Accurate and Insightful Closed-Form Prediction of Subthreshold SRAM Hold Failure Rate

Léopold Van Brandt, Roghayeh Saeidi, *Member, IEEE*, David Bol, *Senior Member, IEEE*, and Denis Flandre, *Senior Member, IEEE*

***Abstract*— The failure probabilities of industrial SRAM cells fall below the ppm (10^{-6}) range, disqualifying the computational-intensive Monte-Carlo simulations for efficient robustness assessment. Starting from a novel two-dimensional threshold voltage imbalance representation, we propose a new methodology for fast and accurate prediction of the subthreshold SRAM hold stability failure rate. The probability is derived in a closed form which involves the transistor threshold-voltage standard deviations and only requires the two quick DC extractions of the worst- and best-case static noise margins. We validate our approach on a Six-Transistor (6T) bitcell in 28 nm Fully Depleted Silicon-On-Insulator (FD-SOI) CMOS technology. Our method turns out to be especially insightful for comparative and sensitivity analyses, for instance to study the effect of supply voltage downscaling or temperature variations. Finally, we show that the achieved accuracy and the capability of estimating extremely low failure probabilities (down to 10^{-9}), combined with the important gain in simulation and post-processing cost, makes our methodology attractive compared to other recent modelling works.**

***Index Terms*— Subthreshold SRAM modelling, static noise margins, failure probability, statistical variability, fully depleted silicon-on-insulator.**

I. INTRODUCTION

THE Six-Transistor (6T) Static Random Access Memory (SRAM) cell is a key building block of integrated memories. In the nanometer era, its excellent scalability property has kept it very competitive for high-density array with fast and low-energy access [1]. Although custom SRAM architectures might be favoured by the designers of Ultra-Low-Power (ULP) microcontroller units [2], standard foundry 6T cells are still widely used [3]. In that regard, the conventional 6T SRAM bitcell remains the ultimate benchmark to compare different approaches of reliability assessment [4]–[20].

In decananometer technologies, multiple variability sources of various physical origins introduce random device-to-device fluctuations in the transistor parameters [21], [22]. The dominant statistical effect on the transistor DC performances is compactly and accurately modelled as a *standard deviation* $\sigma_{V_{\rm th}}$ of the transistor threshold voltage ($V_{\rm th}$), especially in the weak-inversion region [21]. The *accurate prediction* of the *cell failure probability* (or *rate*) is a major industrial concern, in order to guarantee a sufficient array or chip *yield* [9]. Strongly dependent on the architecture [23], [24], technology [4], [25], supply voltage [10], [24], [26] and temperature [10] conditions, the failure-rate metric spans across orders of magnitude. The estimation of probabilities as extremely low as 10^{-6} (i.e. *parts-per-million*, ppm) or even 10^{-9} with limited computational resources and time is a challenge, often referred to as "high-sigma (σ) statistical evaluation" in the circuit reliability literature [11], [12], [18].

In Figure 1, we present a novel representation of the samples of a massive conventional Monte-Carlo simulation of a subthreshold SRAM cell in *hold* (or *retention* or standby) mode. Each tested latch has been classified according to its particular *Static Noise Margin* (SNM) and then coloured as either functional ($SNM > 0$, green) or defective ($SNM \leq 0$, red), and finally mapped according to the imbalance of each inverter, defined by the difference between their pMOS and nMOS threshold voltage deviations arising from statistical variability. Even though the Monte-Carlo simulation was the starting point of our thought and aims at constituting a reliable reference (sometimes referred to as "golden" [13], [18]) for quantitative comparisons, let us emphasize here that it is finally *not* necessarily part of the suggested methodology.

Interestingly, we observe in Figure 1 clearly well-defined and separate pass (green) and fail (red) regions. We will show that they can be cheaply determined with a *deterministic* DC sweep simulation. The boundaries between those regions even seem closely *linear*. We will emphasize that the front lines are moreover cleverly determined by *two special points*, actually corresponding to the *worst-* and *best-case static noise margins* of the latch as they were originally defined by Lohstroh [27]. "Worst case" refers to the situation where both inverters are simultaneously affected by an adverse perturbation, while in the "best case" only one inverter suffers from a stronger variation [27]. Those empirical circuit-related observations, combined with the statistical knowledge of the $V_{\rm th}$ distribution, point to the hope of *predicting the hold failure rate (and the yield) with a closed-form formula*. The methodology developed in this paper applies to a wide family of SRAM bitcells in subthreshold operation, including the conventional 6T cell but also the 8T [23], [28] or the 10T [29] cell which are important

Manuscript received October 22, 2021; revised March 1, 2022; accepted March 24, 2022. Date of publication April 8, 2022; date of current version June 29, 2022. The work of Léopold Van Brandt was supported by the National Fund for Scientific Research (FNRS) of Belgium under FRIA Grant. This article was recommended by Associate Editor C. H. Chang.
(Corresponding author: Léopold Van Brandt.)

The authors are with the ICTEAM Institute, Université Catholique de Louvain, 1348 Ottignies-Louvain-la-Neuve, Belgium (e-mail: leopold.vanbrandt@uclouvain.be).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TCSI.2022.3164680.

Digital Object Identifier 10.1109/TCSI.2022.3164680

1549-8328 © 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: Univ Catholique de Louvain/UCL. Downloaded on August 02,2022 at 08:17:02 UTC from IEEE Xplore. Restrictions apply.

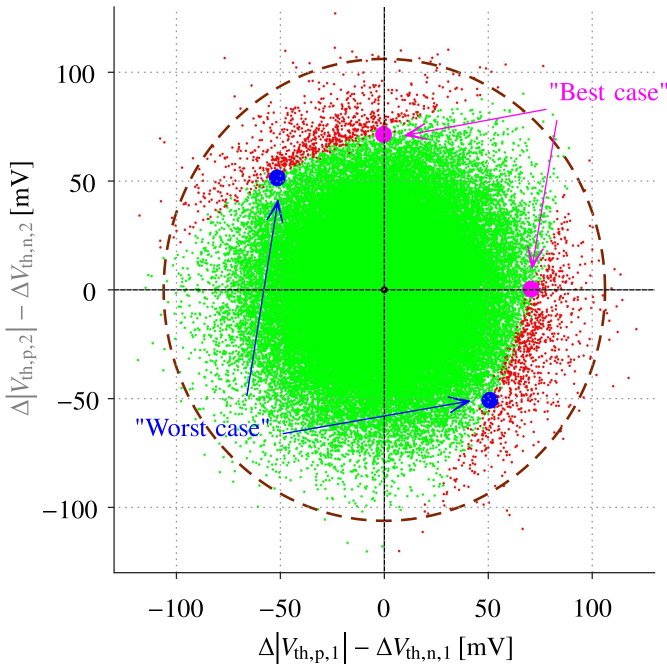

Fig. 1. Monte-Carlo simulation of a latch, based on two cross-coupled inverters, with 100 000 runs. An SNM analysis is performed for each latch which gets classified as either functional ("pass", in green) or defective ("fail", in red). Then, all the latches are mapped into this two-dimensional threshold voltage imbalance representation. 99.9% of the samples are confined in the circle (brown dashed line). The "worst-case" $(\pm 2 \cdot SNM^*, \mp 2 \cdot SNM^*)$ (in blue) and "best-case" $(+2 \cdot \delta V^*, 0)$ and $(0, +2 \cdot \delta V^*)$ (in magenta) static-noise-margin points are plotted and unequivocally determine the front lines between pass and fail regions. Illustrated case: 28nm FD-SOI LVT CMOS inverters with $L_{\mathrm{n}} = L_{\mathrm{p}} = 30$nm, $W_{\mathrm{n}} = 80$nm, $W_{\mathrm{p}} = 160$nm; $V_{\mathrm{DD}} = 150$mV; $T = 120°C$.

ULP architectures [1], [3], since they all use the same cross-coupled inverter pair (latch) for data retention [24].

This rest of the paper is structured as follows. Section II reviews the alternative methodologies to Monte-Carlo simulations and the recent SRAM modelling efforts. In Section III, we analyse our two-dimensional threshold voltage imbalance representation, with a rigorous modelling approach supported with SPICE simulations. Because they play a substantial role in our approach, SNM-related concepts and simulations will be comprehensively reviewed and clarified in Section III-B, especially with the support of the traditional butterfly plot. The analytical prediction of the failure probability is the purpose of Section IV and the related Appendix A. Our methodology is notably applied to evaluate the robustness of a state-of-the-art standard foundry SRAM bitcell implemented in 28 nm Fully Depleted Silicon-On-Insulator (FD-SOI) CMOS technology in Section IV-B. The effect of the supply voltage downscaling is readily obtained without significant computational overhead. Failure rates of the order of ppm and lower are accurately estimated and we show that we outperform the previous approaches of the literature regarding low cumulated simulation and computational cost in Section IV-C. Section V draws conclusions of this paper.

II. Related Works

The most general tool to assess the impact of statistical variability on the functionality of a given design is the *Monte-Carlo simulation*. The principle is to simulate a very large number of random SRAM cells in order to get an numerical empirical estimate of the failure probability. The assessment of the very low industrial bitcell failure probability would require millions or billions runs, which is prohibitively expensive [9]. The order of magnitude of the actual failure rate may not be known *a priori*, so that the required number of samples is generally unknown too, which constitutes another severely annoying practical issue. Works addressing this problem can be roughly divided into two distinct categories: enhanced Monte-Carlo methods on one hand; SRAM statistical modelling (to which our contribution belongs) on the other hand.

A. Importance Sampling in All Its Forms

Enhanced Monte-Carlo methods may rely on extreme value theory [9], or on *importance sampling* (IS) in order to significantly accelerate the simulations, such as lately: gradient IS [12]; high-dimensional IS [13]; "adaptive" (dynamic) IS [17]; Bayesian IS [14], [16], [18]. The number of trials is reduced as concentrated in the distribution tails, but still vary between a few hundreds and a few thousands, mostly depending on the targeted failure rate and desired accuracy. Those statistics-oriented methods are general, but do not exploit the knowledge of the SRAM behaviour and do not provide the designer with insight or intuition about the effect of relevant parameters (i.e. supply voltage or temperature) on the failure rate.

B. SRAM Analytical Modelling Works

The topic used to be well covered. In [4], a 6T SRAM SNM stochastic model was developed based on the outdated transregional long-channel transistor model. In [7], Calhoun analytically studied 65 nm subthreshold SRAM. The used transistor model ignores the *Drain-Induced Barrier Lowering* (DIBL) effect, that we expect to introduce significant discrepancies in more advanced technologies [30]. The voltage transfer characteristics (VTC) of the inverters were derived but do not lead to a direct analytical solution for the SNM. The computational cost of the forced numerical way was later quantified in more than one hour by [10]. The same [10] has developed a fully analytical model for near- and sub-threshold SRAM cells and performed subsequent statistical analyses of the read SNM. The equations rely on the accurate weak-inversion model, DIBL included, but on the other hand must be linearized to achieve a closed-form expression of the SNM.

A non-Monte-Carlo methodology, based on the multi-dimensional response surfaces of the SNM as a function of the numerous statistical input parameters, was proposed in [11]. Although mathematically very sound, robust and general (as not restricted to Gaussian inputs), the final formulation has no closed-form solution for the distribution function of the SNM.

