TESI DI DOTTORATO

UNIVERSITÀ DEGLI STUDI DI NAPOLI "FEDERICO II"

DIPARTIMENTO DI INGEGNERIA BIOMEDICA, ELETTRONICA
DELLE TELECOMUNICAZIONI

DOTTORATO DI RICERCA IN
INGEGNERIA ELETTRONICA E DELLE TELECOMUNICAZIONI

DESIGN AND VERIFICATION OF IGTBS RUGGED IN FORWARD AND REVERSE CONDITIONS

LUCA MARESCA

Il Coordinatore del Corso di Dottorato	Il Tutore
Ch.mo Prof. Niccoló RINALDI	Ch.mo Prof. Andrea IRACE

A. A. 2011–2012

To Enzo, Giuseppina and Maria

Acknowledgments

It would not have been possible to write this thesis without the help and support of the kind people around me.

First of all, I want to thank my supervisor, Professor Andrea Irace, for giving me the opportunity to join his research group. Without his constant support and encouragement, his precious suggestions and teachings, I would not achieved this goal. Many thanks to Professor Giovanni Breglio for his support and advice during my PhD. Thanks to Professor Paolo Spirito for his priceless teaching and for his inexhaustible enthusiasm for the research.

A special thanks to Michele, Martina, Matteo and Andrea, the people with whom I shared a great part of my time in the past three years. Thanks for their support, availability, friendship and the very good time we spent together.

Thanks to the Professors of the Electronic group of the DIETI, Daliento, Napoli and Rinaldi for the interesting discussions we had.

A great thanks to all friends of the DIETI, Grazia, Mariangela, Stefano, Maurizio, Pierluigi, Gianlorenzo, Fabio, Alessandro, Orlando, Enzo, Davide, Salvatore, Nicola, Marino, Giuseppe, Domenico, Giorgio and Michele for the pleasant moments spent together and their availability.

Many thanks to the old and new friends which have supported me and made me feel: Diego, Massimo, Marco, Fabio P., Carlo, Fabio N., Claudio, Agostino, Sabatino, Roberto, Giorgio, Riccardo, Giovanni, Mattia and all those are mentioned here.

I want to express my deep gratitude to my family for its spur to trust in my choices.

Finally, I want to thank my beloved Luisiana for her patience and continuous support.

Contents

List of Figures

Introduction

The optimization and control of the energy is one of the major issue in the modern society since the environmental concerns are becoming the more and more of crucial relevance. This led to an enormous increase of those applications based on electrical power because of their low impact on environment. Typical fields are transport, energy generation and transfer, and industries and common targets in everyday life are elevators, trains, subways and, especially in the last decade, automotive. Semiconductor power devices are becoming widely used since they are necessary to drive electrical motors. The efficiency and the reliability of these systems is strongly related to the power semiconductor devices adopted. Nowadays, IGBTs are the most used power devices for Middle/High power applications because of their excellent features. Actually, they have the advantages of both the power MOSFET and BJT, that is the easy driving of the first device and the low on-state voltage drop of the latter. However, the request of higher compactness and higher power capacity leads to the increases of the heat density and joule heating. These factors are the success key of those devices with a superior reliability in harsh conditions. Two conditions are electrically and thermally critical for the IGBT structure: the short-circuit and the avalanche breakdown. The former is a stressful unwanted operative condition which occurs because of some fail in the electric system where they are inserted, while the latter condition is particularly relevant since it can occur during the switch of the device and it is becoming more and more relevant in the wide area semiconductor devices. This work deals with the overall design parameters which affect the ruggedness of the device when operates in these conditions and the different trade-off among the reverse and forward performances parameters are defined and critically analyzed. This goal has been pursued by means of both theoretical, simulative and experimental approaches.

Thesis contents

The thesis deals with the design, analysis and development of IGBTS rugged in short-circuit conditions as well as in avalanche conditions. The experimental setups adopted to characterize the devices are reported as well.
The outline of the thesis is the following:

Chapter 1 deals with the introduction to the IGBT device structure and its basic modeling. After an historical introduction to the development of the IGBT structure, the physics of the operating mode is illustrated. The basic models to describe the electrical behavior of the device in DC conditions are recalled together with a description of the physical phenomena which occurs during the turn-ON and turn-OFF. The causes of the latch-up phenomenon are addressed and an overview of the FBSOA, RBSOA and SCSOA is reported. Finally the state of art IGBT designs are illustrated.

Chapter 2 deals with TCAD simulators and relevance into the semiconductor device development and the simulation strategy adopted to simulate multi-cellular power semiconductor devices. Firstly, an overview of the TCAD simulators is reported and their relevance in the development and analysis of semiconductor devices is highlighted. The simulation flow-chart in Sentaurus TCAD is reported together with a brief description of the tools involved to generate the structure, to make the mesh, to simulate the structure and to visualize the results. Follows the simulation strategy adopted to correctly simulate a multi-cellular IGBT and an automatized calibration procedure is presented. The latter is based on the minimization of the error between the experimental terminal curves (transfer characteristics and output characteristics) and the simulated ones. This is achieved by applying the minimization algorithm of MATLAB to manage the calibrate a suitable set of physical parameters and to run the Sentaurus TCAD simulations. The calibration is achieved in temperature as well.

Chapter 3 deals with the short-circuit simulation and the design parameters which affect the short-circuit capability of a PT-IGBT structure. The short-circuit capability is fundamental when an IGBT design must validated. Firstly the investigated structure has been calibrated by means of the calibration procedure presented in the previous Chapter and the experimental setup to measure the terminal characteristics are shown. Therefore, the design parameters which affect the short-circuit capability are investigated by

means of a parametric analysis of: cell-pitch, channel length, contact-trench features and Emitter diffusion modulation. For all of the above parameters the performances in on-state are evaluated by the trade-off among the voltage drop in on-sate, the turn-OFF losses and the short-circuit capability. The variation of two parameters, the channel length and the Emitter diffusion modulation, is investigated and the trade-off between the saturation current and the on-state voltage drop is reported. Finally, an optimized structure is presented. The modulation of the Emitter diffusion is adopted together with a shallow contact-trench to improve the short-circuit capability and the latch-up immunity, at the same time. The proposed design is validated by the experimental characterization of the samples in which the new design is adopted.

Chapter 4 deals with the avalanche breakdown phenomenon in the power semiconductor devices and the design parameters which affect the ruggedness in this condition are defined. The Negative Differential concept is illustrated and its relevance in the failure of semiconductor devices in avalanche conditions is highlighted. Firstly, the subject has been analytically investigated by modeling the I-V curve in blocking mode of the vertical PNP of an IGBT. The model is based on the calculation of the mobile charge into the depleted region (which defines the Electric-Field shape) by means of the charge control model to define the Beta of the transistor at different current levels. Therefore, the I-V curves in avalanche conditions are calculated starting from the physical characteristics of the structure. The results are compared to the results of the TCAD simulations performed on a test structure. The avalanche phenomenon has been investigated by means of the TCAD simulations as well to keep into account, in a detailed way, the effect of the design parameters on both the active area and termination I-V curves in avalanche conditions. More in detail the effect of the lifetime and the design of the buffer layer have been characterized. TCAD simulations have been performed to analyze the filamentation phenomenon in large area devices. The current and temperature dynamics in large area PNP structure (the vertical PNP in the IGBT structure) has been simulated in UIS conditions. The correlation between the current dynamics in the structure and the voltage drop on the structure is addressed. To analyze the dynamics of the current in avalanche condition for a real device (where active area and termination are present), a graphical method to find the voltage drop on the overall structure is presented and applied to explain the experimental evidences form the UIS test performed on different families of

IGBT devices. From the experimental point of view, the behavior in avalanche conditions has been characterized by means of suitable electrical tests, such as UIS, together with the Lock-In thermography. Finally, the design parameters which affect the avalanche of an IGBT are defined and the guidelines to improve the ruggedness of the structure are reported.

Scientific contributions

The main scientific contributions of this work are the following:

- Development of an automatic calibration procedure for IGBT, based on the terminal characteristics of device.
- Definition of an optimized design for a Trench-IGBT structure with the adoption of Emitter modulation technique and a swallow recessed Emitter contact.
- Development of a physical model to trace the I-V curve of the vertical PNP of an IGBT, starting from the physical features of the structure.
- Analysis of the filamentation phenomenon in UIS condition in the vertical PNP of the IGBT structure. The structure has a width of $600\mu m$ and the Periodic Boundary Conditions have been adopted.
- Development of a graphical method to analyze the parallel of two dipoles driven in current.

Chapter 1

Insulated Gate Bipolar Transistor

The Insulated Gate Bipolar Transistor (IGBT) is nowadays the most used silicon power device in the range of the medium/high power applications. Typical field of applications are: automotive, trains, variable speed refrigerators, air-conditioners and even stereo systems with switching amplifiers. Basically it is a three-terminal device which combines the advantages of both the Bipolar transistors and the MOSFET. Actually it has the simple Gate-drive characteristics of the power MOSFET devices and the high-current and low saturation-voltage capabilities of the power BJT. From the physical point of view, the IGBT is a four-layer structure (P-N-P-N) with a MOSFET structure which control the current flux of the entire device. The first proposal for an IGBT structure was taken by a Japanese researcher named Yamagami, who in a 1968 Japanese patent application proposed a MOS controlling a positive-negative-positive-negative (PNPN) semiconductor device without regenerative action[1]. From the experimental point of view, the mode of operation proposed by Yamagami was discovered by B. J. Baliga, which implemented a vertical structure with a V-Groove Gate region[2]. The same mode of operation was reported in [3] where the structure is called Insulated-Gate Rectifier (IGR), in [4] where the structure is called Insulated-Gate Transistor (IGT) and in [5] where the Conductivity-Modulated Field-Effect Transistor (COMFET) structure is reported. The same operation mode was proposed by Plummer with its four-layer device (SCR) filed in 1978, issued on 1980 [6] and reissued in 1990 [7]. In 1980 Hans W. Becke and Carl F. Wheatley proposed a structure similar to the IGBT, called Power MOSFET With an Anode Region [8],

in which it is claimed that the structure exhibits a Latch-up immunity in all the operative range of the device. The Latch-up phenomenon has been the main limitation for the commercialization of the new device. The first devices capable to experimentally sustain high currents and high voltages was presented by Baliga et al. [4] and by J.P. Russell et al. [5]. However, this first generation of devices suffered of Latch-up weakness and they showed a slow switching behavior. The latter drawback was overcame by using the electron irradiation lifetime killing technique [4], [9], which reduces the lifetime in the structure and the stored charge is removed faster during the turn-OFF of the device. In [10] it is shown the effective device capability to work at high temperatures. The robustness of the new device for the entire operation range was demonstrated for the first time in 1984. Actually, in [11] it was demonstrated the Short-Circuit withstanding capability of the device of a 1200V IGBT which has been tested with a bus voltage of 600V for 25 μs. When the speed was increased and the Latch-up suppressed for the interest range, the IGBT was introduced in the market of the power semiconductor devices from the Toshiba company. It was demonstrated that the product of the operating current density and the collector voltage exceeded the theoretical limit of bipolar transistors, $2 \cdot 10^5 W/cm^2$, and reached $5 \cdot 10^5 W/cm^2$[11], [12].

1.1 Basic design

In Fig. 1.1 a sketch of a basic IGBT structure is reported for a Planar Non-Punch-Through IGBT (NPT-IGBT) and a Planar Punch-Through IGBT (PT-IGBT). These designs are only two of the several structures which have been presented in literature to implement the device. The overall device is usually organized like a multi-cellular structure that is obtained by repeating the elementary cell many times and by extruding it. This solution returns a very high aspect ratio for the MOSFET part of the device (W/L). The different regions of the device are well visible. As mentioned before the structure shows a P-N-P-N structure, that in normal conditions is not active like a thyristor. More in detail, the N+ region is the Emitter diffusion, which is the Source of the MOSFET part of the device, the P-Body region is the Body of the MOSFET part of the device and, at the same time, it is the Collector of the vertical PNP structure present into the IGBT. The doping profile of the P-Body region at the Silicon/Oxide interface defines the Threshold voltage of the MOSFET part of the device. The Base of the PNP structure (the Drain of the MOSFET part) is called Drift-Layer and it has a doping concentration in the order of 10^{13}-10^{14}

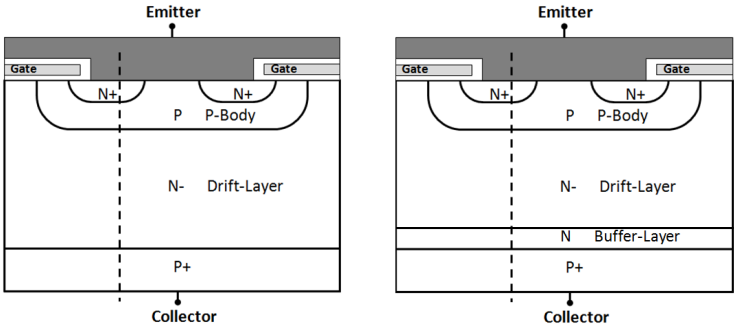
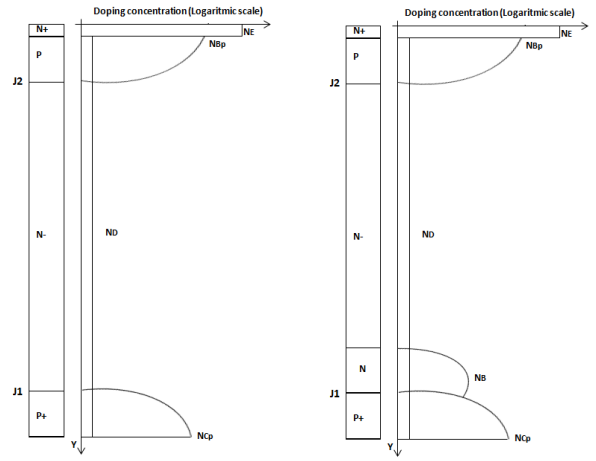


Figure 1.1: (a) Sketch of an elementary of a Planar NPT-IGBT structure and (b) a Planar PT-IGBT structure.

cm^{-3}. The P+ region is the Collector of the IGBT (the Emitter of the vertical PNP) and it is also called *hole-injecting layer*. It has to be pointed out that when the current level becomes very high, up to the latching current, the parasitic NPN structure (N+-P-N-) of the IGBT becomes active and the MOSFET part is not able to switch off any more the current into the device. In the PT-IGBT structure the Buffer-Layer is present as well (See Fig. 1.1(b)).

In Fig. 1.2 the doping profiles of the structures sketched in Fig. 1.1 are reported. The NPT structure is usually made by starting from a low doped silicon wafer with a thickness which depends on the blocking voltage of the device. The back of the structure is obtained by doping diffusion up to the junction J_1. For the PT design the process is different. More in detail the N- layer is epitaxially grown on a P+ silicon wafer and the thickness of the epitaxial layer defines the blocking voltage of the device. The top designs of the two geometries in Fig. 1.1 are made with the same technological steps. The Planar design reported in Fig. 1.1 is also called Self-Aligned since both the N+ and the P diffusion are made by aligning them to the border of the Gate. More in detail, after the deposition of the Gate oxide and the Polysilicon Gate, the P region is made by means of ionic implantation and the Gate border works like a mask for the dopant implantation and the lateral diffusion defines the channel region. In a second step the N+ region is made and it is aligned to the Gate as well as the P region and the difference in depth of the N+ and the P region defines the channel length of the MOSFET part. Therefore the top part of the device is similar to that of a Double-Diffused power MOSFET (DMOS).



(a) (b)

Figure 1.2: (a) Doping profile along the dashed line for the Planar NPT-IGBT structure in Fig. 1.1(a) and (b) the doping profile along the dashed line for the Planar PT-IGBT structure in Fig. 1.1(b).

1.2 Device operation and modeling

The presence of the P+ Collector layer is the key of the success of the IGBT since it allows the conductivity modulation into the Drift-layer in on-state conditions. To simplify the analysis the device behavior, it is supposed that the Emitter terminal is connected to ground. Whether the Gate-Emitter voltage is kept to zero and the Collector voltage is increased, the Collector/Drift junction is forward biased, while the P-Body/Drift junction is reversely biased and the current flow is limited to the leakage current of the junction until the open-base breakdown voltage of the vertical PNP is reached. This is called forward blocking condition. Whether the Gate-Emitter voltage is fixed to zero and a negative potential is applied to the Collector terminal, the P-Body/Drift junction is directly biased and the Collector/Drift junction is reversely biased. This is called reverse blocking condition. In general, the blocking voltages relative to the two illustrated conditions can be very different for the different families of IGBT structures. Only for NPT structure the forward blocking capability is equal to the reverse blocking capability because of the symmetry of the structure; since both the Collector region and the P-Body region have a doping

concentration some order of magnitude higher than that of the Drift-Layer, the blocking voltage of the Collector/Drift junction and that of the P-Body/Drift junction are similar. For the PT structure the forward blocking voltage capability is sensibly higher than that in reverse blocking conditions basically because of the high doping concentration of the Collector/Buffer junction. The on-state condition is achieved when a Gate-Emitter voltage higher than the Threshold voltage is applied. Whether a positive voltage is applied to the Collector of the device, the electron current can flow through the MOSFET channel and it is injected into the Drift-Layer and the Collector/Drift junction is directly biased and the holes are injected into the Drift-Layer. In this condition the double injection occurs into the Drift-Layer because of both the holes injected from the Collector/Drift and the electrons coming from the N-MOSFET channel. The electrons and the holes recombine into the Drift-Layer to keep the neutrality and the exceeding holes are collected by the Collector of the PNP. When the current density increases, the density of the mobile charges becomes higher than the doping concentration of the Drift-Layer. This effect is named *conductivity modulation* and it decreases the resistance of the Drift-Layer. Therefore the current density of an IGBT is much higher than that of a Power MOSFET. Because of the overall doping profile, the P-Body/Drift junction con not be directly biased in on-state, therefore the vertical PNP does not work in saturation. However, when the voltage drop across the inversion layer becomes comparable to the difference between the Gate voltage and Threshold voltage, the MOSFET goes in pinch-off condition, limiting the current. When the current is limited by the MOSFET, the overall current is limited as well, therefore the Collector current saturates. The main limitation in switching frequency for IGBTs is related to the turn-off of the vertical PNP. As mentioned before, during the on-state the conductivity modulation phenomenon occurs into the Drift-Layer and a certain amount of charge is stored. When the Gate voltage is driven to zero to switch-off the IGBT, when the channel closes part of the stored charge is still present into the Drift-Layer. Therefore, the remaining charge is swept out by the recombination phenomenon and a certain time is needed to remove the charge (typically there are not external contacts with the Drift-Layer region). During the charge removal time the device exhibits the typical Collector current tail which differentiates the IGBT from the power MOSFET. As mentioned before to increase the speed of the IGBT, especially in PT-IGBT designs, lifetime killing techniques are adopted to increase the recombination rate into the structure (e.g. [4]). In the PT-IGBT structure the lifetime is also reduced by means of the doping concentration of the Buffer-

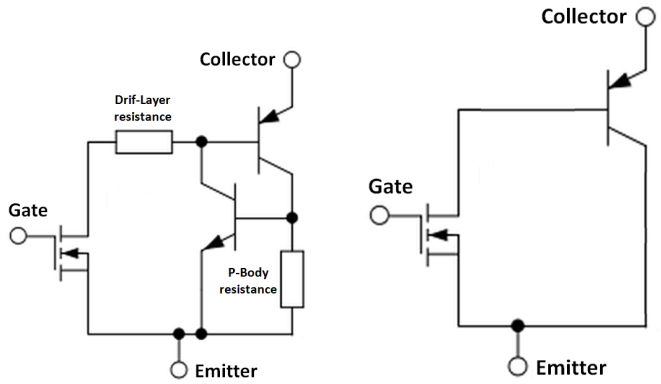
layer. In particular the more the doping concentration increases, the more the lifetime decreases with the consequent reduction of the holes injection into the Drift-Layer. The latter effect reduces the stored charge into the Drift-Layer and the current tail duration. The main drawback is the increase of the Voltage drop in on-state which increases when the lifetime decreases. Usually in NPT structures the lifetime killing technique are not adopted, therefore these devices have a low V_{ON} but a lower speed. An IGBT model must be able to describe with a suitable accuracy the behavior the device in a circuit and in a reasonable time. IGBT modeling can be pursued by three approaches[13]:

1. Analytical Device and Circuit Simulations - The behavior of the device is described like the connection of different discrete devices (e.g. MOSFET and P-i-n or MOSFET and BJT). This approach provides a physical understanding of the device behavior and allows to calculate the output curves of the device in a generic circuit. Great attention has to be paid when dynamical simulations are performed.

2. Numerical device simulation - Part of the device structure is recreated and the terminal characteristics are calculated by solving the Drift-Diffusion equations which regulate the physics of the semiconductor devices. The semiconductor equations can be coupled to the thermal equations keep into account the electro-thermal effect in the structure. The precision of the results is very high, but the exact knowledge of all the design parameters of the structure have to be known. This simulation usually need long time to run.

3. External behavior modeling - In this approach the only information are the terminal curves without any concern for the internal structure of the device. This approach is used for very complex circuits with a large number of devices.

Typically, the terminal characteristics of the IGBT are analyzed by using the first approach proposed above. In Fig. 1.3(a) the equivalent circuit of the structure is sketched and the resistance of the P-Body layer (responsible of the latching of the device) is reported together with the resistance of the Drift-Layer region. When different techniques are adopted to suppress the Latch-up phenomenon, the parasitic NPN transistor con be removed from the circuital schematic together with the resistance of the P-Body Layer and a simplified circuit can be considered (See Fig. 1.3(b)). A simpler schematic circuit, a MOSFET with a P-i-N power diode is series, is also adopted for the



(a) (b)

Figure 1.3: (a) A schematic of the IGBT structure and (b) a simplified version without the parasitic NPN transistor.

description of the IGBT. However this approach does not keep into account the hole injection into the Drift-Layer.

1.3 DC characteristics

By means of the schematic in Fig. 1.3(b), the DC characteristics of a PT-IGBT are evaluated in this section. Basically, the $I_C - V_{GE}$ and $I_C - V_{CE}$ families of curves are needed to characterize the behavior of the device in on-state. The analysis of the transfer characteristic, $I_C - V_{GE}$, can be divided is two regions: the sub-threshold region and the high-current region. In an IGBT the $I_C - V_{GE}$ characteristic in the sub-threshold region is dominated by two contributes: the leakage current of the reversely biased P-body/Drift junction, enhanced by the gain of the vertical PNP [14], and the diffusion current which flows through the weakly depleted region at the gate oxide/P-body interface.The second contribution to the current in the sub-threshold region is due to the diffusion current due to the weakly depleted region at the silicon/dioxide interface. By using the extended expression of the surface potential [15], the sub-threshold slope, in semilogarithmic scale, can be evaluated by the following equation:

$$S_{S-th} = \left[ln(10) \frac{KT}{q} \left(1 + \frac{C_D}{C_{OX}} \right) \right]^{-1} \tag{1.1}$$

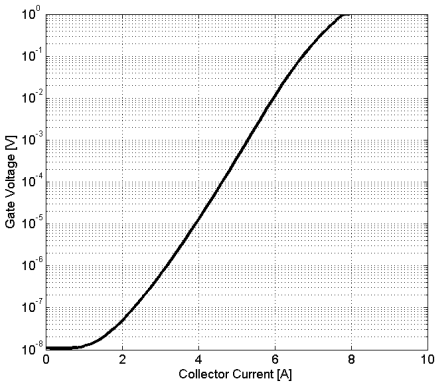


Figure 1.4: The sub-threshold characteristic of a PT-IGBT ($V_{CE} = 10V$).

where C_D is the depletion region specific capacitance, K is the Boltzman constant and C_{OX} is the oxide specific capacitance. The sub-threshold region slope decreases by increasing the temperature by the KT/q term. The two contributes described above are both present for each V_{GE} value in the sub-threshold curve. At low V_{GE} values, the P-body/Drift leakage current is predominant and the collector current is almost constant and related to the Collector voltage applied (since the depletion region reduces, it slightly decreases by increasing the gate voltage). By increasing the gate voltage, the diffusion current, due to the weak depletion region below the gate oxide, becomes higher assuming an exponential behavior (linear in semilogarithmic scale). When the strong inversion occurs and the channel is formed ($V_{GS} > V_{th}$), the current flowing in the channel is due to drift transport and the characteristic is not linear anymore. An experimental sub-threshold curve is reported in FIG. 1.4.

When the Gate voltage reaches the threshold voltage, an electron current, which is injected into the Drift-Layer, starts to flow through the channel region. At this point the electron current coming from the channel feeds the Base of the bipolar transistor. The bipolar transistor activation leads to an enhancement of the collector current (compared to the one of a power-MOSFET with the same rating and geometrical features). In a first-order analysis, the collector current, assuming a pinch-off condition for the MOSFET part of the T-IGBT structure, can be evaluated by [14]:

$$I_C = \frac{\mu_n C_{OX} W_{CH}}{2L_{CH}\left(1 - \alpha_{PNP}\right)} \left(V_{GE} - V_{th}\right)^2 \tag{1.2}$$

where W_{CH} is channel width of the MOSFET, α_{PNP} is the current gain of the vertical PNP and L_{CH} is the channel length of the MOSFET part. The main difference between the two designs in §1.1 from the point of view of the saturation current is related to the variation of the current gain of the vertical PNP. Actually the current gain of the PNP can be calculated by $\alpha_{PNP} = \gamma_E \alpha_{T,PNP}$, where the γ_E and $\alpha_{T,PNP}$ are the Emitter efficiency and the transport factor of the vertical PNP, respectively. These two values are strongly dependent on the design of the IGBT [14].

The on-state voltage drop (V_{ON}) for an IGBT can be modeled like the sum of three contribution:

$$V_{ON} = V_{P+N} + V_{NB} + V_{MOSFET} \tag{1.3}$$

where V_{P+N} is the voltage due to the forward biased Collector/Buffer junction, V_{NB} is the voltage drop in the Drift-Layer region and V_{MOSFET} is the contribute due to the MOSFET part in the structure. For a PT-IGBT structure the voltage due to the forward biased Collector/Buffer junction can be calculated by:

$$V_{P+N} = \frac{KT}{q} ln\left(\frac{p_0 N_{D,NB}}{{n_i}^2}\right) \tag{1.4}$$

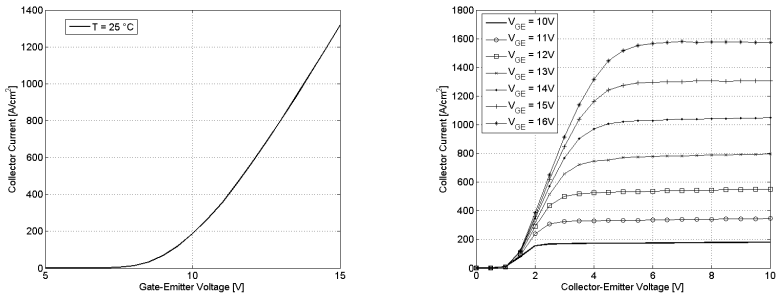
where p_0 is the hole concentration at the P+ Collector/Buffer junction (J1) and $N_{D,NB}$ is the doping concentration of the Buffer-Layer. The contribute due to the MOSFET region can be evaluated by:

$$V_{MOSFET} = \frac{J_C L_{CH} p}{2\mu_n C_{OX} (V_G - V_{TH})}. \tag{1.5}$$

where p is the pitch of the elementary cell (the pitch is considered here as the distance between two consecutive Gates).Whether the doping profile and the geometrical features of the structure are defined, V_{NB} is the only contribute which is very sensitive to the lifetime (τ_{HL}) in the Drift-Layer and it can be evaluated by summing two contributes:

$$V_{NB} = V_{NB1} + V_{NB2} \tag{1.6}$$

where



(a) (b)

Figure 1.5: (a) Transfer characteristic and (b) output characteristics of a PT-IGBT structure.

$$V_{NB1} = \frac{2L_a J_C sinh(W_N/L_a)}{qp_0\left(\mu_n+\mu_p\right)} \left\{tanh^{-1}\left[e^{-(W_{ON}/L_a)}\right] - tanh^{-1}\left[e^{-(W_N/L_a)}\right]\right\} \quad (1.7)$$

and

$$V_{NB2} = \frac{KT}{q}\left(\frac{\mu_n-\mu_p}{\mu_n+\mu_p}\right) ln\left[\frac{tanh\left(W_{ON}/L_a\right) cosh\left(W_{ON}/L_a\right)}{tanh\left(W_N/L_a\right) cosh\left(W_N/L_a\right)}\right]. \quad (1.8)$$

where $L_a = \sqrt{D_a \tau_{HL}}$ is the ambipolar diffusion length, D_a is the diffusion coefficient, τ_{HL} is the high injection level lifetime, W_{ON} is the depletion layer depth, W_N is the length of the not depleted region of the Drift-layer and J_C is the Collector current density. The first term is the most sensitive to the high-injection level lifetime in the drift-layer since it is at the denominator in the term. In Fig. 1.5 both the experimental transfer characteristic and the output characteristics of a PT-IGBT are reported.

1.4 Dynamic switching characteristics

The great success of the IGBTs is due to their DC features as well as their driving facility. Actually the MOSFET part of the device allows the use a

simplified driving circuitry. The vertical PNP of the structure allows to achieve very low losses in ON-state conditions, but, at the same time, it slows down the dynamics of the overall devices and increases the switching losses. In this section the turn-ON and the turn-OFF waveforms are reported together with a brief description of the physical mechanisms which occurs.

1.4.1 Turn-ON

The turn-ON signal is applied to the Gate of the MOSFET part of the IGBT. This is one of the most important features of the IGBT since it is controlled in voltage rather than current (as happens in BJTs). However differently from the power MOSFET, the IGBT takes a longer time to turn-ON because of the time needed to force both the MOSFET and the BJT in conduction state. To turn ON the IGBT, the input capacitance between Gate and Emitter is charged to a voltage V_{GE} higher than the threshold voltage V_{th}. Since the switching applications are the most relevant field of the IGBTs, the turn-ON mechanism is strictly related the Free-Wheeling Diode (FWD) used in the circuit (See Fig. 1.6).

To evaluate the turn-ON characteristics it is supposed that the circuit time constant is $>>$ the switching time of IGBT used in a way that the I_L current is constant. When the IGBT is off, the current flows into the FWD and the voltage drop across the L_0 inductance is almost zero [16],[13].

The switching waveforms of an IGBT in a clamped inductive circuit are shown in Fig. 1.7. The IGBT turn-ON switching performance is dominated by its MOSFET structure. During the time interval 1 the Gate current charges the constant input capacitance with a constant slope until the Gate-Emitter voltage reaches the threshold voltage V_{th} of the device. During the time interval 2 load current is transferred from the diode into the device and increases to its steady-state value. The Gate voltage rise time and IGBT transconductance determine the current slope and as a result the time interval 2. When the Gate-Emitter voltage reaches $V_{GG(on)}$ that will support the steady-state Collector current, Collector-Emitter voltage starts to decrease. After this there are two distinct intervals during IGBT turn-ON. In the first interval the Collector-to-Emitter voltage drops rapidly as the Gate-Drain capacitance C_{gd} of the MOSFET portion of IGBT discharges. At low Collector-Emitter voltage C_{gd} increases. A finite time is required for high-level injection conditions to set in the Drift-Layer region. The PNP-transistor portion of IGBT has a slower transition to its on-state than the MOSFET. The Gate voltage starts rising again only after the transistor comes out of its saturation region into the linear region, when

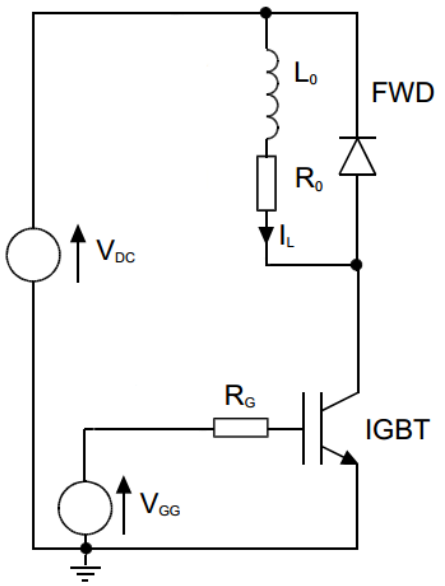


Figure 1.6: Typical circuit for the evaluation of the turn-ON and turn-OFF losses.

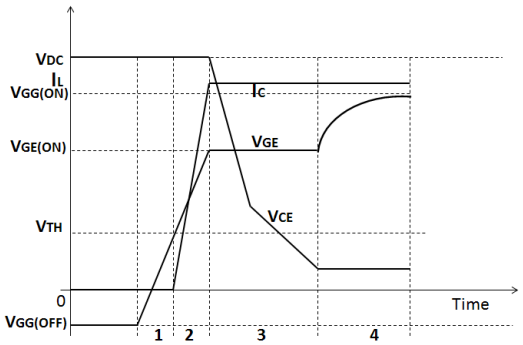


Figure 1.7: Turn-ON waveforms for a NPT-IGBT.

complete conductivity modulation occurs and the Collector-Emitter voltage reaches its final ON-state value.

1.4.2 Turn-OFF

The turn-OFF in IGBTs is characterized by the presence of a current tail after switch-off of the MOSFET part of the device. This tail is the main source of the losses and it limits also the the maximum switching frequency at which the device can work. Depending on the application, the lifetime into the structure can be modified by acting on the Collector/Drift design or by acting on the lifetime of the structure by means on lifetime killing techniques. In general, during turn-OFF the stored charge must be evacuated; this sets rate of rise in collector voltage dV_{CE}/dt, as the charge is extracted and the depletion layer expands . Fig. 1.8 shows a typical NPT IGBT inductive turn-OFF waveform, with the five phases of turn-OFF described as follows [17],[14]:

1. The gate voltage V_{GE} falls to a level set by the MOSFET channel current (approximately proportional to the collector current I_C) and the MOSFET channel is in linear operation during this phase.

2. The MOSFET channel is now in saturated operation. The Miller capacitance C_{GC} charges through the gate and the collector voltage V_{CE} starts to rise.

3. When the accumulation layer under the gate has disappeared, the Miller capacitance decreases suddenly and V_{CE} increases sharply. dV_{CE}/dt is now limited by the rate at which the depletion layer can expand.

4. When V_{CE} has reached approximately the supply voltage V_{dc}, the freewheel diode can turn ON. This allows I_C to fall to a level I_{CT} set by the remaining stored charge, and the collector V_{CE} exhibits the classic overshoot from discharging the stray inductance.

5. The current tail now begins, and I_C is now set only by the remaining stored charge and recombination rate. The tail current is strongly dependent on the high-level carrier lifetime τ in the drift region and the charge remaining in the Drift-Layer.

The mentioned turn-off mechanism, however, becomes more complicated when the RBSOA has to be defined. Actually when silicon power devices turn-off in harsh conditions (over-voltage, over-current, over-temperature and so on

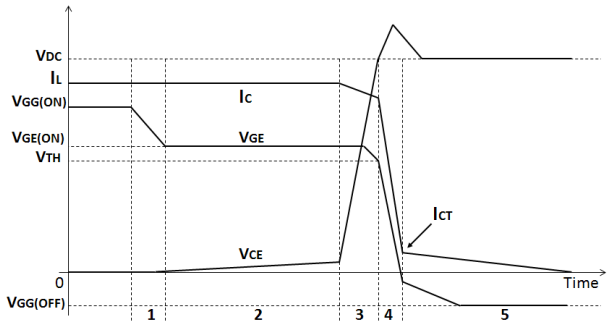


Figure 1.8: Turn-OFF waveforms for a NPT-IGBT.

conditions) complex phenomena appear which can lead to the destruction of the device [18], [19], [20], [21]. In all of these works the main problem is the presence of a current inhomogeneity, which can lead to a premature failure of the device since a great amount of power is dissipated in a reduced area of the device. The current filamentation is due, basically, to the avalanche generation and it is triggered when the power electrical system subjects the device to a fast dynamical transient [22] and the phenomenon is accentuated for high current densities. The failure analysis in power p-i-n diodes has been investigated in many works (e.g. [23],[24]) and the dynamic avalanche phenomenon has been addressed like the main cause of failure since it introduces a strong instability into the device during the turn off and leads a current filamentation on both the cathode and anode side. Actually, the dynamic avalanche phenomenon has been investigated also in thyristor-like devices (IGBTs and GTOs) and the current crowding has been addressed like a detrimental factor for for the ruggedness of the devices (e.g. [18], 12,[25]). The dynamic avalanche can be outlined in this way: when the accumulation layer at silicon/oxide interface goes to zero(after the Miller's plateau), the channel closes and the current falls down abruptly together with the collector voltage which rises up to the bus voltage steeply. During this phase the depleted drift-layer(at the P-body/Drift junction) increases with a certain velocity which defines the dV_{CE}/dt. At the same time, then, coexist two regions in the Drift-Layer: A depleted region and a region where the carriers are still stored (so called plasma). In general a second depleted region can appear between the plasma region and the collector region of the device [26]. In has to be pointed out that the collector voltage increase is strictly related to both the inductive load and the parasitic inductance

of the circuit. The latter contribute is the only one present at the turn-off after a Short-circuit test [27]. As mentioned before a depleted region at P-body/Drift junction rises up during the channel closure, but at the same time a great flux of holes passes through this region, changing the net charge in that region. In harsh conditions, by fixing the depleted region thickness at a certain time, it can happen that the electric field steepness increases so much during the time laps described before that a premature achievement of the critical electrical field can occur and a strong carriers generation due to the impact ionization phenomenon occurs. When the latter condition is achieved the dynamic avalanche is occurring into the device and a strong instability influences the carriers distribution along the whole area of the device. Usually the voltage at which the dynamic avalanche occurs is much lower than the static one, and a typical marker of the phenomenon is a flat collector voltage waveform. In literature the dynamic avalanche topic has been tackled from different points of view. In [17], [28] and [22] the turn-off phenomenon is investigated and an analytical model for the dV_{CE}/dt behaviour during the the turn-off/dynamic avalanche phase is proposed. A complete modeling of the collector voltage dynamic is reported in [17] where the temperature dependence of dV_{CE}/dt is included into the model in order to complete the theoretical analysis introduced by [28], [22]. Some interesting experimental guidelines to enhance the reliability of the IGTBs in avalanche conditions are reported in [29] where it is highlighted that a great improvement of the ruggedness can be achieved by making an active area with a lower breakdown voltage, compared to the termination one. However, as observed by the authors, the restrictions imposed by the on-state performances are usually stronger that that in avalanche conditions. In [27] an experimental investigation of the dynamic avalanche effect on the SCSOA is presented and general guidelines to improve the reliability during the short-circuit are reported. In [30] an experimental investigation of the devices failure has been used to define the geometrical parameters which enhance the ruggedness in hard-switching conditions. TCAD simulators have been widely used to investigate the failures due to the dynamic avalanche effect (e.g.[18], [19], [20]), and, together with the support of the experimental characterization of the devices in stressful conditions, the failure causes have been explained. In [18] the filamentation phenomenon is investigated for GTO devices and it is shown like the dynamic avalanche phenomenon gives an instability to both the current distribution into the device and to the collector voltage. In [19] the failures for high power IGBT modules are reported and explained by means of both the TCAD simulations and analytical approach. It is highlighted that

the geometrical features of the die strongly affect the ruggedness of the device in over-current and over-temperature turn-off conditions. In [31],[32] the cell termination is analyzed when the dynamic avalanche occurs. In [21] by means of TCAD simulation has been shown that the current filamentation, in different families of silicon power devices, has a crucial role into the failure of them and some parameters of the elementary cell are linked to the ruggedness of the device. TCAD simulators recently have been used to investigate the physics involved into the turn-off of large area devices [33]. In this work a 3D wafer level TCAD simulation has been performed and some of the most important phenomenon which occurs during the turn-off of a GCT are investigated. The dynamic avalanche phenomenon is investigated as well. The last work has shown that it possible to use TCAD simulator also for very large areas, however it has to be pointed out that the analysed structure is quite simple from the geometrical point of view compared to the complexity of a Trench-IGBT structure which needs a huge number of nodes for a reliable description of a real structure.

