PVT Variations Aware Robust Transistor Sizing for Power-Delay Optimal CMOS Digital Circuit Design

by

Prateek Gupta, Shirisha Gourishetty, Harshini Mandadapu, Zia Abbas

in

IEEE International Symposium on Circuits and Systems
(*ISCAS 2019*)

Report No: IIIT/TR/2019/-1

Centre for VLSI and Embeded Systems Technology
International Institute of Information Technology
Hyderabad - 500 032, INDIA
May 2019

PVT Variations Aware Robust Transistor Sizing for Power-Delay Optimal CMOS Digital Circuit Design

Prateek Gupta, Shirisha Gourishetty, Harshini Mandadapu, Zia Abbas
Center for VLSI and Embedded Systems Technology (CVEST)
International Institute of Information Technology, Hyderabad (IIIT-H) Hyderabad, India - 500032
Email: (prateek.g, gourishetty.shirisha)@research.iiit.ac.in, h.mandadapu@students.iiit.ac.in, zia.abbas@iiit.ac.in

***Abstract*—Enormous increase in process variations (due to progressive CMOS technology scaling) along-with the temperature and supply voltage variations are severely degrading the fabrication outcome of digital circuits i.e. circuits are not accomplishing the specification bounds of the required performances. Therefore, process and operating variations aware optimization has become a very essential task in VLSI design. Moreover, many specifications in a circuit have challenging trade-offs, hence demand effective optimization skills. With this vision, this paper presents optimization algorithm based robust transistor sizing for various nanoscale CMOS digital circuits. The objective is to minimize the static i.e. leakage power without degrading the operating frequency (i.e. keeping the propagation delays in bound) and area. The reported results are shown for 32nm CMOS Metal gate High-k model parameters, however methodology is equally valid for further scaled technology nodes. The Overall reduction in leakage power obtained is up to 88% keeping bound on the critical path delay. The temperature range and supply voltage has been taken between $-55^{\circ}C$ to $+125^{\circ}C$ (for automotive applications) and 0.90V to 1.10V ($\pm 10\%$ variations) respectively at 3-sigma design.**

***Index Terms*—Particle Swarm Optimization algorithm (PSO), Artificial Bee Colony algorithm (ABC), Transistor sizing, Power, Delay, Area, HSPICE**

I. Introduction

The need for the more and more number of devices, in turn higher numbers of functions in a single Integrated Circuit (IC) and greater operating speed are the prime reason for the continuous downsizing of CMOS technology [1]. However, such progressive miniaturization of device dimensions in CMOS circuits has also imported an enormous increase in process variability [2], [3] and therefore, performances figures like propagation delays and power dissipation notably the leakage power are severely degraded [4]. The detailed discussion on leakage mechanism can be found in [5], [1], [6], [7].

As a result, the efficient functioning of an IC depend on the optimization of circuits performance such as operating frequency (i.e. propagation delays), power dissipation (leakage and dynamic), and area. The impact of statistical variations on circuit performance was insignificant in previous CMOS technologies (0.5m or higher) [1] and therefore the performance was predictable over the device life time. However, it has become very difficult to maintain the same manufacturing control over the designs, which resulted in the enormous increase in statistical variations, hence enforcing a significant threat in achieving the desired timing and/or power criteria eventually degrading the fabrication outcome of ICs [1], [8], [9]. Additionally, operating variations such as supply voltage and temperature are already exist to significantly vary the circuit performance from their intended values. As a result, all the sources of process and operating variations combine together can take the circuit operation out of the needed operation. Therefore, circuit optimization has become a essential as well as complex task in IC design. With this view, the proposed work presents the optimization algorithms (ABC and PSO) based robust transistor sizing capable of searching the circuit design space to find the optimal sizing (width and length) of all the transistors in the respective CMOS digital circuits in order to minimize the power (without degrading the delays) and area, considering operating variations as for the automotive application specifications and process variations for 32nm CMOS technology node.

