Semiconductors, Vol. 39, No. 6, 2005, pp. 716–721. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 39, No. 6, 2005, pp. 748–753.

PHYSICS OF SEMICONDUCTOR DEVICES

A New Memory Element Based on Silicon Nanoclusters in a ZrO_2 Insulator with a High Permittivity for Electrically Erasable Read-Only Memory

V. A. Gritsenko*^, K. A. Nasyrov, D. V. Gritsenko*, Yu. N. Novikov*, A. L. Aseev*, J. H. Lee***, J.-W. Lee***, and C. W. Kim*****

**Institute of Semiconductor Physics, Siberian Division, Russian Academy of Sciences, Novosibirsk, 630090 Russia*
^e-mail: grits@isp.nsc.ru
***Institute of Automation and Electrometry, Siberian Division, Russian Academy of Sciences, Novosibirsk, 630090 Russia*
****Samsung Advanced Institute of Technology, P.O. Box 111, Suwon 440-600, Republic of Korea*
Submitted August 11, 2004; accepted for publication October 14, 2004

Abstract—The write and erase function and the data retention characteristics of a memory element designed to be used in electrically erasable read-only memory and based on a silicon–oxide–(silicon dot)–oxide–polysilicon structure, in which either a SiO_2 insulator or a ZrO_2 high-permittivity insulator are used as blocking oxides, are simulated. It is established that the use of the high-permittivity insulator gives rise to a number of effects: spurious injection from *poly*-Si is reduced; the electric field in the tunneling oxide increases; it becomes possible to increase the thickness of the tunneling insulator and, consequently, to increase the data retention time; and lower voltages for the write and erase functions can be used. Programming with a pulse of ±11 V possessing a width of 10 ms makes it possible to retain a memory window of ~3 V for 10 years. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

At present, the market for electrically alterable read-only memory (EAROM or flash memory) devices exceeds the markets for random-access memory devices and microprocessors taken together [1]. EAROM devices with a floating gate are prevalent in the nonmilitary market [2]. For custom applications, radiation-resistant silicon–oxide–nitride–oxide–silicon (SONOS) devices based on the memory effect in silicon nitride are used [3]. It is predicted that memory elements with a floating gate will be superseded by the SONOS structures in the development of terabit (10^{12} bits per crystal) EAROM devices [4].

One disadvantage of EAROM devices with a floating gate is that the bit transfer rate cannot be increased from the gigabit range to the terabit range. At present, a tunneling oxide with a thickness of ~7.0 nm and located between the silicon substrate and the floating gate is used in gigabit-capacity EAROM devices. A decrease in the channel length is accompanied by a decrease in the tunneling-oxide thickness. However, a further decrease in the thickness of the tunneling oxide leads to a rapid drain of electrons from polysilicon to the substrate because of the resulting degradation (increase) of the tunneling-oxide conductivity during reprogramming (alternation of pulses with different polarities at the gate). As a result of this reprogramming, traps are formed in the tunneling insulator and, as a consequence, the leakage current (stress-induced leakage current (SILC)) increases [5]. The leakage current in the tunneling oxide (SILC) gives rise to the problem of information retention (retention of the charge in the floating gate), which is required for ten years at 85°C for civil applications and at 125°C for special applications.

At present, EAROM devices based on the conducting nanoclusters in an insulator are actively being developed for terabit memory elements [6–12]. Semiconductor (Si, Ge, Si_xGe_y) or metal clusters with sizes of 1–10 nm act as the memory medium in these EAROM devices. A major advantage of EAROM devices based on nanocrystals in comparison with those based on a floating gate is that the nanoclusters are isolated from each other in the direction parallel to the silicon–insulator interface. As a result, defects (voids) in the tunneling insulator do not give rise to a drain of the complete charge accumulated in these nanoclusters.

