IEEE Journal on Exploratory Solid-State Computational Devices and Circuits

Received 1 October 2021; revised 22 October 2021; accepted 25 October 2021.
Date of publication 27 October 2021; date of current version 13 December 2021.

Digital Object Identifier 10.1109/JXCDC.2021.3123783

Cryogenic Operation of 3-D Flash Memory for Storage Performance Improvement and Bit Cost Scaling

TOMOYA SANUKI[1] (Member, IEEE), YUTA AIBA[1], HITOMI TANAKA[1], TAKASHI MAEDA[1], KEIICHI SAWA[1], FUMIE KIKUSHIMA[2], and MASAYUKI MIURA[3]

[1]Institute of Memory Technology Research and Development, Kioxia Corporation, Yokkaichi 512-0906, Japan
[2]Advanced Memory Development Division, Kioxia Corporation, Yokkaichi 512-0906, Japan
[3]Memory Division, Kioxia Corporation, Yokohama 247-8585, Japan

CORRESPONDING AUTHOR: T. SANUKI (tomoya.sanuki@kioxia.com)

ABSTRACT This report introduces the cryogenic operation and storage performance of 3-D flash memory. The cell transistor characteristics and the basic functionalities, including read and program and erase (P/E) operations, are investigated at an extremely low temperature of 77 K cooled by the liquid nitrogen. The cell transistor has a steep subthreshold slope at 77 K compared with at 300 K, and the read operation is fully functional even though the saturation current becomes relatively small. P/E operations at 77 K are not much different from those at 300 K, and the same incremental-step pulse programming (ISPP) and incremental-step pulse erase (ISPE) methods are applicable. In addition, the storage performance and the reliability characteristics of 3-D flash memory, such as the read noise, the disturb characteristics, the data retention, and the cycle endurance, are also investigated. While there is no degradation in the disturb characteristics, the read noise at 77 K was significantly minimized compared with that at 300 K. The data retention characteristics are about three times improved, and the cycle endurance is about ten times improved at 77 K compared with at 300 K. These storage performance improvements and high-reliability characteristics at cryogenic operation enable us to achieve the ultramultilevel cell. We show the successful demonstration of 6-bit per cell (HLC) and discuss the impact of additional cooling cost and the possibility of future storage devices beyond HLC. The cryogenic operation of 3-D flash memory can greatly improve the storage performance and opens the door for potential new applications and the bit cost scaling for future storage devices.

INDEX TERMS 3-D flash memory, 6-bit per cell (HLC), 77 K, bit cost scaling, cryogenic, immersion cooling, liquid nitrogen, NAND.

I. INTRODUCTION

Since NAND flash memory, one of the lowest cost and the highest bit density nonvolatile memory, was announced 30 years ago [1], the continuous increase in bit density and bit cost scaling has been achieved including evolutional conversion from 2-D planar to 3-D flash memory [2], [3]. NAND flash memory and SSD have expanded their applications not only to consumer electronics but also to industrial, scientific, medical, and social infrastructures and are used in a wide variety of fields. In general, NAND flash memory is assumed to operate at a temperature range from 0 °C to 85 °C. In some special applications, such as automotive devices, it is required to operate at a wide temperature range from −40 °C to 125 °C. However, it is not expected to operate at a temperature lower than −40 °C and is not guaranteed as a product specification.

In recent years, there has been an increasing demand for a memory that operates at a very low temperature, and one of the applications is quantum computers. Since many of the currently proposed quantum computers operate at the mK range, the computer systems to control the quantum computers are also required to operate at relatively low temperatures. So far, reports of cryogenic operations are limited to CMOS logic devices, such as CPUs [4], [5] and DRAMs [6]–[8].

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

In addition, it has been reported that the memory operation at 77 K (−196 °C) cooled by the liquid nitrogen instead of 4.2 K or lower is cost-effective and sufficient to support superconducting operation [7].

Another possible application of cryogenic operation is in the aerospace field [9]. Computers and electronic devices used in satellites, planetary probes, lunar bases, and so on are exposed to cryogenic environments, and these devices are required to operate at extremely low temperatures. A number of commercial-off-the-shelf (COTS) components, such as semiconductor chips, sensors, and discrete components, are already used in aerospace. With regard to the nonvolatile memory, there have been several reports about the characteristics of the serial flash memory IC [10] and SONOS transistor [11] at cryogenic temperature. However, there have been no reports about the cryogenic operation of flash memory that has a large bit capacity, including both 2-D planar and 3-D.

While the quantum computer and the aerospace application inevitably require the operation at cryogenic temperature, we are able to use it for boosting the semiconductor device performance. For CPU and logic devices, there was a report of early practices of cryogenic supercomputers [12]. Recently, improvements in the tradeoff between the switching speed and the power consumption at the state-of-the-art Fin FET device at a cryogenic temperature of 77 K have been reported [13]. For DRAM, it has been reported that the data retention characteristics, one of the most important memory performance, are significantly improved at 77 K [8]. For read and write latency, the power consumption and the refresh operation are also improved; then, we are able to apply the cryogenic operation to boost DRAM performance [14], [15]. As one of the emerging memory devices, cryogenic operation of floating body cells (FBC) has been reported [16]. The disadvantage of relatively short data retention time of FBC characteristics is drastically improved at extremely low temperature, and it may become a candidate for future memory devices to replace DRAM. As another emerging memory device, STT-MRAM at the cryogenic operation of 77 K has been reported that it outperforms six-transistor static random access memory (6T-SRAM) in terms of both dynamic and static power, as well as read access latency [17]. CMOS compatible two-transistor gain cell (2T-Gain Cell) [18], [19] and the processing-in-memory (PIM) [20] were also investigated and resulting significant improvements at the cryogenic temperature.

