A New Design-for-Test Technique for SRAM Core-Cell Stability Faults

Alexandre Ney, Luigi Dilillo, Patrick Girard, Serge Pravossoudovitch, Arnaud Virazel, Magali Bastian Hage-Hassan, Vincent Gouin

A New Design-for-Test Technique for SRAM Core-Cell Stability Faults

A. Ney L. Dilillo P. Girard
S. Pravossoudovitch A. Virazel
LIRMM – University of Montpellier /CNRS
161, rue Ada – 34392 Montpellier Cedex 5, France
Email: {ney, girard, pravo, virazel}@lirmm.fr
URL: http://www.lirmm.fr/~w3mic

M. Bastian V. Gouin
Infineon Technologies France
2600, route des Crêtes – 06560 Sophia-Antipolis, France
Email: {magali.bastian, vincent.gouin}@infineon.com
URL: http://www.infineon.com

***Abstract*—Core-cell stability represents the ability of the core-cell to keep the stored data. With the rapid development of semiconductor memories, their test is becoming a major concern in VDSM technologies. It provides information about the SRAM design reliability, and its effectiveness is therefore mandatory for safety applications. Existing core-cell stability Design-for-Test (DfT) techniques consist in controlling the voltage levels of bit lines to apply a weak write stress on the core-cell under test. If the core-cell is weak, the weak write stress induces the faulty swap of the core-cell. However, these solutions are costly in terms of area and test application time, and generally require modifications of critical parts of the SRAM (core-cell array and/or the structure generating the internal auto-timing). In this paper, we present a new DfT technique for stability fault detection. It consists in modulating the word line activation in order to perform an adjustable weak write stress on the targeted core-cell for stability fault detection. Compared to existing DfT solutions, the proposed technique offers many advantages: programmability, low area overhead, low test application time. Moreover, it does not require any modification of critical parts of the SRAM.**

I. INTRODUCTION

Nowadays, more than 50% of the System on a Chip (SoC) area is used to embed different kinds of memory. This is confirmed by the Semiconductor Industry Association roadmap, which forecasts that in 2013 more than 90% of the overall SoC area will be composed of memories [1]. In addition, memories are designed to exploit the technology limits in order to reach the highest storage density and high speed access. The main consequence is that memory devices are statistically more likely to be affected by manufacturing defects impacting the SRAM reliability.

The core-cell stability (ability of the core-cell to maintain its content) is a major concern to evaluate the SRAM design reliability as it represents the SRAM sensitivity to process variations and operating conditions. In current technologies, the decreasing feature sizes and the power supply scaling down make hard to maintain an acceptable core-cell stability as subtle defects appear more and more frequently. These defects are mainly due to bad vias or contact inducing resistive paths [2].

Existing DfT solutions for stability fault detection can be classified into two categories: Single and Programmable threshold methods. One of the well-known single threshold techniques, the Weak Write Test Mode (WWTM) [3], applies a weak overwrite stress to detect weak core-cells. Single-threshold techniques are tuned based on the best available pre-silicon simulation data. To achieve an acceptable test quality versus test yield tradeoff, such techniques may require multiple post-silicon design interactions. On the other hand, when the weak write stress is programmable, the test quality versus test yield tradeoff can be adjusted without design interactions and can be based only on the results of the post-silicon testing. Programmable threshold methods have been described in [4, 5, 6]. These DfT techniques consist in using the content of one or more core-cells in the column of the core-cell under test to modulate the bit line voltage levels and then use these levels to perform a weak write stress. These solutions are effective but require a long test time and some modifications of the internal timing structure to allow the selection of all core-cells of a column at a time.

In this paper we propose a new programmable DfT technique for stability fault detection. It consists in controlling the word line duration in order to apply a weak write stress on the targeted core-cell just after a write operation. This new DfT technique offers many advantages. It is programmable depending on the word line duration. It introduces a minor area penalty as it only requires the modification of the word line enable signal. Finally, its resulting test application time is very low compared to existing programmable stability fault detection techniques.

