Long Distance Mismatch  $\sigma(V_T)$  verschiedener Technologien [XFAB]

## 3.5 Verteilungsfunktion der Parameter

Die aus einer Vielzahl von statistischen unabhängigen Einzelereignissen resultierende Variation einer Größe entspricht i.A. einer Normalverteilung. Diese wird durch Angabe eines Mittelwertes  $\mu$  und einer Standardabweichung  $\sigma$  charakterisiert. Die Abweichungen zwischen zwei identisch entworfenen Transistoren können in guter Näherung als normalverteilt angenommen werden. Der Mittelwert  $\mu$  der Normalverteilung wird vom sogenannten systematischen Mismatching beeinflusst. Ursachen für systematisches Mismatching sind:

- Photomaskentoleranzen und Linsenabweichungen
- variierende Abscheidungen bei CVD-Prozessen<sup>1</sup> oder Metallisierung (als Folge von Temperaturschwankungen und mechanischem Stress)
- variierende Ätzraten (als Folge von Temperaturschwankungen und mechanischem Stress)

Die Streuung  $\sigma$ , die im Wesentlichen von der Transistorweite  $W$ , der Transistorlänge  $L$  sowie dem Abstand  $D$  zwischen den betrachteten Bauelementen abhängig ist, wird durch das sogenannte stochastische Mismatching beeinflusst. Ursachen für stochastisches Mismatching sind:

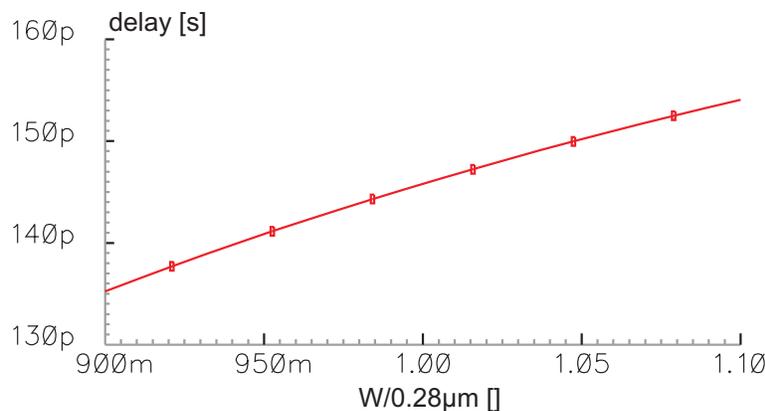
- lokale Schwankungen in der Dotierungskonzentration (inhomogene Ionenimplantation und Dotierungs-Diffusion)
- Ladungshäufung  $\Rightarrow$  Ladungsmobilität wird lokal gestört (Dotierungs Clusterbildung)
- inhomogene Korngrenzenverteilung (Oxidkörnung)
- zufällige geometrische Schwankungen des Bauteils (Oxiddicke, Polysiliziumdicke, effektive Kanalbreite)
- Ätzrauigkeit (bei nicht exakt sauberen Oberflächen)

---

<sup>1</sup>Chemical-Vapour-Deposition = Chemische Gasphasenabscheidung

### 3.6 Relevante Technologieparameter

Technologieparameter sind physikalische Eigenschaften des Chips, die während des Integrationsprozesses festgelegt werden. Dazu gehören die Dicke des Feld- und des Gateoxids. Ebenso sind Dotierungsdichten (N- wie P-Dotierung) relevante Parameter. Aus diesen physikalischen Parametern lassen sich Parameter des CMOS-Transistormodells ableiten, wie die Thresholdspannung  $V_{T0}$  und die Steilheit des Transistors  $\beta_0$  bzw. die abgeleitete Transkonduktanz  $gm$ . Diese Transistorparameter sind Proportionalitätsfaktoren zum Lade- und Entladestrom über die Drain-Source-Strecke und somit Performance-relevant. Die Erfassung von Technologieparametern ist aufwändig. Auch die abgeleiteten Größen  $V_{T0}$  und  $\beta_0$  sind schwierig zu messen. Zusätzlich wird die Performance nicht unwesentlich vom geometrischen Mismatch beeinflusst. Durch Ungenauigkeiten in der Maskenherstellung und Unterdiffusion bei der Prozessierung entstehen Abweichungen der Transistordimensionen von den eigentlich erwünschten Abmessungen (siehe Kapitel 3.6.1).

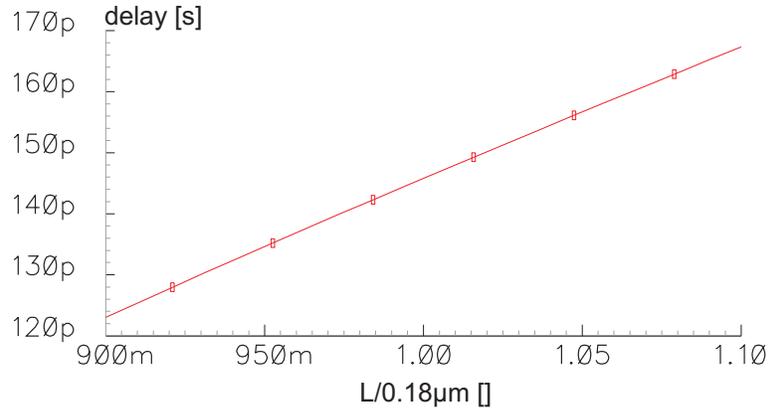


**Abbildung 3.5:** Delay eines digitalen Treibers abhängig von der Weitenvarianz

Hierbei spielt eine Abweichung der Transistorweite eine weniger ausgeprägte Rolle, da Vergrößerung der Gatekapazität und ansteigender Drain-Source-Strom zueinander umgekehrt proportional sind. Abbildung 3.5 zeigt das simulierte delay einer minimal dimensionierten digitalen Treiberkette aus fünf seriellen Invertern, für Abweichungen der Transistorweiten innerhalb  $\pm 10\%$ . Die Abweichung vom Standardwert des Delays (bei  $W = 0.28\mu\text{m}$ ) beträgt hier  $\pm 5.88\%$ . Interessant bei Abbildung 3.5 ist die umgekehrte Proportionalität von Transistorweite und Verzögerungszeit. Es wäre eigentlich zu erwarten, dass mit steigendem  $\beta$  und damit auch steigendem  $I_D$  das delay sinkt. Durch die Vergrößerung der effektiven Transistorweite steigt aber auch  $C_{GS}$ . Die fall- bzw. risetime eines Inverters ist aber nach Gleichung 2.3 proportional zu  $2C_L/\beta$ , die Erhöhung der

Lastkapazität über dem größeren Ladestrom  $I_D$  dominiert.

Eine Abweichung der Transistorlänge beeinflusst die Schaltungsperformance in stärkerem Maße, was Abbildung 3.6 veranschaulicht. Hierbei ist die Abweichung vom Standardwert des Delays (bei  $W = 0.18\mu m$ ) nun  $\pm 12.22\%$ .



**Abbildung 3.6:** Delay eines digitalen Treibers abhängig von der Längenvarianz

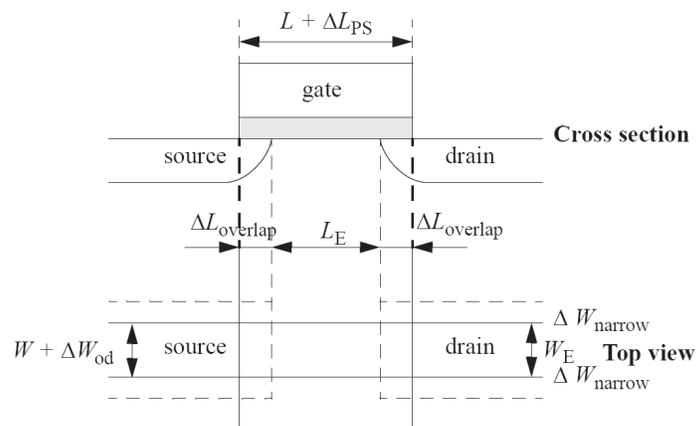
Für eine Evaluierung der Chip-spezifischen Performance ist die genaue Determinierung einzelner Performance-beeinflussender Größen nicht notwendig. Es genügt, die Performance direkt zu messen. Für digitale Schaltungen ist die maximale Clock-Frequenz ein Indiz für die Performance. Diese wird durch die Laufzeit des längsten Pfads bestimmt (critical path delay).

### 3.6.1 Geometrisches Mismatch

Mit geometrischem Mismatch ist die Abweichung der realen Transistordimensionierung von der geplanten Dimensionierung gemeint. Die Ursachen für diese Abweichungen liegen im Herstellungsprozess. Kanallänge und -weite werden durch die Masken für das Gate (Polysilizium) und die Drain und Source-Gebiete (n+, bzw. p+) determiniert. Diese Masken unterliegen Abweichungen in ihrer absoluten Größe. Ebenso kommt es zur Reduktion der effektiven Kanallänge durch laterale Unterdiffusion von source bzw. drain doping Ionen, ebenso wie zur Reduktion der Kanalweite durch laterale Diffusion der den Kanal begrenzenden Doping Ionen. Abbildung 3.7 [37] verdeutlicht den Unterschied zwischen geplanten und realen Kanaldimensionen. Für die sich ergebende effektive Kanallänge und Weite gilt:

$$L_E = L - \Delta L = L + \Delta L_{PS} - 2\Delta L_{overlap} \quad (3.6)$$

$$W_E = W - \Delta W = W + \Delta W_{OD} - 2\Delta W_{narrow} \quad (3.7)$$

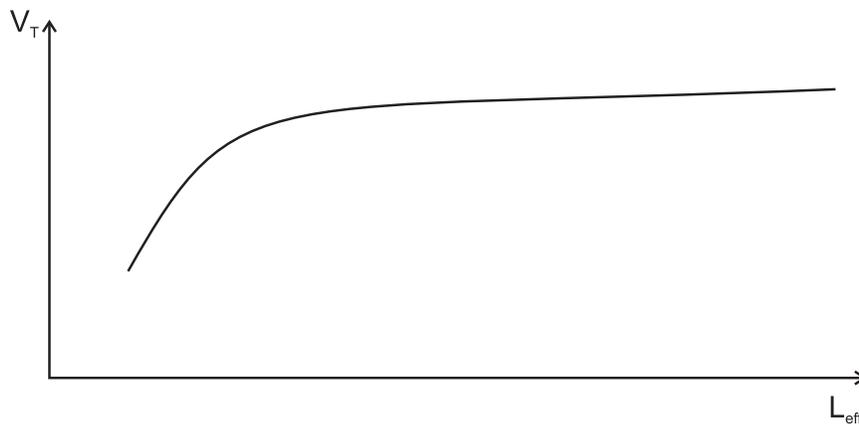


**Abbildung 3.7:** Varianzen der Kanaldimensionen eines MOS-Transistors

Hierbei sind:

- $L_E$ ,  $W_E$ : effektive Kanallänge bzw. -weite
- $\Delta L_{PS}$ ,  $\Delta W_{OD}$ : Differenz zwischen gelayouteter und realer Gatelänge bzw. -weite
- $\Delta L_{overlap}$ : effektive Kanallängenverkürzung durch Unterdiffusion der Drain- und Source Gebiete
- $\Delta W_{narrow}$ : effektive Kanalweitenreduktion durch Unterdiffusion der den Kanal begrenzenden doping Ionen

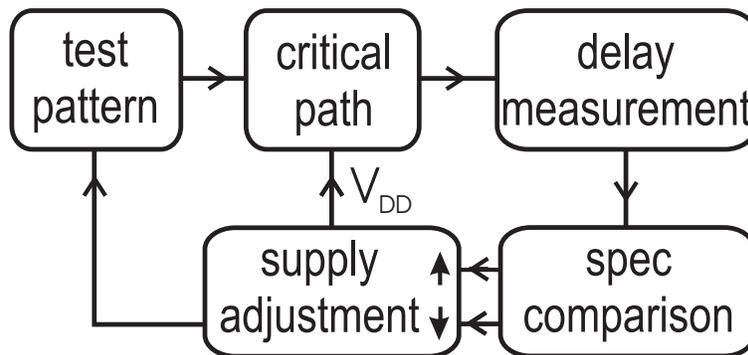
Die die effektive Kanaldimension beeinflussenden Faktoren sind unabhängig voneinander und über den Chip zufällig verteilt. Selbst zueinander benachbarte Transistoren können in unterschiedliche Richtungen variieren [47]. Diese zufällige Streuung der effektiven Kanaldimension ist für minimal dimensionierte Transistoren besonders folgenschwer, da die Thresholdspannung für kurze Kanallängen von dieser stark abhängig ist (short-channel-effects [32] [15] [40]) und kleiner wird (Abbildung 3.8. Da Kanallänge und Thresholdspannung den Drain-Source-Strom entscheidend beeinflussen, entsteht für minimale Transistoren eine nicht vorhersagbare statistische Varianz in der Lade- bzw. Entladezeit. Digitale Schaltungen bestehen i.d.R. aus minimal dimensionierten Transistoren. Das delay eines Pfads unterliegt somit einer Unsicherheit basierend auf dem geometrischen Mismatch der daran beteiligten Transistoren.



**Abbildung 3.8:** Thresholdspannungsbeeinflussung durch Kurzkanaleffekte

### 3.7 Critical-Path-Delay als Performanceindikator

Ein verbreiteter Weg zur Messung der Performance einer digitalen Schaltung ist die Betrachtung des Delays des längsten digitalen Pfads. Somit wird die maximale Berechnungszeit eines digitalen Wertes bestimmt. Abbildung 3.9 zeigt ein Konzept zur Einstellung der Versorgungsspannung aufgrund des Critical-Path-Delay.



**Abbildung 3.9:** Konzept zur Versorgungsspannungseinstellung aufgrund des Critical-Path-Delay

Der kritische Pfad wird zunächst bei maximaler Versorgungsspannung mit einem Teststimulus beaufschlagt. Hierzu kann der kritische Pfad über Multiplexer in die Teststruktur implementiert werden (was während der Dauer der Versorgungsspannungseinstellung bedeutet, dass die Schaltung blockiert ist). Alternativ kann der kritische Pfad kopiert und möglichst raumnah zum Originalpfad platziert werden, oder aber kann eine Teststruktur (z.B. eine Inverterkette) benutzt werden, die in ihrem Delay dem kritischen Pfad entspricht.

Die Verarbeitungsgeschwindigkeit wird gemessen und mit der Spezifikation verglichen. In der Regel ist das Delay des kritischen Pfads kürzer, als es die Spezifikation verlangt. Der Designer trägt schließlich die Verantwortung dafür, dass die Spezifikation unter allen Umständen mindestens erfüllt wird. Deswegen wird im ersten Durchlauf die Versorgungsspannung um einen diskreten Wert erniedrigt und die Messung erneut initiiert. Dieser Vorgang wird solange wiederholt, bis der Vergleich critical-path-delay vs. Spezifikation den Befehl zur Erhöhung der Versorgungsspannung erteilt. Somit ist die niedrigste Versorgungsspannung evaluiert, welche zur Einhaltung der Spezifikation führt.

Bei veränderten Umgebungsvariablen (insbesondere Temperatur) muss dieser Vorgang wiederholt werden.

Die Evaluierung des kritischen Pfads geschieht mittels Simulationstools. Das Delay eines digitalen Gatters ist aber sensibel gegenüber Mismatch. Hierbei hat das geometrische Mismatch, also abweichende W zu L-Verhältnisse, einen starken

Einfluss, was Abbildung 3.6 zeigt (siehe auch Kapitel 3.6.1). Das Mismatch der physikalischen Parameter beeinflusst das Delay zwar auch, jedoch können diese Parameter für kleine bis mittelgroße Chips als global angenommen werden. Das geometrische Mismatch ist hingegen zufällig. Die Empfindlichkeit gegenüber dem geometrischen Mismatch steigt mit sinkenden absoluten Transistordimensionen, da die absoluten Dimensionsfehler, bedingt durch Unterdiffusion und Masken- bzw. Linsenungenauigkeit und auch Positionierungsfehler, absolut gleich bleiben.

Auch ist das Delay eines digitalen Gatters stark abhängig von der Temperatur. Die Gattertemperatur setzt sich zusammen aus der globalen Umgebungstemperatur und der von der Schaltung selbst entwickelten Temperatur bedingt durch Leistungsumsatz. Dieser ist direkt proportional zur Schaltaktivität, so dass besonders aktive Schaltungsteile übermäßig warm werden.

All diese Faktoren führen dazu, dass ein einmalig festgestellter kritischer Pfad nicht eindeutig ist. Schlimmer noch, er kann durch eine heterogene Temperaturverteilung OnChip während des Betriebes wechseln. Ein System, das einen festen kritischen Pfad zur Evaluierung der niedrigsten Versorgungsspannung benutzt, kann so unter Umständen eine zu niedrige Versorgungsspannung generieren, was zu einer Nichteinhaltung der Spezifikationen führen kann. Als zusätzliche Schwierigkeit erweist sich die Messung des Delays eines typischen Pfads. Zwar ist dies die theoretisch längste Verzögerung, die im Design vorkommt, in modernen Designs ist das aber immer noch eine sehr kurze Zeitperiode. Um diese zuverlässig zu erfassen, sind hohe Anstrengungen notwendig.

### 3.8 Ringinverterkette als Performanceindikator

Die in Kapitel 3.7 angesprochenen Schwierigkeiten bei der Evaluierung des kritischen Pfads lassen sich umgehen, indem Teststrukturen verwendet werden, deren Sensibilität gegenüber dem geometrischen Einfluss kleiner ist. Eine solche, frei skalierbare Teststruktur ist die Ringinverterkette (Abbildung 3.10).

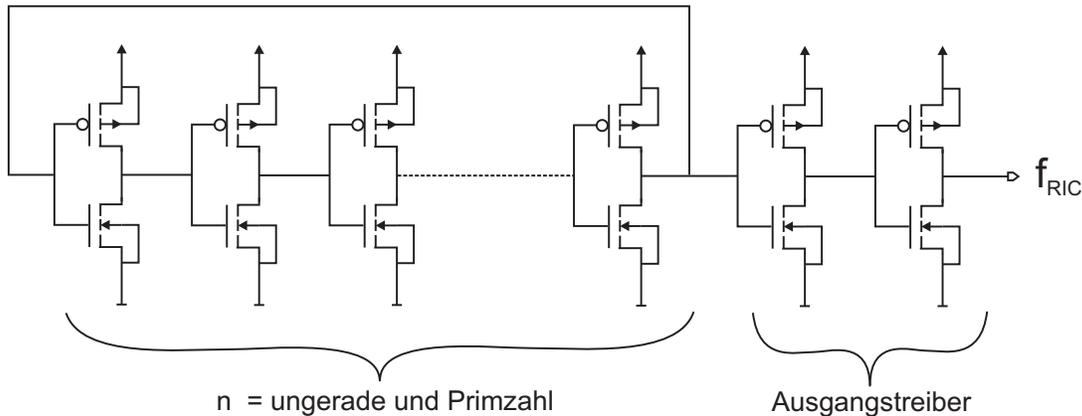


Abbildung 3.10: Ringinverter Prinzip

Die Anzahl der Inverter sollte eine Primzahl sein, um unerwünschte Oberwellen innerhalb des selbstschwingenden Systems zu unterdrücken. Die Frequenz, mit der die Ringinverterkette (im Folgenden RIC genannt) schwingt, wird durch die Laufzeit vom ersten bis zum letzten Inverter bestimmt, ist also abhängig von der Dimensionierung und der Anzahl der Inverter. Das Ausgangssignal wird über einen Ausgangstreiber (vergl. Abb.3.10) ausgekoppelt, der ein sauberes Rechtecksignal erzeugt.

Der große Vorteil, der sich durch die Verwendung der frei dimensionierbaren RIC ergibt, ist die Möglichkeit, die Inverter im Hinblick auf minimale Sensibilität gegenüber geometrischem Mismatch zu dimensionieren. Eine Simulation der Laufzeit einer minimal dimensionierten Inverterkette aus 5 Invertern mit, in den Grenzen der WorstCase Parameter ( $\Delta W = \pm 40nm$   $\Delta l = \pm 25nm \equiv 14\%$  von  $W_{min}, L_{min}$ ) variierenden Transistordimensionen, führt zu einer Laufzeit zwischen 141ps und 224ps. Wegen des lokalen, quasi zufälligen Charakters des geometrischen Mismatches, ist die Laufzeit eines minimal dimensionierten Gatterpfads nicht charakteristisch für die Gesamtheit der Pfade OnChip. Dies führt zu der getroffenen Aussage, dass kritische Pfade in komplexen digitalen Schaltungen nicht eindeutig bestimmbar sind.

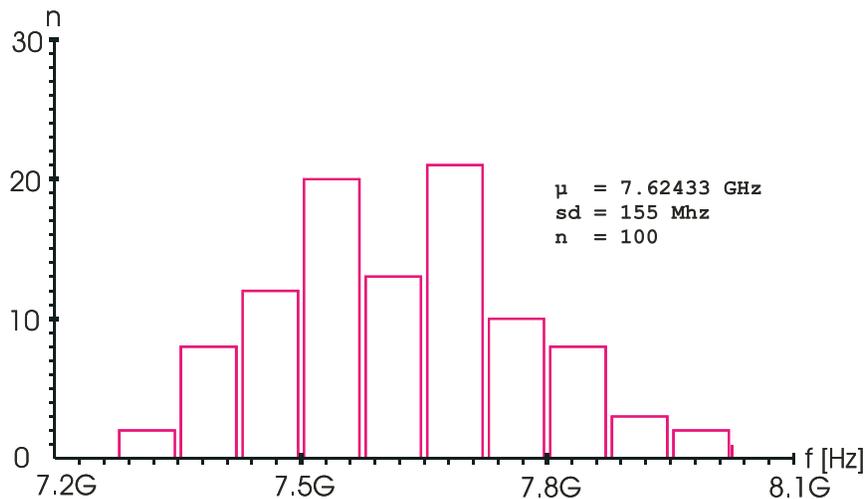
Eine Ringinverterkette kann frei dimensioniert und somit der Einfluss des zufälligen geometrischen Mismatches durch Verwendung großer Transistoren minimiert werden. Die RIC-Frequenz ist somit nur abhängig von Chip-globalen Parametern, deren Einfluss in gleicher Weise auch für die eigentliche Digitalschaltung

des zu betrachtenden Chips vorhanden ist. Somit kann zwischen zufälligem geometrischen Mismatch und systematischem Chipparametern zuverlässig getrennt werden und innerhalb der WorstCase Szenarien für das geometrische Mismatch die gewonnene Information dazu genutzt werden, eine Chip-individuelle Versorgungsspannung zu generieren.

### 3.8.1 Dimensionierung der Inverterkette

Um eine möglichst kleine Sensibilität vom geometrischen Mismatch zu erreichen, werden statistische Untersuchungen (Monte Carlo Simulationen) mit unterschiedlichen Ringinverterkette-Dimensionierungen durchgeführt. Dabei werden nur die geometrischen Abmessungen der Transistoren variiert, denn Änderungen der Ringinverterketten-Frequenz aufgrund des Mismatch physikalischer Parameter und der Temperatur sind ja erwünscht.

Abbildung 3.11 zeigt die statistische Auswertung (Histogramm) für die kleinst mögliche Realisierung einer Ringinverterkette, bestehend aus drei minimal dimensionierten Invertern.



**Abbildung 3.11:** Histogramm einer RIC bestehend aus drei minimal dimensionierten Invertern

Die Abweichung vom Mittelwert ist hier recht groß ( $\approx 2\%$ ), ein weiterer Nachteil ist die sehr hohe Schwingfrequenz (im Mittel  $\overline{f_{RIC}} = 7.62GHz$ ), welche eine elektronische Auswertung im selben Chip unmöglich macht.

In Abbildung 3.12 wird die Länge der Ringinverterkette auf die Anzahl von  $n = 127$  minimal dimensionierte Inverter erhöht. Dadurch ist eine Verkleinerung der Ausgangsfrequenz um den Faktor  $\frac{127}{3} = 42.34$  zu erwarten. In der Tat liegt die mittlere Frequenz bei  $f_{RIC} = 172MHz$ , was einem Faktor von 44.3 entspricht. Eine kleinere Streuung der Ausgangsfrequenz der Ringinverterkette wird erreicht, indem die Dimensionierung der einzelnen Transistoren vergrößert wird. Abbildung 3.13 verdeutlicht dies für eine Ringinverterkette mit dreizehn groß dimensionierten Invertern. Die Frequenz mit  $\overline{f_{RIC}} = 5.767MHz$  liegt hierbei in einem völlig problemlosen Bereich, die Standardabweichung entspricht hier  $\approx 0.2\%$  vom Mittelwert und ist somit um das Hundertfache besser als in Abbildung 3.11.

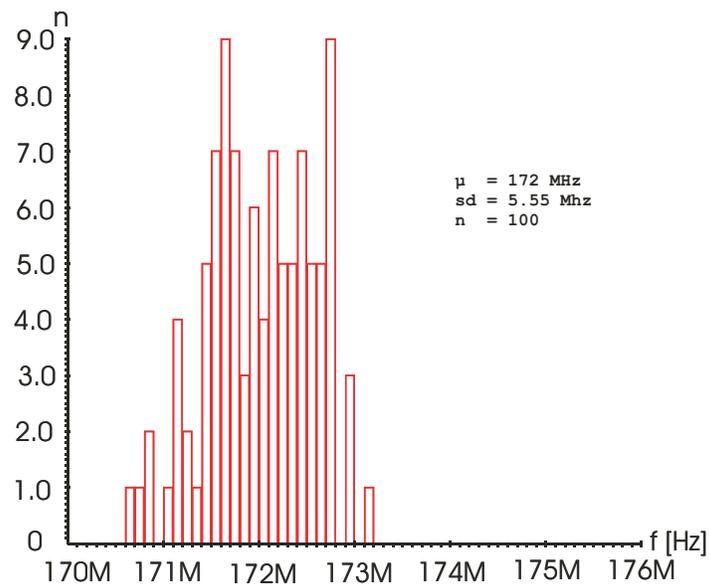


Abbildung 3.12: Histogramm einer RIC bestehend aus 127 minimal dimensionierten Invertern

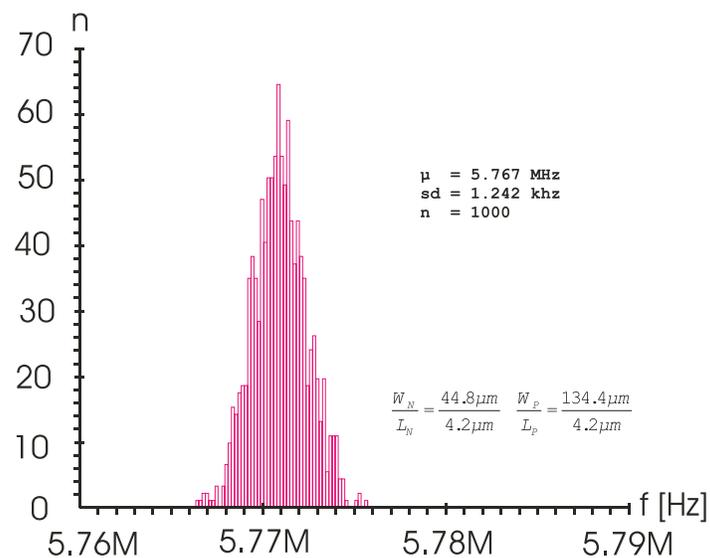


Abbildung 3.13: Histogramm einer RIC bestehend aus 13 groß dimensionierten Invertern

### 3.9 Evaluierung der RIC-Schwingfrequenz

Die Aufgabe der Auswertung der RIC-Schwingfrequenz ist die Generierung einer typischen Ausgangsspannung. Da sowohl die Frequenz der Ringinverterkette, als auch die maximale Schaltfrequenz von den auftretenden Werten der physikalischen Parameter und auch von der Chiptemperatur abhängt, sind RIC-Frequenz und Chip-Performance ( $Clock_{max}$ ) miteinander korreliert. Diese Korrelation soll dazu genutzt werden, um für die gerade auftretenden Parameter eine typische Ausgangsspannung zu generieren. Da der Designer die Schaltung auf WorstCase Parameter dimensioniert, heißt das für die meisten Fälle, dass die Schaltung die geforderte Performance (Spezifikation) auch mit einer kleineren Spannung, als der Standard-Versorgungsspannung, erreichen kann. Abbildung 3.14 zeigt eine typische Verteilung der physikalischen Parameter einer Chipproduktion.

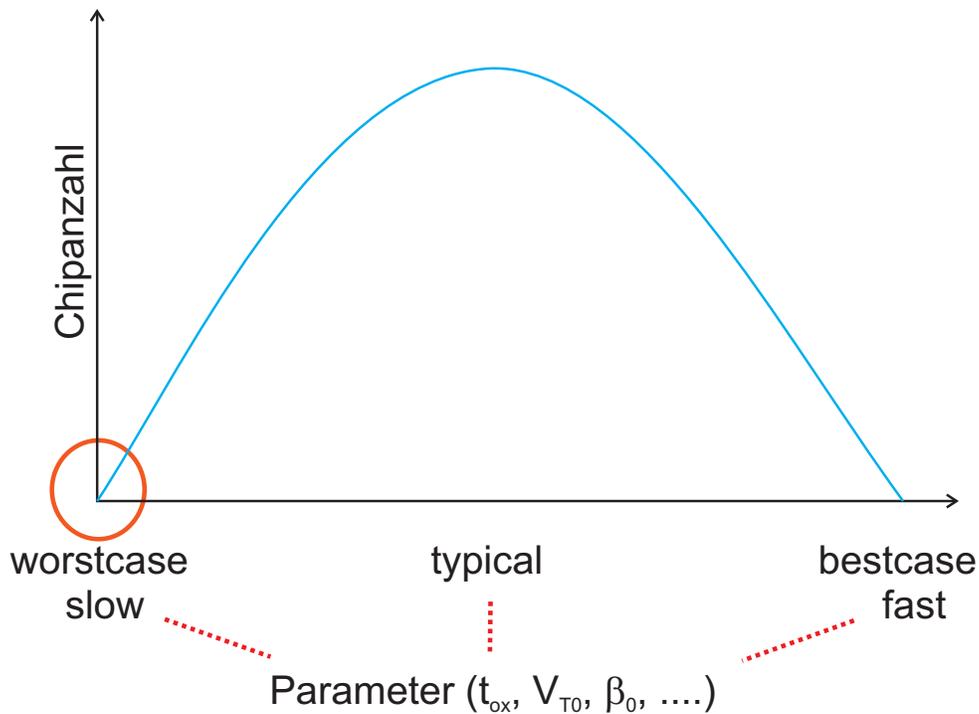


Abbildung 3.14: Verteilung der physikalischen Parameter einer Chipproduktion über einen längeren Zeitraum

Aufgrund des beträchtlichen Einflusses der Versorgungsspannung (siehe Übersicht in Abbildung 3.15) auf jede Verlustleistung (vergl. Kapitel 1), ist die Reduzierung der Versorgungsspannung die am meisten erfolgversprechende Methode zur Reduzierung der Gesamtverlustleistung.

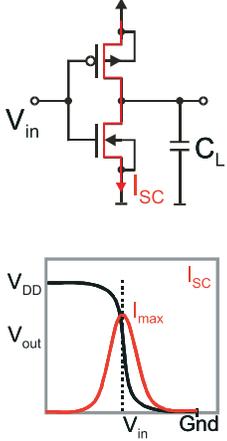
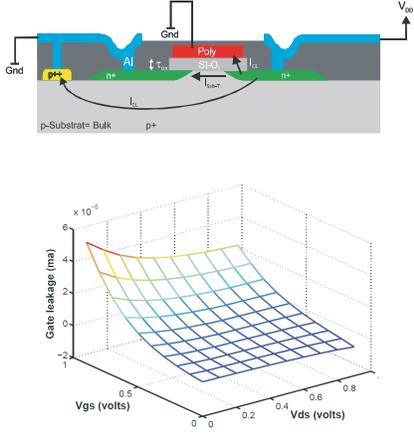
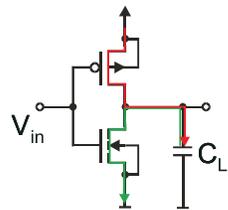
Kurzschluss-Verlustleistung	Leakage-Verlustleistung	Dynamische-Verlustleistung
$P_{SC} = \frac{\beta}{12} (V_{DD} - 2V_{th})^3 \frac{\tau}{T_{clk}}$	$I_{DL} = I_S \cdot \left( e^{\frac{V_D}{V_{therm}}} - 1 \right)$ $I_{DSL} = \mu_0 \cdot \frac{\epsilon_{si}}{\tau_{ox}} \cdot V_{therm} \cdot e^{1.8} \cdot \frac{W}{L} \cdot e^{\frac{V_{GS} - V_{therm}}{n \cdot V_{therm}}} \cdot \left( 1 - e^{\frac{-V_{DS}}{V_{therm}}} \right)$	$P_{dym, cap} \sim f_{clock} \cdot \alpha \cdot C_L \cdot V_{DD}^2$
		

Abbildung 3.15: Überblick über die Verlustleistungen in CMOS-Schaltungen

### 3.9.1 Auswertung durch Frequenzangleichung eines VCO

Das Ziel der Auswertung der RIC-Frequenz ist immer, eine Spannung zu generieren, die zu der RIC-Frequenz umgekehrt proportional ist. Dieses Ziel kann durch verschiedene Methoden erreicht werden. In diesem Fall soll die Eignung eines spannungsgesteuerten Oszillators (VCO) hierzu untersucht werden. Abbildung (3.16) zeigt den prinzipiellen Aufbau der Messschaltung mit digitaler Auswertung: Die Frequenz eines VCO, dessen Schwingfrequenz mit steigender Steuerspannung abnimmt, wird iterativ an die der Sensorinverterkette angenähert. Hierzu wird mit der niedrigsten Steuerspannung (analog zur höchsten zu erwartenden RIC-Frequenz) begonnen. Mit jedem Iterationsschritt wird die Steuerspannung um einen festen Betrag erhöht, bis die Frequenz des VCO gerade unter der der RIC liegt. Die dann anliegende Steuerspannung ist umgekehrt proportional zur RIC-Frequenz. Um die Auswirkung des Fehlers durch die gequantelte Steuerspannungserhöhung zu minimieren, ist der Betrag der Spannungserhöhung im Rahmen der Spezifikationen klein zu halten.

Die Messschaltung besteht aus zwei Schwingkreisen (RIC und VCO), zwei digitalen Zählern (counter1 und counter2) mit Auswertung (AND-Gatter) und einem Spannungsinkrementierer (U incrementer).

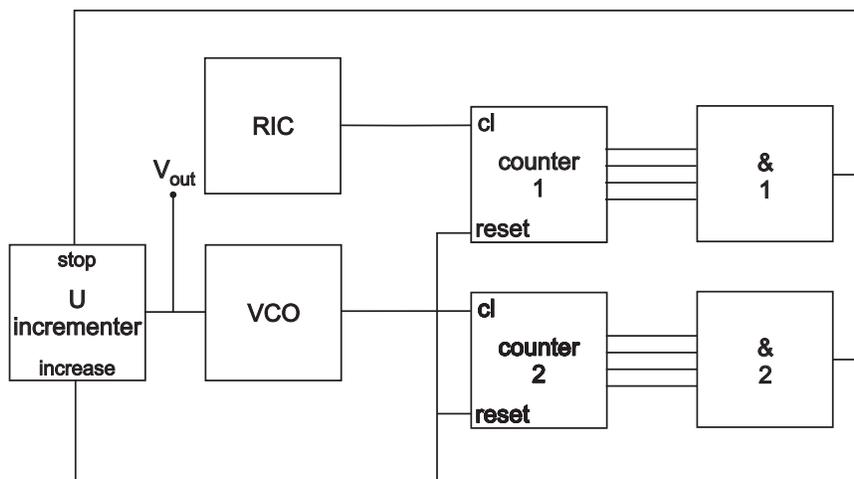


Abbildung 3.16: Schematischer Aufbau der Auswerteschaltung mit VCO

Zu Beginn der Auswertung sind beide Counter leer (alle Bits 0), d.h. die AND Gatter sind low am Ausgang. Der VCO ist, wie später gezeigt wird, so bemessen, dass er mit der niedrigsten angestrebten Ausgangsspannung ( $V_{min}$ ) die höchste, von der RIC generierte, Ausgangsfrequenz liefert.  $V_{min}$  ist somit der Startwert des Spannungsinkrementierers. In nahezu allen Fällen ist die anliegende RIC-Frequenz niedriger, so dass der Counter2 zuerst voll (alle Bits 1) ist und somit das an ihn anliegende AND-Gatter nach high schaltet. Hierdurch werden beide

Counter zurückgesetzt. Das Ausgangssignal des AND-Gatters ist also ein kurzer Puls, durch den auch der Spannungsinkrementierer das Steuersignal erhält, seine Ausgangsspannung um einen festen Betrag zu erhöhen. Die Auswertung findet nun erneut statt; dieser Vorgang wiederholt sich solange, bis die Frequenz der RIC höher ist als die des VCO. In diesem Fall schaltet das AND-Gatter von Counter 1 zuerst auf High und bewirkt ein Festhalten der Ausgangsspannung  $V_{out}$  des Spannungsinkrementierers. Die dann anliegende Ausgangsspannung ist umgekehrt proportional zur Schwingfrequenz der Ringinverterkette und ein Maß für die einzustellende Versorgungsspannung.

### 3.9.2 Sensorinverterkette des Messsystems

Zur Verifikation der Eignung des Messsystems wurde der Frequenzbereich willkürlich auf  $\approx 100\text{kHz}$  festgelegt. Dementsprechend wurde die Sensorinverterkette dimensioniert. Die Kenndaten der Sensorinverterkette sind in Tabelle (3.3) aufgelistet.

Kenndaten der Sensorinverterkette @27°C	
Versorgungsspannung	$V_{DD} = 1.8\text{ V}$
Inverteranzahl	7
Transistorweite $W_{PMOS}$	$0.84\ \mu\text{m}$
Transistorweite $W_{NMOS}$	$0.28\ \mu\text{m}$
Transistorlänge $L_{PMOS}$	$45\ \mu\text{m}$
Transistorlänge $L_{NMOS}$	$45\ \mu\text{m}$
Frequenz $f$ (FAST)	$95.13\ \text{kHz}$
Frequenz $f$ (TYPICAL)	$93.66\ \text{kHz}$
Frequenz $f$ (SLOW)	$91.93\ \text{kHz}$
Mittelwert $\mu(f_{OUT})$	$93.64\ \text{kHz}$
Standardabweichung $\sigma(f_{OUT})$	$0.12\ \text{kHz}$

**Tabelle 3.3:** Kenndaten der Sensorinverterkette

Die Schwingfrequenz der Ringinverterkette hängt in stärkerem Maße als von den physikalischen Parametern (*slow, typ, fast*) von der Temperatur ab. Dieser Effekt ist typisch für die Performance von Digitalschaltungen in der betrachteten Technologie.

### 3.9.3 Aufbau und Wirkungsweise des VCO

Der VCO besteht aus einem CMOS-Schmitt-Trigger, zwei gesteuerten Stromquellen, die jeweils eine Kapazität laden, und aus verschiedenen CMOS-Invertern, die das Signal buffern und in ein sauberes Rechtecksignal formen. Der Spannungsteiler am Eingang des VCO, realisiert aus zwei als Dioden geschalteten NMOS-Transistoren, um eine gute Temperaturkonstanz zu erreichen, sorgt für eine Verdopplung des Eingangsspannungsbereiches. Abbildung (3.17) zeigt die einzelnen Schaltungsteile des Gesamtschaltbildes des VCO (Dimensionierung siehe Anhang A.1).

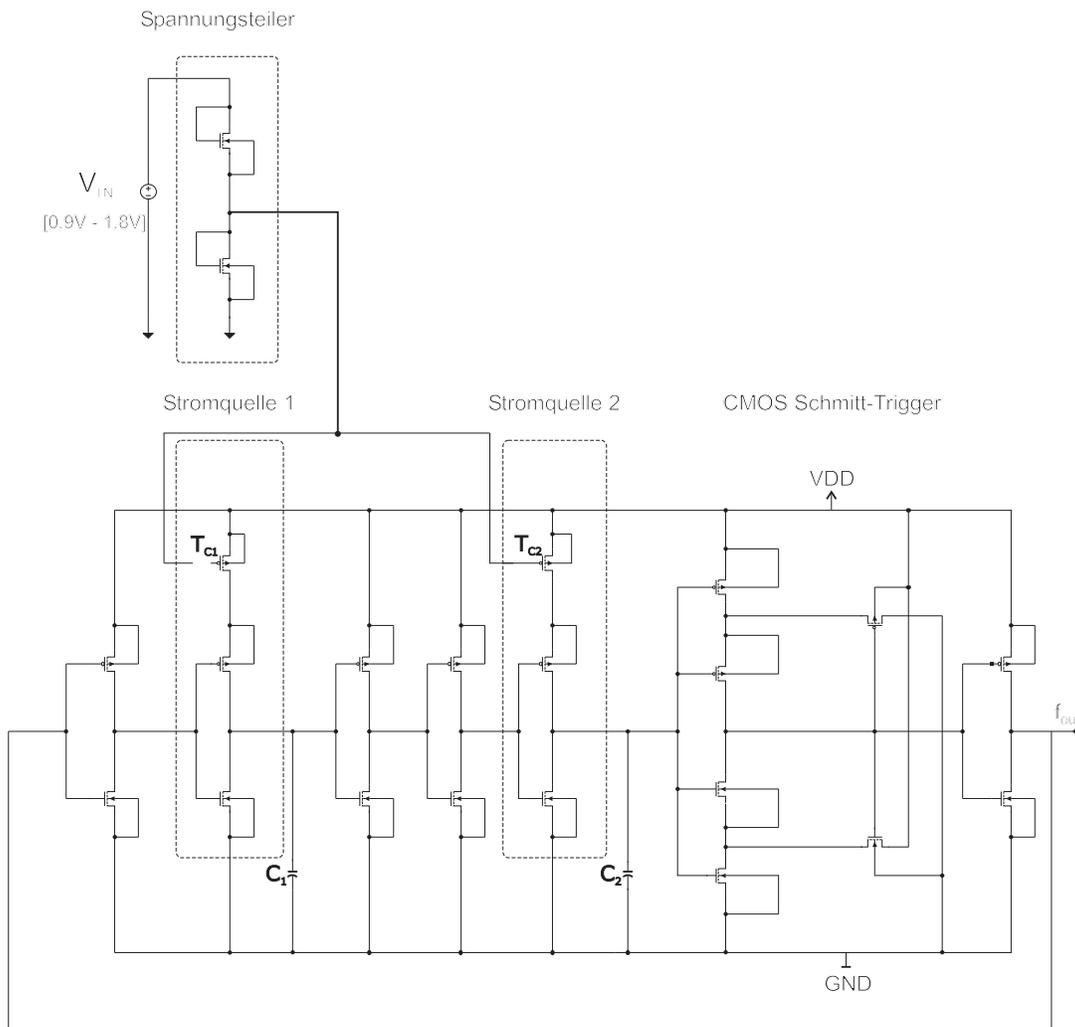


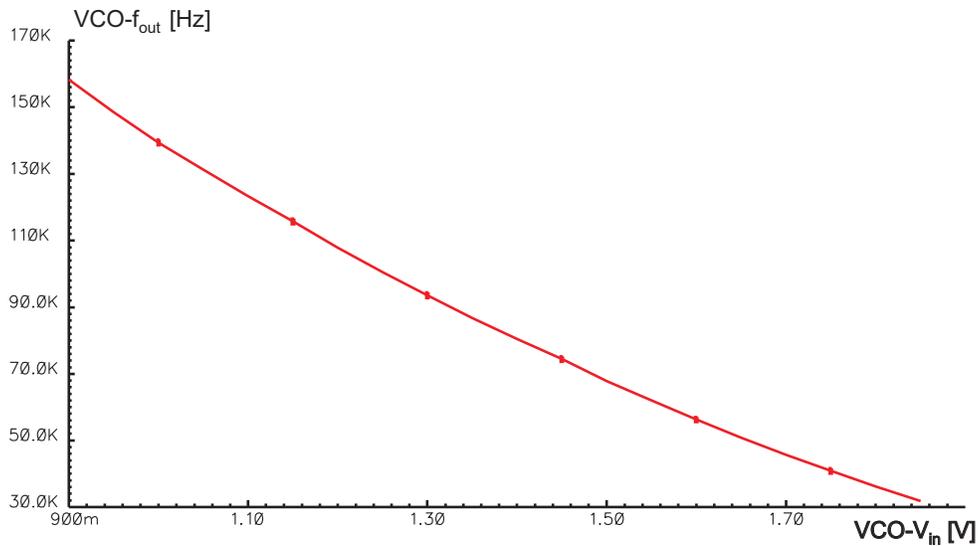
Abbildung 3.17: Schaltbild des VCO

Mit steigender Eingangsspannung  $V_{IN}$  verringert sich der Stromfluss durch die PMOS-Transistoren  $T_{C1}$ ,  $T_{C2}$ , und der Aufladevorgang der Kapazitäten  $C_1$ ,

$C_2$  wird verlangsamt. Die Ausgangsfrequenz  $f_{OUT}$  des VCO nimmt ab. Dies trägt der Forderung nach der umgekehrten Proportionalität von Steuerspannung und VCO-Schwingfrequenz Rechnung. Tabelle 3.4 zeigt den Eingangsspannungsbereich und den Bereich der Ausgangsfrequenz des VCO. Die Frequenz  $f_{OUT}$  weist in Abhängigkeit von der Eingangsspannung  $V_{IN}$  eine hinreichend gute Linearität auf (siehe Abbildung 3.18). Die Dimensionierung erfolgt mit Hinblick auf den Spannungs- und Frequenzbereich der zu versorgenden Schaltung bzw. der verwendeten Ringinverterkette.

typische Eingangsspannung	$0.9\text{ V} < V_{IN} < 1.8\text{ V}$
Ausgangsfrequenzbereich für $V_{in}$	$158.29\text{ kHz} > f_{OUT} > 36.22\text{ kHz}$

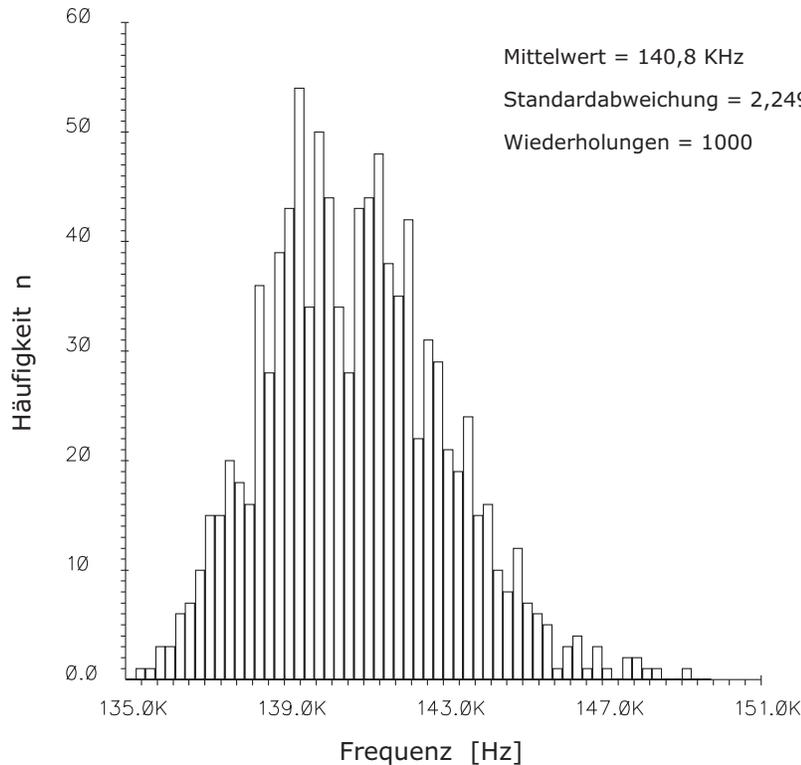
**Tabelle 3.4:** Arbeitsbereich des VCO



**Abbildung 3.18:** Ausgangsfrequenz des VCO in Abhängigkeit der Eingangsspannung @27°C

### 3.9.4 Monte–Carlo Analyse des VCO

Die Ausgangsfrequenz des VCO wird mittels der Monte–Carlo–Methode auf statistische Schwankungen, verursacht durch geometrisches Mismatch, untersucht. Alle statistischen Untersuchungen werden bei einer Eingangsspannung von  $U_{IN} = 1\text{ V}$  durchgeführt. Untersuchungen für andere Eingangsspannungen zeigen keine signifikanten Unterschiede.



**Abbildung 3.19:** Histogramm der VCO-Frequenz bei der Eingangsspannung  $U_{IN} = 1\text{ V}$

Frequenz $f$ (FAST)	154.79 kHz
Frequenz $f$ (TYPICAL)	140.62 kHz
Frequenz $f$ (SLOW)	128.97 kHz
Mittelwert $\mu(f_{OUT})$	140.8 kHz
Standardabweichung $\sigma(f_{OUT})$	2.25 kHz

**Tabelle 3.5:** Prozessabhängigkeit der Ausgangsfrequenz des VCO aus Abbildung 3.17

Die Abhängigkeit der Ausgangsfrequenz des VCO von Prozessparameterschwankungen ist in Tabelle (3.5) zusammengefasst. Ist die Standardabweichung, hervorgerufen durch geometrisches Mismatch, noch akzeptabel, ist die Abhängigkeit der Ausgangsfrequenz von den Prozessparametern (Schwankung von  $\pm 10\%$ ) für eine Auswertung der RIC-Frequenz wesentlich zu hoch. Mit geändertem Design kann diese Abhängigkeit aber minimiert werden. Die Entwicklung eines solchen VCO ist aber sehr technologiespezifisch, bei einer Migration zu einer anderen Technologie muss diese Entwicklungsarbeit i.d.R. erneut vollführt werden. Das macht das Konzept mit VCO unattraktiv. Der Vollständigkeit halber wird im Folgenden die Gesamtschaltung dennoch kurz beschrieben.

### 3.9.5 Der Spannungsinkrementierer

Abbildung (3.20) zeigt den Spannungsinkrementierer. Über das Signal  $A = 1$  wird an  $T_1$  ein Puls mit fester Pulsbreite erzeugt. Die Pulsbreite wird über das Delay  $\tau$  bestimmt. Für die Dauer des Pulses wird die Kapazität  $C$  über den Stromspiegel ( $T_2$  und  $T_3$ ) mit einem konstanten Strom aufgeladen. Wegen des linearen Strom-Spannungszusammenhangs

$$U_C = U_{C0} + \frac{1}{C} \cdot \int_{t_1}^{t_2} i_C \frac{\partial}{\partial t} \quad (3.8)$$

, dem konstanten Strom  $i_C = I_C$  und der Pulsbreite  $\tau = t_2 - t_1$ , erhöht sich bei jedem Puls die Spannung über den Kondensator um:

$$\Delta U_C = \frac{I_C \cdot \tau}{C} \quad (3.9)$$

Diese Spannung wird über den Ausgangsbuffer entkoppelt und steht als Referenzspannung zur Verfügung.

