Mitigation of Sense Amplifier Degradation Using Skewed Design

Daniël Kraak Mottaqiallah Taouil Said Hamdioui
Delft University of Technology
Mekelweg 4, 2628 CD Delft, The Netherlands
{D.H.P.Kraak, M.Taouil, S.Hamdioui}@tudelft.nl

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

***Abstract*—Designers typically add design margins to semiconductor memories to compensate for aging. However, the aging impact increases with technology downscaling, leading to the need for higher margins. This results into a negative impact on area, yield, performance, and power consumption. As an alternative, mitigation schemes can be developed to reduce such impact. This paper proposes a mitigation scheme for the memory's sense amplifier (SA); the scheme is based on creating a skew in the relative strengths of the SA's cross-coupled inverters during design. The skew is compensated by aging due to unbalanced workloads. As a result, the impact of aging on the SA is reduced. To validate the mitigation scheme, the degradation of the sense amplifier is analyzed for several workloads. The experimental results show that the proposed mitigation scheme reduces the degradation of the sense amplifier's critical figure-of-merit, the offset voltage, with up to 26%.**

***Index Terms*—sense amplifier, memory, aging, mitigation**

I. Introduction

Aggressive downscaling has been the main contributor to the improved performance and functionality of Integrated Circuits (ICs) over the past decades. However, due to several challenges, both the rate and benefits of downscaling have started to slow down [1]. One major challenge of downscaling is that it worsens the reliability of ICs due to increased time-dependent variability [2]; this consists of variations that occur during the operational lifetime of the IC. It includes environmental variations, such as fluctuations in the supply voltage and temperature, and *aging* variations due to, for instance, Bias Temperature Instability (BTI) [3]. Traditionally, designers use worst-case guard-banding to tolerate this variability, meaning extra design margins are added to guarantee a reliable operation of the IC during its intended lifetime. However, these extra margins cause penalties in area, performance, and power consumption. As time-dependent variability increases with downscaling, higher margins are required and, thus, these penalties increase. Alternatively, designers can incorporate *mitigation schemes* into the design that reduce the impact of time-dependent variability. In this work, we address aging in the *Sense Amplifier* (SA) of semiconductor memories. The SA speeds up read operations by amplifying a small voltage signal generated by the memory cells. Hence, the SA is a crucial component for the memory's read operation. Degradation of the SA may lead to the incorrect read-out of the memory cell [4, 5]. Moreover, several publications have reported that the SA is very sensitive to aging [6, 7]. For example, its offset voltage metric may degrade up to 100% [7]. Hence, understanding the impact of aging on the SA and providing mitigation schemes are crucial aspects of reliable memory design.

This work was supported through the projects TRACE (CATRENE, Grant 16ES0488K-16ES0502, 16ES0737) and PRYSTINE (ECSEL, Grant 783190).

Prior work has mainly focused on mitigating the impact of aging on the memory cell array [8–11]. Most of these works try to balance the probability of writing zeroes and ones to the memory cells, as this minimizes their degradation. Contrary to the memory cells, the peripheral circuitry has received significantly less attention. Prior work on the peripheral circuitry has mainly focused on estimating the impact of aging; examples are the SAs [12] and the address decoders [13]. To our best knowledge, only two publications exist on mitigation schemes for the SAs [5, 14]. In [5] a framework is proposed that finds a memory array granularity (i.e. number of banks) that results in a minimized SA degradation for a target application. This scheme is, however, not directly applied to the SA circuit. Hence, additional mitigation schemes that directly target the SA can still be added to further reduce its aging. The authors of [14] propose the Input Switching Sense Amplifier (ISSA); it is based on periodically switching the SA's differential inputs to create a balanced workload. A drawback of this scheme is that it requires additional control circuitry to enable the switching. Moreover, the output value of the SA still needs to be inverted in case the inputs are switched, resulting in a non-negligble penalty for the memory's overall delay.

