ELSEVIER

Microelectronics Reliability 44 (2004) 1817–1822

MICROELECTRONICS RELIABILITY

www.elsevier.com/locate/microrel

The Impact of CMOS technology scaling on MOSFETs second breakdown: Evaluation of ESD robustness

O. Semenov[a], H. Sarbishaei[a], V. Axelrad[b] and M. Sachdev[a]

[a] Dept. of Electrical and Computer Engineering, University of Waterloo, Waterloo, Canada N2L 3G1
[b] Sequoia Design Systems, 137 Chapman Rd., Woodside, CA, USA 94062

Abstract

The impact of CMOS technology scaling on the second breakdown of ESD protection devices has been investigated using 2-D simulations and analytical calculations. It is shown that the second breakdown trigger current (It_2) can not be reliably used as an ESD robustness criterion in sub-0.18 um ESD protection devices. When a technology feature size is reduced, the doping of drain and drain extension regions is significantly increased. Thus, the ESD device failure due to the self-heating effect occurs without the second snapback region in high current I-V curve and It_2 current can not be properly extracted. Instead of It_2 current criterion, we propose to use the maximum failure temperature criterion.
© 2004 Elsevier Ltd. All rights reserved.

1. Introduction

As CMOS technology is scaled into the sub-0.18 um region, one major concern is the ability to design highly reliable electrostatic discharge (ESD) protection circuits. Thin oxide N-MOSFET is the most commonly used structure in ESD protection circuits implemented in advanced CMOS processes. In deep submicron technologies, there are many concerns regarding ESD robustness due to thinner gate oxides (30 - 40 Å), smaller channel lengths, shallower junctions (≈ 0.1 um), higher implantation doping of source/drain extensions ($\sim 1\times10^{21}$ cm^{-3}) and use of fully silicided technology. Since, the N-MOSFET should be capable of dissipating all the ESD energy by itself without intrinsic damage, the high-current performance of the N-MOSFET determines the ESD capability of a given protection circuit design. Typical target specification of ESD stress level is 2 kV and above [1].

In advance CMOS processes, the ESD failure is initiated by the heating at the junction edge after current localization due to second (thermal) breakdown [2]. In previous investigations, the second breakdown trigger current (It_2) has been used as the robustness monitor [1,2]. This current is related to ESD threshold voltage by a multiplying factor of between 1500 and 2000 Ohms. Higher It_2 current indicates higher ESD robustness. However, our simulations show that the second breakdown trigger current criterion of ESD robustness is impractical for sub-0.18 um CMOS technologies. For these technologies, the ESD failure due to the self-heating effect occurs without the second snapback in current-voltage characteristics of N-MOSFET and It_2 current can not be properly detected from these measurements or simulations. Instead of It_2 current criterion, we propose to use the maximum failure temperature criterion. The upper limit of failure temperature in ESD device is the melting temperature of silicon and the lower limit of failure temperature can be the temperature of interconnects and contacts degradation in the device under ESD stress.

The paper is organized as follows. In Section 2, we review typical criteria of ESD robustness and discuss its limitations. The ESD device structures used in our

0026-2714/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.microrel.2004.07.091

research are given in Section 3. The circuit and device simulation results (temperature distribution and current-voltage characteristics of analyzed structures under ESD conditions) are considered in Section 4. In Section 5, using the theory of second breakdown in MOSFETs, we provide the explanation of simulation results and show the conditions for the second breakdown occurrence in deep submicron devices. The paper is concluded in Section 6.

2. Criteria of ESD Robustness

The following ESD robustness criteria are widely used: the second breakdown current (It_2) [1,2], the leakage current (I_{off}) measured after ESD stress [3], and the failure temperature or the current at failure temperature [4].

2.1 It2 current criterion

A typical current-voltage (I-V) curve of an N-MOSFET with grounded gate, source and substrate is shown in Fig. 1. The transistor goes into avalanche followed by the snapback (parasitic bipolar) action at the voltage Vt_1 and the current It_1. During ESD event, the device operates mostly in the snapback mode, where V_{sp} is the clamping snapback voltage. As the current increases further, a second breakdown occurs at the voltage Vt_2 and the current It_2. In this operating region, current filament formation is enhanced and device failure eventually occurs. The It_2 current can be easily measured. However, the actual damaging ESD current usually drives devices well beyond the trigger point of the second breakdown [4]. Hence, this criterion gives the pessimistic estimation of ESD robustness.

