Design, Simulation and Fabrication of Insulated Gate Bipolar Transistors (IGBT)

Tejas K. Jhaveri

Abstract- **This project serves as a study to determine the feasibility of the current CMOS toolsets and processes available at Semiconductor & Microsystems Fabrication Laboratory (SMFL) for the fabrication of whole wafer power devices. Several designs and devices were explored. The Insulated Gate Bipolar Transistor (IGBT) is a device widely used for high power electronic applications and was selected for this study. This device has bipolar current flow and a MOS gate thus combining advantages of both the Double diffused MOS (DMOS) and Power Bipolar junction transistor.**

Prototypes consisting of transistors with varying densities, gate lengths and gate widths were fabricated to characterize these devices Attempts were made to study the effect of field oxide thickness on breakdown voltage. Photomask were designed in mentor graphics. Process was designed to obtain required power rating. The design was simulated in ATHENA to verify the process conditions.

The IGBTs were fabricated as per the design on standard 4" high resistivity n-type wafers. 8 device wafers and 4 control wafers were used for the process, which involved 4 mask levels. Two additional wafers were processed with this lot to obtain DMOS. The fabricated devices were tested to determine electrical characteristics. The IV characteristics obtained for both the DMOS and IGBT exhibit field effect. However this field effect is in parallel with a parasitic conductance and limits the transistor from turning off. It is also observed that the devices are operational as depletion mode devices instead of enhancement mode devices. The shortcomings of the current process and device designed have been listed and required modifications have been suggested.

Index Terms— **Power Electronics, IGBT**

I. INTRODUCTION

Power Electronics is the use of electronic components for the rectification and control of power. Transistors are the basic building blocks of most electronic components. Transistors that are used in power electronics need to support high power. Thus a special transistors that have the capability to work at the desired current and voltage ratings have been developed for the purpose. These power transistors have found profound applications in various industries varying from automotive to defense. It is desired to study such power transistors at RIT. The double-diffused MOS (DMOS) and power bipolar is generally preferred for high voltage application, whereas the power BJT is preferred in power e electronics where faster switching is desired. However a new device that combine advantages of both these devices, has gained popularity in the market since it was introduced in 1984. Insulated Gate Bipolar Transistor (IGBT) has the advantage of both reduced current loss due to an insulated gate and faster switching due to bipolar current flow. This paper discussed the development for the design and fabrication of IGBT at RIT using the current CMOS toolset and process available at the Semiconductor and Microsystems Fabrication Laboratory (SMFL).

Power transistors generally have larger areas than conventional CMOS devices. It is discovered that the yield of manufacturing is improved by dividing this large areas into smaller areas connected in parallel. To lay the groundwork for fabrication large area devices with high yield, it is critical to obtain a robust device design. This project scrutinizes the performance of an individual Insulated Gate Bipolar Transistor (IGBT). For the purpose each cell on the wafer would include IGBT of different gate lengths. To estimate the

II. DESIGN

A. Device Design

Scaling trends for power transistors are quite different compared to CMOS devices. It is experienced that as power requirement increase the frequency requirement decrease. The simple IGBT device structure is shown below:

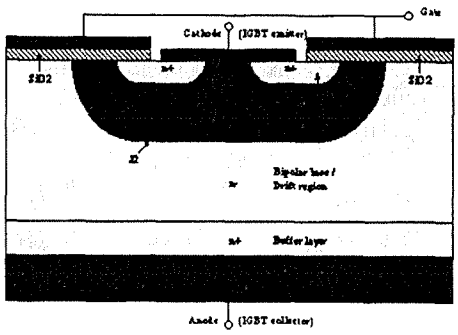

Fig.1: A typical IGBT structure

The transistor is formed by two or more P-N junctions. The current flow in this device is bipolar, thus providing higher current densities due to reduced bulk resistance. Minority carriers are injected into the bulk region from the drain (emitter), which improves the carrier lifetime, thus reducing the resistance. The gate region is formed under the polysilicon by a double diffusion process. In the absence of gate voltage, a large voltage is required at the collector to provide breakdown at the junction J2. This is known as the forward breakdown voltage and is determined by the doping levels of the n– bulk region and the p base region. When a gate voltage higher than the threshold voltage of the device is applied, inversion of the p-channel under the polysilicon occurs, and current flow is achieved from emitter to collector. Device characteristics in standard mode of operation is shown below:

