Weak Cell Detection in Deep-Submicron SRAMs: A Programmable Detection Technique

Weak Cell Detection in Deep-Submicron SRAMs: A Programmable Detection Technique

Andrei Pavlov, *Member, IEEE*, Manoj Sachdev, *Senior Member, IEEE*, and José Pineda de Gyvez, *Member, IEEE*

***Abstract*—Embedded SRAM bit count is constantly growing limiting yield in systems-on-chip (SoCs). As technology scales into deep sub-100-nm feature sizes, the increased defect density and process spreads make stability of embedded SRAMs a major concern. This paper introduces a digitally programmable detection technique, which enables detection of SRAM cells with compromised stability [with data retention faults (DRFs) being a subset]. The technique utilizes a set of cells to modify the bitline voltage, which is applied to a cell under test (CUT). The bitline voltage is digitally programmable and can be varied in wide range, modifying the pass/fail threshold of the technique. Programmability of the detection threshold allows tracking process variations and maintaining the optimal tradeoff between test quality and test yield. The measurement results of a test chip presented in the paper demonstrate the effectiveness of the proposed technique.**

***Index Terms*—Design for testability, memory fault diagnosis, memory testing, SRAM cell stability, weak write test mode.**

I. Introduction

EMBEDDED memories can occupy up to 70% of the total area of modern systems-on-chip (SoCs) [1]. Owing to the higher robustness compared to DRAMs, embedded SRAMs are often used in SoC applications. However, due to the high packing density, SRAMs are often the yield limiters in SoCs [2]. Increased process spreads of modern scaled-down technologies and non-catastrophic defect related sensitivity to environmental parameters introduce mismatches in an SRAM cell and cause stability degradation in SRAMs [3]. ITRS-2003 [4] predicts "greater parametric yield loss with respect to noise margins" for high density circuits. Moreover, the proximity of SRAMs to noisy digital logic blocks in embedded applications makes SRAM cells more vulnerable to noise and coupling effects.

Even under nominal conditions the design of robust SRAMs in scaled geometries has become a challenging task. Process variations and mismatch impose fabrication specifications for parametric yield and an appropriate number of σ of static noise margin (SNM) over which SRAM of a certain bit count should work correctly. For instance, a 4 MB (32 Mb) cache SRAM with error correction circuitry (ECC) contains over 38 million cells. Limiting a design to one unstable cell in 38 million requires operation over a greater than 5σ of SNM parametric variation tolerance only due to process variations [5]. Hence, even with the ability to repair a few cells with the worst mismatch, variations must be acceptable to 5σ or beyond to achieve reasonable yield for such arrays. Subtle defects may further deteriorate the cell stability. Cells with marginal parameter matching have to be identified and preferably repaired by replacing them with the redundant cells.

Detection of data retention faults (DRFs) and stability faults (SFs) has been a time consuming and expensive effort. We will refer to the cells causing such faults as to *weak cells*. The Pause Test (or Data Retention Test, DRT) test traditionally used for weak cell detection is time consuming, requires elevated temperatures, may have insufficient fault coverage for DRFs, and is ineffective for most stability faults. Therefore, several weak cell detection design for test (DFT) techniques have been proposed. However, most of these techniques lack easy digital programmability of the pass/fail threshold, which we are trying to address in this paper.

This paper is organized as follows. Section II discusses the impact of process variations and non-catastrophic defects on SRAM cell stability. Section III presents the existing weak cell detection strategies. Section IV explains the concept of programmable weak cell detection threshold used in the proposed DFT and Section V describes the proposed implementation of the weak cell detection scheme. Test chip measurement results are presented in Section VI. Section VII presents a summary of the main contributions of this work.

II. SRAM Cell Stability

A typical six-transistor (6T) SRAM cell is presented in Fig. 1(a), where $Q1$ and $Q2$ are the driver, $Q3$ and $Q4$ are the load, and $Q5$ and $Q6$ are the access transistors, respectively.

To quantify the level of weakness of such cells, we apply the concept of static noise margin (SNM). SNM is defined as maximum possible square between the normal and mirrored voltage transfer characteristics (VTCs) [6]. SNM in the read-access mode is deteriorated by the pulling action of the saturated access transistor $Q6$ as shown in the equivalent circuit of a read-accessed cell in Fig. 1(b). Effectively, a CMOS inverter $Q2-Q4-Q6$ [Fig. 1(a)] is turned into a ratioed inverter with the logic "0" level above the ground potential [Fig. 1(b)]. The solid and the dotted lines in Fig. 1(c) show the VTCs of an SRAM cell in data retention mode and in the read-accessed mode, respectively. The SNM of a read-accessed cell is significantly smaller than the SNM of the same cell in the data retention mode, as shown in Fig. 1(c).

Due to mismatches and/or defects in the cell, the VTCs of the equivalent inverters forming an SRAM cell can be different.

Manuscript received March 18, 2005; revised June 1, 2006.

A. Pavlov is with Intel Corporation, Hillsboro, OR 97124 USA (e-mail: andrei.pavlov@intel.com).

M. Sachdev is with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: msachdev@ece.uwaterloo.ca).

J. Pineda de Gyvez is with Philips Research Laboratories, 5656 AA Eindhoven, The Netherlands (e-mail: jose.pineda.de.gyvez@philips.com).

Digital Object Identifier 10.1109/JSSC.2006.881554

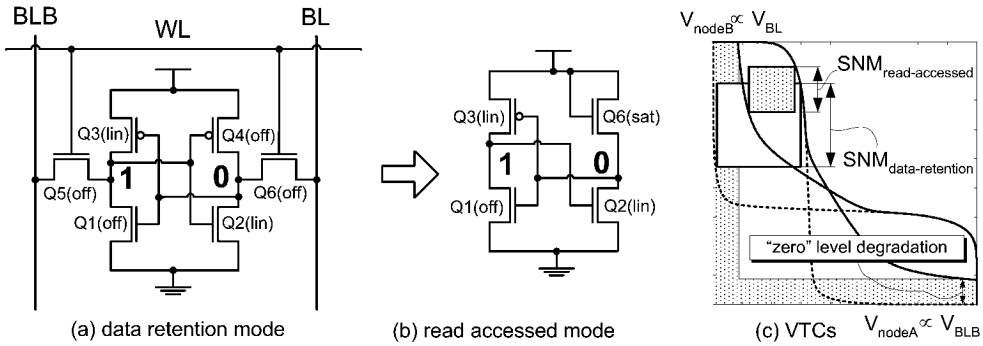

Fig. 1. (a) Six-transistor (6T) SRAM cell in the retention mode when $V_{BL} = V_{BLB} = V_{\rm DD}, V_{WL} = 0$ and (b) its equivalent circuit in read-access mode when $V_{BL} = V_{BLB} = V_{\rm DD}, V_{WL} = V_{\rm DD}$; (c) VTCs of a 6T SRAM cell in the retention and the read-accessed modes.

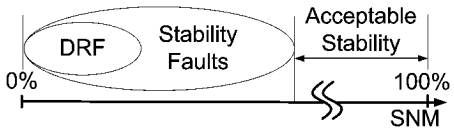

Fig. 2. Relationship between the SNM, data retention and stability faults in an SRAM cell.

Thus, the worst case SNM can be defined as the side of the smaller of the two largest squares between the two VTCs of a cell in the read-accessed mode (such as the $\rm SNM_{weak}$ in Fig. 7). In this paper we assume this definition of the SNM.

Depending on the severity of the SNM degradation, the stability problems in SRAM cells can be classified into data retention faults (DRFs) and stability faults (SFs), with the former being a subset of the latter, as shown in Fig. 2. For extremely low values of the SNM, the cell is likely to flip its state, i.e., it fails to retain its data demonstrating a DRF. If the SNM is sufficient to handle the nMOS off-state leakage current that discharges the storage node, under normal conditions it can retain its data as long as the power is supplied to the cell. However, under the adverse conditions such as the reduced supply voltage, elevated temperature, increased coupling noise, etc., i.e., the conditions contributing to the further SNM degradation, this cell may become so unstable as to flip its state. And finally, the cells with the SNM high enough to withstand the worst possible case scenario are outside of the oval representing the stability faults. We will refer to the cells inside and outside of the larger oval as *weak* and *good*, respectively.

