Statistical Design and Optimization of SRAM Cell for Yield Enhancement

Saibal Mukhopadhyay, Hamid Mahmoodi, and Kaushik Roy
Dept. of Electrical and Computer Engineering, Purdue University, West Lafayette, IN-47907, USA
<sm, mahmoodi, kaushik>@ecn.purdue.edu

ABSTRACT

In this paper, we have analyzed and modeled the failure probabilities of SRAM cells due to process parameter variations. A method to predict the yield of a memory chip based on the cell failure probability is proposed. The developed method is used in an early stage of a design cycle to minimize memory failure probability by statistically sizing of SRAM cell.

1. INTRODUCTION

In nano-scaled devices, the random variations in the number and location of dopant atoms in the channel region of the device cause random variations in transistor threshold voltage [1–3], known as "Random (or discrete) Dopant Fluctuation" (RDF). The impacts of random dopant effect are most pronounced in minimum-geometry transistors commonly used in area-constrained circuits such as SRAM cells [4]. This can result in the threshold voltage mismatch between the neighboring transistors in a cell, resulting in the failure of a cell. Since these failures are caused by the variation in the device parameters, these are known as the *parametric failures*. A failure in any of the cells in a column (or row) of the memory will make that column (or row) faulty. If the number of faulty columns (or rows) in a memory chip is larger than the number of available redundant columns (or rows), then the chip is considered to be faulty. Hence, the failure probability of a cell is directly related to the yield of a memory chip.

The parametric variations, and in particular the Vt fluctuation due to RDF, is a strong function of the size of the different transistors in the cell (channel length (L), width (W)), collectively called as the cell parameters. Hence, the failure probability of a memory can be reduced by optimally designing these parameters. However, any such optimization has to consider its impact on the overall area and the leakage of the SRAM array. Moreover, the memory organization (i.e. # of row (N_{ROW}) and # of row (N_{COL}), # of redundant column (N_{RC})) will also have a strong impact on the memory failure probability. Hence, a statistical design and optimization of the SRAM cell and memory organization is very important to reduce the memory failure and improve the yield in nano-scaled SRAM.

In this paper, we have developed a methodology to optimize the parameters of an SRAM cell and the memory organization to reduce the memory failure probability (constrained by the overall memory area and leakage power) and improve the yield in nanometer regime. The method is developed considering the on-die Vt variation, but can be extended to consider on-die L and W variation.

In our SRAM cell (Fig.1), we have used transistors of 50nm gate length (*Leff=25nm*) designed using MEDICI [6-7]. In our analysis, we have used the short channel MOSFET theory to model the currents and threshold voltage considering the device geometry and doping profile [3,5] (Fig. 2).

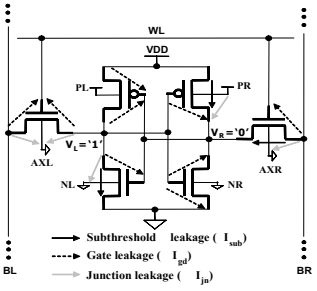


Fig. 1: SRAM Cell

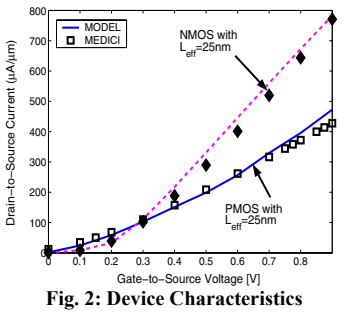
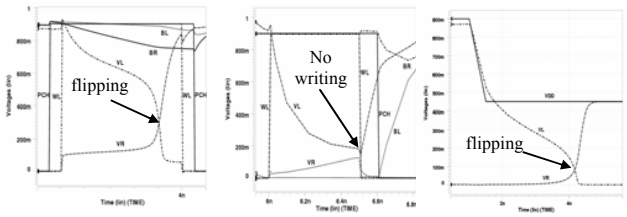


