Conduction mechanism under quasibreakdown of ultrathin gate oxide

Y. D. He, Hao Guan, M. F. Li, Byung Jin Cho, and Zhong Dong

Citation: Appl. Phys. Lett. **75**, 2432 (1999); doi: 10.1063/1.125038
View online: http://dx.doi.org/10.1063/1.125038
View Table of Contents: http://apl.aip.org/resource/1/APPLAB/v75/i16
Published by the American Institute of Physics.

Conduction mechanism under quasibreakdown of ultrathin gate oxide

Y. D. He, Hao Guan, M. F. Li,[a] and Byung Jin Cho
Department of Electrical Engineering, CICFAR, National University of Singapore, Singapore 119260

Zhong Dong
Chartered Semiconductor Manufacturing, Ltd., Singapore 738406

(Received 18 February 1999; accepted for publication 19 August 1999)

The conduction mechanism under quasibreakdown of ultrathin gate oxide has been studied systematically in both *n* and *p* metal–oxide–semiconductor field effect transistors (MOSFETs) with a 3.7 nm gate oxide. The carrier separation experiment is conducted to investigate the evolutions of gate, source/drain, and substrate currents before and after quasibreakdown. It is shown that after quasibreakdown, the substrate current and the source-drain current versus the gate voltage curves are surprisingly analogous to those curves observed in fresh MOSFET with a gate oxide of direct tunneling thickness. This strongly supports the quasibreakdown model based on the local physically damaged region by which the effective oxide thickness is reduced. When direct tunnelings of conduction band electrons, valence band electrons and holes through the effectively thinned gate oxide are taken into account, the major experimental observations in the quasibreakdown can be explained in a unified way. © *1999 American Institute of Physics.* [S0003-6951(99)03542-1]

With the development of ultralarge scale integrated circuits, the gate oxide thickness has been scaled down to less than 5 nm, and the integrity of the ultrathin gate oxide has become one of the most important reliability issues. The idea that the breakdown of oxides starts at a locally damaged region near the Si/SiO_2 interface has been discussed very extensively for thicker oxides.[1] Compared to the thicker gate oxide, a new anomalous degradation and failure mode had been reported over the past few years, referred to as quasibreakdown (QB),[2] or *B*-mode SILC,[3,5] or soft breakdown.[4] The QB is characterized by high leakage current at low field, the gate signal fluctuations, and the increase of random telegraph noise or $1/f$ noise.[2–8] Several models have been proposed to explain the conduction mechanism under QB, such as the physical damage model near the anode interface,[2] the multiple trap assisted tunneling model,[3] the variable range hopping conduction model,[5] and so on. Although there is a consensus that the QB is a localized effect in a very small area,[2–8] there is no generally accepted model of the QB conduction mechanism that gives an overall explanation of major experimental observations under the QB. In this letter, we investigate the QB conduction mechanism systematically in both *n* and *p* metal–oxide–semiconductor field effect transistors (MOSFETs) by constant current stress and carrier separation measurements.[9] All experimental results strongly support the quasibreakdown model based on the local physically damaged region (LPDR) by which the effective oxide thickness is reduced at that local area.[2]

The MOSFETs used in our experiments are dual gate complementary MOS devices, with p^+-polycrystalline silicon (polysilicon) gate for *p*-MOSFETs and n^+-polysilicon gate for *n*-MOSFETs, fabricated by using 0.18 μm technology with a 37 Å gate oxide. The transistor size under test was 50 μm$\times$0.5 μm in channel width and length. Constant 50 nA current stresses (200 mA/cm^2 in current density) for both polarities, i.e., gate and substrate injections, were applied to *n*- and *p*-MOSFETs grounding source/drain and substrate. Static current–voltage ($I-V$) characteristics were monitored using an HP4156A model semiconductor parameter analyzer. In order to understand the conduction mechanism, the carrier separation experiment[9] was conducted to measure the gate current I_g, the sum of the source and drain currents $I_{s/d}$, and the substrate current I_{sub} separately before and after QB. All currents flowing into the device are taken as positive and $I_g+I_{s/d}+I_{sub}=0$ if there is no other leakage channel.

