Abstract—Low power consumption, stability, and PVT-tolerance in Static Random Access Memories (SRAM) is essential for nanoscale System-on-Chip (SoC) designs. In this paper, a novel design flow is presented for optimizing a figure of merit called Power to Static-Noise-Margin (SNM) Ratio (PSR). The minimization of PSR results in power minimization and SNM maximization of nano-CMOS SRAM circuits which are mutually conflicting objectives. A 45 nm single ended 7-Transistor SRAM is used as an example circuit for demonstrating the effectiveness of the optimal design flow presented in this paper. Worst case temperature analysis is performed on a baseline SRAM circuit for all three Figures of Merit (FoMs): power, SNM, and PSR. After accurate characterization of the FoMs for worst case temperature and process variation analysis, the baseline SRAM circuit at worst case temperature is subjected to a polynomial regression based optimization algorithm. Simulation results demonstrate that the optimal SRAM design is PVT-tolerant with optimized power consumption, SNM and PSR.

Keywords—PVT, Nano-CMOS, SRAM, Power, SNM

I. INTRODUCTION

Power consumption is an important factor to be considered in SRAM design when targeted for embedded systems. Different design methods have been proposed, such as dual-threshold voltage (V\text{TH}) assignment, and decrease in supply voltage, which reduces the dynamic power quadratically and reduces the leakage power linearly [1]. However, substantial problems have been noted when the traditional six-transistor SRAM cell is subjected to ultra-low voltage supply as it demonstrates poor stability [2]. As process technology scales deeper, the demand for increased integration density and improved device performance is also increasing [3]. This has resulted in very fast but high power dissipation computation modules such as arithmetic logic units. These units increase the chip temperature by developing local hot spots. Increased leakage also causes self-heating due to sub-threshold leakage current [4]. For the nanoscale circuits, process variation is the most important design challenge for maintaining circuit yield.

The Read Static Noise Margin (RSNM) is defined as the minimum DC noise voltage which is required to flip the state of the SRAM cell during the read operation. It is measured as the length of the side of the largest square that is fitted inside the lobes of the butterfly curve of the SRAM. In this paper, RSNM is treated as a measure of performance. As SRAMs are scaled, sufficient SNM becomes difficult to maintain mainly due to increased variability. In particular, mismatch between the voltage transfer characteristics (VTCs) of the two halves of the cell increases due to dopant fluctuations. In many embedded systems, millions of minimum-size SRAM cells are tightly packed making SRAMs the densest circuitry on chip [3]. These areas are sensitive to manufacturing defects and process variations.

Based on the above discussion, it can be concluded that power consumption and SNM for nano-CMOS SRAMs are major design constraints in the presence of process variation and thermal effects. This paper focuses on this problem. The novel contributions of this paper are as follows:

1) A novel design flow for optimization of three figures of merit (FoMs) (power, SNM and PSR) in nanoscale CMOS SRAM circuits is proposed. To test the effectiveness of this flow, a 7-transistor SRAM cell is designed using 45 nm CMOS technology and is subjected to the proposed methodology.
2) Process variation analysis is performed exhaustively for thermal, process and device geometry effects; this paper concentrates on geometry effects due to their high impact on performance but the proposed flow can be easily adjusted for inclusion of the other effects.
3) For the optimization of the SRAM, a polynomial regression based algorithm is proposed which achieves the conflicting targets of power reduction along with SNM improvement, through optimizing the PSR.

The current literature is rich in several efforts targeting efficient SRAM design using different technology ranges and circuit topologies. Table I summarizes those research efforts. However, it is a non-trivial task to simultaneously reduce power dissipation while maintaining SNM of the circuit at different temperatures.

The paper is organized as follows: Section II presents the proposed design flow for PVT-optimal SRAM design. The baseline 7-Transistor (7T) SRAM circuit design, temperature effect analysis, and PVT analysis using geometric device parameters are discussed in section III. Section IV highlights the optimization details of polynomial regression for SRAM design. Section V has the conclusions and future research.
II. PROPOSED METHODOLOGY FOR PVT-OPTIMAL SRAM

The proposed design flow that accounts for process and voltage variations, and temperature to obtain PVT-tolerant SRAM is shown in figure 1.

![Flowchart](image)

Initially, the logical design is performed for the 7T cell. The 7T single-ended SRAM (SE-SRAM) is well known for its potential of low active power and leakage dissipations [2]. From this design, baseline parameters are measured such as average power consumption (including subthreshold leakage and gate oxide leakage) and SNM. The circuit design is performed using a specific nanoscale CMOS technology, e.g., 45 nm considered in this paper. Standard sizes are taken for transistors in the baseline design where the length of PMOS and NMOS transistors \(L_p\) and \(L_n\) is 45 nm and the corresponding widths \(W_p\) and \(W_n\) are taken as \(8 \times L_p\) and 4 \times L_n\), respectively. The figures of merit considered are the average power consumption of the SRAM circuit, and SNM. Both FoMs are simultaneously optimized through a third FoM (which we introduce in this paper), called PSR. PSR is defined as the power (including leakage) over SNM ratio; by minimizing PSR, the average power (including leakage) dissipation is minimized and performance (SNM) is maximized.