Poorly formed pMOS transistors and/or contacts, shown as $R1$, $R2$ and $R3$ in Fig. 3, can cause SFs in an SRAM cell. A break or a weak open modelled by $R1$ creates a symmetric defect when both the data node pull-up paths of a cell have a highly resistive connection to the power supply. Infinite value of $R1$ corresponds to an open in the cell's supply or to the situation when both pMOS transistors are missing. Resistive opens represented by $R2$ and $R3$ create an asymmetric defect in the left-hand and right-hand pull-up paths of the cell, respectively, and can be modelled by resistive connection of $Q3$ and $Q4$ sources to nodes A and B, respectively.

Conditions and defect resistance range of a DRF detection by means of the DRT are illustrated in Fig. 4 on an example of an asymmetric defect $R3$. DRT test is conducted by reading

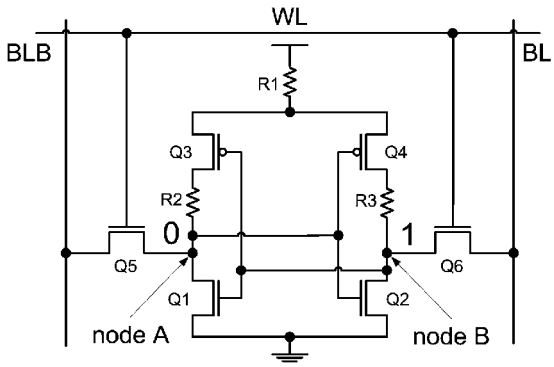

Fig. 3. SRAM cell schematic with resistors in place of potential weak opens that can cause stability faults (SFs).

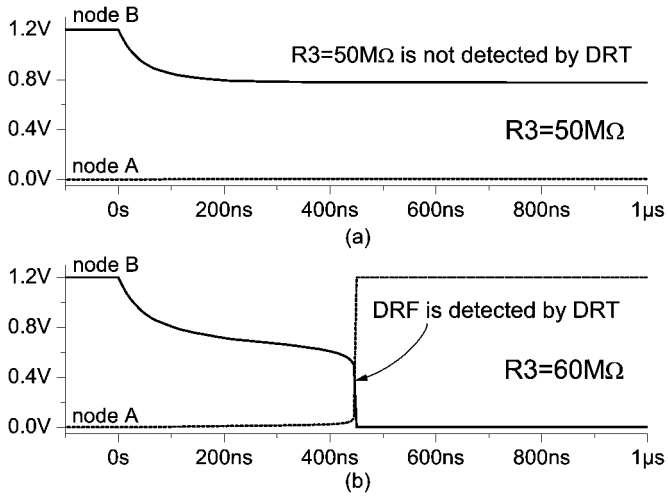

Fig. 4. Data retention fault due to the discharge of node B by the off-state current of $Q2$ (simulation results for CMOS 0.13-μm technology, $V_{\rm DD} = 1.2$ V, $T = 150\,^{\circ}$C). (a) A resistive open $R3 = 50$ MΩ is insufficient to flip the cell, whereas (b) $R3 = 60$ MΩ causes a DRF.

the SRAM array after a pause on the order of 100 ms to determine whether the background data has changed [7]. An asymmetric resistive open defect with resistance of 50 MΩ and below [Fig. 4(a)] is not detected even at the elevated temperature of 150 °C and a pause of more than 100 ms. However, when $R3 \geq 60$ MΩ, the off-state current of $Q2$ is sufficient to gradually discharge the node B capacitance and the cell under test flips at 450 ns, destroying the stored data. If the same test is conducted

at the room temperature, the lowest detected value of $R3$ will be 2.75 GΩ and the test time necessary to detect it will also be over 60 times longer. Obviously, the detection range of the DRT is insufficient to reliably identify many manufacturing defects that cause poor cell stability. Elevated temperatures help to improve the detection range by about 45 times at the cost of the increased test time. However, resistive opens of about 50 MΩ are still considered as strong opens [8]. Unless special tests are applied, cells with such defects will pass the standard tests and an SoC with highly unstable and unreliable SRAM cells will be shipped to the customer.

Subtle defects caused by the process disturbances can also reduce the stability of the cell. The likelihood of resistive bridges grows as the critical area shrinks with scaling, and resistive break defects are likely to appear in place of poor or absent contacts, vias or silicide [7], [9]. While these defects can be non-catastrophic, i.e., not causing a functional fault, they can have a serious impact on the cell stability as shown in Fig. 5(a) and (b), respectively. The cell weakness can be caused by various factors, including resistive defects (resistive breaks and bridges), excessive process shifts, mask misalignment, transistor mismatch, etc. [3]. An SF may occur due to any electrical disturbance such as power supply noise, read/write cell disturbs, i.e., during the normal operation of the SRAM. These adverse conditions, especially combined, can cause a weak cell to flip its state easily and corrupt its contents.

As the technology is scaled in the decananometer region, process parameter spread becomes one of the main contributors to the wider distribution of the SNM figures even in a defect-free chip. The granularity of the electric charge and the atomicity of matter begin to introduce substantial variation in number and position of dopant atoms in MOSFET channel, t_{ox} becomes equivalent to several atomic layers with one to two atomic layers roughness [10]. For sub-100-nm CMOS SRAMs, 6σ deviations of the SNM only due to intrinsic device fluctuations are projected to exceed the nominal SNM [11]. Randomness of channel dopant distribution will become a major source for the SNM deterioration. To maintain reasonable SNM and yield of future bulk CMOS SRAMs, cell ratios may have to be increased from the typical $r = (W/L)_{\text{driver}}/(W/L)_{\text{access}} = 2$ towards the higher ratios [12], which counter-balances the scaling advantages of decananometer technologies with respect to the area of the embedded SRAM cores. For instance, even a single transistor V_{TH} variation can cause significant SNM degradation [Fig. 6(a)]. Furthermore, when more than one V_{TH} is deviating from a designed value [Fig. 6(b)], the impact on the cell stability becomes even stronger.

Inadequate SNM due to undetected resistive defects may indicate intermittent stability and possible long term reliability issues in SRAM cells, while such cells will successfully pass the regular march tests. For certain applications requiring extreme reliability and low PPM levels (e.g., automotive, where SoC chips control such systems as ABS, stability control; life support systems; servers, etc.) detecting all weak defects and possibly unstable cells is crucial. For instance, the new dual-core 90-nm Itanium-family microprocessor [13] is employing the Programmable Weak Write Test Mode [14] for stability testing of its 24 MB on-chip L3 SRAM cache. The identified defective

