

DOE-ILP Assisted Conjugate-Gradient Optimization of High- κ /Metal-Gate Nano-CMOS SRAM

Saraju P. Mohanty¹ and Elias Kougiianos²

NanoSystem Design Laboratory (NSDL, <http://nsdl.cse.unt.edu>)

Computer Science and Engineering¹ and Engineering Technology²

University of North Texas, Denton, TX, USA.^{1,2}

Email-ID: saraju.mohanty@unt.edu¹ and elias.kougiianos@unt.edu²

Abstract

Low power consumption and stability in Static Random Access Memories (SRAMs) is essential for embedded multimedia and communication applications. This paper presents a novel design flow for power minimization of nano-CMOS SRAMs, while maintaining their stability. A 32 nm High- κ /Metal-Gate SRAM has been used as example circuit. The baseline SRAM circuit is subjected to power minimization using a dual- V_{th} assignment based on a novel combined Design of Experiments and Integer Linear Programming (DOE-ILP) approach. However, this leads to a 15% reduction in the Static Noise Margin (SNM) of the SRAM cell, which is an indicator of the stability of the cell. The conjugate gradient optimization overcomes this SNM degradation while reducing the power consumption. The final SRAM design shows 86% reduction in power consumption (including leakage) and 8% increase in the SNM compared to the baseline design. The variability analysis of the optimized cell is performed considering the effect of 12 parameters. SRAM arrays of different sizes are constructed to demonstrate the feasibility of the proposed SRAM cell. To the best of the authors' knowledge, this is the first study which makes use of Design of Experiments, Integer Linear Programming and conjugate gradient method for simultaneous stability and power optimization in High- κ /Metal-Gate SRAM circuits.

Index Terms

High κ /Metal-Gate, Nanoscale CMOS (Nano-CMOS), Static Random Access Memory (SRAM), Power Consumption, Dual- V_{th} , Leakage Dissipation, Static Noise Margin (SNM)

I. INTRODUCTION

The demand for reliable, high performance and high functional integration density digital circuits and systems has made the scaling of CMOS devices inevitable. Such devices are susceptible to variety of leakage components and process variation due to nanoscale manufacturing [1], [2], [3]. Memory is one of the driving forces behind the fast growth of nanoscale-CMOS (nano-CMOS) technology. In processor-based Systems-on-Chip (SoCs), memories occupy an increasing part of the area and are the main sinks of power consumption. The trend in scaled down nanoscale technologies is toward an increased contribution of static power consumption, a major problem for the most common Static Random Access Memory (SRAM) application, cache memories. Low-power SRAM design is crucial since it consumes a large fraction of total power and die area in high-performance processors [4].

The stability of embedded SRAMs is one of the important aspects for designers. It has become increasingly challenging to maintain an acceptable Static Noise Margin (SNM) in embedded SRAMs while scaling the minimum feature size and supply voltage of the SoC. Furthermore, process variation has also become a concern at nanoscale technologies because their precise control is exceedingly difficult and the increased process variations are translated into a wider distribution of transistor and circuit characteristics. Any asymmetry in the SRAM cell structure due to process variation renders the affected cells less stable. Under adverse operating conditions such cells may inadvertently flip and corrupt the stored data. The SNM can serve as a figure of merit in stability evaluation of SRAM cells [5].

The *novel contributions* of this paper are as follows:

- 1) A novel design flow is proposed for power minimization and stability maximization in nanoscale SRAMs.
- 2) A novel Design of Experiments (DOE) - Integer Linear Programming (ILP) approach is proposed for SRAM circuit power minimization.
- 3) The Static Noise Margin (SNM) of the SRAM cell is maximized using a conjugate gradient based algorithm.
- 4) The effectiveness of the methodology is shown by implementing the SRAM cell with a High- κ /Metal Gate (HKMG) 32 nm CMOS technology.
- 5) Process variation analysis of the optimal SRAM is conducted considering 12 device parameters and demonstrates the robustness of the design.
- 6) SRAM arrays of different sizes are constructed using the proposed power and stability optimized SRAM cell to study their feasibility.

The rest of the paper is organized as follows: Related prior research is discussed in Section II. Design of the HKMG 10-transistor SRAM (10T-SRAM) is discussed in Section III. Section IV discusses the proposed design flow. Section V presents the combined DOE-ILP based power minimization step in the design flow. Section VI highlights the SNM maximization step using a conjugate gradient approach. Section VII studies the effect of process variation in the proposed SRAM cell, followed by conclusions in Section VIII.

II. RELATED PRIOR RESEARCH

Recent SRAM cell topologies and optimization techniques are discussed in this section. In [6], a design flow is presented for simultaneous P3 (power, performance and process) optimization of a 7T-SRAM cell. In [7], a multi-level wordline driver scheme is proposed for hold-time power reduction and stability enhancement. In [8], a dual-voltage word line technique have been proposed for low leakage SRAM. In [1], a cell is proposed that increases read/write stability under large variations. In [9], the static and dynamic enhancement of bit-cell stability is explored by using a word-line modulation technique to enhance SNM. In [10], a DOE-ILP based methodology is proposed for dual- V_{th} assignment in a 7T-SRAM cell. A 10T-SRAM cell at low voltage and fast readout operation is proposed in [11]. A Schmitt-trigger based SRAM proposed in [12] provides better read and write ability compared to the standard 6T-SRAM cell, also achieving process variation tolerance. A 9T-SRAM cell is proposed in [13], [14] for simultaneously enhancing read stability and reducing power consumption. In [2], a self-repairing SRAM is proposed to reduce parametric failures due to process variations. A methodology is proposed in [15] to analyze the stability of an SRAM cell in the presence of random fluctuations in device parameters. In [4], the authors present a method based on dual- V_{th} and dual- T_{ox} assignment for low power design of SRAMs while maintaining performance. In [16], a process variation aware SRAM-based cache architecture is proposed.

In the current paper, a new methodology for power reduction and stability maximization is proposed for low power design, with a 10T-SRAM cell as example circuit. In [17], the authors present a 10T-SRAM topology, which is tolerant to process variation induced read failure present in traditional 6T-SRAM cells; this design has been chosen for the methodology presented in this paper. A comparison of the results of this paper with existing literature demonstrate that our design is low power and high stability. This archival journal paper is based on our conference publication [18]. The journal paper includes considerable additional material, such as functional simulation, elaboration of the optimization methods, and details of the process variation and current component analysis.

