FOREWORD

Special Section on Fundamentals and Applications of Advanced Semiconductor Devices

This special section is focusing on fundamentals and application of advanced semiconductor devices. The authors who gave their presentations at the Asia-Pacific Workshop on Advanced Semiconductor Devices (AWAD) mainly submitted their papers to this special section. The 30th anniversary of AWAD was successfully held in 2023 with 93 papers and 170 attendees at Tokyo Institute of Technology, Yokohama, Japan. This year, excellent papers related to the MONOS type flash memory with ferroelectric blocking layer, pin photodetectors of Ge on Si, and Ohmic contacts to AlGaN/GaN structures were accepted in this special section, which are important topics for the advanced semiconductor devices. The structure and fabrication process of semiconductor devices such as CMOS and NAND flash memory have been markedly changed and developed in last decade. The development of those devices has cultivated their novel applications such as AI and quantum computing. I believe this special section would be helpful for the readers to understand the current topics and future prospects of the advanced semiconductor devices.

I appreciate all the authors for their contributions to the special section. I also thank all the reviewers and editorial committee members for their great effort on the publication procedure.

Special Section Editorial Committee Members
Secretaries:
Tomoyuki Suwa (Tohoku Univ.), Takuya Futase (Western Digital)
Members:
Tatsuya Usami (Rapidus), Hiroya Ikeda (Shizuoka Univ.), Taiji Noda (Panasonic), Takuji Hosoi (Kwansei Gakuin Univ.), Hiroshi Okada (Toyohashi Univ. Technology), Taizo Sadoh (Kyushu Univ.), Hiroshige Hirano (Tower Partners Semiconductor), Kazuyuki Omori (Renesas), Yukiharu Uraoka (NAIST), Yuzo Ohno (Tsukuba Univ.), Seiya Kasai (Hokkaido Univ.), Nobuaki Kobayashi (Nihon Univ.), Kiwamu Sakuma (Kioxia), Takahiro Shinada (Tohoku Univ.), Michihiko Suhara (Tokyo Metropolitan Univ.), Hajime Tanaka (Osaka Univ.), Takashi Noguchi (Univ. of Ryukyus), Shintaro Nomura (Tsukuba Univ.), Toshiro Hiramoto (The Univ. of Tokyo), Naoki Matsuda (FC-Cubic), Tomohide Terashima (Kyushu Univ.), Seiichi Miyazaki (Hiroshima Univ.), Takahiro Mori (AIST)

Shun-ichiro Ohmi (Tokyo Tech), Guest Editor-in-Chief

Shun-ichiro Ohmi *(Member)* received his B.S. degree in Material Engineering from Nagoya University, Nagoya, Japan, in 1991, M.S. and Ph. D. degrees in Applied Electronics from the Tokyo Institute of Technology, Yokohama, Japan, in 1993 and 1996, respectively. In 1996, he joined Tokyo Institute of Technology, where he has been engaged in the research of Si and organic semiconductor devices and processes. From 1997 to 1999, he was a Post Doctoral Technical Staff with the Lucent Technologies Bell Laboratories, where he was engaged in the research of $TiSi_2$ and $CoSi_2$ salicide processes. In 2000, he was a visiting researcher with North Carolina State University where he was engaged in the research of high-k gate insulator. Dr. Ohmi has served as a chair of Silicon Devices and Materials (SDM) Technical Committee of IEICE since 2022, and he is a member of the IEEE Electron Devices Society, the Japan Society of Applied Physics, and a Fellow of the Institute of Electrical Engineers of Japan.

Copyright © 2024 The Institute of Electronics, Information and Communication Engineers

PAPER *Special Section on Fundamentals and Applications of Advanced Semiconductor Devices*

Digital/Analog-Operation of Hf-Based FeNOS Nonvolatile Memory Utilizing Ferroelectric Nondoped HfO_2 Blocking Layer

Shun-ichiro OHMI[†a)], ***Member***

SUMMARY In this research, we investigated the digital/analog-operation utilizing ferroelectric nondoped HfO_2 (FeND-HfO_2) as a blocking layer (BL) in the Hf-based metal/oxide/nitride/oxide/Si (MONOS) nonvolatile memory (NVM), so called FeNOS NVM. The Al/$HfN_{0.5}$/$HfN_{1.1}$/HfO_2/p-Si(100) FeNOS diodes realized small equivalent oxide thickness (EOT) of 4.5 nm with the density of interface states (D_{it}) of 5.3×10^{10} $eV^{-1}cm^{-2}$ which were suitable for high-speed and low-voltage operation. The flat-band voltage (V_{FB}) was well controlled as 80–100 mV with the input pulses of ±3 V/100 ms controlled by the partial polarization of FeND-HfO_2 BL at each 2-bit state operated by the charge injection with the input pulses of +8 V/1–100 ms.
key words: *ferroelectric nondoped HfO_2, metal/oxide/nitride/oxide/Si, nonvolatile memory, partial polarization, charge trap*

1. Introduction

Metal-oxide-nitride-oxide-Si (MONOS) nonvolatile memories (NVM) are widely investigated not only for storage memory but for in-memory computing applications [1], [2]. Utilizing the high-k (HK) thin films in MONOS NVM is effective to reduce the operation voltage and improve the operation speed [3], [4]. The memory window (MW) of MONOS NVM is necessary to be increased even when the operation voltage is decreased. In order to increase the MW, metal-ferroelectrics-nitride-oxide-Si (MFNOS) structure was proposed utilizing $Sr_{0.7}Bi_{2.3}Nb_2O_9$ (SBN) as a ferroelectric blocking layer (BL) for further improvement of memory characteristics of MONOS NVM [5]. However, the thickness of SBN was 100 nm to obtain the ferroelectric characteristics, and it was hard to be scaled although the relative dielectric constant (ε_r) was high as 1000.

Since the HfO_2 thin film crystallized in the metastable orthorhombic phase was reported to show ferroelectric characteristics [6], the applications of ferroelectric HfO_2 in the MONOS structure have been attracting much attention because of its Si process compatibility, and the HfO_2 shows ferroelectric characteristics even bellow the thickness of 10 nm which is suitable for device scaling [7], [8]. The ferroelectric HfO_2 is effective to increase MW which is similar to Ref. [3].

We have proposed the digital/analog-operation utilizing ferroelectric nondoped HfO_2 (FeND-HfO_2) as a BL in the Hf-based MONOS structure, which is called FeNOS NVM, as shown in Fig. 1 (a) [9]–[12]. The FeND-HfO_2 was able to be formed when the nitrogen concentration of HfN_x CTL was x = 1.1. The Hf-based FeNOS stacked structures from the HK-HfO_2 tunneling layer (TL) to the $HfN_{0.5}$ gate electrode layer are able to be deposited in a sputtering chamber by reactive sputtering process without exposing to the air. The FeNOS NVM is expected to realize the analog control of threshold voltage (V_{TH}) by the partial polarization of FeND-HfO_2 BL along with the multi-bit/cell operation by the charge trap in the HK-$HfN_{1.1}$ CTL through a HK-HfO_2 TL as shown in Fig. 1 (b). The polarization switching is able to be controlled at low-voltage and the switching speed is quite fast, while the charge trap and detrap operations are performed at high-voltage.

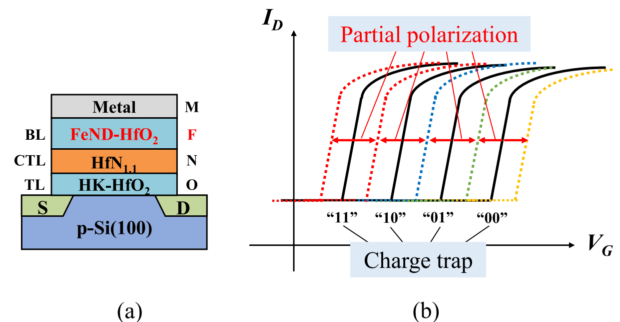


Fig. 1 (a) Schematic cross-section of the FeNOS NVM and (b) schematics of V_{TH} control in FeNOS NVM. The partial polarization of FeND-HfO_2 BL realizes the analog control of V_{TH} (dotted lines) along with the multi-bit/cell operation of the charge trap in the $HfN_{1.1}$ CTL (solid lines).

In this paper, we have investigated the fabrication process of Hf-based FeNOS diode, and the digital/analog-operation of Hf-based FeNOS diode was examined by controlling the pulse input conditions [13].

2. Experimental Procedure

Figure 2 shows the fabrication process for the FeNOS diodes. The schematic cross-sections and the plane-view of the fabricated FeNOS diodes are also shown.

For the fabrication of FeNOS diodes, lightly doped p-Si(100) (10–30 Ωcm) substrates were cleaned by sulfuric-peroxide mixture (SPM) and diluted HF (DHF) solutions. After the 100 nm thick field SiO_2 formation on p-Si(100) substrates, active area was patterned. Some of the FeNOS diodes were fabricated without field oxide. Then, the Hf-based FeNOS structures of $HfN_{0.5}$ (gate electrode, 10 nm)/FeND-

Manuscript received September 19, 2023.
Manuscript revised April 3, 2024.
Manuscript publicized June 3, 2024.
[†]Tokyo Institute of Technology, Yokohama-shi, 226–8502 Japan.
a) E-mail: ohmi@ee.e.titech.ac.jp
DOI: 10.1587/transele.2023FUP0002

Copyright © 2024 The Institute of Electronics, Information and Communication Engineers

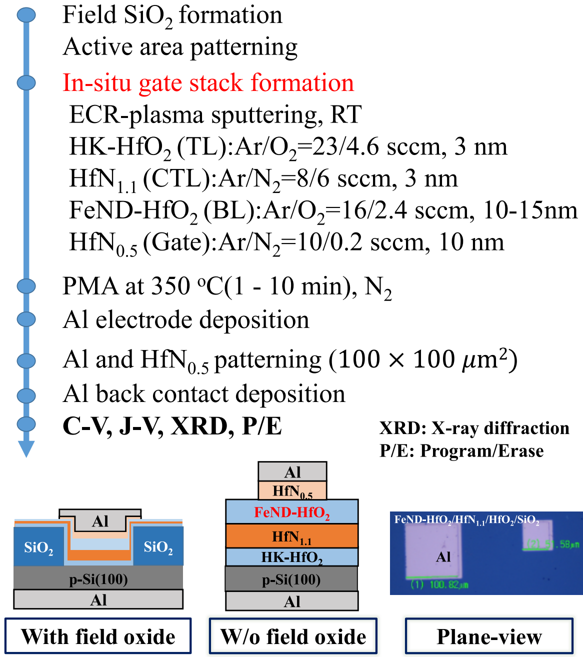


Fig. 2 Fabrication process for $Al/HfO_{0.5}/HfO_2/HfN_{1.1}/HfO_2/Si(100)$ FeNOS diodes. Schematic cross-sections and plane-view were also shown.

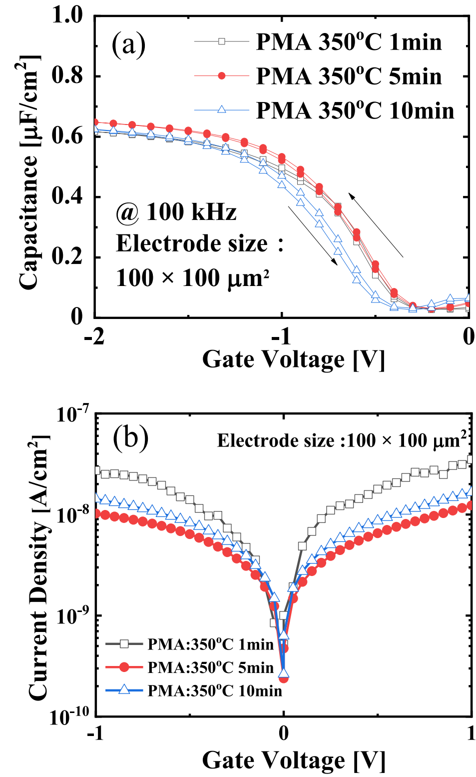


Fig. 3 PMA duration dependence of (a) C-V (100 kHz) and (b) J-V.

HfO_2 (10–15 nm)/$HfN_{1.1}$ (3 nm)/HK-HfO_2 (2 nm)/Si(100) were in-situ deposited by the electron cyclotron resonance (ECR)-plasma sputtering at room temperature (RT) followed by the post-metallization annealing (PMA) at 350°C/1–10 min in N_2 ambient. For the HK-HfO_2 TL deposition, the Ar/O_2 flow ratio was 23/4.6 sccm, while it was 16/2.4 sccm for the FeND-HfO_2 BL deposition. The Ar/N_2 flow ratio for $HfN_{1.1}$ CTL was 8/6 sccm, while it was 10/0.2 sccm for the $HfN_{0.5}$ gate electrode deposition. Next, Al top contact was evaporated, and the gate electrode was patterned by wet etching with the size of 100 × 100 μm².

The FeNOS diode structures were evaluated by C-V, J-V, and program/erase (P/E) measurements utilizing HP 4284A and Agilent 4156C, respectively. The density of interface states (D_{it}) was extracted by Terman method at midgap [14]. The equivalent oxide thickness was extracted from the C-V measurement by considering the quantum effect [15]. The charge centroid (Z_{eff}) for the charge trap operation was evaluated utilizing HP8110A, Keithley 6517A, and KEYSIGHT DAQ970A [16]. The crystallinity was evaluated by the x-ray diffraction (XRD).

3. Results and Discussion

Figure 3 shows the PMA duration dependence of the C-V and J-V characteristics for the $Al/HfN_{0.5}/HfN_{1.1}$(10 nm)/HfO_2/p-Si(100) FeNOS diodes. As shown in Fig. 3 (a), the minimum EOT of 4.5 nm was obtained with negligible hysteresis by the PMA at 350°C/5 min. The D_{it} was extracted as 5.3×10^{10} $eV^{-1}cm^{-2}$. The leakage current was decreased to 1×10^{-8} A/cm² at $V_G = -1$ V by the PMA at 350°C/5 min compared to the PMA at 350°C/1 min as shown in Fig. 3 (b).

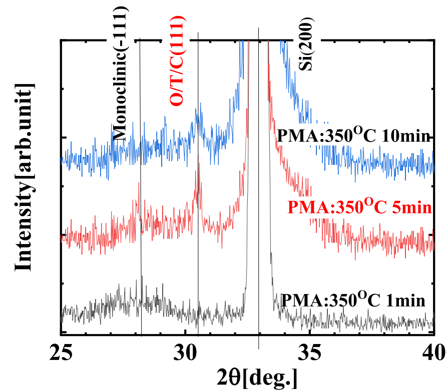
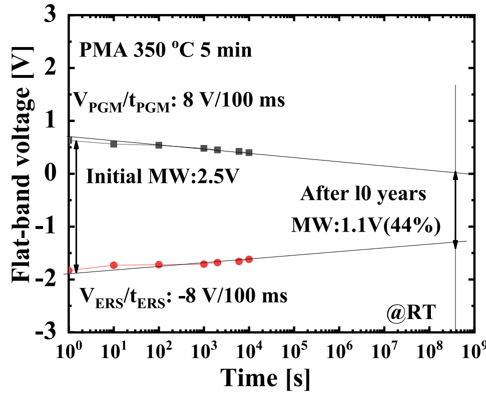


Fig. 4 PMA duration dependence on the XRD patterns for FeNOS structures. PMA was carried out at 350°C/1–10 min.

The leakage current was increased in case of the PMA at 350°C/10 min so that the PMA with long duration seemed to degrade the film quality even at the low annealing temperature such as 350°C. Figure 4 shows the XRD patterns of FeNOS structures. The peak intensity of orthorhombic HfO_2(111) was found to be increased by the PMA at 350°C/5 min, while it was decreased by the PMA at 350°C/10 min. Therefore, the 350°C for 5 min seemed to be the optimum PMA condition for the FeNOS structures.

Figure 5 shows the retention characteristic for the charge trap operation of FeNOS diode with PMA at 350°C/5 min. The schematic measurement sequence was also shown. The P/E input pulses, V_{PGM}/t_{PGM} and V_{ERS}/t_{ERS}, were



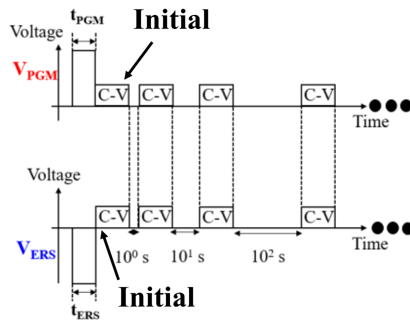


Fig. 5 Retention characteristic for charge trap operation of FeNOS diode with PMA at 350°C/5 min. The input pulses were ±8 V/100 ms for charge trap operation.

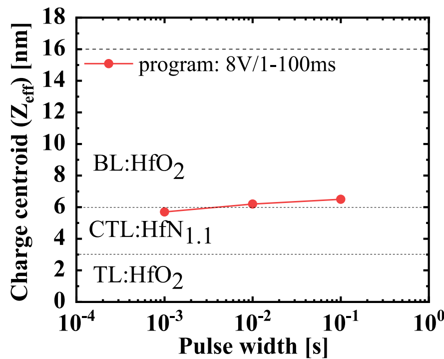
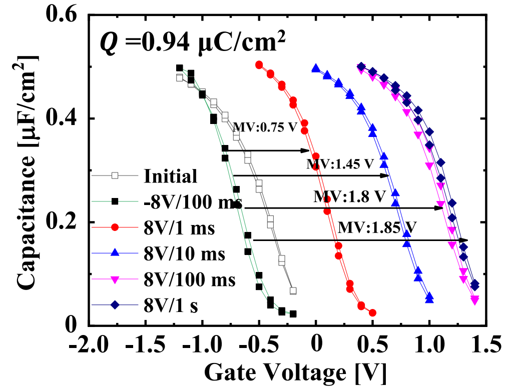


Fig. 6 Pulse width dependence on Z_{eff}.

V_{PGM}/t_{PGM}: 8 V/100 ms and V_{ERS}/t_{ERS}: −8 V/100 ms, respectively. The measurements were carried out until 10^4 s. The initial MW of 2.5 V was observed after P/E input pulses were applied. The estimated MW of 1.1 V after 10 years was obtained which was 44% compared with the initial MW of 2.5 V. This result suggested that reliability of the obtained memory characteristics was good enough even though the annealing temperature was low as 350°C.

Next, the charge centroid (Z_{eff}) was evaluated by changing the program pulses as V_{PGM}/t_{PGM}: 8 V/1–100 ms. Figure 6 shows the pulse width dependence on the Z_{eff} of FeNOS diode. The Z_{eff} was extracted utilizing the following equation,



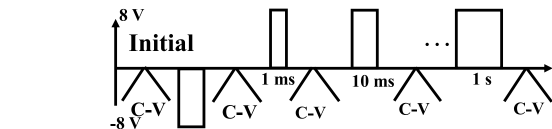


Fig. 7 The charge trap operation utilizing program pulse of V_{PGM}/t_{PGM} = 8 V/1 ms–1 s. The C-V characteristics were measured at 100 kHz. The schematic measurement sequence was also shown.

$$Z_e = \frac{\varepsilon_{ox}\Delta V_{FB}}{\int_0^{V_{FB}} C(V)dV + Q_m}$$

where Q_m is the measured charge, ε_{ox} is the dielectric constant of HfO_2 BL, and V_{FB} is the flat-band voltage.

As shown in Fig. 6, the Z_{eff} was located at the interface of FeND-HfO_2 BL and $HfN_{1.1}$ CTL even for the program pulse of V_{PGM}/t_{PGM}: 8 V/1 ms. Interestingly, the Z_{eff} was not markedly changed for the longer pulse such as V_{PGM}/t_{PGM}: 8 V/100 ms. This is probably because the density of trap sites in the $HfN_{1.1}$ CTL is large enough to accept the charge injection by the program conditions.

Finally, the charge trap and partial polarization operations were examined utilizing Al/$HfN_{0.5}$/$HfN_{1.1}$(15 nm)/HfO_2/p-Si(100) FeNOS diodes. Figure 7 shows the charge trap operation utilizing program pulses of V_{PGM}/t_{PGM}: 8 V/1 ms–1 s. As shown in Fig. 7, 2 bit/cell operation was demonstrated by the input pulses of V_{PGM}/t_{PGM}: 8 V/1–100 ms after the initialization by the input pulse of V_{ERS}/t_{ERS}: −8 V/100 ms with the maximum MW of 1.85 V. Negligible hysteresis was observed for each C-V characteristic after the P/E operations. When the input pulse of 8 V/1 s was applied, the MW was almost same with that of after 8 V/100 ms was applied so that the maximum available charge densities in the $HfN_{1.1}$ CTL was estimated as 0.94 μC/cm^2. From the obtained results, the margin of V_{FB} between each state is large enough so that the further multi-bit/cell operation such as 3 bit or 4 bit/cell operation seems to be available for the FeNOS fabricated in this research.

Next, the V_{FB} control by the partial polarization of FeND-HfO_2 BL was examined utilizing P/E pulses of V_{PGM}/t_{PGM}: −3 V/100 ms and V_{ERS}/t_{ERS}: 8 V/100 ms at ‘11’ and ‘01’ states of charge trap operations. Figure 8 clearly shows that the precise V_{FB} control by the partial po-

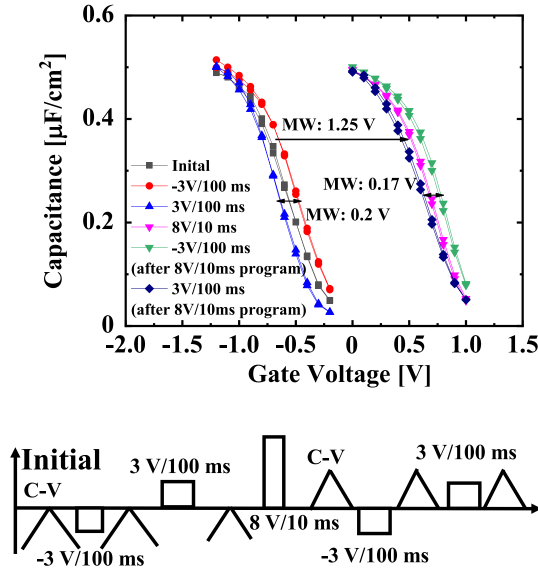


Fig. 8 The partial polarization operation utilizing P/E pulses of ±3 V/100 ms. The C-V characteristics were measured at 100 kHz. The schematic measurement sequence was also shown.

larization. The erase pulse caused the negative V_{FB} shift at each state, while the program pulses made V_{FB} shifted to the positive direction. The V_{FB} shift was approximately 80–100 mV. The MW of charge trap operation is 1.8 V so that 18–22 states control would be realized by the partial polarization operation.

