

Subthreshold SCL for Ultra-Low-Power SRAM

DIPLOMA THESIS TECHNICAL UNIVERSITY OF CRETE Department of Electronics and Computer Engineering

For the acquisition of the Diploma Degree in Electrical Engineering , sector of
Microelectronics

by

ILIAS CHLIS

The board of Examiners:

Assistant Professor **Matthias Bucher** (Thesis Director)

Professor **Apostolos Dollas**

Professor **Kostas Kalaitzakis**

HELLENIC

REPUBLIC



TECHNICAL UNIVERSITY OF CRETE
DEPARTMENT OF ELECTRONICS AND COMPUTER ENGINEERING

Greece
Spring 2010

ABSTRACT

This thesis begins by presenting a theoretical analysis of the operation of the Source-coupled Logic (SCL) or MOS current-mode logic (MCML) circuits for implementing mixed-mode circuits. Subsequently, an SCL-based static random-access memory operating in Weak Inversion (subthreshold) region is implemented to demonstrate the performance of this topology for ultra-low-power consumption and low-activity-rate mixed-signal circuits. The SRAM cells are using PMOS drain-bulk connected transistors as loads, the operation of which is demonstrated through measurements performed on PMOS transistors of various geometries.

ACKNOWLEDGEMENTS

First of all , I owe many thanks to the team of graduate students of the Microelectronics lab of the Technical University of Crete , namely to Agelos Antonopoulos , George Kargakis and Anna-Maria Chalkiadaki who contributed to this work providing valuable help both in the theoretic field and with the tools used for measurements and simulation.

Also , I wish to express my gratitude to Assistant Professor Matthias Bucher for accepting me in his research team , and for his guidance in the field of the circuit design and modeling using the MOS transistor , his knowledge for which is truly profound.

Finally , my deepest appreciation goes to Professor Apostolos Dollas for his inspiring teaching of VLSI circuits , which prompted me to follow this domain at a professional level.

CONTENTS

<u>Chapters</u>	<u>Topics</u>	<u>Pages</u>
1	Introduction	5
2	Source-Coupled logic	6
2.1	General Structure	6
2.2	Performance	11
2.3	Implementation of the current source	16
2.4	Implementation of the load resistances	17
2.5	Implementation of the replica bias	23
2.6	Implementation of the SCL inverter	31
3	Measurements	38
3.1	Measurements on the Drain-Bulk Connected PMOS used as loads	38
4	Implementation of the SCL SRAM	51
4.1	About SRAMs	51
4.2	The 6T SRAM cell	52
4.3	Peripheral circuitry	55
4.4	Low power SRAM design	59
4.5	An alternative SRAM cell	62
4.6	Simulation results	64
4.7	Conclusion	84
REFERENCES		84

1. INTRODUCTION

During the last two to three decades, the once emerging trend towards portable devices that would dissipate less power for active and static operation, has now become the foundation stone of today's research. Mobile phones, laptops, notebook computers, netbooks, tablet PCs and video cameras are only a small part of the large spectrum of applications where power dissipation is the primary design issue. Due to the fact that the energy density of batteries has increased at a relatively low pace, more and more features have to be supported by at most, the same limited amount of power budget and in many cases less than that. Therefore, the circuit designer faces the nontrivial task of reducing power consumption, if possible without degradation of other characteristics like speed and robustness.

The MOS transistor, with its tremendous improvements in downscaling and hence in functionality and speed performance, constitutes the main driving force of modern consumer electronics. Operation in the regions of Weak (Subthreshold) and Moderate Inversion that, with a few brilliant exceptions, were previously ignored or even unknown, are now becoming the very desired regions of operation for the MOS transistor, more importantly because they offer minimum possible operating voltage and thereby minimum P_{dyn} for a given P_{stat} [1]. Other extremely important attributes of those regions include higher gate (g_{mg}), drain (g_{md}), source (g_{ms}) and bulk (g_{mb}) transconductances and higher DC voltage gain ($g_{\text{mg}}/g_{\text{md}}$) [2].

Nowadays, most integrated circuits digital, analog or mixed-signal, are built using CMOS technology, which, for the past few decades has been continuously scaled down, providing the implementation of more complex circuits and systems. Technology scaling, however, has made some of the secondary non-ideality effects in CMOS devices more pronounced, particularly leakage currents, which have been increased considerably, therefore reducing the power efficiency of deep-sub-micron CMOS technologies [3]. Additionally, with conventional CMOS logic, the reduction in the power consumption comes at the expense of the speed performance, so that the operation is tightly related to the supply voltage, which has an impact not only on the speed and power, but also on the noise performance [4].

On the other hand, in Subthreshold Source-coupled Logic (STSC) circuits, there is the possibility of reducing the bias current below the subthreshold leakage current of CMOS circuits, thus making the power-delay performance of this type of circuits comparable to the CMOS ones [3]. Also, in STSC the output swing, propagation delay and hence, the performance, are mostly independent of the supply voltage [4]. Furthermore, compared to the conventional CMOS circuits with rail-to-rail voltage swing and currents on the order of nanoamperes, STSC circuits with reduced voltage swing and currents on the order of picoamperes exhibit a more efficient dynamic power consumption. Finally, SCL circuits in general, present features that make them suitable for a mixed-signal environment, including immunity to supply noise due to differential structures, low crosstalk due to small output-voltage swings, and low generated supply noise due to constant current flowing across the supply rails [5].

2. SOURCE-COUPLED LOGIC

I. GENERAL STRUCTURE

An SCL gate consists of a NMOS Network of one or more Source-Coupled Pairs , two loads that can be active or passive resistors and one constant current source at the common connection node of the NMOS Network , as shown in figure 1. The output load resistance R , converts the bias current I_{ss} shown in figure 2 , back to the voltage domain in order to drive the subsequent SCL gates.

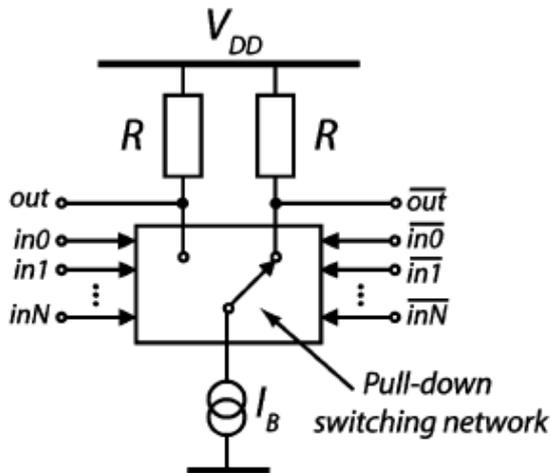


Figure 1 – Picture from [5]

The general topology of figure 1 can be used to implement both combinational and sequential gates , through a proper connection of the source coupled-pairs[6]. The simplest SCL gate is the inverter seen in figure 2.

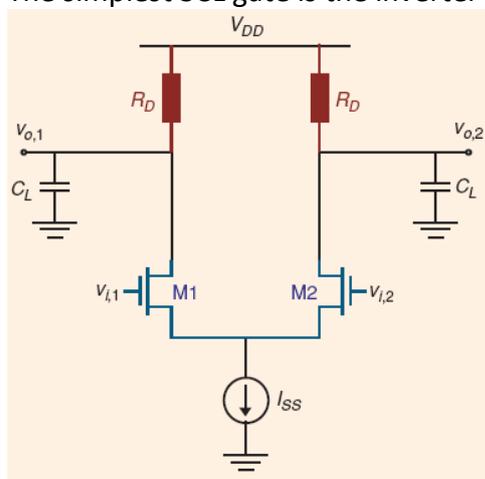


Figure 2 - Picture from [6]

In the SCL inverter , the NMOS network consists of only one Source-coupled pair. We consider v_i as the input voltage , where $v_i = v_{i1} - v_{i2}$ and v_o as the output voltage , where $v_o = v_{o1} - v_{o2}$. The bias current I_{ss} is steered by transistors $M_1 - M_2$, to one of the two resistors , according to the value of the input differential voltage v_i :

$$\begin{array}{l}
 \text{when } v_i \text{ is high} \left\{ \begin{array}{l} v_{o1} = V_{DD} - R \cdot I_{ss} \quad , \text{ M1 is on (} v_{i1} \text{ is high)} \\ v_{o2} = V_{DD} \quad \quad \quad \quad , \text{ M2 is off (} v_{i2} \text{ is low)} \end{array} \right\} \Rightarrow v_o = - R \cdot I_{ss} \\
 \text{when } v_i \text{ is low} \left\{ \begin{array}{l} v_{o1} = V_{DD} \quad \quad \quad \quad , \text{ M1 is off (} v_{i1} \text{ is low)} \\ v_{o2} = V_{DD} - R \cdot I_{ss} \quad , \text{ M2 is on (} v_{i2} \text{ is high)} \end{array} \right\} \Rightarrow v_o = + R \cdot I_{ss}
 \end{array}$$

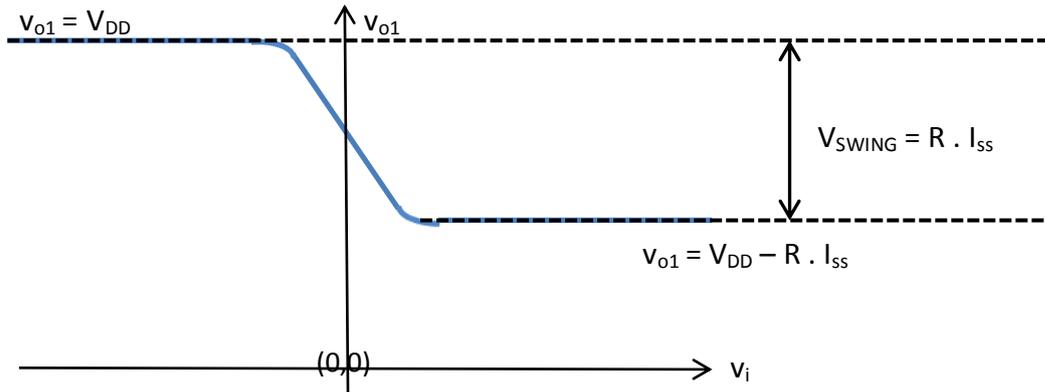


Figure 3

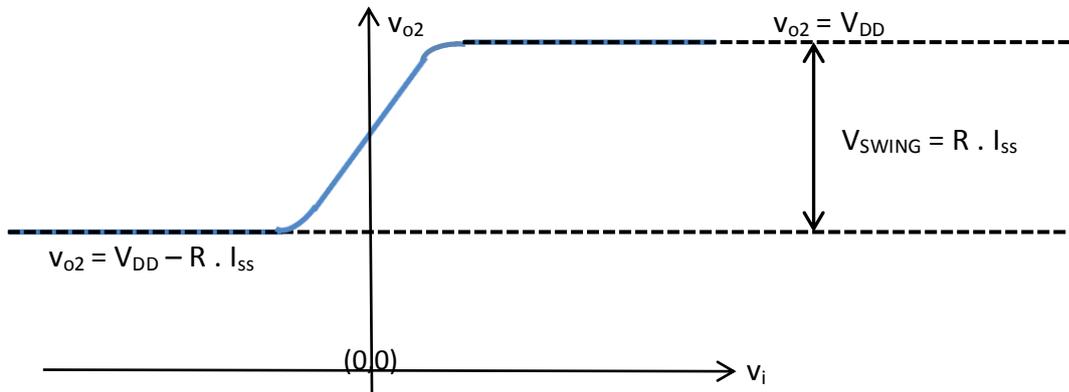


Figure 4

This behavior is shown in figures 3 and 4. Hence , the logic swing of $v_o = v_{o1} - v_{o2}$ is $V_{SWING} = R \cdot I_{SS}$, as shown in figure 5. Compared to the CMOS logic where the signal swing is equal to V_{DD} , in SCL the current needed for charging and discharging the parasitic capacitances is less[3].

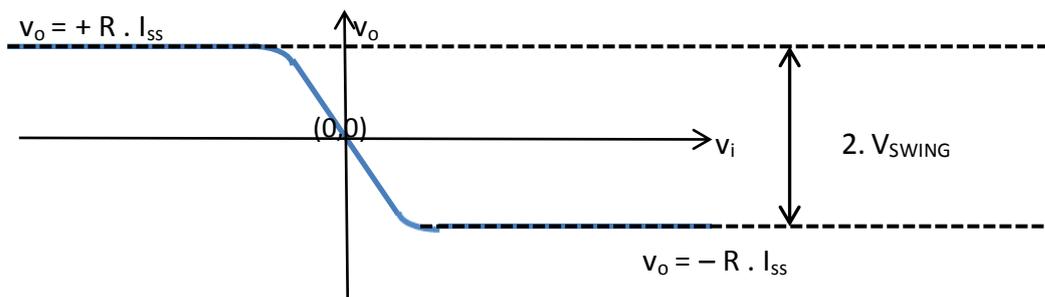


Figure 5

The voltage swing of the difference of the voltage of the output nodes v_o (V_{SWING}) should be high enough to switch completely the input differential pair of the next stage. Equivalently , the gain of each SCL circuit should be high enough to be used as a logic circuit with acceptable noise margin[3]. Thus , V_{SWING} should be larger than $\sqrt{2} \cdot V_{ov} = \sqrt{2}(V_{GS} - V_t) = \sqrt{2n} \cdot V_{DSsat}$ when the input transistors M1 and M2 are in Strong Inversion[8]-[9] as shown in figure 6 , and larger than $4nU_T$ when the input

transistors M1 and M2 are in Weak Inversion[8]-[10]-[17] as shown in figure 7 , where $U_T=KT/q$ is the thermodynamic voltage ($U_T=25.8\text{mV}$ at 300K) and n is the slope factor. Therefore:

$$V_{\text{SWING}}=R \cdot I_{\text{SS}} = \begin{cases} \sqrt{2}n \cdot V_{\text{DSsat}} & , \text{ when M1 , M2 in Strong Inversion} \\ 4nU_T & , \text{ when M1 , M2 in Weak Inversion} \end{cases}$$

and

$$2 \cdot V_{\text{SWINGmin}} = \begin{cases} 2\sqrt{2}n \cdot V_{\text{DSsat}} & , \text{ when M1 , M2 in Strong Inversion} \\ 8nU_T & , \text{ when M1 , M2 in Weak Inversion} \end{cases}$$

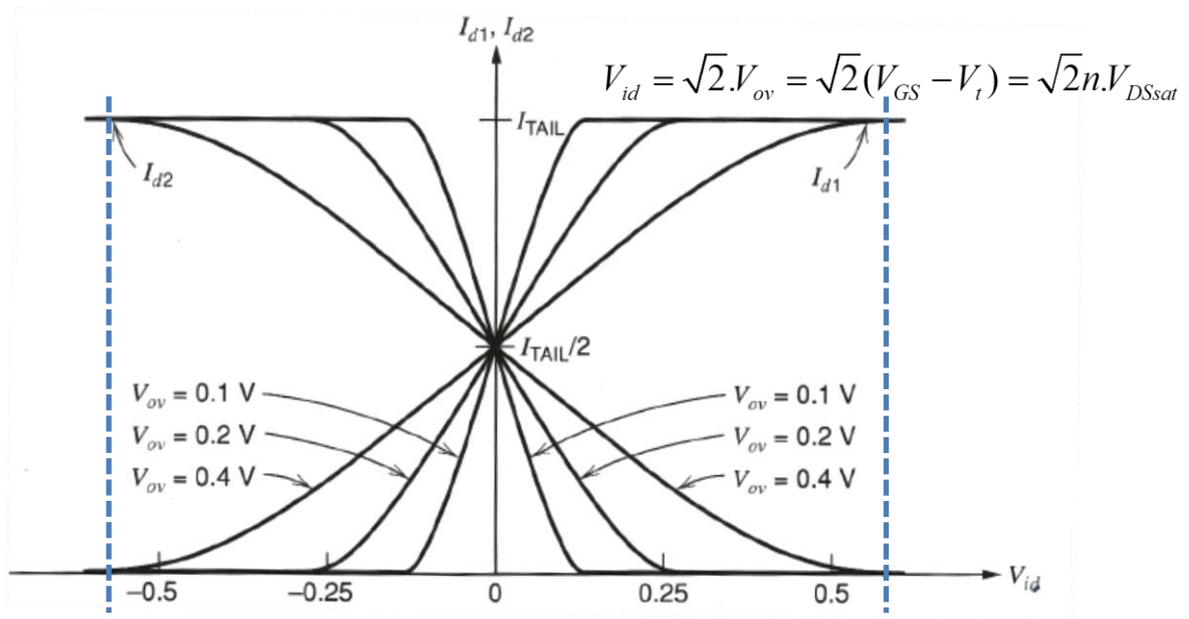


Figure 6 – Picture from [9]

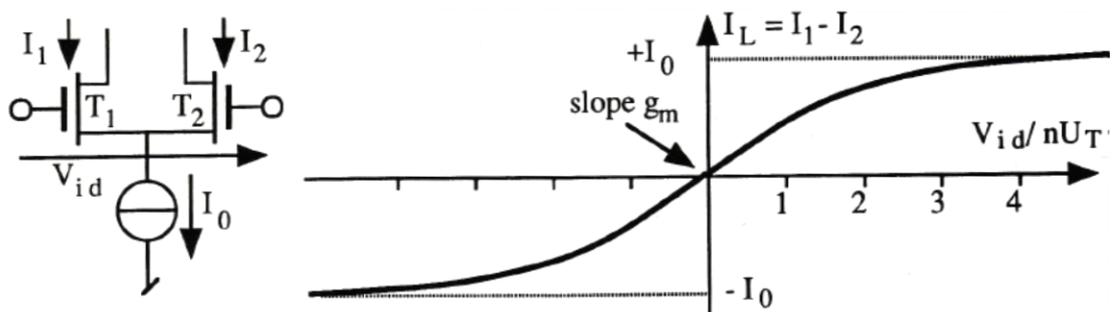


Figure 7 – Picture from [17]

The voltage drop of $R \cdot I_{\text{SS}}$ is in the order of a few hundreds of mV (when the input NMOS transistors are in the subthreshold regime can be as low as 150mV at 300K assuming $n=1.5$) and I_{SS} can range from a few pA to a few nA in ultra-low power applications. Hence , the load resistance R , is in the order of hundreds of M Ω or a few G Ω , which can only be implemented through the use of an active element[7]-[8] , as passive resistive elements mainly achieved by diffused layers(n^+ , p^+ and well) or by polysilicon (first poly or , when available , second poly) can reach resistance

values only up to a few $k\Omega$ [11]. Furthermore, passive resistors suffer a great deal of process variations due to doping, lithography and etching[12].

The logic function is implemented by the NMOS switching network in such a way that, for each input combination, the current is allowed to flow in only one of the two branches[5]. The SCL inverter's NMOS network is built taking as a differential input the difference of the input of M1 minus the negation of the input of M2:

$$v_i = v_{i1} - \overline{v_{i2}}$$

, as in figure 8:

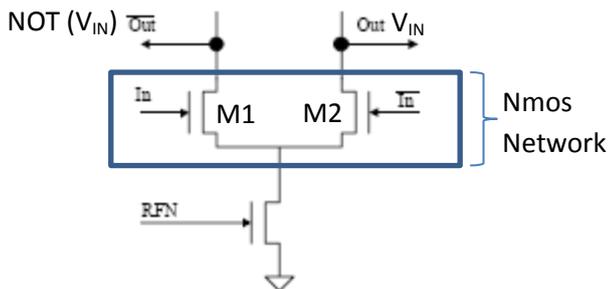


Figure 8 – Picture from [13]

In the same way, we can build a more complex gate. For example, the NMOS network of a NAND gate $F = \overline{A \cdot B}$ can be built by taking the following differential input:

$$v_i = A_1 \cdot B_1 - \overline{A_2 \cdot B_2}$$

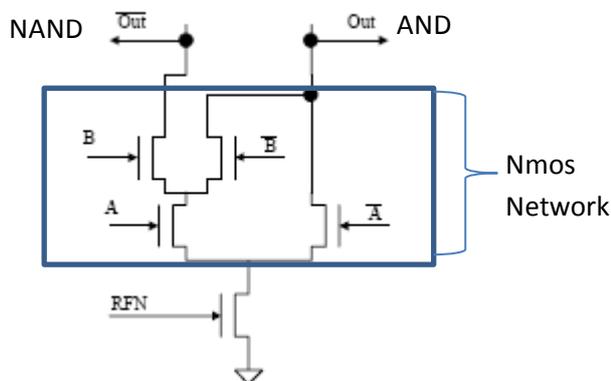


Figure 9 – Picture from [13]

In general, an arbitrary combinational function can be built by implementing an NMOS network having all transistor paths associated with the 2^n possible inputs and then properly connecting each of the upper drain branch to $\overline{v_o}$ or v_o to set the output to the desired value. Each of the 2^n possible input values is associated with the unique product $X_1 \cdot X_2 \dots X_n$ being equal to 1, in which each variable is complemented if the correspondent input bit is 0 (for example, input 0110 is associated with product $\overline{X_1} \cdot X_2 \cdot X_3 \cdot \overline{X_4}$). The product $X_1 \cdot X_2 \dots X_n$ is in turn associated with a unique active path consisting of transistors driven by the correspondent variables. Therefore, an unambiguous correspondence between input values and stacked transistors' paths exists.

Such an NMOS network has a tree-like structure with a source coupled pair connected to each drain node of transistors lying at the lower level, thereby doubling the number of transistors in a logic level compared to the lower level[53]. An example is shown in figure 9b for an arbitrary 3-variable function $F(X_1, X_2, X_3)$.

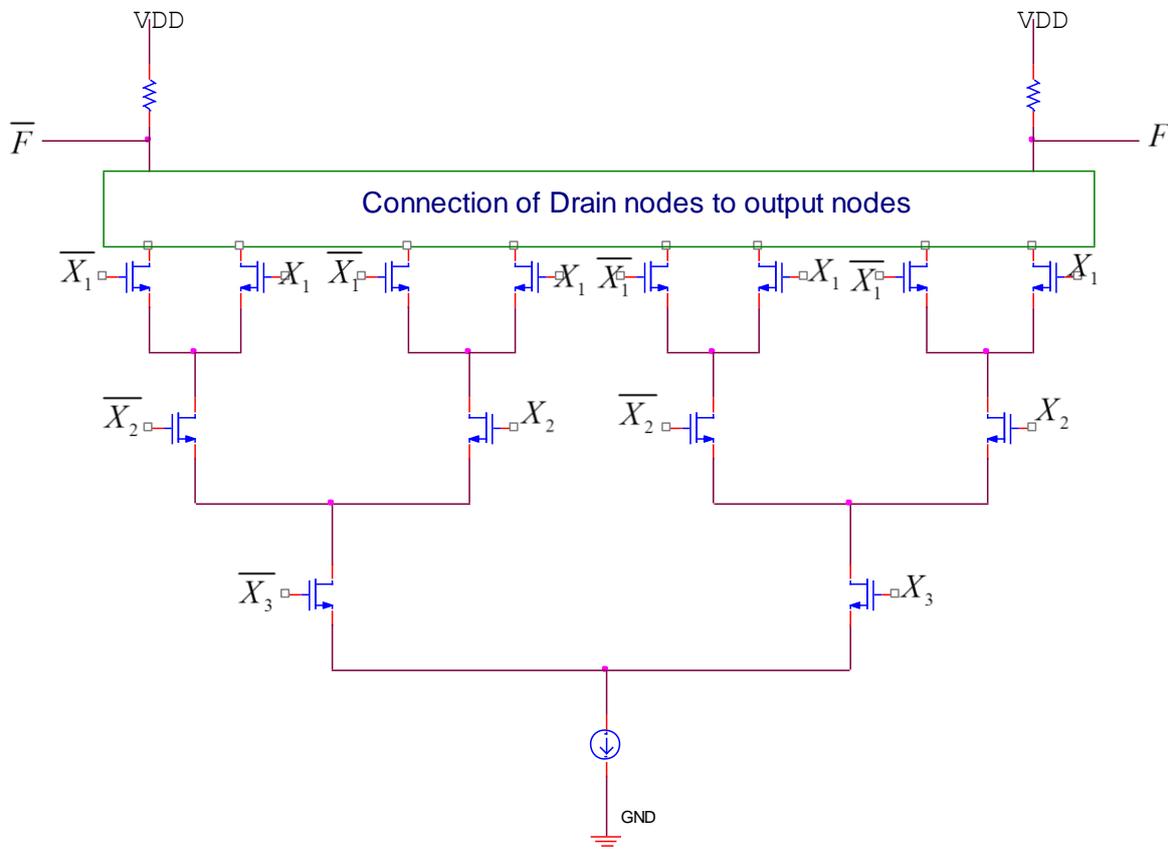


Figure 9b – General topology implementing an arbitrary 3-variable function

To map the value of the Boolean function F to the series-gate output voltage v_o , a proper choice of connections of drain nodes of the highest transistors to one of the two output nodes is needed.

For example, if we want to implement a 3-variable function given a truth table, we must first express this function as a sum of products through a Karnaugh map. Let's say that after expressing the function as a sum of products we have that:

$$F = X_1 \cdot X_2 \cdot X_3 + X_1 \cdot X_2 \cdot \overline{X_3} + X_1 \cdot \overline{X_2} \cdot X_3$$

Then, the SCL gate implementing F is the one in figure 9c. A drawback of this topology is the use of series – connected transistors, called series gating, which results to additional voltage drops across the NMOS transistors causing a small increase in supply voltage and also a reduction in voltage gain, due to the fact that some transistors move into the resistive region[54].

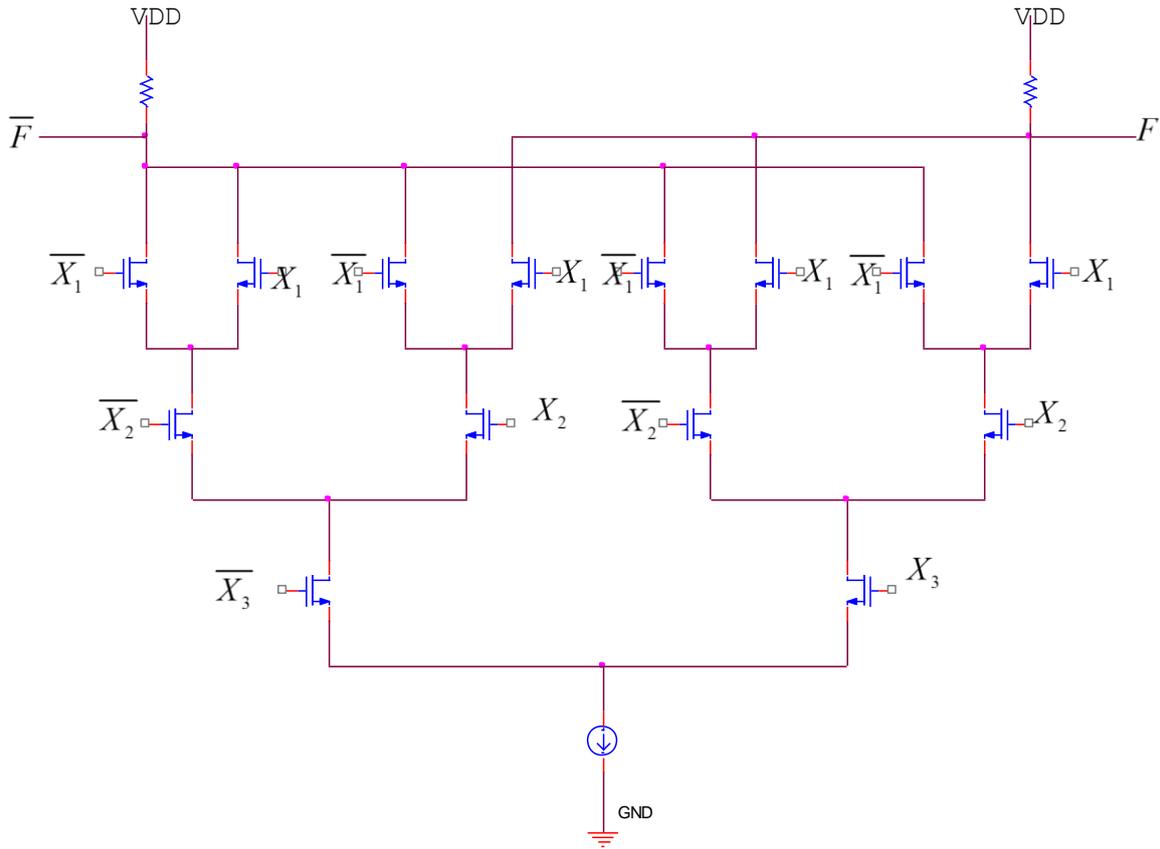


Figure 9c – The SCL gate implementing function F

II. PERFORMANCE

The total propagation delay , the Power consumption , the Power-Delay product , the Energy consumption and the Energy Delay product of a chain of N identical SCL gates , all having capacitance C , respectively is [8]-[13]-[15] :

$$D_{SCL} = N \cdot \ln 2 \cdot R \cdot C = N \cdot \frac{\ln 2 \cdot C \cdot V_{SWING}}{I_{SS}}$$

$$P_{SCL} = N \cdot I_{SS} \cdot V_{DD}$$

$$PD_{SCL} = P_{SCL} \cdot D_{SCL} = N \cdot I_{SS} \cdot V_{DD} \cdot N \cdot \frac{\ln 2 \cdot C \cdot V_{SWING}}{I_{SS}} = N^2 \cdot \ln 2 \cdot C \cdot V_{SWING} \cdot V_{DD}$$

$$E_{SCL} = N \cdot C \cdot V_{DD} \cdot V_{SWING}$$

$$ED_{SCL} = E_{SCL} \cdot D_{SCL} = N^2 \cdot C \cdot V_{SWING} \cdot V_{DD} \cdot \frac{N \cdot C \cdot V_{SWING}}{I_{SS}} = \frac{N^3 \cdot C^2 \cdot V_{DD} \cdot V_{SWING}^2}{I_{SS}}$$

As a result , as opposed to the CMOS logic gates , the current draw of SCL gates is constant over time and independent of switching activity[13]. That is , the power consumption P_{SCL} of each SCL gate depends only on I_{SS} and not on the operating frequency($1/D_{SCL}$). Moreover , since the power consumption P_{SCL} and the total propagation delay D_{SCL} of each SCL gate only depend on I_{SS} which can be controlled very precisely , SCL logic gates exhibit very low sensitivity to the process variations[16].

In comparison , the total propagation delay , the Power consumption , the Power-Delay product , the Energy consumption and the Energy Delay product of a chain of N identical CMOS gates , all having capacitance C , respectively is [13]-[14]-[15]:

$$D_{CMOS} = \frac{N.C.V_{DD}}{\frac{K}{2} \cdot (V_{DD} - V_t)^\alpha}$$

$$P_{CMOS} = N.C.V_{DD}^2 \cdot \frac{1}{D_{CMOS}}$$

$$PD_{CMOS} = P_{CMOS} \cdot D_{CMOS} = N.C.V_{DD}^2$$

$$E_{CMOS} = C \cdot V_{DD}^2$$

$$ED_{CMOS} = E_{CMOS} \cdot D_{CMOS} = N^2 \cdot 2 \cdot \frac{C^2}{k} \cdot \frac{V_{DD}^2}{(V_{DD} - V_t)^\alpha}$$

Where $K = \mu \cdot C'_{ox} \cdot \frac{W}{L}$ and $\alpha \approx 1.3$

By comparison of the SCL and CMOS circuits Energy-Delay product , we observe that in contrary with the CMOS circuits , SCL circuits do not have a theoretical minimum to the energy-delay product. Therefore , the ED_{SCL} can be arbitrarily reduced by increasing the current I_{SS} for a given C , V_{DD} and V_{SWING} . However , in practice this cannot be achieved for very large currents because of a possible deterioration of the robustness of the circuitry[13].

Another important observation is that unlike conventional digital CMOS circuits where neglecting the leakage current there is no static power consumption , in SCL topology each cell consumes significant static power due to the bias current I_{SS} which exists even when not switching. Taking this under consideration , as the activity(or duty) rate reduces , the power efficiency of SCL topology in comparison with the CMOS one degrades quickly[3].

Despite the fact that the previous argument is correct in older technologies where the static power consumption of the CMOS logic circuits is negligible , it does not hold in the newer ultra-deep sub-micron (UDSM) technologies , where the static power consumption of the CMOS logic circuits due to leakage currents , is considerable. Thus , in these technologies , even in low activity rates , SCL circuits can exhibit better power-delay performance in comparison to the conventional CMOS topology[3].

Additionally , one of the main advantages of the SCL topology is the possibility of reducing the signal swing V_{SWING} , hence reducing the current needed for charging

and discharging the parasitic capacitances compared to the CMOS topology where the signal swing is equal to V_{DD} [3].

Finally, the differential operation of SCL circuits can improve immunity from noise such as substrate noise. The digital switching noise traveling through power lines such as supply bounce can also be reduced because the signal voltage swing V_{SWING} and the drive current I_{SS} of SCL circuits are constant under the influence of supply bounce. As a result, the jitter of SCL oscillators is about 65% that of CMOS oscillators[15].

Using the previously mentioned relations, it can be shown that the total power consumption of a chain of N identical SCL gates is:

$$P_{diss,SCL,N} \approx \ln 2 \cdot N^2 \cdot V_{DD,SCL} \cdot V_{SWING} \cdot C \cdot f_{op}$$

which is increasing quadratically with the logic depth and linearly with the operating frequency. Thus, the device parameters and especially the threshold voltage V_t do not influence the speed-power consumption tradeoff in the SCL topology[16].

Correspondingly, including leakage current, the total power consumption of an array of N identical conventional CMOS gates is[16]:

$$P_{diss,CMOS,N} = V_{DD} \cdot \sqrt{\frac{1}{T} \int_0^T i_{DD}^2(t) dt} \Rightarrow$$

$$\Rightarrow P_{diss,CMOS,N} \approx N \cdot I_{leak} V_{DD} \sqrt{1 + \frac{\alpha \cdot \eta}{3} \left(\frac{\gamma^2}{N^2} + \frac{\gamma}{N} - 2 \right)}$$

Where $a = \frac{f_{op}}{f_{max}}$ is the activity(or duty) rate,

$f_{max} = \frac{1}{2 \cdot t_d}$ is the maximum operation frequency of a single gate,

$$\gamma = \frac{I_{peak}}{I_{leak}}, \quad f_{op} = \frac{1}{T} \quad \text{and} \quad \eta = \begin{cases} N/2 & , \text{ if } N \text{ even} \\ (N+1)/2 & , \text{ if } N \text{ odd} \end{cases}$$

The peak current (I_{peak}) and leakage current (I_{leak}) drawn from supply by the logic cell, both depend on V_{DD} and the size ration of devices. Also, I_{peak} depends on the transition time at the input of the gate as shown in figure 10[16].

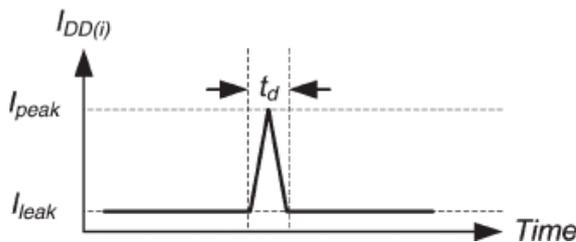


Figure 10 – Picture from [16]

Based on the equation for the power dissipation of CMOS gates, it can be found that for activity rates smaller than a critical activity rate a_c the subthreshold leakage power consumption will be dominant, while for higher activity rates, the dynamic

power consumption comprises the main part of the power consumption[16]. Also , as seen in figure 11 , for activity rates larger than a_c , the power dissipation increases proportionally to \sqrt{a} in a constant V_{DD} . However , by reducing the activity rate , the power consumption will be dominated by the leakage current:

$$P_{diss,CMOS} \approx V_{DD} \cdot \sqrt{I_{leak}^2 + \gamma \cdot \alpha} \Rightarrow P_{diss,CMOS|\alpha \rightarrow 0} \approx V_{DD} \cdot I_{leak} \quad [30]$$

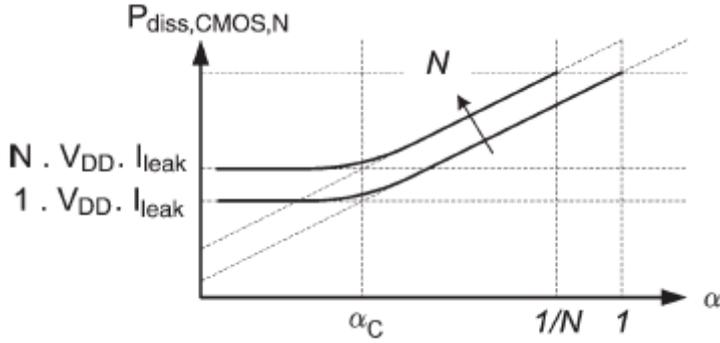


Figure 11 – Picture from [16]

According to [16] , a_c is proportional to $\frac{1}{\gamma^2}$ and therefore increases quadratically with reducing γ . This means that in more advanced CMOS technologies where leakage currents are more pronounced , the contribution of I_{leak} to the total power consumption will be more evident and a_c will be higher[16]. This is shown in figure 12.

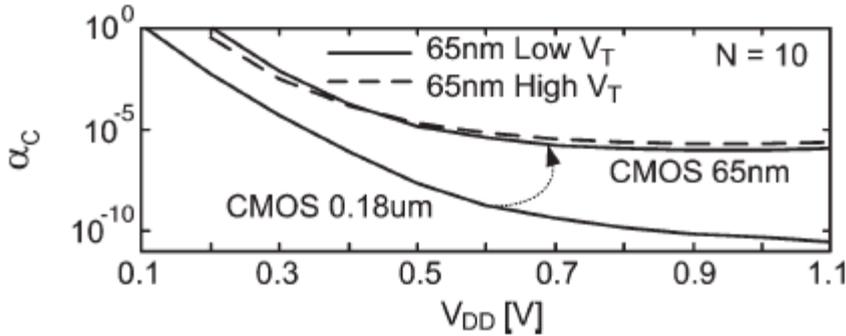


Figure 12 – Picture from [16]

In conclusion , while conventional CMOS topology shows a very good power efficiency for a very wide range of applications and activity rates due to its negligible static power consumption as long as leakage is not dominant , for nanometer-scale CMOS technologies however , where the off-state (subthreshold) leakage of each transistor can reach nA-levels , the SCL topology with its controllable tail bias current I_{SS} can offer reduced power consumption well below the subthreshold leakage of CMOS , while maintaining a significant advantage over CMOS topologies[30].

The range of frequencies over which a chain of SCL gates offers a better power efficiency over a chain of CMOS gates is given in figure 13. Both chains are loaded with the same capacitance and both operate in the subthreshold regime.

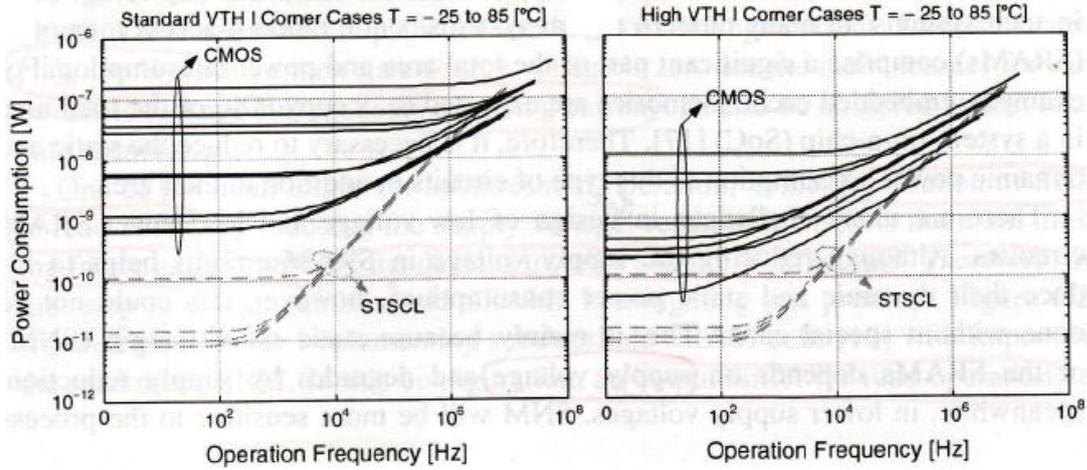


Figure 13 – Picture from [3]

As seen in figure 13 , the overall dissipation of the CMOS chain at very low operating frequencies is limited by the leakage current which can be reduced by lowering the supply voltage V_{DD} , yet a dramatic reduction is not possible because the operational robustness diminishes as the current-drive capability of CMOS gates drops exponentially with the supply voltage[3].

Also , despite the fact that the leakage power dissipation of CMOS circuits can also be reduced significantly by using HVT transistors , this has a serious impact on their operation speed. SCL gates too can be constructed using HVT transistors so as to control the tail bias current I_{SS} , without this having any detrimental effects on switching speed[3].

On the other hand , the lower limit for SCL-based circuit power consumption is the stand-by current I_{SS} that can be as low as a few pico-Amperes[3]-[8]-[30].

Another important issue also seen in figure 13 , is the very wide variation of leakage and dynamic consumption in CMOS topology which can be as high as two orders of magnitude and which is mainly due to the exponential dependence of the residual channel current I_{leak} in the subthreshold regime , on the device threshold voltage V_t , as seen in the following equation:

$$I_{leak} \approx I_{subth} \approx \mu \cdot C_{ox} \frac{W}{L_e} U_T^2 \cdot e^{-\frac{\Delta V_t}{n \cdot U_T}} \cdot e^{-\frac{-V_{TO} + \bar{\eta} \cdot V_{DD}}{n \cdot U_T}} \quad [3]$$

Where $\bar{\eta}$ is the DIBL coefficient expressed as:

$$\bar{\eta} = \frac{1}{2 \cdot \cosh \frac{L_{eff}}{2 \cdot L_t}} \quad , \quad \text{in which} \quad L_t = \sqrt{\frac{\epsilon_{Si} t_{ox} W_{dm}}{\epsilon_{ox} \cdot K}}$$

is a characteristic length

and K is a fitting parameter.

Thus , leakage current I_{leak} highly depends on temperature (through U_T), on the threshold voltage variation (ΔV_t) and on the supply voltage V_{DD} , due to the DIBL effect modeled by $\bar{\eta}$.

III. IMPLEMENTATION OF THE CURRENT SOURCE

The current source providing the tail (bias) current I_{SS} can be implemented as a common NMOS current mirror :

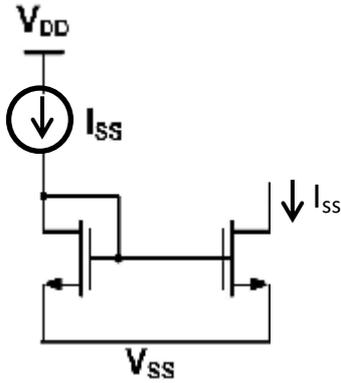


Figure 14

A possible implementation in Pspice simulated using EKV v2.6 indicative data for $0.5\mu\text{m}$ CMOS provided by [18] is shown in figures 15 and 16. M1 and M2 have a W/L ratio equal to $1(0.5\mu\text{m}/0.5\mu\text{m})$.