1.5 Latch-up

One of the problems that the IGBT structure has shown at the beginning of its development is the Latch-up phenomenon. As mentioned before, the IGBT has a parasitic P+NPN+ structure which is normally switched off. However, when the Collector current reaches a critical value, the Emitter/P-Body junction can become forward biased and the thyristor structure starts to conduct without the control of the Gate. The thyristor structure activation can lead to the destruction of the device and suitable countermeasures have to taken.

In Fig. 1.9 the hole path below the N+ Emitter for a Planar-IGBT is reported. When the device is in on-state conditions, the electron current flows only into the MOSFET channel, while the holes are collected by the Collector of the vertical PNP. These holes are collected all over the P-Body/Drift junction and part of them flow below the N+ diffusion and the resistance of the P-Body layer causes a voltage drop at the junction. Even if the N+ emitter and the P-Body are connected by the metal layer on the surface, when the current density achieves a critical value, the voltage drop in a certain point (X) of the Emitter/P-Body junction is high enough to activate the parasitic NPN (N+/P-Body/Drift-Layer) of the structure. At this point the thyristor is active and it there is no way to switch-off the IGBT. If I_p is the hole current and R_p is the resistance of the P-Body, therefore

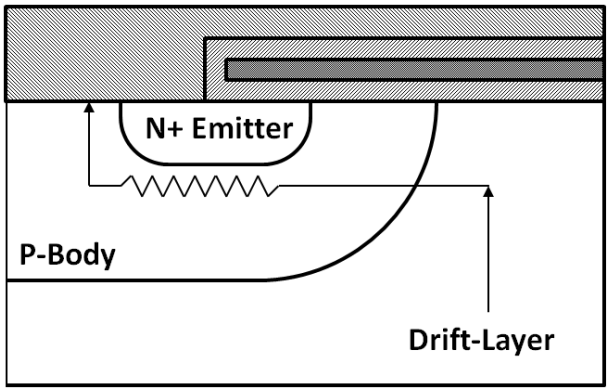


Figure 1.9: Hole path in a Planar-IGBT.

$$V_X = R_p I_p \tag{1.9}$$

but for equation the hole current can be related to the electron current by

$$I_p = \frac{\alpha_{PNP} I_n}{(1 - \alpha_{PNP})} \tag{1.10}$$

where α_{PNP} is the current gain of the PNP. The Emitter current can be written like:

$$I_E = I_p + I_n = \frac{I_n}{(1 - \alpha_{PNP})}. \tag{1.11}$$

By means of equations (1.9), (1.10) and (1.11)

$$V_X = \alpha_{PNP} R_p I_{CE}. \tag{1.12}$$

The Latch-up phenomenon occurs when $\alpha_{NPN} + \alpha_{PNP} = 1$ and this condition is achieved when a voltage equal to the Built-in potential drops at the Emitter/P-Body junction. The resistance R_p is proportional to the length of the N+ region (L_E) and to the sheet resistance of the P-Body region (ρ_{sp}). Therefore, the current at which the Latching occurs is [13]:

$$I_{L,SS} \propto \frac{V_{bi}}{\alpha_{PNP} \rho_{sp} L_E}. \tag{1.13}$$

Even if this expression gives only a rough evaluation of the current density at which the Latch-up occurs, it indicates the guidelines to improve robustness

of the device. Basically it possible to act on the design of the N+ Emitter region by L_E, on the doping profile by $\rho_{sp}L_E$ and on the current gain of the vertical PNP by α_{PNP}. All of these parameters are strictly related to the overall design of the device, but the main solutions to the problem are concentrated to the reduction of R_p. The structure in Fig. 1.9 is very weak from the point of view of the Latch-up since the doping profile of the P-Body region is constant. The P-Body doping concentration defines the threshold voltage of the MOSFET structure and it can not be increased to reduce the sheet resistance (the sheet resistance reduction implies an increase of the threshold voltage). The typical values of doping profile for the P-Body region are too low and an high value for the sheet resistance is achieved, therefore the voltage drop needed to turn-on the thyristor is achieved for low currents because of the high value of the resistance. Different solutions to the problem have been proposed in literature to suppress the Latch-up. In [34] it is reported one of the most effective solutions to the problem. Actually the P-Body region is made with two doping diffusion; one defines the threshold voltage of the MOSFET structure and the other reduces the sheet resistance along the path of the holes. The main drawback is the accuracy needed to avoid that the additive diffusion for the Latch-up phenomenon reaches the oxide interface and changes the threshold voltage of the device. In [35] the adoption of a buried layer below the N+ diffusion is proposed. This technique eliminates the drawback of the previous technique and doubles the current at which the Latching occurs. In [11] a great great improvement of the Latch-up phenomenon is achieved by making only one N+ diffusion per Gate. In this way a part of the current flows below the N+ region on one side and on in the other side the current flows through the P-Body region but not below the N+ region since it is not present. In [36] to overcome the problem of the higher doping concentration which can be caused by the previous techniques an additive N-layer is made below the channel to reduce the net doping concentration. A great improvement of the Latch-up immunity has been achieved when the Trench-IGBT structure has been introduced. Actually, in this structure the MOSFET channel is vertical and the hole flux is drastically reduced compared to the planar design. The Latch-up ruggedness can be improved by adopting the same principle in [34]; a P+ region is made in the middle of the P-Body contact to reduce the sheet resistance (See Fig. 1.10). In [37] a reduction of the Latch-up phenomenon is achieved by using a recessed contact for the Emitter region.

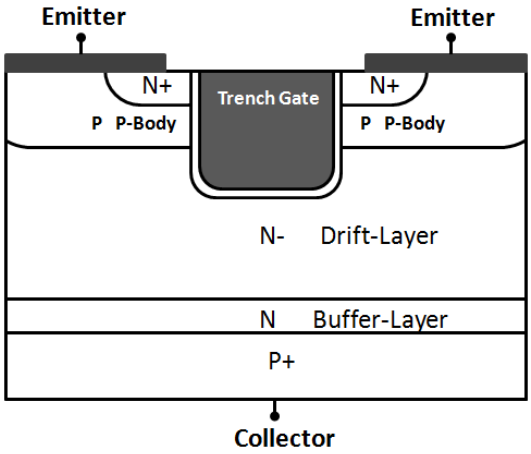


Figure 1.10: Sketch of a Trench-IGBT structure.

1.6 Safe Operating Area

An important parameter which defines the ruggedness of a power device is its Safe Operating Area (SOA) which is the area in the I-V plane where the Collector current and the Collector/Emitter voltage can lay without a reduction of the reliability of the device or destruction of the device. Typically for low collector currents the maximum voltage is limited by the open-base breakdown voltage of the vertical PNP. It has to be pointed out that the avalanche breakdown, which define the maximum voltage that the device can sustain, can occur at the termination of the device as well as in the active area of the device, depending on the design of the two regions (this subject will be analyzed in detail in Chapter 4) and on the electro-thermal interaction between the two regions. For low voltages, the maximum current for a power device is limited by the dissipated power, however for IGBTs the main cause of failure is related to the Latch-up phenomenon. As shown in §1.4, during the turn-OFF and turn-ON of the device the Collector voltage and the Collector current can be high at the same time, with an high power dissipation into the structure. In these conditions the avalanche-induced second breakdown can lead to the destruction of the device. This phenomenon defines the Forward-Biased Safe Operating Area (FBSOA) when the device is turned on and the Reverse-Biased Safe Operating Area (RBSOA) when the device is turned off. The reliability of an IGBT is also related to the Short-Circuit withstanding capability of the device, which the maximum time which the device can stand the Short-Circuit condition be-

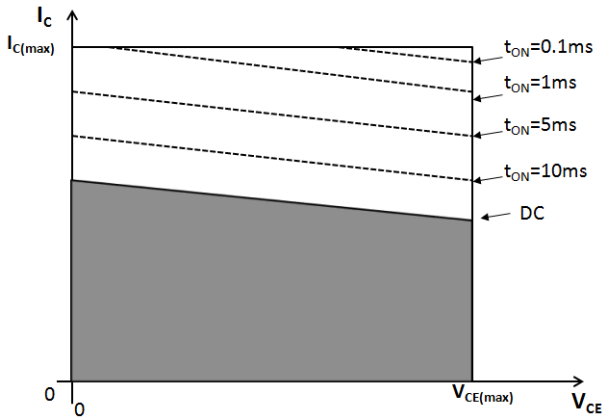


Figure 1.11: SOA of an IGBT.

fore the failure of the device occurs. The Short-Circuit Safe-Operating-Area defines this feature of the device.

In Fig. 1.11 a sketch of the SOA an IGBT is reported. As it is visible the SOA increases when a pulsed condition occurs: the more the t_{on} decreases, the more the maximum Collector voltage and Collector current increases. Anyway the maximum Collector voltage and Collector current are limited by the avalanche breakdown phenomenon and the Latch-up.

1.6.1 FBSOA

The maximum Collector-Emitter blocking voltage when the Collector current is saturated defines the FBSOA. In this condition a great amount of mobile charge flows into the Drift-Layer and, because of the high electric field into the depletion region, the electrons and the holes are accelerated to their saturation velocity, v_{sn} and v_{sp} respectively. Therefore, the electrons and holes density can be calculated starting from the electron and holes currents

$$n = \frac{J_n}{qv_{sn}}, p = \frac{J_p}{qv_{sp}} \tag{1.14}$$

and the total positive charge into the depleted Drift-Layer can be evaluated by

$$Q^+ = N_D - \frac{J_n}{qv_{sn}} + \frac{J_p}{qv_{sp}} \tag{1.15}$$

where N_D is the doping concentration of the Drift-Layer. In FBSOA Q^+ is higher than in low current conditions (typically p¿¿n) and because of the Poisson low, the slope of the electric field increases into the depletion region and the critical Electric Filed increases as well. In this condition the breakdown voltage decreases and a rough estimation of the breakdown voltage can be obtained by [13]

$$BV_{SOA} = \frac{5.34 \times 10^{13}{}^{0.75}}{(Q^+)} \quad . \tag{1.16}$$

The transistor effect has to be taken into account as well, therefore the condition $\alpha_{PNP}M = 1$ has to be satisfied in avalanche conditions, where α_{PNP} can be approximated by

$$\alpha_{PNP} = \left[cosh\left(\frac{W}{L_a}\right)\right]^{-1} \tag{1.17}$$

and by the Miller relation [38] the multiplication factor M can be calculated by

$$M = \left[1 - \left(\frac{V}{BV_{SOA}}\right)^n\right]^{-1} \tag{1.18}$$

where W is the undepleted Base and $n = 4 - 6$. From the above equations can be concluded that the breakdown voltage of the active area can be lower than that of the termination and the limitation in voltage occurs at a lower voltge.

1.6.2 RBSOA

The RBSOA is crucial in turn-OFF conditions. It has to be pointed out that in turn-off conditions the only carrier present into the depletion region are the holes, therefore the Q^+ value defined before con be evaluated by

$$Q^+ = N_D + \frac{J_p}{qv_{sp}} \tag{1.19}$$

but since the only carriers present into the device are holes, it can be written that $J_p = J_C$ and

$$Q^+ = N_D + \frac{J_C}{qv_{sp}} \tag{1.20}$$

From the latter equation it is clear that the phenomenon shown in turn-ON is enhanced and the breakdown voltage in turn-OFF becomes lower than that in turn-ON, limiting further the maximum Collector voltage.

1.6.3 SCSOA

When the Short-Circuit condition is achieved on the device, the maximum time it can stay in this condition defines the Short-Circuit Safe-Operating-Area. The power dissipated per unit of area during the Short-Circuit can be easily calculated since the current is the saturation current of the device ($J_{C,SAT}$) and the voltage is forced by the external circuit (V_CS). The power per unit of are a can be calculated by

$$P_D = J_{C,SAT} V_{CS} \tag{1.21}$$

In [39] it is demonstrated that the maximum temperature at which an IGBT can operate id 700K. At this temperature the high leakage current overcomes the built-in potential and the current con not be switched off anymore. If it supposed that the power is dissipated uniformly in the die, the maximum time which the device can withstand the Short-Circuit condition can be evaluated by [14]

$$t_{SCSOA} = \frac{(T_{CR} - T_{HS})\, W_{Si} C_V}{J_{C,SAT} V_{CS}} \tag{1.22}$$

where $T_{CR} = 700K$, T_{HS} is the heat sink temperature, W_{Si} is the thickness of the silicon wafer and C_V is the volumetric specific heat of the silicon. The relation highlights that the SCSOA increases by reducing the saturation current of the Collector or by reducing the Collector voltage. By using realistic values for the different parameters it has been shown that equation (1.22) gives an overestimation of the SCSOA because the power is not dissipated uniformly into the structure and the heat equations have to be taken into account to evaluate the temperature distribution along the device. By solving the temperature diffusion equation, the temperature distribution in the time and the space can be evaluated by

$$T(y,t) - T_{HS} = (R_T t)\, e^{-(y/L_t)} \tag{1.23}$$

where L_T is the thermal diffusion length ($40\mu m$ for silicon) and R_T can be evaluated by

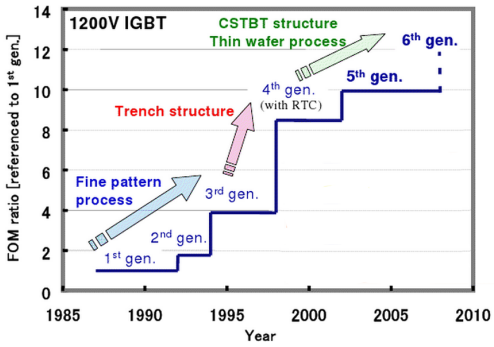


Figure 1.12: Development of the IGBT technology.

$$R_T = \frac{dT}{dt} = \frac{K_T J_{C,SAT} V_{CS}}{W_{Si} C_V} \tag{1.24}$$

where K_T keeps into account the non-uniformity of the temperature in the structure. The last equations give a good estimation of the of the SCSOA of the device, however it is only an estimation since the exact evaluation of the SCSOA of the device is strongly dependent on the design of the structure and an accurate evaluation can be obtained by means of Finite Elements Method (FEM) simulations.

1.7 State-of-art designs

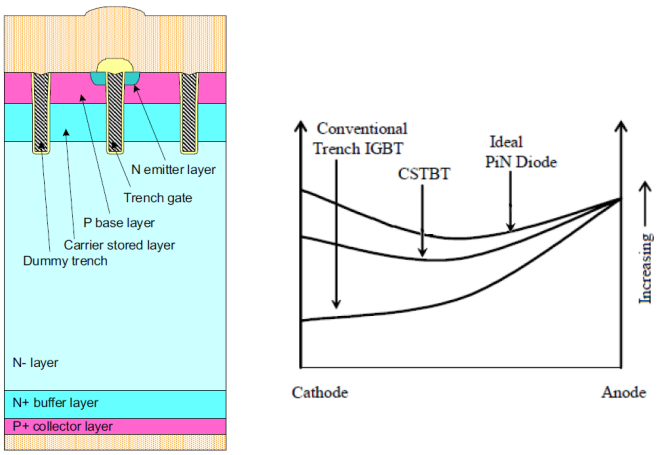
In this section a brief review of the different generations of IGBTs are reported together with the last design proposed to improve the performances.

As mentioned before, the great success of IGBTs is due to their high robustness and reliability in harsh conditions. It has to be pointed out that the success of the IGBT modules is also related to the strong effort to the development of smaller packages (e.g. [40]) and more reliable power modules. However the higher development has been achieved in the definition of new IGBT design with higher performances. In Fig. 1.12 the development of the IGBT technology is reported and the Figure Of Merit (FOM) is the comparative parameter, defined like

$$FOM = \frac{J_C}{V_{ON} E_{off}} \tag{1.25}$$

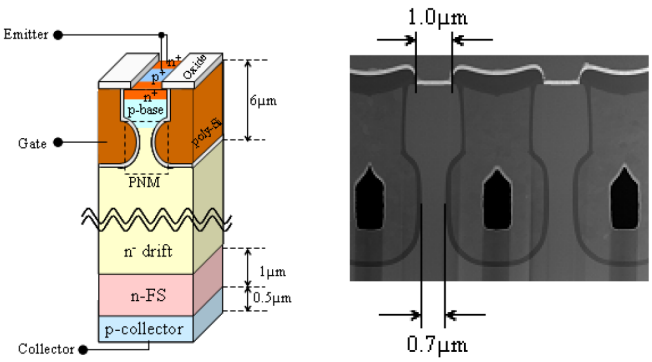
where the J_C is the device rated Collector current density, V_{ON} is the Collector voltage drop at the rated Collector current density and E_{off} is the turn-OFF switching energy at $T_j = 400K$. The first three generations were based on the Planar technology (See Fig. 1.1) and the main development have been focused on the improvement of the superficial structure. More in detail a finer surface pattern and shallow diffusion processing technologies have been adopted to improve the performances [41]. Basically a reduction of the V_{ON} is achieved by reducing the cell pitch and keeping constant the active area. In this way the channel width of the MOSFET part increases, then the current density increases as well. The main drawback of this solution is the enhancement of the effect of the parasitic JFET (in the Planar design the JFET region is located below the Gate between the two adjacent P-Body regions) which increases the voltage drop in on-state. A great improvement of the performances was achieved by adopting the Trench-IGBT design (See Fig. 1.10), passing from the third to the fourth generation[42]. This technology allowed to drastically reduce the V_{ON} because of the higher current density per unit of area. This means that the chip area was reduced compared to the same device in planar technology. However this solution has two main drawbacks: A reduction of the SCSOA and an increase of the Gate capacitance. The former problem was solved by increasing the cell pitch of the structure. In order to allow in a simple way the increase of the cell pitch, the PCM (Plugged Cell Merged) technology was developed to link alternatively the Gate polysilicon the Emitter contact. The technique reduces the Gate capacitance as well, however the V_{ON} increases again. In Fig. 1.13 the conventional Carrier Stored Trench Bipolar Transistor (CSTBT) structure is reported together with the carriers profile when the device is in on-state condition. The CSTBT design is compared to the Trench-IGBT and the P-i-n Diode structures. As it is visible, the PCM technology is adopted, but an N layer is added below the P-Body region. This additive layer increases the carrier density at the Cathode side and the V_{ON} decreases because of the higher conductivity modulation at the Cathode side. The last improvements for the CSTBT technology are reported in [43]

A recent improvement of the IGBT performances [44] was achieved from the structure reported in Fig. 1.14, called Partially Narrow Mesa (PNM) structure. The particular shape of the Trench-Gate allows an extreme injection enhancement which drastically reduces the V_{ON} of the device and improves the $V_{ON} - E_{off}$ trade-off.



(a) (b)

Figure 1.13: (a) Cross section of a conventional CSTBT structure (fifth generation) and (b) carriers profiles in on-state in the CSTBT structure compared to the Conventional Trench-IGBT structure and the P-i-n Diode.



(a) (b)

Figure 1.14: (a) Schematic 3D view of the PNM-IGBT and (b) SEM section of the structure.

Chapter 2

Simulation framework and strategy

The importance of an accurate silicon device design is increasing due to the necessity of achieving an optimized management of the energy into electronic systems. TCAD simulators are a well-established tool for both industry and research in the field of the semiconductor devices because of their advantages in supporting the development of new devices and the investigation of their failures. The predictive capability of TCAD simulators is well proved in CMOS and VLSI systems [45], [46] as well as in power semiconductor devices field [47], [48]. However, the advantages related to the adoption of TCAD simulators are subordinated to an accurate calibration of the physical models used in the numerical resolution of the semiconductor equations. In this chapter the Sentaurus TCAD[49] environment is briefly introduced by reporting the typical flow-chart of the simulations, together with the most relevant physical model adopted to perform the simulations. After the overview of the TCAD simulator, the strategy for the simulation of a multi-cellular IGBT is presented. Finally, a procedure, assisted by minimization algorithms of MATLAB©, for the electro-thermal calibration of Trench-Insulated Gate Bipolar Transistor (T-IGBT) elementary cell in Sentaurus TCAD is proposed. The procedure is based on the fitting of the simulated isothermal terminal curves ($I_C - V_{GE}$ and $I_C - V_{CE}$ characteristics) with the experimental ones.

2.1 TCAD simulators

The great progress in the field of the semiconductor technologies has allowed the production of devices with increasing performances in the time and the TCAD simulators have been one of the key factor of this growth[50]. The Technology Computer-Aided Design (TCAD) tool is referred to the use of computer simulations to develop and optimize semiconductor processing technologies and devices. TCAD tools are used by most of the semiconductor companies in the different technology steps of the development cycle. At the beginning of the technology development, TCAD tools allow to explore product design alternatives such as engineering the structure to enhance on-state performances or the turn-off losses even if the experimental data are not readily available. During the process integration stage, TCAD tools enable to do simulation split runs such as Design of Experiment (DOE) to comprehensively characterize and optimize the process, which saves time and money by reducing experimental runs on real wafers. Another great advantage of the TCAD simulator is the possibility to simulate characteristics which can not be experimentally measured (e.g. the I-V curve in avalanche condition of a small portion of device, especially in presence of NDR (See §4.1)).Different TCAD simulators are present on the market (e.g. Atlas[51]) right now and in this work the Synopsys TCAD software is adopted. It solves fundamental, physical partial differential equations, such as diffusion and transport equations, to model the structural properties and electrical behavior of semiconductor devices. This deep physical approach gives TCAD simulation predictive accuracy for a broad range of technologies. Therefore, TCAD simulations are used to reduce the costly and time-consuming test wafer runs when developing and characterizing a new semiconductor device or technology. TCAD tools are ideal to optimize process modules and integration by fully exploring the process parameter and, at the same time, the number of experimental wafers and development cycles are reduced.

2.1.1 Simulation flow-chart

In Fig. 2.1 the typical simulation flow-chart in Sentaurus TCAD is reported. With a TCAD simulator a real semiconductor device is represented in the simulator as a *virtual* device whose physical properties are discretized onto a nonuniform *grid* (or *mesh*) of nodes. Therefore, a virtual device is an approximation of a real device. The first step for the simulation of a structure in a TCAD simulator is the definition of the geometry. This can be made in differ-

ent ways. Actually one possibility is the definition of the geometry by means of process emulators in which the technological steps are emulated(e.g. diffusion, oxidation, ionic implantation and so on). This solution is very effective in reproducing the real structure, especially when the technological processes are calibrated. The main drawback is that usually the structure which is obtained by technological emulation can not be easily modified and a parametric analysis is difficult to perform since each variation of a parameter must be correlated in a correct way to different technological steps (usually the final doping profile is the superposition of the effect of different technological steps). A more convenient way to perform a parametric analysis of the design of a device is the reconstruction of the geometrical features from experimental characterizations (e.g. SEM sections of test structure) and the reconstruction of the doping profiles from both the experimental characterizations (e.g. by means of Spreading Resistance Profiling (SRP)) and the emulation of the technological steps. The doping reconstruction can be achieved by the superposition of elementary analytical doping profiles (e.g. Gaussian or Error-Function profiles). The latter solution gives a lower quality definition of the doping profiles, however once a suitable reconstruction of the doping profiles is achieved, the parametrization of the structure can be performed for both the geometry and the doping profiles. This approach has been followed to perform the parametric analysis in Chapter 3 and Chapter 4. The first step for the simulation of a device is the definition of the geometry and in Sentaurus TCAD this can be made by means of the *Sentaurus Structure Editor* software. In this environment the geometry, the materials and the doping profiles can be defined and 2D or 3D structures can be defined. As mentioned before, the Sentaurus TCAD solves the equations by means of the FEM, therefore a suitable discretization of the domain, the Mesh, is compulsory to achieve a predictive simulation in a reasonable time. In general the definition of a good mesh can be very time consuming, especially for 3D structures where the number of nodes can diverge to achieve a good predictability of the simulations. Therefore, it is compulsory to achieve a good balance between the reliability of the simulation and the number of nodes which define the simulation time. Some basic rules have to be adopted to improve the efficiency of the of the mesh. Basically, the mesh must be densest in those regions where a strong spatial variation of some physical parameters occurs. More in detail, the density of nodes must be increased where the following conditions are expected:

- High current density (MOSFET channels, bipolar base regions)

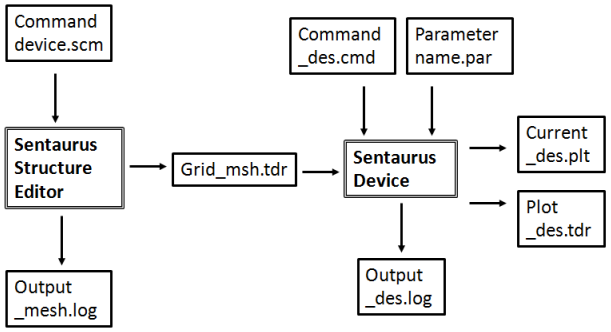


Figure 2.1: An overview of the simulation flow in Sentaurus TCAD.

- High electric fields (MOSFET channels, MOSFET drains, depletion regions in general)
- High charge generation (single event upset (SEU) alpha particle, optical beam).

In Sentaurus TCAD the meshing of the structure can be performed by means of the tools *Sentaurus MESH* and *Noffeset3D*. The former tool is a modular Delaunay (A mesh is a Delaunay mesh if the interior of the circumsphere (circumcircle for 2D) of each element contains no mesh vertices[49]) mesh generator, which generates high-quality spatial discretizations for complex 2D and 3D devices. It contains two mesh generation engines: an axis-aligned mesh generator. The axis-aligned mesh generator produces Delaunay meshes, which are suitable for use in *Sentaurus Device*. In 2D, the meshes contain triangles only, while in 3D the meshes comprise tetrahedra. Sentaurus Mesh reads the input geometry from a file stored in the TDR format. The latter tool, *Noffeset3D*, is a tool that constructs Delaunay meshes using surface-adapted, anisotropic, mesh layers. It uses an advancing front method to create layers of elements from designated surfaces and interfaces, and fills the remainder with an isotropic mesh. Noffset3D uses a sequence of algorithms to produce the final mesh. These two tools are integrated in *Sentuarus Structure Editor* and the overall structure, can be defined by means of script file with an extension **.scm*. The output of the *Sentuarus Structure Editor* is **msh.tdr* file and a **msh.log* file. The former file is the structure used to perform the simulations by means of *Sentaurus Device* while the latter one contains all of the information about the meshing procedure. Once the structure has been defined, the simulation is performed by means of *Sentaurus Device* which needs two files

to run a simulation: a **.cmd* file in which the simulation steps are defined and a **.par* file where all of the parameters of the different physical parameters are defined. *Sentaurus Device* simulates numerically the electrical behavior of a single semiconductor device in isolation or several physical devices combined in a circuit. The latter condition, named Mixed-mode, allows to simulate a single device together with an external circuit or more than one device connected to other devices or more than one devices connected to an external circuit. The elements of the external circuit are described by SPICE circuit models.Terminal currents, voltages, and charges are computed by means of a set of physical device equations that describes the carrier distribution and conduction mechanisms. The **.cmd* contains all of the commands to run a simulation. This file contains six sections in this sequence: *File*, *Electrode*, *Physics*, *Plot*, *Math* and *Solve*. In the *File* section the input and output files which are involved to perform the simulations are defined. In the *Electrode* section the electrodes (defined in *Sentaurus Structure Editor*) are assigned together with their physical properties (See §2.1.2). In the *Physics* section the models adopted to describe the physical phenomena are specified (See §2.1.2). In the *Plot* section are defined the physical parameters which have to be reported into the output files of *Sentaurus Device*, e.g. electron and hole densities, quasi-Fermi potentials, current densities and so on. In the *Math* section the settings for the solvers and convergence strategy are defined. Finally, in the *Solve* section the sequence of solutions to be obtained are defined. The outputs of a simulation are reported in both the **.tdr* file and **.plt* file. The former file reports for all the parameters (defined in the *Plot* section) into the structure in specific conditions (defined in the *Solve* section). Finally, in the **.plt* the specified terminal characteristics of the device are reported. The visualization of the terminal curves can be performed by *Sentaurus Visua*l or by *Inspect*, while the reading of the **.tdr* can be performed only by *Sentaurus Visual* (2D or 3D plots).

2.1.2 Physics

When an electrical device simulation have to be performed, at the very least must compute the electrostatic potential. Actually in all the semiconductor both the mobile charge (electrons and holes) and fixed charges (ionized dopants or traps) have a central role since they determine the electrostatic potential and at the same time they are affected by the electrostatic potential. When all of the contacts are at the same potential, the simplest simulation which can be performed needs the computation of the electrostatic potential

and the quasi-Fermi potentials. The solution of the Poisson equation gives the potential

$$\nabla \cdot (\epsilon \nabla \phi) = -q\,(p - n + N_D - N_A) - \rho_{trap} \tag{2.1}$$

where ϵ is the electrical permittivity, q is the elementary electric charge, n and p are the electrons and holes densities, N_D is the concentration of the ionized donors, N_A is the concentration of the ionized acceptors and ρ_{trap} is the charge due to the traps. If the Boltzmann statistics is assumed, the electron and holes densities can be computed by means of the quasi-Fermi potentials

$$n = N_C e^{\frac{E_{F,n} - E_C}{kT}} \tag{2.2}$$

$$p = N_V e^{\frac{E_V - E_{F,p}}{kT}} \tag{2.3}$$

where N_C and N_V are the effective density-of-states, $E_{F,n}$ and $E_{F,p}$ are the quasi-Fermi levels for electrons and holes, E_C and E_V are the conduction and the valance band edges. To keep into account the carriers transport into the semiconductors, the continuity equations which describe the charge conservation are written in the following form

$$\nabla \cdot \boldsymbol{J_n} = qR_{net} + q\frac{\partial n}{\partial t} \tag{2.4}$$

$$-\nabla \cdot \boldsymbol{J_p} = qR_{net} + q\frac{\partial p}{\partial t} \tag{2.5}$$

where R_{net} is the net recombination rate, $\boldsymbol{J_n}$ is the electron current density and $\boldsymbol{J_p}$ is the hole current density. In this system of equations the unknown terms are ϕ, n and p. The different transport models adopted in the field of the semiconductor devices differ for the equations adopted to compute $\boldsymbol{J_n}$ and $\boldsymbol{J_p}$. Depending on the features of devices which has to be simulated, the following transport models can be used [49]:

- *Drift-Diffusion* - Isothermal simulation, suitable for low-power density devices with long active regions.
- *Thermodynamic* - The self-heating effect is kept into account. It is suitable for devices with low thermal exchange and in particular for high-power density devices with long active regions.

- *Hydrodynamic* - This model accounts for energy transport of the carriers. It is suitable for devices with small active regions.
- *Monte Carlo* - It solves the Boltzmann equation for a full band structure.

The power semiconductor devices are characterized by structures which can be well simulated by adopting the first two models since they are devices with long active regions. In the following a brief review of the Drift-Diffusion and Thermodynamic models are reported together with the mobility and the generation/recombination models adopted for the simulations performed in Chapter 3 and Chapter 4.

Drift-Diffusion model

For the Drift-Diffusion model the current densities of the electrons and holes are given by:

$$\boldsymbol{J_n} = \mu_n \left(n \nabla E_C - 1.5 n k T \nabla ln m_n\right) + D_n \left(\nabla n - n \nabla ln \gamma_n\right) \tag{2.6}$$

$$\boldsymbol{J_p} = \mu_p \left(p \nabla E_V - 1.5 p k T \nabla ln m_p\right) - D_p \left(\nabla p - p \nabla ln \gamma_p\right) \tag{2.7}$$

These equations are general and are valid for both the Boltzmann and Fermi statistics. The first term takes into account the contribution due to the spatial variations of the electrostatic potential, the electron affinity, and the band gap. The remaining terms take into account the contribution due to the gradient of concentration, and the spatial variation of the effective masses m_n and m_p. Since the Boltzmann statistic will be adopted for the simulation of power devices, γ_n and γ_p are equal to one. The diffusivities, D_n and D_p are calculated by means of the Einstein relation

$$D_n = \frac{kT\mu_n}{q} \tag{2.8}$$

$$D_p = \frac{kT\mu_p}{q} \tag{2.9}$$

However, more complicate expressions for the diffusivities can be adopted to simulate particular kind of devices. In the power devices the Einstein relations are suitable for the description of the behavior and the last two relations can be simplified in the following form

$$\boldsymbol{J_n} = -nq\mu_n \nabla \Phi_n \tag{2.10}$$

$$\boldsymbol{J_p} = -pq\mu_p \nabla \Phi_p \tag{2.11}$$

where Φ_n and Φ_p are the electron and hole quasi-Fermi potentials.

Thermodynamic model

In the thermodynamic model [52] the equations (2.10) and (2.11) are generalized to include the temperature gradients like a driving term for the carriers:

$$\boldsymbol{J_n} = -nq\mu_n \left(\nabla \Phi_n + P_n \nabla T\right) \tag{2.12}$$

$$\boldsymbol{J_p} = -pq\mu_p \left(\nabla \Phi_p + P_p \nabla T\right) \tag{2.13}$$

where P_n and P_p are the absolute thermoelectric powers [53] and T is the lattice temperature.

Mobility and Generation/Recombination models

The generation/recombination phenomenon and carriers Mobility are the most important physical parameters which affect the carriers dynamic into a semiconductor device and an accurate evaluation of them is compulsory to achieve a reliable numerical simulation of the device.

- *Recombination*

  Generation-recombination processes imply the exchange of carriers between the conduction band and the valance band. The carriers Recombination affects most of the characteristics of a semiconductor device and an accurate evaluation of the lifetime into the structure becomes crucial when Lifetime Killing techniques are adopted[54]. The R_{net} parameter in (2.4) and (2.5) keeps into account the Recombination rate in each point of the structure and it evaluated by the Shockley-Read-Hall (SRH) relation

$$R_{net}^{SRH} = \frac{np - n_{i,eff}^2}{\tau_p \left(n + n_1\right) + \tau_n \left(p + p_1\right)} \tag{2.14}$$

where

$$n_1 = n_{i,eff}^2 e^{\frac{E_{trap}}{kT}} \tag{2.15}$$

and

$$p_1 = n_{i,eff}^2 e^{\frac{-E_{trap}}{kT}} \tag{2.16}$$

where E_{trap} is the difference between the defect level and the intrinsic level and $n_{i,eff}$ is the effective intrinsic concentration. τ_n and τ_p are the minority carriers lifetime at high injection levels and they are dependent on the temperature, the electric field and the doping concentration. In power semiconductor devices, the temperature effect and the doping effect are the only relevant for the a correct simulation of the device. The temperature dependence is expressed by an exponential law

$$\tau\left(T\right) = \tau_0 e^{C\left(\frac{T}{300K}-1\right)} \tag{2.17}$$

while the doping dependence is modeled by the Scharfetter relation [55], [56], [57]

$$\tau_{dop}\left(N_{A,0}+N_{D,0}\right) = \tau_{min} + \frac{\tau_{max}-\tau_{min}}{1+\left(\frac{N_{A,0}+N_{D,0}}{N_{ref}}\right)^{\gamma}}. \tag{2.18}$$

where $N_{A,0}$ is the acceptor concentration, $N_{D,0}$ is the donor concentration and τ_{max} and τ_{min} are the maximum and the minimum lifetime, respectively.

The Auger recombination effect [58] also contributes to the lifetime of the device and it is particularly relevant at high carriers density. Therefore it is relevant when the extrinsic recombination effects are extremely low. The recombination contribute due to the Auger effect is expressed like

$$R_{net}^A = \left(C_n n + C_p p\right)\left(np - n_{i,eff}^2\right) \tag{2.19}$$

where C_n and C_p are the temperature-dependent Auger coefficients.

- *Generation*

  The generation of carriers in semiconductor devices is due to two main phenomena: the thermal generation and the generation by impact ionization. The former phenomenon is mainly responsible for the leakage current of a reversely biased junction. When a p/n junction is reversely polarized, in the depletion region (for low current densities) the mobile charges are negligible. Therefore in the SRH model $n = p = 0$ and

$$R_{net}^{SRH} = \frac{-n_{i,eff}^2}{\tau_p (0 + n_1) + \tau_n (0 + p_1)} \tag{2.20}$$

  The Recombination has a negative sign, therefore it can be addressed like a generation ($-R_{SRH} = U$) due to the temperature since [59]

$$U = \frac{n_i}{2\tau_{sc}} \tag{2.21}$$

  where τ_{sc} is the space-charge generation lifetime and

$$\tau_{sc} = \left\{ exp\left(\frac{E_t - E_i}{kT}\right)\right\}\left(\frac{\tau_{p0}}{2}\right) + \left\{ exp\left(\frac{E_i - E_t}{kT}\right)\right\}\left(\frac{\tau_{n0}}{2}\right) \tag{2.22}$$

  From the last equation it is clear that the generation lifetime is not dependent on the doping concentration in the depletion region.

  As mentioned before the other generation contribution is due to the impact ionization phenomenon which occurs when the electric field reaches high values and the avalanche breakdown voltage occurs. This therm contributes to the evaluation of R_{net} in in (2.4) and (2.5) as well. More in detail the avalanche breakdown occurs whether the width of a space charge region is greater than the mean free path between two ionizing impacts. The reciprocal of the mean free path is called the ionization coefficient α. With these coefficients for electrons and holes, the generation rate can be expressed as:

$$G_{ii} = \alpha_n n \nu_n + \alpha_p p \nu_p \tag{2.23}$$

  Different models are adopted in literature to evaluate the ionization coefficients [49] and in the simulations performed in Chapter 3 and Chapter

4, University of Bologna impact ionization model [60] has been adopted. The generation coefficient is evaluated by

$$\alpha\left(F_{ava},T\right)=\frac{F_{ava}}{a\left(T\right)+b\left(T\right)exp\left[\frac{d(T)}{F_{ava}+c(T)}\right]} \tag{2.24}$$

where F_{ava} is the driving force (which can be the electric field in the direction of the current flux), a, b, c and d are polynomials of T:

$$a\left(T\right)=\sum_{k=0}^{3}a_k\left(\frac{T}{1K}\right)^k \tag{2.25}$$

$$b\left(T\right)=\sum_{k=0}^{10}b_k\left(\frac{T}{1K}\right)^k \tag{2.26}$$

$$c\left(T\right)=\sum_{k=0}^{3}c_k\left(\frac{T}{1K}\right)^k \tag{2.27}$$

$$d\left(T\right)=\sum_{k=0}^{3}d_k\left(\frac{T}{1K}\right)^k \tag{2.28}$$

- *Mobility*

  The Drift-Diffusion model is widely adopted also because of the great number of mobility models which keep into account different parameters which can degrade it. The simplest model for the mobility is the temperature dependence, where the mobility have a constant value which changes only whit temperature (only the phonon scattering is kept into account). However the mobility depends also on: doping concentration, carrier-carrier scattering [61], degradation at the interface and on the electric field. Each dependence is modeled with an expression which gives a mobility value and the combination of the different models is achieved by the Matthiessen's rule

$$\frac{1}{\mu}=\frac{1}{\mu_1}+\frac{1}{\mu_2}+\frac{1}{\mu_3}+... \tag{2.29}$$

The doping dependence of the mobility has been kept into account by using the Arora's model[62]:

$$\mu_{dop} = \mu_{min} + \frac{\mu_d}{1 + \left(\frac{N_{A,0}+N_{D,0}}{N_0}\right)^{A^*}} \tag{2.30}$$

$$\mu_{min} = A_{min} \cdot \left(\frac{T}{300K}\right)^{\alpha_m} \tag{2.31}$$

$$\mu_d = A_d \cdot \left(\frac{T}{300K}\right)^{\alpha_d} \tag{2.32}$$

$$N_0 = A_N \cdot \left(\frac{T}{300K}\right)^{\alpha_N} \tag{2.33}$$

$$A^* = A_a \cdot \left(\frac{T}{300K}\right)^{\alpha_a} \tag{2.34}$$

The main advantage of this model is the easy calibration.