In the case of the nonvolatile memory, the device performance to be boosted is the storage characteristics, unlike CMOS logic devices, working memory devices, and embedded memory devices. It would be desirable to improve the storage performance, such as the data retention and the cycle endurance, rather than the switching speed, the power consumption, and the latency. For the 3-D flash memory, we previously reported that the cryogenic operation of the 3-D flash memory is fully functional and has the potential for ultramultilevel operation by improving the read noise characteristics, the data retention, and the cycle endurance [21]. We also reported the failure bit analysis and the cell transistor characteristics of our state-of-the-art 3-D flash memory at the cryogenic temperature and discussed the advantage of cryogenic operation for scaling in future 3-D flash memory [22]. In this article, we review these results and explain, in more detail, the characteristics of 3-D flash memory at cryogenic temperature. We propose that the cryogenic operation of the 3-D flash memory is advantageous not only for the application to new fields but also for the improvement of storage performance and the bit cost scaling of future storage devices.

II. COOLING WITH LIQUID NITROGEN

Our experimental system for cryogenic operation of 3-D flash memory uses immersion cooling with liquid nitrogen. The measurement system consists of a 3-D flash memory chip, a memory controller, and a host PC. Only the 3-D flash memory chip is immersed into the liquid nitrogen of 77 K, and the memory controller and the host PC are put in the room temperature environment of 300 K. The memory controller executes firmware to control the memory chip and sends the input and output commands and the data. The 3-D flash memory chip receives commands issued from the controller and performs read, and program and erase (P/E) operations in the same way as it does when operated at room temperature. We use the state-of-the-art 3-D flash memory chip that has highly stacked WL layers and the cell transistors with precisely optimized structures, including the charge trap (CT) layer. The package device of 3-D flash memory chips was newly developed so that it can operate in the liquid nitrogen of 77 K. Especially, the wire bonding in the package device has a hollow structure to minimize the stress in order to operate within a wide temperature range from 300 K of room temperature to 77 K of cryogenic temperature. In addition, the memory controller is put at room temperature because it is not optimized to operate at the cryogenic temperature. We expect that it can be operated at extremely low temperatures together with memory chips in the future.

III. BASIC CHARACTERISTICS AND FUNCTIONALITY

A. CELL TRANSISTOR CHARACTERISTICS

The typical cell transistor with word-line (WL) and the selected gate–source (SGS) transistor of 3-D flash memory consist of a thin polysilicon channel. There have been reports on the low-temperature characteristics of polysilicon channel transistors [23]; however, the transistor characteristics of the 3-D flash memory at the cryogenic temperature were not well understood. The temperature dependence of the I–V curve of the cell transistor is shown in Fig. 1. The subthreshold slope (S.S) of the cell transistor is significantly improved at 77 K and about $\times 1/2$ compared to 300 K. Steep S.S at the cryogenic temperature is obtained even at polysilicon channel transistors as same as a crystal-silicon channel transistor. This experimental result gives evidence of the predicted model for improvement of read noise characteristics, which is discussed later in this article. It is found that the saturated current of the cell transistor is decreased at cryogenic temperature

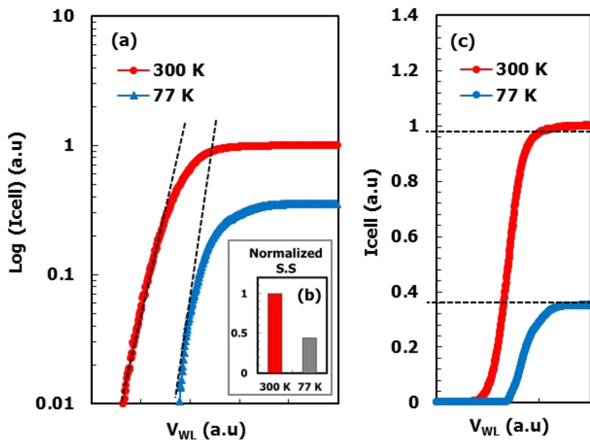

FIGURE 1. *I*–V curves of cell transistor measured at 300 K and 77 K for (a) log plot, (b) normalized S.S, and (c) linear plot. S.S is significantly improved at 77 K and about ×1/2 compared with 300 K. The saturated current of the cell transistor decreases to about ×1/3 at 77 K compared with 300 K.

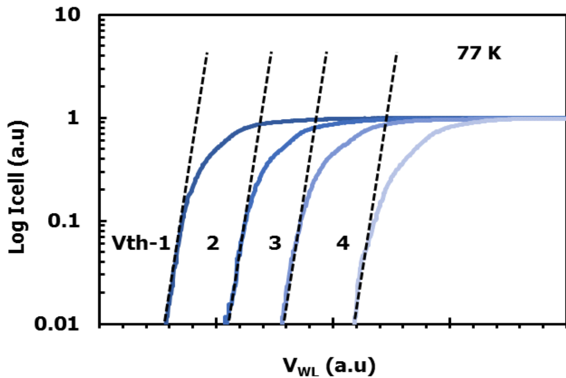

FIGURE 2. Log plot of *I*–V curves of cell transistor for several programmed Vth at 77 K. Steep S.S due to the cryogenic operation is maintained within wide Vth range.

and about ×1/3 at 77 K compared to 300 K. The impact of the decreased saturated current is considered as follows. In the read operation, the actual read current to be sensed is much smaller than the saturation current of the cell transistor because a limited overdrive voltage is applied to the selected gate (WL). The read sensing margin is not directly affected by the saturated current and further affected by the S.S of the cell transistor. Since the S.S is significantly improved at the cryogenic temperature, the read sensing margin is also improved even with the reduction of the saturation current. In fact, in our experiment, the read and verify operations are fully functional without any modification of read bias condition or timing.

Fig. 2 shows programmed cell Vth dependence of the *I*–V characteristics. The saturated current does not depend on the number of stored charges, and the S.S is also independent of programmed cell Vth. Steep S.S of the cell transistor at the cryogenic temperature is maintained within wide range of programmed Vth, and it is a critically important characteristic to realize tight Vth distribution for the ultramultilevel cell.

B. SGS TRANSISTOR CHARACTERISTICS

The SGS transistor of the 3-D flash memory has a different structure compared with the cell transistor. There is no CT layer between the channel and the gate electrode. It has a vertical crystal-silicon channel and is placed at the bottom side of string cells. The temperature dependence of the *I*–V curve of the SGS transistor is shown in Fig. 3. As same as the *I*–V curves of the cell transistor, S.S is slightly better at 77 K than at 300 K. It means that the cutoff characteristics of the SGS transistor are improved at 77 K.