Paper [6] postulated that the SNM is a strictly linear function of all the transistor V_{th} deviations, and hence Gaussian itself. This assumption is harshly questioned in [7], [28]. Reference [5] directly worked on voltage distributions in exactly the same spirit as [6], notably to predict hold and read failure

Authorized licensed use limited to: Univ Catholique de Louvain/UCL. Downloaded on August 02,2022 at 08:17:02 UTC from IEEE Xplore. Restrictions apply.

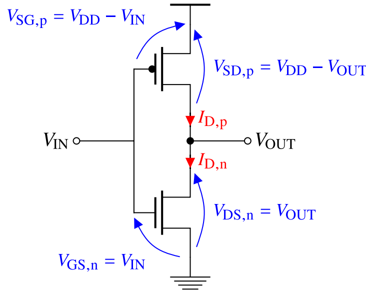

Fig. 2. CMOS inverter: notations for voltages and currents.

probabilities. Recently, the device-oriented statistical model of the read SNM of [15] took into account the (possibly non-linear) dependencies on the various transistor parameters (V_{th}, W and L). The drawback is the huge number of SPICE simulations required to calibrate the final equation: 217 for a 6T SRAM cell; reduced to 145 for a latch.

Paper [28] went beyond the conventional multi-linear model of the SNM for subthreshold SRAM cells. To our best knowledge, [28] is the first and only published successful attempt in a *reduction of the dimensionality* of the problem. The SNM is written as a general, non-linear function of only the two variations in transition voltages of the VTCs. The distribution of the SNM is finally derived in a closed form. Strictly speaking, only one circuit simulation is needed with this approach, in order to obtain the nominal VTC. Yet, non-trivial numerical post-processing is required to compute the crucial "mapping function" and extract the useful information to be plugged into the SNM distribution; the computational cost of the operation was not quantified. Furthermore, the assessment of the failure rate was not addressed.

In two recent twin works, Gupta achieved a nearly instantaneous closed-form prediction of the *read-* [19] and *write-access* [20] failure probability. The analyses of the read and write modes necessarily and specifically involve dynamic metrics, respectively the read- and the write-access time [5], [6], [19], [20]. The present work focuses on the hold mode for which the SNM is the most relevant metric [7], [31].

III. A Two-Dimensional Threshold Voltage Imbalance Representation

Starting from a classical Monte-Carlo simulation, Section I outlined our novel two-dimensional threshold voltage imbalance representation of the robustness of the latch (Figure 1). The first part of the current section aims at building such original representation step by step, given the subthreshold operation of the CMOS inverter. Afterwards, we explain how the pass and fail regions can be determined by simple DC simulation and we emphasize valuable observations toward analytical failure rate prediction.

A. Subthreshold Operation of the CMOS Inverter

1) Inverter DC Voltage Transfer Curve: Most generally, the single nodal equation of the static CMOS inverter circuit (Figure 2) is obtained by applying Kirchhoff's Current Law at the output node of the inverter:

$$I_{D,n} = I_{D,p} \tag{1}$$

where $I_{D,n}$ and $I_{D,p}$ are the currents flowing *into* the nMOS transistor drain terminal and *out* of the pMOS transistor drain terminal, respectively. V_{DD} always denotes the supply voltage.

a) Subthreshold current model of the transistors: In weak inversion, the exponential current model [3]

$$I_{D,n} = I_{0,n} \exp\left(\frac{V_{GS} - (V_{th,n} - \eta_n V_{DS})}{n_n \phi_T}\right) \times \left(1 - \exp\left(\frac{-V_{DS}}{\phi_T}\right)\right) \tag{2}$$

$$I_{D,p} = I_{0,p} \exp\left(\frac{V_{SG} - (|V_{th,p}| - \eta_p V_{SD})}{n_p \phi_T}\right) \times \left(1 - \exp\left(\frac{-V_{SD}}{\phi_T}\right)\right) \tag{3}$$

can boast very close agreement with experimental measurements or with the full compact model, provided that all the parameters, that are the reference currents $I_{0,n}/I_{0,p}$, the threshold voltages $V_{th,n}/V_{th,p}$, the body factors n_n/n_p and the DIBL coefficients η_n/η_p, are accurately extracted. In (2) and (3), $\phi_T = kT/e$ denotes the thermal voltage.

The node voltage differences are sketched in Figure 2. Substituting (2) and (3) in (1), taking the $\ln(\cdot)$ of both sides and rearranging the terms yields

$$\ln\left(\frac{I_{0,n}}{I_{0,p}}\right) + \frac{|V_{th,p}|}{n_p \phi_T} - \frac{V_{th,n}}{n_n \phi_T} - \frac{(1+\eta_p)}{n_p}\frac{V_{DD}}{\phi_T} + \left(\frac{1}{n_n} + \frac{1}{n_p}\right)\frac{V_{IN}}{\phi_T} + \left(\frac{\eta_n}{n_n} + \frac{\eta_p}{n_p}\right)\frac{V_{OUT}}{\phi_T} = \ln\left(\frac{1 - \exp\left(\frac{-(V_{DD} - V_{OUT})}{\phi_T}\right)}{1 - \exp\left(\frac{-V_{OUT}}{\phi_T}\right)}\right). \tag{4}$$

Equation (4) merely confirms and extends the subthreshold inverter model of [7], [32], this time by including the DIBL effect in (2) and (3). Even though the function $V_{OUT} = h(V_{IN})$ cannot be explicitly isolated, the inverse transfer relationship $V_{IN} = h^{-1}(V_{OUT})$ is readily obtained from (4) and determines the *inverter DC voltage transfer curve (VTC)*.

Equation (4) will be extensively used thereafter in order to provide rigorous mathematical support for our two-dimensional threshold voltage imbalance representation already introduced in Figure 1, as well as intuition about it. However, we emphasize that such analytical model will *never* be used for quantitative reliability assessment, notably the failure rate prediction, unlike [10] which had to resort to linearization of (4) to be able to continue. Our unique reference will always remain the SPICE simulations performed with the full transistor compact model. The model (4) of the inverter aims at insightfully constructing and explaining the reliability-oriented representation of the latch.

b) Impact of the nMOS/pMOS transistor imbalance: Depending on the technology and the sizing of the transistors (choice of a W_p/W_n ratio), the nMOS and the pMOS transistors still might not be perfectly balanced. Such imbalance is captured by the first two terms in (4), those involving:

Authorized licensed use limited to: Univ Catholique de Louvain/UCL. Downloaded on August 02,2022 at 08:17:02 UTC from IEEE Xplore. Restrictions apply.

- the ratio of the reference currents $I_{0,\text{n}}/I_{0,\text{p}}$;
- the weighted *difference between the threshold voltages*

$$\frac{n_\text{n}|V_{\text{th,p}}| - n_\text{p}V_{\text{th,n}}}{n_\text{n}n_\text{p}\phi_T},$$

Very interestingly, (4) also teaches us that: it is only this difference that affects the inverter transfer curve, and *not* the individual values of the threshold voltages, in accordance with intuition; the effect of $n_\text{n}|V_{\text{th,p}}| - n_\text{p}V_{\text{th,n}}$ is rigorously *linear*, assuming of course that both transistors remain in the weak inversion region over the entire voltage range $[0, V_\text{DD}]$. This last condition sets the limit of validity of (4).

2) Impact of Static Variability: When the transistors are subject to static variability, the threshold voltages $V_{\text{th,n}}$ and $|V_{\text{th,p}}|$ shift from their nominal values $V_{\text{th0,n}}$ and $|V_{\text{th0,p}}|$ by some variations that we denote $\Delta V_{\text{th,n}}$ and $\Delta|V_{\text{th,p}}|$:

$$V_{\text{th,n}} \equiv V_{\text{th0,n}} + \Delta V_{\text{th,n}} \tag{5}$$
$$|V_{\text{th,p}}| \equiv |V_{\text{th0,p}} + \Delta|V_{\text{th,p}}|. \tag{6}$$

Clearly, a positive $\Delta|V_\text{th}|$ weakens the transistor, while a negative $\Delta|V_\text{th}|$ strengthens it compared to its nominal operation. In general, $\Delta V_{\text{th,n}}$ and $\Delta|V_{\text{th,p}}|$ takes totally independent and different values. When we perform a Monte-Carlo simulation, the threshold voltages are independently sampled from their respective Gaussian distributions.

We study the impact of those threshold voltage shifts with a deterministic approach. Equation (4) predicts an *horizontal shift of the inverter transfer curve* quantified as

$$\Delta V_\text{IN} = -\frac{\left(n_\text{n}\,\Delta|V_{\text{th,p}}| - n_\text{p}\,\Delta V_{\text{th,n}}\right)}{n_\text{n}\ + n_\text{p}}. \tag{7}$$

A similar but less general equation was provided in [28], even though they did not include the DIBL effect in the equations. We first visit some important and instructive special cases, before drawing a general conclusion about the role of the threshold voltage imbalance.

a) Same variation on both transistors: An insightful particular case occurs when $\Delta|V_{\text{th,p}}| = -\Delta V_{\text{th,n}} = \pm\delta V$, which corresponds to a DC voltage $\pm\delta V$ applied at the input of the inverter. The variation is thus common to both transistors but the effect on their respective threshold voltages is opposite in absolute value. The impact on the VTC (7) reduces to

$$\Delta V_\text{IN} = \mp\delta V \tag{8}$$

irrespective of the values of n_n/n_p. We note that the shift (8) is precisely the opposite of half of the *threshold voltage imbalance* Δ introduced by the variability:

$$\Delta \equiv \Delta|V_{\text{th,p}}| - \Delta V_{\text{th,n}} = \pm 2\cdot\delta V. \tag{9}$$

b) Only one transistor affected: The extreme cases consist in attributing such an imbalance entirely to either the nMOS or the pMOS transistor only. For instance, considering $\Delta V_{\text{th,n}} = \mp 2\cdot\delta V$ and $\Delta|V_{\text{th,p}}| = 0$, we compute with (7)

$$\Delta V_\text{IN} = \mp\frac{2\,n_\text{p}}{n_\text{n} + n_\text{p}}\,\delta V. \tag{10}$$

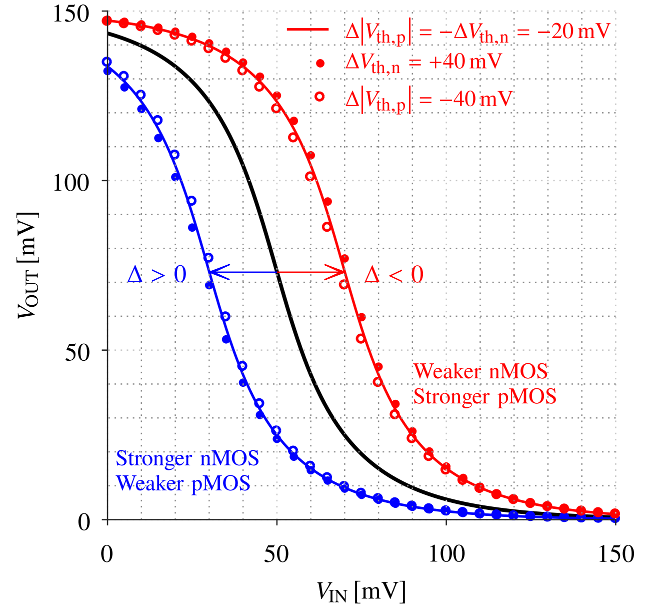

Fig. 3. Inverter DC transfer voltage curve. Horizontal shift of the curve subsequent to the introduced threshold voltage imbalance Δ defined in (9). Illustrated case: inverter defined in Figure 1.

