P3 (Power-Performance-Process) Optimization of Nano-CMOS SRAM using Statistical DOE-ILP

Garima Thakral[1], Saraju P. Mohanty[2], Dhruva Ghai[3], Dhiraj K. Pradhan[4]
Computer Science and Engineering, University of North Texas, USA.[1,2,3]
Department of Computer Science, University of Bristol, UK.[4]
Email-ID: saraju.mohanty@unt.edu[2], pradhan@compsci.bristol.ac.uk[4].

Acknowledgment: This research is supported in part by NSF awards CCF-0702361 and CNS-0854182.

Summary and Conclusions

- A novel design flow is presented for simultaneous P3 (power minimization, performance maximization and process variation tolerance) optimization of nano-CMOS circuits.
- For demonstration of the effectiveness of the flow, a 45 nm single-ended 7-transistor SRAM is used as example circuit.
- The SRAM cell is subjected to a dual-V_{Th} assignment based on a novel statistical Design of Experiments-Integer Linear Programming (DOE-ILP) approach.
- Experimental results show 44.2% power reduction (including leakage) and 43.9% increase in the read static noise margin compared to the baseline design.
- The process variation analysis of the optimized cell is carried out considering the variability in 12 parameters.

P3 (Power-Performance-Process) Optimization of Nano-CMOS SRAM using Statistical DOE-ILP

Garima Thakral[1], Saraju P. Mohanty[2], Dhruva Ghai[3], and Dhiraj K. Pradhan[4]

Department of Computer Science and Engineering, University of North Texas, USA.[1,2,3]

Department of Computer Science, University of Bristol, UK.[4]

Email-ID: saraju.mohanty@unt.edu[2], pradhan@compsci.bristol.ac.uk[4].

Abstract

- *In this research paper, a novel design flow is presented for simultaneous P3 (power minimization, performance maximization and process variation tolerance) optimization of nano-CMOS circuits.*
- *The SRAM cell is subjected to a dual-VTh assignment based on a novel statistical Design of Experiments-Integer Linear Programming (DOE-ILP) approach.*
- *A 45nm single-ended 7-transistor SRAM is used as example circuit.*
- *Experimental results show 44.2% power reduction (including leakage) and 43.9% increase in the read static noise margin compared to the baseline design.*
- *The process variation analysis of the optimized cell is carried out considering the variability effect in 12 device parameters.*

Introduction, Motivation, and Contributions

- In case of nanoscale circuit maintaining power, performance and process variation simultaneously becomes a major design challenge.
- Different design methods have been proposed like decrease in supply voltage, which reduces the dynamic power quadratic ally and reduces the leakage power linearly. However, substantial problems have been noted when the traditional six-transistor.
- SRAM cell is subjected to ultra-low voltage supply as it gives poor stability. In this paper a well-established process-level technique, called dual-*VTh* (threshold voltage) is used for reduction of power consumption.
- Read static noise margin (SNM) is defined as the minimum DC noise voltage which is required to flip the state of the SRAM cell during the read operation. The "read SNM" here is treated as a measure of performance.
- An 8 × 8 SRAM array is constructed using P3 optimized SRAM cell to study the feasibility of P3-optimal SRAM array construction.