Fig. 2: Device Characteristics



(a) read failure **(b) write failure** **(c) hold failure**

Fig. 3: Unstable read, write, and hold failures in SRAM cell

2. DISTRIBUTION OF THE INTRINSIC Vt VARIATION

In an SRAM cell, the threshold voltage (Vt) fluctuations (δVt) of the cell transistors are considered as six *independent Gaussian random variables* (mean=0) [1]. The standard deviation of the Vt fluctuation (σ_{Vt}) depends on the manufacturing process, doping profile, and the transistor sizing. In the proposed method, σ_{Vt} for a minimum sized transistor (σ_{Vt0}) is an input parameter and the dependence of σ_{Vt} on the transistor size is given by [3]:

$$\sigma_{Vt} = \sigma_{Vt0}\sqrt{(L_{min}/L)(W_{min}/W)} \qquad (1)$$

3. MODELING OF SRAM FAILURE

The failures due to parametric variations in a SRAM cell are principally caused by: (1) An increase in the access time of the cell resulting in a violation of delay requirement, defined as *access time failure*, (2) Destructive read (i.e. flipping of the stored data in a cell known as the *read failure*) and/or unsuccessful write (inability to write to a cell defined as the *write failure*), resulting in a dynamic stability failure (Fig. 3) and, (3) The destruction of the cell content in the standby mode with the application of a pre-specified (designed) lower supply voltage *(V_{HOLD}),* known as *hold-stability failure* (Fig. 3). In a die, failures are principally caused by the mismatch in the device parameters (*L, W, Vt)* of different transistors (intra-die) in the cell. Such device mismatch changes the strength of the different transistors resulting in different failure events. The principal source of the device mismatch is the intrinsic fluctuation of the Vt of different transistors due to RDF [1]. Hence, in this work we have considered the Vt variation due to RDF as the major source of intra-die variation. The proposed method can also be extended to include L and W variation.

3.1. Modeling Methodology

In this section, we will summarize the key mathematical bases used to estimate the failure probabilities. Let us consider $y=f(x_1, \ldots, x_n)$ as a function, where $x_1, \ldots, x_n$ are *independent Gaussian random variables* with mean $\eta_1,\ldots,\eta_2$ and standard deviation (STD) $\sigma_1,\ldots,\sigma_2$. The mean (μ_y) and the STD (σ_y) of the random variable y can be estimated as (using multi-variable Taylor-series expansion) [8]:

$$\mu_y = f(\eta_1,\ldots,\eta_n) + \frac{1}{2}\sum_{i=1}^{n}\left.\frac{\partial^2 f(x_1,\ldots,x_n)}{\partial(x_i)^2}\right|_{\eta_i}\sigma_i^2$$
$$\sigma_y^2 = \sum_{i=1}^{n}\left(\left.\frac{\partial f(x_1,\ldots,x_n)}{\partial(x_i)}\right|_{\eta_i}\right)^2\sigma_i^2 \qquad (2)$$

Assuming the Probability Distribution Function (PDF) of y to be also Gaussian ($N_y(y{:}\mu_y, \sigma_y)$), the probability of ($y > Y_0$) is given by:

$$P[y>Y_0] = \int_{y=Y_0}^{\infty} N_y(y:\mu_y,\sigma_y)dy = 1 - \int_{y=-\infty}^{Y_0} N_y(y)dy = 1-\Phi_y(Y_0) \qquad (3)$$

where, Φ_y is the Cumulative Distribution Function (CDF) of y.

Let us assume $y=f(x_1,\ldots,x_n)$ and $z=g(x_1,\ldots,x_n)$ are two Gaussian random variables $N_y(y{:}\mu_y, \sigma_y)$ and $N_z(y{:}\mu_z, \sigma_z)$, respectively. The probability of ($y > Y_0$ & $z > Z_0$) is given by:

$$\begin{aligned} P[(y>Y_0)\&(z>Z_0)] &= 1-P[(y\le Y_0)+(z\le Z_0)] \\ &= 1-\{P[y\le Y_0]+P[z\le Z_0]-P[(y\le Y_0)\&(z\le Z_0)]\} \\ &= \{P[y>Y_0]+P[z>Z_0]-1\}+\Phi_{y,z}(Y_0,Z_0) \end{aligned} \qquad (4)$$

where, $\Phi_{y,z}(y,z)$ is the joint CDF of y and z. In order to evaluate $\Phi_{y,z}(y,z)$ the correlation coefficient between y and z needs to be computed. The correlation coefficient (ρ,) is given by:

