NAWI Graz
Natural Sciences

TU Graz

Bernd Schuscha, BSc

Evaluation of electro-thermal influenced reliability effects in microelectronics

MASTER'S THESIS

to achieve the university degree of

Diplom-Ingenieur

Master's degree programme:

Technical Physics

submitted to

Graz University of Technology

Supervisor

Univ.-Prof. Ph.D. Peter Hadley

Institute for Solid State Physics

in cooperation with

ams AG (R&D EDA)

Graz, August 2021

AFFIDAVIT

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present master's thesis.

Graz, 13.07.2021 B. Schuscha

Date, Signature

Acknowledgement

I would first like to thank my supervisor, Univ.-Prof. Peter Hadley, head of the Institute of Solid State Physics from the Technical University of Graz. His door was always open and I would like to thank him for providing guidance and feedback throughout this thesis.

I would like to acknowledge my colleagues from the Technology R&D EDA department at ams AG in Premstätten, managed by Ehrenfried Seebacher, it was a delight working with them. I would particularly like to single out my supervisors Alexander Steinmair and Chong Jin. They were both kind supervisors to whom I could always come with questions.

I would also like to thank Klaus Cerny and Evgeny Vlasov from Cadence Design Systems. Without their continuous technical support and them providing the licenses the project would not have been possible.

Abstract

Temperature management is import for every high power device. Applying a temperature profile not only has an immediate effect on the electrical device characteristics but also causes long term changes through degradation. The two most important degradation effects which depend on electrical stress are Hot-Carrier-Injection (HCI) and Negative-Bias-Temperature-Instability (NBTI). This thesis aims to test one commercial solution, the Legato tool, to simulate the coupling of such effects.

To test this, simulations of test circuits have been performed. Typical reliability models used for the degradation effects were implemented, the model data was provided by ams AG. Because the tool only couples thermal and electrical properties, a provisional interface to the reliability simulation was made to evaluate whenever it is necessary in the future. A MOS driver was chosen as the main test circuit. The results for the MOS driver show a linear temperature dependence on the duty cycle, which suggests the heat sink at the bottom is functioning well. In addition, a spatially-dependence of electrical and reliability characteristics could be found, which is rooted in the temperature distribution of the chip.

The results show the importance of the device position of a chip, but also that reliability effects could be used to compensate for a difference in device location. Furthermore, the limit of the tool could be found which traded usability with customizability.

Kurzfassung

Temperatur Management ist in Mikroelektronik-Anwendungen mit hoher Leistungsdichte von großer Bedeutung. Dies ist, aufgrund der Temperaturabhängigkeit von elektrische Eigenschaften von Schaltungen oder Bauteilen. Dabeo ist dies nicht nur einen kurzzeitigen und reversible Character, sondern es kann durch Degradierung (Alterung) auch zu dauerhaften Veränderung von Bauteilen kommen. Die zwei bedeutendsten Effekte in MOSFETs sind hierfür Hot-Carrier-Injection (HCI) und Negative-Bias-Temperature-Instability (NBTI). Das Ziel dieser Arbeit ist die Testung einer kommerziellen Lösung, das „Legato Tool", welches eine Kopplung solcher Effekte erlaubt.

Um dies zu testen, wurden Simulationen von Testschaltungen durchgeführt. Für die Degradationseffekte wurden typische Modelle implementiert, bereitstellt von der ams AG. Da das Tool nur thermische und elektrische Eigenschaften koppelt, wurde eine vorläufige Schnittstelle zur Degradierungssimulation erstellt. Dies erfolgte für eine Bewertung, ob eine Berücksichtigung von Temperatureffekten bei der Degradierungssimulation erforderlich ist. Für die Testung wurde ein MOS-Treiber als Haupttestschaltung gewählt. Die Ergebnisse für diesen MOS-Treiber zeigen eine lineare Temperaturabhängigkeit vom Duty-Cycle, was auf eine gute Funktion des Wärmesenke an der Unterseite des Bauteils schließen lässt. Außerdem konnte ein Ortabhängigkeit des Elektrischen- und Degradierungverhalten festgestellt werden, das in der Temperaturverteilung des Chips begründet ist.

Die Ergebnisse zeigen die Bedeutung der Geräteposition eines Mikroelektronikchips, aber auch, dass Degradierungeffekte verwendet werden könnten, um einen Unterschied vom Geräteposition auszugleichen. Darüber hinaus konnte die Grenze des Tools gefunden werden, das Benutzerfreundlichkeit mit Anpassbarkeit tauscht.

Contents

Acknowledgement iii

Abstract iv

Kurzfassung v

1 Introduction 1
1.1 Basics of MOSFETs 1
1.1.1 Different Regimes, MOSFET Model and Characteristic Curves 2
1.2 Power-MOSFET 3
1.2.1 Multi-Finger MOSFET 4
1.2.2 LDMOS 4
1.3 Test Circuits 5
1.3.1 29V HV pMOS 5
1.3.2 Current Mirror 6
1.3.3 MOS Driver 8

2 Electro-Thermal Simulation 10
2.1 Temperature Behavior of MOSFETs 10
2.1.1 Thermal Runaway 11
2.2 Multi-Scale Simulation 18
2.3 Simulation Method 19
2.3.1 Power Input Profile Legato vs Typical Standard FEM Simulation 20
2.3.2 Thermic Connection to the Environment 21
2.3.3 Heat Equation 24
2.4 Simulation Results 26
2.4.1 Electro Thermal Simulation of a Point Source with the 180 nm Technology Layer Stack 27

2.4.2 Temperature Behaviour of a Single p-MOSFET 29
2.4.3 Electro-Thermal Simulation of a Current Mirror 30
2.4.4 Influence of Simulation Variables for the MOS Driver Simulation 32
2.4.5 Electro-Thermal Simulation of a MOS driver 34

3 Ageing Simulation 40
3.1 Reliability 40
3.1.1 Reliability Basics 40
3.2 Degradation Mechanism of MOSFET 41
3.2.1 Hot Carrier Injection (HCI) 42
3.2.2 Negative Bias Temperature Instability (NBTI) 43
3.3 Reliabilty Modeling 44
3.3.1 Takeda Reliability Model 44
3.3.2 HCI Model 45
3.3.3 NBTI model 46
3.3.4 Implementation 46
3.3.5 AC Stress and the BERT Algorythm 47
3.4 Simulation Results 49
3.4.1 29V HV pMOS Reliability Simulation Results 50
3.4.2 Current Mirror Reliability Simulation Results 51
3.4.3 MOS Driver Reliability Simulation Results 52

4 Conclusion 55
4.1 Legato Tool for Electro-Thermal Simulation 55
4.2 Electro-Thermal Simulation Results 56
4.2.1 Current Mirror 56
4.2.2 MOS Driver 56
4.3 Electro-Thermal Aware Reliability Simulation Results 56

Bibliography 58

List of Tables

1.1 Layouts 8

2.1 Used layer stack 26
2.2 Thermal constants used for the simulation 27
2.3 Layout variants of the current mirror 32
2.4 Results for the symmetric layout 32
2.5 Electro thermal simulation result 37
2.6 Comparisson between Grp0 and Grp3 39

3.1 Current mirror aging with temperature profile results for T = 25 °C 51
3.2 Current mirror aging with temperature profile results for T = -40 °C 52
3.3 Electro-thermal reliability simulation result 53

List of Figures

1.1 Schematic of a p-MOSFET 2
1.2 Schematic of an p-MOSFET 3
1.3 MOSFET with three fingers 4
1.4 Cross-section of an LDMOS 5
1.5 Circuit diagram for the p-MOSET test circuit 6
1.6 Circuit diagram for the current mirror test circuit 7
1.7 Layout of the current mirror 7
1.8 Most important components of the MOS driver 8
1.9 Left: Most important components of the MOS driver group; Right: group Layout 9

2.1 Charge carrier concentration for n-doped silicon with a doping concentration of 10^{17} $1/cm^3$ 13
2.2 Left: Conducvity; Right: Resistivity 14
2.3 Temperature dependence of the diode current at V = 0.5 V . . . 15
2.4 Output characteristics and transfer characteristics of the 29V HV pMOS transistor depending on the temperature 16
2.5 Sketch of transfer characteristic visible in figure 2.4; Inversion range of the IDS in dependence of the VGS 17
2.6 Partial derivation of IDS over T 18
2.7 Block diagramm for the general simulation flow 20
2.8 Top: power profile of a standard FEM simulation; Bottom: Legato tool power profile 21
2.9 Left: $\frac{R_{th}}{A}$ for radiation and convection; Right: $\frac{R_{th,ges}}{A}$ region for different convection modes 24
2.10 Equilibrium temperature distribution of a point source with $P = 1.94\ W$ 27
2.11 Temperature development of different elements/points with time 28
2.12 Tempeature profile at $y = 1875\ \mu$m for 40 ms at equilibrium . . 29

2.13 Left: drain-source current for the operating points of the MOS driver and the current mirror; Right: Perceptual change of the drain-source current with the current at 25 °C as a baseline . . 30
2.14 Temperature profile of the current mirror 31
2.15 Left: mean finger temperature; right: finger current 31
2.16 Output current for different uniform temperatures 32
2.17 Temperature dependence of the thermal resistance for Grp0 and Grp3 33
2.18 Temperature dependence of the thickness of the glue layer of Grp0 and Grp3 33
2.19 Temperature dependence of the thickness of the substrate of Grp0 and Grp3 34
2.20 Temperature distribution for different duty cycles 35
2.21 MOS current for different uniform temperatures and for the associated temperature profile for $T_{umg} = 29$ and $D = 10$ % . 36
2.22 Left: settling current with uniform temperature; Right: maximum current with uniform temperature 36
2.23 Left: pulse times, switch on and switch off; Right: pulse length 37
2.24 Temperature distribution if Grp0 or Grp3 is used with D = 10 % 38
2.25 Electro thermal simulation result of the MOS current for Grp0 and Grp3 38

3.1 Failure rate over time, the typical bathtub curve 41
3.2 Schematic of the HCI [21] 42
3.3 Schematic of the RD model of NBTI 44
3.4 Reliability simulation workflow 49
3.5 Left: Change with operation point with temperature; Right: Change of Operation point due to NBTI ageing for an age time of 1 year 50
3.6 Change in output of the current mirror test circuit for age time of 10 years and uniform temperature; Left: current mirror outputs fresh and aged; Right: percentual change, equal for both outputs because of uniform temperature 51
3.7 Fresh vs aged results; Left: settle current with uniform temperature; Right: maximum current with uniform temperature . . . 52
3.8 Fresh vs aged results; Left: pulse times, switch on and switch off; Right: pulse length 53

3.9 Driver current fresh and aged for different groups 54

1 Introduction

1.1 Basics of MOSFETs

The Metal-Oxide-Semiconductor Field-Effect-Transistor or MOSFET for short is a transistor widely used in microelectronics. [1] As the name suggests, the controlling of the transistor works with an electric field instead of a current flow. This allows a low power controlling of the transistor. The manipulation with an electric field is achieved via a capacitor at the connection between the source and drain called ”gate”.

The capacitor is built with the bulk connected to metal or for better processability with poly-silicon. A nonconductive layer is used as dielectrics, and the doped bulk silicon is connected to the bulk electrode. This structure allows change in resistance between the drain and source electrode by several orders of magnitude while keeping the current through the gate orders of magnitude lower than the drain-source-current.

The area under the source- and drain-electrodes is p-doped (in the case of a p-MOSFET) while the rest of the bulk is n-doped, which results in a nonconduction depletion layer around the source and drain. The area near the silicon oxide between source and drain is called the ”channel”. Here most of the charge transport is taking place. The bulk also has a connection to prevent charging of the bulk; for that reason, it is often grounded. [1][2][3] This structure can be also seen in figure 1.1.

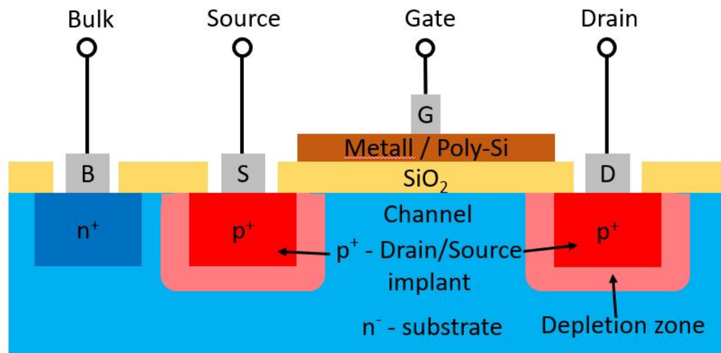

Figure 1.1: Schematic of a p-MOSFET

1.1.1 Different Regimes, MOSFET Model and Characteristic Curves

As explained, the MOSFET device uses the capacitor between gate and channel to manipulate the resistance between source and drain. This is done via the electric field of the capacitor. If the gate is biased negatively in the case of a p-MOSFET, an electric field is induced, which draws bulk holes to the interface between bulk and silicon oxide. As a result, a hole enriched channel between the p doped source and drain regions develops. Overall three different regimes can be seen if a MOSFET is biased in this way:

If the gate bias is smaller than a specific voltage called the threshold V_{th} voltage, this is called the subthreshold regime. In this regime, the applied bias is too small to fully invert the channel. Figure 1.2 shows the depletion zone in this regime for a p-MOSFETs. It is characterized by a small current between source and drain. This current can be approximated with 0.

When the gate bias is increased by V_{th}, the operation area must also be divided with source-drain-bias. One where $V_{DS} > V_{GS} - V_{th}$, there the electric field is big enough for an inversion layer to be built, which means the two p-doped areas are also connected with a p-doped channel. That means the mobile holes can contribute to charge transfer. Figure 1.2 shows this conducting p-type channel for a p-MOSFET in the linear regime. In this region, the current goes linearly with the applied gate-bias and is called linear regime. The other case is $V_{DS} < V_{GS} - V_T$, there the source-drain current saturates, saturated regime.

Here the holes accumulate at the interface, which cannot contribute to the drain current. [1][2][3] Figure 1.2 shows this pinch-off for the saturation of an p-MOSFET.

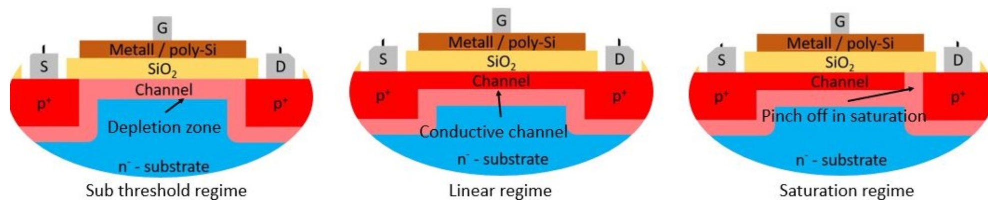

Figure 1.2: Schematic of an p-MOSFET

With the assumption that the inversion layer charge is proportional with the applied voltage this can also put in the form of equations[2]:

$$
\begin{aligned}
I_D &= 0 & &\begin{cases} V_{GS} > V_{th} \end{cases} \\
I_{D,lin} &= -\mu_p \cdot C_{ox} \frac{W}{L} (V_{GS} - V_{th} - \frac{V_{DS}}{2}) V_{DS} & &\begin{cases} V_{GS} \leq V_{th} \\ V_{DS} > V_{GS} - V_{th} \end{cases} \\
I_{D,sat} &= -\mu_p \cdot C_{ox} \frac{W}{2L} (V_{GS} - V_T)^2 & &\begin{cases} V_{GS} \leq V_{th} \\ V_{DS} \leq V_{GS} - V_{th} \end{cases}
\end{aligned}
\tag{1.1}
$$

I_D	Drain current [A]	V_{th}	Threshold voltage [V]
$I_{D,lin}$	I_D linear regime [A]	W	Transistor width [μm]
μ_p	Electron Mobility [$\frac{m^2}{Vs}$]	L	Channel length [μm]
V_{GS}	Gate-source potential [V]	C_{ox}	Oxid capacitance [F]
V_{DS}	Drain-source potential [V]		

This is one of the simplest models for the MOSFET device.