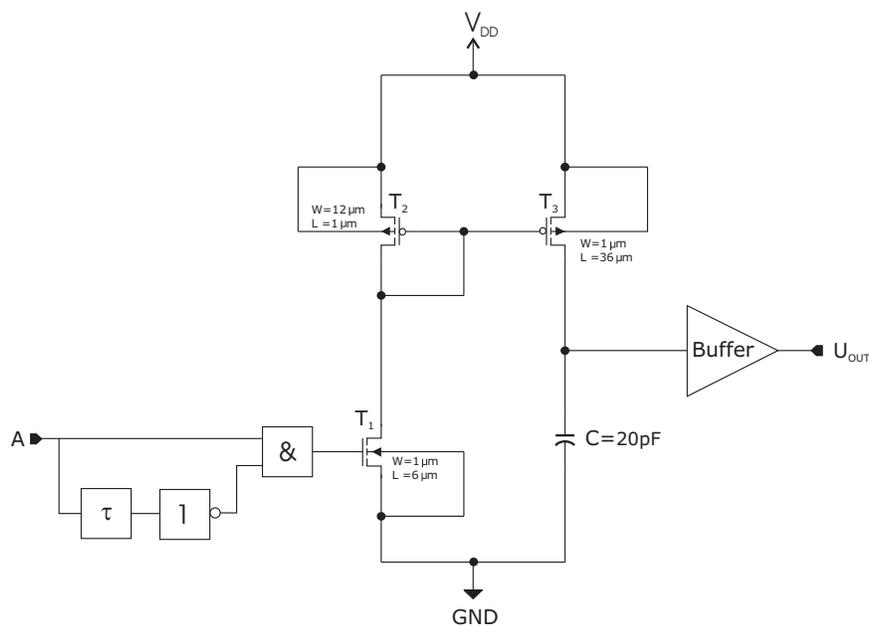
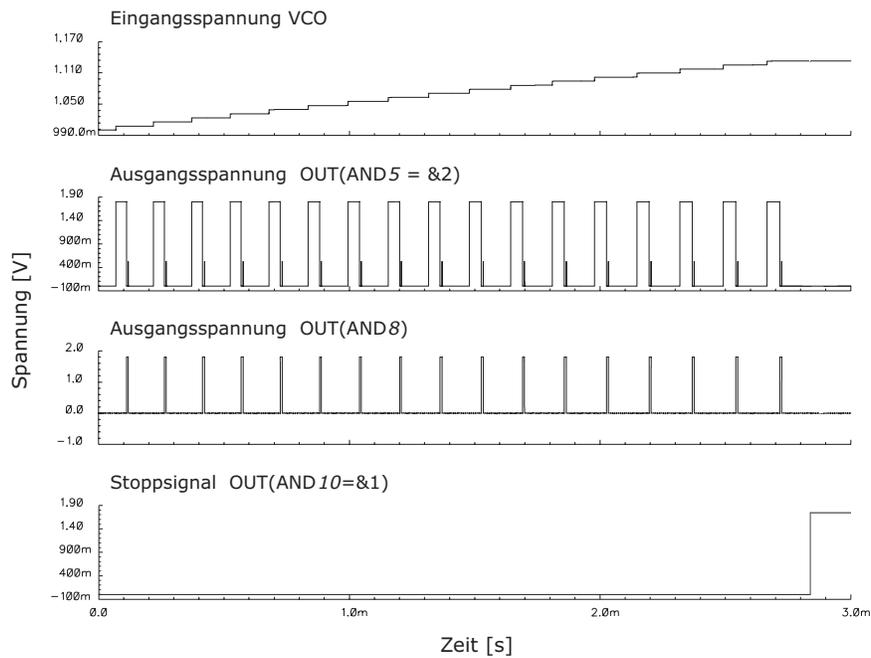


Abbildung 3.20: Spannungsinkrementierer

Für die Simulation der Schaltung wird die Kapazität mit einer Spannung  $V_{min}$  zum Zeitpunkt  $t = t_0$  beaufschlagt.

### 3.9.6 Simulationsergebnisse der Auswerteschaltung mit VCO

Das Gesamtsystem der Auswerteschaltung mit VCO ist im Anhang (A.2) nochmals ausführlich dargestellt. Auf tiefere Erklärungen wird wegen der eingeschränkten Eignung des VCO an dieser Stelle verzichtet. Abbildung 3.21 zeigt die Inkrementierung der VCO-Steuerspannung bis zum Endwert (Stopsignal=1).



**Abbildung 3.21:** Spannungsverlauf der Ausgangsgrößen

Das Signal  $Out(AND8)$  (analog der Schaltung aus Abbildung A.2) erzeugt eine Zeitspanne von  $40\mu s$ , in der das System abwartet, bis sich der VCO auf eine neue Frequenz eingeschwungen hat, bevor erneut mit der Auswertung der RIC-Frequenz begonnen wird.

## 3.10 Alternative Auswerteschaltung mit VCO

### 3.10.1 Funktionsprinzip der Schaltung

An dieser Stelle soll noch kurz eine alternative Methode zur Evaluierung der RIC-Frequenz mit Hilfe eines VCO vorgestellt werden. Das Prinzip ähnelt dem in Kapitel 3.9.1 dargestellten. Statt die Frequenzsignale in einer Zählerschaltung zu erfassen, arbeiten nun beide Signale, die des VCO und die des RIC, auf einen Spannungsinkrementierer. Die Frequenz des VCO wird auch hier iterativ erniedrigt (durch U-incrementer 1). Die Dauer einer Iteration wird hierbei durch eine Referenzspannung bestimmt, bis zu deren Wert der Spannungsinkrementierer, der von der RIC angesteuert wird, ansteigt (U-incrementer 2). Die Höhe der Referenzspannung ist hierbei willkürlich, deswegen könnte sie auch durch einen Spannungsteiler zur optimalen Ansteuerung des Komparators 1 zum Einsatz kommen. Da die VCO-Frequenz am Anfang der Messung in jedem Fall höher ist, als die höchste vorkommende RIC-Frequenz, ist hierdurch sichergestellt, dass am Ende der ersten Messung die Ausgangsspannung des VCO-Spannungsinkrementierers höher ist als die des RIC-Spannungsinkrementierers.

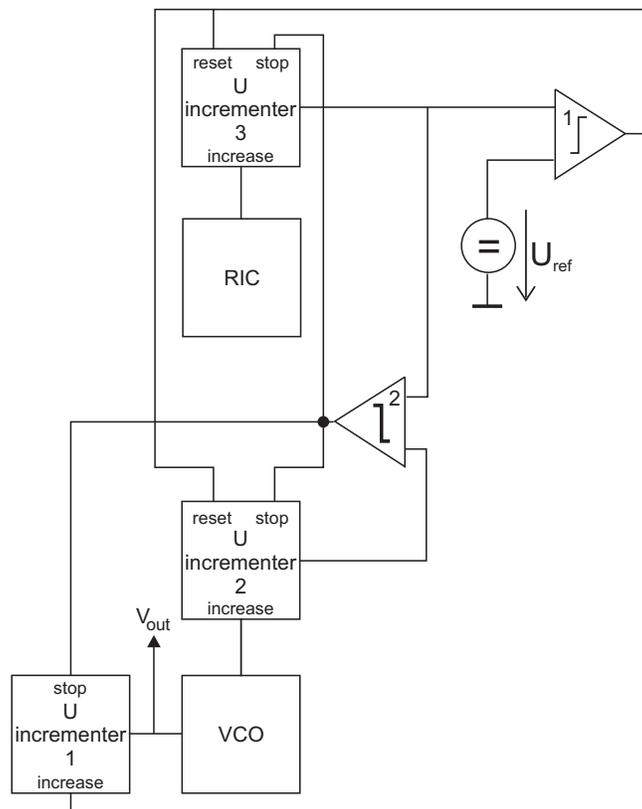
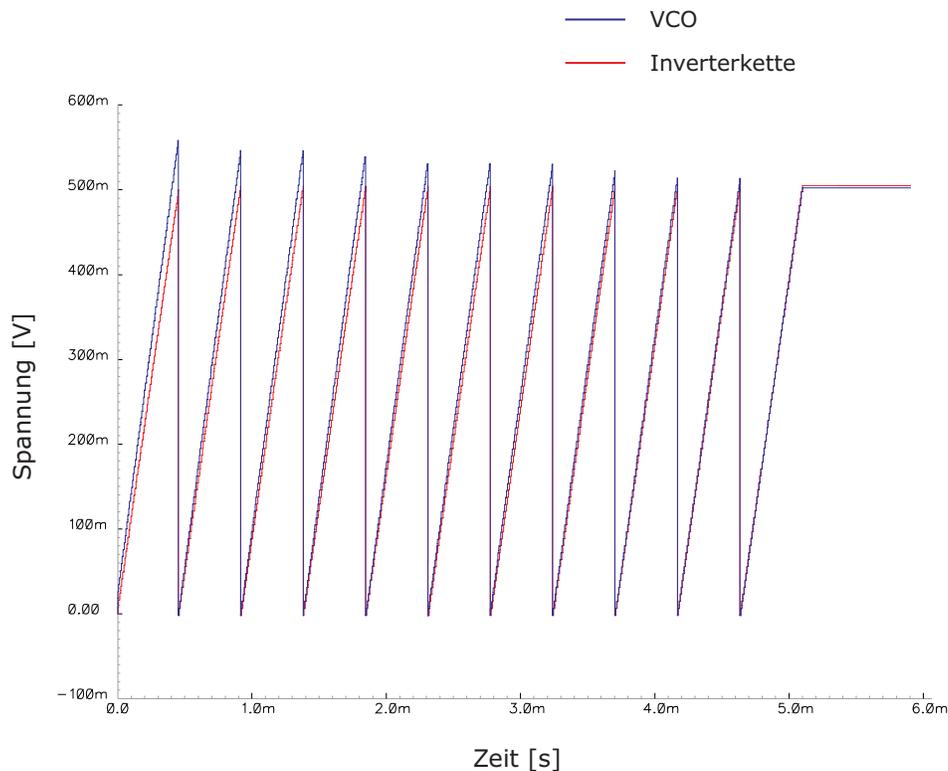


Abbildung 3.22: Aufbau der Messschaltung mit analoger Auswertung

Somit bleibt das *stop*-Signal von Komparator 2 aus, U-incrementer 3 erhöht seine Spannung bis zum Referenzwert, wonach Komparator 1 sowohl das *reset*-Signal für beide U-incrementer 2 und 3, als auch das *increase*-Signal für U-incrementer 1 gibt. Der Ablauf der Messung wird somit erneut initiiert, allerdings mit erhöhter Ausgangsspannung  $V_{out}$ , was die Frequenz des VCO erniedrigt. Der Vorgang wiederholt sich solange, bis die Frequenz des VCO gerade kleiner wird, als die der RIC, wodurch der Komparator 2 das *stop*-Signal erzeugt. Die dann anliegende Ausgangsspannung ist charakteristisch für die RIC-Frequenz und zu ihr umgekehrt proportional. Abbildung 3.22 zeigt den schematischen Aufbau der Messschaltung.

Die vollständige Schaltung mit dem Simulationsergebnis wesentlicher Signale ist im Anhang Abbildung A.3 dargestellt. In Abbildung 3.23 zeigt sich die iterative Annäherung der zeitlichen Spannungsverläufe der beiden Spannungsincrementierer der RIC und des VCO, bis die Frequenz des VCO gerade kleiner ist, als die der RIC (hier für den Fall einer Referenzspannung von 500mV).

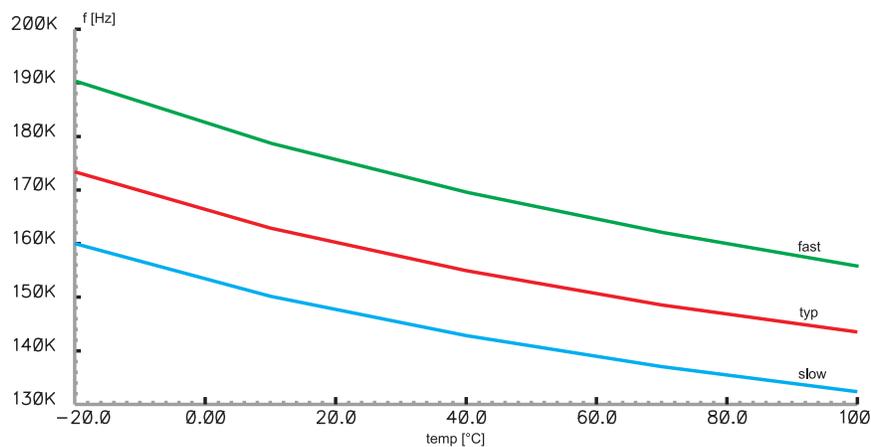


**Abbildung 3.23:** Simulationsergebnis der beiden Spannungsincrementierer der RIC und des VCO

Ein Vorteil dieser Auswerteschaltung ist das Fehlen der digitalen Zähler. Somit wird Fläche gespart, was aber mit der Verwendung zweier zusätzlicher Spannungsinkrementierer relativiert wird. An diese werden nun höhere Matchinganforderungen gestellt, da deren Ausgangssignale miteinander verglichen werden. Jedoch kommt es auch hier nicht auf absolutes Matching an, relatives Matching der Transistoren und Kapazitäten ist bekanntlich in integrierten CMOS-Prozessen leichter zu erreichen. Das Matchingproblem des VCO bleibt aber auch hier symptomatisch.

### 3.11 Beurteilung der Schaltungskonzepte mit VCO

Die Messprinzipien stellen grundsätzlich funktionierende Analyseverfahren dar, mit dem Prozessparameterschwankungen auf einem Chip untersucht werden können. Ein entscheidendes Manko macht die Verwendung der Messverfahren mit VCO fraglich. Die Ausgangsfrequenz des VCO unterliegt Mismatch-bedingten Schwankungen. Von der Struktur her ähnlich der Ringinverterkette reagiert die Ausgangsfrequenz des VCO sowohl auf geometrisches Mismatch, physikalische Prozessschwankungen als auch auf Schwankungen der Temperatur (vergl. Abbildung 3.24).



**Abbildung 3.24:** Schwingfrequenz des VCO @  $V_{in} = 0.9\text{V}$  in Abhängigkeit der Temperatur und der Prozessgrenzen

Die Entwicklung anderer VCO-Schaltungen, die diesen Schwankungen weniger unterliegen, ist prinzipiell möglich, stellt aber einen nicht unerheblichen Entwicklungsaufwand dar, der für jede Technologiemigration erneut erbracht werden muss. Prinzipiell sind die vorgestellten Messverfahren durchaus geeignet. Die entscheidenden Baugruppen sollten hierbei auf optimale Matching- und Temperatur-Unempfindlichkeit hin designed werden. Zu den kritischen Baugruppen gehören hierbei neben dem VCO auch Komparatoren (Offset) und die Stromquellen (Paarigkeit) der Spannungsinkrementierer, sowie die verwendeten Kondensatoren. Da die Optimierung dieser Baugruppen einen nicht unerheblichen Designaufwand bedeutet und diese immer nur für eine bestimmte Technologie durchgeführt werden kann, wird der Nutzen der Messverfahren auf VCO-Basis kritisch beurteilt. Eine optimale Evaluierung der Schwingfrequenz kann nur digital erfolgen.

## 3.12 Digitale Erfassung der Ringinverterschwingfrequenz

Die analoge Auswertung der Ringinverterschwingfrequenz leidet an einer zu grossen Abhängigkeit des Auswertergebnisses vom geometrischen- und Prozessmismatch. Ein digitales System ist grundsätzlich auch von diesen Variationen betroffen, jedoch wirkt sich dieses Mismatch nur in der Verarbeitungsgeschwindigkeit aus. Dies erklärt z.B. die Klassifizierung von Speicherchips oder Prozessoren hinsichtlich ihrer maximalen Taktrate. Für die Anwendung der Auswertung der Schwingfrequenz einer Ringinverterkette ist nur Sorge zu tragen, dass die minimale Verarbeitungsgeschwindigkeit ausreichend ist. Die Schwingfrequenzen der RIC sind hierbei in einem unproblematischen Bereich.

Die Aufgabe des Systems ist es, die Schwingfrequenz der RIC zu detektieren und diese weiterzuverarbeiten. Da Frequenz und Zeit über die Periodendauer verknüpft sind, bietet sich zur Frequenzerfassung ein digitaler Zähler an. Die Tiefe dieses Zählers hängt einerseits von der zu erwartenden maximalen Frequenz ab, andererseits von der Zeit, in der der Ausgang des RIC auf den Zähler gegeben wird. Um eine Variable festzulegen, wird ein Zähler mit 16 flankengesteuerten JK-Flipflops verwendet. Folgende Rechnung bestimmt die Zeit, bis der Zähler bei gegebener Frequenz überläuft:

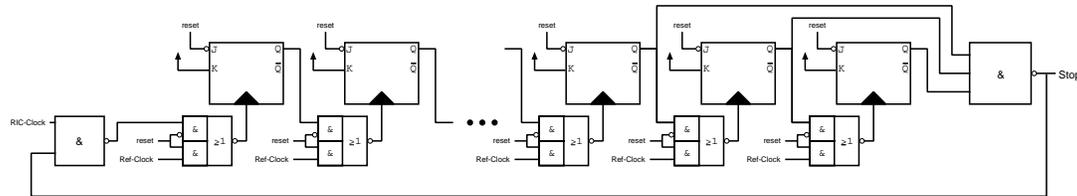
$$f_{max} = 8MHz \Rightarrow T = \frac{1}{8MHz} = 125ns \quad (3.10)$$

Der voll belegte 16 Bit Zähler  $1111\ 1111\ 1111\ 1111_2$  entspricht  $65535_{10}$  Schwingungsdauern mit jeweils einer fallenden Flanke. Es ergibt sich so die Auswertzeit mit:

$$t_{ev} = 65535 \cdot 125ns = 8.191875ms \quad (3.11)$$

### 3.12.1 Erzeugen eines definierten Zeitintervalls

Die Auswertzeit  $t_{ev}$  wird ihrerseits ebenfalls über das Zählen einer Referenzclock verwirklicht. Auch hier gibt es einen willkürlichen Zusammenhang zwischen Zählertiefe und Referenzclock. Es wird ein 14 Bit Zähler mit 14 flankengesteuerten JK-Flipflops verwendet. Die drei höchstwertigen Bits werden hierbei über eine NAND-Verknüpfung zur Generierung eines *STOP*-Signals herangezogen. Abbildung 3.25 zeigt das symbolische Schaltbild der Referenzzeitbestimmung.



**Abbildung 3.25:** Schaltung zur Erzeugung eines Stoppsignals nach einer Referenzzeit

Das Stoppsignal wechselt von *High* nach *Low*, wenn der Zähler wie folgt belegt ist: 1110 0000 0000 00<sub>2</sub>. Dies entspricht 14336<sub>10</sub> Schwingungsdauern. Nach dem Umschalten des Stoppsignals nach *Low* ist das *NAND2* Gatter am Eingang des Zählers aus Abbildung 3.25 für die Referenzclock nicht mehr transparent. Der Zähler verbleibt somit in seinem Istzustand.

Mit Gleichung 3.11 folgt für die erforderliche Schwingungsdauer:

$$T_{ref} = \frac{16.383875ms}{14336} \approx 571.4ns \quad (3.12)$$

Damit ergibt sich für die Referenzfrequenz:

$$f_{ref} = \frac{1}{T_{ref}} = 1750013Hz \quad (3.13)$$

Es wird eine Referenzfrequenz von 1.75MHz gewählt. Die Abweichung von der berechneten Frequenz wirkt sich nur unwesentlich auf die Referenzzeit aus; die resultierende Grenzfrequenz der RIC wird nicht unterschritten.

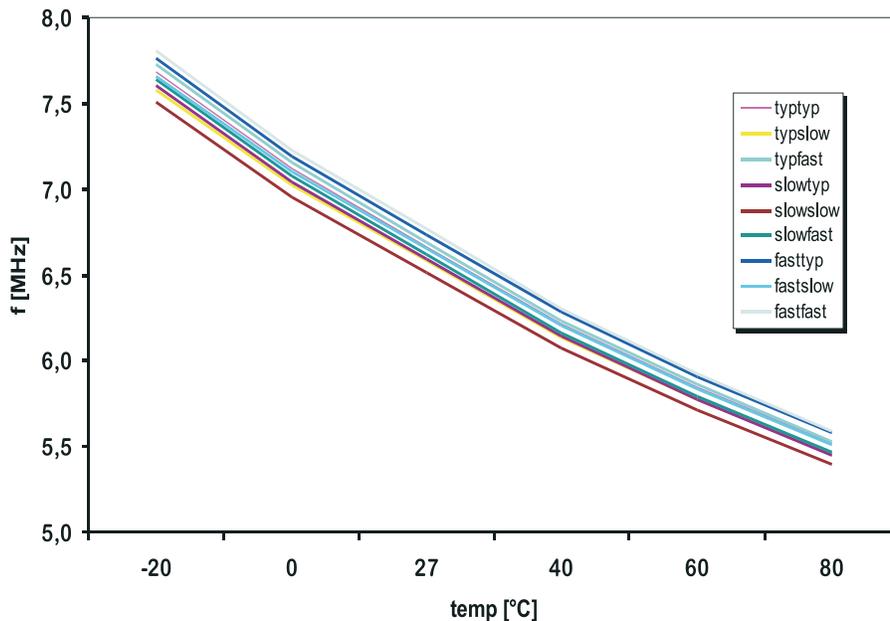
Für das korrekte erneute Generieren des Stoppsignals ist es notwendig, alle Ausgänge der JK-FlipFlops des Zählers auf *Low* zu setzen. Dies geschieht über ein *reset*-Signal. Nach Abbildung 3.25 öffnet ein *reset* = 0-Signal die Flipflop-Kette des Zählers, indem die Multiplexer umschalten.

Jedes einzelne Flipflop wird nun von der Referenzclock getaktet. Zusätzlich werden alle J-Eingänge digital 1. Somit werden nach der ersten steigenden Flanke der Referenzclock (was am FlipFlop durch die Invertierung durch das AOI-Gatter des Multiplexers eine fallende Flanke ist) alle FlipFlop Ausgänge auf

digital 0 gesetzt. Das *reset*-Signal muss folglich mindestens für die Dauer einer Periode der Referenzclock 0 sein. Ist das *reset*-Signal wieder 1, werden die Flipflops wieder zu einem digitalen Zähler verschaltet.

### 3.12.2 Auswerten des Zählerstands

Innerhalb der Referenzzeit werden die Schwingungen der Ringinverterkette gezählt. Je nach den gerade herrschenden Chipkonditionen (Temperatur, physikalische Prozessparameter) variiert der Zählerstand. Bei der Performance zuträglichen Chipkonditionen wird ein hoher Zählerstand erreicht, da die Frequenz der Ringinverterkette hoch ist. Abbildung 3.26 zeigt den Einfluss der Temperatur und der physikalischen Parameter (typslow bedeutet hierbei: typische NMOS-Transistoren, langsame PMOS-Transistoren) auf die Schwingfrequenz des Ringinverters aus Abbildung 3.13.



**Abbildung 3.26:** Einfluss der Temperatur und der physikalischen Parameter auf die RIC-Frequenz

Bemerkenswert ist, dass die Temperatur einen dominanten Einfluss auf die Ringinverterschwingfrequenz und somit auch auf das Delay einer digitalen Schaltung auf gleichem Substrat hat. Die Schwingfrequenz und ihre digitale Abbildung als Zählerstand sind also Messgrößen für die zu erwartende Performance.

Ziel ist es, den Zählerstand in eine spezifische Versorgungsspannung zu überführen. Abbildung 3.27 zeigt das Gesamtkonzept, welches dieses Ziel realisiert. Die fallenden Flanken der Ringinverterkette werden von einem Zähler gezählt, bis das Systemstoppsignal aktiv wird, nämlich nach Ablauf der Zeit, die durch den Referenztakt und den Zählerstand eines zweiten Zählers definiert wird. Somit ist das Wesentliche schon geschehen, der Zählerstand repräsentiert den momentanen Chipzustand (physikalische Parameter und Temperatur). Die Weiterverarbeitung des Zählerstands kann nun auf verschiedene Art und Weise erfolgen. In Abbildung 3.27 sind die **nicht** im Testchip (siehe auch Anhang, Abbildung A.24) implementierten Systemblöcke gestrichelt verbunden bzw. in Klammern gesetzt. Der Zählerstand (die neun höchstwertigen Bits) wird als digitales Wort einem Subtrahierer zugeführt und dort von einem spezifischen digitalen Wort subtrahiert (digital assignment in Abbildung 3.27). Diese spezifische Vorbelegung realisiert zum einen die erwünschte umgekehrte Proportionalität von Zählerstand und zu generierender Referenzspannung, zum anderen legt sie die Grenzen dieser Referenzspannung fest ( $V_{out-min}$  und  $V_{out-max}$ ). Diese Grenzen sind von den Spezifikationen und der jeweiligen Anwenderschaltung abhängig. Der Designer muss diese durch Simulationen mit maximal abweichenden Chipparametern (also slow- und fast-Konditionen bei  $temp_{min}$  und  $temp_{max}$ ) bestimmen. Das Ergebnis der Subtraktion wird einem Digital-Analog-Wandler zugeführt, der aus dem digitalen Wort eine analoge Spannung generiert, die als Referenzspannung für einen DC/DC-Konverter (Kapitel 4) benutzt wird.

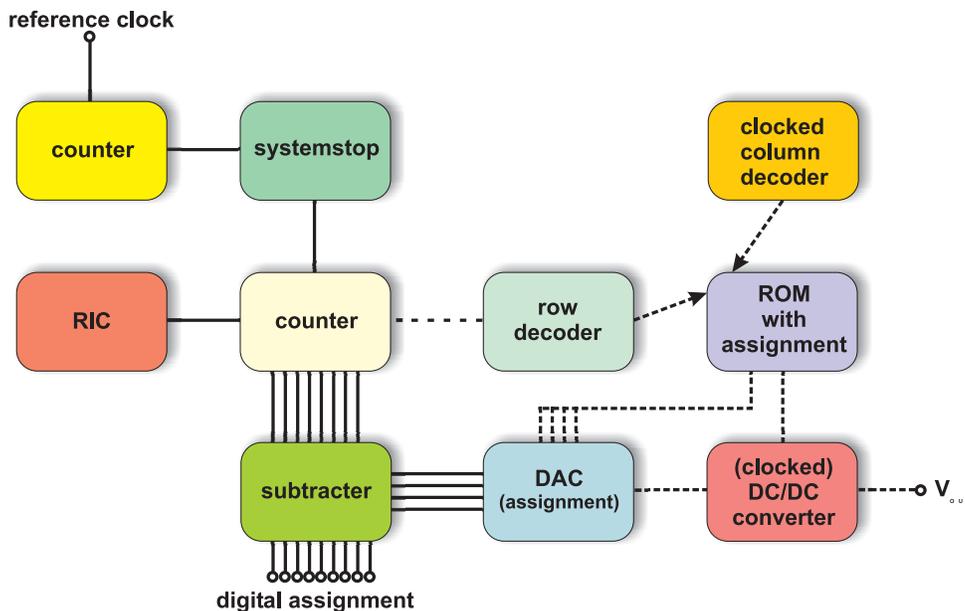
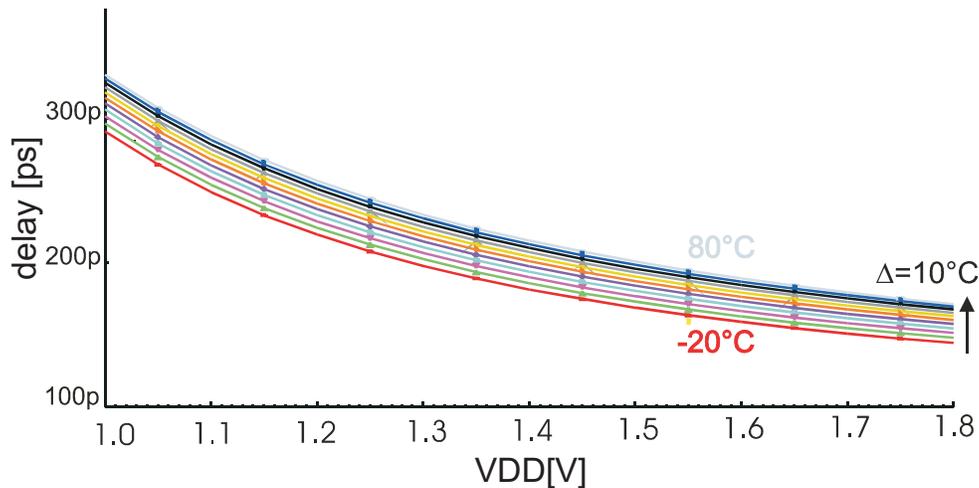


Abbildung 3.27: Konzept zum Performance Driven Supply

Als problematisch bei dieser Vorgehensweise erweist sich das nichtlineare Verhalten der Performance digitaler Schaltungen bei Verringerung der Versorgungsspannung. Abbildung 3.28 zeigt das simulierte Delay einer Gatterkette aus fünf seriellen Gattern abhängig von der Versorgungsspannung, der Temperatur und den physikalischen Parametern.



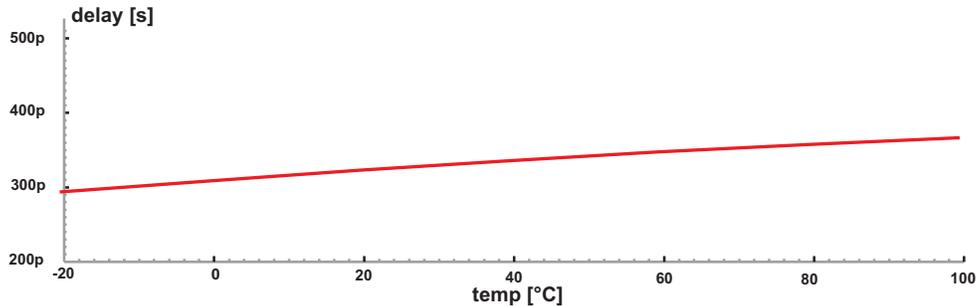
**Abbildung 3.28:** Delay einer Gatterkette mit Best-Case Parametern

Deutlich ist das stärker ansteigende Delay der Gatterkette bei niedrigen Versorgungsspannungen zu erkennen. Dies ist eine Folge der niedrigeren Drain-Source-Spannung aller Transistoren bei niedrigerer Versorgungsspannung. Beim Umschaltvorgang wird damit der lineare Bereich der Transistorausgangskennlinie dominanter, wodurch der Lade- und Entladestrom immer kleiner wird. Diese Nichtlinearität wird bei der Wandlung des Zählerstands nicht berücksichtigt. Infolge dessen muss bei der Berechnung der digitalen Vorbelegung ein Sicherheitsüberschuss an Spannung mit einkalkuliert werden. Dieser kann bei Verwendung des später erläuterten ROMs with assignment (Abb. 3.27) entfallen.

Im Folgenden soll diese Vorgehensweise verdeutlicht werden. Um die Spezifikationen unter allen Umständen einhalten zu können, muss bei der Simulation der Anwenderschaltung von absolut schlechtesten Chip-Konditionen ausgegangen werden. Das bedeutet:

- maximale Temperatur
- physikalische Parameter *slow*
- geometrisches Mismatch maximal ungünstig
- daraus resultierend: critical-path Evaluierung

Das Delay des somit evaluierten critical-path wird nunmehr unter den o.a. Konditionen mit Standardversorgungsspannung bestimmt. Abbildung 3.29 zeigt diese Simulation in Abhängigkeit der Temperatur. Ausgehend von einer maxi-



**Abbildung 3.29:** Delay einer Gatterkette mit Worst-Case Parametern

malen Chiptemperatur kann der Designer ein Delay von 368ps garantieren. Dieser Wert sollte innerhalb der Spezifikation liegen, ist dies nicht der Fall, muss die Schaltung überarbeitet werden bzw. die Versorgungsspannung innerhalb der erlaubten Größen erhöht werden. Das Design erfolgt also unter dem Aspekt maximaler Ausbeute, möglichst alle Chips einer Fertigung sollen den Spezifikationen genügen. Es bleibt also für die Mehrzahl der Chips ein Spielraum, die Versorgungsspannung zu reduzieren, nämlich für den Fall, dass nicht maximal ungünstige Chip-Parameter vorliegen, was fast immer der Fall ist. Für die Gatterkette aus Abbildung 3.29 könnten mit dem o.a. Konzept mit linearem Zusammenhang zwischen Ringinverterfrequenz und Ausgangsspannung exemplarisch die in Abbildung 3.30 dargestellten Versorgungsspannungen generiert werden (blaue Linie: Gatterverzögerung mit individueller Versorgungsspannung; rote Linie: Gatterverzögerung mit fester Versorgungsspannung  $V_{DD} = 1.8V$ ). Dabei ist der Zusammenhang zwischen Ringinverterfrequenz und Ausgangsspannung durch die lineare Beziehung in Gleichung 3.14 gegeben.

$$V_{DD} = -3V \cdot f_{RIC} \cdot \frac{10^{-6}}{Hz} + 3.5V \quad (3.14)$$

Zu beachten ist hierbei, dass die berechneten Versorgungsspannungen in Abbildung 3.30 mit dem vorgestellten Konzept nicht in beliebiger Genauigkeit generiert werden können, da es durch die Quantisierung nur diskrete Ausgangsspannungen geben kann. Mit dem linearen Zusammenhang zwischen Ringinverterkette und Ausgangsspannung (Gleichung 3.14) ist eine Equalisation des Gatterdelays nicht möglich. Im angegebenen Beispiel steigt das Gatterdelay durch die kleine Versorgungsspannung im Bereich kleiner Temperaturen über den Wert des WorstCase aus Abbildung 3.29. Eigentlich müsste der Zusammenhang aus Gleichung 3.14 verändert werden (was einer divergierenden digitalen Vorbelegung des Subtrahierers aus Abbildung 3.27 entspräche), um das Delay nicht

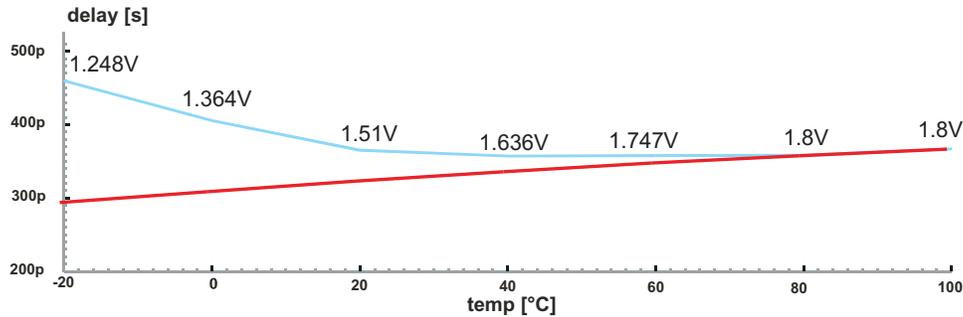


Abbildung 3.30: Delay einer Gatterkette mit Worst-Case Parametern mit variabler Versorgungsspannung

unter den Wert von  $368ps$  fallen zu lassen.

Eine bessere Anpassung der Versorgungsspannung kann erreicht werden, in dem der nichtlineare Zusammenhang von Gatterdelay und Versorgungsspannung berücksichtigt wird. Ein hier nicht weiter betrachteter Weg wäre es, den Digital-Analog-Wandler mit einer solchen Nichtlinearität zu beaufschlagen. Eine Ansteuerlogik könnte die digitalen Eingangssignale eines linearen DACs entsprechend der Nichtlinearität aus dem Zählerstand generieren. Ein flexiblerer Weg ist in Abbildung 3.27 gestrichelt dargestellt. Der Zählerstand, bzw. die relevanten Bits des Zählers, die nach Ablauf der Referenzzeit die Chip-Konditionen repräsentieren, werden nun direkt als Reihen-Adresse eines ROMs verwendet. In diesem ROM sind typische Versorgungsspannungen vom Designer als digitales Wort gespeichert, die nach der Digital-Analog-Wandlung als analoge Referenzspannungen wiederum einem DC/DC-Konverter zugeführt werden. Da das ROM frei beschreibbar ist, kann jegliche Nichtlinearität berücksichtigt werden. Die Abbildung 3.31 zeigt dies exemplarisch für die Gatterkette, die auch schon in den Abbildungen 3.29 und 3.30 betrachtet worden ist.

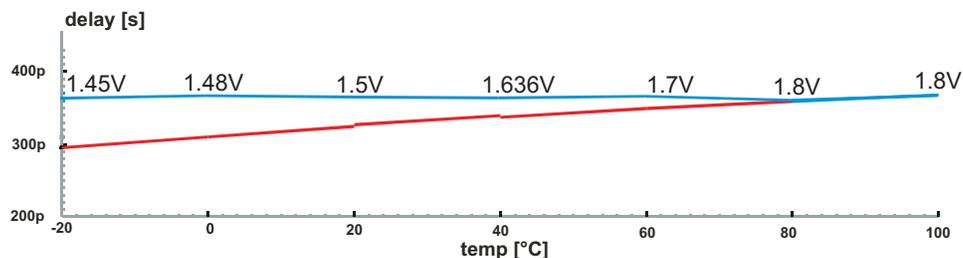


Abbildung 3.31: Delay einer Gatterkette mit Worst-Case Parametern mit optimierter Versorgungsspannung

Durch die nun individuellen Versorgungsspannungen gelingt die Equalisation

des Gatterdelays nahezu perfekt und ist in der Realität nur von der Auflösung des DA-Wandlers (und damit auch von der Speichertiefe des ROMs) abhängig. Eine weitere Alternative, die in Abbildung 3.27 dargestellt wird, ist das völlige Umgehen des DA-Wandlers. Der Zählerstand wird nun als Zeilenadresse eines ROM-Speicherzellenfelds verwendet. Die einzelnen Speicherzellen der entsprechend ausgewählten Zeile werden nun aber über den in Abbildung 3.27 *clocked – column – decoder* mit einer festen Frequenz periodisch ausgelesen. Der Designer kann somit statt digital kodierter Referenzspannungen nun Duty-Cycle-Waveforms im ROM ablegen und durch Verwendung von getakteten DC/DC-Wandlern spezifische Versorgungsspannungen generieren (Kapitel 4).

### 3.13 Einsparpotential des Systems

Die Frequenz der Ringinverterkette ist ein Maßstab für die individuelle Chipperformance, die durch die Temperatur und die chipspezifischen physikalischen Parameter ( $V_{T0}$ ,  $\beta_0$ , etc.) bestimmt wird. Ein weiterer Faktor, der die Performance einer digitalen Schaltung entscheidend mitbeeinflusst, ist das geometrische Mismatch der Transistoren. Dieses wird u.A. durch Unterdiffusion der Drain- und Source-Gebiete unter das Gate, sowie durch Maskenungenauigkeiten hervorgerufen.

Das geometrische Mismatch ist ein lokales Mismatch, d.h. die Abweichungen eines Transistors sind weder in Betrag noch in Vorzeichen mit denen seines Nachbarn korreliert. Darüber hinaus existiert eine Abhängigkeit des geometrischen Mismatches von der Ausrichtung (also vertikale oder horizontale Ausrichtung). Diese Umstände führen zu einer Ungenauigkeit in der Bestimmung des Delays einer digitalen Schaltung, welcher nur durch die Annahme des WorstCase, also maximal ungünstig abweichenden Weite zu Längen-Verhältnisses, zu begegnen ist. Die Ringinverterkette ist auf möglichst kleines geometrisches Mismatch hin dimensioniert, damit ihre Ausgangsfrequenz nicht zufällig abweicht. Es verbleibt also, trotz indirekter Messung der Temperatur und der physikalischen Prozessparameter durch Auswertung der Ringinverterfrequenz, eine Ungewissheit in der Korrelation mit der Chipperformance, hervorgerufen durch das geometrische Mismatch. Diese Ungewissheit könnte nur durch Beobachtung der Schaltung und Auswertung der realen Performance eliminiert werden. Durch die Ungewissheit in der Evaluierung des kritischen Pfads müsste diese Beobachtung aber mehr als den per Simulation evaluierten kritischen Pfad erfassen, was ein erheblicher Aufwand wäre.

Das vorgestellte System muss unter der Annahme von WorstCase geometrischem Mismatch vom Designer kalibriert werden (also das ROM-Speicherzellenfeld entsprechend belegt, bzw. die digitale Vorbelegung entsprechend gestaltet werden). Dadurch wird automatisch bei der Mehrzahl der Chips Energie verschwendet, da WorstCase Parameter nur bei einem sehr kleinen Prozentsatz aller hergestellten Chips vorzufinden sind. Abbildung 3.32 macht dies deutlich.

Die Spezifikationen ( $delay \leq 368ps$ ) wären auch mit deutlich reduzierter Versorgungsspannung zu erreichen, durch das unabhängige geometrische Mismatch ist aber der Sicherheitsabstand notwendig. Dieser Sicherheitsoverhead legt Überlegungen zum Einsparpotential des Systems nahe.

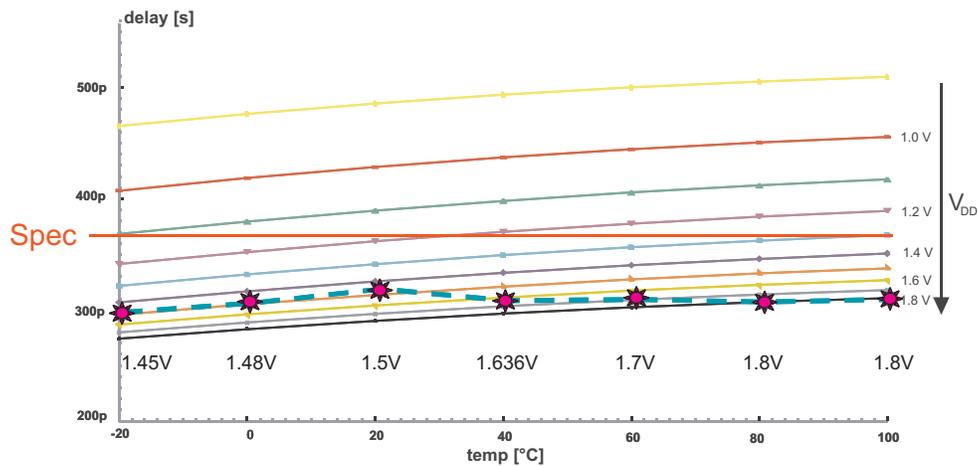


Abbildung 3.32: Delay der Gatterkette aus Abbildung 3.31 mit Best-Case Parametern mit den für Worst-Case Mismatch bestimmten Versorgungsspannungen

Die Grafiken 3.33 bis 3.35 geben einen Überblick über das Einsparpotential des Systems, basierend auf der Gatterkette, die in den bisherigen Simulationen in diesem Kapitel zur Anwendung kam.

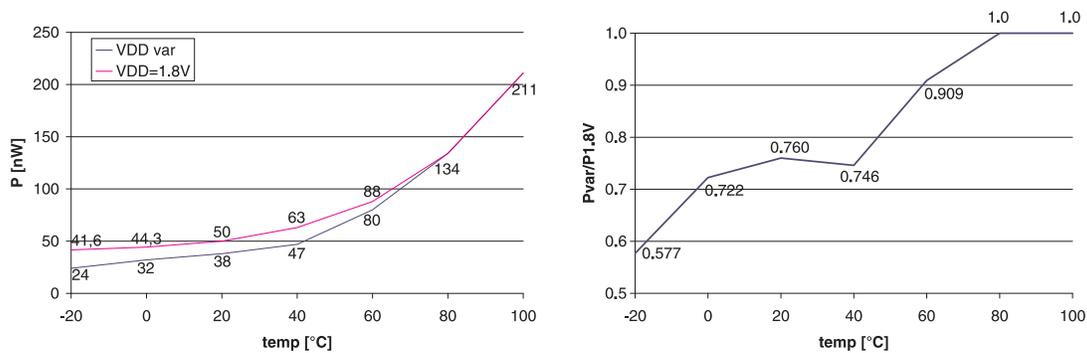
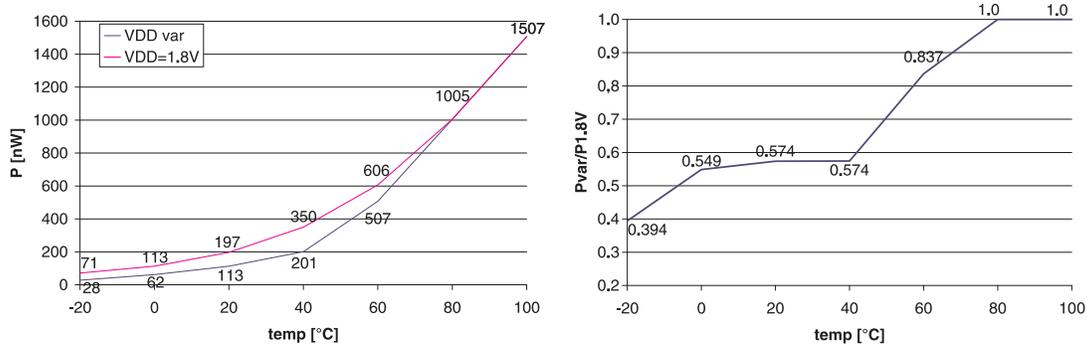
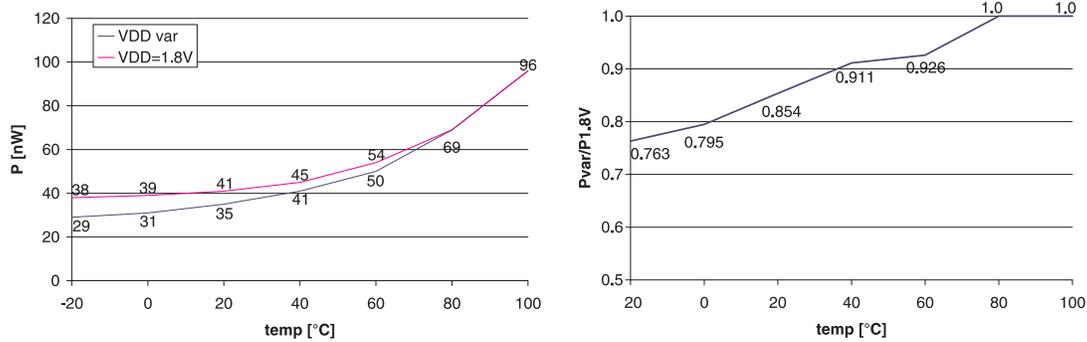


Abbildung 3.33: Verlustleistung einer Gatterkette mit fester und variabler Versorgungsspannung im typischen Fall (Verhältnis rechts)

Obwohl die Versorgungsspannungen aufgrund angenommener WorstCase Parameter bestimmt worden sind, wird auch bei anderen Parametersätzen viel Lei-



**Abbildung 3.34:** Verlustleistung einer Gatterkette mit fester und variabler Versorgungsspannung bei Best-Case-Mismatch (Verhältnis rechts)



**Abbildung 3.35:** Verlustleistung einer Gatterkette mit fester und variabler Versorgungsspannung bei WorstCase-Parametersatz (Verhältnis rechts)

stung eingespart. Ein genauer Wert für diese Einsparung kann nicht angegeben werden, bei einer Gausverteilung der Parametervariation kann von einer durchschnittlichen Einsparung von  $\approx 20\%$  gegenüber der gleichen Charge identischer Schaltungen mit fester Versorgungsspannung ausgegangen werden.

Folgende Überlegungen liegen dieser Abschätzung zugrunde. Das System liefert eine Aussage über die globalen Prozessparameter und die Temperatur, also neben der Dimensionierung der Transistoren die wesentlichen Faktoren für die Schaltgeschwindigkeit.

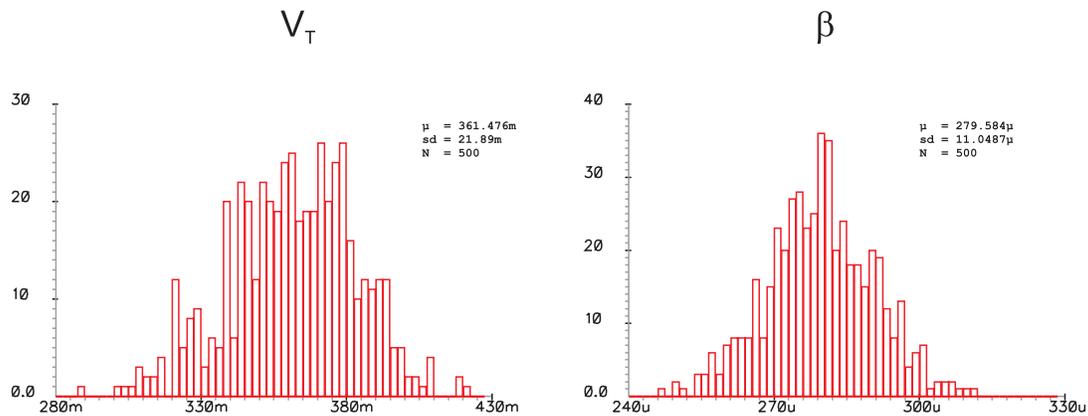


Abbildung 3.36: Mismatch eines minimalen NMOS-Transistors

Abbildung 3.36 zeigt das Mismatch der Parameter  $V_T$  und  $\beta$  bei geometrischer Varianz ( $W$ ,  $L$  und  $t_{ox}$ ) für einen minimal dimensionierten NMOS-Transistor. Es ist eine große Spreizung der Parameter ersichtlich, welche durch die minimalen Transistordimensionen bedingt ist.

