DIPLOMARBEIT

Thermal Imaging of Smart Power DMOS transistors in the thermally unstable regime

ausgeführt zum Zwecke der Erlangung des akademischen Grades eines Diplom-Ingenieurs unter der Leitung von

a.o. Univ.-Prof. Dipl.-Ing. Dr. Dionyz Pogany
am Institut für Festkörperelektronik (E362)

eingereicht an der Technischen Universität Wien
an der Fakultät für Elektrotechnik und Informationstechnik
von

Georg HABERFEHLNER
Matrikelnummer 0326526
Raiffeisenstraße 27, 3300 Amstetten

Wien, im Juli 2009

Kurzfassung

Double-diffused MOS (DMOS) Transistoren, welche in Smart Power Technologien hergestellt werden, sind als Leistungstransistoren weit verbreitet. Hohe Leistungsdichten unter Kurzschlussbedingungen können zu starkem Temperaturanstieg führen und machen die Temperatur in den Bauelementen zu einem wesentlichen limitierenden Faktor für die Integrationsdichte. Vor allem in größeren DMOS Transistoren kann ungleichmäßige Verteilung des Stromes unterhalb des so genannten Temperaturkompensationspunktes (Temperature Compensation Point, TCP) zu Regionen mit erhöhter Temperatur und dadurch zur Zerstörung des Bauelements führen.

Um Informationen über die Temperaturverteilung in den untersuchten Transistoren zu erhalten wurden interferometrische Methoden angewandt. Bei diesen optischen Interferenzmethoden wird der temperaturabhängige Brechungsindex mit Hilfe eines nicht absorbierten Laserstrahls direkt in der aktiven Zone gemessen.

In den durchgeführten Experimenten wurde eine neue, adaptierte Variante eines zweidimensional transienten interferometrischen Abbildungsverfahrens (2-dimensional transient interferometric mapping, TIM) verwendet. Mit Hilfe von zwei Superlumineszenzdioden und digitalen Infrarotkameras wird die Temperaturverteilung in zwei Dimensionen zu zwei Zeitpunkten während eines einzigen elektrischen Pulses aufgezeichnet. Im Verlauf dieser Arbeit wurde das Messverfahren für zwei Kameras automatisiert, so dass nun schnelle Messungen mit geringem Aufwand möglich sind.

Unter thermischer Belastung ist die Aktivierung eines parasitären Bipolartransistors, der in der DMOS Struktur vorhanden ist, eine der wichtigsten Ursachen für das Versagen von Bauelementen. In dieser Arbeit wurde die Temperaturverteilung in mehreren DMOS Transistoren unterhalb und oberhalb des TCP gemessen. Im Speziellen wurde die Aktivierung des parasitären Bipolartransistors experimentell beobachtet, so dass die Größe der thermisch instabilen Region festgestellt werden konnte.

Abstract

Smart power double-diffused MOS (DMOS) transistors are widely used power devices. Under short-circuit conditions the power dissipation in such devices can be very high, which makes temperature increase due to self-heating a major limiting factor for increasing integration densities. Especially in relatively large DMOS transistors, inhomogeneous current flow due to positive thermo-electric feedback below the so-called temperature compensation point can lead to hot regions causing failure of the device.

Interferometric methods are employed to get information about the temperature in DMOS transistors. With optical interferometry changes of the temperature dependent refractive index are recorded directly in the active region by the observation of the phase shift of a non-absorbed optical beam.

For the experiments a new version of the 2-dimensional transient interferometric mapping method was used. This setup uses two superluminescent diodes and two focal plane array cameras, which allow high accuracy 2-dimensional thermal mapping at two time instants during a single stress pulse. During work on this thesis the control of the setup with two cameras was developed. An automated procedure controlled by a PC is now available, which allows fast measurements.

An important failure mechanism in DMOS transistors under thermal stress is the triggering of a parasitic bipolar transistor inherent to the DMOS structure by thermal carrier generation. In this work the temperature distribution in multiple DMOS transistors is analyzed below and above temperature compensation point. In particular the thermal activation of the parasitic bipolar transistor is observed experimentally and the size of the region, where thermal runaway occurs is determined.

Acknowledgments

I want to thank all the people who supported me during this thesis. I really enjoyed the opportunity to work in the device characterization group at the Institute for Solid State Electronics. The surroundings and the support I experienced made the work on this thesis a very interesting and memorable time.

First of all I want to express my gratitude to Prof. Dr. Dionyz Pogany for his excellent supervision of my work, for being always available for discussion and for sharing his profound knowledge, making things easily understandable.

I want to thank the members of the device characterization group, Dr. Sergey Bychikhin, Michael Heer and Wasinee Mamanee for including me in the group and for helping me to overcome all problems. To Sergey I am especially obliged for his patience in explaining and discussing various theoretical and practical aspects of the experiments. To Michael I want to express my gratitude for his excellent support with programming and measurement instruments and for various ideas and proposals for the practical use of the setups. I want to thank the technicians Wolfgang Schramek and Johannes Prinzinger.

I am grateful to Infineon Technologies Munich for providing the samples for the experiments. My special thanks go to Alevtina Podgaynaya for explaining and discussing the devices.

I am very thankful to my family and friends for their moral support during my studies. Thanks go to my sister Christa and my brother Peter for proof reading this thesis and I want to especially thank my parents, who provided excellent background conditions throughout my studies.

Contents

Chapter 1

Introduction

Smart power technologies combine bipolar, CMOS and double diffused MOS (DMOS) transistors on a single substrate. Therefore they are also known as BCD (Bipolar, CMOS, DMOS) technologies. This allows the combination of logic, analog circuits and power applications and provides high flexibility.

Smart power chips are used in a wide field of applications [1, 2]. At low currents and moderate voltages they are used for display drives. An application requiring low voltages but higher currents are power supplies for computers or telecommunications. On the other end of the spectrum, applications requiring high voltages and high currents are motor controls, traction for transportation systems or high voltage DC systems. One major field for the use of smart power technology, which is the focus of this work, is the automotive sector. Here data from multiple sensors has to be processed and actuators have to be driven at high powers. Compared to other applications of smart power technology, the maximum voltage ratings of these applications are relatively moderate (< 100V).

As output stages for power applications DMOS transistors are used. DMOS transistors provide high forward currents and at the same time they can sustain high voltages and have the high input impedance of MOS technologies.

Similar to CMOS or BiCMOS technologies, the integration density in smart power chips has increased rapidly since they were first introduced. This leads to a high power dissipation at increasingly smaller areas, especially in DMOS transistors. Therefore the temperature in these devices has become a very important issue and is a major limiting factor for a further increase of the integration density.

Especially in larger devices the temperature distribution can become highly inhomogeneous and has to be taken into account for a correct estimation of the thermal

safe operating area (SOA) [3, 4]. Experimental knowledge of the device behavior under thermal stress is vitally important for the development and verification of simulation tools [5, 6].

Several methods exist to get experimental information about the temperature distribution in semiconductors, as many properties of a material depend on temperature [7]. Infrared thermography measures the temperature dependent radiation emitted by a device. In liquid crystal thermography the sample is coated with a layer of liquid crystals and temperature dependent properties of the liquid crystals are analyzed. These two methods measure only surface temperature and suffer from a low time resolution. Methods for retrieving temperature information directly in the active region are for example internal infrared deflection [8], which measures the gradient of the temperature dependent refractive index by deflection of a non-absorbed laser beam or Raman spectroscopy [9] where phonon frequencies are measured by analyzing the frequency shift of backscattered light. A method with a very high spatial resolution is scanning thermal microscopy [10], which is a variation of atomic force microscopy with a temperature sensitive tip. However this method is time consuming and requires expensive instrumentation.

The method employed for thermal characterization in this work is optical interferometry. This method measures changes in the refractive index, which is influenced by temperature, directly in the active region. These changes are detected with an interferometer, which provides high sensitivity on changes in the optical path of a non-absorbed laser beam probing the device. The spatial resolution is governed by diffraction limits and a very high time resolution can be achieved with this method [11], which is useful especially for the analysis of electrostatic discharge effects.

At low gate voltages the drain current in DMOS transistors increases with temperature, which can lead to failure due to a local hot spot [12]. This positive thermoelectrical feedback can lead to the activation of a parasitic bipolar transistor inherent to the DMOS structure by thermally generated free carriers or directly to thermal runaway of the body-drain pn-junction [3, 5, 6, 13]. The area where triggering of the parasitic bipolar transistor takes place is retrieved experimentally.

For this purpose an adapted interferometric setup is used with a moderate time resolution in the microsecond range, sufficient for stress pulse durations between 100µs and a few milliseconds. With these relaxed requirements on time resolution it is possible to increase the sensitivity of the setup. Using this setup DMOS transistors

are examined under short circuit conditions at stress pulses in the millisecond range up to destruction of the devices.

1.1 Outline of this work

Chapter 2 describes smart power DMOS transistors. Their types and analytical models are presented. A special focus rests on temperature dependent properties and on thermal failure. Some of the examined devices have built-in temperature sensors. Their function is explained.

In chapter 3 the optical methods used in the experiments are discussed. The principles of optical interferometry are addressed generally and more specifically with respect to the used setups. The optical principles of two different setups are presented in this chapter. The 2D SLD setup is used for thermal imaging at two time instants during a single stress pulse and was improved during work on this thesis. With the scanning heterodyne setup a device can be scanned along its x- and y-axis to get phase shift transients at multiple positions of the device. Under the examined pulse durations, deformation of the sample is also an issue and is discussed at the end of this chapter.

The control units for the 2D SLD setup have been adapted and improved to allow imaging at two time instants during a single pulse. This is addressed in chapter 4. The electrical connections of the measurement instruments are explained and special focus is put on the timing of the instruments. Furthermore the principles of the developed control software are explained.

In the fifth chapter preparation of the samples for the measurements is addressed. This includes the calibration of built-in temperature sensors.

In chapter 6 the results measured on three different DMOS devices are presented. A small vertical DMOS transistor is analyzed at multiple conditions with both measurement setups. In these measurements the performance of the SLD setup is tested. With the scanning heterodyne setup experimental information about the 2-dimensional power dissipation in the small device is retrieved. A larger DMOS device is examined under destructive conditions. Thermal runaway in the device is recorded experimentally. Finally similarly processed n-channel and a p-channel lateral DMOS transistors are compared.

Chapter 7 summarizes this thesis with conclusions and an outlook.

Chapter 2

Smart power DMOS transistors

All the devices examined in this thesis are DMOS transistors fabricated in a smart power technology. Smart power technologies integrate bipolar, CMOS and DMOS transistors in one substrate. In this chapter the layout of DMOS devices is described and analytical models are given. The issue of thermal SOA is addressed. As some of the investigated transistors feature built-in temperature sensors, the principles of those sensors are described.

2.1 Types of DMOS transistors

The term DMOS transistor is derived from double-diffused MOS transistor, named after the processing of body and source region. DMOS transistors are designed to sustain high powers and to drive high currents at a low on-resistance R_{on}. DMOS transistors are characterized by their maximum drain-source voltage rating. In this work DMOS transistors aimed at automotive applications are examined, which are guaranteed to sustain overvoltages of around 60V.

DMOS transistors are realized as vertical devices (VDMOS) as well as as lateral devices (LDMOS). A schematic of a vertical n-channel DMOS transistor is shown in Fig. 2.1. "Vertical" refers to the main direction of the current flow in the active region. In an n-channel VDMOS, the n^+-doped drain region is a buried layer, which is contacted by an n^+-sinker. The n^+-doped source and p-doped body are biased by common contacts. The gate opens the channel in the relatively narrow p-body region, which allows the current flow between source and drain through a lightly n-doped epi-layer. At high drain-source voltages the largest electrical field is applied in this epi-layer. Therefore the length and doping of this region determines

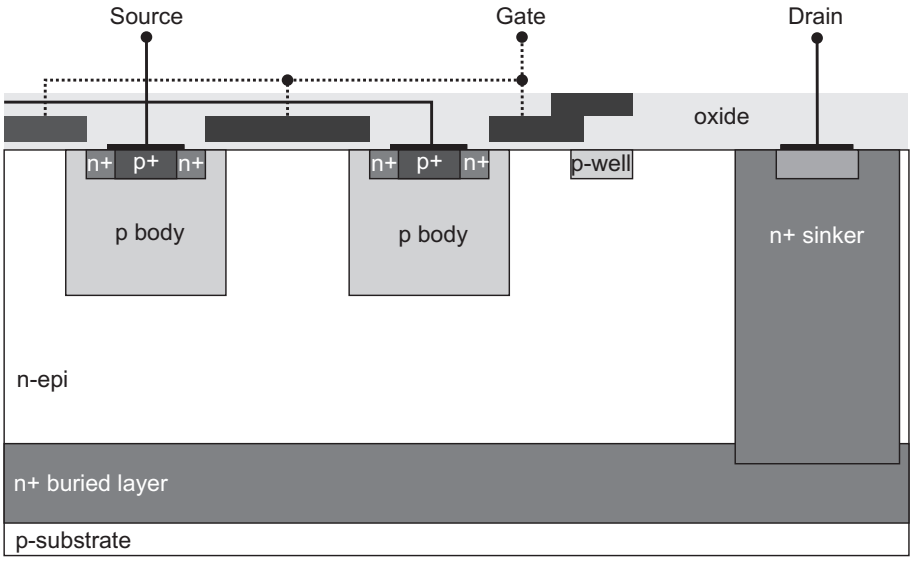


Figure 2.1: Schematic of a vertical n-channel DMOS transistor

the breakdown voltage of the device and guarantees the high robustness of DMOS transistors on V_{ds}. VDMOS devices usually consist of many parallel source/body cells, which are referred to as source field. This allows the creation of large area devices providing high forward currents.

The spacing between neighboring cells has to be optimized to minimize the value of $R_{\mathrm{on}} \times A$. A small spacing between the cells leads to a high R_{on}, if the n-epi layer between the cells gets too thin. Too large spacing increases the total area of the device without significant improvement of the on-resistance. Therefore $R_{\mathrm{on}} \times A$ becomes large.

The concept of a lateral n-channel DMOS transistor is shown in Fig. 2.2. In these devices the current flows along the surface between the drain and source region. In comparison to the vertical transistor, the source/body and gate areas are built in a similar way. However the n-doped drain does not consist of a buried layer, but of highly doped diffusion regions at the surface. Similar to the vertical DMOS, many cells are used in parallel. An advantage of the LDMOS concept is that the length of the epi-layer can be scaled and therefore the breakdown voltage can be scaled without changing process parameters and influencing other devices.

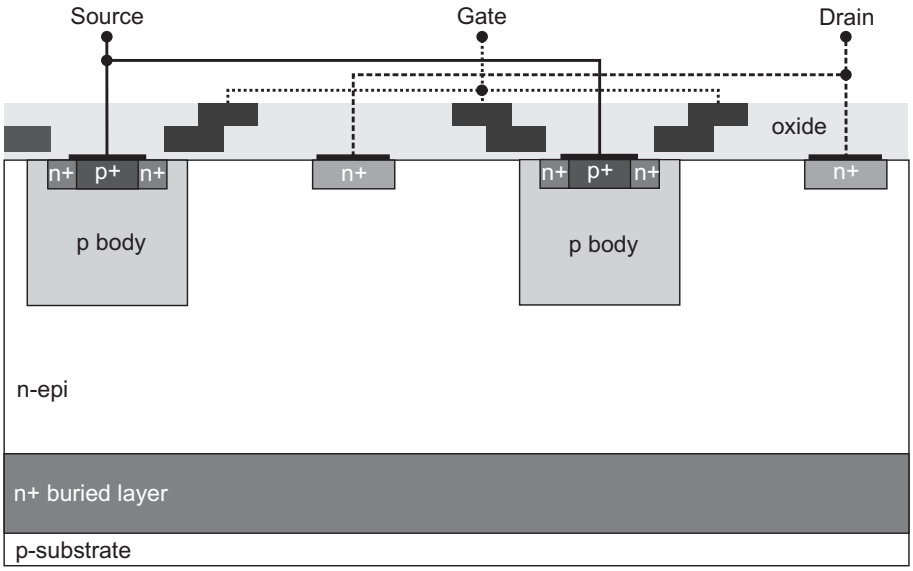


Figure 2.2: Schematic of a lateral n-channel DMOS transistor

2.2 DMOS models

The model for the drain current I_d of a MOS transistor in saturation [12, 14] is given by

$$I_\mathrm{d}(T) = \frac{\mu_\mathrm{n}(T) C_\mathrm{ox} W}{2L} \left(V_\mathrm{gs} - V_\mathrm{th}(T)\right)^2 \tag{2.1}$$

V_gs is the gate-source voltage, V_th is the threshold voltage, W is the channel width, L the channel length and C_ox the gate oxide capacitance per unit area. C_ox can be calculated from the oxide permittivity ϵ_ox and thickness t_ox as

$$C_\mathrm{ox} = \frac{\epsilon_\mathrm{ox}}{t_\mathrm{ox}} \tag{2.2}$$

The threshold voltage [15] of a DMOS transistor is given as

$$V_\mathrm{th}(T) = \phi_\mathrm{ms} + 2\phi_\mathrm{f} + \frac{\sqrt{2\epsilon_\mathrm{s} q N_\mathrm{max}(2\phi_\mathrm{f})}}{C_\mathrm{ox}} - \frac{Q_\mathrm{ss}}{C_\mathrm{ox}} \tag{2.3}$$

In (2.3) ϕ_ms is the metal to semiconductor work function difference, ϕ_f is proportional to the difference between the Fermi level in an intrinsic and the doped semiconductor, ϵ_s is the permittivity of the semiconductor, N_max is the maximum bulk impurity concentration and Q_ss is the surface state charge density. The electron charge $q = 1.6 \times 10^{-19}\mathrm{C}$.

The squareroot term in (2.3) is known as bulk charge Q_b as

$$Q_\mathrm{b} = -\sqrt{2\epsilon_\mathrm{s} q N_\mathrm{max}(2\phi_\mathrm{f})} \tag{2.4}$$

Concerning the behavior at different temperatures, the main temperature dependence of V_th comes from the influence of ϕ_f

$$\phi_\mathrm{f} = \frac{kT}{q} * \ln \frac{N}{n_\mathrm{i}(T)} \tag{2.5}$$

with the temperature dependent intrinsic concentration n_i

$$n_\mathrm{i}(T) = 3.87 \times 10^{16} * T^{3/2} * \exp(-E_\mathrm{g}/2kT) \tag{2.6}$$

N is the doping concentration, E_g is the bandgap energy, the Boltzmann constant $k = 1.38 \times 10^{-21}\mathrm{J/K}$.

From the previous equations the temperature dependence of the threshold voltage can be calculated as

$$\frac{dV_\mathrm{th}}{dT} = \left(\frac{\phi_\mathrm{f}}{T} - \frac{k}{q}\left(\frac{E_\mathrm{g}}{2kT} - \frac{3}{2}\right)\right) * \left(2 + \frac{Q_\mathrm{b}}{2\phi_\mathrm{f} C_\mathrm{ox}}\right) \tag{2.7}$$

The second temperature dependent term in (2.1) is the mobility of electrons μ_n. Its temperature dependence can be described as

$$\mu_\mathrm{n}(T) = \mu_\mathrm{n}(T_0)\left(\frac{T}{T_0}\right)^{-\chi} \tag{2.8}$$

χ is an exponent between 1 and 2.5 depending on the body doping profile, T_0 is a reference temperature.

The differentiation of (2.1) gives the temperature dependence of I_d

$$\frac{dI_\mathrm{d}}{dT} = I_\mathrm{d} * \left(\frac{1}{\mu_\mathrm{n}}\frac{d\mu_\mathrm{n}}{dT} - \frac{2}{V_\mathrm{gs} - V_\mathrm{th}}\frac{dV_\mathrm{th}}{dT}\right) \tag{2.9}$$

In (2.9), there are two concurrent parts. These parts result from the decrease of V_th and of mobility with rising temperature which are competing processes. Therefore there exists a point, when these two parts cancel each other, the condition for which is

$$\frac{dI_\mathrm{d}}{dT} = 0 : \frac{1}{\mu_\mathrm{n}}\frac{d\mu_\mathrm{n}}{dT} = \frac{2}{V_\mathrm{gs} - V_\mathrm{th}}\frac{dV_\mathrm{th}}{dT} \tag{2.10}$$

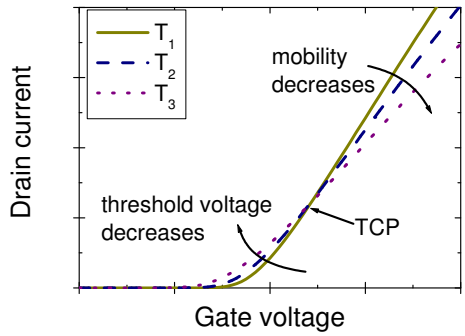


Figure 2.3: Transfer characteristics of a DMOS transistor at three temperatures ($T_1 < T_2 < T_3$)

This point is called temperature compensation point (TCP).

The influence of temperature on the transfer characteristics can be seen in Fig. 2.3. At gate voltages below TCP, the drain current increases with increasing temperature as the decrease of threshold voltage has a larger influence than the decrease of mobility, whereas above TCP the influence of the mobility is higher, therefore the current drops with increasing temperature.

2.3 Thermal failure in DMOS devices

In this work the focus of analysis is on short circuit conditions. It is assumed that a load is short circuited while it is driven by a DMOS transistor (see Fig. 2.4). In this case the supply voltage of the battery is applied directly to the DMOS until the error is recognized and handled. This leads to a high power dissipation in the device for relatively long times, which causes high temperatures in the transistor.

