

ANALYSIS OF SRAM RELIABILITY UNDER
COMBINED EFFECT OF TRANSISTOR AGING,
PROCESS AND TEMPERATURE VARIATIONS IN
NANO-SCALE CMOS

A thesis work submitted to the faculty of

San Francisco State University

In partial fulfillment of the requirements for

The degree

Master in Sciences

In

Engineering

By

Harwinder Singh

San Francisco State University, California

December, 2009

Copyright by

Harwinder Singh

2009

Certification of Approval

I certify that I have read Analysis of SRAM Reliability under Combined Effect of Transistor Aging, Process and Temperature Variations in Nano-Scale CMOS by Harwinder Singh, and that in my opinion this work meets the criteria for approving a thesis submitted in partial fulfillment of the requirement for the degree: Master of Sciences in Engineering at San Francisco State University.

Hamid Mahmoodi

Assistant Professor, Electrical and Computer Engineering

Hao Jiang

Professor, Electrical and Computer Engineering

ANALYSIS OF SRAM RELIABILITY UNDER COMBINED EFFECT OF TRANSISTOR AGING, PROCESS AND TEMPERATURE VARIATIONS IN NANO-SCALE CMOS

By

Harwinder Singh

San Francisco State University, California

As dimensions of MOS devices have been scaled down, new reliability problems are coming into effect. One of these emerging reliability issues is aging effects which result in device performance degradation over time. NBTI (Negative biased temperature instability) and PBTI (Positive biased temperature instability) are well known aging phenomenon which are limiting factors for future scaling of devices. NBTI results in V_t (threshold voltage) degradation of PMOS and PBTI results V_t degradation in the NMOS due to the generation of trapped charges. In nano-scale CMOS technologies, process (threshold voltage) and temperature variations are also crucial reliability concerns. On the other hand, NBTI and PBTI are also dependent on temperature and threshold voltage. In this project, the combined effect of transistor aging, process and temperature variations on the reliability of the 6T SRAM (Static Random Access Memory) in 32nm CMOS technology is analyzed in terms its performance metrics: SNM (Static Noise Margin), Write Margin, Access time and leakage. We analyzed that performance of SRAM degrades over time under aging effect. It is observed that low V_t SRAM at high temperature suffers most performance degradation. Due to the degradation of SRAM performance, number of faulty cells rises over time. We also observed the effect of supply voltage variation of the SRAM. We observed that

performance degradation is minimal at nominal supply voltage. Our results show that adaptive supply voltage can be used to save power over time in active mode.

I certify that the Abstract is a correct representation of the content of this thesis.

Chair, Thesis Committee

Date

ACKNOWLEDGEMENTS

I would like to thank my thesis advisor Dr. Hamid Mahmoodi for his invaluable guidance and support towards the completion of my thesis. I would also like to thank my committee member Dr. Hao Jiang for his timely support and feedbacks. I am also thankful to the school of engineering for providing me the resources required for completing my project. Finally, I would like to thank the whole team of NeCRC (Nano-electronics Computing and Research Center), my family and friends for their support.

Table of Contents

Abstract

Acknowledgments

1. Introduction
2. SRAM cell and operation
3. SRAM cell failures
4. Performance metrics of SRAM cell
 - a) SNM
 - b) Write Margin
 - c) Access Time
 - d) Leakage
5. Model of NBTI under Temperature and Process variations
6. Impact of NBTI on SRAM with Conventional SiO₂ technology
7. Combined effect of NBTI, Process and Temperature Variation on SRAM with Conventional SiO₂ technology
8. Aging effect in High K CMOS technology
9. Combined effect of Aging Effect and PVT (Process, Voltage and Temperature) Variations on 6T SRAM in High K CMOS technology
10. Design solution to mitigate transistor aging effect
11. Conclusion
12. Reference

List of Figures

Fig. 1: V_t shift for the three technology corners: Low V_t transistors, Nominal V_t transistors, High V_t transistors at room temperature (25°C) and worst case temperature (100°C).

Fig. 2: Schematic diagram of SRAM where pull up PMOS PL is stressed.

Fig. 3: SNM degradation after 10 years. SNM degradation is considerably higher in low V_t cell and at high temperature.

Fig. 4: Statistical distribution of SNM due to intra-die V_t variations for the low inter-die V_t chip at time 0 and after 2 years. Number of faulty cells increase over time as SNM distribution shifts to left.

Fig. 5: Percentage rise of faulty cells due to SNM degradation increases over time at 25°C and 100°C . It is observed that number of faulty increase (8.2% over 2years) only at the high temperature.

Fig. 6: Write margin slightly degrades over time. However, this degradation of write margin is practically negligible.

Fig. 7: SRAM cell leakage current for low V_t inter-die chip over time. NBTI does not impact leakage current.

Fig. 8: SRAM cell read current for low V_t inter-die chip over time. NBTI does not impact read current.

Fig 9: Percentage V_t shift in PMOS due to NBTI for the three technology corners: Low V_t transistors, Nominal V_t transistors, High V_t transistors at room temperature (25°C) and worst case temperature (100°C) for high K dielectrics.

Fig 10: Percentage V_t shift for NMOS and PMOS due to PBTI and NBTI respectively with two supply voltage level 0.5V and 1V for high K dielectrics.

Fig. 11: SNM degradation over 2 years. Initially SNM degradation is considerably higher in low voltage SRAM cell and at high temperature. Over long period of time, high voltage SRAM cell shows more SNM degradation and at high temperature.

Fig 12: SNM degradation over time for the three technology corner SRAM 6T cell: low, nominal, high V_t . Low V_t cell shows more SNM degradation as compared to high V_t .

Fig 13: Write margin degrades over time and degradation is considerably higher in high voltage SRAM cell and at high temperature.

Fig 14: SRAM cell access time degradation for low V_t inter-die chip over time. Access time degradation is considerably higher in low voltage SRAM cell and at high temperature.

Fig 15: SNM degradation for read and hold mode at different supply voltages. It can be observed that SNM degradation is minimal both cases at 0.8V supply voltage.

Fig. 16: Access time Degradation at different supply voltage. It can be observed that access time degradation is minimal at 1V.

Fig 17: Read SNM at time 0 and after 2 years at different voltage levels.

Fig. 18 Hold SNM at time 0 and after 2 years at different voltage levels.

Fig 19: Access time at time 0 and after 2 years at different voltage levels.

List of Tables

- TABLE I. PERCENTAGE SNM DEGRADATION FOR LOW VT, NOMINAL VT AND HIGH VT TECHNOLOGY CORNERS AT 25°C AND 100°C OVER 2 YEARS
- TABLE II. PERCENTAGE OF FAULTY CELLS DUE TO SNM DEGRADATION AGING AND PROCESS VARIATIONS AT 100°C
- TABLE III. PERCENTAGE WRITE MARGIN DEGRADATION FOR LOW VT, NOMINAL VT AND HIGH VT TECHNOLOGY CORNERS AT 25°C OVER 2 YEARS
- TABLE IV. PERCENTAGE OF FAULTY CELLS DUE TO WRITE MARGIN DEGRADATION UNDER AGING AND PROCESS VARIATIONS AT 100°C (FOR TARGET WRITE MARGIN OF 0.3V)

1.INTRODUCTION

At the nano scale technology, reliability becomes a serious concern. Even though the chips achieve great performance at the nano scale technology, their operation becomes more unreliable. The whole reliability concept covers many chip reliability problems like leakage, time dependent degradation, manufacturing imperfection, temperature variations. The researchers in industry and in academic institutes are continuously working to provide effective solutions for these reliability problems. Some solutions provided great relief to VLSI technology from these reliability issues. For example, replacement of SiO₂ by high-K dielectrics provided 25%-100% reduction in the gate leakage [9]. However, many reliability issues especially time dependent degradation are limiting factors for the future of technology scaling. NBTI (Negative Bias Temperature Instability) and PBTI (Positive Bias Temperature Instability) are well known aging effects which cause the threshold voltage degradation of PMOS and NMOS transistors over time. NBTI results from interface trapped charges from the broken Si-H bonds at the interface [4]. PBTI results from oxide trapped charges [13]. NBTI and PBTI resultant degradation not only depends on supply voltage and temperature but also threshold voltage and other technology parameters of the MOS transistor which results in more threshold voltage degradation with further scaling. MOS becomes a slower switch with threshold voltage degradation which leads to undesirable operation of circuits consisting MOS transistors. As a result, some high performance application might fail over time. In addition to transistor aging effect, process variations (threshold voltage) and temperature variations also play a crucial role in the circuit operation. Process variation mainly results from the imperfection of manufacturing process which results in variation in threshold voltage from chip to chip (inter-die variation) and transistor to transistor (intra-die variation). These variations are caused by many reasons like variations in channel length, channel width, channel doping, oxide thickness, line-edge roughness, and random dopant fluctuations. Thus transistors or chips' performances might be different from each other due to process variations which results in an undesirable output. To account for inter die V_t variations, we considered three different chips: low V_t chip, nominal V_t chip, high V_t chip. Depending on

the work load and operating condition, temperature might be different on different point within single chip or chip to chip. To account for temperature change, we considered two temperature levels: room temperature (25°C) and worst case temperature (100°C).

The effect of NBTI aging on SRAM reliability with SiO₂ dielectrics has been studied [1, 6, 7, 8]. In [1, 6, 7, 8] the behavior of NBTI and its impact on the 6T SRAM cell has been studied in different technologies (70nm [1], 32nm [6], 35nm [7], 70nm and 90nm [8]). The effect of aging (NBTI and PBTI) is also studied by many researches [14, 15, 16]. However, the results from these papers show the impact of aging (NBTI and PBTI) alone on the SRAM without considering the variations in threshold and temperature. Hence the combined effect of aging with temperature and process variations has not been researched. Given the strong dependence of BTI aging effect (NBTI and PBTI) on temperature and threshold voltage, the impact of BTI can be greatly different under variations in threshold voltage and temperature. In this work, we analyze and report the combined effect of BTI, threshold voltage and temperature variations on the SRAM reliability in 32nm CMOS technology with SiO₂ dielectrics and with high K dielectrics. It is observed that: (1) V_t abruptly increases initially and afterwards V_t shift is very small, even for prolonged time; (2) Low V_t transistors age faster than high V_t transistors; and (3) Both NBTI and PBTI resultant V_t degradation are more significant at higher temperature. In case of SiO₂ CMOS technology, (4) for the worse case, SNM and write margin degrades over time whereas leakage and access time do not get affected from NBTI. Along with these observations, we also quantified our results in terms of number of faulty cells in SRAM array. It is observed that: (5) number of faulty cells rises over time (8.2% rise in faulty cells for the inter-die nominal V_t chip over 2 years) due to SNM degradation; (6) rise in the number of faulty cells over time due to write failures under NBTI effect is practically negligible. In case of high K CMOS technology, (7) for the worse case, SNM, write margin and access time degrades over time whereas leakage does not get effect from the NBTI or PBTI. We also observed the effect of supply voltage variation of the SRAM. By using two supply voltage levels, we observed that (8) Access time degradation is more in case of low supply voltage chip whereas in case of

write operation, degradation is more with high supply voltage; (9) In case of SNM, initially low voltage SRAM cell shows more degradation as compared to high voltage SRAM cell. Over long time, high voltage SRAM cell shows more SNM degradation. We analyzed that (10) even though the sensitivity of performance metrics (SNM and access time) towards supply voltage is different, the low and high voltage circuits suffer more degradation as compared to nominal voltage circuits. We observed that (11) in case of read SNM, adaptive supply voltage technique can be used to save power over time to maintain target read SNM where power overhead is 22.22%. However, (12) in case of hold SNM and access time is power overhead is practically negligible, so adaptive supply voltage technique is not viable. Since read SNM and access time are coupled together, in active mode, power can be saved by using the adaptive supply voltage scheme over time.