2.2 Leakage current (I_{off}) criterion

It_2 corresponds to the current that flows through the device just prior to device failure. However, in many cases, the device exhibits an increased drain-to-substrate leakage current even before it reaches the second breakdown point in the I-V curve. Because of this, the leakage current is used as a monitor of ESD robustness. Typically, the I_{off} criterion is based on the specified value of leakage current after the ESD stress. If the leakage current is higher than the specified value, which is typically 10 nA, 100 nA or 1 uA, then the device is considered failed according to this failure criterion.

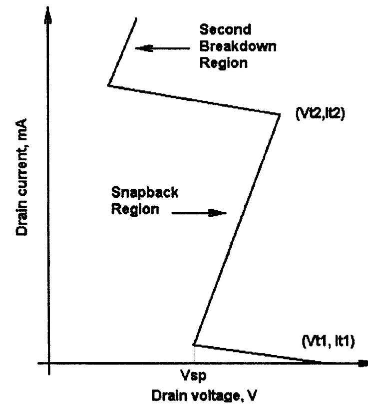

Fig.1. Typical I-V characteristics of N-MOSFET in the high current mode.

For example, the failure criterion during the product qualification and/or screening tests is typically stated as 1 uA [3]. However, the experimental data show that before the ESD device destruction by current filament a soft leakage current occurs in N-MOSFET because of gate oxide degradation. The typical range of soft leakage current in ESD N-MOSFET is 400 nA - 750 nA [5]. Hence, the I_{off} criterion can not be always used as a reliable indicator of actual ESD device destruction.

2.3 Failure temperature criterion

Destruction of an ESD device occurs at the threshold voltage, at which the maximum temperature reaches the melting point of silicon (1412 °C) [4] (typically in the gate-to-drain overlap region) or the melting point of metallization (660 °C for aluminum based metallization and 1034 °C for copper based metallization) [6]. With respect to this ESD robustness criterion, the current I_{1412} is the ESD threshold current (and $\approx$ 1500 x I_{1412} is the ESD threshold voltage) that will raise the temperature at the hottest spots to 1412 °C [4], the melting temperature of silicon. Depending on the failure mechanism, which is typical for a given ESD device, the I_{1034} or I_{660} current limit can be also used. The principal disadvantage of this criterion is the necessity to measure the temperature rise in the device during an ESD event on a nanosecond time scale. Several techniques were developed for the measurement of temperature dynamics under nanosecond high current stress, such as IR thermal interferometry [7], optical pyrometry [8] and backside laser interferometry [9].

3. ESD Devices under Investigation

For the analysis of impact of technology scaling on the second (thermal) breakdown in ESD devices, the following transistor structures were studied in this work: (1) 0.5 um LDD-N-MOSFET, (2) 0.18 um silicided N-MOSFET and (3) LDMOS transistor.

Electrothermal simulation has been introduced to general-purpose commercially available device simulation in the early 90-ties by TMA [10]. Validity of physical models such as mobility, impact ionization rates, etc. has been confirmed by numerous industrial applications and is generally believed to extend to approximately 600-700K. Since then a number of other TCAD companies also developed similar electrothermal models, including Silvaco, ISE and SEQUOIA. We used 2-D "SEQUOIA ESD" simulation software, which was developed by Sequoia Design Systems for the characterization of an ESD event [11]. This simulator has built-in device synthesis, mesh generation, device simulation, circuit-device mixed-mode simulation and lattice self-heating simulation models. The basic parameters of ESD devices used for simulations are given in Table 1 and Table 2. The analyzed transistor structures are presented in Fig. 2 (a,b,c). Note that LDMOS power transistors are commonly used for ESD protection of high voltage pins in Smart Power Technology [12]. Therefore we included an LDMOS transistor in our research.