The most important reason for thermal failure in DMOS devices is the active biasing of the parasitic bipolar transistor inherent to the DMOS structure. A VDMOS with a schematic of its parasitic bipolar transistor is depicted in Fig. 2.5.

The npn-transistor can be triggered by biasing of the body [13, 14]. At high temperatures a current is thermally generated at the reverse biased drain-body junction. This current causes a voltage drop on the body resistance, which gets high enough to trigger the parasitic transistor at high temperatures. The triggering of the bipolar transistor leads to an exponential increase of drain current, which causes thermal runaway and failure of the device.

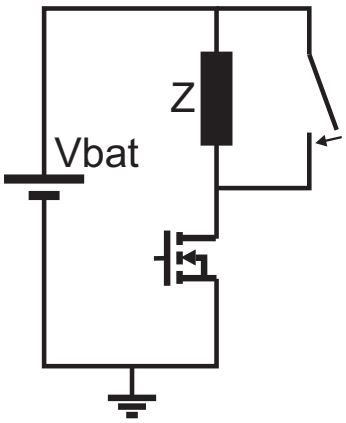


Figure 2.4: Short circuit condition for a low side DMOS driver

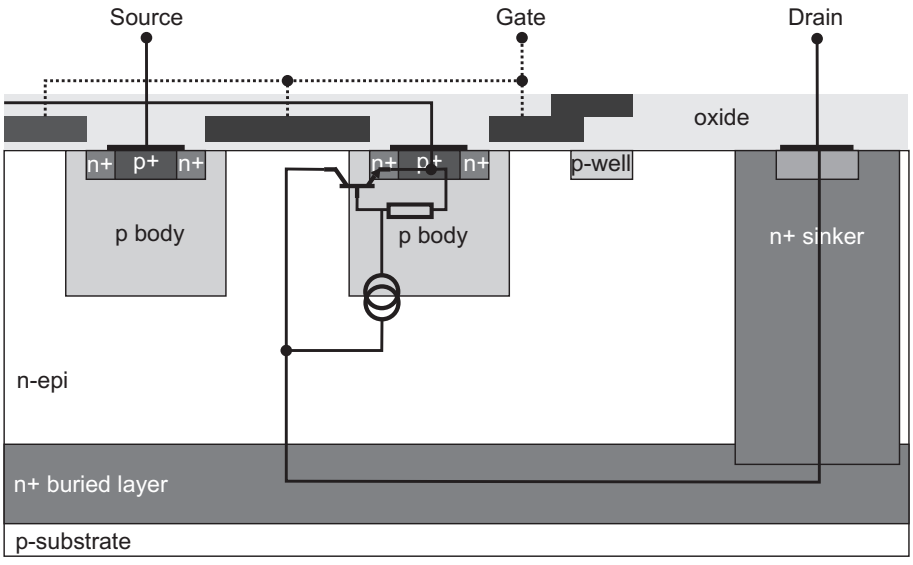


Figure 2.5: VDMOS transistor with parasitic bipolar transistor

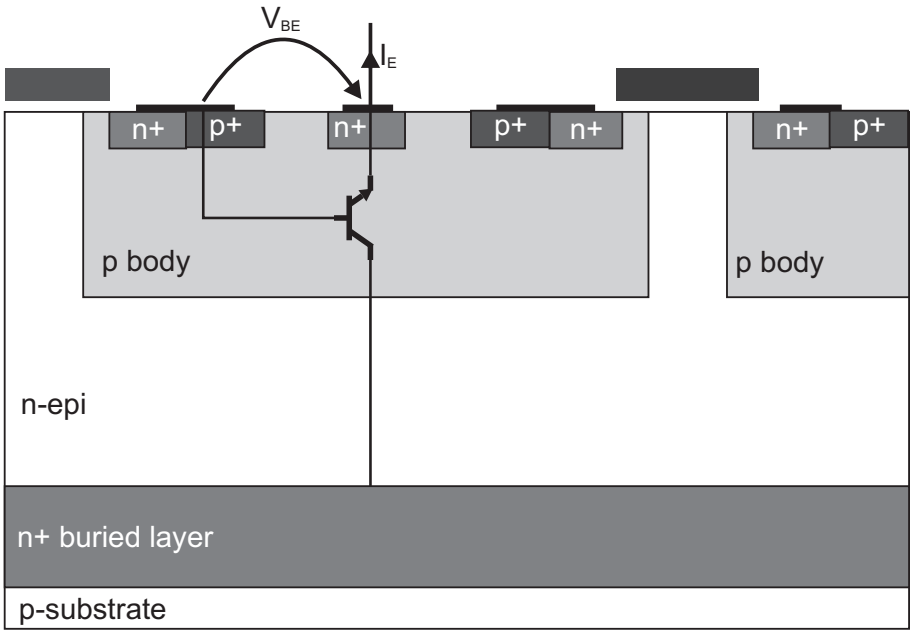


Figure 2.6: VDMOS transistor with built-in temperature sensor

2.4 Temperature sensors

In some of the investigated VDMOS devices, diodes or bipolar transistors are included as temperature sensors. An example of such a sensor is shown in Fig. 2.6 [5]. In a VDMOS array consisting of many cells one of these cells is exchanged for a sensor. The sensor consists of an npn-transistor. The base is built by the body region of the DMOS and is therefore connected to the drain. The n^--doped epi-layer works as the collector and an additional n^+-diffusion region works as the emitter. This diffusion region can be contacted separately, while base and collector are tied to the potential of source and drain of the DMOS device.

The base-emitter junction of the sensor is forward biased at a constant emitter current I_E. The voltage between base and emitter V_BE is measured. As this voltage depends nearly linearly on temperature, the temperature in the base-emitter junction can be retrieved. The influence of the collector voltage is very small and can be neglected, which allows temperature measurements also on the activated DMOS transistor. Therefore in the rest of this thesis only the pn-diode structure will be regarded for all temperature sensors.

Chapter 3

Optical Techniques

In this chapter the optical setups used for thermal characterization of power DMOS transistors are described. First the principles of optical interferometry for mapping of temperature are presented, then the used optical setups are discussed.

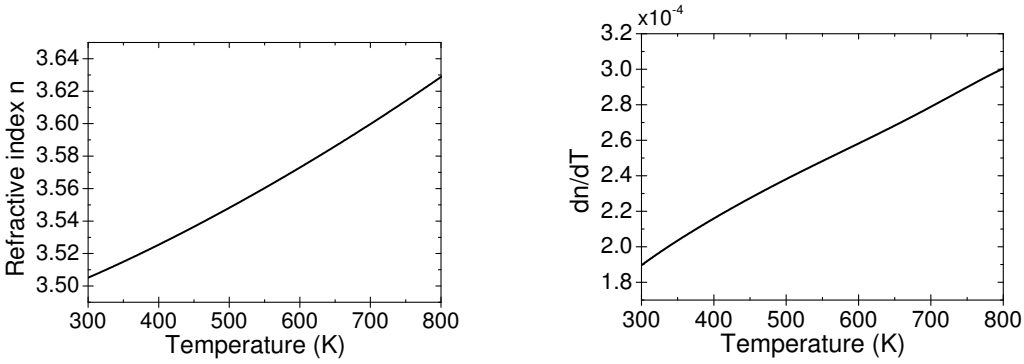
The 2D SLD setup has been developed based on an existing setup [16]. In this chapter the optical principles of the SLD setup are discussed. Improvements of the control of the setup are discussed separately in the next chapter. In addition to the 2D SLD setup an existing scanning heterodyne interferometer [17] has been used for further measurements and to compare the performance of the two setups. The principle of this setup is explained briefly. The methods for extraction of the phase shift from the measured data are given for both setups.

At millisecond stress pulses deformation due to thermal expansion of the samples during the stress becomes an issue and is therefore addressed at the end of this chapter.

3.1 Optical Interferometry

The refractive index of a material depends on temperature and carrier density. These mechanisms are known as thermo-optical and plasma-optical effects. So with a local change of temperature $\Delta T(x, y, z, t)$, electron density $\Delta N_{\mathrm{e}}(x, y, z, t)$ and hole density $\Delta N_{\mathrm{h}}(x, y, z, t)$, the refractive index $n(x, y, z, t)$ varies.

Fig. 3.1 shows the temperature dependence of the refractive index in silicon for a wavelength of 1.31µm. Also dn/dT varies with temperature (see Fig. 3.1b) but this is neglected in the following discussion.



(a) $n(T)$ **(b)** $dn/dT(T)$

Figure 3.1: Temperature dependence of the refractive index (after [18])

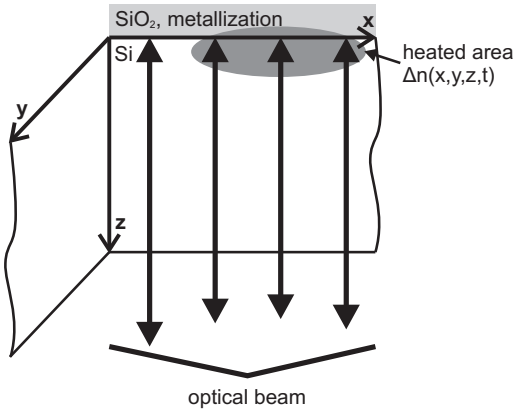


Figure 3.2: Probing by an optical beam

In optical interferometry the variation of the refractive index is measured by an optical beam propagating from the back side of a sample to the top side, where it is reflected, so it leaves the sample again on the back side (see Fig. 3.2).

Therefore the phase shift of an optical beam caused by the thermo- and plasma-optical effects can be calculated as

$$\Delta\varphi(x,y,t) = 2 * \frac{2\pi}{\lambda} \int (\Delta n_{\mathrm{T}}(x,y,z,t) + \Delta n_{\mathrm{e}}(x,y,z,t) + \Delta n_{\mathrm{h}}(x,y,z,t))dz \quad (3.1)$$

where the factor 2 at the beginning is caused by the double passage of the beam.

At the investigated pulse lengths the influence of the plasma-optical effect is much smaller than the thermo-optical effect, so in this work the contribution of free

carriers is neglected and only the contribution by temperature is examined. In this case and assuming dn/dT to be independent of temperature (3.1) simplifies to

$$\Delta\varphi(x,y,t) = \frac{4\pi}{\lambda}\frac{dn}{dT}\int_0^l (\Delta T(x,y,z,t))dz \tag{3.2}$$

with $\frac{dn}{dT} \approx 1.9 \times 10^{-4}$ at $T = 300\text{K}$ (see Fig. 3.1b) and l being the thickness of the sample.

To get a relation between power dissipation and phase shift, the thermal diffusion equation can be used as [19]

$$c_\text{V}\frac{\partial\Delta T(x,y,z,t))}{\partial t} = P_\text{3D}(x,y,z,t) + \kappa\left(\frac{\partial^2\Delta T}{\partial x^2} + \frac{\partial^2\Delta T}{\partial y^2} + \frac{\partial^2\Delta T}{\partial z^2}\right) \tag{3.3}$$

In this equation P_3D is the three-dimensional power dissipation density, κ is the thermal conductivity of the material and c_V is the volume thermal capacity.

Combining (3.2) and (3.3) gives the 2-dimensional power dissipation

$$\begin{aligned} P_\text{2D}(x,y,t) = &\left(\frac{dn}{dT}\right)^{-1}\frac{\lambda}{4\pi}* \\ &\left(c_\text{V}\frac{\partial\Delta\varphi(x,y,t)}{\partial t} - \kappa\left(\frac{\partial^2\Delta\varphi(x,y,t)}{\partial x^2} + \frac{\partial^2\Delta\varphi(x,y,t)}{\partial y^2}\right)\right) + j_\text{th,z}(x,y,0,t) \end{aligned} \tag{3.4}$$

P_2D is given as the integral of P_3D along the z-axis and $j_\text{th,z}(x,y,0,t)$ gives the thermal power transfer from silicon to top layers of the device, like SiO_2, passivation or metalization layers. In the experimental estimation of P_2D in Section 6.1.3, this term is neglected.

3.2 2D SLD setup

A setup for 2-dimensional transient interferometric mapping (TIM) at two time instants has been developed. This setup uses the same principles as a previous setup for imaging at two time instants, which was designed mainly for analysis of electrostatic discharge (ESD) protection devices [20].

The advantage of the new system is its high compactness, avoiding large water-cooled high energy laser sources, and a high low-frequency mechanical stability. Together with the high pulse to pulse stability of the used superluminescent diodes (SLDs) this improves the phase accuracy of the measurements. The time resolution

under single shot operation is 20µs which is limited by the speed of the used acousto-optic modulators (AOMs), the single-pulse detection limit of the cameras and the SLD power. Though this time resolution is much lower than the ns resolution of the previous setup it is sufficient for thermal imaging of stress pulses in the time scale of 100µs to a few milliseconds.

In 2D TIM the phase shift in the whole device is mapped during a single stress pulse at a given time instant. For this purpose two interferograms are recorded, one reference interferogram before the electrical stress and a stress interferogram during the stress. The optical principles of the 2D SLD setup are described first. After that, the fast Fourier transform (FFT) method used for phase shift extraction from interferograms is explained.

3.2.1 Optical setup

A schematic of the optical setup can be seen in Fig. 3.3. The SLD setup is based on holographic interferometry using a Michelson interferometer. The used 2D TIM system for 2D thermal imaging was originally developed for measurements of small repetitive signals [16]. The enhancement of the existing setup includes the implementation of a second interferometer allowing to measure two images at two InGaAs focal plane array (FPA) IR cameras.

Two SLDs serve as light sources for the setup. The SLDs emit in the infrared at a wavelength of $\lambda = 1.3\mu m$. The emitted power is 9mW. The use of SLDs is necessary as a short coherence length is required to get interference only from a single reflecting surface. If laser diodes were used instead of SLDs, their long coherence length would lead to interference not only from the portion of the light reflected from the top side layers of the device, but also from the back side surface or from multiple reflections inside the sample.

As the SLDs themselves cannot be pulsed, the beams are chopped by AOMs. Inside the AOM an acoustic wave modulates the optical beam and creates several higher order beams with slightly shifted wavelengths at different angles. In this setup the shift of the wavelength is not employed and has no significant impact, but the angular dependence of the optical beams on the acoustic frequency is used. The frequencies of the AOMs are set in such a way that only the 1st-order beams can pass through an iris, whereas zero- and higher order beams are blocked. The iris is located after a polarizing beam splitter which combines the two beams.

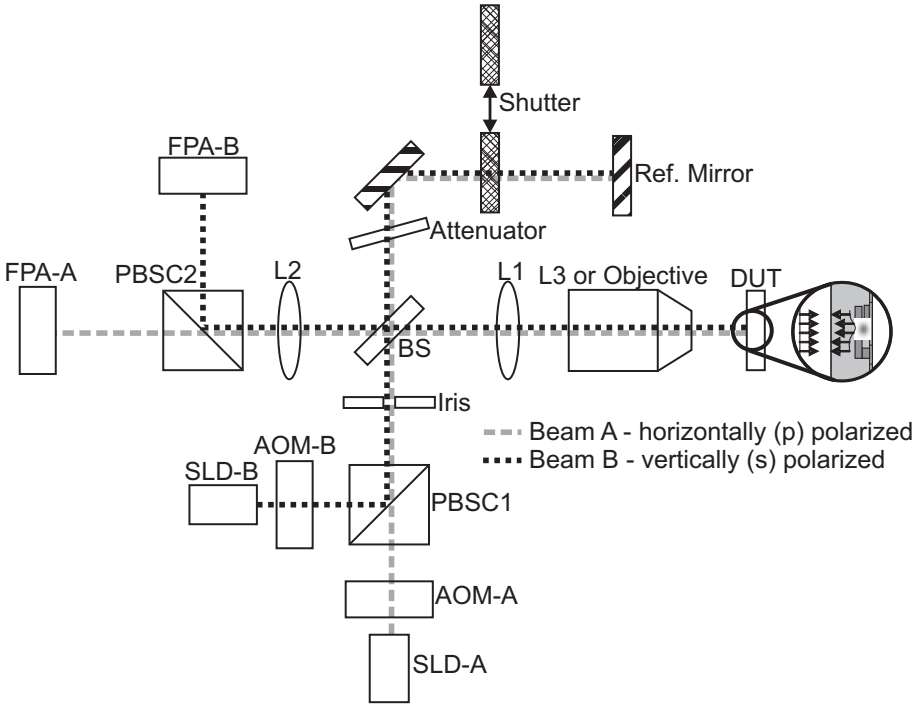


Figure 3.3: Schematic of the 2D SLD setup, abbreviations are: SLD: Superluminescent diode, PBSC: Polarizing beam splitter cube, L: Lens, BS: Beam splitter, FPA: Focal plane array camera, DUT: Device under test

This first polarizing beam splitter cube (PBSC) is used to polarize the beams orthogonally. Horizontally (p-) polarized light is transmitted, whereas vertically (s-) polarized light is reflected at an angle of 90°. Therefore only the p-polarized portion of the light from SLD A and the s-polarized portion of the light from SLD B enter the interferometer.

After passing through the iris a beam splitter divides the combined orthogonally polarized 1st-order beams on a reference and a probe branch of an equal-path Michelson interferometer. The reference beams are reflected by a mirror, which is located on a z-stage for controlling the length of the reference path.

The probe beams illuminate the device after passing either through a lens system, or through a lens and a microscope objective. A microscope or the lens has to be chosen to match the requirements in terms of field of view and spatial resolution.

The probe and reference beams are again combined by the beam splitter and split into the orthogonally polarized parts by a second PBSC. Two FPA cameras record the interferograms.

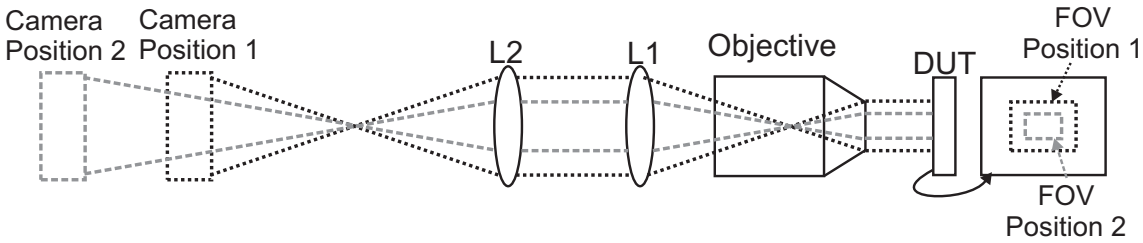


Figure 3.4: Zooming by changing the position of the camera. For clarity beam splitters and reference path are not shown, only one camera is depicted

Field of view and spatial resolution

Different microscope objectives or a lens system can be used to achieve the desired requirements in terms of field of view (FOV) and spatial resolution. The SLD setup can be used with three microscopes, a Mitutoyo 50x, a Mitutoyo 20x and a Mitutoyo 10x microscope and with a lens system.

The spatial resolution is limited by two factors. The first factor is the diffraction limit, which is given as

$$d_{\text{difr}} = \frac{1.22\lambda}{2N_{\text{A}}} \tag{3.5}$$

and depends on the numerical aperture N_{A} of the objective. The numerical aperture of the Mitutoyo 50x microscope is 0.42, which corresponds to a diffraction limit of 1.9µm.

The second limiting factor is the camera pixel size, which also depends on the used microscope and on the position of the camera, as some zooming is possible by changing the camera position. The principle of zooming is depicted in Fig. 3.4. For a larger FOV the camera has to be moved closer to lens L2, for zooming and a higher resolution it has to be moved away.

At the highest zoom the pixel size using the Mitutoyo 50x microscope objective is $1\mu\text{m} \times 1\mu\text{m}$ at a FOV of $250\mu\text{m} \times 300\mu\text{m}$, therefore under large magnifications the limiting factor is the diffraction limit of 1.9µm. The largest possible field of view using the lens system is approximately 2.5mm × 3mm.

Aligning the optical setup

To achieve good results, the optical setup has to be well aligned. The two orthogonally polarized beams have to be overlapping, the device has to be in focus, the optical length of the probe and reference path has to be equal to get interference and the fringes of the interference pattern have to meet certain conditions in terms

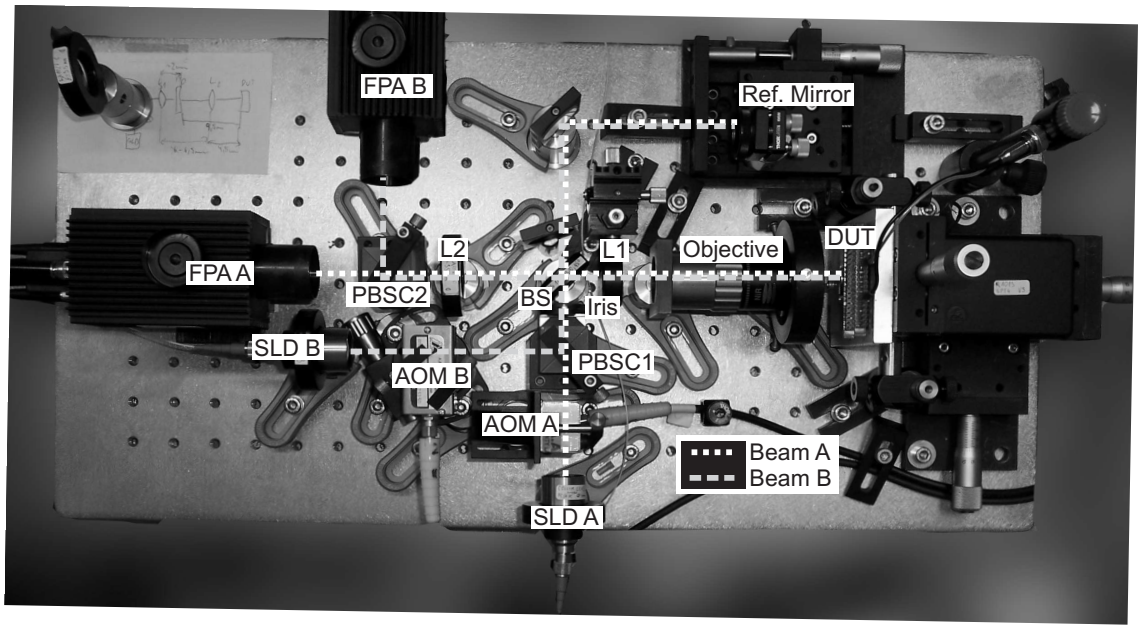


Figure 3.5: Photograph of the SLD setup, microscope objective is installed, abbreviations are: SLD: Superluminescent diode, PBSC: Polarizing beam splitter cube, L: Lens, BS: Beam splitter, FPA: Focal plane array camera, DUT: Device under test

of thickness and orientation. A photograph of the SLD setup, in which the different components and positioners are visible, is shown in Fig. 3.5.