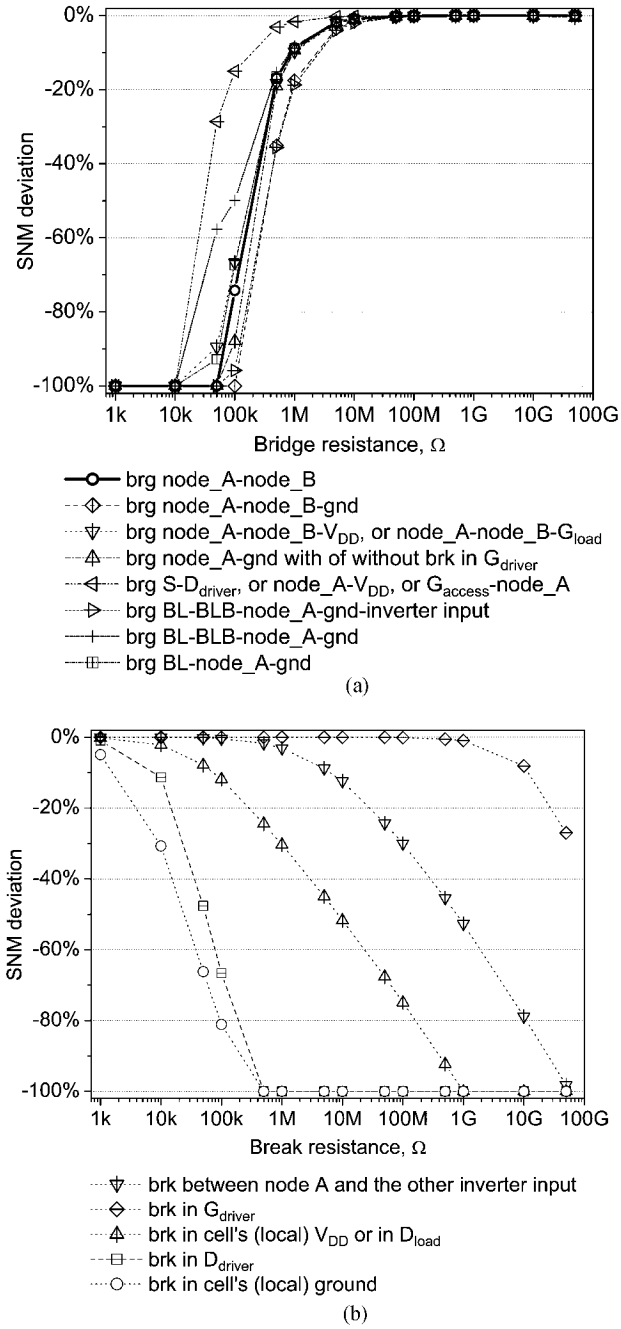

Fig. 5. SRAM cell SNM as a function of non-catastrophic (a) bridge and (b) break defect resistance; simulation results in CMOS 0.13-μm.

cells are then repaired by redundant ones. Thus, cell stability test becomes a necessity for successful debug, test, and repair to ensure low Parts Per Million (PPM) levels.

III. Weak SRAM Cell Detection Strategies

To obtain acceptable noise margins in the deep-submicron technologies with reduced supply voltages, bitlines in a vast majority of SRAM designs are precharged to the full supply voltage. However, reading a 6T SRAM cell with bitlines precharged and equalized at full V_{DD} may not detect several types of defects causing DRF or stability faults, e.g., a missing p-channel in pull-up transistors, poor or absent vias to the

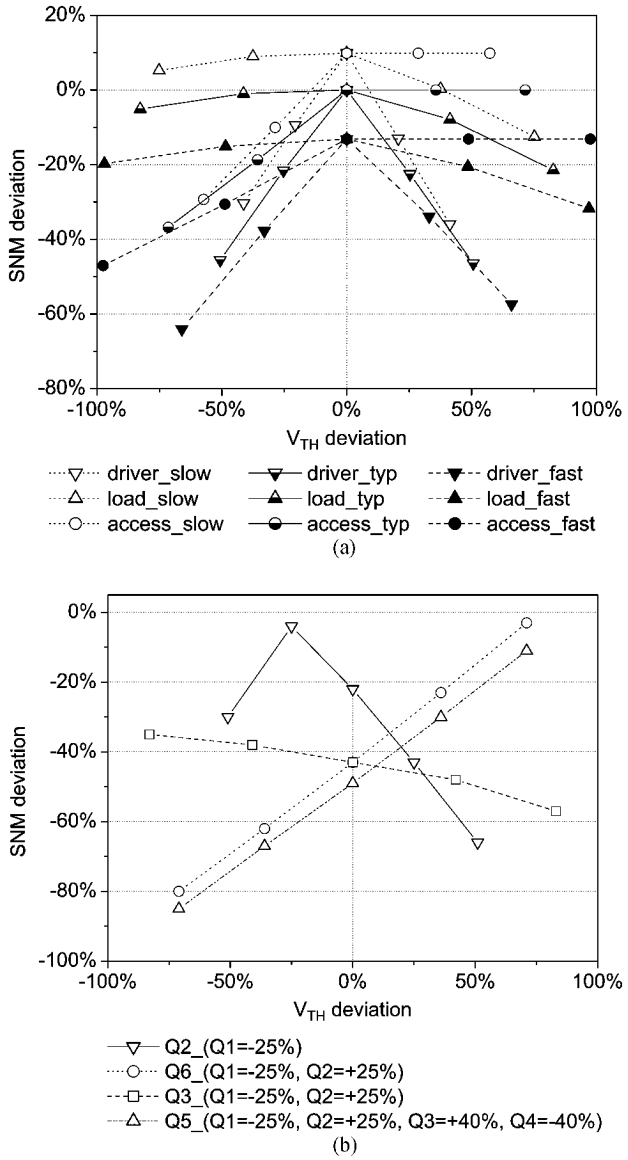

Fig. 6. SRAM cell SNM as a function of V_{TH} variation of (a) one of the transistors and (b) more than one transistor at once; simulation results in CMOS 0.13-μm.

pull-up transistors (an SRAM cell in this case will act as a "good" 4T DRAM cell). Detection of such cells in SRAM arrays may require a data retention test (DRT). However, the DRT can take significant time leading to a more expensive test. Moreover, for stricter PPM levels, some cells may require excessively long DRT pause times, reduced supply voltage and high temperature, and even then the DRT may still miss a wide range of resistive defects. Detection of such defects as poorly formed vias and contacts, shorts with nonzero resistances, gross mismatches in cell transistors require special test conditions or stresses to be applied to make certain that parametric faults (i.e., stability faults) are reliably detected.

Standard suite of test methods often lack sensitivity to detect parametric failures [15]. To provide a better stability fault coverage, several DFT techniques for weak cell detection have been proposed in literature [3], [7], [9], [14], [16]–[19]. These techniques exploit the fact that the state-restoring feedback of a weak cell is weaker or absent and thus they are more susceptible to write or read disturbs. Initially, most of the weak detection techniques were targeting the detection of the DRF. However, with scaling more subtle defects than completely open connections in load transistors can remain undetected and lead to field failures, which often are intermittent and hard to diagnose. Detection of parametric stability faults needs to strike a balance between the yield loss of over-testing the cells on one hand and and the test escapees due to under-testing the cells on the other.

Ideally, a cell stability test technique should satisfy all of the following conditions:

- selectivity and fault coverage: Weak cell detection techniques must not cause the normal cells to fail, i.e., the normal cells $(6\text{–}7\sigma)$ should be able to withstand the application of the test stress. At the same time, the weak cells should not escape the test.
- speed, power, and area: The added test circuitry should not affect the operating speed or power, and take as little area as possible.
- test time: Preferably, no extra test cycles should be added to reduce the expensive tester time spent on each chip.

Some of the existing weak detection techniques satisfy the above criteria better than others. A more detailed analysis of the existing techniques and their implementation complexities is presented next.

A. Previous Art

As process parameter spreads continue to grow with technology scaling, one of the most challenging to satisfy criterion is the selectivity of a weak cell detection technique. With respect to selectivity and fault coverage, all the techniques for weak cell detection can be categorized into single and programmable detection threshold (or weak write stress) techniques. Testing SRAM cells with a fixed weak write stress can lead to under- or over-testing of the targeted defects in SRAM cells due to poor process tracking characteristics. 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 iterations to account for the process changes following the initial design. Whereas, if the weak write stress is programmable, the test quality versus test yield tradeoff can be adjusted without the design iterations and can be based on only on the results of the post-silicon testing. Obviously, programmable detection threshold techniques are superior in terms of time to market and test yield loss minimization.