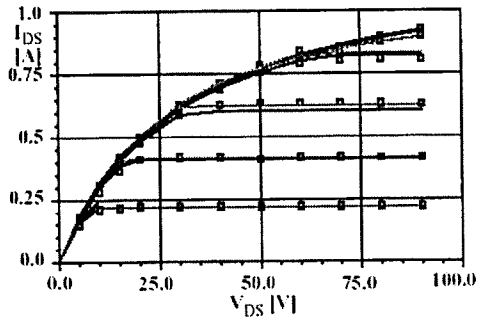

Fig 2: I-V characteristics of IGBT in forward operating mode.

The twp important parameters for the design are breakdown voltage and saturation current. There are two effects that can bring about breakdown: (1) avalanche breakdown, which is dependent solely on the doping of the n- drift region or (2) punchthrough which occurs, when the applied reverse bias causes the depletion region between the junctions J2 and J3 to combine. The n-buffer region is highly doped to improve the DC punchthrough by extending the depletion region only in the P-emitter region under reverse bias. Standard devices in the Industry have voltage-handling capacities of 300V, 600V and 1200V. Hence devices of voltages up to 400V were designed. Using the graph for breakdown voltage vs. doping concentration for abrupt p-n junction a doping of 8E15 cm^{-3} is required for avalanche breakdown at 400V. Hence to fabricate these devices wafers with ρ= 4-5 Ω/cm are required. It was decided to use wafers with ρ= 8-12 Ω/cm, since they are standard wafers used at SMFL.

Doping levels and junction depths at J1 andJ2 can be modeled to obtain breakdown below 400V. The depletion region junction depth equation is shown below:

$$x_p = \sqrt{\frac{2\varepsilon_s}{q}\frac{N_d}{N_a}\frac{1}{N_a + N_d}(-V)} \qquad (1)$$

Assuming Φ = 0, where
V= voltage applied
N_D= n-type doping concentration
N_A= p-type doping concentration
x_n=n-side depletion region depth
x_p=p-side depletion region depth
ε_s= emissivity of silicon

We can get an equation for the depletion junction depth for the p-doped region of J2. When x_p= junction width of double diffused region, punchthrough would occur. Hence the doping concentration in the p-type region is calculated to be 7.75E15. These were done for an abrupt junction but practically p-n junctions are continuous. Hence a range of junction depths were selected and their corresponding punchthrough voltages are shown in the table below:

	1.5	2	2.5
Punchthrough voltage	234.30	416.55	650.85

Table 1: Punchthrough voltage for different effective gate lengths with Na = 8E15 and Nd =5E14 (rho = 8-10).

The doping in the source is selected to be much higher that in the p-well region. It is selected to be around 1E18.

The gate oxide must be designed to handle up to 50V. Assuming breakdown strength of 10MV/cm oxide thickness of 500A is required. Since this device is going to be operated at DC, the time delay due to increased capacitance is not critical. Hence the gate oxide is designed for 700A.

The gate threshold voltage is calculated by:

$$V_T = \frac{\sqrt{4\varepsilon_s kTN_a \ln\left(\frac{N_a}{N_d}\right)}}{\frac{\varepsilon_{ox}}{t_{ox}}} + \frac{2kT}{q}\ln\left(\frac{N_a}{N_d}\right) \qquad (2)$$

where,
N_d = doping concentration in drift region
N_a = doping concentration in well region
T_{ox} =gate oxide thickness

Using the doping concentrations0 calculated above, V_T = 1.8V. Note that this is the threshold voltage at room temperature and neglects oxide charge effects. As more current passes through the device and it heats up, the threshold voltage may increase.

The next critical step for the design was to determine the field oxide thickness. The breakdown voltage for LTO is lower than thermally grown oxide. Hence it is guessed about 2.5 MeV/cm. A minimum field oxide thickness to sustain up to 400V is 1.6um.

