Influence of Si-Nanocrystal Distribution in the Oxide on the Charging Behavior of MOS Structures

Y. Liu, T. P. Chen, C. Y. Ng, L. Ding, M. S. Tse, S. Fung, and Ampere A. Tseng

***Abstract*—In this brief, the electrical characteristics of MOS structures with specially designed distributions of Si nanocrystals (nc-Si) embedded in the oxides were investigated, and very different behaviors in the electrical characteristics are observed as a result of the difference in the nc-Si distribution. For the case of nc-Si being confined in a layer in the oxide close to the gate, charging and discharging the nc-Si lead to a shift in the flat-band voltage, and a small difference in the location of the peak concentration of nc-Si can yield a large difference in the memory programming characteristics. However, for the case of nc-Si distributed throughout the oxide with a high concentration, charging and discharging the nc-Si cause a modulation in the capacitance magnitude instead of the flat-band voltage shift. The difference between the two cases highlights the importance of the nc-Si distribution in the memory device applications of the nc-Si.**

***Index Terms*—Charging behavior, nanocrystal, Si-nanocrystal (nc-Si) distribution.**

I. Introduction

Si nanocrystals (nc-Si) embedded in the gate oxide of MOS structures have been studied extensively due to their applications in memory devices [1]–[14]. For the nonvolatile memory application, the nc-Si is normally confined in a narrow layer in the gate dielectric near the substrate [1]–[12]. Charging/discharging in the nc-Si leads to the flat-band voltage shifts in the MOS capacitance–voltage $(C-V)$ characteristics. There are few detailed studies so far on the electrical characteristics of MOS structures with different depth distributions of the nc-Si in the gate oxide. In this paper, the nc-Si with different depth distributions were embedded in the gate oxide by an ion implantation technique. $C-V$ and time-domain-capacitance measurements were conducted to study the electrical characteristics of the MOS structures. Indeed, it is observed that the charge trapping in the nc-Si with different depth distributions in the gate oxide leads to different behaviors in the MOS electrical characteristics. It is found that the charge trapping/detrapping in the nc-Si confined in the gate oxide leads to flat-band-voltage shifts, while charge trapping/distrapping in the nc-Si distributed throughout the gate oxide causes a modulation in the capacitance magnitude. The differences observed in this work highlight the importance of the nc-Si depth distribution in the memory device applications of nc-Si.

II. Sample Fabrication

In the fabrication of the MOS structures with different nc-Si distributions in the oxide, first a 30-nm SiO_2 film was thermally grown on p-type (100) Si wafers in dry oxygen at 950 °C. The Si^+ ions were implanted to the SiO_2 films at 1, 2, or 14 keV. For the 1 and 2 keV implantations, the implant dose is 8×10^{16} cm^{-2}, while for the 14 keV implantation the dose is 3×10^{16} cm^{-2}. Thermal annealing was carried out at 1000 °C in N_2 ambient for 1 h to induce nc-Si formation. Most of the nc-Si (90%) was distributed in a narrow region in the oxide close to the gate (i.e., 1–7 and 2–12 nm underneath the oxide surface for the 1 and 2 keV-implanted samples, respectively). The oxide layer for the 14 keV-implanted sample was thinned at a rate of 10 nm/min in a HF solution of concentration of 50:1 after the ion implantation. About 17-nm SiO_2 layer was removed such that the nc-Si was distributed throughout the gate oxide. Secondary ion mass spectroscopy measurement has confirmed that the nc-Si was distributed throughout the oxide in this sample (i.e., the 14 keV-implanted sample after the oxide thinning). For all samples, a 20-nm aluminum layer was then deposited to form the gate electrodes. The wafer backside was coated with a layer of aluminum with the thickness of about 1 μm after removing the backside oxide. Finally, a metal alloying process was conducted at 425 °C in the N_2 atmosphere to form ohmic contacts. The 1 and 2 keV-implanted samples are used to study the charge trapping in nc-Si confined in a layer close to the gate with different depths in the oxide. The 14 keV-implanted sample after oxide thinning is used for the study of the charge trapping in nc-Si distributed throughout the gate oxide. The X-ray diffraction measurement shows that the mean size of the nc-Si is 3–4 nm [14]. The experiment of charging/discharging the nc-Si by applying a constant voltage to the MOS structures was performed with a Keithley 4200 semiconductor characterization system, and $C-V$ and capacitance–time $(C-T)$ measurements were carried out with an HP4284A LCR meter at the frequency of 1 MHz.

