Electron mobility dependence on annealing temperature of W/HfO_2 gate stacks: The role of the interfacial layer

A. Callegari,[a] P. Jamison, D. Neumayer, and F. Mc. Feely
IBM Semiconductor Research and Development Center (SRDC), IBM T. J. Watson Research Center, P.O. Box 218, Yorktown Heights, New York 10598

J. Shepard
IBM Semiconductor Research and Development Center (SRDC), 2070 Route 52, Hopewell Junction, New York 12533

W. Andreoni, A. Curioni, and C. Pignedoli
IBM Research GmbH-Zurich Research Lab, 8803 Rueschlikon, Switzerland

(Received 15 August 2005; accepted 6 December 2005; published online 25 January 2006)

Interfacial layers between the high-*k* dielectric and Si surface have a very important role to play in the achievement of high electron mobilities required in the next generations of high performance silicon technologies. W/HfO_2 gate stacks formed on $SiO_2/SiON$ interfacial layers subjected to various process conditions were characterized by electrical measurements such as electron mobility, inversion layer thickness, and leakage reduction with respect to standard SiO_2 technology. The required high electron mobilities were obtained only when the W/HfO_2 gate stack was annealed at high temperature. Electrical data and physical analysis of the stack suggest intermixing of the HfO_2 with the interfacial layer and formation of a silicate layer which may lessen the effects of phonon scattering. Densification of the interfacial layer by spike annealing or interfacial layer stabilization by nitrogen plasma inhibits such reactions and lower electron mobilities were obtained. Also, no electrical performance advantages are seen by thinning the HfO_2 down to 1.5 nm. *© 2006 American Institute of Physics*. [DOI: 10.1063/1.2163980]

I. INTRODUCTION

Electron mobility degradation remains one of the main obstacles to implementing metal/high-*k* dielectric in complementary metal oxide semiconductor technology.[1–3] Remote phonon scattering[4] and remote charge scattering[5] have been predicted to severely degrade mobility in *n*-type metal-oxide-semiconductor devices fabricated with high-*k* materials, especially for HfO_2. Here, we review some of the work done on W/HfO_2 gate stacks. We show how mobility is largely dependent on interfacial layer preparation and stack annealing temperature. A high temperature anneal was found to be necessary to achieve high electron mobility on W/HfO_2 gate stacks. At the same time interfacial layer preparation has a critical role to play in the achievement of these high peak electron mobilities. A reaction between the HfO_2 and interfacial layer rather than interfacial layer regrowth appears to be the main reason for mobility enhancement. Also, thinning of the HfO_2 to about 1.5 nm did not yield to any electrical improvement.

For all these stacks, independently of mobility, charge trapping was low[6–8] giving less than 40 mV shift at a stress voltage of 1.5 V for an extrapolated time of ten years. This topic will not be covered here. Mostly, we will review process-electrical data and their relation to electron mobility.

[a]Author to whom correspondence should be addressed; electronic mail: sandro@us.ibm.com

II. EXPERIMENT

W/HfO_2 ($W \sim 30$ nm and $HfO_2 \sim 25$ nm) gate stacks were formed by metalorganic chemical vapor deposition on thin (~1 nm) $SiO_2/SiON$ interfacial layers on bulk Si substrates. N-type field effect transistors (N-FETs) were fabricated by using a simple non self-aligned gate process[2] to decouple the influence of process integration on mobility; $20 \times 5\ \mu m^2$ FETs with channel doping of $\sim 4 \times 10^{17}$ boron/cm^3 were used for this study. The inversion charge used in mobility measurement was determined by using the split capacitance-voltage method. The FETs were characterized by measuring electron mobilities, inversion layer thickness (T_{inv}), and leakage reduction by the current-voltage method and compared to a SiON base line.