To get overlapping beams the frequency of the AOMs can be adjusted. As AOM A is positioned orthogonally to AOM B, the frequency of AOM A changes the vertical position of beam A, whereas the frequency of AOM B influences the horizontal position of beam B. The overlapping of the beams can be checked for example by diverting the beam in the reference branch with an additional mirror, so it is not reflected back into the interferometer. Then the position of the beams can be checked at a far distance and the AOM frequencies can be set accordingly. Another possibility to check the beam position is to remove PBSC2 and observe both beams on camera A. Now the position of both beams can be corrected together with the inclination of the reference mirror until the intensity on the camera reaches a maximum for both beams.

The distance between lens 1 and lens 3 or a microscope objective has to match the focal distances of the components to achieve a parallel beam illuminating the sample. This can be checked by removing the stage for the DUT and observing the diameter of the beam at different distances to the optical system. If the beam is not

parallel, the position of lens 1 has to be adjusted if a microscope objective is used. If the lens system is used, alternatively the position of lens 3 can be corrected.

For focusing on the device, the reference branch has to be closed, the IR image of the device has to be visible on the camera. The z position of the stage holding the sample has to be moved until a sharp image of the device can be seen on the camera. This is the case if the focal distance of the microscope or lens 3 matches the optical path to the active region of the device.

After the device is in focus the length of the reference path has to be changed by moving the reference mirror until the optical distance of the reference and probe path match. When the paths match, an interference pattern is visible on the camera. However it is also possible that interference from a part of the beam reflected at other layers, e.g. from the device back side, occurs. Interference from the active region can be recognized, as geometrical features of the device influence the interference pattern.

Finally the thickness of the fringes has to be set up accordingly by adjusting the inclination of lens 1. This can be done by adjusting the screws of the positioner holding lens 1. The fringes should have an inclination of 45°. Recommendations for the fringe thickness are given at the end of Section 3.2.2. After having adjusted these screws it might be necessary to refocus on the device, as the optical length of the probe path can be changed slightly.

3.2.2 Phase shift extraction

In the SLD setup interferograms are acquired. To extract the phase shift from these interferograms a FFT method is applied [21, 22]. This method has been developed for the 2D TIM setup with pulsed lasers and a MATLAB software exists for fast processing of the interferograms [23].

The extraction sequence is depicted in Figs. 3.6 and 3.7. During the measurement two interferograms are acquired, a reference interferogram (Fig. 3.6a) and a stress interferogram (Fig. 3.6b). The intensity of each interferogram can be written as

$$i(x, y) = a(x, y) + b(x, y) \cos(\varphi(x, y)) \tag{3.6}$$

$a(x, y)$ is the background intensity, whereas $b(x, y)$ is the intensity of the fringes. In complex form (3.6) can be rewritten as

$$i(x, y) = a(x, y) + c(x, y) + c^*(x, y) \tag{3.7}$$

A two-dimensional Fast-Fourier transform (FFT) is applied to the interferograms, which leads to spectra as depicted in Figs. 3.6c and 3.6d. The FFT of (3.7) is given as

$$I(u, v) = A(u, v) + C(u, v) + C^*(u, v) \tag{3.8}$$

In the depicted spectra the different contributions can be clearly distinguished, $A(u, v)$ as a cross in the center, $C(u, v)$ and $C^*(u, v)$ as crosses in the upper-left and lower-right corner of the spectra.

From these spectra the components of $A(u, v)$ and $C^*(u, v)$ have to be eliminated. This is done by a masking process as depicted in Figs. 3.6e and 3.6f. Additionally also high frequency parts of the spectrum are masked to reduce noise.

The inverse Fast Fourier transform (IFFT) of the remaining spectrum provides a complex function $c(x, y)$ (see Figs. 3.6g and 3.6h). This function allows the retrieval of the phase of each interferogram (see Figs. 3.7a and 3.7b) as

$$\varphi(x, y) = \arctan \frac{\Im(c(x, y))}{\Re(c(x, y))} \tag{3.9}$$

Now the phase of the stress and reference image are subtracted (see Fig. 3.7c).

$$\Delta\varphi'(x, y) = \varphi_{\text{stress}}(x, y) - \varphi_{\text{ref}}(x, y) \tag{3.10}$$

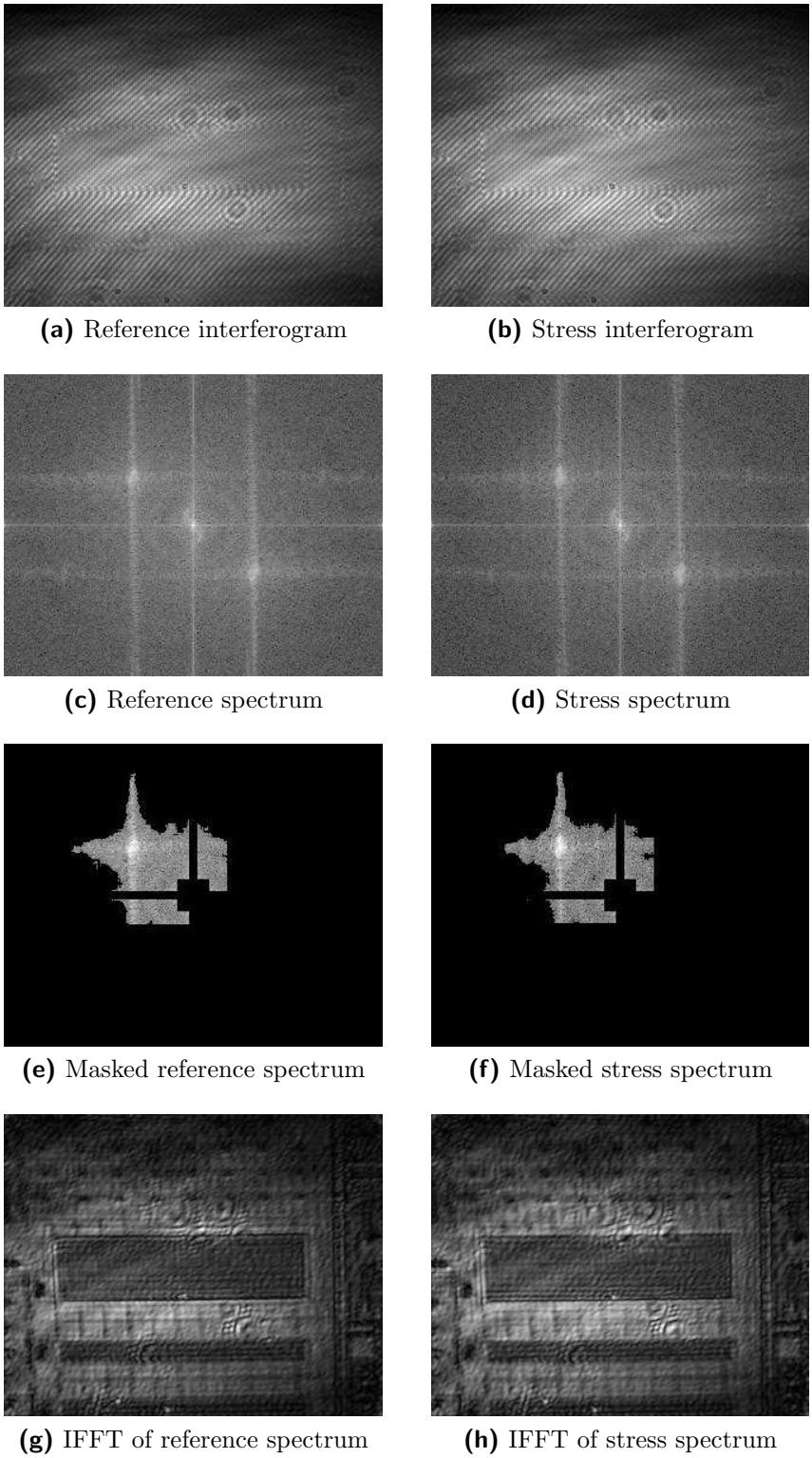
In this way only the phase shift caused by changes of the refractive index during the stress is recorded, while other contributions to the phase, like geometry of the device, cancel each other in the subtraction.

After the subtraction a $n * 2\pi$ uncertainty remains,

$$\Delta\varphi(x, y) = \Delta\varphi'(x, y) + n * 2\pi \tag{3.11}$$

This uncertainty is taken care of by unwrapping processes. In a first automated step neighboring points are set to similar values. This process provides Fig. 3.7d.

The resulting image can already correspond to the correct phase shift, though usually a post-processing is necessary. Steps of the phase have to be unwrapped



(a) Reference interferogram

(b) Stress interferogram

(c) Reference spectrum

(d) Stress spectrum

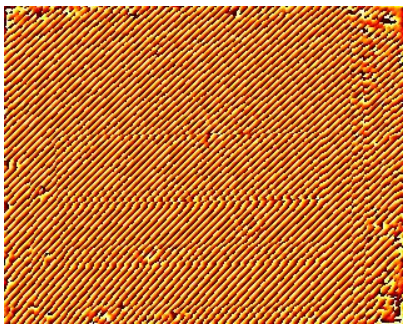
(e) Masked reference spectrum

(f) Masked stress spectrum

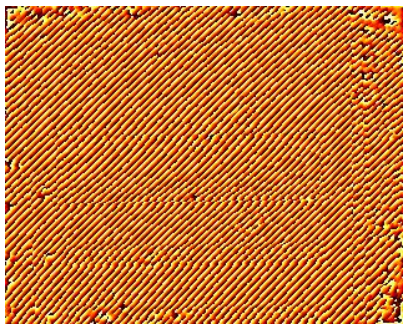
(g) IFFT of reference spectrum

(h) IFFT of stress spectrum

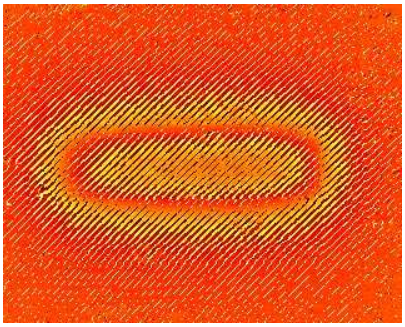
Figure 3.6: Processing of interferograms



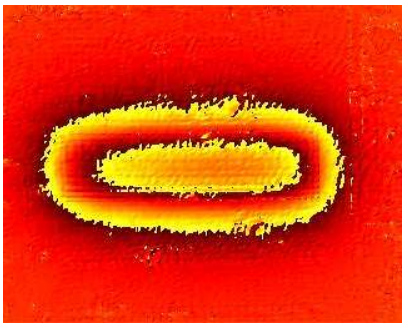
(a) Reference phase



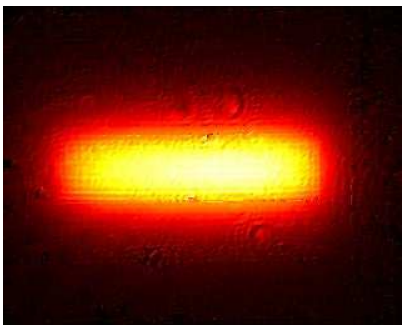
(b) Stress phase



(c) Subtracted wrapped phase

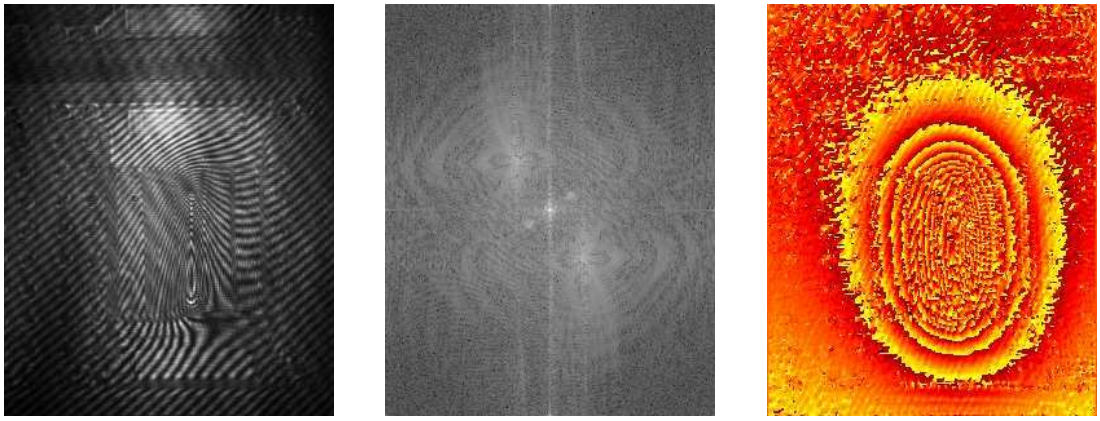


(d) Phase shift after automated unwrapping



(e) Phase shift after manual unwrapping

Figure 3.7: Processing of interferograms (cont.)



(a) Stress interferogram **(b)** Stress spectrum **(c)** Phase shift before unwrapping

Figure 3.8: Extraction of phase at a high phase shift gradient

manually and in some cases, depending on the orientation of the fringes and on which part of the spectrum is utilized, the resulting phase has to be inverted. Finally a phase shift image like Fig. 3.7e is retrieved.

Limitations of phase shift extraction

Problems with the correct extraction of the phase shift occur at high gradients of the phase shift (see Fig. 3.8). In this case two problems occur. Either the distance between neighboring fringes can get too thin, so that two fringes are interpreted as a single fringe, or fringes get wider, until singularities appear. Both of these effects are visible in Fig. 3.8a.

In the spectrum (Fig. 3.8b) too wide fringes lead to an overlapping of $C(u, v)$ with its conjugated complex part $C^*(u, v)$ [22]. Therefore correct filtering is not possible anymore, which leads to incorrect extraction of the phase shift. A condition which guarantees that the overlapping remains minimal is that the wavenumber of the fringes k_{Fi} or equivalently the fringe thickness λ_{Fi} have to fulfill the condition

$$k_{\mathrm{Fi}} = \frac{2\pi}{\lambda_{\mathrm{Fi}}} > \max |\mathrm{grad}\varphi| \tag{3.12}$$

in relation to the maximum phase gradient $\max |\mathrm{grad}\varphi|$. An interpretation of this condition is that the inclination of the reference plane has to be higher than the phase gradient.

The effect of fringes which get too thin is visible in the interferogram but is not directly visible in the spectrum. This effect is caused by aliasing. If the fringe thickness is too thin, the FFT algorithmus interprets the actual fringe frequency as a lower frequency. With the fringe thickness of the stress interferogram $\lambda_{\mathrm{Fi,stress}}$ the sampling criterion can be written as

$$2d < \lambda_{\mathrm{Fi,stress}} \tag{3.13}$$

d is the distance between two points which can be distinguished. This distance can be limited either by the camera pixel size d_{px}, which depends on the used microscope and gives the distance between two pixels in the recorded image or by the diffraction limit d_{difr}, which depends on the numerical aperture of the objective (see (3.5))

$$d = \max\left\{d_{\mathrm{px}}, d_{\mathrm{difr}}\right\} \tag{3.14}$$

$\lambda_{\mathrm{Fi,stress}}$ is connected to original fringe thickness λ_{Fi} and the maximum of the phase gradient by the wavenumber of the different terms as

$$\frac{2\pi}{\lambda_{\mathrm{Fi,stress}}} = k_{\mathrm{Fi,stress}} = k_{\mathrm{Fi}} + \max\left|\mathrm{grad}\varphi\right| = \frac{2\pi}{\lambda_{\mathrm{Fi}}} + \max\left|\mathrm{grad}\varphi\right| \tag{3.15}$$

(3.13) and (3.15) can be combined to form a condition for the minimum fringe thickness as

$$k_{\mathrm{Fi}} = \frac{2\pi}{\lambda_{\mathrm{Fi}}} < \frac{\pi}{d} - \max\left|\mathrm{grad}\varphi\right| \tag{3.16}$$

As both the maximum (3.12) and the minimum fringe thickness (3.16) are governed by the gradient of the phase shift, there exist conditions for which the phase shift cannot be retrieved correctly. In this case a smaller field of view has to be used, as this reduces the value of d.

The fringe thickness, which allows to resolve the highest phase gradient for a given d, can be calculated from (3.12) and (3.16) by calculating the boundary conditions of both inequalities as

$$\max\left|\mathrm{grad}\varphi\right| = k_{\mathrm{Fi}} = \frac{\pi}{d} - \max\left|\mathrm{grad}\varphi\right| \tag{3.17}$$

This leads to a maximum of the resolvable phase shift gradient of

$$\max\left|\mathrm{grad}\varphi\right| = \frac{\pi}{2d} \tag{3.18}$$

	Lens LA1027-C	**Mitutoyo 10x**	**Mitutoyo 20x**	**Mitutoyo 50x**
$d_{\text{difr}}(\mu\text{m})$	2.6	3	2	1.9
$\max\lvert\text{grad}\varphi\rvert\ (\frac{\text{mrad}}{\mu\text{m}})$	604	524	785	827
$\lambda_{\text{Fi}}(\mu\text{m})$	10.4	12	8	7.6

Table 3.1: Conditions for highest resolvable phase shift gradient limited by diffraction

at a fringe thickness of

$$\lambda_{\text{Fi}} = 4d \tag{3.19}$$

In Tab. 3.1 $\max|\text{grad}\varphi|$ and λ_{Fi} for the different objectives and the lens system for the diffraction limited case are given. If the pixel size is larger than the diffraction limit, (3.18) and (3.19) have to be applied to the pixel size, which gives a fringe thickness of $\lambda_{\text{Fi}} = 4\text{Pixel}$ and $\max|\text{grad}\varphi| = 1.6\frac{\text{rad}}{\text{Pixel}}$.

3.3 Scanning heterodyne setup

Measurements on the DMOS transistors were also done on a scanning heterodyne interferometer [17]. This setup allows to measure the evolution of phase shift over time in single points of a device. To get information about the whole device similar measurements have to be done for each point. This is achieved by an x-y stage which moves the sample between the measurements to scan the beam along the sample.

As a stress pulse has to be applied for every single point, measurements on the scanning heterodyne setup are time consuming in comparison to the SLD setup and it is also not suited for destructive measurements. However the phase accuracy of the results is higher and the phase shift is measured over time, whereas in the SLD setup it can be only measured at one or two time instants.

3.3.1 Setup principle

A schematic of the heterodyne scanning setup is shown in Fig. 3.9. A laser diode, emitting at $\lambda = 1.3\mu\text{m}$, is used. The laser beam is split by an acousto-optic modulator (AOM) to multiple refracted beams. Two of those beams, usually the first- and second order refracted beams, are used as probe and reference beams. The AOM causes not only diffraction of the beams, but also changes the frequency of

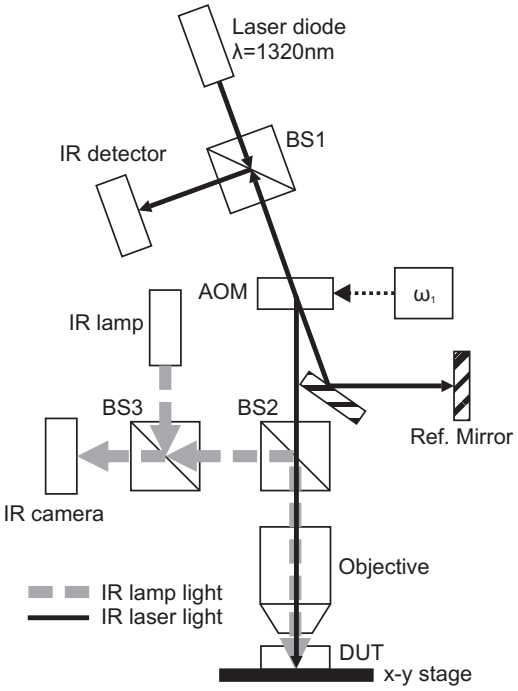


Figure 3.9: Schematic of the heterodyne scanning setup, abbreviations are: AOM: acousto-optic modulator, BS: beam splitter, DUT: device under test

the different beams. The shift of the frequency depends on the acoustic frequency driving the AOM and on the order of the beams.

The reference beam is reflected by a mirror. The probe beam is focused on a single small spot of the device, where the refractive index influences its phase. The reflected beams go back the same way and are combined again by a second passage of the acousto-optic modulator. This combined beam illuminates a detector. The interference of the probe and reference beam causes a signal on the detector of the form

$$s(t) = I_{\mathrm{P}}(t) + I_{\mathrm{R}}(t) + 2 * \sqrt{I_{\mathrm{P}}(t) I_{\mathrm{R}}(t)} * \cos(2\Delta\omega * t + \Delta\varphi(t)) \tag{3.20}$$

In (3.20) I_{R} and I_{P} are the intensities of the reference and the probe beam, $\Delta\omega = \omega_{\mathrm{P}} - \omega_{\mathrm{R}}$ is the beating frequency between the two beams, which has to be multiplied by 2, as the AOM is passed twice and $\Delta\varphi(t) = \varphi_{\mathrm{P}}(t) - \varphi_{\mathrm{R}}(t)$ is the phase shift between probe beam and reference beam over time.