III. HIGH- κ /METAL-GATE DESIGN AND SIMULATION OF THE 10T CELL

A. High- κ / metal-gate CMOS Compact Model

For the design presented in this paper, we have used the 32 nm HKMG CMOS model [3]. The use of high- κ MOS was motivated by the fact that it keeps gate leakage under control. At the same time, prior researches in SRAM have not used high- κ MOS as evident from the discussion in Section II. In the models, which are based on BSIM4/5, two methods are used to account for the dielectric:

- 1) The parameter in the model file that denotes relative permittivity (EPSROX) is changed.
- 2) The equivalent oxide thickness (EOT) for the dielectric under consideration is calculated.

Using these steps, the EOT is calculated so as to keep the ratio of relative permittivity over dielectric thickness constant as follows:

$$T_{ox}^* = \left(\frac{\kappa_{SiO_2}}{\kappa_{gate}} \right) T_{gate}, \quad (1)$$

where κ_{gate} is the relative permittivity and T_{gate} is the thickness of the gate dielectric material other than SiO_2 , while κ_{SiO_2} is the dielectric constant of SiO_2 (=3.9). For example, for a $\kappa_{gate} = 21$ to emulate a HfO_2 based dielectric, the EOT is calculated to be 0.9 nm.

B. SRAM Figure-of-Merit (FoM) Models

The total power of a nano-CMOS circuit can be defined as the summation of dynamic power and subthreshold leakage. By using high- κ metal-gate SRAM we are eliminating gate leakage. Thus, we can calculate the power dissipation by the following expression:

$$P_{total} = P_{dynamic} + P_{subthreshold}, \quad (2)$$

where $P_{dynamic}$ is the dynamic power consumed by the transistors and $P_{subthreshold}$ is the subthreshold leakage. Both dynamic power and subthreshold leakage are calculated from SPICE simulations.

The static noise margin (SNM) of a SRAM cell is expressed as [5]:

$$SNM = V_{th} - \left(\frac{1}{k+1} \right) \times \left[\left(\frac{V_{dd} - \left(\frac{2r+1}{r+1} \right) V_{th}}{1 + \left(\frac{r}{k(r+1)} \right)} \right) - \left(\frac{V_{dd} - 2V_{th}}{1 + \left(\frac{r}{q} \right) + \sqrt{\left(\frac{r}{q} \right) \left(1 + 2k + \frac{r}{q} k^2 \right)}} \right) \right], \quad (3)$$

where r is the ratio of (β_d/β_a) which is the cell ratio or in other terms the ratio of driver transistor (W/L) to the access transistor (W/L). Similarly, q is the ratio of (β_p/β_a) , i.e. the ratio of load transistor (W/L) to the access transistor (W/L), V_{th} is the threshold voltage, k is defined as $\left(\frac{r}{r+1} \right) \left[\sqrt{\frac{r+1}{r+1-V_s^2/V_r^2}} - 1 \right]$, V_s is $V_{dd} - V_{th}$ and V_r is $V_s - \left(\frac{r}{r+1} V_{th} \right)$. SNM is defined as the length of the largest square that is

fitted inside the smallest lobe of the butterfly curves [4]. Eqn. 3 is presented here in order to show the dependency of SNM on V_{th} even though initially introduced for 6T-SRAM [5].

C. Logical Design of 10-Transistor SRAM Design

The 10T-SRAM cell transistor topology is shown in Fig. 1. This topology has been shown to be process variation tolerant in [17]. The 10T-SRAM cell is composed of two inverters connected back to back in a closed loop fashion in order to store the 1-bit information, and three transmission gates (TGW, TGR, and TGH) for the read, write and hold states, respectively.

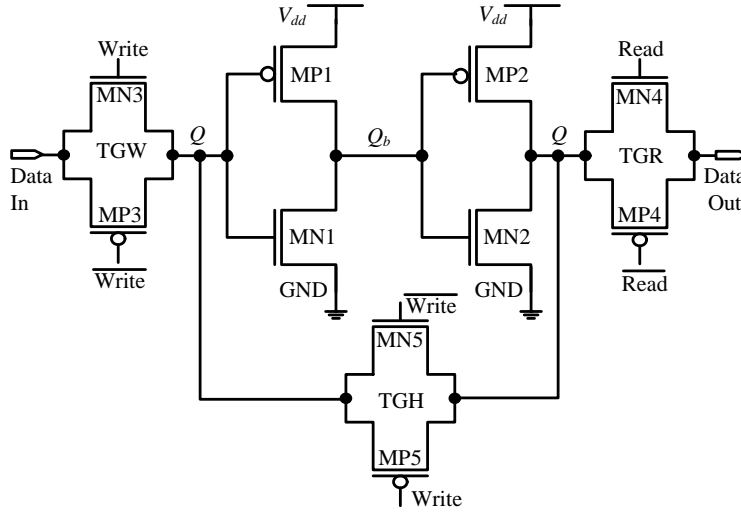


Fig. 1. The 10T-SRAM cell with transistors labeled [17].

The simulation set up for SNM measurement is shown in Fig. 2(a). Two equal voltage sources V_N with opposite polarity are applied between the two inverters of the SRAM cell. The voltage sources V_N are considered static noise sources. These voltages are swept from 0 to 0.5 V or more until the cell storage data flips. The butterfly curve for the baseline 10T-SRAM cell is shown in Fig. 2(b).

The 10T-SRAM cell initiates the read operation with the read ($Read$ and \overline{Read}) nodes. In read 1 operation, $Read$ enables the transmission gate TGR that provides a path for the Q and $Data out$ nodes. The $Read$ node goes to high level and so does Q . The transmission gate TGR as ‘ON’ carries dynamic current. In the read operation the transmission gate TGH is also in the ‘ON’ state, thus carrying dynamic current. Transistors MN1 and MP2 will have dynamic current being in the ‘ON’ state and transistors MP1 and MN2 will have subthreshold current as they are in the ‘OFF’ state, as shown in Fig. 3(a). During the read 0 operation, shown in Fig. 3(b), transistors MP1 and MN2 carry dynamic current in ‘ON’ state,