Wird die Überschlagsformel (Gleichung 2.3) zur Berechnung des Delays eines Inverters (stellvertretend für CMOS-Gatter) mit den statistischen Daten aus Abb. 3.36 beaufschlagt, ergibt sich für das Delay folgende Verteilung mit für das  $W/L$ -Verhältnis typischen normalverteilten Varianzen ( $W:\mu = 0.3\mu m$ ,  $\sigma = 20nm$  und  $L:\mu = 0.1675\mu m$ ,  $\sigma = 12.5nm$ ):

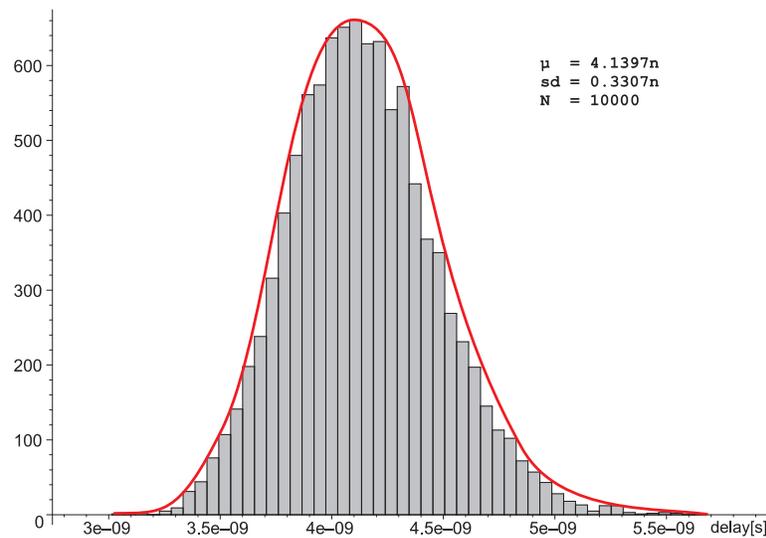
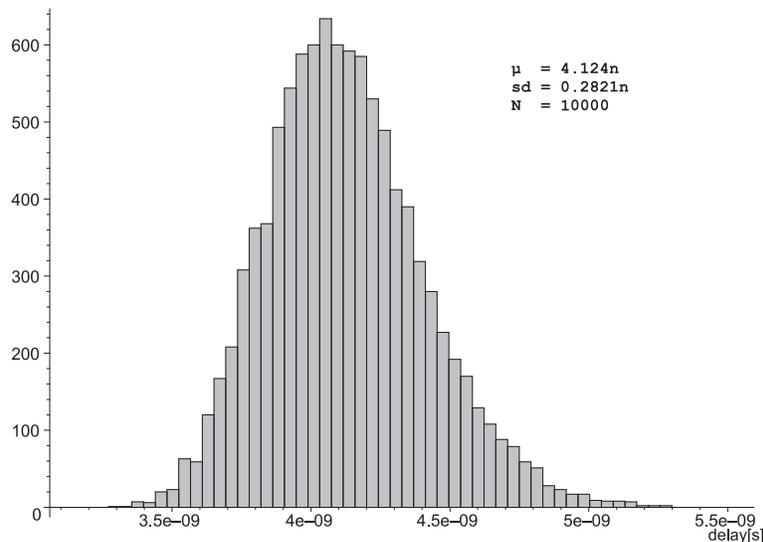


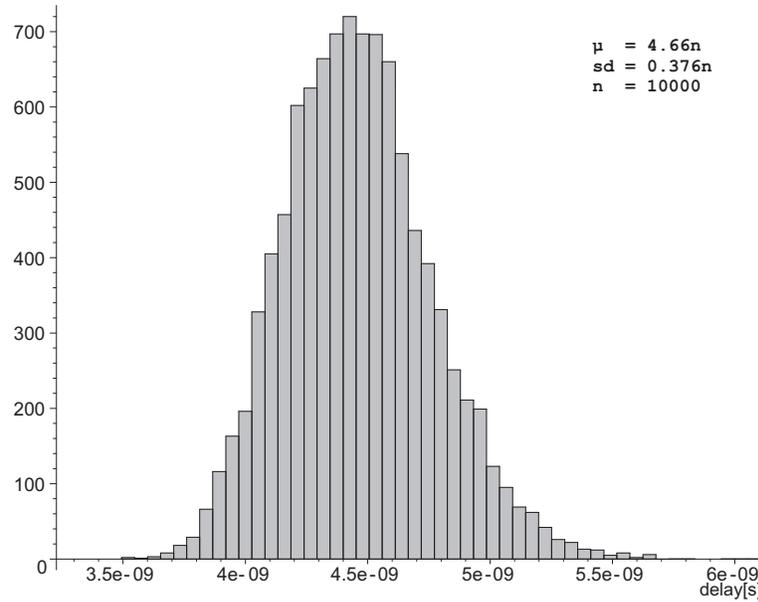
Abbildung 3.37: Varianz der falltime eines CMOS-Inverters

Es stellt sich eine Normalverteilung (10000 Werte pro Variable) des delays ein (rote Linie in Abb. 3.37) mit dem Mittelwert  $\mu = 4.1397ns$  und der Standardabweichung  $\sigma = 0.3307ns$ . Der Inverter ist hierbei mit  $C_L = 1pF$  belastet, die Versorgungsspannung hat den Standardwert von  $V_{DD} = 1.8V$ . Der Designer kann hier also für eine sehr große Wahrscheinlichkeit ( $P_{failure} = 0.1\%$ ) ein Gatterdelay von  $t_f = 5.65ns$  garantieren. Durch die Information, die die Auswertung der RIC-Frequenz dem Designer gibt, ist der Einfluss der globalen Parameter  $\beta_0$  und  $V_{T0}$  nicht mehr zufällig. Entsprechend ändert sich die Verteilung des delays, wenn diesen Parametern feste Werte zugewiesen werden:



**Abbildung 3.38:** Varianz der falltime eines CMOS-Inverters  $V_T, \beta = normal$

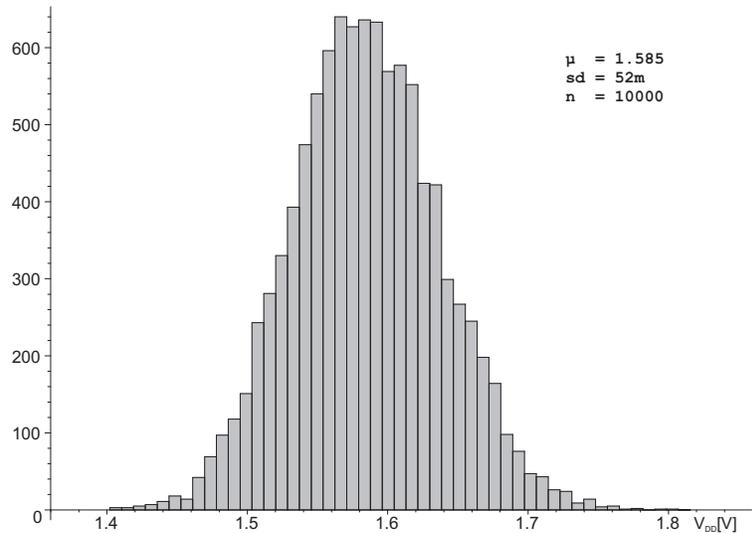
In Abbildung 3.38 sind  $V_T$ , und  $\beta$  typische Werte zugewiesen. Durch das Fehlen der "schlechten" Parametervarianz für diese Variablen verbessert sich der worst-case für das delay auf  $t_f = 5.3ns$  (Mittelwert  $\mu = 4.124ns$ , Standardabweichung  $\sigma = 0.2821ns$ ). Signifikant ist die reduzierte Standardabweichung. Der Designer kann also mit Kenntnis der Parameter durch die RIC-Frequenz die Versorgungsspannung verringern, um die gleiche Verzögerung zu garantieren, wie ohne Kenntnis der Parameter  $V_T$ , und  $\beta$  ( $t_{ox}$  variiert ebenso global und wirkt sich auf die RIC-Frequenz aus). In diesem Fall kann die Versorgungsspannung von  $V_{DD} = 1.8V$  auf  $V_{DD} = 1.65V$  reduziert werden, was Abbildung 3.39 zeigt.



**Abbildung 3.39:** Varianz der falltime eines CMOS-Inverters  $V_T, \beta = normal$ ,  $V_{DD} = 1.65V$

Durch die Zusatzinformation, die die RIC-Frequenz dem Designer liefert, wird die Variation der Chip-Performance zwar nicht grundsätzlich kleiner, der Einfluss des systematischen Mismatches und der Temperatur ist aber nun kalkulierbar. Musste der Designer bislang den Worst-Case aller streuenden Parameter annehmen, ist das nun nur noch für das geometrische Mismatch erforderlich. Die zu erwartende Einsparung durch Versorgungsspannungs-Reduktion errechnet sich durch die Annahme normalverteilter Parameter für  $V_T$ ,  $\beta$  und  $t_{ox}$ . Ausgehend von Gleichung 2.3 wird für absolute WorstCase Parameter aller Variablen die maximale Verzögerung bestimmt. Dies ist der Wert, den der Designer unter allen Umständen garantieren kann. Ausgehend von diesem Wert und der statistischen Verteilung der globalen Chipparameter kann nun die jeweilige Versorgungsspannung errechnet werden, die für jede Parametervariation zur gleichen Verzögerung führt. Abbildung 3.40 zeigt die Verteilung dieser errechneten Versorgungsspannung für WorstCase Transistorgeometrie und der Verzögerung von  $t_f = 5.65ns$  abgeleitet aus Abbildung 3.38. Bezogen auf den Mittelwert von  $\sigma = 1.585V$  und die dominante quadratische Abhängigkeit des Leistungsumsatzes von der Versorgungsspannung berechnet sich die durchschnittliche Einsparung zur Versorgung mit der Standardversorgungsspannung  $V_{DD-norm} = 1.8V$  zu:

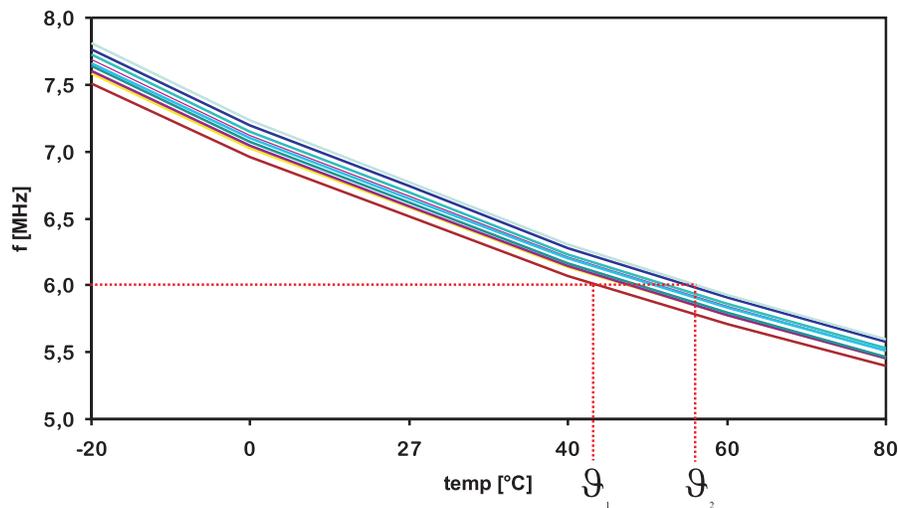
$$\frac{(1.585V)^2}{1.8V^2} = 0.77 \equiv 23\% \quad (3.15)$$



**Abbildung 3.40:** Versorgungsspannung für konstantes Delay bei WorstCase Transistorgeometrie

Dieser Wert beruht auf einer simplen Modellierung des Delays und einer Chiptemperatur von  $27^{\circ}\text{C}$ . Für die Angabe der garantierten Performance der Schaltung müssen aber auch für die Chiptemperatur WorstCase-Parameter angenommen werden, also maximale Betriebstemperatur. In diesem Falle sinkt die garantierte Verzögerung. Das vorgestellte System erfasst aber auch die Temperatur, so dass die Versorgungsspannung im Mittel (mit variierender Temperatur) ungefähr gleich bleibt. Ein hier nicht mit einkalkulierter zusätzlich positiver Effekt ist die durch die verringerte Versorgungsspannung kleinere Erwärmung des Chips. Diese Effekt trägt nochmals zu einem geringeren Energieumsatz bei, da der schlechteren Leitfähigkeit der MOS-Transistoren zu höheren Temperaturen hin nicht unnötig hoch gegengesteuert werden muss.

Eine weitere Einschränkung ist bei der Verwendung des Systems zu beachten, dessen Auswirkungen allerdings im Vergleich mit dem durch unvermeidliches geometrisches Mismatch hervorgerufenen Sicherheitsoverhead gering ist. Eine bestimmte Frequenz der Ringinverterkette kann aufgrund unterschiedlicher Kombinationen von Chiptemperatur und physikalischen Chipparametern auftreten. Abbildung 3.41 zeigt dies exemplarisch für eine Schwingfrequenz von  $f_{RIC} = 6MHz$ . Die Schwingfrequenz von 6MHz kann je nach Parametersatz zwischen  $\vartheta_1 = 45^\circ C$  im Falle von WorstCase-Parametern und  $\vartheta_2 = 55^\circ C$  im Falle von BestCase-Parametern auftreten.



**Abbildung 3.41:** Schwingfrequenz einer Ringinverterkette abhängig von der Temperatur und dem Parametersatz

Es ist von großer Bedeutung, dass die Korrelation von Ringinverterschwingfrequenz und Gatterdelay (also Performance) möglichst eindeutig ist. Um dieses zu untersuchen, werden mit einer Gatterkette Simulationen durchgeführt, die abhängig von der Temperatur, der Versorgungsspannung und dem Parametersatz das Gatterdelay evaluieren. Somit werden jeweils die maximale und die minimale Versorgungsspannung bestimmt, die zum Erreichen eines bestimmten Gatterdelays notwendig ist (Spezifikation). Die maximale und minimale Versorgungsspannung sollten nur wenig divergieren, damit die Korrelation zwischen Ringinverterschwingfrequenz und daraus abgeleiteter Versorgungsspannung möglichst präzise ist. Die Abbildung 3.42 zeigt das Delay der Gatterkette für WorstCase- und BestCase-Parameter (ohne Abweichung der Transistordimensionen, die nicht systematisch sind) in Abhängigkeit der Versorgungsspannung und mit der Temperatur als Scharvariable. Bei einer angenommenen Spezifikation von 200ps liegt der Unterschied der beiden evaluierten Versorgungsspannungen bei ca. 20mV.

Je nach Spezifikation kann diese Divergenz variieren, liegt jedoch im Beispiel aus Abbildung 3.42 nie höher als  $30\text{mV}$ . Der Designer muss diese Divergenz durch Simulation evaluieren und bei der Kalibrierung des Systems berücksichtigen.

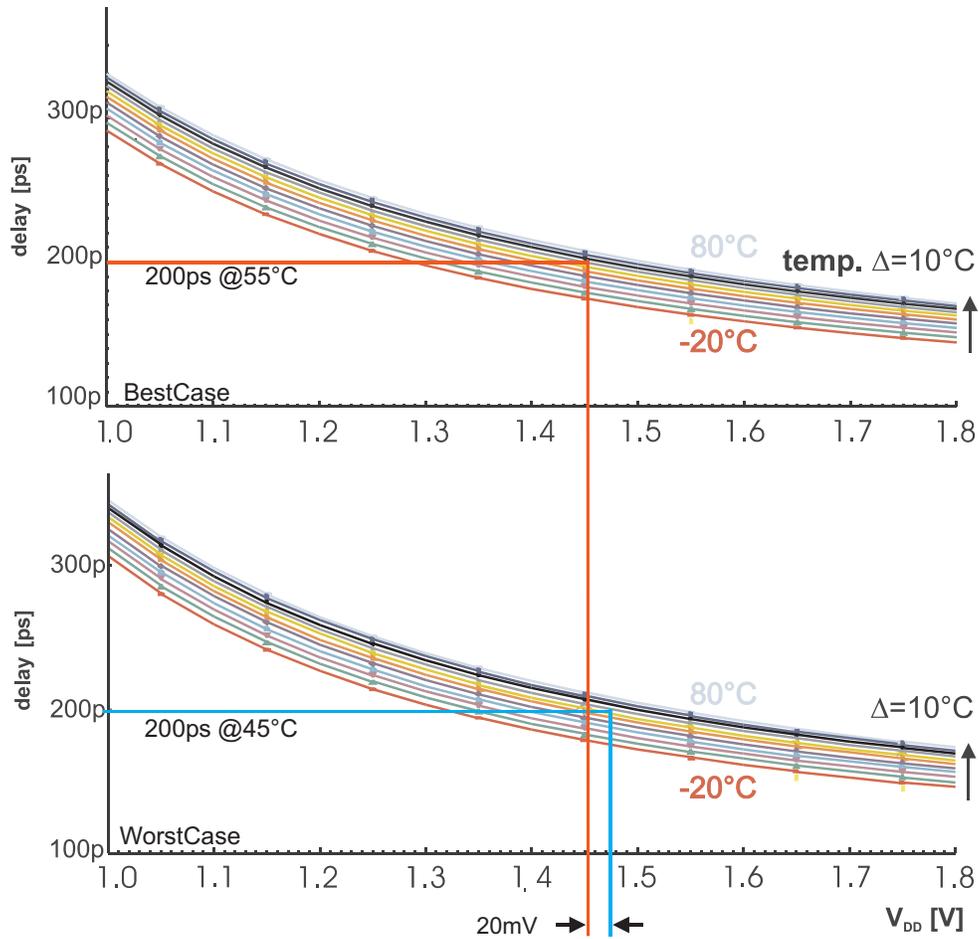


Abbildung 3.42: Gatterdelay abhängig von der Versorgungsspannung, der Temperatur und dem Parametersatz

## 3.14 Erweiterung zum Performance-Driven-Supply-System

### 3.14.1 Digital Analog Konverter mit optimiertem Mismatchverhalten

Um den Zählerstand in eine Referenzspannung zu wandeln, wird ein Digital-Analog-Wandler implementiert. Hierbei wird besonderen Wert auf einfache Implementierung und gutes Mismatch- bzw. Temperaturverhalten gelegt.

Durch den besonderen Einsatzzweck haben die Geschwindigkeit des DACs, die Treiberfähigkeit des Ausgangs, die absolute Genauigkeit und PSRR verminderte Priorität. Die besonderen Spezifikationen führen zu einem Konzept mit Widerstandsspannungsteilern, die über MOS-Schalter angewählt werden. Die Abbildung 3.43 zeigt das Schaltbild des implementierten 4bit-Digital-Analog-Wandlers.

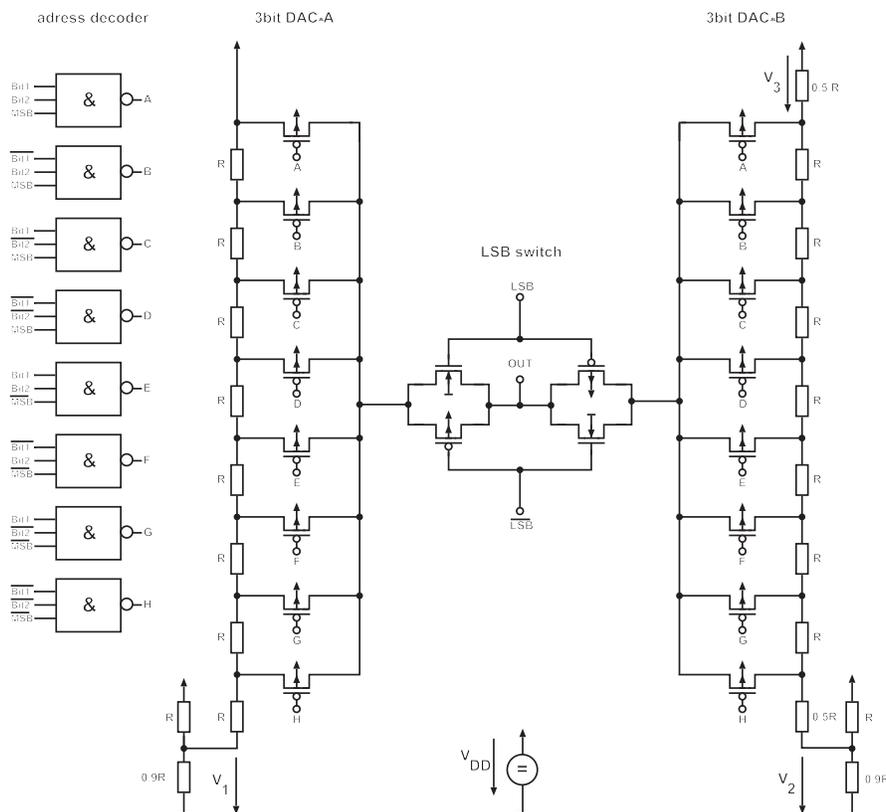
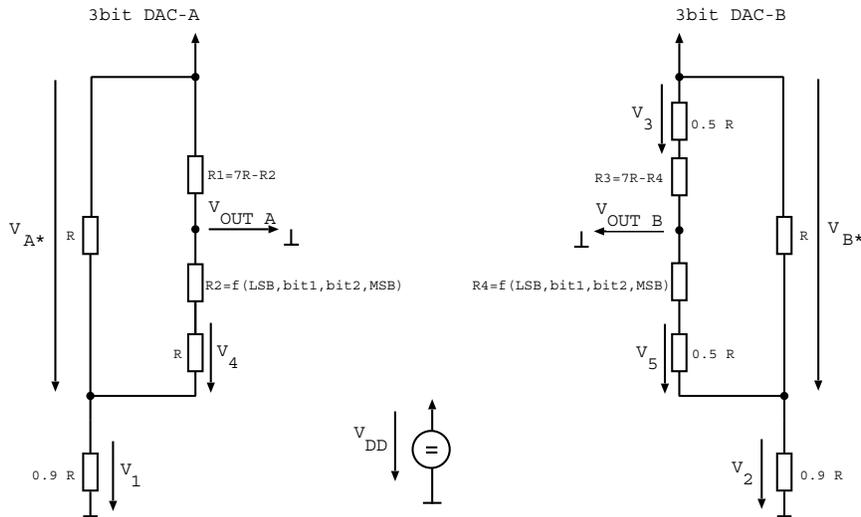


Abbildung 3.43: 4bit Digital-Analog Konverter

Der DAC besteht aus zwei seriellen Widerstandsreihen mit jeweils acht Abgriffen, die über PMOS-Schalter bzw. Transmissiongates mit dem Ausgangsknoten verbunden werden. Somit stellt jede Widerstandsreihe einen 3bit DAC dar (Abbildung 3.43 3bit DAC-A und 3bit DAC-B). Die Widerstände sind überwiegend identisch. Durch die beiden Spannungsteiler, bestehend aus den Widerständen  $R$  und  $0.9R$  in Abbildung 3.43, ergeben sich, abhängig von dem digitalen Eingangswert, folgende Ersatzschaltungen (Abbildung 3.44):



**Abbildung 3.44:** 3bit DAC Ersatzschaltbilder gemäß Abbildung 3.43

Es handelt sich hierbei um zwei Spannungsteiler mit variablem Abgriffpunkt. Durch die beiden Transmissiongates um den LSB switch in Abbildung 3.43 kann dem Ausgangsknoten *out* entweder  $V_{OUTA}$  oder  $V_{OUTB}$  zugewiesen werden.

Es gilt hier

$$R1 + R2 = 7R = R3 + R4 \quad (3.16)$$

, wodurch für die Teilspannungen  $V_{A*}$  und  $V_{B*}$  folgt:

$$V_{A*} = V_{B*} = V_{DD} - V_{DD} \cdot \frac{0.9R}{0.9R + \frac{8R \cdot R}{R+8R}} = V_{DD} - V_{DD} \cdot 0.5031 \approx \frac{V_{DD}}{2} \quad (3.17)$$

Für die Spannung  $V_{OUTA}$  gilt:

$$V_{OUTA} = V_{A*} \cdot \frac{R + R2}{8R} + V_1 \quad (3.18)$$

und für  $V_{OUTB}$ :

$$V_{OUTB} = V_{B*} \cdot \frac{0.5R + R_4}{8R} + V_2 \quad (3.19)$$

mit:

$$V_1 = V_2 = V_{DD} \cdot \frac{0.9R}{0.9R + \frac{8R \cdot R}{R+8R}} = V_{DD} \cdot 0.5031 \quad (3.20)$$

$R_2$  und  $R_4$  sind die resultierenden Widerstände als Serienschaltung von  $nR$ , wobei  $n$  abhängig von der jeweiligen 3-bit Adresse ist (Auswahl des jeweiligen Schalttransistors). Es gilt also:

$$R_2 = (n \cdot R, n = 0..7) \quad (3.21)$$

und

$$R_4 = (n \cdot R, n = 1..8) \quad (3.22)$$

Für die Grenzspannungen der jeweiligen Ausgänge ergibt sich:

$$V_{OUTA_{max}} = V_{DD} \quad (3.23)$$

$$V_{OUTB_{max}} = \frac{7.5}{8} \cdot V_{B*} + V_{DD} - V_{B*} = V_{DD} - \frac{0.5}{8} \cdot V_{B*} \approx \frac{31}{32} \cdot V_{DD} \quad (3.24)$$

und:

$$V_{OUTA_{min}} = \frac{1}{8} \cdot V_{A*} + V_{DD} - V_{A*} = V_{DD} - \frac{7}{8} \cdot V_{A*} \approx \frac{18}{32} \cdot V_{DD} \quad (3.25)$$

$$V_{OUTB_{min}} = \frac{0.5}{8} \cdot V_{B*} + V_{DD} - V_{B*} = V_{DD} - \frac{7.5}{8} \cdot V_{B*} \approx \frac{17}{32} \cdot V_{DD} \quad (3.26)$$

Es handelt sich hierbei also um zwei 3bit-digital-analog Wandler mit zueinander um die halbe Auflösung ( $1/2$  LSB) verschobenen Kennlinien mit einem DC-Offset von  $\approx V_{DD}/2$ . Das LSB der Eingangsdigitalwerte entscheidet, ob entweder DAC-A ( $LSB = 1$ ) oder DAC-B ( $LSB = 0$ ) ausgewählt wird. Somit wird eine Auflösung von 4bit erreicht. Die gleiche Auflösung könnte auch mit einer einzigen Widerstandsreihe, bestehend aus 16 identischen Widerständen, erreicht werden. Der Vorteil des vorgestellten Konzepts ist, dass bei gleichem Querstrom, also auch gleiche Treiberfähigkeit und gleichem SNR, die Widerstände hier doppelt so groß gewählt werden können. Somit reduziert sich der Fehler

durch Widerstandsmismatch nochmals. Realistisch betrachtet ist dies jedoch ein eher marginaler Vorteil. Ein weiterer Vorteil besteht in der Adressdekodierung, die hier nur aus acht *NAND*-Gattern mit drei Eingängen, vier Invertern und zwei Transmissionsgates besteht. Bei einem alternativen Widerstands-DAC mit einer Widerstandsreihe aus 16 Widerständen wäre der Aufwand für die Adressdekodierung wesentlich größer (16 *NAND* mit vier Eingängen).

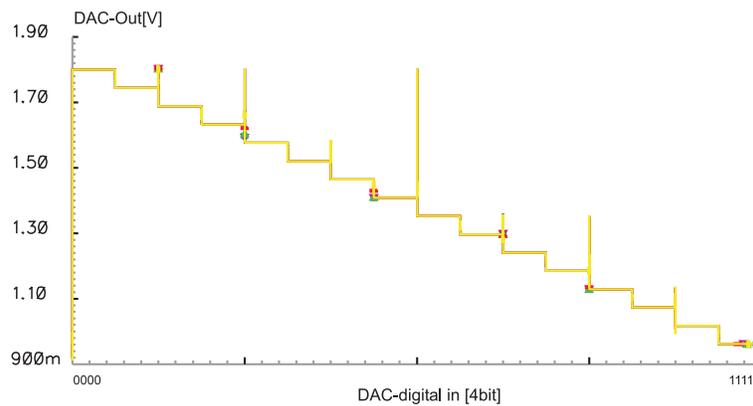
Die Adressdekodierung erfolgt in dem Sinne, dass für hohe Zählerstände niedrige Ausgangsspannungen generiert werden. Tabelle 3.6 verdeutlicht dies für alle möglichen Zählerstände.

MSB-Bit2-Bit1-LSB	A	B	C	D	E	F	G	H	DAC	$V_{out}$
0000	1	1	1	1	1	1	1	0	DAC-B	$17/32 \cdot V_{DD}$
0001	1	1	1	1	1	1	1	0	DAC-A	$18/32 \cdot V_{DD}$
0010	1	1	1	1	1	1	0	1	DAC-B	$19/32 \cdot V_{DD}$
0011	1	1	1	1	1	1	0	1	DAC-A	$20/32 \cdot V_{DD}$
0100	1	1	1	1	1	0	1	1	DAC-B	$21/32 \cdot V_{DD}$
0101	1	1	1	1	1	0	1	1	DAC-A	$22/32 \cdot V_{DD}$
0110	1	1	1	1	0	1	1	1	DAC-B	$23/32 \cdot V_{DD}$
0111	1	1	1	1	0	1	1	1	DAC-A	$24/32 \cdot V_{DD}$
1000	1	1	1	0	1	1	1	1	DAC-B	$25/32 \cdot V_{DD}$
1001	1	1	1	0	1	1	1	1	DAC-A	$26/32 \cdot V_{DD}$
1010	1	1	0	1	1	1	1	1	DAC-B	$27/32 \cdot V_{DD}$
1011	1	1	0	1	1	1	1	1	DAC-A	$28/32 \cdot V_{DD}$
1100	1	0	1	1	1	1	1	1	DAC-B	$29/32 \cdot V_{DD}$
1101	1	0	1	1	1	1	1	1	DAC-A	$30/32 \cdot V_{DD}$
1110	0	1	1	1	1	1	1	1	DAC-B	$31/32 \cdot V_{DD}$
1111	0	1	1	1	1	1	1	1	DAC-A	$V_{DD}$

**Tabelle 3.6:** Adressdekodierung des 4bit DAC aus 3.43

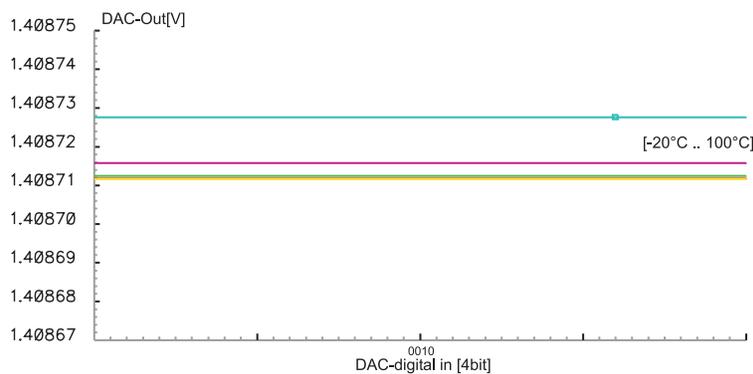
Durch die Verwendung von Spannungsteilern, deren Werte sich jeweils auf den Quotienten von Widerstandswerten beziehen, sind Abweichungen in den realen Widerstandswerten weniger einflussreich. Darüber hinaus sind die Widerstände in der Mehrzahl identisch bzw. lassen sich aus Parallelschaltung identischer Widerstände realisieren. Die Abweichung der Widerstände von dem idealen Wert sind bei der Prozessierung sehr ähnlich, so dass diese sich bei der Quotientenbildung ausmitteln. Unter der realistischen Annahme, dass alle Widerstände gleiche Temperatur besitzen, wirkt sich auch die Temperaturdrift der Widerstandswerte bei der Quotientenbildung wenig aus. Der Einfluss der Schalttransistoren ist minimal, da über sie nur im Umschaltzeitpunkt Strom fließt (Laden eines Transistorgates). Somit fällt keine Spannung über ihnen ab.

Die folgenden Abbildungen bestätigen das gute Mismatchverhalten der Schaltung. Abbildung 3.45 zeigt die Ausgangsspannung des DACs nach Abbildung 3.43 für alle Eingangskombinationen mit der Temperatur als Scharvariable ( $-20^{\circ}\text{C} \leq temp \leq 100^{\circ}\text{C}$ ). Die Peaks beim Wechsel der Adressierung in Abbildung 3.45 sind bedingt durch kurzzeitig überlappende Eingangspiegel der einzelnen Bits (durch nicht unendliche Flankensteilheit). Sie treten in der Realität auch auf, sind aber für den Verwendungszweck unerheblich, da als Referenzspannung die stationäre Ausgangsspannung des DAC verwendet wird, weswegen keine schaltungstechnischen Maßnahmen zu ihrer Unterdrückung unternommen werden.



**Abbildung 3.45:** Ausgangsspannung des DACs nach Abbildung 3.43 bei verschiedenen Spannungen in Abhängigkeit der Temperatur

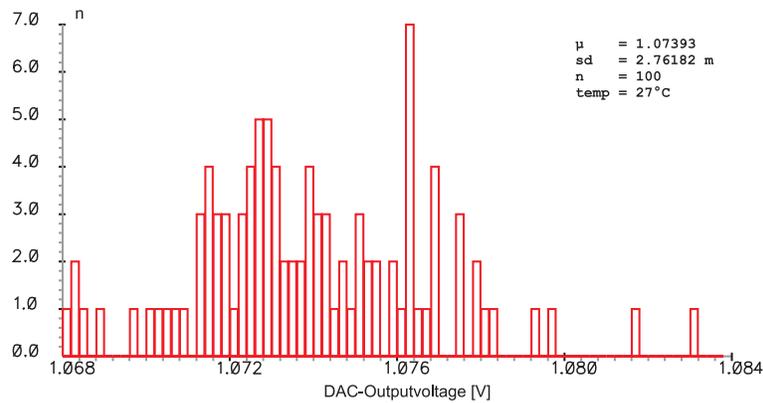
Abbildung 3.46 macht das sehr gute Temperaturverhalten deutlicher. Hier wird zur besseren Kenntlichkeit ein Ausgangswert der Abbildung 3.45 vergrößert dargestellt. Die Abweichung liegt hierbei unter einem Promille.



**Abbildung 3.46:** Vergrößerung der Darstellung aus Abbildung 3.45

Die Simulationsergebnisse, die in den Abbildungen 3.45 und 3.46 dargestellt sind, gelten für typische Prozessparameter. Simulationen mit maximal abweichenden Prozessparametern zeigen nur marginale Unterschiede.

Die Abbildung 3.47 macht den sehr kleinen Einfluss des geometrischen Mismatches deutlich. Hier wird das Histogramm einer MonteCarlo-Simulation mit 100 Durchläufen gezeigt, bei dem das geometrische Mismatch variiert wird.



**Abbildung 3.47:** Statistische Verteilung der Ausgangsspannung des DACs aus Abbildung 3.43

### 3.15 Gesamtsystem und Testchip

Die Erfassung der RIC-Frequenz als Parameter einer Performance-spezifischen Versorgungsspannungseinstellung ist vergleichsweise trivial. Durch das Zählen der fallenden Flanken der Schwingung des RIC ist das Ergebnis der Evaluierung unabhängig von Nichtsymmetrien der high- und low-time der RIC-Schwingung. Die Periode, während derer die RIC-Frequenz gezählt wird, wird durch die Referenzzeit (siehe Kapitel 3.12.1) bestimmt. Die Abbildung 3.48 zeigt einen Überblick über das Konzept des Gesamtsystems, welches in einem Testchip (siehe Anhang A.24 und [51]) umgesetzt wurde.

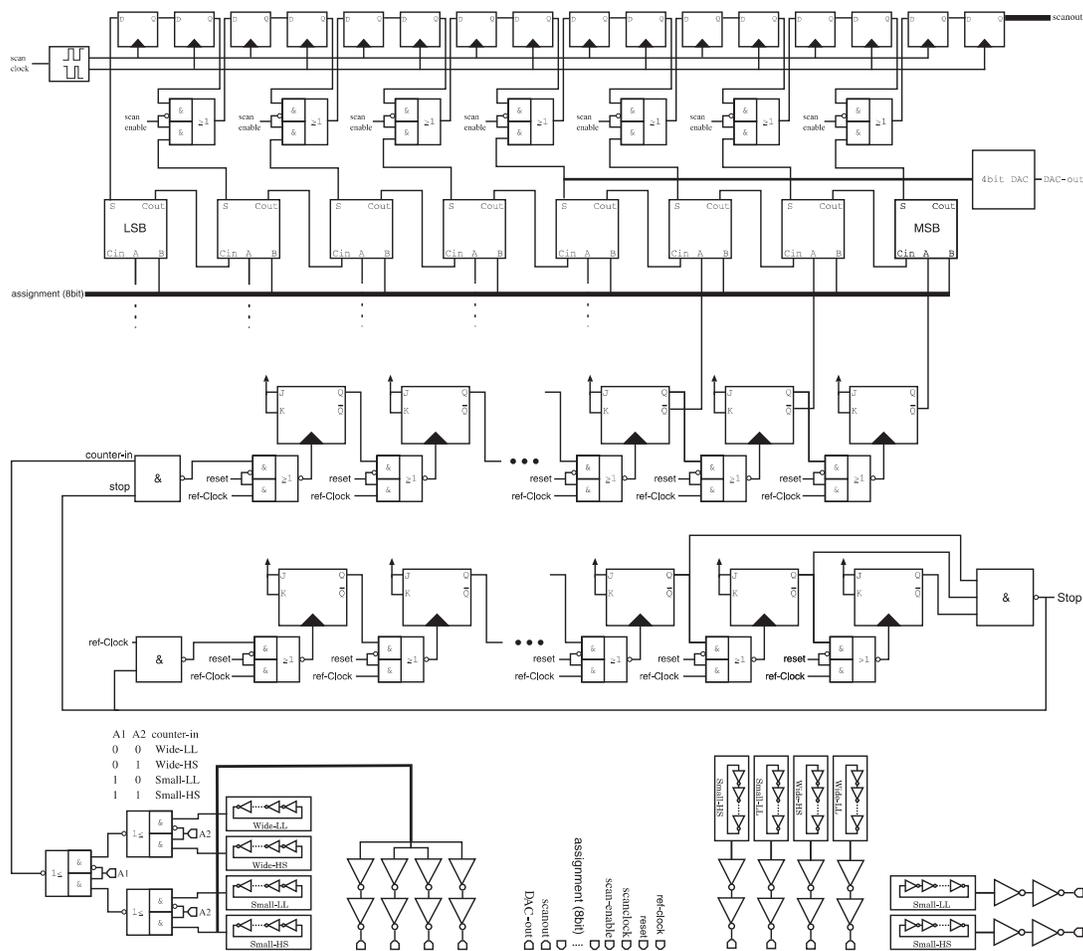


Abbildung 3.48: Systemüberblick

Bei der Fertigung nur weniger Chips (in diesem Fall 50 Dies) kann davon ausgegangen werden, dass alle Chips von einem Wafer stammen. Somit sind global veränderliche Parameter ( $V_{T0}$ ,  $\beta_0$ ) für alle Chips gleich. Dem lokalen geometrischen Mismatch unterliegen Chips eines Wafers jedoch trotzdem. Um Variationen in der Thresholdspannung zu simulieren, wurden Ringinverterketten mit LowLeakage-Transistoren und mit HighSpeed-Transistoren implementiert. Ebenso wurden minimal dimensionierte Ringinverterketten nach Abbildung 3.12 mit LowLeakage-Transistoren und mit HighSpeed-Transistoren implementiert. Zur Untersuchung der Abhängigkeit des geometrischen Mismatches von der Ausrichtung der Aktivgebiete auf dem Substrat sind minimal dimensionierten Ringinverterketten (LowLeakage und HighSpeed) je zweimal senkrecht zueinander auf dem Substrat platziert.

Über Multiplexer können vier verschiedene Ringinverterketten zur Auswertung mit der Auswerteschaltung verbunden werden. Dabei handelt es sich jeweils um LowLeakage- und HighSpeed-Typen mit 13 groß dimensionierten bzw. 127 minimal dimensionierten Invertern (vergleiche dazu Abbildungen 3.12 und 3.13). Die Ausgänge aller 10 Ringinverterketten sind über Treiber von außen abgreifbar.

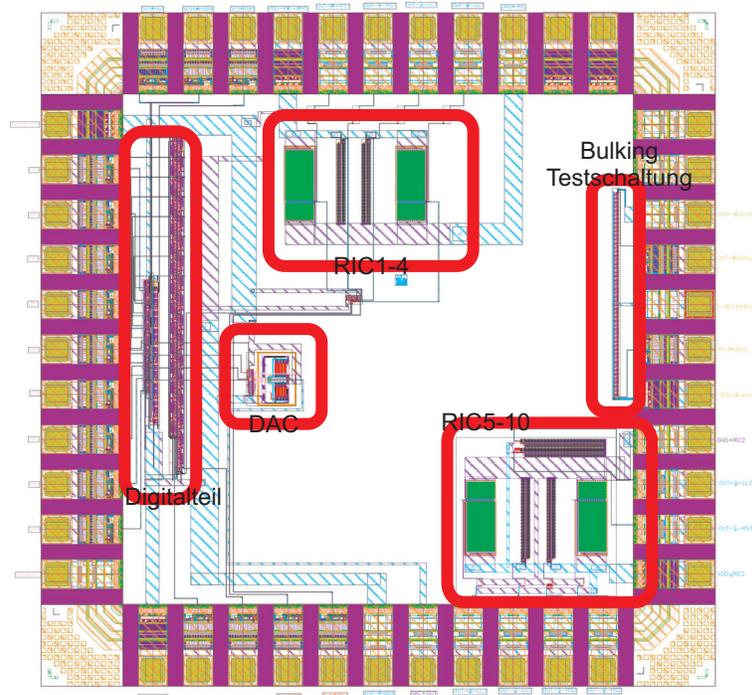


Abbildung 3.49: Layout des Testchips mit Funktionsgruppen

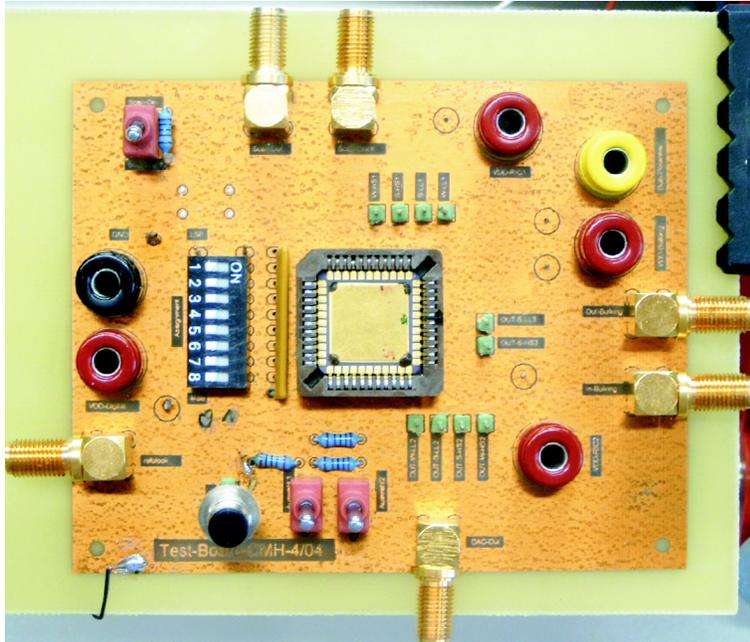


Abbildung 3.50: Testboard

Zur Ausgabe des Zählerstands ist ein Scan-Pfad implementiert, der die acht höchstwertigen Bits seriell zu einem Ausgang führt. Zur Vermeidung von transparenten Zuständen des Scan-Pfad-Schieberegisters, ist dieses mit 16 D-FlipFlops und nicht überlappendem Takt ausgestattet. Der Scan-Takt wird ebenfalls von außen an ein Pad angelegt.

Zusätzlich zu den Schaltungen zum System, das in Abbildung 3.48 dargestellt ist, ist eine Testschaltung implementiert, die Untersuchungen zum Leistungsverbrauch und der Performance digitaler Schaltungen in Abhängigkeit von dem Bulkpotential der PMOS-Transistoren (vergl. Kapitel 2.8) ermöglicht. Das Gehäuse des Testchip ist CLCC44. Abbildung 3.49 zeigt das Layout des Testchips. Ein Testboard (Abbildung 3.50) für den Anschluss des Chips wurde entworfen, das Messungen in einer Klimakammer ermöglicht. Abbildung 3.51 zeigt das Schaltbild des Testboards. Das zugehörige PCB-Layout findet sich in Abbildung A.9 im Anhang.

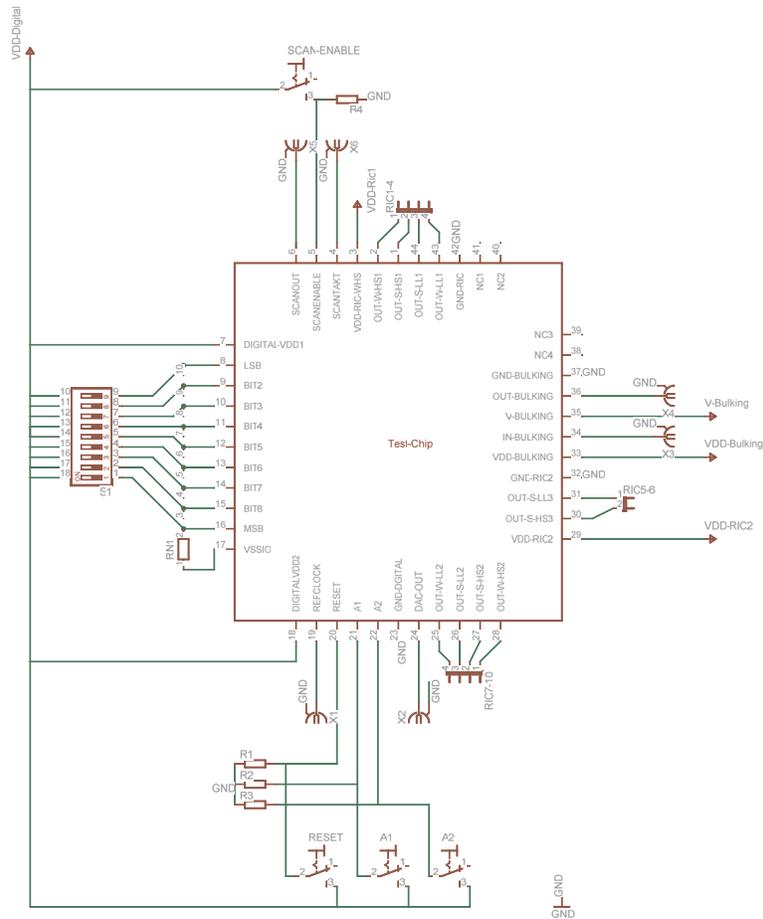
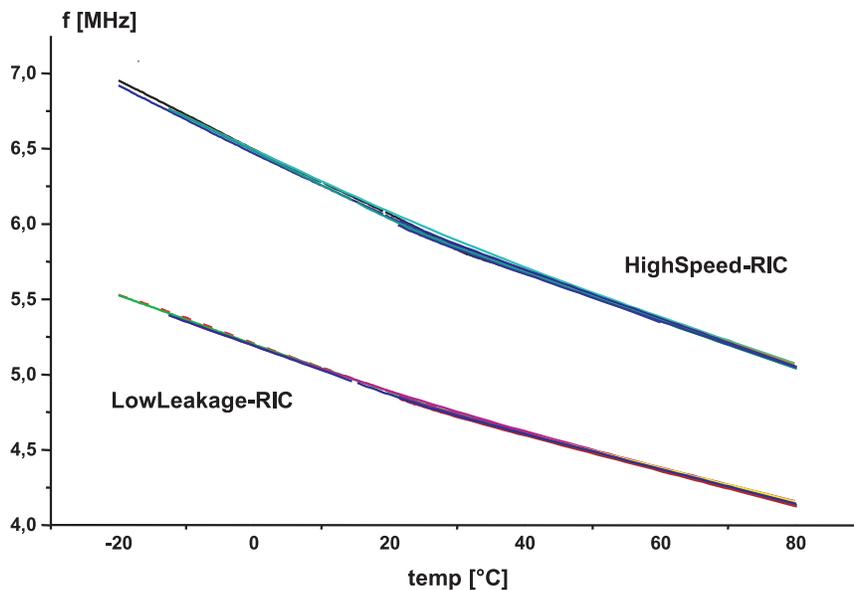


Abbildung 3.51: Testboard-Schaltplan

### 3.15.1 Messergebnisse

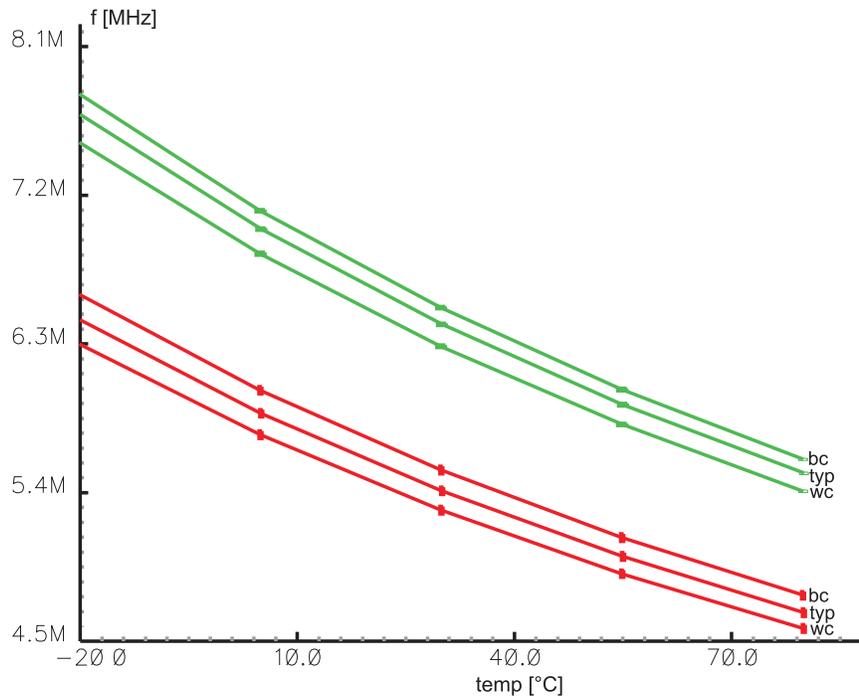
Da nur sehr wenig Zeit für das Design zur Verfügung stand, wurden einige Layoutfehler übersehen, so dass nicht alle Optionen des Konzepts aus Abbildung 3.15 funktionsfähig sind. So können die Zählerketten nicht resetet werden, was eine digitale Auswertung des Zählerstands unmöglich macht. Im Folgenden sind die Messergebnisse der möglichen Messungen dargestellt. Die grundsätzliche Funktionalität der digitalen Schaltung wird in Kapitel 3.15.2 nachgewiesen. Abbildung 3.53 zeigt die Messung der Ringinverterausgangsfrequenz in Abhängigkeit der Temperatur. Die Messungen sind an fünf gehäuteten Chips durchgeführt, mit jeweils vier Ringinverterketten (hälftig mit LowLeakage-Transistoren und hälftig mit High-Speed-Transistoren) bestehend aus 13 groß dimensionierten Invertern.