Silicon dioxide SiO_2 [6–11], silicon nitride Si_3N_4 [10], and hafnium oxide HfO_2 [11, 12] are the typically used tunneling insulators in memory devices. The blocking insulator between the nanoclusters and the gate is typically fabricated of SiO_2 [6–11]. Since silicon dioxide exhibits a low relative permittivity ($\varepsilon \approx 3.9$), a high electric field (and, consequently, a large spurious voltage drop) appears in the blocking insulator in the course of writing or erasing information (reprogramming). A spurious voltage drop across the blocking SiO_2 layer leads to an undesirable increase in the voltage required for this reprogramming process. In addition, a high electric field in the blocking oxide can give rise to a spurious injection of electrons and holes from the poly-

1063-7826/05/3906-0716$26.00 © 2005 Pleiades Publishing, Inc.

silicon gate and, thus, to a decrease in the memory window (in the difference between the threshold voltages of the memory element in the logic states "0" and "1").

Recently, great efforts have been made to replace silicon dioxide (SiO_2) and oxynitride (SiO_xN_y) in complementary metal–oxide–semiconductor (MOS) transistors with alternative insulators known as high-k insulators, which have a high permittivity [13–15]. The most promising alternative insulators are believed to be ZrO_2, HfO_2 ($\varepsilon \approx 25$), Y_2O_3 ($\varepsilon \approx 15$), and Al_2O_3 ($\varepsilon \approx 10$). It has been shown both theoretically and experimentally that the replacement of the blocking SiO_2 insulator in a SONOS memory element with an alternative insulator (for example, ZrO_2 [16–18] or Al_2O_3 [19]) reduces the voltage drop across the blocking insulator and, consequently, lowers the reprogramming voltage in a SONOS EAROM device. The aim of this study was to gain insight into the data retention in EAROM devices based on a silicon–oxide–(silicon dot)–oxide–silicon (SODOS) structure by simulating the processes of reprogramming and retention (spreading) of the charge. Silicon dioxide and zirconium dioxide were used as the blocking insulators.

2. ENERGY DIAGRAMS OF SODOS STRUCTURES WITH DIFFERENT BLOCKING INSULATORS

In Fig. 1, we show energy diagrams of SODOS structures with SiO_2 (on the left) and ZrO_2 (on the right) oxides used as the blocking insulators without applied voltage and with negative or positive potentials applied to polysilicon. The barrier heights for electrons and holes at the Si/SiO_2 interface are 3.14 and 3.8 eV, respectively [20, 21]. The published data on the barrier height for electrons at the Si/ZrO_2 interface differ: 1.95 ± 0.08 eV [22], 2.0 eV [23], and 1.23 eV [24]. The values of 5.4 eV [23] and 5.5 eV [24] have been reported for the band gap of ZrO_2. In this study, we used the value of 5.5 eV in our calculations. Values of 2.0 and 2.4 eV were used for the barrier height for electrons and holes, respectively. Flat-band voltage U_{FB}, which can be measured in a memory element, was used as a control parameter. This voltage is applied to the gate of a memory structure to ensure that the electric field in silicon at the silicon/insulator interface vanishes.

3. THEORETICAL MODEL

In this study, we simulated the memory properties of SODOS structures with a blocking layer that consisted of either silicon dioxide or the alternative ZrO_2 insulator. An intrinsic semiconductor was used for the Si dots (Si nanoclusters). The acceptor concentration in both the silicon substrate and the polysilicon gate was equal to 2×10^{14} cm^{-3}. The thickness of the tunneling oxide was varied in the range from 1.5 to 5.0 nm. The size of the silicon nanoclusters was fixed in all cases and was equal to 5.0 nm. We disregarded both the quantization