Figures 1(a) and 1(b) show the variations of gate voltage V_g and $I_{s/d}$ with stress time for a *p*-MOSFET under constant current stress with -50 nA gate current. Figures 1(c) and 1(d) show similar curves of V_g and I_{sub} with stress time for an *n*-MOSFET under constant current stress with a $+50$ nA gate current. In Fig. 2, each figure shows measured I_g, $I_{s/d}$,

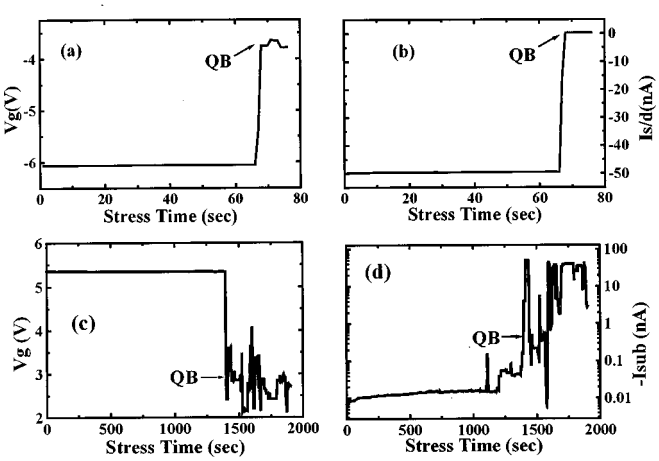

FIG. 1. (a) and (b) The evolution of V_g and $I_{s/d}$ with stress time for a *p*-MOSFET under gate injection mode constant current stress with a stress current of -50 nA; (c) and (d) V_g and I_{sub} vs time for an *n*-MOSFET under substrate injection mode constant current stress with a stress current of $+50$ nA.

[a]Electronic mail: elelimf@nus.edu.sg

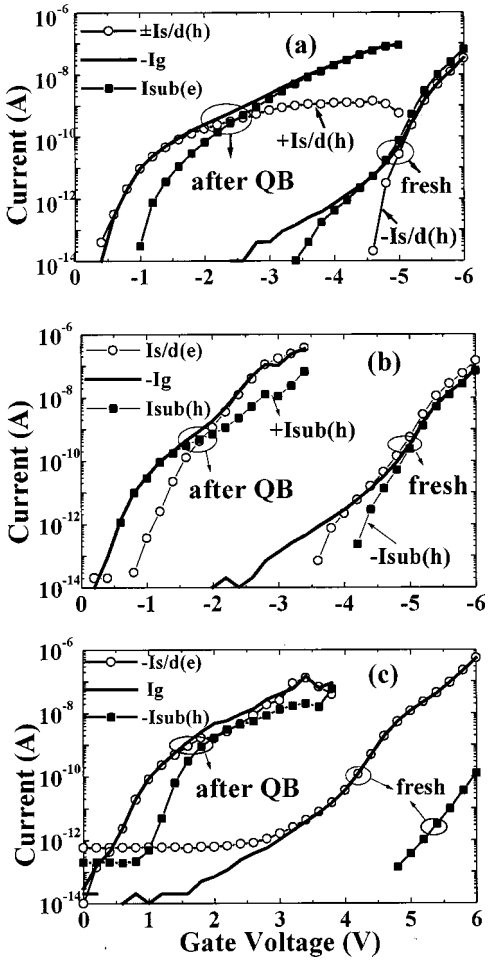

FIG. 2. Carrier separation $I-V$ curves for MOSFETs with a 3.7 nm gate oxide: (a) p-MOSFET, fresh and after QB under gate injection mode stress; (b) n-MOSFET, fresh and after QB under gate injection mode stress, (c) n-MOSFET, fresh and after QB under substrate injection mode stress.

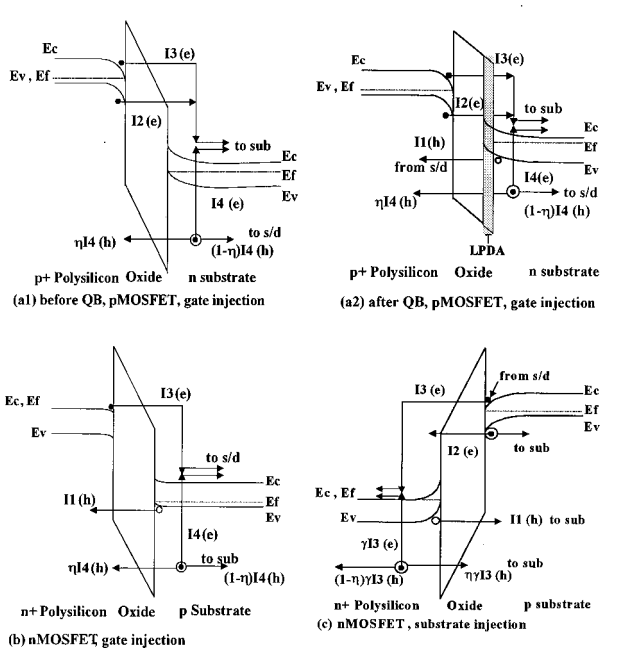

FIG. 3. The energy band diagrams of p-MOSFET under gate injection mode (a1)(a2), and of n-MOSFET under gate injection mode (b) and substrate injection mode (c). (a1) Before QB, oxide is thick, there are only tunneling electron currents I_2 and I_3 and impact ionization induced hole current $(1-\eta)I_4$, here $I_4=\gamma I_3+\gamma' I_2$. γ and γ' are impact ionization quantum yields of I_3 electrons and I_2 electrons respectively. (a2) After QB, the effective oxide thickness is reduced by LPDR. There are electron DT currents I_3 and I_2, hole DT current I_1 and impact ionization induced hole current $(1-\eta)I_4$. LPDR is indicated in (a2) explicitly, but not in (b) and (c).