**Abbildung 3.52:** Messung der Ringinverterfrequenz in Abhängigkeit von der Temperatur

Das Messergebnis bestätigt die Simulation insofern, dass nur eine sehr kleine Varianz der Schwingfrequenzen der einzelnen RIC zu verzeichnen sind ( $\Delta f < 0.5\%$ ). Diesbezüglich gibt es keinen qualitativen Unterschied zwischen den LowLeakage-Ringinverterketten und den HighSpeed-Ringinverterketten. Der Vergleich mit der Simulation (Abbildung 3.41) zeigt zwar tendenziell ähnliche Ergebnisse, in der absoluten Größe jedoch Unterschiede. Die Extraktion der parasitären Kapazitäten der Layouts der RIC und deren Einbeziehung in die Simulation zeigen keinen nennenswerten Effekt, so dass der absolute Unterschied zwischen simulierter und gemessener Schwingfrequenz der RIC nur durch eine

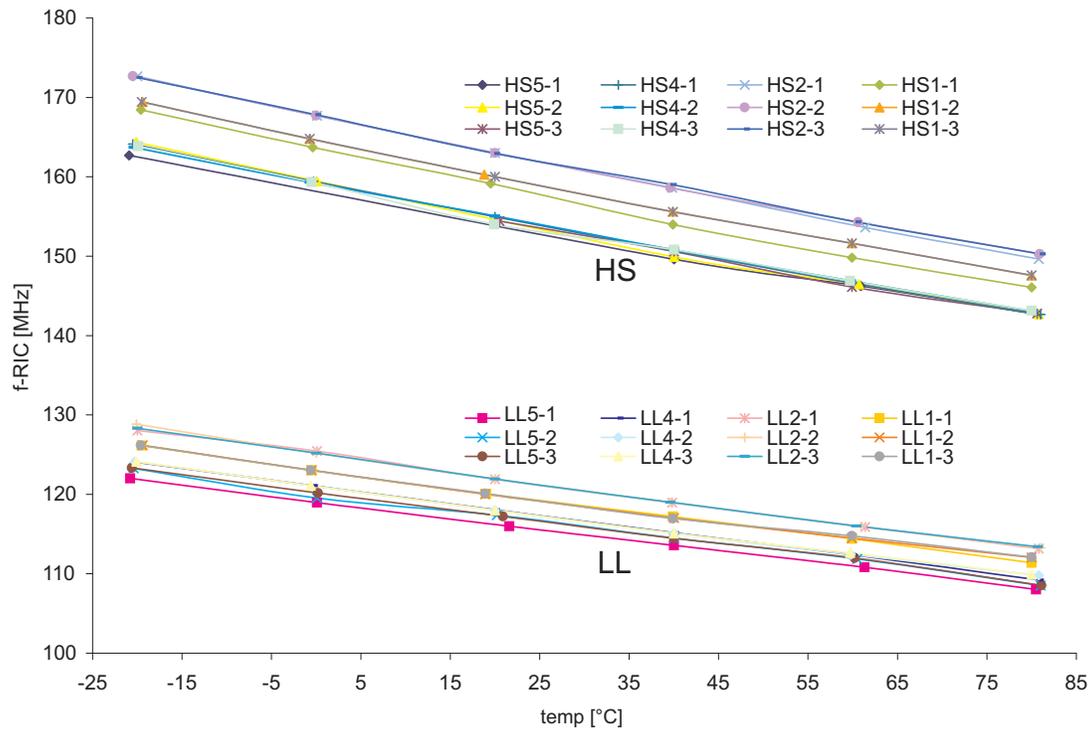
Evolution der Technologie zu erklären ist. Die verwendeten Simulationsmodelle waren bei Herstellung des Testchips vier Jahre alt, so dass diese Möglichkeit in Betracht gezogen werden muss. Eine diesbezügliche Nachfrage bei ST-Microelectronics blieb leider ohne Antwort.



**Abbildung 3.53:** Simulation der Ringinverterfrequenz in Abhängigkeit von der Temperatur analog zur Messung aus Abbildung 3.53

Zusätzlich zu den RICs mit großem Aktivgebiet wurden analog zu Kapitel 3.8.1 6 Inverterketten mit jeweils 127 minimal dimensionierten Transistoren (jeweils 3 in HS und 3 in LL) auf dem Chip integriert. Um eine zusätzliche Aussage über die Abhängigkeit des geometrischen Mismatches von der Ausrichtung des Layouts auf dem Chip zu gewinnen, sind zwei Inverterketten um  $90^\circ$  versetzt angeordnet. Abbildung 3.54 zeigt das Messergebnis für 24 RICs verteilt auf 4 Chips. Die relativ große Streuung der Schwingfrequenzen über die einzelnen Chips ist offensichtlich, was einen großen Unterschied zu der Messung aus Abbildung 3.53 darstellt. Neben Messkurven von RICs, die sehr gut übereinstimmen, gibt es auch Messkurven, sogar innerhalb desselben Chips, die weit auseinander liegen. Signifikante Aussagen zur Abhängigkeit der Schwingfrequenz von der Orientierung der Aktivgebiete sind wegen dieser Streuung nicht möglich (die Kurven in Abbildung 3.54 mit der Endung 2 sind zu jenen mit der Endung 1 oder 3 um  $90^\circ$  versetzt angeordnet). Die Messungen lassen also folgende Schlüsse zu:

- das geometrische Mismatch der Transistoren hat einen relevanten Einfluss auf die Schwingfrequenz von Ringinverterketten, deren Transistoren minimal dimensioniert sind
- Ringinverterketten mit Transistoren großer Aktivgebiete sind gegenüber geometrischem Mismatch unempfindlich
- die geometrische Orientierung der Aktivgebiete bei Ringinverterketten mit minimal dimensionierten Transistoren beeinflusst die Schwingfrequenz nicht signifikant



**Abbildung 3.54:** Messung der Frequenz minimal dimensionierter Ringinverter (n=127) in Abhängigkeit von der Temperatur



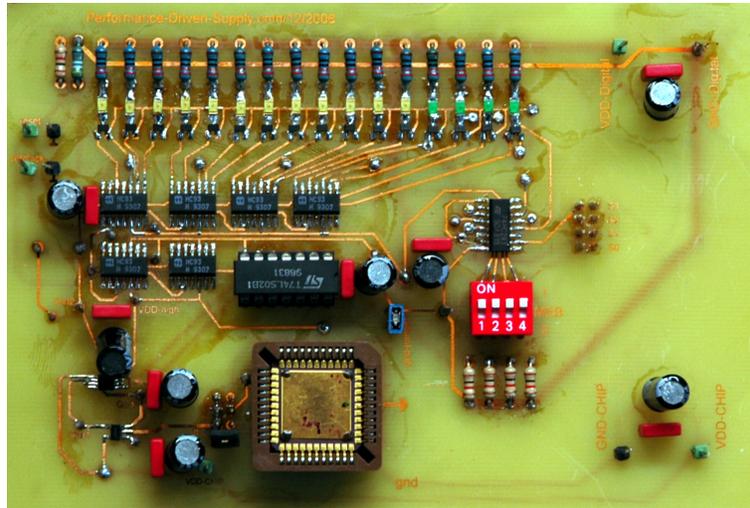


Abbildung 3.56: Platine der diskreten Umsetzung

Die Zeitspanne zwischen reset und MSB=1 des 8-Bit-Zählers definiert die Referenzzeit und ist über die Referenzclock variabel und genau einzustellen. Das MSB des Zählers sorgt, wenn das reset-Signal low ist (über die NOR-Verknüpfung U3b), für die Transparenz des NAND-Gatters U2A. Dieses NAND-Gatter buffert die per Jumper auszuwählenden Signale der verschiedenen Ringinverterketten des Testchips. Für die Dauer der Messung, die definiert ist als Periode zwischen Masterreset 1  $\rightarrow$  0 und MSB=1 des 8-Bit Zählers, ist dieses NAND Gatter für die Ringinvertersignale transparent. Sobald MSB=1 oder reset=1 ist, ist der Ausgang des NAND-Gatters=0. Die gebufferten Signale der Ringinverterkette werden über einen Level-Shifter auf TTL-Niveau angehoben und auf einen 16-Bit-Zähler geleitet. Die Funktion der Schaltung lässt sich also folgendermaßen zusammenfassen: es wird eine zur Referenzclock proportionale Referenzzeit generiert, während der die Flanken des Signals einer auszuwählenden Ringinverterkette des Testchips gezählt werden. Zur optischen Kontrolle des Zählerstands sind alle Zählerausgänge über Schalttransistoren mit LEDs verknüpft (in Abbildung 3.55 sind nur die 4 höchstwertigen Bits dargestellt). Der Zählerstand wird über den Addierer U4 von einer frei wählbaren digitalen Zahl (preselect) subtrahiert. Hierzu wird die Inverterstruktur der LED Ansteuerung ausgenutzt, da invertierte Ausgangs-Signale bei den verwendeten Flip-Flops nicht zur Verfügung stehen. Die vier höchstwertigen Bits werden invertiert zum Addierer geführt, das  $Carry_{in}$  mit 1 vorbelegt, um der Subtraktionsvorschrift

$$A - B = A + \overline{B} + 1 \quad (3.27)$$

gerecht zu werden.

Nachdem am Ausgang des Addierers ein zur Schwingfrequenz der Ringinverterkette umgekehrt proportionales digitales vier-Bit Signal anliegt, wird dieses zur Erzeugung einer analogen Referenzspannung weiterverarbeitet. Dies geschieht über eine Kabelverbindung mit einer zweiten Platine, um den analogen Schaltungsteil vom digitalen räumlich trennen zu können, was praktische Vorteile mit Hinblick auf Platzierung der Platinen in einer Temperatorkammer hat. Der Schaltbild der zweiten Platine findet sich in Abbildung 3.57.

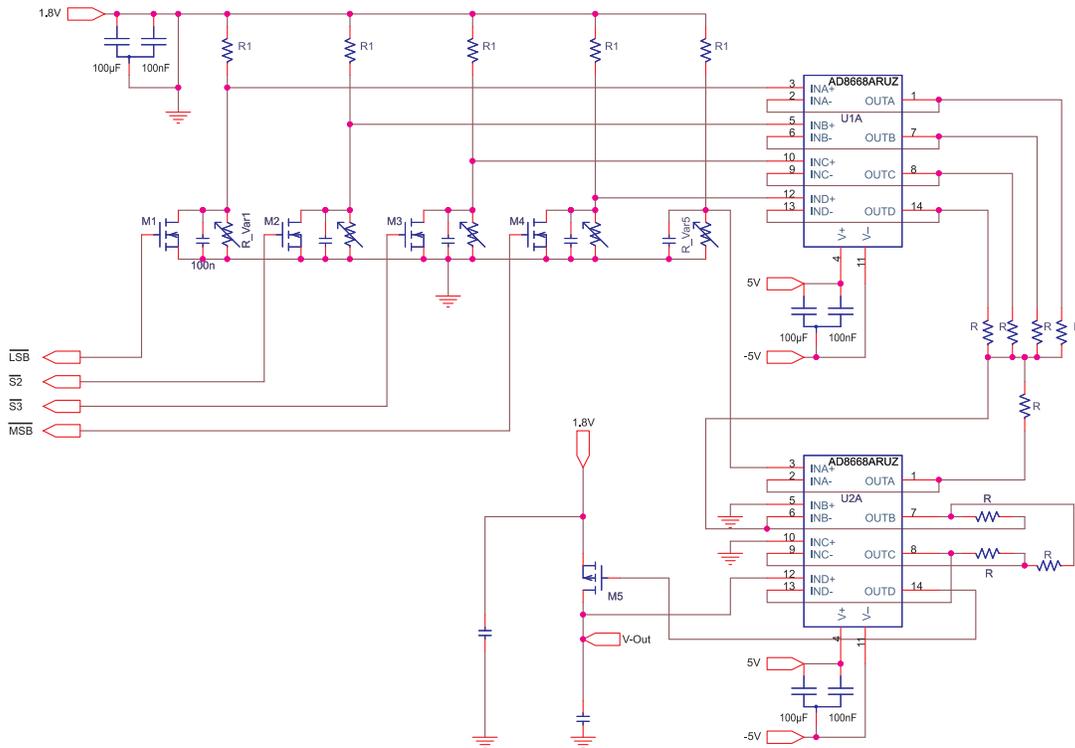


Abbildung 3.57: Schaltplan der DAC-Platine mit VDD-Erzeugung

Die Schaltung aus Abbildung 3.57 besteht aus fünf Spannungsteilern (jeweils  $R_1$  und  $R_{Var-i}$  mit  $i \in [1..5]$ ). Die Spannungen über den Potentiometern  $R_{Var-i}$  werden über die 4fach OPs U1A und U2A (AD8668ARUZ) gebuffert und über ein Widerstandsnetzwerk ( $R$ ) addiert. Die eingestellten Widerstände der Potentiometer bestimmen die jeweiligen Spannungen, die zur Addition geführt werden. Die Spannungen über den Potentiometern  $R_{Var-1-4}$  können durch die Transistoren M2 bis M5 abhängig von den anliegenden Bits (LSB,S2,S3,MSB) nach GND gezogen werden, die Spannung über  $R_{Var-5}$  ist fest und bestimmt die minimale Ausgangsspannung des diskreten DACs, die an die Spezifikation durch Messung bzw. Simulation angepasst wird. Bei einer linearen Ausgangskennlinie werden die Potentiometer so eingestellt, dass die jeweilige Spannung über dem

Potentiometer  $R_{V_{ar-i}}$  zu

$$U_{R_{V_{ar-i}}} = \frac{VDD - U_{R_{V_{ar-5}}}}{2^i} \quad (3.28)$$

wird. Durch eine von Gleichung 3.28 abweichende Einstellung der Potentiometer kann auch eine nichtlineare Kennlinie des DAC realisiert werden. Die Summe der Teilspannungen über den Potentiometern  $R_{V_{ar-i}}$  liegt mit negiertem Vorzeichen am Ausgang von OPB von U2A. Deswegen erfolgt durch OPC von U2A eine Invertierung dieser Spannung, die als Referenzspannung für den Linearregler, bestehend aus OPD von U2A und M1, dient. Alle Spannungen sind durch Kondensatornetzwerke entkoppelt. Am Ausgang V-Out steht die entkoppelte Referenzspannung zur lastunabhängigen Versorgung des Chips bereit. Da die Transistoren M2 bis M5 NMOS Typen sind, müssen die Eingangssignale (LSB,S2,S3,MSB) zuvor invertiert werden. Diese Invertierung erfolgt durch einen separaten Mehrfachinverter und ist hier nicht dargestellt.

### 3.16 Messergebnisse diskrete Lösung

Die Evaluierung der Schwingfrequenz der RIC des Testchips mit der diskreten Umsetzung der Digitalerschaltung funktioniert einwandfrei. Genutzt wird eine low-leakage Ringinverterkette mit dem Frequenzbereich  $4.17MHz \leq f_{RIC} \leq 5.52MHz$  für  $-20^\circ C \leq \vartheta \leq 80^\circ C$ . Die Schaltung benutzt die vier höchstwertigen Bits des Zählers zur Evaluierung der Schwingfrequenz. Der gegebene Frequenzbereich dokumentiert sich bei einmaligem Durchlauf des 16-Bit Zählers aber nicht in den höchstwertigen Bits, was folgende Berechnung zeigt:

Die Zähldauer berechnet sich mit der Zielbelegung 1111 0000 0000 0000 des Zählers bei der Maximalfrequenz  $f_{RIC-max} = 5.52MHz$  zu:

$$T = \frac{1111000000000000_2}{5.52MHz} = \frac{61440}{5.52MHz} = 11.13ms \quad (3.29)$$

Mit dieser Zähldauer hat der 16-Bit Zähler bei der minimalen RIC-Frequenz folgende Belegung:

$$\begin{aligned} counter = (f_{RIC-min} * T)_{10} &= (4.17MHz \cdot 11.13ms)_{10} = 46413_{10} = \\ &1011\ 0101\ 0100\ 1101_2 \end{aligned} \quad (3.30)$$

Bei vollem Frequenzswing der RIC ändern sich somit nur drei der vier höchstwertigen Zählerbits. Damit reduziert sich die Frequenzauflösung des Systems. Der Frequenzunterschied ist also bei der gewählten Zähldauer (und damit der Zählergröße) zu gering. Eine Erweiterung des Zählers ist zwar möglich, erscheint aber nicht praktikabel. Durch mehrmaliges Durchlaufen des gesamten Zählers

vergrößert sich aber der Unterschied der Zählerstände zwischen  $f_{RIC-max}$  und  $f_{RIC-min}$ . Das System wird so ausgelegt, dass  $f_{RIC-max}$  den Zähler vier mal komplett füllt (1111 1111 1111 1111). Das entspricht  $4 \cdot 65535 = 262140$  Flanken, was zu einer Zähldauer von

$$T = \frac{262140}{5.52MHz} = 47.489ms \quad (3.31)$$

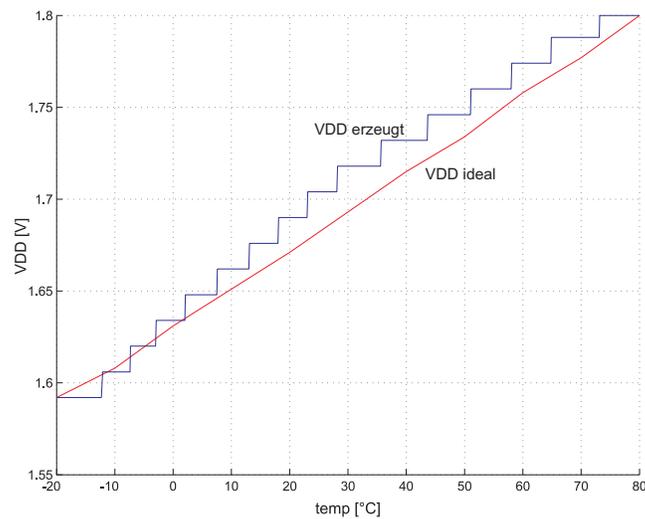
führt. Zur Berechnung der Zählerstände bei anderen RIC-Frequenzen wird diese mit  $T$  multipliziert. Von der daraus resultierenden Anzahl an Flanken subtrahiert sich die Anzahl der komplett gefüllten Zähler, also dreimal 65535. Der Rest als Binärzahl entspricht dem angezeigten Zählerstand. Tabelle 3.7 gibt den Zählerstand (mit den zur Referenzspannungserzeugung genutzten 4 höchstwertigen Bits in rot) für den gemessenen Frequenzbereich an.

Bit16-Bit13	Bit12 ... Bit1	Flanken gezählt	Flanken angezeigt	Frequenz [MHz]
0000	1111 1111 1111	200700	4095	4.215
0001	1111 1111 1111	204796	8191	4.302
0010	1111 1111 1111	208892	12287	4.389
0011	1111 1111 1111	212988	16383	4.476
0100	1111 1111 1111	217084	20479	4.563
0101	1111 1111 1111	221180	24575	4.650
0110	1111 1111 1111	225276	28671	4.737
0111	1111 1111 1111	229372	32767	4.824
1000	1111 1111 1111	233468	36863	4.911
1001	1111 1111 1111	237564	40959	4.988
1010	1111 1111 1111	241660	45055	5.085
1011	1111 1111 1111	245756	49151	5.172
1100	1111 1111 1111	249852	53247	5.259
1101	1111 1111 1111	253948	57343	5.346
1110	1111 1111 1111	258044	61439	5.433
1111	1111 1111 1111	262140	65535	5.520

**Tabelle 3.7:** Zählerstand bei verschiedenen RIC-Frequenzen

Die Frequenzauflösung beträgt 87kHz. Mit dieser Einstellung wird die untere Grenzfrequenz von  $f_{RIC-min} = 4.17MHz$  knapp nicht erfasst, eine 5-Bit Auswertung wäre hier notwendig, ebenso, wenn der Temperaturbereich in den Spezifikationen ausgedehnt würde. Für die Demonstration der Validität des Systems reicht die 4-Bit Auswertung aber aus.

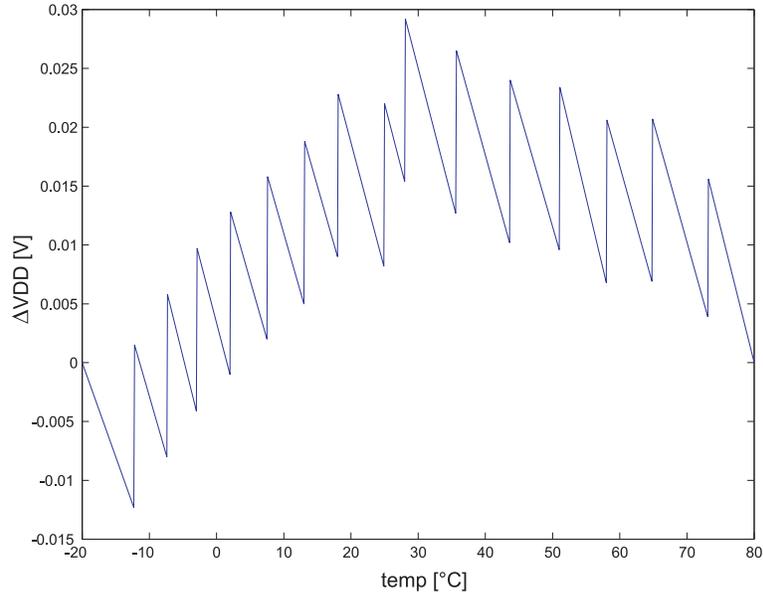
Zur Weiterverarbeitung wird der Zählerstand von der Vorbelegung subtrahiert. Hierbei wird im Beispiel die Vorbelegung 1111 gewählt, so dass Zählerstand und RIC-Frequenz umgekehrt zueinander proportional sind. Zur Evaluierung einer Zustands-spezifischen Versorgungsspannung kann nun wieder eine Simulation herangezogen werden, im praktischen Beispiel wurde die Schwingfrequenz einer minimal dimensionierten Inverterkette als Performanceindex für typische Digitalschaltungen genutzt und deren individuelle Versorgungsspannung über die Temperatur bei gleichbleibender Frequenz messtechnisch bestimmt. Abbildung 3.58 zeigt die für gleichbleibende Performance erforderliche Spannung (rote Linie) und die vom System erzeugte Versorgungsspannung (blaue Linie).



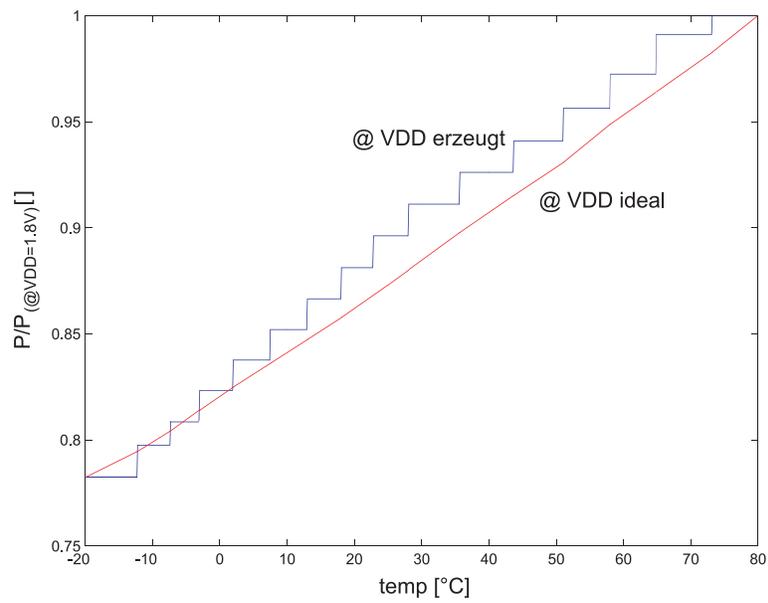
**Abbildung 3.58:** Kennlinie VDD-Erzeugung mit diskretem Aufbau

Der DAC (vergl. Abb 3.57) ist nun mit linearer Kennlinie so eingestellt, dass 16 Ausgangsspannungen zwischen 1.592V (minimale Versorgungsspannung bei  $-20^{\circ}\text{C}$ ) und 1.8V (maximale Versorgungsspannung bei  $80^{\circ}\text{C}$ ) erzeugt werden. Durch den leicht nichtlinearen Zusammenhang zwischen RIC-Schwingfrequenz und Temperatur sind die Treppenstufen der erzeugten Ausgangsspannung in Abb. 3.58 nicht äquidistant. Um die Spezifikation der Digitalschaltung unter allen Umständen einzuhalten, muss die erzeugte Versorgungsspannung (blaue Linie in Abb.3.58) für jede Temperatur eigentlich immer mindestens gleich der optimalen Versorgungsspannung (rote Linie in Abb. 3.58) sein. Dies ist im Beispiel für tiefe Temperaturen knapp nicht der Fall. Es müsste ein etwas größerer Startpunkt (hier 1.592V) gewählt werden. Abbildung 3.59 zeigt den absoluten Fehler zwischen erzeugter und optimaler Versorgungsspannung. Hier wird der durch die Quantisierung bedingte Sägezahnverlauf deutlich. Aussagekräftiger wird der Vergleich der eingesparten Leistung (Abbildung 3.60). Die Kurve des

relativen Leistungsumsatzes für die erzeugte Versorgungsspannung ist idealerweise identisch mit der optimalen, in der Realität kann sie aber nur punktwise mit ihr übereinstimmen. Eine gesteigerte Auflösung der Versorgungsspannung erhöht die Qualität der Versorgungsspannungseinstellung.



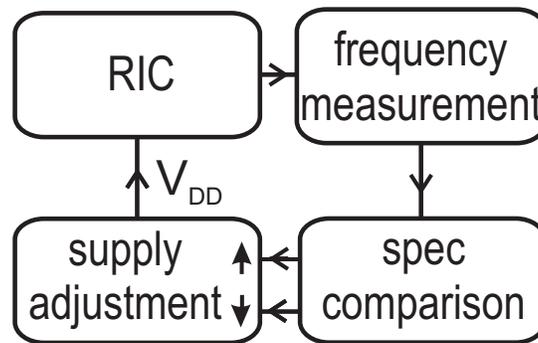
**Abbildung 3.59:** Differenz zwischen erzeugter und optimaler Versorgungsspannung



**Abbildung 3.60:** relativer Leistungsumsatz bei optimaler und erzeugter Versorgungsspannung

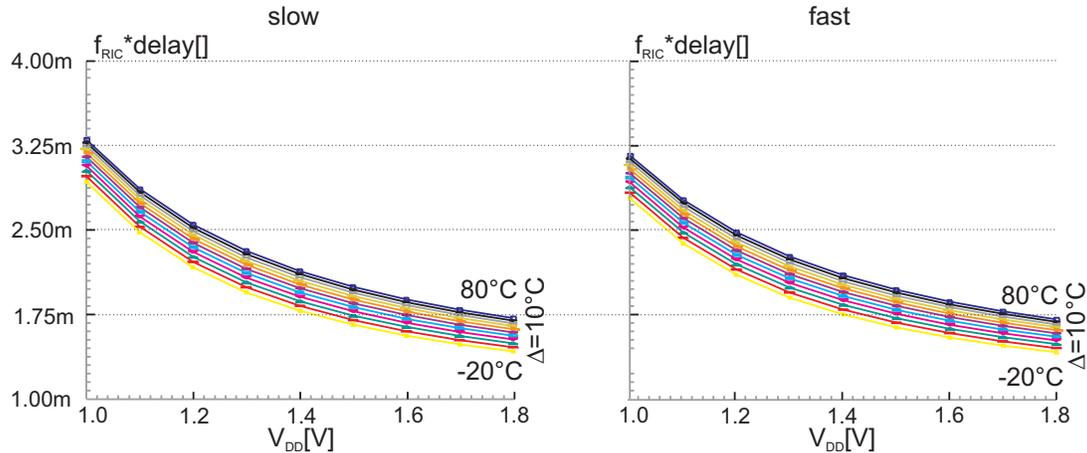
## 3.17 Regelung auf konstante RIC-Frequenz

Bislang arbeitet das System mit konstanter Versorgungsspannung der Ringinverterkette und leitet aus deren Schwingfrequenz eine neue Versorgungsspannung für die eigentliche Applikation ab. Es handelt sich hierbei um eine Steuerung. Ein andere Weg ist die Versorgung der RIC und der Applikation mit derselben Versorgungsspannung, die die Frequenz der RIC konstant erhält. Ein Konzept für ein solches System zeigt Abbildung 3.61 in Analogie zu Abbildung 3.9.



**Abbildung 3.61:** Konzept zur Versorgungsspannungseinstellung mit konstanter RIC-Frequenz

Die Versorgungsspannung wird so geregelt, dass die RIC-Frequenz konstant bleibt. Die Qualität eines solchen Konzepts hängt von der Korrelation der Performance minimal dimensionierter Digitalschaltungen mit der Schwingfrequenz der groß dimensionierten Ringinverterkette ab. Bei einer hypothetisch angenommenen perfekten Korrelation könnte dieses Konzept unabhängig von den gerade vorliegenden Chipparametern und der Temperatur die Performance konstant halten, denn dann ist bei konstanter RIC-Frequenz auch das Delay einer Digitalschaltung konstant. Zur Untersuchung dieser Korrelation werden Simulationen durchgeführt, die das Delay einer Gatterkette zur Frequenz der RIC für verschiedene Temperaturen und maximal abweichenden systematischen Chipparametern in Beziehung setzt (Abbildung 3.62).



**Abbildung 3.62:** Korrelation zwischen RIC-Frequenz und Delay einer Gatterkette

Es ist deutlich zu erkennen, dass weder über die Temperatur, noch über die Versorgungsspannung eine konstante Korrelation zwischen RIC-Frequenz und Delay der Gatterkette herrscht. Positiv ist die relativ geringe Abweichung dieser Korrelation in Abhängigkeit der systematischen Chippparameter (slow, fast) zu bewerten. Im Folgenden soll untersucht werden, zu welchen Verzögerungsschwankungen die Anwendung des Konzepts nach Abbildung 3.61 in diesem konkreten Beispiel im Vergleich zur festen Standardversorgung führt. Abbildung 3.63 zeigt exemplarisch die erzeugten Versorgungsspannungen und die Verzögerungszeiten, die aus diesen erzeugten Versorgungsspannungen folgen. Nicht berücksichtigt ist hier das geometrische Mismatch der Transistoren. Bei einer angenommenen Regelfrequenz der RIC von  $f_{ric} = 5MHz$  erzeugt dieses Konzept eine Performancespreizung von hier maximal  $21ps$ , was einer Abweichung von  $\approx 7\%$  entspricht. Mit der Standardversorgung von  $V_{DD} = 1.8V$  werden  $50ps$  ( $\approx 17\%$ ) erreicht. Die Performancespreizung wird umso größer, je kleiner die Regelfrequenz gewählt wird, d.h. je geringer die Geschwindigkeitsansprüche an die Applikation sind. Die Vorgehensweise des Designers muss die Festlegung einer RIC-Frequenz beinhalten, die unter Berücksichtigung des geometrischen Mismatches die Spezifikationen sicherstellen. Unter Berücksichtigung des geometrischen Mismatches liegt die Performancespreizung drastisch über den o.a. Werten (vergl. Kapitel 3.13). Das in diesem Abschnitt vorgestellte Konzept bietet diesbezüglich keinen Vorteil. Auch hier ist die Verwendung einer minimal dimensionierten Ringinverterkette nicht empfehlenswert, da deren Schwingfrequenz stark abhängig von dem geometrischen Mismatch ist. Im schlechtesten Fall variieren Ringinverterkette und Applikation in entgegengesetzte Richtungen, was zu einer Verdopplung der Performancespreizung führt (mit entsprechendem stark erhöhten Sicherheitsmargin bei der Versorgungsspannungseinstellung). Durch den Einsatz einer vom geometrischen Mismatch unabhängigen Ringinverterkette halbiert sich diese Un-

sicherheit mit entsprechend verkleinertem Sicherheitsmargin. Die Komplexität dieses Systems ist vergleichbar mit dem bereits vorgestellten und realisierten System.

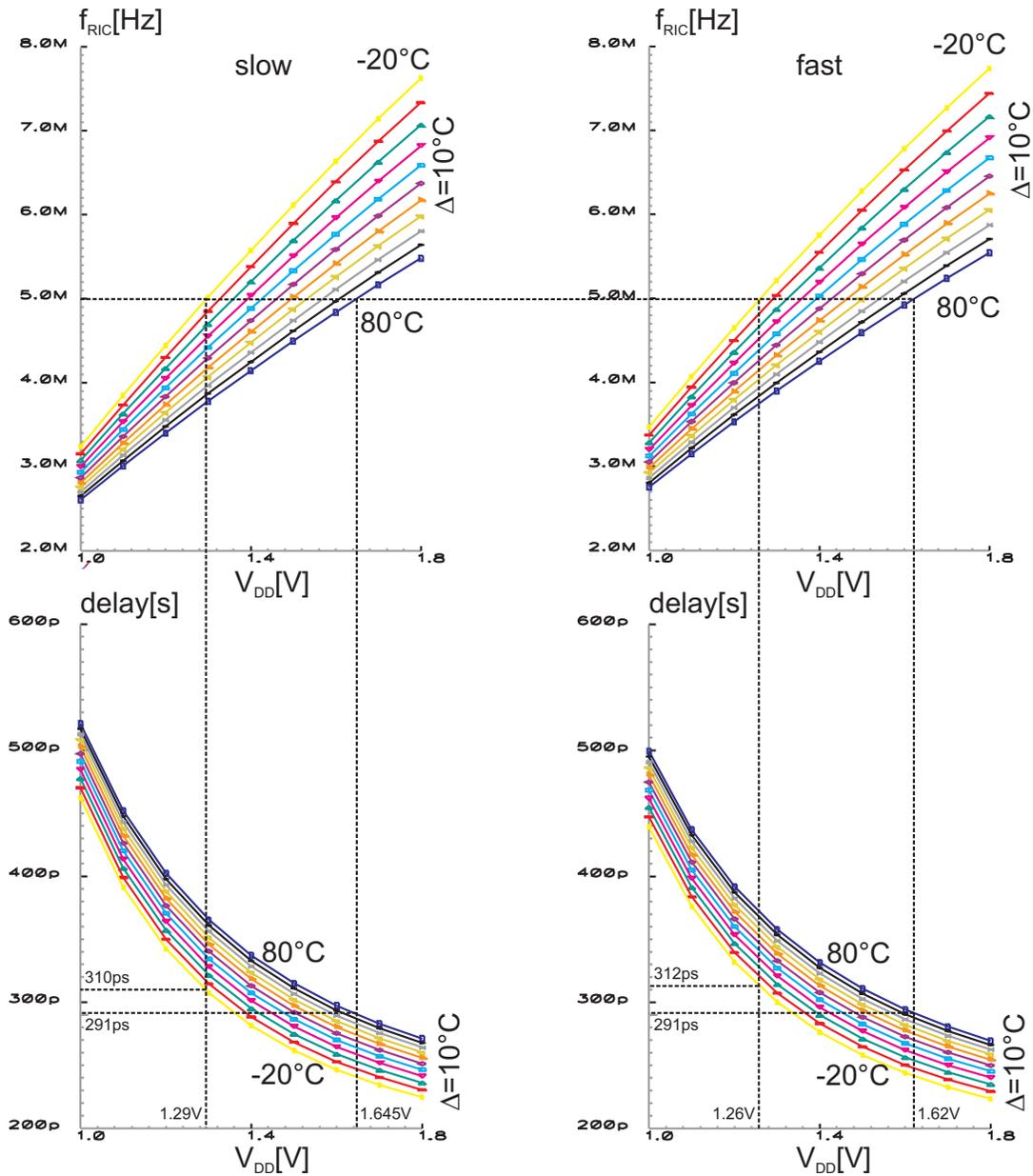


Abbildung 3.63:  $V_{DD}$ -Erzeugung mit konstanter RIC-Frequenz

## 3.18 Beurteilung des Performance-Driven-Supply Systems

Das vorhergehende Kapitel beschreibt die Realisierung eines validen Systems zur parameterabhängigen dynamischen Einstellung der Versorgungsspannung in digitalen integrierten Schaltungen. Die Erfassung der Schwingfrequenz der Ringinverterkette als Sensorgröße ist mit den vorgestellten Schaltungen möglich und ein Standardverfahren, das in jeder Technologie (Strukturgröße) einfach umsetzbar ist. Die Wandlung der Sensorgröße in eine spezifische Versorgungsspannung, die die Einhaltung der Spezifikation gewährleistet, ist eine mehr individuelle Problemstellung. Die vorgestellten und messtechnisch untersuchten Digital-Analog-Konverter für diesen Zweck stellen Vorschläge dar, welche nur eine geringe Komplexität aufweisen. In modernen MOS-Technologien werden i.d.R. Digital-Analog-Konverter als vorentwickelte IP-Blöcke angeboten, auf die zurückgegriffen werden kann. Die Erzeugung einer individuellen Versorgungsspannung mit Hilfe einer direkt erzeugten Referenzspannung erscheint aber angesichts der aufgezeigten nichtlinearen Proportionalität von Chipperformance und Versorgungsspannung in Abhängigkeit der Temperatur als nicht empfehlenswert. Das in Abbildung 3.27 gestrichelt dargestellte System ist hier die bessere Lösung. In einem ROM dauerhaft gespeicherte Information kann im Rahmen des geometrischen Mismatches und der Speichertiefe bzw. Auflösung der Zähler zur RIC-Frequenzauswertung wesentlich genauer auf den individuellen Chip abgestimmt werden. Für höchstwertige Applikationen mit besonderem Fokus auf optimale Versorgungsspannung kann diese Information nicht durch Simulation sondern durch Messung evaluiert werden. Somit entsteht ein für jeden Chip individueller Informationssatz, der aber in einem beschreibfähigen, dauerhaften Speicher (extern, oder intern) abgelegt werden muss. Diese Vorgehensweise ist extrem aufwändig und lohnt sich nur für spezielle Applikationen.

Die Standardvorgehensweise ist die Evaluierung der parameterspezifischen Versorgungsspannung durch WorstCase-Simulationen. Dabei ist es von entscheidender Bedeutung, dass die Simulationsmodelle und die benutzte Technologie auf dem selben Evolutionsstand sind. Von weiterer entscheidender Bedeutung ist die Qualität der DC-DC-Konvertierung. Dies ist aber allen Konzepten zur Reduzierung der Verlustleistung, die an der Versorgungsspannung anknüpfen, gemeinsam und wird in Kapitel 4 aufgegriffen. Interessant erscheint die vorgestellte Technologie, wenn vom Paradigma der unveränderlichen Performance abgerückt wird. In [17] wird ein System des Versorgungsspannung-Hoppings vorgestellt, welches software-gesteuert die Versorgungsspannung in Zeiten niedriger Rechenintensität um einen festen Betrag herunter regelt (siehe Abbildung 3.64). Auch andere Arbeiten befassen sich mit dem Einsatz mehrerer diskreter Versorgungsspannungen [10]. Es gibt also zwei Versorgungsspannungen  $VDD$  und  $VDD_{min}$ . Wird dieses System mit dem vorgestellten System kombiniert, können

beide Versorgungsspannungen, abhängig vom aktuellen Parametersatz, nach unten korrigiert werden. Es ergeben sich somit  $VDD - \Delta V$  und  $VDD_{min} - \Delta V$  und damit weitere Verlustleistungseinsparungen. Abbildung 3.64 zeigt das in [17] veröffentlichte Oszilloskopogramm zweier software-gesteuerter Versorgungsspannungen und der resultierenden Strombelastungen in Abhängigkeit der aktuellen Performanceanforderung.

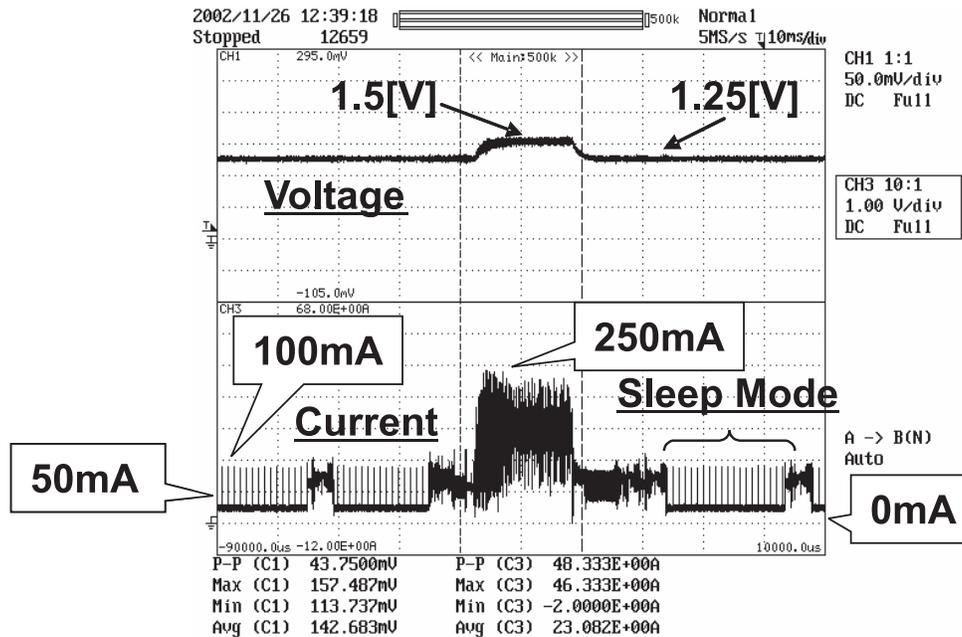


Abbildung 3.64: Versorgungsspannung-Hopping [17]

Für die in Abbildung 3.64 gemessene Applikation wird eine Einsparung von 33% angegeben. Das ungewisse geometrische Mismatch minimal dimensionierter Transistoren ist ein Handicap, das alle Systeme, also auch das vorgestellte, betrifft. Eine Erhöhung der Herstellungsgenauigkeit wirkt sich dementsprechend positiv auf die zu erwartende Einsparung durch Anwendung des vorgestellten Systems aus. Mit jedem Migrationsschritt in Richtung kleiner werdender Strukturgröße ist aber die Einhaltung absoluter geometrischer Dimensionen schwieriger. Die Realisierung von low-power-low-performance Applikationen in Standard-Technologien größerer Strukturgröße erscheint also neben dem kleineren Herstellungspreis auch hinsichtlich des i.d.R. besseren geometrischen Matchings interessant. Die Einsetzbarkeit des vorgestellten Systems ist grundsätzlich unabhängig von Technologie und Gesamtleistungsumsatz. Da die Ringinverterkette alleine schon ein beträchtlicher Verbraucher ist, sollte dieses System nicht dauerhaft arbeiten. Das Auftrennen der Rückkopplung durch ein entsprechendes Transmissiongate sorgt für einen stationären Zustand der Kette, die dann keine dynamische- und short-circuit-Verlustleistung mehr umsetzt. Durch die große

Kanallänge und die beschränkte Anzahl an Invertern ist der statische Stromverbrauch vernachlässigbar. Der Start des Systems kann grundsätzlich softwaregesteuert oder zeit-gesteuert erfolgen. Beides stellt kein Problem hinsichtlich der Implementierung dar. Die durchschnittliche Einsparung durch Anwendung des Systems über eine Chipmarge von  $\approx 20\%$  (für die verwendete Technologie) erweist sich als *zusätzlicher* Beitrag zur Energieeinsparung, und kann eine wertvolle IP für alle Technologien darstellen.

# Kapitel 4

## DC-DC Spannungswandler

Allen Applikationen mit adaptiver Betriebsspannungsversorgung ist die Notwendigkeit einer variabel einstellbaren Versorgungsspannung gemeinsam. Im einfachsten Fall ist dies durch die Bereitstellung serieller Spannungsquellen, etwa durch in Serie geschaltete Batterien oder Akkuzellen, oder durch unterschiedliche Versorgungsspannungen über ein System, gewährleistet. Nachteilig ist hierbei die nur grobe Unterteilung der Versorgungsspannung. In den meisten Fällen wird ein DC-DC Konverter benötigt. Durch ihn ist eine frei wählbare Versorgungsspannung möglich. Die Konzepte zur Spannungskonvertierung lassen sich in zwei Hauptgruppen unterscheiden:

- Linearregler
- Schaltende Regler
  - Induktive Buck-/Boost-Konverter
  - Kapazitive Buck-/Boost-(Ladungspumpe)-Konverter

Kapazitive Konverter werden im Folgenden nicht betrachtet, da sie zumeist mehrere externe Bauelemente benötigen. Durch den Wegfall der Spulen ist diese Bauart als diskreter Wandler besonders interessant ([6], [50], [43]). Spulen sind i.A. teurer als Kondensatoren und bringen einen erhöhten Montageaufwand und zusätzliches Gewicht mit sich. Für mobile Applikationen mit geringem Leistungsumsatz (z.B. Mobilphone) ist diese Art der Wandlung eine interessante Alternative. Problematisch bei kapazitiven DC-DC Konvertern erweisen sich die umfangreichen Schaltvorgänge. Diese ziehen Umladeverluste nach sich, was dem Wirkungsgrad abträglich ist und verursachen hochenergetische, steilflankige Strompulse, die selbst auf kurzen Leitungswegen elektromagnetische Störimpulse nach sich ziehen [4]. Ein penibles, auf EM-Verträglichkeit geprüfetes Layout ist bei solchen DC-DC Konvertern Pflicht. In ähnlicher Form gilt das für alle schaltenden Regler. Ein Zwischending aus schaltendem Regler und Linearregler stellt der in Kapitel 4.2 vorgestellte Regler dar [35].

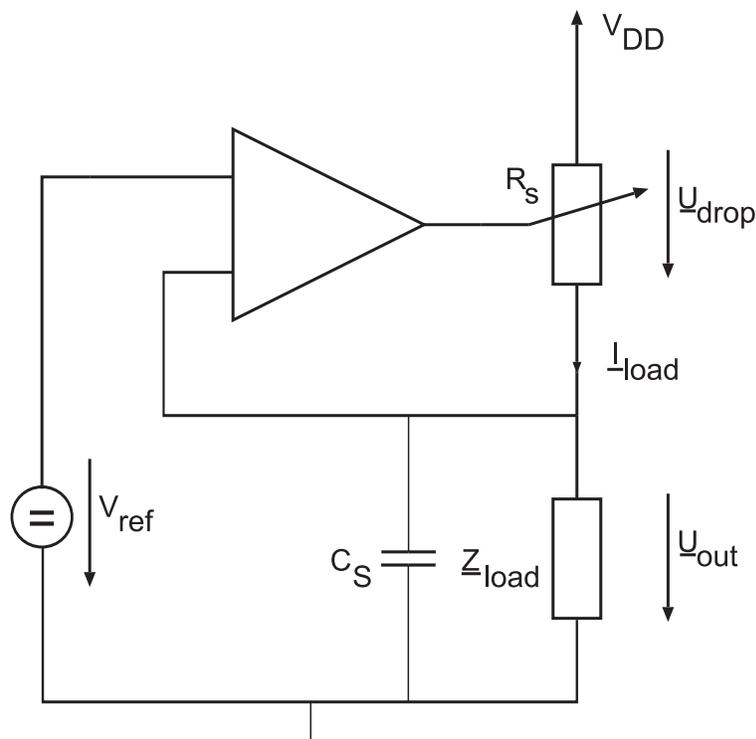
Für das in Kapitel 3.14 vorgestellte System soll nun die Einsetzbarkeit der verschiedenen Spannungswandler untersucht werden. Dabei wird ein besonderer Focus auf niedrigste Energieverluste der Reglerschaltungen gelegt, um auch bei Low-Power-Applikationen einen hohen Wirkungsgrad zu erreichen. Alle vorgestellten Wandler können zusammen mit der eigentlichen Schaltung auf einem ASIC implementiert werden. Externe Bauelemente sollen so weit wie möglich vermieden werden. Bei Realisierung dieser Ziele kann dem Designer eine vorgefertigte IP anhand gegeben werden.

## 4.1 Der Linearregler

Vielfach wird zur Spannungsregelung der Linearkonverter eingesetzt, der aufgrund seiner permanenten Nachführung der Ausgangsspannung eine gute Unterdrückung von Ausgangsspannungsschwankungen bietet. Nach Abbildung 4.1 hängt der Wirkungsgrad  $\eta$  von der Spannung über dem Stellglied  $U_{drop}$  und der Belastung  $I_{load}$  ab. Es gilt für den Wirkungsgrad:

$$\eta = \frac{U_{out} \cdot I_{load}}{P_{ges}} = \frac{V_{DD} - U_{drop} \cdot I_{load}}{P_{ges}} \quad (4.1)$$

$$P_{ges} = V_{DD} \cdot I_{load} + P_{DC-DC} \quad (4.2)$$



**Abbildung 4.1:** Schematischer Aufbau eines linearen DC-DC-Konverters

Für große Dropspannungen und kleine Lasten wird der Wirkungsgrad eines Linearkonverters schlecht, da zum einen immer mehr Leistung im Stellglied umgesetzt wird, zum anderen das Verhältnis der Leistungsaufnahme der Applikation zur Leistungsaufnahme der DC-DC Regelschaltung ungünstig wird. Anzustreben, im Sinne eines guten Wirkungsgrads, sind deshalb ein kleiner Leistungsumsatz der an der Regelung beteiligten Komponenten, wie Differenzverstärker und Referenzspannungsquelle, und kleine Dropspannungen. In CMOS-Prozessen

wird als Stellglied ein P- oder NMOS-Transistor verwendet. Bei kleinen Dropspannungen ist die Ansteuerung des Stelltransistors aber kritisch, da sich der Transistor bei kleinem  $U_{DS}$  im linearen Bereich befindet, wo die Reaktivität des Laststroms in Abhängigkeit von der Steuerspannung immer geringer wird. Dies ist ein Problem bei der Realisierung sogenannter LDO-DC-DC Konverter (Low-Drop-Output Konverter). Wird wegen der ca. 3fach besseren Leitfähigkeit von NMOS-Transistoren gegenüber PMOS-Transistoren gleichen W/L-Verhältnisses ein NMOS-Transistor als Stellglied eingesetzt, wird die Steuerspannung  $U_{GS}$  bei kleinen Dropspannungen sehr klein. Um  $U_{GS}$  zu erhöhen, muss das Gatepotential über  $V_{DD}$  angehoben werden. Dies geschieht durch Ladungspumpen. Der Einsatz von Ladungspumpen bedeutet aber einen nicht unerheblichen Mehraufwand, Platzbedarf und zusätzlichen Leistungsumsatz, weswegen der Einsatz von NMOS-Transistoren als Stellglied nur für hohe Lastströme lohnend erscheint, denn nur dann ist der dreifach niedrigere Platzverbrauch gegenüber einem, für denselben Laststrom dimensionierten, PMOS-Transistor von Vorteil. Regelungstechnisch bietet der Einsatz eines NMOS-Stelltransistors den Vorteil einer gewissen Selbstgegenkopplung, hängt doch seine Steuerspannung  $U_{GS}$  von der Ausgangsspannung  $U_{out}$  in der Weise zusammen, dass eine Erhöhung von  $U_{out}$  eine Verminderung von  $U_{GS}$  bewirkt. Der vorgestellte Linearkonverter für Low-Power Applikationen soll folgendes Anforderungsprofil erfüllen:

- Versorgungsspannung  $V_{DD} = 1.8V$
- Ausgangsstrom  $I_{load} \leq 2mA$
- Ausgangsspannung  $V_{out} \geq 1,0V$ ; LowDrop Fähigkeit
- hohe Konstanz der Ausgangsspannung
- $\rightarrow$  hohe PSRR, hohe LRR (**L**oad **R**ejection **R**atio)
- hoher Wirkungsgrad  $\rightarrow$  niedriger Leistungsverbrauch der Komponenten, LDO-Fähigkeit

Die Gesamtleistung in Abhängigkeit von der Ausgangsspannung des Linearkonverters ( $P_{ges-DCDC}$ ) berechnet sich unter Vernachlässigung des Eigenverbrauchs der Regelschaltung  $P_{DCDC}$  wie folgt:

$$P_{ges-DCDC} = V_{DD} \cdot |I_{load}| \quad (4.3)$$

$$|I_{load}| = \frac{|U_{out}|}{|Z_{load-ges}|} \Rightarrow \quad (4.4)$$

$$P_{ges-DCDC} = V_{DD} \cdot \frac{|U_{out}|}{|Z_{load-ges}|} \quad (4.5)$$

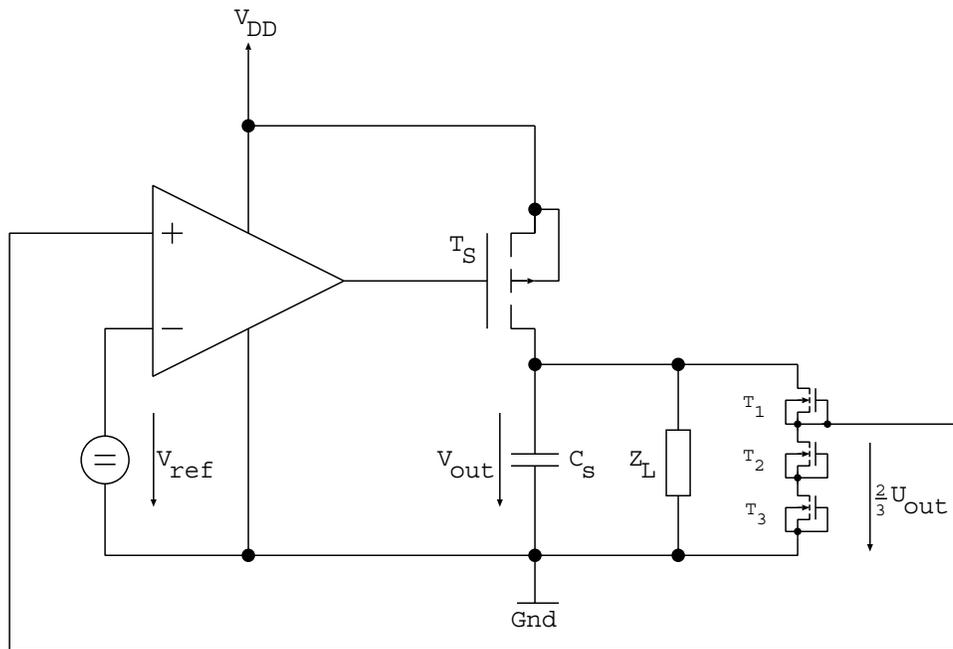
$Z_{load-ges}$  resultiert aus der eigentlichen Last  $Z_{load}$ , nämlich der zu versorgenden Schaltung, und der parallelgeschalteten Siebkapazität  $C_S$ , die die Ausgangsspannung puffert und glättet.  $Z_{load}$  ist größtenteils kapazitiv, da die Last in digitalen CMOS-Technologien überwiegend aus schaltenden Transistoren mit entsprechenden Lastkapazitäten besteht. Für die Leistungersparnis gegenüber der gleichen Applikation mit voller Versorgungsspannung  $V_{DD}$  folgt:

$$\frac{P_{ges-DCDC}}{P_{ges-VDD}} = \frac{V_{DD} \cdot \frac{|U_{out}|}{|Z_{load-ges}|}}{\frac{V_{DD}^2}{|Z_{load-ges}|}} = \frac{|U_{out}|}{V_{DD}} \quad (4.6)$$

Im Folgenden wird auf die komplexe Darstellung von Spannungen und Strömen aus Gründen der Übersichtlichkeit verzichtet. Aus Gleichung 4.6 lässt sich folgern, dass obwohl die in der Applikation umgesetzte Leistung proportional zum Quadrat ihrer Versorgungsspannung  $U_{out}$  skaliert, die Gesamtleistung durch den Leistungsumsatz im Stelltransistor nur linear mit  $U_{out}$  skaliert. Dies stellt einen entscheidenden Nachteil von Linearkonvertern dar.

