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 methodology 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 (NROW) and # of column (NCOL), # of redundant column (NRE)) 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 nano-meter 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 (Lefl=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).

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 (σVt) depends on the manufacturing process, doping profile, and the transistor sizing. In the proposed method, σVt is a minimum sized transistor (σVt0) is an input parameter and the dependence of σVt on the transistor size is given by [3]:

σVt = σVt0 √(Lmin/L) √(Wmin/W)

3. MODELING OF SRAM FAILURE
The failure 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 (Vddstand), 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(x1, ..., x6) as a function, where x1, ..., x6 are independent Gaussian random variables with mean μ1, ..., μ6 and standard deviation (STD) σ1, ..., σ6. The mean (μy) and the STD (σy) of the random variable y can be estimated as (using multi-variable Taylor-series expansion) [8]:

(1)
The above results will be used in this paper to estimate the failure pdf as a Gaussian

\[ \mu = f_0 \sum_{n=1}^{N} \frac{1}{2} \left( \frac{\partial f_0(x)}{\partial x} \right) \delta \]

\[ \sigma^2 = \sum_{n=1}^{N} \left( \frac{\partial f_0(x)}{\partial x} \right)^2 \delta \]

Assuming the Probability Distribution Function (PDF) of \( y \) to be also Gaussian (\( N(y; \mu, \sigma) \)), the probability of \( y > y_0 \) is given by:

\[ P[y > y_0] = \int_{y_0}^{\infty} f(y; \mu, \sigma) \, dy = 1 - \int_{-\infty}^{y_0} f(y; \mu, \sigma) \, dy \]

where, \( f(y; \mu, \sigma) \) is the Cumulative Distribution Function (CDF) of \( y \).

Let us assume \( y = f(x_1, x_2, \ldots, x_n) \) and \( z = g(x_1, x_2, \ldots, x_n) \) are two Gaussian random variables \( N(y; \mu_y, \sigma_y) \) and \( N(z; \mu_z, \sigma_z) \), respectively. The probability of \( y > y_0 \) and \( z > z_0 \) is given by:

\[ P[y > y_0, z > z_0] = -[P[y < y_0] + P[z < z_0] - P[y < y_0, z < z_0]] \]

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

\[ \rho = \frac{E(xy) - E(x)E(y)}{\sigma_x \sigma_y} \]

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

3.2. Read Stability Failure (Rf)

While reading the cell shown in Fig. 1 (\( V_L = 1 \) & \( V_R = 0 \)), due to the voltage divider action between \( AX_L \) and \( AX_R \), the read voltage increases to a positive value \( V_{READ} \). If \( V_{READ} \) is higher than the trip point of the inverter \( P_{NQ} \) (\( V_{TRIP} \)), then the cell flips while reading the cell (Fig. 3(a)). This represents a read failure (\( R_f \)) event. Hence, the read-failure probability (\( P_{RF} \)) is given by:

\[ P_{RF} = P[V_{READ} > V_{TRIP}] \]

\[ V_{TRIP} \text{ and } V_{READ} \text{ can be obtained by solving KCL at node } L \text{ and } R, \text{ respectively. The estimated value of } V_{TRIP} \text{ and } V_{READ} \text{ closely follows the MEDICI simulation result (Fig. 4). Assuming the PDF of } V_{READ} (=N(\mu_{READ}, \sigma_{READ})) \text{ and } V_{TRIP} (=N(\mu_{TRIP}, \sigma_{TRIP})) \text{ as Gaussian distributions with the means the variances obtained using (2) (Fig. 4). } \]

\[ P_{RF} = P[V_{READ} > V_{TRIP}] > 0 = 1 - \Phi_{2}(0) \]

where, \( \mu_{READ} = \mu_{TRIP} - \mu_{READ} \text{ and } \sigma_{READ}^2 = \sigma_{READ}^2 + \sigma_{TRIP}^2 \)

3.3. Write Stability Failure (Wf)

While writing a '0' to a cell storing '1', the node \( V_L \) gets discharged through \( BL \) to a low value \( (V_{S}) \) 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 \( P_{NQ} \) (\( V_{TRIP} \)), within the time when word-line is high (\( T_{WRI} \)), then a write failure occurs (Fig. 3(b)). The write-failure probability (\( P_{WF} \)) is given by:

\[ P_{WF} = P[W_{WRITE} > T_{WRI}] \]

