A 2-bit/Cell Gate-All-Around Flash Memory of Self-Assembled Silicon Nanocrystals

2013 Jpn. J. Appl. Phys. 52 021302

(http://iopscience.iop.org/1347-4065/52/2R/021302)

Japanese Journal of Applied Physics **52** (2013) 021302 **REGULAR PAPER**
http://dx.doi.org/**10.7567/JJAP.52.021302**

A 2-bit/Cell Gate-All-Around Flash Memory of Self-Assembled Silicon Nanocrystals

Hung-Bin Chen[1], Chun-Yen Chang[1], Min-Feng Hung[2], Zih-Yun Tang[2], Ya-Chi Cheng[2], and Yung-Chun Wu[2]*

[1]*Department of Electronics Engineering and Institute of Electronics, National Chiao Tung University, Hsinchu 300, Taiwan*
[2]*Department of Engineering and System Science, National Tsing Hua University, Hsinchu 300, Taiwan*
E-mail: ycwu@ess.nthu.edu.tw

Received October 18, 2012; revised November 22, 2012; accepted November 29, 2012; published online January 21, 2013

This work presents gate-all-around (GAA) polycrystalline silicon (poly-Si) nanowires (NWs) channel poly-Si/SiO_2/Si_3N_4/SiO_2/poly-Si (SONOS) nonvolatile memory (NVM) with a self-assembled Si nanocrystal (Si-NC) embedded charge trapping (CT) layer. Fabrication of the Si-NCs is simple and compatible with the current flash process. The 2-bit operations based on channel hot electrons injection for programming and channel hot holes injection for erasing are clearly achieved by the localized discrete trap. In the programming and erasing characteristics studies, the GAA structure can effectively reduce operation voltage and shorten pulse time. One-bit programming or erasing does not affect the other bit. In the high-temperature retention characteristics studies, the cell embedded with Si-NCs shows excellent electrons confinement vertically and laterally. With respect to endurance characteristics, the memory window does not undergo closure after 10^4 program/erase (P/E) cycle stress. The 2-bit operation for GAA Si-NCs NVM provides scalability, reliability and flexibility in three-dimensional (3D) high-density flash memory applications.

1. Introduction

Flash memory, an nonvolatile memory (NVM), is extensively adopted in portable products owing to its high density and low cost.[1)] Recent efforts have witnessed the aggressive scaling down of flash memory, which follows Moore's law, making device fabrication increasingly difficult. Three-dimensional (3D) memory based on polycrystalline silicon (poly-Si) thin-film transistor (TFT) has subsequently been introduced for ultra-high-density memory.[2,3)]

The silicon–oxide–nitride–oxide–silicon (SONOS) type NVM devices are extensively adopted in flash memory for replacement of conventional floating gate memory, since they have advantages as follows. SONOS NVM is immune to floating-gate coupling, and it exhibits better scaling than the floating gate memory.[4,5)] Moreover, the SONOS NVM dose not suffer from the stress induced leakage current (SILC) due to the discrete charge trapping (CT) property, which further enables the tunneling oxide scaling and reduction of operation voltage.[6)] Recently the nitride read only memory[7)] (NROM) and SONOS NVM[8,9)] apply the spatial programming by channel hot electron injection (CHEI) and erasing by band-to band tunneling induced hot-hole injection (BTBT-HHI), to achieve 2-bit/cell operation for use in system-on-chip (SOP) applications and ultra-high-density data storage flash memory. However, the challenges associated with gate length (L_g) scaling of NROM and SONOS NVM are important to address. The charges have to be kept physically isolated on the source and drain sides, avoiding memory window reduction due to the bit coupling.[10)] In this work, silicon nanocrystal (Si-NC) NVM is used to suppress the bit coupling.

The multi-gate nanowire (NW) structure significantly improves the performance of flash memory, and has progressed in FinFET,[11)] trigate,[9)] Pi-gate,[12)] and gate-all-around[13,14)] (GAA) structures. Among these innovative architectures, the GAA structure achieves the fastest programming and erasing (P/E) speed, due to the highest electric field in the tunneling oxide. Although the GAA structure possesses the best gate control ability, the high electric field in tunneling oxide may result in low data retention. Hence, incorporating discrete Si-NCs in CT layer could reduce the charge leakage for improving the data retention of GAA NVM owing to deep conduction band of NCs. Moreover, Si-NCs are highly compatible with the current flash memory process and absence of metal contamination. Additionally, a local increase of the electric field under the spherical-shaped Si-NCs that improves P/E speed, yet possibly degrades reliability.[15)] Hence, a disk-like or oval-like shape of Si-NC is performed to sustain the reliability of NVM.[16–18)] In this work, poly-Si TFTs GAA NVM embedded Si-NC consisting of ten NWs with different gate lengths are fabricated. The proposed device is characterized by 2-bit operation and its reliability.