This paper proposes the *skewed SA* to mitigate the aging of the memory's SA. The skewed SA has a *skew* in the relative strengths of the SA's cross-coupled inverters. The skew between the cross-coupled inverters is compensated by aging due to unbalanced workloads. As a result, the skewed SA has a lower degradation compared with the standard SA and, thus, it improves the lifetime and reliability of the memory. In short, the contributions of this work are as follows:

1) It proposes the skewed SA design to mitigate the degradation of the memory's SA.
2) It validates the performance of the proposed scheme by investigating, as a case study, the aging impact for different workloads with and without mitigation.
3) It uses an *industrial* $14\,\mathrm{nm}$ FinFET SA design and a *calibrated* BTI aging model for the circuit simulations; thus, it provides accurate and relevant results.

The rest of this paper is organized as follows. Section II provides the background. Section III discusses the proposed mitigation scheme. Section IV presents a case study to validate the mitigation scheme. Finally, Section V concludes this paper.

II. Background

This section first discusses the SA, followed by its metrics.

A. Sense Amplifier

Memories contain SAs that amplify a small voltage generated by the memory cells in order to speed up the read operation. Fig. 1 shows the schematic of the standard latch-type SA. It is a popular design due to its high performance and, therefore, it is used in this work. The operation of the standard latch-type SA consists of two phases, namely the *sensing phase* and the *amplification phase*. In the sensing phase, the SA senses the voltage difference on BL and $\overline{\text{BL}}$ that is generated by the memory cell. In this phase, SAEnable is low, so the pass transistors Mpass and $\overline{\text{MPass}}$ forward the voltage difference on BL and $\overline{\text{BL}}$ to the SA's internal nodes SAOut and $\overline{\text{SAOut}}$. In the amplification phase, the SA amplifies the sensed voltage difference to a full-swing read value. In this phase, SAEnable is high, so the pass transistors disconnect BL and $\overline{\text{BL}}$ from the SA and the voltage difference between SAOut and $\overline{\text{SAOut}}$ is amplified by the cross-coupled inverters of the SA. Finally, the full-swing read value is available at node SAOut and its complementary value at $\overline{\text{SAOut}}$.

B. Sense Amplifier Metrics

Two often-used metrics of the SA are its *sensing delay* and its *offset voltage* [6, 7]. The sensing delay is the delay of the SA to amplify the sensed voltage difference to a full-swing value. The offset voltage is the differential input voltage that results in a differential output voltage of zero for the SA. Due to process variation and *aging* [7], there is typically a mismatch in the strengths of the cross-coupled inverter pairs of the SA. As a result, the offset voltage is never zero. To guarantee a correct readout of the memory cell, the bitline swing should exceed the SA's offset voltage. A lower offset voltage leads to a higher performance (i.e. higher speed and lower power), as it means a lower bitline swing is required. Arguably, the SA's offset voltage is more important than its sensing delay. Typically, the sensing delay is only a small fraction of the memory's total read delay (e.g. $\sim$2% [15]) and, thus, its degradation has a marginal overall impact. Contrarily, generating the required bitline swing takes a considerate amount of time due to the cell's weak drive strength. Therefore, this work focuses on the offset voltage metric of the SA.

Each fabricated SA has a different offset voltage due to the stochastic nature of process variation and aging [7]. Hence, the distribution of the offset voltage should be analyzed. In order to provide a quantifying metric for this distribution, we use the *offset voltage specifation* and the method to determine it introduced in [7]. The method is based on performing Monte Carlo simulations during which process variation and aging are injected into the SA. During each simulation, the offset voltage is measured. Next, a fitted normal distribution of the offset voltage is obtained using the mean and spread of the measured offset voltages. Using this normal distribution, the offset voltage spec for a target failure rate is then calculated;

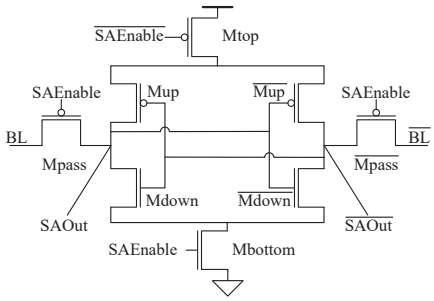


Fig. 1: Latch-type sense amplifier.

i.e. the generated bitline swing should be equal or higher than this spec to guarantee the target failure rate. In this work a target failure rate $f_r = 10^{-9}$ is used, similarly as in [7].