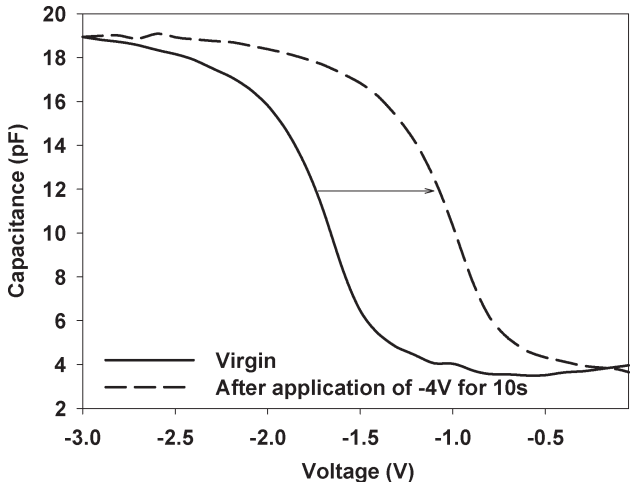

Fig. 1. Effect of charging the nc-Si by electron injection from the gate on the $C-V$ characteristics of an MOS structure with the nc-Si confined in a narrow region in the gate oxide close to the gate. The ion implantation energy is 1 keV.

III. Result and Discussions

Fig. 1 shows the $C-V$ characteristics of an MOS structure with nanocrystals confined in a narrow layer in the gate oxide close to the gate of the 1 keV-implanted sample after a negative voltage (-4 V) was applied to the gate for 10 s. The negative gate voltage caused electron injection from the gate to the nc-Si. Charge trapping in the nc-Si leads to a flat-band-voltage shift of about 0.7 V. As can be seen in Fig. 1, the flat-band voltage is shifted to a more positive voltage, showing

Manuscript received September 16, 2005. This work was supported by the Academic Research Fund from the Ministry of Education, Singapore, under ARC Project RG 1/04. The review of this brief was arranged by Editor R. Shrivastava.

Y. Liu, T. P. Chen, C. Y. Ng, L. Ding, and M. S. Tse are with the School of Electrical and Electronic Engineering, Nanyang Technological University, 639728 Singapore (e-mail: echentp@ntu.edu.sg).

S. Fung is with the Department of Physics, The University of Hong Kong, Hong Kong.

A. A. Tseng is with the Department of Mechanical and Aerospace Engineering, Arizona State University, Tempe, AZ 85287 USA.

Digital Object Identifier 10.1109/TED.2006.870528

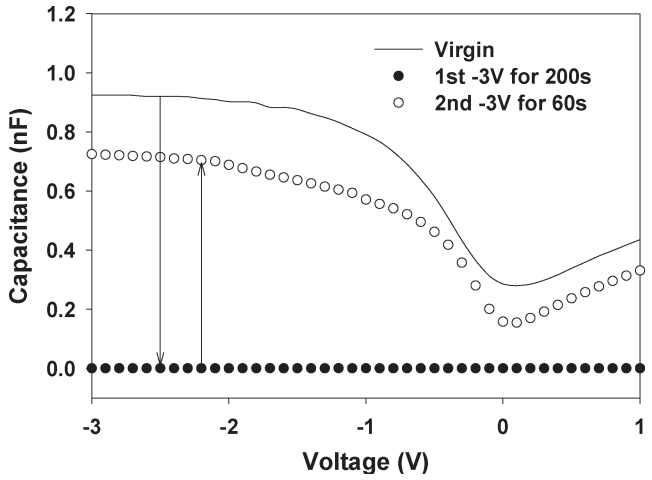

Fig. 2. Effect of charging and discharging the nc-Si on the MOS capacitance for the case of the nc-Si being distributed throughout the oxide. Note that both the charging and discharging occur under the same gate voltage.