### 4.1.1 Topologie des linearen DC-DC Konverters

Der in der vorliegenden Dissertation vorgestellte Linear-Konverter arbeitet mit einem PMOS-Transistor als Stellglied. Durch den verhältnismäßig kleinen Laststrom  $I_{load} \leq 2mA$  lohnt sich die Anwendung eines NMOS Transistors aus in Kap 4.1 angegebenen Gründen nicht. Der schematische Aufbau des vorgestellten Linearreglers ist in Abb. 4.2 zu sehen.



**Abbildung 4.2:** Schematischer Aufbau des vorgestellten Linear DC-DC-Konverters

Der Differenzverstärker wird von der unregulierten Versorgungsspannung  $V_{DD}$  gespeist. Als analoge Komponente arbeitet der Differenzverstärker mit fester Versorgungsspannung sicher innerhalb des Anforderungsprofils. Die Serienschaltung der als Diode verschalteten Transistoren  $T_1$  bis  $T_3$  bilden einen Spannungsteiler, der durch entsprechend kleines  $W/L$ -Verhältnis der Transistoren ( $W$  nicht minimal) die Schaltung kaum zusätzlich belastet. Durch die völlige Gleichheit dieser Transistoren ist ein gutes thermisches Verhalten gegeben, d.h. die Teilerspannungen ändern sich über einen weiten Temperaturbereich nicht. Auch Mismatch (Prozess und  $W/L$ -Verhältnis) hat nur einen geringen Einfluss auf die Ausgangsspannung des Spannungsteilers. Das Teilverhältnis von 2 zu 1 vermindert die maximale Eingangsspannung um  $\frac{1}{3}$ , genauso wie den Betrag der rückgekoppelte Spannungsänderungen am Ausgang. Um einen möglichst hohen Eingangsspannungsbereich zu gewährleisten, wird als Eingangsstufe des Differenzverstärkers eine Doppeldifferenzschaltung, bestehend aus einer NMOS- ( $T_1, T_2$ )

und einer PMOS-Differenzstufe ( $T_3, T_4$ ), gewählt. Die Differenzstufen arbeiten auf Stromspiegel ( $T_7$  und  $T_9, T_8$  und  $T_{10}$  bzw.  $T_5$  und  $T_{11}, T_6$  und  $T_{12}$ ). Der Ausgangsbuffer ( $T_{15}$  und  $T_{16}$ ), dessen Ansteuerung durch den Spannungsabfall über  $T_{13}$  und über  $T_{14}$  erfolgt, entkoppelt die Verstärkerschaltung von der Last. Somit wird der Ausgang des eigentlichen Operationsverstärkers wenig belastet, was einen kleinen Biasstrom ( $i_{bias}$ ) ermöglicht (Dimensionierung und Simulationsergebnisse siehe Kapitel A.4).

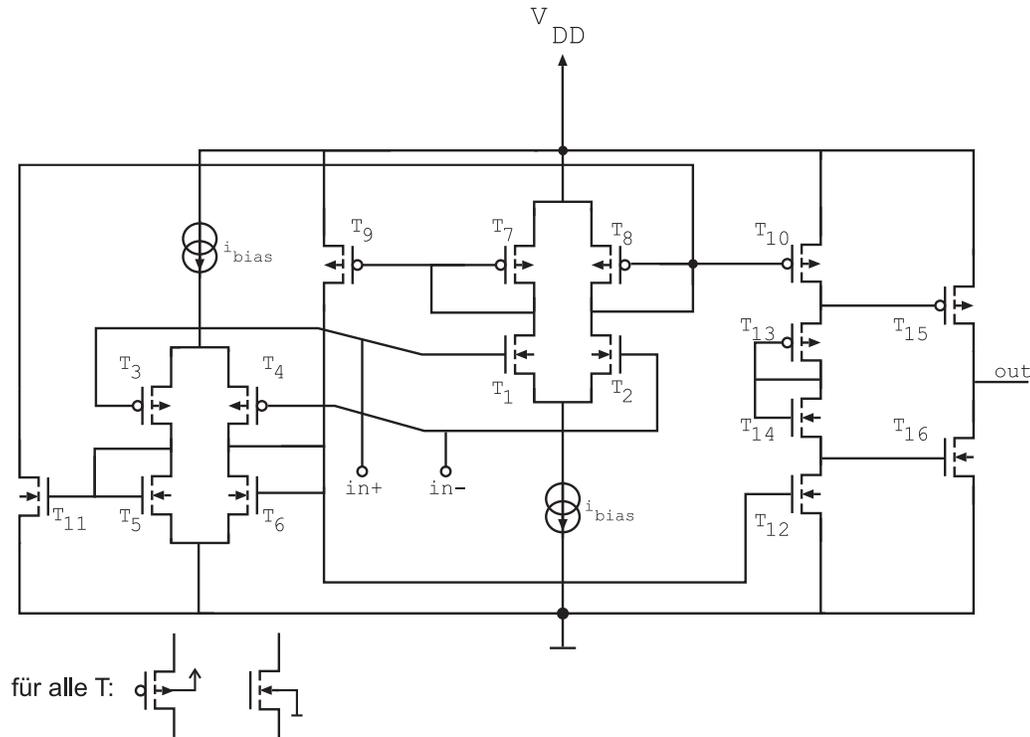


Abbildung 4.3: Schematic des verwendeten Differenzverstärkers

### 4.1.2 Simulationsergebnisse Linearregler

Die Qualität des Linearreglers wird durch Simulation verifiziert. Hierbei wird die Ausgangsspannung bei verschiedenen Randbedingungen betrachtet:

- unterschiedliche Lasten und Belastungssprünge am Ausgang
- Ripple auf der Eingangsspannung
- Temperaturabhängigkeit
- Matchinganalyse
- Abhängigkeit der Regelgüte von der Größe der Siebkapazität

Diese Randbedingungen sind für alle vorgestellten DC-DC-Konverter gleich und charakterisieren die absolute Reglerqualität. Als Lasten für den Regler kommen in der Realität selten reine ohmsche Lasten vor. Bei digitalen Schaltungen überwiegen Ladevorgänge der Lastkapazitäten, also Auf- und Entladen der aktiven Gates, sowie der Leitungskapazitäten. Eine realitätsnahe Belastung wäre prinzipiell eine gemischt komplexe Last mit kleinem seriellen induktivem Anteil, sowie einem ohmschen Serienwiderstand und verhältnismäßig großem kapazitivem Anteil. Eine solche Last stellt an sich aber nicht den WorstCase dar, da sie reaktiv auf den Laststrom wirkt, hängt doch der Laststrom über die Bauteilegleichungen mit der Ausgangsspannung zusammen, die unvermeidbaren Schwankungen unterworfen ist (Ripple). Die härteste aller denkbaren Belastungen ist eine ideale Stromquelle. Die Simulationen werden aus Gründen der Realitätsnähe aber mit reellen Lasten durchgeführt.

In Tabelle 4.1 wird die Ausgangsspannung von einer externen Kapazität  $C_S = 1\mu F$  geglättet. Der Wert der Kapazität ergibt sich aus der Forderung nach möglichst kleinen geometrischen Abmessungen bei möglichst großer Kapazität.

$V_{In}$ [V]	$V_{Out}$ [V]	$V_{Out-Max}$ [V]	$V_{Out-Min}$ [V]	Wirkungsgrad $\eta$
1.8	1.7	1.7	1.7	0.94
1.8	1.6	1.6	1.6	0.88
1.8	1.5	1.5	1.5	0.83
1.8	1.4	1.4	1.4	0.77
1.8	1.3	1.3	1.3	0.71
1.8	1.2	1.2	1.2	0.66

**Tabelle 4.1:** Simulationsergebnisse des Linearreglers mit  $C_S = 1\mu F$  und  $R_L = 1k\Omega$

Durch die verhältnismäßig kleine Last von  $I_L < 2mA$  spielt der Leistungsverbrauch der Reglerschaltung für den Wirkungsgrad eine größere Rolle, als bei

Applikationen mit hohem Strombedarf. Der Leistungsverbrauch der Reglerschaltung setzt sich aus dem Leistungsverbrauch des Differenzverstärkers, sowie der Lade- und Entladeleistung im Gate des Stelltransistors  $T_S$  zusammen. Weitere Verluste entstehen durch den statischen Strom des Spannungsteilers, der die Ausgangsspannung rückkoppelt und der Referenzspannungsquelle, die bei der Betrachtung an dieser Stelle nicht eingeht. Wird der Ausgangsstrom reduziert, wirkt sich das relativ stark auf den Wirkungsgrad aus, wie Tabelle 4.2 zeigt.

$V_{\text{out}}$ [V]	Wirkungsgrad		
	$R_L = 1k\Omega$	$R_L = 5k\Omega$	$R_L = 10k\Omega$
1.7	0.94	0.90	0.86
1.6	0.88	0.85	0.81
1.5	0.83	0.80	0.75
1.4	0.77	0.74	0.70
1.3	0.71	0.68	0.64
1.2	0.66	0.63	0.59

**Tabelle 4.2:** Wirkungsgrad bei verschiedenen Lasten und Ausgangsspannungen

Im Sinne hoch integrierter Baugruppen ist das Verhalten der Regler mit kleinen integrierbaren Kapazitäten interessant. Als integrierte Kapazität mit noch tolerierbarem Flächenbedarf wird  $1,5nF$  angenommen. Die Tabellen 4.1.2 und 4.1.2 zeigen die Simulationsergebnisse für den Linearregler mit  $C_S = 1,5nF$  mit fester Eingangsspannung von  $V_{in} = 1,8V$  mit zwei verschiedenen Lasten.

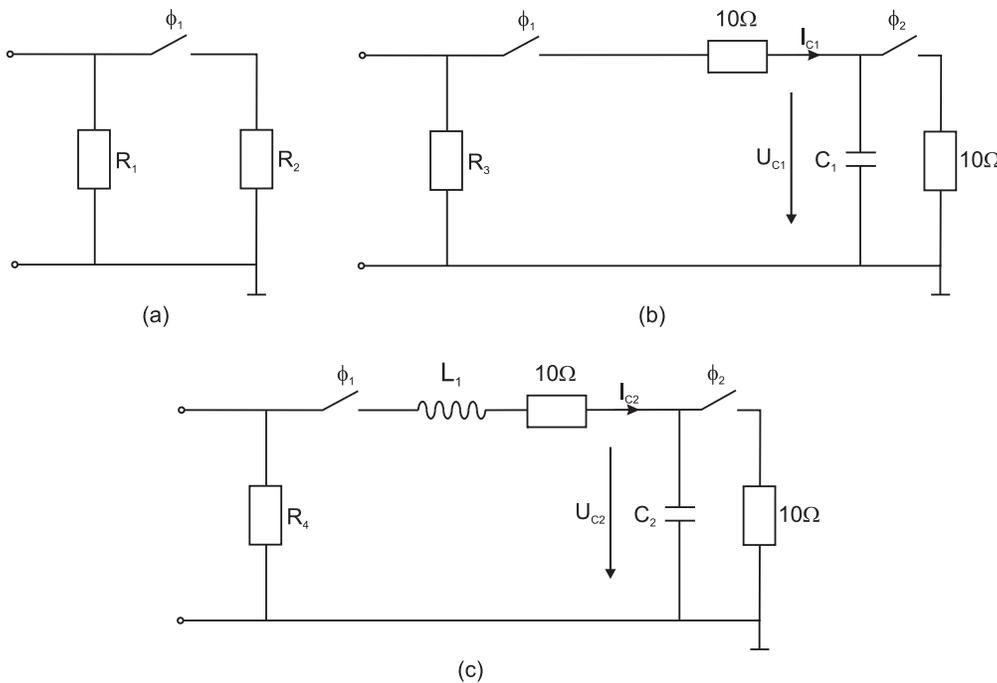
$V_{\text{In}}$ [V]	$V_{\text{Out}}$ [V]	$V_{\text{Out-Max}}$ [V]	$V_{\text{Out-Min}}$ [V]	Wirkungsgrad $\eta$
1.8	1.7	1.720	1.673	0.94
1.8	1.6	1.641	1.570	0.89
1.8	1.5	1.560	1.472	0.84
1.8	1.4	1.465	1.373	0.79
1.8	1.3	1.369	1.274	0.73
1.8	1.2	1.271	1.175	0.68

**Tabelle 4.3:** Variation der Ausgangsspannung bei konstanter Last ( $R_L = 1k\Omega$ )

$V_{In}$ [V]	$V_{Out}$ [V]	$V_{Out-Max}$ [V]	$V_{Out-Min}$ [V]	Wirkungsgrad $\eta$
1.8	1.7	1.712	1.693	0.92
1.8	1.6	1.614	1.593	0.87
1.8	1.5	1.514	1.493	0.82
1.8	1.4	1.413	1.394	0.76
1.8	1.3	1.313	1.294	0.70
1.8	1.2	1.213	1.194	0.65

**Tabelle 4.4:** Variation der Ausgangsspannung bei konstanter Last ( $R_L = 10k\Omega$ )

Simulationen mit realistischen Belastungen und schwankenden Eingangsspannungen geben einen Überblick über das reale Regelverhalten des linearen DC-DC-Konverters. Die folgenden Simulationsergebnisse beziehen sich hierbei auf die Lastelemente, wie sie in Abbildung 4.4 dargestellt sind, bzw. auf Eingangsspannungen, die analog Tabelle 4.5 schwanken.



**Abbildung 4.4:** Lastelemente

Die Takte  $\Phi_1$  und  $\Phi_2$  der Schalter in Bild 4.4 sind nicht überlappend. Die Schaltung simuliert das Verhalten einer CMOS-Digitalschaltung, bei der große Lastkapazitäten ge- und entladen werden. Bei den folgenden Simulationen werden drei verschiedene Eingangsspannungen verwendet.

	Versorgungsspannung
1.	$U_{In}(t) = 1.8 V$
2.	$U_{In}(t) = 1.8 V + 0.1 V \cdot \sin(2 \cdot \Pi \cdot 1 MHz \cdot t)$
3.	$U_{In}(t) = 1.8 V + 0.1 V \cdot \sin(2 \cdot \Pi \cdot 100 MHz \cdot t)$

Tabelle 4.5: Arten von Versorgungsspannungen

Die Last wird zunächst konstant gehalten. Die Angabe der Eingangsspannung bezieht sich auf Tabelle 4.5.

$V_{In}$ [V]	$V_{Out-Max}$ [V]	$V_{Out-Min}$ [V]	Last [k $\Omega$ ]	Wirkungsgrad $\eta$
2.	1.5	1.499	1	0.80
2.	1.5	1.5	5	0.78
2.	1.5	1.5	10	0.77
3.	1.5	1.5	1	0.82
3.	1.5	1.5	5	0.80
3.	1.5	1.5	10	0.78

Tabelle 4.6: Schwankende Eingangsspannungen bei konstanten Lasten

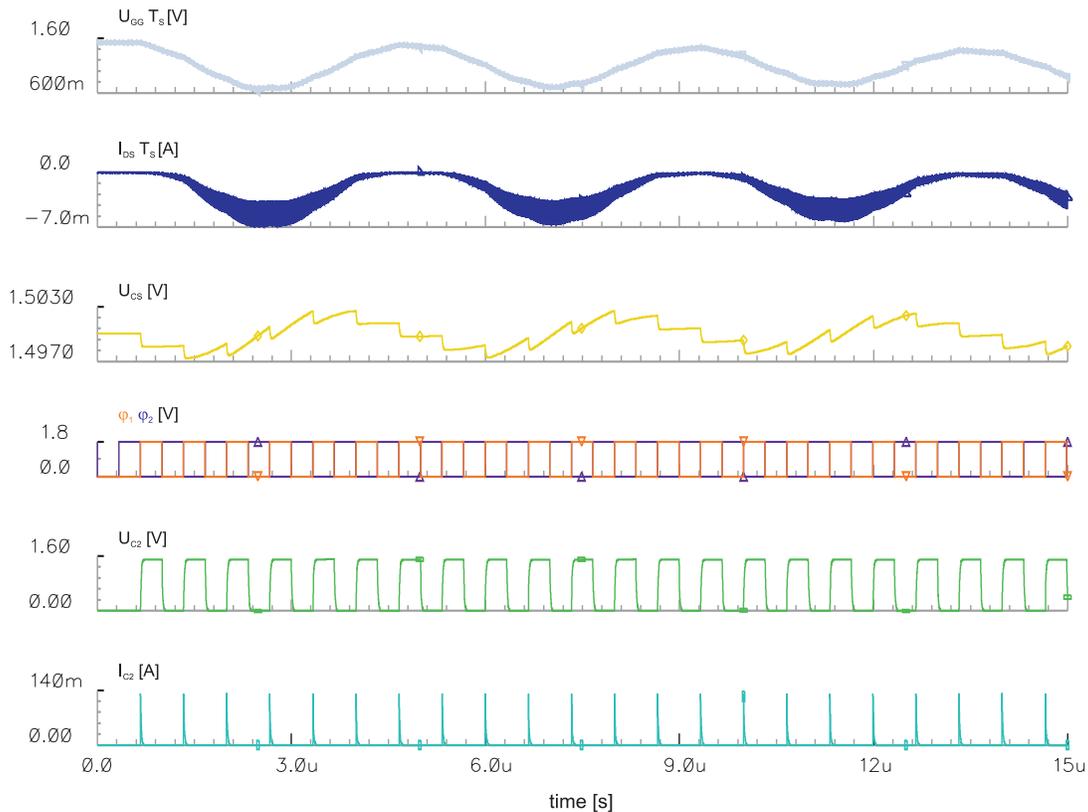
Anhand der Simulationsergebnisse lässt sich die gute Unterdrückung von Versorgungsspannungsschwankungen ( $\leq 100 mV$ ) des Linearreglers feststellen. Die Ausgangsspannung ist von diesen Schwankungen weitgehend unabhängig. Eine unstabilierte Versorgungsspannung wird bei der Regelung gleichzeitig stabilisiert. Die PSRR liegt für Fall 2 bei 46 db und für Fall 3 bei 44 db. Im Vergleich mit den Werten bei konstanter Eingangsspannung ist der Wirkungsgrad bei schwankenden Eingangsspannungen nur wenig kleiner, was mit vermehrter Regelaktivität zu erklären ist. Tabelle 4.7 zeigt Simulationsergebnisse bei wechselnden Lasten (vgl. Abbildung 4.4) und verschiedenen Eingangsspannungen.

$V_{In}$	$V_{Out-Max}$ [V]	$V_{Out-Min}$ [V]	Last	Wirkungsgrad $\eta$
1.	1.5	1.5	$R_1 = R_2 = 2 k\Omega$	0.83
1.	1.5	1.499	$R_3 = 100 k\Omega, C_1 = 1 nF$	0.83
1.	1.5	1.499	$R_4 = 100 k\Omega, C_2 = 1 nF,$ $L_1 = 1 nH$	0.83
2.	1.5	1.499	$R_4 = 100 k\Omega, C_2 = 1 nF,$ $L_1 = 1 nH$	0.81

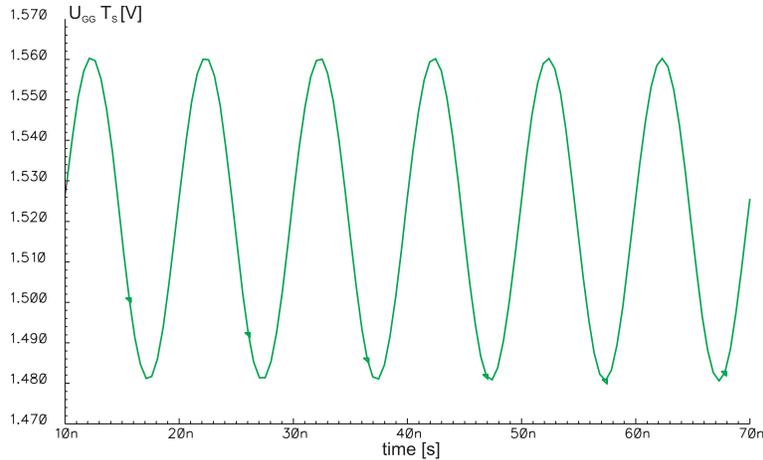
Tabelle 4.7: Wechselnde Last nach Abbildung 4.4

Die Abbildung 4.5 verdeutlicht die typischen Strom- und Spannungsverhältnisse beim Linearregler. Die Taktrate  $\Phi_1$  und  $\Phi_2$  der Schalter in Bild 4.4 beträgt hierbei 1.5MHz. Die Siebkapazität  $C_S = 1\mu F$ , die Referenzspannung  $U_{ref} = 1V$  (Abbildung 4.2), woraus sich eine Ausgangsspannung von  $U_{out} = \frac{3}{2} \cdot U_{ref} = 1.5V$  ergibt. Die Kapazität  $C_2$  wird sehr schnell (durch die nur kleinen Lade- und Entlade-Widerstände von  $10\Omega$ ) ge- und entladen. Der zur Aufladung notwendige Strom fließt fast ausschließlich von der Siebkapazität  $C_S$  zur Lastkapazität  $C_2$ . Durch diesen Ladungstransport sinkt die Spannung über der Siebkapazität ab. Über  $T_S$  wird die Siebkapazität nachgeladen. Während des Ladevorgangs kommt es immer wieder zu großen Stromflüssen zur Lastkapazität, wodurch die Spannung an der Siebkapazität wieder absinkt. Dieses Verhalten resultiert in einer leichten Schwingung der Spannung  $U_{CS}$ .

Die Abbildung 4.6 zeigt das Gate-Potential  $U_{GG}$  des Stelltransistors  $T_S$  vergrößert dargestellt. Hier ist die Reaktion auf den AC-Anteil in der Versorgungsspannung gut zu erkennen ( $V_{DD-AC} = 100mV@100MHz$ ).



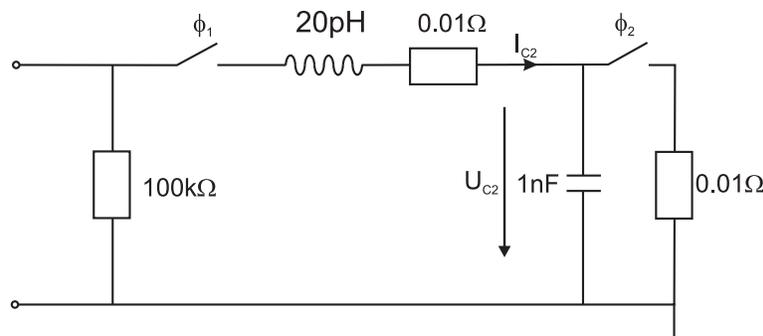
**Abbildung 4.5:** Spannungs- und Stromkurven für eine Beschaltung entsprechend Zeile 4 aus Tabelle 4.7



**Abbildung 4.6:** Gate-Potential  $U_{GG}$  des Stelltransistors  $T_S$  vergrößerte Darstellung aus Abbildung 4.5

### 4.1.3 Einfluss von Strompeaks auf die Regelgüte des linearen DC-DC Konverters

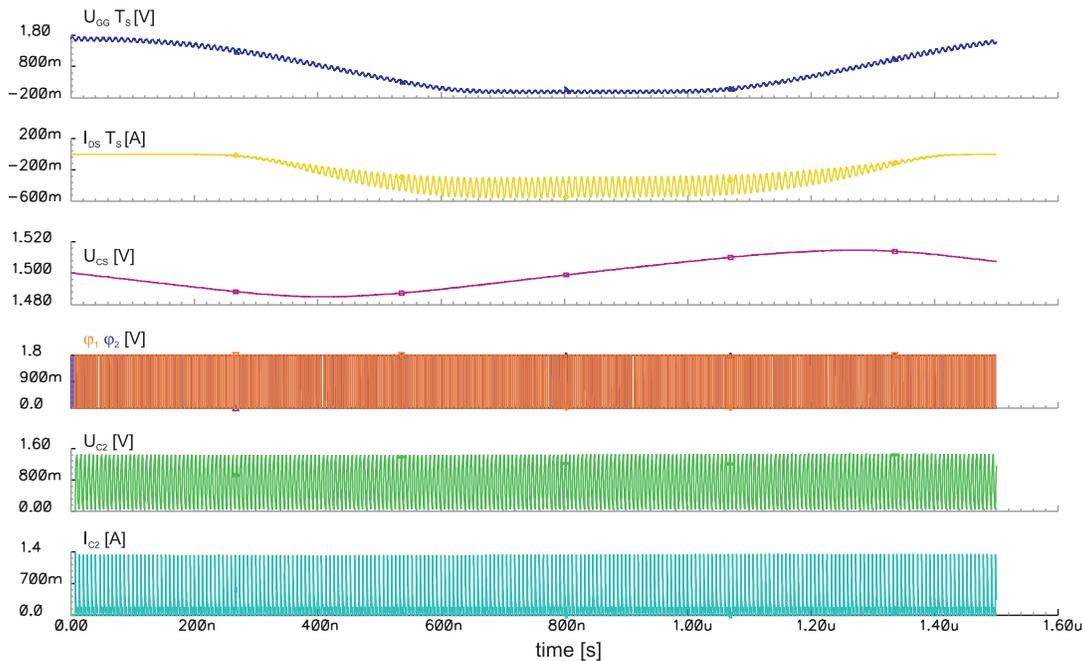
Bei starkem Leistungsverbrauch, mit für komplexe Digitalschaltungen typischen steilflankigen Impulsen im Laststrom, kommt der Linearregler (wie auch andere DC/DC-Konverter) mit einer einzigen zentralen Siebkapazität an seine Grenzen. Die Abbildung 4.8 zeigt die Strom- und Spannungsverläufe beim Linearregler für eine simulierte komplexe Digitalschaltung. Durch die nun geänderten Spezifikationen wird das Lastelement angepasst:



**Abbildung 4.7:** Lastelement zur Simulation einer komplexen Digitalschaltung

$\phi_1$  und  $\phi_2$  takten mit einer Frequenz von  $150\text{MHz}$  und sind nichtüberlappend. Das stellt eine Abweichung von der Realität dar, denn bei realen CMOS-Schaltungen kommt es zu Kurzschlussströmen (vergl. 1.1). Die Siebkapazität ist

mit  $C_S = 1\mu F$  minimal ausgelegt. Kleinere, integrierbare Werte sind in diesem Fall nicht sinnvoll. Durch die nun viel größere Last (durch Parallelschaltung einer Vielzahl von Drain-Source-Strecken mit einer größeren Lastkapazität wegen der Vielzahl von zu be- und entladenden Knoten) ist der Strom, der von der Siebkapazität zur Lastkapazität fließt um Größenordnungen höher. Der Scheitelwert liegt bei  $I_{CS-max} = 1.4A$ . Ein solch hoher Peakstrom mit hoher Steigung ist nicht von außen über Pads und Bonddrähte zu applizieren. Die unvermeidlichen parasitären Induktivitäten stehen dem entgegen. Die Taktfrequenz von  $f_{clock} = 150MHz$  selbst ist vergleichsweise klein. Bei komplexen, schnellen Digitalschaltungen ist es deshalb von entscheidender Bedeutung, die Versorgungsspannung mehrfach lokal (also in unmittelbarer räumlicher Nähe zum Verbraucher, um die seriellen Induktivitäten klein zu halten) durch Siebkapazitäten abzapuffern. Die Anforderungen an diese Stützkapazitäten sind besonders kritisch. Niedrige ESR- und ESL-Werte bei kleinen Abmessungen und hoher Kapazität sind hierbei die Spezifikationen.



**Abbildung 4.8:** Spannungs- und Stromkurven für eine Beschaltung entsprechend Abbildung 4.7

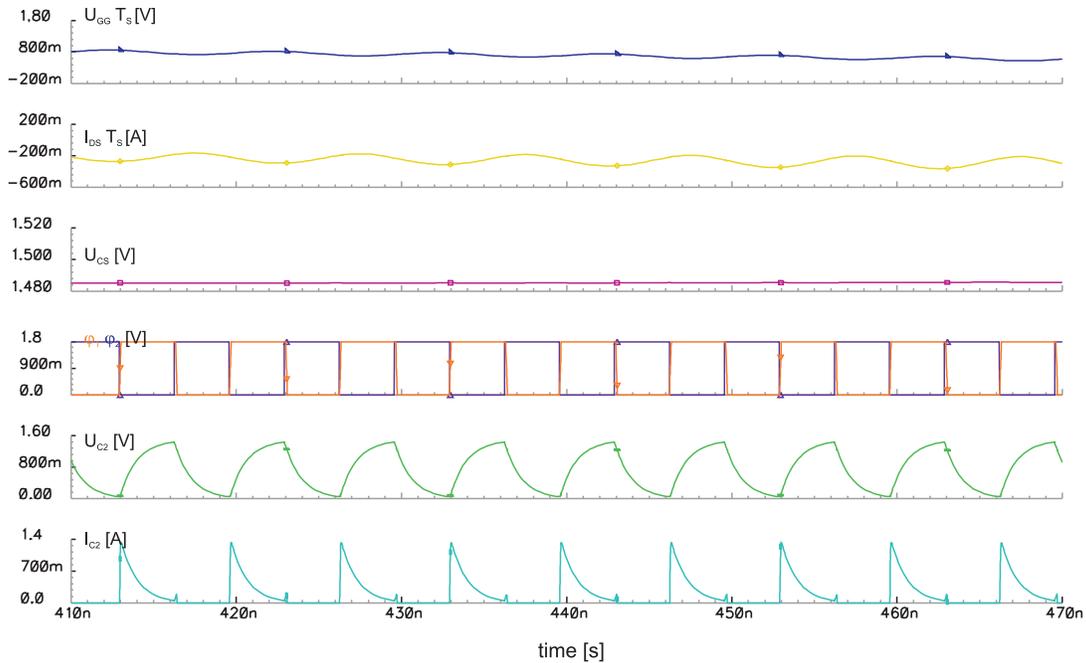


Abbildung 4.9: Vergrößerung aus Abbildung 4.8

Abbildung 4.1.3 zeigt das Package des Intel Pentium 4 Prozessors mit mehreren Stützkapazitäten auf der Unterseite. Die Versorgung solch hochkomplexer Digitalschaltungen ist Thema umfangreicher Untersuchungen. [35]. Integrierte Spannungswandler sind für solche Hochleistungsapplikationen nicht leistungsfähig genug.

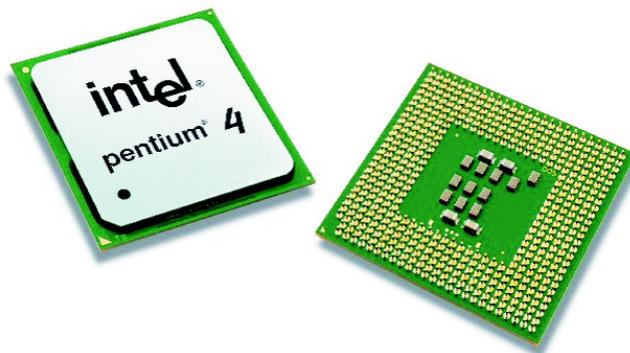
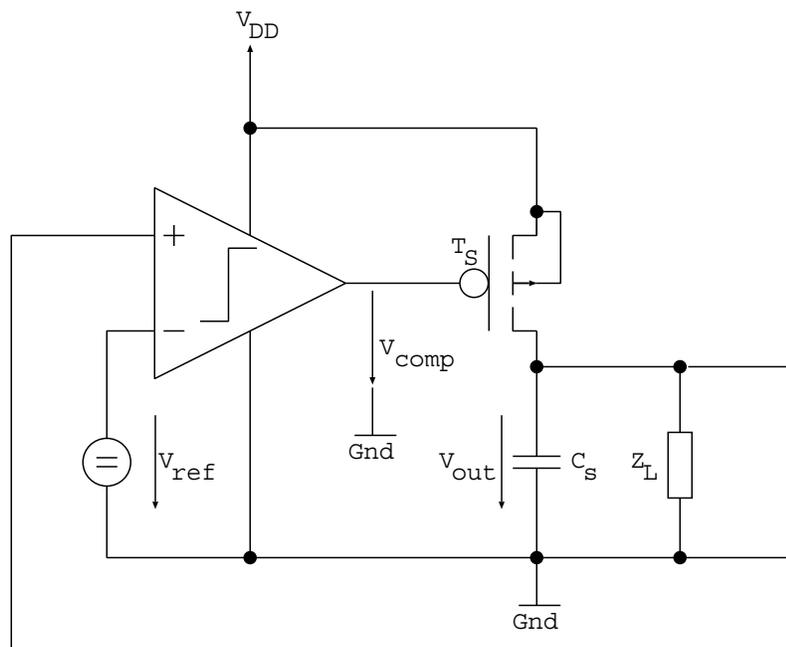


Abbildung 4.10: Package eines Intel-Pentium 4 Prozessors mit Stützkapazitäten

## 4.2 PWM-PFM Regler

Der im Folgenden vorgestellte DC-DC Konverter arbeitet prinzipiell ähnlich wie der Linearregler. Der Differenzverstärker aus Abb.4.1 wird hier durch einen Komparator ersetzt. Dieser vergleicht die (geteilte) Ausgangsspannung mit einer Referenzspannung und schaltet abhängig davon den Transistor  $T_S$  sperrend oder leitend. Das Ausgangssignal des Komparators ist dementsprechend ein Pulsweiten-, Pulsfrequenz-moduliertes digitales Signal (PWM-PFM), dessen Modulationsrate und Frequenz vom momentanen Laststrom  $i_{load}$  und der Dropspannung über  $T_S$  abhängt. Abbildung 4.11 zeigt den schematischen Aufbau dieses Konverters.



**Abbildung 4.11:** Schematischer Aufbau des schaltenden PWM-PFM-Reglers

Der Transistor  $T_S$  arbeitet hier in zwei Zuständen, nämlich entweder im sperrenden Zustand ( $V_{comp} > V_{DD} - V_{T_{T_S}}$ ), oder im bestmöglich leitenden Zustand ( $V_{comp} = Gnd$ ). Die Ladekapazität  $C_S$  fungiert hierbei einerseits als Energiespeicher für den Fall, dass  $T_S$  sperrt, andererseits dämpft sie den Spannungsimpuls als Folge des großen Stromflusses von  $V_{DD}$  über  $T_S$ , wenn  $V_{comp} = Gnd$ . Der Komparator sollte im Sinne eines guten Wirkungsgrades einen kleinen Leistungsverbrauch aufweisen, im Sinne eines niedrigen Ausgangsspannungs-Ripples aber auch ausreichend schnell schalten. Um den Leistungsverbrauch durch die Auf- und Entladung der Gatekapazität von  $T_S$  zu minimieren, ist es vorteilhaft, die maximale Ausgangsspannung des Komparators auf  $V_{comp} \leq V_{DD} - |V_{T_{T_S}}|$  zu

begrenzen. Dennoch wird hier mehr Leistung aufgewendet, als beim Linearkonverter, der das Gate des Stelltransistors angepasst an den Laststrom ansteuert und kleinere Spannungshübe ausführt. Diese Mehrleistung kann unter Umständen im Komparator eingespart werden. Von Vorteil ist hierbei, dass gängige Komparatorstrukturen i.d.R. weniger Fläche in Anspruch nehmen, als leistungsfähige Differenzverstärker. Abbildung 4.12 zeigt die simulierte transiente Wavelform für einen PWM-Konverter.

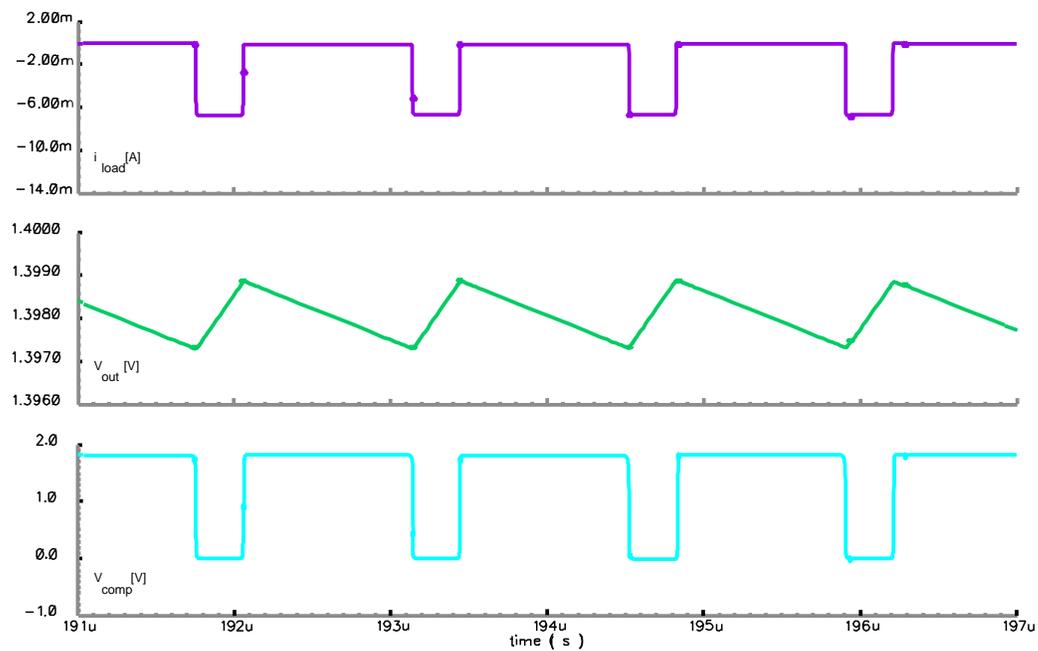
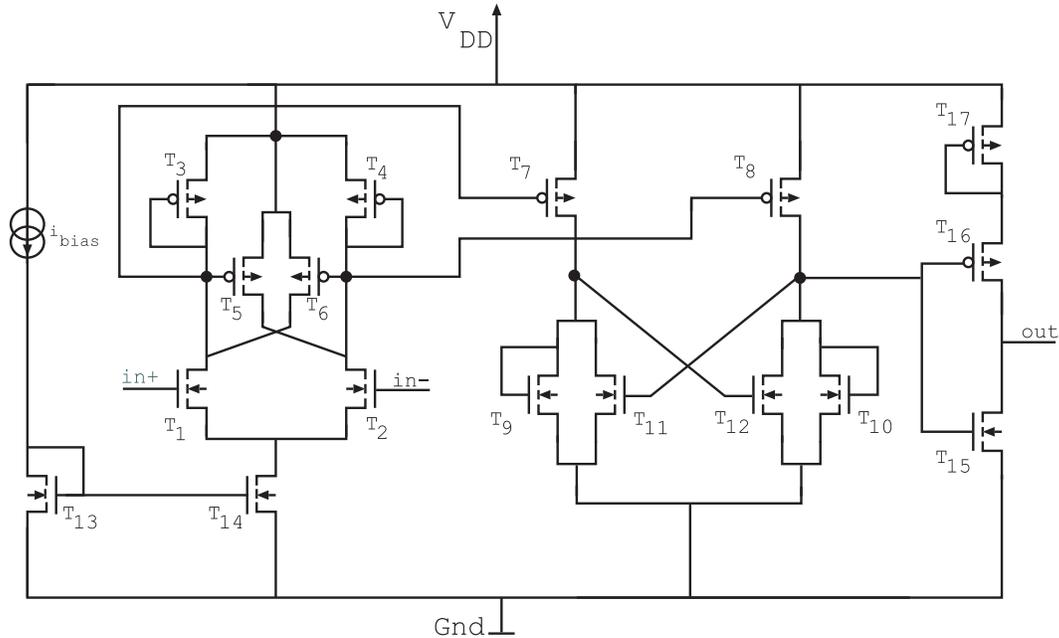


Abbildung 4.12: Transientensimulation PWM-Konverter mit konstanter Last

### 4.2.1 Topologie des PWM-PFM-DC-DC-Konverters

Als Komparator für den PWM-PFM-Konverter wird eine dreistufige Struktur mit zweifacher positiver Rückkopplung gewählt. Diese Struktur vereint niedrigen Leistungsbedarf mit hoher Geschwindigkeit und Auflösung. Abbildung 4.13 zeigt das Schaltbild des Komparators. Die erste Stufe besteht aus einem NMOS Diffe-



**Abbildung 4.13:** Schaltbild des Komparators

renzverstärker ( $T_1, T_2$ ) mit Diodenlast ( $T_3, T_4$ ). Der Arbeitspunkt des Differenzverstärkers wird über den Stromspiegel bestehend aus  $T_{13}$  und  $T_{14}$  eingestellt. Die Transistoren  $T_5$  und  $T_6$  bilden die erste positive Rückkopplung. Die Ausgangssignale des Differenzverstärkers erfahren über die Transistoren  $T_7$  und  $T_8$  eine weitere Verstärkung. Die Last der zweiten Stufe besteht aus den als Diode geschalteten Transistoren  $T_9$  und  $T_{10}$ , unterstützt von der zweiten positiven Rückkopplung ( $T_{11}, T_{12}$ ). Schließlich wird ein Ausgangssignal der zweiten Stufe ausgekoppelt und mit dem Ausgangsinverter ( $T_{15}, T_{16}$ ) erneut verstärkt und geformt. Das Ausgangssignal kann durch den Spannungsabfall über  $T_{17}$   $U_{DS-T_{17}} \approx V_{T_{17}}$  den  $V_{DD}$  Pegel nicht erreichen. Dies ist beabsichtigt, da zum Sperren des Schalttransistors  $T_S$  aus 4.11 ein Gatepotential von  $U_{GG-T_S} > V_{DD} - V_{T-T_S}$  ausreicht. Eine weiteres Aufladen des Gate würde unnötig Leistung umsetzen. Die beiden positiven Rückkopplungsstrukturen sind so dimensioniert, dass eine Rücksetzung des Komparators nicht notwendig ist. Bei zu starker positiver Rückkopplung könnte der Komparator ansonsten in einem Betriebszustand verbleiben.

Diese Maßnahme ermöglicht ein zeitkontinuierliches Auswerten der Eingangssignale, welches für die weitere Verwendung von Vorteil ist.

Der Komparator vergleicht die Ausgangsspannung mit einer Referenzspannung. Unter der Annahme eines Ausgangsspannungsbereichs von  $1V \leq V_{DD} \leq 1.8V$  und der Rückkopplung der Ausgangsspannung über einen Spannungsteiler wie in Abbildung 4.2 ergibt sich ein Eingangsspannungsbereich von  $666mV \leq U_{in} \leq 1200mV$ . Die Ausgangsspannungsschwankungen werden durch den Spannungsteiler ebenfalls mit dem Faktor  $\frac{2}{3}$  zurückgekoppelt. Bei einer erwünschten detektierbaren Amplitude von  $\Delta U_{out} = 1mV$  ergibt die minimale Auflösung des Komparators zu  $666\mu V$  über den gesamten Eingangsspannungsbereich. Diese Spezifikation wird erfüllt (siehe auch Abbildung A.8). Die Simulationsergebnisse und die Dimensionierung des Komparators finden sich in Kapitel A.5

### 4.2.2 Simulationsergebnisse des PWM-PFM-DC-DC-Konverters

Wie schon bei den bereits vorgestellten Schaltungen werden Ausgangsspannungen und Lasten variiert.

$V_{\text{In}}$ [V]	$V_{\text{Out}}$ [V]	$V_{\text{Out-Max}}$ [V]	$V_{\text{Out-Min}}$ [V]	Wirkungsgrad $\eta$
1.8	1.7	1.701	1.699	0.93
1.8	1.6	1.601	1.599	0.87
1.8	1.5	1.502	1.499	0.82
1.8	1.4	1.402	1.399	0.76
1.8	1.3	1.302	1.299	0.71
1.8	1.2	1.203	1.199	0.65

**Tabelle 4.8:** Variation der Ausgangsspannung bei konstanter Last ( $1\text{ k}\Omega$ )

Tabelle 4.8 zeigt die Messwerte bei einer Ausgangslast von  $1\text{ k}\Omega$ . Es wird ein sehr guter Wirkungsgrad bei nur geringen Schwankungen der Ausgangsspannung erreicht. In Tabelle 4.9 sind die Messwerte bei einer Ausgangslast von  $10\text{ k}\Omega$  angegeben. Aufgrund des kleinen Ausgangsstromes ist die maximale Abweichung der Ausgangsspannung vom Sollwert immer  $\leq 1\text{ mV}$ . Allerdings ist der Wirkungsgrad wesentlich schlechter als bei einer Last von  $1\text{ k}\Omega$  (vgl. Tabelle 4.8). Die im Komparator umgesetzte Leistung wird hier dominanter. Auch der PWM-PFM-Konverter wird abschließend bei verschiedenen Bedingungen und einer Ausgangsspannung von  $1,5\text{ V}$  vermessen (Tabelle 4.10). Die Angaben zur Eingangsspannung und zur Last beziehen sich auf Bild 4.4 und Tabelle 4.5.

$V_{\text{In}}$ [V]	$V_{\text{Out}}$ [V]	$V_{\text{Out-Max}}$ [V]	$V_{\text{Out-Min}}$ [V]	Wirkungsgrad $\eta$
1.8	1.7	1.7	1.7	0.80
1.8	1.6	1.6	1.6	0.75
1.8	1.5	1.501	1.5	0.70
1.8	1.4	1.4	1.4	0.64
1.8	1.3	1.301	1.3	0.59
1.8	1.2	1.2	1.2	0.54

**Tabelle 4.9:** Variation der Ausgangsspannung bei konstanter Last ( $10\text{ k}\Omega$ )

$V_{In}$	$V_{Out}$ Max[V]	$V_{Out}$ Min[V]	Last	Wirkungsgrad $\eta$
1.	1.502	1.5	$R_1 = R_2 = 2\text{ k}\Omega$	0.81
1.	1.502	1.499	$R_3 = 100\text{ k}\Omega, C_1 = 1\text{ nF}$	0.82
1.	1.502	1.498	$R_4 = 100\text{ k}\Omega, C_2 = 1\text{ nF},$ $L_1 = 1\text{ nH}$	0.82
2.	1.502	1.498	$R_4 = 100\text{ k}\Omega, C_2 = 1\text{ nF},$ $L_1 = 1\text{ nH}$	0.83

Tabelle 4.10: Messungen bei wechselnden Bedingungen

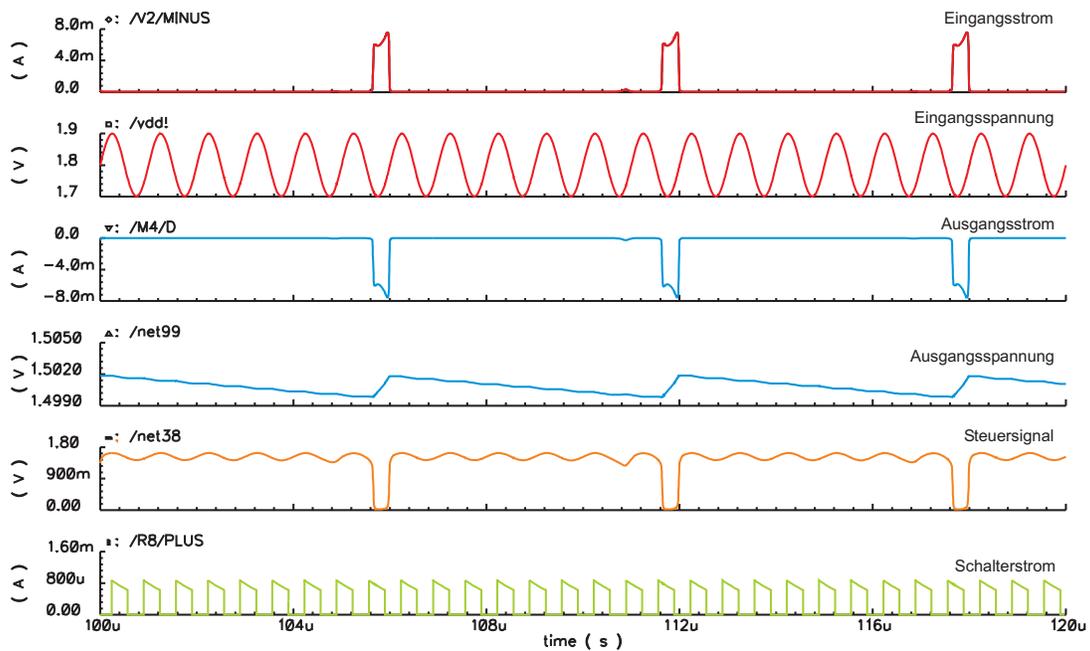


Abbildung 4.14: Spannungs- und Stromkurven für eine Beschaltung entsprechend Zeile 4 aus Tabelle 4.10

## 4.3 Der Buck-Konverter

Beim Abwärts- oder Buck-Konverter ist es nur möglich, die Eingangsspannung zu reduzieren. Der wesentliche Vorteil von Buckkonvertern gegenüber Linearreglern ist die Verwendung eines zweiten Energiespeichers, nämlich einer Induktivität. Wird beim Linearregler überschüssige Energie in Form von Wärme in einem variablen Vorwiderstand  $T_S$  umgesetzt, wo sie dem System verloren geht, ist diese überschüssige Energie bei induktiven Buckkonvertern im magnetischen Feld einer Spule gespeichert und wird dem System in einer Taktphase gepulst wieder zugeführt. Dies erhöht prinzipiell den Wirkungsgrad bei großen Dropspannungen und erreicht theoretisch 100%.