The paper is organized as follows. Section 2 provides an overview about memory core-cell stability faults. Section 3 details the proposed DfT technique. Section 4 presents the performances of the proposed solution. Finally, concluding remarks are given in Section 5.

II. SRAM CORE-CELL STABILITY FAULT

In this section we provide an overview on memory core-cell stability faults. First, we illustrate how a resistive-open defect may be a possible origin of a stability fault. Then, we provide the state-of-the-art of existing techniques for stability fault detection.

978-3-9810801-5-5/DATE09 © 2009 EDAA

A. Defect origin

Figure 1 illustrates the schematic of the well known 6T SRAM core-cell. Node S is the true storage node whereas node SB is the complementary one. As mentioned in the introduction, the high memory density makes these blocks to be more and more prone to defects, such as malformations in vias or contacts, inducing resistive paths. Thus, the state-restoring feedback (i.e. the inverter loop) of the weak core-cell is weaker or non-working at all. Consequently, the core-cell is prone to write or read disturb faults.

A defect in the drain of transistor Mtp2, as shown in Figure 1, may produce a stability fault, especially when the core-cell is storing a logic '0'. In presence of this defect, the strength of Mtp2 is reduced and the pull-up of node SB is not correctly acted. Then, the strength of transistor Mtn1 is reduced as its gate voltage level is reduced, inducing a difficulty to maintain a logic '0' on node S. Consequently, a stability fault may occurs as during the lifetime of a SoC, memories become weaker due to parasitic phenomenon such as DC noise, coupling effects or NBTI (Negative Bias Temperature Instability) [7].

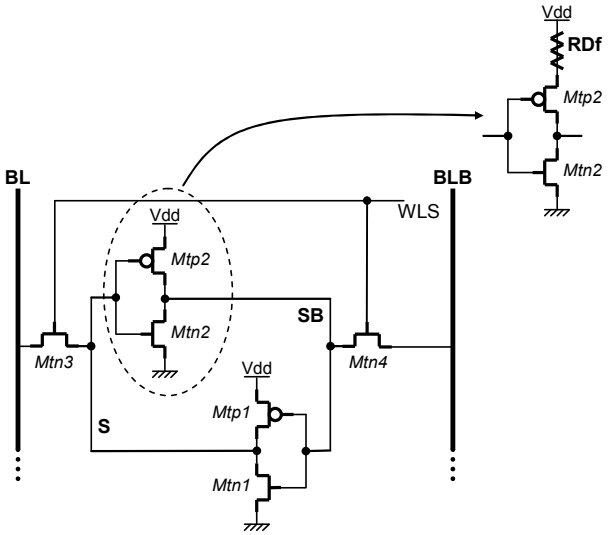

Figure 1: 6T SRAM core-cell

Moreover, it has been shown in [5] that resistive bridges and threshold voltage mismatches, which are more and more likely to occur in VDSM technologies using very aggressive design rules, may also be the root cause of stability faults.

B. State-of-the-art

As shown above, a stability fault may appear during the life time of the memory. Consequently, it cannot be detected using a functional test approach. DfT techniques have been proposed to enhance the detection of core-cell stability fault. They consist in controlling the bit line voltage levels to act a weak write stress on the core-cell under test. If the core-cell is affected by a stability fault, this weak write stress provokes the faulty swap of the core-cell and a straightforward read operation will allow the observation.

The most known DfT solution for stability fault detection is called Weak Write Test Mode (WWTM) [3, 8, 9]. It consists in adding a structure in the core-cell array (for each column or each write driver) to modulate the bit line voltage levels. This solution is effective but it offers a single detection threshold (determined by the size of the structure).

Another single threshold technique is presented in [10], where authors propose to separate the power supply of the memory array from that of the periphery. Thus, by lowering the supply voltage of the memory array, they are able to detect weak cells.