1) Single Threshold Methods: One of the well-known techniques, the Weak Write Test Mode (WWTM) [7], applies a weak overwrite stress to detect weak cells. Weak write circuit can be a stand-alone as in [7] or integrated into a write driver [17]. While the WWTM makes use of weaker write driver transistors, underdriving the access transistors using lower wordline voltage during a write operation [18] can also be used to apply weak write stress. Conversely, an elevated wordline voltage can be used to apply the test stress and detect a weak cell [19]. Kuo *et al.* [16] suggested a soft defect detection (SDD) technique based on the fact that defect-free inverters of an SRAM cell will provide certain read current upon access. If there is a break in cell's

connections, the read current is insufficient or absent. Another approach proposed Kwai *et al.* is to separate the power supply of the memory array from that of the periphery [9]. With the corresponding isolated terminal the memory array can be operated at a lower voltage, making it susceptible to read or write disturb. However, this test alone cannot guarantee detection of all DRF and is mostly used for process development. Moreover, having a separate pad for each of tens of memory arrays in modern SoCs may be impractical.

Many of the single threshold methods can be altered to enable analog control of the weak overwrite stress. However, analog levels are more difficult to control on the global chip level if implemented internally or more pad- and tester-demanding if controlled at the ATE level.

2) Programmable Threshold Methods: Large process spreads of modern deep-submicron technologies necessitated the arrival of digitally programmable techniques for weak cell detection, which can track process variation and/or can target the detection of cells with various degree of weakness. Selvin *et al.* proposed one of the possible extensions of the WWTM [14]. By switching the bias-setting transistors with a decoder, the overwrite stress applied to a Cell Under Test (CUT) can be varied and adjusted to track the process variations.

A technique using the ratio of read currents of SRAM cells within a column to create a weak write stress to the CUT is described in [3]. The ratio of read currents is defined by the number of the cells carrying "0"s in the total number of cells n forming the ratio. One of the cells among the n is the CUT. After enabling of all n wordlines at once, the bitlines will be partially discharged. The side with the larger number of "0"s will discharge the corresponding bitline deeper. The degree of the bitline discharge is proportional to the ratio of read currents on the n cells discharging each bitline. If the CUT is carrying a "1" and the true bitline has been discharged deeper than the complementary, a weak overwrite stress will be applied to the CUT. A CUT with poor stability will flip its state, whereas a stable CUT will withstand the test stress. By changing the number of "0"s on the same side where the CUT is carrying a "1", one can digitally program the overwrite stress applied to the CUT and thus detection threshold of the technique. The resolution of the method can be refined by increasing of the total number of cells n comprising the read current ratio and/or by changing the pulsewidth of the pulse enabling all n cells.

B. Proposed Technique

In this paper, we introduce a new digitally programmable DFT technique capable of detecting SRAM cells with a varying degree of stability degradation. Similarly to [3], it utilizes a ratio of "0"s and "1"s in a group of n cells. However, the CUT in the proposed technique is outside of the n cells comprising the ratio. The weak overwrite stress is created by connecting the CUT to partially discharged *floating* bitlines. We named the proposed technique the read current ratio technique (RCRT). The RCRT was implemented in CMOS 0.18-μm technology and measurement results from the fabricated test chip are presented.

IV. Proposed Weak Cell Detection Concept

To illustrate the concept of programmable detection, let us consider the voltage transfer characteristics (VTCs, a.k.a. an "eye diagram") of a good SRAM cell (solid lines) and a weak SRAM cell (dashed lines) presented in Fig. 7. In most cases, a weak cell is likely to have an asymmetrical eye diagram. In other words, the SNM of each of the data nodes of a weak cell is not the same, as illustrated in Fig. 7. The shape of the VTCs, which defines the SNM, depends on various factors. Fluctuations of V_{TH} and L_{eff}, presence of defects, and poorly formed contacts and vias can make the driving strength of one of the inverters in a cell weaker. This results in the shift of the metastability point Z of the cell. During a read operation, the logic "0" level stored in the weaker node of the weak cell (point B_{weak} in Fig. 7) will degrade more compared to a good cell (point B_{good}). The shift in the VTC on the weaker node will result in a smaller SNM. Moreover, the metastability point of the cell also shifts towards the weaker node. This effectively makes this node more susceptible to weak overwrite.

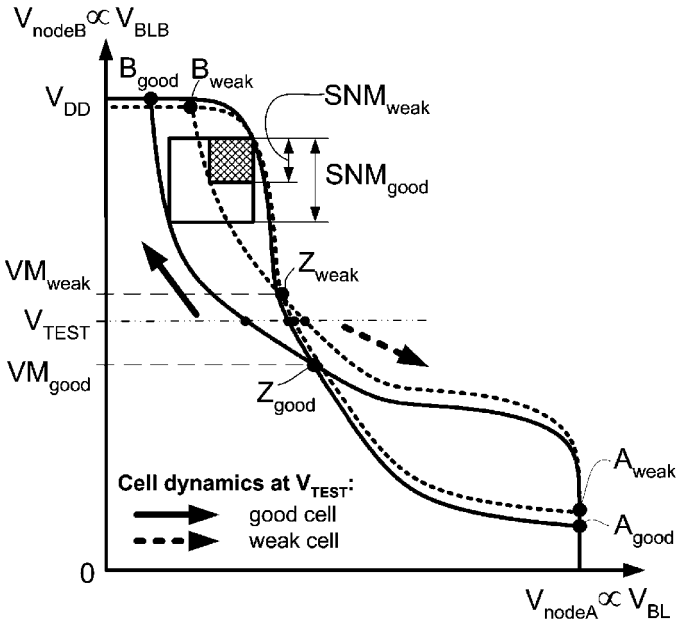

Fig. 7. Choice of V_{TEST} with respect to the metastable points of a good cell VM_{good} and a weak cell VM_{weak}.

We will be considering the worst case, i.e., the data node with the smaller SNM (node B). Axes in Fig. 7 represent node A and node B voltages, which in turn, are proportional to the corresponding bitline voltages V_{BL} and V_{BLB}. VM_{good} and VM_{weak} represent the metastability points of a good and a weak cell, respectively. Points A_{good}, B_{good} (A_{weak}, B_{weak}) on the transfer characteristic represent the stable states and Z_{good}, (Z_{weak})—the metastable states of the good (weak) cell, respectively. If node A or node B of an SRAM cell is driven beyond the metastable point, then the cell will flip its state. As it is apparent from Fig. 7, the SNM of the weak cell is significantly smaller than SNM of the good cell $(\text{SNM}_{\text{weak}} < \text{SNM}_{\text{good}})$.

Let us assume that node B of an SRAM cell has state "1", the bitlines are pre-charged to a known value (e.g., V_{DD}) and we have means to manipulate the voltage on node B and reduce it to V_{TEST}, as shown in Fig. 7. Note that V_{TEST} crosses the eye-diagram of a good cell (solid lines) above the metastable point Z_{good}, whereas it crosses the eye-diagram of the weak

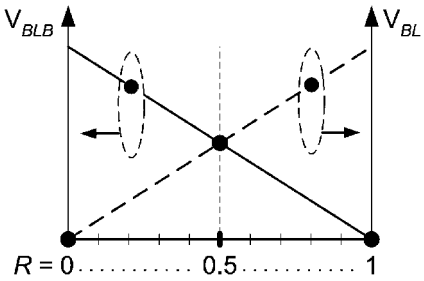

Fig. 8. Definition of the programmable ratio R in the proposed detection technique (RCRT). $R = (\text{number of cells with state 0 in a set of } n \text{ cells})/n$.

cell below its metastable point Z_{weak}. Upon removal of the test stimulus V_{TEST}, node B of the good cell will retain its state "1" while node B of the weak cell will flip to state "0" as shown by a solid and dashed arrow in Fig. 7, respectively. In other words, the weak cell will have been overwritten by the voltage level V_{TEST}. The implementation of this principle is described in detail in Section V.