The remainder of this thesis is organized as follows. Section 2 includes the introduction to 6T SRAM cell and its operations. Section 3 briefly describes failures in the SRAM cell. Section 4 discusses the performance metrics of SRAM cell. Section 5 describes the modeling of NBTI under the process and temperature variations. In Section 6, explains the effect of NBTI on the SRAM with conventional SiO₂ dielectrics. Section 7 shows the combined effect of NBTI, process variation and temperature variation on the performance metrics of the SRAM. Section 8 discusses the aging effects in high K CMOS technology. Section 9 includes the analyses of performance metrics of SRAM with High K dielectrics under the effect of transistor aging effects (NBTI and PBTI) and PVT (Process, voltage and temperature) variations. Design solution to mitigate transistor aging is given in section 10. Finally, Section 11 concludes the thesis.

2.SRAM CELL AND ITS OPERTATION

SRAM (Static Random Access Memory) is a type of semiconductor memory, widely used in various applications in the industry as well as personal electronic applications to store data. Since it is static memory, it does not require to be refreshed periodically. SRAM is available in many forms depending on the number of transistors (T) used per cell (4T, 6T, 8T, 10T etc.). In this project, conventional 6-

Transistor (6T) SRAM cell is used.

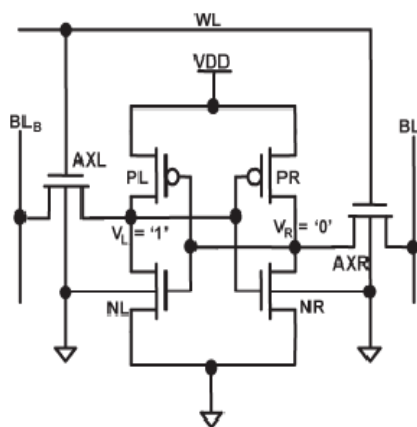


Fig. 2: Schematic diagram of SRAM where pull up PMOS PL is stressed.

Fig. 2 shows the schematic diagram of the 6T SRAM cell. It has 2 pull up PMOS and 2 NMOS pull down transistors as two cross coupled inverters and two 2 NMOS access transistors to access the SRAM cell during Read and Write operations. Both the bitlines (BL and BL_B) are used to transfer the data during the read and write operations in a differential manner. To have better noise margin, the data signal and its inverse is provided to BL and BL_B respectively. The data is stored as two stable states, at storing points V_R and V_L, and denoted as 0 and 1. SRAM mainly works in three modes:

- 1) Read operation – During the read operation, the data stored in SRAM cell is being read. In this operation, both bitlines are precharged to logical 1 and word line is asserted which enables the access transistors. If cell stores the logical 0 (Fig. 2), the BL will get discharged through access

transistor AXR and pull down transistor NR to logical 0 and BL_B will stay precharged, so values of storing nodes get transferred to the bitlines. In the opposite case when cell stores the logical 1, then BL_B get discharged and BL remains precharged. For proper read operation, cell drive transistors should be stronger than the access transistors, otherwise data may be flipped during the read operation.

- 2) Write operation – During the write operation, data is being written into the SRAM cell. Data (0 or 1) which we want to write is applied to the BL and inverted value of data is applied to the BL_B . In case of writing 0, 0 would be applied to the BL and 1 would be applied to the BL_B and then word line WL would be asserted. If cell already contains the data 0 (at storing node V_R) then no charging/discharging operation will take place. If cell stores 1, then 1 will be discharged to 0 through access transistor AXR and 0 (at storing node V_L) will be pulled up to 1 by pull up transistor PL. For proper write operation, input driver access transistors should be designed to be more powerful than relative cell pull-up transistor so that they can easily overwrite the data stored in the cross coupled inverters.
- 3) Standby – In standby mode, cell remains in idle states. Word line is not asserted so the access transistors remain OFF. And cross coupled inverter continuously reinforce each other to hold the stored data.

3. SRAM CELL FAILURES

SRAM cell can have four types of failures during its operations:

- 1) Read Failure – Read failure is known as the flipping of data during read operation. During read operation, the voltage at node R (see fig. 2) rise to some positive voltage ($V_R > 0$) due to the resistor divider action between the access transistor AXR and pull down transistor NR. As a result of this voltage rise, pull down transistor NL turns ON weakly and sub threshold current through NL results in the discharging of node L. As voltage of node L discharges, the strength of NR decreases which results in increase in the V_R . Due to this cyclic action, V_R rise above the trip point of the left inverter (PL-NL), and it flips the state of cell[5].
- 2) Write Failure – The failure to write data into a SRAM cell is known as write failure. In case the cell is storing 0 at the storing node R, if we want to store 1(see fig. 2), 1 would be applied to the bit line BL and 0 would be applied to bit line BL_B . The node L would be discharged to the lower voltage ($V_L < 1$) by the resistor divider action of the pull up transistor PL and access transistor AXL. If V_L cannot be reduced below the trip point of right inverter (PR-NR) during the time V_{WL} is 1, the write failure occurs. Write failure is mainly caused by slow discharging of 1 due to weakening of the access transistor.
- 3) Access Time Failure – The access time of SRAM cell is defined as the time to produce a predefined voltage difference between bitlines (BL and BL_B), so that sense amplifier can read the state of the cell. If discharging of bitline (BL or BL_B) is slow due to the weakening of pull down and access transistor, the voltage difference between BL and BL_B will be less than the offset of the sense amplifier. Then access time failure occurs.
- 4) Hold Failure – Hold failure is known as the flipping of the cell state in standby mode. As in standby mode, the supply voltage of cell is reduced to save the leakage power consumption, the voltage of node storing 1 get reduced. The hold failure occurs when voltage of the node storing 1

further decreases and becomes lower than the trip point of the right inverter (PR-NR) by high leakage of its pull down NMOS transistor.

4.PERFORMANCE METRICS OF SRAM

4.1) Static Noise Margin (SNM) – SNM is the measure of stability of the SRAM cell to hold its data against noise. SNM of SRAM is defined as minimum amount of noise voltage present on the storing nodes of SRAM required to flip the state of cell [10]. SNM can be computed as the length of the side of a maximum square nested between the two voltage transfer characteristic (VTC) curves (i.e., for each back-to back inverters) of SRAM cell [6]. SNM can be categorized into two types: Hold SNM and Read SNM. Hold SNM is the SNM of the cell when wordline is LOW or disabled, meaning that the cell is in standby mode. However, the read SNM is more critical than the hold SNM because the SRAM is more vulnerable during read operation when wordline is active. When SNM of cell goes below the pre-defined target SNM of cell, the read failure may occur.

There is two methods to measure the SNM of SRAM cell. First method is a graphical approach in which SNM can be obtained by drawing and mirroring the inverter characteristics and then finding the maximum possible square between them.

The second approach is given in [10] and is quick and simple. This method was developed for the use together with a standard DC simulator. Fig. ?? shows the stylized version of graphical representation of SNM in the two coordinates systems, which are rotated at 45° relative to each other. In the (u, v) system, subtraction of the 'v' values of normal and mirrored inverter characteristics at given 'u' yields curve 'A', which is measurement of diagonal's length. The maximum and minimum curve of 'A' represent the required maximum squares. We implemented the above in SPICE to calculate SNM.

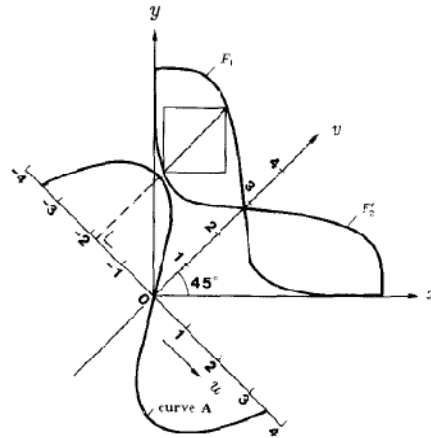


Fig 1 SNM estimation through simulation based on the maximum squares in a 45 degree rotated coordinate system [10]

4.2) Write Margin – Write margin is the measure of the ability to write data into the SRAM cell. In a write operation, two operations take place. Firstly discharging of node storing ‘1’ and pull up transistor PMOS and access transistor NMOS are responsible for the discharging of the node ‘1’. Secondly charging of node ‘0’ through pull up transistor PMOS. Write margin voltage is the maximum noise voltage present on bit lines during successful write operation. When noise voltage exceeds the write margin voltage, then write failure occurs.

4.3) Access Time – Access time is defined as the time required for accessing the data from the SRAM cell. It is determined by time period of Word line ‘WL’ pulse. This is an important metric because it determines the speed of memory circuits. We want to reduce the access time of SRAM cell to make it faster to work with high speed computing systems. The variations due to aging effect, process variations etc. change the behavior of cross coupled inverters in SRAM cell which results in access time failure.

4.4) Leakage – Leakage is defined as the power consumption of SRAM cell in standby mode. To reduce the leakage power, the supply voltage of cell is reduced. The voltage of node storing logical 1 also comes down with lowering of supply voltage. Due to leakage, if the voltage of node storing logical 1 goes below the trip of right inverter, then data flips in the standby mode.

5. MODEL OF NBTI UNDER TEMPERATURE AND PROCESS VARIATIONS

A comprehensive model for NBTI V_t shift is given in [2, 3]. In our research, we simplified models in [2, 3] for the DC stress condition and for $V_{ds}=0$ (as is the case for the PMOS of SRAM cell that is affected by NBTI (PL in Fig. 2)) and came up with the following model

$$\Delta V_t = (1 + m)K_v t^{0.25} + \delta_v$$

$$K_v = A \cdot t_{ox} \cdot \sqrt{C_{ox} (V_{gs} - V_t)} \cdot \exp\left(\frac{E_{ox}}{E_o}\right) \cdot \exp\left(-\frac{E_a}{KT}\right) \quad (1)$$

where t_{ox} is the oxide thickness,

E_{ox} is the electric field across the oxide ($(V_{gs}-V_t)/t_{ox}$).

A, E_a, E_o, δ_v, m and K are constants [2,3], and t is stress time in seconds.

This model shows the dependence of V_t shift on temperature (T) and process (threshold voltage) variation (V_t).