4. Second Breakdown Investigation: Simulation Results

In order to investigate the mechanism of second breakdown phenomenon in ESD devices, we used mixed-mode transient simulations of analyzed transistors. Sequoia Mixed-Mode simulation tool allows to implement relatively small circuits (up to 10-20 devices) including previously designed transistor structures [11]. A schematic of the analyzed circuit with ESD device is shown in Fig. 3. In our simulations, the ESD device width was 50 um. The obtained simulation results are presented in Fig. 4. In this figure, the second snapback can be seen in I-V curves of LDMOS and 0.5 um LDD-N-MOSFET, which is typical for the second/thermal breakdown mode of transistors. However, 0.18 um silicided MOSFET does not show the second snapback behaviour in I-V curve. Note, that all of these devices reached the destructive temperature because of the self-heating effect during ESD event, as it is shown in Fig. 5. Hence, the deep

Table 1.
Basic parameters of MOSFETs used for simulations

ESD Device	0.5 um LDD-N-MOSFET	0.18 um N-MOSFET
Substrate doping, cm^{-3}	5×10^{15}	8×10^{16}
Source/Drain doping, cm^{-3}	8×10^{19}	1×10^{21}
LDD doping, S/D ext. doping, cm^{-3}	1×10^{18}	1×10^{21}
Channel doping, cm^{-3}	6×10^{16}	4×10^{17}
Gate oxide thickness, Å	120	40
Contact-to-Poly spacing, um	0.5	0.5
S/D silicide depth, um	-	0.035
Device width, um	1.0	1.0

Table 2.
Basic parameters of LDMOS used for simulations

ESD Device	LDMOS
Substrate doping, cm^{-3} (p - type)/depth, um	$1 \times 10^{19}/3.0$
Epi-doping, cm^{-3} (p - type)/depth, um	$4 \times 10^{15}/2.0$
S/D doping, cm^{-3} (n - type)/depth, um	$1 \times 10^{21}/0.2$
Channel doping, cm^{-3} (p - type)/depth, um	$1 \times 10^{19}/0.4$
Gate length, um	1.0
Gate oxide thickness, um	0.05
Device width, um	1.0

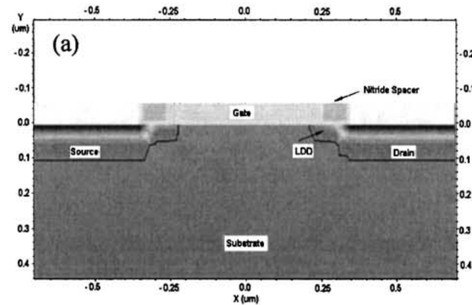

Fig. 2 (a). Cross-section of 0.5 um LDD-N-MOSFET.

submicron transistor can be destroyed without the second snapback behaviour in I-V curve and It_2 current can not be detected directly from these high current measurements. Recently published experimental data obtained for 0.1 um CMOS technology support our simulation results [13]. In the mentioned paper, the It_2 current is defined as the ESD stress level, when the leakage current (I_{off}) of the transistor has increased by 50% from the original value.

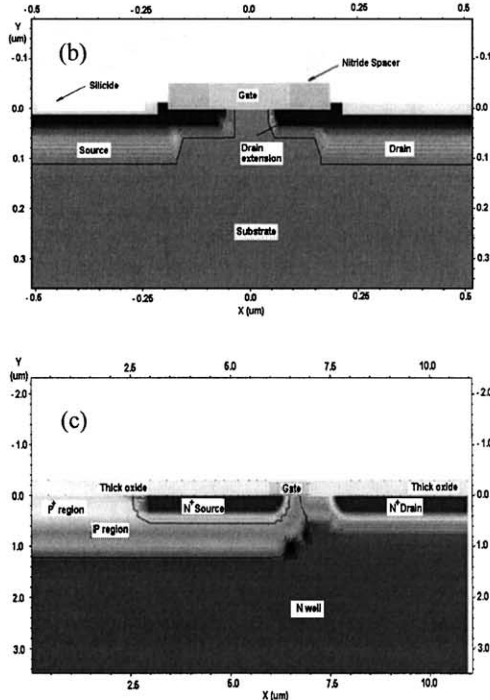

Fig. 2 (b,c). (b) Cross-section of 0.18 um silicided N-MOSFET, (c) Cross-section of Lateral Double Diffused MOSFET (LDMOS).