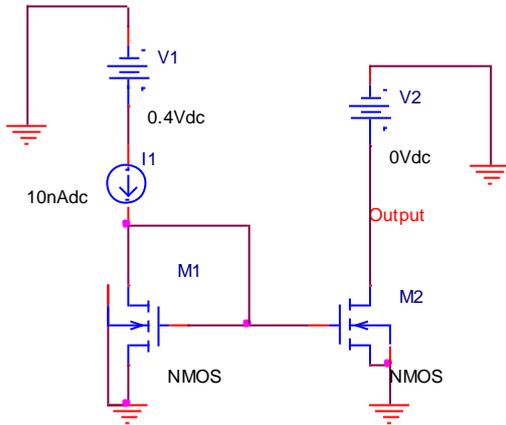


Figure 15

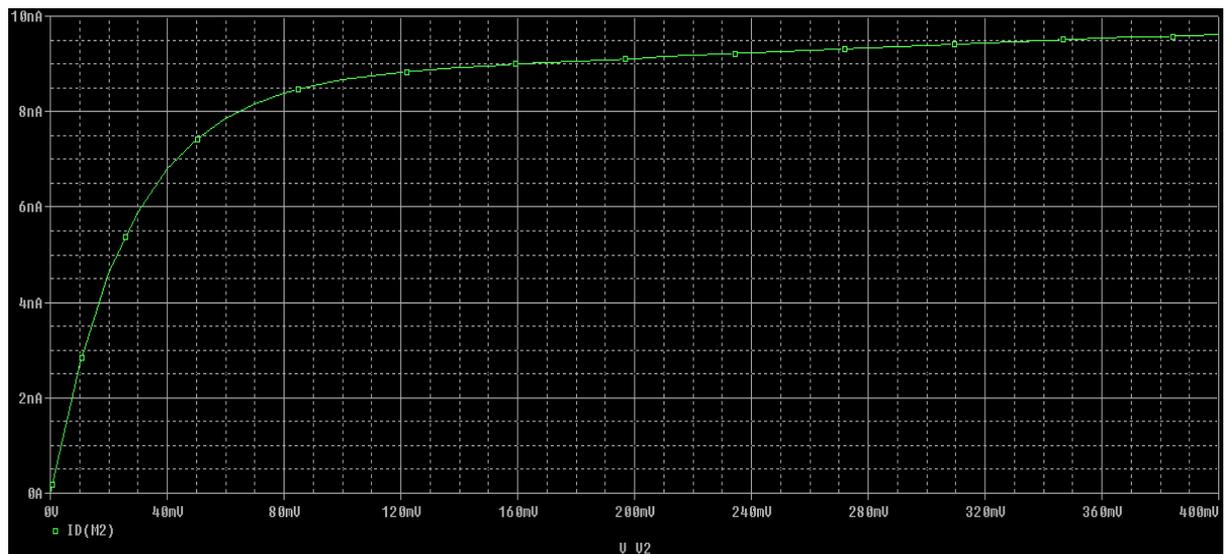


Figure 16

IV. IMPLEMENTATION OF THE LOAD RESISTANCES

Since $V_{SWING} = R \cdot I_{SS}$, the main problem in a low-current SCL circuit is the realization of very large load resistors required for a reasonable output swing, as for example given that $V_{SWING} = 200\text{mV}$ with $I_{SS} = 10\text{nA}$ requires that $R = 20\text{M}\Omega$ [19]. As mentioned before, due to the problems that the implementation of large passive resistors faces, PMOS transistors can be used, that have to be biased in the linear region of their I_{SD} - V_{SD} characteristic.

For conventional PMOS transistors that have their Bulk connected to their Source like the one depicted in figure 18, the linear region of their I_{SD} - V_{SD} characteristic is restricted to the triode region of operation as shown in figures 19a and 19b. According to [22], the limit of this region of operation which is $V_{SD,sat}$ is constant at $4U_T \cong 100\text{mV}$ in Weak Inversion, increases modestly in Moderate Inversion and increases as the square root of drain current or Inversion Coefficient(IC) in Strong Inversion, as illustrated in figures 17 and 19. Figure 17 is referred to NMOS, but can be readily extended to PMOS by changing polarities, whereas in figures 18 and 19 a Pspice simulation has been performed, using a dc parametric sweep of V_{SG} from 150mV to 600mV with a step of 50mV and a dc primary sweep of V_{SD} ranging from 0 to 400mV. The model used for M1 is the EKV v2.6 provided by [18], and the aspect ratio used for M1 is $W/L = 0.5\mu\text{m}/0.5\mu\text{m} = 1$.

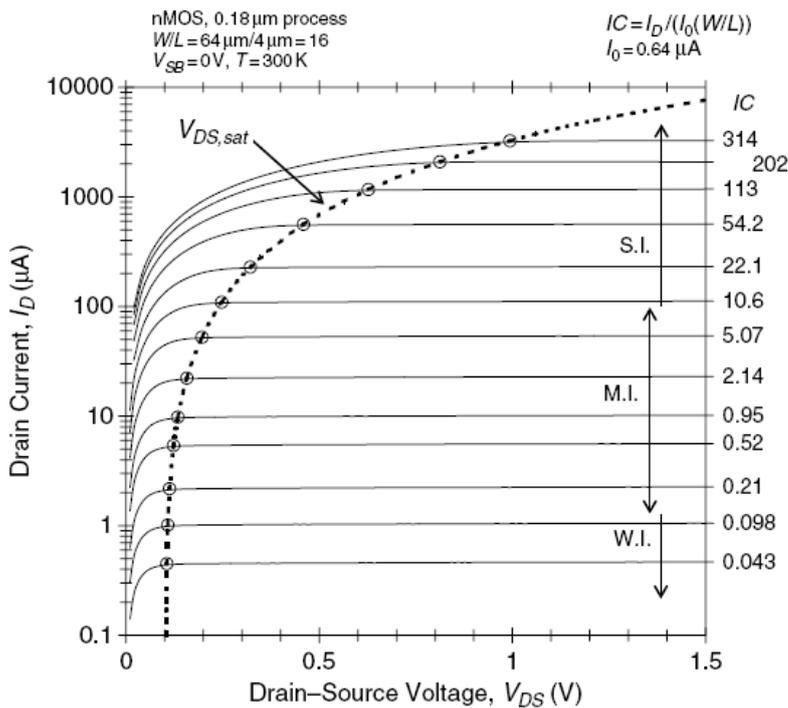


Figure 17 – Picture from [22]

Thus, despite that conventional PMOS transistors biased in Strong Inversion can be used as resistors in a wide area of V_{SD} values, when biased in Moderate or Weak Inversion, the linear region of the characteristic is limited to V_{SD} voltages below the subthreshold saturation voltage $V_{SD,sat} = 3-4U_T$ (75-100mV). Moreover, as seen in figure 17, the upper limit of the linear region is independent of the value of V_{SG} and, therefore, cannot be set by design[5].

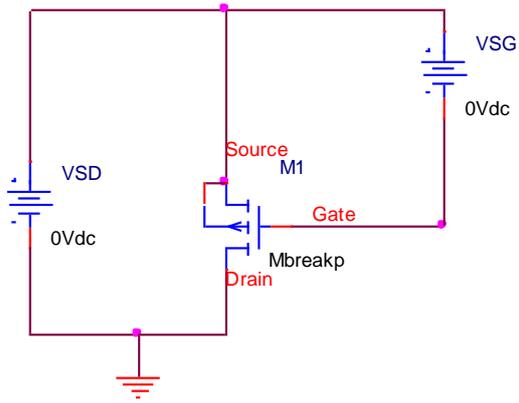


Figure 18

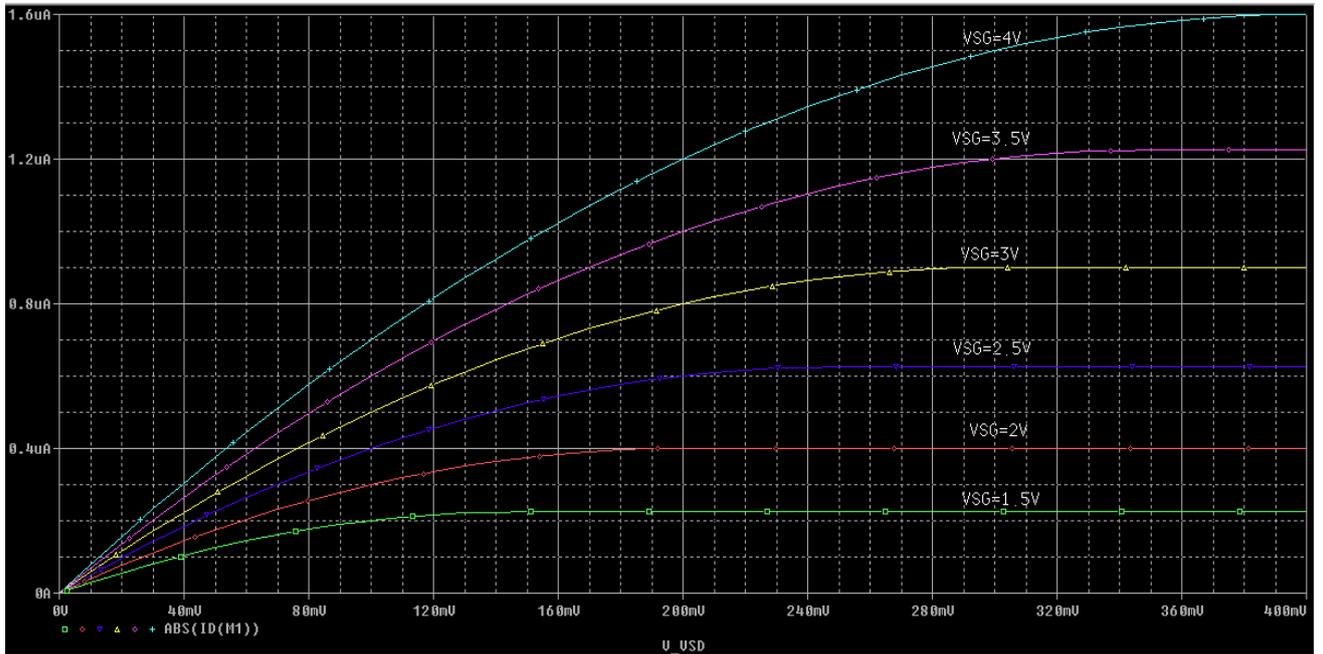


Figure 19a

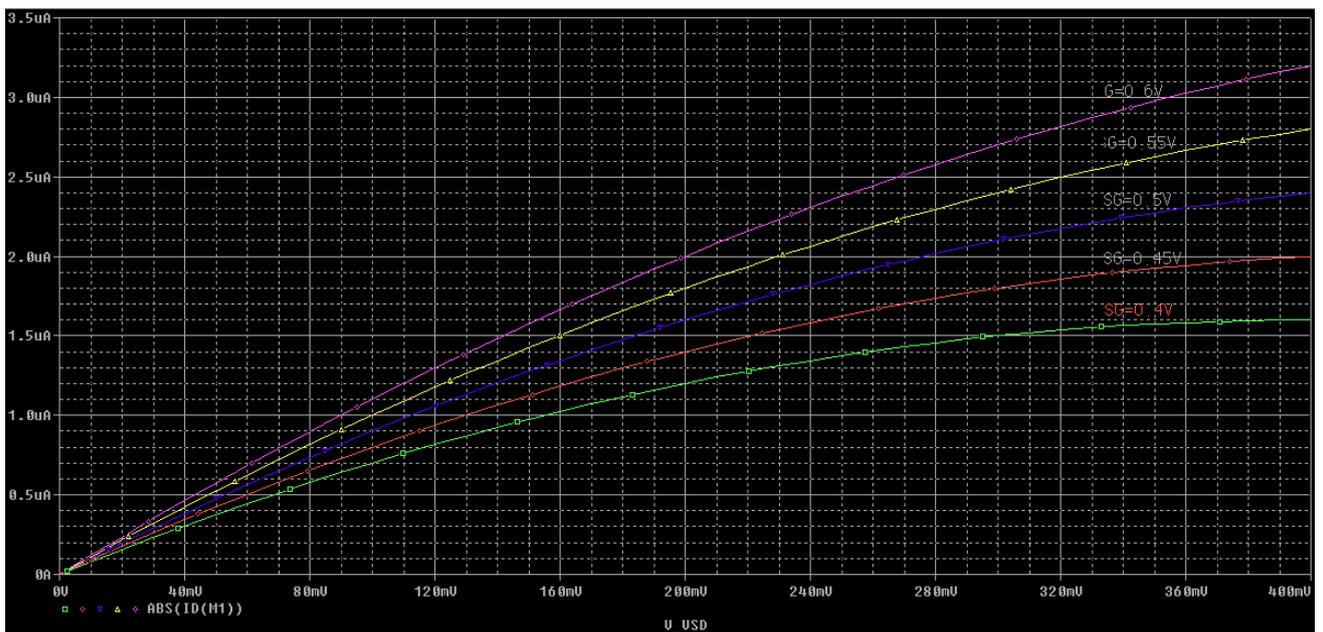


Figure 19b

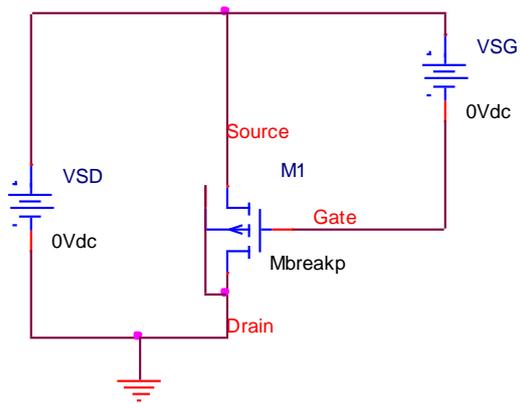


Figure 20

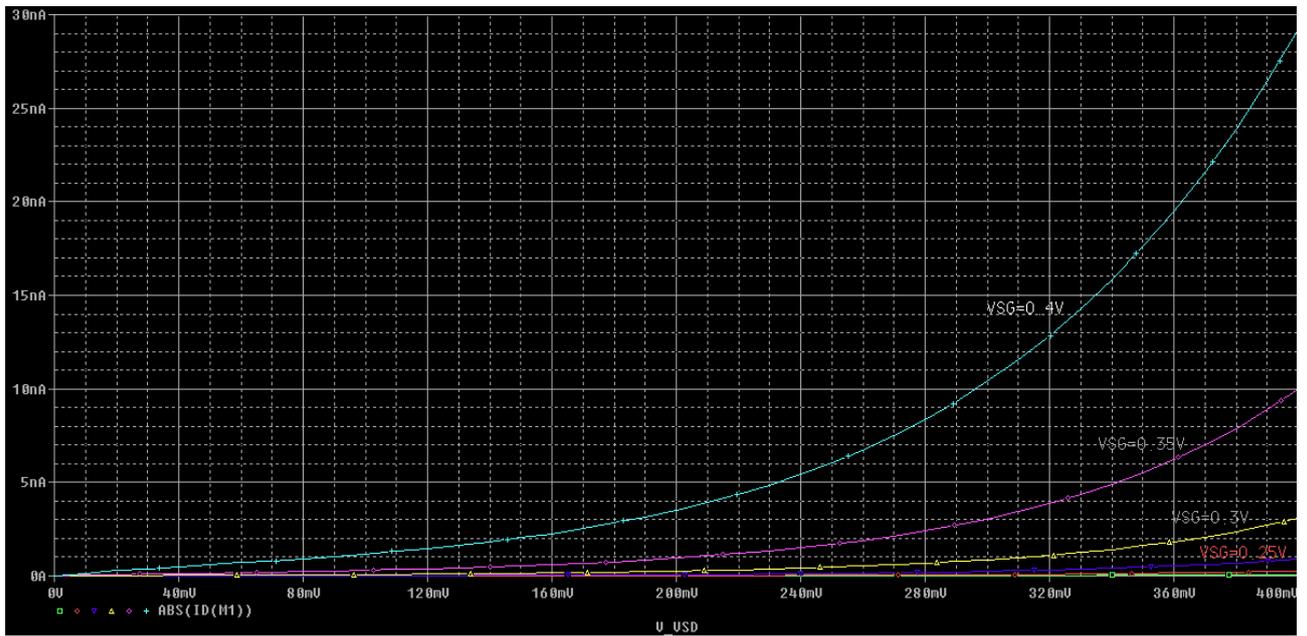


Figure 21a

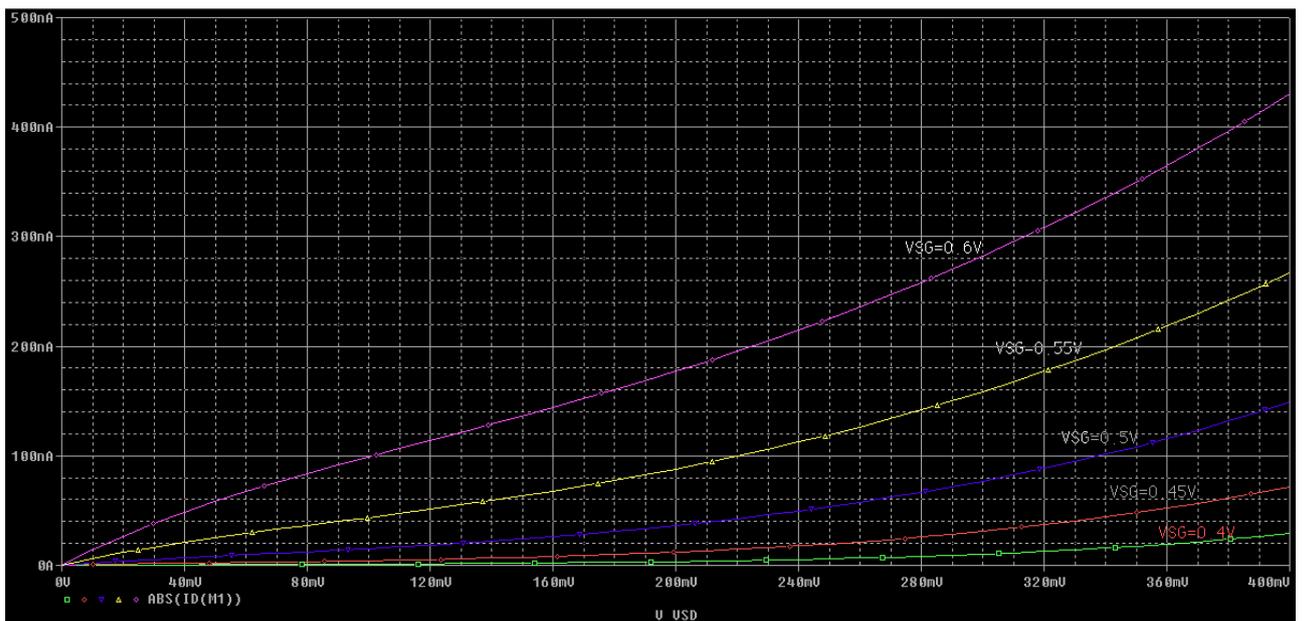


Figure 21b

Consequently , for a conventional PMOS to be biased in Weak Inversion so that we can have a low I_{SD} , the voltage swing V_{SWING} of the SCL gate must be below 100mV which is not enough for robust gate operation , as according to figure 8 , V_{SWING} should be larger than $4nU_T \cong 150mV$.

Additionally , even if we consider operation in the saturation region , as mentioned earlier , the I_{SD} - V_{SD} characteristic is highly nonlinear giving for the subthreshold SCL inverter adopting plain subthreshold PMOS loads $v_o = -\frac{v_i}{n}$ which is a straight line not guaranteeing bistability , thus rendering unfeasible the implementation of SCL gates[5].

The alternative option is the use of a Bulk-Drain connected PMOS load as depicted in figure 20. As observed in figures 21a and 21b simulated again in Pspice using the EKV v2.6 for 0.5 μ m CMOS , this configuration exhibits an approximately linear I_{SD} - V_{SD} characteristic even above V_{SDsat} [5]-[19]-[23]-[24].

Due to the body effect , the threshold voltage V_t depends on the substrate potential which is now V_{DB} , and when V_{SD} increases thereby reducing V_D and V_{DB} (see figure 22) , then V_t decreases , causing the drain current to increase and the device to gradually enter the Strong Inversion region. This phenomenon becomes even more pronounced if V_{SG} close to V_t is used[5]. This body effect is expressed by the following relation:

$$V_t = V_{TO} + (n-1).V_{DB}$$

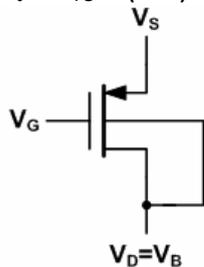


Figure 22 – Picture from [19]

Using a CMOS p-substrate technology , this topology can be used only for PMOS devices since they are implemented in n-wells (see figure 23). Also , each PMOS must be confined in its own n-well [3] , so that each PMOS can have a different value of $V_D(=V_B)$.

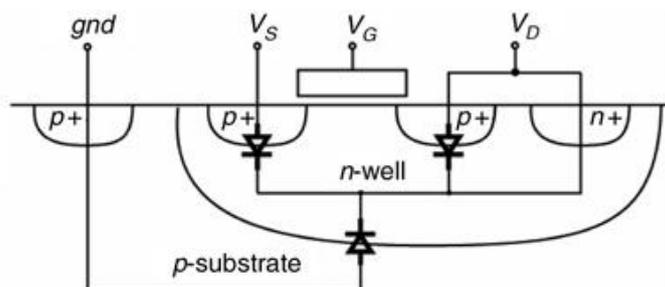


Figure 23 – Picture from [23]

The diode between the source p⁺ - diffusion and the n-well is maintained off by keeping the voltage V_{SD} below the diode turn-on voltage of about 600 mV.

In an SCL gate operating in Weak Inversion , two Bulk-Drain PMOS loads should be used , where the Drain of the first PMOS is v_{o1} , whereas the Drain of the second PMOS is v_{o2} , as shown in figure 24.

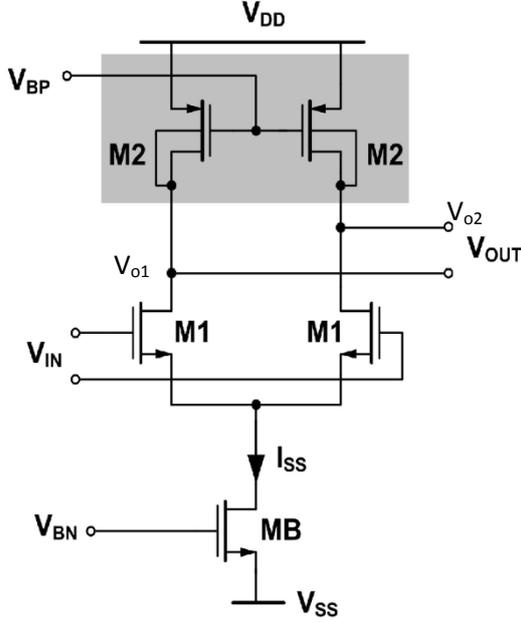


Figure 24 – Picture from [20]

Because of the behavior of this load device which maintains the linearity of the I_{SD}-V_{SD} characteristic going from the Weak through Medium , to Strong Inversion , it is necessary to model the drain current using the EKV model , which is the MOS model with the most precise description of the transition of the drain current between those regions.

The drain current from the Weak to Strong Inversion region according to the EKV model is [26] :

$$I_{SD} = I_F - I_R \Rightarrow$$

$$\Rightarrow I_{SD} = I_S \cdot \ln^2 \left(1 + e^{\frac{V_{SGt}}{2 \cdot n \cdot U_T}} \right) - I_S \cdot \ln^2 \left(1 + e^{\frac{V_{SGt}}{2 \cdot n \cdot U_T}} \cdot e^{-\frac{V_{SD}}{2 \cdot U_T}} \right)$$

Where : $V_{SGt} = V_{SG} - V_t(V_{SD}) = V_{SG} - (V_{TO} + (n-1) \cdot V_{SD})$

and $I_S = \frac{W}{L} \cdot 2 \cdot n \cdot \mu \cdot C_{ox} \cdot U_T^2$

The drain current of the Bulk-Drain connected PMOS transistor biased in the Weak Inversion region , is given by the EKV model [3]-[8]-[10] as:

$$I_{SD} = I_0 \cdot e^{\frac{V_{BG} - V_{TO}}{n_p U_T}} \left(e^{\frac{-V_{BS}}{U_T}} - e^{\frac{-V_{BD}}{U_T}} \right) , \quad \text{where:} \quad I_0 = 2 \cdot n_p \cdot \mu \cdot C_{ox} \cdot \frac{W}{L_e} U_T^2$$

Here , $V_{BD} = 0$, hence :

$$I_{SD} = I_0 \cdot e^{\frac{V_{BG}-V_{TO}}{n_p U_T}} \left(e^{\frac{-V_{BS}}{U_T}} - 1 \right) \quad \Rightarrow \quad I_{SD} = I_0 \cdot e^{\frac{V_{DG}-V_{TO}}{n_p U_T}} \left(e^{\frac{-V_{SD}}{U_T}} - 1 \right)$$

The output small signal resistance of the Bulk-Drain connected PMOS is [8]:

$$R_{SD} = \left(\frac{\partial I_{SD}}{\partial V_{SD}} \right)^{-1} = \left(\frac{n_p \cdot U_T}{I_b} \right) \cdot \left((n_p - 1) \cdot e^{(n_p - 1) \cdot v_{SD}} + e^{-v_{SD}} \right)^{-1} \Rightarrow$$

$$\Rightarrow R_{SD} = \left(\frac{n_p \cdot U_T}{I_{SD}} \right) \cdot \left(\frac{e^{\frac{v_{SD}}{U_T}} - 1}{(n_p - 1) \cdot e^{\frac{v_{SD}}{U_T}} + 1} \right)$$

Where : $v_{SD} = \frac{V_{SD}}{n_p U_T}$ and $I_b = I_0 \cdot e^{\frac{V_{SG}-V_{TO}}{n_p \cdot U_T}}$

In which $I_0 = 2 \cdot n_p \cdot \mu \cdot Cox \cdot \left(\frac{W}{L_e} \right) \cdot U_T^2$

By inspection of these relations , it can be seen that R_{SD} can be controlled through I_{SD} , by altering the value of V_{SG} , on which R_{SD} depends exponentially. Thus , R_{SD} can be adjusted in a very wide range[3]-[8].

V. IMPLEMENTATION OF THE REPLICA BIAS

The resistivity of the Bulk-Drain connected loads used in an SCL gate, can be controlled through the gate voltage of the PMOS transistors V_{BP} , which should be chosen as low as possible (ideally $V_{BP}=0V$), so as to reduce the size of the PMOS devices[4]. V_{BP} is generated on-chip by a bias circuit such as the one shown in figure 25 for the implementation of an SCL inverter.

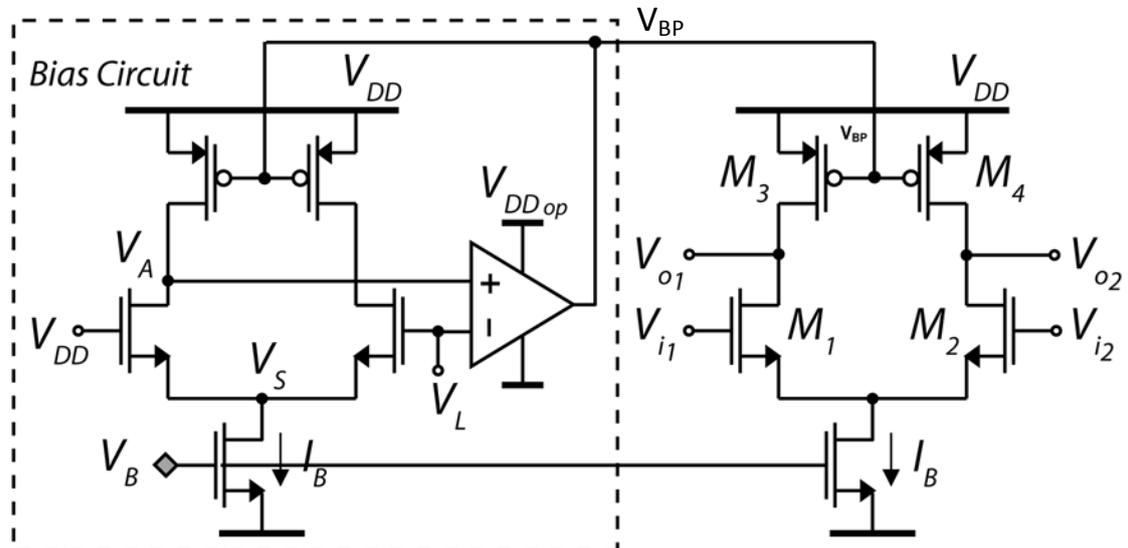


Figure 25 – Picture from [5]-[23]-[26]

The feedback bias circuit which can be shared among several logic gates comprises an identical inverter gate and an single-stage OTA, biased in the Weak Inversion region. The OTA forces V_A to track the desired low logic level V_L by dynamically adjusting V_{SG} and hence the resistance of the PMOS loads. The single-stage OTA that is used, is depicted in figure 26, and avoids any compensation capacitor other than the load itself, which is possible, as the major part of the voltage gain is achieved at the output node[17].

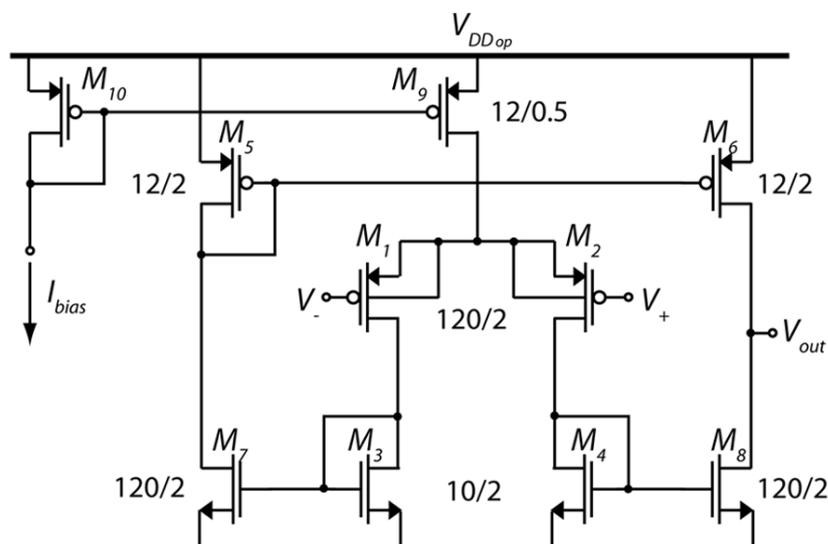


Figure 26 – Picture from [5]

The OTA is designed and simulated in Pspice as shown in figure 27 , presenting a low-frequency open-loop gain given by the following relation[27]:

$$A_{vo} = \frac{g_{m1} \cdot B}{g_{ds6} + g_{ds8}} = \frac{B}{n_1 \cdot U_T \cdot (\lambda_6 + \lambda_8)}$$

Where :

$$B = \frac{W_8 / L_8}{W_4 / L_4} \quad \text{and} \quad g_{ds} = \lambda \cdot I_{DS}$$

Furthermore , the Gain Bandwidth (or unity-gain frequency) is given by:

$$GB = \frac{g_{m1} \cdot B}{C_{BP}} \quad , \quad \text{where} \quad g_{m1} = \frac{I_D}{n_1 \cdot U_T} \quad [27]$$

The 3-dB frequency (and location of the dominant pole too) , is given by:

$$f_{3dB} = \frac{1}{2 \cdot \pi \cdot (r_{o6} \parallel r_{o8}) \cdot C_L} \quad [28] \quad , \quad \text{where} \quad r_{o6} = \frac{1}{\lambda_6 \cdot I_{DS,sat}} \quad \text{and} \quad r_{o8} = \frac{1}{\lambda_8 \cdot I_{DS,sat}}$$

The EKV v2.6 model for 0.5 μ m CMOS taken from [18] and used for the simulation of the OTA , provides us with these parameters:

$$\begin{aligned} \lambda_6 = \lambda_{PMOS} = 1.1 \quad , \quad \lambda_8 = \lambda_{NMOS} = 0.23 \quad , \quad \gamma_{PMOS} = 0.69\sqrt{V} \quad , \\ \phi_{PMOS} = 0.87V \quad , \quad \gamma_{NMOS} = 0.71\sqrt{V} \quad , \quad \phi_{NMOS} = 0.97V \end{aligned}$$

The slope factor for the PMOS transistor M1 , is given by:

$$n_1 = n_{PMOS} = 1 + \frac{\gamma}{2 \cdot \sqrt{V_p + \phi}}$$

and for large device geometries:

$$V_p \approx V_G - V_{TO} + \gamma \cdot \sqrt{\phi} - \gamma \left(\sqrt{V_G + \left(\frac{\gamma}{2}\right)^2} - \frac{\gamma}{2} \right) \quad [31]$$

Where γ is the body effect factor and the parameter ϕ is the approximation of the surface potential ψ_s , in Strong Inversion.

Another approximation is the following one:

$$n \approx 1 + \frac{\gamma}{2 \cdot \sqrt{\phi}} \quad , \quad \text{using which , we have that:}$$

$$n_n \approx 1.36 \quad \text{and} \quad n_p \approx 1.37$$

Using the above relations and by inspection of figure 27a , we can calculate that:
 $r_{o6} \approx 9.1 \cdot 10^6 \Omega = 9.1M\Omega$, $r_{o8} \approx 43.5 \cdot 10^6 \Omega = 43.5M\Omega$, $r_{o6} // r_{o8} = 7.52M\Omega$

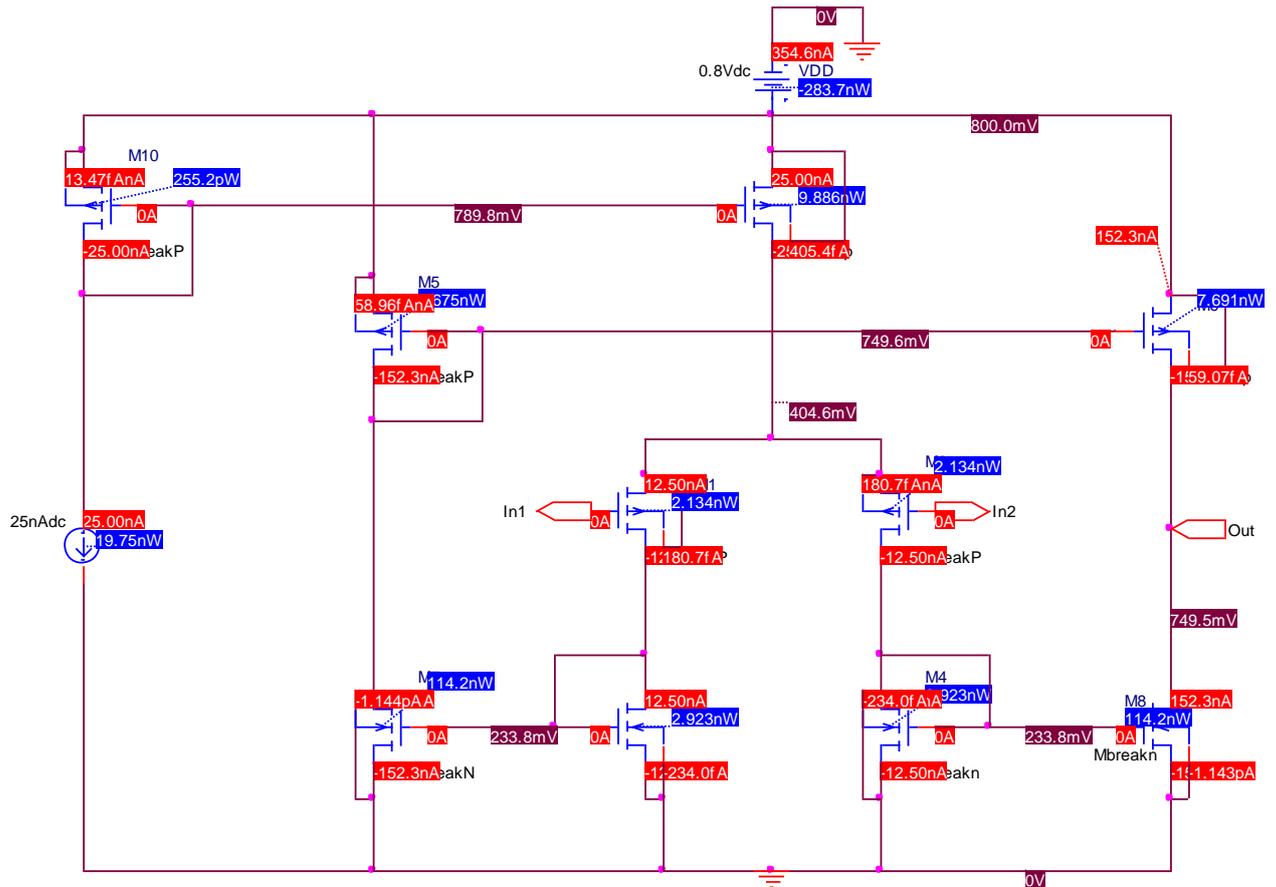


Figure 27b

To be able to simulate the open-loop frequency response of the OTA , we use the scheme proposed in [28] and seen in figure 28 , assuming an output load of 110pF [5]. As this is a PMOS differential input pair OTA , the positive input is the input 2 and the negative input is the input 1.

The generation of the differential voltage for the OTA is from [29]. Also , E_p and E_n are voltage controlled voltage sources , provided by the Analog Library of the Pspice , the gain of which is set to 0.5. The common-mode input voltage V_{CM} of the OTA is set to 0.4V (i.e. to the average of the dc power-supply voltages $V_{DD}=0.8V$ and $GND=0V$) to maximize the available input signal swing[29].

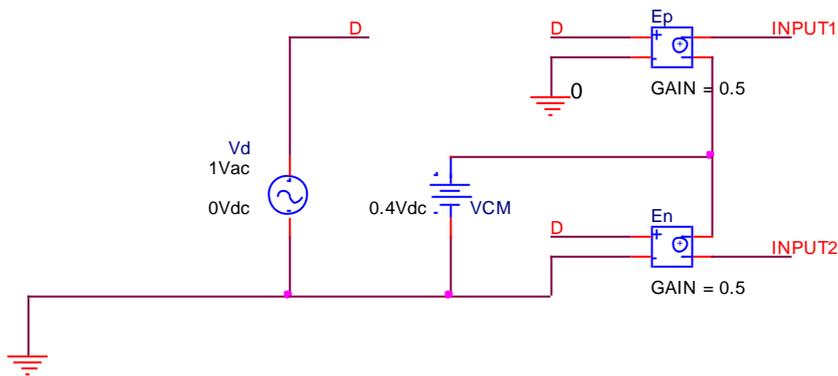
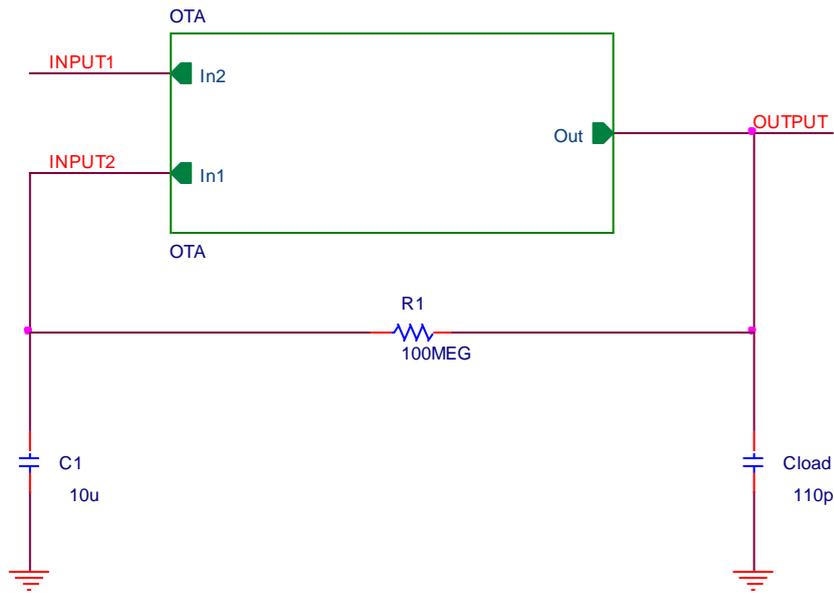


Figure 28

The feedback resistor and capacitor used in figure 28 , form a time constant so large that for all intents and purposes none of the AC output voltage is fed back to the inverting input. However , the DC bias level is fed back so that the OTA biases up correctly (all MOSFETs are operating in the saturation region)[28].