$$\rho = \frac{E(yz)-E(y)E(z)}{\sigma(y)\sigma(z)} = \frac{E(yz)-\mu_y\mu_z}{\sigma_y\sigma_z}$$

$$\begin{aligned} E(yz) &= f(\eta_1,\ldots,\eta_n)g(\eta_1,\ldots,\eta_n) + \frac{1}{2}\sum_{i=1}^{n}\frac{\partial^2(fg)}{\partial(x_i)^2}\sigma_i^2 \\ &= f_0 g_0 + \frac{1}{2}\sum_{i=1}^{n}\left(g_0\frac{\partial^2 f}{\partial x_i^2} + 2\frac{\partial f}{\partial x_i}\frac{\partial g}{\partial x_i} + f_0\frac{\partial^2 g}{\partial x_i^2}\right)\sigma_i^2 \end{aligned} \qquad (5)$$

The above results will be used in this paper to estimate the failure probabilities of different events.

3.2. Read Stability Failure (R_F)

While reading the cell shown in Fig. 1 (V_L='*1*' & V_R='*0*'), due to the voltage divider action between AX_R and N_R, V_R increases to a positive value V_{READ}. If V_{READ} is higher than the trip point of the inverter P_L-N_L (V_{TRIPRD}), then the cell flips while reading the cell (Fig. 3(a)) [9]. This represents a *read failure* (R_F) event. Hence, the read-failure probability (P_{RF}) is given by:

$$P_{RF} = P[V_{READ} > V_{TRIPRD}] \qquad (6)$$

V_{TRIPRD} and V_{READ} can be obtained by solving KCL at node L and R, respectively. The estimated value of V_{TRIPRD} and V_{READ} closely follows the MEDICI simulation result (Fig. 4). Assuming the PDF of V_{READ} (=$N_{RD}(v_{READ})$) and V_{TRIP} (=$N_{TRIP}(v_{TRIP})$) as Gaussian distributions with the means the variances obtained using (2) (Fig. 4). P_{RF} is given by:

$$P_{RF} = P[Z_R \equiv (V_{READ} - V_{TRIPRD}) > 0] = 1-\Phi_{ZR}(0)$$
$$\text{where, } \eta_{ZR} = \eta_{V_{READ}} - \eta_{V_{TRIP}} \text{ and } \sigma_{ZR}^2 = \sigma_{V_{READ}}^2 + \sigma_{V_{TRIPRD}}^2 \qquad (7)$$

3.3. Write Stability Failure (W_F)

While writing a '0' to a cell storing '1', the node V_L gets discharged through BL to a low value (V_{WR}) determined by the voltage division between the PMOS P_L and the access transistor AX_L [9]. If V_L can-not be reduced below the trip point of the inverter P_R-N_R (V_{TRIPWR}), within the time when word-line is high (T_{WL}), then a write failure occurs (Fig. 3b). The write-failure probability (P_{WF}) is given by:

$$P_{WF} = P[(T_{WRITE} > T_{WL})] \qquad (8)$$

where, T_{WRITE} is the time required to pull down V_L from V_{DD} to V_{TRIPWR}. T_{WRITE} is obtained by solving:

$$T_{WRITE} = \begin{cases} \left|\int_{V_{DD}}^{V_{TRIP}} \frac{C_L(V_L)dV_L}{I_{in(L)}(V_L) - I_{out(L)}(V_L)}\right|; & if\,(V_{WR} < V_{TRIPWR}) \\ \infty; & if\,(V_{WR} \ge V_{TRIPWR}) \end{cases} \qquad (9)$$

$I_{in(L)}$ = current into L $\approx I_{dsPL}$, $I_{out(L)}$ = current out of L $\approx I_{dsAXL}$

where C_L is the net capacitance at the node L. V_{WR} can be obtained by solving KCL at node L & R [9]. V_{TRIPWR} can be obtained by solving for trip-point of the inverter P_R-N_R. T_{WRITE} obtained using (9) closely matches the MEDICI simulation result with Vt variation of different transistors (Fig. 5a). Using (2), we can estimate the mean (η_{TWR}) and the standard-deviation (σ_{TWR}) and approximate its pdf as a Gaussian one ($f_{WR}(t_{WR})$) (Fig. 5b). However, most of the write-failures originate from the 'tail' of the distribution function. Hence, to improve the accuracy of the model at the tail region, we can use a non-central F distribution [8]. Using the PDF (Gaussian/non-central F) of $T_{WRITE}(N_{WR}(t_{WR})$, the P_{WF} can be estimated using (3).