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

\[ T_{WRITE} = \int_{V_{DD}}^{V_{TRIP}} \frac{I_{M(3)}(V_{L}) - I_{M(4)}(V_{L})}{I_{M(1)}(V_{L}) - I_{M(2)}(V_{L})} \, \left( V_{TRIP} - V_{READ} \right) \]

\[ \text{If } (V_{READ} \geq V_{TRIP}) \]

\[ \text{If } (V_{READ} < V_{TRIP}) \]

[9]

\[ C_{L} \text{ is the net capacitance at the node } L. \text{ } V_{DD} \text{ can be obtained by solving KCL at node } L \text{ and } R \text{. } V_{TRIP} \text{ can be obtained by solving for trip-point of the inverter } P_{NQ}. \text{ } W_{WRITE} \text{ obtained using (9) closely matches the MEDICI simulation result with } Vt \text{ variation of different transistors (Fig. 5a). Using (2), we can estimate the mean (\( \mu_{TRIP} \)) and the standard-deviation (\( \sigma_{TRIP} \)) and approximate its pdf as a Gaussian one (\( \Phi_{2}(0) \)) (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 \( T_{WRITE}(V_{READ}, \mu_{TRIP}) \), the \( P_{WF} \) can be estimated using (5).

3.4. Access Time Failure (Af)

The cell access time (\( T_{ACCESS} \)) is defined as the time required to produce a pre-specified voltage difference (\( \Delta V_{MIN}=0.1V_{DD} \)) 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:

\[ P_{AF} = P[T_{ACCESS} > T_{MAX}] \]

While reading the cell storing \( V_L = 1 \) and \( V_R = 0 \) (Fig. 1, Fig. 3), bit-line \( BL \) will discharge through \( AX_L \) and \( AX_R \) (by the current \( I_{W} \)).

\[ \text{Fig. 4: (a) Variation of } V_{TRIP} \text{ of } P_{NQ}, (b) Distributions of } V_{READ} \]

\[ \text{Fig. 5: (a) } T_{WRITE} \text{ variation with } \delta \nu_t \text{ (b) distribution of } T_{WRITE} \]

\[ \text{Fig. 6: (a) } T_{ACCESS} \text{ variation with } \delta \nu_t \text{ (b) distribution of } T_{ACCESS} \]

\[ \text{Fig. 7: (a) } V_{READ} \text{ variation with } \delta \nu_t \text{ (b) distribution of } V_{READ} \]
Simultaneously, BL will discharge by the leakage of \( AX \) of all the cells \( (L_n) \) connected to BL. Hence, the access time is given by:

\[
T_{\text{ACCESS}} = \frac{C_{\text{dd}}C_{\text{wl}}}{C_{\text{wl}} - C_{\text{bl}}} I_{\text{SL}} = \frac{C_{\text{dd}}}{I_{\text{SL}} \sum \Delta A} (11)
\]

where, \( N \) is the \# of cells attached to a bit-line (or column), \( C_{\text{dd}}/C_{\text{bl}} \) 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 \( T_{\text{ACCESS}} \) can be approximated as a Gaussian one with the mean \( \mu_{T_{\text{ACCESS}}} \) and the standard deviation \( \sigma_{T_{\text{ACCESS}}} \) obtained from (2) (Fig. 6b). Using the derived PDF \( N_{T_{\text{ACCESS}}} \), \( P_{\text{AF}} \) can be estimated using (3).