### 4.3.1 Funktionsweise des induktiven Buckkonverters

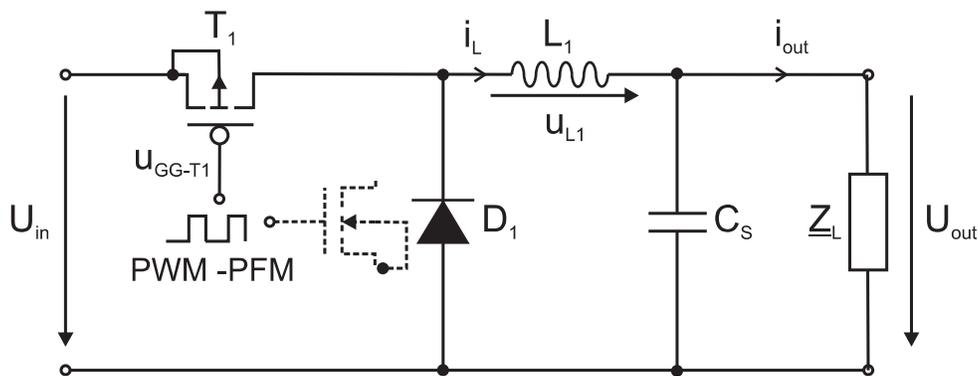


Abbildung 4.15: Prinzipschaltbild eines Abwärts-Wandlers

Der PMOS-Transistor  $T_1$  wird mit einem gepulsten Signal (PWM bzw. PWM-PFM) angesteuert. Ist der Transistor  $T_1$  durchgeschaltet ( $u_{GG-T1} = 0$ ), liegt am LC-Glied näherungsweise die Eingangsspannung  $U_{in}$  an. Der Spulenstrom steigt, wie in Bild 4.16 zu erkennen, linear an. Nach dem Sperren des Transistors  $T_1$  wird die im aufgebauten Magnetfeld der Spule gespeicherte Energie über die Diode  $D_1$  an die Last abgegeben. Das Magnetfeld bricht zusammen, was nach dem Induktionsgesetz eine Spannung induziert, die entgegen der Ursache anliegt (vergl. Abbildung 4.16). Der Spulenstrom vermindert sich wieder. Der Stromfluss geschieht in der Sperrphase von  $T_1$  von der Spule in Richtung Kondensator  $C_S$  (Abbildung 4.15), wodurch die Ausgangsspannung so lange weiter steigt, bis der Strom, den die Spule liefern kann, kleiner ist, als der von der momentanen Belastung ( $Z_L$  und  $C_S$ ) zum Aufrechterhalten der Ausgangsspannung notwendige Strom. Danach ist die Ladungsbilanz von zugeführten Ladungsträgern durch die Spule und abgeführten Ladungsträgern durch die Last negativ, so dass die Ausgangsspannung  $U_{out}$  über  $C_S$  sinkt.

Die fast verlustfreie Überführung von magnetischer Energie der Spule in elektrische Feldenergie von  $C_L$  erklärt den möglichen hohen Wirkungsgrad, der in der Literatur mit bis zu 90% und größer angegeben wird. Der Spulenstrom geht in Abbildung 4.16 nicht auf Null zurück, weshalb hier von kontinuierlichem Betrieb (CCM continuous current mode) gesprochen wird. Wird sämtliche Energie der Spule in der Sperrphase von  $T_1$  in der Last umgesetzt, spricht man vom lückendem Betrieb (DCM discontinuous current mode). Der Spulenstrom erreicht dabei den Nullwert. In diesem Fall könnte Ladung des Kondensators  $C_S$  nicht nur in die Last fließen, sondern auch durch die Spule ( $i_{L1} < 0$ ). Die Diode  $D_1$  verhindert dies. Wird im Interesse eines besseren Wirkungsgrads die Diode durch einen Transistorschalter ersetzt ( $R_{on} \ll U_D$ ), muss für diesen Fall Sorge getragen werden, dass keine Ladung des Kondensators direkt über Spule und Transistor fließen kann. Die Funktion der Diode muss elektronisch nachgebildet werden (siehe dazu Kapitel 4.4.2).

Die Ausgangsspannung des Abwärts-Wandlers ergibt sich idealisiert betrachtet zu

$$U_{out-buck} = \frac{t_{ein}}{T} \cdot U_{in} = D \cdot U_{in} \quad (4.7)$$

mit  $D = \frac{t_{ein}}{T}$  (Dutycycle).

Die Ausgangsspannung  $U_{out-buck}$  kann nach Gleichung 4.7 alleine über das Tastverhältnis des Steuersignals (PWM) eingestellt werden. Somit wäre bei einer konstanten Eingangsspannung eine Steuerung theoretisch möglich, da die Last nach Gleichung 4.7 keinen Einfluss hat. In der Realität ergeben sich Verluste durch Schaltwiderstände der Transistoren, die lastabhängig sind, sowie der Flussspannung der Diode. Deshalb ist eine Regelung der Ausgangsspannung der einzig sinnvolle Weg.

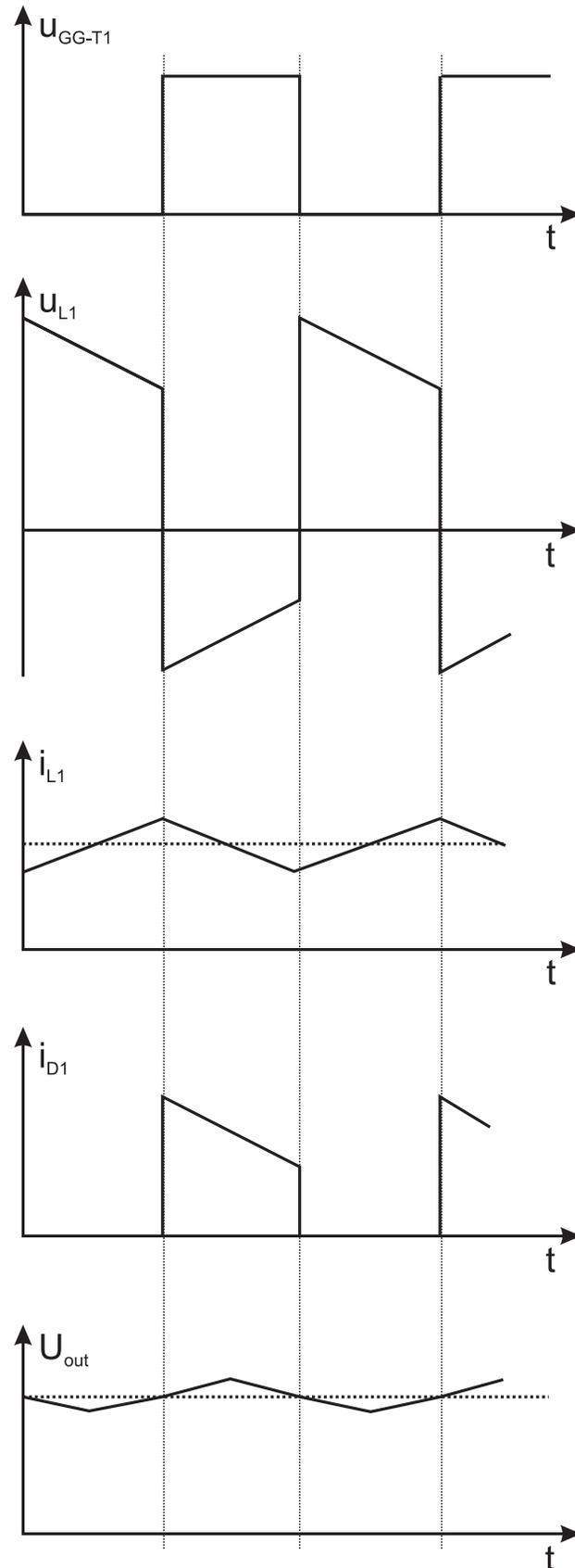


Abbildung 4.16: Ströme und Spannungen des Buck-Konverters aus Abb. 4.15

## 4.4 Buck-Konverter mit variablem Taktverhältnis

Das Konzept dieses Buck-Konverters hat eine geringe Komplexität. Das Steuersignal  $u_{GG}$  wird hierbei durch einen Komparator erzeugt, welcher die Ausgangsspannung ständig mit einer Referenzspannung vergleicht (Bild 4.17). Das Ausgangssignal des Komparators ist dementsprechend ein PWM-PFM-Signal. Die Rückkopplung des Ausgangssignals erfolgt wie bei den schon beschriebenen linearen DC-DC-Wandlern über einen Spannungsteiler im Verhältnis 2 zu 1. Fällt die an den Komparator zurückgekoppelte Ausgangsspannung unter die Referenzspannung, so erzeugt der Komparator  $V_{DD}$  am Ausgang. Über den Inverter, welcher zum Treiben des Schalttransistors benötigt wird, wird der PMOS-Transistor ( $T_1$ ) durchgeschaltet, und der Strom durch die Spule  $L_1$  steigt an. Überschreitet die Ausgangsspannung den Sollwert, so wird der Transistor gesperrt und der Spulenstrom nimmt wieder ab. Das Tastverhältnis der Transistorsteuerung wird also an die jeweilige Belastung angepasst und die Ausgangsspannung dementsprechend auf den Sollwert geregelt. Die Topologie dieses Konverters ähnelt der des in Kapitel 4.2 vorgestellten geschalteten Linearkonverters. Der Vorteil hier besteht in der Verwendung der Spule als Energiespeicher. Sie reduziert die Dropspannung über  $T_1$  für den Fall, dass  $T_1$  leitet und speist die Last, für den Fall, dass  $T_1$  nicht leitet und trägt somit zum höheren Wirkungsgrad bei.

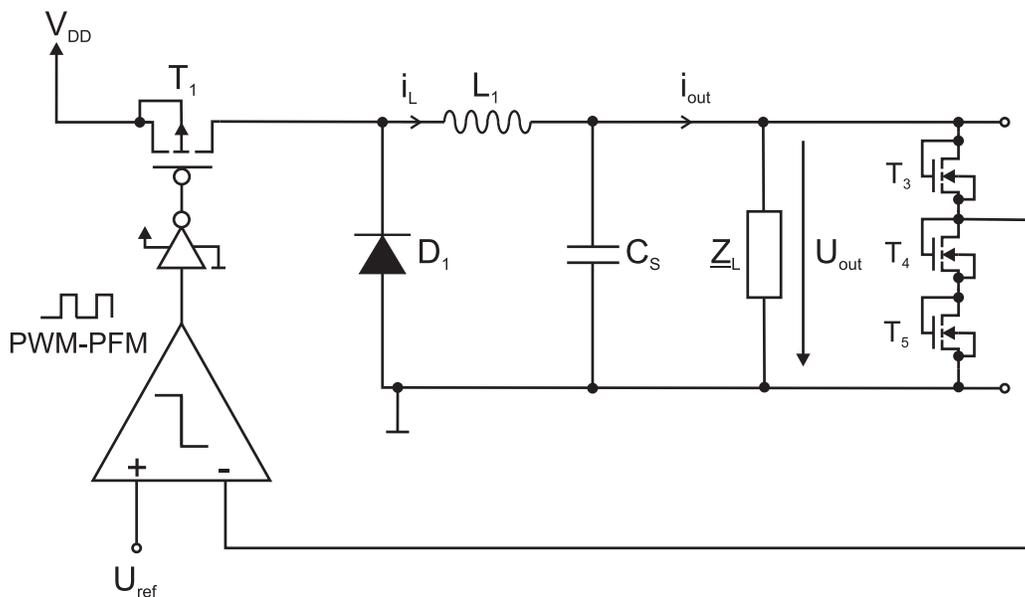


Abbildung 4.17: Buck-Konverter mit Komparator zur PWM-Erzeugung

### 4.4.1 Simulationsergebnisse

Eine Simulation bei konstanter Last ( $1\text{ k}\Omega$ ) liefert die in Tabelle 4.11 aufgeführten Messwerte.

$V_{\text{In}}$ [V]	$V_{\text{Out}}$ [V]	$V_{\text{Out-Max}}$ [V]	$V_{\text{Out-Min}}$ [V]	Wirkungsgrad $\eta$
1.8	1.7	1.701	1.698	0.94
1.8	1.6	1.601	1.599	0.89
1.8	1.5	1.501	1.499	0.87
1.8	1.4	1.402	1.399	0.84
1.8	1.3	1.302	1.299	0.79
1.8	1.2	1.203	1.199	0.75

**Tabelle 4.11:** Variation der Ausgangsspannung bei konstanter Last ( $1\text{ k}\Omega$ )

Über den gesamten Ausgangsspannungsbereich sind die Schwankungen der Ausgangsspannung sehr klein. Der Wirkungsgrad ist bei hohen Ausgangsspannungen vergleichsweise gut, fällt jedoch bis auf 75% bei  $1.2\text{ V } V_{\text{Out}}$  ab. Bei Variation der Last ergeben sich die in Tabelle 4.12 gezeigten Wirkungsgrade.

$V_{\text{Out}}$ [V]	Wirkungsgrad			
	$600\ \Omega$	$1\text{ k}\Omega$	$5\text{ k}\Omega$	$10\text{ k}\Omega$
1.7	0.94	0.94	0.82	0.72
1.6	0.90	0.89	0.77	0.65
1.5	0.87	0.87	0.73	0.6
1.4	0.83	0.84	0.69	0.55
1.3	0.80	0.79	0.62	0.50
1.2	0.76	0.75	0.59	0.44

**Tabelle 4.12:** Wirkungsgrad bei verschiedenen Lasten und Ausgangsspannungen

Aufgrund der Dimensionierungen der Bauteile, welche den bestmöglichen Wirkungsgrad liefern, ist dieser Konverter in der Lage, eine Last bis zu  $600\ \Omega$  zu treiben. Bei  $600\ \Omega$  ist der Wirkungsgrad nahezu identisch mit dem bei einer Ausgangslast von  $1\text{ k}\Omega$ . Wird die Last von  $1\text{ k}\Omega$  auf  $5\text{ k}\Omega$  verringert, so fällt der Wirkungsgrad im Durchschnitt um 14.3%. Verkleinert man die Last auf  $10\text{ k}\Omega$ , sinkt er um weitere 12.6%. Der Wirkungsgrad des Konverters verschlechtert sich stark mit sinkender Last. Die doch recht deutliche Abnahme des Wirkungsgrades bei kleineren Lastströmen beruht auf dem Leistungsverlust durch die Diode  $D_1$ . Ist  $T_1$  gesperrt, so fließt der gesamte Ausgangsstrom über  $D_1$ . An der Diode fällt eine Spannung von ca.  $0.7\text{ V}$  ab, was einen Leistungsumsatz in ihr zur

Folge hat. Um diese Verluste zu verringern, wird im folgenden Abschnitt eine Schaltung mit einem Transistor als „aktive Diode“ vorgestellt. Das Verhalten des Buck-Konverters mit variablem Taktverhältnis bei wechselnden Lasten und Schwankungen in der Versorgungsspannung wird im Anschluss bei einer Ausgangsspannung von 1.5 V untersucht. Die Simulationsergebnisse finden sich in Tabelle 4.13. Die in Tabelle 4.13 und allen folgenden Tabellen Lastkombinationen und Eingangsspannungen beziehen sich auf die Abbildung 4.4 und Tabelle 4.5.

$V_{In}$	$V_{Out-Max}$ [V]	$V_{Out-Min}$ [V]	Last	Wirkungsgrad $\eta$
1.	1.501	1.499	$R_1 = R_2 = 2 k\Omega$	0.85
1.	1.501	1.5	$R_1 = R_2 = 10 k\Omega$	0.65
1.	1.5	1.5	$R_1 = R_2 = 20 k\Omega$	0.50
1.	1.501	1.5	$R_3 = 100 k\Omega, C_1 = 1 nF$	0.75
1.	1.501	1.5	$R_4 = 100 k\Omega, C_2 = 1 nF,$ $L_1 = 1 nH$	0.75
1.	1.501	1.5	$R_4 = 20 k\Omega, C_2 = 1 nF,$ $L_1 = 1 nH$	0.77
2.	1.5	1.5	$R_4 = 100 k\Omega, C_2 = 1 nF,$ $L_1 = 1 nH$	0.73

**Tabelle 4.13:** Messungen bei wechselnden Bedingungen

### 4.4.2 Buck-Konverter mit variablem Taktverhältnis und aktiver Diode

In Bild 4.18 ist die Diode durch einen Transistor mit eigener Steuerlogik ersetzt worden. Diese muss durch eine entsprechende Ansteuerung des Transistors  $T_2$  die Funktionsweise der Diode nachbilden.

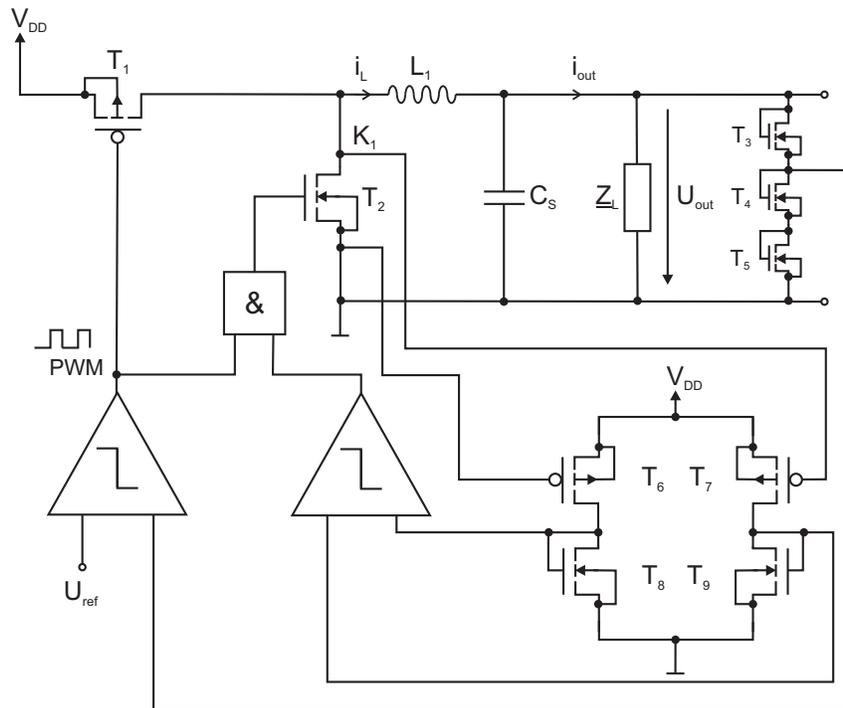


Abbildung 4.18: Buck-Konverter mit aktiver Diode

Das bedeutet, dass der Transistor  $T_2$  nur dann leiten darf, wenn  $T_1$  gesperrt ist und  $U_{DS}$  von  $T_2$  negativ ist. Das Potential an  $K_1$  ergibt sich in der Sperrphase von  $T_1$  zu:

$$U_{K_1} = 0 - i_L \cdot R_{DS-T_2} \quad (4.8)$$

Die Abfrage eines negativen Potentials am Knoten  $K_1$  ist hierbei problematisch, da bei der verwendeten Technologie keine symmetrische Versorgungsspannung zum Einsatz kommt. Deshalb muss sowohl das Ground-Potential, als auch das Potenzial am Knoten  $K_1$  über analoge Level-Shifter ( $T_6$ ,  $T_8$  und  $T_7$ ,  $T_9$ ) an die Eingänge des Komparators geführt werden. Simulationen zeigen jedoch, dass durch den Einsatz von Level-Shiftern, trotz der Verwendung eines Transistors statt der Diode, der Wirkungsgrad weiter sinkt, da in diesen ein

kontinuierlicher Strom fließt. Es liegt hierbei ein Trade-Off zwischen Leistungsumsatz und Schnelligkeit in den Level-Shiftern vor. Der zusätzliche Komparator ist ein weiterer Leistungsumsetzer, der den Wirkungsgrad bei kleinen Lastströmen deutlich begrenzt. Bei höheren Lastströmen ist mit diesem Konzept eine Verbesserung des Wirkungsgrades zu erzielen. Die Beschränkung auf Low-Power Applikationen macht aber eine weitere Verfolgung dieses Konzepts sinnlos.

## 4.5 Buck-Konverter mit Duty-Cycle Generator

Eine weitere Möglichkeit, die Duty-Cycle einzustellen, ist die Verwendung eines Duty-Cycle-Generators. Zur Erzeugung eines variablen Taktverhältnisses wird ein Sägezahn-generator (chopper) entwickelt. Eine Kapazität wird über eine Konstantstromquelle aufgeladen. Da für die Spannung, die über einem Kondensator anliegt, für konstanten Ladestrom

$$U_C(t) = \frac{1}{C} \cdot I \quad (4.9)$$

gilt, ergibt sich über dem Ladekondensator eine linear mit der Zeit ansteigende Spannung, deren Steigung von der Größe des Ladestroms und der Größe der Kapazität abhängt. Die Spannung über dem Ladekondensator wird detektiert und einem Komparator zugeführt, der sie mit einer Referenzspannung vergleicht. Beim Erreichen der Referenzspannung schaltet der Komparator um - ( $V_{comp1-out} = high$ ), woraufhin  $T_1$  leitend wird und die Ladekapazität vollständig entlädt. Da der Komparator die Ladespannung des Kondensators permanent vergleicht, ist es notwendig, dass beim Umschalten des Komparators dieser nicht sofort wieder zurückschaltet, sobald die Spannung des Ladekondensators nur wenig unterhalb der Schaltschwelle liegt. Dies wird durch eine Verzögerung bestehend aus einer Inverterkette und einem NAND-Gatter erreicht (analog zu Abbildungen 3.20). Somit ergibt sich eine periodische Sägezahnspannung, deren Frequenz von der Referenzspannung, dem Ladestrom und der Ladekapazität abhängt. Abbildung 4.19 zeigt diese Schaltung schematisch.

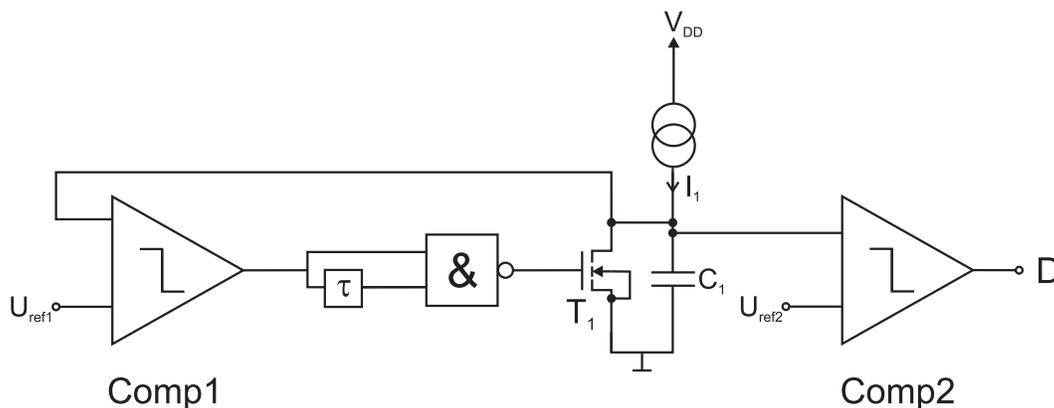
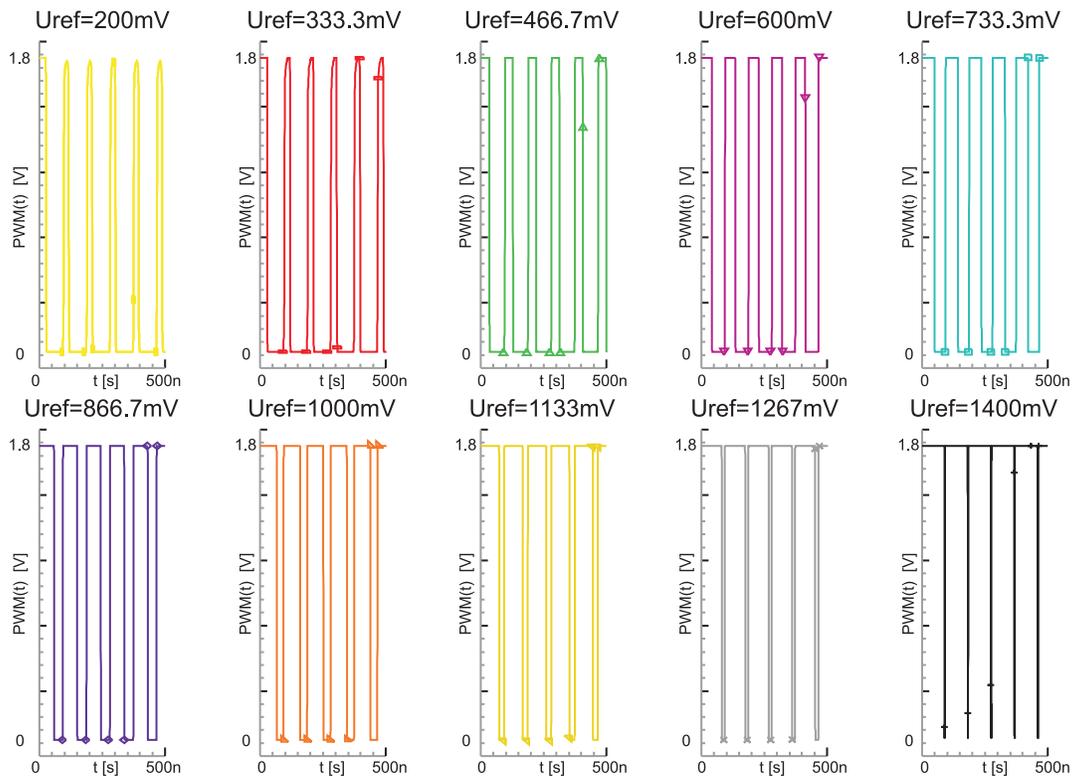


Abbildung 4.19: PWM-Generator

Zur Erzeugung des PWM-Signals (D) wird diese Sägezahnspannung mit einem zweiten Komparator mit einer zweiten Referenzspannung ( $U_{ref2}$ ) verglichen. Dieser Vergleich resultiert in einem PWM-Signal, deren Taktverhältnis von der Referenzspannung (die nicht größer sein darf, als die maximale Sägezahnspan-

nung) abhängt. Abbildung 4.20 zeigt das Ausgangssignal der Schaltung nach Abbildung 4.19 für verschiedene Referenzspannungen  $U_{ref2}$ .



**Abbildung 4.20:** Ausgangssignal des PWM-Generators für verschiedene Steuerspannungen

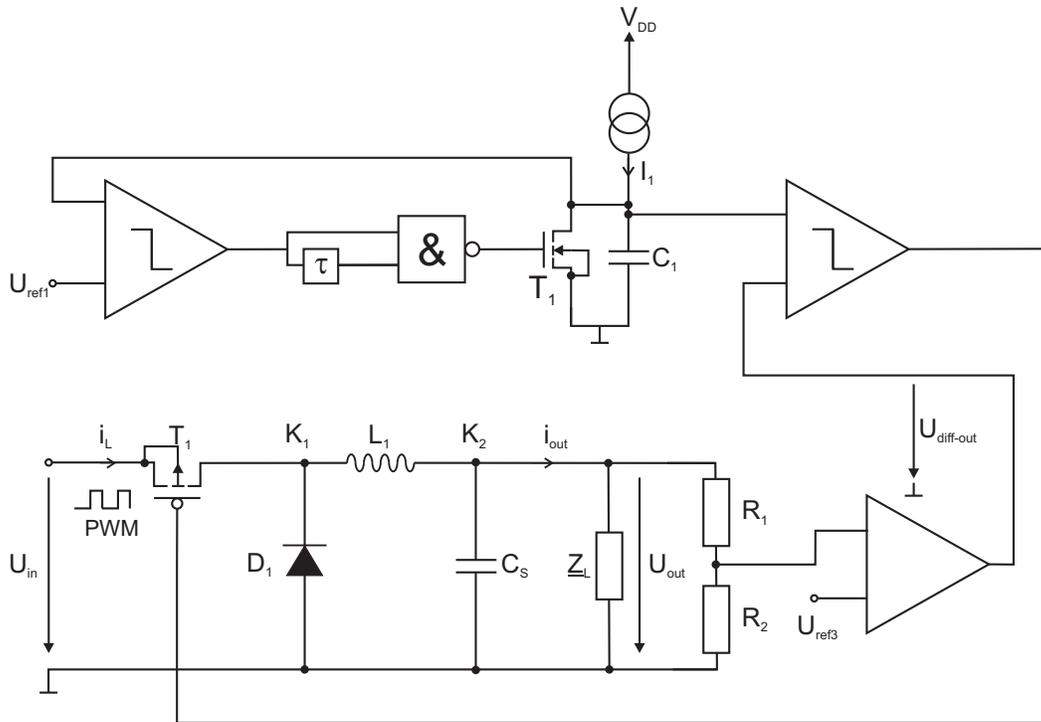


Abbildung 4.21: Schema des Buckkonverters mit Duty-Cycle-Generator

Der Duty-Cycle-Generator wird im Folgenden dazu genutzt, einen geschlossenen Regelkreis aufzubauen. Abbildung 4.21 verdeutlicht den Aufbau. Eine Simulation mit ohmscher Belastung von  $R_L = 1k\Omega$  einer Induktivität der Spule von  $L_1 = 100\mu H$  und einer Siebkapazität von  $C_s = 1\mu F$  zeigt Abbildung 4.22. Nach einem Einschwingvorgang ist hier gut das periodische Ansteuersignal für die PWM-Erzeugung ( $U_{diff-out}$ ) und die Reaktion des Duty-Cycle-Generators zu beobachten. Die Periodizität der Vorgänge resultiert aus den statischen Randbedingungen wie konstante Eingangs- und Ausgangsspannung ( $U_{in} = 1.8V$  und  $U_{out} = 1.4V$ ), sowie aus der konstanten ohmschen Last. Der Ripple ergibt sich hierbei zu  $U_{ripple} = 9mV_{pp}$ , was die Vergrößerung der Abbildung 4.22 verdeutlicht (Abbildung 4.23). Anders, als in den beiden vorangehenden Konzepten, handelt es sich um ein reines PWM-Signal, da die Frequenz des Signals vom Chopper vorgegeben ist.

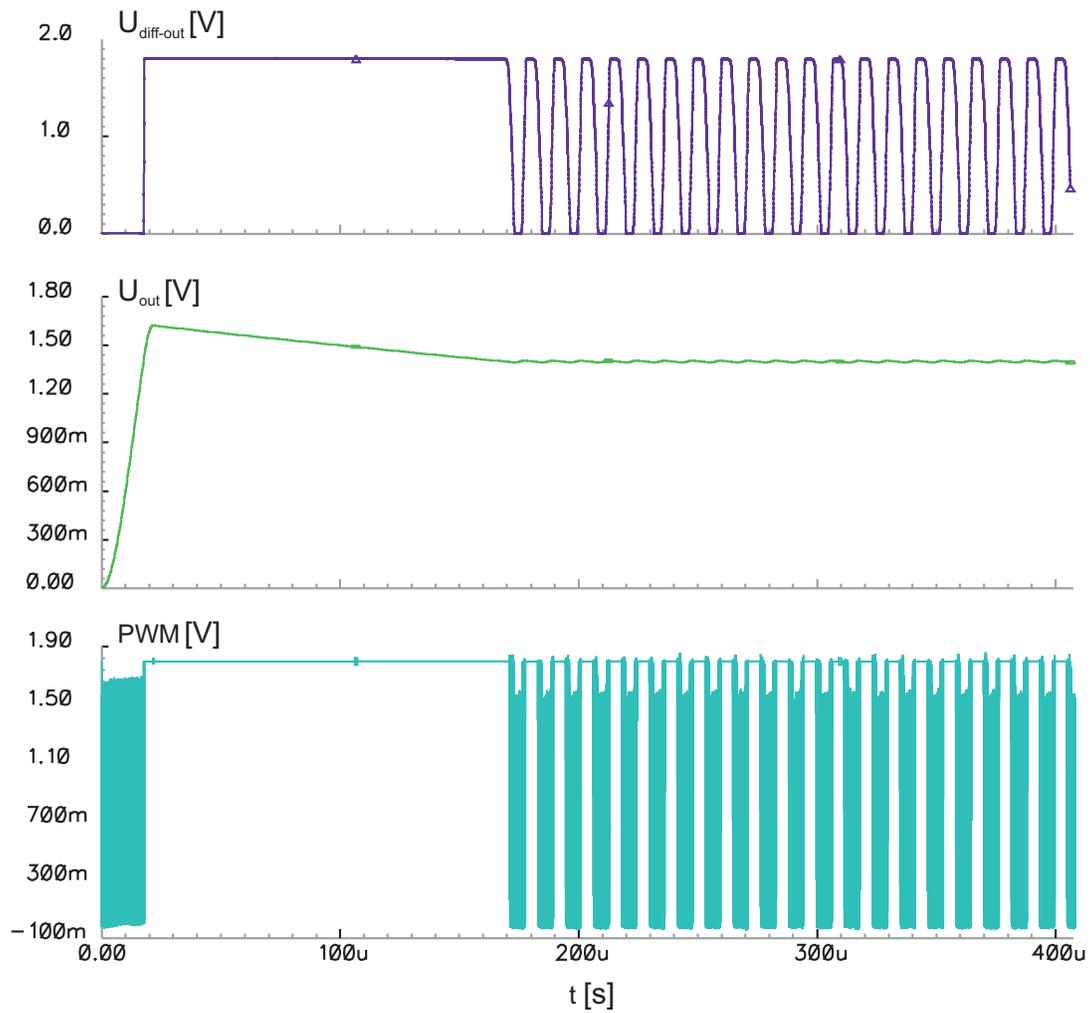


Abbildung 4.22: Simulation des Buckkonverters nach Abbildung 4.21

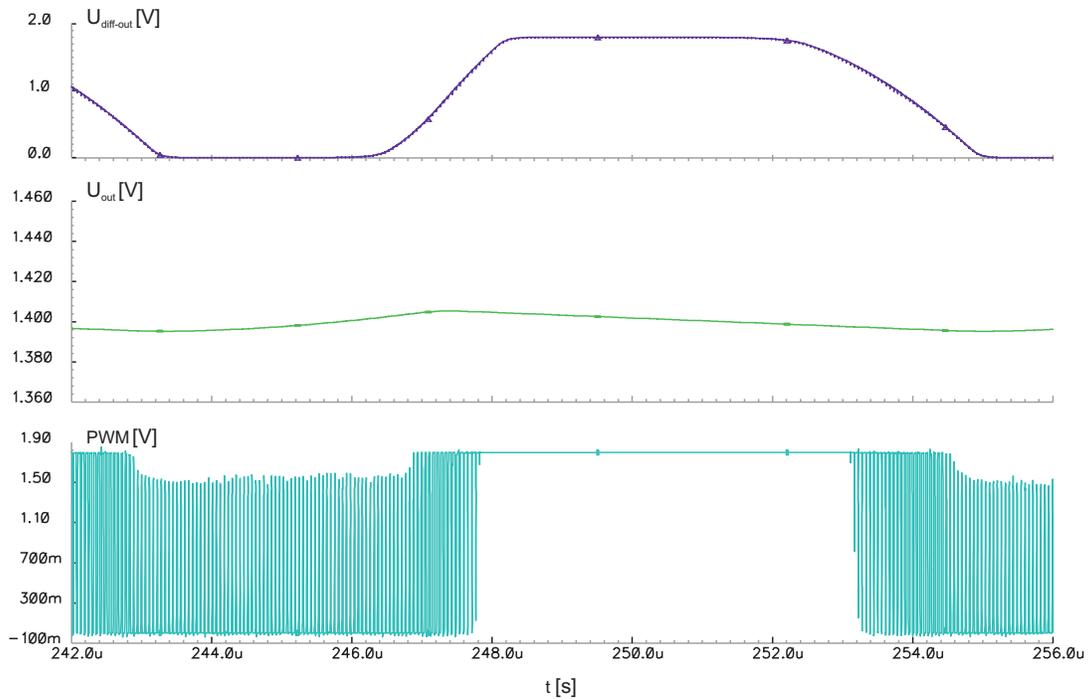


Abbildung 4.23: Vergrößerung der Abbildung 4.22

Ein entscheidender Nachteil des Konzepts ist auch hier die Verwendung der Diode  $D_1$ . Wie jedoch schon in Kapitel 4.4.2 beschrieben, ist die Ansteuerung eines Transistors, der die Diode  $D_1$  ersetzt, problematisch. Ein anderer Nachteil für Systeme mit kleiner Leistungsaufnahme ist die hohe Komplexität des Systems. Die Ansteuerung über Chopper und mehreren Komparatoren setzt selbst eine nicht unerhebliche Leistung um. Der Wirkungsgrad des Systems ist im Vergleich mit den anderen gezeigten schaltenden und über einen großen Ausgangsspannungsbereich sogar im Vergleich zum Linearregler schlechter. Aus diesem Grund wird auf eine weitere Untersuchung des Systems verzichtet.

### 4.5.1 Buck-Konverter mit alternierenden Takten

Der in Bild 4.24 dargestellte Abwärts-Wandler arbeitet mit einem festen Taktverhältnis. Dieses wird einmal in seiner ursprünglichen Form mit einem Tastverhältnis von 1 zu 7 (1 Takt High-Phase und 7 Takte Low-Phase) und einmal in invertierter Form (Taktverhältnis 7 zu 1) auf ein AOI-Gatter (AND-OR-INV) geschaltet. Über einen Komparator wird je nach Bedarf eines der beiden Signale ausgewählt. Es wird derselbe Komparator wie bei der in Kapitel 4.4 beschriebenen Realisierungsform verwendet. Da die Ausgangsstufe des Komparators nun ein AOI-Gatter treibt, wird dem Transistor  $T_1$  in Abbildung 4.24 ein Treiberinverter vorgeschaltet. Die gewählten Taktverhältnisse orientieren sich an dem vorgesehenen Ausgangsspannungsbereich und werden mit Hilfe von Simulationen mit festem Taktverhältnis evaluiert.

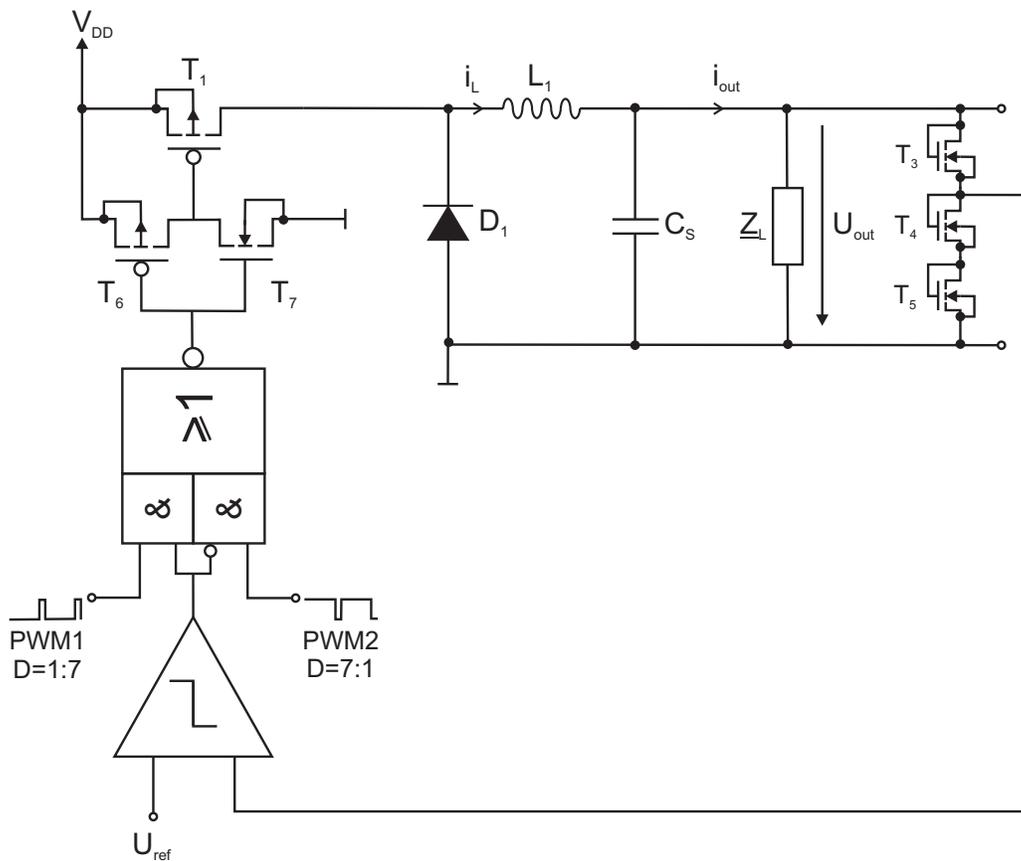


Abbildung 4.24: Buck-Konverter mit alternierenden Takten

Die PWM-Signale werden mit einem 3-bit-Zähler und einem NAND-Gatter realisiert (Bild 4.25).



## 4.6 Der Boost-Konverter

### 4.6.1 Die Funktionsweise des Boost-Konverters

Für den Boost-Konverter werden die gleichen Bauelemente wie für den Buck-Konverter benötigt. Jedoch erhält man durch eine andere Anordnung eine Ausgangsspannung, die größer als die angelegte Eingangsspannung ist.

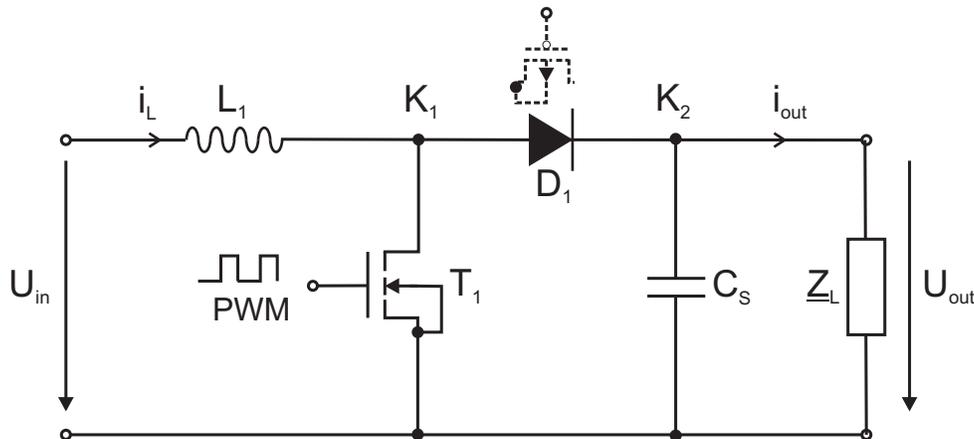
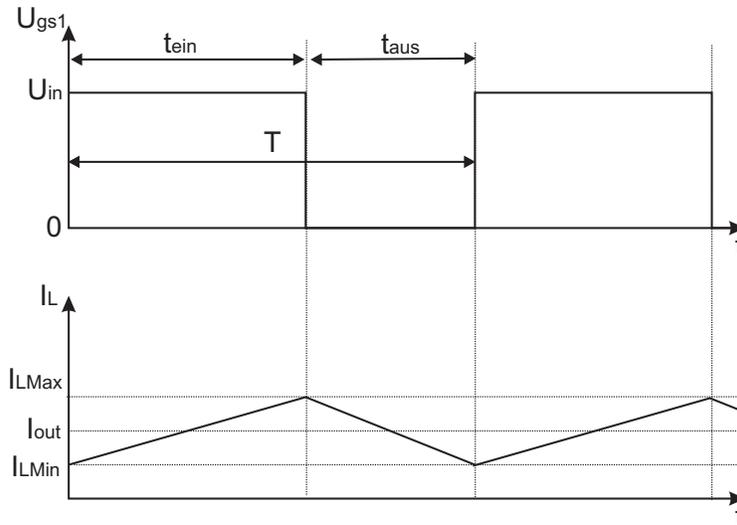


Abbildung 4.26: Prinzipschaltbild eines Aufwärts-Wandlers

Leitet  $T_1$ , so fließt ein Strom durch die Spule  $L_1$ , welcher gemäß  $\frac{di}{dt} = \frac{U_L}{L_1}$  mit  $U_L \approx U_{in}$  unter Vernachlässigung der Transistorspannung  $U_{DS1}$  linear ansteigt. In der Sperrphase von  $T_1$  kehrt sich die Spannung über der Spule  $L_1$  gemäß dem Induktionsgesetz  $U_i(t) = -L \cdot \dot{I}(t)$  um, und addiert sich zu der Eingangsspannung  $U_{in}$ . Die Diode  $D_1$  schaltet nun durch, da der Knoten  $K_1$  (Abb. 4.26) nicht mehr durch den Transistor  $T_1$  mit dem Massepotential verbunden ist. Der Spulenstrom fließt durch die Parallelschaltung von Lastimpedanz  $Z_L$  und Glättungskapazität  $C_S$ . Der Strom  $i_L$  nimmt linear ab, und die in der Spule gespeicherte Energie wird dem Ausgang zugeführt. Der Wirkungsgrad solcher Wandler wird in der Literatur, vergleichbar mit dem Buck-Konverter, zu 90% und größer angegeben. Die Ansteuerung des Transistors erfolgt, analog zum Abwärts-Wandler, üblicherweise mit einem PWM-Signal. Die Ausgangsspannung ergibt sich zu:

$$U_{out} = \frac{T}{t_{aus}} \cdot U_{in} = \frac{1}{1-D} \cdot U_{in} \quad (4.10)$$

Prinzipiell kann auch hier die Diode  $D_1$  durch einen Transistor ersetzt werden, welcher komplementär zu  $T_1$  ausgeführt wird (PMOS) und über dasselbe Signal gesteuert wird. Allerdings muss zusätzlich darauf geachtet werden, dass kein Strom vom Ausgang zurückfließt (vgl. Sperrverhalten der Diode). Der beschriebene Aufbau eines Boost-Konverters würde in dieser Form (Abbildung



**Abbildung 4.27:** Steuersignal und idealisierte Stromkurve eines Aufwärts-Wandlers

4.26) nur schlechte Ergebnisse aufweisen. Analog zum Buck-Konverter sind aufgrund der geringen Spannungen die Verluste an den verschiedenen Bauteilen nicht vernachlässigbar und die Ausgangsspannung nicht unabhängig von der Last. Daher ist auch bei diesem Konverter eine Regelung der Ausgangsspannung notwendig. Da der gesamte Ausgangsstrom über die Diode fließt, und eine typische Diodenspannung im Bereich von  $U_D = 0.7V$  liegt, stellt dies einen enormen Leistungsverbrauch im Bezug auf eine Eingangsspannung von  $U_{in} = 1.2V$  dar. Die gewählte Eingangsspannung ist gleich der Zellenspannung gängiger Akkumulatoren.

Es ist unter dem Aspekt guten Wirkungsgrads zwingend notwendig, die Diode durch einen PMOS-Transistor mit geringem  $U_{DS}$  zu ersetzen. Bei der Ansteuerung des Transistors genügt es jedoch nicht, diesen mit demselben Steuersignal wie den NMOS-Transistor zu beschalten. Es muss sichergestellt werden, dass bei lückendem Spulenstrom kein Strom vom Ausgang zum Eingang zurückfließt. Dazu muss eine entsprechende Steuerlogik entwickelt werden.

### 4.6.2 PMOS-Ansteuerlogik

Um eine solche Logik entwerfen zu können, müssen zunächst die Anforderungen an diese festgestellt werden. Wie in Kapitel 4.6.1 beschrieben, können zwei Arbeitsphasen unterschieden werden. In der ersten Phase ist der Transistor  $T_1$  durchgeschaltet und der Spulenstrom  $i_L$  steigt linear an. Die Spule wird geladen. In der zweiten Phase ist  $T_1$  gesperrt. Die an der Spule anliegende Spannung kehrt sich gemäß Induktionsgesetz um (vgl. Kapitel 4.6.1). Die Energie, die als magnetisches Feld in der Spule gespeichert ist, hält den Strom  $i_L$  aufrecht, der nun über die Diode  $D_1$  zum Ausgang fließt. Der Strom nimmt linear ab. Fällt die Spannung am Knoten  $K_1$  (Abbildung 4.26) unter den Wert  $U_{out} + U_D$ , sperrt die Diode  $D_1$ , und es wird verhindert, dass Ladung vom Ausgang zum Eingang zurückfließt. Die Steuerlogik muss also folgende Bedingungen erfüllen:

1.  $T_2$  darf nur durchschalten, wenn  $T_1$  gesperrt ist
2.  $T_2$  darf nur durchschalten, wenn  $U_{DS-T_2}$  negativ ist

Eine Steuerlogik, die diesen Anforderungen Rechnung trägt, ist in Bild 4.28 dargestellt. Über einen Komparator werden die Spannungen an den Knoten  $K_1$  und  $K_2$  miteinander verglichen. Damit diese Spannungen in einem für den Komparator günstigen Spannungsbereich liegen, werden diese über einen Spannungsteiler ( $T_5, T_6, T_7$  und  $T_8$ ) mit den Eingängen des Komparators verbunden. Eine gute Paarigkeit von  $T_5$  und  $T_8$  bzw.  $T_6$  und  $T_7$  ist hierbei anzustreben. Ist die Spannung am Knoten  $K_1$  größer als die am Knoten  $K_2$ , schaltet der Komparator den Ausgang nach *high* (Stromfluss vom Ausgang zum Eingang wird verhindert). Um die zweite Bedingung zu erfüllen, wird dieses Signal zusammen mit dem invertierten Steuersignal des NMOS-Transistors auf ein AND-Gatter geführt. Der Ausgang des AND-Gatters ist also *low*, wenn sowohl Forderung 1, als auch Forderung 2 erfüllt sind. Da der Transistor  $T_2$  ein PMOS-Transistor ist, welcher beim Anlegen von *low* leitet, wird das Signal des AND-Gatters an einen Inverter gegeben, welcher gleichzeitig als Treiber für  $T_2$  fungiert. Damit der Transistor vollständig gesperrt werden kann, muss der Inverter mit  $U_{out}$  versorgt werden. Würde man diesen mit der Eingangsspannung versorgen, so würde sich bei einer angenommenen Ausgangsspannung von 1.8 V und einer Eingangsspannung von 1.2 V ein  $U_{GS}$  von 0.6 V ergeben und der Transistor  $T_2$  würde durchschalten. Wegen der Treiberfähigkeit des Inverters ist dieser hier explizit notwendig, weswegen die digitale Funktion AND und Inverter nicht zu einem NAND-Gatter zusammengefasst wird. Auf eine Reduzierung des maximalen Eingangspegels von  $T_2$  wurde hier verzichtet, da die Ansteuerung des Treiberinverters durch die variable Versorgungsspannung kritisch ist.