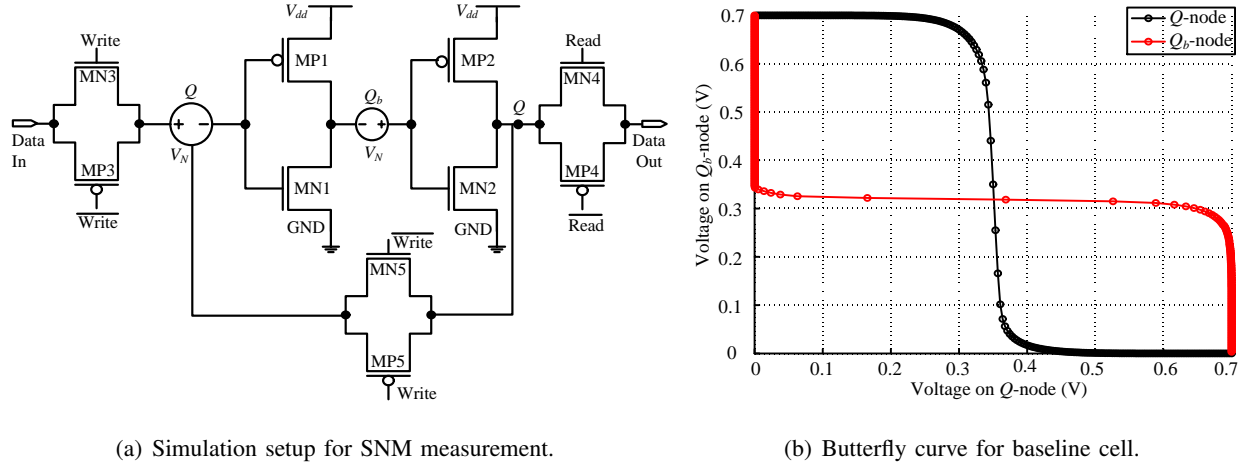


Fig. 2. SNM measurement for the 10T-SRAM.

whereas transistors MN1 and MP2 have subthreshold current in ‘OFF’ state. Subthreshold current flows through the transmission gate TGW at the write node while transmission gates TGR and TGH will have dynamic current.

In the write 1 operation, shown Fig. 3(c), the *Write* signal goes high and the transmission gate TGW connects the *Data in* node to node *Q*. When the write node goes high the transmission gate TGW forces node *Q* to the same level as the *Data* line. The transistors which are in the ‘ON’ state will have dynamic current whereas the transistors which are in the ‘OFF’ state carry subthreshold current. Thus, for write 1, transistors MP1 and MN2 are ‘OFF’ carrying subthreshold current whereas transistors MN1 and MP2 will carry dynamic current in the ‘ON’ state. Similarly in write 0 operation (Fig. 3(d)) dynamic current will flow in transistors MP1 and MN2 whereas transistors MN1 and MP2 will have subthreshold current.

The instantaneous behavior of all current components during Write and Read modes is shown in Fig. 4. The results account for dynamic current, subthreshold current and total current consumption. Further, average currents are tabulated. With the help of this table we can easily analyze the importance of each current component.

The nominal values for average power in different modes of operation such as read and write are presented in Table I. P_{dyn} is considered as the capacitive switching power in case of read/write operations while P_{sub} is subthreshold power which is present when the transistors are in the OFF state. The dynamic power due to the transition of bitlines and the charging and discharging of their capacitances [17] is also included. The power and SNM results obtained for the baseline design with supply voltage (V_{dd}) of 0.7 V are presented in Table II.

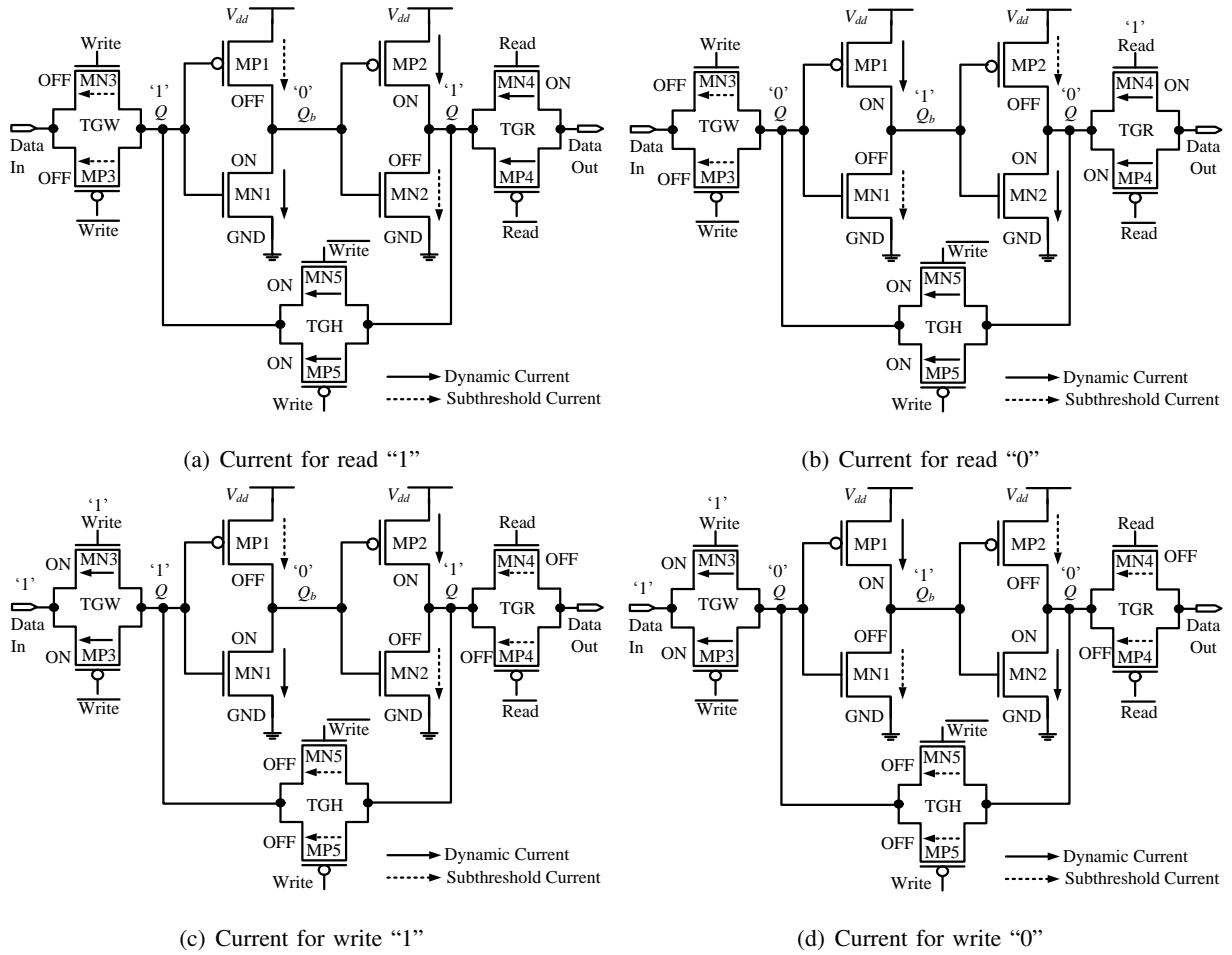
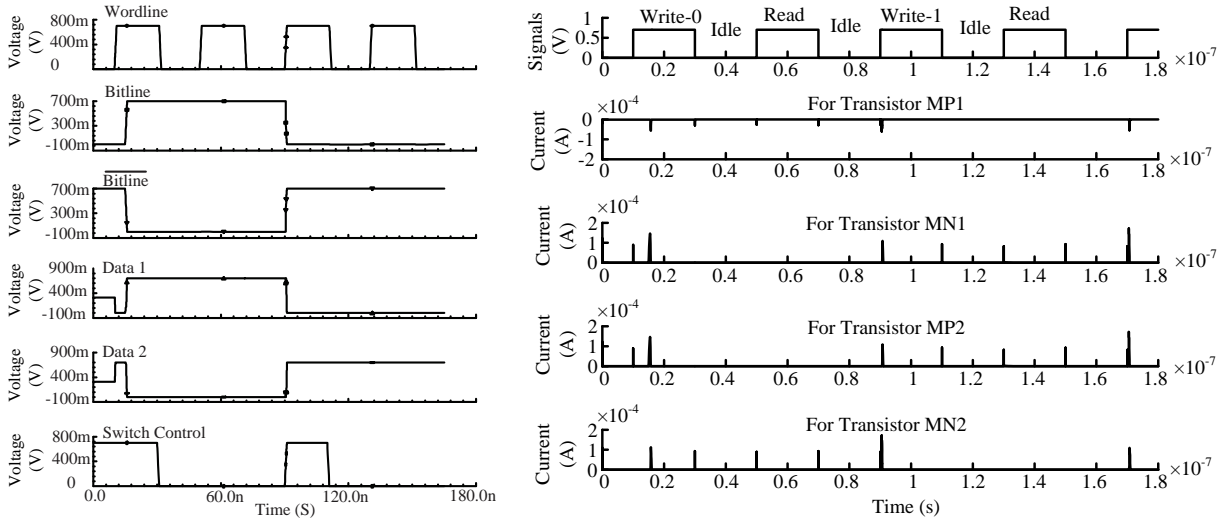


