Fluctuation Limits on Scaling of CMOS SRAMs

Azeez Bhavnagarwala and James Meindl
MiRC, Georgia Institute of Technology
Atlanta, GA 30332

Abstract†

Reductions in CMOS SRAM cell static noise margin (SNM) due to intrinsic threshold voltage fluctuations in uniformly doped minimum geometry cell transistors are reported. Standard deviations in SNM due to random dopant placement alone are projected to exceed 17% of the nominal SNM for CMOS generations beyond 100nm posing severe barriers to scaling of supply voltage and channel length in SRAM dominated CMOS ASICs and microprocessors.

1. Introduction

With scaling of MOSFET dimensions, microscopic variations [1-3] in number and location of dopant atoms in the channel region of the device induce increasingly limiting electrical deviations (Fig. 1, [3]) in device characteristics. These atomic level intrinsic fluctuations cannot be eliminated by external control of the manufacturing process and are most pronounced in minimum geometry transistors commonly used in area-constrained circuits such as SRAM cells [4]. Intrinsic fluctuations are independent of transistor location on a chip with the threshold mismatch between neighboring cell transistors due to intrinsic fluctuations typically contributing to larger reductions in SNM than macroscopic manufacturing related variations [5]. To quantify the joint impact of intrinsic device fluctuations in each cell transistor on SRAM cell stability and scaling, short-channel physical and stochastic models that integrate the unique contributions of each cell transistor on cell SNM are necessary. These are reported in this work for the first time and are engaged to project the scaling limits of 6T CMOS SRAMs.

2. A 6T CMOS SRAM SNM Model

The SNM of a CMOS SRAM cell is defined [6,7] as the minimum DC noise voltage necessary at both of the two cell storage nodes, during a read access, to flip the state of a cell. The side of the smaller of the two maximum squares nested between the static characteristics of the two cell inverters (Fig 2b) thus equals cell SNM. Analytical expressions for cell SNM are derived by approximating the cell static characteristics FI and FII in the neighborhood of the corners of the maximum square, **B** & **A'** using (1)-(2) in the Appendix, and then solving for the side of the square, S (3) in this neighborhood, when the slopes of static characteristics that touch the diagonally opposite corners of the nested square (4) (Fig. 3) become equal. This implicit SNM model is derived using the physical, short-channel *Transregional* MOSFET drain current model [8,9] that considers threshold roll-off and high field

† This work was supported by the Defense Advanced Research Project Agency (Contract: F3361595C1623) and the Semiconductor Research Corporation (SJ-374-001)

effects on carrier mobility. Verified with HSPICE simulations for a 0.18 μm industrial CMOS process in Fig. 4, this SNM model is used to assess the individual sensitivities of SNM to V_t variations in each transistor in Fig. 5 for the 50nm generation. This plot reveals cell SNM to be *most sensitive* to threshold voltage fluctuations in the access NFETs and *least sensitive* to the fluctuations in the pull-up PFET device.

3. Stochastic Distributions of Cell SNM

Using the 'cube model' [3], a MOSFET may be viewed as an array of MOS capacitors separating the source from the drain. The probability density function (pdf) of effective doping concentration $f(n_a)$, the doping density required for a uniformly doped MOSFET without fluctuations to achieve a specific threshold voltage, is derived from a fundamental device analysis in [10]. Based on the effective doping concentration, the threshold voltage distribution density function F_{Vts} is also derived in [10], approximated as gaussian (7) and is plotted in Fig. 1 for 1997 NTRS generations. The threshold voltages of the cell transistors may be viewed as independent random variables and the SNM, given by (5) as a stochastic function of these. Given their statistical independence, total fluctuation in SNM is modeled as the sum of the fluctuations in SNM produced individually by threshold voltage fluctuations in each of the cell transistors (8). The joint pdf of cell SNM is modeled by convolving [11] the individual independent pdfs (9) of SNM, $f(SNM_{ij})$, produced by fluctuations in threshold voltage in each of the cell transistors Mij (10). Cell SNM decreases as device threshold voltages are lowered, with increases in temperature. The joint pdfs for SNM (Fig. 6) thus use the input data from Table I, at high temperature (100°C), as this constitutes the worst case for cell stability.

