Fall 11-19-2014

Detection of Variable Retention Time in DRAM

Neraj Kumar
Portland State University

Detection of Variable Retention Time in DRAM

by

Neraj Kumar

A thesis submitted in partial fulfillment of the
requirements for the degree of

Master of Science
in
Electrical and Computer Engineering

Thesis Committee:
W. Robert Daasch, Chair
Glenn Shirley
Marek Perkowski

Portland State University
2014

Abstract

This thesis investigates a test method to detect the presence of Variable Retention Time (VRT) bits in manufactured DRAM. The VRT bits retention time is modeled as a 2-state random telegraph process that includes miscorrelation between test and use. The VRT defect is particularly sensitive to test and use conditions. A new test method is proposed to screen the VRT bits by simulating the use conditions during manufacturing test. Evaluation of the proposed test method required a bit-level VRT model to be parameterized as a function of temperature and voltage conditions. The complete 2-state VRT bit model combines models for the time-in-state and for the retention-time including miscorrelation. A copula is used to model the effect of miscorrelation between test and use. The proposed VRT test algorithm runtime is estimated as a function of VRT test coverage, test temperature and test voltage.

Acknowledgements

I wish to express my thanks to Prof. Robert Daasch, for his advice and encouragement. I also thank my father Nand Kishore Prasad, my whole family and loved ones, for their constant support and having faith in me.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Problem Statement

This thesis detects the presence of failing variable retention time (VRT) bits in manufactured DRAM specification. The VRT bits are sensitive to test and use environments (In Section 4.2, the sensitivity of the VRT bits to test and use environments is detailed). The VRT bits failing in test behave as hard error but the VRT bits failing in use behave as soft error. There is a need of a test algorithm which mimics the use condition to screen these bad VRT bits at the manufacturing test.

This thesis extends the previous study of DRAM retention time (See Section 1.2 for the definition of retention time) at PSU's ECE ICDT laboratory [4]. In this thesis, the effects of time-in-state of variable retention time will be added to PSU's DRAM retention time study. A model of time-in-state of the VRT bits will be presented as a function of environmental conditions such as temperature and voltage. This thesis models the VRT bits by the combination of Weibull model for marginal distribution of retention time, Copula model for retention time miscorrelation and time-in-state model (See Chapter 3 for the details of models). A DRAM array of 1Gb was assembled from the combination of marginal distribution of retention time model, Copula model of miscorrelation in retention times and time-in-state model. The assembled DRAM array was used to develop and evaluate a test algorithm for the VRT detection at temperature and voltage set points.

1.2 DRAM and Variable Retention Time

In the fast growing computer world, memory is the key to fast and improved performance of any computer system. Random Access Memory is used primarily in computer systems because of the faster read and write operations. Static Random Access Memory (SRAM) and Dynamic Random Access Memory (DRAM) are two types of Random Access Memory. Each bit of SRAM has 6 transistors. For a DRAM bit only one transistor and a tiny capacitor is required. Although DRAMs have longer access time but because of the higher density of bits on DRAM chips, cost per bit of DRAM is very low compared to SRAM. DRAMs are widely used in all computer systems as main memory. In recent years, DRAMs are also used in wireless mobile devices such as smart phones, tablets, modems and printers. Increasing density of bits on DRAM chips has also helped to reduce the size of electronic devices.

Figure 1.1 shows a DRAM subarray. The horizontal lines (labeled as Row0, Row1) in Figure 1.1 are row lines or word lines of DRAM subarray. The vertical lines (labeled as bit0, bit1) are bit lines of DRAM subarray. Each bit of a DRAM is a combination of one CMOS transistor and one capacitor. Information is stored as charge on the capacitor. DRAM bits connected to a single row line make one word. To write on a DRAM bit, the bitlines are pre-charged and the desired row line is selected to turn on the transistor. Then the capacitor is charged high or low. To read a DRAM bit, the bitlines are pre-charged to a different voltage and the row line is selected to turn the transistor on. Then charge on the capacitor is sensed by the sense amplifier through bitlines.

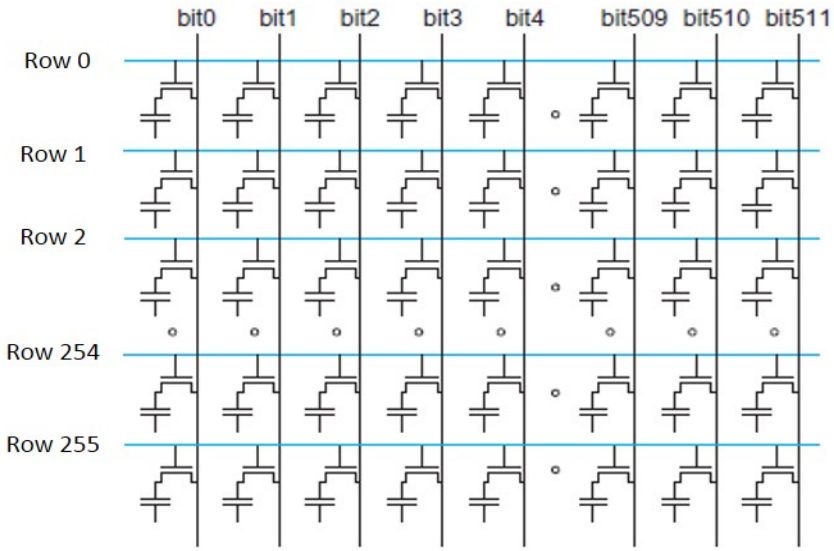


Figure 1.1: DRAM Subarray [1]

The capacitor leaks charge, so a DRAM bit can retain data for a certain time. After some time, the stored data on the DRAM bits become invalid. Each bit's data retention time is known as its characteristic data retention time. The characteristic retention times of DRAM bits vary from a few milliseconds to seconds. To retain the data for a longer duration, until a bit is read or written again, DRAM bits are refreshed periodically. Refresh is a simple operation where data is read from a bit and written back to the same DRAM bit. Figure 1.2 shows the leakage of charge on a DRAM bit and periodic refreshes to retain the data. When data is stored to any DRAM bit, it is stored as charge on its capacitor. Stored charge on a capacitor leaks with time. Before the remaining charge on a capacitor goes below the threshold, a refresh signal refreshes the bit and charge on the capacitor of the DRAM bit is restored again.

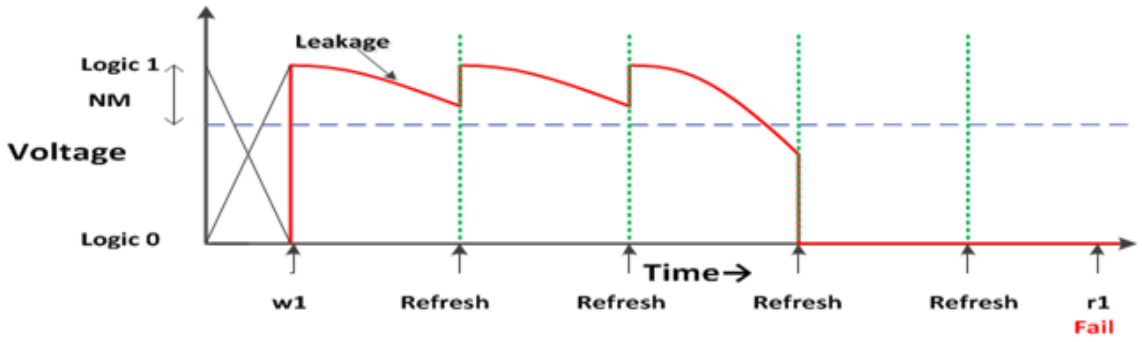


Figure 1.2: Refresh mechanism in DRAM

The time duration between two refresh signals is called refresh cycle time. Generally, refresh cycle time is less than the shortest retention time of all DRAM bits, so refresh mechanism can refresh all DRAM bits before they lose their charge from a capacitor.

In 1987, Yaney [8] reported for the first time that some bits in DRAM memory have variable retention time due to one or more atomic level defect [8]. Such defective bits are different from the normal bits because they have multiple data retention times. In this thesis, DRAM bits having single retention time at a particular environmental condition (temperature, voltage, etc.) are named as single retention time (SRT) bits and bits having variable retention time are named as variable retention time (VRT) bits. In Section 3.1, this thesis models the VRT bits with a 2-state Retention Telegraph Noise (RTN) model.

The VRT bits have two retention times associated with two retention states. The VRT bits have the longest retention time in maximum retention state and the shortest retention time in minimum retention state. Figure 1.3 shows a 2-state

VRT waveform of a VRT bit. The horizontal axis is time in hours and the vertical axis is retention time in seconds. The y-axis in the Figure 1.3, represents the maximum and minimum retention states. The VRT bit has maximum retention time of about 10sec in maximum retention state and minimum retention time of about 1sec in minimum retention state. The time spent by the VRT bit in one retention state is plotted on the x-axis.

The VRT bits stay for a longer time in one retention state. Retention time of DRAM bits at manufacturing test is measured only once. So, a VRT bit may be in maximum retention state or minimum retention state during test.

The duration of use of a DRAM bit is usually in years. During Use, VRT bits will definitely come to minimum retention state, where the VRT bits will have retention time less than refresh cycle time. When retention time of a DRAM bit is smaller than the refresh cycle time, data will be lost from the VRT bits before bits get refreshed. Loss of data from a DRAM bit will result as the faulty operation of a computer system and hence, reduces the performance of the system.

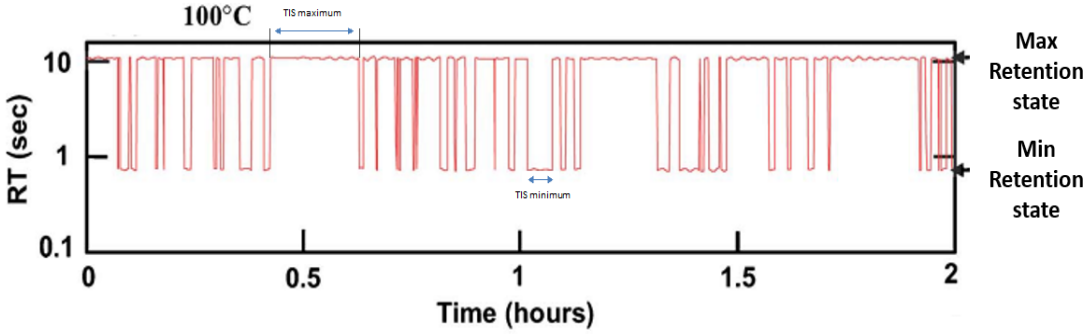


Figure 1.3: 2-state VRT waveform at 100°C [2]

1.3 Thesis Organization

In this thesis, Chapter 2 reviews the origin of VRT behavior and defects causing VRT behavior in DRAM bits. Chapter 2 also highlights the PSU ECE ICDT laboratory retention time study and retention time marginal distribution model as a function of temperature, supply voltage and bias voltage. Chapter 3 explains the step-by-step strategy to assemble a DRAM array with VRT bits at some proportion by including the concept of time-in-state of the VRT bits. The strategies discussed in Chapter 3 will be implemented in Chapter 4. The results driven by the strategies are presented in Chapter 4. A new VRT test algorithm for improved VRT screening will be presented in Chapter 4 and then the figures-of-merit of this test algorithm will be evaluated. The conclusions and recommendations are presented in Chapter 5.

Chapter 2

Background

This chapter presents the detailed summary of the origin of the variable retention time in DRAMs. The defects causing the VRT phenomenon will be discussed in this chapter. This chapter also talks about the DRAM study at PSU ECE ICDT laboratory (See Section 2.3 for the details) [4].

2.1 History: Variable Retention Time

The variable retention time defect has been present in DRAM memories since the beginning and has been observed in every technology node (See the definition of technology nodes at [9]). The VRT was reported for the first time in 1987 by Yaney [8]. The VRT defect and the impact of the VRT defect on the performance of the DRAM was explained in detail in 1992 by Restle [3]. In the DRAM retention time experiment, Yaney [8] found that some bits in DRAM memory were showing variable retention time (Figure 2.1). In Figure 2.1, the x-axis is time in seconds and the y-axis is retention time (μs) of a VRT bit. The VRT bit stays in one retention state for minutes before it makes the transition to the other retention state. The VRT bit in Figure 2.1 has the longest retention time of around 3000μs in the maximum retention state and the shortest retention time of 1300μs in the minimum retention state.

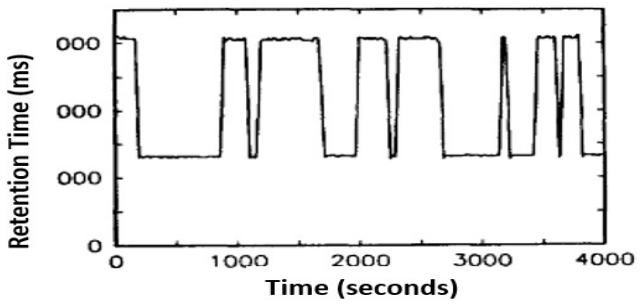


Figure 2.1: Waveform of single bit retention states with time,T=77°C. [3]

The retention time and the time spent by a VRT bit in one retention state was observed to be varying with environmental temperature and supply voltage. Figure 1.3 and Figure 2.1 shows the behavior of a VRT bit at two different environmental conditions.

Using Figure 2.1, the average time spent by a VRT bit in one retention state (either maximum or minimum) is the sum of time when a VRT bit is in one retention state (either maximum or minimum) divided by the total time. In Figure 2.1, the average time spent by the VRT bit in the maximum retention state (τ_{max}) would be the sum of the time (8 peaks in maximum retention state), when the VRT bit was in the maximum retention state, divided by the total time 4000 seconds. Similarly, the average time spent by the VRT bit in the minimum retention state (τ_{min}) would be the sum of the time (8 peaks in minimum retention state), when the VRT bit was in the minimum retention state, divided by the total time 4000 seconds. Following is the example of calculating τ_{max} and τ_{min} from Figure 2.1. For example, the first entry for τ_{max} of 130 is the first peak in Figure 2.1.

$$\tau_{max}(sec) = \left(\frac{130 + 145 + 670 + 270 + 380 + 20 + 140 + 120}{8}\right) = 235$$

$$\tau_{min}(sec) = \left(\frac{700 + 40 + 210 + 25 + 500 + 210 + 10 + 140}{8}\right) = 228$$

The thermal activity of τ_{max} and τ_{min} are shown in Figure 2.2(a). The τ_{max} and τ_{min} vary with change in environmental temperature. The average time spent by a VRT bit in the maximum retention state (τ_{max}) and the average time spent in the minimum retention state (τ_{min}) are the functions of environmental conditions such as temperature and voltage. The transition rate between retention states with a change in temperature is shown in Figure 2.2(b). The transition rate between retention states increases with increasing temperature [3] [8] (Figure 2.2(b)).

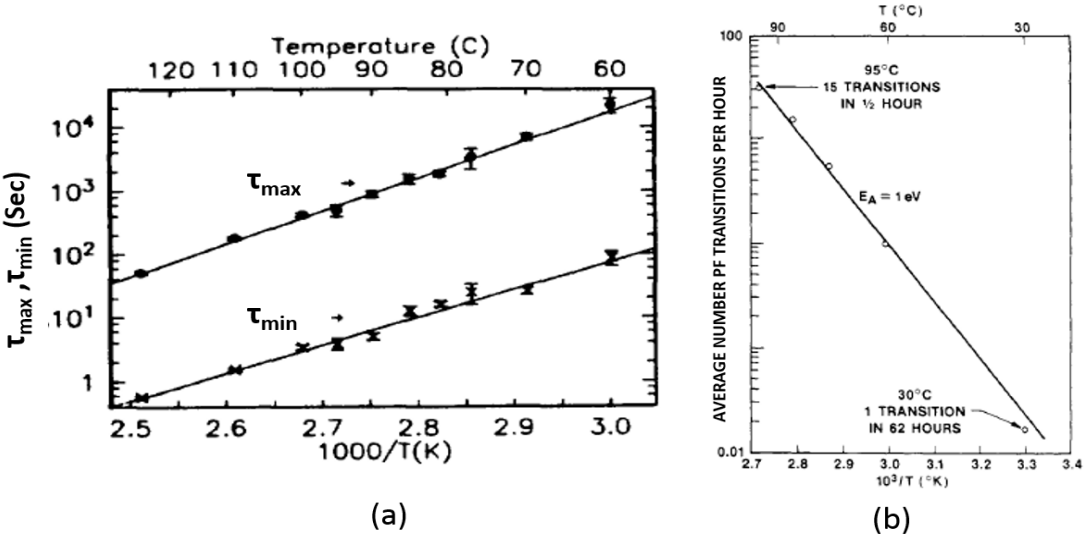


Figure 2.2: (a) Thermal activity of average time in retention states [3], (b) Transition rate between retention states with temperature [8]

The VRT phenomenon can be described also with the help of the bistable energy diagram. The activation energy level diagram is shown in Figure 2.3. The activation

energy for the maximum retention state ranges from *0.93*±*0.03*eV to *1.02*±*0.02*eV and for the minimum retention state from *0.87*±*0.03*eV to *1.01*±*0.07* eV. Generally, the activation energy for the maximum retention state is greater than the activation energy of the minimum retention state.