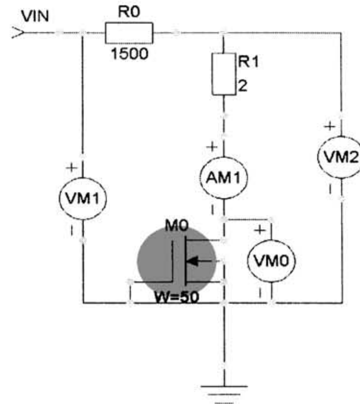

Fig. 3. The circuit schematic used for transient mixed-mode simulations.

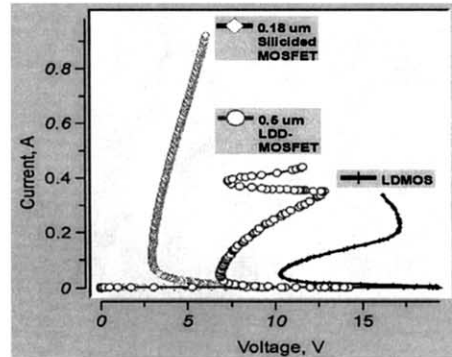

Fig. 4. I-V characteristics of analyzed ESD devices.

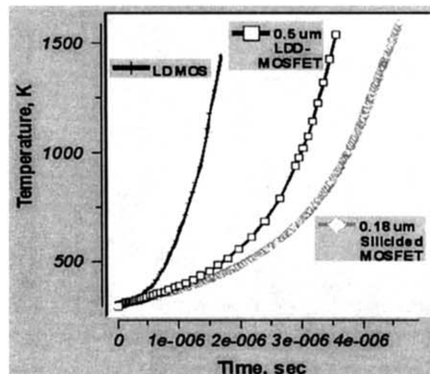

Fig. 5. Temperature rise in transistors during ESD

5. Analysis of Simulation Results

Second breakdown is a thermal phenomenon. There is the critical junction temperature at which second breakdown occurs. The theory of second breakdown suggests that there are two important thermal effects, which occur in the device. The first effect is the adiabatic heating, in which the entire input power is dissipated in the junction region; and the second effect is the thermal diffusion, in which thermal energy is distributed over a much larger region named as a hot spot [14].

Mobile carriers in a device near second breakdown are supplied by several physical effects, including thermal

generation, impact ionization and drift-diffusion conduction. The thermal generation component is a strong function of lattice temperature, and increases rapidly at higher temperatures. In the hot spot, lattice temperature reaches a value at which the total mobile carrier density becomes as large or higher as the donor or acceptor densities of the respective part of the space charge region.

To explain the existence of second breakdown phenomenon in 0.5 um LDD-NMOSFET and LDMOS transistor and its absence in 0.18 um silicided MOSFET, we investigated the increase of carrier density due to thermal generation in these devices during the ESD stress. The simulation results are depicted in Fig. 6 (a,b,c). In this figure, "Donors" is the doping profile in source (S), drain (D), LDD or S/D extension regions and "n" is the concentration of thermal generated charge carriers. In a gate-grounded ESD device, the hot spot is located in the gate-to-LDD/drain extension overlap region, because in this location the electric field has a maximum value. Hence, we should expect that in this region the total concentration of mobile carriers (dominated by thermal generation at high temperatures) would exceed the donor concentration when second/thermal breakdown occurs. From Fig. 6, we can conclude that the concentration of thermally generated charge carriers ("n") exceeds the donor concentration (background doping) in 0.5 um LDD-N-MOSFET and LDMOS transistor at the critical temperature in hot spot (approximately 1400K). However, the carrier concentration ("n") in 0.18 um silicided MOSFET does not exceed the donor concentration in a hot spot. This result explains the existence of second breakdown in the first two devices and its absence in the last one.