By varying the V_{TEST} value, one can test for a given degree of cell weakness. All the cells, which flip at the node voltage above V_{TEST} are deemed "weak" as they have inadequate SNM as illustrated in Fig. 7. The rest of the cells is assumed to have acceptable stability. Making V_{TEST} programmable provides the flexibility to change the pass/fail threshold and/or to track the process variation without post-silicon design iterations to account for a change in the process parameters and/or the target quality levels.

V. Proposed Implementation

The concept of programmable threshold is implemented using a set of n SRAM cells in a given column. Existing cells in the column or external cells can be utilized for this purpose. Let R be the ratio of cells having state "0" to the total number of cells in a set of n cells (Fig. 8). We assume that the rest of the cells in a set n have state "1". Initially, BL and BLB are precharged to V_{DD}. By manipulating the value of R, and simultaneously accessing n cells, we can manipulate the bitline voltage. For instance, if the number of cells carrying zeroes and ones is equal ($R = 1/2$ in Fig. 8), then, provided the n cells have the same driving strength, $V_{BLB} = V_{BL}$. Now, suppose that $R > 1/2$. That will cause V_{BLB} to be less than V_{BL}. Respectively, if $R < 1/2$, then V_{BLB} will be higher than V_{BL}.

Now, if we write a certain ratio of "0"s and "1"s to a set of n cells, disable precharge and then simultaneously enable n wordlines, we can reduce V_{nodeA} or V_{nodeB} to a given V_{TEST} value. Then, if after activation of WL_CUT $V_{\text{TEST}} > VM_{\text{weak}}$, the regenerative property of the CUT will restore the stable state and the CUT will not flip. This situation is similar to a non-destructive read operation with incompletely precharged bitlines. The target range for V_{TEST} is such that $VM_{\text{good}} < V_{\text{TEST}} < VM_{\text{weak}}$. Within this range, weak cells with poor SNM will flip and be detected, whereas the cells with satisfactory stability will withstand this stress. This is the selectivity condition of weak cell detection. And finally, when $V_{\text{TEST}} < VM_{\text{good}}$, even the good cells will flip. Therefore, by programming the ratio R, we can adjust the pass/fail threshold of the RCRT cell stability test. The programmable stress settings can also be used for raster scan during defect debug and diagnosis.

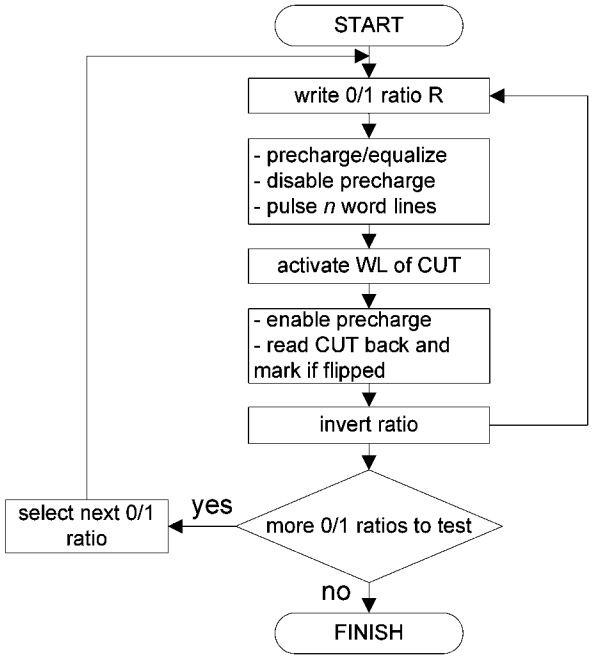

Fig. 9. Flow diagram of the RCRT.

The flow diagram shown in Fig. 9 depicts the sequence of steps necessary to implement the proposed digitally programmable weak cell detection technique. An inverse of the current 0/1 ratio is necessary to detect the weak cells that may flip in the opposite direction.

One of the possible implementations of the RCRT is shown in Fig. 10. The weak cell detection starts by determining the minimal acceptable cell stability. The cells with the SNM below that minimum must flip when connected to the bitlines partially discharged after pulsing the wordlines of the n cells in the column with the necessary ratio R. Access transistors of each side of the n cells share a common gate and a common bitline nodes. The other terminal of each of the access transistors is connected either to the ground or to V_{DD} through the corresponding driver or load transistors of their corresponding cells. The resulting potential on each of the bitlines is a function of the chosen ratio R of the cells carrying "0"s and "1"s connected to the bitline and the pulsewidth of the wordline pulse that enables the n cells forming this ratio. Assuming equal cell read currents, bitline potentials will be equal when 50% of the n cells are in state "0" and the other 50% of cells are in state "1" because the path resistance to the ground and V_{DD} is the same, i.e., $R = 0.5$ (Fig. 8).

When the n wordlines are enabled at once, the capacitance of each bitline discharges according to the time constant created by the corresponding equivalent path. If the bitlines discharge too much, then upon the enabling of WL_{CUT} pulse V_{TEST} can drop below the metastable point VM_{good}. In this case even the good cells with acceptable SNM will flip. To prevent the situation when V_{TEST} is equal or below the metastable point VM_{good}, the pulsewidth of the pulse enabling n wordlines should be shortened or the ratio R should be reduced. For $R = 0.5$, the

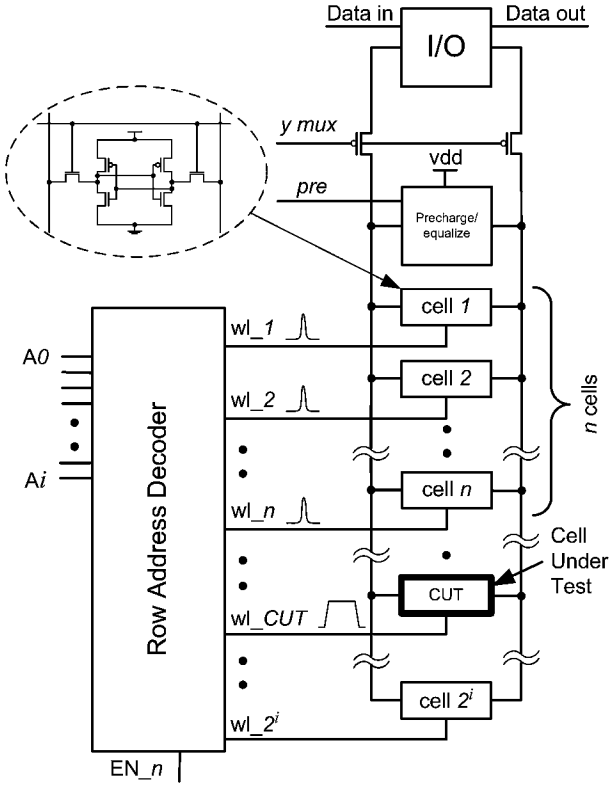

Fig. 10. Implementation of the RCRT.

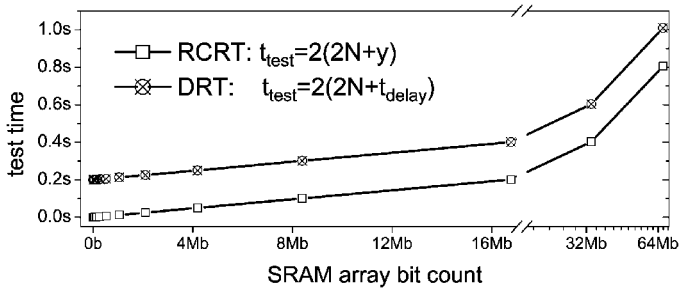

Fig. 11. RCRT test time and DRT test time as a function of the bit count (assuming $t_{\rm cycle} = 3$ ns, $t_{\rm delay} = 100$ ms).