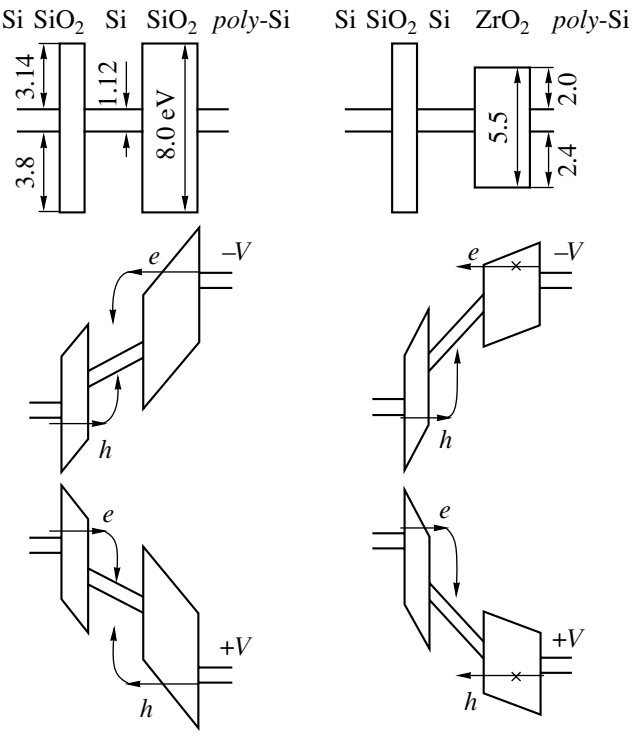

Fig. 1. An energy diagram of SODOS structures with a blocking insulator composed of SiO_2 (on the left) and ZrO_2 (on the right): (at the top) without an applied voltage, (in the middle) with a negative voltage applied to the gate, and (at the bottom) with a positive voltage applied to the gate. The energies are expressed in electronvolts. For simplicity, the voltage drops across the silicon substrate and the silicon gate are not shown.

of the electron and hole spectra, as well as the effect of the Coulomb blockade in the Si dots. The thickness of the blocking SiO_2 layer was 5.0 nm, and that of the blocking ZrO_2 layer was 8.0 nm. We used a one-dimensional two-band model in which electron injection from the negatively biased electrode and hole injection from the positively biased electrode were taken into account (Fig. 1).

We used a modified Fowler–Nordheim formula in calculations of the tunneling current. If the fields and insulator thickness are such that $F_{ox}d_{ox} > \Phi$ (here, d_{ox} is the insulator thickness, Φ is the height of the potential barrier at the Si–insulator interface, and F_{ox} is the electric field in the insulator), the tunneling occurs through a triangular barrier; as a result,

$$j = AF_{ox}^2 \exp\left\{-\frac{4}{3}\frac{\sqrt{2m^*}}{\hbar e}\frac{\Phi^{3/2}}{F_{ox}}\right\}. \quad (1)$$

Here, $A = 2.2 \times 10^{-6}$ A/V^2 and e is the elementary charge. We assumed that the tunneling effective masses m^* for electrons and holes in SiO_2 and ZrO_2 were equal to $0.5m_e$ [20].

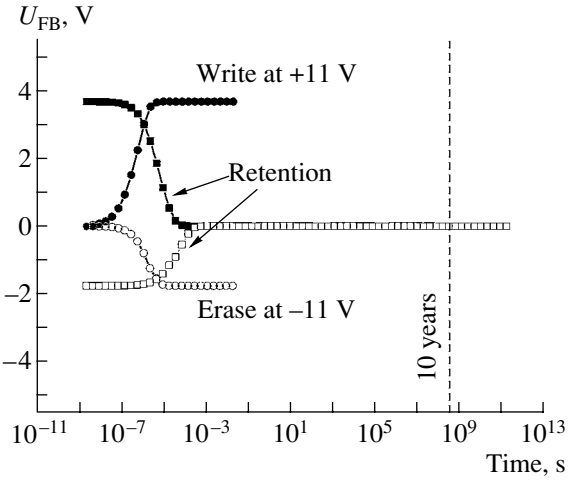

Fig. 2. Characteristics of the write and erase processes in a SODOS structure in which 5.0-nm-thick SiO_2 is used as the blocking insulator. The filled circles represent the accumulation of negative charge, and the open circles represent the accumulation of positive charge. The amplitudes of the write and erase pulses are +11 and –11 V, respectively, and the duration of the pulses is 10 ms. The filled and open squares illustrate the process of charge drain in the shorted state. The Si nanoclusters are 5.0 nm thick, and the tunneling-oxide thickness is 1.5 nm.