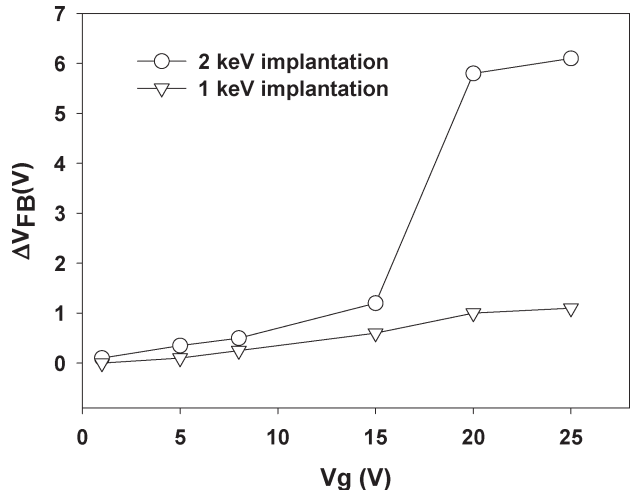

Fig. 3. Comparison of the dependence of the flat-band voltage shift ($\Delta V_{\rm FB}$) on the charging voltage between the 1 and 2 keV ion implantations for the case of the nc-Si being confined in a narrow region in the oxide. The charging time is 1 s.

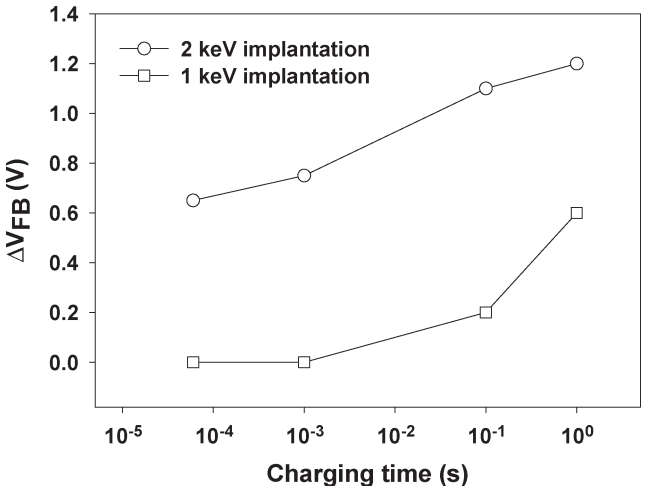

Fig. 4. Comparison of the dependence of the flat-band voltage shift ($\Delta V_{\rm FB}$) on the charging time between the 1 and 2 keV ion implantations for the case of the nc-Si being confined in a narrow region in the oxide. The charging voltage is positive 15 V.

electron trapping in the nc-Si due to the electron injection from the gate under the negative gate voltage. The situation here is similar to the flat-band-voltage shift caused by electron trapping in the nc-Si as a result of electron injection from the Si substrate in a conventional nc-Si memory device [2], [3].

Fig. 2 shows the $C-V$ characteristics of the MOS structure with nc-Si distributed throughout the gate oxide (i.e., the 14 keV-implanted sample after oxide thinning) obtained by sweeping the voltage from -3 to 1 V. As shown in Fig. 2, the application of a -3 V for 200 s reduces the MOS capacitance to an extremely low value ($<$ 700 fF) that is close to the measurement limit of our $C-V$ measurement system; however, after a second application of -3 V for 60 s, the capacitance is recovered but is lower than the virgin capacitance. The reduction in the capacitance is due to the charging up in the nc-Si. The application of -3 V for 200 s leads to the charging up of most of the nc-Si. The charged nc-Si does not respond to the small ac signal (15 mV) as the Coulomb energy is $\sim$ 90 mV for a nc-Si with a size of 4 nm in this paper. Therefore, the capacitance of the nc-Si decreases, and thus the total MOS capacitance is reduced. On the other hand, the trapped charge in the nc-Si is released by the second application of -3 V for 60 s, leading to the recovery of the capacitance, as shown in Fig. 2. However, the capacitance after the recovery is still lower than that of the virgin case. The partial recovery of the capacitance indicates that not all the trapped charges due to the first application of -3 V are released during the second application of the voltage. Fig. 2 clearly shows that the charge trapping in the nc-Si distributed throughout the gate oxide results in a magnitude modulation in the MOS capacitance. This situation is very different from the flat-band-voltage shift shown in Fig. 1 for the charge trapping in the nc-Si confined in a layer. Therefore, it is obvious that the different nc-Si distributions lead to very different charging effects on the $C-V$ characteristics.