Fig. 3. Current (including leakage) paths for the 10T cell during read and write operations.

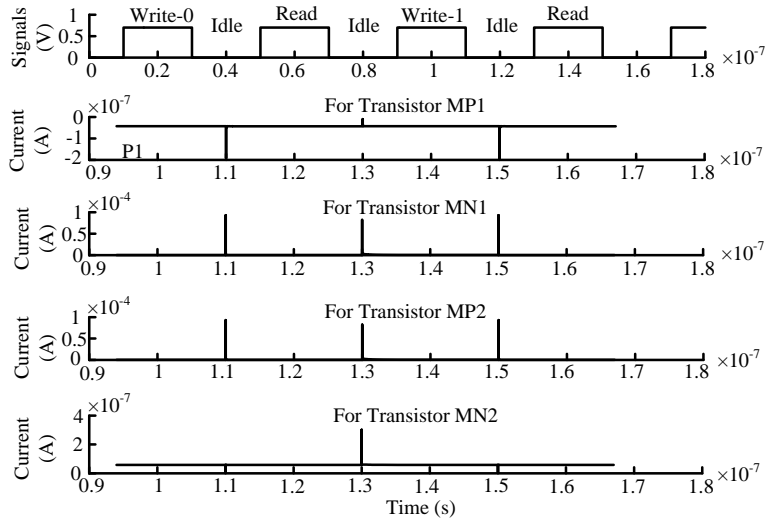
TABLE I
NOMINAL VALUES FOR AVERAGE POWER IN READ AND WRITE OPERATIONS.

Operation	I_{sub} (nA)	I_{dyn} (nA)	Access Time (ns)	Average Power (nW)
Write '1'	201.5	897.5	5.64	769.3
Write '0'	202.41	67.85	0.02	189.20
Read '1'	118.00	129.30	15.56	173.14
Read '0'	71.90	141.10	14.56	149.11



(a) Functional simulation

(b) Dynamic current



(c) Subthreshold current

Fig. 4. Simulation results of the 10T-SRAM cell showing the dynamic and subthreshold leakage currents.

TABLE II
EXPERIMENTAL RESULTS OF THE BASELINE 10T-SRAM CELL

Figure of Merit	Experimental Value
Average power P_{SRAM}	2.27 μ W
SNM	271 mV

IV. THE PROPOSED DOE-ILP BASED HKMG SRAM DESIGN FLOW

The primary purpose of the design flow is power minimization (including leakage), while maximizing stability. The proposed design flow is shown in Algorithm 1. The input to the flow is a baseline cell design that meets specifications with minimum sized transistors. The design optimization specifications are as follows: minimize power (leakage) consumption for designer defined values of static noise margin (SNM).

The figures of merit (FoMs) under consideration are measured for the baseline design. We have considered a dual threshold voltage (dual- V_{th}) assignment for power reduction of the transistors of the SRAM. This is achieved using a Design of Experiments-Integer Linear Programming (DOE-ILP) based approach. The purpose of using DOE is that it is one of the most efficient ways to understand relationship between input factors and response [19]. For determining the setting of input factors which optimize the response, we use ILP, which solves the linear equations, and ensures minimum power consuming configuration of the cell. However, this results in degradation in the stability (SNM) of the SRAM and it fails to meet the specifications. The baseline 10T-SRAM cell is then subjected to a DOE based approach using a 2-Level Taguchi L_{12} array. The factors are the 10 V_{th} states of the 10 transistors of the cell (Fig. 1). Each factor can take a high V_{th} state (+1) or a nominal V_{th} state (-1). The L_{12} array consists of a total of 12 experimental runs.

In order to improve the stability, the minimum power configuration is subjected to another sophisticated method, conjugate-gradient based optimization for SNM maximization. The conjugate gradient method is a suitable methodology for the target objectives compared to other methodologies [20].

V. DOE-ILP APPROACH FOR MINIMUM POWER AND LEAKAGE CONFIGURATION

The baseline 10T-SRAM cell is subjected to a DOE [21], [22] approach using a 2-Level Taguchi L_{12} array. The factors are the V_{th} states of the 10 transistors (Fig. 1), and the response is the average power consumption ($f_{P_{SRAM}}$). Each factor can take a high V_{th} state (+1) or a nominal V_{th} state (-1). The overall optimization steps that use DOE and ILP (DOE-ILP) are shown in Algorithm 2.