For the mapping of a device, the sample is located on an x-y stage and is moved for measurements to multiple points. In this way line scans along one direction or 2D scans along both direction for measurements in the whole device can be done.

In addition to the optics used for the interferometer an IR lamp and a camera are used to acquire IR images of the device. This makes it possible to locate and focus on the device and to set up the position of the scans.

3.3.2 Phase shift extraction

For the extraction of the phase shift over time from the detector signal a FFT method similar to the method applied for the extraction of phase shift over space for 2D images is used (see Section 3.2.2). (3.20) can be rewritten as

$$s(t) = b_{\mathrm{dc}}(t) + b(t) + b^*(t) \tag{3.21}$$

with

$$b(t) = \frac{1}{2}\sqrt{I_{\mathrm{P}} I_{\mathrm{R}}} * \exp(j(2\Delta\omega * t + \Delta\varphi(t))) \tag{3.22}$$

$b^*(t)$ is the conjugate complex of $b(t)$ and b_{dc} is the DC component of the signal and low frequency signal disturbance, much slower than $2\Delta\omega$.

The Fast Fourier transform of $b(t)$ is calculated as

$$S(u) = B_{\mathrm{dc}}(u) + B(u) + B^*(u) \tag{3.23}$$

This function is bandpass filtered to remove B_{dc} and B^* as well as high frequency noise and low frequency disturbance.

From the inverse Fast Fourier transform $b'(t)$ of this filtered function, the phase shift can be calculated as

$$\Delta\varphi'(t) = \arctan\left(\frac{\Im(b'(t))}{\Re(b'(t))}\right) - 2\Delta\omega * t \tag{3.24}$$

There exists still an $n * 2\pi$ uncertainty in these phase values, so unwrapping methods are applied to get the actual phase shift over time $\Delta\varphi(t)$.

3.4 Deformation of the sample

In measurements under millisecond stress pulses, thermal expansion causes a deformation of the sample [24]. This gives rise to an undesired contribution to the measured phase shift, which has to be eliminated if it is large enough to influence

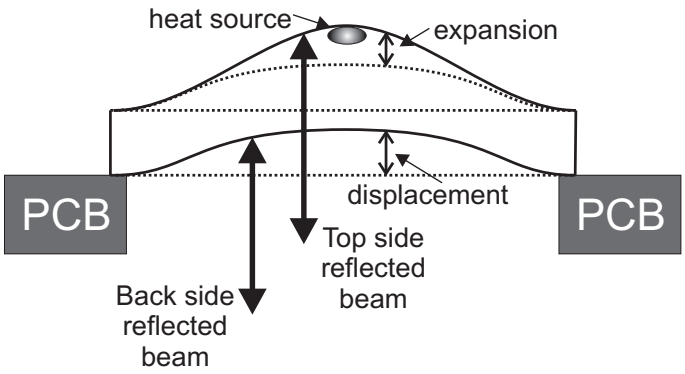


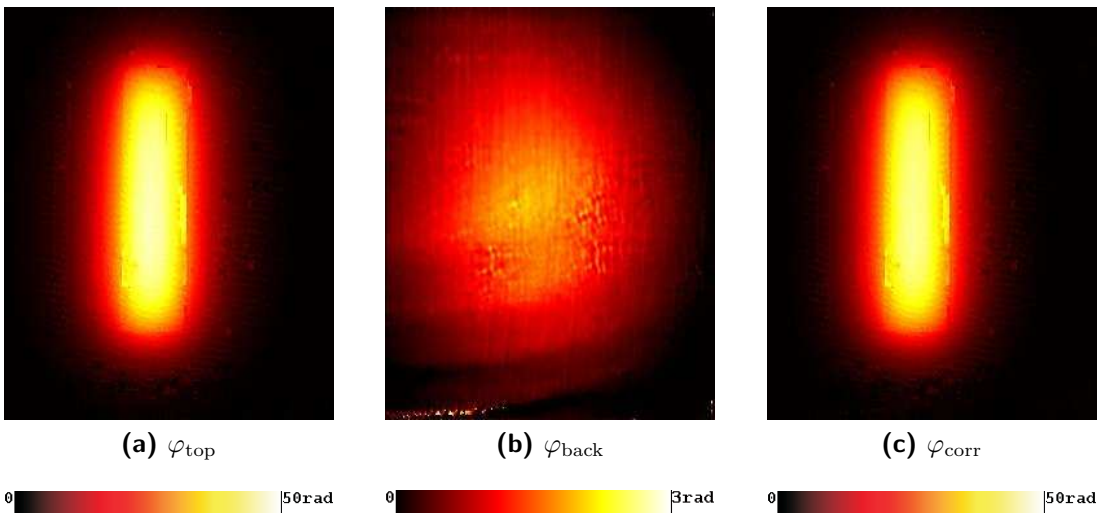
Figure 3.10: Schematic of the sample deformation

the measurement substantially. The mechanisms which contribute to the deformation of the sample are displacement and expansion of the sample (see Fig. 3.10).

Of these two components the contribution by displacement can be determined experimentally and compensated if necessary. First the phase shift is retrieved as usual, with the optical beam focused on the top side surface (φ_{top}). Then a second measurement under similar conditions is done, in which interference from the portion of the beam reflected at the back side surface is recorded. The phase shift determined by this second measurement (φ_{back}) is proportional to the displacement of the sample. Now the corrected phase shift can be calculated by subtracting the two retrieved values

$$\varphi_{\text{corr}} = \varphi_{\text{top}} - \varphi_{\text{back}} \tag{3.25}$$

An example for the correction of a 2D phase image is depicted in Fig.3.11. In this case the contribution of the displacement is relatively small, therefore the correction has small impact on the phase shift image.



(a) φ_{top} **(b)** φ_{back} **(c)** φ_{corr}

Figure 3.11: Correction of thermal displacement

Chapter 4

Control of the 2D SLD setup

In this chapter the instrumentation and software used for control of the 2D SLD setup are described. Based on the existing setup for one camera, the electrical connection scheme and the measurement software were adapted to allow measurements with two cameras. A special focus in this chapter rests on the timing of the stress pulse and image acquisition.

4.1 Electrical setup

The electrical connections of the 2D SLD setup are shown in Fig. 4.1. Control software on a PC is used to set 2 delay generators (DGs), to read out channels from an oscilloscope, to set and read out the FPA cameras and to control the mechanical shutter for the reference path. The delay generators and the oscilloscope are connected to the PC via GPIB (IEEE 488), the FPA cameras are directly connected to the PC by a digital interface. The shutter is controlled by the PC with a Meilhaus Redlab ME-1208 USB Box.

With the help of the delay generators, the timing of the pulse generator, the AOMs and the cameras is controlled. All the outputs of the DGs are set to TTL and 50Ω. As three channels are required for the timing of the stress pulse and the two optical pulses, the two delay generators have to be synchronized. Therefore DG1 is internally triggered and triggers DG2 externally with the AB channel.

The CD channel of DG1 is connected to the pulse generator, which controls the electrical stress of the device. As the input of the used pulse generator is a high impedance input, the cable is terminated by a 50Ω resistor. Depending on the function of the used pulse generator, the CD channel of DG1 is either used only as

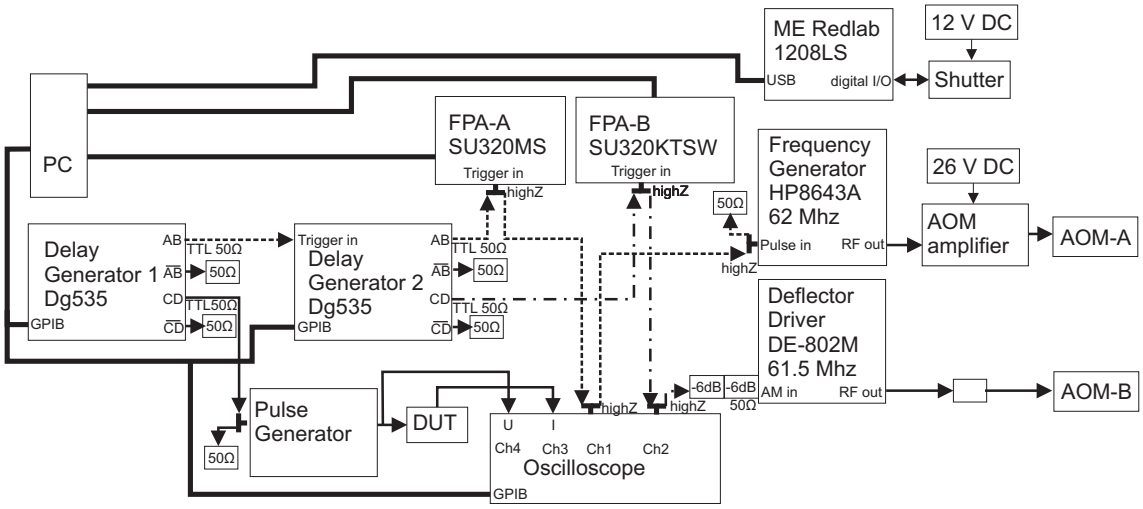


Figure 4.1: Electrical connections for the 2D SLD setup

a trigger for the stress pulse, or it can also determine the length of the stress pulse. If the pulse generator does not support control of the pulse length by an external source, the length has to be set manually on the pulse generator.

The cameras and AOMs are triggered by the AB and CD outputs of DG2. The AB output is used to control AOM A as well as camera A. The cable is connected first to the trigger input of camera A, then to channel 1 of the oscilloscope and finally to an HP8643A frequency generator. As all the inputs are high impedance inputs, a 50Ω resistor terminates the cable. The frequency generator creates a sinusoidal signal of approximately 62MHz. This signal is amplified and serves as input for AOM A.

The CD output of DG2 controls Camera B and AOM B. From the DG the cable leads first to the trigger input of camera B, then to channel 2 of the oscilloscope and finally to the AM input of a DE-802M deflector driver. The signal has to be attenuated by two 6dB attenuators before the deflector driver as the AM input is limited to a peak to peak voltage of 1V. The deflector driver creates a sinusoidal signal of approximately 61.5MHz, which can be directly connected to AOM B as the signal amplitude is sufficient to drive the AOM.

If only one beam is used, the corresponding cable has to be connected directly to the AB channel of DG1 instead of the trigger input of DG2. DG2 is not required in this case.

4.2 Timing

The timing of the image acquisition, namely of the AOMs and the cameras is a critical issue of the setup. As an example, the timing during a single stress pulse of 150µs is given in Fig 4.2. The duration of the optical pulse for beam A and B is set to 20µs, the end of the optical pulses is set to 75µs for beam A and to 150µs for beam B. The depicted signals are discussed from top to bottom.

The top graph gives the signal on the output of channel AB of DG1. This signal is set to high without any delay to the trigger event. The transition to low takes place after the specified pulse duration, which is in this case 150µs. The output of the pulse generator follows the signal of the delay generator without any significant delay. In this case it is assumed that the length of the stress pulse is determined by the trigger signal. If this is not the case the end of the stress pulse could be at a different time instant than the transition of channel AB of DG1.

The timing of camera A and AOM A is controlled by channel AB of DG2. The transition from low to high of this channel is set to happen 9µs before the beginning of the optical pulse. In this case this corresponds to 46µs after the trigger event. The signal stays high for 20µs, the defined optical pulse duration. The HP8634A frequency generator has an internal delay of 1µs, so the output of the frequency generator is slightly delayed to the output of the delay generator.

The frequency generator drives AOM A. The modulation of the optical beam starts 3µs after the RF signal is applied at one side of the beam. After 8µs the central part of the beam is modulated and after 13µs the opposite side of the beam is modulated. The same delays apply when the RF signal is turned off. This leads to an overall delay of $4 - 14$µs between the signal of the delay generator and the beginning of the optical pulse. For camera timing the average value of when the central part of the beam is modulated is used, so in this case the central part of the optical beam is modulated from 55µs to 75µs after the trigger event.

The length of the exposure of camera A is set to be controlled by the width of the trigger signal. However the minimum exposure time of camera A (model: SU320MS) is 98µs. Thus on camera A the exposure time is significantly longer than the optical pulse duration. The camera also has an internal delay of 0.5µs, so the exposure starts 0.5µs after the transition of the delay generator signal. However this delay is lower than the delay of the frequency generator and the AOM, so the whole optical pulse is recorded by the camera. A drawback of the long exposure time is

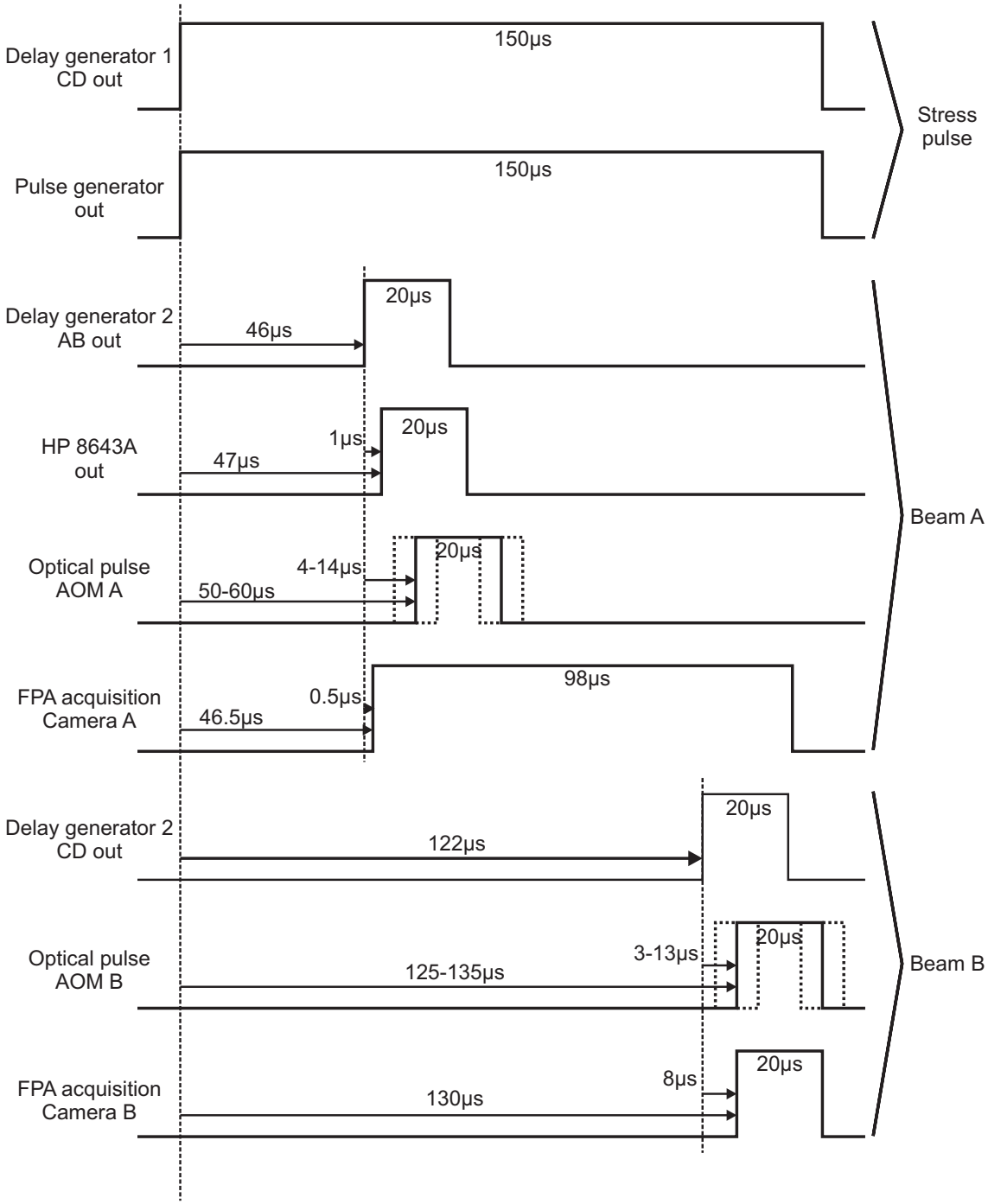


Figure 4.2: Timing for the 2D SLD setup

that more dark current is accumulated at the FPA, which causes higher background noise.

The timing of camera B and AOM B is controlled by channel CD of DG2. The transition to high of this channel happens 8µs before the beginning of the optical pulse, which corresponds to 122µs after the trigger event. The delay of the DE-802M deflector driver is neglectable, so it is not depicted. Thus the modulation of AOM B starts $3 - 13$µs after the transition of the DG signal, so the central part of the optical beam is modulated from 130µs to 150µs after the trigger event.

In contrast to camera A, the newer model used as camera B (SU320KTSW) has a minimum exposure time of only 8.8µs. Therefore the exposure time is controlled by the length of the DG signal. To get the highest possible overlap between the optical pulse and the camera exposure, internal delay of the camera is set to 8µs. This can be done during the setup of the camera. Due to the lower exposure time compared to camera A, camera B records only during the desired interval from 130µs to 150µs and has lower noise than camera A.

4.3 Instrument control

The control software for the 2D SLD setup with 2 cameras [25] was improved, based on an existing software for 1 camera. The software is programmed in LABVIEW, which is well suited for the task of controlling multiple measurement instruments. A target of the software development was to create a program which would work for both cameras as well as for each camera alone should the second camera, or instruments used for controlling the second camera, not be available. When starting the program, one therefore has to select whether one camera or both cameras are to be used.

The program is used for acquisition of IR images as well as for controlling the stress sequence. As these two processes are independent, they are discussed separately. The principles of the control are the same for one camera and for two cameras, so in the following discussion this is not distinguished.

4.3.1 IR image acquisition

In the SLD setup, the light of the superluminescent diodes is used for acquisition of the interferograms as well as for taking infrared images. Therefore the AOMs have

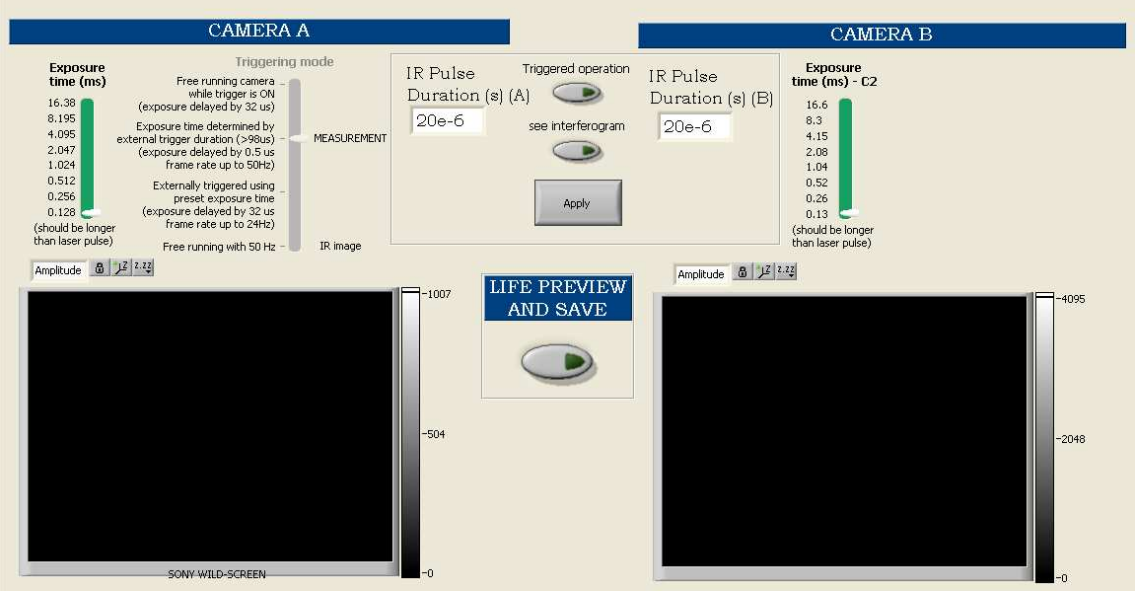


Figure 4.3: Front panel for control of IR image acquisition

to be controlled also for the IR image acquisition, to allow the light from the SLDs to illuminate the device.

The IR image acquisition is used to take back side infrared images without fringes, but also to prepare the setup for a stress sequence, to set up the fringe thickness and contrast.

The front panel of the LABVIEW program for controlling IR image acquisition is shown in Fig. 4.3. It can be controlled whether the device should be illuminated continuously or whether the AOMs and cameras should be triggered. As the beams can illuminate the device only if a signal is applied to the AOMs, the delay generators have to be set accordingly.

In case of continuous illumination, the output signal of the delay generator is set to a duration of 999ms at a frequency of 1Hz, so the signal is not actually continuous, but is not active for 1ms each second. However this is sufficient for the image acquisition, only sometimes a slight flickering of the image is visible on the screen. The camera is set to free running mode. Images on camera A are acquired at a rate of 50Hz, on camera B at a rate of 100Hz. The exposure time of the cameras can be chosen to be set between 130µs and 16ms.

Due to the high intensity of the SLDs, the image can be overexposed in the continuous mode, even at the shortest exposure time. Therefore the triggered mode was introduced, which works similar to the image acquisition during the stress sequence.