In the case of nc-Si being confined in a layer in the oxide, even a small difference in the depth location of the peak concentration of nc-Si, which can be realized by varying the Si^+ implantation energy, can yield a significant difference in the memory programming characteristics. This situation is clearly demonstrated by the dependence of the flat-band-voltage shift ($\Delta V_{\rm FB}$) on the charging voltage (Fig. 3) and the charging time (Fig. 4) for a small difference in the nc-Si depth distribution between the 1 and 2 keV ion implantations. As shown in Fig. 3, for both ion implantations, no flat-band-voltage shift is observed when the magnitude of the gate voltage is less than 1 V, and the flat-band-voltage shift increases with the magnitude of the gate voltage. However, the dependence of the flat-band-voltage shift on the charging voltage for the 2 keV ion implantation is much stronger than that for the 1 keV ion implantation, and the 2 keV ion implantation has a much larger $\Delta V_{\rm FB}$ than the 1 keV ion implantation when the magnitude of the charging voltage is larger than 15 V. On the other hand, the two ion implantations also show a large difference in the dependence of $\Delta V_{\rm FB}$ on the charging time, as shown in Fig. 4. The charging experiment of this figure is carried out with the gate voltage of 15 V. For the 1 keV ion implantation, a charging time of shorter than 1 ms does not lead to an obvious flat-band voltage shift, and 1 s of charging can produce a $\Delta V_{\rm FB}$ of only 0.6 V. In contrast, for the 2 keV ion implantation, 60 μs of charging has led to a $\Delta V_{\rm FB}$ of 0.65 V, and the $\Delta V_{\rm FB}$ increases with the charging time. The difference in the memory programming characteristics shown in Figs. 3 and 4 between the 1 and 2 keV ion implantations highlights the importance of the nc-Si distribution in a memory device performance. The difference is actually related to the difference in the distribution of the electron trapping in the gate oxide as a result of the difference in the nc-Si distribution. If the electron trapping is closer to the gate-oxide/Si

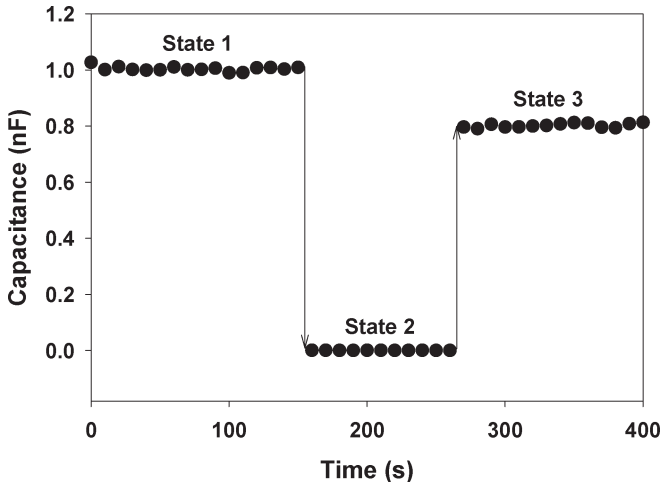

Fig. 5. Capacitance modulations in the $C-T$ measurement under a constant gate voltage of -3 V for the case of the nc-Si being distributed throughout the oxide.

substrate interface, then a larger $\Delta V_{\rm FB}$ will be produced. The 2 keV ion implantation leads to an nc-Si distribution deeper towards the interface, and thus the charge trapping is more effective in shifting the flat-band voltage.

Fig. 5 shows the time-domain-capacitance characteristic for the case of nc-Si distributed throughout the oxide. As can be seen in this figure, the MOS structure shows capacitance magnitude modulations, which is very different from the case of the nc-Si being confined in a layer in the gate oxide. Such a $C-T$ measurement is carried out under a constant gate voltage of -3 V. After about 150 s of the application of the voltage, the MOS capacitance was reduced drastically from the initial value of $\sim$ 1.0 nF to $\sim$ 600 fF. In other words, the electrical state of the MOS device was switched from a high-capacitance state (State 1) to a low-capacitance state (State 2). As discussed early, the low capacitance corresponds to a charged state of the nc-Si, while the high capacitance corresponds to a discharged state of the nc-Si. Therefore, State 1 represents the discharged state and State 2 represents the charged state. State 2 can be maintained for $\sim$ 110 s. Then, the MOS device was switched to State 3 with a capacitance of $\sim$ 0.8 nF, which is much larger than that of State 2 but lower than that of State 1. This indicates that not all the trapped charges associated with the State 2 have been released. Thus, only a partial recovery of MOS capacitance can be achieved. If the time scale is extended, switching between different states, which is a random event, can be observed continuously.