From the DOE the following predictive equation are obtained:

$$\begin{aligned} \widehat{f_{P_{SRAM}}}(nW) = & 2192.4 + 223.9 \times x_1 + 243.7 \times x_2 \\ & + 902.8 \times x_3 - 1352.5 \times x_4 + 211.9 \times x_5 \\ & - 29.2 \times x_6 - 179.1 \times x_7 + 92.6 \times x_8 \\ & - 128.2 \times x_9 - 170.72 \times x_{10}, \end{aligned} \quad (4)$$

Algorithm 1 The design flow for power and stability optimal HKMG based 10T-SRAM cell.

- 1: Design the baseline 10T-SRAM HKMG cell.
 - 2: Measure the power and read SNM of the 10T-SRAM cell.
 - 3: Use DOE-ILP approach to identify dual- V_{th} minimum power configuration.
 - 4: Assign high V_{th} to the transistors to get minimum power configuration.
 - 5: Parameterize minimum power configuration for parameter set D , where $D = (W, L$ of load, driver, and access transistors).
 - 6: Use conjugate gradient method to optimize SNM of the cell.
 - 7: Obtain optimal 10T-SRAM cell with minimum power and improved SNM.
 - 8: Measure optimal power and SNM.
 - 9: Run process variation analysis for the optimal cell considering 12 device parameters.
 - 10: Construct $M \times N$ arrays from the resulting optimal 10T-SRAM cells.
-

Algorithm 2 DOE assisted ILP to obtain minimum power dual- V_{th} HKMG 10T-SRAM cell.

- 1: Measure the power dissipation and read SNM of the 10T-SRAM cell.
 - 2: Setup experiment for transistors of cell using 2-Level Taguchi L_{12} array, where factors are the V_{th} states of the transistors and the response is average power consumption $\widehat{f_{P_{SRAM}}}$.
 - 3: **for** Each 1:12 experiments of 2-Level Taguchi L_{12} array **do do**
 - 4: Perform simulations and record the average response $\overline{f_{P_{SRAM}}}$.
 - 5: **end for**
 - 6: Form predictive equation $\widehat{f_{P_{SRAM}}}$.
 - 7: Solve $\widehat{f_{P_{SRAM}}}$ using integer linear programming (ILP).
 - 8: Obtain the solution set: P_{SRAM} .
 - 9: Assign high V_{th} to the transistors of the solution set.
 - 10: Obtain the minimum power configuration SRAM circuit.
-

where, x_i represents the V_{th} state of transistor i (Fig. 1). From this, we formulate an ILP problem:

$$\begin{aligned} \min \quad & \widehat{f_{P_{SRAM}}} \\ \text{s.t.} \quad & x_i \in \{1, -1\}, \forall i \in \{1, \dots, 10\}, \end{aligned} \tag{5}$$

where the constraints ‘+1’ and ‘-1’ represent coded values for high V_{th} and nominal V_{th} states, respectively. We form the predictive equations for power (f_{PWR}) and read SNM (f_{RSNM}) based on the

performance maximization. From the DOE-ILP method we successfully achieved our objective of minimum power consumption of the cell. We subject the minimum power configuration 10T-SRAM cell to conjugate gradient based SNM maximization, where the parameter set takes on different values, till the specifications are met [23]. The parameters for optimization are: $\{W_{pl}, L_{pl}, W_{nd}, L_{nd}, W_{pa}, L_{pa}, W_{na}, L_{na}\}$ the widths and lengths of the load, driver and access transistors, respectively. The steps of this optimization phase are shown in Algorithm 3.

Algorithm 3 SNM optimization in minimum power configuration 10T-SRAM

- 1: **Input:** Minimum power configuration cell, baseline model file, high-threshold model file, objective Set $F = [SNM, P_{SRAM}]$, stopping criterion S , parameter set $D = [W_{pl}, L_{pl}, W_{nd}, L_{nd}, W_{pa}, L_{pa}, W_{na}, L_{na}]$, lower parameter constraint C_{low} , upper parameter constraint C_{up} .
 - 2: **Output:** Optimized objective set F_{opt} , optimal parameter set D_{opt} for $|S| \leq \beta$. {where $1\% \leq \beta \leq 5\%$ }
 - 3: Run initial simulation with initial guess of D .
 - 4: **while** ($C_{low} < D < C_{up}$) **do**
 - 5: Use conjugate gradient method to generate new set of parameters $D' = D \pm \delta D$.
 - 6: Compute $F = [SNM, P_{SRAM}]$.
 - 7: **if** ($|S| \leq \beta$) **then**
 - 8: **return** $D_{opt} = D'$.
 - 9: **end if**
 - 10: **end while**
 - 11: Using D_{opt} , simulate SRAM cell.
 - 12: Record F_{opt} .
-

Our objective is SNM maximization and P_{SRAM} minimization. The algorithm initially starts with a guess of D followed by iterations, improving the guess each time until it is close enough to the objective set of F_{opt} with the stopping criterion S which here is within $\pm\epsilon$, with ϵ a designer specified error margin, in percentage. We have taken $\epsilon \leq 5\%$. The steps are shown in Algorithm 3. The algorithm satisfies the stopping criterion S with the output of optimized objective set F_{opt} and the optimal values of the design variable set D_{opt} along with the upper and lower parameter constraints.

The optimization algorithm converged in 9 iterations with each iteration lasting for 4 minutes for the

given specifications. The final values of the parameter set for the SNM optimal cell are shown in Table IV. The Table also contains the initial guess of design parameters used during the optimization. For a 32 nm node using standard practices the following sizes are selected as initial lower-limit values: $L = 32$ nm, $W = 4 L$. The upper-limit value is set 10 times larger than the lower-limit W ; this is motivated by the baseline design values that roughly met the specification. However, in practice this will depend on the experience of the design engineer. The results obtained after the optimization are presented in Table V.

TABLE IV
OPTIMIZED VALUES OF THE PARAMETER SET.

D	C_{low}	C_{up}	D_{opt}
W_{pl}	128 nm	1.28 μm	1.18 μm
L_{pl}	32 nm	1.28 μm	1.28 μm
W_{nd}	128 nm	1.28 μm	1.28 μm
L_{nd}	32 nm	1.28 μm	32.28 nm
W_{pa}	128 nm	1.28 μm	1.28 μm
L_{pa}	32 nm	1.28 μm	74.8 nm
W_{na}	128 nm	1.28 μm	1.28 μm
L_{na}	32 nm	1.28 μm	32 nm

TABLE V
FINAL OPTIMIZATION RESULTS OF THE 10T-SRAM.