1.2 Power-MOSFET

Depending on the application the basic MOSFET design was adopted to get a higher functionality. For high current/ high power applications, the relatively

high parasitic gate resistance had to be reduced. In the examples in this thesis, this is done in two ways.

1.2.1 Multi-Finger MOSFET

Here the width W of the MOSFET is split up nf-times, where nf is the number of fingers. Instead of a single transistor, nf parallel-connected smaller ones with the width W/nf are gotten, each drain and source is used doubled. The resulting structure is shown in figure 1.3. This has the advantage of a layout with an aspect ratio ($\frac{W}{L}$) closer to one, which is easier to use. And it enhances the AC capabilities of the transistor by reducing the parasitic capacitance because of the parallel connection. [4]

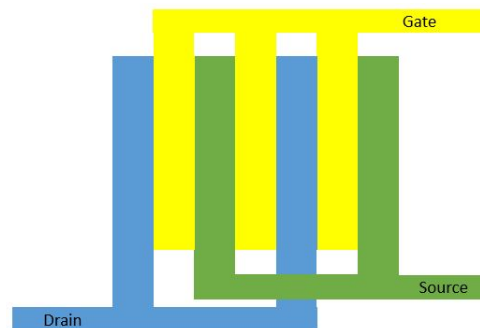

Figure 1.3: MOSFET with three fingers

1.2.2 LDMOS

An asymmetric variation of the standard MOSFET used in high voltage applications is called the Laterally Diffused MOSFET (LDMOS). It is designed for a low on-resistance and high blocking voltage. The cross-section can be seen in figure 1.4. The channel is short and not like a standard MOSFET a p^+-n^--p^+ but a p^+-n-p-p^+ transition in which the drain side is relatively weak doped. This results in a large depletion layer between the n-p and a small one between p^+-n. The consequence of this is a high blocking voltage with a short channel length because the diffused p-region results in a low on-resistance

and high current capability. The device got its name from the deep n-type region around the source and gate, made via diffusion. [1]

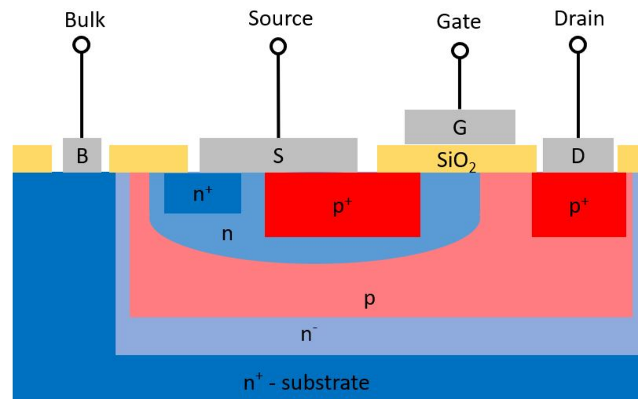

Figure 1.4: Cross-section of an LDMOS

1.3 Test Circuits

In this thesis, electro-thermal circuit simulation and its application in aging is investigated. For this reason, five example circuits were chosen. Because the investigation of the circuit extends over more than one chapter, this section will give an overview of the circuit and how they can be investigated.

1.3.1 29V HV pMOS

The p-MOSFET of the current source in the MOS driver is a multi finger MOSFET with a width of 400 nm and a length of 1.8 um. It was chosen as a test circuit to assess if a finger resolved aging analysis adds value and for the investigation of the influence on boundary conditions of the thermal simulation.

The circuit diagram for this can be seen in figure 1.5. The drain-source voltage will be held constant while the gate will be pulsed. This simulates the operation conditions into the driver circuit.

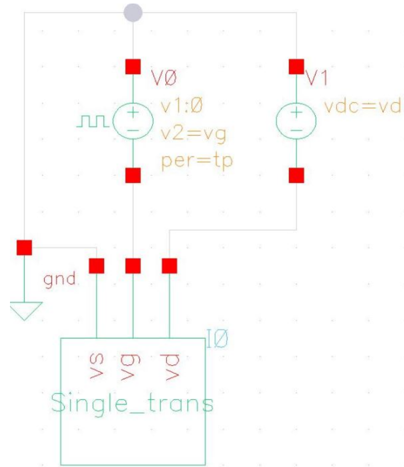

Figure 1.5: Circuit diagram for the p-MOSET test circuit

1.3.2 Current Mirror

A current mirror is an elementary circuit that is used to generate currents out of a reference current. Application of it is analog-digital converter or current sources.

The goal was to demonstrate the layout dependent electrical performance of the circuit. There are multiple reasons why this structure was chosen. It is widely used in different variations, and it can be build up out of the same device blocks. Furthermore, the devices can be split up and placed independently in the layout. This is done to investigate the impact of the layout on the overall CM behaviour.

For this it is necessary to complete an electrothermal analysis and an electrothermal aware aging analysis. The circuit diagram can be seen in figure 1.6.

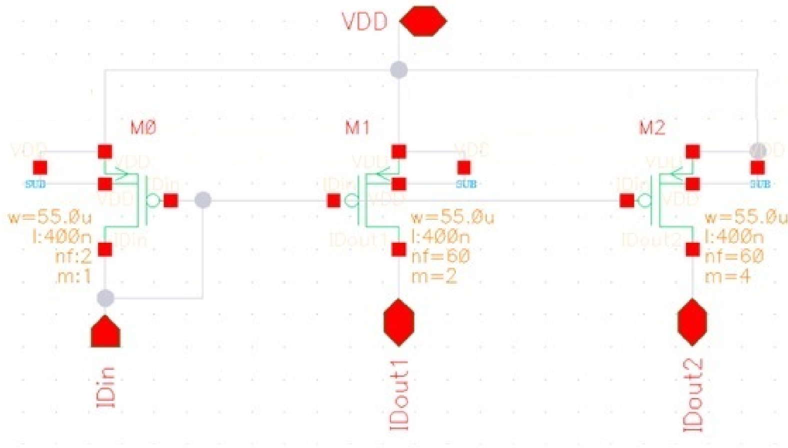

Figure 1.6: Circuit diagram for the current mirror test circuit

The transistors of the two outputs are divided into blocks of 60 fingers. Output 1 (M1) has two blocks, and output 2 (M2) has four. The way in which these six blocks are arranged can be seen in figure 1.7.

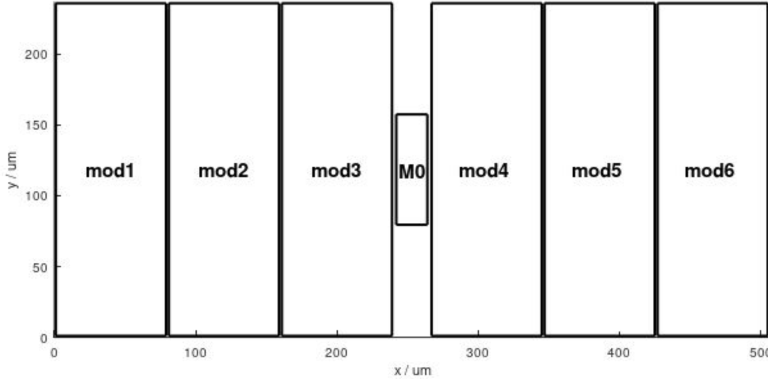

Figure 1.7: Layout of the current mirror

The division of the blocks for the two layouts can be seen in table 1.1.

Table 1.1: Layouts

Layout	M1	M2
1 (symetric)	mod2, mod5	mod1, mod3, mod 5, mod6
2 (staggerad)	mod1, mod2	mod3, mod4, mod 5, mod6

1.3.3 MOS Driver

The MOS driver is part of a used case. It can supply pulses of a few hundred mA up to several A. This pulse can have a pulse length of hundred ns to several us.

As can be seen in figure 1.8 it consists of X identical groups next to a controlling and feedback area on the right side which is used to regulate the output (figure 1.8). The groups can be divided into four parts. The driver bias, which feeds into the current source driver, these two modules control the output current. The current source, the primary power consumer of the circuit and its main components, x 29V HV pMOS connected in parallel, can be electronically inserted and removed. After that, a output can be selected with a multiplexer (MUX). Under the multiplexer and in the controlling and feedback area, there are temperature sensors from which the temperature can be read out electronically.

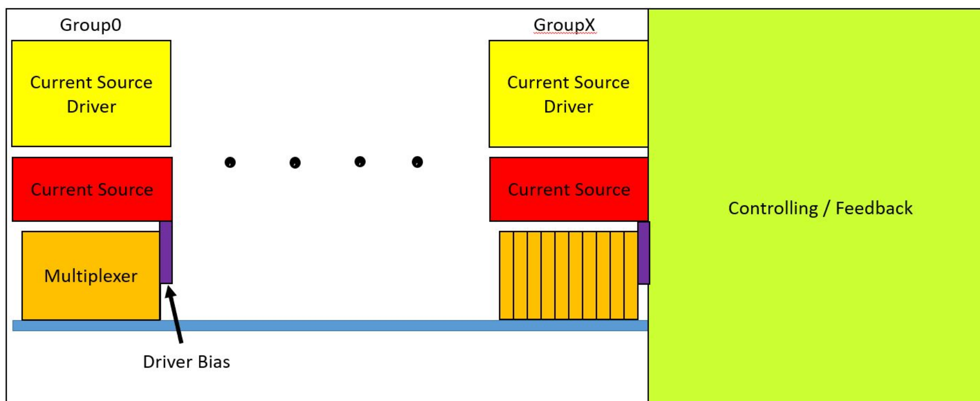

Figure 1.8: Most important components of the MOS driver

In figure 1.9 a simplified schematic of one such module can be seen. The

output current is mainly produced with the drain-source resistance of the 29V HV pMOS. The high voltage on the source contact of the transistor is kept constant.

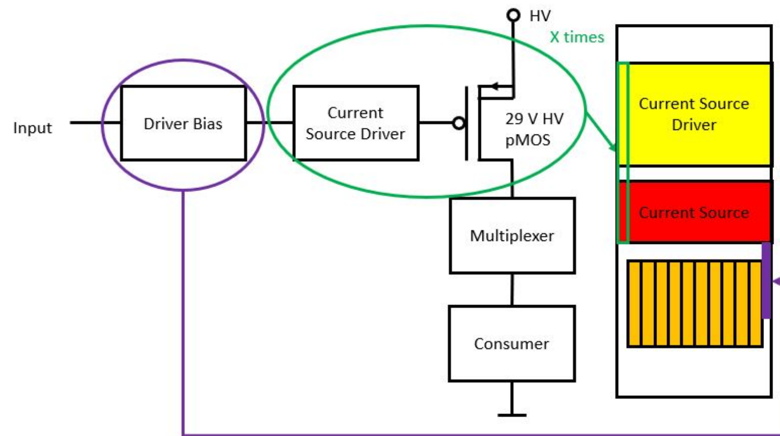

Figure 1.9: Left: Most important components of the MOS driver group; Right: group Layout

The goal is to find a relation between pulse length and channel temperature. The relationship between temperature and boundary condition will be analysed. This is necessary to give a confidence interval to be able to compare the result with measurement data from the temperature sensors. A comparison between the performance change of two groups will then be made. Lastly, the newly implemented temperature-sensitive aging workflow will be tested with this circuit. For this, an NBTI and HCI model was implemented. The result will be compared with the standard aging workflow.

2 Electro-Thermal Simulation

2.1 Temperature Behavior of MOSFETs

A temperature change in the MOSFET area results in updated device characteristics. This is because of the temperature dependence of the band-gap E_g and an increase in phonons density. This higher phonon density results in more charge carrier scattering with the lattice, which decreases electron mobility. The outcome of this is that both the absolute value of the threshold voltage [5] and the charge carrier mobility decrease with temperature. As a simplified model, the threshold voltage can be assumed to depend linear on temperature.[6][7]

$$V_{th}(T) = V_{th}(T_0) + \alpha_{Vth}(T - T_0) \tag{2.1}$$

α_{Vth}	Changing rate of threshold voltage
with:	$sign(\alpha_{Vth}) = -sign(V_{th}(T_0))$

and for the mobility, a power law can be assumed [6][7]

$$\mu(T) = \mu(T_0)(T/T_0)^{\alpha_\mu} \tag{2.2}$$

$\mu(T_0)$	Charge carrier mobility $[\frac{\mathrm{cm}^2}{\mathrm{Vs}}]$
α_μ	Constant dependent on the scattering in the inversion layer []
with	$\alpha_\mu < 0$ (typical values Si-Technology: -1.2 to -2.42) [6]

If (2.1) and (2.2) are substituted in (1.1) to get a temperature dependence for the first order equation of $I_{D,sat}$ of a PMOS.[7]

$$I_{D,sat} = -\mu_0(T/T_0)^{\alpha_\mu} \cdot C_{ox}\frac{W}{2L}(V_{GS} - V_{T0} - \alpha_{VT}(T - T_0))^2 \tag{2.3}$$

2.1.1 Thermal Runaway

Temperature-dependent device operation is inherent to every electrical device. This can be from a negative, in which the power decreases, or positive character, which leads to a power increase. Because consumed power also influences the temperature of a device, this results in a feedback loop. Negative feedback loops could be described as "self-cooling," while a positive feedback loop between consumed power and an increase in temperature is called thermal runaway. The temperature drift resulting from this feedback loop depends on the heat dissipation in the device area.

In MOSFETs, it is strongly gate-voltage dependent on which of the two processes occurs. If V_{GS} exceeds a specific range, the device reduces its current with rising temperature and acts as a negative feedback loop which reduces its power consumption. On the other side, if V_{GS} falls below a specific range, the opposite happens. The temperature increase is proportionate to the current increase. Thus, a positive feedback loop is in place.

Gradual channel approximation

The temperature dependence of the drain-source voltage of a MOSFET eq. 2.3 can be used to show this effect. By derivation the temperature dependency of the drain-source current can be obtained using the following formula:

$$\frac{\partial I_{D,sat}}{\partial T} = I_{D,sat} \cdot \left(\frac{\alpha_\mu}{T} - \frac{2\alpha_{VT}}{V_{GS} - V_{T0} - \alpha_{VT}(T - T_0)} \right) \tag{2.4}$$

For thermal runaway, the second term must be greater than zero.

$$\left(\frac{\alpha_\mu}{T} - \frac{2\alpha_{VT}}{V_{GS} - V_{T0} - \alpha_{VT}(T - T_0)} \right) > 0 \tag{2.5}$$

Which can be reshaped into

$$V_{GS} > V_{T0} - \left(\frac{2}{\alpha_\mu} + 1 \right) T\alpha_{VT} - \alpha_{VT}T_0 \tag{2.6}$$

With the fact in mind that with a PMOS, both $I_{D,sat}$ and V_{GS} are negative, that means thermal runaway can only occur with a small gate voltage. That also means because $P \propto -V_{GS}$ the overall power consumption and the gradient is low in this range of gate voltage which again implies that thermal runaway is unlikely.

The point in which $\frac{\partial I_{D,sat}}{\partial T}$ is equal to zero is called zero temperature coefficient (ZTC) point and only exists if α_μ is exactly -2 if not, this point expands to a small range where the derivation changes signs.

The same thing can, of course, be done with the equation of an NMOS with the same result.

Thermal activated carrier

If failures in microelectronics are analysed, it is found that silicon melts sometimes. This means that thermal runaway leads sometimes to temperatures above the melting temperature of Silicon (1414 °C). This section demonstrates a thermal runaway mechanism that can lead to device failure. The onset of it is where there are more thermally activated carriers than there are carriers due to doping, hence in the range of 600 - 8000 °C depending on the doping. The initial heating to this point might be caused by some other mechanism such as impact ionization, but if it is reached, the resistivity of the silicon drops and thermal runaway is initiated.