4. Conclusions

In this paper, we have investigated the digital/analog-operation of Hf-based FeNOS diode. The low-voltage input pulse operation was found to control the partial polarization, and the V_{FB} shifts of approximately 80–100 mV were realized without causing the charge trap and/or detrap in the $HfN_{1.1}$ CTL. The V_{FB} control by the partial polarization is also applicable for the V_{TH} adjustment after the NVM fabrication. In conclusion, Hf-based FeNOS NVM is a promising memory device not only for storage memory but the in-memory computing applications.

Acknowledgments

This work was partially supported by JSPS KAKENHI Grant Number 19H00758, NEDO, JST, and CASIO Foundation. The author thanks Mr. W. Zhang and Mr. E.K. Hong for their help in this research.

References

[1] W. Shim and S. Yu, "Technological design of 3D NAND-based compute-in-memory architecture for GB-scale deep neural network," IEEE Electron Dev. Lett., vol.42, no.2, pp.160–163, 2021. DOI: 10.1109/LED.2020.3048101

[2] C.-C. Hsieh, H.-T. Lue, Y.-C. Li, S.-N. Hung, C.-H. Hung, K.-C. Wang, and C.-Y. Lu, "Chip demonstration of a high-density (43Gb) and high-search-bandwidth (300Gb/s) 3D NAND based in-memory search accelerator for ternary content addressable memory (TCAM) and proximity search of hamming distance," VLSI Symp., Tech. Dig., T15-1, 2023.

[3] S. Kudoh and S. Ohmi, "Multi-level 2-bit/cell operation utilizing Hf-based MONOS nonvolatile memory," 76th Device Res. Conf., Conf. Dig., pp.157–158, 2018.

[4] S. Ohmi, Y. Horiuchi, H. Morita, A. Ihara, and J.Y. Pyo, "HfN multi charge trapping layers for Hf-based metal-oxide-nitride-oxide-Si nonvolatile memory," Jpn. J. Appl. Phys., vol.60, SBBB03, 2021. DOI: 10.35848/1347-4065/abe09f

[5] Y. Seo, H.-M. An, M.Y. Song, and T.G. Kim, "Charge trap flash memory using ferroelectric materials as a blocking layer," Appl. Phys. Lett., vol.100, 173507, 2012. DOI: 10.1063/1.4705411

[6] T. Böscke, J. Müller, D. Bräuhaus, U. Schröder, and U. Böttger, "Ferroelectricity in hafnium oxide thin films," Appl. Phys. Lett., vol.99, 102903, 2011. DOI: 10.1063/1.3634052

[7] H. Ji, Y. Wei, X. Zhang, and R. Jiang, "Improvement of charge injection using ferroelectric Si:HfO_2 as blocking layer in MONOS charge trapping memory," IEEE J. Electron Devices Soc., vol.6, pp.121–125, 2018. DOI: 10.1109/JEDS.2017.2785304

[8] E.J. Shin, G. Lee, S. Kim, J.H. Chu, and B.J. Cho, "Dual-mechanism memory combining charge trapping and polarization switching for wide memory window flash cell," IEEE Electron Dev. Lett., vol.44, no.7, pp.1108–1111, 2023. DOI: 10.1109/LED.2023.3282366

[9] S. Ohmi, H. Morita, M. Hayashi, A. Ihara, and J.Y. Pyo, "Ferroelectric nondoped HfO_2 blocking layer formation for Hf-based FeNOS analog memory applications," 78th Device Res. Conf., Conf. Dig., pp.67–68, 2021. DOI: 10.1109/DRC52342.2021.9467182

[10] J.W. Shin, M. Tanuma, J. Pyo, and S. Ohmi, "Ultrathin ferroelectric nondoped HfO_2 for MFSFET with high-speed and low-voltage operation," 80th Device Res. Conf., Conf. Dig., pp.73–74, 2021.

[11] W. Zhang, M. Tanuma, J.W. Shin, and S. Ohmi, "PMA condition dependence on the FeNOS diodes with ferroelectric non-doped HfO_2 blocking layer," Asia-Pacific Workshop on Fundamentals and Applications of Advanced Semiconductor Devices, Workshop Dig., pp.142–143, 2022.

[12] J.-W. Shin, M. Tanuma, and S. Ohmi, "MFSFET with 5nm thick ferroelectric nondoped HfO_2 gate insulator utilizing low power sputtering for Pt gate electrode deposition," IEICE Trans. Electron., vol.E105-C, no.10, pp.578–583, Oct. 2022. DOI: 10.1587/transele.2021FUP0003

[13] S. Ohmi, E.K. Hong, S. Awakura, and Y. Sekiguchi, "Digital/analog mixed-operation of Hf-based FeNOS nonvolatile memory with ferroelectric nondoped HfO_2 blocking layer," 2023 Asia-Pacific Workshop on Advanced Semiconductor Devices, Workshop Dig., pp.118–119, 2023.

[14] L.M. Terman, "An investigation of surface states at a silicon/silicon oxide interface employing metal-oxide-silicon diodes," Solid-State Electronics, vol.5, no.5, pp.285–299, 1962. DOI: 10.1016/0038-1101(62)90111-9

[15] S. Saito, K. Torii, M. Hiratani, and T. Onai, "Analytical quantum mechanical model for accumulation capacitance of MOS structures," IEEE Electron Device Lett., vol.23, no.6, pp.348–350, 2002. DOI: 10.1109/LED.2002.1004231

[16] S. Ohmi, Y. Horiuchi, and S. Kudoh, "Improvement of Hf-based metal/oxide/nitride/oxide/Si nonvolatile memory characteristics by Si surface atomically flattening," Jpn. J. Appl. Phys., vol.59, SGGB10, 2020. DOI: 10.35848/1347-4065/ab70ad

Shun-ichiro Ohmi received his Ph.D. degree in Applied Electronics from the Tokyo Institute of Technology, Yokohama, Japan, in 1996. In 1996, he joined Tokyo Institute of Technology, where he has been engaged in the research of Si and organic semiconductor devices and processes. From 1997 to 1999, he was a Post Doctoral Technical Staff with the Lucent Technologies Bell Laboratories, where he was engaged in the research of $TiSi_2$ and $CoSi_2$ salicide processes. In 2000, he was a visiting researcher with North Carolina State University where he was engaged in the research of high-k gate insulator. Dr. Ohmi is a member of the IEEE Electron Devices Society, the Japan Society of Applied Physics, and a Fellow of the Institute of Electrical Engineers of Japan.

BRIEF PAPER *Special Section on Fundamentals and Applications of Advanced Semiconductor Devices*

Reduced Peripheral Leakage Current in Pin Photodetectors of Ge on n^+-Si by P^+ Implantation to Compensate Surface Holes

Koji ABE[†], **Mikiya KUZUTANI**[†], **Satoki FURUYA**[†], **Jose A. PIEDRA-LORENZANA**[†], **Takeshi HIZAWA**[†], ***Nonmembers*, *and* Yasuhiko ISHIKAWA**[†a)], ***Member***

SUMMARY A reduced dark leakage current, without degrading the near-infrared responsivity, is reported for a vertical pin structure of Ge photodiodes (PDs) on n^+-Si substrate, which usually shows a leakage current higher than PDs on p^+-Si. The peripheral/surface leakage, the dominant leakage in PDs on n^+-Si, is significantly suppressed by globally implanting P^+ in the i-Si cap layer protecting the fragile surface of i-Ge epitaxial layer before locally implanting B^+/BF_2^+ for the top p^+ region of the pin junction. The P^+ implantation compensates free holes unintentionally induced due to the Fermi level pinning at the surface/interface of Ge. By preventing the hole conduction from the periphery to the top p^+ region under a negative/reverse bias, a reduction in the leakage current of PDs on n^+-Si is realized.

key words: *Ge epitaxial layer, near-infrared photodetector, Si photonics, peripheral leakage current*

1. Introduction

A near-infrared (NIR) photodiode (PD) using a Ge epitaxial layer on Si (or Si-on-insulator, SOI) is one of the fundamental building blocks in Si photonics [1]. Ge exhibits a substantial optical absorption coefficient (>1000 cm^{-1}) at the NIR optical communication wavelengths (1.3–1.6 μm) as well as the good compatibility with the Si CMOS process. A vertical pin structure of Ge PD has been integrated with a Si optical waveguide on an SOI wafer for the on-chip optical communication [2]–[7], although a lateral pin structure has recently attracted interests regarding ultrahigh frequency response over 50 GHz [8], [9] to enhance the communication capacity. The vertical pin structure also plays a significant role in free-space/normal-incidence PD arrays required in applications such as multicore-fiber communications [10] and NIR sensing including the eye-safe light detection and ranging (LiDAR) [11]. In vertical pin PDs, an n^+-Ge/i-Ge/p^+-Si structure (PD on p^+-Si) has been commonly used because of the dark leakage current lower than that in an inversely ordered structure of p^+-Ge/i-Ge/n^+-Si (PD on n^+-Si) [12], [13]. However, the inverted configuration of PD on n^+-Si is convenient for the connection to the next stage circuit such as a transimpedance amplifier, which is usually designed to receive a positive photocurrent from the p contact. As previously reported by the authors [14], the higher leakage current in the PDs on n^+-Si is ascribed to the peripheral/surface leakage, rather than the leakage due to the thermal generation of carriers via gap defects of threading dislocations in the i-Ge layer on Si, which is dominant in the PDs on p^+-Si [15]. The peripheral leakage in the PDs on n^+-Si is probably derived from free holes [14] unintentionally induced due to the Fermi level pinning at the surface/interface of Ge [16], [17]. These holes induce an electrical conduction from the periphery to the top p^+ region of the PD on n^+-Si under a negative/reverse bias, whereas such a conduction is suppressed in the case of the PDs on p^+-Si because of the top n^+ region electrically isolated from the peripheral holes via a pn junction. One way to suppress the peripheral leakage in the PDs on n^+-Si is to use a mesa structure [14], while the footprint of the photodetection area is reduced because inclined {113} facet sidewalls surround a Ge mesa on (001) Si by a selective growth [1].

In this work, a reduction in the leakage current of non-mesa/planar Ge PDs on n^+-Si is realized by a P^+ implantation globally in the i-Si cap layer on the i-Ge epitaxial layer before locally forming the top p^+ region of the pin junction. The peripheral holes are compensated, leading to a reduced leakage current comparable to PDs on p^+-Si.

2. Holes Induced at Ge Surface/Interface

To verify the presence of free holes at the Ge surface/interface, the Hall effect measurements were performed for Ge epitaxial layers grown by ultrahigh-vacuum chemical vapor deposition (UHV-CVD) with 9%-GeH_4/Ar as a source gas. Here, an undoped Ge layer of 500 nm in thickness was prepared on a 6-inch bonded SOI wafer having a 250-nm-thick top (001) p-Si layer (approximately 10 Ω cm) and a 3-μm-thick buried SiO_2 layer. As in previous studies [13]–[15], a low/high temperature two-step growth was carried out to obtain a Ge layer uniform in thickness, i.e., a buffer layer (~50 nm) of elemental Ge was grown at a low temperature of 370°C, followed by a growth at an elevated temperature of 600°C. After the growth, the wafer was cut into pieces, and several different thicknesses of Ge were prepared by a wet etching in a H_2O_2 solution. All samples showed the p-type conduction, independent of the thickness, and the sheet resistances were as low as 20 kΩ/sq., which is approximately one order of magnitude lower than 400 kΩ/sq. for the top Si layer of the SOI substrate. Thus, the observed p-type conduction is ascribed to holes in Ge.

Manuscript received August 29, 2023.
Manuscript revised February 7, 2024.
Manuscript publicized May 15, 2024.

[†]Department of Electrical and Electronic Information Engineering, Toyohashi University of Technology, Toyohashi-shi, 441–8580 Japan.
a) E-mail: ishikawa@ee.tut.ac.jp

DOI: 10.1587/transele.2023FUS0001

Copyright © 2024 The Institute of Electronics, Information and Communication Engineers

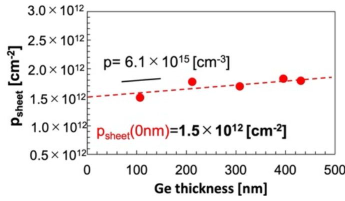


Fig. 1 Sheet hole densities of the Ge epitaxial layer as a function of the Ge thickness.

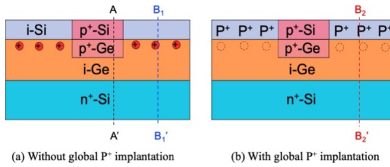


Fig. 2 Schematic cross-sections of PDs. (a) PD without the global P^+ implantation and (b) PD with the global P^+ implantation to compensate the free holes at the Ge surface.

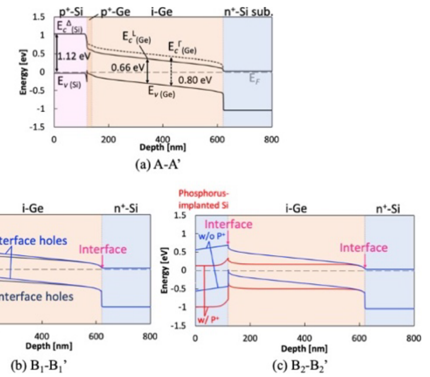


Fig. 3 Calculated energy band diagrams for (a) the main part of p^+-Si (120 nm)/p^+-Ge (20 nm)/i-Ge (480 nm)/n^+-Si PD (along the line A-A' in Fig. 2 (a)), (b) the i-Si/i-Ge/n^+-Si structure comparing the presence and absence of the holes at the top and bottom interfaces with the sheet density of 2×10^{12} cm^{-2} (along the line B_1-B_1' in Fig. 2 (a)), and (c) the n-Si (P^+ implanted)/i-Ge/n^+-Si structure (along the line B_2-B_2' in Fig. 2 (b)) compared with the i-Si/i-Ge/n^+-Si structure. A uniform P^+ density of 2.0×10^{17} cm^{-3} in the top Si layer was assumed in (c).

Figure 1 shows the sheet hole density as a function of the Ge thickness. The hole density slightly decreased with decreasing the Ge thickness. The slope, or the hole density per unit volume in the undoped Ge layer was estimated as low as 6×10^{15} cm^{-3}. More importantly, there is an offset in the sheet hole density at the Ge thickness of 0 nm, which is approximately 1.5×10^{12} cm^{-2}. This indicates the presence of free holes at the top Ge surface and/or bottom Ge/Si interface, although the densities cannot be divided between the top surface and bottom interface.

The peripheral leakage in the PDs on n^+-Si is attributed to the holes at the top Ge surface [14], as shown in Fig. 2 (a). These holes easily induce a leakage current from the periphery to the top p^+ region of the PD under a negative/reverse bias. Typical calculated energy band diagrams along the lines A-A' and B_1-B_1' of Fig. 2 (a) are shown in Figs. 3 (a) and 3 (b), respectively. Here, the bandgap narrowing of a few 10 meV in Ge was ignored, which is induced by the tensile in-plane lattice strain of about 0.2% in Ge due to the thermal expansion mismatch with the Si substrate [1], [14], [15]. Similar to the main part of p^+-Si/p^+-Ge/i-Ge/n^+-Si PD in Fig. 3 (a), the diagram at the periphery in Fig. 3 (b) shows that the Fermi level is located near the valence band maximum of Ge at the top i-Si/i-Ge interface in the presence of the interface holes with the sheet density of 2.0×10^{12} cm^{-2}. It is mentioned that the holes at the bottom interface are compensated with electrons in the n^+-Si substrate, i.e., there is no significant difference in the band diagram at around the bottom interface between the presence and absence of the interface holes, as shown in Fig. 3 (b).

To suppress the peripheral leakage by the free holes at the top i-Si/i-Ge interface, an implantation of P^+ in a Si cap layer on Ge should be effective to compensate the holes, as shown in Fig. 2 (b). The band diagram along the line B_2-B_2' of Fig. 2 (b) is shown in Fig. 3 (c). Here, a uniform P^+ density of 2.0×10^{17} cm^{-3} was assumed in the top n-Si layer of 120 nm in thickness. This corresponds to the sheet P^+ density of 2.4×10^{12} cm^{-2}, almost equal to the interface hole density of 2.0×10^{12} cm^{-2}. In contrast to the absence of the P^+ implantation, the Fermi level is not located near the valence band maximum of Ge but located near the midgap. This indicates the compensation of the interface holes, potentially preventing the peripheral leakage.

3. PD Preparation

Ge PDs on n^+-Si were prepared as follows. As the starting substrate, an Sb-doped n^+-Si (001) wafer was used, whose resistivity was 0.008–0.020 Ω cm. First, an undoped Ge epitaxial layer of 500 nm in thickness was grown on n^+-Si by UHV-CVD. The growth condition was exactly the same as that for the Hall effect measurements in the previous section. Subsequently, a Si cap layer of 120 nm in thickness was grown at 600°C for protecting the fragile surface of Ge using a source gas of 10%-Si_2H_6/Ar. A post-growth annealing was performed at 800°C for 10 min in N_2 to reduce the threading dislocation density in Ge [15].

To compensate holes at the Ge surface (Si cap/Ge interface), a P^+ implantation was performed in the Si cap layer over the entire surface, or globally. Based on the Hall effect measurements, a dose of 2.0×10^{12} cm^{-2} was used, although a different dose of 4.0×10^{12} cm^{-2} was also used for comparison. The acceleration voltage was 35 kV, leading to the P^+ distribution mostly in the Si cap layer, as simulated in Fig. 4. Then, to form a vertical pin junction, an implantation of B^+ and BF_2^+ was performed locally with resist masks. The B^+ and BF_2^+ were implanted at 10 kV with doses of 2.0×10^{14} cm^{-2} and 1.0×10^{14} cm^{-2}, respectively. As shown in Fig. 4, because of the difference in the penetration depth, the B^+ ions were implanted across the Si cap/Ge interface to form the p region in the Ge layer, while the BF_2^+ ions were distributed near the surface of the Si cap layer, realizing a low resistance with a metal contact. The implanted area, cor-

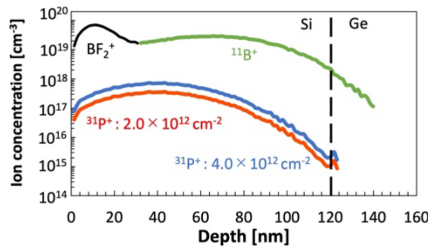


Fig. 4 Simulated distributions of P^+, B^+ and BF_2^+ ions.

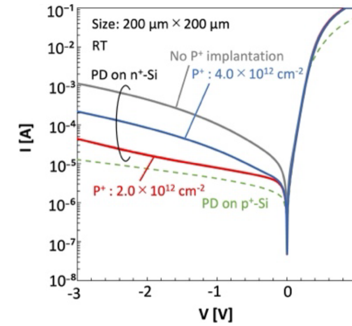


Fig. 5 Typical I-V characteristics of the fabricated Ge PDs.

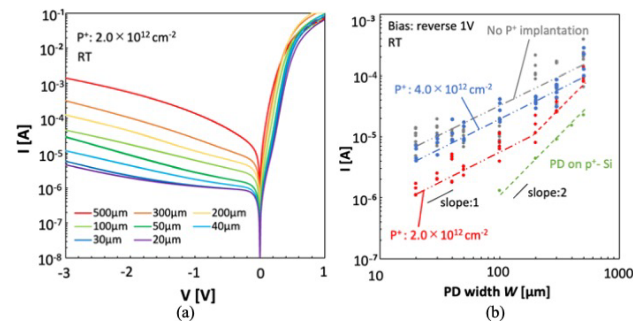


Fig. 6 (a) Typical I-V characteristics of the Ge PDs on n^+-Si (the P^+ implantation dose of 2.0×10^{12} cm^{-2}) with different PD widths, and (b) the leakage current at the reverse voltage of 1 V as a function of the PD width.

responding to the junction/PD area, was square-shaped with different widths W of 20–500 μm. An activation annealing was performed at 600°C for 5 min in N_2. Finally, metal electrodes of Al/Ti (with an illumination window opening) were formed. In addition to the PDs on n^+-Si, Ge PDs on p^+-Si were prepared as references.

Current-voltage (I-V) characteristics were measured at room temperature (RT) under dark to investigate the leakage current. I-V curves under a light illumination of 1.23 mW at 1550 nm were also measured. Furthermore, responsivity spectra were obtained at 1455–1640 nm.

4. Results and Discussion

Figure 5 shows typical I-V characteristics at RT for PDs on n^+-Si with and without the P^+ implantation together with a reference PD on p^+-Si. The width of the squared-shaped PD was 200 μm. Rectifying diode properties were obtained, although the reverse leakage current was different between the PDs. It is important that the P^+ implantation in the Si cap layer of the PDs on n^+-Si successfully reduced the dark leakage current. The lower dose of 2.0×10^{12} cm^{-2} more efficiently reduced the leakage current, which was comparable to that in the PD on p^+-Si.

To examine the leakage mechanism, I-V characteristics for different PD widths were obtained. A typical width dependence is shown in Fig. 6 (a) for the PDs on n^+-Si with the P^+ implantation dose of 2.0×10^{12} cm^{-2}. The reverse leakage decreased with decreasing the PD width. In Fig. 6 (b), the leakage current at the reverse voltage of 1 V was plotted as a function of the width W. The leakage current I should be composed of peripheral and areal components, and is expressed in the case of the square-shaped PD as

$$I = J_{periphery} 4W + J_{area} W^2, \quad (1)$$

where $J_{periphery}$ and J_{area} correspond to the peripheral leakage current per unit length and the one per unit area, respectively. In the logarithmic plot of Fig. 6 (b), the slope of 2 corresponds to the leakage current dominated by the area of the pin junction, whereas the slope of 1 corresponds to the leakage current dominated by the peripheral length. For the reference PDs on p^+-Si, the slope was almost equal to 2, indicating that the leakage current is dominated by the areal leakage. As in the previous report [15], the areal leakage is derived from the thermal generation of carriers via gap defects of threading dislocations in the i-Ge layer.

Table 1 A comparison of leakage current of the PDs at 1-V reverse bias.

Sample		$J_{periphery}$ [mA/cm]	J_{area} [mA/cm²]
PD on p^+-Si		N.A.	9.3
PD on n^+-Si	No P^+ implantation	3.4	N.A.
	P^+ : 2.0×10^{12} cm^{-2}	0.6	30.0
	P^+ : 4.0×10^{12} cm^{-2}	2.0	N.A.