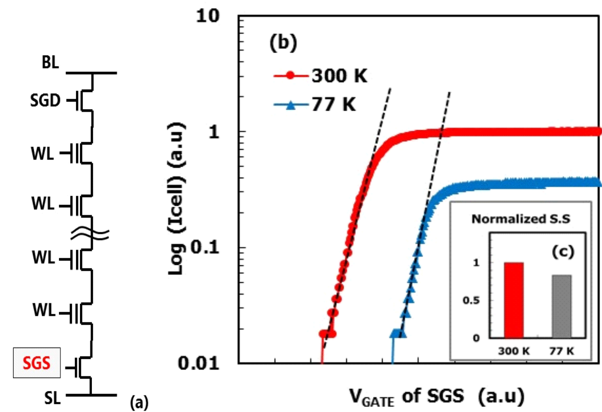

FIGURE 3. (a) Illustration of the SGS, (b) log plots of *I*–V curves of the SGS transistor, and (c) normalized S.S measured at 300 K and 77 K. As same as *I*–V curves of cell transistor, S.S of SGS is slightly better at 77 K than at 300 K.

C. CROSS TEMPERATURE EFFECT

We obtained the cross temperature effect of the cell Vth characteristics when the cell Vth was observed at one temperature and reobserved at another temperature. Fig. 4 shows that cell Vth difference between programmed at 300 K and measured Vth at 77 K (a), and programmed at 77 K and measured Vth at 300 K (b). The cell Vth becomes high while reducing temperature and the cell Vth difference as the cross temperature effect between 300 K and 77 K is determined about 0.75 V. This observed cross temperature effect is not large and within the manageable range. It is indicated that we are able to use 3-D flash memory with a wide temperature range between normal operating temperature around room temperature and the extremely low temperature of cryogenic temperature. In particular, when the 3-D flash memory is cooled by the immersion cooling of liquid nitrogen or another similar method, there is no need to care about the cross temperature effect because the entire 3-D flash memory is always cooled down at constant cryogenic temperature.

D. PROGRAM AND ERASE CHARACTERISTICS

Controlling the Vth of cell transistors by the P/E operation is an essential feature of the 3-D flash memory. Fig. 5 shows the time dependence of the cell Vth by applying a constant

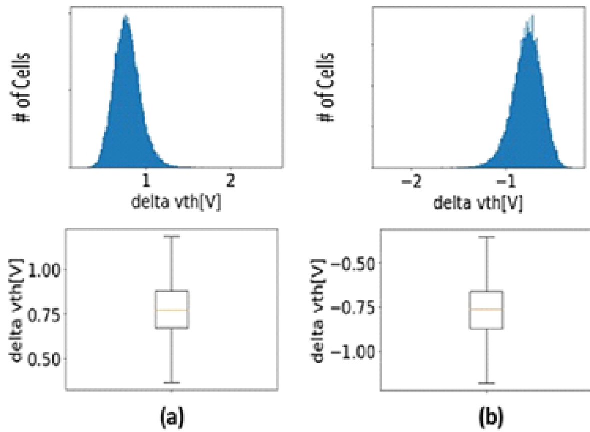

FIGURE 4. Cells are (a) programmed at 300 K and measured Vth at 77 K and (b) programmed at 77 K and measured Vth at 300 K. Cell Vth becomes high while reducing temperature and Vth difference as the cross temperature effect between 300 K and 77 K is determined to be about 0.75 V.

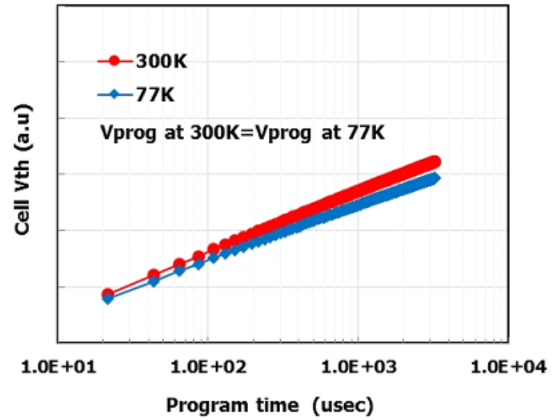

FIGURE 5. Program characteristics of cell Vth with a constant program voltage at 300 K and 77 K. The same program voltage is applied at both 300 K and 77 K.

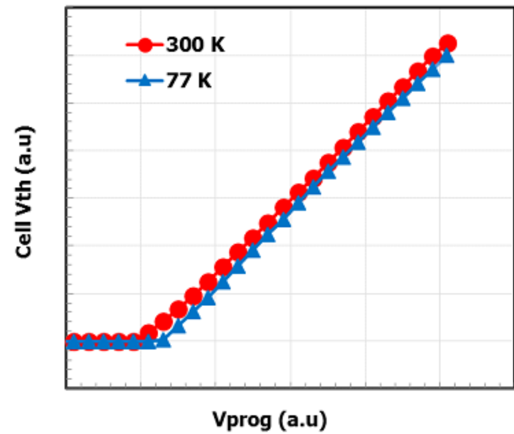

FIGURE 6. Program characteristics of cell Vth by applying ISPP at 300 K and 77 K. The program slope at 77 K has an almost equivalent value compared with 300 K.

program voltage. It is confirmed that the cell Vth rises as applying program voltage even at 77 K. By applying the same program voltage and time, there seems to be a difference in cell Vth in the long program time domain of 1 ms or more. However, the cell Vth difference is relatively small in the short program time domain that is normally used for the program operation.

The incremental-step pulse programming (ISPP) method is usually applied to 3-D flash memory, which can provide a tight Vth distribution control for the multilevel cell. Fig. 6 shows the program characteristics when the ISPP method is applied at 77 K and 300 K. The program slope of the ISPP method at 77 K is linear and has an almost equivalent value of slope compared to 300 K. It is clearly indicated that the same program operation is applicable at the cryogenic temperature.

The observed temperature dependence of the program characteristics at the cryogenic temperature is explained by carrier behavior during operations. Programming of the cell transistor is caused by the F-N tunneling by the electrons; therefore, the program operation has a small temperature dependence due to F-N tunneling. We can expect the same characteristics for program operation even at the cryogenic temperature.