The Proposed Methodology for P3 Optimality

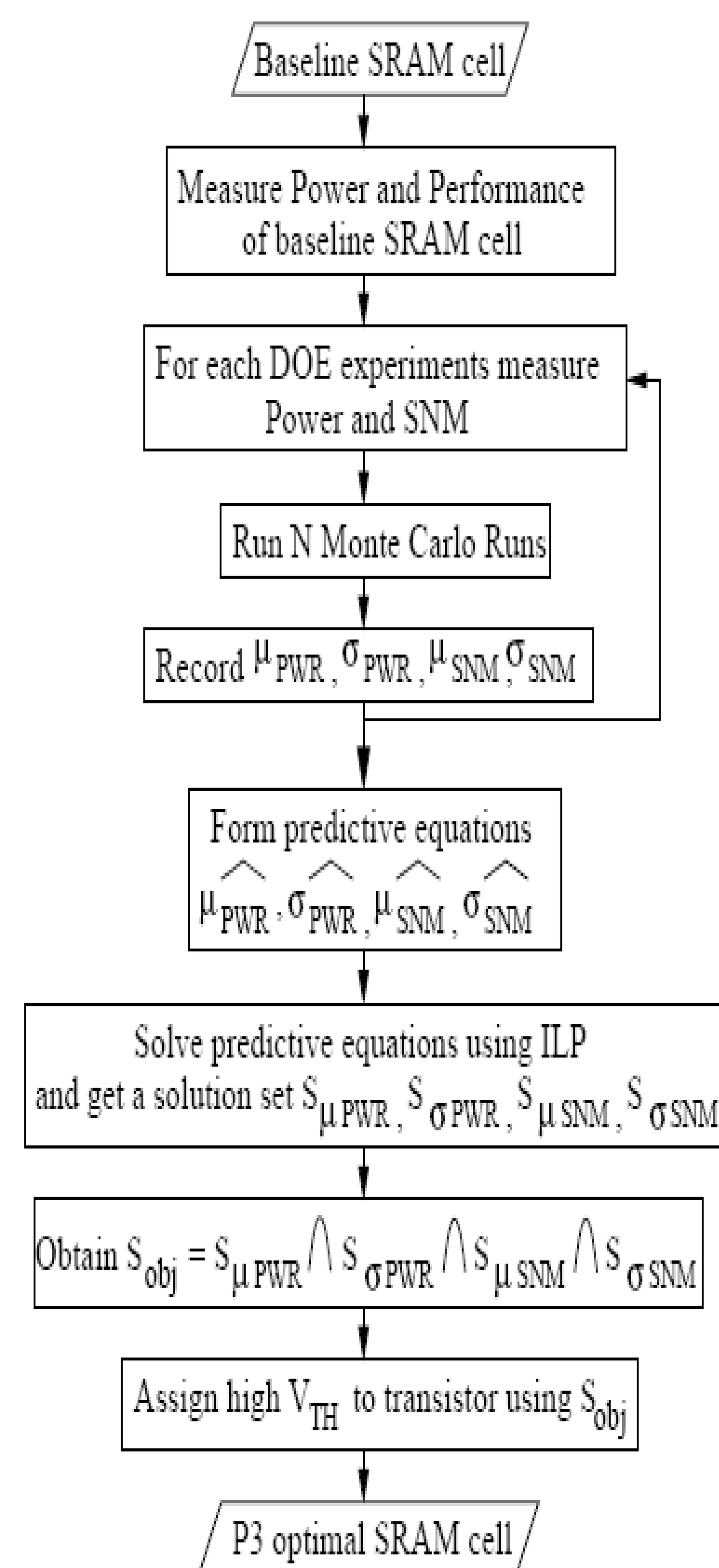

Figure 1. Proposed flow for P3-Optimal SRAM.

Figure 1 presents novel design flow for P3-optimal SRAM with reduced power dissipation, increased performance (i.e. SNM), and process-variation awareness.

Design of Seven Transistor SRAM

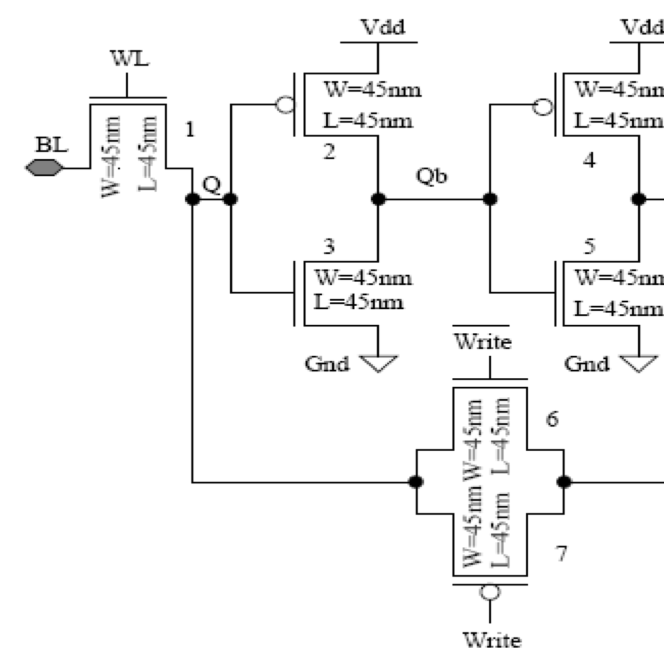

Figure 3. A 7-transistor SRAM cell [12].