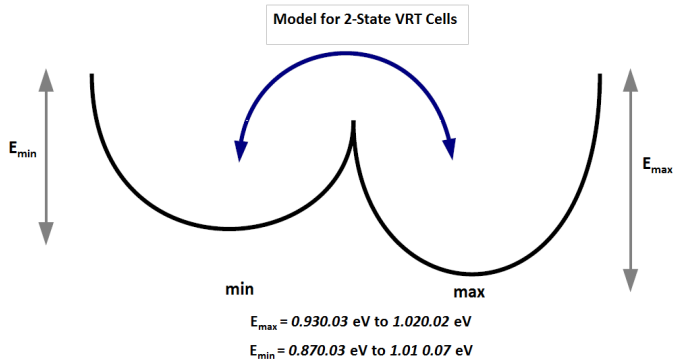


Figure 2.3: Energy diagram of VRT bits [3]. The data is from Restle [3]

The variability in the retention time of a DRAM bit is caused by the defect related to different leakage rates at the capacitor [10]. The silicon defect in the DRAM bit causing variable retention time is shown in Figure 2.4. The higher the leakage rate is at the capacitor, the retention time of that DRAM bit would be lower. There are many leakage mechanisms at the capacitor node but junction leakage is the one which is responsible for VRT phenomenon [10]. The fluctuation in the leakage current is caused by the silicon valence oxygen complexes. This atomic defect possesses two Si dangling bonds with different energy levels in the energy band gap. The activation energy levels are changed when any strain is provided. In the good state the VRT bit has a stable orientation with attached Si and oxygen bond. In the bad state, the VRT bit has an unstable orientation because of the dangling Si bond. In Figure 2.4, the two levels represent two different junction leakage currents at the capacitor; hence two different retention times.

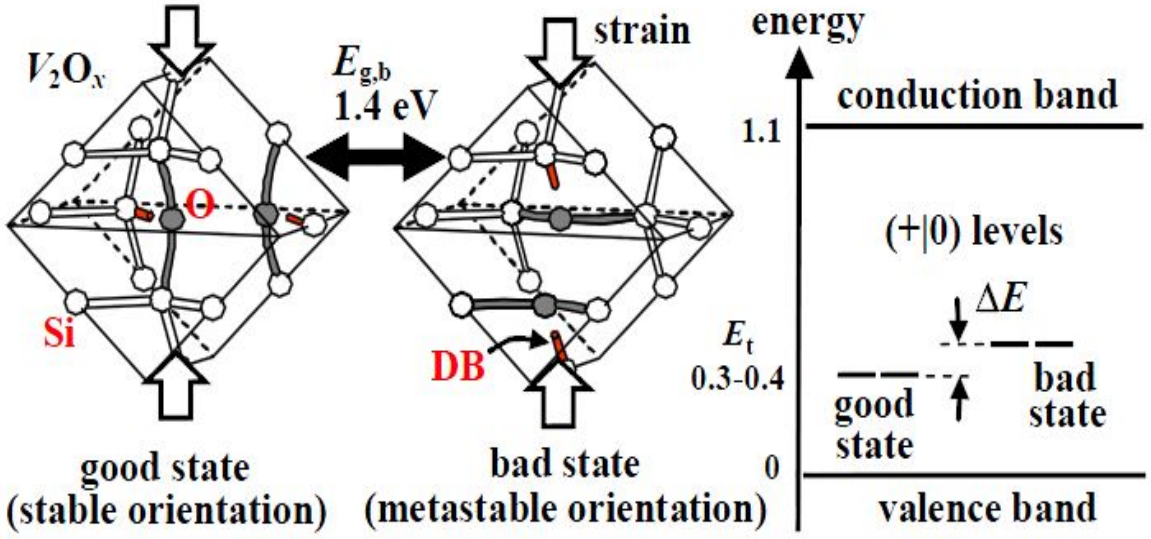


Figure 2.4: Silicon defect [2]

2.2 Modeling of VRT Phenomenon

This VRT phenomenon is a simple Markov process which has no memory for the preceding transitions or history. The VRT phenomenon is an example of Random Telegraph Noise (RTN) which will be discussed in detail in Chapter 3.

2.3 Two-state VRT (mis)correlation

Previously, at Portland State University ECE ICDT laboratory, a retention time experiment was done for embedded DRAM (eDRAM) [4]. The eDRAMs are DRAMs integrated with processors on the same die. The eDRAMs are logic process builders. The eDRAM bits have a metal stack capacitor.

At ICDT laboratory, the eDRAM arrays were tested on a Credence Quartet tester with temperature controlled by a Silicon Thermal Powercool LB300-i controller.

Temperature was measured by a calibrated sensor on the silicon die. Retention time of the eDRAM bits were measured at 12 different steps. The bits having data retention time greater than retention time set were pass and the bits having retention time less than retention time set were fail. The physical location (x,y) of each failing and passing bit in the array was recorded for each retention test. Pass/fail was recorded at 12 different retention steps for each eDRAM bit for 18 different environmental conditions. The environmental conditions were:

- Temperatures: 105° C, 115° C, 125° C
- Supply Voltage,Vd: 0.8 V, 1.0 V, 1.2 V
- Substrate bias,Vp: 0.4 V, 0.45 V

Retention times were measured in twelve steps starting from 262 μs to 2659 μs. Five repetitions of the retention time measurement were done for each bit to capture any variability in the retention time. There was a time interval between the two iterations of retention time measurement. Time interval between measurements ensures that both retention states of VRT bits have been detected.

Figure 2.5 is the table for the pass/fail data record experiment at different environmental conditions. The left half of the table represents the environmental conditions and the right half of the table is the record of pass and fail. In the table (Figure 2.5), identity column presents the unique identification of the eDRAM bit by recording skew, chip, macro, PX and PY. The column environmental condition is the combination of Vp,Vd and temperature. In the results of test column, the maximum retention time (IRetMax) and minimum retention time (IRetMin) were recorded in retention time steps. The variability in the retention time of a bit was

also recorded (IRetDelta). Retention times were measured in loop for five times. Pass is represented by 0 and fail by 1.

Identity					Environmental Condition			Results of Test			
skew	chip	macro	PX	PY	VP	VDD	temp	IRetMin	IRetMax	IRetDelta	LoopGroups
4	1	0	1	238	0.4	0.85	125	10	11	1	000000000001 000000000011 000000000001 000000000011 000000000001
4	1	0	1	238	0.4	1	125	10	11	1	000000000011 000000000011 000000000011 000000000011 000000000001
4	1	0	1	238	0.4	1.2	125	10	11	1	000000000001 000000000011 000000000011 000000000001 000000000001
4	1	0	1	238	0.45	0.85	125	8	9	1	000000000111 000000001111 000000000111 000000000111 000000001111
4	1	0	1	238	0.45	1	125	9	9	0	000000000111 000000000111 000000000111 000000000111 000000000111
4	1	0	1	238	0.45	1.2	125	10	11	1	000000000011 000000000001 000000000001 000000000011 000000000001
4	1	0	16	520	0.4	0.85	105	4	4	0	000011111111 000011111111 000011111111 000011111111 000011111111
4	1	0	16	520	0.4	0.85	115	3	3	0	000111111111 000111111111 000111111111 000111111111 000111111111
4	1	0	16	520	0.4	0.85	125	2	2	0	001111111111 001111111111 001111111111 001111111111 001111111111
4	1	0	16	520	0.4	1	105	3	3	0	000111111111 000111111111 000111111111 000111111111 000111111111
4	1	0	16	520	0.4	1	115	2	2	0	001111111111 001111111111 001111111111 001111111111 001111111111
4	1	0	16	520	0.4	1	125	1	1	0	011111111111 011111111111 011111111111 011111111111 011111111111
4	1	0	16	520	0.4	1.2	105	2	2	0	001111111111 001111111111 001111111111 001111111111 001111111111
4	1	0	16	520	0.4	1.2	115	1	1	0	011111111111 011111111111 011111111111 011111111111 011111111111
4	1	0	16	520	0.4	1.2	125	1	1	0	011111111111 011111111111 011111111111 011111111111 011111111111
4	1	0	16	520	0.45	0.85	105	3	3	0	000111111111 000111111111 000111111111 000111111111 000111111111
4	1	0	16	520	0.45	0.85	115	2	2	0	001111111111 001111111111 001111111111 001111111111 001111111111
4	1	0	16	520	0.45	0.85	125	1	1	0	011111111111 011111111111 011111111111 011111111111 011111111111
etc.	etc.	etc.	etc.	etc.	etc.	etc.	etc.	etc.	etc.	etc.	etc.

Figure 2.5: Retention test pass/fail data record of a DRAM bit at 18 different environmental conditions.

Figure 2.6 shows a table for the count of bits in different retention time bins at one environmental condition. In the PSU ICDT DRAM experiment, the size of eDRAM was 48 Mb, which is the sample size here. The retention times of bits for a sample size of 48 Mb were measured twice. The last 3 rows in the table describe the marginal distribution of retention time for the first retention time measurement and the first three columns in the table describe the marginal distribution of retention time for second retention time measurement.

The retention time measurements were divided into 12 steps from 264 μs to 2658 μs, each column on the axis is the record of fail for the given retention time. The retention time set is represented by r (μs). The count of the bits which failed

retention test for retention time set r (μs) is represented by N. Cumulative N is the sum of fails at all retention tests less than r (μs).

$$N = \sum_{i=1}^{x} N_i$$

Fail fraction in part per million (ppm) for the given retention time r (s) is listed in the last row on both of the axis.

For example, for retention time 700μs, 38 bits have same retention time in both measurements. In 1st retention time measurement (x-axis of Figure 2.6), 61 bits have retention time of 700μs. The cumulative number of bit having retention time 700μs is 109 (sum of 5, 13, 30 and 61). The cumulative fraction (F) of bits having retention time 700μs is 2.2.

2nd R.T.				T=125 C Vp=0.45 V Vd=1.2 V SS=48,750,000												
F = Cum N/SS (PPM)	Cum N	N	r (µs)													
			2658	0	0	0	0	1	0	1	1	1	1	18	69	N/A
33.6	1639	273	2442	0	0	0	0	0	0	0	0	1	14	104	95	59
28	1366	343	2222	0	0	1	0	0	0	1	2	10	97	92	120	20
21	1023	237	2006	0	0	0	0	0	0	1	5	46	83	86	14	2
16.1	786	211	1786	0	0	0	0	0	1	4	59	80	56	6	2	3
11.8	575	156	1571	0	0	0	0	0	1	29	68	53	2	1	2	0
8.6	419	134	1351	0	0	0	0	0	36	56	38	2	1	1	0	0
5.8	285	109	1135	0	0	0	0	15	71	20	2	1	0	0	0	0
3.6	176	69	915	0	0	0	12	38	18	0	0	0	1	0	0	0
2.2	107	61	700	0	0	8	43	9	0	0	0	0	0	0	0	1
0.9	46	27	480	0	1	19	6	0	0	0	0	1	0	0	0	0
0.4	19	13	264	0	11	2	0	0	0	0	0	0	0	0	0	0
0.1	6	6	0	5	1	0	0	0	0	0	0	0	0	0	0	0

0	264	480	700	915	1135	1351	1571	1786	2006	2222	2442	2658	r (µs)	1st R.T.
5	13	30	61	63	127	112	175	195	255	308	302		N	
5	18	48	109	172	299	411	586	781	1036	1344	1646		Cum N	
0.1	0.4	1	2.2	3.5	6.1	8.4	12	16	21.3	27.6	33.8		F = Cum N/SS(PPM)	

Figure 2.6: Retention time data collected in PSU experiment at T=125° C and Voltage=1.2V [4]

Similarly, retention time marginal distribution data was collected for 18 different environmental conditions in PSU's DRAM retention time study. The collected retention time data was fitted to Weibull distribution to present a model for marginal distribution of retention time for the given environmental condition. The cumulative probability distribution of Weibull distribution is given by Eq 2.1. Size parameter of Weibull distribution is represented by α and the shape parameter by β.

$$F(r) = 1 - Exp\left[-\left(\frac{r}{\alpha}\right)^{\beta}\right] \tag{2.1}$$

where r = retention time, α = scale parameter and β = shape parameter.

The fit of data to the Weibull distribution can be visually accessed by plotting Weibit plot. A Weibit plot is a plot of ln(-ln(1-F(r)) versus ln(r). A straight line is expected on the Weibit plot if the data is a good fit for Weibull distribution (See Section 3.1.2 for the details of fitting).

Figure 2.7 shows the miscorrelation in the retention time of DRAM bits. Retention time of DRAM bits recorded in the first measurement is plotted on the x-axis and retention time recorded in the second measurement is plotted on the y-axis. The good bits have the same retention time at both measurements and termed as static retention time (SRT) bits. The SRT bits are populated on the diagonal of this plot. The bits having different retention times at two measurements lie off diagonal, these bits are VRT bits.

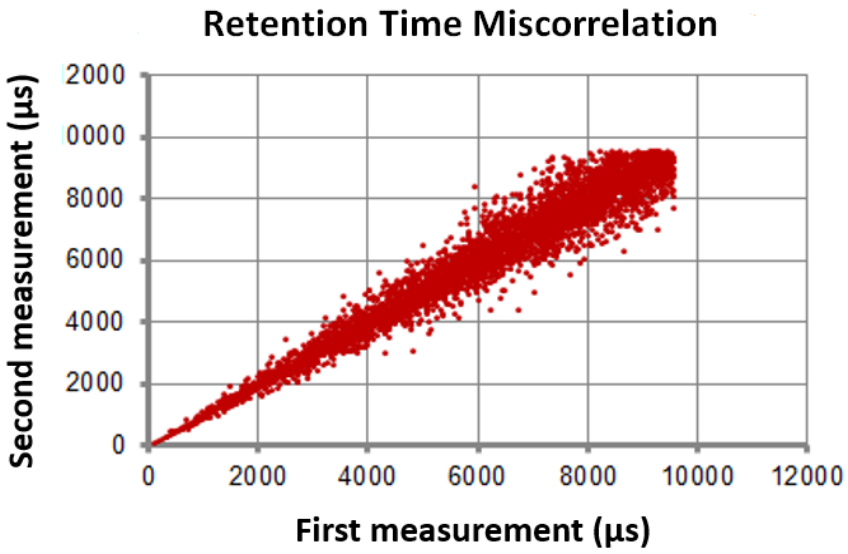


Figure 2.7: Retention time miscorrelation

The retention time experiment at PSU did not include the concept of TIS of VRT bits.

Chapter 3

Strategy: Identifying the Test for Detecting VRT Phenomenon

This chapter assembles a DRAM array of 1Gb with VRT bits in the proportion observed in Chapter 2, 18% ([4]). The assembled DRAM array allows the evaluation of the possible methods to screen the VRT bits in the next chapter.

This thesis presents a time-in-state model (See Section 3.1.3) for the VRT bit. This thesis also integrates the time-in-state model with marginal distribution of retention time model (See Section 3.1.2 for details) and (mis)correlation model (See section 3.2 for details) to get a complete VRT bit model.

3.1 VRT Bit Model

This section introduces the necessary elements to model a VRT bit. The VRT bit has four attributes: maximum retention time, minimum retention time, average time-in-state maximum and average time-in-state minimum. The VRT model models the attributes of a VRT bit by combining the time-in-state model, marginal distribution of retention time model and (mis)correlation model. This section describes each model in detail.