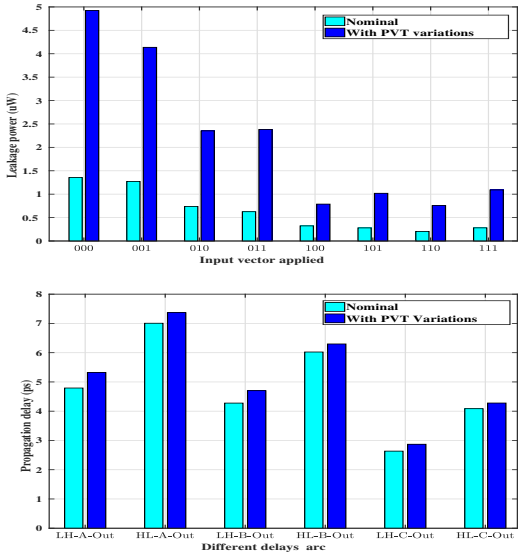


Fig. 1: Impact of PVT variations on performances of NOR3 cell

As an example, Fig. 1 depicts the impact of process, temperature and supply voltage (PVT) variations together on leakage power and the propagation delays in a 3-input NOR standard cell. It can be clearly observed from the Fig.1, the average leakage power increases up to 350% with PVT variations, while the impact is relatively small on the propagation delays.

978-1-7281-0397-6/19/$31.00 ©2019 IEEE

Many Evolutionary Algorithms (EAs) have been proposed in the past for optimization purposes which are population dependent heuristic algorithms. Most commonly used EAs are Genetic Algorithm (GA) and Particle swarm optimization (PSO) [10], [11]. PSO is generally preferred over GA because it allows greater diversity and exploration over a single population. PSO also offers faster convergence over GA. To overcome such concerns and to obtain the PVT invariant robust transistor sizing of various digital circuits, we have implemented two swarm based algorithms named; Particle swarm optimization (PSO) algorithm [11] and Artificial bee colony algorithm (ABC) [12].

The ability of PSO can be judged by two properties; Exploration and Exploitation, to ensure the performance of the algorithm [13]. PSO starts with a group of particles with randomized positions and velocity. Let $x_i(t)$ be the position of i^{th} particle in the search space at time instance t. During every iteration the position of the particle can be changed by adding velocity $v_i(t)$ to it.

$$x_i(t) = x_i(t) + v_i(t) \tag{1}$$

$$v_i(t) = wv_i(t-1) + c_1r_1(p_i(t) - x_i(t)) + c_2r_2(p_g(t) - x_i(t)) \tag{2}$$

In (2), $p_i(t), p_g(t)$ represents i^{th} particle personal best and swarm global best positions at instance t; c_1 and c_2 represents random forces in direction of $p_i(t)$ and $p_g(t)$; r_1 and r_2 are random numbers between [0,1]; w is the inertia weight parameter. High value of w leads to better exploration and low value of w leads to better exploitation.
The algorithm has to run for a certain number of iterations and the p_g position at the end of the process gives the optimum position. In this work, number of particles and epochs considered as 20 and 500 respectively.

ABC is swarm based meta heuristics algorithm based on foraging behaviour of honey bee colonies. Colony consists of three different classes of bees namely employed, onlooker and scout bees. Initialization of food sources is given by the eq. 3

$$x_{mn} = l_n + rand(0,1)\big(u_n - l_n\big) \tag{3}$$

where m=1,...F (F is the number of food sources, taken as 20), n is the number of parameters to be optimized, l_n and u_n are the lower and upper bound of the x_{mn} respectively. Employed bees arbitrarily chooses a food source and create a neighborhood of it using the eq. 4

$$u_{mn} = u_{mn} + \phi_{mn}(x_{mn} - x_{kn}) \tag{4}$$

where x_k is a randomly selected food source, n is a randomly chosen parameter index and ϕ_{mn} is a random number within the range [1,1]. Then they return to their honeycomb and dance on this area. Onlooker bees evaluate the dance of the employers with the probability of selecting high profitability of a food source. The probability (is a function of profitability of the source) is given by the eq. 5

$$p_m = fit_m(x_m) / \sum_{j=1}^{n} fit_m(x_m) \tag{5}$$

The scout bees discover a new food source after the maximum limit (limit taken as 100) of trail is reached.