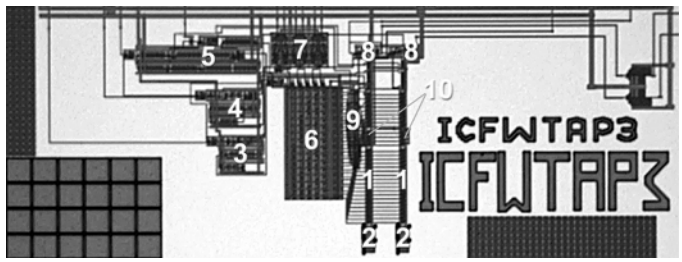

Fig. 12. RCRT test chip microphotograph.

bitline voltage will be approximately equal. If $R \neq 0.5$, the corresponding path resistances to $V_{\rm DD}$ and the ground will be different and thus the bitlines discharged to different levels. After the bitlines have been preconditioned, the wordline of the CUT is activated. Thus, we effectively read the CUT but with the bitlines precharged below the standard value of $V_{\rm DD}$. If the CUT is weak, i.e., has an inadequate SNM, then reading it with partially discharged bitlines will cause it to flip. By controlling the degree of the bitline discharge we can shift the pass/fail threshold of the test. Both the ratio R and the pulsewidth $wl_1 - wl_n$ pulse can be digitally reprogrammed to set a new weak cell detection threshold.

In practice, one is free to use various arrangements to form ratio R. Ratio R can be formed either by the regular cells from the same column, or by external dedicated cells, or by a combination of the above. To improve the resolution of the proposed detection technique, the number of cells n forming the ratio R can be increased. Larger n will also help to mitigate the effect of the possible read current mismatch among the n cells. To further improve the reliability of the RCRT, two groups of n cells can be used in each column. In this case, either group of n cells can be used to test the cells in the column. Moreover, one group of n cells can be used to ensure the stability of the cells comprising the other group of n cells.

Higher capacitance of the bitlines will provide higher detection accuracy. Therefore, the proposed DFT technique may be more attractive for larger SRAM instances with more capacitive bitlines.

If x is the total number of cells in an SRAM array, n is the number of cells used to form ratio R and $t_{\rm cycle}$ is the cycle time, then the test time for each ratio R can be estimated as $t_{\rm test\_RCRT} = 2(2N+y)$ where $N = x \cdot t_{\rm cycle}$ and $y = n \cdot t_{\rm cycle}$. The test time $2N + y$ is doubled to check the other side of the cells under test with an inverted ratio R.

The DRT test time can be estimated as $t_{\rm test\_DRT} = 2(2N + t_{\rm delay})$, where $t_{\rm delay}$ is the delay time for retention test of each data node. Fig. 11 illustrates the RCRT test time advantage over the DRT test time. Effectively, the DRT test time for small SRAM instances exceeds the RCRT by $2t_{\rm delay}$. However, the relative test time advantage decreases once the bit-count related test time component N starts to offset the $2t_{\rm delay}$ of the DRT.

VI. Measurement Results

To verify the proposed DFT technique a test chip has been designed and fabricated in CMOS 0.18-μm technology. A microphotograph of the test chip is shown in Fig. 12. The test chip comprises an asynchronous SRAM with two columns of 32 cells each (1) with extra 200 fF capacitors to imitate more capacitive bitlines (2) connected to each of the bitlines; self-timed control blocks to provide read (3), write (4) and test (5) timing; address decoder (6); address transition detector (7); sensing and writing circuitry (8); wordline switches (9) and a set of weak cells (10).

To demonstrate the detection capabilities of the RCRT, we used nine regular SRAM cells to form the ratio R. To enable nine wordlines simultaneously, our test chip row address decoder was modified to include nine switches (9) on the first nine wordlines. In practice, the switching function can be performed by two-input OR gates between the post-decoder and the wordline buffers. When activated by a pulse coming from the test timing block, all nine wordlines are pulled up simultaneously. The pulse can be formed locally by a simple one-shot circuit. The width of the pulse activating all nine wordlines of the n

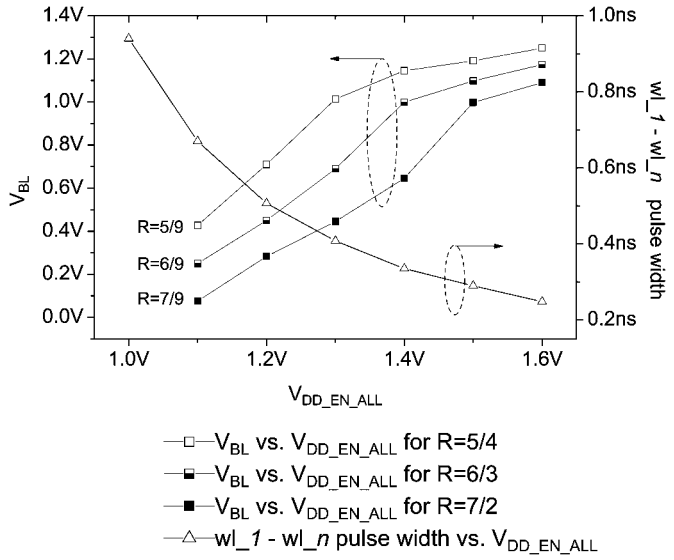

Fig. 13. Bit line voltage and the pulsewidth of $wl\_1 - wl\_n$ pulse as a function of $V_{DD\_EN\_ALL}$ (post-layout simulation results).

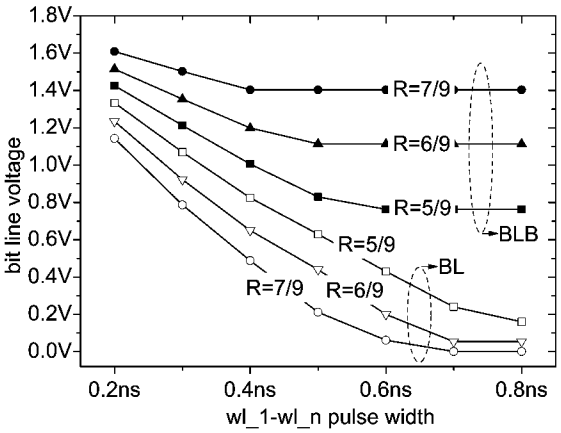

Fig. 14. Bit line voltage as a function of $wl\_1 - wl\_n$ the pulsewidth (post-layout simulation results).

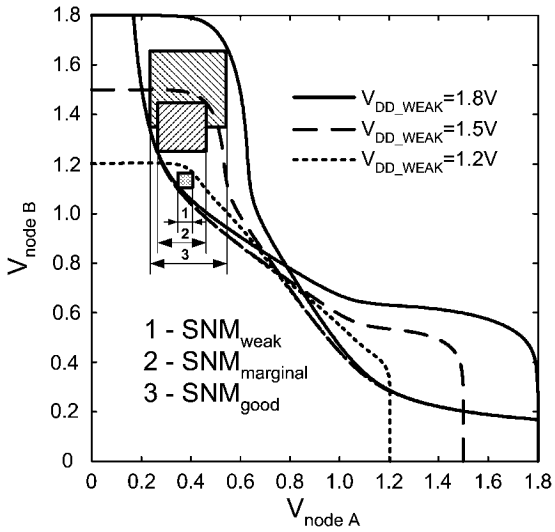

Fig. 15. Dependence of the VTC shape and the SNM on the cell supply voltage (V_{DD_WEAK}).

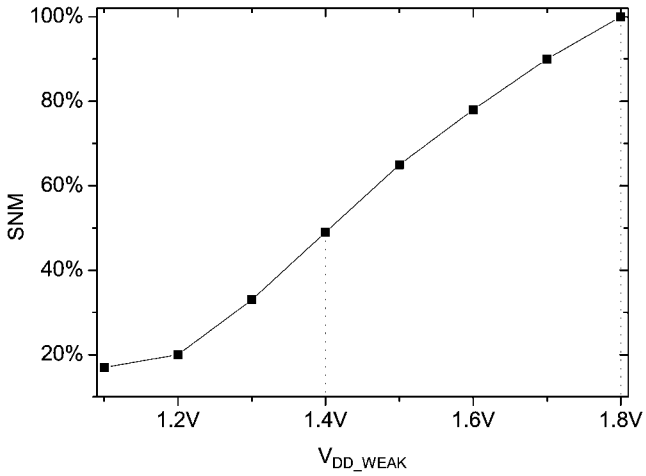

Fig. 16. SNM as a function of the cell supply voltage V_{DD_WEAK} (post-layout simulation results).

cells in the ratio R must be short enough, so that the wordline is deactivated before the cells forming ratio R have flipped.