3.1.1 RTN Model

The VRT phenomenon is an example of Random Telegraphic Noise (RTN) model. The VRT phenomenon in DRAM can be modeled by the RTN model. Time-in-state in the RTN model of the VRT bit is memoryless process. The memoryless

process is special because it does not remember the time of the last transition. Figure 3.1 shows a 2-state (maximum and minimum state) RTN waveform of a VRT bit at 105°C and 1.2V [2]. The RTN waveform of the VRT bit in Figure 3.1 has two important elements: time-in-states (x-axis) and retention times (y-axis).

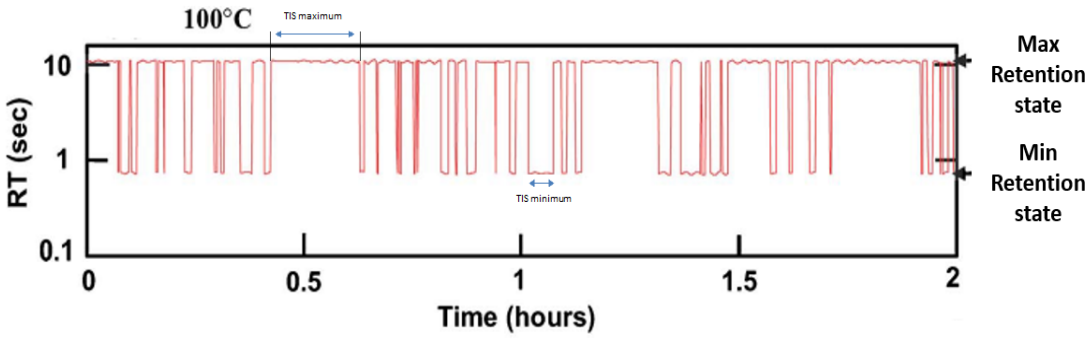


Figure 3.1: 2-state RTN waveform at T = 105°C, Voltage = 1.2V [2]

In the RTN model, the individual samples are drawn from exponential distribution for time-in-state maximum and time-in-state minimum. The time-in-state maximum and the time-in-state minimum of a VRT bit can be generated by sampling the cumulative exponential distribution. The RTN model synthesizes the x-axis of Figure 3.1 which are the time-in-states (time-in-state maximum and time-in-state minimum) of a VRT bit.

Eq 3.1 shows the general expression of the cumulative exponential distribution. Eq 3.1 can be solved to get the sample value of the cumulative exponential distribution (Eq 3.2).

$$F(t) = 1 - \exp(-t/\lambda) \tag{3.1}$$

$$t = -\lambda * ln(1 - F(t)) \tag{3.2}$$

where t = time and λ = parameter of distribution.

Eq 3.2 can be parameterized to get the distribution of time-in-state maximum and the distribution of time-in-state minimum. A sample value of time-in-state maximum can be obtained from Eq 3.3 and a sample value of time-in-state minimum can be obtained from Eq 3.4.

$$TIS_{max} = -\tau_{max} * \ln(1 - rand) \tag{3.3}$$

$$TIS_{min} = -\tau_{min} * \ln(1 - rand) \tag{3.4}$$

where $rand$ = random number between 0 and 1, τ_{max} = average time-in-state maximum and τ_{min} = average time-in-state minimum.

With the help of Eq 3.3 and Eq 3.4, 100 sampled values were generated for time-in-state maximum and time-in-state minimum respectively. The 100 sampled values of time-in-state maximum and time-in-state minimum were plotted to get the VRT waveform of a VRT bit (Figure 3.1).

Figure 3.2 shows the histogram of measured data of time-in-state maximum (Figure 3.2(a)) and time-in-state minimum (Figure 3.2(b)) of a VRT bit [5]. The superimposed lines (red line) on the histograms are the exponential fitted model. To model the time-in-state with RTN, the histogram in Figure 3.2 should be a good fit to the exponential model. The histogram data of time-in-state maximum and time-in-state minimum are visually good fits to the exponential model [5] [8]. The author did not provide any statistical curve fitting data. The RTN model can be used to model the time-in-state of a VRT bit.

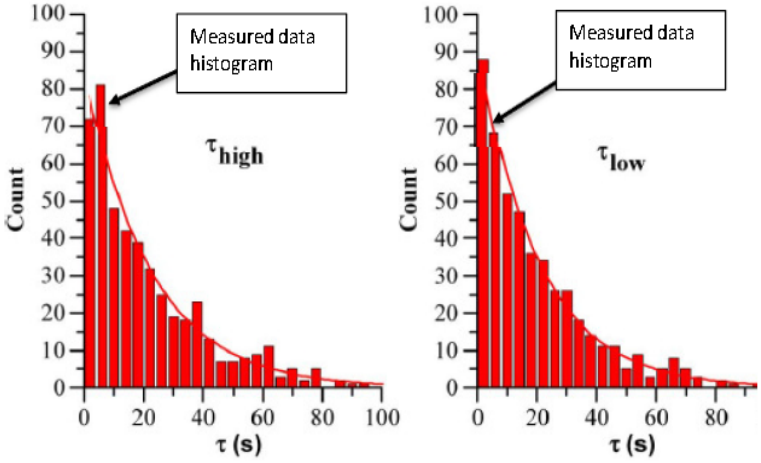


Figure 3.2: Histogram of Time-in-state maximum and time-in-state minimum [5]

The two retention times of a VRT bit can be obtained by sampling the marginal distribution model of retention time (See Section 3.1.2 for details).

3.1.2 Fitting Marginal Distribution of Retention Time

The samples of retention time pair can be obtained by sampling the marginal distribution of retention time model. In [4], a Weibull model on retention time was presented to sample the pair of independent retention times of a VRT bit. A VRT bit's retention time is sampled from the empirical cumulative distribution function (Eq 3.5a) for the given environmental condition. [4].

In [4], the temperature and voltage dependence of the shape and size parameters of the Weibull model on retention time were investigated. The values of β and α of the Weibull model for a particular environmental condition are obtained by plotting the Wiebit plot. Figure 3.3 is a Weibit plot for T=125°C and voltage

1.2V. The x-axis of the Weibit plot is the log of retention times and the y-axis is Weibit (Eq 3.5b). The value of β is obtained from the slope of the line and the value of α is obtained from the intercept (Eq 3.5c).

$$F(r) = 1 - exp\left[-\left(\frac{r}{\alpha}\right)^{\beta}\right] \tag{3.5a}$$

$$W = ln(-ln(1 - F)) \tag{3.5b}$$

$$W = \beta lnr - \beta ln\alpha \tag{3.5c}$$

where r = retention time, α = scale parameter and β = shape parameter.

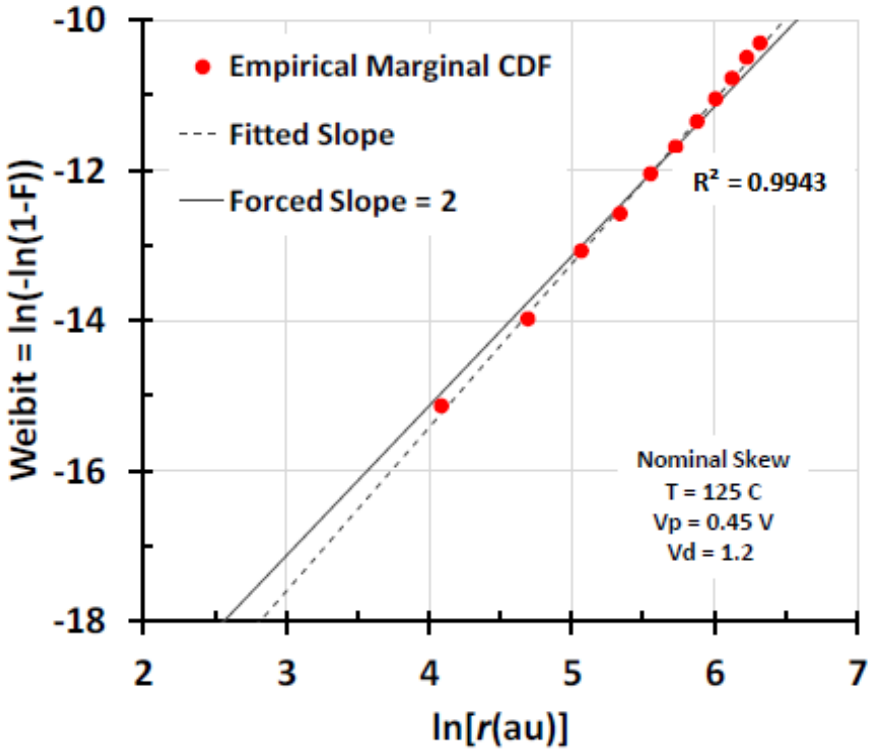


Figure 3.3: Weibit plot for retention time marginal distribution at T = 125°C and Voltage = 1.2V

In [4] similarly, pairs of β and α were obtained for each of the 18 environmental conditions by fitting each marginal distribution retention time data to a Weibit plot. In [4], the studies showed that the value of shape parameter (β) was constant for different environmental conditions. The scale parameter (α) was well fitted by Arrhenius temperature and exponential voltage (Eq 3.6) [4]. The fitted parameters are listed in Table 3.1.

$$\alpha = \alpha_0 * Exp(a(V_p - V_{pref})) * Exp((V_d - V_{dref})) * Exp\left(\frac{Q}{k_B}\left(\frac{1}{T} - \frac{1}{T_{ref}}\right)\right) \quad (3.6)$$

Parameter	$\ln \alpha_0$	a	b	V_{pref}(V)	V_{dref}(V)	Q(eV)	T_{ref}(°C)
Fitted Value	11.55	-5.85	-1.57	0.45	1.2	0.60	125

Table 3.1: Parameters and their fitted values for $ln\alpha$ model [4].

3.1.3 Fitting RTN Time-in-state Model

The average time-in-state data of a VRT bit at different temperatures and voltages were collected from the Kim, et. al paper (Figure 3.4) [5]. Figure 3.4 plots the values of τ_{max} and τ_{min} of a VRT bit at different voltages and temperatures. The solid circles (black) are τ_{max} obtained by simulation and the solid squares are τ_{min} obtained by simulation. The circles and the squares (red) represent the measured τ_{max} and τ_{min} respectively for a VRT bit. The left side in Figure 3.4 is the variation of τ_{max} and τ_{min} with the changing voltages and the right side is the variation in τ_{max} and τ_{min} with changing temperature.

The y-axis in the plots are average time-in-state. In these plots, the value of τ_{max} or τ_{min} is given for either temperature or voltage. So, it took some effort to interpolate the reference voltage and temperature for the data collected from Kim's paper.

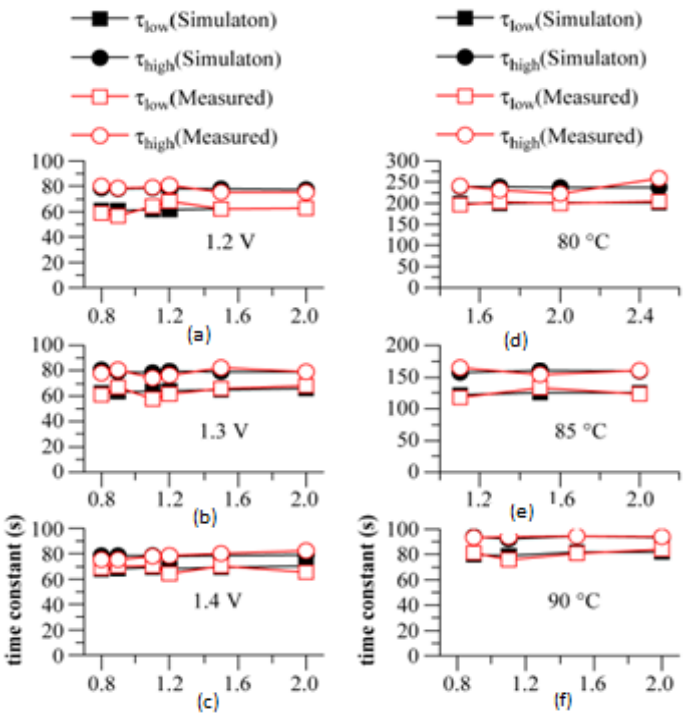


Figure 3.4: Average time-in-state maximum (τ_{max}) and average time-in-state minimum (τ_{min}) at different voltages and different temperatures [5].

To get the final average value of τ_{max} and τ_{min} irrespective of delay between two measurements, the average of τ_{max} and τ_{min} for different delays was taken.

For example, the values of τ_{max} measured with different delays at 1.2V (Figure 3.4(a)) were 80sec, 76sec, 76sec, 81sec, 75sec and 75sec. The average of these values (77sec) are taken to get the final τ_{max}. The average value (77sec) of τ_{max} for 1.2V is shown in Table 3.2(a). Similarly, all other values of τ_{max} and τ_{min} have been calculated for different voltages and different temperatures.

The average values of τ_{max} and τ_{min} at different voltages and temperatures collected from Figure 3.4 are shown in Table 3.2. Table 3.2(a) shows the values of average

τ_{max} and average τ_{min} at different voltages obtained from plots 3.2(a), 3.2(b) and 3.2(c). Table 3.2(b) shows the values of average τ_{max} and average τ_{min} at different temperatures obtained from plots 3.2(d), 3.2(e) and 3.2(f).

The reference temperature and reference voltage in Kim's paper for this data collection were 93°C and 1.4V respectively.

Voltage (V)	τ_{max} (Sec)	τ_{min}(Sec)
1.2	63	77
1.3	63	78
1.4	68	77
1.5	63	78
1.6	64	75

(a)

Temp(°C)	τ_{max} (Sec)	τ_{min} (Sec)
80	208	228
85	127	158
90	80	95
93	68	78
97	43	55

(b)

Table 3.2: Average values of τ_{max} and τ_{min} at different voltages (a) and at different temperatures (b) collected from Figure 2.3 [5]. Reference temperature 93°C, reference voltage 1.4V

Arrhenius temperature and exponential voltage gave an excellent fit for the average time-in-state model (Eq 3.7). Eq 3.8 and Eq 3.9 shows the logarithmic time-in-state maximum and logarithmic time-in-state minimum model respectively. The

description of the parameters is given in Table 3.3.

$$\tau = A * exp\left[B(V - Vref)\right] * exp\left[\frac{Q_{ref}}{k_B}\left(\frac{1}{T} - \frac{1}{T_{ref}}\right)\right] \quad (3.7)$$

$$ln\tau_{max} = lnA_{max} + B_{max}(V - V_{ref}) + \frac{Q_{ref_{max}}}{k_B}\left(\frac{1}{T} - \frac{1}{(T_{ref})}\right) \quad (3.8)$$

$$ln\tau_{min} = lnA_{min} + B_{min}(V - V_{ref}) + \frac{Q_{ref_{min}}}{k_B}\left(\frac{1}{T} - \frac{1}{(T_{ref})}\right) \quad (3.9)$$

Parameter	Description
A_{max}(Temperature co-efficient for $ln\tau_{max}$)	76.76
A_{min} (Temperature co-efficient for $ln\tau_{min}$)	64.06
B_{max} (Voltage co-efficient for $ln\tau_{max}$)	-0.05
B_{min} (Voltage co-efficient for $ln\tau_{min}$)	0.04
$Q_{ref_{max}}$ (Activation energy for $ln\tau_{max}$ in eV)	0.98
$Q_{ref_{min}}$ (Activation energy for $ln\tau_{min}$ in eV)	0.98
$k_B(eVK^{-1})$	$8.617 * 10^{-5}$
V_{ref} (Reference voltage in Volt)	1.4
T_{ref} (Reference temperature in Kelvin)	366

Table 3.3: Description of parameters of time-in-state model.

In Kim's [5] paper, the time-in-state measurements were done in the range of 80°C to 97°C. The VRT screening requires the screening of the VRT bits outside this range (80°C to 97°C) as well.

With the help of the parameters in Table 3.3, the average time-in-state (τ_{max} and τ_{min}) of the VRT bits were extrapolated to both higher and lower temperatures and voltages.