Fig. 7 shows the time dependence of the cell Vth by applying a constant erase voltage. It is confirmed that the cell Vth is reduced by applying the erase voltage even at 77 K. However, a relatively higher erase voltage of about +2 V than 300 K is necessary to obtain the same cell Vth at 77 K. If we use the same erase voltage, relatively longer time to apply erase voltage is necessary.

Fig. 8 shows the characteristics of the cell Vth when the incremental-step pulse erase (ISPE) method is applied. The effective erase voltage is higher at 77 K than 300 K. By increasing the erase voltage, almost the same erase operation is applicable at the cryogenic temperature.

The observed temperature dependence of the erase characteristics at the cryogenic temperature is also explained by carrier behavior during erase operations. Erasing of the cell transistor is caused by the combination of the F-N tunneling by the holes and the Poole–Frenkel (P-F) by the electrons. The P-F depends on the excitation of the carrier and has temperature dependence. Therefore, the erase operation becomes less effective at low temperatures due to the P-F. By increasing erase voltage, we can obtain the same erase characteristics at cryogenic temperature.

IV. PERFORMANCE AND RELIABILITY

A. READ NOISE CHARACTERISTICS

The read Vth difference, obtained by applying multiple times of the read operation, is known as the read noise characteristics. The distribution of read Vth difference (σ) is determined by fitting the cumulative probability for one of the median cells, and the histogram of 16-KB cells is shown in Fig. 9. The read Vth difference is improved at a high temperature of 358 K (85 °C) compared to 300 K (27 °C), and this

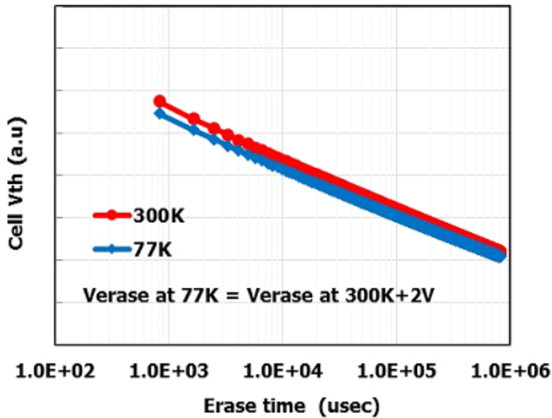

FIGURE 7. Erase characteristics of cell Vth with a constant voltage at 300 K and 77 K. Relatively higher erase voltage is applied at 77 K compared with 300 K.

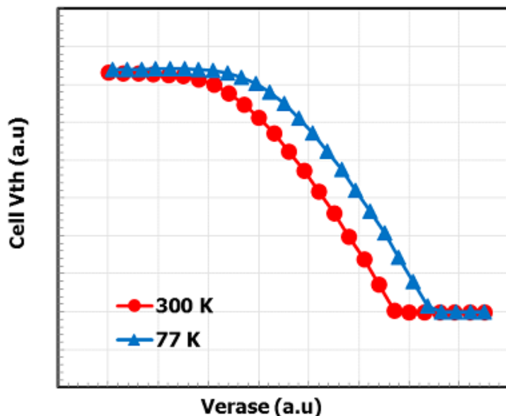

FIGURE 8. Erase characteristics of cell Vth by applying ISPE at 300 K and 77 K. The operational erase voltage is shifted to the upper side at 77 K.

behavior is consistent with previous reports [24], [25]. Within the temperature range between 0 °C and 85 °C, lower temperature becomes the worst case. On the other hand, the read Vth difference is significantly improved at 77 K compared with both 300 K and 358 K.

The mechanism of the read noise improvement at cryogenic temperature is explained as follows. Even at cryogenic temperature, the subthreshold slope becomes steeper while reducing temperature, as shown in Figs. 1 and 2. During the read operation, a limited overdrive voltage (read voltage) is applied to the selected WL, and the sensing current is determined by the cell transistor characteristics in this region. It means that the actual read current is much smaller than the saturation current and much affected by the S.S characteristics of the cell transistor. Then, the read noise is considered to be improved due to steep S.S, and the sensing current is sufficient enough to increase the signal-to-noise ratio at the cryogenic temperature.

B. DISTURB CHARACTERISTICS

The program disturb characteristics at the inhibit cell are one of the critical issues for the 3-D flash memory. Fig. 10 shows

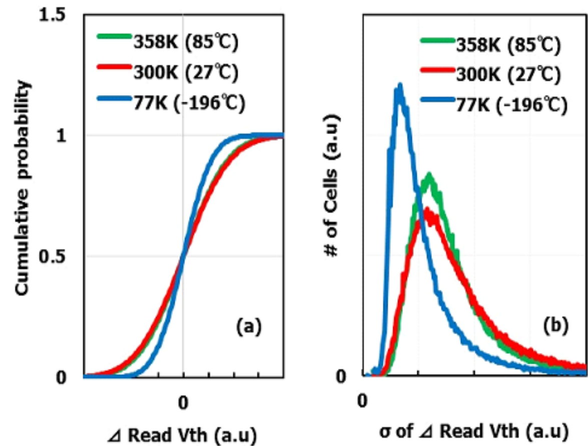

FIGURE 9. (a) Cumulative probability of read Vth difference for one of the median cells at 358 K, 300 K, and 77 K. The distribution of read Vth difference (σ) is determined by fitting. (b) Histogram of σ of read Vth difference for 16-KB cells. Read noise characteristics are clearly improved at 77 K compared with both 300 K and 358 K.

the program disturb characteristics obtained at 300 K and 77 K. The program voltage is applied to selected WL and constant voltage (V_{PASS}) applied to unselected WL. While cell Vth of the programmed cell is increased by program voltage bias, cell Vth of inhibit cell shows no change at both 300 K and 77 K. There is no significant difference in the disturb characteristics between the room temperature and the cryogenic temperature. This behavior is also important to achieve multilevel cells, especially 6-bit per cell (HLC) and beyond. Because, in order to increase the number of bits per cell, multiple times program pulses are necessary to apply to the cell, it is necessary to have enough disturb characteristics to minimize Vth distribution of each program level.