Fig. 13 shows that depending on the number of cells carrying "0" among the nine cells comprising ratio R, the same pulsewidth of the pulse enabling all nine wordlines ($wl\_1 - wl\_n$) will discharge the corresponding bitline to a greater or a lesser extent. To be able to control the pulsewidth of $wl\_1 - wl\_n$ pulse in the test chip, we utilized an external voltage $V_{DD\_EN\_ALL}$, which supplies the delay line in a one-shot circuit and thus modulating the width of the produced pulse. Bit line potentials as a function of $wl\_1 - wl\_n$ pulsewidth are shown in Fig. 14. The required pulsewidth can also be specified and fixed by proper sizing of inverters in the one-shot delay chain. In this case, adjusting of the detection threshold is done only by reprogramming ratio R.

After the bitlines have been discharged to a certain extent, defined by the chosen ratio R and the pulsewidth of n wordlines, the wordline of the cell under test is activated and the partially discharged floating bitlines apply the test stress. From Fig. 15 we can see that a weak (defective) cell has significantly lower SNM than a good (defect-free) cell. Therefore, a weak cell will flip when read-accessed with a lower V_{BL} applied to the node storing a "1" and be detected, whereas a good cell will withstand this test stress.

To imitate weak cells in the test chip we used several cells with a separate supply voltage V_{DD_WEAK} (Fig. 12 (10)), which can be adjusted independently from the $V_{\rm DD}$ of the rest of the chip. Fig. 16 shows that reducing $V_{\rm DD}$ of an SRAM cell reduces the SNM of the cell and thus—the cell stability. For instance, to simulate a CUT with 50% of SNM, we need to reduce the power supply voltage of that CUT from 1.8 V to approximately 1.4 V.

Test stimuli for the test chip were provided by Agilent 93000 SoC series tester. We swept V_{DD_WEAK} and $V_{DD\_EN\_ALL}$ from 0.9 V to 1.8 V. For each of the combinations of V_{DD_WEAK} and $V_{DD\_EN\_ALL}$ we wrote a predetermined ratio R of "0"s and "1"s into the n cells. After applying the test sequence described in Section V we registered whether the

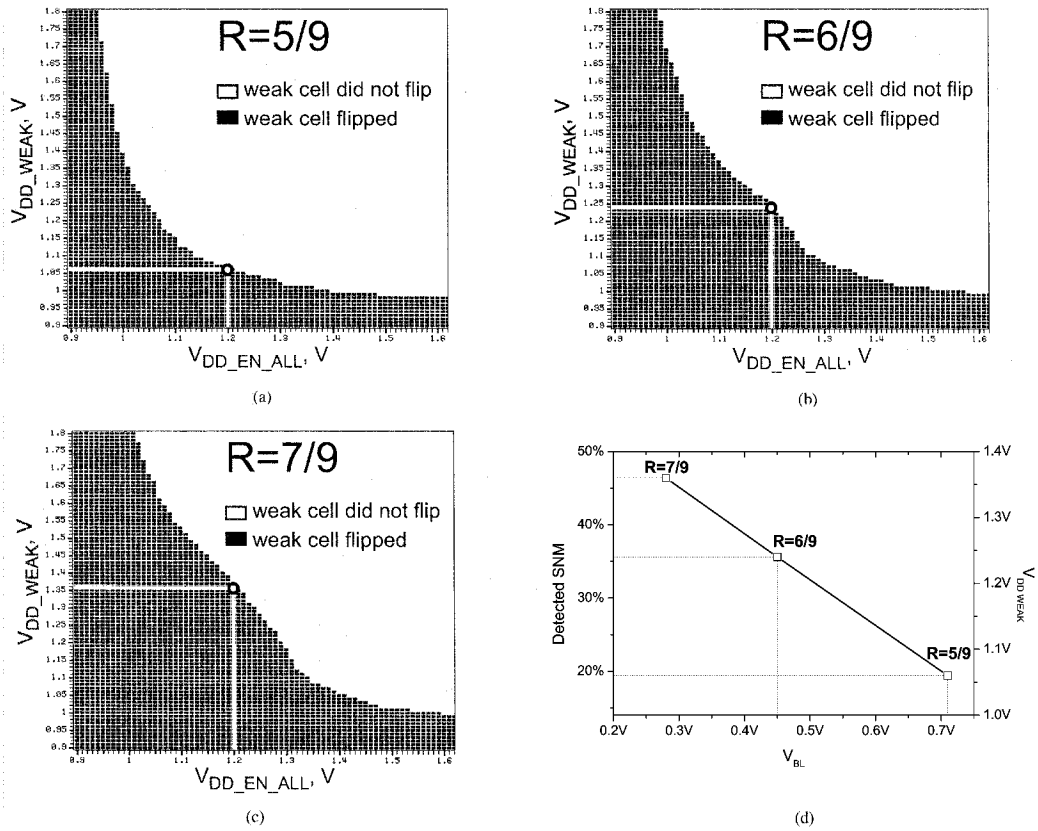

Fig. 17. Shmoo plots for ratios (a) $R = 5/9$, (b) $R = 6/9$, and (c) $R = 7/9$. Circled dot represents the detection threshold for $V_{DD\_EN\_ALL} = 1.2$ V. (d) Summary for $V_{DD\_EN\_ALL}$ fixed at 1.2 V that corresponds to 500 ps pulsewidth of $wl_1 - wl_n$ pulse (see Fig. 13). Programming R from $5/9$ to $7/9$ changes the weak cells' detection threshold from 18% to 46% of the nominal SNM.

TABLE I
DETECTION CAPABILITIES OF THE PROPOSED TECHNIQUE

Ratio R	5/9	6/9	7/9
V_{BL} (at $V_{DD\_EN\_ALL}$=1.2V)	0.71V	0.45V	0.28V
Detected V_{DD_WEAK}	1.06V	1.24V	1.36V
Detected normalized SNM	0.18	0.36	0.46

weak CUT has flipped. Fig. 17 presents the Shmoo plots for ratio R when five out of nine (a), six out of nine (b), and seven out of nine (c) cells carry "0"s. The black rectangles represent the combinations of V_{DD_WEAK} and $V_{DD\_EN\_ALL}$ at which the CUT flipped and the white rectangles present the combinations where the CUT maintained its data. From analyzing the Shmoo plots it can be seen that for every fixed value of $V_{DD\_EN\_ALL}$ the detected degree of cell weakness defined by V_{DD_WEAK} will be different depending on the chosen ratio R. For example, if $V_{DD\_EN\_ALL}$ is fixed at 1.2 V, then for $R = 5/9$, 6/9 and 7/9, the CUT will flip its state after the application of the proposed test sequence at $V_{DD\_WEAK} = 1.06V$, 1.24 V and 1.36 V respectively (white lines and circled dots in Fig. 17(a)–(c)). That corresponds to the bitline voltage levels of 0.71 V, 0.45 V and 0.28 V, respectively.

The measurement results for this case are summarized in Table I and in Fig. 17(d). It shows that by programming ratio R to be 5/9, 6/9 and 7/9 and applying the proposed RCRT test sequence, the weak cells detection threshold can be programmed from 18% to 46% of the nominal SNM.