The baseline 7-transistor SRAM cell is shown in Figure 3. The SRAM cell operates on a single bit line instead of the traditional two bit lines as in case of 6-transistor SRAM cell performing both read and write operations.

Statistical DOE-ILP Optimization Algorithm

Algorithm 1 P3 optimization in nano-CMOS SRAM

1: **Input:** Baseline PWR and SNM of the SRAM cell, Baseline model file, High-threshold model file.
2: **Output:** Optimized objective set $f_{obj} = [f_{PWR}, f_{SNM}]$ optimal SRAM cell with transistors identified for high V_{Th} assignment.
3: Setup experiment for transistors of SRAM cell using 2-Level Taguchi L-8 array, where the factors are the V_{Th} states of transistors of SRAM cell, the response for average power consumption is $\widehat{\mu_{PWR}}, \widehat{\sigma_{PWR}}$ and the response for read SNM is $\widehat{\mu_{SNM}}, \widehat{\sigma_{SNM}}$.
4: **for** Each 1:8 experiments of 2-Level Taguchi L-8 array **do**
5: Run 100 Monte Carlo runs
6: Record μ_{PWR}, σ_{PWR} and μ_{SNM}, σ_{SNM}
7: **end for**
8: Form linear predictive equations $\widehat{\mu_{PWR}}, \widehat{\sigma_{PWR}}$ for power $\widehat{\mu_{SNM}}, \widehat{\sigma_{SNM}}$ for SNM.
9: Solve $\widehat{\mu_{PWR}}$ using ILP: Solution set $S_{\mu PWR}$.
10: Solve $\widehat{\sigma_{PWR}}$ using ILP: Solution set $S_{\sigma PWR}$.
11: Solve $\widehat{\mu_{SNM}}$ using ILP: Solution set $S_{\mu SNM}$.
12: Solve $\widehat{\sigma_{SNM}}$ using ILP: Solution set $S_{\sigma SNM}$.
13: Form $S_{obj} = S_{\mu PWR} \cap S_{\sigma PWR} \cap S_{\mu SNM} \cap S_{\sigma SNM}$.
14: Assign high V_{Th} to transistors based on S_{obj}.
15: Re-simulate SRAM cell to obtain optimized objective set.

- Algorithm 1 discusses a well-established process-level technique, called dual-*VTh* (threshold voltage) which is used for reduction of power consumption.
- It is a very important to choose appropriate transistors for high-*VTh* assignment, thus, the statistical DOE-ILP methodology is proposed.
- Further, ILP is useful for optimizing the linear objective function subjected to constraints and obtain a bound on the optimal value to solve the predictive equations formed using DOE. Minimum sized transistors are taken for the baseline design.

Power and SNM Measurement

Figure 10 shows the comparison of baseline and P3 optimized SRAM cell power and SNM for various values of Vdd.

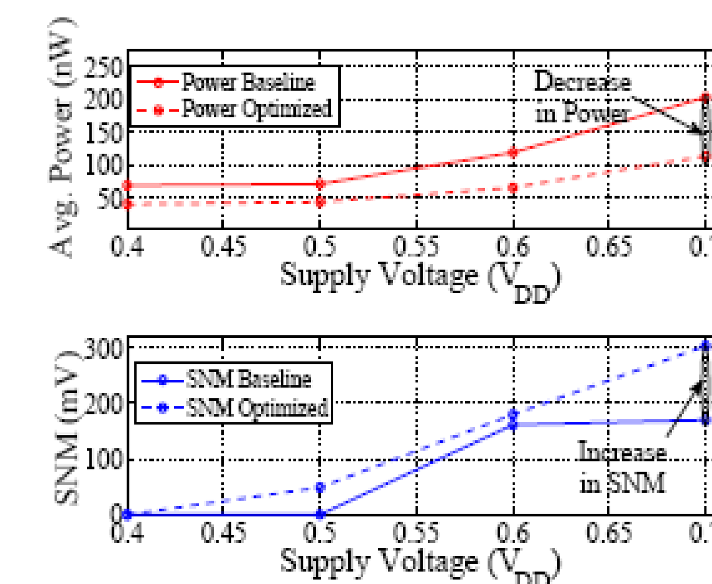

Figure 10. Power and read SNM comparison.