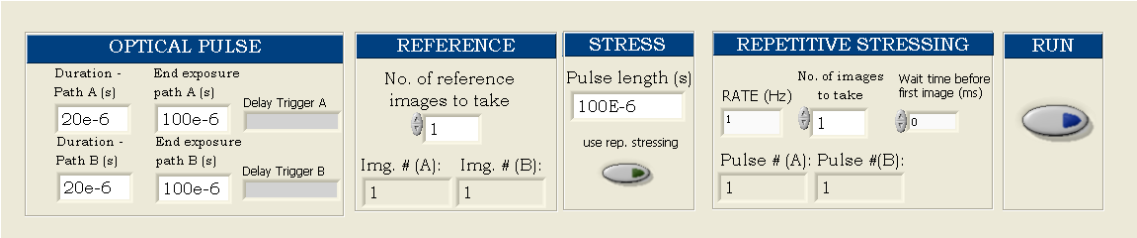


Figure 4.4: Front panel for setup of stress sequence

In this case the camera is triggered and the duration of the optical pulse can be set, usually to a pulse duration between 20µs and 120µs. This leads to reduced optical energy during the exposure time and therefore to a darker image. The brightness can be controlled by the pulse duration. The frame rate is set to 50Hz on both cameras.

To switch between acquisition of IR images without fringes and interferograms, a shutter in the reference path is used. The shutter is controlled by the software. During IR image acquisition one can choose, whether the interferogram is to be shown. If the reference path is closed, no fringes will appear. This mode can be used to get IR images of the device and to focus on the desired region.

With an opened reference path, the contrast and thickness of the fringes can be set up. In this case it is usually necessary to run the image acquisition triggered, as both the reference and the probe beam illuminate the camera. In this way it is also possible to see exactly what the reference interferogram will look like if the optical pulse duration is set to the value which will be used in the stress sequence.

4.3.2 Stress sequence

During a stress sequence the desired electrical stress is applied to the device and reference and stress interferograms as well as the channels of the oscilloscope are recorded. Fig. 4.4 shows the front panel for setup of the stress sequence.

The program flow during a stress sequence is depicted in Fig. 4.5. There are two different modes possible for stressing. The standard mode for single stressing is depicted in black in the image and is explained first. For repetitive stressing, which is the second mode, a few additional steps are introduced, which are grayed out in Fig. 4.5. They are explained in the end of this section.

At the beginning of the stress sequence the delay generators are set up for the reference pictures, the duration of the optical pulses is set to the desired value,

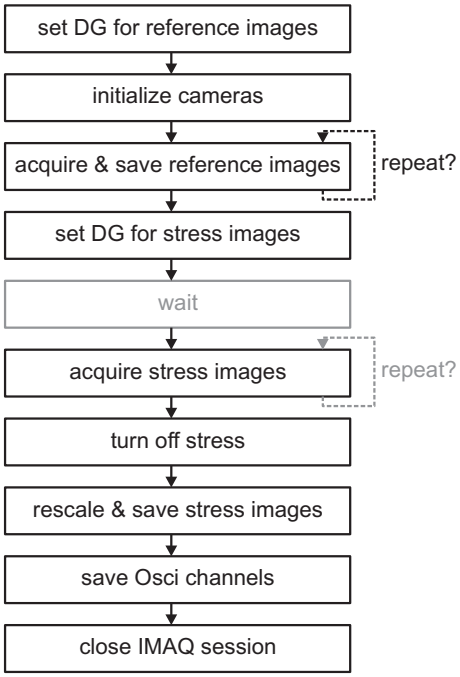


Figure 4.5: Stress sequence

but the stress pulse duration is still set to 0. Then the cameras are initialized and the reference images are taken. It would also be possible to take more than one reference image. While this option was useful in the 2D laser setup, it is actually never necessary in the SLD setup, since pulse to pulse instabilities are reduced substantially.

After reference image acquisition, the DGs are set for the stress images, so the delays of the cameras are set accordingly and the duration of the stress pulse is set. The stress images are acquired and the stress is turned off immediately after image acquisition to allow no additional stress pulses. Only after turning off the stress are the images processed and saved. The channels of the oscilloscope are read out and finally the IMAQ sessions for image acquisition have to be closed.

In addition to the single stressing mode a mode for repetitive stressing was introduced. This mode allows on the one hand the acquisition of multiple stress images under similar conditions and on the other hand the examination of devices which are repetitively stressed for some time, usually a few seconds. In Fig. 4.5 the additional steps during the stress sequence are gray colored.

Fig. 4.6 shows the timing for repetitive image acquisition. A delay can be set between the setting of the delay generator for the stress images and the actual image acquisition. During this delay the device is already repetitively stressed with the

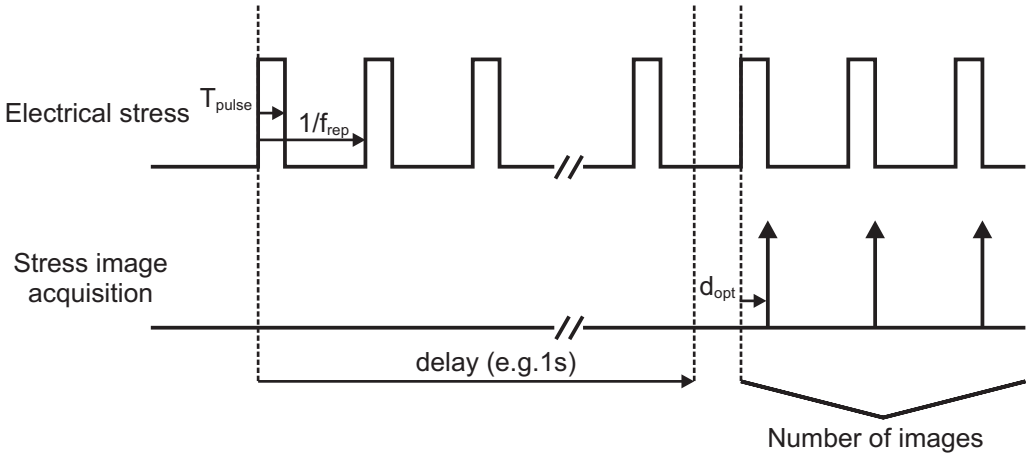


Figure 4.6: Timing for repetitive stressing

desired settings, but the first stress image is only recorded after the delay (e.g. if the frequency of the stress is set to 100Hz, the stress duration to 1ms and the delay to 1s, the device is stress for 1s with pulses of 1ms and 9ms breaks between the single pulses).

After this delay the stress images are acquired. The delay and duration of the image acquisition is set relative to the beginning of a single electrical stress pulse as usual. Furthermore it is also possible to record multiple stress images at the different stress pulses. However in this case there exists a limitation by the maximum frequency of the cameras, which is 50Hz for camera A and 100Hz for camera B. If a higher frequency is chosen, there will not be an image acquisition at every stress pulse.

Chapter 5

Experiment preparation

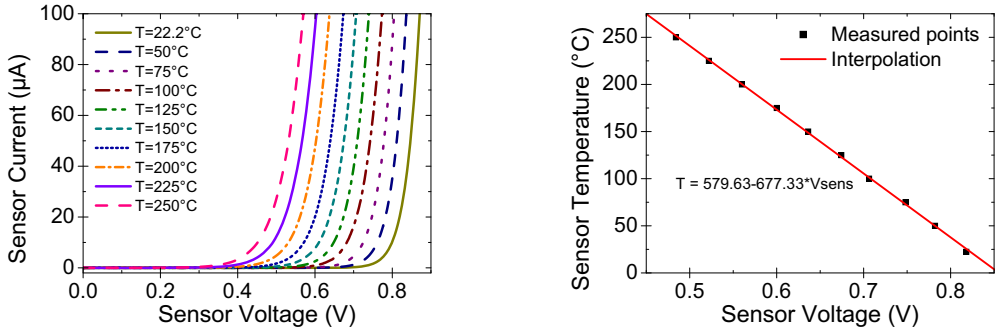
Before the measurements can be performed, the samples have to be prepared. This includes calibration of the temperature sensors and mounting of the devices in a way that allows optical investigations and electrical contacting of the device. The electrical connection scheme for the measurements is also presented in this chapter.

5.1 Sensor calibration

For accurate temperature retrievals a calibration of the sensor has to be performed (see Fig. 5.1). For this purpose the device is mounted on a thermochuck and the I-V characteristic of the diode is measured at multiple temperatures (see Fig. 5.1a). From these curves the voltage values at one current level are plotted over the corresponding temperature. In Fig. 5.1b this is done for a forward current of 20µA. The measured points can be interpolated linearly, which provides a function for the relation between the forward voltage and the sensor temperature at a constant current.

5.2 Mounting of the samples

The investigated devices have to be located on dices of a size of around 1cm × 1cm. These samples are back side polished, to reduce influences of surface roughness on the optical signal. After polishing, the dices are mounted on specially designed printed circuit boards (PCBs). Those PCBs have a hole in the middle. The dice is glued to the PCB in a way that makes the investigated device accessible from both the top side for bonding and the back side for optical investigations. The PCB has



(a) IV-characteristics at multiple temperatures

(b) Calibration curve for a current of 20µA

Figure 5.1: Calibration of a diode used as temperature sensor

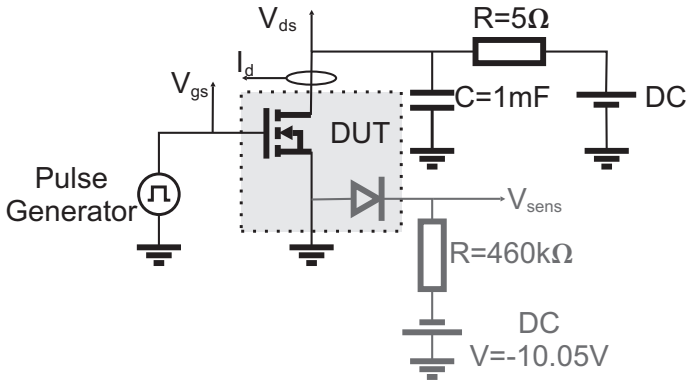


Figure 5.2: Electrical device connection

holes which allow fixing the PCB to a stage by screws. On one side of the PCB a plug can be mounted for electrical contacts.

5.3 Electrical device connection

A similar scheme for the electrical connection is used for all devices. The connections are shown in Fig. 5.2. The connections used for actual stressing of the device are colored black. Additional connections necessary for driving sensor diodes are grayed out.

The drain-source voltage V_{ds} of the transistor is kept at a constant level by a capacity of 1mF which is charged by a DC source over a 5Ω resistor. The capacity of 1mF ensures that for a pulse duration of 1ms and a current of 1A the voltage drop on the drain in not higher than 1V. This is an acceptable condition for the performed measurements. The gate-source voltage V_{gs} is pulsed by a pulse generator. In all

experiments a HP 214B pulse generator was used for this purpose. The HP 214B can create stress pulses up to 100V amplitude and 10ms duration.

V_{ds} and V_{gs} are measured by an oscilloscope with high impedance voltage probes, the drain current I_{d} is measured by an Agilent N2774A current probe connected to an oscilloscope channel. This probe has a bandwidth ranging from DC to 50MHz.

For measurements on the temperature sensor a constant current has to be applied to the sensor diode and the forward voltage is measured. A stable current supply is created by using a DC source and a large resistor connected in series to the diode. The resistance of the diode is small in comparison, therefore the variation of the current can be neglected. V_{sens} is measured with a high impedance voltage probe.

Chapter 6

Results

In this chapter measurements made on different DMOS devices under stress pulses in the millisecond range are presented. In the first section, a relatively small vertical DMOS transistor, named IFX R2016_06 is investigated. In these measurements the performance of the SLD setup is also analyzed and enhanced. Using the heterodyne scanning setup, the 2-dimensional power dissipation in the device us estimated.

Due to the effective cooling of the small device destruction was not possible, so a larger area VDMOS (IFX R2012) is examined in the thermally unstable regime, where triggering of the parasitic bipolar transistor and thereby destruction of the device takes place. These processes are investigated with the SLD setup.

Furthermore the SLD setup is used to compare the phase distribution in a large area lateral NDMOS to a lateral PDMOS (IFX R2053) of the same type.

If measurements on the SLD setup are made, in many cases only the cross sections in important regions of the device are depicted and not the 2D images which these cross sections are derived from. In the cross sections the values of the phase shift are visible much better and comparison of different measurements is easier.

6.1 Small VDMOS

An IR image of the small DMOS device (IFX R2016_06) taken in the SLD setup is shown in Fig. 6.1. The coordinate system used in this section is indicated. The device consists of 5 small fingers. The overall size of the active region of the device is approximately 43µm × 36µm. In some of the investigated devices a temperature sensor is located in the center of the active region.

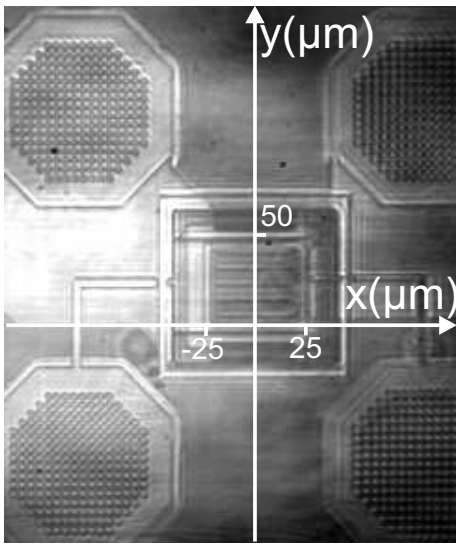


Figure 6.1: Backside IR image of small DMOS device, the µm-scale is retrieved approximately from the backside IR image using the known size of the bondpads

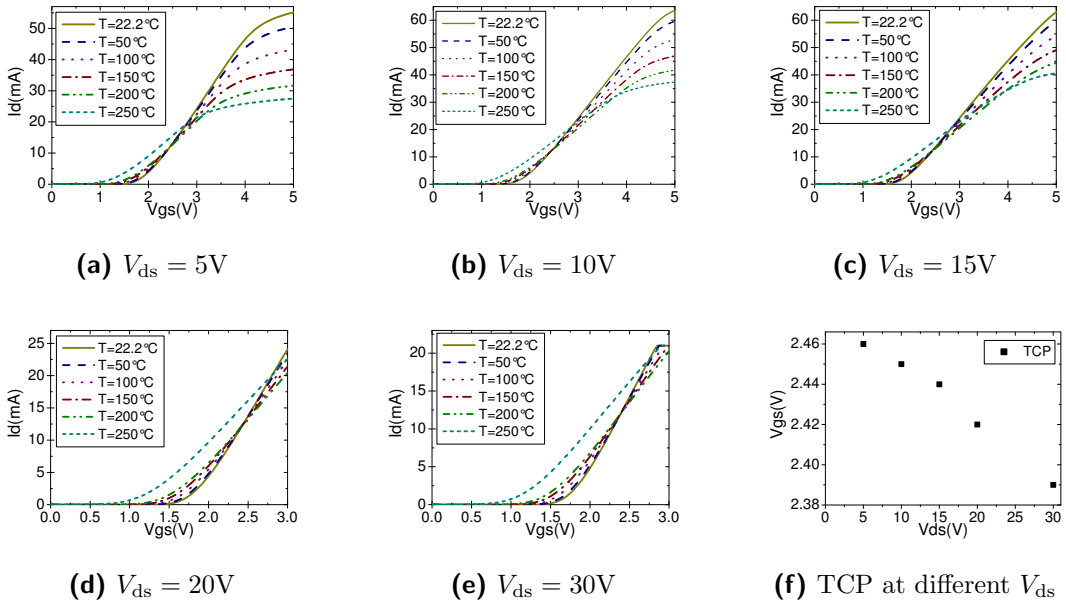
6.1.1 Determination of the TCP

The TCP of the vertical DMOS transistor is determined by measurements on a thermochuck. Transfer characteristics of the transistor are recorded at different drain voltages and temperatures, so the TCP is retrieved at multiple drain voltages between 5V and 30V (see Fig. 6.2).

Due to a problem with the needles contacting the device, the first transistor was destroyed between the measurements at 200°C and 250°C, so the curves at 250°C were measured on a different device, which accounts for the bad fit of those curves to the curves at lower temperatures. In Fig. 6.2f V_{gs}(TCP) retrieved from the other graphs in Fig. 6.2 at different drain voltages is depicted. All the values are between 2.39V and 2.46V, so for the next measurements the TCP on these devices was assumed to be at $V_{\mathrm{gs}} = 2.4V$.

6.1.2 TIM measurements

Measurements on the SLD setup are made to estimate the phase distribution under µs to ms pulses and to characterize the setup. The device is analyzed at and above TCP.



(a) $V_{\mathrm{ds}} = 5\mathrm{V}$ **(b)** $V_{\mathrm{ds}} = 10\mathrm{V}$ **(c)** $V_{\mathrm{ds}} = 15\mathrm{V}$

(d) $V_{\mathrm{ds}} = 20\mathrm{V}$ **(e)** $V_{\mathrm{ds}} = 30\mathrm{V}$ **(f)** TCP at different V_{ds}

Figure 6.2: Estimation of the TCP

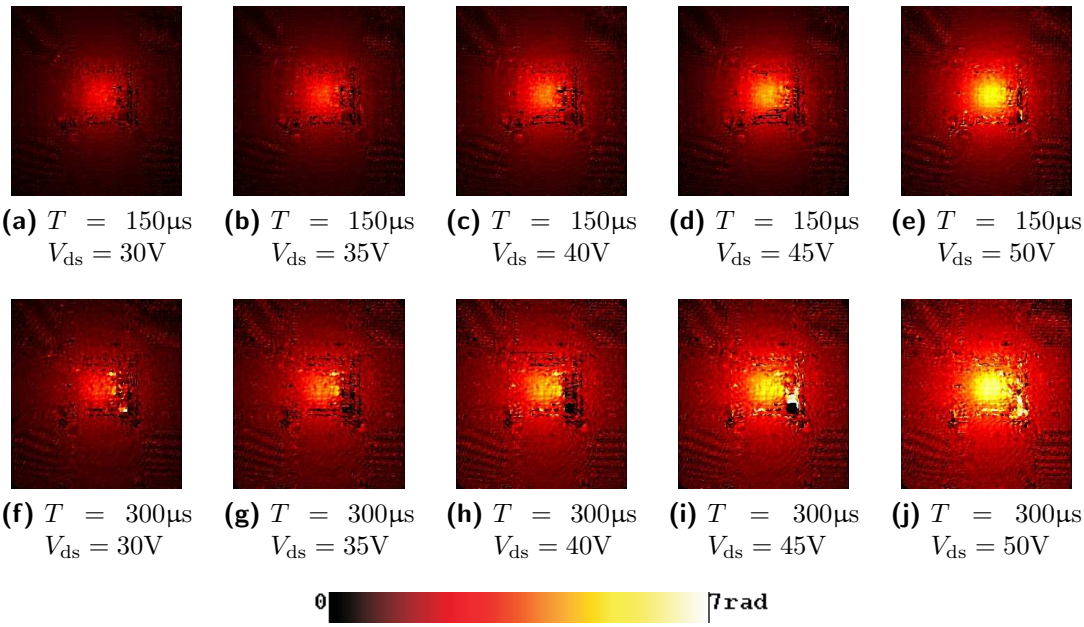
Measurements at TCP

Measurements at TCP are done at five different levels for the drain voltage. While V_{gs} is kept constant at 2.4V, V_{ds} is set to 30V, 35V, 40V, 45V and 50V. At TCP I_{d} is slightly below 15mA which corresponds to a power dissipation between 450mW and 750mW for the different drain voltage levels. Stress images are taken with camera B at 150µs and with camera A at 300µs (see Fig. 6.3.)

Measurements above TCP

Phase shift images above TCP are shown in Fig. 6.4. These images are taken at $V_{\mathrm{gs}} = 5\mathrm{V}$ and drain voltages of $V_{\mathrm{ds}} = 30\mathrm{V}$, $V_{\mathrm{ds}} = 40\mathrm{V}$ and $V_{\mathrm{ds}} = 50\mathrm{V}$. Measurements are made at two time instants during one stress pulse. The images at camera A are recorded at 1ms, the images on camera B at 2ms.

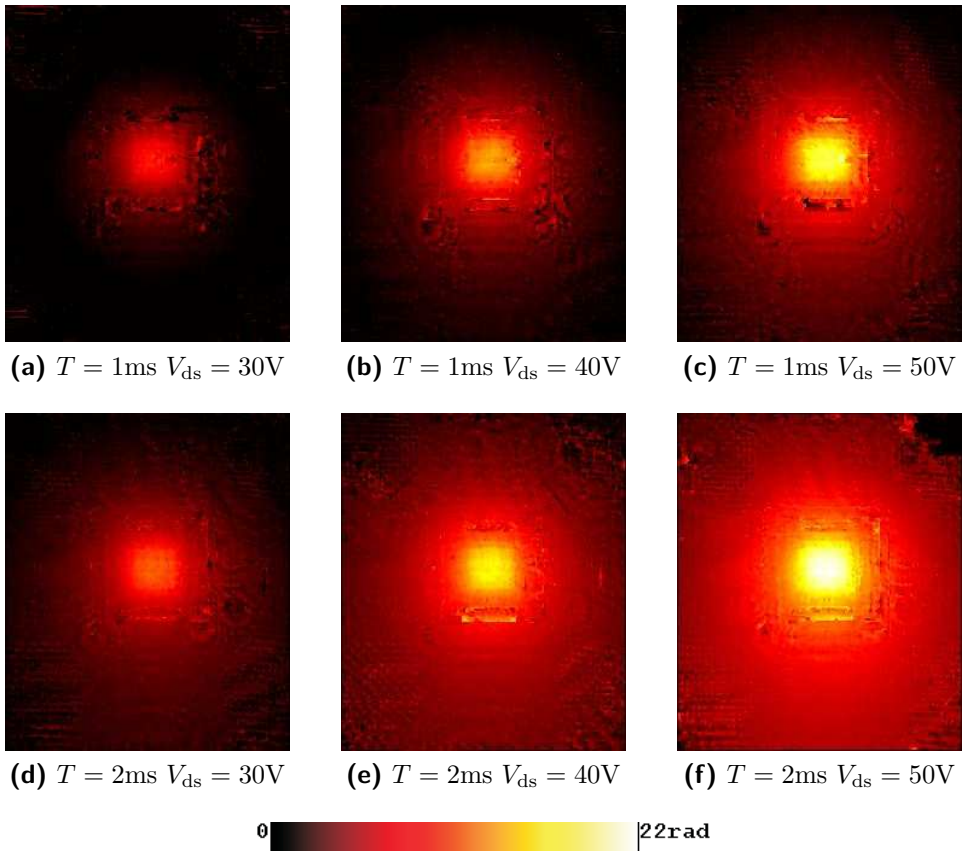
These measurements are performed at high power, V_{ds} is already close to breakdown and V_{gs} is set to have maximum I_{d}. Even under these condition, the device is not destroyed and sustains even longer pulse durations up to 10ms. This is due to the relatively small size of the device, which allows good lateral heat removal from the active region.



