# Delft University of Technology

## Mitigation of sense amplifier degradation using input switching

Kraak, Daniël; Agbo, Innocent; Taouil, Mottaqiallah; Hamdioui, Said; Weckx, Pieter; Cosemans, Stefan; Catthoor, Francky; Dehaene, Wim

**Citation (APA)**
Kraak, D., Agbo, I., Taouil, M., Hamdioui, S., Weckx, P., Cosemans, S., Catthoor, F., & Dehaene, W. (2017). Mitigation of sense amplifier degradation using input switching. In *Proceedings of the 2017 Design, Automation & Test in Europe Conference & Exhibition (DATE)* (pp. 858-863). IEEE. https://doi.org/10.23919/DATE.2017.7927107

**Important note**
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

# Mitigation of Sense Amplifier Degradation Using Input Switching

Daniël Kraak  Innocent Agbo  Mottaqiallah Taouil  Pieter Weckx[1,2]  Stefan Cosemans[1]  Francky Catthoor[1,2]
Said Hamdioui                                          Wim Dehaene[2]
Delft University of Technology
Faculty of Electrical Engineering, Mathematics and CS
Mekelweg 4, 2628 CD Delft, The Netherlands
{D.H.P.Kraak, I.O.Agbo, M.Taouil, S.Hamdioui}@tudelft.nl

[1]imec vzw., Kapeldreef 75, B-3001, Leuven, Belgium
[2]Katholieke Universiteit Leuven, ESAT, Belgium
{Pieter.Weckx, Francky.Catthoor}@imec.be
wim.dehaene@esat.kuleuven.be

*Abstract*—To compensate for time-zero (due to process variation) and time-dependent (due to e.g. Bias Temperature Instability (BTI)) variability, designers usually add design margins. Due to technology scaling, these variabilities become worse, leading to the need for bigger design margins. Typically, only worst-case scenarios are considered, which will not present the actual workload of the targeted application. Alternatively, mitigation schemes can be used to counteract the variability. This paper presents a run-time design-for-reliability scheme for memory Sense Amplifiers (SAs); SAs are an integral part of any memory system and are very critical for high performance. The proposed scheme mitigates the impact of time-dependent variability due to aging by using an on-line control circuit to create a balanced workload. The simulation results show that the proposed scheme can reduce the most critical figures-of-merit, namely the offset voltage shift and the sensing delay of the SA with up to ∼40% and ∼10%, respectively, depending on the stress conditions (temperature, voltage, workload).

*Index Terms*—Mitigation, Offset voltage, zero-time variability, run-time variability, SRAM sense amplifier, sensing delay

## I. INTRODUCTION

The downscaling of CMOS technology over the past decades has significantly improved the feature size of integrated circuits. This downscaling invokes major challenges w.r.t. the device reliability [1], [2]. This is mainly caused by two sources: manufacturing and operational usage [1]. Due to imperfections in the manufacturing process, similarly produced devices will suffer from process variations, resulting in a difference in their characteristics from the intended ones. This process variation is referred to as *time-zero variability*. Variations that occur during the lifetime include *environmental variations* (such as supply voltage fluctuations and temperature variations) and *aging variations* due to, for instance, *Bias Temperature Instability* [3]; these variations are referred to as *time-dependent variability*. The impact of time-zero and time-dependent variability becomes more severe with CMOS scaling [1], [4]. If counter-measures are not taken, devices may fail [5]. Traditionally, designers use guardbanding [6], which means extra design margins are added, so the circuit is guaranteed to still function correctly under worst-case variations during the targeted lifetime. Due to the increased impact of time-zero and time-dependent variability, a bigger guardband is necessary in order to guarantee high product quality. This will negatively affect the area, speed, power

consumption, and/or yield of the design, especially because the workload dependence is not properly incorporated. In particular, only worst-case workloads are incorporated where the correlations present in representative actual workloads are lost. Alternatively, appropriate mitigation schemes can be used to effectively counteract the variability. These mitigation schemes can avoid the worst-case assumptions and instead fully incorporate the realistic conditions. They can even adapt to changing workload conditions by employing on-line control circuits. This paper focuses on the mitigation of the impact of aging on the offset voltage of the memory's Sense Amplifier (SA). The SA is very important for high performance memories, as it forms an integral and critical part of the read path delay. The SA behaviour influences the memory delay in two ways. First, a larger SA *offset* requires a larger bitline swing, which means more time must be allocated for the bitline discharge; failing to provision for sufficient swing results in failures in the field. Second, the *delay* from SA trigger to SA output (sensing delay) is on the cricital path. Note that the SA offset voltage is *at least* as important as the SA sensing delay. Therefore, understanding the impact of workload-dependent aging on the memory SA offset voltage and providing appropriate mitigation schemes is an important part of designing a robust and reliable memory system.