Parameter	Value
10T-SRAM average power P_{SRAM}	314.5 nW
10T-SRAM SNM	295 mV

It shows 86.15% power reduction over the baseline design and 8% improvement in SNM (shown in Fig. 6(a)). Finally, if the results of Section V and VI are compared, in Section V we achieved our aim of minimum average power consumption but degradation of SNM was also observed. Hence, by applying the conjugate gradient algorithm in Section V, SNM was improved keeping the power consumption the same.

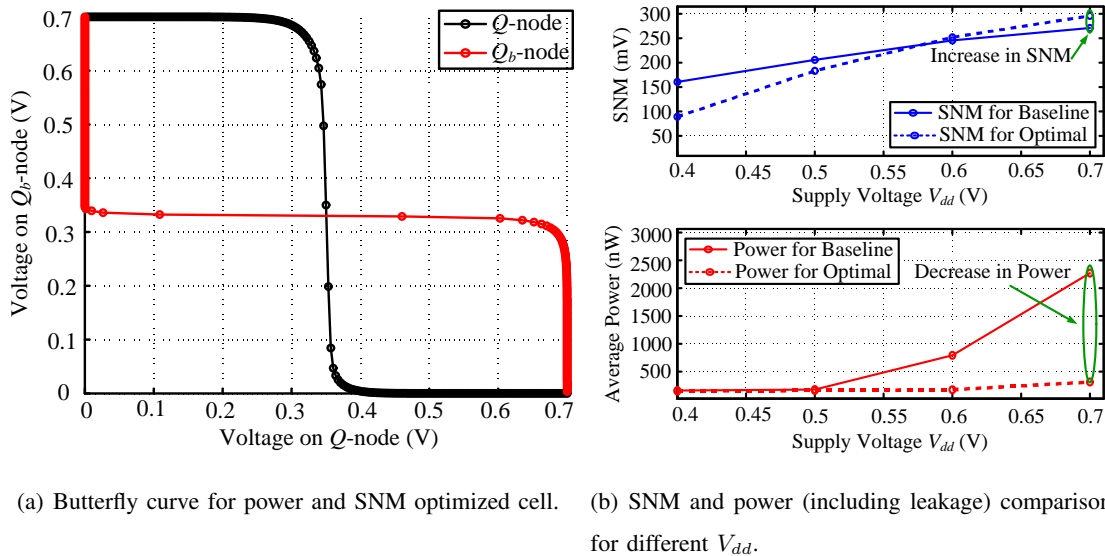


Fig. 6. Results for power and SNM optimal cell. The power dissipation has been reduced significantly with increase in the SNM.

As per the design flow, we then construct arrays of any size $M \times N$ array using the optimized cell, as shown in Fig. 7. The construction of the array verifies the feasibility of creating the functional SRAM arrays from the individual 10T-SRAM cells [24], [2]. The following sizes were designed and simulated: (1) 32×32 (1 Kb), (2) 64×64 (4 Kb), and (3) 128×128 (16 Kb). The average total power dissipation of the 10T-SRAM arrays for the above sizes were $17.5 \mu\text{W}$, $69.7 \mu\text{W}$, and $276.3 \mu\text{W}$, respectively. The SNM values were almost the same for all three test case 10T-SRAM arrays.

VII. PROCESS VARIATION ANALYSIS OF 10T-SRAM CELL

The threshold voltage variation is strongly related to device geometry (length, width, oxide thickness, etc.) and doping profile. We have exhaustively evaluated the SNM through 1000 Monte Carlo simulations to ensure there is no process variation induced failure in the SNM. The setup is depicted in Fig. 8(a). 12 process parameters, which affect the threshold voltage variation, are considered for variability: (1) T_{gaten} : NMOS gate dielectric thickness (nm), (2) T_{gatep} : PMOS gate dielectric thickness (nm), (3) L_{na} , (4) L_{pa} , (5) W_{na} , (6) W_{pa} , (7) L_{nd} , (8) W_{nd} , (9) L_{pl} , (10) W_{pl} , (11) N_{chn} : NMOS channel doping concentration (cm^{-3}), (12) N_{chnp} : PMOS channel doping concentration (cm^{-3}). It may be noted that statistical information about these parameters may not be provided by the foundry. They are identified based on various published works [25]. The aim is to make the data characterization as accurate as possible for the state-of-the-art nanoscale technology.

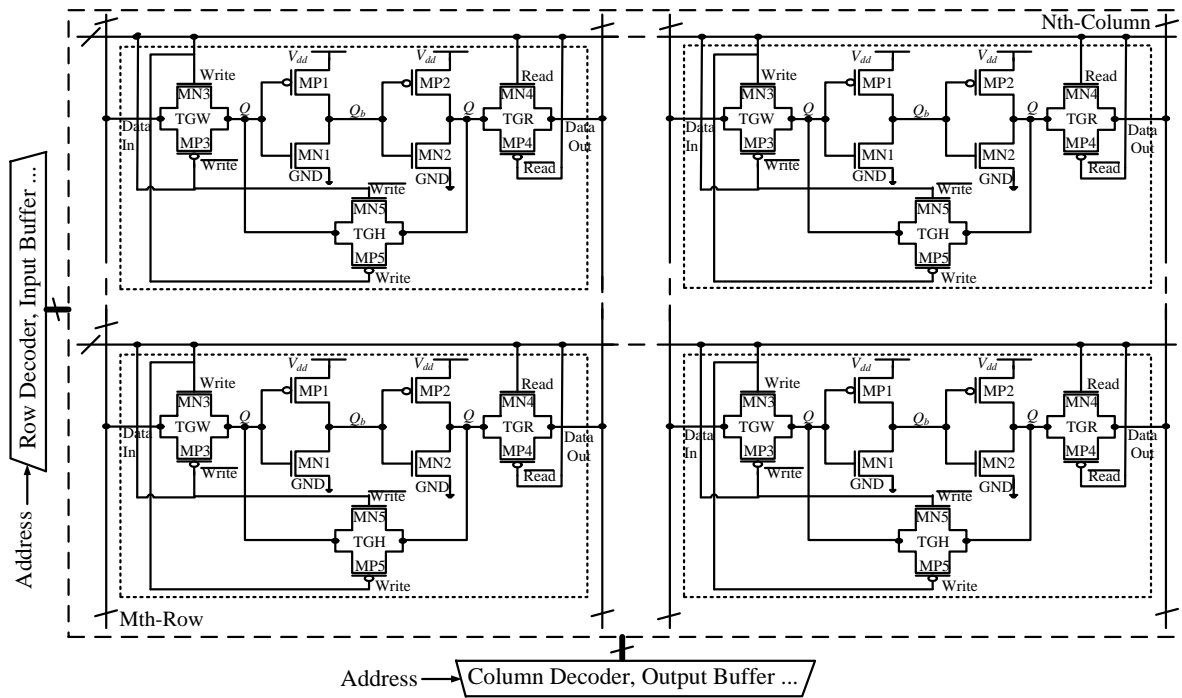


Fig. 7. Schematic representation of the $M \times N$ array constructed using the proposed 10T-SRAM cells.