However, the body factors are in practice relatively close, $n_\text{n} \approx n_\text{p}$, and we therefore again find $\Delta V_\text{IN} \approx \mp\delta V$, that is

$$\Delta V_\text{IN} \approx -\frac{1}{2}\left(\Delta|V_{\text{th,p}}| - \Delta V_{\text{th,n}}\right) \equiv -\frac{1}{2}\Delta. \tag{11}$$

c) Illustration and discussion: The former mathematical analysis is best graphically illustrated and supported with simulations. The nominal VTC of the inverter has been depicted in black in Figure 3. Then, depending on the sign of the threshold voltage imbalance Δ defined in (9), the VTC is horizontally shifted to the left ($\Delta > 0$, blue) or to the right ($\Delta < 0$, red). The coloured full lines correspond to the two mixed cases $\Delta|V_{\text{th,p}}| = -\Delta V_{\text{th,n}} = \pm 20\,\text{mV}$ resulting in a total $\Delta = \pm 40\,\text{mV}$. The consequent shift of the VTC is $\Delta V_\text{IN} = \mp 20\,\text{mV}$ in excellent agreement with (8). Finally, the coloured full (•) and open (◦) dots represent the VTCs obtained when the nMOS or the pMOS transistor alone is affected by a threshold voltage shift of 40 mV in one direction or the other. We can observe that both for positive and negative imbalances, the three curves almost merge, with an horizontal difference between them quantified as small as 1 mV. The way the total threshold voltage imbalance $\Delta = \pm 40\,\text{mV}$ is distributed between the two transistors does *not* significantly impact the VTC.

In summary, we have claimed and shown that the effect of the *two* individual nMOS and pMOS threshold voltage variations on the DC operation of the subthreshold CMOS inverter is very generally and accurately captured by *one* single quantity, the threshold voltage imbalance Δ, which horizontally shifts the transfer curve according to (11).

B. Subthreshold Operation of the CMOS Latch

An SRAM cell in retention mode may be modelled as a *latch* made of two cross-coupled inverters (Figure 4), provided

Authorized licensed use limited to: Univ Catholique de Louvain/UCL. Downloaded on August 02,2022 at 08:17:02 UTC from IEEE Xplore. Restrictions apply.

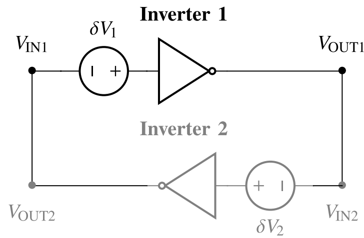

Fig. 4. SRAM cell in retention mode, modelled as a latch made of two back-to-back inverters. The static voltages sources of variation are included.

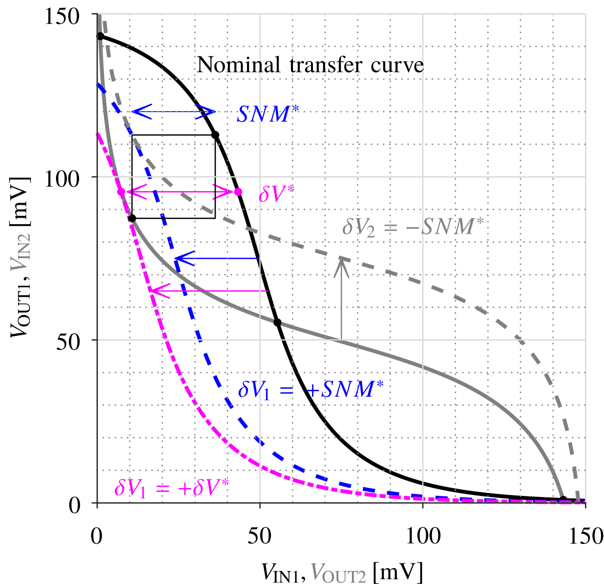

Fig. 5. Transfer characteristics of the two inverters of the latch. Butterfly plot and geometrical representation of the SNM ("worst-case" scenario; blue dashed line for the VTC of Inverter 1, gray dashed line for Inverter 2). The "best-case" scenario is also plotted, when only Inverter 1 is affected by $+\delta V^*$ (magenta dashed-dotted line), resulting in a defective latch. Illustrated case: same as Figure 1. $SNM^* = 26$mV; $\delta V^* = 36$mV.

that the leakage current in the two cut-off access transistors can be neglected. This should be guaranteed in hold mode, even in presence of variations, by designing a zero or even negative word line voltage (WL) on the minimum-size access transistors. The critical hold SNM is moreover reported to improve in specifically-designed subthreshold SRAM, such as 10T cells, even when adjacent cells are written [3], [29].

1) Worst-Case and Best-Case Static Noise Margins: The static stability of the structure is better understood based on the traditional butterfly plot of Figure 5 [31]. The latch is considered as *functional* as long as the number of cross points of the mirrored transfer curves is exactly equal to *three*: the top-left and bottom-right *stable* points corresponding to the two well defined high and low logic levels taken by the data in retention; the middle cross point in the gain region called the metastable point.

We are interested in quantifying the robustness of the cell against DC voltage sources of variations, which we assume to be applied at the input of each inverter as sketched in Figure 4. We note that this is precisely the situation where the nMOS and the pMOS transistors are affected by a common noise source, whose impact is described by (8) and (9). We also know from the analyses of Section III-A that this threshold voltage imbalance approach accurately captures the effect of more general device-to-device variability on the inverter transfer curve. Now, we study the impact of the *imbalance between the two inverters* on the hold stability of the latch.

An important special case occurs when *both* inverters are *adversely* affected, that is by the same voltage perturbation, yet with opposite signs. If $\delta V_1 = -\delta V_2 = SNM^*$ (dashed lines in Figure 5), the VTC of Inverter 1 horizontally shifts to the left, while that of Inverter 2 vertically shifts up, both according to (8). Two roots of the latch equation coincides (the number of cross points reduces to *two*) and the latch gets *defective* [27]. The acronym SNM attributed to this particular value of static variation stands for *Static Noise Margin* while the * means nominal. Pioneer works of Lohstroh demonstrated the remarkable equivalence between this definition of the SNM, emanating from static-stability considerations, and the more well-known geometrical interpretation of the SNM [27]: the width of the largest inscribed square between the normal and the mirrored VTCs (also illustrated in Figure 5). This provides a robust pratical extraction method that has been followed up by List and Seevinck [31] and widely used in the device and digital circuit reliability literature since then [33].

The hereabove described case is usually referred to as "worst-case" scenario, in the sense that when both inverters are simultaneously affected by an adverse variation, a moderate value of the variation (exactly the nominal SNM^*) is sufficient to cause the latch to become defective. The small variations occurring statistically more frequently, this failure case is the worst to be considered by the circuit designer. Less known in the literature is the so-defined "best-case" scenario where only *one* gate suffers from some variation [27]. A larger positive voltage, that we will denote by δV^*, needs to be applied to either inverter in order to reach a degenerate $SNM = 0$ situation (dash-dot line in Figure 5). Let us emphasize that there is, in general, no direct geometrical relationship between the two noise margins SNM^* and δV^*. These two quantities are independent and separately extracted by DC simulation.

a) Generalization: So far, we have reviewed the two important special cases of imbalance in the latch structure: equally but adversely distributed between the two inverters as $\pm SNM^*$ in the worst-case scenario; "all (δV^*) in one inverter" in the best-case scenario. More generally, any intermediate defective case between the "worst" and the "best" ones is possible in the presence of statistical variability. The local parameter mismatch between the nMOS and pMOS transistors, included in (4), quantifies the imbalance of the individual inverter (9) and the resulting shift of the VTC (11). The impact on the functionality of latch then also depends on the behaviour of the second inverter. In any case though, *the SRAM hold failure criterion is that the SNM of the latch drops to zero*, as a consequence of some severe total imbalance between the two cross-coupled inverters.

The values taken by the two key metrics SNM^* and δV^* are strongly dependent on the technology, the transistor sizing, the supply voltage V_{DD} and the temperature T. Those parameters indeed influence the VTCs of the inverters, whose positions relative to each other determine the particular SNM of a given configuration. It may seem intuitive that SNM^* and δV^*

Authorized licensed use limited to: Univ Catholique de Louvain/UCL. Downloaded on August 02,2022 at 08:17:02 UTC from IEEE Xplore. Restrictions apply.

follow correlated trends, for instance when V_{DD} is downscaled or when T changes. Unfortunately, the accurate analytical model (4) of the subthreshold CMOS inverter does not lead to general closed-form expressions. The modelling of SNM^* and δV^* as well as of their interplay goes beyond the scope of this paper.

The inequality $SNM^* \leq \delta V^*$ is trivial: if *both* inverters are adversely affected by the variation, this voltage is less than the one which should be applied to *one single* inverter. Less obvious is an upper bound for δV^* which can still be geometrically derived. For all the practical SRAM cell designs, the following inequality holds

$$SNM^* \leq \delta V^* \leq 2 \cdot SNM^* \tag{12}$$

because the inscribed square of side length SNM^* always covers more than half of the difference between the normal and the mirrored transfer curves along the critical direction regarding stability (the horizontal for the left lobe of the butterfly in Figure 5).

2) Threshold Voltage Imbalance Representation: Acting on the highlighted importance of the relative imbalance between the two inverters, we present a novel reliability-oriented representation, or classification of some set of identically designed latches but all affected by statistical variability. Figure 1 was briefly introduced in the end of Section I. Each latch is located according to its particular inverter imbalances, now shortly denoted by Δ_1 and Δ_2 following definitions (5)-(6) and (9). For the sake of understanding and illustration, samples of a Monte-Carlo simulation were collected and mapped in Figure 1. However, the pass and fail regions as well as the borders separating them can easily be obtained with a deterministic approach as we explain right below. Our goal is indeed to get rid of the computational intensive Monte-Carlo analyses and to provide a quick estimate of the failure probability.

a) Pass and fail regions through a deterministic approach: The failure coloured map outlined in Figure 1 may alternatively be efficiently drawn by a *DC sweep simulation* using the testbench of Figure 4. The two independent DC sources δV_1 and δV_2 are swept over a sufficient voltage range, for instance but not restricted to $[-V_{DD}, +V_{DD}]$.[1] Remembering (9), the corresponding varying inverter imbalances are given by

$$\Delta_{1,2} = 2 \cdot \delta V_{1,2}. \tag{13}$$

In Figure 6, to each (Δ_1, Δ_2) point is virtually associated an SRAM cell. Just like the Monte-Carlo samples, each latch undergoes an SNM analysis, is classified and coloured as either functional (green) or defective (red) accordingly, before being mapped into the two-dimensional representation.

b) Relations and differences to former works: Our representation might recall the traditional $|V_{th,p}|$-$V_{th,n}$ plot formerly presented for instance in [34], [35]. We nevertheless emphasize that the $|V_{th,p}|$-$V_{th,n}$ plot solely outlines the effect of the global,

[1]The necessary range may not be *a priori* known and we might want to perform a second more refined simulation *a posteriori*, which really does not constitute an issue given the speed of a DC sweep simulation.