3.5. Hold Stability Failure \( (H_S) \)

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 \( (V_{DDH_{\text{MIN}}}) \), without destroying the data, is higher than the designed stand-by mode supply voltage \( (V_{DDH}) \). Thus, the probability of hold-stability failure \( (P_{\text{HF}}) \) is given by:

\[
P_{\text{HF}} = P[V_{DDH_{\text{MIN}}} > V_{DDH}] \]

(12)

Lowering the \( V_{DD} \) of the cell (say \( V_{DDH} \) represents the cell \( V_{DDH} \) 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 \( V_L < V_{THP} \) of \( P_{\text{NP}} \). Hence, \( V_{DDH_{\text{MIN}}} \) can be obtained by numerically solving:

\[
V_L(V_{DDH_{\text{MIN}}}) = V_{THP} \]

(13)

The estimated value of \( V_{DDH_{\text{MIN}}} \) using (13) closely follows the values obtained from MEDICI simulation (Fig. 7a). The distribution of \( V_{DDH_{\text{MIN}}} \) can be approximated as a Gaussian one with mean \( \mu_{V_{DDH_{\text{MIN}}}} \) and variance \( \sigma_{V_{DDH_{\text{MIN}}}} \), \( \mu_{V_{DDH_{\text{MIN}}}} \) is approximated as a Gaussian one with mean \( \mu_{V_{DDH}} \) and variance \( \sigma_{V_{DDH}} \) (Fig. 7a).

3.6. Estimation of Overall Cell Failure Probability \( (P_F) \)

The overall failure probability is given by:

\[
P_F = P_{\text{ Fail}} = P_{\text{AF}} + P_{\text{NF}} = P_{\text{AF}} + P_{\text{HF}} + P_{\text{NF}}
\]

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 \( N_{RED} \) number of redundant columns, is defined as the probability that more than \( N_{RED} \) (i.e. at least \( N_{RED}+1 \)) columns fail. Hence, \( P_{COL} \) and \( P_{MEM} \) can be given by:

\[
P_{COL} = 1 - (1 - P_{\text{NF}})^N \quad \text{and} \quad P_{MEM} = N_{MEM} \sum_{i=1}^{N_{MEM}} \binom{N}{i} (1 - P_{\text{NF}})^i P_{\text{AF}}^i
\]

(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 \( V_{\text{READ}}, V_{\text{READ}}^{\text{HOLD}}, V_{\text{THR}} \) and \( V_{\text{HOLD}} \) 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 \( (\text{small } W_m) \) reduces \( P_{\text{HF}} \) \( (V_{\text{READ}} \) decreases); however, it increases \( P_{\text{AF}} \) and \( P_{\text{NF}} \) (Fig. 8) and has very small impact on \( P_{\text{AF}} \). Reducing the strength of the PMOS pull-up transistors \( (\text{by decreasing } W_m) \) reduces \( P_{\text{AF}} \) \( (\text{reducing } I_{\text{LEAK}}) \), but increases \( P_{\text{HF}} \) \( (\text{increasing } V_{\text{READ}}) \). \( P_{\text{NF}} \) does not depend strongly on PMOS strength \( (\text{Fig. 8}) \). \( P_{\text{HF}} \) improves with an increase in \( W_m \) as the node \( L \) is more strongly coupled to the supply voltage \( (V_L \rightarrow V_{DDH}) \). Increasing \( W_m \) increases the strength of pull-down NMOSs \( (N_L \& N_R) \). This reduces \( P_{\text{HF}} \) \( (V_{\text{READ}} \rightarrow V_{TH_0}) \) and \( P_{\text{NF}} \) by increasing the strength of \( N_L \) (Fig. 8). Increase in width of \( N_L \) has little impact on \( P_{\text{HF}} \). Although, it slightly increases the nominal value of \( V_{\text{READ}} \), the reduction of \( W_m \) of \( N_L \) \( (\text{see } (1)) \) tends to reduce \( V_{\text{READ}} \) and hence \( P_{\text{HF}} \). However, a higher increase in the \( V_{\text{READ}} \) of \( P_{\text{NF}} \) initially reduces \( P_{\text{HF}} \) but with an increase in \( W_m \). However, a higher increase in \( W_m \) reduces \( V_L \) (from the applied \( V_{DDH} \)) due to an increase in the leakage of \( N_L \). Consequently, a very high \( W_m \) increases the \( P_{\text{HF}} \).