and I_{sub} vs V_g curves for the fresh device and for the device after the gate oxide QB. The corresponding band diagrams of the devices are illustrated in Fig. 3. Figures 1(a), 1(b), 2(a), and 3(a1), 3(a2) show the case of a p-MOSFET under the gate injection mode. In this case, I_{sub} measures the electron current and $I_{\text{s/d}}$ measures the hole current. The mechanism of three current components of fresh device in Fig. 2(a) is indicated in Fig. 3(a1). From Fig. 1(b) we can see that the magnitude of $I_{\text{s/d}}$ before QB is nearly identical to the gate current ($I_{\text{s/d}}\approx I_g=-50\,\text{nA}$) and the polarity is negative, which means that holes are coming out through the source/drain. This indicates that the electrons of current I_3 and I_2, when tunneled to the substrate, produce electron–hole pairs by impact ionization with a quantum yield $\gamma\sim 1$.[10] After QB, however, the sign of $I_{\text{s/d}}$ is changed to positive. The three curves after QB in Fig. 2(a) show surprising analogy with three curves of fresh p^+-polysilicon gate p-MOSFET with a 2.5 nm gate oxide which is in the direct tunneling (DT) regime reported by Shi *et al.*[11] Our curves after QB in Fig. 2(a) are almost identical with the curves in Fig. 7 in Ref. 11, with slightly different voltage and current scaling, which can be ascribed to the different oxide thickness and area. This analogy gives a very strong support of the QB model proposed in Ref. 2 that after QB, a LPDR region at the Si/SiO_2 interface reduces the effective SiO_2 thickness at that local area and gives rise to DT current. Using this model, we can explain all of our $I-V$ curves in a unified way and also clarify some muddles which have never been understood. In Fig. 2(a), in the low voltage regime ($|V_g|<2$ V), $I_{\text{s/d}}$ is positive and $I_{\text{s/d}}\approx -I_g$, implying that hole DT from the substrate LPDR to the gate [I_1 in Fig. 3(a2)] is the dominating process. When $|V_g|>2.5$ V, I_{sub} starts to supersede $I_{\text{s/d}}$ and becomes the dominant component of I_g. I_{sub} consists of electrons tunneling from the gate (I_3+I_2), and impact ionization induced electron current (I_4). These findings are consistent with DT of electrons and holes in fresh p-MOSFET with a 2.5 nm gate oxide in Ref. 11. Additionally, as shown in Fig. 2(a), when $|V_g|$ increases from 2 to 5 V, the increase of $I_{\text{s/d}}$ tends to slow down and even turns to decrease, showing a downward U turn of the $I_{\text{s/d}}$ curve. In this region when $|V_g|$ increases, the number and the quantum yield γ of injected electrons increase,[10] therefore more electron–hole pairs [I_4 in Fig. 3(a2)] are generated in the substrate region. When the major portion of the generated holes diffuse out of the source/drain ([$1-\eta$] I_4 in Fig. 3(a2), here η is the probability of the impact-ionized hot hole crossing over the oxide barrier and reaching the opposite side), the corresponding source/drain current is negative, as illustrated in Fig. 3(a2). This negative source/drain current will compensate the positive source/drain current I_1, causing a downward U turn of $I_{\text{s/d}}$ curve and even the change of the sign of $I_{\text{s/d}}$ from positive to negative, when $[1-\eta]I_4$ overcompensates the hole tunneling current I_1. This downward U turn of $I_{\text{s/d}}$ is also consistent with $I_{\text{s/d}}$ vs the V_g curve in fresh p-MOSFET with gate oxide in the DT regime reported in Ref. 11.

The band diagrams of p-MOSFET in Fig. 3(a) and n-MOSFET in Fig. 3(b) are similar when negative V_g is applied to both devices. However, in the n-MOSFET case, $I_{s/d}$ measures the electron current and I_{sub} measures the hole. Therefore, applying this LPDR-DT model to n-MOSFET, we may predict that all $I-V$ curves for n-MOSFET can be roughly derived from those for p-MOSFET by swapping $I_{s/d}$ and I_{sub}. This is confirmed by experimental curves in Fig. 2(b). Six curves in Fig. 2(b) are analogous to six curves in Fig. 2(a), when $I_{s/d}$ and I_{sub} are swapped each other with different voltage and current scaling, due to different area and thickness of the LPDR. Therefore, the same explanation of Fig. 2(a) can be made for the n-MOSFET curves in Fig. 2(b) as well.