Under second breakdown conditions, the conductivity of hot spot region is determined by the concentration of intrinsic charges. At the critical temperature, the hot spot is the region of intrinsic semiconductor in ESD protection device. Hence, the quasi-Fermi level in this region must be located close to the middle of the band gap. Note that the position of quasi-Fermi level depends on the background doping in the drain region as well. We can estimate the dependency of quasi-Fermi level in the drain region of NMOS transistors on temperature and doping concentration using the Eq. (1) [15]. It is assumed that the donor concentration in drain region is much higher than the acceptors concentration ($N_d >> N_a$).

$$E_f - E_i \approx kT \ln \frac{N_d + \sqrt{N_d^2 + 4n_i^2}}{2n_i} \text{ [eV]} \qquad (1)$$

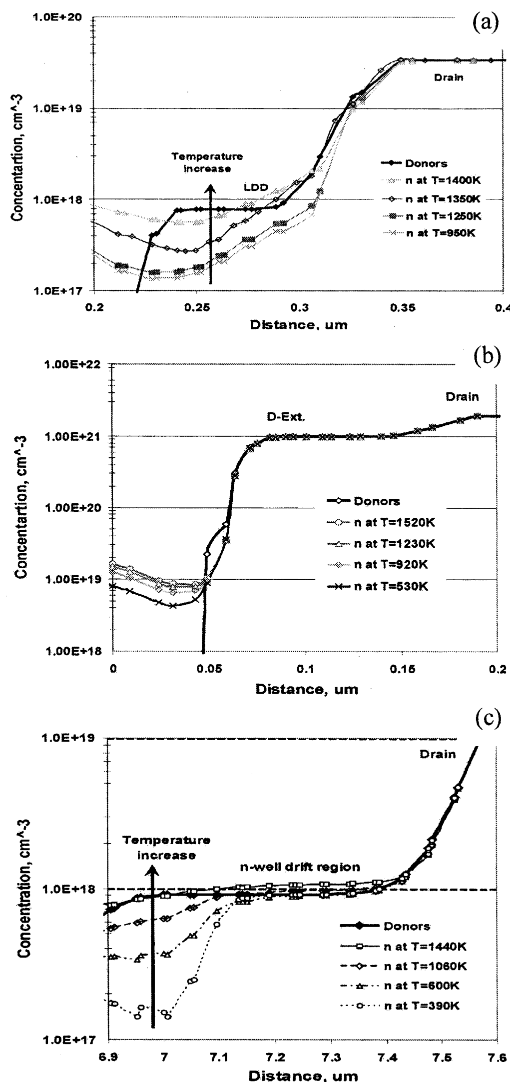

Fig. 6. Thermal generation due to the self-heating effect under ESD event in (a) 0.5 um LDD-N-MOSFET, (b) 0.18 um silicided N-MOSFET and (c) LDMOS transistor.

In Eq. 1, E_f is the quasi-Fermi level, E_i is the middle of semiconductor band gap, n_i is the concentration of intrinsic carriers, T is the absolute temperature and k is the Boltzmann constant.

The intrinsic carrier concentration is temperature dependent, as shown in Eq. (2) [16].

$$n_i(T) = 3.88 \cdot 10^{16} T^{\frac{3}{2}} e^{\frac{-7000}{T}} \text{ [cm}^{-3}\text{]} \qquad (2)$$

The calculation results of $(E_f - E_i)$ are presented in Fig. 7.

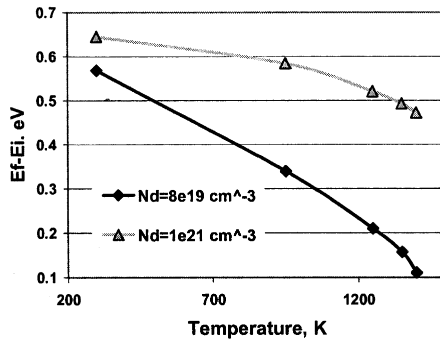

Fig. 7. Quasi-Fermi level vs. Temperature and Doping.

From this figure, we can conclude that the semiconductor in hot spot region becomes intrinsic ($(E_f - E_i)$ is close to zero) only if the silicon in this region is low doped. Because the LDD region in 0.5 um N-MOSFET and channel region in LDMOS transistor are relatively low doped, we can observe the second/thermal breakdown in high current I-V curves (Fig. 4). This conclusion agrees with simulation results depicted in Fig. 6 (a, c). Note that the 0.18 um silicided N-MOSFET has a very high doping in its drain extension region and this device can not reach intrinsic conductivity in hot spot region even at very high temperatures. Hence, it can not have the second snapback behaviour in high current I-V curve (Fig. 4) due to the thermal breakdown.

