Stability and Yield-Oriented Ultra-Low-Power Embedded 6T SRAM Cell Design Optimization

Adam Makosiej[1], Olivier Thomas[3], Andrei Vladimirescu[1,2], and Amara Amara[1]
[1]Institut Supérieur d'Electronique de Paris, France
[2] Berkeley Wireless Research Center, UC Berkeley
[3] CEA, LETI, MINATEC, F-38054, Grenoble

Abstract: This paper presents a methodology for the optimal design of CMOS 6T SRAM ultra-low-power (ULP) bitcells minimizing power consumption under strict stability constraints in all operating modes. An accurate analytical SRAM subthreshold model is developed for characterizing the cell behavior and optimizing its performance. The proposed design approach is demonstrated for an SRAM implemented in a 32nm CMOS UTBB-FDSOI technology. Stable operation in both read and write is obtained for the optimized cell at V_{DD}=0.4V. Moreover, in the optimization process the standby and active power were reduced up to 10x and 3x, respectively.

I. INTRODUCTION

With the constant scaling of Silicon technologies the reduction in leakage power has become one of the main challenges of modern integrated circuit (IC) design. In today's systems-on-a-chip (SOC) very often most of the chip area is taken by embedded SRAM, which leads in some cases to the leakage power dominating the overall power consumption. Therefore, for ultra-low-power design, suppressing leakage current becomes crucial. Sub-threshold operation is a particularly attractive solution for SRAMs, as lowering supply voltage does not only reduce the leakage in retention, but also reduces dynamic power consumption in active mode. However, the ever increasing parameter variation caused by constant technology scaling makes the application of this approach difficult due to the problems with maintaining sufficient stability of the memory.

Previous work on this subject focused mostly on fast assessment of the lowest applicable supply voltage in retention, for which the cell can still retain its data, also known as the Data Retention Voltage (DRV). This problem was first investigated by the authors in [1], where a direct equation for DRV is presented. This work was further extended in [2], where the simple equations describing the voltage transfer curves (VTCs) of SRAM half-cells in retention and read modes are included. The authors also note, that for proper DRV evaluation the focus needs to be on the tail of the distribution and demonstrate the equations describing the PDF and CDF in this region. The presented equations are based on the distribution of HSNMH or HSNML, Hold Static Noise Margin High and Low, taken as upper or lower square between the butterfly curves. This work is further extended in [3-4], where more advanced and efficient algorithms for DRV estimation are presented. Finally, the same approach as in [3-4] was applied in [5] to active modes to estimate the minimum applicable supply voltage (V_{MIN}) in read and write operations.

In this work instead of evaluating the DRV and V_{MIN} for the different operation modes for a specific technology, we make a general-case analysis of the best cell conditions for which the optimum balance between those values is met. As a first step an analytical model is developed allowing an accurate estimation of the Static Noise Margin (SNM) [6] for retention (HSNM), read (RSNM) and write (WSNM). The basic equations for retention and read were presented in [3]. Here, we additionally include the DIBL, body factor and all cell voltages as parameters and demonstrate the equation for WSNM. Thus a set of equations is obtained allowing an accurate evaluation of SRAM stability in all operation modes and an assessment of write assist techniques influence.

The proposed model is applied to estimate the best operating conditions of the SRAM cell from a stability perspective in the three operation modes, and for developing a CAD optimization procedure for the SRAM design. In this work, the proposed approach is applied to the 32nm UTBB-FDSOI [7] process; however, the presented methodology is universal.

II. SRAM TECHNOLOGY AND PERFORMANCE METRICS

The UTBB-FDSOI device [3] (Fig.1) consists of an undoped Silicon thin film on a thin Buried Oxide (BOX) layer of thickness T_{BOX} (10nm<T_{BOX}<30nm) covering a highly doped Back Plane (BP) (Fig.1). Reducing the BOX thickness and doping the BP (i) boosts the channel electrostatic control (ii) gives the possibility of obtaining a V_T modulation by applying different kind of BP doping using a single gate stack work function and (iii) results in a very high body factor for V_T adjustment, reaching 60-70mV/V for T_{BOX} = 25 nm.

Random Dopant Fluctuations (RDF) are the most important factor in process variations in CMOS bulk devices. Since in UTBB-FDSOI the thin film is undoped and the V_T is modified through the application of a different BP and/or body bias, the standard deviation σ_{VT} is expected to be almost half that of typical bulk with an A_{VT} of 1.1mVμm [8]. An additional feature is the availability of multiple V_Ts, such as high-V_T (HVT), standard-V_T (SVT) and low-V_T (LVT).