2. Experiment

The device was fabricated on a 6-in. silicon wafer with a 400-nm-thick SiO_2. A 50-nm-thick undoped amorphous Silicon (a-Si) layer was deposited by low-pressure chemical vapor deposition (LPCVD) and then solid-phase crystallized (SPC) at 600 °C for 24 h. After the active layer with NWs had been defined, the SiO_2 under the NWs is hollowed by buffered oxide etching (BOE) solution for forming the suspended NWs. The 12-nm-thick layers of tetraethylorthosilicate (TEOS) oxide were deposited by LPCVD as a tunneling oxide layer. Then, a hybrid Si_3N_4 (3 nm)/a-Si (2 nm)/Si_3N_4 (3 nm) discrete trap layer was deposited by LPCVD, and the cell underwent furnace annealing at 1050 °C for 30 min. The 12-nm-thick layers of TEOS oxide were deposited by LPCVD as a blocking oxide layer. Next, A 250-nm-thick in-situ n^+ poly-Si was deposited by LPCVD for filling the space under the suspended NWs. After the gate was patterned, the source/drain (S/D) were formed by phosphorus implantation and activated at 1050 °C for 1 s in nitride ambient. A 200-nm-thick TEOS oxide passivation layer was deposited. Finally, the 300-nm-thick Al–Si–Cu metallization was performed, and sintered in hydrogen ambient at 400 °C for 30 min to obtain good contact.

3. Results and Discussion

Figure 1(a) schematically depicts the proposed GAA NWs structure, and Fig. 1(b) shows the tilted-view scanning electron microscopy (SEM) image of suspended NWs with SiO_2/Si_3N_4/Si-NCs/Si_3N_4/SiO_2 layers. The multi-NWs

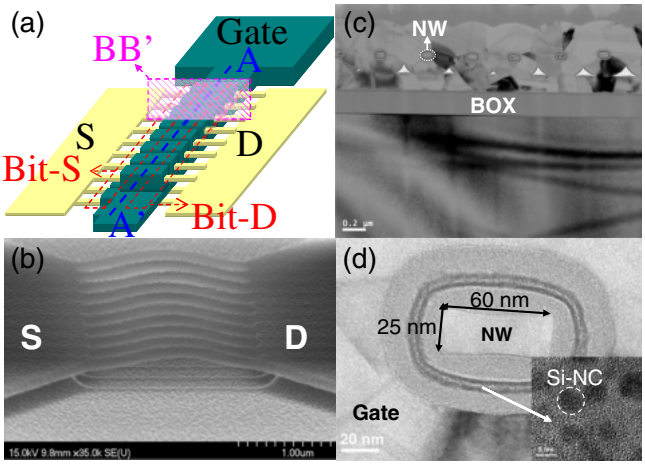

Fig. 1. (Color online) (a) Illustration of the Si-NCs GAA NWs NVM. (b) SEM image of the suspending NWs channel. (c) TEM image of Si-NCs GAA NWs NVM structure along AA′ direction. (d) Magnification TEM image of (c). The effective channel width is (25 nm + 60 nm) × 2 = 170 nm. An inset image shows the plane view of the Si-NCs image. Each NCs are 5–7 nm.