(a) $T = 150\mu s$ $V_{ds} = 30V$ **(b)** $T = 150\mu s$ $V_{ds} = 35V$ **(c)** $T = 150\mu s$ $V_{ds} = 40V$ **(d)** $T = 150\mu s$ $V_{ds} = 45V$ **(e)** $T = 150\mu s$ $V_{ds} = 50V$

(f) $T = 300\mu s$ $V_{ds} = 30V$ **(g)** $T = 300\mu s$ $V_{ds} = 35V$ **(h)** $T = 300\mu s$ $V_{ds} = 40V$ **(i)** $T = 300\mu s$ $V_{ds} = 45V$ **(j)** $T = 300\mu s$ $V_{ds} = 50V$

Figure 6.3: Phase shift images on small VDMOS at TCP



(a) $T = 1ms$ $V_{ds} = 30V$ **(b)** $T = 1ms$ $V_{ds} = 40V$ **(c)** $T = 1ms$ $V_{ds} = 50V$

(d) $T = 2ms$ $V_{ds} = 30V$ **(e)** $T = 2ms$ $V_{ds} = 40V$ **(f)** $T = 2ms$ $V_{ds} = 50V$

Figure 6.4: Phase shift images on small VDMOS above TCP

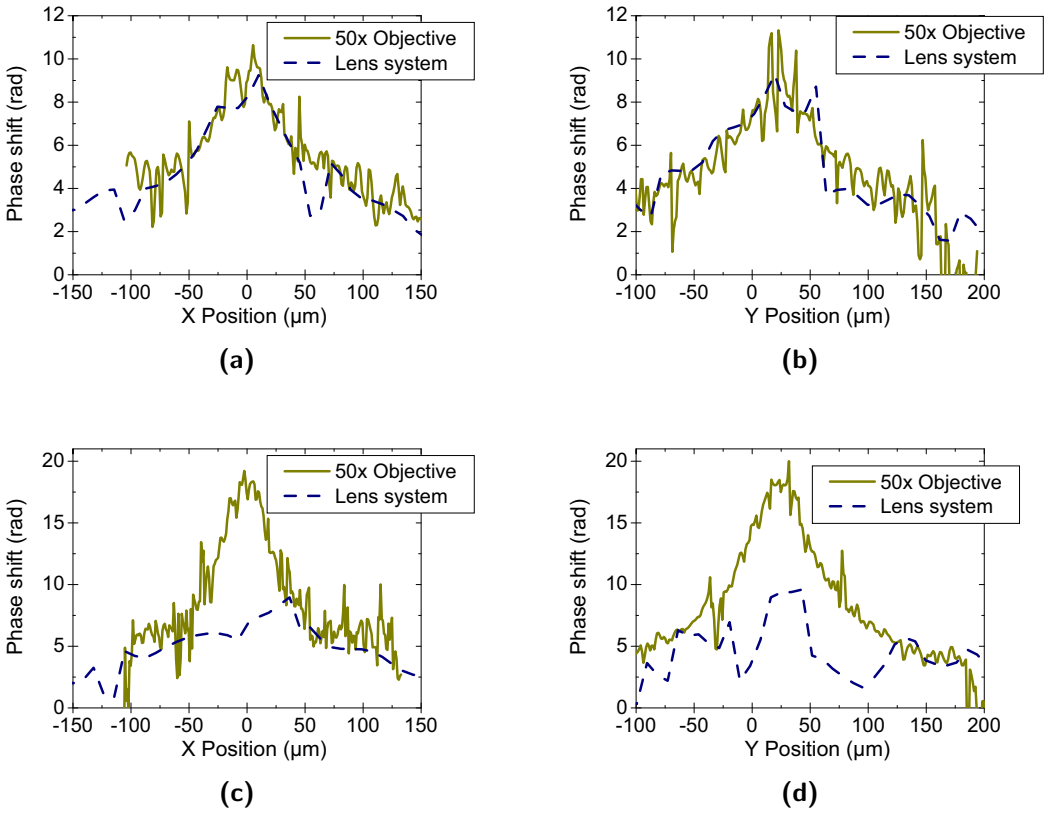


Figure 6.5: Comparison of measurements on lens system to microscope system

Comparison of lens and microscope system

To see the effect of the spatial resolution of the setup, similar measurements are made with the lens system and with the 50x-microscope objective in the SLD setup at two different conditions (see Fig. 6.5).

At a low power of $P = 1.14$W and a long pulse duration of 1.5ms (see Figs. 6.5a and 6.5b) the lens and microscope system provide matching results. In this case the effect of the pixel size and the diffraction limit of the lens system is small enough to resolve the phase shift gradient correctly. The only visible effect is the different distance between neighboring measured points.

At higher power $P = 2.48$W and a pulse duration of 400µs the effects of the spatial resolution can be seen clearly (see Figs. 6.5c and 6.5d). The gradient of the phase shift is too high to be retrieved correctly with the lens system, unwrapping is not possible anymore. This results in lower absolute values of the phase shift in the active part of the device.

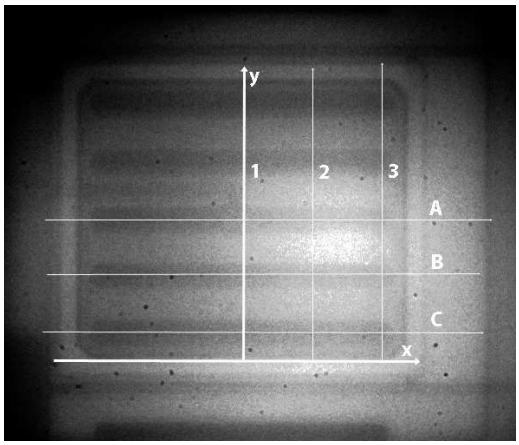


Figure 6.6: IR image and position of scans

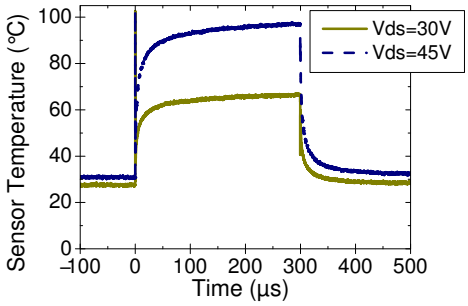


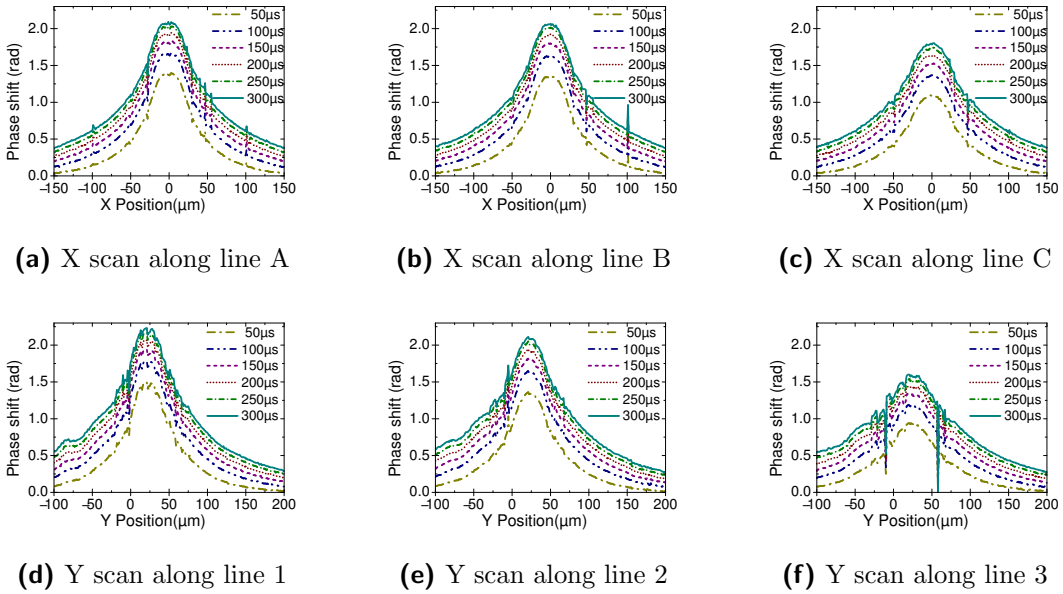
Figure 6.7: Temperature on a built-in temperature sensor

6.1.3 Scanning measurements

Measurements with the heterodyne scanning setup are made to estimate the 2-dimensional power dissipation (P_{2D}) in the device. Fig. 6.6 shows a backside IR image of the device acquired with an immersion oil objective and the position of the line scans.

The measurement is made at temperature compensation point, as this should lead to homogeneous power dissipation in the device, independent of the local temperature. So $V_{gs} = 2.4$V, the length of the pulses is $T = 300\mu$s and the scans are done at two different drain-source voltages of $V_{ds} = 30$V and $V_{ds} = 45$V. At $V_{ds} = 30$V the drain current is $I_d = 12$mA, which corresponds to $P = 360$mW, at $V_{ds} = 45$V the drain current is $I_d = 14.5$mA, so the power is $P = 650$mW.

Additionally the temperature on a built-in sensor located in the center of the device is measured. The retrieved temperatures for the 2 voltage levels are given in Fig. 6.7.



(a) X scan along line A **(b)** X scan along line B **(c)** X scan along line C

(d) Y scan along line 1 **(e)** Y scan along line 2 **(f)** Y scan along line 3

Figure 6.8: Scanning measurements on small VDMOS device, scan positions are given in Fig. 6.6. $V_{\mathrm{gs}} = 2.4\mathrm{V}$, $V_{\mathrm{ds}} = 30\mathrm{V}$, $P = 360\mathrm{mW}$

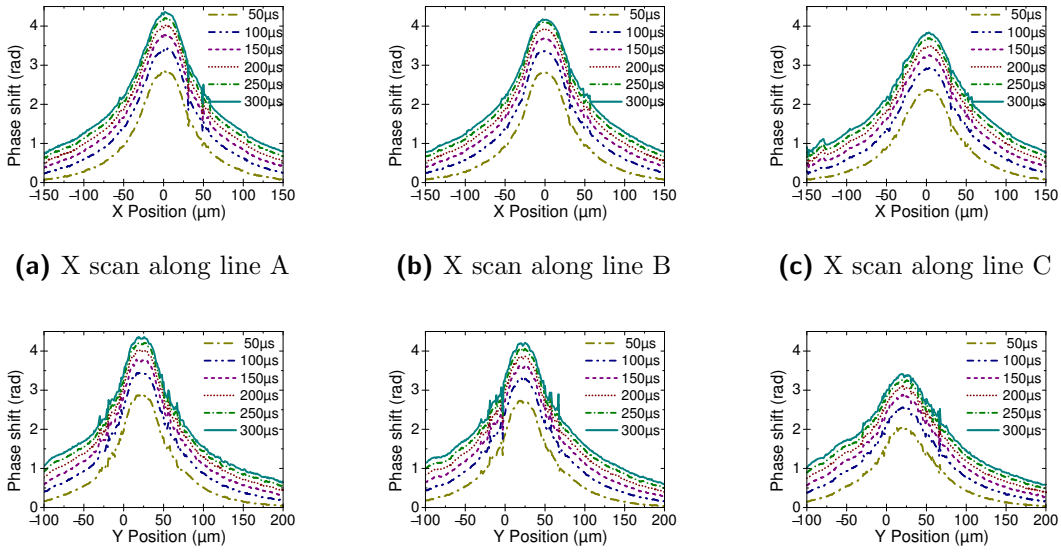
Fig. 6.8 shows the extracted phase shift at different time instants for $P = 360\mathrm{mW}$ and Fig. 6.9 shows the phase shift for $P = 650\mathrm{mW}$.

Comparison to SLD measurements

The measurements on the scanning heterodyne interferometer are compared to SLD measurements performed under similar conditions to verify the phase shift values retrieved with the SLD setup. Cross sections of the 2D phase shift images are compared to the scans (see Fig. 6.10).

The SLD measurements are made with two cameras during a 300µs stress pulse. The image on camera B is taken at 150µs, the image on camera A at 300µs. Gate voltage is set to TCP, $V_{\mathrm{gs}} = 2.4\mathrm{V}$, drain voltage is set to $V_{\mathrm{ds}} = 45\mathrm{V}$, corresponding to $P = 650\mathrm{mW}$.

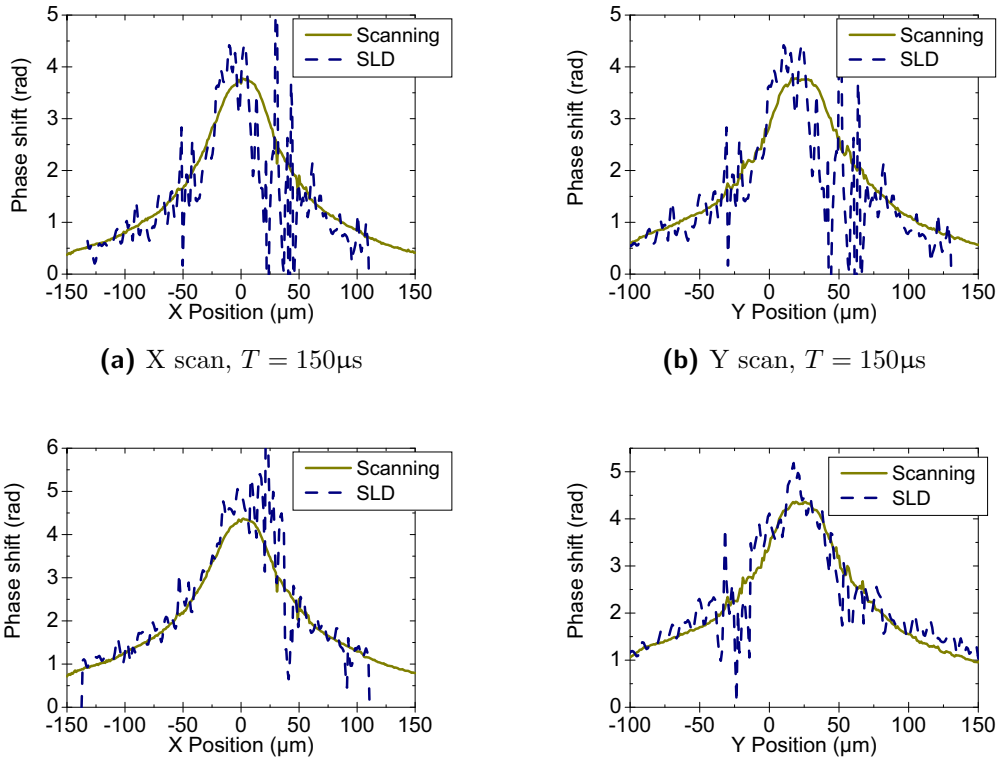
The curves from the SLD measurements match the scanning measurements, though the noise is quite high. This is due to the relatively low amplitude of the phase and due to optical distortions, which are quite high in this measurement. Still the values of the SLD and the scanning measurements are in the same range.



(a) X scan along line A **(b)** X scan along line B **(c)** X scan along line C

(d) Y scan along line 1 **(e)** Y scan along line 2 **(f)** Y scan along line 3

Figure 6.9: Scanning measurements on small VDMOS device, scan positions are given in Fig. 6.6. $V_{\mathrm{gs}} = 2.4\mathrm{V}$, $V_{\mathrm{ds}} = 45\mathrm{V}$, $P = 650\mathrm{mW}$



(a) X scan, $T = 150\mu\mathrm{s}$ **(b)** Y scan, $T = 150\mu\mathrm{s}$

(c) X scan, $T = 300\mu\mathrm{s}$ **(d)** Y scan, $T = 300\mu\mathrm{s}$

Figure 6.10: Comparison of measurements on SLD setup to scanning measurements

Calculation of 2D power dissipation

P_{2D} is estimated from the scanning measurements. The formula for the calculation of P_{2D} is given in (3.4). For the calculation of P_{2D} the first-order derivative of the phase shift in time and the second-order derivative in space along the x- and y-axis have to be known. Therefore P_{2D} can be calculated only at the crossing points of the x-scans and y-scans, where all these components are available.

As the measured data suffers from noise, which has a strong impact on the derivatives, Savitzky-Golay filtering is applied to the data before the derivatives are calculated. The result of the P_{2D} calculations for $V_{gs} = 2.4\text{V}$, $V_{ds} = 45\text{V}$, $P = 650\text{mW}$ and a pulse length of $T = 300\mu\text{s}$ in the 9 crossing points of the line scans are given in Fig. 6.11. In addition to the overall P_{2D}, the contributions of the single space and time-derivatives are given.

As the power dissipation in the device should be homogeneous at TCP, an expected value of the P_{2D} can be calculated as

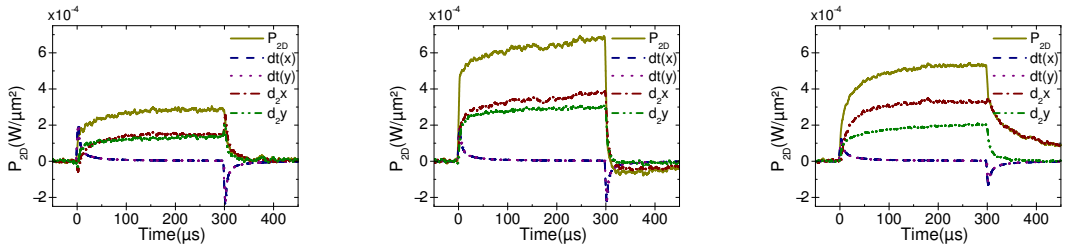
$$P_{2D} = \frac{P}{A} \tag{6.1}$$

where P is the total power dissipated in the device and A is the active area. At a power of 650mW and an approximate active area of $43\mu\text{m} \times 36\mu\text{m}$ this corresponds to $P_{2D} = 4.2 \times 10^{-4} \frac{\text{W}}{\mu\text{m}^2}$.

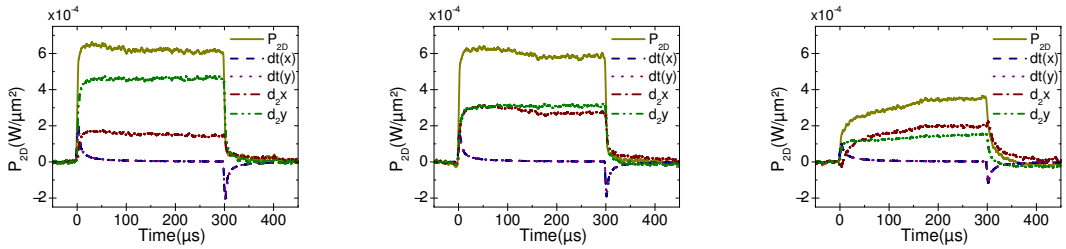
The values retrieved for P_{2D} in Fig. 6.11 are in the range of $2 - 8 \times 10^{-4} \frac{\text{W}}{\mu\text{m}^2}$. It can be seen that the noise in the contribution of the time derivative is low and also that the values for this contribution from the x-scans and y-scans match well. The uncertainty of the results is caused mainly by the contribution of the second-order space derivative and is due to the high impact of noise on higher order derivatives. Therefore for accurate retrievals of P_{2D} under similar conditions, a better signal to noise ratio is necessary.