The rest of the paper is organized as follows: Section II describes methodology and implementation of the algorithms through an example circuit. Section III discusses the results obtained followed by conclusion in Section IV.

II. Methodology and Circuit Implementation

To size each and every transistor in the basic digital cells, we have applied the PSO and ABC algorithm on various basic circuits and varied the width (W) and length (L) of all transistors with a step of 1nm in the range as mentioned in Table II and compared the obtained leakage values with the values at initial sizing. The reported algorithms are applied with the goal to minimize the average of leakage power for all possible input combinations in the given circuit as a fitness function while considering constraints for propagation delay as highest initial delay [14]. All the delay arcs have been calculated at 0.35fF capacitive load, which is equivalent to 4 minimum sized CMOS Inverter (FO4) load in the targeted Metal Gate High-K CMOS 32nm technology node [15], [16]. The initial sizing of digital cells are considered for equal low-to-high and high-to-low delays in the targeted technology node [8].

We have considered $\pm 10\%$ variations at 3σ in 10 process parameters for both the nMOS and pMOS devices as reported in Table I. The reported process parameters in Table I are as follows: channel-length offset parameter $(lint)$, channel-width offset parameter $(wint)$, physical gate equivalent oxide thickness $(toxp)$, electrical gate equivalent oxide thickness for both NMOS $(toxe_n)$ and PMOS $(toxe_p)$, nominal gate oxide thickness for gate dielectric tunneling current model only for both NMOS $(toxref_n)$ and PMOS $(toxref_p)$, junction depth (xj) and channel doping concentration at depletion edge for zero body bias for NMOS $(ndep_n)$ and PMOS $(ndep_p)$ and their dependence on the circuit's performances can be referred from [17], [18].

TABLE I: Process variation considered ($32\mathrm{nm}$ tech.)

Sr. No.	Process parameter	Lower deviation	Nominal	Higher deviation
1	$lint$	2.35e-09	2.7e-09	3.05e-09
2	$wint$	4.34e-09	5e-09	5.59e-09
3	$toxp$	5.74e-09	6.5e-09	7.26e-09
4	$toxe_n$	7.96e-10	9e-10	10.28e-10
5	$toxe_p$	7.97e-10	9.2e-10	10.37e-10
6	$toxref_n$	7.96e-10	9e-10	10.28e-10
7	$toxref_p$	7.97e-10	9.2e-10	10.37e-10
8	xj	1.22e-08	1.4e-08	1.57e-08
9	$ndep_n$	5.77e+18	6.5e+18	7.26e+18
10	$ndep_p$	2.45e+18	2.8e+18	3.12e+18

At $32\mathrm{nm}$ technology node, the normal condition of operation (NCO) and worst condition of operation (WCO) are defined as follows:

- At NCO: {temperature, supply voltage, process} = $\{25^oC,\ 1.0V,$ worst leakage process$\}$
- At WCO: {temperature, supply voltage, process} = $\{125^oC,\ 1.10V,$ worst leakage process$\}$ (for leakage

TABLE II: Design Parameters Bounds

Technology node	Design Parameter (nm)	Lower	Initial	Upper
32nm	Length_NMOS	32	32	35
	Length_PMOS	32	32	35
	Width_NMOS	64	[8]	900
	Width_PMOS	64	[8]	900

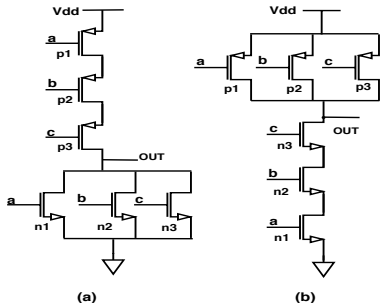


Fig. 2: Schematic of a) NOR3 cell b) NAND3 cell

computation) and {temperature, supply voltage, process} = $\{-55$ to 125^oC, $0.90V$, worst delay process$\}$ (for delays computation).