If the temperature in a semiconductor is raised, electrons can be thermally activated and go from the valence band to the conduction band, which produces a free electron and a hole. The resulting density of the carrier of n-doped silicon with a donor density $10^{17} \frac{1}{\mathrm{cm}^3}$ can be seen in figure 2.1. The carrier density is almost constant until 750 °C because of the donors and then approaches the intrinsic carrier density.

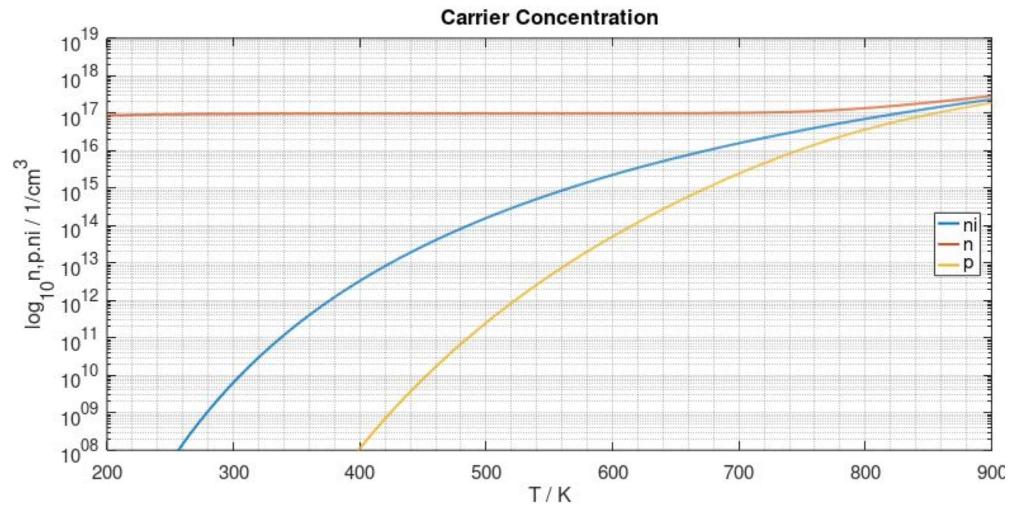

Figure 2.1: Charge carrier concentration for n-doped silicon with a doping concentration of 10^{17} $1/\text{cm}^3$

With the carrier density a conductivity and resistivity can be defined:

$$\sigma = \frac{1}{\rho} = e(p\mu_p + n\mu_n) \tag{2.7}$$

σ	Conductivity $[\frac{1}{\Omega\text{cm}}]$	σ	Resistivity $[\Omega\text{cm}]$
n	electron density $[\frac{1}{\text{cm}^3}]$	p	hole density $[\frac{1}{\text{cm}^3}]$
μ_n	electron mobility $[\frac{\text{cm}^2}{\text{Vs}}]$	μ_p	hole mobility $[\frac{\text{cm}^2}{\text{Vs}}]$

Together with the model of the temperature dependence of the mobility from eq. (2.2) a conductivity/resistivity of silicon can be calculated. The resistivity is, at constant voltage, proportional to the consumed power. If you look at figure 2.2 that means that above 750 °C, the silicon is prone to thermal runaway. This form of the conductivity curve is dominated at low temperature by the temperature dependence of the mobility while the carrier density stays roughly constant, whereas at high temperatures the carrier density dominates. This leads to a maximum in resistivity curve, which is dependent on the doping concentration.

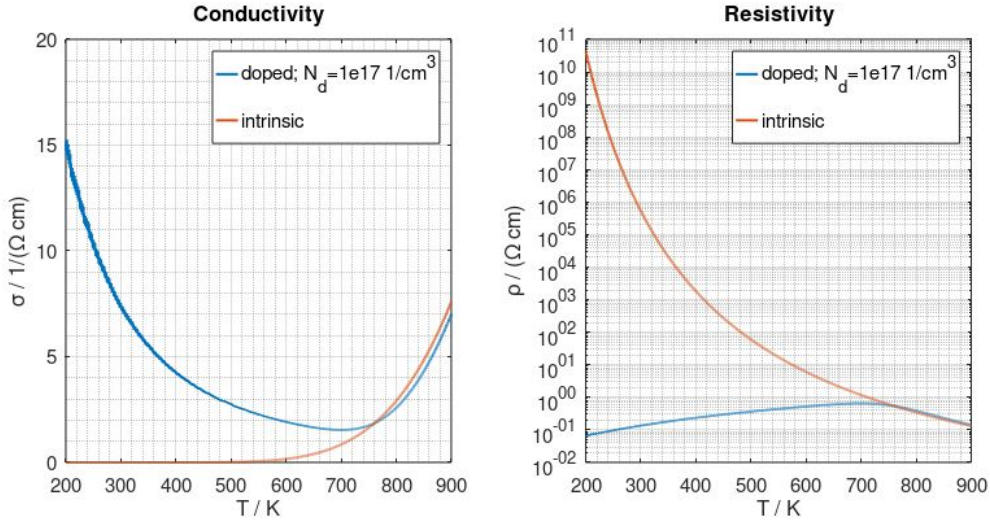

Figure 2.2: Left: Conducvity; Right: Resistivity

In a MOSFET, the current path through the channel is, as demonstrated, above thermal stable. Another possible paths are when the current goes through the reverse biased drain-bulk-diode and then either to the bulk or through the forward-biased bulk-source diode. A diode can be described with the diode equation:

$$I(V) \; = \; I_S \left(\exp(\frac{eV}{k_B T}) - 1 \right) \tag{2.8}$$

With an saturation current:

$$I_S \; = \; A\sqrt{e} n_i^2 \left(\frac{\sqrt{\mu_p k_B T}}{\sqrt{\tau_p} N_d} + \frac{\sqrt{\mu_n k_B T}}{\sqrt{\tau_n} N_a} \right) \tag{2.9}$$

τ_n Minority carrier lifetime for electrons [s] τ_p Minority carrier lifetime for holes [s]

If V is much larger than $\frac{k_B T}{e}$ ($\frac{k_B 200}{e} = 0.077$ to $\frac{k_B 900}{e} = 0.017$) the reserve biased current can be approximated as the saturation current I_S and as therefore independent of the applied voltage. This means the current of the other

paths is limited by the saturation current of the drain-bulk diode. In figure 2.3, the temperature dependence of the diode current of a forward and a reserve biased diode can be seen. The current rises with temperature, which leads to higher power consumption and thermal runaway.

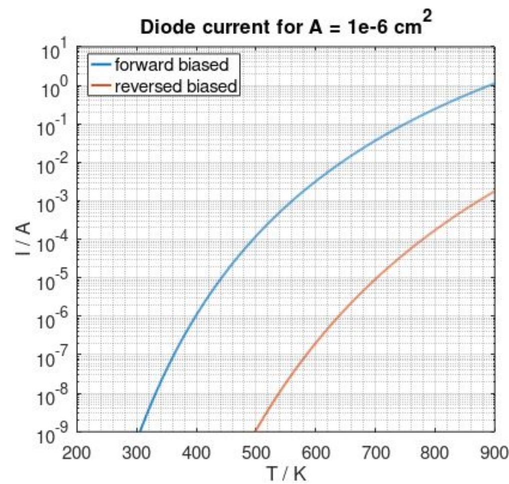

Figure 2.3: Temperature dependence of the diode current at V = 0.5 V

29V HV pMOS

While a typical SPICE (Simulation Program with Integrated Circuit Emphasize) model of a MOSFET has the effect described in chapter 2.1.1 the thermal runaway is because of thermally activated carrier not modeled. Figure 2.4 shows the result of a simulation of the output transconductance curve of the 29V HV pMOS depending on the temperature with a width of $W = 55\ \mu$m and a length of $L = 400$ nm at a drain voltage of $V_{DS} =$ -20 V.

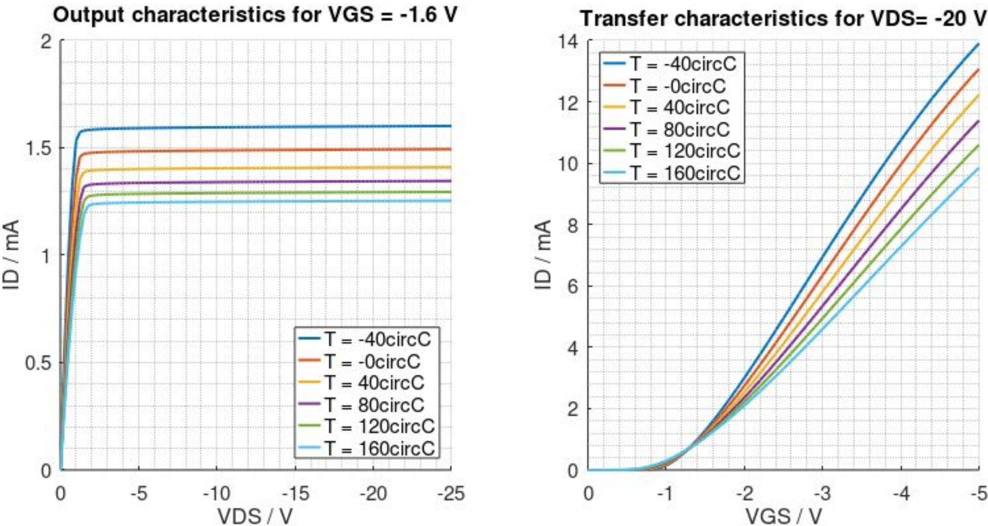

Figure 2.4: Output characteristics and transfer characteristics of the 29V HV pMOS transistor depending on the temperature

Just as derived in the gradual channel approximation, it has an intersection range from V_{GS} -1.26 V to 1.29 V. A Sketch for better visibility can be seen in figure 2.5.

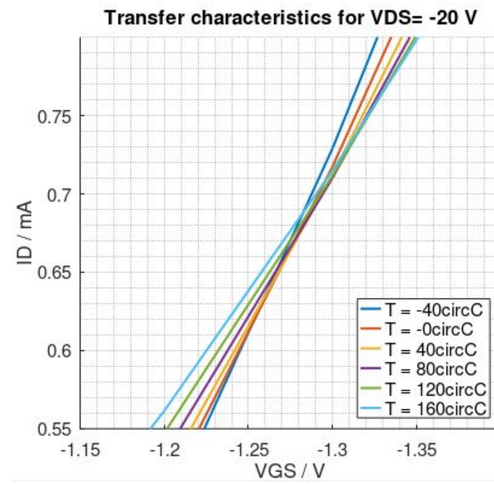

Figure 2.5: Sketch of transfer characteristic visible in figure 2.4; Inversion range of the IDS in dependence of the VGS

The other result derived from the gradual channel approximation is that sign of $\frac{\partial I_{D,sat}}{\partial T}$ is not or is only weakly dependent on the drain-source voltage. In figure 2.6 the numerical derivation of the output and transconductance curve can be seen. Only the derivation of the transconductance curve has a range where the derivation is rising.

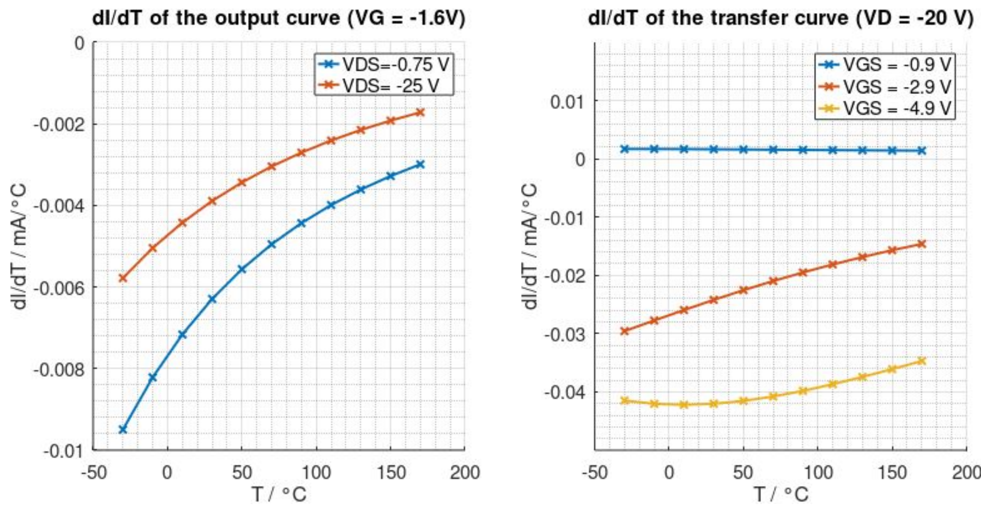

Figure 2.6: Partial derivation of IDS over T

2.2 Multi-Scale Simulation

The behaviour of an electronic device like a MOSFET can be described as dependent on the number of scattering events during the transport. For a low number or no scattering events, the Drift-Diffusion equation can be taken, and in the case of many scattering processes, the Boltzmann transport equation can be used. These two equations can be solved with numerical methods like Finite Element Methods (FEM) for a low number of devices. The problem arises with the simulation of complex devices consisting of thousands or millions of single devices. If this is attempted, the number of grid points needed is determined by the length cubed, which means the numerical method quickly reaches its limits.

To counter this, the assumption is made that devices do not influence each other electrically on a chip. This assumption is needed either way to make chip development possible. With his in mind, single device models are made via empirical data and the numerical solution of the earlier equation. To give examples a couple of examples, HiSIM2[8] uses the surface-potential, and BSIM4[9] is a modification of the result from the gradual channel approximation combined with several empirical relations. Another example of this

would be a solution with the gradual channel approximation of a MOSFET on page 3.

These models translate the inner working of the devices to a relationship between currents and voltages on different terminals. With this, the devices on a chip can be seen as a location-independent network, which allows the representation via conductivity matrix G_{ij}. The electrical behaviour of the circuit can be gotten with:

$$G_{ij} \cdot V_j \; = \; I_i \tag{2.10}$$

With these equations, the calculation of the behaviour of large circuits is possible by combining the device equations to a set of coupled (differential) equations for which standard numerical solution methods are available (eq. Newton-Rapson). [10]

2.3 Simulation Method

The simulation is done with the Legato Reliability Solution, a tool from Cadence. This tool combines the SPICE analog electronic circuit simulator [10] spectre(@TM) with a FEM simulation for the heat equation via a co-simulation. As the SPICE simulation has inherently unconsidered of device positions, the circuit layout is needed to map the consumed power of the devices to a cuboid FEM model. This 3D model consists of a layered structure with an active material, instances, wires, and a passive dielectric material. The connection to the environment in this tool is modeled by a constant thermal resistance and thermal capacity for each side, which simulates a package. This constant thermal connection is a disadvantage of this tool. An analysis of the dependence of the result on the thermal resistance is undertaken in chapter 2.4.5.

With this information, the FEM model is solved. This result can be plugged in the electrical simulation again to update the temperature dependent model parameter of the compact models to iterate until the solution converges.

The simulation flow can be seen in figure 2.7.

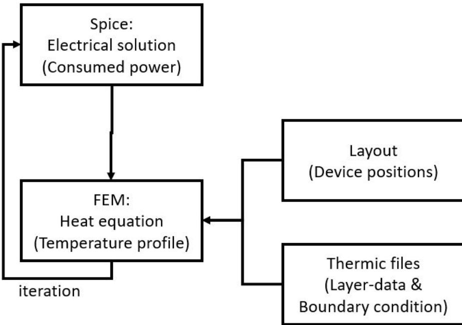

Figure 2.7: Block diagramm for the general simulation flow

2.3.1 Power Input Profile Legato vs Typical Standard FEM Simulation

If this workflow is compared with a standard FEM simulation, the advantage is the detailed power profile obtained from the SPICE simulation. An example of this can be seen in figure 2.8.