3.4. Access Time Failure (A_F)

The cell access time (T_{ACCESS}) is defined as the time required to produce a pre-specified voltage difference ($\Delta_{MIN} \approx 0.1 V_{DD}$) between two bit-lines (bit-differential). If due to Vt variation, the access time of the cell is longer than the maximum tolerable limit (T_{MAX}), an access time failure is said to have occurred. The probability of access time failure (P_{AF}) of a cell is given by:

$$P_{AF} = P(T_{ACCESS} > T_{MAX}) \qquad (10)$$

While reading the cell storing V_L='*1*' and V_R='*0*' (Fig.1, Fig.3), bit-line BR will discharge through AX_R and N_R (by the current I_{BR}).

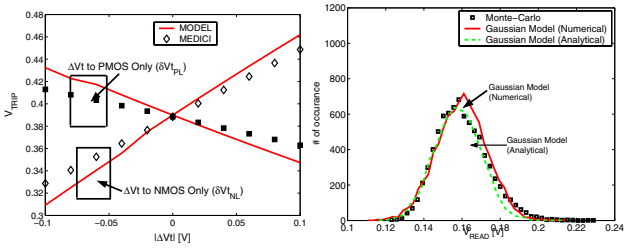


Fig. 4: (a) Variation of V_{TRIP} of P_L-N_L (b) Distributions of V_{READ}

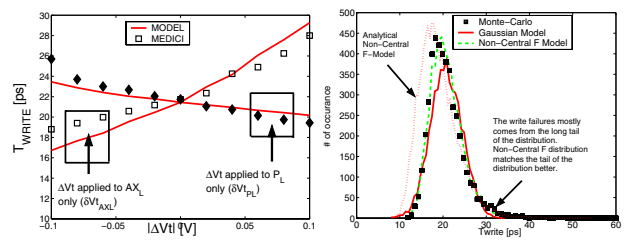


Fig. 5: (a) T_{WRITE} variation with δVt (b) distribution of T_{WRITE}

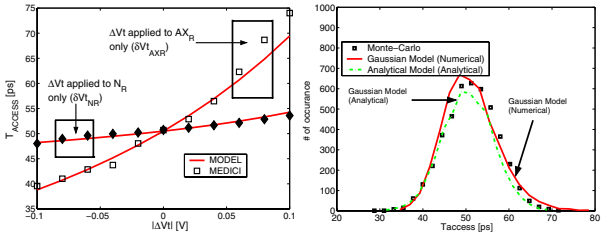


Fig. 6: (a) T_{ACCESS} variation with δVt (b) distribution of T_{ACCESS}

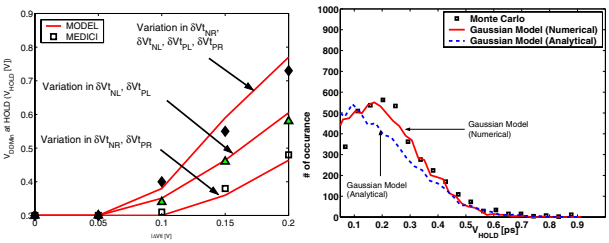


Fig. 7: (a) V_{DDHmin} variation with δVt (b) distribution of V_{DDHmin}

Simultaneously, BL will discharge by the leakage of AX_L of all the cells (I_{BL}) connected to BL. Hence, the access time is given by:

$$T_{ACCESS} = \frac{C_{BR}C_{BL}\Delta_{MIN}}{C_{BL}I_{BR} - C_{BR}I_{BL}} = \frac{C_B\Delta_{MIN}}{I_{dsatAXR} - \sum_{i=1,..,N} I_{subAXL\,(i)}} \quad (11)$$

where, N is the #of cells attached to a bit-line (or column), $C_{BR}/_{BL}$ is the bit-line capacitance (assumed to be equal). The access time given by (11) closely follows the MEDICI simulation result (Fig. 6a). The PDF of T_{ACCESS} can be approximated as a Gaussian one with the mean (η_{TAC}) and the standard deviation (σ_{TAC}) obtained from (2) (Fig. 6b). Using the derived PDF ($N_{TACCESS}(t_{ACCESS})$), P_{AF} can be estimated using (3).