Ambient temperature analysis is performed on the baseline SRAM circuit for different temperatures, namely 27\(^\circ\)C, 50\(^\circ\)C, 75\(^\circ\)C, 100\(^\circ\)C and 125\(^\circ\)C. The SRAM circuit is then subjected to process variation analysis for geometric, process and on-chip parameters, but the most significant parameters, that is the geometric parameter set [16] is considered here and applied for process variation analysis. The device parameters taken in the geometric set are \(W_n\) and \(W_p\) the widths of the NMOS and PMOS devices, respectively, with \(L_n = L_p\) held constant at 45 nm. However, the proposed methodology can easily accommodate the other parameter sets.

The baseline SRAM circuit observed for worst-case ambient temperature is then subjected to a polynomial-regression based optimization for all three FoMs.

III. PVT ANALYSIS OF A 7T SRAM CELL

A. Circuit Design of the 7T cell

The baseline circuit design of the 7T SRAM cell is shown in figure 2. Different topological configurations for SRAM cells have been considered in order to address the read stability problem in nano-CMOS based traditional 6-transistor (6T)
cells. The 7T topology solves the problem of reduced SNM and also, is suitable for the ultra-low voltage regime. The cell operates on a single bit line instead of the traditional two bit lines as in the case of the 6T cell which performs both read and write operations. It has a read and write access transistor (transistor 1), two inverters (transistors 2, 3, 4 and 5) which are connected back to back in a closed loop fashion in order to store a bit of information, and a transmission gate (transistors 6 and 7). This storage cell has two stable states which are denoted “0” and “1”. Additional transistors serve to control the access to a storage cell during read and write operations. Thus, it typically takes seven transistors to store one bit.

![Cell Diagram](image)

**Fig. 2.** The 7T SRAM cell with sizes shown for 45 nm technology.

**B. FoM Measurements of the 7T cell**

1) **Average Power Measurement:** Power dissipation occurs in various forms, including dynamic power, subthreshold leakage, and gate leakage [17]. It is essential to study the dissipation profile of CMOS circuits and formulate the optimization problem accordingly to obtain the desired optimal circuits. The total power in the nano-CMOS circuit of the SRAM is measured as the sum of dynamic current, subthreshold leakage current and gate-oxide leakage current. The cell retains its data for a certain duration of time before it is shut down. Hence, the leakage current becomes an important issue as it affects the total power dissipation. The total power dissipation is calculated as $P_{\text{total}} = P_{\text{dynamic}} + P_{\text{sub}} + P_{\text{gate}}$, where $P_{\text{dynamic}}$ is the dynamic power consumption, $P_{\text{sub}}$ is the subthreshold leakage in transistors in the “OFF” state and $P_{\text{gate}}$ is the gate leakage flowing through the transistors. [17]:

2) **SNM Measurement:** The read SNM measurement model is described in this section. The noise margin is defined using the input voltage to output voltage transfer characteristics (VTC). It is the maximum spurious signal that is accepted by the device when used in a system while maintaining correct operation [3]. Noise is present long enough for the circuit to react, i.e. the noise may be static or DC. A static noise margin is implied if the noise is a DC source. The first model for SNM measurement is discussed in [18] and [3]. This method is used for our SNM measurement. Figure 3 shows the experimental set-up. It consists of the two inverters (I and II) in feedback and voltage sources $V_N$ acting as static noise sources which are placed in adverse direction to the input of the inverters of the SRAM circuit in order to obtain the worst case SNM [19]. In order to obtain the butterfly curve shown in figure 4, the voltages are varied to and from nodes $Q$ and $Q_b$, alternatively. Table II shows power and SNM results for the baseline design. The butterfly curve for baseline design, from which the SNM is measured, is shown in figure 4.

![BUTTERFLY CURVE](image)

**Fig. 3.** Setup for the measurement of the SNM.

3) **PSR Measurement:** A parameter called PSR is introduced in this paper. PSR is defined as the power (including leakage) over SNM ratio. Simultaneous average power dissipation minimization and SNM maximization can be achieved by minimizing the PSR parameter.