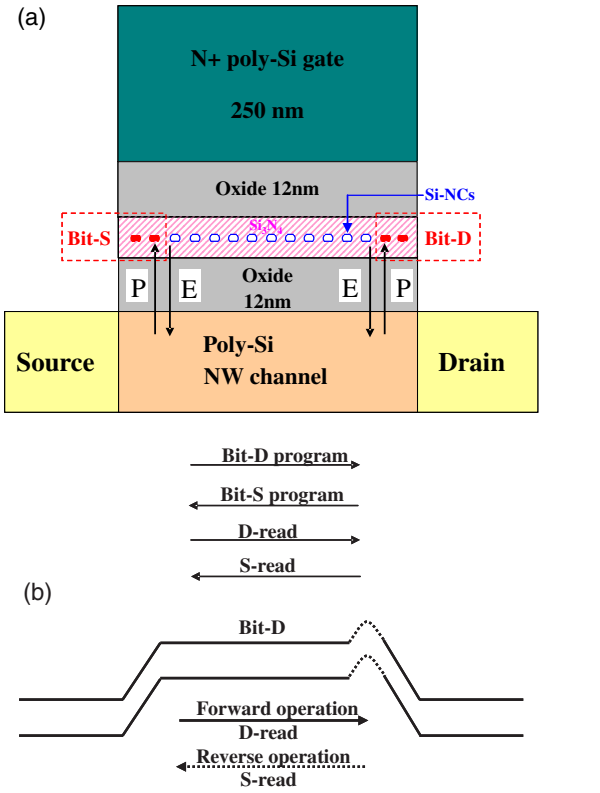

Fig. 2. (Color online) (a) Cross-section view plot of Si-NCs GAA NWs NVM structure including the SiO_2 (12 nm)/Si_3N_4 (3 nm)/Si-NCs (2 nm)/Si_3N_4 (3 nm)/SiO_2 (12 nm) stacked layer with 2-bit storage and the program and read schemes. (b) Band diagram of the cell under forward and reverse operations with Bit-D.

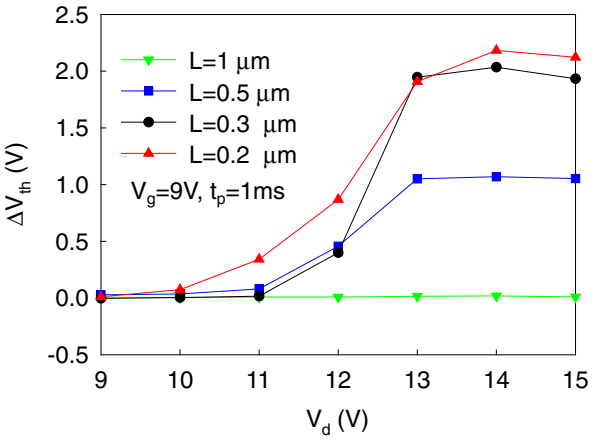

Fig. 3. (Color online) Gate lengths versus efficiency of the CHEI operation.

can obtain a high driving current and the length of the NWs is 2 μm between the S/D. The NWs slightly distort upwards, due to the accumulation of thermal stress in the suspended NWs before the surrounding-gate formation.[19] Figure 1(c) shows a transmission electron microscopy (TEM) photograph along AA′ direction, and Fig. 1(d) presents the enlarged image of Fig. 1(c). It clearly shows that the four corners of the NWs are surrounded by the control gate. The dimension of each NW is 25 nm in height and 60 nm in width; therefore, the effective channel width of one NW is 170 nm. The inset of Fig. 1(d) shows the plane-view of Si-NCs, which are deposited and annealed on the control wafer. During high-temperature thermal annealing, the ultrathin a-Si film (2 nm) self-assembles to form the Si-NCs of around 5–7 nm in diameter with a disk-like shape,[16,17] in which the density of Si-NCs is around 10^{12} cm^{-2}. It is important to note that when the size of the Si-NCs shrinks to the 4.5 nm, the quantum confinement and Coulomb block effect become significant, which usually cause the detrimental effect of NVM performance on data retention and P/E speed.[19,20]

Figure 2(a) presents the cross-section view plot of the poly-Si GAA TFT NVM cell structure with Si-NCs and its 2-bit (bit-S and bit-D) effect operating illustration. For sensing the programmed drain side of the cell and identify the memory state, the forward and reverse operation is performed as shown in Fig. 2(b).