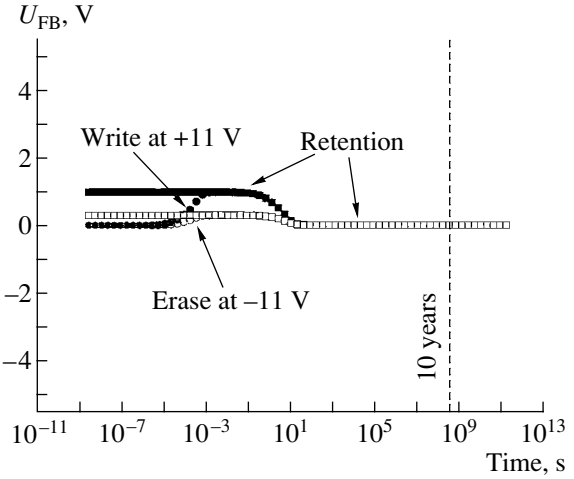

Fig. 3. Characteristics of the write and erase processes in a SODOS structure where 5.0-nm-thick SiO_2 is used as the blocking insulator. The filled and open circles represent the accumulation of negative and positive charges, respectively. The amplitudes of the write and erase pulses are +11 and –11 V, respectively; the duration of the pulses is 10 ms. The filled and open squares illustrate the process of charge drain in the shortened state. The thickness of the Si nanoclusters is 5.0 nm, and the tunneling-oxide thickness is 2.5 nm.

In our calculations, we used the following formula for the tunneling current in the case of a trapezoidal barrier ($F_{ox}d_{ox} < \Phi$):

$$j = AF_{ox}^2 \exp\left\{-\frac{4}{3}\frac{\sqrt{2m^*}}{\hbar e}\frac{(\Phi^{3/2} - (\Phi - F_{ox}d_{ox})^{3/2})}{F_{ox}}\right\}. \quad (2)$$

4. COMPARISON OF THE MEMORY CHARACTERISTICS OF SODOS STRUCTURES WITH SiO_2 AND ZrO_2 AS THE BLOCKING INSULATORS

We carried out the first simulation of the write and erase process and the retention of charge in SODOS structures with the SiO_2 blocking insulator at a tunneling-oxide thickness of 1.5 nm (Fig. 2). The duration of the write and erase pulses was 10 ms in all cases. The positive potential +11 V at the polysilicon gate gave rise to the injection of electrons from the silicon substrate through the tunneling oxide into the Si nanoclusters, the accumulation of negative charge, and a flat-band voltage shift to +4 V. In this case, the memory window amounted to +6 V. In the write and erase mode, we observed a spurious injection of charge through the blocking oxide in accordance with the Fowler–Nordheim mechanism. The drain of negative and positive charges from the Si nanoclusters in the information-retention mode (shortened state) occurred in 10^{-5} s. The drain of electrons and holes from a nanocluster occurs owing to tunneling through the tunneling oxide to the silicon substrate. The discharge due to the tunneling of charge carriers through the blocking oxide into the gate is negligible as a result of the large thickness of this oxide.

When attempting to increase the information-retention time to ten years ($\sim 3 \times 10^8$ s), the most obvious step is to increase the tunneling-oxide thickness so as to suppress the electron and hole tunneling from an Si dot to the substrate. However, an increase in the tunneling-oxide thickness to 2.5 nm at the reprogramming-pulse amplitude of ±11 V leads to a decrease in the memory window to ~1 V (Fig. 3). The 2.5-nm-thick tunneling oxide sharply reduces the electron injection and completely suppresses the hole injection from the substrate. An erase pulse with an amplitude of –11 V does not result in the accumulation of positive charge; however, an insignificant accumulation of negative charge is observed. This behavior is caused by spurious electron injection from the polysilicon gate. An increase in the tunneling-oxide thickness to 2.5 nm results in a slowing down of the charge drain and to an increase in the retention time to ~10 s. However, this value is more than eight orders of magnitude smaller compared to the information-retention time required of an EAROM (ten years). A further increase in the tunneling-oxide thickness is accompanied by a drastic decrease in the memory window, due to an exponential decrease in the injection current of electrons and holes into the Si nanoclusters.

Figure 4 illustrates the dependence of the memory window on the tunneling-oxide thickness in the write and erase mode for voltages of ±8 and ±11 V. It is worth noting that the memory window increases by 2–3 V as the write (erase) voltage decreases from 11 to 8 V. This effect is related to a decrease in the spurious injection through the blocking insulator. It can be seen from