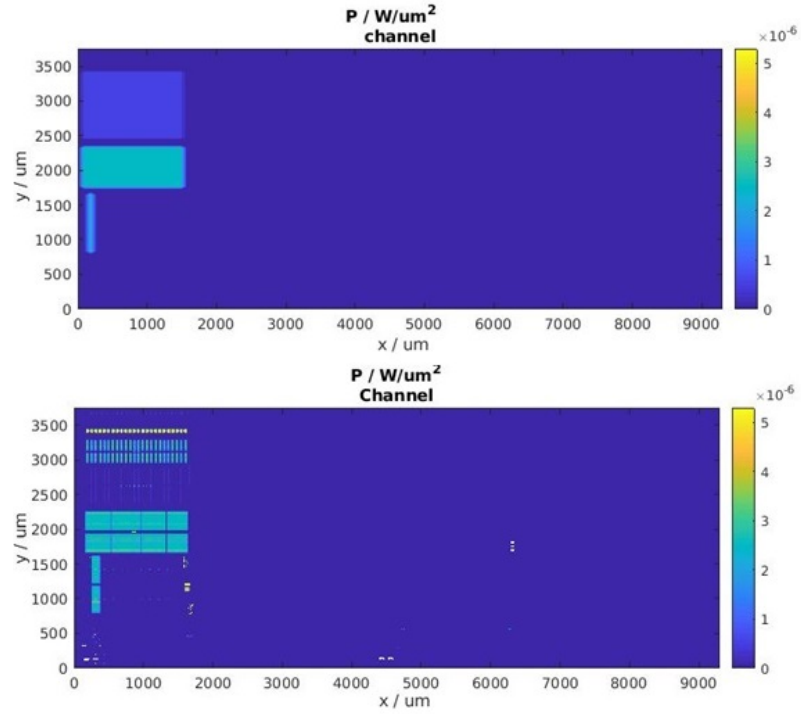

Figure 2.8: Top: power profile of a standard FEM simulation; Bottom: Legato tool power profile

2.3.2 Thermic Connection to the Environment

As described above the environment connection must be inputted via heat capacity C and thermal resistance R_{th}.

$$C = \frac{dQ}{dT} \tag{2.11}$$

$$R_{th} = \frac{\Delta T}{\dot{Q}} \tag{2.12}$$

C	heat capacity [$\frac{J}{K}$]	R_{th}	thermal resistance [$\frac{K}{W}$]
Q	heat energy [J]	$\dot{Q}$	heat transfer rate [W]
ΔT	temperature difference [K]		

For environment connection, three cases would be defined.

Heat Sink

In the case of a heat sink, the boundary temperature of the chip is equal to the ambient temperature, which is equal to $R_{th} = C_p = 0$.

Radiation and Convection

The second one is no package, connection with thermal radiation, and natural or forced convection. Here, on the one hand, the power dissipation via radiation is described with the Stefan-Boltzmann law.

$$\dot{Q} = \epsilon \, A \, \sigma \, (T^4 - T_a^4) \tag{2.13}$$

σ	Stefan-Boltzmann constant [$\frac{W}{m^4K^2}$]	ϵ	emissivity []
T	temperature [K]	T_a	ambient temperature [K]
A	surface area [m^2]		

The only variable here is the emissivity of the surface which describes the effectiveness of emitting but also absorbing radiation. It is a number between 0, perfect reflection or transmission of all incoming radiation, and 1, perfect absorption. The difficulty here is that this value depends not only on the material and material structure but also on the topology of the surface. There are three surface material used in this thesis:

Material		Reference
SiO_2 coated on Si:	$\epsilon = 0.88$ to 0.9	[11]
Fr-4:	$\epsilon = 0.8$ to 0.9	[12]
Copper oxide:	$\epsilon = 0.6$ to 0.7	[13]

On the other hand convective heat transfer can be described with

$$\dot{Q} = k \, A \, \Delta T \tag{2.14}$$

k heat transfer coefficient [$\frac{W}{m^2K}$]

In this formula, all dependencies are put into the transfer coefficient such as temperature difference, thermal proprieties of the fluid, boundary conditions e.g. room height, and the fluid velocity and angle of this velocity to the surface. The heat transfer coefficient is definition by related to the thermal resistance as,

$$R_{th,conv} = \frac{1}{A\ k} \tag{2.15}$$

Because of the many system dependencies, an accurate value can only be obtained for each case with computational fluid dynamics. There are also some more sophisticated empirical equations for specific cases [14] [15], but they are only valid for these defined conditions. Because of this, it was decided to take a more rudimentary approach to experimentally find relations and make a simulation with the lowest and highest heat transfer coefficient.

In principle, there are two types of convective heat transfers: Convective heat transport is called natural, where the motion of the fluid is only caused by buoyancy resulting from the temperature dependency on density variation of the fluid. Here the heat transfer coefficient can be described with the empirical formula.

$$k_{conv,nat} = C(\Delta T)^n \tag{2.16}$$

Here n is 0.25 and C depending on the direction between 1.319 and 3.158.[16]

On the other hand, if the external fluid flow is present, it is called forced convection. Here the heat transfer coefficient can be approximated for air linearly with air velocity, with a and b as constants. [17]

$$k_{conv,forced} = a + b\ v \tag{2.17}$$

v fluid velocity
with constant set $a_1 = 5.7$, $b_1 = 3.8$ or $a_2 = 2.8$ $b_2 = 3.0$.

The thermal resistances from radiation and convection add up like two electrical resistances. In figure 2.9 the equation of this chapter for the thermal resistances are visualised ($\frac{R_{th}}{A} = \frac{1}{k}$).

$$R_{ges} = \frac{1}{A\ k_{ges}} = \frac{1}{\frac{1}{R_{rad}} + \frac{1}{R_{conv}}} \tag{2.18}$$

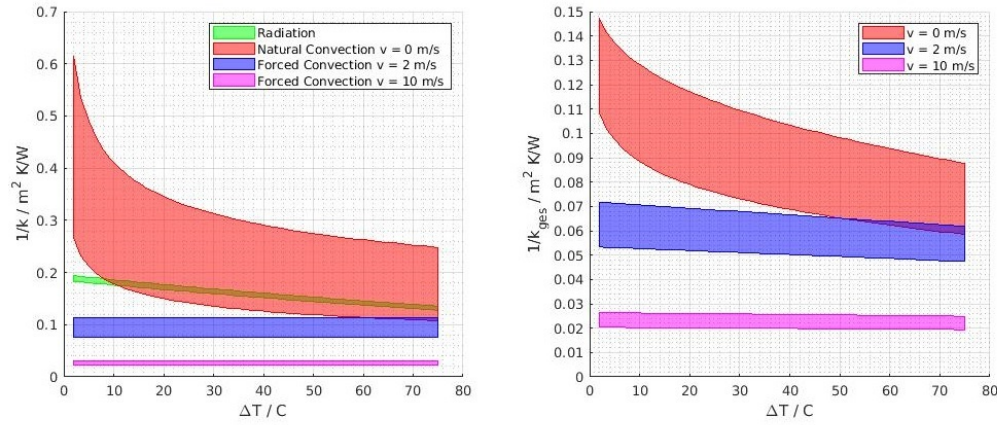

Figure 2.9: Left: $\frac{R_{th}}{A}$ for radiation and convection; Right: $\frac{R_{th,ges}}{A}$ region for different convection modes

2.3.3 Heat Equation

The inhomogeneous heat equation 2.19 describes the time development of the temperature in the presence of a heat source and is solved in a dynamical run of the tool.

$$\frac{\partial T(\vec{x},t)}{\partial t} - \frac{\lambda(\vec{x})}{c(\vec{x})\rho(\vec{x})}\vec{\nabla}^2 T(\vec{x},t) = \frac{\omega(\vec{x},t)}{c(\vec{x})\rho(\vec{x})} \tag{2.19}$$

λ	thermal conductivity $[\frac{\mathrm{W}}{\mathrm{m\,K}}]$	c	specific heat capacity $[\frac{\mathrm{J}}{\mathrm{kg\,K}}]$
ρ	density $[\frac{\mathrm{kg}}{\mathrm{m}^3}]$	$\dot{Q}$	heat transfer rate [W]
ω	heat source density $[\frac{\mathrm{J}}{\mathrm{s\,m}^3}]$		

Example Solution: 1D

One way to analytically solve the equation given in eq. 2.19 is via folding the Green's function of eq. 2.25 [18] for one dimension:

$$T(x,t) \;=\; \int_{-\infty}^{\infty} f(\zeta)G(x,\zeta,t)d\zeta \;+\; \int_0^t \int_{-\infty}^{\infty} \frac{\omega(\zeta,\tau)}{c\,\rho} G(x,\zeta,t-\tau)d\zeta dt \quad (2.20)$$

Where $f(x)$ is the initial temperature distribution at $t \;=\; 0$ and G is the Green's function. For a domain of x from $-\infty$ to ∞:

$$G(x,\zeta,t) \;=\; \frac{1}{2\sqrt{\pi a t}} \exp(-\frac{(x-\zeta)^2}{4at}) \quad (2.21)$$

where a is the thermal diffusivity $a \;=\; \frac{\lambda}{c\rho}$with a point power source of $\omega \;=\; \delta(x) \cdot C$ and a starting temperature profile of $f(x) \;=\; 0$ first part of 2.20 is zero and the rest can be solved with the result

$$T(x,t) \;=\; \frac{C}{2\lambda\sqrt{\pi}} \left[\sqrt{4at} \exp(-\frac{x^2}{4at}) \;-\; |x|\sqrt{\pi} \left(1 \;-\; \mathrm{erf}(\frac{|x|}{2\sqrt{at}})\right)\right] \quad (2.22)$$

For both therms, similar to the diffusion equation, a characteristic length and time scale can be defined. The diffusion length l_{diff} and characteristic time t_{diff}:

$$l_{diff} \;=\; 2\sqrt{at} \quad (2.23)$$

$$t_{diff} \;=\; \frac{\Delta x^2}{4a} \quad (2.24)$$

These two parameters can be used to describe the transition to the equilibrium. If it holds in a system for every point $t \;>>\; t_{diff}$ and $x \;<<\; l_{diff}$ a static solution can be assumed.

Example Solution: Equilibrium Solutions

In the case of $\frac{\partial T(\vec{x},t)}{\partial t} = 0$ the heat equation for the equilibrium conditions is obtained, which is a Poisson equation.

$$-\vec{\nabla}^2 T(\vec{x}) = \frac{\omega(\vec{x})}{\lambda(\vec{x})} \tag{2.25}$$

In the case of cylindrical and spherical coordinates the equation becomes:

cylindrical symmetry	spherical symmetry
$\frac{d}{dr}\left(r\frac{dT}{dr}\right) + \frac{r\omega}{\lambda} = 0$	$\frac{d}{dr}\left(r^2\frac{dT}{dr}\right) + \frac{r^2\omega}{\lambda} = 0$
with the solutions	
$T = -\frac{\omega r^2}{4k} + A\ln r + B$	$T = -\frac{\omega r^2}{6k} + \frac{A}{r} + B$

2.4 Simulation Results

As the experiment circuits first, a single p-MOSFET was chosen to introduce effects without the influence of the circuit. Next, the influence of different layouts on a chip will be shown with a current mirror. Finally the influence of a temperature profile on the operation of a circuit will be shown in the example of the MOS driver.

The layer stack used in this simulation can be seen in table 2.1

Table 2.1: Used layer stack

Layer name	layer thickness / μm	active material	passive material
180 nm technology layer stack	7.36	silicon	silicon oxide
Channel	0.33	silicon	
Substrate	350	silicon	
Glue	100	glue	
PCB	500	TR4	
Glue	50	glue	
Heatsink	1500	aluminum	

The thermal conductivity λ and the volumetric heat capacity $\rho \cdot c_p$ at room temperature for the materials listed in table 2.1 can be seen in 2.2.

Table 2.2: Thermal constants used for the simulation

Material	λ / $\frac{\mathrm{W}}{\mathrm{m\,K}}$	$\rho \cdot c_p$ / $\frac{\mathrm{J}}{\mathrm{m^3K}}$
silicon	148	1658960
silicon oxide	12.4	2440000
glue	1.6	1658960
TR4	30	2090000
aluminum	204	2434498

2.4.1 Electro Thermal Simulation of a Point Source with the 180 nm Technology Layer Stack

In order to judge if a dynamical electro-thermal simulation to equilibrium can be achieved, the power input of the MOS driver at $D = 2\,\%$ of $P = 1.94\,W$ was taken and placed in the middle of a chip at the same size as the MOS driver (9300 μm x 3750 μm). Then a dynamical simulation was made with a 101x101 grid. The associated equilibrium profile of the discussion can be seen in figure 2.10.

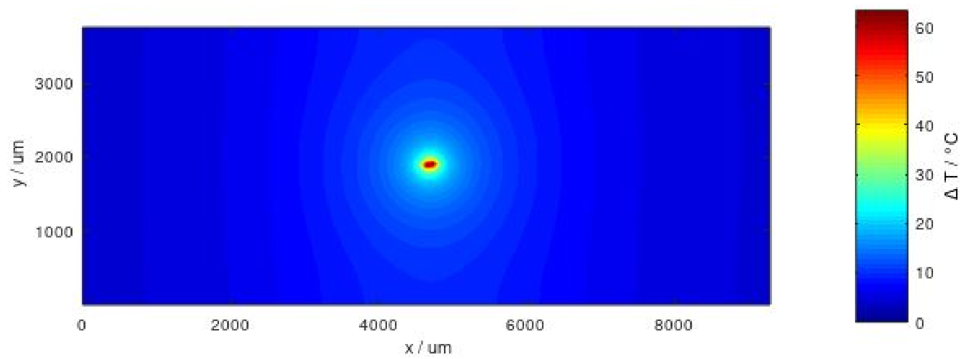

Figure 2.10: Equilibrium temperature distribution of a point source with $P = 1.94\,W$

In figure 2.11 the temperature development with time can be seen for selected grid points. The middle point, the adjacent points left, right, and diagonal of

it, and the border points were taken. It can be seen that each element initially has a heating-up phase, which then turns into a process where the heating rate of the elements is roughly the same. This can be seen in figure 2.12 where a temperature profile at y=1875 μ at a time of 30 ms and equilibrium can be seen. After this 40 ms, the chip warms up at an almost constant rate until the equilibrium is reached at roughly 700 ms. In comparison the characteristic time t_{diff} defined in eq. 2.24 with a thermal diffusivity $a = 86 \frac{\text{mm}^2}{\text{s}}$ and the furthest away from the power source is $t_{diff} = 62$ ms , which is twice as high as the measured 30 ms for the heat to reach the same border. The equilibrium solutions for radial and spherical symmetry are plotted in the figure for comparison.