Programmable stability fault detection has been described in [4, 5, 6]. These DfT techniques consist in using one or more core-cells per column to modulate the bit line voltage levels. In [5], the authors present the Word Line Pulsing Technique (WLPT). It consists in applying multiple read operations on a core-cell (without precharge) to reduce the level on the bit line 'BL' (if the core-cell stores a logic '0') or on the complementary bit line 'BLB' (if the core-cell stores a logic '1'). During each read action BL or BLB is reduced of about ΔBL. Then, these resulting voltage levels are used to act a weak write stress on the core-cell under test.

DfT techniques proposed in [4, 6] consist in using some core-cells belonging to the same column than the targeted one to act the weak write. First, these core-cells are written with a certain background. Then, they are connected to the bit line at the same time. For example, if half of these core-cells contain a logic '0' and the remaining ones a logic '1', the resulting voltage levels on BL and BLB reach about Vdd/2. The weak write is then performed on the core-cell under test. Consequently, depending on the background the voltage levels on BL and BLB can be chosen making the solution programmable.

These programmable DfT solutions are effective but they require a long test time as all core-cells of a column have to be written to program the desired threshold detection level. Moreover, they require modifications of the internal timing to allow the selection of all core-cell of a column at a time.

March test procedures have been developed in order to detect degradations of core-cell stability [11, 12, 13, 14]. These test procedures have two main drawbacks. Firstly, as they consist in stressing core-cells, long March elements including breaks are used. Their resulting test application time becomes prohibitive. Secondly, depending on the defect size, a defective core-cell can escape such a March test.

III. Proposed DfT technique

In this section we present a new DfT technique for stability fault detection. First, we expose the principle, based on the Read Equivalent Stress (RES) proposed in [14] and next, we detail its implementation.

A. Principle

As mentioned above, existing DfT techniques for stability fault detection generally require a non negligible area overhead to implement a structure for modulating the bit line voltage levels. In order to eliminate this problem, we propose to use the Read Equivalent Stress (RES) principle proposed in [14]: when a core-cell is selected for an operation (read or write), the pre-charge circuit of its bit line is normally turned off. For the bit

lines that are not involved in the operation, the pre-charge circuit is commonly left 'on'. With the pre-charge active and the word line being high on the unselected columns, the core-cells fight against the pre-charge circuit. This principle is depicted in Figure 2. A read/write operation is performed on the core-cell in gray. All the other core-cells to the same word line undergo a stress due to the pre-charge circuits still 'on' state.

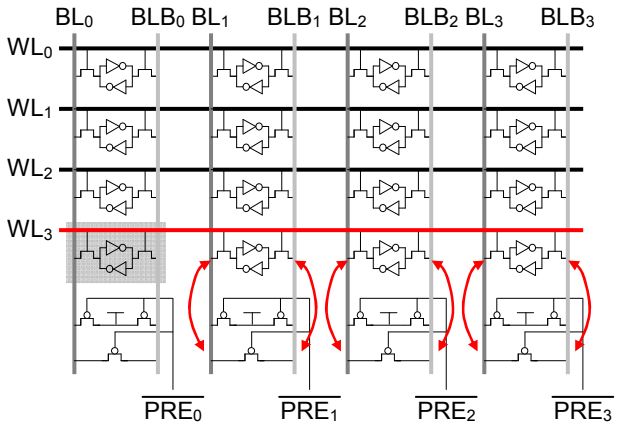

Figure 2: A 4x4 memory array

Although efficient, this technique is not able to detect small defects inducing stability faults. In order to show how the proposed technique may be more effective than the application of the RES principle, let us consider the waveforms in Figure 3. This simulation has been obtained using a core-cell designed in 45nm technology affected by a 10MΩ resistive open defect in the drain of transistor Mtp2 (see Figure 1). First, a w0 operation is performed in that core-cell that initially contains a logic '1'. From a logic point of view, node S reaches a logic '0' while node SB reaches a logic '1'. Nevertheless, due to the defect, the voltage level on node SB does not reach Vdd and remain a little bit higher than half Vdd.