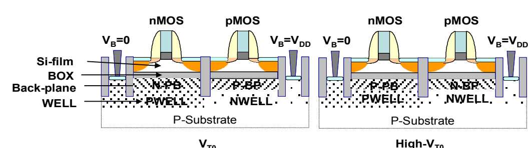

Fig 1. Cross section view of Multi-V_T UT2B-FDSOI NMOS and PMOS device configurations

978-3-9810801-8-6/DATE12/©2012 EDAA

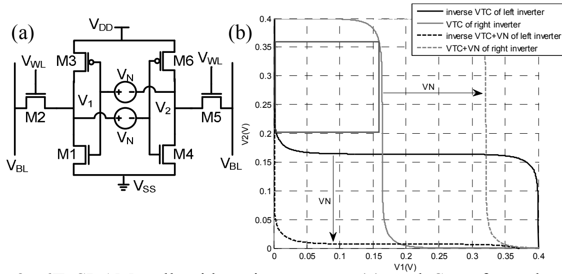

Fig.2. 6T SRAM cell with noise sources (a) and Set of regular butterfly curves and butterfly curves shifted by VN (SNM) value, and a graphical representation of SNM (b)

Static noise margin (SNM) is the key performance metric for SRAM cells and was first introduced in [6]. It can be described as the largest value of noise voltage between the two inverters in a 6T memory cell (Fig. 2a), for which the cell can still retain its data; it is represented graphically in Fig. 2b as the largest square that can fit between the "butterfly curves", which are obtained from a direct and an inverse voltage-transfer curve (VTC) of each cell inverter. This approach applies to both HSNM and RSNM evaluation. WSNM however, is defined as the side of the smallest square that can fit between the "butterfly curves", where the write VTC is obtained for the bitline set to GND.

The SNM model is implemented in Matlab and applied to optimize yield by maximizing μ/6σ of the SNM in the presence of local, Gaussian, V_T variation of ±3σ. The value of 6σ corresponds to roughly one cell in 505 million failing and is a typical yield requirement for modern SRAM design.

III. SUBTHRESHOLD SNM MODEL

A. Overview

For evaluation of the Voltage-Transfer Characteristic (VTC) of an inverter in sub-threshold, the exponential transistor current equation is used:

$$I_D = I_0 \exp\left(\frac{V_{GS}-V_T}{nV_{th}}\right)\left(1-\exp\left(-\frac{V_{DS}}{V_{th}}\right)\right) \quad (1)$$

$$V_{th} = kT/q \quad (2)$$

$$n = S/(V_{th}\ln(10)) \quad (3)$$

where I_0 is the drain current for $V_T = V_{GS}$ and n is the sub-threshold factor calculated based on the sub-threshold slope S. This above equation is valid for both bulk and SOI CMOS circuit operation. For simplicity, both in equations and throughout this paper, we will treat PMOS parameters as positive values.

B. SNM equations

Equations for read and retention modes were derived using Eq. 1 and applying it to Kirchoff equations. The D, A and L indexes correspond to Driver, Access and Load transistors, respectively, and indexes 1 and 2 refer to the left and right half cells. The remaining parameters are as follows: η- DIBL, γ- body factor, I- transistor current for $V_T = V_{GS}$, S- subthreshold slope, V_{th}- thermal voltage, V_{BN}- NMOS body bias, V_{BL}- read bitline voltage, V_{BBL} - write bitline voltage. In the read equation a linear dependence is assumed between V_{BN} and V_T.

For retention mode we can write $i_{D_1} = i_{L_1}$ and $i_{D_2} = i_{L_2}$ since both access transistors are off. Eq. 4 represents the equation for the inverse VTC curve:

$$V_1 = \frac{(S_{D1}S_{L1})}{\ln 10(S_{D1}+S_{L1})}\left(\ln\frac{I_{L1}}{I_{D1}} + ln\left(\frac{1-\exp\left(\frac{V_2-V_{DD}}{V_{th}}\right)}{1-\exp\left(\frac{V_{SS}-V_2}{V_{th}}\right)}\right)\right) - V_2\frac{S_{L1}\eta_{D1}+S_{D1}\eta_{L1}}{(S_{D1}+S_{L1})} + V_{SS}\frac{S_{L1}(1+\eta_{D1})}{(S_{D1}+S_{L1})} + V_{DD}\frac{S_{D1}(1+\eta_{L1})}{(S_{D1}+S_{L1})} + \frac{S_{D1}S_{L1}}{S_{D1}+S_{L1}}\left(\frac{V_{TD1}}{S_{D1}} - \frac{V_{TL1}}{S_{L1}}\right) \quad (4)$$