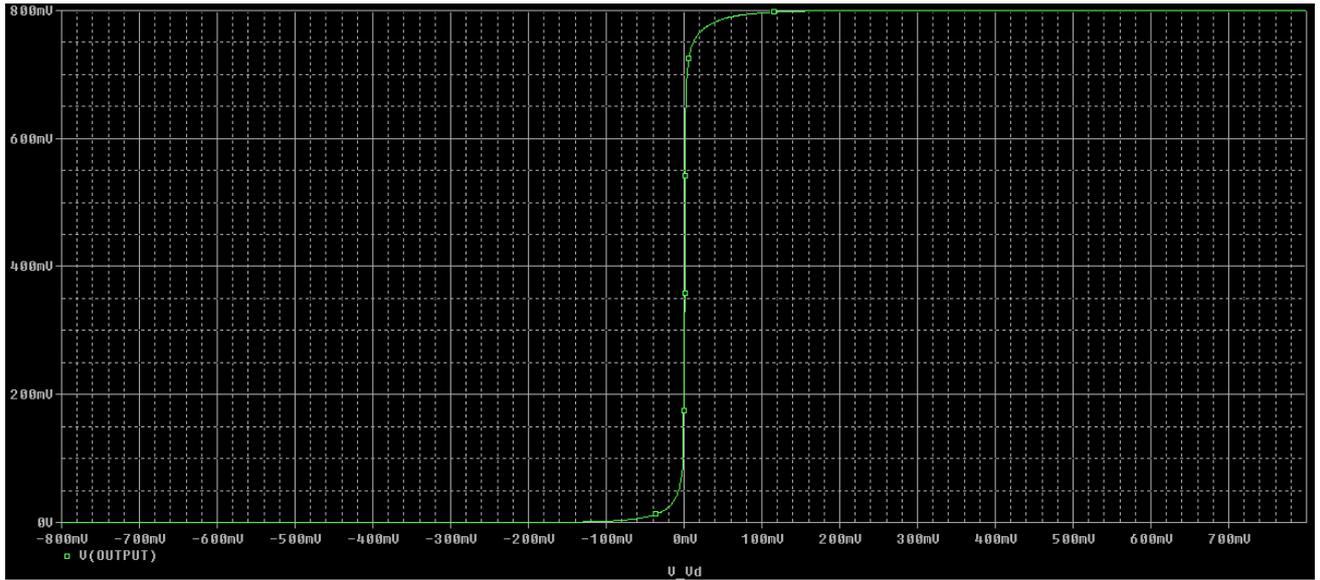


Figure 29a

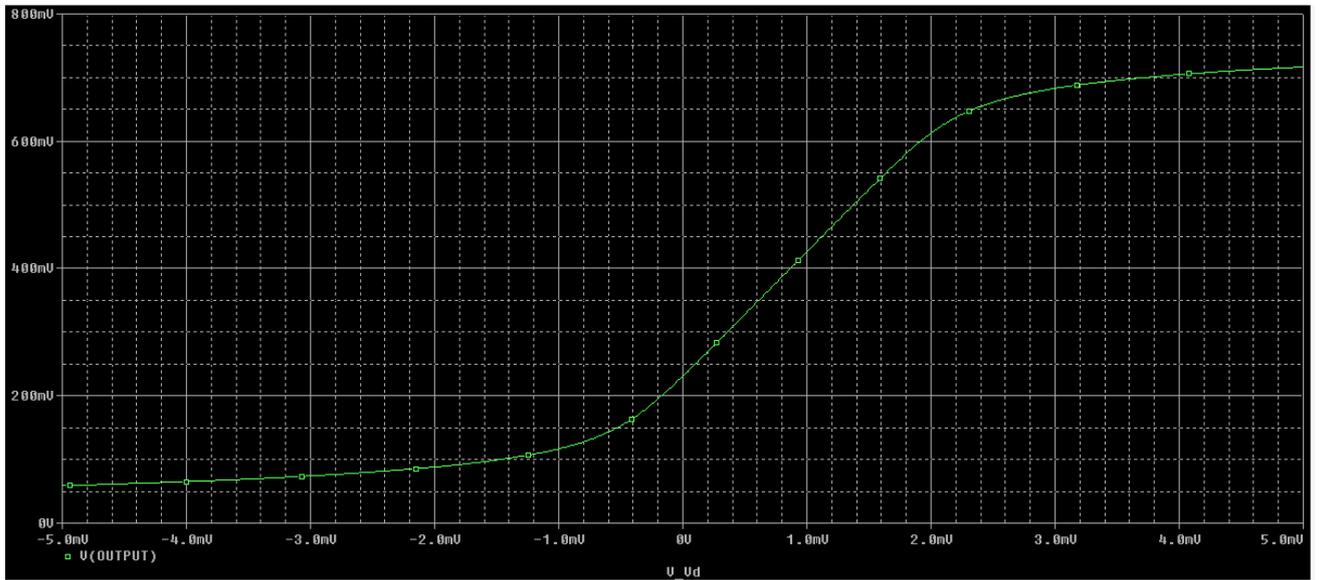


Figure 29b

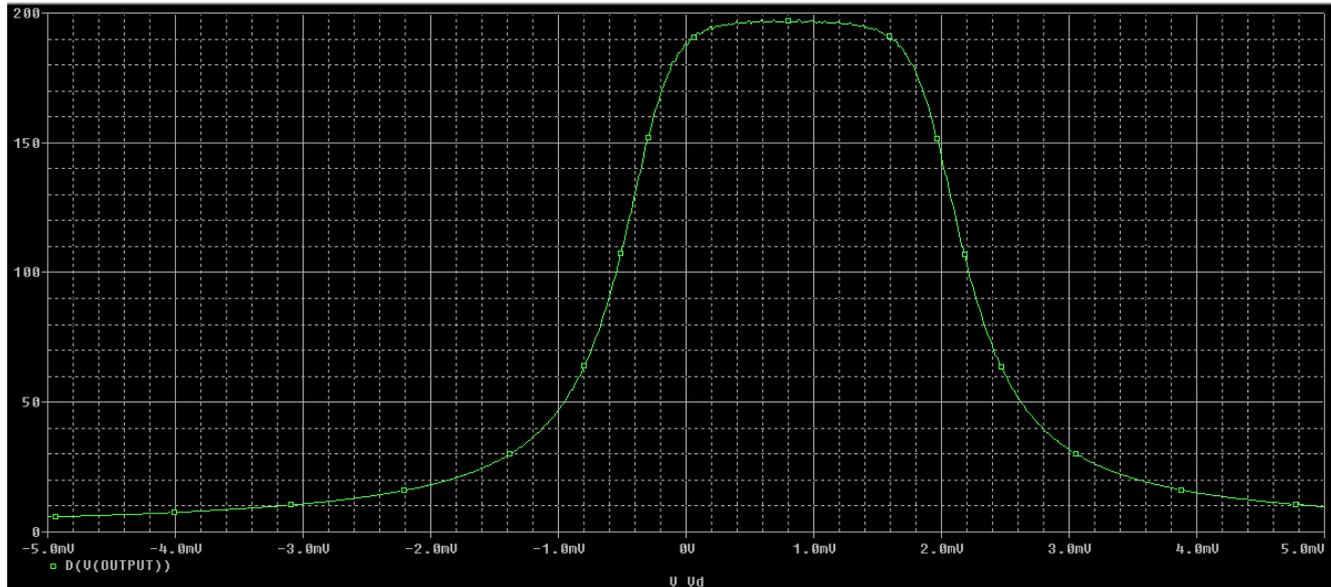


Figure 29c

To compute the large-signal differential transfer characteristic of the OTA , we perform a dc-analysis simulation in Pspice , with the differential voltage input V_d swept over the range $-0.8V$ to $0.8V$ and we plot the corresponding output voltage as seen in figure 29a. The slope of this characteristic(i.e. , $\frac{\partial V_{OUT}}{\partial V_{IN}}$) corresponds to the

differential gain of the amplifier. To examine the high-gain region more closely , the dc analysis is repeated with V_d swept over the range $-5mV$ to $5mV$ at increments of $10\mu V$ and the resulting differential dc transfer characteristic is plotted in figure 29b.

The linear region of the large-signal differential characteristic is bounded approximately by $V_d=-1mV$ and $V_d=1mV$. Over this region , the output level changes from $V_{OUT}=150mV$ to about $V_{OUT}=650mV$ in a linear fashion. Thus , the **output voltage swing** for this amplifier is between $150mV$ and $650mV$. We also observe from figure 29b that $V_d \approx 866\mu V$ when $V_{OUT}=400V$. Therefore , the amplifier has an **input offset voltage** V_{os_input} of $-866\mu V$ as this is by convention , the negative value of the x-axis intercept of the large-signal differential transfer characteristics[29]. This corresponds to an **output offset voltage** of $A_d \cdot V_{os_input} \approx -200 \cdot 866 \cdot 10^{-6} = 173.2mV$, as the **DC differential gain** is about 200 as seen in figure 29c. This voltage offset is inherent in the design and is not the result of component or device mismatches. Thus , it is usually referred to as a systematic offset[29].

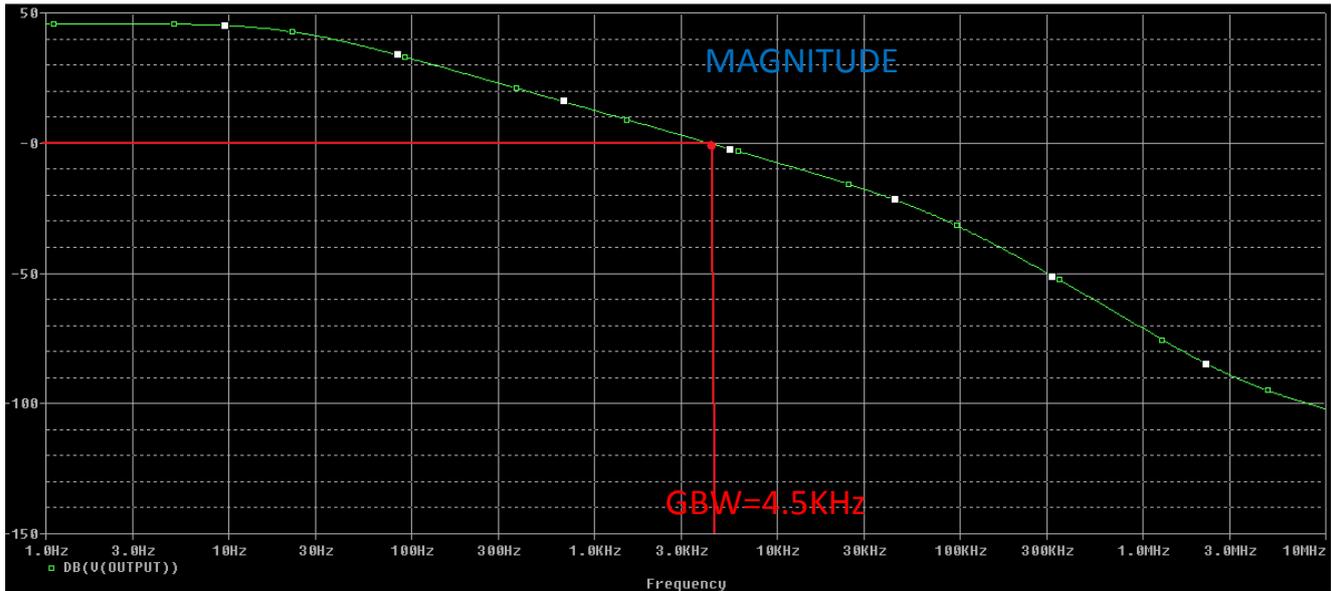


Figure 30

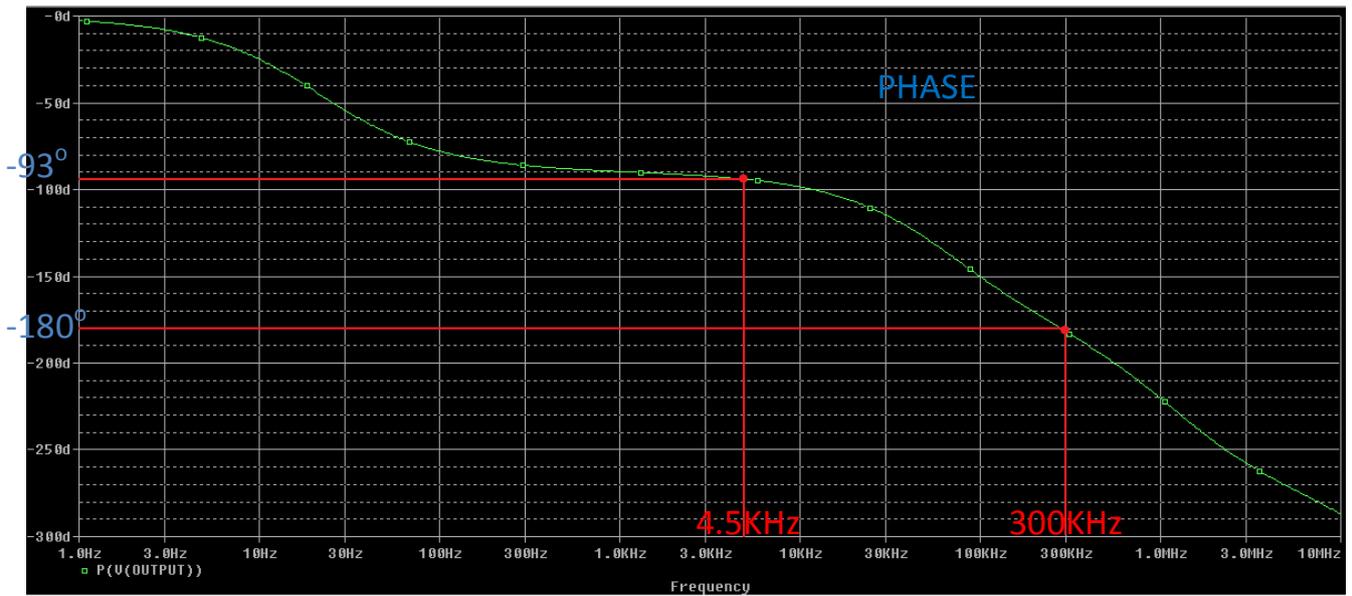


Figure 31

The open-loop frequency response of the OTA is depicted in figures 30 and 31. As seen, $A_d \approx 48\text{dB}$ and the Phase Margin is about 87° .

In sum:

- $A_d \approx 48\text{dB}$
- $PM \approx 87^\circ$
- $F_{3\text{dB}} \approx 30\text{Hz}$
- $GBW \approx 4.5\text{KHz}$
- $V_{\text{SWING}} \approx 500\text{mV}$
- $V_{\text{os_input}} \approx -866\mu\text{V}$
- $V_{\text{os_output}} \approx 173.2\text{mV}$

VI. IMPLEMENTATION OF THE SCL INVERTER

The SCL inverter is implemented and simulated in Pspice , following the configuration proposed in [4]-[5] , as seen in figures 32 , 33 and 34. The model used for simulation is the EKV v2.6 for 0.25 μ m CMOS from [36] and the generation of the differential voltage for the inverter is again from [29].

The three parameters governing SCL gates design are the differential voltage swing V_{SWING} , the bias current I_{SS} and the noise margins NM. Noise margins are related to the small-signal voltage gain A_{dSCL} and together with I_{SS} impose a minimum size on the gate width of the differential pairs. On the other hand , the differential voltage swing is the product of the bias current and of the equivalent on-resistance of the PMOS loads and hence relates to the gate width of the loads. Analytically , the sizing procedure is as follows:

1. We first chose a value for I_{SS} , for example $I_{SS}=1nA$.

2. Next , the gate width of the PMOS loads W_{PMOS} is chosen so that with the largest load size R , we can have the maximum possible voltage swing V_{SWING} , through the relations that connect $W_{PMOS} - R - V_{SWING} - I_{SS}$ as given by [3]:

$$V_{SWING} = R.I_{SS} \Rightarrow$$

$$\Rightarrow V_{SWING} = \left[\left(\frac{n_p \cdot U_T}{I_b} \right) \cdot ((n_p - 1) \cdot e^{(n_p - 1) \cdot v_{SD}} + e^{-v_{SD}})^{-1} \right] \cdot I_{SS}$$

with I_{SS} being fixed to the given value (1 nA) and R being dependent of W_{PMOS}
Where

$$v_{SD} = \frac{V_{SD}}{n_p \cdot U_T} \quad , \quad I_b = I_0 \cdot e^{\frac{V_{SG} - V_{TO}}{n_p \cdot U_T}} \quad ,$$

$$I_0 = 2 \cdot n_p \cdot \mu \cdot Cox \cdot \left(\frac{W}{L_e} \right) \cdot U_T^2$$

3. Then , the gate width of the NMOS differential pairs W_{NMOS} is computed by keeping R stable , and changing the value of V_{SWING} through I_{SS} , in order to have the maximum possible noise margins using again the relations given by [3] that connect $NM - A_v - I_{SS}$:

$$\frac{NM}{V_{SWING}} = \sqrt{1 - \frac{1}{A_v}} - \frac{1}{A_v} \cdot \tanh^{-1} \cdot \left(\sqrt{1 - \frac{1}{A_v}} \right) \Rightarrow$$

$$\Rightarrow NM = \left[\sqrt{1 - \frac{1}{A_v}} - \frac{1}{A_v} \cdot \tanh^{-1} \cdot \left(\sqrt{1 - \frac{1}{A_v}} \right) \right] \cdot V_{SWING} \Rightarrow$$

$$\Rightarrow NM = \left[\sqrt{1 - \frac{1}{A_v}} - \frac{1}{A_v} \cdot \tanh^{-1} \cdot \left(\sqrt{1 - \frac{1}{A_v}} \right) \right] \cdot R \cdot I_{SS}$$

with R being fixed to the value calculated at step 2 , and I_{SS} being dependent of W_{PMOS} [38].

4. Finally , the gate lengths L_{PMOS} , L_{NMOS} of the transistors are usually fixed to the minimum allowed value , in order to save silicon area[38].

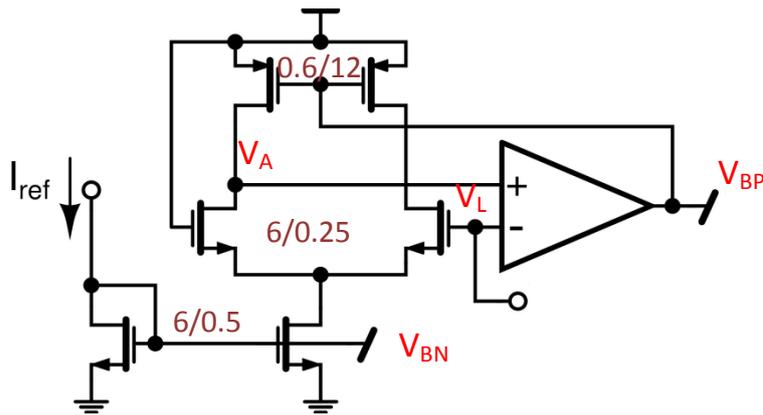


Figure 32 – Picture from [4]

V_L is given the value of $V_{DD} - V_{SWING}$ by an external voltage source, so that ideally $V_A = V_L = V_{DD} - V_{SWING}$.

Also, $I_{REF} = 1\text{nA}$ and $V_{DD} = 0.4\text{V}$, whereas $V_{SWING} = 0.2\text{V}$. The adoption of logic swings larger than 100mV reduces the impact of the offset on the noise margin still guaranteeing a robust operation of the gate[5].

Thus, the resistance of the Drain-Bulk connected PMOS devices is:

$$V_{SWING} = R \cdot I_{SS} \Rightarrow R = \frac{0.2}{1.10^{-9}} = 200\text{M}\Omega$$

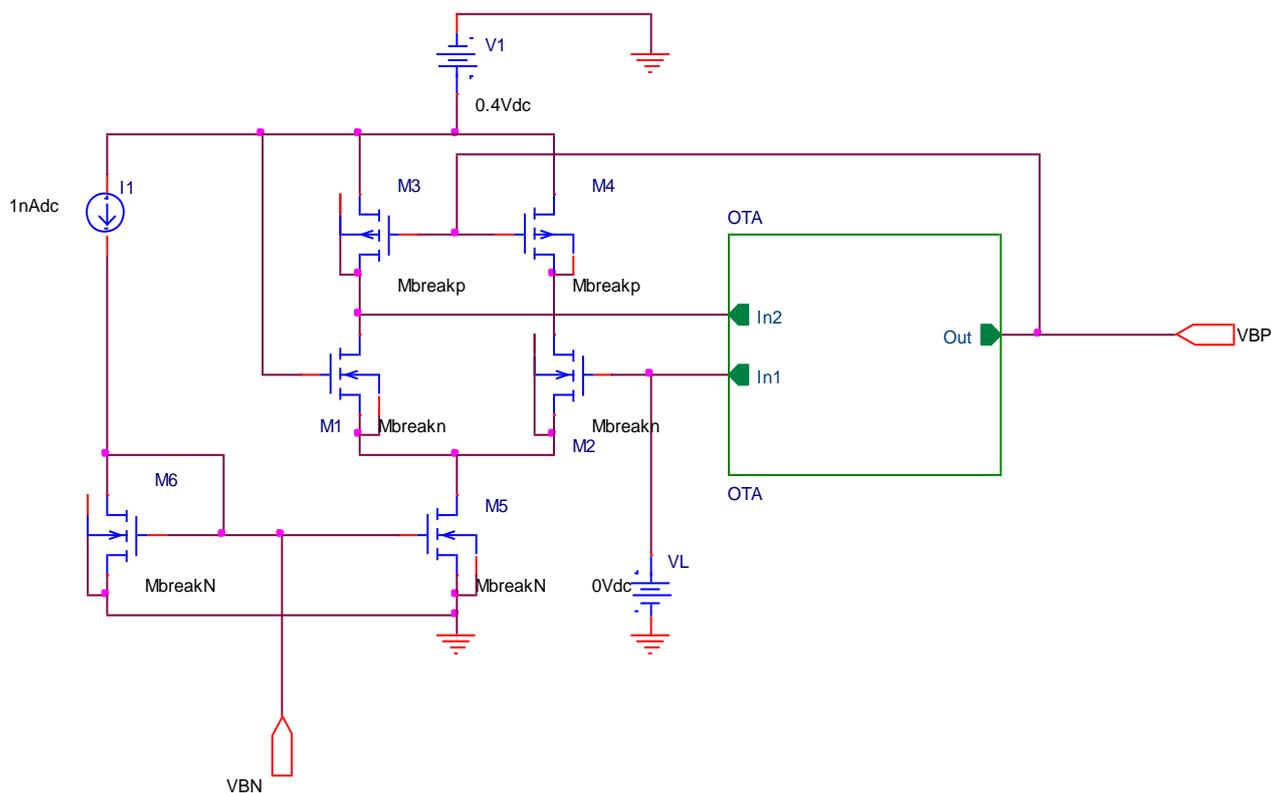


Figure 33

The common-mode input voltage V_{CM} of the SCL inverter seen in figure 34 is set to 0.2V (i.e. to the average of the dc power-supply voltages $V_{DD} = 0.4\text{V}$ and $\text{GND} = 0\text{V}$) to maximize the available input signal swing[29].

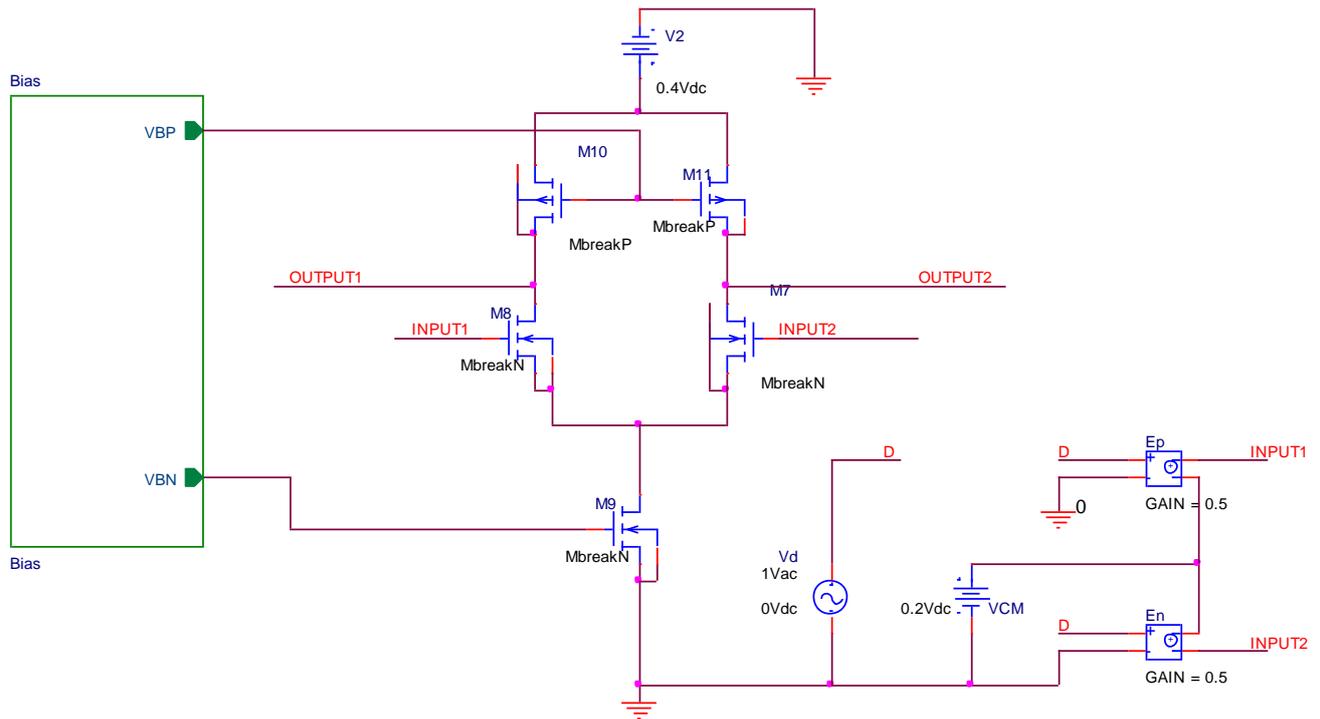


Figure 34

The replica bias circuit seen in figures 32 and 33 , should be well matched to the SCL gates in order for its operating point to have a very low deviation[8]. The role of this circuit is to control the resistance of the load devices and thus adjust the output voltage swing V_{SWING} with respect to the tail bias current I_{SS} , while at the same time tracking the variations on temperature and supply voltage and hence compensating their effect on the circuit performance[16].

The simulated DC transfer characteristics are depicted in figures 35 and 36.

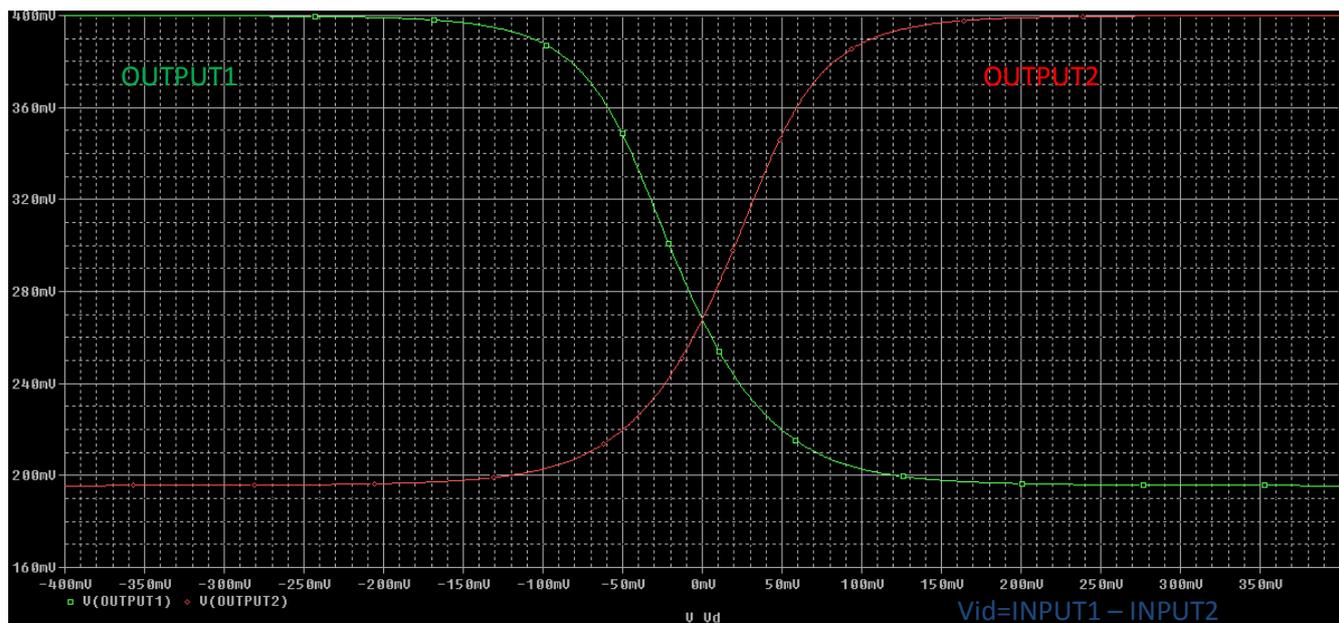


Figure 35

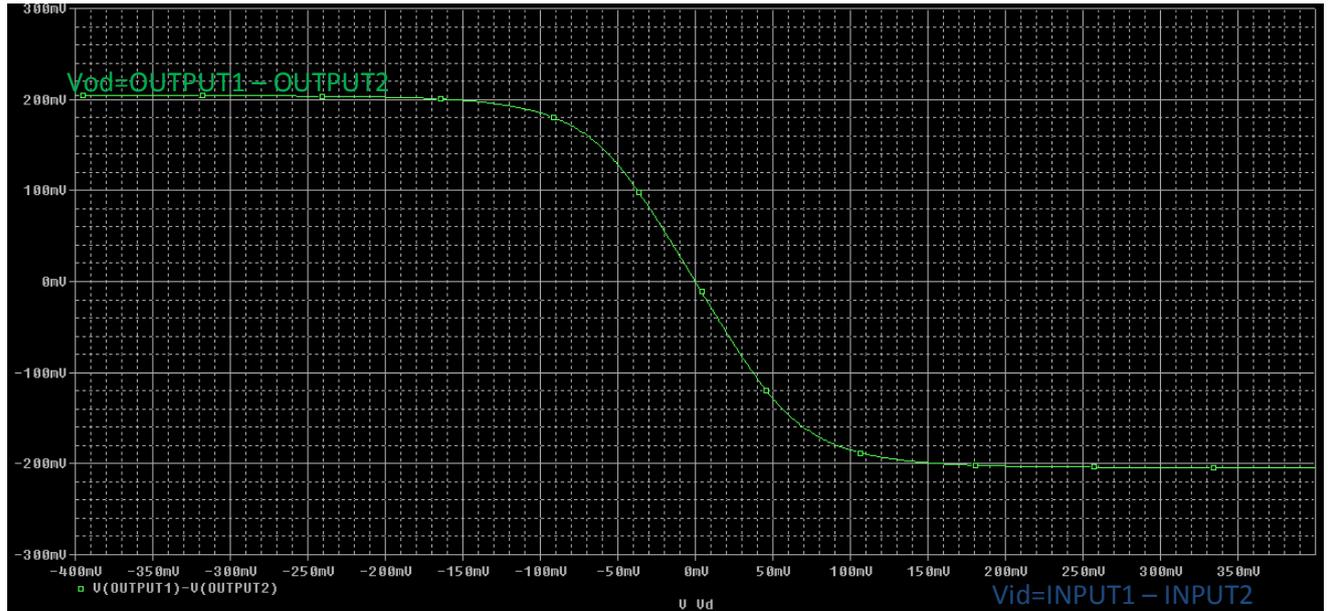


Figure 36

The simulated DC differential gain is seen in figure 37 and is:

$$-A_{dSCL}(V_{id}) = \frac{V_{SWING}}{2.n.U_T} \cdot \frac{1}{\cosh^2 \left(\frac{V_{id}}{2.n.U_T} \right)} \quad [5]$$

The maximum value of the differential mode gain A_V occurs for $V_{id}=0V$:

$$|A_{dSCL}|_{\max} = \frac{V_{SWING}}{2.n.U_T} \quad [5]$$

With:

$$n_n \approx 1 + \frac{\gamma}{1 + \sqrt{2.\phi_{NMOS}}} \approx 1.21 \quad \text{and} \quad n_p \approx 1 + \frac{\gamma}{1 + \sqrt{2.\phi_{PMOS}}} \approx 1.26$$

, as from the model used , it is given that $\gamma_{PMOS} = 0.6\sqrt{V}$, $\phi_{PMOS} = 0.87 \text{ v}$,

$$\gamma_{NMOS} = 0.5\sqrt{V} \quad , \quad \phi_{NMOS} = 0.95 \text{ v}$$

Therefore ,

$$|A_{dSCL}|_{\max} = \frac{0.2}{2 * 1.25 * 25.8 * 10^{-3}} \approx 3.1$$

(In technologies with a smaller n , a larger A_V can be achieved[5].)

An important observation is that in order to have two stable logic states ,

$|A_{dSCL}|_{\max}$ has to be larger than one. This condition defines the absolute minimum for the voltage swing at the output of the SCL inverter :

$$V_{SWING} = R.I_{SS} > 2.n.U_T \quad , \quad \text{which using EKV v.2.6 for } 0.25\mu\text{m CMOS} \text{ , is about } 64.5\text{mV.}$$

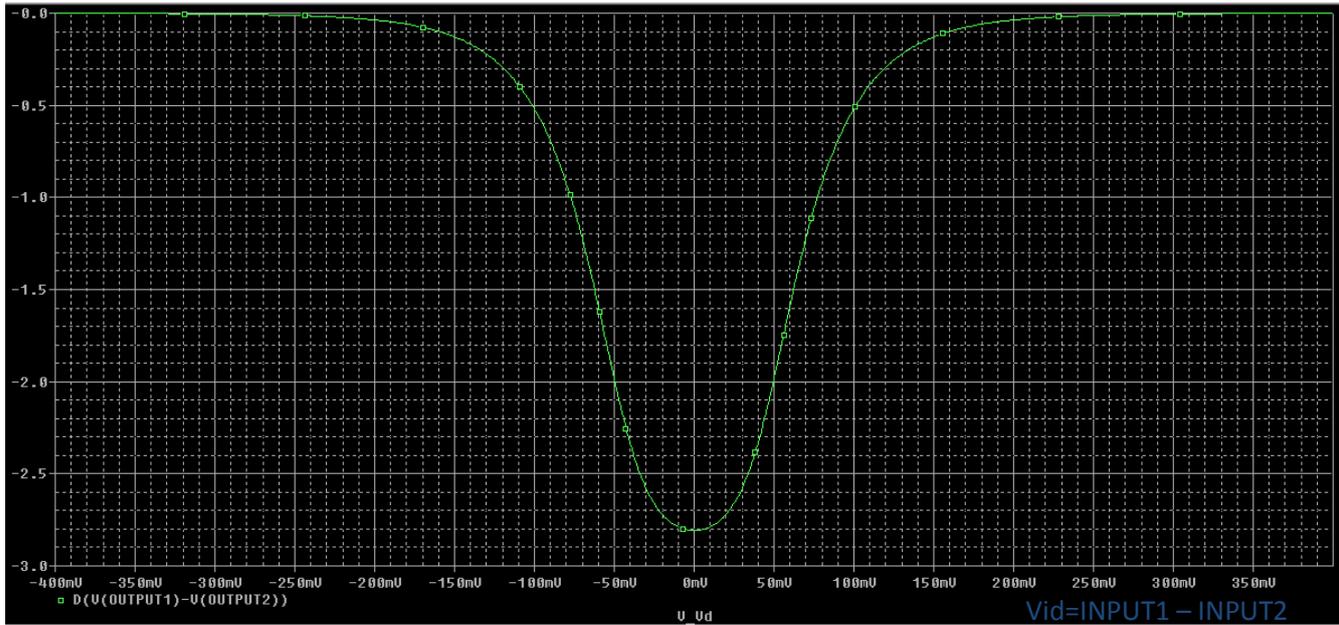


Figure 37

The noise margins of the SCL inverter biased in the subthreshold region are defined as follows:

$$\text{Low Noise Margin: } NM_L = V_{IL,max} - V_{OL,max}$$

$$\text{High Noise Margin: } NM_H = V_{OH,min} - V_{IH,min} \quad [32]$$

Where : $V_{IL,max}$ = maximum Low input voltage

$V_{OL,max}$ = maximum Low output voltage

$V_{OH,min}$ = minimum High output voltage

$V_{IH,min}$ = minimum High input voltage

$V_{IL,max}$, $V_{OL,max}$, $V_{OH,min}$ and $V_{IH,min}$ are defined at the unity gain point where the slope is -1, as shown in figure 38, where on the first left y axis the DC transfer characteristic of the inverter is plotted, whereas on the second left y axis the gain or slope ($\frac{\partial V_{out}}{\partial V_{in}}$) of the inverter is plotted.

According to [7], the SCL inverter operating in the subthreshold region, presents an output differential voltage v_o given by :

$$v_o = v_{o1} - v_{o2} = -V_{SWING} \tanh\left(\frac{v_i}{2.n.U_T}\right) \quad , \quad \text{where} \quad v_i = v_{i1} - v_{i2}$$

Evaluating the points where the DC characteristic presents a slope (or DC gain) of -1, we have that:

$$V_{IH,min} = -V_{IL,max} = n.U_T \cdot \ln\left(2 \cdot \frac{V_{SWING}}{n.U_T} - 2\right) \quad \text{and}$$

$$V_{OH,min} = -V_{OL,max} \approx V_{SWING} - 1.15.n.U_T$$

Where $V_{IH,min} = -V_{IL,max}$ and $V_{OH,min} = -V_{OL,max}$ is assumed due to the symmetry of the DC characteristic which is seen in figure 36 and is given by the previous equation for v_o .

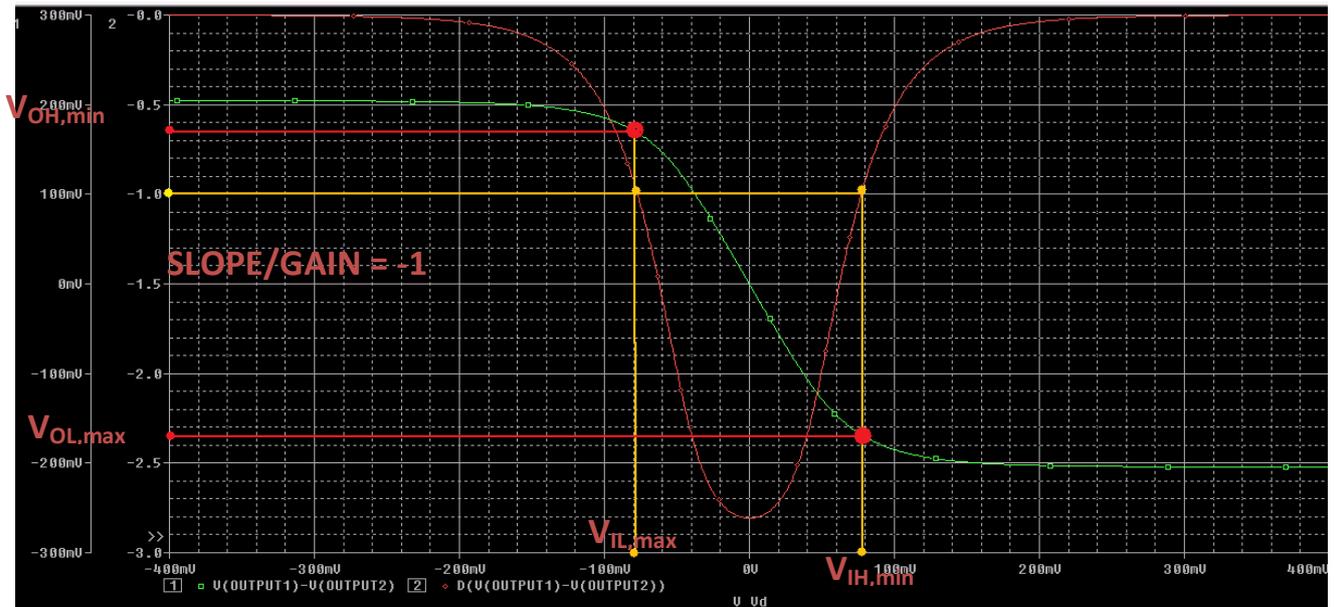


Figure 38

As a result , the static noise margins , which are defined as the maximum values of dc disturbance such as offsets and mismatches due to processing and variations in operating conditions in the input of the inverter , that can be tolerated by the inverter before its changing state [33]-[34] , are:

$$NM_H = |V_{OH,min} | - |V_{IL,max} | \approx NM_L = |V_{OL,max} | - |V_{IH,min} | \Rightarrow$$

$$\Rightarrow NM_H \approx NM_L = V_{SWING} \cdot g \left(\frac{V_{SWING}}{n \cdot U_T} \right)$$

Where $g(x) = 1 - \frac{1}{x} \cdot [1.15 + \ln(2x - 2)] \approx 0.08 \cdot (x - 1)$, depicted in

figure 39

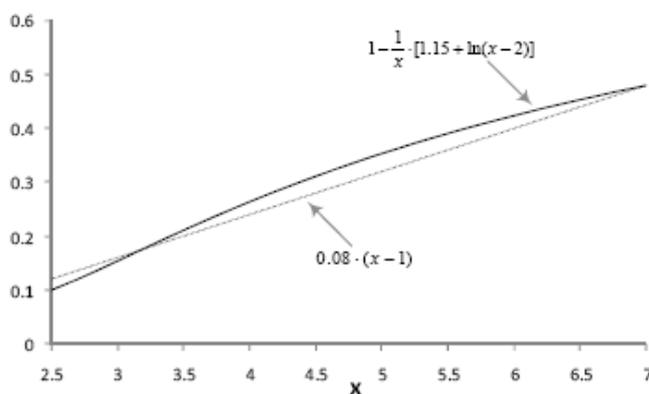


Figure 39

Therefore , for $V_{SWING} = 0.2V$, $n \approx 1.25$, then

$$g \left(\frac{0.2}{1.25 * 25.8 * 10^{-3}} \right) = g(6.2) = 1 - \frac{1}{6.2} \cdot [1.15 + \ln(12.4 - 2)] \approx 1 - 0.56 = 0.44$$

And $NM_H \approx NM_L \approx 0.2 * 0.44 = 88mV$

The static noise margins for the inverter can also be represented graphically on its DC characteristic through the so called “butterfly curves”. The SNMs are then defined as the length of the side of the largest square that can be embedded inside the lobes of the butterfly curve[33]-[35]. This is shown in figure 40, which uses two identical copies of the same SCL inverter connected in a latch structure, where the output of the first inverter is connected to the input of the second inverter and vice versa. For our inverter, this graph can be produced by plotting V_{in} vs V_{out} and at the same graph V_{out} vs V_{in} , as shown in figure 41, which is produced with the use of MATLAB, by combining the results of figure 42 in the same plot.