The observation of the modulation in MOS capacitance in the case that nc-Si is distributed throughout the gate oxide can be explained in the following. The total capacitance C of the MOS structure can be expressed as

$$C = \frac{(C_{\rm nc} + C'_{\rm ox})\, C_D}{(C_{\rm nc} + C'_{\rm ox}) + C_D} \tag{1}$$

where $C_{\rm nc'}$, $C'_{\rm ox}$, and C_D are the capacitance of nanocrystals, the remaining gate-oxide capacitor and the Si-depletion layer, respectively. As C_D is very large under accumulation at -3 V, the total capacitance is approximately equal to $(C_{\rm nc} + C'_{\rm ox})$. $C'_{\rm ox}$ is much smaller than the conventional gate-oxide capacitance $(C_{\rm ox})$ without the nc-Si because of the existence of a huge amount of implanted Si atoms in the oxide. Actually $C'_{\rm ox}$ is found to be less than $\sim$ 700 fF. Furthermore, as there is a high concentration of nc-Si distributed throughout the gate oxide, the nanocrystals easily respond to the small ac signal in the capacitance measurement if they are uncharged or neutral (i.e., no charges trapped in them). The virgin $C_{\rm nc}$ (i.e., the nc-Si capacitance before charge trapping) should be very large due to the existence of a high concentration of nc-Si. Therefore, the MOS capacitance C is approximately equal to $C_{\rm nc}$, and it will be very large when most of the nc-Si are uncharged. Obviously, the capacitance reduction can be explained in terms of the charge trapping in the nc-Si. With charge trapping in the nanocrystals, the charged nc-Si does not respond to the small ac signal (15 mV) as the Coulomb energy is $\sim$ 90 mV for a nc-Si with a size of 4 nm in this brief. The nanocrystal capacitance $C_{\rm nc}$ should decrease as the number of uncharged nanocrystals responding to the small ac signal in the capacitance measurement is reduced. If all the chargeable nanocrystals are charged up, then $C_{\rm nc}$ should approach zero. Thus, the MOS capacitance is very small as it is now determined by the remaining gate-oxide capacitance $C'_{\rm ox}(< 700$ fF). On the other hand, if the charges trapped in the nc-Si are released, then $C_{\rm nc}$ should increase leading to an increase in the MOS capacitance. This explains the experimental results shown in Figs. 2 and 5.

IV. Conclusion

In conclusion, MOS structures with nc-Si embedded in the oxide were synthesized by the ion implantation technique. It is found that the nc-Si distribution has a strong influence on the MOS $C-V$ characteristics. For the case of nc-Si confined in a layer in the oxide, charging and discharging the nc-Si lead to a shift in the flat-band voltage, and a small difference in the location of the peak concentration of the nc-Si can yield a significant difference in the memory programming characteristics. However, for the case of the nc-Si distributed throughout the oxide with a high concentration, charging and discharging the nc-Si cause the modulation in the capacitance magnitude instead of the flat-band-voltage shift. The difference between the two cases highlights the importance of the nc-Si distribution in the memory device applications of the nc-Si.

References

[1] Y. H. Kwon, C. J. Park, W. C. Lee, D. J. Fu, Y. Shon, T. W. Kang, C. Y. Hong, H. Y. Cho, and K. L. Wang, "Memory effects related to deep levels in metal-oxide-semiconductor structure with nanocrystalline Si," *Appl. Phys. Lett.*, vol. 80, no. 14, pp. 2502–2504, Apr. 2002.

[2] S. Tiwari, F. Rana, H. Hanafi, A. Hartstein, E. F. Crabbé, and K. Chan, "A silicon nanocrystals based memory," *Appl. Phys. Lett.*, vol. 68, no. 10, pp. 1377–1379, Mar. 1996.

[3] S. Tiwari, F. Rana, K. Chan, L. Shi, and H. Hanafi, "Single charge and confinement effects in nano-crystal memories," *Appl. Phys. Lett.*, vol. 69, no. 9, pp. 1232–1234, Aug. 1996.

[4] K. Han, I. Kim, and H. Shin, "Characteristics of P-channel Si nano-crystal memory," *IEEE Trans. Electron Devices*, vol. 48, no. 5, pp. 874–879, May 2001.

[5] P. Normand, E. Kapetanakis, D. Tsoukalas, G. Kamoulakos, K. Beltsios, J. Van Den Berg, and S. Zhang, "MOS memory devices based on silicon nanocrystal arrays fabricated by very low energy ion implantation," *Mat. Sci. and Eng. C*, vol. 15, no. 34, pp. 145–147, Aug. 2001.