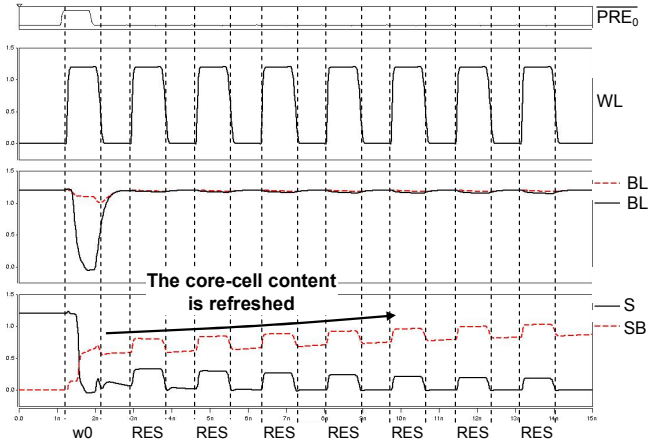

Figure 3: Weakness of RES action

Then, seven operations (read or write) are performed on another core-cell belonging to the same word line. The core-cell under test undergoes seven RES as it fights against the pre-charge circuit during each operation. These RESs stress the core-cell under test, but their action is reduced by restoration phases between two operations. In fact, during these phases, the core-cell is unselected (isolated from the bit lines) and consequently it has enough time to compensate the stress undergone during the previous RES. This is shown by the waveform of core-cell nodes S and SB, in Figure 3. This means that, the RES action is not sufficient to detect the presence of a resistive-open defect with small value, while it can be efficient for higher resistance value that may lead to other fault models such as dynamic Read Destructive Faults or dynamic Data Retention Faults. In the case of stability faults, we have to detect also low restive-open defect.

The proposed DfT technique consists in improving the effectiveness of RES action. It consists in writing the core-cell under test with an opposite value to its previous value and then, to let the word line enabled during a certain time. Consequently, the core-cell under test has directly to fight against the pre-charge circuit (switched-on), without having any refresh time. This stress has more impact than RES action and weak core-cells are therefore not able to maintain their value if the word line remains enabled during a sufficient amount of time. In the next sub-section we show the hardware implementation of the DfT module allowing to keep the word line signal 'on' just after the write operation.

B. Implementation

In SRAM memories, word line signals are generated by a word line decoder. Figure 4 depicts the scheme of a 2-bit word line decoder. It is based on NOR-gates; NAND and NOT gates are present for synchronization and buffering respectively. The WLEN signal activates the selected word line (WLi) depending on the address (Ai).

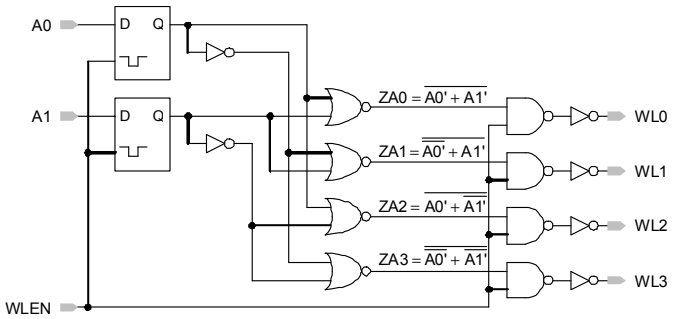

Figure 4: A 2-bit NOR-based word line decoder

In order to maintain the selected word line enabled just after a write operation, we can introduce an extra hardware on the WLEN signal path. The proposed hardware solution is depicted in Figure 5.