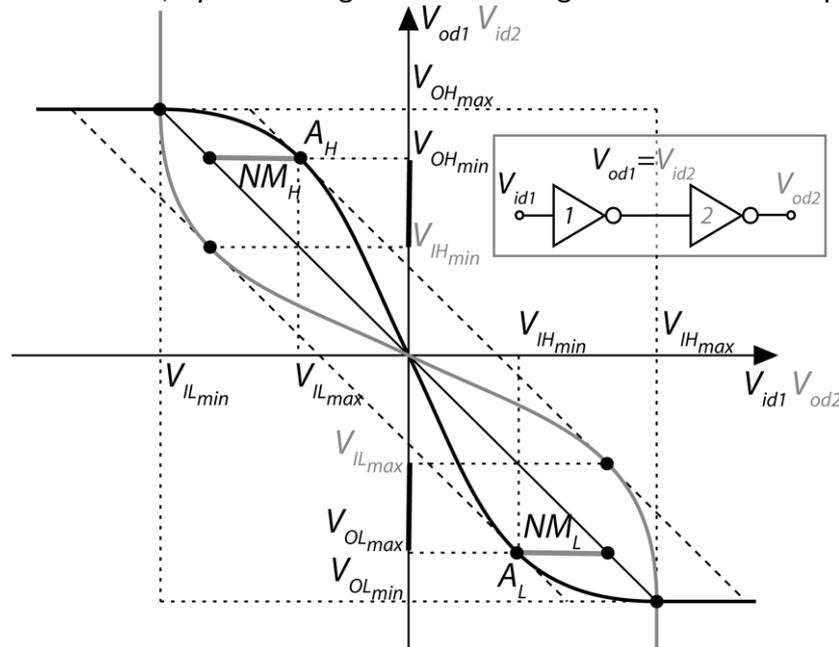


Figure 40 – Picture from [5]

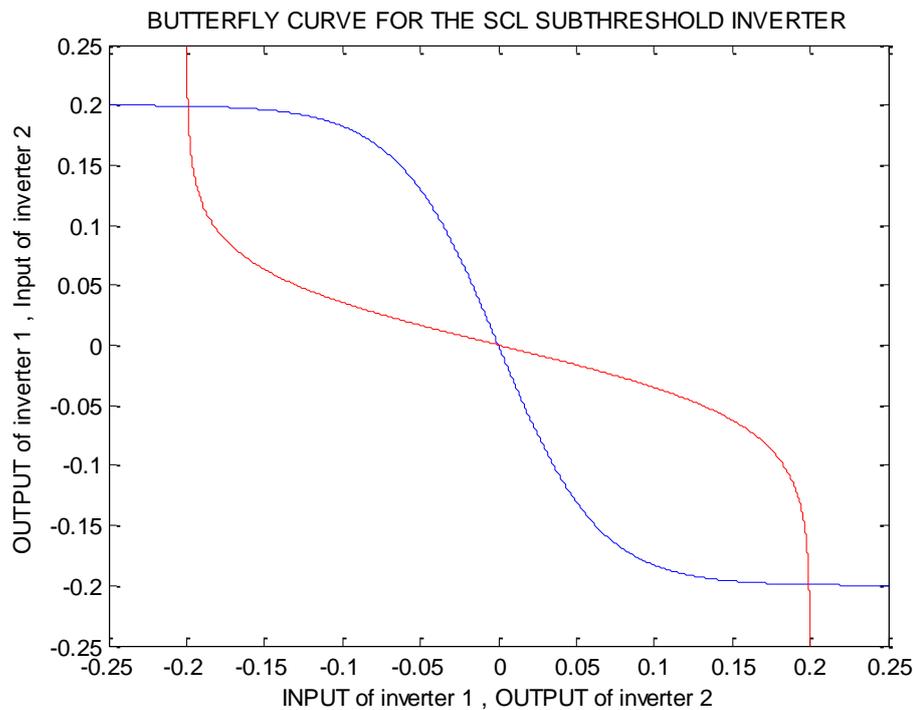


Figure 41

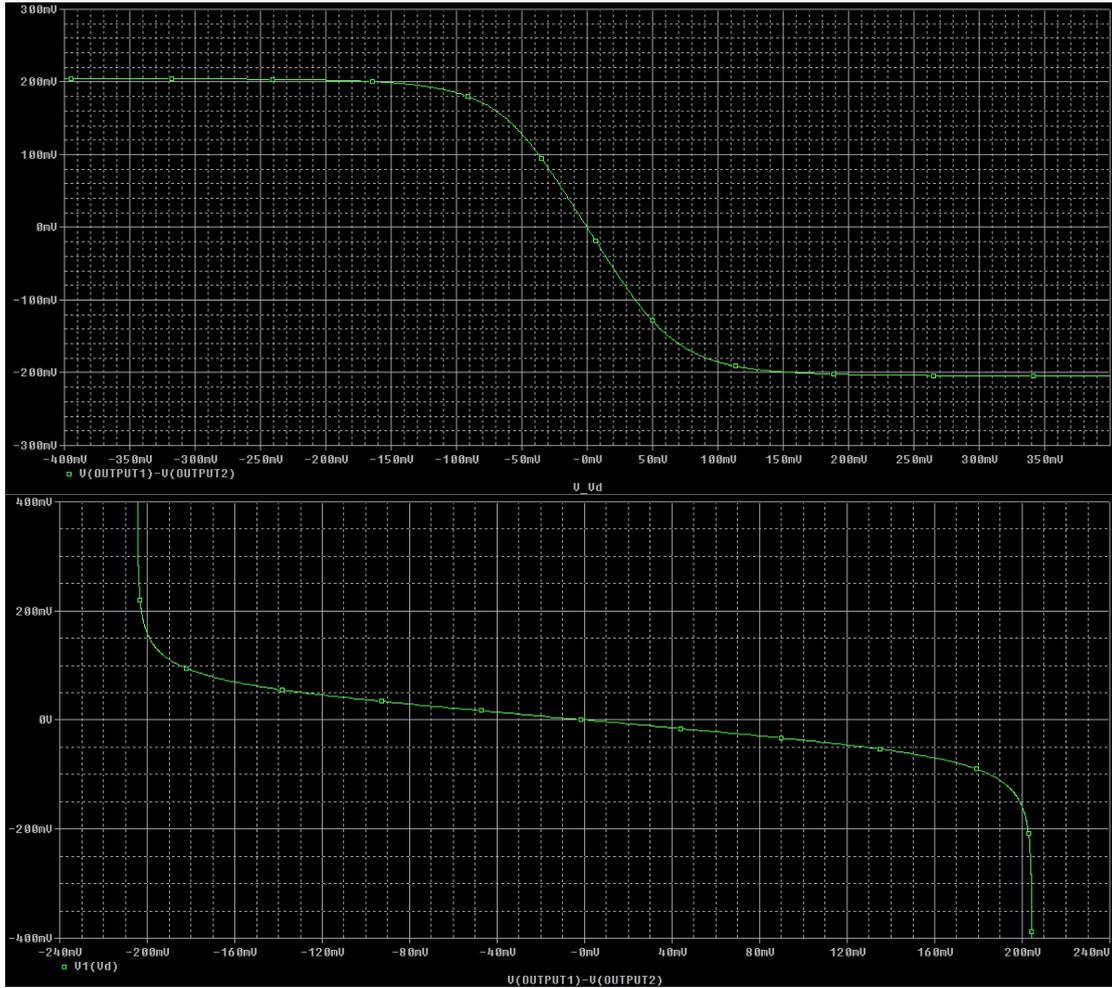


Figure 42

3. MEASUREMENTS

I. MEASUREMENTS ON THE DRAIN-BULK CONNECTED PMOS USED AS LOADS

In order to prove the previously mentioned characteristics of the Bulk – Drain connected loads, measurements have been made on a 180nm wafer through IC-CAP, using prober Cascade Microtech SUMMIT 10600 and HP 4145A Semiconductor Parameter Analyzer. Both structures, conventional PMOS and the Bulk – Drain connected one have been measured with the Source-Drain voltage V_{SD} ranging from 0V to -280mV with a stepsize of -10mV, and the Source-Gate voltage V_{SG} ranging from -250mV to -350mV with a stepsize of -25mV.

The drain current of the PMOS transistors biased in the Weak Inversion region, is given by the EKV model [3]-[8]-[10] as:

$$I_{SD} = I_0 \cdot e^{\frac{V_{BG} - V_{TO}}{n_p U_T}} \left(e^{\frac{-V_{BS}}{U_T}} - e^{\frac{-V_{BD}}{U_T}} \right), \quad \text{where:} \quad I_0 = 2 \cdot n_p \cdot \mu \cdot C_{ox} \cdot \frac{W}{L_e} U_T^2$$

For the conventional device where $V_{BS} = 0$, the relation becomes:

$$I_{SD} = I_0 \cdot e^{\frac{V_{BG} - V_{TO}}{n_p U_T}} \left(1 - e^{\frac{-V_{BD}}{U_T}} \right)$$

Whereas for the Bulk-Drain connected device where $V_{BD} = 0$, the relation becomes:

$$I_{SD} = I_0 \cdot e^{\frac{V_{BG} - V_{TO}}{n_p U_T}} \left(e^{\frac{-V_{BS}}{U_T}} - 1 \right) \quad \Rightarrow \quad I_{SD} = I_0 \cdot e^{\frac{V_{DG} - V_{TO}}{n_p U_T}} \left(e^{\frac{-V_{SD}}{U_T}} - 1 \right)$$

The output small signal resistance and the output conductance of the conventional and the Bulk-Drain connected PMOS are given as :

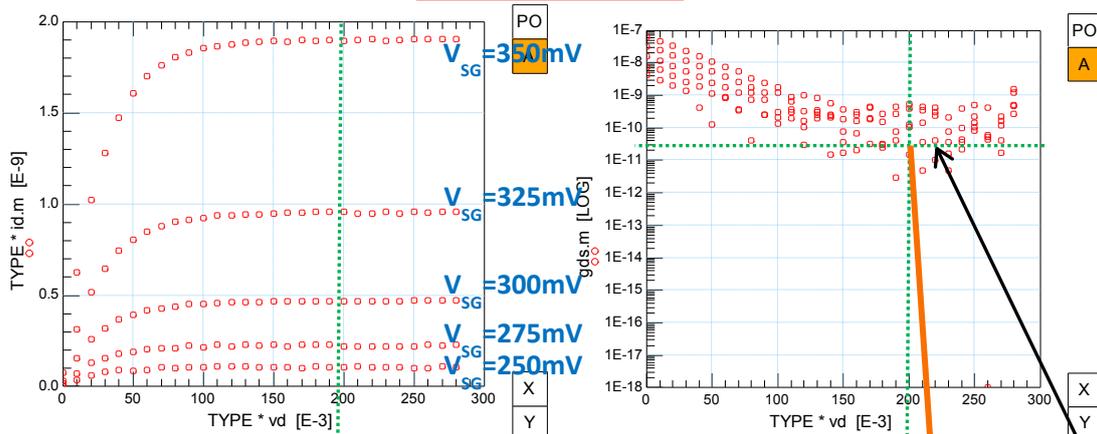
$$g_{SD} = \frac{\partial I_{SD}}{\partial V_{SD}} \quad , \quad R_{SD} = \frac{1}{g_{SD}} = \left(\frac{\partial I_{SD}}{\partial V_{SD}} \right)^{-1}$$

In the Bulk-Source connected (conventional) PMOS, its Bulk and Source are tied to Ground (0V) while its negative Drain voltage changes for various negative values of the Gate voltage. In the Bulk-Drain connected PMOS, its Bulk voltage is synchronized to the negative Drain voltage changing for various negative values of the Gate voltage, while its Source is tied to Ground (0V).

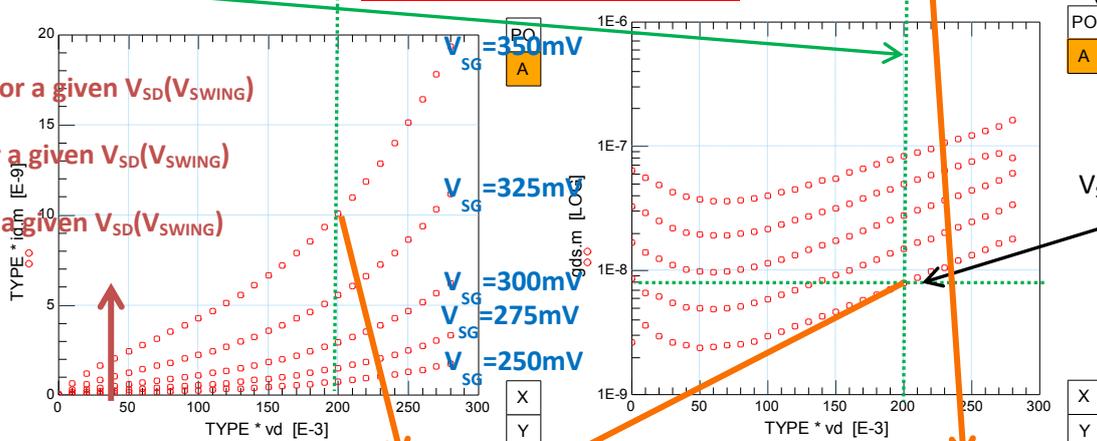
The results for the Drain Current and the Output Conductance, are given as follows:

-----W/L=10μm/10μm-----

Conventional PMOS



Bulk-Drain connected PMOS



BULK-DRAIN CONNECTED PMOS

For $V_{SWING}=200mV$ and $V_{SG}=250mV$ then
 $I_{SD}=775.5pA$, $g_{sd} \approx 8nS$ and $R_{sd}=125M\Omega$
 Therefore, $I_{SS}=1.6nA$

CONVENTIONAL PMOS

For $V_{SWING}=200mV$ and $V_{SG}=250mV$ then
 $I_{SD}=108.6pA$, $g_{sd}=125pS$ and $R_{sd}=8G\Omega$

As we want to maintain the desired output voltage swing V_{SWING} at very low bias current levels, it is necessary to increase the load resistance value in inverse

proportion to reducing tail bias current according to the relation $R_L = \frac{V_{SWING}}{I_{SS}}$. This

resistance should be controlled very accurately based on the I_{SS} value. As a result, a well controlled high resistivity load device with a very small area is required.

As seen, in the case of the conventional PMOS load, the linear region of their I_{SD} - V_{SD} characteristic is restricted to the triode region of operation, whereas in the case of the bulk-drain connected PMOS load, the linear region extends to the saturation region which begins at $V_{SD} > 4 \cdot U_T = 104\text{mV}$.

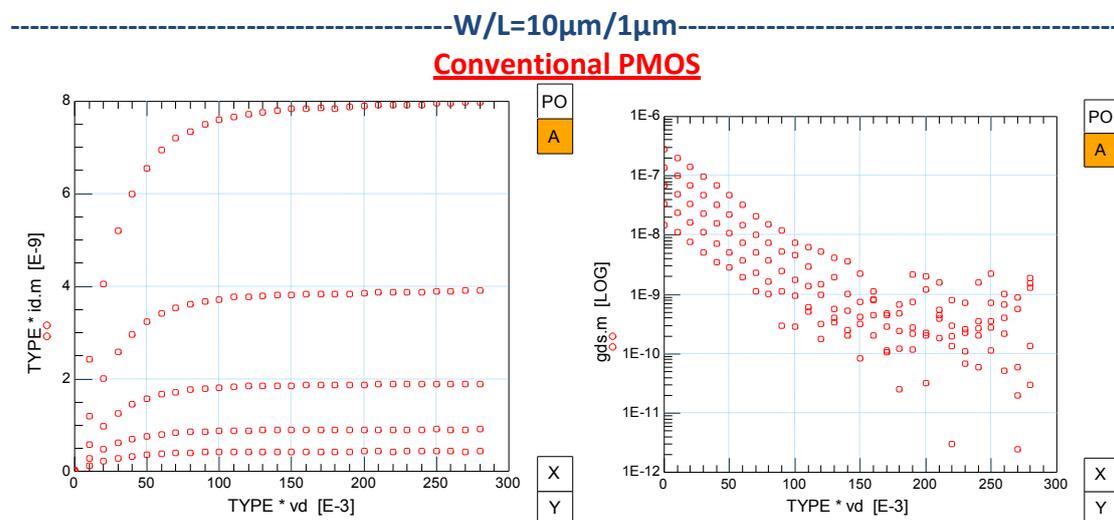
Hence, in the first case the PMOS transistor acts as a current source with almost infinite output impedance even for deep-submicron devices. Thus, for the range of resistivity we want, conventional PMOS devices biased in triode region cannot be utilized, since the required channel length of the transistor in order to have a V_{SWING} larger than at least 150mV-200mV so that we can obtain an adequate Noise Margin, would be impractically large. This can be proven using the following relations provided by [22]-[55]:

$$V_{SD,sat} = 2 \cdot U_T \cdot \sqrt{IC} + 4 \cdot U_T \quad \text{with} \quad IC = i_f = \frac{I_D}{I_{spec}} = \frac{I_D}{I_0 \cdot \frac{W}{L}} = \frac{I_D}{(2 \cdot n \cdot \mu \cdot C_{ox} \cdot U_T^2) \cdot \frac{W}{L}}$$

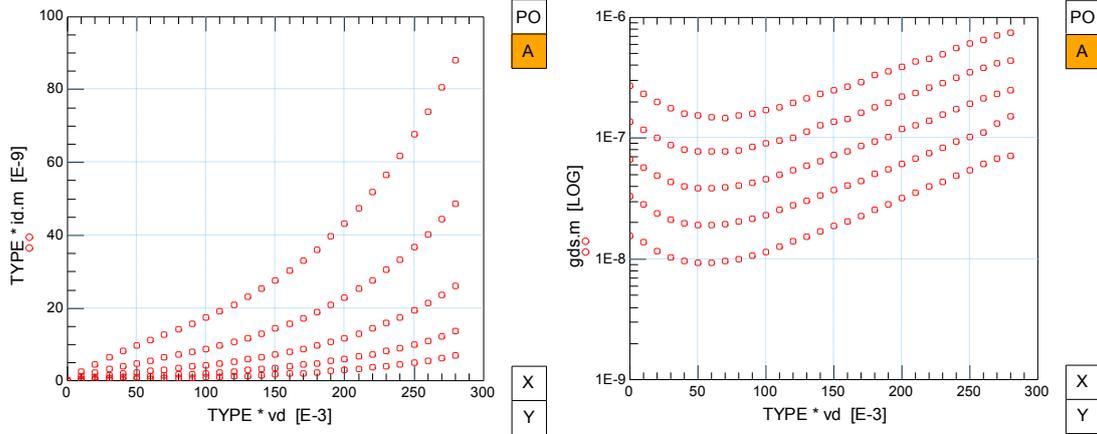
Thus, for constant W , I_D , I_0 , then $V_{SD,sat} \propto \sqrt{L}$

In the second case, the PMOS transistor acts as a finite and easily controllable high valued resistance. Thus, it is possible to implement a very high resistivity load device using a single minimum size PMOS device.

The value of V_{SG} is tuned through the replica bias circuit, which controls the resistance of the load devices and hence adjusts the output voltage swing V_{SWING} with respect to the tail bias current I_{SS} [16].

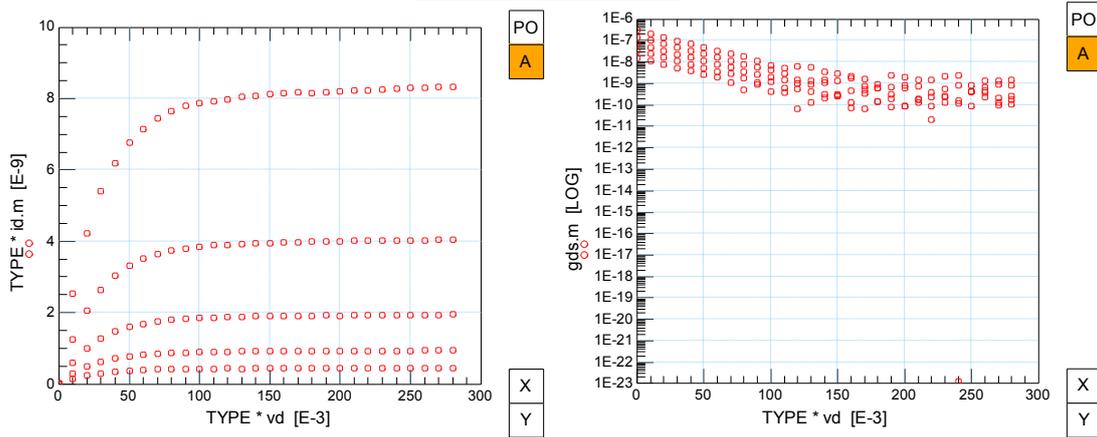


Bulk-Drain connected PMOS

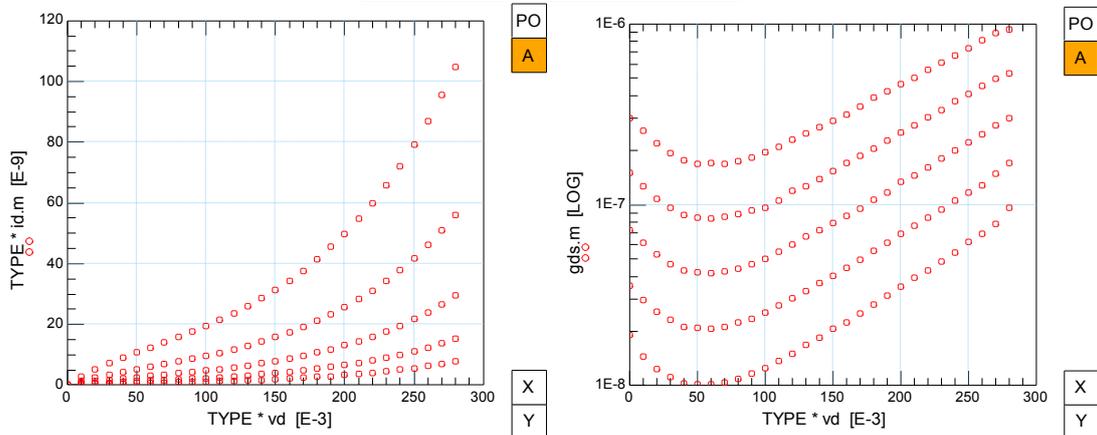


-----W/L=10 μ m/0.5 μ m-----

Conventional PMOS

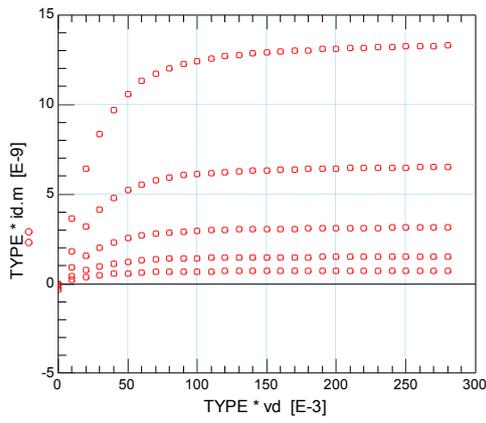


Bulk-Drain connected PMOS



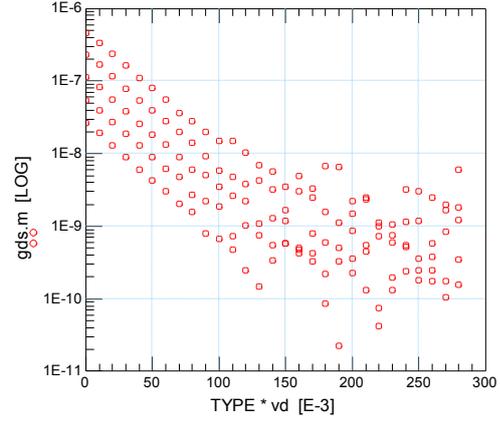
-----W/L=10 μ m/0.3 μ m-----

Conventional PMOS



PO
A

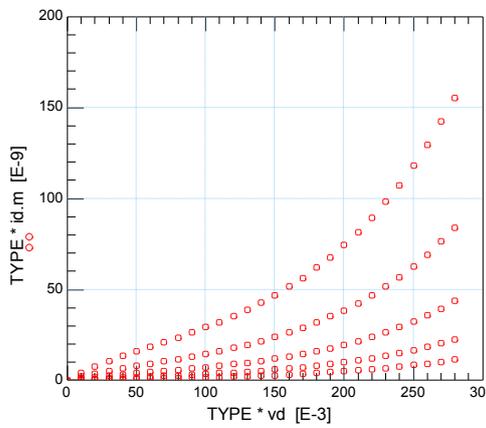
X
Y



PO
A

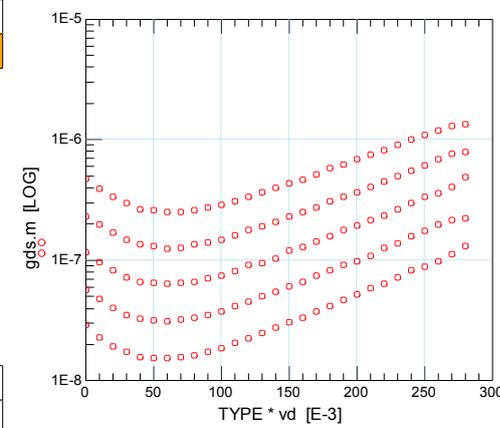
X
Y

Bulk-Drain connected PMOS



PO
A

X
Y

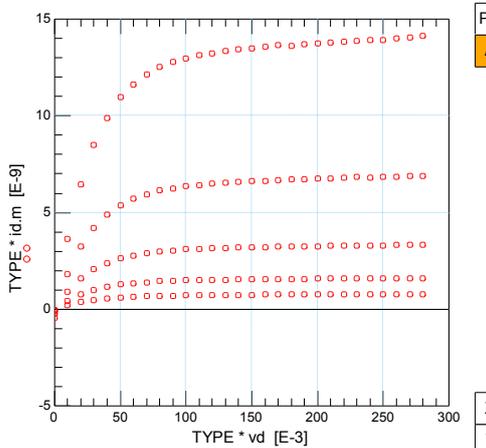


PO
A

X
Y

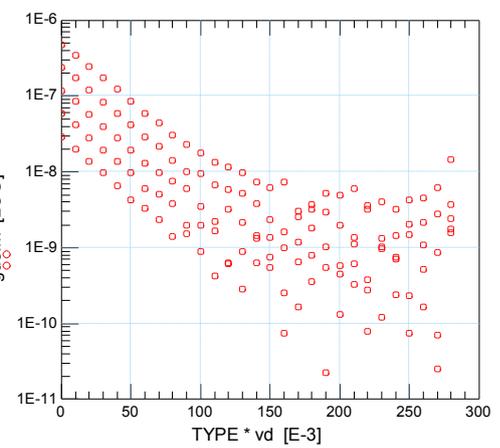
-----W/L=10 μ m/0.24 μ m-----

Conventional PMOS



PO
A

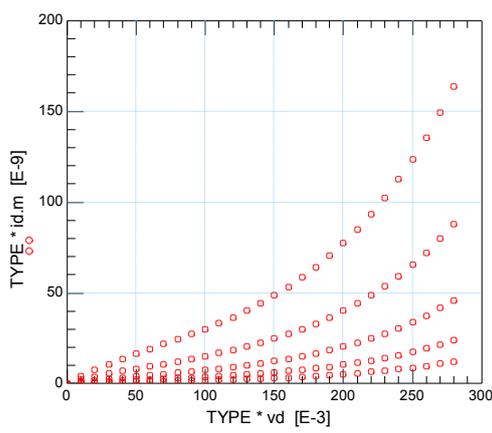
X
Y



PO
A

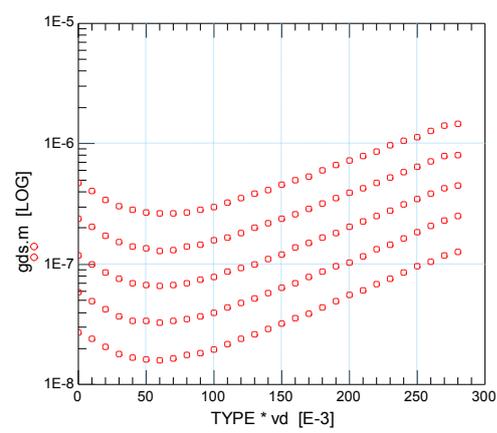
X
Y

Bulk-Drain connected PMOS



PO
A

X
Y

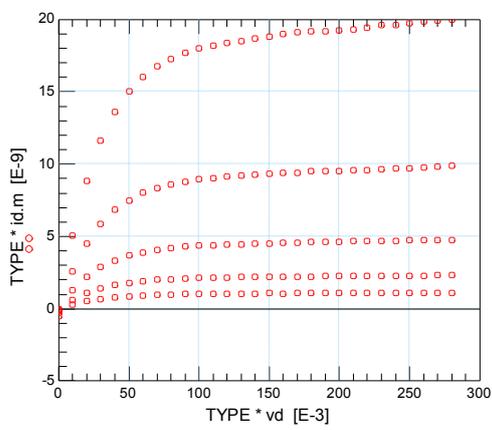


PO
A

X
Y

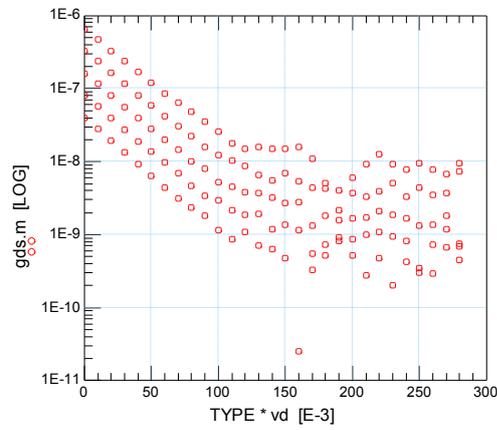
-----W/L=10 μ m/0.22 μ m-----

Conventional PMOS



PO
A

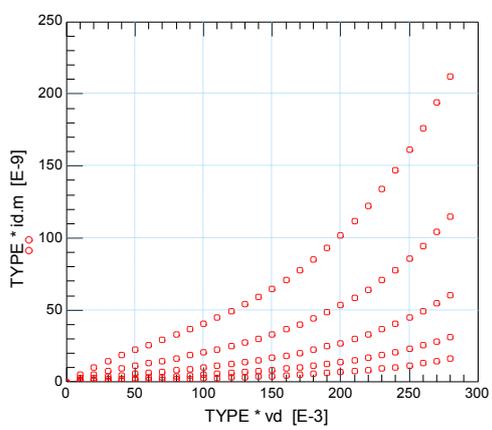
X
Y



PO
A

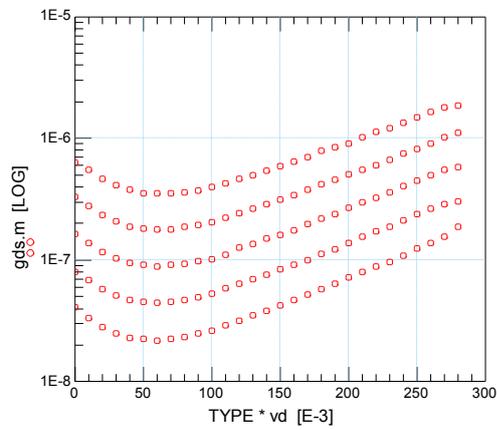
X
Y

Bulk-Drain connected PMOS



PO
A

X
Y

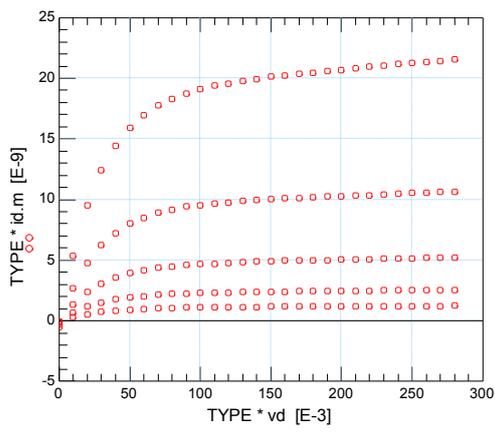


PO
A

X
Y

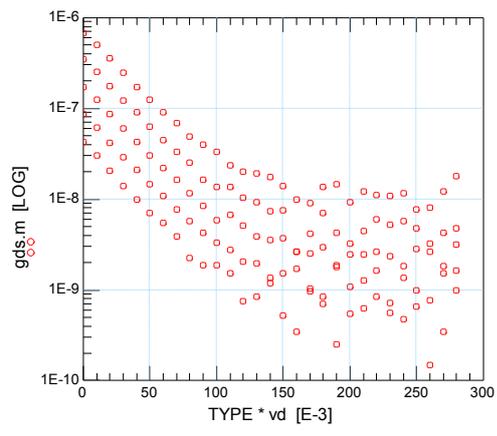
-----W/L=10 μ m/0.2 μ m-----

Conventional PMOS



PO
A

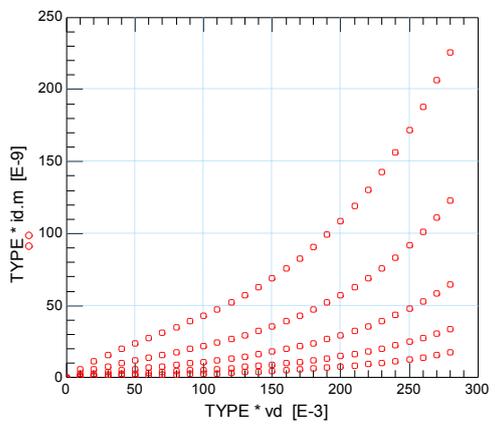
X
Y



PO
A

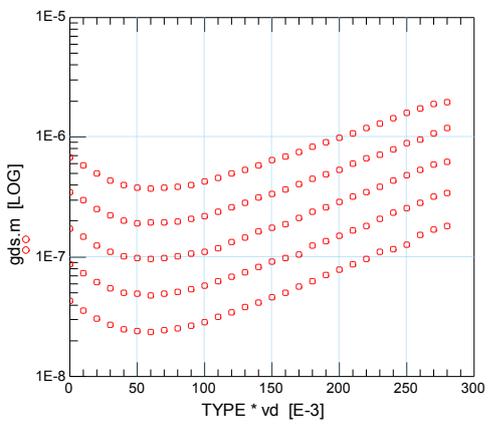
X
Y

Bulk-Drain connected PMOS



PO
A

X
Y

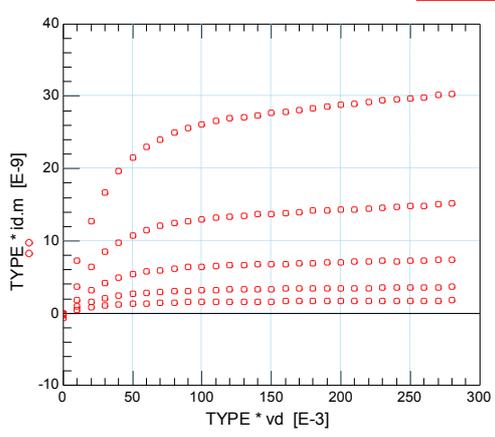


PO
A

X
Y

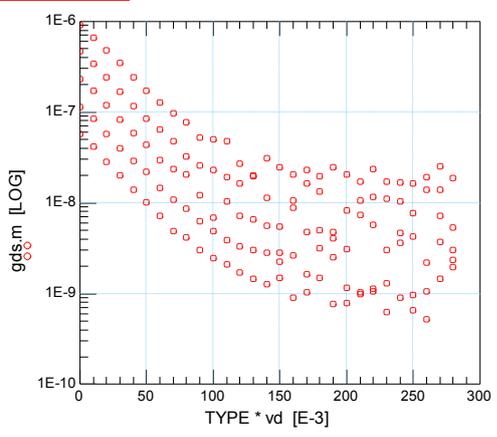
-----W/L=10 μ m/0.18 μ m-----

Conventional PMOS



PO
A

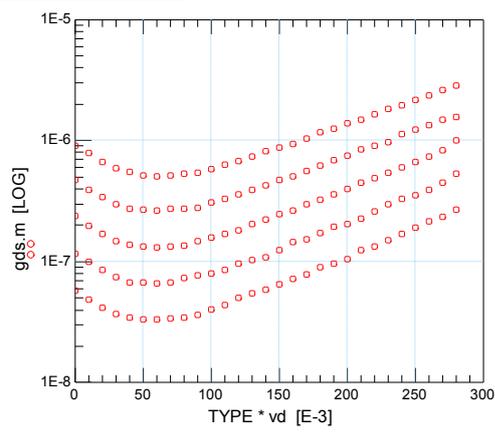
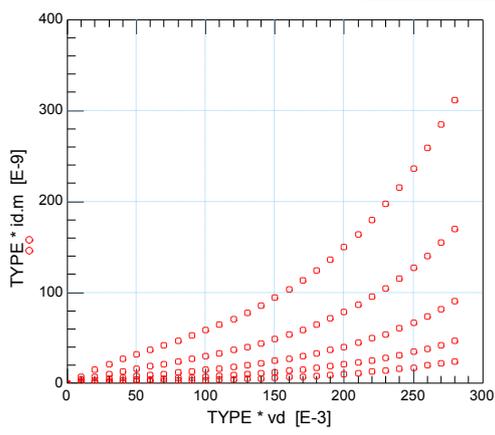
X
Y



PO
A

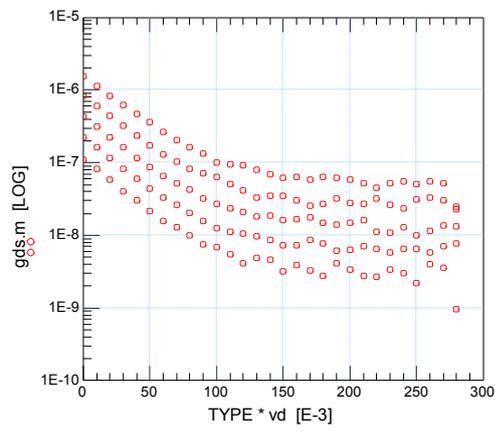
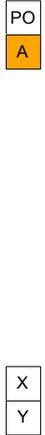
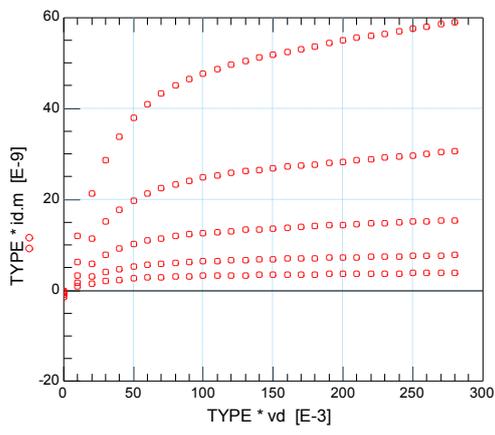
X
Y

Bulk-Drain connected PMOS

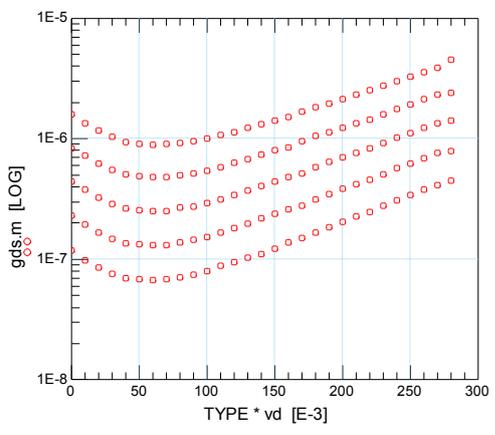
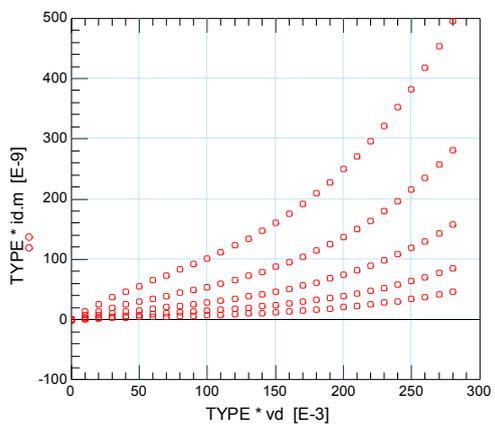