In contrast, the PDs on n^+-Si without the P^+ implantation as well as those with the P^+ implantation dose of 4.0×10^{12} cm^{-2} showed the slope of approximately 1, indicating a high peripheral leakage. However, the P^+ implantation with the dose of 2.0×10^{12} cm^{-2} showed the slope of 2 for the width as large as 200 μm or larger, despite the slope of 1 for the smaller ones. This is a clear evidence of the reduction in the peripheral leakage by the appropriate P^+ implantation. As summarized in Table 1, $J_{periphery}$ was significantly reduced from 3.4 mA/cm (without the P^+ implantation) to 0.6 mA/cm by implanting P^+ with the dose of 2.0×10^{12} cm^{-2}. However, no significant reduction in the leakage current was observed for the higher dose of 4.0×10^{12} cm^{-2}. This is ascribed to the formation of electron channels because of an overdose to compensate the holes. The implantation dose and/or depth should be optimized for further reducing the leakage current.

To examine the effect of the global P^+ implantation on the photodetection efficiency, I-V curves were obtained under an NIR light illumination, as shown in Fig. 7. Independent of the P^+ implantation, the light illumination at the wavelength of 1550 nm increased the reverse current due to the photocurrent, although at higher reverse voltage (V <

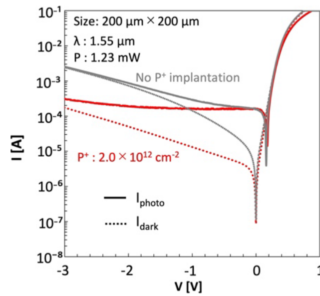


Fig. 7 Typical I-V characteristics of the 200-µm-wide Ge PDs on n^+-Si with and without an NIR light illumination (1.23 mW at 1550 nm).

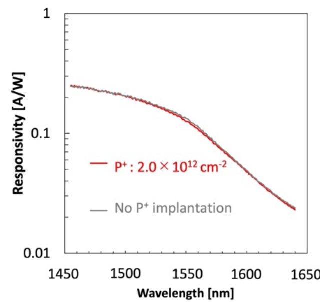


Fig. 8 Typical responsivity spectra at the reverse voltage of 1 V.

−2 V), the current increase was smeared for the PD without the P^+ implantation due to a rapid increase in the leakage current. The responsivity at 1550 nm was estimated to be 0.13 A/W, which is reasonable for the Ge absorption layer as thin as 500 nm. Furthermore, the responsivity spectra at 1455–1640 nm in Fig. 8 were identical between the PDs with and without the P^+ implantation. Therefore, no degradation in the spectral responsivity occurs by the global P^+ implantation, while reducing the dark leakage current in the Ge PDs on n^+-Si.

5. Conclusion

A reduced dark leakage current, without degrading the NIR responsivity, was realized for a vertical pin structure of the Ge PDs on n^+-Si substrate. In particular, the peripheral/surface leakage was significantly suppressed by globally implanting P^+ in the i-Si cap layer protecting the fragile surface of i-Ge epitaxial layer before locally implanting B^+/BF_2^+ for the top p^+ region. The P^+ implantation compensates holes at the top Ge surface, preventing the hole conduction from the periphery to the top p^+ region under a negative/reverse bias.

Acknowledgements

This work was partly supported by JSPS KAKENHI Grant Numbers JP21H01367 and JP22H01555.

References

[1] J. Michel, J. Liu, and L.C. Kimerling, “High-performance Ge-on-Si photodetectors,” Nat. Photon., vol.4, no.8, pp.527–534, Aug. 2010.

[2] D. Ahn, C.-Y. Hong, J. Liu, W. Giziewicz, M. Beals, L.C. Kimerling, and J. Michel, “High performance, waveguide integrated Ge photodetectors,” Opt. Express, vol.15, no.7, pp.3916–3921, April 2007.

[3] T. Yin, R. Cohen, M.M. Morse, G. Sarid, Y. Chetrit, D. Rubin, and M.J. Paniccia, “31 GHz Ge n-i-p waveguide photodetectors on silicon-on-insulator substrate,” Opt. Express, vol.15, no.21, pp.13965–13971, Oct. 2007.

[4] L. Vivien, J. Osmond, J.-M. Fédéli, D. Marris-Morini, P. Crozat, J.-F. Damlencourt, E. Cassan, Y. Lecunff, and S. Laval, “42 GHz pin Germanium photodetector integrated in a silicon-on-insulator waveguide,” Opt. Express, vol.17, no.8, pp.6252–6257, April 2009.

[5] K.-W. Ang, T.-Y. Liow, M.-B. Yu, Q. Fang, J. Song, G.-Q. Lo, and D.-L. Kwong, “Low thermal budget monolithic integration of evanescent-coupled Ge-on-SOI photodetector on Si CMOS platform,” IEEE J. Select. Topics Quantum Electron., vol.16, no.1, pp.106–113, Jan./Feb. 2010.

[6] S. Park, T. Tsuchizawa, T. Watanabe, H. Shinojima, H. Nishi, K. Yamada, Y. Ishikawa, K. Wada, and S. Itabashi, “Monolithic integration and synchronous operation of germanium photodetectors and silicon variable optical attenuators,” Opt. Express, vol.18, no.8, pp.8412–8421, April 2010.

[7] H. Nishi, T. Tsuchizawa, R. Kou, H. Shinojima, T. Yamada, H. Kimura, Y. Ishikawa, K. Wada, and K. Yamada, “Monolithic integration of a silica AWG and Ge photodiodes on Si photonic platform for one-chip WDM receiver,” Opt. Express, vol.20, no.8, pp.9312–9321, April 2012.

[8] J. Fujikata, M. Noguchi, K. Kawashita, R. Katamawari, S. Takahashi, M. Nishimura, H. Ono, D. Shimura, H. Takahashi, H. Yaegashi, T. Nakamura, and Y. Ishikawa, “High-speed Ge/Si electro-absorption optical modulator in C-band operation wavelengths,” Opt. Express, vol.28, no.22, pp.33123–33134, Oct. 2020.

[9] S. Lischke, A. Peczek, J.S. Morgan, K. Sun, D. Steckler, Y. Yamamoto, F. Korndörfer, C. Mai, S. Marschmeyer, M. Fraschke, A. Krüger, A. Beling, and L. Zimmermann, “Ultra-fast germanium photodiode with 3-dB bandwidth of 265 GHz,” Nat. Photon., vol.15, no.12, pp.925–931, Dec. 2021.

[10] T. Umezawa, T. Sakamoto, A. Kanno, N. Yamamoto, and T. Kawanishi, “High Speed 2-D Photodetector Array for Space and Mode-Division Multiplexing Fiber Communications,” J. Lightwave Technol., vol.36, no.17, pp.3684–3692, Sept. 2018.

[11] C.-L. Chen, S.-W. Chu, B.-J. Chen, Y.-F. Lyu, K.-C. Hsu, C.-F. Liang, S.-S. Su, M.-J. Yang, C.-Y. Chen, S.-L. Cheng, H.-D. Liu, C.-T. Lin, K.P. Petrov, H.-W. Chen, K.-C. Chu, P.-C. Wu, P.-T. Huang, N. Na, and S.-L. Chen, “An Up-to-1400 nm 500 MHz Demodulated Time-of-Flight Image Sensor on a Ge-on-Si Platform,” 2020 IEEE Int. Conf. Solid-State Circuits (ISSCC), 5.3, Feb. 2020.

[12] G. Masini, L. Colace, G. Assanto, H.-C. Luan, and L.C. Kimerling, “High-performance p-i-n Ge on Si photodetectors for the near infrared: from model to demonstration,” IEEE Trans. Electron Devices, vol.48, no.6, pp.1092–1096, June 2001.

[13] K. Ito, T. Hiraki, T. Tsuchizawa, and Y. Ishikawa, “Waveguide-integrated vertical pin photodiodes of Ge fabricated on p^+ and n^+ Si-on-insulator layers,” Jpn. J. Appl. Phys., vol.56, no.4S, 04CH05, April 2017.

[14] Y. Ishikawa, K. Ito, K. Noguchi, and T. Hiraki, “Near-infrared pin Photodetectors of Ge Epitaxial Layers for Si Photonics,” IEICE Trans. Electron. C, vol.J101-C, pp.374–380, Oct. 2018.

[15] Y. Ishikawa and S. Saito, “Ge-on-Si photonic devices for photonic-electronic integration on a Si platform,” IEICE Electron Express, vol.11, no.24, 20142008, Oct. 2014.

[16] J. Tersoff, “Schottky barrier heights and the continuum of Gap States,” Phys. Rev. Lett., vol.52, no.6, pp.465–468, Feb. 1984.

[17] H. Hasegawa and H. Ohno, “Unified disorder induced gap state model for insulator–semiconductor and metal–semiconductor interfaces,” J. Vac. Sci. Technol. B, vol.4, no.4, pp.1130–1138, July 1986.

BRIEF PAPER *Special Section on Fundamentals and Applications of Advanced Semiconductor Devices*

Electrical and X-Ray Photoelectron Spectroscopy Studies of Ti/Al/Ti/Au Ohmic Contacts to AlGaN/GaN

Hiroshi OKADA[†a)], ***Member***, **Mao FUKINAKA**[†], ***and*** **Yoshiki AKIRA**[†], ***Nonmembers***

SUMMARY Effects of Al thickness in Ti/Al/Ti/Au ohmic contact on AlGaN/GaN heterostructures are studied. Samples having Al thickness of 30, 90 and 120 nm in Ti/Al/Ti/Au have been investigated by electrical and X-ray photoelectron spectroscopy (XPS) depth profile analysis. It is found that thick Al samples show lower resistance and formation of Al-based alloy under the oxidized Al layer.

key words: *AlGaN/GaN heterostructure, ohmic contact, Ti/Al/Ti/Au, XPS*

1. Introduction

Wide band gap semiconductors such as GaN and related materials are promising materials for next-generation high-efficiency power electronic devices. Particularly, AlGaN/GaN heterostructures for high electron mobility transistor (HEMT) have the potential for high-speed and low sheet resistance devices due to the high mobility and high sheet carrier density two-dimensional electron gas (2DEG) at the interface. Low resistance ohmic contact formation is a crucial aspect in the fabrication of AlGaN/GaN devices as parasitic resistance resulting from the contact resistance directly affects the on-resistance of HEMTs. Among the various metal combinations used for ohmic contacts in AlGaN/GaN [1], Ti/Al/Ti/Au [2] is a widely used combination. A critical parameter in determining the characteristics of these ohmic contacts is the thickness of the metal layers. Thin metal layers offer advantages of cost and productivity, however, they may come with the disadvantage of increased contact resistance due to the influence of surface oxidation on the metal electrodes. In Ti/Al/Ti/Au, the thickness of Al seems a key parameter of the contact because a sufficient amount of Al-based alloy should be formed with overcoming the Al oxidation during/after the annealing. Due to the high-temperature annealing required to achieve good ohmic contacts with the AlGaN/GaN, the diffusion and intermixing (alloying) of metals and semiconductor materials are inevitable. An approach involving electrical and chemical analysis for samples with varying Al thickness is valuable for comprehending the formation of low resistance ohmic contacts. In this study, we investigate the Al thickness dependence on the Ti/Al/Ti/Au ohmic contact characteristics on AlGaN/GaN. Samples having Al thickness of 30, 90, and 120 nm in Ti/Al/Ti/Au on AlGaN/GaN were fabricated and characterized by electrical measurement and X-ray photoelectron spectroscopy (XPS).

2. Experimental

Figure 1 shows a schematic cross-section of AlGaN/GaN heterostructure used in this study. AlGaN/GaN heterostructure grown on Si (111) substrate grown by metal organic chemical vapor deposition (MOCVD) at NTT-AT was purchased and used. Ti/Al/Ti/Au ohmic electrode pattern was formed by the photolithography and lift-off process. Before the ohmic contact evaporation, the samples were treated with HCl:H_2O=1:3 at room temperature for 2 min to remove native surface oxide followed by a rinse in deionized water. Successful evaporation of Ti, Al, Ti and Au was carried out in this order without breaking the vacuum by a multi hearth e-beam evaporation system with base pressure below 5×10^{-5} Pa. The evaporation thickness of each metal was monitored by a quartz crystal thickness monitor, which was calibrated by measuring the step height of the single layer by the surface stylus meter. In this study, samples of A30, A90 and A120 having Al thickness of 30, 90 and 120 nm, respectively, were prepared and characterized. All samples have been thermally annealed at 900°C for 1 min in N_2 ambient in a rapid thermal annealing (RTA) furnace.

For evaluation of contact resistance of the fabricated samples based on the circular transmission line model (c-TLM) [3], a series of two-terminal resistances of various spacing, d, between the ohmic electrodes were obtained from current-voltage (I–V) characteristics measured by Keithley 4200 semiconductor characterization system.

XPS depth profile analysis was performed with the PHI

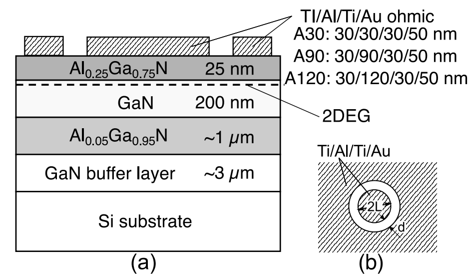


Fig. 1 (a) Cross-section of AlGaN/GaN heterostructure used in this study. Samples of A30, A90 and A120 having Al thickness of 30, 90 and 120 nm, respectively, were prepared and characterized. (b) Schematic of sample of circular transmission line model (c-TLM) structure.

Manuscript received September 2, 2023.
Manuscript revised February 13, 2024.
Manuscript publicized June 4, 2024.

[†]Toyohashi University of Technology (TUT), Toyohashi-shi, 441–8580 Japan.
a) E-mail: okada@las.tut.ac.jp

DOI: 10.1587/transele.2023FUS0002

Copyright © 2024 The Institute of Electronics, Information and Communication Engineers

Quantera SXM-CI (ULVAC PHI) using a monochromatic Al Kα ($h\nu$ = 1486.7 eV) X-ray source. Focused X-ray beam of 20 μm diameter irradiated at the center of the ohmic contact having 200 × 1000 μm. Sputtering for the depth profile analysis was performed with Ar^+ at 2 kV in 1 × 1 mm area for 90 s for each sputtering step. This sputtering condition is capable of an etching rate of 30 nm for the SiO_2. Because the sputtering rate depends on the target material, the depth may not proportional to the sputtering time for complex metal alloy, direct interpretation of the sputtering step to the depth is difficult. In this study, we estimate only the final etching depth roughly based on the appearance of the Ga-N bonding components in XPS Ga3d spectra. Since the Ga-N bonding components in the Ga3d spectra were observed at the final etching steps, a sufficient final etching depth more than 100 nm can be roughly estimated.

3. Results and Discussion

I–V characteristics of the fabricated samples were analyzed to investigate the dependence of the contact resistance on the Al thickness. Figure 2 presents a summary of the total resistance, R_T, obtained from the *I–V* measurements plotted against the spacing d in c-TLM samples of A30, A90 and A120. As shown in Fig. 2, we observed a reduction of the total resistance for the equivalent spacing with an increase in Al thickness. Based on the c-TLM model and using the sheet resistance of the 2DEG, R_S = 395 Ω/sq, from the same AlGaN/GaN wafer determined using the Van der Pauw method, we calculated specific contact resistance of 6×10^{-4}, 4×10^{-4}, 1×10^{-5} Ωcm^2 for A30, A90 and A120, respectively. These results demonstrate that the specific contact resistance are dependent on the Al thickness. Thinner Al contacts exhibits a disadvantage in achieving low contact resistance. This finding is consistent with the observations made by Meer *et al.*, who also noted that thicker Al layers result in smaller contact resistance [4].

To further explore the impact of Al thickness in ohmic contacts, we conducted an XPS depth profile analysis. In XPS measurement, we verified the binding energy (BE) values by referring the C1s peak at 285 eV of the surface contamination layer. Figure 3 shows Au4f spectra of annealed samples confirming the reproducibility of peak energy position. For all the samples, we observed distinct peak for $Au4f_{7/2}$ at 84.4 ± 0.2 eV, accompanied by $Au4f_{5/2}$ at around 88.1 eV. The slight shift in the peak position of $Au4f_{7/2}$ from the BE of pure Au of 83.98 eV is attributed to the presence of Au-Al alloy [5], a phenomenon also reported in TEM observations [6].

In Fig. 4, the depth profile of metals (Ti, Al, Au, Ga) obtained from XPS peak area and the atomic sensitivity factor is presented. It is worth mentioning that the profile for A90 (not shown) resembled that of A120. Due to the thicker Al in A120, there is noticeable increase in the fraction of Al compared with A30. Interestingly, the order of the layers observed in the depth profile does not align with the sequence of evaporation. For instance the top layer of Au extends into deeper region, a phenomenon attributed to the dynamic metallurgical process during RTA. Notably, in A30, a distinct

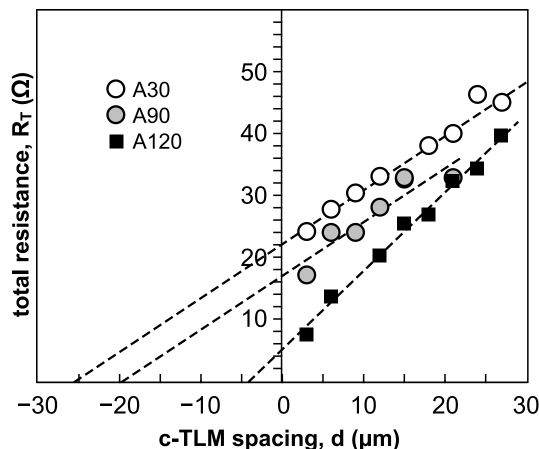


Fig. 2 Total resistance versus spacing (d) in c-TLM structure for sample A30, A90 and A120. Dashed lines represent fitted line using the least-square method.

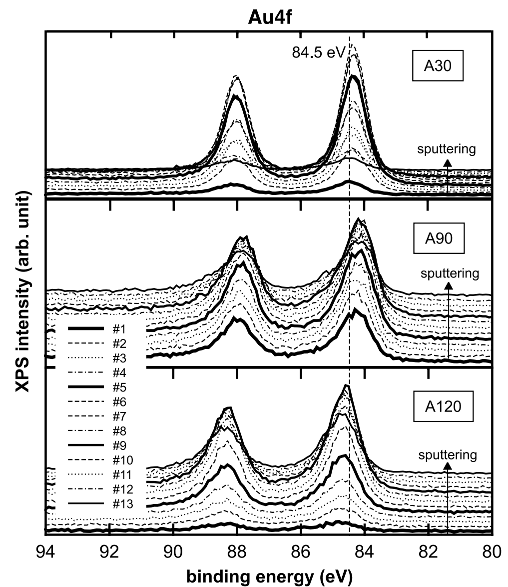


Fig. 3 Evolution of Au4f spectra of A30, A90 and A120. Each spectrum was shifted for clarity from lower to upper following the sputtering order.

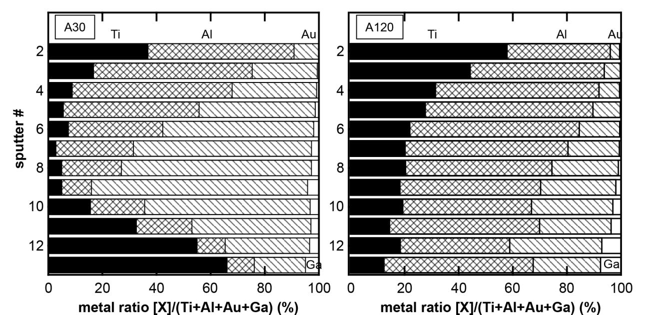


Fig. 4 Depth profile of metals in A30 and A120.

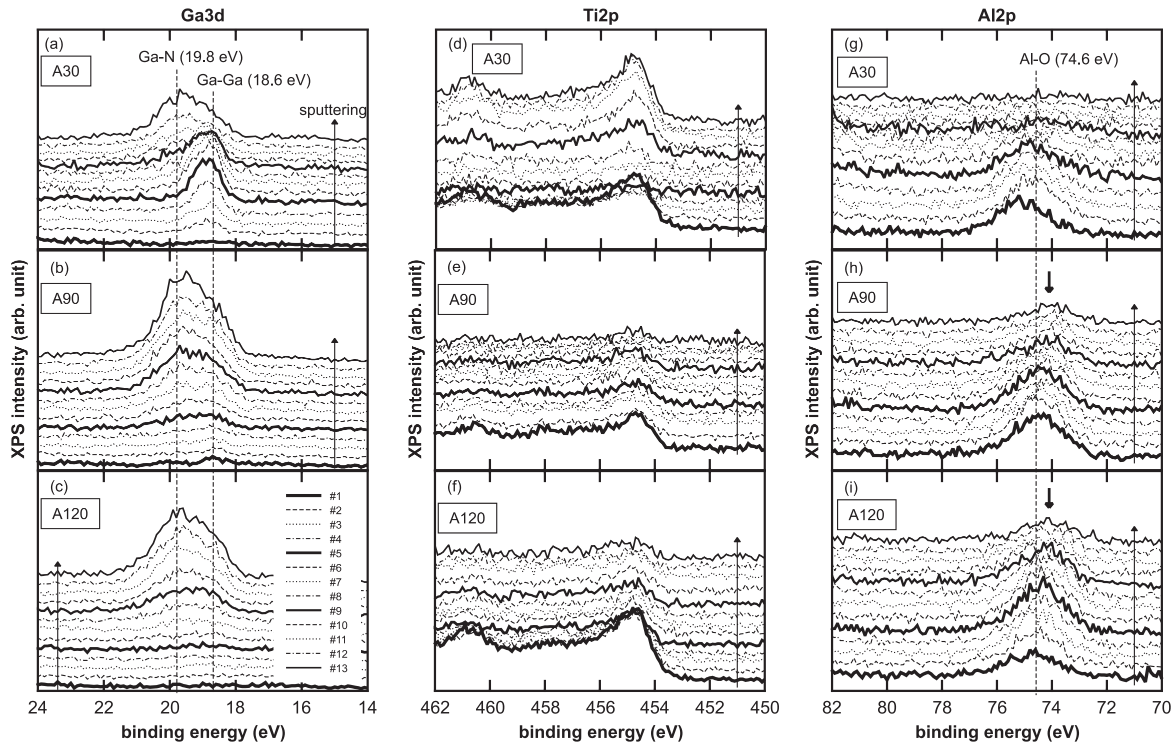


Fig. 5 XPS spectra of Ga3d, Ti2p and Al2p of sample A30, A90 and A120. For clarity, the spectra are shifted in vertically upward direction following the sputtering order.

distribution of Ti is observed in both shallower and deeper regions. However, no such accumulation of Ti in the deeper region is observed in A120. This observation is consistent with findings in Refs. [7], [8], which suggest that Ti-GaN exhibits aggressive reactivity, while the presence of Al can mitigate this reaction. This aligns with the present profile of A120, which shows a relatively even distribution of Ti.