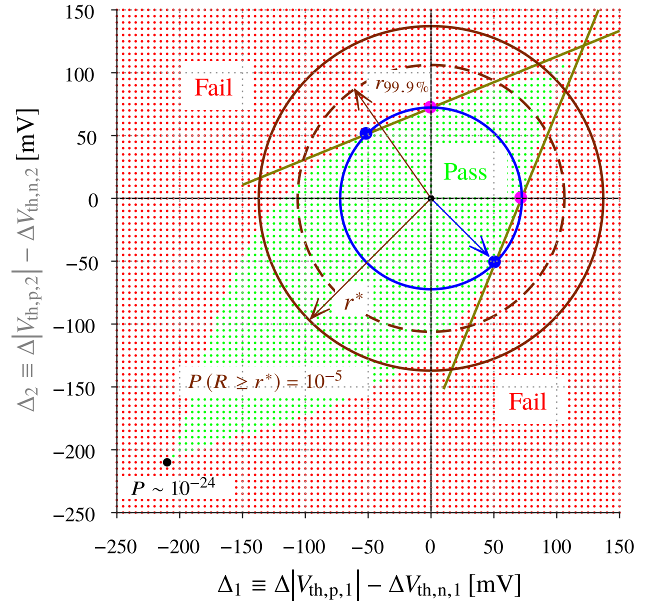

Fig. 6. Assessment of the pass (green) and fail (red) regions with the two-variable *deterministic* DC sweep simulation. Unlike Monte-Carlo, the deep tails of the full diamond shape can be observed. The blue equiprobability circle is defined by the $(\pm 2 \cdot SNM^*, \mp 2 \cdot SNM^*)$ "worst-case" (blue) points but here nearly passes through the $(+2 \cdot \delta V^*, 0)$ and $(0, +2 \cdot \delta V^*)$ "best-case" (magenta) points. Two other equiprobability circles (in brown) are drawn: the 99.9%-circle (dashed line, like in Figure 1), defined by (18)-(19); the outer 10^{-5}-circle of radius r^* (full line) delimiting *in probability* (according to (18) and (20)) the area *observable* with the 100 000 trials of the Monte-Carlo simulation of Figure 1. The bulk of defective samples of Figure 1 lie in the vicinity of the worst-case and best-case special points. Within this statistically relevant region for the failure rate to be estimated, the boundary closely follows the linear front line (defined in (23) further). Extracted $\sigma_{V_{th,n}} = 18mV$, $\sigma_{V_{th,p}} = 22mV$. $r_{99.9\%} = 106mV$; $r^* = 137mV$.

systematic imbalance between nMOS and pMOS transistors, while our representation allows to describe the local, random imbalance between the two inverters of the SRAM cell.

The threshold voltage imbalance representation notably relies on a *reduction of the dimensionality* of the problem. The SNM is therein assumed to be a *two*-dimensional function of the inverter imbalances only:

$$SNM = f(\Delta_1, \Delta_2) \tag{14}$$

The reduced formulation (14) can be interpreted as a *projection* of the full four-dimensional response surface of the SNM as a function of the *four* individual threshold voltage variations. This overly general multi-dimensional framework, adopted by [11], does not lead to a closed-form solution, neither does it give any insight about the impact of the inverter imbalance inherent to the design on the failure rate. The representation of Figure 6 can also be seen as a contour line plot of the function (14). The inside pass region correspond to the latches whose SNM is strictly positive, while the outside fail region contains all those whose SNM drops below zero.

Our approach presents similarities to [28]'s two-dimensional formulation, even though the bivariate SNM function (14) does eventually not need to be explicitly extracted in our case. Instead of considering the two inverter imbalances, Zheng [28] rather attempts to write the SNM as a function

Authorized licensed use limited to: Univ Catholique de Louvain/UCL. Downloaded on August 02,2022 at 08:17:02 UTC from IEEE Xplore. Restrictions apply.

of the variations in transition voltages (defined at $V_{DD}/2$) of the two inverter VTCs, before coming down to the extraction of a single-variable mapping function based on a geometrical argument on the butterfly plot. In our opinion, the construction is only very briefly and qualitatively demonstrated in [28], and not exhaustively supported by SPICE simulations.

c) Comparison with monte-carlo: Within the systematic approach, the pass and fail regions can be fully and accurately determined. Unlike the Monte-Carlo simulation, the full diamond or closing-eye shape of the central pass region is now clearly highlighted. Depending on the examined case and the desired grid resolution, the number of evaluated DC points ranges from tens to a few thousands, i.e. a gain of several orders of magnitude compared to the Monte-Carlo brute force. Furthermore, the bottom-left area of Figure 6 was even not observed in Figure 1. The reason why is the limited number of Monte-Carlo trials (although already considerable) that prevents from having access to the deep tails of the V_{th} distributions, and thus to the faraway regions of the threshold voltage imbalance representation.

d) Statistical considerations: More quantitatively, it can be shown that the random variable associated to *radius*

$$R = \sqrt{\Delta_1^2 + \Delta_2^2} \tag{15}$$

follows a *Rayleigh distribution* [36, p. 10], since $V_{th,n}$ and $V_{th,p}$ are normally distributed (central limit theorem) as well as their difference (property of the Gaussian random variables). The probability distribution function of R reads

$$f_R(r) = \frac{r}{\sigma^2}\exp\left(-\frac{r^2}{2\sigma^2}\right) \quad \text{for } r \geq 0 \tag{16}$$

where

$$\sigma^2 = \sigma^2_{V_{th,n}} + \sigma^2_{V_{th,p}} \tag{17}$$

is the variance of either threshold voltage imbalance Δ. The probability that R *exceeds* some distance r is given by the *complementary cumulative distribution function* calculated as

$$\begin{aligned} P(R \geq r) &= 1 - P(R \leq r) = 1 - \int_0^r f_R(r)\,\mathrm{d}r \\ &= \exp\left(-\frac{r^2}{2\sigma^2}\right). \end{aligned} \tag{18}$$

Probability (18) is thus the probability that one arbitrary latch exhibits transistor variations which send it further than r from the origin $(0, 0)$ of Figure 1 or Figure 6.

For example, the *equiprobability circle* (blue) passing through the worst-case points is defined by $r = \sqrt{2}\cdot 2\cdot SNM^*$ and has been plotted in Figure 6. The word "equiprobability" stands as a consequence of the total radial symmetry of the Rayleigh probability density function (16). We recall that the factor 2 arises from (13). For the illustrated case, we compute with the help of (17) and (18) the probability that some latch is outside the circle: $P\left(R \geq \sqrt{2}\cdot 2\cdot SNM^*\right) = 41\%$. This large value justifies why the 100 000 Monte-Carlo trials were, for the current case, found to be sufficient to abundantly populate the areas near the worst- and best-case points in Figure 1 and so to empirically estimate the failure rate as performed in Section IV-A.1. A potentially troubling observation might be that the "best case" seems almost as probable, in the statistical sense, as the "worst case" in Figure 6. This is explained by the fact that $\sqrt{2}\cdot SNM^*$ is *for this case* tightly close to δV^* and could well be purely fortuitous.

e) Observable region: As intuitively expected, the extent of the region accessible to the Monte-Carlo simulation is limited by the number of samples. We can nevertheless notice in Figure 1 that there is no strict or well defined boundary because of the granularity and the stochastic nature of the simulation. The *average radius of observation* must therefore be rigorously *defined in probability*. Generally, we may want to know where the bulk, say 99.9%, of the latches (like those generated in the Monte-Carlo simulation) are, *in average*, confined. Provided with (18), we need to solve

$$P(R \geq r_{99.9\%}) = (1 - 99.9\%)\cdot N_{\text{latches}} \tag{19}$$

for $r_{99.9\%}$. The computed equiprobability circle is drawn in both Figure 1 and Figure 6. The statistical interpretation of (19) is intuitive with the Monte-Carlo simulation of Figure 1: the latches are concentrated around the origin and the largest part of them inside the 99.9%-circle, while the extreme four cardinal points are logically sparsely or not at all populated. Any "high σ" circle could be similarly defined and plotted.

In average, only one latch among 100 000 is located at or beyond a point of probability as low as $1/N_{\text{latches}} = 10^{-5}$. Hence, the ultimate outer limit r^* may be defined by

$$P\left(R \geq r^*\right) = 1/N_{\text{latches}}. \tag{20}$$

The corresponding circle is also depicted in Figure 6 and distinguishes, in probability, the areas accessible or hidden to the Monte-Carlo simulation performed with a selected finite number of samples N_{latches}. The probability of occurrence of the faraway bottom-left tip of the diamond-shaped pass region is computed to be as derisory as $\sim 10^{-24}$, which prohibits access with any conventional massive Monte-Carlo approach, yet without practical consequence.

A key remark is, however, that the gridded coloured map of Figure 6 does not *explicitly* carry any statistical information, while Figure 1 does. With the DC sweep simulation, the Δ's are indeed uniformly gridded or sampled regardless of their true Gaussian distribution. Different areas of Figure 6 have therefore substantially different statistical contributions even though the points are unweighted by the simulation, those close to the origin are obviously the most likely. Fortunately, the knowledge of the Rayleigh distribution (16) has led to the construction of the equiprobability circles which contain the statistical information. It should be very clear that, according to (16), the points lying *inside* the observation circle are those most frequently measured in practice as in Figure 1 and hence those driving the value of the failure rate. This to say that the conventional Monte-Carlo analysis remains the reference both in terms of statistical intuition and accuracy. The deterministic approach, simple in mind as a DC sweep simulation, can actually be reinterpreted as a very special case of importance sampling [12], [37]. The V_{th} distributions have temporarily been modified for a quick and insightful access to the tail behaviours, here to determine the full failure region

Authorized licensed use limited to: Univ Catholique de Louvain/UCL. Downloaded on August 02,2022 at 08:17:02 UTC from IEEE Xplore. Restrictions apply.

and the borders. Because the threshold voltage distributions are assumed to be known, this information can be combined with the results of the DC simulation in order to reconstruct the proper statistics, that is to *compute and predict the failure rate*. This is precisely the purpose of Section IV.

f) Front lines and nonlinearities: One interesting feature of the proposed two-dimensional threshold voltage imbalance representation is to visually separate the defective latches from all the functional ones and to point out the borderline variations. The boundary between the *pass and fail regions* is well defined in Figure 6. About half of the pass diamond is included in the observation circle. In this bulk area, which strongly dominates the failure rate calculation, the border is reasonably *linear*. This general feature will also be confirmed for other (ppm failure rate) practical cases. Remarkably, the front lines are unequivocally determined by the *two single special points* attributed to the worst- and best-case scenarios reviewed in Section III-B.1. This fundamental observation is highly powerful: the failure region, as well as the *failure probability*, are promised to be quickly *predicted with only two insightful quantities* SNM^* *and* δV^*, extracted with *two very cheap DC simulations* and butterfly analyses.

The whole pass region roughly takes the shape of a diamond, approximated by a piecewise linear shape. Such closely multi-linear behaviour of the boundary, interpreted as the $SNM = 0$ contour line, could have been anticipated on the basis of previous works suggesting a linear evolution of the SNM with *moderate* variations of the individual transistor threshold voltages [4], [6], [10].

