Bipolar effects in snapback mechanism in advanced n-FET transistors under high current stress conditions

Pragati Singh1, Rudra Sankar Dhar2, Srimanta Baishya3 and Amitabh Chatterjee3

1 Department of Electronics & Communication Engineering, National Institute of Technology Mizoram, Aizawl-796012, India
2 Department of Electronics and Communication Engineering, National Institute of Technology Silchar, India
3 Department of Electronics and Communication Engineering, Shiv Nadar University, India

E-mail: prags.b2@gmail.com

Keywords: snapback, breakdown, high current, ZRAM

Abstract

This work models high current snapback behavior in n-FET transistors with bottom body contact under high current stress at the drain for ZRAM (Zero capacitor RAM). We analyze 2D current flow in n-FET near the pinch-off region and relate the results to the S-shaped snapback characteristics under high injection avalanche generated carriers conditions. The role of surface bipolar effects on the first snapback phenomenon in the GG-NMOS (Gate-Grounded NMOS) is investigated. A novel physical insight of the bipolar activity is modeled through current flow and barrier lowering at the source-substrate junctions. Moreover, the coupling of electron and hole injection is described. Finally, a model for surface potential inside the substrate before and after snapback is derived and compared with TCAD simulation results. A very good agreement is observed between the model and the TCAD results.

1. Introduction

Physics of ZRAM is based on parasitic bipolar formation in advanced nano-meter scaled FET transistors. As we go to post CMOS era, ZRAM is a critical device for memory circuit in microprocessors. The lack of clear understanding of the bipolar physics has lead to inaccurate model development for a critical feature of snapback memories. However, the first snapback behavior is not clearly understood [1, 2]. Figure 1 shows the schematic of a GG-NMOSFET, along with its I-V characteristics, when a current ramp is applied at the drain terminal. The snapback behavior in such devices is primarily related to the bipolar turn-on mechanism [3]. However, the electrostatic coupling between source(S), drain (D) and gate terminals has not been clearly addressed.

In the body of NMOSFET bipolar turn on is taking place which is causing snapback as shown in figure 2, where source terminal is behaving like emitter of the BJTs, drain terminal is behaving like collector of the BJTs and body or substrate is behaving as base of the BJTs. Sufficiently large voltage across the collector and base causing impact ionization, with some of the generated carriers then acting as the initiating current as they flow into the base. Once this initiating current flows into the base, the transistor turns on and the collector voltage decreases to the snapback holding voltage shown in current voltage characteristic section of figure 1. This voltage happens at the point where the processes of base current generation and the bipolar transistor turning on are in balance. The collector-emitter current of the bipolar transistor decreases the collector voltage, resulting in a lower electric field. It leads to a smaller impact ionization or avalanche current and thus smaller base current, which weakens the bipolar action.

Traditionally in an NMOS, lateral parasitic bipolar turn-on mechanism under snapback has been described by crowding of electrons, which is manifested as the base push-out/Kirk effect [3–7]. The critical electrostatics of hole crowding near ‘pinch-off’ has been overlooked [8]. Moreover, potential build-up due to 2-D current flow in the n-FET devices, so far lacks clear physical insight [9]. Earlier work involves studying the effects of different parameters on bipolar turn-on mechanisms [10–12], but the critical electrostatics of different regions of the NMOSFET under high current stressed condition and their mutual coupling are often overlooked. Also accurate
modeling of important parameters such as \( (V_{t1}; I_{t1}) \), \( V_h \) and \( (V_{t2}; I_{t2}) \) is more involved \[13–18\]. In this work, we describe the fundamental mechanisms of current flow in the 2D-FET, based on the microscopic features of snapback and the impact of 2D current flow, which determines the I-V characteristics. The analysis is critical in understanding the nano-meter scale n-FET structures under high current conditions. Moreover, the translation of these TCAD models into CAD environments for building circuit simulation models will be useful for the circuit designer. In this article, we report snap-back features during bipolar turn-on and model the high current injection phenomenon. In section 2, we describe the process of snapback and surface bipolar activity using two-dimensional TCAD simulations under high current stress conditions. The crowding of holes at the drain contact and the injection of holes in the substrate and coupling with the source region is discussed. In section 3, analytical models of the surface potential is derived for the bulk region. TCAD result is compared to the derived analytical expression in section 4. Finally, the mutual dependence and coupling of the source and drain are analyzed.