A lot of work has been published on the impact of aging and their countermeasures for memory cell arrays [7]–[11]. However, very limited work has been done on the characterization and workload-dependent mitigation of aging in memory peripheral circuits such as SAs. In [12], a tunable SA is presented to compensate for within-die variations. In [13], the offset voltage is monitored using a physical circuit (on-chip) to estimate the yield. In [14], the authors propose an accurate method to estimate the impact of both time-zero and time-dependent variability on the SA offset voltage; it considers the SA offset voltage dependency on temperature, voltage, and *workload*, but the mitigation is not the focus here. Prior work mainly focuses on mitigating the SA offset voltage due to time-zero variability. Run-time mitigation schemes for workload-driven time-dependent variability have not been researched.

This paper proposes a run-time design-for-reliability scheme for the SA in order to mitigate time-dependent variability. The scheme is based on providing additional input to the

SA resulting into what we refer to as *Input Switching Sense Amplifier (ISSA)*; the ISSA switches its inputs periodically in order to create an on-line control-based balanced workload. Thanks to the balanced workload, the impact of time-dependent variability on the SA offset voltage is minimized. To the best knowledge of the authors, this is the first work that addresses the workload-dependent mitigation of the SA offset voltage degradation for time-dependent variability. The contribution of this paper is as follows:

- A new mitigation scheme (i.e., ISSA) for time-dependent variability is proposed.
- Investigation of BTI impact on offset voltage specification and sensing delay of the ISSA for several workloads, varying supply voltages, and varying temperatures.
- Comparison between the most critical figures-of-merit, namely the offset voltage specification and the sensing delay, of the ISSA and the normal SA under BTI.

The rest of the paper is organized as follows: Section II provides a background and discusses BTI, the targeted standard-latch type SA, and the used method to calculate the offset voltage specification. Section III discusses the proposed methodology. Section IV evaluates the results. Finally, Section V concludes the paper.

## II. BACKGROUND

This section first discusses the BTI aging model. Thereafter, the standard latch-type SA, which is used in this work, and, finally, the method to determine the offset voltage specification are discussed.

### A. Bias Temperature Instability

Several aging mechanisms exist; e.g., Bias Temperature Instability (BTI) [3], Hot Carrier Injection [15], and Time Dependent Dielectric Breakdown [16]. BTI is considered to be the most important of them and, therefore, it is the focus of this paper [17], [18]. BTI takes place inside the MOS transistors and causes an increment in the threshold voltage ($V_{th}$). The $V_{th}$ increase happens under *negative gate stress* for PMOS transistors, which is referred to as Negative BTI (NBTI). For NMOS transistors it happens under *positive gate stress*, which is referred to as Positive BTI (PBTI).

To model BTI, this paper uses the atomistic model presented in [19]; it incorporates the dependency on the workload and is based on the capture of traps during stress and relaxation phases of BTI. Each trap contributes to the threshold voltage shift $\Delta V_{th}$. The $\Delta V_{th}$ of the transistor is the accumulated result of all gate oxide defect traps. The probabilities of the defects capture $P_C$ and emission $P_E$ are defined by [20] as follows:

$$P_C(t_{STRESS}) = \frac{\tau_e}{\tau_c + \tau_e}\left\{1 - exp\left[-(\frac{1}{\tau_e} + \frac{1}{\tau_c})t_{STRESS}\right]\right\} \quad (1)$$

$$P_E(t_{RELAX}) = \frac{\tau_c}{\tau_c + \tau_e}\left\{1 - exp\left[-(\frac{1}{\tau_e} + \frac{1}{\tau_c})t_{RELAX}\right]\right\} \quad (2)$$

Here, $\tau_c$ and $\tau_e$ are the mean capture and emission time constants, $T_{STRESS}$ the stress period, and $T_{RELAX}$ the relaxation period. The impact of temperature is also included in the model [19], [21].
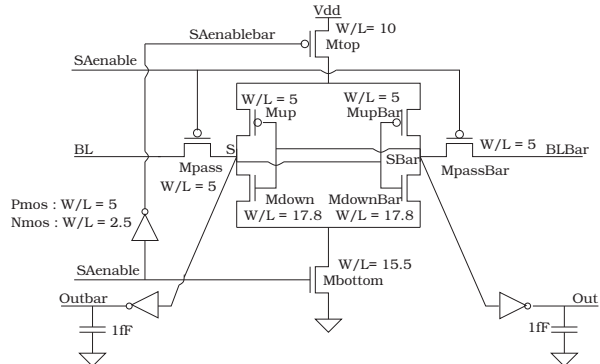


Fig. 1: Standard Latch-Type Sense Amplifier

### B. Sense Amplifier

The standard latch-type SA shown in Figure 1 will be used as a case study in this work; in principle, the proposed scheme can be applied to other types of SAs, such as look-ahead type SA [22], double-tail latch-type SA [23], etc. The working principal of the SA of Figure 1 is as follows: during read operations it amplifies a small voltage difference between bitlines BL and BLBar. Its operation consists of two stages. In the first stage, when SAenable is low, the voltage swing on BL and BLBar is passed to the internal nodes S and SBar of the SA. In the second phase, when SAenable is high, the amplification takes place by the cross-coupled inverters. During this amplification the pass transistors disconnect the internal nodes from BL and BLBar. The cross coupled inverters get current through Mtop and Mbottom and produce the outputs Out and Outbar.