The dependence of the Mobility on the electric-field has been kept into account by the Canali model [63]. Basically when the electric field becomes high, the carrier drift velocity is no longer proportional to the electric field and saturates at a velocity v_{sat}. The electric field dependence of the mobility is evaluated by

$$\mu(F) = \frac{(\alpha+1)\,\mu_{low}}{\alpha + \left[1 + \left(\frac{(\alpha+1)\mu_{low}F_{hfs}}{\nu_{sat}}\right)^{\beta}\right]^{1/\beta}} \tag{2.35}$$

where μ_{low} is the mobility at low electric field, ν_{sat} is the saturation velocity, F_{hfs} is the driving force and β is a fitting parameter which has a temperature dependence modeled like:

$$\beta = \beta_0 \left(\frac{T}{300K}\right)^{\beta_{exp}}. \tag{2.36}$$

Boundary conditions

In order to correctly solve differential equations which describe the semiconductors and to define an unique solution, a correct definition of the boundary conditions is compulsory. By default contacts on the semiconductors are Ohmic with a resistance of 0.001Ω when connected to a circuit node. The equilibrium and the charge neutrality are assumed to the Ohmic contact, therefore:

$$n_0 p_0 = n_{i,eff}^2 \tag{2.37}$$

$$n_0 - p_0 = N_D - N_A \tag{2.38}$$

When the Boltzmann statistics is applied, these conditions become:

$$\phi = \phi_F + \frac{kT}{q} asinh\left(\frac{N_D - N_A}{2n_{i,eff}^2}\right) \tag{2.39}$$

and

$$n_0 = \sqrt{\frac{(N_D - N_A)^2}{4} + n_{i,eff}^2} + \frac{N_D - N_A}{2}, \tag{2.40}$$

$$p_0 = \sqrt{\frac{(N_D - N_A)^2}{4} + n_{i,eff}^2} - \frac{N_D - N_A}{2} \tag{2.41}$$

where n_0 and p_0 are the electron and hole concentration at the equilibrium and ϕ_F is the Fermi potential at the contact. By default $n = n_0$ and $p = p_0$ is assumed at the contact, however, when shallow doping profiles are adopted, it can be necessary to specify the electron or hole recombination velocity at the interface and the following conditions are assumed:

$$\boldsymbol{J_n} \cdot \hat{n} = q\nu_n (n - n_0) \tag{2.42}$$

$$\boldsymbol{J_p} \cdot \hat{n} = -q\nu_p (p - p_0) \tag{2.43}$$

where $\hat{n}$ is the unit vector in the direction of the outher normal. For contact on insulators, like Gate contacts, the electrostatic potential is computed like:

$$\phi = \phi_F - \Phi_{MS} \tag{2.44}$$

where ϕ_F id the Fermi potential and Φ_{MS} is the workfunction difference between the metal and an intrinsic reference semiconductor.

For all of the boundaries that are not contacts, the ideal Neumann boundary conditions are applied:

$$\epsilon \nabla \phi = 0 \tag{2.45}$$

and

$$\boldsymbol{J_n} \cdot \hat{n} = 0, \boldsymbol{J_p} \cdot \hat{n} = 0. \tag{2.46}$$

From the thermal point of view, for free thermally insulated surfaces the following condition is applied:

$$\kappa \hat{n} \cdot \nabla T = 0 \tag{2.47}$$

while in thermally conductive interfaces the thermally resistive boundary conditions are adopted:

$$\kappa \hat{n} \cdot \nabla T = \frac{T_{ext} - T}{R_{th}} \tag{2.48}$$

Finally, for the simulations of periodic device structures, for example lines or arrays of cells, the periodic boundary conditions (PBC) can be used. Instead of building a periodic simulation domain, the PBC can be applied to the vertical boundaries of the device. In this work the Mortar PBC approach is adopted [49]. In the MPBC approach, both sides of the PBC interface are effectively glued together on one side (the mortar side), while on the other side (the nonmortar side), a weak continuity condition for the equation potential (mostly, the solution variable) is imposed. Therefore, the MPBC approach is not symmetric with respect to the selection of the mortar side if the mesh is nonconforming. The side where the accuracy requirements are larger (and, typically, the mesh is finer) should be selected as the mortar side. The equation potential for the carrier continuity equations are the corresponding quasi-Fermi potentials, while for the other equations, the solution variable is chosen. The MPBC approach does not require a current model across the PBC interface.

2.2 Simulation strategy for multi-cellular structures

As mentioned in the previous Chapter, the IGBTs are organized like multi-cellular devices, where an elementary cell is repeated and extruded. The multi-

cellular structure of the device allows to treat the lateral boundary of the elementary cell with reflective (or ideal Neumann) boundary conditions and, from the thermal point of view, the they are treated with adiabatic conditions §2.1.2. If the structure does not show a multi-cellular design, the latter boundary condition gives an overestimation of the temperature in the device [64] and a non uniform current distribution can occur [65]. To achieve a good thermal calibration of the elementary cell it is mandatory to specify accurately the thermal boundaries of the elementary cell at the extremities [66],[67]. In particular, the boundary conditions at the two extremities of the cell have to be specified by accounting for the terminal structure (metal layers and solder layer) in the simulated geometry or by reporting the equivalent thermal network at the terminations. For the electrical terminals of the cell the hypothesis of Ohmic contact is adopted.

2.3 Electrical calibration of the elementary cell on IGBT

The calibration procedure is based on the fitting of both the simulated $I_C - V_{GE}$ (in sub-threshold region and at high current level) and $I_C - V_{CE}$ characteristics on the experimental ones, at two temperatures. In the following part, the parameters chosen to fit the experimental curves, for the different models, are reported. As shown in equation (1.1), the sub-threshold current slope is due to the ratio between the oxide capacitance and the depletion layer capacitance. In order to fit the slope of the sub-threshold characteristic (in semi-logaritmic scale), the oxide capacitance, which depends on both the oxide permittivity and its thickness, has been varied. To change the oxide capacitance, the oxide thickness is used like variable parameter to fit the sub-threshold slope. Since the threshold voltage is dependent on both the oxide thickness and the oxide fixed oxide charge, an iterative approach has been adopted to achieve the value of oxide thickness and fixed charge at the interface which fit the simulated curves on the experimental ones. When the gate voltage exceeds the threshold voltage, the collector current can be evaluated by using equation (1.2). It is essential to point out that, in pinch-off regime, the collector current is mainly regulated by the electron mobility in the channel together with the current gain of the vertical PNP (mainly dependent on the minority carriers lifetime in the structure). To fit the experimental $I_C - V_{GE}$ characteristics, all of the hole parameters are left the default ones, while the following parameters have been chosen for electrons (and the corresponding temperature scaling param-

eters): N_0 in equation (2.30) together with β in equation (2.35). Finally, the threshold voltage is fitted by changing the oxide fixed charge concentration (The calibration procedure is applied to electron irradiated devices). Temperature dependence of N_0 and β has been considered by adopting equation (2.33) and equation (2.36), respectively. Lifetime killing techniques are extensively used [54] to enhance dynamic features in power silicon devices to reduce the switching losses, due to the stored charge in the Drift-layer. Depending on the lifetime killing technique, different energy levels into the forbidden band are introduced, making more complex a modeling of them on the carriers recombination. However, silicon power devices mainly work at high injection levels for minority carriers and, in most of the cases, a dominant energy level (E_{trap}) regulates the recombination rate in the structure [68]. If no lifetime-killing techniques are adopted, the trap energy can be set at $E_t = E_i$. In this work it is assumed that the effect of the lifetime killing treatment introduces a dominant energy level in the forbidden band and [59] has been adopted for the recombination rate in the IGBT structure. As shown in equations (1.7) and (1.8), the minority carriers lifetime strongly affects the voltage drop in the on state, therefore τ_{p0} and τ_{n0} in [59] are chosen to fit the experimentally measured voltage drop of the IGBT structure. The ratio between τ_{p0} and τ_{n0} is proportional to capture cross section and it is available from the literature (i.e. [54]). The doping dependence of the minority carrier lifetime (for both electrons and holes) has been kept in account by using the Scharfetter model (equation (2.18)) and the reference doping concentration (N_{ref}) has been calibrated. The temperature dependence has been calibrated by fitting C in (2.17). Finally, the leakage current at $V_{GE} = 0V$, in the $I_C - V_{GE}$ curve ($V_{CE} = 10V$), is fitted by changing the trap energy level of the recombination center in, E_t in [59].

In Fig. 2.2 the flow chart of the calibration procedure is reported. After the two initial steps in which the elementary cell structure is defined and the device are characterized, the sub-threshold characteristic at T = $25\,^{\circ}\mathrm{C}$, in the linear region, is fitted acting on the oxide thickness and the fixed oxide charge concentration. In order to reduce the time to achieve the calibration of the structure, a minimum finder of unconstrained multivariable function that uses the derivative-free method (the MATLABs function *fminsearch*) has been adopted to manage the fitting procedure of the different curves. A function which is able to face with the TCAD environment has been implemented. More in detail, starting from an initial guess for the fitting parameters (specific for each curve) the function inserts them into the parameter file of the TCAD simulator and launches the simulation. When the simulation is terminated, the error

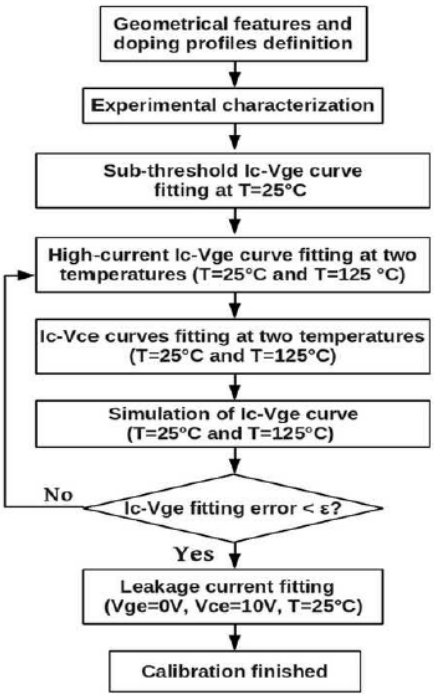


Figure 2.2: Flow-chart of the calibration procedure.

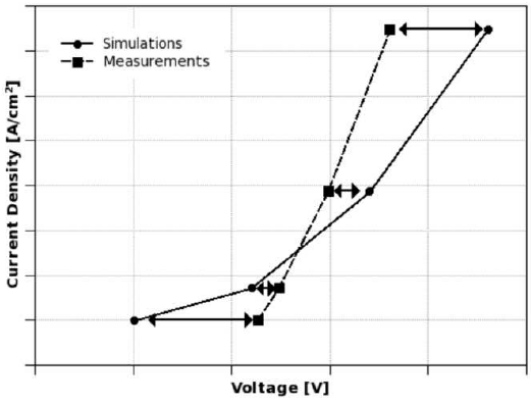


Figure 2.3: The modality of calculation of the error is sketched.

between the simulated and the experimental curve is the output of the function and it becomes the next input of the minimum finder algorithm. At this point the minimum finder algorithm provides the next vector of parameters for the simulator. This iterative procedure finishes when the *TolX* parameter of the *fminsearch* function is achieved. The error, for each curve is evaluated by considering the root-mean-square of the differences between the simulated and the measured curves on the X axis (See Fig. 2.3).

The root-mean-square is calculated by considering the difference between the abscissas of the points at the same current level. This modality has been adopted to optimize the calibration procedure, especially when the sub-threshold characteristic is fitted. More in detail the error has been evaluated:

1. In the linear region of the Sub-threshold $I_C - V_{GE}$ characteristic (in semi-logaritmic scale).

2. For all of the points of the High-current $I_C - V_{CE}$ characteristic.

3. In the linear region of the $I_C - V_{CE}$ characteristic.

As mentioned in the previous section, both the $I_C - V_{GE}$ and the $I_C - V_{CE}$ characteristics are dependent on the inversion layer mobility and the minority carrier lifetimes. Therefore the fitting of one of the two curves can be achieved with different combinations of mobility and lifetime values. To overcome to this problem, both the $I_C - V_{GE}$ and the $I_C - V_{CE}$ curves have been used to calibrate the two parameters, at two temperatures, by using an iterative approach. Starting from an initial set of electron mobility and lifetime values,

the $I_C - V_{GE}$ characteristics fitting is performed at T = 25 °C and T= 125 °C, by acting on the inversion layer mobility; first the curve is fitted at T = 25 °C varying N_{ref} and β, then the relative temperature scaling parameter is fitted at T= 125 °C. In the next step the $I_C - V_{CE}$ curve is fitted at T = 25 °C and T= 125 °C by acting on the minority carriers lifetime (τ_{p0} and τ_{n0}) in the structure. More in detail, starting from the electron mobility reached in the previous step, τ_{p0} and τ_{n0} are fitted at T = 25 °C, then the relative temperature scaling parameters are fitted at T= 125 °C. Since the minority carrier lifetime variation of the second step leads to a variation of the $I_C - V_{GE}$ characteristic, in the next step a simulation of the $I_C - V_{GE}$ characteristic, with the new parameters achieved in the two previous steps, is performed and the error is evaluated. Therefore, Each iteration of the procedure is composed of three steps and it finishes at the iteration **k** when the error at the third step is lower than ϵ, which is the error at the first step of the iteration **k** multiplied by 2. The ϵ value defines the accuracy of the calibrated structure. The last step of the procedure is the leakage current calibration, which is achieved by acting on the energy level of the trap.

Chapter 3

Design influence on the Short-Circuit capability

The increasing demand of power semiconductor devices with an high reliability, especially in application fields like automotive or high power converters, has increased the necessity of a deep comprehension of the failure mechanisms which lead to a possible premature destruction of the device in both nominal and fault conditions. The Short-Circuit is one of the most critical conditions which a device can be subjected to and it is of great interest the investigation of the features of the device structure which can enhance the capability in this extreme condition. Basically two types of fault conditions are possible: 1) the Short-Circuit occurs *during* the on-state, 2) the Short-Circuit condition is already present when the D.U.T. is turned on [69], [70]. Both conditions are very stressful for the device and in this chapter the latter condition is investigated by means of TCAD simulations. Basically, the Short-Circuit condition is very harsh because of the high voltage and high current that at the same time occur in the device which, by means of the Joule effect, lead to a quick increase of the temperature into the structure (with a distribution dependent on the geometry of the device) [71], [66], [72], [67], [73]. A correct TCAD simulation of the Short-Circuit test basically requires the coupling of the electrical and thermal equations which describe the semiconductor physics, together with a pondered definition of structure which which will be simulated. Actually, in this chapter the investigated structure is multi-cellular and the boundary conditions allow to define an elementary cell which can be considered representative of the overall device in on-state (§2.3). In this chapter the Short-Circuit capability of a Punch-Through IGBT will be investigated.

First the Device under test will be experimentally characterized and, by means of the TCAD simulator, the failure phenomenon will be marked. Finally the an optimized structure is proposed and characterized from the point of view of the electrical performances by means of the TCAD simulations and some experimental characterization on the optimized samples will be shown.

3.1 TCAD structure definition and calibration

As mentioned before the short-Circuit condition is very harsh and it can lead to the destruction of the device before the triggering of the protection circuit. The Short-Circuit capability is strongly dependent on the features of the structure and, more in detail both the geometrical features and the doping profiles of the elementary cell affects the behavior of the device. Therefore, an accurate calibration of the structure is needed for a quantitative evaluation of the device operation in harsh conditions. In this section the calibration procedure presented in §2.3 is adopted to calibrate a Punch-Through IGBT structure with 600 V of Breakdown Voltage.

3.1.1 Structure definition

The analysis and development of a PT-IGBT structure rugged in Short-Circuit condition has been tackled starting from a latch-up free structure, provided by a leader company in the field of the power semiconductor devices. The device has a Blocking voltage of 600 V. The aim of this section is the definition of the elementary cell and its calibration. In order to adopt the calibration procedure presented in (§2.3) a detailed characterization of the elementary cell is needed to extract the geometrical features and the doping profiles. These information have been extracted by both the Scanning Electron Microscopy(SEM) section and the Spreading Resistance Profiling (SRP) techniques. In Fig. 3.1 the SEM section of the top of the device is reported. Since the total active area can be obtained by repeating the structure in Fig. 3.1 and extruding it, the symmetry allows the definition of the elementary cell as highlighted by the rectangle. The particular symmetry allows to define an elementary cell which has a width of half pitch and this allows to work with a TCAD model with a lower number of nodes as it will be shown in the following of this work. The above structure is known as Self Aligned Structure [74].From the geometrical point of view, the Gate-Trench and the Contact-Trench can be discerned: the former stays no the right side of the rectangle and the latter is on the opposite side. Therefore,

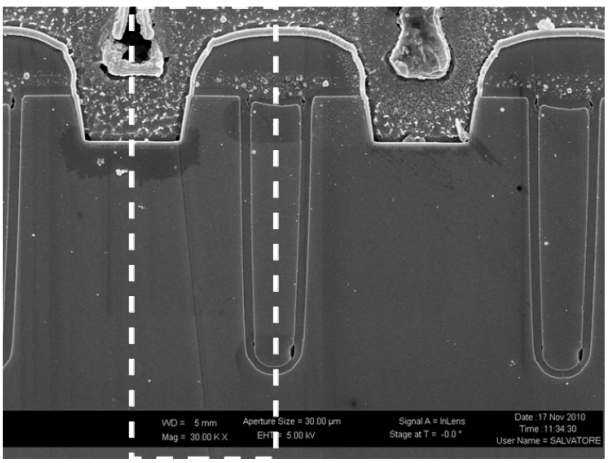


Figure 3.1: SEM section of the analyzed structure.

the geometrical dimensions of the elementary cell are extracted by means of this measurement. In this structure, then, a Contact-Trench is adopted and its main role is the suppression of the latch-up phenomenon [75]. The main advantage of a deep contact-trench is the reduced path for the holes which moves toward the Emitter contact passing below the Emitter diffusion. When the Contact-Trench is not present, a longer path is covered by the holes to achieve the silicon/metal interface, with an higher voltage drop at the P-Body/Emitter junction. By means of simulations it will be shown that a deep Contact-Trench dramatically reduces the hole flow just below the P-Body/Emitter junction because of the very high recombination rate of the silicon/metal contact. However, as it will be shown in the next chapter, this solution increases the latch-up immunity but, on the other hand, it reduces the reliability of the overall device in avalanche conditions. Because of the Punch Through design, the rest of the structure is composed by a low doped N-Drif layer epitaxially grown on a P-type wafer. Therefore, the geometrical features of the elementary cell are defined by the above characterization, but a confirmation of these data is given also from the analysis of the mask layouts of the technological steps involved to realize the devices.

The definition of the doping profiles is a more complicate task. As mentioned in (§2.3), the doping profiles are obtained or by SRP measurements, or by process emulators. Most of the information about the doping profile of the elementary cell has been carried out by means of the process emulator

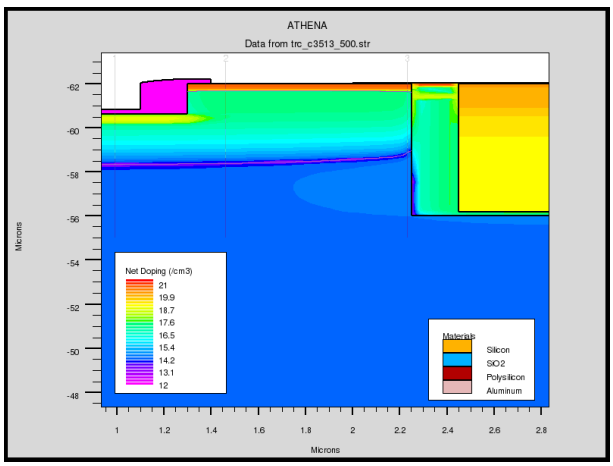


Figure 3.2: Doping profile map of the elementary by means of the ATHENA process emulator.

ATHENA® (**Silvaco** international). In Fig. 3.2 the doping profiles of the elementary cell is obtained by the emulation of the process involved in the device realization. The axis scale allows a better visualization of the entire structure and as it has been shown before the Contact-Trench and the Trench-Gate are well visible. The main difference between the SEM section and the emulated structure is the bottom of the Gate-Trench which is squared in the emulated one, but from the point of view of the on-state characterization this difference is not relevant. From the technological point of view, the P-Body diffusion is made before than the Trench-Gate structure. This temporal sequence is well visible if the P-Body doping profile is observed in the proximity of the Silicon/Oxide interface; because of the Boron segregation [76] into the oxide, the Boron concentration reduces along horizontal direction, close to the interface. This effect generate a curvature of the P-Body/Drift junction in the proximity of the Silicon/Oxide interface. Together with the process emulation tool, the SRP data have been used to define the final doping profile of the structure. Two SRP measurements have been provided to evaluate more in detail both the P-Body and the back of the device (Drift, Buffer-Layer and Collector region). In Fig.3.3 the SRP profile of the P-Body diffusion is reported, while in Fig. 3.4 The Drift layer, the Buffer Layer and the collector region are visible, respectively starting from the left.

As mentioned before, the doping profile by SRP has been divided in two

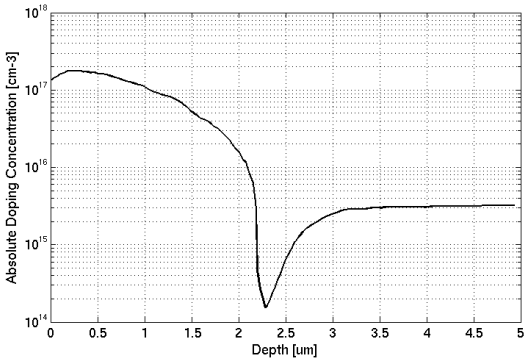


Figure 3.3: SRP profile of the P-Body diffusion.

step in order to get an higher resolution for the two specific regions. Actually the P-Body region is not well defined in Fig.3.4 but, because of the higher width of the back diffusion, the overall profile is well defined.

It has to be pointed out that the SRP measurements are affected by an high uncertainty which force an accurate evaluation of the final doping profile which has to be adopted for the device simulation. More in detail the overall doping profile has been obtained mainly from the results of the process emulator and the SRP measurements have been used to check the accuracy and the coherence of the emulated profiles. The Doping profile of the elementary cell structure recreated in the TCAD environment has been obtained like the superposition of the Gaussian or error functions. The reconstructed profiles are reported in Fig. 3.5 and 3.6.

In Fig. 3.5 the reconstructed structure is reported. From the geometrical point of view the main difference between the real structure and the one reported in figure is the slope of the vertical trench: while the real structure shows a small reduction of the width from the top to the bottom, the reconstructed one has a vertical Gate-Trench. However this geometrical difference has not a relevant impact on the simulations results in both on-state and avalanche condition. On the other hand, the curvature and the oxide thickness of the bottom of the Gate-Trench has a great influence when carriers generation due to impact ionization occur. Therefore, an accurate reconstruction of this region of the device has been done. In Fig. 3.6 the absolute doping profile along the red line in Fig. 3.5 is reported. The four regions of the IGBT are well visible and, starting from the left side, the following layers are visible: Emitter, P-Body, Drift, Buffer, and Collector. Depending on the analysis, both the 2D and the

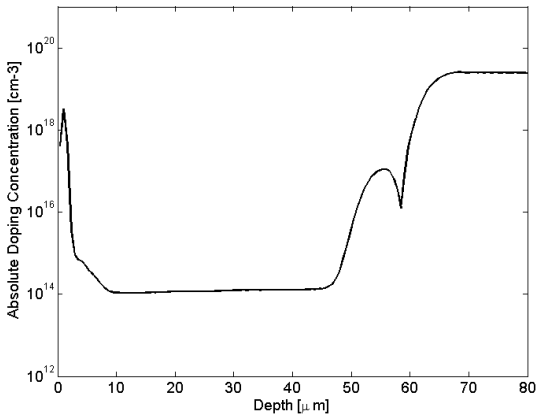


Figure 3.4: SRP profile of the overall structure.

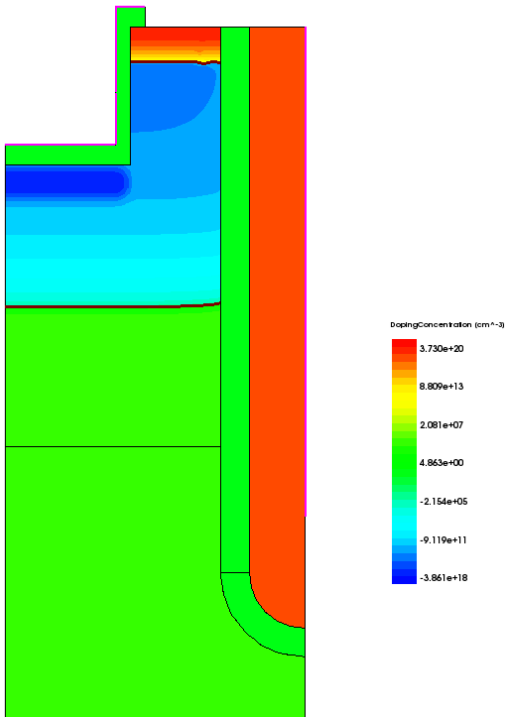


Figure 3.5: 2D PT-IGBT Elementary cell implemented in Sentaurus TCAD.

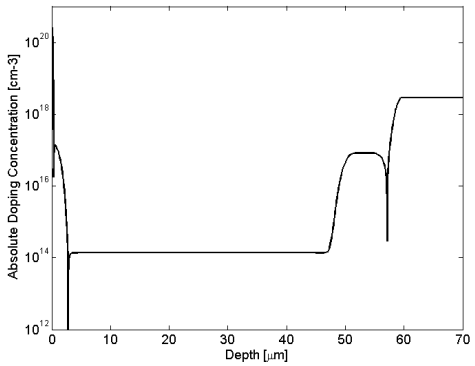


Figure 3.6: Doping profiles of the overall structure.

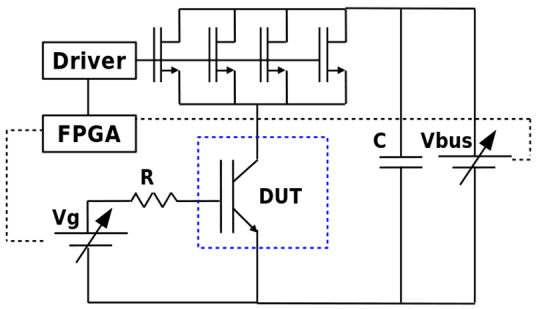


Figure 3.7: Experimental setup of the pulsed curve-tracer.

3D structures have been simulated.

3.1.2 Terminal characteristics

By means of the experimental setup illustrated in Fig. 3.7, the terminal characteristics of the device have been measured at two temperatures (T = 25 °C and T= 125 °C). The pulsed regime is compulsory since the self-heating effect during a DC measurement would make the measurements not isothermal. The experimental setup in Fig. 3.7 is a custom setup and it work in the following way: the Gate voltage is fixed by the generator V_g, the Collector voltage is fixed by the V_{bus} generator and the V_{bus} voltage is applied in a controlled way by the four power MOSFET in parallel. The timing and the control of both the generators and the MOSFET driving is managed by an FPGA which is con-

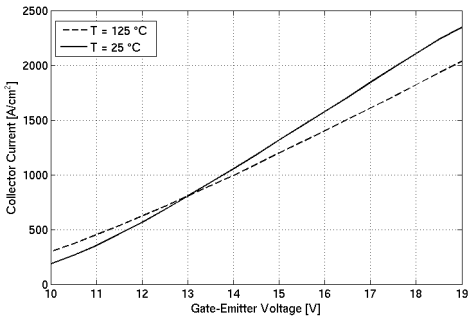


Figure 3.8: Experimental transfer characteristics of the device at two temperatures (T = $25\,^{\circ}\mathrm{C}$ and T= $125\,^{\circ}\mathrm{C}$) with $\mathrm{V}_{CE} = 10V$.

figurable by a PC. The $I_C - V_{GE}$ sub-threshold characteristic, because of the low currents, is not affected by self-heating effect and it has been performed by using a signal-devices curve tracer (handled by an HP4142B parameter analyzer). In Fig. 3.8 the transfer characteristics of the device are reported for T = $25\,^{\circ}\mathrm{C}$ and T= $125\,^{\circ}\mathrm{C}$. The Temperature Compensation Point (TCP) of the device has a value of 13.2 V and, since it is lower than the operative value of the Gate-Emitter voltage in on-state (15 V), the device exhibits a negative temperature coefficient for the saturation current which implies the temperature stability of the device and the possibility of paralleling different devices. Therefore it can be asserted that in this region the behavior of the device is mainly regulated by the MOS part of the structure [77]. In Fig. 3.9 and in Fig. 3.10 the output curves of the device are reported, at T = $25\,^{\circ}\mathrm{C}$ and T= $125\,^{\circ}\mathrm{C}$ respectively, at different Gate-Emitter voltages. In Fig. 3.9 the self-heating effect is visible, since the Collector current density decreases when the Collector voltage increases (the more the collector voltage increases, the more the energy dissipated during the tester pulse is higher, with the consequent reduction of the mobility in the MOS channel). In Fig. 3.11 the schematic of the short-circuit test together with the waveforms of device in short-circuit conditions are reported.

3.1.3 Structure calibration

In many works the numerical simulators are used to analyze the power device behavior at different test conditions [66], [78], [72], and it is highlighted the

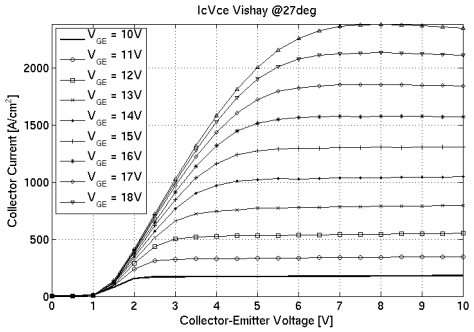


Figure 3.9: The experimental output characteristics of the device at T = 25 °C.

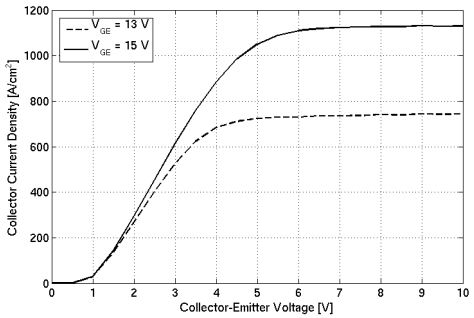
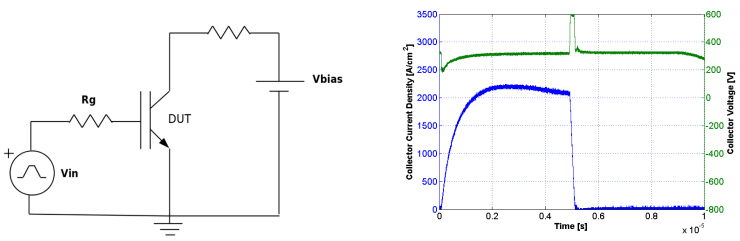


Figure 3.10: The experimental output characteristics of the device at T= 125 °C.

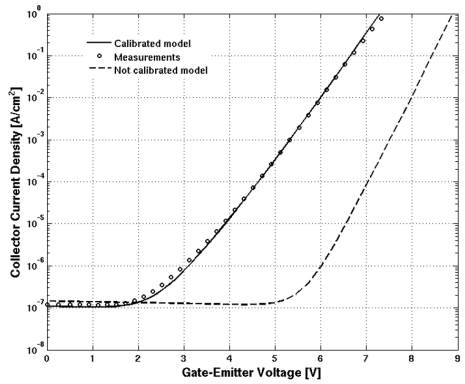


(a) (b)

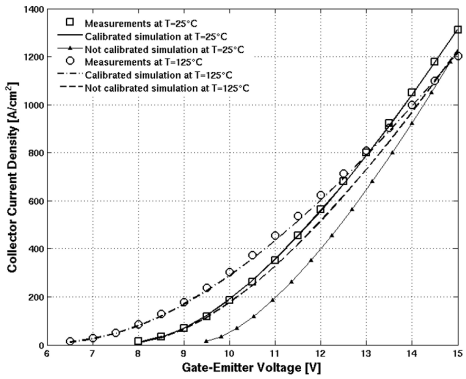
Figure 3.11: (a) Short-circuit test schematic and (b) the experimental curves in short-circuit conditions performed on a $6\mu m$ pitch sample.

importance of an accurate calibration of the parameter models. However, to our knowledge, a detailed description of the procedure adopted to calibrate the structure of silicon power devices in a TCAD environment is reported by only few works [79], [80]. The calibration procedure presented in [79] achieves a good agreement between the simulated and measured terminal characteristics, but the Internal Laser Deflection technique adopted for the measurement of carriers concentration and the temperature profile requires a sophisticated characterization setup and, moreover, a simulation of the device subjected to stressful condition, like the short-circuit one, is not presented. In [80] the calibration of the models, implemented in semiconductor numerical simulator, on a real device is based on the doping profile variation to fit the experimental curves. A set of parameters chosen to calibrate the models is not clarified. The strength of the calibration procedure presented in this work is that the only information given by the isothermal characteristics, supported the calibration algorithm presented in the previous section, are sufficient to calibrate accurately the physical models involved in the electro-thermal simulations. It has to be pointed out that even if a good matching between the experimentally evaluated doping profiles and the ones defined in the elementary cell implemented in TCAD is achieved, the accuracy of the reconstructed doping profiles in the critical parts of the structure can differ from those of the real device. However, if the reconstructed doping profiles are close the real ones, the variations adduced on the physical parameters from the calibration procedure compensate also the small mismatches among the doping profiles. The reliability of the proposed procedure in (§2.3) has been applied to the presented structure, in which the electron irradiation lifetime killing technique is applied.

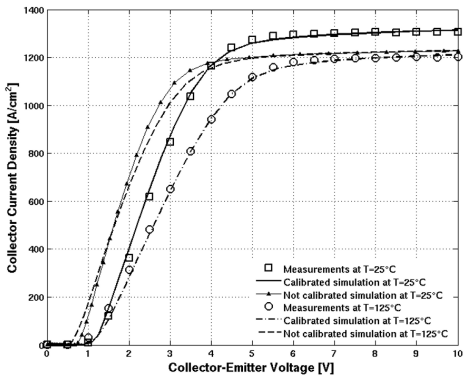
In Fig. 3.12 the curves relative to the calibrated and the not calibrated structures are reported together with the experimental ones and the relative fitted parameters are reported in Table 3.1. In the first step (Fig. 3.12a), the linear region of the sub-threshold curve has been calibrated by fitting the simulated curve on the experimental one, in the Collector current density range $J \in (10^{-5}, 1) A/cm^2$. The leakage current at $V_{GE} = 0$ V has been calibrated in the last step of the calibration procedure and it has been achieved with $E_{trap} = (E_t - E_i) = 0,14eV$, where E_{trap} is the distance of the trap level from the intrinsic Fermi level. The not calibrated curve exhibits two incongruities: the slope of the characteristic does not match with the experimental one and the threshold voltage is 1.5 V higher. An overall analysis of the simulated curves reported in Fig. 3.12b and 3.12c is needed since they are the result of an iterative fitting procedure which involves, in different steps, ten pa-



(a)



(b)



(c)

Figure 3.12: Measured and simulated terminal curves after the calibration procedure for the Trench PT-IGBT with a breakdown voltage of 600 V. The curves without calibration are reported as well. (a) Subthreshold region (T= 25 °C and $V_{CE} = 10V$). (b) $I_C - V_{GE}$ characteristics ($V_{CE} = 10V$). (c) IC-VCE characteristics ($V_{GE} = 15V$).

Table 3.1: Calibrated Parameters for the PT-IGBT structure

Symbol	Default Value	Calibrated Parameter
$N_{ref}(e)$	$1.2500E+17cm^{-3}$	$8.0603E+17cm^{-3}$
$\beta(e)$	1.109	1.585
$\alpha_N(e)$	2.5	2.4
$\beta_{EXP}(e)$	0.66	0.66
τ_{p0}	$3.0000E-6s$	$3.7864E-8s$
τ_{n0}	$1.0000E-5s$	$8.2922E-8s$
$N_0(e)$	$1.0000E+16cm^{-3}$	$5.0000E+18cm^{-3}$
$N_0(h)$	$1.0000E+16cm^{-3}$	$5.0000E+18cm^{-3}$
$C(e)$	2.55	0.01
$C(h)$	2.55	1.11
E_{trap}	$0.00eV$	$0.14eV$

rameters ($N_{ref}(e)$,$\beta(e)$,$\alpha_N(e)$,$\beta_{EXP}(e)$,τ_{p0},τ_{n0},$N_0(e)$,$N_0(h)$,C(e),C(h)). A good fitting accuracy has been achieved for both the families of curves at T = 25 °C as well as T= 125 °C, proving the reliability of the set of models and parameters adopted in the calibration procedure. The $I_C - V_{CE}$ characteristics have been calibrated only for $V_{GE} = 15V$, but it has to be pointed out that a good calibration of the $I_C - V_{GE}$ characteristics implies a good fitting of the saturation current for the $I_C - V_{CE}$ curves. In Table 3.1 the parameters involved into the calibration of the curves in Fig. 3.12b and 3.12c are reported and are compared to the default values of the different models. The first four parameters are relative to the $I_C - V_{GE}$ characteristics and the most relevant variation, compared to the default values, is observed for $N_{ref}(e)$ and $\beta(e)$. As mentioned before, the mobility into the depletion region of the channel is not only dependent on the doping concentration at the interface, but it depends also on the features of the silicon/oxide interface and the doping segregation in the oxide which modifies the doping profile at the interface. Therefore, the couple of parameters $N_{ref}(e)$ and $\beta(e)$ implicitly keeps into account for all of the above effects. The electron irradiation technique is widely used to reduce the lifetime in power devices to achieve higher performance in turn-off conditions. However, the more the lifetime is reduced, the more the V_{ON} increases and the saturation current decreases because of the detrimental effect on the current gain of the vertical PNP. In on-state the Drift layer works in high-injection conditions and the voltage drop is mainly due to the ambipolar lifetime (τ_a), which is the sum of τ_{p0} and τ_{n0}. The default value for τ_a is

$1.3E-5s$, which is a typical value for epitaxial layers, is drastically reduced to $\tau_a = 1.2078E-7s$ because of the effect of the defects introduced into the structure with the electron irradiation. The value of the ambipolar lifetime in the analyzed structure is also confirmed by the experimental current waveform in short-circuit conditions (See Fig. 3.11), which does not exhibit a significant tail at the turn-off, sign of a reduced current gain of the vertical PNP and a very low lifetime [14]. The two doping references for the Scharfetter relation are two orders of magnitude higher than the default one basically because of the uniformity of the defects introduced into the entire structure, which has an effect dominant compared to the effect due to the doping concentration. From the point of view of the temperature-scaling parameters, a suitable fitting of the I_C-V_{GE} curves is achieves by keeping the default values. Lifetime techniques improve the device speed, however they complicate the prediction of the device behavior at high temperatures [81]. This phenomenon is evident for the temperature scaling parameters of τ_{p0} and τ_{n0}, $C(h)$ and $C(e)$ respectively. Both the minority carriers exhibit a lower dependence with the temperature compared to the default ones. In Fig. 3.12b and 3.12c the curves of the not calibrated structure are reported as well. In Fig. 3.12b two main differences are evident between the calibrated and the not calibrated curves: the threshold voltage (as shown in Fig. 3.12a) and the different Temperature Compensation Point (TCP). While the calibrated structure shows a TCP at $V_{GE} = 13V$, the not calibrated structure shows a TCP at $V_{GE} = 15V$, as it is confirmed in Fig. 3.12b. The higher value of the TCP for the not calibrated structure gives an higher instability when the device works in stressful conditions like the short circuit. The effect of the ambipolar lifetime in the Drift layer is highlighted in Fig. 3.12c where the VON difference between the calibrated and the not calibrated curves is evident (the voltage difference between the two curves is in the order of 0.6 V). When an accurate calibration of the physical models is achieved (implicitly, a calibration of the technological environment, used to make the device, is achieved), quantitative evaluations of carriers and temperature distribution in the structure can be obtained by performing electro-thermal simulations in stressful conditions, like the short-circuit test, as well as in new structures in which the technological environment is kept constant. To evaluate the reliability of the calibration procedure, the calibrated structure has been simulated in turn-off and short-circuit condition. Results have been compared to the experimentally measured ones. The schematic of the experimental setup used to evaluate the turn-off energy of the devices under test is reported in Fig. 1.6. The circuit riported in Fig. 1.6 does not accounts for all of the parasitic

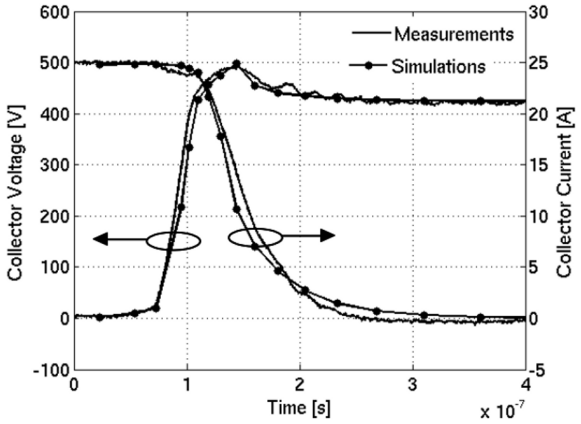


Figure 3.13: Simulated and measured curves relative to the Turn-off performed on the Trench PT-IGBT.

elements. An accurate evaluation of all of the parasitic elements have been performed and the complete circuit has been imported in Sentaurus TCAD. The behavior of the D.U.T. in turn-off conditions has been evaluated by using the Mixed-Mode modality. Since the turn-off waveforms are particularly sensitive to all of the parameters of the freewheeling diode, C_{j0} and tt (spice parameters of the diode) have been calibrated in order to match the curves at the turn-off for the calibrated structure. The above parameters have been chosen since they are the most relevant in the definition of the waveforms at the turn-off. As shown in Fig. 3.13, a good matching between experimental curves and simulations has been achieved for the reference structure.