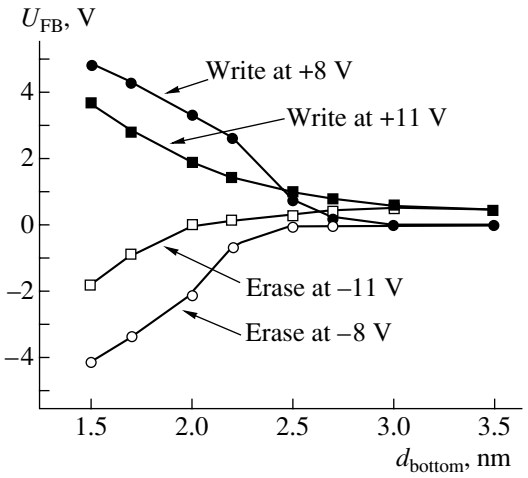

Fig. 4. The dependence of the memory window in a SODOS structure on the tunneling-oxide thickness in the write mode. A 5.0-nm-thick SiO_2 layer was used as the blocking insulator. Voltage pulses with amplitudes of +8 V (filled circles) and +11 V (filled squares) were used for writing and pulses with amplitudes of –8 V (open circles) and –11 V (close squares) were used for erasing. The duration of the pulses was 10 ms. The tunneling-oxide thickness was varied from 1.5 to 3.5 nm.

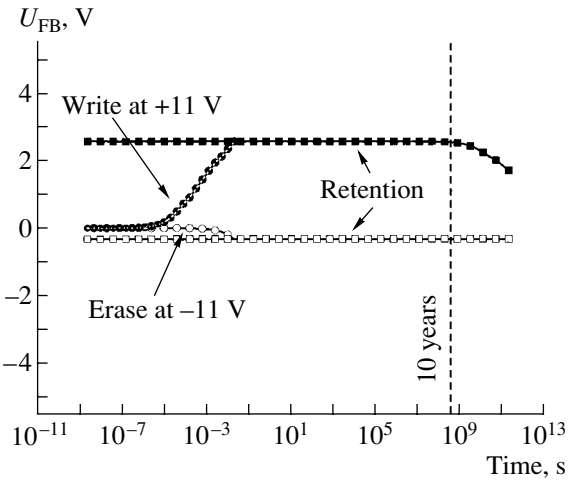

Fig. 5. Characteristics of the write and erase processes in a SODOS structure in which 8.0-nm-thick ZrO_2 is used as the blocking insulator. The filled and open circles represent the accumulation of negative and positive charge, respectively. Voltage pulses with a duration of 10 ms and amplitudes of +11 V (–11 V) were used for the write (erase) operations. The filled and open squares illustrate the process of charge drain in the shortened state. The Si nanocluster thickness was 5.0 nm, and the SiO_2 tunneling-oxide thickness was 5.0 nm.

Fig. 4 that the memory window decreases to less than 1 V if the tunneling-oxide thickness exceeds 2.5 nm.

Thus, the analysis shows that the memory effect (accumulation of electrons and/or holes in SODOS structures) manifests itself only in a situation where the thickness of the bottom oxide is smaller than that of the top oxide. In this case, a charge is accumulated in the silicon dots owing to the higher current of the direct tunneling through the bottom oxide. If the values of the thickness of the bottom and blocking oxide layers are equal to each other and exceed the tunneling length in the insulator, the SODOS structure does not exhibit a memory effect. This circumstance is related to the fact that the current of the electron injection from the substrate into the silicon dots is equal to the current of the electron injection from the silicon dots into the gate.

Replacement of the blocking layer made of SiO_2 by one made of ZrO_2 makes it possible to obtain a flat-band shift of +2.5 V at a tunneling-oxide thickness of 5.0 nm and programming-pulse amplitude of +11 V (Fig. 5). An erase pulse with an amplitude of –11 V leads to only an insignificant accumulation of positive charge (–0.4 V). However, in contrast to SiO_2, a spurious injection of electrons is not observed when ZrO_2 is used as the blocking insulator. The use of the ZrO_2 insulator and a relatively thick tunneling oxide makes it possible to greatly slow down the drain of electrons and obtain a memory window of 3 V even after ten years (Fig. 5).