-----W/L=10 μ m/0.16 μ m-----

Conventional PMOS

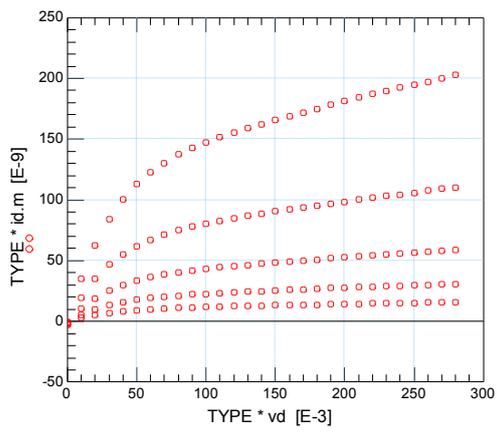


Bulk-Drain connected PMOS



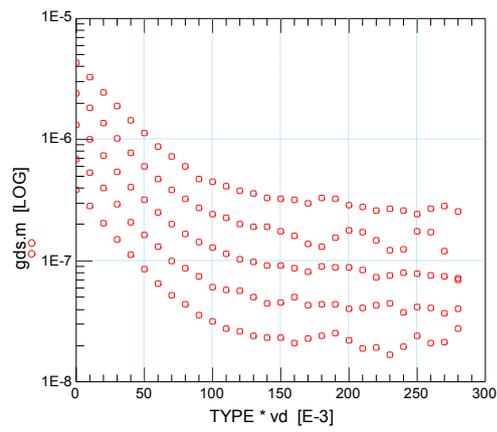
-----W/L=10 μ m/0.14 μ m-----

Conventional PMOS



PO
A

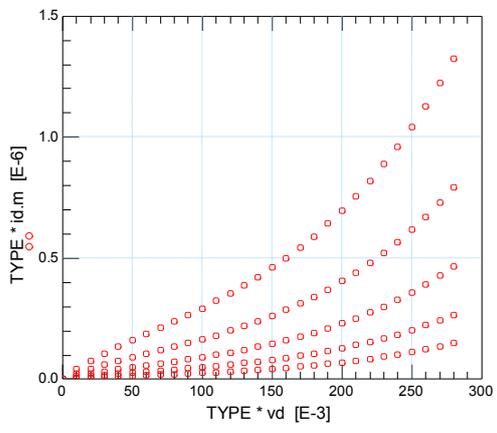
X
Y



PO
A

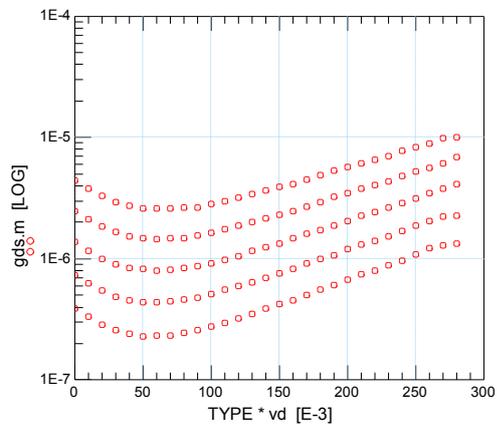
X
Y

Bulk-Drain connected PMOS



PO
A

X
Y

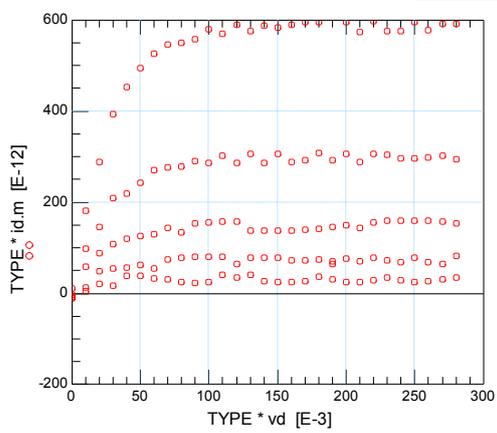


PO
A

X
Y

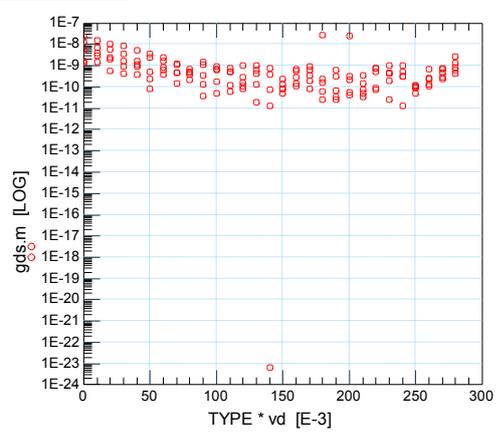
-----W/L=2 μ m/10 μ m-----

Conventional PMOS



PO
A

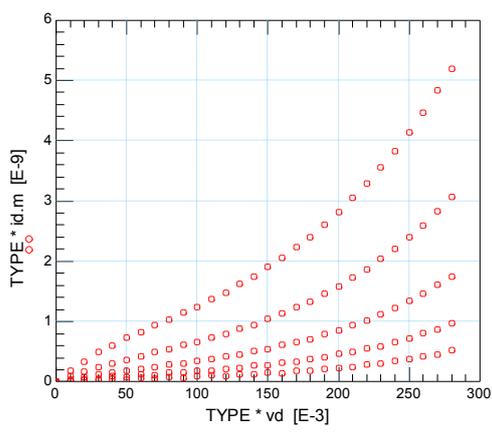
X
Y



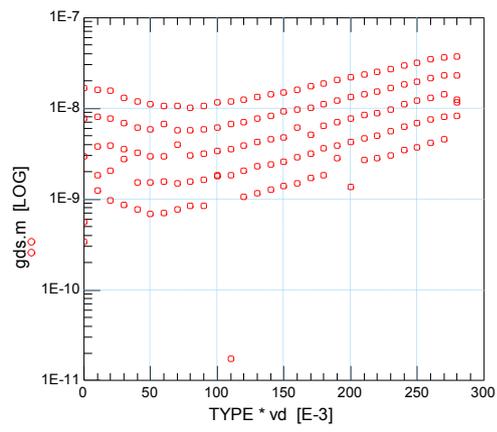
PO
A

X
Y

Bulk-Drain connected PMOS



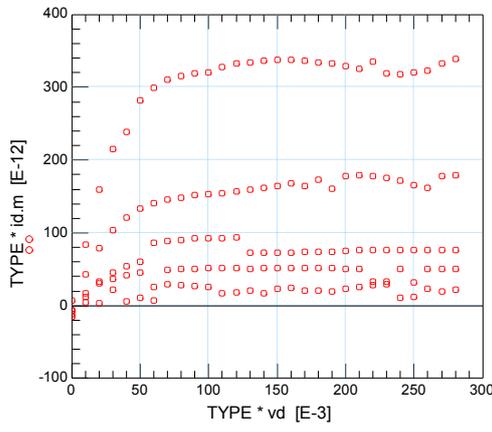
PO
A
X
Y



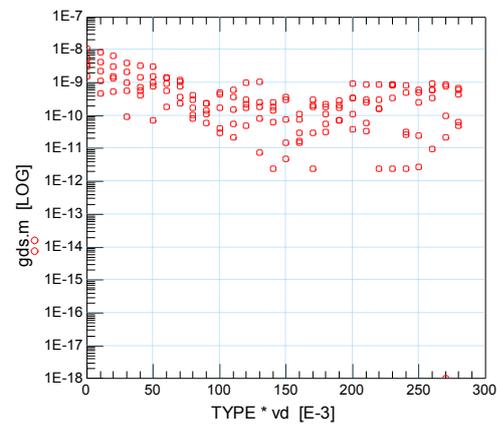
PO
A
X
Y

-----W/L=1μm/10μm-----

Conventional PMOS

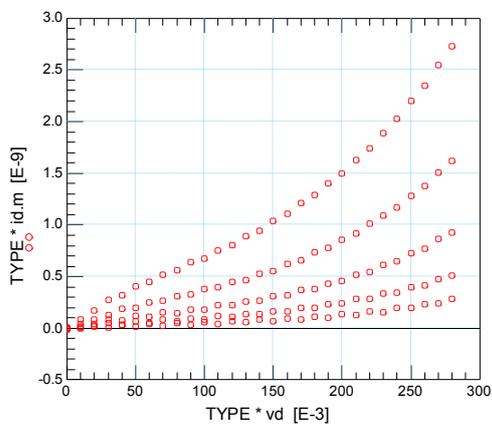


PO
A
X
Y

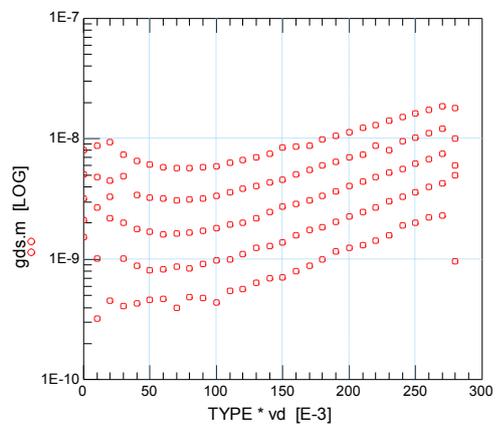


PO
A
X
Y

Bulk-Drain connected PMOS



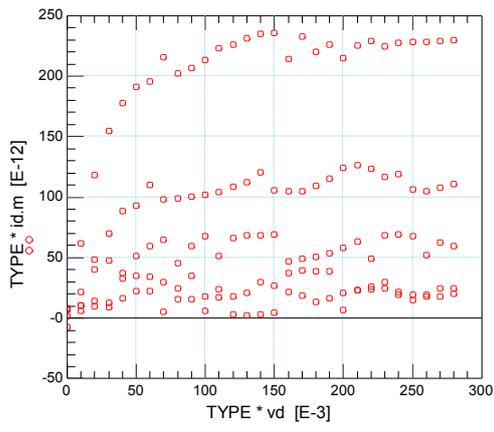
PO
A
X
Y



PO
A
X
Y

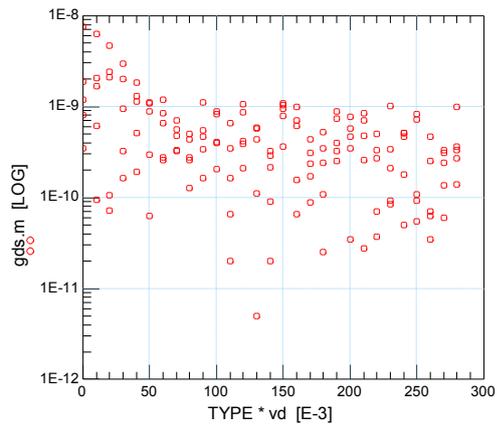
-----W/L=0.5μm/10μm-----

Conventional PMOS



PO
A

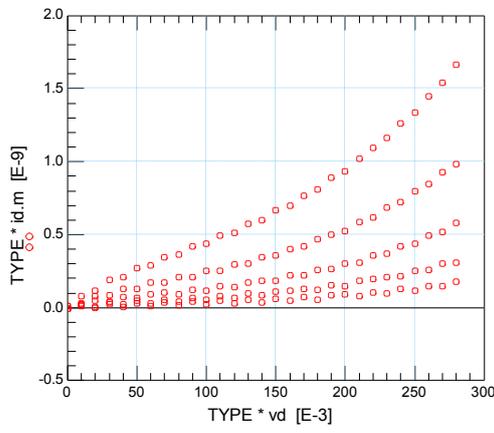
X
Y



PO
A

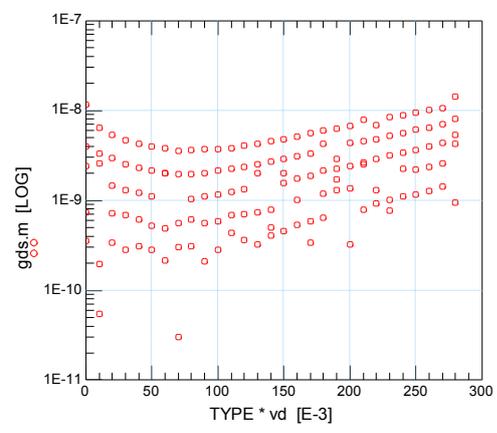
X
Y

Bulk-Drain connected PMOS



PO
A

X
Y

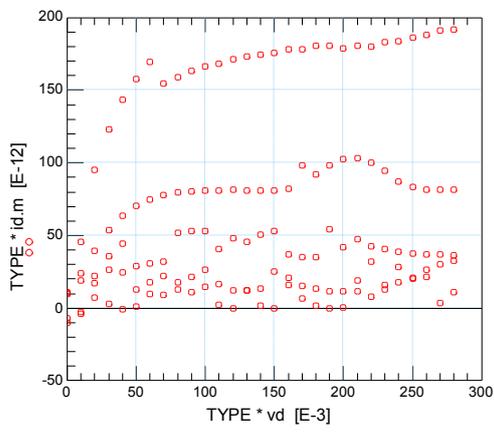


PO
A

X
Y

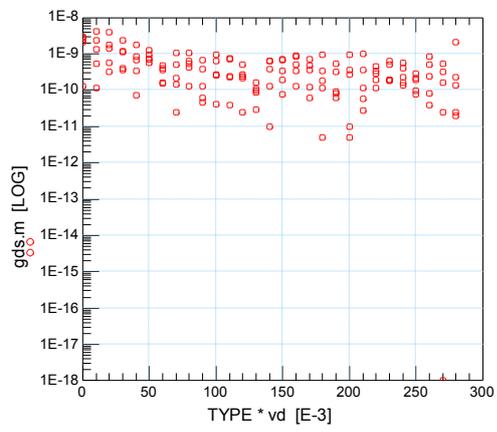
-----W/L=0.3μm/10μm-----

Conventional PMOS



PO
A

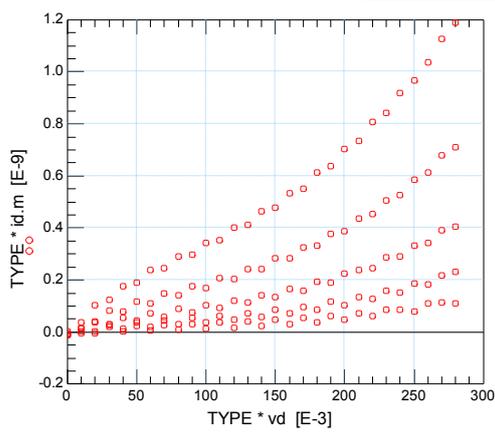
X
Y



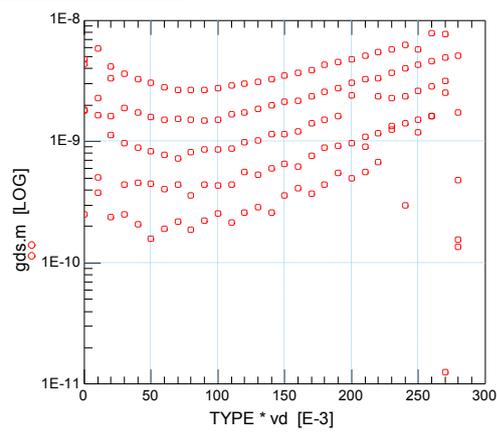
PO
A

X
Y

Bulk-Drain connected PMOS



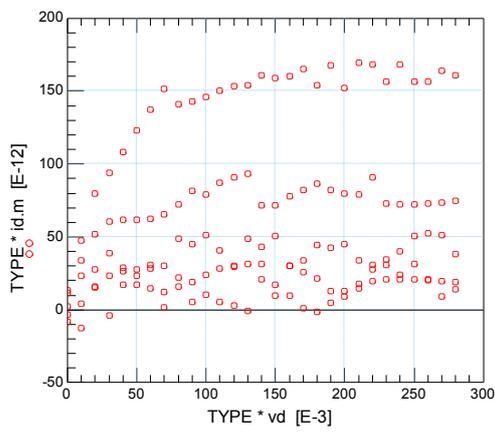
PO
A
X
Y



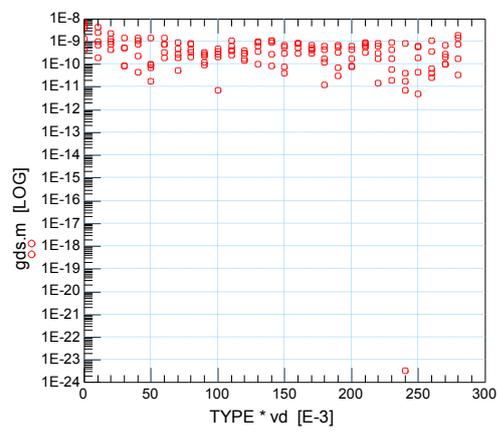
PO
A
X
Y

-----W/L=0.25μm/10μm-----

Conventional PMOS

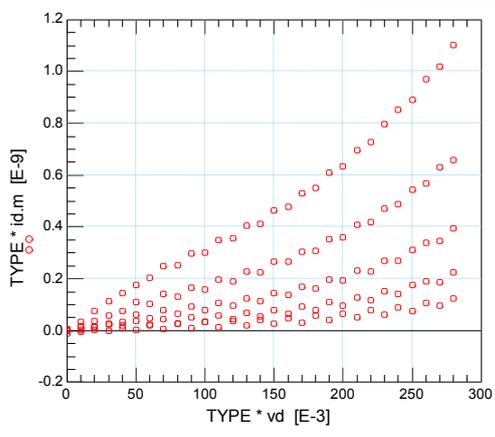


PO
A
X
Y

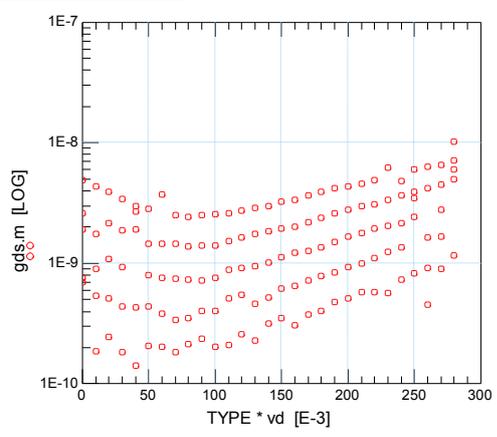


PO
A
X
Y

Bulk-Drain connected PMOS



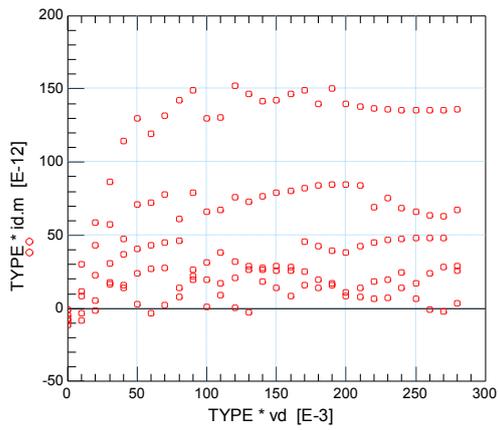
PO
A
X
Y



PO
A
X
Y

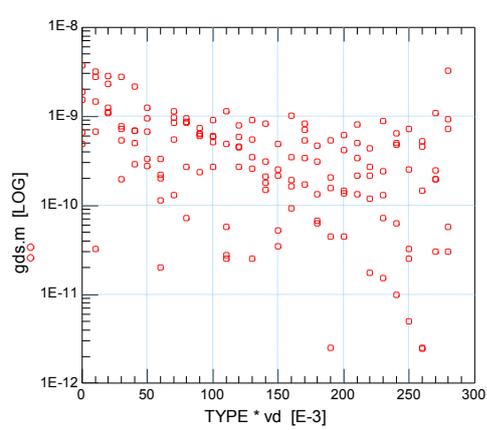
-----W/L=0.22μm/10μm-----

Conventional PMOS



PO
A

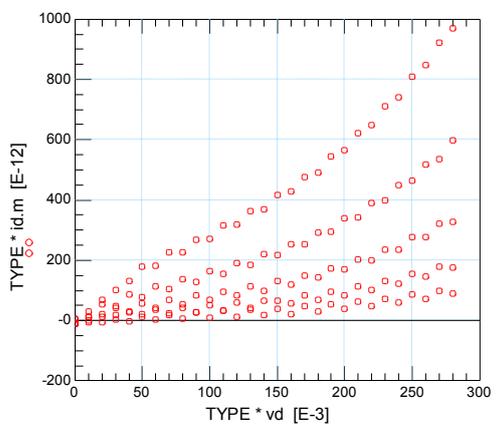
X
Y



PO
A

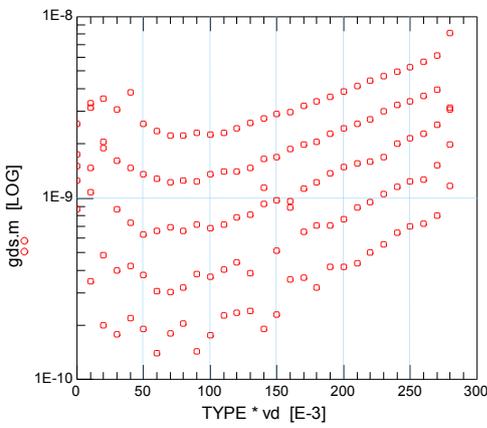
X
Y

Bulk-Drain connected PMOS



PO
A

X
Y

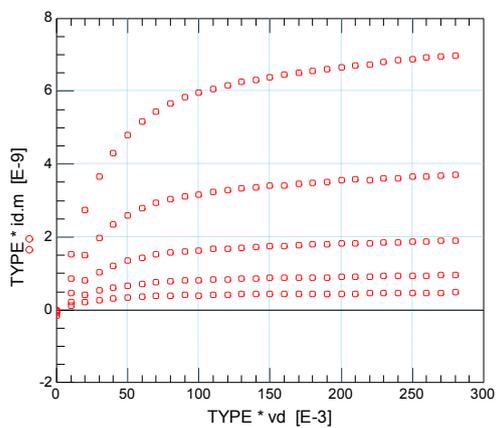


PO
A

X
Y

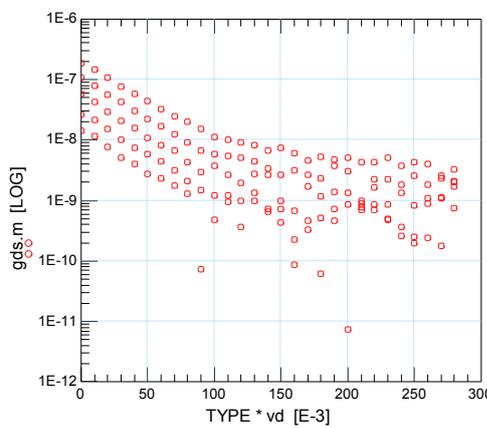
-----W/L=0.22μm/0.18μm-----

Conventional PMOS



PO
A

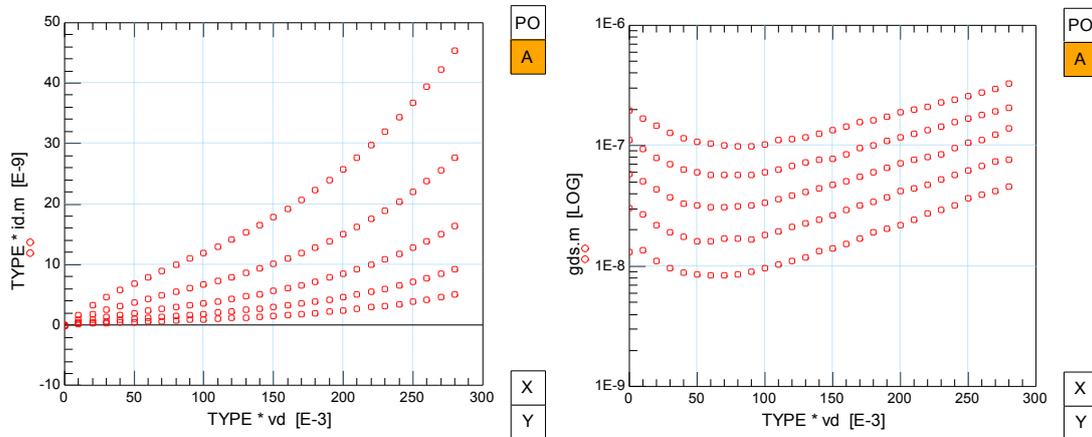
X
Y



PO
A

X
Y

Bulk-Drain connected PMOS



4. IMPLEMENTATION OF THE SCL SRAM

I. ABOUT SRAMs

SRAM is a volatile memory , which means that it retains its data as long as power is applied. Its name fully expanded is Static Random Access Memory , with “Static” meaning that it uses some form of feedback to maintain its state without any need of refreshing the stored charge at the node capacitances of its cells , while “Random Access” means that it can be accessed with an address and that it has a latency independent of the address[32].

Embedded SRAM , is definitely the workhorse for on-chip data storage owing to its robust operation , high speed and low power consumption relative to other options[45].

One of the most important design objectives concerning SRAM , is cell size minimization. A smaller cell allows the number of bits per unit area to be increased and thus , decreases cost per bit , while reduced cell area can indirectly improve the speed and power consumption due to the reduction of the associated cell capacitances[40].

Another equally important consideration is cell stability , which , expressed by the Static Noise Margin(SNM) , determines the Soft Error Rate(SER) and the sensitivity of the memory to process tolerances and operating conditions[33].

SER which is quite an issue for deeply voltage scaled SRAM , refers to errors occurred when an alpha particle or cosmic ray strikes a memory node and causes data loss. It is a concern for sub-threshold memory and for modern SRAM in general , as each bitcell uses less charge to store its data owing to smaller capacitance and lower voltage. Therefore , the bits are more easily upset by cosmic rays and alpha particles[41]-[45].

M6. In the first case , M2 opposes this operation , thus M2 must be weaker than M5 so that \bar{Q} can be pulled low enough. Likewise , in the second case , M4 opposes the operation , thus it should be weaker than M6 so that Q can be pulled low enough. This is called **writability**[32].

To ensure both **read stability** and **writability** , the transistors of the SRAM cell must satisfy ratio constraints. The drivers M1 , M3 must be the strongest ones , while the access transistors M5 , M6 are of intermediate strength and the loads M2 , M4 must be weak.

Thus , keeping the channel length minimum so as to achieve good layout density , the pulldowns M1 , M3 could have a ratio $W/L=4/1$, the access transistors M5 , M6 a ratio of $W/L=2/1$ and the pullups M2 , M4 a ratio of $W/L=1.5/1.5$ [32].

The **read stability** and **writability** of the cell are quantified by the hold margin , the read margin and the write margin , which are determined by the static noise margin of the cell in its various modes of operation. A cell should have two stable states during hold and read operation , and only one stable state during write.

The static noise margin (SNM) measures how much noise can be applied to the inputs of the two cross-coupled inverters before a stable state is lost (during hold or read) or a second stable state is created (during write)[32].

During hold operation , that is when the cell is holding its state and is being neither read nor written , the access transistors are off and do not affect the circuit behavior. An example of the butterfly curve in that case is shown in figure 44a , and presents two stable states (with one output low and the other high) and one metastable state(with $V_1=V_2$)[32].

During read operation , the bitlines are initially precharged and the access transistors tend to pull the low node up. This is due to the voltage dividing effect across the access transistor M5 (or M6) and the drive transistor M1 (or M3) , and distorts the voltage transfer characteristic causing its lower half for each inverter when its V_{in} is high to pull upwards relative to its original position squashing the lobes of the butterfly plot on one end as shown in figure 44b. As a result , for the 6T SRAM cell , the read margin is smaller than the hold margin[32]-[42]-[49]. Again , we have two stable states (with one output low and the other high) and one metastable state (with $V_1=V_2$). The schematic of the 6T bitcell at the onset of a read access for the case of a 0 stored at node Q is shown in figure 44d.

Finally , during write operation , the access transistors must drive the cell to a monostable condition , which corresponds to a negative SNM as shown in figure 44c. If the butterfly curve maintains bistability , this means that the write attempt has failed[32]-[45].

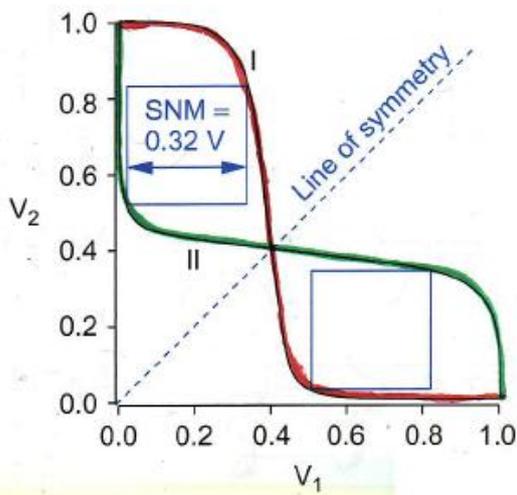


Figure 44a – Picture from [32]

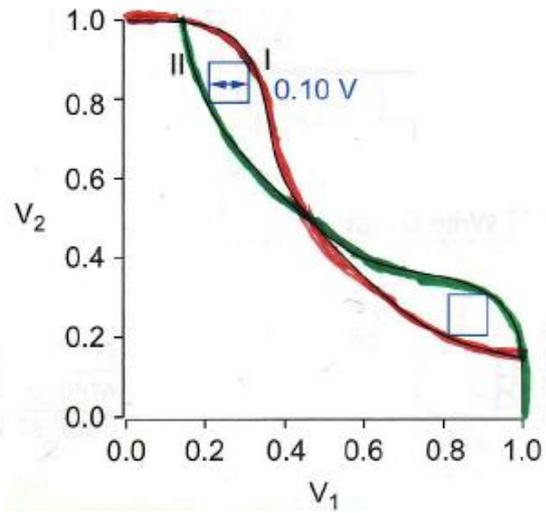


Figure 44b – Picture from [32]

Hold, read and write SNMs decrease with lowering of V_{DD} . Additionally, if the cell is imbalanced, for example due to transistor sizing or process variations, one lobe of the butterfly plot is smaller than the other and in that case, the SNM is the length of the side of the largest square that fits inside the smallest of the two lobes. This indicates that the bitcell is more susceptible to losing one particular data value[45].

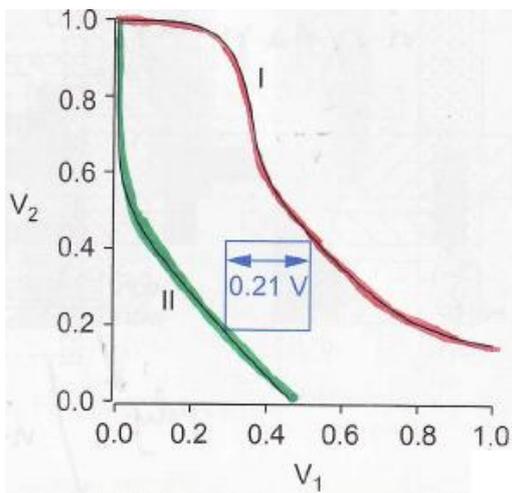


Figure 44c – Picture from [32]

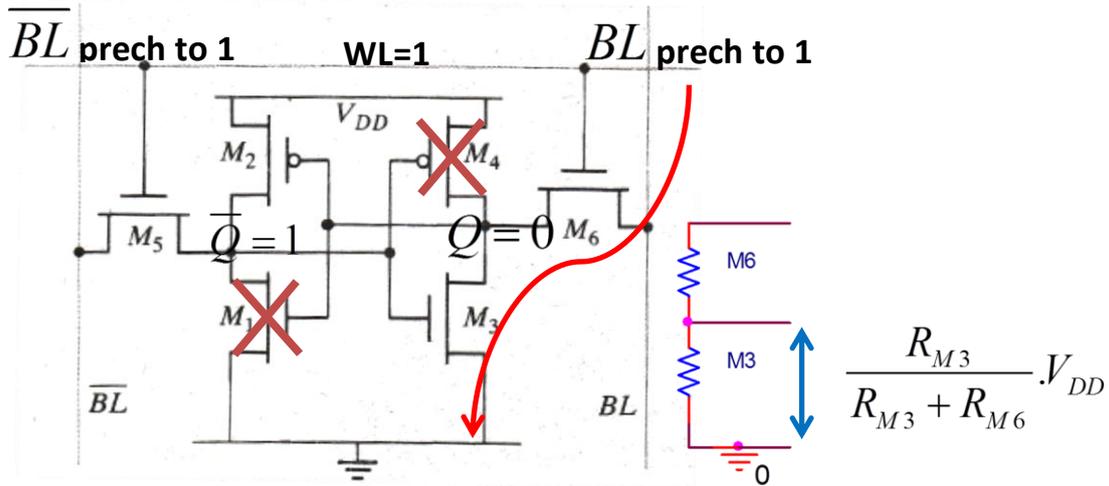


Figure 44d

III. PERIPHERAL CIRCUITRY

A memory array contains 2^n words of 2^m bits each with each bit being stored in a memory cell. The organization of a small memory array containing 16 4-bit words ($n=4$, $m=2$) and using the simplest design with one row per word and one column per bit is shown on the left side of figure 45a. In order to avoid a tall, skinny layout that would be hard to fit in the chip floorplan and slow because of the long vertical wires, the array can be folded into fewer rows of more columns. After folding, each row of the memory contains 2^k words, so the array is organized as 2^{n-k} rows and 2^{m+k} columns or bits. On the right side of figure 45a, the array is organized in a two-fold way with each row of the memory containing $2^k=2$ words ($k=1$) and the array being physically organized as $2^{n-k}=2^{4-1}=8$ rows of $2^{m+k}=2^{2+1}=8$ bits [32].

In order to choose the specific bitcell we want to access, we use the row decoder so as to activate the relevant wordline for a pair of words, and the column decoder to pick the desired word.

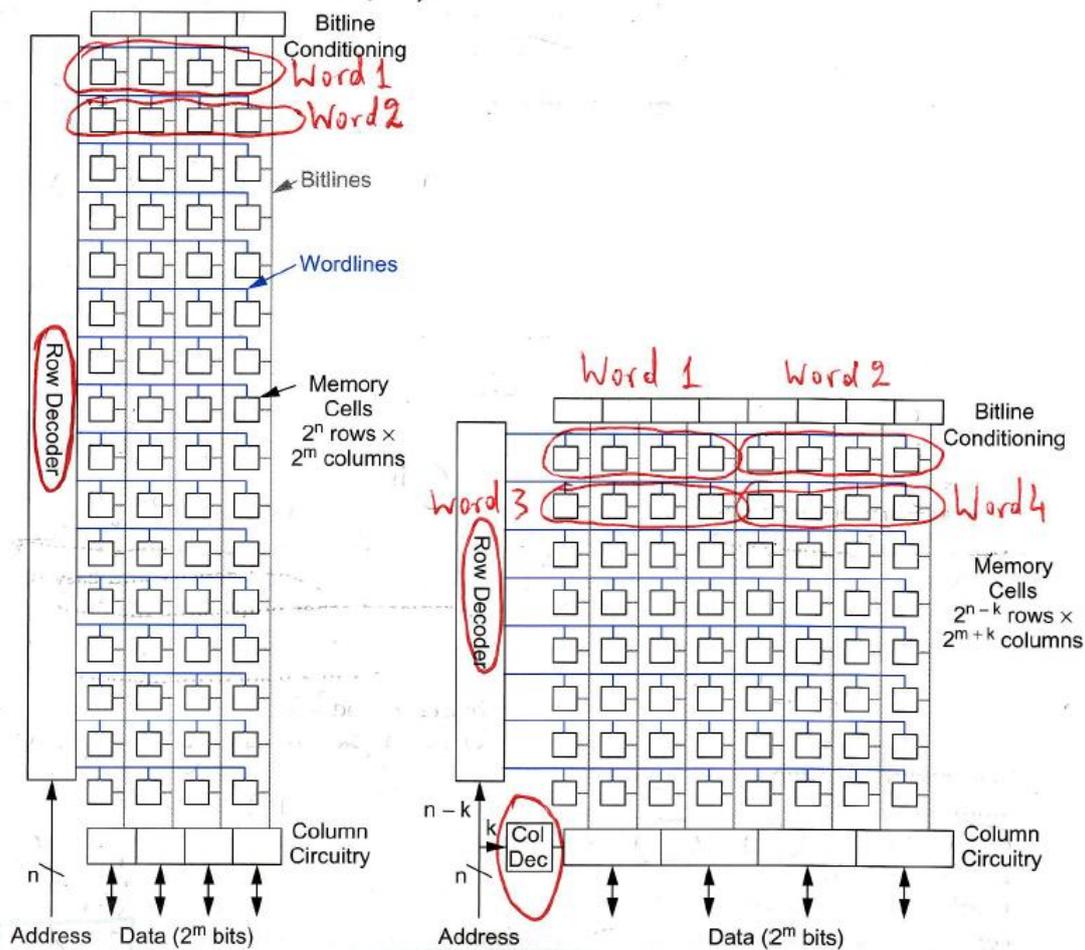


Figure 45a – Picture from [32]

The row decoder in the folded structure of figure 45a is a 3-to-8 decoder. Two possible ways of the implementation of a simple 3-to-8 ($n-k$ to 2^{n-k}) decoder are shown in figure 46. The left implementation is a 3-bit parallel type decoder, whereas the right implementation is a 3-bit tree-type decoder. In both, for each possible input condition, one and only one output signal will be at logic 1[46].

Another view of the SRAM array architecture is shown in figure 45b.

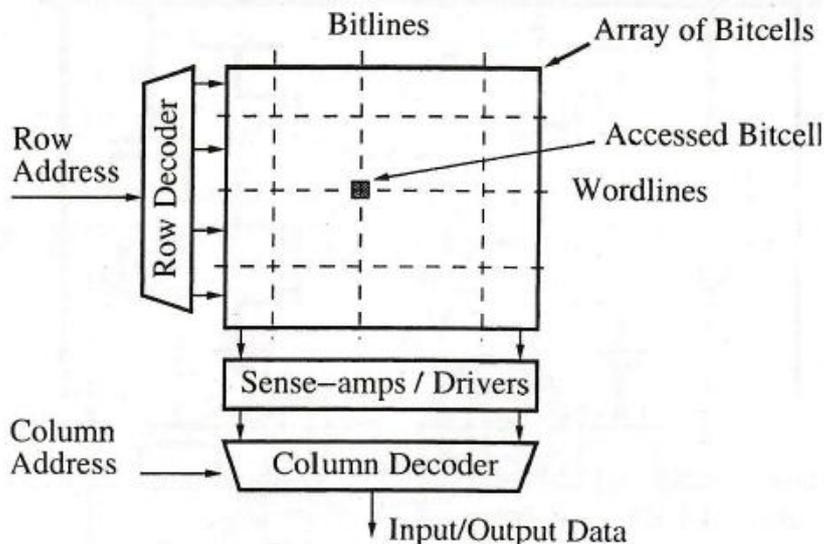


Figure 45b – Picture from [49]

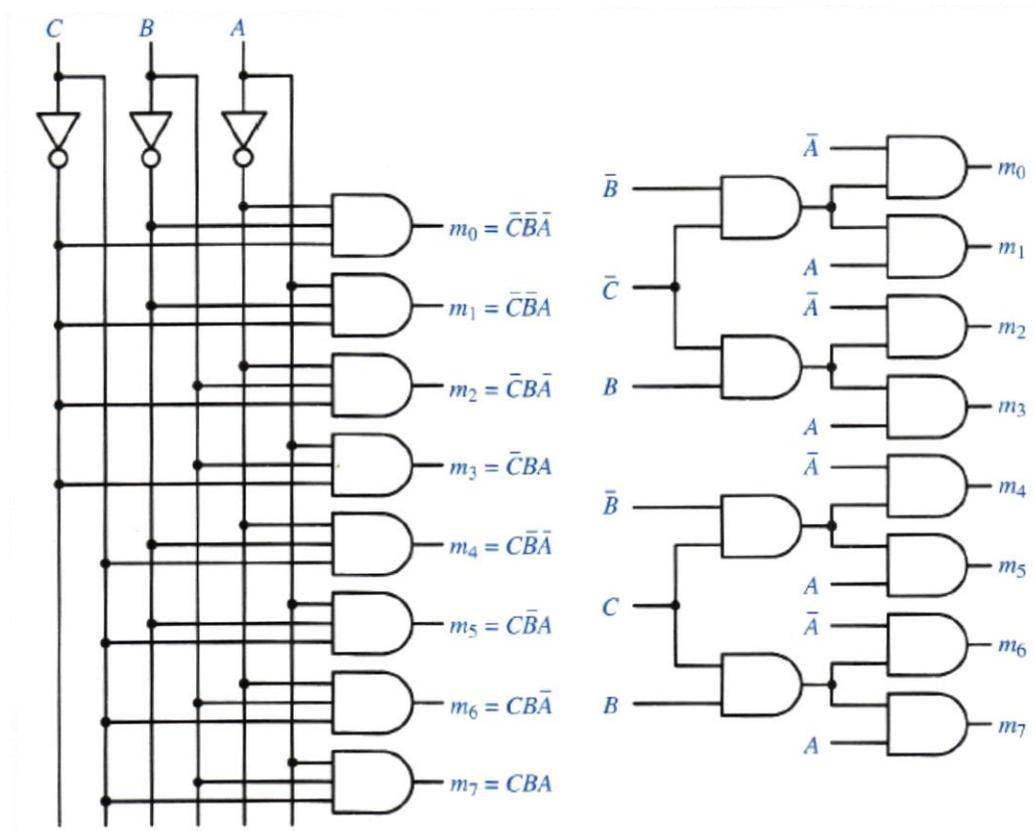


Figure 46 – Picture from [46]

The column decoder in the folded structure of figure 45a controls a multiplexer in the column circuitry to select $2^m=4$ (as $m=2$) bits from the row as the data to access. In general $2^k:1$ column multiplexers may be required to extract 2^m bits from the 2^{m+k} bits of each row. The column multiplexers can either act as their own tree decoder as in figure 47a, or require a separate column decoder to generate select signals as in figure 47b. The choice of figure 47b is faster because data from the bitline must propagate through only one series transistor. Column decoding takes place in parallel with row decoding so it does not impact delay[47].

Also, column multiplexing is helpful because the bit pitch of each column is so narrow that it can be difficult to lay out a sense amplifier for each column. Moreover, placing sense amplifiers after the column multiplexers reduces the number of those amplifiers required in the array[47].

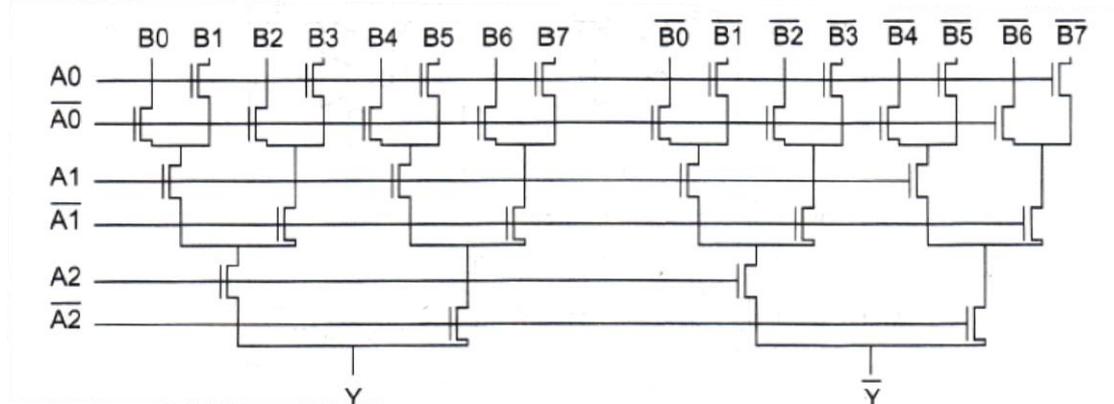


Figure 47a – Picture from [47]

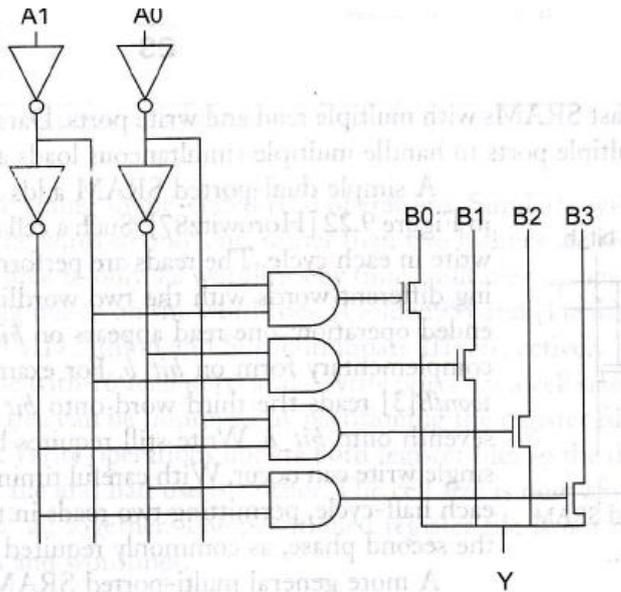


Figure 47b – Picture from [47]

Subsequently, the bitline conditioning circuitry is used to precharge the bitlines high before operation. As seen in figure 48a, a simple conditioner consists of a pair of PMOS transistors[43].

Sense amplifiers are very susceptible to differential noise on the bitlines because they detect small voltage differences. If bitlines are not precharged long enough, residual voltages on the lines from the previous read or write operation may cause pattern-dependent failure[32]. Equalization, is a necessary operation in order to prevent the sense amplifier from making erroneous excursions when turned on and is critical when the bit lines are precharged through PMOS pull-ups, since the precharge value can differ due to the variations in the device threshold[43]. In figure 48b, an equalizer transistor can be added to the bitline conditioning circuits to reduce the required precharge time by ensuring that bit and bit_b are at nearly equal voltage levels even if they have not precharged quite all the way to V_{DD} [47].

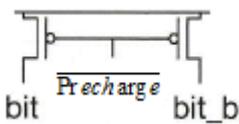


Figure 48a – Picture from [32]

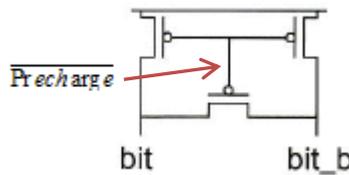
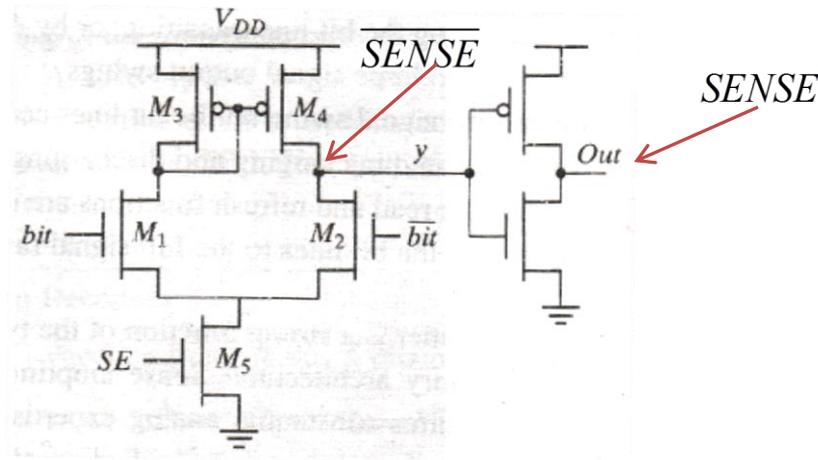


Figure 48b – Picture from [32]

A crucial circuit needed for the reading operation in an SRAM, is the sense amplifier, which takes small-signal differential inputs (i.e. the bit-line voltages) and amplifies them to a large-signal single-ended output. An example of such a circuit is given in figure 49, where amplification is accomplished with a single stage, based on the current mirroring concept. The input signals bit and \overline{bit} are heavily loaded and their swing is small as the small memory cell drives a large capacitive load during reading. The inputs are fed to the differential input devices M1 and M2 and transistors M3 and M4 act as an active current mirror load. The sense amplifier is conditioned by the sense amplifier enable signal SE, which is initially low and is

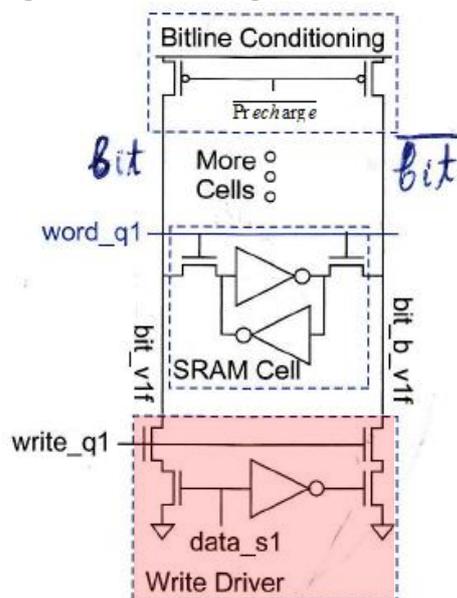
enabled once the read operation is initiated. The gain of such a differential-to-single ended amplifier is given by[43]:

$$A_{sense} = -g_{m1} \cdot (r_{o2} || r_{o4})$$

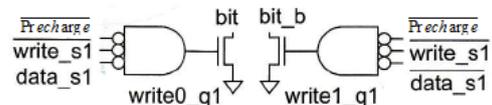


Picture 49 – Figure from [43]

Finally , the write driver which pulls the bitline or its complement low to write the cell , consists of a pair of transistors on each bitline for the data and the write enable as seen in figure 50a , or a single transistor driven by the appropriate combination of signals as seen in figure 50b.