3.5. Hold Stability Failure (H_F)

In the stand-by mode, the V_{DD} of the cell is reduced to reduce the leakage power consumption. However, if the lowering of V_{DD} causes the data stored in the cell to be destroyed, then the cell is said to have failed in the hold-mode [10] (Fig. 3c). Hence, for a hold-failure event, the minimum supply voltage that can be applied to the cell in the hold-mode (V_{DDHmin}), without destroying the data, is higher than the designed stand-by mode supply voltage (V_{HOLD}). Thus, the probability of hold-stability failure (P_{HF}) is given by:

$$P_{HF} = P\left[V_{DDHmin} > V_{HOLD}\right] \quad (12)$$

Lowering the V_{DD} of the cell (say V_{DDH} represents the cell V_{DD} at the hold mode) reduces the voltage at the node storing '1' (V_L in Fig. 1). Due to leakage of N_L, V_L will be less than V_{DDH} for low V_{DDH}. The hold-failure occurs if $V_L < V_{TRIP}$ of P_R-N_R. Hence, V_{DDHmin} can be obtained by numerically solving:

$$V_L\left(V_{DDHmin}, \delta Vt_{PL}, \delta Vt_{NL}\right) = V_{TRIP}\left(V_{DDHmin}, \delta Vt_{PR}, \delta Vt_{NR}\right) \quad (13)$$

The estimated value of V_{DDHmin} using (13) closely follows the values obtained from MEDICI simulation (Fig. 7a). The distribution of V_{DDHmin} can be approximated as a Gaussian one with mean and variance obtained using (6) (Fig. 7b). Using the Gaussian pdf for V_{DDHmin} the P_{HF} can be estimated (using (3)).

3.6. Estimation of Overall Cell Failure Probability (P_F)

The overall failure probability is given by:

$$\begin{aligned} P_F &= P[Fail] = P[A_F + R_F + W_F + H_F] = P_{AF} + P_{RF} + P_{WF} + P_{HF} \\ &- P[A_F R_F] - P[A_F W_F] - P[A_F H_F] - P[R_F W_F] - P[R_F H_F] - P[W_F H_F] \\ &+ P[A_F R_F W_F] + P[A_F R_F H_F] + P[R_F W_F H_F] + P[W_F H_F A_F] - P[All] \end{aligned} \quad (14)$$

An accurate estimate of the probability of joint events is possible by constructing the joint PDF representing two events using the procedure given in (4). We have also assumed that probabilities of simultaneous occurrence of more than two events are negligible (≈0). The estimated probabilities match the Monte-Carlo results very closely.

3.7. Estimation of Column and Memory Failure Probability

The failure probability of column (P_{COL}) (or row (P_{ROW})) is defined as the probability that any of the cells (out of N cells) in that column (or row) fails. Assuming a column redundancy, the probability of failure of a memory chip (P_{MEM}) designed with N_{COL} number of columns and N_{RC} number of redundant columns, is defined as the probability that more than N_{RC} (i.e. at least $N_{RC}+1$) columns fail. Hence, P_{COL} and P_{MEM} can be given by:

$$P_{COL} = 1-(1-P_F)^N \text{ and } P_{MEM} = \sum_{i=N_{RC}+1}^{N_{COL}} \binom{N_{COL}}{i} P_{COL}^{i} (1-P_{COL})^{N_{COL}-i} \quad (15)$$

The exact estimation of the different failure probabilities requires numerical solutions of the KCL at different nodes. In order to reduce the computation complexity, analytical models of different failure probabilities were also obtained using simplified long-channel current equations. The distributions of V_{READ}, V_{TRIPRD}, T_{ACCESS}, T_{WRITE} and V_{HOLD} using the analytical models are also shown in Figs. 4-7.