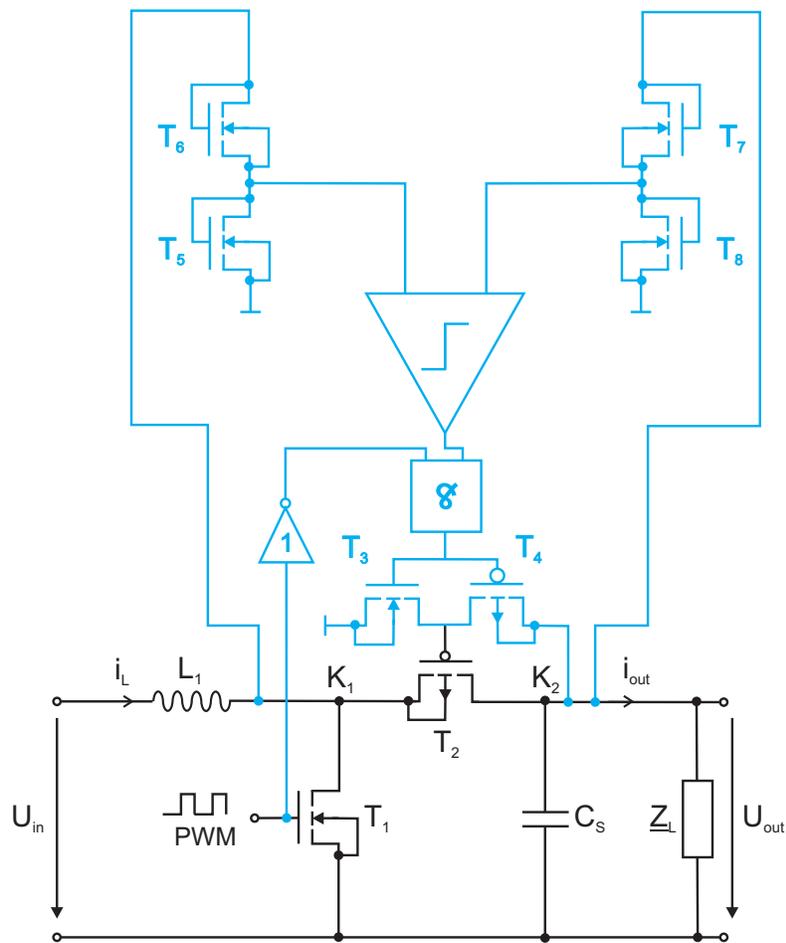


Abbildung 4.28: Steuerlogik der „aktiven Diode“ (blau)

### 4.6.3 Boost-Konverter mit alternierenden Takten

Beim Boost-Konverter mit alternierenden Takten wird, je nachdem ob die Ausgangsspannung unter oder über dem Sollwert liegt, eines von zwei Rechtecksignalen zur Ansteuerung der Transistoren  $T_1$  und  $T_2$  verwendet. Die Auswahl des Steuersignals erfolgt durch einen Komparator. Das Ausgangssignal wird in seiner ursprünglichen Form auf den Eingang B und in invertierter Form auf den Eingang C des AOI-Gatters geführt. Abbildung 4.29 zeigt den schematischen Aufbau dieses Boost-Konverters. Die Referenzspannung muss  $2/3$  der gewünschten Ausgangsspannung der Schaltung betragen, da diese über einen Spannungsteiler im Verhältnis 1 zu 2 an den Signaleingang des Komparators geschaltet wird. Liegt die Ausgangsspannung über der Sollspannung, schaltet der Komparator seinen Ausgang auf *low*. Über das AOI-Gatter wird demnach das Steuersignal mit einem Tastverhältnis von 1 zu 7 (vgl. Kapitel 4.5.1) ausgewählt. Dieses wird anschließend zweimal invertiert (1. AOI-Gatter, 2. Inverter) und dann als Gatesignal des Transistors  $T_1$  verwendet. Mit  $U_{out} = \frac{T}{t_{aus}} \cdot U_{in}$  ergibt sich bei einer Eingangsspannung von 1.2 V (Zellenspannung), unter Vernachlässigung jeglicher Verluste, eine minimale Ausgangsspannung von  $U_{Out} = \frac{8}{7} \cdot 1.2 V = 1.37 V$ . Analog ergibt sich, bei einem Taktverhältnis von 1 zu 1 (Systemtakt), idealisiert betrachtet, die maximale Ausgangsspannung zu  $U_{out} = \frac{2}{1} \cdot 1.2 V = 2.4 V$ . Die Ansteuerung des PMOS-Transistors  $T_2$  erfolgt gemäß Kapitel 4.6.2. Die verwendete Siebkapazität hat eine Größe von  $1 \mu F$ .

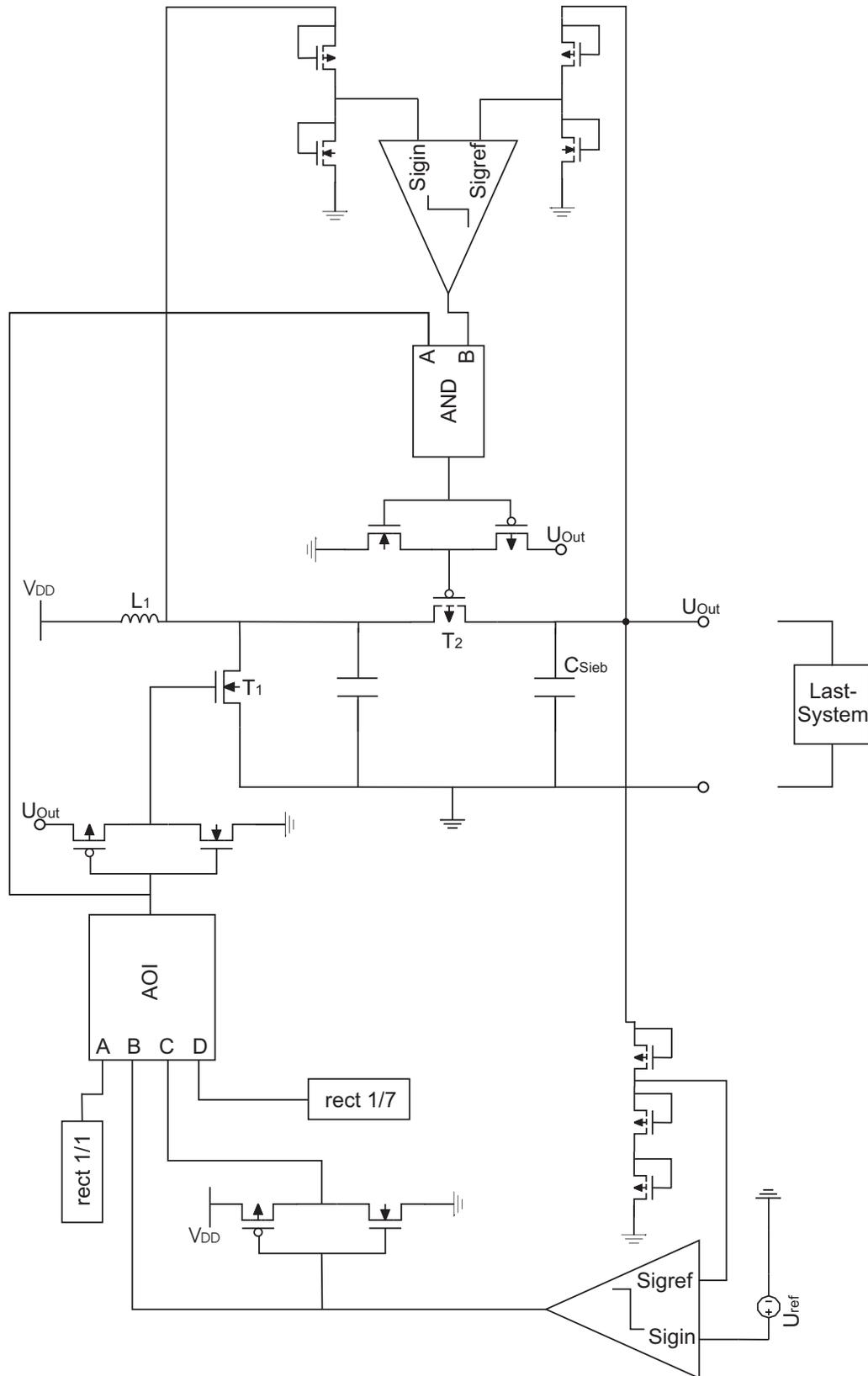


Abbildung 4.29: Boost-Konverter mit alternierenden Takten

#### 4.6.4 Simulationsergebnisse

Bei einer Eingangsspannung von 1.2 V und einer Last von 1 k $\Omega$  erhält man die in Tabelle 4.15 aufgeführten Ergebnisse.

$V_{\text{In}}$ [V]	$V_{\text{Out}}$ [V]	$V_{\text{Out-Max}}$ [V]	$V_{\text{Out-Min}}$ [V]	Wirkungsgrad $\eta$
1.2	1.8	1.8	1.788	0.82
1.2	1.7	1.703	1.696	0.83
1.2	1.6	1.606	1.597	0.83
1.2	1.5	1.510	1.497	0.84
1.2	1.4	1.416	1.397	0.83
1.2	1.3	1.323	1.299	0.84

**Tabelle 4.15:** Variation der Ausgangsspannung bei konstanter Last (1 k $\Omega$ )

Im Gegensatz zur idealen Rechnung lässt sich bei einer Last von 1 k $\Omega$  wegen der Verlustspannungen über den Schalttransistoren auch eine Ausgangsspannung kleiner als 1.37 V realisieren. Der Wirkungsgrad bleibt über den gesamten Ausgangsspannungsbereich nahezu konstant, was einen großen Vorteil darstellt. Verringert man die Last auf 5 k $\Omega$ , ergeben sich die Daten in Tabelle 4.16.

$V_{\text{In}}$ [V]	$V_{\text{Out}}$ [V]	$V_{\text{Out-Max}}$ [V]	$V_{\text{Out-Min}}$ [V]	Wirkungsgrad $\eta$
1.2	1.8	1.801	1.799	0.76
1.2	1.7	1.701	1.699	0.75
1.2	1.6	1.601	1.599	0.77
1.2	1.5	1.502	1.5	0.78
1.2	1.4	1.402	1.4	0.79
1.2	1.3	1.365	1.365	0.80

**Tabelle 4.16:** Variation der Ausgangsspannung bei konstanter Last (5 k $\Omega$ )

Bei dieser Belastung erreicht man 1.3 V Ausgangsspannung nicht mehr. Die minimale Ausgangsspannung liegt bei 1.365 V. Dies ist eine Folge der lastabhängigen Schaltverluste.

Bei wechselnden Bedingungen ergeben sich folgende Ergebnisse (Tabelle 4.17). Hierbei ist  $V(t) = 1.2 \text{ V} + 0.1 \text{ V} \cdot \sin(2 \cdot \Pi \cdot 1 \text{ MHz} \cdot t)$ .

Schwankungen der Eingangsspannung bis 100 mV haben keine Auswirkung auf den Wirkungsgrad. Die Beschreibung der verschiedenen Lasten bezieht sich auf Bild 4.4.

$V_{In}$	$V_{Out-Max}$ [V]	$V_{Out-Min}$ [V]	Last	Wirkungsgrad $\eta$
1.2	1.509	1.498	$R_1 = R_2 = 2 k\Omega$	0.85
1.2	1.511	1.494	$R_3 = 100 k\Omega, C_1 = 1 nF$	0.79
1.2	1.511	1.494	$R_4 = 100 k\Omega, C_2 = 1 nF,$ $L_1 = 1 nH$	0.79
$V(t)$	1.509	1.494	$R_4 = 100 k\Omega, C_2 = 1 nF,$ $L_1 = 1 nH$	0.80

Tabelle 4.17: Messungen bei wechselnden Bedingungen

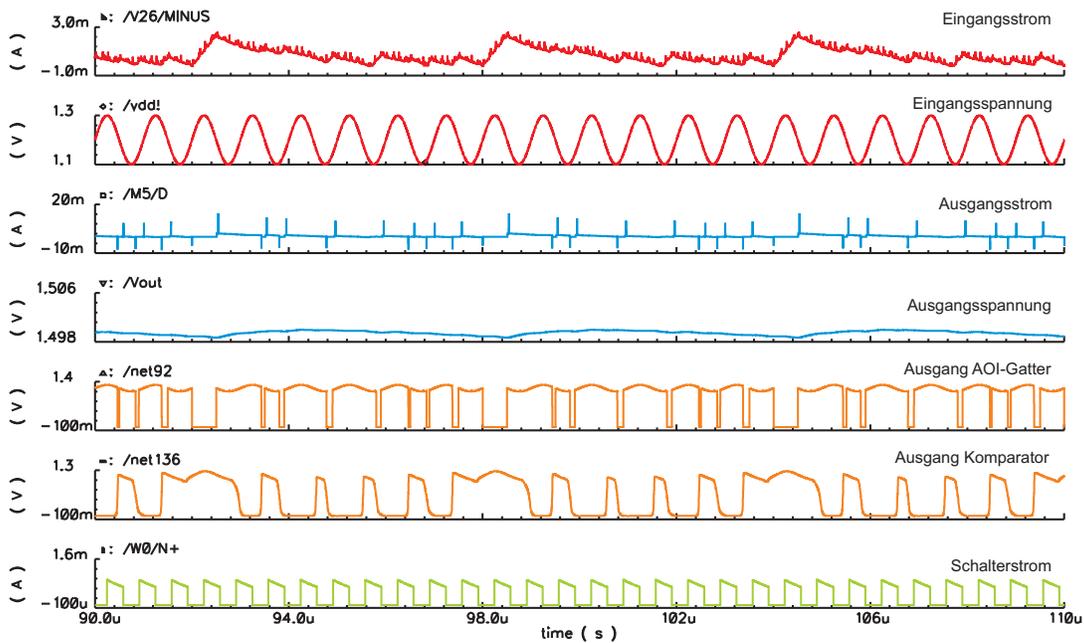


Abbildung 4.30: Spannungs- und Stromkurven für eine Beschaltung entsprechend Zeile 4 aus Tabelle 4.17

### 4.6.5 Boost-Konverter mit variablem Taktverhältnis

Ein Duty-Cycle-Generator, wie in Kapitel 4.5, wird im folgenden Schritt dazu benutzt, um das Taktverhältnis eines Boost-Konverters zu regeln. Hierzu wird eine Regelschleife aufgebaut, bei der die Ausgangsspannung des Boostkonverters mit einer Referenzspannung mit einem Differenzverstärker verglichen wird. Der Ausgang des Differenzverstärkers wirkt nun als Steuerspannung für den PWM-Generator. Somit wird durch den Regelkreis das Taktverhältnis permanent nachgeregelt, so dass sich eine stabilisierte Ausgangsspannung einstellt. Abbildung 4.31 zeigt den schematischen Aufbau des geregelten Boostkonverters.

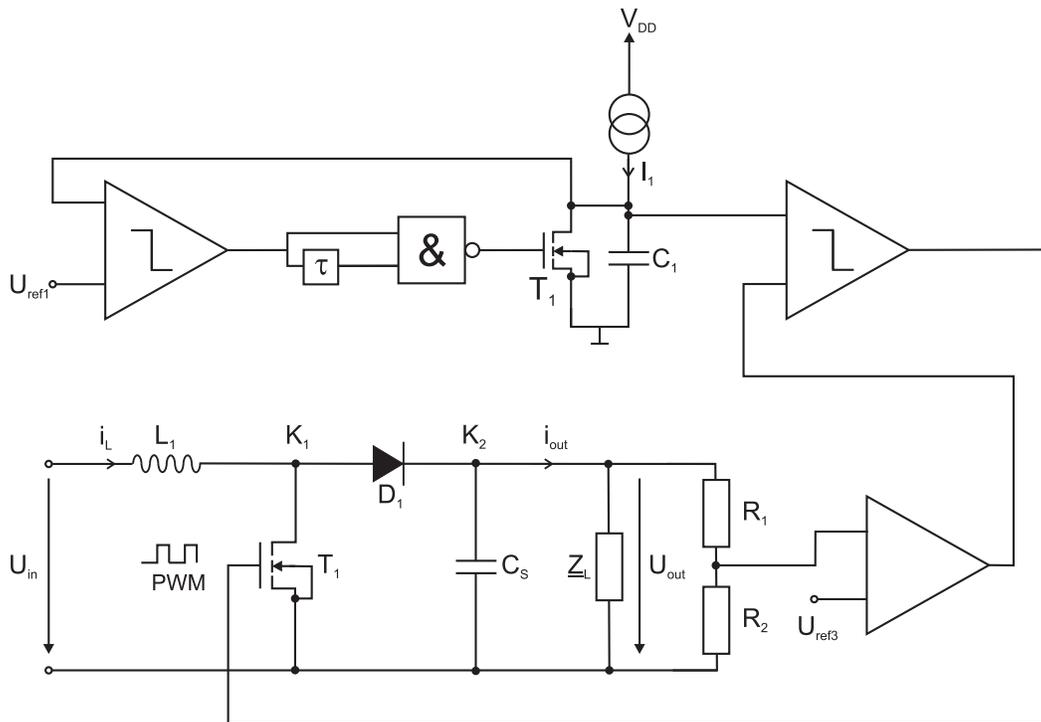
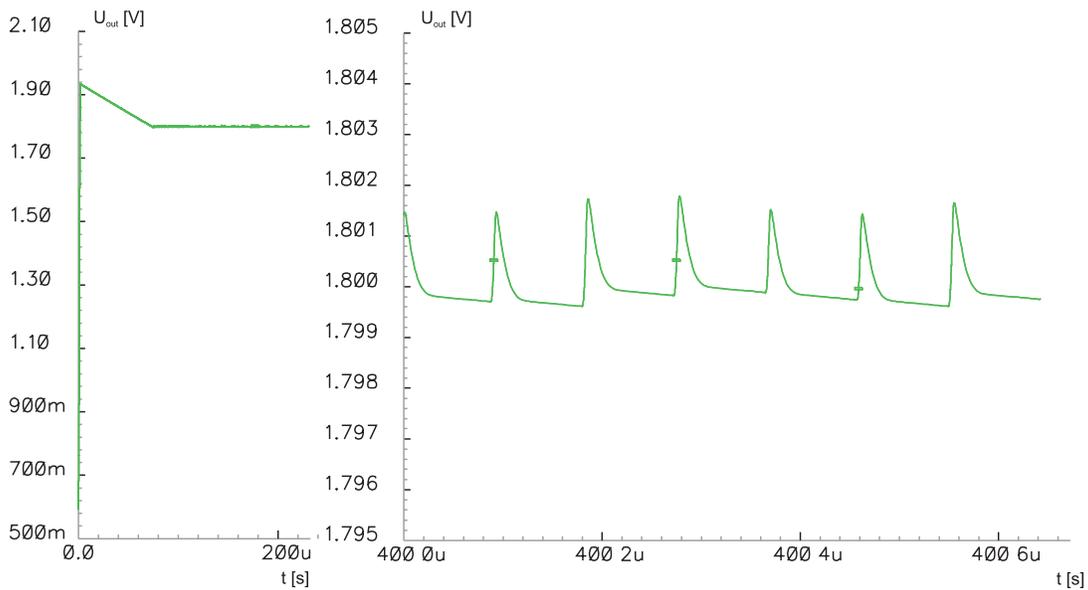


Abbildung 4.31: Schema des geregelten Boost-Konverters mit variablem PWM-Steuersignal

$R_1$  und  $R_2$  bilden hierbei einen Spannungsteiler, der in der Realität aus NMOS-Dioden bestehen sollte, um den Differenzverstärker besser ansteuern zu können. Folglich muss  $U_{ref3}$  den halben Wert der angestrebten Ausgangsspannung annehmen. In einer ersten Simulation ist zur Evaluierung des Leistungsverlusts in der Diode diese noch nicht durch eine aktive Diode (vergl. Kapitel 4.6.2) ersetzt. Abbildung 4.32 zeigt die Transientensimulation der Schaltung nach Abbildung 4.31 mit  $L = 1\text{mH}$ ,  $R = 1\text{k}$ ,  $U_{in} = 1.2\text{V}$ ,  $U_{ref3} = 0.9\text{V}$  und  $C_s = 1\mu\text{F}$ . Nach einem Überschwinger beim start-up der Schaltung regelt die Schaltung auf eine Ausgangsspannung von  $U_{out} = 1.8\text{V}$  mit bemerkenswert kleinem Ripple, der in Abbildung 4.31 rechts zu  $U_{ripple-pp} = 2\text{mV}$  abzulesen ist.

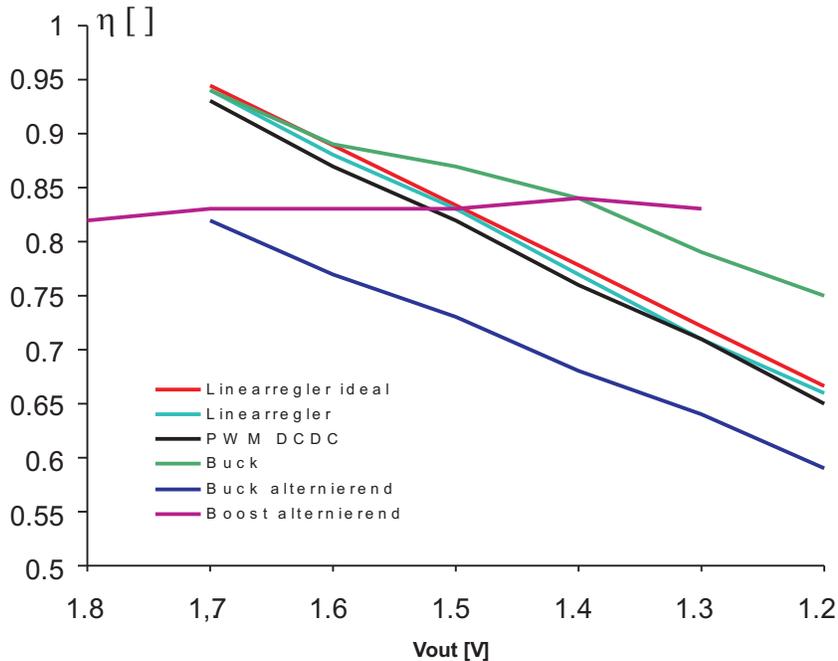


**Abbildung 4.32:** Transientensimulation der Schaltung nach Abbildung 4.31

Bei dem angestrebten Ausgangsspannungsbereich von  $U_{out} \leq 1.8\text{V}$  ist die Diodenspannung der Diode  $D_1$  (Abb. 4.31) in einer Größenordnung, die gute Wirkungsgrade verhindert. Die Evaluierung des Wirkungsgrads für die angegebenen Betriebsparameter bestätigt das. Ein Ersetzen der Diode durch einen PMOS-Transistor mit entsprechender Ansteuerlogik könnte hierbei für Abhilfe sorgen. Da für die Komplexität des Systems und den Leistungsverbrauch der Ansteuerung dasselbe gilt, wie für den Buck-Konverter, wird aber auf eine tiefere Betrachtung dieser Standardschaltung verzichtet.

## 4.7 Abschließender Vergleich der DC-DC- Spannungswandlerkonzepte

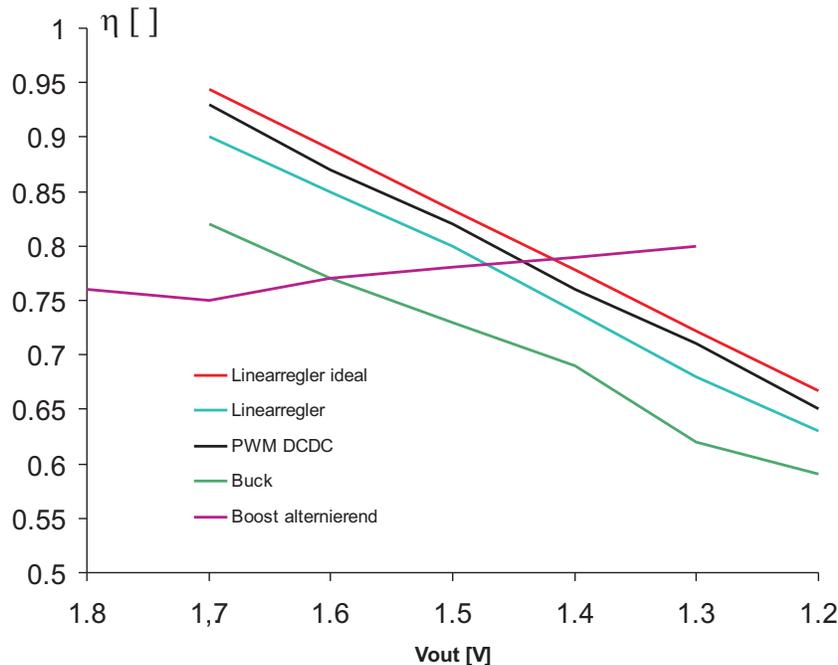
Die Besonderheit bei den hier vorgestellten DC-DC-Konvertern ist die verhältnismäßig kleine Ausgangsleistung, wie sie bei Low-Power-Konzepten typisch ist. Das drängt den Leistungsverbrauch der Komponenten der Regelung in den Vordergrund. Abbildung 4.33 verdeutlicht für eine ohmsche Last von  $1k$  den Wirkungsgrad der Konzepte. Dabei gilt für alle Abwärtswandler  $U_{in} = 1.8V$  und für die Aufwärtswandler  $U_{in} = 1.2V$ .



**Abbildung 4.33:** Wirkungsgradvergleich der verschiedenen Wandlerkonzepte bei einer Last von  $R_L = 1k\Omega$

Es ist offensichtlich, dass der Buck-Konverter mit alternierenden Takten untauglich ist. Als vorteilhaft erweist sich der Buck-Konverter aus Kapitel 4.4, der einen signifikanten Vorsprung beim Wirkungsgrad erreicht bei geringer Komplexität des Systems. Der vorgestellte Linearregler erweist sich bei der angegebenen Belastung nahe am für Linearregler theoretischen Optimum, was sich natürlich bei kleinerer Belastung (siehe Abbildung 4.34) relativiert. Für geringe Dropspannungen ist der Linearregler sogar zu bevorzugen, da er dann hohen Wirkungsgrad, unerreichte Regelgüte mit unkritischem EMV-Verhalten kombiniert. Der PWM-PFM-DC-DC Konverter (Kapitel 4.2) zeigt hinsichtlich des Wirkungsgrades keine Vorteile. Einzig der geringfügig kleinere Flächenbedarf bei der Integration sprechen für die Verwendung dieses Konzepts. Der konstant ho-

he Wirkungsgrad des Boost-Konverters mit alternierenden Takten macht dieses Konzept interessant, vor allem für batteriebetriebene Systeme, ist doch dieses System in der Lage, die in der Batterie oder dem Akkumulator gespeicherte Energie besonders gut auszunutzen, weil es die im Laufe der Entladung absinkende Zellenspannung auf eine konstante Ausgangsspannung wandeln kann.



**Abbildung 4.34:** Wirkungsgradvergleich der verschiedenen Wandlerkonzepte bei einer Last von  $R_L = 5k\Omega$

Bei noch kleinerer Last wird der Einfluss des Leistungsumsatzes in der Regelschaltung selbst sehr deutlich. Der Buck-Konverter fällt nun stark im Wirkungsgrad ab, der Ruhestrom des Komparators und die Umladeleistung des Schaltransistors wirken sich nun aus. Alle Regler verzeichnen, wie zu erwarten, Einbußen beim Wirkungsgrad. Die DC-DC-Regelung bei solch kleinen Lastströmen ist nicht mehr sinnvoll. Die Regelgüte, also Ripplespannung, Reaktion auf Lastsprünge und PSRR, sind für alle getesteten Konverter völlig ausreichend. Bei genügen hoher Siebkapazität ist die Versorgungsspannungsqualität sogar für analoge Schaltungen ausreichend. Eine Entscheidung für oder gegen ein Wandlerkonzept hängt von der jeweiligen Applikation ab. Eine Pauschalaussage für ein System kann es nicht geben.

Die vorgestellten Spannungswandler sind für kleine Leistungen konzipiert. Für die eingesetzten Baugruppen, wie Komparator und Differenzverstärker, können bei vielen Prozessen vom Hersteller vorentwickelte IPs benutzt werden. Im be-

nutzten Prozess stehen keine analogen IPs zur Verfügung, weshalb diese selbst entwickelt wurden. Spannungsregler für Leistungsapplikationen, wie Prozessoren, sind i.d.R. extern aufgebaut. Eine parameterspezifische Versorgungsspannungseinstellung muss in diesem Fall den Referenzwert über Pins dem Regler übergeben.

# Systemintegration der vorgestellten Konzepte

Die vornehmlichste Aufgabe des Designers ist die Einhaltung der Spezifikationen. Für den Designer stellt sich zuerst die Frage nach der Technologie. Diese Entscheidung ist aber zusätzlich von Kostenfragen bestimmt, dennoch muss die Technologie für die Einhaltung der Spezifikationen geeignet sein. Steht die Technologie fest, können zuerst systematische Methoden zur Verringerung der Verlustleistung eingesetzt werden (vergl. Kapitel 3.1). Das beinhaltet möglichst effizientes Coding der Hardwarebeschreibung (VHDL, Verilog), Benutzen von effizienten Algorithmen und optimales Mappen des Hardwarecodes auf die benutzte Technologie, um so wenig Hardware wie möglich mit so wenig kapazitiven Schaltverlusten wie möglich zu erreichen. Schaltungstechnische Methoden, wie z.B. die Reduktion des Signalhubs (LVDS) auf Busleitungen zwischen hochaktiven Funktionsteilen reduzieren weiter die dynamische Verlustleistung. Zur Herabsetzung der statischen Verlustleistung, gerade bei Batterie-betriebenen Schaltungen, bietet sich die Methode der Bulkpotentialerhöhung an, die mit einer Reduzierung der Versorgungsspannung auf ein Minimum (zur Aufrechterhaltung des Status) nochmals kombiniert werden kann (vergl. Kapitel 2.8). Ein eminent wichtiger Punkt zur Reduktion der Verlustleistung ist die Festlegung der maximalen Versorgungsspannung. Es ist sehr wahrscheinlich, dass die Standardversorgungsspannung selbst für den absoluten WorstCase aller Parameter ausreicht, damit die Performance der Applikation die Spezifikation übertrifft. Die Reduktion der maximalen Versorgungsspannung auf die minimal notwendige ist schon ohne weitere Maßnahme i.d.R. ein großer Faktor. Das Konzept der Performance-Driven-Supply ermöglicht es, zusätzlich Leistung einzusparen, in dem es für jeden Chip eine individuelle Versorgungsspannung einstellt. Wichtig ist hierbei, dass bei ungünstig variierenden Chips die DC-DC Wandlung auch deaktiviert werden kann, damit es nicht zu unnötigem Verlust in der Spannungsconversion kommt. Die Schwelle, ab welcher evaluierten Versorgungsspannung sich ein DC-DC-Wandlung lohnt, hängt vom Wirkungsgrad des DC-DC-Wandlers ab, der wiederum lastabhängig ist. Auch diese ist eine Applikations-individuelle Grenze, die vom Designer durch Simulation evaluiert werden muss. Abbildung 4.35 illustriert die beschriebenen Maßnahmen.

S P E Z I F I K A T I O N E N	Technologie	$V_{T0}$ $\beta_0$ , $t_{ox}$ , $I_{min}$ $f_{trans}$ CMOS, SOI ...	Komplexität/Fläche Kosten/Fläche Parameter Mismatch	V E R L U S T L E I S T U N G
	Versorgungsspannung	$VDD_{max}$ $VDD_{min}$ DC-DC intern / extern	Performance Driven Supply Bulkpotentialerhöhung Supply Voltage Hopping	
	System	Algorithmen alternative Schaltungstechnik	HDL-Code Software Mapping	

**Abbildung 4.35:** Übersicht der Verlustleistungs-minimierenden Maßnahmen im Systementwurf

Es kann dabei in die drei Gebiete *Technologie*, *Versorgungsspannung* und *System* unterteilt werden, die miteinander in vielfältiger Weise, wie beschrieben, verknüpft sind. Allen Maßnahmen, die innerhalb der Teilgebiete zu finden sind, ist gemeinsam, dass die Verlustleistung reduziert wird, ohne die Spezifikationen zu verletzen. Dies ist das Ziel eines verlustarmen Entwurfs. Dem Designer sind somit vielfältige Möglichkeiten gegeben. Es entsteht eine umfangreiche Kombinationsvielfalt der beschriebenen Maßnahmen. Letztendlich entscheidet der Designer, unter Berücksichtigung der konkreten Applikation, welche Maßnahmen getroffen werden müssen.

# Zusammenfassung und Ausblick

Die vorliegende Arbeit beschreibt Methoden und Schaltungen zur Reduzierung der Verlustleistung in integrierten CMOS-Schaltungen. Dabei werden die statische und dynamische Verlustleistung betrachtet. Zur Verminderung der statischen Verlustleistung werden Konzepte und Schaltungen vorgestellt, die Einfluß auf die Thresholdspannung einzelner Transistoren nehmen bzw. durch Power-Gating den statischen Strom reduzieren. Dadurch wird eine Reduzierung des Leakagestromes, und damit der statischen Verlustleistung erzielt. Die vorgestellten Maßnahmen sind problemlos innerhalb einer Applikation realisierbar. Insbesondere die Potentialerhöhung der N-Wanne im N-Well CMOS Prozess während Idle-Phasen erweist sich als sehr geeignet, den StandBy-Leistungsumsatz zu verringern. Der zusätzliche Aufwand für diese Maßnahme ist vergleichsweise gering. Konzepte, die dynamisch Einfluß auf die Bulk-Spannungen nehmen, werden als layouttechnisch zu aufwendig erachtet, oder sind aufgrund der dabei angewendeten Technologie (SOI) nur in Ausnahmefällen lohnend. Die Bulkpotentialerhöhung in Idle-Phasen erweist sich als probates Mittel, den statischen Stromfluss im zeitlichen Mittel zu reduzieren. Kombiniert mit der Anwendung von HighSpeed-Transistoren mit reduzierter Versorgungsspannung, so dass die Performance der Applikation an jene angeglichen wird, die mit LowLeakage-Transistoren und Standardversorgungsspannung betrieben wird, vereint diese Vorgehensweise die stark reduzierte dynamische Leistungsaufnahme mit reduzierter statischer Leistungsaufnahme.

Auf die Problematik des Gateleakages wird in der vorliegenden Arbeit nicht tiefergehend eingegangen. Fakt ist jedoch, dass bei modernen deep-submicron Technologien diese Verlustquelle immer dominanter wird. Eine Betrachtung der Möglichkeiten zur Reduzierung des Gateleakages ist Thema vieler Veröffentlichungen auf dem Gebiet des LowPower-Designs und als Forschungsthema so umfangreich, dass es alleine schon Thema einer Dissertation sein kann.

Zur Reduzierung der dynamischen Verlustleistung wird als dominanter Faktor die Versorgungsspannung verringert. Hierbei werden Faktoren, wie die Spezifikationen der Schaltung, Herstellungsvariationen der physikalischen Parameter des Chips und die Temperatur als Parameter für eine dynamische Versorgungs-

spannungseinstellung benutzt. Dieses System evaluiert die Schwingfrequenz einer Ringinverterkette, welche von den Chip-spezifischen Parametern abhängt. Um Mismatch in Folge von geometrischen Abweichungen der Transistoren zu minimieren, werden die Transistoren der Ringinverterkette großzügig dimensioniert. Dadurch wird die Ausgangsfrequenz der Ringinverterkette zuverlässig vorher-sagbar in Abhängigkeit der physikalischen Parameter des Chips und der Tem-peratur. Es werden mehrere Systeme vorgestellt, von welchen eines im Rahmen eines Testchips realisiert wurde, die die Ringinverterschwingfrequenz in eine ge-eignete Referenzspannung wandeln. Es stellt sich dabei heraus, dass die Version mit individuell vom Designer in ein ROM gespeicherten Referenzspannungen bzw. PWM-Mustern, das größte Einsparpotential hat. Eine entscheidende Ein-schränkung erfährt das System durch das unvermeidliche und nicht erfassbare geometrische Mismatch, welches heterogen über die Chipfläche ist. Dieses geome-trische Mismatch muss in die Bestimmung der individuellen Referenzspannun-gen mit einbezogen werden. Die geometrische Genauigkeit eines Prozesses wird somit zu einem Maß für das mittlere Einsparpotential des vorgestellten Systems.

Im letzten Teil der vorliegenden Arbeit werden verschiedenen DC-DC-Konverter vorgestellt. Hauptaugenmerk liegt hierbei auf der Eignung der Konzepte für re-lativ kleine Ausgangsleistungen, wie sie in ASICs der mobilen Kommunikation, der Medizintechnik und anderen batteriegetriebenen Systemen vorkommen. Als kritisch erweist sich hierbei der Leistungsumsatz in den Komponenten des DC-DC-Wandlers, der bei den kleinen Lastleistungen nun viel stärker ins Gewicht fällt, als bei hohen Lastleistungen. Die Systeme sind deswegen möglichst wenig komplex. Auf eine Betrachtung rein kapazitiver Konverter wurde u.a. auch des-wegen verzichtet.

Die vorliegende Arbeit gibt einen weitgefächerten Einblick in Möglichkeiten zur Reduzierung der Verlustleistung. Anhand eines Testchips wurde versucht, die theoretischen Aussagen zu verifizieren. Aufgrund der Fehlfunktion des Test-chips konnte dies in den Grundzügen zwar gelingen, eine völlige Funktionalität aber nicht nachgewiesen werden. Deswegen wurde eine diskrete Umsetzung des Systems entwickelt, mit der die Funktionalität des vorgestellten Systems ge-zeigt wurde. Ein Redesign würde das Problem lösen. Als vorteilhaft, gerade in Verbindung mit schaltenden DC-DC-Reglern (Buck- oder Boost) wäre die Realisierung des Systems nach Abbildung 3.27 mit getaktetem ROM. Die Rea-lisierung eines solchen Systems mit einer typischen Standardapplikation über mehrere Runs einer Technologie könnte wertvolle Aussagen über das mittlerer Einsparpotential machen. Dies wäre aber mit einem enormen zeitlichen und vor allem finanziellen Aufwand verbunden, der im Rahmen dieser Dissertation nicht aufgebracht werden kann. Eine weitere Möglichkeit, die Genauigkeit des Systems zu erhöhen, könnte die Implementierung eines zusätzlichen Tempera-tursensors sein. Darüberhinaus wären Untersuchungen anderer Technologien im

---

Hinblick auf geometrisches Mismatch und systematisches Mismatch interessant, um einen Überblick über das Einsparpotential des Systems über eine Vielzahl von Prozessen zu erlangen. Das erwähnte durchschnittliche Einsparpotential von  $\approx 20\%$  als zusätzlicher Energiegewinn ist nämlich nur für die betrachtete Technologie gültig. Mit einer Technologie mit verbessertem geometrischen Mismatch könnte dieser Gewinn eventuell sogar noch gesteigert werden.

# Anhang A

## A.1 Dimensionierung VCO

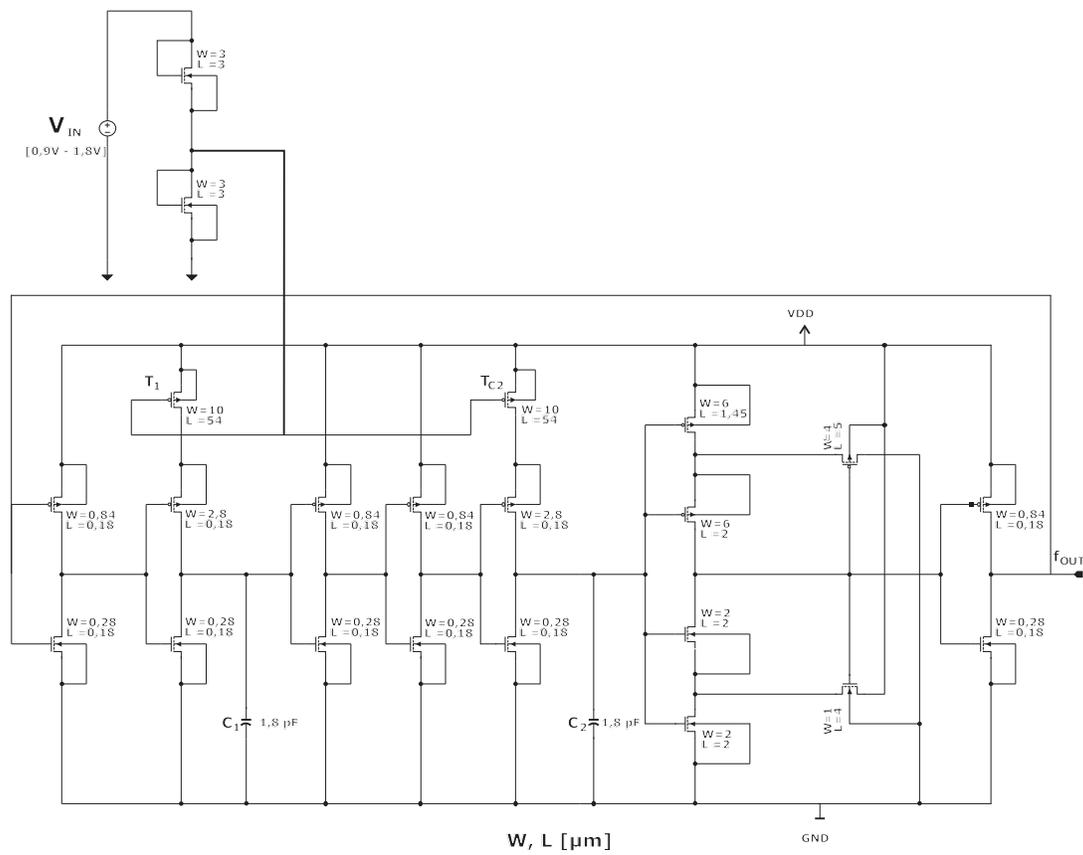


Abbildung A.1: Dimensionierung der VCO-Schaltung



## A.3 Gesamtsystem alternative Auswerteschaltung mit VCO

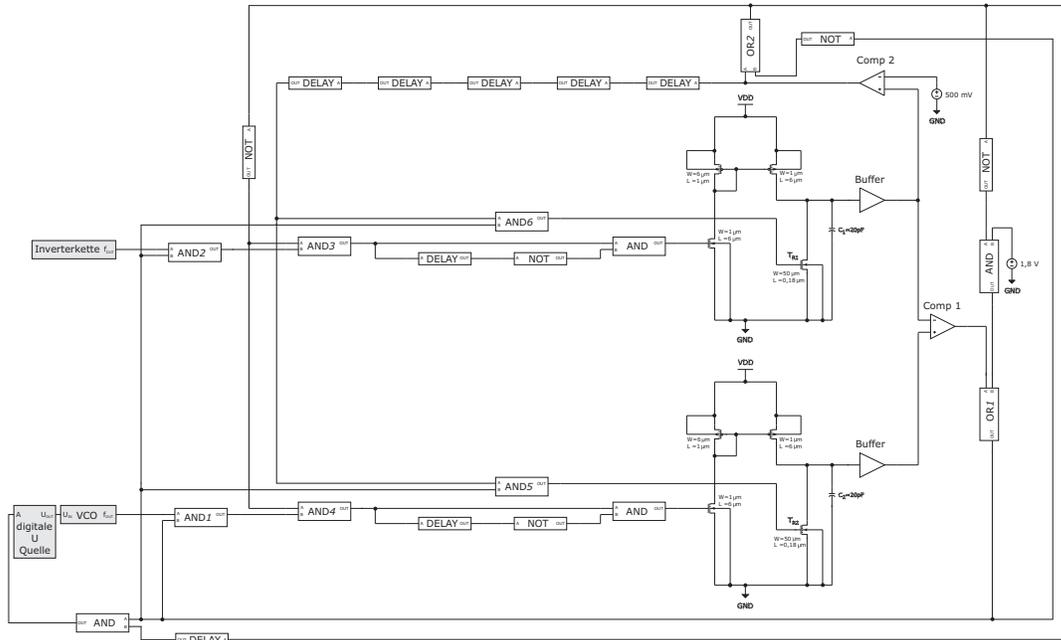


Abbildung A.3: Schematischer Aufbau der analogen Auswerteschaltung

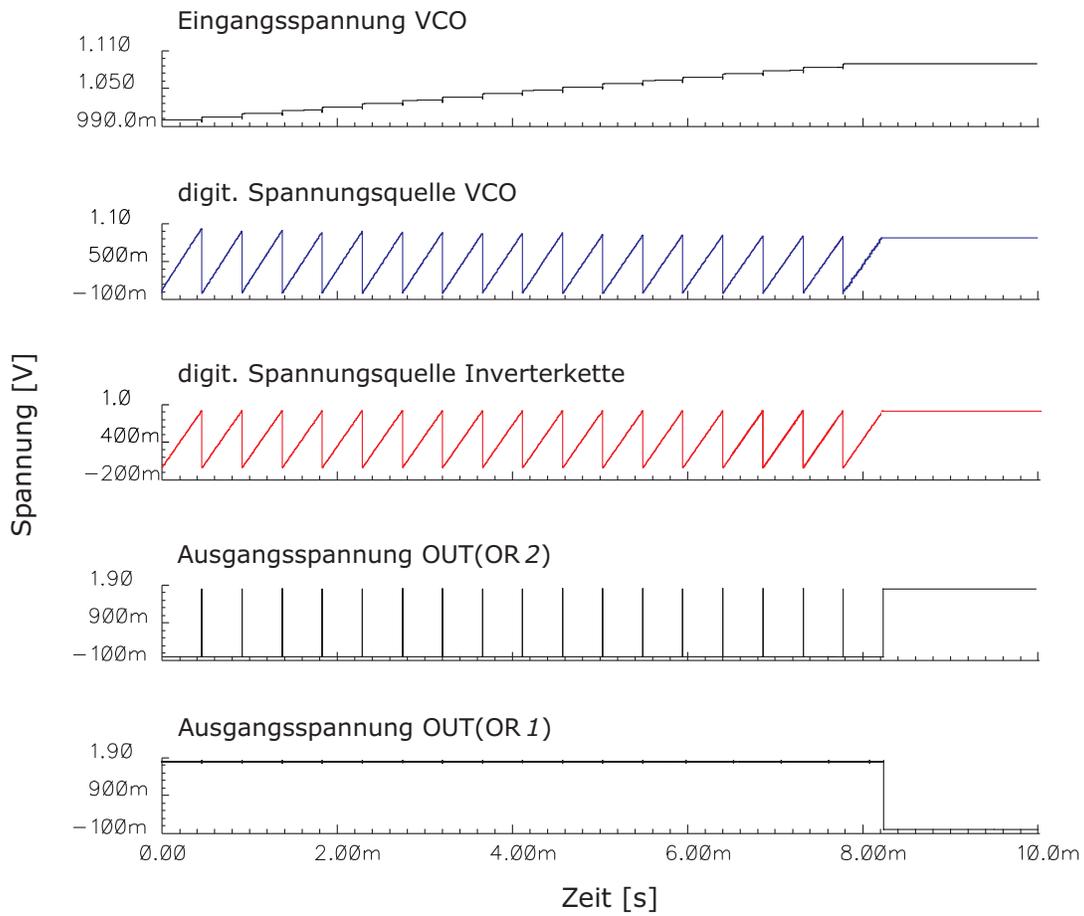


Abbildung A.4: Spannungsverläufe für eine Referenzspannung von 900 mV

## A.4 Differenzverstärker Simulationsergebnisse

Abb. A.5 zeigt das Bodediagramm des Differenzverstärkers aus Abb. 4.3 mit Transistorlast (offener Regelkreis,  $I_{DS}$  von  $T_S = 1mA$ )  
 $U_{inDC} = 1V$ ;  $V_{DM} = 78dB$ ;  $GBW = 34MHz$ ; Phasenreserve =  $72^\circ$