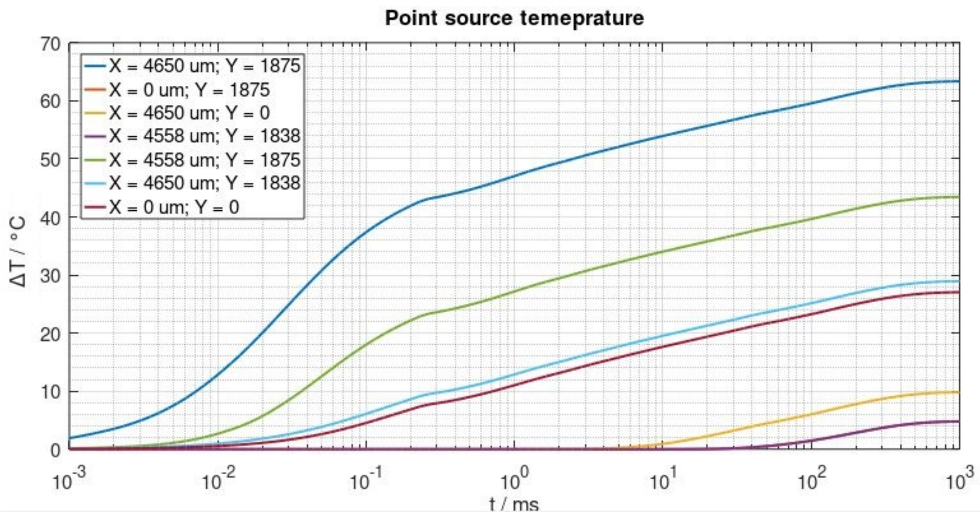

Figure 2.11: Temperature development of different elements/points with time

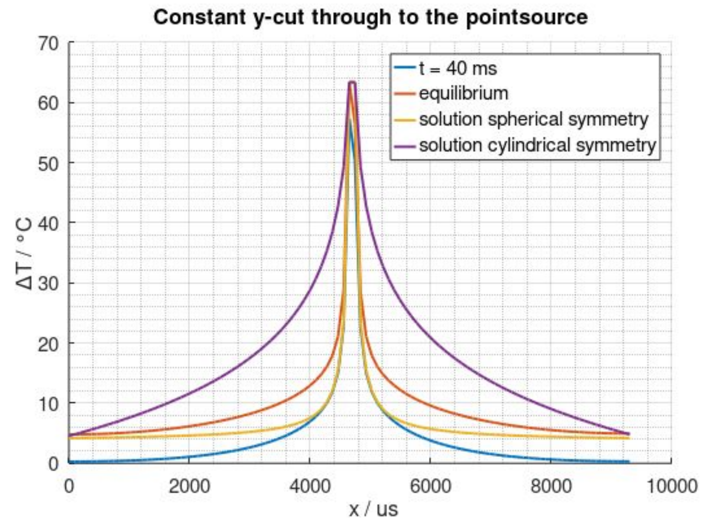

Figure 2.12: Tempeature profile at $y = 1875\ \mu$m for 40 ms at equilibrium

2.4.2 Temperature Behaviour of a Single p-MOSFET

To analyse the behaviour of other test circuits, a temperature dependent simulation of the main transistor (29V HV pMOS) with one finger was made. The results of this can be seen in figure 2.13.

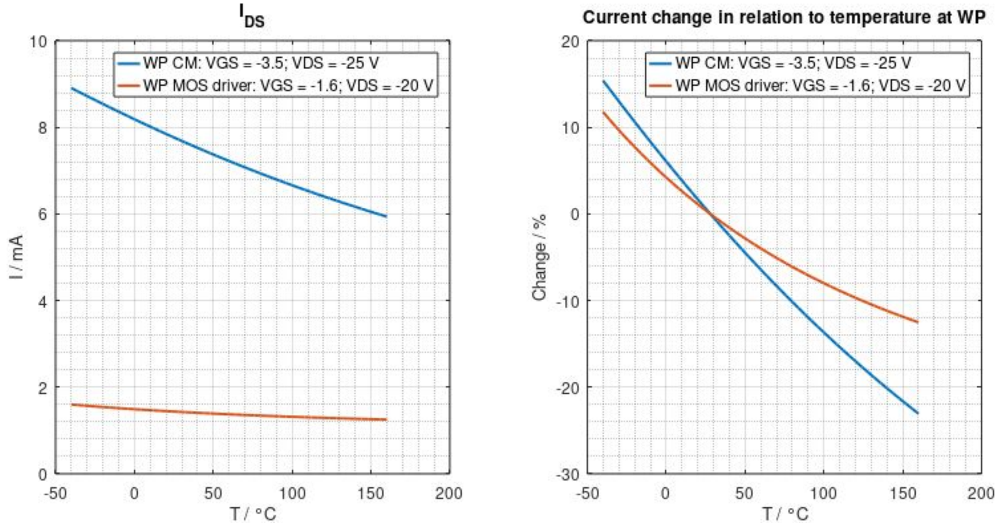

Figure 2.13: Left: drain-source current for the operating points of the MOS driver and the current mirror; Right: Perceptual change of the drain-source current with the current at 25 °C as a baseline

As expected, the current is falling with temperature and the relative change is proportional to the gate voltage. With this also the behaviour of a multi finger transistor can be calculated which is:

$$I_{DS,ges} = \sum_{i=0}^{nf} I_{DS,i}(T_i) \tag{2.26}$$

2.4.3 Electro-Thermal Simulation of a Current Mirror

The current mirror was simulated with an input pulse of 10 mA and a duty cycle from 10 %. The supply voltage was chosen with -25 V. The difference between the two test layouts is only the swapped modules that means that the temperature profile is identical and can be seen in figure 2.14.

For the electrical simulation, the tool takes the mean temperature of the position of the associated transistor. The temperature taken for the simulation and the resulting current can be seen in figure 2.15.

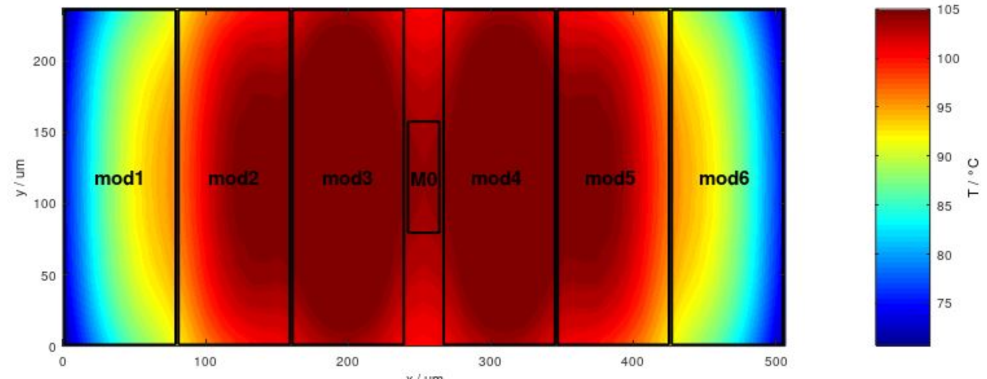

Figure 2.14: Temperature profile of the current mirror

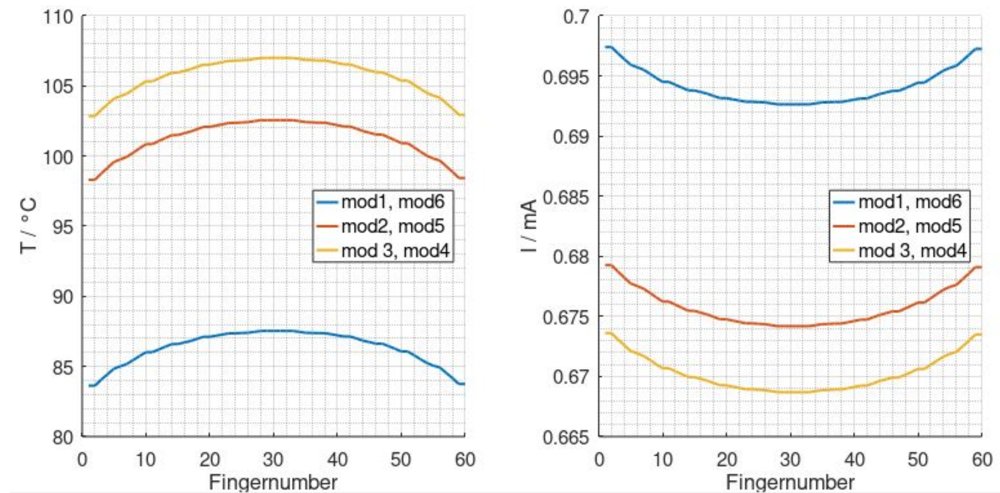

Figure 2.15: Left: mean finger temperature; right: finger current

The division between the two output transistors can be seen in table 2.3. The resulting output currents for the two layouts are listed in table 2.4.

For comparison, the circuit was also simulated with different uniform temperatures ($T(x,y,z) = const$). The resulting currents can be seen in figure 2.16.

Table 2.3: Layout variants of the current mirror

Layout	**M1**	**M2**
1 (symetric)	mod2, mod5	mod1, mod3, mod 4, mod6
2 (staggerad)	mod1, mod2	mod3, mod4, mod 5, mod6

Table 2.4: Results for the symmetric layout

	T_{umg} = **25 °C**	T_{umg} = **-40 °C**
$IDout1$ (symmetric Layout)	-0.828	-0.720
$IDout2$ (symmetric Layout)	-1.636	-1.454
$IDout1$ (staggered Layout)	-0.810	-0.738
$IDout2$ (staggered Layout)	-1.635	-1.457

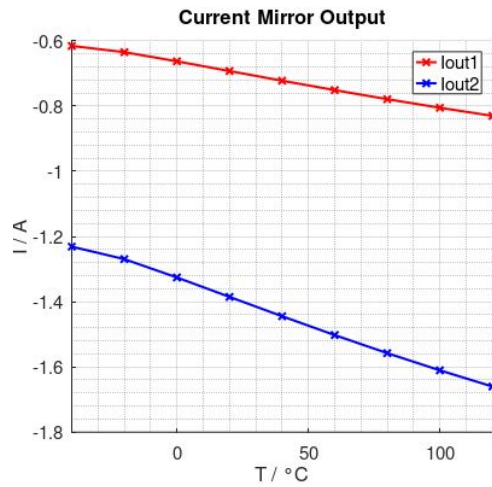

Figure 2.16: Output current for different uniform temperatures

2.4.4 Influence of Simulation Variables for the MOS Driver Simulation

For the influence of the simulation variables, the three most important results are listed below. The thermal resistance has only limited influence on the temperature, as it can be seen in figure 2.17. This indicates that most of the heat dissipation of the chip is done through the heat sink at the bottom.

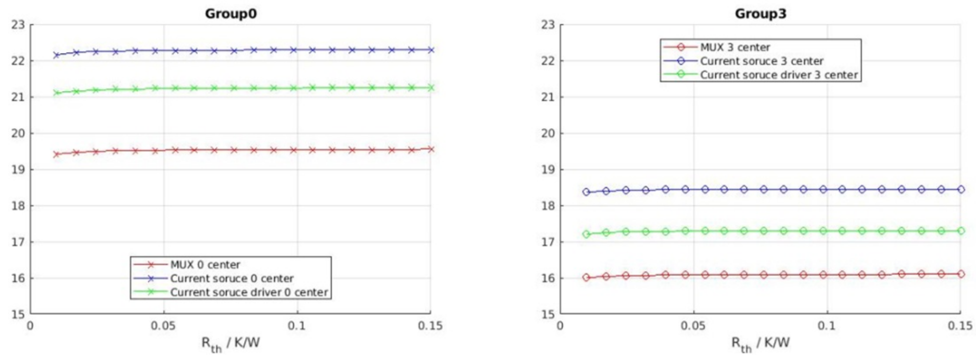

Figure 2.17: Temperature dependence of the thermal resistance for Grp0 and Grp3

The glue thickness shows a linear dependence on the temperature.

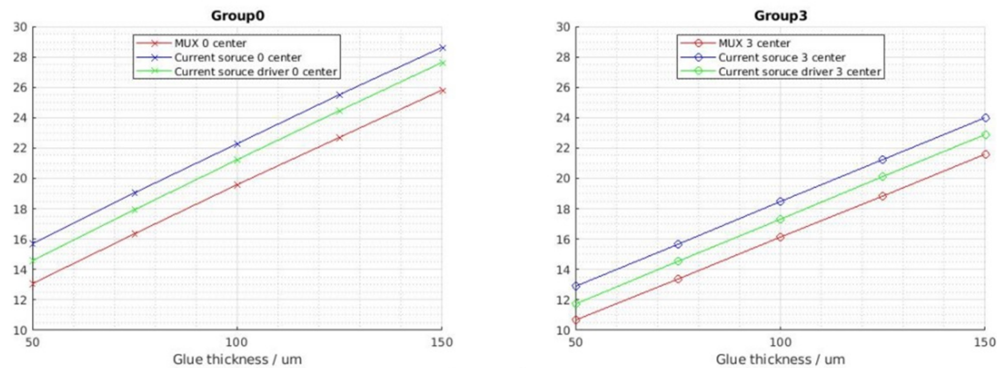

Figure 2.18: Temperature dependence of the thickness of the glue layer of Grp0 and Grp3

The substrate causes dependent on the thickness two effects, a heat spread which causes a temperature decrease, and thermal resistance to the heat sink, which causes a temperature increase. If these two effects for some substrate thickness, a minimum in temperature occurs.

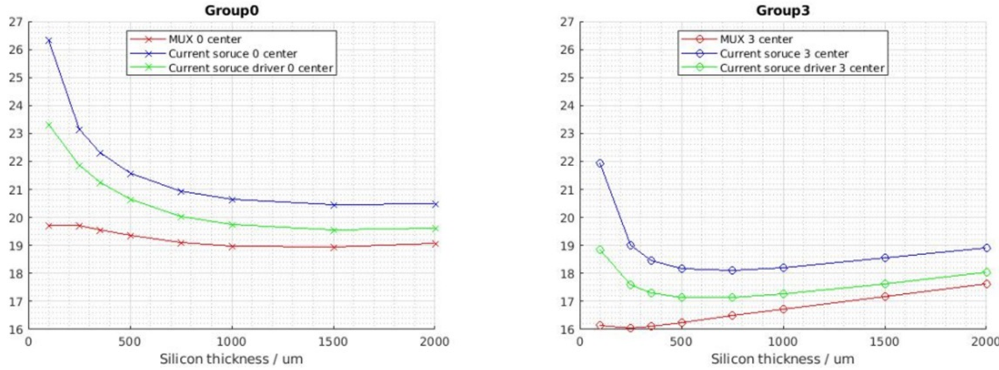

Figure 2.19: Temperature dependence of the thickness of the substrate of Grp0 and Grp3

2.4.5 Electro-Thermal Simulation of a MOS driver

For the simulation of the MOS driver an inverse heat transfer coefficient of $\frac{1}{k} = 0.01\ \frac{\mathrm{m}^2\ \mathrm{K}}{\mathrm{W}}$ was taken. The boundary temperature for the bottom side of the heat sink was set at room temperature. Then first temperature profile depended on the duty cycle was simulated at an environment temperature of $T = 29$ °C and $T = -40$ °C. For a duty cycle of $D = 10$ % the electrical output of this electro-thermal simulation can then be compared to the simulation of uniform temperature.

Thermal Results of the MOS Driver

The result of electro-thermal simulation for different duty cycles can be seen in figure 2.20. For lower duty cycles, a hotspot at $x = 1580$ μm and $y = 1200$ μm was identified. The maximum temperature is linear dependent on the duty cycle.

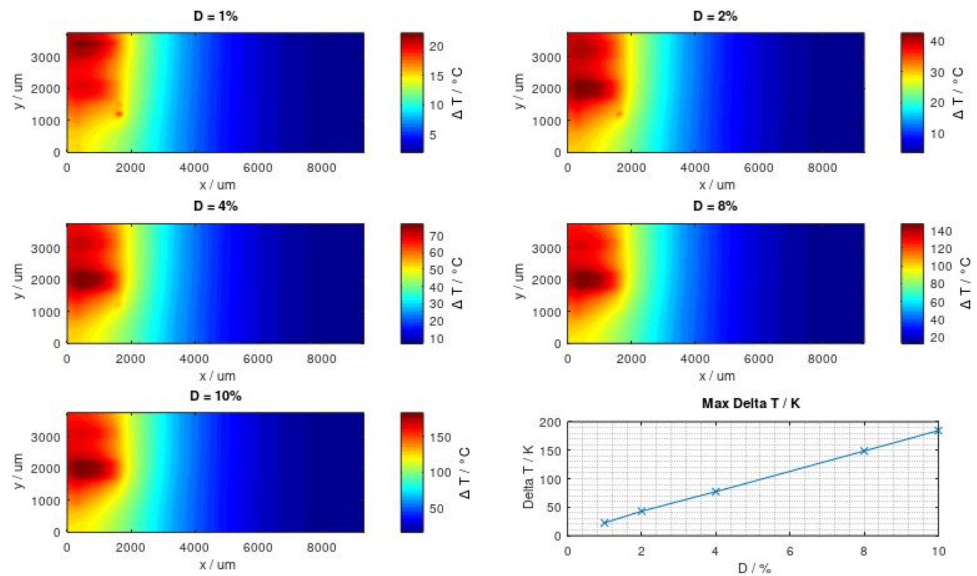

Figure 2.20: Temperature distribution for different duty cycles

Electrical Results of the MOS Driver

For the electro-thermal result, the pulse at $D = 10$ % is compared against the simulation with different uniform temperatures. The overall pulse can be seen in figure 2.21. Also, five different parameters were compared: the maximum current, the settling current at the end of the pulse, the on-time and off-time, which is defined when the current falls below -4 A, and the pulse length, which is the difference between the two.