3.2 Copula Model for Miscorrelation in Two Retention Times

The samples of retention time obtained from the Weibull model (Eq 3.5) are statistically independent of each other. In [4], the miscorrelation in the retention times were studied for different environmental conditions. In [4], the two independent retention times of a VRT bit were correlated using the Clayton Copula. In the Copula approach, the two correlated numbers (x,y) are generated between 0 and 1 to get the two correlated retention times of a bit.

Clayton Copula density function is given by the Eqn 3.10. Clayton Copula has two random variables, x and y.

$$C(x,y) = \left(x^{-\theta} + y^{-\theta} - 1\right)^{\frac{-1}{\theta}} \qquad (0,\infty) \tag{3.10}$$

where θ = (0,∞)

To get the correlated pair of x and y, Eqn 3.11 is derived from Eqn 3.10. In Eqn 3.11, w is an independent random variable uniformly distributed on [0,1],

$$w = \frac{\partial C(x,y)}{\partial x} = x^{-\theta+1}\left(x^{-\theta} + y^{-\theta} + 1\right)^{\frac{\theta+1}{\theta}} \tag{3.11}$$

Eqn 3.11 is solved to get Eqn 3.12. A sampled correlated pair (x,y) can be obtained from Eqn 3.12. In Eqn 3.12, x and w are independent random numbers, uniformly distributed on [0,1].

$$y = \left[\left(w^{\left(\frac{-\theta}{\theta+1}\right)} - 1\right) * x^{-\theta} + 1\right]^{\left(\frac{-1}{\theta}\right)} \tag{3.12}$$

where w = random [0, 1], x = [0,1] and θ = correlation constant.

The 4000 pairs of correlated numbers are obtained by Clayton Copula model (Eqn 3.12). The sampled pairs of correlated numbers are plotted on Figure 3.5. The x-axis is the first number (x) of the correlated pair and the y-axis is the second number (y) of the correlated pair.

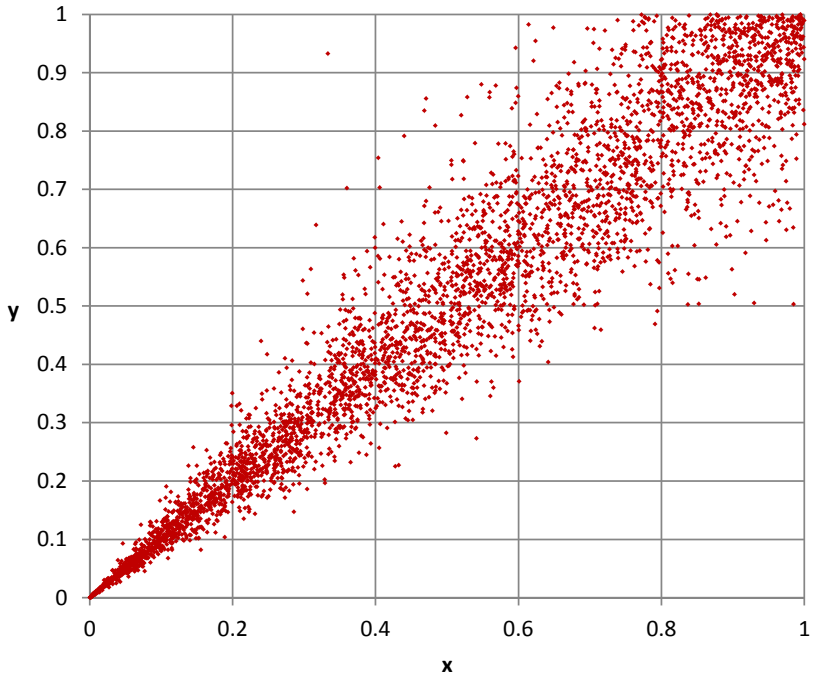


Figure 3.5: A 4000 sample scatterplot of the 2D (bivariate) Clayton copula([4]).

A pair of two retention times of a VRT bit can be obtained with the help of correlated pairs (x, y) returned from the Clayton Copula model (Eq 3.12).

The cumulative correlated pairs can be used to generate a value from a selected marginal distribution. The first retention time is obtained by using the value of x in the inverted Weibull marginal distribution (Eq 3.13) and the second retention time is obtained by using the value of y in the same inverted Weibull marginal distribution.

$$r = exp\left[(ln(\alpha) + \frac{ln(-ln(1 - F(r)))}{\beta}\right] \tag{3.13}$$

The 4000 pairs of retention time are obtained from the Weibull equation by putting the values of x, y pairs obtained from the Clayton Copula model. The 4000 pairs of Clayton dependent retention times are plotted on Figure 3.6. The x-axis in Figure 3.6 is the 1st retention time and the y-axis is the 2nd retention time. Most of the VRT bits have retention times much larger than the test limits. This thesis focuses on the bits which fall on the tail of the Copula distribution (Figure 3.6). To get the DRAM bits at the tail, the Clayton Copula is sampled with rejection. Some of the random samples do not meet the requirements. For the DRAM retention time, the sample from the Clayton is rejected if do not meet the requirements.

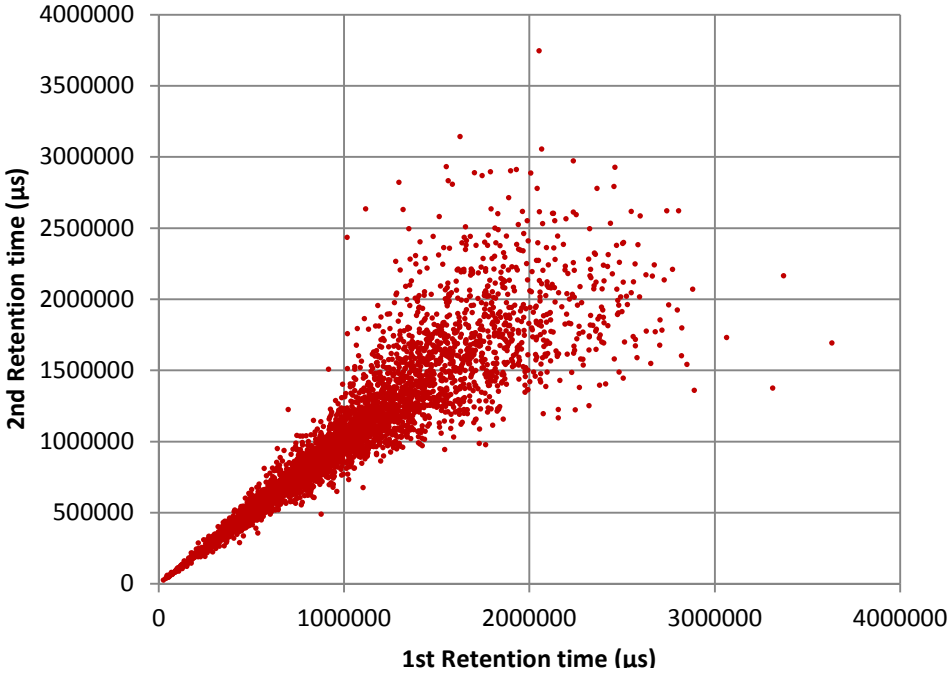


Figure 3.6: A Clayton dependent 4000 sample scatterplot with Weibull marginal retention time distributions.

Figure 3.7 is the distribution of the tail of Figure 3.6. The Clayton copula is sampled for 100PPM on both x and y axis to get the truncated distribution of retention time at the tail of Figure 3.6.

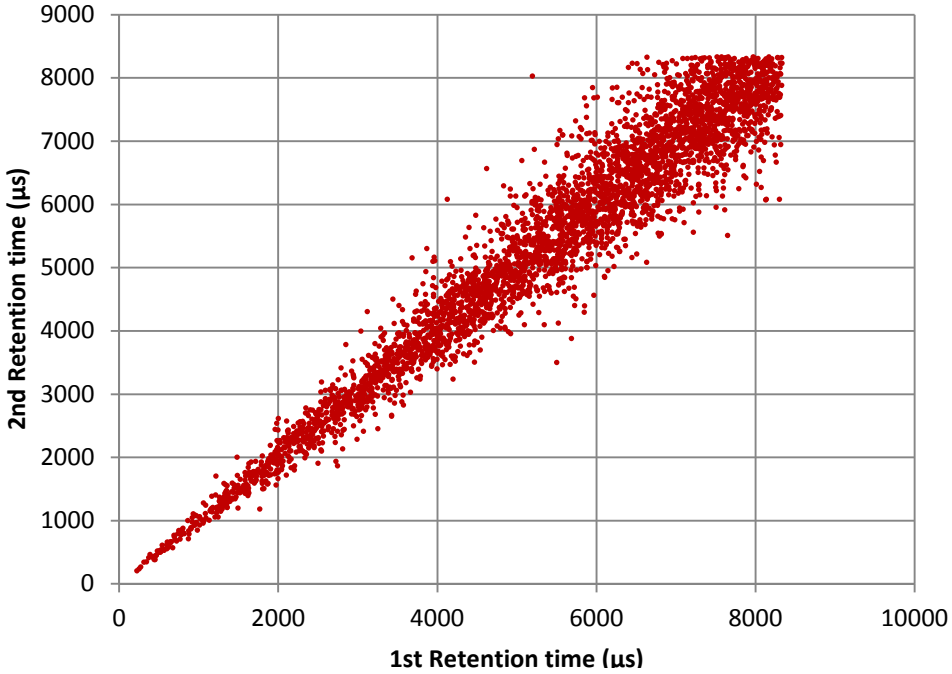


Figure 3.7: Marginal distribution of retention time obtained by sampling with rejection.

The copula based approach of modelling the retention time of VRT bits gives a pair of correlated retention times. The correlated retention time pairs will be used to synthesize VRT bits and then a DRAM array with VRT bits.

3.3 Synthesized VRT Bit

The VRT bit could be characterized by the four attributes: maximum retention time, time-in-state maximum, minimum retention time and time-in-state minimum. The parameterized model discussed in the earlier sections can be used to generate a VRT bit.

3.4 DRAM Memory Array

In the manufacturing testing of DRAM, though each bit of DRAM is screened, but DRAMs are built in arrays. The strategy to screen the VRT bits in DRAM has to be associated with the DRAM arrays instead of DRAM bits. The DRAM arrays are assembled from the DRAM bits. The remaining of this section talks about the necessary steps for assembling an array and introduces the mechanisms of testing an array.

By using the independent time-in-state and the Copula dependent retention times, a common size 1Gb DRAM array will be assembled. The 1Gb DRAM array will have independent bits including VRT bits and SRT bits. Notice that the VRT bits are off diagonal in Figure 3.6 and the SRT bits are on the diagonal.

The 1Gb DRAM array will be used to evaluate the screening of VRT bits at different temperatures and voltages. The bits in the DRAM array will be screened as pass/fail. The bits (including VRT bits) having retention time longer than test retention time will be passes otherwise fails. The largest fraction of the bits in the DRAM array have a minimum and maximum retention time much longer than Test and Use limits of the screen. The screen for the VRT bits will be evaluated for the smallest (smaller retention time) few percent of all bits in the DRAM array.

A thorough survey of different existing memory tests will be done. The screen for the VRT bits will be evaluated for different March tests. The screen will also be evaluated for IFA-9 and IFA-13 test (Section 4.6).

A new test algorithm for VRT screening will be introduced (Section 4.11). A set, pause, and repeat test algorithm will be developed using the concept of time-in-state to effectively screen the bad VRT bits in the DRAM array.

The figures-of-merit for the presented VRT test algorithm will be calculated at different temperatures, voltages, test limits and test runtimes. A tradeoff will be done between temperatures, voltage, test limit and test runtime to get the desired VRT coverage from the new test algorithm.

Chapter 4

Results: VRT Test and Detection

Chapter 3 discussed about synthesizing the VRT bits and DRAM arrays. This chapter talks about the testing of synthesized VRT bit and DRAM arrays for pass and fail for specific test and use conditions.

4.1 VRT Bit

Each VRT bit in a DRAM has four attributes: maximum retention time, minimum retention time, average time-in-state maximum and average time-in-state minimum.

4.1.1 Two Retention Times

Each VRT bit in a DRAM has maximum retention time and minimum retention time. Figure 4.1 shows a 2-state waveform of a VRT bit with maximum and minimum retention times at 100°C and 1.2V. The x-axis in Figure 4.1 is time in hours and the y-axis is the retention time in seconds. The VRT bit in Figure 4.1 keeps changing its retention time from maximum retention time to minimum retention time and from minimum retention time to maximum retention time. Notice that the maximum and minimum retention times of a VRT bit remain constant at the given environmental condition.

4.1.2 Time-in-state

The VRT bit has two retention states: maximum retention state and minimum retention state. The VRT bit has maximum retention time in the maximum retention state and minimum retention time in the minimum retention state.

The amount of time spent in maximum retention state is termed as the time-in-state maximum and the amount of time spent in minimum retention state is termed as time-in-state minimum. The maximum and minimum retention times of a VRT bit are always constant but the time-in-state is variable.

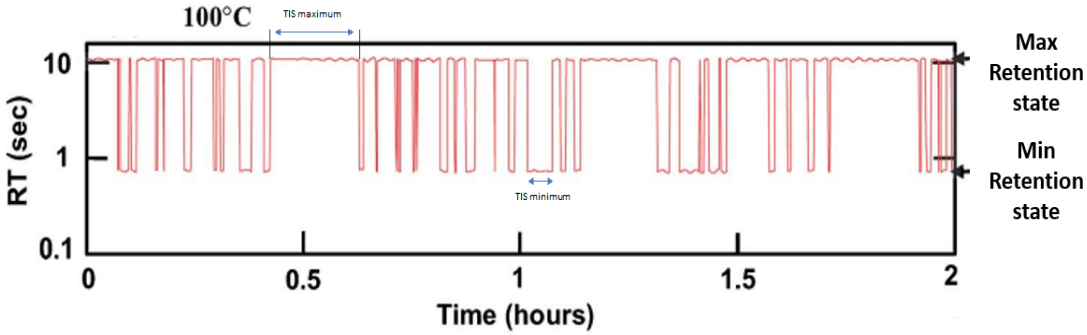


Figure 4.1: 2-state VRT waveform at T = 100°C, Volatge = 1.2V [2]

4.2 Test and Use

The variable retention times of a VRT bit produce miscorrelation at the retention time measurement during manufacturing test and use period. Figure 4.2 shows an example of the behavior of a VRT bit during test and use. The two black arrows represent the instantaneous measurement of retention time during test. Test 1 detects the VRT bit in maximum retention state and Test 2 detects the VRT bit in minimum retention state. The red arrows represent the measurements of

retention time during use.

Test is an instantaneous measurement of retention time, so the VRT bit could be in either maximum retention state or minimum retention state during test. The period of use of a DRAM bit is normally in years and the retention time of the VRT bit is measured billions of times. The VRT bit surely comes to minimum retention state during use.

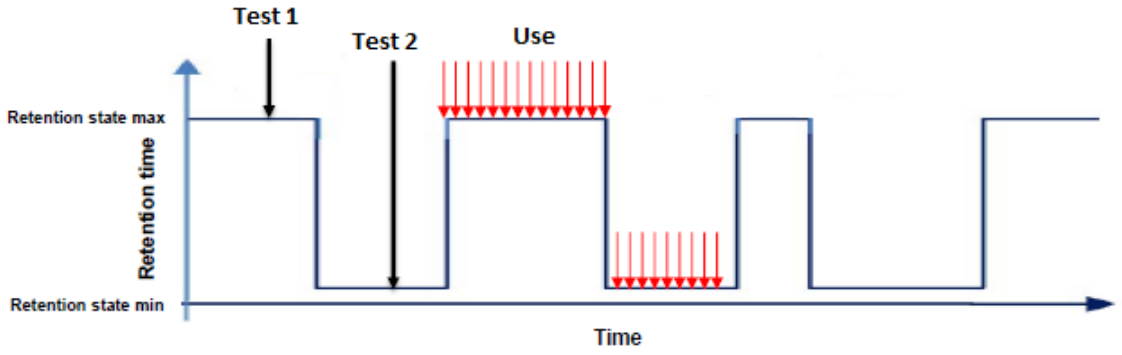


Figure 4.2: Example: Retention time of the VRT bit at test and use. Black arrows: Instantaneous measurement of retention time at test. The continuous red arrows are retention time measurements in use.