The methodology can be easily inferred through an example of CMOS 3-input NOR (NOR3) cell and 3-input NAND (NAND3) cell, each consist of 6 transistors in mirror fashion as shown in Fig. 2. The average leakage power for the all possible eight input combination i.e. from 000, .. , 111 is minimized while keeping the bounds of all low to high and high to low delay arcs as the highest delay value of the initial condition for respective NCO and WCO condition. Table III compares the leakage and delay performances of a NOR3 cell obtained using the reported algorithms with their initial performance results. It also shows the optimal W/L (width/length) ratio obtained for the respective MOSFET. Table IV reports the same for a NAND3 cell. Here, $Leak\_pqr$ represents the leakage power at the input vector *pqr* and T_{plh_i} and T_{phl_i} represent the low to high and high to low delay arc respectively for the input 'i'. The worst delay is observed to occur at $-55°C$ and $+125°C$ for NOR3 cell and NAND3 cell respectively. As shown in table III, the highest delay value for NOR3 cell at NCO and WCO are 7.36ps and 8.24ps. These delay values are taken as the bounds for all delay arcs. Compared to the results at initial sizing, the percentage reductions in average leakage power for NOR3 cell at NCO and WCO are around 65% and 68% respectively using PSO algorithm. Whereas, using ABC algorithm, the reduction is around 65% and 88% respectively. Similarly, as shown in table IV, for NAND3 cell, the leakage results obtained for NCO and WCO are 76% and 74% respectively using PSO, and around 76% and 75% restively using ABC.

III. Results & Discussions

The procedure is also applied on 45nm in addition to 32nm MGK technology on multiple CMOS standard digital cells. Due to space limitations, the author has not shown the individual performances, transistor sizing and the schematics of all the digital circuits.

TABLE III: Results obtained for CMOS NOR3 cell at 32nm technology

Performances	at Initial		with PSO		with ABC	
	NCO	WCO	NCO	WCO	NCO	WCO
$Leak\_000$ (μW)	4.92	12.0	1.48	3.81	1.49	2.09
$Leak\_001$ (μW)	4.13	19.3	1.26	5.40	1.23	2.02
$Leak\_010$ (μW)	2.35	14.5	0.87	3.80	0.88	1.44
$Leak\_011$ (μW)	2.38	6.58	0.96	2.23	0.95	1.25
$Leak\_100$ (μW)	0.78	11.8	0.32	4.64	0.33	0.77
$Leak\_101$ (μW)	1.01	3.88	0.37	1.42	0.36	0.42
$Leak\_110$ (μW)	0.75	3.07	0.28	1.09	0.29	0.30
$Leak\_111$ (μW)	1.09	2.41	0.43	0.80	0.42	0.41
T_{plh_a} (ps)	5.32	5.64	7.36	7.60	7.35	7.60
T_{phl_a} (ps)	**7.36**	**8.24**	5.68	6.47	5.5	6.28
T_{plh_b} (ps)	4.70	4.99	7.36	7.52	7.20	7.35
T_{phl_b} (ps)	6.29	6.88	5.14	5.75	5.04	5.64
T_{plh_c} (ps)	2.87	3.0	6.27	6.17	6.23	6.18
T_{phl_c} (ps)	4.27	4.61	4.40	4.91	4.22	4.75
$(W/L)_{p1}$	768/32		328/32		326/32	
$(W/L)_{p2}$	768/32		152/32		178/32	
$(W/L)_{p3}$	768/32		200/32		173/33	
$(W/L)_{n1}$	128/32		64/32		64/33	
$(W/L)_{n2}$	128/32		64/32		64/33	
$(W/L)_{n3}$	128/32		64/32		64/32	