Each of these process parameters is assumed to have a Gaussian distribution with mean (μ) taken as the nominal value specified in the Predictive Technology Model (PTM) [3] and standard deviation (σ) as 10% of the mean. The effect of process variation on the butterfly curve is presented in Fig. 8(b). The distributions for “SNM High” and “SNM Low” extracted from the Monte Carlo simulations are presented in Fig. 8(c), where “SNM High” is the higher SNM and “SNM Low” is the lower SNM due to asymmetry in the cell, for each Monte Carlo run. However, “SNM Low” is treated as the actual SNM. The distribution of average power of the 10T-SRAM cell is shown in Fig. 8(d). The average power distribution is observed to be lognormal in nature. The corresponding statistical data are summarized in Table VI.

VIII. CONCLUSIONS AND FUTURE RESEARCH

We have presented a methodology for cell-level optimization of SRAM power and stability. A 32 nm HKMG 10T SRAM cell subjected to the proposed methodology has shown 86% reduction in power (including leakage) and 8% increase in stability (SNM). A novel DOE-ILP approach has been used for power minimization, and the conjugate gradient method is used for SNM maximization. The effect of process variation of 12 parameters on the proposed cell is evaluated, and it is found to be process variation

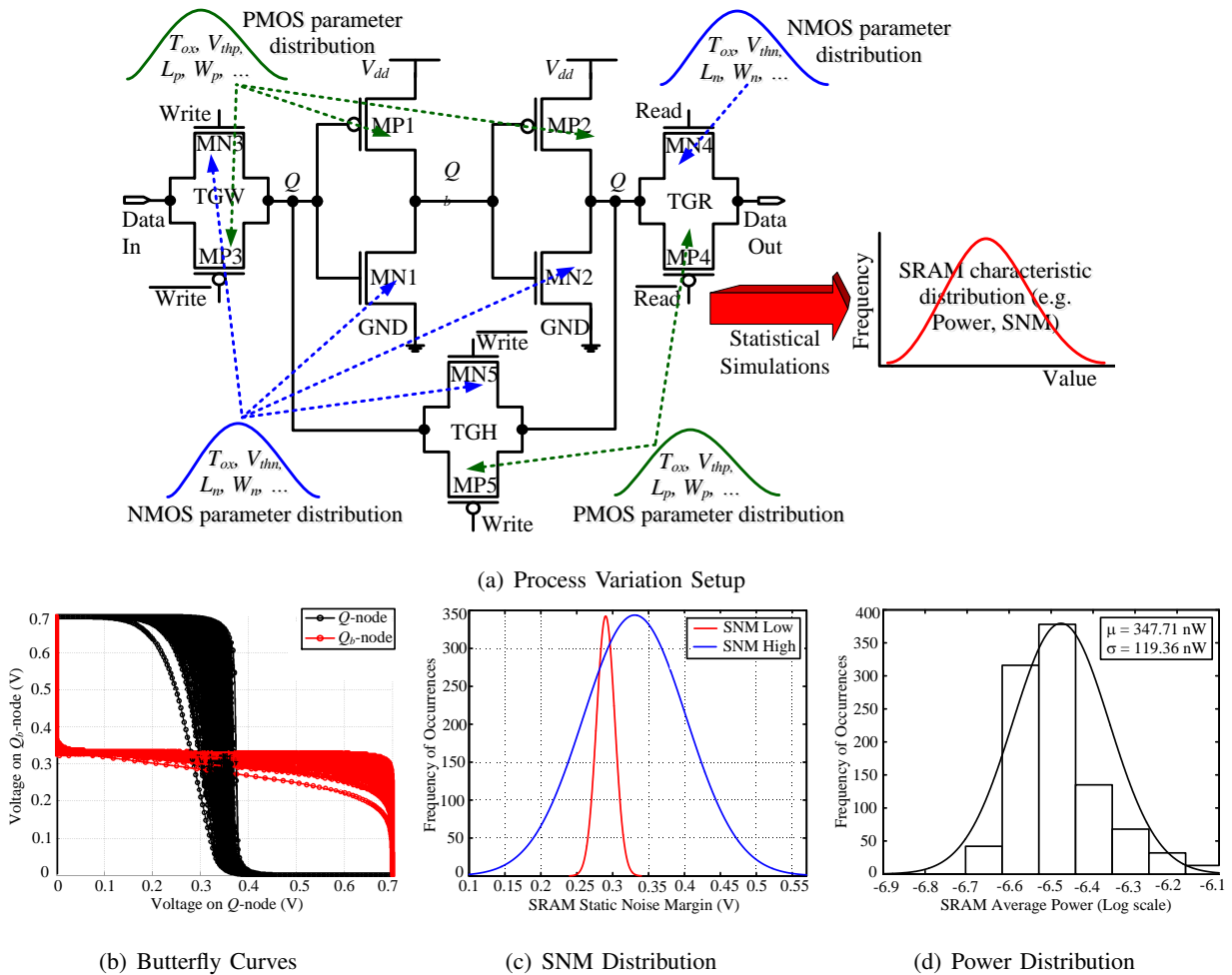


Fig. 8. Process variation analysis setup, butterfly curves, SNM statistical distribution, and power statistical distribution for the optimized 10T-SRAM cell under process variations.

TABLE VI
STATISTICAL DISTRIBUTION FOR SNM AND POWER.

10T-SRAM FoM	Mean (μ)	Standard Deviation (σ)
SNM High	330.7 mV	71.9 mV
SNM Low	290.3 mV	12.7 mV
Power	347.71 nW	119.36 nW

tolerant. Arrays have been constructed using the optimized cell and data for average power are presented. Area is an important design consideration for SRAMs. However, for a meaningful area estimation there is a need for layout design using physical design kits. For high- κ /metal-gate technology used in this paper there is no physical design library available for academic use. So, a scope for future research is layout-level simultaneous optimization of power, performance, and area.

ACKNOWLEDGMENTS

This research is supported in part by NSF awards numbers CNS-0854182 and DUE-0942629. The authors would like to acknowledge the help of Garima Thakral, graduate of the University of North Texas. This archival journal paper is based on our conference publication [18].