Figure 4.3 shows an example of the detection of minimum retention state with the help of refresh mechanism in the DRAM. If a VRT bit makes transition to the minimum retention state then the rate of leakage increases at the capacitor.

After writing Logic 1 (w1) to a capacitor, the stored charge on the capacitor gradually leaks with time. The refresh signal restores the charge on the capacitor before the charge goes below the threshold. When a VRT bit makes the transition to minimum retention state, suddenly the leakage rate at the capacitor increases and the charge at the capacitor goes below the threshold before the refresh signal

arrives. The next refresh signal senses the charge at the capacitor as Logic 0 and restores Logic 0 at the capacitor. The following refresh signals will keep restoring Logic 0 on the capacitor. When the data at the capacitor is read (r1), the expected data does not match with the stored data on the capacitor and fail is detected.

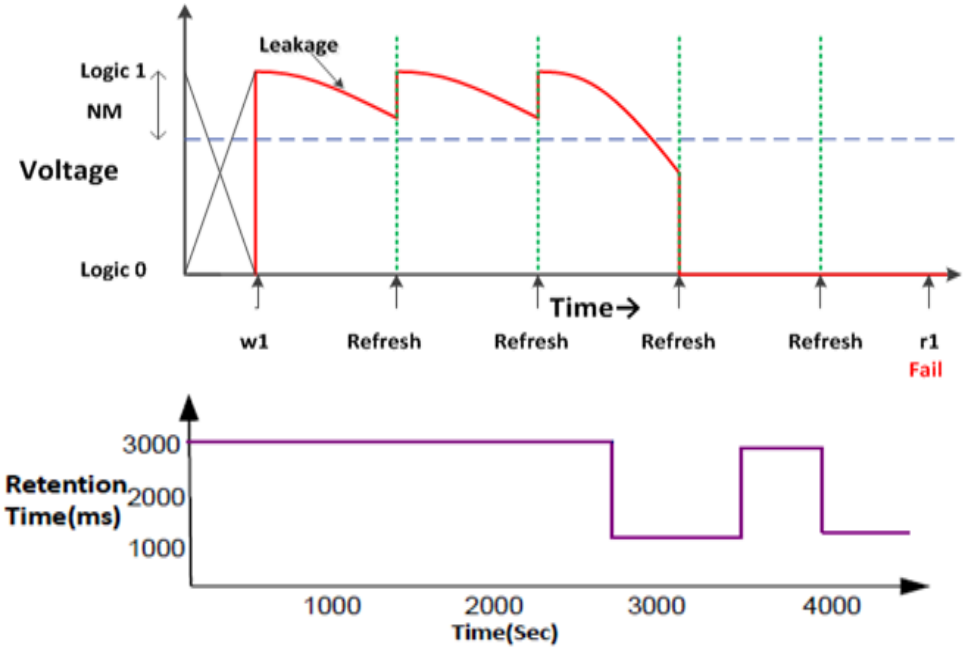


Figure 4.3: Example: Detection of minimum retention state with the help of refresh mechanism in DRAM

The refresh cycle time is programmable and can be adjusted according to the testing requirements.

4.2.1 Time-in-state Dependence on Temperature

The time-in-state dependence on temperature was calibrated in Chapter 3 with the help of the time-in-state model (Eq 3.6). The time-in-state maximum at the given temperature can be predicted with the help of Eq 3.7 and the time-in-state

minimum can be predicted with the help of Eq 3.8. The fitted parameters of the time-in-state model are listed in Table 3.3.

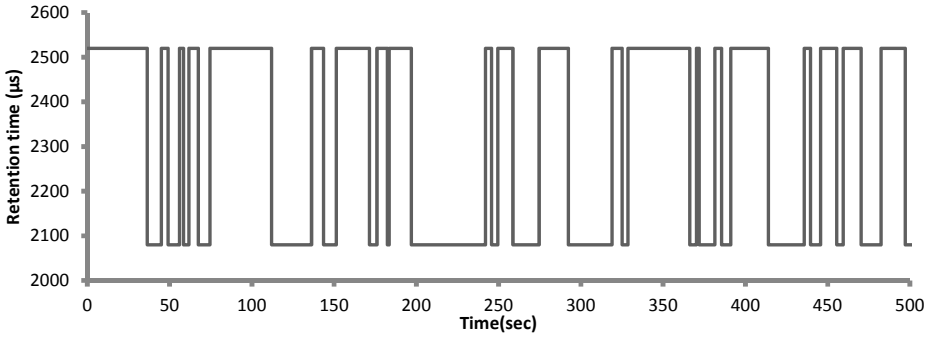
The VRT waveform of a VRT bit at different temperatures can be generated with the help of the RTN model of retention time described in Chapter 3. The parameters of the RTN model (maximum retention time, τ_{max}, minimum retention time, τ_{min}) were calculated at different temperatures to plot the VRT waveforms.

Figure 4.4 shows the waveform of a VRT bit at two different temperatures but the same voltage 1.2V. Figure 4.4a shows the VRT waveform with average time-in-state maximum (τ_{max}) 13sec and average time-in-state minimum (τ_{min}) 11sec at 115°C. The maximum retention time is 2520μs and the minimum retention time is 1080μs.

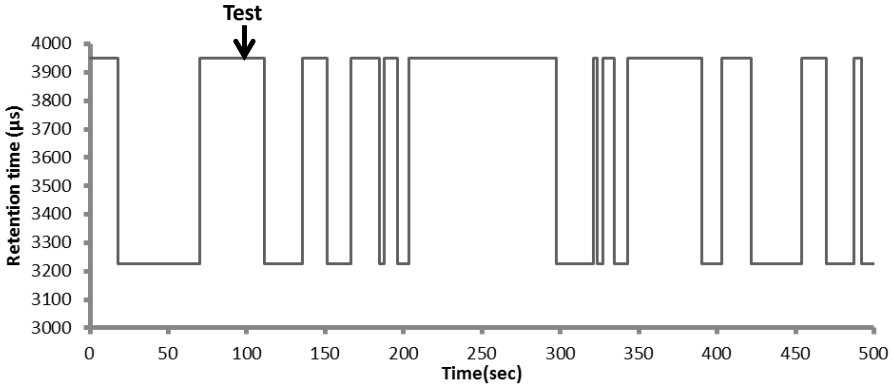
Figure 4.4b shows the VRT waveform with average time-in-state maximum (τ_{max}) 29sec and average time-in-state minimum (τ_{min}) 24sec at 105°C. The maximum retention time is 3950μs and the minimum retention time is 3220μs.

The maximum and minimum retention times of a VRT bit are decreasing with the rise in temperature. The rate of transition between the two retention states is increasing as well with the rising temperature. The higher transition rate decreases the average time-in-state (τ of a VRT bit).

Notice that the maximum and minimum retention times of the VRT bit in Figure 4.4 are also decreasing with the increasing temperature.



(a) VRT waveform at T= 115°C, τ_{max} = 13sec, τ_{min} = 11sec, Max. R.T = 2520μs, Min. R.T. = 2080μs



(b) VRT waveform at T= 105°C, τ_{max} = 29sec, τ_{min} = 24sec, Max. R.T = 3950μs, Min. R.T. = 3220μs

Figure 4.4: Synthetic VRT waveform at two different temperatures, Voltage =1.2V

4.2.2 Time-in-state Dependence on Voltage

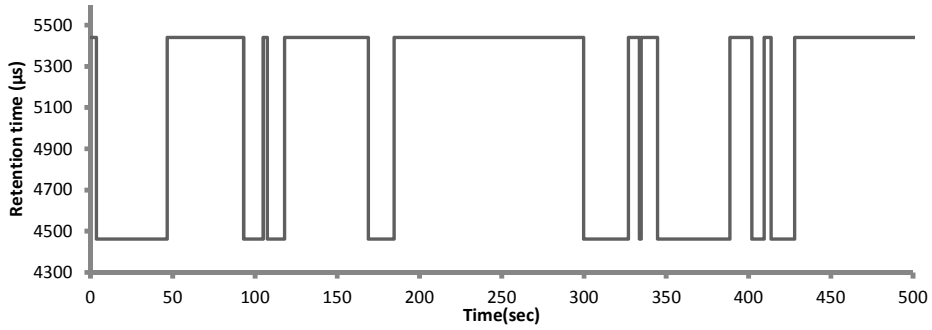
The time-in-state dependence on voltage was also calibrated in Chapter 3 with the help of the time-in-state model (Eq 3.6). The time-in-state maximum of a VRT bit at the given voltage can be predicted with Eq 3.7 and the time-in-state minimum can be predicted with Eq 3.8. The parameters of the time-in-state model are listed in Table 3.3.

The VRT waveforms of a VRT bit at different voltages were generated with the help of the RTN model of retention time described in Chapter 2. The parameters of the RTN model (maximum retention time, τ_{max}, minimum retention time, τ_{min}) were calculated to plot the VRT waveforms at different voltages.

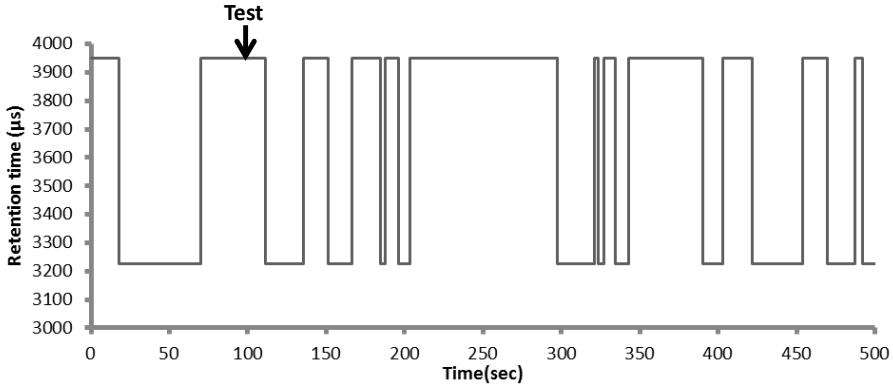
Figure 4.5 shows the waveform of a VRT bit at two different voltages but the same temperature 105°C. Figure 4.13a shows the VRT waveform with average time-in-state maximum (τ_{max}) 29sec and average time-in-state minimum (τ_{min}) 24sec at 1V. The maximum retention time is 5430μs and the minimum retention time is 4460μs.

Figure 4.13b shows the VRT waveform with average time-in-state maximum (τ_{max}) 29sec and average time-in-state minimum (τ_{min}) 24sec at 1.2V. The maximum retention time is 3950μs and the minimum retention time is 3220μs.

Figure 4.5 and Figure 4.4 suggest that the time-in-state of the VRT bit is a strong function of temperature and a weak function of voltage. Figure 4.5 and Figure 4.4 also suggest that the retention time of the VRT bit is a strong function of both temperature and voltage.



(a) VRT waveform at Voltage = 1V, τ_{max} = 29sec, τ_{min} = 24sec, Max. R.T = 5430μs, Min. R.T. = 4460μs



(b) VRT waveform at Voltage = 1.2V, τ_{max} = 29sec, τ_{min} = 24sec, Max. R.T = 3950μs, Min. R.T. = 3220μs

Figure 4.5: Synthetic VRT waveform at two different voltages, T=105°C

4.3 VRT Bits in a DRAM Array

4.3.1 Two Retention times

With the help of the PSU's model for miscorrelation in retention times, two retention times of 100 bits at the given environmental condition were generated synthetically. The two retention times of the 100 VRT bits at 105°C and 1.2V are plotted in Figure 4.6. The x-axis of Figure 4.6 is the 1st retention time in

micro-seconds and the y-axis is the 2nd retention time in micro-seconds. Each dot in the figure represents the 1st and 2nd measurements of the retention time of a VRT bit.

The VRT bit inside the red circle in Figure 4.6 has its own attributes like maximum retention time, minimum retention time, time-in-state maximum and time-in-state minimum. The Waveform of the VRT bit (inside red circle) is shown in Figure 4.4b. The VRT bit inside the blue circle has maximum retention time 5200 μs and minimum retention time 3200μs. Similarly, the other VRT bits in Figure 4.6 have their own attributes and VRT waveforms.

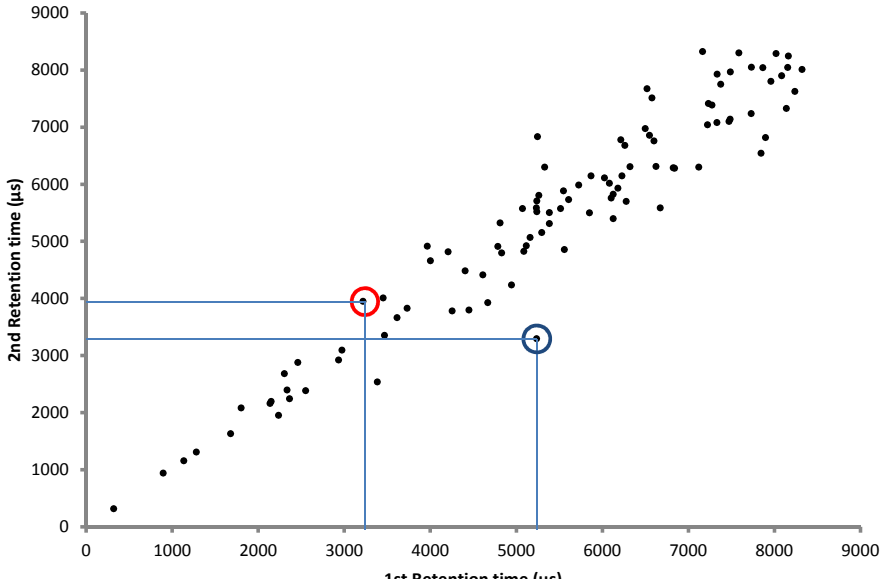


Figure 4.6: Two retention time measurements of 100 VRT bits at 105°C and 1.2V. The VRT bit inside the red circle represents the VRT bit in Figure 4.4b. The VRT bit inside the blue circle has maximum retention time 5200μs and minimum retention time 3200μs.

4.4 Test and Use

To represent the behavior of the VRT bits during test and use, the two retention time measurements of the 100 VRT bits at 105°C and 1.2V in Figure 4.6 are plotted in Figure 4.7 with the help of the test and use model. The x-axis is use and the y-axis is test.

Test is an instantaneous measurement of retention time. A VRT bit has maximum or minimum retention time at test and always minimum retention time at use (Figure 4.2).

The VRT bit in the red circle in Figure 4.7 is the same bit shown in Figure 4.4b and in Figure 4.7 inside the red circle. At test (the black arrows in Figure 4.1), the red circled VRT bit was in maximum retention state, so the retention time at test in Figure 4.7 is the maximum retention time and retention time at use is the minimum retention time. The VRT bit in the blue circle was also in maximum retention state during test, so maximum retention time was recorded during test. Similarly, the retention time of other VRT bits were measured at test with the help of their VRT waveforms. If the black arrow (Test 1 in Figure 4.4b) falls on the maximum retention state of the VRT waveform then the VRT bit has maximum retention time at test and if the black arrow (Test 2 in Figure 4.4b) falls at the minimum retention state of the VRT waveform then the VRT bit has minimum retention time at test (Eq 4.1). Notice that VRT bits have always minimum retention time at use (Eq 4.2).

$$Test = \begin{cases} \text{Max(1st R.T., 2nd R.T.)}, & \text{Test 1 (Figure 4.4b)} \\ \text{Min(1st R.T., 2nd R.T.)}, & \text{Test 2 (Figure 4.4b)} \end{cases} \quad (4.1)$$

$$Use = \text{Min(1st R.T., 2nd R.T.)}, \quad \text{Always} \quad (4.2)$$

In the transformation of retention time physics (Figure 4.6) to retention time engineering (Figure 4.7), the VRT bits which were below diagonal got shifted to either above diagonal or on the diagonal depending on the retention time at test. The VRT bits which have maximum retention time at test are above diagonal in Figure 4.7 and the VRT bits which have minimum retention time at test fall on the diagonal.

The time-in-state of the VRT bits remain unaffected of this transformation to test and use retention time. The RTN waveforms of the VRT bits also remain unchanged of this transformation.