Figure 3 shows the CHEI with various gate lengths for the GAA Si-NCs NVM device. To perform CHEI, moderate V_g and V_d values are applied in a device. The transporting electron in the channel must gain adequate energy for surmounting the Si/SiO_2 barrier, and is locally confined in the CT layer near the drain junction. However, the grain boundary of the poly-Si channel decelerates the transporting electron. For the same V_d, a shorter L_g device has a larger lateral electric field, which heats up transporting electrons and reduces the energy loss due to the less grain boundary scattering in the poly-Si channel. According to Fig. 3, CHEI cannot be performed in the device with $L_g = 1\,\mu m$; however CHEI is significant in the device with $L_g < 0.5\,\mu m$. By using spatial programming and discrete-charge-trapping property of the discrete NCs, the GAA NCs NVM can perform 2-bit per cell. The following 2-bit operations are performed in the GAA NCs NVM with $L_g = 0.5\,\mu m$ because the drain-induced barrier lowering (DIBL) is significant in the device $L < 0.5\,\mu m$, thus limiting the memory window.[9] To avoid interference with the other bit,

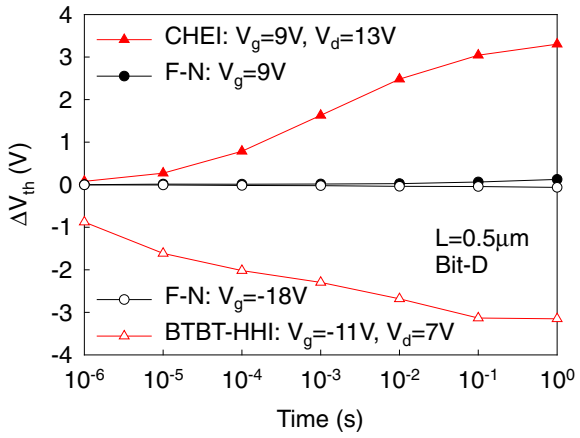

Fig. 4. (Color online) CHEI/BTBT-HHI and FN tunneling program and erase characteristics of bit-D for Si-NCs GAA NVM. The device gate length is 0.5 μm.

the Fowler–Nordheim (FN) current component should be eliminated.

Figure 4 shows the P/E characteristics based on CHEI and BTBT-HHI mechanism for GAA Si-NCs NVM devices. A FN tunneling condition with the same gate voltage (9 V) is compared with the CHEI. The device can obviously not be programmed by FN tunneling at $V_g = 9$ V, yet can be programmed by CHEI at $V_g = 9$ V and $V_d = 13$ V. This observation indicates that the CHEI condition should dominate the gate current to ensure the 2-bit operation. A -18 V of gate voltage, same voltage difference between the gate and drain side, cannot erase in our GAA NCs NVM; however, at $V_g = -11$ V and $V_d = 7$ V, the purposed device can be erased based on BTBT-HHI. Under such an erase condition, the gate current is dominated by BTBT-HHI. BTBT-HHI mechanism is performed for the erasing operation. As the drain voltage is relatively high, with a gate and grounded source, the large electric field and a large band bending exceeding the Si energy band gap at the drain-oxide interface near the n^+ drain junction side leads to BTBT. Some of the holes generated by BTBT are injected into tunneling oxide directly by the electric field towards the trapping layer; others are drawn by the lateral electric field towards the source side, and the accelerated holes can gain sufficient energy, progressing hot-hole injection, to overcome the valence-band offset at the silicon–oxide interface.[21)]

Table I depicts the bias conditions of poly-Si TFTs GAA Si-NCs NVM with CHEI programming ($V_g = 9$ V, $V_d = 13$ V, $t_p = 0.7$ ms), BTBT-HHI erasing ($V_g = -11$ V, $V_d = 7$ V, $t_e = 1$ ms), and reading status. The unique feature of 2-bit per cell operation is based on self-assembly Si-NCs localized charge storage. Based on these CHEI and BTBT-HHI conditions, Fig. 5 plots the I_d–V_g performance of a 2-bit operation for the GAA NCs NVM with four states — (0, 0), (0, 1), (1, 1), and (1, 0). For instance, the symbol "(1, 0)" represents bit-S in the programming state and bit-D in the erasing state. The term "D-read", for example, refers to a situation in which reading voltage is applied at the drain side. The reading voltage is 3 V, and the memory window is approximately 1 V. In addition to the requirement that it must be adequately large to extend the junction depletion region in order to screen out the localized trapped charges, the reading voltage should not influence the channel potential created by the other bit. The off current rises when a read voltage is performed at the bit side full of electrons, because of the gate induced drain leakage current (GIDL) effect.

Table I. The 2-bit operation mechanism and bias conditions utilized GAA Si-NCs poly-Si NVM with $L_g = 0.5$ μm.