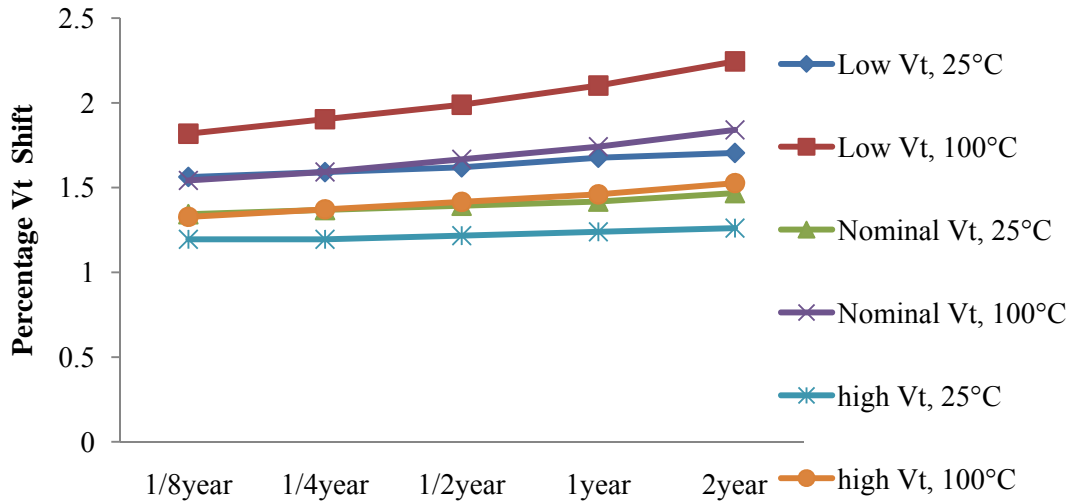


Fig. 1: V_t shift for the three technology corners: Low V_t transistors, Nominal V_t transistors, High V_t transistors at room temperature (25°C) and worst case temperature (100°C) for SiO_2 dielectrics.

Fig. 1 shows the percentage of V_t shift in three process corners: Low V_t transistors, Nominal V_t transistors and High V_t transistors at two temperatures: room temperature (25°C) and worst case (100°C) for SiO₂ dielectrics. These results show that V_t shift is greater at high temperature and the low technology corner. This is due to the dependence of K_v factor in Eq. 1 on the temperature and threshold voltage.

6.IMPACT OF NBTI ON SRAM WITH CONVENTIONAL SiO₂ CMOS TECHNOLOGY

In Fig. 2, the 'ON' PMOS PL is stressed and experiences the NBTI V_t shift. This V_t shift can cause the change in the performance metrics of the SRAM: SNM (Static Noise Margin), Write Margin, Access time and Leakage. Among these performance metrics, SNM (Static Noise Margin) degrades over time under stressed conditions. The reason is, under the stress conditions, the trip point of the left inverter becomes lower due to the increased V_t of PL, and hence the cell can be more easily flipped during the read operation.

Write margin improve over time for the first write cycle because slower PL will help the discharging of the '1' node. However, for the worse case which is second write cycle to write the initial state back to the cell, write margin degrades because the other pull up transistor PR has not been stressed (has low V_t) which is not helpful in the discharging of '1'. Moreover, the inverter PL-NL trip point has reduced which makes the second write more difficult. For the worse case when SRAM remains under continuous stress, Leakage is not impacted by NBTI because leakage is decided by the current flowing through the OFF transistor which is not affected by NBTI. Similarly in worse case, Access time is not impacted by NBTI because access time is determined by discharging of BL through NMOS access transistors and NMOS pull down transistors which are not impacted by NBTI.

7.COMBINED EFFECT OF NBTI, PROCESS AND TEMPERATURE VARIATION WITH CONVENTIONAL SiO₂ TECHNOLOGY

Although the individual effect of NBTI has been studied and verified by many researchers [1, 6, 7,8], in reality, not only aging effect degrades the reliability of SRAM, temperature and process variations also play crucial roles. Moreover, as observed in the previous section, NBTI is dependent on the temperature and V_t process parameter. In this work, we analyzed the combined effect of NBTI, the process and temperature variation on the reliability of SRAM cell, in terms of Static Noise Margin (SNM), write margin and leakage of 6T SRAM cell in the 32nm CMOS process. Three types of chips are considered to incorporate the effect of inter die variations: Low V_t , Nominal V_t and High V_t transistors. For intra die process variations, which are mostly caused by random dopant fluctuations, we applied random values of V_t to all 6 transistors of SRAM and for many SRAM cells in a chip.

SNM is affected by the static process variations and the environmental variations such as temperature as well as aging effects such as NBTI. In this work, the combined effect of NBTI, temperature and process variations on the SNM over the life-time of the SRAM is analyzed for the worse case where SRAM remains under continuous stress (PL is under continuous stress, see Fig. 2). Fig. 3 shows the SNM degradations of low V_t , nominal V_t and high V_t transistors for room temperature (25°C) and worst case temperature (100°C). The results show that SNM degradation is more significant at higher temperature for all process corners. At room temperature, SNM degradation is negligible at different process corners (less than 1%), however, at higher temperature, SNM degradation is more in case of low V_t cells as compared to high V_t cells. Hence, it can be concluded that SRAM chips that are shifted to low V_t process corners will experience faster SNM degradation as a result of NBTI compared to the chips that are shifted to high V_t corners. Table 1 shows the results for SNM degradation for the low V_t , nominal V_t and high V_t technology corners at 25°C and 100°C over 2 years.

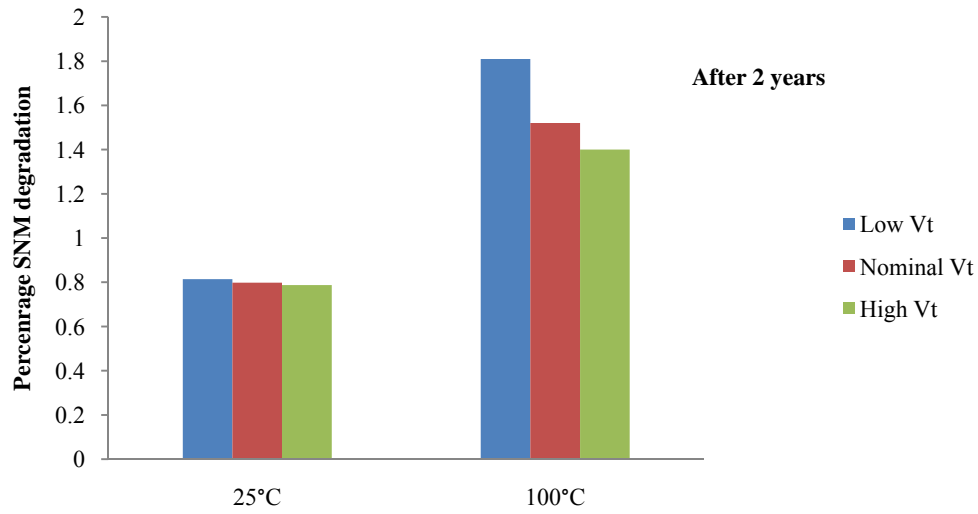


Fig. 3: SNM degradation after 2 years. SNM degradation is considerably higher in low V_t cell and at high temperature.

TABLE I. PERCENTAGE SNM DEGRADATION FOR LOW V_T , NOMINAL V_T AND HIGH V_T TECHNOLOGY CORNERS AT 25°C AND 100°C OVER 2 YEARS

	Time(Years)	1/8year	1/4year	1/2year	1year	2year
Low	25°C	7.45E-01	7.59E-01	7.73E-01	8.00E-01	8.14E-01
	100°C	1.46E+00	1.53E+00	1.60E+00	1.70E+00	1.81E+00
Nominal	25°C	7.30E-01	7.40E-01	7.59E-01	7.69E-01	7.98E-01
	100°C	1.27E+00	1.31E+00	1.38E+00	1.44E+00	1.52E+00
High	25°C	7.43E-01	7.43E-01	7.61E-01	7.78E-01	7.87E-01
	100°C	1.22E+00	1.26E+00	1.30E+00	1.34E+00	1.40E+00

In reality, V_t variation has a statistical nature showing some average and some standard deviation which result in statistical variations in SNM. Hence, a more realistic SNM measurement is the SNM statistical distribution which can be obtained using the Monte Carlo simulations. Fig. 4 shows the intra-die statistical variation of the SNM degradation for low inter-die V_t chip at 100°C. It clearly shows that the SNM distribution shifts towards lower values. From the statistical SNM distribution and by setting a target SNM (45mV), we can estimate the percentage of faulty cells. Table 2 shows the percentage of faulty cells for all three inter-die process corners: low V_t , nominal V_t , and high V_t transistors at 100°C. Fig. 5 shows the percentage rise of faulty cells due to the SNM degradation for low inter-die V_t technology corner at 25°C and 100°C. It is observed that number of faulty increase (8.2% over 2years) only at the high temperature. Even though SNM degrades slightly at 25°C but this degradation does not cause any new failures.

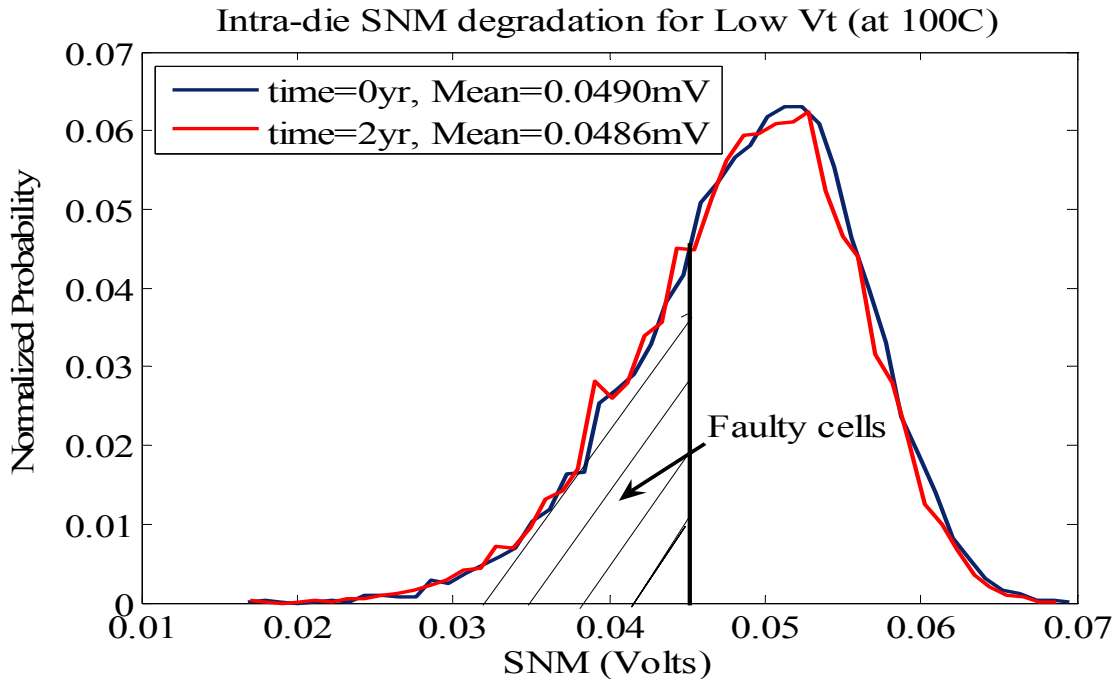


Fig. 4: Statistical distribution of SNM due to intra-die V_t variations for the low inter-die V_t chip at time 0 and after 2 years. Number of faulty cells increase over time as SNM distribution shifts to left.