Picture 50a – Figure from [32]



Picture 50b – Figure from [32]

IV. LOW POWER SRAM DESIGN

SRAMS typically form a dominant portion of the area and power of a system. The driving metric for an SRAM has been area for a long time due to the large number of cells in SRAM arrays. However , power is becoming increasingly important to the point of rivaling area as the driving metric , due to the fact that as higher levels of the cache hierarchy move on-chip , the power dissipation of the SRAM memory is growing relative to that of other components on the chip[45].

This is particularly true concerning the leakage component of chip power. As the SRAM must remain powered on to hold its data, the large number of transistors in on-die SRAM will constantly draw leakage power. This leakage power can dominate the standby power and active leakage power budgets in low-power applications, and become an appreciable fraction of the total dissipation in others[45].

Therefore, energy and leakage power reduction through low-voltage operation is highly desirable. However, due to **technology scaling**, **voltage reduction** and **various forms of variations**, there is a tradeoff between power and robustness, the latter being expressed by the hold, read and write SNMs.

The traditional 6-transistor(6T) SRAM cell, relies on ratioed device sizing to set the relative device strengths required for reading and writing. Since sizing ($\frac{W}{L}$) changes current I_D linearly while V_t variation has an exponential impact in sub-threshold I_D , variation can easily overwhelm the effect of sizing to cause bit-cell failures[42]. According to [49]-[50], process variations will limit standard 90nm and 65nm SRAMs to around 0.7V operation.

Variations in the bitcell transistors, the impact of which increases for DSM(Deep Sub-micrometer) devices because of the smaller transistor channel area, caused by phenomena such as global process variation, random doping mismatch, and temperature changes, degrade the SNM. In the subthreshold region, the consequences of these variations become even worse, as the delay and the drain current depend exponentially on the threshold voltage which varies greatly. The sensitivity of the SNM to threshold voltage mismatch may be lower in Weak Inversion, however, due to reduced V_{DD} , the absolute value of SNM decreases[49].

Since standard write operation depends on a carefully balanced ratio of currents, processing variation makes this ratio difficult to maintain as V_{DD} decreases, leading to errors during write access[49]. The effect of reduced write margin during operation in Weak Inversion, is shown in figure 51, where the SNMs for write access versus temperature and process corners(TT,WW,SS,WS and SW) are depicted for an 65-nm process at $V_{DD}=0.3V$ and $V_{DD}=0.6V$. At $V_{DD}=300mV$, the writing fails for large regions of process corner and temperature.

In this 65-nm process, due to lower mobility, PMOS is weaker than an iso-sized NMOS at nominal V_{DD} , but the PMOS current in Weak Inversion is larger than an iso-sized NMOS[49]. This, makes write functionality more challenging, as for the write operation, the PMOS loads of the cell should be weaker than the NMOS drivers so that we can force \bar{Q} to ground if we want to write a 1, or Q to ground if we want to write a 0[49].

The general trend showing an improvement of write operation at higher temperature occurs because the PMOS transistors(loads) weaken relatively to NMOS(drivers) as temperature rises. Also, as seen from the right part of figure 51, as V_{DD} increases, the write margin improves. The supply voltage of 0.6V is well above the V_t of both types of transistor and the PMOS has weakened relative to the NMOS because the mobility starts to dominate the differences in V_t . However, even at 0.6V, the write margin is barely negative for the worst-case corner (Weak NMOS, Strong PMOS), and this plot does not account for local V_t variation. As a result, $V_{DD}=0.6V$ is the best case voltage for which we can expect traditional write operations to work for a sub-threshold memory in this 65nm process[41]-[49].

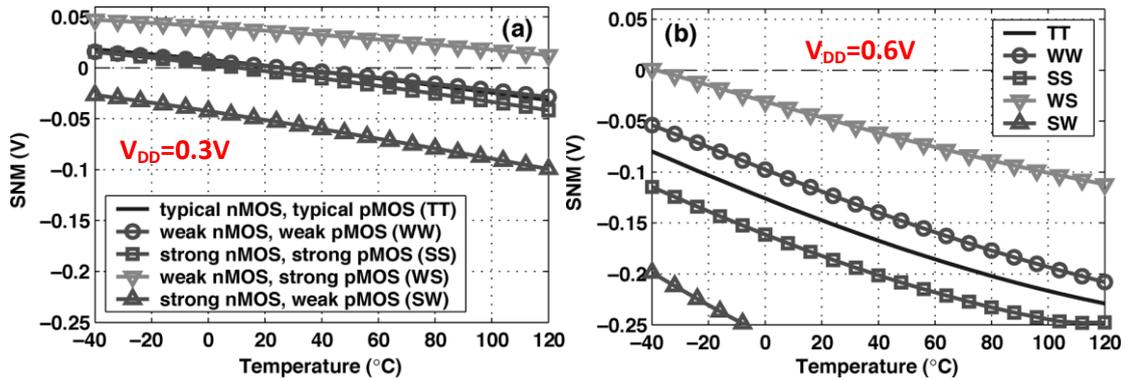


Figure 51 – Picture from [41]

The same effects stand for the case of the hold and read SNMs as seen in figure 52. In figure 52a where zero read or hold SNMs mean that with the presence of a noise signal of even a very small amplitude, the cell will flip, it is obvious that the typical read and hold SNMs degrade with technology scaling and with voltage reduction. This means that it is harder to make a robust array in a scaled technology, and that lowering supply voltage to reduce power degrades the cell stability. Furthermore, figure 52a confirms that the read SNM is quite a bit smaller than the hold SNM[45].

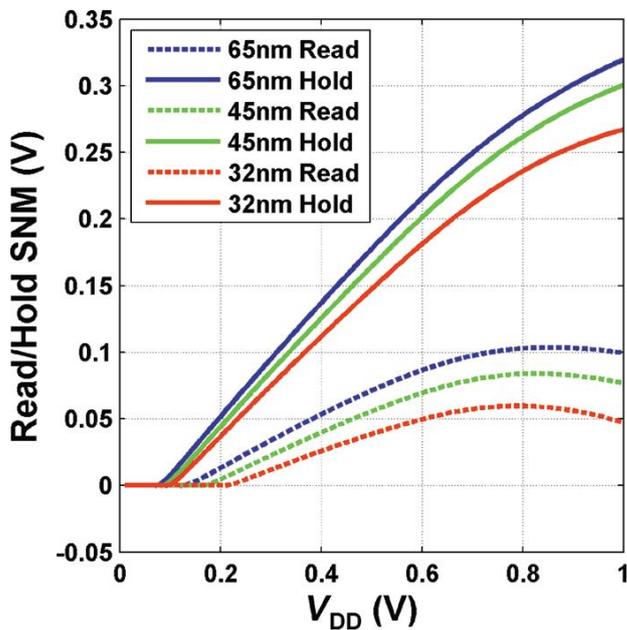


Figure 52a – Picture from [45]

Variations make things even worse, as shown in figure 52b, where distributions of the read SNM for the different technology nodes are plotted. The tails of these distributions correspond to cells with vanishingly small noise margin, indicating that those cells will be quite unstable during a read access even in traditionally safe SRAM architectures. For the 32nm technology, a substantial number of cells exhibit an SNM at (or below) 0, indicating a read upset even in the absence of other noise sources. This degradation of stability means that SRAM circuits/architectures must change if basic reading stability is to be maintained[45].

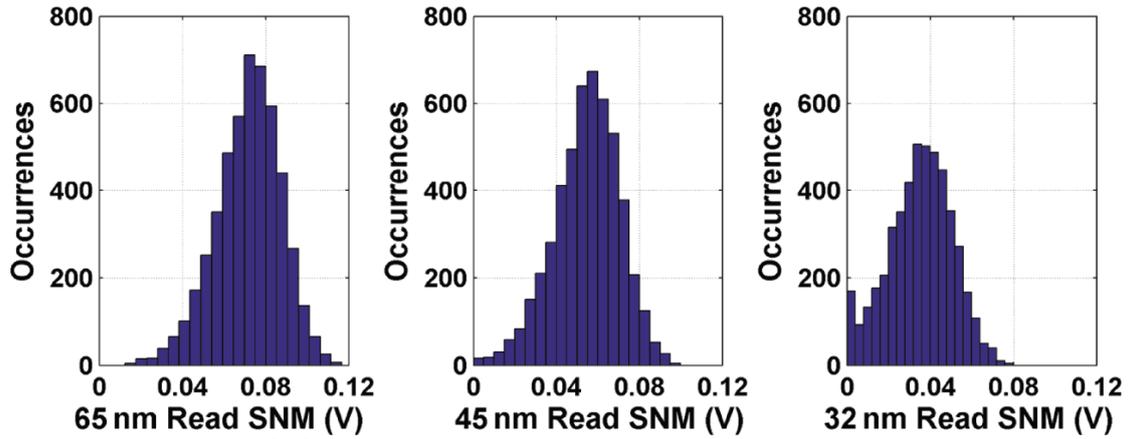


Figure 52b – Picture from [45]

Thus, in sub- V_t , the use of nonratioed static logic styles is necessary as conventional 6T SRAM cells do not function reliably because the ratio constraints for read stability and writability cannot be guaranteed, especially in light of threshold variations. Moreover, the poor ratio of I_{ON} (active current) to I_{OFF} (leakage current) limits the number of cells that can be connected to a local bitline. Nominally, the I_{ON}/I_{OFF} of devices in a circuit operating at the minimum energy voltage is between 10^3 - 10^4 , whereas that in strong inversion is approximately 10^7 with degradation in drain current, due to variation, severely reducing this ratio even further[32]-[48]-[51].

In sum, the ratioed operation, during hold, read and write, leaves the 6T SRAM highly susceptible to both variation and manufacturing defects[48].

V. AN ALTERNATIVE SRAM CELL

An alternative to the conventional 6T SRAM cell, based on the previously mentioned Source-coupled logic, is the 9T SCL SRAM cell, seen in figure 53a. The core of this cell is the SCL inverter, seen again in figure 53b.

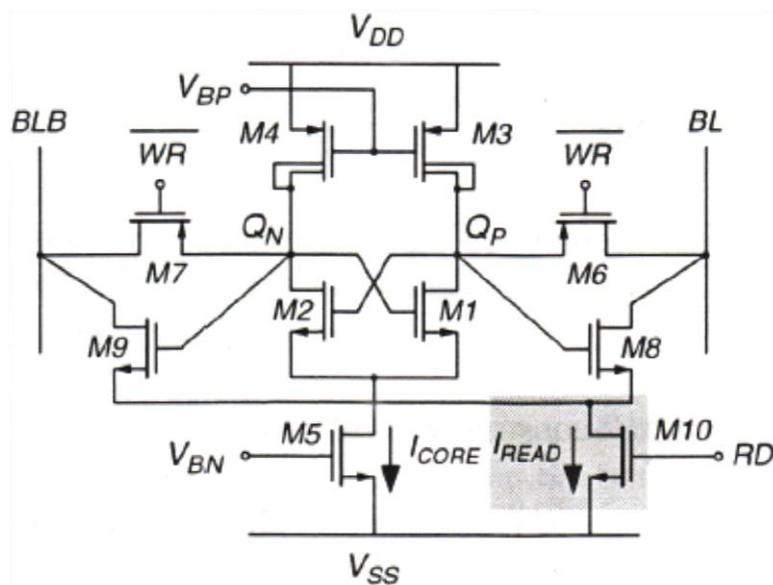


Figure 53a – Picture from [3]

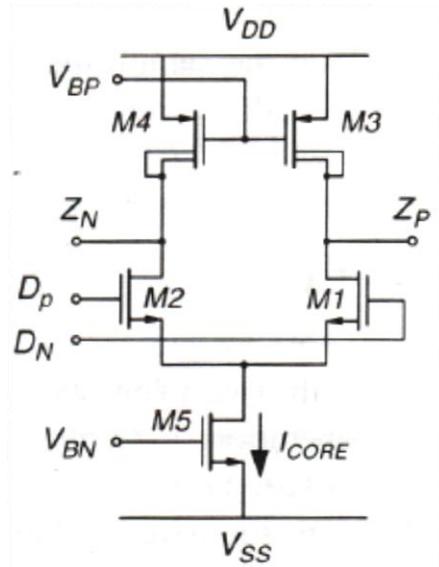


Figure 53b – Picture from [3]

This cell, exhibits very low stand-by dissipation in idle state, and allows robust read and write operations at frequencies that are significantly higher than those achievable in CMOS-based topologies[30].

Its core is based on a cross-coupled SCL inverter to construct the positive feedback needed to store the data. The loads used are the previously discussed Drain-bulk connected PMOS transistors, the gate voltage for which is again provided by the replica bias circuit, which also provides the bias for the generation of I_{CORE} .

According to [30], the supply voltage for this cell can be reduced to 350mV without degrading the static noise margin of the cell.

The write operation is performed by pre-charging BL and BLB nodes to the desired voltage levels, and then turning on the access transistors M6 – M7, by asserting $\overline{WR} = 0$, in order to charge/discharge the output nodes QP and QN of the memory core. After turning off the access transistors ($\overline{WR} = 1$), the positive feedback in the cell will preserve the new state. During the write operation, RD=0 and thus, the pull-down transistor M10 is off[3].

The read operation is performed by using the open-drain differential pair formed by M8 – M9, driven by the tail bias transistor M10 which is external to the cell and shared by the cells on a word-line as illustrated in figure 54. During the read cycle, M10 is turned on (RD=1) and conducts the current I_{READ} , which is steered to one of the output branches of BL/BLB depending on the stored data on the core. This output current is detected by a current-mode sense amplifier and will be converted to voltage. Therefore, the speed of the read operation is completely independent of the core tail bias current I_{CORE} and depends only on I_{READ} as well as the parasitic capacitances of the nodes BL/BLB[3].

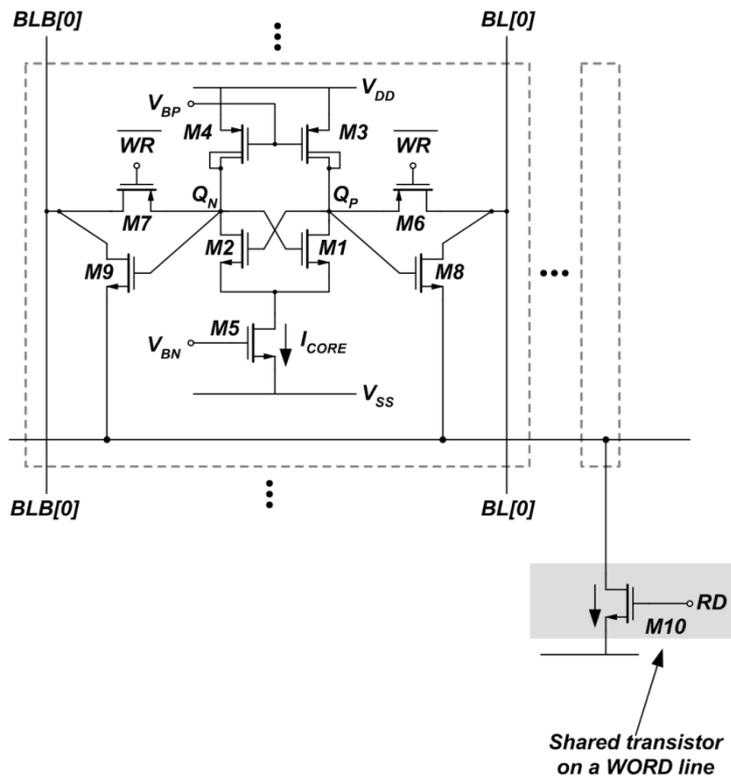


Figure 54 – Picture from [30]

VI. SIMULATION RESULTS

A 64-bit (8-byte) SRAM array using this alternative cell is designed using Cadence Virtuoso Schematic Editor and simulated using Cadence Spectre. The technology library used is the TSMC for 90nm CMOS. The minimum drawn length allowed is 100nm, whereas the minimum drawn width allowed is 150nm. The model used for the simulation, is the BSIM3 v.4 for the typical NMOS – typical PMOS case, provided by TSMC.

For the construction of the bias voltage V_{BP} for the memory cells, the single stage OTA is used again, with the same sizing. The current source used in the bias circuit has a value of 10pA as seen in figure 55c.

According to the BSIM3 v.4 model for the specific TSMC technology used, the ac simulation using the topology of figure 55a derived from [28], gives us, as illustrated in figure 55b:

DC gain ≈ 32

Phase Margin $\approx 95^\circ$

An SCL inverter is also simulated using the measurement topology from [29], the transfer characteristics of which for $V_{SWING}=200mV$ are seen in figures 56b and 56c. The sizing used here is the one used for the SRAM cell and includes smaller Widths and Lengths for the NMOS and PMOS transistors, due to the area constraints.

The SRAM cell itself seen in figure 57a, uses PMOS loads with ratio $W/L=150nm/150nm$, PMOS access transistors with ratio $W/L=200nm/100nm$, differential NMOS network transistors (drivers) with ratio $W/L=400nm/100nm$ and a tail NMOS transistor with $W/L=400nm/200nm$. Concerning the NMOS transistors for the Read operation, the open drain differential pair uses a $W/L=200nm/100nm$ and the - shared by the cells on a word-line - tail bias transistor a $W/L=200nm/200nm$.

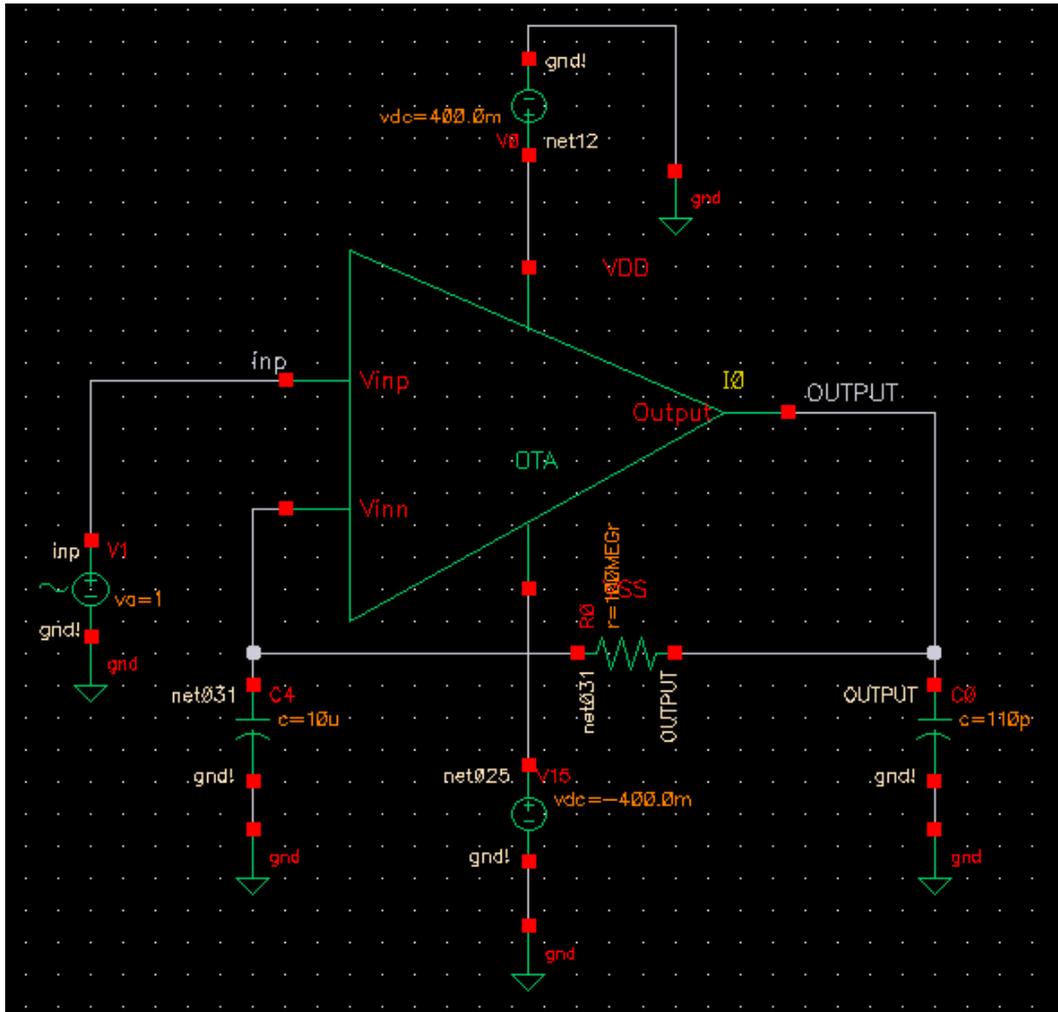


Figure 55a – The measurement topology for the simulation of the frequency response of the single stage OTA

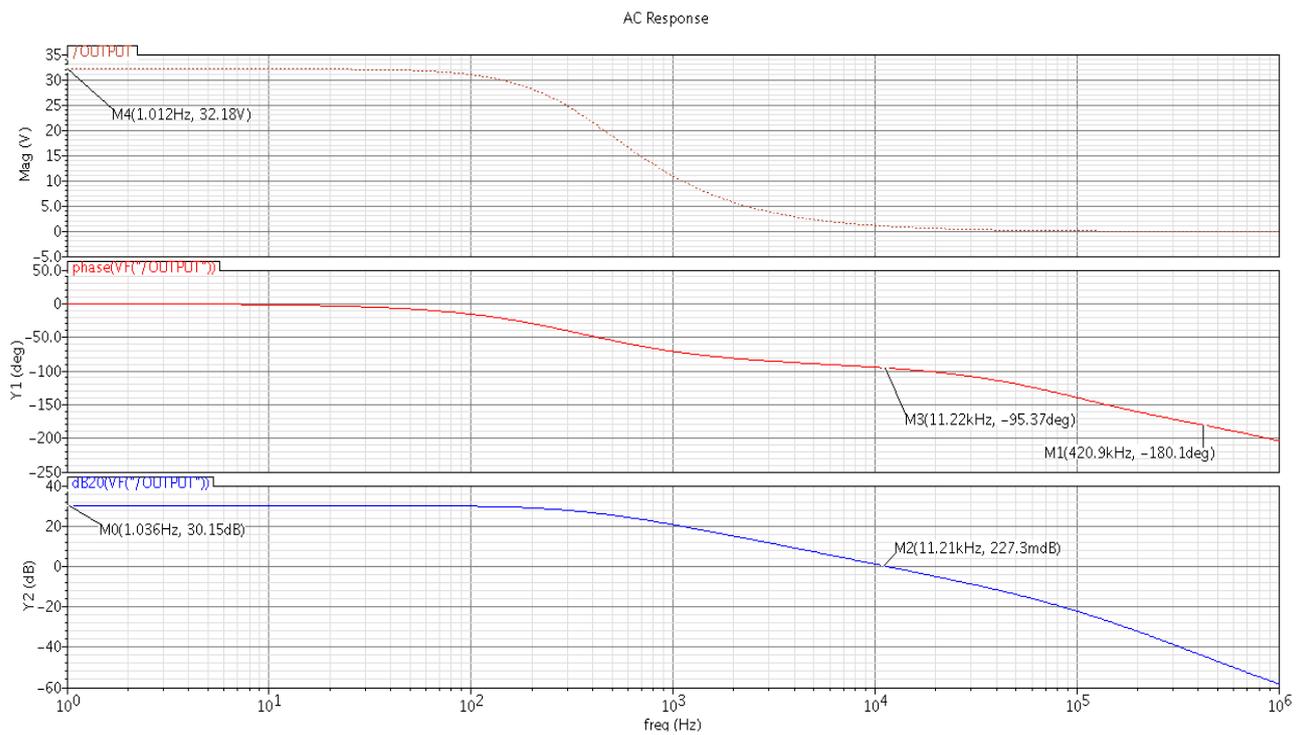


Figure 55b – The frequency response of the single stage OTA

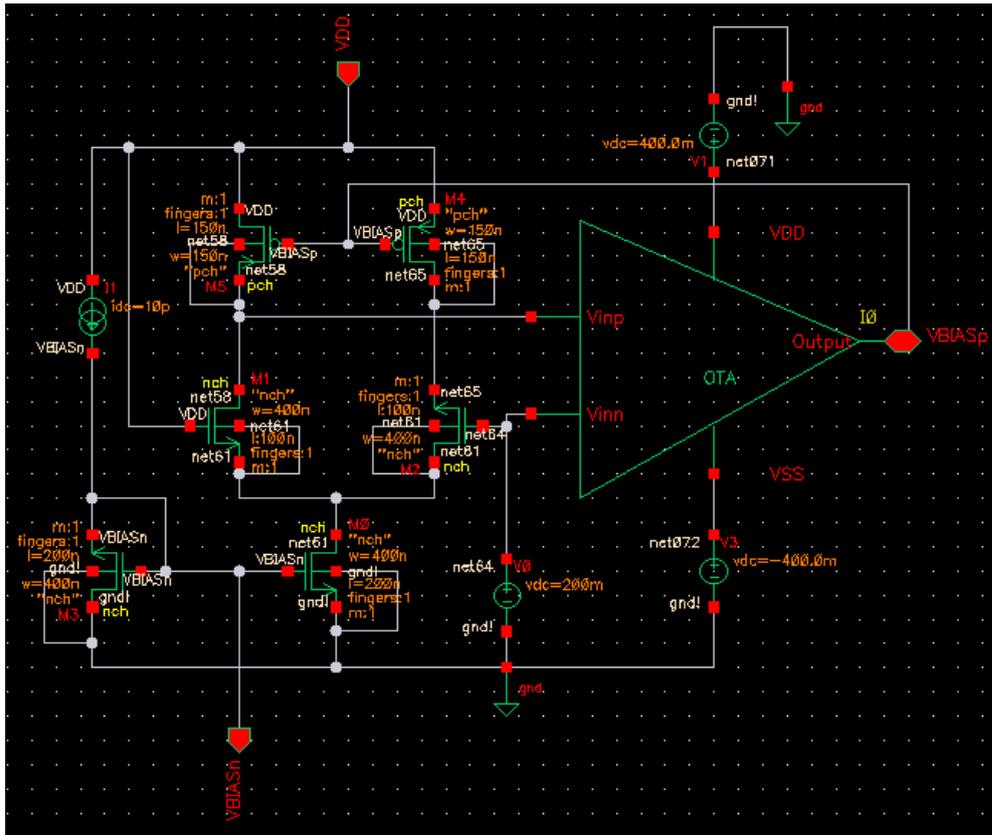


Figure 55c – The bias voltage generator

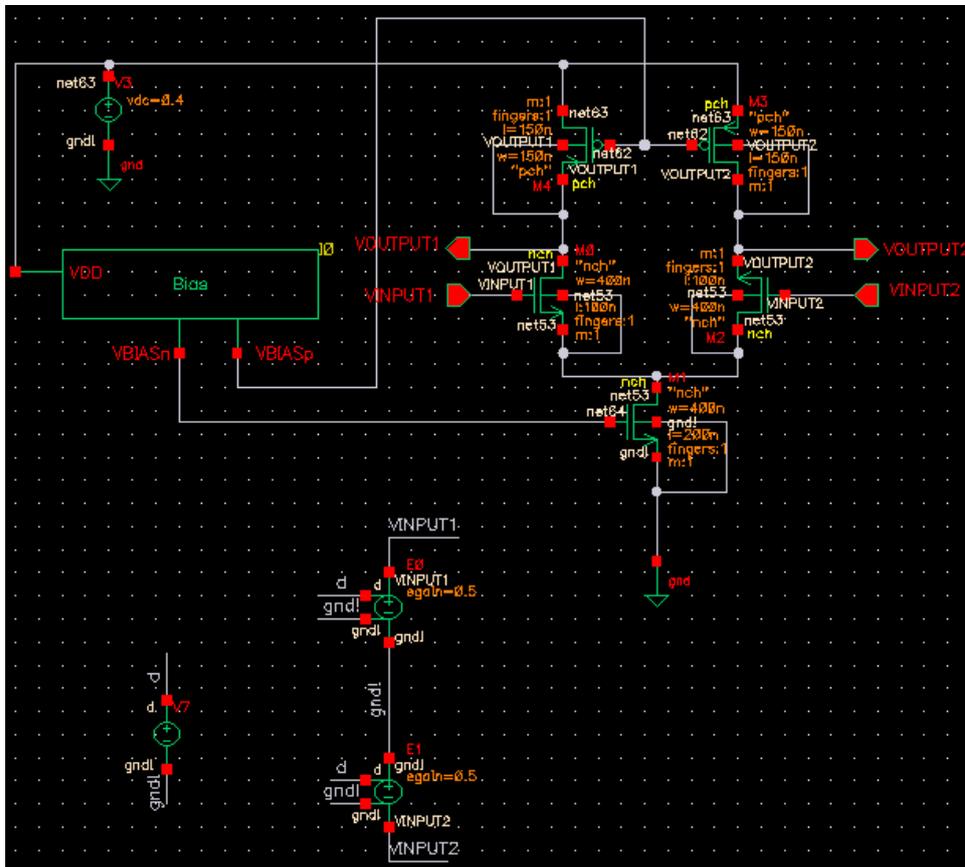


Figure 56a – The SCL inverter

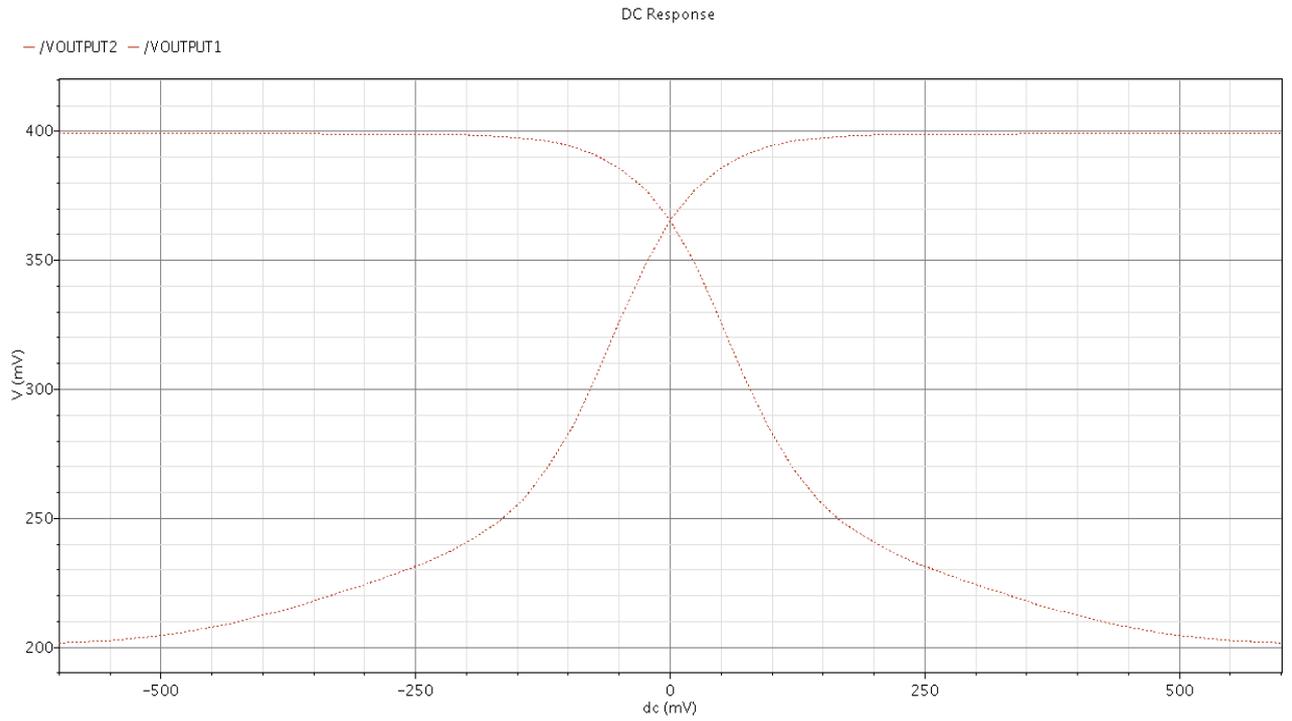


Figure 56b – Output Characteristics of the SCL inverter

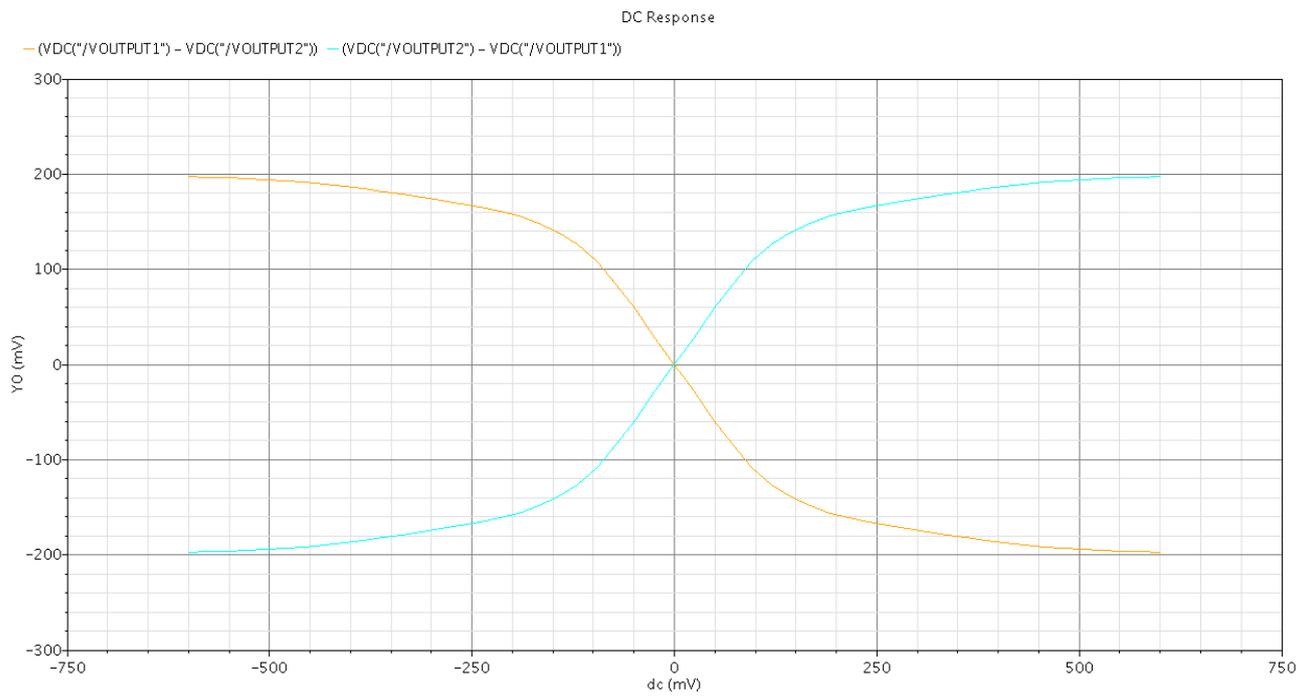


Figure 56c – Output Characteristics of the SCL inverter

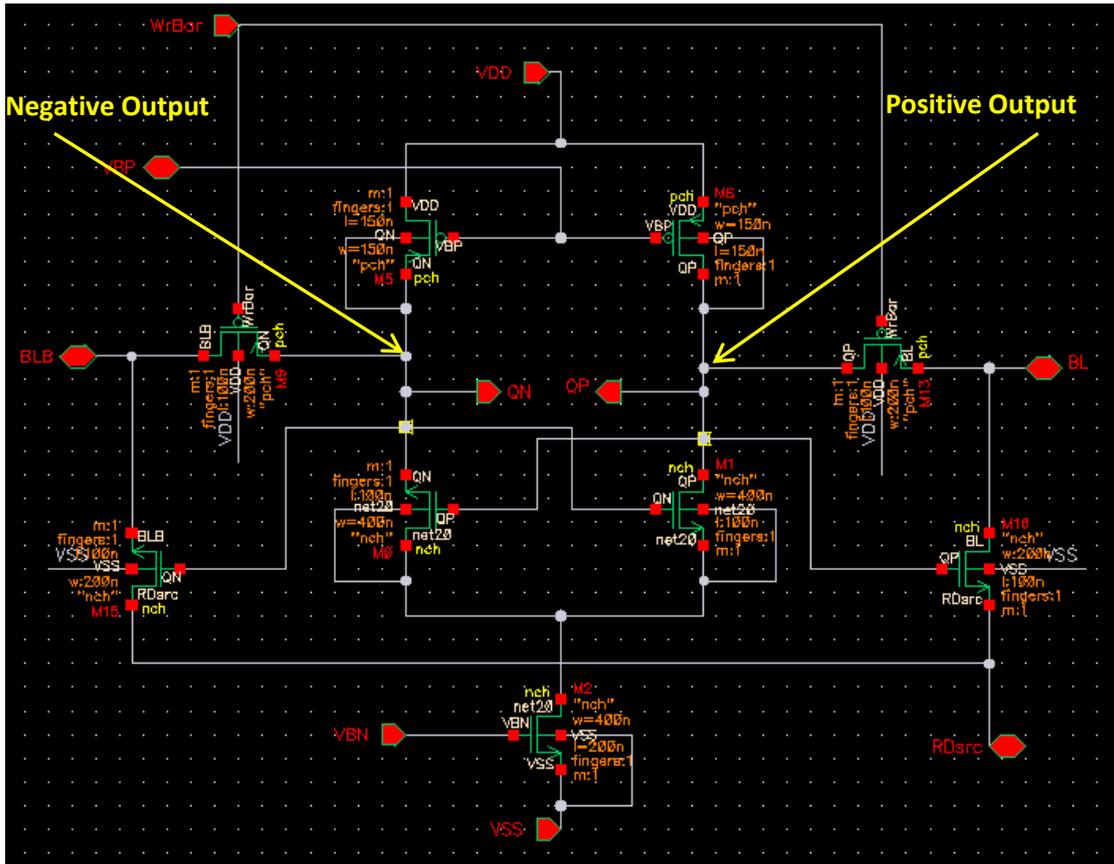


Figure 57a – The SCL SRAM cell

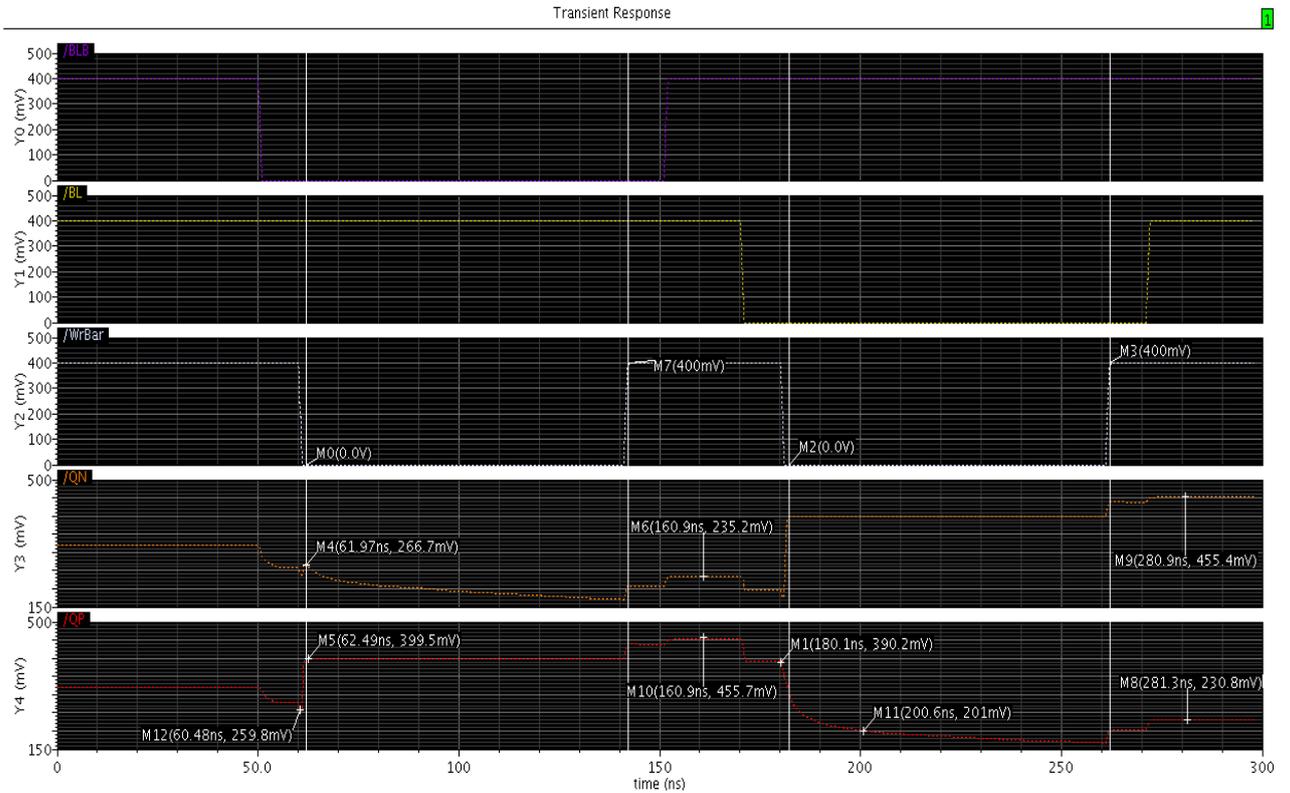


Figure 57b – Transient simulation of the SCL SRAM cell

The transient simulation of the SCL SRAM cell is shown in figure 57b for the case where $V_{DD}=0.4V$ and $V_{SWING}=0.2V$. A 1 is written into the cell when BL equals V_{DD} and BLB goes to ground at $t=50ns$, whereas a 0 is written into the cell when BLB equals V_{DD} and BL goes to ground at $t=170ns$. In the first case where the cell stores a 1, Q_p , which is the internal node of the cell storing a positive value, stabilizes at about 455mV, while Q_n which is the internal node of the cell storing a negative value, stabilizes at about 235mV. The difference between those values approaches the value of V_{SWING} and is about 220mV. Equivalently, in the second case where the cell stores a 0, Q_p stabilizes at about 230mV, while Q_n at about 455mV. The difference between the voltage levels of Q_p and Q_n , again approaches V_{SWING} and is about 235mV.

To implement the Row Decoder, the 3-bit tree-type decoder derived from [46] is used, with the use of NAND and NOR gates instead of the AND gates proposed, as seen in figures 58a, 58b and 58c.

In figure 58c, part of the transient simulation for the Row Decoder is seen. The outputs of the decoder are related to the inputs as follows:

- When ABC="000" then $m_0='1'$
- When ABC="001" then $m_1='1'$
- When ABC="010" then $m_2='1'$ etc.