III. TEMPERATURE ANNEALING OF GATE STACK

In this first set of experiments, N-FETs were fabricated using a $W/HfO_2/SiO_2$ stack. The stack was then annealed at different temperatures ranging from 400 to 1000 °C followed by a 450 °C forming gas anneal. Electrical data are summarized in the three charts in Fig. 1. In the first chart (a), peak mobilities as a function of annealing temperature are compared against a poly-Si reference at about the same T_{inv}(~2 nm). Electron mobilities are found to increase with annealing temperature.[9] At the highest annealing temperature the peak mobility degradation is about 8% when compared with poly-Si. At high field (~1 MV/cm) no electron mobility degradation was found.[9] The second chart [Fig. 1(b)] shows that the inversion layer thickness increases with an-

 © 2006 American Institute of Physics

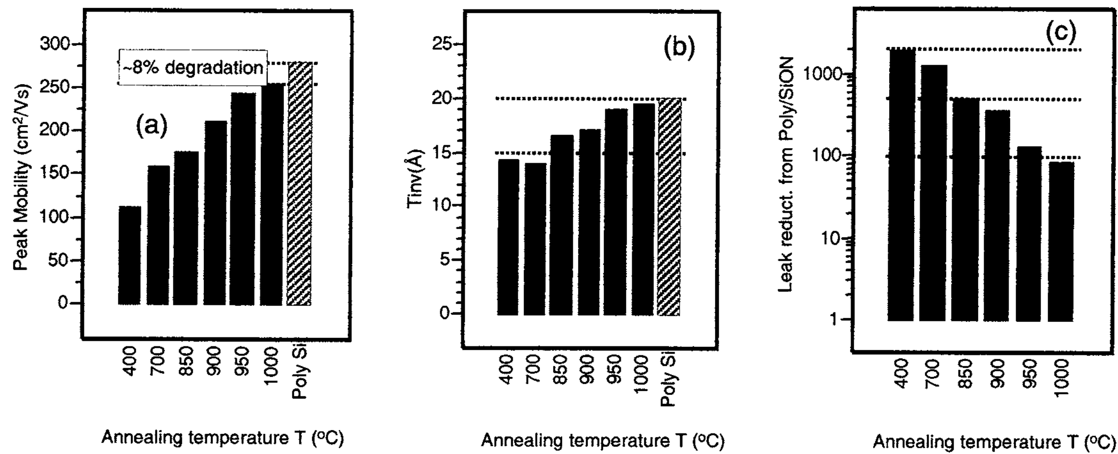

FIG. 1. (a) Electron mobility, (b) inversion layer thickness, and (c) leakage reduction charts of a W(=20 μm)$\times$L(=5 μm) N-FET as a function of annealing temperature. Gate stack is $W/HfO_2/SiO_2$.

nealing temperature. So, at a first glance, it appears that interfacial layer regrowth may be responsible for the higher electron mobilities observed at higher annealing temperature.[8,10] However, this conclusion appears to be inconsistent with the third chart (c) which shows that leakage reduction also decreases with increasing annealing temperature, suggesting that a reaction is occurring between the HfO_2 and the thin SiO_2 interfacial layer. Hf was found to diffuse in the interfacial layer by electron energy loss spectroscopy on similar stacks.[9] Also, our "*ab initio*" calculations are consistent with silicate formation at the interface.[9,11,12] This is shown in Fig. 2. In the center, the dielectric constant of HfOSi is plotted against the Hf content.[11] At low Hf content (left side) a stable SiO_2 structure with Hf embedded is formed and at high Hf content (right side) the stable structure is HfO_2 with Si embedded.[12] Thus, as a result of a high annealing temperature a stable, low Hf content, interfacial HfSiO layer may form on the Si surface.

This silicate formation may reduce the effect of phonon scattering[4] and enable higher electron mobilities. Furthermore, a reduction of interface states ($\sim 5 \times 10^{10}$ charges/cm^2) was found after annealing the stack at high temperature, which is also consistent with the higher mobility values.[9] However, interface states alone ($\sim 2 \times 10^{11}$ charges/cm^2) are insufficient to explain the strong peak mobility degradation (~ 100 cm^2/V s) exhibited by stacks which received only low annealing temperatures.[9] At low annealing temperatures interfacial silicates may not form. Phonon scattering or remote charge scattering rather than interface states may better explain mobility degradation.[4,5] Also, clusters of Hf atoms were found to protrude into the interfacial SiO_2 layer in a $TiN/HfO_2/SiO_2$ gate stack annealed at 900 °C, possibly increasing the apparent permittivity of this interfacial layer compared to a pure SiO_2 layer[13] consistent with our data.