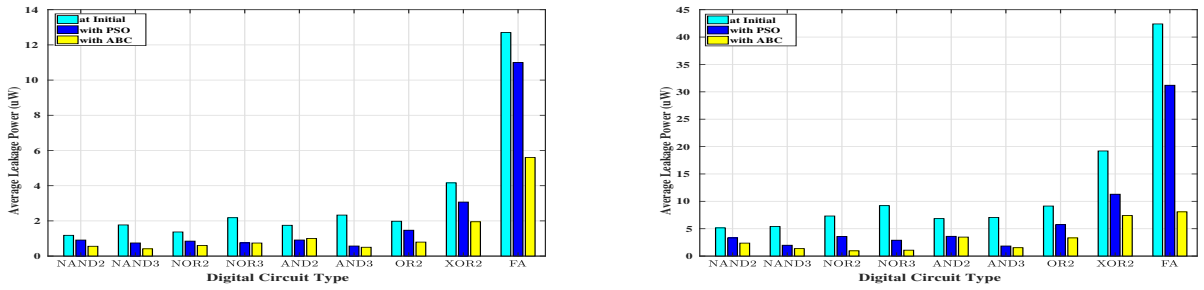
TABLE IV: Results obtained for CMOS NAND3 cell at 32nm technology

Performances	at Initial		with PSO		with ABC	
	NCO	WCO	NCO	WCO	NCO	WCO
$Leak\_000$ (μW)	1.73	3.75	0.53	1.17	0.47	1.03
$Leak\_001$ (μW)	1.13	2.52	0.28	0.63	0.33	0.74
$Leak\_010$ (μW)	1.30	2.80	0.40	0.89	0.34	0.74
$Leak\_011$ (μW)	0.68	2.13	0.15	0.38	0.20	0.57
$Leak\_100$ (μW)	2.04	3.61	0.53	1.03	0.47	0.86
$Leak\_101$ (μW)	1.51	3.12	0.29	0.54	0.35	0.69
$Leak\_110$ (μW)	2.50	4.98	0.55	0.98	0.47	0.81
$Leak\_111$ (μW)	3.26	20.2	0.56	2.97	0.67	5.46
T_{plh_a} (ps)	**6.64**	**7.31**	5.82	6.40	6.70	7.40
T_{phl_a} (ps)	3.11	3.33	6.57	7.05	6.53	7.02
T_{plh_b} (ps)	5.53	6.07	6.18	6.80	6.68	7.39
T_{phl_b} (ps)	3.11	3.29	6.09	6.48	6.70	7.13
T_{plh_c} (ps)	4.01	4.50	5.19	5.70	6.69	7.40
T_{phl_c} (ps)	2.66	2.78	6.10	6.44	6.53	6.91
$(W/L)_{p1}$	256/32		156/34		90/32	
$(W/L)_{p2}$	256/32		126/34		81/32	
$(W/L)_{p3}$	256/32		126/34		71/32	
$(W/L)_{n1}$	384/32		126/34		88/32	
$(W/L)_{n2}$	384/32		156/34		71/32	
$(W/L)_{n3}$	384/32		66/34		64/33	

The overall leakage power reduction obtained in 2 and 3 input NAND, NOR gate, 28-transistor full adder circuit, OR gate, XOR gate etc using reported swarm based algorithms are displayed in Fig. 3 and their corresponding taken delay bounds are shown in Fig. 4 a) which are the highest delay values at initial sizing for every circuit for both NCO and WCO conditions. Also, the total power (includes both static and dynamic power) for all the basic cells have been reported in the Table VI. The obtained optimal sizing for basic cells are used to optimize the various complex cells such as C17 cell, Parity Checker, C432, 4-bit multiplier, 32 bit ripple carry adder etc. and their results are reported in Table V.