For low area overhead cost, an asynchronous solution has been adopted. It is made with five CMOS gates and an adjustable delay chain. This delay chain can be implemented by using different buffer chains selectable with a multiplexor for example. The TEST signal allows to switch-on the DfT module and then to increase the duration of the word line enable signal. Waveforms in Figure 6 show the basic functioning of this module.

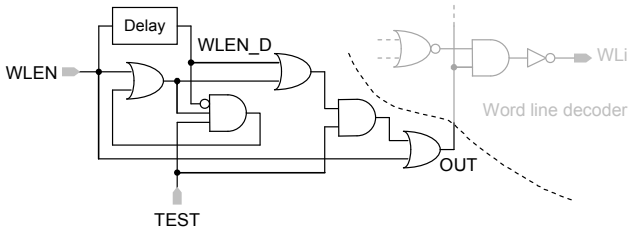

Figure 5: DfT module implementation

During the first word line activation, as the TEST signal is 'off', the output of the DfT module does not change the WLEN duration (T). Then, during the two following word line activation, the TEST signal is 'on'. In these cases, the output of the DfT module provides a WLEN duration increased by the additional delay (T + d1 or T + d2).

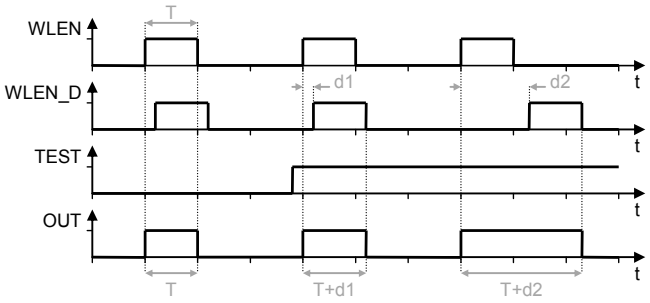

Figure 6: Functioning of the DFT module

Compared to existing DfT solutions for stability fault detection, the proposed technique presents many advantages:

- It is a programmable DfT solution, as the core-cell stability detection threshold is adjustable depending on the selected additional delay.
- It is easy to implement as it does not require any modification of the core-cell array, supply voltage organization and internal timing structure.
- It is a low cost DfT technique as only few CMOS gates and a delay chain are required independently of the memory size.
- It allows a fast core-cell stability fault sensitization as no data background is required.

IV. PERFORMANCES OF THE PROPOSED SOLUTION

The proposed DfT technique has been implemented and tested on an SRAM designed with a 45nm technology. As first result to demonstrate the effectiveness of the proposed solution, we consider again a core-cell affected by a 10 MΩ defect in the pull-up of Mtp2 transistor (see Figure 1). The PVT conditions are: process corner typical, power supply 1.2V and temperature of 25°C. Previously, it has been shown that the RES action has failed to detect a resistive-open defect with the considered resistance value (see Figure 3).

Waveforms in Figure 7 present two electrical simulations for two word line enabling duration: WL_Enabling_1 (doted lines) and WL_Enabling_2 (solid lines). Simulations start by a w0 operation on the core-cell. This core-cell initially contains a logic '1'. During the w0 operation, the precharge circuit of the selected column is switched-off. At the end of the w0 operation node S has reached 0v while node SB reaches a level a little bit higher than half Vdd.