Figure 5 presents the evolution of XPS spectra for Ga3d, Ti2p and Al2p from samples A30, A90 and A120. For clarity, the spectra are shifted in a vertically upward direction following the sputtering order. In the case of Ga3d, all samples exhibit a distinct peak at 19.8 eV in the spectra from the deeper region, indicating Ga-N bonds. Notably, during the middle of the evolution of Ga3d peak in A30, as shown in Fig. 5 (a), an pronounced peak at 18.6 eV emerges. This peak is attributed to Ga metal out diffusion from AlGaN/GaN structure during the RTA process [7], [9], [10]. In Figs. 5 (b) and (c), Ga3d spectra of A90 and 120, respectively, also suggest the presence of Ga metal in deeper region, as indicated by the broadened signal. However, the signal attributed to Ga metal components is less pronounced compared with the A30 in Fig. 5 (a).

In Figs. 5 (d), (e), and (f), the spectra of Ti2p exhibit a similar feature with peak around 455 eV, as well as its spin splitting peak at around 463 eV. These peaks are characterized by their broad nature and higher BE side tail. It is apparent that these spectra consist of multiple components, including Ti-N at BE of 456 eV, Ti-O at 458 eV [7] together with TIAlN at 455 eV [11]. These components appear consistently across all samples, regardless of the Al thickness.

In Figs. 5 (g), (h) and (i), the spectra of Al2p close to the surface showed a peak at 74.6 eV for all the samples, suggesting the presence of Al oxidation in the form of AlO_x [9], [12]. Particularly, the Al2p peak in A30 remained at the same BE, while the peak from the deeper region almost disappeared in the spectra near sputter #9. This is consistent with the higher resistance in A30, where the Al oxide is the dominant phase in the contact. On the other hand, thicker Al samples exhibited a departure from the BE peak at 74.6 eV towards lower BE values in the deeper region, as indicated by the arrow in Fig. 5 (h). This behavior is more pronounced in A120, as shown in Fig. 5 (i). For the thick Al samples, Al alloys are thought to be formed in the deeper region, where the influence of surface oxidation is less pronounced. Garbe *et al.* [13] reported that the formation of the metallic Al_3Ti alloy in annealed Al/TiN on GaN with thick Al, as observed in their TEM and XPS characterization with appearance of BE peak at 72.4 eV, which is 2 eV lower than that of Al nitride/oxide at 74.5 eV. TiAl alloy formation is also reported in TEM observation [14]. However, in the present study, the BE shift of 0.5 eV from the Al oxide peak position in A90 and A120 cannot be explained by the formation of Al_3Ti. A plausible explanation for observed BE shift in A90 and A120 is the formation of TiAlN. Greczynski *et al.* [15] have reported that $Ti_{1-x}Al_xN$ BE varies slightly from 74.1 with $x = 0.25$ to 74.4 eV with $x = 0.83$. Chawla *et al.* [16] have described that $Ti_{1-x}Al_xN$ exhibits lower resistivity for the film with $x < 0.83$ due to the insulating nature of AlN. Fur-

ther detailed studies are needed for the exact identification of the Al alloy under the oxidized surface, however, the formation of observed Al alloy in thick Al samples A90 and A120 plays an important role in reducing the contact resistance. The present results indicate that Ti/Al/Ti/Au ohmic contacts with thin Al have a disadvantage for achieving low contact resistance, and the situation improves with an increase in Al thickness.

Acknowledgments

This work was carried out using facilities of Venture Business Laboratory (VBL), Electronics-Inspired Interdisciplinary Research Institute (EIIRIS), and Cooperative Research Facility Center (CRFC) in Toyohashi University of Technology. This work was supported by JSPS KAKENHI Grant Number 20K04579, Suzuki foundation, Nitto foundation and Naito foundation.

References

[1] G. Greco, F. Iucolano, F. Roccaforte, "Ohmic contacts to gallium nitride materials," Appl. Surf. Sci., vol.383, pp.324–345, 2016.

[2] J. Hilsenbeck, W. Rieger, E. Nebauer, W. John, G. Tränkle, J. Würfl, A. Ramakrishan, and H. Obloh, "AlGaN/GaN HFETs with new ohmic and Schottky contacts for thermal stability up to 400°C," Phys. Stat. Sol. (a), vol.176, no.1, pp.183–187, 1999.

[3] D.K. Schroder, "Semiconductor Material and Device Characterization," pp.150–157, John Wiley & Sons, New Jersey, 2006.

[4] M. Meer, A. Rawat, K. Takhar, S. Ganguly, and D. Saha, "Interface dynamics in ohmic contact optimization on AlGaN/GaN heterostructure by the formation of TiN," Microelectronic Eng., vol.219, 111144, 2020.

[5] H. Piao, M.S. Fuller, D. Miller, and N.S. McIntyre, "A study of thin film Au-Al alloy oxidation in ambient air by X-ray photoelectron spectroscopy (XPS), X-ray absorption near edge structure (XANES), and secondary ion mass spectroscopy (SIMS)," Appl. Surf. Sci., vol.187, no.3-4, pp.266–274, 2002.

[6] A.N. Bright, P.J. Thomas, M. Weyland, D.M. Tricker, C.J. Humphreys, and R. Davies, "Correlation of contact resistance with microstructure for Au/Ni/Al/Ti/AlGaN/GaN ohmic contacts using transmission electron microscopy," J. Appl. Phys., vol.89, no.6, pp.3143–3150, 2001.

[7] M. Nozaki, J. Ito, R. Asahara, S. Nakazawa, M. Ishida, T. Ueda, A. Yoshigoe, T. Hosoi, T. Shimura, H. Watanabe, "Synchrotron radiation X-ray photoelectron spectroscopy of Ti/Al ohmic contacts to n-type GaN: Key role of Al capping layers in interface scavenging reactions," Appl. Phys. Express, vol.9, no.10, 105801, 2016.

[8] B. Van Daele, G. Van Tendeloo, W. Ruythooren, J. Derluynj, M.R. Leys, and M. Germain, "The role of Al on Ohmic contact formation on n-type GaN and AlGaN/GaN," Appl. Phys. Lett., vol.87, no.6, 061905, 2005.

[9] A. Akkaya, and E. Ayyıldız, "Depth profile study on Ti/Al bilayer Ohmic contacts to AlGaN/GaN," Materials Today: Proc., vol.46, pp.6939–6946, 2021.

[10] M. Piazza, C. Dua, M. Oualli, E. Morvan, D. Carisetti, F. Wyczisk, "Degradation of TiAlNiAu as Ohmic contact metal for GaN HEMTs," Mocroelectron. Reliab., vol.49, no.9-11, pp.1222–1225, 2009.

[11] A. Rizzo, L. MIrenghi, M. Massaro, U. Galietti, L. Capodieci, R. Terzi, L. Tapfer, and D. Valerini, "Improved properties of TiAlN coatings through the multilayer structure," Surf. Coat. Technol., vol.235, pp.475–483, 2013.

[12] S. Kaciulis, L. Pandolfi, S. Viticoli, M. Peroni, A. Passaseo, "Characterization of ohmic contacts on GaN/AlGaN heterostructures," Appl. Surf. Sci., vol.253, no.3, pp.1055–1064, 2006.

[13] V. Garbe, J. Weise, M. Motylenko, W. Münchgesang, A. Schmid, D. Rafaja, B. Abendroth, and D.C. Meyer, "Au-free ohmic Ti/Al/TiN contacts to UID n-GaN fabricated by sputter deposition," J. Appl. Phys., vol.121, no.6, 065703, 2017.

[14] G. Greco, F. Giannazzo, F. Iucolano, R.L. Nigro, and F. Roccaforte, "Nanoscale structural and electrical evolution of Ta- and Ti-based contacts on AlGaN/GaN heterostructures," J. Appl. Phys., vol.114, no.8, 083717, 2013.

[15] G. Greczynski, L. Hultman, and M. Odèn, "X-ray photoelectron spectroscopy studies of $Ti_{1-x}Al_xN$ ($0 \le x \le 0.83$) high-temperature oxidation: The crucial role of Al concentration," Surf. Coat. Technol., vol.374, pp.923–934, 2019.

[16] V. Chawla, R. Chandra, R. Jayaganthan, "Effect of phase transformation on structural, electrical and hydrophobic properties of nanocomposite $Ti_{1-x}Al_xN$ films," J. Alloys Compd., vol.507, no.2, pp.L47–L53, 2010.

PAPER

Computer-Aided Design of Cross-Voltage-Domain Energy-Optimized Tapered Buffers

Zhibo CAO[†], **Pengfei HAN**[†], ***Nonmembers***, ***and*** **Hongming LYU**[†,††,†††a)], ***Member***

SUMMARY This paper introduces a computer-aided low-power design method for tapered buffers that address given load capacitances, output transition times, and source impedances. Cross-voltage-domain tapered buffers involving a low-voltage domain in the frontier stages and a high-voltage domain in the posterior stages are further discussed which breaks the trade-off between the energy dissipation and the driving capability in conventional designs. As an essential circuit block, a dedicated analytical model for the level-shifter is proposed. The energy-optimized tapered buffer design is verified for different source and load conditions in a 180-nm CMOS process. The single-V_{DD} buffer model achieves an average inaccuracy of 8.65% on the transition loss compared with Spice simulation results. Cross-voltage tapered buffers can be optimized to further remarkably reduce the energy consumption. The study finds wide applications in energy-efficient switching-mode analog applications.
key words: *tapered buffer, level-shifter, analytical model, low-power circuit, switching-mode, gate driver, short-circuit power*

1. Introduction

Rail-to-rail tapered buffers are extensively applied as gate drivers in a variety of switching-mode circuits such as power amplifiers (PAs) [1]–[4], and DC-DC converters [5]–[8].

While tapered buffers in digital applications emphasize energy and delay [9], [10], gate drivers focus on the energy dissipation for a given drive capability. Research on digital buffers sheds valuable insights into the design of analog gate drivers. For example, Cherkauer et al. provided analytical expressions for the propagation delay, the power dissipation, the physical area, and the system reliability of tapered buffers with fixed ratios between successive stages [11]. In addition, it has been shown that the minimal delay buffer design is contradictory to the purpose of minimizing energy consumption alone [12]–[14].

The driving strength of tapered buffers is manifested by the output transition time, τ_{tb}, for a certain load. And the total energy dissipation of tapered buffers during transition (denoted as E_{tot}) consists of the dynamic losses and the short-circuit losses of each stage (denoted as E_{dyn} and E_{sc}, respectively). While the former can be easily calculated by the corresponding load capacitance, the estimation for E_{sc} is extremely difficult [15], [16]. Although it is claimed that tapered buffers with fixed tapering factors feature relatively low total E_{sc} [15], [17], variably changing the dimensions of each buffer stage can further reduce the energy consumption. E_{dyn} and E_{sc} decrease with a lower supply voltage, which implies the use of a lower V_{DD} domain for power constraint applications. Kursun et al. introduced a low-voltage-swing gate drive technique for the tapered buffer to reduce E_{dyn} [18]. In a buck converter with a 660 MHz switching frequency, a half-rail swing technique for the tapered buffer is employed to reduce the switching losses [19]. While at the same power supply, Frustaci et al. proposed the tapered-V_{TH} buffer design involving transistors with different V_{THs}. While high-V_{TH} devices in the front-end stages reduces the leakage and dynamic power, low-V_{TH} devices in the back-end stages favors the driving capabilities [20]. The applications of tapered buffers often involve different voltage domains in which the frontier and posterior stages operate at different voltages. For example, in DC-DC converters and power amplifiers, the digital control systems are typically powered at a relatively low voltage supply for power reduction and the final stage at a high voltage supply for enhanced driving capabilities. Therefore, a tapered buffer operating at multiple voltage domains is expected to strike a balance between the power consumption and the driving capability.

The analytical model of tapered buffers enables the estimation of its performance, and optimization can be achieved through computer-aided computation methods. For example, Liu et al. applied an iterative optimization algorithm for to improved design metrics such as area, delay, and power [21]. Overeem et al. generated a large quantity of random solutions and select the optimal one [22].

This work aims for the energy-optimized tapered buffer (EOTB) design that meets the output transition time specification (denoted as $\tau_{tb,spec}$) for a given load capacitance, while consuming minimal energy consumption. Cross-V_{DD} energy-optimized tapered buffers (CV-EOTBs) are carefully studied, which consists of two V_{DD} domains and a level-shifter. A computer-aided traversal optimization scheme is proposed to achieve the optimal solution.

The rest of the paper is organized as follows. Section 2 proposes the analytical models of the output transition time (denoted as τ_{out}) and E_{sc} of CMOS inverters. Section 3 introduces the analytical models of τ_{out} and E_{sc} of level-shifters, which serve as a crucial step for CV-EOTB designs. Section 4 presents the design strategy for single-V_{DD} EOTBs

Manuscript received October 14, 2023.
Manuscript revised February 5, 2024.
Manuscript publicized April 9, 2024.
[†]School of Information Science and Technology, ShanghaiTech University, Shanghai 201210, China.
[††]Shanghai Engineering Research Center of Energy Efficient and Custom AI IC, Shanghai 201210, China.
[†††]State Key Laboratory of Advanced Medical Materials and Devices, ShanghaiTech University, Shanghai 201210, China.
a) E-mail: lvhm@shanghaitech.edu.cn
DOI: 10.1587/transele.2023ECP5049

Copyright © 2024 The Institute of Electronics, Information and Communication Engineers

based on the proposed analytical models. Section 5 discusses the design of CV-EOTBs. Section 6 summaries the performances of the tapered buffers achieved through the proposed optimization procedures. Finally, Sect. 7 concludes the paper. In addition, Appendix A presents the parameter values of the analytical inverter models of the inverter. Appendix B presents four design cases of single-V_{DD} EOTBs(Case #B1, Case #B2, Case #B3, Case #B4).

2. Transition Time and Short-Circuit Energy of Inverters

Figure 1 shows the circuit schematic of a CMOS inverter. The transition time is defined as the duration of the voltage transition between 10% V_{DD} and 90% V_{DD}. An accurate inverter model for τ_{out}, should involve the input transition time, τ_{in}, the load capacitance, C_L, and the transistor sizes. It has been pointed out that the τ_{out} models without considering τ_{in} are not accurate, especially for slow input transitions [23]. Thus, Dutta et al. propose a general formulation of τ_{out} involving τ_{in} by curve fitting between the infinitely fast and slow input scenarios [23]. Maurine et al. derive an explicit model of τ_{out} for fast and slow input cases, respectively [24]. It indicates that τ_{out} with slow inputs is proportional to the square root of τ_{in}.

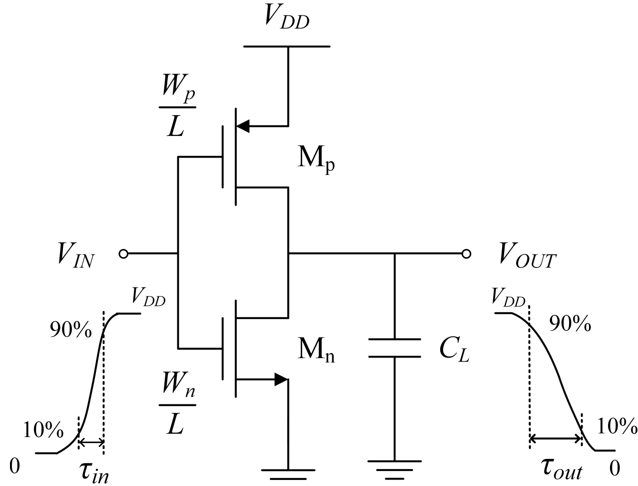


Fig. 1 CMOS inverter.

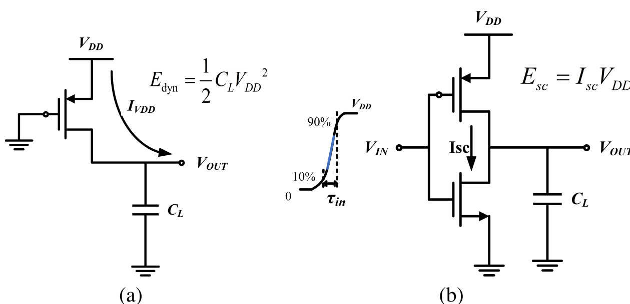


Fig. 2 Illustration of (a) E_{dyn} and (b) E_{sc}.

The switching loss of inverters consists of E_{dyn} and E_{sc} [16], [25]–[28] which are illustrated in Fig. 2 (a) and (b), respectively. When the output voltage of the inverter flips from low to high, the dynamic energy loss equals $\frac{1}{2}C_L V_{DD}^2$ as shown in Fig. 2 (a). On the other hand, during the input transition, there is a phase when both the NMOS and PMOS transistors conducts (indicated by the blue line). The associated energy loss is refered to as E_{sc}. Veendrick et al. derived an approximate expression of E_{sc} for an unloaded inverter and investigate the impact of C_L. The study assumed that with equal τ_{in} and τ_{out}, E_{sc} would only be a small portion (typically less than 20%) of the total transition loss of an inverter. Thus, it proposed fixed-ratio tapered buffer (FTB) designs which feature equalized τ_{in} and τ_{out}. However, the study has not considered variable gate lengths. For inputs with different τ_{in}, transistors in an inverter will experience different operating regions. Therefore, Bisdounis et al. classified four catagories of input ramps and constructed E_{sc} models respectively [29]. Concise models of an inverter is extremely desirable for EOTB designs.

2.1 Analytical Model for Inverter Output Transition Time

The proposed analytical model for τ_{out} is dependent on C_L, τ_{in} and transistor sizes. It is assumed that while NMOS and PMOS transistors have the same gate length, L, the widths of the PMOS transistor, W_p, and NMOS transistor, W_n, enjoy a fixed ratio S, i.e.,

$$W_p = SW_n \tag{1}$$

It is also assumed that $|V_{TH,p}| = V_{TH,n} = V_{TH}$. The propagation delay of the inverter, t_d, can be employed as an intermediary variable that correlates τ_{in} and τ_{out} which is expressed as follows [30],

$$t_d = \left(\frac{1}{2} - \frac{1 - \nu_{TH}}{1 + \alpha}\right)\tau_{in} + \frac{C_L V_{DD}}{2I_{D0}}, \nu_{TH} = \frac{V_{TH}}{V_{DD}} \tag{2}$$

where α is the speed saturation index, and I_{D0} is the saturation current for $V_{GS} = V_{DS} = V_{DD}$.

t_{ds} is the step response of an inverter and is expressed as,

$$t_{ds} = \frac{C_L}{\mu C_{ox}(W_n/L)} \cdot \frac{2V_{DD}}{7/4V_{DD}^2 + V_{TH}^2 - 3V_{DD}V_{TH}} \tag{3}$$

For a given τ_{in}, τ_{out} can be derived as follows [31],

$$\tau_{out} = 2t_{ds}\frac{1 - \nu_{TH}}{0.5 + \frac{t_d}{\tau_{in}} - \nu_{TH}} \tag{4}$$

Therefore, by combining Eqs. (2), (3) and (4), τ_{out} can be expressed as follows,

$$\tau_{out} = \frac{C_L \tau_{in}}{p_1(\frac{W_n}{L})\tau_{in} + p_2 C_L} \tag{5}$$

where p_1 and p_2 are the composite fitting parameters depending on V_{TH}, α and V_{DD}, etc. The unit of p_1 is fF/ns and

p_2 is unitless. Units of all quantities used in the proposed analytical models are listed in Table 1.

2.2 Analytical Model for Inverter Short-Circuit Energy

E_{sc} exists due to the direct current path from V_{DD} to V_{SS} during the transition. Previous research has revealed that τ_{in}, C_L, and transistor sizes directly affect E_{sc}, therefore, are involved in the E_{sc} model.

According to [32], a simplified E_{sc} formula for inverters with relatively slow inputs is expressed as follows,

$$E_{sc}(\tau_{in} \gg \tau_{out}) = k_{sc1}(\frac{W_n}{L})\tau_{in} \tag{6}$$

where k_{sc1} is a composite model parameter which is determined according to V_{DD} and the normalized drain saturated voltage.

For inverters with fast inputs, the E_{sc} model is expressed as,

$$E_{sc}(\tau_{in} \ll \tau_{out}) = \frac{k_{sc2}(\frac{W_n}{L})^2}{C_L}\tau_{in}^2 \tag{7}$$

where k_{sc2} is a composite model parameter determined by V_{TH} and α [32].

A modified E_{sc} model is proposed by combining Eqs. (6) and (7) as follows.

$$E_{sc} = \frac{1}{1/E_{sc}(\tau_{in} \ll \tau_{out}) + 1/E_{sc}(\tau_{in} \gg \tau_{out})} \tag{8}$$

It is simplified as,

$$E_{sc} = \frac{(\frac{W_n}{L})^2\tau_{in}^2}{k_1 C_L + k_2(\frac{W_n}{L})\tau_{in}} \tag{9}$$

where k_1 and k_2 are composite fitting parameters. Their units are ns^2/pJ/fF and ns/pJ, respectively.

2.3 Sub-Models and Parameter Extraction

As pointed out by previouts research [29], [31], τ_{out} and E_{sc} vary significantly for different τ_{in}. Therefore, it is necessary to split the models into sub-models according to τ_{in} for more accurate evaluations. These sub-models are denoted as $M_{0.1,1}$, $M_{1,10}$, $M_{10,100}$, $M_{100,1000}$, $M_{1000,10000}$, $M_{10000,100000}$, respectively, where the subscripts represent their respective ranges of τ_{in} in ns. Each sub-model enjoys the same expression but with different model parameters. All the sub-model parameters are extracted through curve-fitting based on SPICE simulations. The results are listed in Appendix A.