[6] N. Takahashi, H. Ishikuro, and T. Hiramoto, "Control of coulomb blockade oscillations in silicon single electron transistors using silicon nanocrystal floating gates," *Appl. Phys. Lett.*, vol. 76, no. 2, pp. 209–211, Jan. 2000.

[7] R. Ohba, N. Sugiyama, J. Koga, K. Uchida, and A. Toriumi, "Influence of channel depletion on the carrier charging characteristics in silicon nanocrystal floating gate memory," *Jpn. J. Appl. Phys.*, vol. 39, no. 3A, pp. 989–993, Mar. 2000.

[8] I. Kim, S. Han, K. Han, J. Lee, and H. Shin, "Silicon nanocrystal memory cell with room-temperature single electron effects," *Jpn. J. Appl. Phys.*, vol. 40, no. 2A, pp. 447–451, Feb. 2001.

[9] G. Ammendola, V. Ancarani, V. Triolo, M. Bileci, D. Corso, I. Crupi, L. Perniola, C. Gerardi, S. Lombardo, and B. DeSalvo, "Nanocrystal memories for FLASH device applications," *Solid State Electron.*, vol. 48, no. 9, pp. 1483–1488, Sep. 2004.

[10] P. Dimitrakis, E. Kapetanakis, D. Tsoukalas, D. Skarlatos, C. Bonafos, G. Ben Asssayag, A. Claverie, M. Perego, M. Fanciulli, V. Soncini, R. Sotgiu, A. Agarwal, M. Ameen, C. Sohl, and P. Normand, "Silicon nanocrystal memory devices obtained by ultra-low-energy ion-beam synthesis," *Solid State Electron.*, vol. 48, no. 9, pp. 1511–1517, Sep. 2004.
[11] J. Carreras, B. Garrido, and J. R. Morante, "Improved charge injection in Si nanocrystal non-volatile memories," *Microelectron. Reliab.*, vol. 45, no. 5/6, pp. 899–902, May/Jun. 2005.
[12] R. A. Rao, R. F. Steimle, M. Sadd, C. T. Swift, B. Hradsky, S. Straub, T. Merchant, M. Stoker, S. G. H. Anderson, M. Rossow, J. Yater, B. Acred, K. Harber, E. J. Prinz, B. E. White, Jr., and R. Muralidhar, "Silicon nanocrystal based memory devices for NVM and DRAM applications," *Solid State Electron.*, vol. 48, no. 9, pp. 1463–1473, Sep. 2004.
[13] Y. Liu, T. P. Chen, M. S. Tse, H. C. Ho, and K. H. Lee, "Charging effects of silicon nanocrystals in gate oxide near gate on MOS capacitance," *Electron. Lett.*, vol. 39, no. 16, pp. 1164–1166, Aug. 2003.
[14] Y. Liu, T. P. Chen, C. Y. Ng, M. S. Tse, S. Fung, Y. C. Liu, S. Li, and P. Zhao, "Charging effect on electrical characteristics of MOS structures with Si nanocrystal distribution in gate oxide," *Electrochem. Solid-State Lett.*, vol. 7, no. 7, pp. G134–G137, 2004.

From Wafer-Level Gate-Oxide Reliability Towards ESD Failures in Advanced CMOS Technologies

A. Kerber, M. Röhner, C. Wallace, L. O'Riain, and M. Kerber

***Abstract*—The impact of a short-term electrical overstress on dielectric reliability is thoroughly investigated using a dedicated experimental equipment for measurement times in the millisecond and microsecond range. Based on significant statistics, it is confirmed that the dielectric stress-induced damage is cumulative in nature. In addition, the voltage acceleration is shown to follow the power-law model towards the time range of electrostatic discharge and furthermore the breakdown statistics remain unchanged. These results justify the assessment of a short-term electrical overstress in advanced CMOS technologies by using a conventional reliability prediction methodology.**

***Index Terms*—Breakdown statistics, dielectric reliability, electrostatic discharge (ESD), voltage acceleration.**

I. Introduction

The gate-oxide reliability assessment is primarily focused on the prediction of the end of lifetime failure probability at operation conditions. The two important parameters for the evaluation are the voltage acceleration and the spread in the time to breakdown (TBD). These parameters are extensively discussed in the literature for stress times from 1 to 10 000 s and less frequently up to 10^7 s. In the nanosecond to microsecond timeframe, only a few reports are available [1], [2] and their statistical basis is not sufficient to validate reliability models in the microsecond regime. The short-stress times, however, are of particular interest for the prediction of the failure probability due to a short-term electrical overstress or electrostatic discharge (ESD), caused by the penetration of voltage spikes into logic cores of multiple-gate-oxide processes.