The short-circuit condition is a very critical condition for the device [82] and a correct simulation of the elementary cell can be achieved by:

1. reporting the realistic boundary conditions of the elementary cell [66](including the oxide layers and the metallic contacts with their real thickness)

2. adopting a more accurate model of the thermal parameters of the silicon.

In Fig. 3.14 the top of the elementary cell is reported and it is visible the presence of both the oxide on the top of the trench and the contact trench with a thickness of $3\mu m$ above the zero of the y-axis (the end of the silicon part

of the device). In Fig. 3.15 experimental measurements on the silicon heat capacity is reported [83] together with the approximated polynomial function. As it is visible, the heat capacity increase when the temperature increases; if the thermal power dissipated into a region of the structure is kept constant, the more the temperature increases into the silicon, the more the increase of temperature is reduced. The polynomial function, which approximates the heat capacity as function of the temperature, is expressed like:

$$C_L = \alpha + \beta T + \lambda T^2 + \chi T^3 \tag{3.1}$$

By default Sentaurus TCAD models the heat capacity constant with the temperature with $\alpha = 1.63 J/(K cm^3)$ and the rest of the parameters are fixed to zero. A suitable fitting of the experimental data (up to 1000K) has been achieved by setting: $\alpha = 0.6029 J/(K cm^3)$, $\beta = 5.117e-03 J/(K^2 cm^3)$, $\lambda = -6.096e-06 J/(K^3 cm^3)$ and $\chi = 2.572e-09 J/(K^4 cm^3)$.

In Fig. 3.16 the effect of the variation of the thermal boundary conditions on the short-circuit simulation is reported . By default, in Sentaurus TCAD all of the surfaces of the structure are defined adiabatic. When electro-thermal simulations are performed on power devices a not negligible amount of power is dissipated and self-heating effects can lead to a strong variation of the performance of the device. In structure analyzed in this first part of the report a contact trench solution is adopted and the removal of the metal layer from the structure introduces a great error into the results of the simulations. Basically, if the metal layer is removed and the contact (the Emitter contact in this case) is only electrically defined, the contact is adiabatic from the thermal point of view, confining the energy only in the silicon (there is an infinite thermal resistance between the silicon boundaries and the external) and leading to faster increase of the mean temperature into the structure. When the metal layer, together with oxide on the top of the cell, are reported, a more quantitative information can be achieved since a more detailed modeling of the thermal impedance of the structure is achieved. This effect is well visible in Fig. 3.16. The main effect of the metal layer on the waveform is an higher capability of the device in short-circuit conditions, basically because of the lower energy stored into the silicon (The current rise in the short circuit simulations occurs because of the thermal latch-up triggered by the leakage current which increases with the temperature). A not negligible effect of the heat capacity modeling is also visible in Fig.3.16. The fourth curve in the legend is relative to a structure in which the metal layer thickness is fixed at $3\mu m$ while the heat capacity dependence with the temperature has been evaluated by (3.1).

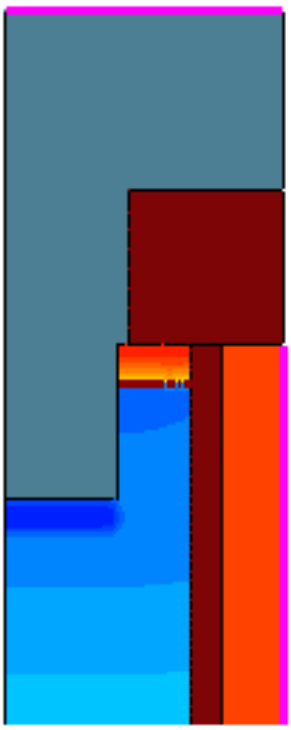


Figure 3.14: Top of the elementary cell with the addition of the metal and oxide regions upon the silicon level.

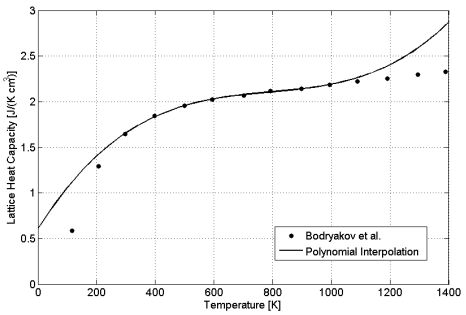


Figure 3.15: The experimental behavior of the heat capacity of silicon is reported together with the approximating polynomial function.

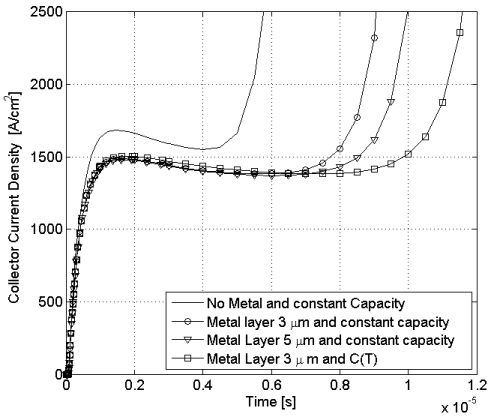
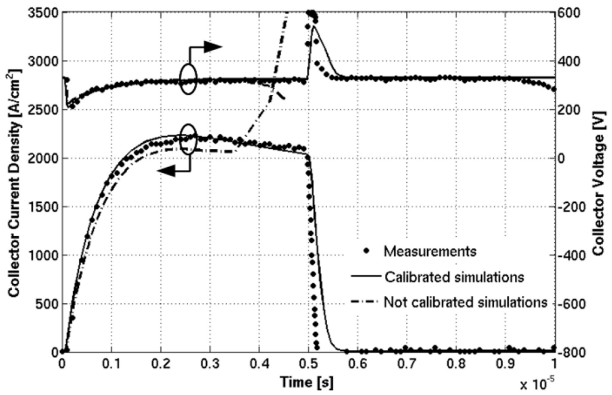
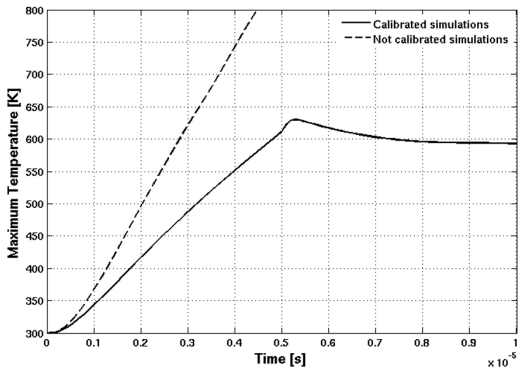


Figure 3.16: Effect of the thermal boundary conditions on the short-circuit simulation.

As mentioned before, the increase in temperature increases the heat capacity, reducing the temperature increase in the device. This effect increases again the short-circuit capability of the structure up to a value coherent with the experimental data (The typical value of Ton at which the device fails is in the order of $7.5\mu s$, as shown in the previous section). The Short-Circuit test simulations have been performed by including the temperature dependence of the silicon heat capacity and in Fig. 3.17a the experimental curve together with the calibrated simulations and the not calibrated simulations are reported. From the point of view of the calibrated curves a good matching between the experimental data and the simulations is achieved in all of the parts of the waveforms (the parasitic elements of the experimental setup have been evaluated in detail in order to have in simulations comparable time constants of the experimental setup). In the same figure the short-circuit simulation performed on the not calibrated structure is reported and a very different behavior is observed with the same thermal boundary conditions are the same. Even if the saturation current is lower for the not calibrated structure, after $3\mu s$ the collector current diverges because of thermally induced latch-up. The latter condition is enhanced mainly because of the higher voltage of the TCP (at $V_{GE} = 15V$) which causes a thermal instability, then a faster increase of the temperature in the structure. The faster increase of the temperature in the structure (reported in detail in Fig. 3.17b) is also due to the higher current gain of both the vertical



(a)



(b)

Figure 3.17: A comparison between the experimental and the simulated (calibrated and not calibrated) curves relative to the short circuit are reported in (a). The maximum temperature in the calibrated and the not calibrated structure are reported in (b).

PNP and the parasitic NPN structure (N+ Emitter - P-Body Drift layer), which reduces the latch-up capability of the overall structure.

3.2 Short-circuit capability analysis

In this section the effect of the geometrical and doping features of the structure will be carried out. As mentioned before, a trade-off among the parameters rises when the short-circuit capability is changed. Therefore, it is necessary to evaluate the overall performances of the device which will be investigated by considering the trade-off among the following parameters:

1. Turn-off energy (E_{off})
2. V_{ON}
3. Short-circuit capability

The Short-Circuit capability is defined starting from the failure mode in Fig. 3.18. When t_{on} is lower than the critical value at which the devices fails, the Collector Current Density shows a negative slope because of the mobility reduction into the channel. When the critical t_{on} is achieved, the Collecotr current density waveform exhibits a change of the slope (becomes positive) and, because of the thermal latch-up which is triggered by the high temperature into the structure, the device can not switch off. If the failure occurs, a switch in series to the V_{bias} generator (this switch is used to interrupt the current flow into the D.U.T.) acts like a Device Under Test (D.U.T.) protection and the slope-change of the collector current during the switch-off was chosen as the test interruption criterion. The slope change represents the switching on of the parasitic thyristor which causes the failure of the device if the series switch does not work. As it can be seen in the Fig. 3.18, at $t = 21\mu s$ the current exhibits a slope change and the D.U.T. fails at $t = 30\mu s$. In order to prevent the physical destruction of the D.U.T., the protection must be activated before the failure. Therefore, the Short-Circuit capability has been defined like the maximum t_{on} achievable by the D.U.T. before the Collector current slope assumes a positive value. In the following sections, the effect of the geometrical features and the doping profiles on the Short-Circuit capability will be analyzed by means of the TCAD simulations.

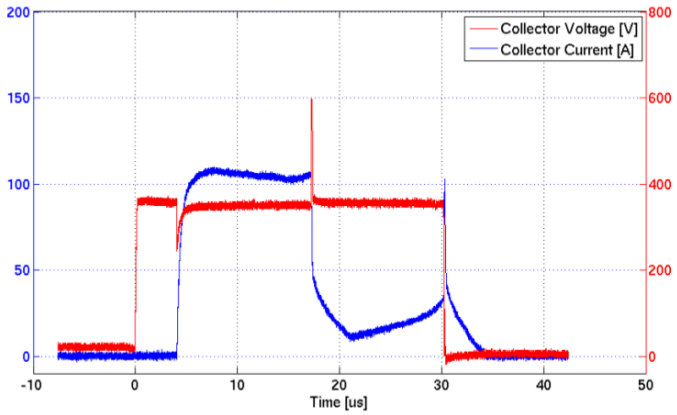


Figure 3.18: An example of Short-Circuit failure for the analyzed class of devices.

3.2.1 Cell pitch effect

In this section the effect of the elementary cell pitch on the performance of the PT-IGBT has been analyzed and three values have been chosen: 4, 5 and 6 μm. The comparison among the three design has been done by keeping constant the overall area of the die. In Fig. 3.19 simulated structures are reported.

Even if the 3D structures can be obtained by the extrusion of the 2D structure into the z-direction ,the simulations have been performed with the 3D structures in order to verify the reliability of the 2D simulations. From the comparison of the 2D and the 3D simulations (E_{off}, V_{ON} and short-circuit capability) it can be concluded that the differences are negligible, from the point of view of the pitch effect. In Table 3.2 the high injection values of lifetime for the holes and electrons are reported (τ_{p0} and τ_{n0}).

In Fig. 3.20 the $V_{ON} - E_{off}$ curves for the three designs is reported. The behavior of the three designs has been simulated at high temperature as well (T=125 °C). As expected, the V_{ON} value decreases when the lifetime increases. E_{off} has the opposite behavior since it increases when the lifetime increases (the more the lifetime increases, the more the stored charge into the drift-layer takes longer time to be swept out because of the lower recombination rate). Moreover, if the technology is fixed (pitch and irradiation dose) and the temperature is increased, because of the increase of the lifetime into the structure (the calibration procedure highlights a positive temperature co-

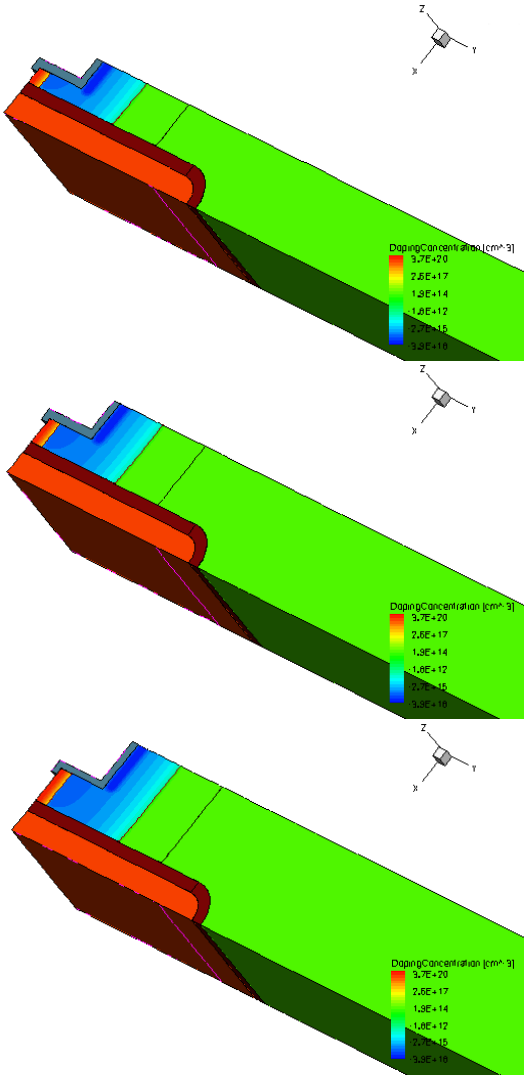


Figure 3.19: The simulated structures have a cell pitch of 4, 5 and 6 μm, in succession.

Table 3.2: Carriers high injection lifetime values

τ_{p0}	τ_{n0}	Notes
1.3698E-8	4.7945E-8	$LowE_{off}$
1.7925E-8	6.2739E-8	
2.2569E-8	7.8991E-8	$Calibrated$
2.3456E-8	8.2098E-8	
3.0694E-8	1.0743E-7	
4.0165E-8	1.4058E-7	
5.2559E-8	1.8395E-7	
6.8777E-8	2.4072E-7	$HighE_{off}$
9.0000E-8	3.1500E-7	

efficient for the lifetime of the minority carriers), the V_{ON} is reduced and the turn-off energy increases. From the comparison of the three designs, the following conclusions can be addressed:

1. If the E_{off} level is fixed, the V_{ON} decreases when the cell pitch is reduced.

2. If the V_{ON} is fixed, the E_{off} increases if the cell pitch increases.

3. The above behavior shown in the previous two points are still valid at higher temperature.

4. If the lifetime into the structure is fixed, the E_{off} does not change (See the points into the circles in Fig. 3.20) This is due to the fact that if the stored charge into the drift-layer does not change (the turn-off test has been performed by switching-off a current density of $260A/cm^2$ as shown in Fig. 3.20a).

The overall evaluation of the performances of the different designs have been done by considering the results of the simulations in short-circuit condition at different lifetimes: the calibrated lifetime together with the $HighE_{off}$ and $LowE_{off}$ values defined in Table 3.2 have been adopted to evaluate the affect of the lifetime on the short-circuit capability.

In Fig. 3.21 the D.U.T. is simulated in short-circuit conditions and the lifetime is set at the values $LowE_{off}$, $CALIBRATION$ and $HighE_{off}$, in ascending order. The following conclusions can be addressed:

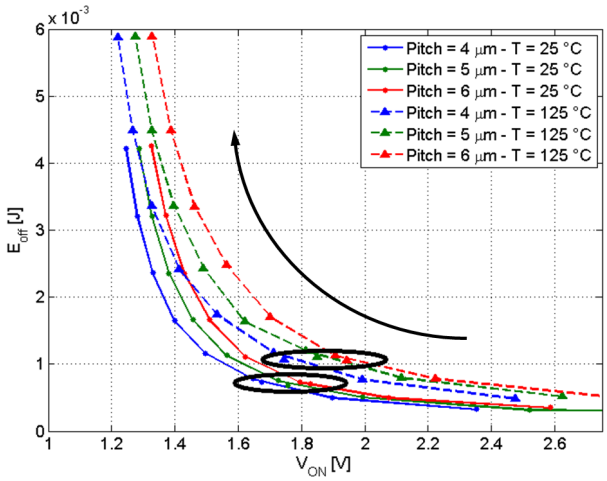


Figure 3.20: Simulated $V_{ON} - E_{off}$ curves of the three designs at T = 27 °C and T= 125 °C. Lifetime increases along the direction of the arrow and the values in the circles are relative to the calibrated lifetime value.

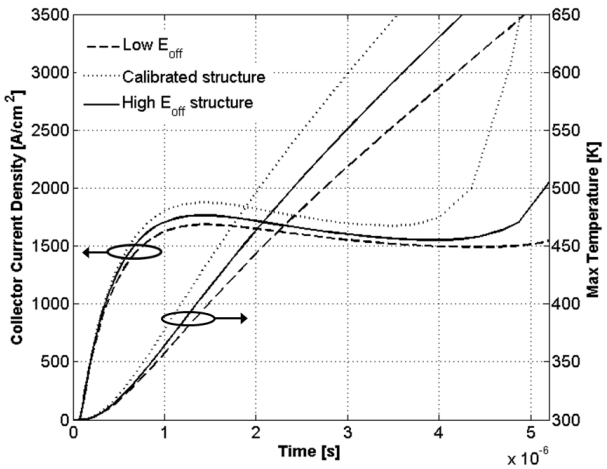


Figure 3.21: The cell pitch is kept constant at 6 μm and the lifetime is varied.

1. The short circuit capability increases when the lifetime in the structure decreases (The legend of the figure refers to the turn-off energy which increases when the lifetime increases) - This phenomenon is due to the effect of the lifetime on the gain of the vertical PNP. More in detail, the more the lifetime is increased, the more the gain increases. As reported in [14], the beta of the vertical PNP (in the Beta value has been evaluated for the NPN structure and (3.2) has been obtained by using the complementary notation for the subscripts) can be descripted by the relation:

$$\beta = \frac{D_{pB} L_{nE} N_{AE}}{D_{nE} W_B N_{DB}} \left(\frac{n_{iB}}{n_{iE}} \right)^2 \tag{3.2}$$

   where D_{pB} is the diffusion constant of holes in the Base, L_{nE} is the diffusion length of the electrons in the Emitter,N_{AE} is doping concentration of the Emitter,D_{nE} is the diffusion constant of the electrons in the Emitter, W_B is the Base width, N_{DB} is the doping concentration of the Base, n_{iB} is the intrinsic concentration in the Base and n_{iE} is the intrinsic concentration of the Emitter. The Emitter and Base region are referred to the vertical PNP structure. This simplified equation for the Beta (3.2) gives easily the relation between the Beta and the lifetime. Since D_{pB} are D_{nE} are dependent only on N_{AE} and N_{DB} , L_{nE} is the only parameter dependent on the ambipolar lifetime of the structure. More in detail it is proportional to the root square of the ambipolar lifetime into the Emitter region of the vertical PNP. The short-circuit capability is reduced when the lifetime increases because of the increase of the PNPN structure efficiency which enhance the thermal latch-up phenomenon.

2. It has to be noticed that even if the collector current variation is reduced compared to that introduced from the cell pitch variation (See Fig. 3.22), the position of the collector current plateau moves faster to the right (higher short-circuit capability) compared to the increase due to the cell pitch variation. Moreover the slope of the collector current after the plateau region increases when the lifetime increases. This phenomenon is due to the different Beta of the structure which increases when the lifetime increases and, at the same time, increases the thermal run-away of the PNPN structure (the higher slope of the collector current during the latch-up for devices with higher lifetime is an additive proof of the

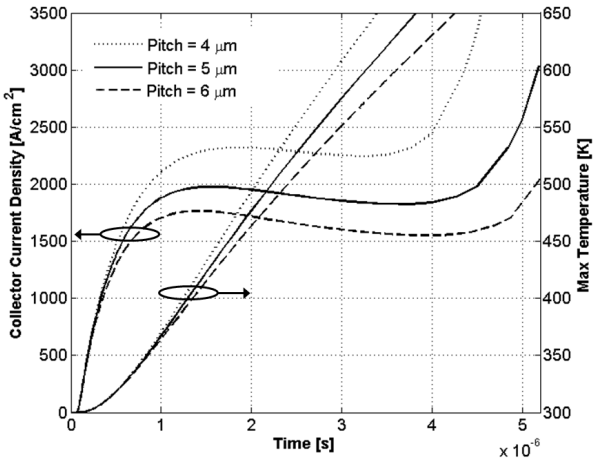


Figure 3.22: The lifetime of the structure is kept constant and the cell pitch is varied (lifetime is the $CALIBRATED$ value).

illustrated phenomenon).

In Fig. 3.22 the lifetime of the structure has been fixed to the calibrated one and the cell pitch has been varied to evaluate the effect on the short-circuit capability. From this graph the following conclusions can be addressed:

1. The short-circuit capability decreases if the cell pitch decreases - Since the active area of the die is kept constant, when the cell pitch is reduced the channel depth of the MOS part of the IGBT is proportionally increased and an higher total current (the saturation current) flows into the Die. In this way the mean value of the power dissipated into the structure increases and the temperature increases faster. This phenomenon is evident in Fig. 3.22 where, in dashed lines, the maximum temperature into the structure is reported.

2. The point in which the slope of the collector current changes its sign can be addressed like the maximum Ton which the D.U.T. can sustain before the thermal latch-up cancels the current control capability of the MOS part of the IGBT. The more the cell pitch is increased, the more the point moves to the right (the device has an higher short-circuit capability).

3. The slope of the curve after the plateau does not change with the cell pitch - This is due to the vertical PNP structure which is not changed by

changing the cell pitch and, since the current gain of the PNP depends on both the doping profiles and the lifetime in the structure (From (3.2)), does not change in a relevant way. The increase of the current after the plateau is enhanced by the increase of the temperature into the structure that, in general, increases the current gain of the PNPN structure of the IGBT.

The experimental data highlighted a technological drawback for the deposition of the emitter metal layer into the contact. More in detail, the metal layer, which is used to make the Emitter contact, is not homogeneous and exhibits vacancies upon the contact trench region (As shown in Fig. 3.1). The dies with a cell pitch of $5\mu m$ have shown a better coverage of the metal layer. The latter considerations, together with the simulation results shown in this section have been the support for the definition of the $5\mu m$ design like the standard one.

3.2.2 Contact trench effect

The experimental characterization of the devices in avalanche conditions (this subject will be explained in §4.3.3) have shown how the weakness of the contact between the Emitter metal layer and the silicon, especially into the Contact-Trench region, drastically reduces the ruggedness of the active area in avalanche conditions. In this section the effect of the contact trench depth on the short circuit and latch-up capabilities are investigated. With latch-up it is defined the condition in which both the vertical PNP and the parasitic NPN of the IGBT are activated, not allowing the turn-off of the current. This condition is detrimental for the device since it can lead to the destruction of the device because of the thermal run-away.

The most important parameter which affect the latch-up triggering is the resistivity of the P-Body layer below the N+ diffusion of the Emitter (See §1.5)(a lower value of doping at the junction of the P-Body region means an higher resistivity, then a lower latch-up current). Moreover it is also relevant the part of the total current which flows below the Emitter diffusion. The latch-up current of the structure with a reduced depth of the contact has been evaluated.

In Fig. 3.24 the structure without contact trench and the one with the contact trench are reported. As mentioned before the latch-up phenomenon is dependent on the spreading resistance of the P-Body region below the Emitter N+ region (the effect of the spreading resistance for the vertical part of the N+/P-Body junction can be neglected since usually, in order to have a good ohmic contact, the P+ region in Fig. 3.23 has a doping two order of magni-

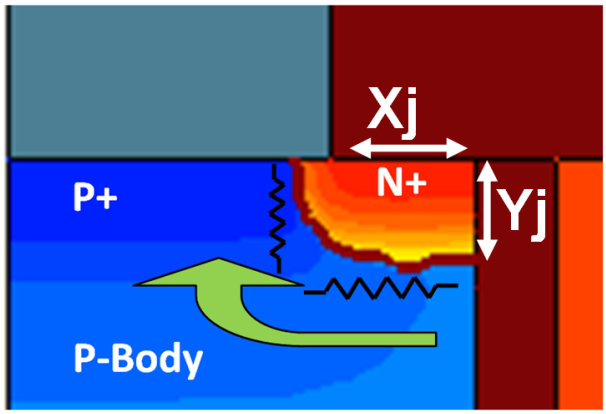
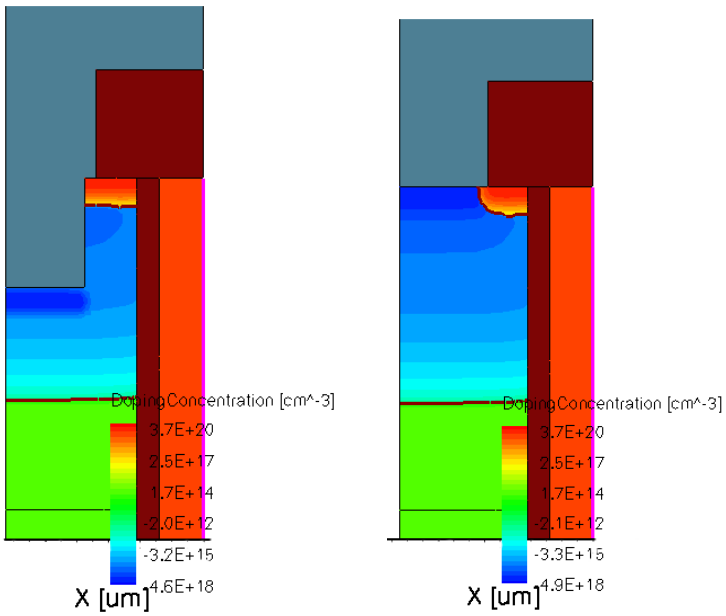


Figure 3.23: Region of the IGBT responsible for initiation of latching. The arrow indicates the holes flow which brushes the P-Body/Emitter junction.



(a) (b)

Figure 3.24: Cross section of the structure with the contact trench(a) and without contact trench (b).

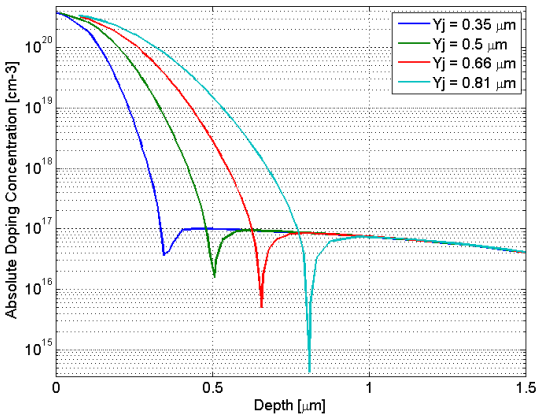


Figure 3.25: Doping profiles used to simulate the effect of the Emitter depth (Y_j in FIG. 3.23) on the Latch-up capability of the structure.

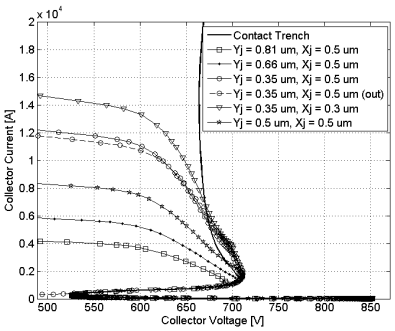


Figure 3.26: I_C-V_{CE} curves in avalanche conditions (solid-dotted lines) and output curve relative to the design with $X_j = 0.5\mu m$ and $Y_j = 0.5\mu m$.

tude higher than the region below the N+ diffusion) and an investigation of the effect of the geometrical features of the N+ region on the latch-up current has been performed by using the TCAD simulator. In Fig. 3.23 the two analyzed geometrical parameters of the N+ region are reported. X_j is the width of the window of the mask for the N+ region (basically the lateral diffusion component is not considered) and Y_j is the depth of the junction at the interface silicon/oxide. In Table 3.3 the simulated values are reported: for $X_j = 0.5\mu m$ (which is the same width of the design with the contact trench) the values $Y_j = 0.35\mu m$, $0.5\mu m$, $0.66m$ and $0.81\mu m$ have been analyzed.

Table 3.3: Geometrical features of the analyzed structures

$X_j(\mu m)$	$Y_j(\mu m)$
0.5	0.35, 0.5, 0.66, 0.81
0.3	0.35

For $X_j = 0.35\mu m$, the value $Y_j = 0.35\mu m$ has been tested. In Fig. 3.25 the doping profile along the silicon/oxide interface is reported for the different values of Y_j. In Fig. 3.26 the I_C-V_{CE} curves relative to the above values of X_j and Y_j are reported. The horizontal segments mark the current level at which the Latch-up starts. The simulations show that:

1. The latch-up current increases when Y_j decreases - The peak of the P-Body profile is close to the surface, therefore the doping concentration decreases when moving from the top to the bottom of the P-Body region (the spreading resistance has the opposite trend). Since the latch-up current decreases when the spreading resistance increases, the Latch-up current decreases when Y_j increases.

2. If Y_j is fixed, the Latch-up current increases if X_j decreases - When Y_j is fixed the the total resistance which the holes have to cross increases if the X_j increases and the Latch-up current decreases. This is shown in Fig. 3.26 by the black segment level ($X_j = 0.5\mu m$ and $Y_j = 0.35\mu m$) and the level of the red segment with the squares at the boundaries ($X_j = 0.3\mu m$ and $Y_j = 0.35\mu m$).

3. The output curve ($V_{GE} = 15V$) of the structure with $X_j = 0.5\mu m$ and $Y_j = 0.5\mu m$ is reported as well to highlight that the Latch-up current does not depend on the working condition of the device.

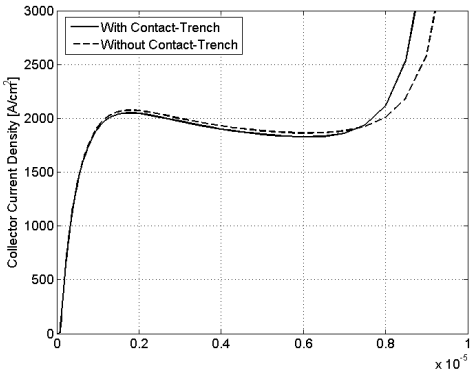


Figure 3.27: Short-circuit waveforms relative to the two structure in Fig. 3.24. The device without the Contact-Trench has been simulated with $X_j = 0.5\mu m$ and $Y_j = 0.35\mu m$.

4. In the worst case the Latch-up current is in the order of 1000A which is a value two orders of magnitude higher than the normal operative conditions of the devices under test.

5. The Structure with a deep contact trench (the one of the calibrate structure) does not exhibit a latch-up current up to over 20000A.

From the above considerations comes out that the latch-up capability is reduced when the contact trench is removed, however this effect does not reduce the ruggedness of the devices in normal operative conditions. The two structures, with and without Contact-Trench, have been compared in Short-Circuit to evaluate the trade-off due to the Contact-Trench. In Fig. 3.27 the waveforms relative to the two designs is reported. As it is visible, the change of the top design of the device has a negligible effect on the short-circuit capability. By means of simulations, the V_{ON} and the E_{off} values does not change by acting on the geometrical features of the Contact-Trench.

3.2.3 Channel Length effect

As shown before, the maximum sustainable t_{on} of the device with an higher lifetime is lower and an improvement of the short-circuit capability can be achieved by reducing the saturation current of the device (to reduce the overall power dissipated during the short-circuit). Since the saturation current is

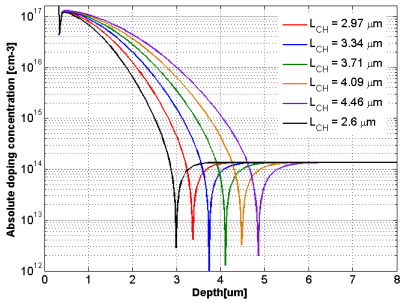
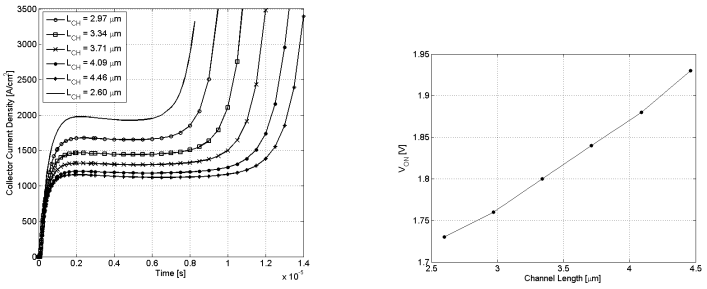


Figure 3.28: Doping profiles of the Body region at the Silicon/Oxide interface.

inversely proportional to the channel length (See §1.3), the $5\mu m$ devices have been characterized in simulation by evaluating V_{ON}, E_{off} and Short-Circuit capability by varying the channel length for two different values of lifetime (the $CALIBRATION$ and $HIGHEoff$ values of Table 3.2). From the point of view of the E_{off}, as shown into the previous section, if the cell pitch is fixed together with the lifetime the variation of the E_{off} with the channel length is negligible, since the stored charge into the Drift-layer is fixed (The turn-off current is fixed). The Body doping profile has been obtained from the superposition of two Gaussian profiles (See §3.1.1). In order to evaluate in a simplified way the qualitative effect on the saturation current of the D.U.T. for different values of Channel Length, a single Gaussian profile has been adopted and the channel length has been varied until the saturation current have matched the saturation current of the calibrated structure, keeping the same doping profile peak (to have the same threshold voltage). In Fig. 3.28 the investigated doping profiles of the Body at the Silicon/Oxide interface are reported.

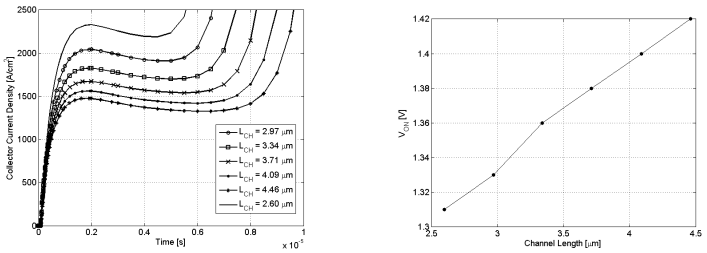
The simulation results are reported in Fig. 3.29 and Fig. 3.30. The short-circuit test has been performed for six values of the channel length (See Fig. 3.28) and two levels of lifetime ($HIGHEoff$ and $CALIBRATED$). From Fig. 3.29 and 3.30 it can be addressed that:

1. V_{ON} increases if the lifetime decreases
2. The V_{ON} proportionally increases with the channel length (Fig. 3.29b and 3.30b) and the slope of the Channel length versus V_{ON} curve increases when the lifetime decreases.



(a) (b)

Figure 3.29: (a) Short-circuit waveforms relative to different channel length values ($CALIBRATED$ structure).(b) The channel length versus the V_{ON} is reported.



(a) (b)

Figure 3.30: (a) Short-circuit waveforms relative to different channel length values ($HIGHEoff$ lifetime value).(b) The channel length versus the V_{ON} is reported.

3. The short circuit capability is enhanced when the channel length is increased. Basically this effect is due to the lower power dissipated into the structure because of the lower saturation current of the IGBT.

3.2.4 Emitter diffusion modulation

As mentioned in the previous sections, the short-circuit capability of the analyzed devices is mainly regulated by the saturation current and the lifetime of the device. One of the possible ways to reduce the saturation current is the modulation of the Emitter diffusion region (N+), [84], [85]. In this kind of solution, the Self Aligned Structure is modified in a way that the Emitter diffusion is modulated along the direction orthogonal to the section in Fig. 3.5 and the Emitter Modulation percentage is here defined like the ratio between the width of the N+ diffusion along the z-direction at the Silicon/Oxide interface (the only part which injects electron directly into the channel) of the elementary cell and the depth along the z-direction of the elementary cell (See Fig. 3.31(a)). Therefore the structure without Emitter-Modulation has an Emitter Modulation percentage of 100%.

Table 3.4: V_{ON} and J_{sat} of the analyzed structures.