The diamond is perfectly symmetrical around the 45° line crossing the origin. This is because the latch is made of two identically designed CMOS inverters. The diamond also presents an axial symmetry in the orthogonal direction, but the corresponding diagonal does not always cross the origin: for the illustrated case in Figure 6, it lies in the third (bottom-left) quadrant. The center of gravity of the diamond is here approximately located in $(-50, -50)$, mV, as a consequence of the original imbalance between the nMOS and the pMOS transistors of the nominal inverter (noticeable in Figure 5: the metastable point is way below $V_{DD}/2$, revealing that the nMOS is significantly stronger than the pMOS). More generally, the deep bottom-left area (relative to the center of symmetry and not to the origin of the representation) contains the latches characterized by a weak nMOS and a strong pMOS, and conversely for the top-left area. This is why in the representation of Figure 6, there is an intermediate segment where the boundary smoothly bends, marking the transition between the two triangles of the pass diamond.

The observed deviations from the linear behaviour of the borders near the extreme bottom-left and top-right tips of the diamond can be attributed to the nonlinearity of the SNM with the largest threshold voltage variations in subthreshold SRAM, formerly reported in [7], [10], [28].

IV. Failure Rate Prediction

This section is dedicated to the calculation of the failure probability of the latch. We take a closer look at the first case study of Figure 1 and Figure 6. Empirical estimates of the fabrication yield and the failure rate are obtained from the intensive Monte-Carlo simulation which serves as a reference whose accuracy is statistically established. We next confirm that the dominant part of the defective latches lies outside the front lines. The corresponding probability is computed with a closed-form formula. The modelling error and the sources of discrepancy are identified and described.

TABLE I

Empirical Probabilities From Figure 7 and Failure Rate Prediction. The Colors and the Symbols Refer to the Points of Figure 7

Type	Symbols	Fraction [%]
Yield	• + •	98.155
Failure Rate (Monte-Carlo)	• + •	1.845
P (outside the front lines)	• + •	1.994
Closed-form formula P^*_{fail} (24)		1.978
P (*functional outside* the front lines)	•	0.251
P (*defective inside* the front lines)	•	0.102

We also apply our methodology to assess the functionality of a realistic ULP SRAM cell operating this time in nominal corner and temperature conditions, giving a failure probability of the order of the ppm. The effect of the supply voltage downscaling is readily studied.

We end up with a discussion of the performance of our approach in comparison with former works, in terms of estimated failure probability, relative accuracy, number of simulations and related computational cost.

A. Two-Dimensional Representation Methodology

1) Monte-Carlo Simulation as a Reference: We wish to estimate the yield of a given latch design. To that purpose, the principle of the standard Monte-Carlo analysis consists in simulating a very large number of latches (N_{latches}). Each latch is understood as a *Bernoulli trial*, that is a random experience with exactly two possible outcomes, “success” (functional) or “failure” (defective) [37], depending on the resulting SNM value. Within this theoretical framework, the cell failure rate is nothing but the probability of failure P_{fail} in a Bernoulli trial, i.e. the probability that some arbitrary, randomly drawn latch is defective. We form a numerical empirical *estimate* *Failure Rate* of the exact P_{fail} by counting the number of defective latches, $N_{\text{defective}}$:

$$P_{\text{fail}} \approx \mathit{Failure\ Rate} = \frac{N_{\text{defective}}}{N_{\text{latches}}}. \tag{21}$$

$\mathit{Yield} = 1 - \mathit{Failure\ Rate}$ is the corresponding cell yield.

The counted values of the *Yield* and the *Failure Rate* (21) are recorded in Table I. The estimator (21) is easily proved to be both unbiased and asymptotically consistent [37], which ensures that they converge to the true probability when $N_{\text{latches}} \to \infty$. For the current case, $N_{\text{latches}} = 100\,000$ guarantees that the probabilities are empirically estimated with an accuracy and a confidence level both superior to 95%, which constitutes an upper standard in the dedicated literature [12],

Authorized licensed use limited to: Univ Catholique de Louvain/UCL. Downloaded on August 02,2022 at 08:17:02 UTC from IEEE Xplore. Restrictions apply.

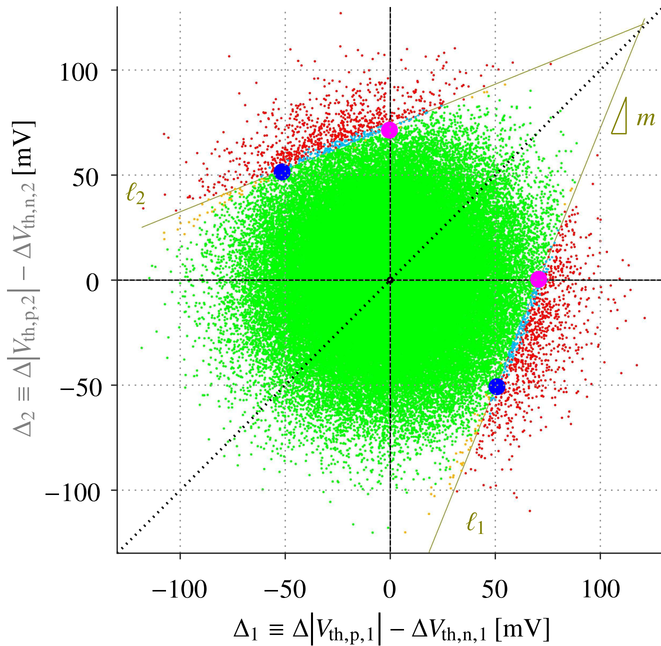

Fig. 7. Classification of the 100 000 samples of a Monte-Carlo simulation toward the calculation and the prediction of the failure rate, and the assessment of the related modelling error: Red •, defective latches. ↪ some are covered by orange •, defective *inside* the front lines (khaki). Green •, functional latches. ↪ some are covered by deep-sky-blue •, functional *outside* the front lines. Illustrated case: same as Figure 1 and Figure 6. $m = 2.47$.

[14], [17], [18], [37]. We will therefore assume that the Monte-Carlo estimate for the failure rate reported in Table I *is* the true probability for further comparisons with our model.

2) Closed-Form Prediction of the Failure Rate: Figure 7 maps the latches in the two-dimensional threshold voltage imbalance representation of Section III. Table I then compares the failure rate (21) to the fraction of the samples that lie *outside the front lines* ℓ_1 *and* ℓ_2. The relative difference between these two *empirical* probabilities is less than 8% and quantitatively confirms the conclusions of Section III-B.2: the failure probability can be accurately and efficiently calculated by counting the latches outside the two linear borders.

a) The slope m: Either front line is uniquely determined by the two special points respectively defined by the quantities SNM^* and δV^*. It is convenient to define a *slope* parameter,

$$m = \frac{SNM^*}{\delta V^* - SNM^*}. \tag{22}$$

actually the slope of ℓ_1. ℓ_2 is the mirror image of ℓ_1 about the 45° axis of reflection, so its slope is $1/m$. Let us focus for instance on the front line ℓ_1. When the two inverter imbalances Δ_1 and Δ_2 verify the line equation

$$\Delta_2 = m \cdot \Delta_1 - m \cdot (2 \cdot \delta V^*). \tag{23}$$

the SNM of the affected latch is equal to zero.

It is hard to make very general comments on m, for the reasons mentioned in the paragraph about SNM^* and δV^* of Section III-B.1. The inequality (12) gives us $m \geq 1$. The value of m mostly depends on the difference of strength between the nMOS and the pMOS transistors of the nominal inverter: the more severe the imbalance, the closer δV^* to SNM^* and hence the larger m. Even if the latch is made of hypothetical perfectly symmetrical CMOS inverters, m might still be greater than unity (e.g. $m = 1.5$), depending on the butterfly wing profile.

b) Failure rate formula: The line equations are straightforwardly written with only two parameters among SNM^*, δV^* and m, see (23) as an example. Next, the Gaussian distributions of $\Delta V_{\mathrm{th,n}}$ of $\Delta V_{\mathrm{th,p}}$ are entirely known, given the $\sigma_{V_{\mathrm{th,n}}}$ and $\sigma_{V_{\mathrm{th,p}}}$ of the technology. Either threshold voltage imbalance Δ_1 or Δ_2 is also normally distributed with a variance previously computed in (17). The proper combination of all those ingredients smartly leads to a *closed-form expression of the failure rate*:

$$P^*_{\mathrm{fail}} = \mathrm{erfc}\left(\sqrt{\frac{2\,m^2}{1+m^2}}\,\frac{\delta V^*}{\sqrt{\sigma^2_{V_{\mathrm{th,p}}} + \sigma^2_{V_{\mathrm{th,n}}}}}\right). \tag{24}$$

The full derivation of (24) is proposed in Appendix A. As noticed in Table I, the formula (24) can boast very accurate prediction of the true probability. The minor modelling error is carefully analysed thereafter.

It is really worth pointing out that (24) can be applied as soon as the worst- and best-case static noise margins SNM^* and δV^* are available (which requires two DC simulations per corner). The threshold-voltage standard deviations $\sigma_{V_{\mathrm{th,p}}}$ and $\sigma_{V_{\mathrm{th,n}}}$ can be extracted from the transistor statistical model of technology for the dimensions of the given bitcell design. m is indeed readily computed with (22) before being plugged into (24). In that regard, a key remark is that although the two-dimensional representation led to the closed-form prediction of the failure rate, the final and highly useful result (24) does not explicitly require to go back to this representation for every SRAM case study. The failure probability can be estimated by means of only the two cheap DC simulations described in Section III-B.1. We nevertheless believe that the threshold voltage imbalance representation *remains* a valuable tool, especially through the computational efficient deterministic approach presented in Section III-B.2, because it does provide the circuit designer a quick visual insight about the impact of local inverter imbalances on the cell reliability, relatively to the statistical variability of the technology thanks to the Rayleigh's equiprobability circles (Figure 6). Our methodology also offers a classification, a map of the results of Monte-Carlo simulations (Figure 1 and Figure 7), if these are still desired.

The erfc function already appeared in [8] in an attempt of estimating the tail of the data retention voltage (DRV) distribution. Wang *et al.* [8] relied on the assumption that the single-wing SNM is Gaussian and exploited the subsequent statistical model formerly developed in [7]. The formulas of [8] do explicitly involve the SNM distribution and thereby still require a thousand-point Monte-Carlo simulation to extract the mean and the standard deviation of the SNM. Our own approach overcomes this step as only the *nominal* static noise margins and the $\sigma_{V_{\mathrm{th}}}$'s of the technology stand in (24).

3) Accuracy and Modelling Error: The relative error on the failure probability is defined by

$$\varepsilon \equiv \frac{Failure\ Rate - P^*_{\mathrm{fail}}}{Failure\ Rate}. \tag{25}$$

Authorized licensed use limited to: Univ Catholique de Louvain/UCL. Downloaded on August 02,2022 at 08:17:02 UTC from IEEE Xplore. Restrictions apply.

The numerical values of Table I yield −7.2%, meaning that the true *Failure Rate* is slightly overestimated by the closed-form prediction P^*_{fail}. We can insightfully understand the sources of such an acceptable error thanks to Figure 7. We learn that the piecewise linear approximation of the failure boundary is fair even though imperfect as there are a few:

- in orange, defective latches lying *inside* the front lines;
- in deep sky blue, functional latches *outside* the front lines.