VII. CONCLUSION

A new digitally programmable DFT technique for detection SRAM cells with compromised stability (i.e., inadequate SNM) has been introduced and a test chip measurement results proving the proposed detection concept have been presented. The proposed technique exceeds the detection range of the Data Retention Test and allows to program the pass/fail threshold of the cell stability test. The technique utilizes a set of n cells to modify the bitline voltage, which is then applied to a cell under test (CUT). The bitline voltage can be programmed by changing the ratio of "zeroes" and "ones" written in the set of n cells and simultaneously enabling their wordlines. Stability of the CUT can be determined by connecting the CUT to the partially discharged floating bitlines. The detection resolution of the technique can be further improved by increasing the number of cells forming the ratio.

ACKNOWLEDGMENT

The authors would like to thank P. van de Steeg from Philips Semiconductors and , R. Salters and M. Azimane from Philips Research Labs for fruitful technical discussions, R. van Veen and B. Kruseman from Philips Research Labs for help with the measurements, and Canadian Microelectronics Corporation (CMC) for providing the chip fabrication service.

REFERENCES

[1] J. Jayabalan and J. Povazanec, "Integration of SRAM redundancy into production test," in *Proc. IEEE Int. Test Conf. (ITC)*, Oct. 2002, pp. 187–193.

[2] Y. Zorian, "Embedded memory test and repair: Infrastructure IP for SoC yield," in *Proc. IEEE Int. Test Conf. (ITC)*, Oct. 2002, pp. 340–349.

[3] A. Pavlov, M. Sachdev, and J. Pineda de Gyvez, "An SRAM weak cell fault model and a DFT technique with a programmable detection threshold," in *Proc. IEEE Int. Test Conf. (ITC)*, Oct. 2004, pp. 1106–1115.

[4] International Technology Roadmap for Semiconductors 2003 [Online]. Available: http://public.itrs.net/

[5] R. Heald and P. Wang, "Variability in sub-100-nm SRAM designs," in *IEEE/ACM Int. Conf. Computer Aided Design (ICCAD)*, Nov. 2004, pp. 347–352.

[6] J. Lohstroh, E. Seevinck, and J. de Groot, "Worst-case static noise margin criteria for logic circuits and their mathematical equivalence," *IEEE J. Solid-State Circuits*, vol. SC-18, no. 6, pp. 803–807, Dec. 1983.

[7] A. Meixner and J. Banik, "Weak write test mode: An SRAM cell stability design for test technique," in *Proc. IEEE Int. Test Conf. (ITC)*, Nov. 1997, pp. 1043–1052.

[8] R. R. Montanés, J. Pineda de Gyvez, and P. Volf, "Resistance characterization of open defects," *IEEE Des. Test Comput.*, vol. 19, no. 5, pp. 18–26, Sep.-Oct. 2002.

[9] D.-M. Kwai, H.-W. Chang, H.-J. Liao, C.-H. Chiao, and Y.-F. Chou, "Detection of SRAM cell stability by lowering array supply voltage," in *Proc. 9th Asian Test Symp. (ATS 2000)*, Dec. 2000, pp. 268–273.

[10] A. Asenov, A. Brown, J. Davies, S. Kaya, and G. Slavcheva, "Simulation of intrinsic parameter fluctuations in decananometer and nanometer-scale MOSFETS," *IEEE Trans. Electron Devices*, vol. 50, no. 9, pp. 1837–1852, Sep. 2003.

[11] A. J. Bhavnagarwala, X. Tang, and J. D. Meindl, "The impact of intrinsic device fluctuations on CMOS SRAM cell stability," *IEEE J. Solid-State Circuits*, vol. 36, pp. 658–665, Apr. 2001.

[12] B. Cheng, S. Roy, and A. Asenov, "The impact of random doping effects on CMOS SRAM cell," in *Proc. ESSCIRC*, Leuven, Belgium, 2004, pp. 219–222.

[13] J. Wuu, D. Weiss, C. Morganti, and M. Dreesen, "The asynchronous 24 MB on-chip level-3 cache for a dual-core Itanium-family processor," in *IEEE ISSCC Dig. Tech. Papers*, Apr. 2005, pp. 488–489.

[14] E. B. Selvin, A. R. Farhang, and D. A. Guddat, "Programmable weak write test mode," U.S. Patent 6,778,450, Aug. 17, 2004.

[15] J. Segura, A. Keshavarzi, J. Soden, and C. Hawkins, "Parametric failures in CMOS ICs—A defect-based analysis," in *Proc. IEEE Int. Test Conf. (ITC)*, Oct. 2002, pp. 90–98.

[16] C. Kuo, T. Toms, B. Neel, J. Jelemensky, E. Carter, and P. Smith, "Soft-defect detection (SDD) technique for a high-reliability CMOS SRAM," *IEEE J. Solid-State Circuits*, vol. 25, no. 1, pp. 61–67, Feb. 1990.

[17] D. R. Weiss, J. Wuu, and R. J. Reidlinger, "Integrated weak write test mode (WWWTM)," U.S. Patent 6,192,001, Feb. 20, 2001.

[18] W. Schwarz, "Data retention weak write circuit and method of using same," U.S. Patent 5,835,429, Nov. 10, 1998.

[19] R. H. W. Salters, "Device with integrated SRAM memory and method of testing such a device," U.S. Patent 6,757,205, Jun. 29, 2004.

Andrei Pavlov (S'01–M'05) received an Engineer Diploma in 1993 and a Candidate of Sciences Postgraduate Degree in electrical engineering in 1998 from Moscow State Institute of Electronic Technology (Technical University), Russia, and the Ph.D. degree in electrical engineering in 2005 from the University of Waterloo, ON, Canada, working on a joint project on design, test, and reliability issues in embedded SRAMs with Philips Research Laboratories, Eindhoven, The Netherlands.

From 1999 to 2001, he was with Dalsa Corporation, Canada, working in Application-Specific Contracts R&D on test aspects of custom CCD and CMOS image sensors and cameras. Currently, he is with the Advanced Memory Circuits and Technology Integration group at Intel Corporation, Hillsboro, OR. He is the author or co-author of eight papers and the inventor or co-inventor of two patents. His research interests include DFT techniques for SRAM test, memory design and test, and power-efficient and robust mixed-signal circuit design.

Manoj Sachdev (M'87–SM'97) received the B.E. degree (with Honors) in electronics and communication engineering from the University of Roorkee, India, and the Ph.D. degree from Brunel University, U.K.

He was with Semiconductor Complex Limited, Chandigarh, India, from 1984 to 1989m where he designed CMOS Integrated Circuits. From 1989 to 1992, he worked in the ASIC division of SGS-Thomson at Agrate (Milan), Italy. From 1992 to 1998, he worked at Philips Research Laboratories, Eindhoven, The Netherlands, where he researched several aspects of VLSI testing and manufacturing. He is currently a Professor in the Electrical and Computer Engineering Department at University of Waterloo, Canada. His research interests include low-power and high-performance digital circuit design, mixed-signal circuit design, test and reliability issues of integrated circuits. He has written a book, contributed book chapters, and has published over 100 technical articles in conferences and journals. He holds more than 10 granted and several pending U.S. patents in the area of VLSI circuit design and test.

Dr. Sachdev received the Best Paper Award for his paper in the European Design and Test Conference 1997, and an Honorable Mention Award for his paper in the International Test Conference 1998.

José Pineda de Gyvez (S'88–M'90) received the Ph.D. degree from the Eindhoven University of Technology, Eindhoven, The Netherlands, in 1991.

From 1991 to 1999, he was a Faculty Member in the Department of Electrical Engineering, Texas A&M University, College Station. Since 1999, he has been a Principal Scientist with Philips Research Laboratories, Eindhoven. He was a co-editor of *IC Manufacturability: The Art of Process and Design Integration* (IEEE Press, 1998). His research interests are in the general areas of design for manufacturability and analog signal processing.

Dr. Pineda has been an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—PART I and an Associate Editor for Technology of the IEEE TRANSACTIONS ON SEMICONDUCTOR MANUFACTURING.