	$V_{ON}[V](atJ_C = 260A/cm^2)$	$J_{SAT}[A/cm^2](atV_{CE} = 10V)$
SALT	1.716	1685
SAHT	1.181	1996
25LT	1.711	1874
25HT	1.181	2366
20LT	2.274	386
20HT	1.449	536
50LT	1.830	913
50HT	1.247	1169

In Fig. 3.31 and Fig. 3.32 the structures adopted to investigate the emitter modulation technique are reported together with the detail of the section of the top of the cell. More in detail in Fig. 3.31 the modulation of the Emitter diffusion (N+) is obtained by defining it like a stripe orthogonal to the trench-gate. In this design the saturation current is limited by the reduced length in the z-direction of the N+ diffusion which modulates the electron current into the MOSFET channel, with the corresponding modulation of the overall current of the device (See §1.3). On the other hand, in Fig. 3.32 it is reported

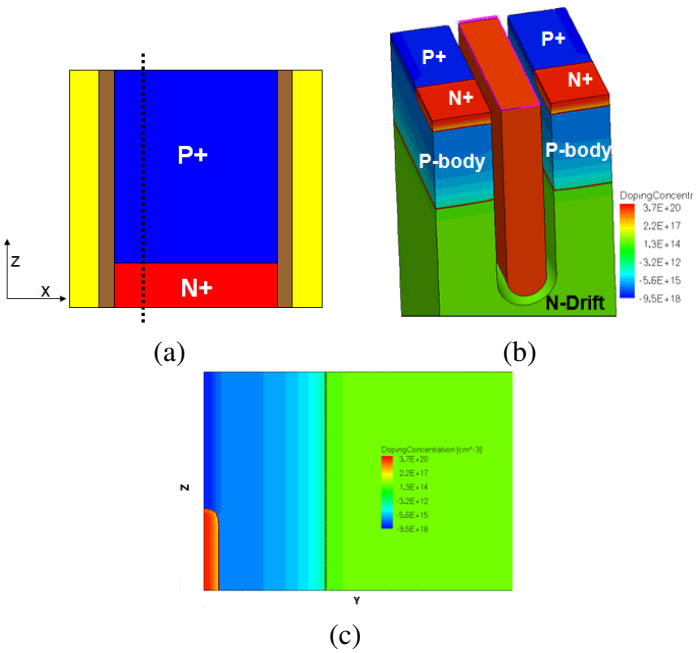


Figure 3.31: The geometrical details of the *Orthogonal Emitter 20%* design. (a) Top view of the cell, (b) 3D structure and (c) section along the cut in (a)

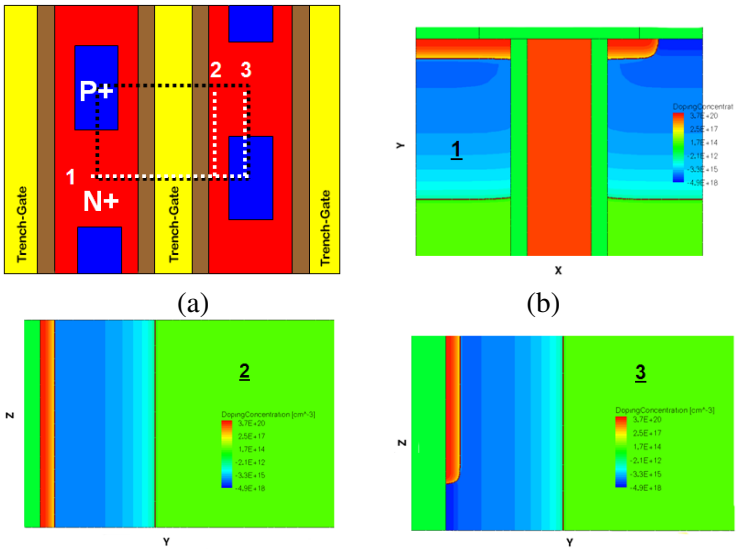


Figure 3.32: The geometrical details of the *P-Body contact 25%* design. (a) Top view of the cell, (b) section along the cut number 1 in (a), (c) section along the cut number 2 in (a) and section along cut number 3 in (a). In (a) the simulated 3D elementary cell area is highlighted by the dashed black rectangle.

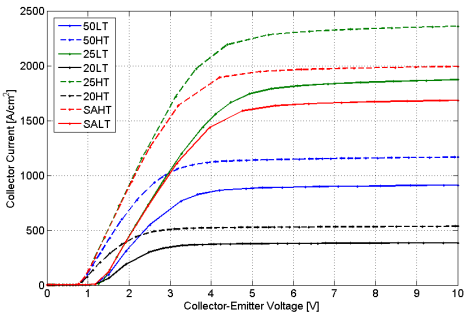


Figure 3.33: Output curves at ambient temperature for the investigated designs.

the structure in which the Emitter diffusion id modulated in a way that the P-Body contact (P+) is extended in a reduced area of the top of the device structure. For all of the designs the contact trench has been removed and the P-Body layer has a secondary doping implantation at the surface with the same depth of the N+ diffusion in order to get at the same time a ohmic contact and better latch-up capability (the vertical distributed resistance in Fig. 3.23 is reduced by increasing the doping concentration). As preliminary step, the above elementary cells have simulated in static conditions to get the output curves to evaluate the V_{ON} (See Fig.3.33 and Table 3.4). The next step has been the simulations of the four design in short-circuit conditions at two level of lifetime into the structure (See Fig.3.34). The following notation has been adopted in Fig. 3.33 and Fig. 3.34:

- **SALT** - *Self Aligned structure* - structure with (See Fig. 3.5) with a Low level of lifetime into the structure.
- **SAHT** - *Self Aligned structure* - with an High level of lifetime into the structure.
- **25LT** - *P-Body Contact 25%* - structure with a Low level of lifetime into the structure and a P-Body Contact modulation of 25%.
- **25HT** - *P-Body Contact 25%* - structure with a High level of lifetime into the structure and a P-Body Contact modulation of 25%.
- **20LT** - *Orthogonal Emitter 20%* - structure with a Low level of lifetime and an Emitter diffusion modulated at 20%.
- **20HT** - *Orthogonal Emitter 20%* - structure with a High level of lifetime and an Emitter diffusion modulated at 20%.
- **50LT** - *Orthogonal Emitter 50%* - structure with a Low level of lifetime and an Emitter diffusion modulated at 50%.
- **50HT** - *Orthogonal Emitter 50%* - structure with a High level of lifetime and an Emitter diffusion modulated at 50%.

where the Low level of lifetime corresponds to the *CALIBRATION* lifetime level in Table 3.2 and the High level of lifetime corresponds to a lifetime level 10 times higher than the *CALIBRATION* level.

After the preliminary simulations from the data reported in Fig.3.33, Fig. 3.34 and in Table 3.4 the following considerations can be addressed:

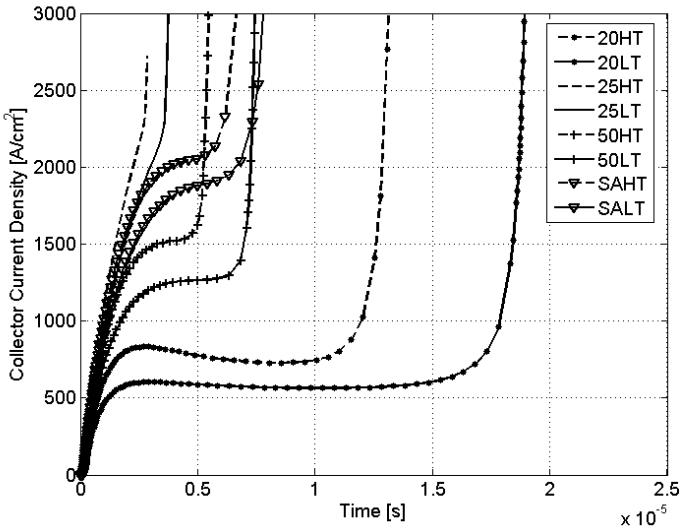


Figure 3.34: Short-Circuit curves at ambient temperature for the investigated designs.

1. The P-body modulation causes a saturation current increment.

2. The *Orthogonal Emitter* solution shows a good saturation current modulation since the saturation is roughly proportional to the emitter modulation factor.

3. Devices with the modulated P-body contact area have a reduced short-circuit capability because of two phenomena:

   - The high saturation current;
   - The wide N+ diffusion enhances the latch-up phenomenon which is accentuated by the high temperature achieved during the test.

4. The *Orthogonal Emitter* solution shows an improved capability for the Short-Circuit test since the saturation current is reduced, but it increases the V_{ON} of the device. It has to be pointed out that for high values of the emitter area the short-circuit capability is reduced anyway, even if the saturation current is lower. This is due to the wide emitter diffusion (at high emitter modulation values) which increases the current path below

the N+ diffusion and enhances the latch-up phenomenon. At ambient temperature, as shown before, the latch-up can be considered negligible for the proposed structures, but the high temperatures achieved during the short-circuit test enhances the current gain of the parasitic NPN transistor and the longer path triggers in advance the phenomenon. As a solution of this drawback, a shallow contact trench (with a depth in the order of $0.4\mu m$) can be adopted.

3.3 An optimized structure

In the previous sections the effect on the Short-Circuit capability of the geometrical and technological parameters of the single cell have investigated and the trade-offs among V_{ON}, E_{off} and short-circuit capability have been defined and quantified by means of the TCAD simulations. In this section a the effect of the modification of both the Emitter Modulation and the Channel length will be investigated. It has been shown that even if the saturation current is reduced, the Short-Circuit capability can be not improved as well, therefore three advanced designs are evaluated from the point of view of the Short-Circuit capability.Finally, the experimental results on the optimized samples will be shown.

3.3.1 An advanced Design of the cell

As mentioned before, the emitter modulation is effective in reducing the saturation current as well as the channel length modulation. The Emitter modulation technique has been investigated more in detail when the channel length is varied to achieve an optimum design that enhances both the Short-Circuit capability and the V_{ON} at the same time.

In Fig. Fig. 3.35 the three doping profiles of the channel region adopted to investigate the Short-Circuit capability of the new designs are reported. In the following the parameters used into the simulations:

1. Channel length (L_{CH}) = 2.5, 3 and 3.5
2. Emitter Modulation = 35, 50 and 75%

In Fig. 3.36 the V_{ON} voltage versus the Saturation current density is reported for the three emitter modulation percentages when the channel length is

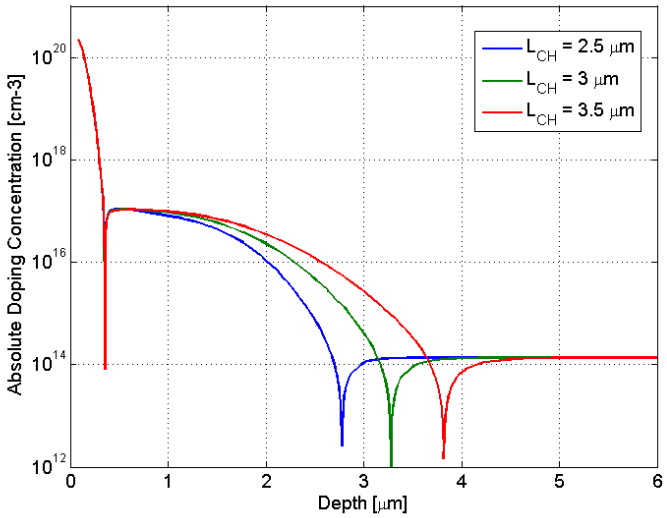


Figure 3.35: The three investigated P-Body profiles.

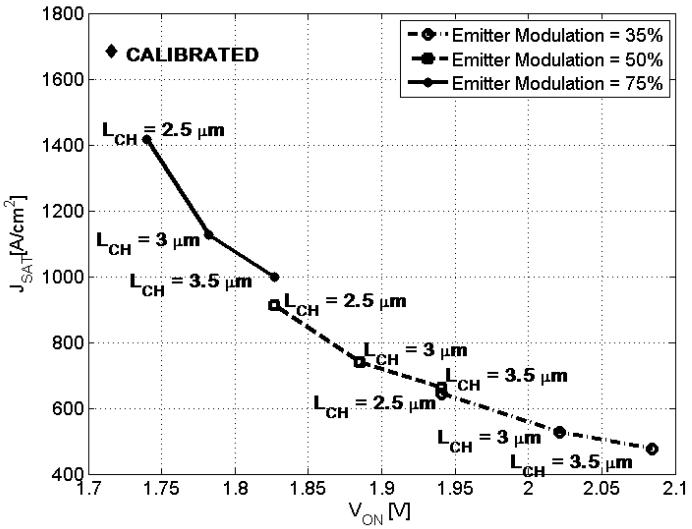


Figure 3.36: The V_{ON} voltage versus the Saturation current density is reported for the different designs for the level of lifetime. V_{ON} is evaluated at $J_C = 260A/cm^2$ and J_{sat} at $V_{CE} = 10V$

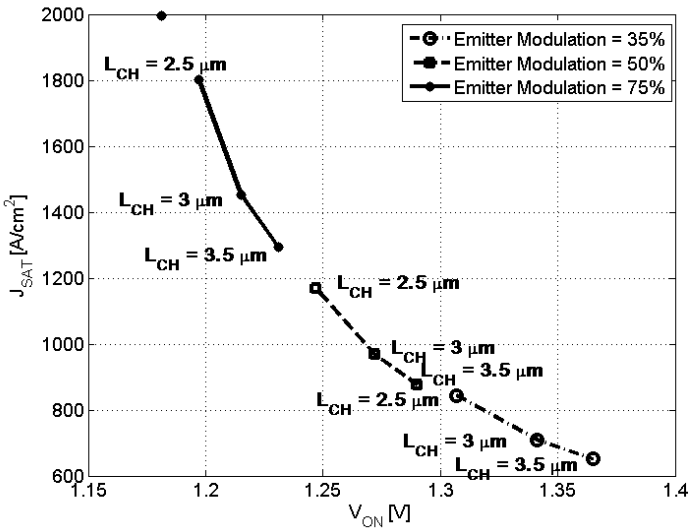


Figure 3.37: The V_{ON} voltage versus the Saturation current is reported for the different designs for the High level lifetime. V_{ON} is evaluated at $J_C = 260A/cm^2$ and J_{sat} at $V_{CE} = 10V$

varied. The lifetime level is that of the highly irradiated devices (low lifetime level).

In Fig. 3.37 the same analysis of Fig. 3.36 is reported for devices with a low level of irradiation (the lifetime is evaluated by multiplying the Low level by 10). By means of the data presented up to this section, the following conclusions have been addressed:

1. The *Orthogonal Emitter* modulation shows a trade-off between the saturation current density and the V_{ON}.

2. All of the points on the V_{ON}-J_{sat} curve lie on an hyperbolic-like curve. This behavior is not modified when the lifetime level is changed and if a point is chosen on this curve it can be achieved by acting in an equivalent way on the Channel Length or on the Emitter Modulation.

3. The above considerations suggest a saturation current modulation made by only adopting the Emitter modulation and keeping constant the channel length. This solution is preferred since it implicates the modification of only one geometrical feature of the elementary cell, the emitter modulation.

In Fig. 3.38 the three investigated structures with a different top design are reported. All of them are a derivation of the structure in Fig. 3.31 with an Emitter Modulation of 50% and they differs for the N+ diffusion design. Basically, in all of them the maximum carriers path below the N+ diffusion is reduced in order to improve the latch-up capability. More in detail, in Fig. 3.38a a shallow Contact-Trench is inserted into the top of the cell. The aim of this design is to enhance the Latch-Capability of the cell by reducing the maximum path for the carriers below the N+ diffusion. The adoption of the shallow contact trench (Depth = 0.4) has two advantages:

1. The shallow depth of the contact trench reduces the non uniformities observed in the metal layer of the Emitter(See §3.2.1).

2. The maximum path for the holes is drastically reduced since the distance between the Contact-Trench anche Trench-Gate is 0.5 .

In Fig. 3.38b the same principle of the design (a) is proposed but the interruption of the N+ Diffusion is achieved by making a N+ diffusion which is modulated and at the same time it is interrupted in the middle of the cell, where

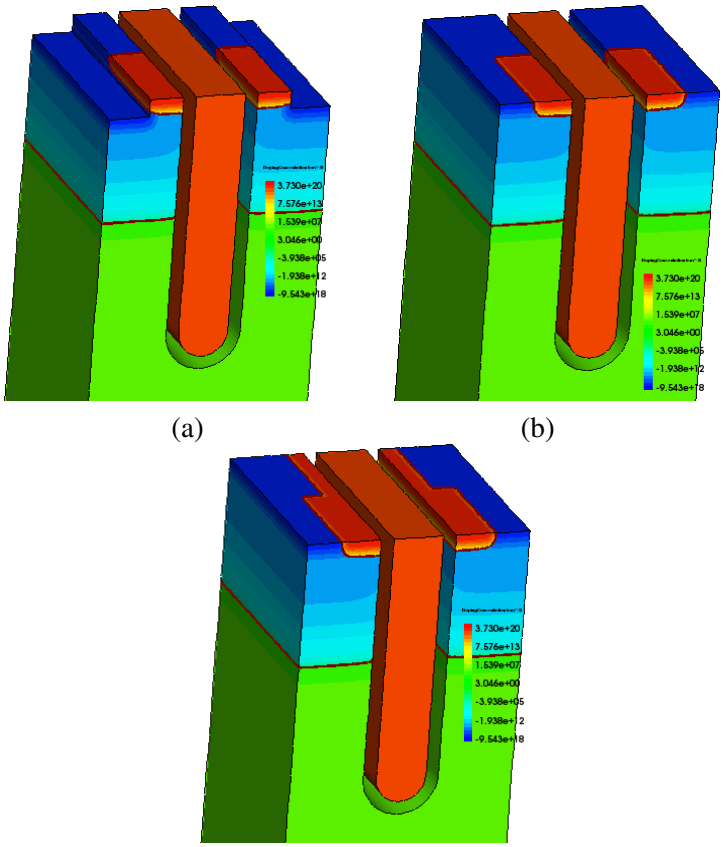


Figure 3.38: The 3D structures of the advanced designs are reported.

the P+ (P-Body contact) is extended. Finally, in Fig. 3.38c a *weak Emitter Modulation* is adopted by modulating only Silicon/metal contact region (this is achieved by masking the Emitter contact with an Oxide layer) together with a reduction of the N+ layer where the oxide mask is deposited. In Fig. 3.39 the Short-Circuit curves relative to the three designs proposed are reported, where the structure A is the one in Fig. 3.38a, the structure B is the one in Fig. 3.31 and the structure C is the one in Fig. 3.38c. The results concerning the structure 3.38bis not reported since it has the same behavior of the structure A. as it can be seen the best design is the A which has the Shallow Contact-Trench. As it can be seen the Collector Current diverges before in the structure B and, as shown from the simulations, this is due to the premature Latch-up which occurs because of the high extension of the N+ region, enhanced by the temperature increase during the Short-Circuit. The structure C shows has a premature divergence of the Collector current basically because even if the N+ diffusion is reduced in width in half part of the structure, the low resistance of the N+ region gives a good path to the electron current, therefore the latter is not modulated.

3.3.2 An advanced Design of the cell: experimental results

By means of the simulative results shown in this chapter different samples have been realized to evaluate the effective improvement of the Short-Circuit capability of the optimized structure presented before. Finally, the optimized elementary cell has the following features:

- The channel length is fixed at the one defined in §3.1.3.
- The cell pitch is 5 μm.
- The N+ emitter region has been modulated and a design with 50% of Emitter Modulation.
- Two design have been defined for the top of the cell: one without the contact trench (See Fig. 3.38b) and the other one with a *Shallow* Contact trench (See Fig. 3.38a).

In Fig. 3.40 the experimental values for the saturation current are reported for different Emitter modulation percentage (100%, 70% and 50%) at $V_{GE} = 8V$ and $V_{GE} = 10V$. As it is visible, the saturation current decreases proportionally to the Emitter modulation percentage. The devices with

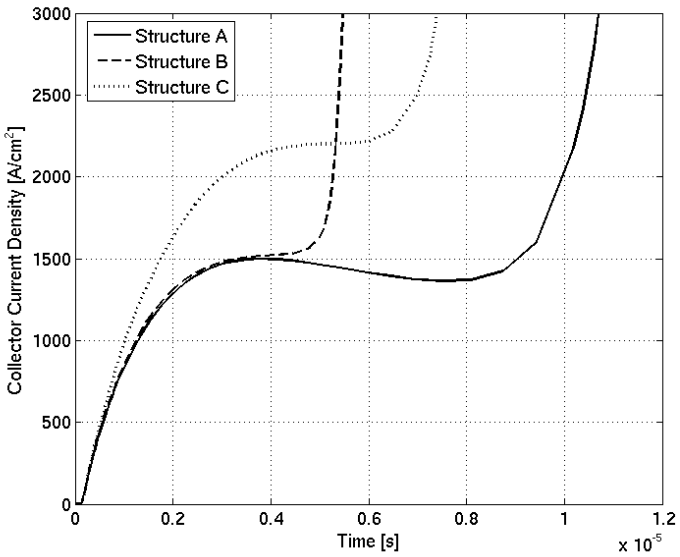
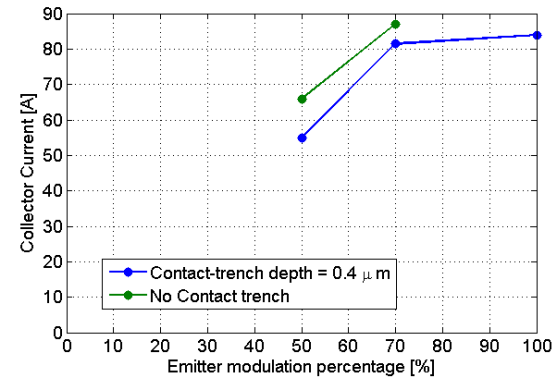
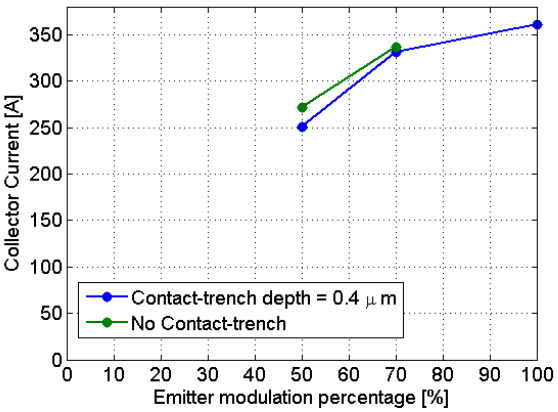


Figure 3.39: The V_{ON} voltage versus the Saturation current density is reported for the different designs for the level of lifetime. V_{ON} is evaluated at $J_C = 260A/cm^2$ and J_{sat} at $V_{CE} = 10V$

an Emitter modulation percentage of 100% have a saturation current which is lower compared to the extrapolation from the data with an Emitter modulation of 50% and 70%. This is attributed to the different batch for the devices with Emitter modulation of 100% and those with the other values of emitter modulation. An increase of the saturation current for the devices without Contact-Trench is observable.



(a)



(b)

Figure 3.40: Saturation current for the structure with and without Contact-Trench in function of the Emitter modulation percentage for $V_{GE} = 8V$ (a) and $V_{GE} = 10V$ (b).

Chapter 4

Guidelines for the design of IGBTs rugged in avalanche conditions

The IGBT are widely used in switching applications where the avalanche phenomenon can occur and an higher robustness of the active silicon power devices is requested. The avalanche breakdown phenomenon has been studied since the beginning of the silicon device development. Great effort has been given to the physical mechanisms which lead to the presence of Negative Differential Resistance (NDR) in the blocking I-V curves of different families of semiconductor devices. The subject has been investigated by means of both the analytical, experimental and simulative point of view. However a definitive explanation of the physical mechanisms which regulate the behavior in avalanche conditions is not present, especially for semiconductor power devices and in detail for IGTBs.In this chapter the avalanche breakdown phenomenon is introduced together with the concept of NDR. After the explanation of a physical model to evaluate the reverse I-V curve of the vertical PNP of the IGBT structure the avalanche breakdown voltage is investigated in the structure presented in the previous chapter by means of the TCAD simulator. Finally the design parameters which affect the I-V curves of both the active region and the termination are evaluated (in avalanche and on-state conditions) and some guidelines to improve the robustness of the overall device in avalanche conditions are defined.

4.1 Negative Differential Resistance

When the electric field of a reversely biased p/n junction reaches the critical electric field, the impact ionization generation phenomenon occurs in a reduced layer of the depletion region and the voltage drop at which this happens is called Breakdown Voltage[14]. Usually the avalanche breakdown phenomenon is considered electrically and thermally stable, since the Breakdown voltage increases when the temperature increases into the structure. This means that, especially in large area devices (e.g. power semiconductor devices), the temperature distribution on the entire area of the device is uniform after a certain lapse of time. Basically, if in a certain region of the device the temperature increases because of a current increase, locally the breakdown voltage increases, leading to a reduction of the local current density. However, this explanation of the electro-thermal behavior in avalanche conditions is not exhaustive, since both experimental and theoretical evidences have shown that semiconductor device can exhibit a current focusing which can drastically reduce the reliability of the device [86],[87],[88]. This phenomenon is due to the presence of a (NDR) into the I-V curve of the elementary structure. It has been theoretically shown by means of thermodynamical principles that when a generic structure exhibits a NDR region into the I-V curve in current-controlled conditions, the current becomes inhomogeneous in planes orthogonal to the current flow direction. More in detail the current arranges in filaments in the direction of the current flow[89]. If the device has a wide area and each vertical portion of the structure has an I-V curve in avalanche conditions with a NDR branch, from a more physical point of view this phenomenon can be addressed in this way: starting from an uniform distribution of the current at a current density value in the range of the NDR region, a infinitesimal local increase of the current causes an infinitesimal reduction of the local voltage. The local reduction of the voltage drop causes the attraction of the neighboring carriers with the relative increase of the current. In this way a strong current focusing occurs, with the relative temperature increase. As mentioned before, the temperature increase induces an increase of the Breakdown Voltage which, in case of Positive Differential Resistance, spreads the current in the surrounding part of the filament. This is not strictly true when operating in NDR. Actually this condition acts like a positive feedback which counteract the negative feedback of the temperature [86]. As shown in different works[25],[90],[21], the focusing of the current does not lead always to the premature destruction of the device since the local temperature increase leads to the filament movement in colder regions. The NDR phenomenon is due to the interaction between the

mobile charge and the electric field into the depletion region by means of the Poisson equation. The first work focused on the Avalanche breakdown subject has been presented by McKay[91] where the basic equations which model the impact ionization phenomenon has been derived from the charge discharge theory in the gas and it has been applied to a silicon p/n junction. It has been derived the multiplication factor (M) to keep into account the overall generation into the depleted region starting from a single carrier. When $M \rightarrow \infty$, the avalanche breakdown occurs in the device and the current diverges. After McKay, many works have been presented for the modeling of I-V curves in avalanche conditions [38],[92]. Even if the empirical relation derived by Miller has shown a great coherence with the experimental results, its validity is confirmed up to the trigger of the impact ionization phenomenon. Therefore a more accurate analysis of the I-V curve behavior at high current densities in avalanche conditions has been presented in [93],[94],[95]. In [93] and [94] it shown that even if the structure is a bulk one (without p/n junctions), the NDR phenomenon occurs because of the electric field warping due to the strong interaction between electric field and mobile charges. In [95] the NDR modeling has been presented for a P-i-n structure, deriving an analytical model for the reverse I-V curve. As mentioned before, the presence of the NDR in the reverse I-V curve leads to the filamentation of the current into the structure and in works like [96], an estimation of the filament size is reported from the analytical point of view starting from the physical features of a P-i-n structure. The radius of the filament is defined by the balance between the phenomenon which leads to an increase of the current into the filament and the lateral diffusion phenomenon which occurs around the filament region. It has to be pointed out that in most of these works some important lacks are present. Basically all of the models are based on strong simplifications of the equations system, e.g. [95]:

1. The carrier speed in the space charge layer is independent of the electric field
2. The ionization rate of electrons is equal to that of the holes
3. The relation between the ionization rate and the electric field is given by the relation $\alpha = aexpbE$
4. The impurity densities of the P and N regions are much higher than that of the intrinsic region
5. The intrinsic region is completely covered by the space charge region

6. Carrier recombination is negligible
7. Diffusion of carriers is negligible.

These assumptions allow to formulate a closed analytical form of the problem, but most of the information on the structure of the specific device is lost and a general description on the phenomenon is not achievable. Since the analytical approach has shown its strong limits, the behavior of the I-V curve in avalanche conditions and its effect on the current distribution in the area of the device has been investigated by means of different approaches. The most common tool for the analysis of the avalanche breakdown phenomenon in complex semiconductor devices are TCAD simulators [25],[90],[21]. Actually, they are the only way to accurately trace the reverse I-V curve of complex semiconductor devices and to analyze the carriers fluxes of two/three dimensional structures. Many phenomena have been observed by means of the TCAD simulations in avalanche conditions: in [90] a destructive effect has been observed at the termination of an IGBT and the design parameters responsible of the failure are addressed. It has to be highlighted that the impact ionization effect can occur even at a collector voltage well below the static breakdown voltage because of the Dynamic avalanche phenomenon (See §1.4.2). In [21] the filamentation phenomenon is investigated in P-i-n structures. More in detail the 1D structure is extruded and a 2D simulation is performed in turn-off conditions. During the plasma removal into the intrinsic region because of the high current which flows into the depletion region, the steepness of the electric field can increase in a way that the critical electric field is reached at lower voltage values and the dynamic avalanche phenomenon occurs together with the appearance of filaments on both the anode and the cathode sides. A more general investigation of both the electric field and the carriers distribution in to semiconductor devices is presented in [97]. An investigation of the I-V curves in avalanche conditions is reported in for different families semiconductor power devices. In [98] the limiting factors of the safe operating area are analyzed by means of TCAD simulations and the presence of current filaments is confirmed. The main drawback of the TCAD approach is the reduced area which can be analyzed. TCAD simulators recently have been used to investigate the physics involved into large area devices [33]. In this work a 3D wafer level TCAD simulation has been performed and some of the most important phenomenon which occurs during the turn-off of a GCT are investigated. The dynamic avalanche phenomenon is investigated as well. The last work has shown that it possible to use TCAD simulator also for very large areas, however it has to be pointed out that the analyzed structure is quite simple from the geometrical

point of view compared to the complexity of a Trench-IGBT structure which needs a huge number of nodes for a reliable description of a real structure. A different approach to the effect of the NDR on large area devices has been presented in different works, e.g.[99], [100], [101], [102]. In these works the effect of the NDR is treated by means of the non-linear spatio-temporal dynamics theory. The core of this theory is a reaction-diffusion approach for the description of the current and voltage dynamics into the wide area devices in a generic circuit. The theoretical results have shown a good coherence with the experimental observations. As shown by means of the TCAD simulations, the current filaments can move from their original position because of the temperature effect. A deep analytical investigation of the solutions stability in the system has been carried out. The weak point of this approach is the adoption of the reaction-diffusion modeling to describe the pattern which can occur in the structure in avalanche conditions. To use this kind of approach, the *kinetic function* of the device (1D) has to be defined in order to get realistic results. However for most of the structures it is too complicate to derive the *kinetic function* and an exhaustive simulation of the behavior in avalanche conditions has been achieved only for simplified structures. Recently an alternative approach has been proposed to investigate the current and voltage dynamics in large area devices [65]. In this work a 3D electro-thermal simulator for high power devices operating in avalanche condition is proposed. The simulation is based on two coupled systems: a 3D-FEM thermal simulator and a 2D electrical solver and it is capable of simulating a large number of macro-cells composing a wide-area power devices operating in avalanche condition. The electrical solver uses a SPICE-like algorithm with a look-up-table description for every cell, while The thermal problem is solved by a finite element method (FEM) in an iterative loop with the electrical simulator. The Unclamped Inductive Switching (UIS) has been performed to analyze the current dynamics into a Trench-IGBT structure. Up to now this approach is the only one effective in simulating large area devices.

4.2 Analytical model in avalanche conditions

In this section the analytical model to trace the I-V curve in avalanche conditions is presented for an open-base PNP, where the Base doping profile is not uniform. The mentioned structure is representative of the vertical PNP of the IGBT structure. The model is based on a linear approximation of the electric field into the depletion region (the steepness depends on the mobile carriers

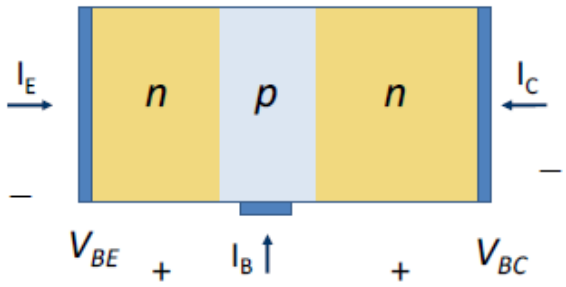


Figure 4.1: Reference current and voltage for an *npn* structure.

flowing into the depletion region) and an accurate evaluation of the β of the PNP is achieved by means of the Gummel and Poon model [103]. The scope of this model is linked to the need of a tool which gives a quick estimation of the I-V curve in avalanche conditions and the correlation with the design parameters which characterize the structure under test. Moreover, a correct estimation of the I-V curve in avalanche conditions can improve the predictability of the simulations performed on large area power devices [65].

4.2.1 Gummel and Poon Model

In this section the Gummel and Poon model [103] is recalled, since the proposed model is based on a charge-control approach. This model is widely used to describe the *bjt* behavior since it has a good predictability in very general hypothesis. Actually it works for devices with non uniform doping profiles and for any injection level. The 1D reference structure is reported in Fig. 4.1 where the conventions for voltages and currents are reported.

Starting from the electron and hole current equations in semiconductors:

$$J_n = q\mu_n nE + qD_{nB}\frac{dn}{dx} \tag{4.1}$$

$$J_p = q\mu_p pE - qD_{pB}\frac{dp}{dx} \tag{4.2}$$

where the electric field can be obtained from (4.2) and it can be substituted in (4.1). The condition $h_{FE} << \frac{n\mu_n}{p\mu_p}$ is always verified for low injection levels and it is still valid at high injection levels if $h_{FE} << \frac{\mu_n}{\mu_p}$. Therefore the collector current can is(for a *npn* structure) like:

$$I_C = \frac{I_s}{q_b}\left[\left(e^{\frac{V_{BE}}{V_T}} - 1\right) - \left(e^{\frac{V_{BC}}{V_T}} - 1\right)\right] \tag{4.3}$$

I_s is the reverse saturation current and it can be written as:

$$I_S = \frac{qA\left\langle D_{nB}\right\rangle n_i^2}{\int p_0 dx} \tag{4.4}$$

where the integral is extended to the Base neutral region and q_b can be calculated by the quadratic equation:

$$q_b = q_1 + \frac{q_2}{q_b} \tag{4.5}$$

.

The solution of equation (4.5) gives:

$$q_b = \frac{q_1}{2}\left(1 + \sqrt{1 + 4\frac{q_2}{q_1^2}}\right) \tag{4.6}$$

where q_1 is the sum of the zero polarization charge and the charge due to the junction capacitance. It keeps into account the Early effect, since V_{AF} and V_{AR} are the forward and reverse Early voltage, respectively.

$$q_1 = 1 + \frac{V_{BE}}{V_{AR}} + \frac{V_{BC}}{V_{AF}} \tag{4.7}$$

q_2 represents the Base excess carriers due to the diffusion capacitance, that is dependent on current and voltage:

$$q_2 = \frac{I_s}{Q_{B0}}\left[\tau_{BF}\left(e^{\frac{V_{BE}}{V_T}} - 1\right) - \tau_{BR}\left(e^{\frac{V_{BC}}{V_T}} - 1\right)\right] \tag{4.8}$$

where Q_{B0} is the is the charge associated to the holes in Base when it is not polarized. τ_{BF} and τ_{BR} are the transit time in Base in forward and reverse condition respectively.

In Fig. 4.2 the charges due to the junction capacitance are reported in detail. In forward conditions the above relations can be simplified in the following way:

$$q_1 = 1 + \frac{V_{BE}}{V_{AR}} \cong 1 \tag{4.9}$$

$$q_2 = \frac{I_s}{Q_{B0}}\tau_{BF}\left(e^{\frac{V_{BE}}{V_T}} - 1\right) \tag{4.10}$$

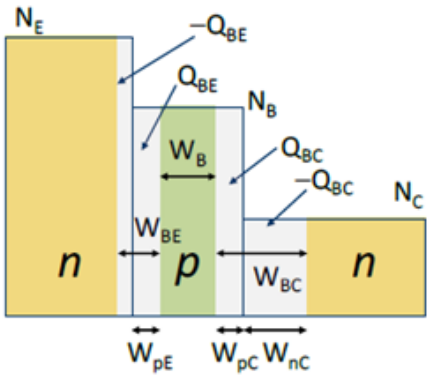


Figure 4.2: Charges due to the junction capacitance.

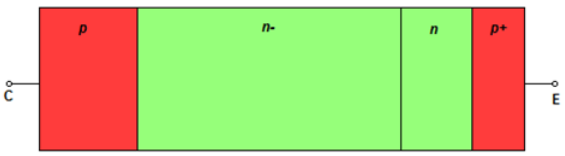


Figure 4.3: Vertical PNP of an IGBT.

$$I_C = \frac{I_s}{0.5 \cdot \left(1 + \sqrt{1 + 4q_2}\right)} \left(e^{\frac{V_{BE}}{V_T}} - 1 \right) \tag{4.11}$$

4.2.2 The proposed model

In this section the proposed physical model to trace the I-V curve of a PNP in avalanche condition is explained. In Fig. 4.3 a sketch of the investigated PNP structure is reported.

When the IGBT works in forward blocking modality, the breakdown voltage occurs between the Emitter and the Collector of the vertical PNP structure of the IGBT (See Fig. 4.3). In this condition the Emitter/Base junction is directly biased, while the Collector/Base junction is reversely biased. This structure in avalanche conditions (when the MOS channel is closed) works like a PNP in open-base configuration. The β value can be obtained by considering the PNP working in forward biasing condition. It has to be pointed out that the vertical PNP structure of an IGBT is not optimized to achieve an high β

value, since the base width is quite long and the doping concentration, especially in the Drift-Layer, is very low if compared to signal *bjt*. Therefore, the high injection condition can be achieved at low current levels, with the relative reduction of the β. In the general case, depending on the design of the IGBT, the electric field in avalanche conditions can have a shape which can be[104]:

- Trapezoidal - The electric field is almost constant into the Drift-layer and it truncated by the Buffer-layer (Punch Through design).
- Triangular - The electric field goes to zero into the Drift-layer. In this case the Buffer-layer is not present (Non Punch Through design)
- Quasi-triangular - The Drift-layer width is engineered in a way that the electric field falls to zero into the low doped Field-Stop layer and the overall shape is quasi-triangular (Field Stop design)

Therefore it can be firstly concluded that the Base of the PNP is not uniform from the point of view of the doping profile. As it will be illustrated ahead, the Breakdown Voltage of the structure is not constant when the current density changes. Actually, especially at high current density levels, the mobile charge flow into the depletion region warps the electric field and a reduction or increase of the depletion region occurs. This effect changes the geometrical features of the undepleted layer of the base. For example, even if nominally the electric field for a Field-Stop structure is quasi-triangular, it can happen that for a certain current level the hole current density is higher than the electron current density. In this case the slope of the electric field increases with the relative reduction of the depletion region. This change for the depleted region implies an increase of the width of the neutral base and, therefore, a change of the β. It can happen that the electric field becomes zero before than the transition between Drift-layer and Field-Stop layer. The β calculation in the latter case requires a modeling which keeps into account the not uniform doping profile. The following assumption have been adopted to formulate the model:

- The structure is 1D and the temperature effect is not kept into account.
- The current level is forced and it is supposed that the impact ionization phenomenon is present.
- Carriers move into the depletion region only because of the drift phenomenon at the saturation velocity.

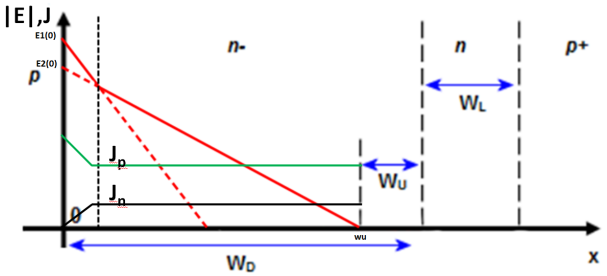


Figure 4.4: Electric Field, current densities and geometrical parameters for the avalanche model.

- The Electric-Field is approximated with a linear behavior.
- The recombination phenomena are neglected into the neutral region.
- The impact ionization phenomenon is confined in reduced zone at the Collector/Base junction. Therefore, since the Electric field drives the electrons toward the Emitter region and the holes into the opposite direction, the depletion region has been divided in two regions: the first one is the region in which the impact ionization occurs and the hole current is assumed egual to the total forced current; in the second region the impact ionization is not present and the carriers move at the saturated velocities.

The first step for the implementation of the model is the definition of the Electric Field shape when a certain current is applied. The Electric Field is defined by the Poisson equation:

$$\frac{dE\left(x\right)}{dx} = \frac{q}{\epsilon_s}\rho \tag{4.12}$$

where q is the electron charge, ϵ_s is the silicon permittivity and ρ is the charge density. Usually the Electric Field shape is calculated keep into account only the fixed charge of the background doping concentration. However, when the current density increases, the charge density has to include the mobile charge as well by means of the equation $\rho = (N_D + p - n)$.