In other work,[14] Hf diffusion was not claimed in a $TiN/HfO_2/SiO_2$ gate stack. In that case, a Si-rich SiO_x layer forms on the Si surface after annealing but this appears to be inconsistent with high mobilities obtained after high temperature annealing since fixed charges in the stacks should

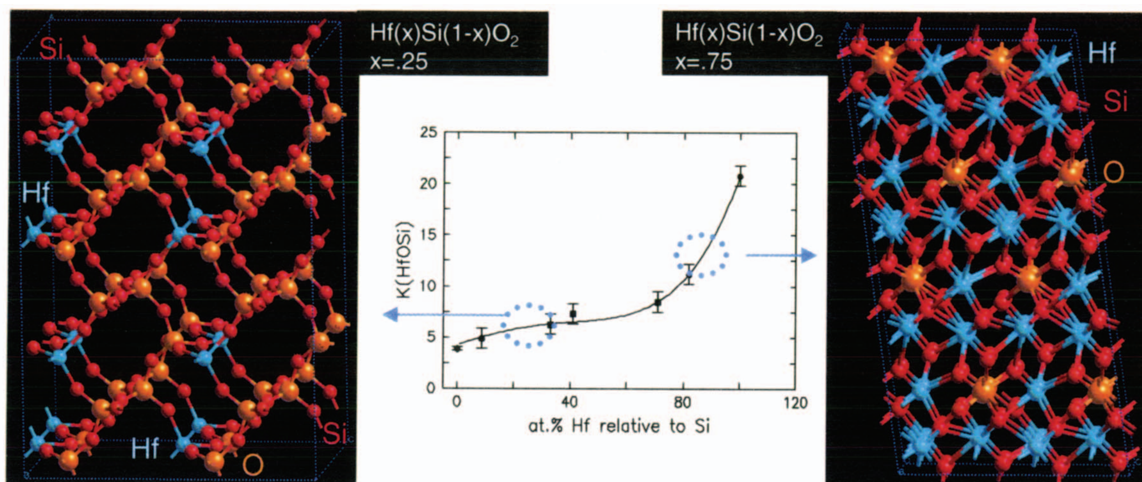

FIG. 2. (Color) Hf silicate structures which may form after a high temperature annealing of $W/HfO_2/SiO_2$. Center is a plot of silicate dielectric constant as a function of Hf at. % relative to Si (see Ref. 11). At low Hf content the silicate structure is a SiO_2 matrix with Hf embedded (left). At high Hf content the silicate structure is a HfO_2 matrix with Si embedded (right).