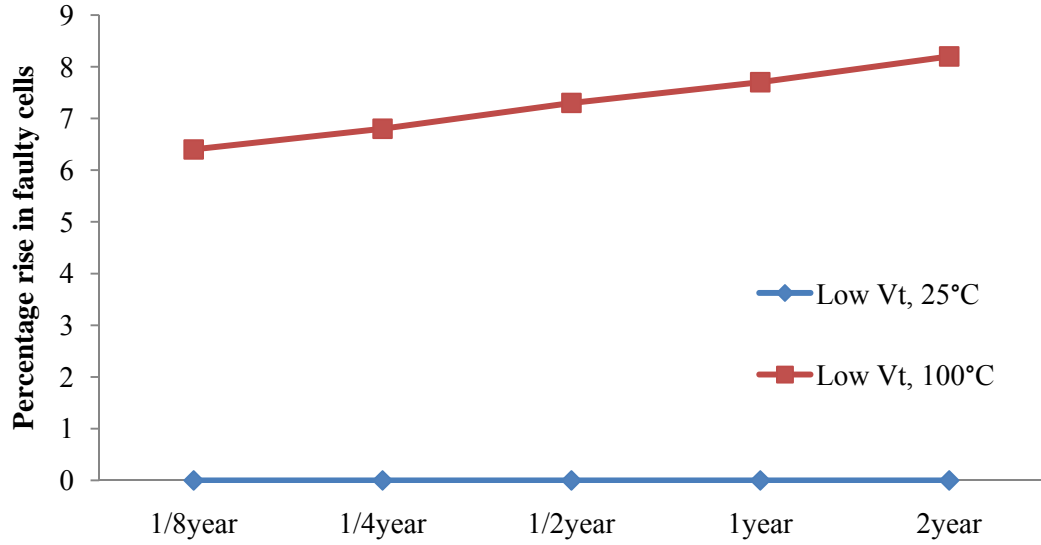


Fig. 5: Percentage rise of faulty cells due to SNM degradation increases over time at 25°C and 100°C. It is observed that number of faulty increase (8.2% over 2years) only at the high temperature.

TABLE II. PERCENTAGE OF FAULTY CELLS DUE TO SNM DEGRADATION AGING AND PROCESS VARIATIONS AT 100°C

Faulty cells (%)	0yr	1/8yr	1/4yr	1/2yr	1yr	2yr
Low Vt	26.56	28.26	28.37	28.52	28.63	28.74
Nominal Vt	1.49	1.68	1.69	1.7	1.71	1.71
High Vt	0.09	0.1	0.1	0.1	0.1	0.1

While writing a “0” to a cell storing “1,” the node V_L gets discharged through BL_B to a low value (V_{WR}) determined by the voltage division between the PMOS PL and the access transistor AXL . If V_L cannot be reduced below the trip point of the inverter $PR-NR$ within the time when word-line is high (T_{WL}), then a write failure occurs. The discharging current (I_L) at node L is the difference in the ON currents of the access transistor AXL (I_{AXL}) and the PMOS PL (I_{PL}) (i.e., $I_L = I_{AXL} - I_{PL}$)[6]. Because under NBTI PMOS becomes slower, I_{PL} goes down. So write margin improves in first write cycle but write margin degrades under NBTI effect in the worse case which is the second write cycle to write the original state back to the cell.

For the second write cycle SRAM, pull up transistor PR is not aged and PL is aged, making it slow which is not helpful in the discharging of ‘1’ from node V_R . Moreover, decreased PL strength reduces the trip voltage of the inverter $NL-PL$ which makes the pull up operation at node VR slightly harder, thereby causing slight decrease in write margin. However, this write margin decrease is insignificant as shown in Fig. 6. Write margin degradation is practically negligible and does not cause any new write failures over time. Table 2 shows results for percentage write margin degradation for low V_t , nominal V_t and high V_t technology corners at 25°C and 100°C over 2 years. Table 3 shows the number of faulty cells for all three process corners: low V_t , nominal V_t , and high V_t chips at 100°C. It is observed that low V_t chips show very few faulty cells as compared to high V_t chips. The reason is that high ON current at low V_t cells helps the fast charging and discharging of nodes which results in a faster write operation, hence resulting in less number of write failures. It is observed that there is no significant increase in number of write failures over time.

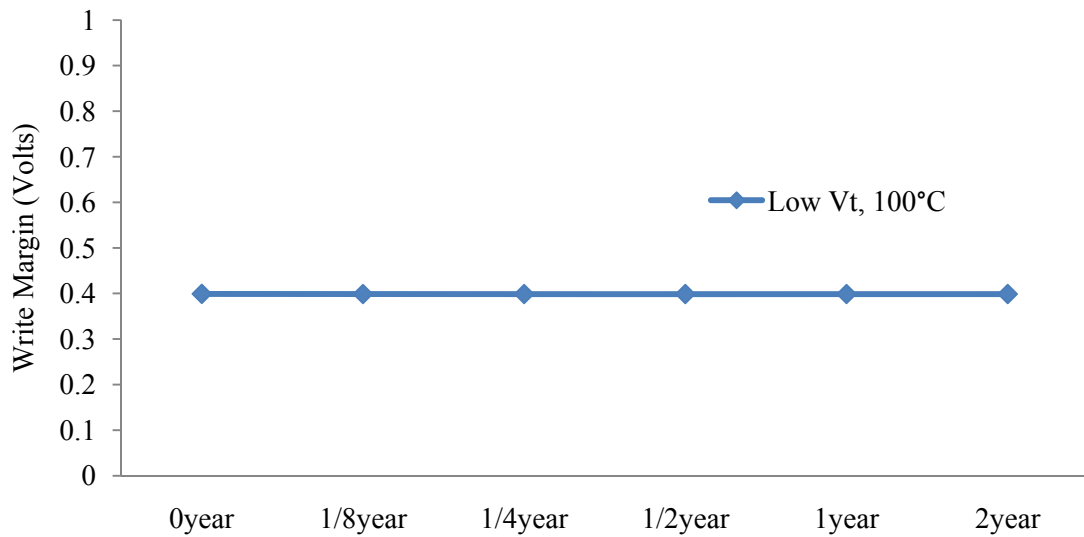


Fig. 6: Write margin slightly degrades over time. However, this degradation of write margin is practically negligible.

TABLE III. PERCENTAGE WRITE MARGIN DEGRADATION FOR LOW V_T , NOMINAL V_T AND HIGH V_T TECHNOLOGY CORNERS AT 25°C OVER 2 YEARS

	Time(Years)	1/8year	1/4year	1/2year	1year	2year
Low	25°C	0.0555864	0.0555864	0.0555864	0.055586	0.055586
	100°C	0.0751842	0.0751842	0.0751842	0.075184	0.082703
Nominal	25°C	0.0561698	0.0561698	0.0561698	0.05617	0.05617
	100°C	0.0541961	0.0541961	0.0541961	0.054196	0.054196
High	25°C	0.0093926	0.0093926	0.0093926	0.009393	0.009393
	100°C	0.0058065	0.0058065	0.0058065	0.005807	0.005807

TABLE IV. PERCENTAGE OF FAULTY CELLS DUE TO WRITE MARGIN DEGRADATION UNDER AGING AND PROCESS VARIATIONS AT 100°C (FOR TARGET WRITE MARGIN OF 0.3V)

Faulty cells (%)	0yr	1/8yr	1/4yr	1/2yr	1yr	2yr
Low Vt	0	0	0	0	0	0
Nominal Vt	0.1	0.1	0.1	0.1	0.1	0.1
High Vt	2.9	2.9	2.9	2.9	2.9	2.9

Leakage power is defined as the power consumption during the OFF state and is determined by the current flowing through OFF MOS transistors. Since the OFF transistor is not affected by the NBTI, leakage is not impacted by NBTI. For the worse case, where cell will remain in the initial state over lifetime, leakage does not change over time. Fig. 7 shows the leakage current for the low V_t SRAM cell at 100°C over time. It is observed that leakage does not change over time. This observation however is valid for the worst case situation where the cell state does not change over the lifetime. If the cell state changes, there will be some leakage reduction over time due to increased V_t .

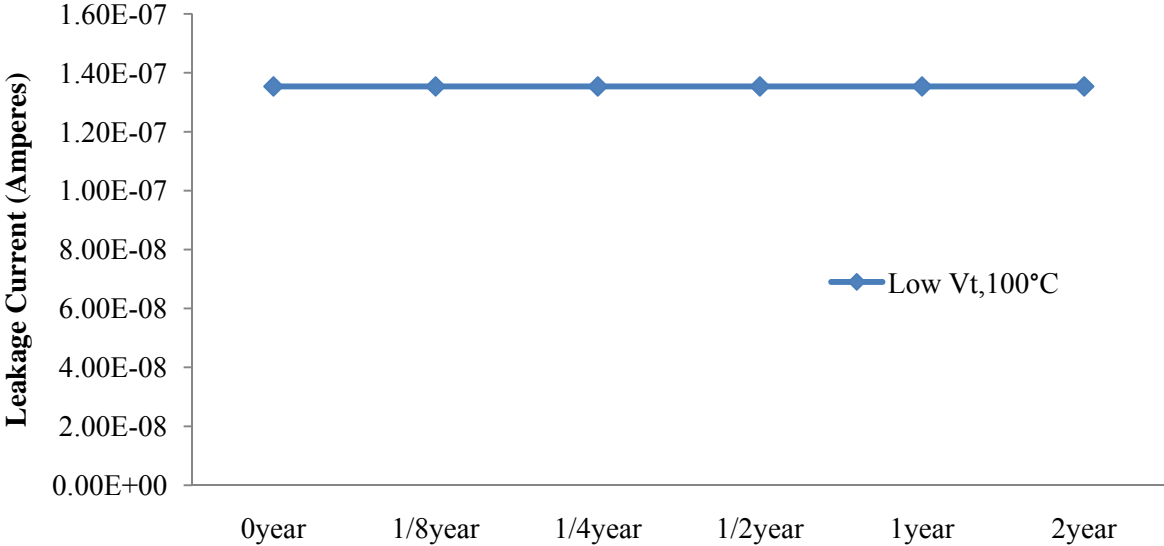


Fig. 7: SRAM cell leakage current for low V_t inter-die chip over time. NBTI does not impact leakage current.