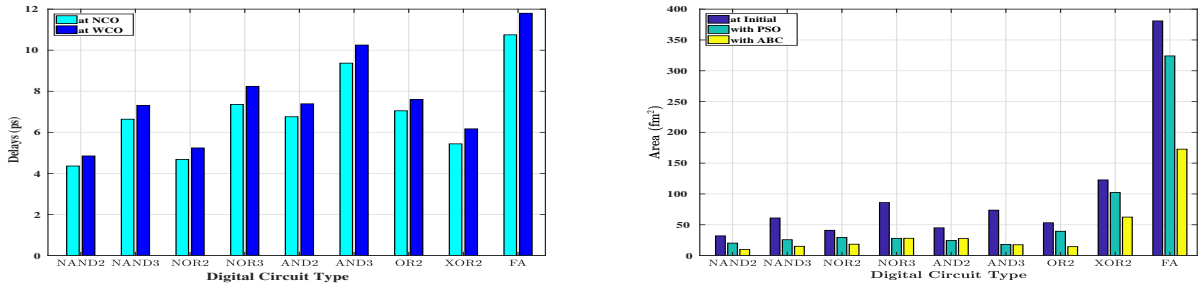
TABLE V: Average leakage power obtained for benchmark circuits at 32nm technology

Sr. No.	Digital Circuit Type	Average Leakage Power (μW)					
		at Initial		with PSO		with ABC	
		NCO	WCO	NCO	WCO	NCO	WCO
1	$C17$	10.4	33.8	6.82	23.9	3.26	11.8
2	*Parity Checker*	5.44	5.65	4.91	4.98	3.30	3.51
3	$C432$ (*Interrupt Controller*)	173	242	168	240	163	247
4	$4-Bit\ Multiplier$	12.0	12.2	9.3	10.5	6.27	7.63
5	$32-Bit\ Ripple\ Carry\ Adder$	406	1246	282	845	182	476
6	$32-Bit\ Carry\ Skip\ Adder$	750	2330	549	1608	374	986
7	$32-Bit\ Carry\ Select\ Adder$	810	2540	593	1705	382	1006



(a) (b)

Fig. 3: Results obtained in various basic CMOS digital circuits at 32nm technology a) at NCO b) at WCO.



(a) (b)

Fig. 4: a) Initial highest delays taken as bound for in basic CMOS digital circuits at 32nm technology b) Area obtained for basic CMOS digital circuits (32nm technology)

TABLE VI: Total average power obtained for Basic Digital cells at 32nm technology

Sr. No.	Digital Circuit Type	Total Average Power (μW)					
		at Initial		with PSO		with ABC	
		NCO	WCO	NCO	WCO	NCO	WCO
1	$NAND2$	0.37	0.60	0.26	0.39	0.19	0.28
2	$NAND3$	0.68	0.95	0.27	0.31	0.22	0.26
3	$NOR2$	0.43	0.25	0.34	0.18	0.21	0.12
4	$NOR3$	0.95	0.59	0.34	0.49	0.33	0.22
5	$AND2$	0.48	2.55	0.24	0.41	0.28	1.24
6	$AND3$	1.06	1.47	0.30	0.39	0.28	0.34
7	$OR2$	0.64	2.20	0.47	1.64	0.27	0.86
8	$XOR2$	0.83	1.53	0.54	0.94	0.37	0.60
9	28-T *Full Adder*	7.4	7.8	4.5	4.9	3.21	3.15

The overall optimized area in basic cells using reported algorithms have been displayed in Fig. 4 b).

IV. Conclusion

In this manuscript, we have implemented swarm based Particle Swarm Optimization algorithm and Artificial Bee Colony algorithm on various digital logic circuits to obtain PVT invariant robust transistor sizing for minimum average leakage power and area without giving any penalty to propagation delay. The average leakage power reduction is achieved up to 88%. The computation time to complete the simulation of single digital circuit using HSPICE tool is around one hour for 500 iterations. The simulations have been carried out on a computer having 12 core Xeon processor and 32GB RAM.

REFERENCES

[1] S. Bhunia and S. Mukhopadhyay, *Low-power variation-tolerant design in nanometer silicon*. Springer, 2010.