4. SENSITIVITY ANALYSIS OF FAILURE PROBABILITY

Fig. 8 shows that a weak access transistor (small W_{nax}) reduces P_{RF} (V_{READ} decreases); however, it increases P_{AF} and P_{WF} (Fig. 8) and has very small impact on P_{HF}. Reducing the strength of the PMOS pull-up transistors (by decreasing W_p) reduces P_{WF} (reducing I_{dsPL}), but increases P_{RF} (lowers V_{TRIPRD}). P_{AF} does not depend strongly on PMOS strength (Fig. 8). P_{HF} improves with an increase in W_p as the node L is more strongly coupled to the supply voltage ($V_L \rightarrow V_{DDH}$). Increasing W_{npd} increases the strength of pull-down NMOSs (N_L & N_R)). This reduces P_{RF} ($V_{READ} \downarrow$) and P_{AF} by increasing the strength of N_R (Fig. 8). Increase in width of N_R has little impact on P_{WF}. Although, it slightly increases the nominal value of T_{WRITE}, the reduction of σ_{VT} of N_R (see (1)) tends to reduce σ_{TWRITE} and hence P_{WF} remains almost constant. An increase in the V_{TRIP} of P_R-N_R initially reduces P_{HF} with the increase in W_{npd}. However, a higher width of N_L reduces V_L (from the applied V_{DDH}) due to an increase the leakage of N_L. Consequently, a very high W_{npd} increases the P_{HF}.

5. STATISTICAL OPTIMIZATION OF SRAM CELL

5.1. SRAM Yield Estimation Model

To estimate the yield, we have used Monte-Carlo simulations for *inter-die* distributions of L, W and Vt (assumed to be Gaussian).For each inter-die values of the parameters (say L_{INTER}, W_{INTER} and Vt_{INTER}) we estimate P_F, P_{COL} and P_{MEM} considering the intra-die distribution of δVt. Finally, the yield is defined as:

$$Yield = 1 - \left(\sum_{INTER} P_{MEM}\left(L_{INTER}, W_{INTER}, Vt_{INTER}\right) \Big/ N_{INTER} \right) \quad (16)$$

where, N_{INTER} is the total number of inter-die Monte-Carlo simulations (i.e. total number of chips).

In order to maximize the yield P_{MEM} needs to be minimized. This requires optimum design of the cell configurations (i.e. length and width of transistors) and the number of redundant columns (N_{RC}). However, such an optimization of P_{MEM} has to consider the impact on the total leakage and the total area (A_{MEM}).

5.2. Estimation of Leakage in SRAM

The total leakage in a cell is principally consist of the subthreshold leakage, the gate leakage, and the junction BTBT leakage through different transistors in the cell (Fig. 1). The leakage current expressions presented in [5] are used to evaluate the different leakage components. The total memory leakage ($I_{LeakMem}$) is the summation of leakage of individual cells.

5.3. Area Estimation of SRAM

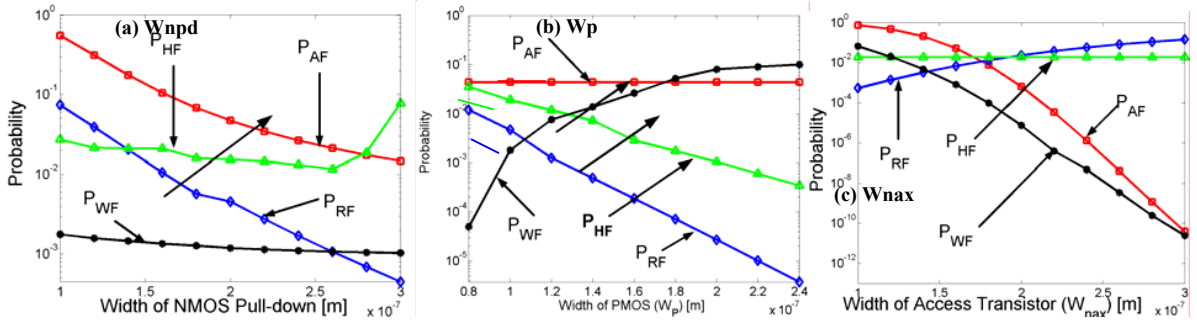


Fig. 8: Variation of Cell Failure Probabilities with Cell structure

The total memory area is given by:

$$A_{actual} = N_{ROW} N_{COL} A_{cell}; A_{redundant} = N_{ROW} N_{RC} A_{cell}$$
$$A_{MEM} = A_{actual} + A_{redundant} = N_{ROW} (N_{COL} + N_{RC}) A_{cell} \quad (17)$$

where, A_{actual} is the required memory area (given by the memory size) and $A_{redundant}$ is the area overhead of the redundant columns. A_{cell} is the cell area computed using the layout given in [11].