In Fig. 6, we show the dependence of the memory window on the tunneling-insulator thickness for a

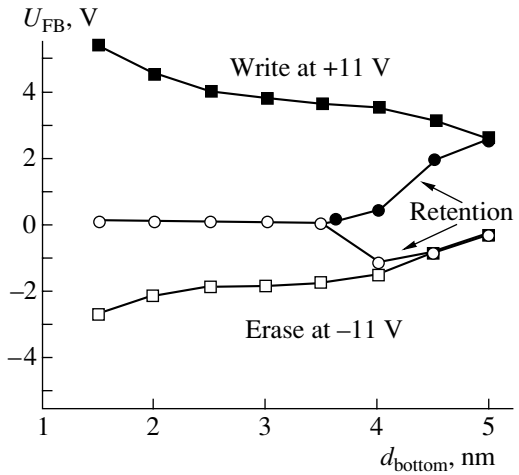

Fig. 6. The dependence of the memory window in a SODOS structure on the thickness of the tunneling (bottom) insulator in the writing and retention modes. A 8.0-nm-thick ZrO_2 layer was used as the blocking insulator. Voltage pulses with a duration of 10 ms and amplitudes of +11 V (filled squares) and –11 V (open squares) were used for the write and erase operations. The tunneling-oxide thickness was varied from 1.5 to 5.0 nm. The memory window (the filled and open circles) is shown after ten years of storage.

SODOS structure with the ZrO_2 blocking insulator. The amplitudes of the write and erase pulses were +11 V and –11 V, respectively. An increase in the tunneling-oxide thickness from 1.5 to 5.0 nm leads to a decrease in the memory window from 8 to 3 V. The memory window for a SODOS structure charged preliminarily with electrons or holes after ten years is also shown in Fig. 6.

The measured memory window is observed for ten years if the tunneling-oxide thickness exceeds 3.5 nm (Fig. 6).

5. DISCUSSION

The replacement of a blocking insulator composed of SiO_2 by one composed of ZrO_2 with a higher permittivity in memory devices based on nanoclusters leads to two effects: (i) the electric field in the tunneling insulator increases and (ii) the field in the blocking insulator decreases. An increase in the field in the tunneling insulator gives rise to an exponential increase in the injection current of electrons and holes from silicon into the nanoclusters, which leads to an increase in the memory window. In addition, it becomes possible to reduce the amplitude and/or the duration of the programming pulse for a fixed memory window. The effect of an increase in the field in the tunneling oxide makes it possible to increase the thickness of this oxide and, thus, suppress the charge drain into the silicon substrate in the mode of information retention.

At present, memory devices based on nanoclusters and insulators with a high permittivity are being studied extensively. For example, Lee *et al.* [12] studied devices with silicon nanoclusters and with HfO_2 used as both the tunneling insulator and the blocking insulator. Similar devices with Si_xGe_{1-x} nanoclusters were studied by Kim *et al.* [11]. However, the effect of an increase in the electric field in the tunneling oxide in the aforementioned structures was not observed. It is noteworthy that the first memory device based on silicon nanoclusters deposited on silicon oxide and coated with silicon nitride was apparently fabricated by Belyĭ *et al.* [25]. In this device [25], the effect of an increase in the field in the tunneling oxide was observed due to the relatively high permittivity of Si_3N_4 ($\varepsilon = 7$) (to be compared with $\varepsilon = 3.9$ in SiO_2).

In the model used in this study, it is assumed that the charge-carrier injection through an insulator with a high permittivity is limited by the Fowler–Nordheim tunneling mechanism. However, in actual insulators with a high permittivity (for example, in Al_2O_3 or HfO_2), there always is a high concentration of traps; therefore, the electrical conductivity is limited by the ionization of defects according to either the Frenkel model or the multiphonon mechanism. Thus, in order to design the devices we have suggested properly, it is necessary to develop the technology of alternative insulators with a low concentration of traps and low leakage currents.

6. CONCLUSIONS

Using the example of ZrO_2, we showed that the application of alternative insulators in the blocking layer in silicon–oxide–(silicon dot)–oxide–silicon (SODOS) structures leads to a number of advantages. These consist in the following:

(i) The electric field in the tunneling oxide increases and, as a result, there is increase in the injection current. This effect makes it possible to use a thicker tunneling insulator at a fixed memory window, which enables an increase in the charge-retention time.

(ii) Lower write and erase voltages can be used for an unchanged SODOS configuration.

(iii) The speed of the response can be increased without changing either the SODOS configuration or the write-and-erase voltages.