The corresponding probabilities are also recorded in Table I. Both are naturally quite small, by one order of magnitude less than the failure probability, in agreement with the error calculated above. Moreover, we notice that given the large number of Monte-Carlo trials, the *statistical error* is negligible compared to *modelling error*: on the failure rate:

$$P\,(\text{outside the front lines}) \approx P^*_{\text{fail}} \text{ for } N_{\text{latches}} \to \infty, \quad (26)$$

and the modelling error is then *exactly* identified as

$$\begin{aligned} &\mathit{Failure\ Rate} - P^*_{\text{fail}} \\ &\quad = P\,(\textit{defective inside}) - P\,(\textit{functional outside})\,. \quad (27) \end{aligned}$$

The validity of (27) can be empirically verified using the values of Table I. The accuracy of the prediction (24) may therefore be *qualitatively* assessed by visual inspection of Figure 7. As long as the number of "outliers" remains moderate compared to the total number of defective latches, the relative and absolute errors, (25) and (27), remain acceptable.

To conclude this part, we attribute the little modelling discrepancy mostly to the non-linear deviations of the borders between the pass and fail regions. We also remind that the two-dimensional threshold voltage imbalance representation of Figure 7 is a projection of the full four-dimensional surface, so that we cannot exclude that some latches, characterized by strong individual threshold voltage variations, wrongly land upon the projection and join the rank of outliers.

B. Application to a Realistic SRAM Cell

1) Description of the Design: Generally, we have seen reported in the literature *fabricated* SRAM cells whose functionality down to $V_{\text{DD}} = 200\,\text{mV}$ or lower is experimentally proven [3], [24], [26]. In 28 nm FD-SOI technology, the *Single P-Well* (SPW) 6T architecture was presented as a compact solution for high-density memory arrays operating over a wide supply voltage range [25], [38]. The SPW bitcell has been integrated in foundry SRAM macros that are used in various digital and mixed-signal systems on a chip, for instance [2]. Close to 400 mV minimum operating voltage and 300 mV DRV are achieved [38].

The 28 nm FD-SOI SPW cell constitutes a very generic study case of competitive foundry SRAM in a state-of-the-art industrial technology. We propose to apply our methodology in order to assess the low failure probability of the subthreshold design. We assume minimal transistor dimensions $L_{\text{n}} = L_{\text{p}} = 30\,\text{nm}$ and $W_{\text{n}} = W_{\text{p}} = 80\,\text{nm}$, together with $V_{\text{PW}} \equiv V_{\text{B}} = 0$ for ultra-low voltage operation [38], however without any loss of generality. The effect of aggressive supply voltage downscaling, down to the ultimate limits currently reported in the literature, is also readily studied without additional computational overhead.

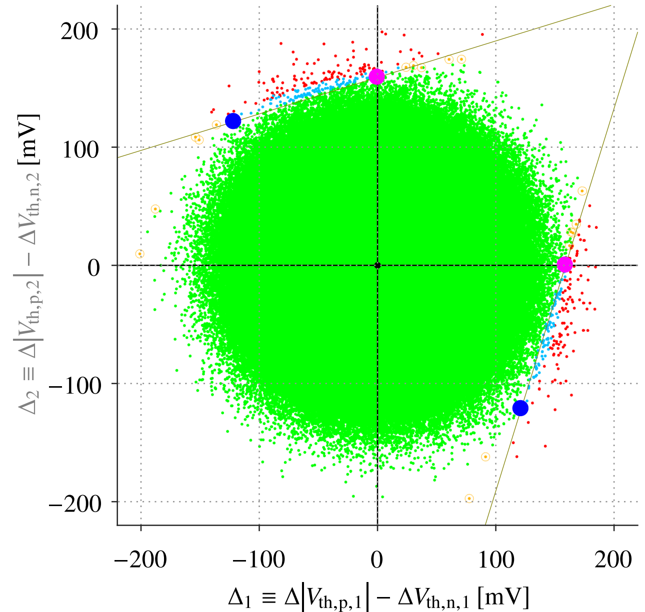

Fig. 8. Two-dimensional threshold voltage imbalance representation of the *five million* samples of a Monte-Carlo analysis for the SPW latch operating at $V_{\text{DD}} = 200\text{mV}$ and room temperature, in TT corner. Red + orange •, defective latches (*Failure Rate* = 47ppm) ↪ orange •, defective latches *inside* the front lines (khaki) (3.4ppm). Green + deep-sky-blue •, functional latches (cell *Yield* = 99.9953%). ↪ deep-sky-blue •, functional *outside* the front lines (33.8ppm). P (outside the front lines) = 77.4ppm (P^*_{fail} = 81ppm). Estimated $\sigma_{V_{\text{th,n}}} = 19\text{mV}$; $\sigma_{V_{\text{th,p}}} = 34\text{mV}$. $SNM^* = 61\text{mV}$; $\delta V^* = 79.5\text{mV}$. $m = 3.24$.

2) Prediction of the Failure Probability: Despite the detailed knowledge of the SPW SRAM bitcell under study and of the technology, as described in Section IV-B.1, the circuit designer had no efficient and fast analytical tool to predict the failure rate at given operating conditions (say $V_{\text{DD}} = 200\,\text{mV}$ at the start). Our closed-form formula (24) precisely aims to compensate for such breach. Following the proposed methodology, the worst- and best-case static noise margins are quickly extracted from two independent DC sweep simulations (see Section III-B.1 and especially the testbench presented in Figure 4) as $SNM^* = 61\,\text{mV}$ and $\delta V^* = 79.5\,\text{mV}$. Equation (22) yields $m = 3.24$ and we are able to provide an *a priori estimate* of the failure probability using (24): $P^*_{\text{fail}} = 8.1 \times 10^{-5} = 81\text{ppm}$. The prediction in itself is almost instantaneous, and the short total computation *time* only depends on the static-noise-margin simulations and extraction procedure (which are both software and user specific, hence hardly generally quantified; however, the total time will never exceed a few minutes).

a) Verification with monte-carlo and two-dimensional representation: In Figure 8, we show the two-dimensional threshold voltage representation of a massive conventional Monte-Carlo simulation of the SPW SRAM bitcell. The required huge number of latch samples, *five millions*, was selected following the accuracy and confidence guidelines of [12], [17], [37] A statistical accuracy and confidence both of 90% can be achieved provided that the number of simulated latches is larger than $271/P_{\text{fail}}$ [12]. The true failure rate P_{fail} is of course *unknown a priori*, but the analytical prediction P^*_{fail}

Authorized licensed use limited to: Univ Catholique de Louvain/UCL. Downloaded on August 02,2022 at 08:17:02 UTC from IEEE Xplore. Restrictions apply.

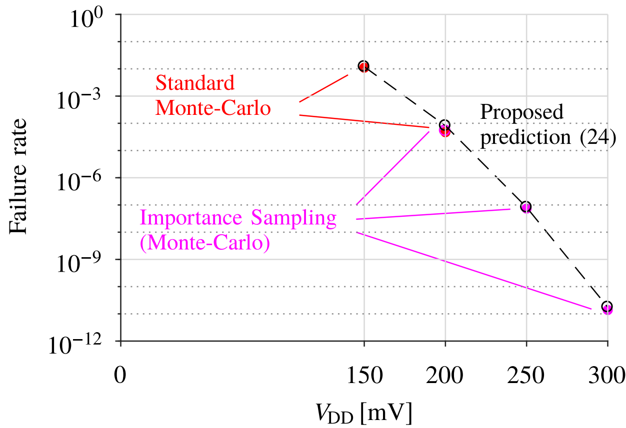

Fig. 9. Effect of the supply voltage downscaling on the failure rate of the single P-well SRAM cell.

could hopefully reasonably be plugged in the rule of thumb instead. A bit more than *three* millions is found, and *five* thus offers a little leeway to guard against the potential difference between P_{fail} and P^*_{fail} as well as against the statistical chance.

The numerical results of the Monte-Carlo analysis are summarized in the caption of Figure 8. The actual failure probability is empirically estimated by $Failure\ Rate$, here computed as 47ppm. It appears that the analytical formula (24) ($P^*_{\text{fail}} = 81$ppm) succeeds in correctly predicting the order of magnitude of a low $\sim$ ppm-range failure rate. For this study case, the relatively large discrepancy is explained, with the help of the representation of Figure 8, not as much by the low targeted probability as by the very pronounced nonlinearity of the border between the pass and fail regions: the closed-form formula slightly *over*estimates the number of defective cases, as the Monte-Carlo simulation reveals several latches that are *functional outside* the front lines, in the area near and between the two special points.

b) Discussion: We nevertheless emphasize that the observed modelling error must be put in a circuit and system design perspective: the cell failure probability is a metric that literally spans across orders of magnitude, between a few % and down to 10^{-9} or much lower. This is the reason why the calculation of the absolute error must be considered as a very harsh test, yet instructive. To the authors' opinion, providing a rigorously supported *upper bound* to the failure rate (that is, with a relative error not much greater than 50%) of some unknown realistic design *is* a *useful* result and the provided analytical estimation tool fullfills its function in that regard. At the light of Figure 8, ways of improvement and modelling refinement *are* possible if desired. For instance, the nonlinearities of the boundary could be captured by a two piecewise-linear approximation while remaining in the closed-form, analytical domain as in (24). A non-linear curve fitting approach would not eventually yield a compact analytical solution.

3) Effect of the Supply Voltage Variation: As a direct application of our methodology, we now propose to predict the impact of the supply voltage downscaling on the failure probability of the SPW latch cell. The trend is depicted in Figure 9 and is very similar to the hold [8] or read failure probability versus V_{DD} plotted for example in [10]. The steep, nearly exponential increase of the failure rate (or, equivalently, a drop of the cell yield) with reduced V_{DD} is mainly explained by the severe degradation of the noise margins at low voltages [1], [7]. Everywhere in the literature, the SNM is shown to decrease quite linearly with V_{DD} for *subthreshold* SRAMs [1], [7], [10], [23]. Thus, at the light of our prediction formula (24), it actually appears that the data points and the dashed curve of Figure 9 closely follow the complementary error function erfc$(\cdot)$ behaviour.

a) Comparison with monte-carlo with and without importance sampling: The predictions emanating from our closed-form formula are compared with Monte-Carlo estimates of the failure rate in Figure 9. The 200 mV exactly corresponds to the case treated in details and insightfully discussed in Section IV-B.2. At 150 mV, the probability lies in the % range and can thus also be easily checked with a standard Monte-Carlo simulation of lower computation time (100 000 samples).

Obviously, the verification of the smallest failure probabilities, such as those falling far below 10^{-9} around 300 mV, is the most challenging because the conventional Monte-Carlo is prohibitively expensive for those cases. Like [28], we then decided to implement a special method of importance sampling in which the distorted sampling functions of the V_{th}'s are Gaussian distributions with *inflated* standard deviations, with 1000 samples for all the three cases $V_{\text{DD}} = 200$ mV, 250 mV and 300 mV, while the inflation factor of the $\sigma_{V_{\text{th}}}$'s was selected as 1.5, 2 and 2.5, respectively. One should carefully keep in mind that, in absence of conventional Monte-Carlo results for the lowest failure rates, we do not have totally trustworthy reference values. We nevertheless think that the excellent matching between the standard Monte-Carlo and the IS methods at $V_{\text{DD}} = 200$ mV gives faith in the following comparisons.