Figures 1(c), 1(d), 2(c), and 3(c) are the case of substrate injection mode on n-MOSFET. In Fig. 3(c), it is shown that $I_{s/d}$ corresponds to the conduction band electron tunneling current I_3, while I_{sub} corresponds to the hole current $I_2+I_1+\eta\gamma I_3$. Here I_2 is actually the substrate valence band electron tunneling current. When the substrate valence band electrons tunnel through the oxide to the gate, there are ''cold'' holes left in the substrate that will diffuse to the substrate and can be measured as negative substrate hole current. Before QB, the oxide is thick so I_2 and $I_1 \approx 0$. The small I_{sub} corresponds to $\eta\gamma I_3$ [$=2$ pA at $V_g=5.4$ V, in Fig. 2(c)] as indicated in Fig. 3(c). From the data of Fig. 2(c), η is less than 10^{-3} at $V_g=5.4$ V if we assume $\gamma \approx 1$. After QB, I_2+I_1 suddenly increases to the range of several tens of nA at $V_g \approx 2-4$ V without changing the polarity as indicated in Figs. 1(d) and 2(c), because the valence electrons at the substrate side (I_2) and the holes at the gate side (I_1) tunnel directly through the oxide due to the reduced effective oxide thickness. This naturally explains the sudden increase of I_{sub} by 3 orders of magnitude after QB when n-MOSFET is stressed under substrate injection mode, as reported but not understood in Refs. 2 and 8.

We have used the carrier separation method by measuring I_g, $I_{s/d}$, and I_{sub} before and after QB, for n-MOSFETs and p-MOSFETs under the substrate and gate injection mode. In all cases, the signs and the magnitudes of all current components can be clearly explained by this unified model that the direct tunnelings of conduction band electrons, valence band electrons and holes through the effectively thinned gate oxide due to formation of the LPDR region after QBs are taken into account. For instance, we can predict that $I_{s/d}$ will increase 2 to 3 orders of magnitude after QB for the case (and only for this case) of p-MOSFET stressed by substrate injection. This has never been reported previously, however, it has been confirmed in our experiment. The detailed experimental results and the analysis of all cases will be published elsewhere.

In summary, systematic carrier separation experiments were conducted to investigate the evolutions of gate, source/drain, and substrate currents before and after gate oxide quasibreakdown in MOSFETs. We found that after quasibreakdown, the substrate current and source/drain current versus gate voltage curves are surprisingly analogous to those curves observed in fresh MOSFET with a gate oxide of direct tunneling thickness. This strongly supports the quasibreakdown model based on the LPDR by which the effective oxide thickness is reduced in that area. When direct tunnelings of conduction band electrons, valence band electrons and holes through thinned gate oxide are considered, the major experimental observations of QB can be explained in a unified and natural way.

This work was supported by Singapore NSTB Research Grant No. NSTB/17/2/3 and the National University of Singapore Research Project No. 3982754.

[1] I. C. Chen, S. Holland, and C. Hu, IEEE Trans. Electron Devices **32**, 413 (1985).
[2] S. H. Lee, B. J. Cho, J. C. Kim, and S. H. Choi, Tech. Dig. Int. Electron Devices Meet., 605 (1994).
[3] K. Okada, S. Kawasaki, and Y. Hirofuji, Extended Abstracts of 1994 International Conference on Solid State Devices and Materials, Yokohama, Japan, 1994, p. 565.
[4] M. Depas, T. Nigam, and M. M. Heyns, IEEE Trans. Electron Devices **43**, 1499 (1996).
[5] K. Okada and K. Taniguchi, Appl. Phys. Lett. **70**, 351 (1997).
[6] T. Yoshida, S. Miyazaki, and M. Hirose, Extended Abstracts of 1996 International Conference on Solid State Devices and Materials, Yokohama, Japan, 1996, p. 539.
[7] T. Tomita, H. Utsunomiya, T. Sakura, Y. Kamakura, and K. Taniguchi, IEEE Trans. Electron Devices **46**, 159 (1999).
[8] F. Crupi, R. Degraeve, G. Groeseneken, T. Nigam, and H. E. Maes, IEEE Trans. Electron Devices **45**, 2329 (1998).
[9] B. Eitan and K. Kolodny, Appl. Phys. Lett. **43**, 106 (1983).
[10] C. Chang, C. Hu, and R. W. Brodersen, J. Appl. Phys. **57**, 302 (1985).
[11] Y. Shi, T. P. Ma, S. Prasad, and S. Dhanda, IEEE Trans. Electron Devices **45**, 2355 (1998).