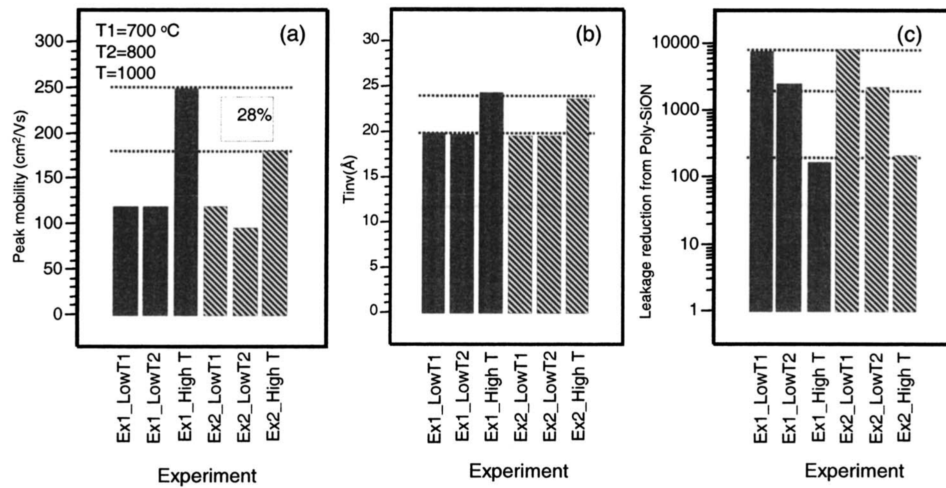

FIG. 3. (a) Electron mobility, (b) inversion layer thickness, and (c) leakage reduction charts of a W(=20 μm)$\times$L(=5 μm) N-FET as a function of annealing temperature; Gate stack is W/HfO_2/SiON. In Ex2 the SiON layer was spike annealed at 1100 °C, in Ex1 the SiON was not spike annealed.

increase. Thus, Hf diffusion in the SiO_2 layer for TiN/HfO_2/SiO_2 gate stacks still appears to be an open question.

It may also be speculated that the reduction in leakage reduction with temperature is due to an increased HfO_2 crystallinity [Fig. 1(c)], thus, enhancing leakage paths through the grain boundaries. But this, as it will be shown below, is not consistent with the effect of high temperature processes in which the HfO_2 and interfacial layer reactions were inhibited and leakage reduction was preserved despite high-*k* grain growth.

IV. SPIKE ANNEAL OF INTERFACIAL LAYER

To better understand the relation between the HfO_2 and interfacial layer on electron mobility, an experiment was performed where the interfacial layer only was spike annealed at 1100 °C before HfO_2 deposition.[15] This was compared with a monitor wafer which did not receive the spike anneal. The spike anneal should improve interfacial layer and therefore the gate stack electrical properties.[15] The interfacial layer was SiO_2 about 1.3 nm thick and it was lightly nitridized.

The electrical-process summary is shown in Figs. 3 and 4. In experiment 1 (Ex1) the interfacial layer was not spike annealed, in experiment 2 (Ex2) it was spike annealed. Low $T1$ and $T2$ refer to temperatures of 700 and 800 °C, respectively, and high T to 1000 °C. At low stack annealing temperatures peak mobilities are severely degraded [Fig. 3(a)] as stated above. Only after the stack is annealed at about 1000 °C is a large mobility improvement observed. There is no difference in peak mobility for Ex1 when compared with Ex2 at low annealing temperatures. Only after annealing the whole stack at high temperature are significant differences observed. The sample in which the interfacial layer was spike annealed showed a 28% degradation in peak mobility. T_{inv} increased from 2.0 to 2.4 nm after annealing, identically for both Ex1 and Ex2, indicating that interfacial layer regrowth was the same for both experiments [Fig. 3(b)]. This again shows that in addition to any regrowth, there must be some reaction at the HfO_2/SiON interface which is responsible for the differences in electron mobility. Leakage reduction compared to a poly-SiON reference also follows the same pattern as described above. As the annealing temperature increases, leakage reduction decreases [Fig. 3(c)].

Contrary to our expectations, peak mobility was found to be degraded when compared with a monitor where the interfacial layer was not spike annealed.[15] We suggest that a denser and less reactive interfacial layer is formed after spike annealing which limits the intermixing with the HfO_2 and therefore the formation of the silicate layer.

To complete the electrical-process matrix threshold voltages and subthreshold slopes are shown in Figs. 4(a) and 4(b), respectively. Note that Vt decreases for both Ex1 and Ex2 with annealing temperature. In general, in Ex2 threshold voltages are higher than in Ex1. At low annealing temperatures, Vt appears to be higher for Ex2 than for Ex1, suggesting that the differences in Vt shifts between the two experiments are due to charges located at the HfO_2 bottom interfaces since the W/HfO_2 is the same in both cases.