2. 2D device simulation and physics of high current injection

The structure of the GG-NMOS transistor used in the study is shown in figure 1. The 2D device characteristics were simulated using the Sentaurus TCAD device simulator with default parameter coefficients. Here, We...
performed 2D transient device simulations and investigated the mechanisms of snapback under high current injection, when a current ramp is applied to the drain contact. In the simulations, the current ramp is adjusted to study the dependence of snapback on the rise time of the pulse.

The application of a current ramp to the drain terminal of the GG-NMOS causes the electric field to increase in the drain-bulk depletion region. Avalanche breakdown is triggered at the $p/n^-$ junction (shown in figure 1), which results in avalanche-generated holes moving towards the bulk region and avalanche-generated electrons being collected by the drain terminal. Impact ionization peak is maximum at A, where the electric field is highest due to the space charge of the mobile carriers. Figure 3 shows pre and post snapback flow lines indicating Post snapback current flowlines smoothen when recombination current dominates. Whereas, Point A in I-V characteristic related to the 2-D structure showing pre snapback flow lines and point B post snapback flow lines in figure 3. The space charge of the mobile carriers pushes the carriers across the drain contact, where they accumulate - (shown in the band diagram of figure 4). The flow of holes across the substrate influences the barrier height on the source side, primarily within the bulk (SB) region, while the gate region is more or less shielded from the influence of the flow of holes flowing towards the substrate. Moreover, the space charge due to holes are initially imaged by the bulk contact towards the bottom. Thus injected excess holes are not imaged on the source side due to coulombic shielding (see figure 4(a)). Furthermore, the hole accumulation at the drain-bulk junction increases the perturbation in electric field and change in barrier height is observed, exhibiting the coupling between the source and bulk regions. The coupling can also be understood through the potential variation in the bulk region as shown in figure 4(b) in a 2D device structure. Thus, due to establishment of the electric field, holes can start flowing towards the bulk contact and the hole-current density increases in the bulk region which is visible in the (contour profile) buildup of hole in figure 4(a). Hole injection into the substrate leads to perturbation of potential (as shown in the figure 4(b) and in the process leads to electron injection below the surface. Now electron injection from the source end leads to compensation of space charge due to holes and in the process triggers snapback as shown in the figure 4(c). Since the drain-bulk depletion region is under the influence of space charge limited (SCL) transport, it acts as a current-controlled two dimensional SCL resistor and exhibits a ballast action due to hole crowding, which leads to potential build-up across the $p/n^-$ junction and influences the electric field in the bulk region (figure 4). Moreover, the hole ballast action due to excess hole current shows a linear characteristic and it is worth noting that in the absence of avalanche-generated holes, the voltage drop is entirely across the depleted junction.

2.1. Diffusing injected minority carriers across the surface
From the above discussion, we observe two dimensional electrostatic coupling between the source and bulk, as excess holes are injected from the drain junction. However, the crowded mobile charges are imaged primarily by the bulk contact; thus, the bands bend in the y-direction (see figure 4 band diagram). As the hole injection...
becomes stronger, it also perturbs the potential in the x-direction. The barrier to the minority carriers (i.e., electrons) near the source is determined both by the electrostatic coupling of the gate and the substrate with the source, which leads to injection of electrons into the channel below the oxide.

1. Bulk injected electrons on the source side reduce the width of the depleted channel. The holes accumulate on the drain side.

2. Initially, in the absence of any injection from the source, diffusion of carriers dominates. In order for the drift to be dominant the band bending in the semiconductor should increase continuously towards the drain side. In the absence of constant gate voltage along the channel, the band bending in the oxide has to decrease by the same amount.

3. In order to decrease the oxide voltage, the surface injected charges should decrease faster than the depletion charge.