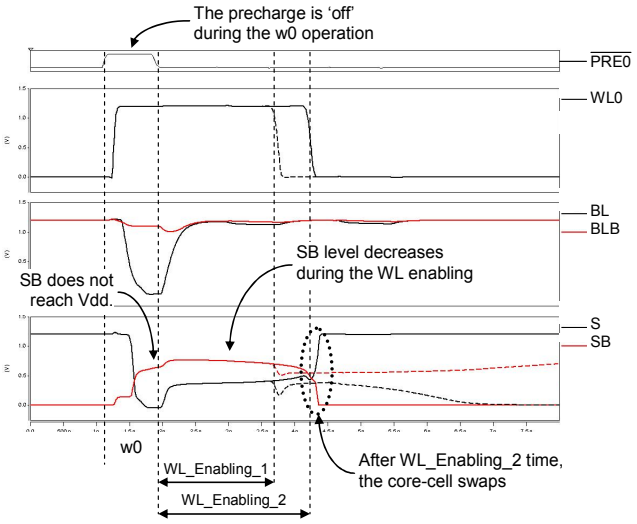

Figure 7: Effectiveness of the word line enabling technique

Then, instead of the normal un-selection of word line WL0, this remains active during a certain time in order to endure the stress on the weak core-cell. During this period, the level of SB decreases and S increases. After WL_Enabling_1 time, both nodes have not reached the threshold of Vdd/2 and, consequently, the weak core-cell does not flip. On the other hand, the WL_Enabling_2 time is large enough to make the faulty swap of the weak core-cell. The core-cell stability fault is therefore sensitized. A straightforward r0 operation will observe the fault effect (not illustrated in Figure 7).

Now, we compare the effectiveness of the proposed solution with the RES action. Figure 8 presents these comparisons. They have been done using an 8k x 16 SRAM memory designed with a 45 nm technology. Typical PVT conditions are used (typical process, 1.2V supply voltage, 25°C). The gray curve provides the minimum defect size detected with the proposed technique as a function of the word line enabling time. For example, when the word line remains enabled 4.2 ns after the write operation, the minimum defect size detected (i.e. inducing the faulty swap of the weak core-cell) is about 6.5 MΩ.

In Figure 8, it is also reported the minimum defect sizes detected by RES actions (black crosses). For example, 2 RES require 4.3 ns to be applied and detect a minimum defect size of about 40 MΩ. These data clearly demonstrate the interest of the proposed DfT solution as it allows the detection of defect sizes at least 6 times lower than that achieved with RES actions in a shorter test application time.

Finally, we compare the effectiveness of the proposed solution with the one presented in [6]. Both DfT techniques are compared in terms of test application time. Minimum defect sizes are not compared as both solutions do not used the same SRAM technology: 0.18 µm in [6] and 45 nm technology in our experiments.

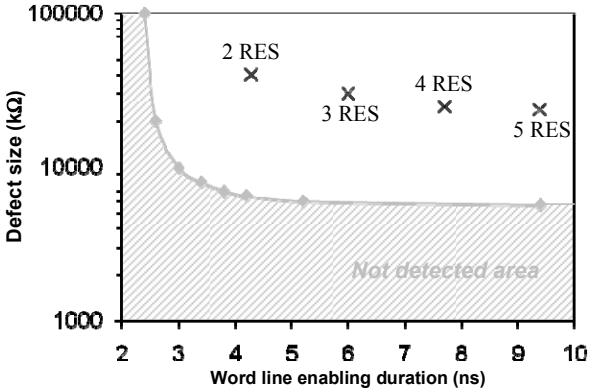

Figure 8: Word line enabling technique versus RES action

Here after, we have computed the resulting test application time (TTPAV) of the solution proposed in [6]:

TTPAV	=	t	for the core-cell initialization
	+	n x t	to write the background
	+	t	to modulate the bit line voltages
	+	t	to write the targeted core-cell ⇒ **fault sensitization**
	+	t	to read the targeted core-cell ⇒ **fault observation**
TTPAV	=	(n + 4) x t	(1)

with n being the number of core-cells written to fix the data background and t the cycle time of the SRAM. This number has to be large enough to obtain a good granularity on the bit line voltage levels. For example, authors in [6] use nine core-cells as background. Consequently, TTPAV becomes:

TTPAV = 13 x t (2)

Then, we have computed the test application time requires by our solution (TTWLET) as follows:

TTWLET	=	t	for the core-cell initialization
	+	t	to write the targeted core-cell
	+	WL_Enabling	the word line remains closed ⇒ **fault sensitization**
	+	t	to read the targeted core-cell ⇒ **fault observation**
TTWLET	=	3 x t + WL_Enabling	(3)

In order to compare the test application times we have to formulate the 'WL_Enabling' time as a function of the cycle time (t) of the SRAM. Let us consider again the gray curve of Figure 8. It appears that after about 5 ns (3 clock cycles) the minimum detectable value of a resistive-open defect is reached. Then, TTWLET becomes:

TTWLET = 6 x t (4)

Consequently, our DfT solution is two times faster than the programmable stability fault DfT technique proposed in [6].