### C. Offset Voltage Specification

The method to calculate the offset voltage specification is taken from [14]. It considers both the effects of time-zero (i.e., local process variation) and time-dependent variability (i.e. variation due to aging, temperature, and voltage) on the offset voltage. The offset voltage is determined by Monte Carlo simulation. During each simulation the offset voltage of one specific sample is determined using a binary search on its inputs. From the Monte Carlo simulations the average and standard deviations of the samples are calculated. The offset voltage of SAs typically follows a normal distribution and the relation between this distribution and failure rate is as follows [14]:

$$\int_{V_{in}=-V_{Offset}}^{V_{Offset}} \mathcal{N}(\mu_{MC}, \sigma_{MC}) = 1 - f_r \quad (3)$$

Here, $V_{in}$ represents the input voltage of the SA, $V_{offset}$ the offset voltage specification, $\mathcal{N}$ a normal distribution of the offset voltage with mean $\mu_{MC}$ and standard deviation $\sigma_{MC}$. In this equation, all SA instantiations that require an input offset voltage outside the range [-$V_{offset}$, $V_{offset}$] result in a failure. Using Equation 3, it is possible to calculate the offset voltage specification of the SA for a certain failure rate. In this work, a failure rate target of $f_r$=$10^{-9}$ is assumed;
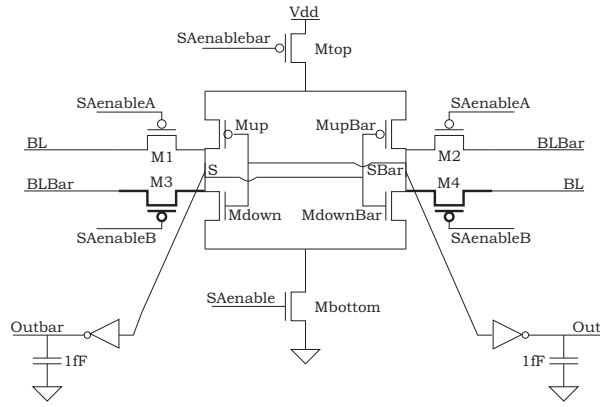
Fig. 2: Input Switching Sense Amplifier

TABLE I: Truth table for SAenableA and SAenableB

| Switch | SAenableBar | SAenableA | SAenableB |
|--------|-------------|-----------|-----------|
| 0 | 0 | 1 | 1 |
| 0 | 1 | 0 | 1 |
| 1 | 0 | 1 | 1 |
| 1 | 1 | 1 | 0 |

thus, targeting an application with *high reliability* requirement. This required failure rate leads to a $V_{offset} = 6.1 \cdot \sigma_{MC}$ (roughly $6\sigma$) for a distribution with a mean of 0 [14]. Besides process variation, the distribution of the offset voltage depends on temperature, supply voltage, and the used workload and stress time. Therefore, the offset voltage specification will be different when these conditions vary. Hence, we clearly need an on-line control approach to effectively deal with this. This work uses the same approach as in [14] to solve Equation 3 numerically to obtain the offset voltage specification.

## III. PROPOSED METHODOLOGY

It has been shown that the SA suffers from the largest increase in offset voltage when an *unbalanced* workload (reading more ones than zeros, or vice versa), is applied for a long lifetime application [14]. When mostly zeros (ones) are read, transistors Mdown (MdownBar) and MupBar (Mup) are the most stressed. This results in an increased $V_{th}$ for these transistors, leading to a shift in the required offset voltage of the SA. This increased offset voltage negatively impacts the time needed to produce the read result as more time is needed to discharge the bitlines. It is also shown in [14] that creating a *balanced* workload (i.e., a workload where the amount of read 0s and read 1s is equal) leads to the minimum impact on the offset voltage. This can be explained by the fact that a lower input voltage is required when the cross-coupled inverter pair is balanced. Therefore, this paper proposes the Input Switching Sense Amplifier (ISSA), where the SA switches its inputs periodically in order to create an on-line control-based balanced workload at its internal nodes; this is in contrast to the normal Non Switching Sense Amplifier (NSSA). Next, the ISSA design will be presented and, subsequently, its control logic.
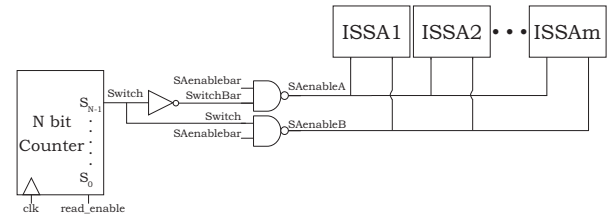