The two current parameters can be seen in figure 2.22. The settling current has its minimum at 100°C and falls at higher temperatures, while the maximum current decreases with rising temperature but has a plateau at 150 °C.

The time characteristics of the pulse are depicted in figure 2.23. Both the start and end time of the pulse is linear with the temperature while the pulse length varies by 0.1 ns.

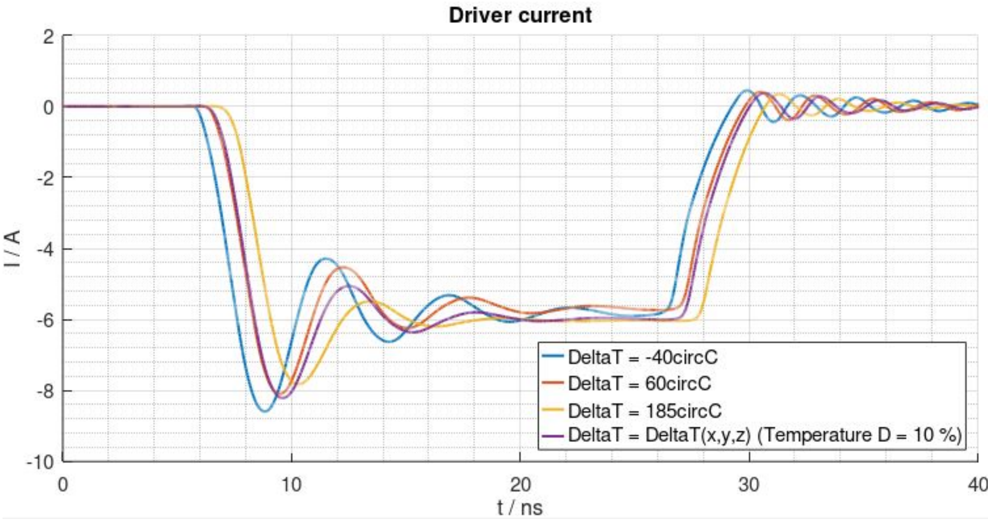

Figure 2.21: MOS current for different uniform temperatures and for the associated temperature profile for $T_{umg} = 29$ and $D = 10$ %

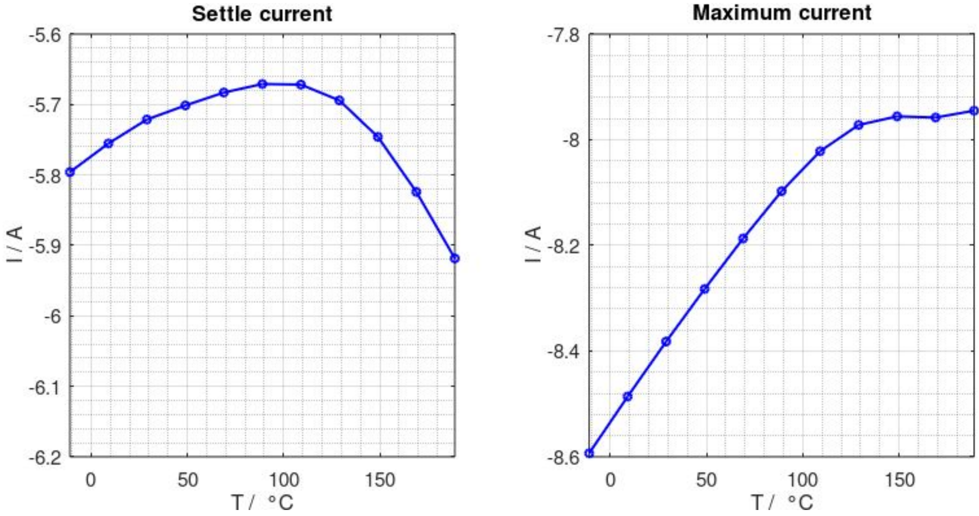

Figure 2.22: Left: settling current with uniform temperature; Right: maximum current with uniform temperature

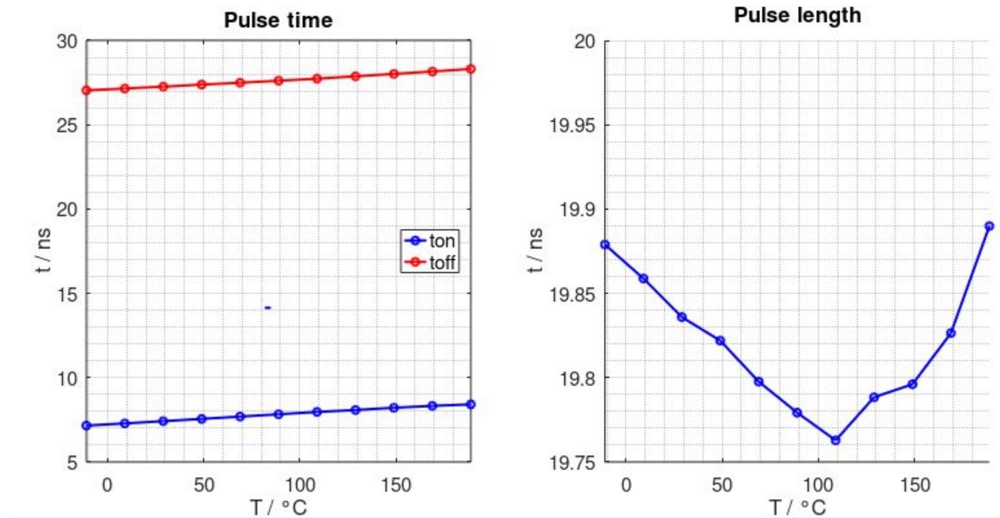

Figure 2.23: Left: pulse times, switch on and switch off; Right: pulse length

The result of the electrical characteristics of the electro-thermal simulation can be seen in figure 2.5

Table 2.5: Electro thermal simulation result

	$-I_{max}$ / A	$-I_{settling}$ / A	t_{start} / ns	t_{end} / ns	t_{PL} / ns
Grp0, T_{umg} = -40 °C	8.09	5.76	7.22	27.01	19.80
Grp0, T_{umg} = 29 °C	8.22	5.98	7.87	27.78	19.92
Grp0, T_{umg} = -40 °C	8.31	5.78	6.95	26.79	19.84
Grp0, T_{umg} = 29 °C	8.13	5.74	7.70	27.49	19.79
Grp3, T_{umg} = -40 °C	8.61	5.87	6.68	26.58	19.98
Grp3, T_{umg} = 29 °C	8.38	5.77	7.42	27.25	19.83

Comparison between Groups

While the electrical construction of the groups of the MOS driver is the same, the difference in power dissipation because of the different positions causes them to behave differently. The different temperature profile for ten % duty cycle is shown in figure 2.24 and the resulting pulse is shown in figure 2.25.

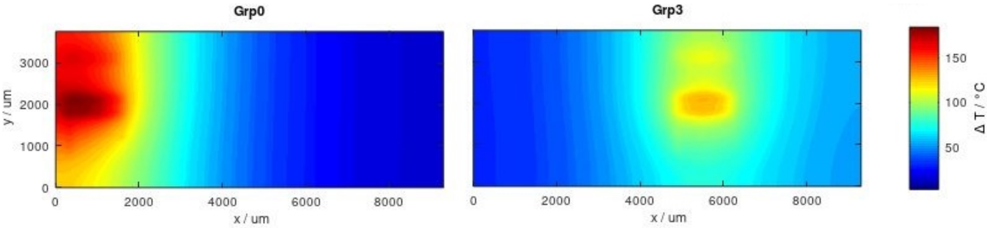

Figure 2.24: Temperature distribution if Grp0 or Grp3 is used with D = 10 %

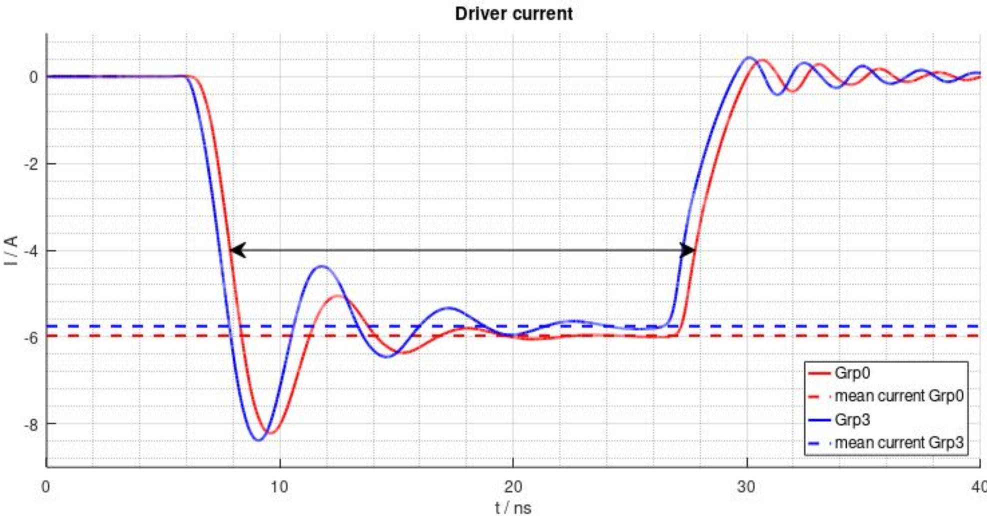

Figure 2.25: Electro thermal simulation result of the MOS current for Grp0 and Grp3

The characteristics resulting from this are listed in table 2.6.

Table 2.6: Comparisson between Grp0 and Grp3

	Grp0	**Grp3**	**Grp0 - Grp3**
T_{max} / °C	184	130	54
$T_{mean,Grp}$ / °C	120	97	23
P / W	14.33	13.99	0.34
$-I_{max}$ / A	8.21	8.38	-0.17
$-I_{settling}$ / A	5.97	5.77	0.20
t_{start} / ns	19.91	19.83	0.08
t_{PL} / ns	7.87	7.42	0.45

Here it can be seen that even if electrically identical circuit parts are used their behaviour changes dependent on the layout around of them.

3 Ageing Simulation

3.1 Reliability

3.1.1 Reliability Basics

Reliability [19] [20] is the quality of a device to have defined properties for a certain amount of time until it has inevitable degrades and it ends its lifetime. This degradation is a statistical process because of the slight process variation in the production of devices, a difference in stress, and the probabilistic nature of physical processes responsible for degradation.

In general, the equation which describes survivors after a specific time is given by:

$$\frac{dM}{dt} = -\lambda(t)M(t) \tag{3.1}$$

M(t) describes the number of survivors at the specific time t and $\lambda(t)$ is the rate of failure to this particular time.

Typically this failure rate has the form of a bathtub curve shown in figure 3.1. In the figure, three regions can be seen. The failure rate is high at the beginning because not all of the significant production errors have been found, however it falls while this happens. Then there is a long time with a constant small failure rate. This is the area where devices degrade slowly, which is an interesting area for this thesis. At the rapid rise of failure rate, in the end, is the wear out and marks the end of the lifetime of a device.

In addition to the failure rate also a degradation rate k can be defined which describes how fast the physical process of degradation is progressing. This is

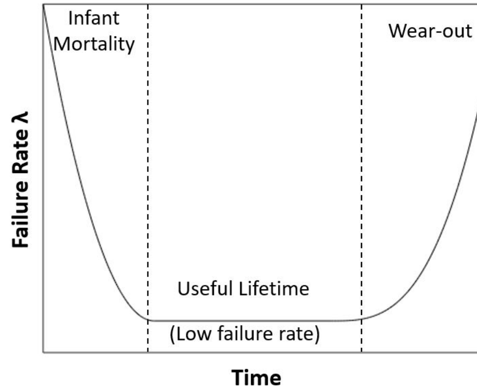

Figure 3.1: Failure rate over time, the typical bathtub curve

related to the time to failure (TF) in the way that failure is reached when a critical parameter p_{krit} is exceeded.

$$p_{krit} = p_0 \cdot e^{-\int_0^{TF} k(t)dt} \tag{3.2}$$

With this, a TF for a device can be calculated, which is the time until p_{krit} is reached.

The standard reliability time frames are often around ten years, therefore the constants for reliability models can not be obtained from regular measurements following has to be done. Extrapolation from shorter time frames under the assumption that the degradation mechanism does not change. And accelerated testing, a greater than usual stress is applied so that the devices age faster. The result is then calculated back to normal stresses. This works under the assumption that the ageing mechanism does not change with stress.

3.2 Degradation Mechanism of MOSFET

The shift of the characteristic curves characterises the reliability performance of a MOSFET. Different stresses on mechanisms lead to the shifting of each device until the circuit can no longer perform its task. In the simulation of this

thesis, two of these degradation mechanisms will be considered. For both of them, Si-H bounds at the Si-SiO_2 interface play an important role. These bonds exist because, in modern CMOS technology, the material for the gate dielectric is almost always amorphous silicon dioxide. This amorphousness results in dangling bonds at the interface, which, if left alone, would be electrically active and capture the charge. To prevent this, hydrogen is used to bind these dangling bonds.

3.2.1 Hot Carrier Injection (HCI)

Hot carriers emerge in a MOSFET trough the acceleration of carriers by the drain-source voltage. These hot electrons can produce through impact ionization an electron whose momentum is directed to the dielectric (figure 3.2). This is non-uniform because the electrons first need the build up the required kinetic energy.

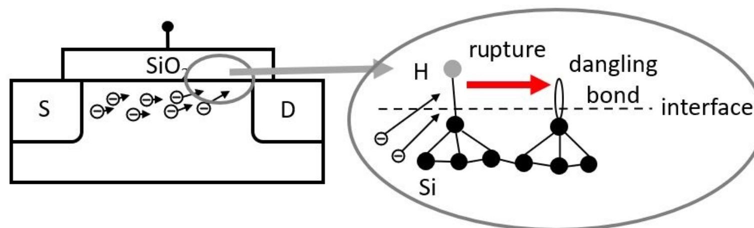

Figure 3.2: Schematic of the HCI [21]

There are three different damaging modes of hot carriers on MOSFET:

- interface trap creation: The dissociation energy of the Si-H bond at the interface is as low as 2eV. [22] If such a bond is broken, it can act as a trap for electrons.
- hot electron injection: If the momentum of an electron is directed to the bulk oxide and it can overcome the potential barrier of 3.1 eV, it can be injected into the oxide.
- hole injection: This is the same as electron injection; potential barrier 4.8 eV

The occurrence of these modes dependents on the type of MOSFET (n,p) and the voltage between Gate and Source (V_G), and they influence the MOSFET characteristics due to their electrostatic effects. [23][24]

There are three different ways to generate an interface trap: First, the electron has the required energy to break the Si-H bond. This is the governing mechanism if the gate-drain voltage is higher than the needed 2 V. Second, an electron can obtain the necessary energy through exchange mechanisms (electron-electron scattering). Third subsequent collisions of lower energy electrons can excitation to bond until it breaks. These two mechanisms are more relevant for $V_D < 2V$. The interplay between these two cases gives a dependence on channel length. The HCI is related to the maximum average energy for longer channel devices, and for short channel length, the carrier flux is more important. A consequence of this is a different temperature dependent on the degradation because of the HCI effect. For devices with higher V_D and longer channels, the degradation decreases with increased temperature, while for devices V_D and short channels, the degradation becomes more severe. [22]

3.2.2 Negative Bias Temperature Instability (NBTI)

NBTI, first discovered in 1966 [25], is a fundamental degradation mechanism in MOSFET devices. The effect occurs because of a build-up of positive charges in the Si-SiO_2 interface and the Sio_2 dielectric when a negative gate bias stress is applied. This results in a shift of device parameters like the threshold voltage V_{th} or the saturated source-drain current I_{Dsat} with time and is enhanced with higher temperatures. [26] [27] Another aspect of NBTI is that some of this shift will recover if the stress is turned off. While the mechanism can be observed in both p- and n-channel devices, it is much more prominent in p-channel MOSFETs [28][26], for this reason NBTI is mainly considered an issue of p-channel MOSFETs.