In read analysis it can be assumed that when input voltage value is low and thus the voltage of the internal node V_1 (or V_2) is at V_{DD}, the current through the access transistor is negligible. As the input voltage increases, both NMOS transistors become dominant and the PMOS current can be omitted instead. Therefore, the read plot is a piecewise combination of the curve obtained from Eq. 4 and Eq. 5. The latter equation corresponds to read mode and is obtained with the afore-mentioned assumption that at the onset of read operation, the current through the load PMOS can be neglected. Therefore, it represents the inverse VTC obtained for $i_{D_1} = i_{A_1}$:

$$V_1 = -\frac{S_{D1}}{\ln 10}\left(\ln\frac{I_{D1}}{I_{A1}} + ln\left(\frac{1-\exp\left(\frac{V_{SS}-V_2}{V_{th}}\right)}{1-\exp\left(\frac{V_2-V_{BL}}{V_{th}}\right)}\right)\right) - V_2\frac{S_{D1}+S_{A1}\eta_{D1}+S_{D1}\eta_{A1}+\gamma_{A1}\eta_{D1}}{S_{A1}} + V_{WL}\frac{S_{D1}}{S_{A1}} + V_{BL}\frac{S_{D1}}{S_{A1}}\eta_{A1} + V_{SS}(1+\eta_{D1}+\gamma_{D1}) + V_{BN}\frac{S_{D1}\gamma_{A1}-S_{A1}\gamma_{D1}}{S_{A1}} + S_{D1}\left(\frac{V_{TD1}}{S_{D1}} - \frac{V_{TA1}}{S_{A1}}\right) \quad (5)$$

It can also be noted that bitline voltage V_{BL} and wordline voltage V_{WL} have been used as parameters. Normally both are simply set to V_{DD} unless stated otherwise.

WSNM evaluation uses the read curve and write-mode equation curve obtained with the assumption, that the current

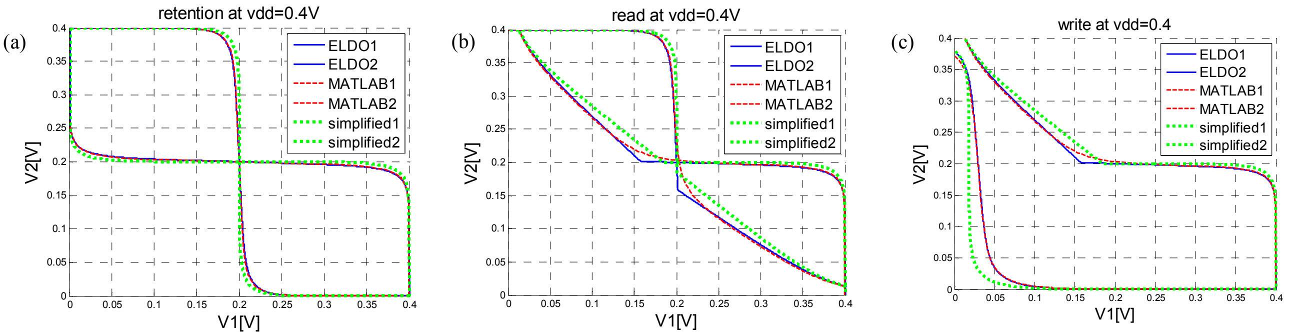

Fig.3 Comparison of equation (dashed), simple equation (dotted) and SPICE (solid) butterfly curves for retention (a), read (b) and write (c) at V_{DD}=0.4

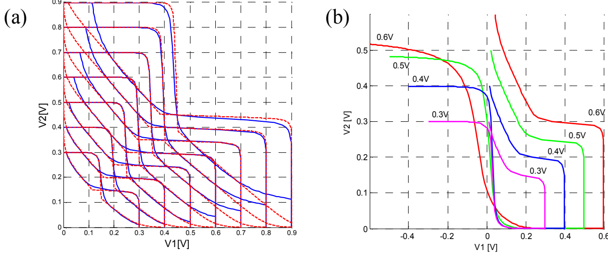

Fig.4 Comparison of curves based on the equations (dashed, red) and compact model UTBB-FDSOI model (solid, blue) for read (a) and illustration of write curve shape modification with V_{DD} changes (b)

flowing through the driver NMOS is negligible with the bitline set to GND. Contrary to the read case the final plot is not a piecewise combination of two equations, because in the input voltage range where the driver NMOS current would become a factor, the cell is already flipped as it can be noted in Fig. 3c. Hence, only one equation models the write curve:

$$V_1 = \frac{S_{L2}}{\ln 10}\left(\ln\frac{I_{L2}}{I_{A2}} + ln\left(\frac{1-\exp\left(\frac{V_2-V_{DD}}{V_{th}}\right)}{1-\exp\left(\frac{V_{BBL}-V_2}{V_{th}}\right)}\right)\right) - V_2\frac{S_{L2}\eta_{A2}+S_{A2}\eta_{L2}}{S_{A2}} - V_{WL}\frac{S_{L2}}{S_{A2}} + V_{BBL}\frac{S_{L2}}{S_{A2}}(1+\eta_{A2}+\gamma_{A2}) + V_{DD}(1+\eta_{L2}) - V_{BN}\frac{S_{L2}}{S_{A2}}\gamma_{A2} + S_{L2}\left(\frac{V_{TA2}}{S_{A2}} - \frac{V_{TL2}}{S_{L2}}\right) \quad (6)$$