	Bit-D			Bit-S		
	V_g	V_d	V_s	V_g	V_d	V_s
Program (V) CHEI	9	13	0	9	0	13
Erase (V) BTBT-HHI	−11	7	0	−11	0	7
Read (V)	×	0	3	×	3	0

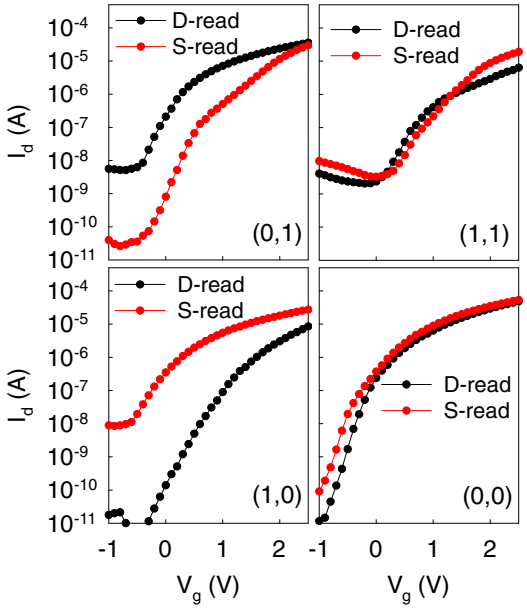

Fig. 5. (Color online) GAA Si-NCs NVM ($L_g = 0.5$ μm) under CHEI ($V_g = 9$ V and $V_d = 13$ V in 0.7 ms) and BTBT-HHI ($V_g = -11$ V and $V_d = 7$ V in 1 ms).

Figures 6(a) and 6(b) show the V_{th} shift of two bits during bit-D programming and erasing characteristics, respectively, of GAA NCs NVM. In Fig. 6(a), when the bit-D is programming and erasing, bit-S is slightly disturbed. Owing to the exclusion of the FN tunneling current, bit-D does not affect the bit-S. Compared with previous studies with tri-gate SONOS NVM,[9)] GAA Si-NCs NVM indeed further improve the 2-bit cell performance. One reason is that the GAA structure of outstanding gate control ability reduces the operation voltage and P/E pulse time. Another reason is that the Si-NCs can localize electrons in the narrow region, which assists the gate length scaling.

Figure 7 displays the endurance characteristics of the bit-D by CHEI and BTBT-HHI P/E cycles. Although the V_{th} value of the S-read shifts slightly upwards, the memory window does not undergo closure after 10^4 P/E cycles. This finding can be explained as follows. First, the interface between Si-NCs and nitride contains numerous trapping

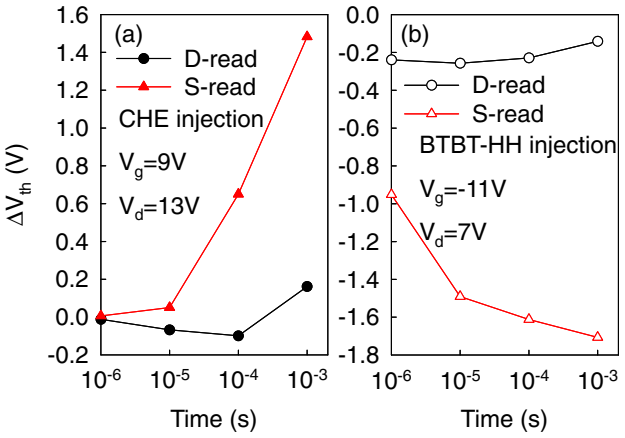

Fig. 6. (Color online) 2-bit operation P/E speed by (a) CHE injection and (b) BTBT-HH injection. The reading voltage is 3 V and device gate length is 0.5 μm.

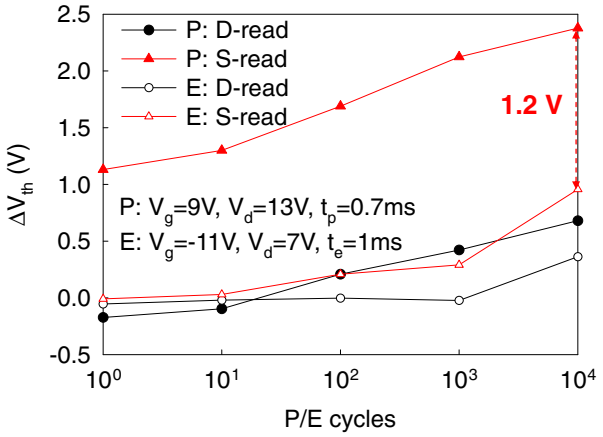

Fig. 7. (Color online) Endurance characteristics of the bit-D.

sites, which can store the tunneling charges.[22–24] Second, the NCs device has a higher hole tunneling current than in the SONOS device because the Si-NCs have deeper valence bands. Hence, most of the electrons are trapped in the NCs or its interface, and the tunneling holes can efficiently recombine with stored electrons.