(iv) A decrease in the voltage drop across the blocking insulator leads to a reduction in the spurious injection of electrons and holes from the polysilicon gate into the write or erase modes.

The simulation of the charge-retention mode showed that the retention time of the charge is no longer than 10 s if SiO_2 is used as the blocking insulator (for any thickness of the blocking and tunneling oxides). If 8.0-nm-thick ZrO_2 is used as the blocking insulator and 5-nm-thick SiO_2 is used as the tunneling insulator, it becomes possible to attain a memory window of ~3 V, even after ten years at room temperature.

REFERENCES

1. S. M. Sze, in *Future Trends in Microelectronics*, Ed. by S. Luryi, J. Xu, and A. Zaslavsky (Wiley, New York, 1999), p. 135.
2. P. Pavan, R. Bez, P. Olivo, and E. Zanony, Proc. IEEE **85**, 1248 (1997).
3. J. Bu and M. H. White, Solid-State Electron. **45**, 113 (2001).
4. *International Technology Roadmap*, http://public.itrs.net/ (2003).
5. K. Komiya and Y. Omura, J. Appl. Phys. **92**, 2953 (2002).
6. Y. I. Hanafi, S. Tiwari, and I. Khan, IEEE Trans. Electron Devices **43**, 1553 (1996).
7. Y.-C. King, T.-J. King, and C. Hu, IEEE Electron Device Lett. **20**, 409 (1999).
8. B. De Salvo, G. Gibaudo, G. Pananakakis, *et al.*, IEEE Trans. Electron Devices **48**, 1789 (2001).
9. Z. Liu, C. Lee, V. Narayanan, *et al.*, IEEE Trans. Electron Devices **49**, 1614 (2002).
10. M. She and T.-J. King, IEEE Trans. Electron Devices **50**, 1934 (2003).
11. D.-W. Kim, T. Kim, and S. K. Banerjee, IEEE Trans. Electron Devices **50**, 1823 (2003).
12. J. J. Lee, X. Wang, W. Bai, *et al.*, IEEE Trans. Electron Devices **50**, 2067 (2003).
13. G. D. Wilk, R. M. Wallace, and J. M. Anthony, J. Appl. Phys. **89**, 5243 (2001).
14. E. P. Gusev, E. Cartier, D. A. Buchanan, *et al.*, Microelectron. Eng. **59**, 341 (2001).
15. G. Ya. Krasnikov, *Structural and Technological Features of Submicron MOS Transistors* (Tekhnosfera, Moscow, 2002) [in Russian].

16. V. A. Gritsenko, K. A. Nasyrov, and Yu. N. Novikov, in *Proceedings of 12th Workshop on Dielectrics in Microelectronics (WODIM)* (Grenoble, France, 2002), p. 179.

17. V. A. Gritsenko, K. A. Nasyrov, Yu. N. Novikov, and A. L. Aseev, Mikroélektronika **32** (2), 69 (2003) [Russ. Microelectronics **32**, 69 (2003)].

18. V. A. Gritsenko, K. A. Nasyrov, Yu. N. Novikov, *et al.*, Solid-State Electron. **47**, 1651 (2003).

19. C. Lee, S. Hur, Y. Shin, *et al.*, in *Abstracts of 2002 International Conference on Solid State Devices and Materials* (Nagoya, Japan, 2002), p. 162.

20. V. A. Gritsenko, E. E. Meerson, and Yu. N. Morokov, Phys. Rev. B **57**, R2081 (1997).

21. V. A. Gritsenko, *Design and Electronic Properties of Amorphous Insulators in Silicon MIS Structures* (Nauka, Novosibirsk, 1993) [in Russian].

22. V. V. Afanas'ev, M. Houssa, A. Stesmans, *et al.*, Microelectron. Eng. **59**, 335 (2001).

23. M. Houssa, M. Tuominen, M. Naili, *et al.*, J. Appl. Phys. **87**, 8615 (2000).

24. S. Miyazaki, M. Narasaki, M. Ogasawara, and M. Hirose, Microelectron. Eng. **59**, 373 (2001).

25. V. I. Belyĭ, V. V. Voskoboĭnikov, A. S. Ginovker, *et al.*, Mikroélektronika **2**, 182 (1971).

Translated by A. Spitsyn