Table 1 Units of quantities in the analytical models

Quantities	Units
Transition Time	ns
Capacitance	fF
Energy	pJ
Transistor Size	μm

3. Transition Time and Short-Circuit Energy Models of Level-Shifters

Figure 3 (a) shows the schematic of a cross-coupled level-shifter between the low V_{DD} domain, V_{DDL}, and the high V_{DD} domain, V_{DDH}. As the input signal V_{IN} turns from V_{SS} to V_{DDL}, M_{N1} switches on so that node Q_1 is pulled down. Thereafter, M_{P4} gradually turns on to pull up node Q_2 to V_{DDH}, so that M_{p3} is switched off. An input falling transition leads to a similar operation process as shown in Fig. 3 (b).

To ensure the successful operation, NMOS transistors should have a larger driving strength than PMOS transistors at the rising input (Fig. 3 (a)). Therefore, the aspect ratio of NMOS transistors is larger than that of the PMOS transistor.

To model the output transition time ($\tau_{out,ls}$)and the transition loss of a level-shifter, the following assumptions are made.

(a) The input signal for the level-shifter features an identical rising and falling transition time denoted as $\tau_{in,ls}$.

(b) M_{N1} and M_{N2} have an identical gate width of $W_{n,ls}$ and M_{P3} and M_{P4} have an identical gate width of $W_{p,ls}$. Both NMOS and PMOS transistors feature an identical gate length of L_{ls}.

(c) The internal inverter in the level-shifter is assumed to perform an ideal inversion, i.e., the input waveforms of M_{N1} and M_{N2} are completely symmetric with opposite phases.

(d) The level-shifter is assumed to be open-loaded since it seldom directly drives the actual load.

These assumptions can greatly simplify the analysis and renders CV-EOTBs to satisfy the engineering requirement.

3.1 Output Transition Time Balancing Scheme

A primary requirement for level-shifter designs is to achieve equalized output rising and falling transition times, i.e., $\tau_{r,ls} = \tau_{f,ls}$. Figure 3 (a) shows the operation upon an input rising transition. The arrows indicate the transient current flow directions. Since M_{N2} is turned off before M_{P4} is turned on, there is no contention and the output rising transition is calculated as,

$$C_{Q2}\frac{dV_{Q2}}{dt} = I_{MP4} \tag{10}$$

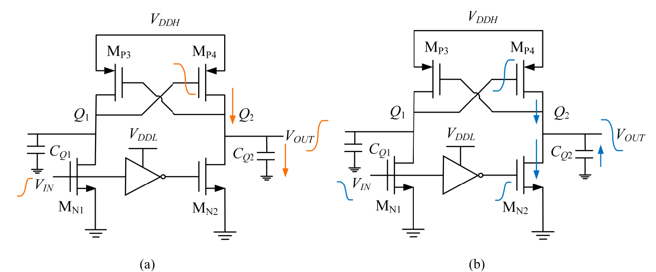


Fig. 3 The operation of a cross-coupled level-shifter upon (a) an input rising transition (b) an input falling transition.

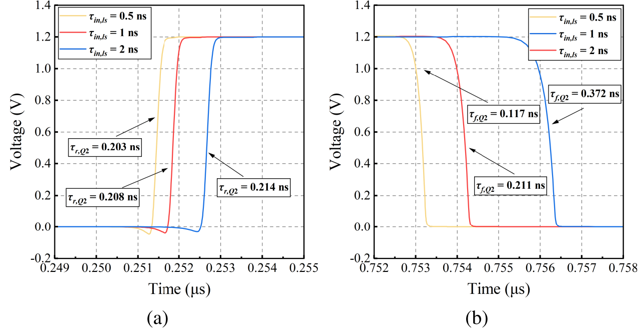


Fig. 4 The uneven impacts on the output transition times due to input transition time variations. Output (a) rising and (b) falling waveforms

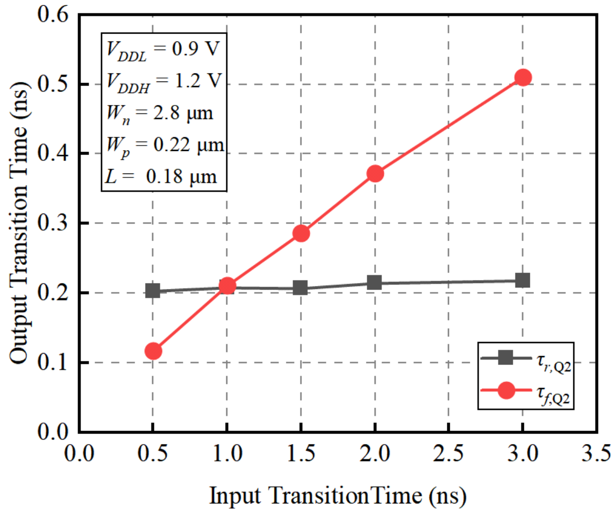


Fig. 5 Output transition times with varying input transition times.

where C_{Q2} is the capacitance at node Q_2.

Figure 3 (b) shows the operation of the level-shifter upon an input falling transition. In this case, there is a contention between M_{P4} and M_{N2}. Hence, the output voltage satisfies the following relationship,

$$I_{MN2} + C_{Q2}\frac{\mathrm{d}V_{Q2}}{\mathrm{d}t} = I_{MP4} \tag{11}$$

It is observed from equations and that $\tau_{in,ls}$ has uneven impacts on $\tau_{r,ls}$ and $\tau_{f,ls}$. Figure 4 and 5 demonstrate a set of exemplary simulation results, where $V_{DDL} = 0.9V, V_{DDH} = 1.2V, W_{n,ls} = 2.8\mu m, W_{p,ls} = 0.22\mu m$,and $L_{ls} = 0.18\mu m$. When $\tau_{in,ls}$ varies from 0.5 ns to 2 ns, $\tau_{f,ls}$ changes from 0.117 ns to 0.372 ns (Fig. 4 (b)), while $\tau_{r,ls}$ only experiences a relatively small change from 0.203 ns to 0.214 ns (Fig. 4 (a)).

Figure 5 summarizes the dependence of output transition times on $\tau_{in,ls}$. It shows that $\tau_{f,ls}$ features a linear dependence on $\tau_{in,ls}$, while $\tau_{r,ls}$ shows a relative independency. As $\tau_{in,ls}$ increases, level-shifter sizes must be adjusted accordingly to guarantee $\tau_{r,ls} = \tau_{f,ls}$. It is achieved by increasing $W_{n,ls}/W_{p,ls}$. The relationship between $W_{n,ls}/W_{p,ls}$ and $\tau_{in,ls}$ can be concluded as,

Table 2 Parameters for the sizing factor model at different V_{DDL} and V_{DDH} combinations

$V_{DDH}(V)$	$V_{DDL}(V)$	$s_1(ns^{-1})$	s_2
1.8	1.5	6.389	6.278
1.8	1.2	14.167	4.056
1.5	1.2	9.722	2.778
1.8	0.9	12.89	7.629
1.5	0.9	14.167	4.444
1.2	0.9	10.278	0.167

$$\frac{W_{n,ls}}{W_{p,ls}} = s_1\tau_{in,ls} + s_2 \tag{12}$$

where $s_1(ns^{-1})$ and s_2 are the fitting parameters which can be obtained from SPICE simulations.

It is also empirically observed that the $W_{n,ls}/W_{p,ls}$ ratio is inversely proportional to L_{ls}. Therefore, a sizing factor, β, is proposed as follows,

$$\beta = \frac{W_{n,ls}L_{ls}}{W_{p,ls}L_{min}} = s_1\tau_{in,ls} + s_2 \tag{13}$$

Table 2 enlists the values of s_1 and s_2 for different V_{DDL} and V_{DDH}.

3.2 Short-Circuit Energy Model

The switching loss of a level-shifter comprises the dynamic and the short-circuit energies. While the dynamic energy has a similar form as that of an inverter, its short-circuit energy, $E_{sc,ls}$, is heavily dependent on $\tau_{in,ls}$ and transistor sizes. The impact of $\tau_{in,ls}$ is proportional to the aspect ratio of the PMOS transistors. Therefore, a scale factor, k, is introduced as follows,

$$k = \frac{W_{p,ls}/L_{ls}}{W_{p,min}/L_{min}} \tag{14}$$

where $W_{p,min}$ is the minimum width of PMOS transistors.

$E_{sc,ls}$ features its linear dependence on the product of $\tau_{in,ls}$ and k, and is modeled as follows,

$$E_{sc,ls} = r_1 k\tau_{in,ls} + r_2 \tag{15}$$

where $r_1(pJ/ns)$ and $r_2(pJ)$ are curve-fitting parameters. Table 3 lists the values of r_1 and r_2 for different V_{DDL} and V_{DDH}. The accuracies of the fitting models in Eqs. (13) and (15) are intuitively illustrated in Fig. 6, which shows the fitting results for $V_{DDH} = 1.8V$ and $V_{DDL} = 0.9V$. The fitting functions are remarkably close to the SPICE simulation results.

$\tau_{out,ls}$ can be calculated as either the rising or the falling transition times since they are equal. The output resistance can be expressed as,

$$R_{eq,p} = R_{eq,p0}/k \tag{16}$$

where $R_{eq,p0}$ is the output resistance of the PMOS transistor with $W_{p,min}$ and L_{min}. The output transition time of the level-shifter is proportional to $R_{eq,p}C_{Q2}$. For example, the

Table 3 Parameters for $E_{sc,ls}$ model at different V_{DDL} and V_{DDH} combinations

$V_{DDH}(V)$	$V_{DDL}(V)$	$r_1(pJ/ns)$	$r_2(pJ)$
1.8	1.5	0.1096	0.00029
1.8	1.2	0.0878	−0.00143
1.5	1.2	0.0346	0.021
1.8	0.9	0.0459	0.1233
1.5	0.9	0.0342	0.0328
1.2	0.9	0.01292	0.0222

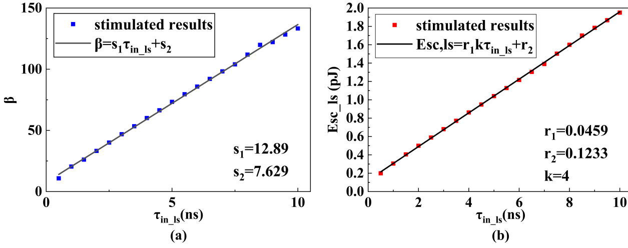


Fig. 6 Comparisons between the SPICE simulation results of (a) β and (b) $E_{sc,ls}$ with the fitting functions for $V_{DDH} = 1.8V$ and $V_{DDL} = 0.9V$.

transition time from 10% V_{DDH} and 90% V_{DDH} takes about $2.2R_{eq,p}C_{Q2}$.

4. Design Strategy for Single-V_{DD} EOTBs

Performance evaluation of tapered buffers relies on the estimation of the equivalent load capacitances of each stage as shown in Fig. 7. In a tapered buffer, the load capacitance of the i-th stage ($i \geqslant 1$), C_i, is composed of the output capacitance of the i-th stage, $C_{out,i}$, and the input gate capacitance of the (i+1)-th stage, $C_{g,i+1}$. For each stage, a fixed process-dependent ratio, γ, exists between the output and input capacitances, i.e., $C_{out,i} = \gamma C_{g,i}$.

Therefore, C_i can be expressed as [32],

$$C_i = (1 + S)(W_{n,i+1}L_{i+1} + \gamma W_{n,i}L_i)C_{ox} \tag{17}$$

where the widths of the NMOS and PMOS transistors in the i-th stage are represented as $W_{n,i}$ and $W_{p,i}$, respectively, with a ratio of S. Their gate lengths are denoted as L_i.

Figure 7 shows the circuit schematic of a single-V_{DD} N-stage tapered buffer where R_s is the source resistance and $C_{L,tb}$ is the load capacitance. The output transition time and short-circuit energy of each stage can be calculated according to Eqs. (5) and (9). The design of a single-V_{DD} EOTB is essentially to minimize the overall transition loss, E_{tot}, while meeting the output transition time specification, $\tau_{tb,spec}$, i.e.,

$$\begin{cases} E_{eotb} &= \min E_{tot} = \min \sum_{i=1}^{N}(E_{dyn,i} + E_{sc,i}) \\ \tau_{tb} &= \tau_N \in [\tau_{tb,spec} - \Delta\tau, \tau_{tb,spec} + \Delta\tau] \end{cases} \tag{18}$$

where E_{eotb} is the total energy dissipation of the EOTB, while $E_{dyn,i}$ and $E_{sc,i}$ represent the dynamic and short-circuit energy dissipation of the i-th stage in the tapered buffer, respectively. τ_N is the output transition time of the N-th stage, and $\Delta\tau$ is the acceptable error margin of the

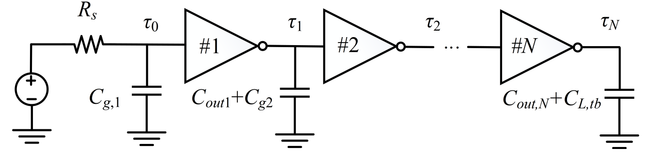


Fig. 7 Circuit model of a single-V_{DD} N-stage tapered buffer.

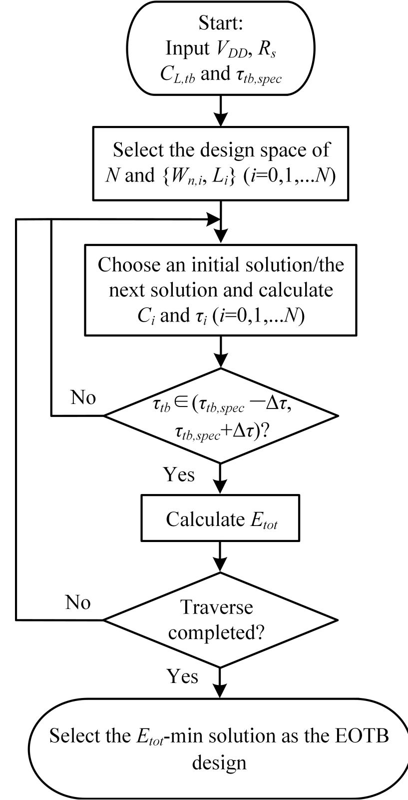


Fig. 8 Traversal algorithm for the EOTB design.

output transition time, τ_{tb}. A computer-aided traversal algorithm is proposed accordingly as shown in Fig. 8. $\tau_{tb,spec}$ is determined according the application requirements, i.e., the driving capability of the tapered buffer. And, E_{eotb} is the minimum energy consumption among that of all feasible designs that satisfy $\tau_{tb,spec}$. It is worth mentioning that the ideality of the optimal design is determined by the range of the design set and the step size of the travesal algorithm. There is a trade-off between the accuracy and the algorithm execution time.

5. Design Strategy of CV-EOTBs

As E_{sc} increases with V_{DD} [33], it is favorable to adopt a lower V_{DD} at the front-end stages of tapered buffers to reduce power consumption and a higher V_{DD} at the back-end stages to ensure driving capabilities as depicted in the circuit schematic in Fig. 9. Such implementations require cross-V_{DD} level-shifters and are referred to as CV-EOTBs in this work.

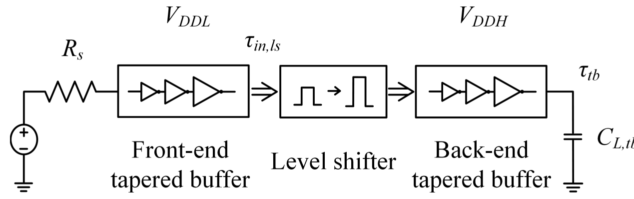


Fig. 9 Circuit schematic of a cross-V_{DD} tapered buffer.

5.1 CV-EOTB Design

The CV-EOTB design should start from the sizing of the level-shifter as it influences the load capacitance of the front-end buffer (denoted as $C_{L,fe}$) and the source resistance of the back-end buffer (denoted as $R_{s,be}$). $C_{L,fe}$, i.e., input capacitance of the level-shifter, is calculated as follows,

$$C_{L,fe} = W_{n,ls} L_{ls} C_{ox} \tag{19}$$

$R_{s,be}$, i.e., $R_{eq,p}$ of the level-shifter, is determined by k. $\tau_{in,ls}$ and k determine $E_{sc,ls}$ according to Eqs. (15). It also partially determines the size of the NMOS transistor according to Eqs. (13). To simplify the algorithm, when k >1, L_{ls} is chosen to be the minimum size and $W_{p,ls}$ is increase to satisfy k. Conversely, when k <1, $W_{p,ls}$ is chosen as the minimum size and L_{ls} is increased.

Therefore, a traversal algorithm based on $\tau_{in,ls}$ and k is employed to achieve the optimized CV-EOTBs design. The algorithm flow is shown in Fig. 10. As $C_{L,fe}$ and $R_{s,be}$ can be calculated for any combination of $\tau_{in,ls}$ and k, the CV-EOTB design problem is transformed to the low-energy design of the front-end and back-end buffers, respectively. Their switching losses are is denoted as $E_{fe,eotb}$ and $E_{be,eotb}$, respectively, while that of the level-shifter is denoted as E_{ls}. The design constraints for front-end and back-end EOTBs are summarized in Table 4.

5.2 CV-EOTB Design Example

A design case (CV-EOTB Design #1) is presented to exemplify the proposed CV-EOTB design procedure. The design constraints are the following: $V_{DDL} = 1.2V, V_{DDH} = 1.8V, R_s = 100M\Omega, C_{L,tb} = 200fF$,and $\tau_{tb,spec} = 1ns$. Through the optimization algorithm, k and $\tau_{in,ls}$ are determined to be 0.123 and 1 ns, respectively. β is calculated to be 18.22 according to Eqs. (13). Therefore, $W_{n,ls}, W_{p.ls}$ and L_{ls} are determined to be 0.49 μm,0.22 μm and 1.47 μm, respectively. The front-end and back-end EOTB designs are obtained according to the EOTB design procedure in Fig. 8. The optimized solutions are presented in Table 5 and Table 6, respectively.

In this design, $E_{fe,eotb}, E_{be,eotb}$ and E_{ls} are estimated to 0.51 pJ, 0.71 pJ, and 0.06 pJ, respectively, which sum to be 1.28 pJ(denoted as $E_{cv,eotb}$). In contrast, transistor-level SPICE simulation leads to $E_{fe,eotb}, E_{be,eotb}$ and E_{ls} of 0.86 pJ, 0.77 pJ and 0.09 pJ, respectively, which sum to be 1.72 pJ. Both results are remarkably close which validates the

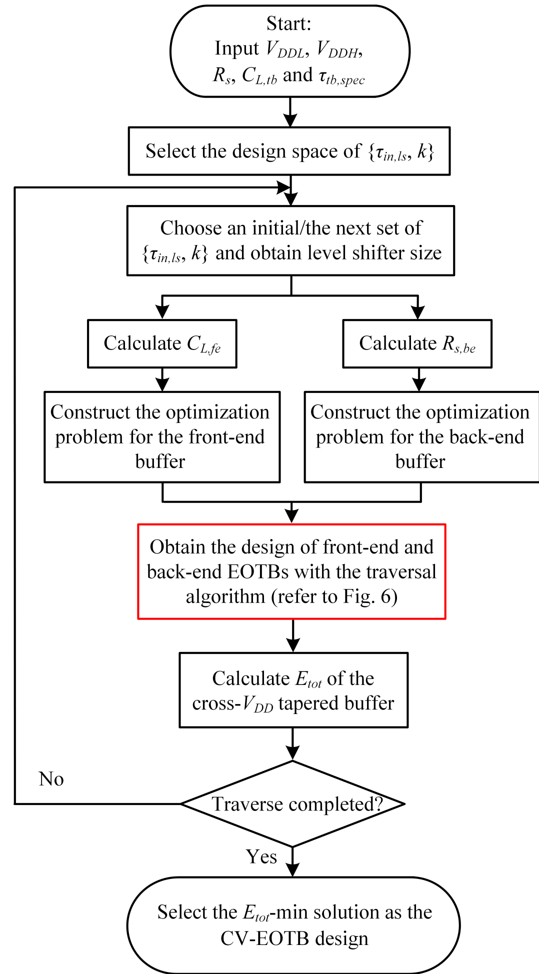


Fig. 10 Optimization flow of the CV-EOTB.

Table 4 Design constraints for front-end and back-end EOTBs in a CV-EOTB design

External conditions	Front-end EOTB	Back-end EOTB
Source resistance	R_s	$R_{s,be}$
Load capacitance	$C_{L,fe}$	$C_{L,tb}$
Output transition time specification	$\tau_{in,ls}$	$\tau_{tb,spec}$

Table 5 Front-end EOTB design in CV-EOTB Design #1

i	$W_{n,i}(\mu m)$	$L_i(\mu m)$
1	0.22	0.5
2	0.3	0.5
3	0.4	0.58

Table 6 Back-end EOTB design in CV-EOTB Design #1

i	$W_{n,i}(\mu m)$	$L_i(\mu m)$
1	0.22	0.4
2	0.3	0.3
3	0.6	0.18

proposed design procedure.

6. Performances of EOTBs and CV-EOTBs

The proposed single- V_{DD} EOTB design strategy is applied in four design cases(Case #B1, Case #B2, Case #B3, Case #B4) each with different design constraints as listed in Appendix B. The estimated τ_{tb} and E_{eotb} results are compared with those obtained from SPICE simulations, and the average errors of this four cases are only 16.43% and 8.65%, respectively. These four examples encompass V_{DD} from 0.9V to 1.8V, R_s from 30MΩ to 1GΩ, C_L from 0.5pF to 5pF, and $\tau_{tb,spec}$ from 1ns to 20ns, which are typical values in practical applications.

As a key variable in the optimization procedure, the effect of N is demonstrated in an example in Fig. 11. For $V_{DD} = 1.8V, R_s = 50M\Omega, C_{L,tb} = 5pF$, and $\tau_{tb,spec} = 8ns$, the design of EOTBs for $N = 3$ and $N = 4$ are shown in Fig. 11 (a) and 11 (b), respectively. The 4-stage EOTB achieves a lower E_{eotb} thanks to the reduction of E_{sc}. It is also worth noting that, for an N-stage EOTB, E_{sc} of the first stage and E_{dyn} of the last stage usually take a dominant part.