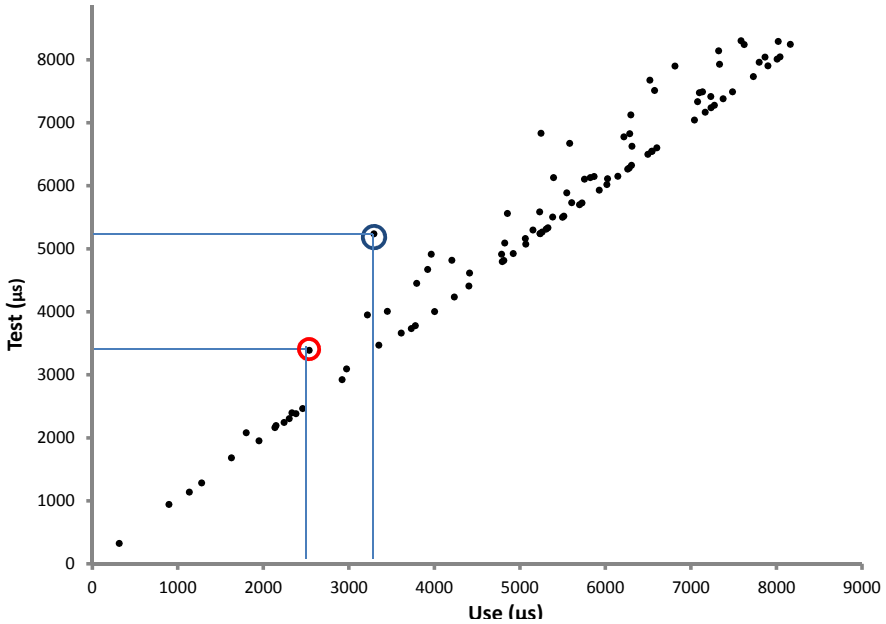


Figure 4.7: Transformation of retention time measurements to test and use model (Eq 4.1 and Eq 4.2) at 105°C and 1.2V. The VRT bit inside the red circle (Figure 4.4b) is at the same place after transformation and the VRT bit inside the blue circle got shifted.

4.4.1 Time-in-state

Time-in-states (time-in-state maximum and time-in-state minimum) were also calculated for the 100 VRT bits shown in Figure 4.6. Figure 4.8 shows the transition of VRT bits after different time intervals. The x-axis in Figure 4.8 is use in micro-seconds and the y-axis is test in micro-seconds.

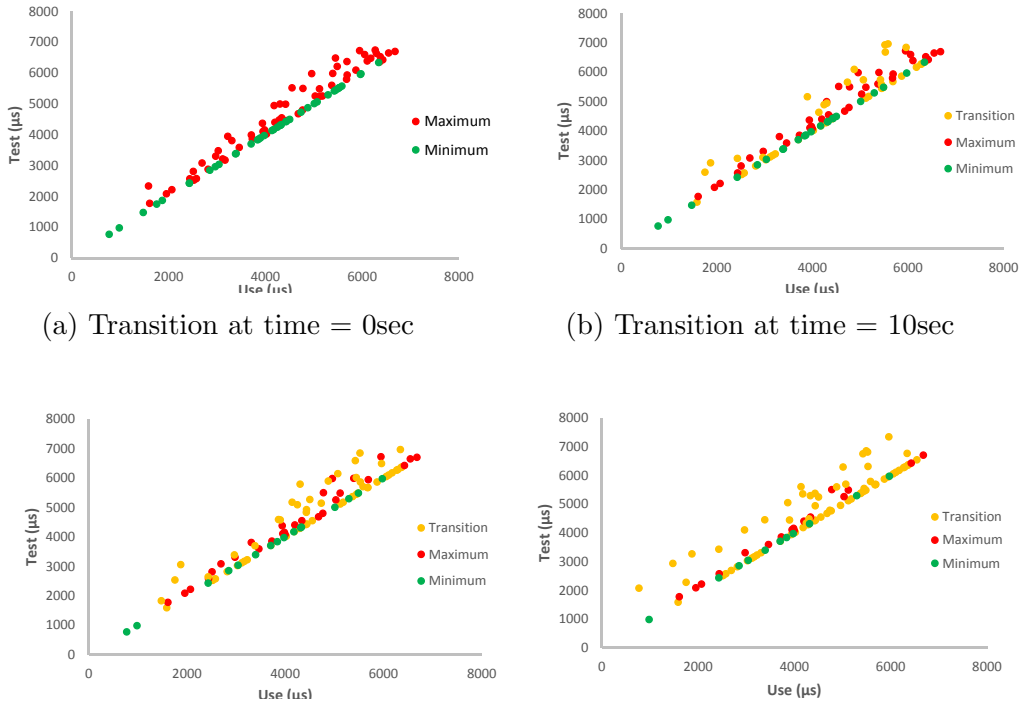
At t=0, a fraction of the 100 VRT bits were found in maximum retention state during test and the rest of the VRT bits were found in minimum retention state. The VRT bits found in the maximum retention state are shown with the red colored dots and the VRT bits found in the minimum retention state are shown with the green colored dots in Figure 4.8.

At t=10sec, the retention times of the VRT bits were measured again. In the duration of 10sec, some of the VRT bits have made the transition to minimum retention state from maximum retention state. Similarly, some of the VRT bits have made transition to maximum retention state from minimum retention state. Some of the VRT bits have made two transitions in the duration of 10sec and have come back to the same retention state. It is also possible that some VRT bits have made more than two transitions in the duration of 10sec.

The VRT bits, which were in maximum retention state at t=0sec, and after 10sec are still in maximum retention state, are shown with the red colored dots. The VRT bits, which were in minimum retention state, and after 10sec are still in minimum retention state, are shown with the green colored dots. The VRT bits which have made at least one transition are shown with orange colored dots.

The retention times of the VRT bits were measured again at t=20sec and at t=30sec. The VRT bits which are still in maximum retention state are represented with red colored dots and the VRT bits which are still in minimum retention state are represented with green colored dots. The VRT bits which have made at least one transition are shown with the orange colored dots.

In Figure 4.8, the number of VRT bits making the transition is increasing with time. At t=30sec, the majority of the VRT bits have made at least one transition and the minimum retention time of the VRT bits have been recorded at test.



(a) Transition at time = 0sec

(b) Transition at time = 10sec

(c) Transition at time = 20sec

(d) Transition at time = 30sec

Figure 4.8: Transition of the VRT bits at different time intervals. T=105°C and Voltage =1.2V

4.4.2 Time-in-state Dependence on Temperature

Time-in-state of a VRT bit is a strong function of temperature (Figure 4.4, Table 3.2(b)). The average time-in-state maximum and average time-in-state minimum gets shorter with rising temperature. The number of transitions between retention states at higher temperature increases due to shorter average time-in-state.

Figure 4.9 shows the transition of the VRT bits in 10sec at three different temperatures. The x-axis in Figure 4.9 is the use axis and the y-axis is the test axis. Each

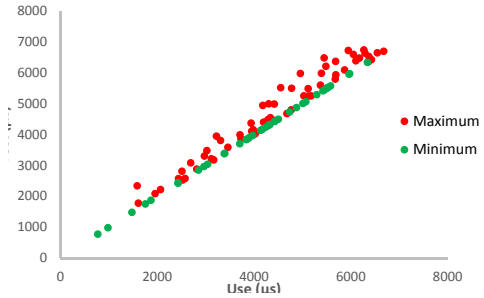
dot in Figure 4.9 represents a single VRT bit which has its own attributes. The VRT bits which remained in the maximum retention state after 10sec are represented with the red dots. The VRT bits which remained in the minimum retention state after 10sec are represented with the green dots. The VRT bits which made transition to other retention state in the duration of 10sec are represented with yellow dots.

Figure 4.9(a) shows the VRT bits at t=0sec and Figure 4.9(b) shows the transition of the VRT bits in the duration of 10sec at 105°C. In Figure 4.9(b), the yellow dots represent the VRT bits which have made the transition to other retention state in the duration of 10sec.

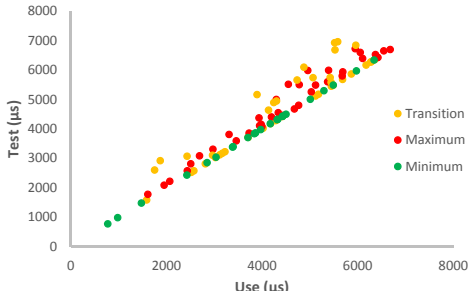
Figure 4.9(c) shows the VRT bits at t=0 and Figure 4.9(d) shows the transition of the VRT bits in the duration of 10sec at 105°C. Similarly, Figure 4.9(e) shows the VRT bits at t=0sec and Figure 4.9(f) shows the transition of the VRT bits in the duration of 10sec at 125°C.

The fraction of bits in maximum retention state at t=0sec is the same for all three temperatures. The number of VRT bits making transition (yellow dots) is increasing with the increasing temperature. Notice that the maximum and minimum retention time of the VRT bits are also decreasing with increasing temperature.

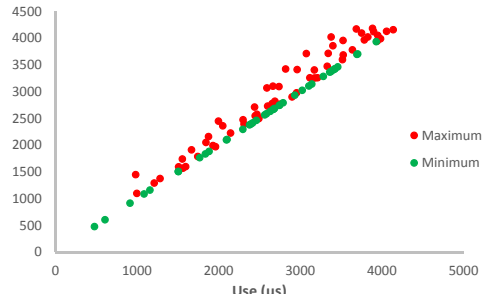
Since time-in-state of the VRT bits is highly sensitive to temperature, testing of VRT bits at different temperatures will give different results. The retention time of the VRT bits is also sensitive to temperature, so test limits should be adjusted accordingly during testing.



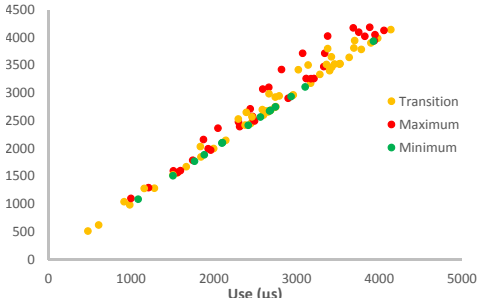
(a) t=0sec, T=105°C, τ_{max} = 29sec, τ_{min} =24sec



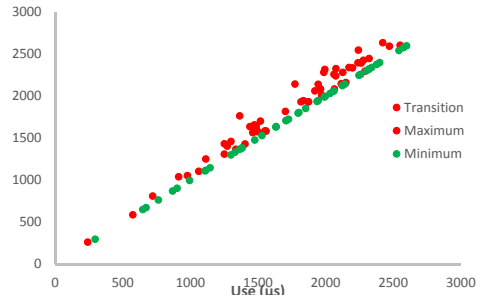
(b) t=10sec, T=105°C, τ_{max} = 29sec, τ_{min} =24sec



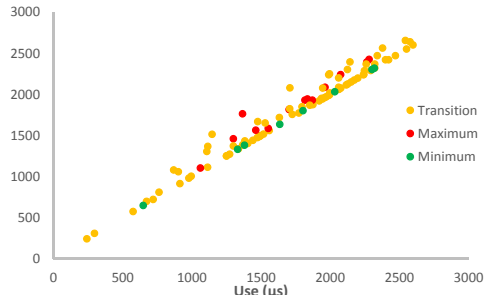
(c) t=0sec, T=115°C, τ_{max} = 13sec, τ_{min} =11sec



(d) t=10sec, T=115°C, τ_{max} = 13sec, τ_{min} =11sec



(e) t=0sec, T=125°C, τ_{max} = 6sec, τ_{min} = 5sec



(f) t=10sec, T=125°C, τ_{max} = 6sec, τ_{min} =5sec

Figure 4.9: Transitions of the VRT bits in 10sec at three different temperatures. Voltage = 1.2V

4.4.3 Time-in-state Dependence on Voltage

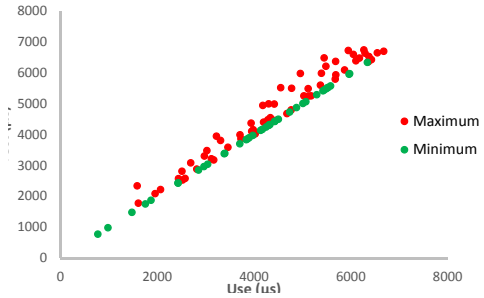
Time-in-state of a VRT bit is a weak function of voltage (Figure 4.5, Table 3.2 (a)). The average time-in-state maximum and average time-in-state minimum do not change much with rising voltage. The number of transitions between retention states at different voltages remains unaffected.

Figure 4.10 shows the transition of VRT bits in 10sec at three different voltages. The x-axis in Figure 4.10 is the retention time during use and the y-axis is the retention time observed during test. Each dot in Figure 4.10 represents a single VRT bit which has its own attributes. The VRT bits which remained in the maximum retention state after 10sec are represented with the red dots. The VRT bits which remained in the minimum retention state after 10sec are represented with the green dots. The VRT bits which made transition to other retention state in the duration of 10sec are represented with yellow dots.

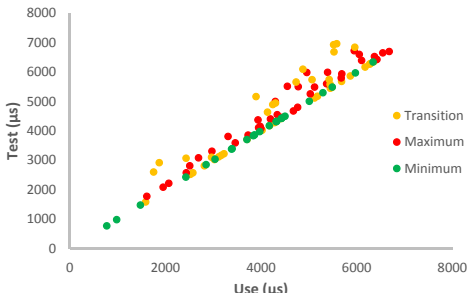
Figure 4.10(a) shows the VRT bits at t=0sec and Figure 4.10(b) shows the transition of VRT bits in the duration of 10sec at 0.85V. In Figure 4.9(b), the yellow dots represent the VRT bits which have made transitions to other retention state in the duration of 10sec.

Figure 4.10(c) shows the VRT bits at t=0 and Figure 4.10(d) shows the transition of the VRT bits in the duration of 10sec at 1V. Similarly, Figure 4.10(e) shows the VRT bits at t=0sec and Figure 4.10(f) shows the transition of VRT bits in the duration of 10sec at 1.2V.

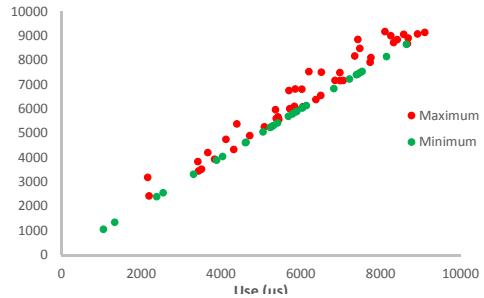
The fraction of bits in maximum retention state at t=0sec is the same for all three voltages. Since the time-in-state is a weak function of voltage, the number of VRT bits making transition (yellow dots) is not changing for different voltages. Notice that the maximum and minimum retention times of VRT bits are decreasing with increasing voltage.



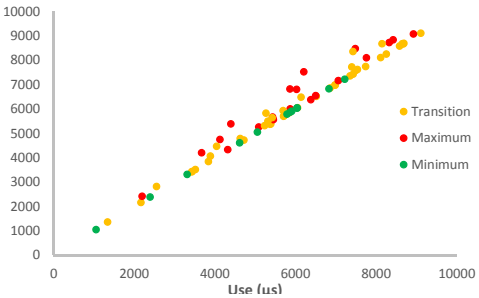
(a) t=0sec, Voltage =0.85V, τ_{max} = 29sec, τ_{min} = 24sec



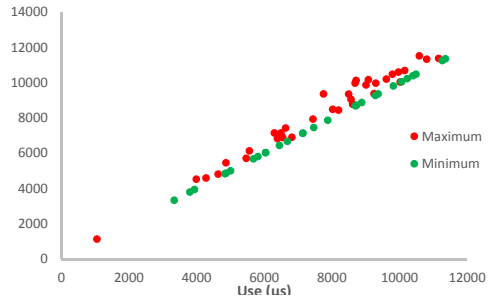
(b) t=10sec, Voltage=0.85V, τ_{max} = 29sec, τ_{min} = 24sec



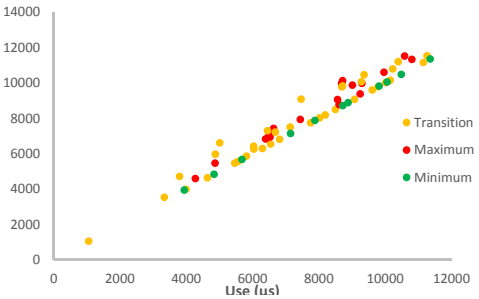
(c) t=0sec, Voltage = 1V, τ_{max} = 29sec, τ_{min} = 24sec



(d) t=10sec, Voltage=1V, τ_{max} = 29sec, τ_{min} = 24sec



(e) t=0sec, Voltage = 1.2V, τ_{max} = 29sec, τ_{min} = 24sec



(f) t=10sec, Voltage = 1.2V, τ_{max} = 29sec, τ_{min} = 24sec

Figure 4.10: Transition of the VRT bits in 10sec at three different voltages. T = 105°C

4.5 Number of Transitions

The rate of transition of VRT bits from maximum to minimum retention state and minimum to maximum retention state increases with time and temperature (Figure 4.8, Figure 4.9). Table 4.1 lists the count of the 100 VRT bits which made transition in the given time intervals and temperatures.