4. Therefore, the change in band bending in the semiconductor and the oxide is negligible as one approaches the drain.

5. However, it is not possible because the device is far away from weak inversion region (in absence of gate potential). The maximum voltage drop in the channel is much less than $\phi_p$, which keeps the entire channel away from weak inversion.

6. The combination of very small voltage drop in the channel, coupled with the small charge in the weak inversion layer, makes drift current negligibly small in the channel.

Therefore, one can observe that in the absence of gate bias the injected carriers primarily diffuse below the oxide in the bulk and even if sufficient holes are injected from the drain region (large barrier lowering), the diffusion component remains dominant. However, During the entire cycle of bipolar turn-on mechanism, the first
snapback behavior is observed. The simulated results show complete bipolar turn-on process and a potential build-up as injection of hole increases and therefore injection of electron from source side, as shown in figure 4.

2.2. Salient features of surface and bulk bipolar turn-on & bipolar gain of the device

Current injection from the source end is maximum across the bulk, where not only the barrier height is minimum but also the base transport factor is maximum as shown in figure 4.

1. However, the minority electrons are swept across the surface where initially the electrons are attracted to the depleted channel.

2. Now, the carrier gradient is maximum primarily because the electron concentration is maximum along the surface and slowly the carrier concentration increases towards drain. Thus the carriers which are injected in the bulk are swept along the surface. As the bipolar injection increases, the surface carriers efficiently shield the injected bulk electrons and in the process they begin to flow in the bulk.

3. Similarly, towards the drain side, holes accumulate near the surface, shielding the gate from attracting holes, which prefer to stay below the gate on the right side. To summarize, injected electrons accumulate along the surface from the source end and holes accumulate from the drain end. Slowly, the substrate below the gate attains neutrality and injected electrons begin to flow in the bulk. As the bipolar structure turns on, the activity slowly shifts to the bulk where the injection has started at first and the recombination is maximum. Once excess carrier build-up stops, the voltage buildup stops and one can clearly observe the onset of negative resistance, which is taken as the point where the device begins to filament first in 2D and subsequently 3D.

4. The current injection is maximum in the bulk. However, electrons are swept along the surface where the gradient is maximum. Now the carrier gradient is maximum primarily because the electron concentration is maximum at the surface.

5. Where it was assumed that the semiconductor is non-degenerate and that the difference between the electron and hole quasi Fermi energies in electron volt equals the applied voltage in volt. High injection of carriers causes to violate one of the approximations made in the derivation of the ideal diode characteristics, namely that the majority carrier density equals the thermal equilibrium value. Excess carriers will dominate the electron and hole concentration and can be expressed in the following way.

\[ n_p p_n = n_p^2 e^{\frac{V}{V_T}} \]  
\[ n_p = -\frac{n_p^2 e^{\frac{V}{V_T}}}{2} + \sqrt{\frac{n_p^2 e^{\frac{V}{V_T}} + 4N_n}{4N_n}} \approx \sqrt{\frac{n_p^2 e^{\frac{V}{V_T}}}{4N_n}} + 1 \]

3. Analytical modeling

The schematic cross section of the GG-NMOS transistor is shown in figure 5, where L is the length of the channel and W is the width of the device. To develop the analytical model for the first snapback mechanism in the n-FET devices, we described the surface bipolar turn-on mechanism. The device is divided into three regions: (i) source region (ii) drain region, and (iii) bulk region. The carrier injection electrostatics are summarized in figure 4, showing the nature of the potential distribution, along with the flow of holes and electrons in the bulk region. The coupled electrostatics of the drain and bulk can be determined by the amount of hole and electron injection in the bulk region. In order to determine the potential distribution in the bulk region, a cubic polynomial expression for potential is taken as an approximation.