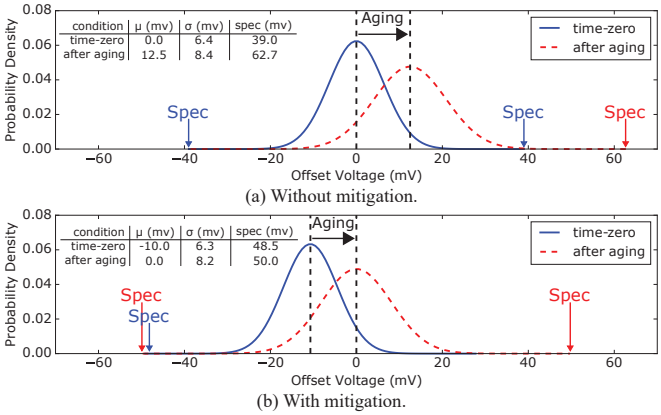
III. Proposed Mitigation Scheme

This section first introduces the concept of the mitigation scheme. Next, it discusses the skewed SA design.

A. Concept

It was demonstrated in [7] that the SA's offset voltage has the highest degradation in case the SA is stressed by an *unbalanced* workload; i.e. it reads more ones than zeros or vice versa. The impact on the fitted normal distribution of the SA's offset voltage due to an unbalanced workload (dominated by read 1), is illustrated in Fig. 2a. Here, the distributions are shown before aging (time-zero) and after aging. In addition, the mean (the striped vertical bars) and the corresponding offset voltage spec are indicated for each distribution. Note that these results have been obtained from a calibrated model at SPICE level, so these results should be sufficiently accurate. The figure reveals that the mean of the distribution shifts away from zero due to aging. This happens because unbalanced workloads cause an unequal stress for the SA's cross-coupled inverters. For instance, in case the SA reads mainly ones, transistors $\overline{\text{Mdown}}$ and Mup in Fig. 1 are stressed the most. Due to the shift of the mean, the offset voltage spec degrades significantly for unbalanced workloads. This needs to be taken into account during design, meaning sufficient bitline swing margin should be added to the memory. Alternatively, methods can be developed to mitigate the degradation, which is the focus of this work.

To mitigate the aforementioned degradation, this paper proposes to shift the SA's mean offset voltage at time-zero (i.e. during design). The effect of this scheme is illustrated in Fig. 2b; the mean of the distribution at time-zero is now shifted to $-10.0\,\text{mV}$. However, with aging the distribution's mean shifts towards zero. This results in a lower offset spec compared to the case without mitigation from Fig. 2a. Note that in case an opposite workload is applied (mainly zeros are read, instead of ones), the distribution would shift to the left, resulting in a higher spec compared with the case without mitigation. Hence, this method can be efficiently applied if the workload has a certain degree of predictability. It is also worth noting that this method causes a higher spec at time-zero due to the shift of the mean. However, the spread of



(a) Without mitigation.

(b) With mitigation.

Fig. 2: Degradation of the offset voltage (a) without and (b) with mitigation.

the offset voltage also increases with aging [7]. As a result, the spec at time-zero is still lower than the spec after aging in Fig. 2b. Hence, the increasing spread due to aging is able to compensate for the shift of the mean at time-zero up to a certain degree. In addition, prior works have shown that the bitline swing stays roughly constant with aging [6, 15].

To create a shift in the mean offset voltage at time-zero, we propose the *skewed SA*. Its design is discussed next.