Manuscript received September 22, 2005; revised January 6, 2006. The review of this brief was arranged by Editor G. Groeseneken.

A. Kerber, M. Röhner, and M. Kerber are with the Corporate Reliability Methodology Department, Infineon Technologies, Munich 81739, Germany (e-mail: Andreas.Kerber@infineon.com).

C. Wallace and L. O'Riain are with the Infineon Technologies, Dresden 01099, Germany.

Digital Object Identifier 10.1109/TED.2006.870517

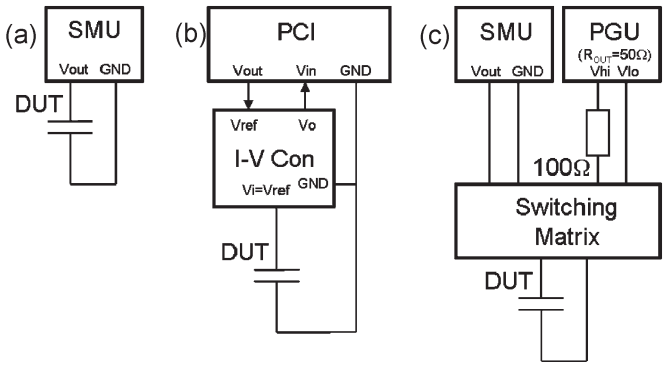

Fig. 1. (a) Schematic drawing of the reliability test setup including a conventional SMU, (b) PCI card with an $I-V$ converter, and (c) PGU stress unit combined with an SMU sense unit. For the PGU–SMU setup, the alternating connections were made through a switching matrix.

II. Experimental

n-channel MOSFETs were fabricated using a state-of-the-art CMOS process with a gate-oxide thickness of 1.5, 2.2 and 5.2 nm and test structures down to 0.12 μm^2. The gate-oxide reliability tests were carried out at 140 °C using several experimental setups optimized to study the dielectric breakdown from conventional wafer-level test times down to microsecond breakdown times. In particular, a conventional source measurement unit (SMU) setup, a fast measurement setup based on a peripheral computer interconnect (PCI) card [3] and a combined setup of a pulse generator unit (PGU) for stressing and an SMU for sensing were used (Fig. 1). For the PGU–SMU setup, the alternating connection of the device under test (DUT) to the stress and sense unit was provided by a switching matrix. All experiments were carried out in inversion mode and the stress bias was applied to the gate of the n-channel MOSFET while source, drain, and substrate were connected to a ground potential. For breakdown times from 1 to 10 000 s, the conventional SMU setup was used. In order to record breakdown times in the millisecond regime, a PCI-card setup was developed where the stress bias ($V_{\rm out}$) is set by a digital-to-analog converter to the reference input ($V_{\rm ref}$) of an external current–voltage ($I-V$) converter. The stress voltage and current for the DUT are provided by the input terminal of the converter, which is at the same potential as the reference terminal ($V_i = V_{\rm ref}$). The output voltage (V_o) of the $I-V$ converter is proportional to the logarithm of the current flowing at the input terminal and recorded by the analog-to-digital converter of the PCI card. A time resolution of ~ 20 μs is achieved with this setup [3]. For microsecond breakdown times the stress pulse was provided by a PGU and the intermediated monitor current recorded with the SMU. In order to prevent fusing of the metal connections during the breakdown event, a 100-Ω resistor was put in series to the standard 50-Ω output resistance of the PGU.

The TBD distributions based on a constant voltage stress (CVS) are provided in Fig. 2 for the PCI-card setup and the SMU setup with and without an intermediate monitoring step. As can be seen, the TBD distributions obtained with the conventional SMU and the PCI-card setup are neither dependent on the measurement setup nor on the stress procedure, which validates the use of the PCI-card setup for dielectric-reliability tests.

III. Results and Discussion

At first, the voltage acceleration and breakdown statistics were determined for the 1.5-nm gate oxide based on current time traces