Similarly, at high annealing temperature, Vt decreases in both experiments. Again, the Vt shift appears associated with the bottom interfaces (HfO_2/SiON and/or SiON/Si) since the subthreshold slopes are also decreasing with annealing temperature indicating a reduction of interface states. Both Vt and subthreshold slope are higher for Ex2 than for Ex1 suggesting again that the reaction between the HfO_2 and the interfacial layer is inhibited by the spike anneal of the interfacial layer. It should be noted that if the mobility degradation shown in Fig. 3 (Ex2) after 1000 °C anneal is due to charge, this charge must be located at the SiO_2/Si or HfO_2/SiO_2 interface since the W/HfO_2 was deposited in similar fashion on both Ex1 and Ex2. This charge is also responsible for the higher $Vt \sim 0.95$ V observed for Ex2 when compared with Ex1 ($Vt \sim 0.8$ V) after 1000 °C anneal (Fig. 4). Thus, attempts to use unpassivated and/or fixed charges to modulate threshold voltages can lead to significant mobility degradation. Here, charges in the HfO_2 appear not to affect mobility.

V. PLASMA NITRIDATION OF INTERFACIAL LAYER

In this section we show again the importance of the preparation of the interfacial layer on electron mobilities. We will also demonstrate that thinning the HfO_2 to about 1.5 nm

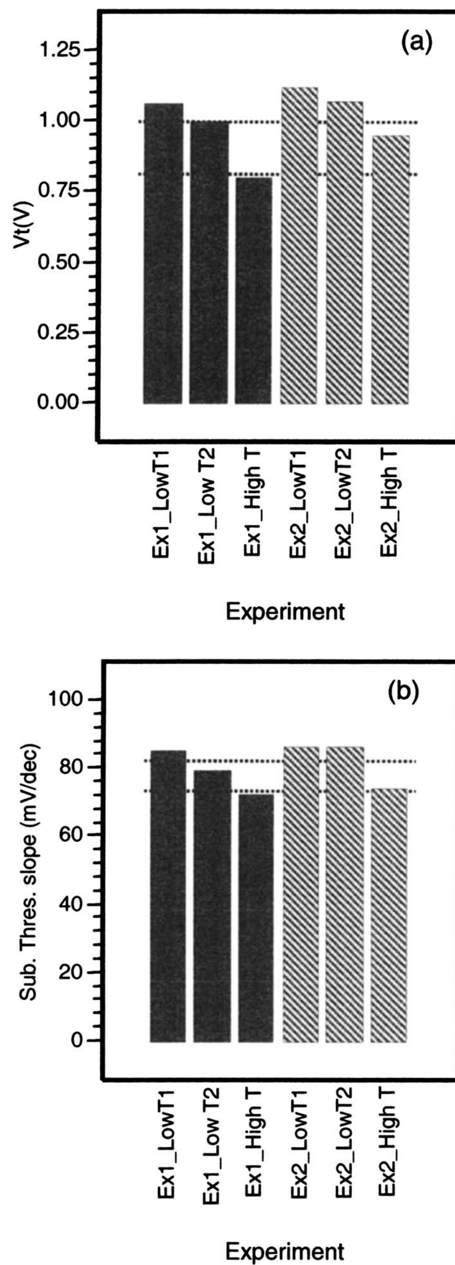

FIG. 4. (a) Threshold voltage and (b) subthreshold slopes charts of a W(=20 μm)$\times$L(=5 μm) N-FET as a function of annealing temperature. Gate stack is W/HfO_2/SiON. In Ex2 the SiON layer was spike annealed at 1100 °C, in Ex1 the SiON was not spike annealed.

does not produce any electrical performance gain. Here, we compare the results of a large process matrix in which only the SiO_2 and the SiON interfacial layers have been plasma nitridized (PN) using a remote low power radio frequency (rf) plasma. Process conditions are labeled 1_PN, 2_PN, 3_PN, and 4_PN and shown in Fig. 5(a). The HfO_2 oxide was not nitridized. Electrical-process results are summarized in Fig. 5.