5.4. Optimization Procedure

The optimization problem can be stated as:

Minimize $P_{MEM}=f(\boldsymbol{X})$

where $\boldsymbol{X} \equiv [L_{nax}, W_{nax}, L_{npd}, W_{npd}, L_{pup}, W_{pup}, N_{RC}]$

Subject to:

$$I_{LeakMem} \leq I_{LeakMax} \quad (18)$$

$A_{MEM} \leq$ Maximum Area (A_{MAX})

$E[T_{AC}]=\mu_{TACCESS} \leq$ Maximum access time mean ($\mu_{TAC\text{-}MAX}$)

For all the parameters: $\{X_{MIN}\} \leq \{X\} \leq \{X_{MAX}\}$

This is essentially a non-linear optimization problem with non-linear constraints. The upper bound on the mean access time is given to ensure that robustness of the memory has not been achieved by significantly sacrificing performance. Fig. 9 shows the basic steps of the optimization process. The upper bound of N_{RC} is determined using (17) as shown below:

$$N_{RCmax} = \frac{A_{MEM} - N_{ROW} N_{COL} A_{cellmin}}{A_{cellmin} N_{ROW}} = \frac{A_{MEM}}{A_{cellmin} N_{ROW}} - N_{COL} \quad (19)$$

The minimization of P_F in step 5-7 requires the estimation of the joint probabilities given in (4) which are computationally expensive. However, it should be noted that:

$$P_F = P[A_F + R_F + W_F + H_F] \leq P_{AF} + P_{RF} + P_{WF} + P_{HF} = P_{FMOD} \quad (20)$$

Hence, instead of minimizing P_F we try to minimize P_{FMOD}. The above problem can be solved using *Lagrange Multiplier* based algorithm [12]. The Lagrangian formulation of the above problem is shown below:

$$\text{Minimize } \boldsymbol{f(X)}=P_{FMOD}$$
$$\text{where } \boldsymbol{X} \equiv [L_{nax}, W_{nax}, L_{npd}, W_{npd}, L_{pup}, W_{pup}]$$
$$\text{Subject to: } h_1(X) = (A_{cell} / A_{cellmax}(i))\text{-}1 \leq 0$$
$$h_2(X) = (I_{LeakMem} / I_{LeakMax})\text{-}1 \leq 0$$
$$h_3(X) = (\mu_{TAC} / \mu_{TAC\text{-}MAX\,r})\text{-}1 \leq 0$$

In this work, we have considered the discrete variable space (to account for the minimum limit on the lithographic controllability of L and W) for L_{nax}, W_{nax}, L_{npd}, W_{npd}, L_{pup} and W_{pup}. To solve the discrete space Lagrangian problem, we have used the Discrete Lagrangian Method (DLM) described in [12,13].

6. OPTIMIZATION RESULTS

The optimization methodology described earlier is used to optimize the cell structure (from [14]) and the use of redundancy to minimize the memory failure probability (Table-I). To improve the beta ratio between the pull-up PMOS and access transistor, the original cell was designed with a longer PMOS. However, a weaker PMOS tends to increase the read failure. Hence, the optimization reduces the length of the PMOS and uses the extra area in the pull-down NMOS, thereby reducing the access failure. The proposed optimization algorithm allows to trade-off between the redundancy area and the active cell area. Reducing the number of redundant column allows more area for each of the actual cells. This reduces the failure probability of the cells, thereby reducing P_{MEM}. On the other hand, from (15) it can be observed that reducing N_{RC} will tend to increase P_{MEM}. Fig. 10 shows the variation of P_{MEM} with the variation of N_{RC} considering constant A_{MEM}. It can be observed that increasing the redundancy beyond a certain point increases the memory failure probability. It should be further noted that with the application of a higher value of the σ_{Vt0}, the optimized value of the redundancy (that minimizes failure probability) reduces. This indicates that with larger amount of variations, design of robust cell is more effective in reducing the failure probability (improving yield) as compared to increasing number of redundant columns. Hence, it can be concluded that a statistical analysis of effectiveness of the redundancy is necessary to improve the memory failure probability.