Concerning the sizing of the transistors constituting the NAND and NOR gates of each Sub Decoder, PMOS have a ratio of $W/L=400nm/100nm$ and NMOS a ratio of $W/L=200nm/100nm$. As seen for the example of the output m_1 , the delay between the selection of the input and the activation of the respective output is about 3ns.

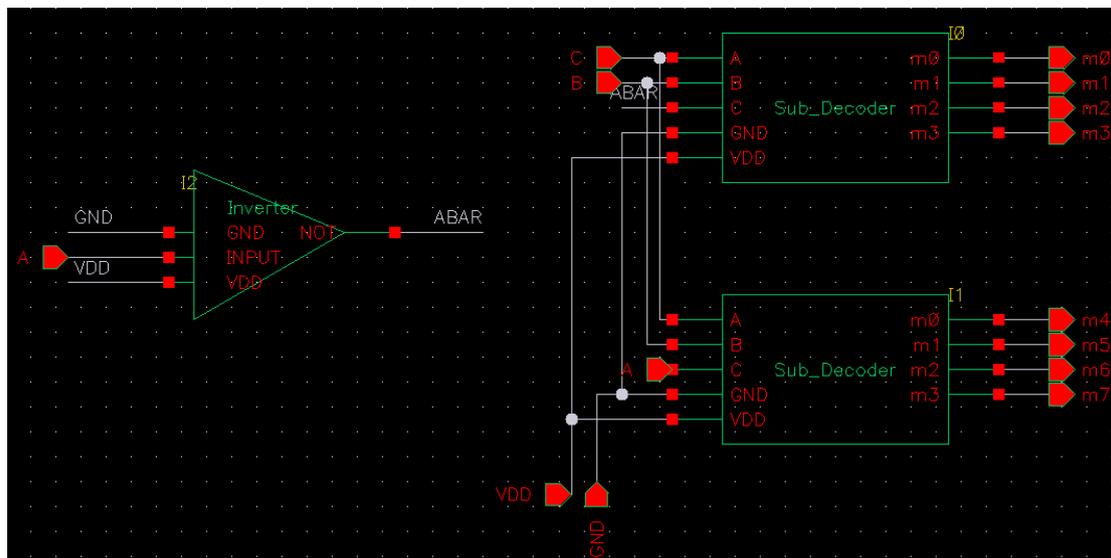


Figure 58a – The Row Decoder top level hierarchical module

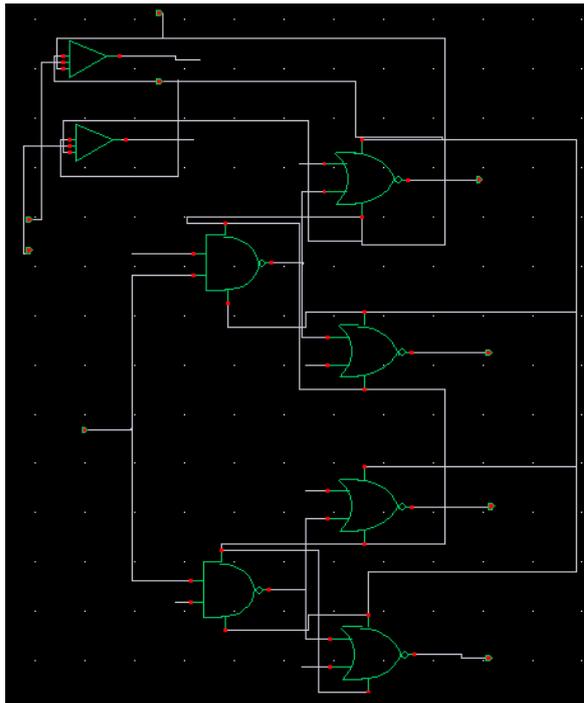


Figure 58b – The subdecoder module

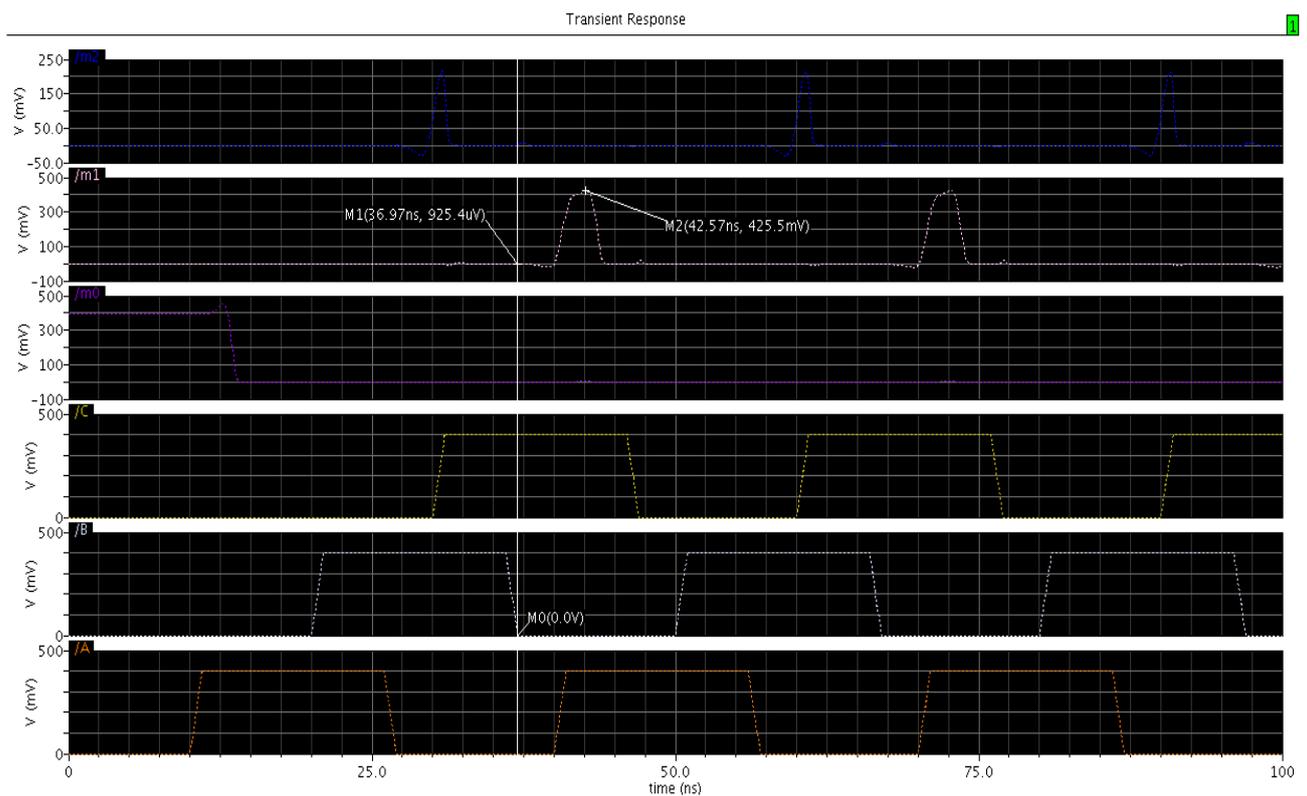


Figure 58c – Part of the transient simulation of the Row Decoder

Next , we refer to the bitline conditioning circuitry used to precharge the bitlines high before each write or read operation , including the equalizer transistor , as seen in figure 59a. We use this circuit for each column of our memory. For a quicker precharge operation , we use a W/L ratio of $1.6\mu\text{m}/100\text{nm}$.

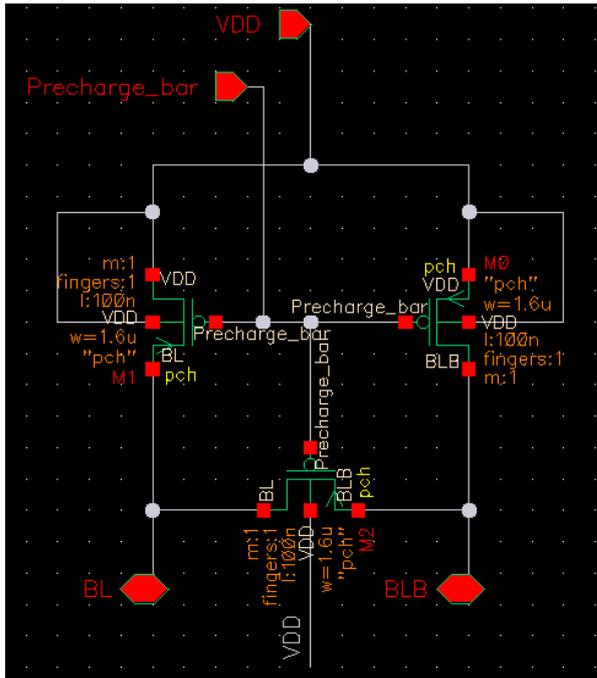


Figure 59a – Bitline conditioning circuitry along with the equalizer transistor

The memory array architecture we use, is the folded structure seen from [32] with each row including 2 words, 4 bits each. Thus, we have 8 rows and 8 columns. To extract the 4 bits of the specific word we access, we use four 2-to-1 column multiplexers. This is shown in the following scheme:

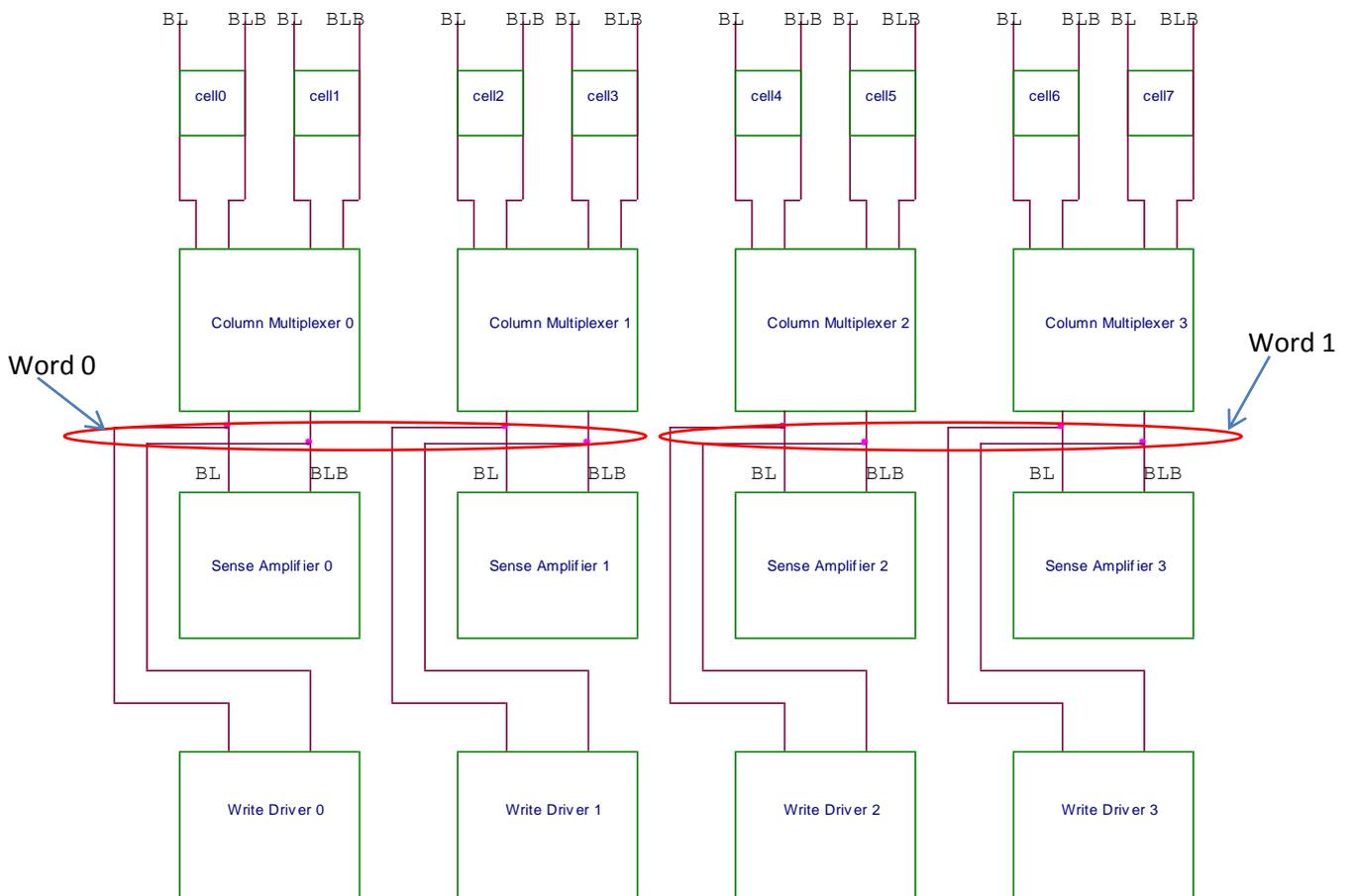


Figure 60a – Column Multiplexing

The column multiplexer is shown in figure 60b. For a better performance , we use CMOS pass gates instead of only NMOS or only PMOS transistors. This keeps the BL or BLB that we want to keep to V_{DD} during a read or a write operation from being quickly discharged to ground. The sizing used is $W/L=800nm/100nm$ for NMOS and $W/L=1.6u/100nm$ for PMOS.

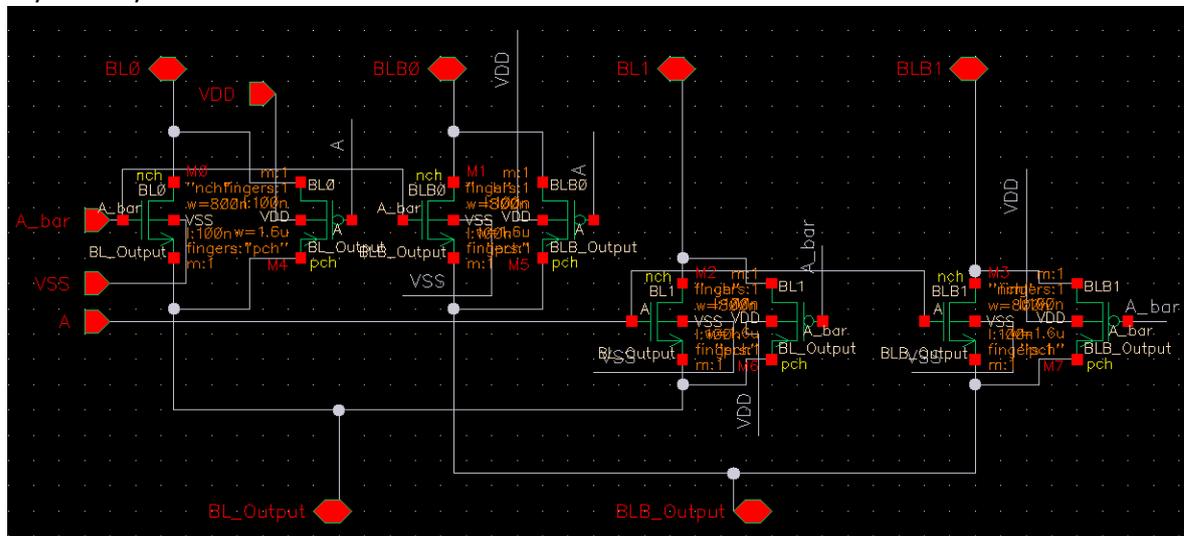


Figure 60b – The 2-to-1 column multiplexer

As depicted in figure 60a , at the output of the column multiplexer two circuits are connected , namely the Sense Amplifier which performs (or speeds up) the read operation and the Write Driver which performs the write operation.

In this particular implementation seen in figure 61a , we use the SCL variation of the latch based structure proposed in [43] where instead of the traditional CMOS inverter , we use the SCL inverter.

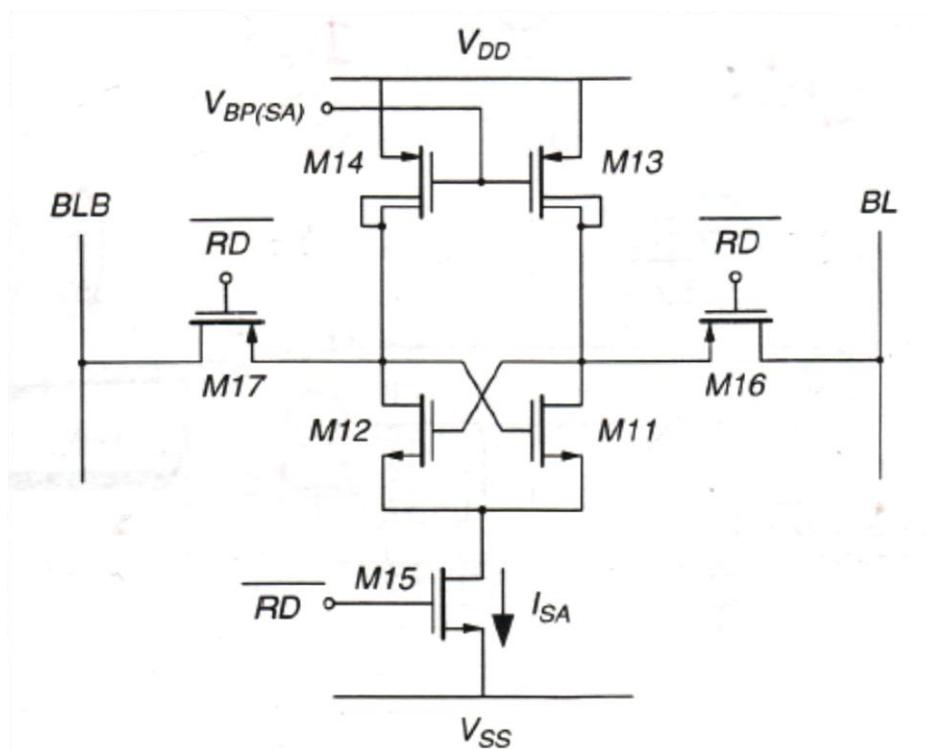


Figure 61a – Picture from [3]

During the hold or write modes ($RD=0$) where M16 and M17 are off and M15 is on, the Sense Amplifier is isolated from the memory and operates as a latch keeping the latest data that has been previously read. When $RD=1$, M15 turns off and M16 and M17 turn on. As a result, tail bias current will be switched off and the PMOS loads of the Sense Amplifier (transistors M13 and M14), the NMOS open drain differential pair (transistors M8 and M9 seen in figure 54) of each SRAM cell and the tail bias transistor M10 (seen in figure 54) which is external to the cell and shared by the cells on a word-line, will construct a single stage amplifier and the output of the memory cell will be amplified[3].

At the output of our SCL Sense Amplifier which is seen in figure 61b, we connect a CMOS inverter for two reasons:

1. First, to get the positive value stored into the cell because due to the way the Read operation works, in order to read a 1, after a precharge, BL discharges to 0 and BLB stays to 1. Thus, the SCL Sense Amplifier will read a 0 as its positive output goes to 0 and its negative output to 1. On the other hand, in order to read a 0, after a precharge, BLB discharges to 0 and BL stays to 1. Thus, the Sense Amplifier will read a 1 as its positive output goes to 1 and its negative output to 0. This is seen in figure 61c.
2. Second, to get an output with voltage levels ranging from rail to rail, i.e. from ground to V_{DD} .

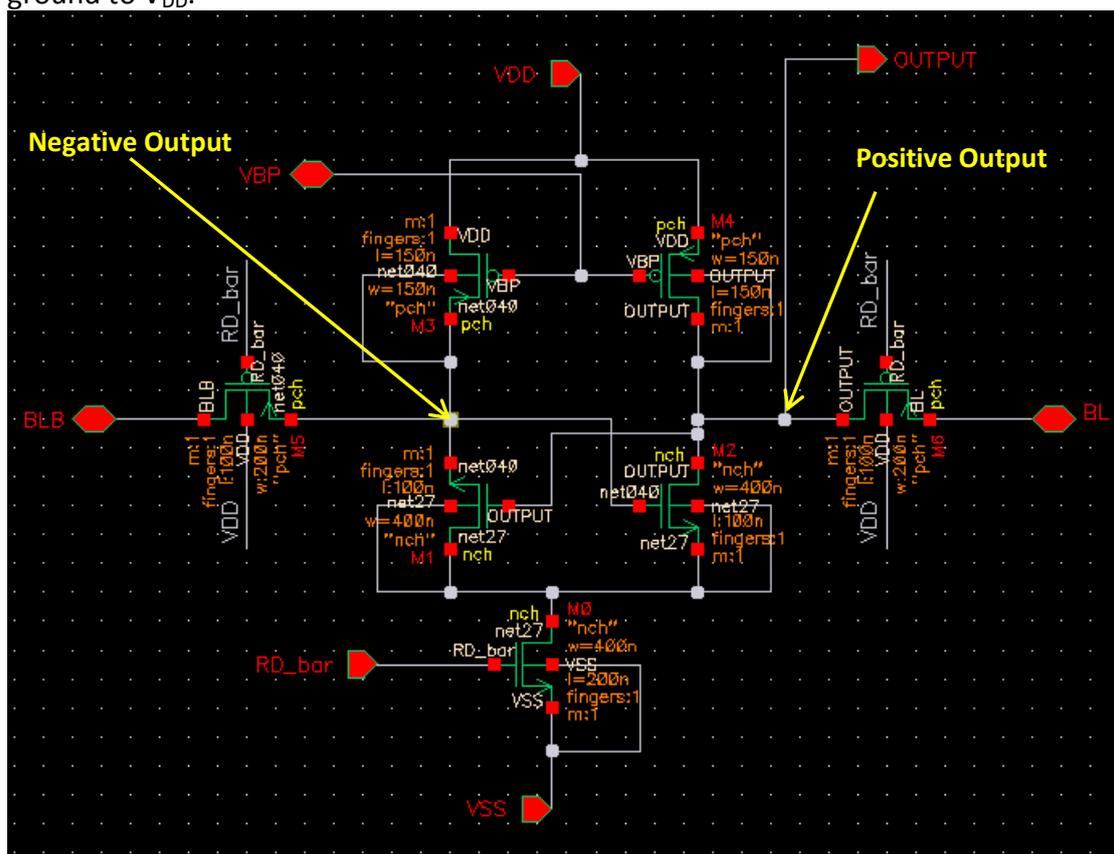


Figure 61b – The SCL Sense Amplifier

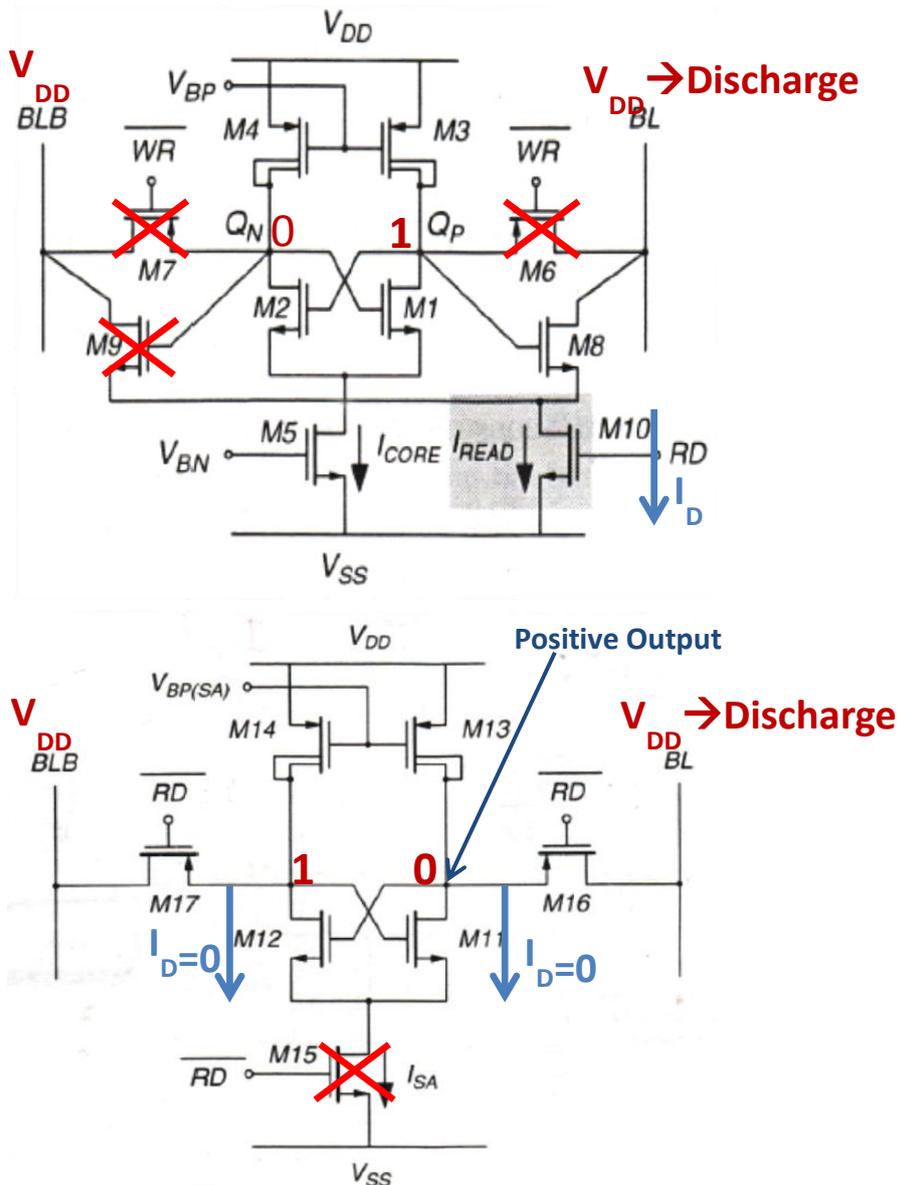


Figure 61c – Read of 1 from the selected memory cell using the sense amplifier

The last part of the peripheral circuitry of our memory is the Write Driver. We adopt the implementation from [40] shown in figures 62a and 62b. This circuit writes the input data in and its complement buffered by inverters 2 and 3 to the bit lines BL and BLB through two transmission gates TG1 and TG2. WE and its complementary WEB are used to activate TG1 and TG2 and discharge BL or BLB through the NMOS transistors in inverter 2 or 3[40]. In order to write a 1 into the chosen cell, first we perform a precharge of both bitlines and then, BLB discharges to 0 while BL is kept to 1 (i.e. to 400mV), while in order to write a 0 into the chosen cell, first we perform a precharge of both bitlines and then, BL discharges to 0 while BLB is kept to 1.

The sizing of the Write Driver transistors is as follows:

NMOS: $W/L=800\text{nm}/100\text{nm}$, PMOS: $W/L=1.6\mu\text{m}/100\text{nm}$

Because of the fact that only one Write Driver is needed for two SRAM columns (i.e. here for 16 memory cells), the area impact of a large Write Driver is not multiplied by the number of cells in the column and thus it can have a larger size

than minimal[40]. This is also true for the other shared circuits that are used only once or in small numbers , like the Sense Amplifier , the Precharge and Equalization unit , the Row Decoder and the Column Multiplexer.

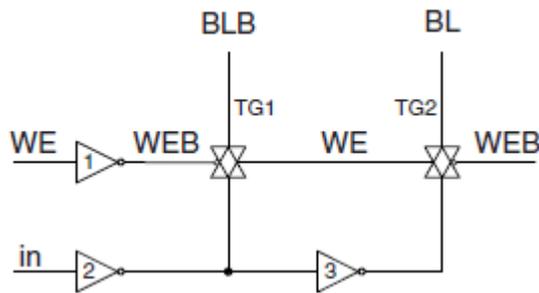


Figure 62a – Picture from [40]

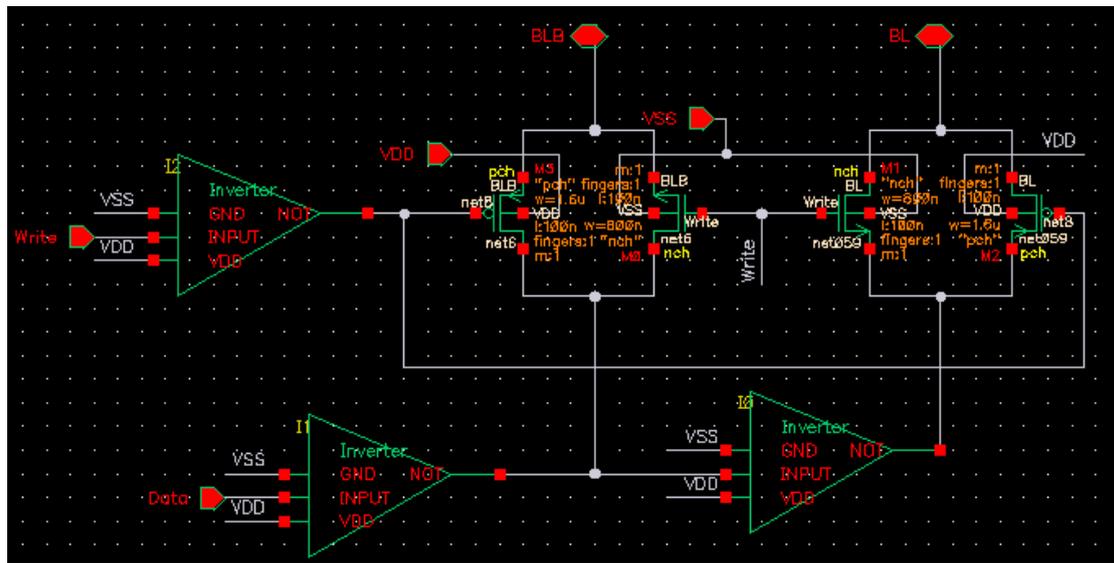


Figure 62b – The Write Driver

The complete circuit of the SRAM memory is seen in figure 63a. In figure 63b , we show the part of the memory where the control signals for the read operation are generated. We use an AND gate (NAND with an inverted output) with inputs the signal Read and the signal WordlineX , where X is the row (0 to 7) that we access a given time. When WordlineX=1 and Read=1 , then the external to the cell and shared on a word-line NMOS transistor M10 turns on and the read operation begins.

In figure 63c , we show the part of the memory where the control signals for the write operation are generated. We use a NAND gate with inputs the signal Write and WordlineX. When Write=1 and WordlineX=1 , then the output of the NAND gate which we call WriteX , where X the row (0 to 7) that we access a given time , takes the value of 0. Thus , the PMOS access transistors used in the SRAM cell that we want to write , turn on.

In figure 63d , the bitline conditioning circuitry used to precharge the bitlines high before operation along with the equalizer PMOS transistor is seen , connected together with the cells of each column of the memory. We need one such circuitry for each column , that is 8 in total.

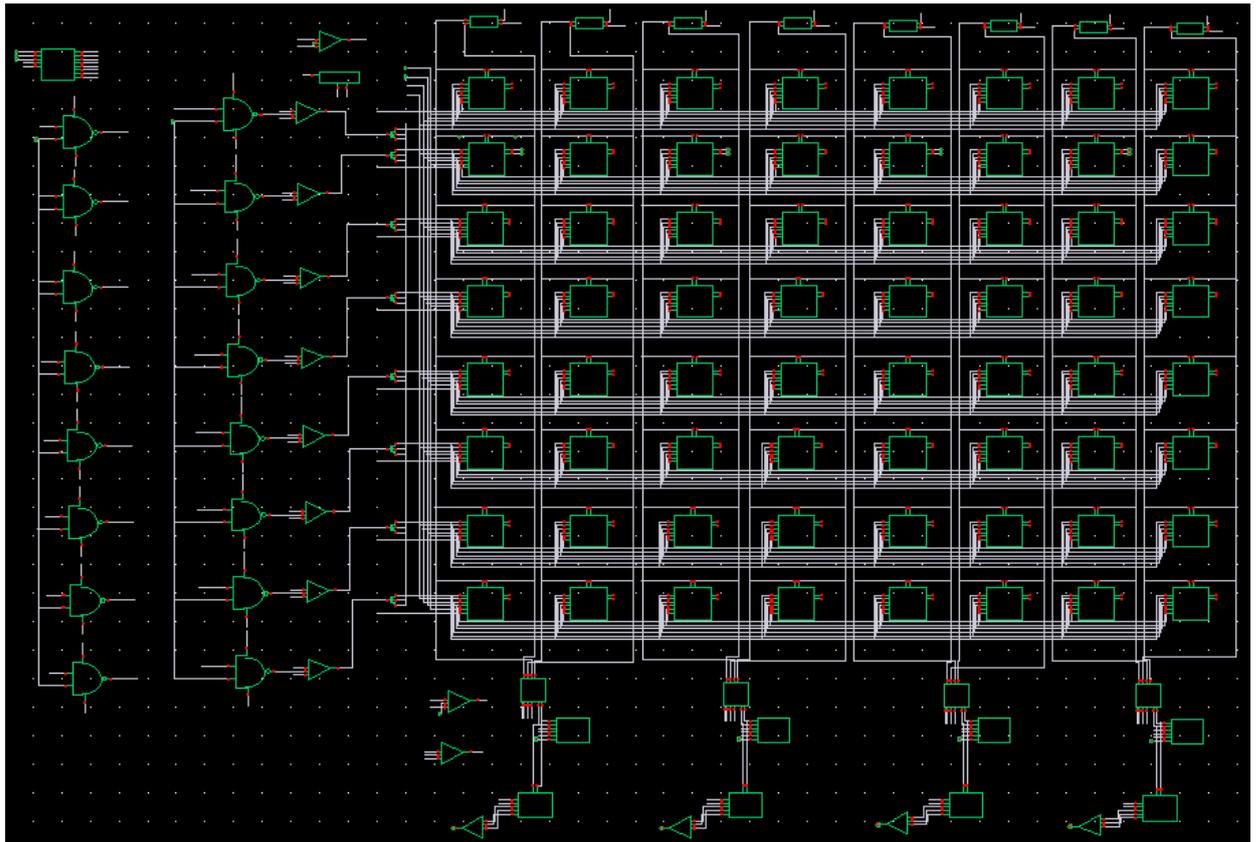


Figure 63a – The complete SRAM memory

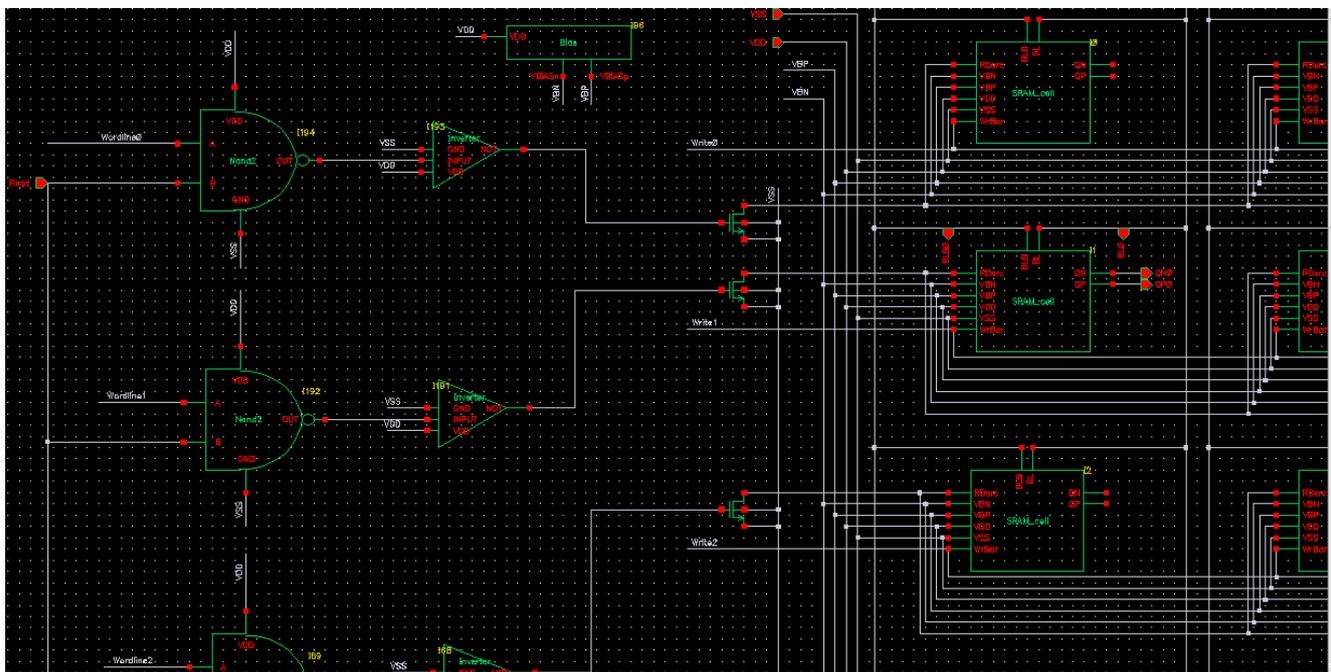


Figure 63b – The generation of the signals for the read operation

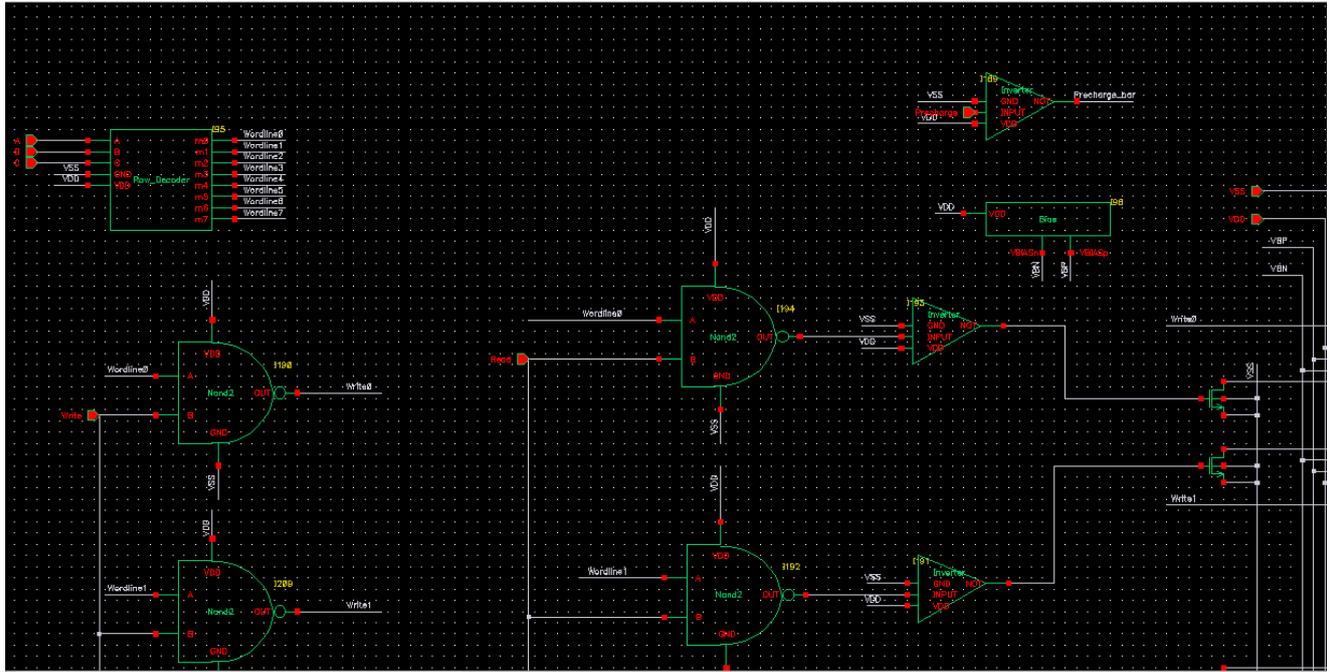


Figure 63c – The generation of the signals for the write operation

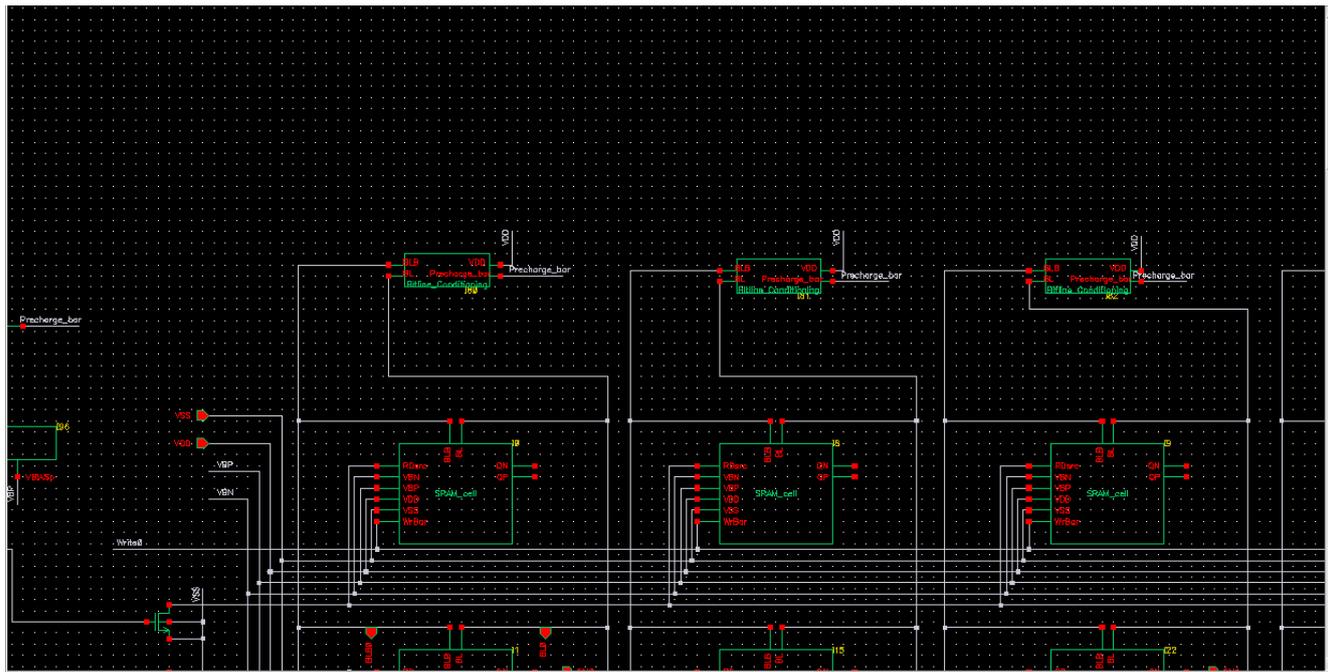


Figure 63d – The bitline conditioning and equalizer circuitry connected to the rest of the memory

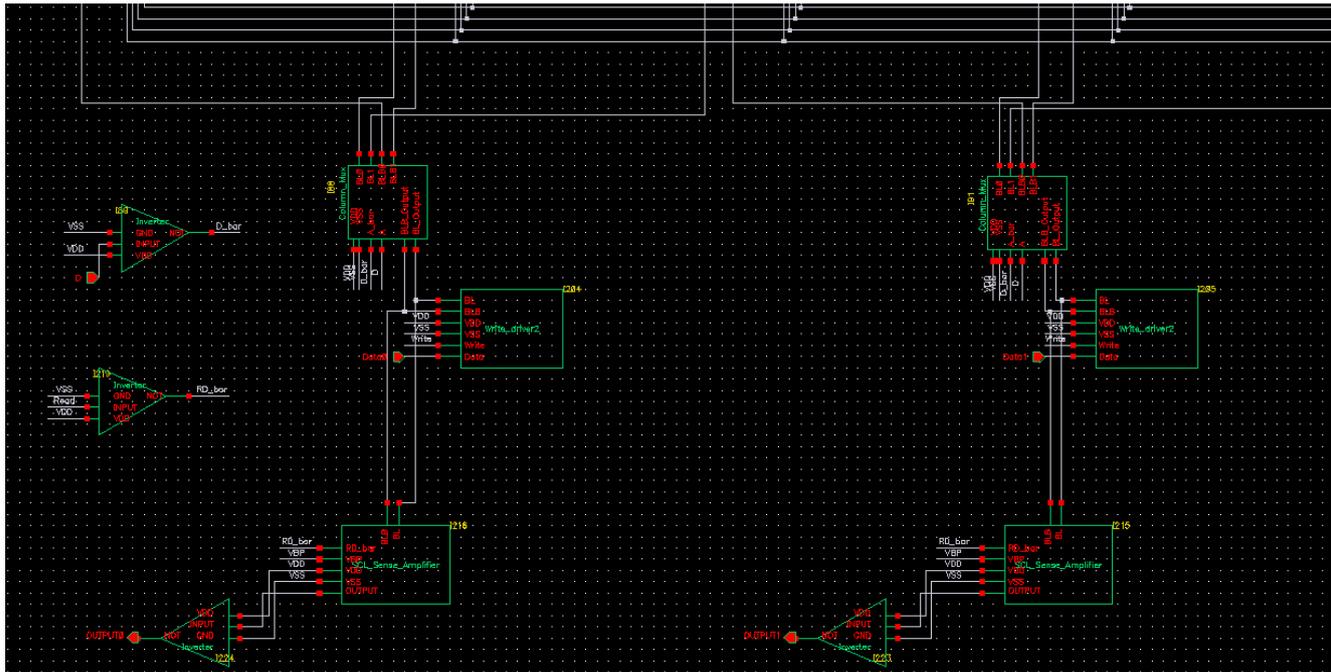


Figure 63e – The Column Multiplexers , Write Drivers and Sense Amplifiers connected to the rest of the memory

The connection of the Column Multiplexers , Write Drivers and Sense Amplifiers with the rest of the memory , according to the scheme of figure 60a , is seen in figure 63e. In total , we need 4 Column Multiplexers , 4 Write Drivers and 4 Sense Amplifiers.