For the read disturb characteristics, the raw bit error rate (RBER) degradation at low temperatures has been reported recently [26]. In this study, we evaluate the cell characteristics mostly in the main distribution, and there seems to be no clear degradation in both program and read disturb. We think that it is important to evaluate the detailed characteristics of the tail bit, which causes an increase in RBER, and the difference between low temperature around −40 °C and extremely low temperature of 77 K as further investigation.

C. DATA RETENTION

The data retention characteristics are greatly improved at cryogenic temperatures and one of the advantages for the 3-D flash memory [21], [22]. Fig. 11 shows the data retention characteristics at 300 K and 77 K. The delta cell Vth shows cell Vth reduction after programming multilevel cells. Assuming the worst case of the data retention, the target cells are set to the highest Vth-level, and the neighboring cells are set to the erase level [27]. It can be seen that the reduction of cell Vth is significantly minimized at 77 K.

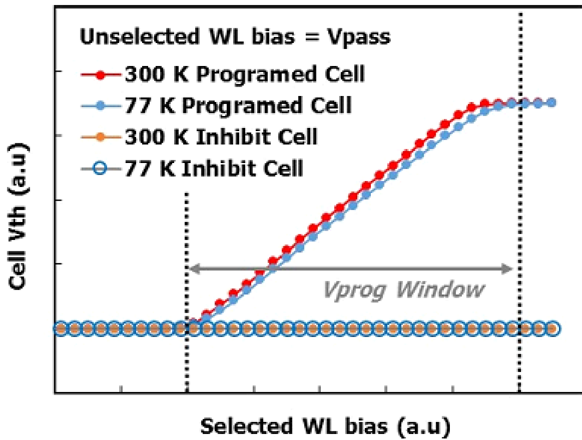

FIGURE 10. Program disturb characteristics when the program voltage bias is applied to the programmed cell and inhibit cell at selected WL. The same V_{PASS} bias is induced at unselected WLs. There is no degradation in the V_{PROG} window at both 300 K and 77 K.

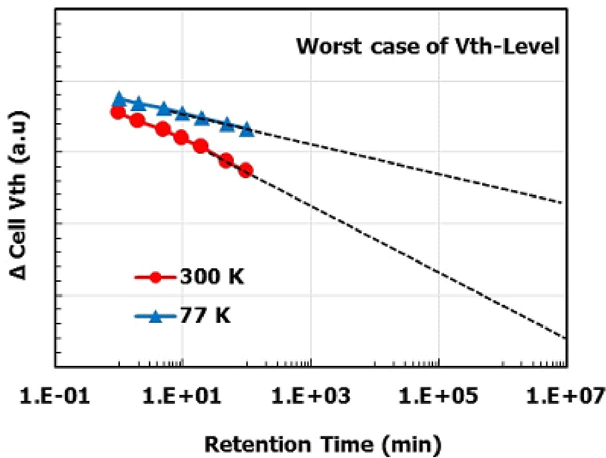

FIGURE 11. Data retention characteristics at 300 K and 77 K. The shift of cell Vth due to charge loss is suppressed at 77 K, and the data retention is significantly improved.

It is explained that the charge loss, the main factor to decide the data retention characteristics, is caused by the thermal excitation of the trapped carriers, and it is efficiently suppressed at the cryogenic temperature. In more detail, the mechanism of the charge loss at the target cell depends on the Vth-level of neighboring cells and has lateral and vertical directions [28], [29]. While both vertical and lateral charge losses are improved at cryogenic temperature, the improvement in the lateral direction is larger than in the vertical direction [22]. At high temperatures, the charge transfer due to the P-F effect becomes more active, and the charge loss in the lateral direction increases. At low temperatures, the thermal excitation of the P-F effect is suppressed, and the data retention characteristics are dominated by the charge transfer in the vertical direction. In addition, the suppression of lateral charge loss enables scaling of the pitch of the WL stack, which is another advantage of the cryogenic operation

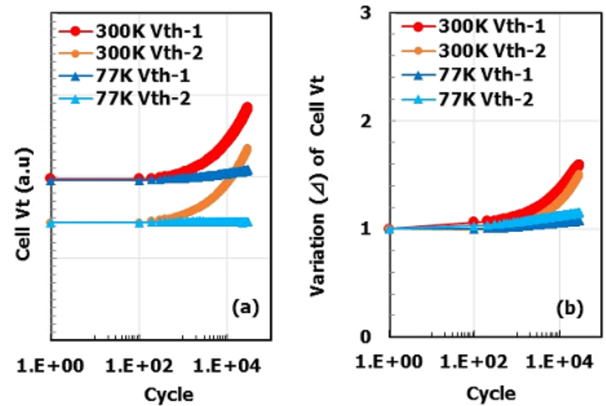

FIGURE 12. Cycle endurance characteristics of the program/erase cycle at 300 K and 77 K. No clear degradation at 30k cycles is observed at 77 K.

from the bit scaling point of view for future 3-D flash memory.

D. CYCLE ENDURANCE

We have discovered that the cycle endurance is also greatly improved by the cryogenic operation [21]. Fig. 12 shows the endurance characteristics due to P/E cycles at 300 K and 77 K. The applying P/E pulse is one cycle with fixed voltage and time, and two states of cell Vth are plotted. At 300 K, the cell Vth and its variation rise rapidly at over P/E 1k cycles. On the other hand, at 77 K, there is only a slight increase in cell Vth and variation even at P/E 30k cycles. The cycle endurance is improved over ten times higher at 77 K than 300 K.

We can explain a possible mechanism of these cycle endurance improvements at the cryogenic temperature as follows. During the erase operation, the holes are injected from the channel region into the CT layer through the tunnel oxide. The energy of hot holes at the channel decreases with decreasing temperature because of no electric field at the channel region. The low energy of holes prevents the cause of defects in the tunnel oxide [30]–[32], and holes have a larger impact than electrons [33], [34]. The origin of the defects is known to Si–H bonds in the tunnel oxide [32]. Furthermore, low temperature suppresses the hydrogen migration in the tunnel oxide. Consequently, the defects in the tunnel oxide that traps charges can be less generated at the cryogenic temperature.