1. Determine the bounds on N_{RC} and initialize:
   $i=0$, $N_{RC}(0)=N_{RCmin}$, $P_{MEMmin}= 1$
2. ***repeat***
3. Determine maximum allowable cell area:
4. $A_{cellmax}(i) \leq A_{MAX} / [N_{ROW} (N_{RC}(i) + N_{COL})]$
5. Solve the following optimization problem
6. *Minimize* P_F (L_{nax}, W_{nax}, L_{npd}, W_{npd}, L_{pup}, W_{pup})
7. Subject to: $A_{cell} \leq A_{cellmax}(i)$; $P_{LeakPower} \leq P_{LeakMin}$; $\mu_{TAC} \leq \mu_{TAC\text{-}MAX}$.
8. Calculate $P_{COL}(i)$ and $P_{MEM}(i)$ using the optimum P_F and $N_{RC}(i)$.
9. ***If*** $P_{MEM}(i) < P_{MEMMIN}$ ***Then*** $P_{MEMmin} := P_{MEM}(i)$
10. ***until*** $N_{RC}(i)=N_{RCmax}$

Fig. 9: Optimization procedure

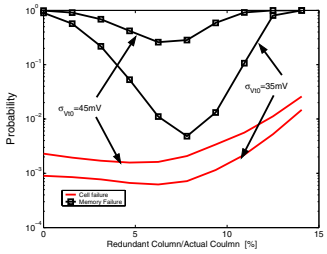


Fig. 10: impact N_{RC} on memory yield.

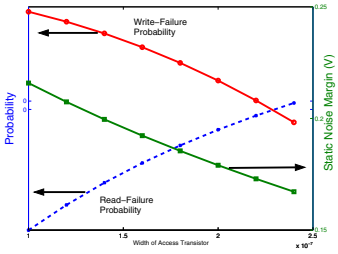


Fig. 11: Impact of SNM on failure probability of SRAM cell

Table I: optimization results

	β_{nax}/β_p	β_{npd}/β_{nax}	P_F	P_{MEM}	I_{Leak}	T_{AC}	*Yield*
Cell [11]	1.5	1.36	2.6e-3	0.034	851µA	55ps	47%
Opt. Cell	1.2	1.48	3.4e-5	0.001	950µA	46ps	95%

The static noise margin (SNM) of a cell is often used as a measure of the robustness of an SRAM cell against flipping [1]. However, an increase in SNM makes the cell difficult to write by increasing its data holding capability, which increases write failures (Fig. 11). Consequently, an increase in the SNM does not necessarily reduce the overall failure probability. Using the proposed models, it is observed that SNM does not have a strong relationship with the parametric failure of the memory.

7. CONCLUSION

In this work, we have proposed a semi-analytical method to estimate the failure probability of an SRAM cell and memory due to parameter variations. The derived models have been used to predict the yield of memory at an early stage of a design. The proposed models are used for statistical optimization of memory design, which is necessary for maximizing yield in nano-meter regimes.

REFERENCES

[1] A. Bhavnagarwala, et. al., IEEE JSSC, vol. 36, pp. 658-665, April 2001.
[2] X. Tang, et. al., *IEEE Trans. VLSI Syst.*, vol. 5, pp. 369–376, Dec. 1997.
[3] Y. Taur and T. H. Ning, *Fundamentals of Modern VLSI Devices,* New York: Cambridge Univ. Press, 1998.
[4] D. Burnett, et. al., *Symp. VLSI Tech.*, pp. 15–16, June 1994.
[5] S. Mukhopadhyay, et. al., DAC, pp. 169 -174, June 2003.
[6] http://www-mtl.mit.edu/Well/
[7] MEDICI: 2-D device simulation program, Synopsys Inc.
[8] A. Papoulis, *Probability, Random Variables and Stochastic Process*
[9]A. Chandrakasan, *Design of High-Performance Microprocessor Circuits*
[10] H. Quin, et. al., Symp. On Quality Elect. Design, pp. 55-60, Mar. 2004.
[11] R.W. Mann, IBM journal of R&D, pp. 553-566, Sep. 2003.
[12] B. W. Wah, et. al., Int. Conf. on Tools with AI, 2000
[13] Edwin K. P. Chong, An introduction to Optimization
[14] A. Agarwal, et. al., IEEE JSSC, vol. 38, pp. 319-328, Feb. 2003.