At present, no consensus for a theory of the underlying physical mechanisms was found. The main focus of this thesis is not NBTI modelling, so the main approach will be introduced shortly below.

Reaction Diffusion Model

The reaction-diffusion model approaches NBTI as the overlap of two processes. At the interface, the Si-H bond can be broken through an electrochemical reaction with an electric field which results in a dangling bond and a mobile hydrogen atom. The hydrogen atom is then transported into the dielectric through diffusion and the dangling bond can catch a hole which leads to a positively charged interface. If the electric field is turned off the concentration gradient of hydrogen leads to a back transport of the hydrogen atoms through diffusion. At the interface they can recombine with a dangling bound while emitting a hole, which explains the recovery.[27] An illustration of this can be seen in figure 3.3.

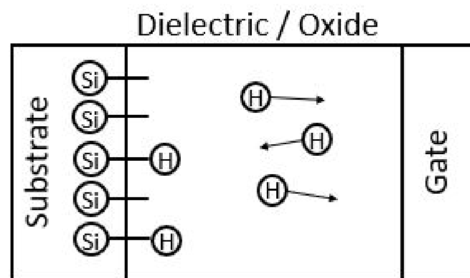

Figure 3.3: Schematic of the RD model of NBTI

3.3 Reliabilty Modeling

3.3.1 Takeda Reliability Model

In 1983 E. Takeda and N. Suzuki published "An Empirical Model for Device Degradation due to Hot Carrier Injection". While devices scaled down and materials and processes change, this model is still used as an industry standard. The reason for this is the easy set up process, the relative accuracy, and the computational inexpensiveness.

They found the relationship between the degradation of a parameter eg. the threshold voltage (V_{th}) can be expressed with a simple power function:

$$\Delta Par = A \cdot t^n \tag{3.3}$$

ΔPar	change design parameter
A	model constant for scaling
t	aging time [s]
n	model constant for exponent

And while n is nearly independence on V_G which suggest that n changes according of the hot carrier mechanism, for A a strong dependency on V_D was found:

$$A \propto e^{\frac{-\alpha}{V_D}} \tag{3.4}$$

VD	Drain-source voltage [V]	α	constant [V]

This is the same dependence as the peak substrate current, which can be taken as a measure for the number of electron-hole pairs generated. From it can be deduced that A is the number of excess carriers due to impact ionization. This model makes it possible to predict long therm degradation with short therm measurements by fitting the model and extrapolating.

3.3.2 HCI Model

Based on the Takeda model the quality department at ams AG in Prämstätten has implemented their own empirical model:

$$\Delta Par = \Delta_0 \cdot \left(\frac{t}{t_{ref}}\right)^n \tag{3.5}$$

ΔPar	change design parameter
Δ_0	modelling constant for scaling the model[]
t	stress time [s]
t_{ref}	scaling factor for including temperature, geometry and stress voltage [s]

The scaling factor t_{ref} is a variation of the voltage dependency of the Takeda model, an Arrhenius equation for the temperature dependency, and an experimental relation for the transistor length.

$$t_{ref} = A \cdot e^{\frac{B}{V_D}} \cdot L^C \cdot e^{-\frac{E_a}{k_b T}} \qquad (3.6)$$

k_B Boltzmann constant [$\frac{\mathrm{eV}}{\mathrm{K}}$]
A model constant for scaling all dependencies [s]
V_D DC Drain-source voltage [V]
B model constant for voltage scaling [V]
L transistor length [μm]
C model constant for lenght scaling []
T transistor temperature [K]
Ea model constant for temperature scaling [eV]

3.3.3 NBTI model

For NBTI a power model can be used [29]. At ams AG Premstätten, this is combined with an Arrhenius model for the temperature dependence and an exponential model for voltage dependence.

$$\Delta V_{th0} = A \cdot e^{k_V V_G} \cdot e^{-\frac{E_a}{k_b T}} \cdot t^n \qquad (3.7)$$

k_B Boltzmann constant [$\frac{\mathrm{eV}}{\mathrm{K}}$]
A model constant for scaling all dependencies [s]
V_G DC Gate-source voltage [V]
k_V model constant for voltage scaling [$\frac{1}{\mathrm{V}}$]
T transistor temperature [K]
Ea model constant for temperature scaling [eV]

3.3.4 Implementation

No HCI measurement data was available for the used technology. In order to test the workflow, ams AG decided to use an HCI model used for similar technology. For this technology, the degradation of $GMmax$, currents in the linear regime $IDlin$ and saturated regime $IDsat$, the threshold voltage in the linear regime $VTlin$, and the saturated regime $VTsat$ were available. For NBTI,

it is amused that it is fully described with a threshold voltage shift in the linear regime $VTlin$. Here the model from the original technology is used. The disadvantage of this method is that the models for the design parameter are entirely independent of one another, which is not the case in reality. When the threshold voltage is shifted, all other parameters are changed with it. In order to use this degradation model in SPICE simulations, the change in design parameters must be translated into parameters of the SPICE model BSIM4 [9], which is the transistor model. This is essentially a multidimensional fit. The following three parameters were used for this:

- **VTH0**:
  Long-channel treshhold voltage, give the treshhold voltage without short-channel effects; change for treshhold voltage
- **VSAT**:
  Velocity Saturation; changes IDsat
- **ua**:
  Coefficient of first-order mobility degradation due to vertical field

3.3.5 AC Stress and the BERT Algorythm

The previously introduced models are only valid to test constant stress, therefore a formalism must be introduced to map any time dependent stress on a time independent one. [29] [30] With constant stress the parameter change due to HCI and NBTI are modeled with a power law:

$$\Delta Par(t) \; = \; C \cdot (f(\zeta_i) \cdot t)^n \tag{3.8}$$

C	model parameter
n	exponential model parameter
ζ_i	arbitrary stress
ΔPar	model parameter change

But when stress is time dependent a quasi-static equivalent can be defined.

$$\Delta Par(t) \; = \; C \cdot [\int_0^t f(\zeta_i(\hat{t}))d\hat{t}]^n \tag{3.9}$$

A variable can be defined, which includes all the stress parameters in an integral.

$$AGE = \int_0^t f(\zeta_i(\hat{t}))d\hat{t} \tag{3.10}$$

and because of the integral and if $f(\zeta_i(\hat{t}))$ is periodic, following relationship must be also valid:

$$AGE(N \cdot t) = N \cdot AGE(t) \tag{3.11}$$

This allows degradation calculation based on simulation in a much shorter timeframe. The rate of aging is, per definition, just the integrant.

$$AGErate = f(\zeta_i(\hat{t})) \tag{3.12}$$

The workflow with this reliability simulation can be seen in figure 3.4. The first simulation, must be a transient one which is putting the stress on the devices. During this simulation, the age rate is integrated over the period-time, which gives the AGE of one period time. The age over the use time (eq. ten years) is calculated via multiplication from this AGE. The device degradation can then be calculated. These degradation values are then stored somewhere to produce an aged netlist that can then be tested.

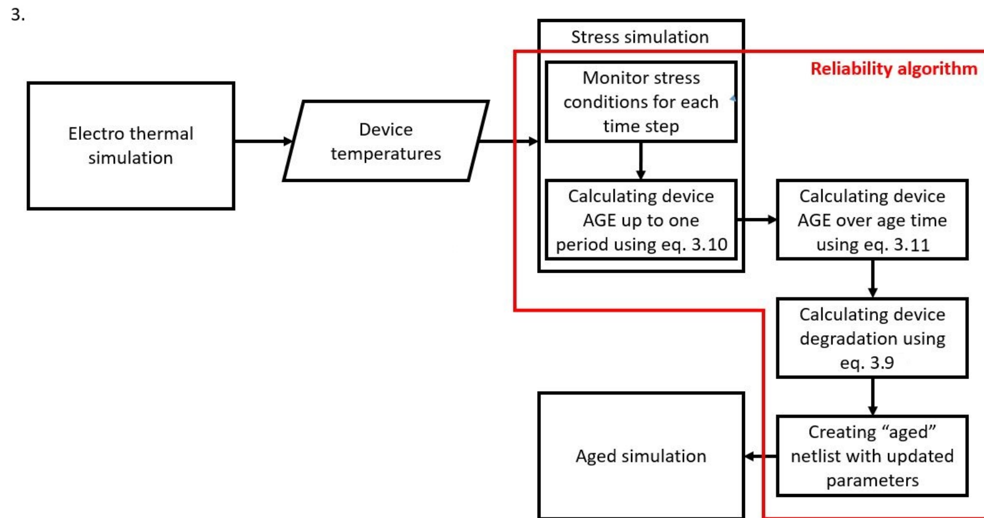

Figure 3.4: Reliability simulation workflow

3.4 Simulation Results

The model and algorithms explained above were applied to the test circuits. The only available degradation model for the 180nm technology were NBTI models for the PMOS because, as it will be shown below. It was decided together with ams AG to implement a model from another technology for the main transistor of the MOS driver. Degradation of NMOS, and the other hand, was not taken into account as the implementation of multiple degradation models is time-consuming and goes beyond the scope of the thesis. This means, therefore, that all of the following investigations are only qualitative.

An age time of ten years was chosen because this is a typical product lifetime. For the two test circuits, a temperature scan was undertaken, which later serves as a comparison for the ageing with temperature profile.

3.4.1 29V HV pMOS Reliability Simulation Results

The single transistor test circuit is used to gauge the variations of the working points of the test circuits. Because of the relatively uniform temperature profile, a comparison with the temperature profile dependent on aging did not show any significant changes. However, the circuit can still be used to visualize the expected change of the working points for the other two test circuits.

For NBTI aging, an aging model that only shifts the threshold voltage in the linear regime is used, resulting in a minor shift in the current. This can be seen in figure 3.5. On the left, the temperature dependence of the working point for both the MOS driver current source transistor and the current mirror test circuit is pictured. On the right side, the change of current due to NBTI aging for ten years can be seen. Shifts range from 0.05 % up to 0.45 %, which means if one is interested in current output, the NBTI ageing only plays a minor role.

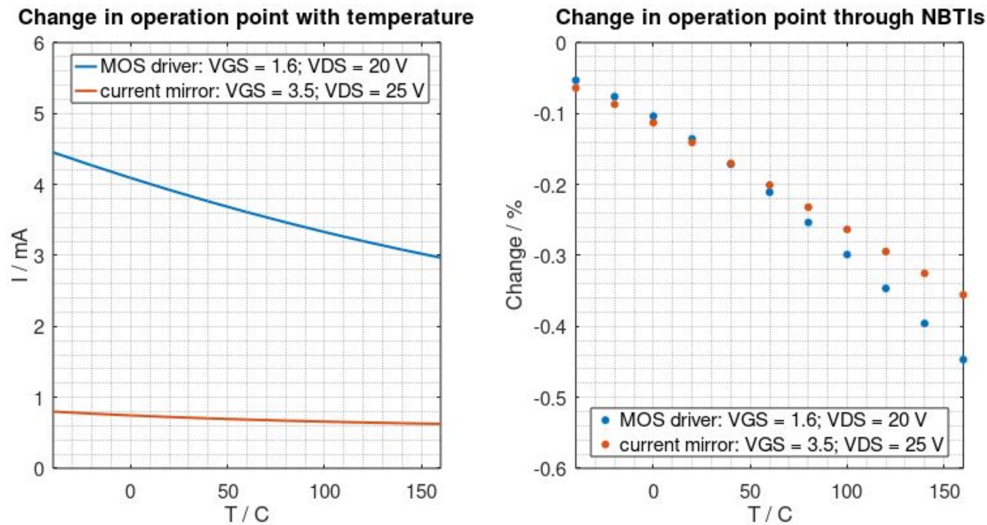

Figure 3.5: Left: Change with operation point with temperature; Right: Change of Operation point due to NBTI ageing for an age time of 1 year

3.4.2 Current Mirror Reliability Simulation Results

In the reliability simulation of the current mirror, no sizeable difference in waveforms, except the pulse height, could be found. The aging with uniform temperature is used as a benchmark and is shown in figure 3.6.The two transistors on the output side experience the same stress conditions, therefore they age at the same rate, which results in an overlap of curves if the percentual change is considered. This changes if the temperature profile is considered.

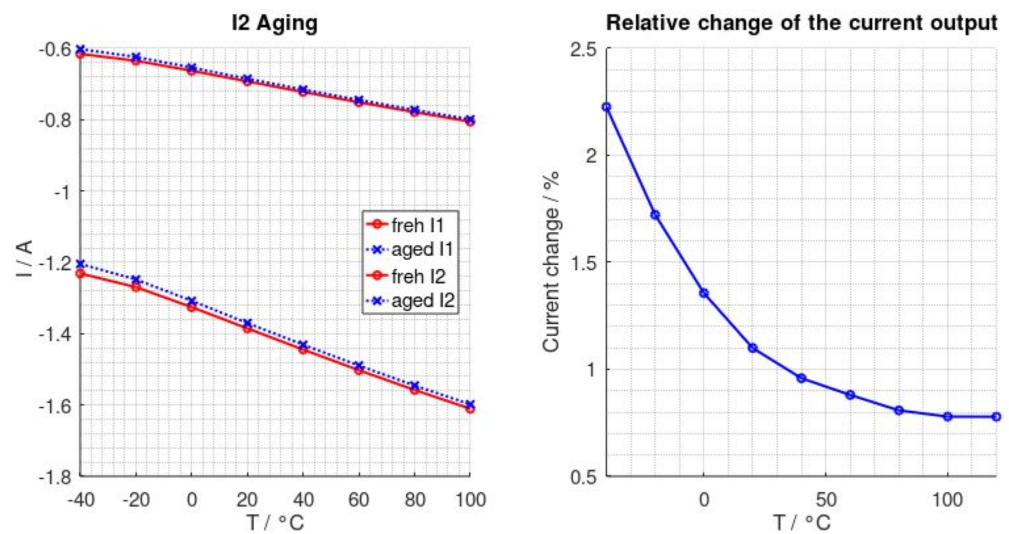

Figure 3.6: Change in output of the current mirror test circuit for age time of 10 years and uniform temperature; Left: current mirror outputs fresh and aged; Right: percentual change, equal for both outputs because of uniform temperature

With the inclusion of the temperature profile and an environment temperature of $T_{env1} = 25°\text{C}$ and $T_{env2} = -40°\text{C}$ the following values for the pulse current were obtained:

Table 3.1: Current mirror aging with temperature profile results for T = 25 °C

Layout	$I1$ fresh / A	$I2$ fresh / A	$\Delta I1$ / %	$\Delta I2$ / %
Common Centroid Layout	-0.84	-1.64	0.781	0.778
Row-Layout	-0.81	-1.64	0.778	0.781

Table 3.2: Current mirror aging with temperature profile results for T = -40 °C

Layout	$I1$ fresh / A	$I2$ fresh / A	$\Delta I1$ / %	$\Delta I2$ / %
Common centroid layout	-0.72	-1.44	1.02	1.04
Row-Layout	-0.74	-1.46	0.97	0.98

While the influence of the layout on the fresh simulation is visible, it seems that ageing did not to influence this further.

3.4.3 MOS Driver Reliability Simulation Results

The comparison between the implementation of the electro-thermal aware reliability workflow simulation with uniform temperature were done, the result of this investigation can be seen in data in figure 3.7 and 3.8. Because of the high temperature, the aging influence on the current characteristics correlates with the increasing in temperature. The impact on the start and stop time of the pulse is unexpected for the current change and the resulting change in the RC network.