E. MULTILEVEL CELL AND 6-BIT PER CELL

In order to increase the bit density and reduce bit cost, the multilevel cell is one of the critical ways for the 3-D flash memory. The multilevel cell requires tight cell Vth distribution and enough cell Vth window. The cell Vth window is mainly determined by three factors: the read noise, the data retention, and the cycle endurance. As described above, all these characteristics are sufficiently improved at the cryogenic operation. Therefore, we can obtain excellent

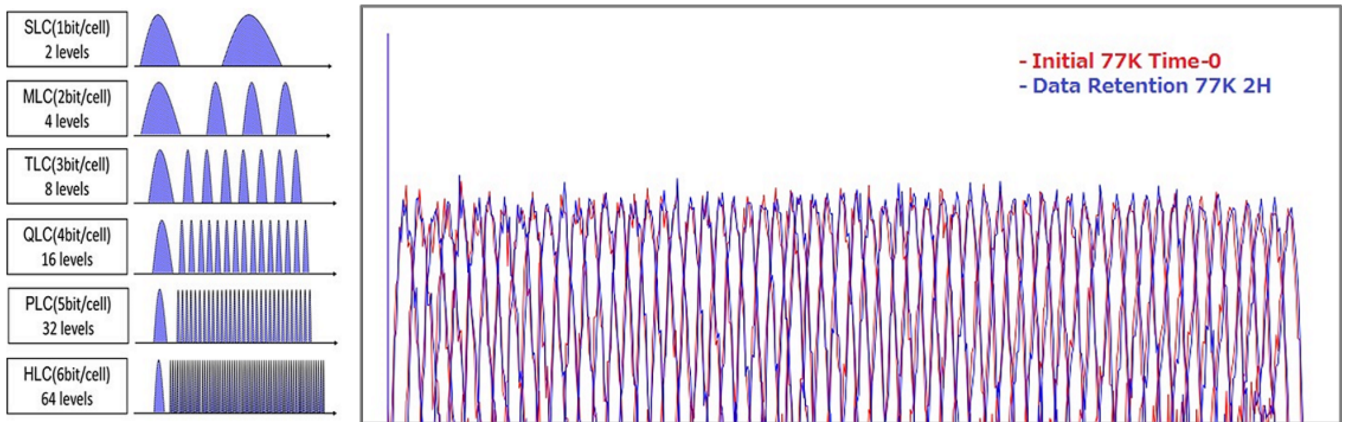

FIGURE 13. (a) Illustration of Vth distribution from 1-bit per cell (SLC), 2-bit per cell (MLC), 3-bit per cell (TLC), 4-bit per cell (QLC), 5-bit per cell (PLC), and HLC. (b) Vth distribution of successful demonstration of HLC at initial and after data retention obtained at the cryogenic temperature of 77 K.

TABLE 1. Assumption of storage device and condition of its operation to estimate the cooling power and cost.

	Parameter	Value
Storage Device	Type	SSD
	Bit per Die	4 Tbit/Die
	Package Structure	8 Dies stack
	Multi-Level-Cell	6-bit (HLC)
	Total Bit Capacity	4 TByte
Operation	Tera Byte Written	800 TBW
	Read : Write Ratio	50 : 1
	Average Duty	30%
	Duration	3 Years
Cooling	Temperature	77 K
	Cooling Power Efficiency (Output/Input)	10%

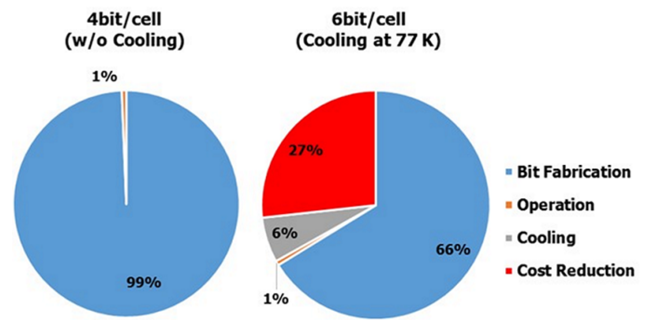

FIGURE 14. Estimated cost for bit fabrication, operating power consumption, and cooling power. The total cost of HLC with cooling at 77 K is sufficiently reduced compared with 4-bit per cell without cooling around 300 K.

performance of the multilevel cell. Moreover, increasing the number of bits per cell is achievable. Fig. 13 shows the successful demonstration of HLC by using our state-of-the-art 3-D flash memory. The tight cell Vth distribution and enough cell Vth window are obtained at initial and after data retention of 120 min at 77 K. There is no significant difference between initial and after data retention due to the improvement of the cryogenic operation.

Beyond HLC, to achieve 7-bit per cell (7LC) or 8-bit per cell (8LC or OLC), cell Vth distribution needs to be much smaller. The read noise may be lowered by the single-crystal channels [35]–[37] or new channel materials. The data retention characteristics are known to be better for the floating gate (FG) type than the CT type [38], and there is still room to optimize the tunneling oxide layer for extremely low temperatures. Although cycle degradation is sufficiently small at cryogenic temperatures, a system-level approach to reduce the write amplification factor (WAF) is also effective [39], and it can relax the requirement of cycle times instead of improving the cycle endurance. The cryogenic operation is fully compatible with these emerging technologies that are expected to be applied in the future. We believe that the cryogenic operation is essential to realize ultramultilevel cells beyond HLC.

Since the current bit density of the 3-D flash memory strongly depends on the number of highly stacked WL layers, there is a huge concern about increasing chip fabrication cost. The ultramultilevel cell by the cryogenic operation is possible to increase bit density and reduce bit cost without increasing chip fabrication cost. Then, boosting the number of the multilevel cell may become a new scaling factor for future 3-D flash memory.

F. COOLING COST ESTIMATION AND BIT COST SCALING

The cooling cost is a unique and additional challenge for cryogenic operation because it requires cooling power and cost. The bit cost scaling of boosting the ultramultilevel cell by the cryogenic operation is effective if the cooling cost is small enough compared with chip fabrication cost. We estimate the cooling cost by considering the power consumption for cooling the storage devices. We assume the workloads of a typical storage device and estimate the average power consumption. These assumptions are shown in Table 1. We assume that SSDs are cooled at the constant cryogenic temperature all the time within three years of operation. Fig. 14 shows the bit fabrication cost, operating

TABLE 2. Summary of application fields and advantages of cryogenic operation of the 3-D flash memory.