Access time is defined as the time required for accessing the data from the SRAM cell. Access time is not impacted by NBTI because access time is determined by discharging of BL through NMOS access transistors and NMOS pull down transistors which are not impacted by NBTI. Figure 8 shows results for the read current for the low V_t SRAM cell at 100°C over time. It is observed that read current does not show any change over time. Hence, access time is not affected by aging effect.

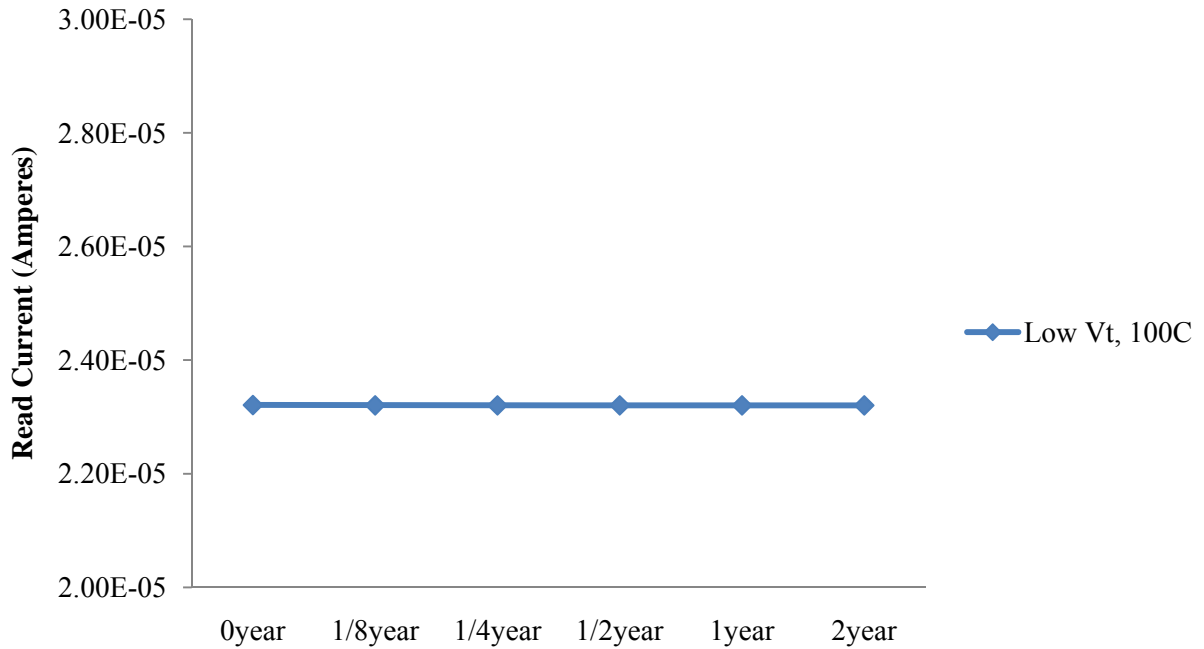


Fig. 8: SRAM cell read current for low V_t inter-die chip over time. NBTI does not impact read current.

8. AGING EFFECT IN HIGH K CMOS TECHNOLOGY

With the continuous scaling of SiO₂ gate dielectric film in conventional CMOS technology, gate leakage becomes intolerably high. After many intense research efforts, replacement of SiO₂ dielectric with high K gate dielectrics is found as an optimal solution. While HfO₂-based (and other high-candidates) show a desired effect of significantly reduced gate tunneling (leakage) current, there are still a number of fundamental issues, such as fixed charge, reduced channel mobility and trapped charge, which have to be understood and solved for successful high- integration into the silicon CMOS technology [11,12]. In high K technology, both NMOS and PMOS experience threshold voltage shift over time due to charge trapping in ON (stressed) condition. The threshold voltage shift is due to negative bias temperature instability (NBTI) in PMOS transistor and positive bias temperature instability (PBTI) in NMOS transistors. The model given in equation 1 is used to obtain threshold voltage shift values at 1/4 year, 1/2 year, 1 year and 2 year.

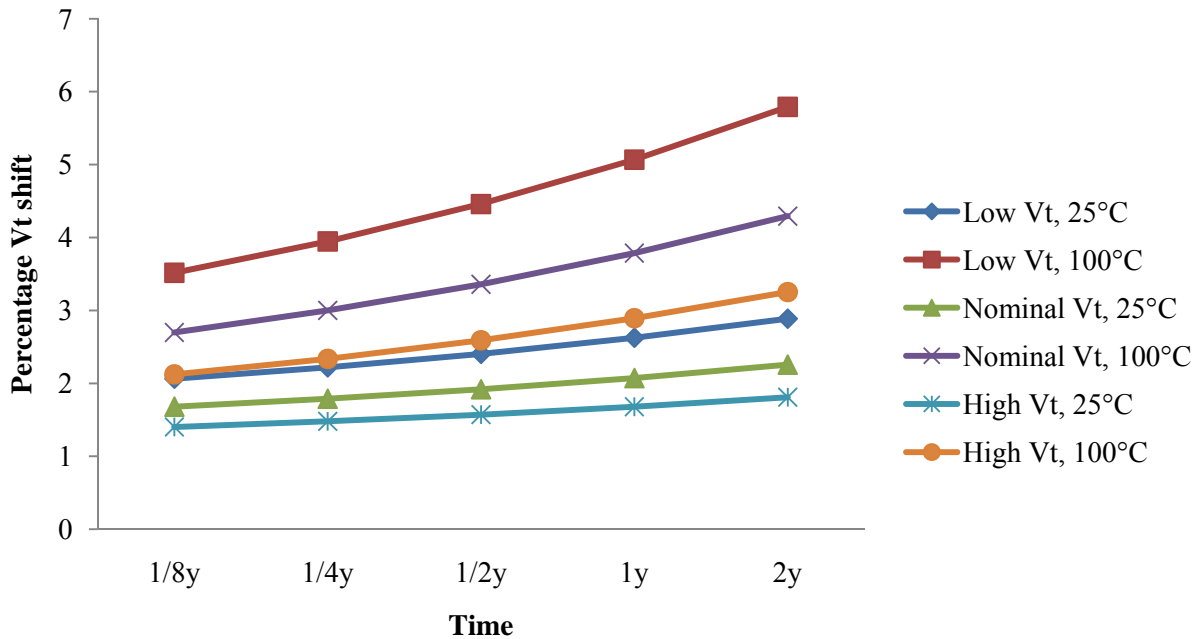


Fig 9: Percentage Vt shift in PMOS due to NBTI for the three technology corners: Low Vt transistors, Nominal Vt transistors, High Vt transistors at room temperature (25°C) and worst case temperature (100°C) for high K dielectrics.

Fig. 9 shows results for percentage V_t shift in PMOS due to NBTI for the three technology corners: Low V_t transistors, Nominal V_t transistors, High V_t transistors at room temperature (25°C) and worst case temperature (100°C) for high K dielectrics. These results show that NBTI resultant V_t shift is greater at high temperature and the low technology corner. This is due to the dependence of K_v factor in Eq. 1 on the temperature and threshold voltage. Figure 10 shows the Percentage V_t shift for NMOS and PMOS due to PBTI and NBTI respectively with two supply voltage level 0.5V and 1V for high K dielectrics. This results in V_t shift to be higher in case of high supply voltage (1V) as compared to low supply voltage (0.5V) in both NMOS and PMOS. The PBTI results in more severe V_t degradation as compared to NBTI, especially in case of high supply voltage. It is observed that in high K technology, PBTI is a bigger reliability treat as compared to NBTI.

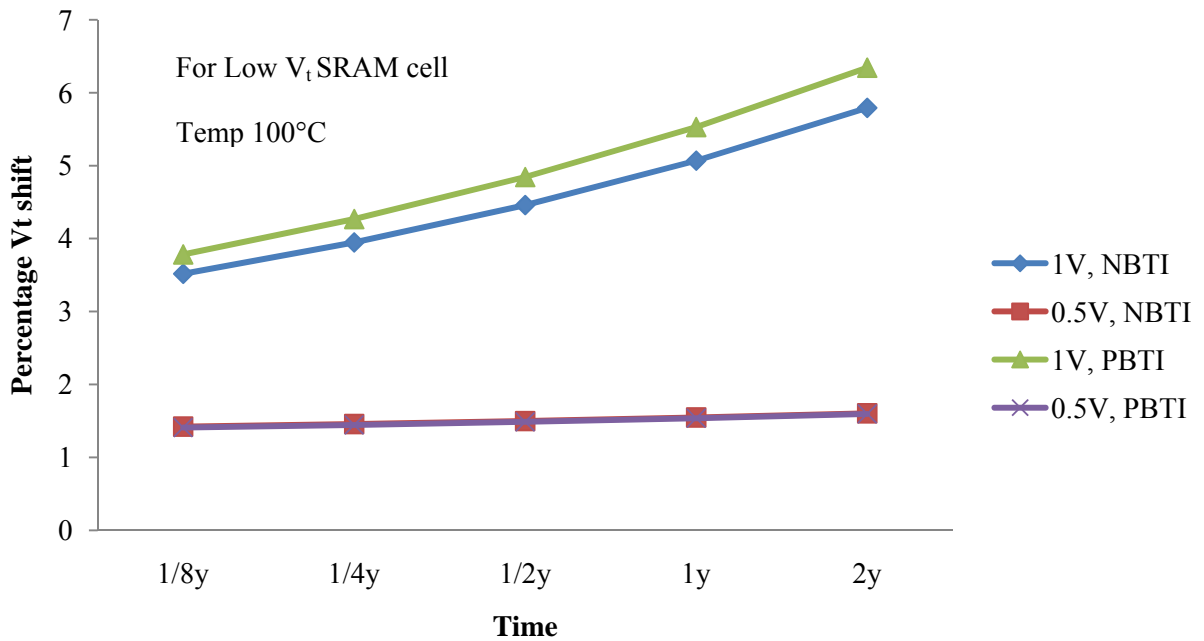


Fig 10: Percentage V_t shift for NMOS and PMOS due to PBTI and NBTI respectively with two supply voltage level 0.5V and 1V for high K dielectrics.