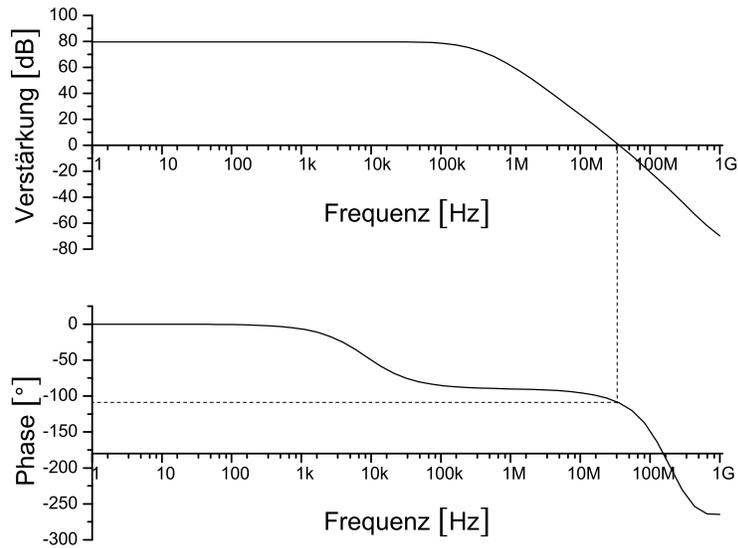


Abbildung A.5: Bodeplot des Differenzverstärkers aus Abbildung 4.3

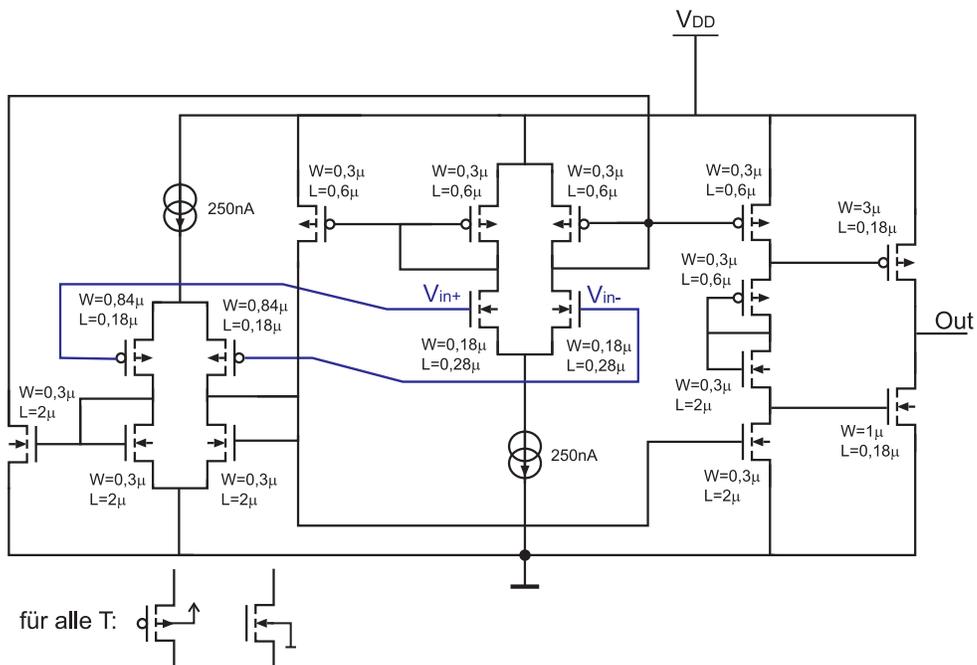


Abbildung A.6: Dimensionierung des Differenzverstärkers aus Abb. 4.3





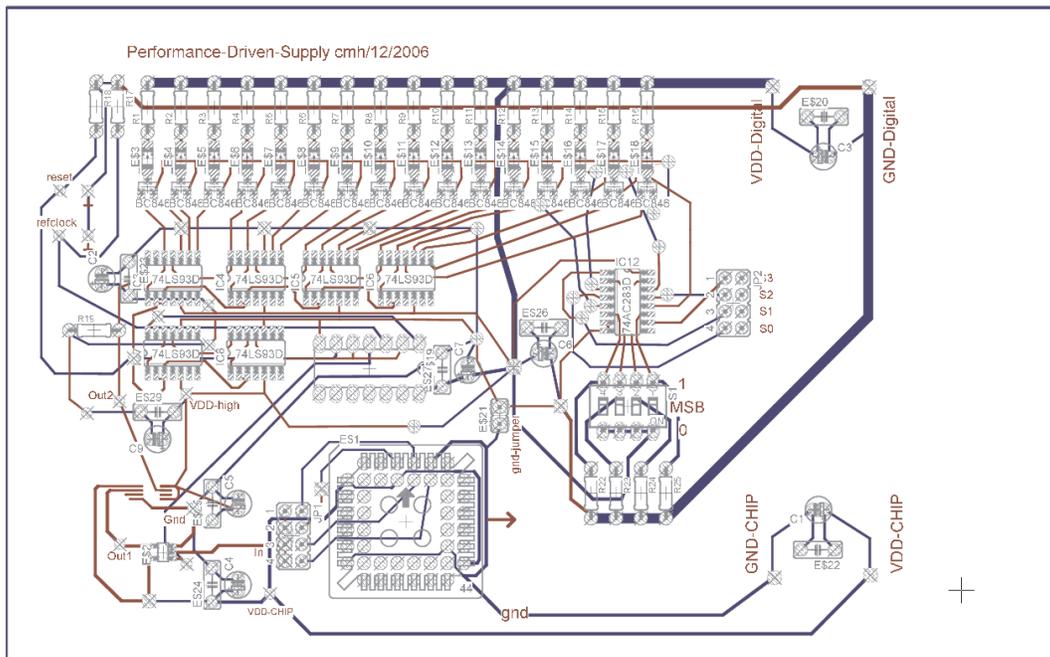


Abbildung A.10: Layout der diskreten Umsetzung

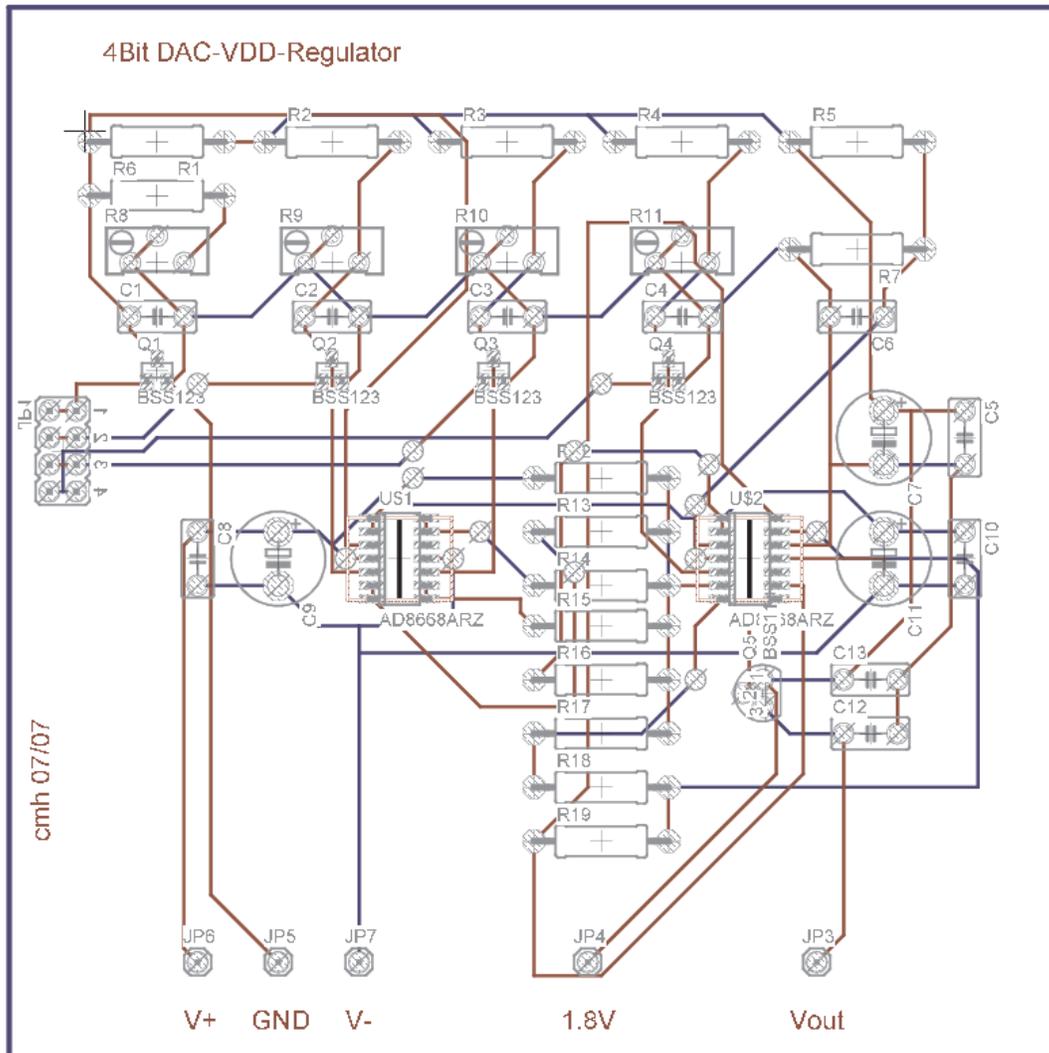


Abbildung A.11: Layout des diskreten 4Bit-DAC mit variabler Kennlinie

## A.7 Schaltpläne und Layouts des Testchips

Im Folgenden werden Schaltpläne und Layouts des entwickelten Testchips gezeigt. Für die ausgewählte Technologie stand kein digitaler Flow zur Verfügung. Deshalb sind die gezeigten Layouts und Standardzellen selbst entwickelt.

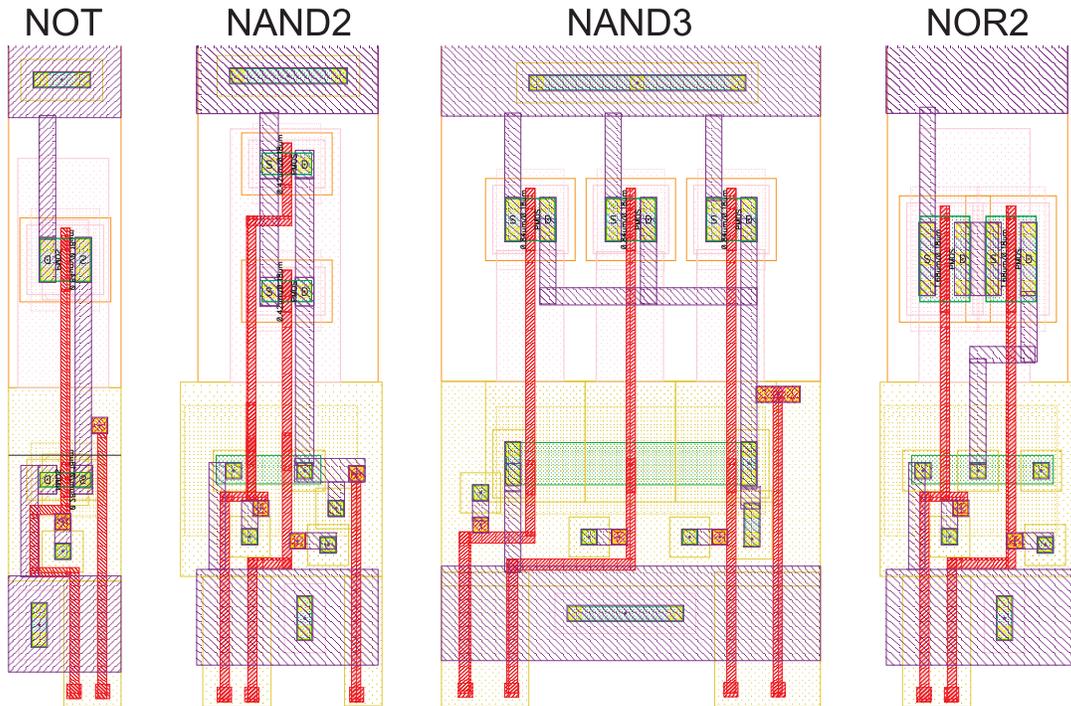


Abbildung A.12: Layout CMOS-Standardzellen

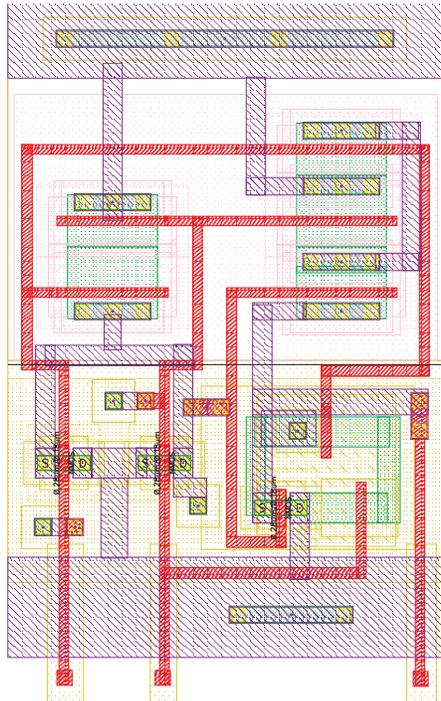
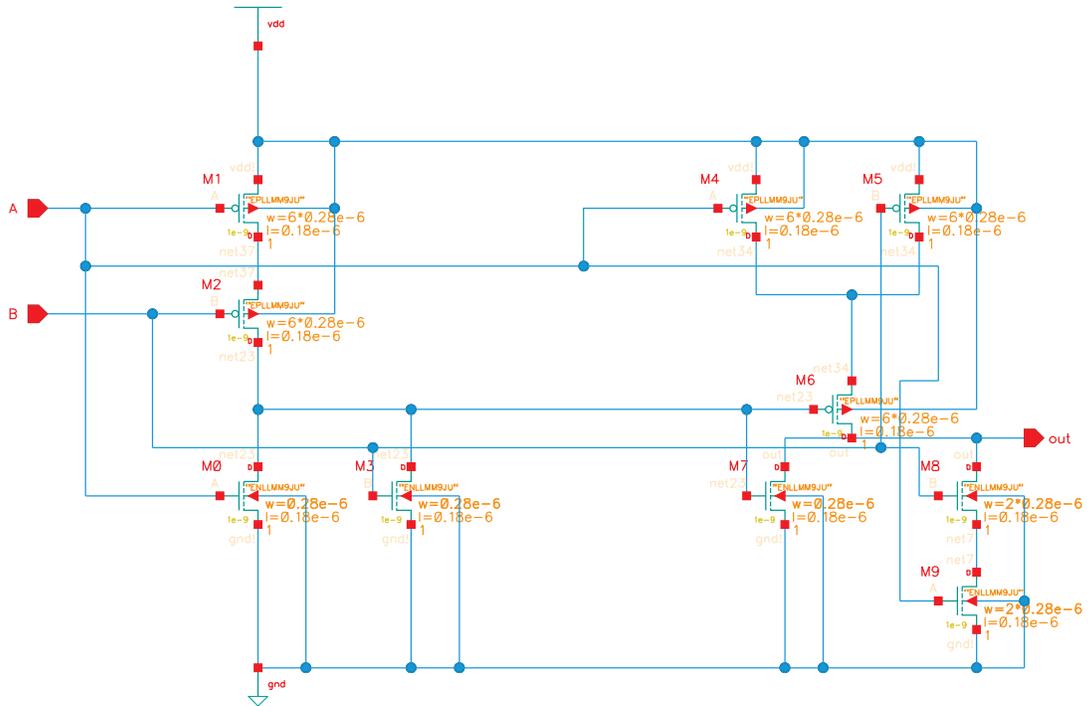


Abbildung A.13: Schaltplan und Layout CMOS-XOR2

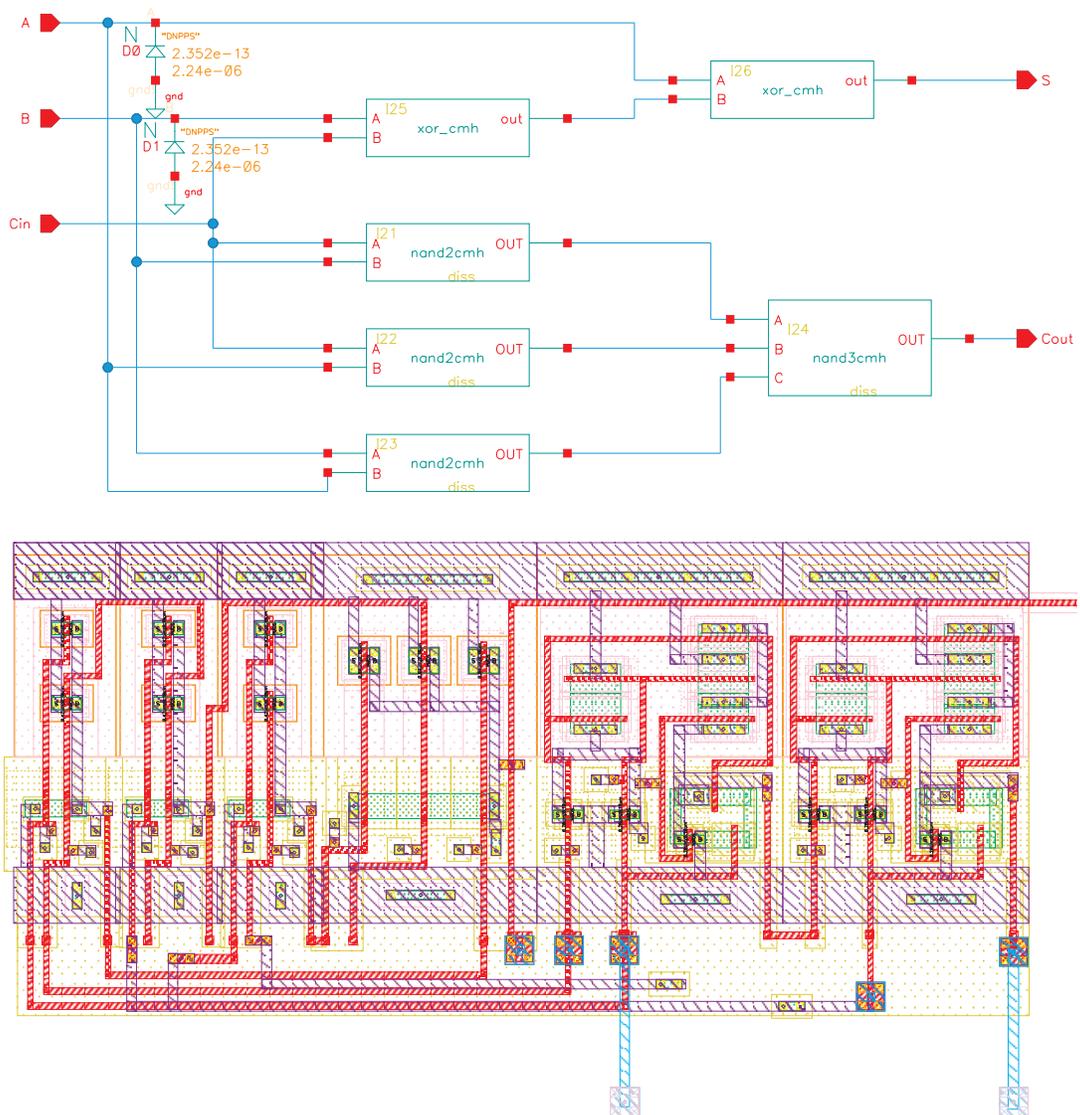


Abbildung A.14: Schaltplan und Layout 2Bit Volladdierer

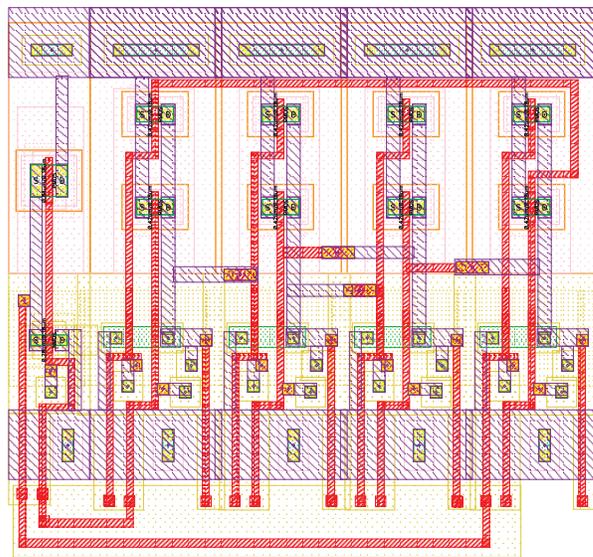
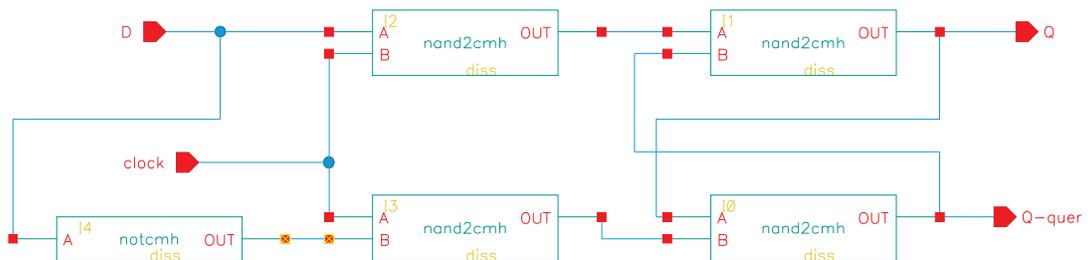


Abbildung A.15: Schaltplan und Layout D-FlipFlop

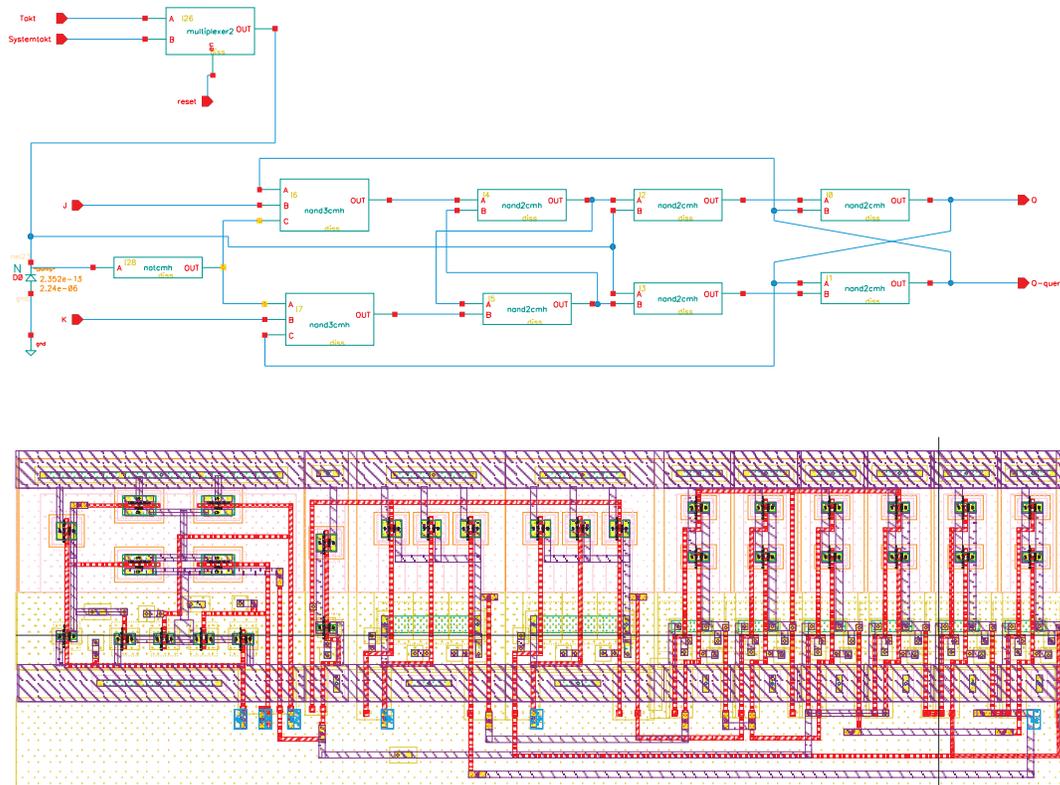


Abbildung A.16: Schaltplan und Layout JK-FlipFlop mit asynchronem Reset

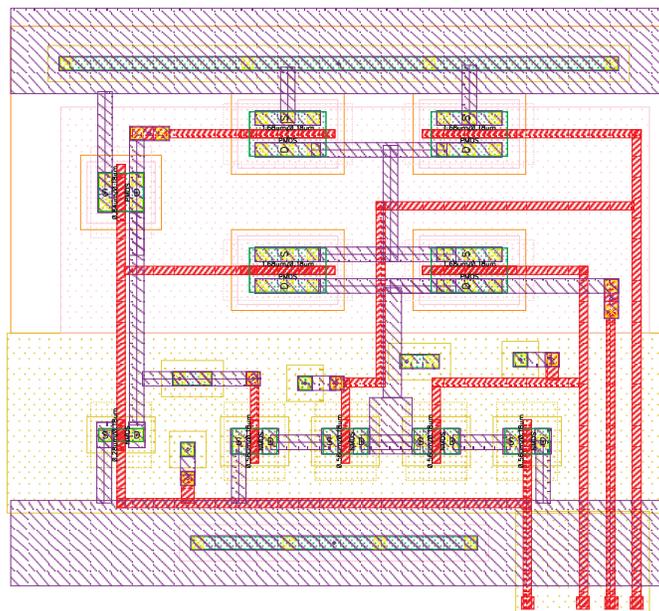
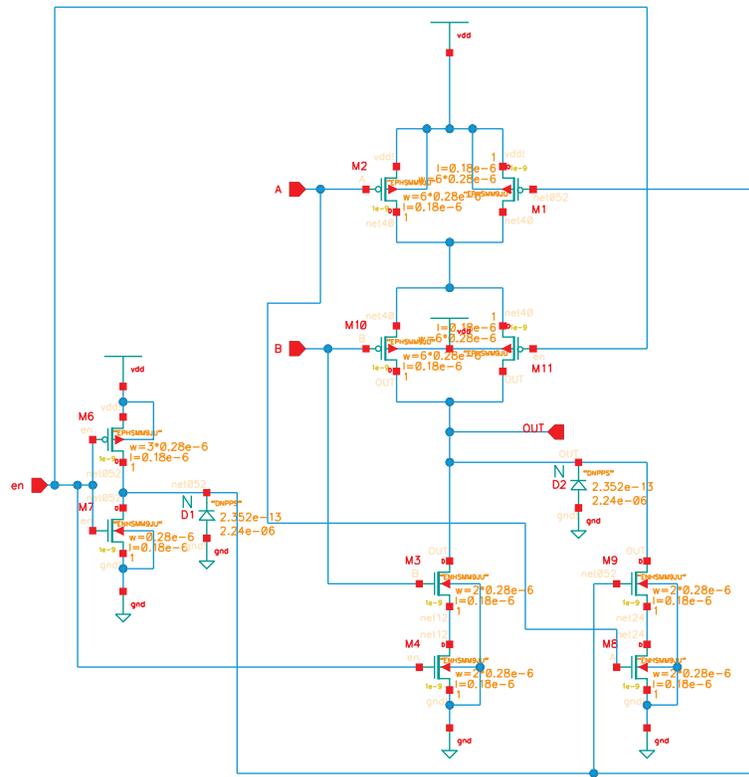


Abbildung A.17: Schaltplan und Layout 2-Input-Multiplexer

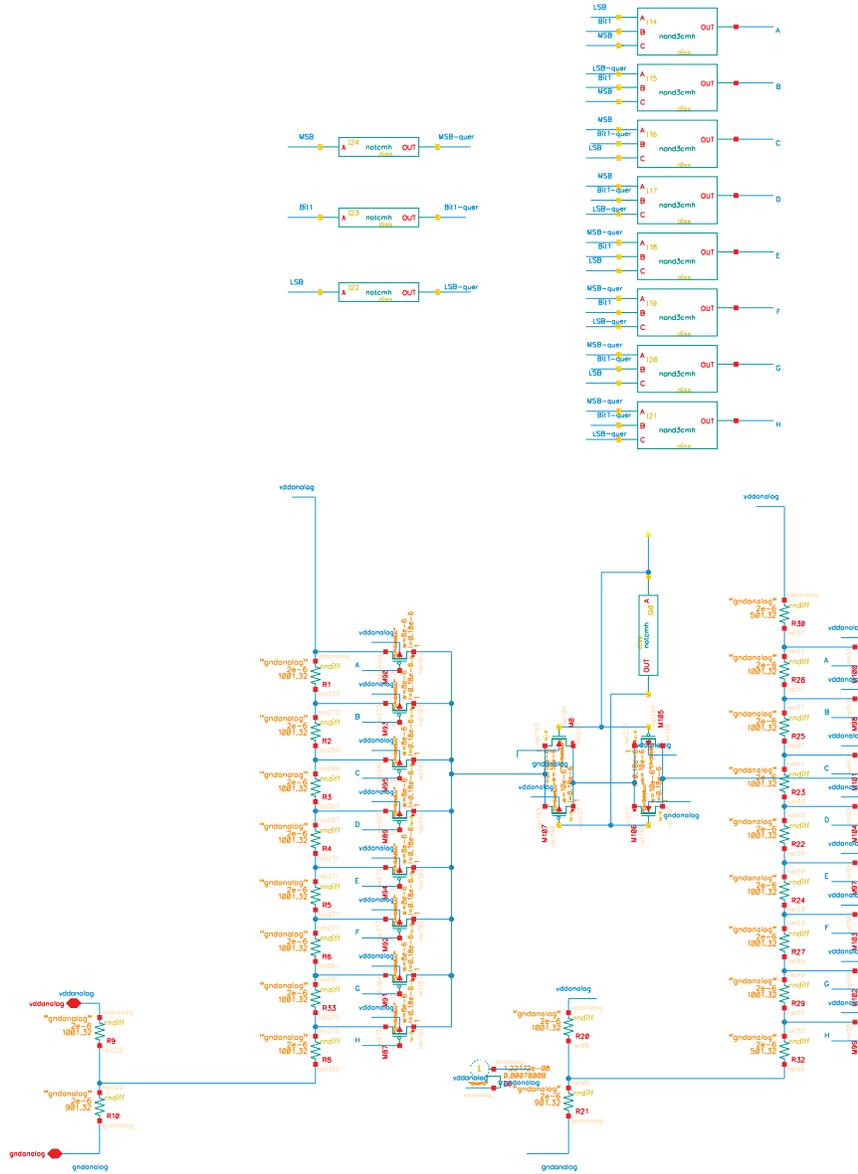


Abbildung A.18: Schaltplan 4-Bit Widerstands-DAC

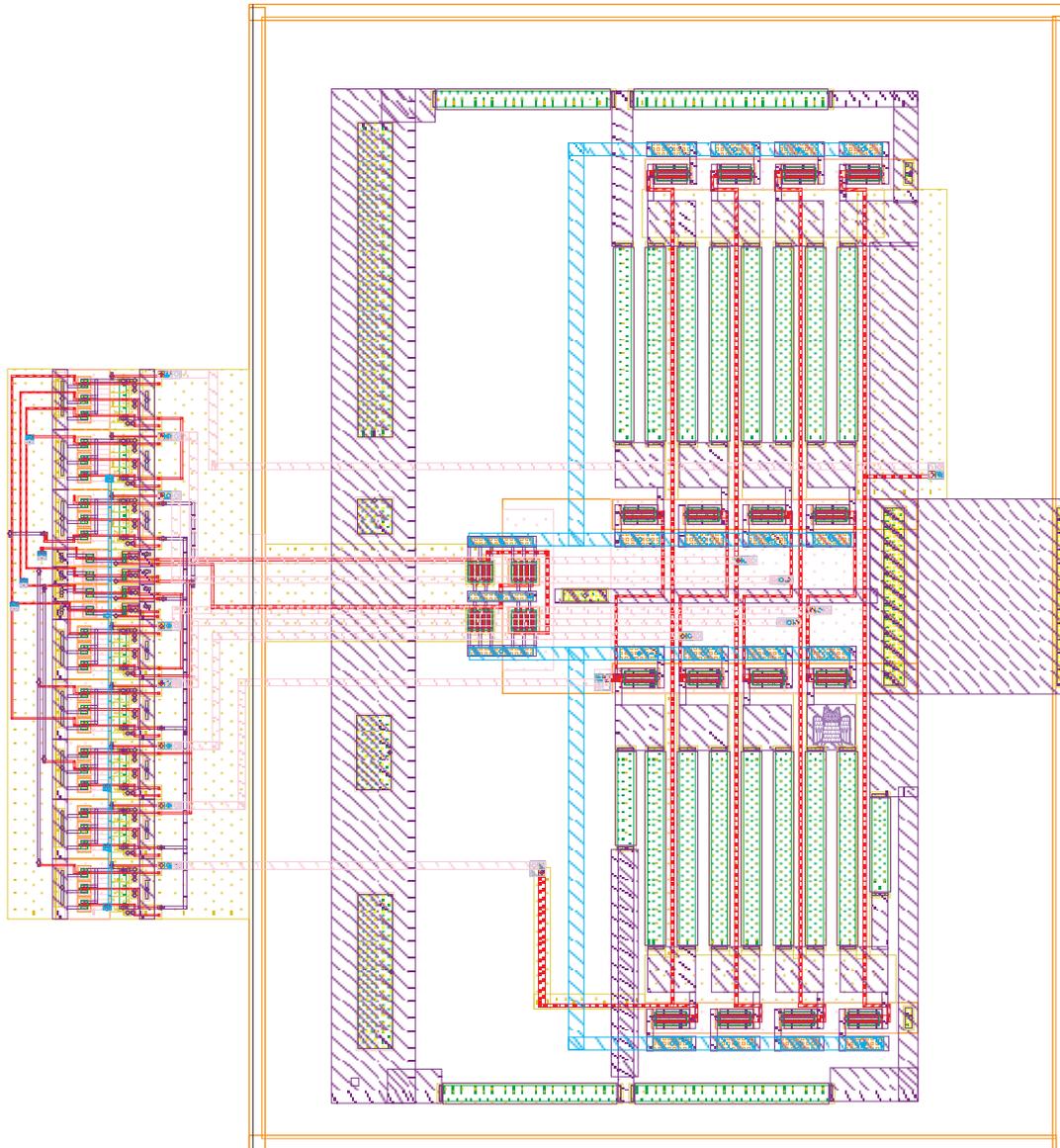
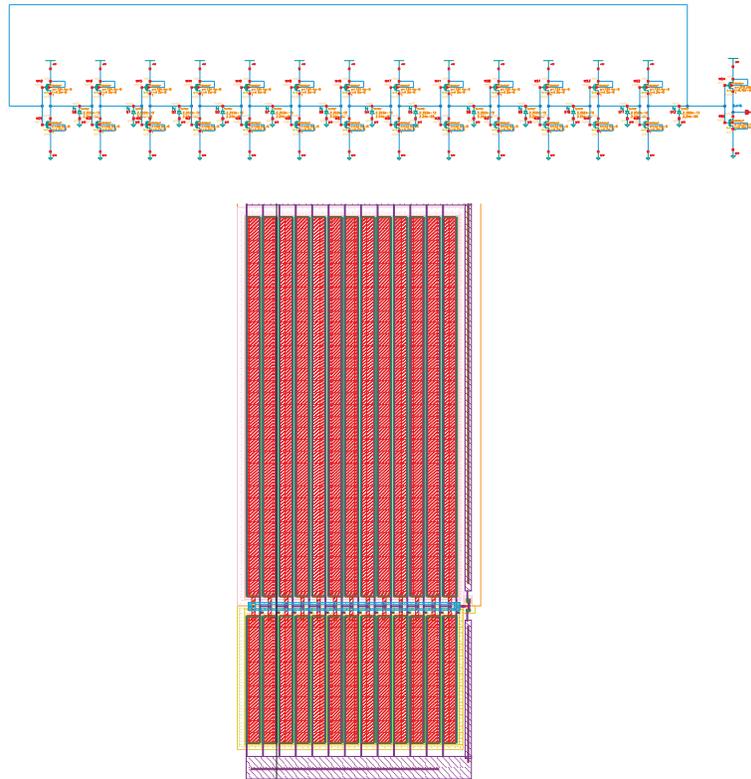
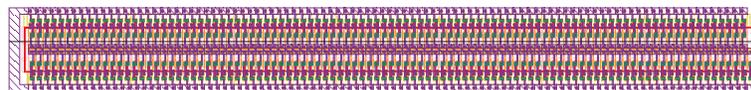


Abbildung A.19: Layout 4-Bit Widerstands-DAC

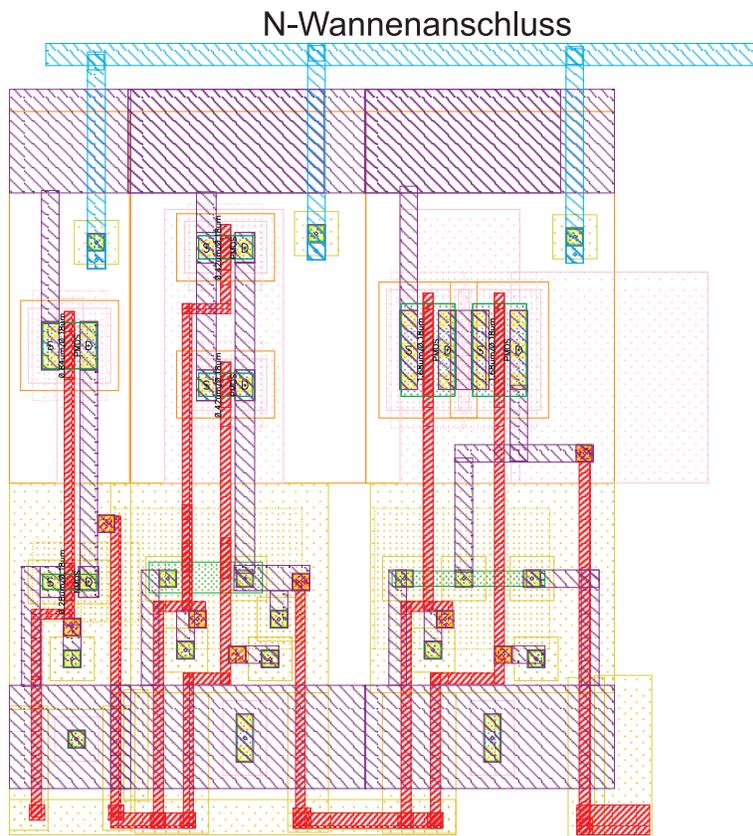


**Abbildung A.20:** Schaltplan und Layout RIC, groß dimensioniert



Der Schaltplan zu obiger Abbildung ist analog zu dem aus Abbildung A.20, jedoch mit 127 Invertern und wird aus Gründen der Übersichtlichkeit nicht mit angezeigt.

**Abbildung A.21:** Layout RIC, minimal dimensioniert



Die in Kapitel 2.8 beschriebene Testschaltung besteht aus 100 seriellen Strukturen (Inverter, Nand2 und Nor2) der obigen Abbildung.

**Abbildung A.22:** Layout Teststruktur für Bulksteuerung



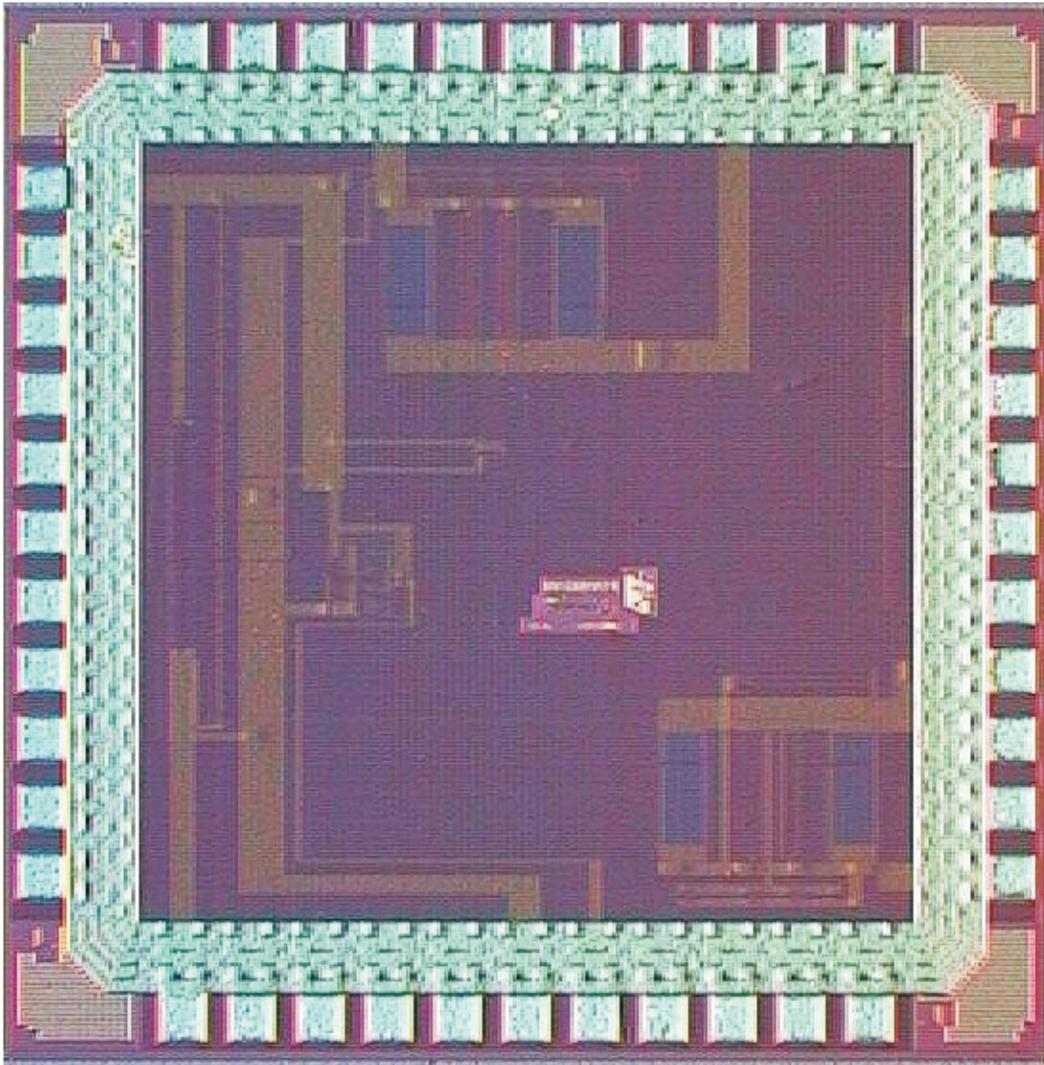


Abbildung A.24: Chipfotografie

# Literaturverzeichnis

- [1] J. Burns, A. Drake, R. Brown. *Dynamic-Threshold Logic for Low-Power VLSI Design* . University of Michigan, 2001.
- [2] S. J. Abou-Samra and A. Guyot. Performance/Complexity Space Exploration : Bulk vs. SOI . In *International Workshop on Power And Timing Modeling Optimization and Simulation*, 1998.
- [3] Douglas Adams. *So long, and Thanks for All the Fish* . Random House Publishing Group, 1999.
- [4] Holger Arweiler. Messungen an kapazitiven Buck-Konvertern. *Studienarbeit am Lehrstuhl für Mikroelektronik, Universität des Saarlandes*, 2007.
- [5] International Semiconductor Industry Association. *International Technology Roadmap For Semiconductors 2002 Update*, 2003.
- [6] Shao Bin, Yang Yujia, Hong Zhiliang, and Xu Chihao and Dirk Killat. A Capacitive Step-Down Converter Using a Linear Mode Pre-Regulator for Improved Load Regulation. *International Conference on Solid-State and Integrated-Circuit Technology, Shanghai, 2006*, 2006.
- [7] Y. Manoli, C. Hitzelberger. Low-Power digital design in CMOS technology. *European Solid-State Circuits Conference (ESSCIRC 2000) MEDEA Project Workshop, Stockholm, 2000*, 2002.
- [8] Y. Manoli, C. Hitzelberger, R. Hakenes, S. Gross. A Microcontroller Embedded ASIC for an Implantable Electro-Neural Stimulator. In *ESSCIRC01*, 2001.
- [9] A.P. Chandrakasan. *Low Power Digital CMOS Design* . Kluwer Academics Publisher, 1995.
- [10] Yin-Shuin Kang, Chingwei Yeh, Min-Cheng Chang. Algorithms Promoting the Use of Dual Supply Voltages for Power-Driven Designs. In *20th Anniversary Conference on Advanced Research in VLSI*, 1998.

- [11] Ira Miller, Dan Fitzpatrick. *Analog Behavioral Modeling With The Verilog-A Language*. Kluwer Academic Publisher, 1998.
- [12] Dennis Sylvester, Dongwoo Lee. Gate Oxide Leakage Current Analysis and Reduction for VLSI Circuits. *IEEE TRANSACTIONS ON VERY LARGE SCALE INTEGRATION (VLSI) SYSTEMS, VOL. 12, NO. 2, FEBRUARY 2004*, 2004.
- [13] Andreas G.Andreou, Edegar Sánchez-Sinencio. *Low-Voltage/Low-Power Integrated Circuits and Systems*. IEEE Press, 1998.
- [14] K. M. Cao et al. BSIM4 Gate Leakage Model Including Source-Drain Partition. *EDM Meeting 2000, IEDM Technical Digest, San Francisco, CA, pp. 815-818, December 10-13, 2000*, 2000.
- [15] M. Steyaert et al. Threshold voltage mismatch in shortchannel MOS transistors. *ELECTRONICS LETTERS 1st September 1994 Vol. 30 No. 18*, 1994.
- [16] S. Martin et al. Combined Dynamic Voltage Scaling and Adaptive Body Biasing for Low Power Microprocessors under Dynamic Workload. In *ICCAD*, 2002.
- [17] S. Misaka et al. Application to cellular phones. *Frequency-Voltage Cooperative Power Reduction for Multi-tasking Multimedia Applications, COOL-Chips VI, Vol.I, pp.103-116, April 2003*, 2003.
- [18] S. Mutoh et al. 1-V Power Supply High-Speed Digital Circuit Technology with Multithreshold Voltage CMOS. *III journal of Solid-State Circuits, Vol. 30, No.8*, 1995.
- [19] Siva Narendra et al. Full-chip subthreshold leakage power prediction and reduction techniques for sub-0.18- $\mu\text{m}$  CMOS. *IEEE JOURNAL OF SOLID-STATE CIRCUITS, VOL. 39, NO. 2, FEBRUARY 2004*, 2004.
- [20] D. Fotty. *MOSFET Modeling With Spice*. Prentice Hall PTR, 1997.
- [21] Andrea Ghetti. Characterization and Modeling of the Tunneling Current in  $\text{Si}/\text{SiO}_2/\text{Si}$  structures with ultra thin oxide layer. In *Presentation Infos 12<sup>th</sup> Bio-annual Conference*, 2001.
- [22] Chenming Hu. Gate Oxide Scaling Limits and Projection. *Technical Digest of International Electron Devices Meeting*, pp. 319-322, December 1996., 1996.
- [23] IBM. *SOI Technology, IBM's Next Advance In Chip Design*. IBM Online Publication.

- [24] Intel. Transistor Elements for 30nm Physical Gate Length and Beyond . *Intel Technology Journal Volume06 Issue02 ISSN 1235766X*, 2002.
- [25] D.Antoniadis, J. Kao, A. Chandrakasan. *Transistor Sizing Issues and Tools For Multi-Threshold CMOS Technology*. Design Automation Conference, 1998.
- [26] R. A. Hadaway, K. R. Lakshmikumar and M. A. Copeland. Characterization and Modeling of Mismatch in MOS Transistors for Precision Analog Design . *IEEE Journal of Solid-State Circuits*, 1986.
- [27] Nikhil M. Kriplani. Transistor Modeling using Advanced Circuit Simulator Technology . *A dissertation submitted to the Graduate Faculty of North Carolina State University in partial fulfillment of the requirements for the Degree of Master of Science*, 2002.
- [28] Guenter Lau. Matching of Primitive Devices for Analog and Mixed Signal Applications. *MOS-AK meeting, 21 Oct. 2002*, 2002.
- [29] C. Hitzelberger, M. Becker and Y. Manoli. LowVoltage Concepts. *IEEE International Conference on Computer Design, Workshop on Low Power Design Methodologies*, 2002.
- [30] A. C. J. Duinmaijer, M. J. M. Pelgrom and A. P. G. Welbers. Matching Properties of MOS Transistors . *IEEE Journal of Solid-State Circuits*, 1989.
- [31] Vivian Ma. SOI VS CMOS FOR ANALOG CIRCUIT . *Tutorial Papers on Analog Circuits, University of Toronto*, 2001.
- [32] B.s. NATARAJ and JENDRA Kuwr. Threshold-Voltage Variations in VLSI MOSFETS Due to Short Channel Lengths. *IEEE JOURNAL OF SOLID-STATE CIRCUITS, VOL. SC-22, NO. 5, OCTOBER 1987*, 1987.
- [33] E. J. Nowak. Maintaining the benefits of CMOS scaling when scaling bogs down . *IBM journal of research and development, Volume 46, Numbers 2/3, 2002*, 2002.
- [34] Hyungcheol Shin, Oonho Gil Minkyu Je, Jongho Lee. A high speed and low power SOI inverter using active body-bias. In *ISLPED 1998: 59-63*, 1998.
- [35] Philips-Semiconductor. *The MOS model, level 903* , 1998.
- [36] Philips-Semiconductor. *The MOS model, level 1101* . Online Publication, 2002.

- [37] Philips-Semiconductor. *The MOS model, level 1102*. Online Publication, 2006.
- [38] Douglas R. Holberg, Phillip E. Allen. *Analog Circuit Design - 2nd Edition*. Oxford University Press, 2002.
- [39] Jan M. Rabaey. *Digital Integrated Circuits - A Design Perspektive*. Prentice Hall, 1996.
- [40] Behzad Razavi. *Design of Analog CMOS Integrated Ciruits*. McGraHill Higher Education, 2001.
- [41] Yiannos Manoli, Rolf Hakenes. A Segmented Gray Code for Low-Power Microcontroller Address Buses. In *EUROMICRO 99 Workshop on Digital System Design*, 1999.
- [42] Kaushik Roy. *Low Power CMOS Design, Lectures*. Electrical and Computer Engineering, Purdue University, 2000.
- [43] K. Phang, S. Nicolson. STEP-UP vs STEP-DOWN DC/DC converters for RF-powered systems. In *ISCAS*, 2004.
- [44] Arijit Raychowdhury Saibal Mukhopadhyay and Kaushik Roy. Accurate Estimation of Total Leakage Current in scaled CMOS Logic Circuits Based an Compact Current Modeling. *Proceedings of the 40th Conference on Design automation*, pages 169–174, 2003.
- [45] Kai Schmerer. AMD Athlon 64 mit Cool n Quiet . <http://www.zdnet.de>, 2003.
- [46] Tobias Schneider. Konzepte zur Reduzierung der statischen Verlustleistung in digitalen C-MOS Schaltungen . *Studienarbeit am Lehrstuhl für Mikroelektronik, Universität des Saarlandes*, 2003.
- [47] Vivek De, Sean T. Ma, Ali Keshavarzi and John R. Brews. A Statistical Model for Extracting Geometric Sources of Transistor Performance Variation. *IEEE TRANSACTIONS ON ELECTRON DEVICES, VOL. 51, NO. 1, JANUARY 2004*, 2004.
- [48] U. Grünebaum und Jürgen Oehm. *General Analysis of Mismatch Effects*, 1999. Anleitung zum Matching Analyse Tool *GAME*.
- [49] John P. Uyemura. *CMOS Logic Circuit Design*. Kluewe Academic Publishers, 1999.
- [50] A.K.P. Viraj. Analysis of switched capacitor DC-DC step down converter . In *ISCAS*, 2004.

- 
- [51] C. Hitzelberger, Y. Manoli, C. Xu. Parameter Specific Supply Voltage Scaling by Evaluating the Frequency of an On Chip Ring-Inverter-Chain. In *IEEE International Conference On Semiconductor Electronics*, 2004.
- [52] F. Gerfers, Y. Manoli, J. Peters. *Einführung in den Entwurf von VLSI-Systemen*, 1996.