Poisson’s equation in the bulk region is:

\[ \frac{\partial^2 \psi(x, y)}{\partial x^2} + \frac{\partial^2 \psi(x, y)}{\partial y^2} = -\frac{J_T}{\epsilon_{Si} V_{Sat}} \]

where, \( \psi(x, y) \) is the electrostatic potential in the channel region, \( \epsilon_{Si} \) is the permittivity of silicon, \( J_T = J_h + J_e \) is the total injected hole and electron current density in the bulk, and \( V_{Sat} \) is the saturation velocity. The potential variation inside the MOSFET can be approximated by the cubic polynomial potential function as described in [19]

\[ \psi(x, y) = a_0(x) + a_1(x)y + a_2(x)y^2 + a_3(x)y^3 \]
The coefficients $a_0(x)$, $a_1(x)$, $a_2(x)$, and $a_3(x)$ are functions of $x$ only. To find these coefficients, we use the boundary conditions as mentioned below.

(1) The electric displacement at the silicon body/gate oxide interface is continuous.

\[
\frac{\partial \psi(x, y)}{\partial y} \bigg|_{y=0} = \frac{\varepsilon_{\text{ox}}}{\varepsilon_{\text{Si}}} \times \left( \frac{\psi_f(x)}{t_{\text{ox}}} - \frac{(V_G - V_{FB})}{t_{\text{ox}}} \right)
\]

where, $\psi_f(x)$ is the electrostatic potential at the silicon body/gate oxide interface, $\varepsilon_{\text{ox}}$ represents the gate oxide dielectric constant, $t_{\text{ox}}$ is the thickness of the gate oxide layer, $V_G$ is the applied gate voltage and $V_{FB}$ is the flat band voltage.

(2) The electrostatic potential at the back side of the silicon body is zero since the substrate is grounded.

\[
\psi(x, t_b) = 0
\]

(3) The electric field at the back side of the silicon body is zero.

\[
\frac{\partial \psi(x, y)}{\partial y} \bigg|_{y=t_b} = 0
\]

(4) The source end potential is simply the built-in potential $V_{bi}$ between the n-type source and p-type channel since the source is grounded.

\[
\psi(x, 0) = V_{bi}
\]

(5) The drain end potential is the built-in potential plus the drain bias voltage $V_{DS}$.

\[
\psi(x, L) = V_{bi} + V_{DS}
\]

Here, we have assumed an abruptly doped source/drain-to-channel junction and corner effects [20] are neglected.

Now differentiating (4) and substituting the boundary conditions 1 and 2 in it, we can find the values of the coefficients $a_0(x)$, $a_1(x)$ and $a_2(x)$ as

\[
a_0(x) = \psi_f(x)
\]

\[
a_1(x) = \frac{\varepsilon_{\text{ox}}}{\varepsilon_{\text{Si}}} \times \left( \frac{\psi_f(x)}{t_{\text{ox}}} - \frac{(V_G - V_{FB})}{t_{\text{ox}}} \right)
\]

Figure 5. Schematic cross sectional view of Gate Grounded NMOS (GGNMOS) Transistor.
\[ a_2(x) = \frac{3a_1(x) - 2a_2(x)t_b}{t_b^3} \]
\[ a_3(x) = \frac{a_2(x)t_b - 2a_1(x)}{t_b^3} \]  

### 3.1. Potential distribution at the surface

Substituting (4) along with the obtained coefficient values in (3) for \( y = 0 \), we can get

\[ \frac{\partial^2 \psi_f(x)}{\partial x^2} - C_{1,f} \psi_f(x) = C_{2,f} \]  

The values of \( C_{1,f} \) and \( C_{2,f} \) are obtained by putting the values of \( a_0(x) \), \( a_1(x) \), \( a_2(x) \), and \( a_3(x) \) in (4) and then substituting (4) in the (3) and by setting \( y = 0 \) for surface of the device,

\[ C_{1,f} = \frac{6}{t_b^2} - \frac{4C_{ox}}{t_b \varepsilon_{Si}} \]
\[ C_{2,f} = -\frac{J_T}{\varepsilon_{Si} V_{Sat}} - \frac{4C_{ox} (V_G - V_{T})}{t_b \varepsilon_{Si}} \]  