REFERENCES

- [1] T. Azam, B.Cheng, and D. Cumming, "Variability Resilient Low-power 7T-SRAM Design for Nano-Scaled Technologies," in *Proceedings of the 11th IEEE International Symposium on Quality Electronic Design (ISQED)*, 2010, pp. 9–14.
- [2] S. Mukhopadhyay, K. Kim, H. Mahmoodi, and K. Roy, "Design of a Process Variation Tolerant Self-Repairing SRAM for Yield Enhancement in Nanoscaled CMOS," *Solid-State Circuits, IEEE Journal of*, vol. 42, no. 6, pp. 1370–1382, june 2007.
- [3] W. Zhao and Y. Cao, "New Generation of Predictive Technology Model for sub-45nm Design Exploration," in *Proceedings of the International Symposium on Quality Electronic Design*, 2006, pp. 585–590.
- [4] B. Amelifard, F. Fallah, and M. Pedram, "Reducing the Sub-threshold and Gate-tunneling Leakage of SRAM Cells using Dual- V_t and Dual- T_{ox} Assignment," in *Proceedings of the Design Automation and Test in Europe*, 2006, pp. 1–6.
- [5] E. Seevinck, F. J. List, and J. Lohstroh, "Static Noise Margin Analysis of MOS SRAM cells," *IEEE Journal of Solid-State Circuits*, vol. 22, no. 5, pp. 748–754, October 1987.
- [6] S. P. Mohanty, J. Singh, E. Kougianos, and D. K. Pradhan, "Statistical DOE-ILP based Power-Performance-Process (P3) Optimization of nano-CMOS SRAM," *Elsevier The VLSI Integration Journal*, vol. 45, no. 1, pp. 33–45, 2012.
- [7] F. Moradi, G. Panagopoulos, G. Karakonstantis, D. Wisland, H. Mahmoodi, J. Madsen, and K. Roy, "Multi-Level Wordline Driver for Low Power SRAMs in Nano-scale CMOS Technology," in *Proceedings of the IEEE 29th International Conference on Computer Design (ICCD)*, 2011, pp. 326–331.
- [8] B. Alorda, G. Torrens, S. A. Bota, and J. Segura, "Static and Dynamic Stability Improvement Strategies for 6T CMOS Low-Power SRAM," in *Proceedings of the Design, Automation, and Test in Europe*, 2010, pp. 429–434.
- [9] F. Moradi, D. T. Wisland, H. Mahmoodi, Y. Berg, and T. V. Cao, "New SRAM design using Body Bias Technique for Ultra Low Power Applications," in *Proceedings of the International Symposium on Quality Design*, 2010, pp. 468–471.
- [10] G. Thakral, S. P. Mohanty, D. Ghai, and D. K. Pradhan, "A Combined DOE-ILP Based Power and Read Stability Optimization in Nano-CMOS SRAM," in *Proceedings of the 23rd IEEE International Conference on VLSI Design*, 2010, pp. 45–50.
- [11] S. Okumura, Y. Iguchi, S. Yoshimoto, H. Fujiwara, H. Noguchi, K. Nii, H. Kawaguchi, and M. Yoshimoto, "A 0.56-V 128kb 10T SRAM Using Column Line Assist (CLA) Scheme," in *Proceedings of the International Symposium on Quality Electronic Design*, 2009, pp. 659–663.

- [12] J. P. Kulkarni, K. Kim, S. P. Park, and K. Roy, "Process Variation Tolerant SRAM Array for Ultra Low Voltage Applications," in *Proceedings of the Design Automation Conference*, 2008, pp. 108–113.
- [13] S. Lin, Y. B. Kim, and F. Lombardi, "A Low Leakage 9T SRAM Cell for Ultra-Low Power Operation," in *Proceedings of the ACM Great Lakes symposium on VLSI*, 2008, pp. 123–126.
- [14] Z. Liu and V. Kursun, "High Read Stability and Low Leakage Cache Memory Cell," in *Proceedings of the International Symposium on Circuits and Systems*, 2007, pp. 2774–2777.
- [15] K. Agarwal and S. Nassif, "Statistical Analysis of SRAM Cell Stability," in *Proceedings of the Design Automation Conference*, 2006, pp. 57–62.
- [16] A. Agarwal, B. Paul, S. Mukhopadhyay, and K. Roy, "Process Variation in Embedded Memories: Failure Analysis and Variation Aware Architecture," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 9, pp. 1804–1814, sept. 2005.
- [17] J. Singh, J. Mathew, S. P. Mohanty, and D. K. Pradhan, "A Nano-CMOS Process Variation Induced Read Failure Tolerant SRAM Cell," in *Proceedings of the International Symposium on Circuits and Systems*, 2008, pp. 3334–3337.
- [18] G. Thakral, S. P. Mohanty, D. Ghai, and D. K. Pradhan, "A DOE-ILP Assisted Conjugate-Gradient Approach for Power and Stability Optimization in High- κ /Metal-Gate SRAM," in *Proceedings of the 20th ACM/IEEE Great Lakes Symposium on VLSI*, 2010, pp. 323–328.
- [19] S. R. Schmidt and R. G. Launsby, *Understanding Industrial Design of Experiments: 4th edition*. Air Academy Press, 1994.
- [20] M. R. Hestenes and E. Stiefel, "Methods of Conjugate Gradients for Solving Linear Systems," *Journal of Research of the National Bureau of Standards*, vol. 49, no. 6, pp. 409–436, Dec 1952.
- [21] D. Ghai, S. P. Mohanty, and E. Kougianos, "Variability-aware optimization of nano-CMOS Active Pixel Sensors using design and analysis of Monte Carlo experiments," in *Proceedings of the International Symposium on Quality Electronic Design*, 2009, pp. 172–178.
- [22] E. Kougianos and S. P. Mohanty, "Impact of Gate-Oxide Tunneling on Mixed-Signal Design and Simulation of a Nano-CMOS VCO," *Elsevier Microelectronics Journal*, vol. 40, no. 1, pp. 95–103, January 2009.
- [23] S. X. D. Tan, C. J. R. Shi, and J.-C. Lee, "Reliability-Constrained Area Optimization of VLSI Power/Ground Networks via Sequence of Linear Programmings," *IEEE Transactions on CAD of Integrated Circuits and Systems*, vol. 22, no. 12, pp. 1678–1684, 2003.
- [24] S. O. Toh, "Nanoscale SRAM Variability and Optimization," University of California at Berkeley, Tech. Rep. UCB/EECS-2011-144, December 16 2011.
- [25] T. Mizuno, J. Okamura, and A. Toriumi, "Experimental Study of Threshold Voltage Fluctuation Due to Statistical Variation of Channel Dopant Number in MOSFETs," *IEEE Transactions on Electron Devices*, vol. 41, no. 11, pp. 2216–2221, November 1994.