9.COMBINED EFFECT OF AGING EFFECT AND PVT (PROCESS, VOLTAGE AND TEMPERATURE) VARIATIONS ON 6T SRAM IN HIGH K CMOS TECHNOLOGY

Although the effect of time dependent V_t degradation on SRAM cell with high-k dielectrics due of NBTI (in case of PMOS) and PBTI (in case of NMOS) has been studied and verified by many researchers [14, 15, 16]. In this work, we analyzed the combined effect of aging effect (NBTI and PBTI), Voltage, process and temperature variation on the reliability of SRAM cell, in terms of Static Noise Margin (SNM), write margin and access time of 6T SRAM cell (shown in Fig. 2) in 32nm CMOS process. To incorporate inter-die threshold voltage variation (process variations), three different chips are used: low V_t chip, nominal V_t chip and high V_t chip. Two temperatures, 25°C and 100°C, are considered to incorporate the temperature variations. In this work, we considered two types of SRAM chips - low voltage SRAM cell (0.5V) and high voltage SRAM cell (1V). Two different supply voltages are considered to observe the effect of supply voltage on performance of SRAM cell under aging effect. Even though at low supply voltage, stress induced V_t shift will be less (see Fig. 10), but sensitivity of cell towards variation will be more. So by using two supply voltage levels, we observed that the sensitivity of cell towards variations has dominance over stress induced V_t degradation over time in case of access time, whereas in case of write operation stress induced V_t degradation is more crucial. In case of SNM, initially low voltage SRAM cell shows more degradation as compared to high voltage SRAM cell. Over long time, high voltage SRAM cell shows more SNM degradation.

Figure 11 shows the percentage degradation of SNM of low V_t , SRAM cell for two temperatures, one room temperature (25°C) and second high temperature (100°C) with two supply voltages – low voltage SRAM cell (0.5V) and high voltage SRAM cell (1V). SNM degrades over time due to stress induced V_t mismatch. Initially, SNM degradation is more in case of low voltage SRAM cell as compared to high voltage SRAM cell because low voltage SRAM cell is more sensitive to variations as compared to high voltage SRAM cell. But over prolonged time, V_t shift is considerably more in case of high supply

voltage as compared to low supply voltage. As per simulation results obtained it is evident that high voltage SRAM cell shows more SNM degradation for long period of time. To integrate the inter-die V_t variation, three types of cells: low, nominal, high V_t have been considered in this work. Fig. 12 shows the SNM degradation for low, nominal, and high V_t cells. It can be observed that low SNM degradation as compared to high V_t because low V_t cell experience greater V_t shift over time as compared to high V_t cell.

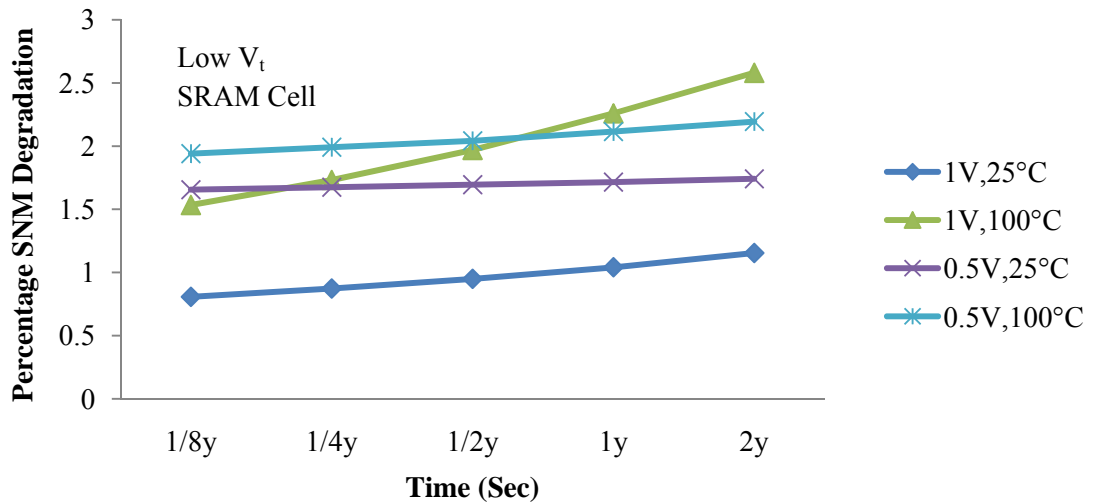


Fig. 11: SNM degradation over 2 years. Initially SNM degradation is considerably higher in low voltage SRAM cell and at high temperature. Over long period of time, high voltage SRAM cell shows more SNM degradation and at high temperature.

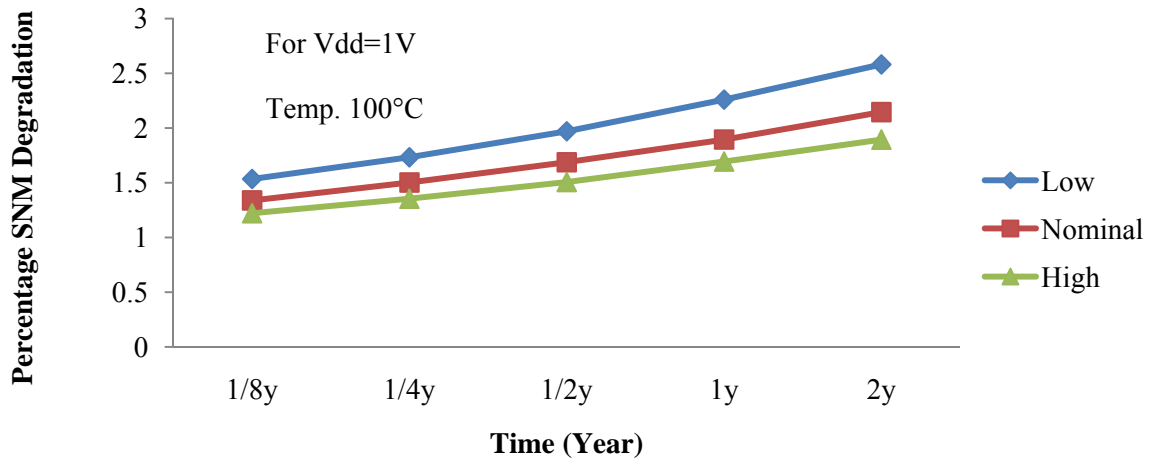


Fig 12: SNM degradation over time for the three technology corner SRAM 6T cell: low, nominal, high V_t . Low V_t cell shows more SNM degradation as compared to high V_t .

In this work, we observe the write margin of SRAM cell for worse case which is the second write cycle (when $V_R = 1$ and $V_L = 0$; PR and NL are stressed) to write the original state back to the cell. For the second write cycle SRAM, pull up transistor PL is aged and pull down transistor NR is aged. The aged PL is not helpful in the discharging operation of node $V_R = 1$, but it makes the pull up operation at node V_L harder. Moreover aged NL impedes the discharging of '1' at node V_R , so write margin degrades over time. Fig. 13 shows the percentage degradation of write margin with low V_t SRAM cell for two temperatures: room temperature (25°C) and high temperature (100°C) with two supply voltage: low voltage SRAM (0.5V) and high voltage SRAM (1V). It is observed that write margin degradation is more in case of high voltage SRAM cell (4.4% degradation in write margin) as compared to low voltage SRAM cell and degradation is more significant at higher temperature. Hence, it can be concluded that high voltage SRAM will experience faster SNM degradation as a result of high stress induced V_t variations as compared to the low voltage SRAM cell.

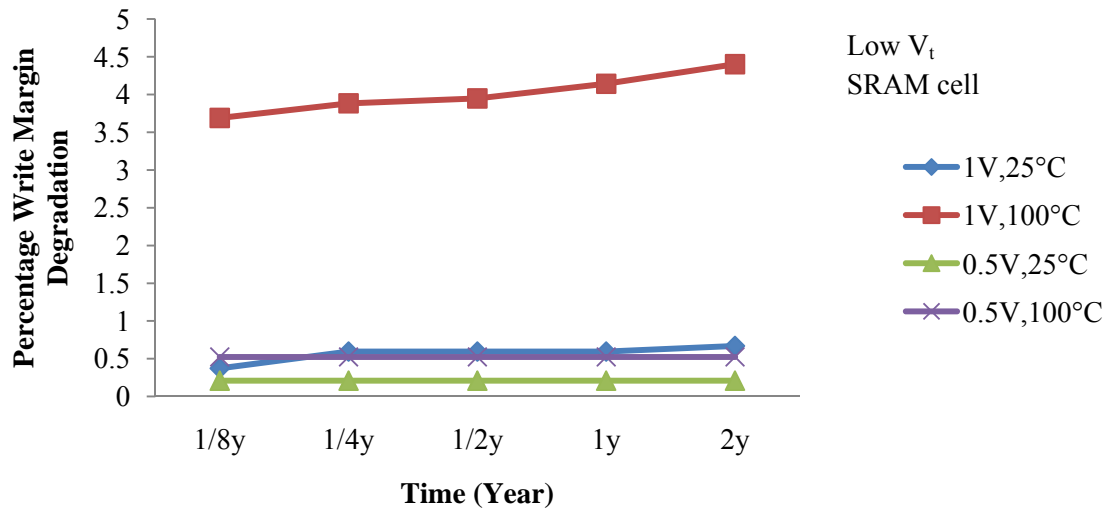


Fig 13: Write margin degrades over time and degradation is considerably higher in high voltage SRAM cell and at high temperature.

The cell access time is defined as the time required for producing a pre specified voltage difference ($\Delta_{MIN} \approx 0.1V_{DD}$) between two bit-lines (bit-differential)[8]. The access time is determined by discharging of BL through NMOS access transistors and NMOS pull down transistors. For the worse case, transistors PL and NR remains under stress. Aged NR makes discharging of BL harder which causes access time degradation in SRAM cell. Fig. 14 shows percentage degradation of access time in low V_t cell considering two temperatures: - room temperature (25°C) and worst case temperature (100°C) with two supply voltage:- low voltage SRAM (0.5V) and high voltage SRAM (1V). The results show that access time degradation is more significant at higher temperature regardless of supply voltage and SNM degradation is more in case of low voltage SRAM with 2.7% degradation in read current compared to high voltage SRAM cell. Hence, it can be concluded that low voltage SRAM will experience faster SNM degradation as a result of high sensitivity towards variations as compared to high voltage SRAM.

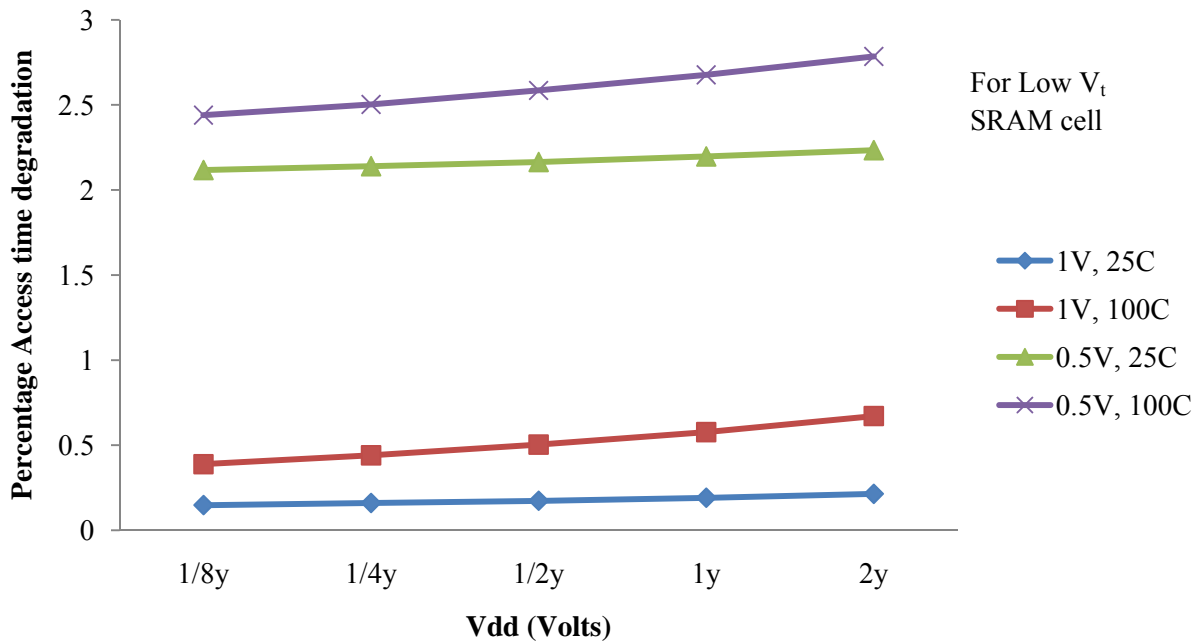


Fig 14: SRAM cell access time degradation for low V_t inter-die chip over time. Access time degradation is considerably higher in low voltage SRAM cell and at high temperature.