V. CONCLUSIONS

In this paper we have proposed a low cost DfT solution for stability fault testing. It consists in controlling the word line duration in order to act a long stress on the core-cell under test strong enough to induce a faulty swap. The presented solution, called Word Line Enabling Technique, requires a low area DfT module on the word line enable signal controlling the address decoder. Compared to existing solutions, it offers the advantages of low area overhead, programmability, and low test application time. Moreover, it does not require any modification of critical SRAM parts such as the core-cell array and the internal timing structure.

REFERENCES

[1] Semiconductor Industry Association (SIA), "International Technology Roadmap for Semiconductors (ITRS)", 2005 edition.

[2] J.C.M. Li, T. Chao-Wen and E.J. McCluskey, "Testing for Resistive Opens and Stuck Opens", Proc. of IEEE International. Test Conference, pp. 1049-1058, 2001.

[3] A. Meixner and J. Banik, "Weak Write Test Mode: an SRAM Cell Stability Design for Test Technique", Proc. of IEEE International Test Conference, pp. 1043-1052, 1997.

[4] A. Pavlov, M. Sachdev and J.P. de Gyvez, "An SRAM Weak Cell Fault Model and a DFT Technique with a Programmable Detection Threshold", Proc. of IEEE International Test Conference, pp. 1106-1115, 2004.

[5] A. Pavlov, M. Azimane, J. P. de Gyvez and M. Sachdev, "Word line Pulsing technique for stability fault detection in SRAM cells", Proc. of IEEE International Test Conference, paper 33.1, 2005.

[6] A. Pavlov, M. Sachdev and J. P. de Gyvez, "Weak cell detection in Deep-submicron SRAMs: A programmable detection technique", IEEE Journal of Solid-State Circuits, Vol. 41, N° 10, October 2006, pp. 2334-2343.

[7] S.V. Kumar, K.H. Kim and S.S. Sapatnekar, "Impact of NBTI on SRAM Read Stability and Design for Reliability", Proc. of IEEE International Symposium on Quality Electronic Design, pp. 210-218, 2006.

[8] W. Schwarz, "Data Retention Weak Write Circuit and Method of using Same", U.S. Patent 5835429, November 10, 1998.

[9] D. R. Weiss, J. Wuu and R. J. Riedlinger, "Integrated Weak Write Test Mode", U.S. Patent 6192001, February 20, 2001.

[10] D.-M. Kwai, "Detection of SRAM Cell Stability by Lowering Array Supply Voltage", Proc. of IEEE Asian Test Symposium, pp. 268-273, 2000.

[11] R. Dekker, F. Beenker and L. Thijssen, "A Realistic Fault Model and Test Algorithms for Static Random Access Memories", IEEE Transaction on Computer, Vol. 9, N° 6, June 1990, pp. 567-572.

[12] A.J. van De Goor, "Testing Semiconductor Memories: Theory and Practice", John Wiley & Sons, West Sussex England, 1991.

[13] R. Rajsuman, "An Algorithm and Design to test Random Access Memories", Proc. of IEEE International Symposium on Circuits and Systems, pp. 439-442, 1992.

[14] L. Dilillo, P. Girard, S. Pravossoudovitch, A. Virazel, S. Borri and M. Hage-Hassan, "Dynamic Read Destructive Faults in Embedded SRAMs: Analysis and March Test Solution", Proc. of IEEE European Test Symposium, pp. 140-145, 2004.