Fig. 3: Control Logic for ISSAs

### A. Input Switching Sense Amplifier

Figure 2 depicts the structure of the ISSA. A second pair of pass transistors, M3 and M4, is added compared to the standard latch-type SA (Figure 1). Pass transistor M3 (M4) connects BLBar (BL) to S (SBar). This makes it possible to switch the inputs of the SA, and connect BLBar to S and BL to SBar. This requires, however, additional control circuitry. Pass transistors M1 and M2 are controlled by signal SAenableA and pass transistors M3 and M4 by signal SAenableB. When SAenableA is low/enabled, pass transistors M1 and M2 forward the voltage level on BL and BLBar to the internal nodes. Note that in this case the SA operates in the same way as that of the standard latch-type SA; the signal SAenableB is disabled/high (i.e., M3 and M4 are off).

When the inputs of the SA are switched (i.e., SAenableB is low/enabled and SAenableA high/disabled), the SA will effectively read the opposite value. Hence, by controlling this switching, it is possible to balance the amount of zeros and ones read by the internal nodes of the SA. This will lead to a more balanced workload for the SA and mitigate the degradation of the offset voltage at the penalty of a limited area overhead (discussed in Section IV-C). It is worth noting that in the case that the inputs of the SA are switched, the final read value needs to be inverted (e.g., using additional circuitry).

### B. Control Logic

Figure 3 illustrates the control logic. Two NAND gates are used to generate SAenableA and SAenableB from the original SAenable(bar) and the Switch signal. The Switch signal is generated by an N-bit counter (updated only during reads, controlled by read_enable) and used to decide when the inputs of the SA should be swapped; for example each $2^{N-1}$ reads. Table I contains the truth table for SAenableA and SAenableB. When Switch is low (high), only SAenableA (SAenableB) is able to change its value. SAenableB (SAenableA) is always high in this case, to make sure the corresponding pass transistors M3 and M4 (M1 and M2) are switched off.

## IV. SIMULATION RESULTS

In this section the performed experiments and the obtained results are presented.

### A. Performed Experiments

The circuits of the NSSA (Figure 1) and the ISSA (Figure 2) are implemented using the 45nm PTM high-performance
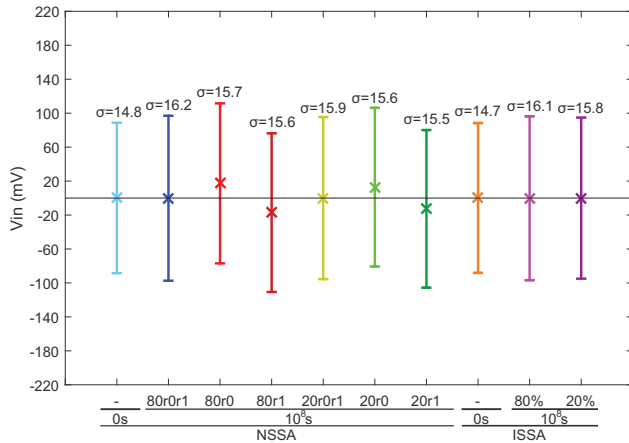
*2017 Design, Automation and Test in Europe (DATE)*

Fig. 4: Workload impact on offset voltage

| Scheme | Time (s) | Workload | Offset Voltage | | | Delay (ps) |
|---|---|---|---|---|---|---|
| | | | $\mu$ (mV) | $\sigma$ (mV) | spec. (mV) | |
| NSSA | 0 | - | 0.1 | 14.8 | 90.2 | 13.6 |
| NSSA | $10^8$ | 80r0r1 | -0.2 | 16.2 | 99.0 | 14.2 |
| NSSA | $10^8$ | 80r0 | 17.3 | 15.7 | 111.5 | 14.3 |
| NSSA | $10^8$ | 80r1 | -17.2 | 15.6 | 110.6 | 14.0 |
| NSSA | $10^8$ | 20r0r1 | -0.08 | 15.9 | 97.2 | 14.1 |
| NSSA | $10^8$ | 20r0 | 12.8 | 15.6 | 106.3 | 14.2 |
| NSSA | $10^8$ | 20r1 | -12.7 | 15.5 | 105.5 | 14.0 |
| ISSA | 0 | - | 0.1 | 14.7 | 89.9 | 13.9 |
| ISSA | $10^8$ | 80% | -0.2 | 16.1 | 98.3 | 14.5 |
| ISSA | $10^8$ | 20% | -0.09 | 15.8 | 96.6 | 14.3 |

library [24]. As a case study, an 8-bit counter is used for the control logic of the ISSA; the most significant bit of the counter is used to generate the Switch signal. As a result, the inputs of the SA are swapped each 128 reads. The circuits are simulated using Spectre and for each Monte Carlo simulation 400 iterations are performed. Two sets of experiments are performed to compare the performance of the NSSA with the ISSA:

- **Impact on offset voltage spec.:** In this experiment the combined effect of time-zero and time-dependent variability on the offset voltage is analyzed. During the experiment, the impact of six workloads (which will be defined later), supply voltage variations (-10% $V_{dd}$, nom. $V_{dd}$=1.0V, +10% $V_{dd}$), and temperature variations (25°C, 75°C, 125°C) are analyzed.
- **Impact on sensing delay:** In this experiment, the impact of time-dependent variability on the sensing delay is analyzed. The sensing delay is the time needed for the SA to complete its operation. It is the time between the activation of the SA (when SAEnable rises to 50% of $V_{dd}$) and when the result is produced at the output (when Out or Outbar rises to 50% of $V_{dd}$). With this experiment the impact of the added circuitry on the sensing delay is analyzed for the ISSA.

The six different workloads used to evaluate the impact of workload are: 80r0r1, 80r0, 80r1, 20r0r1, 20r0, and 20r1; the first number (80 or 20) indicates the activation rate of the SA. For example, for 80, it is assumed that 80% of the time a read operation is performed. The 80% activation rates are used to mimic a read intensive application and the 20% activiation rates mimic a less read intensive one. After the activation rate, the read sequence is indicated; three read sequences are used: r0r1 (50% of the reads are 0 and 50% of the reads are 1), r0 (all reads are 0), and r1 (all reads are 1). The sequences r0 and r1 are used to mimic unbalanced workloads and the sequence r0r1 to mimic a balanced workload. Note that for the ISSA all three workloads 80r0, 80r1, and 80r0r1 are compiled by the design-for-reliability scheme into the same balanced workload 80r0r1; hence, we denote these with just the activation rate

80%. The same applies to the three workloads 20r0, 20r1, and 20r0r1.

### B. Simulation Results

**Impact on offset voltage spec.**
Three different experiments were performed in order to investigate the workload dependency, the supply voltage dependency, and the temperature dependency.

**Workload dependency -** Figure 4 shows the offset voltage distributions for different workloads at nominal temperature and $V_{dd}$. The averages (i.e., denoted by the 'x' marker) and the +/-6$\sigma$ values (i.e., denoted by the edges of the vertical lines) of the distributions are shown. The $\mu$, $\sigma$, and corresponding offset voltage specification of the distributions can be also found in Table II. It can be seen that the distributions of the NSSA shift up or down for unbalanced workloads (80r0, 80r1, 20r0, 20r1), leading in all cases to a higher required offset voltage specification. This means additional time will be needed to appropriately discharge the bitlines, which makes the overall memory slower. The ISSA, however, significantly reduces this shift; e.g. the offset voltage spec. of the NSSA is 111.5mV for the sequence 80r0 at t=$10^8$s, while this is just 98.3mV for the ISSA; a reduction of ~12%. Hence, less time is needed for the bitline discharge, which makes the overall memory faster. What can be noticed too is that the spread ($\sigma$) of the distributions of the NSSA increase with aging. This increase in spread is only marginally dependent on the workload; for instance, the $\sigma$ is 15.9mV for the sequence 20r0r1 and 16.2mV for the sequence 80r0r1 at t=$10^8$s. Note that the difference in spreads between those of the NSSA and those of the ISSA are marginal; hence, the design-for-reliability scheme of the ISSA does not impact the spread. Nevertheless, the ISSA always brings the mean ($\mu$) of the distributions to 0.

**Voltage dependency -** Figure 5 shows the voltage dependency of the offset voltage at nominal temperature for t=$10^8$s. The $\mu$, $\sigma$, and corresponding offset voltage specification of the distributions can also be found in Table III. For the NSSA, it can be seen that the increase in offset voltage specification is significant at higher $V_{dd}$ (up to ~35% for unbalanced workloads at t=$10^8$s) and is about 3x more than that at lower $V_{dd}$. For the ISSA, however, the increase in the offset voltage
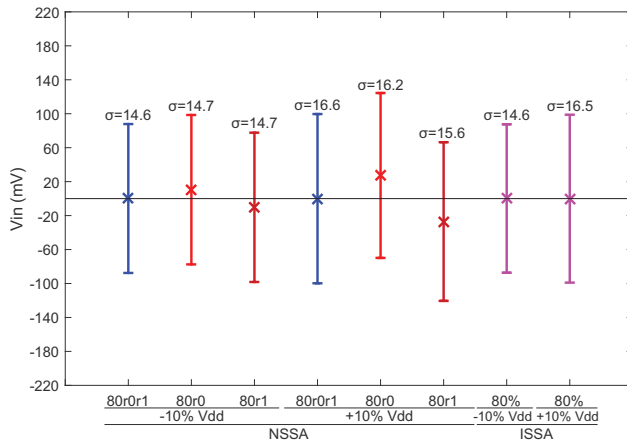
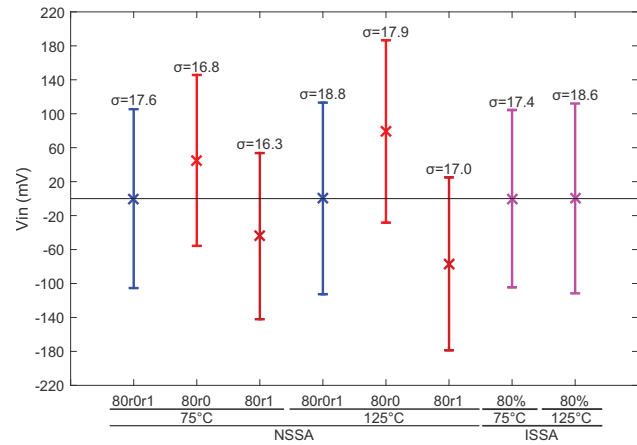Fig. 5: Voltage impact on offset voltage at t=$10^8$s