B. Skewed Sense Amplifier Design

Fig. 3 depicts the design of the skewed SA. An additional PMOS transistor is added either to node SAout or $\overline{\text{SAout}}$ compared with the standard latch-type SA from Fig. 1. Note that this transistor is turned off; i.e. the gate and source are hard-wired to V_{DD}. The purpose of this transistor is to create an additional capacitance. As a result, the inverter will be slower at driving its node during a read-operation. Hence, this creates a shift in the mean of the SA's offset voltage. The addition of transistor Mskew mitigates workloads that favor read ones, since this makes the SA weaker at reading zeros. On the other hand, the addition of $\overline{\text{Mskew}}$ mitigates workloads that favor read zeros.

The impact of the additional capacitance on the mean offset voltage of the skewed SA can be mathematically modelled as follows. At the start of the SA's operation, pull-down transistors Mdown and $\overline{\text{Mdown}}$ discharge the pre-charged nodes SAOut and $\overline{\text{SAOut}}$. In case the offset voltage is supplied at the inputs of the SA, both pull-down transistors will discharge the nodes to the SA's metastable point. For a SA with a balanced cross-coupled inverter pair, its metastable point lies at $\frac{V_{DD}}{2}$. In case we assume the drain-current of the pull-down transistors is constant, the discharge to the metastable point can be described for a conventional SA as follows [16]:

$$\frac{I_d \cdot \Delta t}{C_L} = \frac{V_{DD}}{2} \tag{1}$$

Here, I_d is the drain current of the pull-down transistors, C_L is the load capacitance at nodes SAOut/$\overline{\text{SAOut}}$, and Δt the time to discharge the load by $\frac{V_{DD}}{2}$. In case of the skewed SA, one of the internal nodes has a higher capacitance. Hence, the connected pull-down transistor will be slower at discharging it. In case both pull-down transistors are matched, this node requires a lower initial voltage in order to discharge to the metastable point in the same time Δt as the other pull-down

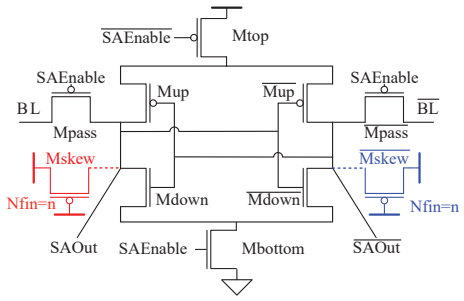


Fig. 3: Design of the skewed sense amplifier.

transistor. The required voltage difference between the nodes to achieve this, is the offset voltage V_{os}. Based on (1), the discharge of the skewed SA to its metastable point can now be described as follows:

$$\frac{I_d \cdot \Delta t}{C_L + n \cdot C_d} + V_{os} = \frac{I_d \cdot \Delta t}{C_L} = \frac{V_{DD}}{2} \tag{2}$$

Here, n is the skewing factor, which denotes the number of fins of the skew transistor, and C_d the drain capacitance of a single fin. By solving (2) for V_{os}, the following equation can be derived to estimate the offset voltage of the skewed SA:

$$V_{os} = \frac{V_{DD}}{2} - \frac{V_{DD}}{2} \cdot \frac{C_L}{C_L + n \cdot C_d} \tag{3}$$

From (3), it follows that the created shift in offset voltage depends on the skewing factor n, e.g., a higher skewing factor results in a higher offset voltage shift. Hence, an appropriate skewing factor should be selected which must be based on the expected aging due to the target workload.

IV. Validation via a Case Study

This section presents a case study to validate the mitigation scheme. It first discusses the experimental setup, followed by the experimental results.