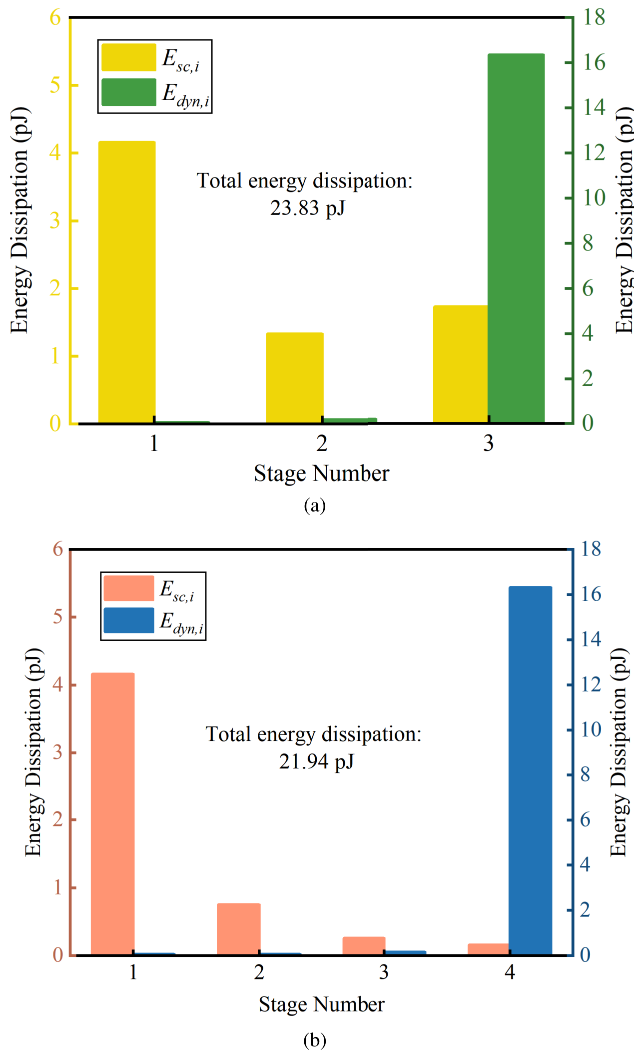


Fig. 11 Energy distribution in EOTBs. (a) The 3-stage EOTB design. (b) The 4-stage EOTB design.

Figure 12 plots the optimized designs of a 3-stage EOTB for $V_{DD} = 1.8V, R_s = 100M\Omega, C_{L,tb} = 3pF$ and $\tau_{tb,spec} \in (5ns, 15ns)$. Tapered buffers with random optimized sizes are demonstrated for comparison, in which the gate widths and lengths have been randomly modified (always with $W_{n,1} < W_{n,2} < W_{n,3}$). For any $\tau_{tb,spec}$, the EOTB design achieves the lowest transition loss.

Figure 13 (a) and 13 (b) displays the size, E_{sc}, and E_{dyn} of each stage of an optimized single-V_{DD} EOTB design, where $V_{DD} = 1.8V, R_s = 100M\Omega, C_{L,tb} = 3pF, \tau_{tb,spec} =$

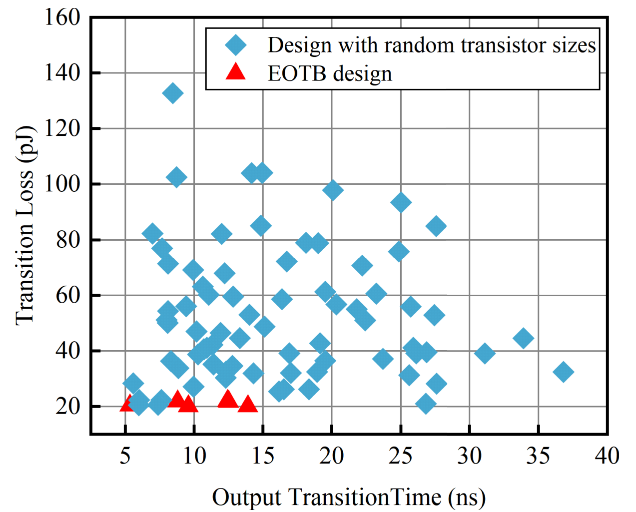


Fig. 12 Performance of the EOTB design compared to the design with random gate sizes.

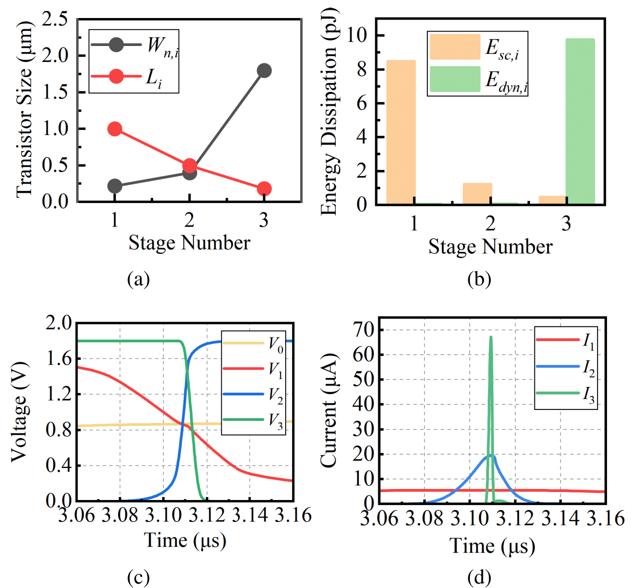


Fig. 13 A 3-stage EOTB design. (a) Buffer sizes. (b) Short-circuit and dynamic energy dissipation of all stages. (c) Transient voltage waveforms of all stages. (d) Transient current waveforms of all stages.

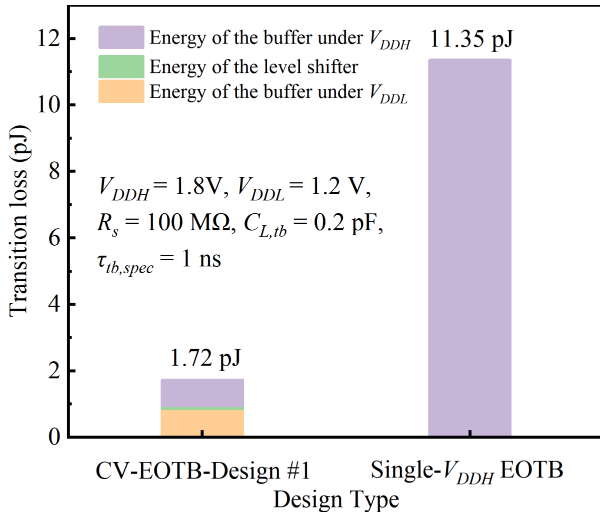


Fig. 14 Transition losses in CV-EOTB-Design #1 and the single-V_{DD} EOTB.

Table 7 Process corner simulations of CV-EOTB-Design #1

Process corner	TT	SS	FF	SF	FS
$E_{cv,eotb}(pJ)$	1.72	1.26	2.65	1.78	1.75
$\tau_{tb}(ns)$	0.87	0.99	0.77	0.88	0.87

$5ns$. Figure 13 (c) and 13 (d) demonstrate the corresponding transient voltage and the transient current waveforms, respectively.

Based on the accurate modeling of EOTBs, the design of CV-EOTBs is transformed to the that of the front-end and back-end EOTBs along with a level shifter. Figure 14 displays the transition losses of CV-EOTB-design #1 and the single-V_{DD} EOTB design assuming the same constraints and the supply of V_{DDH}. For the single-V_{DD} EOTB, E_{sc} alone is estimated to be 10.68 pJ, according to SPICE simulations and E_{eotb} is estimated to be 11.35 pJ. In contrast, CV-EOTB-design #1 reduces the transition loss by 84.8%. Corner simulations are performed for CV-EOTB-design #1. It shows that the effects of potential process variations are limited. Table 7 enlists $E_{cv,eotb}$ and τ_{tb} of CV-EOTB-Design #1 with different process corners.

7. Conclusion

The low-power design of taper buffers is crucial for a series of switch-mode analog applications. This paper proposes the optimization procedure for taper buffers based on analytical models of τ_{out} and E_{sc} of CMOS inverters and level-shifters.

Computer-aided traversal algorithms for the single-V_{DD} EOTB and CV-EOTB optimization are proposed with the latter being transformed to the front-end and back-end single-V_{DD} designs, respectively. The proposed design procedures are validated against SPICE simulation results manifesting their effectiveness in addressing practical engineering problems.

References

[1] G. Tochou, A. Cathelin, A. Frappé, A. Kaiser, and J. Rabaey, “Impact of forward body-biasing on ultra-low voltage switched-capacitor RF power amplifier in 28 nm FD-SOI,” IEEE Transactions on Circuits and Systems II: Express Briefs, vol.69, no.1, pp.50–54, 2021.

[2] A. Abuelnasr, M. Amer, M. Ali, A. Hassan, B. Gosselin, A. Ragab, and Y. Savaria, “Delay Mismatch Insensitive Dead Time Generator for High-Voltage Switched-Mode Power Amplifiers,” IEEE Transactions on Circuits and Systems I: Regular Papers, vol.70, no.4, pp.1555–1565, 2023.

[3] H. Zhang, S. Karmakar, L.J. Breems, Q. Sandifort, M. Berkhout, K.A.A. Makinwa, and Q. Fan, “A high-linearity and low-EMI multi-level class-D amplifier,” IEEE Journal of Solid-State Circuits, vol.56, no.4, pp.1176–1185, 2020.

[4] R.J. Bootsman, et al., “High-power digital transmitters for wireless infrastructure applications (a feasibility study),” IEEE Transactions on Microwave Theory and Techniques, vol.70, no.5, pp.2835–2850, 2022.

[5] A. Mishra and V. De Smedt, “A novel hybrid buck-boost converter topology for Li-ion batteries with increased efficiency,” 2020 27th IEEE International Conference on Electronics, Circuits and Systems (ICECS), pp.1–4, IEEE, 2020.

[6] H. Ma, G. Namgoong, E. Choi, and F. Bien, “Instantaneous power consuming level-shifter for improving power conversion efficiency of buck converter,” IEEE Transactions on Circuits and Systems II: Express Briefs, vol.66, no.7, pp.1207–1211, 2018.

[7] P.A.E. Divino, M.P. Prunes, and K.O. Maglinte, “Single-Inductor Multiple-Output (SIMO) Boost DC-DC Converter using Ripple-Based Control for Ultra-Low Power Indoor Light Energy Harvesting Applications,” 2019 IEEE 11th International Conference on Humanoid, Nanotechnology, Information Technology, Communica-tion and Control, Environment, and Management (HNICEM), pp.1–6, IEEE, 2019.

[8] M.W. Kim and J.J. Kim, “A PWM/PFM dual-mode DC-DC buck converter with load-dependent efficiency-controllable scheme for multi-purpose IoT applications,” Energies, vol.14, no.4, p.960, 2021.

[9] S. Ma and P. Franzon, “Energy control and accurate delay estimation in the design of CMOS buffers,” IEEE Journal of Solid-State Circuits, vol.29, no.9, pp.1150–1153, 1994.

[10] C.W. Kang, S. Abbaspour, and M. Pedram, “Buffer sizing for minimum energy-delay product by using an approximating polynomial,” Proceedings of the 13th ACM Great Lakes symposium on VLSI, pp.112–115, 2003.

[11] B.S. Cherkauer and E.G. Friedman, “A unified design methodology for CMOS tapered buffers,” IEEE Transactions on Very Large Scale Integration Systems, vol.3, no.1, pp.99–111, 1995.

[12] S.R. Vemuru and A.R. Thorbjornsen, “Variable-taper CMOS buffers,” IEEE Journal of Solid-State Circuits, vol.26, no.9, pp.1265–1269, 1991.

[13] H. Wang, M. Miranda, A. Papanikolaou, F. Catthoor, and W. Dehaene, “Variable tapered pareto buffer design and implementation allowing run-time configuration for low-power embedded SRAMs,” IEEE Transactions on Very Large Scale Integration Systems, vol.13, no.10, pp.1127–1135, 2005.

[14] J.-S. Choi and K. Lee, “Design of CMOS tapered buffer for minimum power-delay product,” IEEE Journal of Solid-State Circuits, vol.29, no.9, pp.1142–1145, 1994.

[15] H.J.M. Veendrick, “Short-circuit dissipation of static CMOS circuitry and its impact on the design of buffer circuits,” IEEE Journal of Solid-State Circuits, vol.19, no.4, pp.468–473, 1984.

[16] S. Turgis and D. Auvergne, “A novel macromodel for power estimation in CMOS structures,” IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol.17, no.11, pp.1090–1098, 1998.

[17] A.J. Stratakos, S.R. Sanders, and R.W. Brodersen, “A low-voltage

CMOS DC-DC converter for a portable battery-operated system," Proceedings of 1994 Power Electronics Specialist Conference-PESC'94, vol.1, pp.619–626, IEEE, 1994.

[18] V. Kursun, S.G. Narendra, V.K. De, and E.G. Friedman, "Low-voltage-swing monolithic dc-dc conversion," IEEE Transactions on Circuits Systems II: Express Briefs, vol.51, no.5, pp.241–248, 2004.

[19] M. Alimadadi, S. Sheikhaei, G. Lemieux, S. Mirabbasi, W.G. Dunford, and P.R. Palmer, "A fully integrated 660 MHz low-swing energy-recycling DC–DC converter," IEEE Transactions on Power Electronics, vol.24, no.6, pp.1475–1485, 2009.

[20] F. Frustaci, P. Corsonello, and M. Alioto, "Tapered-V TH CMOS buffer design for improved energy efficiency in deep nanometer technology," 2011 IEEE International Symposium of Circuits and Systems (ISCAS), pp.2075–2078, IEEE, 2011.

[21] S. Liu, S.O. Memik, and Y.I. Ismail, "A comprehensive tapered buffer optimization algorithm for unified design metrics," 2011 IEEE International Symposium of Circuits and Systems (ISCAS), pp.2277–2280, IEEE, 2011.

[22] W.T. Overeem, M.S.O. Alink, and B. Nauta, "Inverter Chain Buffer Optimization for N-path Filter Switch Drivers and Validation through Simulations in 22nm FD-SOI Technology," 2023 IEEE International Symposium on Circuits and Systems (ISCAS), pp.1–5, IEEE, 2023.

[23] S. Dutta, S.S.M. Shetti, and S.L. Lusky, "A comprehensive delay model for CMOS inverters," IEEE Journal of Solid-State Circuits, vol.30, no.8, pp.864–871, 1995.

[24] P. Maurine, M. Rezzoug, N. Azemard, and D. Auvergne, "Transition time modeling in deep submicron CMOS," IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol.21, no.11, pp.1352–1363, 2002.

[25] S.R. Vemuru and N. Scheinberg, "Short-circuit power dissipation estimation for CMOS logic gates," IEEE Transactions on Circuits Systems I: Fundamental Theory and Applications, vol.41, no.11, pp.762–765, 1994.

[26] T. Tanzawa, "On-chip high-voltage generator design," Springer, 2013.

[27] S. Turgis, N. Azemard, and D. Auvergne, "Explicit evaluation of short circuit power dissipation for CMOS logic structures," Proceedings of the 1995 International Symposium on Low Power Design, pp.129–134, 1995.

[28] A. Hirata, H. Onodera, and K. Tamaru, "Estimation of short-circuit power dissipation and its influence on propagation delay for static CMOS gates," 1996 IEEE International Symposium on Circuits and Systems. Circuits and Systems Connecting the World. ISCAS 96, vol.4, pp.751–754, IEEE, 1996.

[29] L. Bisdounis, S. Nikolaidis, and O. Loufopavlou, "Propagation delay and short-circuit power dissipation modeling of the CMOS inverter," IEEE Transactions on Circuits and Systems I: Fundamental theory and applications, vol.45, no.3, pp.259–270, 1998.

[30] T. Sakurai and A.R. Newton, "Alpha-power law MOSFET model and its applications to CMOS inverter delay and other formulas," IEEE Journal of Solid-State Circuits, vol.25, no.2, pp.584–594, 1990.

[31] D. Auvergne, J.M. Daga, and M. Rezzoug, "Signal transition time effect on CMOS delay evaluation," IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications, vol.47, no.9, pp.1362–1369, 2000.

[32] J.M. Rabaey, A.P. Chandrakasan, and B. Nikolic, Digital Integrated Circuits, Prentice Hall, Englewood Cliffs, 2002.

[33] K. Nose and T. Sakurai, "Analysis and future trend of short-circuit power," IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol.19, no.9, pp.1023–1030, 2000.

Appendix A: Sub-Model Parameters of Inverters

Table A·1 Sub-model parameters with 1.8 V V_{DD} and 180-nm technology

Sub-model	$p1$	$p2$	$k1$	$k2$
$M_{0.1,1}$	14.163	1.601	0.0896	1.9094
$M_{1,10}$	19.2	0.868	0.1202	2.0268
$M_{10,100}$	0.892	6.91	0.1202	2.0268
$M_{100,1000}$	0.366	5.718	1.2884	1.7576
$M_{1000,10000}$	−0.0025	14.606	0	42.373
$M_{10000,100000}$	0.00017	14.116	0	42.373

Table A·2 Sub-model parameters with 1.5 V V_{DD} and 180-nm technology

Sub-model	$p1$	$p2$	$k1$	$k2$
$M_{0.1,1}$	13.395	1.425	−0.1028	5.8222
$M_{1,10}$	18.536	0.745	0.324	7.2744
$M_{10,100}$	1.409	6.612	1.9342	5.7942
$M_{100,1000}$	0.613	4.933	4.8886	5.7568
$M_{1000,10000}$	−0.001	27.526	0	142.86
$M_{10000,100000}$	0.0003	25.92	0	142.86

Table A·3 Sub-model parameters with 1.2 V V_{DD} and 180-nm technology

Sub-model	$p1$	$p2$	$k1$	$k2$
$M_{0.1,1}$	11.68	1.23	−0.6332	18.7282
$M_{1,10}$	16.998	0.586	25.336	27.9666
$M_{10,100}$	1.575	5.064	17.7742	39.4496
$M_{100,1000}$	0.752	4.13	36.3968	35.4432
$M_{1000,10000}$	0.02	30.129	0	555.56
$M_{10000,100000}$	−0.0002	16.747	0	555.56

Table A·4 Sub-model parameters with 0.9 V V_{DD} and 180-nm technology

Sub-model	$p1$	$p2$	$k1$	$k2$
$M_{0.1,1}$	12	0.5199	−1.2254	32
$M_{1,10}$	14.88	0.2582	140.88	−577.2
$M_{10,100}$	0.6946	3.651	−3571.4	1313.2
$M_{100,1000}$	0.3676	3.163	−219.6	947.8
$M_{1000,10000}$	0.01611	8.838	0	8333
$M_{10000,100000}$	−0.006087	32.33	0	8333

Appendix B: Single-V_{DD} EOTB Design Cases

Table A·5 Case #B1:180-nm technology, V_{DD}=1.8 V, R_s=100 MΩ, $C_{L,tb}$=1 pF, $\tau_{tb,spec}$=2 ns

i	$W_{n,i}(\mu m)$	$L_i(\mu m)$	$\tau_{i-1,est}(ns)$	$\tau_{i-1,sim}(ns)$	$E_{i,est}(pJ/op)$	$E_{i,sim}(pJ/op)$
1	0.22	0.9	1647	1552	9.6	8.68
2	0.3	1	113	170	0.102	0.54
3	0.4	0.7	18	9.08	0.0836	0.11
4	1	0.18	2.41	1.12	3.31	3.31
			2.14	2.69	13.1	12.6

Table A·6 Case #B2:180-nm technology, V_{DD}=1.5 V, R_s=1 GΩ, $C_{L,tb}$=0.5 pF, $\tau_{tb,spec}$=5 ns

i	$W_{n,i}(\mu m)$	$L_i(\mu m)$	$\tau_{i-1,est}(ns)$	$\tau_{i-1,sim}(ns)$	$E_{i,est}(pJ/op)$	$E_{i,sim}(pJ/op)$
1	0.22	0.5	7323	7259	28.2	38
2	0.3	0.8	302.4	907	0.191	2
3	0.5	0.3	30.67	34.95	0.153	0.305
4	4	0.18	3.474	3.223	1.508	1.538
			1.001	0.958	30.07	41.87

Table A·7 Case #B3:180-nm technology, V_{DD}=1.2 V, R_s=30 MΩ, $C_{L,tb}$=5 pF, $\tau_{tb,spec}$=10 ns

i	$W_{n,i}(\mu m)$	$L_i(\mu m)$	$\tau_{i-1,est}(ns)$	$\tau_{i-1,sim}(ns)$	$E_{i,est}(pJ/op)$	$E_{i,sim}(pJ/op)$
1	0.22	1.1	604.1	501.9	0.142	0.21
2	0.3	0.5	57.3	31.5	0.0277	0.052
3	2	0.18	7.23	4.25	7.22	7.25
			8.42	10.04	7.39	7.51

Table A·8 Case #B4:180-nm technology, V_{DD}=0.9 V, R_s=200 MΩ, $C_{L,tb}$=2 pF,$\tau_{tb,spec}$=20 ns

i	$W_{n,i}(\mu m)$	$L_i(\mu m)$	$\tau_{i-1,est}(ns)$	$\tau_{i-1,sim}(ns)$	$E_{i,est}(pJ/op)$	$E_{i,sim}(pJ/op)$
1	0.22	1.2	4394	3204	0.178	0.119
2	0.4	0.8	466	171.7	0.0168	0.04
3	3	0.38	63.9	27.1	1.63	1.65
			16.7	13.4	1.82	1.81

Zhibo Cao received the B.S. degree in electrical engineering from Tianjin University and is now pursuing the M.S. degree at Shanghaitech University. His research interests include low-power design techniques for analog ICs.

Pengfei Han received the B.S. degrees in Electrical Information Engineering from Liaoning University in 2022. He is now with Shanghaitech University to pursue a master degree.His research interest is power management ICs.

Hongming Lyu received the B.S. degree from University of Electronic Science and Technology of China and the Ph.D. degree from Tsinghua University in 2010 and 2015, respectively. From 2015 to 2020, he was a Postdoctoral Research Fellow successively with University of California San Diego (UCSD), Rice University and University of California Los Angeles (UCLA). Since 2020, he has been a tenure-track assistant professor and the principle Investigator of Biomedical Integrated Circuits and Microsystems Laboratory (BICML) at ShanghaiTech University. His research focuses on analog and mixed-signal integrated circuits for biomedical applications. He is a member of IEEE and IEICE.