After solving the above second order differential equation using the boundary conditions (8) and (9), we have the gate-oxide/silicon-body potential distribution:

\[ \psi_f(x) = K_{1,f} \exp\left(\sqrt{C_{1,f}}x\right) + K_{2,f} \exp\left(-\sqrt{C_{1,f}}x\right) - \delta \]  

where

\[ K_{1,f} = \frac{C_{2,f} (e^{\sqrt{C_{1,f}}L} - 1)}{C_{1,f} (e^{\sqrt{C_{1,f}}L} - 1)} + \frac{V_{th} (e^{\sqrt{C_{1,f}}L} - 1) + V_{DS} e^{\sqrt{C_{1,f}}L}}{e^{2\sqrt{C_{1,f}}L} - 1} \]
\[ K_{2,f} = \frac{C_{2,f} e^{\sqrt{C_{1,f}}L} (e^{\sqrt{C_{1,f}}L} - 1)}{C_{1,f} (e^{\sqrt{C_{1,f}}L} - 1)} + \frac{e^{\sqrt{C_{1,f}}L} (V_{th} (e^{\sqrt{C_{1,f}}L} - 1) - V_{DS})}{e^{2\sqrt{C_{1,f}}L} - 1} \]
\[ \delta = \frac{C_{2,f}}{C_{1,f}} \]

This potential distribution expression has been verified with Sentaurus device simulator results, as shown in figure 6 and very good agreement between the two is found.

### 3.2. Turn-on mechanism

The electron and hole current density inside the bulk region after bipolar turn-on can be determined by considering the bipolar action of the device. Due to the application of a pulse signal at the drain terminal a hole current will flow toward the substrate contact. This current will increase the potential inside the bulk region and when a sufficient forward bias is established between source and bulk, electron current from source to drain starts to flow and bipolar turn on may occur. This electron current \( (I_S) \) will cause hole current \( (I_{sub}) \) to flow from bulk to source and drain current \( (I_D) \) to flow from bulk to drain.

\[ I_T = I_S + I_D + I_{sub} = 0 \]  

where,

\[ I_S = I_{so} \left( e^{\frac{V_{th}}{NT}} - 1 \right) \]

where, \( I_{so} \) is the reverse saturation current due to diffusion of holes in the source of the MOSFET. The hole current \( I_{sub} \) is a function of the avalanche multiplication factor ‘M’ and electron current \( I_S \).

\[ I_{sub} = (M - 1) I_S \]

The multiplication factor can be approximated as;

\[ M = \frac{1}{1 - \alpha x_d} \]

where \( x_d = \sqrt{\frac{2q}{\varepsilon_{Si} N_D} (V_{th} + V_{DS})} \) is the width of the depletion region and \( \alpha \) is given as;

\[ \alpha = A e^{-\frac{E}{F}} \]

A and B are the ionization coefficients and \( E \) is the electric field in the high field region.
4. Conclusions

In this paper, we model the high current ambipolar mechanism of n-FET devices under high current stressing. The formation of bipolar junction within Gate Grounded NFET for snapback based memory (ZRAM) application has been explained. The application of an high current pulse to the drain contact, which results in hole crowding and the injection of holes to the substrate is discussed. Moreover, the coupling mechanism of the drain-substrate junction to the source-substrate junction is modelled and used to explain the bipolar turn-on mechanism. The coupling between electron and hole injection and the electrostatics of different regions of the GG-NMOSFET is discussed.

Acknowledgments

Authors would like to thank Mr Sourav Roy, graduate student University of Utah for assisting in preparation of the manuscript.

ORCID iDs

Pragati Singh  https://orcid.org/0000-0002-2363-5717
Rudra Sankar Dhar  https://orcid.org/0000-0002-6571-3808