A. Experimental Setup

As a a case-study, we investigate the impact of aging on the latch-type SA from Fig. 1. For this analysis, a 14 nm FinFET design at industrial-strength is used. The SA circuit is simulated using Spectre. Here, Monte Carlo simulations are performed during which the effects of aging and process variation are injected. For the aging, the calibrated atomistic BTI model from [3] is used. The used stress condition for the aging is three years at 125°C and nominal V_{DD}. This extreme condition corresponds to a device with a long lifetime. To simulate process variation, Pelgrom's model is used [17]. For each workload, 1,000 Monte Carlo simulations are performed. During each Monte Carlo iteration, the offset voltage of the SA is measured. After the Monte Carlo simulations have finished, the corresponding offset voltage spec (discussed in Section II) of the offset distribution is calculated. Subsequently, the SA is replaced by the skewed SA from Fig. 3 with an appropriate skewing factor n. Here, a skewing factor is chosen that results in a mean offset voltage shift that is closest to the mean shift due to aging without mitigation. Note that this shift is performed in the opposite direction. Next, the impact of aging

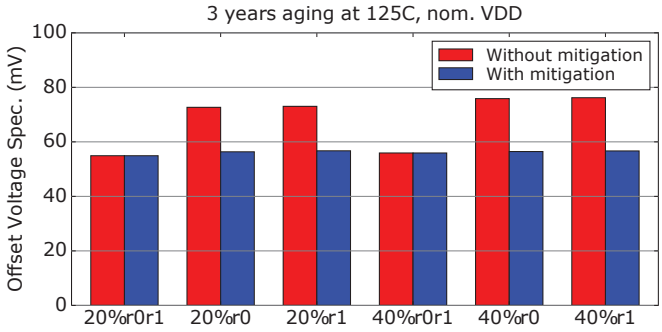


Fig. 4: Degadradation of the offset voltage spec for several workloads.

is also evaluated for this skewed SA in order to evaluate its mitigation capability.

Six different workloads are used for the aging; they are named 20%r0r1, 20%r0, 20%r1, 40%r0r1, 40%r0, and 40%r1. Here, the first number (20% or 40%) indicates the activation rate of the SA. For instance, 20% indicates that 20% of the cycles, the SA performs a read operation. After the activation rate, the pattern of the read values is indicated. Three patterns are used: r0r1 (50% of the reads are 0 and 50% of the reads are 1), r0 (all reads are 0), r1 (all reads are 1). The sequence r0r1 is used to mimic a balanced workload, while r0 and r1 are used to mimic an unbalanced workload.

B. Experimental Results

Fig. 4 shows the offset voltage spec after three years of aging without and with mitigation for the different workloads. The figure reveals the following:

- The mitigation scheme reduces the degradation of the offset voltage spec significantly. The highest observed offset voltage spec without mitigation is $76.2\,\mathrm{mV}$, while it is only $56.7\,\mathrm{mV}$ with mitigation; a reduction of $\sim$26%. In addition, the benefit of the mitigation is strongly workload-dependent. The mitigation scheme gives the highest benefits for unbalanced workloads (i.e. 20%r0, 20%r1, 40%r0, and 40%r1). This is due to the fact that unbalanced workloads shift the mean of the offset voltage distribution, which the mitigation scheme compensates for. On the other hand, the mitigation scheme gives no advantage for balanced workloads (i.e., 20%r0r1 and 40%r0r1). The reason for this is that balanced workloads do not shift the mean of the offset voltage distribution. Hence, the degradation is already minimal for balanced workloads and, thus, the mitigation scheme cannot reduce it further.
- The mitigation scheme reduces the dependency of the offset spec's degradation on the workload significantly; it varies only between $54.9\,\mathrm{mV}$ and $56.7\,\mathrm{mV}$ with mitigation, while it varies between $54.9\,\mathrm{mV}$ and $76.2\,\mathrm{mV}$ without mitigation. The dependency on the workload for the mitigation scheme is low, due to the fact that it compensates for the shift of the mean offset voltage. Note that this mean offset voltage shift has the highest impact on the degradation of the offset voltage spec [7].
- In case no mitigation scheme is applied, the degradation of the offset voltage spec depends strongly on the workload; the offset voltage spec varies between $54.9\,\mathrm{mV}$ and $76.2\,\mathrm{mV}$. The degradation is mainly impacted by the pattern of the read values (i.e. balanced versus unbalanced) and it is highest for unbalanced workloads. For example, workload 20%r0r1 gives an offset spec of only $54.9\,\mathrm{mV}$, while workload 20%r1 gives a spec of $73.0\,\mathrm{mV}$. Unbalanced workloads result in a much higher degradation due to the shift in the mean offset voltage [7]. In addition, the degradation of the offset spec has a slight dependence on the activation rate of the workload. A higher activation rate results in a higher degradation, since the SA is stressed more. For instance, workload 20%r1 gives an offset spec of $73.0\,\mathrm{mV}$, while 40%r1 gives an offset spec of $76.2\,\mathrm{mV}$.