6.2 Destructive measurements on large VDMOS

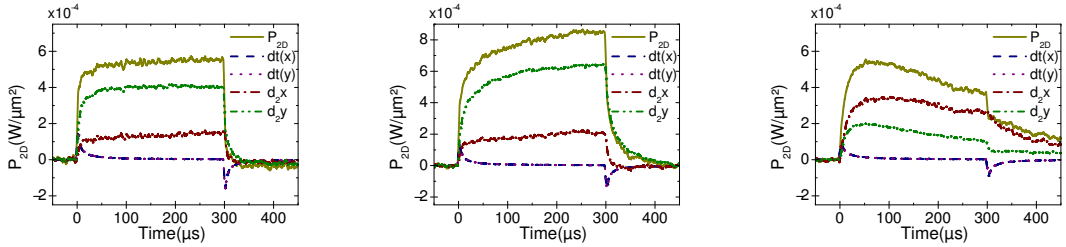
A large area VDMOS device (IFX R2012) is analyzed at conditions below temperature compensation point (TCP) [26]. This device consists of 5 fingers which have common source and drain but individual contacts for the gate. In the experiments only the central finger is activated, the gate contacts of the other fingers are grounded. The size of one finger is approximately $1\text{mm} \times 200\mu\text{m}$. Some of the



(a) A1 ($x = 0\mu m$, $y = 23\mu m$) **(b)** A2 ($x = 13\mu m$, $y = 23\mu m$) **(c)** A3 ($x = 24\mu m$, $y = 23\mu m$)



(d) B1 ($x = 0\mu m$, $y = 13\mu m$) **(e)** B2 ($x = 13\mu m$, $y = 13\mu m$) **(f)** B3 ($x = 24\mu m$, $y = 13\mu m$)



(g) C1 ($x = 0\mu m$, $y = 7\mu m$) **(h)** C2 ($x = 13\mu m$, $y = 7\mu m$) **(i)** C3 ($x = 24\mu m$, $y = 7\mu m$)

Figure 6.11: P_{2d} retrieved from scanning measurements

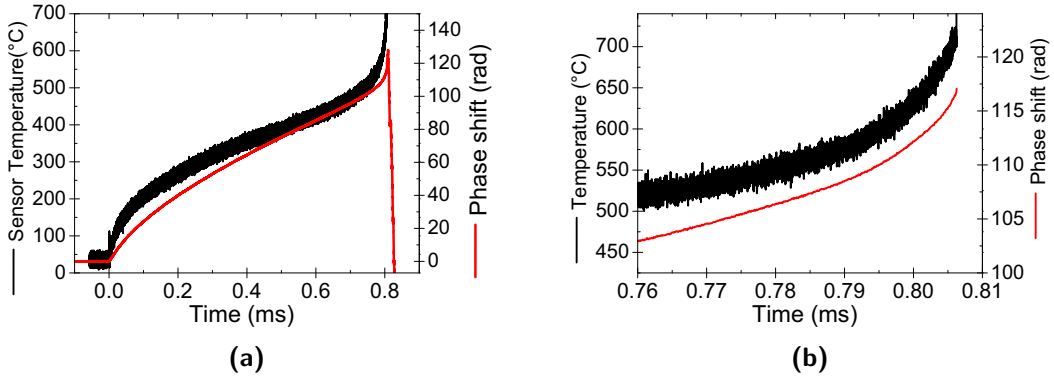


Figure 6.12: Sensor temperature and phase shift (a) over the whole pulse (b) shortly before destruction of the device

devices contain temperature sensors. One of these sensors is located exactly in the center of the device and is used in the measurements.

Measurements are performed with the heterodyne setup and with the SLD setup. With the heterodyne setup is was possible to see the evolution of the phase over time during a destructive pulse in one point in the center of the device. However, the size of the region where the the parasitic bipolar transistor becomes activated can be only observed by 2D imaging. This was done by the observation of destructive pulses with the 2D SLD setup. The heterodyne scanning setup was additionally used to calibrate the absolute phase values in the SLD measurements.

6.2.1 Transient during a destructive pulse

Measurements in the heterodyne setup were performed during a destructive pulse for a single point. For this purpose the laser is positioned a few µm away from the center, so the phase shift is not measured exactly at the position of the temperature sensor, but in the active region. V_{ds} is set to 50V, $V_{\mathrm{gs}} = 2.2\mathrm{V}$. The pulse duration is set to a long duration of 1.5ms to make sure destruction occurs during the pulse. Temperature is measured at a sensor located exactly in the center of the device.

As the stress is applied, destruction happens around 800µs after the beginning of the pulse. The sensor temperature and the recorded phase shift are depicted in Fig. 6.12. The temperature increase due to positive thermo-electric feedback is clearly visible. Therefore also the phase shift shows an increasing slope before destruction.

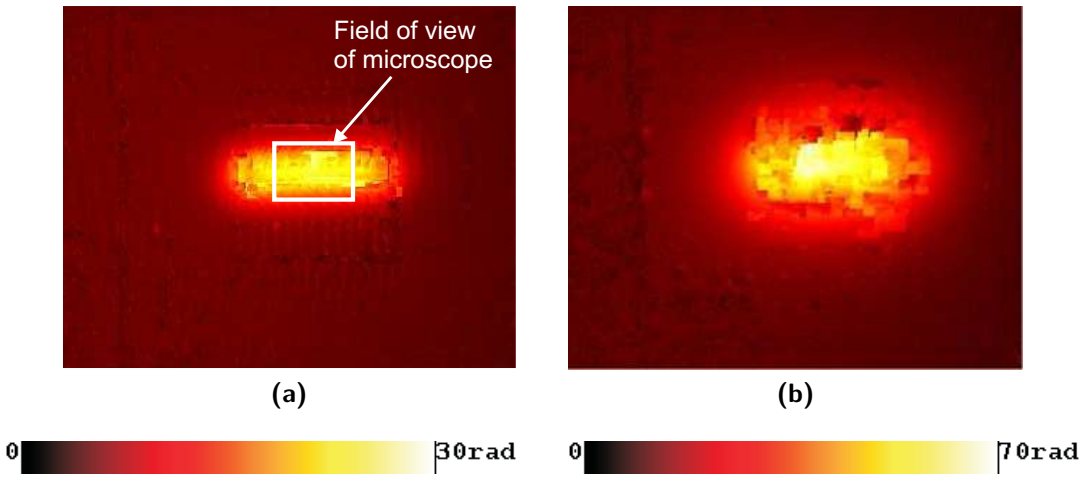


Figure 6.13: Phase shift images of large area VDMOS taken with the lens system, the field of view (FOV) of the 50x microscope objective is indicated

The shape of the phase shift and temperature transients immediately before destruction is depicted in more detail in Fig. 6.12b. At destruction of the transistor, the sensor is also destroyed and the reflectivity of the sample is lost, therefore no information about the temperature after destruction is available.

6.2.2 Destructive measurements on the SLD setup

Measurements on the SLD setup were made to get experimental information about the size of the area, where the parasitic bipolar transistor is triggered. The phase shift in the device is first measured with the lens system under a large field of view (see Fig. 6.13). It is found that the spatial resolution of the lens system is too small for a correct estimation of the high gradient of the phase shift. The unwrapping of the image is still satisfying in the case of a 200µs stress pulse (see Fig. 6.13a), but for a stress pulse of 900µs the uncertainty of the retrieved phase shift is very high due to unwrapping problems (see Fig. 6.13b).

As the target of the measurements was to resolve the region of thermal runaway, further detailed investigations are performed under larger magnification with the 50x objective. V_{ds} is set to 50V, the gate is pulsed at a voltage of $V_{\mathrm{gs}} = 2.2\mathrm{V}$. The pulse duration is increased until destruction takes place. Recorded waveforms of the drain current and the gate voltage during the destructive pulse are shown in Fig. 6.14. Furthermore the periods when the stress images on both cameras are taken during this pulse are depicted.

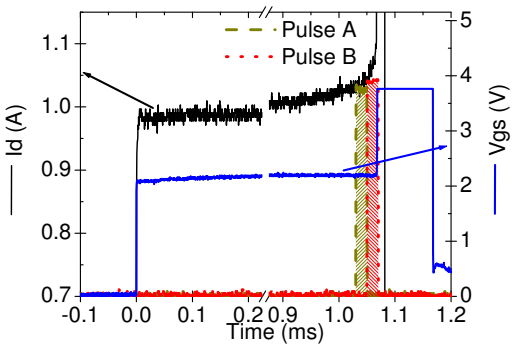
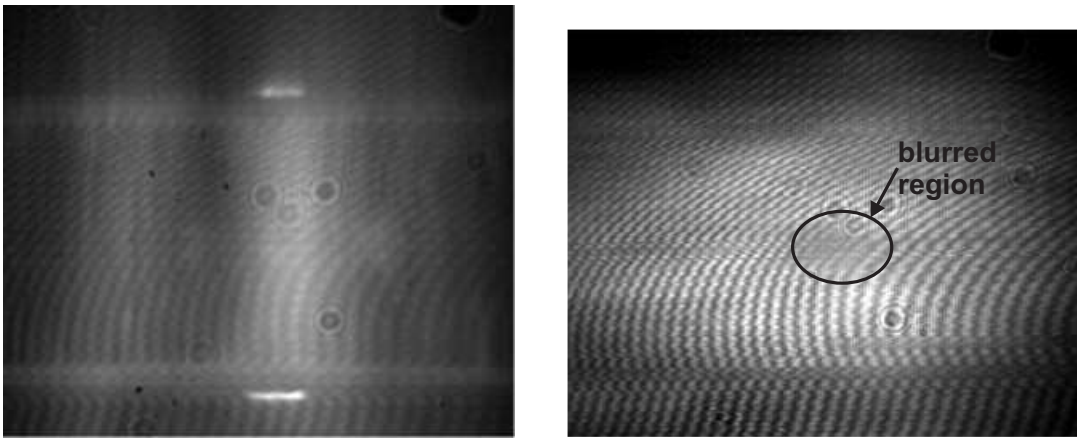


Figure 6.14: Osci channels during destructive pulse



(a) Camera A, 1040µs **(b)** Camera B, 1060µs

Figure 6.15: Stress interferograms taken during a destructive pulse

Bright spots on camera A

Fig. 6.15 shows the stress interferograms on both cameras taken during the destructive pulse. In the image taken with camera A (Fig. 6.15a) a pronounced feature are two bright spots on both sides of the active finger. These spots are also visible in other measurements but always only during destructive pulses and only on camera A, never on camera B. Therefore is appears most likely that these spots are not caused by the optical beam of the SLD, but by some other light source at a later time instant, during the 98µs exposure time of camera A.

During and after destruction of the device, very high currents flow through the device and also neighboring, originally inactive fingers are destroyed. So these spots could be caused by light emission due to high current or temperature during the spreading of the destruction to neighboring fingers. The position of the spots corre-

sponds to the area between the contact of the n-sinker and the p-well of the inactive fingers. At this position a high electric field is applied, as the p-well is connected to the gate of the inactive fingers, which is grounded, and the n-sinker is at the potential of the drain, which is biased at 50V.

Blurring in the interferogram

In Fig. 6.15b the stress interferogram on camera B can be seen. This image is acquired shortly before destruction of the device. In the center of the interferogram a blurred region can be seen. This blurring probably occurs due to a rapid change in the phase shift and partial fringe intensity averaging during the 20µs-long optical pulse or due to absorption by free carriers. So it is possible to extract the size of the thermal runaway area even from the interferogram.

Phase shift during destruction

Phase shift images at different pulse lengths, up to destruction, can be seen in Fig. 6.16. All these images are taken with camera B, as extraction of the phase shift from the images on camera A close to destruction is not possible accurately due to the mentioned bright spots in the interferogram.

Due to an absence of phase reference in the zoomed image (i.e. the whole zoomed area is heated) there arises a $N \times 2\pi$ uncertainty in the absolute phase value in the image ($N = 1, 2, \ldots$). Therefore the extracted phase profile maxima are calibrated using measurements with the heterodyne interferometer recorded in the central part of the device. Due to different thermal conditions caused by the sample mounting and a high sensitivity of the time to failure on the exact settings of gate and drain voltage, destruction in the SLD setup occurred at a different time instant than in the measurements on the heterodyne setup discussed in the previous section.

Therefore the calibration is done by extrapolation of a non-destructive pulse (see Fig. 6.17). The accuracy of this calibration is lower for later time instants close to destruction, as the actual maximum phase shift is higher than the prediction due to the onset of thermal instability. This is especially true for the measurement at $T = 1060$µs. In this case a constant phase shift was added to get higher phase shift than in measurements at earlier time instants in all regions of the device.

In Fig. 6.16d increased phase shift is visible in the center of the device, but as it can hardly be recognized in the 2D images, cross sections are done in the central

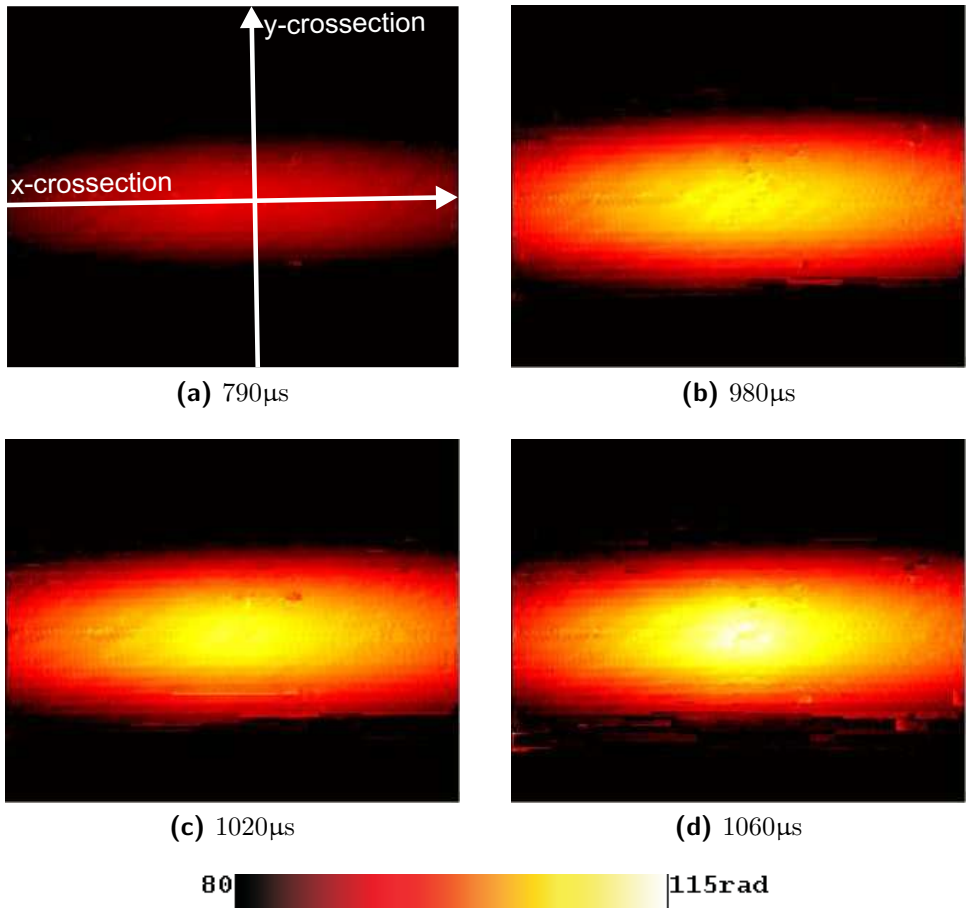


Figure 6.16: Phase shift images at different time instants, scale starts from 80 rad to emphasize central region, the arrows in (a) indicate the position of the cross sections in Fig. 6.18.

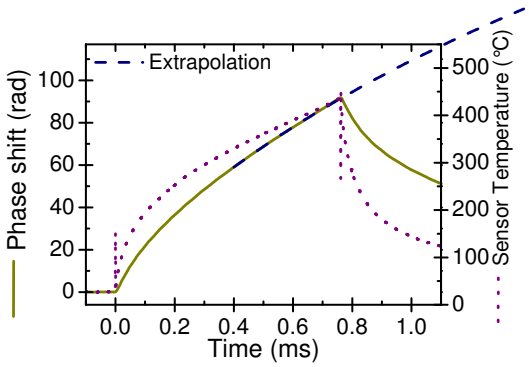
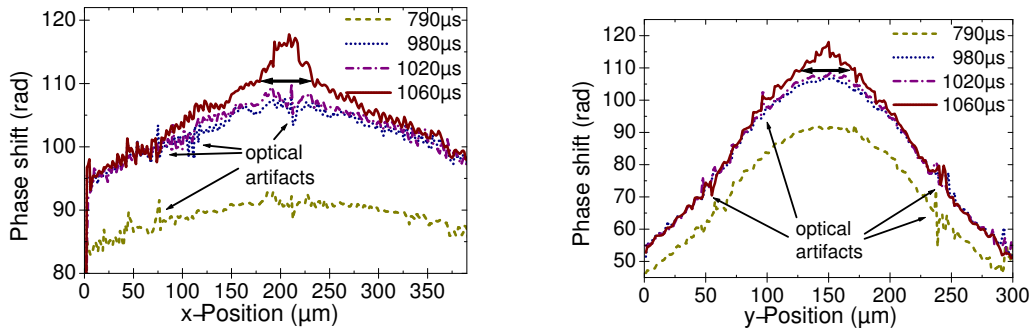


Figure 6.17: Phase shift from scanning measurements, extrapolation of phase shift for longer pulse durations, temperature retrieved from built-in temperature sensor



(a) X-Cross section **(b)** Y-Cross section

Figure 6.18: Phase profiles extracted from Figs. 6.16a-6.16d, the region with increased phase shift is indicated by double arrows

part of all images in Fig. 6.16 along the horizontal and vertical axis. These cross sections are depicted in Fig. 6.18.

A pronounced increase of the phase shift is visible in the central region of the cross sections at 1060µs. The size of this thermal runaway area is approximately 40µm × 50µm, which is important information for thermal model calibration [5].

Final destruction signature

As soon as thermal runaway occurs the device is not controllable by the gate potential anymore, leading to a catastrophic destruction. The current flows until the on-chip metalization or bond wires are destroyed. As a result the damaged area is finally larger than the area in which the parasitic bipolar transistor is triggered, it even extends to neighboring, originally inactive finger blocks. In Fig. 6.19 the backside IR image after destruction is shown, Fig. 6.20 shows a front side visible light photograph, a damage in the power metal is visible.

6.3 Comparison of NLDMOS and PLDMOS

Measurements on large area lateral DMOS transistors are performed to compare the phase shift profile in a n-channel LDMOS (IFX R2053_NLD60V) to a p-channel LDMOS transistor (IFX R2053_PLDM) of the same area and technology. Tests at the manufacturer showed that destruction on these devices under milisecond stress did not take place at the center of the device, as would be expected, but at a place a few hundred µm away from the center (see Fig. 6.21). The devices are tested under

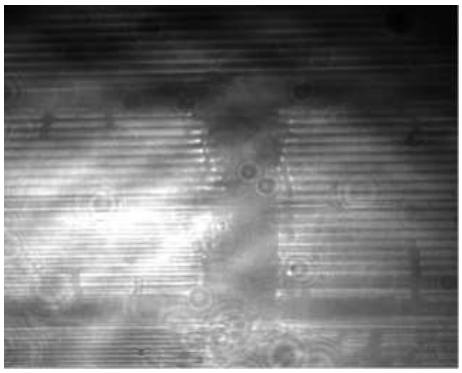


Figure 6.19: Back side IR image after destruction

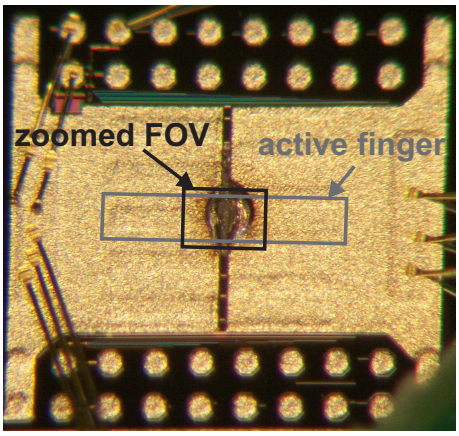


Figure 6.20: Front side visible light photograph after destruction, field of view (FOV) of the 50x objective and the region of the active finger are indicated

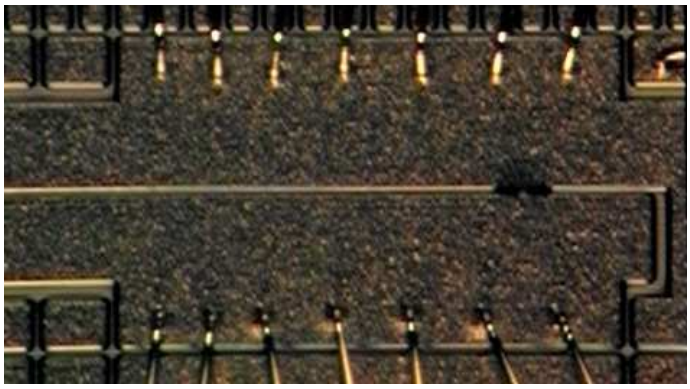
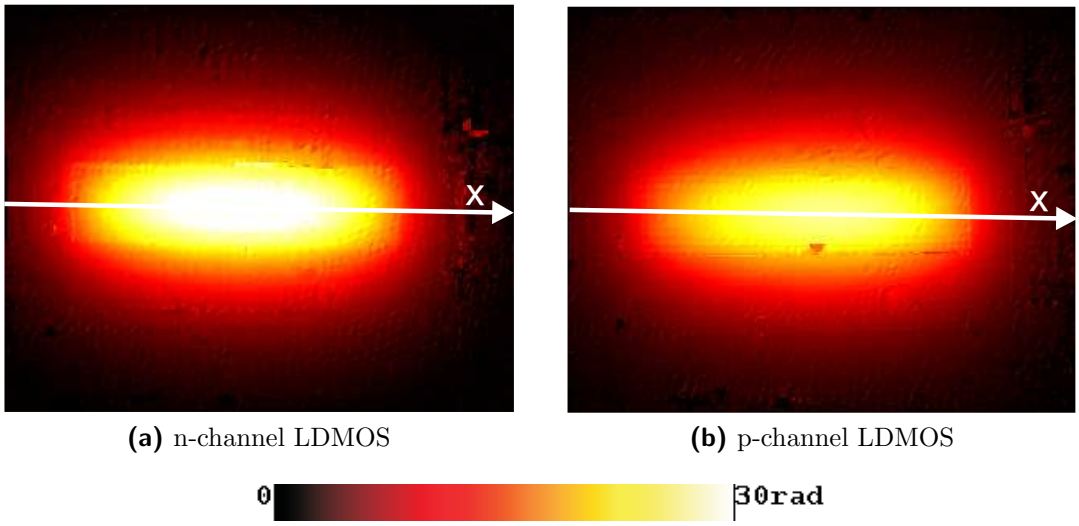


Figure 6.21: Destruction signature of LDMOS transistor



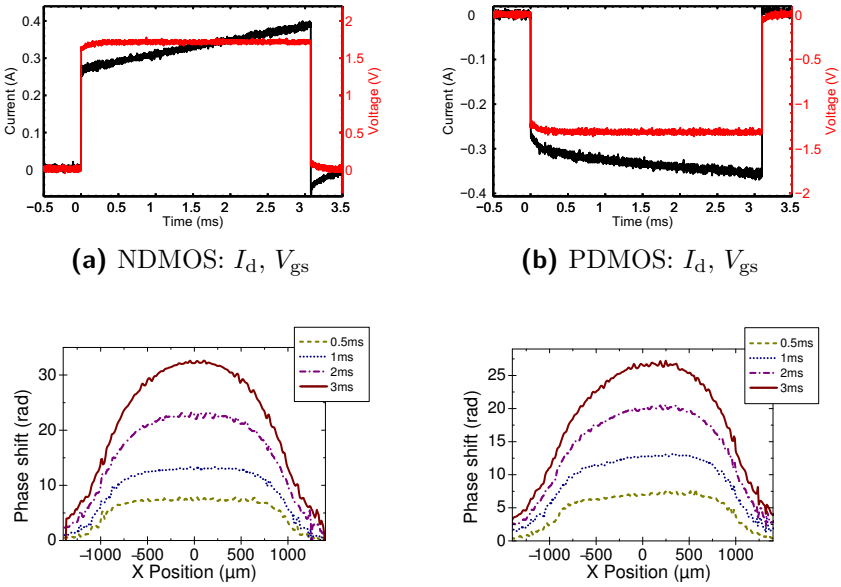
(a) n-channel LDMOS **(b)** p-channel LDMOS

Figure 6.22: 2D phase images at $T = 3$ms, $V_{ds} = \pm 50$V and (a) $V_{gs} = 1.7$V, (b)$V_{gs} = -1.3$V. The arrows indicate the position of the cross sections in the following images

short circuit conditions at pulse durations ranging from 100µs to 3ms. Measurements are performed at different gate voltages below and above TCP.