PAPER

Measuring SET Pulse Widths in pMOSFETs and nMOSFETs Separately by Heavy Ion and Neutron Irradiation

Jun FURUTA[†a)], ***Member***, **Shotaro SUGITANI**[†], **Ryuichi NAKAJIMA**[†], ***Student Members***, **Takafumi ITO**[†], ***Nonmember***, ***and*** **Kazutoshi KOBAYASHI**[†], ***Senior Member***

SUMMARY Radiation-induced temporal errors become a significant issue for circuit reliability. We measured the pulse widths of radiation-induced single event transients (SETs) from pMOSFETs and nMOSFETs separately. Test results show that heavy-ion induced SET rates of nMOSFETs were twice as high as those of pMOSFETs and that neutron-induced SETs occurred only in nMOSFETs. It was confirmed that the SET distribution from inverter chains can be estimated using the SET distribution from pMOSFETs and nMOSFETs by considering the difference in load capacitance of the measurement circuits.

key words: *single event transient, pulse width, heavy ion, high energy neutron*

1. Introduction

Single event effects are one of a significant issue for circuit reliability since they transiently flip the output of transistor, resulting in circuit malfunction. This malfunction is called soft error [1]–[4]. Single event effects are caused by a charged particle. A charged particle generates electron-hole pairs in semiconductor along its track and the generated electrons are collected to the diffusion region of off-state nMOSFET by the electric field in the depletion layer. Thereby, the output of the off-state nMOSFET is flipped transiently. When a charged particle passes through a logic circuit, its output is inverted for a time depending on the amount of collected charge, which is called single event transient (SET). When a charged particle passes through a storage element, its stored data can be flipped by collected charge, which is called single event upset (SEU). The SEU is more critical to the circuit reliability because incorrect outputs due to SET can only lead to malfunctions when they are captured by flip-flops (FFs). However, SET is more frequently captured by FFs as the clock frequency increases [5], [6]. Therefore, in advanced technology, measuring SET rates and their pulse widths has become important for calculating the soft error rate and estimating the reduction of error rate by low-pass filters [7], [8]. Furthermore, measuring the SET pulse width distribution can be used to evaluate the distribution of the amount of collected charge generated by radiation strike, since the SET pulse width depends on the amount of collected charge.

Previous researches on SETs have reported numerous characteristics, such as temperature characteristics, dependence on well contact density, driving strength and logic circuit structure [9]–[14]. However, most of these studies measured SET pulse widths using CMOS inverter chains. Hence, it is impossible to determine whether the measured SET was generated from an nMOSFET or a pMOSFET. Since the electrons and holes generated by a radiation strike are collected by nMOSFETs and pMOSFETs, respectively, the SETs produced in nMOSFETs can have different characteristics from those in pMOSFETs. Therefore, separated measurement of SETs from nMOSFETs and from pMOSFETs is important to obtain detailed characteristics of single event transients. To address the above problem, S. Jagannathan et al. measured SET pulse widths from nMOSFET and pMOSFET using partially-duplicated logic chains [12]. Their results showed that pMOSFET had 3.9 times more SETs than nMOSFET. However, only a few reports have measured SET pulse widths using this method.

In addition to a lack of distinction between SETs from nMOSFET and pMOSFET, there were issues with the measurement accuracy of SET pulse width [15]–[17]. SET pulse width can be modulated during propagation through logic gates since SET pulse widths are generally measured using a long logic chain [18]. This propagation-induced pulse modulation (shrinking/broadening) is caused by the difference between the rise and fall propagation times. Thus, the target logic structures need to be considered to achieve high measurement accuracy of SET pulse widths. Additionally, it is crucial to measure the same SET characteristics using different measurement circuits in order to verify the SET characteristics that depend on specific circuit structures.

In this paper, SET pulses widths from nMOSFET (nSET pulse widths) and pMOSFET (pSET pulse widths) were measured to separate single event effects on nMOSFET and pMOSFET. We show that the impact of single-event effects is different for nMOSFETs and pMOSFETs by performing the above measurements, indicating that separate measurements are essential for accurate SET modeling and estimating the soft error rate of a circuit. In addition, we clarify the effects of the differences in the measurement circuits by comparing the measurement results of nSETs and pSETs with SET pulse widths measured by the conventional method using inverter chains. The rest of this paper is organized as follows. Section 2 introduces structures of Target logic circuits and time-to-digital converter. The measurement results by Kr ions and neutron irradiation are shown in

Manuscript received October 3, 2023.
Manuscript revised February 9, 2024.
Manuscript publicized April 10, 2024.
[†]Graduate School of Science and Technology, Kyoto Institute of Technology, Kyoto-shi, 606–8585 Japan.
a) E-mail: furuta@vlsi.es.kit.ac.jp
DOI: 10.1587/transele.2023ECP5045

Copyright © 2024 The Institute of Electronics, Information and Communication Engineers

Sect. 3. We conclude this paper in Sect. 4.

2. SET Pulse Width Measurement Structures

2.1 Conventional Measurement Structures

Typically, SET pulse widths are measured by a logic gate chain and a time-to-digital converter (TDC) such as vernier delay line [19], [20]. When a SET pulse is generated by radiation passing through one logic gate, the SET propagates through the logic gate chain and is input to the TDC. Thus, SET pulse width can be measured by storing the SET pulse width as a digital signal. The more the number of stages in the logic gate chain under measurement, the more area-efficient the measurement can achieve.

Figure 1 shows target logic chains to measure nSET and pSET separately, which was proposed by S. Jagannathan et al. Since the inverters are duplicated, a SET generated on the inverters does not change output of NOR gate or NAND gate. Thereby, SETs generated on off-state nMOSFETs in NAND gates are only measured by the N-hit target circuit, and we can measure nSETs and pSETs separately.

SET pulse width measurement using a logic gate chain has a problem due to propagation-induced pulse modulation (PIPM) effect [15]. Because of the difference in propagation delay time of the logic circuit between rising and falling signals, the pulse width increases or decreases linearly as the SET pulse passes through the logic circuit. Thus, the SET pulse width depends on the position in the logic gate chain through which the radiation passes. PIPM effects can occur even in inverter chains due to negative bias temperature instability (NBTI) [17], and it is particularly significant in logic circuits with different even and odd numbered stages, as in the block in Fig. 1. To reduce the PIPM effect, short logic gate chains connected to the inputs of a multi-input OR gate are used for the measurement target.

2.2 Target Structures

SET pulse width measurement circuits were implemented in a 65 nm bulk process. Their simplified schematic is shown in Fig. 2. It consists of 27 unit circuits as the SET source, duplicated 27-input NOR (or AND) gates and a time-to-digital converter (TDC). In this circuit, the output inversion of an unit circuit caused by a SET is transmitted to inputs of the TDC by the 27-input NOR gates, and the pulse width is saved as a digital signal by the TDC.

Figure 3 shows the target unit structures used to measure nSET widths, pSET widths and SET widths from the inverters. Figure 3 (a) and (b) are based on victim circuit structures reported in Ref. [21]. The structure for nSET was composed of one on-state pMOSFET and 30 off-state nMOSFETs connected in parallel. In this structure, only SET pulses from nMOSFETs were measured because single event effect is caused on off-state MOSFETs. Additionally, the SET pulse width does not change unintentionally by PIPM effect, therefore, SET pulse widths from nMOSFET and pMOSFET can be measured accurately and individually using the unit structures as shown in Fig. 3 (a) and (b). Furthermore, all target unit structures contain an equal number of off-state MOSFETs, so the area efficiency of the measurement circuit does not decrease when nSETs and pSETs are measured separately.

The SETs from the inverter chain (SET_{inv}) cannot separate nSETs and pSETs because the inverter chain contains 15 off-state pMOSFETs and 15 off-state nMOSFETs. SET_{inv} were measured for comparison with SETs calculated using nSETs and pSETs obtained from the circuits in Fig. 3 (a) and (b), respectively.

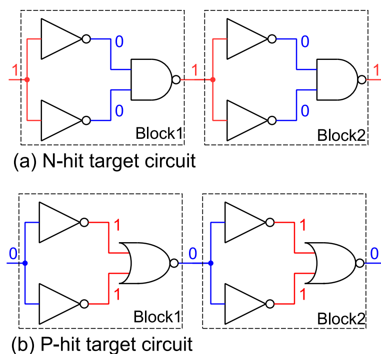


Fig. 1 Schematics of two of the blocks of (a) N-hit target circuit and (b) P-hit target circuit proposed by S. Jagannathan et al [12].

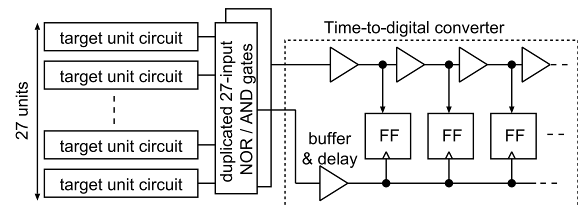


Fig. 2 Simplified schematic diagram of SET pulse width measurement circuit.

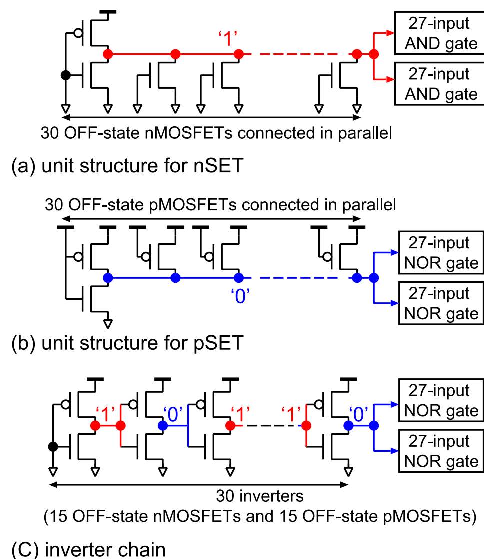


Fig. 3 Schematic diagrams of target unit circuits.

Figure 4 shows the structure of the 27-input NOR gate. It consists of 3-input NOR gates and 3-input NAND gates to reduce fanout less than four.

2.3 Time-to-Digital Converter

The TDC was implemented based on a snapshot circuit introduced in Ref. [11]. Figure 5 shows a detailed schematic diagram of the implemented TDC. Figure 6 shows timing chart of the TDC when a SET is simultaneously input to the two input ports, PULSE_IN and TRIG_IN. Input pulse width can be calculated by multiplying the delay time of the buffer and the number of positive-edge triggered FFs that capture flipped output values of the buffers. Measured SET pulse width is stored in the FFs as a sequence of '0'.

The TDC has two input ports to measure pulse width, PULSE_IN is for the input of the pulse to be measured, and TRIG_IN is used as a trigger to save the output values of the buffers. The pulse width is not stored unless the SET pulse is simultaneously propagated to the two input ports. Therefore, SETs caused in the TDC and the 27-input NOR/AND gates were not captured by FFs, and SETs caused in the target unit circuits are only measured and stored by the TDC. Therefore, the TDC does not need to be duplicated and achieves better area efficiency than conventional TDCs for SET pulse width measurement.

2.4 Fabricated Test Chip

150 SET pulse width measurement circuits were implemented in a 65 nm twin-well bulk process. We implemented single-finger (1x) and two-finger (2x) transistors versions of the three circuits shown in Fig. 3 as the evaluation targets. Two-finger transistors duplicate on-current, while the drain area is quite similar to that of a single-finger transistor. The gate widths of the nMOSFETs and pMOSFETs in Fig. 3 (a)–(c) are 390 nm and 520 nm per finger, respectively. The gate width ratio between pMOSFETs and nMOSFETs is 1.33, which is determined by optimizing the delay time and area in the standard cell library [22].

The simplified layouts of the implemented target circuits are partially shown in Fig. 7. Since there are 30 off-state MOSFETs in Fig. 3 (a) and (b), it is possible that a charged particle may affect multiple off-state MOSFETs at the same time. This phenomenon is called charge sharing or multiple-node charge collection [23]. Charge collection in multiple MOSFET connected in parallel can increase the SET pulse width significantly, making it impossible to measure SET pulse widths from a MOSFET. To mitigate this problem, only two off-state MOSFETs in the unit circuit were placed in the same well and an array of well-contacts were inserted between them to suppress multiple-node charge collection and parasitic bipolar effect. Additionally, the unit circuits are placed 1 μm from each other to avoid charge sharing between the two unit circuits and to prevent two SET pulses from occurring simultaneously [24].

Figure 8 shows calibration results of the fabricated TDCs in different chips by inputting square wave from a variable-stage ring oscillator and frequency divider circuit. The generated square wave was input to first stage of a 30-stage inverter chain. In this calibration, it was not possible to check the TDC operation using pulses of under 600 ps. However, resolution of the TDC determined by the delay time of the 30 buffers which does not differ significantly from buffer to buffer. Moreover, the calibration results agreed well with circuit-level simulation results. The resolution of the TDCs was about 35 ps when supply voltage was 1.2 V, which is less than 10% error from the value obtained by the circuit-level simulation. Therefore, we estimate that the TDC maintains a resolution of 35 ps even with input pulses below 600 ps.

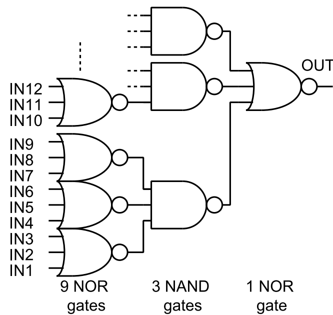


Fig. 4 The structure of the implemented 27-input NOR gate.

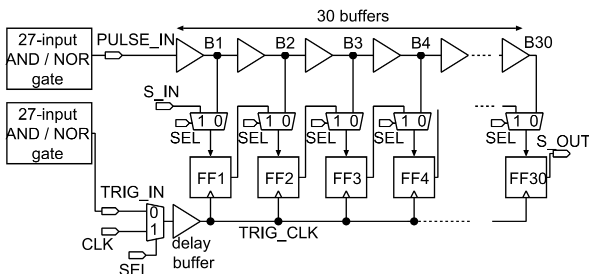


Fig. 5 Detailed schematic diagram of the implemented TDC. 31 multiplexers are inserted to perform 30 stage shift register which is used to initialize the TDC and to read the measurement.

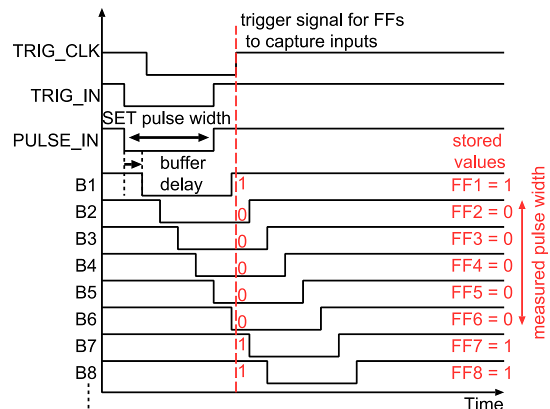


Fig. 6 Timing chart of the TDC when a SET is simultaneously input to the two input ports, PULSE_IN and TRIG_IN. The delay time of the multiplexers and the setup time of the FFs are ignored in this timing chart.

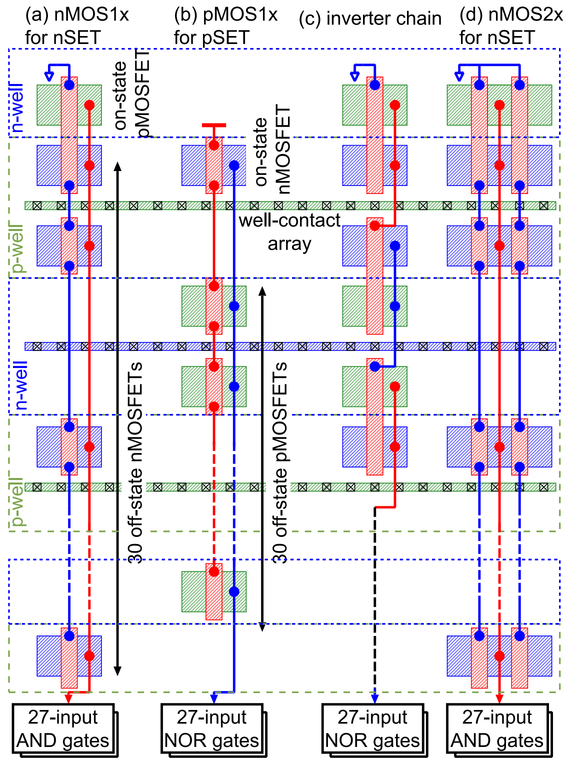


Fig. 7 Simplified layouts of target unit structures: (a) unit structure for nSET with one finger, (b) unit structure for pSET with one finger, (c) inverter chain with one finger and (d) unit structure for nSET with two fingers.

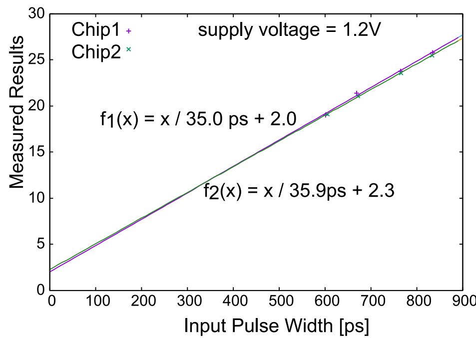


Fig. 8 Calibration results of the implemented TDCs.

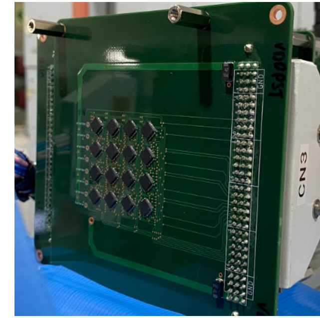


Fig. 9 Simultaneous measurement of 32 test chips using two PCB boards in the neutron irradiation test.

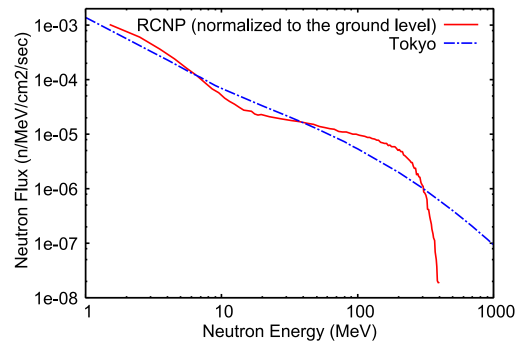


Fig. 10 Neutron spectrum at Tokyo and that from spallation neutron source at RCNP.

3. Experiment Results

3.1 Experiment Setup

Heavy ion tests were performed at Cyclotron and Radioisotope Center (CYRIC), Tohoku University, Japan. The chip was irradiated perpendicularly with $^{84}Kr^{17+}$. The energy and linear energy transfer (LET) values of the Kr ions are 322 MeV and 40 MeVcm^2/mg. Total fluence of the Kr ions was about 8,000,000 ions/cm^2.

The flux of heavy ions varies greatly with position in space. Therefore, the probability of heavy-ion-induced SETs is generally expressed using the cross-section (CS) calculated from the following equation [25].

$$CS\,[\mathrm{cm}^2/\mathrm{target}] = \frac{\mathrm{N_{SET}}}{\mathrm{F_{ION}} \times \mathrm{N_{target}}}, \quad (1)$$

where, N_{SET} is the number of measured SETs, F_{ION} is ion fluence per cm^2, N_{target} is the number of measurement target (off-state nMOSFETs, pMOSFETs or inverters).

High energy neutron tests using spallation neutron source were performed at Research Center for Nuclear Physics (RCNP), Osaka University, Japan. Figure 10 shows neutron spectrum at Tokyo and that from spallation neutron source at RCNP. In order to observe as many SET pulses as possible, two test boards with 16 chips each (32 in total) were measured simultaneously (Fig. 9). Chips were tested to a total fluence of 5×10^9 n/cm^2, which is equivalent to the total fluence of neutrons in 400,000 years at ground level (terrestrial neutron flux = 12.9 n/cm¹2/hr [26]). The probability of neutron-induced SETs is expressed using failure in time (FIT) rate. FIT rate is the number of SETs that can be expected in 10^9 hours.

3.2 SET Pulse Widths from pMOSFET and nMOSFET

Figures 11, 12 and 13 show the pulse width distributions of Kr-ion-induced nSET, Kr-ion-induced pSET and neutron-induced nSET, respectively. Since pSET was not induced by neutron tests, only the distribution of nSET pulse widths is shown in Fig. 13. Total number of SETs, total probability of SETs and average SET pulse width are summarized in Tables 1 and 2.

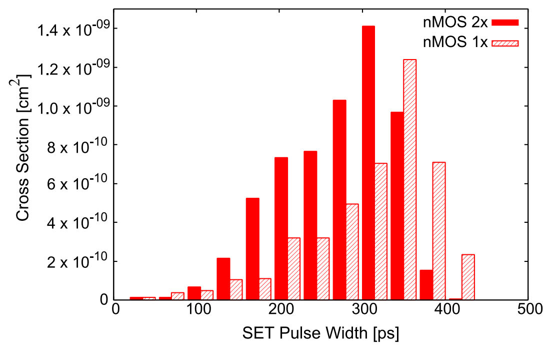


Fig. 11 Distribution of nSET pulse widths by Kr-ion irradiation.

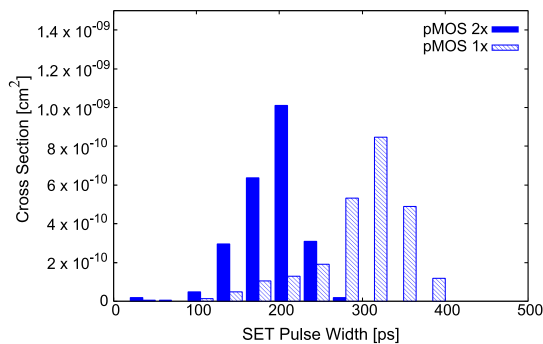


Fig. 12 Distribution of pSET pulse widths by Kr-ion irradiation.

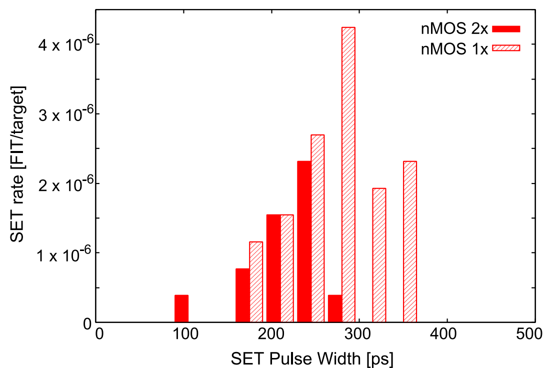


Fig. 13 Distribution of nSET pulse widths by neutron irradiation.