In Fig. 4.4 the sketch of the structure is reported together with the approximated Electric Field and current densities of electron and hole is reported. W_U is the undepleted part of the Drift-layer, W_D is the Drift-layer thickness, W_L is

the thickness of the Buffer layer and the wu is the abscissa at which the Electric Field E_2 goes to zero (this value can be higher than W_D and in that case the Electric Filed is truncated at the n-/n transition region). The depletion region is divided into two regions. In the first region, at the Collector/Base junction, because of the impact ionization phenomenon the holes are the dominant carrier present in this zone. Therefore the charge density can be approximated by $\rho = N_D + p = N_D + \frac{J}{qv_{psat}}$, where J is the total current forced into the structure. In the second region the impact ionization effect does not occur and both carriers move with their respective saturated velocity. In this region the net density of charge can be calculated by $\rho = N_D + p - n = N_D + \frac{J_p}{qv_{psat}} - \frac{J_n}{qv_{nsat}}$. The ratio between holes and electrons into the second region is defined by the β of the PNP structure. By means of the above approximations into the depletion region, the Electric Filed into the first region can be written as:

$$E_1(x) = E_1(0) + \frac{q}{\varepsilon_s}\left(N_D + \frac{J}{qv_{ps}}\right)x \tag{4.13}$$

where [15]

$$E_1(0) = -\frac{4 \cdot 10^5}{1 - \frac{1}{3}\log_{10}\left[\left(N_D + \frac{J}{qv_{ps}}\right)/10^{16}\right]} \tag{4.14}$$

By means of equation (4.14), it has been kept into account the critical Electrical Field dependence on the effective charge density at the Collector/Base junction. In the same way the Electric field into the second region can be obtained by:

$$E_2(x) = E_2(0) + \frac{q}{\varepsilon_s}\left(N_D + \frac{J}{q(\beta+1)}\left(\frac{\beta}{v_{ps}} - \frac{1}{v_{ns}}\right)\right)x \tag{4.15}$$

where [15]

$$E_2(0) = -\frac{4 \cdot 10^5}{1 - \frac{1}{3}\log_{10}\left[\left(N_D + \frac{J}{q(\beta+1)}\left(\frac{\beta}{v_{ps}} - \frac{1}{v_{ns}}\right)\right)/10^{16}\right]} \tag{4.16}$$

As mentioned before the β defines the ratio between holes and electrons into the second region of the depleted part of the device. More in detail the β of the structure is defined at the Emitter/Buffer-layer junction, but, as reported before, the recombination phenomenon is neglected in the Base and the fluxes

ratio does not change. The electron current can be considered the Base current of the PNP since it is the electron flux injected into the Emitter by the Base region. The Collector current can be approximated with the hole current which flows into the second region of the depleted region. Therefore the β of the structure can be evaluated like the ratio $\frac{J_C}{J_B}$, where J_C is evaluated by (4.9)-(4.11) for a PNP. J_B is the sum of the reverse saturation current of the Base/Emitter junction and the generation current in the SCR region of the same junction. The Emitter current can be evaluated by:

$$J_E = \left(q\frac{D_{pL} n_i^2}{N_L W_L} + q\frac{D_{nE} n_{iE}^2}{N_E W_E} \right) \left(e^{\frac{V_{EB}}{V_T}} - 1 \right) + \left(q\frac{W n_i}{2\tau_{SC}} \right) \left(e^{\frac{V_{EB}}{2V_T}} - 1 \right). \tag{4.17}$$

From the last equation the voltage drop at the Emitter/Base junction is obtained and it is used to evaluate J_C and J_B. The second term of the equation is due to the generation current in the Emitter/Base junction. The ideality factor of the junction has been fixed at 2, the depletion layer width has been fixed to the value at zero polarization. As mentioned before, depending on the current density level, the depletion layer can reduce its size and part of the Drift-layer can become neutral. Therefore the Base charge at zero polarization (Q_{B0}), the Base Time of flight (τ_{bF}) and the reverse saturation current density (J_{SS}) have to be modified in the charge control model to keep into account the not constant doping profile of the Base:

$$Q_{B0} = \begin{cases} qN_L W_L & \omega_U \geq W_D \\ q\left(N_L W_L + N_D\left(W_D - \omega_U\right)\right) & \omega_U < W_D \end{cases} \tag{4.18}$$

$$\tau_{BF} = \begin{cases} \frac{W_L^2}{2D_{pL}} & \omega_U \geq W_D \\ \frac{W_L^2}{2D_{pL}} + \frac{(W_D - \omega_U)^2}{2D_{pU}} & \omega_U < W_D \end{cases} \tag{4.19}$$

$$J_{SS} = \begin{cases} \frac{qD_{pL} n_i^2}{N_L W_L} & \omega_U \geq W_D \\ \frac{q\langle D_p \rangle n_i^2}{N_L W_L + N_D (W_D - \omega_U)} & \omega_U < W_D \end{cases} \tag{4.20}$$

with

$$\langle D_p \rangle = \frac{1}{\left(W_L + \left(W_D - \omega_U\right)\right)} \left(W_L D_{pL} + \left(W_D - \omega_U\right) D_{pU}\right) \tag{4.21}$$

where D_{pL} and D_{pU} are the diffusivity into the Buffer-Layer and into the undepleted Drift-layer respectively. ω_U is obtained by:

$$\omega_U = -\frac{E_2(0)}{A_2} \tag{4.22}$$

where A_2 is the slope of $E_2(x)$ in (4.15). Finally it can be concluded that: The electric field depends on the forced current by the Electric Field peak (Eq. (4.14) and (4.16)) and by β; β depends on the forced current and ω_U; ω_U depends on the Electric Field shape. The breakdown voltage for low current levels (where the mobile charge are negligible compred to the fixed charge) has been evaluated by following the approach in [14] to keep into account the β effect. More in detail a corrective factor for the maximum Electric Filed has been evaluated by:

$$\alpha_{PNP} M = \frac{\alpha_{PNP}}{1 - \left(\frac{V_C}{BV_{pp}}\right)^n} = 1 \tag{4.23}$$

where BV_{PP} is the Breakdown Voltage of a planar abrupt p/n junction where the n-region has the same doping concentration of the Drift-Layer. V_C is the Collector voltage, however when the Electric Field is truncated by the Buffer layer, the effective value of V_C has to be evaluated by the approach in [14]. More in detail the it has to be calculated the equivalent Non Punch Through voltage by:

$$V_{NPT} = \frac{\varepsilon_s}{2qN_D}\left(\frac{V_C}{W_D} + \frac{qN_D W_D}{2\varepsilon_s}\right). \tag{4.24}$$

4.2.3 Results

The solutions of the presented set of equations has been solved by means of a Spice solver and the results have been compared to the TCAD simulations. In Fig. 4.5 it is reported the FS structure used to validate the model by means of the TCAD simulations.

More in detail the Drift-layer has a width $W_D = 160\mu m$ and a doping concentration $N_D = 9E13cm^{-3}$. The Emitter has a doping concentration $N_E = 9E17cm^{-3}$ and $W_E = 0.5\mu m$. The Buffer-Layer parameters have been varied in order the verify the predictability of the model and the I-V curves have been traced for different combinations of the Buffer-Layer width and doping concentration in the range of the values reported into the literature [105] [106] (See Table 4.1).

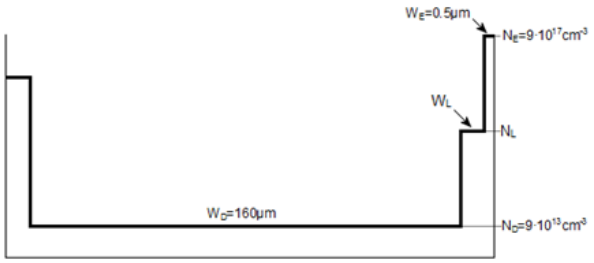


Figure 4.5: Reference structure for the comparison between the TCAD simulations and the model.

Table 4.1: Buffer-Layer parameters.

$W_L\,(\mu m)$	$N_L\,(cm^{-3})$
1	1E15
5	1E15
10	1E15
1	5E15
5	5E15
10	5E15
1	1E16
5	1E16
10	1E16

In Fig. 4.6 the most relevant parameters of the I-V curve in avalanche conditions are highlighted. The Breakdown Voltage defines the voltage rate of the device, while the NDR slope defines the robustness of the device in avalanche conditions, as it will be shown in the next sections.

In the following the results of the comparison between the TCAD simulations and the model are reported.

From the illustrated results in Fig. 4.7, 4.8 and 4.9 it can be concluded that the model shows a good predictability when the buffer layer has a thickness of $5m$ and $10m$ from both the point of view of the Breakdown voltage and the I-V curve slope in the NDR region. The validity of the model form the point of view of the maximum current density is mainly limited by the high injection regime which occurs at low current levels because of the very low doping concentration into the Drift-Layer [107]. Actually in a Field-Stop structure at low

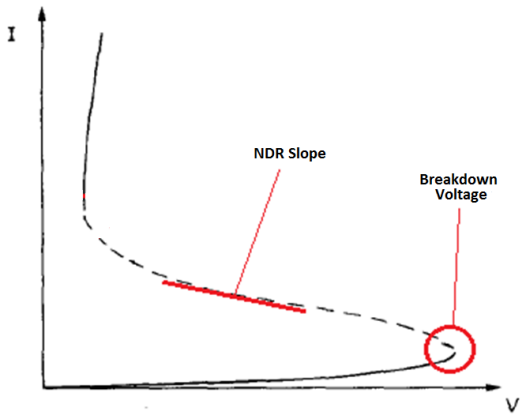
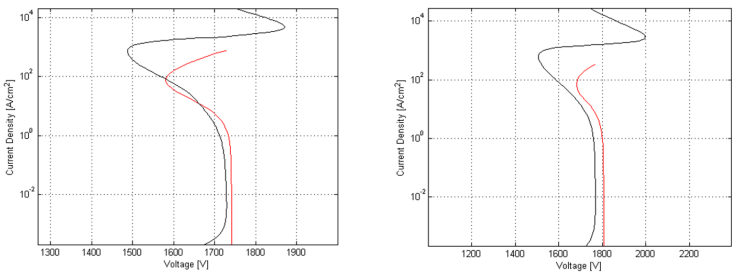


Figure 4.6: Relevant parameters of the I-V curve in avalanche conditions.



(a) (b)

Figure 4.7: Comparison between the numerical simulations (black curves) and the model curves (red curves) are reported for $W_L = 10\mu m$. (a) $N_L = 5E15cm^{-3}$ and (b) $N_L = 1E16cm^{-3}$.

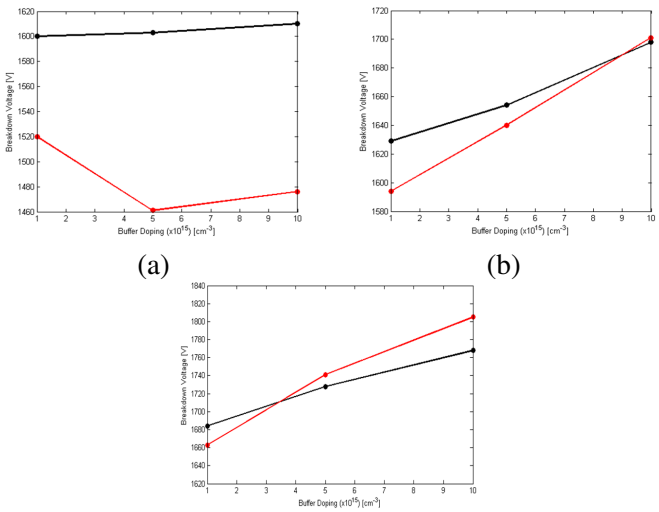


Figure 4.8: Comparison between the numerical simulations (black curves) and the model curves (red curves) of the Breakdown Voltage as a function of N_L. (a) $W_L = 1m$, (b) $W_L = 5\mu m$ and (c) $W_L = 10m$.

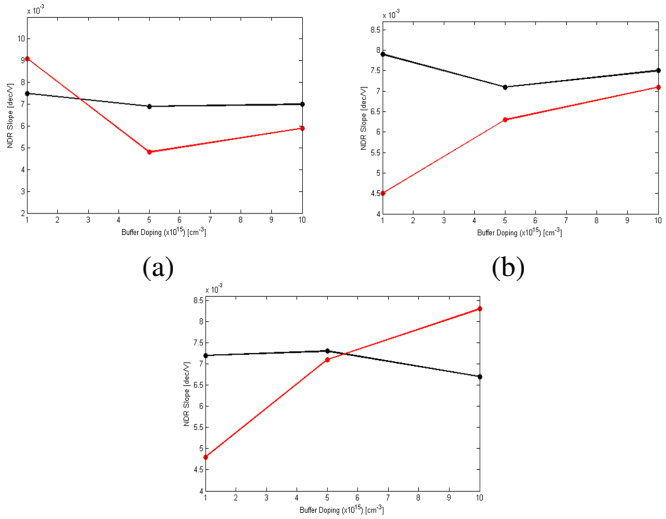


Figure 4.9: Comparison between the numerical simulations (black curves) and the model curves (red curves) of the NDR slope as a function of N_L. (a) $W_L = 1m$, (b) $W_L = 5\mu m$ and (c) $W_L = 10m$.

current densities the Electric-Field has a quasi-triangular shape and the it has a low value at the Drift-Layer/Field-Stop Layer. However as soon as the mobile charge becomes comparable to the Drift-Layer doping concentration, whether the β of the structure is still greater than one, the depletion region shrinks and part of the Drift-Layer becomes undepleted and the high injection regime is reached. The validity of the model is also reduced when the Buffer-Layer has thickness of $1\mu m$. In this case as soon as a small part of the Drift-Layer becomes undepleted, the high injection phenomenon occurs in most of the neutral part and the overall predictability is drastically reduced.

4.3 TCAD simulations in avalanche conditions

As mentioned before (§4.1), the avalanche breakdown is a strongly unstable phenomenon and an evaluation of the behavior of complex structures can be achieved only by means TCAD simulators. Actually when the 2D or 3D effects occurs the analytical approach becomes very complex and only a strongly approximated evaluation can be achieved. A TCAD simulation approach is mandatory when the geometrical features of the structure can not be kept into account (e.g. Trench-Gate structure, Field-Plates, more complicate termination structures) in an analytical way. When an IGBT works in avalanche conditions, differently from the on-state, the current can flow either in the termination or in the active area, depending on the features of the two regions. In this section different design parameters of the structure presented in the previous chapter (See Fig. 3.5) will be analyzed to evaluate their effect on the avalanche I-V curves of the termination and the active area. In Fig. 4.10 the analyzed termination is reported. As it can be seen, the Floating Field Ring (FFR) technique is adopted by means of three FFR which have a junction depth different from the one of the P-Body in the active region. A Field Plate is adopted on the last FFR to warp the rest of the Electric Field at the end of the termination. On the left side the Gate bus is visible above the hookup ring of the termination.

As it will be shown, the improvement of the avalanche ruggedness can reduce the on-state performances, therefore an accurate evaluation of the tradeoff among the improvements in the avalanche conditions and those in the on-state have to be taken into account. Usually the avalanche phenomenon occurs at the termination (as confirmed by the investigated structure), therefore the effect of the lifetime, of the Buffer-Layer and of the Floating Field Ring structure will be analyzed and their effect on the on-state performances will be evaluated. The effect of a bad Emitter contact, due to the recessed solution, will

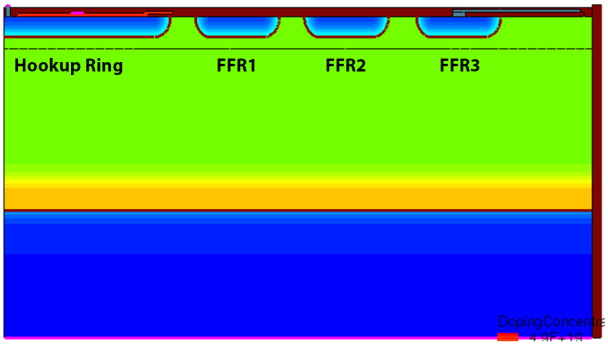


Figure 4.10: Investigated termination structure.

be investigated. Finally, the interaction between the active area will be analyzed and the guideline for an improvement of the avalanche ruggedness of the overall device will be shown.

4.3.1 Lifetime effect

As shown into the previous chapter, the lifetime of the carriers into the structure strongly affects the on-state performances and this is due mainly to the vertical PNP of the IGBT structure. In this section the effect of the lifetime on the reverse I-V curves of both the active area and the termination will be investigated. The reference structure is the one presented in the previous chapter for the active area (See Fig. 3.5) and the reference lifetime is the one reported for the calibrated structure in Table 3.1. It has to point out that the data in Table 3.1 refer to a structure where the lifetime has been strongly reduced by means of the electron irradiation lifetime technique. Therefore three lifetime levels have been investigated for this analysis: the calibrated level reported in Table 3.1 (CALIBRATED), ten times the calibrated value (10-CAL)and hundred times the calibrated value (100-CAL). The TCAD simulations have been performed by using the UniBo2 [108], [109], [110] model to keep into account the impact ionization rate dependence on the Electric Field and on the temperature (This model will be adopted for all of the avalanche simulations in this work). In Fig. 4.11 the I-V avalanche curve of the termination region are reported for T = 300K and T = 400K and for the three levels of lifetime defined before. In Fig. 4.12 the I-V curves of the active region are reported for the same conditions of Fig.4.11.

From the TCAD simulations the following considerations cab be ad-

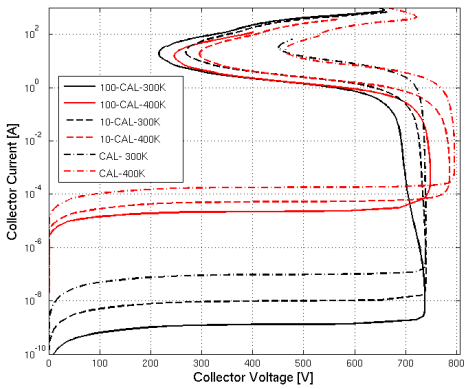


Figure 4.11: I-V curves of the termination region for different lifetime levels and two temperatures.

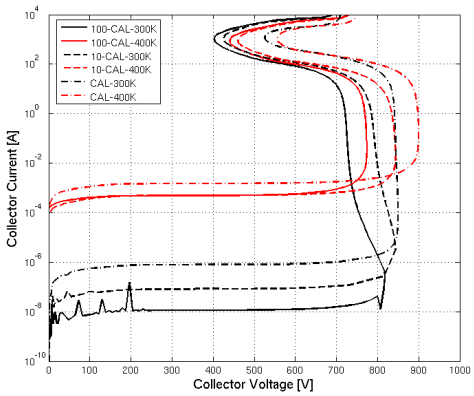


Figure 4.12: I-V curves of the active area region for different lifetime levels and two temperatures.

dressed:

1. The breakdown voltage (at low currents) in the active region is dependent on the lifetime at ambient temperature and it increases when the lifetime level decreases.The same behavior occurs when the I-V curves are simulated at T=400K. This behavior can be explained be means of the equation (4.23). When the lifetime increases, the α_{PNP} increases and the avalanche condition is reached at lower Collector Voltage.

2. The breakdown voltage (at low currents) in the termination region is not dependent on the lifetime at ambient temperature. At T=400K the breakdown voltage increases and it increases when the lifetime level decreases. This behavior can be partially explained be means of the equation (4.23). When the lifetime increases, the α_{PNP} increases and the avalanche condition is reached at lower Collector Voltage. However, the termination structure is strongly affected by 2D multiple coupled phenomena and a simplified dissertation can not explain the behavior. This aspect is illustrated in Fig. 4.13 where the electron flux into the termination is reported for structures with the three lifetime levels (CAL, 10-CAL and 100-CAL) at the same current level. As it can be seen, for the lowest level of lifetime the current flows mainly on the curvature of the Hookup Ring. When the lifetime level increases a secondary path arises at the interface Si/Ox. The latter condition is a more complex behavior which can be investigated only by means of the TCAD simulator.

3. At high current levels, when the lifetime level increases, the current level of the transition knee between the vertical region and the NDR region decreases. At the same time, up to current below the knee region, the steepness of the I-V curve decreases.

4. For both the termination and the active region the current level at which the transition from the NDR to the Positive Differential Resistance (PDR) increases when the lifetime level decreases.

5. From the point of view of the On-State performances, the analysis shown in §3.2.1 is complete.

4.3.2 Buffer layer effect

The Buffer-Layer features have a great impact on the avalanche behavior of the IGBTs [111] as well as other structures [48]. The Buffer-Layer importance

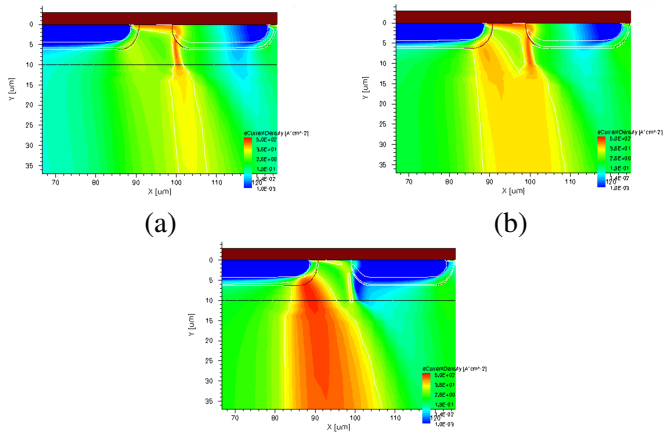


Figure 4.13: Electron flux in the termination region (Between the Hookup ring and the first FFR) at the same current level.(a) 100-CAL lifetime level, (b) 10-CAL lifetime level and (c) CAL lifetime level.

is related to the its capability to regulate the current level at which the double injection occurs as well as its effect on the β of the vertical PNP of an IGBT. Usually the Buffer-Layer features are the same for the active area and the termination. Therefore, in this section the Buffer-Layer effect will be investigated in avalanche as well as in on-state conditions. In Fig. 4.14 the I-V curves in avalanche conditions are reported for different Buffer-Layers, with a lifetime level *10-CAL*. The I-V curves are compared the one of the calibrated structure. As it will be illustrated, the structure with a lower lifetime exhibits am higher avalanche ruggedness, therefore the I-V curve of the calibrated structure is the reference for the ones with an higher value of the lifetime.

From the results reported in Fig. 4.14 it can be concluded that:

1. The slope of the NDR region increases when the Buffer-Layer width or the doping concentration increase.

2. The current level at which the knee between the vertical region of the I-V curve and the NDR region increases when the Buffer-Layer width or the doping concentration increase. This effect is equivalent to the lifetime reduction in the structure and basically is due to the increases of the recombination rate in the Buffer region.

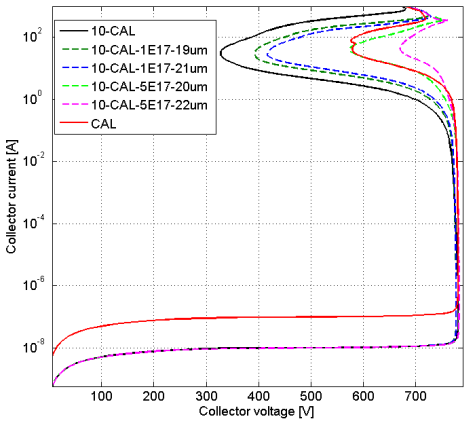
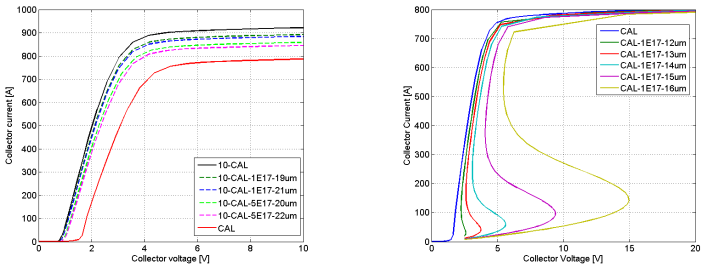


Figure 4.14: I-V curves of the termination region for different Buffer configurations.

3. The Breakdown Voltage does not change when the Buffer-Layer features change.

The I-V curves in the active area follow the behavior of the termination. Even if the ruggedness in avalanche conditions improves with the increase of the Buffer-Layer doping concentration or width, these modifications of the design provoke a variation of the on-state performances. In Fig. 4.15, the output curves (at $V_{GE} = 15V$) are reported for two lifetime levels. In Fig. 4.15(a) the output curves are reported for a lifetime level ten times higher than the calibrated one. As it can be visible, when the Buffer-Layer width or the doping concentration increases the V_{ON} of the device increases and in Table 4.2 the voltage drops are reported for the different structures together with the turn-off losses. N_B is the doping concentration of the Buffer-Layer and W_B is the width of the Buffer-Layer. From Table 4.2 it is clear that the turn-off energy decreases when the Buffer-Layer width or doping concentration increases and this is due basically because of the higher recombination rate into the Buffer which reduces the time to sweep out the charge stored into the Drift-Layer.

In Fig. 4.15(b) the output curves relative to the structure with the calibrated lifetime level are reported. The Buffer-Layer features are not the same of that in Fig. 4.15(a) because the anomalous behavior for the V_{ON} becomes too relevant. Therefore a smaller variation of the Buffer-Layer width is in-



(a) (b)

Figure 4.15: Output curves of the active area for the (a) Calibrated lifetime level and for (b) a lifetime level ten times higher.

Table 4.2: Buffer-Layer effect on the on-state parameters

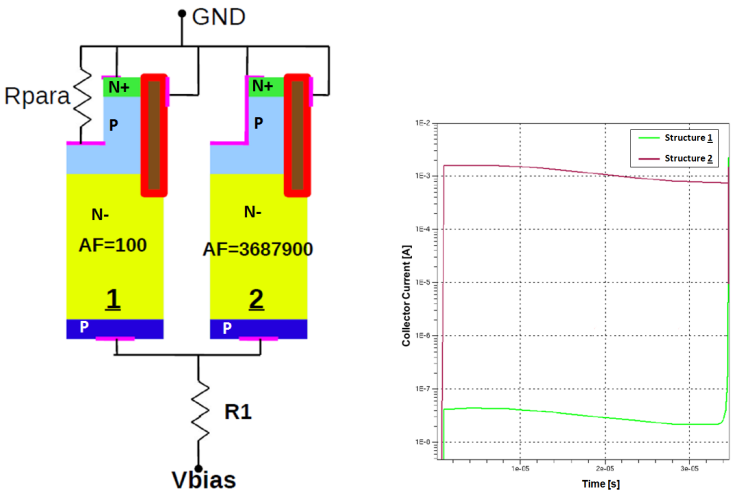
Design	$N_B\ (cm^{-3})$	$W_B\ (\mu m)$	$V_{ON}(V)$	$E_{off}(mJ)$
10-CAL	1E17	10	1.134	41
10-CAL-1E17-19um	1E17	19	1.175	26
10-CAL-1E17-21um	1E17	21	1.190	24
10-CAL-5E17-20um	5E17	20	1.266	13
10-CAL-5E17-22um	5E17	22	1.311	11
CAL	1E17	10	1.865	5

vestigated. When the Buffer-Layer width increases, a NDR is visible in the output curve and the more the width increases, the more this phenomenon is relevant. This is due to the variation of the current level at which the conductivity modulation phenomenon occurs into the Drift-Layer. Actually, in the Buffer-Layer region the lifetime decreases because of the high doping concentration and the more the Buffer-Layer increases, the more the holes injected by the Collector recombine. Therefore the current level at which the conductivity modulation occurs increases and the voltage drop between Collector and Emitter stays high. When the the conductivity modulation occurs the V_{CE} reduces and a NDR region rises up.

4.3.3 Effect of a non-ideal P-Body contact

When the device works in avalanche conditions, even if the structure does not exhibits instabilities like NDR regions, the presence of defects into the structure can reduce the reliability of the overall structure. In this section the effect of a bad Emitter contact is analyzed for the structure presented in Chapter 3. As reported in Fig. 3.5, the structure has a recessed Emitter Contact which has the task to improve the Latch-up capability (See §3.2.2). The simulation results presented in the previous chapter demonstrate that the recessed trench technology adopted to make the Emitter contact is effective in improve the Latch-up robustness and the more the depth increases, the more the Latch-up capability improves. However the simulations do not keep into account the eventuality of a bad contact between the metal and the silicon in the recessed area. From the technological point of view the presence of the Contact-Trench provokes some imperfections in the metal layer upon this region. This effect is well visible in Fig. 3.1 where just upon the Contact-Trench region a pronounced vacancy in the metal layer is visible. Together with this defect, more accurate SEM sections have shown that in some areas of the device the contact between metal and silicon is not perfect, especially into the recessed area. The latter drawback has been investigated by means of TCAD simulations.

In Fig. 4.16(a) a sketch of the simulated structure is reported. It is supposed that a bad contact is present in a reduced area of the device and the ration between the bad-connected area and the well connected area is 100/3687900. In the structure 1 the bad contact between the metal and the silicon in the recessed region is schematized with a resistance R_{PARA} which has a value of 2E7 Ω. R_1 is an high value resistance which coupled to an high voltage generator has been used to emulate a current generator. A current pulse of $750\mu A$ has been applied to the parallel of the two structures with $t_{on} = 8ms$. As it is visible, after a certain transient, the Collector Current of the structure 1 quickly rises up to the forced current because of the Latch-up phenomenon which occurs of the voltage drop on the R_{PARA} resistance. The strong focusing of the current leads to a quick increases of the local temperature a premature failure of the device can occur. It has be pointed out that the Latch-up phenomenon is electrical since the temperature increases is negligible before it occurs and the current rises after $35\mu s$. These results suggest that the bad contact condition have to be avoided and from the technological point of view this has been implemented either by reducing the depth of the Contact-Trench or by removing it and acting on the Emitter diffusion to reduce the Latch-up phenomenon (See §3.3).



(a) (b)

Figure 4.16: (a) A sketch of the performed TCAD simulation in mixed-mode. (b) Collector currents of the two structures during the avalanche test.

4.3.4 TCAD simulations in large area structures

As shown in to the previous sections, both the active area and the termination have a pronounced NDR region in the reverse I-V curves, even at high temperatures. In §4.1 the effect of the NDR region in wide area is investigated by means of the results shown up to now in the literature. All of simulations performed in 2D large area structures shows that, especially when a large current flows into the structure, the current focus and on or more filaments rise up in the section of the device. This is observed mainly in dynamic avalanche conditions for P-i-n structures, GTO and Trench-IGTBs. However, it has to be pointed out that in all of the investigated structures some elements which can trigger the filamentation of the current are present. In IGBT structure the Trench-Gate structures are the main element which can trigger the focusing of the current; in P-i-n structures planar doping profiles are adopted, but reflective boundary conditions are adopted to model the edge of the structure. The latter hypothesis is a trigger cause for the filamentation since the normal component of the current on the border of the structure is zero. This condition enhances the strong positive feedback which occurs for the electrons in avalanche break-

down conditions. In this section it is shown that even in the ideal case in which the border are not present in a wide area structure, the current focuses and and it is shown that when the current density is in the range of the NDR even a small variation of the mesh can trigger the filamentation phenomenon. Since the Emitter diffusion (N+) becomes relevant only when the Latch-up phenomenon occurs, the investigated structure is the vertical PNP of the device in Fig. 3.5, without the Trench-Gate and without the Emitter diffusion. The doping profile is the one reported in Fig. 3.6 without the Emitter Diffusion; this structure has been extruded in the x-direction for $600\mu m$. The Mortar Periodic Boundary Condition (MPBC approach)[49] has been adopted for the vertical edges of the structure. In the MPBC approach, both sides of the PBC interface are effectively glued together on one side (the mortar side), while on the other side (the nonmortar side), a weak continuity condition for the equation potential (mostly, the solution variable) is imposed. Therefore, the MPBC approach is not symmetric with respect to the selection of the mortar side if the mesh is nonconforming. The side where the accuracy requirements are larger (and, typically, the mesh is finer) should be selected as the mortar side. The equation potential for the carrier continuity equations are the corresponding quasi-Fermi potentials, while for the other equations, the solution variable is chosen. The MPBC approach does not require a current model across the PBC interface. This condition has the advantage to eliminate the possibility of a current focusing trigger due to geometrical constraints. From the thermal point of view the adiabatic condition is set for all of the borders. The structure has two terminals (the Collector and the Emitter contacts of the IGBT). In Fig. 4.17 the I-V curve of the elementary cell is reported. A strong NDR region is visible and it starts at $J_C = 2 \cdot 10^{-4} A/cm^2$. In Fig. 4.18 the waveforms of the the UIS test (See §4.4.1) performed on the structure are reported. The Collector Current decreases linearly from $11,3mA$ to zero in $4.25\mu s$. It has to be pointed out that in the UIS test the maximum current is achieved by keeping in on-state the device until the current level is reached. In the investigated structure this is not feasible since the Base contact is not present. A first quick ramp is forced to achieve the maximum current into the structure (faster than the electrical phenomena which trigger the filament formation into the structure). Afterwords the current decreases linearly as reported in Fig. 4.18. Differently from a real UIS test, a stored charge at the beginning of the test is not present into the Drift-Layer. Therefore the simulations does not keep into account the Dynamic avalanche effect.

The Collector Voltage exhibits a complex behavior. To understand the re-

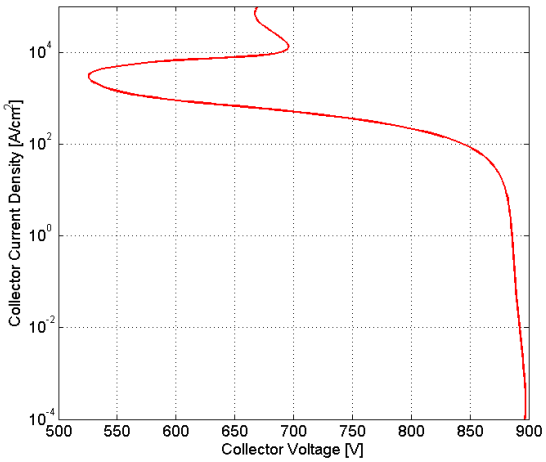


Figure 4.17: I-V curve in avalanche conditions of the elementary cell.

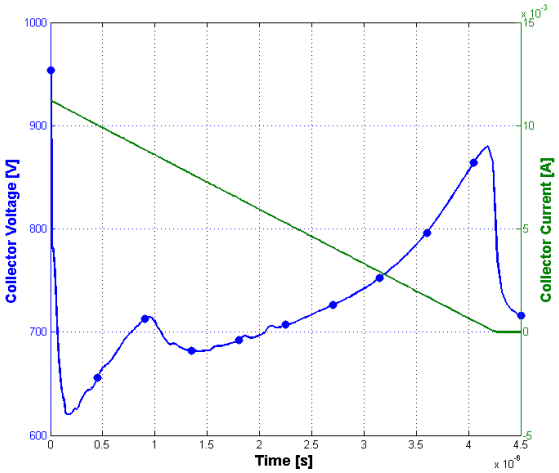


Figure 4.18: Collector Current and Collector Voltage of the structure during the UIS test.

lation between the Collector Voltage behavior and the phenomena which occur in the structure, the electron current density and the temperature distribution in the structure are reported in Fig. 4.19 and Fig. 4.20, respectively, where each frame is relative to the dots on the Collector Voltage waveform in Fig. 4.18.

From the simulations results a complex dynamics occurs when the electrical and thermal phenomena are coupled in a wide area structure. Soon after the beginning of the current decrease, the Collector Voltage abruptly decreases and meanwhile the current filaments. This behavior does not change even if the simulation is performed in isothermal conditions, proving that the current filamentation is intrinsic in structures with NDR regions. As mentioned before, the boundary conditions adopted to simulate the structure eliminate preferential regions for the triggering of the current focusing, therefore a denser mesh is made in a reduced area of the structure. The filamentation occurs where the mesh is modified. By following the temporal sequence, after the initial reduction of the Collector Voltage, the temperature into the filament increases until t_2 and the V_{CE} increases. The core of the filament shows a complex distribution of the electron current, which is organized in vertical regions with higher density. At a certain time the filament start to move toward the right side of the structure and, at the same time, the V_{CE} decreases again until t_3. From t_3 to t_9 the Collector Voltage basically increases and this due to the reduction of the current (because of the NDR) together with the increases of the mean value of the temperature in the structure (it is well known that the breakdown voltage increases with the temperature). It has to be pointed out that from t_3 to t_9 the Collector Voltage does not increases monotonically, but small spikes occurs (See at t_4) in correspondence of the reduction of the vertical region with higher current density. For instance from t_4 to t_5 the regions with high current density reduce from two to one. The filament moves into the direction of the temperature gradient. From t_5 to t_8 the current filament passes from the right side to the left side and this due to MPBC boundary condition which basically glues the two vertical borders. As clarified before this condition is needed to demonstrate that at current focuses even if the ideal condition is achieved. In other simulations where the reflective condition is adopted, the filament moves until reaches the border. In this case the voltage drop decreases because of the restriction of the region in which the current flows. At the same time the temperature increases and after a certain time the it reaches a critical value at which the filament start to move in the opposite direction, with the reduction of the Collector voltage drop.

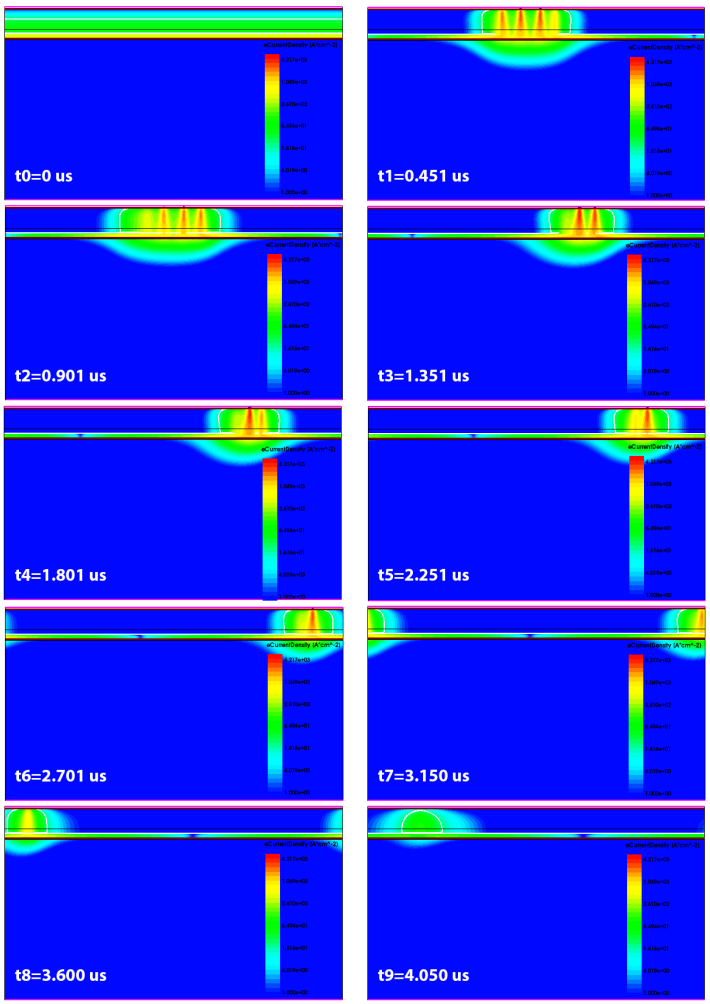


Figure 4.19: Electron current density during the UIS test.