Figure 8 shows the bit-D retention at 85 °C. The memory window is maintained for up to ten years. The inset shows the schematic oval-like shape of Si-NC surrounded by Si_3N_4 layer. The electric field in the radius component (E_r) can be approximately given as the following[25] to state the phenomenon qualitatively:

$$E_r \cong E_0\left[1+\frac{2a^3}{r^3}\left(\frac{\varepsilon_{\text{Si-NC}}-\varepsilon_{\text{SiN}}}{\varepsilon_{\text{Si-NC}}+\varepsilon_{\text{SiN}}}\right)\right]\cos\theta+\frac{\sum_i Q_i}{4\pi\varepsilon_{\text{SiN}}r^2}. \quad (1)$$

Here, $\varepsilon_{\text{Si-NC}}$ and ε_{SiN} are dielectric constant of Si-NCs and Si_3N_4, a is the radius of Si-NCs, r is the distance from the center of Si-NCs, θ is the angel from the horizontal plane and $\sum_i Q_i$ is the total charges in Si-NCs. Base on Eq. (1), Si-NCs is covered with high dielectric constant material, indicating the E_r declines. This phenomenon suppresses the stored charge migration laterally. In addition, the stored

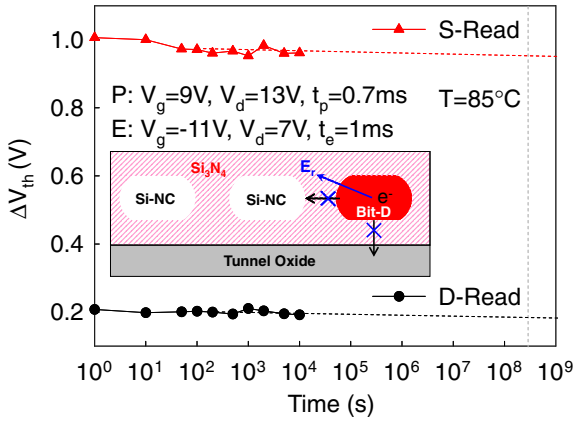

Fig. 8. (Color online) Bit-D retention by CHE programming at 85 °C. Si-NCs GAA NWs NVM shows excellent charge confinement.

charge would not leak vertically, owing to the deep quantum well of Si-NC, which effectively suppresses direct tunneling of electrons from the CT layer into the channel.

4. Conclusions

The 2-bit effect of GAA NVM with Si NCs through self-assembly processes is investigated. The experimental results reveal that the GAA Si-NCs NVM performs clear 2-bit effect with gate length of 0.5 μm by CHE programming and channel hot holes erasing. At large gate length of 1 μm, FN tunneling occurred and dominated which resulted in the absence of the 2-bit effect. In the programming and erasing characteristics studies, the GAA structure can reduce operation voltage and shorten pulse time. In the retention characteristics studies, the Si-NCs of confining electrons in the narrow region assist the gate length scaling and lateral migration. Such GAA Si-NCs poly-Si TFT NVM is, thus, highly promising for use in future SOP and 3D stacked high-density NAND flash memory applications.

Acknowledgments

The authors would like to thank the National Science Council of Republic of China, Taiwan, for supporting this research (NSC 100-2221-E-007-088-MY2). The National Nano Device Laboratories (NDL) is greatly appreciated for its technical support.