C. Model accuracy

Fig. 3 compares the butterfly curves obtained from the above equations (MATLAB) to the simplified version of the equations [2] (without DIBL and the influence of the body factor during read) and SPICE simulations which are used as reference. The SPICE FDSOI compact model used is a surface potential based model using similar approach as the PSP model. The model's core has been finely developed to take into account the intrinsic effects of ultra-thin undoped FDSOI MOSFET, such as the interface coupling. Moreover, it includes a charge model, short channel effects, DIBL, narrow-width effect, GIDL/GISL current, self-heating effects, and gate currents. The device parameters have been obtained from experimental data.

In order to properly evaluate the SRAM cell stability, it is crucial to perform statistical simulations with the transistor V_Ts varying in the range defined by the process. As a result, in some iterations the V_T shift may lead to a part of the cell transistors to operate at V_{GS}>V_{DD}. For this reason and also to evaluate the utility of the model for higher supply voltages, it is important to analyze its accuracy in the above-subthreshold region. The results show a good agreement between the proposed model and SPICE simulations. The inaccuracy of model [2] is strongly improved by including the electrostatic device parameters. Fig. 4a shows the good behavior of the model on a large range of voltages (V_{TN}=451mV, $|V_{TP}|$=415mV). Above 0.8V, a misalignment can be observed due to the high-gain area of the VTC curve being at the edge of the subthreshold range for both transistors. The read part of the curve shows an ever-increasing inaccuracy with V_{DD} on the edges of the direct and inverse read VTCs. However, since the RSNM is evaluated as the side of the largest square fitted between the butterfly curves, it can be noted, that the part containing the lower-left corner of such a square will always be positioned in the part which remains very accurately simulated even up to V_{DD}=0.9V.

The WSNM evaluation with the presented model is strictly limited to the subthreshold region in the current implementation. This is caused by the fact that at the onset of the write operation both transistors are already at V_{GS}=V_{DD}. Moreover, the strength ratio between access NMOS and load PMOS is the key factor for WSNM and it will vary significantly depending if the transistors operate above- or below threshold. As plotted in Fig. 4b, with the increase of V_{DD} the whole shape of the write curve changes, which is not modeled by Eq. 6. In order to better illustrate the modification of the shape of the write curve, Fig 4b plots the VTC curves for a range of the input voltage between ±V_{DD}. It can be noted, that for V_{DD}>0.4V, the transition slope of the non-inverted VTC decreases leading to an increase of WSNM, as it is evaluated as the side of the smallest square fitted between both curves. Since this behavior is not properly modeled by Eq. 6, the WSNM will be strongly underestimated for the cell operating above subthreshold. Care has to also be taken for statistical simulations performed at the edge of the subthreshold region, as it will occur in some iterations, due to random V_T variations, either NMOS or PMOS may operate in the strong inversion. This in turn, will lead to an underestimation of WSNM and finally of the stability.

IV. Optimum Cell Condition Evaluation

A. Overview

In order to estimate the best cell conditions defined in this work as the optimum V_{TN}/V_{TP} ratio, a series of statistical simulations using the equation-based model were performed. For each V_{TN} ,$|V_{TP}|$ set a Monte Carlo (MC) analysis was run,

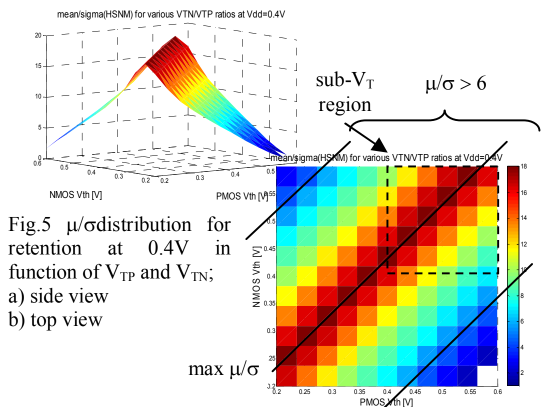

Fig.5 μ/σdistribution for retention at 0.4V in function of V_{TP} and V_{TN};
a) side view
b) top view