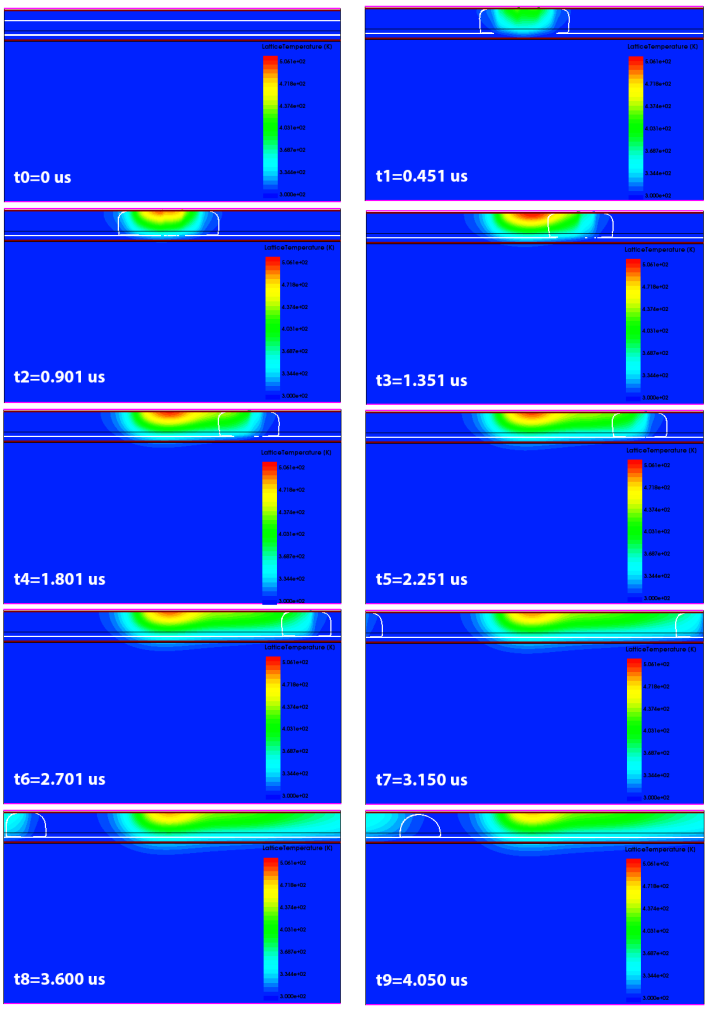


Figure 4.20: Temperature distribution during the UIS test.

4.4 Guidelines for an avalanche rugged IGBT structure

Usually, the avalanche ruggedness of a power device is sacrificed to achieve the best performances in on-state [29]. However, when a device exhibits a strong weakness in avalanche conditions, whether the avalanche phenomenon occurs for a very short time (in the order of some hundreds of nanoseconds) the device can fail. An experimental characterization of the structure presented in the second Chapter has been performed from the thermal and electrical point of view in avalanche conditions to evaluate its robustness in avalanche conditions. The analysis has been carried out for devices (on wafer) with a pitch of $4m$, $5\mu m$ and $6\mu m$ to investigate the current dynamics during the avalanche breakdown. In the following the experimental setups adopted to characterize the device in avalanche conditions are briefly reported. Finally, a critical analysis of the parameters which affect the avalanche ruggedness of the structures in avalanche conditions is reported.

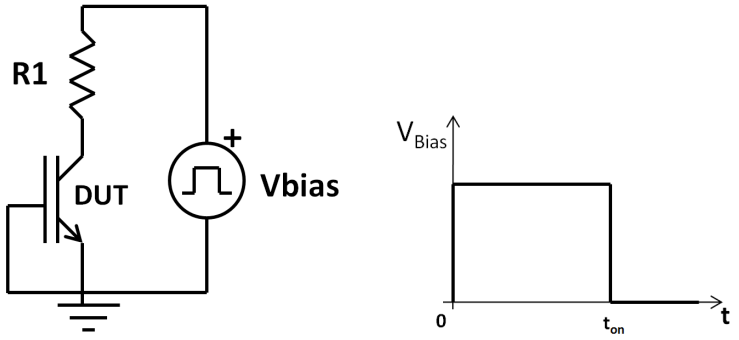
4.4.1 Electro-thermal characterization setup

In this section a brief review of the experimental setups adopted to characterize the behavior of the devices in avalanche conditions are reported.

Electrical characterization setup

Two kind of test have been adopted to evaluate the behavior of the investigated devices in avalanche conditions. The first test circuit is reported in Fig. 4.21. This circuit is used to bring in avalanche conditions the device and it is used in conjunction with the thermographic setup. V_{Bias} is a pulsed generator and it combined to the high resistance value of R_1 to drive in current the device. Once a suitable value of the V_{Bias} in ON is achieved, the test is performed by applying the high voltage to the series of R_1 and the device in blocking state for a lapse of time t_{on}. This circuit can not force an high current into the device because of the high value of the resistance, how ever it forces the avalanche into the device for low currents. The Lock-In thermography is used in conjunction with this test to evaluate the temperature distribution over the surface of the device and visualize a possible focusing of the current in a specific region of the Die.

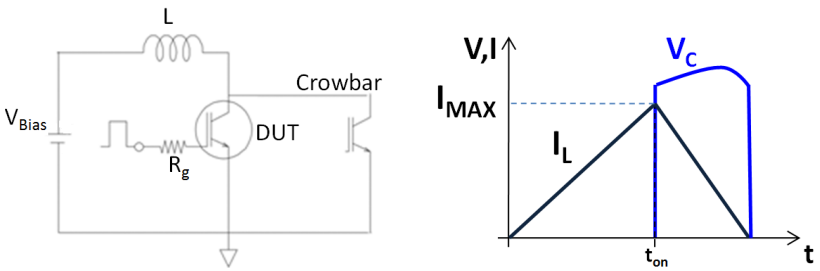
The second circuit test is reported in Fig. 4.22(a) and it is the Unclamped Inductive Switching (UIS) test, one of the most important characterization tool



(a) (b)

Figure 4.21: (a) The schematic of the avalanche test and (b) a sketch of the Bias voltage during the test.

used in the power devices research [112],[113], [114], [115]. The test is usually adopted to define the maximum energy that the device can sustain in avalanche conditions, since a current is forced by the inductor when the device is in blocking state. Generally, device ruggedness is defined as the amount of energy that can be absorbed prior to device failure[116]. The energy limit is correlated to the physics and to the design of the device in a complex way and the electro-thermal phenomena have a crucial role in this dynamic [117]. Therefore, during the test complex phenomena such as an inhomogeneous current distribution on the device area and (if a positive feedback occurs) a subsequent current filamentation [118], [119]. The filamentation phenomenon is an extremely uneven distribution of the current density in the multicellular devices, with a small number of cells carrying most of the load current. This occurs when a power device presents an intrinsic electro-thermal instability, where increased current can lead to increased temperature, and increased temperature in turn leads to exponentially increased current. A crucial role is played by the presence of a NDR (See §4.1) since it is a source of strong positive feedback for the current filamentation. It has been demonstrated by simulations [120] that the current crowding results in a drop of the collector voltage. This phenomenon, in multicellular power devices, can drastically reduce the device performance and its maximum power rating can be much smaller than the sum of the power ratings of the cells in the multi-cellular structure. Therefore, electro-thermal device simulations and experimental measurements are crucial to understand if a semiconductor device is prone to instability and filamentation phenomena



(a) (b)

Figure 4.22: (a) The schematic of the UIS test and (b) a sketch of the Collector voltage and current during the test.

during the turn off under UIS conditions.

As it reported in Fig. 4.22(a), the UIS test circuit consist in a D.U.T. with an inductive load (L), without the free-weeling diode (FWD) (the unclamped name comes from the absence of the FWD). From the operative point of view, the test can be divided in two phases: the charge of the inductor at the desired current and the discharge of the inductor through the device in blocking state. A sketch of the waveforms which occur during the test are reported in Fig. 4.22(b). During the charge phase, the D.U.T. is driven in conduction and the current start to increase. Because of the inductive load, the current increases linearly with the law:

$$\frac{di_C}{dt} = \frac{V_{Bias}}{L} \tag{4.25}$$

When the current reaches the I_{max} value, fixed by the t_{on} duration, the device is turned OFF. The current in an inductor can have a discontinuity for the derivative of the current, but it can not have discontinuities of the first order (a current jump between two values), therefore, when the device turns off the current in the inductor start to force its current into the device in blocking state, otherwise in avalanche conditions. When this happens, the Collector voltage rises up to the breakdown voltage (which is higher than V_{Bias}) and the current into the inductor start to decrease with a slope:

$$\frac{di_C}{dt} = \frac{V_{BR} - V_{Bias}}{L} \tag{4.26}$$

The current flows into the device until all of the stored energy into the inductor is dissipated into the device. A simplified evaluation of the energy

dissipated in the power device during the test (the dissipation in the parasitic components are neglected) can be obtained by:

$$E = \frac{1}{2} L I_{MAX}^2 \frac{V_{BR}}{V_{BR} - V_{CC}} \tag{4.27}$$

It has to be pointed out that the instant in which the dissipated power and the junction temperature are maximum are different. Actually, the maximum dissipation of power occurs soon after t_{on} since the current is maximum and the voltage is maximum as well. On the other hand, from the point of view of the maximum junction temperature, it is dependent on many factors and it occurs in not well established instant of time between t_{on} and the instant in which the current becomes zero. The uncertainty is due to the the balance between the generated heat and the dissipated heat into the structure, which is strongly dependent on the way how the measurement is performed. The inductance value also changes the kind of stress which occurs in the device [121]. Actually, whether a small value of inductance is used, even of the I_{max} is quite high, the time in which the device is in avalanche is low and the temperature increase is low as well. In this condition the electrical stress becomes higher than that due to the thermal phenomenon. On the other hand, when the inductance is higher, even if the I_{max} value is low, the transient is longer and the thermal phenomenon is dominant on the electrical one. Finally, as reported in Fig. 4.22(a), a *Crowbar* is inserted into the circuit to control the Energy in avalanche condition which is dissipated into the D.U.T. [122].

Lock-In Thermography

In this section the Lock-In Thermography (LIT) system based on a thermocamera is introduced. Nowadays LIT is widely used in a variety field of research as a powerful investigation method of electron components such as solar cells, diodes, MOSFETs, ICs etc [123], [124], [125]. Of course thermal measurements with lock-in correlation are very helpful to the characterization of microelectronic devices [126] (i.e. to obtain knowledge about the internal physics mechanism) and also a suitable method for the failure analysis (FA) [127], [128]. In general terms, the Lock-In principle is the best choice if the signals has to be extracted from statistical noise. The only necessary assumption to use this technique is that the signal of interest can be amplitude-modulated. This is because the aim of Lock-In principle is to evaluate only the oscillating part of the detected signal and filter all that falling out of the Lock-In frequency. The key point of this process is that statistical noise is un-

correlated with the recovering signal. Based on this hypothesis, after a process so called spectral inversion (in which a demodulation is performed), we have a dc noise free component easily obtainable by a low pass filtering. With reference to a thermal measurement [129] on an electronic device, suppose that an harmonic electrical power is introduced (e. g. switching on and off the supply voltage at certain Lock-In frequency), the thermal signal on the surface of the device can be seen as follow (harmonics of higher order are omitted):

$$T_{x,y} = T_0 + A_{x,y} cos\left(2\pi f_{LOCK_IN} t + \phi_{x,y}\right) + n_{x,y}\left(t\right) \tag{4.28}$$

In this expression it is present a constant term T_0 that is the equilibrium temperature on the surface, an harmonic term at f_{LOCKIN} and a noise term. In analog lock-in detection, this $T_{x,y}\left(t\right)$ is multiplied with a reference signal of the same frequency by a mixer, and the result is passed through a lowpass filter. Because the phase of the signal is unknown and quantity of interest, quadrature detection is performed, which requires two reference signals that are a quadrature apart (cosine and sine). Since multiplying by sinusoids causes frequency shifts in the spectral components (spectral inversion), the in-quadrature ($S^{90}_{x,y}$) and in-phase ($S^{0}_{x,y}$) component will have the $A_{x,y}$ and dc components shifted in the spectrum. There will be a dc component, which is due to the amplitude and phase of the original sinusoidal signal, a component at the original modulation that has an amplitude of dc, and a component at twice the modulation frequency with an amplitude of one-half $A_{x,y}$. The measured amplitude $A_{x,y}$ and the phase $\phi_{x,y}$ can be calculated from the dc terms of $S^{90}_{x,y}$ and $S^{0}_{x,y}$. In order to obtain the dc terms, one can filter out the unwanted ac terms by convolving $S^{90}_{x,y}$ and $S^{0}_{x,y}$ with a low-pass filter:

$$\begin{cases} S^{0}_{x,y} = \frac{1}{t_{int}} \int_0^{t_{int}} 2T_{x,y}\left(t\right) sin\left(2\pi f_{LOCK-IN} t\right) dt \\ S^{90}_{x,y} = \frac{1}{t_{int}} \int_0^{t_{int}} 2T_{x,y}\left(t\right) cos\left(2\pi f_{LOCK-IN} t\right) dt \end{cases} \tag{4.29}$$

After the ltering process one has the two components commonly referred as the real and imaginary parts of the measured sinusoid. In a digital realization of the Lock-In algorithm, as in the case of thermography based on a thermocamera, an Analog-Digital-Converter (ADC) is used to digitize the input signal leading to a set of samples. Then the whole lock-in correlation procedure is performed numerically as shown in the following formulas:

$$S^{0,90}_{i,j} = \frac{1}{N} \sum_{n=1}^{N} T_{i,j}\left(nt_c\right) K^{0,90}_{n} \tag{4.30}$$

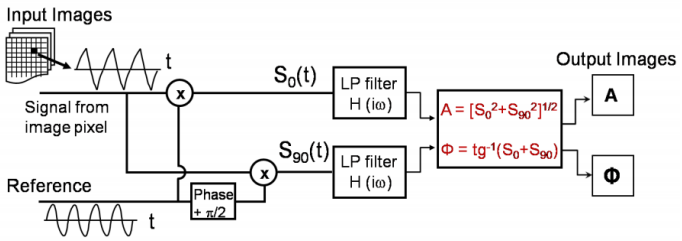


Figure 4.23: Schematic of the Lock-In elaboration for thermal images.

In this case the correlation is realized through a sampled version of sine and cosine function:

$$\begin{cases} K_n^0 = 2sin\left(2\pi f_{LOCK-IN}\left(n-1\right)t_c\right) \\ K_n^{90} = 2cos\left(2\pi f_{LOCK-IN}\left(n-1\right)t_c\right) \end{cases} \tag{4.31}$$

The phase-independent amplitude and the phase can easily be retrieved by:

$$\begin{cases} A_{i,j} = \sqrt{\left(S_{i,j}^0\right)^2 + \left(S_{i,j}^{90}\right)^2} \\ \varphi_{i,j} = arctan\left(\frac{S_{i,j}^{90}}{S_{i,j}^0}\right) \end{cases} \tag{4.32}$$

The whole lock-in elaboration is schematically illustrate in the Fig. 4.23. The schematic of implemented LIT system is depicted in Fig. 4.24.

The core of the experimental apparatus is a Thermo-camera which is synchronized by suitable electronic to the device under test (D.U.T.) and connected to a PC used for data acquisition. All the instruments are controlled via MATLAB©code to acquire the frames sequence, to realize numerical Lock-In elaboration and to produce different kinds of outputs such as amplitude and phase image, temperature proles across sections of the D.U.T. etc.

In terms of noise performance, for a Lock-In process, it is possible to demonstrate that, given a thermo-camera with f_{frame} sample-rate of the thermal images and low-pass lter integration time t_{int}, the average amplitude noise level after a certain acquisition time is:

$$\langle A_{noise} \rangle = \frac{2}{\sqrt{f_{frame} t_{int}}} NETD_{cam} \tag{4.33}$$

where the $NETD$ is the noise equivalent temperature difference of the used thermo-camera. For the realized setup the sensitivity is in the range of

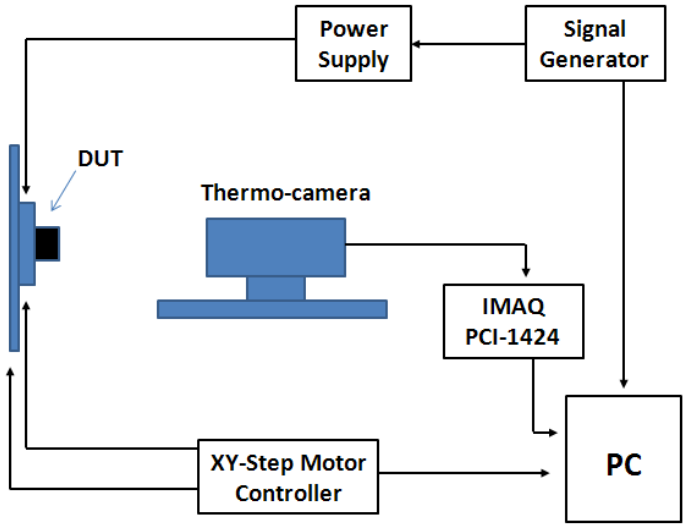


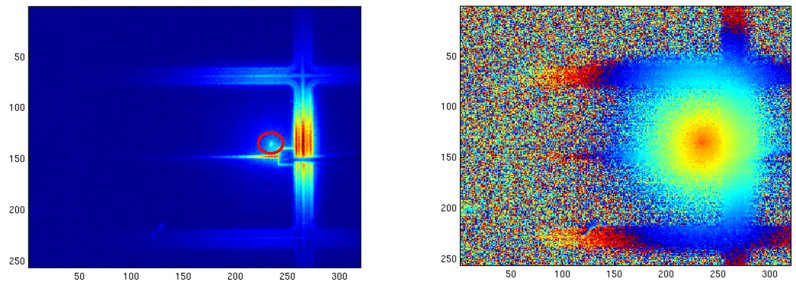
Figure 4.24: Sketch of the Lock-In apparatus.

$100\mu K$ after about 10 min of Lock-In data acquisition.

In Fig. 4.25 the two main output of the LOCK-IN analysis are illustrated:

1. *Amplitude map* (Fig. 4.25(a)) In this map the amplitude of the temperature on the die is reported. It is important to point out that this is an integral information, since the intensity value for each pixel is the integral of the signal in the acquisition period. For example in Fig. 4.25(a) most of the temperature increase is located in a reduced area close to the gate pad (the dot into the red circle). It has to be pointed out that the map is not calibrated to the emissivity of the die; this implies that even if the temperature into the red circle is higher than the red region, the emissivity of the metal (active area) is lower than that of the termination region and an apparent higher temperature is shown in the figure.

2. *Phase map* (Fig. 4.25(b)) In this map it is reported the time shift between the initial area in which the temperature start to rise up and the borders of the area in which the temperature changes during the measurement. This information is crucial for the interpretation of results like in Fig. 4.25(a). As mentioned in the previous point, the amplitude of the temperature (not calibrated in emissivity) into the red circle is lower than that into the red area at the termination region, but the phase



(a) (b)

Figure 4.25: The two main outputs of the Lock-In: (a) amplitude and (b) phase.

map demonstrates that the temperature start to increase into the red circle and during the experiment (during the acquisition time) the heat diffuses laterally.

4.4.2 A critical analysis of the avalanche breakdown in an IGBT

The two tests in avalanche conditions reported above have been performed on the samples. In Fig. 4.26 the thermal map of a device with a pitch of $6\mu m$ is reported when the test in Fig. 4.26 is performed with the current level fixed at $I = 550\mu A$ and $t_{ON} = 8ms$. This test leads to the failure of only the $6\mu m$ pitch devices and the failure modality is the one reported in the Fig. 4.26. The current focus in the active area and, whether the t_{ON} is increased the device breaks because of the high temperature in the hot spot.

UIS is a more stressful test and it is widely adopted from the power semiconductor companies to characterize the reliability of the device in avalanche conditions. Basically it is used to measure the maximum energy which can be dissipated from the device in avalanche conditions before its destruction. Normally it is supposed that the current density is constant all over the area of the device, but this is not true when the device exhibits a NDR (See §4.1). This effect is clearly visible in Fig.4.27 where both the electrical waveforms and the thermal map are reported for a device with pitch = $5\mu m$. The result is representative of the main failure modality for all of the three families of devices. From the electrical point of view, as soon as the device enters in avalanche condition (the Collector Voltage is clamped at a constant value) after some hundreds of nanoseconds (300ns in the specific case) the Collector volt-

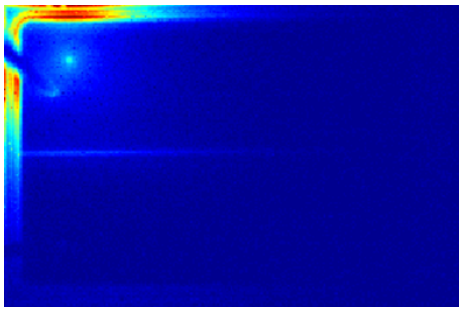
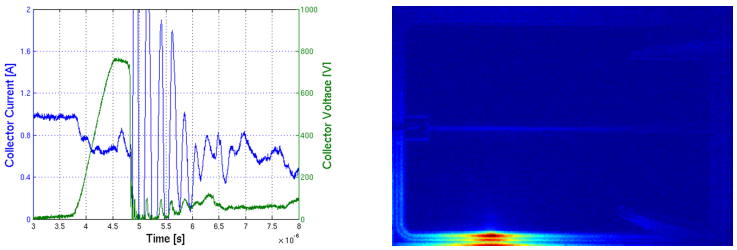


Figure 4.26: Temperature map of a device with a pitch of $6\mu m$ during the first typology of avalanche test.

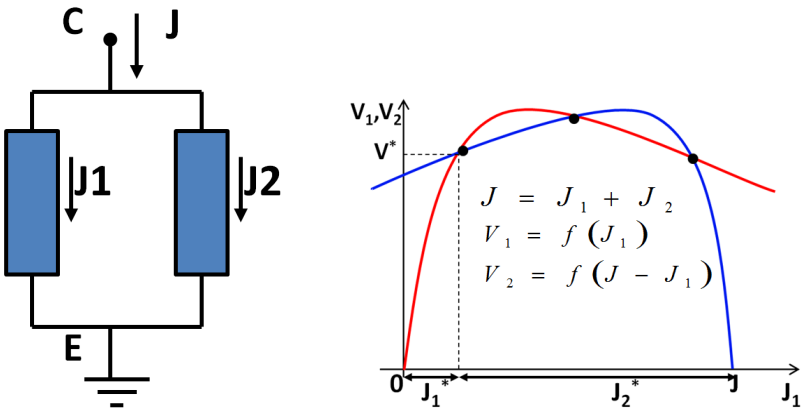


(a) (b)

Figure 4.27: Failure during the UIS test ($L = 14mH$ and $I_{max} = 1A$). (a) Electrical waveforms and (b) temperature map.

age collapses and the device fails. From the thermal point of view it clearly visible a strong focusing of the current in the termination region of the device. Usually the termination region is engineered in a way that the current flows along all perimeter of the device when the avalanche phenomenon occurs, but in this case it well visible the current filamentation.

The two illustrated failure modalities are mainly due to the strong NDR that both the active area and the termination region exhibits. The active area and the termination region are very different from the point of view of the design as well as from the point of view of the electrical behavior (See Fig. 4.11 and 4.12). The termination area and the active area region can be schematized like two dipoles in parallel where the current is forced and the overall voltage drop is dependent on the area where the current flows and its evolution in the time follows the complex dynamics shown in the previous section. From the



(a) (b)

Figure 4.28: (a) Schematization of the active area and the termination. (b) Graphical method to evaluate the voltage drop on the two dipoles.

statical point of view, the region in which the current start to flow can be evaluated by schematizing the active region and the termination like two dipoles (See Fig. 4.28(a)) and graphically finding the solution (See Fig. 4.28(b)).

The graph in Fig. 4.28(b) is realized in the following way: it supposed that the I-V curve in avalanche conditions of both the active area and the termination can be traced (the aim of the graph is to evaluate the initial division of the current between the two branches and an initial uniformity of the current is supposed. After a certain transient the current filaments whether the NDR is present into the I-V curve) and it can be written that $V_1 = f_1(J_1)$ and $V_2 = f_2(J_2)$. Since the two dipoles are in parallel configuration, $J = J_1 + J_2$ and it can be written that $V_2 = f_2(J_2) = f_2(J - J_1)$, where J is the total current forced in the system. To graphically found the solution the V-I curve of the first dipole has to be reported together with the one of the second dipole translated in the positive direction by J and rotated by 180 degrees. The intersection of the two curves are the solutions of the problem since the sum of the teo currents is always equal to the total current J. When the I-V curves have not NDR regions, only one solution is possible. However, when the characteristics have one or more NDR regions the solutions of the problem become multiple and a strong instability occurs in the system. The illustrated graphical technique can be applied to the structures analyzed in the previous sections. In Fig. 4.29 the I-V curves in reverse conditions are reported for the termination

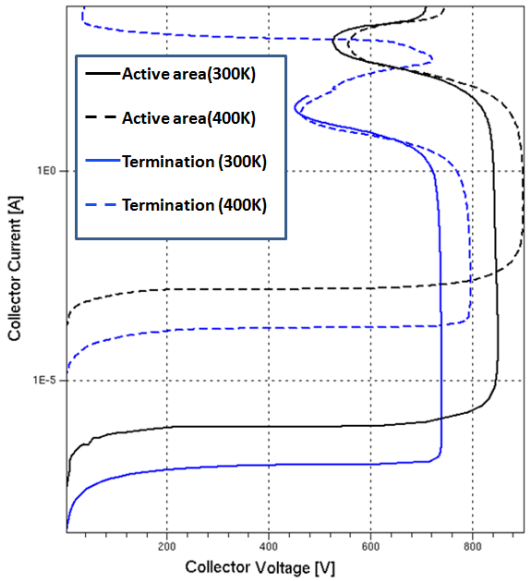


Figure 4.29: Reverse I-V curves of the active region and the termination region at $T = 300K$ and $T = 400K$.

and the active area region, scaled on the area of the two regions, at two temperatures. By adopting the graphical analysis shown above, it clear that the current flows only into the termination region.

The curves in Fig. 4.29 show a strong NDR region therefore the dynamic of the avalanche failure con be explained by observing that initially the current start to flow into the termination area because of the lower breakdown voltage and because of the absence of the intersections between the two curves. When the current start to flow into the termination area, which has a strong NDR region in its I-V curve, the filamentation phenomenon occurs because of the dynamics shown in §4.3.4. The critical temperature at which the filament start to move is strongly dependent on the features of the structure under test. From the experimental measurements it clear that filament does not move from its initial position and the high increase of the temperature leads to the destruction of the area and the failure of the device. On the other hand, the failure modality reported in Fig. 4.26 can be explained by combining the results of this section to those of §4.3.3. The devices with the pitch value of $6\mu m$ have shown a bad metal coverage and a reduced reliability of the emitter contact. In this devices

the effect of the premature Latch-up is dominant and focus of the current in active area occurs. It has to be pointed out that the Latch-up phenomenon is enhanced by the strong NDR of the active area. When the current starts to flow in a reduced area of the device for the phenomenon shown in §4.3.3, the positive feedback due to the NDR region speeds up the current increase in the failed region. Therefore the main action to get an overall increased robustness of the device in avalanche conditions is the reduction of the effect of the NDR. This goal can be achieved by increasing the current level at which the I-V curve passes from the vertical region to the NDR region, by increasing the slope in the NDR and by reducing the current level at which the transition from the NDR to the PDR occurs. From the simulation results and the experimental evidences it can be concluded that the ruggedness in avalanche conditions is achieved when:

1. The lifetime into the structure is reduced. The Buffer-Layer width is increased.

2. The Buffer-Layer doping concentration increases.

As mentioned before all of these measures are effective in increasing the reliability in avalanche conditions, but they also affect the on-state performances of the device and an accurate evaluation of the benefits have to be taken into account.

Conclusions

This dissertation has presented the analysis of the aspects regarding the ruggedness in forward and reverse conditions in IGBTs, the most used power semiconductor devices used in the applications of Middle/High power range. This was achieved by means of modeling, simulation by TCAD and experimental analysis. Firstly the IGBT structure and its operation mode has been introduced together a brief history of the development of its design up to now. The basic compact modeling of the IGBT is presented as well for the DC characteristics and a detailed description of the dynamics of the turn-OFF and turn-ON are illustrated. The Latch-up phenomenon is illustrated as well as the SOA concept in forward, reverse and short-circuit conditions. An overview of the Sentaurus TCAD simulator is reported together with the main information on the flow of simulation. The TCAD simulator has been the main tool adopted in this work to investigate the behavior of the IGBT structure in harsh conditions. After the definition of the simulation strategy for multicellular structures (as for modern IGBTs), a new calibration procedure has been presented for a Trench-IGBT structure and it has been automatized by means of the MATLAB©function *fminsearch* to minimize the error between the experimental terminal characteristics and the simulated ones. The results show that the calibration procedure is effective in predict the behavior of the device also from a dynamical point of view since a good match is achieved for the electrical waveforms for both the turn-OFF test and the short-circuit test. The calibrated structure has been the starting point to analyze the effect of the design parameters on the short-circuit capability of a PT-IGBT structure. The cell-pitch, the channel length, the contact-trench and the Emitter diffusion modulation have been parametrically analyzed to find the trade-off among the voltage drop in on-state, the turn-OFF loss and the short-circuit capability. It has been pointed out that an Emitter modulation with a N+ design orthogonal to the Trench-Gate is effective in reduce the saturation current, but the Latch-up robustness is reduced when a large area of the N+ diffusion is

adopted. An optimized design has been proposed to overcome both the Latch-up and short-circuit capability problem. this has been achieved by modulating the N+ diffusion with a percentage of 50% and by using shallow contact trench with a depth of $0.4\mu m$. the design has been made and the new devices have been experimentally characterized. The experimental data confirm the simulation results. The reliability of an IGBT is also related to its robustness in avalanche conditions. Actually, the analysis of the failures in avalanche conditions are becoming an important research field since the large area power devices emphasizes the current focusing phenomenon. Therefore, the avalanche phenomenon has been investigated from the experimental, analytical and simulative point of view. A physical analytical model has been developed to trace the I-V curve in avalanche conditions of the vertical PNP of the IGBT, starting from the geometrical features of the structure. The model is based on the calculation of the flux of electron and holes in the depletion region by means of a charge control model for the Beta of the BJT, at the different current levels. The limits of validity of the model are shown as well. From the simulative point of view, TCAD simulations of both the active area and the termination of structure analyzed in the third Chapter have been performed. The buffer design and the lifetime in the structure have been defined as the main design parameters which affect the ruggedness in avalanche conditions. TCAD simulation in mixed mode have been performed to simulate the UIS test for a large area device with the PNP structure of the IGBT (the N+ diffusion of the Emitter is neglected since its activation implies the Latch-up of the device, therefore the failure). The correlation between the voltage drop on the structure and the carriers and temperature distribution are reported. A graphical tool to analyze the interplay between termination and active region in avalanche region has been developed. It is a formally and mathematically correct way to solve graphically the voltage drop on the parallel of two dipoles when a total current is forced. Finally, it is shown that the avalanche ruggedness can be improved by reducing the lifetime in the structure or by increasing the doping concentration of the Buffer layer or by increasing the thickness of the Buffer layer. The effect of these modification of the structure have been evaluated from the point of view of the on-state performances.

Appendix A

FFR technique applied to Silicon Photomultipliers

In many fields of science it is fundamental to detect very weak optical signals [130],[131],[132]. Photomultiplier tubes (PMTs) have been the first devices suitable for the detection of low intensity optical signals. However PMTs have two main drawbacks: they are very sensitive to magnetic fields and they have an high cost [133]. Silicon photomultipliers (SiPms) are interesting candidates to replace PMTs since they are not affected by the above limitations and, in addition, they have the advantages of the solid-state technology. SiPms are basically composed by a big number of Avalanche Photodiodes operating in Geiger mode (GM-APDs), connected in parallel (See Fig. A.1). Each GM-APD cell is a p-n asymmetrical junction (D) reversely biased at a voltage (VBIAS) higher than the breakdown voltage (VBR), through a high ballast resistor (RL). The overvoltage $V_O = (V_{BIAS} V_{BR})$ strongly affects the SiPm performance. A GM-APD cell can detect a single photo-generated electron-hole pair and its principle of operation can be explained by studying the temporal evolution of the phenomenon. Let us suppose that the p-n junction is not biased. If the junction is suddenly reversely biased at voltage above VBR, the impact ionization does not occur until a carrier is injected in the depletion region. The electric field at the junction is very high and when a carrier is injected in the depletion region, there is a certain probability that an avalanche is triggered because of the impact ionization phenomenon. If the avalanche occurs, the depletion capacitance is rapidly discharged and a sharp current pulse, electronically measurable, is generated [134]. Since the avalanche is self-sustaining, it is necessary a quenching circuit which quenches

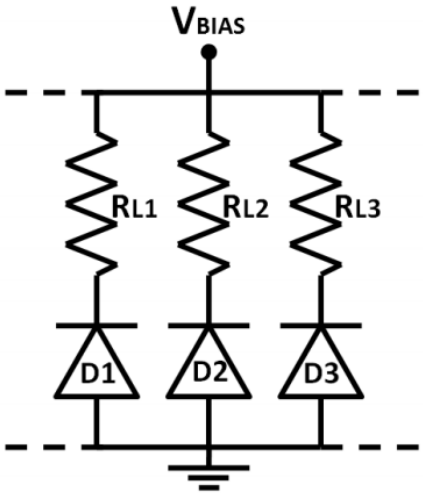


Figure A.1: Equivalent circuit of a SiPm.

the avalanche by lowering the bias voltage below the VBR and, then, it must restore the bias voltage to detect another incident photon [135]. Each GM-APD has the same current response when it is activated by incident photons, therefore the output current of SiPm is proportional to the number of cells activated at that moment. Photon detection efficiency (PDE) is the probability to detect single photon incoming on the photodetector surface. PDE depends on the structure of the single cell of the SiPm; in particular it depends on geometrical fill factor, the triggering probability and the quantum efficiency (QE). Existing SiPms show a low PDE at short wavelengths, especially in the blue range because of the photon absorption which occurs mainly at the surface where the recombination phenomenon is predominant. Floating field ring technique is widely used to reduce the premature breakdown in n+/p APD structure [136].

A.1 Photon detection efficiency

For a SiPm, PDE can be evaluated from the following:

$$PDE = QE \cdot P_{Trigger} \cdot \epsilon \tag{A.1}$$

Where QE is the quantum efficiency, ϵ is the geometrical fill factor and $P_{Trigger}$ is the probability that a photo-generated carrier triggers an avalanche.

The QE is the probability that an incoming photon generates an electron-hole pair into the active volume of the SiPm. This parameter can be expressed like the product of three terms: the probability that an incident photon is not reflected, the probability that a photo-generated electron-hole pair does not recombine on the device surface and the probability that an electron-hole pair is photo-generated into the active volume of the SiPm. QE strongly depends on the wavelength and it assumes low values in the blue range, since the photo-generation occurs mainly close to the surface where recombination phenomenon is predominant [137]. $P_{Trigger}$ depends on the position where the electron-hole pair is photo-generated inside the depleted region and it also depends on the overvoltage (V_O) [138]. is a critical parameter in SiPm since it represents the ratio between the active area and the total area. Each single cell has an inactive area due to the presence of: the guard ring structures, the structures for the suppression of the optical crosstalk and the quenching resistor. The latter is integrated on each cell and it is made in polysilicon. Even if the above structures have a size in the order of some micrometers, they limit the geometrical fill factor, especially if the single cells has dimension in the order of $30x30\mu m^2$ [139]. An increasing of the geometrical fill factor could be achieved by increasing the size of the single cell. However the dark count rate limits the maximum area of a SiPm to few mm^2, therefore the size of the single cell depends on the application. The number of incident photon in Cherenkov telescope is very low, it is required an high PDE and the single cell have to be as big as possible. In positron emission tomography (PET) the number of incident photons on the surface of the SiPm is in the order of the 10^3 at the same time [133]. Since each cell cannot detect further photon during the avalanche pulse, the latter application requires an high number of small cells to avoid the saturation effect.

A.2 APD structure

The first APD prototype appeared in the early 1960s, during the pioneering studies of the Micro-plasma phenomenon [134], [140]. A p/n junction, surrounded by a guard ring, was the first structure investigated (Fig. A.2). The guard ring structure was introduced in order to prevent the premature breakdown at the main junction termination and to confine the avalanche region in the central area of the device [141]. The above structure has a very low detection efficiency mainly related to the presence of the guard ring structure. The structure in Fig. A.2, exposed to a picoseconds laser pulse, has a time

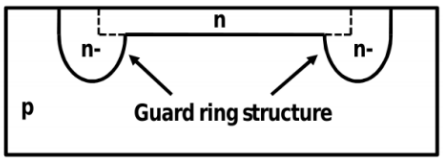


Figure A.2: APD structure used to investigate the micro-plasma phenomenon. .

response which exhibits a long diffusion tail that cannot be fitted by a simple exponential function. Furthermore the shape of the diffusion tail is wavelength dependent [142].

The optical efficiency of the APD was improved by the reach-through APD (RAPD) structure [143]. Starting from the top, the structure is n+/p/π/p+ (Fig. A.3). Since the doping concentration of the p-layer (at the junction) is typically 2-3 order of magnitude higher than of the π-layer (low doped p-layer), the breakdown voltage of the n+/π junction is higher than the one of the n+/p junction. The above condition confines the avalanche region in a thin layer at n+/p junction. In operative conditions the depleted layer reaches the substrate (p+) through the p and π regions. The low electric field in the π region drifts the photo-generated electrons to the n+/p junction, where they can trigger an avalanche because of the high electric field. Therefore in the RAPD structure the absorption region is separated from the multiplication region. The n+/p/π/p+ structure exhibits a low gain fluctuation since the photo-generated electrons pass through the same multiplication region, regardless of the position in the depletion layer where they are photo-generated. The absorption of the short wavelength photons ($<$ 400 nm) occurs in a thin region close to the surface. Whereupon the n+ layer has to be as shallow as possible.

A.3 Floating Field Ring technique

The addition of a floating field ring (FFR) to the planar junction structure mitigates the rising of the electric field due to the curvature of the main junction border. If the distance between the main junction and the FFR, labelled WS in Fig. A.4, is optimized, the FFR is effective in supporting a portion of the applied voltage. By solving the Poissons equation in cylindrical coordinates, it can be obtained an analytical expression for the difference in voltage between the main junction and the FFR [14]:

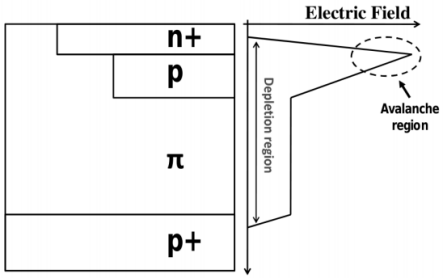


Figure A.3: RAPD structure and electric field distribution in the active area. Photon-electrons generated into the p and π regions are driven to the high field region where they can trigger an avalanche.

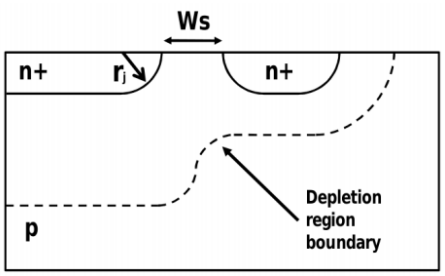


Figure A.4: Impact of floating field ring on the depletion region boundary of a planar junction.

$$V_M - V_{FFR} = \frac{qN_D}{2\epsilon_S}\left[\left(\frac{r_j^2 - W_S^2}{2}\right) + r_D^2 ln\left(\frac{W_D}{r_j}\right)\right] \qquad \text{(A.2)}$$

Where V_M is the main junction voltage and V_{FFR} is the FFR voltage.

Since the potential along the FFR is constant, the potential gradient at the border of the junction is reduced. Therefore, the depletion region is laterally extended and the electric field is mitigated at the junction termination. This technique does not need a further technological step since it can be implemented by modifying the mask design of the main junction. The superficial distribution of the electric field, starting from the main junction edge, is plotted in Fig. A.5. Simulations also show that by fixing the applied voltage of the main junction, the electric field peaks in the FFR structure are lower than the

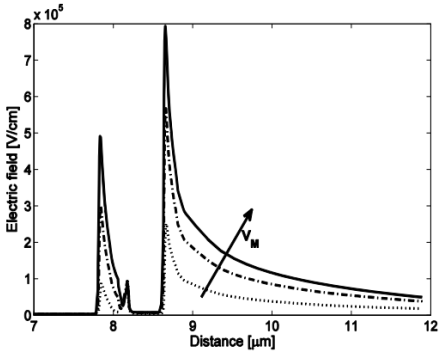


Figure A.5: Superficial electric field profile in cylindrical junction with single floating field ring. The electric field rises by increasing V_M.

one without the FFR structure.