Application	Advantage of Cryogenic Operation
Quantum Computing	· Large Storage Capacity near the Qubit · Low Thermal Leakage by I/O · Simple Structure of Memory&Storage System
Aerospace Electronics	· Wide Range of Operating Temperature · High Cycle Endurance and Long-life Usage · Not Specialized, COTS Component
High Performance Storage	· Small Read (Verify) Noise and Short Latency · Long Data Retention · High Cycle Endurance
Bit Density & Bit Cost Scaling	· Enabling Ultra-Multi Level Cell · Cooling Cost is Smaller than Bit Fabrication Cost · Countermeasure of Increasing WL layers · Continuing of Bit Density and Bit Cost Scaling

cost, and cooling cost for 4-bit per cell without cooling as reference and HLC with cooling at 77 K. Compared with the bit fabrication cost, the operating cost is negligible, and the cooling cost is less than 10%. Compared with the cryogenic operation of other semiconductor devices, such as CPU and DRAM, flash memory has a relatively long standby time and small duty from the workload point of view. Therefore, the storage device, such as flash memory, is suitable for cryogenic operation because of the relatively small cooling cost.

V. CONCLUSION

We introduced the cell characteristics and the storage performance of 3-D flash memory at the cryogenic operation. The cell transistor and the SGS transistor show better subthreshold characteristics at 77 K than at 300 K. The basic functionality of program/erase and read operations is observed. ISPP and ISPE methods are completely applicable. Moreover, the cryogenic operation can bring us small read noise, long data retention, and high cycle endurance characteristics as the storage performance improvement.

Potential new applications or usage to apply cryogenic operation of 3-D flash memory are summarized in Table 2. For quantum computing, we can put a large storage capacity near the qubit, and it may improve the functionality and controllability of the quantum computer. A simple structure of memory and storage system is constructed with low thermal leakage by I/O connection. For aerospace electronics, a wide range of operating temperatures may have huge advantages because there are many mission-critical tasks in this field. High cycle endurance brings long-life usage of storage devices, these components are not specialized, and the COTS component is available. In addition to quantum computing and aerospace electronics, we can simply apply the cryogenic operation to boost storage performance. It has excellent storage performance improvements of small read (verify) noise, long data retention, and high cycle endurance. For future storage devices, the bit cost scaling can be achieved without huge manufacturing costs for increasing more WL layers. The cryogenic operation enables us to achieve ultramultilevel cells of HLC and beyond. We can continue to increase the bit density and reduce the bit cost. The cryogenic operation of 3-D flash memory enables new applications and can greatly improve future computing and storage devices.

REFERENCES

[1] F. Masuoka, M. Momodom, Y. Iwata, and R. Shirota, “New ultra high density EPROM and flash EEPROM with NAND structure cell,” in *IEDM Tech. Dig.*, Dec. 1987, pp. 552–555.

[2] H. Tanaka *et al.*, “Bit cost scalable technology with punch and plug process for ultra high density flash memory,” in *Proc. IEEE Symp. VLSI Technol.*, Jun. 2007, pp. 14–15.

[3] H. Aochi, “BiCS flash as a future 3D non-volatile memory technology for ultra high density storage devices,” in *Proc. IEEE Int. Memory Workshop*, May 2009, pp. 1–2.

[4] W. Chakraborty, K. Ni, J. Smith, A. Raychowdhury, and S. Datta, “An empirically validated virtual source FET model for deeply scaled cool CMOS,” in *IEDM Tech. Dig.*, Dec. 2019, pp. 39.4.1–39.4.4.

[5] M. Shin *et al.*, “Low temperature characterization of 14 nm FDSOI CMOS devices,” in *Proc. 11th Int. Workshop Low Temp. Electron. (WOLTE)*, Jul. 2014, pp. 29–32.

[6] S. S. Tannu, D. M. Carmean, and M. K. Qureshi, “Cryogenic-DRAM based memory system for scalable quantum computers,” in *Proc. Int. Symp. Memory Syst.*, Oct. 2017, pp. 189–195.

[7] F. Ware *et al.*, “Do superconducting processors really need cryogenic memories?” in *Proc. Int. Symp. Memory Syst.*, Oct. 2017, pp. 183–188.

[8] F. Wang, T. Vogelsang, B. Haukness, and S. C. Magee, “DRAM retention at cryogenic temperatures,” in *Proc. IEEE Int. Memory Workshop (IMW)*, May 2018, pp. 1–4.

[9] M. Song, J. Y. Yang-Scharlotta, M. Ashtijou, and M. Mojarradi, “Evaluation of commercial-off-the-shelf (COTS) electronics for extreme cold environments,” in *Proc. IEEE Aerosp. Conf.*, Mar. 2018, pp. 1–12.

[10] F. R. Ihmig, S. G. Shirley, and H. Zimmermann, “Batch screening of commercial serial flash-memory integrated circuits for low-temperature applications,” *Cryogenics*, vol. 71, pp. 39–46, Oct. 2015.

[11] L. Fan *et al.*, “Cryogenic characterisation of 55 nm SONOS charge-trapping memory in AC and DC modes,” *Electron. Lett.*, vol. 56, no. 4, pp. 199–201, Feb. 2020.

[12] D. M. Carlson *et al.*, “The ETA 10 liquid-nitrogen-cooled supercomputer system,” *IEEE Trans. Electron Devices*, vol. 36, no. 8, pp. 1404–1413, Aug. 1989.

[13] H. L. Chiang *et al.*, “Cold CMOS as a power-performance-reliability booster for advanced FinFETs,” in *Proc. IEEE Symp. VLSI Technol.*, Jun. 2020, pp. 1–2.

[14] G.-H. Lee, D. Min, I. Byun, and J. Kim, “Cryogenic computer architecture modeling with memory-side case studies,” in *Proc. 46th Int. Symp. Comput. Archit.*, Jun. 2019, pp. 774–787.

[15] G.-H. Lee, S. Na, I. Byun, D. Min, and J. Kim, “CryoGuard: A near refresh-free robust DRAM design for cryogenic computing,” in *Proc. ACM/IEEE 48th Annu. Int. Symp. Comput. Archit. (ISCA)*, Jun. 2021, pp. 637–650.