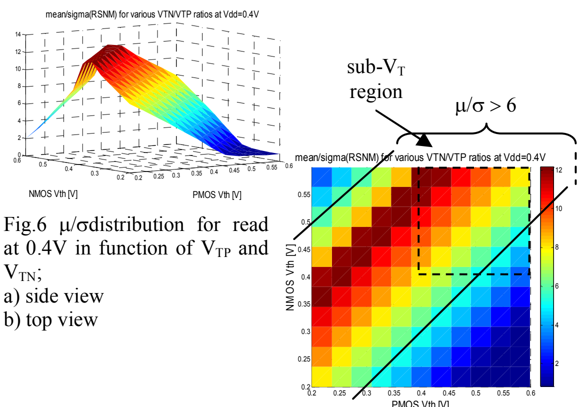

Fig.6 μ/σdistribution for read at 0.4V in function of V_{TP} and V_{TN};
a) side view
b) top view

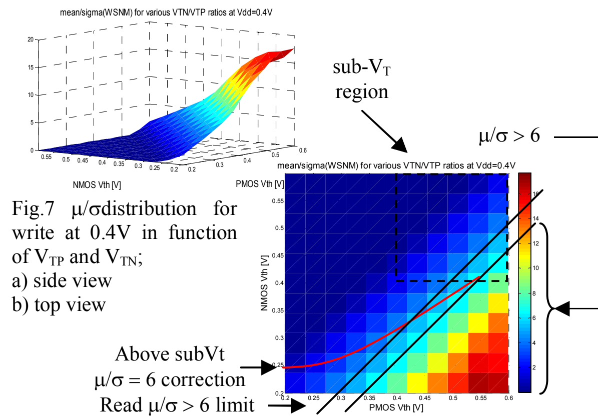

Fig.7 μ/σdistribution for write at 0.4V in function of V_{TP} and V_{TN};
a) side view
b) top view

the mean (μ) and standard deviation (σ) noise margin were extracted, and the μ/σ ratio was plotted on the Z axis (Figs.5-7) [9]. For SRAM, the μ/σ should be higher than 6 to meet the stability criteria. The cell used for simulations was sized with W/L of 198n/38n for driver NMOS, 153n/45n for access NMOS and 63n/38n for load PMOS. Fig.5 depicts the result for retention. Since the simulation was performed at V_{DD}=0.4V, the region of μ/σ>6 in the top-view plot is very large and the maximum μ/σ value is found for a V_T difference $\Delta V_{THOLD}=V_{TN}-|V_{TP}|\simeq 36mV$.

As shown in Fig. 6, the read operation also maintains the required stability margin in a wide range of the V_{TN}/V_{TP} ratio with the maximum shifted towards high V_{TN}. It is important to note the linear decay of the read μ/σ as a function of V_{TN} and $|V_{TP}|$ below the optimum and towards low values of V_{TN} and high values $|V_{TP}|$. This can be explained by the fact that in this region σ remains constant and μ varies linearly. This means, that from MC simulation it is possible to estimate the read stability for other cases as well. Naturally this applies only as long as the evaluation is made in the linear μ/σ range in read.

Fig. 7 depicts the same set of plots for write. Ensuring proper stability in this operation is the main challenge in subthreshold, as highlighted by the small region where μ/σ>6, see Fig. 7 top view. Comparing the read and write operations it can be noted that the regions where μ/σ>6 are mutually exclusive with each other, making it impossible to ensure stable operation in both active modes in subthreshold voltage.

B. Write assist techniques

There are a number of techniques to boost the stability during the write operation. In this work we focus on 3 of them: (1) increase V_{SS} above GND, (2) under-drive the bitline (V_{BL}) below GND and (3) increase the wordline voltage (V_{WL}) above V_{DD}. Naturally, both (1) and (3) will lead to a decrease of read stability in all cells on the same row (3) or column (1), therefore the influence of both on read has to be evaluated as well. Due to the imperfect accuracy of the equation-based write model above subthreshold, the following analysis was made using ELDO for better demonstration. Fig. 8 depicts the behavior of read and write μ/σ for each technique in a |100mV| offset range from the nominal value of the corresponding voltages, applied to the cell with V_{TN}=451mV and $|V_{TP}|$=415mV (optimized for retention). Clearly, the application of assist techniques can improve write stability significantly. It can be noted, that setting V_{BL}=-0.1V gives the write μ/σ=5.66, while maintaining read μ/σ>9. As presented in Fig. 8, the balance between these values can be further adjusted by increasing either V_{SS} or V_{WL}. Techniques (1) and (3) are also efficient however the opposite influence on the other active mode makes it difficult to find the proper balance to satisfy both, especially if global and temporal variations are considered. Fig. 9 depicts the results of the same write assist technique influence analysis for another V_T set, for which the initial read μ/σ=6 (V_{TN}=451mV, $|V_{TP}|$=528mV). In such a case, the write μ/σ for |offset=0| is equal to 4.5 and can be easily increased to 6 by, for example, setting V_{BL} to -40mV.