Two HfO_2 thicknesses, 1.5 and 2.5 nm, were investigated. All the samples shown in Fig. 5 were annealed at 1000 °C. The electron mobility chart shows that all the interfacial layers which received PN conditions have reduced peak electron mobilities of about 100 cm^2/V s despite the high temperature anneal. Only the interfacial layers which did not receive PN (No N_15 Å HfO and No N_25 Å HfO) have peak mobilities above 200 cm^2/V s with the thinnest HfO_2 sample being ~10% degraded. The 0.4–0.5 nm interfacial layer regrowth seen in the previous sections is not observed for most of the plasma nitrided interfacial layers [Fig. 5(b)]. Also, there is no gain in T_{inv} by thinning the HfO_2 to 1.5 nm. T_{inv} could not be scaled below ~1.5 nm for both HfO_2 thicknesses for the processes used here. For the PN_25 Å HfO samples, leakage reduction is about 100$\times$ higher when compared with PN_15 Å HfO samples [Fig. 5(c)].

Thus, there is not much improvement in T_{inv} or electron mobility by thinning the HfO_2 below 2.5 nm. Also, leakage reduction is much worse for the thin 1.5 nm HfO_2 film. Likely, thin HfO_2 films may not be continuous and therefore may exhibit poor electrical properties.

We conclude that stabilization of the interfacial layer by plasma nitridization inhibits the reaction between HfO_2 and interfacial layer, preserving low mobility, preventing interfacial layer regrowth, and maintaining higher leakage reduction. Note that after high temperature annealing, leakage reduction is much higher for interfacial layers which received PN compared with interfacial layers which did not receive PN [Figs. 1(c) and Fig. 5(c)]. Since the HfO_2 film and likely grain formation should be the same in both experiments, the difference in leakage reduction at high annealing temperature is probably related to different stack interfacial compositions and not to leakage paths through grain boundaries.

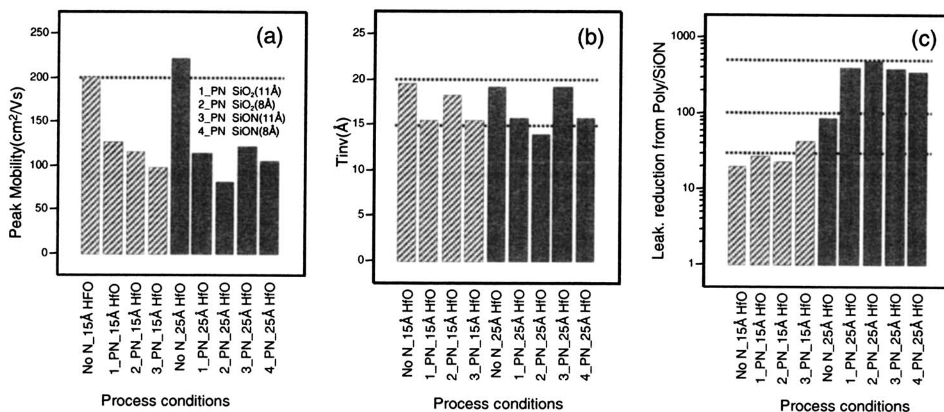

FIG. 5. (a) Electron mobility, (b) inversion layer thickness, and (c) leakage reduction charts of a W(=20 μm)$\times$L(=5 μm) N-FET as a function of plasma nitridation process on different interfacial layers as shown in (a). The gate stack was annealed at 1000 °C. Data are shown for two different HfO_2 thicknesses, 1.5 and 2.5 nm.

VI. SUMMARY

In summary, electron mobility degradation in W/HfO_2 gate stacks is largely dependent on interfacial layer formation on the Si surface. A reaction between the interfacial layer and the HfO_2 and the formation of a silicate layer adjacent to the Si surface after high temperature annealing can explain the high electron mobilities. This silicate layer may lessen the effect of phonon scattering, thus enabling high electron mobility. If such a reaction is inhibited by stabilizing the interfacial layer by spike anneal or by plasma nitridation, degraded electron mobilities are obtained.

ACKNOWLEDGMENTS

The authors would like to thank Dr. M. Igeta from Tokyo Electronics (TEL) corporation for his help with sample preparation and M. Chudzik for critical reading of the manuscript.