Leakage power is defined as the power consumption during “OFF” state and determined by the current flowing through “OFF” MOS transistors. In the worse case, pull up transistor PL and pull down transistor NR are stressed and experience V_t degradation. Since the “OFF” transistors (PR and NL) are not affected by BTI. Thus, leakage is not impacted by aging effect.

10. DESIGN SOLUTION TO MITIGATE TRANSISTOR AGING EFFECT

In this section, we will discuss the design solution to mitigate the transistor aging effect. In the previous section, we observed the effect of two different supply voltages on the performance metrics of SRAM cell under the combined effect of transistor aging effect, and process and temperature variations. We analyzed the supply voltage is an important parameter as it effects the sensitivity of circuits and also determines the stress induced threshold voltage shift.

In this work, we analyzed the effect of transistor aging on the performance metrics of SRAM: SNM (read SNM and hold SNM) and access time at different supply voltage levels. It idea was to figure out the optimal supply voltage which results in minimal degradation in SRAM performance. Furthermore, we also investigated that adaptive supply voltage scheme can be applied (in case of read SNM) to save the power consumption while maintaining the reliability of the cell over lifetime.

Fig 15 shows the SNM degradation for read and hold mode at different supply voltages. It can be observed that SNM degradation is minimal for both cases at 0.8V supply voltage. SNM degradation is more at low supply voltage levels because at low supply voltages, SRAM is more sensitive towards the variations. SNM degradation is also high in case of high supply voltages because high stress induced threshold voltages shifts. Whereas, at 0.8V SRAM cell is less sensitive towards variations as compared to low supply voltages and stress induced V_t shift is also small as compared to high supply voltages. As a result, SNM degradation is minimal at the 0.8V. In case of access time, degradation is minimal at 1V (See fig. 16). However, the sensitivity of performance metrics (SNM and access time) towards supply voltage is different; the low voltage circuits suffer more degradation. It is also observed that at nominal supply voltage (0.9V), the degradation is nearly minimal in case of both performance metrics (SNM and access time).

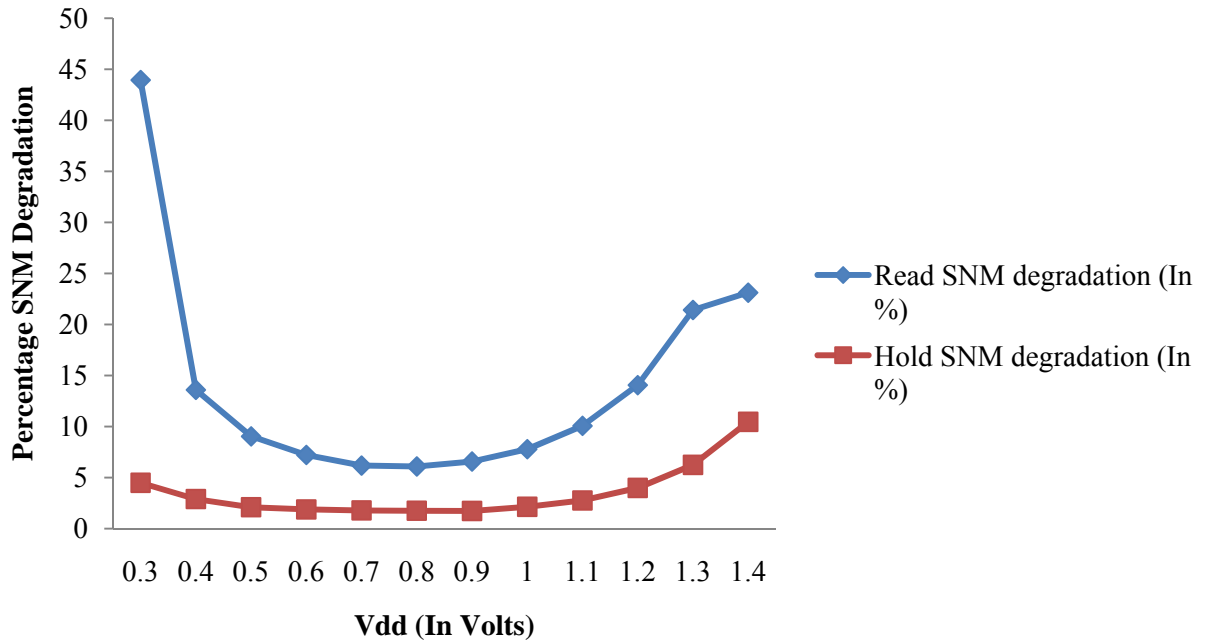


Fig 15: SNM degradation for read and hold mode at different supply voltages. It can be observed that SNM degradation is minimal in both cases at 0.8V supply voltage.

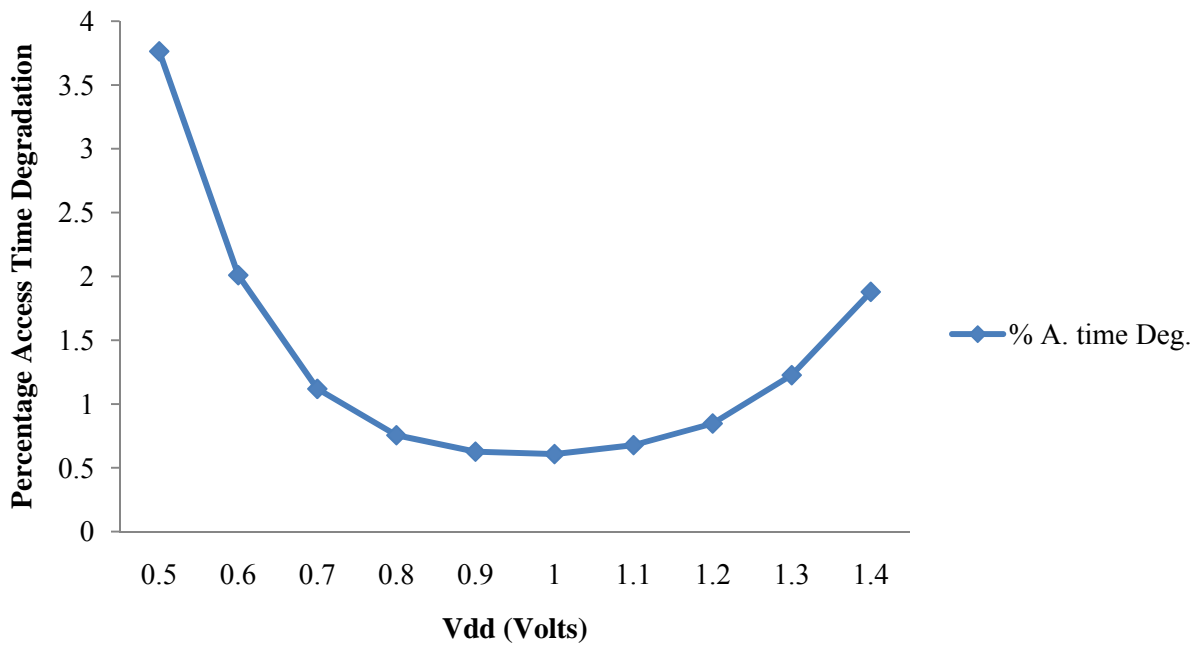


Fig. 16: Access time Degradation at different supply voltage. It can be observed that access time degradation is minimal at 1V

Figure 17 shows read SNM at time 0 and after 2 years at different voltage levels. To maintain the target SNM of 100mV over 2 year, supply voltage should be 0.83V. However, at time 0, the target SNM of 100mV can be achieved at 0.785V. The power overhead of using single voltage (0.83V) from time 0 is 22.22%. Considering if the power overhead is more than 10%, adaptive supply voltage scheme is an attractive solution to save power over time; in case of read SNM adaptive supply voltage scheme can be used to save power.

In case of hold SNM, to maintain the target SNM 100mV over 2 years, the required supply voltage is 0.407V instead of 0.4V at time 0 (see Fig 18). The power overhead is small (5.7%). So the adaptive supply voltage scheme is not viable.

In case of access time, to maintain the target access time of 0.468 ns over 2 years, the required supply voltage is 0.903V instead of 0.9V at time 0 (see Fig 19). The power overhead is small (1.66%). So the adaptive supply voltage scheme is not viable.

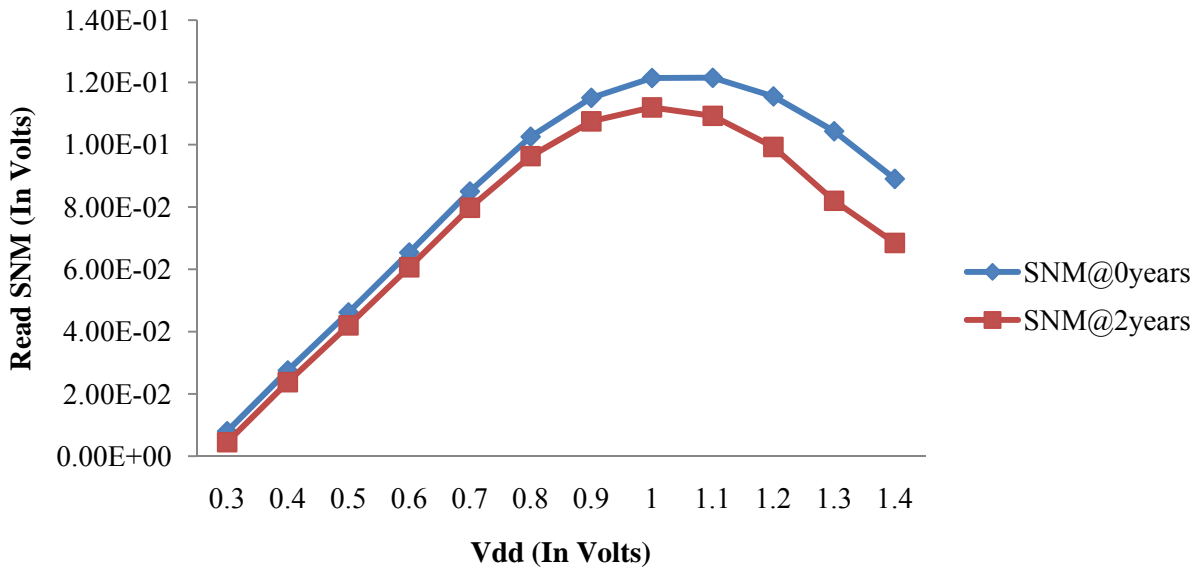


Fig 17: Read SNM at time 0 and after 2 years at different voltage levels.