[2] K. Kuhn, C. Kenyon, A. Kornfeld, M. Liu, A. Maheshwari, W.-k. Shih, S. Sivakumar, G. Taylor, P. VanDerVoorn, and K. Zawadzki, "Managing process variation in intel's 45nm cmos technology." *Intel Technology Journal*, vol. 12, no. 2, 2008.

[3] J. Jaffari and M. Anis, "Statistical thermal profile considering process variations: Analysis and applications," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 27, no. 6, pp. 1027–1040, 2008.

[4] R. X. Gu and M. I. Elmasry, "Power dissipation analysis and optimization of deep submicron cmos digital circuits," *IEEE Journal of solid-state circuits*, vol. 31, no. 5, pp. 707–713, 1996.

[5] A. Agarwal, S. Mukhopadhyay, A. Raychowdhury, K. Roy, and C. H. Kim, "Leakage power analysis and reduction for nanoscale circuits," *IEeE Micro*, vol. 26, no. 2, pp. 68–80, 2006.

[6] Z. Abbas, A. Mastrandrea, and M. Olivieri, "A voltage-based leakage current calculation scheme and its application to nanoscale mosfet and finfet standard-cell designs," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 22, no. 12, pp. 2549–2560, 2014.

[7] Z. Abbas and M. Olivieri, "Impact of technology scaling on leakage power in nano-scale bulk cmos digital standard cells," *Microelectronics Journal*, vol. 45, no. 2, pp. 179–195, 2014.

[8] Z. Abbas and M.Olivieri, "Optimal transistor sizing for maximum yield in variation-aware standard cell design," *International Journal of Circuit Theory and Applications*, vol. 44, no. 7, pp. 1400–1424, 2016.

[9] Z. Abbas, M. Olivieri, and A. Ripp, "Yield-driven power-delay-optimal cmos full-adder design complying with automotive product specifications of pvt variations and nbti degradations," *Journal of Computational Electronics*, vol. 15, no. 4, pp. 1424–1439, 2016.

[10] K.-F. Man, K.-S. Tang, and S. Kwong, "Genetic algorithms: concepts and applications [in engineering design]," *IEEE transactions on Industrial Electronics*, vol. 43, no. 5, pp. 519–534, 1996.

[11] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *MHS'95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science*. Ieee, 1995, pp. 39–43.

[12] D. Karaboga, "An idea based on honey bee swarm for numerical optimization," Technical report-tr06, Erciyes university, engineering faculty, computer engineering department, Tech. Rep., 2005.

[13] D. P. Rini, S. M. Shamsuddin, and S. S. Yuhaniz, "Particle swarm optimization: technique, system and challenges," *International journal of computer applications*, vol. 14, no. 1, pp. 19–26, 2011.

[14] P. Gupta, H. Mandadapu, S. Gourishetty, Z. Abbas, "Robust Transistor Sizing for Improved Performances in Digital Circuits using Optimization Algorithms", 20th International Symposium on Quality Electronic Design, California, USA.

[15] D. M. Harris, N. Wei, Z. Wang, A. Fikes, and A. Thaker, "Evaluation of predictive technology models," *Microelectronics journal*, vol. 80, pp. 7–17, 2018.

[16] Predictive technology model (ptm) [ONLINE]. available: http://ptm.asu.edu/.

[17] Bsim 4.6.1 mosfet model, users manual, tech. rep., EECS Department, University of California, Berkeley, 2007, available: http://www.srware.com/xictools/docs/model_docs/bsim4.6.1/bsim461 _manual.pdf.

[18] X. Xi, M. Dunga, J. He, W. Liu, K. M. Cao, X. Jin, J. J. Ou, M. Chan, A. M. Niknejad, C. Hu *et al.*, "Bsim4. 3.0 mosfet model," *Dept. Elect. Eng. Comput. Sci., Univ. California, Berkeley, CA, Tech. Rep*, vol. 94720, p. 30, 2003.