**C. Temperature Analysis of the 7T cell**

Aggressive scaling in nano-CMOS technology has resulted in increased chip density. Different portions of an SRAM circuit may experience different temperature profiles depending on their proximity to other logic units [4]. The impact of ambient temperature variation (measured at 25°C, 50°C, 75°C,
100°C and 125°C) is plotted for the 7T cell for all three FoMs, that is average power, SNM and PSR in figure 5. The increase in leakage in the circuit increases the temperature (ambient temperature and on-chip temperature) because of the strong temperature dependence on subthreshold leakage flow.

As observed from figure 5, the SNM is degraded as the temperature increases. As a result, the PSR is analyzed on the basis of power and SNM. It may be noted that the PSR is the ratio of power and SNM with units nW/mV and mV, respectively. Therefore, the two quantities are normalized for unified analysis and optimization.

D. Process Variation Analysis of the 7T cell

Process variation is one of the most important and serious problems pervasive to nano-CMOS technology. Its effects on the FoMs of the 7T cell are analyzed. Process engineers bring the process in control and minimize the sources of extrinsic variation. Thus, variation in the process will translate to variability in power and performance of the entire chip [17].

This paper highlights the geometric parameter set (sizes of the load and access transistors) which are most significant. 1000 Monte Carlo simulations are run for each FoM, taking into account the geometric device parameters (channel widths of NMOS and PMOS). Each of these process technology library and standard deviation presented in figure 9(b).

This paper highlights the geometric parameter set (sizes of the load and access transistors) which are most significant. 1000 Monte Carlo simulations are run for each FoM, taking into account the geometric device parameters (channel widths of NMOS and PMOS). Each of these process technology library and standard deviation presented in figure 9(b).

IV. PVT-TOLERANT SRAM OPTIMIZATION ALGORITHM
A. Polynomial regression based optimization algorithm

The polynomial regression based algorithm is the heart of the PVT-optimal design flow. As shown in algorithm 1, the baseline SRAM cell is taken as the input along with the baseline model file. The design is initially identified for worst-case ambient temperature. The three FoMs are measured at different temperature ranges, i.e. 25°C, 50°C, 75°C, 100°C and 125°C. The baseline 7T cell undergoes process variation at 125°C identified as the worst-case ambient temperature. On-chip analysis is being carried out in ongoing research and will be presented in a future publication. The proposed algorithm fast converges to a solution using few resources.

Algorithm 1 PVT-tolerant SRAM optimization using polynomial regression based approach.

1: Input: Baseline power and SNM of the SRAM circuit.
2: Output: Optimized FoMs (Power, SNM and PSR) with transistors identified for optimized parameter assignment.
3: Identify worst-case ambient temperature for defined FoMs (Power, SNM and PSR) of the SRAM circuit.
4: Generate power dissipation profile of SRAM design by measuring average power dissipation including leakage.
5: for Each range of $W_n$ and $W_p$ of transistors in SRAM do

6: Perform simulations of the SRAM circuit.
7: Record power, SNM and PSR.
8: end for
9: Generate surface plots for all the fOMs using polynomial regression.
10: Form normalized polynomial equations: $\hat{f}_{PWR}$ for power, $\hat{f}_{SNM}$ for SNM and $\hat{f}_{PSR}$ for PSR.
11: Minimize $\hat{f}_{PWR}$ using analytical differentiation.
12: Maximize $\hat{f}_{SNM}$ using analytical differentiation.
13: Minimize $\hat{f}_{PSR}$ using analytical differentiation.
14: Assign optimized parameters (e.g. $W_n$ and $W_p$) for the NMOS and PMOS transistors.
15: Re-simulate SRAM circuit to obtain optimized FoMs.

The objective function for PSR is defined as $\hat{f}_{PSR} = (\hat{f}_{PWR}/\hat{f}_{SNM})$, where the “hat” indicates that the quantity has been normalized by division with the maximum value in its range.

Three surface plots are generated by fitting simulation data to quadratic polynomials of the form:

$$\hat{f}_X = \sum_{i,j=0}^{2} \alpha_{ij} W_n^i W_p^j,$$

where $X$ is PWR, SNM or PSR and $\alpha_{ij}$ is the matrix of coefficients obtained during the polynomial regression. We chose polynomial regression because it is efficient, reliable and allows for very fast design exploration.

Once the analytical polynomials of the form (1) are obtained, optimal values of the vector $x = [W_n, W_p]^T$ are obtained from:

$$\frac{\partial f_X}{\partial W_n} = \frac{\partial f_X}{\partial W_p} = 0$$

$$\frac{\partial^2 f_X}{\partial W_n^2} > 0, \quad \frac{\partial^2 f_X}{\partial W_p^2} < 0,$$

where the $>$ criterion is used for minimization and the $<$ criterion is used for maximization.
B. Power optimality: $\tilde{f}_{PWR}$

The design space is spanned through parametric simulations for a range of $W_n$ and $W_p$ and the results are used in a least squares fit to a polynomial of the form (1). The coefficient matrix obtained is as follows:

$$\alpha_{ij} = \begin{bmatrix} 1.13 \times 10^{-6} & 5.02 \times 10^{-8} & -1.76 \times 10^{-9} \\ 7.81 \times 10^{-9} & 1.3 \times 10^{-10} & 1.33 \times 10^{-11} \\ -4.07 \times 10^{-9} & -8.54 \times 10^{-12} & 0 \end{bmatrix} \quad (4)$$

The surface plot corresponding to this polynomial is shown in figure 6.