Fig. 6: Temperature impact on offset voltage at t=$10^8$s

TABLE III: Voltage impact on offset voltage and delay

| Scheme | Time (s) | Workload | Supply Voltage | Offset Voltage | | | Delay (ps) |
|---|---|---|---|---|---|---|---|
| | | | | $\mu$ (mV) | $\sigma$ (mV) | spec. (mV) | |
| NSSA | 0 | - | -10% | 0.1 | 14.5 | 88.6 | 17.2 |
| NSSA | 0 | - | +10% | 0.8 | 15.0 | 91.6 | 11.3 |
| NSSA | $10^8$ | 80r0r1 | -10% | 0.1 | 14.6 | 89.3 | 17.6 |
| NSSA | $10^8$ | 80r0r1 | +10% | -0.07 | 16.6 | 101.5 | 12.0 |
| NSSA | $10^8$ | 80r0 | -10% | 10.5 | 14.7 | 98.5 | 17.7 |
| NSSA | $10^8$ | 80r0 | +10% | 27.3 | 16.2 | 124.4 | 12.2 |
| NSSA | $10^8$ | 80r1 | -10% | -10.3 | 14.7 | 98.2 | 17.3 |
| NSSA | $10^8$ | 80r1 | +10% | -27.0 | 15.6 | 120.4 | 11.9 |
| ISSA | 0 | - | -10% | 0.1 | 14.5 | 88.5 | 17.4 |
| ISSA | 0 | - | +10% | 0.08 | 14.9 | 91.1 | 11.6 |
| ISSA | $10^8$ | 80% | -10% | 0.1 | 14.6 | 89.0 | 17.8 |
| ISSA | $10^8$ | 80% | +10% | -0.07 | 16.5 | 100.7 | 12.3 |

TABLE IV: Temperature impact on offset voltage and delay

| Scheme | Time (s) | Workload | Temp. (°C) | Offset Voltage | | | Delay (ps) |
|---|---|---|---|---|---|---|---|
| | | | | $\mu$ (mV) | $\sigma$ (mV) | spec. (mV) | |
| NSSA | 0 | - | 75 | 0.09 | 15.1 | 92.2 | 17.1 |
| NSSA | 0 | - | 125 | 0.08 | 15.3 | 93.6 | 21.3 |
| NSSA | $10^8$ | 80r0r1 | 75 | -0.03 | 17.6 | 107.3 | 19.2 |
| NSSA | $10^8$ | 80r0r1 | 125 | 0.2 | 18.8 | 114.9 | 25.7 |
| NSSA | $10^8$ | 80r0 | 75 | 45.0 | 16.8 | 145.6 | 19.9 |
| NSSA | $10^8$ | 80r0 | 125 | 79.1 | 17.9 | 186.5 | 29.0 |
| NSSA | $10^8$ | 80r1 | 75 | -44.2 | 16.3 | 142.0 | 18.3 |
| NSSA | $10^8$ | 80r1 | 125 | -76.8 | 17.0 | 178.6 | 23.5 |
| ISSA | 0 | - | 75 | 0.08 | 15.0 | 91.6 | 17.5 |
| ISSA | 0 | - | 125 | 0.08 | 15.2 | 92.9 | 21.7 |
| ISSA | $10^8$ | 80% | 75 | -0.02 | 17.4 | 106.3 | 19.5 |
| ISSA | $10^8$ | 80% | 125 | 0.2 | 18.6 | 113.9 | 26.0 |

specification does not exceed ∼10% at higher $V_{dd}$ and ∼0.5% at lower $V_{dd}$ compared to the NSSA at t=0s. This shows the superiority of the ISSA design. Note that at a higher $V_{dd}$ the spread is higher in all cases and that the shift of the distribution in case of unbalanced workloads is higher.