The current (I) through the IGBT structure is found by solving the following equation iteratively:

$$V_D = \frac{kT}{q}\ln\left[\frac{Id}{2qWZD_aN_dF\left(\frac{d}{L_a}\right)} + \frac{IL}{\mu_n C_{ox} Z V_G}\right] \quad (3)$$

where,
V_D=Voltage at drain
V_G= voltage at gate
W = gate half length
Z = gate width
L = gate length
d = half thickness of wafer
L_a= ambipolar diffusion length
D_a= ambipolar diffusion coefficient
μ_n= electron mobility in n-channel
F(d/L) = Function obtained from reference 1

B. *Mask design*

To perform a thorough analysis of IGBTs it is necessary to study the effect of gate dimensions. The length of the gate would be varied through the different process splits. However the width of the gates was varied on the mask to obtain devices with gate widths of 4000, 3000, 2000 and 1000um. Each of these four different devices were made as individual devices and as a set of multiple devices connected in parallel to determine the scalability of current for whole wafer device application. The mask also includes several test patterns to determine channel region resistance, p-well bulk resistance, breakdown voltage of junction J2. The layout of mask is shown below:

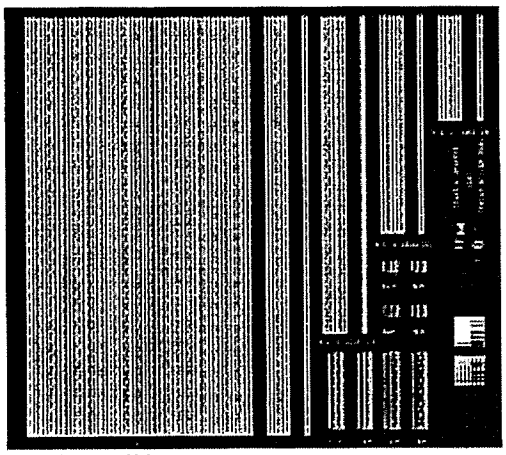

Fig 3: IGBT cell layout

C. *Process Design & Simulation*

The steps involved in processing are listed below. A thorough flow chart of the process is shown in the appendix:

i. Start with bare n-type Si wafers
ii. Backside p-type spin on dopant and predep
iii. Etch oxide
iv. RCA clean
v. Deposit 700 A gate oxide (dry)
vi. Deposit 5000A polysilicon
vii. Level 1 photolithography
viii. Etch poly
ix. P-well implant and diffusion in N_2
x. Level 2 photolithography
xi. Source implant
xii. Source diffusion
xiii. LPCVD oxide
xiv. Level 3 photolithography
xv. Etch contacts through oxide
xvi. Back side polysilicon etch
xvii. RCA clean
xviii. Deposit Aluminum on both sides
xix. Metal definition and etch
xx. Sinter

Once the process parameters requirements are set it is necessary to determine the process parameters to be used to obtain the required output. This was achieved through a simulation of the process using ATHENA. To reduce the time for simulation only the top of the device, where the bulk of the processing was performed, was simulated.

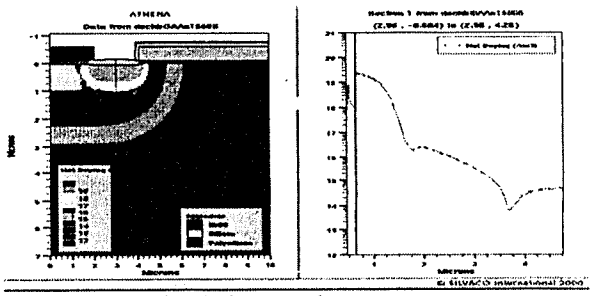

Fig 4: Process simulation template

III. EXPERIMENTAL

The mask designed using CAD tools was fabricated at SMFL using the MEBES III E-beam system. A total of 10 device wafers and 5 control wafers were fabricated in the process. The wafers obtained were standard 4" wafers with resistivity ranging from 8-12 ohm/cm. All the standard 4" processing tools available at SMFL were utilized to fabricate the devices. The ten device wafers were processed with some modifications in each of the runs. The table below summarizes the processes that each of the wafers went through:

TABLE 2: THE DESIGN OF EXPERIMENT

Wafer	Device	Effective gate length (um)	Field oxide thickness (A)
D1,D4,D8	IGBT	2.5	17500
D2, D5	IGBT	1.5	17500
D3, D7	IGBT	2	10000
D6	IGBT	1.5	10000
D9	DMOS	2	10000
D10	DMOS	2.5	17500

Since the double diffused process forms the gate, either the first or the second drive-in process can be varied to obtain the different gate lengths. Since the first drive-in process requires an overnight drive in it was decided to alter the second drive-in step. Simulations were done to determine appropriate diffusion temperatures and times. For the two implant and diffusion steps, the doping concentration in the p-well region was targeted to be around 8E15 with a junction depth of 3.5 um. Simulation suggested that a drive in time of 9hrs at 1100C is required, after B^{11} implant @60KeV with dose of 1E13. The second implant was done at 120KeV with a dose of 1E15. Simulations results for the three different drive-in temperatures and times are summarized in the table below:

TABLE 3: PROCESS CONDITIONS FOR SECOND DRIVE-IN

Effective gate length (um)	Drive in Temperature (C)	Drive in Time (min)
1.5	1000	20
2	1000	1
2.5	950	5

There were a few hiccups during fabrication. During the etching of polysilicon to from the gates, the recipe used etched all the polysilicon, 700A of oxide and 2500A silicon. When this effect was scrutinized, it was realized that the etch recipe was etching at a rate double of the characterized etch rate. It is suggested that this could be attributed to a RIE lag effect. The etch was characterized on a smaller dimensions but was used to etch much larger features.

IV. TESTING

The fabricated devices were tested on the HP 4145 curve tracer for low voltage characteristics. This test equipment has a current and voltage compliance of 100mA and 20V respectively. Hence only low voltage testing was performed on this test station. IV curves for diode, DMOS and IGBT were obtained from testing.

V. RESULTS & ANALYSIS

There were a few hiccups during fabrication. During the etching of polysilicon to from the gates, the recipe used etched all the polysilicon, 700A of oxide and 2500A silicon. When this effect was scrutinized, it was realized that the etch recipe was etching at a rate double of the characterized etch rate. It is suggested that this could be attributed to a RIE lag effect. The etch was characterized on a smaller dimensions but was used to etch much larger features.

During testing it is always recommended to test the least complicated device first. Diodes were tested first. The IV characteristics of the diode is shown in Fig.5.

The diode was tested for –20V to +20V and it displays both forward and reverse bias characteristics. The current saturation observed in forward bias is an artifact of the current compliance of the test equipment

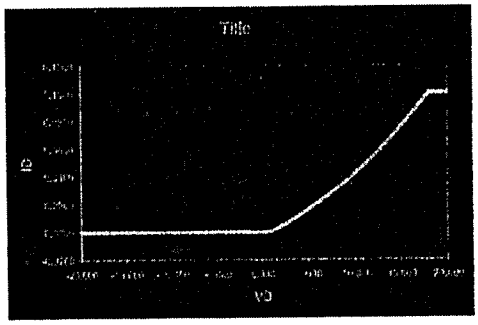

Fig 5: Diode IV characteristics

The next devices to be tested was a DMOS. The device was tested for source to drain voltage of 0 to 20V. At a gate voltage of 0 V the devices were ON, hence it was realized that the devices were depletion mode devices and not enhancement mode devices. This could be attributed to oxide charge effects. Since the p-well is lightly doped, it would oxide charges would keep the p-channel in inversion. Thus these devices were tested for gate voltages from 0 to –10V with –1V increments.

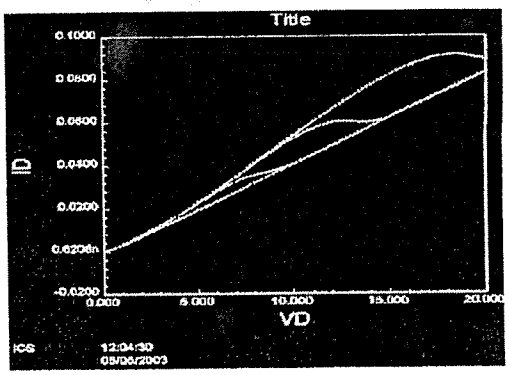

Fig 6: DMOS IV characteristics

The IV characteristics do show a field effect. It is observed that the saturation current decreases with decreasing gate voltage (as it should in a depletion mode device). But there is some sort of conductive path parallel to the FET that is purely dependent on the drain to source bias. Thus the current is dominated by the FET or resistive component, depending on the applied biases at the drain and gate. The resistance of this path is calculated to be ~25Ω.