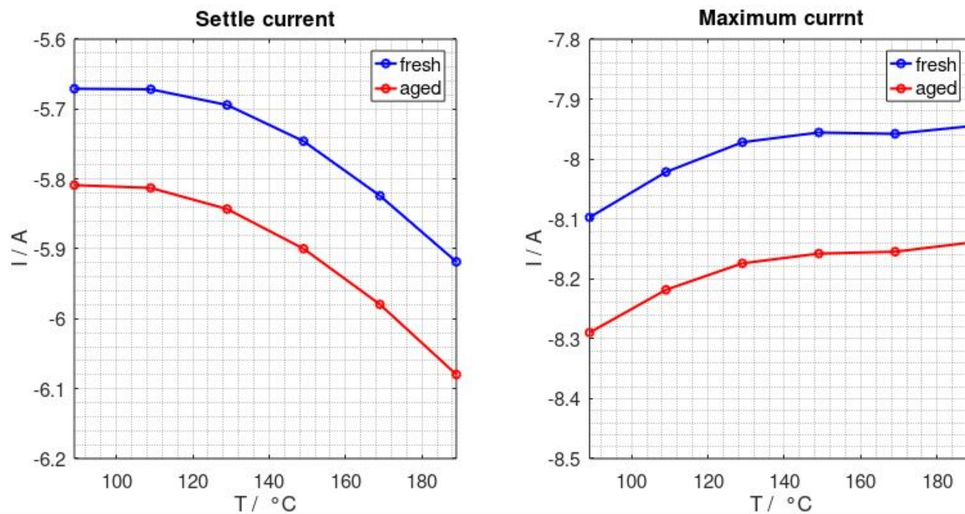

Figure 3.7: Fresh vs aged results; Left: settle current with uniform temperature; Right: maximum current with uniform temperature

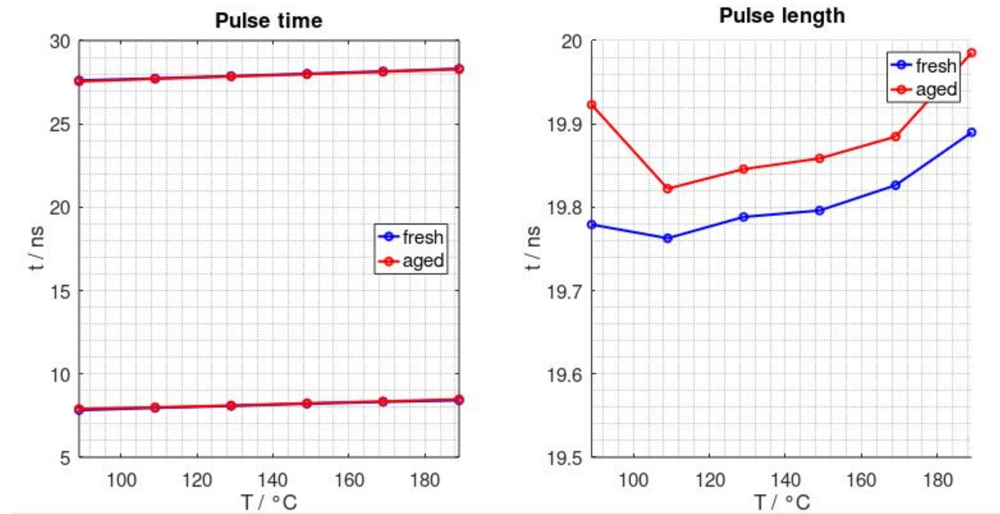

Figure 3.8: Fresh vs aged results; Left: pulse times, switch on and switch off; Right: pulse length

For the electro-thermal aware reliability simulation, group 0 was simulated for the case that one or two MOS drivers are in use and group 3 for the same comparison as in chapter 2.4.5. The changes in the pulse characteristics can be seen in table 3.3.

Table 3.3: Electro-thermal reliability simulation result

	$-I_{max}$ / A	$-I_{settling}$ / A	t_{start} / ns	t_{end} / ns	t_{PL} / ns
Grp0, 1 driver, T_{umg} = -40 °C	0.21	0.15	0.04	0.04	0.00
Grp0, 1 driver, T_{umg} = 29 °C	0.20	0.16	0.04	0.04	0.00
Grp0, 2 driver, T_{umg} = -40 °C	0.14	0.098	0.02	0.02	0.00
Grp0, 2 driver, T_{umg} = 29 °C	0.20	0.14	0.03	0.03	0.00
Grp3, 1 driver, T_{umg} = 29 °C	0.06	0.04	0.00	0.00	0.00

For settling current and maximum current, no rule of behaviour could be found. If only the high current path (power source transistor - multiplexer) is considered, the expectation would be that the change decreases as the temperature increases. This was compensated with a shift in the operation point of the current source transistor. Also, the pulse length shows no change, which is unexpected because of the difference in the RC network. At least

some change was anticipated.

Comparison between Groups

When comparing the electro-thermal aware aging between the outpust of group 0 and group 3, the results show that ageing equalizes the two groups between the two groups.

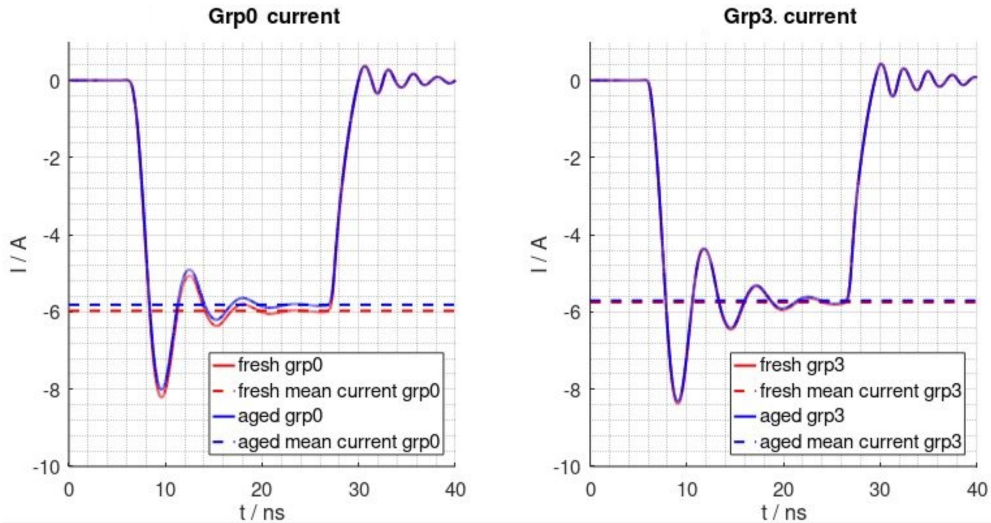

Figure 3.9: Driver current fresh and aged for different groups

4 Conclusion

4.1 Legato Tool for Electro-Thermal Simulation

The Legato tool is the primary tool used in this thesis. It is a tool designed to be easy to use and does only require limited theoretical knowledge about FEM simulation. The main advantage of the tool is direct communication with the SPICE simulation. This communication allows for a transient simulation of the temperature distribution of the chip, which is done via a co-simulation SPICE and thermal simulator. Even if this is not possible, for example, if the time to equilibrium is too long, an iteration between electrical and thermal solution can be made. This information has the advantage that a much more sophisticated power profile is used than if the chip is modeled. This is shown in chapter 2.4.5 on page 37. As for the drawbacks, two main disadvantages could be identified in the course of the thesis. First, the connection to the environment is modeled in this tool with six constant thermal resistances and six thermal capacities. This could lead to problems with more complicated packages. The second disadvantage is the direct coupling between electrical and thermal simulation. It is estimated in chapter 2.4.1 at page 27, that the time to thermal equilibrium for the MOS driver is around the magnitude of 700 ms. With a time step size of the SPICE simulation in the ps range, an electrical simulation up to 700 ms is impossible. A solution to this would be to choose a thermal time step and assume that the input power does not change during one step. This would allow that only one period of the electrical simulation has to be calculated per time step.

4.2 Electro-Thermal Simulation Results

4.2.1 Current Mirror

The primary test circuit of this thesis, the MOS driver, is essentially a high-performance current source as a second test circuit, therefore a current mirror was chosen. This circuit demonstrated that a schematic could perform differently depending on the layout, see table 2.4 on page 32.

4.2.2 MOS Driver

For the MOS driver, it was found that the main heat dissipation is done through the heat sink at the bottom (figure 2.17 on page 33).

This means it is possible to disregard the disadvantage of the Legato tool (that the thermal resistance can only be set to a constant), as previously discussed in chapter 2.3.2 on page 21. Because of this dependency on the heat sink for energy dissipation, the chip's temperature is highly dependent on the glue thickness, which badly conducts heat and acts as a heat break. The simulated temperature profiles in figure 2.20 on page 35 shows a linear dependency of the duty cycle. For lower values, a hotspot at the driver bias can be seen. While the electrical output resulting from this electro-thermal simulation can be distinguished from the uniform temperature simulation, the overlying curve form is similar. No failures could be found. This is also the case when comparing the performance of the different electrically identical groups of the device. The curve form is similar, but a difference in timing and current magnitude can be seen in figure 2.25 on page 38.

4.3 Electro-Thermal Aware Reliability Simulation Results

No HCI reliability model for the 180 nm technology was available at the time of the thesis, therefore a model from a similar technology was implemented. This means that the result of the chapter is only qualitative. It is demonstrated on the current mirror that the layout dependence of the degradation cannot be

excluded. For the MOS driver, the degradation result cannot only be derived from the degradation of the current source transistor. This is because there are also matched transistors in the driver bias and the current source driver module. If they age; this results in a change of the operating point. Also, in the comparison between group 0 and group 3, it is found that the ageing functions in the cause of the used model as compensation and equalizes the difference between these groups. When considering the actual reliability behaviour of the circuit, no conclusion could be made because of the usage of the estimated reliability model.

Bibliography

[1] B. Van Zeghbroeck. *Principles of Semiconductor Devices*. 2011. URL: http://ecee.colorado.edu/~bart/book/book/contents.htm (cit. on pp. 1, 3, 5).

[2] S.M. Sze Kwok K. Ng. *Physics of Semiconductor Devices*. First. John Wiley and Sons, 2006. ISBN: 978-0471143239 (cit. on pp. 1, 3).

[3] S.M.Sze. *Semiconductor Devices: Physics and Technology*. Second. Wiley India Pvt. Limited, 2008. ISBN: 978-8126516810 (cit. on pp. 1, 3).

[4] Y. Cheng C. Enz. "MOS transistor modeling for RF IC design." In: *IEEE Journal of Solid-State Circuits* 35 (2 2000), p. 186 (cit. on p. 4).

[5] et al. N. C. Dao A. E. Kass. "An enhanced MOSFET threshold voltage model for the 6–300 K temperature range." In: *Microelectronics Reliability* 69 (2017), p. 36 (cit. on p. 10).

[6] B. Meinerhagen Y. Wenger. "Low-Voltage Current and Voltage Reference Design Based on the MOSFET ZTC Effect." In: *IEEE Transactions on Circuits and Systems I: Regular Papers* 66 (9 2019), p. 3445 (cit. on p. 10).

[7] A. Allam I. M. Filanovsky. "Mutual Compensation of Mobility and Threshold Voltage Temperature Effects with Applications in CMOS Circuits." In: *IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—I: FUNDAMENTAL THEORY AND APPLICATIONS* 48 (7 2001), p. 876 (cit. on p. 10).

[8] et al. M. Miura-Mattausch. *HiSIM2 3.0.0 User's Manual*. Hiroshima University and STARC, 2017 (cit. on p. 18).

[9] et al. T. H. Morshed. *BSIM4.6.4 MOSFET Model*. Department of Electrical Engineering and Computer Sciences University of California, Berkeley, 2009 (cit. on pp. 18, 47).

[10] Andrei Vladimirescu. *The SPICE Book*. First. John Wiley and Sons, 1994. ISBN: 978-0471609261 (cit. on p. 19).

[11] et. al. N. M. Ravidra. "Temperature-Dependet Emissivity of Silicon-Related Materials and Structures." In: *IEEE Transaction on semiconductor manufactorin* 11 (1 1998), p. 30 (cit. on p. 22).

[12] M. S. Bohn F. Kreith. *Principles of Heat Transfer*. Forth. Harper and Row, 1986. ISBN: 978-1-4757-0911-7 (cit. on p. 22).

[13] M. A. Bramson. *Infrared Radiation: A Handbook for Applications*. 1st ed. Springer US, 1968. ISBN: 978-1-4757-0911-7 (cit. on p. 22).

[14] L. M. Jiji. *Heat Convection*. second. Springer Berlin, 2010. ISBN: 978-3642029707 (cit. on p. 23).

[15] A. Jaffer. *Convection from a rectangular plate*. 2019. URL: `https://people.csail.mit.edu/jaffer/SimRoof/Convection/` (cit. on p. 23).

[16] A. J. N. Khalifa. "Natueral convective heat transfer coefficient - a review I. Isolated vertical and horizontal surfaces." In: *Energy Conversion and Managment* 42 (2001) (cit. on p. 23).

[17] E. Sartori. "Convection coefficient equation for forced air flow over flat surfaces." In: *Solar Energy* 80 (2006), p. 1063 (cit. on p. 23).

[18] Kahsay Godifey Wubneh. "Solving Fundamental Solution of Non-Homogeneous Heat Equation with Dirichlet Boundary Conditions." In: *Bulletin of Mathematical Sciences and Applications* 22 (10 2020), p. 1 (cit. on p. 24).

[19] J. W. McPherson. *Reliability Physics and Engineering*. Second. Spriner, 2013. ISBN: 978-3-319-00121-0 (cit. on p. 40).

[20] J. W. McPherson. *Applied Reliability*. Thrid. Chapman and Hall, 2011. ISBN: 978-1-584-88466-8 (cit. on p. 40).

[21] Bianka Ullmann. "Mixed Negative Bias Temperature Instability and Hot-Carrier Stress." PhD thesis. Vienna University of Technology, 2018 (cit. on p. 42).

[22] Tibor Grasser Stanislav Tyaginov. "Modeling of Hot-Carrier Degradation: Physics and Controversial Issues." In: *IIRW Final Report*. 2012, p. 206 (cit. on pp. 42, 43).

[23] Brain Doyle Kaizad Mistry. "AC versus DC Hot-Carrier Degradation in n-Channel MOSFETs." In: *IEEE Transactions on Electronic Devices* 44.1 (1993), p. 96 (cit. on p. 43).

[24] Rudi Bellens Paul Heremans. "Consistent Model for the Hot-Carrier Degradation in n-Channel and p-Channel MOSFET's." In: *IEEE Transactions on Electronic Devices* 35.12 (1988), p. 2194 (cit. on p. 43).

[25] Y. Miura and Y. Matukura. "Investigation of Silicon-Silicon Dioxide Interface Using MOS Structure." In: *Japanese Journal of Applied Physics* 5 (1966), p. 180 (cit. on p. 43).

[26] Robert Entner. "Modeling and Simulation of Negative Bias Temperature Instability." PhD thesis. Vienna University of Technology, 2007 (cit. on p. 43).

[27] N. Parihar S. Mahapatra. "A review of NBTI mechanisms and models." In: *Microelectronics Reliability* 81 (2018), p. 127 (cit. on pp. 43, 44).

[28] C. Parthasarathy V. Huard M. Denais. "NBTI Degradation: From Physical Mechanisms to Modelling." In: *Microelectronics Reliability* 46 (2006), p. 1 (cit. on p. 43).

[29] et al. A. J. Scholten D. Stephens. "The Relation Between Degradation Under DC and RF Stress Conditions." In: *IEEE Transactions on Electron Devices* 58 (8 2011), p. 2721 (cit. on pp. 46, 47).

[30] et al. R.H. Tu. "Berkeley Reliability Tools - BERT." In: *IEEE Transactions on computer added design of integrated curcuits and systems* 12 (10 1993), p. 1524 (cit. on p. 47).