1) R. Bez, E. Camerlenghi, A. Modelli, and A. Visconti: Proc. IEEE **91** (2003) 489.
2) E. K. Lai, H. T. Lue, Y. H. Hsiao, J. Y. Hsieh, S. C. Lee, C. P. Lu, S. Y. Wang, L. W. Yang, K. C. Chen, J. Gong, K. Y. Hsieh, J. Ku, R. Liu, and C. Y. Lu: Symp. VLSI Technology Dig. Tech. Pap., 2006, p. 46.
3) H. T. Lue, T. H. Hsu, Y. H. Hsiao, S. P. Hong, M. T. Wu, F. H. Hsu, N. Z. Lien, S. Y. Wang, J. Y. Hsieh, L. W. Yang, T. Yang, K. C. Chen, K. Y. Hsieh, and C. Y. Lu: Symp. VLSI Technology Dig. Tech. Pap., 2010, p. 131.
4) 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, and R. Muralidhar: Solid-State Electron. **48** (2004) 1463.
5) B. De Salvo, C. Gerardi, R. van Schaijk, S. Lombardo, D. Corso, C. Plantamura, S. Serafino, G. Ammendola, M. van Duuren, P. Goarin, W. Y. Mei, K. van der Jeugd, T. Baron, M. Gély, P. Mur, and S. Deleonibus: IEEE Trans. Device Mater. Reliab. **4** (2004) 377.

6) M. H. White, D. A. Adams, and B. Jiankang: IEEE Circuits Devices Mag. **16** (2000) 22.
7) B. Eitan, P. Pavan, I. Bloom, E. Aloni, A. Frommer, and D. Finzi: IEEE Electron Device Lett. **21** (2000) 543.
8) Y. H. Lin, C. H. Chien, T. H. Chou, T. S. Chao, and T. F. Lei: IEEE Trans. Electron Devices **54** (2007) 531.
9) Y. C. Wu, M. F. Hung, C. W. Chang, and P. W. Su: Appl. Phys. Lett. **92** (2008) 163506.
10) H. T. Lue, T. H. Hsu, M. T. Wu, K. Y. Hsieh, R. Liu, and C. Y. Lu: IEEE Trans. Electron Devices **53** (2006) 119.
11) P. Xuan, M. She, B. Harteneck, A. Liddle, J. Bokor, and T. J. King: IEDM Tech. Dig., 2003, p. 26.4.1.
12) S. C. Chen, T. C. Chang, P. T. Liu, Y. C. Wu, C. C. Ko, S. Yang, L. W. Feng, S. M. Sze, C. Y. Chang, and C. H. Lien: Appl. Phys. Lett. **91** (2007) 213101.
13) C. Li, M. Bescond, and M. Lannoo: Appl. Phys. Lett. **97** (2010) 252109.
14) A. S. Verhulst, B. Sorée, D. Leonelli, W. G. Vandenberghe, and G. Groeseneken: J. Appl. Phys. **107** (2010) 024518.
15) G. Gay, D. Belhachemi, J. P. Colonna, S. Minoret, P. Brianceau, D. Lafond, T. Baron, G. Molas, E. Jalaguier, A. Beaurain, B. Pelissier, V. Vidal, and B. De Salvo: Appl. Phys. Lett. **97** (2010) 152112.
16) M. Dai, K. Chen, X. Huang, L. Wu, L. Zang, F. Qiao, W. Li, and K. Chen: J. Appl. Phys. **95** (2004) 640.
17) L. Zhang, K. Chen, L. Wang, W. Li, J. Xu, X. Huang, and K. Chen: J. Phys.: Condens. Matter **14** (2002) 10083.
18) H. B. Chen, Y. C. Wu, L. C. Chen, J. H. Chiang, C. K. Yang, and C. Y. Chang: IEEE Trans. Electron Device Lett. **33** (2012) 537.
19) T. Takagahara and K. Takeda: Phys. Rev. B **46** (1992) 15578.
20) T. W. Kim, C. H. Cho, B. H. Kim, and S. J. Park: Appl. Phys. Lett. **88** (2006) 123102.
21) B. Eitan, P. Pavan, I. Bloom, E. Aloni, A. Frommer, and D. Finzi: IEEE Trans. Electron Device Lett. **21** (2000) 543.
22) J. D. Choe, S. H. Lee, J. J. Lee, E. S. Cho, Y. Ann, B. Y. Choi, S. K. Sung, J. No, I. Chung, K. Park, and D. Park: Jpn. J. Appl. Phys. **46** (2007) 2197.
23) M. She and T. J. King: IEEE Trans. Electron Devices **50** (2003) 1934.
24) M. F. Hung, Y. C. Wu, and Z. Y. Tang: Appl. Phys. Lett. **98** (2011) 162108.
25) D. K. Cheng: *Field and Wave Electromagnetics* (Addison-Wesley, Reading, MA, 1983).