**Temperature dependency -** Figure 6 shows the temperature dependency of the offset voltage at nominal $V_{dd}$ for t=$10^8$s. The results are shown for temperatures of 75°C and 125°C. The $\mu$, $\sigma$, and corresponding offset voltage specification of the distributions can also be found in Table IV. The impact of temperature on the offset voltage specification is much higher than that of $V_{dd}$. For the NSSA, it can be seen that the increase in offset voltage specification is significant at 125°C (up to ∼99% for unbalanced workloads at t=$10^8$s) and is about 2x more than that at 75°C. For the ISSA, however, the increase in the offset voltage specification does not exceed ∼22% at 125°C and ∼16% at 75°C compared to the NSSA at t=0s. Compared to the NSSA at t=$10^8$s and 125°C for unbalanced workloads, the ISSA reduces the offset voltage with about 40%; hence, the ISSA performs significantly better.

**Impact on sensing delay**
For the impact on sensing delay, the workload dependency, voltage supply dependency, and temperature dependency were

also investigated. The average delays can be found for each of these experiments in the last column of Tables II, III, and IV, respectively.

When looking at workload impact (Table II), the NSSA shows the largest increase in delay for unbalanced workloads, reaching in the worst-case 14.3ps. For the ISSA the delay is 14.5ps; hence, the delay overhead is negligible as compared with the gain the ISSA achieves w.r.t. the offset voltage specification.

When looking at the impact of supply voltage (Table III), the sensing delay is the highest at -10% $V_{dd}$. In this case, the highest sensing delay at t=$10^8$s for the NSSA is 17.7ps, while this is 17.8ps for the ISSA. This stresses again the fact that although the ISSA requires some extra circuitry for its implementation, its delay overhead is negligible, while its added value w.r.t. the offset voltage specification is significant.

When looking at the temperature dependency (Table IV), it can be noticed that the degradation of the sensing delay is significant at high temperatures. This is also illustrated in Figure 7, which shows the delay increase due to aging for both the NSSA (two workloads) and the ISSA at 125°C. It can be seen that the delay of the NSSA increases so fast for the sequence 80r0 that it becomes even higher than that of the ISSA; at t=$10^8$s the delay of the ISSA is ∼10% lower than that

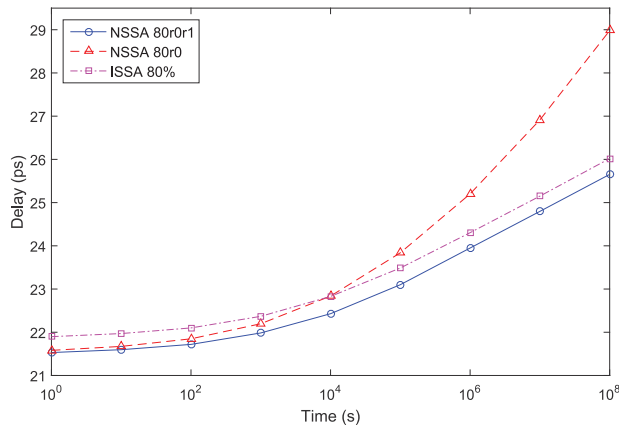Fig. 7: Delay versus aging at T=125°C

of the NSSA. This indicates that not only the delay overhead of the ISSA is negligible (see also Table IV), but also that in some cases it even performs better than the NSSA.

Overall, we can conclude from the sensing delay experiments that the additional hardware of the ISSA adds a negligible delay overhead and, furthermore, under high stress (unbalanced workloads, high temperature) and for long lifetime applications ($t > 10^8$s), the ISSA will even perform better.

*C. Discussion*

This paper proposes a run-time scheme to mitigate the degradation of the SA; it is based on the creation of a balanced workload for the SA. The results show that the proposed scheme significantly reduces the offset voltage increase (up to ~40%) due to time-dependent variability. This means less time is needed to discharge the bitlines, which leads to a faster memory. This experiment assumed a random input pattern, which is a reasonable assumption. Therefore, this design-for-reliability leads to a more optimal and reliable design and provides a good alternative to guardbanding.

The question is now: what is the impact of the scheme in terms of delay, area, and energy overhead? First, when looking at the delay overhead of the scheme, the results show it is negligible; under high stress and for long lifetime applications the scheme even performs better. The final read value still needs to be inverted, when the inputs of the ISSA are swapped. However, this adds negligible overhead.

Second, the implementation of the control logic needs one counter and three extra gates. These, however, can be shared by multiple columns of SAs. The area overhead is, therefore, very marginal as the area of a memory is mainly dominated by the cell matrix (typically > 70%), followed by the address decoders.

Third, the energy overhead is also negligible. The scheme consists of one counter and a couple of gates (shared by multiple columns of SAs). The counters are active only during the read operations; hence, during memory write operations, the counters are not consuming any dynamic power.

In conclusion, the proposed scheme has a large gain with only negligible overheads.