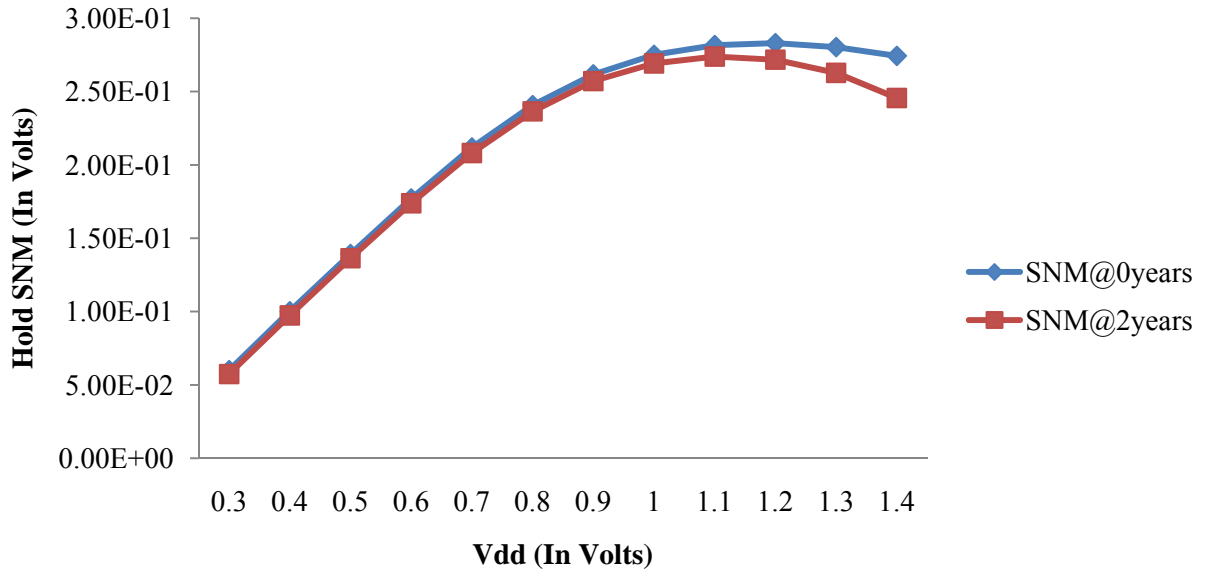


Fig. 18: Hold SNM at time 0 and after 2 years at different voltage levels.

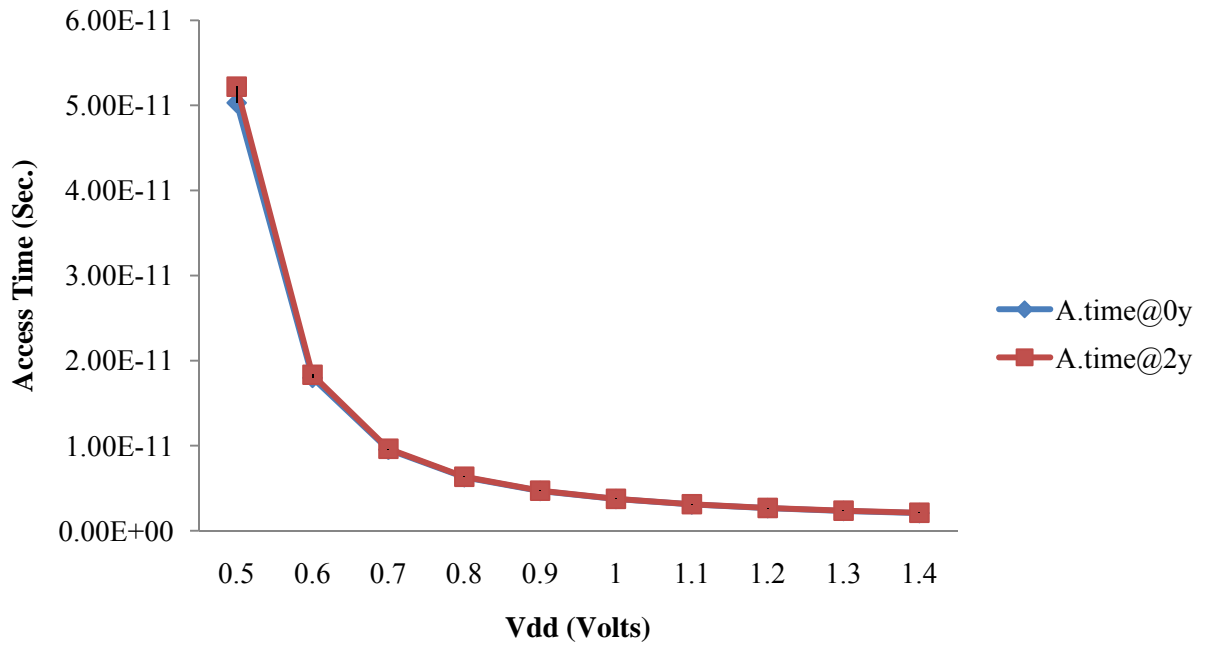


Fig 19: Access time at time 0 and after 2 years at different voltage levels.

11. CONCLUSION

In this work, the combined effect of the transistor aging, process (V_t) and temperature variations is discussed on SiO_2 and High K CMOS technology. It is observed that low V_t transistors age at faster rate than high V_t transistors and stressed induced V_t degradation rate is more significant at higher temperature. In case of SiO_2 CMOS technology, for the worse case, SNM and write margin degrades over time whereas leakage and access time do not get affected from the NBTI. Along with these observations, we also quantified our results in terms of number of faulty cells in SRAM array. It is observed that number of faulty cells rises over time (8.2% rise in faulty cells for the inter-die nominal V_t chip over 2 years) due to SNM degradation whereas rise in the number of faulty cells over time due to write failures under NBTI effect is practically negligible. Low V_t SRAM chip shows more increase in number of faulty cells over time as compared to the high V_t SRAM chip because SNM degradation due to NBTI is more significant in the case of low V_t transistors as compare to high V_t transistors. However at room temperature, NBTI does not cause any new cell failure due to SNM degradation because at the room temperature, SNM degradation is practically insignificant. In case of high K CMOS technology, for the worse case, SNM, write margin and access time degrades over time whereas leakage does not get affected from the NBTI or PBTI. We also observed the effect of supply voltage variation of the SRAM. By using two supply voltage levels, we observed that access time degradation is more in case of low supply voltage chip whereas in case of write operation, degradation is more with high supply voltage. In case of SNM, initially low voltage SRAM cell shows more degradation as compared to high voltage SRAM cell. Over long time, high voltage SRAM cell shows more SNM degradation. We analyzed that even though the sensitivity of performance metrics (SNM and access time) towards supply voltage is different, the low and high voltage circuits suffer more degradation as compared to nominal voltage circuits. We observed that in case of read SNM, adaptive supply voltage technique can be used to save power over time to maintain target read SNM where power overhead is 22.22%. However, in case of hold SNM and access time is power overhead is practically negligible, so adaptive supply voltage technique is not viable. Since

read SNM and access time are coupled together, in active mode, power can be saved by using the adaptive supply voltage scheme over time.

12. REFERENCES

- [1] K. Kang, et. al., "Impact of Negative-Bias Temperature Instability in Nanoscale SRAM Array: Modeling and Analysis," *IEEE Transactions on Computer Aided Design of Integrated Circuits and Systems*, vol. 26, no. 10, pp. 1770-1781, Oct. 2007
- [2] Rakesh Vattikonda, Wenping Wang, Yu Cao, "Modeling and Minimization of PMOS NBTI Effect for Robust Nanometer Design" in *Design Automation Conference*, July 2006.
- [3] B. C. Paul, et. al., "Impact of NBTI on the Temporal Performance degradation of digital circuits," *IEEE Electron Device Letters*, vol. 26, no. 8, Aug 2005.
- [4] Predictive Technology Models (PTM): <http://www.eas.asu.edu/~ptm/>
- [5] S. Mukhopadhyay, et. al.; "Modeling of Failure Probability and Statistical Design of SRAM Array for Yield Enhancement in Nanoscaled CMOS", *IEEE transactions on computer-aided design of integrated circuits and systems*, vol. 24, no. 12, December 2005
- [6] Kunhyuk Kang et. al., "Characterization of NBTI induced Temporal Performance Degradation in Nano-Scale SRAM array using I_{DDQ} ," *INTERNATIONAL TEST CONFERENCE*
- [7] Yangang Wang et. al., "Impact of NBTI on the Performance of 35nm CMOS Digital Circuits,"
- [8] Sanjay V. Kumar et. al., "Impact of NBTI on SRAM Read Stability and Design for Reliability" *Proceedings of the 7th International Symposium on Quality Electronic Design (ISQED'06)*
- [9] Jeffrey Hicks et. al., "45nm Transistor Reliability," *Intel Technology Journal*, Volume 12, Issue 2, 2008
- [10] E. Seevinck, et. al., "Static-noise margin analysis of MOS SRAM cells," *IEEE Journal of Solid-State Circuits*, Vol SC-22 No. 5, October 1987.
- [11] E. P. Gusev, et. al., "Ultrathin high-K gate stacks for advanced CMOS devices," in *IEDM Tech. Dig.*, 2001, p. 451.
- [12] S. Zafar, A. Kumar, et. al., "Threshold voltage instabilities in high-k gate dielectric stacks," *IEEE transactions on device and materials reliability* vol. 5, no. 1, pp. 45-64, March 2005.

- [13] S. Zafar et. al., "Charge trapping related threshold voltage instabilities in high permittivity gate dielectric stacks", *J. Appl. Phys.* 2003, vol. 93, pp. 9298-9303.

- [14] Aditya Bansal et. al., "Impacts of NBTI and PBTI on SRAM Static/Dynamic Noise Margins and Cell Failure Probability" ??????????

- [15] Aditya Bansal et. al., "Impact of NBTI and PBTI in SRAM Bit-cells: Relative Sensitivities and Guidelines for Application-Specific Target Stability/Performance," ????????

- [16] J.C. Lin et. al., "Time Dependent Vccmin Degradation of SRAM Fabricated with High-k Gate Dielectrics" ??????????????????