4. CMOS SRAM Stability Projections

Expressed as a percentage of the SNM, the standard deviation in SNM (Table-II) due to intrinsic fluctuations increases alarmingly from 3% (8mV) for the 250nm generation to 26% (16mV) for the 50nm generation imposing restrictive limits on the percentage of array cells that can hold data.

5. Summary

Intrinsic fluctuations in device V_t are projected to limit CMOS SRAM scaling. Device and circuit techniques that suppress device fluctuations due to random placement of dopant atoms will be will become increasingly necessary to enable normal SRAM operation.

References:

[1] R. W. Keyes, *App. Phys.,* 8, pp. 251-259, 1975.

[2] T. Mizuno et al, *Symp. VLSI Tech.,* Jun. 1993, pp. 41-42

[3] J. Meindl et al, *ISSCC Dig. of Tech Pprs* Feb. 1997, pp. 232-233

[4] K. Itoh, "*Low Power Memory Dsgn.,"* Tutorial, *1997 ISLPED*

[5] D. Burnett et al, *Symp. VLSI Tech.,* June 1994, pp. 15-16.

[6] J. Lohstroh et al, *IEEE JSSC,* Vol. SC-18, No 6, Dec. 1983, pp 803-806.

[7] E. Seevinck et al, *IEEE JSSC*, Vol. SC-22, No. 5, Oct 1987, pp. 748-754

[8] B. Austin, et al, *Proc. 11th IEEE ASIC Conf.,* pp. 125-129, Sept. 1998.

[9] A. Agrawal, et al, *Proc. 23rd ESSDERC,* pp. 919-926, Sept. 1993.

[10] X. Tang, et al, *IEEE TVLSI Sys.,* Vol.5, pp. 369-376, Dec. 1997.

[11] P. Peebles Jr., New York, NY, *McGraw Hill*, 1987, Chapt. 3-4.

Appendix:

Each SRAM cell transistor is identified by 2 subscripts M_{ij}, with $i \in a,n,p$ identifying a transistor as access, pull-down or pull-up and $j \in R,L$ identifying the left or right side of the cell.

*In the neighborhood of **B** and **A'** respectively (Fig 2b).*

(1a) $I_{nR} \cong I_{pR}$;**(1b)** $I_{nL} \cong I_{aL}$

(2a) $$F1 = V_L + V_{tpR} + \left(V_{dd} - V_L - V_{tpR}\right)\sqrt{1 - \frac{2I_{nR}}{\frac{W_{pR}\mu_{op}C_{ox}}{L\left(1+\theta_p\left[V_{dd}-V_L-V_{tpR}\right]\right)}\left(V_{dd}-V_{in}-V_{tpR}\right)^2}}$$

(2b) $$FII \cong V_R = \frac{I_{aL}}{\frac{W_{nL}\mu_{on}C_{ox}V_L}{L\left(1+\theta_n\left[V_{dd}-V_{tnL}\right]\right)\left(1+\frac{V_L}{LE_{cn}}\right)}} + \frac{V_L}{2} + V_{tnL}$$

Substituting. (2) in (3) and simultaneously solving (3) and (4) numerically for S=SNM we get (5): SNM_L corresponding to maximum side of the square between **B** and **A'**:

Side of square in Fig. 3:

(3) $S = FI(V_L) - FII(V_L - S)$

Maximum square:

(4) $$\left.\frac{\partial S}{\partial V_L}\right|_{S=SNM,\ V_L=V_x} = \frac{\partial FI(V_L)}{\partial V_L} - \frac{\partial FII(V_L - SNM)}{\partial V_L} = 0$$