6. Conclusion

In this paper, we analyzed the impact of CMOS technology scaling on the second breakdown in sub-0.18 um ESD protection devices. Our simulation results and analytical calculations show that the second breakdown trigger current (It_2) criterion of ESD robustness is impractical for ultra deep submicron technologies, which have very high doping concentrations in drain and drain extension regions. For these technologies, ESD failure due to the self-heating effect occurs without a second snapback in high current I-V characteristics of ESD devices and It_2 current can not be properly detected from these measurements or simulations. Instead of It_2 current criterion, we propose to use the maximum failure temperature criterion. In accordance with this criterion, the protection device has higher ESD robustness if it reaches the critical temperature (silicon melting temperature or temperature of catastrophic degradation of metallization) at higher ESD stress and has higher ESD threshold current.

References

[1] A. Amerasekera and C. Duvvury. The impact of technology scaling on ESD robustness and protection circuit design. IEEE Trans. on Component, Packaging, and Manufacturing Tech., Part A, vol. 18, No 2, 1995, p. 314-320.

[2] A. Amerasekera and R.A. Chapman. Technology design for high current and ESD robustness in a deep submicron CMOS process. IEEE Electron Dev. Letters, vol. 15, No. 10, 1994, p. 383-385.

[3] G. Meneghesso, S. Santirosi, E. Novarini, C. Contiero, E. Zanoni. ESD robustness of smart-power protection structures evaluated by means of HBM and TLP tests. IEEE IRPS, 2000, p. 270-275.

[4] D.L. Lin. ESD sensitivity and VLSI technology trends: thermal breakdown and dielectric breakdown. EOS/ESD Symp., 1993, p. 73-81.

[5] T. Wadaıo Study of the soft leakage current induced ESD on LDD transistor. Microelectronics Reliability, vol. 36, No. 11/12, 1996, p. 1707-1710.

[6] S.H. Voldman. The impact of technology scaling on ESD robustness of aluminum and copper interconnects in advanced semiconductor technologies. IEEE Trans. on Component, Packaging, and Manufacturing Tech., Part C, vol. 21, No 4, 1998, p. 265-277.

[7] K. Murakami, K. Takita, K. Masuda. Measurement of lattice temperature during pulsed-laser annealing by time-dependent optical reflectivity. Jpn. J. Appl. Phys., vol.20, No. 12, 1981, p. L867-L870.

[8] I.P. Herman. Real-time optical thermometry during semiconductor processing. IEEE J. on Selected Topics in Quantum Electronics, vol. 1, No. 4, 1995, p. 1047-1053.

[9] C. Furbock, N. Seliger, D. Pogany, M. Litzenberger, E. Gornik, M. Stecher, H. Gossner, W. Werener. Backside laser probe characterization of thermal effects during high current stress in smart power ESD protection devices. Proc. of IEDM, 1998, p. 691-694.

[10] V. Axelrad and R. Klein. Electrothermal simulation of an IGBT. Int. Symp. on Power Semicond. Dev. and ICs, 1992, p.158-159.

[11] SEQUOIA Device Designer User's Guide, Sequoia Design Systems, http://www.sequoiadesignsystems.com

[12] M. Mergens, W. Wilkening, S. Mettler, H. Wolf, A. Sticker, W. Fichner. Analysis and compact modeling of lateral DMOS power devices under ESD stress conditions. EOS/ESD Symp., 1999. p. 1A.1.1-1A.1.9.

[13] A. Salman, R. Gautier, E. Wu, P. Riess, C. Putman, M. Muhammad, M. Woo, D. Ioannou. Electrostatic discharge induced oxide breakdown characterization in a 0.1 um CMOS technology," IEEE IRPS, 2002, p. 170-174.

[14] H. Melchior and M.J.O. Strutt. Secondary breakdown in transistors. Proceedings of the IEEE, April 1964, p.439-440.

[15] Y.S. Chung. Junction temperature induced thermal snapback breakdown of MOSFET device. IEEE Electron Device Lett., vol. 23, No. 10, 2002, p. 615-617.

[16] R.B. Adler, A.C. Smith, R.L. Longini. Introduction to semiconductor physics. John Wiley & Sons Inc., New York, 1964, p. 128.