A.4 FFR technique applied to the RAPD structure

The floating field ring technique has been applied to a conventional RAPD structure. In this structure the avalanche phenomenon occurs at the n+/p junction and the remaining n+/π abrupt junction is used for the shaping of the electric field on the edge of the elementary cell. The n+ well overlap spacing affects the overall fill factor of the SiPm since it is one of the contributes to the inactive area. The floating field ring technique has been applied to the n+/π termination to reduce the electric field out of the cell, keeping constant the n-well size. The single cell structure has been defined by using the criteria for a conventional RAPD structure for the blue photons detection [139]. In Fig. A.6 a sketch of the 2D structure simulated in Sentaurus TCAD [49].

Different models have been used to correctly simulate the device. The Shockley-Read-Hall has been included for the thermal bulk and surface recombination. It has been also included the Auger recombination. Since APDs are based on the impact ionization effect, it has been included the University of Bologna impact ionization model. The n+ has a peak doping concentration of $10^{19}cm^{-3}$ and its thickness is $100nm$. The p-layer has a peak doping concentration of $10^{17}cm^{-3}$ with a depth of $900nm$. The π-layer doping concentration is in the order of few $10^{14}cm^{-3}$. The n+ well is $2\mu m$ larger than

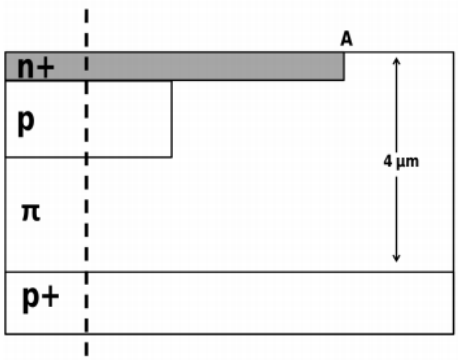


Figure A.6: A sketch of the border of a conventional RAPD structure.

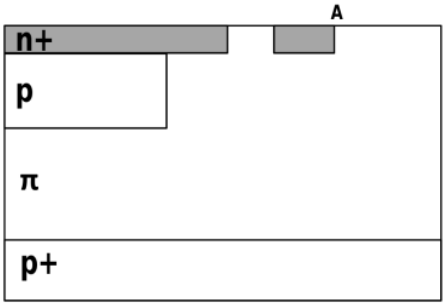


Figure A.7: Floating field ring applied to the conventional RAPD structure.

the p-layer. By simulations of the above structure in operative conditions, the electric field along the dashed line in Fig. A.6 shows a strong peak close to the n+/p junction, rapidly decreases in p-layer and finally reaches a low constant value in the π-region, confirming its correct behavior. The simulated breakdown voltage is 23 V. In Fig. A.7 a schematic view of the floating field ring applied to the conventional cell.

The distance between the external border of the field ring A and the p-layer is equal to the n+ overlapping space of the structure in Fig. A.6. The minimum FFR width has been found equal to 450 nm. The effect of the FFR has been simulated by changing the distance between the main n+/π junction and the FFR. The optimum distance has been found to be 250 nm. By using the latter geometrical values, the maximum reduction of the electric field, starting from the external border of the FFR, has been obtained. The modified structure shows a breakdown voltage of 23 V, meaning that the FFR only affects the

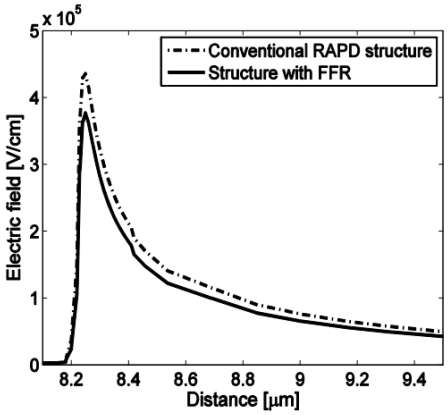


Figure A.8: Superficial electric field profile starting from the border (A) of the two investigated structures.

electric field shape. In Fig. A.8 it is shown the superficial electric field profile toward the external part of the cell, starting from the A point of both the structures. The FFR structure exhibits an electric field which is 10% lower than the conventional structure. Once the minimum n+ overlapping size is fixed to achieve a suitable shaping of the electric field on the border of the cell, it can be applied the FFR technique, which reduces the electric field out of the cell by a factor of 10% keeping constant the area of the single cell. In absence of structure for the optical crosstalk suppression, the minimum distance among the cells in the SiPm is limited by the lateral electric field overlapping. Therefore by fixing the maximum electric field acceptable in the middle of the region between a cell and its neighbour, the above geometry allows a lower distance, improving, then, the geometrical fill factor of the SiPm.

Bibliography

[1] N. S.-. i. t. d. b. Yamagami's patent can be searched by inputting "Kind code" B.

[2] B. Jayant Baliga, "Enhancement- and depletion-mode vertical-channel m.o.s. gated thyristors," *Electronics Letters*, vol. 15, no. 20, pp. 645–647, 1979.

[3] B. Baliga, M. Adler, P. Gray, R. love, and N. Zommer, "The insulated gate rectifier (igr): A new power switching device," in *Electron Devices Meeting, 1982 International*, vol. 1982, pp. 264–267.

[4] B. Baliga, "Fast-switching insulated gate transistors," *Electron Device Letters, IEEE*, vol. 4, no. 12, pp. 452–454, Dec 1983.

[5] J. P. Russell, A. Goodman, L. Goodman, and J. M. Neilson, "The comfet 8212;a new high conductance mos-gated device," *Electron Device Letters, IEEE*, vol. 4, no. 3, pp. 63–65, Mar 1983.

[6] J. D. Plummer, "Monolithic semiconductor switching device," Apr. 22 1980, uS Patent 4,199,774.

[7] ——, "Monolithic semiconductor switching device," May 1 1990, uS Patent RE33,209.

[8] H. W. Becke and C. F. Wheatley Jr, "Power mosfet with an anode region," Dec. 14 1982, uS Patent 4,364,073.

[9] A. Goodman, J. P. Russell, L. Goodman, C. Nuese, and J. Neilson, "Improved comfets with fast switching speed and high-current capability," in *Electron Devices Meeting, 1983 International*, vol. 29, pp. 79–82.

[10] B. Baliga, "Temperature behavior of insulated gate transistor characteristics," *Solid-State Electronics*, vol. 28, no. 3, pp. 289 – 297, 1985.

[11] A. Nakagawa, H. Ohashi, M. Kurata, H. Yamaguchi, and K. Watanabe, "Non-latch-up 1200v 75a bipolar-mode mosfet with large aso," in *Electron Devices Meeting, 1984 International*, vol. 30, pp. 860–861.

[12] A. Nakagawa, Y. Yamaguchi, K. Watanabe, H. Ohashi, and M. Kurata, "Experimental and numerical study of non-latch-up bipolar-mode mosfet characteristics," in *Electron Devices Meeting, 1985 International*, vol. 31, 1985, pp. 150–153.

[13] V. K. Khanna, *Insulated Gate Bipolar Transistor IGBT Theory and Design*. Wiley-IEEE Press, 2003.

[14] B. Baliga, *Fundamentals of Power Semiconductor Devices*. 1st ed. New York: Springer-Verlag, 2008.

[15] S. M. Sze and K. K. Ng, *Physics of semiconductor devices*. Wiley-interscience, 2006.

[16] M. H. Rashid, *Power electronics handbook*. Academic Pr, 2001.

[17] A. Bryant, S. Yang, P. Mawby, D. Xiang, L. Ran, P. Tavner, and P. Palmer, "Investigation into igbt dv/dt during turn-off and its temperature dependence," *Power Electronics, IEEE Transactions on*, vol. 26, no. 10, pp. 3019–3031, Oct.

[18] J. Oetjen, R. Jungblut, U. Kuhlmann, J. Arkenau, and R. Sittig, "Current filamentation in bipolar power devices during dynamic avalanche breakdown," *Solid-State Electronics*, vol. 44, no. 1, pp. 117 – 123, 2000.

[19] X. Perpina, J.-F. Serviere, J. Urresti-Ibanez, I. Cortes, X. Jorda, S. Hidalgo, J. Rebollo, and M. Mermet-Guyennet, "Analysis of clamped inductive turnoff failure in railway traction igbt power modules under overload conditions," *Industrial Electronics, IEEE Transactions on*, vol. 58, no. 7, pp. 2706–2714, July.

[20] T. Raker, H. P. Felsl, F. Niedernostheide, F. Pfirsch, and H. Schulze, "Limits of strongly punch-through designed igbts," in *Power Semiconductor Devices and ICs (ISPSD), 2011 IEEE 23rd International Symposium on*, May, pp. 100–103.

[21] R. Baburske, F. Niedernostheide, E. Falck, J. Lutz, H. Schulze, and J. Bauer, "Destruction behavior of power diodes beyond the soa limit,"

in *Power Semiconductor Devices and ICs (ISPSD), 2012 24th International Symposium on*, June 2012, pp. 365–368.

[22] T. Ogura, H. Ninomiya, K. Sugiyama, and T. Inoue, "Turn-off switching analysis considering dynamic avalanche effect for low turn-off loss high-voltage igbts," *Electron Devices, IEEE Transactions on*, vol. 51, no. 4, pp. 629–635, April.

[23] H. P. Felsl, E. Falck, F. J. Niedernostheide, S. Milady, D. Silber, and J. Lutz, "Electro-thermal simulation of current filamentation in 3.3-kv silicon p+- n– n+diodes with differenth edge terminations," in *Power Semiconductor Devices and IC's, 2006. ISPSD 2006. IEEE International Symposium on*, June, pp. 1–4.

[24] R. Baburske, B. Heinze, F. Niedernostheide, J. Lutz, and D. Silber, "On the formation of stationary destructive cathode-side filaments in p+-n x2212;-n+ diodes," in *Power Semiconductor Devices IC's, 2009. ISPSD 2009. 21st International Symposium on*, June, pp. 41–44.

[25] H.-J. Schulze, F.-J. Niedernostheide, F. Pfirsch, and R. Baburske, "Limiting factors of the safe operating area for power devices," *Electron Devices, IEEE Transactions on*, vol. 60, no. 2, pp. 551–562, Feb.

[26] J. Lutz and R. Baburske, "Dynamic avalanche in bipolar power devices," *Microelectronics Reliability*, vol. 52, no. 3, pp. 475 – 481, 2012, ¡ce:title¿Special section on International Seminar on Power Semiconductors 2010¡/ce:title¿.

[27] K.-H. Oh, Y.-C. Kim, K.-H. Lee, and C.-M. Yun, "Investigation of short-circuit failure limited by dynamic-avalanche capability in 600-v punchthrough igbts," *Device and Materials Reliability, IEEE Transactions on*, vol. 6, no. 1, pp. 2–8, March.

[28] S. Lefebvre and F. Miserey, "Analysis of cic npt igbt's turn-off operations for high switching current level," *Electron Devices, IEEE Transactions on*, vol. 46, no. 5, pp. 1042–1049, May 1999.

[29] M. Otsuki, Y. Onozawa, S. Yoshiwatari, and Y. Seki, "1200v fs-igbt module with enhanced dynamic clamping capability," in *Power Semiconductor Devices and ICs, 2004. Proceedings. ISPSD '04. The 16th International Symposium on*, May, pp. 339–342.

[30] M. Rahimo, U. Schlapbach, A. Kopta, and S. Linder, "An assessment of modern igbt and anti-parallel diode behaviour in hard-switching applications," in *Power Electronics and Applications, 2005 European Conference on*, 2005, pp. 10 pp.–P.10.

[31] X. Perpina, I. Cortes, J. Urresti-Ibanez, X. Jorda, J. Rebollo, and J. Millan, "Edge termination impact on clamped inductive turn-off failure in high-voltage igbts under overcurrent conditions," in *Power Semiconductor Devices and ICs (ISPSD), 2011 IEEE 23rd International Symposium on*, May 2011, pp. 112–115.

[32] I. Corts, X. Perpi, J. Urresti, X. Jord, and J. Rebollo, "Study of layout influence on ruggedness of npt-igbt devices by physical modelling," *Microelectronics Reliability*, vol. 52, no. 910, pp. 2471 – 2476, 2012.

[33] N. Lophitis, M. Antoniou, F. Udrea, I. Nistor, M. Arnold, T. Wikstrom, and J. Vobecky, "Experimentally validated three dimensional gct wafer level simulations," in *Power Semiconductor Devices and ICs (ISPSD), 2012 24th International Symposium on*, June, pp. 349–352.

[34] B. Baliga, M. Adler, P. Gray, and R. love, "Suppressing latchup in insulated gate transistors," *Electron Device Letters, IEEE*, vol. 5, no. 8, pp. 323–325, Aug 1984.

[35] S. Eranen and M. Blomberg, "The vertical igbt with an implanted buried layer," in *Power Semiconductor Devices and ICs, 1991. ISPSD '91., Proceedings of the 3rd International Symposium on*, Apr 1991, pp. 211–214.

[36] T. Chow, B. Baliga, and D. Pattanayak, "Counterdoping of mos channel (cdc)-a new technique of improving suppression of latching in insulated gate bipolar transistors," *Electron Device Letters, IEEE*, vol. 9, no. 1, pp. 29–31, Jan.1988.

[37] T. Laska, F. Pfirsch, F. Hirler, J. Niedermeyr, C. Schaffer, and T. Schmidt, "1200 v-trench-igbt study with square short circuit soa," in *Power Semiconductor Devices and ICs, 1998. ISPSD 98. Proceedings of the 10th International Symposium on*, Jun 1998, pp. 433–436.

[38] S. L. Miller, "Avalanche breakdown in germanium," *Phys. Rev.*, vol. 99, pp. 1234–1241, Aug 1955.

[39] H. Hagino, J. Yamashita, A. Uenishi, and H. Haruguchi, "An experimental and numerical study on the forward biased soa of igbts," *Electron Devices, IEEE Transactions on*, vol. 43, no. 3, pp. 490–500, Mar 1996.

[40] H.-R. Chang, J. Bu, H. Hauenstein, M. Wittmann, J. Marcinkowski, M. Pavier, S. Palmer, and J. Tompkins, "200 kva compact igbt modules with double-sided cooling for hev and ev," in *Power Semiconductor Devices and ICs (ISPSD), 2012 24th International Symposium on*, June 2012, pp. 299–302.

[41] G. Majumdar, J. Yamashita, H. Nishihara, Y. Tomomatsu, N. Soejima, M. Tabata, and H. Hagino, "A new generation high speed low loss igbt module," in *Power Semiconductor Devices and ICs, 1992. ISPSD '92. Proceedings of the 4th International Symposium on*, pp. 168–171.

[42] E. R. Motto and J. F. Donlon, "The latest advances in industrial igbt module technology," in *Applied Power Electronics Conference and Exposition, 2004. APEC'04. Nineteenth Annual IEEE*, vol. 1. IEEE, 2004, pp. 235–240.

[43] T. Takahashi, Y. Tomomatsu, and K. Sato, "Cstbt (iii) as the next generation igbt," in *Power Semiconductor Devices and IC's, 2008. ISPSD '08. 20th International Symposium on*, May, pp. 72–75.

[44] M. Sumitomo, J. Asai, H. Sakane, K. Arakawa, Y. Higuchi, and M. Matsui, "Low loss igbt with partially narrow mesa structure (pnm-igbt)," in *Power Semiconductor Devices and ICs (ISPSD), 2012 24th International Symposium on*, June 2012, pp. 17–20.

[45] M. Pinto, C. Rafferty, R. Smith, and J. Bude, "Ulsi technology development by predictive simulations," in *Electron Devices Meeting, 1993. IEDM '93. Technical Digest., International*, 1993, pp. 701–704.

[46] H. Masuda, H. Sato, K. Tsuneno, K. Aoyama, T. Nakamura, H. Kunitomo, and K. Kajigaya, "Tcad strategy for predictive vlsi memory development," in *Electron Devices Meeting, 1994. IEDM '94. Technical Digest., International*, 1994, pp. 153–156.

[47] G. Breglio, A. Irace, E. Napoli, P. Spirito, K. Hamada, T. Nishijima, and T. Ueta, "Study of a failure mechanism during uis switching of planar

pt-igbt with current sense cell," *Microelectronics Reliability*, vol. 47, no. 911, pp. 1756 – 1760, 2007.

[48] A. Irace, G. Breglio, P. Spirito, A. Bricconi, D. Raffo, and L. Merlin, "Effect of a buffer layer in the epi-substrate region to boost the avalanche capability of a 100v schottky diode," *Microelectronics Reliability*, vol. 46, pp. 1784 – 1789, 2006.

[49] T. Sentaurus, "Synopsys," *Inc., F-2011.09 edition*, 2011.

[50] R. Minixhofer, "Tcad as an integral part of the semiconductor manufacturing environment," in *Simulation of Semiconductor Processes and Devices, 2006 International Conference on*, 2006, pp. 9–16.

[51] T. Silvaco, "Manual, atlas," *USA: Silvaco Int'l Co*.

[52] G. Wachutka, "Rigorous thermodynamic treatment of heat generation and conduction in semiconductor device modeling," *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 9, no. 11, pp. 1141–1149, Nov 1990.

[53] H. B. Callen, "Thermodynamics and an introduction to thermostatistics," *Thermodynamics and an Introduction to Thermostatistics, 2nd Edition, by Herbert B. Callen, pp. 512. ISBN 0-471-86256-8. Wiley-VCH, August 1985.*, vol. 1, 1985.

[54] B. Baliga and E. Sun, "Comparison of gold, platinum, and electron irradiation for controlling lifetime in power rectifiers," *Electron Devices, IEEE Transactions on*, vol. 24, no. 6, pp. 685–688, Jun.

[55] J. G. Fossum, "Computer-aided numerical analysis of silicon solar cells," *Solid-State Electronics*, vol. 19, no. 4, pp. 269 – 277, 1976.

[56] J. Fossum and D. Lee, "A physical model for the dependence of carrier lifetime on doping density in nondegenerate silicon," *Solid-State Electronics*, vol. 25, no. 8, pp. 741 – 747, 1982.

[57] J. Fossum, R. Mertens, D. Lee, and J. Nijs, "Carrier recombination and lifetime in highly doped silicon," *Solid-State Electronics*, vol. 26, no. 6, pp. 569 – 576, 1983.

[58] R. Hacker and A. Hangleiter, "Intrinsic upper limits of the carrier lifetime in silicon," *Journal of Applied Physics*, vol. 75, no. 11, pp. 7570–7572, 1994.

[59] V. K. Khanna, "Physical understanding and technological control of carrier lifetimes in semiconductor materials and devices: A critique of conceptual development, state of the art and applications," *Progress in Quantum Electronics*, vol. 29, no. 2, pp. 59 – 163, 2005.

[60] S. Reggiani, E. Gnani, M. Rudan, G. Baccarani, C. Corvasce, D. Barlini, M. Ciappa, null, M. Denison, N. Jensen, G. Groos, and M. Stecher, "Measurement and modeling of the electron impact-ionization coefficient in silicon up to very high temperatures," *Electron Devices, IEEE Transactions on*, vol. 52, no. 10, pp. 2290–2299, 2005.

[61] S. C. Choo, "Theory of a forward-biased diffused-junction p-l-n rectifier 8212;part i: Exact numerical solutions," *Electron Devices, IEEE Transactions on*, vol. 19, no. 8, pp. 954–966, 1972.

[62] N. Arora, J. R. Hauser, and D. Roulston, "Electron and hole mobilities in silicon as a function of concentration and temperature," *Electron Devices, IEEE Transactions on*, vol. 29, no. 2, pp. 292–295, Feb.

[63] C. Canali, G. Majni, R. Minder, and G. Ottaviani, "Electron and hole drift velocity measurements in silicon and their empirical relation to electric field and temperature," *Electron Devices, IEEE Transactions on*, vol. 22, no. 11, pp. 1045–1047, Nov.

[64] V. d'Alessandro and N. Rinaldi, "A critical review of thermal models for electro-thermal simulation," *Solid-State Electronics*, vol. 46, no. 4, pp. 487–496, 2002.

[65] M. Riccio, G. D. Falco, L. Maresca, G. Breglio, E. Napoli, A. Irace, Y. Iwahashi, and P. Spirito, "3d electro-thermal simulations of wide area power devices operating in avalanche condition," *Microelectronics Reliability*, vol. 52, no. 910, pp. 2385 – 2390, 2012.

[66] S. Pendharkar, M. Trivedi, and K. Shenai, "Electrothermal simulations in punchthrough and nonpunchthrough igbt's," *Electron Devices, IEEE Transactions on*, vol. 45, no. 10, pp. 2222 –2231, oct 1998.

[67] M. Otsuki, Y. Onozawa, H. Kanemaru, Y. Seki, and T. Matsumoto, "A study on the short-circuit capability of field-stop igbts," *Electron Devices, IEEE Transactions on*, vol. 50, no. 6, pp. 1525 – 1531, june 2003.

[68] B. G. Svensson and M. Willander, "Generation of divacancies in silicon irradiated by 2 x2010;mev electrons: Depth and dose dependence," *Journal of Applied Physics*, vol. 62, no. 7, pp. 2758–2762, Oct.

[69] R. Chokhawala, J. Catt, and L. Kiraly, "A discussion on igbt short-circuit behavior and fault protection schemes," *Industry Applications, IEEE Transactions on*, vol. 31, no. 2, pp. 256 –263, mar/apr 1995.

[70] M. Barnes, E. Blackmore, G. Wait, J. Lemire-Elmore, B. Rablah, G. Leyh, M. Nguyen, and C. Pappas, "Analysis of high-power igbt short circuit failures," *Plasma Science, IEEE Transactions on*, vol. 33, no. 4, pp. 1252 – 1261, aug. 2005.

[71] M. Trivedi and K. Shenai, "Investigation of the short-circuit performance of an igbt," *Electron Devices, IEEE Transactions on*, vol. 45, no. 1, pp. 313 –320, jan 1998.

[72] ——, "Failure mechanisms of igbts under short-circuit and clamped inductive switching stress," *Power Electronics, IEEE Transactions on*, vol. 14, no. 1, pp. 108 –116, jan 1999.

[73] J. Lutz and T. Basler, "Short-circuit ruggedness of high-voltage igbts," in *Microelectronics (MIEL), 2012 28th International Conference on*, may 2012, pp. 243 –250.

[74] N. Thapar and B. Baliga, "An experimental evaluation of the on-state performance of trench igbt designs," *Solid-State Electronics*, vol. 42, no. 5, pp. 771 – 776, 1998.

[75] I.-Y. Park and Y.-I. Choi, "Trench cathode tigbt with improved latch-up characteristics," *Physica Scripta*, vol. 1999, no. T79, p. 337, 1999.

[76] A. S. Grove, O. Leistiko, and C. T. Sah, "Redistribution of acceptor and donor impurities during thermal oxidation of silicon," *Journal of Applied Physics*, vol. 35, no. 9, pp. 2695 –2701, sep 1964.

[77] D. Dibra, M. Stecher, S. Decker, A. Lindemann, J. Lutz, and C. Kadow, "On the origin of thermal runaway in a trench power mosfet," *Electron*

Devices, IEEE Transactions on, vol. 58, no. 10, pp. 3477 –3484, oct. 2011.

[78] A. Benmansour, S. Azzopardi, J. Martin, and E. Woirgard, "Failure mechanisms of trench igbt under various short-circuit conditions," in *Power Electronics Specialists Conference, 2007. PESC 2007. IEEE*, june 2007, pp. 1923 –1929.

[79] R. Thalhammer, G. Deboy, E. Knauf, E. Kuhbandner, and G. Wachutka, "Calibration of electrothermal power device models using combined characterization techniques," in *Power Semiconductor Devices and IC's, 1997. ISPSD '97., 1997 IEEE International Symposium on*, may 1997, pp. 181 –184.

[80] J. B. P. W. A.R. Brown, A. Asenov and D. Crees, "Calibration of the numerical simulations in the design of high temperature igbt," in *Available from ¡http://www.elec.gla.ac.uk/devmod/papers/isps94/isps.html¿,*, 2001.

[81] G. Deboy, H. Hulsken, H. Mitlehner, and R. Rupp, "A comparison of modern power device concepts for high voltage applications: field stop-igbt, compensation devices and sic devices," in *Bipolar/BiCMOS Circuits and Technology Meeting, 2000. Proceedings of the 2000*, 2000, pp. 134 –141.

[82] A. Irace, G. Breglio, P. Spirito, R. Letor, and S. Russo, "Reliability enhancement with the aid of transient infrared thermal analysis of smart power mosfets during short circuit operation," *Microelectronics Reliability*, vol. 45, no. 911, pp. 1706 – 1710, 2005.

[83] V. Bodryakov and V. Zamyatin, "Analysis of thermodynamic functions of silicon at high temperatures," *High Temperature*, vol. 38, pp. 698–704, 2000.

[84] H. Yilmaz, "Cell geometry effect on igt latch-up," *Electron Device Letters, IEEE*, vol. 6, no. 8, pp. 419 – 421, aug 1985.

[85] C. M. Yun, H. C. Kim, K. H. Lee, J. I. Kim, and T. H. Kim, "Comparison of stripe and cellular geometry for short circuit rated trench igbt," in *Power Semiconductor Devices and ICs, 2000. Proceedings. The 12th International Symposium on*, 2000, pp. 275 –278.

[86] A. Irace, P. Spirito, M. Riccio, and G. Breglio, "Voltage drops, sawtooth oscillations and hf bursts in breakdown current and voltage waveforms during uis experiments," in *Power Semiconductor Devices and ICs (ISPSD), 2012 24th International Symposium on*, June, pp. 165–168.

[87] M. Riccio, A. Irace, G. Breglio, P. Spirito, E. Napoli, and Y. Mizuno, "Electro-thermal instability in multi-cellular trench-igbts in avalanche condition: Experiments and simulations," in *Power Semiconductor Devices and ICs (ISPSD), 2011 IEEE 23rd International Symposium on*, May, pp. 124–127.

[88] F.-J. Niedernostheide, B. S. Kerner, and H.-G. Purwins, "Spontaneous appearance of rocking localized current filaments in a nonequilibrium distributive system," *Phys. Rev. B*, vol. 46, pp. 7559–7570, Sep 1992.

[89] B. Ridley, "Specific negative resistance in solids," *Proceedings of the Physical Society*, vol. 82, no. 6, p. 954, 1963.

[90] X. Perpina, I. Cortes, J. Urresti-Ibanez, X. Jorda, and J. Rebollo, "Layout role in failure physics of igbts under overloading clamped inductive turnoff," *Electron Devices, IEEE Transactions on*, vol. 60, no. 2, pp. 598–605, Feb.

[91] K. G. McKay, "Avalanche breakdown in silicon," *Phys. Rev.*, vol. 94, pp. 877–884, May 1954.

[92] W. Fulop, "Calculation of avalanche breakdown voltages of silicon p-n junctions," *Solid-State Electronics*, vol. 10, no. 1, pp. 39 – 43, 1967.

[93] A. Caruso, P. Spirito, and G. Vitale, "Negative resistance induced by avalanche injection in bulk semiconductors," *IEEE Transactions on Electron Electronics*, vol. 9, pp. 578–587, 1974.

[94] P. Hower and V. Krishna Reddi, "Avalanche injection and second breakdown in transistors," *Electron Devices, IEEE Transactions on*, vol. 17, no. 4, pp. 320–335, Apr 1970.

[95] H. Egawa, "Avalanche characteristics and failure mechanism of high voltage diodes," *Electron Devices, IEEE Transactions on*, vol. 13, no. 11, pp. 754–758, Nov 1966.

[96] M. Muller and H. Guckel, "Negative resistance and filamentary currents in avalanching silicon p+-i-n+junctions," *Electron Devices, IEEE Transactions on*, vol. 15, no. 8, pp. 560–568, Aug.

[97] R. Baburske, J. Lutz, and B. Heinze, "Effects of negative differential resistance in high power devices and some relations to dmos structures," in *Reliability Physics Symposium (IRPS), 2010 IEEE International*, May, pp. 162–169.

[98] H.-J. Schulze, F.-J. Niedernostheide, F. Pfirsch, and R. Baburske, "Limiting factors of the safe operating area for power devices," *Electron Devices, IEEE Transactions on*, vol. 60, no. 2, pp. 551–562, Feb.

[99] A. Alekseev, S. Bose, P. Rodin, and E. Schöll, "Stability of current filaments in a bistable semiconductor system with global coupling," *Phys. Rev. E*, vol. 57, pp. 2640–2649, Mar 1998.

[100] P. Rodin, "Theory of traveling filaments in bistable semiconductor structures," *Phys. Rev. B*, vol. 69, p. 045307, Jan 2004.

[101] F.-J. Niedernostheide, B. S. Kerner, and H.-G. Purwins, "Spontaneous appearance of rocking localized current filaments in a nonequilibrium distributive system," *Phys. Rev. B*, vol. 46, pp. 7559–7570, Sep 1992.

[102] D. Pogany, S. Bychikhin, E. Gornik, M. Denison, N. Jensen, G. Groos, and M. Stecher, "Moving current filaments in esd protection devices and their relation to electrical characteristics," in *Reliability Physics Symposium Proceedings, 2003. 41st Annual. 2003 IEEE International*, March-4 April, pp. 241–248.

[103] H. K. Gummel and H. C. Poon, "An Integral Charge Control Model of Bipolar Transistors," 1970.

[104] A. Alessandria, L. Fragapane, and G. Morale, "Design considerations on field-stop layer processing in a trench-gate igbt," in *Power Electronics and Applications, 2009. EPE '09. 13th European Conference on*, Sept., pp. 1–6.

[105] T. Yong, C. Ming, and W. Bo, "New methods for extracting field-stop igbt model parameters by electrical measurements," in *Industrial Electronics, 2009. ISIE 2009. IEEE International Symposium on*, July, pp. 1546–1551.

[106] X. Kang, A. Caiafa, E. Santi, J. Hudgins, and P. Palmer, "Characterization and modeling of high-voltage field-stop igbts," *Industry Applications, IEEE Transactions on*, vol. 39, no. 4, pp. 922–928, July-Aug. 2003.

[107] N. Rinaldi, "Modeling of minority-carrier transport in semiconductor regions with position-dependent material parameters at arbitrary injection levels," *Electron Devices, IEEE Transactions on*, vol. 43, no. 8, pp. 1256–1263, Aug 1996.

[108] E. Gnani, S. Reggiani, M. Rudan, and G. Baccarani, "Extraction method for the impact-ionization multiplication factor in silicon at large operating temperatures," in *Solid-State Device Research Conference, 2002. Proceeding of the 32nd European*, September, pp. 227–230.

[109] S. Reggiani, M. Rudan, E. Gnani, and G. Baccarani, "Investigation about the high-temperature impact-ionization coefficient in silicon," in *Solid-State Device Research conference, 2004. ESSDERC 2004. Proceeding of the 34th European*, Sept., pp. 245–248.

[110] S. Reggiani, E. Gnani, M. Rudan, G. Baccarani, C. Corvasce, D. Barlini, M. Ciappa, W. Fichtner, M. Denison, N. Jensen, G. Groos, and M. Stecher, "Experimental extraction of the electron impact-ionization coefficient at large operating temperatures," in *Electron Devices Meeting, 2004. IEDM Technical Digest. IEEE International*, Dec., pp. 407–410.

[111] P. Vanmeerbeek, J. Roig, F. Bogman, P. Moens, A. Villamor-Baliarda, and D. Flores, "Enhancing the robustness of a multiple floating field-limiting ring termination by introducing a buffer layer," in *Power Semiconductor Devices and ICs (ISPSD), 2012 24th International Symposium on*, June, pp. 357–360.

[112] J. P. Phipps and K. Gauen, "New insights affect power mosfet ruggedness," in *Applied Power Electronics Conference and Exposition, 1988. APEC'88. Conference Proceedings 1988., Third Annual IEEE*. IEEE, 1988, pp. 290–298.

[113] R. R. Stoltenburg, "Boundary of power-mosfet, unclamped inductive-switching (uis), avalanche-current capability," in *Applied Power Electronics Conference and Exposition, 1989. APEC'89. Conference Proceedings 1989., Fourth Annual IEEE*. IEEE, 1989, pp. 359–364.

[114] K. Fischer and K. Shenai, "Dynamics of power mosfet switching under unclamped inductive loading conditions," *Electron Devices, IEEE Transactions on*, vol. 43, no. 6, pp. 1007–1015, 1996.

[115] E. D'Arcangelo, A. Irace, G. Breglio, and P. Spirito, "Experimental characterization of temperature distribution on power mos devices during unclamped inductive switching," *Microelectronics and reliability*, vol. 44, no. 9-11, pp. 1455–1459, 2004.

[116] P. Hower, C.-Y. Tsai, S. Merchant, T. Efland, S. Pendharkar, R. Steinhoff, and J. Brodsky, "Avalanche-induced thermal instability in ldmos transistors," in *Power Semiconductor Devices and ICs, 2001. ISPSD'01. Proceedings of the 13th International Symposium on*. IEEE, 2001, pp. 153–156.

[117] A. Kwan, K. Teasdale, N. Nguyen, J. Ambrus, and T. McDonald, "Improved soa analysis for trench mosfets using the spirito approach," in *9th Annual Automotive Electronics Reliability Workshop*, 2004.

[118] V. Axelrad, J. Rollins, and S. Motzny, "Current Filamentation and Thermal Instability in a Power BJT Array Cell," in *Solid State Device Research Conference ESSDERC'93*, 1993, pp. 339–342.

[119] T. Shoji, M. Ishiko, T. Fukami, T. Ueta, and K. Hamada, "Investigations on current filamentation of igbts under undamped inductive switching conditions," in *Power Semiconductor Devices and ICs, 2005. Proceedings. ISPSD'05. The 17th International Symposium on*. IEEE, 2005, pp. 227–230.

[120] B. Fatemizadeh, D. Silber, M. Fullmann, and J. Serafin, "Modeling of ldmos and ligbt structures at high temperatures," in *Power Semiconductor Devices and ICs, 1994. ISPSD'94. Proceedings of the 6th International Symposium on*. IEEE, 1994, pp. 137–142.

[121] G. Breglio, A. Irace, E. Napoli, M. Riccio, and P. Spirito, "Experimental detection and numerical validation of different failure mechanisms in igbts during unclamped inductive switching," *Electron Devices, IEEE Transactions on*, vol. 60, no. 2, pp. 563–570, 2013.

[122] L. Rossi, M. Riccio, E. Napoli, A. Irace, G. Breglio, and P. Spirito, "A novel uis test system with crowbar feedback for reduced failure energy

in power devices testing," *Microelectronics Reliability*, vol. 50, no. 911, pp. 1479 – 1483, 2010.

[123] "Lock-in contact thermography investigation of lateral electronic inhomogeneities in semiconductor devices," *Sensors and Actuators A: Physical*, vol. 71, no. 12, pp. 46 – 50, 1998.

[124] J. Rakotoniaina, O. Breitenstein, and M. Langenkamp, "Localization of weak heat sources in electronic devices using highly sensitive lock-in thermography," *Materials Science and Engineering: B*, vol. 91, pp. 481–485, 2002.

[125] O. Breitenstein, J. Rakotoniaina, and M. A. Rifai, "Quantitative evaluation of shunts in solar cells by lock-in thermography," *Progress in Photovoltaics: research and Applications*, vol. 11, no. 8, pp. 515–526, 2003.

[126] M. Riccio, A. Irace, G. Breglio, L. Rossi, M. Barra, F. Di Girolamo, and A. Cassinese, "Current distribution effects in organic sexithiophene field effect transistors investigated by lock-in thermography: Mobility evaluation issues," *Applied Physics Letters*, vol. 93, no. 24, pp. 243 504–243 504, 2008.

[127] G. Breglio, A. Irace, M. Riccio, P. Spirito, K. Hamada, T. Nishijima, and T. Ueta, "Detection of localized uis failure on igbts with the aid of lock-in thermography," *Microelectronics Reliability*, vol. 48, pp. 1432–1434, 2008.

[128] M. Riccio, A. Irace, and G. Breglio, "Lock-in thermography for the localization of pre-breakdown leakage current on power diodes," in *Proc. PRIME'09.*, 2009, pp. 208–211.

[129] O. Breitenstein, W. Warta, and M. Langenkamp, *Lock-in thermography: Basics and use for evaluating electronic devices and materials*. Springer, 2010, vol. 10.

[130] S. Korpar, R. Dolenec, K. Hara, T. Iijima, P. Krian, Y. Mazuka, R. Pestotnik, A. Stanovnik, and M. Yamaoka, "Measurement of cherenkov photons with silicon photomultipliers," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 594, no. 1, pp. 13 – 17, 2008.

[131] V. Andreev, V. Balagura, B. Bobchenko, P. Buzhan, J. Cvach, M. Danilov, E. Devitsin, V. Dodonov, B. Dolgoshein, G. Eigen, L. Filatov, E. Garutti, M. Groll, R.-D. Heuer, A. Ilyin, M. Janata, I. Kacl, V. Kantserov, V. Kaplin, A. Karakash, S. Klemin, V. Korbel, V. Kozlov, H. Meyer, R. Mizuk, V. Morgunov, E. Novikov, S. Nmeek, R. Pschl, I. Polk, E. Popova, A. Raspereza, S. Reiche, V. Rusinov, F. Sefkow, P. Smirnov, S. Smirnov, Y. Soloviev, E. Tarkovsky, A. Terkulov, V. Tikhomirov, . . Valkr, J. Weichert, and J. Zlek, "A high-granularity scintillator calorimeter readout with silicon photomultipliers," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 540, no. 23, pp. 368 – 380, 2005.

[132] A. Otte, J. Barral, B. Dolgoshein, J. Hose, S. Klemin, E. Lorenz, R. Mirzoyan, E. Popova, and M. Teshima, "A test of silicon photomultipliers as readout for pet," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 545, no. 3, pp. 705 – 715, 2005.

[133] D. Renker, "Geiger-mode avalanche photodiodes, history, properties and problems," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 567, no. 1, pp. 48 – 56, 2006.

[134] R. H. Haitz, "Model for the electrical behavior of a microplasma," *Journal of Applied Physics*, vol. 35, no. 5, pp. 1370–1376, May.

[135] S. Cova, M. Ghioni, A. Lacaita, C. Samori, and F. Zappa, "Avalanche photodiodes and quenching circuits for single-photon detection," *Applied optics*, vol. 35, no. 12, pp. 1956–1976, 1996.

[136] E. Gramsch, O. Pcheliakov, I. Chistokhin, and E. Tishkovsky, "Evaluation of a junction termination extension apd for use with scintillators," in *Nuclear Science Symposium Conference Record, 2006. IEEE*, vol. 2, 29 2006-Nov. 1, pp. 1220–1223.

[137] B. E. Saleh and M. C. Teich, *Fundamentals of photonics*, 2007.

[138] W. G. Oldham, R. Samuelson, and P. Antognetti, "Triggering phenomena in avalanche diodes," *Electron Devices, IEEE Transactions on*, vol. 19, no. 9, pp. 1056–1060, Sep.

[139] C. Piemonte, "A new silicon photomultiplier structure for blue light detection," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 568, no. 1, pp. 224 – 232, 2006.

[140] R. Mcintyre, "Theory of microplasma instability in silicon," *Journal of Applied Physics*, vol. 32, no. 6, pp. 983–995, Jun.

[141] R. H. Haitz, "Mechanisms contributing to the noise pulse rate of avalanche diodes," *Journal of Applied Physics*, vol. 36, no. 10, pp. 3123–3131, Oct.

[142] M. Ghioni, A. Gulinatti, I. Rech, F. Zappa, and S. Cova, "Progress in silicon single-photon avalanche diodes," *Selected Topics in Quantum Electronics, IEEE Journal of*, vol. 13, no. 4, pp. 852–862, July-aug.

[143] H. Ruegg, "An optimized avalanche photodiode," *Electron Devices, IEEE Transactions on*, vol. 14, no. 5, pp. 239–251, May.