C. Cell sizing

The write stability is clearly the most critical parameter for subthreshold operation. Since it is not possible to obtain μ/σ>6 in all operation modes regardless of the V_{TN}/V_{TP} ratio, modifying the cell sizing may be an attractive solution. Three cell sizings (W/L) are considered: (1 - default)- 63n/38n, 198n/38n, 153n/45n (load, driver, access), (2)- 63n/38n, 198n/38n, 198n/45n (load, driver, access) and (3)- 63n/45n, 198n/45n, 198n/38n (load, driver, access). The sizing is modified under the condition that the area of the cell should not increase and hence, the sum of NMOS gate lengths must be ≤83n and NMOS widths must be ≤198n. Due to process considerations also the PMOS and driver NMOS gate lengths should be equal. The main reason for such analysis is that since write is the main limitation for subthreshold operation, the stability in this mode has to be boosted. This can be obtained by increasing the strength of access NMOS as compared to the load PMOS. Cell (2) obtains this by increasing the width of the access transistor. In case (3), the length of the inverter transistors is increased with the

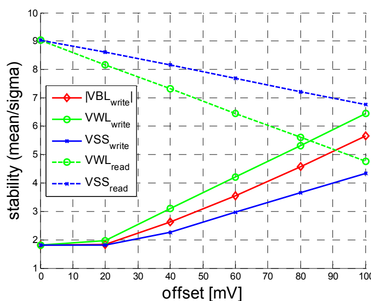

Fig.8 Assessment of write assist techniques on read and write μ/σ (from ELDO simulations) for V_Ts adjusted for top retention stability

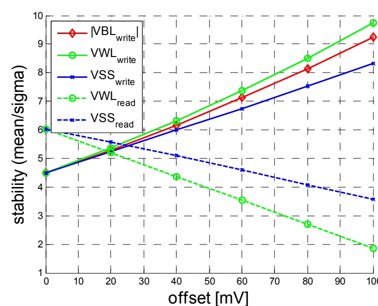

Fig.9 Assessment of write assist techniques on read and write μ/σ (from ELDO simulations) for V_Ts adjusted for initial read μ/σ=6

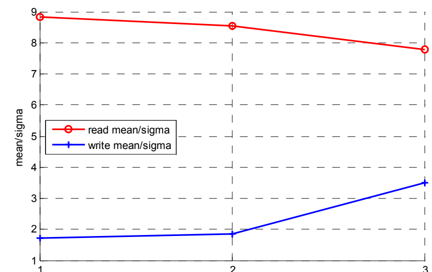

Fig.10 Assessment of read and write μ/σ behavior for (1), (2) and (3) cell setups

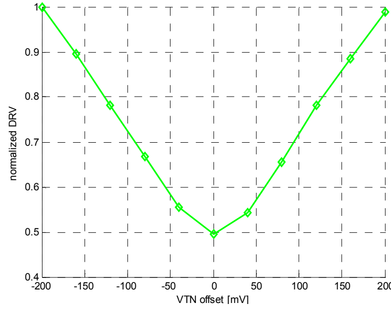

Fig.11 Normalized DRV In function of V_{TN} offset from the optimum V_{TN}/V_{TP} ratio

simultaneous decrease of access transistor length. However, as the access transistor becomes significantly stronger (in particular for (3)), as compared to the default case (1), the magnitude of the bit line factor (NBL) has to be taken into account as well. NBL is defined as the ratio between cell current during read and the access transistor leakage on the side storing "0" and should be at least 10x higher than the number of cells in the column. Indeed, the NBL for (3) drops by 50% as compared to (1), but still remains very high at 100k. Due to the strong influence of temperature on leakage the similar analysis was also performed for T=85℃. As expected, the NBL dropped significantly reaching as low as 13k for (3) and 27k for (1). These values are still sufficient but the magnitude of the change clearly shows how for some technologies this can become a major issue. The values of μ/σ for read and write for cases (1,2,3) are depicted in Fig.10. It can be noted that compared to (1), (3) has twice higher write stability of 3.5 with the read stability dropping from 8.83 to 7.8 and still maintaining the sufficient margin for further write assist techniques application.

V. SRAM CELL OPTIMIZATION

A. Overview

The ULP SRAM cell design is focused on reducing leakage, as typically it is the main contributor to the overall power consumption. Therefore minimizing DRV becomes the main concern. Fig. 11 depicts the normalized DRV value in function of the V_{TN} offset (relative to V_{TN}=451mV for which minimum DRV is obtained) and a fixed $|V_{TP}|$=415mV for the typical cell sizing. It should be noted that DRV is very sensitive to the V_{TN}/V_{TP} ratio increasing by roughly 50% for $|V_{TN}$ offset$|$=110mV. Achieving the minimum DRV should therefore be the main objective of the optimization process.