5. STATISTICAL OPTIMIZATION OF SRAM CELL

5.1. SRAM Yield Estimation Model

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

\[
\text{Yield} = \frac{1}{N_{\text{INTER}}} \sum_{i=1}^{N_{\text{INTER}}} \frac{N_{\text{MEM}}(\text{intra-die } W_{\text{INTER}}, \sigma_{\text{INTER}})}{N_{\text{MEM}}(\text{inter-die } L_{\text{INTER}}, W_{\text{INTER}}, \sigma_{\text{INTER}})}
\]

(16)

where, \( N_{\text{INTER}} \) is the total number of inter-die Monte-Carlo simulations \( (\text{i.e. total number of chips}) \). In order to maximize the yield \( P_{\text{MEM}} \) needs to be minimized. This requires optimum design of the cell configurations \( (\text{i.e. length and width of transistors}) \) and the number of redundant columns \( (N_{\text{RED}}) \). However, such an optimization of \( P_{\text{MEM}} \) has to consider the impact on the total leakage and the total area \( (A_{\text{TOTAL}}) \).

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 BJT leakage through different transistors in the cell \( (\text{Fig. 1}) \). The leakage current expressions presented in [5] are used to evaluate the different leakage components. The total memory leakage \( (I_{\text{LEAKAGE}}) \) is the summation of leakage of individual cells.

5.3. Area Estimation of SRAM

\[
\text{FIG. 8: Variation of Cell Failure Probabilities with Cell structure}
\]
The total memory area is given by:

\[ A_{\text{total}} = N_{\text{ROW}} \times N_{\text{COL}} \times A_{\text{cell}} \]

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

5.4. Optimization Procedure

The optimization problem can be stated as:

Minimize \( P_{\text{FAM}}(X) \)

where \( X = [L_{\text{cell}}, W_{\text{cell}}, L_{\text{pull-up}}, W_{\text{pull-up}}, L_{\text{pop}}, W_{\text{pop}}, N_{\text{RC}}] \)

Subject to:

\[ 1 \leq A_{\text{bot}} \leq A_{\text{top}} \]
\[ A_{\text{ME}} \leq \text{Maximum Area (A_{MAX})} \]
\[ E_{\text{MAX}} \leq \text{Maximum access time mean (E_{MAX})} \]

For all the parameters: \( \{X_{\text{min}} \leq X \leq X_{\text{MAX}} \} \)

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 \( N_{\text{RC}} \) is determined using (17) as shown below:

\[ N_{\text{RC}} = \frac{A_{\text{cell}}}{A_{\text{cell}}} - \frac{N_{\text{ROW}} \times N_{\text{COL}}}{A_{\text{cell}}} = \frac{A_{\text{ME}}}{A_{\text{ME}}} \] (19)

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

\[ P_{F} = P_{\text{cell}} + P_{\text{pull-up}} + P_{\text{pop}} + P_{\text{pull-up}} = P_{\text{FAMOD}} \] (20)

Hence, instead of minimizing \( P_{F} \) we try to minimize \( P_{\text{FAMOD}} \). The above problem can be solved using Lagrange Multiplier based algorithm [12]. The Lagrangian formulation of the above problem is shown below:

Minimize \( f(X) = P_{\text{FAMOD}} \)

where \( X = [L_{\text{cell}}, W_{\text{cell}}, L_{\text{pull-up}}, W_{\text{pull-up}}, L_{\text{pop}}, W_{\text{pop}}] \)

Subject to:

\[ h_{x}(X) = (A_{\text{cell}} / A_{\text{cell}}) - 1 \leq 0 \]
\[ h_{y}(X) = (A_{\text{ME}} / A_{\text{ME}}) - 1 \leq 0 \]
\[ h_{z}(X) = (E_{\text{MAX}} / E_{\text{MAX}}) - 1 \leq 0 \]

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 \( L_{\text{cell}}, W_{\text{cell}}, L_{\text{pull-up}}, W_{\text{pull-up}}, L_{\text{pop}} \) and \( W_{\text{pop}} \). 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 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 \( P_{\text{FAMOD}} \). On the other hand, from (15) it can be observed that reducing \( N_{\text{RC}} \) will tend to increase \( P_{\text{FAMOD}} \). Fig. 10 shows the variation of \( P_{\text{FAMOD}} \) with the variation of \( N_{\text{RC}} \) considering constant \( A_{\text{ME}} \). 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 \( \text{ov}_{\text{cell}} \), 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.

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