To minimize the power consumption, $\tilde{f}_{PWR}$ is analytically minimized. The power optimal results are shown in Table II.

C. SNM optimality: $\tilde{f}_{SNM}$

Similarly, the design space is spanned through parametric simulations for a range of $W_n$ and $W_p$ and the data fit to a polynomial of the form (1). The coefficient matrix obtained is as follows:

$$\alpha_{ij} = \begin{bmatrix} 150.9 & 0.73 & 0.06 \\ -1.67 & -0.07 & 0.07 \\ 0.21 & -0.15 & 0 \end{bmatrix} \quad (5)$$

The surface plot corresponding to this polynomial is shown in figure 7.

To maximize the SNM, $\tilde{f}_{SNM}$ is analytically maximized. The results for SNM optimality are shown in Table II.

D. PSR optimality: $\tilde{f}_{PSR}$

For PSR optimality, the method is similar to that of power optimality. The equations are formed by normalizing the values. The normalization is performed by division of each data by the maximum value in the range of the function. Normalized data enables directly accommodating different units. The design space is spanned through parametric simulations for a range of $W_n$ and $W_p$ and the data fit to a polynomial of the form (1). The coefficient matrix obtained is:

$$\alpha_{ij} = \begin{bmatrix} 0.94 & 0.05 & 0 \\ 0 & 0 & 0.01 \\ 0 & -0.01 & 0 \end{bmatrix} \quad (6)$$

The surface plot is shown in figure 8.

To minimize the PSR, $\tilde{f}_{PSR}$ is analytically minimized. The results for PSR optimality are shown in Table II.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Baseline Design Power Optimality</th>
<th>SNM Optimality</th>
<th>PSR Optimality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Power</td>
<td>1.03μW</td>
<td>1.03μW</td>
<td>1.25μW</td>
</tr>
<tr>
<td>SNM</td>
<td>150.1mV</td>
<td>150.1mV</td>
<td>154mV</td>
</tr>
<tr>
<td>PSR</td>
<td>18.94</td>
<td>18.94</td>
<td>20.84</td>
</tr>
</tbody>
</table>

E. Analysis and comparison of the optimal design

The analysis and comparison of the baseline and optimal designs for the 7T cell at worst-case ambient temperature are now considered. Process variation is again conducted on the optimal SRAM design using device parameters $W_n$ and $W_p$ for 1000 Monte Carlo runs and the results for the three FoMs are shown in figures 9, 10 and 11. Thus, it is observed that optimal SRAM design is PVT-Tolerant.

A fair comparison of this research with related research presented in table I is not possible as the various designs differ in terms of the technology node, topology, and optimization objective. The current work does not need complex dual-$V_{TH}$ technology which is expensive for fabrication [1], [12], [15].
The current paper uses polynomial regression optimization compared to DOE-ILP in [1], [15], which is scalable to any size of SRAM circuit and will be able to optimize SRAM arrays much faster than the ILP which is of exponential complexity. Furthermore, the DOE-ILP based methodology proposed in [15] for dual-$V_T$ assignment, requires seven runs while in this work we optimize the same FOMs with only one iteration. The proposed methodology while demonstrated for $W_n$ and $W_p$ design parameters can be easily customized for other design parameters including $L_n$, $L_p$, $V_{Th}$, $T_{ox}$, as per the designers choice, without affecting the complexity and simulation/optimization time.

V. CONCLUSIONS AND FUTURE RESEARCH

A novel polynomial regression based methodology has been proposed for optimizing three FoMs (power, SNM and PSR) of nano-CMOS SRAM circuits. This methodology converges fast to an optimal solution. The optimization has been performed at cell level. A single-ended 7T cell at 45 nm has been subjected to the proposed approach which results in optimized power while increasing performance through a single metric, PSR. It is observed that the optimal SRAM is PVT tolerant. The advantage of using polynomial regression over other methods is that it is much faster, accurate and reliable. This research work, considers $W_n$ and $W_p$ device parameters, however future research will involve $L_n$, $L_p$, $T_{ox}$, and $V_{Th}$. As part of extension of this research, on-chip temperature analysis is under consideration. Also, array-level optimization of SRAM with mismatch and process variation will be considered as part of the design flow.

REFERENCES