Table 1 Measurement results of SET pulses by Kr-ions irradiation.

target	finger	total number of SETs	total CS of SETs [cm^2/target]	avg. SET width [ps]
nMOSFET	2x	956	5.9×10^{-9}	270
	1x	703	4.3×10^{-9}	310
pMOSFET	2x	382	2.4×10^{-9}	190
	1x	404	2.5×10^{-9}	300

Table 2 Measurement results of SET pulses by neutron irradiation.

target	finger	total number of SETs	total SET rate [FIT/target]	avg. SET width [ps]
nMOSFET	2x	14	5.4×10^{-6}	150
	1x	36	1.4×10^{-5}	200
pMOSFET	2x	0	–	–
	1x	0	–	–

SETs from off-state nMOSFET have larger CS and longer average pulse width than those from off-state pMOSFET. Additionally, there were no pSET events during the neutron tests. These results are due to differences in electron and hole mobility. The electrons that are collected in the off-state nMOSFET had higher mobility than holes, and more charge is collected in the nMOSFET diffusion region. Furthermore, pMOSFET on-current is smaller than nMOSFET on-current in the measured structures. Thus, longer SETs were more frequently generated from nMOSFETs than from pMOSFETs. The nSETs have a greater impact on circuit reliability than the pSETs, at least when the transistor sizes evaluated in this paper are used in standard cells.

The nSETs and pSETs also have different degrees of dependence on the number of fingers (drive strength). In the Kr-ion tests, increasing the number of fingers reduced the pulse widths of the nSETs and pSETs by 10% and 40%, respectively. It is assumed that the pulse widths of pSETs are reduced by the on-current from nMOSFETs, which doubles with the increase in the number of fingers. In the case of nSETs, the increase in the number of fingers also doubles the on-current from pMOSFETs, but we assume that twice the number of off-state nMOSFETs increases the amount of radiation-induced parasitic bipolar current, so that the effect of twice the on-current from pMOSFETs become smaller [10]. These results are consistent with the fact that the current amplification factor of a lateral npn bipolar transistor is greater than that of a lateral pnp bipolar transistor in the 65 nm process design kit. Based on the above measurements, separate measurements of nSET and pSET are essential in SET modeling because the characteristics of nSET and pSET are different.

In a previous study, pSETs were reported to be approximately 3.9 times higher than nSETs [12], however, this is not consistent with our results. We assume that this is due to the difference in the gate width of the pMOSFETs. In Ref. [12], the gate widths of sensitive pMOSFET and nMOSFET were 1.3 μm and 400 nm, respectively, while, in our target circuits they were 520 nm and 390 nm. There is more than twice the difference in the gate width of the pMOSFETs.

In this experiment, we also measured the single event upset (SEU) rates on standard FFs using Kr ions and neutrons. The CS of Kr-ion-induced SEU rate was 1.50×10^{-8} cm^2/FF and the neutron-induced SEU rate was 5×10^{-4} FIT/FF. Neutron-induced SET rate was about two order of magnitude lower than the SEU rate. In contrast, CS of Kr-ion-induced SET was 1/6–1/4 of the SEU rate, which is a significant ratio since combinational circuits are basically composed of a large number of logic circuits. SET protection is also necessary in space where heavy ions exist. As can be inferred from the SEU rates on standard FFs, 851 and 533 SEUs on FFs in the TDC were also observed in the accelerated tests with neutrons and Kr ions, respectively. However, as shown in the calibration results (Fig. 8), SET measurement data is always stored in the FFs as three or more consecutive zeros, making it easily distinguishable from the SEUs.

3.3 nSET and pSETs versus SETs from Inverter Chains

SETs from a 30-stage inverter chain (Fig. 3 (c)) occur in 15 off-state nMOSFETs and 15 off-state pMOSFETs. Therefore, the SET cross section and error probability of the inverter chain, CS_{inv} and P_{inv} can be calculated from the following equations.

$$CS_{\mathrm{inv}} = \frac{CS_{\mathrm{pSET}} + CS_{\mathrm{nSET}}}{2}, \quad (2)$$

$$P_{\mathrm{inv}} = \frac{P_{\mathrm{pSET}} + P_{\mathrm{nSET}}}{2}. \quad (3)$$

In this section, we compare measured SETs from the inverter chain and SETs calculated from the measurement results of nSET and pSETs.

The measured and calculated SET rates in inverter chain are summarized in Tables 3 and 4. For both neutron and heavy ion tests, the average $\mathrm{SET_{inv}}$ pulse widths were about 100 ps shorter than the average $\mathrm{SET_{inv}}$ pulse width calculated by nSET and pSET. There are two possible reasons for this: one is the difference in load capacitance, and the other is multiple node charge collection in the inverter chain. In the nSETs and pSETs measurement circuits, 30 off-state MOSFETs provided a large junction capacitance and load capacitance are several times larger than those in the inverter chain. Thereby, the recovery speed from radiation-induced inversion became slower. The voltage variation due to an ion strike is also slower due to the large load capacitance. However, compared to the increase in recovery time, the time increase in radiation-induced voltage change is relatively small since the drift current due to a radiation strike has a large peak value [27], [28]. In addition, the drift is effective for a longer time due to the large load capacitance, and more significant charges are collected. As a result, longer SET pulses were more likely to occur in the nSETs and pSETs experiment circuits. Another reason could be pulse quenching, which is a reduction of SET pulse widths in the inverter chains due to multiple node charge collection [29]–[31]. Pulse quenching is caused by charge collecting in the on-state MOSFETs and the collected charge suppresses the inversion time (SET pulse width). Additionally, since the charge generated by radiation strikes change the well potential, it increases the on-current of the MOSFET as a forward body bias. Thus, the recovery speed increases from the SET inversion and the SET pulse width shortens. These effects are less likely to occur in nSETs and pSETs measurement circuits because there are only one on-state MOSFET and it is separated from most of the off-state MOSFETs (Fig. 7).

Measured and calculated distributions of the Kr-ion-induced $\mathrm{SET_{inv}}$ pulse widths are shown in Figs. 14 and 15. The distributions calculated by nSETs and pSETs were adjusted by subtracting the difference in average pulse widths to consider the effects described above. The calculated pulse width distributions were consistent with the measurement results using the inverter chain. In particular, the calculated distribution reproduced the measurement result of two-finger transistors that had two peaks: the peak around 100 ps is composed of nSETs and pSETs, while the other peak is composed of only nSETs (Fig. 15). The above results indicate that the proposed nSET and pSET measurement circuits can be used to analyze SET in inverters. For more accurate measurement, use a circuit that matches the load capacitance of the inverter chain and to consider multiple node charge collection in the inverter chain.

Table 3 Measurement and calculated results of Kr-ion-induced SETs from inverters.

	finger	total number of SETs	total CS of SETs [$\mathrm{cm^2}$/target]	avg. SET width [ps]
measured $\mathrm{SET_{inv}}$	2x	725	4.5×10^{-9}	180
	1x	711	4.4×10^{-9}	160
calculated $\mathrm{SET_{inv}}$	2x	–	4.2×10^{-9}	250
	1x	–	3.4×10^{-9}	310

Table 4 Measurement and calculated results of neutron-induced SETs from inverters.

	finger	total number of SETs	total SET rate [FIT/target]	avg. SET width [ps]
measured $\mathrm{SET_{inv}}$	2x	4	1.5×10^{-6}	40
	1x	13	5.0×10^{-6}	80
calculated $\mathrm{SET_{inv}}$	2x	–	2.7×10^{-6}	150
	1x	–	7.0×10^{-6}	200

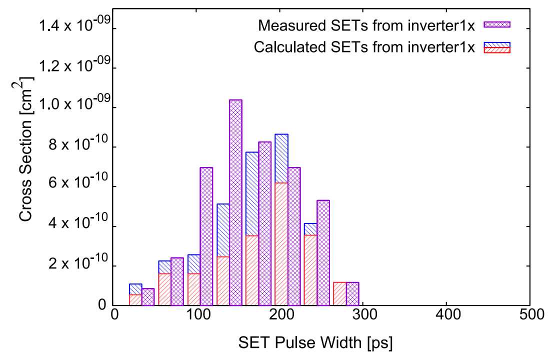


Fig. 14 Measured and calculated distribution of SET pulse widths from inverter 1x. Red and blue bars represent nSET and pSET, respectively.

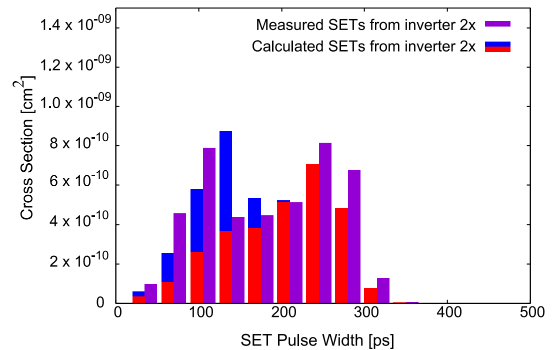


Fig. 15 Measured and calculated distribution of SET pulse widths from inverter 2x. Red and blue bars represent nSET and pSET, respectively.

4. Conclusion

SET pulse widths for pMOSFETs were measured separately from those for nMOSFETs in a 65 nm bulk process using off-state pMOSFETs or nMOSFETs connected in parallel. Measurement results show that nMOSFETs had twice the the heavy-ion-induced SET rate as pMOSFETs and neutron-induced SETs were only observed from nMOSFETs. In addition, the nSETs and pSETs also have different degrees of dependence on the number of fingers. In the Kr-ion tests, increasing the number of fingers reduced the pulse widths of the nSETs and pSETs by 10% and 40%, respectively. Based on the above results, separate measurements of nSET and pSET are essential in SET modeling because the characteristics of nSET and pSET are different. On the ground level, it is less effective for circuit reliability to protect pMOSFETs from single event effects when the transistor sizes of the nMOSFET and pMOSFET are roughly equivalent.

This measurement also compared SET from inverter chains with nSETs and pSETs, and observed that nSETs and pSETs had larger average SET pulse width than SET from inverters, due to differences in load capacitance and multi-node charge collection. However, it was confirmed that the peaks in the pulse width distribution of SETs from inverter chains could be reproduced by the pulse width distribution calculated from measurement results of nSETs and pSETs. More accurate nSETs and pSETs measurements can be achieved by reducing the number of MOSFETs connected in parallel, thereby reducing load capacitance. By reducing the load capacitance of the target circuits, the SET distribution of an inverter can be calculated from pSET and nSET without measuring the average SET of an inverter.

Acknowledgments

The authors would like to thank to Cyclotron and Radioisotope Center (CYRIC), Tohoku University and Research Center for Nuclear Physics (RCNP), Osaka University for heavy ion and high energy neutron experiments. This work was supported through the activities of VDEC, The University of Tokyo, in collaboration with Cadence Design Systems, NIHON SYNOPSYS G.K. and Siemens Electronic Design Automation Japan K.K.

References

[1] T. Karnik and P. Hazucha, "Characterization of soft errors caused by single event upsets in CMOS processes," IEEE Trans. Dependable Secure Comput., vol.1, no.2, pp.128–143, 2004.
[2] H. Kobayashi, N. Kawamoto, J. Kase, and K. Shiraish, "Alpha particle and neutron-induced soft error rates and scaling trends in SRAM," IEEE International Reliability Physics Symposium, pp.206–211, 2009.
[3] N. Seifert, P. Slankard, M. Kirsch, B. Narasimham, V. Zia, C. Brookreson, A. Vo, S. Mitra, B. Gill, and J. Maiz, "Radiation-induced soft error rates of advanced CMOS bulk devices," IEEE International Reliability Physics Symposium, pp.217–225, 2006.
[4] B. Narasimham, V. Chaudhary, M. Smith, L. Tsau, D. Ball, and B. Bhuva, "Scaling trends in the soft error rate of SRAMs from planar to 5-nm finfet," 2021 IEEE International Reliability Physics Symposium (IRPS), pp.1–5, 2021.
[5] N.N. Mahatme, I. Chatterjee, B.L. Bhuva, J. Ahlbin, L.W. Massengill, and R. Shuler, "Analysis of soft error rates in combinational and sequential logic and implications of hardening for advanced technologies," IEEE International Reliability Physics Symposium, pp.1031–1035, 2010.
[6] B. Gill, N. Seifert, and V. Zia, "Comparison of alpha-particle and neutron-induced combinational and sequential logic error rates at the 32nm technology node," IEEE International Reliability Physics Symposium, pp.199–205, 2009.
[7] P. Eaton, J. Benedetto, D. Mavis, K. Avery, M. Sibley, M. Gadlage, and T. Turflinger, "Single event transient pulsewidth measurements using a variable temporal latch technique," IEEE Trans. Nucl. Sci., vol.51, no.6, pp.3365–3368, 2004.
[8] E. Keren, S. Greenberg, N.M. Yitzhak, D. David, N. Refaeli, and A. Haran, "Characterization and mitigation of single-event transients in Xilinx 45-nm SRAM-based FPGA," IEEE Trans. Nucl. Sci., vol.66, no.6, pp.946–954, 2019.
[9] M.J. Gadlage, J. Ahlbin, B. Narasimham, V. Ramachandran, C.A. Dinkins, N.D. Pate, B.L. Bhuva, R.D. Schrimpf, L.W. Massengill, R.L. Shuler, and D. McMorrow, "Increased single-event transient pulsewidths in a 90-nm bulk CMOS technology operating at elevated temperatures," IEEE Trans. Device Mater. Rel., vol.10, no.1, pp.157–163, 2010.
[10] J. Furuta, R. Yamamoto, K. Kobayashi, and H. Onodera, "Evaluation of parasitic bipolar effects on neutron-induced SET rates for logic gates," IEEE International Reliability Physics Symposium, pp.SE.5.1–SE.5.5, 2012.
[11] Y. Yanagawa, K. Hirose, H. Saito, D. Kobayashi, S. Fukuda, S. Ishii, D. Takahashi, K. Yamamoto, and Y. Kuroda, "Direct measurement of SET pulse widths in 0.2-μm SOI logic cells irradiated by heavy ions," IEEE Trans. Nucl. Sci., vol.53, no.6, pp.3575–3578, 2006.
[12] S. Jagannathan, M.J. Gadlage, B.L. Bhuva, R.D. Schrimpf, B. Narasimham, J. Chetia, J.R. Ahlbin, and L.W. Massengill, "Independent measurement of SET pulse widths from n-hits and p-hits in 65-nm CMOS," IEEE Trans. Nucl. Sci., vol.57, no.6, pp.3386–3391, 2010.
[13] C. Jianjun, C. Shuming, C. Yaqing, and L. Bin, "Characterization of single-event transient pulse quenching among dummy gate isolated logic nodes in 65 nm twin-well and triple-well CMOS technologies," IEEE Trans. Nucl. Sci., vol.62, no.5, pp.2302–2309, 2015.
[14] B. Narasimham, B.L. Bhuva, R.D. Schrimpf, L.W. Massengill, M.J Gadlage, W.T. Holman, A.F. Witulski, W.H. Robinson, J.D. Black, J.M. Benedetto, and P.H. Eaton, "Effects of guard bands and well contacts in mitigating long SETs in advanced CMOS processes," 9th European Conference on Radiation and Its Effects on Components and Systems, pp.1–6, 2007.
[15] V. Ferlet-Cavrois, P. Paillet, D. McMorrow, N. Fel, J. Baggio, S. Girard, O. Duhamel, J.S. Melinger, M. Gaillardin, J.R. Schwank, P.E. Dodd, M.R. Shaneyfelt, and J.A. Felix, "New insights into single event transient propagation in chains of inverters—Evidence for propagation-induced pulse broadening," IEEE Trans. Nucl. Sci., vol.54, no.6, pp.2338–2346, 2007.
[16] M.J. Gadlage, J.R. Ahlbin, B.L. Bhuva, N.C. Hooten, N.A. Dodds, R.A. Reed, L.W. Massengill, R.D. Schrimpf, and G. Vizkelethy, "Alpha-particle and focused-ion-beam-induced single-event transient measurements in a bulk 65-nm CMOS technology," IEEE Trans. Nucl. Sci., vol.58, no.3, pp.1093–1097, 2011.
[17] R. Harada, Y. Mitsuyama, M. Hashimoto, and T. Onoye, "Impact of NBTI-induced pulse-width modulation on SET pulse-width measurement," IEEE Trans. Nucl. Sci., vol.60, no.4, pp.2630–2634, 2013.
[18] N. Mikami, T. Nakauchi, A. Oyama, H. Kobayashi, and H. Usui, "Role of the deep parasitic bipolar device in mitigating the single event transient phenomenon," IEEE International Reliability Physics

Symposium, pp.936–939, 2009.
[19] C.L.-M. de Boissac, F. Abouzeid, V. Malherbe, G. Gasiot, P. Roche, and J.-L. Autran, "Influence of supply voltage and body biasing on single-event upsets and single-event transients in UTBB FD-SOI," IEEE Trans. Nucl. Sci., vol.68, no.5, pp.850–856, 2021.
[20] R. Harada, Y. Mitsuyama, M. Hashimoto, and T. Onoye, "Set pulse-width measurement eliminating pulse-width modulation and within-die process variation effects," IEEE International Reliability Physics Symposium, pp.SE.1.1–SE.1.6, 2012.
[21] Z. Li, L. Berti, J. Wouters, J. Wang, and P. Leroux, "Characterization of the total charge and time duration for single-event transient voltage pulses in a 65-nm CMOS technology," IEEE Trans. Nucl. Sci., vol.69, no.7, pp.1593–1601, 2022.
[22] D.S. Kung and R. Puri, "Optimal P/N width ratio selection for standard cell libraries," 1999 IEEE/ACM International Conference on Computer-Aided Design, Digest of Technical Papers (Cat. no.99CH37051), pp.178–184, 1999.
[23] L.W. Massengill, O.A. Amusan, S. Dasgupta, A.L. Sternberg, J.D. Black, A.F. Witulski, B.L. Bhuva, and M.L. Alles, "Soft-error charge-sharing mechanisms at sub-100nm technology nodes," 2007 IEEE International Conference on Integrated Circuit Design and Technology, pp.1–4, 2007.
[24] J. Chen, J. Yu, P. Yu, B. Liang, and Y. Chi, "Characterization of the effect of pulse quenching on single-event transients in 65-nm twin-well and triple-well CMOS technologies," IEEE Trans. Device Mater. Rel., vol.18, no.1, pp.12–17, 2018.
[25] J.S. Kauppila, T.D. Loveless, R.C. Quinn, J.A. Maharrey, M.L. Alles, M.W. McCurdy, R.A. Reed, B.L. Bhuva, L.W. Massengill, and K. Lilja, "Utilizing device stacking for area efficient hardened SOI flip-flop designs," 2014 IEEE International Reliability Physics Symposium, pp.SE.4.1–SE.4.7, 2014.
[26] JEDEC Standard, "Measurement and reporting of alpha particle and terrestrial cosmic ray-induced soft errors in semiconductor devices," JESD89B, Sept. 2021.
[27] Y.M. Aneesh and B. Bindu, "A physics-based single event transient pulse width model for CMOS VLSI circuits," IEEE Trans. Device Mater. Rel., vol.20, no.4, pp.723–730, 2020.
[28] D.A. Black, W.H. Robinson, I.Z. Wilcox, D.B. Limbrick, and J.D. Black, "Modeling of single event transients with dual double-exponential current sources: Implications for logic cell characterization," IEEE Trans. Nucl. Sci., vol.62, no.4, pp.1540–1549, 2015.
[29] J.R. Ahlbin, L.W. Massengill, B.L. Bhuva, B. Narasimham, M.J. Gadlage, and P.H. Eaton, "Single-event transient pulse quenching in advanced CMOS logic circuits," IEEE Trans. Nucl. Sci., vol.56, no.6, pp.3050–3056, 2009.
[30] K. Zhang, R. Yamamoto, J. Furuta, K. Kobayashi, and H. Onodera, "Parasitic bipolar effects on soft errors to prevent simultaneous flips of redundant flip-flops," 2012 IEEE International Reliability Physics Symposium (IRPS), pp.5B.2.1–5B.2.4, 2012.
[31] J.R. Ahlbin and P. Gadfort, "Impact of ultra-low voltages on single-event transients and pulse quenching," 2014 SOI-3D-Subthreshold Microelectronics Technology Unified Conference (S3S), pp.1–3, 2014.

Jun Furuta was born in Gifu, Japan, in 1986. He received the B.E., M.E., and Ph.D. degrees in electrical and electronic engineering in informatics from Kyoto University, Kyoto, Japan, in 2009, 2011, and 2014, respectively. In 2014, he joined the Kyoto Institute of Technology as a specially appointed Assistant Professor, where he was then promoted to an Assistant Professor in 2017.

Shotaro Sugitani was born in Kyoto, Japan in 1998. He received the B.E. degree in Electronics and System Engineering from Kyoto Institute of Technology, Kyoto, Japan in 2022. He is currently a master course student at Kyoto Institute of Technology. He is a student member of IEEE, IPSJ and IEICE.

Ryuichi Nakajima was born in Hiroshima, Japan in 1997. He received the B.E. and the M.E. degrees in Electronics and System Engineering from Kyoto Institute of Technology, Kyoto, Japan in 2021 and 2023, respectively. He is currently a doctoral course student at Kyoto Institute of Technology. He is a student member of IEEE, IPSJ and IEICE.

Takafumi Ito was born in Shiga, Japan in 1998. He received the B.E. and the M.E. degrees in Electronics and System Engineering from Kyoto Institute of Technology, Kyoto, Japan in 2021 and 2023, respectively. He is a member of IPSJ.

Kazutoshi Kobayashi received his B.E., M.E., and Ph.D. in Electronic Engineering from Kyoto University, Japan in 1991, 1993, 1999, respectively. Starting as an Assistant Professor in 1993, he was promoted to an Associate Professor in the Graduate School of Informatics, Kyoto University in 2001, and stayed in that position until 2009. For two years during this time, he acted as an Associate Professor of VLSI Design and Education Center (VDEC) at the University of Tokyo. Since 2009, he has been a Professor at Kyoto Institute of Technology. While in the past he focused on reconfigurable architectures utilizing device variations, his current research interest is in improving the reliability (Soft Errors, Bias Temperature Instability and Plasma Induced Damage) of current and future VLSIs. He started a research project related to gate drivers for GaN and SiC transistors since 2013. In 2022, he was nominated for one of the project managers of Moonshot Goal 6 "Realization of a fault-tolerant universal quantum computer that will revolutionize economy, industry, and security by 2050" and organized his project team named "Development of Scalable Highly Integrated Quantum Bit Error Correction System" abbreviated as QUBECS. He was the recipient of the IEICE best paper award in 2009, the IRPS best poster award in 2013, the ICICDT best paper award in 2019 and the IEICE Electronics Society award in 2021.