Our formula succeeds in predicting the correct order of magnitude of the failure rate for all the supply voltage values. Furthermore, the effect of the voltage downscaling seems accurately captured by the quick predictive model. The uses of our methodology are very diverse and not restricted to the study of the supply voltage downscaling. One could of course exploits it to quickly *benchmark* different technologies for instance, as performed in [30], [32]. Generally, the proposed analytical approach appears really suitable and efficient for *comparative* evaluations as well as *sensitivity* estimates.

C. Discussion

The purpose of this section is to compare our methodology of closed-form failure rate prediction, which has been natively derived from experimental observations in two-dimensional threshold voltage imbalance representation, with other analytical or semi-analytical, that is non-Monte-Carlo, approaches reviewed from the literature. The quantitative comparison criteria are obviously the computational cost (which includes both the required number of DC simulations and the post-processing when it is significant), the accuracy as well as the lowest assessable (and reported) failure probabilities.

Authorized licensed use limited to: Univ Catholique de Louvain/UCL. Downloaded on August 02,2022 at 08:17:02 UTC from IEEE Xplore. Restrictions apply.

TABLE II

BENCHMARK OF DIFFERENT ANALYTICAL PREDICTION METHODS OF THE HOLD AND READ FAILURE RATES REGARDING SIMULATION AND PROCESSING COST

Reference	Condensed description of the methodology	# DC simulations	Required additional processing
	Conventional Monte-Carlo	$\sim 5\times 10^{6\dagger}$	
	Importance sampling	~ 1000	
This work	*Closed-form formula* (24) involving nominal SNM^* and δV^*	**2**	Noise margin extractions from traditional butterfly analyses
Saeidi [10]	Full analytical model of the SRAM hold/read operation mode	2	Accurate calibration of the parameters of the subthreshold current model (2)-(3)
Calhoun [7] Wang [8]	Statistical model of the SNM and DRV	$\sim 1000 - 5000$	Thousand-sample Monte-Carlo to estimate the μ and the σ of the model
Agarwal [6] (Mukhopadhyay [5])	*Linear* SNM (node voltage V) model and closed-form formula	5	Extraction of the sensitivities of the SNM (voltage V) to the four ΔV_{th}'s
Bang [15]	Non-linear transistor-aware SNM statistical model	145	Many coefficients to be calculated with a method of least squares
Zheng [28]	SNM non-linear formulation of reduced dimensionality	$1^\ddagger$	Computation of the mapping function $g(s)$ and piecewise linear approximation

† For Monte-Carlo, the realistic ~ ppm case study of Figure 8 (Section IV-B) is assumed.
‡ Claimed in [28] but not clarified.

1) Computational Cost: In Table II, we compare the gain in number of simulations of our methodology both with standard Monte-Carlo and importance sampling and with the other former analytical-inspired works reviewed in Section II.

Our formula (24) provides the circuit designer a quick estimate of the failure probability and requires only two prior DC sweep simulations to determine SNM^* and δV^*. The enormous reduction of simulation time with respect to Monte-Carlo methods is trivial and has been the central motivation for all the contributions collected in Table II, ours included. Although mathematically very sound, Calhoun [7]'s statistical model and Wang [8]'s subsequent work, takes us away from the dream of prediction the failure rate within a few seconds, due to the required thousand-point Monte-Carlo calibration. Bang [15]'s device-oriented statistical model does not require any Monte-Carlo simulation, but the number of parameters to be calculated (145 for the latch) is very large.

As emphasized in Table II, a few other analytical approaches can get by with a limited number of DC simulations. In the case of [6], the sensitivities of the SNM with respect to the transistor threshold voltage deviations should be re-extracted for each potential temperature or supply voltage variation, which could make architectural and/or technological study like Figure 9 or [30], [32] quite tedious and lengthy.

The full analytical model developed in [10] does even *not* involve any DC simulation of the SRAM cell and SNM extraction. The reason why is that the SNM is there available as an explicit symbolic function of all the transistor parameters; so are its mean and standard deviation and thereby its distribution to be eventually integrated. The two DC simulations indicated in Table II refer to the full $I_{\text{D,n}}(V_{\text{GS}}, V_{\text{DS}})$ and $I_{\text{D,p}}(V_{\text{SG}}, V_{\text{SD}})$ characteristics at given V_{DD} and T. The combined simulation and post-processing cost of [10]'s methodology is thus conditioned by the need for a very accurate calibration of the transistor compact model.

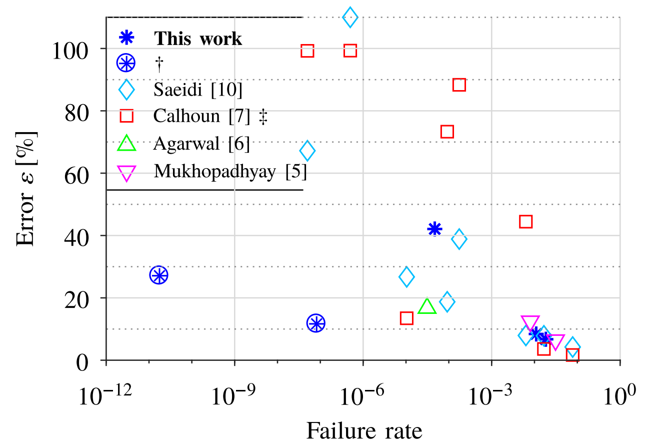

Fig. 10. Benchmark of the accuracy of different analytical prediction methods of the hold and read failure rates. The relative error from (25) has been computed for all the case studies across various targeted probabilities. For all the points, the true failure rate was always estimated from a conventional Monte-Carlo simulation. † Only for the ⊛ (blue), $V_{\text{DD}} = 250mV$ and $V_{\text{DD}} = 300mV$ extracted from Figure 9, the reference value relies on importance sampling (Section IV-B.3). ‡ The data for Calhoun [7] directly come from [10, Table I].

It is worth mentioning that all [5], [6], [10] predictions rely on linearised models, so that their accuracy and their validity for the lowest failure rates should carefully be inspected.

2) Accuracy and Evaluation Scope: Because it seemed difficult to fairly compare the accuracy of different methods applied to substantially different case studies, we propose the Figure 10 where to relative error (previously defined in (25)) is plotted versus the actual failure rate to be estimated. The Monte-Carlo analysis yielded the reference estimate, and we have limited this benchmark to the competitive analytical works recorded in Table II.

Clearly, all the works offer comparable good accuracy level for the largest %-range failure rates. Anyway, those cases, an example of which is the one firstly presented in Section IV-A of the current paper, should generally be

Authorized licensed use limited to: Univ Catholique de Louvain/UCL. Downloaded on August 02,2022 at 08:17:02 UTC from IEEE Xplore. Restrictions apply.

considered as academic cases to verify the sanity of a methodology with a reasonably intensive Monte-Carlo simulation.

The assessment of the lower probabilities, down to the ppm or below, is not systematically addressed in the literature, despite its huge practical interest, for sure because it is more challenging all in terms of accuracy, computational resources and verification ability. Our methodology was shown to successfully predict the order of magnitude of the smallest probabilities falling between 10^{-6} and 10^{-9}. Figure 10 suggests that the methodology even offers better accuracy than Calhoun and Chandrakasan [7] and Saeidi *et al.* [10]'s models. We may regret the absence of [15], [28] in the benchmark, missing because those papers focus on the modelling of the distributions but never report or compare failures probabilities.

V. Conclusion

In this paper, we have presented a novel non-Monte-Carlo methodology that can analytically predict the failure rate of subthreshold SRAM cells in retention mode. The proposed approach originally relies on a two-dimensional threshold voltage imbalance representation, whose mathematical soundness and physical relevance were supported by both subthreshold CMOS inverter modelling and SPICE simulations. A closed-form formula was derived and dispenses the circuit designer with Monte-Carlo simulations or approximate SNM statistical model. Rather, the method appeals to the traditional static noise margins and thereby offers insight and intuition about the SRAM stability, notably regarding the impact of the transistor imbalance and supply voltage or temperature variations. The methodology was successfully applied to a state-of-the-art SRAM cell in FD-SOI technology: the accuracy of the estimation of probabilities smaller than the ppm has been verified with reference Monte-Carlo simulations, without or with inflated standard deviation importance sampling. We have shown that we are competitive with the recent works in terms of accuracy while outperforming most of them regarding total computation cost and reported lowest assessable failure rates.

Appendix A
Derivation of the Failure Rate Formula

The two front lines equations read

$$\ell_1 \equiv \Delta_2 = m \cdot \Delta_1 - m \cdot (2 \cdot \delta V^*); \tag{28}$$
$$\ell_2 \equiv \Delta_2 = \frac{1}{m} \cdot \Delta_1 + 2 \cdot \delta V^*. \tag{29}$$

The fail region is made up of two distinct areas whose statistical contribution is significant, respectively below ℓ_1 and above ℓ_2. To each is associated an *event* and a probability:

$$\begin{aligned} P_1 &\equiv P\left((\Delta_1, \Delta_2) \text{ is below } \ell_1\right) \\ &= P\left(\Delta_2 \le m \cdot \Delta_1 - m \cdot (2 \cdot \delta V^*)\right) \end{aligned} \tag{30}$$
$$\begin{aligned} P_2 &\equiv P\left((\Delta_1, \Delta_2) \text{ is above } \ell_2\right) \\ &= P\left(\Delta_2 \ge (1/m) \cdot \Delta_1 + 2 \cdot \delta V^*\right). \end{aligned} \tag{31}$$

The failure rate is then predicted as

$$P^*_{\text{fail}} = P_1 + P_2. \tag{32}$$

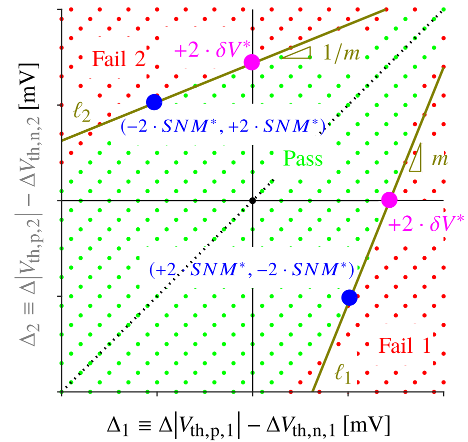

Fig. 11. Schematic two-dimensional representation for the derivation of the failure rate formula.