B. DRV optimization

As demonstrated in Fig.11, the V_{TN} offset from the perfectly balanced cell inverter case is an important metric for DRV optimization. Since the V_{TN}-$|V_{TP}|$ value for which the minimum DRV is obtained is approximately constant as depicted in Fig. 6, it would be interesting to be able to analytically evaluate this value. Starting with a typical subthreshold current equation and assuming NMOS and PMOS strength equality, the following equation for V_{TN} can be obtained for $V_{GS}=V_{DD}/2$:

$$V_{TN} = V_{DD}/2\left((1+\eta_n) - \frac{S_n}{S_p}(1+\eta_p)\right) + \frac{S_n}{\ln 10}\ln\frac{I_n}{I_p} + \frac{S_n}{S_p}|V_{TP}| \quad (7)$$

where S, η, and I are the subthreshold slope, DIBL and transistor current for $V_T = V_{GS}$, respectively. It can be noted, that V_{TN} has a weak dependence on V_{DD} based on the ratio of DIBL parameters and subthreshold slopes. The V_{TN} value obtained with Eq.7 shows a good correlation with the ELDO result with an error of less than 0.5%.

C. Body biasing

As presented in the previous section, in order to improve the stability in active modes while at the same time ensuring the lowest possible DRV, two different V_{TN}/V_{TP} ratios are required. These ratios can be obtained through body biasing (BB), in particular for 32-nm UTBB-FDSOI, where the body factor is as high as 60-70mV/V. The range of V_T modification is limited by the maximum positive voltage, which can be applied to the back terminal. Looking at Fig.1 it can be noted, that special care has to be taken not to forward bias the PN junction between the P-Substrate and the N-Well. As there is no real limitation to the applied reverse body bias, with the exception of the additional complexity due to the need to generate a high negative voltage, the biasing range can be assumed in the interval ($-V_{DDNOM}$,$+0.5V_{DDARRAY}$) for the NMOS and ($0.5V_{DDARRAY}$,$2V_{DDNOM}$) for the PMOS, where V_{DDNOM} corresponds to the nominal V_{DD} and $V_{DDARRAY}$ to the SRAM cell (array) supply voltage. Under this assumption, the total possible magnitude of V_T adjustment is in the range of approximately 100mV, depending on the V_{DD} value. Naturally, the body coefficient for bulk technologies is much lower, significantly limiting the possibility of V_T modification through body biasing.

D. Stability-oriented optimization

In order to illustrate the optimization process, let us consider two different cases: (1) high V_T transistors in the UTBB-FDSOI process with $|V_{TP}|$=542mV and V_{TN}=444mV and (2) the opposite V_T setup with V_{TN}=542mV and $|V_{TP}|$=444mV. The SRAM cell has the default sizing of 63n/38n, 198n/38n, 153n/45n (load, driver, access). Furthermore, the assumption is that we want to obtain stable operation for V_{DD}=0.4V. In order to minimize the DRV the $\Delta V_{THOLD}=V_{TN}-|V_{TP}|$ should be at approx. 36mV for both cases (Eq. 7). Knowing the magnitude of the UTBB-FDSOI body factor and the applicable BB values given in the previous section, it can be noted that for (1) the top achievable $\Delta V_{THOLD}\simeq$-10mV, whereas for (2) the $\Delta V_{THOLD}\simeq$36mV is obtainable. Assuming that different BB values can be used for standby and active modes, the following optimization steps assume initial "unbiased" V_T values. As depicted in Figs.5-7, for case (2) a high initial read stability can be expected, whereas for (1) it can be unacceptably low. A MC analysis leads to the following read statistical parameters: μ/σ=5.37 for (1), and μ/σ=10.65 for (2). Since the μ/σ for (1) is <6, it has to be boosted and the most efficient way to do this is through body biasing. Increasing V_{TN} by 20mV, which corresponds to roughly -300mV for NMOS BB, leads to an increase of μ/σ to 6, reaching the required stability level. As a final step, the write μ/σ is evaluated at 4.44 (vs. 4.46 from ELDO simulation;

0.5% error) and an assist technique of under-driving the bitline by 47mV is applied to increase this value to 6, thus finalizing the optimization process. Since the write assist technique biases the access NMOS at the edge of subthreshold, as described in section III c, an underestimation of write stability can be expected. Indeed, in the ELDO simulation based on the UTBB-FDSOI compact model, the target write stability is met for 7mV less applied to the bitline. This error is however very predictable and does not affect the whole optimization process.