[16] J.-H. Bae *et al.*, “Characterization of a capacitorless DRAM cell for cryogenic memory applications,” *IEEE Electron Device Lett.*, vol. 40, no. 10, pp. 1614–1617, Oct. 2019.

[17] E. Garzon, R. De Rose, F. Crupi, A. Teman, and M. Lanuzza, “Exploiting STT-MRAMs for cryogenic non-volatile cache applications,” *IEEE Trans. Nanotechnol.*, vol. 20, pp. 123–128, 2021.

[18] E. Garzon, Y. Greenblatt, O. Harel, M. Lanuzza, and A. Teman, “Gain-cell embedded DRAM under cryogenic operation—A first study,” *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 29, no. 7, pp. 1319–1324, Jul. 2021.

[19] R. Saligram, S. Datta, and A. Raychowdhury, “CryoMem: A 4K-300K 1.3 GHz eDRAM macro with hybrid 2T-gain-cell in a 28 nm logic process for cryogenic applications,” in *Proc. IEEE Custom Integr. Circuits Conf. (CICC)*, Apr. 2021, pp. 1–2.

[20] S. Resch, H. Cilasun, and U. R. Karpuzcu, “Cryogenic PIM: Challenges & opportunities,” *IEEE Comput. Archit. Lett.*, vol. 20, no. 1, pp. 74–77, Jan. 2021.

[21] Y. Aiba *et al.*, “Cryogenic operation of 3D flash memory for new applications and bit cost scaling with 6-bit per cell (HLC) and beyond,” in *Proc. 5th IEEE Electron Devices Technol. Manuf. Conf. (EDTM)*, Apr. 2021, pp. 1–3.

[22] Y. Aiba *et al.*, "Bringing in cryogenics to storage: Characteristics and performance improvement of 3D flash memory," in *Proc. IEEE Int. Memory Workshop (IMW)*, May 2021, pp. 1–4.

[23] H. Chen *et al.*, "Performance of GAA poly-Si nanosheet (2 nm) channel of junctionless transistors with ideal subthreshold slope," in *Proc. Symp. VLSI Technol.*, Jun. 2013, pp. T232–T233.

[24] D. Resnati, A. Goda, G. Nicosia, C. Miccoli, A. S. Spinelli, and C. M. Compagnoni, "Temperature effects in NAND flash memories: A comparison between 2-D and 3-D arrays," *IEEE Electron Device Lett.*, vol. 38, no. 4, pp. 461–464, Apr. 2017.

[25] A. Goda, "3-D NAND technology achievements and future scaling perspectives," *IEEE Trans. Electron Devices*, vol. 67, no. 4, pp. 1373–1381, Apr. 2020.

[26] F. Chen *et al.*, "Temperature impacts on endurance and read disturbs in charge-trap 3D NAND flash memories," *Micromachines*, vol. 12, no. 10, p. 1152, Sep. 2021.

[27] Y. Kong *et al.*, "Comprehensive investigations on data pattern dependences in charge-trap (CT) 3D NAND flash memory," in *Proc. IEEE 15th Int. Conf. Solid-State Integr. Circuit Technol. (ICSICT)*, Nov. 2020, pp. 1–3.

[28] C. Woo *et al.*, "Modeling of charge loss mechanisms during the short term retention operation in 3-D NAND flash memories," in *Proc. Symp. VLSI Technol.*, Jun. 2019, pp. T214–T215.

[29] B. Choi *et al.*, "Comprehensive evaluation of early retention (fast charge loss within a few seconds) characteristics in tube-type 3-D NAND flash memory," in *Proc. IEEE Symp. VLSI Technol.*, Jun. 2016, pp. 1–2.

[30] D. J. DiMaria and J. H. Stathis, "Non-Arrhenius temperature dependence of reliability in ultrathin silicon dioxide films," *Appl. Phys. Lett.*, vol. 74, no. 12, pp. 1752–1754, Mar. 1999.

[31] E. Y. Wu *et al.*, "The effect of change of voltage acceleration on temperature activation of oxide breakdown for ultrathin oxides," *IEEE Electron Device Lett.*, vol. 22, no. 12, pp. 603–605, Dec. 2001.

[32] Y. Higashi *et al.*, "Mechanism of gate dielectric degradation by hydrogen migration from the cathode interface," *Microelectron. Rel.*, vol. 70, pp. 12–21, Mar. 2017.

[33] S. Lombardo, J. H. Stathis, B. P. Linder, K. L. Pey, F. Palumbo, and C. H. Tung, "Dielectric breakdown mechanisms in gate oxides," *J. Appl. Phys.*, vol. 98, no. 12, Dec. 2005, Art. no. 121301.

[34] A. Subirats, A. Arreghini, R. Degraeve, D. Linten, G. Van Den Bosch, and A. Furnemont, "In depth analysis of post-program VT instability after electrical stress in 3D SONOS memories," in *Proc. IEEE 8th Int. Memory Workshop (IMW)*, May 2016, pp. 1–4.

[35] A. Subirats *et al.*, "Trap reduction and performances improvements study after high pressure anneal process on single crystal channel 3D NAND devices," in *IEDM Tech. Dig.*, Dec. 2018, pp. 25.4.1–25.4.4.

[36] G. Congedo *et al.*, "Analysis of performance/variability trade-off in macaroni-type 3-D NAND memory," in *Proc. IEEE 6th Int. Memory Workshop (IMW)*, May 2014, pp. 1–4.

[37] H. Miyagawa *et al.*, "Metal-assisted solid-phase crystallization process for vertical monocrystalline Si channel in 3D flash memory," in *IEDM Tech. Dig.*, Dec. 2019, pp. 28.3.1–28.3.4.

[38] P. Kalavade, "4 bits/cell 96 layer floating gate 3D NAND with CMOS under array technology and SSDs," in *Proc. IEEE Int. Memory Workshop (IMW)*, May 2020, pp. 1–4.

[39] K. Takeuchi, "Green high performance storage class memory & NAND flash memory hybrid SSD system," in *Proc. IEEE/ACM Int. Symp. Low Power Electron. Design*, Aug. 2011, pp. 369–370.