The x-axis of Table 4.1 is temperature and the y-axis is the test runtime. Each cell in the table represents the number of VRT bits which made transition in the given test runtime and temperature.

The rate of transition of the VRT bits can be increased by increasing the temperature or by allowing more time to make transition.

Time	**Temp 105°C**	**Temp 115°C**	**Temp 125°C**	**Temp 130°C**
10 sec	29	52	79	89
20 sec	50	77	95	99
30 sec	64	89	99	100

Table 4.1: Number of VRT bits making transition in the given test runtime and temperature. 105°C (τ_{max} = 29sec, τ_{min} = 24sec), 115°C (τ_{max} = 13sec, τ_{min} = 11sec), 125°C (τ_{max} = 6sec, τ_{min} = 5sec)

4.6 Survey of Existing DRAM Test Programs

4.6.1 March Tests

March tests are common type of tests for DRAMs and SRAMs. Different March tests have been designed to address a set of fault models effectively. Table 4.2 shows the examples of March tests for different fault models of a DRAM. For example, MATS+ March test covers SAF and AF. A test program based on the March tests gives higher memory fault coverage and minimal test runtime.

Algorithm	Fault coverage								O.C.
	SAF	AF	TF	CFin	CFid	CFdyn	SCF	Linked faults	
MATS	All	Some							4.n
MARCHC-	All	All	All	All	All	All	All		10.n
MARCHY	All	All	All	All				All TF linked with CFins	8.n

Table 4.2: Fault models covered by different March Test Algorithms [6].

The right column in the table is the operation count of each March test. Operation count is the sum of the total number of reads and writes operations in a March test. Description of read and write operations of different March tests are given in Table 4.3. Table 4.4 describes the notations of March test.

Algorithm	Description
MATS	$\{\Updownarrow (w0); \Updownarrow (r0, w1); \Updownarrow (r1)\}$
MARCHC-	$\{\Updownarrow (w0); \Uparrow (r0, w1); \Uparrow (r1, w0); \Downarrow (r0, w1); \Downarrow (r1, w0); \Updownarrow (r0)\}$
MARCHA	$\{\Updownarrow (w0); \Uparrow (r0, w1, w0, w1); \Uparrow (r1, w0, w1); \Downarrow (r1, w0, w1, w0); \Downarrow (r0, w1, w0)\}$

Table 4.3: Description of different March Test Algorithms [6].

Notation	Instruction
$\Uparrow$	Increasing address order
$\Downarrow$	Decreasing address order
$\Updownarrow$	Address order is either increasing or decreasing
w0	write logic 0 to all the bits in DRAM array
w1	write logic 1 to all the bits in DRAM array
r0	read logic 0 from all the bits in DRAM array
r1	read logic 1 from all the bits in DRAM array

Table 4.4: Notations of March Test

The March test is made up of March elements. March elements in the March test are separated by semicolons. One element is applied to each DRAM bit in DRAM array before moving to the next element of the March test.

MATS+ March test is chosen for example, to explain the notations of the March tests. MATS+ March test has three March elements and five operation counts since the sum of total reads and writes is five in the test pattern. The first March

element ⇳(w0) specifies to fill out the memory subarray with logic 0 in either increasing or decreasing order of row lines. The second March element ⇑(r0,w1) specifies to read the expected logic 0 and then write logic 1 to the row 0 and the process is repeated in increasing row order till row 255. The third March element ⇓(r1,w0) asks to read expected logic 1 and write logic 0 to the row 255 and the process is repeated in decreasing row order till row 0. Fail is detected if read logic from a bit is not the same as expected. In MATS+ March test, while r0 operation, if the read logic from a bit is not logic 0, which is expected, then fail is detected. Similarly, while r1 operation, if read logic from DRAM bits is not logic 1, which is expected, then fail is detected and a DRAM bit is marked as a defective bit.

4.6.2 IFA-9 and IFA-13 March Test

To address the data retention faults in DRAMs, the IFA-9 and IFA-13 March tests were proposed by Dekker [7]. Fault models covered by the IFA-9 and IFA-13 March tests are listed in Table 4.5. In these March tests a new element 'Delay' was added to detect any data retention fault. Delay was inserted between write and read instructions. Table 4.6 shows the description of the IFA-9 and IFA-13 March tests. To detect the data retention fault, first, all the bits in a DRAM array are written with logic 1 (charge on capacitor) and then a certain amount of delay is applied before the data is read from the bits. All the refresh mechanisms are turned off while delay is applied.

In IFA March tests, layout of DRAM chips is analyzed using the inductive fault analysis method. The IFA-9 March test has operation count of 12.n + delays. The IFA-13 March test has operation count of 16.n + delays. The IFA-9 March test

and the IFA-13 March test have the same number of March elements. The IFA-13 March test provides fault coverage for SOF as well.

Algorithm	Actual physical defect fault coverage							O.C.
	SAF	TF	AF	SOF	SCF	CFid	DRF	
IFA-9	All	All	All		All	All	All	12.n + Delays
IFA-13	All	All	All	All	All	All	All	16.n + Delays

Table 4.5: Fault models covered by IFA-9 and IFA-13 March test [7].

Algorithm	Description
IFA-9	$\{\Uparrow (w0); \Uparrow (r0, w1); \Uparrow (r1, w0); \Downarrow (r0, w1); \Downarrow (r1, w0); Delay; \Uparrow (r0, w1); Delay; \Uparrow (r1)\}$
IFA-13	$\{\Uparrow (w0); \Uparrow (r0, w1, r1); \Uparrow (r1, w0, r0); \Downarrow (r0, w1, r1); \Downarrow (r1, w0, r0); Delay; \Uparrow (r0, w1); Delay; \Uparrow (r1)\}$

Table 4.6: Description of IFA-9 and IFA-13 March test [7].

None of the March tests understands the concept of time-in-state. There is a need of a special test pattern which stresses the concept of time-in-state to screen the VRT bits effectively during manufacturing test.

4.7 Proposed VRT Test Algorithm

The following modifications have been done to the MARCHC- test (Table 4.3):

- More than one write-read is required to detect the VRT behavior, so test pattern would have repeated write-read operations.

- Capacitors of DRAM bits should always be charged to track the change in retention time.
- Address order must be preserved during entire test runtime. Preserving the address order ensures the same time interval between write and read for all bits.
- During the VRT test, the refresh mechanism of the DRAM could be on or off depending on the number of rows in the array. If the specified refresh cycle time of the DRAM is shorter than the time to write all rows of the DRAM then the refresh mechanism should be on.
- The runtime of the VRT test pattern should be long enough to allow the VRT bits to change their retention state to achieve the desired fault coverage at a given condition.
- The VRT test pattern should be run at the end of all other memory tests. Running the VRT pattern in the end ensures that fails are retention fails.

The proposed VRT algorithm is shown in the Figure 4.11. In the test algorithm, K is the number of pairs of write-read. The number K depends on the time required to achieve the desired fault coverage at a given temperature.

$$\underbrace{\{\Uparrow (w1); \Uparrow (r1); \Uparrow (w1); \Uparrow (r1); \ldots\ldots\ldots \Uparrow (w1); \Uparrow (r1)\}}_{K}$$

Figure 4.11: VRT Test Algorithm

The VRT bits which passed at all reads (read data matches with expected data) during VRT testing are considered passing bits and the VRT bits which failed even at one read during VRT testing are considered failing bits. The VRT bits making transition to the maximum retention state do not matter since it has already been recorded as fail in the minimum retention state. Table 4.7 shows an example of screening the VRT bits with the help of the proposed VRT test algorithm. The columns in the table are the bit addresses and the rows in the table are the reads during VRT testing. Tick in Table 4.7 for Bit number 1 and read number 5 represents a fail. Similarly, each tick in the Table 4.7 represent the fail during read from a bit.

In Table 3.7, the DRAM bits which pass at all reads during test are considered passing bits. There are three possible cases for always passing bits. First, the bit has single retention time (SRT bits) which is longer than test limit. Second, the bit is a VRT bit and has minimum retention time longer than test limit. Third, the bit is a VRT bit and has minimum retention time less than test limit but never made transition to minimum retention state during test.

Reads⟶

← Bit Address	1	2	3	4	5	6	7	8	9	10	11	12	Bit
0													PASS
1					√	√	√	√	√				FAIL
2													PASS
3	√	√	√	√	√	√	√	√	√	√	√	√	FAIL
4													PASS
5								√	√	√	√	√	FAIL
6	√	√	√	√	√	√	√	√	√	√	√	√	FAIL
7													PASS
N	√	√	√	√	√	√							FAIL

Table 4.7: Example: Screening of VRT bits with VRT test algorithm. Ticks represent the fails during read from DRAM bit.

In Table 3.7, the DRAM bits which pass at some reads and fail at other reads are bad VRT bits and considered failing bits. These VRT bits pass reads in maximum retention state but fail when they make transition to minimum retention state.

In Table 3.7, the DRAM bits which fail all reads during test are also considered failing bits. There are three possible cases for always failing bits. First, the DRAM bit has single retention time which is smaller than the test limit. Second, the bit is a VRT bit and has maximum retention time smaller than the test limit. Third, the bit is a VRT bit and has maximum retention time longer than the test limit but never made transition to maximum retention state during test.

The length of the introduced VRT test algorithm should be long enough to allow VRT bits to make transitions during test and should have enough write-read pairs (K) to detect the transition made by the VRT bits.

4.8 Determination of the VRT Test Length (K)

The runtime of the VRT test algorithm is determined based on the average time-in-state at the given temperature. The number of refresh signals is calculated for the target test runtime. The refresh cycle time at test can be programmed according to the test set-point.

The time required to perform each write-read operation depends on the size of the DRAM array. The value of K is chosen to accommodate the number of refresh signals during targeted test runtime at the given temperature.

4.9 Figures-of-merit

By introducing the concept of the use set-point (use refresh cycle time) and the test set-point (test refresh cycle time), the VRT bits can be characterized into four categories.

Figure 4.12 shows a scatter plot of VRT bits at test and use. The horizontal line (blue) on the Figure 4.12 is the test-set. The vertical line (green) on the Figure 4.12 is the use-set.

During use, the VRT bits having retention time longer than the use set-point are passing use and the VRT bits having retention time smaller than the use set-point

are failing use(Figure 4.12).

During test, the VRT bits having retention time longer than the test set-point are passing test and the VRT bits having retention time smaller than the test set-point are failing test (Figure 4.12).

The VRT bits which fail both in use and test are termed as Fails. The VRT bits which fail in use and pass in test are termed as Overkills (OK). The VRT bits which pass both in use and test are termed as Good bits. The VRT bits which fail in use but pass in test are termed as Test Escapes (TE)(Figure 4.12).

Figure 4.12 shows the characterization of 100 VRT bits into four quadrants by introducing test set-point and use set-point. The test set-point was chosen to be 1200μs and use set-point 1000μs.

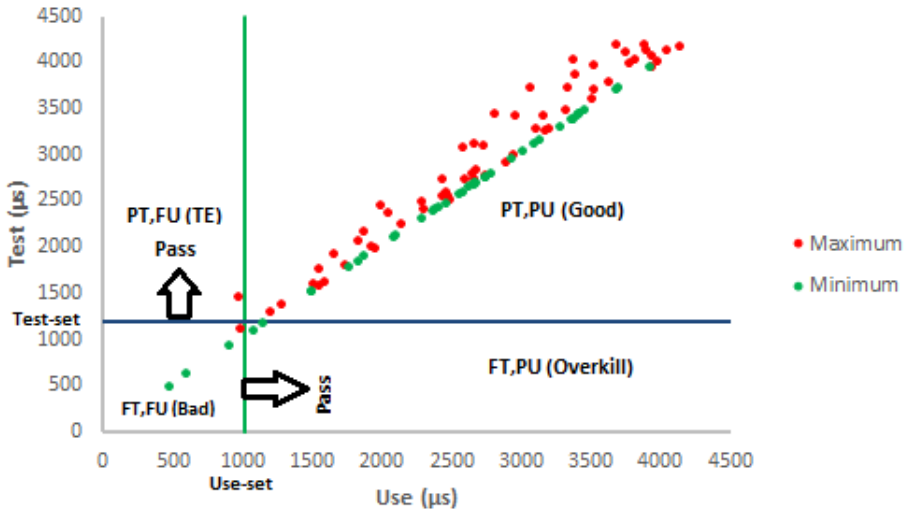


Figure 4.12: Figures-of-merit of a test program. Test-set (Blue horizontal line) is 1200μs and use-set (Green vertical line) is 1000μs

The count of Pass, Fail, Overkill (OK) and Test Escape (TE) in Figure 4.12 is listed in Table 4.8.

Pass	Fail	Overkill	Test Escape
93	4	2	1

Table 4.8: Count of Pass, Fail, Overkill and Test Escape of 100 VRT bits in Figure 4.12.Test set-point= 1200μs, use set-point = 1000μs

Similarly, 4000 VRT bits were generated 10 times with the PSU's retention time marginal distribution model at T=125°C and voltage =1.2V. Table 4.9 lists the average count of Pass, Fail, Overkill (OK) and Test Escape (TE) for the given test set-point (1200μs) and use set-point (1000μs) at 125°C.

FOM	Avg. Count	Std. Dev.
Pass	2464	42
Fail	1146	24
Overkill	386	12
Test Escape	4	1

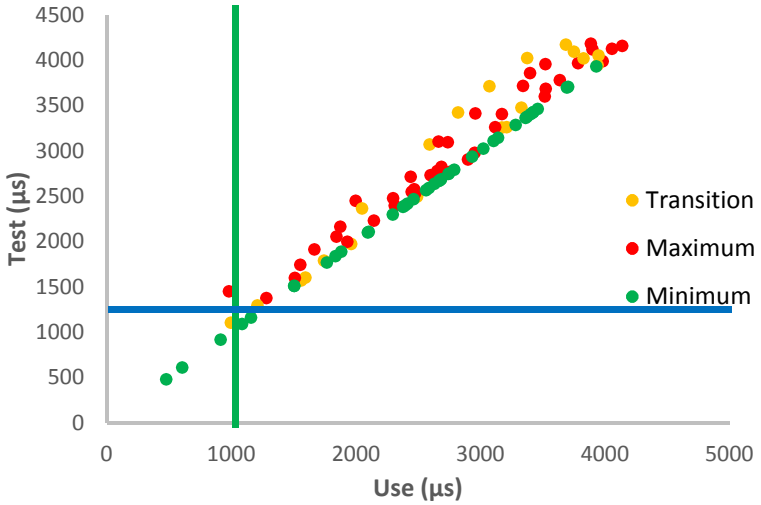
Table 4.9: Count of Pass, Fail, Overkill and Test Escape. Average of 10 samples of 4000 VRT bits were taken. T=125°C, voltage =1.2V. Test set-point =1200μs, use set-point = 1000μs

4.10 Revised Figures-of-merit after Including Time-in-state

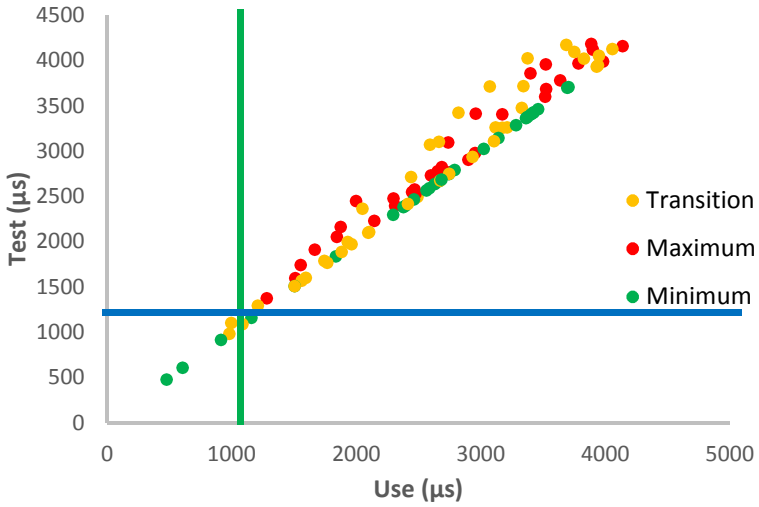
Figure 4.12 shows the retention states of VRT bits after one write and read (K=1). After K=1, the number of bits in each quadrant are counted (Good, Bad, Overkill and TE).