The SLD setup is used for the measurements. In Fig. 6.22 examples for retrieved 2D phase shift images are depicted. In the following, however, only cross sections along the device will be shown, as the phase shift values can be compared best in this way and as the phase shift distribution along the x-axis of the device is the main interest of this study.

To achieve comparable values on the NDMOS and PDMOS, drain voltage is set to an equal value on the device (i.e. positive on NDMOS, negative on PDMOS), whereas the gate voltage is set to get approximately equal drain current. Due to a



(a) NDMOS: I_d, V_gs **(b)** PDMOS: I_d, V_gs

(c) NDMOS: Cross sections **(d)** PDMOS: Cross sections

Figure 6.23: Waveforms and phase shift profiles below TCP, $V_\mathrm{ds} = \pm 50\mathrm{V}$, $V_\mathrm{gs}(\mathrm{NDM}) = 1.7\mathrm{V}$, $V_\mathrm{gs}(\mathrm{PDM}) = -1.3\mathrm{V}$

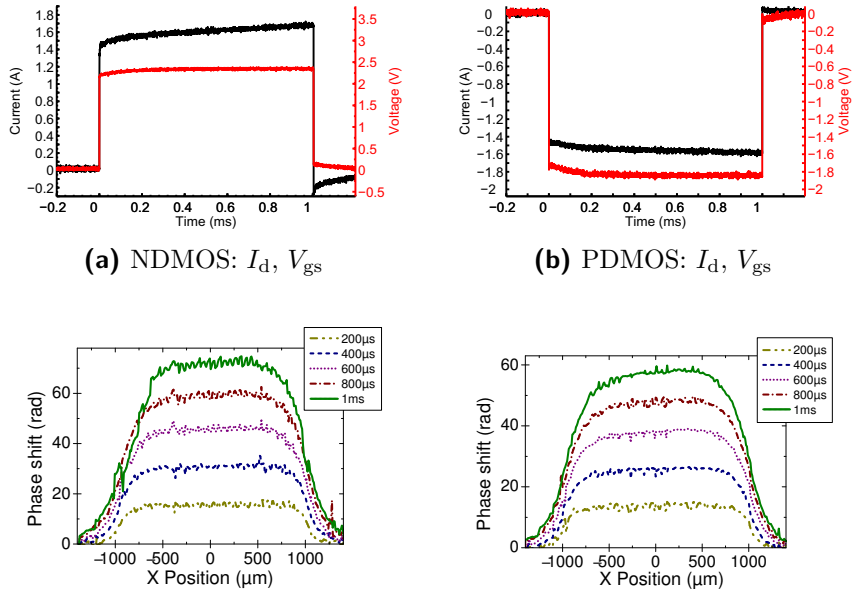
different temperature dependence of the drain current, this is not exactly possible, but still the shape of the phase shift profiles can be well compared.

Fig. 6.23 shows measurements relatively far below TCP. V_ds is set to $\pm 50\mathrm{V}$, $V_\mathrm{gs} = 1.7\mathrm{V}$ on the NDMOS and -1.3V on the PDMOS. Figs. 6.23a and 6.23b show the waveforms of I_d and V_gs during the pulse, Figs. 6.23c and 6.23d the corresponding phase shift at different time instants.

In these measurements the phase shift profile of the NDMOS looks nearly homogeneous, whereas a pronounced shift of the maximum in the PDMOS towards the right side of the device can be seen. Compared to the following measurements, the power in the device is relatively low, as V_ds could not be increased further.

In Fig. 6.24 measurements at higher gate voltages, but still below TCP, are shown. V_ds is set again to $\pm 50\mathrm{V}$, $V_\mathrm{gs} = 2.3\mathrm{V}$ on the NLDMOS and $-1.8\mathrm{V}$ on the PLDMOS. The current increase with temperature is higher in the NLDMOS, thus the maximum phase shift is higher at later time instants, as V_gs is set to have equal I_d at the start of the pulse.

In Figs. 6.24c and 6.24d inhomogeneous phase distribution can be seen in both NDMOS and PDMOS, but the shift of the maximum appears to be more pronounced on the PDMOS.



(a) NDMOS: I_{d}, V_{gs}

(b) PDMOS: I_{d}, V_{gs}

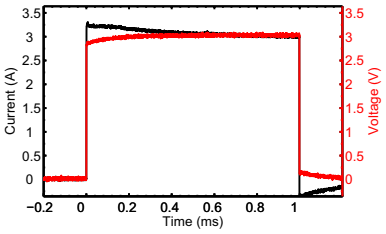
(c) NDMOS: Cross sections

(d) PDMOS: Cross sections

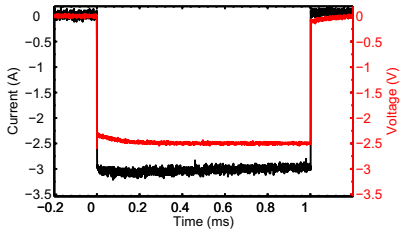
Figure 6.24: Waveforms and phase shift profiles slightly below TCP, $V_{\mathrm{ds}} = \pm 50\mathrm{V}$, $V_{\mathrm{gs}}(\mathrm{NDM}) = 2.3\mathrm{V}$, $V_{\mathrm{gs}}(\mathrm{PDM}) = -1.8\mathrm{V}$

Investigations above TCP are made at two different conditions. Measurements are made close to the TCP at $V_{\mathrm{ds}} = \pm 35\mathrm{V}$ and $V_{\mathrm{gs}} = 3\mathrm{V}$ on the NDMOS and $V_{gs} = -2.5V$ on the PDMOS (see Fig. 6.25) and far above TCP at $V_{\mathrm{ds}} = \pm 15\mathrm{V}$ and $V_{\mathrm{gs}} = 4.5V$ on the NDMOS and $V_{gs} = -3.9V$ on the PDMOS (see Fig. 6.26).

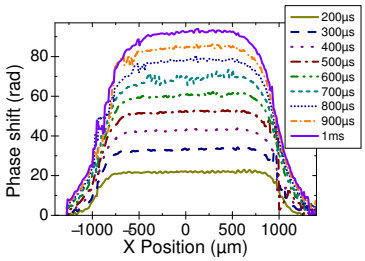
Similar to the measurements slightly below TCP, the phase shift distribution is inhomogeneous with a maximum on the right side of the device. The inhomogeneities are slightly higher on the PDMOS devices.



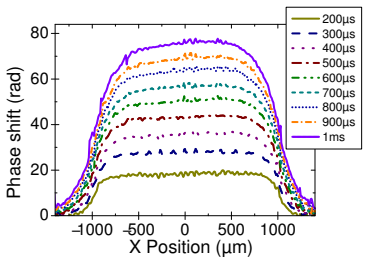
(a) NDMOS: I_{d}, V_{gs}



(b) PDMOS: I_{d}, V_{gs}

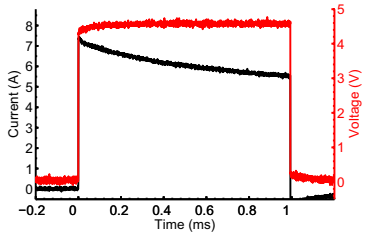


(c) NDMOS: Cross sections

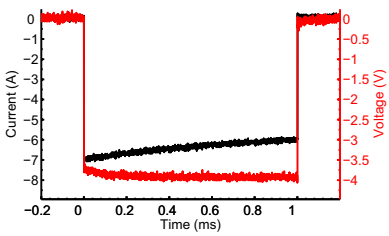


(d) PDMOS: Cross sections

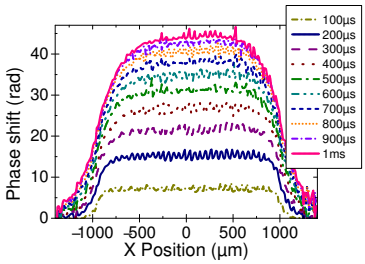
Figure 6.25: Waveforms and phase shift profiles above TCP, $V_{\mathrm{ds}} = \pm 35\mathrm{V}$, $V_{\mathrm{gs}}(\mathrm{NDM}) = 3\mathrm{V}$, $V_{\mathrm{gs}}(\mathrm{PDM}) = -2.5\mathrm{V}$



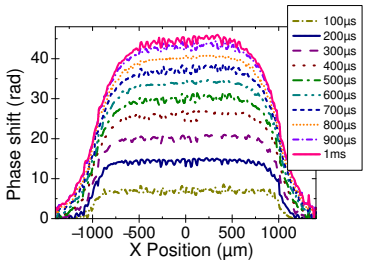
(a) NDMOS: I_{d}, V_{gs}



(b) PDMOS: I_{d}, V_{gs}



(c) NDMOS: Cross sections



(d) PDMOS: Cross sections

Figure 6.26: Waveforms and phase shift profiles far above TCP, $V_{\mathrm{ds}} = \pm 15\mathrm{V}$, $V_{\mathrm{gs}}(\mathrm{NDM}) = 4.5\mathrm{V}$, $V_{\mathrm{gs}}(\mathrm{PDM}) = -3.9\mathrm{V}$

Chapter 7

Conclusions and Outlook

The feasibility of the new 2D SLD setup for measurements at two time instants has been shown. The control system of the setup allows fast measurements. Only a few adjustments to the optics have to be made before the measurements and only a few parameters have to be set, while most of the measurement procedure is automated.

The 2D SLD setup is a useful additional setup to other interferometric setups. In comparison to the 2D TIM setup with pulsed lasers and 5ns time resolution, the SLD setup is much more feasible for the study of longer pulses. While ESD pulses can only be analyzed with the pulsed laser setup, the SLD setup shows superior phase accuracy for millisecond pulses, due to much lower noise and higher pulse to pulse stability. Compared to the heterodyne scanning setup, the SLD setup allows much faster measurements than the time consuming 2D scans. Though the phase shift accuracy of the heterodyne setup is still higher, the accuracy of the 2D SLD setup is often sufficient. In particular the feasibility of the 2D SLD setup for the characterization of unrepeatable behavior under long stress pulses was demonstrated.

The phase shift distribution in multiple DMOS transistors was mapped. In particular it was possible to observe the triggering of the parasitic bipolar transistor inherent to the DMOS structure experimentally and to get information about the size of the area where the onset of thermal runaway occurs. These results can be used for the verification and development of simulation tools. In a comparative study between an n-channel and a p-channel LDMOS transistor, an inhomogeneous phase distribution in the device was observed. This inhomogeneous distribution of temperature in the device is the reason that the destruction in this device under thermal stress does not occur in the center of the active region.

Further improvement in the automation of the 2D SLD setup would still be possible and comfortable for the user. In particular it would be useful to adapt the automated software for the 2D laser setup [27] to work also with the SLD setup. In this way additional functions like automated measurements at multiple stress levels and measurements of transfer or output characteristics between the single measurements would be possible.

Using such an automated setup and a pulse generator controlling the pulse width it would be possible to increase the stress duration step by step up to a desired pulse length or up to device destruction. This could be used to take image series of the evolution of the phase shift over time. Eventually the processing of multiple images acquired in this way could be made very efficient by implementing a new automated unwrapping process, which unwraps the phase based on 2D phase images at earlier time instants, similar to the unwrapping of transients in the heterodyne setup. However, it would be necessary to verify if the quality of the 2D phase shift images is good enough to allow such an unwrapping algorithm.

Another interesting future task could be to try to retrieve the 2-dimensional power dissipation directly from 2D images, if the noise is low enough. This is feasible, as the contribution of the time derivative at long pulse duration is low and therefore good information about the space derivative would be sufficient to estimate P_{2D}.

In addition to further studies on DMOS transistors the SLD setup is also well suited to investigate other types of silicon devices like bipolar transistors or IGBTs. Furthermore studies of III/V devices could be also of interest.

Bibliography

[1] B.J. Baliga. An overview of smart power technology. *IEEE Transactions on Electron Devices*, 38(7):1568–1575, July 1991.

[2] B.J. Baliga. Trends in power semiconductor devices. *IEEE Transactions on Electron Devices*, 43(10):1717–1731, October 1996.

[3] M. Denison, M. Pfost, K.-W. Pieper, S. Markl, D. Metzner, and M. Stecher. Influence of inhomogeneous current distribution on the thermal SOA of integrated DMOS transistors. *The 16th International Symposium on Power Semiconductor Devices and ICs (ISPSD). Proceedings*, pages 409–412, May 2004.

[4] V. Khemka, V. Parthasarathy, Ronghua Zhu, A. Bose, and T. Roggenbauer. Detection and optimization of temperature distribution across large-area power MOSFETs to improve energy capability. *IEEE Transactions on Electron Devices*, 51(6):1025–1032, June 2004.

[5] M. Pfost, J. Joos, and M. Stecher. Measurement and simulation of self-heating in DMOS transistors up to very high temperatures. *The 20th International Symposium on Power Semiconductor Devices and ICs (ISPSD). Proceedings*, pages 209–212, 2008.

[6] P. Spirito, G. Breglio, and V. d'Alessandro. Modeling the onset of thermal instability in low voltage power MOS: an experimental validation. *The 17th International Symposium on Power Semiconductor Devices and ICs (ISPSD). Proceedings*, pages 183–186, May 2005.

[7] J. Kölzer, E. Oesterschulze, and G. Deboy. Thermal imaging and measurement techniques for electronic materials and devices. *Microelectronics Engineering*, 31(1-4):251–270, 1996.

[8] G. Deboy, G. Sölkner, E. Wolfgang, and W. Claeys. Absolute measurement of transient carrier concentration and temperature gradients in power semiconductor devices by internal IR-laser deflection. *Microelectronics Engineering*, 31(1-4):299–307, 1996.

[9] M. Kuball, J. M. Hayes, M. J. Uren, I. Martin, J. C. H. Birbeck, R. S. Balmer, and B. T. Hughes. Measurement of temperature in active high-power AlGaN/GaN HFETs using Raman spectroscopy. *IEEE Electron Device Letters*, 23:7–9, 2002.

[10] E. Oesterschulze and M. Stopka. Imaging of thermal properties and topography by combined scanning thermal and scanning tunneling microscopy. *Microelectronics Engineering*, 31(1-4):241–248, 1996.

[11] D. Pogany, V. Dubec, S. Bychikhin, C. Fürböck, A. Litzenberger, G. Groos, M. Stecher, and E. Gornik. Single-shot thermal energy mapping of semiconductor devices with the nanosecond resolution using holographic interferometry. *IEEE Electron Device Letters*, 23(10):606–608, October 2002.

[12] G. Dolny, G. Nostrand, and K. Hill. Characterization and modeling of the temperature dependence of lateral DMOS transistors for high-temperature applications of power ICs. pages 789–792, December 1990.

[13] A. Castellazzi, V. Kartal, R. Kraus, N. Seliger, M. Honsberg-Riedl, and D. Schmitt-Landsiedel. Hot-spot meaurements and analysis of electro-thermal effects in low-voltage power-MOSFET's. *Microelectronics Reliability*, 43(9-11):1877–1882, 2003. Proceedings of the 14th European Symposium on Reliability of Electron Devices, Failure Physics and Analysis.

[14] Marie Denison. *Single Stress Safe Operating Area of DMOS Transistors integrated in Smart Power Technologies*. PhD thesis, Universität Bremen, November 2004.

[15] M.D. Pocha, A.G. Gonzalez, and R.W. Dutton. Threshold voltage controllability in double-diffused-MOS transistors. *IEEE Transactions on Electron Devices*, 21(12):778–784, December 1974.

[16] V. Dubec, S. Bychikhin, D. Pogany, E. Gornik, T. Brodbeck, and W. Stadler. Backside interferometric methods for localization of ESD-induced leakage current and metal shorts. *Microelectronics Reliability*, 47(9-11):1539–1544, 2007.

Proceedings of the 18th European Symposium on Reliability of Electron Devices, Failure Physics and Analysis.

[17] M. Litzenberger, C. Fürböck, S. Bychikhin, D. Pogany, and E. Gornik. Scanning heterodyne interferometer setup for the time-resolved thermal and free-carrier mapping in semiconductor devices. *IEEE Transactions on Instrumentation and Measurement*, 54(6):2438–2445, December 2005.

[18] J. A. McCaulley, V. M. Donnelly, M. Vernon, and I. Taha. Temperature dependence of the near-infrared refractive index of silicon, gallium arsenide, and indium phosphide. *Physical Review B*, 49(11):7408–7417, March 1994.

[19] D. Pogany, S. Bychikhin, M. Litzenberger, E. Gornik, G. Groos, and M. Stecher. Extraction of spatio-temporal distribution of power dissipation in semiconductor devices using nanosecond interferometric mapping technique. *Applied Physics Letters*, 81(15):2881–2883, 2002.

[20] V. Dubec, S. Bychikhin, M. Blaho, M. Heer, D. Pogany, M. Denison, N. Jensen, M. Stecher, G. Groos, and E. Gornik. Multiple-time-instant 2D thermal mapping during a single ESD event. *Microelectronics Reliability*, 44(9-11):1793–1798, 2004. Proceedings of the 15th European Symposium on Reliability of Electron Devices, Failure Physics and Analysis.

[21] V. Dubec, S. Bychikhin, D. Pogany, E. Gornik, N. Jensen, M. Stecher, and G. Groos. Error analysis in phase extraction in a 2D holographic imaging of semiconductor devices. volume 5290, pages 233–242. SPIE, 2004.

[22] Viktor Dubec. *Advanced optical interferometric methods for nanosecond mapping of semiconductor devices under high energy pulses*. PhD thesis, Technische Universität Wien, October 2005.

[23] `profi.m`. MATLAB program, Technische Universität Wien, 2009.

[24] S. Bychikhin and D. Pogany. Report on long pulse observation of DMOS transistors: improved methodology of extraction of useful signal. Internal report, Technische Universität Wien, May 2008.

[25] `SLD_TIM_2camera_ver1_SLD.vi`. LABVIEW program, Technische Universität Wien, 2009.

[26] G. Haberfehlner, S. Bychikhin, V. Dubec, M. Heer, A. Podgaynaya, M. Pfost, M. Stecher, E. Gornik, and D. Pogany. Thermal imaging of smart power DMOS transistors in the thermally unstable regime using a compact transient interferometric mapping system. *submitted and accepted for the 20th European Symposium on Reliability of Electron Devices, Failure Physics and Analysis (ESREF)*, 2009.

[27] M. Heer, V. Dubec, M. Blaho, S. Bychikhin, D. Pogany, E. Gornik, M. Denison, M. Stecher, and G. Groos. Automated setup for thermal imaging and electrical degradation study of power DMOS devices. *Microelectronics Reliability*, 45(9-11):1688–1693, 2005. Proceedings of the 16th European Symposium on Reliability of Electron Devices, Failure Physics and Analysis.

Acronyms

2D 2-dimensional

AM amplitude modulation

AOM acousto-optic modulator

BCD bipolar CMOS DMOS

BiCMOS bipolar CMOS

CMOS complementary MOS

DG delay generator

DMOS double-diffused MOS

ESD electrostatic discharge

FFT fast Fourier transform

FPA focal plane array

IFFT inverse fast Fourier transform

IMAQ image acquisition

IR infrared

LDMOS lateral double-diffused MOS

MOS metal oxide semiconductor

NDMOS n-channel DMOS

$\mathbf{P_{2D}}$ 2-dimensional power dissipation

P_{3D} 3-dimensional power dissipation

PCB printed circuit board

PDMOS p-channel DMOS

PBSC polarizing beam splitter cube

RF radio frequency

SLD superluminescent diode

TCP temperature compensation point

TIM transient interferometric mapping

TTL transistor transistor logic

VDMOS vertical double-diffused MOS

List of Figures