## V. CONCLUSION

This work clearly shows that run-time mitigation schemes can be a very good alternative to the traditional guardbanded designs. Not only they can provide optimal design solutions, but they can even extend the lifetime of the devices by balancing the workload. This is extremely important for cutting edge technology as they suffer from reduced lifetime and increased failure rate.

## REFERENCES

[1] S. Borkar, "Microarchitecture and design challenges for gigascale integration," in *Intl. Sympos. on Microarchitecture*, Dec 2004, pp. 3–3.

[2] S. Hamdioui *et al.*, "Reliability challenges of real-time systems in forthcoming technology nodes," in *Design, Automation Test in Europe Conference Exhibition (DATE), 2013*, March 2013, pp. 129–134.

[3] B. Kaczer *et al.*, "Atomistic approach to variability of bias-temperature instability in circuit simulations," in *IRPS*, April 2011, pp. XT.3.1–XT.3.5.

[4] J. Srinivasan *et al.*, "The impact of technology scaling on lifetime reliability," in *DSN*, June 2004, pp. 177–186.

[5] V. Huard *et al.*, "From bti variability to product failure rate: A technology scaling perspective," in *IRPS*, April 2015, pp. 6B.3.1–6B.3.6.

[6] K. Bowman *et al.*, "Circuit techniques for dynamic variation tolerance," in *DAC*, July 2009, pp. 4–7.

[7] B. Cheng *et al.*, "Impact of nbti/pbti on sram stability degradation," *IEEE Electron Device Letters*, vol. 32, no. 6, pp. 740–742, June 2011.

[8] A. Gebregiorgis *et al.*, "Aging mitigation in memory arrays using selfcontrolled bit-flipping technique," in *20th Asia and South Pacific Design Automation Conference*, Tokyo, Japan, November 2015.

[9] M. Khan *et al.*, "Bias temperature instability analysis of finfet based sram cells," in *DATE*, Dresden, Germany, March 2014.

[10] S.K. Krishnappa *et al.*, "Comparative bti reliability analysis of sram cell designs in nano-scale cmos technology," in *Quality Electronic Design (ISQED), 2011 12th International Symposium on*, March 2011, pp. 1–6.

[11] P. Pouyan *et al.*, "Process variability-aware proactive reconfiguration technique for mitigating aging effects in nano scale sram lifetime," in *2012 IEEE 30th VLSI Test Symposium (VTS)*, April 2012, pp. 240–245.

[12] M.H. Abu-Rahma *et al.*, "Characterization of sram sense amplifier input offset for yield prediction in 28nm cmos," in *2011 IEEE Custom Integrated Circuits Conference (CICC)*, Sept 2011, pp. 1–4.

[13] J. Vollrath, "Signal margin analysis for dram sense amplifiers," in *Electronic Design, Test and Applications, 2002. Proceedings. The First IEEE International Workshop on*, 2002, pp. 123–127.

[14] I. Agbo *et al.*, "Quantification of sense amplifier offset voltage degradation due to zero- and run-time variability," in *IEEE Computer Society Annual Symposium on VLSI*, Pittsburgh, U.S.A., July 2016.

[15] F. Cacho *et al.*, "Hot carrier injection degradation induced dispersion: Model and circuit-level measurement," in *2011 IEEE International Integrated Reliability Workshop Final Report*, Oct 2011, pp. 137–141.

[16] K.B. Yeap *et al.*, "A realistic method for time-dependent dielectric breakdown reliability analysis for advanced technology node," *IEEE Transactions on Electron Devices*, vol. 63, no. 2, pp. 755–759, Feb 2016.

[17] S. Bhardwaj *et al.*, "Predictive modeling of the nbti effect for reliable design," in *CICC*, Sept 2006, pp. 189–192.

[18] R. Newhart, "Early reliability modeling for aging and variability in silicon system, ibm view, ermavss workshop," in *DATE*, 2016.

[19] B. Kaczer *et al.*, "Origin of nbti variability in deeply scaled pfets," in *IRPS*, May 2010, pp. 26–32.

[20] M. Toledano-Luque *et al.*, "Response of a single trap to ac negative bias temperature stress," in *IRPS*, April 2011, pp. 4A.2.1–4A.2.8.

[21] T. Grasser *et al.*, "Analytic modeling of the bias temperature instability using capture/emission time maps," in *Electron Devices Meeting (IEDM), 2011 IEEE International*, Dec 2011, pp. 27.4.1–27.4.4.

[22] T. Asano *et al.*, "Low-power design approach of 11fo4 256-kbyte embedded sram for the synergistic processor element of a cell processor," *IEEE Micro*, vol. 25, no. 5, pp. 30–38, Sept 2005.

[23] D. Schinkel *et al.*, "A double-tail latch-type voltage sense amplifier with 18ps setup+hold time," in *2007 IEEE International Solid-State Circuits Conference. Digest of Technical Papers*, Feb 2007, pp. 314–605.

[24] "Predictive technology model," http://ptm.asu.edu/.