This table below presents that the dual- *VTh* assignment in SRAM shows 44.2% power reduction and 43.9% increase in read SNM over the baseline design.

Optimization	Parameter	Value	Change
S_{obj}	Average power P_{SRAM}	$113.6\,nW$	44.2%
S_{obj}	SNM	$303.3\,mV$	43.9%

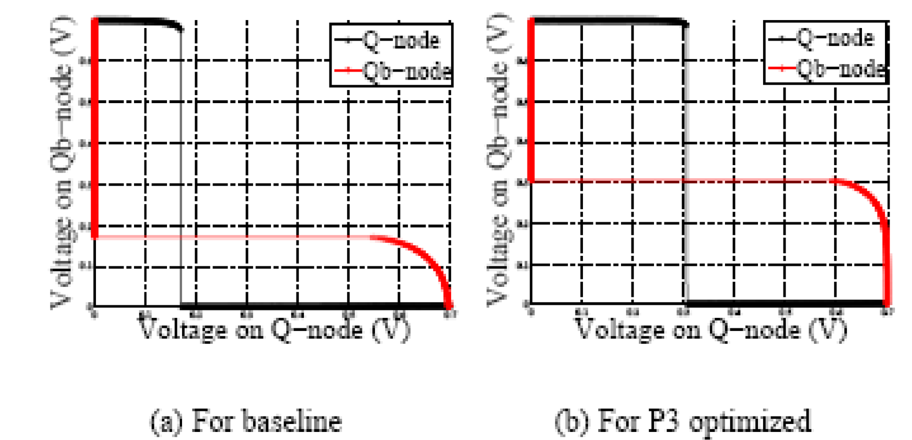

Figure 9. Butterfly curves of the SRAM.

The optimized butterfly curve is shown in Fig. 9(b)

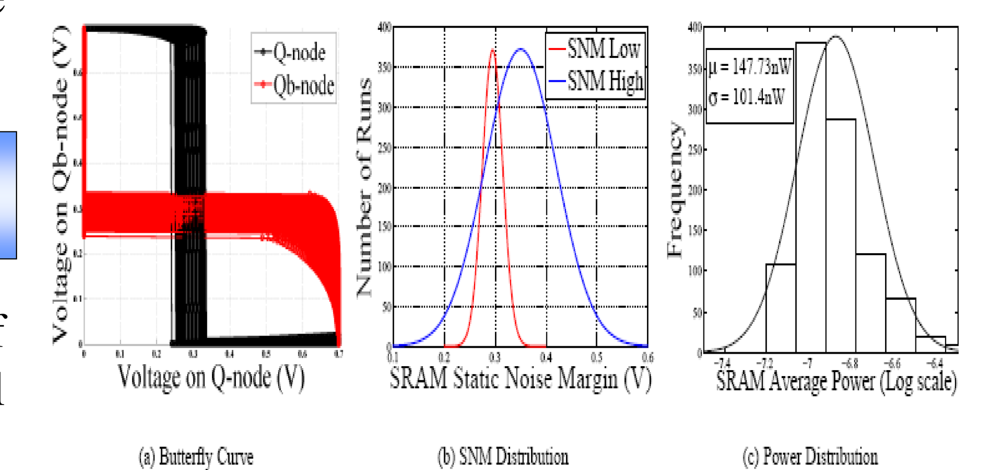

Fig. (a) shows the effect of process variations on the butterfly curve of the P3 optimized SRAM. Fig. (b) shows the distributions for "SNM High" and "SNM Low" extracted from the Monte Carlo simulations. "SNM Low" is treated as the actual SNM. Fig. (c) shows the distribution of average power of the P3 optimized SRAM cell under process variations.

Summary, Conclusions, and Future Works

- A statistical DOE-ILP approach has been presented in this paper for simultaneous P3 optimization of SRAM cell.
- As part of extension of this research, a P4 optimal methodology is under consideration, where the 4th "P" would be parasitics. Thermal effects will also be incorporated in the future which will lead to what is envisioned as P4VT optimal; V stands for voltage and T stand for temperature.
- Also, array-level optimization of SRAM with mismatch and process variation will be considered as part of the design flow.

This research is supported in part by NSF awards CCF-0702361 and CNS-0854182

UNT UNIVERSITY OF NORTH TEXAS Discover the power of ideas

VLSI Design and CAD Laboratory (VDCL)

University of BRISTOL