Because the latch is made of two identical inverters, the two front lines and fail regions are mirror images of each other about the 45° axis of reflection (dotted line in Figure 11). Consequently, we have $P_1 = P_2$. This could also be proved by direct calculation. We choose to present the derivation of P_1. We rewrite (31) by isolating the random terms

$$P_1 = P\left(\Delta_2 - m \cdot \Delta_1 \le -m \cdot (2 \cdot \delta V^*)\right) \tag{33}$$

and we define the random variable $X \equiv \Delta_2 - m \cdot \Delta_1$. X is a linear combination of the four independent threshold voltage variations (zero mean), so it is also normally distributed:

$$\begin{aligned} X &= \Delta|V_{\text{th,p,2}}| - \Delta V_{\text{th,n,2}} - m \cdot \Delta|V_{\text{th,p,1}}| + m \cdot \Delta V_{\text{th,n,1}} \\ &\sim \mathcal{N}\left(0, \left(1+m^2\right)\left(\sigma^2_{V_{\text{th,n}}} + \sigma^2_{V_{\text{th,p}}}\right)\right). \end{aligned} \tag{34}$$

The cumulative distribution function of the normalized random variable $Z \equiv X/\sigma_X$ is expressed as

$$F_Z(z) \equiv P(Z \le z) = \Phi\left(\frac{z-0}{\sigma_X}\right) \tag{35}$$
$$= \frac{1}{2}\left(1 + \text{erf}\left(\frac{z}{\sqrt{2}\,\sigma_X}\right)\right), \tag{36}$$

where in (35) Φ denotes the cumulative distribution function of the *standard* normal distribution, while the *error function*

$$\text{erf}(x) \equiv \frac{2}{\sqrt{\pi}} \int_0^x \exp\left(-u^2\right) \mathrm{d}u \tag{37}$$

was introduced in (36). Either form (35) or (36) is strictly equivalent as the two functions are tabulated and readily available in all the popular scientific calculation tools such as MATLAB®, the R software or Microsoft® Excel. Our own preference goes to the erf formulation (36) all through this work. From (33), we successively get

$$\begin{aligned} P_1 &= P\left(Z \le -\frac{2\,m\,\delta V^*}{\sigma_X}\right) = \frac{1}{2}\left(1 - \text{erf}\left(\frac{2\,m\,\delta V^*}{\sqrt{2}\,\sigma_X}\right)\right) \\ &= \frac{1}{2}\,\text{erf}\mathbf{c}\left(\frac{\sqrt{2}\,m\,\delta V^*}{\sigma_X}\right). \end{aligned} \tag{38}$$

The *complementary error function* $\text{erfc}(\cdot) \equiv 1 - \text{erf}(\cdot)$ has been introduced in (38), while erf was formerly defined in (37).

Authorized licensed use limited to: Univ Catholique de Louvain/UCL. Downloaded on August 02,2022 at 08:17:02 UTC from IEEE Xplore. Restrictions apply.

The failure rate is finally obtained combining (32), (34) and (38):

$$P^*_{\text{fail}} = 2 \cdot P_1 = \text{erfc}\left(\frac{\sqrt{2}\, m\, \delta V^*}{\sqrt{\left(1+m^2\right)\left(\sigma^2_{V_{\text{th,n}}} + \sigma^2_{V_{\text{th,p}}}\right)}}\right). \quad (39)$$

References

[1] A. P. Chandrakasan *et al.*, "Technologies for ultradynamic voltage scaling," *Proc. IEEE*, vol. 98, no. 2, pp. 191–214, Feb. 2010.

[2] D. Bol *et al.*, "SleepRunner: A 28-nm FDSOI ULP cortex-M0 MCU with ULL SRAM and UFBR PVT compensation for 2.6–3.6-μW/DMIPS 40–80-MHz active mode and 131-nW/kB fully retentive deep-sleep mode," *IEEE J. Solid-State Circuits*, vol. 56, no. 7, pp. 2256–2269, Jul. 2021.

[3] M. Alioto, "Ultra-low power VLSI circuit design demystified and explained: A tutorial," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 59, no. 1, pp. 3–29, Jan. 2012.

[4] 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, no. 4, pp. 658–665, Apr. 2001.

[5] S. Mukhopadhyay, H. Mahmoodi, and K. Roy, "Modeling of failure probability and statistical design of SRAM array for yield enhancement in nanoscaled CMOS," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 24, no. 12, pp. 1859–1880, Dec. 2005.

[6] K. Agarwal and S. Nassif, "Statistical analysis of SRAM cell stability," in *Proc. 43rd Annu. Conf. Design Autom. (DAC)*, 2006, pp. 57–62.

[7] B. H. Calhoun and A. P. Chandrakasan, "Static noise margin variation for sub-threshold SRAM in 65-nm CMOS," *IEEE J. Solid-State Circuits*, vol. 41, no. 7, pp. 1673–1679, Jul. 2006.

[8] J. Wang, A. Singhee, R. A. Rutenbar, and B. H. Calhoun, "Statistical modeling for the minimum standby supply voltage of a full SRAM array," in *Proc. 33rd Eur. Solid-State Circuits Conf. (ESSCIRC)*, Sep. 2007, pp. 400–403.

[9] A. Singhee and R. A. Rutenbar, "Statistical blockade: Very fast statistical simulation and modeling of rare circuit events and its application to memory design," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 28, no. 8, pp. 1176–1189, Aug. 2009.

[10] R. Saeidi, M. Sharifkhani, and K. Hajsadeghi, "Statistical analysis of read static noise margin for near/sub-threshold SRAM cell," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 61, no. 12, pp. 3386–3393, Dec. 2014.

[11] P. Weckx *et al.*, "Non-Monte-Carlo methodology for high-sigma simulations of circuits under workload-dependent BTI degradation—Application to 6T SRAM," in *Proc. IEEE Int. Rel. Phys. Symp.*, Jun. 2014, p. 5D-2.

[12] T. Haine, J. Segers, D. Flandre, and D. Bol, "Gradient importance sampling: An efficient statistical extraction methodology of high-sigma SRAM dynamic characteristics," in *Proc. Design, Automat. Test Eur. Conf. Exhib. (DATE)*, Mar. 2018, pp. 195–200.

[13] M. Wang, C. Yan, X. Li, D. Zhou, and X. Zeng, "High-dimensional and multiple-failure-region importance sampling for SRAM yield analysis," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 25, no. 3, pp. 806–819, Mar. 2017.

[14] M. Wang *et al.*, "Efficient yield optimization for analog and SRAM circuits via Gaussian process regression and adaptive yield estimation," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 37, no. 10, pp. 1929–1942, Oct. 2018.

[15] B.-J. Bang, H. Kwon, Y. H. Kim, K. Cho, and H.-S. Kim, "Statistical modeling of read static noise margin for 6-transistor SRAM cell," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2019, pp. 1–4.

[16] D. D. Weller, M. Hefenbrock, M. S. Golanbari, M. Beigl, J. Aghassi-Hagmann, and M. B. Tahoori, "Bayesian optimized mixture importance sampling for high-sigma failure rate estimation," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 39, no. 10, pp. 2772–2783, Oct. 2020.

[17] X. Shi, H. Yan, J. Wang, J. Zhang, L. Shi, and L. He, "An efficient adaptive importance sampling method for SRAM and analog yield analysis," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 39, no. 12, pp. 4999–5010, Dec. 2020.

[18] D. D. Weller, M. Hefenbrock, M. Beigl, and M. B. Tahoori, "Fast and efficient high-sigma yield analysis and optimization using kernel density estimation on a Bayesian optimized failure rate model," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 41, no. 3, pp. 695–708, Mar. 2022.

[19] S. Gupta and B. H. Calhoun, "Dynamic read VMIN and yield estimation for nanoscale SRAMs," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 68, no. 3, pp. 1171–1182, Mar. 2021.

[20] S. Gupta and B. H. Calhoun, "Dynamic write VMIN and yield estimation for nanoscale SRAMs," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 68, no. 12, pp. 5038–5048, Dec. 2021.

[21] G. Roy, A. R. Brown, F. Adamu-Lema, S. Roy, and A. Asenov, "Simulation study of individual and combined sources of intrinsic parameter fluctuations in conventional nano-MOSFETs," *IEEE Trans. Electron Devices*, vol. 53, no. 12, pp. 3063–3070, Dec. 2006.

[22] M. H. Abu-Rahma and M. Anis, "A statistical design-oriented delay variation model accounting for within-die variations," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 27, no. 11, pp. 1983–1995, Nov. 2008.

[23] L. Chang *et al.*, "Stable SRAM cell design for the 32 nm node and beyond," in *Dig. Tech. Papers. Symp. VLSI Technol.*, Jun. 2005, pp. 128–129.

[24] J. P. Kulkarni, K. Kim, and K. Roy, "A 160 mV robust Schmitt trigger based subthreshold SRAM," *IEEE J. Solid-State Circuits*, vol. 42, no. 10, pp. 2303–2313, Oct. 2007.

[25] R. Ranica *et al.*, "FDSOI process/design full solutions for ultra low leakage, high speed and low voltage SRAMs," in *Proc. Symp. VLSI Technol.*, Jun. 2013, pp. T210–T211.

[26] B. Zhai, S. Hanson, D. Blaauw, and D. Sylvester, "A variation-tolerant sub-200 mV 6-T subthreshold SRAM," *IEEE J. Solid-State Circuits*, vol. 43, no. 10, pp. 2338–2348, Oct. 2008.

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

[28] N. Zheng and P. Mazumder, "Modeling and mitigation of static noise margin variation in subthreshold SRAM cells," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 64, no. 10, pp. 2726–2736, Oct. 2017.

[29] I. J. Chang, J.-J. Kim, S. P. Park, and K. Roy, "A 32 kb 10T sub-threshold SRAM array with bit-interleaving and differential read scheme in 90 nm CMOS," *IEEE J. Solid-State Circuits*, vol. 44, no. 2, pp. 650–658, Feb. 2009.

[30] D. Bol, R. Ambroise, D. Flandre, and J. D. Legat, "Interests and limitations of technology scaling for subthreshold logic," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 17, no. 10, pp. 1508–1519, Oct. 2009.

[31] E. Seevinck, F. J. List, and J. Lohstroh, "Static-noise margin analysis of MOS SRAM cells," *IEEE J. Solid-State Circuits*, vol. 22, no. 5, pp. 748–754, Oct. 1987.

[32] S. Hanson, M. Seok, D. Sylvester, and D. Blaauw, "Nanometer device scaling in subthreshold logic and SRAM," *IEEE Trans. Electron Devices*, vol. 55, no. 1, pp. 175–185, Jan. 2008.

[33] J. R. Hauser, "Noise margin criteria for digital logic circuits," *IEEE Trans. Educ.*, vol. 36, no. 4, pp. 363–368, Nov. 1993.

[34] M. Yamaoka, K. Osada, R. Tsuchiya, M. Horiuchi, S. Kimura, and T. Kawahara, "Low power SRAM menu for SOC application using Yin-Yang-feedback memory cell technology," in *Symp. VLSI Circuits Dig. Tech. Papers*, Jun. 2004, pp. 288–291.

[35] F. Olivera and A. Petraglia, "Analytic modeling of static noise margin considering DIBL and body bias effects," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2017, pp. 1–4.

[36] M. K. Simon, *Probability Distributions Involving Gaussian Random Variables: A Handbook for Engineers and Scientists*. Berlin, Germany: Springer, 2007.

[37] R. Srinivasan, *Importance Sampling: Applications in Communications and Detection*. Springer, 2002.

[38] O. Thomas *et al.*, "Dynamic single-p-well SRAM bitcell characterization with back-bias adjustment for optimized wide-voltage-range SRAM operation in 28 nm UTBB FD-SOI," in *IEDM Tech. Dig.*, San Francisco, CA, USA, Dec. 2014, pp. 3–4.

Authorized licensed use limited to: Univ Catholique de Louvain/UCL. Downloaded on August 02,2022 at 08:17:02 UTC from IEEE Xplore. Restrictions apply.