The VRT bits keep making transition from one retention state to other retention state. At the second read (K=2), some of the VRT bits have made transitions to other retention state. The number of bits are counted again in the four quadrants. The yellow dots are those VRT bits which have made transitions to other retention state(Figure 4.13(a)).

At the third read (K=3), more VRT bits have made transitions to other retention state and now represented with yellow dots. The number of dots in each quadrant is counted again (Figure 4.13(b)).



(a) Retention states of the VRT bits after K=2



(b) Retention state of VRT bits after K=3

Figure 4.13: Retention state of VRT bits after different number of reads.

Multiple reads (k) allow VRT bits to change their retention state. The reads are continued till most of the dots on the plot become yellow. The maximum and minimum retention time of the yellow dots get recorded. If the minimum retention state of the VRT bit is greater than test limit then it is screened as passing bits; otherwise, fail.

Figure 4.14 shows the transition of VRT bits from maximum to minimum retention state. The Good bits which make transition to minimum retention state during test are now either Overkill (rectangle 2 to rectangle 5 in Figure 4.14) or Pass (rectangle 3 in Figure 4.14) at test. If the minimum retention of a Good bit is less than the test set-point then the Good bit is overkill after transition to minimum retention state.

The TE bit which makes transition to minimum retention state is fail now (rectangle 1 to rectangle 4 in Figure 4.14).

The VRT bits making transition from minimum retention state to maximum retention state do not matter because Test Escape and Overkill come from only going to minimum retention state from maximum retention state.

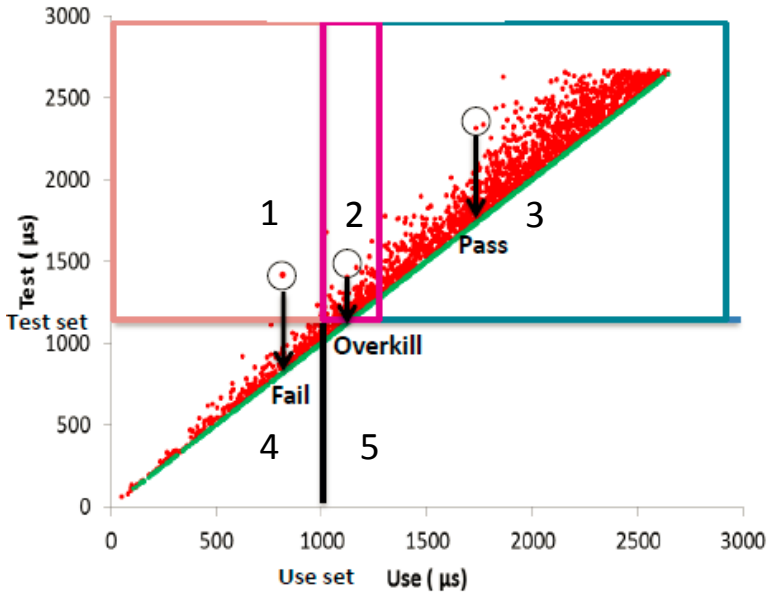


Figure 4.14: Transition of the VRT bits during test. Sample size: 4000. T=125°C, Voltage =1.2V

The proposed VRT test algorithm screens the VRT bits in rectangle 1 and 2 of Figure 4.14. The VRT bits in rectangle 3 of Figure 4.14 will always pass the test because even the minimum retention time of these VRT bits is greater than the test set-point. The VRT bits in rectangle 4 will always fail the VRT test because even the maximum retention time of these VRT bits is smaller than the test set-point.

By using the Monte-carlo method, the DPM rate of Overkill and TE was calculated for the VRT test algorithm of 5sec runtime. The DPM was calculated for the sample size of 4000 VRT bits. The process was repeated 10 times to get the average DPM of Overkill and Test Escape. Table 4.10 lists the DPM of Overkill

and Test Escape for test set-point =1200μs and use set-point = 1000μs at 125°C and 1.2V.

FOM	Avg. DPM Rate
Overkill	3 DPM
Test Escape	0.8 DPM

Table 4.10: Average DPM of Overkill and TE for 10 samples of 4000 bits.

Table 4.11 shows the number of Overkills (OK), and Test Escapes (TE) at 4 different test set-points, 4 different test runtimes and 3 different temperatures. The size of the DRAM used to calculate the figures-of-merit was 1.2 Gb and the refresh cycle time was 2000μs (Chapter 3.4).

The retention time of the VRT bits decreases with change in voltage or temperature. The settings of test also changes with the change in temperature. The test set-point and use set-point are adjusted accordingly with the temperature. The equivalent test set-point and use set-point can be calculated by Eq 4.3.

$$r_{115°C} = \left(\frac{\alpha_{115°C}}{\alpha_{105°C}}\right) * r_{105°C} \tag{4.3}$$

where $r_{105°C}$ = retention time at °C , $\alpha_{115°C}$ = scale parameter at 115°C and $\alpha_{105°C}$ = scale parameter at 105°C.

Table 4.11(a) shows the count of OK and TE at 125°C, Table 4.11(b) shows the count of OK and TE at 115°C and Table 4.11(c) shows the count of OK and TE at 105°C. The test set-points at 105°C in Table 4.11(c) are 2000μs, 2100μs, 2200μs

and 2400μs. The test set-points in Table 4.11 (a) at 125°C and Table 4.11(b) at 115°C are equivalent to 105°C. For example, the test set-point 789μs at 125°C is equivalent to the test set-point 2000μs at 105°C.

In Table 4.11, TE is improving with increasing the test runtime. The TE's bits are also decreasing with increasing test set-point. The TE can be significantly improved by increasing the temperature. Improvement in TE always comes at the cost of higher OK.

Test runtime⟶

← Test set-point(μs)	5sec		10sec		20sec		30sec		
	OK	TE	OK	TE	OK	TE	OK	TE	
789μs	0	116	0	53	0	11	0	2	T=125°C
829μs	359	58	395	26	419	5	424	1	
868μs	778	18	835	8	872	2	879	2	
947μs	1521	9	1577	4	1614	1	4098	0	

(a) 125°C

Test runtime⟶

← Test set-point(μs)	5sec		10sec		20sec		30sec		
	OK	TE	OK	TE	OK	TE	OK	TE	
1250μs	0	149	0	102	0	48	0	22	T=115°C
1312μs	258	108	286	74	318	35	333	16	
1375μs	557	54	612	37	676	17	706	8	
1500μs	1759	7	1838	5	1930	2	1972	1	

(b) 115°C

Test runtime⟶

← Test set-point(μs)	5sec		10sec		20sec		30sec		
	OK	TE	OK	TE	OK	TE	OK	TE	
2000μs	0	218	0	183	0	129	0	91	T=105°C
2100μs	206	116	229	98	265	69	290	49	
2200μs	566	37	604	31	662	22	703	15	
2400μs	1463	13	1511	11	1585	8	1637	6	

(c) 105°C

Table 4.11: Test Escape vs Overkill as a function of test set-point, test runtime and temperature.

Chapter 5

Conclusions and Recommendations

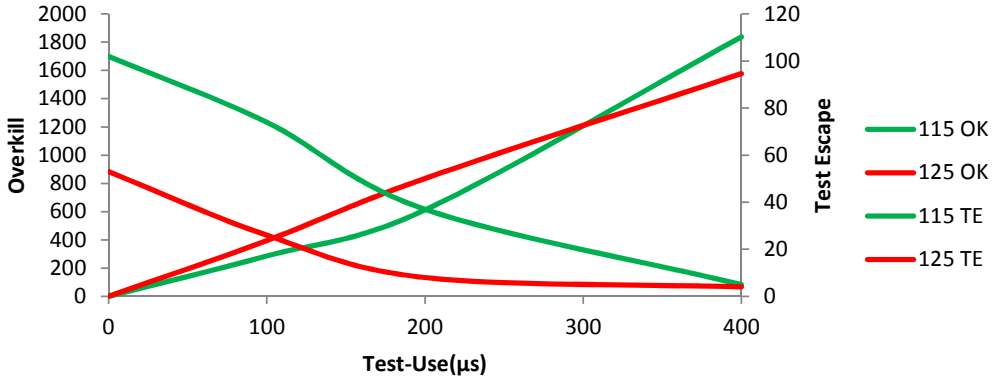
Time-in-state of the VRT bits has been studied to identify the bad VRT bits during the manufacturing test. The VRT defect is particularly sensitive to the difference between test and use conditions. The retention times of the VRT bits are measured instantaneously during test. The test can capture either the minimum or the maximum retention time of the VRT bits. In use, the retention times of the VRT bits are measured millions of times during the lifetime of the DRAM bit. The minimum retention time of the VRT bits is always recorded in use. The introduced VRT test algorithm simulates the use condition to screen the failing bits during the manufacturing test.

The introduced test algorithm tries to record both the maximum and minimum retention times of the VRT bits by taking the repeated measurements like in use. The time to see both the maximum and minimum retention time during test decided the test runtime. The retention times of the VRT bits are sensitive to both voltage and temperature. The settings of test change with voltage and temperature. A weibull distribution was found to be a good fit to the measured retention times. The temperature and voltage dependence were determined for the Weibull distribution parameters. The retention time of the VRT bits decreases with voltage and temperature, so test set-point and use set-point are changed accordingly [4].

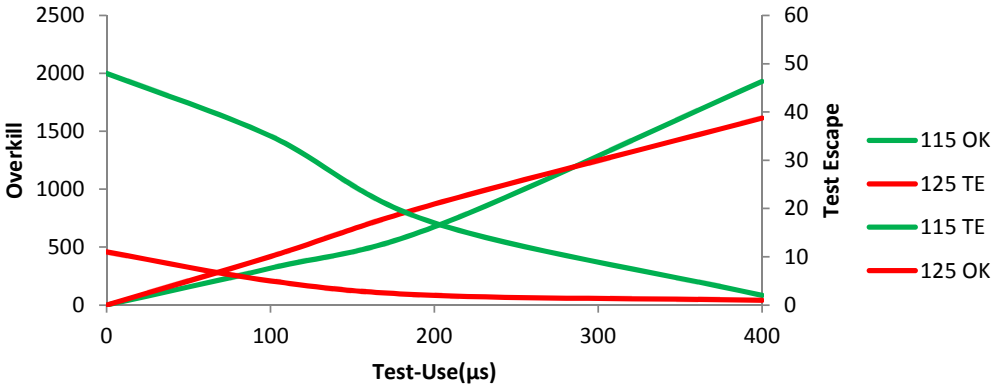
The test runtime is an important factor which decides the cost of test and DRAM. The exponential distribution was found to be a good fit for the time-in-state data

obtained from the Kim paper [5]. The time-in-state of the VRT bits was insensitive to voltage [Table 3.2a]. The change in voltage did not help to reduce the test runtime to see both the maximum and minimum retention times of the VRT bits. The time-in-state was found to be a very strong function of temperature. The time-in-state of the VRT bits decreases with the increasing temperature [Table 3.2b]. The smaller time-in-state of VRT bits at higher temperature allows the test to see both the minimum and maximum retention time of the VRT bits in smaller runtime. The test escape of the VRT bits improves by more than 50% by increasing the temperature from 105°C to 125°C for the test set-point 2400μs, use set-point 2200μs and test runtime 10sec [4].

There is a tradeoff between overkill, test escape and test runtime. In the manufacturers point of view, the overkill can be decreased by decreasing the difference between test and use set-points. The overkill can also be decreased by running the test at a lower temperature with a smaller test runtime. In the users point of view, the test escape can be decreased by increasing the difference between test set-point and use set-point. The test escapes are also significantly decreased by running test at a higher temperature with a longer test runtime. Figure 5.1 shows the tradeoff curve between the Test escapes and Overkills. The x-axis is the difference between the test set-point and use set-point. The primary y-axis is the overkill and the secondary y-axis is the Test Escape. The recommended use set-point is 2000μs, test set-point 2075μs, temperature 125°C and test runtime 20sec for optimal results.



(a) TE vs OK for test runtime 10sec



(b) TE vs OK for test runtime 20sec

Figure 5.1: TE vs OK for test runtime 10sec and 20sec at temperatures 115°C and 125°C

The test escape improves with the increase in temperature but it should not go over 20-30°C than the specified operational temperature in the datasheet. The excessive temperature can burn the chip or might introduce some new fail mechanism.

The introduced VRT test algorithm should be run at the end of the test flow after running the other DRAM tests. Running the VRT test at the end confirms that the failed bits are failing VRT bits not the other functional fails.

This thesis extended the previous study of DRAM retention time at PSU's ECE ICDT laboratory [4] by adding the time-in-state of the VRT bit. A model of time-in-state of the VRT bits was presented as a function of environmental conditions such as temperature and voltage. This thesis modeled the VRT bit by the integrating of Weibull model for marginal distribution of retention time, Copula model for retention time mis(correlation) and the presented time-in-state model. A DRAM array of 1Gb was assembled from the combination of marginal distribution of retention time model, Copula model of miscorrelation in retention times and time-in-state model. The assembled DRAM array was used to develop and evaluate a test algorithm for the VRT detection at temperature and voltage set points. The time-in-state of VRT bit was found to be a strong function of temperature and weak function of voltage. Temperature can be used to increase the VRT fault coverage.

References

[1] N. H. Weste and D. M. Harris, "Cmos vlsi design." Pearson/Addison Wesley, 2005, ch. 12, p. 523.

[2] T. Umeda, K. Okonogi, K. Ohyu, S. Tsukada, K. Hamada, S. Fujieda, and Y. Mochizuki, "Single silicon vacancy-oxygen complex defect and variable retention time phenomenon in dynamic random access memories," *Applied physics letters*, vol. 88, no. 25, pp. 253 504–253 504, 2006.

[3] P. Restle, J. Park, and B. Lloyd, "Dram variable retention time," in *Electron Devices Meeting, 1992. Technical Digest., International.* IEEE, 1992, pp. 807–810.

[4] C. Shirley and W. Daasch, "Copula models of correlation: A dram case study," *Computers, IEEE Transactions on*, vol. PP, no. 99, pp. 1–1, 2013 DOI:10.1109/TC.2013.129. [Online]. Available: http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6532289

[5] H. Kim, B. Oh, Y. Son, K. Kim, S. Cha, J. Jeong, S. Hong, and H. Shin, "Characterization of the variable retention time in dynamic random access memory," *Electron Devices, IEEE Transactions on*, vol. 58, no. 9, pp. 2952–2958, 2011.

[6] M. Bushnell and V. D. Agrawal, "Essentials of electronic testing for digital, memory, and mixed-signal vlsi circuits." Springer, 2000, ch. 9.

[7] R. Dekker, F. Beenker, and L. Thijssen, "Fault modeling and test algorithm development for static random access memories," in *Test Conference, 1988. Proceedings. New Frontiers in Testing, International.* IEEE, 1988, pp. 343–352.

[8] D. Yaney, C. Lu, R. Kohler, M. Kelly, and J. Nelson, "A meta-stable leakage phenomenon in dram charge storagevariable hold time," in *Electron Devices Meeting, 1987 International*, vol. 33. IEEE, 1987, pp. 336–339.

[9] http://www.itrs.net/Links/2007ITRS/IRCPosition.html [Online; accessed 07-Oct-2014].

[10] J. Yu and K. Aflatooni, "Leakage current in dram memory cell," in *University/Government/Industry Microelectronics Symposium, 2006 16th Biennial.* IEEE, 2006, pp. 191–194.