(5) $$SNM_L = V_X + V_{tpR} + \left(V_{dd} - V_X - V_{tpR}\right)\sqrt{1 - \frac{2I_{nRx}}{\frac{W_{pR}\mu_{op}C_{ox}}{L\left(1+\theta_p\left[V_{dd}-V_X-V_{tpR}\right]\right)}\left(V_{dd}-V_X-V_{tpR}\right)^2}}$$
$$-\left[\frac{I_{aLx}}{\frac{W_{nL}\mu_{on}C_{ox}\left(V_X - SNM_L\right)}{L\left(1+\theta_n\left[V_{dd}-V_{tnL}\right]\right)\left(1+\frac{V_X-SNM_L}{LE_{cn}}\right)}} + \frac{V_X - SNM_L}{2} + V_{tnL}\right]$$

SNM_R corresponding to the side of square between **B'** and **A** in Fig. 2b is obtained by swapping right (R) and left (L) subscripts in (5)

(6) $SNM = \min\left[SNM_L, SNM_R\right]$

Threshold voltage and cell SNM stochastic distributions:

(7) $$f\left(V_{Tij}\right) \cong \frac{1}{\sqrt{2\pi\sigma^2_{V_{Tij}}}} e^{-\frac{\left(V_{Tij}-V_{T0}\right)^2}{2\sigma^2_{V_{Tij}}}}$$

(8) $$\Delta SNM_L = \Delta SNM_{aL} + \Delta SNM_{nL} + \Delta SNM_{nR} + \Delta SNM_{pR}$$

(9) $$f\left(SNM_{ij}\right) = \frac{f\left(V_{Tij}\right)}{\frac{\partial (SNM_L)}{\partial V_{Tij}}}$$

(10) $$f\left(SNM_L\right) = f\left(SNM_{aL}\right) * f\left(SNM_{nL}\right) * f\left(SNM_{nR}\right) * f\left(SNM_{pR}\right)$$

Table I: 1997 NTRS input parameters for calculating threshold voltage and SNM distributions due to intrinsic fluctuations [3]

year	'97	'01	'06	'12
L (nm)	250	150	100	50
T_{ox} (nm)	4.5	2.8	1.7	0.8
V_{dd} (V)	2.2	1.5	1.0	0.5
V_{th} (V)	0.45	0.35	0.30	0.2
N_a (cm^{-3})	6e17	9e17	2e18	5e18
σVth (mV)	19	25	28	32

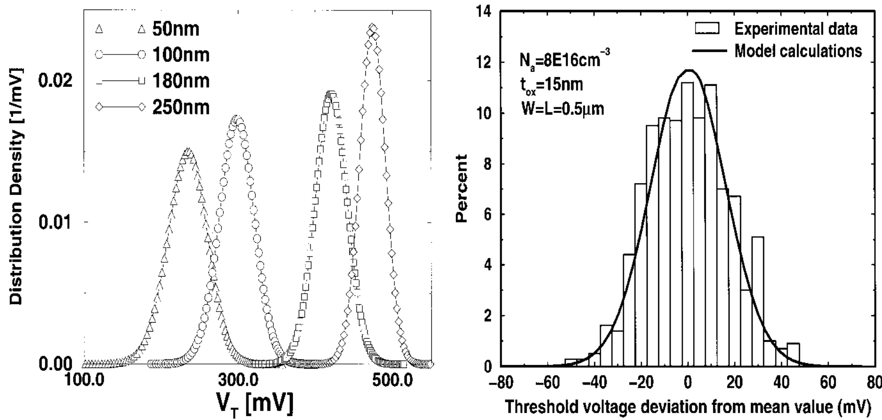
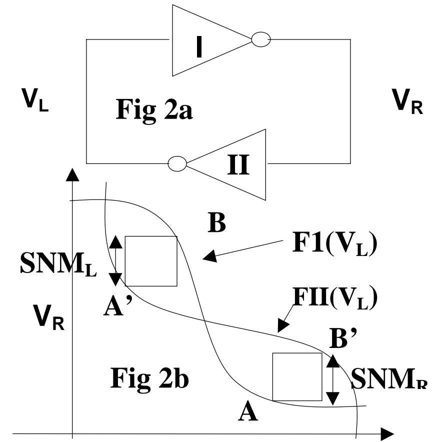