Proceeding now to the simulation of the complete SRAM memory , the transient simulation results are shown in figures 64b , 64c and 64c. In this transient simulation , we access the cells within the red circles in figure 64a , which is the word “0010”.

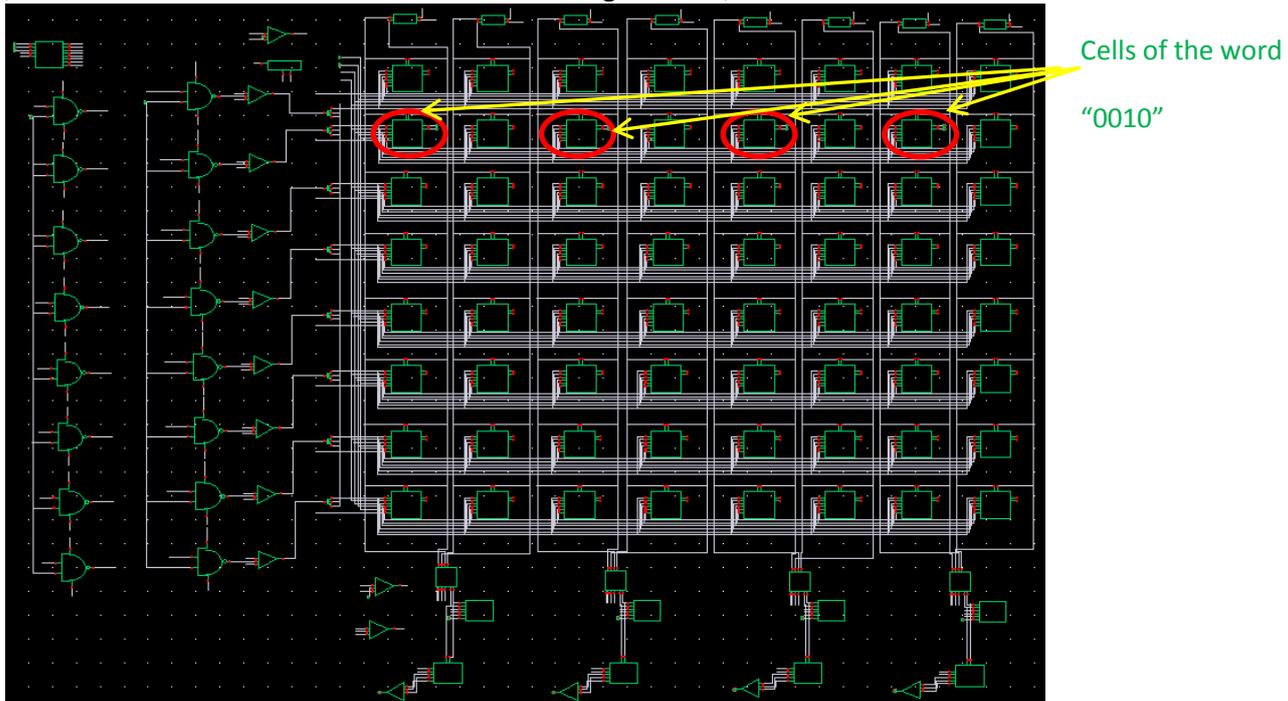


Figure 64a – Accessing the cells of the word 0010

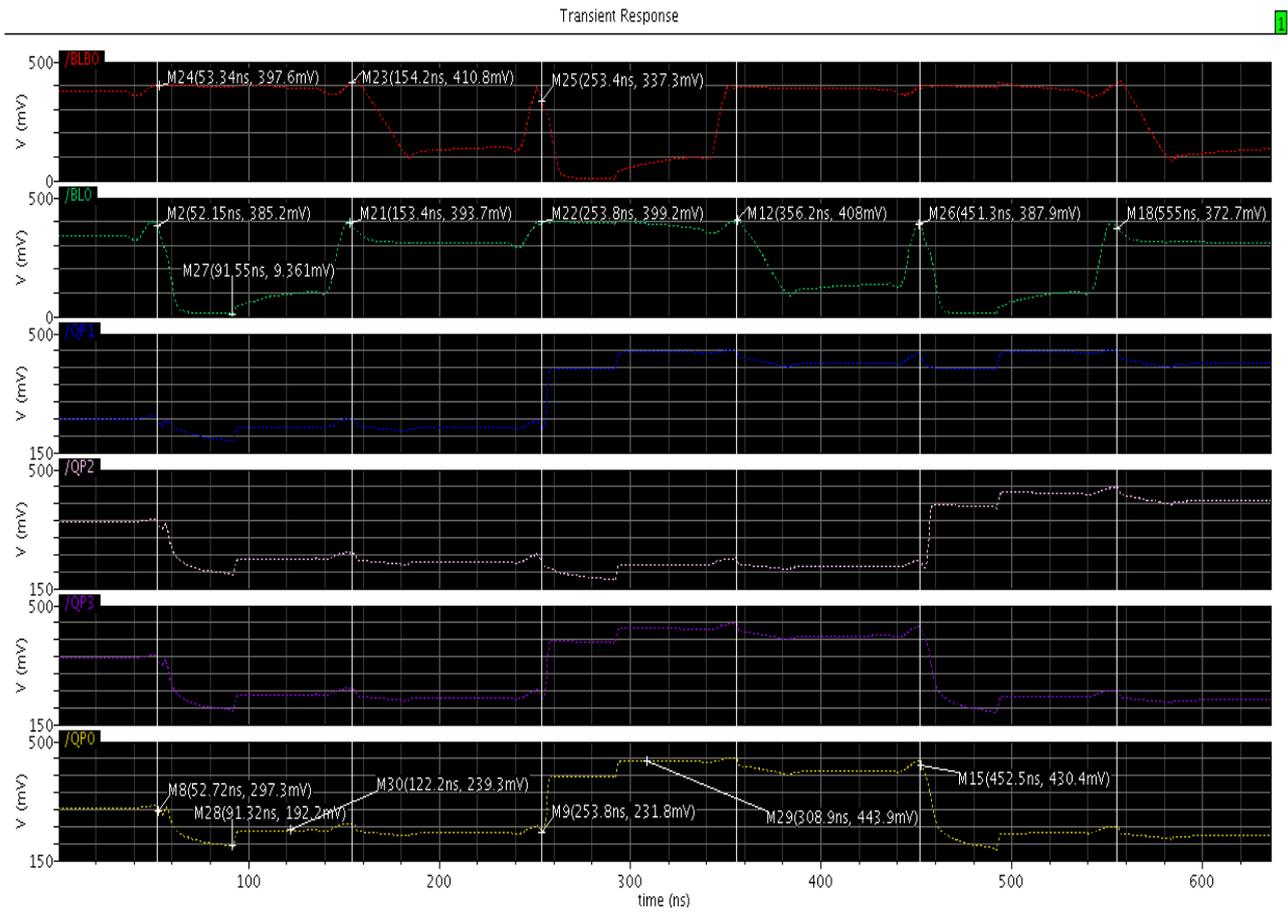


Figure 64b – Transient simulation of the SRAM memory , part 1

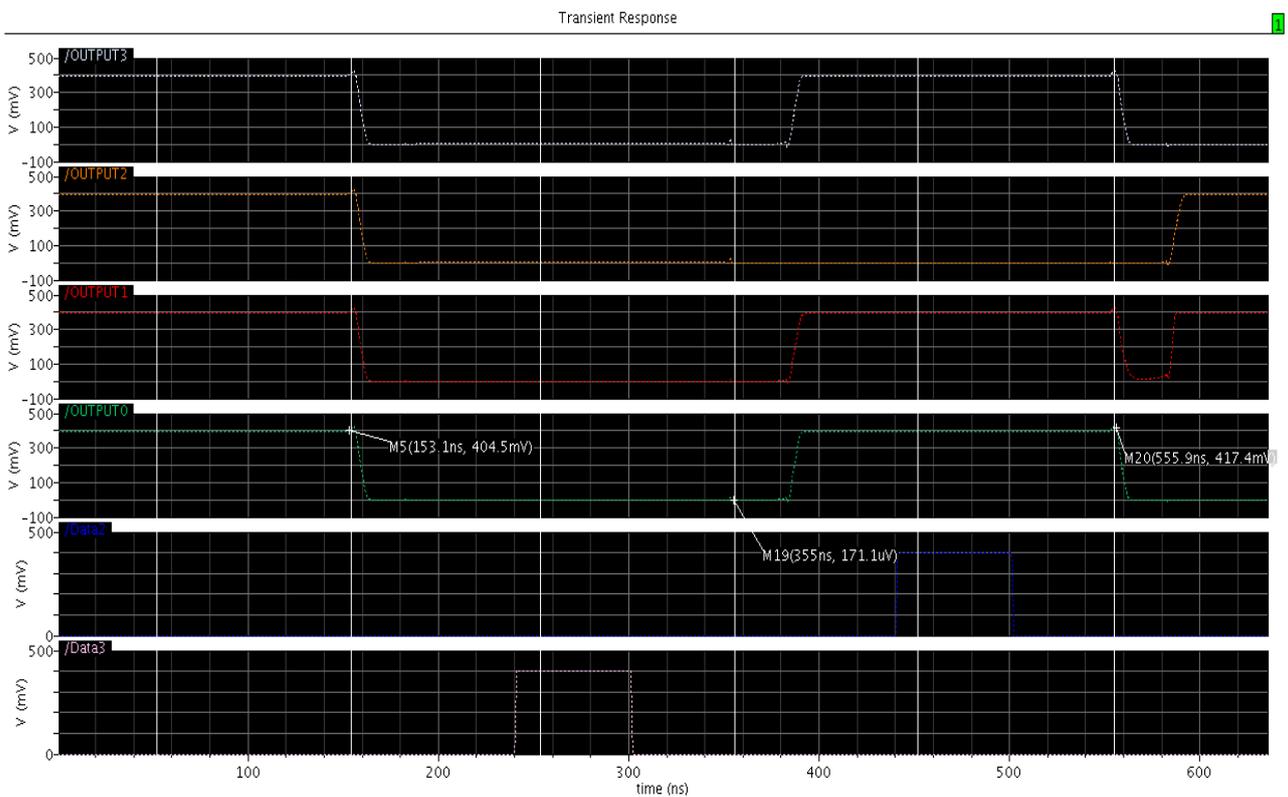


Figure 64c – Transient simulation of the SRAM memory , part 2

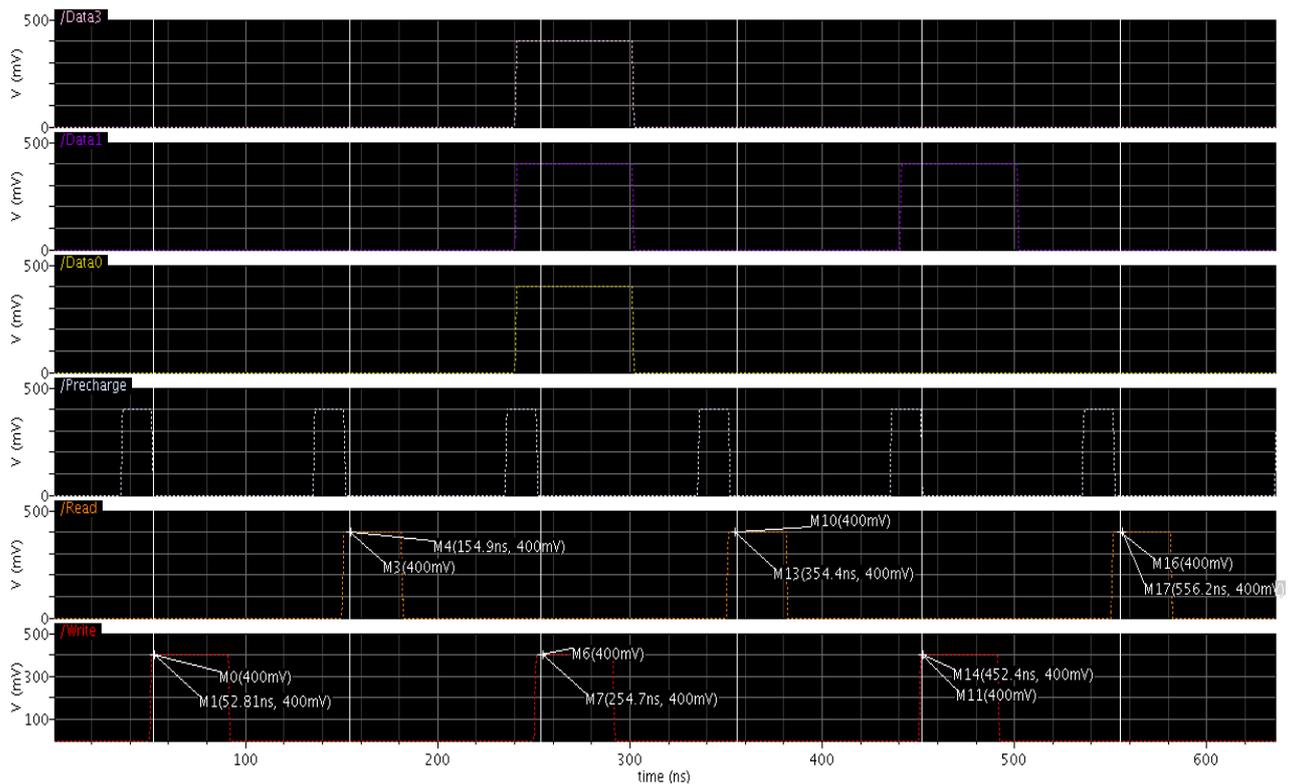


Figure 64d – Transient simulation of the SRAM memory , part 3

During this transient simulation , we provide the input signals Write , Read , Precharge , A , B , C , D , Data0 , Data1 , Data2 , and Data3.

The signals A , B , C and D , are the signals of the address of the word that we are accessing. Since we are accessing the word “0010” those signals take the value: A=’0’, B=’0’, C=’1’, D=’0’.

The signals Data0 , Data1 , Data2 and Data3 , are the signals of the data that we want to write to the SRAM memory during the write operation. In this particular simulation we perform a write operation three times. The first time we write into the selected word the value “0000” , the second time the value “1101” and the third time the value “0110”. Therefore , the first time Data0=’0’, Data1=’0’, Data2=’0’ and Data3=’0’, the second time Data0=’1’, Data1=’1’, Data2=’0’ and Data3=’1’ and the third time Data0=’0’, Data1=’1’, Data2=’1’ and Data3=’0’.

The signal Write is given the value ‘1’ only when we perform a write operation , it is given the value ‘0’ otherwise.

The signal Read is given the value ‘1’ only when we perform a read operation , it is given the value ‘0’ otherwise.

The signal Precharge is given the value ‘1’ only when we perform a precharge operation , it is given the value ‘0’ otherwise.

The output signals that we check are BL0 , BLB0 , QP0 , QP1 , QP2 , QP3 , OUTPUT0 , OUTPUT2 , OUTPUT3 and OUTPUT4.

The signal BL0 is the (positive) bitline BL of the first cell of the word we access , while BLB0 is the (negative) bitline BLB of the first cell of the word we access.

The signals QP0 , QP1 , QP2 , QP3 are the positive outputs of the cells of the word we access. These signals show at any given time the values stored into the cells of this specific word.

The signals OUTPUT0 , OUTPUT1 , OUTPUT2 , OUTPUT3 , are the outputs of the Sense Amplifiers that , when Read=1 , provide us with the value stored into the cells of the word at the moment we perform the read operation.

According to [3] , the value stored within the SCL SRAM cells signifying 0 , should be about $V_{DD}-V_{SWING}$, which in our case is about 200mV , as we use $V_{DD}=400\text{mV}$ and $V_{SWING}=200\text{mV}$. Also , the value stored signifying 1 , should be V_{DD} , that is 400mV for our design.

As derived from the transient simulation , indeed the difference between the voltage levels representing values 0 and 1 , is close to V_{SWING} ($444\text{mV}-239\text{mV}=205\text{mV}$) , however the voltage level representing 0 is a bit larger than $V_{DD}-V_{SWING}$ (239mV) and the voltage level representing 1 is a bit larger than V_{DD} (444mV). This is due to statistical dopant fluctuations that affect V_t and cause a nonzero output offset voltage of about 40-45mV for our SCL memory cell[47].

The inverted outputs of the Sense Amplifier (OUTPUT0 through OUTPUT3) indeed provide us with the correct values written within the cells of the accessed word.

As seen in figure 6e which is part of the previous transient simulation , even though a greater discharge of the highly capacitive bitlines is required for a write operation , a write operation can be carried out faster than a read operation[40]. This is because during the Write Driver has much larger current driving capability than that of the cell[39].

As previously mentioned , the speed of the read operation is completely independent of the core tail bias current I_{CORE} shown in figure 53a and depends only on I_{READ} (again shown in figure 53a) as well as on the parasitic capacitances at the nodes BL and BLB. Thus , in order to increase the speed of the read operation , which is the main speed limiting factor of this memory , it is necessary to increase I_{READ} , which can be achieved by increasing the voltage swing at the gate of the NMOS transistor M10 (again shown in figure 53a)[30].

From figure 64e it can be deduced that the minimum time needed for the completion of the write operation is about 33 ns , while for the completion of the read operation is about 40 ns.

As a result , including the minimum time for the precharge (and equalization) operation which is 15 ns , we have that , for a write access we need at least 48 ns in total , whereas for a read access we need at least 55 ns in total. This means a maximum frequency of 10.4 MHz for a write access and a maximum frequency of 9.1 MHz for a read access.

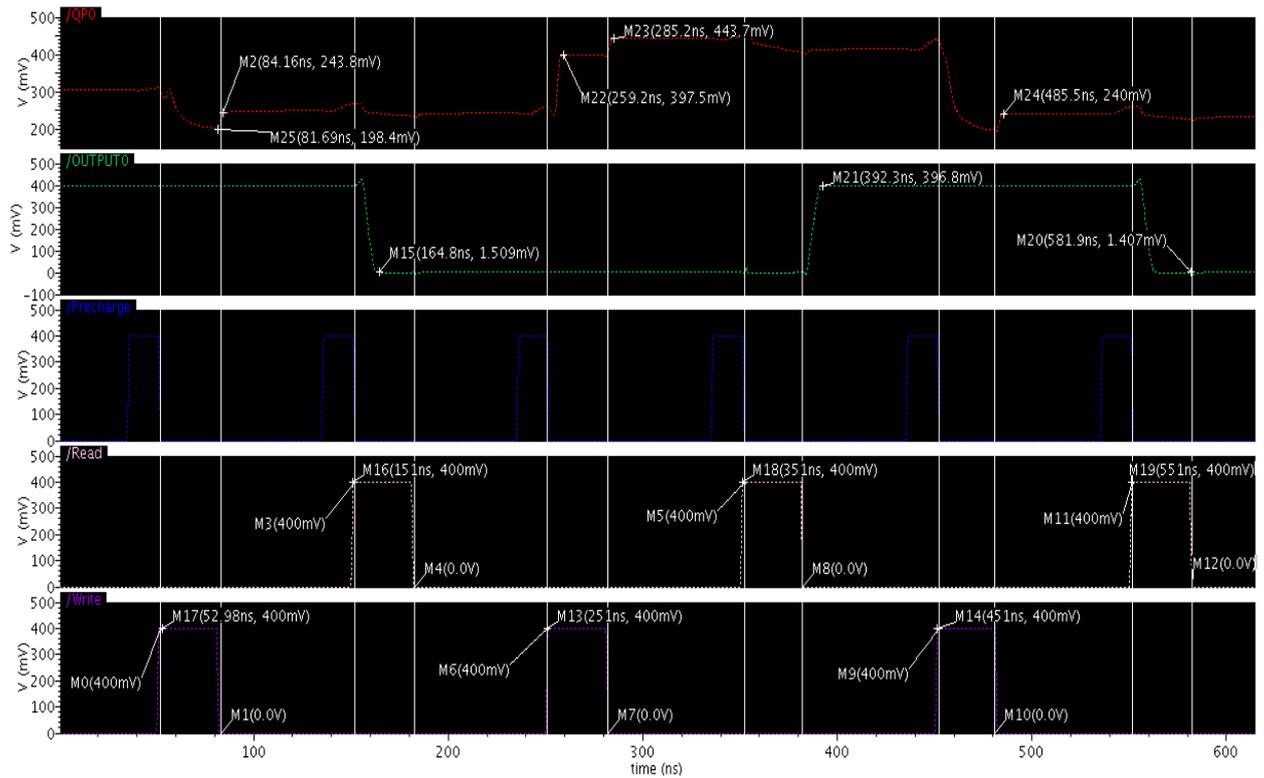


Figure 64e – the time needed for a write and a read operation for the typical NMOS typical PMOS corner

For the case of the slow NMOS slow PMOS corner , from figure 64f , we see that the minimum time needed for the completion of the write operation is about 120 ns , while for the completion of the read operation is about 190 ns.

As a result , including the minimum time for the precharge (and equalization) operation which is now 60 ns , we have that , for a write access we need at least 180 ns in total , whereas for a read access we need at least 250 ns in total. This means a maximum frequency of 2.78 MHz for a write access and a maximum frequency of 2 MHz for a read access.

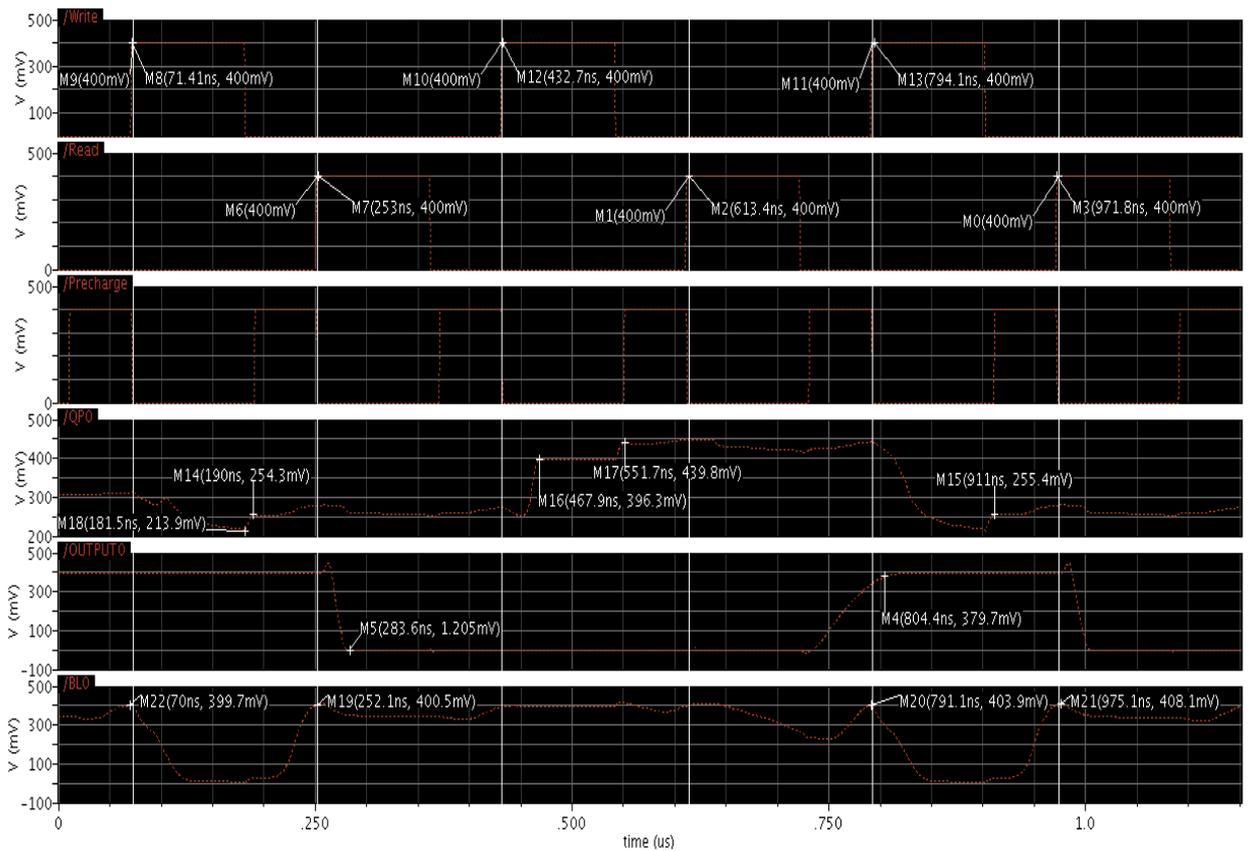


Figure 64f – the time needed for a write and a read operation for the slow NMOS slow PMOS corner

As far as the power consumption is concerned , in order to measure it , we need to be able to measure current from the power supply[52]. In figure 65 we plot the supply current $i_{DD}(t)$ and the instantaneous power $P(t)$ for the whole memory drawn from the power supply , where: $P(t)=i_{DD}(t) \cdot V_{DD}$. Also , the total peak power used within the time interval of our simulation is marked and is $P_{peak} = \max(I_{DD}(t) \cdot V_{DD}) = 4.583 \cdot 10^{-6} \text{ W} \approx 4.6 \mu\text{W}$. It is interesting to notice that the instantaneous power consumed , takes its maximum value at the time of the write operation.

The total average power used over the time interval of this transient simulation which is 700 ns , is the energy consumed over that time interval , divided by time[52]:

$$P_{avg} = \frac{E}{t} = \frac{1}{t} \cdot \int_0^t i_{DD}(t) \cdot V_{DD} dt = \frac{V_{DD}}{t} \cdot \int_0^t i_{DD}(t) dt = \frac{0.4}{700 \cdot 10^{-9}} \cdot \int_0^{700 \cdot 10^{-9}} i_{DD}(t) dt$$

which using the Cadence Virtuoso Analog Environment’s Calculator , is computed as $2.016 \cdot 10^{-7} \text{ W} \approx 0.2 \mu\text{W}$.

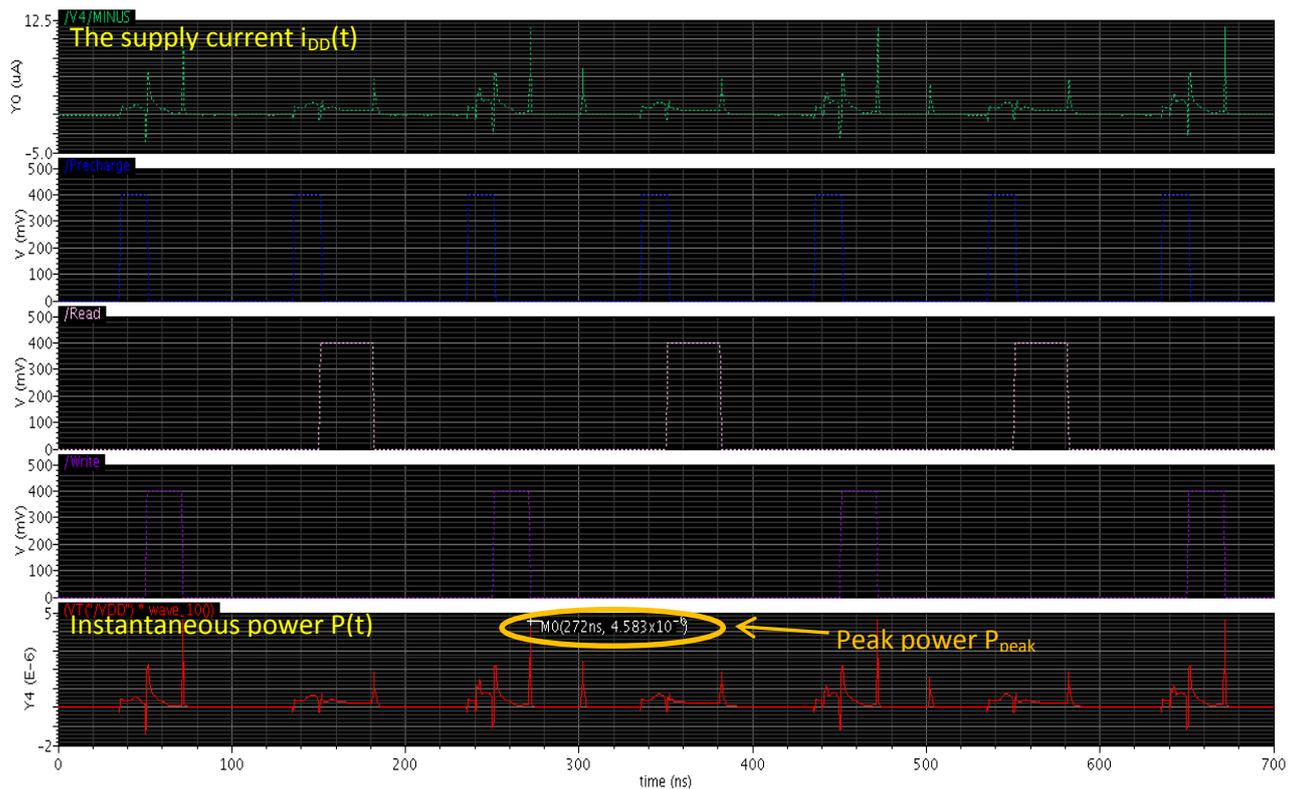


Figure 65 – The instantaneous current and power consumption of the memory array

5. CONCLUSION

In this work , the potential of subthreshold SCL circuits as an alternative solution for implementing ultra-low-power digital systems is explored. After a theoretical examination of the operation of the SCL circuits , and experimental measurements on their Drain-bulk connected PMOS load devices , a 9T SRAM memory cell is developed based on the SCL topology and its operation is demonstrated through the implementation of a small (64-bit) SRAM memory.

REFERENCES

- [1] Eric Vittoz
Low-Power CMOS Circuits Technology , Logic Design and CAD Tools , Chapter 16
“Weak Inversion for Ultimate Low-Power Logic” , CRC Editions
- [2]M.Bucher , D.Kazazis , F.Krummenacher , D.Binkley , D.Foty , Y.Papananos
“Analysis of Transconductances at All Levels of Inversion in Deep Submicron CMOS” , in
Proceedings of 9th International Conference on Electronics , Circuits and Systems (September
25-18 , 2002) , vol.3 , pp. 1183-1186
- [3] Armin Tajalli , Yusuf Leblebici

- “Extreme Low-Power Mixed Signal IC Design” , Springer Editions
- [4]Stephane Badel , Yusuf Leblebici
 “Breaking the Power-Delay Tradeoff: Design of Low-Power High-Speed MOS Current-Mode Logic Circuits Operating with Reduced Supply Voltage” , in IEEE International Symposium on Circuits and Systems(ISCAS) , 2007 , pp. 1871 - 1874
- [5]F.Cannillo , C.Toumazou , T.Sverre Lande
 “Nanopower Subthreshold MCML in Submicrometer CMOS Technology” , in IEEE Transactions on Circuits and Systems , vol.56 , August 2009 , pp.1598 – 1611
- [6]Massimo Alioto , Gaetano Palumbo
 “Power-Aware Design Techniques for Nanometer MOS Current-Mode Logic Gates: A Design Framework” , in IEEE CIRCUITS AND SYSTEMS MAGAZINE , fourth quarter 2006 , pp.41-59
- [7]Massimo Alioto , Yusuf Leblebici
 “Analysis and Design of Ultra-Low Power Subthreshold MCML Gates” , in ISCAS 2009 , pp.2557 – 2560
- [8]Armin Tajalli , Elizabeth Brauer , Yusuf Leblebici , Eric Vittoz
 “Subthreshold Source-Coupled Logic Circuits for Ultra-Low-Power Applications” , in IEEE JOURNAL OF SOLID-STATE CIRCUITS , vol.43 , no.7 , July 2008 , pp.1699 – 1710
- [9]P.R.Gray , P.J.Hurst , S.H.Lewis and R.G.Meyer , Analysis and Design of Analog Integrated Circuits , 4th edition. New York: Wiley , 2000
- [10]C.Enz , E.Vittoz , Charge-Based MOS Transistor Modeling: The EKV Model for Low-Power and RF IC Design. New York: Wiley , 2006
- [11]Franco Maloberti
 Analog Design for CMOS VLSI Systems. 2001 Kluwer Academic Publishers , Boston
- [12]Stephane Badel
 “MOS Current-Mode Logic Standard Cells for High-Speed Low-Noise Applications” , Phd Thesis , July 1 , 2008 , EPFL , Switzerland
- [13]Jason Musicer , Jan Rabaey
 “MOS Current Mode Logic for Low Power , Low Noise CORDIC Computation in Mixed-Signal Environments” , in ISLPED , 2000 , pp.102-107
- [14]Anantha P.Chandrakasan , Robert W.Broderson
 “Minimizing Power Consumption in Digital CMOS Circuits” , in Proceedings of the IEEE , Vol.83 , no.4 , April 1995 , pp.498-523
- [15]Masayuki Mizuno , Koichiro Furuta , Hiroyuki Igura , Hitoshi Abiko , Kazuhiro Okabe , Atsuki Ono and Hachiro Yamada
 “A GHz MOS Adaptive Pipeline Technique Using MOS Current-Mode Logic” , in IEEE JOURNAL OF SOLID-STATE CIRCUITS , vol.31 , no.6 , June 1996
- [16]Armin Tajalli , Yusuf Leblebici
 “Leakage Current Reduction Using Subthreshold Source-Coupled Logic” , in IEEE Transaction on Circuits and Systems-II , vol.56 , no.5 , 2009 , pp.347-351
- [17]Eric Vittoz
 Design of Analog-Digital VLSI Circuits for telecommunications and signal processing , Chapter 3 “Micropower Techniques” , Prentice Hall , pp.53 – 93
- [18] <http://ekv.epfl.ch/site/ekv>
- [19]Armin Tajalli , Eric Vittoz , Yusuf Leblebici
 “Ultra Low Power Subthreshold MOS Current Mode Logic Circuits Using a Novel Load Device Concept” , in 33rd European Solid State Circuits Conference(ESSCIRC) , 2007 , pp.304-307
- [20]Armin Tajalli , Frank Gurkaynak , Yusuf Leblebici
 “Improving the Power-Delay Product in SCL Circuits Using Source Follower Output Stage” , in IEEE International Symposium on Circuits and Systems(ISCAS) 2008 , pp.145-148
- [21]C.A.Mead , Analog VLSI and Neural Systems
 Reading , MA: Addison-Wesley , 1989

- [22]David M. Binkley , Tradeoffs and Optimization in Analog CMOS Design
Wiley
- [23] F. Canillo , C. Toumazou and T.S. Lande
“Bulk-Drain connected load for subthreshold MOS current-mode logic” , in ELECTRONIC LETTERS , 7th June 2007 , vol.43 , no.12
- [24]A. Tajalli , E.Vittoz , Y.Leblebici , E.J. Brauer
“Ultra-low power subthreshold current-mode logic utilising PMOS load device” , in ELECTRONIC LETTERS , 16th August 2007 , vol.43 , no.17
- [25]Christian C.Enz , Francois Krummenacher and Eric A. Vittoz
“An Analytical MOS Transistor Model Valid in All Regions of Operation and Dedicated to Low-Voltage and Low-Current Applications” , in Analog Integrated Circuits and Signal Processing , pp.83-114 , 1995
- [26]F.Cannillo , C.Toumazou
"Nano-power subthreshold current-mode logic in sub-100 nm technologies" , in ELECTRONIC LETTERS , 10th November 2005 , vol.41 , no.23
- [27] Philip E.Allen , Douglas R.Holberg
CMOS Analog Circuit Design , second edition , Oxford University Press , 2002
- [28]R. Jacob Baker
CMOS Circuit Design , Layout , and Simulation , revised second edition , IEEE Press Series on Microelectronic Systems , 2007
- [29]Adel S. Sedra , Kenneth C. Smith
Microelectronic Circuits , Fifth Edition , Oxford University Press , 2004
- [30]Armin Tajalli and Yusuf Leblebici
“Subthreshold SCL for Ultra-Low-Power SRAM and Low-Activity-Rate Digital Systems” , in Proceedings of ESSCIRC , 2009 , pp.164-167
- [31]Matthias Bucher , Cristophe Lallement and Christian C.Enz
“An Efficient Parameter Extraction Methodology for the EKV MOST Model” , in proceedings of the IEEE int. Conference on Microelectronic Test Structures , vol.9 , pp.145-150 , 1996
- [32]Neil H.E.Weste , David Money Harris
CMOS VLSI Design a Circuits and Systems Perspective , fourth edition , Addison Wesley , 2011
- [33]Evert Seevinck , Frans List , Jan Lohstroh
“Static-Noise Margin Analysis of MOS SRAM Cells” , in IEEE Journal of Solid-State Circuits , vol. sc-22 , no.5 , pp.748-754 , october 1987
- [34]Lohstroh , Seevinck , J.de Groot
“Worst-case static noise margin criteria for logic circuits and their mathematical equivalence” , in IEEE J.Solid-State Circuits , vol.SC-18 , no.6 , pp.803-807 , 1983
- [35]Benton H. Calhoun , Anantha P.Chandrakasan
“Static Noise Margin Variation for Sub-threshold SRAM in 65-nm CMOS” , in IEEE Journal of Solid-State Circuits , vol.41 , no.7 , July 2006 , pp.1673 – 1679
- [36]Matthias Bucher
Design of Analog CMOS Circuits , coursenotes , 2008
- [37]Julio Pimentel , Fabio Salazar , Marco Pacheco , Yosef Gavriel
“Very-Low Power Analog Cells in CMOS” , in proc. Of the 43rd IEEE Midwest Symp. On Circuits and Systems , 2000 , pp.328-331
- [38]Stephane Badel , Ilhan Hatirnaz , Yusuf Leblebici
“Semi-Automated Design of a MOS Current-Mode Logic Standard Cell Library from Generic Components” , in Research in Microelectronics and Electronics , 2005 PhD
- [39]Sung-Mo Kang , Yusuf Leblebici
CMOS Digital Integrated Circuits Analysis and Design , McGraw-Hill , 3rd edition , 2002
- [40]Andrei Pavlov

- CMOS SRAM Circuit Design and Parametric Test in Nano-Scaled Technologies , Springer , 2008
- [41]Benton Highsmith Calhoun , Anantha Chandrakasan
"A 256-kb 65-nm Sub-threshold SRAM Design for Ultra-Low-Voltage Operation" , in IEEE Journal of Solid-State Circuits , vol.42 , no.3 , March 2007 , pp.680-688
- [42]Joyce kwong , Anantha P.Chandrakasan
"Advances in Ultra-Low-Voltage Design" , in IEEE Solid-State Circuits Society , October 2008 issue
- [43]Jan M. Rabaey , Anantha Chandrakasan , Borivoje Nikolic
Digital Integrated Circuits , A Design Perspective , Prentice Hall , second edition , 2003
- [44]Benton H. Calhoun , Anantha Chandrakasan
"Analyzing Static Noise Margin for Subthreshold SRAM in 65nm CMOS" , in Proceedings of ESSCIRC , Grenoble , France , 2005 , pp.363 – 366
- [45]Jan Rabaey
Low Power Design Essentials , Springer 2009
- [46]Victor P. Nelson , H.Troy Nagle , Bill D. Carroll , J.David Irwin , Digital Logic Circuit Analysis and Design , Prentice Hall 1995
- [47] Neil H.E.Weste , David Harris , Ayan Banerjee
CMOS VLSI Design a Circuits and Systems Perspective , third edition , Addison Wesley , 2006
- [48]Naveen Verma , Joyce Kwong and Anantha P. Chandrakasan
"Nanometer MOSFET Variation in Minimum Energy Subthreshold Circuits" , in IEEE transactions on electron devices , vol.55 , no.1 , january 2008 , pp.163 - 174
- [49]Alice Wang , Benton Highsmith Calhoun , Anantha P. Chandrakasan
Sub-Threshold Design for Ultra Low-Power Systems , Springer , 2006
- [50]M. Yamaoka , N.Maeda , Y.Shinozaki , Y.Shimazaki K. Nii , S. Shimada , K. Yanagisawa , and T. Kawahara
"Low-Power Embedded SRAM Modules with Expanded Margins for Writing" , in IEEE International Solid-State Circuits Conference (ISSCC) Digest of Technical Papers , Feb. 2005 , pp.480-481
- [51]J. Chen , L.T.Clark , and Y.Cao
"Ultra-low voltage circuit design in the presence of variations" , in IEEE circuits devices Mag. , pp.12-20 , Nov./Dec. 2005
- [52]Erik Brunvand
Digital VLSI Chip Design with Cadence and Synopsys CAD Tools , Addison – Wesley , 2010
- [53]Massimo Alioto , Gaetano Palumbo
Model and Design of Bipolar and MOS Current-Mode Logic , Springer , 2005
- [54]Kurt Hoffman
System Integration , from transistor design to large scale integrated circuits , Wiley , 2004
- [55]M.Bucher , C.Lallement , C.Enz , F.Theodoloz and F.Krummenacher
"The EPFL-EKV MOSFET model equations for simulation , version 2.6" Technical Report , EPFL , July 1998 , Revision II