[1] V. Ku, R. Amos, A. Steegen, C. Cabral, V. Narayanan, P. Jamison, P. Nguyen, Y. Li, M. Gribelyuk, Y. Wang, J. Cai, A. Callegari, F. McFeely, F. Jamin, K. Wong, E. Wu, A. Chou, D. Boyd, H. Ng, M. Ieong, C. Wann, R. Jammy, and W. Haensch, Extended Abstracts, SSDM, Tokyo, Japan, 2003, p. 730.

[2] A. Callegari, P. Jamison, B. H. Lee, D. Neumayer, V. Narayanan, S. Zafar, E. Gousev, C. D'Emic, D. Lacey, M. Gribelyuk, C. Cabral, A. Steegen, V. Ku, R. Amos, Y. Li, P. Nguyen, F. Mc. Feely, G. Singco, J. Cai, S.-H. Ku, Y. Y. Wang, C. Wajda, D. O'Meara, H. Shinriki, and T. Takahashi, in Ref. 1, p. 808.

[3] L. Ragnarson, W. Tsai, A. Kerber, P. J. Chen, E. Cartier, L. Pantisano, S. De Gendt, and M. Heyns, in Ref. 1, p. 46.

[4] M. V. Fischetti, D. Neumayer, and E. Cartier, J. Appl. Phys. **90**, 4587 (2001).

[5] M. Hiratani, S. Saito, Y. Shimamoto, and K. Torii, Jpn. J. Appl. Phys., Part 1 **41**, 4521 (2002).

[6] A. Callegari, P. Jamison, D. Neumayer, V. Narayanan, S. Zafar, D. Lacey, M. Gribelyuk, R. Amos, F. Mc. Feely, C. Wajda, D. O'Meara, H. Shinriki, and T. Takahashi, Proc. WoDiM 2002, edited by IMEP, Grenoble, France, 2002, p. 57.

[7] E. P. Gusev, V. Narayanan, S. Zafar, C. Cabral, Jr., E. Cartier, N. Bojarczuk, A. Callegari, R. Carruthers, M. Chudzik, C. D'Emic, E. Duch, P. Jamison, P. Kozlowski, D. La Tulipe, K. Maitra, F. R. McFeely, J. Newbury, V. Paruchuri, and M. Steen, International Electron Devices Meeting, San Francisco, 2004, p. 729.

[8] S. Zafar (unpublished).

[9] A. Callegari, P. Jamison, E. Cartier, S. Zafar, E. Gusev, V. Narayanan, C. D'Emic, D. Lacey, F. Mc. Feely, R. Jammy, M. Gribelyuk, J. Shepard, W. Andreoni, A. Curioni, and C. Pignedoli, in Ref. 7, p. 825.

[10] M. Hiratani, S. Saito, Y. Shimamoto, S. Tsujikawa, Y. Matsui, O. Tonomura, K. Torii, J. Yugami, and S. Kimura, in Ref. 1, p. 700.

[11] A. Callegari, E. Cartier, M. Gribelyuk, H. F. Okorn-Schmidt, and T. Zabel, J. Appl. Phys. **90**, 6466 (2001).

[12] C. A. Pignedoli, A. Curioni, and W. Andreoni (submitted).

[13] M. P. Agustin, L. R. C. Fonseca, J. C. Hooker, and S. Stemmer, Appl. Phys. Lett. **87**, 121909 (2005).

[14] G. Bersuker, J. Peterson, J. Barnett, A. Korkin, J. H. Sim, R. Choi, B. H. Lee, J. Greer, P. Lysaght, and H. R. Huff, Proc. 2005 ECS Spring meeting, 2005, Vol. 2005-05, p. 141.

[15] A. Callegari, P. Jamison, S. Zafar, D. Lacey, F. R. McFeely, J. Shepard, E. P. Gusev, E. Cartier, and R. Jammy, in Ref. 14, p. 319.

Journal of Applied Physics is copyrighted by the American Institute of Physics (AIP). Redistribution of journal material is subject to the AIP online journal license and/or AIP copyright. For more information, see http://ojps.aip.org/japo/japcr/jsp