References

[1] Duvvury C and Boselli G 2004 Esd and latch-up reliability for nanometer cmos technologies Electron Devices Meeting, 2004. IEDM Technical Digest. IEEE International (IEEE) pp 933–6
[2] Duvvury C, Briggs D, Rodrigues J, Carvajal F, Young A, Redwine D and Smayling M 1995 Efficient npn operation in high voltage nmosfet for esd robustness Electron Devices Meeting, 1995. IEDM ’95., International (IEEE) pp 345–8
[3] Shrivastava M, Schneider I, Baghini M S, Gossner H and Rao V R 2010 On the failure mechanism and current instabilities in resurf type dmemos device under esd conditions Reliability Physics Symposium (IRPS), 2010 IEEE International (IEEE) pp 841–5
[4] Hayden J, Burnett D and Nangle J 1991 A comparison of base current reversal and bipolar snapback in advanced npn bipolar transistors IEEE Electron Device Lett. 12 407–9
[5] Chatterjee A, Duvvury C and Banerjee K 2003 New physical insight and modeling of second breakdown lst phenomenon in advanced esd protection devices Electron Devices Meeting, 2005. IEDM Technical Digest. IEEE International (IEEE) pp 195–8
[6] Chatterjee A, Brewer F, Gossner H, Pendharkar S and Duvvury C 2010 Robust high current esd performance of nano-meter scale denmos by source ballasting Reliability Physics Symposium (IRPS), 2010 IEEE International (IEEE) pp 853–6

[7] Chatterjee A, Pendharkar S, Gossner H, Duvvury C and Banerjee K 2008 3d device modeling of damage due to filamentation under an esd event in nanometer scale drain extended nmos (de-nmos) Reliability Physics Symposium, 2008. IRPS 2008. IEEE International (IEEE) pp 639–40

[8] Chatterjee A, Shrivastava M, Gossner H, Pendharkar S, Brewer F and Duvvury C 2011 An insight into the esd behavior of the nanometer-scale drain–extended nmos device–part i: Turn–on behavior of the parasitic bipolar IEEE Transactions on Electron Devices 58 309–17

[9] Chatterjee A, Shrivastava M, Gossner H, Pendharkar S, Brewer F and Duvvury C 2011 An insight into esd behavior of nanometer-scale drain extended nmos (denmos) devices: Part ii (two-dimensional study–biasing & comparison with nmos) IEEE Trans. Electron Devices 58 318–26

[10] Duvvury C and Amerasekera A 1996 State-of-the-art issues for technology and circuit design of esd protection in cmos ics Semicond. Sci. Technol. 11 833

[11] Hower P and Merchant S 1999 Snapback and safe operating area of Idmos transistors Electron Devices Meeting, 1999. IEDM’99. Technical Digest. International (IEEE) pp 193–6

[12] Chung Y Y and Baird B 2000 Electrical-thermal coupling mechanism on operating limit of Idmos transistor Electron Devices Meeting, 2000. IEDM’00. Technical Digest. International (IEEE) pp 83–6

[13] Vassilev V, Groeseneken G, Bock K and Maes H 1999 A compact mosfet breakdown model for optimization of gate coupled esd protection circuits Solid-State Device Research Conf., 1999. Proceeding of the 29th European (IEEE) vol 1, pp 600–3

[14] Chatterjee A, Pendharkar S, Lin Y-Y, Duvvury C and Banerjee K 2007 A microscopic understanding of nanometer scale denmos failure mechanism under esd conditions Electron Devices Meeting, 2007. IEDM 2007. IEEE International (IEEE) pp 181–4

[15] Reggiani S et al 2005 A new numerical and experimental analysis tool for esd devices by means of the transient interferometric technique IEEE Electron Device Lett. 26 916–8

[16] Mergens M, Wilkening W, Mettler S, Wolf H, Stricker A and Fichtner W 1999 Analysis and compact modeling of lateral dmos power devices under esd stress conditions Electrical Overstress/Electrostatic Discharge Symposium Proc., 1999 (IEEE) pp 1–10

[17] Diaz C H, Kang S-M and Duvvury C 1994 Circuit-level electrothermal simulation of electrical overstress failures in advanced mos i/o protection devices IEEE Trans. Comput. Aided Des. Integr. Circuits Syst. 13 682–93

[18] Dutton R W 1975 Bipolar transistor modeling of avalanche generation for computer circuit simulation IEEE Trans. Electron Devices 22 334–8

[19] Young K 1989 Analysis of conduction in fully depleted SOI MOSFETs IEEE Trans. Electron Devices 36 504–6

[20] Fossum J, Yang J-W and Trivedi V 2003 Suppression of corner effects in triple-gate MOSFETs IEEE Electron Device Lett. 24 745–7