V. Conclusion

This paper proposed the skewed SA design to mitigate the impact of aging on the memory's SA. It is based on creating a skew in the relative strengths of the SA's cross-coupled inverters at design time. The skew is compensated by aging due to unbalanced workloads, which results in a reduced overall impact of aging on the SA. Our case study revealed that the mitigation scheme reduces the degradation of the offset voltage spec with up to 26%. Hence, our proposed mitigation scheme can be applied in order to extend the lifetime and improve the reliability of memories.

References

[1] S. Borkar, "Design challenges of technology scaling," *IEEE Micro*, vol. 19, no. 4, Jul. 1999.
[2] J. Srinivasan *et al.*, "The impact of technology scaling on lifetime reliability," in *DSN*, Jun. 2004.
[3] B. Kaczer *et al.*, "Origin of nbti variability in deeply scaled pfets," in *IRPS*, May 2010.
[4] E. I. Vatajelu *et al.*, "Analyzing the effect of concurrent variability in the core cells and sense amplifiers on sram read access failures," in *DTIS*, Mar. 2013.
[5] A. Listl *et al.*, "Sram design exploration with integrated application-aware aging analysis," in *DATE*, Mar. 2019.
[6] J. Kinseher *et al.*, "Analyzing the effects of peripheral circuit aging of embedded sram architectures," in *DATE*, Mar. 2017.
[7] I. Agbo *et al.*, "Quantification of sense amplifier offset voltage degradation due to zero-and run-time variability," in *ISVLSI*, Jul. 2016.
[8] S. V. Kumar *et al.*, "Impact of nbti on sram read stability and design for reliability," in *ISQED*, Mar. 2006.
[9] S. Khan *et al.*, "Bias temperature instability analysis of finfet based sram cells," in *DATE*, Mar. 2014.
[10] P. Pouyan *et al.*, "Adaptive proactive reconfiguration: A technique for process-variability- and aging-aware sram cache design," *IEEE TVLSI*, vol. 23, no. 9, Sep. 2015.
[11] A. Gebregiorgis *et al.*, "Aging mitigation in memory arrays using self-controlled bit-flipping technique," in *ASPDAC*, Jan. 2015.
[12] R. Menchaca *et al.*, "Impact of transistor aging effects on sense amplifier reliability in nano-scale cmos," in *ISQED*, Mar. 2012.
[13] S. Khan *et al.*, "Impact of partial resistive defects and bias temperature instability on sram decoder reliablity," in *DTS*, Dec. 2013.
[14] D. Kraak *et al.*, "Mitigation of sense amplifier degradation using input switching," in *DATE*, Mar. 2017.
[15] D. Kraak *et al.*, "Degradation analysis of high performance 14nm finfet sram," in *DATE*, Mar. 2018.
[16] B. Wicht *et al.*, "Yield and speed optimization of a latch-type voltage sense amplifier," *IEEE JSSC*, vol. 39, no. 7, Jul. 2004.
[17] M. J. M. Pelgrom *et al.*, "Matching properties of mos transistors," *IEEE Journal of Solid-State Circuits*, vol. 24, no. 5, Oct. 1989.