Figure 1a,b: Threshold voltage distribution functions for representative 1997 NTRS generations and verification of models in [3] with experimental data



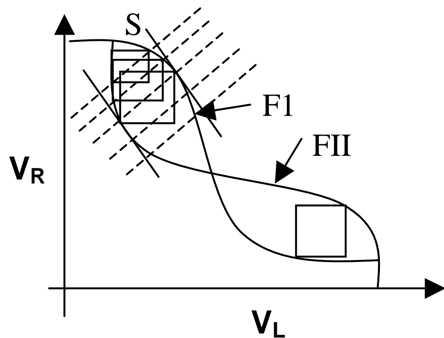


Figure 3: The side of the nested square, S given by (3) reaches it's maximum when (4) is satisfied

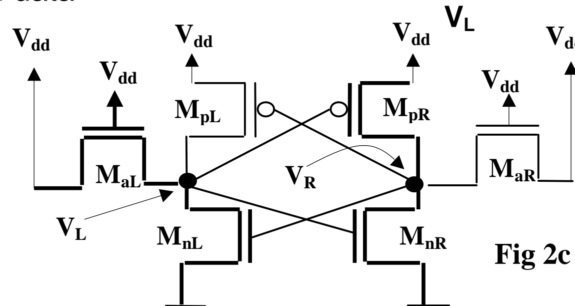


Figure 2 a, b, c: During read access, the set of 4 highlighted transistors contributes to SNM_L. A left-right symmetrical set contributes to SNM_R

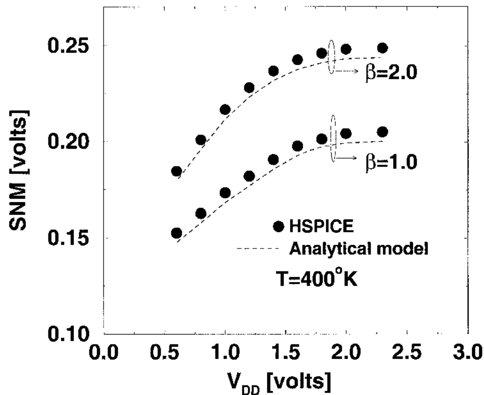


Figure 4. SNM_L dependence on V_{dd} using analytical model and HSPICE.

Table-II: Calculated variations in SNM at 400°K, β=1

F	250 nm	100 nm	70 nm	50 nm
SNM (mV)	256	113	82	60
σ SNM (mV)	8	12	14	16

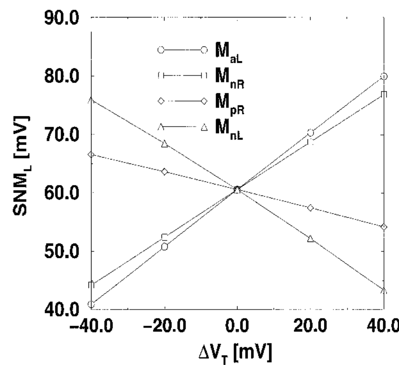


Figure 5: Sensitivities of SNM_L to variations in device threshold voltage of each cell transistor (50nm). β=1

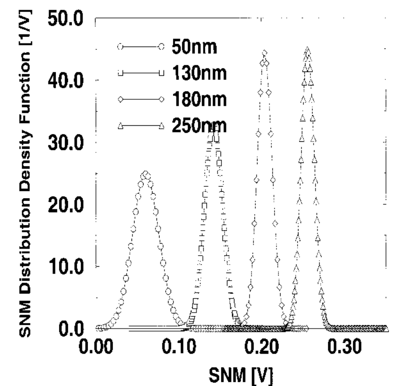


Figure 6: Distribution density function in SNM due to variations in each cell transistor. β=1.