In case (2) the initial read μ/σ is much larger than 6. The goal of the optimization for such a case is to bring down the read μ/σ to the minimum required value of 6, thus increasing the write stability and reducing the need for aggressive assist technique application. Therefore, since it will be clearly difficult to obtain write μ/σ>6, either the cell size can be modified or the body bias impact can be considered. Starting with the latter it can be assumed that a full-range reverse and forward bias should be applied to PMOS and NMOS, respectively, giving V_{TN}=525mV and $|V_{TP}|$=514mV. For such a case, the read μ/σ is equal to 8.15, meaning the size optimization needs to be applied as well. The transistor sizes are modified as in case (3) of section IV c, leading to a read μ/σ=7. Since there is some read stability overhead, all write-assist techniques are applicable in some range. However, for simplicity only bitline under-driving is used. The application of this technique by -60mV allows increasing the write μ/σ to 6 in ELDO and to 5.47 in the equation-based evaluations, showing a 9% error in the latter. The stability target with the equation-based approach is met for the bitline voltage of -74mV.

Should body biasing be unavailable or its influence be very limited, the optimization would be simplified to choosing the proper initial V_{TN}/V_{TP} ratio to ensure low DRV, resizing the cell to optimize write, and compensate between read and write stability through assist techniques. The obtained optimized cell setups were tested using UTBB-FDSOI compact models in ELDO. The results are summarized in Table I. Clearly, the optimized cases present a much better DRV and significantly lower leakages, while maintaining sufficient stability margins.

For the purpose of this analysis, the proposed equation set was implemented in MATLAB; however, another similar software or a standalone program can be used as well.

TABLE I SUMMARY OF OPTIMIZATION RESULTS

	Initial (1)	Optimized (1)	Initial (2)	Optimized (2)
DRV	334mV	212mV	197mV	179mV
I_{LEAK} at DRV	1.2pA	116fA	247fA	65fA
I_{LEAK} in active	1.3pA	628fA	328fA	96fA
Read μ/σ	5.3	6	9.5	7
Write μ/σ	5.3	6	3.54	6

VI. CONCLUSIONS

In this paper, we presented a methodology for the optimal design of CMOS 6T SRAM ultra-low-power (ULP) bitcells minimizing power consumption under strict stability constraints in all operating modes. The presented equation set allows an accurate analysis of SRAM stability in all operation modes. Since all cell node voltages are included in the equations, an assessment of the impact of assist techniques (in the accuracy range) is possible as well. Based on the given equation set, an optimization algorithm is described. As demonstrated in section V d and Table I, through the application of appropriate optimization techniques, the standby and active leakages can be reduced in case (1) as much as 10x and 2x,respectively. In the second analyzed case, the initial write stability was very low; application of the optimization process allowed achieving a sufficiently high μ/σ while lowering active and standby leakage by 3x and 4x, respectively. Application of the proposed methodology allowed operation at a supply voltage as low as 0.4V while maintaining sufficiently high yield. Moreover, the necessity of write assist techniques was minimized requiring under-driving the bitline by only 40mV for case (1) and 60mV for case (2). The presented methodology is easily implemented in MATLAB, another similar tool or as a standalone program, using the presented equations. The error on statistical WSNM estimation is not a major issue either, since it is very predictable and it always leads to an underestimation of the result. Therefore, the application of the write assist technique in the range given by the equation will lead to obtaining a higher than necessary stability only, in the worst case. The presented results show an excellent efficiency of our optimization process for ULP design, as well as it highlights the attractiveness of UTBB-FDSOI due to its high body factor and low process variability. Should a technology be used where the influence of body bias is very limited, a similar approach can be applied, with the exception of BB step.

References:

[1] Hulfang Qin et al. "SRAM leakage suppression by minimizing standby supply voltage", 5th International Symposium on Quality Electronic Design, 2004.

[2] Calhoun, B.H. et al., "Static noise margin variation for sub-threshold SRAM in 65-nm CMOS" , IEEE Journal of Solid-State Circuits, Volume: 41, 2006

[3] J. Wang et al., "Statistical modeling for the minimum standby supply voltage of a full SRAM array," in Proc. Eur. Solid State Circuits Conf., 2007, pp. 400–403.

[4] J. Wang et al., "Two Fast Methods for Estimating the Minimum Standby Supply Voltage for Large SRAMs", IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol 29, 12, December 2010

[5] J. Wang et al., "Minimum Supply Voltage and Yield Estimation for Large SRAMs Under Parametric Variations", IEEE Transactions on Very Large Scale Integration (VLSI) Systems 2010

[6] E. Seevinck et al., "Static-Noise Margin Analysis of MOS SRAM Cells" IEEE Journal of Solid-State Circuits, VOL. SC-22, NO. 5, OCTOBER 1987

[7] J-P. Noel, et al., "UT2B-FDSOI Device Architecture Dedicated to Low Power Design Techniques," ESSDERC 2010

[8] J. Mazurier, et al., "On the variability in Planar FDSOI technology: from MOSFETs to SRAM cells", IEEE Transactions on Electron Devices 2011

[9] A. Makosiej, et al., "An SNM Estimation and Optimization Model forULP sub-45nm CMOS SRAM in the Presence of Variability," NEWCAS 2010