The IGBT was tested under similar conditions as the DMOS. A similar FET parallel with resistance effect is observed for the IGBT. The resistance of the conducting path is ~20Ω. The decrease in resistance is due to the bipolar current flow in the device.

It is also observed that under similar conditions, the saturation current of the IGBT is much higher than that of the DMOS. This is the drive behind the development of the IGBT.

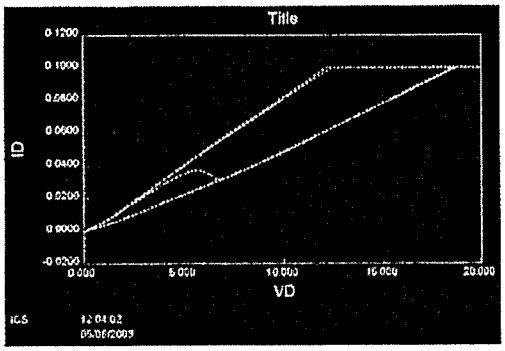

Fig 7: IGBT IV characteristics

The parasitic conductive path is a cause of concern for the devices. It could be attributed to the low resistivity substrates used. To solve this problem three modifications to the process are suggested: (a) Use high resistivity substrates, (b) Use a thinner drift region by using a n-epi layer on p-type substrate and (c) Use higher doping concentration in p-well region

VI. CONCLUSIONS

Development of power semiconductor devices has been initialized at RIT with the first IGBT fabricated at SMFL. Although these devices have serious shortcomings an operational field effect has been demonstrated. This filed effect is plagued by a parallel parasitic conductance channel that hampers the complete functionality of the device. The causes of failure are yet under analysis and few modifications to the process are required. It is necessary to use high resistivity n-epi on p-type wafers to obtain a shorter drift region with lower carrier lifetime. A higher doping in the p-well region is necessary to avoid the formation of parallel parasitic resistance and compensating for high gate oxide charge.

FURTHER WORK

First thing to be done is alter the process as suggested. Adding CV test structures in chip would be beneficial to characterize the oxide charge effects. Eventually it is desired to use knowledge from this experiment obtain whole wafer device. Finally, research in the field of alternative gate materials to extend IGBTs for higher frequency applications would be done.

ACKNOWLEDGEMENTS

The author would like to acknowledge Dr. Santosh Kurinec and Dr. Karl Hirshcman to advise the project. The useful comments and guidance provided by both have been invaluable. The assistance and support from all the staff at SMFL is highly appreciated. The author is thankful to Dr. Sean Rommel for assistance with electrical testing and characterization and Jon Reese for help with ATHENA simulation.

The author would like to thank his father Kamlesh Jhaveri, mother Charu Jhaveri and brother Mihir Jhaveri for always being supportive of his academic and endeavors.

REFERENCES

[1] J. Baliga, Modern power Devices, Wiley, New York, 1987

[2] M Anderson et al, A Comparative Study of

[3] Physical and Subcircuit Models for MOS-Gated Power Devices, rf.rfglobalnet.com/library/Papers/files/5/ispsd94paper.htm

[4] Streetman & Banerjee, Solid state electronic devices, Prentice hall, New Jersey, 2000.

[5] The Insulated Gate Bipolar Transistor (IGBT), http://www.elec.gla.ac.uk/groups/dev_mod/papers/igbt/igbt.html

Tejas Jhaveri is a native of Bombay, India. He is currently a 4th yr student and expected to graduate in May 2004. He has worked as a development engineer co-op with Photronics Inc, Austin (TX) and as a undergraduate research co-op at RIT under Dr. Grande. He has been awarded as RIT's distinguished Outstanding Undergraduate Scholar. He is also a primary author of a paper on "Rotation Induced Measurement Error by a CD SEM", published at SPIE's Microlithography 2003.